

www.mientayvn.com

Dịch tiếng anh chuyên ngành khoa học tự nhiên và kỹ thuật.

Dịch các bài giảng trong chương trình học liệu mở của học viện MIT, Yale.

Tìm và dịch tài liệu phục vụ cho sinh viên làm seminar, luận văn.

Tại sao mọi thứ đều miễn phí và chuyên nghiệp ???

Trao i tr c tuy n t i:

www.mientayvn.com/chat_box_toan.html

Chương I

MỘT SỐ MÔ HÌNH VÀ PHƯƠNG PHÁP TỐI ƯU

1. Mô hình quy hoạch tuyến tính

1.1. Các bước cần thiết khi áp dụng phương pháp mô hình hoá

- Trước hết phải khảo sát, phát hiện vấn đề cần giải quyết.
- Phát biểu các điều kiện ràng buộc, mục tiêu của bài toán dưới dạng định tính. Sau đó lựa chọn các biến quyết định / các ẩn số và xây dựng mô hình định lượng (còn gọi là mô hình toán học).
- Thu thập số liệu, xác định phương pháp giải quyết.
- Định ra quy trình giải / thuật giải. Có thể giải mô hình bằng cách tính toán thông thường. Đối với các mô hình lớn, gồm nhiều biến và nhiều điều kiện ràng buộc cần lập trình và giải mô hình trên máy tính.
- Đánh giá kết quả. Trong trường hợp phát hiện thấy có kết quả bất thường hoặc kết quả không phù hợp với thực tế, cần kiểm tra và chỉnh sửa lại quy trình giải hoặc mô hình.

- Triển khai các phương án tìm được trên thực tế.

Các thuật ngữ sau thường gặp khi áp dụng phương pháp mô hình hoá:

- Ứng dụng toán / Toán ứng dụng (*Mathematical Applications* hay *Applied Mathematics*).
- Vận trù học (*Operations Research* viết tắt là *OR*).
- Khoa học quản lí (*Management Science* viết tắt là *MS*)

1.2. Mô hình quy hoạch tuyến tính

Phát biểu mô hình

Với mục đích tìm hiểu bước đầu, xét mô hình toán học sau đây, còn gọi là mô hình quy hoạch tuyến tính hay bài toán quy hoạch tuyến tính (BTQH TT), mà trong đó chúng ta muốn tối ưu hoá (cực đại hoá hay cực tiểu hoá) hàm mục tiêu:

$$z = c_1x_1 + c_2x_2 + \dots + c_nx_n \rightarrow \text{Max (Min)}$$

với các điều kiện ràng buộc:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \leq b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \leq b_2 \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \leq b_m \\ x_1, x_2, \dots, x_n \geq 0 \text{ (điều kiện không âm)} \end{cases}$$

Ví dụ: $z = 8x_1 + 6x_2 \rightarrow \text{Max}$

với các ràng buộc:

$$\begin{cases} 4x_1 + 2x_2 \leq 60 \\ 2x_1 + 4x_2 \leq 48 \\ x_1, x_2 \geq 0 \end{cases}$$

Cần tìm các giá trị của các biến quyết định x_1, x_2 để các ràng buộc được thoả mãn và hàm mục tiêu đạt giá trị lớn nhất.

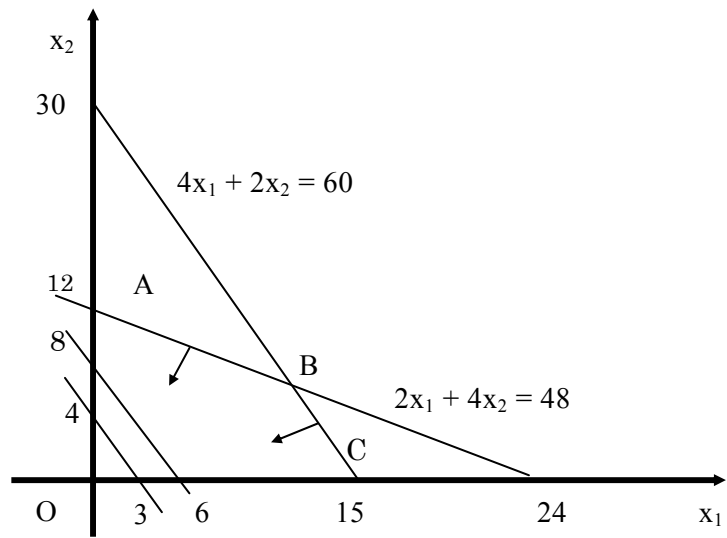
Bài toán này có ý nghĩa kinh tế như sau: Giả sử một xí nghiệp sản xuất hai loại sản phẩm I và II. Để sản xuất ra một đơn vị sản phẩm I cần có 4 đơn vị nguyên liệu loại A và 2 đơn vị nguyên liệu loại B, các chỉ tiêu đó cho một đơn vị sản phẩm loại II là 2 và 4. Lượng nguyên liệu dự trữ loại A và B hiện có là 60 và 48 (đơn vị). Hãy xác định phương án sản xuất đạt lợi nhuận lớn nhất, biết lợi nhuận trên mỗi đơn vị sản phẩm bán ra là 8 và 6 (đơn vị tiền tệ) cho các sản phẩm loại I và II.

