

Phan Quốc Khánh

VẬN TRỪ HỌC

*vận dụng...
trừ tính...
vận dụng...
trừ tính...
vận dụng...
trừ tính...
vận dụng...
trừ tính...
vận dụng...
trừ tính...
vận dụng...
trừ tính...
vận dụng...
trừ tính...*



NHÀ XUẤT BẢN GIÁO DỤC

GS.TSKH. Phan Quốc Khánh

VẬN TRÙ HỌC

(Tái bản lần thứ nhất)

NHÀ XUẤT BẢN GIÁO DỤC

Chịu trách nhiệm xuất bản :

Chủ tịch HĐQT kiêm Tổng Giám đốc NGÔ TRẦN ÁI
Phó Tổng Giám đốc kiêm Tổng biên tập NGUYỄN QUÝ THAO

Tổ chức bàn thảo và chịu trách nhiệm nội dung:

Phó Tổng Giám đốc kiêm Giám đốc NXBGD tại TP.HCM
VŨ BÁ HÒA

Biên tập lần đầu và tái bản:

TS. NGUYỄN CAO THẮNG

Trình bày bìa :

VŨ CÔNG MINH

Minh họa :

GS. TSKH. PHAN QUỐC KHÁNH

TS. NGUYỄN CAO THẮNG

LỜI NÓI ĐẦU

Hơn ba chục năm giảng dạy và nghiên cứu về tối ưu hóa – vận trù học với nhiều tâm huyết và kinh nghiệm học hỏi được từ lao động kiên trì, từ thầy và bạn, đến nay chúng tôi mới quyết tâm viết quyển sách *VẬN TRÙ HỌC* này. (Chúng tôi đã lưỡng lự nhiều lần, đã soạn bài và dạy không ít, nhưng luôn muốn sách phải có nét mới và đáp ứng nhu cầu độc giả, nhất là dùng để học cho sinh viên và tham khảo cho kỹ sư và các nhà khoa học trẻ trong mọi lĩnh vực.)

Vận trù học có thể dùng cho tất cả mọi người vì nó bàn về cách vận dụng khoa học, mà nền tảng là toán học, để trù tính mọi việc. Để đọc sách này thực chất là chỉ cần kiến thức hết lớp 12, vì những đòi hỏi cao hơn đều được trình bày đủ mức để sử dụng. Với độc giả đã có kiến thức về Giải tích, Đại số tuyến tính và Lý thuyết xác suất thuộc ba học kì đầu của chương trình đại học và cao đẳng, thì việc hiểu sâu vấn đề sẽ thuận lợi hơn.

Chúng tôi đã cố gắng viết để các chương tương đối độc lập nhau. Bạn đọc có thể chỉ đọc phần mình quan tâm, vì đối tượng của vận trù học rất rộng như tên của nó cũng đã thể hiện. Tuy vậy chúng tôi cũng chú ý nêu lên quan hệ logic giữa các chương, mục vì nó rất cần để hiểu kỹ và nắm hết cái hay của một lĩnh vực khoa học, nhằm nghiên cứu và ứng dụng. Chính vì sự rộng lớn của vận trù học, chúng tôi cố gắng giải quyết mâu thuẫn giữa việc muốn viết nhiều và kỹ càng, toàn diện (các sách mới về vận trù học trên thế giới thường dày hơn nghìn trang) và khuôn khổ sách mà Nhà xuất bản Giáo dục cho phép để phù hợp

tình hình Việt Nam. Có lẽ sách đã bao hàm được những nội dung cơ bản của vận trù học. Trong đó chúng tôi muốn tạo một nét riêng, khác các sách khác kể cả trên thế giới, là đi sâu vào bài toán vận tải, vì ở Mi chẳng hạn người ta thấy là 85% các bài toán quy hoạch tuyến tính trong thực tế, phần nhiều không phải "vận tải", ở dạng mô hình toán học này.

Rất nhiều người đã góp phần hình thành quyển sách này. Trước hết chúng tôi phải nhắc đến công của các đồng nghiệp, nhất là thuộc Bộ môn Tối ưu và Hệ thống, trường Đại học khoa học Tự nhiên, Đại học Quốc gia TP. Hồ Chí Minh. Tác giả rất cảm ơn Học viện Công nghệ Bưu chính Viễn thông đã mời viết bài giảng về toán chuyên ngành cho quản trị kinh doanh mà nội dung là tiền đề của sách này. Lời cảm tạ chân thành cũng dành cho Hội đồng liên trường đại học dùng tiếng Pháp (C.I.U.F - C.U.D. / C.U.I) của Vương quốc Bỉ, đã tài trợ cho Chương trình hợp tác quốc tế "Optimisation – Mathématiques Appliquées" trong suốt năm năm qua mà tác giả là đồng chủ nhiệm cùng GS. Nguyễn Văn Hiến thuộc Đại học Namur. Sự hợp tác nghiên cứu và đào tạo rất hiệu quả này đã giúp nhiều cho việc hình thành sách.

Sự tri ân trực tiếp nhất là gửi đến TS. Nguyễn Cao Thắng, đã biên tập sách, Nhà xuất bản Giáo dục đã xuất bản sách và bạn đọc những người cổ vũ tôi và sẽ góp ý sửa đổi những thiếu sót chắc chắn còn nhiều của quyển sách.

TP. Hồ Chí Minh, 9.1.2002

PHAN QUỐC KHÁNH

MỞ ĐẦU

Vận trù (Operations research) có nghĩa là *Vận dụng* khoa học mà nền tảng là toán học để *trù tính* mọi việc, theo chủ ý của GS. Hoàng Tuy, người chọn thuật ngữ tiếng Việt này từ đầu thập niên 60 thế kỷ hai mươi. Nghiên cứu vận trù xuất hiện trên thế giới trước đó nhiều thập kỉ. Phát triển và ứng dụng thực sự vận trù đầu tiên là trong quân sự thời kì chiến tranh thế giới thứ hai, sớm nhất là ở Anh trong việc dùng radar phòng không chống tàu ngầm. Sau chiến tranh, vận trù càng được phát triển với sự áp dụng ngày càng rộng trong các lĩnh vực rất đa dạng như kinh tế, kinh doanh, công nghệ, quản lí hành chính, sản xuất, xây dựng, vận tải, viễn thông, quân sự, y tế, chính trị, giáo dục, đào tạo, Thuật ngữ vận trù đôi khi cũng được dùng gần như đồng nghĩa với *khoa học quản trị* (management science) và cũng rất gần với *chọn quyết định* (decision making). Điểm nổi bật là các bài toán vận trù thường nhằm tìm nghiệm tối ưu, tức là chọn quyết định tốt nhất theo một mục tiêu nào đó. Vì vậy, vận trù học cũng rất gần với tối ưu hóa. Nhưng nó liên quan đến rất nhiều lĩnh vực khoa học như lý thuyết kinh tế, xác suất thống kê, công nghệ thông tin ... và cả vật lí, kĩ thuật. Tuy vậy, vẫn phải khẳng định đây là một bộ phận của toán ứng dụng vì phương pháp và ngôn ngữ toán học là chủ đạo.

Các bước nghiên cứu và ứng dụng một bài toán vận trù điển hình là như sau.

1. Xác định vấn đề cần giải quyết và thu thập dữ liệu thống kê của quá khứ.
2. Lập mô hình toán học.
3. Xây dựng các thuật toán giải bài toán đã mô hình hóa bằng ngôn ngữ thuận lợi cho việc lập trình cho máy tính.

4. Tính toán thủ và điều chỉnh mô hình nếu cần.

5. Áp dụng giải các bài toán thực tế.

Các bài toán thực tế được giải quyết bằng các phương pháp của vận trù học thường là cỡ lớn, nên việc áp dụng vận trù ngày càng hiệu quả với sự phát triển vượt bậc của công nghệ thông tin và nói riêng là năng lực của máy tính.

Vì sự đa dạng của nghiên cứu và áp dụng vận trù học, người ta thường lập các *đội vận trù* (operations research team, viết tắt là O.R team) có chuyên gia nhiều lĩnh vực, được chủ trì bởi chuyên gia về tối ưu hóa toán học. Áp dụng vận trù thậm chí phải tính đến cả yếu tố tâm lý bên cạnh phương pháp khoa học. Chẳng hạn trong *bài toán thang máy* (elevator problem). Người ta đã áp dụng lý thuyết xếp hàng (xem Chương 10) để tính toán, khi thấy nhiều người than phiền phải chờ lâu ở thang máy, và thấy thời gian chờ không đáng kể. Nghiên cứu kỹ hơn người ta phát hiện ra rằng nguyên nhân là chính việc đứng phía trước và cả trong thang máy rất tẻ nhạt. Sáng kiến được đưa ra là lắp các gương lớn. Quả nhiên mọi người mãi ngắm mình và người khác trong gương và không còn lời than vãn về chờ đợi nữa.

100%
ĐÃ ĐƯỢC
SỬ DỤNG
TRÊN 100

CHƯƠNG 1

QUY HOẠCH TUYẾN TÍNH. PHƯƠNG PHÁP ĐƠN HÌNH

1.1 QUY HOẠCH TUYẾN TÍNH VÀ CÁCH GIẢI HÌNH HỌC

Quy hoạch tuyến tính (linear programming) là phần quan trọng nhất, cơ sở nhất của tất cả các lĩnh vực khoa học dùng các phương pháp toán học để tối ưu hoá hoạt động như quy hoạch toán học, vận trù học, toán kinh tế, Mô hình toán học đơn giản, phổ dụng và đẹp đẽ này được G.B. Danzig đưa ra cùng phương pháp đơn hình (simplex method) để giải cho trường hợp tổng quát vào năm 1947, khi ông lãnh đạo một tập thể các nhà khoa học Mỹ giải quyết bài toán lập kế hoạch cho không quân Mỹ. Ngay sau khi được công bố, quy hoạch tuyến tính đã được dùng để mô tả toán học và giải quyết rất nhiều bài toán thực tế thuộc các lĩnh vực rất khác nhau. Ta hãy xét hai bài toán điển hình được mô hình hoá bằng quy hoạch tuyến tính.

THÍ DỤ 1.1.1 *Bài toán kế hoạch sản xuất* - (Production planning problem). Công ty Reddy Mikks sản xuất sơn nội thất và sơn ngoài trời. Nguyên liệu gồm 2 loại A và B với trữ lượng là 6 tấn và 8 tấn tương ứng. Để sản xuất một tấn sơn nội thất cần 2 tấn nguyên liệu A và 1 tấn nguyên liệu B. Hai số tương ứng của sơn ngoài trời là 1 tấn và 2 tấn. Qua tiếp thị được biết nhu cầu thị trường là như sau (cho một ngày):

- Nhu cầu sơn nội thất không hơn nhu cầu sơn ngoài trời quá 1 tấn;

- Nhu cầu cực đại của sơn nội thất là 2 tấn.

Giá bán sỉ là 2000 USD một tấn sơn nội thất và 3000 USD một tấn sơn ngoài trời.

Vấn đề là cần sản xuất mỗi ngày như thế nào để doanh thu là lớn nhất.

Gọi x_1 và x_2 là số lượng (tính theo tấn) sơn ngoài trời và sơn nội thất tương ứng cần sản xuất trong ngày. Đây sẽ là các *biến* (variable) hoặc *phương án* (alternative) của bài toán. Khi đó doanh thu trong ngày sẽ là

$$z = 3x_1 + 2x_2$$

và được gọi là *hàm mục tiêu* (objective function).

Các ràng buộc trên biến x_1, x_2 , sẽ gọi là các *ràng buộc* (constraint) của bài toán, là như sau. Nguyên liệu sử dụng không được quá trữ lượng:

$$x_1 + 2x_2 \leq 6 \quad (\text{nguyên liệu A}),$$

$$2x_1 + x_2 \leq 8 \quad (\text{nguyên liệu B}).$$

Sản xuất không nhiều hơn nhu cầu thị trường:

$$x_2 - x_1 \leq 1,$$

$$x_2 \leq 2.$$

Sản lượng phải là số thực không âm:

$$x_1 \geq 0, x_2 \geq 0.$$

Ta gọi phương án (x_1, x_2) là *chấp nhận được* (feasible) nếu nó thỏa mọi ràng buộc. Khi đó (x_1, x_2) cũng gọi là *nghiệm chấp nhận được* (feasible solution).

Vậy bài toán trở thành:

Tìm phương án chấp nhận được làm cực đại hàm mục tiêu z và

được viết ở dạng toán học như sau

$$\begin{aligned}\max z &= 3x_1 + 2x_2, \\ x_1 + 2x_2 &\leq 6, \\ 2x_1 + x_2 &\leq 8, \\ -x_1 + x_2 &\leq 1, \\ x_2 &\leq 2, \\ x_1 \geq 0, \quad x_2 &\geq 0.\end{aligned}$$

THÍ DỤ 1.1.2. *Bài toán vận tải* - (Transportation problem). Hàng hóa được vận chuyển từ m kho đến n cửa hiệu bán lẻ. Lượng hàng ở kho i là $s_i \geq 0$ (tấn), $i = 1, \dots, m$ và cửa hiệu j có nhu cầu $d_j \geq 0$ (tấn), $j = 1, \dots, n$. Cước vận chuyển một tấn hàng từ kho i đến cửa hiệu j là c_{ij} đồng. Giả sử tổng hàng ở các kho và tổng nhu cầu bằng nhau:

$$\sum_{i=1}^m s_i = \sum_{j=1}^n d_j. \quad (1.1)$$

Bài toán đặt ra là lập kế hoạch vận chuyển để tiền cước là nhỏ nhất, với điều kiện là mỗi cửa hàng đều nhận đủ và mỗi kho đều trao hết hàng.

Gọi lượng hàng vận chuyển từ kho i đến cửa hàng j là x_{ij} , thì kế hoạch vận chuyển, tức là phương án theo định nghĩa chung, là ma trận (x_{ij}) cấp $m \times n$. Dạng toán học của bài toán là

$$\begin{aligned}\min z &= \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij}, \\ \sum_{j=1}^n x_{ij} &= s_i, \quad i = 1, \dots, m,\end{aligned} \quad (1.2)$$

$$\sum_{i=1}^m x_{ij} = d_j, j = 1, \dots, n,$$

$$x_{ij} \geq 0, i = 1, \dots, m, j = 1, \dots, n.$$

Mô hình này gọi là *mô hình vận tải đóng* (closed transportation model). Nếu không có giả thiết (1.1) và ràng buộc (1.2) đổi lại là

$$\sum_{j=1}^n x_{ij} \leq s_i, \text{ tức là các kho có thể không trao hết, thì mô hình được}$$

gọi là *mô hình vận tải mở* (open transportation model).

1.1.1 Định nghĩa quy hoạch tuyến tính

Các thí dụ trên thuộc các lĩnh vực khác nhau nhưng được đưa về dạng toán học có nét chung là:

Phải xác định các *biến quyết định* (decision variable), gọi tắt là *biến*, hoặc *phương án* (alternative), thỏa mãn các ràng buộc sao cho làm cực đại hoặc cực tiểu hàm mục tiêu. Hơn nữa cả hàm mục tiêu và các ràng buộc đều tuyến tính (bậc nhất) theo biến quyết định.

Nhận xét thêm là việc tìm cực đại có thể dễ dàng chuyển thành tìm cực tiểu và ngược lại, vì $\max z = -\min(-z)$. Do đó các thí dụ trên đều là trường hợp riêng của bài toán

$$\begin{aligned} \min z &= c^T x, \\ a_i^T x &\geq b_i, & i \in M_1, \\ a_i^T x &\leq b_i, & i \in M_2, \\ a_i^T x &= b_i, & i \in M_3, \\ x_j &\geq 0, & j \in N_1, \\ x_j &\leq 0, & j \in N_2, \end{aligned} \quad (1.3)$$

ở đây M_1, M_2, M_3, N_1 và N_2 là tập hợp chỉ số nào đó, c^T là chuyển vị (transpose) của vectơ c , x , a_i và c là các vectơ n thành phần, b_i là các số thực. Ta luôn quy ước vectơ là vectơ cột, vậy c^T là vectơ hàng. Chú ý là $\min z = c^T x$ cũng thường viết gọn là $\min c^T x$.

Ràng buộc đẳng thức có thể viết tương đương ở dạng bất đẳng thức như sau:

$$a_i^T x = b_i \Leftrightarrow a_i^T x \leq b_i, \quad a_i^T x \geq b_i.$$

Ràng buộc dấu $x_j \geq 0$ là trường hợp riêng của ràng buộc bất đẳng thức dạng $a^T x \geq b$ với $a^T = (0, \dots, 0, 1, 0, \dots, 0)$ và $b = 0$. Ngoài ra, rõ ràng $a_i^T x \leq b_i$ tương đương với $-a_i^T x \geq -b_i$.

Giả sử sau khi thực hiện các biến đổi trên (1.3) có tất cả m ràng buộc dạng $a_i^T x \leq b_i$. Ký hiệu A là ma trận cấp $m \times n$ có các hàng là a_i^T và $b = (b_1, \dots, b_m)^T$, thì bài toán (1.3) có dạng ma trận là

$$\begin{aligned} \min \quad & c^T x, \\ \text{s.t.} \quad & Ax \leq b. \end{aligned} \quad (1.4)$$

Ở đây bất đẳng thức vectơ $Ax \leq b$ được hiểu là m bất đẳng thức theo thành phần, tức là $a_i^T x \leq b_i$, $i = 1, \dots, m$. Vậy quy hoạch tuyến tính là bài toán có dạng tổng quát là (1.3) hoặc viết ở dạng ma trận (1.4). Dùng lời diễn đạt thì quy hoạch tuyến tính là bài toán tìm cực tiểu (hoặc cực đại) của hàm mục tiêu tuyến tính với các ràng buộc bất đẳng thức và đẳng thức tuyến tính. Tuy vậy, trong nghiên cứu quy hoạch tuyến tính cũng như khi áp dụng nó, người ta thường dùng hai dạng đặc thù

<i>Dạng chính tắc:</i> (canonical form)	$\min c^T x,$ $Ax = b,$ $x \geq 0;$	<i>Dạng chuẩn:</i> (normal form hoặc standard form)	$\min C^T x,$ $Ax \leq b,$ $x \geq 0.$
--	---	---	--

Ở đây A là ma trận cấp $m \times n$ với các hàng là a_i , $i = 1, \dots, m$ và

các cột là $A_j, j = 1, \dots, n$. c và x là vectơ n chiều và b là vectơ m chiều (Ở đây cũng có sự khác nhau về thuật ngữ. Dantzig định nghĩa dạng chính tắc trên đây là standard form và gọi cal form là trường hợp đặc biệt của nó, mà ta sẽ gặp sau.) Việc viết riêng các ràng buộc $x \geq 0$ là có nhiều thuận lợi, nhất là khi xét về hình học. Nếu ở bài toán có các biến tự do (về dấu không có ràng buộc), giả sử là x_j , thì ta thay x_j bằng hai biến mới x_j^+ và x_j^- bởi liên hệ $x_j = x_j^+ - x_j^-$, $x_j^+ \geq 0, x_j^- \geq 0$. Ta đã biết ràng buộc đẳng thức có thể viết tương đương thành ràng buộc bất đẳng thức. Ngược lại, (để đưa về dạng chính tắc) ràng buộc bất đẳng thức có thể chuyển thành đẳng thức bằng cách đưa thêm vào biến bù như sau. Giả sử ta có ràng buộc bất đẳng thức $a_i^T x \leq b_i$. Ta đưa thêm vào biến mới s_i , gọi là *biến bù* (slack variable) như sau:

$$a_i^T x + s_i = b_i,$$

$$s_i \geq 0.$$

$$(\text{Tương tự, } a_i^T x \geq b_i \Leftrightarrow a_i^T x - s_i = b_i, s_i \geq 0.)$$

1.1.2 Giải hình học quy hoạch tuyến tính hai biến

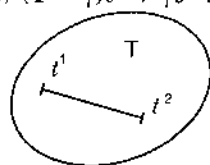
Khi bài toán chỉ có hai biến, ta có thể giải quy hoạch tuyến tính bằng hình học dễ dàng. Chú ý rằng trường hợp riêng này cũng cho phép ta tưởng tượng hình học về bài toán tổng quát. Hãy xét thí dụ 1.1 về kế hoạch sản xuất của công ty Reddy Mikks. Nhận xét rằng mỗi ràng buộc bất đẳng thức $a_i^T x \leq b_i$ cho ta một *nửa không gian đóng* (closed half space) $\{x \in \mathbf{R}^n : a_i^T x \leq b_i\}$ giới hạn bởi *siêu phẳng* (hyperplane) $\{x \in \mathbf{R}^n : a_i^T x = b_i\}$. Trường hợp hai biến, đây là nửa mặt phẳng (đóng) giới hạn bởi đường thẳng. Tập tất cả các nghiệm chấp nhận được gọi là *miền chấp nhận được* (feasible region), hoặc *miền ràng buộc*, là một *tập lồi đa diện* (polyhedron). Tập lồi đa diện là giao

của một số hữu hạn các nửa không gian đóng, tức là tập có dạng $\{x \in \mathbf{R}^n: a_i^T x \leq b_i, i = 1, \dots, m\}$; hoặc ở dạng vectơ $\{x \in \mathbf{R}^n: Ax \leq b\}$. Chú ý rằng chiều của bất đẳng thức ở đây có thể ở dạng khác. Các tập $\{x \in \mathbf{R}^n: Ax \geq b\}$, $\{x \in \mathbf{R}^n: a_i^T x \leq b_i, i = 1, \dots, m; a_j^T x \geq b_j, j = 1, \dots, k; a_p^T x = b_p, p = 1, \dots, s\}$ cũng là tập lồi đa diện. Nếu tập lồi đa diện giới nội ta gọi nó là *đa diện lồi* (polytope). Trong trường hợp hai biến, đây là đa giác lồi. *Tính lồi* (convexity) là một khái niệm quan trọng. Tập $T \subset \mathbf{R}^n$ gọi là *lồi* (convex) nếu bất kì hai điểm $t^1 \in T$ và $t^2 \in T$ đều có đoạn thẳng nối chúng nằm hoàn toàn trong T (xem H.1.1). Chú ý rằng đoạn nối t^1 và t^2 là tập $\{t \in \mathbf{R}^n: t = (1 - \gamma)t^1 + \gamma t^2, 0 \leq \gamma \leq 1\}$, tức là tập tất cả các *tổ hợp lồi* (convex combination) $(1 - \gamma)t^1 + \gamma t^2$ của t^1 và t^2 . Một cách tự nhiên, ta có định nghĩa tổ hợp lồi tổng quát của nhiều điểm.

Điểm trong $\mathbf{R}^n$ có dạng $\sum_{i=1}^m \gamma_i t^i$, ở

đây $\gamma_i \geq 0$ và $\sum_{i=1}^m \gamma_i = 1$, gọi là một

tổ hợp lồi của các điểm $t^i, i = 1, \dots, m$.



tập lồi



tập không lồi

H. 1.1

Nhắc lại, bài toán công ty Reddy Mikks là quy hoạch tuyến tính

$$\max 3x_1 + 2x_2,$$

$$x_1 + 2x_2 \leq 6,$$

$$2x_1 + x_2 \leq 8,$$

$$-x_1 + x_2 \leq 1,$$

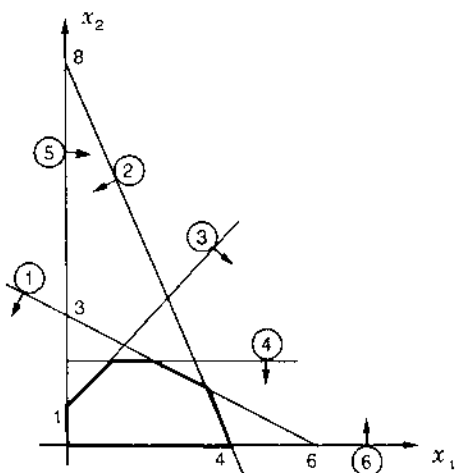
$$x_2 \leq 2,$$

$$x_1 \geq 0, x_2 \geq 0.$$

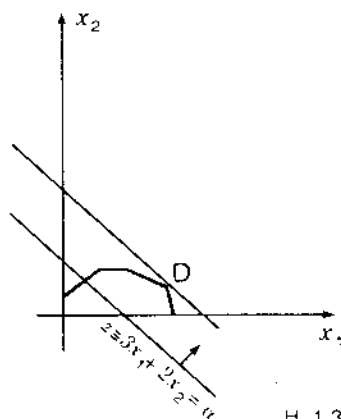
Miền chấp nhận được được biểu diễn ở H.1.2, các số trong vòng

tròn và mũi tên, chẳng hạn ① là chỉ số thứ tự của ràng buộc tương ứng và phía của nửa không gian xác định bởi ràng buộc.

Đường mức (level curve, trong trường hợp nhiều biến hơn thì gọi là *mặt mức* - level surface) của hàm mục tiêu là đường thẳng $z = 3x_1 + 2x_2 = \alpha$. Khi cho α tăng dần ta thấy điểm cuối cùng mà đường mức α còn cắt miền chấp nhận được là đỉnh D (H.1.3). D là giao điểm hai đường (1) và (2):



H. 1.2



H. 1.3

$$x_1 + 2x_2 = 6,$$

$$2x_1 + x_2 = 8.$$

Giải hệ này ta được $x_1 = 3\frac{1}{3}$, $x_2 = 1\frac{1}{3}$, chính là *ng nghiệm chấp nhận được tối ưu* (optimal feasible solution), thường gọi tắt là *ng nghiệm tối ưu*. Giá trị mục tiêu tối ưu (optimal objective value) là $z^* = z(3\frac{1}{3}, 1\frac{1}{3}) = 12\frac{1}{3}$ ngàn USD.

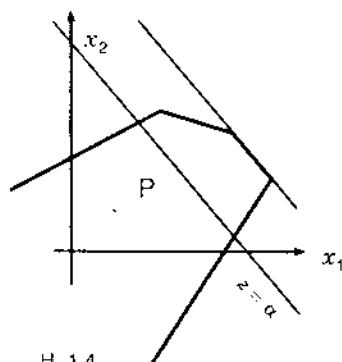
Qua bài toán cụ thể trên ta có thể nhận xét sơ bộ như sau:

a. Miền chấp nhận được của quy hoạch tuyến tính là tập lồi đa diện. Nếu nó giới nội (tức là đa diện lồi) thì quy hoạch tuyến tính có nghiệm tối ưu là một đỉnh. Trường hợp nghiệm tối ưu không duy nhất nhưng miền chấp nhận được (thậm chí có thể không giới nội) có đỉnh

(H.1.4) thì vẫn luôn có nghiệm tối ưu là đỉnh.

b. Trường hợp không có nghiệm tối ưu thì hàm mục tiêu không giới nội trên miền chấp nhận được (H.1.5).

c. Trường hợp miền chấp nhận được không có đỉnh được minh họa ở (H.1.6). Bài toán quy hoạch tuyến tính có thể không có nghiệm tối ưu hoặc có nhưng không có nghiệm tối ưu là đỉnh.

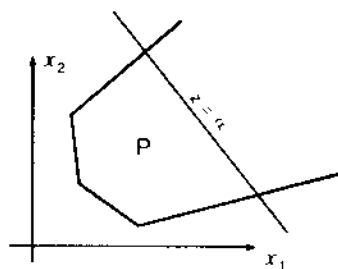


H. 1.4

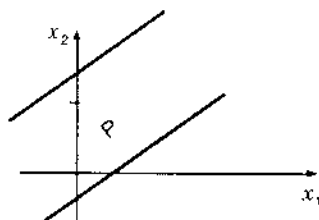
Những nhận xét hình học trên đây vẫn đúng cho cả trường hợp nhiều biến (hơn 2) và cho những gợi ý quan trọng để xây dựng phương pháp giải quy hoạch tuyến tính, ở dạng các thuật toán, bằng ngôn ngữ đại số để có thể lập trình cho máy tính.

Về nguyên tắc, ta có thể giải quy hoạch tuyến tính hai biến, với số ràng buộc bất kì, bằng hình học.

Độc giả có thể xem một số cách giải quy hoạch tuyến tính đơn giản khác trong [Khánh - Nương 2000].



H. 1.5



H. 1.6

1.2 PHƯƠNG PHÁP ĐƠN HÌNH

1.2.1 Một thí dụ

Phương pháp đơn hình được G. Dantzig đưa ra năm 1947 (xem [Dantzig 1949], [Dantzig 1963]) cùng lúc khi ông khai sinh quy hoạch tuyến tính. Trong khoảng bốn chục năm, nó là phương pháp thực sự hiệu quả duy nhất để giải quy hoạch tuyến tính trong thực tế (khi các bài toán thường là cỡ lớn). Trong thập niên cuối vừa qua chỉ có phương pháp điểm trong là cạnh tranh được với phương pháp đơn hình. Với các cách nhìn hiện đại, tư tưởng của phương pháp đơn hình rất đơn giản. Ở đây ta sẽ theo cách tiếp cận của [Chvátal 1983] và [Vanderbei 1996].

Xét quy hoạch tuyến tính

$$\begin{aligned} \min z &= -5x_1 - 4x_2 - 3x_3, \\ 2x_1 + 3x_2 + x_3 &\leq 5, \\ 4x_1 + x_2 + 2x_3 &\leq 11, \\ 3x_1 + 4x_2 + 2x_3 &\leq 8, \\ x_1, x_2, x_3 &\geq 0. \end{aligned} \tag{1.5}$$

Trước hết ta hãy đưa bài toán về dạng chính tắc bằng cách đưa vào biến bù w_1, w_2, w_3 và viết hệ ở dạng chính tắc đặc biệt sau

$$\begin{aligned} \min z &= -5x_1 - 4x_2 - 3x_3, \\ w_1 &= 5 - 2x_1 - 3x_2 - x_3, \\ w_2 &= 11 - 4x_1 - x_2 - 2x_3, \\ w_3 &= 8 - 3x_1 - 4x_2 - 2x_3, \\ x_1, x_2, x_3, w_1, w_2, w_3 &\geq 0. \end{aligned}$$

Để bắt đầu ta cần có một nghiệm chấp nhận được. Để nhận thấy một nghiệm là

$$x_1 = 0, x_2 = 0, x_3 = 0, w_1 = 5, w_2 = 11, w_3 = 8.$$

Giá trị mục tiêu tương ứng là $z = 0$.

Trong bước lặp đầu tiên ta xem có thể cải tiến nghiệm xuất phát thế nào để được nghiệm chấp nhận được mới với mục tiêu z nhỏ hơn. Vì hệ số của x_1 ở hàm mục tiêu là âm, nếu ta tăng giá trị x_1 từ 0 lên giá trị dương càng lớn thì z càng giảm (nếu ta giữ $x_2 = 0, x_3 = 0$). Nhưng khi đó giá trị các biến bù cũng bị thay đổi. Sự thay đổi này không ảnh hưởng hàm mục tiêu nhưng x_1 chỉ được thay đổi sao cho $w_1, w_2, w_3 \geq 0$. Vì $x_2 = 0, x_3 = 0$ nên

$$\begin{aligned} w_1 &= 5 - 2x_1 \geq 0, & x_1 &\leq \frac{5}{2}, \\ w_2 &= 11 - 4x_1 \geq 0, & \Rightarrow x_1 &\leq \frac{11}{4}, \\ w_3 &= 8 - 3x_1 \geq 0, & x_1 &\leq \frac{8}{3}. \end{aligned}$$

Vậy ta lấy $x_1 = \frac{5}{2}$ và khi đó $w_1 = 0$. Ta được nghiệm chấp nhận được mới tốt hơn sau bước một là

$$x_1 = \frac{5}{2}, x_2 = 0, x_3 = 0, w_1 = 0, w_2 = 1, w_3 = \frac{1}{2}.$$

Giá trị mục tiêu là $z = -12,5$.

Nhận xét rằng bước lặp đầu dễ chọn vì trong nghiệm xuất phát có một nhóm biến bằng 0 và các biến khác biểu diễn qua chúng qua các ràng buộc. Do đó để vào bước 2 ta cũng viết lại bài toán ở dạng tương tự, tức là đổi vai trò của x_1 và w_1 . Vậy x_1, w_2, w_3 và z sẽ biểu diễn qua w_1, x_2 và x_3 . Muốn vậy, ở ràng buộc đầu ta giải x_1 ra theo w_1, x_2 và x_3

$$x_1 = \frac{5}{2} - \frac{1}{2}w_1 - \frac{3}{2}x_2 - \frac{1}{2}x_3. \quad (1.6)$$

Để vào bước lặp 3 ta lại viết phương trình ở dạng x_1, w_2, x_3 và z biểu diễn theo w_1, x_2 và w_3 . Thực hiện phép toán trên hàng đối với phương trình cho w_3 ta được

$$x_3 = 1 + 3w_1 + x_2 - 2w_3.$$

Làm tương tự đối với các hàng khác ta được

$$\begin{aligned} z &= -13 + w_1 + 3x_2 + w_3 \\ x_1 &= 2 - 2w_1 - 2x_2 + w_3 \\ w_2 &= 1 + 2w_1 + 5x_2 \\ x_3 &= 1 + 3w_1 + x_2 - 2w_3 \end{aligned} \quad (1.8)$$

Bắt đầu bước 3, ta thấy không có hệ số nào ở hàm mục tiêu là âm nên ta không thể cho biến “độc lập” nào tăng từ 0 thành số dương (và giữ nguyên giá trị 0 cho biến “độc lập” khác) để giảm hàm mục tiêu. Phương pháp đơn hình phải dừng. Nhưng rất tốt là ta thấy ngay nghiệm chấp nhận được ở bước 2 là tối ưu. Thật vậy, hệ (1.8) là tương đương với quy hoạch tuyến tính ban đầu (1.5). Nhận xét (1.8) ta thấy, vì các biến phải không âm nên $z \geq -13$ với mọi nghiệm chấp nhận được của quy hoạch tuyến tính, mà nghiệm đã nhận được đã cho $z = -13$ nên nó là tối ưu. Vậy nghiệm tối ưu là $x_1 = 2, x_2 = 0, x_3 = 1$ và giá trị mục tiêu tối ưu là $z = -13$.

Các hệ phương trình dạng (1.7), (1.8) mà ta lập ở mỗi bước lặp gọi là các *từ vựng* (dictionary). Trừ z ra, các biến nằm ở vế trái các phương trình (tức là biến “phụ thuộc”) ở mỗi bước lặp gọi là *biến cơ sở* ở bước đó (basic variable). Các biến ở vế phải, tức là biến “độc lập”, được gọi là *biến không cơ sở* (nonbasic variable). Nghiệm nhận được khi cho các biến không cơ sở giá trị 0 được gọi là *nghiệm cơ sở* (basic solution). Vậy mỗi từ vựng xác định một nghiệm cơ sở tương ứng.

1.2.2 Thuật toán đơn hình trên từ vựng

Xét quy hoạch tuyến tính ở dạng chuẩn

$$\begin{aligned} \min \sum_{j=1}^n c_j x_j, \\ \sum_{j=1}^n a_{ij} x_j \leq b_i, \quad i = 1, \dots, m, \\ x_j \geq 0, \quad j = 1, \dots, n. \end{aligned} \quad (1.9)$$

Việc đầu tiên là đưa biến bù vào, đặt tên mục tiêu là z và viết quy hoạch tuyến tính ở dạng từ vựng

$$\begin{aligned} z = \sum_{j=1}^n c_j x_j, \\ w_i = b_i - \sum_{j=1}^n a_{ij} x_j, \quad i = 1, \dots, m. \end{aligned} \quad (1.10)$$

Ta thấy ở thí dụ trên đây khi tiến hành thuật toán đơn hình, biến ban đầu và biến bù được xử lý như nhau, không phân biệt. Do đó ta ký hiệu lại thành một bộ biến x :

$$(x_1, \dots, x_n, w_1, \dots, w_m) := (x_1, \dots, x_n, x_{n+1}, \dots, x_{n+m}).$$

Khi đó (1.10) trở thành

$$\begin{aligned} z = \sum_{j=1}^n c_j x_j, \\ x_{n+i} = b_i - \sum_{j=1}^n a_{ij} x_j, \quad i = 1, \dots, m. \end{aligned}$$

Đây là từ vựng xuất phát. Nội dung của thuật toán đơn hình là chuyển từ một từ vựng sang một từ vựng khác với giá trị mục tiêu tốt hơn. Mỗi từ vựng có m biến cơ sở và n biến không cơ sở. Hãy kí hiệu B là tập các chỉ số tương ứng với các biến cơ sở trong các chỉ số $\{1, 2, \dots, n + m\}$ và N là tập chỉ số của các biến không cơ sở. Ở từ vựng xuất

phát thì $N = \{1, \dots, n\}$ và $B = \{n+1, \dots, n+m\}$, nhưng chúng sẽ thay đổi sau mỗi bước. Ở mỗi bước, từ vựng đều có dạng

$$z = \bar{z} + \sum_{j \in N} \bar{c}_j x_j,$$

$$x_i = \bar{b}_i - \sum_{j \in N} \bar{a}_{ij} x_j, \quad i \in B.$$

Ở đây dấu gạch trên đầu ký tự để chỉ rằng đại lượng này thay đổi qua các bước.

Ở mỗi bước lặp, đúng một biến từ không cơ sở trở thành biến cơ sở, được gọi là *biến vào* (entering variable), và đúng một biến cơ sở trở thành biến không cơ sở, gọi là *biến ra* (leaving variable). Biến vào được chọn trong các biến có hệ số mục tiêu (tức hệ số trong hàm mục tiêu) âm để làm giảm hàm mục tiêu. Nếu không có hệ số mục tiêu âm thì nghiệm nhận được ở bước lặp đó là tối ưu. Nếu có nhiều hệ số mục tiêu âm ta được phép lựa chọn. Bây giờ ta chọn một cách tự nhiên là chọn biến có hệ số (âm) nhỏ nhất để hi vọng làm giảm hàm mục tiêu nhiều nhất. (Khi đó ta vẫn còn độ tự do khi có nhiều hệ số bằng nhau.) Quy tắc này còn gọi là quy tắc xoay Dantzig.

Biến ra được chọn để đảm bảo tính không âm của các biến. Giả sử biến vào đã chọn là x_k , tức là giá trị của nó trở thành dương. Khi đó các biến đang là cơ sở sẽ bị thay đổi và bằng

$$x_i = \bar{b}_i - \bar{a}_{ik} x_k, \quad i \in B.$$

x_k được phép lớn đến mức mọi $x_i \geq 0$, $i \in B$ tức là

$$\frac{1}{x_k} \geq \frac{\bar{a}_{ik}}{\bar{b}_i}, \quad i \in B.$$

Rõ ràng x_k lớn nhất thỏa mọi bất đẳng thức này sẽ là

$$x_k = \left(\max_{i \in B} \frac{\bar{a}_{ik}}{\bar{b}_i} \right)^{-1}.$$

Ở đây ta quy ước $\frac{0}{0} = 0$ và ta sẽ xét sau trường hợp có số $\bar{b}_i = 0$ và trường hợp không có tỉ số $\frac{\bar{a}_{ik}}{\bar{b}_i}$ nào dương. Vậy quy tắc chọn biến ra là chọn biến có chỉ số $l \in B$ mà $\frac{\bar{a}_{lk}}{\bar{b}_k} = \left(\max_{i \in B} \frac{\bar{a}_{ik}}{\bar{b}_i} \right)$ (Ta vẫn còn tự do khi có nhiều tỉ số cùng bằng giá trị cực đại này.)

Sau khi chọn biến vào và biến ra, việc chuyển từ vừng sang từ vừng mới là nhờ các phép toán hàng. Toàn bộ việc làm này gọi là *phép xoay* (pivot). Vì có thể có nhiều biến vào và biến ra có thể lấy đều đảm bảo giảm hàm mục tiêu và các biến vẫn không âm, ta sẽ thấy có các quy tắc cụ thể để tránh sự không xác định đó, được gọi là *quy tắc xoay* (pivot rule).

1.2.3 Tìm từ vừng xuất phát

Ở mục trước ta đã trình bày thuật toán đơn hình, nhưng ta chỉ xét trường hợp về phải các ràng buộc đều không âm; nhờ vậy từ vừng xuất phát là chấp nhận được. (Ta quy ước như vậy nếu nghiệm tương ứng của từ vừng, khi cho các biến không cơ sở bằng 0 và đọc ra ngay được các biến khác, là chấp nhận được.)

Nếu quy hoạch tuyến tính chuẩn (1.9) có các $b_i < 0$ thì ta phải tìm từ vừng xuất phát chấp nhận được bằng một bài toán bổ trợ. Người ta tìm *bài toán bổ trợ* (auxiliary problem) cũng là quy hoạch tuyến tính nhưng có 2 tính chất quan trọng sau đây

(i) dễ thấy ngay từ vừng xuất phát chấp nhận được của bài toán bổ trợ này;

(ii) từ vừng tối ưu của nó cho ngay một từ vừng chấp nhận được của quy hoạch tuyến tính gốc đang cần xét.

Bài toán bổ trợ đó là

$$\begin{aligned} \min x_0, \\ \sum_{j=1}^n a_{ij}x_j - x_0 &\leq b_i, \quad i = 1, \dots, m, \\ x_j &\geq 0, \quad j = 0, 1, \dots, n. \end{aligned}$$

Để có một nghiệm chấp nhận được chỉ việc đặt $x_j = 0, j = 1, \dots, n$ và lấy $x_0 > 0$ đủ lớn. Dễ thấy quy hoạch tuyến tính cần xét có nghiệm chấp nhận được khi và chỉ khi bài toán bổ trợ có nghiệm chấp nhận được với $x_0 = 0$. Nhưng đây rõ ràng là nghiệm tối ưu của bài toán bổ trợ. Vậy, quy hoạch tuyến tính cần xét có nghiệm chấp nhận được khi và chỉ khi bài toán bổ trợ có mục tiêu tối ưu bằng 0.

Mặc dù dễ thấy ngay một nghiệm chấp nhận được của bài toán bổ trợ như trên đã thấy, nhưng ta vẫn chưa có một từ vựng chấp nhận được của nó. Làm sao tìm? Ta hãy minh họa qua một thí dụ:

$$\begin{aligned} \min 2x_1 + x_2, \\ -x_1 + x_2 &\leq -1, \\ -x_1 - 2x_2 &\leq -2, \\ x_2 &\leq 1, \\ x_1, x_2 &\geq 0. \end{aligned}$$

Bài toán bổ trợ là

$$\begin{aligned} \min x_0, \\ -x_1 + x_2 - x_0 &\leq -1, \\ -x_1 - 2x_2 - x_0 &\leq -2, \\ x_2 - x_0 &\leq 1, \\ x_0, x_1, x_2 &\geq 0. \end{aligned}$$

Ta đưa vào biến bù và viết từ vựng xuất phát không chấp nhận được của bài toán bổ trợ (vì có vế phải ràng buộc âm):

$$\begin{aligned}\zeta &= \frac{\quad}{\quad} \frac{x_0}{\quad} \\ w_1 &= -1 + x_1 - x_2 + x_0 \\ w_2 &= -2 + x_1 + 2x_2 + x_0 \\ w_3 &= 1 - x_2 + x_0.\end{aligned}$$

Vì từ vệt này chưa chấp nhận được, ta chưa thực hiện bước lặp đầu bằng phép xoay của thuật toán đơn hình, mà cứ lấy x_0 là biến vào và chọn biến ra là biến w_2 "không chấp nhận được nhất":

$$\begin{aligned}\zeta &= 2 - x_1 - 2x_2 + w_2 \\ w_1 &= 1 - 3x_2 + w_2 \\ x_0 &= 2 - x_1 - 2x_2 + w_2 \\ w_3 &= 3 - x_1 - 3x_2 + w_2.\end{aligned}$$

Đây đã là từ vệt chấp nhận được nên ta bắt đầu thuật toán đơn hình, cụ thể phải đưa x_2 vào và w_1 ra khỏi cơ sở, sau khi xoay ta được

$$\begin{aligned}\zeta &= \frac{4}{3} - x_1 + \frac{2}{3}w_1 + \frac{1}{3}w_2 \\ x_2 &= \frac{1}{3} - \frac{1}{3}w_1 + \frac{1}{3}w_2 \\ x_0 &= \frac{4}{3} - x_1 + \frac{2}{3}w_1 + \frac{1}{3}w_2 \\ w_3 &= 2 - x_1 + w_1.\end{aligned}$$

Ở bước lặp tiếp theo ta đưa x_1 vào và x_0 ra khỏi cơ sở:

$$\begin{aligned}\zeta &= 0 + x_0 \\ x_2 &= \frac{1}{3} - \frac{1}{3}w_1 + \frac{1}{3}w_2 \\ x_1 &= \frac{4}{3} - x_0 + \frac{2}{3}w_1 + \frac{1}{3}w_2 \\ w_3 &= \frac{2}{3} + x_0 + \frac{1}{3}w_1 - \frac{1}{3}w_2.\end{aligned}$$

Vì không còn hệ số âm ở hàm mục tiêu, từ vệt này là tối ưu. Để được từ vệt chấp nhận được tương ứng của quy hoạch tuyến tính cần

xét ban đầu, ta bỏ x_0 khỏi từ vựng này (vì giá trị mục tiêu tối ưu là $\zeta = x_0 = 0$) và tính giá trị tương ứng của mục tiêu z :

$$z = 2x_1 + x_2 = 3 + w_1 + w_2.$$

Vậy từ vựng chấp nhận được của quy hoạch tuyến tính cần xét là

$$\begin{aligned}\zeta &= 3 + w_1 + w_2 \\ x_2 &= \frac{1}{3} - \frac{1}{3}w_1 + \frac{1}{3}w_2 \\ x_1 &= \frac{4}{3} + \frac{2}{3}w_1 + \frac{1}{3}w_2 \\ w_3 &= \frac{2}{3} + \frac{1}{3}w_1 - \frac{1}{3}w_2.\end{aligned}\tag{1.11}$$

May mắn, một cách vô tình, ở thí dụ này là từ vựng này đã tối ưu, vì các hệ số mục tiêu đều dương. Nói chung thì ta chỉ được từ vựng xuất phát chấp nhận được.

Thông thường người ta ghép luôn việc giải bài toán bổ trợ vào thuật toán đơn hình và gọi là *pha I* (phase I). Quá trình trực tiếp áp dụng thuật toán đơn hình từ từ vựng xuất phát chấp nhận được nhận từ pha I gọi là *pha II* (phase II). Toàn bộ thuật toán gọi là *thuật toán hai pha* (two - phase algorithm).

1.2.4 Trường hợp mục tiêu không giới nội

Tất nhiên có quy hoạch tuyến tính mà hàm mục tiêu có thể dẫn đến $-\infty$ trong miền chấp nhận được. Trường hợp này sẽ không có nghiệm tối ưu.

Làm sao phát hiện ra là quy hoạch tuyến tính đang xét có hàm mục tiêu không giới nội như vậy? Câu trả lời là: có thể phát hiện được, và ngay trong quá trình thực hiện thuật toán đơn hình.

Ta hãy xét kĩ mấy trường hợp “bất thường” mà khi xây dựng thuật toán đơn hình ở Mục 1.2.2 cho trường hợp “bình thường” thuận lợi ta

đã phải gác lại. Ở đây, ta xét trường hợp không có tỉ số $\frac{\bar{a}_{ik}}{\bar{b}_i}$, $i \in B$, nào dương. (Trường hợp có $\bar{b}_i = 0$ với $i \in B$ sẽ được xét ở Chương 3 dưới đây.) Tỉ số này gặp phải khi tìm biến ra sau khi đã xác định biến x_k là biến vào, tức là tăng từ 0 lên một số dương. Nhớ rằng lúc đó các biến cơ sở là

$$x_i = \bar{b}_i - \bar{a}_{ik}x_k, \quad i \in B.$$

Mà $\bar{b}_i$ và $-\bar{a}_{ik}$ là cùng dấu (vì $\frac{\bar{a}_{ik}}{\bar{b}_i} \leq 0$) và là không âm. Do đó mọi biến cơ sở x_i không thể từ không âm trở thành âm. Vậy biến vào có thể lấy giá trị lớn tùy ý để hàm mục tiêu tiến tới $-\infty$. Lúc này ta nói là hàm mục tiêu không giới nội (dưới), hoặc bài toán không giới nội.

Chẳng hạn, xét từ vừng

$$\begin{aligned} z &= 7 - x_2 + x_1 \\ x_3 &= 2 + 2x_2 - x_1 \\ x_4 &= 6 + 3x_1 \\ x_5 &= 2x_1. \end{aligned} \tag{1.12}$$

Biến vào là x_2 và các tỉ số $\frac{\bar{a}_{ik}}{\bar{b}_i}$ là $-\frac{2}{2}$, $-\frac{0}{6}$ và $\frac{0}{0}$. Vì không có tỉ số nào dương nên bài toán là không giới nội.

1.2.5 Tìm tất cả nghiệm tối ưu

Khi đã biết là có tồn tại nghiệm tối ưu của một quy hoạch tuyến tính đang xét, chẳng hạn khi đã có một từ vừng tối ưu (tức là từ vừng cho ta nghiệm tối ưu), thì vấn đề tìm mọi nghiệm tối ưu cũng là quan trọng hay được đặt ra trong thực tế.

Thuật toán đơn hình có một ưu việt là cho phép ta tính mọi nghiệm tối ưu, nói riêng là xét được tính duy nhất nghiệm một cách đơn giản.

Chẳng hạn, xét từ vựng tối ưu (1.11). Phương trình z (tức là phương trình đầu) cho thấy mọi nghiệm chấp nhận được có giá trị mục tiêu $z = 3$, tức là tối ưu, đều phải thỏa $w_1 = 0$, $w_2 = 0$. Ba phương trình sau cho thấy nếu $w_1 = 0$, $w_2 = 0$ thì bắt buộc $x_2 = \frac{1}{3}$, $x_1 = \frac{4}{3}$ và $w_3 = \frac{2}{3}$. Vậy nghiệm tối ưu là xác định duy nhất.

Ở Mục 1.1.2 ta đã thấy bằng hình học là quy hoạch tuyến tính có thể có vô số nghiệm tối ưu. Thuật toán đơn hình cho phép ta mô tả tất cả các nghiệm tối ưu khá đơn giản như sau.

THÍ DỤ 1.2.1 Giả sử (bằng thuật toán đơn hình) ta nhận được từ vựng tối ưu sau

$$\begin{aligned} z &= 5 \quad + x_3 \dots \\ w_1 &= 3 + x_2 - 2w_2 + 7x_3 \\ x_1 &= 1 - 5x_2 + 6w_2 - 8x_3 \\ w_3 &= 4 + 9x_2 + 2w_2 - x_3. \end{aligned}$$

Dòng đầu tiên cho thấy mỗi nghiệm tối ưu phải thỏa điều kiện $x_3 = 0$ (nhưng không cần $x_2 = 0$ hoặc $w_2 = 0$). Với nghiệm như vậy, các dòng dưới của từ vựng cho ta

$$\begin{aligned} w_1 &= 3 + x_2 - 2w_2, \\ x_1 &= 1 - 5x_2 + 6w_2, \\ w_3 &= 4 + 9x_2 + 2w_2. \end{aligned}$$

Do đó, để nhận được nghiệm tối ưu ta có thể cho x_2 và w_2 giá trị không âm tùy ý miễn là đảm bảo $w_1 \geq 0$, $x_1 \geq 0$ và $w_3 \geq 0$, tức là x_2 và w_2 chỉ cần thỏa

$$-x_2 + 2w_2 \leq 3,$$

$$5x_2 - 6w_2 \leq 1,$$

$$-9x_2 - 2w_2 \leq 4,$$

$$x_2, w_2 \geq 0.$$

(Thực ra bất đẳng thức $-9x_2 - 2w_2 \leq 4$ là thừa vì nó luôn đúng với $x_2 \geq 0, w_2 \geq 0$.)

1.2.6 Dạng bảng của phương pháp đơn hình

Trên đây ta đã dùng từ vựng để ghi kết quả ở mỗi bước lặp của thuật toán đơn hình. Một cách khác, thậm chí phổ biến hơn là dùng bảng biểu diễn như sau. Để tiện lợi ta xét lại thí dụ cũ là quy hoạch tuyến tính (1.5). Ta viết lại dạng chính tắc của bài toán như sau (mục tiêu được ghi ở dạng một phương trình)

$$\begin{array}{rcll} -z - 5x_1 - 4x_2 - 3x_3 & & & = 0 \\ 2x_1 + 3x_2 + x_3 + w_1 & & & = 5 \\ 4x_1 + x_2 + 2x_3 & + w_2 & & = 11 \\ 3x_1 + 4x_2 + 2x_3 & & + w_3 & = 8. \end{array} \quad (1.13)$$

Với quy ước kết quả ở mỗi bước đều viết ở dạng như vậy, ta chỉ cần ghi các hệ số là đủ. Ta sẽ được bảng sau (hoàn toàn xác định bài toán (1.5))

x_1	x_2	x_3	w_1	w_2	w_3		
-5	-4	-3	0	0	0	0	
2	3	1	1	0	0	5	(1.14)
4	1	2	0	1	0	11	
3	4	2	0	0	1	8	

Theo cách viết bảng như vậy thì từ vừng (1.7) đạt được sau bước 1 sẽ là

$$\begin{array}{cccccc|c}
 x_1 & x_2 & x_3 & w_1 & w_2 & w_3 & \\
 \hline
 0 & \frac{7}{2} & -\frac{1}{2} & \frac{5}{2} & 0 & 0 & \frac{25}{2} \\
 1 & \frac{3}{2} & \frac{1}{2} & \frac{1}{2} & 0 & 0 & \frac{5}{2} \\
 0 & -5 & 0 & -2 & 1 & 0 & 1 \\
 0 & -\frac{1}{2} & \frac{1}{2} & -\frac{3}{2} & 0 & 1 & \frac{1}{2}
 \end{array} \quad (1.15)$$

Ta hãy xem bảng (1.15) nhận được từ bảng (1.14) như thế nào. Nói cách khác ta hãy xem bước 1 tiến hành trên bảng (1.14) thế nào để nhận được bảng (1.15).

Để chọn biến vào ta lấy số âm lớn nhất ở hàng đầu không kể số tận cùng bên phải (là giá trị của $-z$). Trong thí dụ này là -5 , và x_1 là biến vào. Khi đó cột đầu được gọi là *cột xoay* (pivot column). Chú ý là nếu không có số âm ở hàng đầu thì ta kết luận bảng hiện tại là tối ưu và viết được ra ngay nghiệm tối ưu tương ứng của quy hoạch tuyến tính. Thuật toán kết thúc.

Tiếp theo là tìm biến ra. Theo lí luận bảng từ vừng ở Mục 1.2.1 ta phải xét các hàng i mà phần tử a_{ik} của nó trên cột xoay k là dương. Nếu mọi phần tử trên cột xoay là âm, thì theo Mục 1.2.4, quy hoạch tuyến tính không có nghiệm tối ưu vì hàm mục tiêu không giới nội dưới trên miền chấp nhận được. Ngược lại, ta phải chọn hàng $i = l$ mà $\frac{\bar{a}_{lk}}{b_l}$ là lớn nhất, tức là tỉ số giữa phần tử trên cột xoay và phần tử cùng hàng nhưng ở cuối hàng là lớn nhất. Trong thí dụ của ta là hàng 1 (không kể hàng mục tiêu trên cùng):

x_1	x_2	x_3	w_1	w_2	w_3	
- 5	- 4	- 3	0	0	0	0
2	3	1	1	0	0	5
4	1	2	0	1	0	11
3	4	2	0	0	1	8

Hàng l được chọn ra được gọi là *hàng xoay* (pivot row) và phần tử ở giao của cột xoay và hàng xoay được gọi là *phần tử trục* (pivot).

Tiếp đến, như đã thấy ở Mục 1.2.1 ta phải thực hiện phép toán trên hàng. Thực hiện điều này trên bảng sẽ là chia hàng xoay cho phần tử trục để biến phần tử trục thành 1; rồi lấy mỗi hàng khác trừ đi hàng xoay mới, nhân với số trên cột xoay của hàng đó để biến cột xoay thành cột các số 0 ngoại trừ phần tử trục. Đến đây ta được bảng (1.15) và kết thúc một bước lặp.

Dạng bảng viết có gọn hơn dạng từ vựng nhưng không tường minh bằng, nên sau đây ta vẫn sử dụng từ vựng.

Các thuật ngữ “phép xoay, cột xoay, hàng xoay” có thể hiểu là liên quan đến việc đưa một biến vào và lấy một biến ra ở mỗi bước lặp tương tự như ta “xoay” từ vựng hoặc bảng đơn hình đang có ở bước đó. Thực ra các thuật ngữ này, cùng với chính tên của thuật toán là “đơn hình” có hình ảnh hình học rõ ràng gắn liền với phép xoay thực sự xem [Khánh - Nường 2000] Mục 4.6.

1.2.7 Hiện tượng suy biến

Ta xét trường hợp “bất thường” còn lại chưa giải quyết, nhưng đã nhắc đến, là trường hợp khi chọn biến ra ta gặp tỉ số $\frac{a_{ik}}{b_i}$ không xác định $\bar{b}_i = 0$.

Ta nói một từ vựng là *suy biến* (degenerate) nếu có $i \in \bar{B}$ (tập chỉ

số của các biến cơ sở) với $b_1 = 0$. Một từ vựng suy biến có thể gây khó khăn khi ta thực hiện tiếp thuật toán đơn hình trên nó, nhưng cũng có thể không. Chẳng hạn, nếu ta nhìn lại từ vựng (1.12)

$$\begin{aligned} z &= 7 - x_2 + x_1 \\ x_3 &= 2 + 2x_2 - x_1 \\ x_4 &= 6 + 3x_1 \\ x_5 &= 2x_1. \end{aligned}$$

Đây là từ vựng suy biến vì $\bar{B} = \{3, 4, 5\}$ và $\bar{b}_5 = 0$. Nhưng như ta đã biết ở

Mục 1.2.4, vì trong các tỉ số $\frac{\bar{a}_{ik}}{\bar{b}_i}$ là $-\frac{2}{2}$, $-\frac{0}{6}$ và $\frac{0}{0}$, không có tỉ số nào dương nên

bài toán là không giới nội. Nếu bây giờ ta gặp bài toán khác có từ vựng hoàn toàn như vậy, chỉ khác là $2x_2$ ở phương trình của x_3 lại là $-2x_2$, thì theo quy tắc chọn biến ra, x_2 là biến ra. Tiếp theo, với phép toán trên hàng để "xoay" ta được từ vựng mới

$$\begin{aligned} z &= 6 + \frac{1}{2}x_3 + 2x_1 \\ x_2 &= 1 + 3x_1 \\ x_4 &= 6 + 3x_1 \\ x_5 &= 2x_1. \end{aligned}$$

Đây vẫn là một từ vựng suy biến, nhưng đã là tối ưu, nên ta được ngay nghiệm tối ưu của quy hoạch tuyến tính $(x_1, x_2, x_3, x_4, x_5) = (0, 1, 0, 6, 0)$ với mục tiêu tối ưu là $z = 6$. Nhận xét rằng ở từ vựng "khác" vừa xét các tỉ số $\frac{\bar{a}_{ik}}{\bar{b}_i}$ là $\frac{2}{2} - \frac{0}{6}$ và $\frac{0}{0}$, không có tỉ số nào là $+\infty$. Chính giá trị $+\infty$ này sẽ gây rắc rối như ta thấy bây giờ.

Một phép xoay được gọi là *phép xoay suy biến* (degenerate pivot)

nếu có tỉ số $\frac{\bar{a}_{ik}}{b_i}$, khi xác định biến ra, là $+\infty$, tức là tử số dương và mẫu bằng 0. Khi đó ta cũng nói là từ vưng suy biến đã sinh ra phép xoay suy biến. Như vậy theo suy luận trên đây, từ vưng suy biến sẽ không gây khó khăn gì cho thuật toán đơn hình nếu nó không sinh ra phép xoay suy biến, còn nếu nó sinh ra phép xoay suy biến thì có thể phiên toái. Ta xét hai thí dụ sau đây.

THÍ DỤ 1.2.1 Xét từ vưng

$$\begin{aligned} z &= -3 + 0,5x_1 - 2x_2 + 1,5w_1 \\ x_3 &= 1 - 0,5x_1 - 0,5w_1 \\ w_2 &= x_1 - x_2 + w_1. \end{aligned}$$

Từ vưng này suy biến. Biến vào là x_2 và các tỉ số để xác định biến ra là 0 và $+\infty$. Vì cực đại trong đó là $+\infty$ nên biến ra là w_2 . Nhưng, dễ thấy tỉ số tương ứng là $+\infty$ dẫn đến hiện tượng là x_2 vừa từ 0 trở thành dương thì w_2 trở thành âm, không chấp nhận được! Vì vậy mặc dù x_2 từ ngoài trở thành biến cơ sở nhưng từ vưng chấp nhận được tiếp theo có nghiệm tương ứng vẫn với $x_2 = 0$. Ta thấy điều đó sau khi thực hiện phép xoay (suy biến):

$$\begin{aligned} z &= -3 - 1,5x_1 + 2w_2 - 0,5w_1 \\ x_3 &= 1 - 0,5x_1 - 0,5w_1 \\ x_2 &= x_1 - w_2 + w_1. \end{aligned} \tag{1.16}$$

Nhận xét rằng, chính do phép xoay là suy biến nên từ vưng mới này vẫn có nghiệm chấp nhận được $(x_1, x_2, x_3, w_1, w_2) = (0, 0, 1, 0, 0)$ và giá trị mục tiêu tương ứng $z = -3$ hoàn toàn như cũ. Nhưng dù sao từ vưng đã thay đổi, hi vọng là phép xoay tiếp theo sẽ cải thiện tình hình, nếu chưa thì phép xoay sau đó Cụ thể, ta thấy từ vưng (1.16) vẫn suy biến, nhưng x_1 là biến vào và x_3 là biến ra đã cho ta phép xoay không suy biến, dẫn đến từ vưng

$$\begin{aligned} z &= -6 + 3x_3 + 2w_2 + w_1 \\ x_1 &= 2 - 2x_3 - w_1 \\ x_2 &= 2 - 2x_3 - w_2 \end{aligned}$$

Từ vệt này là tối ưu.

Trong thực tế hơn nửa thế kỉ áp dụng phương pháp đơn hình người ta thấy thường gặp từ vệt suy biến nhưng sau một số bước lặp luôn có được phép xoay không suy biến làm thoát ra khỏi sự suy biến. Tuy nhiên về mặt lí thuyết, vẫn có thể xảy ra sự cố là sau một số phép xoay suy biến ta lại quay lại từ vệt suy biến đã gặp và thuật toán đơn hình lâm vào một vòng lặp vô hạn, không bao giờ tìm được nghiệm tối ưu. Trường hợp này gọi là *sự xoay vòng* (cycling).

Thí dụ đầu tiên về xoay vòng được tìm thấy ở [Hoffman 1953]. Ở nước ta, thí dụ xoay vòng đầu tiên là của GS Trần Vũ Thiệu (Viện Toán học, Hà Nội), năm 1967. [Marshall - Suurballe 1969] chứng minh rằng quy hoạch tuyến tính (chính tắc) xoay vòng phải có ít nhất sáu biến và ba ràng buộc.

Định lí sau đây khẳng định rằng thuật toán đơn hình luôn kết thúc được sau hữu hạn bước, chỉ trừ trường hợp bị xoay vòng.

ĐỊNH LÍ 1.2.1. *Nếu thuật toán đơn hình kéo dài vô hạn thì nó phải xoay vòng.*

CHỨNG MINH. Một từ vệt được hoàn toàn xác định khi ấn định các biến nào là biến cơ sở, biến nào là không cơ sở. Có tất cả $C_{n+m}^m \equiv \binom{n+m}{m}$ khả năng ấn định khác nhau. Nếu thuật toán kéo dài vô hạn nó phải gặp lại một từ vệt hơn một lần. Ngay từ khi nó lặp lại một từ vệt nào đó, ta phải có sự xoay vòng từ đây. $\square$

CHÚ Ý. Tất cả các phép xoay trong một vòng, khi có hiện tượng xoay vòng, đều phải là phép xoay suy biến và các từ vệt trong vòng có cùng giá trị hàm mục tiêu và cùng nghiệm chấp nhận được tương ứng.

Quy hoạch tuyến tính xoay vòng hiểm đến nỗi hầu hết các thể nghiệm cụ thể phương pháp đơn hình trong thực tế, để giải các bài toán cỡ lớn là chính (các thuật toán, các phần mềm) đều không chú ý đến phòng ngừa xoay vòng. Tuy vậy, chỉ ra các nguyên tắc xoay để phương pháp đơn hình tránh được xoay vòng vẫn là quan trọng. [Bland 1977] đã đưa ra quy tắc xoay sau đây để tránh xoay vòng.

Quy tắc Bland: Biến vào được chọn là biến có chỉ số nhỏ nhất trong các biến có hệ số mục tiêu âm; biến ra được chọn là biến có chỉ số nhỏ nhất trong số các biến có chỉ số $l \in B$ thỏa quy tắc chọn biến ra là

$$\frac{\bar{a}_{lk}}{\bar{b}_l} = \left(\max_{i \in B} \frac{\bar{a}_{ik}}{\bar{b}_i} \right).$$

ĐỊNH LÝ 1.2.2. Thuật toán đơn hình sẽ kết thúc sau hữu hạn bước, tức là không bị xoay vòng, nếu áp dụng quy tắc xoay Bland.

Độc giả có thể xem chứng minh trong [Khánh - Nương 2000]. Phương pháp Bland để chống xoay vòng có ý nghĩa lý thuyết nhưng không tiện áp dụng, nhất là khi viết phần mềm. Người ta hay dùng phương pháp nhiễu loạn và phương pháp từ điển để chống xoay vòng (xem sách nói trên).

1.3 NGHIÊN CỨU HÌNH HỌC QUY HOẠCH TUYẾN TÍNH

1.3.1 Nghiệm cơ sở và điểm cực biên

Ở Mục 1.1.2 ta đã giải quy hoạch tuyến tính hai biến bằng hình học và có nhận xét quan trọng, mà theo cảm nhận trực quan là đúng cả cho trường hợp n biến tổng quát, là nếu quy hoạch tuyến tính có nghiệm tối ưu và miền chấp nhận được (là tập lồi đa diện) có đỉnh thì quy hoạch tuyến tính luôn có nghiệm tối ưu là đỉnh. Chính tính

chất này là cơ sở của phương pháp đơn hình. Nội dung của phương pháp là xuất phát từ một đỉnh của miền chấp nhận được; ở mỗi bước lặp ta đi sang một đỉnh kế nó có giá trị hàm mục tiêu nhỏ hơn. Vì số đỉnh là hữu hạn, sau hữu hạn bước ta sẽ đến đỉnh tối ưu.

Tuy vậy, trong hai chương vừa qua ta lại trình bày phương pháp đơn hình bằng lý luận giải tích, dựa vào biểu diễn từ vựng để lập luận đơn giản hơn. Trong chương này ta sẽ xem xét vấn đề từ hình học và sẽ thấy rõ sự tương đương của các cách tiếp cận.

Trước hết, ta phân tích khái niệm nghiệm cơ sở, nền tảng của phương pháp đơn hình. Ta nhắc lại định nghĩa bằng ngôn ngữ từ vựng đã nói ở Mục 1.2.1.

ĐỊNH NGHĨA 1.3.1. *Nghiệm cơ sở* (basic solution) của quy hoạch tuyến tính ở dạng từ vựng là nghiệm nhận được khi cho các biến không cơ sở giá trị 0 và giải ra các biến cơ sở bằng từ vựng. Ta cũng nói đây là nghiệm cơ sở ứng với từ vựng này. Nghiệm cơ sở thoả thêm các ràng buộc dấu $x_j \geq 0$ được gọi là nghiệm cơ sở chấp nhận được. Lúc này ta cũng nói từ vựng tương ứng là chấp nhận được.

Bây giờ ta xét các định nghĩa khác và sẽ thấy rằng nghiệm cơ sở chấp nhận được là khái niệm hình học gắn liền với quy hoạch tuyến tính không phụ thuộc biểu diễn giải tích của nó. Tổng quát hơn từ vựng là dạng chính tắc của quy hoạch tuyến tính

$$\begin{aligned} \min z &= c^T x, \\ Ax &= b, \\ x &\geq 0, \end{aligned} \tag{1.17}$$

ở đây, A là ma trận cấp $m \times n$, với $m < n$, mà không mất tổng quát giả sử là có $\text{rank } A = m$ (vì nếu có một hàng là tổ hợp tuyến tính của các hàng khác thì bỏ nó đi vẫn được hệ ràng buộc tương đương, tức là xác định cùng một tập lồi đa diện).

ĐỊNH NGHĨA 1.3.2. Nghiệm cơ sở của hệ $Ax = b$ hoặc nói khác là của quy hoạch tuyến tính dạng chính tắc (1.17) là nghiệm nhận được bằng cách cho $n - m$ biến x_j nào đó bằng 0 sao cho giải ra được duy nhất m biến còn lại.

Rõ ràng, Định nghĩa 1.3.1 cho từ vựng chỉ là trường hợp riêng khi $Ax = b$ đã giải sẵn ra m biến theo $n - m$ biến còn lại, tức là $Ax = b$ có dạng đặc biệt

$$x_i = \bar{b}_i - \sum_{j \in N} a_{ij} x_j, \quad i \in B. \quad (1.18)$$

Trở lại trường hợp chính tắc tổng quát $Ax = b$, một cách tự nhiên ta gọi m cột của A ứng với các biến giải ra sau đó là cơ sở (basis) của quy hoạch tuyến tính dạng chính tắc ứng với nghiệm cơ sở đó tức là ta có

ĐỊNH NGHĨA 1.3.3. Cơ sở của nghiệm cơ sở $(x_1^0, \dots, x_n^0)$ của quy hoạch tuyến tính chính tắc (1.17) là hệ m vectơ cột $\{A_i : i \in B\}$ độc lập tuyến tính của A sao cho

$$B \supset I := \{i : x_i^0 \neq 0\}.$$

Bây giờ xét trường hợp suy biến. Ta đã gọi từ vựng là suy biến nếu có $b_i = 0$ (ở cột tự do về phải). Một cách tương đương, nghiệm cơ sở tương ứng cũng gọi là suy biến. Nói cách khác nghiệm cơ sở là *suy biến* (degenerate) nếu nó có hơn $n - m$ thành phần bằng 0, và là *không suy biến* (nondegenerate) nếu nó có đúng $n - m$ thành phần bằng 0. Định nghĩa này về suy biến rõ ràng là áp dụng được cho quy hoạch tuyến tính chính tắc tổng quát, không cần ở dạng từ vựng.

So với định nghĩa cơ sở ta thấy:

Nếu nghiệm cơ sở là không suy biến thì chỉ ứng với một cơ sở duy nhất và nếu biểu diễn quy hoạch tuyến tính ở dạng từ vựng, thì ứng với một từ vựng duy nhất.

Nếu nghiệm cơ sở là suy biến, thì nó có thể có nhiều cơ sở và ứng với nhiều từ vựng khác nhau. Chẳng hạn, khi nó xoay vòng, thì mọi từ vựng trong vòng xoay cùng ứng với một nghiệm cơ sở.

Bây giờ ta định nghĩa nghiệm cơ sở cho quy hoạch tuyến tính tổng quát.

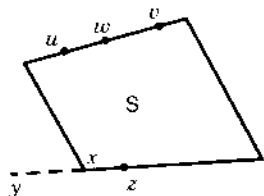
ĐỊNH NGHĨA 1.3.4. Nghiệm cơ sở của quy hoạch tuyến tính $\min c^T x$, $x \in P$, ở đây P xác định bởi các đẳng thức và bất đẳng thức tuyến tính, hoặc gọi khác là nghiệm cơ sở của tập lồi đa diện P , là $x \in \mathbb{R}^n$ thoả mọi ràng buộc đẳng thức (nếu có) và thoả chặt thêm một số ràng buộc bất đẳng thức để có đủ n ràng buộc thoả với dấu = độc lập tuyến tính. Nghiệm cơ sở là không suy biến nếu nó thoả chặt đúng n ràng buộc và là suy biến nếu nó thoả chặt hơn n ràng buộc.

Rõ ràng định nghĩa này quy về Định nghĩa 1.3.2 nếu P có dạng chính tắc vì theo Định nghĩa 1.3.2 thì nghiệm cơ sở thoả chặt m ràng buộc đẳng thức độc lập tuyến tính và $n - m$ ràng buộc bất đẳng thức $x_j \geq 0$, rõ ràng là độc lập tuyến tính với m ràng buộc trên.

Đến đây ta chuyển sang định nghĩa hình học như sau.

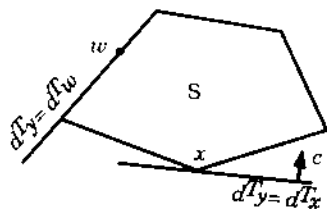
ĐỊNH NGHĨA 1.3.5. (a) Điểm x của tập lồi $S \subset \mathbb{R}^n$ được gọi là *điểm cực biên* (extreme point) của S nếu x không thể là tổ hợp lồi của hai điểm của S khác x , tức là không có $y, z \in S$, $\gamma \in (0, 1)$ để $x = (1 - \gamma)y + \gamma z$.

(b) Điểm x của tập lồi $S \subset \mathbb{R}^n$ được gọi là *đỉnh* (vertex) của S nếu có một siêu phẳng chỉ cắt S ở x và toàn bộ S nằm ở một phía của siêu phẳng, tức là nếu tồn tại $d \in \mathbb{R}^n$, d khác 0, sao cho $d^T x < d^T y$, $\forall y \in S$, y khác x .



H. 1.7 : x là điểm cực biên
 w không phải là điểm cực biên

Định nghĩa trên được minh hoạ bởi H. 1.7 và H. 1.8. Qua đây ta thấy "cực biên" có nghĩa là "biên cùng cực", "biên nhất có thể được". Còn đỉnh thì sao? Định lí sau đây cho câu trả lời.



H. 1.8 : x là đỉnh
 w không phải là đỉnh

ĐỊNH LÝ 1.3.1. Giả sử P là một tập lồi đa diện và $x^* \in P$. Khi đó ta có ba khẳng định tương đương

- (a) x^* là đỉnh;
- (b) x^* là điểm cực biên;
- (c) x^* là nghiệm cơ sở chấp nhận được.

CHỨNG MINH. (a) $\Rightarrow$ (b). Giả sử x^* là đỉnh, tức là có $d \neq 0$ để $d^T x^* < d^T y \forall y \in P, y \neq x^*$. Với hai điểm y và z khác x^* bất kì của P và với $\gamma \in [0, 1]$ bất kì thì $d^T x^* < d^T y, d^T x^* < d^T z$. Do đó $d^T x^* < d^T(\gamma y + (1 - \gamma)z)$. Vì vậy $x^* \neq \gamma y + (1 - \gamma)z$, tức x^* là điểm cực biên.

(b) $\Rightarrow$ (c). Giả sử x^* không phải nghiệm cơ sở chấp nhận được. Ta sẽ chứng tỏ x^* không phải điểm cực biên của P . Giả sử P có biểu diễn là $\{x \in \mathbb{R}^n: a_i^T x \geq b_i; i = 1, \dots, m\}$. Gọi $I = \{i: a_i^T x^* = b_i\}$. Vì x^* không phải là nghiệm cơ sở chấp nhận được, nên trong các vectơ $a_i, i \in I$, không đủ n vectơ độc lập tuyến tính. Do đó tồn tại $d \in \mathbb{R}^n, d \neq 0$ sao cho $a_i^T d = 0, \forall i \in I$. Xét $y = x^* + \varepsilon d$ và $z = x^* - \varepsilon d$. Ta thấy $a_i^T y = a_i^T x^* = b_i, \forall i \in I$. Còn với $i \notin I$ thì $a_i^T x^* > b_i$, nên với ε đủ nhỏ $a_i^T y > b_i$. Vậy $y \in P$ với $\varepsilon > 0$ đủ nhỏ. Tương tự $z \in P$ với $\varepsilon > 0$ đủ nhỏ. Mà $x^* = \frac{1}{2}(y + z)$ nên x^* không phải là điểm cực biên.

(c) $\Rightarrow$ (a). Giả sử $P = \{x \in \mathbb{R}^n: a_i^T x \geq b_i, i = 1, \dots, m\}$ và x^* là nghiệm cơ sở chấp nhận được. Gọi $I = \{i: a_i^T x^* = b_i\}$ và $d = \sum_{i \in I} a_i$. Khi đó

$$d^T x^* = \sum_{i \in I} a_i^T x^* = \sum_{i \in I} b_i.$$

Với bất kì $x \in P$ ta có

$$d^T x = \sum_{i \in I} a_i^T x \geq \sum_{i \in I} b_i. \quad (1.19)$$

Vì có n vectơ độc lập tuyến tính trong các vectơ $a_i, i \in I$, nên hệ $a_i^T x = b_i, i \in I$, có duy nhất nghiệm là x^* . Do đó, với $x \in P, x \neq x^*$ ta có $d^T x > d^T x^*$ (do (1.19)). Vậy x^* là đỉnh. $\square$

Do định lý trên, nghiệm cơ sở chấp nhận được của P , cũng như đỉnh và điểm cực biên, không phụ thuộc vào dạng cụ thể của hệ đẳng thức và bất đẳng thức biểu diễn P (khác với nghiệm cơ sở).

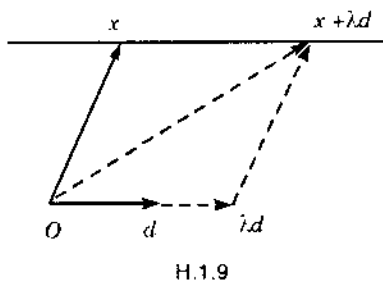
MỆNH ĐỀ 1.3.2. *Mỗi hệ ràng buộc tuyến tính chỉ có hữu hạn nghiệm cơ sở. Do đó nó cũng chỉ có hữu hạn nghiệm cơ sở chấp nhận được.*

CHỨNG MINH. Mỗi nghiệm cơ sở thỏa chất n ràng buộc độc lập tuyến tính. Vì hệ ràng buộc chất này có nghiệm duy nhất, nên nghiệm cơ sở khác nhau ứng với hệ ràng buộc chất độc lập tuyến tính khác nhau. Vì chỉ có số hữu hạn cách chọn n ràng buộc từ số m ràng buộc của hệ, nên chỉ có hữu hạn nghiệm cơ sở. $\square$

1.3.2 Sự tồn tại đỉnh và đỉnh tối ưu

Ở Mục 1.1.2 ta đã biết là tập lồi đa diện có thể không có đỉnh. Bây giờ ta xét điều kiện để tồn tại đỉnh.

Chú ý rằng một đường thẳng trong $\mathbb{R}^n$ qua điểm $x \in \mathbb{R}^n$ có thể biểu diễn ở dạng $\{x + \lambda d: \lambda \in \mathbb{R}\}$, ở đây $d \neq 0, d \in \mathbb{R}^n$ chính là vectơ chỉ phương của đường thẳng (H.1.9).



ĐỊNH LÝ 1.3.3. *Giả sử tập lồi đa diện $P = \{x: a_i^T x \geq b_i, i = 1, \dots, m\}$ khác trống. Khi đó các kết luận sau là tương đương:*

- (a) P có đỉnh;
- (b) P không chứa một đường thẳng nào;
- (c) trong các vectơ $a_1, \dots, a_m$ có n vectơ độc lập tuyến tính.

40 CHƯƠNG 1 QUY HOẠCH TUYẾN TÍNH. PHƯƠNG PHÁP ĐƠN HÌNH

CHỨNG MINH. (a) $\Rightarrow$ (c). Nếu x là đỉnh thì nó thỏa n ràng buộc chặt độc lập tuyến tính trong số $a_i^T x \geq b_i, i = 1, \dots, m$. Khi đó n vectơ a_i tương ứng là độc lập tuyến tính.

(c) $\Rightarrow$ (b). Không mất tổng quát giả sử n vectơ độc lập tuyến tính là $a_1, \dots, a_n$. Giả sử phản chứng là P chứa đường thẳng $\{x + \lambda d: \lambda \in \mathbf{R}\}$, ở đây $d \neq 0$. Khi đó $a_i^T(x + \lambda d) \geq b_i, i = 1, \dots, m$, với mọi λ . Từ đây ta thấy $a_i^T d = 0, i = 1, \dots, m$, (vì giả sử có $a_k^T d > 0$ thì lấy $\lambda \rightarrow -\infty$ ta có $a_k^T(x + \lambda d) \rightarrow -\infty$ mâu thuẫn với bất đẳng thức $a_k^T(x + \lambda d) \geq b_k$ với mọi λ ; tương tự $a_k^T d < 0$ không thể xảy ra). Nhưng $a_1, \dots, a_n$ độc lập tuyến tính nên từ $a_i^T d = 0, i = 1, \dots, n$, suy ra $d = 0$, trái với định nghĩa d .

(b) $\Rightarrow$ (a). Giả sử $x \in P$ và $I := \{i: a_i^T x = b_i\}$. Nếu có n vectơ $a_i, i \in I$, độc lập tuyến tính, thì x là nghiệm cơ sở chấp nhận được, tức là điểm cực biên rồi.

Trường hợp ngược lại thì mọi vectơ $a_i, i \in I$ nằm trong một không gian con thực sự của $\mathbf{R}^n$. Do đó sẽ có $d \in \mathbf{R}^n, d \neq 0$, vuông góc với chúng, tức là $a_i^T d = 0, i \in I$. Xét đường thẳng $\{x + \lambda d: \lambda \in \mathbf{R}\}$. Ta có

$$a_i^T(x + \lambda d) = a_i^T x + \lambda a_i^T d = a_i^T x = b_i, \forall i \in I, \forall \lambda \in \mathbf{R}.$$

tức là các ràng buộc chặt tại x vẫn thỏa chặt trên cả đường thẳng. Nhưng P không chứa đường thẳng nên khi thay đổi λ để đi theo đường thẳng phải đến một giá trị λ^* để qua đó có ràng buộc j (chỉ số j không thuộc I) bị vi phạm. Vậy điểm ranh giới phải thỏa chặt ràng buộc j , tức là $a_j^T(x + \lambda^* d) = b_j$.

Ta sẽ chỉ rằng a_j không phải tổ hợp tuyến tính của các $a_i, i \in I$. Ta có $a_j^T x \neq b_j$ (vì $j \notin I$). Do đó $a_j^T d \neq 0$, tức là d không vuông góc với a_j mà lại vuông góc với mọi $a_i, i \in I$. Vậy a_j không phải tổ hợp tuyến tính của các $a_i, i \in I$. Như vậy đi từ x đến $x + \lambda^* d$, số các ràng buộc chặt độc lập tuyến tính đã thêm ít nhất là một. Cứ tiếp tục quá trình, lại đi từ điểm $x + \lambda^* d, \dots$ cuối cùng ta phải đi đến điểm thỏa chặt n ràng buộc độc lập tuyến tính, chính là một đỉnh. $\square$

HỆ QUẢ 1.3.4. *Tập lồi đa diện ở dạng chính tắc mà khác trống thì luôn có đỉnh.*

CHỨNG MINH. $P = \{x: Ax = b, x \geq 0\}$ nằm trong góc tọa độ dương $\{x: x \geq 0\}$ nên không thể chứa đường thẳng. $\square$

NHẬN XÉT 1.3.1. Quy hoạch tuyến tính có thể có miền ràng buộc không có đỉnh. Để tiến hành thuật toán đơn hình ta đưa nó về dạng chính tắc bằng cách thêm biến bù. Khi đó theo Hệ quả 1.3.4, trong không gian các biến mới (nhiều chiều hơn) tập ràng buộc luôn có đỉnh. Đây là một lý do để ta phải đưa về dạng chính tắc khi tiến hành thuật toán đơn hình. Ở Mục 1.2.3, khi đưa ra bài toán bổ trợ để tìm từ vụng xuất phát, ta đã thừa nhận là luôn có từ vụng này. Hệ quả trên mới chứng minh điều thừa nhận này.

ĐỊNH LÝ 1.3.5 (tồn tại đỉnh tối ưu). *Xét bài toán min $c^T x$ trên tập lồi đa diện $P = \{x: Ax \geq b\}$ khác trống. Giả sử P có đỉnh. Khi đó hoặc hàm mục tiêu không giới nội dưới trên P hoặc P có đỉnh tối ưu.*

CHỨNG MINH. Ta gần như lặp lại chứng minh Định lý 1.3.3, chỉ có điểm khác là khi đi dần về một đỉnh của P , ta làm sao để có thêm tính chất là hàm mục tiêu không tăng.

Ta sẽ nói $x \in P$ là có hạng bằng k nếu ta chỉ tìm được đến k ràng buộc độc lập tuyến tính thỏa chặt tại x . Ta chỉ phải xét trường hợp hàm mục tiêu giới nội dưới trên P . Lấy $x \in P$ bất kỳ và giả sử hạng của x là $k < n$.

Trước hết ta chứng tỏ sẽ có $y \in P$ có hạng lớn hơn x và $c^T y \leq c^T x$. Đặt $I = \{i: a_i^T x = b_i\}$. Vì $k < n$, các vectơ a_i nằm trong một không gian con thực sự của $\mathbb{R}^n$ và sẽ có $d \neq 0$ để $a_i^T d = 0, i \in I$. Không mất tổng quát, giả sử $c^T d \leq 0$ (nếu $c^T d > 0$ thì lấy $-d$ thay cho d).

Nếu $c^T d < 0$, ta xét nửa đường thẳng $\{y = x + \lambda d: \lambda > 0\}$. Ta có $a_i^T(x + \lambda d) = a_i^T x = b_i, i \in I, \lambda > 0$, nên các ràng buộc thỏa chặt ở x cũng thỏa chặt trên mọi điểm của nửa đường thẳng. Vì $c^T(x + \lambda d) \rightarrow -\infty$ khi $\lambda \rightarrow +\infty$ nên

đến giá trị λ^* phải có ràng buộc $j \in I$ nào đó là ràng buộc chặt $a_j^T(x + \lambda^*d) = b_j$ để sau đó $x + \lambda d$ ra ngoài P . Cũng như ở Định lý 1.3.3 ta thấy hạng của $y := x + \lambda^*d$ ít nhất là $k + 1$ vì a_j độc lập tuyến tính với các $a_i, i \in I$. Hơn nữa ta cũng có $c^Ty < c^Tx$ (vì $c^Td < 0$).

Nếu $c^Td = 0$, ta xét cả đường thẳng $\{y = x + \lambda d; \lambda \in \mathbf{R}\}$ và tương tự trên ta sẽ thấy điểm $x + \lambda^*d$, khi đường thẳng đến ranh giới để rời P , có hạng lớn hơn hạng của x . Hơn nữa $c^T(x + \lambda^*d) = c^Tx$.

Cứ tiếp tục quá trình cuối cùng ta sẽ đến một điểm w có hạng n , tức w là một đỉnh của P , mà $c^Tw \leq c^Tx$.

Giả sử $w^1, \dots, w^r$ là các đỉnh của P (vì chỉ có hữu hạn đỉnh) và w^* là đỉnh có $c^Tw^* \leq c^Tw^i, i = 1, \dots, r$. Vì ta đã chỉ ra là với mọi $x \in P$, sẽ có một w^i để $c^Tw^i \leq c^Tx$, nên $c^Tw^* \leq c^Tx, \forall x \in P$. Vậy w^* là đỉnh tối ưu. $\square$

HỆ QUẢ 1.3.6. Bài toán quy hoạch tuyến tính min c^Tx trên tập lồi đa diện khác trống hoặc không giới nội dưới hoặc là có nghiệm tối ưu.

CHỨNG MINH. Ta đưa bài toán về dạng chính tắc tương đương với nó. Khi đó miền ràng buộc luôn có đỉnh và áp dụng được Định lý 1.3.5. Tuy nhiên đỉnh tối ưu theo Định lý 1.3.5 lại chỉ cho ta nghiệm tối ưu của tập lồi đa diện ban đầu, chưa chắc là đỉnh. $\square$

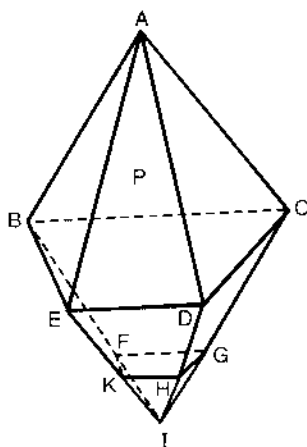
Sự tồn tại nghiệm tối ưu như trên là đặc thù của quy hoạch tuyến tính. Bài toán phi tuyến thậm chí đơn giản, như min $(\frac{1}{x} : x \geq 1)$ cũng không có nghiệm tối ưu mặc dù hàm mục tiêu ở đây lồi, liên tục, giới nội dưới và miền ràng buộc là tập lồi đa diện một chiều.

1.3.3 Hiện tượng suy biến

Theo Định nghĩa 1.3.5, nghiệm cơ sở đã thoả chặt đủ n ràng buộc độc lập tuyến tính. Nếu nó thoả chặt thêm ràng buộc khác nữa, (tất

nhiên không thể độc lập tuyến tính nữa vì không gian là n chiều) tức là quá mức cần thiết mà ta gọi là suy biến. Ý nghĩa hình học rất rõ ràng. Ta hãy lí giải trong $\mathbf{R}^3$ để dễ hình dung.

Trong không gian $\mathbf{R}^3$, nghiệm cơ sở suy biến thỏa hơn 3 ràng buộc chặt (tuyến tính), có nghĩa là nó là giao của hơn 3 mặt phẳng. Chẳng hạn ở tập lồi đa diện trong hình H.1.10, đỉnh A, B, C, D và E là nghiệm cơ sở suy biến chấp nhận được (cũng có thể gọi là *đỉnh suy biến*). F, G, H và K là nghiệm cơ sở không suy biến, chấp nhận được. I là nghiệm cơ sở suy biến (không chấp nhận được).



H. 1.10

THÍ DỤ. (i) Xét tập lồi đa diện ở dạng chính tắc

$$P = \{(x_1, x_2, x_3): x_1 - x_2 = 0; x_1 + x_2 + 2x_3 = 2; x_1, x_2, x_3 \geq 0\}.$$

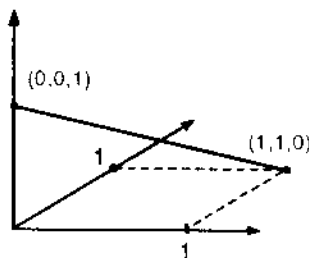
Vì $n = 3, m = 2, n - m = 1$, nên $(1, 1, 0)$ là đỉnh không suy biến còn $(0, 0, 1)$ là đỉnh suy biến. P là đoạn thẳng nối hai đỉnh này (H.1.11).

Nếu ta biểu diễn lại P ở dạng khác (không phải chính tắc):

$$P = \{(x_1, x_2, x_3): x_1 - x_2 = 0; x_1 + x_2 + 2x_3 = 2; x_1 \geq 0, x_3 \geq 0\},$$

thì $(0, 0, 1)$ là đỉnh không suy biến vì nó chỉ thỏa chặt đúng ba ràng buộc (độc lập tuyến tính).

(ii) Thậm chí, việc chuyển dạng chính tắc tổng quát sang dạng chuẩn tổng quát bao giờ cũng làm thay đổi tính không suy biến. Giả sử x^* là đỉnh không suy biến của $P = \{x: Ax = b, x \geq 0\}$, ở đây A là ma trận cấp $m \times n$ có hạng là m . Khi đó có đúng $n - m$ thành phần $x_i^* = 0$. Bây giờ xét P ở dạng chuẩn, $P = \{x: Ax \leq b, -Ax \leq -b, x \geq 0\}$. Tất nhiên x^* vẫn là nghiệm cơ sở chấp nhận được, tức là đỉnh. Nhưng còn tính suy biến thì sao?



H. 1.11

Tại x^* có $n - m$ ràng buộc thỏa chặt và $2m$ ràng buộc bất đẳng thức thỏa chặt. Vậy có $m + n$ ràng buộc thỏa chặt nên x^* là đỉnh suy biến của P ở dạng biểu diễn chuẩn.

1.3.4 Giải thích hình học phương pháp đơn hình

Ta đã biết một số sự kiện hình học sau đây

- Khi thực hiện thuật toán đơn hình ra biểu diễn tập ràng buộc ở dạng chính tắc. Về hình học khi đó không gian các biến là thích hợp để tập ràng buộc luôn có đỉnh.

Ở mỗi bước lặp của thuật toán đơn hình ta chuyển từ một nghiệm cơ sở chấp nhận được sang một nghiệm cơ sở chấp nhận được khác có giá trị mục tiêu tốt hơn. Về hình học có nghĩa là chuyển từ một đỉnh của tập ràng buộc sang đỉnh khác có giá trị mục tiêu tốt hơn.

Nhưng vị trí hai đỉnh đó so với nhau như thế nào? Bây giờ ta xem xét.

Tập các điểm cùng thỏa chặt một ràng buộc tuyến tính, tức là $\{x: a_1^T x = b_1\}$ là một siêu phẳng trong $\mathbf{R}^n$, tức là một diện $n - 1$ chiều. Tập các điểm cùng thỏa chặt hai ràng buộc độc lập tuyến tính là giao của hai siêu phẳng, tức là một diện $n - 2$ chiều. Cứ tiếp tục như vậy ta thấy tập các điểm cùng thỏa chặt $n - 1$ ràng buộc độc lập tuyến tính sẽ là một cạnh (diện một chiều) của tập lồi đa diện. Và cuối cùng tập các điểm thỏa chặt n ràng buộc độc lập tuyến tính là một diện 0 chiều, tức là một điểm mà ta đã gọi là đỉnh hoặc điểm cực biên.

Như vậy hai đỉnh của tập lồi đa diện trong $\mathbf{R}^n$ sẽ gọi là *đỉnh kề nhau* (adjacent) nếu có $n - 1$ ràng buộc độc lập tuyến tính thỏa chặt tại cả hai đỉnh.

Bây giờ xét hai đỉnh ở hai bước lặp liên nhau của thuật toán đơn hình. Khi chuyển sang nhau, chỉ có một biến x_k vào cơ sở, tức là một ràng buộc chặt $x_k = 0$ trở thành thỏa không chặt $x_k > 0$. Đồng thời,

để bù lại, cũng có đúng một biến cơ sở x_i trở thành không cơ sở, tức là có một ràng buộc thỏa không chặt $x_i > 0$ lại thành thỏa chặt $x_i = 0$. Như vậy chỉ đổi chỉ số một ràng buộc chặt (k thành i), tức là còn $n - 1$ ràng buộc độc lập tuyến tính thỏa chặt ở cả hai đỉnh. Vậy hai đỉnh này là kề nhau. Ta có kết luận hình học sau đây.

- Ở mỗi bước lặp của thuật toán đơn hình ta đã đi từ một đỉnh, theo một cạnh (edge) để đến một đỉnh kề có giá trị mục tiêu tốt hơn.

Tiếp theo, ta xét trường hợp suy biến. Chẳng hạn, xét đỉnh suy biến A ở H.1.10 trang 41. A suy biến vì nó thỏa chặt bốn ràng buộc (trong không gian ba chiều). Về hình học, điều này nghĩa là A là giao điểm của bốn mặt phẳng biên của miền chấp nhận được P . Như vậy là quá “mức bình thường”, mà đã suy biến, vì chỉ cần ba mặt phẳng không có cặp nào song song nhau là xác định một giao điểm duy nhất. Tuy vậy, sự suy biến này rất ngẫu nhiên, “tình cờ”. Chỉ một nhiễu nhỏ, vào một mặt phẳng trong bốn mặt chẳng hạn, đã có thể làm bốn mặt không có một điểm chung nữa, mà chỉ từng ba mặt có điểm chung. Khi đó một đỉnh suy biến A trở thành vài đỉnh không suy biến. Chính vì vậy khi nhiều cả m ràng buộc với $\varepsilon_i, i = 1, \dots, m$, khác hẳn cỡ nhau, ta đã loại bỏ được hoàn toàn các đỉnh suy biến. Đây là bản chất của phương pháp nhiễu loạn để tránh xoay vòng.

1.4 PHƯƠNG PHÁP ĐƠN HÌNH CẢI BIẾN

Trong các mục trước ta đã nghiên cứu phương pháp đơn hình dựa trên biểu diễn từ vựng. Từ bước này sang bước khác, ta tính từ vựng mới từ từ vựng cũ. Nếu xem xét kỹ ta sẽ thấy thực ra chỉ một phần nhỏ của từ vựng là cần dùng và tính lại (tức là cập nhật) khi sang bước lặp sau. Đồng thời, các dữ kiện ban đầu của bài toán (đã mất công đưa vào máy) lại chỉ được dùng ở bước lặp đầu. Còn ở mỗi bước lặp sau chỉ dùng đến từ vựng ngay trước đó. Vì vậy người ta đã xây dựng *phương pháp đơn hình cải biên* (revised simplex method) khác

phục hai yếu điểm trên, tức là tận dụng dữ liệu ban đầu và qua các bước chỉ cập nhật những gì cần thiết, không cần tính toán bộ từ vựng. Phương pháp đơn hình trình bày ở các mục trước sẽ có tên đầy đủ là *phương pháp đơn hình tiêu chuẩn* (standard simplex method).

Tuy vậy, mỗi bước lặp của phương pháp đơn hình cải biên chưa chắc đã tốn ít thời gian hơn phương pháp đơn hình tiêu chuẩn. Nhưng với các bài toán cỡ lớn điển hình hay gặp và các *bài toán thưa* (scarce problem), tức là bài toán có phần lớn các hệ số bằng 0, thì phương pháp đơn hình cải biên cho kết quả nhanh hơn hẳn. Vì vậy các phần mềm hiện đại đều dùng một dạng nào đó của phương pháp đơn hình cải biên.

1.4.1 Dạng ma trận của phương pháp đơn hình

Phương pháp đơn hình thực hiện trên dạng chính tắc của quy hoạch tuyến tính. Giả sử ta có quy hoạch dạng chuẩn, khi đó đưa vào biến bù ta viết lại quy hoạch ở dạng chính tắc là

$$\begin{aligned} \min z &= c^T x, \\ [A \ I]x &= b, \\ x &\geq 0, \end{aligned} \quad (1.20)$$

ở đây $x \in \mathbf{R}^n$, $c \in \mathbf{R}^n$, A là ma trận cấp $m \times (n - m)$, $b \in \mathbf{R}^m$, I là ma trận đơn vị cấp $m \times m$ xuất hiện khi đưa biến bù vào. Như vậy vectơ x chứa $n - m$ thành phần là biến ban đầu, m biến bù và cũng vậy c chứa $n - m$ hệ số ban đầu, ứng với vectơ c_I , và m hệ số là 0 ứng với biến bù, kí hiệu là c_{II} , tức là $c = [c_I \ c_{II}]$. Lúc này (1.20) có thể viết lại ở dạng sau, tương tự như dạng bảng ở Mục 1.2.6,

cơ sở	x_I	x_{II}	nghiêm
z	c_I^T	c_{II}^T	0
x_{II}	A	I	b

Bảng 1.1

Thể hiện phương trình của bảng này là

$$\begin{bmatrix} 1 & -c_I & -c_{II} \\ 0 & A & I \end{bmatrix} \begin{bmatrix} z \\ x_I \\ x_{II} \end{bmatrix} = \begin{bmatrix} 0 \\ b \end{bmatrix}. \quad (1.21)$$

So với từ vựng, ma trận A ở đây được viết nguyên, không chuyển vế và đổi dấu; đồng thời ta viết thêm cả cột ứng với x_{II} đang là biến cơ sở.

Khi thực hiện phương pháp đơn hình, nội dung tính toán chính ở mỗi bước lặp là phép xoay, tức là giải hệ phương trình đại số tuyến tính bằng phương pháp Gauss - Jordan. Trong môn học đại số tuyến tính ta đã biết việc giải này là tương đương với tìm nghịch đảo của ma trận. Vì vậy để “cải biên” cho thuật toán chạy nhanh hơn thì nghiên cứu cách nghịch đảo của ma trận là quan trọng. Do đó ta đi vào một mục bổ trợ.

1.4.1.a Nghịch đảo ma trận theo khối

Giả sử có ma trận cấp n $B = A^{-1}$ và có biểu diễn khối như sau

$$A = \left| \begin{array}{c|c} A_{11} & A_{12} \\ \hline (p \times p) & (p \times q) \\ \hline A_{21} & A_{22} \\ \hline (q \times p) & (q \times q) \end{array} \right| \quad B = \left| \begin{array}{c|c} B_{11} & B_{12} \\ \hline (p \times p) & (p \times q) \\ \hline B_{21} & B_{22} \\ \hline (q \times p) & (q \times q) \end{array} \right|$$

(ta có ghi cả cấp của các khối), ở đây giả sử ta biết chắc là A_{11} khả nghịch. Ta có

$$AB = BA = I_n = \left| \begin{array}{c|c} I_p & 0 \\ \hline 0 & I_q \end{array} \right|.$$

Do đó, tách riêng khối khi thực hiện phép nhân ma trận ta sẽ biểu diễn được các khối của ma trận nghịch đảo $B = A^{-1}$ qua các khối của A như sau.

Vì $AB = I_n$ nên

$$A_{11}B_{11} + A_{12}B_{21} = I_p,$$

$$A_{11}B_{12} + A_{12}B_{22} = 0.$$

Vì $BA = I_n$ nên

$$B_{21}A_{11} + B_{22}A_{21} = 0,$$

$$B_{21}A_{12} + B_{22}A_{22} = I_q.$$

Vì A_{11} khả nghịch, ta giải được B_{11} , B_{12} và B_{21} từ ba phương trình đầu; thay B_{21} vào phương trình cuối ta giải nốt ra B_{22} rồi đơn giản các biểu thức ta được

$$B_{11} = A_{11}^{-1} + (A_{11}^{-1} A_{12})D^{-1}(A_{21} A_{11}^{-1}),$$

$$B_{12} = - (A_{11}^{-1} A_{12})D^{-1},$$

$$B_{21} = - D^{-1}(A_{21} A_{11}^{-1}),$$

$$B_{22} = D^{-1},$$

ở đây $D = A_{22} - A_{21} (A_{11}^{-1} A_{12})$.

1.4.1.b Dạng ma trận của bảng đơn hình

Bây giờ xét bảng đơn hình ở một bước lặp bất kì. Giả sử B là tập chỉ số tương ứng với các biến cơ sở và N là tập chỉ số các biến không cơ sở ở bước này. Khi đó ta ký hiệu x_B là vectơ biến cơ sở (m thành phần) và x_N là vectơ biến không cơ sở ($n - m$ thành phần). Để tiện lợi ta cũng ký hiệu B là cơ sở ở bước này, vậy B là ma trận cấp m gồm m cột của ma trận của hệ ở bước này ứng với biến cơ sở. Tương tự, N là các cột của ma trận của hệ ứng với biến không cơ sở. (Ta không sợ nhầm lẫn vì tập chỉ số và ma trận khác hẳn nhau, từ ngữ cảnh sẽ rõ là nói đến cơ sở B hay tập chỉ số B .) Khi đó

$$z = c^T x, \quad \Leftrightarrow \quad z = c_B^T x_B + c_N^T x_N,$$

$$Ax = b. \quad Bx_B + Nx_N = b.$$

$$\Leftrightarrow \begin{cases} z = c_B^T x_B \\ Bx_B = b \end{cases} \Leftrightarrow \begin{bmatrix} 1 & -c_B^T \\ 0 & B \end{bmatrix} \begin{bmatrix} z \\ x_B \end{bmatrix} = \begin{bmatrix} 0 \\ b \end{bmatrix} \quad (1.22)$$

Để tính $[z \ x_B]$ từ (1.22) ta phải tính nghịch đảo ma trận khối $\begin{bmatrix} 1 & -c_B^T \\ 0 & B \end{bmatrix}$. Khi tính xong ta sẽ viết được bảng ma trận cho bước này tương tự Bảng 1.1 cho bước xuất phát. Theo mục trên ta có các khối tương ứng của ma trận nghịch đảo ở đây là

$$D = B - 0(1^{-1} \cdot (-c_B^T)) = B,$$

$$B_{11} = 1 + 1 \cdot (-c_B^T) B^{-1} \cdot 0 = 1,$$

$$B_{12} = - (1 \cdot (-c_B^T)) \cdot B^{-1} = c_B^T B^{-1},$$

$$B_{21} = - B^{-1} \cdot 0 = 0,$$

$$B_{22} = B^{-1}.$$

Do đó

$$\begin{bmatrix} 1 & -c_B^T \\ 0 & B \end{bmatrix}^{-1} = \begin{bmatrix} 1 & c_B^T B^{-1} \\ 0 & B^{-1} \end{bmatrix}.$$

Vậy giải (1.22) ta được

$$\begin{bmatrix} z \\ x_B \end{bmatrix} = \begin{bmatrix} 1 & c_B^T B^{-1} \\ 0 & B^{-1} \end{bmatrix} \begin{bmatrix} 0 \\ b \end{bmatrix} = \begin{bmatrix} c_B^T B^{-1} b \\ B^{-1} b \end{bmatrix} \quad (1.23)$$

Muốn được $[z \ x_B]$ biểu diễn qua các dữ kiện ban đầu của bài toán ta thay $[0 \ b]$ theo (1.21) vào (1.23) :

$$\begin{bmatrix} z \\ x_B \end{bmatrix} = \begin{bmatrix} 1 & c_B^T B^{-1} \\ 0 & B^{-1} \end{bmatrix} \begin{bmatrix} 1 & -c_I^T & -c_{II}^T \\ 0 & A & I \end{bmatrix} \begin{bmatrix} z \\ x_I \\ x_{II} \end{bmatrix} = \begin{bmatrix} c_B^T B^{-1} b \\ B^{-1} b \end{bmatrix}.$$

Biến đổi về giữa bằng cách thực hiện hai phép nhân ma trận ta được

$$\begin{aligned} \begin{bmatrix} z \\ x_B \end{bmatrix} &= \begin{bmatrix} 1 & c_B^T B^{-1} \\ 0 & B^{-1} \end{bmatrix} \begin{bmatrix} z - c_I^T x_I - c_{II}^T x_{II} \\ Ax_I + x_{II} \end{bmatrix} \\ &= \begin{bmatrix} z - c_I^T x_I - c_{II}^T x_{II} + c_B^T B^{-1} Ax_I + c_B^T B^{-1} x_{II} \\ B^{-1} Ax_I + B^{-1} x_{II} \end{bmatrix} \\ &= \begin{bmatrix} z + (c_B^T B^{-1} A - c_I^T) x_I + (c_B^T B^{-1} - c_{II}^T) x_{II} \\ B^{-1} Ax_I + B^{-1} x_{II} \end{bmatrix} \end{aligned}$$

Viết kết quả này vào bảng đơn hình (tương tự như từ (1.21) viết thành Bảng 1.1) ta được

cơ sở	x_I	x_{II}	nhịệm
z	$c_I^T - c_B^T B^{-1} A$	$c_{II}^T - c_B^T B^{-1}$	$c_B^T B^{-1} b$
x_B	$B^{-1} A$	B^{-1}	$B^{-1} b$

Bảng 1.2

NHẬN XÉT: Bảng 1.2 là bảng đơn hình ở bước lặp bất kì ở dạng ma trận. Nó được biểu diễn chỉ qua cơ sở hiện hành B (và B^{-1}) và dữ liệu ban đầu của bài toán A, b, c_I, c_{II}, c_B . Bảng 1.1 (xuất phát) cũng là trường hợp riêng của Bảng 1.2.

1.4.1.c Dạng tích của ma trận ngược

Theo Bảng 1.2 ở mỗi bước, bảng đơn hình chỉ chứa B^{-1} là nghịch đảo của cơ sở hiện hành B là cần tính, còn A, b, c_I, c_{II}, c_B là dữ liệu có ngay ở bài toán ban đầu. Nói một cách khác để chuyển từ một bước đã có, tức là có B^{-1} rồi, sang bước mới ta chỉ cần tính B_{next}^{-1} là nghịch đảo của cơ sở mới B_{next} . Phương pháp “dạng tích của ma trận ngược” sau đây, đề xuất bởi [Dantzig - Orchard Hays 1954], là rất thuận lợi để tính B_{next}^{-1} từ B^{-1} .

Gọi r là chỉ số thứ tự của cột ra khỏi cơ sở B (chú ý là thứ tự tính trong B chứ không phải tính chung trong hệ) và P là vectơ cột vào cơ sở mới. Đặt

$$\alpha := B^{-1}P := (\alpha_1, \dots, \alpha_m)^T,$$

$$I_m := (e_1, e_2, \dots, e_m),$$

với e_i là vectơ cột có thành phần i là 1 và mọi thành phần khác là 0. Ta sẽ chứng minh

$$B_{\text{next}}^{-1} = EB^{-1}, \quad (1.24)$$

ở đây

$$E = (e_1, \dots, e_{r-1}, \xi, e_{r+1}, \dots, e_m),$$

$$\xi = \left(-\frac{\alpha_1}{\alpha_r}, \dots, -\frac{\alpha_{r-1}}{\alpha_r}, \frac{1}{\alpha_r}, -\frac{\alpha_{r+1}}{\alpha_r}, \dots, -\frac{\alpha_m}{\alpha_r} \right)^T.$$

(Người ta chứng minh được là nếu $\alpha_r = 0$ thì B_{next}^{-1} không tồn tại.)

Trước hết hãy nhận xét rằng B_{next} là B được thay đổi một cột thứ r bởi cột P . Do đó dễ kiểm tra trực tiếp được là

$$B_{\text{next}} = BF,$$

với $F = (e_1, \dots, e_{r-1}, \alpha, e_{r+1}, \dots, e_m)$. Đồng thời, từ định nghĩa F thì rõ ràng $EF = I_m$, tức $F = E^{-1}$. Vậy ta có (1.24) vì

$$B_{\text{next}}^{-1} = (BF)^{-1} = F^{-1}B^{-1} = EB^{-1}.$$

THÍ DỤ. Giả sử

$$B = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 4 & 0 & 1 \end{bmatrix}, \quad B^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & \frac{1}{2} & 0 \\ -2 & 1 & 1 \end{bmatrix}.$$

Giả sử đã xác định được cột thứ $r = 3$ của B là ra, vào là $P = (2 \ 1 \ 5)^T$. Khi đó ta tính B_{next}^{-1} như sau.

$$\alpha = B^{-1}P = \begin{bmatrix} \frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & \frac{1}{2} & 0 \\ -2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 5 \end{bmatrix} = \begin{bmatrix} \frac{3}{4} \\ \frac{1}{2} \\ 2 \end{bmatrix},$$

$$\xi = \begin{bmatrix} -\frac{3}{4} \\ \frac{2}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} -\frac{3}{8} \\ -\frac{1}{4} \\ \frac{1}{2} \end{bmatrix}, E = \begin{bmatrix} 1 & 0 & -\frac{3}{8} \\ 0 & 1 & -\frac{1}{4} \\ 0 & 0 & \frac{1}{2} \end{bmatrix},$$

$$B_{\text{next}}^{-1} = EB^{-1} = \begin{bmatrix} 1 & 0 & -\frac{3}{8} \\ 0 & 1 & -\frac{1}{4} \\ 0 & 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & \frac{1}{2} & 0 \\ -2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{5}{4} & -\frac{5}{8} & -\frac{3}{8} \\ \frac{1}{2} & \frac{1}{4} & -\frac{1}{4} \\ -1 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

Gọi B_i là cơ sở ở bước thứ i và giả sử cơ sở xuất phát là $B_0 = I$ (tương ứng với m biến bù chẳng hạn). Gọi E_i là ma trận E ở bước i . Khi đó

$$B_1^{-1} = E_1 B_0^{-1} = E_1, B_2^{-1} = E_2 B_1^{-1} = E_2 E_1, \dots, \\ B_i^{-1} = E_i E_{i-1} \dots E_1. \quad (1.25)$$

Các ma trận E_i thường gọi là eta ma trận. Do đó (1.25) được gọi là *dạng thừa số theo eta* (eta factorization) của ma trận ngược của cơ sở. Về hình thức, (1.25) cho ta B_i^{-1} không phụ thuộc vào các cơ sở trước đó, nhưng xác định các $E_i, E_{i-1}, \dots$ cũng không dễ. Vì vậy khi tính toán người ta vẫn dùng (1.24).

1.4.2 Phương pháp đơn hình cải biên

Phương pháp đơn hình cải biên gốc (primal revised simplex method) có nội dung là ở mỗi bước ta tính B_{next}^{-1} dựa vào nghịch đảo B^{-1} của cơ sở ở bước trước và viết ra bảng ở dạng 1.2 nhưng cũng chỉ cần tính các phần tử cần thiết của bảng. Như vậy phương pháp này có ưu điểm như sau.

(a) Số thông tin cần tính ở mỗi bước và cần lưu trữ để tính cho bước sau ít hơn phương pháp đơn hình tiêu chuẩn.

(b) Phương pháp đơn hình tiêu chuẩn thường dẫn đến sai số làm tròn các dữ liệu một cách tích tụ, không điều chỉnh được. Còn ở phương pháp cải biên, tính B^{-1} ở các bước có thể điều khiển được độ chính xác, giảm hậu quả sai số do làm tròn số.

Để tính B_{next}^{-1} thì một trong các cách tốt là dùng dạng tính của ma trận ngược (1.24). Nhưng phương pháp đơn hình cải biên cũng có thể dùng cách tính ma trận ngược khác.

Các công đoạn ở mỗi bước lặp của phương pháp đơn hình cải biên cũng đúng như ở phương pháp tiêu chuẩn. Giả sử ta có nghịch đảo của cơ sở hiện hành là B^{-1} .

Đoạn 1 (xác định vectơ vào P). Ta cần xét các hệ số của phương trình z ở Bảng 1.2 ứng với các biến không cơ sở. Vậy ta cần tính $c_B^T B^{-1} P_j$ và, với mỗi cột P_j không cơ sở, $c_j - c_B^T B^{-1} P_j := c_j - z_j$. Cột vào là cột P_j có $c_j - z_j$ âm nhất (vẫn xét bài toán tìm minimum). Nếu $c_j - z_j \geq 0$, $\forall j$ thì phương án hiện hành $x_B = B^{-1}b$, $x_N = 0$, là tối ưu với giá trị mục tiêu là $z = c_B^T x_B$.

Đoạn 2 (xác định vectơ ra).

+ Tính các biến cơ sở $x_B = B^{-1}b$.

+ Tính hệ số của biến vào ở các ràng buộc, tức là vectơ $\alpha = B^{-1}P$ (là một cột ở khối $B^{-1}A$ trong Bảng 1.2). Số thứ tự r của vectơ cột ra (ở trong cơ sở B) ứng với chỉ số i đạt maximum trong biểu thức (so sánh Mục 1.2.2)

$$\max_{i \in B} \frac{\bar{a}_{ik}}{\bar{b}_i} = \max_{i \in B} \frac{\alpha_i}{(B^{-1}b)_i},$$

ở đây chỉ số i ở α_i và $(B^{-1}b)_i$ là chỉ thành phần thứ i , nếu giá trị maximum này dương.

Nếu maximum này không dương (tức là $\forall \alpha_i \leq 0$, vì $(B^{-1}b)_i = \bar{b}_i \geq 0, \forall i$, do đây là giá trị x_i của biến cơ sở trong nghiệm chấp nhận được), thì thuật toán dừng với kết luận là bài toán không giới nội (so sánh Mục 1.2.4).

Đoạn 3. Tính $B_{\text{next}}^{-1} = EB^{-1}$.

Nhận xét rằng đoạn 1 và đoạn 2 là tương ứng với việc viết ra bảng đơn hình ở bước hiện hành, nhưng ta viết ở dạng Bảng 1.2 và chỉ viết một số vị trí như sau

cơ sở	x_1	x_2	...	x_j	...	x_n	thử nghiệm
z	$c_1 - z_1$	$c_2 - z_2$		$c_j - z_j$		$c_n - z_n$	
x_B				$B^{-1}P_j$			$B^{-1}b$

THÍ DỤ 1.4.1. Giải bằng phương pháp đơn hình bài toán công ty Reddy Mikks (so sánh Mục 1.1.2). Sau khi đưa biến bù x_3, x_4, x_5, x_6 vào, bài toán có dạng chính tắc là

$$\begin{aligned} \min & -3x_1 - 2x_2, \\ & x_1 + 2x_2 + x_3 = 6, \\ & 2x_1 + x_2 + x_4 = 8, \end{aligned}$$

$$-x_1 + x_2 + x_5 = 1,$$

$$x_2 + x_6 = 2,$$

$$x_1, x_2, \dots, x_6 \geq 0.$$

Ta có dữ liệu xuất phát là

$$x_B = (x_3 \ x_4 \ x_5 \ x_6)^T,$$

$$c_B = (0 \ 0 \ 0 \ 0)^T,$$

$$B = (P_3, P_4, P_5, P_6) = I, \ B^{-1} = I.$$

BƯỚC 1. *Đoạn 1* (xác định vectơ vào). Trước hết tính $c_j - z_j$ cho các vectơ không cơ sở P_1 và P_2 :

$$c_B^T B^{-1} = (0 \ 0 \ 0 \ 0)I = (0 \ 0 \ 0 \ 0),$$

$$\begin{aligned} (c_1 - z_1 \ c_2 - z_2) &= (c_1 \ c_2) - c_B^T B^{-1} (P_1 \ P_2) = \\ &= [-3 \ -2] - [0 \ 0 \ 0 \ 0] \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ -1 & 1 \\ 0 & 1 \end{bmatrix} = [-3 \ -2]. \end{aligned}$$

Vì $-3 < -2$ nên P_1 là vectơ vào cơ sở.

Nhận xét rằng với vectơ P_j ở cơ sở thì $c_j - z_j = 0$ vì $c_j - c_B^T B^{-1} P_j = c_j - c_B^T I P_j = c_j - c_j = 0$.

Đoạn 2 (xác định vectơ ra). Tính các biến cơ sở :

$$x_B = B^{-1}b = Ib = \begin{bmatrix} 6 \\ 8 \\ 1 \\ 2 \end{bmatrix}.$$

Tính hệ số của biến vào ở các ràng buộc

số của biến vào ở các ràng buộc

$$\alpha = B^{-1}P_1 = IP_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \end{bmatrix}$$

Viết ra bảng đơn hình ta thấy rõ là chỉ cần tính có một hàng mục tiêu và hai cột như sau

cơ sở	x_1	x_2	x_3	x_4	x_5	x_6	nhịệm
z	-3	-2	0	0	0	0	
x_3	1						6
x_4	2						8
x_5	-1						1
x_6	0						2

$$\theta := \max_{i \in B} \frac{\alpha_i}{(B^{-1}b)_i} = \max \left\{ \frac{1}{6}, \frac{2}{8}, -\frac{1}{1}, \frac{0}{2} \right\} = \frac{1}{4} > 0,$$

tương ứng số thứ tự $r = 2$ của vector P_4 trong cơ sở $B = (P_3, P_4, P_5, P_6)$.
 Vậy $P_4 = (0, 1, 0, 0)$ là vector ra và biến ra là x_4 .

Đoạn 3. Tính $B_{\text{next}}^{-1} = EB^{-1}$. Chú ý là $\alpha_r = \alpha_2 = 2$ ta có

$$\xi = \begin{bmatrix} \alpha_1 \\ -\alpha_2 \\ \frac{1}{\alpha_2} \\ \alpha_3 \\ -\frac{\alpha_3}{\alpha_2} \\ \alpha_4 \\ -\frac{\alpha_4}{\alpha_2} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ -\left(-\frac{1}{2}\right) \\ 0 \\ \frac{0}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix}$$

$$B_{\text{next}}^{-1} = EB^{-1} = EI = (e_1 \dots e_{r-1} \quad \xi \quad e_{r+1} \dots e_m) = \begin{bmatrix} 1 & -\frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Từ đây ta lại kí hiệu cơ sở này là B để chuẩn bị sang bước mới. Vectơ cơ sở tương ứng là $x_B = (x_3 \ x_1 \ x_5 \ x_6)^T$ và $c_B = (0 \ -3 \ 0 \ 0)^T$.

BƯỚC 2. Đoạn 1. Tính $c_j - z_j$ cho P_2 và P_4 (không cơ sở):

$$c_B^T B^{-1} = [0 \ -3 \ 0 \ 0] \begin{bmatrix} 1 & -\frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = [0 \ -\frac{3}{2} \ 0 \ 0],$$

$$[c_2 - z_2 \ c_4 - z_4] = [-2 \ 0] - [0 \ -\frac{3}{2} \ 0 \ 0] \begin{bmatrix} 2 & 0 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} = \left[-\frac{1}{2} \ \frac{3}{2} \right].$$

Vậy P_2 là vectơ vào (vì $-\frac{1}{2}$ là hệ số âm duy nhất)

Đoạn 2.

$$x_B = B^{-1}b = \begin{bmatrix} 1 & -\frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 8 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 5 \\ 2 \end{bmatrix},$$

$$\alpha = B^{-1}P_2 = \begin{bmatrix} 1 & -\frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{3}{2} \\ \frac{1}{2} \\ \frac{3}{2} \\ 1 \end{bmatrix}$$

Viết ra bảng là

cơ sở	x_1	x_2	x_3	x_4	x_5	x_6	thiêm
z	0	$-\frac{1}{2}$	0	$\frac{3}{2}$	0	0	
x_3		$\frac{3}{2}$					2
x_1		$\frac{1}{2}$					4
x_5		$\frac{3}{2}$					5
x_6		1					2

$$\theta = \max_{i \in B} \frac{\alpha_i}{(B^{-1}b)_i} = \max \left\{ \frac{\frac{3}{2}}{2}, \frac{\frac{1}{2}}{4}, \frac{\frac{3}{2}}{5}, \frac{1}{2} \right\} = \frac{3}{4} > 0,$$

tương ứng số thứ tự $r = 1$ của vectơ P_3 trong cơ sở $B = (P_3, P_1, P_5, P_6)$ và biến ra là x_3 .

Đoạn 3. Tính B_{next}^{-1} . Vì $\alpha_1 = \frac{3}{2}$ ta có

$$\xi = \begin{bmatrix} \frac{1}{3/2} \\ \frac{-1/2}{3/2} \\ \frac{-3/2}{3/2} \\ \frac{-1}{3/2} \end{bmatrix} = \begin{bmatrix} 2/3 \\ -1/3 \\ -1 \\ -2/3 \end{bmatrix}$$

$$B_{\text{next}}^{-1} = \begin{bmatrix} \frac{2}{3} & 0 & 0 & 0 \\ -\frac{1}{3} & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -\frac{2}{3} & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & 0 & 0 \\ -\frac{1}{3} & \frac{2}{3} & 0 & 0 \\ -1 & 1 & 1 & 0 \\ -\frac{2}{3} & \frac{1}{3} & 0 & 1 \end{bmatrix}$$

Ta ký hiệu lại cơ sở mới là B. Vector cơ sở tương ứng là $x_B = (x_2 \ x_1 \ x_5 \ x_6)^T$ và $c_B = (-2 \ -3 \ 0 \ 0)^T$.

BƯỚC 3. Đoạn 1. Tính $c_j - z_j$ cho P_3 và P_4 (không cơ sở):

$$c_B^T B^{-1} = [-2 \ -3 \ 0 \ 0] \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & 0 & 0 \\ -\frac{1}{3} & \frac{2}{3} & 0 & 0 \\ -1 & 1 & 1 & 0 \\ -\frac{2}{3} & \frac{1}{3} & 0 & 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & -\frac{4}{3} & 0 & 0 \end{bmatrix},$$

$$[c_3 - z_3 \ c_4 - z_4] = [0 \ 0] - \begin{bmatrix} -\frac{1}{3} & -\frac{4}{3} & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \left[\frac{1}{3} \ \frac{4}{3} \right].$$

Vì $c_j - z_j > 0$ với mọi j là chỉ số không cơ sở nên cơ sở cuối cùng đã có là tối ưu. Nghiệm tối ưu là

$$\mathbf{x}_B = \begin{bmatrix} x_2 \\ x_1 \\ x_5 \\ x_6 \end{bmatrix} = \mathbf{B}^{-1}\mathbf{b} = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & 0 & 0 \\ -\frac{1}{3} & \frac{2}{3} & 0 & 0 \\ -1 & 1 & 1 & 0 \\ -\frac{2}{3} & \frac{1}{3} & 0 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 8 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{4}{3} \\ \frac{10}{3} \\ 3 \\ \frac{2}{3} \end{bmatrix},$$

$$\mathbf{x}_N = \begin{bmatrix} x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$z = \mathbf{c}_B^T \mathbf{x}_B = \begin{bmatrix} - & 2 & -3 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{4}{3} \\ \frac{10}{3} \\ 3 \\ 3 \\ \frac{2}{3} \end{bmatrix} = -\frac{38}{3}.$$

BÀI TẬP CHƯƠNG 1

1.1. Một máy cán thép có thể sản xuất hai sản phẩm thép tấm và thép cuộn với công suất cho mỗi loại là (nếu chỉ sản xuất một sản phẩm): thép tấm: 200 tấn/giờ, thép cuộn: 140 tấn/giờ. Lợi nhuận bán sản phẩm là: thép tấm 25 USD/tấn, thép cuộn: 30 USD/tấn. Theo tiếp thị, một tuần chỉ tiêu thụ được tối đa 6000 tấn thép tấm và 4000 tấn thép cuộn. Biết rằng máy làm việc 40 giờ một tuần. Vấn đề đặt ra là cần sản xuất mỗi loại sản phẩm bao nhiêu trong một tuần để có lợi nhuận cao nhất. Hãy diễn đạt bài toán thành quy hoạch tuyến tính. Bạn có thể giải bài toán bằng nhận xét trực tiếp không?

1.2. "Bài toán xe đạp" đặt ra như sau. Có n người cùng phải đi quãng đường 10 dặm mà chỉ có một xe đạp một chỗ ngồi. Tốc độ đi bộ của người j là w_j và đi xe đạp là b_j , $j = 1, \dots, n$. Làm sao để thời gian người cuối cùng đến đích là ngắn nhất.

(a) Giải bài toán với $n=3$, $w_1=4$, $w_2=w_3=2$, $b_1 = 16$, $b_2 = b_3 = 12$.