Phương pháp đồ thị

Phương pháp đồ thị có ý nghĩa minh hoạ và giúp hiểu bản chất vấn đề.

Bước 1: Vẽ miền ràng buộc / miền các phương án khả thi, là tập hợp các phương án khả thi (các phương án, nếu nói một cách ngắn gọn). Mỗi phương án được thể hiện qua bộ số (x_1, x_2) còn gọi là véc tơ nghiệm, thoả mãn tất cả các ràng buộc đã có (xem hình I.1).

– Trước hết chúng ta vẽ đồ thị $4x_1 + 2x_2 = 60$ bằng cách xác định hai điểm trên đồ thị: $(x_1 = 0, x_2 = 30)$ và $(x_2 = 0, x_1 = 15)$.



Hình I.1. Phương pháp đồ thị giải bài toán quy hoạch tuyến tính

Đồ thị trên là một đường thẳng chia mặt phẳng làm hai nửa mặt phẳng: một phần gồm các điểm (x_1, x_2) thoả mãn $4x_1 + 2x_2 \leq 60$; một phần thoả mãn $4x_1 + 2x_2 \geq 60$. Ta tìm được nửa mặt phẳng thoả mãn $4x_1 + 2x_2 \leq 60$.

– Tương tự, có thể vẽ đồ thị $2x_1 + 4x_2 = 48$ bằng cách xác định hai điểm thuộc đồ thị $(x_1 = 0, x_2 = 12)$ và $(x_2 = 0, x_1 = 24)$. Sau đó tìm nửa mặt phẳng thoả mãn $2x_1 + 4x_2 \leq 48$.

– Lúc này, giao của hai nửa mặt phẳng tìm được trên cho ta tập hợp các điểm (x_1, x_2) thoả mãn hai ràng buộc đầu tiên. Tuy nhiên, để thoả mãn điều kiện không âm của các biến, ta chỉ xét các điểm nằm trong góc phần tư thứ nhất. Vậy miền các phương án khả thi là miền giới hạn bởi tứ giác OABC (còn gọi là đơn hình vì là miền tạo nên bởi giao của các nửa mặt phẳng).

Bước 2: Trong miền (OABC) ta tìm điểm (x_1, x_2) sao cho

$$z = 8x_1 + 6x_2 \text{ đạt giá trị lớn nhất.}$$

Cách 1: Dùng đường đồng mức. Tùy theo giá trị của x_1, x_2 mà z có những mức giá trị khác nhau.

– Vẽ đường đồng mức: $8x_1 + 6x_2 = c$ ở mức $c = 24$, (ta có thể chọn giá trị c bất kì, nhưng chọn $c = 24$ là bội số chung của 6 và 8 để việc tìm tọa độ các điểm cắt hai trục tọa độ thuận lợi hơn). Dễ dàng tìm được hai điểm nằm trên đường đồng mức này là $(x_1 = 3, x_2 = 4)$ và $(x_2 = 0, x_1 = 3)$. Các điểm nằm trên đường đồng mức này đều cho giá trị hàm mục tiêu $z = 24$.

– Tương tự, có thể vẽ đường đồng mức thứ hai: $8x_1 + 6x_2 = 48$ đi qua hai điểm $(x_1 = 0, x_2 = 8)$ và $(x_2 = 0, x_1 = 6)$. Chúng ta nhận thấy, nếu tịnh tiến song song đường đồng mức lên trên theo hướng của véc tơ pháp tuyến $\vec{n}(8, 6)$ thì giá trị của hàm mục tiêu $z = 8x_1 + 6x_2$ tăng lên.

Vậy giá trị z lớn nhất đạt được khi đường đồng mức đi qua điểm B(12, 6) (tìm được $x_1 = 12, x_2 = 6$ bằng cách giải hệ phương trình $4x_1 + 2x_2 = 60$ và $2x_1 + 4x_2 = 48$).

Kết luận: Trong các phương án khả thi thì phương án tối ưu là $(x_1 = 12, x_2 = 6)$. Tại phương án này, giá trị hàm mục tiêu là lớn nhất $z_{\max} = 8 \times 12 + 6 \times 6 = 132$.

Nhận xét: Phương án tối ưu của bài toán trên (hay các BTQHHT khác, nếu có) luôn đạt được tại một trong các đỉnh của đơn hình hay còn gọi là các điểm cực biên của đơn hình (chính xác hơn, điểm cực biên là điểm thuộc đơn hình, mà không thể tìm được một đoạn thẳng nào cũng thuộc đơn hình nhận điểm đó là điểm trong). Nhận xét trên đây là một định lí toán học đã được chứng minh một cách tổng quát. Nói một cách hình ảnh, muốn đạt được phương án tối ưu cho các BTQHHT thì cần phải “mạo hiểm” đi xét các điểm cực biên của miền phương án.

Cách 2: Từ nhận xét trên, để tìm phương án tối ưu ta chỉ cần so sánh giá trị của hàm mục tiêu tại các điểm cực biên của miền phương án.

Tính giá trị z tại $O(0, 0)$: $z(0, 0) = 0$; tại $A(0, 12)$: $z(0, 12) = 72$; tại $C(15,0)$: $z(15, 0) = 120$; tại $B(12, 6)$: $z(12, 6) = 132 = \text{Max}\{z(O), z(A), z(B), z(C)\}$. Vậy $z_{\text{max}} = 132$.