(b) Chứng minh rằng giá trị mục tiêu tối ưu của quy hoạch tuyến tính

min t

$$t - x_j - x'_j - y_j - y'_j \geq 0, j = 1, \dots, n,$$

$$t - \sum_{j=1}^n y_j - \sum_{j=1}^n y'_j \geq 0,$$

$$w_j x_j - w_j x'_j + b_j y_j - b_j y'_j = 10, j = 1, \dots, n,$$

$$\sum_{j=1}^n b_j y_j - \sum_{j=1}^n b_j y'_j \leq 10,$$

$$x_j, x'_j, y_j, y'_j \geq 0, j = 1, \dots, n.$$

là nhỏ hơn hoặc bằng giá trị mục tiêu tối ưu của bài toán xe đạp.

1.3. Một nhà máy chế biến thịt sản xuất ba loại thịt bò, lợn và cừu, với tổng lượng mỗi ngày 480 tấn thịt bò, 400 tấn thịt lợn và 230 tấn thịt cừu. Mỗi loại đều có thể bán được ở dạng tươi hoặc nấu chín. Tổng lượng các loại thịt có thể nấu chín (để bán) trong giờ làm việc là 420 tấn. Ngoài ra, có thể nấu thêm ngoài giờ 250 tấn (với giá cao hơn). Lợi nhuận thu được trên một tấn được cho bằng bảng sau, với đơn vị là triệu đồng,

	tươi	nấu chín	nấu chín ngoài giờ
bò	8	14	11
lợn	4	12	7
cừu	4	13	9

Thí dụ: Phương án sản xuất sau đây sẽ cho lợi nhuận 9965 triệu đồng

	tươi	nấu chín	nấu chín ngoài giờ
bò	165 tấn	280 tấn	35 tấn
lợn	295 tấn	70 tấn	35 tấn
cừu	55 tấn	70 tấn	105 tấn

Mục đích của nhà máy là tìm phương án sản xuất để làm cực đại lợi nhuận. Hãy phát biểu bài toán ở dạng quy hoạch tuyến tính chuẩn.

1.4. Một xưởng mộc làm bàn và ghế. Để một công nhân làm xong một bàn cần hai giờ

còn một ghế thì cần 30 phút. Khách hàng thường mua nhiều lắm là bốn ghế kèm theo một bàn. Do đó, tỉ lệ sản xuất ghế và bàn nhiều nhất là 4:1. Giá bán một bàn là 135 USD và một ghế là 50 USD. Hãy xác định xem số lượng sản xuất ghế và bàn hàng ngày là bao nhiêu để làm cực đại doanh thu. Biết rằng xưởng có bốn công nhân làm việc 8 giờ một ngày.

1.5. Một nhà máy sản xuất hai kiểu mũ. Thời gian lao động để làm ra một mũ kiểu thứ nhất nhiều gấp hai lần thời gian để làm xong mũ kiểu hai. Nếu sản xuất toàn mũ kiểu thứ hai thì nhà máy làm được 500 mũ một ngày. Thị trường tiêu thụ được trong mỗi ngày nhiều nhất là 150 mũ kiểu một và 200 mũ kiểu hai. Tiến lãi một mũ kiểu một là 8 USD và kiểu hai là 5 USD. Cần sản xuất bao nhiêu mũ mỗi kiểu trong ngày để tổng tiến lãi lớn nhất.

1.6. *Bài toán gà và trứng của Dantzig.* Trong hai tuần một con gà mái đẻ được 12 trứng để bán hoặc ấp được 4 trứng nở ra gà con. Sau 8 tuần thì bán tất cả gà và trứng với giá 0,6 USD một gà (lớn hoặc bé đều cùng giá) và 0,1 USD một trứng. Cần phải bố trí gà đẻ và ấp trứng như thế nào để doanh thu lớn nhất. Phát biểu bài toán cho hai trường hợp:

- a) ban đầu có 100 gà mái và 100 trứng;
- b) ban đầu có 100 gà mái và không có trứng.

1.7. Hãy đưa bài toán

$$\min 2x_1 + 13x_2 - 10I,$$

$$|x_1 + 2I| + |x_2| \leq 5,$$

về bài toán quy hoạch tuyến tính tương đương.

1.8. Giải bằng hình học

$$\begin{aligned} \text{(a)} \quad & \min -x_1 + x_2, \\ & -x_1 - 2x_2 \leq 6, \\ & x_1 - 2x_2 \leq 4, \\ & -x_1 + x_2 \leq 1, \\ & x_1, x_2 \geq 0; \end{aligned}$$

$$\begin{aligned} \text{(b)} \quad & \max 5x_1 + 6x_2, \\ & x_1 - 2x_2 \geq 2, \end{aligned}$$

$$-2x_1 + 3x_2 \geq 2.$$

(x_1, x_2) không bị ràng buộc dấu;

1.9. Chứng minh rằng bài toán tối ưu (với $x \in \mathbb{R}^n$)

$$\max f(x),$$

$$Ax = b,$$

$$x \geq 0,$$

trong đó $f(x) = \min \{c^1x, c^2x, \dots, c^px\}$ có thể đưa về bài toán quy hoạch tuyến tính

$$\max x_{n+1},$$

$$Ax = b,$$

$$x \geq 0,$$

$$x_{n+1} \leq c^1x, x_{n+1} \leq c^2x, \dots, x_{n+1} \leq c^px.$$

1.10. Tìm điều kiện cần và đủ đối với các số thực a và b để quy hoạch tuyến tính

$$\max x_1 + x_2,$$

$$ax_1 + bx_2 \leq 1,$$

$$x_1, x_2 \geq 0;$$

(i) có nghiệm tối ưu;

(ii) có nghiệm chấp nhận được;

(iii) không giới nội.

1.11. Tìm mối liên hệ giữa hai quy hoạch tuyến tính

$$\min c^T x, Ax = b, x \geq 0,$$

và

$$\min \alpha c^T x, Ax = \lambda b, x \geq 0,$$

ở đây $\alpha > 0$ và $\lambda > 0$ là hai số cho trước.

1.12. Giải bằng phương pháp đơn hình

$$(a) \quad \max x_1 + 3x_2 + 2x_3$$

$$x_1 + x_2 + x_3 = 1,$$

$$7x_1 + 2x_2 + 3x_3 \leq 20,$$

$$x_1 + 5x_2 + 4x_3 \leq 30,$$

$$x_1, x_2, x_3 \geq 0;$$

$$(b) \quad \min 2x_1 - x_2 + 3x_3 + 7x_4 - 5x_5,$$

$$x_1 + 2x_2 + x_3 + x_4 + 6x_5 \geq 10,$$

$$2x_1 + 3x_2 + 4x_3 + x_4 + 2x_5 \geq 4,$$

$$3x_1 + 2x_2 + x_3 + 3x_4 + x_5 \leq 8,$$

$$x_1, x_2 \geq 0.$$

$$(c) \quad \max 3x_1 + x_2,$$

$$-x_1 + x_2 \leq -1,$$

$$-3x_1 - x_2 \leq -1,$$

$$4x_1 + 2x_2 \leq 1,$$

$$2x_2 \leq 1,$$

$$x_1, x_2 \geq 0.$$

$$(d) \quad \max 6x_1 + 8x_2 + 5x_3 + 9x_4,$$

$$2x_1 + x_2 + x_3 + 3x_4 \leq 5,$$

$$x_1 + 3x_2 + x_3 + 2x_4 \leq 3,$$

$$x_1, x_2, x_3, x_4 \geq 0;$$

$$(e) \quad \min x_1 - 4x_2 + x_3 + 13x_4 + 23x_5,$$

$$2x_1 + x_2 + x_3 + 4x_4 - x_5 = 6,$$

$$x_1 + 2x_2 + x_3 + x_4 - 4x_5 = 3,$$

$$x_1, \dots, x_5 \geq 0.$$

$$(f) \quad \min -14x_1 - 18x_2 - 16x_3 - 10x_4,$$

$$4.5x_1 + 8.5x_2 + 6x_3 + 20x_4 + x_5 = 6000,$$

$$x_1 + x_2 + 4x_3 + 40x_4 + x_6 = 4000,$$

$$x_1, \dots, x_6 \geq 0.$$

1.13. Giải bằng phương pháp đơn hình

$$(a) \quad \min -2x_1 - x_2,$$

$$2x_1 + x_2 \leq 4,$$

$$2x_1 + 3x_2 \leq 3,$$

$$4x_1 + x_2 \leq 5,$$

$$x_1 + 5x_2 \leq 1,$$

$$x_1, x_2 \geq 0;$$

$$(b) \quad \max 3x_1 + 2x_2,$$

$$-x_1 + 2x_2 \geq -1,$$

$$x_1 - x_2 \leq 2,$$

$$-2x_1 - x_2 \geq -6,$$

$$x_1 \leq 5,$$

$$2x_1 + x_2 \leq 16,$$

$$x_1 + x_2 \leq 12,$$

$$x_1 + 2x_2 \leq 21,$$

$$x_2 \leq 10,$$

$$x_1, x_2 \geq 0;$$

$$(c) \quad \min x_1 + 2x_3 + x_5,$$

$$x_1 + x_2 + x_3 + x_4 + x_5 = 5,$$

$$x_2 + x_3 + x_4 - x_5 = 2,$$

$$x_3 - x_4 + x_5 = 1,$$

$$x_1, \dots, x_5 \geq 0;$$

$$(d) \quad \min 5x_1 - 10x_2,$$

$$-2x_1 + x_2 \leq 1,$$

$$x_1 - x_2 \geq -2,$$

$$3x_1 + x_2 \leq 8,$$

$$-2x_1 + 3x_2 \geq -9,$$

$$4x_1 + 3x_2 \geq 0,$$

$$x_1, x_2 \geq 0.$$

1.14. Giải bằng phương pháp đơn hình cải biến

- (a) $\min 2x_1 + x_2,$
 $3x_1 + x_2 = 3,$
 $4x_1 + 3x_2 \geq 6,$
 $x_1 + 2x_2 \leq 3,$
 $x_1, x_2 \geq 0;$
- (b) $\max 6x_1 - 2x_2 + 3x_3,$
 $2x_1 - x_2 + 2x_3 \leq 2,$
 $x_1 + 4x_3 \leq 4,$
 $x_1, x_2, x_3 \geq 0;$
- (c) $\min -2x_1 - x_2 - 2x_3,$
 $4x_1 + 3x_2 + 8x_3 \leq 12,$
 $4x_1 + x_2 + 12x_3 \leq 8,$
 $4x_1 - x_2 + 3x_3 \leq 8,$
 $x_1, x_2, x_3 \geq 0.$

1.15. Giải bằng thuật toán hai pha

- (a) $\max 2x_1 - 6x_2,$
 $-x_1 - x_2 - x_3 \leq -2,$
 $2x_1 - x_2 + x_3 \leq 1,$
 $x_1, x_2, x_3 \geq 0;$
- (b) $\min -x_1 - 3x_2,$
 $x_1 + x_2 \geq 3,$
 $-x_1 + x_2 \leq -1,$
 $x_1 + 2x_2 \leq 4,$
 $x_1, x_2 \geq 0;$
- (c) $\max x_1 + 3x_2,$
 $-x_1 - x_2 \leq -3,$

$$-x_1 + x_2 \leq -1,$$

$$-x_1 + 2x_2 \leq 2,$$

$$x_1, x_2 \geq 0.$$

1.16. Giải quy hoạch tuyến tính

$$\max \sum_{j=1}^n p_j x_j,$$

$$\sum_{j=1}^n q_j x_j \leq \beta,$$

$$0 \leq x_j \leq 1, j = 1, \dots, n.$$

Ở đây các số $p_j, j = 1, \dots, n$ dương và có tổng bằng 1. Các số q_j cũng có tính chất đó. Hơn nữa, giả sử rằng (gần như không mất tính tổng quát)

$$\frac{p_1}{q_1} < \frac{p_2}{q_2} < \dots < \frac{p_n}{q_n}.$$

Cuối cùng, biết rằng $\beta > 0$ và nhỏ. (Xem Bài tập 1.3 để thấy nguồn gốc của bài toán này.)

CHƯƠNG 2**QUY HOẠCH TUYẾN TÍNH ĐỐI NGẪU.**

Lý thuyết đối ngẫu là một vấn đề trung tâm của lý thuyết tối ưu. Nó giúp phân tích sâu cấu trúc của bài toán tối ưu, cho phép giải thích hình học, đồng thời cũng là công cụ mạnh mẽ để nghiên cứu và xây dựng các thuật toán. Nói riêng, trong quy hoạch tuyến tính, ở chương này ta cũng thấy rõ điều đó.

2.1 BÀI TOÁN MỞ ĐẦU. ĐỊNH NGHĨA**2.1.1 Bài toán mở đầu**

Xét một quy hoạch tuyến tính, chẳng hạn ở dạng chính tắc

$$\begin{aligned} \min c^T x, \\ Ax = b, \\ x \geq 0. \end{aligned} \tag{2.1}$$

Nếu ta biết một nghiệm chấp nhận được x^0 thì rõ ràng ta được một cận trên của mục tiêu tối ưu. Giả sử là có tồn tại nghiệm tối ưu x thì $c^T x \leq c^T x^0$. Tuy ta chưa tìm được x , nhưng nếu ta biết một cận dưới của mục tiêu tối ưu thì đã “khoanh vùng” được giá trị mục tiêu tối ưu. Ta hãy thử tìm một cận dưới.

Xử lý ràng buộc $Ax = b$ là khó nên ta hãy thử “nới lỏng” bằng cách coi như không có ràng buộc này nhưng phạt vào các phương án x không thỏa ràng buộc này (tức $Ax - b \neq 0$) một lượng $y^T(b - Ax)$, ở đây y là vec tơ (m chiều như b) giá phải trả cho sự vi phạm ràng buộc $b - Ax$. Thay cho (2.1) ta giải bài toán

$$\begin{aligned} \min \quad & c^T x + y^T(b - Ax), \\ & x \geq 0. \end{aligned} \quad (2.2)$$

y được gọi là *nhân tử Lagrange* (Lagrange multiplier), $L(x, y) := c^T x + y^T(b - Ax)$ được gọi là *hàm Lagrange* (Lagrange function) hoặc *Lagrangean* của bài toán (2.1).

Đặt $g(y)$ là giá trị tối ưu của bài toán nới lỏng (2.2) (phụ thuộc vec tơ giá y). Vì bài toán nới lỏng bị ràng buộc ít hơn nên ta dự đoán $g(y)$ là cận dưới cho giá trị mục tiêu tối ưu $c^T x$ của bài toán (2.1). Điều này được kiểm chứng dễ dàng:

$$g(y) = \min_{x \geq 0} (c^T x + y^T(b - Ax)) \leq c^T x + y^T(b - Ax) = c^T x,$$

vì x là chấp nhận được của (2.1) nên $b - Ax = 0$. Vậy ta đã có một cận dưới cho giá trị mục tiêu tối ưu của (2.1) là $g(y)$ với vec tơ giá y bất kì. Một cách tự nhiên ta quan tâm bài toán tìm cận dưới sát nhất:

$$\begin{aligned} \max \quad & g(y), \\ & y \in \mathbb{R}^m \end{aligned} \quad (2.3)$$

Bài toán này gọi là *bài toán đối ngẫu* (dual problem) của quy hoạch tuyến tính (2.1). Sau đây, ta sẽ thấy kết quả chính của lý thuyết đối ngẫu của quy hoạch tuyến tính khẳng định giá trị mục tiêu tối ưu của bài toán gốc (primal problem), tức là bài toán (2.1), bằng giá trị mục tiêu tối ưu của bài toán đối ngẫu (2.3) của nó. Nói một cách khác, nếu vec tơ giá y được chọn thích hợp là nghiệm tối ưu của bài toán đối ngẫu thì việc vi phạm ràng buộc $b = Ax$ không làm ảnh hưởng đến

giá trị mục tiêu tối ưu (vì giá trị mục tiêu tối ưu của bài toán nói lỏng (2.2) bằng $c^T x$).

Bây giờ ta hãy biến đổi (2.3) để được dạng tốt hơn. Ta có

$$g(y) = \min_{x \geq 0} (c^T x + y^T (b - Ax)) = y^T b + \min_{x \geq 0} (c^T - y^T A)x.$$

Mặt khác

$$\min_{x \geq 0} (c^T - y^T A)x = \begin{cases} 0, & \text{nếu } c^T - y^T A \geq 0^T, \\ -\infty & \text{nếu khác.} \end{cases}$$

Thay vào (2.3) ta được dạng tương đương của nó

$$\begin{aligned} \max y^T b, \\ y^T A \leq c^T. \end{aligned} \quad (2.4)$$

Vậy bài toán đối ngẫu của quy hoạch tuyến tính (2.1) là quy hoạch tuyến tính (2.4) được lập để dùng từ các dữ kiện của (2.1).

Ta lập bài toán đối ngẫu cho quy hoạch tuyến tính dạng chuẩn một cách tương tự, bằng cách thêm biến bù

$$\begin{aligned} Ax \leq b & \Leftrightarrow Ax + s = b, s \geq 0. \\ & \Leftrightarrow [A \ I] \begin{bmatrix} x \\ s \end{bmatrix} = b, s \geq 0. \end{aligned}$$

Vì với biến bù, mục tiêu trở thành $\min c^T x + 0^T s$, nên theo (2.4), áp dụng cho trường hợp có biến bù này ta được bài toán đối ngẫu của quy hoạch tuyến tính dạng chuẩn là

$$\begin{aligned} \max y^T b, \\ y^T [A \ I] \leq [c^T \ 0^T], \end{aligned}$$

tức là

$$\begin{aligned} \max y^T b, \\ y^T A \leq c^T, \\ y \leq 0. \end{aligned} \quad (2.5)$$

Bây giờ ta xét trường hợp bài toán gốc không ở dạng chính tắc hoặc dạng chuẩn. Khi đó để lập bài toán đối ngẫu, tức là bài toán tìm cận dưới sát nhất của giá trị mục tiêu tối ưu ta cũng lí luận tương tự như trên. Với các ràng buộc đẳng thức, bất đẳng thức và ràng buộc về dấu ta làm đúng như trên. Nếu vec tơ x ở bài toán gốc là tự do (về dấu) thì cũng lí luận như với (2.4) ta có

$$\min_{x \in \mathbb{R}^n} (c^T - y^T A)x = \begin{cases} 0, & \text{nếu } c^T - y^T A = 0^T, \\ -\infty & \text{nếu khác.} \end{cases}$$

Vậy ở bài toán đối ngẫu phải có ràng buộc

$$y^T A = c^T.$$

2.1.2 Định nghĩa bài toán đối ngẫu

Giả sử ma trận A có hàng là a_i^T và các cột là A_j . Bởi lí luận ở Mục 2.1.1, ta dẫn đến định nghĩa tổng quát của bài toán đối ngẫu như sau. Giả sử bài toán gốc có cấu trúc ở bên trái trong bảng dưới đây, Khi đó bài toán đối ngẫu được định nghĩa với cấu trúc tương ứng ở bên phải

$\min c^T x,$ với ràng buộc	$\max y^T b,$ với ràng buộc
$a_i^T x = b_i, i \in M_1,$	$y_i \text{ tự do, } i \in M_1,$
$a_i^T x \leq b_i, i \in M_2,$	$y_i \leq 0, i \in M_2,$
$a_i^T x \geq b_i, i \in M_3,$	$y_i \geq 0, i \in M_3,$

$$\begin{array}{ll} x_j \geq 0, j \in N_1, & y^T A_j \leq c_j, j \in N_1, \\ x_j \leq 0, j \in N_2, & y^T A_j \geq c_j, j \in N_2, \\ x_j \text{ tự do}, j \in N_3, & y^T A_j = c_j, j \in N_3. \end{array}$$

NHẬN XÉT. Mỗi ràng buộc ở bài toán gốc, không kể ràng buộc dấu, ứng với một biến của bài toán đối ngẫu. Mỗi biến của bài toán gốc ứng với một ràng buộc ở bài toán đối ngẫu. Đồng thời các chiều bất đẳng thức có quan hệ trực tiếp với nhau và cho bằng bảng sau đây.

GỐC	min	max	ĐỐI NGẪU
ràng buộc	$= b_i$	tự do	biến
	$\leq b_i$	≤ 0	
	$\geq b_i$	≥ 0	
biến	≥ 0	$\leq c_j$	ràng buộc
	≤ 0	$\geq c_j$	
	tự do	$= c_j$	

Bảng 2.1 Quan hệ giữa biến và ràng buộc của bài toán gốc và bài toán đối ngẫu

Các biến y_i của bài toán đối ngẫu được gọi là *biến đối ngẫu* (dual variable). Ở trên ta đã thấy là biến đối ngẫu cũng chính là nhân tử Lagrange.

Mỗi quy hoạch tuyến tính có nhiều dạng tương đương. Từ mỗi dạng đó, theo định nghĩa trên ta sẽ lập được một bài toán đối ngẫu. Tất nhiên các bài toán đối ngẫu nhận được cũng sẽ tương đương nhau.

Một câu hỏi được đặt ra: bài toán đối ngẫu của bài toán đối ngẫu có khác bài toán gốc không ?

THÍ DỤ 2.1.1. Xét quy hoạch tuyến tính ở bên trái dưới đây ta sẽ lập được bài toán đối ngẫu của nó như ở bên phải.

$$\begin{array}{ll} \min x_1 + x_2 + 3x_3, & \max 5y_1 + 6y_2 + 4y_3, \\ -x_1 + 3x_2 & = 5, \\ & y_1 \text{ tự do}, \end{array}$$

$$\begin{array}{ll}
2x_1 - x_2 + 3x_3 \geq 6, & y_2 \geq 0, \\
x_3 \leq 4, & y_3 \leq 0, \\
x_1 \geq 0, & -y_1 + 2y_2 \leq 1, \\
x_2 \leq 0, & 3y_1 - y_2 \geq 1, \\
x_3 \text{ tự do.} & 3y_2 + y_3 = 3.
\end{array}$$

Bây giờ ta tìm đối ngẫu của bài toán đối ngẫu trên đây. Ta viết lại nó ở dạng bài toán min và nhân cả ba ràng buộc cuối với -1 . Dạng tương đương nhận được ta viết vào bên trái dưới đây và ta lập đối ngẫu của nó ở bên phải.

$$\begin{array}{ll}
\min -5y_1 - 6y_2 - 4y_3, & \max -x_1 - x_2 - 3x_3, \\
y_1 \text{ tự do,} & x_1 - 3x_2 = -5, \\
y_2 \geq 0, & -2x_1 + x_2 - 3x_3 = -6, \\
y_3 \leq 0, & -x_3 \geq -4, \\
y_1 - 2y_2 \geq -1, & x_1 \geq 0, \\
-3y_1 + y_2 \leq -1, & x_2 \leq 0, \\
-3y_2 - y_3 = -3. & x_3 \text{ tự do.}
\end{array}$$

Ta thấy bài toán đối ngẫu ở bên phải trên đây của bài toán đối ngẫu trùng với bài toán gốc.

Bài toán gốc ta xét ở thí dụ trên đã có đủ các kiểu ràng buộc. Vì vậy làm đúng như ở thí dụ trên đối với bài toán tổng quát (chứ không phải bằng số cụ thể) ta chứng minh được định lý sau đây

ĐỊNH LÝ 2.1.1. Nếu ta biến đổi bài toán đối ngẫu về dạng tương đương là tìm cực tiểu (min) và lập đối ngẫu của nó ta nhận được bài toán gốc.

2.2 CÁC ĐỊNH LÝ VỀ ĐỐI NGẪU

2.2.1 Đối ngẫu yếu

Ở Mục 2.1 ta đã thấy đối với bài toán chính tắc thì giá trị mục tiêu $g(y)$ của bài toán đối ngẫu luôn là cận dưới của giá trị mục tiêu tối ưu của bài toán gốc. Định lý sau cho thấy tính chất này đúng với bất kì cặp bài toán tuyến tính đối ngẫu dạng bất kì.

ĐỊNH LÝ 2.2.1 (đối ngẫu yếu). *Nếu x là nghiệm chấp nhận được của quy hoạch tuyến tính gốc và y là nghiệm chấp nhận được của quy hoạch đối ngẫu thì*

$$y^T b \leq c^T x.$$

CHỨNG MINH. Ta đặt

$$u_i = y_i(a_i^T x - b_i),$$

$$v_j = (c_j - y^T A_j)x_j.$$

Theo định nghĩa bài toán đối ngẫu (Bảng 2.1) thì y_i và $a_i^T x - b_i$ cùng dấu, còn $c_j - y^T A_j$ và x_j cùng dấu. Do đó $u_i \geq 0 \forall i$ và $v_j \geq 0 \forall j$.

Nhận xét rằng

$$\sum_i u_i = y^T A x - y^T b,$$

$$\sum_j v_j = c^T x - y^T A x.$$

Do đó

$$0 \leq \sum_i u_i + \sum_j v_j = c^T x - y^T b. \square$$

HỆ QUẢ 2.2.2. (i) Nếu hàm mục tiêu của quy hoạch tuyến tính gốc không giới nội dưới, thì bài toán đối ngẫu không có nghiệm chấp nhận được.

(ii) Nếu bài toán đối ngẫu không giới nội trên (tức là hàm mục tiêu của nó không giới nội trên), thì bài toán gốc không chấp nhận được (tức là không có nghiệm chấp nhận được).

CHỨNG MINH. Do sự tương tự ta chỉ chứng minh (i). Giả sử bài toán gốc không giới nội dưới, tức là tồn tại dãy x^k các nghiệm chấp nhận được để $c^T x^k \rightarrow -\infty$ khi $k \rightarrow \infty$. Giả sử phản chứng là bài toán đối ngẫu có nghiệm chấp nhận được y^0 . Khi đó, do đối ngẫu yếu,

$$(y^0)^T b \leq c^T x^k \quad \forall k.$$

Cho $k \rightarrow \infty$ ta được điều vô lí $(y^0)^T b \leq -\infty$. $\square$

HỆ QUẢ 2.2.3. Giả sử $\bar{x}$ và $\bar{y}$ là nghiệm chấp nhận được tương ứng của hai bài toán đối ngẫu nhau. Nếu $\bar{y}^T b = c^T \bar{x}$ thì $\bar{x}$ và $\bar{y}$ phải là nghiệm tối ưu tương ứng của hai bài toán.

CHỨNG MINH. Xét nghiệm chấp nhận được x bất kì của bài toán gốc. Theo đối ngẫu yếu ta có $\bar{y}^T b \leq c^T x$, và do đó $c^T \bar{x} \leq c^T x$, $\forall x$ chấp nhận được. Vậy $\bar{x}$ là nghiệm tối ưu. Tương tự ta sẽ thấy $\bar{y}$ là nghiệm tối ưu của bài toán đối ngẫu. $\square$

2.2.2. Đối ngẫu mạnh

Bây giờ ta chứng minh kết quả chính nhất của lý thuyết đối ngẫu đã có nhắc đến ở Mục 2.1 là giá trị mục tiêu tối ưu của hai bài toán quy hoạch tuyến tính đối ngẫu nhau là bằng nhau.

ĐỊNH LÝ 2.2.4 (đối ngẫu mạnh). Nếu quy hoạch tuyến tính gốc có nghiệm tối ưu x thì quy hoạch đối ngẫu cũng có nghiệm tối ưu y và

giá trị mục tiêu tối ưu bằng nhau:

$$c^T x = (y)^T b.$$

CHỨNG MINH. Quy hoạch tuyến tính có nhiều dạng tương đương, nhưng ta đã biết là các bài toán đối ngẫu của các dạng tương đương đó cũng là các dạng tương đương. Mà các bài toán tương đương thì có cùng giá trị mục tiêu tối ưu. Vì vậy ta chỉ cần chứng minh định lý cho một dạng cụ thể, chẳng hạn là dạng chuẩn:

$$\begin{aligned} \min \quad & c^T x, \\ \text{Ax} \leq & b, \\ x \geq & 0. \end{aligned}$$

Ta sẽ áp dụng phương pháp đơn hình lên quy hoạch này. Vì có nghiệm tối ưu nên sau hữu hạn bước ta sẽ được từ vùng tối ưu với hàm mục tiêu là

$$z = z^* + \sum_{j=1}^n c_j^* x_j + \sum_{i=1}^m d_i^* w_i, \quad (2.6)$$

ở đây, $c_j^* = 0$ và $d_i^* = 0$ với các biến x_j và w_i là biến cơ sở ở từ vùng tối ưu này. Vì z^* là giá trị mục tiêu tối ưu nên (nó phải tương ứng với nghiệm tối ưu)

$$z^* = \sum_{j=1}^n c_j x_j^*. \quad (2.7)$$

Đặt

$$y_i^* = -d_i^*, \quad i = 1, \dots, m.$$

Ta sẽ chứng minh $y^* = (y_1^*, \dots, y_m^*)$ là nghiệm chấp nhận được của bài toán đối ngẫu với $(y^*)^T b = c^T x^*$. Vì hàm mục tiêu ban đầu và (2.6) là tương đương, ta có

$$\sum_{j=1}^n c_j x_j = z^* + \sum_{j=1}^n c_j^* x_j + \sum_{i=1}^m d_i^* w_i$$

$$\begin{aligned}
&= z^* + \sum_{j=1}^n c_j^* x_j + \sum_{i=1}^m (-y_i^*)(b_i - \sum_{j=1}^n a_{ij} x_j) \\
&= z^* - \sum_{i=1}^m b_i y_i^* + \sum_{j=1}^n (c_j^* + \sum_{i=1}^m y_i^* a_{ij}) x_j.
\end{aligned}$$

Hai vế là biểu diễn của cùng một hàm tuyến tính nên các hệ số tương ứng phải đồng nhất. Do đó

$$z^* = \sum_{i=1}^m b_i y_i^*, \quad (2.8)$$

$$c_j = c_j^* + \sum_{i=1}^m y_i^* a_{ij}. \quad (2.9)$$

Từ (2.7) và (2.8) ta có $(y^*)^T b = c^T x$. Ở từ vựng tối ưu các hệ số của hàm mục tiêu phải không âm, tức $c_j^* \geq 0, j = 1, \dots, n$, và $d_i^* \geq 0, i = 1, \dots, m$. Vì vậy, theo (2.9)

$$\begin{aligned}
&\sum_{i=1}^m y_i^* a_{ij} \leq c_j, \quad j = 1, \dots, n, \\
&y_i^* = -d_i^* \leq 0, \quad i = 1, \dots, m,
\end{aligned}$$

tức là y^* là nghiệm chấp nhận được của bài toán đối ngẫu. Do $(y^*)^T b = c^T x^*$, y phải là nghiệm tối ưu theo Hệ quả 2.2.3. $\square$

NHẬN XÉT 2.2.1 (a) Sau định lý đối ngẫu yếu ta mới chỉ biết $(y^*)^T b \leq (c)^T x^*$ (nếu hai bài toán có nghiệm tối ưu). Khi xét các cặp bài toán đối ngẫu nhau cho quy hoạch phi tuyến tổng quát người ta cũng đạt được sự so sánh giá trị mục tiêu tối ưu ở dạng đối ngẫu yếu như vậy. Nếu dấu bất đẳng thức chặt xảy ra ta nói có lỗ hổng đối ngẫu (duality gap). Đối với quy hoạch phi tuyến muốn lỗ hổng này bằng 0, bài toán gốc phải là bài toán lồi, tức là các

tập và hàm trong dữ liệu phải thỏa các giả thiết nhất định về lỗi. Định lý 2.2.4 về đối ngẫu mạnh nói rằng đối với quy hoạch tuyến tính chấp nhận được thì lỗ hổng đối ngẫu luôn bằng 0. Điều này dễ hiểu vì quy hoạch tuyến tính là trường hợp riêng của quy hoạch lồi.

(b) Hệ quả 2.2.2 có thể "ghép vào" Định lý 2.2.4 như sau. Quy hoạch gốc không giới nội được quy ước (khi nói đến đối ngẫu) như là mục tiêu tối ưu bằng $-\infty$. Khi đó hệ quả nói rằng quy hoạch đối ngẫu không chấp nhận được. Điều này có thể quy ước coi là giá trị mục tiêu của bài toán đối ngẫu (bài toán max) luôn bằng $-\infty$. Như vậy ta cũng coi như lỗ hổng đối ngẫu bằng 0, tức là ở dạng đối ngẫu mạnh. Tương tự, trường hợp bài toán đối ngẫu không giới nội có thể coi như giá trị mục tiêu tối ưu của bài toán bằng nhau ở $+\infty$.

(c) Theo quy ước mở rộng cách nói như ở (b) liệu có phải quy hoạch tuyến tính (không cần chấp nhận được) luôn có lỗ hổng đối ngẫu bằng 0 hay không? Thí dụ sau đây cho câu trả lời phủ định.

THÍ DỤ 2.2.1. Cặp quy hoạch đối ngẫu sau đây cùng không chấp nhận được, tức là không có nghiệm chấp nhận được,

$$\begin{array}{ll} \min -2x_1 + x_2, & \max y_1 - 2y_2, \\ x_1 - x_2 \leq 1, & y_1 \leq 0, \\ -x_1 + x_2 \leq -2, & y_2 \leq 0, \\ x_1 \geq 0, & y_1 - y_2 \leq -2, \\ x_2 \geq 0, & -y_1 + y_2 \leq 1. \end{array}$$

Trường hợp này có thể quy ước coi là mục tiêu tối ưu của bài toán min là $+\infty$ và của bài toán max là $-\infty$. Vậy ta có lỗ hổng đối ngẫu là rộng vô cùng.

(d) Lý thuyết đối ngẫu cho một phương tiện tuyệt vời để kiểm tra nghiệm tối ưu. Giả sử bạn phải giải một quy hoạch cỡ lớn rất khó để giúp bà giám đốc quyết định chiến lược cạnh tranh của công ti. Sau khi áp dụng

phương pháp đơn hình và tìm được nghiệm tối ưu x^* , làm sao bạn thuyết phục được bà giám đốc tính đúng đắn của nó. Vì quyết định kinh tế dựa vào x^* là mang tính sống còn nên giám đốc cần sự đảm bảo. Kiểm tra lại cả quá trình tính toán chẳng, đây là việc không tưởng! Bạn hãy áp dụng lý thuyết đối ngẫu. Theo chứng minh Định lý đối ngẫu mạnh 2.2.4, khi giải quy hoạch gốc bằng phương pháp đơn hình bạn cũng đồng thời có nghiệm tối ưu của bài toán đối ngẫu là y^* , với $y_i^* = -d_i^*$, $i = 1, \dots, m$. Bạn chỉ cần chỉ định lý đó cho bà giám đốc và kiểm tra cho bà ta thấy x^* là chấp nhận được (để), y^* là chấp nhận được cho bài toán đối ngẫu (cũng dễ như vậy) và $(y^*)^T b = c^T x^*$ (còn dễ hơn).

(c) Lý thuyết đối ngẫu còn giúp xây dựng các thuật toán liên quan đến tính đối ngẫu mà trong nhiều trường hợp cho lời giải nhanh hơn phương pháp đơn hình thông thường. Thí dụ đơn giản như trường hợp chưa biết nghiệm cơ sở chấp nhận được của bài toán gốc nhưng lại biết nghiệm cơ sở chấp nhận được của bài toán đối ngẫu thì ta có thể thay cho việc làm cả hai pha đối với bài toán gốc bằng cách áp dụng chỉ pha II cho bài toán đối ngẫu.

2.2.3 Độ lệch bù

Định lý đối ngẫu mạnh cho ta tiêu chuẩn để kiểm tra tính tối ưu như ở Nhận xét 2.2.1 (d). Nói cách khác nó cho ta điều kiện tối ưu: $c^T x = (y)^T b$. Điều kiện này biểu thị qua giá trị hàm mục tiêu. Từ định lý đối ngẫu cũng suy được ra điều kiện tối ưu ở dạng biểu thị qua các dữ kiện của bài toán. Điều kiện này thường gọi là *điều kiện độ lệch bù* (complementary slackness). Trong kinh tế định lý này thường có tên là *định lý cân bằng* (equilibrium theorem) vì trạng thái tối ưu có ý nghĩa là trạng thái cân bằng kinh tế.

ĐỊNH LÝ 2.2.5 (điều kiện độ lệch bù). *Giả sử x và y là nghiệm chấp nhận được của bài toán gốc và bài toán đối ngẫu, tương ứng. Khi đó*

x và y là tối ưu khi và chỉ khi

$$y_i(a_i^T x - b_i) = 0 \quad \forall i, \quad (2.10)$$

$$(c_j - y^T A_j)x_j = 0 \quad \forall j. \quad (2.11)$$

CHỨNG MINH. Ở chứng minh Định lý 2.2.1 ta đã đặt $u_i = y_i(a_i^T x - b_i)$ và $v_j = (c_j - y^T A_j)x_j$ và thấy là theo định nghĩa bài toán đối ngẫu ta có $u_i \geq 0, v_j \geq 0 \quad \forall i, \forall j$. Đồng thời ta cũng có

$$c^T x - y^T b = \sum_i u_i + \sum_j v_j.$$

Theo định lý đối ngẫu mạnh, nếu x và y là tối ưu thì $c^T x = y^T b$. Do đó $u_i = v_j = 0 \quad \forall i, \forall j$. Ngược lại, nếu $u_i = v_j = 0 \quad \forall i, \forall j$ thì $c^T x = y^T b$. Theo Hệ quả 2.2.3, x và y cũng là tối ưu. $\square$

Các điều kiện nêu ở Định lý 2.2.5 gọi là điều kiện độ lệch bù vì từ đó ta có

$$a_i^T x < b_i \quad \Rightarrow \quad y_i = 0, \quad (2.12)$$

$$y^T A_j < c_j \quad \Rightarrow \quad x_j = 0. \quad (2.13)$$

(2.12) nghĩa là ràng buộc ở bài toán gốc “lệch” khỏi thỏa chặt thì biến đối ngẫu tương ứng “bù” lại phải không lệch khỏi thỏa chặt, tức là bằng 0 (và tương tự cho điều kiện (2.13)). Điều này rất tự nhiên nếu nói theo ngôn ngữ bài toán mở đầu ở Mục 2.1. Một ràng buộc không thỏa chặt tại điểm tối ưu thì có thể bỏ đi mà không ảnh hưởng đến giá trị mục tiêu tối ưu, và do đó không cần phải đặt giá y_i khác không cho sự vi phạm ràng buộc đó.

NHẬN XÉT 2.2.1 (a) Nếu bài toán gốc ở dạng chính tắc thì (2.10) luôn thỏa với mọi x chấp nhận được nên điều kiện độ lệch bù cho quy hoạch chính tắc chỉ còn là

$$(c_j - y^T A_j)x_j = 0 \quad \forall j.$$

(b) Điều kiện độ lệch bù có thể cho ta có ngay nghiệm tối ưu của quy hoạch tuyến tính khi biết nghiệm tối ưu của bài toán đối ngẫu. Hãy xét thí dụ sau

THÍ DỤ 2.2.2. Xét cặp bài toán đối ngẫu

$$\begin{array}{ll} \min 13x_1 + 10x_2 + 6x_3, & \max 8y_1 + 3y_2, \\ 5x_1 + x_2 + 3x_3 = 8, & 5y_1 + 3y_2 \leq 13, \\ 3x_1 + x_2 = 3, & y_1 + y_2 \leq 10, \\ x_1, x_2, x_3 \geq 0. & 3y_1 \leq 6. \end{array}$$

Ngay dưới đây ta sẽ kiểm tra là $x^* = (1, 0, 1)$ là nghiệm tối ưu của bài toán gốc. Trước mắt ta thấy nó chấp nhận được (và giả sử nó tối ưu) và sử dụng điều kiện độ lệch bù (2.11) (chỉ có (2.11) vì bài toán gốc có dạng chính tắc)

$$(c_j - y^T A_j)x_j^* = 0 \quad j = 1, 2, 3.$$

Vì $x_2^* = 0$ nên điều kiện thỏa với $j = 2$. Còn $x_1^* > 0$ và $x_3^* > 0$ nên ta có, theo điều kiện độ lệch bù

$$\begin{aligned} 5y_1 + 3y_2 &= 13, \\ 3y_1 &= 6. \end{aligned}$$

Giải hệ này ta được $y_1^* = 2$ và $y_2^* = 1$. Tính giá trị mục tiêu đối ngẫu ta được $8y_1^* + 3y_2^* = 19$. Vì x^* là chấp nhận được, $y = (y_1^*, y_2^*)$ là tối ưu theo điều kiện độ lệch bù. Giá trị mục tiêu ứng với x^* cũng là 19 nên x^* cũng là tối ưu (của bài toán gốc).

Nhận xét rằng x^* là nghiệm tối ưu cơ sở không suy biến (vì có $2 = n - m$ thành phần dương). Trường hợp ta có một nghiệm tối ưu cơ sở suy biến thì điều kiện độ lệch bù giúp rất ít cho việc xác định nghiệm tối ưu cho bài toán đối ngẫu như người ta có thể chứng minh một cách tổng quát được.

2.3 THUẬT TOÁN ĐƠN HÌNH ĐỐI NGẪU

Ở Mục 2.2, khi chứng minh định lý đối ngẫu mạnh ta đã biết là giải một bài toán quy hoạch tuyến tính thì ta cũng được nghiệm của bài toán đối ngẫu. Ở nhận xét (e) cuối mục đó ta cũng đã chú ý là nếu bài toán gốc khó giải (chẳng hạn chưa có từ vựng chấp nhận được để xuất phát) ta có thể giải bằng cách áp dụng thuật toán đơn hình cho bài toán đối ngẫu. Tất nhiên do có đối ngẫu mạnh, cuối cùng ta vẫn được nghiệm tối ưu của bài toán gốc. Bây giờ ta xét cụ thể điều này qua một thí dụ

THÍ DỤ 2.3.1. Giả sử ta phải giải quy hoạch

$$\begin{aligned} \min \quad & x_1 + x_2, \\ & -2x_1 - x_2 \leq 4, \\ & -2x_1 + 4x_2 \leq -8, \\ & -x_1 + 3x_2 \leq -7, \\ & x_1, x_2 \geq 0. \end{aligned} \quad (2.14)$$

Vì trong các vế phải b_i có số âm nên khi viết thành từ vựng (sau khi đưa biến bù w_1, w_2 , vào) ta được từ vựng cơ sở không chấp nhận được. Nếu muốn dùng đơn hình ta phải qua pha I. Để tránh việc đó ta hãy viết bài toán đối ngẫu của (2.14) và biến đổi nó về dạng chuẩn ở bên phải (y_i mới là $-y_i$ cũ).

$$\begin{aligned} \max \quad & \zeta = 4y_1 - 8y_2 - 7y_3, & \min \quad & -\zeta = 4y_1 - 8y_2 - 7y_3, \\ & y_1 \leq 0, & & 2y_1 + 2y_2 + y_3 \leq 1, \\ & y_2 \leq 0, & & y_1 - 4y_2 - 3y_3 \leq 1, \\ & y_3 \leq 0, & & y_1, y_2, y_3 \geq 0. \\ & & & -2y_1 - 2y_2 - y_3 \leq 1, \end{aligned} \quad (2.15)$$

$$-y_1 + 4y_2 + 3y_3 \leq 1.$$

Dưới đây ta sẽ viết từ vựng và thực hiện thuật toán đơn hình. Ở cột bên trái luôn là quy hoạch gốc, cột bên phải là quy hoạch đối ngẫu. Trước hết ta viết cặp đối ngẫu (2.14) và (2.15) ở dạng từ vựng

$$\begin{array}{ll} z = \frac{x_1}{w_1 = 4} + \frac{x_2}{+ 2x_1 + x_2} & -\zeta = \frac{4y_1}{z_1 = 1 - 2y_1} - \frac{8y_2}{- 2y_2} - \frac{7y_3}{- y_3} \\ w_2 = - 8 + 2x_1 - 4x_2 & z_2 = 1 - y_1 + 4y_2 + 3y_3 \\ w_3 = - 7 + x_1 - 3x_2 & \end{array}$$

Nhận xét là từ vựng đối ngẫu nhận được bằng cách chuyển vị (tức là đổi hàng thành cột) và có thay đổi dấu từ vựng gốc. Việc thay dấu bao gồm giữ nguyên dấu ở hàng mục tiêu và các hằng số (tức là cột vế phải) và đổi dấu ở các hệ số khác (tức là các a_{ij}).

Vì từ vựng gốc không chấp nhận được nhưng từ vựng đối ngẫu chấp nhận được nên ta áp dụng thuật toán đơn hình (với quy tắc xoay Dantzig) lên từ vựng đối ngẫu. Tuy vậy ta vẫn đồng thời thực hiện phép toán tương ứng lên từ vựng gốc để theo dõi mối quan hệ. Dễ thấy biến vào là y_2 và biến ra là z_1 . Vì w_2 (ứng với ràng buộc gốc thứ hai) là biến bù của y_2 và x_1 là biến bù của z_1 , nên cặp w_2, x_1 sẽ là biến vào/ra tương ứng ở quy hoạch gốc. Nhưng x_1 đang ở ngoài cơ sở nên phải là biến vào (ngược với z_1) và w_2 đang ở trong cơ sở phải là biến ra (ngược với y_2).

Thực hiện phép xoay đã chọn cho cả hai từ vựng ta được

$$\begin{array}{ll} z = 4 + 0,5w_2 + 3x_2 & -\zeta = - 4 + 12y_1 + 4z_1 - 3y_3 \\ w_1 = 12 + w_2 + 5x_2 & y_2 = 0,5 - y_1 - 0,5z_1 - 0,5y_3 \\ x_1 = 4 + 0,5w_2 + 2x_2 & z_2 = 3 - 5y_1 - 2z_1 + y_3 \\ w_3 = - 3 + 0,5w_2 - x_2 & \end{array}$$

(Chú ý là hằng số hàng mục tiêu ở hai từ vựng là trái dấu.)

Thực hiện bước đơn hình tiếp theo trên từ vừng đối ngẫu ta thấy y_3 vào và y_2 ra. Tương ứng ở từ vừng gốc là w_2 vào và w_3 ra:

$$\begin{array}{rcl} z & = & 7 + w_3 + 4x_2 \\ w_1 & = & 18 + 2w_3 + 7x_2 \\ x_1 & = & 7 + w_3 + 3x_2 \\ w_2 & = & 6 + 2w_3 + 2x_2 \end{array} \qquad \begin{array}{rcl} -\zeta & = & -7 + 18y_1 + 7z_1 + 6y_2 \\ y_3 & = & 1 - 2y_1 - z_1 - 2y_2 \\ z_2 & = & 4 - 7y_1 - 3z_1 - 2y_2 \end{array}$$

Đến đây ta thấy cả hai từ vừng đều đã tối ưu. Chú ý rằng ở mỗi bước quy luật về chuyển vị và đổi dấu ở hai từ vừng đều như nhau.

Tất nhiên ta chỉ cần giải quy hoạch đối ngẫu rồi áp dụng định lý đối ngẫu mạnh cũng được ngay lời giải cho bài toán gốc: mục tiêu tối ưu là $z = \zeta = 7$, nghiệm tối ưu là $(x_1^*, x_2^*) = (d_1^*, d_2^*) = (7, 0)$. (Hãy so sánh với $y_1^* = -d_1^*$ ở chứng minh định lý đối ngẫu mạnh!) Nhưng bằng nhận xét mối quan hệ trên đây giữa từ vừng gốc và từ vừng đối ngẫu cũng như giữa hai phép xoay ta thấy có thể làm ngược lại là chỉ viết dãy từ vừng gốc nhưng lại thực hiện phép xoay tương ứng với phép xoay đơn hình ở từ vừng đối ngẫu (mà không cần viết dãy từ vừng đối ngẫu). Thuật toán này gọi là *thuật toán đơn hình đối ngẫu* (dual simplex algorithm).

THUẬT TOÁN ĐƠN HÌNH ĐỐI NGẪU

(thực hiện trên từ vừng gốc mỗi bước lặp như sau)

- Biến ra là biến có hằng số về phải âm nhất (tức là âm và modul lớn nhất). Nếu không có biến như vậy (mà từ vừng đối ngẫu là chấp nhận được, tức là các hệ số hằng mục tiêu gốc đều không âm) thì từ vừng hiện hành đã là tối ưu.

- Để chọn biến vào ta xét tỉ số dương lớn nhất giữa các hệ số hàng ra đã chọn và hệ số tương ứng ở hàng mục tiêu. Biến ứng với tỉ số dương lớn nhất là biến vào.

- Phép xoay được thực hiện hoàn toàn như phép xoay đơn hình.

GIẢI THÍCH HÌNH HỌC THUẬT TOÁN ĐƠN HÌNH ĐỐI NGẪU

Ta đã định nghĩa thuật toán đơn hình đối ngẫu là thuật toán thực hiện trên quy hoạch gốc khi từ vùng xuất phát của nó không chấp nhận được, tương ứng với thuật toán đơn hình thực hiện trên quy hoạch đối ngẫu với từ vùng xuất phát chấp nhận được.

Ta cũng đã nhận xét là ở mỗi bước lặp, nếu ta viết cả hai từ vùng gốc và đối ngẫu ra thì thấy chúng có tính chuyển vị lẫn nhau với sự đổi dấu chỉ ở phần tử góc trên bên trái (giao của hàng mục tiêu và cột số hạng tự do) và ở các hệ số a_{ij} , còn ở hàng mục tiêu và cột số hạng tự do thì dấu không đổi (chỉ chuyển vị).

Về mặt hình học ta hãy xem thuật toán có ý nghĩa thế nào đối với bài toán gốc.

- Đỉnh xuất phát là đỉnh không chấp nhận được, tức là giao của các diện của tập lồi đa diện chấp nhận được, nhưng nằm ngoài tập này. Điều này rõ ràng vì từ vùng xuất phát không chấp nhận được, tức là nó xác định một nghiệm cơ sở không chấp nhận được.

- Qua mỗi bước lặp, đúng một biến vào cơ sở và một biến ra khỏi cơ sở. Việc này có nghĩa là đi từ một đỉnh sang đỉnh kế với nó như ở Mục 1.3.4 đã giải thích. Bước lặp tương ứng ở quy hoạch đối ngẫu làm mục tiêu ζ tăng ($-\zeta$ giảm) mà ζ luôn bằng giá trị mục tiêu z của bài toán gốc. Vậy qua mỗi bước lặp của thuật toán đơn hình đối ngẫu, mục tiêu gốc z tăng lên (chứ không cải thiện hơn). Tuy vậy, nói chung là đỉnh mới sẽ “gần” vào miền chấp nhận được hơn, vì các hệ số mục tiêu âm ở quy hoạch đối ngẫu “ít đi” (ta đã chọn hệ số âm nhất cho biến vào!) tức là các số âm trong cột số hạng tự do của từ vùng gốc “ít đi”, đồng nghĩa với sự “giảm tính không chấp nhận được” của đỉnh tương ứng với từ vùng.

- Thuật toán ngừng khi từ vùng đối ngẫu tối ưu tức là hệ số hàng mục tiêu đều không âm (còn cột số hạng tự do luôn không âm ở mỗi bước lặp). Lúc đó từ vùng gốc có cột số hạng tự do lần đầu tiên trở

thành không âm, tức là đỉnh chấp nhận được. Còn hệ số hàng mục tiêu của nó là cột số hạng tự do của từ vựng đối ngẫu nên đã luôn không âm qua mọi bước lặp. Vậy từ vựng gốc cũng là tối ưu.

Tóm lại, thuật toán đơn hình đối ngẫu xuất phát từ một đỉnh không chấp nhận được, nhưng tương ứng với một đỉnh chấp nhận được của quy hoạch đối ngẫu. Qua mỗi bước lặp, ta chuyển sang đỉnh kề “gần hơn” theo nghĩa tiến dần về miền chấp nhận được, với sự trả giá là mục tiêu kém đi (tức là tăng lên). Đỉnh đầu tiên vào được miền chấp nhận chính là đỉnh tối ưu. Ta thấy rõ sự đối lại với *thuật toán đơn hình gốc* (primal simplex algorithm), ở đây ta thêm chữ “gốc” để nhấn mạnh sự khác nhau. Ở thuật toán đơn hình gốc ta đi từ một đỉnh chấp nhận được, qua các đỉnh kề nhưng luôn luôn ở trong miền chấp nhận được, sao cho qua mỗi bước hàm mục tiêu giảm, cho đến đỉnh tối ưu thì dừng.

2.4 PHA MỘT DỰA TRÊN TÍNH ĐỐI NGẪU

Nhờ thuật toán đơn hình đối ngẫu, trường hợp từ vựng xuất phát chưa chấp nhận được nhưng từ vựng đối ngẫu tương ứng chấp nhận được, ta sẽ tránh được pha I bổ trợ mà tiến hành ngay thuật toán.

Nếu cả từ vựng đối ngẫu cũng không chấp nhận được thì sao? Ngay sau đây ta sẽ thấy rằng thuật toán đơn hình đối ngẫu cũng cho ta một pha I hay không kém gì pha I xét ở Mục 2.3. Ý tưởng đơn giản như sau. Ta cải biên bài toán gốc để đối ngẫu với bài toán cải biên có từ vựng xuất phát chấp nhận được. Muốn vậy, vì cột số hạng tự do (quyết định về tính chấp nhận được) của từ vựng đối ngẫu là các hệ số hàng mục tiêu gốc, ta chỉ việc cải biên hàm mục tiêu gốc để các hệ số đều không âm.

Đã có từ vựng đối ngẫu (cải biên) chấp nhận được rồi thì ta thực hiện được thuật toán đơn hình đối ngẫu trên bài toán cải biên và đi đến từ vựng tối ưu của bài toán cải biên. Vì bài toán cải biên có chung miền ràng buộc với bài toán ban đầu nên từ vựng cuối cùng này là

chấp nhận được của bài toán ban đầu.

Lúc này ta có thể tiến hành pha II là thuật toán đơn hình gốc trên quy hoạch ban đầu. Thuật toán hai pha này được gọi là *thuật toán đối ngẫu - gốc* (dual - primal algorithm) hoặc thật đầy đủ là *thuật toán đơn hình hai pha đối ngẫu - gốc* (dual - primal two - phase simplex algorithm). Nhận xét thêm rằng pha I của thuật toán đối ngẫu - gốc cũng giúp ta phát hiện được bài toán gốc là không chấp nhận được. Thật vậy, nếu bài toán đối ngẫu của bài toán gốc cải biên là không giới nội thì bài toán gốc cải biên, và do đó bài toán gốc cũng vậy, là không chấp nhận được.

Ta cũng thấy rằng có thể thực hiện hai pha theo thứ tự ngược lại như sau để được thuật toán mới. Ta cải biên cột số hạng tự do của từ vựng gốc để từ vựng này trở thành chấp nhận được. Khi đó ta tiến hành được thuật toán đơn hình gốc để nhận được từ vựng tối ưu của bài toán cải biên. Từ vựng đối ngẫu tương ứng là từ vựng tối ưu của bài toán đối ngẫu của bài toán cải biên, nhưng nó vẫn bảo đảm là từ vựng chấp nhận được của bài toán đối ngẫu ban đầu (vì cải biên cột số hạng tự do ở từ vựng gốc xuất phát tương ứng là cải biên hàng mục tiêu của từ vựng đối ngẫu; riêng tính chấp nhận được không bị ảnh hưởng). Sau khi thực hiện pha I như vậy trên từ vựng gốc, pha II sẽ là thuật toán đơn hình trên từ vựng đối ngẫu chấp nhận được vừa có để đi đến từ vựng đối ngẫu tối ưu. Từ vựng tương ứng bên gốc sẽ là từ vựng tối ưu cần tìm. Thuật toán hai pha này được gọi là *thuật toán gốc - đối ngẫu* (primal - dual algorithm) hoặc có tên đầy đủ là *thuật toán đơn hình hai pha gốc - đối ngẫu* (primal - dual two - phase simplex algorithm).

BÀI TẬP CHƯƠNG 2

2.1. Tìm đối ngẫu của quy hoạch

$$\max c^T x,$$

$$a \leq Ax \leq b,$$

$$l \leq x \leq u.$$

2.2. Xét quy hoạch tuyến tính chính tắc với A có các hàng độc lập tuyến tính. Chứng minh hoặc tìm phản thí dụ cho mỗi điều khẳng định sau đây.

(a) Giả sử x là nghiệm cơ sở chấp nhận được. Giả sử với mỗi cơ sở tương ứng x , nghiệm cơ sở tương ứng của bài toán đối ngẫu là không chấp nhận được. Khi đó giá trị mục tiêu tối ưu nhỏ hơn (hẳn) $c^T x$.

(b) Bài toán đối ngẫu của bài toán bổ trợ ở pha I của phương pháp đơn hình hai pha luôn chấp nhận được.

(c) Giả sử y_i là biến đối ngẫu tương ứng ràng buộc đẳng thức thứ i . Loại bỏ ràng buộc đẳng thức thứ i ở bài toán gốc là tương đương với đưa thêm vào ràng buộc $y_i = 0$ ở bài toán đối ngẫu.

2.3. Giải các quy hoạch ở Bài tập 1.15 (a, b) bằng thuật toán đơn hình đối ngẫu.

2.4. Xét quy hoạch

$$\max \sum_j p_j x_j,$$

$$\sum_j q_j x_j \leq \beta,$$

$$x_j \leq 1, j = 1, \dots, n,$$

$$x_j \geq 0, j = 1, \dots, n.$$

Giả sử các p_j, q_j dương và $\sum_j p_j = \sum_j q_j = 1$. Giả sử

$$\frac{p_1}{q_1} < \frac{p_2}{q_2} < \dots < \frac{p_n}{q_n}$$

và $\beta > 0$ đủ nhỏ. Kí hiệu

$$k = \min \{j: q_j + 1 + \dots + q_n \leq \beta\},$$

y_0 là biến đối ngẫu tương ứng ràng buộc chứa β và y_j là biến đối ngẫu tương ứng ràng buộc $x_j \leq 1$.

Dùng lý thuyết đối ngẫu chứng minh rằng nghiệm tối ưu của bài toán gốc và đối ngẫu tương ứng là

$$x_j = \begin{cases} 0 & j < k, \\ \frac{\beta - q_{k+1} - \dots - q_n}{q_k} & j = k, \\ 1 & j > k; \end{cases} \quad y_j = \begin{cases} \frac{p_k}{q_k} & j = 0, \\ 0 & 0 < j \leq k, \\ q_j \left(\frac{p_j}{q_j} - \frac{p_k}{q_k} \right) & j > k. \end{cases}$$

2.5. Giả sử A là ma trận vuông đối xứng. Xét quy hoạch

$$\min c^T x, Ax \geq c, x \geq 0.$$

Chứng minh rằng nếu x thỏa $Ax = c, x \geq 0$ thì x là nghiệm tối ưu.

2.6. Xét quy hoạch $\min c^T x, Ax \leq b$. Ta đã biết điều kiện độ lệch bù là $y(a_i^T x - b_i) = 0, i = 1, \dots, m$, ở đây y_i là biến đối ngẫu. Điểm x^0 được gọi là thỏa *điều kiện độ lệch bù chặt* (strict complementary slackness condition) nếu x^0 thỏa điều kiện độ lệch bù và

$$\text{nếu } a_i^T x^0 = b_i, \text{ thì } y_i > 0.$$

Giả sử x^0 là nghiệm cơ sở tối ưu không suy biến và thỏa điều kiện độ lệch bù chặt. Chứng minh rằng x^0 là nghiệm tối ưu duy nhất.

2.7. Xét quy hoạch gốc $\min c^T x, Ax \geq b, x \geq 0$. Hãy viết quy hoạch đối ngẫu ở dạng bài toán tìm cực tiểu. Tìm điều kiện về A, b, c để bài toán đối ngẫu trùng với bài toán gốc. Cho thí dụ cụ thể thỏa điều kiện này.

2.8. Xét cặp quy hoạch tuyến tính chính tắc và đối ngẫu của nó.

(a) Chứng minh rằng nếu một trong hai quy hoạch có nghiệm tối ưu duy nhất và không suy biến thì quy hoạch kia cũng vậy.

(b) Giả sử ta có một cơ sở tối ưu không suy biến của quy hoạch gốc và biết rằng một hệ số mục tiêu của biến cơ sở là 0. Khi đó khẳng định rằng (a) suy ra điều gì? Có cơ sở tối ưu khác hay không?

CHƯƠNG 3

BÀI TOÁN DÒNG TRÊN MẠNG

Nhiều bài toán quy hoạch tuyến tính có thể quy về bài toán làm cực tiểu phí tổn vận chuyển hàng trong một mạng (gồm các nút và các cung đường) sao cho đảm bảo được các nhu cầu ở một số nút khi đã biết nguồn cung cấp tại một số nút khác. Các bài toán như vậy được gọi là các *bài toán dòng trên mạng* (network flow problem) hoặc *bài toán chuyển vận* (transshipment problem). Đây là lớp bài toán quan trọng nhất và hay gặp nhất trong quy hoạch tuyến tính. Lớp này bao gồm các bài toán quen thuộc trong thực tế như: bài toán vận tải, các bài toán mạng điện và mạng giao thông, các bài toán quản lý và phân bố vật tư, bài toán bổ nhiệm, bài toán kế hoạch tài chính, bài toán đường ngắn nhất, bài toán dòng lớn nhất....

Vì cũng là quy hoạch tuyến tính nên các bài toán dòng trên mạng có thể được giải bằng bất kỳ một thuật toán cho quy hoạch tuyến tính nào, chẳng hạn bằng các kiểu thuật toán đơn hình đã xét. Tuy nhiên, tận dụng những cấu trúc đặc biệt của các bài toán mạng sẽ làm cho phương pháp đơn hình được đơn giản và thực hiện nhanh hơn khi áp dụng các thuật toán đơn hình tổng quát rất nhiều. Đó là nội dung của Chương 3 này.

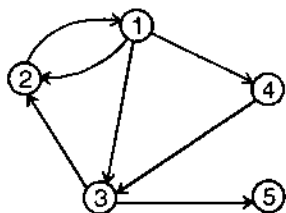
3.1 ĐỒ THỊ VÀ CÂY

3.1.1 Đồ thị

Đồ thị vô hướng (undirected graph) $\mathcal{G}$ là một cặp hai tập $(\mathcal{N}, \mathcal{A})$, ở đây $\mathcal{N}$ là tập các nút (node) và $\mathcal{A}$ là tập các cung (arc) không có hướng. Cung không có hướng, hoặc cũng gọi là *cung vô hướng* (undirected arc) là một cặp không kể đến thứ tự hai nút khác nhau i và j nào đó, $i \in \mathcal{N}, j \in \mathcal{N}$ kí hiệu là (i, j) . Vậy $(i, j) = (j, i)$ trong đồ thị vô hướng. Cung cũng còn gọi là *cạnh* (edge). Cung (i, j) được gọi là *nối* các nút i và j . **Cấp** (degree) của một nút là số cung nối tới nó. **Cấp của đồ thị** là cấp cực đại trong các cấp của các nút của nó. Một **đường đi** (path) từ nút i_1 đến nút i_t là bộ t nút khác nhau $i_1, i_2, \dots, i_t$ sao cho $(i_k, i_{k+1}) \in \mathcal{A}$ với $k = 1, \dots, t-1$. **Chu trình** (cycle hoặc loop) là bộ t nút $i_1, i_2, \dots, i_t$ sao cho $i_1, \dots, i_{t-1}$ là một đường đi, $i_t = i_1$ và có ít nhất ba nút khác nhau (tức là $t - 1 \geq 3$). Đồ thị vô hướng được gọi là **liên thông** (connected) nếu ứng mỗi cặp $i, j \in \mathcal{N}$ đều có một đường đi từ i đến j . Số các nút của đồ thị thường kí hiệu là $|\mathcal{N}|$ hoặc n , còn số các cung là $|\mathcal{A}| = m$.

Đồ thị có hướng (directed graph hoặc digraph) cũng là một cặp

$(\mathcal{N}, \mathcal{A}) := \mathcal{G}$, nhưng mỗi cung là một cặp có thứ tự. Vậy $(i, j) \neq (j, i)$. Nhưng đồ thị không được chứa cung "tự nối" (i, i) . Thí dụ ở H 3.1 là một đồ thị có hướng $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ với $\mathcal{N} = \{1, 2, 3, 4, 5\}$, $\mathcal{A} = \{(1, 2), (2, 1), (1, 3), (3, 2), (1, 4), (4, 3), (3, 5)\}$. Ta sẽ nói là cung (i, j) đi từ nút i đến nút j . Cung này là nối tới cả nút i và j . Đồ thị vô hướng tương ứng với một đồ thị có hướng là đồ thị nhận được khi không



H. 3.1

tính đến hướng trên các cung nữa. Đồ thị có hướng là *liên thông* nếu đồ thị vô hướng tương ứng là liên thông.

Mỗi đường đi trong đồ thị vô hướng tương ứng đều gọi là một *đường đi* trong đồ thị có hướng. Nhưng đồ thị có hướng có thể chứa cả hai cung (i, j) và (j, i) , nên để xác định một đường đi phải nói rõ cả dãy nút $i_1, \dots, i_t$ và dãy cung $a_1, \dots, a_{t-1}$. Khi đó nếu một cung a_k có dạng “thuận” $a_k = (i_k, i_{k+1})$ thì ta nói a_k là *cung tiến* (forward arc) trong đường đi này. Ngược lại nếu $a_k = (i_{k+1}, i_k)$ thì a_k là *cung lùi* (backward arc). *Chu trình* cũng định nghĩa như ở đồ thị vô hướng, nhưng ở đây cho phép chu trình chỉ gồm hai nút khác nhau. Một đường đi hoặc chu trình được gọi là *có hướng* (directed path và directed cycle hoặc circuit) nếu nó chỉ chứa các cung tiến. Thí dụ ở hình H.3.1 thì 1, (1, 3), 3, (3, 2), 2, (2, 1), 1 là một chu trình có hướng. Nhưng 3, (1, 3), 1, (1, 2), 2, (3, 2), 3 là một chu trình không có hướng vì (1, 3) và (3, 2) là các cung lùi.

3.1.2 Cây

Đồ thị vô hướng được gọi là một *cây* (tree) nếu nó liên thông và không chứa chu trình. Mỗi nút cấp 1 của cây gọi là một *lá* (leaf).

ĐỊNH LÝ 3.1.1 (i) *Mỗi cây có hơn một nút thì sẽ có lá.*

(ii) *Đồ thị vô hướng là một cây khi và chỉ khi nó liên thông và có $|A| - 1$ cung.*

(iii) *Với bất kỳ hai nút $i \neq j$ trên cây đều có tồn tại duy nhất một đường đi từ i tới j .*

(iv) *Nếu thêm một cung mới vào một cây thì đồ thị mới nhận được có đúng một chu trình.*

CHỨNG MINH. (i) Giả sử phản chứng là không có lá, tức là mọi nút đều có cấp lớn hơn 1 (nếu có nút cấp 0 thì đồ thị không liên thông). Vì vậy giả sử ta

đi theo một cung nối tới một nút thì sẽ có cung khác nối tới đó để ta đi tiếp. Vì số nút hữu hạn, sẽ đến lúc ta quay lại nút đã qua, tức là có chu trình. Điều này vô lý vì đồ thị là cây.

(ii) "*chỉ khi*". Rõ ràng cây một nút có $|A| - 1$ cung. Giả sử cây có hơn một nút. Theo (i) nó có ít nhất một lá. Bỏ đi lá này và cung duy nhất nối tới nó thì đồ thị mới vẫn là một cây (vì không thể có chu trình mới và việc bỏ lá và một cung như trên cũng không thể làm mất tính liên thông). Lại tiếp tục bỏ đi một lá và một cung từ cây mới này. Quá trình sẽ thực hiện được $|A| - 1$ lần cho đến khi ta nhận được một nút duy nhất (không còn cung nữa). Vì ở mỗi lần ta bỏ đi đúng một cung, nên cây ban đầu phải có $|A| - 1$ cung.

"*khi*". Giả sử đồ thị là liên thông và có $|A| - 1$ cung. Nếu đồ thị có chu trình thì ta bỏ một cung trong chu trình mà đồ thị vẫn liên thông. Tiếp tục quá trình cho đến khi không còn chu trình mà đồ thị vẫn liên thông, tức là một cây. Vì trong quá trình vừa làm ta không bỏ nút nào nên cây cuối cùng vẫn có $|A|$ nút. Theo phần "*chỉ khi*" nó phải có $|A| - 1$ cung, tức là bằng số cung ban đầu. Điều này lại có nghĩa là chưa có cung nào bị bỏ đi cả. Vậy đồ thị ban đầu đã là một cây.

(iii) Giả sử có hai đường đi khác nhau từ i tới j . Khi đó xét tất cả các nút và cung của hai đường đi này ta phải có một hoặc nhiều chu trình. Điều này mâu thuẫn với định nghĩa cây.

(iv) Giả sử ta bổ sung cung vô hướng (i, j) vào một cây. Theo (ii) đồ thị nhận được có $|A|$ cung và không thể là cây. Do đó đồ thị có chu trình. Mọi chu trình được tạo do việc thêm cung (i, j) đều phải chứa cung (i, j) và một đường từ i tới j . Vì (theo (iii)) chỉ có đúng một đường đi từ i tới j nên chỉ có đúng một chu trình được tạo. $\square$

Cây bao trùm (spanning tree), hoặc cũng gọi là *cây tối đại* (maximal tree), của đồ thị liên thông, vô hướng $G = (A, A)$ là một cây trong G mà chứa tất cả các nút của G . Vậy cây bao trùm là cây T có dạng $T = (A, A_1)$ với $A_1 \subset A$.

ĐỊNH LÝ 3.1.2. Giả sử $G = (N, A)$ là một đồ thị vô hướng liên thông và $A_0 \subset A$. Nếu các cung của A_0 không tạo thành một chu trình nào thì có thể bổ sung thêm các cung của đồ thị G vào A_0 để được tập mới A_1 sao cho (N, A_1) là một cây bao trùm.

CHỨNG MINH. Nếu chính G là một cây thì lấy $A_1 = A$ (ta được chính G là cây bao trùm). Giả sử G có chứa một chu trình thì chu trình này phải có cung nằm ngoài A_0 . Ta bỏ đi một cung của chu trình và không thuộc A_0 . Nếu còn chu trình ta lại tiếp tục quá trình. Cuối cùng ta được một đồ thị liên thông (N, A_1) không còn chu trình, tức là một cây bao trùm. Vì các cung của A_0 không bị bỏ đi lần nào nên $A_0 \subset A_1$. $\square$

3.2 PHÁT BIỂU BÀI TOÁN DÒNG TRÊN MẠNG

3.2.1 Bài toán dòng trên mạng

Mạng (network) là một đồ thị có hướng $G = (N, A)$ cùng với một số thông tin về số lượng bổ sung như: các số thực b_i , $i \in N$, biểu thị *nguồn từ ngoài* (source hoặc external supply) hoặc gọi tắt là *nguồn* (supply); các số $u_{ij} \geq 0$ (có thể vô hạn) biểu thị *tải nặng* (capacity) của cung $(i, j) \in A$ và các số thực c_{ij} biểu thị *cước phí* (cost) một đơn vị của dòng (chuyển vận) trên cung (i, j) (cost per unit).

Ta chọn biến của bài toán là số thực x_{ij} biểu thị (lượng của) dòng trên cung (i, j) .

Nút i sẽ gọi là một *nguồn* (source) nếu $b_i > 0$, khi đó b_i gọi là *nguồn cung* hoặc gọi tắt là *nguồn* (supply). Nếu $b_i < 0$ thì nút i được gọi là một *điểm hút* (sink) và $|b_i|$ được gọi là *nhu cầu* (demand) của nút i .

Một cách tự nhiên ta đặt hai điều kiện sau đây.

(a) Luật bảo toàn dòng : (tổng lượng) dòng vào nút i phải bằng (tổng lượng) dòng ra khỏi nút i :

$$b_i + \sum_{j \in I(i)} x_{ji} = \sum_{j \in O(i)} x_{ij}, \quad \forall i \in \mathcal{N}, \quad (3.1)$$

ở đây

$$I(i) := \{j \in \mathcal{N} : (j, i) \in \mathcal{A}\},$$

$$O(i) := \{j \in \mathcal{N} : (i, j) \in \mathcal{A}\}$$

là tập tất cả các nút j có cung xuất phát từ đó tới i trong đồ thị, và tập các nút j có cung từ i đi tới, tương ứng.

(b) Dòng trên cung là không âm và không vượt quá *tải năng* (capacity) của cung :

$$0 \leq x_{ij} \leq u_{ij}, \quad \forall (i, j) \in \mathcal{A}. \quad (3.2)$$

Mọi vectơ x có các thành phần là x_{ij} , $(i, j) \in \mathcal{A}$, được gọi là một *dòng* (flow). Dòng x thỏa mãn (3.1), (3.2) được gọi là *dòng chấp nhận được* (feasible flow).

Lấy tổng cả hai vế (3.1) theo mọi $i \in \mathcal{N}$ ta được

$$\sum_{i \in \mathcal{N}} b_i = 0. \quad (3.3)$$

Điều này có nghĩa là tổng dòng từ bên ngoài vào mạng phải bằng tổng dòng từ mạng ra ngoài. Rõ ràng nếu điều này không thỏa thì bài toán là không chấp nhận được (vô nghiệm).

Mục tiêu của bài toán làm cực tiểu cước phí dòng trên mạng là

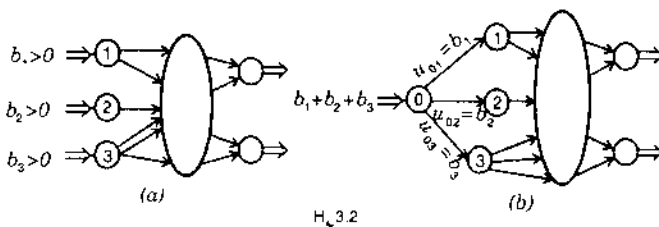
$$\min \sum_{(i,j) \in \mathcal{A}} c_{ij} x_{ij}, \quad (3.4)$$

ở đây cực tiểu lấy trên mọi dòng chấp nhận được. Như vậy là ta nhận

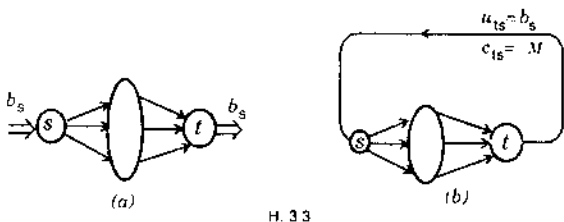
được quy hoạch tuyến tính (3.1), (3.2), (3.4). Nếu $u_{ij} = \infty$ với mọi $(i, j) \in \mathcal{A}$ thì ta có bài toán với tải năng không hạn chế (uncapacitated problem). Ngược lại thì ta có bài toán với tải năng hạn chế (capacitated problem). Bài toán thứ nhất là một quy hoạch tuyến tính dạng chính tắc.

3.2.2 Quy gọn bài toán

Có nhiều dạng khác nhau của bài toán dòng trên mạng mà tất cả đều tương đương nhau. Ở đây ta nêu vài cách chuyển dạng bài toán.



(a) Mọi bài toán dòng trên mạng có thể đưa về dạng có đúng một nguồn và đúng một điểm hút. Thật vậy, giả sử mạng có ba nguồn như ở H.3.2 (a). Ta thêm vào nút 0 với nguồn cung là tổng $b_1 + b_2 + b_3$ và ba cung $(0, j)$, $j = 1, 2, 3$, với tải năng $u_{0j} = b_j$. Khi đó bắt buộc dòng b_j , $j=1,2,3$, phải đi trên cung $(0, j)$. Vì vậy bài toán mới ở hình (b) là tương đương và chỉ có một nút 0 là nguồn. Với điểm hút ta làm tương tự.



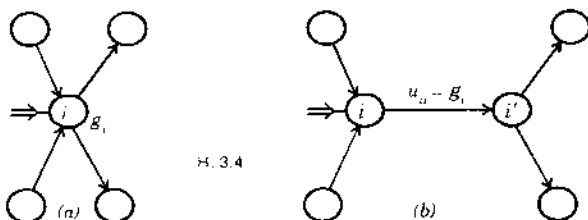
(b) Mọi bài toán dòng trên mạng có thể đưa về dạng không có nguồn hoặc không có điểm hút. Theo (a) có thể giả sử bài toán chỉ có một nguồn và một hút, giả sử có dạng ở H.3.3. Ta thêm vào cung mới (t, s) với tải năng

$u_{ts} = b_s$ và cước phí đơn vị $c_{ts} = -M$ với $M > 0$ và rất lớn. Khi đó nghiệm tối ưu của bài toán mới phải có $x_{ts} = b_s$ (vì cước phí trên cung là âm với modul lớn) hoặc bài toán là vô nghiệm. Vậy ta đã được bài toán tương đương ở H.3.3 (b) không có nguồn và không có hút.

(c) *Nguồn năng của nút* (node capacity). Giả sử trong thực tế phải có cận trên g_i cho (tổng lượng) dòng có thể vào nút i :

$$b_i + \sum_{j \in I(i)} x_{ji} \leq g_i.$$

Ta có thể loại bỏ ràng buộc về nguồn năng của nút này khỏi bài toán bằng cách tách nút i thành hai nút i và i' bởi cung (i, i') với tải năng $u_{ii'} = g_i$ (xem H.3.4).



(d) *Cận dưới của dòng trên cung*. Trong thực tế có thể nảy sinh ràng buộc về cận dưới của dòng trên cung ở dạng $x_{ij} \geq d_{ij}$. Khi đó ta có thể đưa lại về trường hợp $d_{ij} = 0$ với mọi $(i, j) \in \mathcal{A}$ (xem Bài tập 3.2).

3.2.3 Dạng ma trận của bài toán dòng trên mạng

Viết ràng buộc (3.1) về luật bảo toàn dòng ở dạng ma trận $Ax = b$, ở đây $x = (x_{ij})$, $b = (b_i)$, $i \in \mathcal{N}$, $(i, j) \in \mathcal{A}$, ta sẽ thấy ma trận A phải có cấp $n \times m$, ở đây $n = |\mathcal{N}|$ là số nút và $m = |\mathcal{A}|$ là số cung. Đồng thời, phần tử của A có dạng

$$a_{ik} = \begin{cases} 1 & \text{nếu } i \text{ là nút đầu của cung } k, \\ -1 & \text{nếu } i \text{ là nút cuối của cung } k, \\ 0 & \text{các trường hợp khác.} \end{cases}$$

Vậy mỗi cột của A có đúng hai phần tử khác 0 là 1 và -1; chỉ nút đầu và nút cuối của cung tương ứng với cột này. Khi xét các hàng ta sẽ thấy hàng i ứng với nút i và ràng buộc sẽ là

$$a_i^T x = \sum_{j \in O(i)} x_{ij} - \sum_{j \in I(i)} x_{ji} = b_i,$$

ở đây a_i^T là hàng i của ma trận A . Như vậy là các phần tử khác 0 (là 1 hoặc -1) trên hàng i của A ở cột nào thì có nghĩa là cung tương ứng cột đó có nối tới nút i (1 ứng với cung đi ra từ nút i , -1 ứng với cung đi vào nút i). Ma trận A được gọi là *ma trận nối nút - cung* (node - arc incidence matrix) hoặc gọi tắt là *ma trận nối*.

THÍ DỤ. Xét đồ thị có hướng ở H.3.1. Giả sử ta chọn thứ tự các cung là (1, 2), (2, 1), (3, 2), (4, 3), (1, 4), (1, 3), (3, 5). Khi đó ma trận nối là A .

$$A = \begin{bmatrix} 1 & -1 & 0 & 0 & 1 & 1 & 0 \\ -1 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}.$$

Nhận xét rằng tổng tất cả các hàng của A là vectơ 0. Vì vậy n hàng của A là phụ thuộc tuyến tính. Vì vậy ta có thể bỏ đi một hàng (nếu bài toán là chấp nhận được).

3.2.4 Hoàn lưu

Vectơ dòng x (chấp nhận được hoặc không chấp nhận được) thỏa mãn

$$Ax = 0$$

sẽ được gọi là một *hoàn lưu* (circulation). Vậy hoàn lưu là một dòng chạy quanh trong mạng, không có nguồn từ ngoài và cũng không có điểm hút vì $b = 0$.

Cho một chu trình C thì sẽ có một hoàn lưu tương ứng với C một

cách tự nhiên, kí hiệu là h^C , có các thành phần như sau

$$h_{ij}^C = \begin{cases} 1 & \text{nếu } (i, j) \in F, \\ -1 & \text{nếu } (i, j) \in B, \\ 0 & \text{trường hợp khác,} \end{cases}$$

ở đây F và B là tập các cung tiến và tập các cung lùi, tương ứng, của chu trình C . Vectơ dòng h^C được gọi là *hoàn lưu đơn giản* (simple circulation) ứng với chu trình C . Rõ ràng là

$$Ah^C = 0,$$

tức là h^C thực sự là một hoàn lưu. Cước phí của dòng h^C dọc theo chu trình C , tức là

$$c^T h^C = \sum_{(i,j) \in F} c_{ij} - \sum_{(i,j) \in B} c_{ij},$$

được gọi là *cước phí của chu trình C* .

Giả sử x là một vectơ dòng, θ là một số thực. Khi đó $x + \theta h^C$ được gọi là vectơ dòng nhận được từ x khi đưa thêm θ đơn vị của dòng vào chu trình C . Lúc này cước phí sẽ thêm một lượng là $\theta c^T h^C$ (bằng θ nhân với cước phí của chu trình C).

Trên H.3.5 là một chu trình và hoàn lưu đơn giản tương ứng. Giả sử ta chọn chiều của chu trình là theo thứ tự các nút 1, 2, 3, 4, 5. Khi đó các cung (4, 3) và (1, 5) là cung lùi và sẽ có dòng là -1. Chú ý là tại mỗi nút luật bảo toàn dòng đều được thỏa.

3.3 THUẬT TOÁN ĐƠN HÌNH MẠNG

Xét bài toán dòng trên mạng với tải năng không hạn chế

$$\min c^T x,$$

$$Ax = b,$$

$$x \geq 0,$$

ở đây A là ma trận nối của đồ thị có hướng $G = (N, A)$. (Bài toán có tải năng hạn chế sẽ xét ở cuối Mục 3.3.) Nhớ lại rằng ma trận A có cấp $n \times m$, với $n = |N|$, $m = |A|$, tức là chúng ta có n ràng buộc đẳng thức và m biến (ngược với kí hiệu ở quy hoạch tuyến tính tổng quát xét trước đây).

Chúng ta sẽ trình bày *thuật toán đơn hình mạng* (network simplex algorithm) của [Dantzig 1951], đồng thời so sánh với thuật toán đơn hình tổng quát ở Mục 1.2. Ta luôn có giả thiết như sau đây.

GIẢ THIẾT 3.3.1

$$(a) \sum_{i \in N} b_i = 0;$$

(b) đồ thị G là liên thông.

Nếu không có giả thiết (a) thì bài toán là vô nghiệm. Giả thiết (b) cũng tự nhiên vì nếu G không liên thông thì bài toán có thể được tách thành các bài toán con để xét độc lập nhau.

Nhớ lại rằng các hàng của ma trận nối A là phụ thuộc tuyến tính (vì tổng các hàng bằng 0). Do đó ràng buộc cuối cùng (bảo toàn dòng tại nút n) có thể bỏ đi mà không ảnh hưởng đến tập chấp nhận được. Ma trận $\tilde{A}$ cấp $(n-1) \times m$ nhận được khi bỏ hàng cuối của A được gọi là *ma trận nối thu gọn* (truncated incidence matrix). Nhận xét rằng cột của $\tilde{A}$ sẽ có dạng như sau. Các cột của A tương ứng với cung dạng (i, n) chỉ chứa một phần tử khác 0, cụ thể là bằng 1, ở hàng i . Tương tự, các cột ứng với cung có dạng (n, i) chỉ chứa một phần tử khác 0, bằng -1, ở hàng i . Còn lại, mọi cột khác của $\tilde{A}$ đều chứa hai phần tử khác không là 1 và -1. Đặt $\tilde{b} = (b_1, \dots, b_{n-1})$ thì ta chỉ cần xét ràng buộc $\tilde{A}x = \tilde{b}$. Hệ quả 3.3.2 cho thấy $\tilde{A}$ có các hàng độc lập (tức là có hạng bằng $n-1$).

3.3.1 Nghiệm cây và nghiệm cơ sở

ĐỊNH NGHĨA 3.3.1. Một vectơ dòng x thỏa $Ax = b$ được gọi là một *nghiệm cây* (tree solution) ứng với cây bao trùm T (của bài toán dòng trên mạng) nếu $x_{ij} = 0$ với mọi $(i, j) \notin T$.

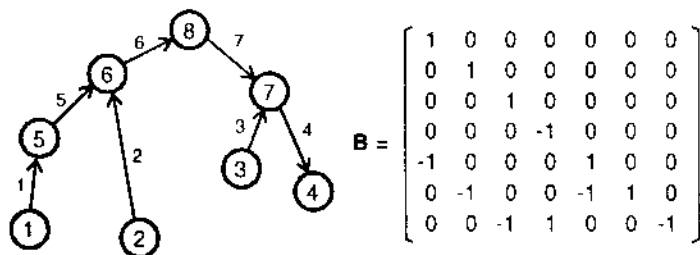
Nghiệm cây thỏa cả điều kiện $x \geq 0$ được gọi là *nghiệm cây chấp nhận được* (feasible tree solution).

ĐỊNH LÝ 3.3.1. Luôn có duy nhất một nghiệm cây ứng với một cây bao trùm cho trước.

CHỨNG MINH. Ta phải chứng minh rằng cho trước cây bao trùm T thì hệ

$$\begin{aligned} \bar{A}x &= \bar{b}, \\ x_{ij} &= 0 \text{ với mọi } (i, j) \notin T, \end{aligned}$$

luôn có nghiệm duy nhất. Giả sử B là ma trận cấp $(n-1) \times (n-1)$ chứa $n-1$ cột ứng với các cung của T . Gọi x_T là một phần gồm $n-1$ thành phần x_{ij} , $(i, j) \in T$, của x . Ta phải chỉ rằng $Bx_T = \bar{b}$ có nghiệm duy nhất, tức là B là ma trận bình thường. Ta đánh số lại các nút và cung trên T bắt đầu từ một lá theo một đường đi tới nút thứ n (là nút đã bỏ ra khỏi ma trận thu gọn $\bar{A}$). Nút này sẽ gọi là *nút gốc* (root node) hoặc *gốc* (root) của cây. Ta cũng gán cho mỗi cung $(i, j) \in T$ số tự nhiên min $\{i, j\}$ (xem H.3.6). Việc đánh số lại các nút và cung gây ra việc sắp lại thứ tự các hàng và cột của B nhưng không ảnh hưởng tới việc B là ma trận bình thường hay không.



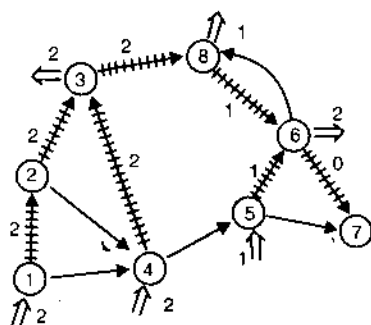
H.3.6 Đánh số các nút và cung, và ma trận B tương ứng.

Sau khi đánh số lại, cột i của B ứng với cung i , là cung có dạng (i, j) hoặc (j, i) với $j > i$. Do đó mọi phần tử khác không của cột i là ở hàng i hoặc j . Vì $j > i$, không có phần tử khác 0 nào ở trên đường chéo và phần tử ở đường chéo là khác 0. Vậy $\det B \neq 0$. $\square$

HỆ QUẢ 3.3.2. Nếu đồ thị G liên thông thì $\tilde{A}$ có tất cả $n - 1$ hàng là độc lập tuyến tính.

CHỨNG MINH. Vì G liên thông nên theo Định lý 3.1.2 sẽ có cây bao trùm T với ma trận tương ứng là B (như ở định lý trên đây). Vì B có $n - 1$ cột độc lập tuyến tính nên $\tilde{A}$ cũng có $n - 1$ cột độc lập tuyến tính. Do đó $\tilde{A}$ phải có $n - 1$ hàng độc lập tuyến tính. $\square$

Trong thực tế để giải hệ $\tilde{A}x = \tilde{b}$ ta đi từ các lá, sau khi đã đặt dòng trên các cung bên ngoài cây bằng 0, rồi đi dần về gốc n (xem hình H. 3.7).



H. 3.7

Theo luật bảo toàn dòng ở các nút, khi đi từ lá là nút 1 và 4 ta có ngay $x_{12} = 2$, $x_{43} = 2$. Tiếp theo là $x_{23} = 2$. Đến nút 3 ta phải có $x_{38} = 2$. Tiếp đến $x_{56} = 1$ và $x_{67} = 0$. Tại nút 6 ta có ngay $x_{86} = 1$. Nhận xét rằng ta đã được một nghiệm cây chấp nhận được.

So sánh với thuật toán đơn hình tổng quát áp dụng cho bài toán dòng trên mạng với $n - 1$ ràng buộc (có ma trận là $\tilde{A}$) ta thấy B chính là cơ sở. Nghiệm cây là nghiệm có các thành phần ngoài các cột của B (tức là các thành phần ngoài cơ sở), $x_{ij} \notin T$, đều bằng 0, là nghiệm cơ sở tương ứng. Vậy ta đi đến kết luận quan trọng : cây bao trùm trong bài toán mạng đóng vai trò cơ sở trong thuật toán đơn hình tổng quát và nghiệm cây là nghiệm cơ sở. Điều ngược lại cũng đúng theo định lý sau.

ĐỊNH LÝ 3.3.3. (i) Một đồ thị con của G là một cây bao trùm khi và chỉ khi các cột trong $\tilde{A}$ tương ứng với các cung của nó lập thành một cơ sở.

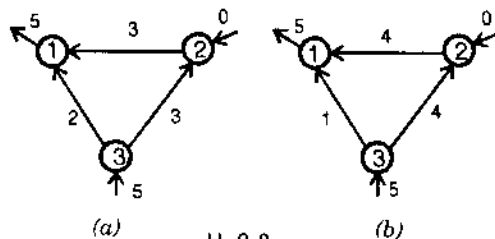
(ii) Một vectơ dòng là một nghiệm cơ sở khi và chỉ khi nó là nghiệm cây.

CHỨNG MINH. (i) Ta chỉ còn phải chứng minh “khi”. Giả sử có một đồ thị con mà các cột tương ứng trong $\tilde{A}$ lập thành cơ sở. Khi đó đồ thị con này phải có $n-1$ cung. Ta chỉ còn phải chứng minh nó chứa tất cả các nút. Nếu nó không chứa một nút i thì trong ma trận (con của $\tilde{A}$) tương ứng sẽ có hàng ứng với i gồm chỉ các số 0, vô lý vì ma trận này là cơ sở tức là ma trận có định thức khác 0.

(ii) Ở phần này ta chỉ phải chứng minh nốt “chỉ khi”. Giả sử dòng x không phải là nghiệm cây. Ta phải chứng tỏ x không phải nghiệm cơ sở. Nếu $Ax \neq b$ thì x không phải nghiệm cơ sở. Ta còn phải xét trường hợp $Ax = b$. Đặt $S := \{(i, j) \in \mathcal{A} : x_{ij} \neq 0\}$. Có hai trường hợp.

Trường hợp đầu là S không tạo thành chu trình nào cả. Theo Định lý 3.1.2 sẽ có cây bao trùm T chứa S (và gồm $n-1$ cung). Vì $x_{ij} = 0$ với mọi $(i, j) \notin T$ nên dòng x là một nghiệm cây tương ứng với T , mâu thuẫn.

Trường hợp thứ hai là S có chứa chu trình C . Gọi h^C là hoàn lưu đơn giản ứng với C và xét dòng $x + h^C$. Vì $Ax = b$ và $Ah^C = 0$ nên $A(x + h^C) = b$. Ngoài ra, nếu $x_{ij} = 0$ thì cung $(i, j) \notin C$ nên $h_{ij}^C = 0$. Do đó các ràng buộc thỏa chặt tại x cũng sẽ thỏa chặt tại $x + h^C$. Vì vậy các ràng buộc thỏa chặt tại x không phải có duy nhất nghiệm. Do đó x không phải là đỉnh, nghĩa là không phải nghiệm cơ sở. $\square$



H. 3.8

H.3.8 minh họa trường hợp dòng x không phải nghiệm cây. Trên hình (a) là phần của dòng x thỏa chặt $Ax = b$, x không phải nghiệm cây vì các cung $(2, 1)$, $(3, 1)$

và (3, 2) tạo thành chu trình C và có tải một dòng khác 0. Hình (b) cho thấy dòng $x + h^C$ vẫn thỏa chặt các ràng buộc mà x thỏa chặt.

Cũng như trong trường hợp quy hoạch tuyến tính tổng quát, nghiệm cơ sở chấp nhận được có thể suy biến. Điều này xảy ra khi có cung $(i, j) \in T$ mà $x_{ij} = 0$. Khi đó, cùng một nghiệm cơ sở chấp nhận được có thể tương ứng nhiều cây. Chẳng hạn, cây ở H.3.7 dẫn tới một nghiệm cơ sở chấp nhận được suy biến vì $x_{67} = 0$. Một cây khác cũng ứng với cùng nghiệm này sẽ nhận được khi thay cung (6, 7) bởi cung (5, 7).

3.3.2 Đối cơ sở. Chọn biến ra

Nhớ lại là trong quy hoạch tuyến tính tổng quát khi đối cơ sở trước hết ta chọn một biến không cơ sở để cho vào cơ sở, tức là cho tăng biến này thành dương lớn đến mức tất cả các biến cơ sở cũ vẫn không âm. Biến ra khỏi cơ sở sẽ là biến bị trở thành 0 đầu tiên. Cụ thể hóa cho bài toán dòng trên mạng là như sau. Chọn một biến không cơ sở tức là chọn một cung $(i, j) \notin T$. Cung này cùng với một số cung của T tạo thành một chu trình. Ta chọn hướng của chu trình sao cho (i, j) là cung tiến. Gọi F và B là tập các cung tiến và tập các cung lùi, tương ứng. Khi tăng giá trị của biến không cơ sở x_{ij} từ 0 lên $\theta > 0$. Các biến cơ sở cũ thay đổi để vẫn thỏa ràng buộc (về bảo toàn dòng). Điều này được thực hiện bởi đây θ đơn vị dòng vòng theo chu trình, cũng tức là x_{kl} sẽ tăng (giảm) bởi θ nếu (k, l) là cung tiến (cung lùi, tương ứng). Các biến dòng (tức là dòng trên từng cung) mới là

$$\hat{x}_{kl} = \begin{cases} x_{kl} + \theta & \text{nếu } (k, l) \in F, \\ x_{kl} - \theta & \text{nếu } (k, l) \in B, \\ x_{kl} & \text{trường hợp khác.} \end{cases} \quad (3.5)$$

Rõ ràng θ lớn nhất có thể được sao cho mọi biến dòng vẫn không âm là

$$\theta^* = \min_{(k,l) \in B} x_{kl}, \quad (3.6)$$

trừ trường hợp $B = \emptyset$ thì ta đặt $\theta^* = +\infty$. Biến x_{kl} đạt cực tiểu ở (3.6)

sẽ trở thành $\hat{x}_{kl} = 0$ và là biến ra. Nếu $x_{kl} = 0$ với một $(k, l) \in B$ nào đó, tức là nghiệm cơ sở chấp nhận được hiện hành là suy biến, thì việc thay cơ sở không làm thay các biến dòng. Chẳng hạn ở H.3.7, nếu x_{57} vào thì x_{67} ra khỏi cơ sở và $\theta^* = 0$.

3.3.3 Sự biến đổi của cước phí. Chọn biến vào

Khi thay đổi cơ sở như trên thì lượng cước phí tăng thêm sẽ là

$$\theta^* \left(\sum_{(k,l) \in F} c_{kl} - \sum_{(k,l) \in B} c_{kl} \right). \quad (3.7)$$

Do đó, tất nhiên x_{ij} chỉ được lấy là biến vào khi lượng (3.7) âm.

Ta đã biết từ phương pháp đơn hình cho quy hoạch tuyến tính tổng quát là, nếu biến vào nhận giá trị θ^* , thì lượng thay đổi của cước phí (tức là hàm mục tiêu) là θ^* nhân với hệ số mục tiêu của biến vào. Hệ số này cũng gọi là *cước phí thu gọn* (reduced cost) của biến vào. So sánh với (3.7) thì cước phí thu gọn ở bước hiện hành của biến vào x_{ij} sẽ là

$$\bar{c}_{ij} = \sum_{(k,l) \in F} c_{kl} - \sum_{(k,l) \in B} c_{kl}. \quad (3.8)$$

Ta thấy vế phải (3.8) lại chính là cước phí của chu trình mà dọc theo nó dòng được tải đi.

Việc áp dụng (3.8) để tính các cước phí thu gọn cũng có bất tiện là phải tính đến từng chu trình một. Bây giờ ta xây dựng một công thức có hiệu quả tính toán hơn cho các cước phí thu gọn hiện hành $\bar{c}_{ij}$. Nhớ lại công thức tổng quát $\bar{c}^T = c^T - p^T \bar{A}$ (xem Mục 1.4.1), ở đây p là vectơ đối ngẫu (dual vector) xác định bởi $p^T = c_B^T B^{-1}$ với B là ma trận cơ sở hiện hành. Mỗi biến p_i , $i = 1, \dots, n-1$, ứng với một hàng i của $\bar{A}$, tức là ứng với nút $i \neq n$. Giả sử rằng (i, j) là cung thứ k của đồ thị. Khi đó thành phần thứ k của $\bar{c}$ và c là $\bar{c}_{ij}$ và c_{ij} , tương ứng. Thành phần thứ k của $p^T \bar{A}$ là tích p với cột k của $\bar{A}$. Cột này có thành

phần là 1 ở hàng i (nếu $i < n$) và -1 ở hàng j (nếu $j < n$). Do đó

$$\bar{c}_{ij} = \begin{cases} c_{ij} - (p_i - p_j) & \text{nếu } i, j \neq n, \\ c_{ij} - p_i & \text{nếu } j = n, \\ c_{ij} + p_j & \text{nếu } i = n. \end{cases} \quad (3.9)$$

(3.9) có thể viết gọn là

$$\bar{c}_{ij} = c_{ij} - (p_i - p_j), \quad \forall (i, j) \in \mathcal{A}, \quad (3.10)$$

nếu ta định nghĩa thêm $p_n = 0$. Để tính $\bar{c}$ theo (3.10) cần tính p . Vì cước phí thu gọn hiện hành của biến cơ sở hiện hành bằng 0 nên từ (3.10) ta có

$$p_i - p_j = c_{ij}, \quad \forall (i, j) \in T \quad (3.11)$$

và $p_n = 0$. Để tính p theo hệ (3.11) ta đi từ nút n , gọi là gốc của cây, tiến về phía các lá để xác định dần mọi thành phần p_i .

3.3.4 Phát biểu thuật toán đơn hình mạng

Một bước lặp ứng với nghiệm cơ sở chấp nhận được hiện hành x có cây tương ứng là T gồm các đoạn sau đây.

Đoạn 1. Tính vector đối ngẫu p bằng cách giải hệ phương trình (3.11) từ gốc đến lá.

Đoạn 2. Tính các cước phí thu gọn $\bar{c}_{ij} = c_{ij} - (p_i - p_j)$ của mọi cung $(i, j) \in T$. Nếu chúng đều không âm thì nghiệm cơ sở chấp nhận được hiện hành là tối ưu và thuật toán kết thúc. Ngược lại thì chọn (i, j) nào đó với $\bar{c}_{ij} < 0$ cho vào cơ sở.

Đoạn 3. Cung vào (i, j) và các cung của T chỉ tạo thành một chu trình. Nếu mọi cung của chu trình đều có hướng như (i, j) thì cước phí tối ưu là $-\infty$ và thuật toán kết thúc.

Đoạn 4. Gọi B là tập các cung của chu trình có hướng ngược với (i, j) . Đặt $\theta^* = \min_{(k,l) \in B} x_{kl}$ và tải θ^* đơn vị dòng theo chu trình. Vector dòng mới được xác định bởi (3.5). Đưa ra khỏi cơ sở biến có giá trị mới bằng 0.

Nếu khó xác định nghiệm cơ sở chấp nhận được xuất phát, ta có thể phải lập và xét bài toán bổ trợ. Chẳng hạn, với mỗi cặp gồm hai nút là nguồn và điểm hút ta có thể đưa vào một cung bổ trợ. Với mạng mới này thì việc tìm một nghiệm cơ sở chấp nhận được là tương đối dễ. Hơn nữa nếu chọn cước phí của các cung bổ trợ là đủ lớn thì giải bài toán là tương đương với giải bài toán ban đầu.

Nhớ lại rằng một bước lặp của phương pháp đơn hình tổng quát cần cỡ $O(mn)$ phép tính. Ta xét xem với phương pháp đơn hình mạng thì sao. Để tính p ta cần cỡ $O(m)$ phép tính. Để tính tất cả $\bar{c}_j$ cần $O(n)$ phép tính. Tiếp theo, để thực hiện đổi cơ sở cần $O(n)$ phép tính. Vì $m \geq n - 1$ nên tính chung một bước lặp cần $O(m)$ phép tính, ít hơn hẳn số $O(mn)$. Trong thực tế tính toán, thời gian tính còn được cải thiện do khéo cập nhật vector p và sử dụng cấu trúc dữ liệu thích hợp để tổ chức tính toán.

Nếu không gặp nghiệm cơ sở chấp nhận được, suy biến thì thuật toán kết thúc sau hữu hạn bước. Nếu có hiện tượng suy biến, thuật toán có thể bị xoay vòng. Có thể tránh xoay vòng bằng các quy tắc chống xoay vòng chung cho quy hoạch tuyến tính hoặc các phương pháp riêng cho mạng. Trong thực tế số bước lặp cần thiết chỉ cỡ $O(m)$, nhưng có những thí dụ cần một số mũ các bước lặp.

Sau đây là một thí dụ về áp dụng (3.8) để thực hiện thuật toán đơn hình cho mạng đơn giản, không cần tính p .

THÍ DỤ 3.3.1. Xét bài toán mạng với tải năng không hạn chế ở H.3.9(a), ở đây các số ở cạnh cung là cước phí của cung. Ở hình (b) là một cây bao trùm

và nghiệm cây chấp nhận được tương ứng mà ta dùng để xuất phát. Ở đây và (c), (d) thì số ở cạnh các cung là dòng. Xét cung (4, 3) (ở ngoài cây xuất phát) tạo thành một chu trình với hai cung (3, 5) và (4, 5) của cây xuất phát. Ta xét (3.8) xem cước phí của chu trình có dấu gì :

$$\bar{c}_{43} = c_{43} + c_{35} - c_{45} = 0 + 0 - 1 = -1 < 0.$$

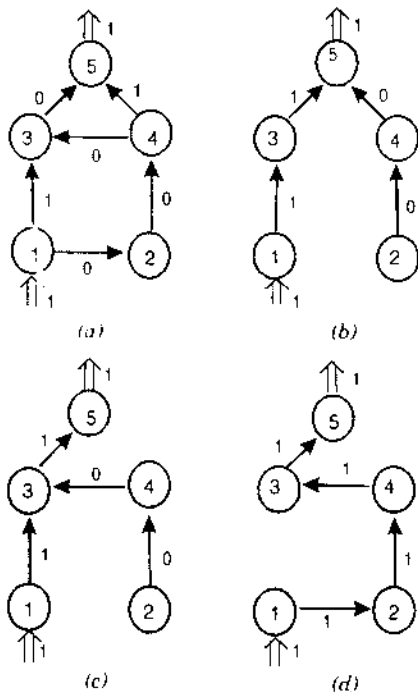
Vì $\bar{c}_{43} < 0$ nên có thể chọn cung (4, 3) vào cây (tức là x_{43} vào cơ sở). Bây giờ xét (3, 5), tức là đẩy θ^* đơn vị dòng theo chu trình để dòng trên cung lùi (4, 5) trở thành 0. Nhưng chính x_{45} đã bằng 0 (nghiệm cơ sở chấp nhận được xuất phát là suy biến) nên $\theta^* = 0$ (so sánh với (3.6)). Do đó, tuy cung (4, 5) ra khỏi cây và ta được cây mới (và tương ứng là nghiệm cơ sở chấp nhận được mới) ở

hình (c), nhưng cước phí không đổi. Chuyển sang bước lập mới, ta lại xét cung (1, 2) (cây ở (c) chỉ thiếu hai cung (1, 2) và (4, 5) so với cả mạng ở hình (a); mà cung (4, 5) vừa được đưa ra ở bước lập trước). Ta lại phải xét dấu $\bar{c}_{12}$:

$$\bar{c}_{12} = c_{12} + c_{24} + c_{43} - c_{13} = 0 + 0 + 0 - 1 = -1 < 0.$$

Do $\bar{c}_{12} < 0$ ta đưa cung (1, 2) vào cây và đẩy θ^* đơn vị dòng theo chu trình 1, 2, 4, 3, 1 để dòng trên cung lùi (1, 3) trở thành 0. Vì thế $\theta^* = 1$ và cung (1, 3) ra khỏi cây. Ta nhận được cây mới ở hình (đ). Để sang bước lập tiếp theo ta lại tính các cước phí thu gọn của các cung ngoài cây ở trạng thái hiện hành. Ta có

$$\bar{c}_{13} = c_{13} - c_{43} - c_{24} - c_{12} = 1 - 0 > 0,$$



H. 3.9

$$\bar{c}_{45} = c_{45} - c_{35} - c_{43} = 1 - 0 > 0.$$

Vậy điều kiện tối ưu đã thỏa (xem đoạn 2 ở thuật toán đơn hình mạng), nên nghiệm ở (d) đã là tối ưu.

3.3.5 Khi nào nghiệm tối ưu có mọi thành phần nguyên?

Một đặc trưng rất quan trọng của bài toán dòng trên mạng là khi dữ liệu của bài toán là số nguyên thì đa số các đại lượng cần quan tâm về bài toán cũng là nguyên. Do đó thuật toán đơn hình có thể được thực hiện chỉ trên các số nguyên. Điều này cho phép tính toán nhanh hơn và nhất là không còn sai số do làm tròn số. Đặc trưng vừa nhắc đến được phát biểu chính xác như sau.

ĐỊNH LÝ 3.3.4. *Xét bài toán dòng trên mạng với tải năng không hạn chế và có đồ thị liên thông.*

(i) *Nghịch đảo B^{-1} của mọi ma trận cơ sở B đều có các phần tử là số nguyên.*

(ii) *Nếu các nguồn b_i đều nguyên thì mọi nghiệm cơ sở đều có mọi thành phần là số nguyên.*

(iii) *Nếu các hệ số cước phí c_{ij} là nguyên, thì mọi nghiệm đối ngẫu cơ sở đều có mọi thành phần là số nguyên.*

CHỨNG MINH. (i) Theo chứng minh Định lý 3.3.1 ta có thể sắp lại các hàng và cột của ma trận cơ sở B để nó có dạng tam giác dưới với các phần tử trên đường chéo là 1 hoặc -1 . Do đó $\det B$ bằng 1 hoặc -1 . Theo quy tắc Cramer, B^{-1} có các phần tử đều là số nguyên.

(ii) Nghiệm cơ sở có dạng $x_T = B^{-1}b$ nên có các thành phần đều là số nguyên.

(iii) Nghiệm đối ngẫu cơ sở là $p^T = c_B^T B^{-1}$, nên có các thành phần đều là số nguyên. $\square$

HỆ QUẢ 3.3.5. Xét bài toán như ở Định lý 3.3.4 với giả thiết thêm là cước phí tối ưu của bài toán là hữu hạn.

(i) Nếu mọi nguồn b_i là nguyên, thì sẽ tồn tại vectơ dòng tối ưu nguyên (tức là có mọi thành phần nguyên).

(ii) Nếu mọi hệ số cước phí c_{ij} là nguyên, thì sẽ tồn tại nghiệm tối ưu nguyên cho bài toán đối ngẫu.

CHỨNG MINH. Vì cước phí tối ưu là hữu hạn, nên thuật toán đơn hình (có sử dụng phương pháp chống xoay vòng) kết thúc sau hữu hạn bước. Sau mỗi bước ta đều nhận được nghiệm cơ sở chấp nhận được cho bài toán gốc và đối ngẫu có các thành phần nguyên, nên nghiệm tối ưu nhận được bằng thuật toán đơn hình cũng có tính chất đó. $\square$

3.4 ĐỐI NGẪU TRONG BÀI TOÁN DÒNG TRÊN MẠNG

Để đơn giản ta chỉ xét bài toán với tải năng không hạn chế và ta vẫn luôn có giả thiết cơ bản ở Mục 3.2.1 về bảo toàn dòng và dòng không âm. Khi đó bài toán dòng trên mạng là một quy hoạch tuyến tính chính tắc

$$\min c^T x,$$

$$Ax = b,$$

$$x \geq 0.$$

Theo quy tắc chung của quy hoạch tuyến tính, bài toán đối ngẫu sẽ là

$$\begin{aligned} \max p^T b, \\ p^T A \leq c^T. \end{aligned} \quad (3.12)$$

Từ dạng của ma trận nối A , các ràng buộc đối ngẫu (3.12) có dạng

đơn giản là

$$p_i - p_j \leq c_{ij}, \quad (i, j) \in \mathcal{A}. \quad (3.13)$$

Giả sử $(p_1, \dots, p_n)$ là nghiệm đối ngẫu chấp nhận được và $\gamma \in \mathbf{R}$. Rõ ràng $(p_1 + \gamma, \dots, p_n + \gamma)$ cũng là nghiệm đối ngẫu chấp nhận được, vì cũng thỏa (3.13). Hơn nữa, xét mục tiêu đối ngẫu ta thấy

$$\sum_{i=1}^n (p_i + \gamma) b_i = \sum_{i=1}^n p_i b_i + \gamma \sum_{i=1}^n b_i = \sum_{i=1}^n p_i b_i,$$

tức là thêm hằng số vào mọi thành phần của vectơ đối ngẫu không ảnh hưởng gì đến tính chấp nhận được và giá trị mục tiêu. Do đó ta luôn có thể coi $p_n = 0$. Điều này hoàn toàn tương đương với việc ta bỏ ràng buộc thứ n ở bài toán gốc, tức là xét $\tilde{A}$ thay cho A .

3.4.1 Ý nghĩa mạng của đối ngẫu

Giả sử ta phải chuyển lượng $b_i > 0$ hàng từ mỗi nút $i = 1, \dots, n-1$ tới nút n theo một mạng của riêng. Khi đó nghiệm tối ưu của bài toán dòng trên mạng cho ta cách vận chuyển tốt nhất.

Lại giả sử có một công ti vận tải, có mở dịch vụ vận chuyển từ mọi nút i đến nút n (theo đường của họ) với giá chuyển một đơn vị hàng là p_i . Nếu (i, j) là một cung trong mạng riêng của ta và phí tổn chuyển một đơn vị hàng trên cung này là c_{ij} , thì ta có thể tự chuyển trên cung này đến j rồi trao cho công ti chuyển nốt đến nút n , với giá đơn vị là $c_{ij} + p_j$. Công ti vận tải biết các vectơ b và c và tìm cách bao hết việc chuyển hàng của ta, kể cả trên cung (i, j) . Khi đó họ phải ra giá để cạnh tranh là $p_i \leq c_{ij} + p_j$ và $p_n = 0$. Với giá đủ hấp dẫn này, coi như ràng buộc, mục đích của họ tất nhiên là làm cực đại doanh

thu $\sum_{i=1}^{n-1} p_i b_i$. Vậy bài toán của công ti chính là bài toán đối ngẫu với

bài toán tự vận tải của ta. Định lý đối ngẫu mạnh của quy hoạch tuyến tính nói rằng doanh thu tối ưu của công ti đúng bằng phí tổn tối ưu khi ta tự vận chuyển bằng mạng riêng. Nói một cách khác nếu hai phía đều tìm cách vận chuyển tối ưu và giá của công ti đặt đúng thì cước phí là như nhau.

Như vậy là ngoài các ý nghĩa trước đây đã trình bày, bằng ngôn ngữ mạng định lý đối ngẫu cũng có ý nghĩa rất rõ ràng.

3.4.2 Đối ngẫu và độ nhạy cảm

Để dễ đối chiếu với lý thuyết chung của quy hoạch tuyến tính, ta bỏ ràng buộc thứ n (ứng với nút n) và giả sử các ràng buộc còn lại $\bar{A}x = \bar{b}$ là độc lập tuyến tính. Giả sử x là nghiệm cơ sở chấp nhận được, tối ưu và không suy biến (tương ứng với một cây) và p là nghiệm đối ngẫu tối ưu. Xét một nút nào đó $i \neq n$ và giả sử nguồn b_i bị nhiễu thành $b_i + \epsilon$, với $\epsilon > 0$ và nhỏ, còn các nguồn khác cho đến b_{n-1} không đổi.

Khi đó điều kiện $\sum_{i=1}^n b_i = 0$ dẫn đến b_n phải trở thành $b_n - \epsilon$. Nhưng

thực ra điều này chỉ ảnh hưởng đến ràng buộc thứ n đã bỏ đi. Giả sử ϵ đủ nhỏ để cơ sở cũ vẫn giữ nguyên là cơ sở, tức là cây không đổi. Khi đó để đảm bảo ràng buộc về bảo toàn dòng, lượng nguồn tăng là ϵ bắt buộc phải chuyển từ nút i đến nút n dọc theo đường duy nhất xác định bởi cây đã có. Giả sử ϵ đủ bé để nghiệm tối ưu bị nhiễu vẫn là tối ưu. Vì giá trị mục tiêu tối ưu của hai bài toán đối ngẫu nhau là bằng nhau, nên lượng thay đổi cũng bằng nhau (vì hai nghiệm bị nhiễu

vẫn tối ưu). Mục tiêu của bài toán đối ngẫu là $\sum_{i=1}^n p_i b_i$, nên lượng thay

đổi là ϵp_i . Vậy ý nghĩa của biến đối ngẫu p_i là: ở nghiệm tối ưu, p_i là tốc độ thay đổi của giá trị mục tiêu tối ưu theo sự thay đổi của vế phải ràng buộc thứ i là b_i . (Thực ra có thể chứng minh được điều này cho quy hoạch tuyến tính tổng quát, không phải chỉ đúng với bài toán

mạng.)

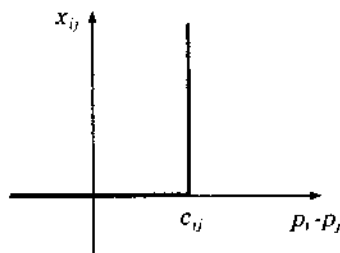
Bằng lý luận tương tự ta sẽ thấy nếu tăng b_i thêm ε giảm b_j đi ε và giữ nguyên các nguồn khác và giả sử cơ sở không đổi và nghiệm vẫn là tối ưu thì mục tiêu tối ưu bị thay đổi là $\varepsilon(p_i - p_j)$ (vẫn giả sử không có suy biến). Điều này nghĩa là $p_i - p_j$ là cước phí tối ưu để chuyển một đơn vị hàng từ nút i đến nút j .

3.4.3 Độ lệch bù

Từ điều kiện độ lệch bù của quy hoạch tuyến tính tổng quát ta suy ra ngay điều kiện độ lệch bù cho bài toán dòng trên mạng:

(a) Nếu $p_i \neq 0$ thì $(Ax)_j = b_j$. Vì bài toán chỉ gồm ràng buộc đẳng thức nên điều này tự động thỏa với mọi dòng chấp nhận được x .

(b) Nếu $x_{ij} > 0$ thì $p_i - p_j = c_{ij}$, tức là ràng buộc đối ngẫu $p_i - p_j \leq c_{ij}$ phải thỏa chặt. “Ý nghĩa vận tải” của điều kiện này là nếu $p_i - p_j < c_{ij}$ (thì $x_{ij} = 0$), thì có cách chuyển hàng từ i tới j rẻ hơn là dùng cung (i, j) do đó dòng trên cung này phải bằng 0 (nếu thỏa điều kiện độ lệch bù). Điều này phù hợp với lý luận ở cuối Mục 3.4.2.



H. 3.10

Từ điều kiện tối ưu cho quy hoạch tuyến tính tổng quát ta đã có: x là nghiệm tối ưu của bài toán dòng (gốc) và p là nghiệm tối ưu của bài toán đối ngẫu khi và chỉ khi x là nghiệm gốc chấp nhận được, p là nghiệm đối ngẫu chấp nhận được và điều kiện độ lệch bù (a), (b) thỏa.

Hình đơn giản H.3.10 thể hiện đầy đủ ràng buộc dấu $x_{ij} \geq 0$, ràng buộc đối ngẫu $p_i - p_j \leq c_{ij}$ và điều kiện (b) về độ lệch bù: với mọi cung (i, j) , cặp số $(p_i - p_j, x_{ij})$ phải nằm trên

đường tô đậm.

Vậy điều kiện tối ưu có thể phát biểu như sau.

ĐỊNH LÝ 3.4.1. *Đối với bài toán dòng trên mạng với tải năng không hạn chế, hai khẳng định sau là tương đương.*

(i) *Các vectơ x và p là nghiệm tối ưu của bài toán gốc và bài toán đối ngẫu, tương ứng.*

(ii) *x thỏa ràng buộc bảo toàn dòng $Ax = b$ và với mọi cung (i, j) , cặp số $(p_i - p_j, x_{ij})$ thỏa quan hệ trên H.3.3.*

3.5 THUẬT TOÁN CHU TRÌNH CƯỚC PHÍ ÂM

Thuật toán đơn hình mạng (gốc) trình bày ở Mục 3.3 dựa vào một ý tưởng cơ bản, là cơ sở của hầu như tất cả các thuật toán gốc khác cho bài toán mạng là: khi đã có một nghiệm gốc chấp nhận được ta tìm một chu trình có cước phí âm và đẩy một dòng theo nó để nhận được nghiệm chấp nhận được cải thiện hơn (tức là ứng với cước phí trên toàn bộ dòng nhỏ hơn).

Ưu điểm của thuật toán đơn hình mạng gốc là cách tìm chu trình cước phí âm rất đơn giản (chọn một cung ngoài cây bao trùm và xét chu trình tạo thành do nó và các cung của cây). Nhược điểm lớn là khi chọn được chu trình có cước phí âm rồi, việc đổi cơ sở vẫn có thể bị suy biến. Đó là khi nghiệm cơ sở ứng với cây bao trùm là suy biến nên $0 \cdot \text{tính theo (3.6) bằng } 0$. Ta không đẩy được dòng nào trên chu trình và sự thay đổi cước phí trên toàn dòng bằng 0 (theo (3.7)). Vậy ta vẫn có nghiệm chấp nhận được mới nhưng không cải thiện gì thêm về hàm mục tiêu.

Trong Mục 3.5 này ta sẽ trình bày thuật toán chu trình cước phí âm. Đây cũng là thuật toán gốc và dựa trên chu trình cước phí âm, nhưng không phải thuật toán đơn hình, vì nó không dựa trên nghiệm cây (tức là nghiệm cơ sở). Ta sẽ thấy chu trình cước phí âm ở đây là bất kì, không cần tạo từ một cung ngoài với các cung của cây bao trùm (ở các bước lặp ta cũng không có cây bao trùm). Thuật toán sẽ kết thúc khi không có chu trình cước phí âm nữa. Còn ở mỗi bước lặp đều có sự cải thiện hàm mục tiêu.

3.5.1 Mô tả thuật toán

Xét bài toán dòng trên mạng với tải năng hạn chế

$$0 \leq x_{ij} \leq u_{ij}, (i, j) \in A.$$

Giả sử C là một chu trình nào đó, với F và B là tập các cung tiến và tập các cung lùi của C , tương ứng. Gọi h^C là hoàn lưu đơn giản tương ứng với C , tức là

$$h_{ij}^C = \begin{cases} 1 & \text{nếu } (i, j) \in F, \\ -1 & \text{nếu } (i, j) \in B, \\ 0 & \text{nếu } (i, j) \notin C. \end{cases}$$

Giả sử x là một vectơ dòng chấp nhận được. Đặt $\theta \geq 0$ đơn vị dòng theo chu trình C ta được vectơ dòng mới $x + \theta h^C$ vẫn thỏa ràng buộc bảo toàn dòng $A(x + \theta h^C) = b$ vì $Ah^C = 0$ (do h^C là hoàn lưu). Để dòng mới vẫn chấp nhận được, tức là $0 \leq x_{ij} + \theta h_{ij}^C \leq u_{ij}$, hay viết rõ là

$$0 \leq x_{ij} + \theta \leq u_{ij} \quad \text{nếu } (i, j) \in F,$$

$$0 \leq x_{ij} - \theta \leq u_{ij} \quad \text{nếu } (i, j) \in B,$$

thì cần và đủ là $\theta \leq u_{ij} - x_{ij}$ nếu $(i, j) \in F$, $\theta \leq x_{ij}$ nếu $(i, j) \in B$. Vậy lượng dòng cực đại có thể đẩy theo chu trình C , ký hiệu là $\theta(C)$, sẽ là

$$\theta(C) = \min \left\{ \min_{(i, j) \in B} x_{ij}, \min_{(i, j) \in F} (u_{ij} - x_{ij}) \right\}. \quad (3.14)$$

Nếu $B = \emptyset$ và $u_{ij} = \infty$ cho mọi cung của chu trình thì không có hạn chế nào đối với θ nên ta đặt $\theta(C) = +\infty$.

Chu trình C có $\theta(C) > 0$, tức là $x_{ij} < u_{ij}$ với mọi $(i, j) \in F$ và $x_{ij} > 0$ với mọi $(i, j) \in B$, được gọi là *chu trình chưa bão hòa* (unsaturated cycle). Nhớ lại là cước phí của chu trình C

$$c^T h^C = \sum_{(i,j) \in F} c_{ij} - \sum_{(i,j) \in B} c_{ij},$$

chính là sự thay đổi cước phí của toàn bộ dòng khi ta đẩy một đơn vị dòng theo chu trình C (điều này cũng dễ thấy trực tiếp). Bây giờ ta có thể mô tả thuật toán.

THUẬT TOÁN CHU TRÌNH CƯỚC PHÍ ÂM

Ở mỗi bước lặp, xuất phát từ dòng chấp nhận được x (không cần là nghiệm cơ sở) ta làm các đoạn sau đây.

1. Tìm chu trình chưa bão hòa có cước phí âm. Nếu không có, thì thuật toán kết thúc.

2. Nếu có chu trình C như vậy, thì xét hai trường hợp :

- (a) nếu $\theta(C) < \infty$, thì ta lấy dòng chấp nhận được mới $x + \theta(C)h^C$ và quay lại đoạn 1 (của bước lặp mới);
- (b) nếu $\theta(C) = +\infty$, thì thuật toán kết thúc và cước phí tối ưu của bài toán là $-\infty$.

Xét cụ thể cách tiến hành và chứng minh thuật toán là đúng ta phải làm rõ các điểm sau đây.

- (i) Làm sao có dòng chấp nhận được ban đầu để xuất phát thuật toán?
- (ii) Tìm chu trình chưa bão hòa có cước phí âm thế nào?
- (iii) Có đúng là khi thuật toán kết thúc ta nhận được nghiệm tối ưu không?

(iv) Có chắc là thuật toán kết thúc sau hữu hạn bước?

Bốn mục nhỏ tiếp theo sẽ trả lời bốn câu hỏi này.

3.5.2 Xuất phát thuật toán

Theo Mục 3.2.2 mọi bài toán dòng trên mạng có thể biến đổi thành bài toán tương đương không có nguồn và điểm hút. Đối với bài toán này thì dòng 0 là chấp nhận được và có thể dùng để xuất phát thuật toán.

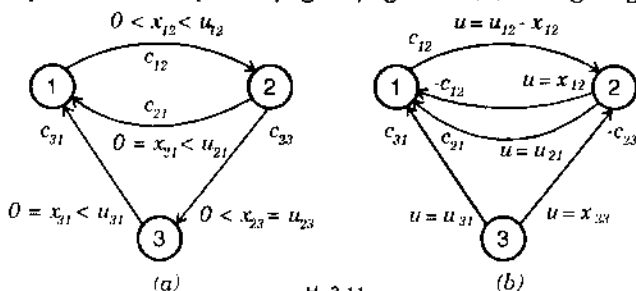
Ngoài ra, ở đầu Mục 3.6 dưới đây sẽ chứng minh là có thể tìm được một dòng chấp nhận được bằng cách giải một bài toán dòng cực đại thích hợp. Khi đó ta có dòng (nói chung khác 0) để xuất phát.

3.5.3 Mạng thặng dư

Giả sử ta có một mạng $G = (A, A)$ và một dòng chấp nhận được x . *Mạng thặng dư* (residual network) định nghĩa cụ thể dưới đây sẽ là một mạng $G = (A, A)$ có cùng các nút, nhưng các cung khác và tải năng khác, liên quan đến phạm vi của dòng có thể đẩy thêm vào các cung của mạng ban đầu. Xét cung (i, j) với tải năng u_{ij} và dòng chấp nhận được hiện hành x_{ij} . Khi đó x_{ij} có thể tăng thêm tới $u_{ij} - x_{ij}$ và giảm đi lượng tới mức x_{ij} . Tương ứng, ta xây dựng mạng thặng dư (ứng với mạng G và dòng chấp nhận được x) có cung (i, j) với tải năng $u_{ij} - x_{ij}$ và có cung (j, i) với tải năng x_{ij} . Như vậy dòng trên cung (j, i) của mạng thặng dư sẽ có vai trò là lượng giảm dòng trên cung (i, j) của mạng ban đầu.

Các yếu tố khác của mạng thặng dư được xác định như sau. Cước phí của các cung sẽ phản ánh sự thay đổi cước phí ở mạng ban đầu. Cụ thể là ta đặt cho mạng thặng dư cước phí của cung (i, j) là c_{ij} và của cung (j, i) là $-c_{ij}$ (vì một đơn vị dòng trên cung (j, i) là một đơn vị giảm dòng trên (i, j) của mạng ban đầu, ứng với sự thay đổi cước phí $-c_{ij}$). Nguồn trên mạng thặng dư ta lấy bằng 0 cho mọi nút. Khi đó mọi dòng chấp nhận được đều là hoàn lưu. Cuối cùng ta bỏ đi mọi cung có tải năng bằng không trong mạng thặng dư.

H.3.11 minh họa cách xác định mạng thặng dư ở (b) tương ứng với mạng và một dòng chấp nhận được đã cho ở (a). Ta thấy mạng thặng dư có thể có hai cung có cùng các nút đuôi và đầu. Điều này mâu thuẫn với định nghĩa chung về đồ thị nhưng không gây phiền toái gì. (Ở (b) ta kí hiệu u cho tải năng của mọi cung.)



H. 3.11

Bây giờ giả sử dòng chấp nhận được x đã có (tương ứng với mạng thặng dư) một vectơ giá $\bar{x}$ sao cho $x + \bar{x}$ là dòng chấp nhận được của mạng ban đầu. Ta sẽ định nghĩa vectơ dòng $\bar{x}$ (của mạng thặng dư) tương ứng với vectơ dòng giá tăng $\bar{x}$ (từ đó có tên “thặng dư”) như sau.

- Nếu $\bar{x}_{ij} > 0$, ta đặt dòng $\bar{x}_{ij}$ của cung (i, j) (trong mạng thặng dư) bằng $\bar{x}_{ij}$. Khi đó $\bar{x}_{ij} = \bar{x}_{ij} \leq u_{ij} - x_{ij}$ tức là $\bar{x}_{ij}$ thỏa ràng buộc tải năng của mạng thặng dư.

- Nếu $\bar{x}_{ij} < 0$ ta đặt dòng $\bar{x}_{ji}$ của cung (j, i) (trong mạng thặng dư) bằng $-\bar{x}_{ij}$. Khi đó $\bar{x}_{ji} = -\bar{x}_{ij} \leq x_{ij}$ (do tính chấp nhận được của $x + \bar{x}$) nên $\bar{x}_{ji}$ thỏa hạn chế tải năng ở mạng thặng dư.

- Tất cả các dòng $\bar{x}_{ij}$ khác chưa được xác định bằng hai cách trên đều được đặt bằng 0.

Ta có các nhận xét sau đây.

(a) $\bar{x}_{ij} \geq 0$ ở mọi cung của mạng thặng dư.

(b) $\bar{x}$ là một hoàn lưu. Thật vậy, ta có $A\bar{x} = 0$ tức là với dòng $\bar{x}$ thì dòng tình (net flow) vào thực sự ở mỗi nút bằng 0. Do đó, theo cách xây dựng $\bar{x}$, dòng tình vào mỗi nút của mạng thặng dư cũng bằng 0.

Dòng của mạng thặng dư thỏa (a) và (b) sẽ được gọi là chấp nhận được nếu nó thỏa mọi hạn chế tải năng.

(c) Cước phí của dòng $\bar{x}$ trong mạng thặng dư bằng $\sum_{(i,j)} c_{ij} \bar{x}_{ij}$ tức là bằng cước phí của dòng gia tăng $\bar{x}$ trong mạng ban đầu.

Như vậy từ dòng gia tăng $\bar{x}$ ta có dòng $\bar{x}$ của mạng thặng dư có các tính chất (a), (b) và (c). Ngược lại, từ một dòng chấp nhận được $\bar{x}$ của mạng thặng dư ta cũng xây dựng được dòng gia tăng $\bar{x}$ của mạng ban đầu sao cho $x + \bar{x}$ chấp nhận được và $c^T \bar{x}$ bằng cước phí của $\bar{x}$ trong mạng thặng dư.

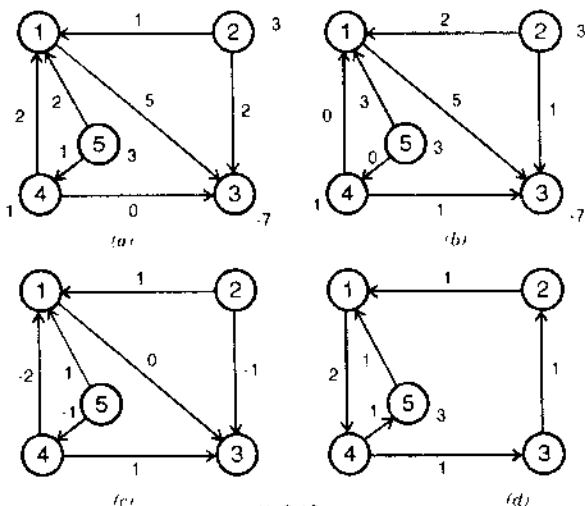
Nhận xét rằng mỗi cung (i, j) trong mạng ban đầu, có tải năng u_{ij} và cước phí c_{ij} đều dẫn tới hai cung trong mạng thặng dư: cung (i, j) với tải năng $u_{ij} - x_{ij}$, cước phí c_{ij} và cung (j, i) với tải năng x_{ij} và cước phí $-c_{ij}$. Nhưng có thể một cung trong đó có tải năng bằng 0 và bị bỏ đi.

Mỗi chu trình chưa bão hòa trong mạng đều có tương ứng một chu trình có hướng trong mạng thặng dư mà mọi cung có tải năng dương. Thật vậy, nếu (i, j) là cung tiến thì do chu trình chưa bão hòa nên $u_{ij} - x_{ij} > 0$ nên (i, j) là cung có tải năng dương trong chu trình tương ứng ở mạng thặng dư. Nếu (i, j) là cung lùi, thì chu trình chưa bão hòa có nghĩa là $x_{ij} > 0$, nên trong mạng thặng dư có cung (j, i) với tải năng x_{ij} . Cung (j, i) sẽ là cung tiến trong chu trình ở mạng thặng dư. Vậy ta có chu trình chỉ gồm cung tiến, tức là chu trình có hướng mà tải năng mọi cung là dương.

Dễ thấy rằng điều ngược lại cũng đúng: mỗi chu trình có hướng (với tải năng dương ở mỗi cung) trong mạng thặng dư đều có ứng với một chu trình chưa bão hòa trong mạng ban đầu. Hơn nữa, cước phí của hai chu trình này trong mạng tương ứng là bằng nhau. Do đó, việc tìm chu trình chưa bão hòa cước phí âm có thể được thực hiện bằng cách tìm chu trình định hướng cước phí âm ở mạng thặng dư.

Trên H. 3.12 là thí dụ về xây dựng dòng chấp nhận được $\bar{x}$ ở mạng

thặng dư (trên hình (d); chỉ vẽ các cung với dòng khác không) tương ứng với sự tăng dòng x (hình (c)) của mạng ban đầu và dòng x chấp nhận được (hình (a)). Tất nhiên dòng $x + \bar{x}$ phải chấp nhận được (biểu diễn ở hình (b)). Các số ở cạnh cung là các dòng. Các số ở cạnh nút là các nguồn.



H. 3.12

3.5.4 Điều kiện tối ưu

Bây giờ ta xét xem điều gì xảy ra khi thuật toán kết thúc. Nếu thuật toán kết thúc vì có chu trình cước phí âm với $\theta(c) = +\infty$ và x là dòng chấp nhận được hiện hành, thì $x + \theta h^c$ là chấp nhận được với mọi $\theta > 0$. Cho θ lớn tùy ý ta thấy cước phí của nghiệm chấp nhận được dạng này là không bị chặn dưới.

Nếu thuật toán kết thúc vì không có chu trình chưa bão hòa cước phí âm thì Định lý 3.5.2 dưới đây nói rằng ta đã có nghiệm tối ưu. Để chứng minh định lý này ta cần bổ đề sau đây.

BỔ ĐỀ 3.5.1 (Định lý phân tích dòng). Giả sử $x \geq 0$ là một hoàn lưu khác 0. Khi đó sẽ có các hoàn lưu đơn giản và có hướng $x^1, \dots, x^k$ và các số dương $\alpha_1, \dots, \alpha_k$ sao cho

$$x = \sum_{i=1}^k \alpha_i x^i.$$

Hơn nữa, nếu x là một vectơ nguyên thì có thể lấy mọi α_i đều nguyên.

CHỨNG MINH. Vì x khác không, sẽ có cung (i, j) với $x_{ij} > 0$. Khi đi hết cung (i, j) , do sự bảo toàn dòng tại nút j , phải có cung (j, l) nào đó với $x_{jl} > 0$ để ta đi tiếp. Làm tiếp quá trình này sẽ phải quay lại một nút đã qua, vì số nút là hữu hạn. Lúc này ta có một chu trình có hướng mang một dòng dương. Giả sử x^1 là hoàn lưu đơn giản tương ứng với chu trình này và α_1 là số nhỏ nhất trong các dòng x_{ij} trên các cung của chu trình. Khi đó vectơ $\hat{x} := x - \alpha_1 x^1$ có mọi thành phần không âm. Hơn nữa $A\hat{x} = Ax - \alpha_1 Ax^1 = 0 - 0$, nên $\hat{x}$ là hoàn lưu. Đồng thời, do định nghĩa của α_1 số thành phần dương của $\hat{x}$ phải ít hơn của x . Ta làm lại với $\hat{x}$ như với x để được hoàn lưu đơn giản có hướng x^2 và $\alpha_2 > 0$, rồi tiếp tục quá trình. Qua mỗi bước, số cung mang dòng dương giảm đi ít nhất một. Do đó sau hữu hạn bước ta được dòng 0. Lúc này ta đã được phân tích $x = \sum \alpha_i x^i$ mà mỗi x^i đều là hoàn lưu đơn giản có hướng (tức là chỉ gồm các cung tiến).

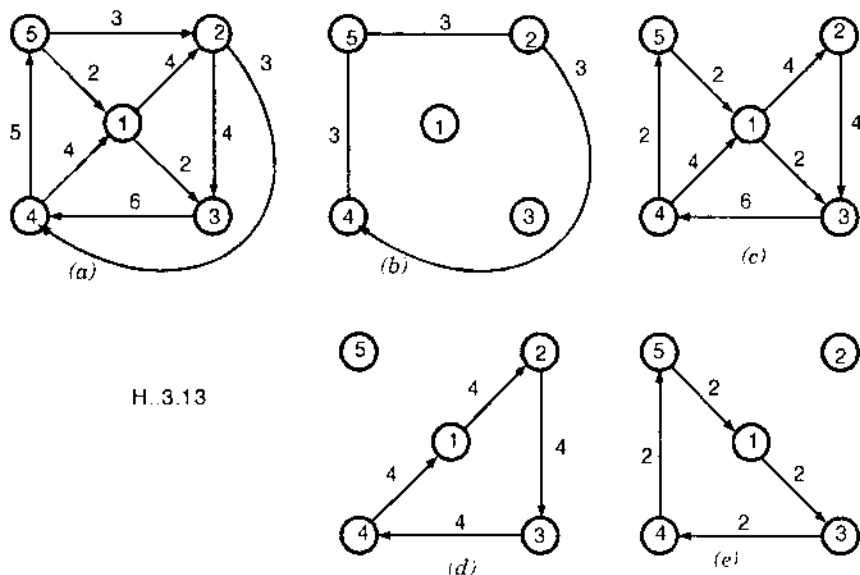
Nếu x là nguyên, thì α_i là số nguyên và $\hat{x}$ là vectơ nguyên. Do đó, theo quy nạp, nếu ta bắt đầu bởi một vectơ dòng nguyên x thì mọi dòng sinh ra trong quá trình trên đều nguyên và mọi α_i là số nguyên. $\square$

ĐỊNH LÝ 3.5.2. *Dòng chấp nhận được x là tối ưu khi và chỉ khi ứng với nó không có chu trình chưa bão hòa có cước phí âm.*

CHỨNG MINH. "*Chỉ khi*". Giả sử có chu trình chưa bão hòa cước phí âm C . Khi đó $x + \theta(c)h^C$ là dòng chấp nhận được có cước phí nhỏ hơn cước phí của x , tức là x chưa tối ưu.

"*Khi*". Giả sử x là dòng chấp nhận được chưa tối ưu. Khi đó có dòng chấp nhận được $\bar{x}$ có cước phí nhỏ hơn, tức là $c^T \bar{x} < 0$. Theo Mục 3.5.3 sẽ có tương ứng trên mạng thặng dư hoàn lưu $\bar{x} \geq 0$, chấp nhận được và có cước phí âm ($= c^T \bar{x}$). Theo Bổ đề 3.5.1 ta có phân tích $\bar{x} = \sum \alpha_i x^i$, với $\alpha_i > 0$ và $\bar{x}^i$ là các hoàn lưu đơn giản có hướng. Vì $\bar{x}$ có cước phí âm, ít nhất một hoàn lưu $\bar{x}^i$ có cước phí âm. Do đó mạng thặng dư chứa chu trình có hướng với cước phí âm. Theo lý luận ở cuối Mục 3.5.3 mạng ban đầu có chu trình chưa bão hòa với cước phí âm. $\square$

H.3.13 minh họa sự phân tích dòng: (a) Một hoàn lưu không âm x (các cung có dòng bằng 0 không cần vẽ); (b) hoàn lưu $\alpha_1 x^1$; (c) dòng còn lại $\hat{x} = x - \alpha_1 x^1$; (d) hoàn lưu $\alpha_2 x^2$; (e) dòng còn lại $\hat{x} = x - \alpha_1 x^1 - \alpha_2 x^2$ là hoàn lưu đơn giản nên ta đặt luôn là $\alpha_3 x^3$. Như vậy ta được phân tích $x = \alpha_1 x^1 + \alpha_2 x^2 + \alpha_3 x^3$.



H.3.13

3.5.5 Sự kết thúc hữu hạn của thuật toán

Để chứng tỏ thuật toán là đúng, tức là thực sự sử dụng được, ta còn phải có tiêu chuẩn đảm bảo nó kết thúc được. Đó cũng là nội dung của định lý sau.

ĐỊNH LÝ 3.5.3. Giả sử mọi tải năng u_{ij} đều nguyên hoặc bằng $+\infty$ và thuật toán chu trình cước phí âm được xuất phát bằng một dòng chấp nhận được nguyên. Khi đó

(i) Các biến dòng trên cung vẫn là nguyên ở mọi bước lặp;

(c) Khi cước phí tối ưu là $-\infty$, thuật toán có thể kết thúc không hữu hạn, ngay cả khi tải năng của các cung là nguyên. Vì vậy ta phải kiểm tra xem cước phí tối ưu có hữu hạn không trước khi bắt đầu thuật toán.

3.6 BÀI TOÁN DÒNG CỰC ĐẠI

3.6.1 Đặt bài toán. Sự tồn tại nghiệm

Một trong các bài toán dòng trên mạng quan trọng nhất phát biểu như sau.

BÀI TOÁN DÒNG CỰC ĐẠI. Cho một đồ thị có hướng $G = (A, \mathcal{A})$ có tải năng các cung là u_{ij} , $(i, j) \in \mathcal{A}$, có thể bằng $+\infty$. Giả sử s và t là hai nút đặc biệt, gọi tương ứng là nguồn và điểm hút. Bài toán đặt ra là tìm dòng lớn nhất có thể chuyển qua mạng từ s tới t .

Phát biểu toán học của bài toán là

$$\begin{aligned} \max \quad & b_s, \\ \text{Ax} = & b, \\ b_t = & -b_s, \\ b_i = & 0, \quad \forall i \neq s, t, \\ 0 \leq x_{ij} \leq & u_{ij} \quad \forall (i, j) \in \mathcal{A}. \end{aligned}$$

Ta thấy mục tiêu ở đây khác với bài toán dòng trên mạng. Giá trị của b_s sẽ gọi là giá trị của dòng chấp nhận được tương ứng.

Bài toán dòng cực đại có thể được đưa về bài toán dòng trên mạng đã xét trên đây bằng cách sau. Trước hết, ta phải đưa vào các c_{ij} . Ta đặt $c_{ij} = 0$, $\forall (i, j) \in \mathcal{A}$. Ta vẽ thêm cung mới (t, s) với $u_{ts} = +\infty$ và $c_{ts} = -1$. Tìm cực tiểu hàm mục tiêu $\sum_{(i,j)} c_{ij}x_{ij} = -x_{ts}$ ở bài toán dòng trên

mạng mới được đặt, do đó, chính là tìm dòng chấp nhận được sao cho x_{ts} là lớn nhất. Nhưng vì bài toán mới đặt không có nguồn và điểm hút nên toàn bộ dòng này phải đi qua mạng để từ s quay lại t . Vậy x_{ts} chính là giá trị nghiệm của bài toán dòng cực đại ban đầu.

Ngược lại, việc tìm dòng chấp nhận được của bài toán dòng trên mạng có tải năng hạn chế có thể đưa về giải bài toán dòng cực đại như sau. Trước hết ta bỏ các hệ số c_{ij} đi và đưa thêm vào nút nguồn s và nút hút t . Ta vẽ thêm cung (s, j) tới mỗi nút j có $b_j > 0$ và đặt tải năng của nó là $u_{sj} = b_j$ và cung (i, t) từ mỗi nút i có $b_i < 0$ và đặt $u_{it} = -b_i$. Vì $\sum b_k = 0$ ta có $\sum u_{sj} = \sum u_{it} := w$. Bây giờ ta đưa vào khái niệm cắt, dùng đến nhiều cho bài toán dòng cực đại.

Một tập con S của tập các nút $\mathcal{A}$ được gọi là một *cắt* (cut) hoặc cụ thể hơn, s.t cắt của bài toán dòng cực đại nếu $s \in S$ và $t \notin S$. *Dung lượng* (capacity) $C(S)$ của cắt S là tổng tải năng của các cung đi từ S ra ngoài phần bù của S , tức là các cung có đuôi thuộc S , đầu không thuộc S . Vậy

$$C(S) = \sum_{(i,j) \in A: i \in S, j \notin S} u_{ij}.$$

Mỗi dòng đi từ s đến t có thể tách thành hai phần là tổng các dòng trên các cung ra khỏi S trừ đi tổng các dòng trên các cung quay ngược lại S . Tức là mỗi vectơ dòng x có một giá trị v (chính là dòng từ s đến t):

$$v = \sum_{j \in S, k \notin S} x_{jk} - \sum_{i \in S, j \in S} x_{ij}. \quad (3.15)$$

Nói riêng, nếu ta lấy $S = \{s\}$ thì $v = \sum_{k \neq s} x_{sk}$ hoặc $S = \mathcal{A} \setminus \{t\}$ thì

$v = \sum_{j \neq t} x_{jt}$, vì ta sẽ luôn giả thiết là trong mạng không có cung dạng

(i, s), tức là cung sẽ đi đến nguồn, và cũng không có cung dạng (t, j), tức là cung ra khỏi điểm hút. Từ (3.15) ta thấy ngay là giá trị của mọi dòng chấp nhận được (từ s đến t) phải thỏa

$$v \leq C(S), \quad (3.16)$$

với mọi cắt S. Vậy dung lượng của bất kì cắt nào cũng là một "cổ chai" mà dòng cực đại không thể vượt quá được.

Bây giờ quay lại bài toán dòng cực đại ta vừa lập từ bài toán dòng trên mạng. Rõ ràng giá trị $w = \sum u_{sj} = \sum u_{it}$ là dung lượng của cắt gồm chỉ có nút s (và cũng là dung lượng của cắt gồm mọi nút, chỉ trừ nút t). Do đó, theo (3.16) mọi dòng chấp nhận được của bài toán dòng cực đại đang xét đều có giá trị không quá w. Cụ thể hơn là đối với dòng x chấp nhận được của bài toán dòng cực đại này thì ba khẳng định sau là tương đương

(i) giá trị của x là w ;

(ii) $x_{sj} = u_{sj}$ với mọi cung (s,j);

(iii) $x_{it} = u_{it}$ với mọi cung (i,t).

Để thấy dòng chấp nhận được của bài toán dòng cực đại (mà có giá trị w) luôn là dòng chấp nhận được của bài toán dòng trên mạng ban đầu (vì ở các điểm nguồn và điểm hút luật bảo toàn dòng được thỏa do (ii) và (iii); còn ở các nút khác thì sự bảo toàn dòng ở hai bài toán là như nhau). Ngược lại, mỗi dòng chấp nhận được của bài toán dòng trên mạng phải thỏa (ii) theo sự bảo toàn dòng ở các điểm nguồn (vì $u_{sj} = b_{sj}$), nên phải là dòng chấp nhận được với giá trị w của bài toán dòng cực đại tương ứng (theo cách đặt của ta trên đây).

Vậy có thể tìm dòng chấp nhận được để xuất phát thuật toán chu trình cực phí âm bằng cách giải một bài toán dòng cực đại xem có đúng dòng cực đại có giá trị w không, như đã nói trước ở Mục 3.5.2. (Chú ý là theo (3.16) dòng chấp nhận được của bài toán dòng cực đại

đang xét mà có giá trị w phải là dòng cực đại.) Cụ thể là : bài toán dòng trên mạng có tải năng hạn chế có nghiệm chấp nhận được khi và chỉ khi bài toán dòng cực đại tương ứng (lập như trên) có dòng cực đại với giá trị w . Định lí sau đây khẳng định một nhận xét trực quan là dòng cực đại trong bài toán dòng cực đại bất kì chính bằng "cổ chai" hẹp nhất. Nó được công bố độc lập nhau bởi [Fulkerson - Dantzig 1955], [Ford - Fulkerson 1956a] và [Elias - Feinstein - Shannon 1956].

ĐỊNH LÍ 3.6.1 (Định lí dòng cực đại cắt cực tiểu). Đối với bài toán dòng cực đại đúng một trong hai trường hợp sau sẽ xảy ra:

- (i) có dòng chấp nhận được với giá trị lớn tùy ý và mọi cắt đều có dung lượng vô hạn;
- (ii) tồn tại dòng cực đại và giá trị của nó là cực tiểu của dung lượng của tất cả các cắt.

CHỨNG MINH. (i) Ta đưa bài toán dòng cực đại về bài toán dòng trên mạng tương ứng theo cách trên đây. Đối với bài toán dòng trên mạng bất kì sẽ có ba khả năng: (a) giá trị mục tiêu có thể lớn tùy ý, do đó có dòng chấp nhận được với giá trị lớn tùy ý; (b) có nghiệm tối ưu; (c) không có dòng chấp nhận được. Với bài toán dòng trên mạng tương ứng bài toán dòng cực đại thì luôn có dòng chấp nhận được là $x_{ij} = 0, \forall (i, j) \in \mathcal{A}$. Vì vậy chỉ còn hai khả năng (a) và (b). Nếu xảy ra (a) thì mọi cắt đều có dung lượng vô hạn theo (3.16).

(ii) Nếu xảy ra (b) thì nghiệm tối ưu phải thỏa điều kiện độ lệch bù

$$x_{ij} = 0 \quad \text{mỗi khi} \quad p_j + c_{ij} > p_i, \quad (3.17)$$

$$x_{ij} = u_{ij} \quad \text{mỗi khi} \quad p_j + c_{ij} < p_i. \quad (3.18)$$

(Đây là mở rộng kết luận ở Mục 3.4.3 khi bài toán có tải năng hạn chế, độc giả có thể thừa nhận hoặc tự chứng minh.) Nói riêng, vì $u_{is} = +\infty$ và $c_{is} = -1$ ta phải có $p_s - 1 > p_i$. Do đó, tập S tất cả các nút k có tính chất $p_k \geq p_s$ phải là một cắt (vì nó không chứa t). Vì $c_{ij} = 0$ cho mọi cung (i, j) , chỉ trừ cung (t, s) , nên

(3.17) suy ra $x_{ij} = 0$ nếu $j \in S$ và $i \notin S$ và (3.18) suy ra $x_{ij} = u_{ij}$ nếu $j \notin S$, $i \in S$. Công thức (3.15) do đó sẽ suy ra là giá trị của dòng x tối ưu của bài toán dòng trên mạng sẽ là dung lượng của S . Vì vậy x là dòng cực đại theo (3.16). Cũng theo công thức này cắt S có dung lượng nhỏ nhất trong tất cả các cắt. $\square$

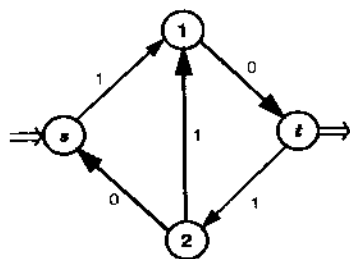
3.6.2 Thuật toán Ford-Fulkerson

Tất nhiên có thể đưa bài toán dòng cực đại về bài toán dòng trên mạng và áp dụng các phương pháp tổng quát là phương pháp đơn hình mạng và phương pháp chu trình cực phí âm. [Ford - Fulkerson 1956a, 1957] đã đưa ra thuật toán sử dụng khái niệm đường đi tăng dòng và dùng thuật toán đánh dấu để tìm đường đi tăng dòng. Thuật toán Ford-Fulkerson tận dụng cấu trúc riêng của bài toán dòng cực đại, có thể chứng minh được từ thuật toán chu trình cực phí âm. Trong mục này ta sẽ chứng minh thuật toán Ford-Fulkerson một cách khác, độc lập với phương pháp chung nói trên.

Cũng cần nói thêm rằng áp dụng thuật toán này ta cũng chứng minh được Định lý 3.6.1 về dòng cực đại cắt cực tiểu (xem [Bertsimas-Tsitsiklis 1997]).

ĐỊNH NGHĨA 3.6.1 (Đường đi tăng dòng). Cho x là một vectơ dòng chấp nhận được của bài toán dòng cực đại. *Đường đi tăng dòng* (augmenting path) là một đường đi từ s tới t (điểm nguồn và điểm hút, tương ứng) sao cho $x_{ij} < u_{ij}$ ở các cung tiến và $x_{ij} > 0$ ở các cung lùi, ở mọi cung của đường đi.

Nhận xét rằng đường đi tăng dòng có vai trò như chu trình chưa bão hòa, nhưng khác với bài toán dòng trên mạng, ở bài toán dòng cực đại ta phải xét các đường đi từ điểm nguồn đến điểm hút. Thực ra cũng có thể gọi đường đi tăng dòng là đường đi chưa bão hòa từ điểm nguồn đến điểm hút. Từ định nghĩa, rõ ràng ta có thể đẩy thêm một dòng, hay nói khác là làm *sự tăng dòng* (flow augmentation) dọc theo đường đi tăng dòng P một lượng thích hợp mà các ràng buộc của



H. 3.14

bài toán vẫn thỏa. Lượng dòng đẩy thêm này chỉ cần không vượt quá

$$\theta(P) = \min \left\{ \min_{(i,j) \in F} (u_{ij} - x_{ij}), \min_{(i,j) \in B} x_{ij} \right\}$$

ở đây, cũng như trước, ta kí hiệu, F và B là tập các cung tiến và tập các cung lùi trên đường đi, tương ứng. Chẳng hạn với bài toán dòng cực đại ở H.3.14, với $u_{ij} = 1, \forall (i, j)$, và các số cạnh cung là dòng trên cung x_{ij} của dòng chấp nhận được x hiện hành, đường đi P có tô đậm là một đường đi tăng dòng và $\theta(P) = 1$.

THUẬT TOÁN FORD-FULKERSON

1. Xuất phát từ một dòng chấp nhận được x .
2. Tìm một đường đi tăng dòng P . Nếu không có thì thuật toán kết thúc. Nếu có, tiếp đoạn 3 dưới đây.
3. Nếu $\theta(P) < \infty$ đẩy $\theta(P)$ đơn vị dòng theo đường đi P để được dòng chấp nhận được mới rồi quay lại đoạn 2.

Nếu $\theta(P) = +\infty$ thuật toán kết thúc.

Trong hai mục nhỏ sau đây ta sẽ chứng minh sự đúng đắn của thuật toán và cách thực hiện cụ thể.

1. *Cách tìm đường đi tăng dòng.* Ta sẽ sử dụng thuật toán đánh dấu có nội dung như sau. Một đường đi thỏa mãn Định nghĩa 3.6.1 về đường đi tăng dòng, nhưng chỉ đi từ s đến nút k nào đó (chưa tới t , nói chung) sẽ được gọi là *đường đi chưa bão hòa* (unsaturated path).

Ta nói nút i là *đã đánh dấu* (i is labeled) nếu ta biết là có một đường đi chưa bão hòa từ s tới i . Bây giờ ta sẽ xét tất cả các nút j có

nối trực tiếp đến nút i (sẽ gọi là ở cạnh nút i) xem chúng có thể được đánh dấu hay không khi i đã đánh dấu. Việc này được gọi là *quét* (scanning) nút i .

Nếu (i, j) có dòng trên cung $x_{ij} < u_{ij}$ thì ta có thể nối thêm cung (i, j) và đường đi chưa bão hòa P từ s đến i để được đường đi chưa bão hòa tới j . Vậy j có thể đánh dấu.

Tương tự nếu (j, i) có $x_{ji} > 0$ thì ta cũng thêm cung này vào P để thấy j có thể đánh dấu.

Gọi I là tập mọi nút đã đánh dấu nhưng chưa được quét. Ta có thuật toán sau để tìm đường đi tăng dòng.

THUẬT TOÁN ĐÁNH DẤU (The labeling algorithm)

Xuất phát với $I = \{s\}$ và s là nút đã đánh dấu duy nhất.

Một bước lặp sẽ có I hiện hành và gồm ba đoạn như sau.

1. Nếu $t \in I$ hoặc $I = \emptyset$, thuật toán kết thúc. Ngược lại thì chọn một nút $i \in I$ để quét và đưa nó ra khỏi I . Xét tất cả các nút cạnh i , tức là xét mọi cung có dạng (i, j) và (j, i) .

2. Nếu $(i, j) \in \mathcal{A}$, $x_{ij} < u_{ij}$ và j chưa đánh dấu thì đánh dấu nó và đưa j vào tập I .

3. Nếu $(j, i) \in \mathcal{A}$, $x_{ji} > 0$ và j chưa đánh dấu thì đánh dấu nó và đưa vào tập I .

Bây giờ ta xét kết quả của thuật toán đánh dấu. Nó có kết thúc hữu hạn hay không? Nhận xét rằng một nút được vào tập I chỉ khi chuyển từ chưa đánh dấu sang được đánh dấu. Do đó một nút chỉ được vào I nhiều nhất là một lần. Mà mỗi bước lặp bỏ một nút ra khỏi I . Do đó, vì số nút của mạng là hữu hạn, thuật toán phải kết thúc hữu

hạn.

Lúc thuật toán kết thúc, kết quả ra sao? Ở đoạn 1 của thuật toán có nói hai trường hợp kết thúc.

(i) Nếu thuật toán kết thúc vì t được đánh dấu, thì có một đường đi chưa bão hòa từ s đến t , tức là một đường đi tăng dòng. Đường đi tăng dòng này được chỉ ra cụ thể như sau. Trong quá trình đánh dấu, khi nút j được đánh dấu lúc quét nút i (đã được đánh dấu trước) thì ta ghi lại nút i là cha của j . Khi thuật toán dừng ở nút t ta đi ngược lại được cha của t , rồi đến cha của nó (ông của t) Đến khi quay lại đến s thì ta khôi phục được đường đi tăng dòng (từ s đến t).

(ii) Nếu thuật toán kết thúc vì $I = \emptyset$ thì sẽ không có đường đi tăng dòng. Thật vậy, gọi S là tập các nút đã đánh dấu cuối cùng. Giả sử phản chứng là có đường đi tăng dòng. Vì $s \in S$ và $t \notin S$ nên phải có hai nút kề nhau trên đường này là i và j sao cho $i \in S$ và $j \notin S$. Vì đây là hai nút kề nhau trên đường đi tăng dòng nên hoặc $(i, j) \in \mathcal{A}$ và $x_{ij} < u_{ij}$ hoặc $(j, i) \in \mathcal{A}$ và $x_{ji} > 0$. Trong cả hai trường hợp ta thấy nút j lại phải thuộc S , vì nó phải được đánh dấu khi quét nút i . Mâu thuẫn này chứng tỏ không có đường đi tăng dòng.

H.3.15 cho thí dụ một bài toán dòng cực đại, đã có một dòng chấp nhận được. Ta hãy tìm đường đi tăng dòng theo thuật toán đánh dấu xuất phát với $I = \{s\}$. Ta quét s và đưa s khỏi I . Nút 1, 2 được đánh dấu.

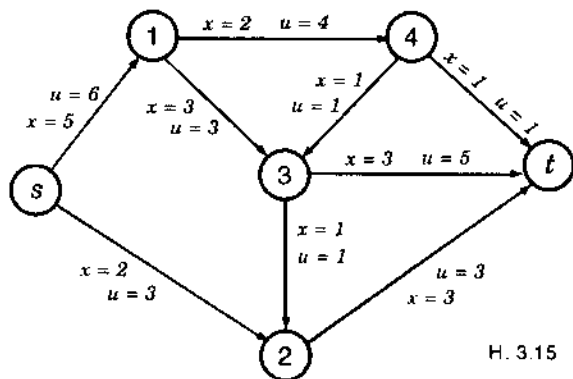
Bước lặp 2, với $I = \{1, 2\}$. Quét nút 1. Nút 4 được đánh dấu.

Bước lặp 3: $I = \{2, 4\}$. Quét nút 4. Không có nút nào được đánh dấu.

Bước lặp 4: $I = \{2\}$. Quét nút 2. Nút 3 được đánh dấu.

Bước lặp 5: $I = \{3\}$. Quét nút 3. Nút t được đánh dấu.

Ta kết luận là có đường đi tăng dòng. Để xác định nó ta lần ngược lại. t được đánh dấu khi quét nút 3. 3 được đánh dấu khi quét 2. 2 được đánh dấu khi quét s . Vậy đường đi tăng dòng là $s, 2, 3, t$.



H. 3.15

Cuối cùng ta xét độ phức tạp tính toán của thuật toán đánh dấu. Mỗi bước được quét một lần là cùng. Mỗi cung được xét khi đầu hoặc đuôi được quét, tức là nhiều nhất là hai lần. Khi xét một cung chỉ cần một số nhỏ, cố định các phép toán số học. Vì vậy độ phức tạp tính toán của thuật toán tỉ lệ với số cung, tức là $O(|A|)$. Vậy, tổng kết lại ta có định lý sau.

ĐỊNH LÝ 3.6.2. Với dòng chấp nhận được của bài toán dòng cực đại hiện hành đã có, thuật toán đánh dấu dừng ở việc đánh dấu t khi và chỉ khi có đường đi tăng dòng. Độ phức tạp tính toán là $O(|A|)$, tức là thuật toán chạy trong thời gian $O(|A|)$.

2. Sự kết thúc của thuật toán Ford - Fulkerson. Nếu thuật toán kết thúc vì $\theta(P) = \infty$, ta đã có một đường đi tăng dòng mà có thể đẩy theo nó lượng dòng lớn tùy ý.

ĐỊNH LÝ 3.6.3. Nếu thuật toán Ford - Fulkerson kết thúc vì không có đường đi tăng dòng thì dòng hiện hành là tối ưu.

CHỨNG MINH. Gọi S là tập các nút đã đánh dấu khi thuật toán kết thúc. Theo định nghĩa về đánh dấu, S gồm các nút i mà có đường đi chưa bão hòa từ s đến i . Vì thuật toán đánh dấu xuất phát từ s nên $s \in S$. Vì không có đường đi tăng dòng nên $t \notin S$. Vậy S là một cắt.

Mỗi cung $(i, j) \in \mathcal{A}$ với $i \in S, j \notin S$ phải có dòng hiện hành $x_{ij} = u_{ij}$ (trái lại thì j phải được đánh dấu). Vậy tổng dòng ra khỏi S chính bằng dung lượng $C(S)$. Hơn nữa, mỗi cung $(i, j) \in \mathcal{A}$ với $i \notin S, j \in S$ phải có $x_{ij} = 0$ (trái lại thì i phải được đánh dấu). Điều này nghĩa là dòng ra khỏi S không thể trở lại S . Do đó, khi thuật toán Ford - Fulkerson kết thúc, giá trị của dòng từ s đến t phải bằng $C(S)$. Nhưng theo (3.16) giá trị của dòng bất kì không được quá $C(S)$. Vậy dòng hiện hành đã tối ưu. $\square$

Định lí sau đây nói về nghiệm nguyên và sự kết thúc hữu hạn.

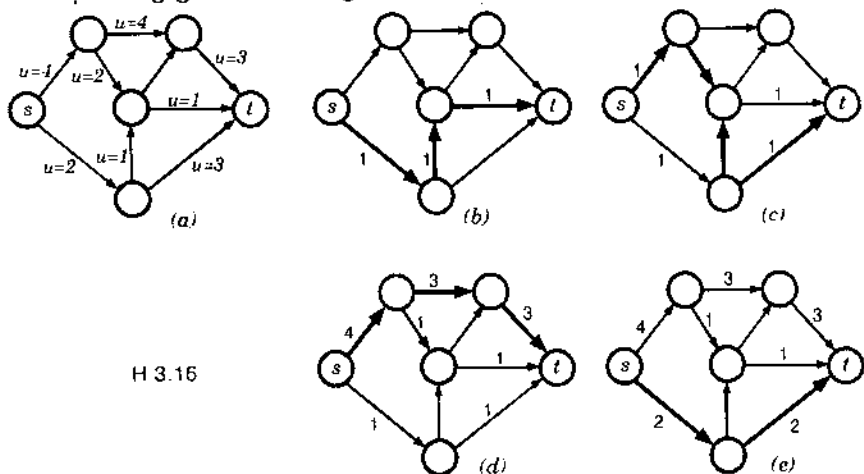
ĐỊNH LÍ 3.6.4. *Giả sử mọi u_{ij} là nguyên hoặc vô hạn và thuật toán Ford - Fulkerson bắt đầu bằng một vectơ dòng chấp nhận được nguyên. Khi đó mọi dòng chấp nhận được qua các bước đều nguyên. Nếu, thêm vào đó, giá trị tối ưu (tức là giá trị dòng cực đại) được biết là hữu hạn, thì thuật toán kết thúc sau hữu hạn bước.*

CHỨNG MINH. Định lí này có thể được chứng minh như hệ quả của Định lí 3.5.3. Ở đây ta chứng minh trực tiếp. Nhận xét rằng nếu dòng hiện hành là nguyên và mọi cung có tải năng u_{ij} nguyên hoặc vô hạn, thì mọi đường đi P đều có $\theta(P)$ nguyên hoặc vô hạn. Vậy tính nguyên của dòng hiện hành được duy trì ở mọi bước lặp. Giá trị dòng nhận được qua mỗi bước lặp tăng ít nhất là 1 (vì $\theta(P)$ nguyên). Nếu giá trị dòng cực đại là nguyên thì thuật toán phải kết thúc sau hữu hạn bước, khi giá trị dòng nhận được đạt đến giá trị này. $\square$

Nếu các u_{ij} đều là số hữu tỉ thì thuật toán Ford - Fulkerson vẫn kết thúc sau hữu hạn bước vì ta có thể nhân các u_{ij} với mẫu số chung nhỏ nhất để được bài toán tương đương với tải năng các cung đều

nguyên. Tuy nhiên người ta tìm được thí dụ cho thấy nếu có u_{ij} vô tỉ, thuật toán có thể không kết thúc được, mặc dù giá trị dòng tối ưu là hữu hạn. Thậm chí có thể giá trị dòng tăng đơn điệu qua các bước lặp và có giới hạn nhưng giới hạn nhỏ hơn hẳn giá trị tối ưu. May mắn rằng với “trường hợp vô tỉ”, nếu dùng các cách chọn đường đi tăng dòng đặc biệt, thuật toán vẫn kết thúc sau hữu hạn bước. Chẳng hạn người ta chứng minh được rằng nếu chọn đường đi tăng dòng có số cung ít nhất có thể được thì thuật toán kết thúc sau $O(|A| \cdot |A|)$ bước lặp. Mỗi bước lặp có độ phức tạp là $O(|A|)$ theo Định lý 3.6.2. Do đó, độ phức tạp của thuật toán là $O(|A| \cdot |A|^2)$.

Ta hãy xét một thí dụ đơn giản về áp dụng thuật toán Ford - Fulkerson. Giả sử ta xuất phát từ dòng chấp nhận được 0 cho mạng ở H.3.16 (a). Ở đây các số ghi cạnh cung là tải năng u_{ij} . Còn các số ghi cạnh cung ở (b), (c), (d) và (e) là dòng trên cung. Đường đi tô đậm ở (b) là một đường đi tăng dòng với $\theta(P) = 1$. Sau khi đẩy lượng dòng $\theta(P)$ theo P ta được dòng ghi ở cạnh các cung của P. Đường đậm ở (c) là một đường đi tăng dòng với $\theta(P) = 1$. Sau khi tăng lượng dòng này ta được dòng ghi ở các cung. Hình (d) cũng tương tự. Dòng ở (e) sau



H 3.16

khi tăng có giá trị là 6. Ở đây không có đường đi tăng dòng nữa. Theo Định lí 3.6.3 thì dòng đã là cực đại. Điều này có thể thấy trực tiếp vì tổng tải năng các cung rời khỏi s bằng 6, nên số này là một "cổ chai" không thể vượt quá.

3.7 BÀI TOÁN ĐƯỜNG NGẮN NHẤT

Bài toán đường ngắn nhất (shortest path problem) là một bài toán quan trọng nảy sinh từ nhiều bài toán thực tế về vận tải, mạng thông tin, điều khiển tối ưu,... Có thể phát biểu một dạng toán học chung cho các bài toán như vậy thành bài toán dòng trên mạng như sau. Cho một đồ thị có hướng $(\mathcal{N}, \mathcal{A})$. Mỗi cung (i, j) có cước phí $c_{ij} \geq 0$ chính là độ dài của cung. Để tìm đường ngắn nhất từ một nút s đến một nút r ta sẽ thấy là cần tính đường ngắn nhất từ nhiều hoặc thậm chí mọi nút khác r tới nút r . Vì vậy ta gọi bài toán đường ngắn nhất là bài toán tìm đường ngắn nhất từ mọi nút trong $\mathcal{N}$ tới một nút $r \in \mathcal{N}$ cho trước, gọi là *nút gốc* (root node). Để đưa về bài toán dòng trên mạng, đặt $b_i = 1$ cho mỗi nút $i \neq r$ và

$$b_r = - \sum_{i \neq r} b_i.$$

Ta có thể giải bài toán đường ngắn nhất bằng thuật toán đơn hình mạng. Đường ngắn nhất từ nút i tới r chính là các cung thuộc cây bao trùm tối ưu T nối từ i đến r (chú ý đây là đường đi có hướng, trong đường đi không có cung lùi). Ta có ngay độ dài đường đi ngắn nhất như sau. Vì

$$p_i^* - p_r^* = c_{ij} \quad \forall (i, j) \in T,$$

nên $\sum c_{ij}$ trên các cung từ i đến r sẽ là $p_i^* - p_r^*$, ở đây p^* là vectơ nghiệm đối ngẫu tối ưu.

Có nhiều cách giải bài toán đường ngắn nhất nhanh hơn thuật toán đơn hình mạng. Ở đây ta chỉ giới thiệu hai thuật toán.

3.7.1 Thuật toán sửa nhãn

Gọi v_i là độ dài đường đi ngắn nhất từ nút i tới nút r . v_i cũng được gọi là *nhãn* (label) của nút i . Rõ ràng

$$v_r = 0.$$

Xét cung (i, j) . Nếu đường ngắn nhất từ i tới r có dùng cung này thì $v_i = c_{ij} + v_j$. Nếu từ i có nhiều cung qua các nút khác nhau, thì đường ngắn nhất phải chọn cung tốt nhất trong đó, tức là

$$v_i = \min \{c_{ij} + v_j : (i, j) \in \mathcal{A}, i \neq r\}. \quad (3.19)$$

(3.19) gọi là *phương trình Bellman* (Bellman's equation). Nó có thể được giải thích như sau. Ở đường đi tối ưu (trong bài toán đang xét là đường đi ngắn nhất), khi xét ở thời điểm bất kì j , bắt buộc từ đoạn j tới đích phải là tối ưu (khi ta chưa xét gì đến đoạn trước, ở đây là cung (i, j)), nên khoảng cách từ j đến đích phải là v_j . Từ đó muốn đường là tối ưu từ i , ta phải tìm cực tiểu $c_{ij} + v_j$ với mọi $(i, j) \in \mathcal{A}$. Đây là nội dung của *nguyên lý tối ưu* (principle of optimality) của Bellman trong *quy hoạch động* (dynamic programming), là một lý thuyết thuật toán phát triển bởi Bellman, để giải rất nhiều bài toán tối ưu, nhất là các bài toán điều khiển.

Từ phương trình Bellman, rõ ràng tập các cung mà đường ngắn nhất tới r có đi qua là

$$\{(i, j) \in \mathcal{A} : v_i = c_{ij} + v_j\}.$$

Tập này có thể chứa các đường ngắn nhất khác nhau từ cùng một nút nào đó tới r . Do đó, nói chung tập này không phải là một cây. Tuy nhiên đường đi dọc theo các cung này tới r sẽ đều là đường ngắn nhất.

Để giải bài toán đường ngắn nhất ta phải giải phương trình Bellman (3.19), là một hệ các phương trình với ẩn là v_i , $i \in \mathcal{N}$.

Ý tưởng của *thuật toán sửa nhãn* (label correcting algorithm) để giải (3.19) là ta đặt trước coi như xấp xỉ nghiệm nhãn v_j , $j \in \mathcal{N}$ thay vào vế phải (3.19) để tính vế trái. Ở bước tiếp theo, kết quả vừa tính được lại thay vào vế phải. Như vậy đây cũng là một *phương pháp xấp xỉ liên tiếp* (method of successive approximation). Cụ thể như sau. Chọn các nhãn xuất phát là

$$v_i^0 = \begin{cases} 0 & \text{nếu } i = r, \\ +\infty & \text{nếu } i \neq r. \end{cases}$$

Ở mỗi bước k ta sửa nhãn bằng phương trình Bellman :

$$v_i^{k+1} = \begin{cases} 0 & \text{nếu } i = r, \\ \min \{c_{ij} + v_j^k : (i, j) \in \mathcal{A}\} & \text{nếu } i \neq r. \end{cases}$$

Ta xét khả năng kết thúc hữu hạn của thuật toán. Thuật toán dừng khi mọi v_i không thay đổi qua một bước. Ta sẽ chứng tỏ thuật toán phải kết thúc sau không quá $n = |\mathcal{N}|$ bước. Chú ý rằng v_i^k chính là độ dài đường ngắn nhất từ i tới r chứa không quá k cung. Vì đường ngắn nhất cần tìm đi qua mỗi nút cùng lắm một lần nên thuật toán không có quá n bước lặp.

Vì ở mỗi bước ta phải nhìn đến mọi cung nên số phép toán cộng hoặc so sánh sử dụng trong thuật toán là cỡ nm .

3.7.2 Thuật toán đặt nhãn (thuật toán Dijkstra)

Thuật toán này, công bố trong [Dijkstra 1959], tỏ ra hiệu quả hơn thuật toán sửa nhãn, qua thực tế sử dụng cho thấy. Ý tưởng của thuật toán là lần lượt tìm đường đi ngắn nhất từ mỗi nút đến nút gốc kí hiệu là n , bắt đầu từ nút có đường này ngắn hơn cả, rồi theo thứ tự nút có đường ngắn hơn làm trước.

Nếu một cung nối i và j không có ta coi là $c_{ij} = +\infty$ để coi là mọi cặp nút đều có cung nối. Vì nghiệm bài toán đối ngẫu được xác định sai khác hằng số cộng nên ta đặt $p_n^* = 0$. Do đó đường đi ngắn nhất từ nút i đến n là p_i^* . Thuật toán đặt nhãn xác định p_i^* và đường ngắn nhất cụ thể có độ dài đó như sau.

Đầu tiên ta tìm nút l có $p_l^* \leq p_i^* \forall i \neq n$. Đó là nút l sao cho

$$c_{ln} = \min_{i \neq n} c_{in},$$

và khi đó $p_l^* = c_{ln}$.

Thật vậy, mỗi đường đi tới n phải có cung cuối cùng có dạng (i, n) mà $c_{in} \geq c_{ln}$. Do đó $p_k^* \geq c_{ln} \forall k \neq n$. Riêng với nút l ta lại có thêm tính chất $p_l^* \leq c_{ln}$. Vậy $p_l^* = c_{ln} \leq p_i^* \forall i \neq n$.

Vậy ta đã được đường đi ngắn nhất từ l đến n . Để tìm tiếp đường ngắn nhất cho nút thứ hai ta tìm cách có mạng mới, bớt đi nút l và làm lại như đã làm ở trên. Xét nút i bất kì. Một trong các cách đi từ i đến n là qua l bằng cung (i, l) . Khi đã ở l , thì để đi ngắn nhất tất phải đi theo cung (l, n) . Ta đổi lại mạng như sau. Thay hai cung (i, l) và (l, n) bởi một cung thêm vào là $(i, n)'$ có độ dài $c_{il} + c_{ln}$. Sau khi làm như vậy cho mọi nút $i \neq n$, $i \neq l$ ta bỏ nút l đi. Chú ý rằng nút i (bất kì) bây giờ có thể có hai cung tới n là (i, n) và $(i, n)'$. Ta sẽ thay hai cung này bằng một cung, cũng kí hiệu là (i, n) nhưng với độ dài là

$$\min \{c_{in}, c_{il} + c_{ln}\}.$$

Như vậy ta đã được một bài toán đường ngắn nhất mới chỉ có $n-1$ nút. Ta lại làm như với nút l ở trên.

Mỗi bước lặp cho ta thêm đường ngắn nhất từ một nút và bài toán mới bớt đi nút đó. Do đó, sau $n-1$ bước thuật toán kết thúc. Tóm tắt như sau.

THUẬT TOÁN ĐẶT NHÃN (label setting algorithm)

1. Tìm nút $l \neq n$ sao cho $c_{in} \leq c_{il} \forall i \neq n$. Đặt $p_l^* = c_{in}$.
2. Với mỗi $i \neq l, n$, đặt c_{in} mới

$$c_{in} := \min \{c_{in}, c_{il} + c_{ln}\}.$$

3. Bỏ nút l từ đồ thị và làm lại đoạn 1 với đồ thị mới.

THÍ DỤ 3.7.1. Xét bài toán ở H.3.17, với $n = 4$ và các số cạnh cung là chiều dài cung. Ở (a) là bài toán đường ngắn nhất. Ta có $l = 3$ và $p_3^* = 1$. Ta thực hiện đoạn 2 là sửa độ dài cung:

$$c_{14} := \min \{ \infty, 9 + 1 \} = 10,$$

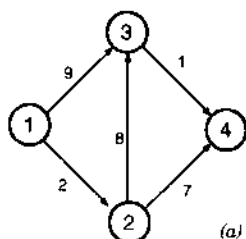
$$c_{24} := \min \{ 7, 8 + 1 \} = 7.$$

Sang đoạn 3 ta bỏ nút 3 và được đồ thị mới ở (b).

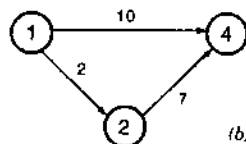
Thực hiện bước lặp hai với mạng mới. Ta lại có $l = 2$ và $p_2^* = 7$. Tiếp theo là sửa độ dài cung:

$$c_{14} := \min \{ 10, 2 + 7 \} = 9.$$

Nút 2 bị bỏ đi. Ta chỉ còn nút 1 là khác n nên p_1^* chính là c_{14} hiện hành, ước là $p_1^* = 9$.



(a)



(b)

H. 3.17

Ta xem xét độ phức tạp tính toán của thuật toán đặt nhãn. Ở mỗi bước lặp ta phải so sánh các c_{in} rồi sửa các độ dài cung c_{in} . Các việc đó chỉ cần $O(n)$ phép toán số học. Ta có $n - 1$ bước, nên độ phức tạp tính toán của cả thuật toán là $O(n^2)$.

Đối với các đồ thị nhiều cạnh, cỡ $O(n^2)$ cạnh, ta thấy thuật toán bất kì để giải bài toán đường ngắn nhất đều cần ít nhất $O(n^2)$ phép tính số học vì mỗi cung phải xét ít nhất một lần. Do đó thuật toán

Dijkstra là tốt nhất rồi.

Đối với các đồ thị thưa, với m nhỏ hơn n nhiều thì người ta chứng minh được độ phức tạp tính toán của thuật toán Dijkstra là $O(m \log n)$.

Đọc giả có thể đọc về thuật toán Bellman - Ford, một thuật toán khác khá thông dụng trong [Bertsimas - Tsitsiklis 1997].

BÀI TẬP CHƯƠNG 3

3.1. Xét đồ thị vô hướng $(\mathcal{A}, \mathcal{A})$. Hai nút i và j trong $\mathcal{A}$ được gọi là *thông nhau* (connected) nếu có một đường đi từ i tới j , kí hiệu là $i \sim j$.

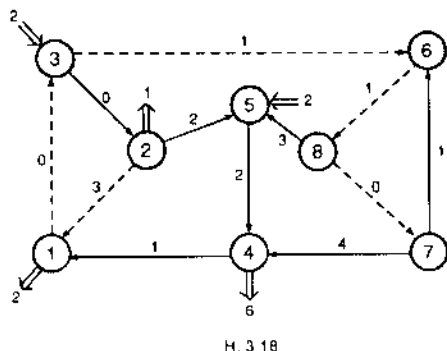
(a) Chứng minh rằng quan hệ $\sim$ trong $\mathcal{A}$ là quan hệ tương đương (tức là thỏa các tính chất phản xạ, đối xứng và bắc cầu).

(b) Phân hoạch $\mathcal{A}$ ra thành các tập $\mathcal{A}_1, \dots, \mathcal{A}_k$ sao cho hai nút thuộc cùng một tập con khi và chỉ khi chúng thông nhau. Chứng minh rằng $\text{rank } A = |\mathcal{A}| - k$, ở đây A là ma trận nối.

3.2. Loại bỏ cận dưới của dòng. Chứng minh rằng nếu trong bài toán dòng trên mạng có ràng buộc về cận dưới dòng $x_{ij} \geq d_{ij}$, thì ta có thể đưa về bài toán tương đương với $d_{ij} = 0$ với mọi $(i, j) \in \mathcal{A}$. Gợi ý: Đặt $\bar{x}_{ij} = x_{ij} - d_{ij}$ và lập bài toán mới cho $\bar{x}_{ij}$. Các b_i sẽ thay đổi thế nào?

3.3. Bài toán thi đấu - tournament problem. Có n đội thi đấu vòng tròn và các trận không được có kết quả hòa. Giả sử x_i là số trận thắng của đội i và X là tập mọi vectơ $(x_1, \dots, x_n)$ kết quả có thể có. Bài toán đặt ra là phải xác định xem một vectơ $(x_1, \dots, x_n)$ bất kì có phải thuộc X hay không. Hãy diễn đạt thành bài toán dòng trên mạng.

3.4. Một công ti khai thác gỗ, có M đơn vị rừng, muốn tìm phương án đốn cây lấy gỗ trong thời gian K năm. Được biết là đơn vị rừng i có a_{ij} tấn gỗ để đốn trong năm j . Công ti cần đáp ứng nhu cầu d_j tấn gỗ trong năm j . Tuy nhiên, trong mỗi năm j , $j = 1, \dots, K$, do hạn chế của công suất, công ti chỉ thu hoạch được nhiều nhất là u_j tấn. Gỗ đã thu hoạch có thể lưu kho để đáp ứng nhu cầu các năm sau với cước là c_j cho một tấn lưu



H. 3.18

từ năm $j - 1$ đến năm j . Giả sử rằng gỗ có thể khai thác nhưng chưa kịp đốn có thể để khai thác vào các năm sau. Diễn đạt bài toán tìm phương án thu hoạch gỗ làm cực tiểu cước phí lưu kho và đảm bảo nhu cầu gỗ như bài toán dòng trên mạng.

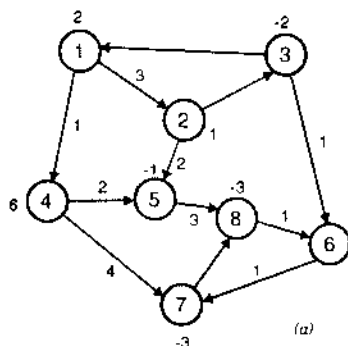
3.5. Xét bài toán dòng trên mạng với tải năng không hạn chế ở H.3.18. Ở đây số bên cạnh các cung là cước phí.

(a) Tìm ma trận A tương ứng.

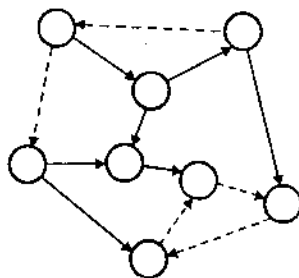
(b) Giải bài toán bằng thuật toán đơn hình mạng, bắt đầu bởi cây gồm các cung có nét liền.

3.6. Giải bài toán dòng trên mạng ở H.3.19(a) xuất phát từ cây bao trùm chỉ ra ở (b) bằng đường liền (số tại nút là nguồn b_i , số dọc theo cung là cước c_{ij} , các số không đã bỏ đi)

3.7. Xét bài toán tìm dòng có cước nhỏ nhất ở H.3.20 với cách ghi các số như ở H.3.19. Nếu nghiệm ứng với cây bao trùm cho bởi đường liền là tối ưu thì hãy chứng minh. Nếu chưa phải thì hãy tìm nghiệm tối ưu, xuất phát từ cây bao trùm này.



(a)



(b)

H. 3.19

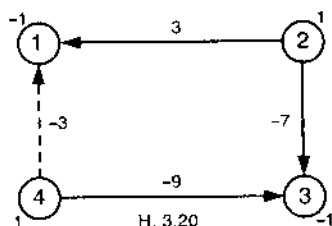
3.8. Phương pháp đơn hình đối ngẫu. Xét bài toán dòng trên mạng với tài năng không hạn chế.

(a) Chứng minh rằng mỗi cây bao trùm xác định một nghiệm cơ sở cho bài toán đối ngẫu.

(b) Giả sử có nghiệm cơ sở chấp nhận được của bài toán đối ngẫu, tương ứng với một cây. Chứng minh rằng nó là tối ưu khi và chỉ khi nghiệm cây tương ứng cho bài toán gốc là chấp nhận được.

(c) Nếu nghiệm cây ở (b) là không chấp nhận được, ta bỏ một cung có dòng âm. Nếu ta đang muốn có dòng đối ngẫu chấp nhận được thì phải chọn cung vào cây như thế nào?

(d) Nhận xét rằng cung vào chia cây thành hai phần. Xét biến đối ngẫu xem phải cập nhật theo thuật toán đơn hình đối ngẫu thế nào. Chứng minh rằng các biến đối ngẫu trong một phần của cây là không đổi còn ở phần kia thì mọi biến đều thay đổi cùng một lượng.



3.9. Thuật toán chu trình cước phí âm với quy tắc cải thiện lớn nhất. Đây là thuật toán chu trình cước phí âm trong đó ta chọn chu trình C với giá trị $\delta(C)|c^T h^C|$ lớn nhất. Giả sử x là dòng hiện hành và x^* là dòng tối ưu.

(a) Chứng minh rằng $x^* - x$ bằng một tổ hợp tuyến tính với các hệ số không âm của nhiều nhất là m hoàn lưu đơn giản, ở đây m là số cung. Hơn nữa, mỗi hoàn lưu đơn giản này đều tương ứng với một chu trình chưa bão hòa. Gợi ý: dùng kết quả Bài tập 3.16.

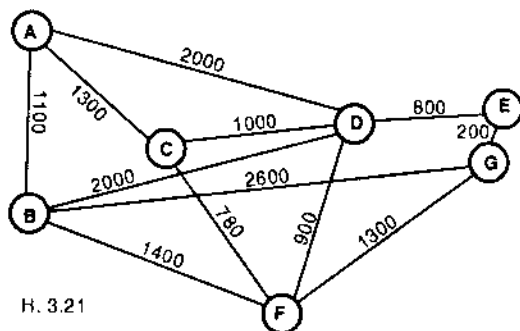
(b) Chứng minh rằng sự cải thiện cước phí ở mỗi bước lặp ít nhất là $\frac{1}{m}(c^T x - c^T x^*)$.

(c) Giả sử dữ kiện của bài toán là nguyên. Chứng minh rằng thuật toán kết thúc sau $O(m \log(mCU))$ bước lặp, ở đây C và U là cận trên l_{ij} và u_{ij} tương ứng.

3.10. Chứng minh rằng có một tương ứng 1-1 giữa các đường đi tăng dòng trong thuật toán dòng cực đại và các chu trình cước phí âm chưa bão hòa trong diễn đạt bài toán dòng cực đại thành bài toán dòng trên mạng.

3.11. Bài toán hôn nhân. Một làng nhỏ có n đàn ông chưa vợ, n cô gái chưa chồng và m bà mối. Bà mối i biết một số người trong đó và có thể mối lái được cho tới b_i cặp hôn nhân trong số người họ biết. Bài toán là xác định số cực đại các cặp hôn nhân. Chứng minh rằng có thể thay bài toán bằng một bài toán dòng cực đại.

3.12. Phải thiết kế mạng thông tin cấp nối các địa điểm chính (có khoảng cách cho trong H. 3.21) của một khu vực thế nào để tổng độ dài dây cáp là nhỏ nhất?



H. 3.21

CHƯƠNG 4

ĐIỀU HÀNH DỰ ÁN BẰNG PHƯƠNG PHÁP PERT-CPM (PHƯƠNG PHÁP SƠ ĐỒ MẠNG LƯỚI)

Dự án (project) là một tập hợp các *hoạt động* (activity) liên quan với nhau và phải được thực hiện theo một thứ tự nào đó cho đến khi hoàn thành toàn bộ các hoạt động. *Hoạt động* được hiểu như là một việc đòi hỏi thời gian, và *nguyên liệu* (resource) để hoàn thành. Trước kia để điều hành dự án người ta thường dùng *biểu đồ Gantt* (Gantt bar chart), là một đồ thị gồm các đường kẻ ngang, biểu thị điểm khởi công và kết thúc hoạt động. Nhược điểm của biểu đồ là không xác định được quan hệ giữa các hoạt động, nên không áp dụng được cho các *dự án lớn* (large-scale project), đòi hỏi *đặt kế hoạch* (planning), *điều hành thực hiện* (scheduling) và *kiểm tra* (controlling) một cách hệ thống và hiệu quả, thậm chí phải tối ưu hóa hiệu quả (về thời gian và tiết kiệm nguyên liệu). Vì vậy, gần như đồng thời vào năm 1956-1958, hai phương pháp kế hoạch, điều hành và kiểm tra dự án đã ra đời. *Phương pháp đường găng* hoặc *phương pháp đường tới hạn* (critical path method, viết tắt là CPM) được E.I. du Pont de Nemours và công ty xây dựng của ông đưa ra. Phương pháp thứ hai có tên là *Kỹ thuật xem xét và đánh giá dự án* (project evaluation and review technique, viết tắt là PERT) là kết quả nghiên cứu của một công ty tư vấn theo đặt hàng của hải quân Mỹ, dùng để điều hành các hoạt động nghiên cứu và phát triển chương trình tên lửa đối cực. Hai phương pháp được hình thành độc lập nhưng rất giống nhau, cùng

nhằm vào mục đích điều hành thời gian là chính. Sự khác nhau chính là trong CPM thời gian ước lượng cho các hoạt động, được coi là *tất định* (deterministic), còn trong PERT có thể là *ngẫu nhiên* (probabilistic). Ngoài ra CPM có tính đến quan hệ thời gian và cước phí, còn PERT tập trung vào thời gian. Ngày nay, khi đã phát triển lên, hai phương pháp được coi là một, dưới một tên chung là *phương pháp điều hành dự án* (project scheduling method) hoặc *phương pháp điều hành dự án PERT-CPM*, hoặc *phương pháp sơ đồ mạng lưới* hoặc *hệ thống kiểu PERT* (PERT-type system). Nó được dùng để thực hiện rất nhiều kiểu dự án, từ xây dựng, lập trình máy tính, sản xuất phim đến vận động tranh cử chính trị hoặc các cuộc giải phẫu thuật phức tạp.

Phương pháp điều hành dự án PERT-CPM gồm ba pha (tức là ba khâu): kế hoạch, điều hành và kiểm tra điều chỉnh. *Pha kế hoạch* có nội dung là lập một *sơ đồ mạng lưới* (arrow network diagram hoặc arrow diagram), tương tự một đồ thị có hướng. Pha này mở đầu bằng việc tách dự án thành nhiều hoạt động riêng và định thời gian hoàn thành cho chúng. Trong mạng, mỗi cung có hướng biểu diễn hoạt động và cả sơ đồ mạng biểu thị mối quan hệ giữa các hoạt động. Mỗi nút biểu thị một *biến cố* hoặc *sự kiện* (event), đánh dấu hoàn thành một *số hoạt động* (activity) là các cung đi vào nút, và bắt đầu các hoạt động ứng với các cung ra khỏi nút.

Pha điều hành (scheduling phase) có nhiệm vụ xây dựng biểu đồ thời gian, chỉ rõ thời điểm bắt đầu và kết thúc của mỗi hoạt động và mối quan hệ giữa các hoạt động. Nói riêng, điều quan trọng là phải tính chính xác các hoạt động *tới hạn*, tức là *găng* (critical), cần chú ý đặc biệt khi thực hiện, để toàn bộ dự án được hoàn thành đúng hạn.

Pha kiểm tra bao gồm việc sử dụng sơ đồ mạng lưới, và biểu đồ thời gian để theo dõi và báo cáo định kì tiến triển của dự án. Nếu cần thì phải phân tích lại và xác định sơ đồ mới cho phần dự án còn lại.

4.1 LẬP SƠ ĐỒ MẠNG LƯỚI

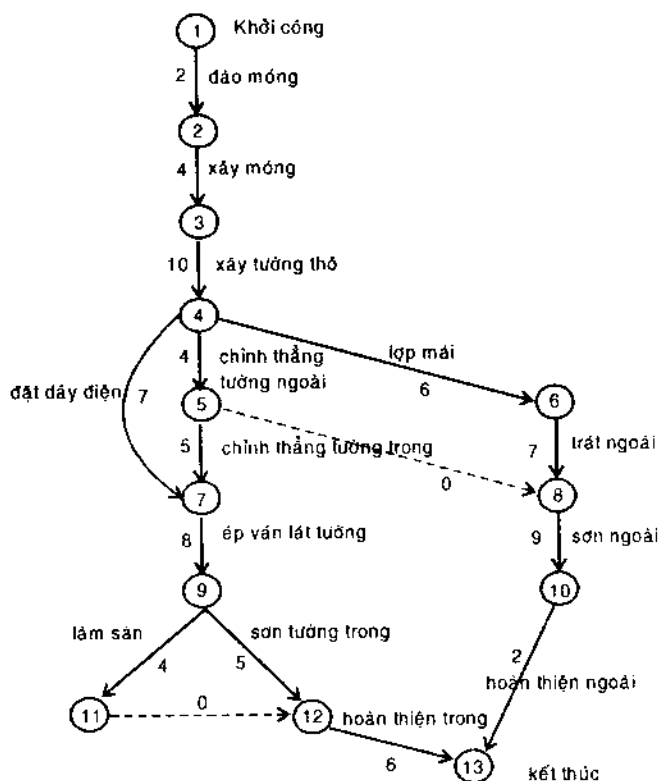
Như trên đã nói, pha đầu của phương pháp PERT-CPM là lập kế hoạch thể hiện ở một sơ đồ mạng lưới, biểu diễn như một đồ thị có hướng. Hãy xét dự án xây một tòa nhà. Việc tách dự án thành các hoạt động như đào đất, xây móng, xây tường thô, lợp mái, đặt đường dây điện... và định thời gian cho từng hoạt động, cũng như xác định hoạt động (cũng gọi là việc) nào trước, việc nào liền ngay trước, việc nào sau... là do kiến trúc sư hoặc kỹ sư xây dựng làm. Dựa vào đó, người quản lý dự án lập được sơ đồ mạng lưới như H. 4.1. Các số bên cạnh cung là thời gian thực hiện hoạt động đó.

Qua sơ đồ mạng lưới ở H. 4.1 (trang sau) ta thấy rõ mối quan hệ giữa các hoạt động về thời gian. Chẳng hạn hoạt động (6, 8) là trát ngoài - phải sau (4, 6) là lợp mái, nhưng độc lập với (5, 7) là chỉnh tường trong. Cũng vậy (4, 7) độc lập với (4, 5) và (5, 7). Ở đây có hai *hoạt động giả* (dummy activity) với thời gian để thực hiện bằng 0 được đưa vào để đảm bảo các quy tắc xây dựng sơ đồ.

- Cung giả (11, 12), kí hiệu bởi đường đứt đoạn, đưa vào để đảm bảo quy tắc không có hai hoạt động cùng biến cố bắt đầu và kết thúc, tức là không có 2 cung có cùng gốc và ngọn (tức là đồ thị là đơn). Việc sơn tường trong và làm sàn có cùng biến cố đầu là nút 9, tức là biến cố lát ván tường xong, và biến cố cuối là nút 12 (làm sàn và sơn tường xong, bắt đầu hoàn thiện trong). Do đó ta phải thêm nút 11 là biến cố giả và cung giả (11, 12).

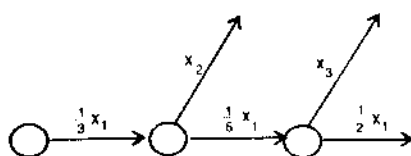
- Cung giả (5, 8) để chỉ rằng hoạt động (4, 5) phải hoàn hành trước khi bắt đầu hoạt động (8, 10) (nếu bỏ cung giả này thì thời điểm làm hai việc là độc lập).

Cung giả này là phục vụ cho quy tắc sơ đồ mạng lưới phải thể hiện đủ quan hệ thứ tự cần có.



H. 4.1

• Nếu quan hệ thời gian có dạng: việc x_2 bắt đầu khi xong $\frac{1}{3}x_1$ việc x_1 , việc x_3 bắt đầu khi xong một nửa x_1 , thì ta phải thêm các nút đánh dấu các biến cố xong $\frac{1}{3}x_1$ và xong $\frac{1}{2}x_1$ đó như ở H. 4.2.



H. 4.2

Tóm lại: Sơ đồ mạng lưới phải là một đồ thị có hướng, đơn, liên thông, không có khuyên (tức là cung có gốc và ngọn cùng là một nút), không có chu trình có hướng (directed cycle), có nút khởi công và nút kết thúc.

4.2 PHÂN TÍCH CÁC CHỈ TIÊU THỜI GIAN. XÁC ĐỊNH ĐƯỜNG GẮNG

Pha điều hành có nhiệm vụ phân tích các chỉ tiêu thời gian và đưa ra các bảng và số liệu cần thiết trên sơ đồ mạng lưới. Nếu trong dự án phải điều hành cả nguyên liệu (hoặc nhân lực) thì phải xét cả các chỉ tiêu đó, ta sẽ nói đến ở mục sau.

4.2.1 Tính các thời điểm

Chỉ tiêu đầu tiên là *thời điểm sớm của biến cố* (earliest time for an event) là thời điểm biến cố xảy ra khi mọi hoạt động trước nó được bắt đầu sớm nhất có thể. Thời điểm sớm của biến cố i thường kí hiệu là E_i . Các E_i được tính theo hướng tăng (forward pass), tức là đi từ nút khởi công theo thứ tự tăng của nút i . Như vậy với nút khởi công 1 thì $E_1 = 0$. Đến nút 2 trong sơ đồ ở H. 4.1 thì E_2 rõ ràng bằng 2 vì biến cố hoàn thành hoạt động (1, 2) phải là $E_1 + t_{12}$, ở đây t_{12} là thời gian thực hiện hoạt động (1, 2). Việc tính $E_3, E_4, E_5, E_6, E_9, E_{10}$ và E_{11} cũng tương tự vì các nút tương ứng chỉ có một cung vào, khi đó

$$E_i = E_j + t_{ji},$$

ở đây j là nút ngay trước i . Chẳng hạn $E_6 = E_4 + t_{46} = 16 + 6 = 22$. Nếu có nhiều cung vào nút, tức là nhiều hoạt động kết thúc tại biến cố, thì từ định nghĩa E_i rõ ràng đây là thời điểm mọi hoạt động đó vừa xong cả, tức là phải lấy maximum của các tổng. Chẳng hạn

$$E_7 = \max \{E_4 + t_{47}, E_5 + t_{57}\} = \max \{16+7, 20+5\} = 25,$$

$$E_8 = \max \{E_5 + t_{58}, E_6 + t_{68}\} = \max \{20+0, 22+7\} = 29.$$

Tổng quát, công thức tính E_i cho mọi trường hợp là

$$E_i = \max_j \{E_j + t_{ji}\},$$

ở đây j là các nút ngay trước i , tức là có cung nối tới i . Các E_i được ghi ở H. 4.3 là số đầu trong ngoặc ở mỗi nút.

Thời điểm muộn (latest time) của biến cố j là thời điểm muộn nhất mọi cung đi vào biến cố j đều hoàn thành mà không làm thay đổi thời điểm kết thúc dự án sớm nhất có thể, kí hiệu là L_j . Đối lại với E_j , các L_j được tính theo hướng lùi (backward pass), tức là đi từ nút kết thúc. Theo định nghĩa, ở nút kết thúc thì $E_n = L_n$, ở thí dụ H. 4.1 là $E_{13} = L_{13} = 44$. Nếu ở biến cố chỉ có một cung ra, tức là một hoạt động được bắt đầu thì, thời điểm muộn là

$$L_j = L_i - t_{ji},$$

tức là thời điểm muộn của nút ngay sau nó trừ đi thời gian thực hiện hoạt động nối hai nút. Các biến cố 12, 11, 10, 8, 7, 6, 3, 2 và 1 ở H. 4.1 là ở trường hợp này. Nếu có nhiều cung ra khỏi biến cố, thì theo định nghĩa ta có

$$L_j = \min_i \{L_i - t_{ji}\}$$

ở đây min theo các nút i ngay sau j và t_{ji} là thời gian thực hiện hoạt động nối (j, i) . Các nút 9, 5, 4 là ở trường hợp này, chẳng hạn

$$L_9 = \min \{L_{11} - t_{9,11}, L_{12} - t_{9,12}\} = \min \{38 - 4, 38 - 5\} = 33.$$

Hãy chú ý sự "đối xứng" của quá trình tính E_i và L_j . Các L_j được ghi ở số thứ 2 trong ngoặc ở mỗi nút trong H. 4.3.

4.2.2 Tính thời gian dự trữ

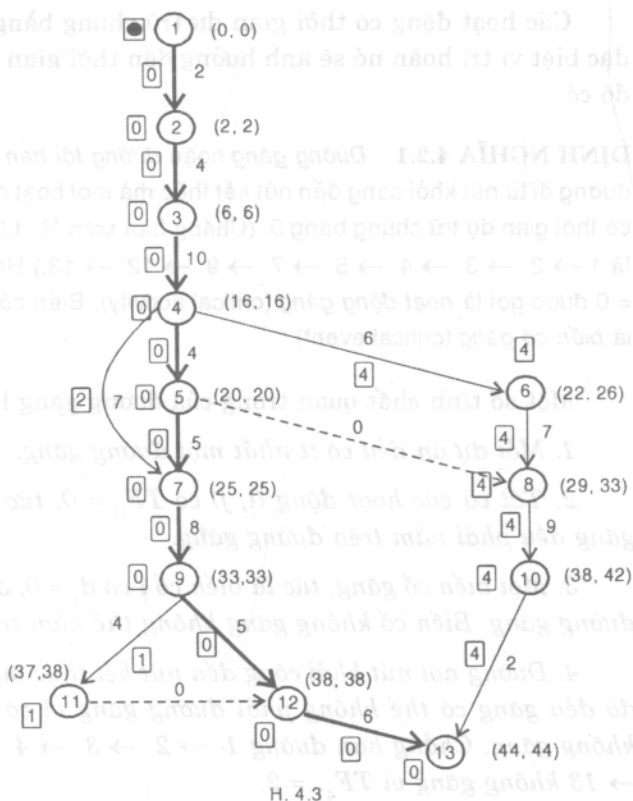
Thời gian dự trữ (slack hoặc float) của một biến cố là hiệu thời điểm muộn và thời điểm sớm của nó: $d_i = L_i - E_i$. Thời gian dự trữ (slack hoặc float) của hoạt động được chia làm hai loại. *Thời gian dự trữ chung* (total float hoặc total slack) của hoạt động (i, j) là

$$TF_{ij} = L_j - E_i - t_{ij}.$$

TF_{ij} chỉ thời gian có thể trì hoãn của hoạt động (i, j) mà không ảnh hưởng đến thời điểm kết thúc cả dự án. Vì nó bằng thời gian tối đa dành cho hoạt động (i, j) là $L_j - E_i$ trừ đi thời gian để thực hiện là t_{ij} . Thời gian dự trữ độc lập (free float hoặc free slack), kí hiệu là FF_{ij} , cũng là hiệu thời gian dành cho (i, j) và thời gian thực hiện là t_{ij} , nhưng với giả thiết là mọi hoạt động đều bắt đầu sớm nhất có thể, vậy

$$FF_{ij} = E_j - E_i - t_{ij}.$$

Trên sơ đồ mạng lưới thì d_i là hiệu hai số trong ngoặc ở nút i, thường được ghi bằng số trong ô vuông cạnh nút. Thời gian dự trữ chung của hoạt động TF_{ij} được ghi trong ô vuông cạnh mỗi cung. Còn thời gian dự trữ độc lập của hoạt động FF_{ij} ít quan trọng hơn, thường không ghi, xem H. 4.3.



H. 4.3

4.2.3 Đường găng (đường tới hạn)

Các hoạt động có thời gian dự trữ chung bằng 0 cần được chú ý đặc biệt vì trì hoãn nó sẽ ảnh hưởng đến thời gian kết thúc dự án. Từ đó có

ĐỊNH NGHĨA 4.2.1 Đường găng hoặc đường tới hạn (critical path) là một đường đi từ nút khởi công đến nút kết thúc mà mọi hoạt động trên đường đều có thời gian dự trữ chung bằng 0. (Chẳng hạn trên H. 4.3 có một đường găng là $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 7 \rightarrow 9 \rightarrow 12 \rightarrow 13$.) Hoạt động (i, j) có $TF_{ij} = 0$ được gọi là *hoạt động găng* (critical activity). Biến cố i có $d_i = 0$ được gọi là *biến cố găng* (critical event).

Một số tính chất quan trọng của đường găng là như sau.

1. Mỗi dự án đều có ít nhất một đường găng.
2. Tất cả các hoạt động (i, j) có $TF_{ij} = 0$, tức là mọi hoạt động găng đều phải nằm trên đường găng.
3. Mọi biến cố găng, tức là biến cố i có $d_i = 0$, đều phải nằm trên đường găng. Biến cố không găng không thể nằm trên đường găng.
4. Đường nối nút khởi công đến nút kết thúc mà mọi biến cố trên đó đều găng có thể không phải đường găng vì có thể có hoạt động không găng. Chẳng hạn đường $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 7 \rightarrow 9 \rightarrow 12 \rightarrow 13$ không găng vì $TF_{47} = 2$.
5. Đường găng là đường dài nhất trong các đường nối nút khởi công đến nút kết thúc.

Điều 5 này là rõ từ định nghĩa vì ở nút khởi công và kết thúc hai thời điểm sớm và muộn trùng nhau và thời gian hoàn thành dự án chính là hiệu thời gian ở hai nút (ở H. 4.3 là $44 - 0$). Đường găng là đường gồm các hoạt động không có dự trữ nên tổng chiều dài, tức là thời gian thực hiện, là toàn bộ thời gian thực hiện dự án (ở H.4.3 là 44), nên phải dài nhất. Trên H. 4.3 đường găng được tô đậm.

Một thí dụ dự án có nhiều đường găng là sơ đồ ở H. 4.3 nhưng với t_{46} thay từ 6 thành 10. Khi đó thời gian dự trữ của các hoạt động (6, 8), (8, 10) và (10, 13) và thời gian dự trữ của các biến cố 6, 8 và 10 đều thay từ 4 thành 0. Lúc này đường $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 6 \rightarrow 8 \rightarrow 10 \rightarrow 13$ là đường găng thứ hai.

Các chỉ tiêu thời gian của dự án ở H. 4.3 được ghi vào Bảng 4.1

biến cố	thời điểm sớm	thời điểm muộn	thời gian dự trữ	Hoạt động	thời gian dự trữ chung
1	0	0	0	(1, 2)	0
2	2	2	0	(2, 3)	0
3	6	6	0	(3, 4)	0
4	16	16	0	(4, 5)	0
5	20	20	0	(4, 6)	4
6	22	26	4	(4, 7)	2
7	25	25	0	(5, 7)	0
8	29	33	4	(6, 8)	4
9	33	33	0	(7, 9)	0
10	38	42	4	(8, 10)	4
11	37	38	1	(9, 11)	1
12	38	38	0	(9, 12)	0
13	44	44	0	(10, 13)	4
				(12, 13)	0

Bảng 4.1. Chỉ tiêu thời gian xây nhà

Ngoài các chỉ tiêu chính nói trên, khi cần các thông tin chi tiết hơn để điều hành dự án, người ta cũng đưa ra một số khái niệm về chỉ tiêu thời gian khác nữa như sau.

Thời điểm khởi công sớm (earliest start) của hoạt động (i, j) là thời điểm sớm của nút gốc: $ES_{ij} = E_i$.

Thời điểm hoàn thành sớm (earliest completion) của hoạt động (i, j) là $EC_{ij} = E_i + t_{ij}$.

Thời điểm khởi công muộn (latest start) của hoạt động (i, j) là $LS_{ij} = L_j - t_{ij}$.

Thời điểm hoàn thành muộn (latest completion) của hoạt động (i, j) là $LC_{ij} = L_j$ tức là thời điểm muộn của nút ngọn.

Nhận xét rằng $EC_{ij} \leq E_j$, $LS_{ij} \geq L_i$. Thật vậy, ta có

$$E_j = \max_k \{E_k + t_{kj}\} \geq E_i + t_{ij} = EC_{ij},$$

vì i cũng là một trong các nút k ngay trước j. Bất đẳng thức thứ hai tương tự.

Thời gian dự trữ của một đường đi (total float of a path) P từ nút khởi công đến nút kết thúc, kí hiệu TF_P , là thời gian có thể kéo dài thêm các hoạt động trên đường này mà không ảnh hưởng đến thời điểm hoàn thành công trình, tức là

$$TF_P = \sum t_{ij}^G - \sum t_{ij}^P := T^G - T^P,$$

ở đây $\sum t_{ij}^G = T^G$ là độ dài đường găng và $\sum t_{ij}^P = T^P$ là độ dài đường P, là tổng thời gian thực hiện các hoạt động trên đường P.

Hệ số găng (critical coefficient) biểu thị mức độ căng thẳng về thời gian của một đường P nối nút khởi công và kết thúc, không phải đường găng G, được định nghĩa là

$$K^P := \frac{T^P - T^{PG}}{T^G - T^{PG}},$$

ở đây T^{PG} là độ dài quãng đường (tức là một phần của đường) mà P trùng với G. Rõ ràng $0 < K^P < 1$ và K^P càng gần 1 thì thời hạn thực hiện các hoạt động không găng trong P càng phải chặt chẽ.

Hai định nghĩa trên đây của đường đi có thể mở rộng cho đường P có nút đầu và cuối trùng với nút trong đường găng, không cần là nút khởi công và kết thúc của cả dự án.

THÍ DỤ 4.2.1. Ở dự án trên H. 4.3, đường găng được tô đậm. Thời điểm hoàn thành sớm $EC_{68} = E_6 + t_{68} = 22 + 7 = 29 = E_8$, $EC_{10,13} = 38 + 2 = 40 < E_{13} = 44$. Thời điểm khởi công muộn $LS_{46} = L_6 - t_{46} = 26 - 6 = 20 > L_4 = 16$. Bây giờ giả sử P là đường đi $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 6 \rightarrow 8 \rightarrow 10 \rightarrow 13$ thì $T^P = \sum t_{ij}^P = 40$ nên thời gian dự trữ của P là $T^G - T^P = 44 - 40 = 4$. Hệ số găng là $K^P = \frac{40}{44} = \frac{10}{11}$ (không có quãng chung với đường găng). Gọi Q là đường $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 7 \rightarrow 9 \rightarrow 12 \rightarrow 13$ thì $T^Q = 42$, $K^Q = \frac{42 - 35}{44 - 35} = \frac{7}{9} < \frac{10}{11}$. Ta thấy mặc dù $T^Q > T^P$ nhưng thời hạn thực hiện các hoạt động không găng trong P lại chặt chẽ hơn hoạt động không găng (4, 7) duy nhất của Q. Nguyên nhân là (4, 7) là không găng duy nhất, nên mọi sự nới lỏng của Q đều dẫn cho hoạt động này.

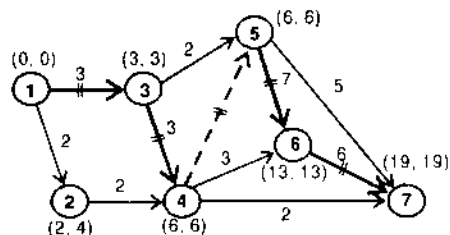
Chú ý rằng các dữ liệu thời gian quan trọng nhất là các chỉ tiêu có trong Bảng 4.1. Ở bảng này cũng cho thấy đường găng (đường gồm các hoạt động găng, tức là có thời gian dự trữ bằng 0).

4.2.4 Biểu đồ thời gian

Một cách truyền thống, bên cạnh sơ đồ lưới và bảng, để theo dõi điều hành thời gian cho dự án là dùng *biểu đồ thời gian* (time chart). Ta hãy xét cách vẽ và sử dụng biểu đồ thời gian qua một thí dụ.

THÍ DỤ 4.2.2. Xét dự án ở H. 4.4, và Bảng 4.2 tương ứng.

(Chú ý là hoạt động giả (4, 5) lại là hoạt động găng.)



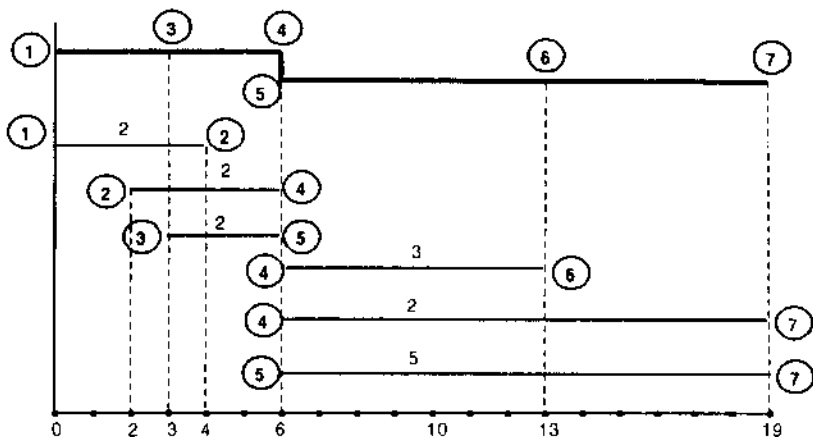
H. 4.4

biến cố	E_i	L_i	d_i	Hoạt động	TF_{ij}
1	0	0	0	(1, 2)	2
2	2	4	2	(1, 3)	0
3	3	3	0	(2, 4)	2
4	6	6	0	(3, 4)	0
5	6	6	0	(3, 5)	1
6	13	13	0	(4, 5)	0
7	19	19	0	(4, 6)	4
				(4, 7)	11
				(5, 6)	0
				(5, 7)	8
				(6, 7)	0

Bảng 4.2

Biểu đồ thời gian cho ở H. 4.5. Ở đây chỉ có trục hoành là thời gian. Cao độ không quan trọng. Ta biểu diễn các hoạt động găng phía trên. Độ dài (thời gian) là cố định, chặt chẽ cho các hoạt động găng. Hoạt động giả (4, 5) có độ dài 0 nên biểu diễn bởi đoạn thẳng đứng.

Mỗi hoạt động không găng biểu diễn ở độ cao khác nhau để nhìn rõ vì các hoạt động này có độ cơ động và được điều hành bằng biểu đồ thời gian.



H. 4.5

Biểu đồ được vẽ từ các E_i và L_i ở Bảng 4.2 (hoạt động gắng hay không gắng thì theo TF_{ij} bằng 0 hay khác 0). Các số không có vòng chỉ thời gian thực hiện của hoạt động. Chẳng hạn hoạt động (1, 2) thực hiện trong 2 đơn vị thời gian, được phép xê dịch trong khoảng thời gian dài 4 đơn vị (từ 0 đến 4). Xét sâu hơn thì sự xê dịch có tự do trong khoảng thời gian này không là phụ thuộc vào $FF_{ij} < TF_{ij}$ hay $FF_{ij} = TF_{ij}$. Tính bổ sung các FF_{ij} vào Bảng 4.2 ta sẽ thấy $FF_{12} = 0 < TF_{12} = 2$, còn tất cả các hoạt động khác có $FF_{ij} = TF_{ij}$. Nếu $FF_{ij} = TF_{ij}$ thì hoạt động (i, j) có thể cơ động tùy ý trong khoảng thời gian vẽ trên biểu đồ. Nếu $FF_{ij} < TF_{ij}$ thì hoạt động (i, j) chỉ được bắt đầu muộn hơn thời điểm khởi công sớm ES_{ij} một khoảng thời gian không quá FF_{ij} thì mới không ảnh hưởng đến các hoạt động ngay sau nó. Áp dụng cho H. 4.5, chỉ có (1, 2) có $FF_{12} = 0 < TF_{12} = 2$. Vì $FF_{12} = 0$, nên theo quy tắc trên, (1, 2) phải bắt đầu ngay lúc thời điểm sớm thì hoạt động ngay sau nó (duy nhất) là (2, 4) mới được xê dịch tùy ý trong khoảng thời gian từ 2 đến 6. Nếu (1, 2) thực hiện lùi lại khoảng từ 1 đến 3 chẳng hạn, thì ảnh hưởng đến hoạt động (2, 4). Mặc dù có $FF_{24} = TF_{24}$ nhưng lúc này nó chỉ còn được xê dịch thực hiện trong khoảng từ 3 đến 6.

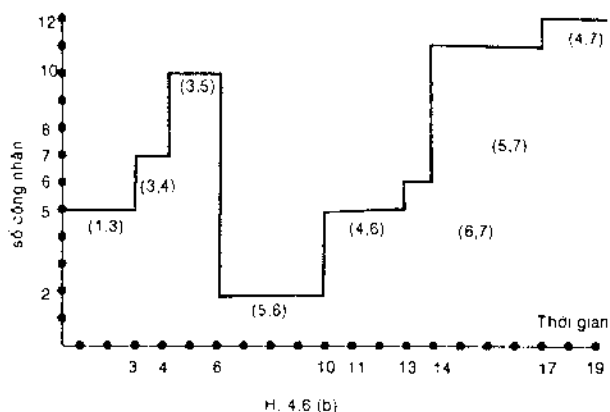
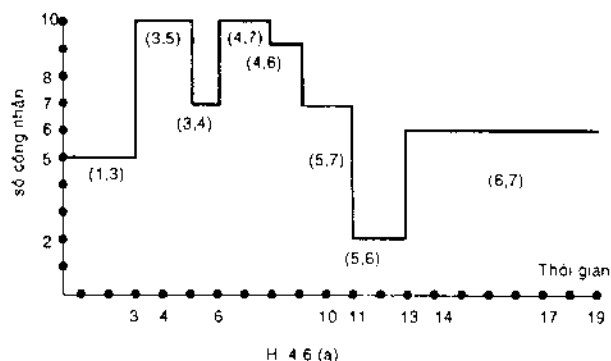
4.3 ĐIỀU KHIỂN NHÂN LỰC

Các hoạt động không gắng được phép xê dịch nhất định, nhất là khi $FF_{ij} = TF_{ij}$. Có thể sắp đặt chúng đáp ứng các yêu cầu khác nữa ngoài thời gian ra, chẳng hạn nhân lực, nguyên liệu, chi phí... Về mặt toán học xử lý yêu cầu loại nào cũng vậy. Ở đây ta nói theo ngôn ngữ nhân lực chẳng hạn.

THÍ DỤ 4.3.1. Giả sử nhân lực cho các hoạt động của dự án ở Thí dụ 4.2.2 đòi hỏi như sau

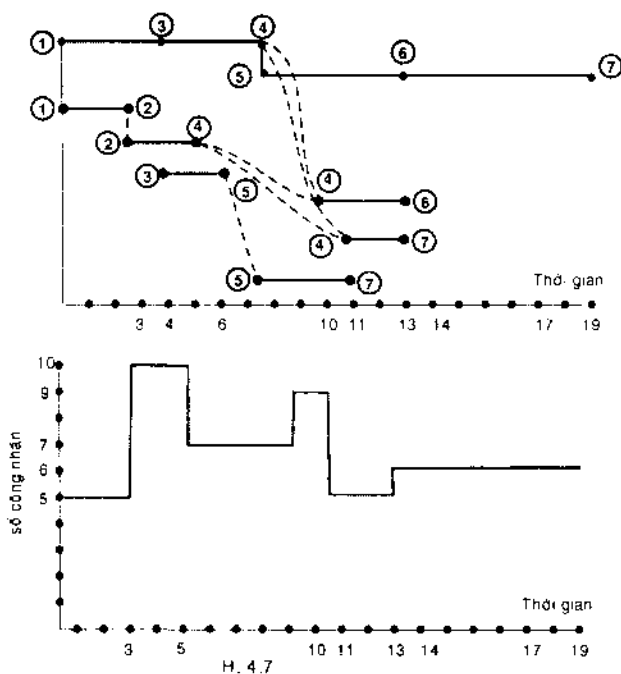
Hoạt động	Số công nhân	Hoạt động	Số công nhân
(1, 2)	0	(4, 6)	2
(1, 3)	5	(4, 7)	1
(2, 4)	0	(5, 6)	2
(3, 4)	7	(5, 7)	5
(3, 5)	3	(6, 7)	6

Chú ý rằng tại thời điểm hai hoạt động cùng tiến hành thì số nhân lực cần là tổng hai số công nhân. Vì vậy cần phải sắp xếp khéo các hoạt động không găng để đòi hỏi tổng công nhân của cả dự án ít (tạm coi là mỗi người biết làm mọi việc). Việc sắp xếp tối ưu là phức tạp, đến nay vẫn chưa có phương pháp toán học để giải quyết tổng quát. Ở đây ta sử dụng biểu đồ thời gian, biểu diễn thêm



nhan lực để sắp xếp theo trực quan. H. 4.6 (a) biểu diễn tổng công nhân cần ở mỗi thời điểm nếu tất cả các hoạt động không găng xếp vào lúc sớm nhất có thể, còn H. 4.6 (b) là tương ứng khi xếp vào lúc muộn nhất có thể. Hai biểu đồ này nên vẽ thẳng dưới H. 4.5 để rõ mốc thời gian của các hoạt động nhưng ở đây ta không vẽ lại H. 4.5 nữa. Sắp đặt sớm nhất ở hình (a) cho thấy ở mỗi thời điểm dự án cần nhiều nhất là 10 công nhân còn ở sắp đặt muộn nhất (b) là 12 công nhân. Ở hai

phương án này, số công nhân cần ở các thời điểm không đều. Theo trực quan ta chỉnh lại từ (a) như sau: chuyển hoạt động (4, 6) đến thời điểm muộn nhất có thể, chuyển (4, 7) đến ngay sau khi (5, 7) kết thúc. Kết quả được vẽ lại ở biểu đồ H. 4.7. (Chú ý là hoạt động (1, 2) và (2, 4) không cần công nhân nên không cần vẽ.)



4.4 HOÀN THÀNH SỚM DỰ ÁN

Trên đây đã xét thời điểm hoàn thành dự án là cố định và xác định các đường găng, phải thực hiện chặt chẽ để dự án hoàn thành đúng thời gian quy định. Nếu muốn giảm thời gian hoàn thành dự án thì làm thế nào? Ta cũng sử dụng đường găng, nhưng phải dựa vào kỹ thuật và công nghệ, chứ không phải quản lý chỉ bằng toán học được nữa. Cụ thể là phải dùng công nghệ mới, tăng vật tư, nhân công... để có thời gian thực hiện các hoạt động ngắn hơn. Nhưng tập trung vào hoạt động nào? Rõ ràng là vào các hoạt động găng. Cụ thể là nếu ta quan tâm đến hạn chế tăng chi phí thì với $(i, j) \in G$, tìm số gia chi phí ΔC_{ij} khi đạt được rút ngắn thời gian thực hiện hoạt động là Δt_{ij}

(tìm bằng thực tế công nghệ, không phải thuần túy toán học). Khi đó sẽ chọn cách tăng chi phí để giảm thời gian sao cho đạt $\min \frac{\Delta C_{ij}}{\Delta t_{ij}}$. Giá sử cực tiểu là $\frac{\Delta C_{ij}^0}{\Delta t_{ij}^0}$. Khi đó độ dài đường găng mới, tức là thời gian hoàn thành dự án mới, là

$$\tilde{T}^G = T^G - \sum \Delta t_{ij}^0,$$

ở đây tổng lấy trên mọi hoạt động găng.

4.5 DỰ ÁN CÓ TÍNH NGẪU NHIÊN

Trong các mục trên ta đã coi thời gian thực hiện các hoạt động t_{ij} là xác định hoàn toàn từ đầu, khi lập sơ đồ mạng lưới. Do đó ta có *mô hình tất định* (deterministic model). Trong thực tế, nhiều yếu tố bất định phải được tính đến, do đó thời gian thực hiện hoạt động (i, j) là một biến ngẫu nhiên (random variable), mà ta chỉ xác định được phân bố xác suất (probability distribution) qua kinh nghiệm và số liệu thống kê. Từ đó dẫn đến phải sử dụng *mô hình ngẫu nhiên* hoặc gọi khác là *mô hình xác suất* (probabilistic model). Việc tính toán các chỉ tiêu để điều hành dự án có hai nhiệm vụ chính. Một là tính *kỳ vọng* (mean hoặc expected value) của các đại lượng cần tính, chẳng hạn thời gian thực hiện hoạt động (activity time), thời gian hoàn thành dự án (project time), và phương sai (variance) của các đại lượng này. Hai là tính xác suất của biến cố nào đó, chẳng hạn biến cố là dự án được hoàn thành trước thời điểm T .

Thời gian thực hiện mỗi hoạt động, thường gọi tắt là thời gian hoạt động, trong mô hình ngẫu nhiên thường được giả thiết là xác định được ba yếu tố sau. *Thời gian lạc quan* (optimistic time) kí hiệu là a , là thời gian cần để làm xong khi hoạt động được thực hiện thuận

(tìm bằng thực tế công nghệ, không phải thuần túy toán học). Khi đó sẽ chọn cách tăng chi phí để giảm thời gian sao cho đạt $\min \frac{\Delta C_{ij}}{\Delta t_{ij}}$. Giả sử cực tiểu là $\frac{\Delta C_{ij}^0}{\Delta t_{ij}^0}$. Khi đó độ dài đường găng mới, tức là thời gian hoàn thành dự án mới, là

$$\bar{T}^G = T^G - \sum \Delta t_{ij}^0,$$

ở đây tổng lấy trên mọi hoạt động găng.

4.5 DỰ ÁN CÓ TÍNH NGẪU NHIÊN

Trong các mục trên ta đã coi thời gian thực hiện các hoạt động t_{ij} là xác định hoàn toàn từ đầu, khi lập sơ đồ mạng lưới. Do đó ta có *mô hình tất định* (deterministic model). Trong thực tế, nhiều yếu tố bất định phải được tính đến, do đó thời gian thực hiện hoạt động (i, j) là một biến ngẫu nhiên (random variable), mà ta chỉ xác định được phân bố xác suất (probability distribution) qua kinh nghiệm và số liệu thống kê. Từ đó dẫn đến phải sử dụng *mô hình ngẫu nhiên* hoặc gọi khác là *mô hình xác suất* (probabilistic model). Việc tính toán các chỉ tiêu để điều hành dự án có hai nhiệm vụ chính. Một là tính *kì vọng* (mean hoặc expected value) của các đại lượng cần tính, chẳng hạn thời gian thực hiện hoạt động (activity time), thời gian hoàn thành dự án (project time), và phương sai (variance) của các đại lượng này. Hai là tính xác suất của biến cố nào đó, chẳng hạn biến cố là dự án được hoàn thành trước thời điểm T .

Thời gian thực hiện mỗi hoạt động, thường gọi tắt là thời gian hoạt động, trong mô hình ngẫu nhiên thường được giả thiết là xác định được ba yếu tố sau. *Thời gian lạc quan* (optimistic time) kí hiệu là a , là thời gian cần để làm xong khi hoạt động được thực hiện thuận

lợi nhất. Thời gian này rất khó đạt được. Theo lí thuyết thống kê, thì đây thực chất là cận dưới của các thời gian xong việc có thể xảy ra, tức là *cận dưới* (lower bound) của phân bố xác suất. *Thời gian bi quan* (pessimistic time), kí hiệu là b , là thời gian cần để xong hoạt động khi tiến hành gặp trục trặc nhất, tức là *cận trên* (upper bound) của phân bố xác suất. *Thời gian hợp lí nhất* (most likely time), kí hiệu là m , là thời gian hiện thực nhất, tức là có xác suất lớn nhất (đỉnh cao nhất của hàm mật độ). Ba lượng trên chưa đủ để xác định phân bố xác suất của thời gian hoạt động, do đó chưa đủ để xác định kì vọng t_e tức là giá trị trung bình theo xác suất, và phương sai σ^2 đặc trưng cho độ lệch khỏi t_e của thời gian hoạt động. Mô hình cần hai giả thiết phù hợp thực tế sau đây.

Giả thiết 1. $b-a$, tức là độ dài khoảng mà thời gian hoạt động có thể lấy, bằng 6 lần *độ lệch chuẩn* (standard deviation), tức là ta có phương sai

$$\sigma^2 = \left[\frac{1}{6}(b-a)\right]^2. \quad (4.1)$$

Điều này đúng cho nhiều biến ngẫu nhiên hay gặp.

Giả thiết 2. Phân bố xác suất của mỗi thời gian hoạt động đều là *phân bố beta* (beta distribution).

Ta hãy nhắc lại vài kiến thức xác suất. Mỗi biến ngẫu nhiên x có hai hàm quan trọng nhất. *Hàm mật độ xác suất* (probability density function) $f(x)$, $a \leq x \leq b$, và *hàm phân bố tích lũy* (cumulative distribution function) $F(X)$, gọi khác là *hàm phân bố*. Ở đây giả thiết là x chỉ lấy giá trị trong $[a, b]$. Ta có các quan hệ sau

$$\int_a^b f(x)dx = 1,$$

$$F(X) = P\{x \leq X\} = \int_a^x f(x)dx,$$

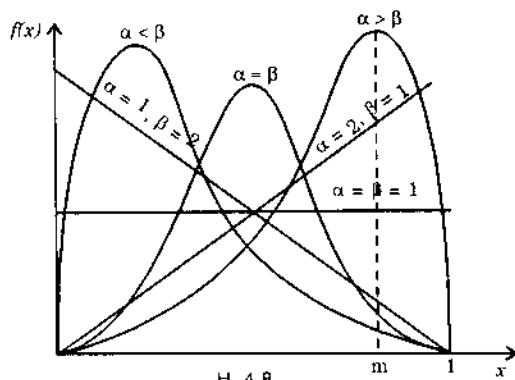
$$x_e = \int_a^b x f(x) dx,$$

$$\sigma^2 = \int_a^b (x - x_e)^2 f(x) dx,$$

ở đây x_e là kì vọng và σ^2 là phương sai của biến ngẫu nhiên x , $P \{ \dots \}$ là xác suất của biến cố $\{ \dots \}$. Mỗi một trong hàm mật độ hoặc hàm phân phối đặc trưng hoàn toàn cho biến ngẫu nhiên. kì vọng và phương sai là các đại lượng quan trọng. Ta cũng nói là hàm mật độ (hoặc hàm phân bố), xác định hoàn toàn phân bố xác suất. *Phân bố beta* (beta distribution) là một trong các phân bố xác suất phổ biến nhất, xác định bởi hàm mật độ sau, nếu $0 \leq x \leq 1$,

$$f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \times \Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1} = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}, \quad (4.2)$$

ở đây α và β là tham số, $\Gamma(\cdot)$ là *hàm đặc biệt gamma* và $B(\cdot, \cdot)$ là *hàm đặc biệt beta*, được định nghĩa bằng tích phân phụ thuộc tham số (xem, chẳng hạn [Khánh 2000]) như sau



H. 4.8

$$\Gamma(\alpha) := \int_0^{+\infty} t^{\alpha-1} e^{-t} dt,$$

$$B(\alpha, \beta) := \int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt.$$

Nếu lấy giá trị trên $[a, b]$ và có phân bố beta thì hàm mật độ nhận được từ (4.2) bằng đổi biến $y = a$

$+ (b - a)x$. Chẳng hạn, hàm mật độ của phân bố beta có dạng như H. 4.8 với $\alpha \geq 1$, $\beta \geq 1$ và $a = 0$, $b = 1$.

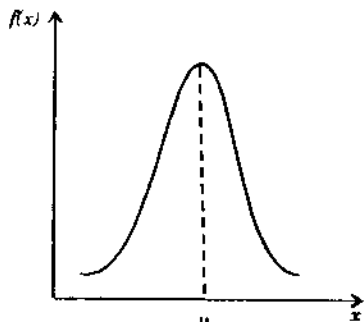
Phân bố chuẩn (normal distribution) là phân bố xác suất phổ biến nhất, định nghĩa bởi hàm mật độ sau

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < +\infty,$$

ở đây tham số μ chính là kì vọng và σ^2 chính là phương sai của biến ngẫu nhiên x có phân bố chuẩn. Khi đó biến

$z = \frac{x - \mu}{\sigma}$ có phân bố là phân bố chuẩn với kì vọng 0, phương sai là 1.

1. Hàm mật độ của phân bố chuẩn có dạng ở H. 4.9.



H. 4.9

Các biến ngẫu nhiên $x_1, \dots, x_n$ được gọi là *độc lập* (independent) nếu

$$P\{x_1 \leq X_1, \dots, x_n \leq X_n\} = P\{x_1 \leq X_1\} \dots P\{x_n \leq X_n\}.$$

ĐỊNH LÝ GIỚI HẠN TRUNG TÂM (central - limit theorem) nói rằng với các điều kiện khá nhẹ, tổng các biến ngẫu nhiên độc lập luôn có phân bố chuẩn (không phụ thuộc vào phân bố của từng biến ngẫu nhiên).

Trở lại mô hình ngẫu nhiên điều hành dự án. Để tính kì vọng t_e của thời gian hoạt động, người ta giả thiết là điểm giữa $\frac{a+b}{2}$ chiếm tỉ trọng bằng nửa điểm hợp lí nhất m . Khi đó

$$t_e = \frac{1}{3} \left[2m + \frac{1}{2}(a + b) \right]. \quad (4.3)$$

THÍ DỤ 4.5.1. Giả sử dự án xây nhà ở H. 4.1 bây giờ có các thời gian hoạt động là ngẫu nhiên có phân bố beta thỏa hai giả thiết trên và xác định được ba mốc thời gian lạc quan, bi quan và hợp lí nhất theo Bảng 4.3. Khi đó phương sai và kì vọng của các thời gian hoạt động, tính theo công thức (4, 1) và (4, 3) được

ghi ở hai cột cuối.

Hoạt động	thời gian lạc quan a	thời gian hợp lý nhất m	thời gian bi quan b	kì vọng le	phương sai σ^2
(1, 2)	1	2	3	2	$\frac{1}{9}$
(2, 3)	2	$3\frac{1}{2}$	8	4	1
(3, 4)	6	9	18	10	4
(4, 5)	1	$4\frac{1}{2}$	5	4	$\frac{4}{9}$
(4, 6)	4	$5\frac{1}{2}$	10	6	1
(4, 7)	3	$7\frac{1}{2}$	9	7	1
(5, 7)	4	4	10	5	1
(6, 8)	5	$6\frac{1}{2}$	11	7	1
(7, 9)	3	9	9	8	1
(8, 10)	5	8	17	9	4
(9, 11)	4	4	4	4	0
(9, 12)	1	$5\frac{1}{2}$	7	5	1
(10, 13)	1	2	3	2	$\frac{1}{9}$
(12, 13)	5	$5\frac{1}{2}$	9	6	$\frac{4}{9}$

Bảng 4.3

Nhận xét rằng cột kì vọng ở Bảng 4, 3, do thí dụ được xây dựng đặc biệt, trùng hoàn toàn với các thời gian hoạt động trong mô hình tất định đã xét ở H. 4.1. Do đó đường găng xây dựng trên các thời gian hoạt động kì vọng trùng với đường găng của mô hình tất định ở H. 4.2 và thời gian của đường găng này là 44.

Tuy nhiên để xác định kì vọng và phương sai của thời gian dự án, ta cần thêm hai giả thiết sau.

Giả thiết 3. Các thời gian hoạt động là các biến ngẫu nhiên độc lập.

Giả thiết 4. Đường găng xây dựng trên các thời gian hoạt động kì vọng, luôn đòi hỏi thời gian (hoàn thành mọi hoạt động trên nó) lớn hơn các đường khác.

Tính thật chi li trong các thí dụ cụ thể thì hai giả thiết 3 và 4 có thể không đúng. Chẳng hạn ở Thí dụ 4.5.1, nếu xảy ra thời gian bị quan ở mọi hoạt động thì đường găng đã tính dài 69 (ngày). Còn đường $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 6 \rightarrow 8 \rightarrow 10 \rightarrow 13$ có thời gian bị quan là 70. Tuy vậy người ta vẫn chấp nhận các giả thiết xấp xỉ này. Khi đó, vì kì vọng và phương sai của tổng các biến ngẫu nhiên là tổng các kì vọng và phương sai nên ta có: Kì vọng và phương sai của thời gian dự án là tổng các kì vọng và phương sai của các thời gian hoạt động trên đường găng (xây dựng theo các kì vọng). Đến đây ta nhận xét rằng một trong các cách áp dụng thực tế là dùng các kì vọng của các biến, rồi áp dụng mọi tính toán và lý luận ở các mục trước vào các kì vọng, thay cho các biến bất định.

Ở Thí dụ 4.5.1 kì vọng và phương sai của thời gian dự án là 44 và 9, vì đường găng là $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 7 \rightarrow 9 \rightarrow 12 \rightarrow 13$.

Bây giờ ta xét vấn đề quan trọng là tính xác suất để dự án hoàn thành trước một thời hạn bắt buộc (deadline). Theo định lý giới hạn trung tâm, thời gian dự án là biến ngẫu nhiên có phân bố chuẩn. Do đó ta tính được xác suất $P(x \leq X)$, thường được tính sẵn để tra theo bảng. Chẳng hạn Bảng A1 ở cuối sách cho biết $P(x \leq x_e + \sigma K_\alpha)$, ở đây σ là độ lệch chuẩn. Do đó K_α là số đơn vị lệch chuẩn.

Thí dụ, hãy tính xác suất để thời gian xây nhà ở Thí dụ 4.5.1 không quá 47 ngày. Ta thấy $47 = 44 + 3.1 = x_e + \sigma.K_\alpha$ nên $K_\alpha = 1$. Theo bảng thì $P\{x \geq 47\} = 0,1587$. Do đó xác suất cần tìm là $\approx 1 - 0,1587 \approx 0,84$.

Phương pháp điều hành dự án có tính ngẫu nhiên trên đây thường

được gọi là *phương pháp ba ước lượng PERT* (PERT three estimate method).

Nếu cần tính các yếu tố thời gian ở các biến cố trung gian (không chỉ thời gian hoàn thành dự án, tức là với biến cố cuối) thì ta lý luận như sau. Trước hết tính kì vọng và phương sai của thời điểm sớm μ_i của biến cố i . Nếu chỉ có một đường từ khởi công đến i thì, do các hoạt động là độc lập kì vọng của μ_i , kí hiệu là $E(\mu_i)$, bằng tổng các kì vọng t_e của thời gian các hoạt động dẫn đến i . Cũng vậy phương sai $\text{Var}(\mu_i)$ bằng tổng các phương sai σ^2 của các hoạt động dẫn đến i . Khi có nhiều đường dẫn đến i thì người ta coi xấp xỉ (để đơn giản) $E(\mu_i)$ và $\text{Var}(\mu_i)$ là tổng các t_e và σ^2 của các hoạt động theo đường đến i có tổng $E(\mu_i)$ dài nhất. Nếu có nhiều đường với cùng $E(\mu_i)$ thì $\text{Var}(\mu_i)$ quy ước lấy lượng của đường có tổng các σ^2 dài nhất.

Bây giờ hãy tính xác suất để biến cố i xảy ra trước thời gian bắt buộc T_i cho trước. Theo định lý giới hạn trung tâm μ_i , tuân theo phân bố chuẩn, ta chỉ việc tra bảng các xác suất ứng với phân bố chuẩn để tính $P\{\mu_i \leq T_i\}$. Cụ thể, để tra bảng, quy về trường hợp đại lượng z có phân bố chuẩn với kì vọng 0 và phương sai 1 như sau:

$$P(\mu_i \leq T_i) = P\left\{\frac{\mu_i - E(\mu_i)}{\sqrt{\text{Var}(\mu_i)}} \leq \frac{T_i - E(\mu_i)}{\sqrt{\text{Var}(\mu_i)}}\right\} = P(z \leq K_i),$$

ở đây $K_i = \frac{T_i - E(\mu_i)}{\sqrt{\text{Var}(\mu_i)}}$ đã biết.

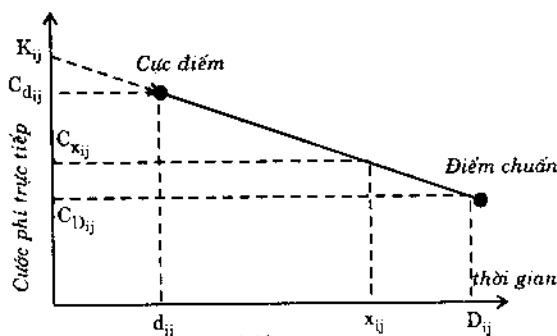
4.6 DỰ ÁN CÓ THỎA HIỆP THỜI GIAN - CƯỚC PHÍ

Trong các mục trước của Chương 4 ta trình bày về các dự án có yêu cầu chủ yếu là điều hành thời gian. Theo ngôn ngữ ban đầu thì đây là phương pháp PERT, các thời gian ở đây có thể xét như các biến

tất định hoặc ngẫu nhiên. Còn phương pháp đường găng CPM thì đặt ngang nhau thời gian và cước phí. Mục tiêu chính của CPM là chọn cách thỏa hiệp thời gian - cước phí ở mỗi hoạt động thế nào để toàn bộ dự án hoàn thành trước một thời hạn bắt buộc và làm cực tiểu cước phí. Bây giờ ta trình bày phương pháp CPM để giải bài toán này.

Trước hết phải giả thiết là biết quan hệ giữa cước phí và thời gian thực hiện mỗi hoạt động (theo ngôn ngữ hình học) tức là biết *đường cong thời gian - cước phí* (time - cost curve) của mỗi hoạt động. Trong mô hình toán học (xấp xỉ thô tình trạng thực tế) người ta giả thiết quan hệ thời gian cước phí là tuyến tính. Do đó chỉ cần biết hai điểm. Người ta chọn hai điểm nút như sau.

Điểm chuẩn (normal point) có tọa độ là thời gian và cước phí của hoạt động khi nó được tiến hành trong điều kiện bình thường, tức là chuẩn, không có cước phí bổ sung tăng cường (như làm ngoài giờ, tăng thiết bị, nhân lực...). *Cực điểm* (crash point) là điểm ứng với thời gian và cước phí khi đầu tư hết mức để thời gian thực hiện hoạt động ngắn nhất có thể. Mọi điểm trung gian giữa điểm chuẩn và cực điểm, tức là mọi cách *thỏa hiệp thời gian cước phí* (time - cost trade - off) đều coi là chấp nhận được, xem H. 4.10.



H. 4.10

Đường cong thời gian - cước phí của hoạt động (i, j)

Các kí hiệu trên H. 4.10 rõ ràng là như sau. D_{ij} và d_{ij} là thời gian chuẩn và thời gian cực điểm. $C_{D_{ij}}$ và $C_{d_{ij}}$ là *cước phí chuẩn* (normal cost) và *cước phí cực điểm* (crash cost), đều của hoạt động (i, j) . Gọi x_{ij} (thời gian thực hiện hoạt động (i, j)) là *biến quyết định* (decision variable) của bài toán mà ta cần tính. Gọi S_{ij} là độ xiên, tức là hệ số góc đường thẳng biểu thị đường cong thời gian - cước phí, tức là

$$S_{ij} = \frac{C_{D_{ij}} - C_{d_{ij}}}{D_{ij} - d_{ij}}.$$

Gọi K_{ij} là tung độ điểm đường thẳng cắt trục tung. Khi đó cước phí của hoạt động (i, j) ứng với thời gian hoạt động x_{ij} rõ ràng là

$$C_{x_{ij}} = K_{ij} + S_{ij} x_{ij}.$$

Tổng cước phí dự án là $C = \sum_{(i, j)} (K_{ij} + S_{ij} x_{ij})$, ở đây tổng lấy theo mọi

hoạt động. Vậy, dạng toán học của bài toán chọn cách thỏa hiệp thời gian - cước phí làm cực tiểu cước phí dự án là

Bài toán: Chọn các x_{ij} để thời gian dự án không quá thời hạn bắt buộc T cho trước và làm cực tiểu cước phí dự án C .

Nhận xét rằng các yếu tố của bài toán đều là tuyến tính, ta có gắng đưa về quy hoạch tuyến tính như sau.

Đưa vào các biến bổ sung y_k là thời điểm sớm E_k của biến cố k . Khi đó quan hệ giữa các biến theo Mục 4.2 là

$$E_k = \max_j \{E_j + t_{jk}\}, \quad (4.4)$$

ở đây max lấy theo các biến cố j ngay trước k , tức là có hoạt động nối (j, k) . kí hiệu $y_k = E_k$, $t_{jk} = x_{jk}$ và viết lại (4.4) ta được

$$y_j + x_{jk} - y_k \leq 0$$

(số ràng buộc là số các biến cố j ngay trước k). Gọi 1 là nút xuất phát

và n là nút kết thúc dự án thì

$$y_1 = 0, y_n \leq T.$$

Ở mục tiêu thì các K_{ij} là hằng. Tóm lại ta được quy hoạch tuyến tính

$$\min \sum_{(i,j)} S_{ij} x_{ij},$$

$$d_{ij} \leq x_{ij} \leq D_{ij},$$

$$y_i + x_{ij} - y_j \leq 0,$$

$$y_1 = 0, y_n \leq T.$$

Để đưa về quy hoạch tuyến tính dạng chuẩn ta làm như sau. Đổi biến $x_{ij} = d_{ij} + x'_{ij}$ thì ràng buộc $x_{ij} \geq d_{ij}$ trở thành ràng buộc dấu $x'_{ij} \geq 0$. Thêm ràng buộc hình thức $y_i \geq 0, \forall i$. Ràng buộc này tự nhiên thỏa do $y_1 = 0$ và $y_j \geq y_i + d_{ij} + x'_{ij}$.

Trường hợp không có thời hạn bắt buộc T cho trước, tức là cần tìm thỏa hiệp tốt nhất giữa tổng cước phí và tổng thời gian dự án, người ta coi T là tham số và giải quy hoạch tuyến tính tham số để được nghiệm tối ưu như hàm của T .

4.7 KIỂM TRA HIỆU CHỈNH DỰ ÁN

Sau khi dùng phương pháp điều hành dự án PERT – CPM xác định được sơ đồ mạng lưới, các biểu đồ và bảng tính các chỉ tiêu và dự án đang được tiến hành, người quản lý luôn phải theo dõi, kiểm tra. Điều kiện lao động thực tế có thể nhiều bất ngờ. Khi cần thiết có thể phải dùng phương pháp PERT – CPM lại, dựa trên các dữ liệu mới, để tính toán cho phần còn lại của dự án. Sau đó điều hành dự án theo các biểu đồ và bảng tính mới.

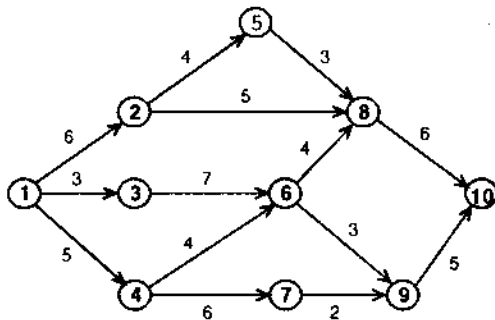
BÀI TẬP CHƯƠNG 4

4.1 Hãy xây dựng sơ đồ mạng lưới gồm các hoạt động A, B, ..., L sao cho các quan hệ sau đây được thỏa.

- (a) A, B và C là các hoạt động đầu tiên và có thể đồng thời.
- (b) A và B trước D. (c) B trước E, F và H. (d) F và C trước G.
- (e) E và H trước I và J. (g) C, D, F và J trước K. (h) K trước L.
- (i) I, G và L là các hoạt động cuối cùng của dự án.

4.2 Nếu thêm các quan hệ sau đây vào dự án ở Bài 4.1 thì sơ đồ phải chỉnh thể nào

- (a) A và B trước G. (b) D trước G. (c) C trước D.



H. 4.11

4.3 Xét sơ đồ mạng lưới ở H. 4.11, ở đây số ở cạnh cung chỉ thời gian hoạt động. Tìm thời điểm sớm E_i , thời điểm muộn L_i và thời gian dự trữ của mỗi biến cố, và thời gian dự trữ chung TF_{ij} của mỗi hoạt động. Đồng thời xác định đường găng.

4.4 Xét dự án ở H. 4.4. Tìm đường găng nếu

- (a) thời gian hoạt động t_{01} mới là 4;
- (b) thời gian hoạt động L_{36} mới là 15.

4.5 Xét dự án ở Thí dụ 4.2.2. Tính các thời gian hoàn thành sớm EC_{ij} , thời gian khởi công muộn LS_{ij} và thời gian dự trữ độc lập FF_{ij} của mọi hoạt động.

4.6 Xét dự án có các thời gian hoạt động cho ở bảng sau

Hoạt động	(1, 2)	(2, 3)	(3, 4)	(4, 5)	(5, 9)	(2, 6)	(6, 7)	(7, 8)	(8, 9)	(4, 7)
thời gian	5	5	5	5	5	6	2	5	6	0

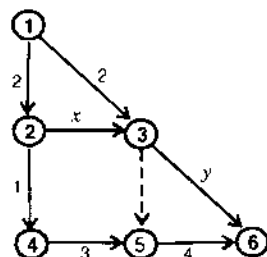
(a) Xây dựng sơ đồ mạng lưới.

(b) Tìm thời điểm sớm E_i , thời điểm muộn L_i và thời gian dự trữ của các biến cố cũng như thời gian dự trữ của các hoạt động. Xác định đường găng.

(c) Nếu bỏ đi quan hệ trước sau xác định bởi hoạt động giả (4, 7), thì thời gian hoàn thành dự án giảm được bao nhiêu ?

4.7 Xét sơ đồ mạng lưới ở H. 4.12, ở đây x và y là các thời gian hoạt động chưa biết. Xác định x và y để đường $1 \rightarrow 2 \rightarrow 3 \rightarrow 6$ là đường găng.

4.8 Các hoạt động của một dự án được cho ở bảng sau



H. 4.12

Hoạt động	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
Hoạt động ngay trước		A	A	B	D	D	D	G	H	D	J	D	L	D	N	O	E	Q	M
Thời gian thực hiện	21	14	14	3	70	70	14	1	1	1	1	7	7	7	14	1	1	1	1

Hãy xây dựng sơ đồ mạng lưới và tính đường găng.

4.9 Giả sử ở thí dụ 4.3.1, các hoạt động (1, 2) và (2, 4) cần 8 và 2 công nhân. Hãy chỉ ra các thay đổi trên H. 4.6 trong mỗi trường hợp sau

(a) cả hai hoạt động được sắp đặt sớm nhất có thể;

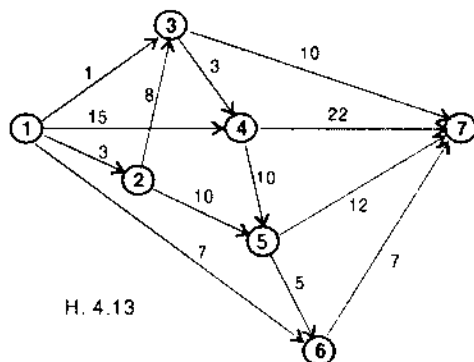
(b) cả hai hoạt động được sắp muộn nhất có thể.

172 CHƯƠNG 4 ĐIỀU HÀNH DỰ ÁN BẰNG PHƯƠNG PHÁP PERT-CPM

4.10 Giả sử trong dự án ở H. 4.13 nhu cầu công nhân được cho bằng bảng dưới

hoạt động	số công nhân	hoạt động	số công nhân	hoạt động	số công nhân
(1, 2)	1	(2, 5)	4	(5, 6)	2
(1, 3)	2	(3, 4)	10	(5, 7)	5
(1, 4)	5	(3, 7)	9	(6, 7)	3
(1, 6)	3	(4, 5)	8		
(2, 3)	1	(4, 7)	7		

Hãy tìm số công nhân ít nhất (như hàm của thời gian dự án) trong quá trình điều hành dự án.



H. 4.13

4.11 Giả sử khi dùng phương pháp ba ước lượng PERT ta có ba ước lượng của một hoạt động là: ước lượng lạc quan bằng 30 ngày, ước lượng hợp lý xảy ra nhất bằng 36 ngày và ước lượng bi quan là 48 ngày. Hãy tính kì vọng và phương sai của thời gian hoạt động này.