Nhận xét: Muốn tìm phương án tối ưu của BTQHHT ta xuất phát từ một điểm cực biên nào đó, tìm cách cải thiện hàm mục tiêu bằng cách đi tới điểm cực biên kề nó. Tiếp tục như vậy cho tới khi tìm được phương án tối ưu. Trong trường hợp BTQHHT có phương án tối ưu thì quy trình giải này bao gồm hữu hạn bước (do số điểm cực biên là hữu hạn).

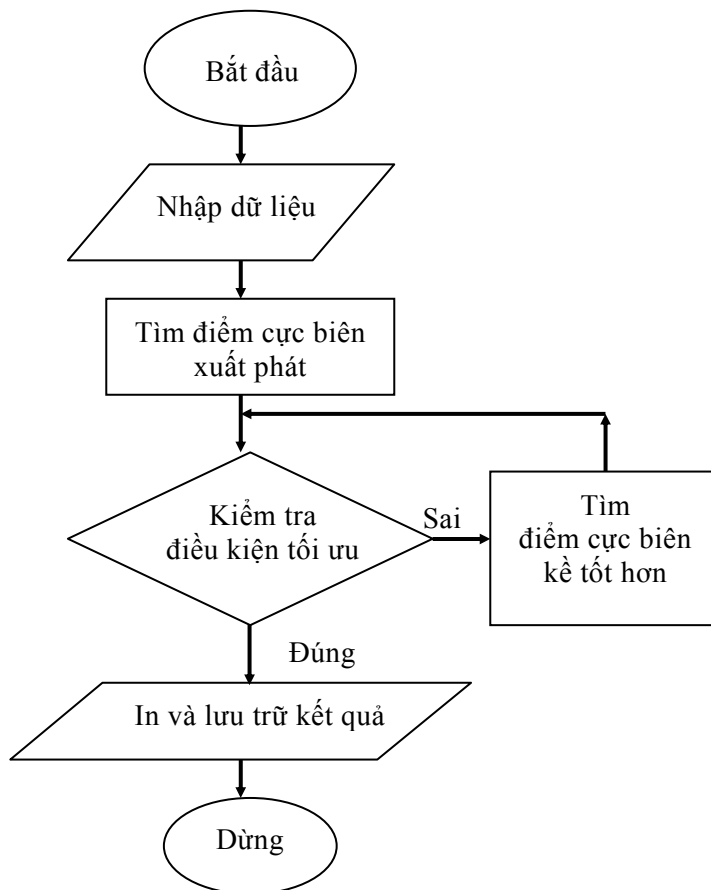
Đối với BTQHHT đang xét, quy trình giải được minh họa như sau:

$O(0, 0) \rightarrow A(0,12) \rightarrow B(12,6)$ dừng
 $z = 0 \rightarrow z = 72 \rightarrow z = 132$

hoặc:

$O(0, 0) \rightarrow C(15, 0) \rightarrow B(12, 6)$ dừng
 $z = 0 \rightarrow z = 120 \rightarrow z = 132$

Sơ đồ khởi



Hình 1.2. Sơ đồ khởi giải BTQHHT

Quy trình giải BTQHHT tổng quát có sơ đồ khôi giản lược như trình bày trên hình I.2. Trong sơ đồ trên, vì mục đích trình bày vấn đề đơn giản, chúng ta không đề cập tới các trường hợp khi BTQHHT có miền phương án là tập rỗng (lúc đó ta không tìm được phương án xuất phát) cũng như khi ta không tìm được điểm cực biên kẻ tốt hơn mặc dù điều kiện tối ưu chưa thoả mãn (lúc đó tập các giá trị hàm mục tiêu z không bị chặn).

1.3. Phương pháp đơn hình

Đây là phương pháp số giải BTQHHT theo sơ đồ trên. Để giải ví dụ đã cho, trước hết chúng ta cần đưa BTQHHT về dạng chính tắc bằng cách thêm vào các biến bù không âm x_3 và x_4 như sau:

$$z = 8x_1 + 6x_2 + 0x_3 + 0x_4 \rightarrow \text{Max}$$

với các ràng buộc:

$$\begin{cases} 4x_1 + 2x_2 + x_3 = 60 \\ 2x_1 + 4x_2 + x_4 = 48 \\ x_1, x_2, x_3, x_4 \geq 0 \end{cases}$$

Cách lập và biến đổi các bảng đơn hình

Để giải BTQHHT dạng chính tắc trên đây, cần lập một số bảng đơn hình như trình bày trong bảng I.1. Trước hết, cần điền số liệu của bài toán đã cho vào bảng đơn hình bước 1:

– Cột 1 là cột hệ số hàm mục tiêu ứng với các biến cơ sở đã chọn. Phương án xuất phát có thể chọn là $x_1 = x_2 = 0$ (đây chính là điểm gốc tọa độ $O(0, 0)$), do đó $x_3 = 60$, $x_4 = 48$). Như vậy tại bước này chúng ta chưa bước vào sản xuất, nên trong phương án chưa có đơn vị sản phẩm loại I hay II được sản xuất ra (chỉ “sản xuất” ra các lượng nguyên liệu dư thừa, ta cũng nói là các “sản phẩm” loại III và IV), và giá trị hàm mục tiêu z tạm thời bằng 0. Các biến bù có giá trị lớn hơn 0 có nghĩa là các nguyên liệu loại tương ứng chưa được sử dụng hết. Ta gọi các biến x_3 và x_4 là các biến cơ sở vì chúng có giá trị lớn hơn 0 còn x_1 và x_2 là các biến ngoài cơ sở vì chúng có giá trị bằng 0. Với bài toán có hai ràng buộc, tại mỗi bước chỉ có hai biến cơ sở.

– Cột 2 là cột các biến cơ sở. Trong cột 3 (cột phương án) cần ghi các giá trị của các biến cơ sở đã chọn.

– Các cột tiếp theo là các cột hệ số trong các điều kiện ràng buộc tương ứng với các biến x_1 , x_2 , x_3 và x_4 của bài toán đã cho.

Bảng I.1. Các bảng đơn hình giải BTQHTT

Hệ số hàm mục tiêu c_j	Biến cơ sở	Phương án	$c_1 = 8$	$c_2 = 6$	$c_3 = 0$	$c_4 = 0$
			x_1	x_2	x_3	x_4
0	x_3	60	4	2	1	0
0	x_4	48	2	4	0	1
Hàng z		$z_0 = 0$	$z_1 = 0$	$z_2 = 0$	$z_3 = 0$	$z_4 = 0$
Hàng $\Delta_j = c_j - z_j$			$\Delta_1 = 8$	$\Delta_2 = 6$	$\Delta_3 = 0$	$\Delta_4 = 0$
8	x_1	15	1	1/2	1/4	0
0	x_4	18	0	3	-1/2	1
Hàng z		$z_0 = 120$	$z_1 = 8$	$z_2 = 4$	$z_3 = 2$	$z_4 = 0$
Hàng $\Delta_j = c_j - z_j$			$\Delta_1 = 0$	$\Delta_2 = 2$	$\Delta_3 = -2$	$\Delta_4 = 0$
8	x_1	12	1	0	1/3	-1/6
6	x_2	6	0	1	-1/6	1/3
Hàng z		$z_0 = 132$	8	6	5/3	2/3
Hàng $\Delta_j = c_j - z_j$			0	0	-5/3	-2/3

Phân tích bảng đơn hình bước 1

– Hệ số ứng với biến x_1 trên hàng thứ nhất là $a_{11} = 4$ có nghĩa là tỉ lệ thay thế riêng giữa một đơn vị sản phẩm loại I và một đơn vị sản phẩm loại III là 4 (giải thích: xét phương trình / ràng buộc thứ nhất $4x_1 + 2x_2 + x_3 = 60$, x_1 tăng một đơn vị thì x_3 phải giảm bốn đơn vị nếu giữ nguyên x_2). Tương tự ta có thể giải thích được ý nghĩa của các hệ số a_{ij} khác cho trên hàng 1 và hàng 2 trong bảng đơn hình bước 1.

– Chúng ta xét hàng z của bảng đơn hình. Để tính z_1 , cần áp dụng công thức $z_1 =$ (cột hệ số của hàm mục tiêu) \times (cột hệ số của biến x_1) $= 0 \times 4 + 0 \times 2 =$ (giá một đơn vị sản phẩm loại III) \times (tỉ lệ thay thế riêng loại I / loại III) $+ ($ giá một đơn vị sản phẩm loại IV) \times (tỉ lệ thay thế riêng loại I / loại IV) $=$ tổng chi phí phải bỏ ra khi đưa thêm một đơn vị sản phẩm loại I vào phương án sản xuất mới $= 0$. Các giá trị z_j , với $j = 1, 2, 3, 4$, được tính tương tự và chính là các chi phí khi đưa một thêm một đơn vị sản phẩm loại x_j vào phương án sản xuất mới. Còn z_0 là giá trị của hàm mục tiêu đạt được tại phương án đang xét: $z_0 =$ (cột hệ số của hàm mục tiêu) \times (cột phương án) $= 0 \times 60 + 0 \times 48 = 0$.

– Trên hàng Δ_j cần ghi các giá trị Δ_j , $j = 1, 2, 3, 4$, tính theo công thức $\Delta_j = c_j - z_j =$ lợi nhuận trên một đơn vị sản phẩm – chi phí trên một đơn vị sản phẩm. Vậy Δ_j là "lãi biên"/một đơn vị sản phẩm khi đưa thêm một đơn vị sản phẩm loại j vào phương án sản xuất mới. Nếu $\Delta_j > 0$ thì hàm mục tiêu còn tăng được khi ta đưa thêm các đơn vị sản phẩm loại j vào phương án sản xuất mới. Có thể chứng minh được Δ_j chính là đạo hàm riêng $\partial z / \partial x_j$ của hàm mục tiêu z theo biến x_j . Như vậy, x_1 tăng lên 1 thì z tăng lên 8 còn x_2 tăng lên 1 thì z tăng lên 6.

Do Δ_1 và Δ_2 đều dương nên vẫn còn khả năng cải thiện hàm mục tiêu khi chuyển sang (hay “xoay sang”) một phương án cực biên kề tốt hơn (quay lại nhận xét ở phần giải bài toán bằng phương pháp đồ thị: điểm cực biên kề của điểm $(0, 0)$ có thể là $A(0, 12)$ hay $C(15, 0)$).