4.12 Xét dự án ở Thí dụ 4.2.2, tức là ở H. 4.4, nhưng trong trường hợp ngẫu nhiên với a , b , m của các hoạt động cho ở bảng dưới đây. Hãy tính các kì vọng t_e và phương sai σ^2 của các hoạt động. Đồng thời cho biết thời gian bắt buộc của các biến cố T_i theo bảng dưới đây. Hãy tính kì vọng $E(\mu_i)$, phương sai $Var(\mu_i)$ và xác suất $P\{\mu_i \leq T_i\}$ để mỗi biến cố hoàn thành trước T_i .

hoạt động	a	b	m	hoạt động	a	b	m	biến cố	T _i
(1, 2)	1	3	2	(4, 6)	1	7	2, 5	1	4
(1, 2)	2	8	2	(4, 7)	1	3	2	2	2
(2, 4)	1	3	2	(5, 6)	6	8	7	3	5
(3, 4)	1	11	1,5	(5, 7)	3	11	4	4	6
(3, 5)	0,5	7,5	1	(6, 7)	4	8	6	5	13
								6	19

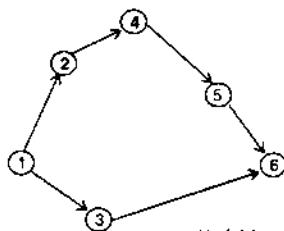
4.13 Xét dự án có tính ngẫu nhiên có sơ đồ mạng lưới dưới đây. Biết ba ước lượng của các hoạt động theo bảng đã cho.

(a) Tính t_e và σ^2 của các hoạt động.

(b) Xác định đường găng theo các t_e .

(c) Tính xác suất hoàn thành dự án trước thời gian 57 (ngày).

(d) Nếu mọi thời gian hoạt động đều bằng thời gian bị quan thì đường nào là đường găng. Tính xác suất để đường găng này hoàn thành trước 57 (ngày).



H. 4.14

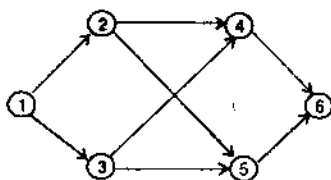
hoạt động	a	m	b
(1, 2)	12	12	12
(1, 3)	15	21	39
(2, 4)	12	15	18
(3, 6)	18	27	36
(4, 5)	12	18	24
(5, 6)	2	5	14

4.14 Dùng phương pháp thỏa hiệp thời gian-cước phí CPM cho dự án ở H. 4.15 tìm cực tiểu cước phí hoàn thành dự án trong 15 (tuần). Điểm chuẩn và cực điểm của các biến cố được cho trong bảng (theo tuần) cùng cước phí chuẩn và cực điểm cụ thể là

Hoạt động	D_{ij}	d_{ij}	$C_{D_{ij}}$	$C_{d_{ij}}$
(1, 2)	5	3	20	30
(1, 3)	3	2	10	20
(2, 4)	4	2	16	24
(2, 5)	6	3	25	43
(3, 4)	5	4	22	30
(3, 5)	7	4	30	48
(4, 6)	9	5	25	45
(5, 6)	8	6	30	44

(a) Lập quy hoạch tuyến tính cho bài toán này.

(b) Giải quy hoạch tuyến tính.



H. 4.15

CHƯƠNG 5

BÀI TOÁN VẬN TẢI

Trong chương này ta sẽ xét một bài toán dòng trên mạng đặc biệt là bài toán vận tải. Đây là bài toán quan trọng nhất trong các bài toán quy hoạch tuyến tính. Ở Mi người ta tổng kết rằng 85% các bài toán quy hoạch tuyến tính gặp trong ứng dụng là bài toán vận tải hoặc mở rộng của nó. Thực ra, ở Chương 3 ta đã thấy bản thân bài toán dòng trên mạng tổng quát có thể được giải thích như một bài toán chuyển vận, sao cho cước phí nhỏ nhất. Thuật ngữ *bài toán vận tải* (transportation problem) là dành cho trường hợp riêng được nêu ở Thí dụ 1.1.2.

5.1 CÁC KHÁI NIỆM CƠ BẢN

5.1.1 Định nghĩa

Ta đã định nghĩa bài toán vận tải (gồm hai trường hợp mô hình vận tải đóng và mô hình vận tải mở) như là một trường hợp riêng của quy hoạch tuyến tính chính tắc ở Thí dụ 1.1.2. Với ngôn ngữ dòng trên mạng, bài toán vận tải là bài toán dòng đặc biệt khi tập các nút được chia làm hai phần rời nhau ($N = S \cup D$, $S \cap D = \emptyset$) sao cho mọi cung đều có đuôi ở S và đầu ở D ; các nút thuộc S là các *nút nguồn* (source) hoặc cũng gọi là *nút cung* (supply node); các nút thuộc D là

đích (destination), cũng gọi là nút cầu (demand node). Đồ thị trong trường hợp này sẽ gọi là đồ thị hai phần (bipartite graph). Ở bài toán vận tải, đồ thị hai phần còn có thêm giả thiết là mỗi nguồn đều có cung nối với mọi đích. (Cũng có khi người ta định nghĩa bài toán vận tải không có giả thiết cuối này, còn bài toán thêm giả thiết này được gọi là bài toán vận tải Hitchcock. Ở đây ta chỉ xét bài toán vận tải Hitchcock, nên không cần từ "Hitchcock".) Người ta ký hiệu phân biệt các nguồn là $s_i \geq 0$ và các nhu cầu là $d_j \geq 0$ (không ký hiệu chung là b_i).

Chú ý thêm rằng bài toán vận tải mở, tức là khi các nguồn có thể không chuyển hết sang các đích ($\sum_{j=1}^n x_{ij} \leq s_i, i = 1, \dots, m$) có thể đưa về bài toán vận tải đóng bằng cách đưa thêm một nút cầu giả $n+1$ với nhu cầu

$$d_{n+1} = \sum_{i=1}^m s_i - \sum_{j=1}^n d_j$$

d_{n+1} chính là tổng lượng hàng còn trong các nguồn, không được chuyển đi.

Nhiều bài toán thực tế có bản chất khác hẳn "vận tải" nhưng có mô hình toán học là bài toán vận tải.

THÍ DỤ 5.1.1 (Bài toán bổ nhiệm - assignment problem).

Giả sử có tập S gồm m người và tập D gồm m việc hoặc chức vụ. Cước phí của việc bổ nhiệm người $i \in S$ vào việc $j \in D$ là c_{ij} . Bài toán đặt ra là tìm cách chia cho mỗi người đúng một việc sao cho tổng cước phí là nhỏ nhất. Đặt biến (biến dòng) là

$$x_{ij} = \begin{cases} 1 & \text{nếu người } i \text{ được việc } j, \\ 0 & \text{trường hợp khác} \end{cases}$$

thì bài toán trở thành

$$\min \sum_{i \in S} \sum_{j \in D} c_{ij} x_{ij},$$

$$\sum_{j \in D} x_{ij} = 1, i \in S \text{ (mỗi người chỉ được đúng một việc),}$$

$$\sum_{i \in S} x_{ij} = 1, j \in D \text{ (mỗi việc chỉ cho đúng một người).}$$

Đây là bài toán vận tải nhưng có thêm yêu cầu là các x_{ij} chỉ lấy giá trị 0 hoặc 1. Khi bỏ yêu cầu này ta được bài toán vận tải. Bài toán bớt yêu cầu đó được gọi là *LP - nới lỏng* (LP - relaxation, LP là từ Linear Programming) của bài toán bổ nhiệm.

Phương pháp đơn hình mạng nêu trong Chương 3 rất ưu việt trong trường hợp này. Vì các s_i và d_j ở bài toán LP - nới lỏng đều nguyên (= 1) nên theo Định lý 3.3.4 mọi nghiệm cơ sở đều nguyên, do đó các thành phần x_{ij} chỉ bằng 0 hoặc 1. Do đó, lời giải LP - nới lỏng cũng chính là lời giải bài toán bổ nhiệm. Chú ý rằng nếu ta áp dụng phương pháp khác, chẳng hạn các phương pháp điểm trong (xem [Khánh - Nương 2000]) để giải LP - nới lỏng thì có thể hội tụ nhanh hơn phương pháp đơn hình, nhưng tính nguyên của nghiệm không được đảm bảo, nên ta chưa được lời giải của bài toán bổ nhiệm.

Bài toán bổ nhiệm cũng có khi được gọi là *bài toán chọn* (choice problem). Nhiều bài toán thực tế đa dạng có mô hình toán học là bài toán bổ nhiệm, chẳng hạn bài toán phân bố hoá khí vào các mục tiêu cần tiêu diệt.

5.1.2 Các bài toán đưa về bài toán vận tải

Ta xét một số trong nhiều bài toán có hình thức khác phức tạp hơn hoặc tổng quát hơn, nhưng có thể đưa về bài toán vận tải.

(1) Bài toán với cung ít hơn cầu

Khi tổng nguồn cung $\sum_{i=1}^m s_i$ nhỏ hơn tổng nhu cầu $\sum_{j=1}^n d_j$, tất nhiên không thể đáp ứng nhu cầu d_j cho mọi nút cầu $j = 1, \dots, n$. Khi đó ở phát biểu bài toán ta phải có các ràng buộc

$$\sum_{i=1}^n x_{ij} \leq d_j, j = 1, \dots, n,$$

thay cho ràng buộc đẳng thức.

Ta có thể đưa về bài toán vận tải đóng như sau. Có hai trường hợp. Trường hợp thứ nhất là có tính đến sự thiệt hại (bằng tiền) khi thiếu một đơn vị hàng ở nút cầu j là $r_j, j = 1, \dots, n$. Khi đó ta đưa thêm một *nút cung giả* (dummy source) $m + 1$, với nguồn giả là

$$s_{m+1} = \sum_{j=1}^n d_j - \sum_{i=1}^m s_i$$

và cước phí $c_{m+1j} = r_j$. Bài toán vận tải đóng nhận được là

$$\min \sum_{j=1}^n \sum_{i=1}^{m+1} c_{ij} x_{ij},$$

$$\sum_{i=1}^{m+1} x_{ij} = d_j, j = 1, \dots, n,$$

$$\sum_{j=1}^n x_{ij} = s_i, i = 1, \dots, m+1,$$

$$x_{ij} \geq 0, i = 1, \dots, m+1, j = 1, \dots, n.$$

Lượng $x_{m+1j} = d_j - \sum_{i=1}^m x_{ij}$ chính là lượng hàng thiếu (không được chở đến) ở nút cầu j .

Trường hợp thứ hai là không tính đến sự thiệt hại cho thiếu hàng ở các nút cầu. Khi đó ta đưa về bài toán vận tải đóng như trên, nhưng đã biết $r_j = c_{m+1j} = 0, j = 1, \dots, n$, nên mục tiêu chỉ còn là như bài toán gốc đang xét:

$$\min \sum_{j=1}^n \sum_{i=1}^m c_{ij} x_{ij}.$$

Hoàn toàn tương tự như vậy, ở bài toán vận tải mở cũng có thể tính đến *cước phí lưu kho* (storage cost) ở các nguồn cho một đơn vị hàng là $c_{in+1}, i = 1, \dots, m$. Khi đưa về bài toán vận tải đóng bằng các đưa vào nút cầu giả $n+1$ như trên đã nói, mục tiêu trở thành

$$\min \sum_{j=1}^{n+1} \sum_{i=1}^m c_{ij} x_{ij}.$$

Như vậy ta chỉ cần xét bài toán vận tải đóng, và sau đây gọi (tất) là bài toán vận tải. Đó là bài toán

$$\min \sum_{j=1}^n \sum_{i=1}^m c_{ij} x_{ij}, \quad (5.1)$$

$$(T) \quad \sum_{j=1}^n x_{ij} = s_i, \quad i = 1, \dots, m, \quad (5.2)$$

$$\sum_{i=1}^m x_{ij} = d_j, \quad j = 1, \dots, n, \quad (5.3)$$

$$x_{ij} \geq 0, \quad i = 1, \dots, m, \quad j = 1, \dots, n. \quad (5.4)$$

(2) Bài toán vận tải có đường cấm

Đây là bài toán vận tải nhưng không phải mỗi nguồn đều có cung nối với mọi đích, có nghĩa là "có đường cấm". Cách đưa về bài toán vận tải là dùng phương pháp M-lớn, tức là phương pháp phạt, như sau.

Gọi E là tập các cung không cấm, tức là tập các (i, j) , $i \in S$, $j \in D$, mà bài toán có thêm điều kiện

$$x_{ij} = 0 \text{ với } (i, j) \notin E. \quad (5.5)$$

Ta đưa bài toán (5.1)-(5.5) về bài toán (T) bằng cách đặt cước vận chuyển mới

$$\bar{c}_{ij} = \begin{cases} c_{ij} & \text{nếu } (i, j) \in E, \\ M & \text{nếu } (i, j) \notin E, \end{cases}$$

ở đây M là một số rất lớn, coi là lớn hơn mọi số gặp phải khi tính toán. Gọi bài toán (T), tức là bài toán (5.1)-(5.4), nhưng với cước phí mới này là $T(M)$ thì ta có

ĐỊNH LÝ 5.1.1. Giả sử $x^* = (x_{ij}^*)_{m,n}$ là phương án vận chuyển tối ưu của bài toán $T(M)$. Khi đó

(i) nếu $x_{ij}^* = 0 \quad \forall (i, j) \notin E$ thì chính x^* là phương án (vận chuyển) tối ưu của bài toán vận tải có đường cấm (5.1)-(5.5);

(ii) nếu tồn tại $(k, l) \notin E$ mà $x_{kl}^* > 0$ thì bài toán vận tải có đường cấm không có nghiệm chấp nhận được.

CHỨNG MINH. (i) Vì mọi phương án x chấp nhận được của $T(M)$ mà $x_{ij} = 0 \quad \forall (i, j) \notin E$ cũng là phương án chấp nhận được của bài toán vận tải có đường cấm và ngược lại nên kết luận là đúng rõ ràng.

(ii) Giả sử phản chứng là có $x = (x_{ij})_{m,n}$ là nghiệm chấp nhận được của bài toán có đường cấm. Khi đó nó cũng là nghiệm chấp nhận được của $T(M)$, với

giá trị mục tiêu là

$$\sum_{i=1}^m \sum_{j=1}^n \bar{c}_{ij} x_{ij} = \sum_{(i,j) \in E} c_{ij} x_{ij} \leq (\max_{(i,j)} c_{ij}) \sum_{i=1}^m \sum_{j=1}^n x_{ij} = (\max_{(i,j)} c_{ij}) \sum_{i=1}^m s_i. \quad (5.6)$$

Giá trị mục tiêu tối ưu của $T(M)$ là

$$\sum_{i=1}^m \sum_{j=1}^n \bar{c}_{ij} x_{ij}^* \geq M x_{kl}^*,$$

lớn hơn giá trị (5,6) vì M là rất lớn, mâu thuẫn với tính tối ưu. $\square$

(3) Bài toán vận tải kèm chế biến trung gian

Giả sử trong mô hình vận tải có một số điểm nguồn, tức là điểm sản xuất, cho ra một số bán sản phẩm cần phải chế biến trước khi đến điểm cầu (điểm đích). Giả sử có k điểm chế biến $\lambda = 1, \dots, k$ với khả năng chế biến a_λ đơn vị sản phẩm, tương ứng. Gọi cước phí vận chuyển một đơn vị bán sản phẩm từ i đến λ là $c'_{i\lambda}$ và chuyển một đơn vị sản phẩm từ λ đến j là $c''_{\lambda j}$. Bài toán đặt ra là lập kế hoạch vận chuyển tất cả các bán sản phẩm qua chế biến đến thỏa mãn tất cả các điểm cầu sao cho cước phí nhỏ nhất.

Gọi $z_{i\lambda j}$ là lượng sản phẩm từ i qua λ rồi đến điểm cầu j . Dạng toán học sẽ là tìm $z = (z_{i\lambda j})_{mkn}$ sao cho

$$\min \sum_{i=1}^m \sum_{\lambda=1}^k \sum_{j=1}^n (c'_{i\lambda} + c''_{\lambda j}) z_{i\lambda j}, \quad (5.7)$$

$$\sum_{\lambda=1}^k \sum_{j=1}^n z_{i\lambda j} = s_i, \quad i = 1, \dots, m, \quad (5.8)$$

$$\sum_{i=1}^m \sum_{\lambda=1}^k z_{i\lambda j} = d_j, j = 1, \dots, n, \quad (5.9)$$

$$\sum_{i=1}^m \sum_{j=1}^n z_{i\lambda j} \leq a_{\lambda}, \lambda = 1, \dots, k, \quad (5.10)$$

$$z_{i\lambda j} \geq 0, i = 1, \dots, m, \lambda = 1, \dots, k, j = 1, \dots, n. \quad (5.11)$$

Ta sẽ cố gắng đưa (5.7)-(5.11) về bài toán vận tải có đường cấm bằng cách đưa vào các biến mới

$$x_{i, n+\lambda} = \sum_{j=1}^n z_{i\lambda j}, i = 1, \dots, m, \lambda = 1, \dots, k, \quad (5.12)$$

$$x_{m+\lambda, j} = \sum_{i=1}^m z_{i\lambda j}, \lambda = 1, \dots, k, j = 1, \dots, n. \quad (5.13)$$

Biến $x_{i, n+\lambda}$ là toàn bộ lượng bán sản phẩm chuyển từ nút i đến λ . Biến $x_{m+\lambda, j}$ là lượng sản phẩm chuyển từ nút chế biến λ đến nút cầu j . Viết theo các biến mới, bài toán (5.7)-(5.11) trở thành

$$\min \sum_{i=1}^m \sum_{\lambda=1}^k c'_{i\lambda} x_{i, n+\lambda} + \sum_{\lambda=1}^k \sum_{j=1}^n c''_{\lambda j} x_{m+\lambda, j}, \quad (5.14)$$

$$\sum_{\lambda=1}^k x_{i, n+\lambda} = s_i, i = 1, \dots, m, \quad (5.15)$$

$$\sum_{\lambda=1}^k x_{m+\lambda, j} = d_j, j = 1, \dots, n, \quad (5.16)$$

$$\sum_{i=1}^m x_{i, n+\lambda} = \sum_{j=1}^n x_{m+\lambda, j} \leq a_{\lambda}, \lambda = 1, \dots, k, \quad (5.17)$$

$$x_{i, n+\lambda} \geq 0, x_{m+\lambda, j} \geq 0, i = 1, \dots, m, \lambda = 1, \dots, k, j = 1, \dots, n. \quad (5.18)$$

Ta biến đổi thêm (5.17) bằng cách thêm vào biến bù

$$x_{m+\lambda, n+\lambda} \geq 0, \lambda = 1, \dots, k$$

để (5.11) viết lại ở dạng hai đẳng thức:

$$\sum_{i=1}^m x_{i, n+\lambda} + x_{m+\lambda, n+\lambda} = a_{\lambda}, \lambda = 1, \dots, k, \quad (5.19)$$

$$\sum_{j=1}^n x_{m+\lambda, j} + x_{m+\lambda, n+\lambda} = a_{\lambda}, \lambda = 1, \dots, k. \quad (5.20)$$

Để thấy rõ lúc này ta đã nhận được bài toán vận tải có đường cấm ta đưa thêm vào các điểm cung (điểm sản xuất) mới $m+1, \dots, m+k$ với lượng hàng tương ứng $a_1, \dots, a_k$ và các điểm cầu mới $n+1, \dots, n+k$ với lượng cầu $a_1, \dots, a_k$, tương ứng. Chúng ta quy định rằng các điểm cung $i, i = 1, \dots, m$, chỉ nối với các điểm cầu $n+\lambda, \lambda = 1, \dots, k$, và các điểm cung mới $m+\lambda, \lambda = 1, \dots, k$, nối với mọi điểm cầu. Đặt cước phí mới là

$$c_{i, n+\lambda} = c'_{i\lambda}, c_{m+\lambda, j} = c''_{\lambda j}, c_{m+\lambda, n+\lambda} = 0, \\ i = 1, \dots, m, \lambda = 1, \dots, k, j = 1, \dots, n.$$

Dễ kiểm tra rằng (5.15), (5.20) đảm bảo toàn bộ lượng hàng sản xuất ra được vận chuyển, và (5.16), (5.19) đảm bảo các nhu cầu đều thỏa. Vậy bài toán (5.14)-(5.20) chính là bài toán vận tải có đường cấm. Ở đây lượng hàng trên đường cấm (bằng 0) không đưa vào ký hiệu nữa.

Ngoài ba bài toán trên, còn nhiều bài toán có thể đưa về bài toán vận tải, chẳng hạn bài toán vận chuyển nhiều loại hàng cùng lúc, trong đó chia thành các nhóm mà các loại hàng trong mỗi nhóm có thể thay thế cho nhau (theo hệ số nào đó). Chẳng hạn một nhóm là than, dầu hỏa, xăng, mazut; nhóm khác là ba loại xi măng khác

nhau... Một bài toán khác là có nhiều phương tiện vận tải kiểu khác nhau. Độc giả có thể xem thêm trong [Golstein-Yudin 1969].

5.1.3 Biểu diễn toán học bài toán vận tải

Ta đã thấy bài toán vận tải là bài toán quy hoạch tuyến tính có dạng (5.1)-(5.4). Như ở Mục 5.1.1, khi định nghĩa ta cũng thấy bài toán vận tải là một bài toán dòng trên mạng đặc biệt, mà đồ thị của nó là đồ thị hai phần, phần S là các nút nguồn và phần D là các nút cầu (đích) rời nhau mà mỗi nguồn đều có nối với mọi đích. Như vậy về biểu diễn cũng như cách giải có thể dùng tất cả các phương pháp chung nhất của quy hoạch tuyến tính, cũng như các phương pháp của lớp các bài toán mạng. Tuy vậy, tận dụng các đặc điểm của bài toán vận tải sẽ thuận lợi hơn, đỡ phức tạp hơn. Chẳng hạn nếu vẽ đồ thị của bài toán vận tải thì các cung cắt nhau rất nhiều nên thực tế không cung cấp được minh họa hình học.

Biểu diễn ma trận của bài toán vận tải (5.1)-(5.4) cũng ít ý nghĩa thực tế nhưng cũng giúp ta phân tích các tính chất lý thuyết. Không gian phương án là $m \cdot n$ chiều nên vectơ hệ số mục tiêu là

$$c = (c_{11}, \dots, c_{1n}, c_{21}, \dots, c_{2n}, \dots, c_{m1}, \dots, c_{mn}).$$

Vectơ phương án là

$$x = (x_{11}, x_{12}, \dots, x_{1n}, x_{21}, x_{22}, \dots, x_{2n}, \dots, x_{m1}, x_{m2}, \dots, x_{mn}).$$

Ma trận A sẽ là (từ (5.2), (5.3))

$$A = \begin{pmatrix} 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & 1 & \dots & 1 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 1 & 1 & \dots & 1 \\ 1 & 0 & \dots & 0 & 1 & 0 & \dots & 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & 1 & \dots & 0 & \dots & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & 1 & \dots & 0 & 0 & \dots & 1 \end{pmatrix}$$

Vectơ vế phải là $b = (s_1, s_2, \dots, s_m, d_1, d_2, \dots, d_n)^T$.

Dạng ma trận của bài toán vận tải (5.1)-(5.4) là

$$\min c^T x,$$

$$Ax = b,$$

$$x \geq 0.$$

Các vectơ cột của ma trận A sẽ ký hiệu là A_{ij} , $i = 1, \dots, m$, $j = 1, \dots, n$.
Các vectơ hàng của A ký hiệu là A_k , $k = 1, \dots, m+n$.

NHẬN XÉT 5.1. (a) Vì ta đã phân biệt nguồn $s_i \geq 0$ và điểm hút với $d_j \geq 0$ mà không ký hiệu chung là b_i với quy ước $b_i < 0$ là điểm hút nên cách biểu diễn ma trận A ở đây khác với cách biểu diễn chung của bài toán dòng trên mạng. (Chú ý là ở đây thứ tự các cột đã ấn định sẵn.)

(b) A có đúng $m + n - 1$ hàng độc lập tuyến tính, tức là $\text{rank } A = m + n - 1$. Chứng minh ở đây đơn giản hơn trường hợp bài toán dòng trên mạng tổng quát. Thật vậy, dễ thấy tổng m hàng đầu của A bằng tổng n hàng cuối của A (và bằng vectơ có các thành phần đều bằng 1). Do đó $\text{rank } A \leq m + n - 1$. Mặt khác, ma trận con vuông cấp $m + n - 1$ của A là giao của các cột $A_{11}, A_{21}, \dots, A_{m-1,1}, A_{m1}, A_{m2}, \dots, A_{mn}$ và các hàng $A_1, A_2, \dots, A_{m-1}, A_{m+1}, \dots, A_{m+n}$ có định thức bằng 1. Vậy $\text{rank } A = m + n - 1$.

Cách biểu diễn thuận tiện nhất bài toán vận tải là dùng bảng gồm $m+1$ hàng, $n+1$ cột, tức là có $(m+1)(n+1)$ ô. Trên các ô $(i, 0)$ ghi các số s_i , $i = 1, \dots, m$, từ nay sẽ gọi là *lượng phát*. Trên các ô $(0, j)$ ghi các số d_j , $j = 1, \dots, n$, gọi là *lượng thu*. Tập các ô $((i, j): i = 1, \dots, m, j = 1, \dots, n)$ gọi là phần chính của bảng vận tải. Sau đây ta dùng thuật ngữ *bảng vận tải* (transportation table) để chỉ phần chính này cho gọn. Ở góc trên bên trái mỗi ô (i, j) ta ghi cước phí c_{ij} chuyển (một đơn vị hàng) từ nút nguồn i , sau đây sẽ gọi là *điểm phát* i , đến điểm hút j , sau đây gọi là *điểm thu* j . Như vậy mỗi ô (i, j) của bảng vận tải

ứng với cột A_{ij} của ma trận A . Ta ghi thành phần tương ứng x_{ij} của phương án đang xét vào góc dưới bên phải ô. Khi đó ta được bảng vận tải tương ứng với phương án x . Các ô (i, j) mà $x_{ij} > 0$ được gọi là ô sử dụng (hoặc ô chọn). Các ô có $x_{ij} = 0$ được gọi là ô loại (hoặc ô không sử dụng). Để đơn giản ta không viết các lượng $x_{ij} = 0$ vào bảng. Tập các ô sử dụng trong bảng vận tải ứng với phương án x thường kí hiệu là $G(x)$:

$$G(x) = \{(i, j): x_{ij} > 0\}.$$

d_j	d_1	...	d_j	...	d_n
s_i	c_{i1}		c_{ij}		c_{in}
s_1	x_{11}		x_{1j}		x_{1n}
...					
s_i	c_{i1}		c_{ij}		c_{in}
	x_{i1}		x_{ij}		x_{in}
...					
s_m	c_{m1}		c_{mj}		c_{mn}
	x_{m1}		x_{mj}		x_{mn}

Bảng vận tải

Ta gọi *dây chuyền* hoặc *đường đi* (path) trong bảng vận tải là một dãy các ô thỏa hai điều kiện:

- (i) hai ô cạnh nhau nằm trong cùng hàng hoặc cùng cột;
- (ii) không có ba ô nào nằm cùng hàng hoặc cùng cột.

Dây chuyền thỏa thêm điều kiện

(iii) ô đầu tiên và ô cuối cùng nằm trên cùng một hàng hoặc một cột được gọi là *chu trình* (loop hoặc closed loop hoặc cycle). Vậy chu trình là một dây chuyền khép kín.

5.2 CÁC TÍNH CHẤT CỦA BÀI TOÁN VẬN TẢI

5.2.1 Sự tồn tại tại phương án tối ưu

ĐỊNH LÝ 5.2.1 (Định lý tồn tại). *Điều kiện cần và đủ để bài toán vận tải có phương án tối ưu là tổng lượng phát bằng tổng lượng thu, tức là*

$$\sum_{i=1}^m s_i = \sum_{j=1}^n d_j > 0, \quad (5.21)$$

(điều kiện cân bằng thu phát).

CHỨNG MINH. *Cần.* Giả sử bài toán vận tải có phương án tối ưu x^* . Khi đó, theo các điều kiện (5.2) và (5.3),

$$\sum_{i=1}^m s_i = \sum_{i=1}^m \sum_{j=1}^n x_{ij}^* = \sum_{j=1}^n \sum_{i=1}^m x_{ij}^* = \sum_{j=1}^n d_j.$$

Đủ. Giả sử có (5.21). Trước hết ta chỉ ra là bài toán là chấp nhận được, tức là có phương án chấp nhận được. Đặt $E = \sum_{i=1}^m s_i = \sum_{j=1}^n d_j$. Gọi $x = (x_{ij})$ là phương án có thành phần

$$x_{ij} = \frac{s_i d_j}{E}, \quad i = 1, \dots, m, j = 1, \dots, n.$$

Ta có

$$\sum_{j=1}^n x_{ij} = \sum_{j=1}^n \frac{s_i d_j}{E} = \frac{s_i}{E} \sum_{j=1}^n d_j = \frac{s_i}{E} \cdot E = s_i, \quad i = 1, \dots, m.$$

Tương tự, cũng có $\sum_{i=1}^m x_{ij} = d_j, j = 1, \dots, n$. Đồng thời, rõ ràng $x_{ij} \geq 0$. Vậy

x là phương án chấp nhận được.

Bây giờ ta chứng tỏ hàm mục tiêu bị chặn dưới trên miền ràng buộc. Từ (5.2) và (5.4) ta có $0 \leq x_{ij} \leq s_i$. Vì vậy

$$\begin{aligned} c_{ij} x_{ij} &\geq 0 \text{ nếu } c_{ij} \geq 0, \\ c_{ij} x_{ij} &\geq c_{ij} \cdot s_i \text{ nếu } c_{ij} < 0. \end{aligned}$$

Do đó

$$\sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \geq \sum_{i=1}^m \sum_{j=1}^n \min \{0, c_{ij} s_i\} = \text{const}.$$

Vì bài toán vận tải là quy hoạch tuyến tính và có hàm mục tiêu bị chặn dưới trên miền ràng buộc khác trống nên tồn tại phương án tối ưu. $\square$

Từ nay về sau ta chỉ xét bài toán vận tải cân bằng thu phát, hoặc gọi tắt là *bài toán vận tải cân bằng* (balanced transportation problem).

5.2.2 Các tính chất về nghiệm cơ sở

Bài toán vận tải cũng là quy hoạch tuyến tính dạng chính tắc nên ta đã biết một phương án chấp nhận được x là nghiệm cơ sở (hay cũng gọi là phương án cơ sở, phương án cực biên) chấp nhận được khi và chỉ khi các cột của ma trận A ứng với các thành phần $x_{ij} > 0$ là độc lập tuyến tính. Ta đã gọi các ô của bảng vận tải ứng với x mà có $x_{ij} > 0$ là các ô sử dụng. Ta quy ước lúc này cũng nói là các ô sử dụng độc lập tuyến tính.

MỆNH ĐỀ 5.2.2. Giả sử tập U các ô trong bảng vận tải có tính chất là mỗi hàng hoặc cột của bảng hoặc là không chứa ô nào của U hoặc chứa ít ra là hai ô của U . Khi đó U phải chứa chu trình.

CHỨNG MINH. Bắt đầu từ một ô (i_1, j_1) bất kỳ của U . Trong hàng i_1 phải có ô (i_1, j_2) nào đó của U . Tiếp tục trong cột j_2 phải có ô khác nữa, chẳng hạn ô (i_2, j_2) . Ta làm tiếp nếu chưa được chu trình. Vì số ô của U là hữu hạn nên sẽ đến lúc ta gặp lại một ô đã đi qua tức là ta đã gặp một chu trình trong U . $\square$

ĐỊNH LÝ 5.2.3. Một tập các ô trong bảng vận tải là độc lập tuyến tính (tức là các cột tương ứng của ma trận A là độc lập tuyến tính) khi và chỉ khi tập này không chứa chu trình.

CHỨNG MINH. Khi. Giả sử phản chứng là tập U các ô trong bảng vận tải không chứa chu trình nhưng là phụ thuộc tuyến tính, tức là tồn tại γ_{ij} , $(i, j) \in U$, không đồng thời bằng 0 sao cho

$$\sum_{(i,j) \in U} \gamma_{ij} A_{ij} = 0. \quad (5.22)$$

Điền các số γ_{ij} vào ô (i, j) tương ứng trong bảng vận tải còn các ô khác ghi số 0. Khi đó theo (5.22) tổng các số ghi trong mỗi hàng của bảng đơn hình bằng 0 (chú ý đến m hàng đầu của ma trận A viết ở Mục 5.1.3). Bây giờ xét tổng các số ghi trong mỗi cột của bảng. Giả sử xét cột 1. Nhìn ma trận A ta thấy cột này ghi các số trong số $\{\gamma_{11}, \gamma_{21}, \dots, \gamma_{m1}\}$ mà có mặt trong (5.22), nên tổng các số này chính là vế trái phương trình thứ $m+1$ trong (5.22) và phải bằng 0 theo (5.22). Tương tự cho đến cột thứ n thì tổng các số chính là vế trái phương trình thứ $m+n$ trong (5.22) nên cũng bằng 0.

Gọi P là tập các ô trong bảng chứa các số được điền khác 0, tức là $P = \{(i, j): \gamma_{ij} \neq 0\}$. Khi đó mỗi hàng và cột trong bảng hoặc không chứa ô nào của P hoặc chứa ít ra là hai ô (vì tổng các số khác 0 đã điền phải bằng 0). Theo Mệnh đề 5.2.2, P phải chứa chu trình. Vì P là tập con của U nên U cũng chứa chu trình, mâu thuẫn.

Chỉ khi. Giả sử phản chứng là U là độc lập tuyến tính nhưng lại chứa một chu trình V . Điền 1 và -1 một cách xen kẽ vào các ô của V và 0 vào các ô còn lại của bảng vận tải. Theo định nghĩa chu trình, tổng các số đã điền trong mỗi hàng và mỗi cột của bảng vận tải phải bằng 0. Như vậy là ta đã có một bộ hệ

số, có các hệ số khác 0 là 1 và -1 , để tổ hợp tuyến tính các cột A_{ij} của ma trận A tương ứng với các ô của U bằng 0. Suy ra U phụ thuộc tuyến tính, mâu thuẫn. $\square$

Từ định lý này ta có ngay

HỆ QUẢ 5.2.4. Một phương án chấp nhận được của bài toán vận tải là phương án cực biên khi và chỉ khi tập các ô sử dụng của nó không chứa chu trình.

Áp dụng định nghĩa phương án cực biên suy biến của quy hoạch tuyến tính tổng quát vào bài toán vận tải ta sẽ gọi phương án cực biên của bài toán vận tải là *không suy biến* hoặc *không thoái hóa* (nondegenerate) nếu số ô sử dụng của nó đúng bằng $m+n-1$. Ngược lại *phương án cực biên suy biến* (degenerate extreme alternative) là phương án cực biên mà số các ô sử dụng ít hơn $m+n-1$.

5.2.3 Các tính chất về chu trình

MỆNH ĐỀ 5.2.5. Trong bảng vận tải gồm $m \cdot n$ ô với $m \geq 2$, $n \geq 2$ mọi tập gồm từ $m+n$ ô trở lên đều phải chứa chu trình.

CHỨNG MINH. Vì hạng rank $A = m + n - 1$ nên trong bảng vận tải mọi tập có $\geq m+n$ ô đều phụ thuộc tuyến tính, và do đó phải chứa chu trình theo Định lý 5.2.3. $\square$

Từ định nghĩa chu trình ta có thủ tục sau đây để nhét ra chu trình từ một tập T có chu trình, chẳng hạn tập có $\geq m+n$ ô.

- Xóa khỏi bảng vận tải mọi dòng và cột chỉ chứa không quá một ô của tập T .
- Lập lại việc làm trên cho bảng mới nhận được cho đến khi thu được bảng mà mỗi dòng và cột đều chứa ít nhất hai ô của T .
- Lấy ra chu trình từ bảng cuối cùng như cách làm ở chứng minh Mệnh đề 5.2.2.

Thủ tục này cũng dùng để phát hiện ra tập các ô không có chu trình. Điều này xảy ra nếu ta xoá hết bảng vận tải mà không dừng lại trước đó.

THÍ DỤ 5.2.1. Tìm chu trình từ các ô có đánh dấu x.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1		x				x				x		x				
2	x			x											x	
3		x	x		x		x	x		x				x		x
4				x								x		x		
5								x	x							

- Xoá cột 1, 3, 5, 6, 7, 9, 15, 16, 11 và 13.
- Xoá hàng 2 và hàng 5.
- Xoá cột 4 và cột 8 ta nhận được bảng mới

	2	10	12	14
1	x	x	x	
3	x	x		x
4			x	x

- Bắt đầu từ ô (1, 2) đi đến ô (1, 10) cùng hàng, tiếp đến ô (3, 10) cùng cột và về ô (3, 2) cùng hàng ta được một chu trình.

Nếu bắt đầu từ ô (1, 10) sang ô (1, 12), đến (4, 12) rồi (4, 14), (3, 14) và về (3, 10) ta được chu trình khác.

Nhận xét rằng số ô trong một chu trình luôn chẵn.

Ta đã biết phương án cực biên là rất quan trọng đối với quy hoạch tuyến tính. Mà phương án chấp nhận được là cực biên khi và chỉ khi

các ô sử dụng của nó không chứa chu trình. Vì vậy định lý sau về "phá chu trình" là quan trọng.

ĐỊNH LÝ 5.2.6. Giả sử x là phương án chấp nhận được mà tập ô sử dụng $G(x)$ có chứa chu trình. Khi đó từ x luôn xây dựng được phương án $\bar{x}$ chấp nhận được với giá trị mục tiêu không lớn hơn mà $G(\bar{x})$ không chứa chu trình.

CHỨNG MINH. Giả sử V là một chu trình trong $G(x)$. Đánh dấu các ô trong V , từ một ô bất kỳ, bởi các dấu $+$ và $-$ xen kẽ. Gọi V^+ (V^-) là tập các ô trong V có dấu $+$ ($-$, tương ứng). Giả sử, để xác định,

$$\sum_{(i,j) \in V^+} c_{ij} \leq \sum_{(i,j) \in V^-} c_{ij}. \quad (5.23)$$

Xây dựng phương án $\bar{x}$ với các thành phần

$$\bar{x}_{ij} = \begin{cases} x_{ij} + \theta & \text{nếu } (i,j) \in V^+, \\ x_{ij} - \theta & \text{nếu } (i,j) \in V^-, \\ x_{ij} & \text{nếu } (i,j) \notin V, \end{cases}$$

trong đó $\theta = \min \{x_{ij} : (i,j) \in V^-\}$. Từ định nghĩa θ rõ ràng $\bar{x}_{ij} \geq 0$. Vì các ô trong V từng đôi thuộc V^+ và V^- xen kẽ nên

$$\sum_{i=1}^m \bar{x}_{ij} = \sum_{i=1}^m x_{ij} = d_j, j = 1, \dots, n, \quad \sum_{j=1}^n \bar{x}_{ij} = \sum_{j=1}^n x_{ij} = s_i, i = 1, \dots, m.$$

Vậy $\bar{x}$ là phương án chấp nhận được. Hơn nữa, giá trị mục tiêu mới là

$$\sum_i \sum_j c_{ij} \bar{x}_{ij} = \sum_i \sum_j c_{ij} x_{ij} + \theta \left(\sum_{(i,j) \in V^+} c_{ij} - \sum_{(i,j) \in V^-} c_{ij} \right) \leq \sum_i \sum_j c_{ij} x_{ij}.$$

Do định nghĩa của θ và $\bar{x}_{ij}$ sẽ có ít nhất một $(i,j) \in V^-$ để $\bar{x}_{ij} = 0$ tức là $(i,j) \notin G(\bar{x})$. Vậy $G(\bar{x})$ không chứa V . Nếu (5.23) có dấu ngược lại thì cải biên định nghĩa $\bar{x}_{ij}$ và θ tương ứng ta cũng có kết quả như vậy. Đến đây nếu $G(\bar{x})$ còn chứa

chu trình ta lại làm tiếp với $\bar{x}$ như với x trên đây. Cuối cùng ta đến phương án chấp nhận được có mục tiêu không lớn hơn ban đầu và có tập ô sử dụng không chứa chu trình. $\square$

MỆNH ĐỀ 5.2.7. Giả sử P là tập gồm $m + n - 1$ ô, không chứa chu trình, trong bảng vận tải gồm $m.n$ ô và $(i, j) \notin P$. Khi đó $P \cup (i, j)$ sẽ chứa một chu trình duy nhất.

CHỨNG MINH. Theo Mệnh đề 5.2.5 thì $P \cup (i, j)$ phải chứa chu trình. Giả sử phản chứng là $P \cup (i, j)$ chứa hai chu trình khác nhau:

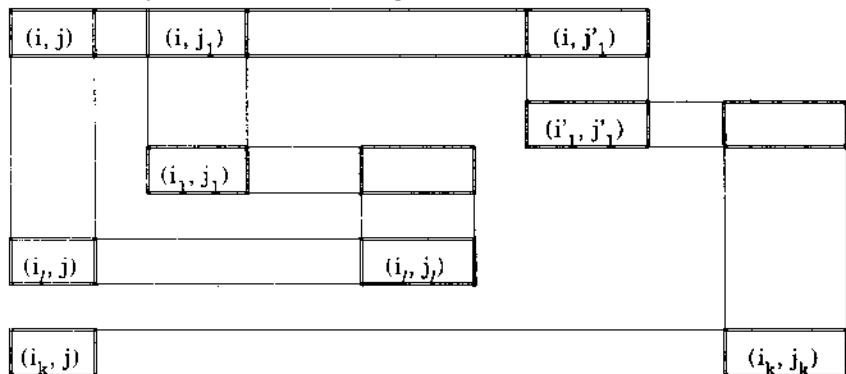
$$V_1 = \{(i, j), (i, j_1), (i_1, j_1), \dots, (i_p, j_p), (i_p, j)\},$$

$$V_2 = \{(i, j), (i, j'_1), (i'_1, j'_1), \dots, (i_k, j_k), (i_k, j)\}.$$

Giả sử để xác định là sự khác nhau thể hiện ít nhất ở chỗ $k \neq l$. Ta lập một chu trình mới không chứa (i, j) , và do đó nằm trong P , như sau

$$V_3 = \{(i, j_1), (i_1, j_1), \dots, (i_p, j_p), (i_p, j), (i_k, j), (i_k, j_k), \dots, (i, j'_1)\}.$$

Điều này mâu thuẫn vì P không chứa chu trình (xem H.5.1). $\square$



H 5.1

MỆNH ĐỀ 5.2.8. Nếu tập P các ô trong bảng vận tải chứa duy nhất một chu trình V_1 và $(i, j) \in V_1$, thì tập $P \setminus (i, j)$ không chứa chu trình.

CHỨNG MINH. Giả sử phản chứng là $P \setminus (i, j)$ chứa chu trình V_2 thì $(i, j) \notin V_2$ nên $V_2 \neq V_1$. Tập P có hai chu trình khác nhau, vô lý. $\square$

NHẬN XÉT 5.2.1. Bài toán vận tải cũng là bài toán dòng trên mạng nên cũng biểu diễn được bằng đồ thị (tuy rằng khi đó các cung trên hình vẽ cắt nhau nhiều nên phức tạp). Vậy khái niệm dây chuyền (đường đi) và chu trình định nghĩa theo ngôn ngữ bảng vận tải trên đây có trùng với định nghĩa theo ngôn ngữ đồ thị không? Dễ thấy là "trùng!", chỉ việc so sánh qua định nghĩa bảng vận tải và đồ thị ta sẽ thấy rõ điều đó. Phương án cơ sở ở bảng vận tải là ứng với cây bao trùm ở đồ thị.

5.3 GIẢI BÀI TOÁN VẬN TẢI. THUẬT TOÁN THẾ VỊ

5.3.1 Tiêu chuẩn tối ưu

Ta sẽ trình bày phương pháp đơn hình mạng áp dụng cho bài toán vận tải theo ngôn ngữ bảng vận tải. Bài toán đối ngẫu của bài toán vận tải (5.1) – (5.4) (tức là đối ngẫu của một quy hoạch tuyến tính chính tắc) sẽ là

$$\max \sum_{i=1}^m s_i p_i + \sum_{j=1}^n d_j q_j, \quad (5.24)$$

$$p_i + q_j \leq c_{ij}, \quad i = 1, \dots, m, j = 1, \dots, n.$$

Dựa trên tính đối ngẫu ta sẽ chứng minh tiêu chuẩn tối ưu sau đây.

ĐỊNH LÝ 5.3.1. (Tiêu chuẩn tối ưu). Phương án chấp nhận được $x = (x_{ij})$ của bài toán vận tải là tối ưu khi và chỉ khi tìm được các giá trị

của biến đối ngẫu p_i và q_j , $i = 1, \dots, m, j = 1, \dots, n$, sao cho

$$p_i + q_j \leq c_{ij} \quad \forall (i, j), \quad (5.25)$$

$$p_i + q_j = c_{ij} \quad \text{nếu } x_{ij} > 0. \quad (5.26)$$

CHỨNG MINH. Khi giả sử có các p_i, q_j thỏa (5.25), (5.26) và $\bar{x} = (\bar{x}_{ij})$ là phương án chấp nhận được bất kỳ. Do (5.26), giá trị mục tiêu tại $\bar{x}$ là

$$\sum_{i=1}^m \sum_{j=1}^n c_{ij} \bar{x}_{ij} = \sum_{i=1}^m \sum_{j=1}^n (p_i + q_j) \bar{x}_{ij}.$$

Do (5.25) giá trị mục tiêu tại $\bar{x}$ được đánh giá như sau

$$\begin{aligned} \sum_{i=1}^m \sum_{j=1}^n c_{ij} \bar{x}_{ij} &\geq \sum_{i=1}^m \sum_{j=1}^n (p_i + q_j) \bar{x}_{ij} = \sum_i p_i \sum_j \bar{x}_{ij} + \sum_j q_j \sum_i \bar{x}_{ij} \\ &= \sum_i p_i s_i + \sum_j q_j d_j = \sum_i p_i \sum_j x_{ij} + \sum_j q_j \sum_i x_{ij} = \sum_i \sum_j (p_i + q_j) x_{ij}. \end{aligned}$$

Vậy $\bar{x}$ là phương án tối ưu.

Chỉ khi. Giả sử $\bar{x}$ là tối ưu. Theo định lý đối ngẫu mạnh, bài toán đối ngẫu cũng có phương án tối ưu $z = (p_1, \dots, p_m, q_1, \dots, q_n)^T$. Do đó z phải chấp nhận được, tức là thỏa (5.25). Theo điều kiện độ lệch bù ta có

$$(c_{ij} - p_i - q_j) x_{ij} = 0 \quad \forall (i, j).$$

Do đó ta có (5.26). $\square$

Qua chứng minh trên ta thấy (5.25), (5.26) có nghĩa là các p_i, q_j là biến đối ngẫu chấp nhận được và cùng với (x_{ij}) thỏa điều kiện độ lệch bù.

Các số p_i, q_j (các biến đối ngẫu) được gọi là các nhân tử

(multipliers) vì từ lý thuyết đối ngẫu tổng quát ta đã biết biến đối ngẫu chính là nhân tử Lagrange. Ở Việt Nam và một số nước như Nga... p_i, q_j được gọi là các *thế vị*. Vì thế phương pháp đơn hình mạng sau đây để giải bài toán vận tải có tên là *phương pháp nhân tử* (method of multipliers) hoặc *phương pháp thế vị*. (Hãy so sánh với phương pháp đơn hình mạng ở Mục 3.3.) Phương pháp này cũng có tên là *phương pháp đơn hình vận tải* (transportation simplex method).

5.3.2 Thuật toán thế vị

- *Bước xuất phát*: Tìm một phương án cơ sở chấp nhận được để xuất phát.

- *Bước lặp tổng quát*: Giả sử x là phương án cơ sở ở bước hiện hành.

(a) Xác định biến vào. Nhớ lại công thức cho cước phí thu gọn hiện hành $\bar{c}$ (tức là hệ số mục tiêu hiện hành) qua dữ liệu ban đầu c, A và biến đối ngẫu hiện hành z (xem Bảng 1.2)

$$\bar{c}^T = c^T - z^T A.$$

Để kiểm tra được với ma trận A của bài toán vận tải, công thức này trở thành

$$\bar{c}_{ij} = c_{ij} - (p_i + q_j) \quad \forall (i, j). \quad (5.27)$$

Theo quy tắc Dantzig biến vào x_{ij} là biến ứng với $\bar{c}_{ij}$ âm nhất, ta phải tính các $\bar{c}_{ij}$. Do đó ta phải tính các biến đối ngẫu p_i, q_j . Vì cước phí thu gọn hiện hành của các biến cơ sở hiện hành phải bằng 0 nên $\bar{c}_{ij} = 0$ với mọi x_{ij} là biến cơ sở hiện hành (ứng với các ô (i, j) của dãy chuyển gồm các ô sử dụng hiện hành). Suy ra

$$p_i + q_j = c_{ij}, \text{ với các biến } x_{ij} \text{ cơ sở}, \quad (5.28)$$

hay viết ở dạng vectơ

$$z^T A_{ij} = c_{ij}, \text{ với } x_{ij} \text{ cơ sở.}$$

Theo Định lý 5.2.3, vì các $\hat{o}(i,j)$ ứng với các biến cơ sở x_{ij} không chứa chu trình nên hệ vec tơ cột tương ứng của A là

$$\{A_{ij}: x_{ij} \text{ là biến cơ sở}\}$$

độc lập tuyến tính. Vậy (5.28) gồm $m+n-1$ phương trình độc lập tuyến tính với nhau để xác định $m+n$ biến p_i, q_j . Do đó có thể cho một biến giá trị tùy ý, thường lấy $p_1 = 0$, rồi giải ra duy nhất các biến p_i, q_j còn lại.

Sau khi có các p_i, q_j ta tính được $\bar{c}_{ij}, \forall (i,j)$, theo (5.27) và chọn $\bar{c}_{i_0 j_0}$ âm nhất để biết biến vào $x_{i_0 j_0}$. Nếu không có hệ số mục tiêu $\bar{c}_{ij}$ nào âm thì phương án hiện hành x là tối ưu.

(b) Xác định biến ra. Tập $\{x_{ij} \text{ là biến cơ sở}\} \cup \{\text{biến vào } x_{i_0 j_0}\}$ gồm $m+n$ phần tử nên tập các $\hat{o}$ tương ứng trong bảng vận tải sẽ có một chu trình duy nhất V . Coi $\hat{o}(i_0, j_0) \in V^+$ rồi xây dựng phương án mới từ x theo thủ tục như trong chứng minh Định lý 5.2.6:

$$\bar{x}_{ij} = \begin{cases} x_{ij} + \theta & \text{nếu } (i, j) \in V^+, \\ x_{ij} - \theta & \text{nếu } (i, j) \in V^-, \\ x_{ij} & \text{nếu } (i, j) \notin V, \end{cases} \quad (5.29)$$

ở đây $\theta = \min \{x_{ij}: (i,j) \in V^-\}$. Khi đó phải có $x_{i_1 j_1} = \theta$ và từ đó $\bar{x}_{i_1 j_1} = 0$ tức là $x_{i_1 j_1}$ là biến ra. Theo Mệnh đề 5.2.8, $\bar{x}$ là phương án cơ sở.

Ở Định lý 5.2.6 ta đã thấy giá trị mục tiêu ở $\bar{x}$ là không lớn hơn giá trị mục tiêu ở x .

NHẬN XÉT 5.3.1

(a) Ở các bước lặp của thuật toán thể vị ta luôn có phương án hiện

hành x là phương án cơ sở chấp nhận được, luôn có điều kiện về độ lệch bù (5.26) thỏa, nhưng các biến đối ngẫu p_i, q_j chưa chấp nhận được. Đến bước nào chúng trở thành chấp nhận được, tức là thỏa (5.25) thì mọi $\bar{c}_{ij}$ không âm và phương án hiện hành $\bar{x}$ là tối ưu.

(b) Nếu bài toán vận tải là không suy biến, tức là tất cả các phương án cơ sở chấp nhận được của nó là không suy biến, thì phương pháp đơn hình không thể xoay vòng, tức là thuật toán thế vị (là thuật toán đơn hình mạng) phải kết thúc sau hữu hạn bước với phương án cơ sở tối ưu.

(c) Rõ ràng ở các bước lặp của thuật toán thế vị ta không phải dùng đến phép tính chia. Do đó nếu dữ liệu ban đầu s_i, d_j là các số nguyên thì theo các cách tìm phương án cơ sở để xuất phát trình bày ở Mục 4.4 dưới đây, phương án xuất phát sẽ có các thành phần nguyên. Vì thế qua các bước lặp, cho đến phương án tối ưu ta luôn có các thành phần phương án đều là các số nguyên.

THÍ DỤ 5.3.1. Giải bài toán vận tải cho bằng bảng H.5.2

d_j	5	15	15	10
s_i				
15	10 x_{11}	0 x_{12}	20 x_{13}	11 x_{14}
25	12 x_{21}	7 x_{22}	9 x_{23}	20 x_{24}
5	0 x_{31}	14 x_{32}	16 x_{33}	18 x_{34}

H.5.2

	1	2	3	4
1	5	10		
2		5	15	5
3				5

phương án cơ sở xuất phát

H.5.3

Bằng phương pháp góc tây-bắc trình bày ở Thí dụ 5.4.1 dưới đây đã tìm được phương án cơ sở để xuất phát cho ở bảng vận tải H.5.3 (phần chính của bảng, và chỉ ghi các x_{ij} là biến cơ sở). Giá trị mục tiêu là

$$f(x) = 10.5 + 0.10 + 7.5 + 9.15 + 20.5 + 18.5 = 410.$$

Bước lập 1

• Giải phương trình (5.28) ứng với các biến cơ sở x_{ij} để tính biến đối ngẫu p_i, q_j (khi đã lấy $p_1 = 0$):

$$x_{11}: \quad p_1 + q_1 = c_{11} = 10,$$

$$x_{12}: \quad p_1 + q_2 = c_{12} = 0,$$

$$x_{22}: \quad p_2 + q_2 = c_{22} = 7,$$

$$x_{23}: \quad p_2 + q_3 = c_{23} = 9,$$

$$x_{24}: \quad p_2 + q_4 = c_{24} = 20,$$

$$x_{34}: \quad p_3 + q_4 = c_{34} = 18.$$

Ta thu được vec tơ đối ngẫu hiện hành $(p_1, p_2, p_3, q_1, q_2, q_3, q_4) = (0, 7, 5, 10, 0, 2, 13)$.

• Viết phương trình (5.27) ứng với các biến không cơ sở x_{ij} để xác định $\bar{c}_{ij}$ âm nhất:

$$x_{13}: \quad \bar{c}_{13} = c_{13} - (p_1 + p_3) = 20 - (0+2) = 18,$$

$$x_{14}: \quad \bar{c}_{14} = c_{14} - (p_1 + q_4) = 11 - (0 + 13) = -2$$

$$x_{21}: \quad \bar{c}_{21} = c_{21} - (p_2 + q_1) = 12 - (7 + 10) = -5,$$

$$x_{31}: \quad \bar{c}_{31} = c_{31} - (p_3 + q_1) = 0 - (5 + 10) = -15$$

$$x_{32}: \quad \bar{c}_{32} = c_{32} - (p_3 + q_2) = 14 - (5 + 0) = 9,$$

$$x_{33}: \quad \bar{c}_{33} = c_{33} - (p_3 + q_3) = 16 - (5 + 2) = 9.$$

Vì $\bar{c}_{31}$ là âm nhất nên x_{31} là biến vào.

• Xác định biến ra như sau. Ô (3,1) ứng với biến vào ghép vào các ô ứng với biến cơ sở tạo thành một chu trình = ((3,1), (1,1), (1,2), (2,2), (2,4), (3,4)) duy nhất. Đánh dấu các ô trong chu trình bằng dấu + và - xen kẽ, bắt đầu từ ô "vào" (3,1) theo chiều tùy chọn, chẳng hạn chiều kim đồng hồ, xem H.5.4. Tính $\theta = \min \{x_{ij} : (i,j) \in V\} = \min \{x_{11}, x_{22}, x_{34}\} = \min \{5, 5, 5\} = 5$.

Vì ba biến cơ sở đều bằng θ , lấy biến nào làm biến ra cũng được, giả sử ta lấy x_{34} ra.

• Tiếp theo ta tính phương án cơ sở mới $\bar{x}$ theo (5.29) và được bảng ở H.5.5.

	1	2	3	4
1	5 $\rightarrow$	10 $+$		
2	$\uparrow$	5 $-$	15 $\rightarrow$	5 $+$
3	x_{31} $+$			5 $-$

H. 5.4

	1	2	3	4
1	0	15		
2		0	15	10
3	5			

phương án cơ sở $\bar{x}$

H. 5.5

Chú ý rằng phương án cơ sở mới $\bar{x}$ là suy biến vì có biến cơ sở bằng 0: $\bar{x}_{11} = \bar{x}_{22} = 0$ (tập ô sử dụng $G(\bar{x})$ có $4 < n + m - 1$ ô), nhưng thuật toán thế vị cứ tiến hành bình thường.

Giá trị mục tiêu mới là

$$\xi(\bar{x}) = 10.0 + 0.15 + 7.0 + 9.15 + 20.10 + 10.5 = 335.$$

Bước lặp 2

• Giải phương trình (4.28) ứng với các biến cơ sở $\bar{x}_{ij}$, vừa nhận được để tính các biến đối ngẫu hiện hành $\bar{p}_i, \bar{q}_j$ (lấy $\bar{p}_1 = 0$). Vì 5 biến cơ sở đầu vẫn như cũ (chỉ đổi biến x_{34} ra và x_{31} vào) nên chỉ phương trình cuối là khác Bước lặp 1:

$$\bar{x}_{31}: \quad \bar{p}_3 + \bar{q}_1 = 0.$$

Do đó 6 biến đối ngẫu vẫn nhận giá trị cũ, chỉ $\bar{p}_3 = -10$ nhận được từ phương trình cuối. Vậy $(\bar{p}_1, \dots, \bar{q}_4) = (0, 7, -10, 10, 0, 2, 13)$.

• Dùng (5.27) ứng với các biến không cơ sở $\bar{x}_{ij}$ để xác định cước phí hiện hành c'_{ij} âm nhất. Các phương trình ứng với hai hàng ô trên vẫn như ở Bước lặp 1 nên vẫn cho ta $c'_{13} = \bar{c}_{13} = 18$, $c'_{14} = \bar{c}_{14} = -2$, $c'_{21} = \bar{c}_{21} = -5$. Chỉ các phương trình ứng hàng ô thứ ba đổi vì $\bar{p}_3 = -10 \neq p_3 = 18$ và không còn phương trình ứng $\bar{x}_{31}$ mà thay bằng $\bar{x}_{34}$:

$$\bar{x}_{32}: \quad c'_{32} = c_{32} - (\bar{p}_3 + \bar{q}_2) = 14 - (-10 + 0) = 24,$$

$$\bar{x}_{33}: \quad c'_{33} = c_{33} - (\bar{p}_3 + \bar{q}_3) = 16 - (-10 + 2) = 24,$$

$$\bar{x}_{34}: \quad c'_{34} = c_{34} - (\bar{p}_3 + \bar{q}_4) = 18 - (-10 + 13) = 15.$$

Do đó $c'_{21} = \bar{c}_{21} = -5$ là âm nhất nên $\bar{x}_{21}$ là biến vào.

• Xác định biến ra: Ô (2,1) tạo thành với các ô cơ sở chu trình duy nhất $U = \{(2,1), (1,1), (1,2), (2,2)\}$. Ta có

$$\theta = \min \{\bar{x}_{11}, \bar{x}_{22}\} = 0,$$

	1	2	3	4
1		15		
2	0	0	15	10
3	5			

nên $\bar{x}_{11}$ và $\bar{x}_{22}$ có thể làm biến ra, ta lấy tùy ý $\bar{x}_{11}$ ra.

• Theo (5.29) ta được phương án cơ sở mới cho ở H.5.6.

phương án cơ sở thu được ở Bước lặp 2

H.5.6

Nhận xét rằng vì $\theta = 0$ nên tuy biến cơ sở đã đổi nhưng nghiệm cơ sở vẫn giữ nguyên. Sự kiện này tương ứng với phép xoay suy biến.

Bước lặp 3

Thực hiện các công đoạn như trên ta được kết quả cho ở H.5.7. Ở đây về bảng vận

	$q_1 = 5$	$q_2 = 0$	$q_3 = 2$	$q_4 = 13$	
$p_1 = 0$	5	15	18	2	15
$p_2 = 7$	0	0	15	10	25
$p_3 = -5$	5	19	19	10	5
	5	15	15	10	

phương án cơ sở thu được ở Bước lặp 3

H. 5.7

tải (phần chính) nhưng có ghi chú rõ giá trị các biến đổi ngẫu nhiên hành ở rìa bên trái và phía trên bảng, cước phí thu gọn hiện hành tính theo (5.27) ở mép dưới bên trái ô và chu trình để xác định biến ra. (Lượng phát và lượng thu ghi ở rìa phải và dưới bảng.)

Bước lặp 4

	$q_1 = 5$	$q_2 = 0$	$q_3 = 2$	$q_4 = 11$	
$p_1 = 0$	5	5	18	10	15
$p_2 = 7$	0	10	15	2	25
$p_3 = -5$	5	19	19	12	5
	5	15	15	10	

phương án cơ sở thu được ở Bước lặp 4

H. 5.8

Theo kết quả của Bước lặp 3, x_{14} vào và x_{24} ra. Ta nhận được nghiệm cơ sở mới ở H.5.8.

Khi tính các cước phí thu gọn mới theo (5.27) ta được các số ghi

ở mép dưới bên trái các ô. Chúng đều dương nên nghiệm hiện hành là tối ưu. Vậy nghiệm cơ sở tối ưu nhận được có các biến cơ sở là $x_{12} = 5$, $x_{14} = 10$, $x_{21} = 0$, $x_{22} = 10$, $x_{23} = 15$, $x_{31} = 5$. Để tính được mục tiêu tối ưu là 315 (tổng cước vận tải).

5.4 TÌM PHƯƠNG ÁN CƠ SỞ ĐỂ XUẤT PHÁT

Để giải được bài toán vận tải, bằng thuật toán thế vị chẳng hạn, ta cần phải có một phương án cơ sở chấp nhận được để xuất phát. Khi giải quy hoạch tuyến tính tổng quát việc này thường chiếm đến gần nửa thời gian giải cả bài toán, vì nó là pha một trong thuật toán đơn hình hai pha. Riêng đối với bài toán vận tải, người ta có nhiều phương pháp riêng, hiệu quả, để tìm phương án xuất phát. Ở đây ta trình bày một số phương pháp hay dùng nhất.

5.4.1 Quy tắc góc tây bắc

Quy tắc góc tây bắc (northwest - corner rule) bắt đầu bằng việc phân phối lượng hàng lớn nhất có thể được vào ô (1,1) (nằm ở góc tây bắc của bảng vận tải). "Có thể được" nghĩa là phải thỏa ràng buộc ((5.2) và (5.3)). Vậy

$$x_{11} = \min \{s_1, d_1\}.$$

Giả sử $x_{11} = s_1$, tức là đã dùng hết cung lượng hàng có ở điểm phát 1, nên hàng 1 trong bảng phải xóa đi. Đồng thời ta chỉnh lại lượng thu ở điểm thu 1, bây giờ chỉ còn là $d'_1 = d_1 - x_{11}$. Nếu $x_{11} = d_1$, thì tương tự ta xóa cột 1 và chỉnh lại lượng phát ở hàng 1 là $s'_1 = s_1 - x_{11}$. Nếu $x_{11} = s_1 = d_1$ thì xóa một trong hai, hàng 1 hoặc cột 1, đều được.

Đối với bảng T' còn lại sau khi xóa và chỉnh số liệu ta lại thực

hiện thủ tục phân phối lượng hàng tối đa có thể được vào góc tây bắc mới. Vì bảng có m hàng, n cột mà mỗi bước ta xóa một nên sau $m+n-1$ bước, thủ tục phải kết thúc. Rõ ràng rằng quy tắc góc tây bắc không chú ý gì đến tính tối ưu, tức là đến cước phí.

THÍ DỤ 5.4.1. Tìm phương án cơ sở xuất phát cho bảng vận tải ở H.5.2, Thí dụ 5.3.1.

$$x_{11} = \min \{s_1, d_1\} = \min \{15, 5\} = 5.$$

Ta xóa cột 1 và chỉnh lại lượng phát còn ở điểm phát 1 là $s'_1 = 15 - 5 = 10$ để được bảng T' ở H.5.9

10			10
			25
			5
15	15	10	

H. 5.9

Bước hai. Ta thấy

$$x'_{11} = \min \{10, 15\} = 10,$$

nên phải xóa hàng 1 và chỉnh lại lượng cần thu ở cột 1 là $15 - 10 = 5$.

Cứ tiếp tục ta sẽ nhận được phương án cơ sở chấp nhận được cho ở H.5.3, mà ta đã dùng để xuất phát thuật toán thế vị ở Thí dụ 5.3.1.

5.4.2 Phương pháp cực tiểu cước phí

Để có đường đi "rẻ hơn", tức là nhiều khả năng gần với nghiệm tối ưu hơn, so với đường đi trên bảng vận tải cho bằng quy tắc góc tây bắc, người ta dùng thủ tục phân phối lượng hàng tối đa vào ô có cước phí nhỏ nhất, thay vì ô ở góc tây bắc. Đó là nội dung *phương pháp cực tiểu cước phí* (the least-cost method). Việc xóa bỏ hàng (hoặc cột) đã phát hết hàng (hoặc tương ứng, thu đủ hàng) và hiệu chỉnh lượng hàng ở cột (hoặc hàng, tương ứng) giữ lại cũng như ở quy tắc góc tây bắc.

THÍ DỤ 5.4.2. Lại xét bảng vận tải ở H.5.2. Ta thấy cước phí nhỏ nhất, là $c_{12} = c_{31} = 0$. Ta chọn một tùy ý trong hai ô, chẳng hạn ô (1,2) để đưa lượng hàng tối đa vào, tức là

$$x_{12} = \min \{s_1, d_2\} = \min \{15, 15\} = 15.$$

Hàng 1 và cột 2 đều thỏa tiêu chuẩn được xóa. Ta chọn tùy ý, chẳng hạn cột 2 và chỉnh lại lượng phát còn ở hàng 1 là $s'_1 = 15 - 15 = 0$. Ở bảng còn lại, ô (3,1) có cước phí nhỏ nhất $c_{31} = 0$ nên lại phân phối tối đa vào đó:

$$x_{31} = \min \{s_3, d_1\} = \min \{5, 5\} = 5.$$

Hàng 3 và cột 1 đều có thể xóa, ta xóa hàng 3 và chỉnh lại lượng thu ở cột 1 là $d'_1 = 5 - 5 = 0$.

Trong các ô còn lại, $c_{23} = 9$ là cước phí nhỏ nhất.

$$x_{23} = \min \{s_2, d_3\} = \min \{25, 15\} = 15.$$

Ta xóa bỏ cột 3 và chỉnh lượng phát ở hàng 2 là $s'_2 = 25 - 15 = 10$.

Tiếp theo, trong các ô còn lại $c_{11} = 10$ là cước phí nhỏ nhất. Ta có

$$x_{11} = \min \{s'_1, d'_1\} = \min \{0, 0\} = 0.$$

Trong hàng 1 và cột 1, xóa tùy ý cột 1. Lượng phát ở hàng 1 đương nhiên vẫn là 0, không phải chỉnh. Biến cơ sở còn lại (nằm ở cột 4 chưa xóa) rõ ràng phải lấy giá trị là $x_{14} = 0$, $x_{24} = 10$. Ta được phương án cơ sở chấp nhận được ở H.5.10. Tổng cước phí của phương án này là $10.0 + 0.15 + 11.0 + 9.15 + 20.10 + 0.5 = 335$.

10	0	0	15	20	11	0	15	
12		7		9	15	20	10	25
0	5	14		16		18		5
	5		15		15		10	

H. 5.10 phương án cơ sở thu được ở bằng phương pháp cực tiểu cước phí

So với phương án cơ sở nhận được bằng quy tắc góc tây bắc ở H.5.3 với cước phí tổng là 410 USD thì phương án này rẻ hơn.

Chọn cước phí nhỏ nhất trên toàn bảng như trên có thể khó nhìn. Để tiện hơn người ta có thể vận dụng thủ tục chỉ chọn ô có cước phí nhỏ nhất ở hàng (cột) đầu. Trường hợp này phương pháp có tên là *phương pháp cực tiểu cước phí theo hàng (cột)*. Để khỏi lẫn, phương pháp chính trên đây có tên cụ thể hơn là *phương pháp cực tiểu cước phí toàn bảng*.

5.4.3 Phương pháp xấp xỉ Vogel

Phương pháp xấp xỉ Vogel (Vogel's approximation method, gọi tắt là VAM) có nội dung ở mỗi bước gồm 3 đoạn như sau.

Đoạn 1. Tính *tiền phạt* (penalty) cho mỗi hàng và mỗi cột. Tiền phạt là hiệu của cước phí thấp nhì và thấp nhất trong hàng (cột).

Đoạn 2. Chọn hàng hoặc cột có tiền phạt lớn nhất. Nếu có hơn một ứng viên thì lấy tùy ý. Lấy ô có cước phí nhỏ nhất trong hàng hoặc cột đã chọn và phân phối vào đó lượng hàng tối đa có thể được theo ràng buộc. Xóa bỏ hàng (hoặc cột) đã phát (hoặc thu) hết hàng và chỉnh lại lượng thu (hoặc phát) của cột (hoặc hàng) còn lại. Nếu có nhiều ứng viên như nhau thì chỉ xóa một.

Đoạn 3. • Nếu còn đúng một hàng hoặc một cột chưa xóa thì ngừng.

• Nếu chỉ còn một hàng (hoặc cột) có lượng phát (thu) dương là chưa xóa thì chọn biến cơ sở trong hàng (hoặc cột) này ở ô có cước phí cực tiểu.

• Nếu tất cả các hàng và cột chưa xóa có lượng hàng (phát hoặc thu) bằng 0, thì chọn biến cơ sở trong đó theo cước phí cực tiểu và cho biến cơ sở giá trị 0. Ngừng.

• Trong trường hợp khác (3 khả năng trên) thì tính lại tiền phạt

cho các hàng và cột còn lại và có lượng hàng (phát hoặc thu) dương, rồi quay lại Đoạn 2.

THÍ DỤ 5.4.3. Lại xét bảng vận tải ở H.5.2, Thí dụ 5.4.1.

tiền phạt

hàng				
10	10	0	20	11
2	12	7	9	20
14	0	5	14	16

5 15 15 10

10 7 7 7

H. 5.11

s_i

15

25

5

tiền phạt
cột

Bước 1 Tính tiền phạt cho mỗi hàng và mỗi cột ta được H.5.11. Sang Đoạn 2, ta thấy hàng 3 có tiền phạt 14 là lớn nhất. Vì $C_{31} = 0$ là cước phí nhỏ nhất trên hàng này, ta phân phối vào ô (3,1) lượng hàng

$$x_{31} = \min \{s_3, d_1\} = 5.$$

Hàng 3 và cột 1 cùng đủ tiêu chuẩn để xóa, ta xóa bất kỳ cột 1. Chính lại $s'_3 = 5 - 5 = 0$.

Bước 2. Tính tiền phạt cho bảng còn lại.

Chú ý rằng hàng 3 có lượng phát bằng 0 nên không cần tính. Hàng 1 và cột 3 có tiền phạt lớn nhất là 11 (xem H.5.12). Ta lấy tùy ý cột 3. Ở đây cước phí nhỏ nhất là $C_{23} = 9$. Ta phân phối vào ô (2,3) lượng hàng

tiền phạt hàng

11

2

	0	20	11
7	7	9	20
15	15	15	10
7	7	11	9

s_i

15

25

tiền phạt cột

H. 5.12

$$x_{23} = \min \{s_2, d_3\} = 15$$

và xóa cột 3. Chính lại lượng phát ở hàng 2 là $s'_2 = 25 - 15 = 10$.

Áp dụng VAM ta được các bước tiếp theo như sau (đọc giả tự kiểm tra lại)

Bước 3. $x_{22} = 10$. Xóa bỏ hàng 2.

Bước 4. $x_{12} = 5$. Xóa cột 2.

Bước 5. $x_{14} = 10$. Xóa cột 1. Còn lại $x_{34} = 0$.

Tổng cước phí là 315 USD đã là mục tiêu tối ưu. Đây là may mắn ở thí dụ này, nên phương án cơ sở chấp nhận được xuất phát đã là tối ưu. Tuy nhiên, người ta thấy rằng phương pháp xấp xỉ Vogel luôn cho phương án đã rất gần tối ưu rồi.

Chú ý thêm rằng trong thuật toán xấp xỉ Vogel trình bày trên đây, nếu có nhiều ứng viên cùng đủ tiêu chuẩn ở một thủ tục nào đó, ta chọn bất kỳ. Tuy nhiên việc chọn này có thể ảnh hưởng lớn đến chất lượng phương án cơ sở nhận được cuối cùng.

Chẳng hạn, ở bước 2 trên đây nếu ta chọn hàng 1 thay vì cột 3 thì tính toán lại các bước sau đó ta sẽ thấy đi đến phương án cơ sở là $x_{12} = 15$, $x_{23} = 15$, $x_{24} = 10$, $x_{31} = 5$ và các biến khác bằng 0. Cước phí tổng cộng là 335 USD, chưa phải tối ưu. Trong [Reinfeld–Vogel 1958] có trình bày quy tắc chọn để được phương án cơ sở tối khi có nhiều ứng viên.

Cả ba phương pháp tìm phương án cơ sở xuất phát trên đây đều thực sự cho ta một phương án cơ sở chấp nhận được. Tính chấp nhận được là rõ ràng vì ta đã đảm bảo các ràng buộc thỏa khi phân phối các lượng hàng vào các ô. Việc phương án đúng là phương án cơ sở tuy đã có thể cảm nhận được, nhưng ta vẫn chứng minh chặt chẽ như sau.

MỆNH ĐỀ 5.4.1. *Các phương án tìm được theo quy tắc góc tây bắc, phương pháp cực tiểu cước phí và phương pháp xấp xỉ Vogel đều là phương án cơ sở chấp nhận được.*

CHỨNG MINH. Ta chứng minh đây là phương án cơ sở, tức là không chứa chu trình. Bảng quy nạp toán học theo k là tổng số hàng và cột của bảng vận tải. Rõ ràng mệnh đề đúng cho bảng vận tải chỉ có một ô, tức là $m = n = 1, k = 2$. Giả sử mệnh đề đúng cho bảng có kích thước $m(n-1)$ hoặc $(m-1)n$. Ta chứng minh nó cũng đúng cho bảng cỡ mn . Giả sử ô được phân phối hàng đầu tiên là (i_1, j_1) . Sau đó ta phải xoá hàng i_1 hoặc cột j_1 . Không mất tổng quát giả sử ta xoá cột j_1 và thu được bảng T' có tổng số hàng và cột là $m+n-1$. Thủ tục sẽ được lặp lại cho bảng T' . Theo giả thiết quy nạp, đối với bảng T' ta được $m+n-2$ ô cơ sở không chứa chu trình. Phương án thu được của thuật toán sẽ có các ô cơ sở (ứng với biến cơ sở) là $m+n-2$ ô này và thêm ô (i_1, j_1) rõ ràng cũng không chứa chu trình, vì trong tập các ô này, ô (i_1, j_1) là ô duy nhất nằm ở cột j_1 . $\square$

5.5 CÁC BÀI TOÁN VẬN TẢI MỞ RỘNG

Trong mục này ta xét một số bài toán vận tải mở rộng và trình bày cách giải chúng. Cách giải chủ yếu là đưa về xét bài toán vận tải đóng "tiêu chuẩn" rồi suy ra nghiệm của bài toán cần giải.

5.5.1 Bài toán vận tải với ràng buộc bất đẳng thức

Trong thực tế cũng hay gặp các bài toán với ràng buộc nhẹ hơn bài toán vận tải (5.1)-(5.4) là các điểm phát không cần phát hết hàng và các điểm thu có thể thu nhiều hơn nhu cầu. Dạng toán học sẽ là

$$\min \sum_{i=1}^n \sum_{j=1}^m c_{ij} x_{ij}, \quad (5.30)$$

$$\sum_{j=1}^m x_{ij} \leq s_i, \quad i = 1, \dots, m, \quad (5.31)$$

$$\sum_{i=1}^m x_{ij} \geq d_j, j = 1, \dots, n, \quad (5.32)$$

$$x_{ij} \geq 0, i = 1, \dots, m, j = 1, \dots, n. \quad (5.33)$$

Bằng cách tương tự như chứng minh Định lý 5.2.1 ta sẽ thu được

ĐỊNH LÝ 5.5.1 (Tồn tại nghiệm tối ưu). *Điều kiện cần và đủ để bài toán vận tải với ràng buộc bất đẳng thức (5.30)-(5.33) có phương án tối ưu là*

$$\sum_{i=1}^m s_i \geq \sum_{j=1}^n d_j. \quad (5.34)$$

Tùy theo điều kiện (5.34) thỏa chặt hay không ta có cách giải bằng thuật toán thế vị như sau.

Nếu (5.34) thỏa chặt, tức là ta có điều kiện cân bằng thu phát, thì có thể chứng minh được rằng mọi phương án chấp nhận được của bài toán (5.30) – (5.33) đều phải thỏa chặt các ràng buộc (5.31), (5.32). Vì vậy, có thể thay các ràng buộc này bởi các ràng buộc đẳng thức tương ứng mà bài toán không đổi. Lúc này ta nhận được bài toán vận tải đóng "tiêu chuẩn" và giải được bằng thuật toán thế vị.

Nếu (5.34) thỏa không chặt, tức là với dấu $>$, thì có thể chứng minh được rằng bài toán vận tải (5.30) – (5.33) luôn có phương án tối ưu thỏa (5.32) với dấu $=$. Vì vậy có thể thay bài toán (5.30)-(5.33) bởi bài toán với dấu $=$ ở (5.32) tức là bài toán vận tải mở. Như đã nói ở Mục 5.1.1, bài toán này có thể được đưa về bài toán đóng bằng cách thêm vào điểm thu giả $n+1$ với lượng thu

$$d_{n+1} = \sum_{i=1}^m s_i - \sum_{j=1}^n d_j,$$

cũng tức là đã thêm biến $x_{in+1} \geq 0$, $i = 1, \dots, m$, với hệ số ở hàm mục tiêu $c_{in+1} = 0$. Lúc này ta có thể giải bằng thuật toán thế vị.

5.5.2 Bài toán vận tải nhanh

Đây là bài toán vận tải với mọi ràng buộc như đã xét, tức là (5.2)–(5.4), nhưng cước phí vận chuyển chính là thời gian vận chuyển cho bằng ma trận

$$\tau = (t_{ij})_{m \times n},$$

trong đó t_{ij} là thời gian vận chuyển hàng từ điểm phát i đến điểm thu j (dù lượng hàng là bao nhiêu, miễn là thỏa ràng buộc, không phải một đơn vị hàng). Bài toán đặt ra là tìm phương án chấp nhận được để vận chuyển nhanh nhất. Như vậy ràng buộc vẫn là (5.2)–(5.4), còn mục tiêu là

$$\min_x \max_{(i,j) \in G(x)} t_{ij}, \quad (5.35)$$

hoặc viết tương đương là

$$\max_{(i,j) \in G(x)} t_{ij} \rightarrow \min.$$

ở đây max lấy theo mọi ô sử dụng của phương án x đang xét, vì vậy để thật rõ cũng hay viết là

$$t_x := \max_{(i,j) \in G(x)} t_{ij}.$$

Min lấy theo mọi phương án chấp nhận được.

Nhận xét rằng bài toán vận tải nhanh, hoặc gọi khác là bài toán vận tải theo thời gian, (5.2), (5.3), (5.4), (5.35) tuy là mở rộng trực tiếp từ bài toán vận tải “tiêu chuẩn”, một trường hợp rất riêng của quy hoạch tuyến tính, nhưng lại không phải quy hoạch tuyến tính. Tuy vậy, do dạng đặc biệt của nó, người ta vận dụng cách giải bài

toán vận tải "tiêu chuẩn" (chẳng hạn thuật toán thế vị) như sau.

Xuất phát. Tìm một phương án cực biên x^0 bằng một trong các cách đã biết. Chú ý rằng nếu phương pháp có liên quan đến cước phí thì ở đây là các thời gian t_{ij} .

Tính

$$t_x^0 = \max_{(i,j) \in G(x^0)} t_{ij},$$

Đoạn 1. Lập bài toán phụ là bài toán vận tải "tiêu chuẩn" với ma trận cước phí $c = (c_{ij})$ xác định bởi

$$c_{ij} = \begin{cases} 0 & \text{nếu } t_{ij} < t_x^0, \\ 1 & \text{nếu } t_{ij} \geq t_x^0. \end{cases}$$

Đoạn 2. Giải bài toán phụ, bằng thuật toán thế vị chẳng hạn, để nhận được phương án cực biên tối ưu x^k sau khi đi qua các phương án $x^1, \dots, x^{k-1}$ (ở các bước lặp của thuật toán thế vị). Nếu trong các ô sử dụng $G(x^k)$ của x^k có ô (i,j) với $c_{ij} = 1$ thì x^k là phương án tối ưu của bài toán vận tải nhanh. Thời gian tối ưu là

$$t^* = \min \{t_x^0, t_{x^1}, \dots, t_{x^k}\}.$$

Nếu, ngược lại, các ô của $G(x^k)$ đều có $c_{ij} = 0$ thì quay lại bước xuất phát với phương án xuất phát là x^k .

THÍ DỤ 5.5.1. Giải bài toán vận tải nhanh ở H.5.13, có 3 điểm phát, 3 điểm thu.

Xuất phát. Trước hết nhận xét là bài toán có dạng cung vượt cầu (phát vượt thu) vì

$$\sum_{i=1}^3 s_i = 180 > 160 = \sum_{j=1}^3 d_j.$$

t_{ij} d_j	40	70	50	20
s_i				
50	3	5	6	0
70	5	1	3	0
60	4	5	8	0

H. 5.13

Vì vậy ta phải đưa vào điểm thu giả 4 với lượng thu

$$d_4 = \sum_{i=1}^3 s_i - \sum_{j=1}^3 d_j = 20, \text{ với } t_{i4} = 0, \\ i = 1, 2, 3.$$

Tiếp theo ta tìm phương án cực biên xuất phát theo phương pháp cực tiểu cước phí toàn bảng và được x^0 cho ở H.5.14.

$$t_x^0 = \max \{3, 6, 1, 3, 8, 0\} = 8.$$

Đoạn 1. Lập bài toán phụ với ma trận cước phí có

$$c_{ij} = \begin{cases} 0 & \text{nếu } t_{ij} < 8, \\ 1 & \text{nếu } t_{ij} \geq 8, \end{cases}$$

tức là

$$C = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

3	5	6	0
40		10	
5	1	3	0
	70	0	
4	5	8	0
		40	20

Phương án cơ sở thu được theo phương pháp cực tiểu cước phí

H. 5.14

Đoạn 2. Giải bài toán phụ bằng thuật toán thế vị xuất phát từ x^0 .

Bước 1. Đoạn 1. Xác định biến vào.

- Tính biến đối ngẫu hiện hành (từ (5.28) cho các biến cơ sở x_{ij}^0)

$$x_{11}^0: \quad p_1 + q_1 = 0,$$

$$x_{13}^0: \quad p_1 + q_3 = 0,$$

$$x_{22}^0: \quad p_2 + q_2 = 0,$$

$$x_{23}^0: \quad p_2 + q_3 = 0,$$

$$x_{33}^0: \quad p_3 + q_3 = 1,$$

$$x_{34}^0: \quad p_3 + q_4 = 0,$$

ta được $p_1 = q_1 = p_2 = q_2 = q_3 = 0$, $p_3 = 1$, $q_4 = -1$.

- Tính cước phí hiện hành $\bar{c}_{ij}$ với các ô (i,j) không cơ sở ở x^0 theo (5.27).

$$\bar{c}_{12} = c_{12} - (p_1 + q_2) = 0 - 0 = 0,$$

$$\bar{c}_{14} = c_{14} - (p_1 + q_4) = 0 - (-1) = 1,$$

$$\bar{c}_{21} = c_{21} - (p_2 + q_1) = 0 - 0 = 0,$$

$$\bar{c}_{24} = c_{24} - (p_2 + q_4) = 0 - (-1) = 1,$$

$$\bar{c}_{31} = c_{31} - (p_3 + q_1) = 0 - 1 = -1,$$

$$\bar{c}_{32} = c_{32} - (p_3 + q_2) = 0 - 1 = -1.$$

Có hai cước phí âm nhất là $\bar{c}_{31} = \bar{c}_{32} = -1$. Ta chọn tùy ý x_{32} vào.

Đoạn 2. Xác định biến ra. Ô $(3,2)$ lập với các ô cơ sở của x^0 một chu trình duy nhất $U = \{(2,2), (2,3), (3,3), (3,2)\}$. Đánh dấu + và - xen kẽ từ ô vào $(3,2)$ ta được $U^- = \{(2,2), (3,3)\}$. Vậy

$$\theta = \min \{x_{22}^0, x_{33}^0\} = 40 = x_{33}^0,$$

nên x_{33} ra và ta nhận được phương án cực biên mới x^1 ở H.5.15.

Bước 2. Đoạn 1. Xác định biến vào.

• Tính biến đối ngẫu hiện hành. So với bước 1 chỉ bỏ phương trình ứng với x_{33}^0 đã ra và thêm phương trình:

$$x_{32}^1: p_3 + q_2 = c_{32} = 0.$$

Giải hệ phương trình mới ta được $p_1 = q_1 = p_2 = q_2 = p_3 = q_3 = 0, q_4 = 0$.

³ 40		⁶ 10	
	¹ 30	³ 40	
	⁵ 40		⁰ 20

phương án cơ sở x^1

H. 5.15

• Tính $\bar{c}_{ij}$ với các ô (i,j) không cơ sở ở x^1 ta thấy so với bước 1 chỉ bỏ phương trình cho $\bar{c}_{32}$ và thêm phương trình cho $\bar{c}_{33}$:

$$\bar{c}_{33} = c_{33} - (p_3 + q_3) = 1 - 0 = 1.$$

Tuy vậy, một số phương trình khác có thay đổi:

$$\bar{c}_{14} = c_{14} - (p_1 + q_4) = 0 - 0 = 0,$$

$$\bar{c}_{24} = c_{24} - (p_2 + q_4) = 0 - 0 = 0,$$

$$\bar{c}_{31} = c_{31} - (p_3 + q_1) = 0 - 0 = 0.$$

Ta đã có $\bar{c}_{ij} \geq 0$ với mọi ô (i,j) không cơ sở nên x^1 đã là phương án tối ưu cho bài toán phụ. Tuy nhiên các ô cơ sở của x^1 đều có $c_{ij} = 0$ nên x^1 chưa phải phương án tối ưu vận tải nhanh. Ta phải xuất phát lại từ x^1 .

Xuất phát. Từ x^1 . Ta có

$$t_x = \max \{3, 6, 1, 3, 5, 0\} = 6.$$

Đoạn 1. Lập bài toán phụ. Ta có

$$c_{ij} = \begin{cases} 0 & \text{nếu } t_{ij} < 6, \\ 1 & \text{nếu } t_{ij} \geq 6, \end{cases}$$

nên ta được

$$C = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Đoạn 2. Giải bài toán phụ bằng thuật toán thế vị, từ x^1 .

Bước 1. Đoạn 1. Xác định biến vào.

- Tính biến đối ngẫu hiện hành (ứng x^1)

$$x_{11}^1: \quad p_1 + q_1 = 0,$$

$$x_{13}^1: \quad p_1 + q_3 = 1,$$

$$x_{22}^1: \quad p_2 + q_2 = 0,$$

$$x_{23}^1: \quad p_2 + q_3 = 0,$$

$$x_{32}^1: \quad p_3 + q_2 = 0,$$

$$x_{34}^1: \quad p_3 + q_4 = 0,$$

ta được $p_1 = q_1 = 0$, $p_2 = p_3 = -1$, $q_2 = q_3 = q_4 = 1$.

- Tính $\bar{c}_{ij}$ với các ô (i,j) không cơ sở ở x^1 :

$$\bar{c}_{12} = c_{12} - (p_1 + q_2) = 0 - 1 = -1,$$

$$\bar{c}_{14} = c_{14} - (p_1 + q_4) = 0 - 1 = -1,$$

$$\bar{c}_{21} = c_{21} - (p_2 - q_1) = 0 + 1 = 1,$$

$$\bar{c}_{24} = c_{24} - (p_2 + q_4) = 0 - 0 = 0,$$

$$\bar{c}_{31} = c_{31} - (p_3 + q_1) = 0 - (-1) = 1,$$

$$\bar{c}_{33} = c_{33} - (p_3 + q_3) = 1 - (-1 + 1) = 1,$$

$$\bar{c}_{34} = c_{34} - (p_3 + q_4) = 0 - (-1 + 1) = 0.$$

Trong hai ứng viên x_{12}^1 và x_{14}^1 ứng cước phí âm nhất ta chọn tùy ý x_{12}^1 vào.

Đoạn 2. Xác định biến ra. Ô (1,2) lập với các ô cơ sở chu trình duy nhất $V = \{(1,2), (1,3), (2,3), (2,2)\}$. Đánh dấu ta được $V^- = \{(1,3), (2,2)\}$

$$\theta = \min \{x_{13}^1, x_{22}^1\} = \min \{10, 30\} = 10 = x_{13}^1,$$

nên x_{13}^1 ra. Ta được phương án cực biên mới x^2 ở H.5.16.

Bước 2. Đoạn 1. Xác định biến vào.

• Tính biến đối ngẫu hiện hành. So với bước trước chỉ bỏ phương trình ứng với x_{13} và thêm phương trình ứng với biến vào x_{12} :

$$x_{12}^2: p_1 + q_2 = c_{32} = 0.$$

3 40	5 10		
	1 20	3 50	
	5 40		0 20

phương án cơ sở x^2

H. 5.16

Giải hệ ta được $p_1 = p_2 = p_3 = q_1 = q_2 = q_3 = q_4 = 0$.

• Tính các $\bar{c}_{ij}$ với (i,j) không cơ sở ta thấy $\bar{c}_{ij} = c_{ij}$, $\forall$ (i,j) không cơ sở, nên đều không âm. Vậy x^2 là phương án tối ưu cho bài toán vận tải phụ.

Nhận thấy $c_{ij} = 0$ tại mọi ô cơ sở của x^2 nên x^2 chưa phải tối ưu thời gian.

Xuất phát lại từ x^2 . Ta có

$$t_x^2 = \max \{2, 5, 1, 3, 5, 0\} = 5.$$

Đoạn 1. Lập bài toán phụ. Ta có

$$c_{ij} = \begin{cases} 0 & \text{nếu } t_{ij} < 5, \\ 1 & \text{nếu } t_{ij} \geq 5, \end{cases} \text{ tức là } C = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}.$$

Đoạn 2. Giải bài toán phụ xuất phát từ x^2 .

Bước 1. Đoạn 1. Xác định biến vào.

• Tính biến đối ngẫu hiện hành ứng x^2 : ta được $p_1 = q_1 = p_3 = q_3 = 0$, $q_2 = q_3 = 1$, $p_2 = -1$.

• Tính $\bar{c}_{ij}$: ta được $\bar{c}_{13} = 0$, $\bar{c}_{14} = 0$, $\bar{c}_{21} = 2$, $\bar{c}_{24} = 1$, $\bar{c}_{31} = 0$, $\bar{c}_{33} = 0$.

Mọi $\bar{c}_{ij} \geq 0$ nên x^2 là phương án tối ưu cho bài toán phụ đang xét. Ta thấy trong các ô cơ sở của x^2 có $c_{12} = 1$, $c_{32} = 1$ nên x^2 cũng là phương án tối ưu của bài toán vận tải nhanh. Thời gian vận tải nhanh tối ưu là

$$t^* = \min \{t_x^0, t_x^1, t_x^2\} = 5.$$

5.5.3 Bài toán lập kho nhận hàng

(a) BÀI TOÁN. Giả sử cần vận chuyển hết hàng từ m điểm phát $i = 1, \dots, m$ với lượng hàng s_i tương ứng tới n điểm thu $j = 1, \dots, n$. Nhưng trong các điểm thu chỉ có các điểm j , với $j \in J_0 \subset J = \{1, \dots, n\}$ cần đúng lượng thu d_j đã cho, còn các điểm thu $j \in J \setminus J_0$ có lượng thu không hạn chế. Biết cước phí vận chuyển một đơn vị hàng từ i đến j là c_{ij} . Vấn đề đặt ra là tìm phương án vận tải hết hàng từ các điểm phát, thỏa các điểm thu $i, j \in J_0$, sao cho tổng cước phí nhỏ nhất. Như vậy ta cũng xác định được lượng thu d_j , $j \in J \setminus J_0$, mà ta phải lập kho chứa.

Mô hình toán học. Ta sẽ đưa bài toán trên về bài toán vận tải mở, cầu vượt cung, bằng cách gán thêm lượng thu d_j , $j \in J \setminus J_0$, như sau

$$d_j = \sum_{i=1}^m s_i, \quad \forall j \in J \setminus J_0.$$

Lúc này phải thêm điểm phát giả $m + 1$ với lượng phát là

$s_{m+1} = \sum_{j=1}^n d_j - \sum_{i=1}^m s_i$, và thêm cước phí vận chuyển từ $m + 1$ đến j , $j = 1, \dots, m$, là $c_{m+1,j} = 0$. Khi đó đã có cung bằng cầu và bài toán lập kho nhận hàng tương đương với bài toán vận tải đóng mở rộng

$$\min \sum_{i=1}^{m+1} \sum_{j=1}^n c_{ij} x_{ij},$$

$$\sum_{j=1}^n x_{ij} = s_i, \quad i = 1, \dots, m+1,$$

$$\sum_{i=1}^{m+1} x_{ij} = d_j, \quad j = 1, \dots, n,$$

$$x_{ij} \geq 0, \quad i = 1, \dots, m+1, j = 1, \dots, n.$$

Sau khi giải bài toán vận tải đóng mở rộng, bằng phương pháp thế vị chẳng hạn, được phương án tối ưu x^* , ta chỉ việc bỏ đi hàng $m+1$, ứng với điểm phát giả, tức là hàng gồm các lượng vận chuyển "hư cấu thêm". Lượng thu d_j^* cần lập kho chứa, $j \in J \setminus j_0$, sẽ là tổng theo cột j các x_{ij}^* sau khi đã bỏ hàng $m+1$:

$$d_j^* = \sum_{i=1}^m x_{ij}^*.$$

Nói riêng, nếu $d_j^* = 0$ thì không cần lập kho tại j này.

(b) THÍ DỤ 5.5.2. Cần vận tải một loại hàng từ bốn nhà máy sản xuất với lượng phát 60, 120, 80, 100 (tấn) tương ứng tới bốn kho. Trong đó hai kho đã ấn định lượng thu là $d_1 = 50$, $d_2 = 90$ (tấn). Còn hai kho có lượng thu không hạn chế. Cước phí (đơn vị hàng) là c_{ij} cho ở bảng H.5.17. Hãy tìm phương án tối ưu và trữ lượng hai kho phải lập để tổng cước vận tải nhỏ nhất.

c_{ij} d_j s_i	50	90		
60	2	3	5	4
120	4	5	8	9
80	3	7	10	11
100	7	7	6	9

Bài toán lập kho 3 và 4

H. 5.17

GIẢI. Gán lượng thu $d_3 = d_4 = \sum_{i=1}^4 s_i = 360$.

Đặt điểm phát giả thứ 5 với lượng phát $s_5 =$

$$\sum_{j=1}^4 d_j - \sum_{i=1}^4 s_i = 860 - 360 = 500 \text{ và các cước}$$

phí $c_{sj} = 0, j = 1, \dots, 4$.

Giải bài toán vận tải đóng mở rộng bằng phương pháp thế vị ta nhận được bảng vận tải tối ưu ở H.5.18.

Bỏ đi hàng cuối ta được phương án vận tải tối ưu x^* . Trừ lượng hai kho phải lập mới là

$$d_3^* = 30 + 30 + 100 = 160,$$

$$d_4^* = 60.$$

Tổng cước vận tải là

$$f^* = \sum_{i=1}^4 \sum_{j=1}^4 c_{ij} x_{ij}^* = 1980.$$

5.5.4 Bài toán điều xe

(a) BÀI TOÁN. Cần vận tải một loại hàng từ m điểm phát với lượng hàng tương ứng $s_i, i = 1, \dots, m$, đến n điểm thu với lượng thu $d_j, j = 1, \dots, n$, tương ứng, thỏa cân bằng thu

d_j s_i	50	90	360	360
60	2 0	3 0	5 0	4 60
120	4 0	5 90	8 30	9 0
80	3 50	7 0	10 30	11 0
100	7 0	7 0	6 100	9 0
500	0 0	0 0	0 200	0 300

Bảng vận tải tối ưu

H. 5.18

phát $\sum_i s_i = \sum_j d_j$. Biết khoảng cách là l_{ij} (giữa điểm phát i và điểm thu j). Giả sử ta không quan tâm cước chuyển hàng mà quan tâm việc điều xe không hàng về các điểm phát để có thể thậm chí chở hàng tiếp. Miễn sao cuối cùng chở đủ hàng đáp ứng cung s_i và cầu d_j , $\forall i, \forall j$ và tổng quãng đường xe chạy không là ngắn nhất.

Như vậy bài toán quy về phát xe không từ các điểm j (thực ra là xe đã chở hàng đến trước đó) về các điểm thu i sao cho tổng quãng đường ngắn nhất. Đây là bài toán vận tải đóng bình thường. Sau khi được phương án tối ưu (chạy xe không) ta kết hợp với phương án phát hàng ở đầu bài đã cho để sắp xếp cụ thể hành trình xe chạy.

THÍ DỤ 5.5.3. Giải bài toán điều xe cho ở bảng H.5.19. Ở đây đã cho một phương án vận tải chấp nhận được, không cần là phương án cơ sở (và cũng không tối ưu).

GIẢI. Chuyển các điểm phát thành điểm thu xe trống và điểm thu thành điểm phát xe trống. Giải bài toán vận tải đóng với mục tiêu

$$f = \sum_{i=1}^m \sum_{j=1}^n l_{ij} x_{ij} \rightarrow \min,$$

$l_{ij} \quad d_j$	81	39	27
s_i			
45	¹² 30	⁸ 15	¹⁰
32	¹⁰ 12	¹⁵	¹³ 20
28	⁷ 18	¹⁰ 10	⁸
42	¹⁵ 21	²⁰ 14	¹⁸ 7

Bài toán điều xe

H. 5.19

bằng phương pháp thế vị ta nhận được bảng điều xe trống tối ưu ở H.5.20.

Kết hợp với phương án phát hàng đã cho ở H.5.19 ta suy ra hành trình điều xe kết hợp cả chở hàng và đi không như sau.

(Gọi điểm phát i là s_i , điểm thu j là D_j)

l_{ij} d_j	s_i	45	32	28	42
81	12	10	7	15	42
39	8	15	10	20	
27	10	13	8	21	18
	6				

Phương án điều xe trống tối ưu
H. 5.20

- $S_1 \xrightarrow[15]{15T} D_2$ (nhận đủ từ S_1)
- $S_4 \xrightarrow[21]{21T} D_1$ (nhận đủ từ S_4)
- $S_2 \xrightarrow[12]{12T} D_1$ (nhận đủ từ S_2)
- $S_3 \xrightarrow[7]{7T} D_1$
- $S_1 \xrightarrow{6T} D_1 \xrightarrow[6]{6} S_2 \xrightarrow{6T} D_3$
 $\uparrow \quad \quad \quad \leftarrow \quad \quad \quad \downarrow$
- $S_1 \xrightarrow{14T} D_1 \xrightarrow{14} S_4 \xrightarrow{14T} D_2$
 $\uparrow \quad \quad \quad \leftarrow \quad \quad \quad \downarrow$
- $S_1 \xrightarrow{3T} D_1 \xrightarrow{3} S_2 \xrightarrow{3T} D_3 \xrightarrow{3} S_3 \xrightarrow{3T} D_2$
 $\uparrow \quad \quad \quad \leftarrow \quad \quad \quad \leftarrow \quad \quad \quad \downarrow$
- $S_2 \xrightarrow{11T} D_3 \xrightarrow{11} S_3 \xrightarrow{11T} D_1$
 $\uparrow \quad \quad \quad \leftarrow \quad \quad \quad \downarrow$
- $S_3 \xrightarrow{7T} D_2 \xrightarrow{7} S_1 \xrightarrow{7T} D_1 \xrightarrow{7} S_4 \xrightarrow{7T} D_3$
 $\uparrow \quad \quad \quad \leftarrow \quad \quad \quad \leftarrow \quad \quad \quad \downarrow$

Giải thích: Ta giải thích hành trình điều xe thứ 5 chẳng hạn. Điều lượng xe chờ đủ 6 (tấn) từ điểm phát 1 đến điểm thu 1, rồi cho tất cả xe chạy không về điểm phát 2 lấy đủ 6 (tấn) chạy đến điểm thu 3 rồi chạy không về điểm phát 1 thì chấm dứt.

Chú ý rằng có thể có nhiều cách điều xe khác nhau từ phương án vận tải ở H.5.19 và phương án điều xe tổng tối ưu ở H.5.20. Đồng thời thứ tự các khoản từ 1 đến 9 ở trên khi điều xe cũng không nhất thiết phải đánh số như vậy.

Qua thí dụ trên ta thấy bài toán điều xe rất dễ gặp trong thực tế. Chẳng hạn sau khi giải bài toán vận tải theo cước phí được phương án tối ưu. Nhưng ở phương án này không tính gì đến việc điều xe thế nào. Nếu coi như toàn bộ các lượng hàng cần chuyển cùng lúc thì nói chung tốn rất nhiều xe và sau đó lượng xe trống quay về cũng rất nhiều. Do đó sau khi có phương án tối ưu cước phí cũng hay phải giải bài toán điều xe.

5.6 PHƯƠNG PHÁP HUNGARY

Từ đầu Chương 4 đến giờ ta mới chỉ nói đến một phương pháp tổng quát để giải bài toán vận tải là phương pháp thế vị (phương pháp nhân tử). Ta đã biết đây chính là phương pháp đơn hình mạng áp dụng vào bài toán vận tải là trường hợp riêng. Đó là phương pháp đơn hình gốc. Mỗi bước lặp tương ứng một phương án cơ sở chấp nhận được, và qua mỗi bước lặp mục tiêu được cải thiện dần cho đến tối ưu thì thuật toán kết thúc.

Trong Mục 5.6 này ta sẽ trình bày phương pháp Hungary, là phương pháp tổng quát, giải bài toán vận tải, phổ biến thứ hai, sau phương pháp thế vị. Ta sẽ thấy đây chính là phương pháp đơn hình đối ngẫu áp dụng vào bài toán vận tải. Ở mỗi bước lặp ta có tương ứng

biến đổi ngẫu nhiên chấp nhận được và điều kiện độ lệch bù thỏa; cho đến bước lặp nào ta được phương án chấp nhận được thì đó chính là phương án tối ưu.

Điều lý thú là phương pháp này được nhà toán học Hungary Egervar công bố trong một bài báo bằng tiếng Hungary năm 1931, tức là 16 năm trước khi quy hoạch tuyến tính và phương pháp đơn hình được khai sinh bởi Dantzig, nhưng thế giới không biết đến cho đến năm 1953, khi nhà toán học Mỹ là Kuhn dịch bài báo ra tiếng Anh và phát triển phương pháp cho bài toán bổ nhiệm. Sau đó phương pháp được mở rộng cho bài toán vận tải. Như vậy là ý tưởng của phương pháp đơn hình đối ngẫu đã ra đời 16 năm trước khi phương pháp đơn hình ra đời, độc lập với nó. Đúng là tư tưởng lớn gặp nhau. Chính Kuhn đã gọi phương pháp là "Hungary" để nhấn mạnh công của nhà toán học nước này.

5.6.1 Phương pháp Hungary cho bài toán bổ nhiệm

(a) Ý tưởng của phương pháp

Ta nhắc lại bài toán bổ nhiệm đã nêu ở Thí dụ 5.1.1: Cần phân m việc cho n người. Tiền công (cước phí) việc i được làm bởi người j là c_{ij} . Tìm phương án để tổng tiền công là nhỏ nhất. Nếu cần thì thêm việc giả hoặc thêm người giả (tùy $m < n$ hay $n < m$), ta có thể coi là $m = n$. Mô hình toán học là như sau. đặt biến của bài toán là

$$x_{ij} = \begin{cases} 0 & \text{nếu việc } i \text{ không phân cho người } j, \\ 1 & \text{nếu việc } i \text{ phân cho người } j. \end{cases}$$

Bài toán trở thành: tìm phương án $x = (x_{ij})$ sao cho

$$\min z = \sum_{i=1}^n \sum_{j=1}^n c_{ij} x_{ij} \quad (5.36)$$

$$\sum_{j=1}^n x_{ij} = 1, i = 1, \dots, n, \quad (5.37)$$

$$\sum_{i=1}^n x_{ij} = 1, j = 1, \dots, n, \quad (5.38)$$

$$x_{ij} = 0 \text{ hoặc } 1. \quad (5.39)$$

Thay điều kiện cuối bởi $x_{ij} \geq 0$ ta được LP-nới lỏng của bài toán bổ nhiệm, chính là bài toán vận tải đặc biệt với $s_i = d_j = 1, \forall i, \forall j$.

MỆNH ĐỀ 5.6.1. Mọi phương án tối ưu của bài toán bổ nhiệm vẫn sẽ là tối ưu cho bài toán mới nhận được khi ta cộng vào mỗi hàng và mỗi cột của ma trận cước phí một hằng số thực (có thể khác nhau).

CHỨNG MINH. Giả sử cộng $-p_i$ vào hàng i và $-q_j$ vào cột j của ma trận cước phí, ta được $c'_{ij} = c_{ij} - p_i - q_j$, và hàm mục tiêu mới là

$$z' = \sum_i \sum_j c'_{ij} x_{ij} = \sum_i \sum_j c_{ij} x_{ij} - \sum_i p_i \sum_j x_{ij} - \sum_j q_j \sum_i x_{ij}.$$

Vì $\sum_j x_{ij} = \sum_i x_{ij} = 1$ nên $z' = z - \text{const}$, ở đây const là hằng số không phụ thuộc phương án. Do đó các phương án làm cực tiểu z thì cũng làm cực tiểu z' và ngược lại. $\square$

Từ mệnh đề trên dẫn đến

ĐỊNH NGHĨA 5.6.1. Hai ma trận $C = (c_{ij})_{m,n}$ và $C' = (c'_{ij})_{m,n}$ được gọi là tương đương (equivalent) ký hiệu là $C \sim C'$ nếu $c'_{ij} = c_{ij} + \alpha_i + \beta_j, i = 1, \dots, m, j = 1, \dots, n$, với α_i, β_j là các số thực bất kỳ.

Định nghĩa này cho ta ý tưởng để giải bài toán bổ nhiệm như sau. Cố gắng tạo ra ma trận tương đương C' với các c'_{ij} bằng 0 ở nhiều

chỗ sao cho lấy $x_{ij} = 1$ vào một số trong các ô đó (còn bằng 0 trong mọi ô khác) ta được một phương án chấp nhận được. Khi đó phương án này phải là tối ưu theo mục tiêu C' vì tổng cước phí bằng 0 là nhỏ nhất có thể rồi. Phương án cũng sẽ là tối ưu theo mục tiêu C do tính tương đương.

Do cấu trúc đặc biệt của bài toán bổ nhiệm, phương án là chấp nhận được khi mỗi hàng và mỗi cột đều có đúng một ô (i,j) với $x_{ij} = 1$, còn lại các x_{ij} khác phải bằng 0. Vì vậy ta cũng nói đây các ô 0 của ma trận vuông C là độc lập (independent) nếu mỗi hàng và mỗi cột của C chỉ có nhiều lắm là một ô của dãy. Tóm lại, để giải bài toán bổ nhiệm (5.36)-(5.39) ta phải tìm ma trận tương đương với C , mà có chứa đầy đủ n ô 0 độc lập. Lúc này chỉ việc đặt $x_{ij} = 1$ vào các ô này ta được phương án tối ưu.

(b) Thuật toán Hungary

Ta sẽ trình bày thuật toán trên một thí dụ. Để có khả năng áp dụng đa dạng, ta xét thuật toán cho bài toán với mục tiêu

$$\max z = \sum_i \sum_j c_{ij} x_{ij}. \quad (5.40)$$

Chẳng hạn c_{ij} là lượng sản phẩm sẽ làm được khi phân máy i cho người j thì ta gặp bài toán $\max z$ như vậy. Độc giả tự suy ra thuật toán cho bài toán tìm min z .

THÍ DỤ 5.6.1. Giải bài toán bổ nhiệm (5.37) – (5.40) với

Bước lập 0 (chuẩn bị). Lấy số cực đại ở mỗi cột trừ đi mỗi số trong cột kết quả ghi vào chỗ cũ trong cột. Ta thu được ma trận mới mà mỗi cột có ít nhất một số 0. Nhưng xét theo hàng thì có

$$C = \begin{bmatrix} 3 & 4 & 2 & 2 & 1 \\ 4 & 5 & 3 & 1 & 3 \\ 4 & 3 & 1 & 1 & 1 \\ 3 & 1 & 2 & 2 & 2 \\ 1 & 3 & 1 & 2 & 1 \end{bmatrix}$$

hàng chưa có 0. Nên ta lại lấy mỗi hàng trừ đi số nhỏ nhất trong hàng. Mỗi hàng

mỗi cột đều có 0. Để chọn ô độc lập trong các ô 0, đánh dấu * vào ô 0 bất kỳ ở cột 1. Tiếp đến cột 2, đánh * vào ô có 0 không cùng hàng với ô có * trước đó. Tiếp tục cho đến cột cuối ta được dãy ô số 0 độc lập, đã đánh dấu *

Bước lặp k+1. Giả sử đã qua k bước lặp và có C_k . Nếu ở C_k có đủ n số 0 với dấu * thì ngừng. đặt $x_{ij} = 1$ tại các ô có * và = 0 tại ô khác ta được phương án tối ưu.

$$C_0 = \begin{bmatrix} \oplus & & & + & \\ 1 & 1 & 1 & 0^* & 2 \\ 0^* \rightarrow 0' & 0 & 1 & 0 & \\ \uparrow & & & & \\ 0' & 2 & 2 & 1 & 2 \\ 1 & 4 & 1 & 0 & 1 \\ 3 & 2 & 2 & 0 & 2 \end{bmatrix} +$$

Trái lại thì đánh dấu + trên các cột có ô 0* *Bước 0 và đoạn 1 của bước 1.* và tách các cột đánh dấu + này ra.

Đoạn 1.

- Nếu trong các phần tử còn lại không tách (ở các cột không có dấu + ở trên) không có số 0 nào, sang đoạn 3.

- Nếu có 0 chưa tách ra thì chia hai trường hợp:

(α) hàng chứa 0 này cũng chứa 0* (đã tách);

(β) hàng chứa 0 này không chứa 0* (đã tách).

Trong trường hợp (α) đặt dấu 0' lên số 0 của hàng và đánh dấu + bên phải hàng để tách hàng ra, và bỏ dấu + trên cột có chứa 0* cùng hàng với 0' này (bằng cách khoanh tròn lại).

Trong trường hợp (β) đánh dấu 0' lên số 0 này và chuyển sang Đoạn 2. Lúc này ta cũng nói đoạn 1 kết thúc kiểu 1A. Ta cứ nghiên cứu lần lượt các số 0 chưa tách, nếu gặp liên tiếp các trường hợp (α) cho đến khi hết số 0 chưa tách thì ta cũng kết thúc Đoạn 1 và chuyển sang Đoạn 3, ta gọi kết thúc kiểu này là 1B (để vẽ sơ đồ khối mô tả thuật toán).

Ở thí dụ này ta được 0' ở ô (2,2) và đánh dấu + hàng 2, xóa + ở cột 1. Khi đó cột 1 trở thành chưa đánh dấu nên lại có 0 ở ô (3,1) ứng trường hợp (β) ta đánh dấu 0' và sang Đoạn 2 (kết thúc kiểu 1A).

Đoạn 2. Đi từ 0' ở hàng không có 0*, tức là ô (3,1) ở đây, vẽ một dây truyền các phần tử của C_k , từ 0' sang 0* trên cùng cột, tiếp đến 0' trên cùng hàng với 0* Dưới đây sẽ chứng minh rằng thủ tục này phải xác định, hữu hạn, và kết thúc tại 0'. Đổi lại các dấu 0' thành 0* trong dây chuyền và bỏ các dấu 0* cũ của dây truyền, đánh dấu lại, cột có 0*. Xóa bỏ tất cả các 0' của C_k và dấu + tách hàng hoặc cột ứng với nó. Trong thí dụ này không có 0' nào như vậy.

Ta thấy sau Đoạn 2 số 0* đã tăng một.
Ta kết thúc Bước lặp 1 và được C_1 .

$$C_1 = \begin{bmatrix} + & \oplus & + \\ 1 & 1 & 1 & 0^* & 2 \\ 0 & 0^* & 0' & 1 & 0 \\ 0^* & 2 & 2 & 1 & 2 \\ 1 & 4 & 1 & 0 & 1 \\ 3 & 2 & 2 & 0 & 2 \end{bmatrix} +$$

Bước lặp 2.

sau Bước lặp 1 và Đoạn 1 lặp 2

Đoạn 1. Đánh dấu 0' ở ô (2,3) khoanh dấu + ở cột 2 thành $\oplus$ (trường hợp (α)).

Đoạn 3. Chọn h bằng số nhỏ nhất trong các c_{ij} ở các ô chưa tách. Ở đây $h = 1$, lấy $c_{ij} - h$ với các hàng chưa tách rồi cộng h vào các cột đã tách rồi ta nhận được $C_k^{(1)}$ (tương đương C_k)

$$\xrightarrow{-h} \begin{bmatrix} + & \oplus & + \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0^* & 0' & 1 & 0 \\ -1 & 1 & 1 & 0 & 1 \\ 0 & 3 & 0 & -1 & 0 \\ 2 & 1 & 1 & -1 & 1 \end{bmatrix} \xrightarrow{+h} \begin{bmatrix} + & \oplus & + \\ 1 & 0 & 0 & 0^* & 1 \\ 1 & 0^* & 0' & 2 & 0 \\ 0^* & 1 & 1 & 1 & 1 \\ 1 & 3 & 0 & 0 & 0 \\ 3 & 1 & 1 & 0 & 1 \end{bmatrix} + = C_0^{(1)}$$

Đoạn 3

Xuất hiện tình huống áp dụng Đoạn 1.

Đoạn 1. Trước hết cột 3 chưa đánh dấu có 0 ở ô (4,3) trên hàng của nó chưa có 0* nên ta đánh dấu 0* ở đó và đánh dấu cột 3 bằng +. Vào đoạn 1, chỉ còn cột 5 chưa đánh dấu, có 0 ở ô (2,5), rơi vào trường hợp (α) ta đánh dấu 0' đánh dấu hàng 2 và khoanh $\oplus$ trên dấu + ở cột 2 (có 0* cùng hàng 2). Cột 2 chưa đánh dấu lại có 0 ở (1,2) là trường hợp α , đánh dấu 0' và + vào hàng 2, khoanh $\oplus$ vào cột 4. Cột 4 trở thành chưa đánh dấu lại có 0 ở (5,4) đánh dấu 0' ở đây

và + vào hàng 5, ô (5,4) rồi vào trường hợp β , sang Đoạn 2.

$$\begin{bmatrix} + & \oplus & + & \oplus & \\ 1 & 0' & 0 & 0^* & 1 \\ 1 & 0^* & 0 & 2 & 0' \\ 0^* & 1 & 1 & 1 & 1 \\ 1 & 3 & 0^* & 0 & 0 \\ 3 & 1 & 1 & 0' & 1 \end{bmatrix} \begin{matrix} + \\ + \\ + \\ + \\ + \\ + \end{matrix}$$

Cuối Đoạn 1, khi sắp sang giai đoạn 2 ở Bước lặp 2

Đoạn 2. Lập dãy truyền từ có 0' ở ô (5,4) như hình vẽ.

Đổi lại các 0' trong dãy truyền thành 0*, bỏ các 0* cũ của dãy truyền ta nhận được ma trận cước phí mới C_2 đã có đủ $n = 5$ số 0*. Thuật toán kết thúc.

$$C_2 = \begin{bmatrix} 1 & 0^* & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 & 0^* \\ 0^* & 1 & 1 & 1 & 1 \\ 1 & 3 & 0^* & 0 & 0 \\ 3 & 1 & 1 & 0^* & 1 \end{bmatrix}$$

$$\begin{bmatrix} + & & & & \\ 1 & 0' & 0 & 0^* & 1 \\ 1 & 0^* & 0 & 2 & 0' \\ 0^* & 1 & 1 & 1 & 1 \\ 1 & 3 & 0^* & 0 & 0 \\ 3 & 1 & 1 & 0' & 1 \end{bmatrix} \begin{matrix} + \\ + \\ + \\ + \\ + \\ + \end{matrix}$$

Đoạn 2 bước lặp 2

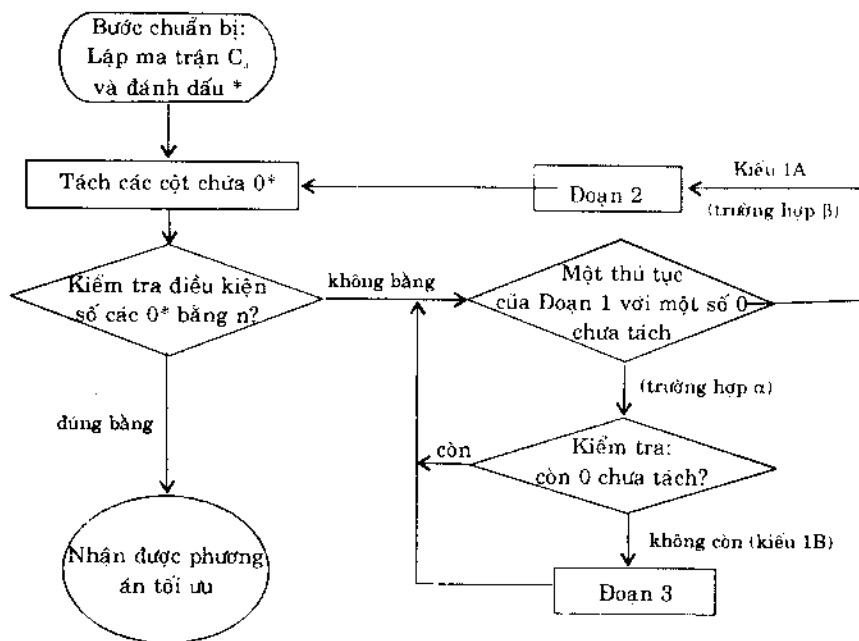
Phương án tối ưu là

$$x_{12} = x_{25} = x_{31} = x_{43} = x_{54} = 1, \quad x_{ij} = 0 \text{ với các } (i,j) \text{ khác.}$$

$$\text{Cước phí tối ưu là } c_{12} + c_{25} + c_{31} + c_{43} + c_{54} = 15.$$

(c) Sơ đồ khối của thuật toán Hungary cho bài toán bổ nhiệm

Thường với các thuật toán phức tạp người ta biểu diễn nó bằng sơ đồ khối. Sơ đồ của thuật toán Hungary được vẽ ở trang sau.



(d) Chứng minh thuật toán Hungary

• *Cơ sở chung.* Ở bước chuẩn bị ta đã xây dựng ma trận $C_0 \sim -C$ mà mọi phần tử không âm. Cho đến bước cuối ta được ma trận $C_l \sim -C$. Giả sử n số 0^* của C_l ở vị trí $1j_1, 2j_2, \dots, nj_n$. Khi đó

$$\sum_{k=1}^n (-C_{kj_k}^l) = \sum_{k=1}^n 0 = 0,$$

còn mọi tổng các phần tử khác của $-C_l$ đều không dương, nên phương án $x_{1j_1} = \dots = x_{nj_n} = 1, x_{ij} = 0 \forall ij$ khác là tối ưu của mục tiêu $-C_l$. Nhưng $-C_l \sim C$, nên phương án này là tối ưu cho mục tiêu C .

NHẬN XÉT 5.6.1. Để thấy rằng nếu ta xét bài toán bổ nhiệm với mục tiêu $\min \sum_i \sum_j c_{ij} x_{ij}$ thì ở bước chuẩn bị ta lấy ma trận C_0 có các phần tử là

$$c_{ij}^0 = c_{ij} - p_i - q_j,$$

$$p_i = \min_j c_{ij}, i = 1, \dots, n,$$

$$q_j = \min_i (c_{ij} - p_i), j = 1, \dots, n.$$

• *Tính đúng đắn của từng công đoạn.*

Bước chuẩn bị và Đoạn 1 là rõ ràng.

Xét Đoạn 2. Khi xây dựng các $0'$ và 0^* trong dây truyền ta có xác định duy nhất vì mỗi hàng có không quá một $0'$ vì đánh dấu ($'$) xong cũng phải tách hàng (bằng dấu $+$), và mỗi cột có không quá một 0^* (vì các 0^* độc lập).

Bây giờ ta chứng tỏ dây truyền không chứa chu trình và do đó dây truyền là hữu hạn. Giả sử dây truyền là

$$(i_1, j_1), (i_2, j_1), \dots, (i_s, j_{s-1}), (i_s, j_s),$$

ở đây $0'$ ở các ô thứ tự lẻ và 0^* ở ô chẵn. Xét hai $0'$ liên nhau (i_t, j_t) và (i_{t+1}, j_{t+1}) . Khi ô (i_t, j_t) có số 0 được đánh dấu $0'$ thì cột j_t là chưa tách. Nhưng cũng cột j_t này lại có 0^* ở ô (i_{t+1}, j_t) (cũng trong dây truyền) nên thực ra, theo thuật toán cột này đã được đánh dấu $+$ và tách. Do đó, theo Đoạn 1, dấu $+$ này đã được khoanh $\oplus$ để xoá đi khi tách hàng i_{t+1} . Điều đó có nghĩa là việc đánh $0'$ ở (i_{t+1}, j_{t+1}) là cùng lúc với việc tách hàng i_{t+1} này và trước khi đánh $0'$ ở (i_t, j_t) .

Từ lý luận cho hai ô có $0'$ liên nhau suy ra cho hai ô (i_λ, j_λ) và (i_μ, j_μ) , $\lambda < \mu$ có $0'$ trong dây truyền ta thấy ô (i_μ, j_μ) được đánh dấu

$0'$ trước $\hat{o}(i_\lambda, j_\lambda)$. Suy ra tất cả các $\hat{o}$ có $0'$ trong dây truyền phải khác nhau.

Bây giờ ta chỉ ra là dây truyền phải kết thúc bằng $0'$. Giả sử phản chứng là dây truyền dừng ở 0^* tại (i_{t+1}, j_t) .

Ô trước đó (i_t, j_t) chứa $0'$ nên cột j_t không tách. Nhưng cột này chứa 0^* ở (i_{t+1}, j_t) . Do đó dấu tách $+$ của nó đã bị xóa khi đánh dấu hàng i_{t+1} bởi đánh dấu $0'$ ở (i_{t+1}, j_{t+1}) nào đó, mâu thuẫn giả thiết là (i_{t+1}, j_t) với 0^* là cuối.

Ta lại chứng tỏ các 0^* trong dây truyền cũng khác nhau. Giả sử có hai $\hat{o}$ với 0^* trùng nhau. Khi đó hai $\hat{o}$ với $0'$ ngay sau đó (trong dây truyền) cũng trùng nhau, vô lý.

Theo thủ tục của Đoạn 2, xóa các $0'$ đổi thành 0^* và bỏ các dấu 0^* đang có trong dây truyền. Như vậy ta đã thêm được một $\hat{o} 0^*$. Ta chứng tỏ các 0^* vẫn độc lập. Với 0^* cũ (tức là không ở dây truyền) ở một $\hat{o}(i, j)$ nào đó, thì hàng i và cột j không có 0^* mới vì 0^* cũ không trong dây truyền. Tất nhiên do tính độc lập của các 0^* cũ, hàng i và cột j cũng không có 0^* cũ khác. Bây giờ xét 0^* mới bất kỳ, giả sử ở $\hat{o}(i_\lambda, j_\lambda)$. Rõ ràng 0^* này độc lập với các 0^* mới khác theo định nghĩa dây truyền được lập ở Đoạn 2. Còn các 0^* cũ, thì nó không có ở hàng và cột đã có 0^* mới như trên đã thấy. Tóm lại các $\hat{o} 0^*$ vẫn độc lập. Vậy Đoạn 2 là kết thúc hợp lý cho mỗi bước lặp.

Xét Đoạn 3. Khi vào Đoạn 3 tất cả các số 0 của C_k đều đã ở hàng và cột tách. Giả sử số 0^* của C_k là $m_k < n$. Nhận xét rằng tổng số hàng tách và cột tách của C_k đúng bằng m_k (xem Đoạn 1). Gọi m'_k là số hàng tách, m''_k là số cột tách và ρ là số phần tử chưa tách của C_k , thì

$$\rho = n(n - m_k) + m'_k m''_k \geq n(n - m_k) \geq n.$$

Vậy là có phần tử chưa tách. Gọi I'_k (I''_k tương ứng) là tập chỉ số các hàng tách (cột tách, tương ứng) của C_k . Sau thủ tục của Đoạn 3 ta được $C_k^{(1)}$ với

$$c_{ij}^{k1} = \begin{cases} c_{ij}^k - h & \text{với } i \in I'_k, j \in I''_k, \\ c_{ij}^k + h & i \notin I'_k, j \notin I''_k, \\ c_{ij}^k & \begin{cases} \text{hoặc } i \in I'_k, j \notin I''_k, \\ \text{hoặc } i \notin I'_k, j \in I''_k, \end{cases} \end{cases}$$

ở đây $h = \min_{i \in I'_k, j \in I''_k} c_{ij}^k$.

Chú ý rằng (theo thủ tục của Đoạn 1) tất cả các số 0 có đánh dấu đều ứng với các ô có $c_{ij}^{k1} = c_{ij}^k$ nên tất cả các số 0 có đánh dấu của C_k đều được chuyển sang C_k^1 (và dấu * cũng vẫn giữ nguyên, dấu ' cũng giữ nguyên). Các phần tử của C_k^1 vẫn không âm. Ngoài ra, trong số các phần tử không được tách ra của C_k^1 có số 0. Tiếp tục quá trình, hoặc sẽ dẫn đến áp dụng Đoạn 2 và tăng thêm một số 0 độc lập, hoặc là tách hết các số 0 và lại quay lại Đoạn 3. Khi đó số chung các hàng và cột tách của C_k^1 không đổi, vẫn là m_k , nhưng số hàng tách tăng, vì các số 0 mới lập của C_k^1 , theo thuật toán, phải nằm ở hàng tách, còn các hàng tách cũ vẫn là tách. Lại thực hiện Đoạn 3 sẽ dẫn đến $C_k^2 \sim C_k^1 \sim C_k$ và hoặc là tăng thêm một số 0 độc lập, hoặc là tách tất cả số 0 của C_k^2 bởi sự tăng hàng tách và giảm cột tách. Lại tiếp tục quá trình cho đến C_k^p mà sau Đoạn 1 là Đoạn 2 ngay. Khi đó ta lại tăng được một số 0 độc lập. Như vậy là Đoạn 3 là đúng đắn. Duy chỉ còn phải chứng minh là số p thủ tục để đến được Đoạn 2 vừa nói là không lớn quá. Cụ thể ta chứng minh

$$p \leq m_k + 1.$$

Khi chuyển từ C_k^m sang C_k^{m+1} (với $\mu = 0, \dots, p-1$) số hàng tách tăng không ít hơn 1, vì vậy nếu sau m_k vòng chỉ áp dụng Đoạn 3 và Đoạn 1, không đến được Đoạn 2 tất cả m_k số 0* của $C_k^{m_k}$ đều nằm ở các hàng tách. Đến lần m_k+1 , hàng có chứa số 0 mới của $C_k^{m_k+1}$ không thể có 0* (đã nằm ở hàng tách cả), nên ta được áp dụng Đoạn 2.

(e) Giải thích phương pháp Hungary như phương pháp đơn hình đối ngẫu

Để tiện so sánh với phương pháp đơn hình đối ngẫu cho bài toán mạng ở Mục 2.3 ta xét bài toán vận tải với mục tiêu min $\sum_i \sum_j c_{ij} x_{ij}$.

Quy hoạch tuyến tính đối ngẫu là

$$\begin{aligned} \max w &= \sum_{i=1}^n p_i + \sum_{j=1}^n q_j \\ p_i + q_j &\leq c_{ij}, \quad i, j = 1, \dots, n. \end{aligned}$$

Một phương án đối ngẫu chấp nhận được rõ ràng là

$$\begin{aligned} p_i &= \min_j c_{ij}, \quad i = 1, \dots, n, \\ q_j &= \min_i (c_{ij} - p_i), \quad j = 1, \dots, n. \end{aligned}$$

Theo Nhận xét 5.6.1, đây chính là cách chọn p_i, q_j để ta được ma trận C_0 với $c_{ij}^0 = c_{ij} - p_i - q_j$ ở bước chuẩn bị. Nếu $c_{ij}^0 = c_{ij} - p_i - q_j = 0$ thì lấy $x_{ij} = 1$ còn $x_{ij} = 0$ với ij khác, điều kiện độ lệch bù vẫn thỏa:

$$(c_{ij} - p_i - q_j) x_{ij} = 0.$$

Do đó nếu phương án $x = (x_{ij})_{n \times n}$ này chấp nhận được thì nó là tối ưu. (x muốn là chấp nhận được thì các ô $c_{ij}^0 = 0$ này phải độc lập.) Nếu x chưa chấp nhận được, ta phải tìm nghiệm đối ngẫu p_i, q_j chấp nhận được khác vẫn thỏa điều kiện độ lệch bù mà cho tương ứng x chấp

nhận được bằng thuật toán đã nêu, thực chất là phủ tất cả các số 0 bằng một số ít nhất hàng và cột. Chứng minh chi tiết điều này khá dài, ta không nêu ở đây.

5.6.2 Phương pháp Hungary cho bài toán vận tải

(a) Ý tưởng của phương pháp

Ta đã biết LP - nơi lỏng của bài toán bổ nhiệm (5.36) - (5.39), tức là thay điều kiện (5.39): $x_{ij} = 0$ hoặc 1 bởi $x_{ij} \geq 0$, chính là bài toán vận tải đặc biệt với $s_i = d_j = 1, \forall i, \forall j$.

NHÂN XÉT 5.6.2. Tập B các phương án chấp nhận được của bài toán bổ nhiệm chính là tập tất cả các đỉnh của V với V là tập các phương án chấp nhận được của LP - nơi lỏng của nó.

Thật vậy, rõ ràng $B \subset V$ và V là tập lồi đa diện. Các ô sử dụng (ij) của B có $x_{ij} = 1$ nằm ở đúng n hàng khác nhau, n cột khác nhau nên mỗi phương án của B đều không thể chứa chu trình và do đó là đỉnh của V (theo đều Hệ quả 5.2.4). Ngược lại, vì bài toán LP - nơi lỏng có các s_i và d_j nguyên nên các đỉnh của V đều có tọa độ nguyên (xem Định lý 3.3.4). Nhưng do $x_{ij} \geq 0$ và (5.37), (5.38), một đỉnh bất kỳ của V là bảng mà mỗi hàng và mỗi cột đều có đúng một phần tử bằng 1, các phần tử khác bằng 0, tức là phần tử của B.

Do nhận xét trên bài toán bổ nhiệm là tương đương với tìm đỉnh của LP - nơi lỏng của nó, là bài toán vận tải. Từ đây dẫn đến ý tưởng để mở rộng phương pháp Hungary ra cho bài toán vận tải tổng quát, có $s_i \neq 1, d_j \neq 1$.

Ở bước lập 0 (bước chuẩn bị) ta xây dựng phương án $x^0 = (x_{ij}^0)$ thỏa

$$\sum_{j=1}^n x_{ij}^0 \leq s_i, i=1, \dots, m,$$

$$\sum_{i=1}^m x_{ij}^0 \leq d_j, j=1, \dots, n, \quad (5.41)$$

(chú ý, tổng quát thì $m \neq n$) sao cho nếu mọi ràng buộc thỏa chặt thì x^0 là phương án tối ưu. Còn nếu có ràng buộc thỏa không chặt thì phải sang bước lặp 1.

Giả sử khi kết thúc bước lặp k , được phương án chấp nhận được x^k . Ta gọi

$$\Delta_k = \sum_{i=1}^m s_i + \sum_{j=1}^n d_j - 2 \sum_{i=1}^m \sum_{j=1}^n x_{ij}^k$$

là *độ lệch (tổng cộng)* của phương án x^k . Thuật toán sẽ dừng khi đến bước lặp $\Delta_d = 0$ (ở mọi bước trước ta luôn có $\Delta_k > 0$). Lúc này x^d là phương án tối ưu (và là đỉnh).

(b) Thuật toán Hungary cho bài toán vận tải

• *Bước lặp 0* (chuẩn bị). Từ ma trận cước phí $C = (c_{ij})$ lập ma trận C_0 bằng cách trừ mọi cột cho phần tử nhỏ nhất của cột, sau đó trừ mọi hàng (với phần tử mới) cho phần tử nhỏ nhất của hàng. Ta được C_0 có mọi phần tử không âm và mỗi hàng, mỗi cột đều có 0. Thủ tục này đúng như ở thuật toán cho bài toán bổ nhiệm. Nhưng tiếp theo ta không cần chọn ô độc lập để đánh dấu 0^* , mà xét ngay phương án tương ứng x^0 , vì ở đây không cần chỉ có một x_{ij} ở mỗi hàng, mỗi cột ở các ô độc lập 0^* nhận giá trị dương ($=1$). Ta chỉ cần trước hết là đảm bảo các ràng buộc bằng cách ta chọn là cho nhận các lượng hàng tối đa vào "góc tây bắc" theo nghĩa ưu tiên vào cột 1 trước, trong cột lại ưu tiên vào ô trên trước. Cụ thể như sau. Gọi i_{kj} ($k = 1, \dots, r_j$) là chỉ

số các ô 0 của cột j trong C_0 , đánh từ trên xuống. Khi đó theo ưu tiên góc tây bắc ta có công thức truy hồi để xác định x_{ij}^0 như sau. Trước hết ưu tiên cột 1, ta có (theo ưu tiên hàng trên)

$$x_{i1}^0 = \begin{cases} 0 & \text{nếu } i \neq i_{\lambda 1}, \lambda = 1, 2, \dots, r_1, \\ \min \left\{ a_i, b_1 - \sum_{\mu=1}^{i-1} x_{\mu 1}^0 \right\} & \text{nếu } i = i_{\lambda 1}. \end{cases}$$

Giả sử đã xác định xong cột $j-1$. Cột j sẽ là

$$x_{ij}^0 = \begin{cases} 0 & \text{nếu } i \neq i_{\lambda j}, \lambda = 1, 2, \dots, r_j, \\ \min \left\{ a_i - \sum_{\mu=1}^{j-1} x_{i\mu}^0, b_j - \sum_{\mu=1}^{i-1} x_{\mu j}^0 \right\} & \text{nếu } i = i_{\lambda j}. \end{cases}$$

Như vậy rõ ràng $x_{ij}^0 \geq 0, \forall i, j$. Nếu $\Delta_0 = 0$ thì x^0 là nghiệm tối ưu. Nếu $\Delta_0 > 0$, chuyển sang bước lặp 1.

Mỗi bước lặp $k > 0$ đều là sự áp dụng có thể nhiều lần đoạn 1, đoạn 3 và khi tới đoạn 2 thì kết thúc. Giả sử đã xong bước lặp k và có $\Delta_k > 0$ và đã có x^k, C_k hiện hành. Mở đầu bước $k+1$, đánh dấu + và tách ra các cột j của C_k có độ lệch bằng 0 (đã nhận đủ lượng hàng):

$$\bar{\delta}_j^k = d_j - \sum_{i=1}^m x_{ij}^k = 0.$$

• **Đoạn 1.** Nếu tất cả 0 của C_k đã được tách, chuyển sang đoạn 3. Ngược lại thì chọn tùy ý một số 0, giả sử ở ô (i, j) . Khi đó tính độ lệch hàng i

$$\delta_i^k = s_i - \sum_{\mu=1}^n x_{i\mu}^k.$$

Có hai trường hợp.

(α) $\delta_i^k = 0$. đặt + bên phải dùng i và tách hàng i , đánh dấu $0'$ vào 0 ở ô (i, j) . Xét tất cả các cột đã tách của C_k nếu ở cột μ có $x_{i\mu}^k > 0$ (trên giao với hàng i đang xét) thì khoanh $\oplus$ xóa dấu tách cột μ và đánh dấu 0^* lên 0 ở ô (i, μ) này của C_k .

(β) $\delta_i^k > 0$. Xét các cột như trên, nếu có $x_{i\mu}^k > 0$ thì đánh dấu $0'$ lên 0 ở ô (i, μ) và chuyển sang đoạn 2.

Phải xét tất cả các số 0 chưa tách khi thấy ở đầu đoạn 1, cho đến khi sang được đoạn 2 thì kết thúc bước lặp, có thể phải làm đoạn 1 và đoạn 3 nhiều lần. Ta cũng gọi kết thúc đoạn 1 bởi sang được đoạn 2 là kết thúc kiểu 1A, còn sang đoạn 3 là kiểu 1B như ở bài toán bổ nhiệm.

• **Đoạn 2.** Lập dãy truyền các $0'$ và 0^* của C_k , bắt đầu từ một 0 chưa tách, ứng với trường hợp $\delta_i^k > 0$, giả sử ở ô (λ_1, μ_1) như sau. Đi theo cột μ_1 lấy $0_{\lambda_2 \mu_1}^*$. Tiếp theo, đi theo hàng để lấy $0_{\lambda_2 \mu_2}'$. Rồi lại theo cột $\mu_2, \dots$ cho đến khi không làm tiếp được nữa. (Thủ tục này giống trường hợp bài toán bổ nhiệm.) Sẽ chứng minh được thủ tục này là xác định đơn trị, dãy truyền không có chu trình (các phần tử đều khác nhau) và luôn kết thúc bằng $0'$:

$$0_{\lambda_1 \mu_1}', 0_{\lambda_2 \mu_1}^*, 0_{\lambda_2 \mu_2}', \dots, 0_{\lambda_s \mu_{s-1}}^*, 0_{\lambda_s \mu_s}'$$

Phương án x^{k+1} xác định bởi

$$x_{ij}^{k+1} = \begin{cases} x_{ij}^k & \text{nếu } (i, j) \text{ không ở dây truyền,} \\ x_{ij}^k + \theta_k & \text{nếu } (i, j) \text{ là } 0' \text{ ở dây truyền, tức là ở thứ tự lẻ,} \\ x_{ij}^k - \theta_k & \text{nếu } (i, j) \text{ là } 0^* \text{ ở dây truyền tức là ở thứ tự chẵn,} \end{cases}$$

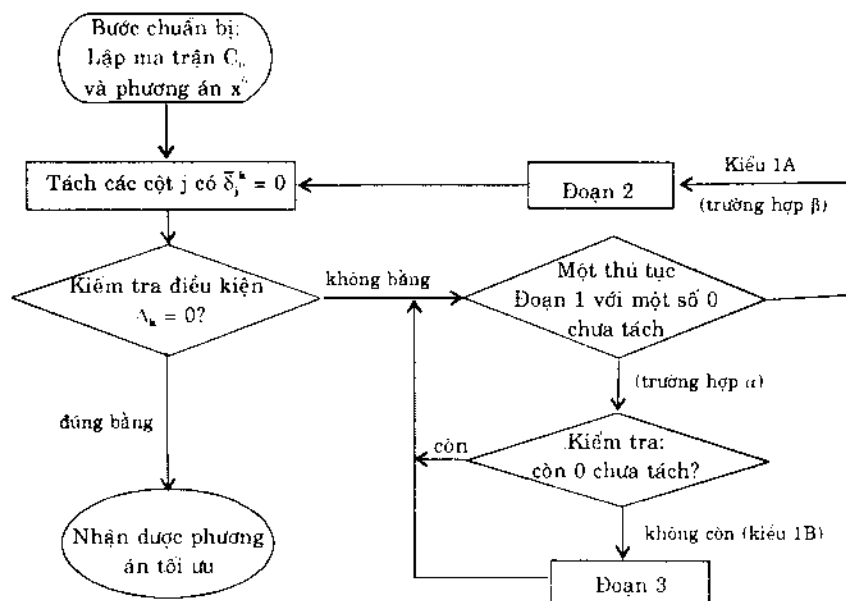
ở đây

$$\theta_k := \min_{2 \leq t \leq s} \left\{ x_{\lambda_t \mu_{t+1}}^k, \delta_{\lambda_1}^k, \bar{\delta}_{\mu_s}^k \right\}.$$

Dưới đây sẽ chứng minh là $x_{ij}^{k+1} \geq 0$, thỏa (5.41) và $\Delta_{k+1} < \Delta_k$.

• **Đoạn 3.** Nội dung là chuyển từ C_k sang C_k^1 , tương đương, với $c_{ij}^{k1} \geq 0$ và có phần tử 0 chưa tách. Gọi h là số nhỏ nhất trong các phần tử chưa tách. Lấy mỗi hàng chưa tách trừ h và lấy mỗi cột đã tách cộng thêm h ta được C_k^1 . Rồi chuyển sang đoạn 1. Dưới đây sẽ chứng minh rằng sau p hữu hạn lần áp dụng đoạn 1 và đoạn 3 ta phải sang được đoạn 2 để kết thúc bước lặp.

(c) Sơ đồ khối của thuật toán Hungary cho bài toán vận tải



(d) Chứng minh thuật toán Hungary cho bài toán vận tải

Bước chuẩn bị và đoạn 1 là xác định và rõ ràng.

• **Đoạn 2.** Chứng minh dãy truyền là xác định hữu hạn, kết thúc ở 0' đúng như cho bài toán bổ nhiệm.

Bây giờ giải thích tính chất của phương án x^{k+1} . Rõ ràng $x_{ij}^{k+1} \geq 0$. Ta kiểm tra điều kiện “chấp nhận được nói lỏng” (5.41). Trước hết ta chỉ ra rằng

$$\begin{aligned} \sum_{j=1}^n x_{ij}^k &= \sum_{j=1}^n x_{ij}^{k+1} \quad \text{nếu } i \neq \lambda_1, \\ \sum_{i=1}^m x_{ij}^k &= \sum_{i=1}^m x_{ij}^{k+1} \quad \text{nếu } j \neq \mu_s. \end{aligned} \quad (5.42)$$

Thật vậy, ở mỗi hàng $i \neq \lambda_1$ cũng như ở mỗi cột $j \neq \mu_s$, hoặc là không có phần tử của dãy truyền đang xét hoặc có hai phần tử liền nhau, khi đó chuyển từ x^k sang x^{k+1} , trừ hàng $i = \lambda_1$ và cột $j = \mu_s$, tất cả x_{ij}^k không đổi, hoặc vừa cộng θ_k vừa trừ θ_k nên vẫn là không đổi. Với $i = \lambda_1$ và $j = \mu_1$ thì

$$\begin{aligned} \sum_{j=1}^n x_{\lambda_1 j}^{k+1} &= \sum_{j=1}^n x_{\lambda_1 j}^k + \theta_k, \\ \sum_{i=1}^m x_{i \mu_1}^{k+1} &= \sum_{i=1}^m x_{i \mu_1}^k + \theta_k. \end{aligned} \quad (5.43)$$

Do định nghĩa θ_k , (5.42) và (5.43), x^k vẫn thỏa (5.41).

Bây giờ chứng minh $\Delta_{k+1} < \Delta_k$. Theo (5.42) và (5.43) thì

$$\begin{aligned}
 \Delta_{k+1} &= \sum_{i=1}^m s_i + \sum_{j=1}^n d_j - 2 \sum_{i=1}^m \sum_{j=1}^n x_{ij}^{k+1} \\
 &= \sum_{i=1}^m s_i + \sum_{j=1}^n d_j - 2 \sum_{i=1}^m \sum_{j=1}^n x_{ij}^k - 2\theta_k = \Delta_k - 2\theta_k.
 \end{aligned}
 \tag{5.44}$$

Nhưng $\theta_k > 0$ vì với $2 \leq t \leq s$ thì $x_{\lambda_t \mu_{t-1}} > 0$ (ở ô 0*), $\delta_{\lambda_1} > 0$ (do điều kiện chuyển sang đoạn 2), $\delta_{\mu_s} > 0$ (điều kiện dây truyền ngừng). Do đó $0 \leq \Delta_{k+1} < \Delta_k$.

• *Đoạn 3.* Sau khi thực hiện đoạn 3, chuyển từ C_k sang C_k^1 , ít nhất một số 0 chưa tách xuất hiện. Do đó ta phải đến Đoạn 1. Sau Đoạn 1 này hoặc là dẫn đến Đoạn 2 (kết thúc kiểu 1A, trường hợp β) hoặc là tăng số hàng tách của C_k^1 (trường hợp α). Chú ý rằng, theo (α), chỉ có hàng i với độ lệch $\delta_i^k = 0$ mới có thể được tách. Giả sử sau một số cặp Đoạn 3 và 1 tất cả các hàng có độ lệch 0 đã được tách. Số các hàng với độ lệch 0 nhỏ hơn m (vì $\Delta_k > 0$). Do đó vẫn còn phần tử chưa tách và chúng nằm ở các hàng có độ lệch dương. Giả sử vẫn chưa sang được Đoạn 2 mà là Đoạn 3, thì phải dẫn tới các phần tử 0 chưa được tách, nằm ở các hàng có độ lệch dương. Vậy thì Đoạn 1 kế tiếp chỉ xử lý một số 0 ở hàng có độ lệch dương (trường hợp β) là phải sang Đoạn 2. Như thế Đoạn 3 được tiến hành trong bước lặp $k+1$ vượt quá số hàng có độ lệch 0 của phương án x^k không quá 1. Vậy sau hữu hạn thủ tục phải kết thúc bước lặp.

Nói riêng, giả sử tất cả s_i và d_j là các số nguyên. Khi đó θ_k nguyên. Theo (5.44), mỗi bước lặp làm giảm độ lệch tổng cộng Δ_k đi ít nhất là 2. Vì vậy qua hữu hạn bước lặp phải dẫn đến $\Delta_d = 0$.

- Chứng minh x^d là nghiệm tối ưu bài toán vận tải.

Trước hết chứng tỏ x^d chấp nhận được. Do (4.41)

$$\delta_i^k \geq 0, \bar{\delta}_j^k \geq 0, \forall i, \forall j, \forall k.$$

Từ các đẳng thức

$$0 = \Delta_d = 2 \sum_{i=1}^m \delta_i^d = 2 \sum_{j=1}^n \bar{\delta}_j^d$$

ta có $\delta_i^d = 0, \bar{\delta}_j^d = 0 \forall i, \forall j$, tức là x^d chấp nhận được.

Chứng minh x^d tối ưu dựa vào ba nhận xét sau.

NHẬN XÉT 5.6.3. Nếu $x_{ij}^r > 0$ thì $c_{ij}^r = 0$.

CHỨNG MINH. Với $r = 0$ thì $x_{ij}^0 > 0$ phải suy ra $c_{ij}^0 = 0$ theo cách xây dựng phương án x^0 ở bước chuẩn bị. Giả sử Nhận xét là đúng với k và ở bước lập $k+1$ sử dụng đoạn 3 p lần. Nếu $x_{ij}^k > 0$ thì theo giả thiết quy nạp, $c_{ij}^k = 0$ và phải nằm hoặc ở hàng tách, hoặc ở cột tách (theo đoạn 1). Do đó theo đoạn 3, $c_{ij}^{k+1} = c_{ij}^k = 0$. Áp dụng đoạn 3 p lần thì $c_{ij}^k = c_{ij}^{k+1} = \dots = c_{ij}^{kp} = c_{ij}^{k+1} = 0$. Bây giờ kiểm tra kết luận với $k+1$. Nếu $x_{ij}^{k+1} > 0$ thì có hai khả năng. Nếu $x_{ij}^k > 0$ thì, theo lý luận vừa rồi $c_{ij}^{k+1} = 0$. Khả năng thứ hai là $x_{ij}^k = 0$, thì ở lập $k+1$ biến này được tăng lên dương nên phải ứng ô (i, j) lẻ trong dây chuyền để xây dựng x^{k+1} . Mà mọi ô trong dây chuyền đều là ô số 0 của $c_k^p = c_{k+1}$. Vậy $c_{ij}^{k+1} = 0$. □

NHẬN XÉT 5.6.4. Với k bất kỳ, ma trận C_k có mọi phần tử không âm.

CHỨNG MINH. Chỉ ở đoạn 3 (và bước chuẩn bị) là biến đổi các c_{ij}^k , nhưng là phép trừ đi các số nhỏ nhất, nên vẫn luôn đảm bảo $c_{ij}^k \geq 0$. □

NHẬN XÉT 5.6.5. Nếu hai ma trận cước phí C và $\bar{C}$ tương đương thì tập nghiệm tối ưu của hai bài toán vận tải tương ứng, với cùng ràng buộc (tức là cùng các s_i, d_j), trùng nhau.

CHỨNG MINH. Do $C \sim \bar{C}$, sẽ có α_i và β_j để

$$c_{ij} = \bar{c}_{ij} + \alpha_i + \beta_j, \forall i, \forall j.$$

Giả sử $x^* = (x^*_{ij})$ là phương án tối ưu của bài toán với ma trận cước phí C , ký hiệu là $T(C)$ và $x = (x_{ij})$ là phương án chấp nhận được của $T(C)$ và $T(\bar{C})$. Ta có

$$\begin{aligned} \sum_i \sum_j c_{ij} x_{ij} &= \sum_i \sum_j (\bar{c}_{ij} + \alpha_i + \beta_j) x_{ij} \\ &= \sum_i \sum_j \bar{c}_{ij} x_{ij} + \sum_i \alpha_i \sum_j x_{ij} + \sum_j \beta_j \sum_i x_{ij} \\ &:= \sum_i \sum_j \bar{c}_{ij} x_{ij} + \gamma, \end{aligned}$$

với γ không phụ thuộc vào phương án x . Vì x^* tối ưu của $T(C)$,

$$\sum_i \sum_j c_{ij} x_{ij} \geq \sum_i \sum_j c_{ij} x^*_{ij},$$

tức là

$$\sum_i \sum_j \bar{c}_{ij} x_{ij} = \sum_i \sum_j c_{ij} x_{ij} - \gamma \geq \sum_i \sum_j c_{ij} x^*_{ij} - \gamma = \sum_i \sum_j \bar{c}_{ij} x^*_{ij}.$$

Vậy x^* là tối ưu của $T(\bar{C})$. $\square$

Chứng minh x^d tối ưu. Xét bài toán vận tải $T(C_d)$ tương đương với $T(C)$ (vì theo đoạn 3, $C_d \sim C$). Theo Nhận xét 4.6.4, $c_{ij}^d \geq 0$ nên với phương án chấp nhận được bất kỳ ta có, theo Nhận xét 4.6.3,

$$\sum_i \sum_j c_{ij}^d x_{ij} \geq 0 = \sum_j \sum_i c_{ij}^d x_{ij}^d.$$

Vậy x^d là tối ưu của $T(C_d)$ và do đó cũng tối ưu của $T(C)$. $\square$

Cần nhấn mạnh lại là nếu các s_i và d_j nguyên thì sau hữu hạn bước phương pháp Hungary dẫn đến nghiệm tối ưu.

THÍ DỤ 5.6.1. Giải bài toán vận tải với ma trận cước phí và lượng hàng phát và thu là

$$C = \begin{bmatrix} 2 & 5 & 9 & 5 \\ 8 & 3 & 5 & 8 \\ 7 & 3 & 1 & 4 \\ 5 & 9 & 7 & 2 \end{bmatrix} \quad (s_1, s_2, s_3, s_4; d_1, d_2, d_3, d_4) = (4, 2, 3, 3; 3, 6, 2, 1).$$

Bước chuẩn bị. Trừ các cột 1, 2, 3, 4 cho 2, 3, 1, 2 tương ứng ta được

$$C_0 = \begin{bmatrix} 0 & 2 & 8 & 3 \\ 6 & 0 & 4 & 6 \\ 5 & 0 & 0 & 2 \\ 3 & 6 & 6 & 0 \end{bmatrix}.$$

Vì cực tiểu mỗi hàng là 0 nên không trừ từ các hàng. Phương án xuất phát là

$$x^0 = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} 4 \\ 1 \\ 3 \\ 2 \end{matrix}$$

$$\begin{matrix} 3 & 6 & 2 & 1 \\ 0 & 1 & 2 & 0 \end{matrix}$$

(hai cột ngoài ghi lượng phát và độ lệch hàng, hai hàng dưới ma trận ghi lượng thu và độ lệch cột). Vì $\frac{1}{2}\Delta_0 = 3 > 0$ nên vào bước lặp 1. (Chú ý 3 là tổng độ lệch hàng, cũng bằng tổng độ lệch cột.)

Bước lặp 1. Tách cột 1 và 4 ở C_0 bằng dấu + vì có độ lệch $\bar{\delta}_1^0, \bar{\delta}_4^0$ (ghi ở dưới ma trận x^0) bằng 0. Các số 0 ở cột tách đánh dấu 0 nếu $x_{ij}^0 > 0$. Còn số 0 chưa tách nên vào đoạn 1.

Đoạn 1. Đánh dấu 0' vào hai chỗ hàng 2, 3 và tách hai hàng này bằng + bên phải. Đánh thêm × vào bên trái hàng không tách. Đã hết 0 chưa tách nên sang Đoạn 3.

Đoạn 3. $h = 2$ nên ta được C_0^1 . Có 0 chưa tách ở ô (1,2) nên quay lại Đoạn 1.

Đoạn 1. Đánh dấu 0' ở ô (1,2). Độ lệch hàng 1 là $\delta_1^0 = 1 > 0$ nên sang Đoạn 2.

Đoạn 2. Dây truyền chỉ có một phần tử 0' ở ô (1,2).

$$C_0 = \begin{array}{cc} & + & + \\ \times \begin{bmatrix} \bar{0} & 2 & 8 & 3 \\ 6 & 0' & 4 & 6 \\ 5 & 0' & 0 & 2 \\ \times & 3 & 6 & 6 & \bar{0} \end{bmatrix} & \xrightarrow{h=2} & C_0^1 = \begin{array}{cc} & + & + \\ \times \begin{bmatrix} \bar{0} & 0 & 6 & 3 \\ 8 & 0' & 4 & 8 \\ 7 & 0' & 0 & 4 \\ \times & 3 & 4 & 4 & \bar{0} \end{bmatrix} & \xrightarrow{\quad} & C_1 = \begin{array}{cc} & + \uparrow & + \\ \times \begin{bmatrix} \bar{0} & 0' & 6 & 3 \\ 8 & 0' & 4 & 8 \\ 7 & 0' & 0 & 4 \\ \times & 3 & 4 & 4 & \bar{0} \end{bmatrix} \end{array} \end{array}$$

Bước lặp 1

$$\theta_0 = \min_{i=1} \left\{ \delta_i^0, \bar{\delta}_i^0 \right\} = \min \{1, 1\} = 1. \text{ Chuyển từ } x^0 \text{ sang } x^1.$$

$$x^0 = \begin{array}{cc} & \uparrow \\ \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 2 & 0 \end{bmatrix} & \begin{matrix} 1 \\ 0 \\ 0 \\ 2 \end{matrix} \end{array} \xrightarrow{\theta_0=1} x^1 = \begin{array}{cc} \begin{bmatrix} 3 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 \end{bmatrix} & \begin{matrix} 0 \\ 0 \\ 0 \\ 2 \end{matrix} \end{array}$$

Bước lặp 1 (Đoạn 2)

(Bên phải và dưới ma trận ghi độ lệch hàng, cột.) Vì $\frac{1}{2}\Delta_1 = 2 > 0$, phải sang Bước lặp 2.

Bước lặp 2. Mở đầu đánh dấu + các cột j có độ lệch $\bar{\delta}_j^1 = 0$ của C_1 tức là cột 1, 2 và 4. (Khác với bài toán bổ nhiệm, mọi dấu tách, 0*, 0' và $\bar{0}$ ở đây được làm lại ở đầu mỗi bước lặp.) Còn 0 ở ô (3,3) chưa tách nên vào đoạn 1.

Đoạn 1. Đánh dấu 0' ở (3,3) và + bên phải hàng. $\delta_3^1 = 0$ nên ta có trường hợp α . Ta có $x_{32}^1 > 0$ nên đánh dấu 0* vào ô (3,2) khoanh $\oplus$ xóa dấu tách cột 2. Lại xuất hiện 0 chưa tách ở (1,2) với $\delta_1^1 = 0$, trường hợp α đánh dấu + hàng 1, 0' ở ô (1,2). Vì $x_{11}^1 > 0$ nên ghi 0* vào ô (1,1) khoanh $\oplus$ vào cột 1. Còn 0 ở (2,2) ứng $\delta_2^1 = 0$, đánh + hàng 2, 0' vào ô (2,2). Bây giờ đã tách hết số 0 của C_1 nên phải vào đoạn 3. (ghi 0 vào các ô có $x_{ij}^1 > 0$.)

Đoạn 3. $h = 3$ nên ta được C_1^1 như dưới đây.

$$C_1 = \begin{bmatrix} \oplus & \oplus & & + \\ \bar{0}^* & \bar{0}' & 6 & 3 \\ 8 & \bar{0}' & 4 & 8 \\ 7 & \bar{0}^* & 0' & 4 \\ 3 & 4 & 4 & \bar{0} \end{bmatrix} \xrightarrow{h=3} C_1^1 = \begin{bmatrix} & & & + \\ 0^* & 0' & 6 & 6 \\ 8 & 0' & 4 & 11 \\ 7 & 0^* & 0' & 7 \\ 0 & 1 & 1 & \bar{0} \end{bmatrix}$$

Lại xuất hiện 0 chưa tách ở (4,1) với độ lệch hàng $\delta_4^1 = 2 > 0$, là trường hợp β nên đánh dấu 0' ở ô (4,1) và phải sang đoạn 2. Vì thế cũng lấy $C_2 = C_1^1$.

Đoạn 2. Lập dây truyền ở C_2 bắt đầu từ 0'

ở (4,1) chưa tách và có $\delta_4^1 = 2 > 0$, đi theo cột lên 0* ở (1,1), lại theo hàng sang 0' ở (1,2), ...

$$C_2 = \begin{bmatrix} \bar{0}^* & 0' & 6 & 6 \\ 8 & 0' & 4 & 11 \\ 7 & 0^* & 0' & 7 \\ 0' & 1 & 1 & \bar{0} \end{bmatrix}$$

Phương án x^2 xác định như sau, ở đây

$$\theta_1 = \min \{3, 3, 2, 2\} = 2, \\ x^1 = \begin{bmatrix} 3 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 \end{bmatrix} \xrightarrow{\theta_1=2} x^2 = \begin{bmatrix} 1 & 3 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Ta thấy $\Delta_2 = 0$ (các độ lệch cột là 0 tất cả !) nên thuật toán kết thúc; x^2 là phương án tối ưu. (Chú ý rằng x^2 không phải phương án cực biên.)

CHÚ Ý CUỐI CÙNG. Phương pháp Hungary thường áp dụng tốt cho bài toán vận tải với các s_j, d_j nguyên. Khi đó số bước lặp là hữu hạn và đánh giá trước được là không quá $\frac{1}{2}\Delta_0$, vì mỗi bước độ lệch giảm được ít nhất là 2 (thông thường là nhỏ hơn $\frac{1}{2}\Delta_0$). Ở Ví dụ 5.6.1 thì $\frac{1}{2}\Delta_0 = 3$ và ta cần 2 bước lặp. Người ta cũng đánh giá được độ phức tạp tính toán mỗi bước lặp qua m và n . Hơn nữa còn đánh giá được mức độ gần tối ưu của kết quả ở mỗi bước lặp để có thể dừng thuật toán nếu ở mức sai số cho phép (xem [Golstein - Yudin 1969], Mục 5.8.2).

Ngoài ra phương pháp này ít nhạy cảm với tính suy biến. Việc lập các dây chuyền cũng khá đơn giản.

Còn có nhiều phương pháp khác để giải bài toán vận tải, chẳng hạn phương pháp Gleyzal [Gleyzal 1958] phương pháp của Williams, [Williams 1962], phương pháp phương án tối ưu có điều kiện của [Lurie 1961] và [Oleynik 1960], phương pháp của [Brudno 1962], Có thể chia các phương pháp ra làm hai loại: Loại 1, điển hình là phương pháp thế vị, là các phương pháp cải thiện dần phương án. Loại 2, điển hình là phương pháp Hungary, là phương pháp giảm dần độ lệch khỏi ràng buộc.

BÀI TẬP CHƯƠNG 5

5.1 Xét bài toán vận tải. Chứng minh rằng

(a) mọi cơ sở B (của bài toán vận tải) có $|\det B| = 1$;

(b) mọi cơ sở là ma trận nửa đường chéo dưới.

(c) nếu các s_j và d_j đều nguyên thì mọi phương án cơ sở chấp nhận được đều có mọi thành phần nguyên.

5.2 Nhu cầu về một loại hàng để hồng trong 4 tháng là 500, 630, 200 và 230 tấn. Khả năng cung cấp trong các tháng liên tiếp là 400, 300, 420 và 380 tấn và giá một tấn tương ứng là 100, 140, 120 và 150 ngàn đồng. Hàng chỉ để được 3 tháng (kể cả tháng nhận hàng). Cước lưu kho một tấn một tháng là 3 ngàn đồng. Viết bài toán ở dạng bài toán vận tải.

5.3 Dùng quy tắc góc tây bắc tìm nghiệm xuất phát. Nghiệm này có suy biến không. (Đưa về bài toán vận tải cân bằng nếu cần.)

$$(a) (s_1, s_2, s_3, s_4) = (10, 5, 4, 6), (d_1, d_2, d_3, d_4) = (10, 5, 7, 3).$$

$$(b) (s_1, s_2, s_3, s_4) = (1, 16, 7, 8), (d_1, d_2, d_3, d_4, d_5) = (3, 4, 5, 2, 8).$$

$$(c) (s_1, s_2, s_3) = (30, 35, 45), (d_1, d_2, d_3, d_4) = (15, 17, 40, 15).$$

5.4 Xét phương án cơ sở ở H.5.5.

(a) Xác định chu trình tương ứng với mỗi biến không cơ sở.

(b) Giả sử mỗi biến không cơ sở đều dùng làm biến vào. Hãy xác định biến ra tương ứng. Mỗi biến không cơ sở vào ở mức nào?

(c) Xác định độ tăng và giảm của giá trị mục tiêu ứng với mỗi trường hợp ở (b).

5.5 Tìm phương án cơ sở xuất phát bằng cả ba phương pháp: góc tây bắc, cực tiểu cước phí và xấp xỉ Vogel cho mô hình vận tải bên phải. Phương án nào cho mục tiêu tốt nhất.

d_j	10	10	10
s_i			
7	1	2	6
12	0	4	2
11	3	1	5

5.6 Làm như Bài 5.5 cho mô hình vận tải bên dưới. Sau đó tìm phương án tối ưu dùng phương án xuất phát tốt nhất trong các phương án vừa có.

d_j	60	60	20	10
s_i				
10	10	20	5	7
20	13	9	12	8
30	4	15	7	9
40	14	7	1	0
50	3	12	5	19

5.7 Giải bài toán không cân bằng bên phải bằng cách dùng phương pháp xấp xỉ Vogel tìm nghiệm xuất phát. Bài toán có thêm một điều kiện là bắt buộc nút cung 4 tải đủ hàng đến nút cầu 1.

	d_j	5	10	15
s_i				
20		5	1	0
10		3	2	4
15		7	5	2
15		9	6	0

5.8 Giải các bài toán bổ nhiệm sau đây bằng phương pháp thế vị và bằng phương pháp Hungary (cho ma trận cước phí)

$$(a) \begin{bmatrix} 3 & 8 & 2 & 10 & 3 \\ 8 & 7 & 2 & 9 & 7 \\ 6 & 4 & 2 & 7 & 5 \\ 8 & 4 & 2 & 3 & 5 \\ 9 & 10 & 6 & 9 & 10 \end{bmatrix}$$

$$(b) \begin{bmatrix} 3 & 9 & 2 & 3 & 7 \\ 6 & 1 & 5 & 6 & 6 \\ 9 & 4 & 7 & 10 & 3 \\ 2 & 5 & 4 & 2 & 1 \\ 9 & 6 & 2 & 4 & 6 \end{bmatrix}$$

$$(c) \begin{bmatrix} 5 & 5 & -2 \\ 7 & 4 & 2 & 3 \\ 9 & 3 & 5 & - \\ 4 & 2 & 6 & 7 \end{bmatrix}$$

dấu - ở c13 chẳng hạn, có nghĩa là above người 1 không thể được phân việc 3).

5.9 Giải bài toán vận tải bên dưới bằng phương pháp thế vị với hai cách tìm nghiệm xuất phát: quy tắc góc tây bắc và phương pháp xấp xỉ Vogel

	d_j	3	3	2	2
s_i					
5		3	7	6	4
2		2	4	3	2
3		4	3	8	5

5.10 Dựa vào thành tích kiểm tra bơi bốn kiểu dưới đây của năm vận động viên thiếu niên dự tuyển, huấn luyện viên cần xếp ai bơi kiểu gì trong đội bơi 200 met tiếp sức hỗn hợp

kiểu	An	Chiến	Hùng	Quang	Tuấn
ngửa	37,7	32,9	33,8	37,0	35,4
ếch	43,4	33,1	42,2	34,7	41,8
bướm	33,3	28,5	38,9	30,4	33,6
sải	29,2	26,4	29,6	28,5	31,1

(a) Phát biểu thành bài toán bổ nhiệm.

(b) Giải bằng phần mềm nào mà bạn biết hoặc tính tay.

5.11 Trong mô hình vận tải bên phải cầu nhiều hơn cung. Biết tiền phạt mỗi đơn vị hàng thiếu là 5, 3 và 2 tương ứng cho nút 1, 2, 3. Tìm phương án tối ưu.

	75	20	50
10	5	1	7
80	6	4	6
15	3	2	5

5.12 Giả sử trong Bài 5.8 (c) có thêm máy (việc) thứ năm với cước tương ứng với bốn người là 2, 1, 2 và 8. Máy mới có thể thay một máy cũ nếu có lợi về kinh tế. Hãy phát biểu bài toán mới như một bài toán bổ nhiệm và giải. Nói riêng, nên thay máy nào không?

5.13 Giải bằng phương pháp Hungary

(a) Bài 5.7;

(b) Bài 5.9;

(c) Bài 5.12.

5.14 Giải bài toán vận tải hàng từ bốn điểm phát đến bốn điểm thu có các lượng cung, cầu (theo tấn) và khoảng cách các điểm (km) cho ở bảng dưới đây. Tìm phương án vận tải để lượng tấn hàng x km là nhỏ nhất.

	120	180	100	150
140	2	4	6	7
550	3	8	9	8
80	6	5	6	9
550	5	7	8	7

CHƯƠNG 6

QUY HOẠCH NGUYÊN

6.1 PHÁC HỌA QUY HOẠCH NGUYÊN

Quy hoạch nguyên (Integer programming, viết tắt IP) là bài toán quy hoạch trong đó tất cả hoặc một phần các biến bị ràng buộc chỉ lấy giá trị nguyên. Tương ứng tên cụ thể cho từng trường hợp là *quy hoạch nguyên hoàn toàn* (pure integer programming) và *quy hoạch nguyên bộ phận* (mixed integer programming, viết tắt là MIP). Tuy vậy, thuật ngữ "quy hoạch nguyên" thực tế chỉ dành cho lĩnh vực toán nghiên cứu lớp bài toán quy hoạch tuyến tính nguyên. Vì vậy hai thuật ngữ được hiểu như nhau. Lớp bài toán này rất phổ biến trong thực tế. Mảng bài toán có vẻ đơn giản nhất mà cũng quan trọng nhất trong lớp này là các bài toán chọn các quyết định (có liên quan nhau nhưng đều ở dạng "có" hoặc "không" (yes - or - no decision). Chẳng hạn bài toán bổ nhiệm ở Thí dụ 5.1.1, biến quyết định sẽ lấy giá trị như sau

$$x_{ij} = \begin{cases} 1 & \text{nếu người } i \text{ được nhận việc } j, \\ 0 & \text{nếu người } i \text{ không nhận việc } j. \end{cases}$$

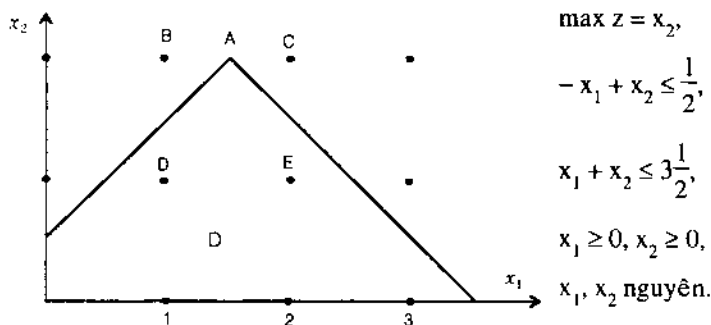
Nhiều bài toán dạng quyết định có - không như vậy, chẳng hạn : có đầu tư không? có mở đại lý ở thành phố A không? có nhận thực hiện dự án không?.... Vì các biến (quyết định) chỉ nhận hai giá trị (biến 0-1) nên bài toán có tên là *quy hoạch nguyên nhị phân* (binary integer

programming, tất là BIP).

Nếu bài toán có n biến và các ràng buộc cho thấy miền chấp nhận được là giới nội trong R^n , thì bài toán "chỉ có" hữu hạn nghiệm chấp nhận được, không như quy hoạch tuyến tính (tổng quát) có miền này là cả một tập lồi đa diện. Tuy vậy, sẽ rất sai lầm nếu nghĩ rằng quy hoạch nguyên dễ giải hơn, hoặc hơn nữa nghĩ rằng có thể *điểm diện* (enumeration) lần lượt các phương án (tức là nghiệm) chấp nhận được để chọn phương án tối ưu. Chẳng hạn, BIP với n biến có 2^n phương án chấp nhận được. Nhưng đây là số khổng lồ nếu n lớn. Trong lịch sử Trung hoa cổ đại đã có minh họa rất hay là nhà vua không đủ thóc, dù vét hết kho trong nước, để thưởng 2^{64} hạt thóc theo ước muốn (được vua ban) của người nghĩ ra cờ tướng làm vua thích thú (bàn cờ có 64 ô!). Nguyên nhân vì miền chấp nhận được, tuy hữu hạn, của quy hoạch nguyên không có các đỉnh để thực hiện phương pháp đơn hình, đã biết là rất hiệu quả. Chính vì vậy quy hoạch tuyến tính dễ giải hơn rất nhiều so với quy hoạch tuyến tính nguyên. Người ta có thể giải mọi quy hoạch tuyến tính với nhiều ngàn biến và ràng buộc bằng máy tính trong vài giờ, nhưng đến nay vẫn không đảm bảo giải được với thời gian tương tự bất kỳ dù chỉ là BIP với 100 biến. (Người ta thấy rằng ở quy hoạch tuyến tính thì độ phức tạp tính toán tăng rõ theo số ràng buộc nhưng với quy hoạch nguyên thì số biến ảnh hưởng quyết định.)

Một ý tưởng tự nhiên để giải IP (quy hoạch nguyên) là giải quy hoạch tuyến tính tương ứng, tức là IP nhưng bỏ đi ràng buộc giá trị biến phải nguyên (quy hoạch tuyến tính này gọi là LP – nói lỏng (LP relaxation) của IP) rồi làm tròn nghiệm tối ưu thành nghiệm nguyên. Nó sẽ xấp xỉ nghiệm tối ưu của IP. Cách này có thể áp dụng trong thực tế, nhưng phải cẩn thận hai nguy cơ sau đây. Một là nghiệm làm tròn có thể thậm chí không chấp nhận được đối với IP như thí dụ sau.

THÍ DỤ 6.1.1. Xét quy hoạch nguyên

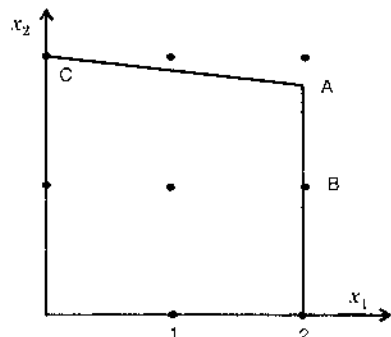


H. 6.1 cho thấy điểm A là nghiệm tối ưu của LP – nơi lỏng, B và C là hai nghiệm làm tròn từ A nhưng không chấp nhận được. Nghiệm tối ưu của IP là D và E.

Nguy cơ thứ hai là nghiệm làm tròn có thể chấp nhận được nhưng có thể có giá trị mục tiêu rất xa mục tiêu tối ưu của IP như thí dụ sau đây cho thấy.

THÍ DỤ 6.2.2. Xét quy hoạch nguyên

$$\begin{aligned}
 \max z &= x_1 + 5x_2, \\
 x_1 + 10x_2 &\leq 20, \\
 x_1 &\leq 2, \\
 x_1 \geq 0, x_2 &\geq 0, \\
 x_1, x_2 &\text{ nguyên.}
 \end{aligned}$$



H.6.2 cho thấy điểm A là nghiệm tối ưu của LP - nơi lỏng (với mục tiêu $z = 11$). B là nghiệm làm tròn cho IP, chấp nhận được và ứng mục tiêu $z = 7$, xa với mục tiêu tối ưu của IP là $z = 10$ đạt tại điểm C.

H. 6.2 Nghiệm tối ưu làm tròn xa nghiệm tối ưu của IP

Bài tập 6.15 cho thấy trường hợp nghiệm làm tròn không chấp nhận được là rất phổ biến. Có thể nghĩ rằng dữ liệu của bài toán đã là xấp xỉ nên tính không chấp nhận được này có thể "tha thứ". Nhưng nhiều bài toán IP thực tế có dữ liệu chính xác. Chẳng hạn ràng buộc kiểu lựa chọn đúng một trong n quyết định "có" còn lại là "không" là ràng buộc với hệ số chính xác

$$x_1 + x_2 + \dots + x_n = 1, x_j = 0 \text{ hoặc } 1.$$

Hơn nữa ở bài toán nhị phân với biến có ý nghĩa là "có" hay "không" thì giá trị không nguyên là vô nghĩa rồi, nên không thể làm tròn đi nữa.

Vì quy hoạch nguyên hay gặp như thế, lại không thể nhận được từ làm tròn nghiệm nhận được từ LP – nói lỏng của nó, nhu cầu tìm các thuật toán riêng là rất lớn, vẫn đang rất được quan tâm nghiên cứu hiện nay. Do giải quy hoạch nguyên khó và tính toán phức tạp nên với các bài toán cỡ lớn cách tiếp cận tốt nhất là dùng các *thuật toán phán đoán* (heuristic algorithm), tuy không đảm bảo đi đến nghiệm tối ưu trong mọi trường hợp, nhưng thường rất hiệu quả. Với các bài toán không lớn quá, đã có nhiều thuật toán giải chính xác. Cách tiếp cận chung là kết hợp việc giải các LP - nói lỏng với *điểm diện ẩn* (implicit enumeration), tức là điểm diện các nghiệm chấp nhận được nhưng ở mỗi bước tìm cách loại bỏ số đáng kể các nghiệm không thể là tối ưu không cần điểm diện. Trước khi trình bày hai phương pháp chính là phương pháp nhánh và cận (branch - and - bound method) và phương pháp mặt phẳng cắt (cutting - plane method) ở các mục sau, Mục 6.2 giới thiệu một số bài toán điển hình dẫn đến quy hoạch nguyên.

Để kết thúc mục này, hãy nhắc lại rằng ở Chương 3 và 5 ta đã thấy nhiều trường hợp may mắn là giải quy hoạch tuyến tính (không nguyên) bằng thuật toán đơn hình ta nhận được nghiệm tối ưu là

nguyên. Đó là bài toán dòng trên mạng với các dữ liệu nguyên, mà trường hợp riêng là các bài toán sau với dữ liệu nguyên: bài toán dòng cực đại, bài toán đường ngắn nhất, bài toán vận tải. Nếu khi giải LP - nơi lỏng của quy hoạch nguyên mà gặp may mắn như vậy thì ta được luôn nghiệm tối ưu của quy hoạch nguyên.

6.2 QUY HOẠCH NGUYÊN ĐIỂN HÌNH TRONG THỰC TẾ

THÍ DỤ 6.2.1 (Bài toán ba lô - knapsack problem). Một nhà thám hiểm chỉ được mang theo một ba lô trọng lượng không quá b . Có n loại vật dụng nên mang theo. Loại vật j có trọng lượng mỗi vật là a_j . Và giá trị sử dụng mỗi vật là c_j . Hỏi ông phải mang thế nào để có giá trị sử dụng lớn nhất?

Gọi x_j là số lượng vật j mang theo thì mô hình toán học của bài toán ba lô này là quy hoạch nguyên sau

$$\max z = \sum_{j=1}^n c_j x_j,$$

$$\sum_{j=1}^n a_j x_j \leq b,$$

$$x_j \geq 0 \text{ và nguyên, } j = 1, \dots, n.$$

Sau đây sẽ gọi $\mathbf{Z}_+$ là tập số nguyên không âm. Về toán học, nếu mục tiêu là min z hoặc ràng buộc là đẳng thức, thì bài toán cũng gọi là bài toán ba lô.

Bài toán ba lô có dạng đặc biệt và đơn giản là chỉ có một ràng buộc hàm, ngoài ràng buộc dấu và tính nguyên. Do đó người ta nghiên cứu được nhiều cách giải riêng, rồi tìm cách đưa các quy hoạch nguyên khác về bài toán ba lô để giải. Chẳng hạn người ta chứng minh được

định lý sau để ghép các ràng buộc thành một ràng buộc.

ĐỊNH LÝ 6.2.1. Nếu $a_{ij} \geq 0$, $b_i > 0$, $i = 1, \dots, m$, $j = 1, \dots, n$ và tất cả nguyên thì luôn có các số nguyên $t_1, \dots, t_m$ để tập nghiệm nguyên không âm của m ràng buộc

$$\sum_{j=1}^n a_{ij}x_j = b_i, \quad i = 1, \dots, m,$$

trùng với tập nghiệm nguyên không âm của một ràng buộc

$$\sum_{i=1}^m t_i \sum_{j=1}^n a_{ij}x_j = \sum_{i=1}^m t_i b_i.$$

(Xem chứng minh, chẳng hạn, trong [Nghĩa 1997], Mục 4.2.)

Bây giờ ta giới thiệu một cách giải của bài toán ba lô

$$\begin{aligned} \min z &= \sum_{j=1}^n c_j x_j, \\ \sum_{j=1}^n a_j x_j &= b, \\ x_j &\in \mathbf{Z}_+, j = 1, \dots, n, \end{aligned} \tag{6.1}$$

bằng cách đưa nó về bài toán đường ngắn nhất (đã biết ba cách giải ở Mục 3.7 là thuật toán đơn hình mạng, thuật toán sửa nhân và thuật toán đặt nhân), với giả thiết $a_j \in \mathbf{Z}_+$, $b \in \mathbf{Z}_+$. Xây dựng đồ thị có hướng $\mathcal{G} = (\mathcal{V}, \mathcal{A})$ với tập nút $\mathcal{V} = \{0, 1, \dots, b\}$ và tập cung

$$\mathcal{A} = \{(i, k): \exists j \in \{1, \dots, n\} \text{ sao cho } k - i = a_j\}.$$

Gán cho cung (i, k) với $k - i = a_j$ độ dài là $d(i, k) = c_j$. Khi đó rõ ràng mỗi đường đi S của đồ thị có độ dài là

$$L(S) = \sum_{(i,k) \in S} d(i, k).$$

ĐỊNH LÝ 6.2.2. Mỗi đường đi S từ 0 đến b trong đồ thị $\mathcal{G}$ đều ứng với một phương án chấp nhận được $(x_1, \dots, x_n)$ nào đó của bài toán ba lô (6.1) và ngược lại. Khi đó độ dài đường S là

$$L(S) = \sum_{j=1}^n c_j x_j.$$

CHỨNG MINH. Giả sử $(x_1, \dots, x_n)$ là một phương án chấp nhận được của (6.1). Khi đó rõ ràng trong $\mathcal{G}$ có đường đi qua các đỉnh

$$0, a_1, 2a_1, \dots, x_1 a_1, x_1 a_1 + a_2, \dots, x_1 a_1 + x_2 a_2, \dots, \sum_{j=1}^n x_j a_j = b,$$

với độ dài $L(S) = \sum_{j=1}^n c_j x_j.$

Ngược lại, giả sử $S = (0, i_1, \dots, i_m = b)$ là một đường đi từ 0 đến b trong $\mathcal{G}$. Ta chứng tỏ mỗi nút i_k trong S đều có dạng

$$i_k = \sum_{j=1}^n a_j p_{kj}, \quad p_{kj} \in \mathbf{Z}_+ \quad (6.2)$$

bằng quy nạp. Theo định nghĩa $\mathcal{G}$, sẽ có chỉ số r để $i_1 - 0 = r$. Khi đó

$$i_1 = \sum_{j=1}^n a_j p_{1j},$$

trong đó $p_{1r} = 1, p_{1j} = 0$ với $j \neq r$.

Giả sử (6.2) đã đúng đến cùng i_{k-1} . Ta chứng minh cho i_k . Theo định nghĩa

$\mathcal{G}$, sẽ có chỉ số t sao cho $d_k - i_{k-1} = a_t$. Do đó

$$i_k = a_t + i_{k-1} = a_t + \sum_{j=1}^n a_j p_{k-1,j} = \sum_{j=1}^n a_j p_{kj}.$$

Với $p_{kj} = p_{k-1,j}$ nếu $j \neq t$ và $p_{kt} = p_{k-1,t} + 1$. Vậy (6.2) đúng với mọi nút i_k trong S .

Nói riêng với nút cuối $i_m = b$, đặt $p_{mj} = x_j$, $j = 1, \dots, n$, ta được

$$b = \sum_{j=1}^n a_j x_j,$$

tức là phương án $(x_1, \dots, x_n)$ chấp nhận được cho (6.1) và có mục tiêu z bằng độ dài đường S tương ứng:

$$z = \sum_{j=1}^n c_j x_j = L(S). \square$$

Một điểm hạn chế của phương pháp đưa về bài toán đường ngắn nhất là đồ thị $\mathcal{G}$ có số nút quá lớn khi b lớn. Người ta còn nhiều phương pháp khác để giải bài toán ba lô.

THÍ DỤ 6.2.2 (Sản xuất có lệ phí cố định: fixed-charge production). Giả sử nhà máy sản xuất n sản phẩm. Chi phí sản xuất sản phẩm j gồm lệ phí cố định K_j không phụ thuộc lượng sản phẩm j và cước phí đơn vị sản phẩm là c_j . Gọi x_j là lượng sản phẩm j sẽ sản xuất thì cước sản xuất sản phẩm j là

$$c_j(x_j) = \begin{cases} K_j + c_j x_j & \text{nếu } x_j > 0, \\ 0 & \text{nếu } x_j = 0. \end{cases}$$

Mục tiêu sản xuất với cước phí cực tiểu sẽ là

$$\min z = \sum_{j=1}^n c_j(x_j).$$

Ta thấy z là hàm phi tuyến của $x_1, \dots, x_n$ vì nó gián đoạn tại các điểm có một hoặc nhiều $x_j = 0$. Mặc dù các ràng buộc về nguyên liệu, thị trường ... là tuyến tính tất cả (chẳng hạn như ở Ví dụ 1.1.1) thì bài toán vẫn là phi tuyến và khó giải. Ta có thể đưa về quy hoạch tuyến tính nguyên bộ phận bằng cách đưa vào các biến phụ nhị phân

$$y_j = \begin{cases} 1 & \text{nếu } x_j > 0, \\ 0 & \text{nếu } x_j = 0. \end{cases} \quad (6.3)$$

Để biểu thị y_j là biến (nhị phân) độc lập, không phụ thuộc vào x_j như trên nữa, ta đưa vào một ràng buộc (cho mỗi j) tuyến tính như sau

$$x_j \leq My_j,$$

ở đây $M > 0$ coi là rất lớn để ràng buộc $x_j \leq \mu$ là thừa. Khi đó mục tiêu và ràng buộc (6.3) trở thành

$$\begin{aligned} \min z &= \sum_{j=1}^n (k_j y_j + c_j x_j), \\ 0 &\leq x_j \leq My_j, \\ y_j &= 0 \text{ hoặc } 1. \end{aligned} \quad (6.4)$$

Thật vậy, nếu $x_j > 0$ thì y_j không thể bằng 0 nên $y_j = 1$. Nếu $x_j = 0$ thì từ (6.4) ta thấy $y_j = 0$ hoặc $y_j = 1$. Nhưng vì $K_j > 0$ (nếu $K_j = 0$ thì không cần đưa vào biến phụ y_j) và mục tiêu là $\min z$ nên ở thuật toán tìm phương án tối ưu luôn lấy $y_j = 0$, vì phương án với $x_j = 0, y_j = 1$ không thể tối ưu. Khi viết đủ các ràng buộc tuyến tính khác vào ta được bài toán quy hoạch tuyến tính nguyên bộ phận.

THÍ DỤ 6.2.3 (Ràng buộc hoặc là - hoặc là: either - or constraints). Thực tế hay gặp trường hợp chẳng hạn một trong hai nguyên liệu đều dùng được khi đó chỉ cần một trong hai ràng buộc trữ lượng cần đưa vào, thí dụ như

$$\text{hoặc } 2x_1 + x_2 \leq 8,$$

$$\text{hoặc } x_1 + 4x_2 \leq 16,$$

(tức là ít nhất một ràng buộc thỏa). Gọi $M > 0$ là số rất lớn cố định và $y = 0$ hoặc 1 là biến bổ trợ thì ràng buộc trên trở thành

$$2x_1 + x_2 \leq 8 + My,$$

$$x_1 + 4x_2 \leq 16 + (1 - y),$$

$$y = 0 \text{ hoặc } 1,$$

và cả bài toán trở thành quy hoạch nguyên bộ phận.

Trường hợp tổng quát hơn là k trong N ràng buộc sau phải thỏa

$$f_1(x_1, \dots, x_n) \leq b_1,$$

....

$$f_N(x_1, \dots, x_n) \leq b_N.$$

Khi đó phải đưa vào N biến nhị phân phụ $y_1, \dots, y_N$ và đòi hỏi k trong N ràng buộc thỏa sẽ có dạng

$$f_1(x_1, \dots, x_n) \leq b_1 + My_1,$$

....

$$f_N(x_1, \dots, x_n) \leq b_N + My_N,$$

$\sum_{i=1}^N$

$$\sum_{i=1}^N y_i = N - K,$$

$i = 1$

$$y_i = 0 \text{ hoặc } 1, i = 1, \dots, N.$$

Cũng có thể chỉ một ràng buộc liên quan một biến như: hoặc là không sản xuất sản phẩm j , tức $x_j = 0$, hoặc phải sản xuất hơn số tối thiểu b_j để bù chi phí xin phép, và không quá d_j theo thông tin tiếp thị. Ràng buộc này sẽ có dạng

$$x_j \geq b_j y_j,$$

$$\begin{aligned}x_j &\leq d_j y_j, \\ y_j &= 0 \text{ hoặc } 1.\end{aligned}$$

Độc giả có thể xem nhiều thí dụ khác trong [Taha 1992], [Hillier - Lieberman 2001] và [Schrijver 1999].

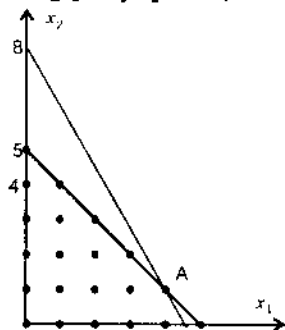
6.3 PHƯƠNG PHÁP NHÁNH-VÀ-CẶN CHO QUY HOẠCH NGUYÊN

6.3.1 Giới thiệu thuật toán trên một thí dụ

Thuật toán đầu tiên giải quy hoạch nguyên là thuật toán mặt phẳng cắt của Gomory công bố năm 1958 (tương tự tư tưởng của Dantzig, Fulkerson và Johnson nêu lên từ 1954). Nhưng từ khi phương pháp nhánh - và - cận (branch - and - bound method) xuất hiện trong [Land - Doig 1960], và nhất là dạng hoàn thiện trong [Dakin 1965], nó trở nên ưu thế rõ rệt. Ta hãy trình bày phương pháp qua một thí dụ trước

THÍ DỤ 6.3.1. Xét quy hoạch nguyên (gọi tắt IP)

$$\begin{aligned}\max z &= 5x_1 + 4x_2, \\ x_1 + x_2 &\leq 5, \\ 10x_1 + 6x_2 &\leq 45, \\ x_1, x_2 &\in \mathbf{Z}_+.\end{aligned}$$



H. 6.3

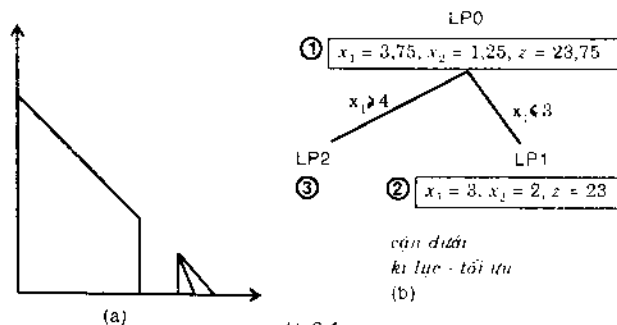
Gọi LP0 là quy hoạch tuyến tính tương ứng ở bước lập 0, tức là thay $x_1, x_2 \in \mathbf{Z}_+$ bằng $x_1, x_2 \geq 0$, có nghiệm tối ưu là $x_1 = 3,75, x_2 = 1,25$ và $z = 23,75$, xem H.6.3, ở đây A là điểm tối ưu của LP0 còn các dấu . là miền chấp nhận được

(cnd) của IP. Lấy bất kỳ một biến mà ở A có giá trị không nguyên, chẳng hạn $x_1 = 3,75$. Vì phần miền chấp nhận được của LP0 thỏa $3 < x_1 < 4$ không thể chứa nghiệm chấp nhận được của IP nên ta có thể bỏ không xét phần này và *chia nhánh* (branching) ra được hai miền và tương ứng là hai bài toán quy hoạch tuyến tính con ở bước 1 tiếp theo:

1. LP1 với miền cnd = cnd LP0, với $x_1 \leq 3$;

2. LP2 với miền cnd = cnd LP0, với $x_1 \geq 4$.

x_1 được gọi là *biến chia nhánh* (branching variable). H. 6.4(a) biểu diễn việc chia miền và (b) là sơ đồ cây của thuật toán ở bước đầu và đã được nghiệm tối ưu của IP (sau khi giải LP1 và lý luận mới hoàn thành H.6.4.(b)).



H. 6.4

Đến đây ta biết là nghiệm tối ưu của IP, phải nằm trong LP1 hoặc LP2 (tức là trong miền xác định của ít nhất một trong hai quy hoạch tuyến tính này). Ta đành giải bất kỳ, chẳng hạn LP₁, là LP0 thêm ràng buộc $x_1 \leq 3$:

$$\max z = 5x_1 + 4x_2,$$

$$x_1 + x_2 \leq 5,$$

$$10x_1 + 6x_2 \leq 45,$$

$$x_1 \leq 3,$$

$$x_1, x_2 \geq 0.$$

LP1 có thể giải bằng cách đưa về dạng chính tắc và dùng thuật toán đơn hình

tiêu chuẩn. Cũng có thể giải bằng phương pháp đơn hình áp dụng cho quy hoạch tuyến tính tổng quát, có chứa cận dạng $l_j \leq x_j \leq u_j$, xem [Khánh - Nường 2000], Chương 9. Nghiệm tối ưu của LP1 là $x_1 = 3$, $x_2 = 2$, $z = 23$, vô tình đã là nguyên. Đây là nghiệm tốt nhất tạm thời cho IP, ta gọi là một *ký lục* (incumbent). Do đó LP1 không thể suy ra nghiệm chấp nhận được tốt hơn cho IP, chỉ còn chờ so sánh với sự nghiên cứu LP2. Ta nói đã *làm chạm đáy* (fathoming) LP1. Việc "chờ so sánh" này cũng quan trọng không kém việc (nếu như) LP1 còn được thao tác tiếp (như LP2, sau đây sẽ thấy) vì nó cho một cận dưới của mục tiêu tối ưu của IP do $z = 23$ đã đạt được (bởi nghiệm nguyên), z tối ưu không thể nhỏ hơn. Vậy ở đây ta đã *đặt cận* (bounding) cho thuật toán. Tiếp theo, nói chung ta phải giải LP2 như sẽ làm dưới đây. Nhưng trong trường hợp may mắn này ta thấy cận dưới đã là $23 = [23, 75]$ ([.] ký hiệu phần nguyên), với 23,75 là mục tiêu tối ưu của LP0. Ở trường hợp đặc biệt là mọi hệ số mục tiêu của LP0, và cũng là của IP, đều nguyên, nên mục tiêu tối ưu của IP, là số nguyên, không thể tốt hơn $[23, 75] = 23$. Vậy ta đã may mắn được một nghiệm tối ưu của IP ngay sau một bước.

Nếu ta không chọn (bất kỳ) LP₁ như trên mà chọn LP₂, thì dù đã chọn (bất kỳ và may mắn một phần) biến chia nhánh là x_1 , thuật toán kéo dài như sau. Giải LP2 (là quy hoạch có được khi thay ràng buộc $x_1 \leq 3$ bởi $x_1 \geq 4$ ở LP1) bằng thuật toán đơn hình ta được $x_1 = 4$, $x_2 = 0,8333$, $z = 23,3333$. Vì $x_2 = 0,8333$ chưa nguyên, ta phải chia nhánh LP2 để được LP3 và LP4 với ràng buộc bổ sung tương ứng $x_2 \leq 0$ và $x_2 \geq 1$, tức là

1. Miền cnd LP3 = cnd LP0 thêm ràng buộc $x_1 \geq 4$, $x_2 \leq 0$;

2. Miền cnd LP4 = cnd LP0 thêm ràng buộc $x_1 \geq 4$, $x_2 \geq 1$.

Đến đây ta thấy hợp cả ba miền cnd của LP1, LP3 và LP4 chứa miền cnd của IP và còn phải xét cả ba. Có thể chọn bất kỳ để làm tiếp. Một trong các cách chọn hay dùng là chọn trong các bài toán con được lập sau cùng, ở đây là LP3 và LP4. Như vậy khi giải bằng phương pháp đơn hình có thể tận dụng kết quả ở lời giải LP2 ngay trước đó, vì so với LP2, ở LP3 và LP4 chỉ có thêm một ràng buộc. Do đó có thể áp dụng nghiên cứu hậu tối ưu lên nghiệm tối ưu của LP2 mà không cần giải lại (xem [Khánh - Nường 2000], Chương 8). Giả sử chọn

ngẫu nhiên LP4 và giải ta thấy LP4 không có nghiệm chấp nhận được. Tái nhiên sau này chẳng cần nhìn đến LP4 nữa, ta cũng nói nó đã chạm đáy. Tiếp theo, còn LP3 và LP1, ta lại được chọn bất kỳ, nhưng theo nguyên tắc mới lập hơn thì ta lấy LP3 và giải được nghiệm tối ưu của LP3 là $x_1 = 4,5$, $x_2 = 0$, $z = 22,5$. Vì x_1 chưa nguyên ta phải chia nhánh để được hai bài toán con:

1. Miền cnd LP5 = cnd LP0, thêm ràng buộc $x_1 \geq 4$, $x_2 \leq 0$, $x_1 \leq 4$;
2. Miền cnd LP6 = cnd LP0, thêm ràng buộc $x_1 \geq 4$, $x_2 \leq 0$, $x_1 \geq 5$.

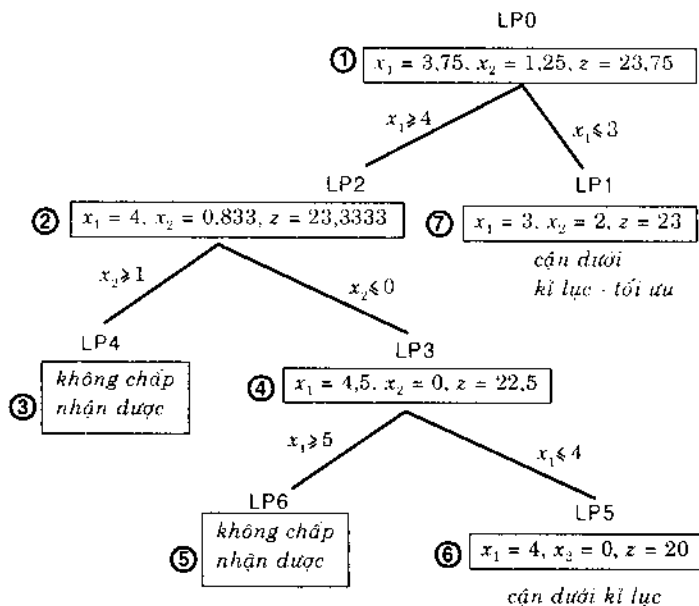
Ta còn ba bài toán con chưa xét LP1, LP5 và LP6, chọn bất kỳ theo nguyên tắc mới hơn xét LP6 ta thấy nó không chấp nhận được, tức là đã chạm đáy. Đến lượt giải LP5, mới hơn LP1, ta được nghiệm tối ưu là $x_1 = 4$, $x_2 = 0$, $z = 20$ thỏa cả ràng buộc về tính nguyên. Đến đây ta được một cận dưới của mục tiêu tối ưu của IP là $z = 20$ và do đó LP5 cũng chạm đáy, cung cấp một cận dưới cũng là một kỷ lục là $z = 20$. Còn lại duy nhất LP1 phải xét (xem có được kỷ lục mới không). Theo trên đã thấy giải LP1 ta được nghiệm tối ưu $x_1 = 3$, $x_2 = 2$, $z = 23$, cũng thỏa tính nguyên nên ta được một cận dưới và LP1 đã chạm đáy, cung cấp một kỷ lục mới. Vì không còn bài toán con nào để xét, kỷ lục mới này là nghiệm tối ưu của IP.

Như vậy cũng với việc chọn x_1 là biến phân nhánh nhưng chọn LP2 để xét thay cho LP1 ta đã phải làm dài nhất trong các khả năng có thể (nếu dùng quy tắc chọn trong các bài toán con mới lập) như H.6.5 trang bên minh họa sơ đồ cây thực hiện thuật toán.

6.3.2 Cách chia nhánh, đặt cận và chạm đáy

Mọi thuật toán nhánh - và - cận đều gồm ba thủ tục chia nhánh, đặt cận và chạm đáy. Mỗi thủ tục được linh động trong cách thực hiện nên có các dạng thuật toán khác nhau. Ta hãy giải thích thêm.

"Chia nhánh" là chọn một trong các bài toán con còn phải xét để chia tiếp ra. "Chọn" có thể lấy bất kỳ. Một quy tắc phổ biến là chọn trong các bài toán con mới lập. Vì việc đặt cận cũng có thể khác nhau,



H. 6.5

có nhiều cách định nghĩa cận (không chỉ là kỷ lục như định nghĩa trên), nên cũng sinh ra cách chọn khác là chọn trong các bài toán con có cận tốt nhất, với hy vọng sớm đi đến kỷ lục mới. "Chia" có nghĩa phổ biến là lấy một biến chia nhánh x_j để chia ra hai bài toán con với ràng buộc thêm $x_j \leq [a]$ và $x_j \geq [a] + 1$ tương ứng. Với bài toán nhị phân thì hai ràng buộc thêm sẽ là $x_j = 0$ và $x_j = 1$ như ở Mục 6.4. Tuy vậy, một số thuật toán chi tiết hơn có thể dùng các chiến thuật nào đó để lấy biến chia nhánh.

"Đặt cận" thường thực hiện bằng giải bài toán nói lỏng. Như phương pháp trên đây thì phải dùng LP - nói lỏng để đặt cận. Với các thuật toán nhánh - và - cận khác, còn có các kiểu bài toán nói lỏng khác. Chẳng hạn thuật toán nhánh - và - cận cho quy hoạch nhị phân ở Mục 6.4 có thể áp dụng *Lagrangian nói lỏng* (Lagrangian relaxation). Định nghĩa tổng quát Lagrangian nói lỏng như sau. Giả sử cần xét

quy hoạch nguyên

$$\begin{aligned}\max z &= c^T x, \\ Ax &\leq b, \\ x_j &\in \mathbf{Z}_+, j = 1, \dots, n.\end{aligned}$$

Ở đây $x \in \mathbf{R}^n$, A cấp $m \times n$. Lấy một $\lambda \geq 0$, $\lambda \in \mathbf{R}^m$ cố định. Khi đó Lagrangian nổi lờng được định nghĩa là bài toán

$$\max_{x \geq 0} L = c^T x - \lambda^T (Ax - b).$$

Rõ ràng ta có quan hệ hai mục tiêu tối ưu $z^* \leq L^*$, tức là L^* là một cận trên (có sát z^* hay không còn phụ thuộc λ đã chọn).

"Chạm đáy" có ba tiêu chuẩn như sau (cho bài toán tìm max).

- *Tiêu chuẩn 1:* Nếu cận của bài toán con $\leq$ mục tiêu kỷ lục, hoặc nói chung hơn là nếu mọi mục tiêu chấp nhận được của bài toán con $\leq$ mục tiêu kỷ lục, thì bài toán này chạm đáy.
- *Tiêu chuẩn 2:* Nếu bài toán con không chấp nhận được thì nó chạm đáy.
- *Tiêu chuẩn 3:* Nghiệm tối ưu của LP - nổi lờng của bài toán con là nguyên thì nó chạm đáy.

Rõ ràng nếu bài toán là tìm min thì ở Tiêu chuẩn 1 chỉ việc đổi dấu thành $\geq$, còn hai tiêu chuẩn kia giữ nguyên dạng.

6.3.3 Phát biểu thuật toán nhánh - và - cận

Xét bài toán quy hoạch nguyên tìm max. Gọi z là cận dưới - kỷ lục. Bước khởi động đặt $i = 0$ và đặt $z = -\infty$. Ở một bước lặp tổng quát có hai đoạn như sau.

Đoạn 1 (chạm đáy/đặt cận). Chọn LP_i là bài toán con sẽ xét. Giải LP_i và xét nó có chạm đáy không.

(a) Nếu LP_i chạm đáy theo Tiêu chuẩn 3 thì cập nhật z để được cận dưới mới. Trái lại, theo tiêu chuẩn khác, thì chọn bài toán con i mới và lặp như trên. Nếu không còn bài toán con chưa xét thì dừng. Nghiệm tối ưu của IP là tương ứng với cận dưới - kỷ lục cuối cùng, nếu có nghiệm như vậy.

(b) Nếu LP_i không chạm đáy, chuyển sang Đoạn 2 để chia nhánh LP_i .

Đoạn 2 (chia nhánh). Chọn một trong các biến x_j mà x_j^* ở nghiệm tối ưu LP_i chưa là số nguyên. Lập hai bài toán con LP - nối lỏng bằng cách thêm ràng buộc tương ứng

$$x_j \leq [x_j^*] \text{ và } x_j \geq [x_j^*] + 1.$$

Chuyển sang Đoạn 1.

NHẬN XÉT. Nếu quy hoạch là tìm min, thì ở thuật toán chỉ việc thay cận dưới z bởi cận trên $\bar{z}$ và ở bước khởi động đặt $\bar{z} = +\infty$. Các chi tiết khác không đổi.

Thuật toán trên cũng áp dụng cho quy hoạch nguyên bộ phận. Nếu biến nào không bị ràng buộc lấy giá trị nguyên thì không được chọn là biến phân nhánh. Giả sử ở Thí dụ 6.3.1, bây giờ x_2 là biến liên tục. Nếu ta chọn LP1 để giải được nghiệm tối ưu $x_1 = 3$, $x_2 = 2$, $z = 23$ thì đây là nghiệm chấp nhận được của quy hoạch nguyên bộ phận MIP. Tuy nhiên ta không thể làm chạm đáy LP1 để được nghiệm vừa có là tối ưu của MIP bằng lí luận như ở Mục 6.3.1 vì z của MIP có thể không nguyên. Ta phải giải LP2 để được nghiệm như ở H.6.5. Nghiệm này chấp nhận được cho MIP nên LP2 chạm đáy theo Tiêu chuẩn 3. $z = 23,3333$ của LP2 là kỷ lục mới, và là nghiệm tối ưu của MIP vì không còn bài toán con nào để xét.

6.4 PHƯƠNG PHÁP NHÁNH - VÀ - CẬN CHO QUY HOẠCH NHỊ PHÂN

Áp dụng cho quy hoạch nguyên nhị phân BIP, phương pháp nhánh - và - cận có một số đặc thù sau đây (cụ thể cho bài toán tìm max).

Ở thủ tục chia nhánh việc chọn bài toán con áp dụng cả quy tắc chọn bài toán mới lập nhất và (nếu cùng mới lập) thì chọn bài toán có cận tốt hơn. Lấy biến chia nhánh thường theo thứ tự chỉ số thấp trước. Với biến chia nhánh là x_j thì ràng buộc thêm cho hai bài toán con tương ứng là $x_j = 0$ và $x_j = 1$.

Ở thủ tục đặt cận, đặt cận cho mọi bài toán con là phần nguyên của mục tiêu tối ưu của LP - nơi lỏng hoặc Lagrange nơi lỏng.

Việc hình thành kỷ lục và thủ tục chạm đáy vẫn như quy hoạch nguyên tổng quát.

THÍ DỤ 6.4.1. Giải quy hoạch nhị phân

$$\begin{aligned}
 \max z &= 9x_1 + 5x_2 + 6x_3 + 4x_4, \\
 6x_1 + 3x_2 + 5x_3 + 2x_4 &\leq 10, \\
 x_3 + x_4 &\leq 1, \\
 -x_1 + x_3 &\leq 0, \\
 -x_2 + x_4 &\leq 0, \\
 x_j &= 0 \text{ hoặc } 1, j = 1, \dots, 4.
 \end{aligned} \tag{6.5}$$

Bước lập 1. Lấy biến chia nhánh là x_1 và lập hai bài toán con với ràng buộc bổ sung vào (6.5) là $x_1 = 0$ và $x_1 = 1$, tương ứng hai bài toán nơi lỏng là LP1 và LP2.

Ta đặt cận bằng cách giải các bài toán nổi lỏng bằng thuật toán đơn hình. LP0 (là (6.5) với ràng buộc nhị phân thay bởi $x_j \leq 1$ và $x_j \geq 0$, $j = 1, \dots, 4$) có nghiệm tối ưu là $(x_1, x_2, x_3, x_4) = (\frac{5}{5}, 1, 0, 1)$, $z = 16,5$. Do đó mọi mục tiêu chấp nhận được của (6.5) không lớn hơn 16. Vì vậy cận cho (6.5) là 16. Cũng giải tương tự ta có

$$\text{LP1: } (x_1, x_2, x_3, x_4) = (0, 1, 0, 1), z = 9 \text{ cận: } 9.$$

$$\text{LP2: } (x_1, x_2, x_3, x_4) = (1, \frac{4}{5}, 0, \frac{4}{5}), z = 16\frac{1}{5}, \text{ cận: } 16;$$

Ta thấy bài toán con 1 đã chạm đáy theo Tiêu chuẩn 3, ở biểu đồ ta viết F(3). Ta được kỷ lục là $(0, 1, 0, 1)$ và $z^* = 9$.

Bước lập 2. Chia nhánh từ bài toán con 2 còn lại với biến chia nhánh x_2 , ta được

$$\text{LP3 } (x_1 = 1, x_2 = 0): (x_1, x_2, x_3, x_4) = (1, 0, \frac{4}{5}, 0), z = 13\frac{4}{5}, \text{ cận } 13.$$

$$\text{LP4 } (x_1 = 1, x_2 = 1): (x_1, x_2, x_3, x_4) = (1, 1, 0, \frac{1}{2}), z = 16, \text{ cận } 16.$$

Cả hai bài toán con 3 và 4 chưa chạm đáy.

Bước lập 3. Vì còn lại hai bài toán con mới bằng nhau, ta chọn bài toán 4 có cận tốt hơn ($16 > 13$). Đến lượt x_3 là biến chia nhánh. Ta có

$$\text{LP5 } (x_1 = 1, x_2 = 1, x_3 = 0): (x_1, x_2, x_3, x_4) = (1, 1, 0, \frac{1}{2}), z = 16, \text{ là cận};$$

LP6 $(x_1 = 1, x_2 = 1, x_3 = 1)$: không có nghiệm chấp nhận được, F(2) (chạm đáy theo Tiêu chuẩn 2).

Bước lập 4. Còn bài toán con 3 và 5. Ta lấy bài toán 5 mới hơn để chia nhánh với biến chia nhánh là x_4 . Nhưng đây là biến cuối cùng nên ta được hai nghiệm cụ thể chứ không phải hai bài toán nữa:

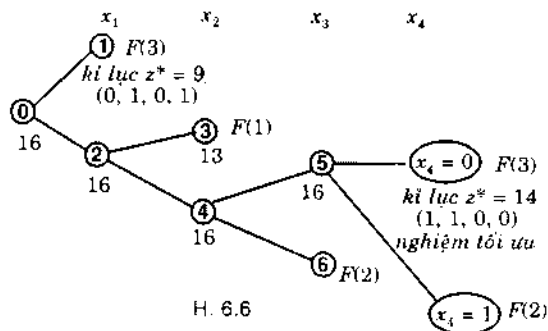
$x_4 = 0$: $(x_1, x_2, x_3, x_4) = (1, 1, 0, 0)$ chấp nhận được, $z = 14$ là cận;

$x_4 = 1$: $(x_1, x_2, x_3, x_4) = (1, 1, 0, 1)$ không chấp nhận được.

Thực hiện hình thức thủ tục chạm đáy ta thấy nghiệm đầu thỏa Tiêu chuẩn 3, và lập kỷ lục mới $(1, 1, 0, 0)$ với $z^* = 14$, nghiệm thứ hai thỏa Tiêu chuẩn 2.

Vì có kỷ lục mới ta kiểm tra Tiêu chuẩn 1 để xét chạm đáy lại với các bài toán còn lại, ở đây là bài toán con 3. Nó thỏa $z \leq z^*$ nên trở thành chạm đáy. Vậy tất cả đã chạm đáy. Nghiệm tối ưu của BIP là kỷ lục mới nhất.

Sơ đồ cây cho ở H.6.6. Ở đây chỉ cần ghi các nghiệm là kỷ lục, các nghiệm tối ưu của LP_i khác không cần ghi vì không liên quan đến chia nhánh như ở H.6.5. Nhưng ta ghi thêm các biến chia nhánh ở hàng trên cùng.



6.5 THUẬT TOÁN CỘNG GIẢI QUY HOẠCH NHỊ PHÂN

Vì quy hoạch nhị phân là quan trọng nhất trong áp dụng thực tế, ta xét thêm một thuật toán nữa cho quy hoạch này, đó là *thuật toán cộng* (additive algorithm). Tên của thuật toán có nguồn gốc từ việc thuật toán chỉ cần hai phép tính số học là cộng và trừ. Ở đây thuật toán được trình bày như một dạng của thuật toán nhánh - và - cận, với một số cải biên trong các thủ tục cụ thể. Để xác định, ta xét bài toán tìm min.

Trước hết, phải đưa về dạng đối ngẫu chấp nhận được, tức là mọi hệ số mục tiêu $c_j \geq 0$ (quy hoạch đối ngẫu chấp nhận được ngay ở từ vựng đầu tiên) và mọi ràng buộc có dạng bất đẳng thức $\leq$. Khi đó đưa biến bù w_i vào ta được

$$\min z = \sum_{j=1}^n c_j x_j,$$

$$\sum_{j=1}^n a_{ij} x_j + w_i = b_i, \quad i = 1, \dots, m,$$

$$x_j = 0 \text{ hoặc } 1, \quad j = 1, \dots, n,$$

$$w_i \geq 0, \quad i = 1, \dots, m.$$

(Nếu có $c_j < 0$ thì thay biến x_j sang biến mới $x'_j = 1 - x_j$ ta được bài toán thỏa yêu cầu trên.)

Nếu bài toán là gốc chấp nhận được, tức là mọi $b_i > 0$, thì rõ ràng nghiệm $x_j = 0, j = 1, \dots, n$ là tối ưu. Giả sử có $b_i < 0$. Ý tưởng của thuật toán cộng như sau. Xuất phát với $x_j = 0, j = 1, \dots, n$. Các w_i ứng với $b_i < 0$ sẽ là $w_i = b_i < 0$ nên nghiệm xuất phát chưa chấp nhận được. Phải tăng một số biến x_j lên 1. Ở các bước sẽ chọn biến thích hợp cho tăng lên 1 sao cho tiến về sự chấp nhận được, tức là $\forall w_i \geq 0$. Ta hãy giới thiệu thuật toán qua thí dụ cụ thể.

THÍ DỤ 6.5.1. Giải quy hoạch nhị phân

$$\max z = 3x_1 + 2x_2 + 5x_3 + 2x_4 + 3x_5,$$

$$-x_1 - x_2 + x_3 + 2x_4 - x_5 \leq 1,$$

$$-7x_1 + 3x_3 - 4x_4 - 3x_5 \leq -2,$$

$$11x_1 - 6x_2 - 3x_4 - 3x_5 \leq -1,$$

$$x_1, \dots, x_5 = 0 \text{ hoặc } 1.$$

Đưa biến bù w_1, w_2, w_3 vào và viết quy hoạch ở dạng bảng để dễ quan sát (dạng bảng này cũng hay gặp, hơi khác ở Mục 1.2.6 là đã đổi dấu, tức là chuyển về, hàng mục tiêu).

x_1	x_2	x_3	x_4	x_5	w_1	w_2	w_3	vế phải
3	2	5	2	3	0	0	0	z
-1	-1	1	2	-1	1	0	0	1
-7	0	3	-4	-3	0	1	0	-2
11	6	0	-3	-3	0	0	1	-1

Bước lặp 0. Xuất phát với $\forall x_j^0 = 0$, $(w_1^0, w_2^0, w_3^0) = (1, -2, -1)$, $z^0 = 0$.

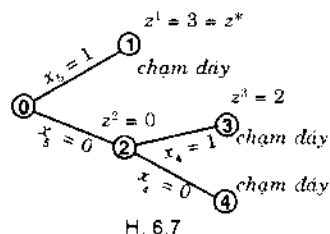
Vì nghiệm chưa chấp nhận được, ít nhất một x_j phải được nâng lên giá trị 1 để các w_i hướng tới không âm. Ta thấy các hệ số của x_3 đều không âm, nếu x_3 tăng từ 0 lên 1 thì để thỏa mãn ràng buộc các w_i sẽ nhỏ đi, tức là càng không chấp nhận được. Vậy x_3 phải bị giữ lại mức 0. Các biến x_1 , x_2 và x_4 đều có hệ số dương nên từng biến riêng không thể cho lên 1 để được $\forall w_i \geq 0$, nhưng một tổ hợp của chúng có thể làm cho $\forall w_i \geq 0$, nên ta chưa loại chúng như x_3 . Còn x_5 có mọi hệ số âm nên cho $x_5 = 1$.

Bước lặp 1 với $x_5 = 1$. Ta được nghiệm chấp nhận được với

$$(w_1^1, w_2^1, w_3^1) = (w_1^0 - (-1), w_2^0 - (-3), w_3^0 - (-3)) = (2, 1, 2), z^1 = 3.$$

Vậy $z^* = z^1 = 3$ là một cận trên, và là kỷ lục. Ta có thể biểu diễn bước lặp 1 như phần đầu của sơ đồ cây ở H.6.7. Và ta cũng lấy x_5 là biến chia nhánh. Bài toán con 1 với ràng buộc $x_5 = 1$ đã chạm đáy.

Bước lặp 2. Bài toán con 2 với $x_5 = 0$, tất nhiên có nghiệm trùng với nghiệm xuất phát. Nhưng thực chất bài toán 2 khác với bài toán 0 vì ở đây đã cố định $x_5 = 0$. Để chia nhánh từ bài toán con 2 còn lại (duy nhất) cũng làm như với bài



toán 0 nhưng có thêm một thông tin là chọn biến cho nâng lên 1 phải dẫn đến kết quả tốt hơn kỷ lục $z^* = 3$. Biến x_3 cũng bị loại như ở bước 1. x_1 có $c_1 = 3$ nên dẫn đến $z = 3$ không tốt hơn z^* cũng bị loại. Còn x_2 và x_4 , như bước 1 đã lý luận, từng biến riêng cho lên 1 chưa dẫn đến nghiệm chấp nhận, nên ta chọn biến nào dẫn đến "độ không chấp nhận được" ít nhất, tức là tổng mức âm ở w_1, w_2, w_3 của biến x_j .

$$v_j = \sum_i \min\{0, w_i - a_{ij}\}$$

ít nhất. Ta có

$$v_2 = 0 + (-2 - 0) + 0 = -2,$$

$$v_4 = (1 - 2) + 0 + 0 = -1.$$

Vậy x_4 là biến phân nhánh.

Bước lặp 3. Cho $x_4 = 1$ dẫn đến nút 3 ở H.6.7 với

$$(w_1^3, w_2^3, w_3^3) = (1 - 2, -2 + 4, -1 + 3) = (-1, 2, 2), z^3 = 2.$$

Ở nút này 3 biến tự do là x_1, x_2 và x_3 , có hệ số mục tiêu là 3, 2 và 5, tương ứng. Nếu nâng lên 1, từng biến sẽ dẫn đến mục tiêu tương ứng là $z = 2+3, 2+2$ và $2+5$, không hơn kỷ lục $z^* = 3$. Vì vậy nút 3 chạm đáy.

Bước lặp 4. Với $x_4 = 0$, ta được nút 4, có

$$(w_1^4, w_2^4, w_3^4) = (1, -2, -1), z^4 = 0.$$

Trong ba biến tự do x_1, x_2, x_3 thì x_3 đã biết là không làm cải thiện tính không chấp nhận được nên bị loại, x_1 có hệ số $c_1 = 3$, nếu đưa lên 1 thì được $z = 3$, không hơn kỷ lục, cũng bị loại. Còn x_2 có hệ số $a_{22} = 0$ cũng không làm thay đổi $w_2^4 = -2$ nếu đưa x_2 lên 1, nên cũng bị loại. Vậy nút 4 đã chạm đáy. Đã xét hết các bài toán con nên thuật toán kết thúc. Nghiệm tối ưu đạt được ở nút 1 với $z^* = 3$.

NHẬN XÉT (a) Trên đây ta đã trình bày thuật toán cộng khác với dạng ban đầu của nó để ghép vào dạng thuật toán nhánh - và - cận. Từ sơ đồ cây ở H. 6.7 ta thấy rất rõ bản chất nhánh - và - cận với ba thủ tục chia nhánh, đặt cận và chạm đáy. Nhưng thực hiện các thủ tục có khác với thuật toán nhánh - và - cận cơ bản ở Mục 6.4. Biến chia nhánh ở đây phải chọn phức tạp chứ không lấy tự nhiên như ở thuật toán cơ bản. Cận ở đây chỉ là kỷ lục và nhận được không phải do giải bài toán nối lỏng. Nghiệm

của các bài toán ở mỗi bước đều là nghiệm nhị phân; tính không chấp nhận được là ở chỗ không thỏa ràng buộc hàm (tức là các w_i âm)

$$\sum_{j=1}^n a_{ij}x_j + w_i \leq b_i \quad (\text{Các nghiệm này là đối ngẫu chấp nhận được.})$$

Tiêu chuẩn chấm dứt là: một bài toán con (nói khác là một nút ở biểu đồ) là chấm dứt nếu thỏa một trong hai tiêu chuẩn:

- Xét tiếp không dẫn đến mục tiêu tốt hơn kỷ lục.
- Xét tiếp không dẫn đến nghiệm chấp nhận được.

(b) Thuật toán cộng cũng có bản chất điểm diện ẩn. Ở đây mỗi khi được một nút chấm dứt là ta đã loại đi rất nhiều trường hợp trong số tất cả 2^n trường hợp. Chẳng hạn nút 1 (ở đó $x_5 = 1$) chấm dứt nên tất cả $2^{5-1} = 16$ trường hợp có $x_5 = 1$ bị loại. Nút 3 chấm dứt ($x_4 = 1$) loại được $2^{5-2} = 8$ trường hợp có $x_5 = 0, x_4 = 1$. Nút 4 chấm dứt loại được $2^{5-2} = 8$ trường hợp. Tổng cộng đã loại đủ $16+8+8 = 2^5 = 32$, tức là điểm diện hết mà thực ra chỉ tính cho 5 trường hợp.

6.6 PHƯƠNG PHÁP MẶT PHẪNG CẮT

6.6.1 Ý tưởng của phương pháp cắt

Ta minh họa ý tưởng của *phương pháp mặt phẳng cắt* (cutting plane method) do Dantzig, Fulkerson và Johnson đưa ra 1954 bằng thí dụ sau.

THÍ DỤ 6.6.1. Xét quy hoạch nguyên

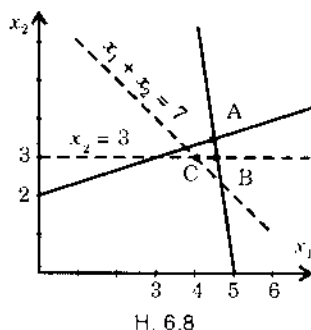
$$\begin{aligned} \max z &= 7x_1 + 9x_2, \\ -x_1 + 3x_2 &\leq 6, \end{aligned}$$

$$7x_1 + x_2 \leq 35,$$

$$x_1, x_2 \geq 0,$$

$$x_1, x_2 \text{ nguyên.}$$

Bỏ đi ràng buộc về tính nguyên, giải quy hoạch tuyến tính (liên tục) tương ứng bằng cách đưa vào biến bù x_3, x_4 ta đi đến bảng tối ưu



cơ sở	x_1	x_2	x_3	x_4	nghiệm
z	0	0	28/11	15/11	63
x_2	0	1	7/12	1/22	7/2
x_1	1	0	-1/22	3/22	9/2

Nghiệm là điểm $A(9/2, 7/2)$ trên H.6.8.

Ý tưởng của phương pháp mặt phẳng cắt là mỗi bước lặp k sẽ dùng một siêu phẳng cắt để thu nhỏ miền chấp nhận được ở bước trước ($k = 1$ là miền của quy hoạch liên tục) thỏa hai yêu cầu sau đây.

1. Mọi phương án chấp nhận được của quy hoạch nguyên không bị cắt đi.
2. Nghiệm tối ưu (chưa nguyên) x^k vừa nhận ở bước này phải bị cắt đi.

Ở thí dụ 6.6.2 ta sẽ áp dụng phương pháp cho quy hoạch nguyên trên đây và sẽ thấy mặt phẳng cắt ở hai bước lặp lần lượt là $x_2 = 3$ và $x_1 + x_2 = 7$. Sau hai bước, tức là thêm lần lượt ràng buộc $x_2 \leq 3$ và $x_1 + x_2 \leq 7$ ta được lần lượt đỉnh tối ưu là $B(4\frac{4}{7}, 3)$ và $C(4, 3)$. C đã nguyên nên là nghiệm tối ưu của quy hoạch nguyên, xem H.6.8.

Năm 1958 Ralph Gomory đã thật sự xây dựng thuật toán đảm bảo được yêu cầu 3: Sau hữu hạn bước nhận được đỉnh tối ưu nguyên. Tất nhiên đó là nghiệm tối ưu của quy hoạch nguyên, xem [Gomory 1958] hoặc [Schrijver 1999], Mục 23.8.

6.6.2 Thuật toán phân số giải quy hoạch nguyên hoàn toàn

Yêu cầu cơ bản để tiến hành *thuật toán phân số* (cho quy hoạch nguyên hoàn toàn) (fractional (pure integer) algorithm) là phải đưa quy hoạch về dạng mọi dữ liệu (tức là các hệ số ràng buộc và mục tiêu cũng như vế phải ràng buộc) đều là số nguyên. Nội dung chính là ở mỗi bước lặp ta sẽ xây dựng một lát cắt với mọi hệ số là phân số (do đó sinh ra tên của thuật toán) và thực hiện một phép xoay đơn hình đối ngẫu. Thuật toán do Gomory đưa ra nên cũng gọi là *thuật toán lát cắt Gomory*.

Bước chuẩn bị. Ta giải quy hoạch tuyến tính nới lỏng (tức là chỉ bỏ điều kiện nguyên) bằng phương pháp đơn hình. Nếu nhận được nghiệm tối ưu đã nguyên thì chính là nghiệm tối ưu của quy hoạch nguyên. Trái lại thì ta được bảng tối ưu của LP nới lỏng như dưới đây (để tiện ta ký hiệu các biến cơ sở ở bảng tối ưu là w_i , biến không cơ sở là x_j ; nhưng khi tiến hành thuật toán không cần phân biệt; dạng bảng này như ở Mục 6.5, có hàng mục tiêu trái dấu với dạng ở Mục 1.2.6).

cơ sở	$w_1 \dots w_i \dots w_m$	$x_1 \dots x_j \dots x_n$	nghiệm
z	$0 \dots 0 \dots 0$	$c_1 \dots c_j \dots c_n$	z_0
w_1	$1 \quad 0 \quad \dots \quad 0$	$a_{11} \quad \dots \quad a_{1j} \quad \dots \quad a_{1n}$	b_1
$\vdots$	$\vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots$	$\vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots$	$\vdots$
w_i	$0 \quad \dots \quad 1 \quad \dots \quad 0$	$a_{i1} \quad \dots \quad a_{ij} \quad \dots \quad a_{in}$	b_i
$\vdots$	$\vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots$	$\vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots$	$\vdots$
w_m	$0 \quad \dots \quad 0 \quad \dots \quad 1$	$a_{m1} \quad \dots \quad a_{mj} \quad \dots \quad a_{mn}$	b_m

Bước 1. Đoạn 1. Xét một phương trình thứ i mà b_i chưa nguyên tức là biến tối ưu w_i nhận được chưa nguyên (có thể là hàng z).

$$w_i = b_i - \sum_{j=1}^n a_{ij}x_j \quad (6.6)$$

và gọi là một *hàng nguồn* (source row). Đặt

$$a_{ij} = [a_{ij}] + \alpha_{ij}, \quad b_i = [b_i] + \beta_i$$

ở đây $[]$ là kí hiệu phần nguyên của một số thực. (6.6) suy ra

$$\beta_i - \sum_{j=1}^n \alpha_{ij}x_j = w_i - [b_i] + \sum_{j=1}^n [a_{ij}]x_j. \quad (6.7)$$

Để các x_j và w_i trở thành nguyên không âm cả thì vế phải (6.7) phải là nguyên và do đó vế trái cũng vậy. Do đó vế trái

$$\beta_i - \sum \alpha_{ij}x_j \leq \beta_i$$

phải nhỏ hơn 1 theo định nghĩa β_i và lại phải nguyên nên

$$\beta_i - \sum \alpha_{ij}x_j \leq 0. \quad (6.8)$$

Như vậy để hướng về tính chấp nhận được là nguyên, xét một ràng buộc (6.6) còn vi phạm, ta dẫn đến phải thêm ràng buộc (6.8), chính là mặt phẳng cắt. Đưa thêm biến bù $S_i \geq 0$ vào, (6.8) trở thành

$$S_i = \sum_{j=1}^n \alpha_{ij}x_j - \beta_i, \quad (6.9)$$

được gọi là *mặt cắt phân số* (fractional cut). Bổ sung ràng buộc (6.9) vào bảng đơn hình ta được bảng mới.

Bảng này là bài toán mới đặt ra nhằm đưa bảng tối ưu liên tục (tức là chưa nguyên) về tính nguyên. Nhưng chính bảng này vừa chưa nguyên vừa chưa đảm bảo ràng buộc dấu.

Đoạn 2. Áp dụng thuật toán đơn hình đối ngẫu tiến tới thỏa ràng buộc dấu (vì điều kiện tối ưu đã đảm bảo, tức là $\forall c_j \geq 0$). Có thể phải xoay nhiều lần, cho đến khi đạt được ràng buộc dấu thỏa. Lúc này, theo lý luận định tính ở trên về mặt cắt cũng đã tiến được một bước về thỏa tính nguyên. Nếu thuật toán đơn hình đối ngẫu dẫn tới không có nghiệm chấp nhận được về dấu thì dừng. Quy hoạch nguyên là vô nghiệm. Trái lại, thì sang bước lập tiếp theo gồm hai đoạn như trên (cắt và xoay đơn hình đối ngẫu).

Gomory đã chứng minh là thuật toán sẽ kết thúc sau hữu hạn bước. Ở đây ta không trình bày. Trong nhiều trường hợp, phụ thuộc vào kiểu dữ liệu, thuật toán mặt phẳng cắt tỏ ra kém hiệu quả hơn thuật toán nhánh - và - cận.

THÍ DỤ 6.6.2. Ta hãy giải cụ thể Thí dụ 6.6.1. Bước khởi động đã thực hiện ở Thí dụ 6.6.1, được nghiệm tối ưu chưa nguyên.

Bước 1. Đoạn 1. Có thể chọn hàng i bất kỳ có b_i không nguyên nhưng thường chọn hàng ứng với $\max \beta_i$. Ở đây $\beta_1 = \beta_2 = \frac{1}{2}$ nên ta lấy bất kỳ phương trình với x_2 :

$$x_2 + \frac{7}{22}x_3 + \frac{1}{22}x_4 = 3\frac{1}{2},$$

tức là

$$x_2 + (0 + \frac{7}{22})x_3 + (0 + \frac{1}{22})x_4 = 3 + \frac{1}{2}.$$

Do đó mặt cắt phân số là

$$S_1 - \frac{7}{22}x_3 - \frac{1}{22}x_4 = -\frac{1}{2}.$$

Bổ sung ràng buộc này vào bảng ta được

cơ sở	x_1	x_2	x_3	x_4	S_1	nghiệm
z	0	0	28/11	15/11	0	63
x_2	0	1	7/22	1/22	0	$3\frac{1}{2}$
x_1	1	0	-1/22	3/22	0	$4\frac{1}{2}$
S_1	0	0	-7/22	-1/22	1	-1/2

Đoạn 2. Thực hiện phép xoay đơn hình đối ngẫu với S_1 ra, x_3 vào ta được

Cơ sở	x_1	x_2	x_3	x_4	S_1	nghiệm
z	0	0	0	1	8	59
x_2	0	1	0	0	1	3
x_1	1	0	0	1/7	-1/7	$4\frac{4}{7}$
x_3	0	0	1	1/7	-22/7	$1\frac{4}{7}$

Bước 2. Đoạn 1. $\beta_1 > \beta_3$ nên lập cắt mới từ phương trình x_1

$$x_1 + (0 + 1/7) x_4 + (-1 + 6/7) S_1 = 4 + \frac{4}{7}$$

ta được mặt cắt

$$S_2 - \frac{1}{4}x_4 - \frac{6}{7}S_1 = -\frac{4}{7}.$$

Bổ sung vào bảng đơn hình dẫn đến

Cơ sở	x_1	x_2	x_3	x_4	S_1	S_2	nghiệm
z	0	0	0	1	8	0	59
x_2	0	1	0	0	1	0	3
x_1	1	0	0	$\frac{1}{7}$	$-\frac{1}{7}$	0	$4\frac{4}{7}$
x_3	0	0	1	$\frac{1}{7}$	$-\frac{22}{7}$	0	$1\frac{4}{7}$
S_2	0	0	0	$-\frac{1}{7}$	$-\frac{6}{7}$	1	$-\frac{4}{7}$

Đoạn 2. Thực hiện phép xoay đơn hình đối ngẫu với S_2 ra và x_4 vào ta được

cơ sở	x_1	x_2	x_3	x_4	S_1	S_2	thử nghiệm
z	0	0	0	0	2	7	55
x_2	0	1	0	0	1	0	3
x_1	1	0	0	0	-1	1	4
x_3	0	0	1	0	-4	1	1
x_4	0	0	0	1	6	-7	4

Bảng này đã cho nghiệm tối ưu (chấp nhận được) nguyên là $x_1 = 4$, $x_2 = 3$, $z = 55$.

Đối sang biến x_1 , x_2 ban đầu dễ thấy hai cắt phân số ở hai bước ở trên chính là

$$S_1 + x_3 = 3 \text{ hay } x_2 \leq 3,$$

$$x_1 + x_2 \leq 7.$$

đúng như H.6.8.

CHÚ Ý 6.6.1. Một bảng đơn hình có thể có nhiều hàng nguồn (tức là có b_i không nguyên) nên có thể sinh ra nhiều cắt. Vậy dùng cắt nào là mạnh nhất tức là vào sâu trong miền chấp nhận được nhất. Định nghĩa toán học về cắt mạnh hơn là như sau. Xét hai cắt

$$\sum_{j=1}^n \alpha_{ij} x_j \geq \beta_i, \quad (6.10)$$

$$\sum_{j=1}^n \alpha_{kj} x_j \geq \beta_k. \quad (6.11)$$

Cắt (6.10) là *mạnh hơn* (stronger) cắt (6.11) nếu $\beta_i \geq \beta_k$ và $\alpha_{ij} \leq \alpha_{kj}$, $\forall j$, và có ít nhất một dấu <. Khó thực hiện tính toán theo định nghĩa này nên thường chọn cắt "tốt hơn" đạt $\max_i \beta_i$ hoặc $\max_i \left\{ \frac{\beta_i}{\sum_{j=1}^n \alpha_{ij}} \right\}$.

THÍ DỤ 6.6.3. Ở Bước 1 của Thí dụ 6.6.2 ta chọn cắt theo $\max_i \beta_i$ nên lấy

bất kỳ phương trình x_2 (vì $\beta_1 = \beta_2 = \frac{1}{2}$). Nếu ta chọn theo quy tắc $\max_i \left\{ \frac{\beta_i}{\sum_{j=1}^n \alpha_{ij}} \right\}$ thì

$$\text{do} \quad \frac{\frac{1}{2}}{\frac{7}{22} + \frac{1}{22}} > \frac{\frac{1}{2}}{\frac{21}{22} + \frac{3}{22}}$$

ta phải chọn phương trình x_2 chứ không phải bất kỳ. Nếu ta lấy phương trình x_1 sẽ dẫn đến cắt $x_2 \leq \frac{10}{3}$ không sâu bằng.

6.6.3 Thuật toán hỗn hợp

Bây giờ xét thuật toán cắt cho quy hoạch nguyên bộ phận. Giả sử w_k là biến nguyên (có ràng buộc nguyên) mà về phải b_k ở bảng tối ưu của LP nới lỏng chưa nguyên. Khi đó ta được hàng nguồn

$$w_k = b_k - \sum_{j=1}^n a_{kj} x_j = [b_k] + \beta_k - \sum_{j=1}^n a_{kj} x_j,$$

tức là

$$w_k - [b_k] = \beta_k - \sum_{j=1}^n a_{kj} x_j \quad (6.12)$$

(chú ý là một số x_j có thể không bị ràng buộc nguyên).

Để xây dựng cắt ta lý luận như sau. Để w_k trở thành nguyên thì hoặc $w_k \leq [b_k]$ hoặc $w_k \geq [b_k] + 1$. Theo hàng nguồn (6.12) thì hai điều kiện này tương đương với

$$\sum_{j=1}^n a_{kj} x_j \geq \beta_k, \quad (6.13)$$

$$\sum_{j=1}^n a_{kj} x_j \leq \beta_k - 1. \quad (6.14)$$

Đặt $J^+ = \{j: a_{kj} \geq 0\}$, $J^- = \{j: a_{kj} < 0\}$

thì (6.13) và (6.14) suy ra

$$\sum_{j \in J^+} a_{kj} x_j \geq \beta_k, \quad (6.15)$$

$$\frac{\beta_k}{\beta_k - 1} \sum_{j \in J^-} a_{kj} x_j \geq \beta_k. \quad (6.16)$$

Vì (6.15) và (6.16) không thể xảy ra đồng thời chúng có thể được ghép thành một ràng buộc ở dạng

$$S_k - \left\{ \sum_{j \in J^+} a_{kj} x_j + \frac{\beta_k}{\beta_k - 1} \sum_{j \in J^-} a_{kj} x_j \right\} = -\beta_k, \quad (6.17)$$

ở đây $S_k \geq 0$ là biến bù. (6.17) là điều kiện cần để w_k là nguyên và gọi là một *cắt hỗn hợp* (mixed cut). Vì Bảng tối ưu cho nghiệm (tối ưu) tương ứng với $x_j = 0$, nên cắt này chưa chấp nhận được. Cũng như ở thuật toán phân số ta dùng thuật toán đơn hình đối ngẫu để tiến tới chấp nhận được.

Nếu tính đến việc một số x_j là biến nguyên ta có thể chứng minh được là phương trình sau xác định một cắt mạnh hơn:

$$S_k = -\beta_k + \sum_{j=1}^n \lambda_j x_j,$$

ở đây

$$\lambda_j = \begin{cases} a_{kj} & \text{nếu } a_{kj} \geq 0 \text{ và } x_j \text{ không nguyên,} \\ \frac{\beta_k}{\beta_k - 1} a_{kj} & \text{nếu } a_{kj} < 0 \text{ và } x_j \text{ không nguyên,} \\ \alpha_{kj} & \text{nếu } \alpha_{kj} \leq \beta_k \text{ và } x_j \text{ nguyên,} \\ \frac{\beta_k}{1 - \beta_k} (1 - \beta_k) & \text{nếu } \alpha_{kj} > \beta_k \text{ và } x_j \text{ nguyên.} \end{cases}$$

Xem chứng minh trong [Taha 1975].

THÍ DỤ 6.6.4. Xét Thí dụ 6.6.1 nhưng bây giờ chỉ x_1 bị ràng buộc nguyên. Ràng buộc ứng với x_1 ở bảng tối ưu LP - nổi lồi (Thí dụ 6.6.2) là

$$x_1 - \frac{1}{22}x_3 + \frac{3}{22}x_4 = 4 + \frac{1}{2},$$

$$J^+ = \{4\}, J^- = \{3\}, \beta_1 = \frac{1}{2}.$$

Do đó cắt hỗn hợp là

$$S_1 - \left\{ \frac{3}{22}x_4 + \left(\frac{\frac{1}{2}}{\frac{1}{2} - 1} \right) \left(-\frac{1}{22}x_3 \right) \right\} = -\frac{1}{2}.$$

tức là

$$S_1 - \frac{1}{22}x_3 - \frac{3}{22}x_4 = -\frac{1}{2}.$$

Bổ sung ràng buộc này vào bảng ta được

cơ sở	x_1	x_2	x_3	x_4	S_1	nhịệm
z	0	0	28/11	15/11	0	63
x_2	0	1	7/22	1/22	0	7/2
x_1	1	0	-1/22	3/22	0	9/2
S_1	0	0	-1/22	-3/22	1	-1/2

Tiếp theo, áp dụng thuật toán đơn hình đối ngẫu với S_1 ra x_4 vào dẫn tới bảng tối ưu cho quy hoạch nguyên bộ phận

cơ sở	x_1	x_2	x_3	x_4	S_1	nhịệm
z	0	0	23/11	0	10	58
x_2	0	1	10/33	0	-1/3	10/3
x_1	1	0	-1/11	0	1	4
x_4	0	0	1/3	1	-22/3	11/3

CHÚ Ý 6.6.2. Ngoài các cắt đã trình bày người ta đã nghiên cứu một số cắt khác. Tuy nhiên, phương pháp mặt phẳng cắt kém hiệu quả hơn phương pháp nhánh - và - cận và chỉ thích hợp cho một số dạng bài toán cỡ nhỏ. Nói chung thì các thuật toán giải quy hoạch nguyên không đảm bảo giải

hiệu quả các bài toán 100 biến, khác xa với thuật toán đơn hình cho quy hoạch tuyến tính (không nguyên), hiệu quả cho cả các bài toán hàng trăm ngàn biến và hàng chục ngàn ràng buộc. Người ta cũng đã phối hợp hai phương pháp nhánh - và - cận và mặt phẳng cắt, chẳng hạn phương pháp nhánh - và - cắt (branch - and - cut method) khá hiệu quả, xem [Hoffman - Padberg 1991].

6.7 QUY HOẠCH NHỊ PHÂN ĐA THỨC

Từ đầu chương đến giờ ta chỉ xét quy hoạch tuyến tính. Bây giờ ta mở rộng xét quy hoạch 0-1 đa thức (zero-one polynomial programming) sau

$$\max z = f(x_1, \dots, x_n),$$

$$g_i(x_1, \dots, x_n) \leq b_i, i = 1, \dots, m,$$

$$x_j = 0 \text{ hoặc } 1, j = 1, \dots, n,$$

ở đây f và g_i là các đa thức, với số hạng thứ k có dạng $d_k \prod_{j=1}^{n_k} x_j^{a_{kj}}$,

$a_{kj} > 0$. Dễ dàng đưa bài toán về quy hoạch 0-1 tuyến tính như sau:

Vì $x_j = 0$ hoặc 1 nên $x_j^{a_{kj}} = x_j$. $\forall a_{kj} > 0$ nên $d_k \prod_{j=1}^{n_k} x_j^{a_{kj}} = d_k x_1 \dots x_{n_k}$. Lại

đặt $y_k = x_1 \dots x_{n_k}$ thì y_k cũng bằng 0 hoặc 1 nên ta đã được quy hoạch

0-1 tuyến tính. Để đưa vào biến y_k ta phải xác định ràng buộc. Vì $y_k = 1$ khi mọi $x_j = 1, j = 1, \dots, n_k$ và $y_k = 0$ trong các trường hợp khác, dễ kiểm tra được ràng buộc này có thể viết ở dạng

$$\sum_{j=1}^{n_k} x_j - (n_k - 1) \leq y_k,$$

$$\frac{1}{n_k} \sum_{j=1}^{n_k} x_j \geq y_k.$$

THÍ DỤ 6.7.1. Đưa quy hoạch sau về quy hoạch 0-1 tuyến tính

$$\max z = 2x_1x_2x_3^3 + x_1^2x_2,$$

$$5x_1 + 9x_2^2x_3 \leq 15,$$

$$x_1, x_2, x_3 = 0 \text{ hoặc } 1.$$

Đặt $y_1 = x_1x_2x_3$, $y_2 = x_1x_2$ và $y_3 = x_2x_3$. Quy hoạch trở thành tuyến tính:

$$\max z = 2y_1 + y_2,$$

$$5x_1 + 9y_3 \leq 15,$$

$$x_1 + x_2 + x_3 - 2 \leq y_1,$$

$$\frac{1}{3}(x_1 + x_2 + x_3) \geq y_1,$$

$$x_1 + x_2 - 1 \leq y_2,$$

$$\frac{1}{2}(x_1 + x_2) \geq y_2,$$

$$x_2 + x_3 - 1 \leq y_3,$$

$$\frac{1}{2}(x_2 + x_3) \geq y_3,$$

$$x_1, x_2, x_3, y_1, y_2, y_3 = 0 \text{ hoặc } 1.$$

BÀI TẬP CHƯƠNG 6

6.1 Đưa hệ ràng buộc sau về một ràng buộc tương đương

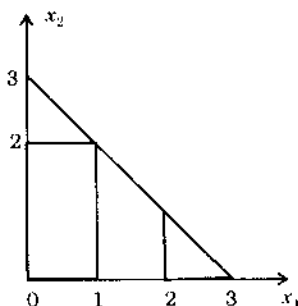
$$3x_1 + 2x_2 \leq 10,$$

$$x_1 + 4x_2 \leq 11,$$

$$3x_1 + 3x_2 \leq 13,$$

$$x_1, x_2 \geq 0, \text{ nguyên.}$$

6.2 Cần phải đưa hàng đến 5 địa điểm bằng 6 đường. Mỗi đường có thể đến một số điểm như sau: đường 1: 1, 2, 3, 4; đường 2: 1, 3, 4, 5; đường 3: 1, 2, 5; đường 4: 2, 3, 5; đường 5: 1, 2, 3 và đường 6: 1, 3, 5. Khoảng cách tới 5 địa điểm tương ứng là 10, 5, 9, 13 và 6 km. Cần phải mang hàng đến mỗi địa điểm một lần sao cho tổng khoảng cách là ngắn nhất. Hãy phát biểu bài toán thành quy hoạch nguyên.



H. 6.9

6.3 Biểu diễn miền ràng buộc không lỗi ở H.6.9 bằng phần được tô đậm bởi các ràng buộc dạng giải tích. Tìm nghiệm tối ưu cho bài toán với ràng buộc này và mục tiêu là

$$\max z = 2x_1 + 3x_2.$$

(Gợi ý: Dùng ràng buộc hoặc là - hoặc là.)

6.4 Giải bằng phương pháp nhánh và cận. Hãy chọn x_1 là biến phân nhánh ở nút 0.

(a) $\max z = 3x_1 + 2x_2,$

$$2x_1 + 2x_2 \leq 9,$$

$$3x_1 + 3x_2 \leq 18,$$

$$x_1, x_2 \in \mathbf{Z}_+.$$

(b) $\max z = x_1 + x_2,$

$$2x_1 + 5x_2 \leq 16,$$

$$6x_1 + 5x_2 \leq 30,$$

$$x_1, x_2 \in \mathbf{Z}_+.$$

$$(c) \quad \min z = 5x_1 + 4x_2,$$

$$4x_1 + 2x_2 \geq 6,$$

$$2x_1 + 3x_2 \geq 8,$$

$$x_1, x_2 \in \mathbf{Z}_+.$$

6.5 Giải bằng phương pháp nhánh và cận các quy hoạch ở Bài tập 6.4 nhưng bây giờ chỉ x_2 bị ràng buộc nguyên.

6.6 Diễn đạt bài toán chất hàng

(cargo - loading problem) sau đây

thành quy hoạch nguyên và giải bằng phương pháp nhánh và cận.

Có 5 loại hàng phải chất lên tàu.

Trọng lượng w_i thể tích v_i và giá trị

r_i đơn vị hàng được cho bằng bảng

bên. Trọng tải và thể tích cho hàng

của tàu là 112 và 109 tương ứng.

Hãy xác định lượng đơn vị các loại

hàng chất lên sao cho giá trị là cực

đại.

loại hàng	w_i	v_i	r_i
1	5	1	4
2	8	8	7
3	3	6	6
4	2	5	5
5	7	4	4

6.7 Xét quy hoạch

$$\max z = -3x_1 + 5x_2,$$

$$5x_1 - 7x_2 \geq 3,$$

$$0 \leq x_j \leq 3, j = 1, 2,$$

$$x_j \text{ nguyên, } j = 1, 2.$$

(a) Giải bằng hình học.

(b) Giải bằng thuật toán nhánh - và - cận (giải các LP nới lỏng bằng hình học hoặc thuật toán đơn hình).

(c) Sử dụng biểu diễn nhị phân cho biến nguyên sau đây để đưa bài toán về quy hoạch nhị phân rồi giải quy hoạch nhị phân này bằng thuật toán nhánh - và - cận. Giả

sử biến nguyên x bị ràng buộc $0 \leq x \leq u$ và N là số nguyên sao cho $2^N \leq u < 2^{N+1}$. Khi đó biểu diễn nhị phân của x là $x = \sum_{i=0}^N 2^i y_i$, y_i bằng 0 hoặc 1.

6.8 (a) Chứng minh rằng mọi nghiệm chấp nhận được của MIP cũng là chấp nhận được của Lagrange nổi lòng.

(b) Chứng minh rằng $z^* \geq z_{\text{LP}}^*$, ở đây z^* và z_{LP}^* tương ứng là giá trị mục tiêu tối ưu của MIP và Lagrange nổi lòng.

6.9 Xét bài toán phân việc (bổ nhiệm) với bảng tiến công (cước phí)

Việc người	1	2	3	4	5
1	39	65	69	66	57
2	64	84	24	92	22
3	49	50	61	31	45
4	48	45	55	23	50
5	59	34	30	34	18

(a) Thiết kế thuật toán nhánh - và - cận giải bài toán bằng việc mô tả cụ thể các thủ tục chia nhánh, đặt cận và làm chậm đáy. (Gợi ý: Ứng với những người chưa được phân việc ở bài toán con hiện hành hãy lập bài toán nổi lòng bằng cách bỏ các ràng buộc rằng mỗi người này đều nhận đúng một việc.)

(b) Dùng thuật toán giải cụ thể bài toán đã cho.

6.10 Giải quy hoạch nhị phân sau bằng thuật toán nhánh - và - cận.

(a) $\max z = 2x_1 - x_2 + 5x_3 - 3x_4 + 4x_5,$

$$3x_1 - 2x_2 + 7x_3 - 5x_4 + 4x_5 \leq 6,$$

$$x_1 - x_2 + 2x_3 - 4x_4 + 2x_5 \leq 0,$$

$$x_j = 0 \text{ hoặc } 1, j = 1, \dots, 5.$$

(b) $\min z = 5x_1 + 6x_2 + 7x_3 + 8x_4 + 9x_5,$

$$3x_1 - x_2 + x_3 + x_4 - 2x_5 \geq 2,$$

$$x_1 + 3x_2 - x_3 - 2x_4 + x_5 \geq 0,$$

$$-x_1 - x_2 + 3x_3 + x_4 + x_5 \geq 1,$$

$$x_j = 0 \text{ hoặc } 1, j = 1, \dots, 5.$$

6.11 Giải bằng thuật toán cộng.

$$(a) \quad \max z = 3x_1 + x_2 + 3x_3,$$

$$-x_1 + 2x_2 + x_3 \leq 4,$$

$$4x_2 - 3x_3 \leq 2,$$

$$x_1 - 3x_2 + 2x_3 \leq 3,$$

$$x_1, x_2, x_3 = 0 \text{ hoặc } 1.$$

$$(b) \quad \max z = 2x_1 + x_2,$$

$$10x_1 + 10x_2 \leq 9,$$

$$10x_1 + 5x_2 \geq 1,$$

$$x_1, x_2 = 0 \text{ hoặc } 1.$$

$$(c) \quad \max z = 20x_1 + 40x_2 + 20x_3 + 15x_4 + 30x_5,$$

$$5x_1 + 4x_2 + 3x_3 + 7x_4 + 8x_5 \leq 25,$$

$$x_1 + 7x_2 + 9x_3 + 4x_4 + 6x_5 \leq 25,$$

$$8x_1 + 10x_2 + 2x_3 + x_4 + 10x_5 \leq 25,$$

$$x_j = 0, \text{ hoặc } 1, j = 1, \dots, 5.$$

6.12 Xét

$$\max z = x_1 + 2x_2,$$

$$x_1 + \frac{1}{2}x_2 \leq \frac{13}{4},$$

$$x_1, x_2 \in \mathbb{Z}_+.$$

Chứng tỏ rằng thuật toán phân số không tìm ra nghiệm chấp nhận được trừ khi các hệ số và vế phải ràng buộc là nguyên. Đưa bài toán về dạng như thế rồi giải.

6.13 Giải bằng thuật toán phân số

$$\max z = 4x_1 + 6x_2 + 2x_3,$$

$$4x_1 - 4x_2 \leq 5,$$

$$-x_1 + 6x_2 \leq 5,$$

$$-x_1 + x_2 + x_3 \leq 5,$$

$$x_1, x_2, x_3 \in \mathbb{Z}_+.$$

So sánh với nghiệm tối ưu nhận được bằng cách làm tròn từ LP nới lỏng.