Thủ tục xoay (pivotal procedure)

Bước 1: Chọn cột xoay là cột có $\Delta_j > 0$ tức là chọn biến x_j làm biến cơ sở mới do x_j tăng kéo theo hàm mục tiêu tăng. Ở đây ta chọn đưa x_1 vào (đánh dấu \surd ở cột Δ_1).

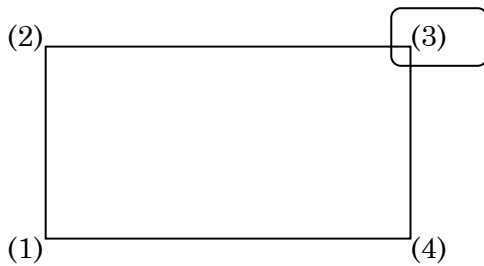
Bước 2: Chọn hàng xoay để xác định đưa biến nào ra khỏi số biến cơ sở (vì tại mỗi bước số biến cơ sở là không thay đổi). Để chọn hàng xoay, ta thực hiện quy tắc “tỉ số dương bé nhất” bằng cách lấy cột phương án $(60 \ 48)^T$ chia tương ứng cho cột xoay $(4 \ 2)^T$ để chọn tỉ số bé nhất. Một điều cần chú ý là ta chỉ xét các tỉ số có mẫu số dương.

Vì $\text{Min}\{60/4, 48/2\} = 60/4$ đạt được tại hàng đầu, nên ta đánh dấu \surd vào hàng xoay là hàng đầu (hàng tương ứng với biến x_3). Do đó cần đưa x_3 ra khỏi các biến cơ sở.

Bước 3: Chọn phần tử xoay nằm trên giao của hàng xoay và cột xoay.

Bước 4: Xoay sang bảng đơn hình mới, xác định các biến cơ sở mới để điền vào cột biến cơ sở, đồng thời thay các giá trị trong cột hệ số hàm mục tiêu. Sau đó, tính lại các phần tử của hàng xoay bằng cách lấy hàng xoay cũ chia cho phần tử xoay để có hàng mới tương ứng.

Bước 5: Các phần tử còn lại của bảng đơn hình mới được tính theo quy tắc “hình chữ nhật”: $(1)_{\text{mới}} = (1)_{\text{cũ}} - (2)_{\text{cũ}} \times (4)_{\text{cũ}} / (3)_{\text{cũ}}$, trong đó (3) là đỉnh tương ứng với phần tử xoay (xem hình I.3).



Chặng hạn: $(1)_{\text{cũ}} = 4, 2_{(\text{cũ})} = 2$
 $(3)_{\text{cũ}} = \text{phần tử xoay} = 4, (4)_{\text{cũ}} = 2$
 $\Rightarrow (1)_{\text{mới}} = 4 - 2 \times \frac{2}{4} = 3.$

Hình I.3. Quy tắc hình chữ nhật

Giải thích: Các bước xoay trên đây chỉ là phép biến đổi tương đương hệ phương trình

$$\begin{cases} 4x_1 + 2x_2 + x_3 = 60 \text{ (a)} \\ 2x_1 + 4x_2 + x_4 = 48 \text{ (b)} \end{cases}$$

để có hệ

$$\begin{cases} x_1 + (1/2)x_2 + (1/4)x_3 = 15 \text{ (a')} \\ \end{cases}$$

$$0x_1 + 3x_2 - (1/2)x_3 + x_4 = 18 \text{ (b')}$$

bằng cách lấy phương trình (a) chia cho 4 (phần tử xoay) để có (a'), rồi lấy (b) trừ bớt $2 \times (a)/4$ để có (b'). Đây chính là nội dung của bước 4 và bước 5. Còn bước 3 sẽ đảm bảo rằng giá trị của các biến cơ sở mới không âm ($x_1 = 15, x_4 = 18$).

Áp dụng thủ tục xoay cho các phần tử nằm trên hàng 1 và 2 của bảng đơn hình bước 1, sau đó tính các giá trị trên hàng z_j và Δ_j tương tự như khi lập bảng đơn hình bước 1, chúng ta sẽ nhận được bảng đơn hình bước 2.

Phân tích bảng đơn hình bước 2

Bảng bước 2 có thể được phân tích tương tự như bảng bước 1. Cần chú ý rằng lúc này ta đang ở vị trí của điểm C(15, 0) vì $x_1 = 15$ còn $x_2 = 0$; giá trị của hàm mục tiêu là $z_0 = 120$ đã được cải thiện hơn so với bước 1. Ta thấy $\Delta_2 = 2 > 0$ nên còn có thể cải thiện hàm mục tiêu bằng cách chọn biến x_2 làm biến cơ sở mới. Thực hiện các bước xoay sang phương án cực biên kế tốt hơn, chúng ta sẽ có bảng đơn hình bước 3.

Phân tích bảng đơn hình bước 3

Tại bảng đơn hình bước 3 ta thấy điều kiện tối ưu đã được thoả mãn ($\Delta_j \leq 0 \forall j=1, 2, 3, 4$) nên không còn khả năng cải thiện phương án. Phương án tối ưu đã đạt được tại $x_1 = 12, x_2 = 6, x_3 = 0, x_4 = 0$, tức là tại điểm cực biên B(12, 6) với giá trị $z_{\max} = 132$.