6.14 Giải bằng thuật toán hỗn hợp

(a) quy hoạch ở Bài 6.13 nhưng chỉ x_1, x_3 là biến nguyên;

$$\begin{aligned} \text{(b)} \quad & \max z = 3x_1 + x_2 + 3x_3, \\ & -x_1 + 2x_2 + x_3 \leq 4, \\ & 4x_2 - 3x_3 \leq 2, \\ & x_1 - 3x_2 + 2x_3 \leq 3, \\ & x_1, x_2, x_3 \geq 0, x_1, x_3 \text{ nguyên.} \end{aligned}$$

6.15 Chứng minh rằng nếu quy hoạch nguyên tuyến tính có ít nhất một ràng buộc đẳng thức thì nghiệm làm tròn từ LP nới lỏng không thể là nghiệm chấp nhận được của quy hoạch nguyên.

CHƯƠNG 7

LÝ THUYẾT CHỌN QUYẾT ĐỊNH VÀ TRÒ CHƠI

Chúng ta đã xác định là VẬN TRÙ HỌC nghiên cứu cách Vận dụng toán học để trừu tượng mọi việc, thực chất tức là để chọn quyết định hành động trong mọi việc, mọi tình huống khác nhau. Vì vậy, có thể nói vận trù học là lý thuyết chọn quyết định. Chúng ta đã thấy nội dung đó bao trùm suốt sáu chương vừa qua, mặc dù thuật ngữ quyết định gần như chưa được dùng, thay vào đó là phương án và nghiệm... Nhận xét rằng, ngoại trừ Mục 4.5 về điều hành dự án có tính ngẫu nhiên bằng phương pháp PERT, cả sáu chương đều nói đến cách chọn quyết định hành động trong điều kiện có thông tin hoàn hảo, tức là có dữ liệu chắc chắn để tính toán. Nói khác là *chọn quyết định trong hoàn cảnh tất định* (decision making under certainty). Trong thực tế thường thiếu thông tin đầy đủ. Tiêu đề "Lý thuyết chọn quyết định" thường dành cho việc nghiên cứu về hành động trong điều kiện thiếu thông tin này. Có thể chia làm hai trường hợp.

(a) Không có dữ liệu chắc chắn nhưng có phân bố xác suất của các dữ liệu. Lúc này ta gọi việc chọn quyết định là *chọn quyết định có mạo hiểm* (decision making under risk).

(b) Chọn quyết định bấp bênh hoàn toàn được gọi là *chọn quyết định bấp bênh* (decision making under uncertainty).

7.1 CHỌN QUYẾT ĐỊNH CÓ MẠO HIỂM

THÍ DỤ 7.1.1. Công ti Goferbroke có một vùng đất mà theo cố vấn địa chất là có dầu với xác suất $\frac{1}{4}$ và nếu có thì việc khai thác sẽ mang lợi 700000 USD, nếu khoan khai thác mà không có thì sẽ mất 100000 USD. Nếu bán vùng đất thì sẽ được 90000 USD. Công ti cần quyết định thế nào?

Rõ ràng đây là một bài toán chọn quyết định có mạo hiểm. Có hai cách tiếp cận, một là *chọn quyết định không thử nghiệm* (decision making without experimentation), tức là sử dụng ngay xác suất đã biết, gọi là *xác suất tiên nghiệm* (prior probabilities), để tính toán. Hai là thử nghiệm lại, tức là thăm dò địa chất thêm để cải thiện thông tin về xác suất, để có *xác suất hậu nghiệm* (posterior probabilities). Cách tiếp cận này là *chọn quyết định có thử nghiệm* (decision making with experimentation).

7.1.1 Chọn quyết định không thử nghiệm

Bài toán của Công ti Goferbroke được biểu diễn bởi *bảng giải thưởng* (payoff table) như sau

trạng thái tự nhiên quyết định	θ_1 (có dầu)	θ_2 (không)
a_1 : khoan dầu	700	- 100
a_2 : bán đất	90	90
xác suất tiên nghiệm	0,25	0,75

Ở đây "giải thưởng" cụ thể là lợi nhuận thu được (nghìn USD). Có các cách tiếp cận khác nhau, dẫn đến các quy tắc giải khác nhau như sau.

1. Quy tắc hợp lý nhất (Maximum likelihood criterion): Chọn quyết định có giải thưởng cao nhất ở trạng thái tự nhiên có hợp lý nhất (cho nên cũng gọi là quy tắc xác suất cực đại).

Trong thí dụ trên thì $P(\theta = \theta_2) \approx 0,75$ cực đại, ở đây θ là ký hiệu trạng thái tự nhiên, là biến ngẫu nhiên nhận giá trị θ_1 hoặc θ_2 . Trong cột θ_2 thì quyết định a_2 có phần thưởng 90 là cao nhất. Vậy lời giải là công ti cần bán đất.

2. **Quy tắc Bayes** (Bayes' decision rule). Tính kì vọng (tức là trung bình) của các giải thưởng cho mỗi quyết định rồi chọn quyết định có kì vọng của giải thưởng lớn nhất.

Trong bài toán đang xét thì

$$E(a_1) = 0,25 \cdot 700 + 0,75 \cdot (-100) = 100,$$

$$E(a_2) = 0,25 \cdot 90 + 0,75 \cdot 90 = 90.$$

Vậy công ty cần quyết định khoan dầu, trái với lời giải theo quy tắc hợp lí nhất.

Quy tắc Bayes còn được gọi là *quy tắc kì vọng* (expected value criterion). Nó hay được dùng vì có ưu việt là kết hợp được mọi thông tin. Quy tắc hợp lí nhất không tính gì đến phần thưởng ở các cột khác ngoài cột trạng thái tự nhiên có xác suất lớn nhất.

3. **Quy tắc kì vọng - phương sai** (Expected value - variance criterion). Ta đã biết là đối với một biến ngẫu nhiên thì hai đặc trưng định lượng quan trọng nhất là kì vọng, tức là giá trị trung bình theo xác suất, và *phương sai* (variance), đặc trưng cho mức độ lệch của biến ngẫu nhiên so với kì vọng, được định nghĩa cho biến ngẫu nhiên x là

$$\sigma^2 = \text{Var } x := E(x - E(x))^2.$$

$\sigma = \sqrt{\text{Var } x}$ được gọi là *độ lệch chuẩn* (standard deviation). Quy tắc Bayes chưa tính đến phương sai, chỉ dựa vào kì vọng. Nhưng kì vọng, tức là giá trị trung bình, sẽ đại diện cho biến ngẫu nhiên - càng ít chính xác nếu phương sai càng lớn. Vì vậy, thay cho chọn biến x_0 có $\max_x E(x)$ như quy tắc Bayes, người ta chọn x_0 có

$$\max_x [E(x) - K\sqrt{\text{Var } x}],$$

ở đây K là hằng đánh giá vai trò của $\text{Var } x$ trong việc xem xét. Người chọn quyết định ít liều lĩnh sẽ chọn K lớn. Vì vậy nhiều người gọi K là *hệ số chống mạo hiểm* (risk aversion factor). Quy tắc này gọi là quy tắc kì vọng - phương sai.

Xét Thí dụ 7.1.1 theo Quy tắc kì vọng - phương sai ta thấy với $K = 1$ thì

$$E(a_1) - K\sqrt{\text{Var } a_1} = 100 - \sqrt{\text{Var } a_1} = 100 - 2\sqrt{3} \cdot 100 \approx -240,$$

$$(\text{Var } a_1 = 0,25 \cdot 600^2 + 0,75 \cdot 200^2 = 120000),$$

$$E(a_2) - K\sqrt{\text{Var } a_2} = 90 - \sqrt{\text{Var } a_2} = 90 - 0 = 90,$$

$$(\text{var } a_2 = 0,25 \cdot 0 + 0,75 \cdot 0 = 0).$$

Vậy công ty cần quyết định bán đất, lại trái với quyết định dựa theo Quy tắc Bayes.

4. Quy tắc đặt mức đòi hỏi (aspiration – level criterion). Quy tắc này không dựa vào tối ưu hóa giải thưởng theo nghĩa xác suất nào đó mà đặt ra một cách hành động chấp nhận được.

THÍ DỤ 7.1.2. Một người rao bán lại xe ô tô. Anh có biết giá thị trường cho xe tương tự. Như vậy thực chất là anh có "cảm thụ" phân phối xác suất của giá xe cũ, nhưng không đủ để xác định chính xác hàm mật độ xác suất. Vì vậy ta không thể áp dụng các quy tắc trên đây. Một cách thích hợp là chủ xe dựa vào cảm thụ trên để đặt *mức đòi hỏi* (aspiration level) và sẽ bán xe cho người đầu tiên trả giá không thấp hơn mức này.

Trong nhiều trường hợp, mặc dù có phân bố xác suất, nhưng quy tắc đặt mức đòi hỏi vẫn được áp dụng phù hợp hơn.

THÍ DỤ 7.1.3. Biết x là nhu cầu một loại hàng trong một tháng, là một biến ngẫu nhiên có mật độ xác suất $f(x)$. Nếu hàng chuẩn bị lúc đầu tháng ít thì thiếu, nếu nhiều quá thì tồn đầu tư quá (để sản xuất, lưu kho...). Nhưng khó định lượng được phí tổn do thiếu hàng. Để cân bằng người ta đặt bài toán xác định lượng hàng đầu tháng I sao cho kì vọng của lượng hàng thiếu không quá A và kì vọng

của lượng hàng thừa không quá B, tức là

$$\int_1^{\infty} (x - 1)f(x) dx \leq A, \quad (7.1)$$

$$\int_0^1 (1 - x)f(x) dx \leq B. \quad (7.2)$$

Tuy vậy, với một cặp (A,B) đã chọn có thể không có I chấp nhận được, tức là thỏa cả hai điều kiện. Chẳng hạn cho $f(x)$ bởi

$$f(x) = \begin{cases} \frac{20}{x^2} & \text{nếu } 10 \leq x \leq 20, \\ 0 & \text{nếu khác.} \end{cases}$$

Khi đó (7.1) và (7.2) trở thành

$$\int_1^{20} (x - 1)f(x) dx \equiv 20 \left(\ln \frac{20}{1} + \frac{1}{20} - 1 \right) \leq A,$$

$$\int_{10}^1 (1 - x)f(x) dx \equiv 20 \left(\ln \frac{10}{1} + \frac{1}{10} - 1 \right) \leq B.$$

Đơn giản đi ta được

$$\ln I - \frac{1}{20} \geq \ln 20 - \frac{A}{20} - 1 = 1,996 - \frac{A}{20}, \quad (7.3)$$

$$\ln I - \frac{1}{10} \geq \ln 10 - \frac{B}{20} - 1 = 1,302 - \frac{B}{20}. \quad (7.4)$$

Nếu chọn $A = 2$, $B = 4$ thì vế phải (7.3) và (7.4) là 1,896 và 1,102 nên các điều kiện này trở thành

$$\ln I - \frac{1}{20} \geq 1,896, \quad (7.5)$$

$$\ln I - \frac{1}{10} \geq 1,102. \quad (7.6)$$

Ta lập bảng tính về trái tương ứng với các I nguyên trong đoạn $10 \leq x \leq 20$ để xác định xấp xỉ I như sau.

I	10	11	12	13	14	15	16	17	18	19	20
$\ln I - \frac{I}{20}$	1,8	1,84	1,88	<u>1,91</u>	1,94	1,96	1,97	1,98	1,99	1,99	1,99
$\ln I - \frac{I}{10}$	1,3	1,29	1,28	1,26	1,24	1,21	1,17	<u>1,13</u>	1,09	1,04	0,99

Qua bảng này ta thấy với $13 \leq I \leq 17$ thì cả hai điều kiện về mức đòi hỏi (7.5) và (7.6) đều thỏa.

7.1.2 Chọn quyết định có thử nghiệm

THÍ DỤ 7.1.4. Giả sử với công ti dầu khí Goferbroke ở Thí dụ 7.1.1 có thêm một khả năng lựa chọn nữa là thăm dò địa chất thêm với phí tổn 30000 USD để có *xác suất hậu nghiệm* chính xác hơn. Cụ thể như sau. Gọi biến ngẫu nhiên S là kết quả thăm dò địa chất: $S = 0$ gần như là không có dầu, $S = 1$, gần chắc là có dầu. Theo đúc kết kinh nghiệm của quá khứ ta biết các xác suất có điều kiện sau đây.

Nếu thực sự là có dầu, tức là $\theta = \theta_1$, thì biết hai xác suất có điều kiện là $P(S = 0 \mid \theta = \theta_1) = 0,4$ và $P(S = 1 \mid \theta = \theta_1) = 0,6$.

Nếu thực sự là không có dầu, tức là $\theta = \theta_2$, thì biết hai xác suất có điều kiện là $P(S = 0 \mid \theta = \theta_2) = 0,8$ và $P(S = 1 \mid \theta = \theta_2) = 0,2$.

Để tính các xác suất hậu nghiệm khi đã biết kết quả thăm dò $S = 0$ hoặc 1 cụ thể, tức là các xác suất có điều kiện $S = 0$ hoặc $S = 1$ của biến có $\theta = \theta_1$ và $\theta = \theta_2$, ta nhắc lại các công thức xác suất có điều kiện.

Giả sử biến ngẫu nhiên θ có n giá trị có thể θ_i , $i = 1, \dots, n$, và s là một giá trị có thể của biến ngẫu nhiên S . Định nghĩa của *xác suất có điều kiện* (conditional probability) là

$$P(\theta = \theta_i \mid S = s) = \frac{P(\theta = \theta_i, S = s)}{P(S = s)}, \quad (7.7)$$

Ở đây $P(\theta = \theta_i, S = s) := P((\theta = \theta_i) \cap (S = s))$, tức là xác suất để hai biến cố $\theta = \theta_i$ và $S = s$ cùng xảy ra. Vì n biến cố $\theta = \theta_i, i = 1, \dots, n$, *tách biệt* (exclusive events) và là hệ đầy đủ, tức là hợp n biến cố này là *biến cố chắc chắn* (certain event) nên rõ ràng

$$P(S = s) = \sum_{i=1}^n P(S = s, \theta = \theta_i). \quad (7.8)$$

Mặt khác, cũng theo định nghĩa (7.7)

$$P(S = s \mid \theta = \theta_i) = \frac{P(S = s, \theta = \theta_i)}{P(\theta = \theta_i)}.$$

Do đó

$$P(S = s, \theta = \theta_i) = P(S = s \mid \theta = \theta_i) P(\theta = \theta_i). \quad (7.9)$$

Từ (7.7), (7.8) và (7.9) ta được

$$P(\theta = \theta_i \mid S = s) = \frac{P(S = s \mid \theta = \theta_i) P(\theta = \theta_i)}{\sum_{j=1}^n P(S = s \mid \theta = \theta_j) P(\theta = \theta_j)}. \quad (7.10)$$

Áp dụng vào thí dụ của ta (7.10) dẫn đến

$$P(\theta = \theta_1 \mid S = 0) = \frac{0,4 \cdot 0,25}{0,4 \cdot 0,25 + 0,8 \cdot 0,75} = \frac{1}{7},$$

$$P(\theta = \theta_2 \mid S = 0) = 1 - \frac{1}{7} = \frac{6}{7},$$

$$P(\theta = \theta_1 \mid S = 1) = \frac{0,6 \cdot 0,25}{0,6 \cdot 0,25 + 0,2 \cdot 0,75} = \frac{1}{2},$$

$$P(\theta = \theta_2 \mid S = 1) = 1 - \frac{1}{2} = \frac{1}{2},$$

Đây là các xác suất hậu nghiệm. Tùy theo kết quả thử nghiệm là $S = 0$ hay $S = 1$ ta được hai trong bốn xác suất này thay thế cho các xác suất tiên nghiệm để áp dụng bốn quy tắc chọn quyết định có mạo hiểm trên đây. Chẳng hạn, áp dụng Quy tắc Bayes ta tính như sau.

(a) Trường hợp thăm dò được $S = 0$, kì vọng của phần thưởng khi chọn quyết định a_1, a_2 tương ứng, sau khi trừ phí tổn thăm dò là

$$E_0(a_1) = \frac{1}{7} \cdot 700 + \frac{6}{7}(-100) - 30 = -\frac{110}{7},$$

$$E_0(a_2) = \frac{1}{7} \cdot 90 + \frac{6}{7} \cdot 90 - 30 = 60.$$

(b) Trường hợp thăm dò thấy $S = 1$, kì vọng phần thưởng là

$$E_1(a_1) = \frac{1}{2} \cdot 700 + \frac{1}{2}(-100) - 30 = 270.$$

$$E_1(a_2) = \frac{1}{2} \cdot 90 + \frac{1}{2} \cdot 90 - 30 = 60.$$

Vậy quyết định tối ưu là nếu kết quả thăm dò lại là không có dầu thì bán đất (được 60 ngàn USD), nếu kết quả thăm dò là có dầu thì khoan dầu với kì vọng thu được 270 ngàn USD.

Tuy vậy, câu hỏi đặt ra là có đáng thử nghiệm lại hay không vẫn còn ngò. Nghiên cứu về *giá trị của thử nghiệm* (value of experimentation) đề cập đến vấn đề này, xem chẳng hạn [Hillier - Lieberman 2001], Mục 20.3.

7.2 CÂY QUYẾT ĐỊNH

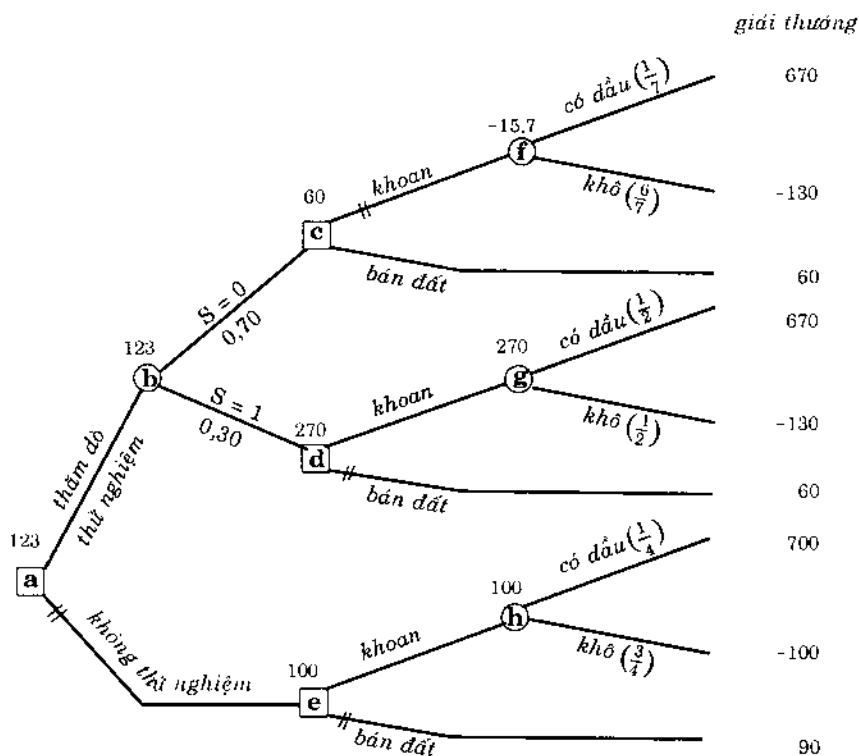
Một cách hiển thị các tính toán khi chọn quyết định có mạo hiểm là biểu diễn *cây quyết định* (decision tree). Ở đây các nút (node) được gọi là các *chạc (cây)* (fork), còn các cung (arc) thì có tên riêng là *cành* (branch). *Chạc quyết định* (decision fork), dạng ô vuông trên H.7.1, chỉ ra quyết định cần chọn ở điểm này của quá trình. *Chạc tình huống* (chance fork), biểu diễn bằng hình tròn, chỉ ra biến cố ngẫu nhiên xảy ra ở thời điểm này.

Xác suất của mỗi biến cố ngẫu nhiên (đi từ chạc tình huống) được

ghi trong ngoặc () bên cạnh cành tương ứng. Chú ý ở các cành từ chạc f và g ta ghi xác suất hậu nghiệm, còn ở cành từ chạc h, không có xác suất hậu nghiệm, ta ghi xác suất tiên nghiệm.

Ở H.7.1 là cây quyết định theo quy tắc Bayes. Khi đó tất cả các số còn lại trên hình đều là các kì vọng của giải thưởng tại điểm tương ứng của quá trình. Tính giải thưởng dựa vào các xác suất và tính từ phải sang:

- tại chạc f: $\frac{1}{7} \cdot 670 + \frac{6}{7}(-130) = -15,7$,
- tại chạc g: $\frac{1}{2} \cdot 670 + \frac{1}{2}(-130) = 270$,



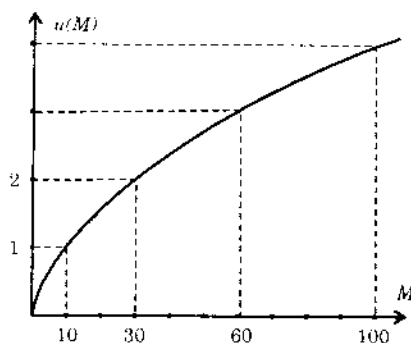
H. 7.1

- tại chạc h: $\frac{1}{4} \cdot 700 + \frac{3}{4}(-100) = 100$,
- tại chạc c, d, e, a ghi các giải thưởng lớn hơn trong hai quyết định đi ra từ chạc đó,
- tại chạc b: $P(S=0) \cdot 60 + P(S=1) \cdot 270 = 0,70 \cdot 60 + 0,30 \cdot 270 = 123$.

Các quyết định có giải thưởng ít hơn ở các chạc quyết định có ghi dấu hai gạch xóa đi. Bây giờ có thể đi theo đường mở từ trái sang để thấy quyết định tối ưu cho cả quá trình: có thăm dò thử nghiệm; nếu kết quả thăm dò là không có dầu thì bán đất, nếu là có dầu thì khoan; kì vọng của giải thưởng là 123.

7.3 LÝ THUYẾT VỀ CÔNG DỤNG

Trong cả hai mục trên ta dùng giải thưởng bằng tiền để tính kì vọng khi chọn quyết định. Trong nhiều tình huống cách tiếp cận này có thể không thích hợp. Chẳng hạn rất nhiều người mua bảo hiểm mặc dù kì vọng của tiền đạt được là rất nhỏ. Vì vậy, tổng quát hơn ta cần có một thang thích hợp đánh giá sự ưa thích hơn của người



H. 7.2

chọn quyết định, là hàm số của giải thưởng M nhưng không nhất thiết là tỉ lệ tuyến tính. Thang này được gọi là *hàm công dụng* hoặc cụ thể hơn *hàm công dụng thay tiền* (utility function for money) $u(M)$.

Một trường hợp điển hình của hàm công dụng cho ở H.7.2. Ở đây ta thấy $u(100) = 4$ và $u(10) = 1$, tức là công dụng của 100 (ngàn USD) chỉ gấp 4 mà không gấp 10 lần công dụng của 10. Ở hàm

$u(M)$ này ta thấy hệ số góc giảm dần khi M tăng. Nếu người chọn quyết định thích mạo hiểm hơn có thể chọn hàm công dụng có hệ số góc tăng.

Ta xét cách xây dựng hàm công dụng khi dùng quy tắc Bayes vì quy tắc này dùng phổ biến nhất. Việc xây dựng hàm công dụng là tùy theo thang ưa thích của người chọn quyết định, nhưng phải tuân theo tính chất sau.

TÍNH CHẤT CƠ BẢN CỦA HÀM CÔNG DỤNG: *Người chọn quyết định không cần phân biệt hai cách hành động có cùng kì vọng của (hàm) công dụng.*

Chẳng hạn, mỗi cặp hành động sau đây đều là như nhau đối với người chọn quyết định có hàm công dụng ở H.7.2.

- (a) "Được 100 (ngàn USD) với xác suất $p = 0,25$ (và không được gì với xác suất 0,75)" là được ưa chọn cũng như "chắc chắn được 10", vì cách hành động đầu có kì vọng $E(\text{công dụng}) = 4 \cdot p = 1$, còn cách hành động sau có công dụng là $u(10) = 1$.
- (b) "Được 100 với xác suất $p = 0,5$ " là không phân biệt với "chắc chắn được 30" ($E(\text{công dụng}) = 4 \cdot 0,5 = 2$, công dụng là $u(30) = 2$).
- (c) "Được 100 với xác suất $p = 0,75$ " là được ưa chọn hoàn toàn giống "chắc chắn được 60" ($E(\text{công dụng}) = 4 \cdot 0,75 = 3$ và công dụng là $u(60) = 3$).

Với hàm công dụng $u(M)$ thay cho giải thưởng tiền là M ta áp dụng quy tắc Bayes hoàn toàn như trước. Chú ý rằng hàm công dụng cũng có thể được xây dựng không phải thay tiền mà thay cho cả hiệu quả khác. Tất nhiên các hiệu quả khó định lượng hoặc thậm chí khó đánh giá thì khó xây dựng được hàm công dụng.

THÍ DỤ 7.3.1. Lại xét bài toán công ti Goferbroke. Giả sử công ti thấy cần có đánh giá hiệu quả khác với tiền thu được là tiền thưởng, tức là phải xây dựng hàm công dụng thay tiền. Trong trường hợp, rời rạc ta phải quy định công dụng cho mọi giải thưởng có thể, theo đánh giá hiệu quả của người chọn quyết định,

dựa vào tính chất cơ bản của hàm công dụng. Cụ thể anh ta phải quy định một loạt xác suất như sau.

Giả sử có hai lựa chọn: "không làm gì", tức là giải thưởng và công dụng đều bằng 0, và "được giải thưởng 700 với xác suất p và được -130 với xác suất $1 - p$ ". Với p nào thì anh coi hai lựa chọn là như nhau? Giả sử theo đánh giá hiệu quả, và theo "mức độ liều" của mình, người quyết định lấy $p = \frac{1}{5}$. Khi đó, kì vọng của công dụng của hai lựa chọn phải bằng nhau:

$$0 = \frac{1}{5}u(700) + \frac{4}{5}u(-130).$$

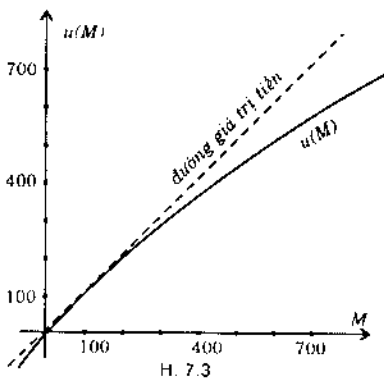
Vì có một phương trình để xác định hai giá trị hàm u ta được quyền lấy tự do một, chẳng hạn $u(-130) = -150$. Từ phương trình suy ra $u(700) = 600$.

Tiếp theo ta dùng hai giá trị của hàm công dụng đã biết để xây dựng tiếp. Để xác định $u(-100)$ ta dùng giá trị gần đó $u(-130)$ và đặt câu hỏi chẳng hạn: giải thưởng -130 với xác suất p và chắc chắn được giải thưởng -100 sẽ được coi là như nhau với xác suất p nào? Giả sử người lập quyết định cho biết $p = 0,7$. Khi đó

$$1. u(-100) = pu(-130) = 0,7(-150) = -105.$$

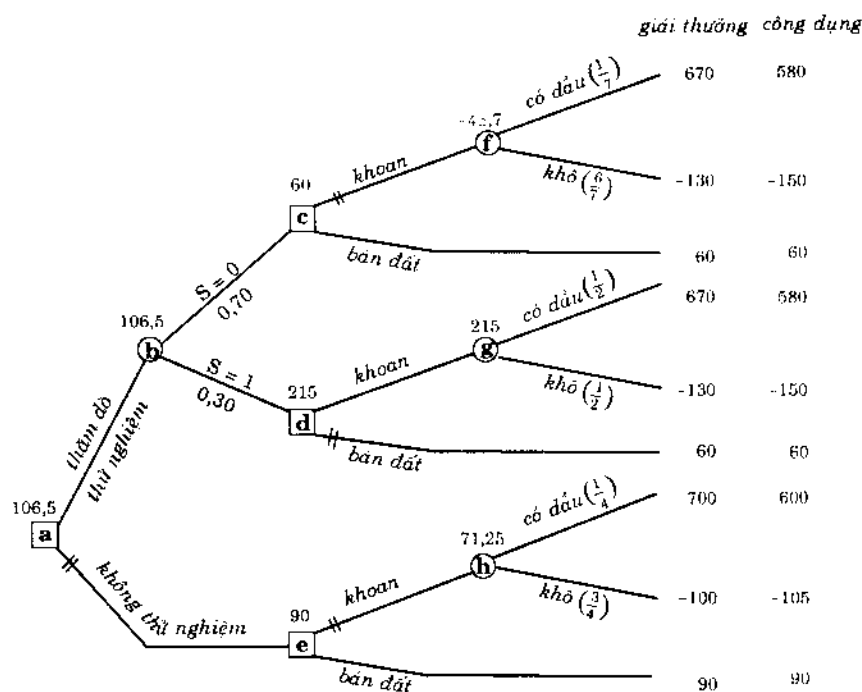
Tiếp theo, anh ta cho biết: nhận 700 với xác suất $p = 0,15$ cũng không khác với nhận 90 chắc chắn. Từ đó ta có

$$u(90) = 0,15, u(700) = 0,15.600 = 90.$$



Nếu đã được một số giá trị của u ta có thể vẽ xấp xỉ đồ thị $u(M)$ để xác định các giá trị cần có khác. Chẳng hạn từ đồ thị ở H.7.3 ta có $u(60) = 60$, $u(670) = 580$. Ta thấy với M càng lớn $u(M)$ càng cách xa đường $u = M$, là đường chéo gọi là *đường giá trị tiền* (monetary value line).

Khi biểu diễn cây quyết định ta ghi thêm cột công dụng bên phải cùng, xem H.7.4. Tất nhiên các chạc cây và cành vẫn như ở H.7.1. Nhưng bây giờ ta dùng cột công dụng để tính các kì vọng công dụng ghi vào các chạc cây. Dựa vào các kì vọng công dụng này các quyết định ở các chạc a, c và d vẫn như khi dùng giải thưởng để xét (H.7.1), nhưng ở chạc c bây giờ là "bán" thay cho "khoan" trước đây. Tuy vậy lần ngược từ cuối về đầu đến chạc a, thì thấy quyết định tối ưu cho cả quá trình vẫn như cũ (ở cuối Mục 7.2). Sở dĩ cây quyết định không khác trước mấy vì hàm công dụng không khác đường giá trị tiền nhiều như H.7.3 chỉ rõ. Nếu người chọn quyết định muốn "ăn chắc", tức là hạn chế mạo hiểm tối đa thì sẽ xây dựng hàm công dụng khác và có thể dẫn tới quyết định bán đất ngay từ đầu.



H. 7.4

7.4 CHỌN QUYẾT ĐỊNH BẮP BÊNH

Có bốn quy tắc chính sau đây, thể hiện mức độ "ăn chắc" hay "tiêu" khác nhau. Chẳng hạn, quy tắc Laplace dựa vào tính lạc quan hơn, tức là liên hơn quy tắc minimax. Về thông tin thì ta cũng biết phần thưởng, $v(a_i, \theta_j)$ ứng với cách hành động a_i và trạng thái tự nhiên θ_j , chỉ thiếu không biết xác suất xảy ra $\theta = \theta_j$, so với điều kiện chọn quyết định có mạo hiểm.

7.4.1 Quy tắc Laplace (Laplace criterion)

Quy tắc này dựa vào *nguyên lý không đủ lý do* (principle of insufficient reason). Vì không đủ lý do để lý luận về xác suất $P(\theta = \theta_j)$ phải khác nhau thế nào đó, nên ta coi là các trạng thái θ_j là đồng khả năng, tức là $P(\theta = \theta_j) = \frac{1}{n}, j = 1, \dots, n$. Đến đây bài toán lập quyết định được đưa về trường hợp có mạo hiểm. Ta áp dụng quy tắc Bayes và chọn quyết định lấy hành động a_j^* đạt được cực đại

$$\max_{a_i} \frac{1}{n} \sum_{j=1}^n v(a_i, \theta_j).$$

THÍ DỤ 7.4.1. Cần quyết định mức lắp đặt một thiết bị giải trí vào đầu ngày lễ. Trạng thái tự nhiên được coi là có bốn mức độ khách đến giải trí 200, 250, 300 hoặc 350. Có bốn mức thiết bị a_i (quyết định hoạt động) được định sẵn phù hợp lý tưởng với bốn mức độ khách đến. Theo kinh nghiệm, nếu xảy ra từng mức khách sẽ có thực tế trạng thái là $\theta_j, j = 1, \dots, 4$ (khác với bốn số chính xác

Mức khách Mức thiết bị	θ_1	θ_2	θ_3	θ_4	$\max_{\theta_j} v(a_i, \theta_j)$
a_1	5	10	18	25	25
a_2	8	7	8	23	23
a_3	21	18	12	21	21
a_4	30	22	19	15	30

trên) dẫn tới phí tổn ứng với từng quyết định a_i theo bảng bên.

Theo Quy tắc Laplace, coi $p(\theta = \theta_j)$

← minimax

$= \frac{1}{4}$ và kì vọng phí tổn cho bốn mức thiết bị là

$$E(a_1) = \frac{1}{4}(5 + 10 + 18 + 25) = 14,5,$$

$$E(a_2) = \frac{1}{4}(8 + 7 + 8 + 23) = 11,5, \leftarrow \text{quyết định}$$

$$E(a_3) = \frac{1}{4}(21 + 18 + 12 + 21) = 18,0,$$

$$E(a_4) = \frac{1}{4}(30 + 22 + 19 + 15) = 21,5.$$

Vậy cần chọn quyết định hành động a_2 .

7.4.2 Quy tắc minimax (minimax criterion)

Đây là quy tắc chắc ăn, cần trọng nhất, dựa trên cách hành động chọn phần thưởng tốt nhất có thể trong khả năng xấu nhất. Chẳng hạn với Thí dụ 7.4.1 thì "phần thưởng" là phí tổn, nên quy tắc này chọn quyết định hành động a_i^* đạt được min max sau đây

$$\min_{a_i} \max_{\theta_j} v(a_i, \theta_j) = 21,$$

theo bảng trên đây, tức là chọn quyết định hành động a_3 , khác với kết quả theo Quy tắc Laplace. Tất nhiên, nếu phần thưởng là "được" thực sự, thì phải chọn a_1^* thỏa

$$\max_{a_i} \min_{\theta_j} v(a_i, \theta_j).$$

7.4.3 Quy tắc minimax - hối tiếc (minimax regret criterion)

Quy tắc minimax dựa vào suy tính quá cần trọng nên dẫn đến hành động thiếu hợp lý trong một số trường hợp.

Chẳng hạn, trong trường hợp ma trận phí tổn đầu trang bên, theo quy tắc minimax, phải chọn quyết định hành động a_2 (với $\min_{a_i} \max_{\theta_j} v(a_i, \theta_j) = 10000$).

	θ_1	θ_2
a_1	11000	90
a_2	10000	10000

$v(a_i, \theta_j)$

Nhưng theo hợp lý trực quan thì ai cũng chọn a_1 để có thể chỉ mất phí tổn 90. Quy tắc minimax-hối tiếc, cũng còn gọi là *quy tắc khe khát*

(savage criterion), dùng ma trận khác gọi là *ma trận hối tiếc* (regret matrix) với các phần tử là $r(a_i, \theta_j)$ thay cho ma trận giải thưởng, ở đây

$$r(a_i, \theta_j) = \begin{cases} v(a_i, \theta_j) - \min_{a_k} v(a_k, \theta_j) & \text{nếu } v \text{ là phí tổn,} \\ \max_{a_k} v(a_k, \theta_j) - v(a_i, \theta_j) & \text{nếu } v \text{ là lợi nhuận.} \end{cases}$$

Về ý nghĩa, $r(a_i, \theta_j)$ là hiệu giữa ô tốt nhất trong cột θ_j và $v(a_i, \theta_j)$, tức là $r(a_i, \theta_j)$ biểu thị sự mất, sự hối tiếc nếu chọn $v(a_i, \theta_j)$ mà bỏ mất lựa chọn quyết định tốt nhất ứng với trạng thái tự nhiên θ_j .

Ở thí dụ với ma trận 2×2 trên đây thì ma trận hối tiếc là như dưới đây. Áp dụng Quy tắc minimax lên bảng ma trận này ta thấy phải chọn a_1 với

$$\min_{a_i} \max_{\theta_j} r(a_i, \theta_j) = 1000.$$

	θ_1	θ_2
a_1	1000	0
a_2	0	9910

$r(a_i, \theta_j)$

Áp dụng vào Thí dụ 7.4.1 ta được bảng ma trận hối tiếc và lời giải như sau, cho thấy phải chọn quyết định a_2 , khác với quy tắc minimax ở Mục 7.4.2.

	θ_1	θ_2	θ_3	θ_4	$\max_{\theta_j} r(a_i, \theta_j)$
a_1	0	3	10	10	10
a_2	3	0	0	8	8
a_3	16	11	4	6	16
a_4	25	15	11	0	25

$\underline{g} \text{ minimax}$

Chú ý rằng dù $v(a_i, \theta_j)$ là phí tổn hay lợi nhuận thì $r(a_i, \theta_j)$ đều

là sự mất nên chỉ có thể áp dụng minimax lên $v(a_i, \theta_j)$, không thể áp dụng maximin.

7.4.4 Quy tắc Hurwicz (Hurwicz criterion)

Để xác định ngữ cảnh ta xét trường hợp $v(a_i, \theta_j)$ là lợi nhuận. Cách chọn quyết định dựa trên sự lạc quan nhất về vận may là chọn a_i^* đạt $\max_{a_i} \max_{\theta_j} v(a_i, \theta_j)$, còn bi quan nhất là chọn a_i^* đạt $\max_{a_i} \min_{\theta_j} v(a_i, \theta_j)$ như quy tắc maximin. Quy tắc Hurwicz cân bằng hai thái cực bằng cách đưa vào tỉ trọng $\alpha \in [0,1]$, tức là chọn a_i^* đạt được

$$\max_{a_i} [\alpha \max_{\theta_j} v(a_i, \theta_j) + (1 - \alpha) \min_{\theta_j} v(a_i, \theta_j)].$$

α được gọi là *chỉ số lạc quan* (index of optimism) được lấy trong $[0,1]$ tùy mức độ "lạc quan" của người lập quyết định. Thường gặp mức trung gian $\alpha = \frac{1}{2}$. Áp dụng Quy tắc Hurwicz với $\alpha = \frac{1}{2}$ cho Thí dụ 7.4.1 ta được bảng sau với lời giải là chọn a_1 hoặc a_2 là như nhau

	$\min_{\theta_j} v$	$\max_{\theta_j} v$	$\frac{1}{2} \min_{\theta_j} v + \frac{1}{2} \max_{\theta_j} v$	
a_1	5	25	15	} $\min_{a_i}$
a_2	7	23	15	
a_3	12	21	16,5	
a_4	15	30	22,5	

Ngoài các quy tắc trên, còn một số quy tắc khác, chẳng hạn quy tắc chọn quyết định hỗn hợp, được gọi là quy tắc *chiến lược hỗn hợp* (mixed strategies), hoàn toàn như đối với lý thuyết trò chơi ở Mục 7.5 dưới đây, có thể xem [Baumol 1977].

7.5 LÝ THUYẾT TRÒ CHƠI

Ở bốn mục trước đã xét trường hợp người chọn quyết định phải đối mặt với sự mạo hiểm hoặc bất định do tự nhiên có các trạng thái mà không biết trước trạng thái nào xảy ra, nhưng tự nhiên không có ý làm tổn hại người chọn quyết định. Thực tế hay gặp tình huống phải chọn quyết định bấp bênh do phải đối mặt với đối thủ "thông minh" và có quyền lợi đối lập với ta, chẳng hạn trong các loại cờ hoặc trò chơi tranh chấp nói chung, trong quân sự hoặc vận động tranh cử. Nghiên cứu về chọn quyết định trong trường hợp đối kháng này có tên riêng là *lý thuyết trò chơi* (game theory). Ở đây cả người chọn quyết định và đối thủ đều gọi là *người chơi* (player). Mỗi người chơi có một tập (hữu hạn hoặc vô hạn, rời rạc hoặc liên tục) cách hành động để lựa chọn gọi là *chiến lược* (strategy). Tương ứng trực tiếp với bài toán chọn quyết định có người chọn quyết định và tự nhiên là *trò chơi hai người* (two - person game). Chúng ta chỉ xét trường hợp đơn giản nhất nhưng cơ bản nhất này. Đồng thời phần thưởng sẽ là cái được của một người và đúng bằng cái mất của người kia. Khi đó ta có *trò chơi tổng zero* (zero - sum game).

Giải một trò chơi có nghĩa là tìm chiến lược tốt nhất cho mỗi người. Hai người chơi thường ký hiệu là A và B, chiến lược của mỗi người, khi ta chỉ xét trò chơi hữu hạn, thường ký hiệu tương ứng là x_i , $i = 1, \dots, m$, và y_j , $j = 1, \dots, n$. Giải thưởng ứng với chiến lược (x_i, y_j) của hai người ký hiệu là a_{ij} và ta cũng viết thành (bảng) ma trận như khi chọn quyết định.

		B			
		y_1	y_2	...	y_n
A	x_1	a_{11}	a_{12}	...	a_{1n}
	x_2	a_{21}	a_{22}	...	a_{2n}
	...	...	...	...	...
	x_m	a_{m1}	a_{m2}	...	a_{mn}

Nghiệm tối ưu, cũng có khi gọi tắt là *nghiệm* của trò chơi, là bộ (x_i^*, y_j^*) có tính chất là nếu một người lấy chiến

lược khác còn người kia giữ nguyên thì phần thưởng cho người đi khác sẽ bị thiệt hơn phần thưởng ứng với (x_1^*, y_j^*) . Giải trò chơi có nghĩa là tìm nghiệm tối ưu.

7.5.1 Giải trò chơi có nghiệm ổn định

1. Chiến lược bị trội hơn

THÍ DỤ 7.5.1. Hai nhà chính trị A và B đang vận động tranh cử một ghế ở nghị viện trong hai ngày cuối quan trọng nhất ở hai thành phố chính P và Q, bằng cách đi ban đêm để dành cả hai ngày ở một thành phố hoặc mỗi ngày một nơi. Phải đặt kế hoạch trước khi không biết kế hoạch của đối phương. Nhưng các cố vấn có thể đánh giá được ảnh hưởng của các cách tranh cử của A và B để dựa vào đó chọn một trong ba: chiến lược 1 là ở mỗi thành phố một ngày,

chiến lược người	người B		
	1	2	3
A	1	2	4
	2	1	0
	3	0	1

chiến lược 2 là cả hai ngày ở P, chiến lược 3 là cả hai ngày ở Q. Công dụng được đánh giá bằng tổng số (đơn vị ngàn) phiếu người A dành được từ B (nếu số âm là B dành được từ A) được cho bằng bảng bên.

Trong trường hợp đặc biệt này ta thấy tình cờ có thể giải đơn giản bằng khái niệm *chiến lược bị trội hơn* (dominated strategy). Ta thấy đối với người A thì chiến lược 3 bị trội hơn bởi chiến lược 1 và 2 vì dù B chọn chiến lược nào thì A cũng chọn 1 hoặc 2 sẽ lợi hơn 3. Vì vậy ta có thể loại bỏ hàng thứ ba. Khi còn hai hàng ta lại thấy đối với người B thì chiến lược 3 bị trội hơn nên lại bỏ cột ba. Khi mà trận còn bốn phần tử ta lại thấy với người A thì chiến lược 2 bị trội hơn, bỏ hàng 2 ta chỉ còn hai phần tử như bảng bên. Đến đây ta lại thấy với người B thì chiến lược 2 bị trội hơn. Cuối cùng bộ chiến lược (1,1) là nghiệm tối ưu của trò chơi với kết quả A thu thêm được 1 (ngàn) phiếu từ B. Số 1 này được gọi là *giá trị của trò chơi* (value of the game). Trò chơi tổng zero mà có giá trị 0 được gọi là *trò chơi đúng* (fair game). Trong nhiều trường hợp dùng chiến lược bị trội có thể giảm được cỡ trò chơi, tuy chưa giải xong như thí dụ trên.

	1	2
1	1	2

2. Chiến lược minimax cho trò chơi ổn định

THÍ DỤ 7.5.2. Xét Thí dụ 7.5.1 nhưng với bảng công dụng được các cố vấn đánh giá khác như dưới đây. Rõ ràng đây là trường hợp chọn quyết định mà đối phương thông minh và cố ý chọn quyết định chống lại nên dùng chiến lược "ăn chắc" nhất là chiến lược minimax đối với người B và maximin đối với người A là hợp lý. Kiểm tra lại ta thấy bộ chiến lược tìm được theo quy tắc này là (2,2) là nghiệm tối ưu của trò chơi. Vì nếu người nào đi trệch mà người kia

đi đúng thì người kia thu lợi hơn giá trị của trò chơi, trong trường hợp này là 0. Đây là trò chơi đúng.

Nguyên nhân bộ chiến lược (2,2) là tối ưu là ở chỗ hai giá trị minimax và maximin này cùng đạt được ở $a_{22} = 0$. Từ đó người nào đi trệch khỏi chiến lược này sẽ chịu thiệt nếu người kia vẫn đi đúng. a_{22} lúc này được gọi là *điểm yên ngựa* (saddle point) của ma trận các giá trị a_{ij} . Nghiệm tối ưu cũng được gọi là *nghiệm ổn định* (stable solution), vì người nào đi trệch sẽ bị thiệt.

7.5.2 Chiến lược hỗn hợp

Ta hãy xét một thí dụ trò chơi không có nghiệm ổn định

chiến lược		người B			
		1	2	3	
người A	1	0	-2	2	-2 ← maximin
	2	5	4	-3	-3
	3	2	3	-4	-4
max		5	4	2	
		↑			minimax

THÍ DỤ 7.5.3. Xét Thí dụ 7.5.1 với bảng công dụng khác như bảng bên cạnh. Ta thấy

$$\begin{aligned}\max_i \min_j a_{ij} &= -2, \\ \min_i \max_j a_{ij} &= 2\end{aligned}$$

nên trò chơi (cũng tức là ma trận của nó) không có điểm yên ngựa. Ta xem hậu quả ra sao. Nếu cả A và B thực hiện đúng chiến lược maximin và minimax thì theo bảng ma trận A sẽ thắng 2 từ B (khác với -2 như A dự tính). B có ma trận trước mắt để suy tính và thấy phải chọn chiến lược 2 để thắng 2 từ A. Đến lượt A, cũng đủ thông minh để tính liền được hai nước cờ để biết B sẽ dùng chiến lược 2, nên sẽ chuyển sang chiến lược 2 để thắng 4. Nhưng B cũng tính được điều này nên quay lại chiến lược 3 để thắng 3. Do đủ thông minh A sẽ quay lại chiến lược 1 để thắng 2. Như vậy ta đã có đủ một vòng và cứ lý luận như vậy sẽ xoay vòng mãi. Bộ chiến lược (1,3) nhận được theo quy tắc minimax ở đây là *ng nghiệm không ổn định* (unstable solution).

Để có lời giải của trò chơi trong trường hợp này, người ta đưa ra khái niệm *chiến lược hỗn hợp* (mixed strategy). Mỗi người chơi không chọn một *chiến lược thuần túy* (pure strategy) như trước đây quy định mà chọn một phân bố xác suất sử dụng tất cả các chiến lược. Ký hiệu

x_i : xác suất người A sử dụng chiến lược i , $i = 1, \dots, m$,

y_j : xác suất người B sử dụng chiến lược j , $j = 1, \dots, n$.

Khi đó cả kế hoạch, hoặc gọi khác là phân bố xác suất ($x_1, \dots, x_m$) và ($y_1, \dots, y_n$) được gọi là chiến lược hỗn hợp của A và B tương ứng, khi trò chơi đang tiến triển, tất nhiên mỗi người dùng một chiến lược thuần túy, được chọn bằng một cách, một "thiết bị" để chọn ngẫu nhiên theo đúng phân bố xác suất.

Quay lại Thí dụ 7.5.3, giả sử A chọn chiến lược hỗn hợp $\left(\frac{1}{2}, \frac{1}{2}, 0\right)$ và B chọn $\left(0, \frac{1}{2}, \frac{1}{2}\right)$. Chẳng hạn với A, chiến lược này có nghĩa là A loại bỏ chiến lược 3 và tung xu để xem chọn chiến lược 1 hay 2 để dùng.

Tìm chiến lược hỗn hợp tối ưu thế nào? Có đúng là dùng chiến lược hỗn hợp để đưa ra lời giải cho trò chơi không ổn định, tức là không có điểm yên ngựa, sẽ có câu trả lời thỏa đáng không? Bây giờ ta xem xét.

Khi đã dùng đến chiến lược hỗn hợp có chứa yếu tố ngẫu nhiên và xác suất, thì giải thưởng cũng là biến ngẫu nhiên, nhận giá trị a_{ij} với xác suất $x_i y_j$. Vậy kì vọng giải thưởng là $\sum_{i=1}^m \sum_{j=1}^n a_{ij} x_i y_j$. Chẳng hạn ở Thí dụ 7.5.3, nếu A chọn chiến lược $\left(\frac{1}{2}, \frac{1}{2}, 0\right)$ và B chọn $\left(0, \frac{1}{2}, \frac{1}{2}\right)$ thì kì vọng giải thưởng (A được và B mất) là $\frac{1}{4}(-2 + 2 + 4 - 3) = \frac{1}{4}$. Với một lần chơi thì A phải lấy ngẫu nhiên với xác suất $\frac{1}{2}$ một chiến lược 1 hoặc 2 cụ thể và B cũng vậy trong chiến lược 2 và 3. Khi đó phần thưởng nói chung là khác $\frac{1}{4}$. Nhưng nếu chơi nhiều lần cùng với chiến lược hỗn hợp $\left(\frac{1}{2}, \frac{1}{2}, 0\right)$ và $\left(0, \frac{1}{2}, \frac{1}{2}\right)$ thì phần thưởng trung bình là $\frac{1}{4}$.

Bây giờ người A nên suy nghĩ thế nào về cách chọn chiến lược hỗn hợp tốt nhất. Vì B là thông minh và hành động với mục đích đối lập nên tốt nhất là A dùng quy tắc maximin đối với kì vọng giải thưởng tức là chọn chiến lược x_i^* ứng với

$$\max_{x_i} \min_{y_j} \sum_{i=1}^m \sum_{j=1}^n a_{ij} x_i y_j := \underline{v}.$$

Còn B cần dùng quy tắc minimax, tức là chọn y_j^* đạt được

$$\min_{y_j} \max_{x_i} \sum_{i=1}^m \sum_{j=1}^n a_{ij} x_i y_j := \bar{v}.$$

May mắn là mặc dù trò chơi là không ổn định khi dùng chiến lược thuần túy, tức là không có điểm yên ngựa cho ma trận giải thưởng:

$$\max_i \min_j a_{ij} < \min_j \max_i a_{ij},$$

nhưng ta lại luôn có $\underline{v} = \bar{v}$, tức là maximin = minimax đối với kì vọng khi dùng chiến lược hỗn hợp như định lý sau khẳng định.

ĐỊNH LÝ 7.5.1 (Định lý minimax). *Nếu dùng chiến lược hỗn hợp, cặp chiến lược tối ưu theo quy tắc minimax là nghiệm ổn định với $\underline{v} = \bar{v} := v$ (gọi là giá trị của trò chơi). Do đó nếu một người đi lệch chiến lược tối ưu thì người đó sẽ nhận giải thưởng kém hơn v .*

(Xem chứng minh cuối Mục 7.5.4.)

Mặc dù trò chơi ban đầu là hữu hạn, khi dùng chiến lược hỗn hợp ta có vô số (ở mức continuum) khả năng nên không tìm chiến lược tối ưu (x_i^*, y_j^*) bằng cách dùng bảng trực tiếp được. Ta sẽ xét ở hai mục sau đây.

7.5.3 Giải hình học trường hợp đơn giản

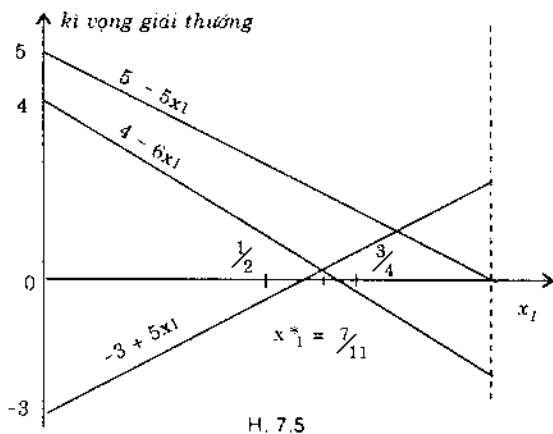
Nếu sau khi gạch bỏ chiến lược bị trội, một trong hai người chơi chỉ còn hai chiến lược thuần túy, thì ta có thể tìm chiến lược hỗn hợp tối ưu bằng hình học như minh họa bởi thí dụ sau.

THÍ DỤ 7.5.4. Xét trò chơi ở Thí dụ 7.5.3, ta thấy chiến lược 3 của người A là bị trội. Bỏ chiến lược này thì A chỉ còn hai chiến lược thuần túy. Mỗi chiến lược hỗn hợp của A có dạng $(x_1, x_2) = (x_1, 1 - x_1)$ vì $x_1 + x_2 = 1$ (độ đo xác suất!). Vì vậy kì vọng giải thưởng là hàm của một biến x_1 nên ta có thể tìm minimax và maximin trên đồ thị. Cụ thể ở thí dụ này ta thấy nếu B dùng chiến lược thuần túy, tức là (y_1, y_2, y_3) bằng $(1, 0, 0)$, $(0, 1, 0)$ hoặc $(0, 0, 1)$ thì kì vọng giải thưởng của A tương ứng là

(y_1, y_2, y_3)	kì vọng giải thưởng A
$(1, 0, 0)$	$0x_1 + 5(1 - x_1) = 5 - 5x_1$
$(0, 1, 0)$	$-2x_1 + 4(1 - x_1) = 4 - 6x_1$
$(0, 0, 1)$	$2x_1 - 3(1 - x_1) = -3 + 5x_1$

Đồ thị của các kì vọng giải thưởng của A là đường thẳng. Nếu B dùng chiến lược hỗn hợp bất kì (y_1, y_2, y_3) thì kì vọng giải thưởng của A sẽ là tổ hợp với tỉ trọng y_1, y_2 và y_3 ba kì vọng trên:

$$y_1(5 - 5x_1) + y_2(4 - 6x_1) + y_3(-3 + 5x_1). \quad (7.11)$$



Như vậy với mỗi x_1 , kì vọng tổ hợp (7.11) là tổ hợp của ba cao độ của ba đường thẳng trên H.7.5, với hệ số $y_j \geq 0$ và $y_1 + y_2 + y_3 = 1$. Chiến lược tối ưu $(x_1^*, 1 - x_1^*)$ của A là ứng với maximin (7.11). Khi tìm min trước (tức là tìm tổ hợp y_j để có (ung độ thấp nhất), thì phải có $y_1 = 0$ vì đường $5 - 5x_1$ luôn cao hơn đường

ứng y_2 là $4 - 6x_1$. Do đó x_1^* là điểm thỏa

$$\max_{0 \leq x_1 \leq 1} \min \{4 - 6x_1, -3 + 5x_1\},$$

tức là $x_1^* = \frac{7}{11}$. Do đó $\underline{v} = 4 - \frac{6 \cdot 7}{11} = \frac{2}{11}$.

Bây giờ tìm chiến lược tối ưu của B là (y_1^*, y_2^*, y_3^*) đạt minimax (7.11). Do đó, với mọi x_1 ,

$$y_1^*(5 - 5x_1) + y_2^*(4 - 6x_1) + y_3^*(-3 + 5x_1) \leq \bar{v}. \quad (7.12)$$

Theo Định lý minimax thì $\bar{v} = \underline{v} = \frac{2}{11}$ và giá trị này phải đạt để (7.12) trở thành đẳng thức khi $x_1 = x_1^* = \frac{7}{11}$, tức là

$$\frac{20}{11}y_1^* + \frac{2}{11}y_2^* + \frac{2}{11}y_3^* = \bar{v} = \frac{2}{11}. \quad (7.13)$$

Vì (y_1, y_2, y_3) là phân bố xác suất nên thay $y_2^* + y_3^* = 1 - y_1^*$ vào (7.13) ta được $\frac{18}{11}y_1^* = 0$. Suy ra $y_1^* = 0$. Thực ra ở đây ta có thể nhận xét tổng quát là mọi đường thẳng j không đi qua điểm maximin $(x_1^*, \underline{v})$ phải ứng với tỉ trọng $y_j^* = 0$. Do đó (7.12) và (7.13) suy ra

$$y_2^*(4 - 6x_1) + y_3^*(-3 + 5x_1) \begin{cases} \leq \frac{2}{11} & \text{với } 0 \leq x_1 \leq 1, \\ = \frac{2}{11} & \text{với } x_1 = \frac{7}{11}. \end{cases}$$

Nhưng y_2^* và y_3^* cố định nên vế trái là một vế của phương trình đường thẳng, phải $\equiv \frac{2}{11}, \forall x \in [0, 1]$. Do đó lấy hai giá trị bất kì của x_1 , chẳng hạn 0 và 1, ta được

$$4y_2^* - 3y_3^* = \frac{2}{11},$$

$$-2y_2^* + 2y_3^* = \frac{2}{11}.$$

Suy ra $y_2^* = \frac{5}{11}, y_3^* = \frac{6}{11}$, và chiến lược tối ưu của B là $\left(0, \frac{5}{11}, \frac{6}{11}\right)$.

7.5.4 Giải trò chơi hai người tổng quát bằng quy hoạch tuyến tính

Xét trò chơi hai người tổng zero với ma trận cấp $m \times n$ tổng quát (định nghĩa ở đầu Mục 7.5). Trước hết tìm chiến lược hỗn hợp tối ưu cho người A. Ta đã biết kì vọng phần thưởng của A là

$$\sum_{i=1}^m \sum_{j=1}^n a_{ij} x_i y_j.$$

Chiến lược $(x_1, \dots, x_m)$ tối ưu phải thỏa điều kiện maximin, nên

$(x_1, \dots, x_m)$ là tối ưu nếu và chỉ nếu

$$\sum_{i=1}^m \sum_{j=1}^n a_{ij} x_i y_j \geq v = \underline{v} \quad (7.14)$$

$\forall (y_1, \dots, y_n)$. Nói riêng với n chiến lược đặc biệt (là chiến lược thuần túy $(y_1 = 1, y_2 = \dots = y_n = 0)$, $\dots$, $(y_1 = \dots = y_{n-1} = 0, y_n = 1)$) (7.14) trở thành hệ n bất đẳng thức

$$\sum_{i=1}^m a_{ij} x_i \geq v, j = 1, \dots, n. \quad (7.15)$$

Ngược lại, từ (7.15) suy ra

$$\sum_{j=1}^n y_j \left(\sum_{i=1}^m a_{ij} x_i \right) \geq \sum_{j=1}^n y_j v = v,$$

vì $\sum_{j=1}^n y_j = 1$. Vậy (7.15) tương đương (7.14). Ngoài ra, nhắc lại là

$(x_1, \dots, x_m)$ tối ưu tất nhiên cũng phải thỏa

$$x_1 + x_2 + \dots + x_m = 1, \quad (7.16)$$

$$x_i \geq 0, i = 1, \dots, m. \quad (7.17)$$

Vậy $(x_1, \dots, x_m)$ là chiến lược hỗn hợp tối ưu của A khi và chỉ khi nó thỏa hệ ràng buộc tuyến tính (7.15), (7.16), (7.17). Để đưa về quy hoạch tuyến tính ta còn phải đưa vào mục tiêu và giải quyết việc v chưa biết. Cả hai việc sẽ được giải quyết khi đặt mục tiêu là $\max x_{m+1}$ và thay v ở (7.15) bởi x_{m+1} . Khi đó rõ ràng x_{m+1} sẽ bằng v tại nghiệm tối ưu. Tóm lại, chiến lược hỗn hợp tối ưu của A là nghiệm tối ưu của quy hoạch tuyến tính

$$\max x_{m+1},$$

$$\sum_{i=1}^m a_{ij} x_i - x_{m+1} \geq 0, j = 1, \dots, n, \quad (7.18)$$

$$\sum_{i=1}^m x_i = 1,$$

$$x_i \geq 0, i = 1, \dots, m.$$

Bây giờ xét người B. Hoặc là viết lại ma trận giải thưởng ở dạng phần được của B và mất của A, tức là đổi dấu, rồi làm đúng như trên, hoặc là lý luận tương tự như trên nhưng từ quan điểm của B trên ma trận ban đầu ta sẽ thấy chiến lược hỗn hợp tối ưu của B là nghiệm tối ưu của quy hoạch tuyến tính đối ngẫu với (7.18)

$$\begin{aligned} & \min y_{n+1}, \\ & \sum_{j=1}^n a_{ij} y_j - y_{n+1} \leq 0, i = 1, \dots, m, \\ & \sum_{j=1}^n y_j = 1, \\ & y_j \geq 0, j = 1, \dots, n. \end{aligned} \quad (7.19)$$

Do đó, theo lý thuyết đối ngẫu của quy hoạch tuyến tính (Chương 2) ta chỉ cần giải một quy hoạch sẽ nhận được cả bộ chiến lược tối ưu.

Nhận xét thêm là ta có thể đưa quy hoạch tuyến tính (7.18) và (7.19) về dạng có cả ràng buộc dấu cho biến còn lại $x_{m+1} \geq 0, y_{n+1} \geq 0$ như sau. Nếu đã biết trước là $v \geq 0$ thì cứ viết thêm các ràng buộc dấu này. Nếu $v < 0$ thì có ba cách giải quyết. Một là thay biến như đã biết x_{m+1} bởi hai biến mới theo quan hệ $x_{m+1} = x_{m+1}^+ - x_{m+1}^-$, $x_{m+1}^+ \geq 0, x_{m+1}^- \geq 0$. Cách hai là viết lại phần thưởng ở dạng ma trận phần thưởng được cho B, tức là đổi dấu, khi đó v cũng đổi dấu. Cách thứ ba, hay dùng nhất, là cộng thêm một số đủ lớn vào mọi phần thưởng để giá trị của trò chơi mới dương.

THÍ DỤ 7.5.5. Xét trò chơi ở Thí dụ 7.5.4. Đưa về quy hoạch tuyến tính

$$\begin{aligned} \max x_3, \\ 5x_2 - x_3 &\geq 0, \\ -2x_1 + 4x_2 - x_3 &\geq 0, \\ 2x_1 - 3x_2 - x_3 &\geq 0, \\ x_1 + x_2 &= 1, \\ x_1 \geq 0, x_2 &\geq 0. \end{aligned} \quad (7.20)$$

Ta cứ bỏ sung ràng buộc $x_3 \geq 0$ và giải theo thuật toán đơn hình sẽ được nghiệm tối ưu $(x_1^*, x_2^*, x_3^*) = \left(\frac{7}{11}, \frac{4}{11}, \frac{2}{11}\right)$ (như vậy việc thêm ràng buộc $x_3 \geq 0$ không ảnh hưởng gì). Vậy chiến lược tối ưu của A là $\left(\frac{7}{11}, \frac{4}{11}\right)$ và giá trị của trò chơi là $v = \frac{2}{11}$. Khi giải bằng thuật toán đơn hình, theo từ vựng tối ưu nhận được ở bước cuối ta cũng nhận được nghiệm tối ưu của quy hoạch tuyến tính đối ngẫu là $(y_1^*, \dots, y_4^*) = \left(0, \frac{5}{11}, \frac{6}{11}, \frac{2}{11}\right)$. Vậy chiến lược tối ưu của B là $\left(0, \frac{5}{11}, \frac{6}{11}\right)$. Tất nhiên chiến lược này cũng có thể nhận được bằng cách giải trực tiếp quy hoạch đối ngẫu của (7.20).

Từ lý luận về đối ngẫu trên đây ta có ngay

CHỨNG MINH ĐỊNH LÝ MINIMAX. Theo Định lý đối ngẫu mạnh của quy hoạch tuyến tính $x_{n+1}^* = y_{n+1}^*$. Nhưng rõ ràng từ định nghĩa của $\underline{v}$ và $\bar{v}$ ta có $\underline{v} = x_{n+1}^*$ và $\bar{v} = y_{n+1}^*$. Vậy $\underline{v} = \bar{v}$. $\square$

7.6 MỘT SỐ NỘI DUNG KHÁC CỦA LÝ THUYẾT TRÒ CHƠI

Lý thuyết trò chơi là một lĩnh vực khá rộng. Ở đây ta mới đề cập đến trò chơi hai người, tổng zero, hữu hạn. Mở rộng trực tiếp là trường hợp *tổng hằng số* (constant-sum), nhưng nói chung khác không, được

xét tương tự. Trường hợp *tổng khác không* (nonzero-sum), nhưng không phải hằng số, phản ánh tình huống, có chứa yếu tố không xung đột, cạnh tranh mà cùng tác động đến cả hai đối thủ. Chẳng hạn, chiến lược quảng cáo của hai công ti cạnh tranh không chỉ ảnh hưởng việc chia thị trường mà cả mở rộng hay thu hẹp thị trường tổng thể.

Tiếp sau trò chơi hai người, rõ ràng trong nhiều tình huống cần xét *trò chơi n người* (n-person game).

Trò chơi chủ yếu được hiểu về tình trạng đối kháng, nhưng cũng phải xét cả trường hợp có yếu tố hợp tác. Đối lại hoàn toàn với trường hợp trò chơi tổng zero là trường hợp *trò chơi hợp tác* (cooperative game). Chẳng hạn hiệp định thương mại giữa hai nước, việc thảo luận giữa chủ và công nhân có thể phát biểu thành trò chơi hợp tác. Trường hợp có nhiều người chơi, có thể phải xét việc một số lập thành *đồng minh* (coalition).

Bên cạnh trò chơi hữu hạn, *trò chơi vô hạn* (infinite game) xét trường hợp có vô hạn chiến lược hành động. Phát triển hơn nữa là trường hợp *biến quyết định liên tục* (continuous decision variable). Chẳng hạn quyết định điều khiển $u(t)$ phụ thuộc thời gian t . Trường hợp liên tục được nghiên cứu nhiều nhất là *trò chơi vi phân* (differential game). Ở đây quá trình chơi phải tuân theo một (hoặc hệ) phương trình vi phân và người chơi chọn điều khiển của mình để quá trình diễn ra làm cực trị hàm mục tiêu theo quyền lợi của mình.

BÀI TẬP CHƯƠNG 7

7.1 Để giải quyết máy hỏng có hai cách; chữa máy hỏng với phí tổn $c_1 = 100$ một máy và bảo trì toàn bộ $n = 50$ máy với cước phí $c_2 = 10$ một máy. Biết xác suất một máy bị hỏng trong chu kì thời gian thứ T là p_T và n_T là biến ngẫu nhiên chỉ số máy hỏng ở chu kì T . Hãy chọn quyết định là chu kì

T	1	2	3	4	5
p_T	0,05	0,07	0,1	0,13	0,18

thứ T^* để bảo trì toàn bộ n máy (vào cuối chu kì) để phí tổn là ít nhất. p_T được cho bởi bảng trang 319. (Ở các chu kì trước T không bảo trì mà sửa máy hỏng.)

Gợi ý: Tính kì vọng phí tổn trong một chu kì EC_T khi chọn chu kì T để bảo trì, chú ý $E(n_T) = np_T$. T^* chọn theo điều kiện $EC_{T-1} \geq EC_T, EC_{T+1} \geq EC_T$.

7.2 Một công ti có bản quyền một loại chip mới. Nếu bán bản quyền thì được 800000 USD. Nếu sản xuất chip để bán máy tính thì có hai trường hợp đồng khả năng: bán lẻ chắc chắn được 1000 máy, tổ chức hệ thống đại lý thì bán được 10000 máy. Phí tổn tổ chức hệ thống là 600000 USD. Lãi bán một máy là 600 USD. Phát biểu thành bài toán chọn quyết định và giải bằng Quy tắc Bayes.

7.3 Giải Bài 7.1 nếu $c_1 = 200$, $c_2 = 15$, $n = 30$ và

$$p_t = \begin{cases} 0,03 & \text{nếu } t = 1, \\ p_{t-1} + 0,1 & \text{nếu } t = 2, \dots, 10, \\ 0,13 & \text{nếu } t = 11, 12, \dots \end{cases}$$

7.4 Các lô hàng có 8, 10, 12 hoặc 14 phần trăm hỏng với xác suất tương ứng 0,4; 0,3; 0,25 và 0,05. Hợp đồng giao hàng cho khách A, B và C có quy định phần trăm hàng hỏng không quá 8, 12 và 14, tương ứng. Nếu hàng hỏng nhiều hơn hợp đồng, công ti bị phạt 100 USD cho 1% hàng hỏng quá mức. Để giảm 1% hàng hỏng trong lô công ti tốn thêm 50 USD khi sản xuất. Nếu các lô hàng không được kiểm trước khi chuyển, thì khách nào phải có ưu tiên nhận hàng nhất.

7.5 Lượng mua bánh mì ở một cửa hàng thực phẩm có xác suất như sau (x là số bánh). Giá một bánh tươi là 4,9 (ngàn đồng). Nếu phải để sang ngày sau thì cước bảo quản là 2,5 và giá giảm 1,5. Hỏi cửa hàng phải đặt mua lượng nào trong bảng dưới.

x	100	150	200	250	300
$p(x)$	0,20	0,25	0,3	0,15	0,10

7.6 Giải Bài 7.1 bằng Quy tắc kì vọng - phương sai với $k = 1$.

Gợi ý: $\text{Var } n_t = np_t(1 - p_t)$.

7.7 Nhu cầu mua hàng có hàm mật độ xác suất như bảng sau.

x	0	1	2	3	4	5
$p(x)$	0,10	0,15	0,40	0,15	0,10	0,10

Xác định mức nhập hàng để xác suất hết hàng bán không quá 0,45. Xác định mức nhập hàng để lượng hàng thiếu và thừa không quá 1 và 2.

7.8 Xét Bài 7.7. Tìm mức nhập hàng để kì vọng lượng hàng thiếu ít hơn kì vọng lượng hàng thừa ít nhất là 1.

7.9 Xét Bài 7.4. Giả sử lấy một lượng hàng $n = 20$ ra kiểm lại trước khi chuyển và phát hiện 4 hỏng. Tính xác suất hậu nghiệm của 4 lô hàng đã nói và dùng nó để xác định khách hàng nào có kì vọng phí tổn ít nhất.

7.10 Xét Bài 7.2. Giả sử có thể nghiên cứu thị trường với phí tổn 400000 USD để xác định mức nào trong hai mức nhu cầu có thể xảy ra. Kinh nghiệm cũ cho thấy nghiên cứu thị trường đúng với xác suất $2/3$.

(a) Tìm xác suất hậu nghiệm của hai mức nhu cầu cho cả hai kết quả nghiên cứu thị trường.

(b) Dùng xác suất hậu nghiệm chọn quyết định.

7.11 Một trường đại học cân nhắc có tổ chức quyền góp trong cựu học sinh tiến để xây một trung tâm thể thao trong trường. Trường xác định rằng nếu đội tuyển bóng đá của trường thắng trong giải vào hè này thì chắc chắn quyền được 300 triệu đồng (trừ phí tổ chức rồi), nếu đội thua thì sẽ lỗ 200 triệu vào tiền tổ chức. Nếu không tổ chức thì không mất gì, được gì.

(a) Trong quá khứ đội tuyển thắng 60% giải. Trường có thể chi đến bao nhiêu để thăm dò thêm khả năng thắng mùa tới?

(b) Giả sử có khả năng thuê chuyên gia mất 10 triệu để đánh giá đúng đến xác suất $\frac{3}{4}$ là đội sẽ thắng hay thua. Xác định xác suất hậu nghiệm cho kết quả thi đấu cho cả hai dự báo của chuyên gia.

(c) Dùng cây quyết định tìm quyết định tối ưu và kì vọng phần thưởng của quyết định tối ưu.

7.12 Nhu cầu mua bánh trong ngày ở một cửa hàng được coi như chỉ có ba trường hợp 100, 120 hoặc 130 với xác suất 0,2 ; 0,3 và 0,5. Giá mua một bánh là 0,6 (USD), bán là 1,05. Nếu để đến hôm sau thì giá còn là 0,55. Dùng cây quyết định để tìm mức nhập bánh tối ưu.

7.13 Dựa vào nhu cầu mua sản phẩm cao hay thấp, một công ti sẽ chọn quyết định lắp dây truyền sản xuất lớn hay nhỏ lúc này và có mở rộng dây truyền vào hai năm sau

322 CHƯƠNG 7 LÝ THUYẾT CHỌN QUYẾT ĐỊNH VÀ TRÒ CHƠI

không, để sản xuất trong tổng cộng 10 năm. Tiếp thị cho thấy trong vòng 10 năm, xác suất nhu cầu giữ nguyên là 0,75 cho nhu cầu cao, 0,25 cho nhu cầu thấp. Lắp dây truyền lớn mất 500 (triệu đồng), nhỏ mất 100 triệu và mở rộng dây truyền mất 420 triệu. Ước lượng doanh thu hàng năm như sau.

(a) Với dây truyền lớn và nhu cầu cao (thấp): 100 triệu (30 triệu);

(b) với dây truyền nhỏ và nhu cầu cao (thấp): 25 triệu (20 triệu);

(c) mở rộng dây truyền nhỏ với nhu cầu cao (thấp): 90 (20).

Dùng cây quyết định tìm quyết định tối ưu và kì vọng tiền lãi tổng cộng 10 năm.

	θ_1	θ_2	θ_3	θ_4	θ_5	7.14 Tìm quyết định cho bài toán có bảng giải thưởng đã cho bằng cả 4 quy tắc: (a) Laplace, (b) minimax, (c) minimax - hối tiếc, (d) Hurwicz với $\alpha = \frac{1}{2}$. (Không biết xác suất xảy ra $\theta = \theta_j$.)
a_1	15	10	0	-6	17	
a_2	3	14	8	9	2	
a_3	1	5	14	20	-3	
a_4	7	19	10	2	0	

7.15 Dùng chiến lược bị trội giải các trò chơi

chiến lược			B			chiến lược			B		
			1	2	3				1	2	3
(a)	A	1	-3	1	2	(b)	A	1	1	2	0
		2	1	2	1			2	2	-3	-2
		3	1	0	-1			3	0	3	-1

7.16 Chứng minh rằng với bất kì ma trận (a_{ij}) thì

$$\max_i \min_j a_{ij} \leq \min_j \max_i a_{ij}.$$

7.17 Các trò chơi (hai người, tổng zero) có ma trận giải thưởng sau đây có điểm yên ngựa không? Tìm $\underline{v}$ và $\bar{v}$, điểm yên ngựa và chiến lược tối ưu nếu có. (Để đơn giản ta không ghi hàng trạng thái θ_j và cột quyết định hàng động a_{ij} .)

$$(a) \begin{bmatrix} 1 & 0 & -3 \\ 6 & -3 & -1 \\ 3 & 1 & 2 \end{bmatrix}$$

$$(b) \begin{bmatrix} -3 & 1 & -2 \\ 4 & -2 & -1 \\ 3 & 2 & 1 \end{bmatrix}$$

$$(c) \begin{bmatrix} 9 & 7 & 8 & 10 \\ 6 & 5 & 12 & 8 \\ 8 & 10 & 5 & 9 \end{bmatrix}$$

$$(d) \begin{bmatrix} 2 & 6 & 1 & 2 \\ 3 & 5 & 4 & 3 \\ 1 & 0 & 2 & 4 \end{bmatrix}$$

$$(e) \begin{bmatrix} 1 & 1 & 1 & 2 \\ 1 & 2 & 0 & 0 \\ 1 & 2 & 1 & 1 \end{bmatrix}$$

$$(g) \begin{bmatrix} 4 & -3 & 7 & -4 & 9 \\ 5 & -2 & -1 & -2 & 0 \\ 6 & -8 & 8 & -2 & -3 \end{bmatrix}$$

7.18 Xác định khoảng thay đổi của p và q để a_{22} là điểm yên ngựa trong trò chơi hai người tổng zero có ma trận giải thưởng đã cho

$$(a) \begin{bmatrix} 1 & q & 6 \\ p & 5 & 10 \\ 6 & 2 & 3 \end{bmatrix}$$

$$(b) \begin{bmatrix} 2 & 4 & 5 \\ 10 & 7 & q \\ 4 & p & 6 \end{bmatrix}$$

7.19 Giải bằng hình học trò chơi dưới đây, dùng chiến lược hỗn hợp.

(a)

		B	
		1	2
A	1	2	4
	2	2	3
	3	3	2
	4	-2	6

(b)

		B			
		1	2	3	4
A	1	2	2	3	-1
	2	2	4	3	2

(c)

		B			
		1	2	3	4
A	1	1	3	-3	7
	2	2	5	4	-6

7.20 Giả sử ma trận của trò chơi thỏa điều kiện $a_{ij} = -a_{ji}$ với mọi $i, j = 1, \dots, n$. Chứng minh rằng giá trị của trò chơi bằng 0.

7.21 Giải trò chơi sau bằng quy hoạch tuyến tính

		B		
		1	2	3
A	1	2	-1	-3
	2	-3	3	-1
	3	-4	-3	3

7.22 (Trò chơi đại tá Blotto). Đại tá Blotto có 2 trung đoàn và đối phương có 3, cần chia ra để đánh chiếm 2 vị trí. Gọi n_1 (n_2) là số trung đoàn ông bố trí ở vị trí 1 (2, tương ứng) và m_1 , m_2 là số trung đoàn đối phương bố trí. "Phần thưởng" của đại tá được coi là như sau. Nếu $n_1 > m_1$ ông nhận $m_1 + 1$ và nếu $n_2 > m_2$ ông nhận $m_2 + 1$. Nếu $n_1 < m_1$ ông mất $n_1 + 1$ và nếu $n_2 < m_2$ ông mất $n_2 + 1$. Nếu số trung đoàn hai bên bằng nhau thì ông nhận 0. Phát biểu bài toán thành trò chơi hai người tổng zero và giải bằng quy hoạch tuyến tính.

CHƯƠNG 8

CÁC PHƯƠNG PHÁP DỰ BÁO

Trong Chương 7 ta thấy việc chọn quyết định hoạt động sẽ thuận lợi hơn khi đã biết phân bố xác suất của các dữ kiện ngẫu nhiên. Các xác suất này thực ra là được suy ra từ kết quả thống kê của quá khứ, dùng làm dự báo cho khả năng xảy ra trong tương lai. Khi các quyết định mang tính định lượng, chẳng hạn phải sản xuất bao nhiêu xe đạp, dựa vào các kết quả bán xe trong các giai đoạn trước, thì việc dự báo định lượng số xe bán được trong một giai đoạn sắp tới từ các số liệu xe bán trước đây là cần thiết khi chọn quyết định sản xuất. Chương này sẽ nghiên cứu các phương pháp dự báo đại lượng ngẫu nhiên trong tương lai.

Có hai kỹ thuật dự báo cơ bản là định tính và định lượng. Kỹ thuật định tính dựa vào phán đoán của các chuyên gia nên được gọi là *kỹ thuật phán đoán* (judgmental technique). Hai kỹ thuật định lượng dựa tương ứng vào hai phương pháp thống kê là *phân tích chuỗi thời gian* (time series analysis) và *phân tích hồi quy* (regression analysis).

8.1 KỸ THUẬT PHÁN ĐOÁN

8.1.1 Phương pháp nhóm chuyên gia

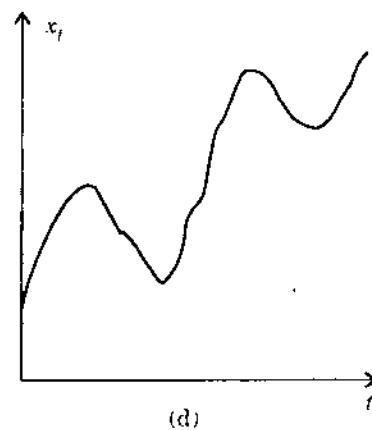
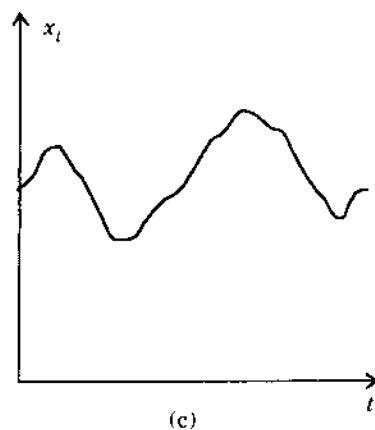
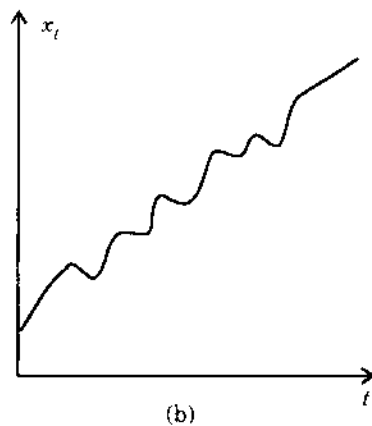
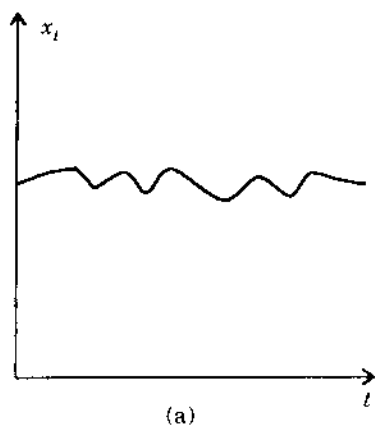
Phương pháp nhóm chuyên gia (expert group method) có nội dung đơn giản là mời một nhóm các chuyên gia họp lại và đưa ra dự báo dựa trên sự bàn bạc nhất trí.

8.1.2 Phương pháp Delphi

Phương pháp Delphi (Delphi method) sử dụng một nhóm chuyên gia, một nhóm thư ký chuyên môn và người chọn quyết định. Đầu tiên nhóm thư ký soạn *phiếu điều tra* (questionnaire) gửi đến từng chuyên gia lấy ý kiến. Dựa vào kết quả, nhóm thư ký soạn phiếu điều tra thứ hai gửi đến các chuyên gia. Khi điền vào phiếu thứ hai này các chuyên gia đã có thông tin về kết quả đợt một để tham khảo. Cuối cùng, dựa vào kết quả phân tích phiếu điều tra thứ hai, người chọn quyết định mới xây dựng dự báo. Nếu cần thiết có thể áp dụng phiếu điều tra hơn hai lần, chẳng hạn khi các ý kiến chuyên gia vẫn còn phân tán.

8.2 DỰ BÁO BẰNG CHUỖI THỜI GIAN

Đây là phương pháp dự báo giá trị x của biến ngẫu nhiên X lấy tại thời điểm $t+1$, khi biết một chuỗi giá trị biến X đã lấy trong một thời kì quá khứ ở các nút cách đều nhau. Chuỗi giá trị này có thể được biểu thị bằng đồ thị, bằng biểu đồ cột hoặc bằng bảng. Thường dùng nhất là đồ thị. Dựa vào đồ thị này ta xác định xấp xỉ mô hình sinh ra chuỗi thời gian, tức là chuỗi các giá trị của biến ngẫu nhiên. Có bốn loại mô hình điển hình. Ở H. 8.1(a) là mô hình mà quá trình sinh ra chuỗi thời gian được biểu thị bởi mức hằng số với nhiễu ngẫu nhiên, được gọi là *mô hình mức hằng số* (constant-level model). Ở (b) là *mô hình khuynh hướng tuyến tính* (linear trend model), tức là quá trình sinh ra chuỗi thời gian, (là chuỗi các giá trị của biến ngẫu nhiên) được biểu thị bởi đường thẳng bị nhiễu ngẫu nhiên. Tương tự, (c) là *mô hình mức hằng số biến động theo mùa* (constant level with seasonal-effects model), đôi khi gọi tắt là *mô hình theo mùa* (seasonal model). Cuối cùng (d) là kết hợp của (b) và (c), tức là *mô hình khuynh hướng tuyến tính biến động theo mùa* (linear trend with seasonal-effects



H. 8.1

model). Ta lần lượt xét kỹ ba mô hình đầu, còn mô hình cuối ta không trình bày vì ý tưởng đã rõ nhưng tính toán phức tạp quá. Khi dự báo một đại lượng ngẫu nhiên, trước hết ta phải quyết định mô hình.

8.2.1 Dự báo theo mô hình mức hằng số

Theo mô hình mức hằng số thì biểu diễn của quá trình sinh ra chuỗi thời gian là

$$X_t = A + e_t,$$

ở đây X_t là biến ngẫu nhiên quan sát tại thời điểm t , A là mức hằng số và e_t là sai số ngẫu nhiên tại t . Ta luôn giả thiết là e_t có kì vọng 0 và phương sai là hằng. Gọi F_{t+1} là dự báo giá trị của biến ngẫu nhiên tại thời điểm $t+1$. Có bốn cách dự báo theo mô hình mức hằng số.

1. *Dự báo theo giá trị cuối* (last - value forecasting procedure). Gọi x_t là giá trị của biến ngẫu nhiên (thực sự) lấy tại t . Phương pháp này lấy

$$F_{t+1} = x_t.$$

Kiểu dự báo này ít chính xác và hay được dùng để so sánh thêm với các dự báo phức tạp hơn.

2. *Dự báo theo trung bình* (average forecasting procedure). Thay cho việc bỏ mọi quan sát trừ quan sát cuối như trên, bây giờ ta lấy trung bình tất cả :

$$F_{t+1} = \sum_{i=1}^t \frac{x_i}{t}.$$

Dự báo này sẽ rất tốt nếu mô hình khá đúng tức là quá trình là ổn định (chỉ còn nhiễu nhỏ e_t). Nhưng thường mô hình thậm chí biến đổi đến phải thay kiểu. Vì vậy phương pháp này thường dùng cho quá trình chưa dài quá, và mọi x_t được dùng là chưa "quá cũ".

3. *Dự báo theo trung bình chuyển động* (moving - average forecasting procedure). Để tránh dữ liệu quá cũ, phương pháp này chỉ lấy trung bình n dữ liệu cuối:

$$F_{t+1} = \sum_{i=t-n+1}^t \frac{x_i}{n}$$

Chú ý rằng dễ dàng cập nhật F_{t+1} để được dự báo tiếp theo vì chỉ bỏ một x_i đầu tiên và thêm một quan sát mới. Phương pháp này còn nhược điểm là trọng số của quan sát cũ nhất x_{t-n+1} cho đến mới nhất x_t là như nhau. Do đó dẫn tới phương pháp sau đây.

4. *Dự báo làm trơn theo lũy thừa* (exponential smoothing forecasting procedure). Phương pháp này dùng công thức

$$F_{t+1} = \alpha x_t + (1 - \alpha) F_t, \quad (8.1)$$

ở đây $\alpha \in (0,1)$ là *hằng số làm trơn* (smoothing constant) mà cách chọn sẽ nói sau. Thay $F_t = \alpha x_{t-1} + (1 - \alpha)F_{t-1}, \dots$, vào ta được

$$F_{t+1} = \alpha x_t + \alpha(1 - \alpha)x_{t-1} + \alpha(1 - \alpha)^2 x_{t-2} + \dots \quad (8.2)$$

Đến đây rõ ràng là hằng số làm trơn cho trọng số nặng nhất đặt lên x_t và càng nhỏ dần đối với các quan sát càng cũ. Còn (8.1) lại cho thấy việc tính là dễ dàng, chỉ dùng đến quan sát và dự báo mới nhất viết lại là

$$F_{t+1} = F_t + \alpha(x_t - F_t),$$

cho thấy dự báo ở thời điểm $t + 1$ là dự báo ở bước trước F_t cộng với sai số dự báo lấy với trọng số α .

Một trong các bất lợi của phương pháp dự báo làm trơn theo lũy thừa là khó xác định α thích hợp. Nếu α nhỏ quá thì dự báo ở các bước thay đổi rất ít và sẽ sai nhiều nếu quá trình là kém ổn định. Nếu α lớn thì có thể dự báo thay mạnh quá. Kinh nghiệm cho thấy thường nên lấy $\alpha \in (0,1; 0,3)$.

8.2.2 Dự báo theo mô hình khuynh hướng tuyến tính

Giả sử ta thấy, theo chuỗi thời gian đã biết, mô hình dự báo xấp xỉ tốt nhất là mô hình khuynh hướng tuyến tính. Biểu diễn toán học sẽ là

$$X_t = A + Bt + e_t,$$

ở đây X_t , A và e_t định nghĩa như ở Mục 8.2.1, còn B là *nhân tố hướng* (trend factor), tức là hệ số góc. Ở mô hình mức hằng số, dự báo cho thời điểm $t+1$ cũng cho ta dự báo vào lúc này cho thời điểm $t+1+m$, $m = 1, 2, \dots$ Nhưng với mô hình khuynh hướng tuyến tính điều này không đúng nữa. Do đó ta cần khái niệm *mức đã trơn* (smoothed level) hoặc *giá trị đã trơn* (smoothed value) được xác định bởi

$$S_t = \alpha x_t + (1 - \alpha)(S_{t-1} + B),$$

ở đây $\alpha \in (0,1)$ là hằng số làm trơn. Vậy mức đã trơn ở t là tổ hợp tuyến tính với hệ số α giá trị x_t và mức trơn tại $t-1$ đã chỉnh theo khuynh hướng một đơn vị thời gian. Bây giờ dự báo tại $t+1$ được lấy là

$$F_{t+1} = S_t + B.$$

Nhưng nhân tố hướng B là chưa biết. B được xác định bằng cách làm trơn theo lũy thừa bởi công thức

$$B_t = \beta(S_t - S_{t-1}) + (1 - \beta)B_{t-1},$$

với β là một hằng số làm trơn khác, chọn tương tự như α . Do đó S_t xác định cụ thể là

$$S_t = \alpha x_t + (1 - \alpha)(S_{t-1} + B_{t-1}).$$

Dự báo trước cho m bước được lấy là

$$F_{t+m} = S_t + mB_t.$$

Các giá trị ban đầu S_0 và B_0 được lấy theo các dữ kiện đã có của quá khứ bằng cách lấy đường thẳng khớp nhất với dữ kiện đã có theo bình

phương tối thiểu như ở Mục 8.3. Tóm lại dự báo ở thời điểm t gồm các đoạn như sau

1. Dùng x_t , S_{t-1} , B_{t-1} tính $S_t = \alpha x_t + (1 - \alpha)(S_{t-1} + B_{t-1})$;
2. Dùng S_t , S_{t-1} và B_{t-1} tính $B_t = \beta(S_t - S_{t-1}) + (1 - \beta)B_{t-1}$;
3. Dự báo của m bước, $m = 1, 2, \dots$, là $F_{t+m} = S_t + mB_t$.

THÍ DỤ 8.2.1. Vào đầu thu một công ti sản xuất xe đạp muốn dự báo lượng xe bán vào mùa đông để có kế hoạch đầu tư. Chuỗi thời gian là lượng xe bán được ở các mùa đông, xuân và hè vừa qua tương ứng là $x_1 = 2800$, $x_2 = 2925$, và $x_3 = 3040$ (vậy t tương ứng với thời điểm là mùa t). Ta cần có dự báo F_5 . Giả sử ta chọn $\alpha = \beta = 0,1$. Qua ba điểm $(1, 2800)$, $(2, 2925)$ và $(3, 3040)$ ta lập được đường thẳng (theo bình phương tối thiểu) và thấy điểm ứng với mùa thu trước là $(0, 2750)$ và hệ số góc là 100. Vậy $S_0 = 2750$ và $B_0 = 100$.

Giá trị đã trơn S_1 (lượng xe bán vào mùa đông trước) là

$$S_1 = 0,1 \cdot 2800 + 0,9 (2750 + 100) = 2845.$$

Giá trị đã trơn của nhân tố hướng ở mùa đông trước là

$$B_1 = 0,1 (2845 - 2750) + 0,9 \cdot 100 = 99,5.$$

Các giá trị đã trơn tương ứng cho mùa xuân là

$$S_2 = 0,1 \cdot 2925 + 0,9 (2845 + 99,5) = 2943,$$

$$B_2 = 0,1(2943 - 2845) + 0,9(99,5) = 99,4.$$

Giá trị đã trơn cho mùa hè là

$$S_3 = 0,1 \cdot 3040 + 0,9(2943 + 99,4) = 3042,$$

$$B_3 = 0,1(3042 - 2943) + 0,9 \cdot 99,4 = 99,4.$$

Do đó, dự báo lượng xe bán cho mùa đông tới là

$$F_{3+2} = 3042 + 2.99,4 = 3241.$$

Nếu ta ở vào thời điểm cuối đông trước thì dự báo cho mùa xuân sau đó là $F_2 = S_1 + B_1 = 2845 + 99,5 = 2945$. Còn vào cuối xuân thì dự báo cho mùa hè đã phải là $F_3 = S_2 + B_2 = 2943 + 99,4 = 3042$.

8.2.3 Dự báo theo mô hình mức hằng số biến động theo mùa

Từ dữ liệu cũ hoặc từ nhân tố xã hội nếu ta thấy biến động theo mùa là hợp lí thì áp dụng mô hình này sẽ chính xác hơn. Chẳng hạn, vào mùa thu có thể bán được nhiều xe đạp hơn vì trời mát. Biểu diễn toán học của mô hình mức hằng số biến động theo mùa là

$$X_t = AI_t + e_t,$$

ở đây có yếu tố mới là I_t , gọi là *chỉ số mùa* (seasonal index) tại thời điểm t . Vì A và I_t là chưa biết, người ta xác định các mức đã trơn theo mô hình này theo công thức sau

$$S_t = \alpha \frac{x_t}{I_{t-p}} + (1 - \alpha)S_{t-1},$$

$$I_t = \gamma \frac{x_t}{S_t} + (1 - \gamma)I_{t-p},$$

ở đây α và γ , đều thuộc khoảng $(0,1)$, là các hằng làm trơn, còn p là số thời điểm trong một chu kì mùa, tức là số mùa trong năm. Dự báo cho thời điểm tới (mùa sắp đến) là

$$F_{t+1} = S_t I_{t-p+1}$$

và cho m mùa sau là, $m = 1, 2, \dots$,

$$F_{t+m} = S_t I_{t-p+m}$$

Nếu $m > p$ thì I_{t-p+m} không tính được, người ta quy ước cho nó giá trị tính được cuối cùng của chỉ số mùa tương ứng. Chẳng hạn cần

dự báo cho mùa đông 2001 nhưng chỉ có dữ liệu cho đến mùa hè 1999 thì lấy chỉ số mùa đông 1998 cho mùa đông 2000. Các giá trị đã tron ban đầu cho thuật toán theo mô hình này cần có cho cả một chu kì mùa, như sẽ thấy trong thí dụ sau.

THÍ DỤ 8.2.2. Xét dự báo bán xe đạp ở Thí dụ 8.2.1 nhưng áp dụng mô hình mới trên đây. Ta có (và cần đến) dữ liệu về lượng xe bán trong hai năm như bảng sau đây.

Mùa	năm ngoái	năm nay
đông	(t = -3) 2786	(t = 1) 2800
xuân	(t = -2) 2928	(t = 2) 2925
hè	(t = -1) 3025	(t = 3) 3040
thu	(t = 0) 3061	

Giả sử cần dự báo cho mùa thu và mùa đông tới, tức là cần F_4 và F_5 . Nhưng để minh họa thuật toán ta cũng tính cả dự báo cho các mùa trước (mà đáng lẽ tính từ trước mùa

đó). Để có xấp xỉ ban đầu cho mức hằng số A là S_0 (theo cách chọn của ta thời điểm 0 ứng với mùa thu trước) ta lấy trung bình bốn mùa năm ngoái:

$$S_0 = \frac{1}{4}(2786 + 2928 + 3025 + 3061) = 2950.$$

Còn các chỉ số mùa ứng với bốn mùa năm ngoái ta lấy bằng $\frac{x_{-t}}{S_0}$, tức là $I_{-3} =$

$$\frac{2786}{2950} = 0,944, I_{-2} = 0,993, I_{-1} = 1,025, I_0 = 1,038 \text{ (vì } X_t = AI_t + e_t \text{ nên khi xấp xỉ}$$

không tính đến nhiễu e_t thì $I_t \approx \frac{X_t}{A})$. Giả sử ta chọn $\alpha = 0,1$ và $\gamma = 0,2$.

Dự báo cho mùa đông năm nay, tính vào thu trước đó, là

$$F_1 = S_0 I_{0-4+1} = 2950 \cdot 0,944 = 2785.$$

Để có dự báo cho mùa xuân, cần giá trị đã tron S_1 và chỉ số mùa đã tron I_1 của mùa đông ngay trước đó:

$$S_1 = \alpha \frac{x_1}{I_{-3}} + (1 - \alpha)S_0 = 0,1 \cdot \frac{2800}{0,944} + 0,9 \cdot 2950 = 2952,$$

$$I_1 = \gamma \frac{x_1}{S_1} + (1 - \gamma)I_{-3} = 0,2 \frac{2800}{2952} + 0,8 \cdot 0,944 = 0,945.$$

Do đó dự báo cho mùa xuân là

$$F_2 = S_1 I_{-2} = 2952 \cdot 0,993 = 2931.$$

Tương tự, các giá trị đã trơn của mùa xuân và dự báo cho mùa hè là

$$S_2 = \alpha \frac{x_2}{I_{-2}} + (1 - \alpha)S_1 = 2951, \quad I_2 = \gamma \frac{x_2}{S_2} + (1 - \gamma)I_{-2} = 0,993,$$

$$F_3 = S_2 I_{-1} = 3025.$$

Đến lượt các giá trị đã trơn của mùa hè và dự báo cho mùa thu ta có

$$S_3 = \alpha \frac{x_3}{I_{-1}} + (1 - \alpha)S_2 = 2952, \quad I_3 = \gamma \frac{x_3}{S_3} + (1 - \gamma)I_{-1} = 1,026,$$

$$F_4 = S_3 I_0 = 3064.$$

Cũng tại điểm cuối hè này, nếu ta dự báo luôn cho mùa đông tới thì có

$$F_5 = S_3 I_1 = 2952 \cdot 0,945 = 2790.$$

8.2.4 Sai số dự báo

Sai số dự báo cho một thời điểm t tất nhiên là

$$E_t = x_t - F_t.$$

Số này được gọi là *sai số dự báo* (forecast error) hoặc *thặng dư* (residual). Nếu có dự báo n thời điểm thì sai số chung được lấy là trung bình của sai số bình phương

$$E_{TB2} = \frac{1}{n}(E_1^2 + E_2^2 + \dots + E_n^2).$$

THÍ DỤ 8.2.3. Tính sai số chung khi dự báo lượng xe bán vào mùa xuân và hè năm nay, sau khi đã có số bán thực x_2 và x_3 theo bảng ở Mục 8.2.3, của ba phương pháp dự báo xét ở các thí dụ trước. Ta có

dự báo theo giá trị cuối: $E_{TB2} = \frac{1}{2}[(2925 - 2800)^2 + (3040 - 2925)^2] = 14425,$

dự báo theo khuynh hướng tuyến tính: $E_{TB2} = \frac{1}{2}[(2925 - 2945)^2 + (3040 - 3042)^2] = 202,$

dự báo theo mùa: $E_{TB2} = \frac{1}{2}[(2925 - 2931)^2 + (3040 - 3025)^2] = 131.$

Như vậy nếu chỉ tính đến hai dự báo F_2 và F_3 thì phương pháp dự báo theo mùa chính xác nhất. Dù sao sai số E_{TB2} cũng cho thông tin cho quá trình dự báo tiếp theo.

8.2.5 Phương pháp Box - Jenkins

Trong các phương pháp dự báo bằng chuỗi thời gian trình bày ở trên, sau khi đã chọn mô hình dự báo thì trong quá trình dự báo không kiểm tra xem mô hình có phù hợp không để chỉnh lại. Phương pháp Box - Jenkins có kết hợp kiểm tra sự phù hợp của mô hình, vì vậy nó rất chính xác nhưng cũng khá phức tạp. Chẳng hạn nó cần ít nhất dữ liệu đã xảy ra của 50 thời điểm. Vì vậy phương pháp này chỉ áp dụng cho dự báo quy mô lớn. Chúng ta không trình bày ở đây; độc giả có thể xem [Hoff 1983].

8.3 DỰ BÁO BẰNG HỒI QUY TUYẾN TÍNH

Phân tích thống kê các dữ liệu thường phát hiện quan hệ giữa hai biến ngẫu nhiên mà ta quan tâm. Ở mục này ta chỉ xét trường hợp quan hệ này là tuyến tính. Nếu ta xác định được một cách xấp xỉ nào đó từ các dữ liệu, đã biết biến ngẫu nhiên Y phụ thuộc biến ngẫu nhiên X tuyến tính ra sao, thì khi có một quan sát thực $X = x$ ta sẽ dùng quan sát đó để dự báo giá trị của Y . Như vậy là ta đã dự báo bằng *hồi quy tuyến tính* (linear regression). Phương pháp dự báo này bao gồm cả việc đánh giá khoảng tin cậy chung quanh giá trị dự báo. Ta lần lượt xét hai vấn đề xác định quan hệ giữa hai biến ngẫu nhiên và tính khoảng tin cậy.

8.3.1 Xác định đường hồi quy

Quan hệ giữa hai biến ngẫu nhiên có thể là một quan hệ hàm xác định. Chẳng hạn ở Mục 8.2.2 khi xét quá trình sinh ra chuỗi thời gian có mô hình khuynh hướng tuyến tính, ta có quan hệ hàm giữa biến ngẫu nhiên X và thời gian t (là biến thứ hai, tất định trong trường hợp này) là

$$X_t = A + Bt + e_t,$$

ở đây e_t là sai số ngẫu nhiên. Khi phân tích hồi quy ta luôn ký hiệu biến phụ thuộc là Y nên quan hệ trên được ký hiệu lại là

$$Y_t = A + Bt + e_t. \quad (8.3)$$

Quan hệ kiểu này là *quan hệ hàm chính xác* (exact functional relationship).

Quan hệ giữa hai biến ngẫu nhiên có thể xuất hiện ở dạng khác.

Chẳng hạn giám đốc một công ty phát hành sách phát hiện qua các dữ liệu (đã xảy ra) trong quá khứ là lượng sách bán Y của mỗi đầu sách tại các cửa hàng phụ thuộc tuyến tính vào lượng sách bán được bằng yêu cầu qua thư X . Nhưng trường hợp này không viết được quan hệ chính xác như trên, mà hiểu là kì vọng của Y là phụ thuộc tuyến tính theo giá trị quan sát thấy x của X , tức là

$$E(Y/X = x) = A + Bx. \quad (8.4)$$

Như vậy ta nói giữa X và Y có một *mức độ quan hệ* (degree of association), hoặc nói (8.4) là một *mô hình mức độ quan hệ* (degree of association model). Thực ra quan hệ hàm (8.3) cũng dẫn đến mô hình mức độ quan hệ vì nếu lấy kì vọng (8.3) ta được

$$E(Y_i) = A + Bx_i.$$

Vì vậy cả hai kiểu quan hệ (8.3) và (8.4) được xử lý như nhau. Để xác định A và B ta dùng phương pháp bình phương tối thiểu và các dữ liệu đã có. Giả sử đã quan sát được n giá trị của X và Y là $x_i, y_i, i = 1, \dots, n$. Ước lượng xấp xỉ bình phương tối thiểu $\bar{y}_i$ ứng với giá trị x_i là $\bar{y}_i = a + bx_i$ với a và b là giá trị của A và B sẽ xác định phải thỏa

$$Q := \sum_{i=1}^n (y_i - \bar{y}_i)^2 \rightarrow \min.$$

Như vậy ta phải tìm cực tiểu của hàm hai biến $Q(a, b)$ theo điều kiện $\frac{\partial Q}{\partial a} = 0, \frac{\partial Q}{\partial b} = 0$. Tính cụ thể ta sẽ được

$$b = \frac{\sum_{i=1}^n x_i y_i - \frac{1}{n} \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sum_{i=1}^n x_i^2 - \frac{1}{n} \left(\sum_{i=1}^n x_i \right)^2} := \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2},$$

$$a = \bar{y} - b\bar{x},$$

$$\text{ở đây } \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i.$$

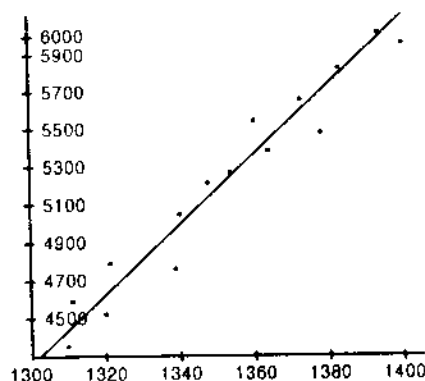
THÍ DỤ 8.3.1. Lượng sách bán được bằng hai cách tại cửa hàng và bằng thư của 15 đầu sách của một công ti phát hành được cho ở bảng bên và được biểu diễn bằng các điểm đen trên H. 8.2. Đường thẳng tìm bằng phương pháp bình phương tối thiểu vẽ trên H. 8.2 được tính như sau

$$\bar{x} = 1353,1, \quad \bar{y} = 5219,3,$$

$$\sum_{i=1}^{15} (x_i - \bar{x})(y_i - \bar{y}) = 214543,9,$$

$$\sum_{i=1}^{15} (x_i - \bar{x})^2 = 11966,$$

$$a = -19041,9, \quad b = 17,930.$$



H. 8.2

	Bán qua thư	Bán cửa hàng
1	1310	4360
2	1313	4590
3	1320	4520
4	1322	4770
5	1338	4760
6	1340	5070
7	1347	5230
8	1355	5080
9	1360	5550
10	1364	5390
11	1373	5670
12	1376	5490
13	1384	5810
14	1395	6060
15	1400	5940

Đường thẳng $\hat{y} = -19041,9 + 17,930x$ được gọi là *đường hồi quy* (regression line). Dùng đường hồi quy ta dự báo được số lượng sách bán tại cửa hàng của bất kì đầu sách nào khác nếu biết lượng bán được theo thư.

8.3.2 Ước lượng khoảng tin cậy

Để tính được các khoảng tin cậy người ta giả sử hai điều kiện sau thỏa

(a) lấy được ngẫu nhiên n cặp $(x_1, Y_1), \dots, (x_n, Y_n)$;

(b) Các giá trị ngẫu nhiên Y_i có phân bố chuẩn với kì vọng $A + Bx_i$ và có cùng phương sai σ^2 . Một giá trị σ^2 hay được dùng (coi như giả thiết) là lượng $s_{y|x}^2$ sau đây (y_i là quan sát được của Y_i)

$$s_{y|x}^2 = \sum_{i=1}^n \frac{(y_i - \bar{y})^2}{n - 2}.$$

Người ta thường dùng ba loại khoảng tin cậy được xét lần lượt sau đây.

1. Khoảng tin cậy cho kì vọng

Lấy dự báo cho kì vọng của biến ngẫu nhiên Y khi đã biết $X = x^*$, tức là $E(Y/X = x^*)$ xấp xỉ là điểm trên đường hồi quy :

$$\bar{y} = a + bx^*.$$

Vấn đề là với xác suất tin cậy là $1 - \alpha$ thì $E(Y/X = x^*)$ sẽ rơi vào khoảng nào quanh $\bar{y} = a + bx^*$. Người ta chứng minh được công thức cho khoảng tin cậy này dựa vào công thức tính xác suất cho biến ngẫu nhiên t có phân bố Student (Student's distribution), một phân bố xác suất hay gặp. Người ta đã tính sẵn xác suất để biến ngẫu nhiên t , có phân bố Student với v bậc tự do, nhận giá trị $\geq t_{v,\beta}$ là bằng β , với $t_{v,\beta}$ cho ở Bảng A_2 cuối sách này. Khi đó người ta chứng minh được rằng với xác suất đảm bảo là $1 - \alpha$ thì $E(Y/X = x^*)$ rơi vào khoảng $(a + bx^* - T_1, a + bx^* + T_1)$ với

$$T_1 = t_{\frac{\alpha}{2}, n-2} s_{y|x} \sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

Nhận xét rằng khoảng tin cậy đảm bảo nhỏ (tức là dự báo chính xác) nếu $x^* = \bar{x}$ và càng rộng nếu x^* xa $\bar{x}$.

THÍ DỤ 8.3.2. Xét việc dự báo lượng sách bán tại cửa hàng ở Thí dụ 8.3.1, ứng với lượng sách bán theo thư đã biết là $x^* = 1400$. Theo bảng bán 15 sách ở Thí dụ 8.3.1 thì tính được $s_{y|x} = 130,5$. Tìm khoảng đảm bảo với xác suất 0,95 $= 1 - \alpha$ là $E(Y|X = 1400)$ rơi vào khoảng đó. Ta đã có, theo Thí dụ 8.3.1,

$$a + bx^* = -19041,9 + 17,930 \cdot 1400 = 6060$$

là dự báo cho $E(Y|X = 1400)$. Theo Bảng A₂ ở cuối sách thì $t_{0,025;13} = 2,160$. Do đó

$$T_1 = 2,160 \cdot 130,5 \sqrt{\frac{1}{15} + \frac{46,9^2}{11966}} = 141.$$

Vậy khoảng tin cậy cần tìm là $(6060 - 141, 6060 + 141) = (5919, 6201)$.

2. Khoảng tin cậy cho giá trị của biến ngẫu nhiên

Trong nhiều trường hợp người ta cần khoảng tin cậy để đảm bảo với xác suất $1 - \alpha$ là giá trị thực sự của biến ngẫu nhiên rơi vào đó.

Người ta chứng minh được khoảng này là $(a + bx^* - T_2, a + bx^* + T_2)$ với

$$T_2 = t_{\frac{\alpha}{2}, n-2} s_{y|x} \sqrt{1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}.$$

Xét Thí dụ 8.3.2 với $x^* = 1400$, $1 - \alpha = 0,95$ thì $T_2 = 315$, tức là khoảng này không xác định được nhỏ như khoảng tin cậy cho kì vọng.

Điều này là hiển nhiên. Khoảng tin cậy cho giá trị cụ thể của biến ngẫu nhiên thường được gọi là *khoảng tiên định* (prediction interval).

3. Khoảng nói lỏng đồng thời

Với việc xác định khoảng tin cậy theo hai kiểu trên, công ti phát hành sách không thể dùng để đánh giá dự báo nhiều lần, cho nhiều sách khác nhau. Vì các kết luận về khoảng tin cậy của dự báo dựa trên cùng một số dữ liệu (chẳng hạn bảng bán 15 sách) nên không độc lập. Cho dù giả sử là độc lập thì áp dụng k lần khẳng định về khoảng tin cậy với xác suất $1 - \alpha$ thì xác suất đúng cả k lần chỉ còn là $(1 - \alpha)^k$. Nếu k lớn thì không áp dụng được nữa.

Để vượt qua khó khăn này người ta dùng một khoảng tin cậy khác, gọi là *khoảng nói lỏng đồng thời* (simultaneous tolerance interval). Khoảng này cho biết với xác suất chắc chắn là $1 - \alpha$, giá trị cụ thể của biến ngẫu nhiên phải rơi vào khoảng này, dù ta có xét k lần với k bằng bao nhiêu cũng được. Người ta tính được khoảng này là $(a + bx^* - T_3, a + bx^* + T_3)$ với

$$T_3 = c^{**} s_{y|x} \sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

ở đây c^{**} đã được tính sẵn theo bảng dưới đây, phụ thuộc thêm cả vào xác suất P để tất cả các lần khẳng định về khoảng tin cậy là chính xác.

Chẳng hạn với thí dụ về bán sách ta đã biết $a + bx^* = 6060$ nếu $x^* = 1400$. Ta muốn tính khoảng nói lỏng đồng thời sao cho với xác suất $P = 0,90$ để ít nhất 0,95.100% lượng sách bán được ở cửa hàng tương ứng với các lượng sách bán theo thư là 1400 rơi vào khoảng $(6060 - T_3, 6060 + T_3)$. Tra bảng c^{**} ứng với $P = 0,90$, $\alpha = 0,05$ (vì $1 - \alpha = 0,95$) và $n = 15$ ta được $c^{**} = \frac{1}{2}(11,447 + 11,803) = 11,625$. Theo công thức ta sẽ có $T_3 = 759$. Ta thấy khoảng để đảm bảo chứa các lượng

342 CHƯƠNG 8 CÁC PHƯƠNG PHÁP DỰ BÁO

sách bán đều rơi vào phải lớn hơn hai khoảng tin cậy ở hai mục trước.

n	$\alpha = 0,50$	$\alpha = 0,25$	$\alpha = 0,10$	$\alpha = 0,05$	$\alpha = 0,01$	$\alpha = 0,005$
P = 0,90						
4	7,471	10,160	13,069	14,953	18,663	23,003
6	5,380	7,453	9,698	11,150	14,014	17,363
8	5,037	7,082	9,292	10,722	13,543	16,837
10	4,983	7,093	9,366	10,836	13,733	17,118
12	5,023	7,221	9,586	11,112	14,121	17,634
14	5,101	7,394	9,857	11,447	14,577	18,232
16	5,197	7,586	10,150	11,803	15,057	18,856
18	5,300	7,786	10,449	12,165	15,542	19,484
20	5,408	7,987	10,747	12,526	16,023	20,140
P = 0,95						
4	10,756	14,597	18,751	21,445	26,760	32,982
6	6,652	9,166	11,899	13,669	17,167	21,266
8	5,933	8,281	10,831	12,484	15,750	19,568
10	5,728	8,080	10,632	12,286	15,553	19,309
12	5,684	8,093	10,701	12,391	15,724	19,619
14	5,711	8,194	10,880	12,617	16,045	20,050
16	5,771	8,337	11,107	12,898	16,431	20,559
18	5,848	8,499	11,357	13,204	16,845	21,097
20	5,937	8,672	11,619	13,521	17,272	21,652
P = 0,99						
4	24,466	33,019	42,398	48,620	60,500	74,642
6	10,444	14,285	18,483	21,215	26,606	32,920
8	8,290	11,453	14,918	17,166	21,652	26,860
10	7,567	10,539	13,796	15,911	20,097	24,997
12	7,238	10,182	13,383	15,479	19,579	24,403
14	7,127	10,063	13,267	15,355	19,485	24,316
16	7,079	10,055	13,306	15,410	19,582	24,467
18	7,074	10,111	13,404	15,552	19,794	24,746
20	7,108	10,198	13,566	15,745	20,065	25,122

Bảng giá trị của c**

BÀI TẬP CHƯƠNG 8

8.1 Giả sử dự báo trước đã có là 2083, giá trị thực tế xảy ra của biến ngẫu nhiên vào thời điểm cuối là 1975 và giá trị cũ nhất (ở dữ liệu) của biến ngẫu nhiên là 1.945. Tính dự báo mới cho thời điểm tiếp theo bằng phương pháp trung bình chuyển động dựa trên bốn quan sát mới nhất.

8.2 Giả sử dự báo trước đã có là 782, giá trị thực tế xảy ra của biến ngẫu nhiên vào thời điểm cuối là 794 và $\alpha = 0,1$. Dùng cách làm trơn theo lũy thừa tính dự báo mới cho thời điểm tiếp theo.

8.3 Giải lại bài toán dự báo về bán xe đạp ở Thí dụ 8.2.1 với các hằng số làm trơn $\alpha = \beta = 0,3$.

8.4 Biết chuỗi thời gian của tỉ lệ thất nghiệp tính vào đầu mỗi quý từ 1990 đến quý 3 năm 1994 là bảng sau đây

quý	% thất nghiệp	quý	% thất nghiệp	quý	% thất nghiệp
I/90	6,9	I/91	8,2	I/92	9,4
II/90	6,7	II/91	7,0	II/92	9,2
III/90	7,9	III/91	7,3	III/92	9,8
IV/90	7,1	IV/91	7,5	IV/92	9,9

quý	% thất nghiệp	quý	% thất nghiệp
I/93	11,4	I/94	8,8
II/93	10,0	II/94	7,6
III/93	9,4	III/94	7,5
IV/93	8,4		

Áp dụng mô hình theo mùa, dùng dữ liệu cho năm 1990 để ước lượng S_0 và chỉ số mùa ban đầu. Dùng cách làm trơn theo lũy thừa với biến động mùa dự báo tỉ lệ thất nghiệp cho bốn quý năm 1991. Lại dùng dự báo cho quý IV/91 để dự báo cho quý IV/94.

Lấy $\alpha = \gamma = 0,1$.

8.5 Sử dụng bảng ở bài 8.4.

(a) Dùng dự báo theo trung bình chuyển động dựa trên ba quý cuối cùng dự báo cho tất cả các quý từ IV/90 cho đến IV/94.

(b) Xét thẳng dư cho các quý từ IV/90 đến III/94 và tính trung bình của sai số bình phương.

8.6 Giả sử có chuỗi thời gian theo bảng sau đây.

t	y	t	y
1	430	6	514
2	446	7	532
3	464	8	548
4	480	9	570
5	498	10	591

(a) Dùng phương pháp bình phương tối thiểu xác định đường hồi quy.

(b) Dùng đường hồi quy dự báo Y_{11} .

(c) Bắt đầu bằng dự báo cho thời điểm 3, dùng kỹ thuật làm trơn với ước lượng đầu 430 và $\alpha = 0,1$, tức là dự báo cho thời điểm 3 = $0,1.446 + 0,9.430$, để dự báo cho mọi thời điểm cho đến 11.

(d) Bắt đầu bằng dự báo một bước cho thời điểm 3, dùng kỹ thuật làm trơn vận dụng cho khuynh hướng với ước lượng ban đầu cho biến ngẫu nhiên là 430 và cho khuynh hướng là 18, hãy dự báo cho mọi thời điểm đến 11. Lấy $\alpha = \beta = 0,2$.

CHƯƠNG 9**XÍCH MARKOV VÀ QUÁ TRÌNH QUYẾT ĐỊNH MARKOV****9.1 QUÁ TRÌNH NGẪU NHIÊN VÀ XÍCH MARKOV****9.1.1 Định nghĩa**

Chúng ta đã nghiên cứu các bài toán vận trù có chứa tính ngẫu nhiên, bất định và đã thấy ứng dụng của khái niệm biến ngẫu nhiên. Nếu cần xét cả quá trình theo thời gian thì biến ngẫu nhiên lại phụ thuộc thời gian, nó lấy ngẫu nhiên giá trị tại thời điểm t , khác tại s . Nói cách khác là ta có tập hợp các biến ngẫu nhiên $\{X_t\}$ phụ thuộc chỉ số thời gian t . Tập hợp này sẽ gọi là quá trình ngẫu nhiên. Rõ ràng các bài toán "động" theo thời gian là rất phổ biến, nên quá trình ngẫu nhiên được ứng dụng rất nhiều.

Nếu chỉ có hữu hạn hoặc đếm được thời điểm t của quá trình thì ta nói là có một *quá trình ngẫu nhiên rời rạc* (discrete stochastic process). Nếu có không đếm được thời điểm thì ta có *quá trình ngẫu nhiên liên tục* (continuous stochastic process), lúc này cũng hay ký hiệu là $\{X(t)\}$. Tập tất cả các giá trị biến ngẫu nhiên có thể lấy ở mọi thời điểm, tức là tập tất cả các *kết quả* (outcome) của quá trình ngẫu nhiên được gọi là *không gian trạng thái* (state space) hoặc *không gian mẫu* (sample space). Chú ý rằng chữ "giá trị" ở đây có thể là định lượng nhưng cũng có thể là định tính như sẽ thấy ở các thí dụ dưới

đây. Chúng ta xét chủ yếu là trường hợp không gian trạng thái có hữu hạn trạng thái $i = 0, 1, \dots, M$, tức là $M+1$ trạng thái.

THÍ DỤ 9.1.1. Xét quá trình thời tiết với các thời điểm là các ngày $t = 0, 1, \dots$, ở đây $t = 0$ là hôm nay, $t = 1$ là ngày mai... Để đơn giản giả sử ta chỉ quan tâm đến hai trạng thái mưa và nắng. Khi đó ta có một quá trình ngẫu nhiên rời rạc (điểm được thời điểm) và có hai trạng thái. Nếu thời tiết là hoàn toàn ngẫu nhiên thì chẳng còn gì để nói. Vì vậy, hy vọng là quá trình tương đối có quy luật. Theo dõi thống kê một thời gian, chẳng hạn 1000 ngày, ta có thể phát biểu quy luật bằng ngôn ngữ xác suất để giải quyết các vấn đề liên quan đến thời tiết sau này. Ở đây để đơn giản ta chỉ dựa trên quan sát 10 ngày và thấy (ở đây chỉ có 2 trạng thái nên ta không ký hiệu X_t mà ký hiệu luôn M_t và N_t từ chữ "mưa" và "nắng")

$$M_0, M_1, M_2, N_3, N_4, N_5, N_6, N_7, N_8, M_9.$$

Mô hình đơn giản nhất về tính ngẫu nhiên của thời tiết được chọn là coi xác suất để mưa trong ngày n bất kỳ là $P(M_n) = \frac{4}{10}$, và nắng là $P(N_n) = \frac{6}{10}$ vì trong 10 ngày quan sát có 4 mưa, 6 nắng. Như vậy là ta coi thời tiết của các ngày dù là liên nhau cũng độc lập. Mô hình này quá đơn giản, thậm chí chẳng cần phát biểu thành "quá trình". Mô hình khác, đơn giản nhất có thể mà khá phù hợp với thực tế mà sau đây sẽ gọi là xích Markov, là coi thời tiết ở ngày n chỉ phụ thuộc vào thời tiết ở ngày $n-1$, không liên quan đến các ngày trước đó, và mức độ phụ thuộc như nhau cho mọi n . Với mô hình thời tiết theo xích Markov này ta chỉ phải xác định bốn xác suất chuyển dựa vào quan sát thống kê. Ta thấy trong dãy quan sát có 3 trường hợp đứng ở một ngày mưa M_{n-1} để xem

trạng thái	M_n	N_n
M_{n-1}	$\frac{2}{3}$	$\frac{1}{3}$
N_{n-1}	$\frac{1}{6}$	$\frac{5}{6}$

ngày hôm sau là mưa hay nắng, cụ thể có M_0M_1 , M_1M_2 và M_2N_3 , tức là có $\frac{2}{3}$ là chuyển sang mưa và

$\frac{1}{3}$ sang nắng. Do đó ta thấy xác suất có điều kiện là

$$P(M_n | M_{n-1}) = \frac{2}{3} \text{ và } P(N_n | M_{n-1}) = \frac{1}{3}. \text{ Tương tự}$$

$P(M_n | N_{n-1}) = \frac{1}{6}$ và $P(N_n | N_{n-1}) = \frac{5}{6}$. Ta viết vào ma trận (như trang trước) gọi là *ma trận truyền* (transition matrix). Giả sử ta làm tròn số như sau. Đồng thời, vì quá trình là "đứng", tức là không phụ thuộc thời điểm n , ta viết lại ma trận truyền như sau

trạng thái	M	N
M	0,6	0,4
N	0,2	0,8

$$= \begin{bmatrix} P_{MM} & P_{MN} \\ P_{NM} & P_{NN} \end{bmatrix}$$

Một trong các mục đích đầu tiên của mô hình là tính xác suất để một ngày nào đó trong tương lai, là mưa hoặc nắng khi biết trạng thái hôm nay. Ta hãy xét trực tiếp (mặc dù sau đây có nêu lý thuyết tổng quát hơn). Các xác suất này có thể tính bằng công thức truy hồi, ngày n theo ngày $n-1$, lần ngược về ngày 0 đã biết. Vì chỉ có hai trạng thái nên ta có

$$P(M_n) = P(M_n \cap M_{n-1}) + P(M_n \cap N_{n-1}),$$

$$P(N_n) = P(N_n \cap M_{n-1}) + P(N_n \cap N_{n-1}).$$

Biểu diễn qua xác suất có điều kiện ta được

$$P(M_n) = P(M_n | M_{n-1})P(M_{n-1}) + P(M_n | N_{n-1})P(N_{n-1}), \quad (9.1)$$

$$P(N_n) = P(N_n | M_{n-1})P(M_{n-1}) + P(N_n | N_{n-1})P(N_{n-1}). \quad (9.2)$$

Thay các xác suất có điều kiện ở bảng ma trận truyền vào ta được

$$P(M_n) = 0,6P(M_{n-1}) + 0,2P(N_{n-1}),$$

$$P(N_n) = 0,4P(M_{n-1}) + 0,8P(N_{n-1}).$$

Thêm hai phương trình luôn đúng cho xác suất (vì chỉ có hai trạng thái) là $P(M_n) + P(N_n) = 1$, $P(M_{n-1}) + P(N_{n-1}) = 1$ vào và rút gọn ta được một công thức truy hồi

$$P(N_n) = 0,4 + 0,4P(N_{n-1}). \quad (9.3)$$

Từ đó tính nốt $P(M_n) = 1 - P(N_n)$. Nếu hôm nay nắng thì $P(N_0) = 1, P(M_0) = 0$. Nếu hôm nay mưa thì $P(M_0) = 1, P(N_0) = 0$. Sử dụng công thức truy hồi ta tính được hai bảng *xác suất quá độ* (transient probability) như sau.

Các xác suất quá độ

khi biết $P(M_0) = 1$

n	$P(M_n)$	$P(N_n)$
0	1	0
1	0,6	0,4
2	0,44	0,56
3	0,376	0,624
4	0,350	0,650
5	0,340	0,660
6	0,336	0,664
7	0,333	0,667
8	0,333	0,667
9	0,333	0,667
10	0,333	0,667

Các xác suất quá độ

khi biết $P(N_0) = 1$

n	$P(M_n)$	$P(N_n)$
0	0	1
1	0,2	0,8
2	0,28	0,72
3	0,312	0,688
4	0,325	0,675
5	0,330	0,670
6	0,332	0,668
7	0,333	0,667
8	0,333	0,667
9	0,333	0,667
10	0,333	0,667

Qua hai bảng xác suất quá độ ta thấy với n nhỏ thì $P(M_n)$ (cũng như $P(N_n)$) trong mỗi bảng thay đổi theo n và khác nhau nhiều ở hai bảng. Nhưng với n lớn thì cố định $P(M_n) = 0,333$ cho mọi n và cả hai bảng, tức là không phụ thuộc vào trạng thái hôm nay. Điều này chứng tỏ tồn tại $\lim_{n \rightarrow \infty} P(M_n) := P(M) = 0,333$, sẽ

được gọi là *xác suất duy trì* (steady-state probability). Nếu biết nó tồn tại thì tìm được ngay từ công thức truy hồi. Chẳng hạn từ (9.3), qua giới hạn ta được $P(N) = 0,4 + 0,4P(N)$. Do đó $P(N) = \frac{2}{3} = 0,667$.