Một số chú ý

- Điều kiện tối ưu cho các BTQH TT dạng Max là $\Delta_j \leq 0 \forall j$.
- Đối với các BTQH TT cần cực tiểu hoá hàm mục tiêu thì điều kiện tối ưu (hay tiêu chuẩn dừng) là $\Delta_j \geq 0 \forall j$ (nếu tồn tại j mà $\Delta_j \leq 0$ thì cần tiếp tục cải thiện hàm mục tiêu bằng cách chọn cột j làm cột xoay...).
- Trong thực tiễn giải các BTQH TT dạng tổng quát có thể xảy ra trường hợp không tìm được phương án xuất phát (tức là không có phương án khả thi, xem thêm mục 1.2). Lúc này có thể kết luận mô hình đã thiết lập có các điều kiện ràng buộc quá chặt chẽ, cần xem xét nới lỏng các điều kiện này.
- Trong trường hợp ta tìm được cột xoay mà không tìm được hàng xoay thì kết luận hàm mục tiêu không bị chặn trên (đối với các BTQH TT dạng Max) hoặc không bị chặn dưới (đối với các BTQH TT dạng Min). Khi đó dừng quá trình giải và kết luận mô hình quy hoạch tuyến tính đã thiết lập không phù hợp với thực tế.

1.4. Giải mô hình quy hoạch tuyến tính bằng các phần mềm tính toán

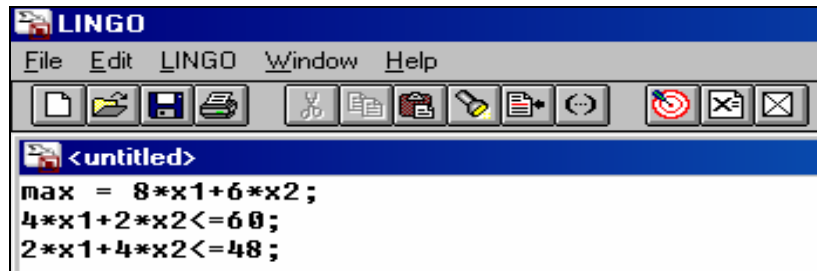
Hiện nay có nhiều phần mềm tính toán giải BTQH TT khá hiệu quả như Excel, Lingo. Những phần mềm này rất thân thiện với người dùng. Tuy nhiên cần nhấn mạnh rằng, việc phát biểu được mô hình bài toán và phân tích, đánh giá được kết quả mới chính là những khâu quan trọng nhất trong phương pháp mô hình hoá. Sau đây, chúng ta dùng phần mềm Lingo để giải ví dụ đã xét ở trên.

$$z = 8x_1 + 6x_2 \rightarrow \text{Max}$$

với các ràng buộc:

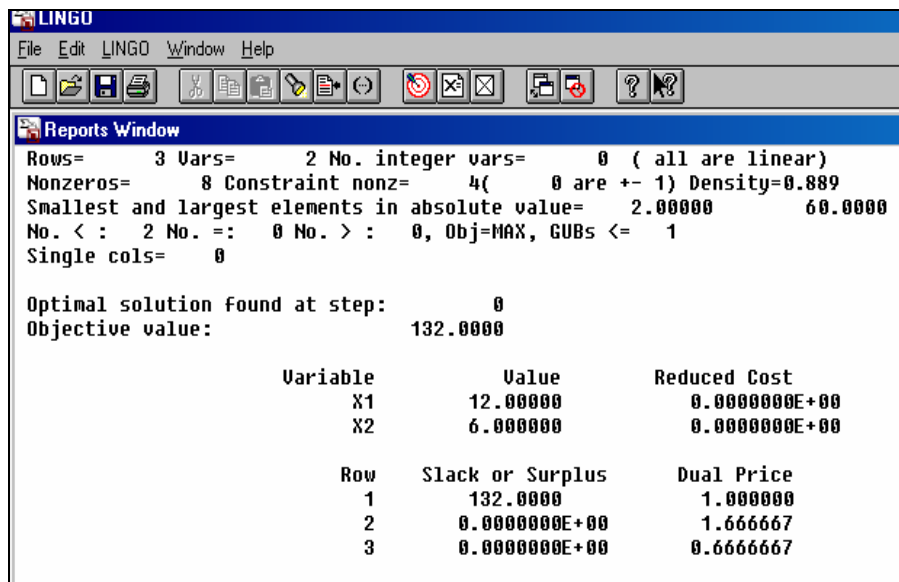
$$\begin{cases} 4x_1 + 2x_2 \leq 60 \\ 2x_1 + 4x_2 \leq 48 \\ x_1, x_2 \geq 0. \end{cases}$$

Để giải bài toán này, chúng ta cần cài đặt Lingo vào trong máy tính. Nhấn vào biểu tượng Lingo trên màn hình để vào cửa sổ Lingo. Sau đó thực hiện các lệnh Lingo: *Menu > New > <Untitled>* và gõ vào các dữ liệu của bài toán như hình I.4.



Hình I.4. Nhập dữ liệu của bài toán quy hoạch tuyến tính trong Lingo

Tiếp theo, cần nhấp chuột vào nút LINGO và giải bài toán để thu được kết quả chi tiết như trên hình I.5.



Hình I.5. Kết quả giải bài toán quy hoạch tuyến tính trong Lingo

Kết quả chi tiết cho ta biết giá trị cực đại của hàm mục tiêu là 132 với phương án tối ưu là: $x_1 = 12$, $x_2 = 6$. Các giá trị tối ưu của các biến đối ngẫu là $y_1 = 5/3$ và $y_2 = 2/3$ (còn gọi là các giá ước định hay giá bóng *Shadow Prices*).

1.5. Một số ứng dụng của phương pháp đơn hình

(Giải các bài toán quy hoạch sản xuất trong lĩnh vực cơ khí và điện lực)

Bài toán phân phối điện năng

Có ba hộ phụ tải cần được cung cấp điện năng từ hai nguồn điện nằm cách xa nhau. Giá thành truyền tải một đơn vị điện năng từ nguồn i đến hộ tiêu thụ j là c_{ij} . Khả năng cung cấp điện năng của mỗi nguồn bị giới hạn bởi trữ lượng hiện có của chúng là A_1 và A_2 . Nhu cầu tiêu dùng của các hộ tiêu thụ là B_1 , B_2 và B_3 . Gọi x_{ij} là lượng điện năng được đưa từ nguồn i tới hộ tiêu thụ j . Cần phải xác định các x_{ij} sao cho tổng chi phí là nhỏ nhất. Như vậy ta có BTQHHTT sau:

$$z = \sum_{i=1}^2 \sum_{j=1}^3 c_{ij} x_{ij} \rightarrow \text{Min}$$

với các điều kiện ràng buộc là:

$$\begin{cases} x_{11} + x_{12} + x_{13} \leq A_1, \\ x_{21} + x_{22} + x_{23} \leq A_2, \\ x_{11} + x_{21} = B_1, \\ x_{12} + x_{22} = B_2, \\ x_{13} + x_{23} = B_3, \\ x_{ij} \geq 0, \quad \forall i = 1, 2 \text{ và } \forall j = 1, 2, 3. \end{cases}$$

Bài toán trên đây (hoặc ở dạng tổng quát hơn) có thể giải được bằng phương pháp đơn hình đã biết hay phương pháp phân phối sẽ được nghiên cứu ở mục 1.3, chương II.

Bài toán phân tải cho máy

Một xí nghiệp có hai loại máy M_1 và M_2 . Các loại máy này có thể sản xuất được ba loại sản phẩm P_1 , P_2 và P_3 với các năng suất là a_{ij} , chẳng hạn máy M_1 sản xuất sản phẩm P_2 với năng suất a_{12} . Mỗi đơn vị sản phẩm mang lại lãi suất c_j với $j = 1, 2, 3$. Mỗi tháng xí nghiệp phải sản xuất sản phẩm loại j không ít hơn b_j đơn vị và không vượt quá d_j đơn vị, $j = 1, 2, 3$. Hãy lập kế hoạch phân tải cho các máy sao cho đạt tổng lợi nhuận lớn nhất.

Để thấy bài toán này dẫn tới BTQHHTT sau:

$$z = \sum_{j=1}^3 c_j \sum_{i=1}^2 a_{ij} x_{ij} \rightarrow \text{Max}$$

với các điều kiện ràng buộc:

$$\begin{cases} a_{11}x_{11} + a_{21}x_{21} \geq b_1, \\ a_{12}x_{12} + a_{22}x_{22} \geq b_2, \\ a_{13}x_{13} + a_{23}x_{23} \geq b_3, \\ a_{11}x_{11} + a_{21}x_{21} \leq d_1, \\ a_{12}x_{12} + a_{22}x_{22} \leq d_2, \\ a_{13}x_{13} + a_{23}x_{23} \leq d_3, \\ x_{11} + x_{12} + x_{13} \leq m_1, \\ x_{21} + x_{22} + x_{23} \leq m_2, \\ x_{ij} \geq 0, i = 1, 2 \text{ và } j = 1, 2, 3. \end{cases}$$

(trong đó m_1 và m_2 là tổng thời gian chạy máy M_1 và M_2).

Bài toán trên đây còn có thể phát biểu một cách tổng quát hơn và vẫn giải được bằng phương pháp đơn hình. Hơn nữa, trong lĩnh vực quy hoạch sản xuất hay quản lí kinh doanh, nói riêng trong ngành cơ khí và điện lực, BTQHTT được ứng dụng rất rộng rãi và mang lại hiệu quả cần thiết.

2. Bổ sung thêm về phương pháp đơn hình

2.1. Đưa BTQHTT về dạng chính tắc

Ví dụ 1: (Trường hợp các ràng buộc đều có dấu \leq)

$$z = 8x_1 + 6x_2 \rightarrow \text{Max}$$

với các ràng buộc:

$$\begin{cases} 4x_1 + 2x_2 \leq 60 \\ 2x_1 + 4x_2 \leq 48 \\ x_1, x_2 \geq 0 \end{cases}$$

Đưa BTQHTT về dạng chính tắc như đã biết bằng cách thêm hai biến bù (*slack variables*) x_3 và x_4 . Ta có BTQHTT dạng chính tắc là:

$$z = 8x_1 + 6x_2 + 0x_3 + 0x_4 \rightarrow \text{Max}$$

$$\begin{cases} 4x_1 + 2x_2 + x_3 = 60 \\ 2x_1 + 4x_2 + x_4 = 48 \\ x_1, x_2, x_3, x_4 \geq 0 \end{cases}$$

Lúc này, trong hệ hai điều kiện ràng buộc đã có đủ hai biến đứng độc lập trong từng phương trình với hệ số $+1$, nên đã có thể tìm được phương án cực biên xuất phát để bắt đầu quá trình giải bài toán. Một cách tổng quát, *BTQHTT dạng chính tắc là bài toán với các biến không âm, các ràng buộc với dấu "=", hệ số về phải của các ràng buộc*

không âm. Ngoài ra, mỗi phương trình bắt buộc phải có một biến đứng độc lập với hệ số +1.