ĐỊNH NGHĨA 9.1.1. Quá trình ngẫu nhiên $\{X_t\}$ được gọi là có *tính chất Markov* (Markovian property) hoặc là *xích Markov* (Markov chain) nếu

$$P\{X_{t+1} = j \mid X_0 = k_0, \dots, X_{t-1} = k_{t-1}, X_t = i\} = P\{X_{t+1} = j \mid X_t = i\}, \quad \forall t, i, j, k_0, \dots, k_{t-1}.$$

Quá trình ngẫu nhiên thỏa điều kiện là *xác suất chuyển* (transition

probability) $P \{X_{t+1} = j \mid X_t = i\}$ không phụ thuộc t , tức là

$$P \{X_{t+1} = j \mid X_t = i\} = P \{X_1 = j \mid X_0 = i\}, \quad t = 0, 1, \dots,$$

được gọi là *quá trình dừng* (stationary process). Lúc đó ký hiệu

$$p_{ij} := P \{X_1 = j \mid X_0 = i\}$$

được gọi là *xác suất truyền dừng*. Nếu xác suất truyền là dừng thì dễ thấy

$$P \{X_{t+n} = j \mid X_t = i\} = P \{X_n = j \mid X_0 = i\} := p_{ij}^{(n)}.$$

$p_{ij}^{(n)}$ được gọi là *xác suất truyền n bước* (n - step transition probability). Ta quy ước viết $p_{ij}^{(1)} = p_{ij}$ như đã ký hiệu ở trên.

Sau đây chỉ xét xích Markov có xác suất truyền dừng nên ta gọi tắt là xích Markov. Giả sử có $M+1$ trạng thái $i = 0, 1, \dots, M$. Khi đó

$$p_{ij}^{(n)} \geq 0, \quad \sum_{j=0}^M p_{ij}^{(n)} = 1 \quad \forall i, j = 0, \dots, M, \quad \forall n = 0, 1, \dots$$

Ta ký hiệu ma trận truyền là

$$P = \begin{bmatrix} p_{00} & p_{01} & \cdots & p_{0M} \\ p_{10} & p_{11} & \cdots & p_{1M} \\ \vdots & \vdots & \ddots & \vdots \\ p_{M0} & p_{M1} & \cdots & p_{MM} \end{bmatrix}$$

9.1.2 Phương trình Chapman - Kolmogorov

Phương trình Chapman - Kolmogorov là phương trình để tính các xác suất truyền n bước, $\forall i, j, n$ và $0 \leq m \leq n$,

$$p_{ij}^{(n)} = \sum_{k=0}^M p_{ik}^{(m)} p_{kj}^{(n-m)}.$$

CHỨNG MINH. $p_{ik}^{(m)} p_{kj}^{(n-m)}$ là xác suất có điều kiện để xuất phát từ trạng thái i quá trình chuyển đến trạng thái k sau m bước rồi chuyển đến trạng thái j sau $n-m$ bước. Do đó lấy tổng mọi xác suất có điều kiện này với mọi trạng thái k có thể là $k = 0, 1, \dots, M$ ta được $p_{ij}^{(n)}$. $\square$

Chú ý rằng với mỗi m thỏa $0 \leq m \leq n$ ta được một phương trình Chapman - Kolmogorov cho $p_{ij}^{(n)}$. Nói riêng với $m = 1$ chẳng hạn thì

$$p_{ij}^{(n)} = \sum_{k=0}^M p_{ik} p_{kj}^{(n-1)}.$$

Với $m = 1$ và $n = 2$ ta có

$$p_{ij}^{(2)} = \sum_{k=0}^M p_{ik} p_{kj}.$$

Ta thấy $\sum_{k=0}^M p_{ik} p_{kj}$ cũng chính là phần tử ở vị trí (i, j) của P^2 . Vậy

$$P^{(2)} = P^2.$$

Dễ thấy bằng phép nhân tiếp là ta có tổng quát

$$P^{(n)} = P^n.$$

Từ đây ta có ngay dạng ma trận của phương trình Chapman - Kolmogorov

$$P^{(n)} = P^{(m)} P^{(n-m)},$$

chỉ là kết quả của tính chất kết hợp của phép nhân ma trận.

Bây giờ ta tính xác suất (không điều kiện) của một trạng thái j trong tương lai ở bước n là $P\{X_n = j\}$. Tương tự như (9.1), (9.2) ở Thí dụ 9.1.1 ta có

$$\begin{aligned} P\{X_n = j\} &= \sum_{i=0}^M P\{X_n = j | X_{n-1} = i\} P\{X_{n-1} = i\} \\ &= \sum_{i=0}^M p_{ij} P\{X_{n-1} = i\}. \end{aligned} \quad (9.4)$$

Nhưng bây giờ có $M+1$ trạng thái, không phải chỉ hai, nên ta không có được công thức truy hồi như (9.3) chỉ chứa một trạng thái. Tuy vậy nếu đã tính trước các xác suất truyền n bước $p_{ij}^{(n)}$ thì ta cũng tính được ngay $P\{X_n = j\}$ chỉ dùng $p_{ij}^{(n)}$ và các xác suất (không điều kiện) của trạng thái hiện tại $P\{X_0 = i\}$:

$$\begin{aligned} P\{X_n = j\} &= \sum_{i=0}^M P\{X_n = j | X_0 = i\} P\{X_0 = i\} \\ &= \sum_{i=0}^M p_{ij}^{(n)} P\{X_0 = i\}. \end{aligned} \quad (9.5)$$

Công thức này cũng hiển nhiên theo định nghĩa của xác suất có điều kiện. Nhưng nó cũng suy từ (9.4) như sau

$$\begin{aligned} P\{X_n = j\} &= \sum_{i=0}^M p_{ij} P\{X_{n-1} = i\} = \sum_{i=0}^M p_{ij} \sum_{k=0}^M p_{ki} P\{X_{n-2} = k\} \\ &= \sum_{k=0}^M P\{X_{n-2} = k\} \sum_{i=0}^M p_{ki} p_{ij} = \sum_{k=0}^M p_{kj}^{(2)} P\{X_{n-2} = k\} \end{aligned}$$

$$= \sum_{i=0}^M p_{ij}^{(2)} P\{X_{n-2} = i\} = \dots = \sum_{i=0}^M p_{ij}^{(n)} P\{X_0 = i\}.$$

Nếu biết trạng thái hiện tại i_0 xảy ra, tức là $P\{X_0 = i_0\} = 1$ thì ta có

$$P\{X_n = j\} = p_{i_0 j}^{(n)}$$

và ta được bảng xác suất quá độ $P\{X_n = j\}$ tương tự bảng ở Thí dụ 9.1.1 ngay từ ma trận xác suất truyền cấp n .

9.1.3 Phân loại trạng thái của xích Markov

Trạng thái j được gọi là *đến được* (accessible) từ trạng thái i nếu $p_{ij}^{(n)} > 0$ với n nào đó. Nếu j đến được từ i và i đến được từ j thì hai trạng thái i và j được gọi là *thông nhau* (communicating), hoặc ta nói i thông với j . Rõ ràng thông nhau là một quan hệ tương đương, tức là: i thông với i , i thông với j thì j thông với i và có tính bắc cầu là i thông với k , k thông với j thì i thông với j (do phương trình Chapman - Kolmogorov). Vì vậy không gian trạng thái được chia thành các lớp tương đương. Mọi trạng thái thông nhau tạo thành một lớp. Nếu cả không gian trạng thái chỉ gồm một lớp tương đương thì xích Markov được gọi là *không thu gọn được* (irreducible).

Gọi f_{ii} là xác suất để quá trình bắt đầu từ trạng thái i sẽ quay lại i (lúc nào đó). Nếu $f_{ii} = 1$, ta nói i là *trạng thái lặp* (recurrent state). Nếu $f_{ii} < 1$ ta nói i là *trạng thái quá độ* (transient state). Trường hợp riêng của trạng thái lặp là $p_{ii} = 1$, tức là xác suất truyền sau một bước đã bằng 1, thì ta nói i là *trạng thái hấp thụ* (absorbing state). Tuy nhiên, để xem một trạng thái i là lặp hay quá độ, thường rất khó tính f_{ii} . Vì vậy thường dùng một số tính chất sau đây thuận tiện hơn.

Trạng thái i là lặp tức là xác suất để quá trình quay lại i bằng một. Nhưng quá trình là Markov tức là quá khứ không có ảnh hưởng

nên lại bắt đầu i thì xác suất quay lại là một. Lập lại lí luận nữa ta thấy i sẽ được quay lại vô hạn lần. Nếu i là trạng thái quá độ thì $1 - f_{ii} > 0$ và người ta chứng minh được kỳ vọng số lần quá trình quay lại i là $\frac{1}{1 - f_{ii}}$, là hữu hạn. Một công thức khác cho kỳ vọng này được lập như sau. Giả sử $X_0 = i$, để xét "số lần" ta đưa ra biến ngẫu nhiên mới

$$B_n = \begin{cases} 1 & \text{nếu } X_n = i, \\ 0 & \text{nếu } X_n \neq i. \end{cases}$$

Khi đó kì vọng số lần quay lại trạng thái i khi xuất phát từ i là

$$E\left(\sum_{n=1}^{\infty} B_n \mid X_0 = i\right) = \sum_{n=1}^{\infty} E(B_n \mid X_0 = i) = \sum_{n=1}^{\infty} P\{X_n = i \mid X_0 = i\} = \sum_{n=1}^{\infty} p_{ii}^{(n)}.$$

Từ đây cũng suy ra i là lặp khi và chỉ khi $\sum_{n=1}^{\infty} p_{ii}^{(n)} = +\infty$. Từ định nghĩa

cũng thấy mọi trạng thái thông nhau, tức là thuộc cùng lớp tương đương, thì cùng lặp hoặc cùng quá độ. Hơn nữa ở xích Markov hữu hạn trạng thái thì rõ ràng không thể mọi trạng thái đều quá độ (vì quá trình tiến đến vô hạn thì phải có trạng thái được qua vô hạn lần). Vì vậy, mọi trạng thái của xích Markov hữu hạn trạng thái không thu gọn được đều là trạng thái lặp.

Chu kì (period) của trạng thái i là số nguyên lớn nhất $t > 1$ sao cho $p_{ii}^{(n)} = 0$ với $\forall n \neq t, 2t, 3t, \dots$. Trạng thái có chu kì được gọi là *trạng thái tuần hoàn* (periodic state). Ngược lại, nếu có hai số nguyên liên nhau s và $s + 1$ để trạng thái i có thể xảy ra ở thời điểm s và $s + 1$ thì i được gọi là trạng thái *không tuần hoàn* (aperiodic), hoặc cũng nói là i có chu kì 1. Dễ thấy là mọi trạng thái trong cùng lớp có cùng chu kì.

Một trạng thái lặp i được gọi là *lặp dương* (positive recurrent) nếu kì vọng thời điểm quá trình quay lại i (lần đầu) là hữu hạn. Nếu kì vọng này là vô hạn thì ta nói trạng thái là *lặp không* (null recurrent). Có thể chứng minh được là ở xích Markov hữu hạn trạng thái, mọi trạng thái lặp là lặp dương.

Trạng thái lặp dương, không tuần hoàn được gọi là *trạng thái ergodic* (ergodic state). Nếu mọi trạng thái là ergodic thì ta nói là có *xích Markov ergodic* (ergodic Markov chain).

THÍ DỤ 9.1.2 (Một bài toán đầu tư). Một cửa hàng bán điện thoại di động có bán một loại máy và nhập hàng từng tuần vào tối chủ nhật để bán cả tuần. Với máy này, cửa hàng áp dụng *sách lược nhập hàng* (s, S) ((s, S) ordering policy): Nếu số máy còn chưa bán được vào chủ nhật là nhỏ hơn s , thì nhập S máy mới. Trái lại, nếu còn $\geq s$ máy thì không nhập mới. Để đơn giản tính toán lấy $s = 1, S = 3$. Chọn thời điểm t là cuối tuần thứ t và biến ngẫu nhiên X_t là số máy (loại đang xét) còn vào thời điểm t . Giả sử D_t là số máy khách muốn mua trong tuần thứ t . D_t là không biết nhưng qua thống kê biết được phân bố xác suất của D_t là như nhau cho mọi t . Khi đó không gian trạng thái gồm bốn trạng thái 0, 1, 2, 3. Đồng thời X_t phụ thuộc vào D_t như sau

$$X_{t+1} = \begin{cases} \max \{3 - D_{t+1}, 0\} & \text{nếu } X_t < 1, \\ \max \{X_t - D_{t+1}, 0\} & \text{nếu } X_t \geq 1, \end{cases} \quad (9.6)$$

và rõ ràng $\{X_t\}$ là xích Markov. Ta hãy tính ma trận truyền P , với giả thiết là mỗi D_t có phân bố Poisson với tham số $\mu = 1$. Phân bố Poisson rất phổ biến trong thực tế, nhất là ở các mô hình về "lượng đến" như số cuộc gọi đến một tổng đài điện thoại, số khách đến một cửa hàng... Ta nhắc lại kiến thức cơ bản về xác suất này.

Trước hết là về phân bố nhị thức. Nó được định nghĩa bởi ba điều cơ bản: có n phép thử độc lập; mỗi phép thử chỉ có hai kết quả (trạng thái) tạm gọi là sắp và ngửa; xác suất xảy ra sắp là π (ngửa là $1 - \pi$). Khi đó biến ngẫu nhiên

X là số lần sắp trong n phép thử có phân bố xác suất gọi là *phân bố nhị thức* (binomial distribution). Vì các phép thử độc lập, rõ ràng ta có phân bố nhị thức là

$$P\{X = r\} = C_n^r \pi^r (1 - \pi)^{n-r}$$

Sở dĩ gọi là "nhị thức" vì $P\{X = r\}$ chính là số hạng thứ $r+1$ trong khai triển nhị thức (đặt $q = 1 - \pi$):

$$1 = (\pi + q)^n = \sum_{r=0}^n C_n^r \pi^r q^{n-r}.$$

Vậy diễn nôm na thì phân bố nhị thức là phân bố của số lần sắp xảy đến. Kỳ vọng của biến có phân bố nhị thức là $n\pi$ sẽ dẫn đến ∞ nếu $n \rightarrow \infty$. Sở dĩ như vậy vì luôn bắt buộc sắp hoặc ngừng phải đến. Nếu "sự đến" là ngẫu nhiên của một đối tượng (không bắt buộc là một trong hai trạng thái phải đến), nhưng khá "chính quy" thì kỳ vọng với mọi n lớn, kể cả $n \rightarrow \infty$, là hằng số. Vì vậy coi là π giảm sao cho $n\pi = \mu$ là hằng. Ở giới hạn khi $n \rightarrow \infty$, biến ngẫu nhiên số lần sắp (nhưng trong vô hạn phép thử) nhận được lúc này được gọi là có *phân bố Poisson* (Poisson distribution). Người ta chứng minh được, với phân bố Poisson,

$$P\{X = r\} = \frac{\mu^r e^{-\mu}}{r!}, \quad r = 0, 1, 2, \dots \quad (9.7)$$

Phân bố Poisson đặc trưng bởi bốn điều: $\frac{P\{X = r+1\}}{P\{X = r\}} = \frac{\mu}{r+1}$; kỳ vọng $EX = \mu$; phương sai $VarX = \mu$; tổng các biến Poisson độc lập là biến Poisson có $E(X_1 + X_2) = EX_1 + EX_2$, $Var(X_1 + X_2) = VarX_1 + VarX_2$.

Khi số khách "đến" lại xét theo chỉ số thời gian là D_t thì ta được quá trình ngẫu nhiên. Quá trình này được gọi là có phân bố Poisson hoặc là *quá trình Poisson* (Poisson process) nếu

$$P(D_t = r) = \frac{(\mu t)^r e^{-\mu t}}{r!}.$$

Lúc này μ được gọi là tham số của quá trình. Ta cũng có $EX_t = \text{Var}X_t = \mu t$. μ sẽ là kì vọng số lần đến trong một đơn vị thời gian hoặc gọi là *tốc độ đến* (arrival rate).

Ở thí dụ này ta giả thiết là ở mỗi thời điểm t , D_t đều là biến ngẫu nhiên có phân bố Poisson có tham số $\mu = 1$. Ta quay lại việc tính ma trận truyền P của xích Markov $\{X_t\}$.

Ta có $p_{00} = P\{X_t = 0 \mid X_{t-1} = 0\}$. Theo (9.6), nếu $X_{t-1} = 0$ thì $X_t = \max\{3 - D_t, 0\}$. Muốn $X_t = 0$ thì D_t phải ≥ 3 . Vậy $p_{00} = P\{D_t \geq 3\} = 1 - P\{D_t \leq 2\} = 1 - \frac{1}{e} - \frac{1}{e} - \frac{1}{e \cdot 2!} \approx 0,080$ (theo (9.7) với $\mu = 1$).

Tiếp theo $p_{10} = P\{X_t = 0 \mid X_{t-1} = 1\} = P\{D_t \geq 1\} = 1 - P\{D_t = 0\} = 1 - \frac{1}{e} \approx 0,632$ (vì nếu $X_{t-1} = 1$ thì $X_t = \max\{1 - D_t, 0\} = 0$ suy ra $D_t \geq 1$).

Tương tự $p_{21} = P\{X_t = 1, X_{t-1} = 2\} = P\{D_t = 1\} \approx 0,368$ và ta được

$$P = \begin{bmatrix} 0,080 & 0,184 & 0,368 & 0,368 \\ 0,632 & 0,368 & 0 & 0 \\ 0,264 & 0,368 & 0,368 & 0 \\ 0,080 & 0,184 & 0,368 & 0,368 \end{bmatrix}$$

Suy ra

$$P^{(2)} = P \cdot P = \begin{bmatrix} 0,249 & 0,286 & 0,300 & 0,165 \\ 0,283 & 0,252 & 0,233 & 0,233 \\ 0,351 & 0,319 & 0,233 & 0,097 \\ 0,249 & 0,286 & 0,300 & 0,165 \end{bmatrix}$$

Giả sử lúc bắt đầu $t = 0$ có 3 máy di động ở cửa hàng và phải tính $P\{X_2 = 3\}$ tức là xác suất để vào tối chủ nhật sau hai tuần còn 3 máy.

Vì $P\{X_0 = 3\} = 1$, theo (9.5) ta có $P\{X_2 = 3\} = 1 \cdot p_{33}^{(2)} = 0,165$.

Ta thấy mọi $p_{ij}^{(2)} > 0$ nên xích Markov $\{X_t\}$ này là không thu gọn được, tức

là mọi trạng thái đều thông nhau. Do đó mọi trạng thái là trạng thái lặp. Vì chỉ có hữu hạn trạng thái nên từ đây suy ra mọi trạng thái là lặp dương. Vì $p_{ii} > 0$ và $p_{ij}^{(2)} > 0, \forall i$, nên ta có xích Markov ergodic.

9.1.4 Xác suất duy trì

Ở Thí dụ 9.1.1 ta đã thấy với xích Markov là quá trình thời tiết thì tồn tại xác suất duy trì, tức là xác suất để ngày n , với n lớn, là mưa (hoặc nắng) là 0,333 (hoặc 0,667) không phụ thuộc vào thời tiết hiện tại. Ở Thí dụ 9.1.2 về đầu tư điện thoại di động, nếu tính ma trận xác suất truyền 8 bước ta có

$$P^{(8)} = P^8 = \begin{bmatrix} 0,286 & 0,285 & 0,264 & 0,166 \\ 0,286 & 0,285 & 0,264 & 0,166 \\ 0,286 & 0,285 & 0,264 & 0,166 \\ 0,286 & 0,285 & 0,264 & 0,166 \end{bmatrix}$$

Ta thấy $p_{ij}^{(8)} = \pi_j > 0$ không phụ thuộc vào trạng thái i ban đầu. Tức là các xác suất duy trì cũng tồn tại. Người ta chứng minh được khẳng định sau về đáng điệu dài hạn của xích Markov.

ĐỊNH LÍ 9.1.1. *Xích Markov ergodic không thu gọn được luôn có các xác suất duy trì $\pi_j = \lim_{n \rightarrow \infty} p_{ij}^{(n)} > 0$ và là nghiệm duy nhất của hệ phương trình*

$$\pi_j = \sum_{i=0}^M \pi_i p_{ij}, \quad j = 0, 1, \dots, M, \quad (9.8)$$

$$\sum_{j=0}^M \pi_j = 1, \quad (9.9)$$

gọi là các **phương trình duy trì** (steady - state equations).

CHỨNG MINH. Ở đây ta không chứng minh sự tồn tại xác suất duy trì mà chỉ lập các phương trình duy trì. Qua giới hạn $n \rightarrow \infty$ ở (9.4) ta được ngay (9.8). Vì không gian trạng thái có $M+1$ trạng thái nên theo định nghĩa xác suất ta có (9.9). $\square$

Độc giả có thể xem chứng minh đầy đủ trong [Cinlar 1975], Định lý 2.1, Chương 6.

Hệ (9.8), (9.9) có $M+2$ phương trình, $M+1$ biến mà có nghiệm duy nhất, nên một trong $M+1$ phương trình đầu là thừa ((9.9) không thể bỏ đi vì $\forall \pi_j = 0$ thỏa (9.8)).

Ở Thí dụ 9.1.2, thay các p_{ij} vào (9.8) và giải (9.8), (9.9) ta sẽ được $\pi_0 = 0,285$, $\pi_1 = 0,285$, $\pi_2 = 0,264$, $\pi_3 = 0,166$, phù hợp với $P^{(8)}$ trên đây.

9.1.5 Kỳ vọng của cước phí trung bình

Giả sử ứng với mỗi trạng thái j có cước phí $C(j)$, $j = 0, \dots, M$. Do đó tại mỗi thời điểm t phải chịu cước phí $C(X_t)$. Đây là biến ngẫu nhiên có $M+1$ trạng thái $C(j)$.

ĐỊNH LÝ 9.1.2. *Xích Markov không thu gọn được và có mọi trạng thái là lặp dương sẽ có kỳ vọng của cước phí trung bình dài hạn (long-run expected average cost (per unit time)) tính được qua các π_j như sau*

$$\lim_{n \rightarrow \infty} E \left(\frac{1}{n} \sum_{t=1}^n C(X_t) \right) = \sum_{j=0}^M \pi_j C(j). \quad (9.10)$$

Ta không chứng minh mà chỉ giải thích. Xích Markov không thu gọn được với mọi trạng thái lặp dương thì chưa chắc có xác suất duy trì. Chẳng hạn xét ma trận chuyển (chỉ hai trạng thái) $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Khi đó $p_{00}^{(n)} = 1$ nếu n chẵn và $p_{00}^{(n)} = 0$ nếu n lẻ. Do đó không tồn tại

$\lim_{n \rightarrow \infty} p_{00}^{(n)}$. Nhưng người ta chứng minh được là với xích Markov thỏa giả thiết Định lí 9.1.2 thì ta có (không phụ thuộc vào i)

$$\lim_{n \rightarrow \infty} \left(\frac{1}{n} \sum_{k=1}^n p_{ij}^{(k)} \right) = \pi_j,$$

ở đây π_j là nghiệm các phương trình duy trì. Dùng giới hạn này sẽ chứng minh được định lý.

Một đại lượng dài hạn khác là *cước phí trung bình (dài hạn) thực tế* ((long-run) actual average cost (per unit time)) $\lim_{n \rightarrow \infty} \left(\frac{1}{n} \sum_{t=1}^n C(X_t) \right)$.

Người ta chứng minh được nó cũng bằng $\sum_{j=0}^M \pi_j C(j)$.

Xét Thí dụ 9.1.2 về đầu tư điện thoại. Giả sử j máy tồn đọng cuối tuần bị một cước phí bảo quản $C(j)$ là $C(0) = 0$, $C(1) = 2$, $C(2) = 8$, $C(3) = 18$. Khi đó kỳ vọng cước phí trung bình dài hạn và cước phí trung bình thực tế dài hạn là

$$\begin{aligned} \lim_{n \rightarrow \infty} E \left(\frac{1}{n} \sum_{t=1}^n C(X_t) \right) &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n C(X_t) \\ &= 0,286 \cdot 0 + 0,285 \cdot 2 + 0,264 \cdot 8 + 0,166 \cdot 18 = 5,67. \end{aligned}$$

9.1.6 Thí dụ áp dụng xích Markov

Một công ti điện thoại đang thì công lắp thiết bị đặc biệt ngoài trời mà trời mưa thì phải ngừng. Đồng thời phải điều nhân công trước, rồi nếu mưa thì báo hoãn nhưng phải trả cước phí 500 (ngàn đồng). Trái lại, nếu không điều nhân công trước mà thấy trời nắng mới thông báo gấp ngay đầu ngày thì phải trả cước phí 1500. Còn một ngày theo đúng kế hoạch thì phải trả 1000. Doanh thu của một ngày lao động là 3000. Tóm lại, lãi của công ti một ngày được cho bằng

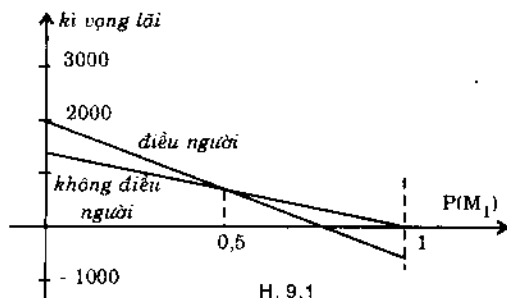
bảng sau đây.

Thời tiết	có điều người trước	không điều người trước
mưa	hoãn việc, lãi bằng - 500	không làm gì, lãi bằng 0
nắng	làm việc, lãi bằng $3000 - 1000 = 2000$	báo gấp, lãi = $3000 - 1500 = 1500$

Giả sử quy luật mưa nắng như ở Thí dụ 9.1.1. Ta xét ba trường hợp điển hình.

Trường hợp 1 (Đặt kế hoạch trước một ngày). Điều kiện thời tiết hôm nay (tức là ngày chọn quyết định) đã biết là M_0 hoặc N_0 . Theo Thí dụ 9.1.1, ta biết xác suất mưa của hôm sau là $P(M_1)$. kì vọng lãi của hôm sau là bảng sau đây

quyết định	kì vọng lãi
điều người	$-500 P(M_1) + 2000(1 - P(M_1))$
không điều người	$0 P(M_1) + 1500(1 - P(M_1))$



Đồ thị hàm kì vọng lãi theo $P(M_1)$ cho hai quyết định được vẽ ở H.9.1. Hai quyết định sẽ cho lãi như nhau tại giao điểm, ứng với $P(M_1) = 0,5$. Theo H. 9.1 nếu $P(M_1) < 0,5$ thì quyết định điều người trước là tối ưu. Theo Thí dụ 9.1.1, ta có kết quả phụ thuộc thời tiết hôm nay như sau.

thời tiết hôm nay	$P(M_1)$	quyết định tối ưu	kì vọng lãi
M_0	0,6	không điều người	$1500 \cdot 0,4 = 600$
N_0	0,2	điều người	$-500 \cdot 0,2 + 2000 \cdot 0,8 = 1500$

Trường hợp 2 (Đặt kế hoạch trước hai ngày). Lúc này xác suất mưa vào

ngày làm việc là $P(M_2)$ cũng đã có ở Thí dụ 9.1.1. Ta có bảng kết quả

thời tiết hôm nay	$P(M_2)$	quyết định tối ưu	kì vọng lãi
M_0	0,44	điều người	$-500.0,44 + 2000.0,56 = 900$
N_0	0,28	điều người	$-500.0,28 + 2000.0,72 = 1300$

Chú ý rằng H. 9.1 biểu diễn kì vọng lãi phụ thuộc quyết định và $P(M_1)$ vẫn giữ nguyên nếu thay $P(M_1)$ bởi $P(M_2)$ hay (PM_n) .

Trường hợp 3 (Đặt kế hoạch trước một tuần). Ta có xác suất mưa vào ngày làm việc là $P(M_1) \approx \frac{1}{3}$, không phụ thuộc thời tiết hôm nay là M_0 hay N_0 . Vì $\frac{1}{3} < \frac{1}{2}$ nên quyết định tối ưu là điều người trước. Kì vọng lãi là

$$-500 \cdot \frac{1}{3} + 2000 \cdot \frac{2}{3} = 1167.$$

Bây giờ ta so sánh ba trường hợp. Vì trong quá trình dài vào một ngày "hôm nay" nào đó ta có thể chọn quyết định theo một trong ba kiểu. Lúc đó $P(M_0)$ (và $P(N_0)$) là chưa biết nhưng có thể coi nó là xác suất duy trì, tức là $P(M_0) = \frac{1}{3}$. Do đó trung bình của kì vọng lãi cho ba trường hợp là :

$$\text{Kế hoạch trước 1 ngày : } 600 \cdot \frac{1}{3} + 1500 \cdot \frac{2}{3} = 1200,$$

$$\text{Kế hoạch trước 2 ngày : } 900 \cdot \frac{1}{3} + 1300 \cdot \frac{2}{3} = 1167,$$

Kế hoạch trước 1 tuần: 1167 (vì chỉ có một quyết định, không phụ thuộc M_0 hay N_0).

Như vậy ta thấy đặt kế hoạch trước chỉ một ngày là tốt nhất. Điều này dễ hiểu vì thông tin về thời tiết hôm nay cho biết khá nhiều về thời tiết ngày làm việc. Nhận xét này càng rõ với quá trình Markov vì chỉ có thời tiết hôm nay ảnh hưởng đến thời tiết tương lai. Chính vì vậy mà nếu tính tiếp trung bình của kì vọng lãi cho các trường hợp lập kế hoạch trước 3, 4, ... ngày ta thấy cùng kết

quả 1167 (vì $P(M_n)$ khi đó đều $< 0,5$).

Qua thí dụ này ta hình dung rằng có khá nhiều cách áp dụng xích Markov để chọn quyết định hành động. Áp dụng như vừa nêu là áp dụng "trực tiếp", không tác động vào làm thay đổi quá trình Markov. Ở phần còn lại của chương này ta sẽ xét các phương pháp chọn quyết định bằng cách tác động hẳn vào quá trình Markov, làm thay đổi ma trận truyền theo mục đích tối ưu đã chọn.

9.2 MÔ HÌNH QUÁ TRÌNH QUYẾT ĐỊNH MARKOV

Ở Mục 9.1 ta đã thấy nhiều quá trình có thể được mô hình hóa như xích Markov, khi coi diễn biến tương lai của quá trình chỉ phụ thuộc vào trạng thái hiện tại, không liên quan đến lịch sử của quá trình. Đây là tính chất Markov, đôi khi cũng gọi là *tính chất không nhớ* (lack-of-memory property). Với đặc trưng này xích Markov là mô hình khá khớp với nhiều quá trình thực tế. Bây giờ, thay vì chỉ nghiên cứu thụ động các tính chất của trạng thái của xích Markov cố định mô tả quá trình, ta tập trung vào thiết kế các sách lược để chọn quyết định thích hợp cho mỗi trạng thái để thay đổi tiến trình của xích Markov với mục đích làm cực trị mục tiêu (cực đại thu hoạch hoặc cực tiểu phí tổn).

9.2.1 Thí dụ mẫu

Một quá trình sản xuất có một máy chính luôn phải đặt ra vấn đề hao mòn và giảm chất lượng sản phẩm, vì cường độ hoạt động cao. Cuối mỗi tháng phải kiểm tra để xếp máy vào một trong bốn trạng thái: 0 là như mới, 1 là hao mòn nhẹ và hoạt động được, 2 là hao mòn lớn nhưng còn hoạt động được, 3 là không hoạt động được. Kết quả thống kê lâu dài cho thấy cứ cuối mỗi tháng khả năng (xác suất) chuyển

trạng thái của máy khi sang tháng sau là theo bảng dưới đây và nó không phụ thuộc vào quá khứ. Gọi X_t là trạng thái của máy vào cuối tháng t thì ta được một xích Markov với ma trận chuyển vừa nói. Vì $p_{33} = 1$ nên trạng thái 3 là hấp thụ, tức là máy đã vào trạng thái 3 thì luôn ở đó, không thoát được. Vì vậy phải thay để tiếp tục sản xuất. Phí tổn thay máy là 4 (triệu). Đồng thời mất một tuần tạm ngừng sản xuất phí tổn mất 2 (triệu). Ngoài ra máy kém cũng làm ra phế phẩm với phí tổn trung bình cho mỗi trạng thái trong một tháng là : trạng thái 0 có phí tổn 0 (triệu), 1 có phí tổn 1 và 2 có phí tổn 3.

	trạng thái	0	1	2	3
Máy có thể được thay trước khi đến trạng thái 3 hoặc có thể được đại tu với phí tổn 2 triệu. Đại tu chỉ có tác dụng cho trạng thái 2 để đưa máy về trạng thái 1 và	0	0	7/8	1/16	1/16
	1	0	3/4	1/8	1/8
	2	0	0	1/2	1/2
	3	0	0	0	1

ngừng sản xuất một tuần gây ra phí tổn 2 triệu.

Đặt tên *quyết định* (decision) như sau: quyết định 1 là không tác động gì, 2 là đại tu và 3 là thay máy. Phí tổn khi áp dụng các quyết định vào các trạng thái được cho ở Bảng 9.1.

quyết định	trạng thái	phí tổn do phế phẩm	phí tổn kỹ thuật	phí tổn ngừng sản xuất	tổng phí tổn
1	0	0	0	0	0
	1	1	0	0	1
	2	3	0	0	3
2	2	0	2	2	4
3	1, 2, 3	0	4	2	6

Bảng 9.1

Ta gọi *sách lược* (policy) là vectơ gồm thành phần là quyết định

sách lược	$d_0(R)$	$d_1(R)$	$d_2(R)$	$d_3(R)$	cho từng trạng thái. Ở thí dụ này,
R_a	1	1	1	3	sách lược là vectơ 4 chiều vì có bốn
R_b	1	1	2	3	trạng thái. Kí hiệu sách lược $R =$
R_c	1	1	3	3	$(d_0(R), d_1(R), d_2(R), d_3(R))$. Rõ ràng
R_d	1	3	3	3	là chỉ có bốn sách lược tất cả. Kí

Bảng 9.2

hiệu R_a là sách lược thay máy ở trạng thái 3 (và không làm gì ở các trạng thái khác), R_b là thay máy ở trạng thái 3 và đại tu ở trạng thái 2, R_c là thay máy ở trạng thái 2 và 3, R_d là thay máy ở 1, 2 và 3. Bốn sách lược được biểu diễn ở Bảng 9.2.

Mỗi sách lược tác động vào xích Markov làm thay đổi ma trận chuyển như sau

trạng thái	0	1	2	3	trạng thái	0	1	2	3
0	0	7/8	1/16	1/16	0	0	7/8	1/16	1/16
1	0	3/4	1/8	1/8	1	0	3/4	1/8	1/8
2	0	0	1/2	1/2	2	0	1	0	0
3	1	0	0	0	3	1	0	0	0

sách lược R_a sách lược R_b

trạng thái	0	1	2	3	trạng thái	0	1	2	3
0	0	7/8	1/16	1/16	0	0	7/8	1/16	1/16
1	0	3/4	1/8	1/8	1	1	0	0	0
2	1	0	0	0	2	1	0	0	0
3	1	0	0	0	3	1	0	0	0

sách lược R_c sách lược R_d

Phí tổn C_{ik} khi áp dụng quyết định k vào trạng thái i được ghi vào bảng sau, lấy từ Bảng 9.1

quyết định	C_{ik} (triệu đồng)		
trạng thái	1	2	3
0	0	-	-
1	1	-	6
2	3	4	6
3	-	-	6

Bảng 9.3

Một cách tự nhiên, coi mục tiêu của quá trình chọn quyết định là làm cực tiểu kì vọng của cước phí trung bình trong tháng, kì vọng theo dài hạn. Theo (9.10) kì vọng $E(C)$ này là

$$E(C) = \sum_{i=0}^M C_{ik} \pi_i, \quad (9.11)$$

ở đây C_{ik} là cước phí khi $d_i(R) = k$ với sách lược R . Mỗi sách lược có một kết quả $E(C)$. Vì chưa có cách giải ta dùng cách điểm diện tất cả bốn sách lược để chọn sách lược tối ưu. Hãy tính cụ thể cho R_a . Trước hết phải tính các xác suất duy trì từ các phương trình duy trì (9.8), (9.9). Theo ma trận truyền của sách lược R_a các phương trình duy trì là

$$\begin{aligned} \pi_0 &= \pi_3, & \pi_1 &= \frac{7}{8}\pi_0 + \frac{3}{4}\pi_1, \\ \pi_2 &= \frac{1}{16}\pi_0 + \frac{1}{8}\pi_1 + \frac{1}{2}\pi_2, & \pi_3 &= \frac{1}{16}\pi_0 + \frac{1}{8}\pi_1 + \frac{1}{2}\pi_2, \\ \pi_0 + \pi_1 + \pi_2 + \pi_3 &= 1. \end{aligned}$$

Giải ra ta được $\pi_0 = \frac{2}{13}$, $\pi_1 = \frac{7}{13}$, $\pi_2 = \frac{2}{13}$, $\pi_3 = \frac{2}{13}$. Với $R = R_a$ thì, theo Bảng 9.1 và 9.3, $C_{0k} = 0$, $C_{1k} = 1$, $C_{2k} = 3$, $C_{3k} = 6$ (C_{ik} là cước phí khi $d_i(R_a) = k$ trong trường hợp này). Thay vào (9.11) ta được $E(C) \approx 1,923$. Tương tự tính cho R_b , R_c và R_d ta được bảng sau đây.

sách lược	$(\pi_0, \pi_1, \pi_2, \pi_3)$	$E(C)$
R_a	$\left(\frac{2}{13}, \frac{7}{13}, \frac{2}{13}, \frac{2}{13}\right)$	$\frac{1}{13}(2.0 + 7.1 + 2.3 + 2.6) = \frac{25}{13} = 1,923$
R_b	$\left(\frac{2}{21}, \frac{5}{7}, \frac{2}{21}, \frac{2}{21}\right)$	$\frac{1}{21}(2.0 + 15.1 + 2.4 + 2.6) = \frac{35}{21} = 1,667$
R_c	$\left(\frac{2}{11}, \frac{7}{11}, \frac{1}{11}, \frac{1}{11}\right)$	$\frac{1}{11}(2.0 + 7.1 + 1.6 + 1.6) = \frac{19}{11} = 1,727$
R_d	$\left(\frac{1}{2}, \frac{7}{16}, \frac{1}{32}, \frac{1}{32}\right)$	$\frac{1}{32}(16.0 + 14.6 + 1.6 + 1.6) = \frac{96}{32} = 3,000$

Ta thấy sách lược R_b , tức là đại tu ở trạng thái 2 và thay máy ở trạng thái 3, là tối ưu với kì vọng của cước phí trung bình là 1,667 triệu đồng.

9.2.2 Mô hình quá trình quyết định Markov

1. Xích Markov (rời rạc) có $M+1$ trạng thái và có K quyết định $k = 1, \dots, K$ (một số quyết định chỉ áp dụng được cho một số trạng thái, không phải tất cả).
2. Quyết định $d_i = k$ áp dụng ở trạng thái i ấn định các xác truyền ở bước tới cho trạng thái j : $p_{ij}(k)$, $j = 0, \dots, M$. Bộ các quyết định cho $M+1$ trạng thái lập thành một sách lược cho quá trình quyết định Markov: $(d_0, d_1, \dots, d_M)$.
3. Có hai loại phí tổn. Quyết định $d_i = k$ áp vào trạng thái i sinh ra *cước phí tức thời* (immediate cost) với kì vọng C_{ik} . Ngoài ra còn có *cước phí sau* (subsequent cost) xảy ra trong tương lai của quá trình.
4. Mục tiêu là tìm sách lược tối ưu làm cực trị một tiêu chuẩn cước phí. Tương đối phổ biến là cần làm cực tiểu kì vọng cước phí trung bình (trong đơn vị thời gian, dài hạn).

Chú ý là trong mô hình sách lược là dừng, nó ấn định quyết định cho mỗi trạng thái không phụ thuộc thời điểm t . Đồng thời sách lược $(d_0, d_1, \dots, d_M)$ cũng là tất định, tức là nếu ở trạng thái i thì quyết định đã xác định hẳn là d_i .

Trong hai mục sau ta sẽ xét hai thuật toán tìm sách lược tối ưu cho quá trình quyết định Markov.

9.3 TÌM SÁCH LƯỢC TỐI ƯU BẰNG QUY HOẠCH TUYẾN TÍNH

Giả sử xích Markov luôn là không thu gọn được. Để tìm cách biểu diễn việc tìm sách lược tối ưu ở dạng quy hoạch tuyến tính trước hết ta biểu diễn lại sách lược R . Ở trên đã định nghĩa R qua bộ quyết định $(d_0(R), \dots, d_M(R))$. Bộ này lại viết tương đương ở dạng ma trận như sau.

$$\begin{array}{c} \text{trạng} \\ \text{thái} \end{array} \begin{array}{c} 0 \\ 1 \\ \vdots \\ M \end{array} \begin{array}{c} \text{quyết định} \\ \begin{array}{c} 1 \quad 2 \quad \dots \quad K \\ \left[\begin{array}{cccc} D_{01} & D_{02} & \dots & D_{0K} \\ D_{11} & D_{12} & \dots & D_{1K} \\ \cdot & \cdot & \cdot & \cdot \\ D_{M1} & D_{M2} & \dots & D_{MK} \end{array} \right] \end{array} \end{array} D_{ik} = \begin{cases} 1 & \text{nếu quyết định } k \text{ được} \\ & \text{chọn ở trạng thái } i, \\ 0 & \text{nếu ngược lại.} \end{cases}$$

Chú ý rằng mỗi hàng chỉ có đúng một số 1. Chẳng hạn sách lược tối ưu R_b ở cuối Mục 9.2.1 có dạng ma trận là

Như vậy bài toán quy về tìm ma trận các D_{ik} tối ưu. Nhưng

D_{ik} chỉ nhận giá trị 0 và 1, mà biến trong quy hoạch tuyến tính phải được nhận giá trị liên tục. Cách khắc phục là

dùng sách lược ngẫu nhiên hóa (randomized policy) như sau (tương tự như chiến lược hỗn hợp ở lý thuyết trò chơi trong Chương 7). Ta

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

xác định lại D_{ik} là xác suất có điều kiện

$$D_{ik} := P(\text{lấy quyết định } k \mid \text{trạng thái đã là } i).$$

Khi đó $0 \leq D_{ik} \leq 1$ và $\sum_{k=1}^K D_{ik} = 1$. Lúc này sách lược vẫn biểu diễn thành ma trận (D_{ik}) như trên, nhưng với D_{ik} là các xác suất có điều kiện và được lấy giá trị liên tục trong $[0,1]$.

Bây giờ ta xác định biến y_{ik} là xác suất không điều kiện

$$y_{ik} := P(\text{trạng thái là } i \text{ và quyết định là } k).$$

Theo công thức xác suất có điều kiện ta có

$$y_{ik} = P(\text{lấy quyết định } k \mid \text{trạng thái đã là } i) \cdot P(\text{trạng thái là } i)$$

$$= D_{ik} \pi_i.$$

$$\text{Do đó } \sum_{k=1}^K y_{ik} = \pi_i \sum_{k=1}^K D_{ik} = \pi_i. \text{ Vì vậy ta tính được } D_{ik} \text{ qua } y_{ik}$$

$$D_{ik} = \frac{y_{ik}}{\pi_i} = \frac{y_{ik}}{\sum_{l=1}^K y_{il}}.$$

Ta chọn biến để xây dựng quy hoạch tuyến tính là y_{ik} . Các ràng buộc do các π_i và D_{ik} phải thỏa sẽ là như sau.

$$1. \text{ Do (9.9) } \sum_{i=0}^M \pi_i = 1 \text{ nên } \sum_{i=0}^M \sum_{k=1}^K y_{ik} = 1.$$

$$2. \text{ Do phương trình duy trì (9.8) } \pi_j = \sum_{i=0}^M \pi_i p_{ij} \text{ nên}$$

$$\sum_{k=1}^K y_{jk} = \sum_{i=0}^M \sum_{k=1}^K y_{ik} p_{ij}(k), \quad j = 0, 1, \dots, M.$$

3. Do $\pi_i \geq 0$, $D_{ik} \geq 0$ nên $y_{ik} \geq 0$.

Mục tiêu là kì vọng cước phí trung bình dài hạn $E(C)$. Theo công thức (9.10), $E(C) = \sum_{i=0}^M \pi_i C(i)$, ở đây $C(i)$ là cước phí khi ở trạng thái

i. Lúc này quyết định k được chọn với xác suất D_{ik} . Theo mô hình quá trình quyết định Markov thì quyết định k áp vào trạng thái i bị cước phí là C_{ik} . Do đó $C(i) = \sum_{k=1}^K C_{ik} D_{ik}$. Vậy hàm mục tiêu là

$$Z = E(C) = \sum_{i=0}^M \pi_i \sum_{k=1}^K C_{ik} D_{ik} = \sum_{i=0}^M \sum_{k=1}^K C_{ik} y_{ik}.$$

Tóm lại, việc tìm sách lược tối ưu quy về quy hoạch tuyến tính

$$\min Z = \sum_{i=0}^M \sum_{k=1}^K C_{ik} y_{ik},$$

$$\sum_{i=0}^M \sum_{k=1}^K y_{ik} = 1,$$

$$\sum_{k=1}^K y_{jk} - \sum_{i=0}^M \sum_{k=1}^K y_{ik} p_{ij}(k) = 0, \quad j = 0, 1, \dots, M, \quad (9.12)$$

$$y_{ik} \geq 0, \quad i = 0, 1, \dots, M, \quad k = 1, \dots, K.$$

Quy hoạch có $K(M+1)$ biến y_{ik} và $M+2$ ràng buộc hàm (nhưng trong $M+1$ ràng buộc cuối có thể bỏ một ràng buộc thừa, cũng như (9.8)).

Khi giải bằng thuật toán đơn hình sẽ nhận được nghiệm tối ưu cơ sở, nên chỉ có cùng lắm là $M+1$ (bằng số ràng buộc hàm) biến cơ sở $y_{ik} > 0$, còn lại biến không cơ sở bằng 0. Người ta lại chứng minh được $\forall i = 0, \dots, M$ có một k để $y_{ik} > 0$. Vì vậy $y_{il} = 0, \forall i, \forall l \neq k$. Do

$$D_{ik} = \frac{y_{ik}}{\sum_{l=1}^K y_{il}} \quad \text{nên nếu đúng } k \text{ này thì } D_{ik} = \frac{y_{ik}}{y_{ik}} = 1, \text{ còn với chỉ số}$$

khác thì $D_{il} = 0, l \neq k$. Vậy sách lược tối ưu lại vẫn là tất định mặc dù quá trình giải ta đã ngẫu nhiên hóa. Thật là may.

THÍ DỤ 9.3.1 (Giải thí dụ mẫu 9.2.1 bằng cách đưa về quy hoạch tuyến tính). Theo Bảng 9.1 một số quyết định chỉ áp dụng cho một số trạng thái, theo đó chỉ cần các biến $y_{01}, y_{11}, y_{13}, y_{21}, y_{22}, y_{23}$ và y_{33} . Cột cước phí tổng cộng ở cuối bảng cho ta các hệ số hàm mục tiêu. Ràng buộc (9.12) là khá phức tạp, ở đây $p_{ij}(k)$ là xác suất truyền từ trạng thái i sang j khi áp quyết định k , được lấy ra dễ nhất là từ Bảng 9.1 và bốn ma trận truyền ứng với bốn sách lược $R_a, \dots, R_d$. Ta viết dần từng ràng buộc (9.12).

$$\text{Với } j = 0 \text{ ta có ràng buộc } y_{01} - \sum_{i=0}^3 \sum_{k=1}^3 y_{ik} p_{i0}(k) = 0. \text{ Ta phải lấy tổng với}$$

mọi trạng thái i và mọi quyết định k , nhưng xác suất dương để truyền từ i sang 0 chỉ có khi dùng quyết định 3, tức là $k = 3$. Với quyết định này cả ba trạng thái $i = 1, 2, 3$ đều có $p_{i0}(3) = 1$. Do đó thay vào ràng buộc ta được

$$y_{01} - (y_{13} \cdot 1 + y_{23} \cdot 1 + y_{33} \cdot 1) = 0.$$

$$\text{Với } j = 1 \text{ ta có ràng buộc } (y_{11} + y_{13}) - \sum_{i=0}^3 \sum_{k=1}^3 y_{ik} p_{i1}(k) = 0. \text{ Ta xét từng } i$$

để viết tổng cuối cùng. Với $i = 0, p_{01}(k) = \frac{7}{8}$ với $k = 1$ vì ở trạng thái 0 chỉ có xác suất chuyển sang trạng thái 1 là $\frac{7}{8}$ với quyết định 1, còn quyết định k khác

không áp dụng cho trạng thái 0. Với $i = 1$ chỉ có $p_{11}(1) = \frac{3}{4}$ còn với k khác thì

$p_{11}(3) = 0$ còn $k = 2$ không áp dụng cho trạng thái 1 nên không có $p_{11}(2)$. Tương tự với $i = 2$ chỉ có $p_{21}(2) = 1$ tức là với $k = 2$, còn với $i = 3$ không có $p_{31}(k)$ nào dương. Vậy với $j = 1$ ràng buộc là

$$y_{11} + y_{13} - \left(\frac{7}{8}y_{01} + \frac{3}{4}y_{11} + 1y_{22} \right) = 0.$$

Làm tương tự với $j = 2$ và $j = 3$. Cuối cùng ta được quy hoạch tuyến tính

$$\min Z = y_{11} + 6y_{13} + 3y_{21} + 4y_{22} + 6y_{23} + 6y_{33},$$

$$y_{01} + y_{11} + y_{13} + y_{21} + y_{22} + y_{23} + y_{33} = 1,$$

$$y_{01} - (y_{13} + y_{23} + y_{33}) = 0,$$

$$y_{11} + y_{13} - \left(\frac{7}{8}y_{01} + \frac{3}{4}y_{11} + y_{22} \right) = 0,$$

$$y_{21} + y_{22} + y_{23} - \left(\frac{1}{16}y_{01} + \frac{1}{8}y_{11} + \frac{1}{2}y_{21} \right) = 0,$$

$$y_{33} - \left(\frac{1}{16}y_{01} + \frac{1}{8}y_{11} + \frac{1}{2}y_{21} \right) = 0,$$

$$\text{mọi } y_{ik} \geq 0.$$

Dùng thuật toán đơn hình ta được nghiệm tối ưu $y_{01} = \frac{2}{21}$, $(y_{11}, y_{13}) = \left(\frac{5}{7}, 0 \right)$, $(y_{21}, y_{22}, y_{23}) = \left(0, \frac{2}{21}, 0 \right)$, $y_{33} = \frac{2}{21}$. Suy ra $D_{01} = 1$, $(D_{11}, D_{13}) = (1, 0)$, $(D_{21}, D_{22}, D_{23}) = (0, 1, 0)$, $D_{33} = 1$. Vậy sách lược R_b là tối ưu, trùng với kết quả giải bằng cách điểm diện toàn bộ ở Mục 9.2.1.

9.4 THUẬT TOÁN CẢI THIẾN SÁCH LƯỢC

9.4.1 Tính toán sơ bộ

Gọi $v_i^n(R)$ là kì vọng tổng cước phí của hệ qua n bước bắt đầu từ trạng thái i và theo sách lược R . Khi đó

$$v_i^n(R) = C_{ik} + \sum_{j=0}^M p_{ij}(k) v_j^{n-1}(R), \quad i = 0, 1, \dots, M, \quad (9.13)$$

ở đây C_{ik} là kì vọng cước phí ở ngay bước đầu, và tổng sau đó là tổng kì vọng cước phí của hệ qua $n-1$ bước còn lại. Ta có $v_i^1(R) = C_{ik}$ với mọi i . Gọi $g(R)$ là kì vọng cước phí trung bình một đơn vị thời gian (dài hạn) ứng với sách lược R và $d_i(R) = k$ thì đã biết

$$g(R) = \sum_{i=0}^M \pi_i C_{ik},$$

và $g(R)$ không phụ thuộc trạng thái xuất phát i . Nếu bỏ qua sai số nhỏ thì có thể coi

$$v_i^n(R) = ng(R) + v_i(R),$$

ở đây $v_i(R)$ là thành phần của tổng cước phí phụ thuộc vào trạng thái đầu. Thay biểu thức $v_i^n(R)$ này và $v_j^{n-1}(R) = (n-1)g(R) + v_j(R)$ vào công thức truy hồi (9.13) ta được

$$g(R) = C_{ik} + \sum_{j=0}^M p_{ij}(k) v_j(R) - v_i(R), \quad i = 0, 1, \dots, M. \quad (9.14)$$

Đây là hệ $M+1$ phương trình của $M+2$ ẩn $g(R), v_i(R)$. Một ẩn được lấy tùy ý ta chọn là $v_M(R) = 0$. Giải hệ này cho mỗi sách lược R rồi chọn sách lược có $g(R)$ nhỏ nhất là sách lược tối ưu. Tuy nhiên ta có thuật toán sau thường cho sách lược tối ưu sau ít bước hơn dùng quy hoạch tuyến tính khá nhiều.

9.4.2 Thuật toán cải thiện sách lược

Khởi tạo: Chọn sách lược R_1 bất kì. Đặt $n = 1$.

Bước lặp n (tổng quát) gồm hai đoạn.

Đoạn 1 (tính giá trị). Dùng $p_{ij}(k)$, C_{ik} và $v_M(R_n) = 0$ giải hệ $M+1$ phương trình (9.14) cho $g(R_n)$, $v_0(R_n)$, ..., $v_{M-1}(R_n)$.

Đoạn 2 (cải thiện sách lược). Tìm sách lược R_{n+1} sao cho $d_i(R_{n+1}) = k$ đạt cực tiểu sau đây cho mọi i

$$\min_{k=1, \dots, K} \left[C_{ik} + \sum_{j=0}^M p_{ij}(k) v_j(R_n) - v_i(R_n) \right] \quad (9.15)$$

Kiểm tra tính tối ưu: Nếu $R_{n+1} = R_n$ thì dừng. R_{n+1} là tối ưu. Nếu chưa có điều này thì đặt $n := n+1$ và sang bước lặp mới.

THÍ DỤ 9.4.1. Ta giải thí dụ mẫu 9.2.1 bằng thuật toán cải thiện sách lược, với $R_1 = R_a = (d_0(R_a), \dots, d_3(R_a)) = (1, 1, 1, 3)$.

Bước lặp 1. Đoạn 1. Lấy $v_3(R_1) = 0$. Giải hệ

$$g(R_1) = \quad + \frac{7}{8}v_1(R_1) + \frac{1}{16}v_2(R_1) - v_0(R_1),$$

$$g(R_1) = 1 \quad + \frac{3}{4}v_1(R_1) + \frac{1}{8}v_2(R_1) - v_1(R_1),$$

$$g(R_1) = 3 \quad + \frac{1}{2}v_2(R_1) - v_2(R_1),$$

$$g(R_1) = 6 + v_0(R_1),$$

$$\text{ta được } g(R_1) = \frac{25}{13} = 1,923, \quad v_0(R_1) = -\frac{53}{13} = -4,077, \quad v_1(R_1) = -\frac{34}{13} = -2,615,$$

$$v_2(R_1) = \frac{28}{13} = 2,154.$$

Đoạn 2. Chọn R_2 để quyết định k ở mọi trạng thái i đạt cực tiểu (9.15). Chú ý rằng ở trạng thái 0 chỉ dùng được quyết định 1 và ở 3 chỉ dùng quyết định 3. Ta chỉ viết (9.15) cho hai trạng thái

trạng thái 1 : $C_{1k} - p_{10}(k).4,077 - p_{11}(k).2,615 + p_{12}(k).2,154 + 2,615$,

trạng thái 2 : $C_{2k} - p_{20}(k).4,077 - p_{21}(k).2,615 + p_{22}(k).2,154 - 2,154$.

Ở trạng thái 1 có hai quyết định 1 và 3. Ta có bảng sau

quyết định	trạng thái 1					giá trị biểu thức (9.15)
	C_{1k}	$p_{10}(k)$	$p_{11}(k)$	$p_{12}(k)$	$p_{13}(k)$	
1	1	0	3/4	1/8	1/8	1,923 (min)
3	6	1	0	0	0	4,538

Vậy quyết định 1 làm cực tiểu biểu thức cần xét, nên $d_1(R_2) = 1$.

Bảng cho trạng thái 2 là

quyết định	trạng thái 2					giá trị biểu thức (9.15)
	C_{2k}	$p_{20}(k)$	$p_{21}(k)$	$p_{22}(k)$	$p_{23}(k)$	
1	3	0	0	1/2	1/2	1,923
2	4	0	1	0	0	-769 (min)
3	6	1	0	0	0	-231

Vậy $d_2(R_2) = 2$. Ta ghi lại R_2 , ma trận truyền và cước phí tương ứng :

sách lược R ₂		ma trận truyền				cước phí		
trạng thái	quyết định	trạng thái	0	1	2	3	trạng thái	C _{ik}
0	1	0	0	7/8	1/16	1/16	0	0
1	1	1	0	3/4	1/8	1/8	1	1
2	2	2	0	1	0	0	2	4
3	3	3	1	0	0	0	3	6

Bước lặp 2. Đoạn 1. Hệ phương trình cần giải ứng với R_2 là

$$g(R_2) = + \frac{7}{8}v_1(R_2) + \frac{1}{16}v_2(R_2) - v_0(R_2),$$

$$g(R_2) = 1 + \frac{3}{4}v_1(R_2) + \frac{1}{8}v_2(R_2) - v_1(R_2),$$

$$g(R_2) = 4 + v_1(R_2) - v_2(R_2),$$

$$g(R_2) = 6 + v_0(R_2).$$

Nghiệm là $g(R_2) = 1,667$, $v_0(R_2) = -4,333$, $v_1(R_2) = -3$, $v_2(R_2) = -0,667$.

Đoạn 2. Ta viết (9.15) cho hai trạng thái có hơn một quyết định

$$\text{trạng thái 1: } C_{1k} - p_{10}(k).4,333 - p_{11}(k).3 - p_{12}(k).0,667 + 3,$$

$$\text{trạng thái 2: } C_{2k} - p_{20}(k).4,333 - p_{21}(k).3 - p_{22}(k).0,667 + 0,667.$$

Khi lập bảng cho hai trạng thái này thì các cột với C_{1k} , C_{2k} , $p_{10}(k)$, ..., $p_{23}(k)$ đã có ở bước lập 1, ta chỉ phải tính cột cho giá trị biểu thức (9.15) như sau

quyết định	giá trị (9.15) cho trạng thái 1	giá trị (9.15) cho trạng thái 2
1	1,667 (min)	3,333
2	-	1,667 (min)
3	4,667	2,334

Vậy sách lược $R_3 = (d_0, d_1, d_2, d_3) = (1, 1, 2, 3) = R_2$. Do đó $R_3 = R_b$ là tối ưu.

9.5 TIÊU CHUẨN VỚI CƯỚC PHÍ GIẢM :

Trong các bài toán thực tế, khi quá trình Markov khá dài thì cước phí có thể phải hiệu chỉnh bởi một *hệ số giảm* (discount factor) $\alpha \in (0,1)$, chẳng hạn do giá trị thời gian của tiền nên cần bổ sung cước vào các giai đoạn tương lai. Lúc này ta vẫn áp dụng hai phương pháp giải như ở Mục 9.3 và 9.4 nhưng có các hiệu chỉnh.

9.5.1 Dừng quy hoạch tuyến tính với tiêu chuẩn cước phí giảm

Ngoài hệ số giảm α , mô hình có thêm các hằng β_j , $j = 0, 1, \dots, M$, thỏa $\beta_j > 0$, $\sum_{j=0}^M \beta_j = 1$. Quy hoạch tuyến tính sẽ là (không có ràng buộc đầu như trước)

$$\begin{aligned} \min Z &= \sum_{i=0}^M \sum_{k=1}^K C_{ik} y_{ik}, \\ \sum_{k=1}^K y_{ik} - \alpha \sum_{i=0}^M \sum_{k=1}^K y_{ik} p_{ij}(k) &= \beta_j, \quad j = 0, 1, \dots, M, \\ y_{ik} &\geq 0, \quad i = 0, 1, \dots, M, \quad k = 1, \dots, K. \end{aligned}$$

Áp dụng thuật toán đơn hình tìm được nghiệm tối ưu (y_{ik}) xong thì sách lược tối ưu ngẫu nhiên hóa sẽ là

$$D_{ik} = P \{ \text{quyết định } k \mid \text{trạng thái } i \} = \frac{y_{ik}}{\sum_{l=1}^K y_{il}}.$$

Người ta cũng chứng minh được trong trường hợp này sách lược tối ưu cũng vẫn là tất định, tức là $D_{ik} = 0$ hoặc 1. Hơn nữa lại không cần giả thiết xích Markov là không thu gọn được như ở Mục 9.3.

THÍ DỤ 9.5.1. Giải thí dụ mẫu 9.2.1 với $\alpha = 0.9$ và $\beta_j = \frac{1}{4}$, $\forall j$. Khi đó quy hoạch tuyến tính tương ứng với quá trình quyết định Markov là

$$\begin{aligned} \min Z &= y_{11} + 6y_{13} + 3y_{21} + 4y_{22} + 6y_{23} + 6y_{33}, \\ y_{01} - 0.9(y_{13} + y_{23} + y_{33}) &= \frac{1}{4}, \\ y_{11} + y_{13} - 0.9\left(\frac{7}{8}y_{01} + \frac{3}{4}y_{11} + y_{22}\right) &= \frac{1}{4}, \end{aligned}$$

$$y_{21} + y_{22} + y_{23} - 0,9 \left(\frac{1}{16} y_{01} + \frac{1}{8} y_{11} + \frac{1}{2} y_{21} \right) = \frac{1}{4},$$

$$y_{33} - 0,9 \left(\frac{1}{16} y_{01} + \frac{1}{8} y_{11} + \frac{1}{2} y_{21} \right) = \frac{1}{4},$$

$$\text{mọi } y_{ik} \geq 0.$$

Dùng thuật toán đơn hình ta được nghiệm tối ưu $y_{01} = 1,210$, $y_{11} = 6,656$, $y_{13} = 0$, $y_{21} = y_{23} = 0$, $y_{22} = 1,067$, $y_{33} = 1,067$. Do đó $D_{01} = D_{11} = D_{22} = D_{33} = 1$, $D_{13} = D_{21} = D_{23} = 0$, tức là sách lược tối ưu vẫn như ở Thí dụ 9.3.1 (do tác động của α và β_j chưa đến mức thay đổi). Nhưng mục tiêu tối ưu là $Z = 17,325$ đã khác trước.

9.5.2 Thuật toán cải thiện sách lược với tiêu chuẩn cước phí giảm

Tương tự ở Mục 9.4.1, gọi $V_i^n(R)$ là kì vọng tổng cước phí có giảm qua n bước, bắt đầu từ trạng thái i và theo sách lược R thì, tương tự (9.13), ta có

$$V_i^n(R) = C_{ik} + \alpha \sum_{j=0}^M p_{ij}(k) V_j^{n-1}(R), \quad i = 0, 1, \dots, M,$$

với $V_i^1(R) = C_{ik}$. Khi n dần đến ∞ thì phương trình truy hồi này dần đến

$$V_i(R) = C_{ik} + \alpha \sum_{j=0}^M p_{ij}(k) V_j(R), \quad i = 0, 1, \dots, M, \quad (9.16)$$

Thuật toán. Khởi tạo: chọn R_1 tùy ý. Đặt $n = 1$.

Bước lặp n (tổng quát). Đoạn 1 (tính giá trị). Thay R_n vào (9.16) và giải ra $V_0(R_n), \dots, V_M(R_n)$.

Đoạn 2 (cải thiện sách lược). Tìm R_{n+1} sao cho $\forall i, d_i(R_{n+1}) = k$ là quyết định đạt được cực tiểu

$$\min_{k=1, \dots, K} \left[C_{ik} + \alpha \sum_{j=0}^M p_{ij}(k) V_j(R_n) \right].$$

Kiểm tra tính tối ưu. Nếu $R_{n+1} = R_n$ thì dừng. R_{n+1} là tối ưu.

9.6 PHƯƠNG PHÁP XẤP XỈ LIÊN TIẾP

Ý tưởng của *phương pháp xấp xỉ liên tiếp* (method of successive approximation) là ở mỗi bước lặp $n = 1, 2, \dots$ tìm sách lược tối ưu cho quá trình quyết định Markov chỉ có đúng n bước, tức là một *quá trình quyết định Markov hữu hạn bước* (finite - period Markov decision process) bằng phương trình Bellman. Khi $n \rightarrow \infty$ các sách lược tối ưu này sẽ dẫn đến sách lược tối ưu cho quá trình quyết định Markov vô hạn. Phương pháp này ưu việt hơn hai phương pháp trình bày trên đây vì nó đơn giản và nhanh. Ở mỗi bước không cần giải hệ phương trình. Nếu quá trình có n bước thì sau đúng n bước ta có sách lược tối ưu. Nếu quá trình là vô hạn thì ta chỉ được xấp xỉ, nhưng sự hội tụ khá nhanh.

Gọi V_i^n là kì vọng tổng cước phí có giảm cho n bước còn lại (bắt đầu từ i) theo sách lược tối ưu. Theo phương trình Bellman, trong quy hoạch động (xem Mục 3.7), V_i^n sẽ nhận được theo hệ thức truy hồi

$$V_i^n = \min_k \left[C_{ik} + \alpha \sum_{j=0}^M p_{ij}(k) V_j^{n-1} \right], \quad i = 0, 1, \dots, M.$$

Giá trị k đạt cực tiểu sẽ là quyết định tối ưu ở bước đầu khi quá trình đi từ i . Khi xuất phát thuật toán, với $n = 1$, vì mọi $V_i^0 = 0$ nên

$$V_i^1 = \min_k C_{ik}, \quad i = 0, 1, \dots, M.$$

THÍ DỤ 9.6.1. Ta giải Thí dụ mẫu 9.2.1 với $\alpha = 0,9$ bằng phương pháp xấp xỉ liên tiếp. Chú ý là các C_{ik} có ở Bảng 9.3.

Bước lặp 1.

$$V_0^1 = \min_k C_{0k} = 0, \text{ ứng với } k = 1;$$

$$V_1^1 = \min_k C_{1k} = 1, \text{ ứng với } k = 1;$$

$$V_2^1 = \min_k C_{2k} = 3, \text{ ứng với } k = 1;$$

$$V_3^1 = \min_k C_{3k} = 6, \text{ ứng với } k = 3.$$

Vậy sách lược tối ưu xấp xỉ đầu tiên là $R_1 = (d_0, d_1, d_2, d_3) = (1, 1, 1, 3)$.

Bước lặp 2.

$$\begin{aligned} V_0^2 &= \min \left\{ 0 + 0,9 \left[\frac{7}{8} \cdot 1 + \frac{1}{16} \cdot 3 + \frac{1}{16} \cdot 6 \right], 4 + 0,9 [1.1], 6 + 0,9 [1.0] \right\} \\ &= 1,294, \text{ ứng với } k = 1. \end{aligned}$$

$$\begin{aligned} V_1^2 &= \min \left\{ 1 + 0,9 \left[\frac{3}{4} \cdot 1 + \frac{1}{8} \cdot 3 + \frac{1}{8} \cdot 6 \right], 4 + 0,9 [1.1], 6 + 0,9 [1.0] \right\} \\ &= 2,688, \text{ ứng với } k = 1; \end{aligned}$$

$$V_2^2 = \min \left\{ 3 + 0,9 \left[\frac{1}{2} \cdot 3 + \frac{1}{2} \cdot 6 \right], 4 + 0,9 [1.1], 6 + 0,9 [1.0] \right\}$$

$= 4,900$, ứng với $k = 2$;

$V_3^2 = 6 + 0.9 [1.0] = 6,000$, ứng với $k = 3$.

Ta được sách lược tối ưu xấp xỉ thứ hai là $R_2 = (1, 1, 2, 3)$. Sách lược này đã trùng với sách lược tối ưu cho quá trình vô hạn nhận được ở Thí dụ 9.5.1, nhưng nếu không giải bằng cách khác (như bằng quy hoạch tuyến tính ở Thí dụ 9.5.1) thì ta không nhận biết được.

Nếu làm tiếp bước lặp 3 thì ta thấy $R_3 = R_2$ (nhưng các $V_i^3 \neq V_i^2$).

BÀI TẬP CHƯƠNG 9

9.1 Xác suất để một người có điện thoại di động mua máy mới theo số năm sử dụng máy cũ được cho ở bảng sau

số năm đã sử dụng	0	1	2	3	4
xác suất thay máy	0,1	0,15	0,22	0,5	1

(đây là xác suất có điều kiện khi đến năm hiện hành máy chưa thay). Hãy biểu diễn thành xích Markov với trạng thái là số năm đã dùng máy. Lập ma trận chuyển. Tính xác suất duy trì.

9.2 (Mô hình giá cổ phiếu ở thị trường chứng khoán). Giả sử cho đơn giản là giá một loại cổ phiếu chỉ có năm mức 1, ..., 5 (triệu đồng) và chỉ tính một lần một ngày (thực tế có nhiều mức và thời điểm gần như xét liên tục). Một người kinh doanh cổ phiếu coi giá X_t tại ngày t là ngẫu nhiên, nhưng tạo thành một xích Markov, để tính việc mua nếu dự đoán giá sẽ tăng và bán nếu dự đoán giảm. Giả sử qua kinh nghiệm đã có ma trận chuyển sau đây.

	1	2	3	4	5
1	0,6	0,2	0,1	0,05	0,05
2	0,25	0,55	0,1	0,05	0,05

3	0,1	0,15	0,5	0,15	0,1
4	0,05	0,05	0,1	0,55	0,25
5	0,05	0,05	0,1	0,2	0,6

(a) Hãy tính kì vọng giá ngày mai và kì vọng lãi nếu mua hôm nay và bán ngày mai cho cả năm trường hợp giá hôm nay.

(b) Tính kì vọng giá và kì vọng lãi nếu mua hôm nay và bán sau hai ngày, cho cả năm trường hợp giá hôm nay.

9.3 Một chất điểm chuyển động trên đường tròn và chỉ dừng ở năm vị trí 0, ..., 4 xếp theo kim đồng hồ, xuất phát ở 0. Ở mỗi bước xác suất chuyển một vị trí theo kim là 0,5 và ngược kim là 0,5. Gọi X_t là vị trí của chất điểm ở bước t và giả sử $\{X_t\}$ là xích Markov.

(a) Tính ma trận chuyển P , $P^{(2)}$ và $P^{(5)}$.

(b) Tính các xác suất duy trì.

9.4 Biết ma trận chuyển của xích Markov. Xác định các lớp tương đương và xét xem lớp là lặp (recurrent) hay không.

$$(a) P = \begin{bmatrix} 0 & 0 & \frac{1}{3} & \frac{2}{3} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

$$(b) P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}$$

9.5 Ma trận chuyển P sẽ được gọi là ngẫu nhiên đúng nếu tổng các phần tử của mỗi cột đều bằng 1. Giả thiết thêm nữa là xích Markov không tuần hoàn, không thu gọn được và gồm $M + 1$ trạng thái. Chứng minh rằng $\pi_j = \frac{1}{M+1}$, $j = 0, \dots, M$.

9.6 Xét bài toán đầu tư ở Thí dụ 9.1.2 nhưng dùng sách lược (s, S) với $s = 2$, $S = 2$.

(a) Tính các xác suất duy trì.

(b) Tính kỳ vọng của cước phí trung bình dài hạn.

9.7 Một người mướn lái xe con đến hãng hàng ngày và rất chú ý giữ xe, cô có ba lựa chọn đầu xe. Ở bãi giữ mất về 5 (ngàn đồng) và không bị trầy xước, chiếm một chỗ

ngoài đường thì xác suất trầy xước là 0,1, chiếm bữa hai chỗ miễn phí thì xác suất trầy xước là 0,02 nhưng xác suất bị phạt 15 ngàn là 0,3. Nếu xe bị trầy xước thì có thể chữa mất 50 ngàn. Nếu cứ đi xe không sửa thì coi là mất thể diện với giá trị 9 ngàn một ngày. Cô muốn chọn sách lược tối ưu về chỗ đậu xe và có chữa xe không khi hỏng.

(a) Định nghĩa trạng thái để phát biểu quá trình quyết định Markov. Tính các C_{ik} .

(b) Xác định các sách lược (dừng, tắt định). Với mỗi sách lược, tìm ma trận truyền và kỳ vọng cước phí trung bình dài hạn qua xác suất duy trì $(\pi_0, \dots, \pi_M)$ chưa biết.

(c) Xác định $(\pi_0, \dots, \pi_M)$ cho mỗi sách lược. Tính các giá trị ở (b) để tìm sách lược tối ưu bằng cách điểm diện toàn bộ.

9.8 Người chơi quần vợt được giao bóng hai lần nếu hỏng cả thì (cước phí) là mất 1 (điểm). Nếu anh cố mạnh hết sức (để ăn ngay) thì xác suất bóng không hỏng là $\frac{3}{8}$, nếu đánh dăm bảo thôi thì xác suất là $\frac{7}{8}$. Nếu đánh mạnh mà không hỏng thì xác suất thắng điểm là $\frac{2}{3}$, nếu đánh dăm bảo mà vào sân thì xác suất được điểm là $\frac{1}{3}$. Ba yêu cầu (a), (b), (c) như Bài 9.7.

Gợi ý : trạng thái 0 là giao bóng xong và có điểm (1 hoặc -1), trạng thái 1 là còn một lần giao bóng.

9.9 Một công ti điện thoại di động xếp mức bán máy mỗi quý ở hai mức cao và thấp. Mức này có phụ thuộc việc có quảng cáo hay không vào quý trước. Giả sử đầu mỗi quý sẽ xác định được chính xác mức bán là cao hay thấp ở quý này và phải quyết định có quảng cáo hay không ở quý này. Cước phí quảng cáo một quý là 100 (triệu đồng). Nếu có quảng cáo thì xác suất mức bán quý sau cao là $\frac{1}{2}$ và $\frac{3}{4}$ tùy mức bán quý này là cao hay thấp. Nếu không quảng cáo thì hai xác suất này là $\frac{1}{4}$ và $\frac{1}{2}$. Lãi một quý (chưa trừ cước quảng cáo) là 400 nếu mức bán cao và 200 nếu thấp. Hãy xác định sách lược quảng cáo làm cực đại kỳ vọng lãi (trừ cước quảng cáo) trung bình dài hạn. Ba yêu cầu (a), (b), (c) như Bài 9.7.

9.10 (a) Phát biểu mô hình quy hoạch tuyến tính cho Bài 9.7.

(b) Dùng thuật toán đơn hình để giải và suy ra sách lược tối ưu.

9.11 Yêu cầu như Bài 9.10 cho bài toán ở Bài 9.8.

9.12. Yêu cầu như Bài 9.10 cho bài toán ở Bài 9.9.

9.13. Dùng thuật toán cải thiện sách lược tìm sách lược tối ưu cho Bài 9.7.

9.14. Yêu cầu như Bài 9.13 cho bài toán ở Bài 9.8.

9.15. Yêu cầu như Bài 9.13 cho bài toán ở Bài 9.9.

9.16. Một người muốn bán nhà. Giả sử mô hình là mỗi tháng có thể có ba người trả giá 600 (triệu đồng), 800 và 1000 với xác suất $\frac{5}{8}$, $\frac{1}{4}$ và $\frac{1}{8}$. Người bán phải quyết định ngay là đồng ý không. Nếu không thì người mua không quay lại. Vì cước phí bảo trì nhà là 60 nên người bán muốn nhanh và chọn hệ số giảm giá $\alpha = 0,95$. Hãy dùng thuật toán cải thiện sách lược tìm sách lược làm cực tiểu kỳ vọng tổng cước phí có giảm. (Gợi ý : ba trạng thái là ba mức giá đề nghị và trạng thái ∞ . Trạng thái ∞ xảy ra khi một giá đề nghị được chấp nhận và quá trình ở lại trạng thái này với cước phí bảo trì 0. Có hai quyết định đồng ý và không.)

9.17. (a) Phát biểu thành quy hoạch tuyến tính cho Bài 9.16.

(b) Giải bằng thuật toán đơn hình và suy ra sách lược tối ưu.

9.18. Tính ba bước lặp theo phương pháp xấp xỉ liên tiếp cho bài toán ở Bài 9.16.

9.19. Xét thí dụ mẫu 9.2.1. Bây giờ giả sử quá trình dùng máy để sản xuất chỉ còn 4 tháng nữa. Sử dụng tiêu chuẩn cước phí giảm với $\alpha = 0,9$ tìm sách lược tối ưu cho bài toán 4 bước.

CHƯƠNG 10**HỆ THỐNG DỊCH VỤ (LÝ THUYẾT XẾP HÀNG)**

Các hệ thống dịch vụ (service system) theo nghĩa rộng, bao gồm rất nhiều hoạt động như: thương mại (các cửa hàng), dịch vụ công cộng (viễn thông, bưu điện, khu vui chơi giải trí), giao thông (trạm xăng, trạm thu phí giao thông), y tế (phòng khám, phòng cấp cứu) ... là một trong các đối tượng sớm nhất của toán ứng dụng nói chung và vận trù học nói riêng. Dù mục đích là lợi nhuận hay phục vụ phi lợi nhuận thì nhiều vấn đề đặt ra được qui về việc thiết kế và điều hành hệ sao cho việc "xếp hàng" chờ đợi được giảm thiểu. Nếu đầu tư ít phương tiện và người phục vụ thì khách phải "xếp hàng" lâu, quy ra tổn thất về tiền (bán được ít hàng, khách sang cửa hàng khác nếu dịch vụ là thương mại hoặc thiệt hại khác nếu là phục vụ công cộng, thậm chí cả sinh mạng như ở dịch vụ y tế). Ngược lại quá nhiều phương tiện và người phục vụ thì "dòng chờ" không có nhưng lại thiệt hại tiền đầu tư mua phương tiện, thuê người nhân rồi. Vì vậy, nghiên cứu hệ thống dịch vụ hay xuất hiện dưới thuật ngữ lý thuyết xếp hàng (queueing theory) hoặc nghiên cứu về dòng chờ (waiting line). Các hệ thống dịch vụ mang tính ngẫu nhiên rất lớn, vì chẳng biết lúc nào "khách" đến mua hàng, cấp cứu hay yêu cầu chữa máy.... Vì vậy ta cũng gặp thuật ngữ hệ thống dịch vụ ngẫu nhiên (stochastic service system).

Phần đầu chương này trình bày lý thuyết xếp hàng, cung cấp các thông tin cần thiết về dòng chờ giúp cho việc lựa chọn quyết định về thiết kế hệ thống dịch vụ, và phần sau xét mô hình chọn quyết định điển hình

10.1 QUÁ TRÌNH NGẪU NHIÊN SINH VÀ TỬ

Đa số các mô hình xếp hàng đều dựa trên *quá trình (ngẫu nhiên) sinh và tử* (birth an death process). Quá trình này lại được định nghĩa thông qua phân bố xác suất dạng hàm số mũ và phân bố Poisson. Vì vậy ta bắt đầu bằng các khái niệm xác suất này.

10.1.1 Phân bố mũ

Nguồn vào (input source) của hệ thống dịch vụ, cũng còn gọi là nguồn khách (calling population), là tổng tất cả các *khách hàng tiềm ẩn* (potential customer) có thể hữu hạn hoặc vô hạn. Nguồn này cung cấp khách theo một quy luật ngẫu nhiên. Qua kinh nghiệm thống kê người ta thấy thường *thời gian giữa hai khách* (interarrival time) là biến ngẫu nhiên T có *phân bố mũ* (exponential distribution) tức là T có hàm phân bố

$$F(t) := P\{T \leq t\} = 1 - e^{-\alpha t}, \quad t \geq 0,$$

ở đây α là tham số. Từ đây suy ra $P\{T > t\} = e^{-\alpha t}$ và hàm mật độ là

$$f_T(t) := \frac{d}{dt}F(t) = \begin{cases} \alpha e^{-\alpha t} & \text{nếu } t \geq 0, \\ 0 & \text{nếu } t < 0. \end{cases}$$

Thời gian phục vụ một khách ở một *phương tiện* (facility) được gọi là *thời gian phục vụ* (service time). Không quá lệch so với thực tế dịch vụ và để thuận lợi về toán học thường người ta giả thiết thời gian phục vụ cũng là biến ngẫu nhiên có phân bố mũ.

Ta hãy nhắc lại một số tính chất cần thiết của phân bố xác suất quan trọng này. Nếu T có phân bố mũ với tham số α thì kỳ vọng là

$E(T) = \frac{1}{\alpha}$ và phương sai là $\text{Var } T = \frac{1}{\alpha^2}$. Có sáu tính chất quan trọng khác như sau.

Tính chất 1. Hàm mật độ $f_T(t)$, với $t \geq 0$, là giảm chặt. (Rõ ràng.)

Tính chất 2 (Tính không nhớ). $P(T > t + \Delta t \mid T > \Delta t) = P(T > t)$. (Suy dễ dàng từ công thức cho xác suất có điều kiện.)

Tính chất 3. Cực tiểu của các biến ngẫu nhiên độc lập có phân bố mũ cũng có phân bố mũ.

Thực vậy, nếu $U = \min \{T_1, \dots, T_n\}$ thì

$$P\{U > t\} = P\{T_1 > t\} \dots P\{T_n > t\} = e^{-\alpha_1 t} \dots e^{-\alpha_n t} = e^{-\sum_{k=1}^n \alpha_k t}$$

tức là U có phân bố mũ với tham số $\sum_{k=1}^n \alpha_k$.

Tính chất 4 (Quan hệ với phân bố Poisson). Gọi $X(t)$, $t \geq 0$, là số lần xảy ra một biến cố mà quãng thời gian giữa hai lần xảy ra liên tiếp là T với phân bố mũ tham số α (lúc $t=0$ bắt đầu đếm số lần). Trong lý thuyết xác suất đã chứng minh là $X(t)$ có phân bố Poisson với tham số αt , tức là

$$P\{X(t) = n\} = \frac{(\alpha t)^n e^{-\alpha t}}{n!}, \quad n = 0, 1, \dots$$

Nói cách khác $X(t)$, $t \geq 0$, là quá trình (ngẫu nhiên) Poisson liên tục với tham số α . α sẽ là kỳ vọng số lần biến cố xảy ra trong một đơn vị thời gian. Chẳng hạn, với giả thiết thời gian giữa hai khách đến, có phân bố mũ thì số khách đã đến ở thời điểm t là $X(t)$ có phân bố Poisson. Số khách đến trung bình trong một đơn vị thời gian sẽ

là tham số α và do đó số đơn vị thời gian trung bình giữa hai khách đến là $\frac{1}{\alpha}$, hoàn toàn phù hợp với Tính chất 4. *Mô hình xếp hàng* (queueing model) với giả thiết là thời gian giữa hai khách có phân bố mũ được gọi là có *quá trình vào Poisson* (Poisson input process) hoặc đơn giản là có *đầu vào Poisson* (Poisson input). Lúc này

$$E\{X(t)\} = \alpha t.$$

Tính chất 5. $P\{T \leq t + \Delta t \mid T > t\} \approx \alpha \Delta t, \forall t > 0, \forall \Delta t \text{ nhỏ.}$

Thực vậy, theo Tính chất 2,

$$P\{T \leq t + \Delta t \mid T > t\} = P\{T \leq \Delta t\} = 1 - e^{-\alpha \Delta t} = \alpha \Delta t - \sum_{n=2}^{\infty} \frac{(-\alpha \Delta t)^n}{n!} \approx \alpha \Delta t.$$

(Tức là $\lim_{\Delta t \rightarrow 0} \frac{P\{T \leq t + \Delta t \mid T > t\}}{\Delta t}$). Về ý nghĩa, Tính chất 2 nói rằng

xác suất để biến cố xảy ra trong khoảng thời gian Δt sắp tới là hằng không phụ thuộc đang xét ở thời điểm t nào (từ khi biến cố cuối xảy ra). Tính chất 5, đi xa hơn, khẳng định là với Δt nhỏ thì xác suất hằng này là xấp xỉ $\alpha \Delta t$. Điều này, cũng phù hợp với Tính chất 4 là số lần xảy ra trung bình trong một đơn vị thời gian là α nên trong Δt đơn vị thời gian phải là $\alpha \Delta t$.

Tính chất 6 (Có thể gộp hoặc tách nhiều loại biến cố độc lập cùng có phân bố Poisson hoặc cùng có phân bố mũ). *Giả sử có n loại khách có thời gian giữa hai khách cùng loại i là T_i có phân bố mũ với tham số λ_i , hay nói cách khác là đầu vào của loại khách i là đầu vào Poisson với tham số λ_i . Khi đó Tính chất 6 nói rằng đầu vào của khách nói chung, không phân loại cũng là đầu vào Poisson. Tham số*

sẽ là $\lambda = \sum_{i=1}^n \lambda_i$, theo Tính chất 3.

Tách biến cố được giải thích cụ thể như sau. Giả sử đầu vào là Poisson có

tham số λ . Giả sử mỗi khách đến có xác suất thuộc loại i là p_i với

$$\sum_{i=1}^n p_i = 1 \text{ và } \lambda_i := p_i \lambda.$$

Tính chất 6 nói rằng đầu vào của khách loại i cũng là đầu vào Poisson với tham số λ_i .

Một áp dụng quan trọng của Tính chất 6 là như sau. Giả sử khách đến nhưng lại bỏ đi với xác suất p còn vào hệ thống dịch vụ với xác suất $1 - p$. Nếu khách đến theo quy luật đầu vào Poisson thì theo Tính chất 6, tính riêng khách vào thì đầu vào cũng là Poisson.

10.1.2 Quá trình sinh và tử của lý thuyết xếp hàng

Ta trình bày quá trình ngẫu nhiên sinh và tử bằng ngôn ngữ của lý thuyết xếp hàng. Gọi *trạng thái* (state) của hệ dịch vụ tại t là số khách trong hệ $N(t)$ ở thời điểm t , $t \geq 0$. Một khách đến được gọi là có một sinh, một khách được phục vụ xong là có một tử. *Quá trình sinh và tử* (birth - and - death process) mô tả xác suất thay đổi của $N(t)$, được định nghĩa bởi ba điều kiện.

Điều kiện 1. Ở trạng thái $N(t) = n$, phân bố xác suất của thời gian đến lần sinh tiếp theo là phân bố mũ với tham số λ_n ($n = 0, 1, \dots$).

Điều kiện 2. Ở trạng thái $N(t) = n$, phân bố xác suất của thời gian cho tới lần tử kế tiếp là phân bố mũ với tham số μ_n .

Điều kiện 3. Hai biến ngẫu nhiên thời gian đến lần sinh kế tiếp và thời gian đến lần tử kế tiếp là độc lập nhau. Việc chuyển trạng thái của hệ sẽ là :

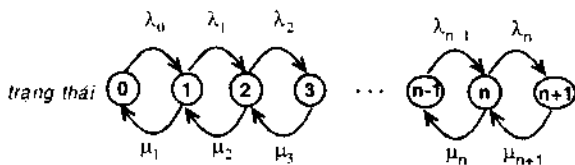
$$n \rightarrow n+1 \text{ (có 1 sinh)}$$

hoặc

$$n \rightarrow n-1 \text{ (có 1 tử),}$$

tùy theo biến ngẫu nhiên nào có giá trị nhỏ hơn (tại trạng thái $N(t) = n$).

Với ba giả thiết này quá trình sinh và tử là một xích Markov với thời gian liên tục. Theo Tính chất 4, λ_n và μ_n là tốc độ sinh và tử trung bình (kỳ vọng số khách trong một đơn vị thời gian). H. 10.1 biểu diễn lược đồ quá trình sinh và tử. Mũi tên là chỉ khả năng chuyển trạng thái. Các λ_i , μ_i ghi bên cạnh là tốc độ chuyển.



H. 10.1

Khi nghiên cứu các mô hình xếp hàng dựa trên quá trình sinh và tử (ở Mục 10.2 sẽ thấy) có bốn đại lượng chính cần tính như sau. Một là kỳ vọng L (tức là trung bình) số khách có mặt trong hệ dịch vụ. Hai là kỳ vọng khách trong hàng L_q (tức là không kể số đang được phục vụ). Gọi $\bar{w}$ là thời gian chờ trong hệ (trong hàng và lúc phục vụ) của mỗi khách, thì kỳ vọng $W := E(\bar{w})$ là số thứ ba cần tính. Gọi $\bar{w}_q$ là thời gian chờ trong hàng thì số thứ tư là $W_q := E(\bar{w}_q)$. [Little 1961] đã chứng minh được là nếu $\lambda_n = \lambda, \forall n$, thì ở *điều kiện duy trì* (steady - state condition) của quá trình xếp hàng ta có

$$L = \lambda W. \quad (10.1)$$

Công thức này gọi là *công thức Little* (Little's formula). Sau đây ta chỉ xét quá trình xếp hàng ở điều kiện duy trì. Nghiên cứu ở *điều kiện quá độ* (transient condition) phức tạp và còn ít có hiệu quả. Cũng với chứng minh của Little ta có

$$L_q = \lambda W_q. \quad (10.2)$$

Nếu λ_n khác hằng số thì ở (10.1) và (10.2) λ được thay bởi tốc độ sinh trung bình xét trong giai đoạn dài là $\bar{\lambda}$.

Nếu thời gian phục vụ một khách trung bình là hằng, tức là $\frac{1}{\mu}$, $\forall n$, thì rõ ràng theo định nghĩa

$$W = W_q + \frac{1}{\mu}. \quad (10.3)$$

10.1.3 Tính các đại lượng cơ bản L , L_q , W , W_q

Gọi $E_n(t)$ là số lần quá trình xếp hàng chuyển vào trạng thái n tính đến thời điểm t và $L_n(t)$ là số lần rời trạng thái n . Vì hai biến cố vào và rời đi phải luân phiên nên

$$|E_n(t) - L_n(t)| \leq 1.$$

Do đó

$$\lim_{t \rightarrow \infty} \left| \frac{E_n(t)}{t} - \frac{L_n(t)}{t} \right| = 0.$$

Mà $\lim_{t \rightarrow \infty} \frac{E_n(t)}{t}$ (và $\lim_{t \rightarrow \infty} \frac{L_n(t)}{t}$) là tốc độ trung bình quá trình chuyển

vào (và rời khỏi) trạng thái n . Vậy ta có *phương trình cân bằng* (balance equation) cho mọi trạng thái n là

$$\text{tốc độ trung bình vào} = \text{tốc độ trung bình ra}.$$

Người ta viết phương trình này (cho mọi n) qua các ẩn số chưa biết là xác suất P_n để hệ phục vụ có đúng n khách và được hệ vô hạn phương trình để xác định các P_n (cùng với phương trình $\sum_{n=0}^{\infty} P_n = 1$).

Rồi từ đó tính được L , L_q , W , W_q như sau.

Trước hết ta viết cụ thể phương trình cân bằng cho trạng thái 0. Quá trình chỉ có thể vào trạng thái 0 từ trạng thái 1 (xem H. 10.1). Xác suất hệ ở trạng thái 1 là P_1 và ở chỗ khác là $1 - P_1$. Tốc độ tử trung bình ở trạng thái 1 là μ_1 chính là tốc độ trung bình vào trạng thái 0 từ trạng thái 1. Còn tốc độ trung bình vào trạng thái 0 từ mọi trạng thái khác là 0. Do đó theo công thức tính kỳ vọng, tốc độ trung bình vào trạng thái 0 là

$$\mu_1 P_1 + 0(1 - P_1) = \mu_1 P_1.$$

Lý luận tương tự sẽ thấy tốc độ trung bình ra khỏi trạng thái 0 là $\lambda_0 P_0$. Do đó phương trình cân bằng cho trạng thái 0 là

$$\mu_1 P_1 = \lambda_0 P_0.$$

Với các trạng thái khác thì đều có hai khả năng vào và hai khả năng ra nên lý luận tương tự như trên ta sẽ được phương trình cân bằng mà mỗi vế đều có hai số hạng. Cuối cùng ta dẫn đến hệ phương trình cân bằng cho quá trình sinh và tử là

trạng thái	tốc độ vào = tốc độ ra
0	$\mu_1 P_1 = \lambda_0 P_0$
1	$\lambda_0 P_0 + \mu_2 P_2 = (\lambda_1 + \mu_1) P_1$
2	$\lambda_1 P_1 + \mu_3 P_3 = (\lambda_2 + \mu_2) P_2$
...	...
$n - 1$	$\lambda_{n-2} P_{n-2} + \mu_n P_n = (\lambda_{n-1} + \mu_{n-1}) P_{n-1}$
n	$\lambda_{n-1} P_{n-1} + \mu_{n+1} P_{n+1} = (\lambda_n + \mu_n) P_n$
...	...

Để dàng giải hệ này từ trên xuống ta được mọi P_n tính qua P_0 như sau

trạng thái

$$0: \quad P_1 = \frac{\lambda_0}{\mu_1} P_0$$

$$1: \quad P_2 = \frac{\lambda_1}{\mu_2} P_1 + \frac{1}{\mu_2} (\mu_1 P_1 - \lambda_0 P_0) = \frac{\lambda_1}{\mu_2} P_1 = \frac{\lambda_1 \lambda_0}{\mu_2 \mu_1} P_0,$$

$$2: \quad P_3 = \frac{\lambda_2}{\mu_3} P_2 + \frac{1}{\mu_3} (\mu_2 P_2 - \lambda_1 P_1) = \frac{\lambda_2}{\mu_3} P_2 = \frac{\lambda_2 \lambda_1 \lambda_0}{\mu_3 \mu_2 \mu_1} P_0,$$

$$\dots \quad \dots \quad \dots$$

$$\begin{aligned} n-1: \quad P_n &= \frac{\lambda_{n-1}}{\mu_n} P_{n-1} + \frac{1}{\mu_n} (\mu_{n-1} P_{n-1} - \lambda_{n-2} P_{n-2}) = \frac{\lambda_{n-1}}{\mu_n} P_{n-1} = \\ &= \frac{\lambda_{n-1} \lambda_{n-2} \dots \lambda_0}{\mu_n \mu_{n-1} \dots \mu_1} P_0, \end{aligned}$$

$$\begin{aligned} n: \quad P_{n+1} &= \frac{\lambda_n}{\mu_{n+1}} P_n + \frac{1}{\mu_{n+1}} (\mu_n P_n - \lambda_{n-1} P_{n-1}) = \frac{\lambda_n}{\mu_{n+1}} P_n = \\ &= \frac{\lambda_n \lambda_{n-1} \dots \lambda_0}{\mu_{n+1} \mu_n \dots \mu_1} P_0, \end{aligned}$$

$$\dots \quad \dots \quad \dots$$

Để đơn giản ta đặt

$$C_n = \begin{cases} \frac{\lambda_{n-1} \dots \lambda_0}{\mu_n \dots \mu_1} & \text{nếu } n = 1, 2, \dots, \\ 1 & \text{nếu } n = 0. \end{cases}$$

Khi đó các xác suất duy trì trên đây có dạng

$$P_n = C_n P_0, \quad n = 0, 1, 2, \dots$$

Cùng với phương trình $\sum_{n=0}^{\infty} P_n = 1$ ta sẽ được $\left(\sum_{n=0}^{\infty} C_n \right) P_0 = 1$. Do đó

$$P_0 = \left(\sum_{n=0}^{\infty} C_n \right)^{-1}.$$

Bốn đại lượng cần tính sẽ biểu diễn được qua các P_n vừa tính như sau.

Theo cách tính kì vọng

$$L = \sum_{n=0}^{\infty} n P_n.$$

Gọi số người được phục vụ trong hệ là s thì ở trạng thái n số người còn trong hàng là $n-s$ vì s người đang được phục vụ rồi. Vì vậy

$$L_q = \sum_{n=s}^{\infty} (n-s) P_n.$$

Theo (10.1) và (10.2) ta có

$$W = \frac{L}{\lambda}, \quad W_q = \frac{L_q}{\lambda}.$$

Lại theo cách tính kì vọng thì

$$\bar{\lambda} = \sum_{n=0}^{\infty} \lambda_n P_n.$$

Chú ý rằng các kết quả trên đây có được là với giả thiết λ_n và μ_n phải thích hợp để quá trình xếp hàng đạt được đến điều kiện duy trì. Điều kiện này thỏa chẳng hạn khi $\lambda_n = 0$ với một $n > 0$ (do đó $\lambda_m = 0$, $\forall m \geq n$), hoặc λ_n và μ_n là hằng sao cho $\frac{\lambda}{s\mu} < 1$. Người ta cũng chứng

minh được là nếu $\sum_{n=0}^{\infty} C_n = \infty$ thì hệ không đạt được điều kiện duy trì.

10.2 CÁC MÔ HÌNH XẾP HÀNG DỰA TRÊN QUÁ TRÌNH SINH VÀ TỬ

Có nhiều mô hình khác nhau dựa trên quá trình sinh và tử do λ_n và μ_n khác nhau. Do Điều kiện 1 và 2 và Tính chất 4 nói trên của quá trình sinh và tử, nên người ta cũng hay nói các mô hình xếp hàng này là có đầu vào Poisson (Poisson input) và thời gian phục vụ kiểu mũ (exponential service time)

10.2.1 Mô hình M/M/1

Chữ M đầu nói rằng đầu vào hệ dịch vụ là Poisson, tức là thời gian giữa hai khách có phân bố mũ, và không phụ thuộc trạng thái n (tức là $\lambda_n = \lambda, \forall n$). Chữ M thứ hai nói rằng thời gian phục vụ cũng kiểu mũ với $\mu_n = \mu$ là hằng. Số 1 nghĩa là hệ chỉ có một người phục vụ. Theo các kết quả tính toán ở cuối Mục 10.1, ta có $C_n = \left(\frac{\lambda}{\mu}\right)^n := \rho^n$, $n = 0, 1, \dots$. Từ nay về sau $\rho := \frac{\lambda}{\mu}$ sẽ gọi là hệ số sử dụng (utilization factor), vì đây là tỉ lệ phần thời gian người phục vụ bận. Ta có

$$P_0 = \left(\sum_{n=0}^{\infty} \rho^n \right)^{-1} = \left(\frac{1}{1-\rho} \right)^{-1} = 1 - \rho,$$

$$P_n = \rho^n P_0 = (1 - \rho) \rho^n, \quad n = 0, 1, \dots$$

Do đó

$$L = \sum_{n=0}^{\infty} n(1 - \rho)\rho^n = (1 - \rho)\rho \sum_{n=0}^{\infty} \frac{d}{d\rho}(\rho^n) = (1 - \rho)\rho \frac{d}{d\rho} \left(\frac{1}{1 - \rho} \right) = \frac{\rho}{1 - \rho} = \frac{\lambda}{\mu - \lambda}.$$

$$L_q = \sum_{n=1}^{\infty} (n-1)P_n = L - 1(1 - P_0) = \frac{\lambda^2}{\mu(\mu - \lambda)}.$$

Chú ý rằng nếu $\lambda \geq \mu$ thì chuỗi $\sum_{n=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^n$ phân kì nên không có P_0

và tính toán trên không đúng nữa. Điều này cũng phù hợp thực tế vì tốc độ khách đến trung bình vượt tốc độ phục vụ trung bình thì khách trong hệ cứ tăng mãi, hệ không thể có điều kiện duy trì. Ta phải giả thiết $\lambda < \mu$.

Bây giờ dùng (10.1) và (10.2) ta có ngay W và W_q như sau

$$W = \frac{1}{\mu - \lambda}, \quad W_q = \frac{\lambda}{\mu(\mu - \lambda)}.$$

Có thể chứng minh được rằng W có phân bố mũ với tham số $\mu - \lambda$: $P\{W > t\} = e^{-(\mu - \lambda)t}$, ($t \geq 0$). Hơn nữa phân bố có điều kiện của W_q , với điều kiện $W_q > 0$, có phân bố mũ với tham số $\mu - \lambda$, tức là

$$P\{W_q > t \mid W_q > 0\} = e^{-(\mu - \lambda)t}, \text{ với } t \geq 0.$$

Người ta cũng tính được phân bố không điều kiện là

$$P\{W_q > t\} = \rho e^{-\mu(1 - \rho)t}, \quad t \geq 0.$$

Công thức này cũng phù hợp với $P\{W_q = 0\} = P_0 = 1 - \rho > 0$.

(Xem [Hillier - Lieberman 2001], Mục 15.6.)

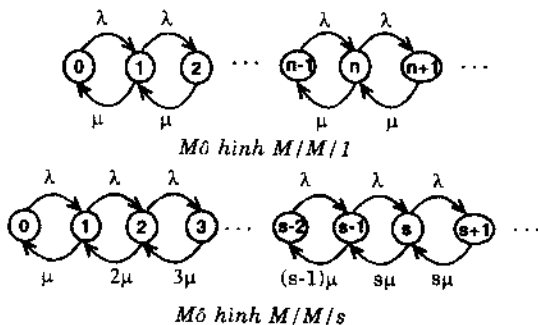
10.2.2 Mô hình M/M/s

Mô hình này tương tự M/M/1 nhưng có s người phục vụ với tốc độ phục vụ trung bình cùng là μ (nếu bận liên tục). Nhớ lại μ_n là tốc

độ phục vụ trung bình của cả hệ khi có n khách trong hệ. Do đó

$$\mu_n = \begin{cases} n\mu & \text{nếu } n \leq s, \\ s\mu & \text{nếu } n \geq s. \end{cases}$$

Vì vậy ta có lược đồ khách vào và ra khi đã được phục vụ xong, cụ thể hoá H. 10.1 cho hai mô hình đang xét, như sau



H. 10.2

Bây giờ ta tính các đại lượng quan trọng cho mô hình M/M/s

$$C_n = \begin{cases} \frac{\left(\frac{\lambda}{\mu}\right)^n}{n!} & \text{với } n \leq s, \\ \frac{\left(\frac{\lambda}{\mu}\right)^s}{s!} \left(\frac{\lambda}{s\mu}\right)^{n-s} = \frac{\left(\frac{\lambda}{\mu}\right)^n}{s! s^{n-s}} & \text{với } n \geq s. \end{cases}$$

Do đó, nếu $\lambda < s\mu$ (tức $\rho = \frac{\lambda}{s\mu} < 1$) thì

$$P_0 = \left(1 + \sum_{n=1}^{s-1} \frac{\left(\frac{\lambda}{\mu}\right)^n}{n!} + \frac{\left(\frac{\lambda}{\mu}\right)^s}{s!} \sum_{n=s}^{\infty} \left(\frac{\lambda}{s\mu}\right)^{n-s} \right)^{-1} = \left(\sum_{n=0}^{s-1} \frac{\left(\frac{\lambda}{\mu}\right)^n}{n!} + \frac{\left(\frac{\lambda}{\mu}\right)^s}{s!} \frac{1}{1 - \frac{\lambda}{s\mu}} \right)^{-1},$$

$$P_n = \begin{cases} \frac{\left(\frac{\lambda}{\mu}\right)^n}{n!} P_0 & \text{nếu } 0 \leq n \leq s, \\ \frac{\left(\frac{\lambda}{\mu}\right)^n}{s! s^{n-s}} P_0 & \text{nếu } n \geq s, \end{cases}$$

$$\begin{aligned} L_q &= \sum_{n=s}^{\infty} (n-s) P_n = \sum_{j=0}^{\infty} j P_{s+j} = \sum_{j=0}^{\infty} j \frac{\left(\frac{\lambda}{\mu}\right)^s}{s!} \rho^j P_0 \\ &= P_0 \frac{\left(\frac{\lambda}{\mu}\right)^s}{s!} \rho \sum_{j=0}^{\infty} \frac{d}{d\rho} (\rho^j) = P_0 \frac{\left(\frac{\lambda}{\mu}\right)^s}{s!} \rho \frac{d}{d\rho} \left(\frac{1}{1-\rho} \right) = \frac{P_0 \left(\frac{\lambda}{\mu}\right)^s \rho}{s! (1-\rho)^2}, \end{aligned}$$

và $W_q = \frac{L_q}{\lambda}$, $W = W_q + \frac{1}{\mu}$, $L = \lambda \left(W_q + \frac{1}{\mu} \right) = L_q + \frac{\lambda}{\mu}$. Tính cụ thể thêm ta sẽ thấy khác với mô hình M/M/1, lúc này $\{i\}$ và $\{i_q\}$ (với điều kiện $i_q > 0$) không còn là biến ngẫu nhiên có phân bố mũ nữa:

$$P\{i > t\} = e^{-\mu t} \left[1 + \frac{P_0 \left(\frac{\lambda}{\mu}\right)^s}{s! (1-\rho)} \left(\frac{1 - e^{-\mu t \left(s-1-\frac{\lambda}{\mu}\right)}}{s-1-\frac{\lambda}{\mu}} \right) \right],$$

$$P\{i_q > t\} = (1 - P\{i_q = 0\}) e^{-s\mu(1-\rho)t},$$

$$P\{i_q = 0\} = \sum_{n=0}^{s-1} P_n.$$

THÍ DỤ 10.2.1. Một trạm sửa chữa thiết bị viễn thông có một người phục vụ (ta dùng số nhỏ để minh họa, có thể hiểu là một dây chuyền ...). Vì có hiện tượng khách quan phải chờ. Giám đốc công ty xét khả năng tăng một người nữa. Ông áp dụng mô hình M/M/s. Thống kê số liệu cũ cho thấy tốc độ khách đến

là 2 người một giờ, tức là thời gian giữa hai khách trung bình là $\frac{1}{\lambda} = \frac{1}{2}$. Thời gian trung bình chữa một bộ máy là 20 phút, tức là $\frac{1}{\mu} = \frac{1}{3}$ (coi đơn vị thời gian là giờ và mỗi khách mang chữa một máy). Phải xét hai trường hợp $s = 1$ và 2. Trong cả hai trường hợp $\rho = \frac{\lambda}{s\mu} < 1$ nên hệ sẽ đạt điều kiện duy trì. Dùng các công thức trên ta tính được bảng sau đây.

	s=1	s=2
ρ	2/3	1/3
P_0	1/3	1/2
$P_n, n \geq 1$	$\frac{1}{3} \left(\frac{2}{3} \right)^n$	$\left(\frac{1}{3} \right)^n$
L_q	$\frac{4}{3}$	$\frac{1}{12}$
L	2	3/4
W_q	2/3	$\frac{1}{24}$
W	1	3/8
$P\{n_q > 0\}$	0,667	0,167
$P\{n_q > t\}$	$\frac{2}{3}e^{-t}$	$\frac{1}{6}e^{-4t}$
$P\{n > t\}$	e^{-t}	$\frac{1}{2}e^{-3t}(3 - e^{-t})$

Qua bảng này ta thấy nên tăng thêm một người phục vụ, vì các số liệu với $s = 2$ tốt hơn nhiều.

Bảng 10.1

10.2.3 Mô hình M/M/1/K và M/M/s/K

Các mô hình này có thêm ký hiệu thứ tư là K để chỉ rằng hàng trong mô hình là hữu hạn, chỉ K khách có chỗ trong mô hình. Hạn chế này trong thực tế có thể do chật hẹp về không gian, cũng có thể do đánh giá theo thống kê là khách thứ K+1 sẽ bỏ sang hệ dịch vụ khác. Do đó ta có

$$\lambda_n = \begin{cases} \lambda & \text{nếu } n \leq K-1, \\ 0 & \text{nếu } n \geq K. \end{cases}$$

Vì $\lambda_n = 0$ với một $n > 0$. Hệ sẽ đạt tới điều kiện duy trì.

Lúc này ta có lược đồ đúng như H.10.2, nhưng "cắt cụt" từ trạng thái $K+1$ trở đi.

1. *Mô hình M/M/1/K*. Ta tính được các đại lượng cần thiết như sau.

$$C_n = \begin{cases} \left(\frac{\lambda}{\mu}\right)^n & \text{nếu } n \leq K, \\ 0 & \text{nếu } n > K. \end{cases}$$

với $\rho = \frac{\lambda}{\mu} \neq 1$ thì

$$P_0 = \left(\sum_{n=0}^K \left(\frac{\lambda}{\mu}\right)^n \right)^{-1} = \frac{1-\rho}{1-\rho^{K+1}}.$$

Do đó

$$P_n = \frac{1-\rho}{1-\rho^{K+1}} \rho^n \text{ với } n = 0, \dots, K,$$

$$L = \sum_{n=0}^K n P_n = \frac{1-\rho}{1-\rho^{K+1}} \rho \sum_{n=0}^K \frac{d}{d\rho} (\rho^n) = \frac{\rho}{1-\rho} - \frac{(K+1)\rho^{K+1}}{1-\rho^{K+1}},$$

$$L_q = L - (1 - P_0)$$

Nếu $\rho = 1$ thì $P_n = \frac{1}{K+1}$ vì khả năng có 0, 1, ..., K khách là như

nhau. Do đó $L = \sum_{n=0}^K n P_n = \frac{1}{K+1} \sum_{n=0}^K n = \frac{K}{2}$.

Nếu $\rho < 1$ thì cho $K \rightarrow +\infty$ ta được các kết quả đã có cho $M/M/1$.

Tiếp theo, ta có $\bar{\lambda} = \sum_{n=0}^{\infty} \lambda_n P_n = \sum_{n=0}^{K-1} \lambda P_n = \lambda (1 - P_K)$ và $W = \frac{L}{\bar{\lambda}}$,

$$W_q = \frac{L_q}{\bar{\lambda}}.$$

2. **Mô hình $M/M/s/K$.** Tất nhiên phải giả thiết $s \leq K$. Công thức cho C_n và P_n chỉ khác ở mô hình $M/M/s$ là $= 0$ với $n > K$. Tiếp theo,

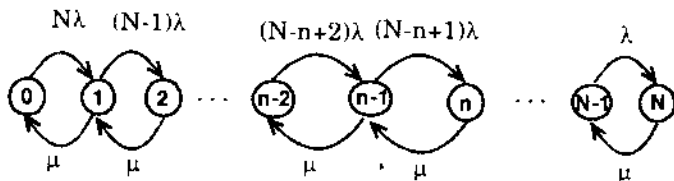
$$P_0 = \left(\sum_{n=0}^s \frac{\left(\frac{\lambda}{\mu}\right)^n}{n!} + \frac{\left(\frac{\lambda}{\mu}\right)^s}{s!} \sum_{n=s+1}^K \left(\frac{\lambda}{s\mu}\right)^{n-s} \right)^{-1},$$

$$L_q = \frac{P_0 \left(\frac{\lambda}{\mu}\right)^s \rho}{s!(1-\rho)^2} [1 - \rho^{K-s} - (K-s)\rho^{K-s}(1-\rho)],$$

$$L = \sum_{n=0}^{s-1} n P_n + L_q + s \left(1 - \sum_{n=0}^{s-1} P_n \right).$$

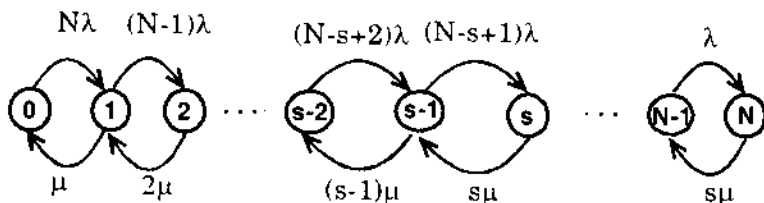
10.2.4 Mô hình $M/M/s$ với nguồn vào hữu hạn

Một thí dụ thực tế áp dụng mô hình này là trường hợp một mạng lưới bảo hành cho các máy hoạt động, mỗi "người phục vụ" phụ trách N máy. Ta gọi N là cơ sở của nguồn vào (size of the input source hoặc size of the calling population). Khi hệ có n khách thì số khách tiềm ẩn là $N - n$. Thời gian giữa hai khách có phân bố mũ λ sẽ có nghĩa là thời gian mỗi người (trong N) ở ngoài hệ có phân bố mũ λ . Số người ngoài hệ là $N - n$ nên phân bố xác suất hiện hành của thời gian cho tới lần vào hệ tiếp đến là cực tiểu thời gian tới lúc vào hệ của $N - n$ người. Theo Tính chất 2 và 3 của quá trình sinh và tử phân bố xác suất này là phân bố mũ với tham số $\lambda_n = (N - n)\lambda$. Do đó lược đồ của mô hình với nguồn hữu hạn là trường hợp riêng của lược đồ ở H.10.1:



$$\lambda_n = \begin{cases} (N-n)\lambda & \text{nếu } n \leq N, \\ 0 & \text{nếu } n \geq N, \end{cases} \quad \mu_n = \mu$$

H. 10.3 (a) Mô hình M/M/1 với nguồn hữu hạn



$$\lambda_n = \begin{cases} (N-n)\lambda & \text{nếu } n \leq N, \\ 0 & \text{nếu } n \geq N, \end{cases} \quad \mu_n = \begin{cases} n\mu & \text{nếu } n \leq s, \\ s\mu & \text{nếu } n \geq s. \end{cases}$$

H. 10.3 (b) Mô hình M/M/s với nguồn hữu hạn

Chú ý rằng $\lambda_n = 0$ với $n = N$ nên hệ dịch vụ sẽ đạt được điều kiện duy trì.

1. *Mô hình M/M/1 với nguồn hữu hạn.* Theo công thức tổng quát ở Mục 10.1.3 ta có

$$C_n = \begin{cases} N(N-1)\dots(N-n+1)\left(\frac{\lambda}{\mu}\right)^n = \frac{N!}{(N-n)!}\left(\frac{\lambda}{\mu}\right)^n & \text{nếu } n \leq N, \\ 0 & \text{nếu } n > N, \end{cases}$$

$$P_0 = \left(\sum_{n=0}^N \frac{N!}{(N-n)!}\left(\frac{\lambda}{\mu}\right)^n \right)^{-1}$$

$$P_n = \frac{N!}{(N-n)!} \left(\frac{\lambda}{\mu}\right)^n P_0, \quad n \leq N,$$

$$L_q = \sum_{n=s}^N (n-1)P_n = N - \frac{\lambda + \mu}{\lambda}(1 - P_0),$$

$$L = \sum_{n=0}^N nP_n = L_q + 1 - P_0 = N - \frac{\mu}{\lambda}(1 - P_0),$$

$$\bar{\lambda} = \sum_{n=0}^{\infty} \lambda_n P_n = \sum_{n=0}^N (N-n)\lambda P_n = \lambda(N - L),$$

$$W = \frac{L}{\bar{\lambda}}, \quad W_q = \frac{L_q}{\bar{\lambda}}.$$

2. Mô hình $M/M/s$ với nguồn hữu hạn ($N \geq s$). Tương tự ta có

$$C_n = \begin{cases} \frac{N!}{(N-n)!n!} \left(\frac{\lambda}{\mu}\right)^n & \text{với } n \leq s, \\ \frac{N!}{(N-n)!s!s^{n-s}} \left(\frac{\lambda}{\mu}\right)^n & \text{với } n = s, \dots, N, \\ 0 & \text{với } n > N, \end{cases}$$

$$P_n = \begin{cases} \frac{N!}{(N-n)!n!} \left(\frac{\lambda}{\mu}\right)^n P_0 & \text{với } n \leq s, \\ \frac{N!}{(N-n)!s!s^{n-s}} \left(\frac{\lambda}{\mu}\right)^n P_0 & \text{với } n = s, \dots, N, \\ 0 & \text{với } n > N, \end{cases}$$

$$P_0 = \left(\sum_{n=0}^{s-1} \frac{N!}{(N-n)!n!} \left(\frac{\lambda}{\mu}\right)^n + \sum_{n=s}^N \frac{N!}{(N-n)!s!s^{n-s}} \left(\frac{\lambda}{\mu}\right)^n \right)^{-1},$$

$$L_q = \sum_{n=s}^N (n-s)P_n, \quad L = \sum_{n=0}^{s-1} nP_n + L_q + s \left(1 - \sum_{n=0}^{s-1} P_n \right).$$

[Bunday - Scraton 1980] đã chứng minh được rằng các công thức cho P_0 , P_n , L_q , L , W và W_q cho hai mô hình với nguồn hữu hạn trên đây vẫn đúng khi thời gian ở ngoài hệ của $N-n$ khách tiềm ẩn không có phân bố mũ mà có phân bố xác suất bất kì với kì vọng $1/\lambda$.

10.2.5 Mô hình với tốc độ khách đến và tốc độ phục vụ phụ thuộc n

Mô hình này phù hợp thực tế vì khi đông khách người phục vụ thường bị sức ép và làm nhanh hơn. Trong bài toán thực tế phức tạp hơn thì có thể giải thích là lúc đông khách có thể phải bố trí nhiều công đoạn sang người phụ cho người phục vụ (chính), chẳng hạn nhiều y tá hỗ trợ cho bác sĩ ...

1. Trường hợp $s = 1$. Thường tốc độ phục vụ tăng khi khách đông. Giả sử

$$\mu_n = n^a \mu_1,$$

ở đây a là hệ số sức ép. (Như vậy các mô hình với tốc độ không phụ thuộc trạng thái trên đây là trường hợp $a=0$.) Lúc này ở mô hình

thường giả sử $\lambda_n = \lambda$. Khi đó $C_n = \frac{(\lambda/\mu_1)^n}{(n!)^a}$. Cũng có thể ngược lại, giả

sử $\mu_n = \mu$ và $\lambda_n = (n+1)^{-b} \lambda_0$, ở đây b cũng là hệ số sức ép (chú ý là

trong thực tế thường λ_n giảm theo n). Lúc này $C_n = \frac{(\lambda_0/\mu)^n}{(n!)^b}$. Các công

thức tính cho trường hợp này tất nhiên là trường hợp riêng của Mục 10.1.3, nhưng cũng đã phải tính xấp xỉ bằng tổng hữu hạn thay cho chuỗi.

Tổng quát hơn có thể giả thiết

$$\mu_n = n^a \mu_1, \quad \lambda_n = (n+1)^{-b} \lambda_0.$$

Khi đó $C_n = \frac{(\lambda_0/\mu_1)^n}{(n!)^c}$ với $c = a + b$. Người ta lập bảng tính gần xấp xỉ các đại lượng cần, nhất là P_0 và L .

2. Trường hợp $s > 1$. Một cách tự nhiên giả sử μ_n và λ_n biến thiên theo $\frac{n}{s}$ (tức là số khách trong hệ trên một người phục vụ). Khi đó, tương tự trường hợp $s = 1$, người ta giả sử

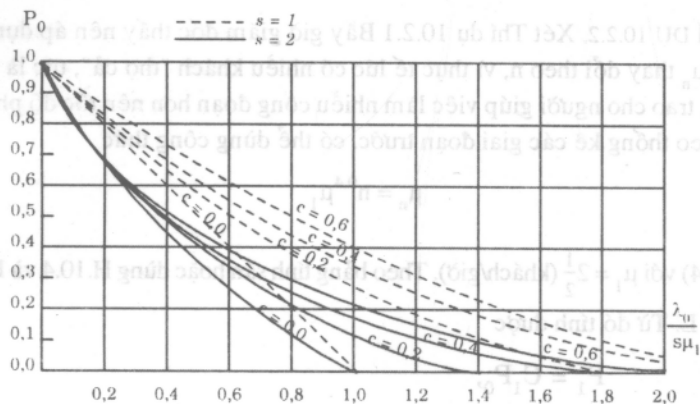
$$\mu_n = \begin{cases} n\mu_1 & \text{nếu } n \leq s, \\ \left(\frac{n}{s}\right)^d s\mu_1 & \text{nếu } n \geq s, \end{cases}$$

$$\lambda_n = \begin{cases} \lambda_0 & \text{nếu } n \leq s-1, \\ \left(\frac{s}{n+1}\right)^b \lambda_0 & \text{nếu } n \geq s-1. \end{cases}$$

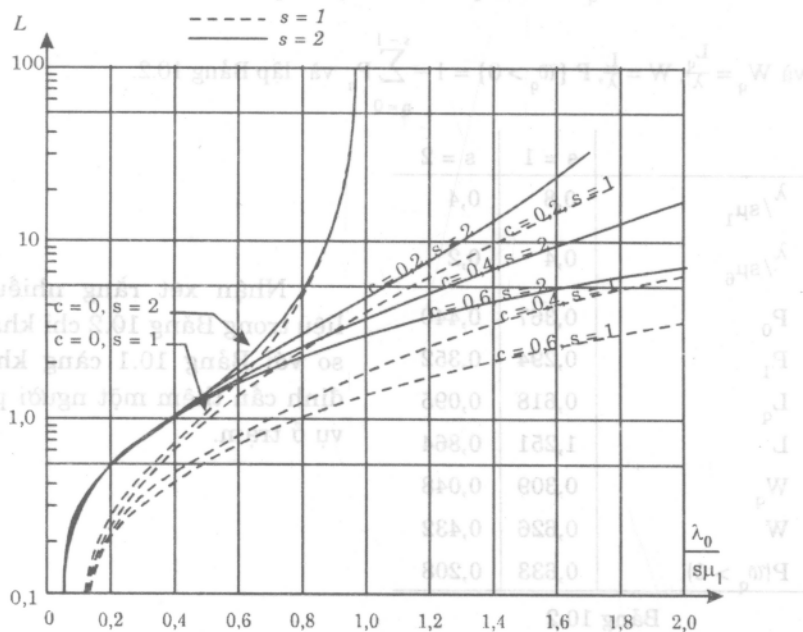
Khi đó theo công thức chung cho quá trình sinh và tử ta có ($c = a+b$)

$$C_n = \begin{cases} \frac{(\lambda_0/\mu_1)^n}{n!} & \text{nếu } n \leq s, \\ \frac{(\lambda_0/\mu_1)^n}{s! \left(\frac{n!}{s!}\right)^c s^{(1-c)(n-s)}} & \text{nếu } n \geq s. \end{cases}$$

Người ta cũng lập bảng sẵn cho P_0 và L . Theo bảng đó ta có đồ thị của P_0 và L phụ thuộc vào c , $\frac{\lambda_0}{\mu_1}$ và s như ở H.10.4 và H.10.5.



H.10.4 Xác suất duy trì của trạng thái không khách trong mô hình phụ thuộc trạng thái



H.10.5 Kỳ vọng duy trì của số khách trong hệ ở mô hình phụ thuộc trạng thái

THÍ DỤ 10.2.2. Xét Thí dụ 10.2.1 Bây giờ giám đốc thấy nên áp dụng mô hình có μ_n thay đổi theo n , vì thực tế lúc có nhiều khách "thợ cả", tức là người phục vụ, trao cho người giúp việc làm nhiều công đoạn hơn nên tốc độ phục vụ tăng. Theo thống kê các giai đoạn trước, có thể dùng công thức

$$\mu_n = n^{0.4} \mu_1$$

(tức $a=0.4$) với $\mu_1 = 2\frac{1}{2}$ (khách/giờ). Theo bảng tính sẵn hoặc dùng H.10.4 và H.10.5 ta có P_0, L . Từ đó tính được

$$P_1 = C_1 P_0,$$

$$L_q = L - (1 - P_0) \text{ với } s = 1,$$

$$L_q = L - P_1 - 2(1 - P_0 - P_1) \text{ với } s = 2,$$

và $W_q = \frac{L_q}{\lambda}$, $W = \frac{L}{\lambda}$, $P\{w_q > 0\} = 1 - \sum_{n=0}^{s-1} P_n$ và lập Bảng 10.2.

	s = 1	s = 2
$\lambda / s\mu_1$	0,8	0,4
$\lambda / s\mu_6$	0,4	0,2
P_0	0,367	0,440
P_1	0,294	0,352
L_q	0,618	0,095
L	1,251	0,864
W_q	0,309	0,048
W	0,626	0,432
$P\{w_q > 0\}$	0,633	0,208

Bảng 10.2

Nhận xét rằng nhiều số liệu trong Bảng 10.2 chỉ khác ít so với Bảng 10.1 càng khẳng định cần thêm một người phục vụ ở trạm.

10.3 CÁC MÔ HÌNH KHÔNG DÙNG QUÁ TRÌNH SINH VÀ TỬ

Các mô hình ở Mục 10.2 được dùng phổ biến nhất, nhưng một số bài toán thực tế không hợp với mô hình dùng quá trình sinh và tử như vậy. Một số mô hình khác đã được nghiên cứu, thường gọi chung là *mô hình xếp hàng với phân bố khác hàm mũ* (queueing models, with nonexponential distribution). Vì các kết quả này rất phức tạp, ở đây ta chỉ giới thiệu sơ qua một số mô hình.

Mô hình M/G/1. Điểm khác ở đây là thời gian phục vụ cho các khách được giả thiết là độc lập và có cùng phân bố xác suất. Phân bố này được giả thiết là tổng quát chỉ cần có kì vọng $\frac{1}{\mu}$ và phương sai σ^2 đã biết. Khi đó Pollaczek và Khintchine đã chứng minh được các công thức đơn giản mang tên họ như sau, nếu $\rho = \frac{\lambda}{\mu} < 1$ (hệ đạt được điều kiện duy trì), xem [Hordijk-Tijms 1976],

$$P_0 = 1 - \rho, L_q = \frac{\lambda^2 \sigma^2 + \rho^2}{2(1 - \rho)}, L = \delta + L_q, W_q = \frac{L_q}{\lambda}, W = W_q + \frac{1}{\mu}.$$

Các công thức này đơn giản một cách bất ngờ so với sự tổng quát của mô hình (ở "vị trí" G) nên mô hình được sử dụng rộng rãi. Tiếc rằng người ta chưa tìm được kết quả tương tự cho trường hợp $s > 1$. Chỉ có hai mô hình sau là trường hợp G cụ thể.

Mô hình M/D/s. Chữ D ở đây ấn định thời gian phục vụ là hằng số cố định. Xem các phương pháp xác định P_n và L chẳng hạn trong [Gross-Harris 1985] trang 344-346.

Mô hình M/E_k/s. Trung gian giữa M/M/s với $\sigma^2 = \frac{1}{2} \left(\frac{1}{\mu} \right)$ (lớn) và

$M/D/s$ với $\sigma^2 = 0$ là mô hình với *phân bố Erlang* (Erlang distribution) cho thời gian phục vụ phân bố Erlang có hàm mật độ là

$$f(t) = \frac{(\mu k)^k}{(k-1)!} t^{k-1} e^{-k\mu t}, \quad t \geq 0,$$

ở đây μ và k dương, k nguyên. Nếu bỏ điều kiện k nguyên thì phân bố cũng có tên là *phân bố gamma* (gamma distribution). Phân bố Erlang có kì vọng là $\frac{1}{\mu}$ và phương sai là $\frac{1}{k\mu^2}$. Vậy tham số k đặc trưng cho phương sai trong quan hệ với μ (E_k là từ tên Erlang và k), thường gọi là *tham số dạng* (shape parameter). Xem [Gross-Harris 1985] về các công thức, bảng và đồ thị.

Mô hình với đầu vào không Poisson. Mô hình GI/M/s giả thiết đầu vào có phân bố bất kì, xem [Gross - Harris 1985], trang 304-320. Các mô hình D/M/s và E_k /M/s cũng được nghiên cứu và lập bảng sẵn, xem [Hillier - Yu - Avis - Fossett - Lo - Reiman 1981].

10.4 MÔ HÌNH XẾP HÀNG ƯU TIÊN

Trên đây đã xét các mô hình mà khách được phục vụ theo nguyên tắc *đến trước phục vụ trước* (first - come - first - served). Nhiều bài toán cần phân loại ưu tiên phục vụ, chẳng hạn ra làm N lớp (lớp 1 được ưu tiên nhất). Trong cùng lớp thì lại áp dụng quy tắc đến trước phục vụ trước. Có hai mô hình có ưu tiên chính. Cả hai mô hình đều giả thiết mỗi lớp ưu tiên đều có đầu vào Poisson và thời gian phục vụ trung bình cho khách mọi lớp là như nhau. Tốc độ khách đến trung bình với mỗi lớp có thể khác nhau.

Mô hình ưu tiên nhường ngay (preemptive priorities model) quy định là khách ở lớp kém ưu tiên nhất, trong số khách đang được phục

vụ rồi, phải nhường ngay chỗ khi có khách ưu tiên hơn vào hệ và ra chờ. Vì tính chất không nhớ của quá trình, ta không phải nhớ thời điểm một khách phải nhường. Khi người này lại đến lượt thì phân bố xác suất của thời gian phục vụ còn lại luôn không đổi. *Mô hình ưu tiên không nhường ngay* (nonpreemptive priorities model) quy định là ai đã được phục vụ thì được làm cho xong, ưu tiên chỉ áp dụng khi gọi khách vào một chỗ vừa được giải phóng.

Theo Tính chất 6 của quá trình sinh và tử thì khi không phân biệt khách nữa, đầu vào chung cũng là đầu vào Poisson. Còn thời gian phục vụ thì đã giả thiết là luôn có cùng phân bố mũ rồi. Do đó khi ta tính tổng số khách (L và L_q) cũng như kì vọng của thời gian chờ (W và W_q) của một khách ngẫu nhiên thì các công thức của mô hình không ưu tiên ở các mục trước vẫn đúng. Chỉ có phân bố xác suất của thời gian chờ là khác. Phân bố này có phương sai lớn vì thời gian chờ của khách ưu tiên cao là ngắn hơn nhiều. Vì vậy, chẳng hạn để xem thiết kế hệ dịch vụ thế nào cho tốt phải tính kì vọng thời gian chờ trong hệ W_k cho từng lớp $k = 1, \dots, N$ và kì vọng số khách trong hệ L_k cho từng lớp (thay cho W, L chung).

10.4.1 Mô hình ưu tiên không nhường ngay

Người ta chứng minh được là với điều kiện $\sum_{i=1}^N \lambda_i < s\mu$ thì

$$W_k = \frac{1}{AB_{k-1}B_k} + \frac{1}{\mu}, \quad k = 1, \dots, N,$$

ở đây

$$A = s! \frac{s\mu - \lambda}{r^s} \sum_{j=0}^{s-1} \frac{r^j}{j!} + s\mu,$$

$$B_0 = 1, B_k = 1 - \frac{1}{s\mu_1} \sum_{i=1}^k \lambda_i, \quad k = 1, \dots, N,$$

λ_i là tốc độ đến trung bình của lớp i , $\lambda = \sum_{i=1}^N \lambda_i$, $r = \frac{1}{\mu}$;

$$L_k = \lambda_k W_k, \quad k = 1, \dots, N.$$

10.4.2 Mô hình ưu tiên nhường ngay

Với $s = 1$, người ta chứng minh được

$$W_k = \frac{1}{\mu B_{k-1} B_k}, \quad k = 1, \dots, N.$$

Với $s > 1$, ta có phương pháp tính dần từ $W_1, W_2, \dots$ như minh họa ở Ví dụ 10.4.1. Số khách trung bình của lớp k trong hệ (kể cả đang được phục vụ) cũng vẫn thỏa công thức Little:

$$L_k = \lambda_k W_k, \quad k = 1, \dots, N.$$

THÍ DỤ 10.4.1. Lại xét thí dụ trạm sửa chữa viễn thông. Bây giờ giám đốc công ty thấy cần phân biệt khách làm ba loại ưu tiên khác nhau vì đây là trạm phục vụ cho mục đích kinh doanh chung của công ty, có máy cần được ưu tiên sửa trước. Giả sử có ba lớp ưu tiên với tốc độ đến trung bình là $\lambda_1 = 0,2$, $\lambda_2 = 0,6$, $\lambda_3 = 1,2$. Ta tính các đại lượng theo các công thức trên và viết vào Bảng 10.3 ba hàng cuối (trang sau) là thời gian trong hàng (không kể thời gian phục vụ, nhưng có thể ở trong hàng nhiều lần do phải nhường chỗ). Ở mô hình với $s = 2$ ba lượng này được tính như sau (trên đây chưa có công thức).

		ưu tiên nhường ngay		ưu tiên không nhường ngay	
		s = 1	s = 2	s = 1	s = 2
A		–	–	4,5	36

B_1	0,933	—	0,933	0,967
B_2	0,733	—	0,733	0,867
B_3	0,333	—	0,333	0,667
$W_1 - \frac{1}{\mu}$	0,024	0,00037	0,238	0,029
$W_2 - \frac{1}{\mu}$	0,154	0,00793	0,325	0,033
$W_3 - \frac{1}{\mu}$	0,033	0,06542	0,889	0,048

Bảng 10.3 Kết quả ở điều kiện duy trì của trạm sửa chữa máy

Vì khách ở lớp ưu tiên 1 không bị ảnh hưởng bởi khách lớp kém hơn nên $W_1 = W$, ở đây W là thời gian chờ trung bình trong hệ không ưu tiên (với một lớp) $M/M/s$ có $s = 2$, $\mu = 3$, $\lambda = \lambda_1 = 0,2$ nên tương tự Bảng 10.1 (ở đó $\lambda = 2$) ta có

$$W_1 = W = 0,33370,$$

$$W_1 - \frac{1}{\mu} = 0,33370 - 0,33333 = 0,00037.$$

Bây giờ xét hai lớp đầu. Khách loại này không bị ảnh hưởng bởi lớp 3 nên thời gian chờ trung bình chung cho hai lớp W_{1-2} lại vẫn tính theo mô hình không ưu tiên với $s = 2$, $\mu = 3$ và $\lambda = \lambda_1 + \lambda_2 = 0,8$ và được

$$W_{1-2} = 0,033937.$$

Với một khách ngẫu nhiên thuộc hai lớp đầu thì xác suất thuộc lớp 1 là

$$\frac{\lambda_1}{\lambda_1 + \lambda_2} = \frac{1}{4} \text{ và thuộc lớp 2 là } \frac{\lambda_2}{\lambda_1 + \lambda_2} = \frac{3}{4}. \text{ Do đó}$$

$$W_{1-2} = \frac{1}{4}W_1 + \frac{3}{4}W_2.$$

Từ đó suy ra

$$W_2 = \frac{4}{3}(0,33937 - \frac{1}{4} \cdot 0,33370) = 0,34126$$

$$\text{và } W_2 - \frac{1}{\mu} = 0,00793.$$

Tiếp theo gọi W_{1-3} là thời gian chờ trung bình trong hệ (nhớ là tính cả thời gian phục vụ) của một khách bất kì (thuộc ba lớp) thì W_{1-3} được tính với mô hình M/M/2 không ưu tiên với $\mu = 3$ và $\lambda = \lambda_1 + \lambda_2 + \lambda_3 = 2$, do đó $W_{1-3} = 0,375$. Xác suất để khách này rơi vào lớp 1, 2, 3 tương ứng là 0,1; 0,3 và 0,6. Do đó

$$W_{1-3} = 0,1W_1 + 0,3W_2 + 0,6W_3.$$

Suy ra

$$W_3 = \frac{1}{0,6}(0,375 - 0,1 \cdot 0,33370 - 0,3 \cdot 0,34126) = 0,39875,$$

$$W_3 - \frac{1}{\mu} = 0,06542.$$

Nhìn Bảng 10.3 ta thấy khi có hai người phục vụ thời gian chờ giảm nhiều, nhất là ở mô hình ưu tiên nhường ngay. Vì vậy giám đốc quyết định tăng một người phục vụ cho trạm.

10.5 MẠNG DỊCH VỤ

Trên đây mới chỉ xét hệ dịch vụ gồm một phương tiện ở đó có s người phục vụ. Khách đã được phục vụ là chỉ qua một phương tiện. Tất nhiên mô hình gồm nhiều phương tiện là rất cần thiết, chẳng hạn một khu du lịch giải trí gồm nhiều "trò" và mỗi khách có thể qua một số trò giải trí theo trình tự khác nhau, cũng có thể bỏ qua một số chỗ không vào. Các mô hình như vậy được gọi là *mạng dịch vụ* (service network) hoặc *mạng xếp hàng* (queueing network), đang được nghiên

cứu nhiều, nhưng khá phức tạp. Ở đây ta chỉ giới thiệu sơ lược.

10.5.1 Chuỗi phương tiện với hàng vô hạn

Quá trình Poisson có tính chất cơ bản sau đây giúp cho việc nghiên cứu mạng đơn giản hơn nhiều [xem Burke 1956].

Tính chất tương đương (equivalence property). Giả sử một phương tiện có s người phục vụ và hàng vô hạn, có đầu vào Poisson với tham số λ . Giả sử mỗi người phục vụ đều có thời gian phục vụ là phân bố mũ với tham số μ (tức là ta có mô hình $M/M/s$) với $s\mu > \lambda$. Khi đó đầu ra duy trì của phương tiện cũng là quá trình Poisson với tham số λ .

Chú ý rằng tính chất tương đương đúng cho cả nguyên tắc đến trước phục vụ trước và cả nguyên tắc ưu tiên phục vụ.

Trường hợp mạng đơn giản nhất là *chuỗi phương tiện với hàng vô hạn* (infinite queues in series). Ở đây có m phương tiện, phương tiện i có s_i , $i = 1, \dots, m$, người phục vụ cùng có thời gian phục vụ với phân bố mũ với tham số μ_i , xếp thành chuỗi cố định; phương tiện đầu tiên có đầu vào Poisson với tham số λ . Khi đó theo tính chất tương đương đầu ra của phương tiện 1, lại là đầu vào của phương tiện 2, cũng là đầu vào Poisson. Vì vậy ta có m mô hình $M/M/s_i$ thực chất là độc lập, có thể tính toán riêng như ở các mục trên. Từ đó suy ra các tính chất cho cả mạng. Chẳng hạn xác suất của biến cố (chung) là có n_1 khách ở phương tiện 1, ..., n_m khách ở phương tiện m là

$$P[(N_1, \dots, N_m) = (n_1, \dots, n_m)] = P_{n_1} \dots P_{n_m}, \quad (10.4)$$

ở đây P_{n_i} được tính theo công thức cho mô hình $M/M/s_i$. (10.4) được gọi là *ng nghiệm dạng tích* (product form solution).

Tiếc rằng nghiệm dạng tích không đúng cho chuỗi phương tiện

với hàng hữu hạn, tức là chỗ xếp hàng ở các phương tiện bị hạn chế. Sự độc lập của các phương tiện trong mạng và nghiệm dạng tích làm cho việc nghiên cứu mạng dễ hơn nhiều nên người ta cũng tập trung tìm các mô hình mạng có các tính chất đó. Một mô hình như vậy được xét bây giờ.

10.5.2 Mạng Jackson

[Jackson 1963] đưa ra mô hình mạng gồm m phương tiện xác định như sau:

1. mỗi phương tiện i có đầu vào từ ngoài mạng là đầu vào Poisson với tham số a_i có s_i người phục vụ cùng với thời gian phục vụ là phân bố mũ với tham số μ_i và có hàng dài vô hạn;
2. mỗi khách rời phương tiện i có xác suất vào phương tiện j là p_{ij} (và rời mạng với xác suất $q_j = 1 - \sum_{j=1}^m p_{ij}$).

Jackson đã chứng minh được rằng mạng như vậy có tính chất là, với điều kiện duy trì, mỗi phương tiện j đều hoạt động như một hệ $M/M/s_j$ độc lập với tốc độ đến trung bình

$$\lambda_j = a_j + \sum_{i=1}^m \lambda_i p_{ij} ,$$

ở đây $s_j \mu_j > \lambda_j$. Do đó, nói riêng ta cũng có nghiệm dạng tích.

10.6 BIỂU DIỄN XẾP HÀNG QUA XÍCH MARKOV RỜI RẠC

Trong các mục trước ta đã mô tả quá trình xếp hàng, hay nói

khác là quá trình dịch vụ, như quá trình ngẫu nhiên liên tục, vì thực tế biến thời gian là liên tục, không thể định ra hai thời điểm thế nào là kế tiếp nhau. Tuy vậy ta có thể dùng cách tiếp cận xấp xỉ để dẫn tới quá trình rời rạc.

Giả sử tốc độ trung bình khách đến là α_n và tốc độ phục vụ trung bình là μ_n tại trạng thái hệ có n khách. Ở đây đơn vị thực tế thường là khách/giờ. Ta chọn một đoạn thời gian đơn vị nhỏ để có thể coi là trong một đoạn chỉ một khách đến hoặc rời hệ (khi phục vụ xong), chẳng hạn 1 giây. Khi đó tốc độ trung bình nói trên theo đơn vị mới là $\frac{\lambda_n}{3600} := \frac{\lambda_n}{m}$, $\frac{\mu_n}{3600} := \frac{\mu_n}{m}$.

Giả sử hàng có độ dài N (khách) hữu hạn. Có thể coi quá trình là Markov, tức là không nhớ, vì nếu ở thời điểm hiện hành hệ ở trạng thái n thì diễn biến tiếp theo chỉ phụ thuộc vào α_n , μ_n không phụ thuộc quá khứ. Vì đã có các thời điểm rời rạc, ta được xích Markov (rời rạc). Do đó thuận tiện nhất là viết ma trận chuyển. Nếu hệ đang ở trạng thái 0 thì xác suất để chuyển sang 1 là xác suất có một khách đến, tức là $\frac{\lambda_0}{m}$. Do đó xác suất để chuyển sang trạng thái 0 là

$p_{00} = 1 - \frac{\lambda_0}{m}$. Rõ ràng $p_{02} = \dots = p_{0N} = 0$. Ta đã được hàng đầu của ma

trận (p_{ij}) . Xét hàng thứ hai, tức là khi hệ đang ở trạng thái có 1 khách.

Xác suất chuyển sang có 0 khách rõ ràng là $p_{10} = \left(1 - \frac{\lambda_0}{m}\right) \frac{\mu_1}{m}$; không khách đến và phục vụ xong 1 khách. Vì m lớn có thể coi $\frac{\lambda_0 \mu_1}{m^2}$ bằng 0

nên $p_{10} = \frac{\mu_1}{m}$. Tương tự xác suất chuyển sang trạng thái 1 và 2 là

$$p_{11} = \left(1 - \frac{\lambda_1}{m}\right) \left(1 - \frac{\mu_1}{m}\right) + \frac{\lambda_1}{m} \frac{\mu_1}{m} = 1 - \frac{\lambda_1}{m} - \frac{\mu_1}{m} + 0,$$

$$p_{12} = \frac{\lambda_1}{m} \left(1 - \frac{\mu_1}{m}\right) = \frac{\lambda_1}{m} + 0.$$

Xác suất khác trong hàng rõ ràng là $p_{13} = \dots = p_{1N} = 0$. Các hàng khác cũng nhận được bằng cách tương tự. Cuối cùng ta được ma trận truyền P có dạng

	0	1	2	3	...	N-2	N-1	N
0	$1 - \frac{\lambda_0}{m}$	$\frac{\lambda_0}{m}$	0	0	0	0	0	0
1	$\frac{\mu_1}{m}$	$1 - \frac{\lambda_1}{m} - \frac{\mu_1}{m}$	$\frac{\lambda_1}{m}$	0	0	0	0	0
2	0	$\frac{\mu_2}{m}$	$1 - \frac{\lambda_2}{m} - \frac{\mu_2}{m}$	$\frac{\lambda_2}{m}$	0	0	0	0
...								
N-1	0	0	0	0	$\frac{\mu_{N-1}}{m}$	$1 - \frac{\lambda_{N-1}}{m} - \frac{\mu_{N-1}}{m}$	$\frac{\lambda_{N-1}}{m}$	0
N	0	0	0	0	0	$\frac{\mu_N}{m}$	$1 - \frac{\mu_N}{m}$	0

Do đó các phương trình duy trì (9.8), (9.9) trong trường hợp này là

$$P_j = \sum_{i=0}^N P_i p_{ij}, \quad j = 0, \dots, N,$$

$$\sum_{i=0}^N P_i = 1,$$

ở đây các xác suất duy trì cần tính P_j là xác suất để hệ có j khách. Theo ma trận truyền $P = (p_{ij})$ ở trên ta có

$$P_0 = P_0 \left(1 - \frac{\lambda_0}{m} \right) + P_1 \frac{\mu_1}{m} \Rightarrow P_1 = \frac{\lambda_0}{\mu_1} P_0,$$

$$P_1 = P_0 \frac{\lambda_0}{m} + P_1 \left(1 - \frac{\lambda_1}{m} - \frac{\mu_1}{m} \right) + P_2 \frac{\mu_2}{m} \Rightarrow P_2 = \frac{\lambda_1}{\mu_2} P_1 = \frac{\lambda_1 \lambda_0}{\mu_2 \mu_1} P_0,$$

...

$$P_N = P_{N-1} \frac{\lambda_{N-1}}{\mu_N} + P_N \left(1 - \frac{\mu_N}{\mu_N}\right) \Rightarrow P_N = \frac{\lambda_{N-1}}{\mu_N} P_{N-1} = \dots = \frac{\lambda_{N-1} \lambda_{N-2} \dots \lambda_0}{\mu_N \mu_{N-1} \dots \mu_1} P_0,$$

$$\sum_{i=0}^N P_i = 1.$$

Thay tất cả P_i theo P_0 vào phương trình cuối ta được đúng như ở Mục 10.1.3, khi tiếp cận theo quá trình Markov liên tục,

$$P_0 = \left(\sum_{i=0}^N C_i \right)^{-1}, \quad P_n = C_n P_0, \quad n = 0, \dots, N,$$

$$C_i = \begin{cases} \frac{\lambda_{i-1} \dots \lambda_0}{\mu_1 \dots \mu_i} & \text{nếu } i = 1, \dots, N, \\ 1 & \text{nếu } i = 0. \end{cases}$$

Do đó các đại lượng cơ bản L, L_q, W, W_q cũng đúng như ở Mục 10.1.3.

THÍ DỤ 10.6.1. Một trò chơi trong khu du lịch có thể phục vụ 50 khách một giờ ($\mu_n = \mu = 50$). Thống kê cho thấy tốc độ trung bình khách đến là 100 khách/giờ. Trò chơi được coi như hệ có một người phục vụ. Lãi thu được từ một khách là 5 (ngàn đồng). Bài toán đặt ra là thiết kế hàng chờ thế nào để lãi hơn. Theo ý kiến chung là có hai cách: hàng chờ lộ thiên và hàng chờ kín để khách không biết có nhiều người trước mình, chỉ khi còn 4 người đăng trước mới bắt đầu nhìn thấy. Người ta thấy rằng thường khách (tiềm năng) đến nhưng thấy đông người chờ thì bỏ đi. Cụ thể thống kê cho thấy tốc độ khách vào như bảng sau

n khách	0	1	2	3	4	5
λ_n khi hàng lộ thiên	100	80	60	40	20	0
λ_n khi hàng kín	100	100	100	100	100	0

Ta áp dụng các công thức vừa có để tính các P_j (với $\mu_j = 50$).

- Hàng lộ thiên. $P_1 = \frac{\lambda_0}{\mu} P_0 = 2P_0$, $P_2 = \frac{\lambda_1}{\mu} P_1 = 3,2P_0$, $P_3 = 1,92P_0$,
 $P_4 = 1,536P_0$, $P_5 = 0,6144P_0$. Do đó
 $P_0(1 + 2 + 3,2 + 1,92 + 1,536 + 0,6144) = 1 \Rightarrow P_0 = 0,0974$.

Như vậy xác suất để trò chơi có khách là $1 - P_0 = 0,9026$. Khi có khách thường xuyên thì 50 khách/giờ và lãi từ mỗi khách là 5 nên lãi trung bình một giờ là

$$50 \cdot 5 \cdot 0,9026 = 225,5 \text{ ngàn đồng/giờ.}$$

- Hàng kín. $P_1 = 2P_0$, $P_2 = 4P_0$, $P_3 = 8P_0$, $P_4 = 16P_0$, $P_5 = 32P_0$. Do đó
 $P_0(1 + 2 + 4 + 8 + 16 + 32) = 1 \Rightarrow P_0 = 0,1587$.

Do đó lãi trung bình là

$$50 \cdot 5 \cdot (1 - 0,1587) = 246,0 \text{ ngàn đồng/giờ.}$$

Như vậy để hàng kín lãi hơn 20,5 ngàn đồng/giờ.

Nhận xét rằng ở đầu Mục 10.6 ta đã chọn m lớn để trong thời gian $1/m$ có thể coi là chỉ xảy ra một biến cố đến hoặc đi khỏi hệ. Qua tính toán tất cả các đại lượng cần P_j , C_j và do đó L , L_q , W , W_q ta thấy chúng không phụ thuộc giá trị m cụ thể, chỉ cần lớn để giả thiết trên là đúng.

10.7 CHỌN QUYẾT ĐỊNH VỀ THIẾT KẾ HỆ DỊCH VỤ

10.7.1 Kiểu quyết định

Ở các mục trước ta đã nghiên cứu các tính chất và tính các đại lượng quan trọng nhất cho các mô hình dịch vụ, tức là các mô hình xếp hàng. Ta cũng đã áp dụng để chọn quyết định liên quan đến thiết kế dịch vụ một cách trực tiếp bằng so sánh. Bây giờ ta tập trung nghiên cứu các cách chọn quyết định tốt nhất về thiết kế hệ. Các tính

huống phải chọn quyết định trong thực tế rất đa dạng, nhưng đa số là tổ hợp của ba quyết định sau:

1. Số người phục vụ s ở 1 phương tiện,
2. Chất lượng của người phục vụ, thể hiện qua tốc độ phục vụ trung bình μ ,
3. Số phương tiện trong hệ n , thể hiện qua tốc độ khách đến mỗi phương tiện λ (vì $n = \lambda_p / \lambda$, ở đây λ_p là tốc độ khách đến của cả hệ, đã biết)

Một số các quyết định kiểu khác ta không xét ở đây, chẳng hạn số chỗ đợi cần bao nhiêu? thiết kế quy tắc ưu tiên thế nào?.... Ta cũng không xét vấn đề điều khiển và điều chỉnh hệ trong mô hình "động" tức là nhu cầu phục vụ thay đổi

Mục tiêu tối ưu là làm cực tiểu kì vọng của tổng cước phí. Tổng cước phí bao gồm cước phục vụ CP (lương người phục vụ, tiền mua phương tiện) và cước chờ đợi CD. Vì ta xét hệ dịch vụ bằng ngôn ngữ lý thuyết xếp hàng, nên ta quan tâm chính đến vấn đề giảm chờ đợi, như các mục trước đã nói (từ đó dẫn đến tăng lợi nhuận). Cước phí chờ đợi không dễ hiểu nên ta xét cụ thể thêm.

10.7.2 Các dạng cước chờ đợi

Có hai dạng cơ bản hay gặp trong thực tế.

1. *Dạng $g(N)$* . Khi khách hàng cũng là người của tổ chức có hệ dịch vụ, chẳng hạn các máy hỏng đợi chữa ở một trạm sửa chữa của công ty, hay công nhân đợi phát dụng cụ ..., thì cước đợi chờ là hàm của số người đợi N (trong hệ), vì khi đợi là ngừng sản xuất. Do đó cước đợi là hàm $g(N)$ tăng theo N . Nhưng N là biến ngẫu nhiên với $P\{N = n\} = P_n$ nên kì vọng cước đợi chờ một đơn vị thời gian của hệ là

$$E(CD) = E(g(N)) = \sum_{n=0}^{\infty} g(n)P_n.$$

Nếu $g(N) = C_D N$ là tuyến tính, ở đây C_D là cước đợi một đơn vị thời gian của mỗi khách, thì

$$E(CD) = C_D \sum_{n=0}^{\infty} nP_n = C_D L.$$

THÍ DỤ 10.7.1 (Bao nhiêu thợ sửa máy?) Một nhà máy thiết bị viễn thông có 10 máy sản xuất một loại linh kiện đắt tiền. Máy khá nhạy cảm, dễ hỏng, nên nhà máy chỉ bố trí 8 “công nhân” để hoạt động 8 máy (“công nhân” theo nghĩa rộng, có thể là đội sản xuất), còn hai máy sẵn sàng thay thế hoặc đang sửa. Nếu có ít hơn 8 máy chạy được thì thiếu mỗi máy sản xuất sẽ mất 400 (ngàn đồng) một ngày. Khi đó cước chờ đợi của hệ là

$$g(n) = \begin{cases} 0 & \text{nếu } n = 0, 1, 2, \\ 400(n-2) & \text{nếu } n = 3, \dots, 10. \end{cases}$$

Muốn tính $E(CD)$ ta cần biết P_n . Khi đó cần bổ sung dữ kiện để xác định bài toán. Giả sử thời gian đến khi máy hỏng là biến ngẫu nhiên có phân bố mũ, với kì vọng là 20 ngày (và ta coi đối với các máy là như nhau). Thời gian chữa một máy cũng là biến ngẫu nhiên có phân bố mũ, với kì vọng là 2 ngày. Nhà máy hiện chỉ có một thợ sửa máy và thấy có hiện tượng máy hỏng phải chờ nên đang xét việc bổ sung một thợ. Áp dụng mô hình M/M/s với nguồn vào hữu hạn ở Mục 10.2.4 với $N = 10$, $\lambda = \frac{1}{20}$, $\mu = \frac{1}{2}$, với chú ý là, thay cho $\lambda_0 = 10\lambda$, $\lambda_1 = 9\lambda$ ở đó, trong trường hợp này $\lambda_0 = 8\lambda$, $\lambda_1 = 8\lambda$. Ta tính được Bảng 10.4 cho $s = 1$ và $s = 2$.

N = n	g(n)	s = 1		s = 2	
		P_n	$g(n)P_n$	P_n	$g(n)P_n$
0	0	0,271	0	0,433	0
1	0	0,217	0	0,346	0
2	0	0,173	0	0,139	0

3	400	0,139	56	0,055	24
4	800	0,097	78	0,019	16
5	1200	0,058	70	0,006	8
6	1600	0,029	46	0,001	0
7	2000	0,012	24	$3 \cdot 10^{-4}$	0
8	2400	0,003	7	$4 \cdot 10^{-5}$	0
9	2800	$7 \cdot 10^{-4}$	0	$4 \cdot 10^{-6}$	0
10	3200	$7 \cdot 10^{-5}$	0	$2 \cdot 10^{-7}$	0
E(CD)		281/ngày		48/ngày	

Bảng 10.4

2. *Dạng $h(w)$* . Khi khách hàng là người ngoài tổ chức có hệ dịch vụ, chẳng hạn ở các hệ dịch vụ thương mại, vận tải, dịch vụ xã hội thì cước đợi chờ thường là hàm của thời gian đợi của một khách w (kể cả thời gian được phục vụ), có dạng hàm $h(w)$, tăng theo w . Khi đó

$$E(CD) = \lambda E(h(w)) = \lambda \int_0^{\infty} h(w)f(w)dw,$$

ở đây $f(w)$ là hàm mật độ của biến ngẫu nhiên w .

THÍ DỤ 10.7.2 (Máy tính nào?). Một cơ quan của bộ tổng tham mưu đang xét việc trang bị lại siêu máy tính. Có hai phương án thuê main frame: máy deep blue của hãng IBM giá 5000 USD, máy của hãng CRAB giá 3750 USD một ngày. Khi chạy các chương trình điển hình của cơ quan trong 24 giờ thì số chương trình chạy xong là biến ngẫu nhiên có phân bố Poisson với kì vọng tương ứng là 30 và 25. Thời gian giữa thời điểm xuất hiện hai chương trình cần chạy là biến ngẫu nhiên có phân bố mũ với kì vọng 0,05 ngày. Các chuyên gia tham mưu đánh giá tổn thất là 500 USD mỗi ngày chương trình phải chờ máy. Ngoài ra, tổn thất do bị gián đoạn công tác tham mưu còn lớn hơn, được đánh giá là tỉ lệ với w^2 , cụ thể là bằng $400 w^2$. Như vậy

$$h(w) = 500 w + 400 w^2$$

Để tính $E(CD)$ ta áp dụng mô hình M/M/1. Vì $f(w) = (\mu - \lambda)e^{-(\mu - \lambda)w}$ nên

$$E(h(w)) = \int_0^{\infty} (500w + 400w^2)(\mu - \lambda)e^{-(\mu - \lambda)w} dw.$$

Ta có

$$\mu - \lambda = \begin{cases} 30 - 20 = 10 & \text{cho hãng IBM,} \\ 25 - 20 = 5 & \text{cho hãng CRAB.} \end{cases}$$

Cuối cùng ta được

$$E(CD) = 20 E(h(w)) = \begin{cases} 20.58 = 1160 & \text{cho hãng IBM,} \\ 20.132 = 2640 & \text{cho hãng CRAB.} \end{cases}$$

TRƯỜNG HỢP TUYỂN TÍNH. Nếu $h(w)$ tuyển tính:

$$h(w) = C_D w,$$

thì

$$E(CD) = \lambda E(C_D w) = C_D (\lambda W) = C_D L,$$

hoàn toàn trùng với trường hợp $g(N) = C_D N$. Vậy kì vọng cước đợi chờ một đơn vị thời gian của cả hệ trường hợp hàm cước phí phụ thuộc tuyến tính vào số khách trong hệ hoặc vào thời gian chờ của mỗi khách là như nhau và tỉ lệ với số khách trung bình L trong hệ.

THÍ DỤ 10.7.3 (Máy trạm dụng cụ?). Một nhà máy cơ khí mới đang được thiết kế. Một vấn đề đặt ra là đặt máy trạm phát và thu dụng cụ khi các thợ cơ khí không cần nữa. Kinh nghiệm cho thấy số thợ cơ khí đến trạm theo phân bố Poisson với kì vọng 120 thợ một giờ (tức $\lambda = 120$ nếu có một trạm) và thời gian một nhân viên trạm phục vụ một thợ cơ khí phân bố mũ với kì vọng $\frac{1}{120}$ giờ (tức là $\mu = 120$). Một thợ nghỉ (chờ dụng cụ) một giờ thì nhà máy thiệt hại 480 (ngàn đồng). Tiền công một nhân viên trạm là 200 một giờ và tiền bảo hành ... một trạm là 160 một giờ.

Ta có ngay $C_D = 480$ và $E(CD) = 480L$.

Để thuận tiện ta sẽ tính nổi L ở mục sau.

10.8 CÁC MÔ HÌNH QUYẾT ĐỊNH

Tổ hợp của ba biến quyết định đã nói là s , μ và λ người ta thường dùng ba mô hình sau đây.

Mô hình 1 (Chọn s). Biết μ , λ và C_s là tiền công một đơn vị thời gian của một người phục vụ. Tìm s sao cho cực tiểu kì vọng cước phí tổng cộng trong một đơn vị thời gian.

$$\min E(TC) = sC_s + E(CD).$$

Chỉ một số giá trị s (số người phục vụ) được xét, nên cách thông thường là tính $E(TC)$ cho mỗi giá trị s rồi so sánh.

THÍ DỤ 10.8.1 (Bao nhiêu thợ sửa máy?). Xét tiếp Thí dụ 10.7.1. Giả sử tiền công một ngày của thợ sửa máy là 280 (ngàn đồng) tức là $C_s = 280$. Khi đó, dùng Bảng 10.4 cho $E(CD)$ ta tính được $E(TC)$ như Bảng 10.5. Theo đó nhà máy vẫn chỉ thuê một thợ.

s	sC_s	$E(CD)$	$E(TC)$
1	280	281	561 $\leftarrow \min$
2	560	48	608
≥ 3	≥ 840	≥ 0	≥ 840

Bảng 10.5

Mô hình 2 (chọn s và μ). Giả sử $f(\mu)$ là tiền công một đơn vị thời gian của người phục vụ khi tốc độ phục vụ trung bình là μ và A là tập giá trị chấp nhận được của μ . Giả sử đã biết λ , $f(\mu)$ và A . Tìm s và μ

sao cho

$$\min_{\mu \in A \text{ và } s} E(TC) = sf(\mu) + E(CD).$$

THÍ DỤ 10.8.2 (Máy tính nào?). Xét Thí dụ 10.7.2. Ta có $A = \{30, 25\}$, $f(30) = 5000$, $f(25) = 3750$, $s \equiv 1$. Do đó

$$E(TC) = \begin{cases} 5000 + 1160 = 6160 & \text{cho máy IBM,} \\ 3750 + 2640 = 6390 & \text{cho máy CRAB.} \end{cases}$$

Do đó cơ quan quyết định thuê máy deep blue của IBM.

Nếu s không cố định như thí dụ trên nhưng A vẫn hữu hạn, thì ta làm hai bước. Bước 1 ta cố định $\mu \in A$, tìm s theo mô hình 1. Bước 2, so sánh các giá trị phụ thuộc μ nhận được.

Nếu A là tập hợp vô hạn, có thể thực hiện hai bước theo thứ tự ngược lại. Bước 1 cố định s tìm $\min_{\mu \in A} E(TC)$. Bước 2 so sánh các cực

tiểu với s khác nhau. Tuy vậy việc giải là khó khăn vì kết quả giải tích cho mô hình xếp hàng với $s > 1$ phức tạp và nhiều khi không có công thức giải tích. Thực tế thường cố gắng chọn tập A hữu hạn làm đại diện và dùng bảng tính sẵn cho các μ thuộc A đó. May mắn rằng với giả thiết khá nhẹ thì cực tiểu theo μ , s luôn đạt được với $s = 1$. Chẳng hạn [Stidham 1970] đã chứng minh được kết quả sau.

Tính tối ưu của một người phục vụ. Giả sử

(i) lấy $s = 1$ thì $\min_{\mu \in R} E(TC)$ đạt cực trị tại $\mu^* \in A$;

(ii) hàm $f(\mu)$ tuyến tính hoặc lõm (tức là $-f(\mu)$ lồi).

Khi đó $\min \{ E(TC) : \mu \in A \text{ và mọi } s \}$ luôn đạt được với $s = 1$. Để hiểu kết quả này ta hãy giải thích ý tưởng. Giả sử tìm được nghiệm tối ưu khác của mô hình 2 là (s^*, μ^*) với $s^* > 1$. Xét nghiệm $(1, s^*\mu^*)$

với một người phục vụ và có cùng công suất. Bảng 10.6 cho thấy tuy cùng công suất nhưng hệ với một người có trường hợp có tốc độ phục vụ trung bình cao hơn và không có trường hợp nào thấp hơn.

N = n	Tốc độ phục vụ trung bình	
	(1, $s^* \mu^*$)	(s^*, μ^*)
n = 0	0	= 0
n = 1, ..., $s^* - 1$	$s^* \mu^*$	> $n \mu^*$
$n \geq s^*$	$s^* \mu^*$	= $s^* \mu^*$

Bảng 10.6

Do tốc độ phục vụ thấp hơn nên $E(CD)$ với (s^*, μ^*) sẽ cao hơn. Còn cước phí phục vụ thì, do giả thiết (ii),

$$s^* f(\mu^*) \geq f(s^* \mu^*).$$

Vì vậy tổng cước $E(TC)$ của hệ với (s^*, μ^*) cũng lớn hơn, tức là tối ưu là hệ một người phục vụ. Trong thực tế người ta áp dụng tính chất này cả khi (i) và (ii) không thỏa.

Mô hình 3 (chọn s và λ). Giả sử biết μ , λ_p (tốc độ khách đến cả hệ), C_s (công một người phục vụ một đơn vị thời gian) và C_f (cước phí một phương tiện, một đơn vị thời gian). Giả sử khách chia đều cho các phương tiện, tức là số phương tiện $n = \lambda_p / \lambda$. Tìm λ và s sao cho đạt $\min E(TC)$.

Có thể hình dung công thức cho $E(TC)$ là

$$E(TC) = n [(C_f + sC_s) + E(CD)].$$

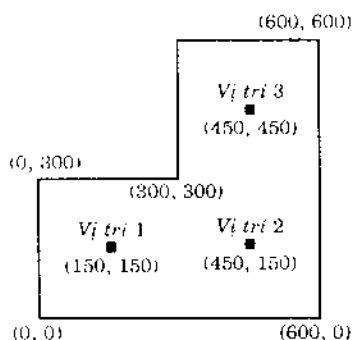
Nếu công thức này đúng thì lý luận như với mô hình 2 sẽ thấy nghiệm tối ưu phải đạt được với $n = 1$. Tuy nhiên nhiều tình huống thực tế cho thấy không thể có nghiệm tối ưu với một phương tiện (chẳng hạn với cao ốc 30 tầng mà thiết kế một nhà vệ sinh dù to mấy đi nữa).

Vậy có gì không thực tế ở công thức trên. Đó là chưa kể đến "cước phí" thời gian đi tới và rời khỏi phương tiện. Gọi T là (biến ngẫu nhiên) thời gian đến và rời (hai chiều) phương tiện của một khách và C_t là cước một đơn vị thời gian này của một khách. Khi đó

$$E(TC) = n [(C_f + sC_s) + E(CD) + \lambda C_t E(T)].$$

Để tìm cực tiểu, thường tính $E(TC)$ cho mỗi n và mỗi s .

THÍ DỤ 10.8.3 (Mấy trạm dụng cụ?). Xét Thí dụ 10.7.3. Giả sử khuôn viên nhà máy cho ở H.10.6. Hai số trong ngoặc là tọa độ các điểm. Có ba vị trí thích hợp cho trạm dụng cụ, công nhân đến trạm bằng các hành lang song song với các trục tọa độ (không có đường đi xiên). Có ba lựa chọn



H. 10.6

(a) một trạm ở vị trí 2,

(b) hai trạm ở vị trí 1 và 3,

(c) ba trạm ở 1, 2 và 3.

Công nhân phân bố đều trong khuôn viên và đi với tốc độ 3 (km/giờ). Khi đó dễ tính được $E(T)$ bằng 0,04; 0,0278 và 0,02 giờ cho mỗi lựa chọn (a), (b) và (c). Ta đã có $\mu = 120$, $\lambda = 120$, $C_f = 160$, $C_s = 200$ và $C_t = 480$. Ta đã có, ở Thí dụ 10.7.3, $E(CD) = 48L$. Áp dụng mô hình M/M/s ta tính được L , do đó có Bảng 10.7. Theo đó ta thấy phải chọn quyết định (c) là ba trạm, đồng thời mỗi trạm chỉ có một nhân viên (lương nhân viên cao quá!), khi đó $\lambda = 40$ cho mỗi trạm.

n	λ	s	L	$E(T)$	$C_f + sC_s$	$E(CD)$	$\lambda C_t E(T)$	$E(TC)$
1	120	1	∞	0,04	360	∞	2304,0	∞
1	120	2	1,333	0,04	560	640,0	2304,0	3504,0
1	120	3	1,044	0,04	760	501,1	2304,0	3565,1
2	60	1	1,000	0,0278	360	480,0	800,0	3280,0

2	60	2	0,534	0,0278	560	256,3	800,0	3232,6
3	40	1	0,500	0,02	360	240,0	384,0	2952,0
3	40	2	0,344	0,02	560	165,1	384,0	3327,3

Bảng 10.7

BÀI TẬP CHƯƠNG 10

10.1. Giả sử các việc đến thực hiện trên một máy theo quá trình đầu vào Poisson với kì vọng 2 việc một giờ. Giả sử máy hỏng và sẽ phải chữa trong một giờ. Tính xác suất để số việc sẽ đến trong một giờ này

- (a) bằng 0, (b) bằng 2 và (c) bằng hoặc lớn hơn 5.

10.2 Xét quá trình sinh và tử với $\mu_n = 2$, $n = 1, 2, \dots$ và $\lambda_0 = 3$, $\lambda_1 = 2$, $\lambda_2 = 1$ và $\lambda_n = 0$ với $n = 3, 4, \dots$

- (a) Vẽ lược đồ của quá trình;
(b) Tính P_0, P_1, P_2, P_3 và P_n với $n = 4, 5, \dots$;
(c) Tính L, L_q, W và W_q .

10.3 Một trạm xăng chỉ có một máy bơm. Xe máy đến theo quá trình Poisson với tốc độ trung bình 15 xe/giờ. Biết rằng nếu ở trạm đang có n xe thì xác suất để một xe khác đến rồi bỏ đi mua chỗ khác là $n/3$ với $n = 1, 2, 3$. Thời gian phục vụ một xe có phân bố mũ với kì vọng 4 phút.

- (a) Vẽ lược đồ cho hệ xếp hàng.
(b) Viết các phương trình cân bằng.
(c) Giải các phương trình này cho xác suất duy trì để hệ có n xe. Kiểm tra rằng nghiệm này trùng với nghiệm nhận được cho quá trình sinh và tử.
(d) Tìm kì vọng thời gian chờ trong hệ kể cả thời gian phục vụ.

10.4 Một siêu thị nhỏ chỉ có một quầy tính tiền, khách ra tính tiền theo quá trình Poisson với tốc độ trung bình 30 khách/giờ. Khi chỉ có một khách ở quầy tính tiền thì chỉ có một nhân

viên thu tiền với thời gian trung bình 1,5 phút. Khi có 2 khách trở lên thì có thêm người phụ giúp và thời gian trung bình còn 1 phút trung bình. Biết thời gian phục vụ có phân bố mũ.

- (a) Vẽ lược đồ cho hệ xếp hàng.
- (b) Tính các xác suất duy trì của số khách trong hệ.
- (c) Tính L , L_q , W và W_q .

10.5 Khách đến một dịch vụ theo đầu vào Poisson với tốc độ trung bình 2 và thời gian phục vụ có phân bố mũ với kì vọng 1/4. Ngoài một khách đang được phục vụ chỉ có chỗ chờ tiện nghi cho ba khách, còn lại ở phòng ngoài. Tính tỉ lệ thời gian tất cả khách chờ đều có chỗ tiện nghi.

10.6 Một dịch vụ sửa xe ô tô đang được thiết kế và phải cân nhắc cần mấy chỗ xe chờ. Ước lượng được rằng xe đến như đầu vào Poisson với tốc độ trung bình 4 phút một xe, nhưng nếu hết chỗ thì xe sẽ bỏ đi. Thời gian sửa một xe có phân bố mũ với kì vọng 3 phút. Tính xác suất để một xe đến phải đi vì hết chỗ nếu

- (a) có 0 chỗ (không kể 1 xe đang rửa); (b) 2 chỗ; (c) 4 chỗ.

10.7 Một xí nghiệp đang được thiết kế sẽ gồm nhiều máy tự động cùng loại và cần tính xem để mỗi công nhân phụ trách mấy máy. Biết rằng thời gian giữa hai lần máy cần công nhân đến và thời gian một lần công nhân phải đứng chỉnh máy đều có phân bố mũ với kì vọng 150 phút và 15 phút, tương ứng. Mỗi công nhân chỉ lo các máy của mình, không thể giúp nhau. Cần đảm bảo sản xuất ít nhất 89% thời gian các máy phải chạy.

- (a) Nhiều nhất là mấy máy được trao cho mỗi công nhân?
- (b) Nếu trao số máy như vậy thì trung bình bao nhiêu phần thời gian công nhân bận điều chỉnh máy?

10.8 Xét mô hình M/G/1.

- (a) So sánh kì vọng thời gian chờ trong hàng W_q nếu thời gian phục vụ có phân bố (i) mũ, (ii) hằng số và (iii) Erlang với độ lệch chuẩn là trung bình cộng của hai trường hợp trước.
- (b) W_q và L_q sẽ thay đổi thế nào cho ba phân bố (i), (ii) và (iii) nếu cả λ và μ được nhân đôi.

10.9 Một xí nghiệp chuyên đại tu một loại thiết bị viễn thông có bốn phần giống nhau và chỉ đại tu được từng thiết bị. Hơn nữa, cần hoàn thành nhanh để trả thiết bị về sản xuất, nên người ta áp dụng cách là mỗi lần chỉ đại tu một trong bốn phần. Với cách này thì các thiết bị phải đến đại tu theo quá trình Poisson với tốc độ trung bình 1/ngày. Thời gian đại tu một phần (trong bốn) có phân bố mũ với trung bình 1/2 ngày. Có người đề nghị cứ đại tu cả bốn phần một lượt. Như thế thời gian đại tu tăng 4 lần, nhưng tốc độ thiết bị phải đến đại tu cũng giảm 4 lần. Hãy so sánh L của hai cách để thấy cách nào hơn.

10.10 Có ba loại "khách" đến một người phục vụ theo quá trình Poisson với tốc độ trung bình là 2,4 và 2 (một ngày). Thời gian phục vụ cũng có phân bố mũ với kì vọng 0,1 (ngày). Thứ tự cần ưu tiên là loại 1 rồi đến 2, 3. Hãy tính kì vọng thời gian chờ (cả thời gian phục vụ) cho ba nguyên tắc xếp hàng

- (a) đến trước phục vụ trước,
- (b) ưu tiên không nhường ngay,
- (c) ưu tiên nhường ngay.

10.11 Xét chuỗi gồm hai phương tiện với hàng vô hạn, mỗi phương tiện có một người phục vụ. Thời gian phục vụ là độc lập, có phân bố mũ với kì vọng 3 và 4 phút, tương ứng. Phương tiện 1 có đầu vào Poisson với tốc độ trung bình 10 khách/giờ.

- (a) Tìm các xác suất duy trì của số khách cho từng phương tiện.
- (b) Tính xác suất để cả hai phương tiện không có khách.
- (c) Tính kì vọng tổng khách trong hệ và kì vọng tổng thời gian chờ (tính cả thời gian phục vụ) của một khách.

10.12 Một trạm xăng có một bơm riêng cho ôtô (xe máy phục vụ riêng), có chỗ đỗ cho 2 xe chờ và 1 xe đang bơm. Biết xe đến theo quá trình Poisson với tốc độ trung bình $\lambda = 20$ xe/giờ và thời gian phục vụ có phân bố mũ với tốc độ trung bình $\mu = 40$ xe/giờ. Hãy chọn đoạn thời gian đơn vị là 1 giây để biểu diễn xếp hàng qua xích Markov rời rạc.

- (a) Tính các xác suất duy trì để có 0, ..., 3 xe trong hệ.
- (b) Giả sử trạm lãi 15 (ngàn đồng) từ một xe ôtô. Tính lãi trung bình của trạm trong một giờ. Nếu có chỗ chờ đầy đủ thì lãi trung bình là bao nhiêu?

(c) Hãy tính lãi trung bình của trạm trong một giờ như là hàm của số chỗ chờ N (kể cả xe đang bơm) và tính cụ thể cho $N = 1, \dots, 6$.

10.13 Giả sử hệ xếp hàng phù hợp mô hình $M/M/1$ với $\lambda = 2$, $\mu = 4$. Tính $E(CD)$ khi cước đợi chờ có dạng (a) $g(N) = 10N + 2N^2$, (b) $h(w) = 25w + w^3$.

10.14 Một siêu thị nhỏ có một quầy tính tiền với một nhân viên lương 16 ngàn/giờ. Khách ra tính tiền theo quá trình Poisson với tốc độ 30 khách/giờ, thời gian tính tiền có phân bố mũ với kì vọng 1,5 phút. Thấy khách phải chờ, chủ siêu thị tính rằng mỗi khách chờ một phút sẽ làm mất 0,08 ngàn đồng cho kinh doanh (vì khách có thể không đến siêu thị nữa). Nếu thuê thêm một người phụ ở quầy 8 ngàn/giờ thì kì vọng thời gian tính tiền còn 1 phút.

(a) Mô hình nào trong ba mô hình quyết định ở Mục 10.8 áp dụng ở đây? Vì sao?

(b) Áp dụng mô hình này xem có nên thêm người phụ không.

10.15 Một đội bốc dỡ xe tải với thời gian xong một xe có phân bố mũ và nếu đội có một người thì kì vọng là 1 giờ. Xe đến theo quá trình Poisson với tốc độ trung bình 1 xe/giờ. Tiến công một người một giờ là 20 (ngàn đồng). Một xe tải phải đứng một giờ ở trạm sẽ thiệt hại 30 ngàn vì không chuyên chở.

(a) Giả sử cường độ bốc dỡ tỉ lệ số người. Cần mấy người để cước phí trung bình một giờ nhỏ nhất.

(b) Nếu giả sử cường độ bốc dỡ tỉ lệ bình phương thì cần mấy người.

10.16 Xét trạm sửa chữa viễn thông với ưu tiên những ngay ở Thí dụ 10.4.1. Giả sử xác định được cước đợi chờ (không kể thời gian chữa máy) là 10 triệu đồng/giờ cho máy ưu tiên 1, 100 ngàn cho máy ưu tiên 2 và 1 ngàn cho loại 3. Giả sử thêm một người phục vụ chỉ tốn 4 ngàn/giờ. Dùng Bảng 10.3, dựa vào kì vọng tổng cước $E(TC)$ xét xem có nên thêm một người không?

TRẢ LỜI CÁC BÀI TẬP

CHƯƠNG 1

1.1 $\max 25x_t + 30x_c$

$$\frac{x_t}{200} + \frac{x_c}{140} = 40,$$

$$x_t \leq 6000,$$

$$x_c \leq 4000,$$

$$x_t, x_c \geq 0.$$

$$(x_t^*, x_c^*) = (6000, 1400), z^* = 192000.$$

1.2 $T := \max_j (x_j + x'_j + y_j + y'_j) \rightarrow \min$

$$w_j x_j - w_j x'_j + b_j y'_j - b_j y'_j = 10, j = 1, \dots, n,$$

$$\sum_1^n b_j y_j - \sum_1^n b_j y'_j \leq 10.$$

$$T \geq \sum_1^n y_j + \sum_1^n y'_j,$$

$$x_j, x'_j, y_j, y'_j \geq 10, j = 1, \dots, n.$$

1.3 $\max (8x_1 + 14x_2 + 11x_3 + 4y_1 + 12y_2 + 7y_3 + 4z_1 + 13z_2 + 9z_3),$

$$x_1 + x_2 + x_3 = 480,$$

$$y_1 + y_2 + y_3 = 400,$$

$$z_1 + z_2 + z_3 = 230,$$

$$x_2 + y_2 + z_2 \leq 420,$$

$$x_3 + y_3 + z_3 \leq 250,$$

$$x_i, y_i, z_i \geq 0, i = 1, 2, 3,$$

(Viết lại 3 ràng buộc đầu thành 6 ràng buộc bất đẳng thức)

1.4 8 bàn, 32 ghế.

1.5 150 mũ kiểu một và 200 kiểu hai

$$1.6. (a) \max \left(870 - 0,8 \sum_{i=1}^4 x_i \right)$$

$$x_1 \geq 75,$$

$$4x_1 + x_2 \geq 175,$$

$$4x_1 + 4x_2 + x_3 \geq 275,$$

$$4x_1 + 4x_2 + 4x_3 + x_4 \geq 375,$$

$$0 \leq x_i \leq 100, x_i \text{ nguyên}, i = 1, 2, 3, 4.$$

$$(b) \max \left(780 - 0,8 \sum_{i=1}^3 x_i \right)$$

$$0 \leq x_i \leq 100, x_i \text{ nguyên}, i = 2, 3, 4.$$

$$1.7 \min 2x_1 + 3(y^+ + y^-)$$

$$-x_1 + (y^+ - y^-) \leq \frac{11}{3},$$

$$x_1 \leq -2,$$

$$y^+ - y^- \geq \frac{10}{3},$$

$$x_1, y^+, y^- \geq 0.$$

$$(y = x_2 - \frac{10}{3})$$

$$1.8 (a) (x_1^*, x_2^*) = \left(-\frac{8}{3}, -\frac{5}{3} \right);$$

$$(b) (x_1^*, x_2^*) = (-10, -6).$$

1.10 (i) $a > 0, b > 0$; (ii) $\forall a, \forall b$; (iii) phủ định của $a > 0, b > 0$.

1.11 Nghiệm chấp nhận được $x \in D_1 \Leftrightarrow \lambda x \in D_2$, hoặc $D_1 = D_2 = \emptyset$.

Nghiệm tối ưu $x^* \in S_1 \Leftrightarrow \lambda x^* \in S_2$ hoặc $S_1 = S_2 = \emptyset$.

Mục tiêu tối ưu $z_2^* = \alpha \lambda z_1^*$.

1.12 (a) $(x_1^*, x_2^*, x_3^*) = (0, 1, 0)$, $z^* = 3$.

(b) Không giới nội dưới

(c) không chấp nhận được;

(d) $(x_1^*, \dots, x_4^*) = (1, 0, 0, 1)$, $z^* = 15$;

(e) $(x_1^*, \dots, x_5^*) = (3, 0, 0, 0, 0)$, $z^* = 3$.

(f) $(x_1^*, \dots, x_6^*) = \left(\frac{4000}{3}, 0, 0, 0, 0, \frac{8000}{3}\right)$.

1.13 (a) $z^* = -2$, $(x_1^*, x_2^*) = (1, 0)$;

(b) $z^* = 12$;

(c) $z^* = \frac{3}{2}$, $x^* = (0, 3, 0, \frac{1}{2}, \frac{3}{2})$;

(d) $z^* = -\frac{55}{2}$, $x^* = \left(\frac{3}{2}, \frac{7}{2}\right)$.

1.14 (a) $x^* = \left(\frac{3}{5}, \frac{6}{5}\right)$, $z^* = \frac{12}{5}$.

1.15 (b) $x^* = (2, 1)$, $z^* = -5$;

(c) không giới nội.

CHƯƠNG 2

2.3 (b) $x^* = (2, 1)$, $z^* = -5$.

CHƯƠNG 3

3.3 Đồ thị gồm n cặp nút. Véc tơ $(x_1, \dots, x_n)$ là kết quả thi đấu khi và chỉ khi nó là nguồn của bài toán dòng sau đây và có tồn tại nhu cầu $(d_1, \dots, d_n)$ để bài toán có dòng chấp nhận được (x_{ij}) . Ràng buộc của bài toán là

$$\sum_1^n x_i = \sum_1^n d_i = \frac{n(n-1)}{2}.$$

$$\sum_{j \neq i} x_{ij} = x_i, i = 1, \dots, n.$$

$$\sum_{j \neq i} x_{ji} = d_i = n - 1 - x_i, i = 1, \dots, n,$$

$$x_{ij} = 0 \text{ hoặc } 1.$$

3.4 x_{ij} là số tấn gỗ khai thác ở rừng i năm j .

$$\sum_{j=0}^{K-1} c_j \left(\sum_{i=1}^M x_{ij} + \dots + \sum_{i=1}^m x_{i1} - \sum_{j=0}^{K-1} d_j \right) \rightarrow \min,$$

$$\sum_{i=1}^M x_{ij} \leq u_j, j = 1, \dots, K,$$

$$\sum_{i=1}^M x_{ij} + \left(\sum_{i=1}^M x_{i,j-1} - d_{j-1} \right) + \dots + \left(\sum_{i=1}^M x_{i1} - d_1 \right) \geq d_j, j = 1, \dots, K,$$

$$0 \leq x_{ij} \leq a_{ij}.$$

3.5 (b) $(x_{32}, x_{21}, x_{41}, x_{54}, x_{74}, x_{76}, x_{87}) = (2, 1, 1, 2, 5, 0, 0)$, $z^* = 28$.

3.6 $(x_{12}, x_{14}, x_{23}, x_{45}, x_{36}, x_{47}, x_{78}) = (1, 1, 2, 1, 0, 6, 3)$, $z^* = 30$.

3.7 $(x_{41}, x_{23}, x_{43}) = (1, 1, 0)$, $z^* = -10$.

3.12 (BA, AC, CF, FD, DE, EG), độ dài cấp 5080.

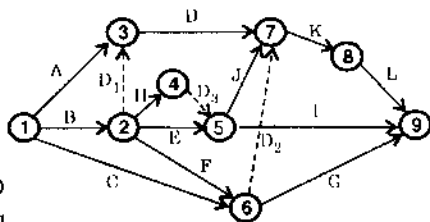
CHƯƠNG 4

4.1 Sơ đồ mạng lưới là

4.2 (a) Thêm cung giả (3,6).

(b) Thêm cung giả giữa ngọn của D và nút 7, rồi nối ngọn của D và nút 6 bằng một cung giả.

(c) Thêm cung giả giữa ngọn của C và nút 6, rồi nối ngọn của C và nút 3 bằng một cung giả.



4.3

biến cố	1	2	3	4	5	6	7	8	9	10
thời điểm sớm	0	6	3	5	10	10	11	14	13	20
thời điểm muộn	0	7	3	6	11	10	13	14	15	20
thời gian dự trữ	0	1	0	1	1	0	2	0	2	0

Đường găng $1 \rightarrow 3 \rightarrow 6 \rightarrow 8 \rightarrow 10$.

4.4 (a) $0 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 6$ và $0 \rightarrow 1 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 6$.

(b) $0 \rightarrow 2 \rightarrow 3 \rightarrow 6$.

4.5

Hoạt động	EC _{ij}	LS _{ij}	FF _{ij}
(1,2)	2	2	0
(1,3)	3	0	0
(2,4)	4	4	2
(3,4)	6	3	0
(3,5)	5	4	1
(4,5)	6	6	0
(4,6)	9	10	4
(4,7)	8	17	11
(5,6)	13	6	0
(5,7)	11	14	8
(6,7)	19	13	0

4.8 (a) Thay đổi về số công nhân là 13 với $0 \leq t < 2,7$ và với $2 < t < 3$, và là 12 với $3 \leq t < 4$.

(b) Thay đổi về số công nhân là 13 với $2 \leq t < 3$, 15 với $3 \leq t < 4$ và 12 với $4 \leq t < 6$.

4.9 Có 4 đường găng với độ dài 111 : (A, B, D, E, Q, R, S), (A, B, D, F, Q, R, S), (A, C, D, E, Q, R, S), (A, C, D, F, Q, R, S).

4.10 Số công nhân ít nhất được xác định bởi các hoạt động găng.

4.11 $t_e = 37$, $\sigma^2 = 9$.

4.12

Hoạt động	t_e	σ^2	Hoạt động	t_e	σ^2
(1,2)	2	0,11	(4,6)	3	1,00
(1,3)	3	1,00	(4,7)	2	0,11
(2,4)	2	0,11	(5,6)	7	0,11
(3,4)	3	2,78	(5,7)	5	1,78
(3,5)	2	1,36	(6,7)	6	0,44

biến cố	đường đi	$E(\mu_i)$	$\text{var}(\mu_i)$	$P(\mu_i, T_i)$
2	1, 2	2	0,11	1,000
3	1, 3	3	1,00	0,159
4	1, 3, 4	6	3,78	0,304
5	1, 3, 4, 5	6	3,78	0,500
6	1, 3, 4, 5, 6	13	3,89	0,987
7	1, 3, 4, 5, 6, 7	19	4,33	0,684

CHƯƠNG 5

5.2

d_j	500	630	200	230
s_i				
400	100	103	106	M
300	M	140	143	123
420	M	M	120	100
380	M	M	M	100
60	0	0	0	0

cung giả

5.3

(a)

10	0			10
	5			5
	0	4		4
		3	3	6
10	5	7	3	

(b)

1						giả	1
2	4	5	2	3			16
				5	2		7
					8		8
3	4	5	2	8	10		

 5.4 (a) $x_{33} \rightarrow x_{23} \rightarrow x_{22} \rightarrow x_{12} \rightarrow x_{11} \rightarrow x_{31} \rightarrow x_{33}$

 (b) x_{33} vào ở mức 5 và x_{31} ra.

 (c) $\bar{C}_{33} = -24$ và sự thăng giá trị mục tiêu là 120.

 5.5 Góc tây bắc : $x_{11} = 7, x_{21} = 3, x_{22} = 9, x_{32} = 1, x_{33} = 10, z = 94$;

 cực tiểu cước phí: $x_{13} = 7, x_{21} = 10, x_{23} = 2, x_{32} = 10, x_{33} = 1, z = 61$;

 xấp xỉ Vogel: $x_{11} = 7, x_{21} = 2, x_{23} = 10, x_{31} = 1, x_{32} = 10, z = 40$, tốt nhất.

 5.6 Phương pháp xấp xỉ Vogel cho nghiệm nhất phát tốt nhất, dẫn đến nghiệm tối ưu sau ba bước lặp : $x_{13} = 10, x_{22} = 20, x_{31} = 30, x_{42} = 30, x_{44} = 10, x_{51} = 30, x_{52} = x_{53} = 10, z = 820$.

 5.7 Có ba cách giải: (1) (đơn giản nhất) bỏ cột đầu và giảm s_4 đi 5; (2) đặt $c_{41} = -M$; (3) đặt $c_{11} = c_{21} = c_{31} = M$. Nghiệm tối ưu là $x_{14} = 10, x_{12} = x_{13} = 5, x_{24} = 10, x_{34} = 15, x_{41} = 5, x_{43} = 10; z = 55$.

5.8 (a) 1-5, 2-3, 3-2, 4-4, 5-1 với cước là 21 (bốn bước lặp nếu giải bằng phương pháp Hungary).

(b) 1-1, 2-2, 3-5, 4-4, 5-3 với cước là 11 (ba bước lặp);

(c) 1-4, 2-3, 3-2, 4-1, với cước phí 14 (năm bước lặp).

 5.9 Góc tây bắc : bốn bước lặp đến nghiệm tối ưu, $x_{11} = 3, x_{12} = 2, x_{22} = 1, x_{23} = 1, x_{33} = 1, x_{34} = 2$.

 Xấp xỉ Vogel: được ngay nghiệm tối ưu $x_{11} = 3, x_{12} = 0, x_{13} = 0, x_{14} = 2, x_{23} = 2, x_{32} = 3$.

5.12 Máy 5 thay máy 1.

5.13 $x_{12} = x_{22} = x_{23} = 10$, $x_{21} = 60$, $x_{31} = 15$, nút 3 thiếu 40 đơn vị hàng, $z = 595$ (năm bước lặp).

5.14 $x_{12} = 140$, $x_{21} = 120$, $x_{32} = x_{33} = 40$, $x_{43} = 60$, $x_{44} = 150$, $z = 2890$.

CHƯƠNG 6

6.1 $1431x_1 + 1746x_2 + 180x_3 + 165x_4 + 242x_5 = 6761$.

6.2 $\min z = 37x_1 + 38x_2 + 21x_3 + 20x_4 + 24x_5 + 25x_6$,

$$x_1 + x_2 + x_3 + x_5 + x_6 \geq 1,$$

$$x_1 + x_3 + x_4 + x_5 \geq 1,$$

$$x_1 + x_2 + x_4 + x_5 + x_6 \geq 1,$$

$$x_1 + x_2 \geq 1,$$

$$x_2 + x_3 + x_4 + x_6 \geq 1,$$

$$x_j = 0 \text{ hoặc } 1, j = 1, \dots, 6.$$

6.3 $(x_1 \leq 1 \text{ và } x_2 \leq 2)$ hoặc $(x_1 + x_2 \leq 3 \text{ và } x_1 \geq 2)$. Dạng tương đương là $x_1 - My \leq 1$, $x_2 - My \leq 2$, $x_1 + x_2 - M(1 - y) \leq 3$, $x_1 + M(1 - y) \leq 2$, $y = 0$ hoặc 1. Nghiệm tối ưu là $(x_1, x_2) = (1, 2)$.

6.4 (a) Nút 0: $(2, 25; 2, 25)$, $z = 11, 25$; nút 1 $(x_1 \leq 2)$: $(2; 2, 4)$, $z = 11, 2$; nút 2 $(x_1 \leq 2, x_2 \geq 3)$: $(1, 3)$; $z = 11$ (cận dưới); nút 3 $(x_1 \geq 3)$: $(3; 1, 5)$; $z = 10, 5$; nút 4 $(x_1 \leq 2, x_2 \leq 2)$: $(2, 2)$; $z = 10$ (chạm đáy). Nghiệm tối ưu ở nút 2.

6.6 Hai nghiệm tối ưu: $(x_1 = 14, x_4 = 19, \text{ hàng khác} = 0)$ và $(x_1 = 15, x_3 = 1, x_4 = 17, \text{ hàng khác} = 0)$; $z = 151$.

6.7 $(x_1, x_2) = (2, 1)$ là nghiệm tối ưu.

6.9

việc	1	2	3	4	5
người	1	3	2	4	5

6.10 (a) $(x_1, \dots, x_5) = (0, 0, 1, 1, 1)$, $z = 6$.

6.11 (a) $x_1 = x_2 = x_3 = 1$, $z = 7$.

(b) không có nghiệm chấp nhận được.

(c) $(x_1, \dots, x_4, x_5) = (1, 1, 1, 1, 0)$, $z = 95$.

6.12 $(x_1, x_2, x_3) = (0, 6, 1)$, $z = 12$.

6.13

	nhịệm làm tròn	nhịệm nguyên
x_1	2 hoặc 3	2
x_2	1	1
x_3	6	6
z	26 hoặc 30	26

6.14 (b) $x_1 = 5$, $x_2 = 2,75$, $x_3 = 3$, $z = 26,75$.

CHƯƠNG 7

7.1 $EC_T = \frac{n}{T} \left(C_1 \sum_{t=1}^{T-1} p_t + C_2 \right)$, quyết định cần chọn là $T^* = 3$ với $EC_{T^*} = 366,7$.

7.2 Quyết định sản xuất chip.

7.3 $T^* = 3$ với $EC_{T^*} = 290$.

7.4 Kỳ vọng phí tổn cho khách A, B và C là 190, 120 và 205 USD, tương ứng. B được ưu tiên nhất.

7.5 Mua 200 bánh, kỳ vọng doanh thu tối ưu là 369,5 ngàn đồng.

7.6 $\min (EC_T + \text{Var } C_T) = n \left\{ \left(\frac{C_1}{T} + \frac{C_2}{T^2} \right) \sum_{t=1}^{T-1} p_t - \left(\frac{C_1}{T} \right)^2 \sum_{t=1}^{T-1} p_t^2 + \frac{C_2}{T} \right\}$. $T^* = 1$ với $EC_1 + \text{Var } C_1 = 500$.

7.7 $1 \geq 2$; $2 \leq t \leq 4$.

7.8 $1 \geq 4$.

7.9 Kỳ vọng phí tổn của A, B và C là 263,7; 93,3 và 168,2 USD. B có ưu tiên nhất.

7.11

(a) tới 80 triệu. (b) bảng bên:

		trạng thái tự nhiên	
		thắng	thua
Dự báo	thắng	0,818	0,181
	thua	0,333	0,667

(c) Thuê chuyên gia. Tổ chức quyền góp nếu chuyên gia dự báo thắng giải. Kỳ vọng giải thưởng là 105 triệu (tức là tiền quyền được).

7.12 130 bánh.

7.13 Lắp dây truyền lớn ngay, kỳ vọng của lãi trong 10 năm là 325 triệu đồng.

7.14 (a) a_4 ; (b) a_2 ; (c) a_2 ; (d) a_4 .

7.15 (a) chiến lược 2; (b) chiến lược 1.

7.17 (a) $\underline{v} = \bar{v} = v = 1$, điểm yên ngựa $a_{32} = 1$;

(b) $\underline{v} = \bar{v} = v = 1$, điểm yên ngựa $a_{33} = 1$;

(c) $\underline{v} = 7$, $\bar{v} = 9$, không có điểm yên ngựa;

(d) $\underline{v} = \bar{v} = v = 3$, điểm yên ngựa $a_{21} = 3$;

(e) $\underline{v} = \bar{v} = v = 1$, điểm yên ngựa $a_{11} = a_{13} = a_{31} = a_{33} = 1$;

(g) $\underline{v} = \bar{v} = v = -2$, điểm yên ngựa $a_{22} = a_{24} = -2$.

7.18 (a) $p \geq 5$ và $q \leq 5$; (b) $p \leq 7$ và $q \geq 7$.

7.19 (a) Chiến lược tối ưu của B là $(y_1^*, 1 - y_1^*) = (\frac{2}{3}, \frac{1}{3})$, $\bar{v} = \frac{8}{3}$; Chiến lược tối ưu của A

là $(x_1^*, \dots, x_4^*) = (\frac{1}{3}, 0, \frac{2}{3}, 0)$, $\underline{v} = \frac{8}{3}$.

(b) Chiến lược tối ưu của A là $(x_1^*, 1 - x_1^*) = (\frac{1}{2}, \frac{1}{2})$ với $\underline{v} = \frac{5}{2}$.

Chiến lược tối ưu của B là $(y_1^*, \dots, y_4^*) = (0, \frac{1}{2}, \frac{1}{2}, 0)$, $\bar{v} = \frac{5}{2}$.

(c) Chiến lược tối ưu của A là $(x_1^*, 1 - x_1^*) = (\frac{1}{2}, \frac{1}{2})$, của B là $(y_1^*, \dots, y_4^*) = (0, 0, \frac{13}{20}, \frac{7}{20})$, $v = \frac{1}{2}$.

7.21 Chiến lược tối ưu của B và A là $(\frac{14}{45}, \frac{11}{45}, \frac{20}{45})$, $(\frac{20}{45}, \frac{11}{45}, \frac{14}{45})$, $v = -\frac{29}{45}$.

7.22 Chiến lược tối ưu của đại tá là $(\frac{1}{5}, \frac{3}{5}, \frac{1}{5})$ và của đối phương là $(\frac{1}{3}, \frac{1}{5}, 0, \frac{7}{15})$.

CHƯƠNG 8

8.1 2090,5.

8.2 783, 2.

8.5

(a)

IV/90	7,167				
I/91	7,233	I/92	7,267	I/93	9,633
II/91	7,733	II/92	8,067	II/93	10,367
III/91	7,433	III/92	8,700	III/93	10,433
IV/91	7,500	IV/92	9,467	IV/93	10,267
				IV/94	7,967
				I/94	9,267
				II/94	8,867
				III/94	8,267

(b) 1,184.

8.6 (a) 410,333 = 17,630t;

(b) 604,267;

(c)

3	431,6	6	445,220	9	468,880
4	434,84	7	452,098	10	478,992
5	439,356	8	460,089	11	490,193

442 TRẢ LỜI CÁC BÀI TẬP

(d)

3	465,52	6	517,405	9	567,864
4	483,075	7	534,236	10	585,671
5	500,196	8	551,212	11	604,329

CHƯƠNG 9

9.1 ma trận chuyển

	0	1	2	3	4
0	0,1	0,9	0	0	0
1	0,15	0	0,85	0	0
2	0,22	0	0	0,78	0
3	0,50	0	0	0	0,5
4	1	0	0	0	0

xác suất duy trì :

$$\pi_0 = 0,281, \pi_1 = 0,253, \pi_2 = 0,215$$

$$\pi_3 = 0,167, \pi_4 = 0,084$$

9.2 (a)

giá hôm nay	kỳ vọng giá ngày mai	kỳ vọng lãi
1	1,75	0,75
2	2,1	0,1
3	3	0
4	3,9	-0,1
5	4,25	-0,75

(b) Xác suất giá sau hai ngày, biết giá hôm nay là 1 ($P\{x_0 = 1\} = 1$)

$$P\{X_2 = j\} = \sum_{i=1}^5 P\{X_2 = j \mid X_1 = i\} P\{X_1 = i \mid X_0 = 1\};$$

$$P\{X_2 = 1\} = 0,425; P\{X_2 = 2\} = 0,250; P\{X_2 = 3\} = 0,140;$$

$$P\{X_2 = 4\} = 0,0925; P\{X_2 = 5\} = 0,0925; \text{ kỳ vọng lãi } 1,1775.$$

$$9.3 \text{ (a) } P = \begin{bmatrix} 0 & 0,5 & 0 & 0 & 0,5 \\ 0,5 & 0 & 0,5 & 0 & 0 \\ 0 & 0,5 & 0 & 0,5 & 0 \\ 0 & 0 & 0,5 & 0 & 0,5 \\ 0,5 & 0 & 0 & 0,5 & 0 \end{bmatrix} \quad p^{(2)} = \begin{bmatrix} 0,5 & 0 & 0,25 & 0,25 & 0 \\ 0 & 0,5 & 0 & 0,25 & 0,25 \\ 0,25 & 0 & 0,5 & 0 & 0,25 \\ 0,25 & 0,25 & 0 & 0,5 & 0 \\ 0 & 0,25 & 0,25 & 0 & 0,5 \end{bmatrix}$$

$$p^{(5)} = \begin{bmatrix} 0,062 & 0,312 & 0,156 & 0,156 & 0,312 \\ 0,312 & 0,062 & 0,312 & 0,156 & 0,156 \\ 0,156 & 0,312 & 0,062 & 0,312 & 0,156 \\ 0,156 & 0,156 & 0,312 & 0,062 & 0,312 \\ 0,312 & 0,156 & 0,156 & 0,312 & 0,062 \end{bmatrix}$$

(b) $\pi_0 = \dots \pi_4 = 0,2$.

9.4 (a) một lớp, lặp.

9.6 (a) $\pi_0 = 0,182$, $\pi_1 = 0,285$, $\pi_2 = 0,368$, $\pi_3 = 0,165$.

(b) 31,42.

9.7 (a) Trạng thái có thể là xe bị trầy xước và không bị trầy xước.

(c) Nếu xe không bị trầy xước, tối ưu là chiếm một chỗ miễn phí, nếu xe bị trầy xước thì chữa.

9.8 (c) Sách lược tối ưu : ở trạng thái 0 thì giao bóng thật mạnh, 1 thì giao đảm bảo.

9.10 (a) $\min Z = 4,5 y_{02} + 5 y_{03} + 50 y_{14} + 9 y_{15}$.

$$y_{01} + y_{02} + y_{03} + y_{14} + y_{15} = 1,$$

$$y_{01} + y_{02} + y_{03} - \left(\frac{9}{10} y_{01} + \frac{49}{50} y_{02} + y_{03} + y_{14} \right) = 0,$$

$$y_{14} + y_{15} - \left(\frac{1}{10} y_{01} + \frac{1}{50} y_{02} + y_{15} \right) = 0,$$

mọi $y_{ik} \geq 0$.

9.11 $\min Z = -\frac{1}{8} y_{01} + \frac{7}{24} y_{02} + \frac{1}{2} y_{11} + \frac{5}{12} y_{12}$.

$$y_{01} + y_{02} - \left(\frac{3}{8} y_{01} + y_{11} + \frac{7}{8} y_{02} + y_{12} \right) = 0,$$

$$y_{11} + y_{12} - \left(\frac{5}{8}y_{01} + \frac{1}{8}y_{02}\right) = 0,$$

$$y_{01} + y_{02} + y_{11} + y_{12} = 1,$$

$$\text{mọi } y_{ik} \geq 0.$$

9.13 Nếu xe chưa trải suốt : chiếm một chỗ miễn phí, nếu xe đã trải : đi chữa.

9.16 Không đồng ý giá 600, đồng ý nếu được đề nghị 800 hoặc 1000.

$$\mathbf{9.17} \quad \min Z = 60(y_{01} + y_{11} + y_{21}) - 600y_{02} - 800y_{12} - 1000y_{22},$$

$$y_{01} + y_{02} - 0,95 \cdot \frac{5}{8}(y_{01} + y_{11} + y_{21}) = \frac{5}{8},$$

$$y_{11} + y_{12} - 0,95 \cdot \frac{1}{4}(y_{01} + y_{11} + y_{21}) = \frac{1}{4},$$

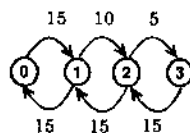
$$y_{21} + y_{22} - 0,95 \cdot \frac{1}{8}(y_{01} + y_{11} + y_{21}) = \frac{1}{8},$$

$$\text{mọi } y_{ik} \geq 0.$$

9.18 Sau ba bước lập ta được ngay sách lược tối ưu của Bài 9.16.

9.19 Từ bước 1 đến bước 3: không làm gì nếu máy ở trạng thái 0 hoặc 1, đại tu nếu ở trạng thái 2 và thay nếu ở 3. Ở bước 4 : không làm gì nếu máy ở 0, 1, 2 và thay nếu máy ở trạng thái 3.

CHƯƠNG 10



10.1 (a) 0,135 , (b) 0,270, (c) 0,0527.

10.3 (a) Hình bên.

$$c) P_0 = \frac{9}{26}, P_1 = \frac{9}{26}, P_2 = \frac{3}{13}, P_3 = \frac{1}{13},$$

$$d) W = 0,11 \text{ giờ.}$$

$$\mathbf{10.4} \quad (b) P_0 = \frac{2}{5}, P_n = \left(\frac{3}{5}\right)\left(\frac{1}{2}\right)^n, (c) L = \frac{6}{5}, L_q = \frac{3}{5}, W = \frac{1}{25}, W_q = \frac{1}{50}.$$

10.5 $\frac{31}{32}$.

10.6 (a) 0,429, (b) 0,154, (c) 0,072.

10.7 (a) ba máy, (b) 0,268.

10.8 (a) W_q (phân bố mũ) $= 2W_q$ (hằng số) $= \frac{8}{5}W_q$ (Erlang).

(b) W_q (mới) $= \frac{1}{2}W_q$ (cũ) và L_q (mới) $= L_q$ (cũ) cho cả 3 phân bố.

10.9 cách cũ $L = 1$, cách mới $L = 13/16$.

10.10 (a) $W = \frac{1}{2}$; (b) $W_1 = 0,20$, $W_2 = 0,35$, $W_3 = 1,10$;

(c) $W_1 = 0,125$, $W_2 = 0,3125$, $W_3 = 1,250$.

10.12 (a) $P_0 = 8/15$, $P_1 = \frac{4}{15}$, $P_2 = \frac{2}{15}$, $P_3 = \frac{1}{15}$;

(b) Lãi một giờ 280 (ngàn đồng), nếu đủ chỗ thì lãi 300 (ngàn);

(c) Lãi trung bình một giờ $= \frac{\left(\frac{\lambda}{\mu}\right) \left[1 - \left(\frac{\lambda}{\mu}\right)^N\right]}{1 - \left(\frac{\lambda}{\mu}\right)^{N+1}}$, với $\frac{\lambda}{\mu} = \frac{1}{2}$. Với $N = 1, \dots, 6$ ta có

N	1	2	3	4	5	6
lãi	200	256,8	280	290,3	295,2	297,6

(lãi trung bình một giờ $= 600 (1 - P_0)$ /giờ).

10.13 (a) $F(CD) = 16$, (b) $E(CD) = 26,5$.

10.14 (a) Mô hình 2 với $s = 1$; (b) nên thêm người phụ.

10.15 (a) hai người, (b) ba người.

10.16 Một người phục vụ: $E(TC) = 62,48$ ngàn/giờ, hai người: $E(TC) = 9,295$ ngàn/giờ. Cần hai người phục vụ.

PHỤ LỤC (CÁC BẢNG TÍNH)

AI : $P\{x \geq x_e + \sigma K\alpha\}$, x có phân bố chuẩn, σ là độ lệch chuẩn, x_e là kỳ vọng

K α	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0,0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0,1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0,2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0,3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0,4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0,5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0,6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0,7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0,8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0,9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1,0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1,1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1,2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1,3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1,4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1,5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1,6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1,7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1,8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1,9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2,0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2,1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2,2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2,3	.0107	.0104	.0102	.0099	.0096	.0093	.0091	.0088	.0086	.0084
2,4	.0082	.0079	.0077	.0075	.0073	.0071	.0069	.0067	.0065	.0063
2,5	.0062	.0060	.0058	.0057	.0055	.0053	.0052	.0050	.0049	.0048
2,6	.0046	.0045	.0044	.0042	.0041	.0040	.0039	.0037	.0036	.0035
2,7	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0028	.0027	.0026
2,8	.0025	.0024	.0024	.0023	.0022	.0021	.0021	.0020	.0019	.0019
2,9	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014	.0013

K α	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9
3	.00135	.0 ⁰ 968	.0 ⁰ 687	.0 ⁰ 483	.0 ⁰ 337	.0 ⁰ 233	.0 ⁰ 159	.0 ⁰ 108	.0 ⁰ 723	.0 ⁰ 481
4	.0 ⁰ 317	.0 ⁰ 207	.0 ⁰ 133	.0 ⁰ 854	.0 ⁰ 541	.0 ⁰ 340	.0 ⁰ 211	.0 ⁰ 130	.0 ⁰ 793	.0 ⁰ 479
5	.0 ⁰ 287	.0 ⁰ 170	.0 ⁰ 996	.0 ⁰ 579	.0 ⁰ 333	.0 ⁰ 190	.0 ⁰ 107	.0 ⁰ 599	.0 ⁰ 332	.0 ⁰ 182
6	.0 ⁰ 987	.0 ⁰ 530	.0 ⁰ 282	.0 ⁰ 149	.0 ⁰ 777	.0 ⁰ 402	.0 ⁰ 206	.0 ⁰ 104	.0 ⁰ 523	.0 ⁰ 260

A 2 : giá trị $t_{v,\beta}$ để $P\{t \geq t_{v,\beta}\} = \beta$, t có phân bố Student với v bậc tự do

β	0,40	0,25	0,10	0,05	0,025	0,01	0,005	0,0025	0,001	0,0005
v										
1	0,325	1,000	3,078	6,314	12,706	31,821	63,657	127,32	318,31	636,62
2	0,289	0,816	1,886	2,920	4,303	6,965	9,925	14,089	22,327	31,598
3	0,277	0,765	1,638	2,353	3,182	4,541	5,841	7,453	10,214	12,924
4	0,271	0,741	1,533	2,132	2,776	3,747	4,604	5,598	7,173	8,610
5	0,267	0,727	1,476	2,015	2,571	3,365	4,032	4,773	5,893	6,869
6	0,265	0,718	1,440	1,943	2,447	3,143	3,707	4,317	5,208	5,959
7	0,263	0,711	1,415	1,895	2,365	2,998	3,499	4,029	4,785	5,408
8	0,262	0,706	1,397	1,860	2,306	2,896	3,355	3,833	4,501	5,041
9	0,261	0,703	1,383	1,833	2,262	2,821	3,250	3,690	4,297	4,781
10	0,260	0,700	1,372	1,812	2,228	2,764	3,169	3,581	4,144	4,587
11	0,260	0,697	1,363	1,796	2,201	2,718	3,106	3,497	4,025	4,437
12	0,259	0,695	1,356	1,782	2,179	2,681	3,055	3,428	3,930	4,318
13	0,259	0,694	1,350	1,771	2,160	2,650	3,012	3,372	3,852	4,221
14	0,258	0,692	1,345	1,761	2,145	2,624	2,977	3,326	3,787	4,140
15	0,258	0,691	1,341	1,753	2,131	2,602	2,947	3,286	3,733	4,073
16	0,258	0,690	1,337	1,746	2,120	2,583	2,921	3,252	3,686	4,015
17	0,257	0,689	1,333	1,740	2,110	2,567	2,898	3,222	3,646	3,965
18	0,257	0,688	1,330	1,734	2,101	2,552	2,878	3,197	3,610	3,922
19	0,257	0,688	1,328	1,729	2,093	2,539	2,861	3,174	3,579	3,883
20	0,257	0,687	1,325	1,725	2,086	2,528	2,845	3,153	3,552	3,850
21	0,257	0,686	1,323	1,721	2,080	2,518	2,831	3,135	3,527	3,819
22	0,256	0,686	1,321	1,717	2,074	2,508	2,819	3,119	3,505	3,792
23	0,256	0,685	1,319	1,714	2,069	2,500	2,807	3,104	3,485	3,767
24	0,256	0,685	1,318	1,711	2,064	2,492	2,797	3,091	3,467	3,745
25	0,256	0,684	1,316	1,708	2,060	2,485	2,787	3,078	3,450	3,725
26	0,256	0,684	1,315	1,706	2,056	2,479	2,779	3,067	3,435	3,707
27	0,256	0,684	1,314	1,703	2,052	2,473	2,771	3,057	3,421	3,690
28	0,256	0,683	1,313	1,701	2,048	2,467	2,763	3,047	3,408	3,674
29	0,256	0,683	1,311	1,699	2,045	2,462	2,756	3,038	3,396	3,659
30	0,256	0,683	1,310	1,697	2,042	2,457	2,750	3,030	3,385	3,646
40	0,255	0,681	1,303	1,684	2,021	2,423	2,704	2,971	3,307	3,551
60	0,254	0,679	1,296	1,671	2,000	2,390	2,660	2,915	3,232	3,460
120	0,254	0,677	1,289	1,658	1,980	2,358	2,617	2,860	3,160	3,373
∞	0,253	0,674	1,282	1,645	1,960	2,326	2,576	2,807	3,090	3,291

HƯỚNG DẪN TRA CỨU

bài toán bố nhiệm (assignment problem) 176

bài toán bổ trợ (auxiliary problem) 22

bài toán chất hàng (cargo-loading problem) 287

bài toán chọn (choice problem) 177

bài toán chuyển vận (transshipment problem) 91

bài toán dòng cực đại (maximum flow problem) 125

bài toán dòng trên mạng (network flow problem) 91

bài toán điều xe 220

bài toán đối ngẫu (dual problem) 70, 72

bài toán đường ngắn nhất (shortest path problem) 136

bài toán hôn nhân (marriage problem) 144

bài toán kế hoạch sản xuất (production planning problem) 5

bài toán thừa (scarce problem) 46

bài toán vận tải (transportation problem) 9

bài toán vận tải cân bằng (balanced) 188

bài toán vận tải có đường cấm 180

bài toán vận tải Hitchcock 176

bài toán vận tải kèm chế biến trung gian 181

bài toán vận tải nhanh 211

bài toán với tài năng hạn chế (capacitated problem) 97

bài toán với tài năng không hạn chế (uncapacitated problem) 97

bảng vận tải (transportation table) 185

biến (variable) 10

biến bù (slack variable) 12

biến chia nhánh (branching variable) 262

biến cố (event) 146

biến cố chắc chắn (certain event) 297

biến cố gang (critical event) 152

biến cố tách biệt (exclusive events) 297

biến cơ sở (basic variable) 19

biến đối ngẫu (dual variable) 73

biến không cơ sở (nonbasic variable) 19

biến quyết định (decision variable) 10, 168

biến quyết định liên tục (continuous decision variable) 319

biến ra (leaving variable) 21

biến vào (entering variable) 21

biểu đồ Gantt (Gantt bar chart) 145

biểu đồ thời gian (time chart) 155

cành (branch) 298

cạnh (edge) 47, 92

cắt (cut) 126

cắt hỗn hợp (mixed cut) 282

cắt mạnh hơn (stronger cut) 280

cận trên (upper bound) của phân bố 161

cấp của nút, của đồ thị (degree) 92

cây (tree) 93

cây bao trùm (spanning tree) 94

cây quyết định (decision tree) 298

cây tối đại (spanning tree) 94

chạc (fork) 298

chạc quyết định (decision fork) 298

chạc tình huống (chance fork) 298

chỉ số lạc quan (index of optimism) 307

chỉ số mùa (seasonal index) 332

chia nhánh (branching) 264

chiến lược bị trội (dominated strategy) 309

- chiến lược hỗn hợp* (mixed strategy) 307
chiến lược thuần túy (pure strategy) 311
chọn quyết định bất định (decision making uncertainty) 292
chọn quyết định có mạo hiểm (decision making under risk) 291
chọn quyết định có thử nghiệm (decision making with experimentation) 292
chọn quyết định không thử nghiệm (decision making without experimentation) 292
chọn quyết định trong hoàn cảnh tất định (decision making under certainty) 291
chu trình (loop hoặc cycle) 92, 186
chu trình chưa bão hòa (unsaturated cycle) 117
chu trình có hướng (directed cycle) 93
chuỗi phương tiện với hàng vô hạn (infinite queuer in series) 413
công thức Little (Little's formula) 389
cột xoay (pivot column) 29
cơ sở (basis) 36
cung (arc) 92, 298
cung lùi (backward arc) 93
cung tiến (forward arc) 93
cung vô hướng (undirected arc) 92
cực điểm (crash point) 167
cước phí (cost) 95
cước phí chuẩn (normal cost) 168
cước phí của chu trình (cost of a cycle) 100
cước phí lưu kho (storage cost) 179
cước phí cực điểm (crash cost) 168
cước phí sau (subsequent cost) 366
cước phí thu gọn (reduced cost) 106
cước phí trung bình thực tế (actual average cost) 359
cước phí tức thời (immediate cost) 366
dạng chính tắc (canonical form) 11
dạng chuẩn (normal form hoặc standard form) 11
dạng ma trận của bảng đơn hình (matrix form) 48
dạng tích của ma trận ngược (product form of an inverse matrix) 50
dạng thừa số eta (eta factorization) 52
dây chuyển (path) 186
dòng (flow) 96
dòng chờ (waiting time) 384
dòng chấp nhận được (feasible flow) 96
dòng tinh (net flow) 120
dung lượng của cdt (capacity) 126
dự án (project) 145
dự án lớn (large scale project) 145
dự báo làm trơn theo lũy thừa (exponential smoothing forecasting) 329
dự báo theo giá trị cuối (last-value forecasting) 328
dự báo theo trung bình (average forecasting) 328
dự báo theo trung bình chuyển động (moving-average forecasting) 328
đa diện lồi (polytope) 13
đặt cận (bomding) 263
đặt kế hoạch (planning) 145
đầu vào Poisson (Poisson input) 387, 394
đích (destination) 176
điểm chuẩn (normal point) 167
điểm cực biên (extreme point) 37
điểm diện (enumeration) 252
điểm diện ẩn (implicit enumeration) 254
điểm hút (sink) 95
điểm phát 185
điểm thu 185
điểm yên ngựa (saddle point) 310
điều hành thực hiện (scheduling) 145

- p>điều kiện độ lệch bù (complementary slackness) 80
p>điều kiện độ lệch bù chặt (strict complementary slackness condition) 90
p>điều kiện cân bằng thu phát 187
p>điều kiện duy trì (transient condition) 389
p>điều kiện quá độ (steady state condition) 389
p>đỉnh (vertex) 37
p>đỉnh kề (adjacent vertex) 44
p>đỉnh suy biến (degenerate vertex) 43
p>định lý giới hạn trung tâm (central-limit theorem) 163
p>định lý cân bằng (equilibrium theorem) 80
p>độ lệch bù 80
p>độ lệch (tổng cộng) của phương án 236
p>độ lệch chuẩn (standard deviation) 161, 293
p>đồ thị có hướng (directed graph) 92
p>đồ thị liên thông (connected graph) 92
p>đồ thị vô hướng (undirected graph) 92
p>đối ngẫu mạnh (strong duality) 76
p>đối ngẫu yếu (weak duality) 75
p>đồng minh (coalition) 319
p>đường cong thời gian cước phí (time-cost curve) 167
p>đường đi (path) 92, 186
p>đường đi chưa bão hoà (unsaturated path) 130
p>đường đi tăng dòng (augmenting path) 129
p>đường gãy (critical path) 152
p>đường giá trị tiền (monetary value line) 302
p>đường hồi quy (regression line) 338
p>đường mức (level curver) 14
p>đường tới hạn (critical path) 152
p>giá trị của trò chơi (value) 309
p>giá trị đã trơn (smoothed value) 330
p>giá trị mục tiêu tối ưu (optimal objective value) 14
p>giá trị thử nghiệm (value of experimentation) 298
p>hàm công dụng (thay tiền) (utility function (for money)) 300
p>hàm đặc biệt beta (beta function) 162
p>hàm đặc biệt gamma (gamma function) 162
p>hàm Lagrange (Lagrange function) 70
p>hàm mục tiêu (objective function) 8
p>hàng nguồn (source row) 277
p>hàng xoay (pivot row) 30
p>hằng số làm trơn (smoothing constant) 329
p>hệ số chống mạo hiểm (risk aversion factor) 294
p>hệ số gãy (critical coefficient) 154
p>hệ số giảm (của cước phí) (discount factor) 375
p>hệ số sử dụng (utilization factor) 394
p>hệ thống dịch vụ (service system) 384
p>hệ thống dịch vụ ngẫu nhiên (stochastic service system) 384
p>hệ thống kiểu PERT (PERT-type system) 146
p>hoàn lưu (circulation) 99
p>hoàn lưu đơn giản (simple circulation) 100
p>hoạt động (activity) 145
p>hoạt động gãy (critical activity) 152
p>hoạt động giả (dummy activity) 147
p>hồi quy tuyến tính (linear regression) 336
p>kết quả (out come) của quá trình ngẫu nhiên 345
p>khách hàng tiềm ẩn (potential customer) 385
p>khoảng nói lòng đồng thời (simultaneous tolerance interval) 341
p>khoảng tiên định (prediction interval) 341
p>khoảng tin cậy cho giá trị của biến ngẫu nhiên 340
p>khoảng tin cậy cho kì vọng 339
p>không gian đóng (closed space) 12

không gian mẫu (sample space) 345
 không gian trạng thái (state space) 345
 không suy biến (nondegenerate) 36
 kì vọng cước phí trung bình (dài hạn)
 (long-run expected average cost) 358
 kì lục (incumbent) 263
 kĩ thuật xem xét và đánh giá dự án
 (PERT, tức project evaluation and review
 technique) 145
 kĩ thuật phán đoán (judgmental technique)
 325
 kiểm tra (controlling) 145
 lá (leaf) 93
 Lagrange nói lỏng (Lagrangian relaxation)
 265
 Lagrangean 70
 làm chạm đáy (fathoming) 263
 lí thuyết xếp hàng (queueing theory) 384
 lí thuyết trò chơi (game theory) 308
 lỗ hổng đối ngẫu (duality gap) 78
 lõm (convex) 13
 LP - nói lỏng (LP-relaxation) 177
 lượng phát 185
 lượng thu 185
 ma trận hối tiếc (regret matrix) 306
 ma trận nổi (incidence matrix) 99
 ma trận nổi thu gọn (trimcated incidence
 matrix) 101
 ma trận tương đương 225
 ma trận truyền (transition matrix) 344
 mạng (network) 95
 mạng dịch vụ (service network) 412
 mạng Jackson (Jackson's network) 415
 mạng thặng dư (residual network) 118
 mạng xếp hàng (queueing network) 412
 mặt cắt phân số (fractional cut) 277
 mặt mức (level surface) 14

miễn chấp nhận được, miễn ràng buộc
 (feasible region) 12
 mô hình khuynh hướng tuyến tính (linear
 trend model) 326
 mô hình khuynh hướng tuyến tính biến
 động theo mùa (linear trend with sea-
 sonal-effects model) 326
 mô hình $M/E_k/s$ 407
 mô hình $M/D/s$ 407
 mô hình $M/M/1$ 394
 mô hình $M/M/1/K$ 399
 mô hình $M/M/s$ 395
 mô hình $M/M/1$ với nguồn hữu hạn 401
 mô hình $M/M/s$ với nguồn hữu hạn 402
 mô hình $M/M/s/K$ 400
 mô hình mức độ quan hệ (degree of asso-
 ciation model) 337
 mô hình mức hằng số (constant-level
 model) 326
 mô hình mức hằng số biến động theo
 mùa (constant-level with seasonal-effects
 model) 326
 mô hình ngẫu nhiên (probabilistic model)
 160
 mô hình tất định (deterministic model) 160
 mô hình theo mùa (seasonal model) 326
 mô hình ưu tiên không nhường ngay
 (nonpreemptive priorities model) 409
 mô hình ưu tiên nhường ngay (preemptive
 priorities model) 408
 mô hình vận tải đóng (closed transportation
 model) 10
 mô hình vận tải mở (open transportation
 model) 10
 mô hình xác suất (probabilistic model) 160
 mô hình xếp hàng (queueing model) 387
 mức đã trơn (smoothed level) 330
 mức đòi hỏi (aspiration level) 294
 mức độ quan hệ (degree of association) 337

- p>
- thử nghiệm cây*
- (tree solution) 102
- thử nghiệm chấp nhận được* (feasible solution) 8
- thử nghiệm chấp nhận được tối ưu* (optimal feasible solution) 14
- thử nghiệm cơ sở* (basic solution) 19, 36
- thử nghiệm cơ sở không suy biến* (nondegenerate basic solution) 36
- thử nghiệm cơ sở suy biến* (degenerate basic solution) 36
- thử nghiệm dạng tích* (product form solution) 413
- thử nghiệm không ổn định* (unstable solution) 311
- thử nghiệm ổn định* (stable solution) 310
- thử nghiệm tối ưu* (optimal solution) 14, 308
- thử nghiệm tối ưu của trò chơi* 308
- nguồn, nguồn cung* (supply, source) 95
- nguồn năng của nút* (node capacity) 98
- nguồn vào* (input source) 385
- nguyên liệu* (resource) 145
- nguyên lý không đủ lý do* (principle of insufficient reason) 304
- nguyên lý tối ưu* (principle of optimality) 137
- nhân tố hướng* (trend factor) 330
- nhân tử Lagrange* (Lagrange multiplier) 70
- nhu cầu* (demand) 95
- nút* (node) 92
- nút cầu* (demand node) 176
- nút cung* (supply node) 175
- nút cung giả* (dummy source) 178
- nút đã đánh dấu* (labeled node) 130
- nút gốc* (root node) 102, 136
- nút nguồn* (source) 175
- nửa không gian đóng* (closed halfspace) 12
- ô chọn* 186
- ô độc lập* (của ma trận) 226
- ô không sử dụng* 186
- ô loại* 186
- ô sử dụng* 186
- pha điều hành* (scheduling phase) 146
- pha I, II* (phase I, II) 25
- phân bố beta* (beta distribution) 161
- phân bố chuẩn* (normal distribution) 163
- phân bố Erlang* (Erlang distribution) 408
- phân bố gamma* (gamma distribution) 408
- phân bố mũ* (exponential distribution) 385
- phân bố nhị thức* (binomial distribution) 355
- phân bố Poisson* (Poisson distribution) 355
- phân bố Student* (Student's distribution) 339
- phân tích chuỗi thời gian* (time series analysis) 325
- phân tích hồi quy* (regression analysis) 325
- phần tử trục* (pivot) 30
- phép toán trên hàng* (row operation) 18
- phép xoay* (pivot) 22
- phép xoay suy biến* (degenerate pivot) 31
- phiếu điều tra* (questionnaire) 326
- phương án* (alternative) 8, 10
- phương án chấp nhận được* (feasible alternative) 8
- phương án cực biên không suy biến* (nondegenerate extreme alternative) 190
- phương án cực biên suy biến* (degenerate extreme alternative) 190
- phương pháp ba ước lượng PERT* (PERT three estimate method) 166
- phương pháp Box-Jenkins* (Box-Jenkins method) 335
- phương pháp cực tiểu cước phí* (least-cost method) 204
- phương pháp cực tiểu cước phí theo hàng (cột)* 206
- phương pháp cực tiểu cước phí toàn bảng* 206
- phương pháp Delphi* (Delphi method) 326

- phương pháp đơn hình cải biên* (revised simplex method) 45
- phương pháp đơn hình cải biên gốc* (primal revised simplex method) 45
- phương pháp đơn hình tiêu chuẩn* (standard simplex method) 46
- phương pháp đơn hình vận tải* (transportation simplex method) 196
- phương pháp đường găng* (CPM, tức critical path method) 145
- phương pháp mặt phẳng cắt* (cutting plane method) 274
- phương pháp nhánh và cắt* (branch-and-cut method) 261
- phương pháp nhánh và cận* (branch-and-bound method) 261
- phương pháp nhân tử* (multiplier) 196
- phương pháp nhóm chuyên gia* (expert group method) 325
- phương pháp sơ đồ mạng lưới* 145
- phương pháp thế vị* (potential) 196
- phương pháp xấp xỉ liên tiếp* (method of successive approximation) 138, 378
- phương pháp xấp xỉ Vogel* (Vogel's approximation method, viết tắt VSM) 206
- phương sai* (variance) 293
- phương tiện* (facility) 385
- phương trình Bellman* (Bellman's equation) 137
- phương trình cân bằng* (balance equation) 390
- phương trình Chapman-Kolmogorov* 349
- phương trình duy trì* (steady-state equations) 357
- quá trình dừng* (stationary process) 349
- quá trình ngẫu nhiên liên tục* (continuous stochastic process) 345
- quá trình ngẫu nhiên rời rạc* (discrete stochastic process) 345
- quá trình Poisson* (Poisson process) 355
- quá trình quyết định Markov* (Markov decision process) 362
- quá trình quyết định Markov hữu hạn bước* (finite-period Markov decision process) 378
- quá trình sinh và tử* (birth and death process) 388
- quá trình vào Poisson* (Poisson input process) 387
- quan hệ hàm chính xác* (exact functional relationship) 336
- quét nút i* (scanning) 131
- quy hoạch 0-1 đa thức* (zero-one polynomial programming) 284
- quy hoạch động* (dynamic programming) 137
- quy hoạch nguyên* (integer programming, viết tắt IP) 251
- quy hoạch nguyên bộ phận* (mixed programming, viết tắt MIP) 251
- quy hoạch nguyên hoàn toàn* (pure programming) 251
- quy hoạch nguyên nhị phân* (binary programming, viết tắt BIP) 251
- quy hoạch nhị phân đa thức* (binary polynomial programming) 284
- quy tắc Bayes* (Bayes' decision rule) 293
- quy tắc Bland* (Bland's rule) 34
- quy tắc đặt mức đòi hỏi* (aspiration - level criterion) 294
- quy tắc góc tây bắc* (northwest - corner rule) 203
- quy tắc hợp lý nhất* (maximum likelihood criterion) 292
- quy tắc Hurwicz* (Hurwicz criterion) 307
- quy tắc khe khát* (savage criterion) 306
- quy tắc kì vọng* (expected value criterion) 293
- quy tắc kì vọng - phương sai* (expected value - variance criterion) 293
- quy tắc Laplace* (Laplace criterion) 304
- quy tắc minimax* (minimax criterion) 305

454 HƯỚNG DẪN TRA CỨU

- p>
quy tắc minimax hối tiếc (minimax regret criterion) 305
quy tắc xoay (pivot rule) 22
quyết định (decision) 363
ràng buộc (constraint) 8
sách lược (policy) 363
sách lược ngẫu nhiên hóa (randomized policy) 367
sách lược nhập hàng (s, S) ((s,S) ordering policy) 354
sai số dự báo (forecast error) 334
siêu phẳng (hyperplane) 12
sơ đồ mạng lưới (arrow network diagram) 146
suy biến (degenerate) 36
sự tăng dòng (flow augmentation) 129
sự xoay vòng (cycling) 33
tải năng (capacity) 95
tập lồi (convex set) 13
tập lồi đa diện (polyhedron) 12
tính chất không nhớ (lack-of-memory property) 362
tính chất Markov (Markovian property) 348
tính chất tương đương (equivalence property) 413
tính lồi (convexity) 13
tham số dạng (shape parameter) 408
thặng dư (residual) 334
thế vị 196
thỏa hiệp thời gian - chi phí (time-cost trade-off) 164
thời điểm hoàn thành muộn (latest completion) của hoạt động 154
thời điểm hoàn thành sớm (earliest completion) của hoạt động 153
thời điểm khởi công muộn (latest start) của hoạt động 153
thời điểm khởi công sớm (earliest start) của hoạt động 153
thời điểm muộn (latest time) của biến cố 150
thời điểm sớm (earliest time) của biến cố 149
thời gian bi quan (pessimistic time) 161
thời gian dự trữ (slack hoặc float) 150
thời gian dự trữ độc lập (free slack, free float) 151
thời gian dự trữ chung (total slack, total float) 150
thời gian dự trữ của đường đi (total float of a path) 154
thời gian giữa hai khách (inter arrival time) 385
thời gian hợp lý nhất (most likely time) 161
thời gian lạc quan (optimistic time) 160
thời gian phục vụ (service time) 385
thời gian phục vụ kiểu mũ (exponential service time) 394
thuật toán cải thiện sách lược (policy-improving algorithm) 372, 377
thuật toán chu trình cực phí âm (negative cost cycle algorithm) 143
thuật toán cộng (additive algorithm) 270
thuật toán Dijkstra (Dijkstra's algorithm) 138
thuật toán đánh dấu (labeling algorithm) 131
thuật toán đặt nhãn (label setting algorithm) 138
thuật toán đối ngẫu gốc (dual-primal algorithm) 88
thuật toán đơn hình đối ngẫu (dual simplex algorithm) 85
thuật toán đơn hình hai pha đối ngẫu - gốc (dual-primal two-phase simplex algorithm) 88
thuật toán đơn hình gốc (primal simplex algorithm) 87
thuật toán đơn hình mạng (network simplex algorithm) 101
thuật toán Ford-Fulkerson (Ford - Fulker-

- son's algorithm) 130
- thuật toán Gomory 276
- thuật toán gốc - đối ngẫu (primal-dual algorithm) 88
- thuật toán hai pha (two-phase algorithm) 25
- thuật toán phân đoán (heuristic algorithm) 254
- thuật toán phân số (fractional (pure integer) algorithm) 276
- thuật toán sửa nhãn (label correcting algorithm) 138
- tiền phạt (penalty) 206
- tổ hợp lồi (convex combination) 13
- tốc độ đến (arrival rate) 356
- trạng thái đến được (accessible state) 352
- trạng thái ergodic (ergodic state) 354
- trạng thái hấp thụ (absorbing state) 352
- trạng thái không tuần hoàn (aperiodic state) 353
- trạng thái lặp (recurrent state) 352
- trạng thái lặp dương (positive recurrent state) 354
- trạng thái lặp không (null recurrent state) 354
- trạng thái quá độ (transient state) 352
- trạng thái thông nhau (communicating states) 352
- trạng thái tuần hoàn (periodic state) 353
- trò chơi đúng (fair game) 309
- trò chơi hai người (two-person game) 308
- trò chơi hợp tác (cooperative game) 319
- trò chơi n người (n -person game) 319
- trò chơi tổng hằng số (constant-sum game) 318
- trò chơi tổng khác không (nonzero-sum game) 319
- trò chơi tổng zero (zero-sum game) 308
- trò chơi vi phân (differential game) 319
- trò chơi vô hạn (infinite game) 319
- từ vựng (dictionary) 19
- từ vựng suy biến (degenerate dictionary) 31
- xác suất có điều kiện (conditional probability) 296
- xác suất duy trì (steady-state probability) 348
- xác suất hậu nghiệm (posterior probability) 292
- xác suất quá độ (transient probability) 348
- xác suất tiên nghiệm (prior probability) 292
- xác suất truyền dừng (stationary transition probability) 349
- xác suất truyền n bước (n -step transition probability) 349
- xích Markov (Markov chain) 348
- xích Markov ergodic 354
- xích Markov không thu gọn được (irreducible) 352

TÀI LIỆU THAM KHẢO

- [Abounadi - Bertsekas - Borkar 2001] (J. Abounadi ...), Learning algorithms for Markov decision processes with average costs, *SIAM J. Control Optimization*, Vol. 40, 681-698.
- [Atamtürk - Johnson - Linderoth - Savelsbergh 2000] (A. Atamtürk ...), A relational modeling system for linear and integer programming, *Operations Research*, Vol. 48, 846-857.
- [Avis - Chvátal 1978] (A. Avis, V. Chvátal), notes on Bland's pivoting rule, *Mathematical Programming Study*, Vol. 8, 24-34.
- [Baumol 1977] (W. J. Baumol), *Economic Theory and Operations Analysis*, Prentice-Hall International, London.
- [Bazaraa - Shetty 1990] (M. S. Bazaraa, C. M. Shetty), *Nonlinear Programming - Theory and Algorithm*, John Wiley and Sons, New York.
- [Beasley 1996] (J. E. Beasley, editor), *Advances in Linear and Integer Programming*, Clarendon Press, Oxford.
- [Bertsimas - Tsitsiklis 1997] (D. Bertsimas, J. N. Tsitsiklis), *Introduction to Linear Optimization*, Athena Scientific, Belmont, Massachusetts.
- [Bixby - Boyd - Rios Mercado 1998] (R. E. Bixby, E. A. Boyd, R. Z. Rios - Mercado eds) *Integer Programming and combinatorial Optimization*, Springer, New York - Berlin.
- [Bland 1977] (R. G. Bland), New finite pivoting rules for the simplex method, *Mathematics of Operations Research* Vol. 2, 103-107.
- [Borkar 2000] (V. S. Borkar), Average cost dynamic programming equations for controlled Markov chains with partial observations, *SIAM J. Control Optimization*, Vol. 39, 673-681.
- [Brudno 1962] (A. L. Brudno), Giải bài toán Vận tải bằng điểm điện, *Áp dụng máy tính trong kinh tế*. NXB Viện hàn lâm Khoa học Liên Xô, 17-38 (tiếng Nga).

- [Burke 1956] (P. J. Burke), The output of a queueing system, *Operations Research*, Vol. 4, 699-704.
- [Chvátal 1983] (V. Chvátal), *Linear Programming*, Freeman and Company, New York.
- [Cook - Cunningham - Pulleyblank - Schrijver 1998] (W. J. Cook, W. H. Cunningham, W. R. Pulleyblank, A. Schrijver) *Combinatorial Optimization*, John Wiley and Sons, New York - Berlin.
- [Dantzig - Orchard Hays 1954] (G. B. Dantzig, W. Orchard - Hays), The product form for the inverse in the simplex method, *Mathematical Tables and Other Aids to Computation* VIII, April, 64-67.
- [[Dantzig - Thapa 1997]](G. B. Dantzig, M. N. Thapa), *Linear Programming*, 1: *Introduction*, Springer, New York.
- [Dantzig 1949] (G. B. Dantzig), Programming in a linear structure, Report of the September 9, 1948 meeting in Madison, *Econometrica* Vol. 17, 73-74.
- [Dantzig 1951] (G. B. Dantzig), Application of the simplex method to a transportation problem, in *Activity Analysis of Production and Allocation*, ed. T. C. Koopmans, John Wiley and Sons, New York, 359-373.
- [Dantzig 1963] (G. B. Dantzig), *Linear Programming and Extensions*, Princeton University Press, Princeton, New Jersey.
- [Dijkstra 1959] (E. Dijkstra), A note on two problems in connection with graphs, *Numerische Mathematik*, Vol. 1, 269-271.
- [Edmonds 1965] (J. Edmonds), Paths, trees and flowers, *Canadian Journal of Mathematics*, Vol. 17, 449-467.
- [Elias - Feinstein - Shannon 1956] (P. Elias, A. Feinstein, C. E. Shannon), Note on maximum flow through a network, *IRE Transaction on Information Theory*, Vol. 2, 117-119.
- [Ford-Fulkerson 1956a] (L. R. Ford, D. R. Fulkerson), Maximal flow through a network, *Canadian Journal of Mathematics*, Vol. 8, 399-404.
- [Ford-Fulkerson 1956b] (L. R. Ford, D. R. Fulkerson), Solving the transportation problem, *Management Science*, Vol. 3, 24-32.

- [Ford-Fulkerson 1957] (L. R. Ford, D. R. Fulkerson), A simple algorithm for finding maximal network flows and an application to the Hitchcock problem, *Canadian Journal of Mathematics*, Vol. 9, 210-218.
- [Ford-Fulkerson 1958] (L. R. Ford, D. R. Fulkerson), Constructing maximal dynamic flows from static flows, *Operations Research*, Vol. 6, 419-433.
- [Fulkerson - Dantzig 1955] (D. R. Fulkerson, G. B. Dantzig), Computation of maximum flow in networks, *Naval Research Logistics Quarterly*, Vol. 2, 277-283.
- [Gass - Saaty 1955] (S. I. Gass, T. Saaty), The computational algorithm for the parametric objective function, *Naval Research Logistics Quarterly* Vol. 2, 39-45.
- [George 2001] (J.M. George), Dynamic control of a queue with adjustable service rate, *Operations Research*, Vol. 49, 720-731.
- [Glazebrook - Niño Mora 2001] (K. D. Glazebrook, J. Niño-Mora), Parallel scheduling of multiclass M/M/m queues : Approximate and Heavy - Traffic Optimization of achievable performance, *Operations Research*, Vol. 49, 609-623.
- [Gleyzal 1958] (A. Gleyzal), Thuật toán giải bài toán Vận tải, *Matematika*, Vol. 2, Số 1 (tiếng Nga).
- [Goberna - López 1998] (M. A. Goberna, M. A. López), *Linear Semi - Infinite Optimization*, John Wiley and Sons, New York.
- [Golstein-Yudin 1969] (E.G. Golstein, D. B. Yudin), Các bài toán quy hoạch tuyến tính kiểu vận tải, *Nauka*, Moskva, (tiếng Nga).
- [Gomory 1958] (R. E. Gomory), Outline of an algorithm for integer solutions to linear programs, *Bulletin of the AMS*, Vol. 64, 275-278.
- [Gross - Harris 1985] (D. Gross, C. M. Harris), *Fundamentals of Queueing Theory*, Wiley, New York.
- [Gupta - Cozzolino 1974] (S. K. Gupta, J. M. Cozzolino), *Fundamental of Operations Research for Management*, Holden-Day, San Francisco.
- [Helmes - Röhl - Stockbridge 2001] (K. Helmes ...), Computing moments of the exit time distributon for Markov processes by linear programming, *Opera-*

- tions Research, Vol. 49, 516-530.
- [Hillier - Yu - Avis - Fossett - Lo - Reiman 1981] (F. S. Hillier, O. S. Yu ...) *Queueing Tables and Graphs*, Elsevier North - Holland, New York.
- [Hillier-Lieberman 2001] (F.S.Hillier, G.J. Lieberman), *Introduction to Operations Research*, McGraw-Hill, New York.
- [Hoff 1983] (J. C. Hoff), *A Practical Guide to Box-Jenkins Forecasting*, Lifetime Learning, Belmont.
- [Hoffman - Padberg 1991] (K. L. Hoffman, M. Padberg), Improving LP - representations of zero-one linear programs for branch-and-cut, *ORSA J. Computing*, Vol. 3, 121-134.
- [Hoffman 1953] (A. G. Hoffman), Cycling in the simplex algorithm, *National Bureau of Standards*, Report 2974.
- [Hordijk - Tijms 1976] (A. Hordijk, H. C. Tijms), A simple proof of the equivalence of the limiting distribution of the continuous - time and embedded process of the queue size in the M/G/1 queue, *Statistica Neerlandica*, Vol. 36, 97-100.
- [Jackson 1963] (J. R. Jackson), Jobshop - like queueing systems, *Management Science*, Vol 10, 131-142.
- [Jeroslow 1973] (R. G. Jeroslow), The simplex algorithm with the pivot rule of maximizing criterion improvement, *Discrete Mathematics*, Vol. 4, 367-377.
- [Karloff 1991] (H. Karloff), *Linear Programming*, Birkhauser, Boston - Berlin.
- [Khánh - Nương 2000] (P. Q. Khánh, T. H. Nương), *Quy Hoạch Tuyến Tính*, Giáo dục, Hà Nội.
- [Khánh 2000], *Toán Chuyên Đề*, NXB Đại học Quốc gia TP. Hồ Chí Minh.
- [Klee - Minty 1972] (V. Klee, G. J. Minty), How good is the simplex algorithm, in *Inequalities III*, ed. O. Shisha, Academic Press, New York, 159-175.
- [Konda - Borkar 1999] (V.R.Konda), Actor - critic - type learning algorithms for Markov decision processes, *SIAM J. Control Optimization*, Vol. 38, 94-123.
- [Lurie 1961] (A. L. Lurie), Thuật toán giải bài toán vận tải bằng xấp xỉ các phương án tối ưu có điều kiện, *Toán học tính toán*, số 7, 151-160 (tiếng Nga).

- [Taha 1997] (H. Taha), *Operations Research. An Introduction*, MacMillan, New York.
- [Vanderbei 1996] (R. J. Vanderbei), *Linear Programming. Foundations and Extensions*, Kluwer, Boston - London.
- [Williams 1962] (A. C. Williams), A treatment of transportation problems by decomposition. *London Soc. Int. Appl. Math.*, Vol. 10, 35-48.
- [Yudin - Nemirovskii 1977] (D. B. Yudin, A. Nemirovskii), Informational complexity and efficient methods for the solution of convex extremal problems, *Maketon* Vol. 13, 25-45.

MỤC LỤC

LỜI NÓI ĐẦU	3
MỞ ĐẦU	5
CHƯƠNG 1 QUY HOẠCH TUYẾN TÍNH. PHƯƠNG PHÁP ĐƠN HÌNH	7
1.1 Quy hoạch tuyến tính và cách giải hình học	7
1.2 Phương pháp đơn hình	16
1.3 Nghiên cứu hình học quy hoạch tuyến tính	35
1.4 Phương pháp đơn hình cải biên	45
Bài tập Chương 1	60
CHƯƠNG 2 QUY HOẠCH TUYẾN TÍNH ĐỐI NGẪU.	69
2.1 Bài toán mở đầu. định nghĩa	69
2.2 Các định lý về đối ngẫu	75
2.3 Thuật toán đơn hình đối ngẫu	83
2.4 Pha một dựa trên tính đối ngẫu	87
Bài tập Chương 2	88
CHƯƠNG 3 BÀI TOÁN DÒNG TRÊN MẠNG	91
3.1 Đồ thị và cây	92
3.2 Phát biểu bài toán dòng trên mạng	95
3.3 Thuật toán đơn hình mạng	100
3.4 Đối ngẫu trong bài toán dòng trên mạng	111
3.5 Thuật toán chu trình cực phí âm	115
3.6 Bài toán dòng cực đại	125
3.7 Bài toán đường ngắn nhất	136
Bài tập Chương 3	141
CHƯƠNG 4 ĐIỀU HÀNH DỰ ÁN BẰNG PHƯƠNG PHÁP PERT-CPM (PHƯƠNG PHÁP SƠ ĐỒ MẠNG LƯỚI)	145
4.1 Lập sơ đồ mạng lưới	147
4.2 Phân tích các chỉ tiêu thời gian, xác định đường găng	149
4.3 Điều khiển nhân lực	157

4.4. Hoàn thành sớm dự án	159
4.5 Dự án có tính ngẫu nhiên	160
4.6 Dự án có thỏa hiệp thời gian - cước phí	166
4.7 Kiểm tra hiệu chỉnh dự án	169
Bài tập Chương 4	170
CHƯƠNG 5 BÀI TOÁN VẬN TẢI	175
5.1 Các khái niệm cơ bản	175
5.2 Các tính chất của bài toán vận tải	187
5.3 Giải bài toán vận tải, thuật toán thế vị	194
5.4 Tìm phương án cơ sở để xuất phát	203
5.5 Các bài toán vận tải mở rộng	209
5.6 Phương pháp HUNGARY	223
Bài tập Chương 5	247
CHƯƠNG 6 QUY HOẠCH NGUYÊN	251
6.1 Phác họa quy hoạch nguyên	251
6.2 Quy hoạch nguyên điển hình trong thực tế	255
6.3 Phương pháp nhánh-và-cận cho quy hoạch nguyên	261
6.4 Phương pháp nhánh - và - cận cho quy hoạch nhị phân	267
6.5 Thuật toán cộng giải quy hoạch nhị phân	270
6.6 Phương pháp mặt phẳng cắt	274
6.7 Quy hoạch nhị phân đa thức	284
Bài tập Chương 6	286
CHƯƠNG 7 LÝ THUYẾT CHỌN QUYẾT ĐỊNH VÀ TRÒ CHƠI	291
7.1 Chọn quyết định có mạo hiểm	292
7.2 Cây quyết định	298
7.3 Lý thuyết về công dụng	300
7.4 Chọn quyết định bấp bênh	304
7.5 Lý thuyết trò chơi	308
7.6 Một số nội dung khác của lý thuyết trò chơi	318
Bài tập Chương 7	319

CHƯƠNG 8 CÁC PHƯƠNG PHÁP DỰ BÁO	325
8.1 Kỹ thuật phán đoán	325
8.2 Dự báo bằng chuỗi thời gian	326
8.3 Dự báo bằng hồi quy tuyến tính	336
Bài tập Chương 8	343
CHƯƠNG 9 XÍCH MARKOV VÀ QUÁ TRÌNH QUYẾT ĐỊNH MARKOV	345
9.1 Quá trình ngẫu nhiên và xích MARKOV	345
9.2 Mô hình quá trình quyết định MARKOV	362
9.3 Tìm sách lược tối ưu bằng quy hoạch tuyến tính	367
9.4 Thuật toán cải thiện sách lược	371
9.5 Tiêu chuẩn với cước phí giảm	375
9.6 Phương pháp xấp xỉ liên tiếp	378
Bài tập Chương 9	380
CHƯƠNG 10 HỆ THỐNG DỊCH VỤ (LÝ THUYẾT XẾP HÀNG)	384
10.1 Quá trình ngẫu nhiên sinh và tử	385
10.2 Các mô hình xếp hàng dựa trên quá trình sinh và tử	394
10.3 Các mô hình không dùng quá trình sinh và tử	407
10.4 Mô hình xếp hàng ưu tiên	408
10.5 Mạng dịch vụ	412
10.6 Biểu diễn xếp hàng qua xích MARKOV rời rạc	414
10.7 Chọn quyết định về thiết kế hệ dịch vụ	418
10.8 Các mô hình quyết định	423
Bài tập Chương 10	427
TRẢ LỜI BÀI TẬP	431
PHỤ LỤC	446
HƯỚNG DẪN TRA CỨU	448
TÀI LIỆU THAM KHẢO	456
MỤC LỤC	462

VẬN TRÙ HỌC

Mã số : 7K538t6 - CPH

In 1.000 bản, khổ 14,3 x 20,3 cm, tại Công ty CP In Sách Giáo Khoa Tp.HCM,
240 Trần Bình Trọng, Quận 5, Tp. HCM. Số XB : 242-2006/CXB/4-488/GD.
Số in : 137GC. In xong và nộp lưu chiểu tháng 04 năm 2006.

SÁCH ĐẠI HỌC MÔN TOÁN NHÀ XUẤT BẢN GIÁO DỤC

BIẾN ĐỔI TÍCH PHẦN

ĐẠI SỐ TUYẾN TÍNH (chương trình A3)

GIẢI TÍCH HÀM

HÀM MỘT BIẾN PHỨC

BÀI TẬP HÀM BIẾN PHỨC

HÌNH HỌC CAO CẤP

BÀI TẬP HÌNH HỌC CAO CẤP

LÍ THUYẾT TÍCH PHẦN

LÍ THUYẾT XÁC SUẤT VÀ THỐNG KÊ

BÀI TẬP XÁC SUẤT VÀ THỐNG KÊ

NHẬP MÔN GIẢI TÍCH

PHÉP TÍNH VI TÍCH PHẦN T1, T2

QUY HOẠCH TUYẾN TÍNH

Đặng Đình Áng

Nguyễn Cao Thắng

Đậu Thế Cấp

Đậu Thế Cấp

Đậu Thế Cấp

Nguyễn Mộng Hy

Nguyễn Mộng Hy

Đặng Đình Áng

Đinh Văn Gắng

Đinh Văn Gắng

Đặng Đình Áng

Phan Quốc Khánh

Phan Quốc Khánh-

Trần Huệ Nương

TOÁN CAO CẤP T1: *Phép tính vi tích phân hàm một biến
và lý thuyết chuỗi*

TOÁN CAO CẤP T2 : *Đại số tuyến tính*

BÀI TẬP TOÁN CAO CẤP T1, T2

Lê Thị Thiên Hương-

Nguyễn Viết Đông-

Nguyễn Anh Tuấn-

Lê Anh Vũ

NHỮNG BẤT NGỜ LÍ THỦ VỚI CƠ HỌC LÍ THUYẾT

TOÁN OLYMPIC CHO SINH VIÊN T1

TOÁN OLYMPIC CHO SINH VIÊN T2

Trịnh Phôi

Trần Lưu Cường

Trần Lưu Cường

**LIÊN HỆ ĐẶT MUA TẠI CÁC CÔNG TY SÁCH - TBTH ĐỊA PHƯƠNG
HOẶC CÁC CỬA HÀNG GIỚI THIỆU SÁCH CỦA NHÀ XUẤT BẢN GIÁO DỤC :**

57 Giảng Võ, Hà Nội

81 Trần Hưng Đạo, Hà Nội

15 Nguyễn Chí Thanh, Tp. Đà Nẵng

231 Nguyễn Văn Cừ, Q.5, Tp. Hồ Chí Minh

240 Trần Bình Trọng, Q.5 Tp. Hồ Chí Minh

Website : www.nxbgd.com.vn

ĐT : (04) 8562499

ĐT : (04) 8222466

ĐT : (0511) 827373

ĐT : (08) 8307141

ĐT : (08) 8323557



Giá : 35.500đ