Ví dụ 2: (Trường hợp có điều kiện ràng buộc với dấu \geq)

$$z = 8x_1 + 6x_2 \rightarrow \text{Max}$$

với các ràng buộc:

$$\begin{cases} 4x_1 + 2x_2 \leq 60 \\ 2x_1 + 4x_2 \geq 48 \\ x_1, x_2 \geq 0 \end{cases}$$

Ta thêm các biến bù x_3 (*slack variable*) mang dấu “+”, x_4 (*surplus variable*) mang dấu “-” để có hệ điều kiện ràng buộc sau:

$$\begin{cases} 4x_1 + 2x_2 + x_3 = 60 \\ 2x_1 + 4x_2 - x_4 = 48 \\ x_1, x_2, x_3, x_4 \geq 0 \end{cases}$$

Phải thêm biến giả x_5 (x_5 gọi là lượng vi phạm của phương trình thứ hai) để được hệ điều kiện ràng buộc

$$\begin{cases} 4x_1 + 2x_2 + x_3 = 60 \\ 2x_1 + 4x_2 - x_4 + x_5 = 48 \\ x_1, x_2, x_3, x_4, x_5 \geq 0 \end{cases}$$

Lúc này, đã có đủ hai biến đứng độc lập trong từng phương trình với hệ số +1, nên đã có thể tìm được phương án cực biên xuất phát để bắt đầu quá trình giải bài toán bằng phương pháp đơn hình với hàm mục tiêu là $z = 8x_1 + 6x_2 + 0x_3 + 0x_4 - Mx_5 \rightarrow \text{Max}$, trong đó $M \approx +\infty$ và biểu thức $-Mx_5$ gọi là lượng phạt (đánh thuế). Bài toán đã được đưa về dạng chính tắc. Lượng vi phạm x_5 càng lớn thì hàm mục tiêu càng giảm, giá trị của hàm mục tiêu chỉ có thể đạt Max khi $x_5 = 0$.

Ví dụ 3: (Trường hợp có biến không dương)

$$z = 8x_1 - 6x_2 \rightarrow \text{Max}$$

với các ràng buộc:

$$\begin{cases} 4x_1 + 2x_2 + x_3 \leq 60 \\ 2x_1 + 4x_2 - x_4 = 48 \\ x_1 \geq 0, x_2 \leq 0, x_3 \geq 0, x_4 \geq 0 \end{cases}$$

Lúc này muốn giải bài toán bằng phương pháp đơn hình ta phải đổi biến $x'_2 = -x_2$. Ta có BTQHTT với các biến đều không âm.

$$z = 8x_1 + 6x'_2 \rightarrow \text{Max}$$

với các ràng buộc:

$$\begin{cases} 4x_1 - 2x'_2 + x_3 \leq 60 \\ 2x_1 - 4x'_2 - x_4 = 48 \\ x_1, x'_2, x_3, x_4 \geq 0 \end{cases}$$

Ví dụ 4: (Trường hợp có biến với dấu tùy ý)

$$z = 8x_1 + 6x_2 \rightarrow \text{Max}$$

với các ràng buộc:

$$\begin{cases} 4x_1 + 2x_2 \leq 60 \\ 2x_1 + 4x_2 \leq 48 \\ x_1 \geq 0, x_2 \text{ dấu tùy ý} \end{cases}$$

Lúc này ta viết biến x_2 dưới dạng $x_2 = x'_2 - x''_2$ với

$$\begin{cases} x'_2 = \max[0, x_2] \\ x''_2 = \max[0, -x_2] \end{cases} \text{ thì đảm bảo } \begin{cases} x'_2 \geq 0 \\ x''_2 \geq 0 \end{cases}$$

Các ràng buộc sẽ là

$$\begin{cases} 4x_1 + 2x'_2 - 2x''_2 + x_3 = 60 \\ 2x_1 + 4x'_2 - 4x''_2 + x_4 = 48 \\ x_1, x'_2, x''_2, x_3, x_4 \geq 0 \end{cases}$$

Bài toán với hàm mục tiêu là: $z = 8x_1 + 6x'_2 - 6x''_2 + 0x_3 + 0x_4$ và các điều kiện ràng buộc trên là BTQH TT dạng chính tắc.

Kết luận: Bao giờ cũng đưa được BTQH TT bất kì (các biến có dấu tùy ý, các ràng buộc có thể $\leq, \geq, =$) về dạng chính tắc.

2.2. Phương pháp đơn hình mở rộng

Phương pháp đơn hình mở rộng còn gọi là phương pháp đánh thuế M được áp dụng để để giải BTQH TT có biến giả.

Ví dụ:

$$z = 8x_1 + 6x_2 \rightarrow \text{Max}$$

với các ràng buộc:

$$(a) \begin{cases} 4x_1 + 2x_2 \leq 60 \\ 2x_1 + 4x_2 \geq 48 \\ x_1, x_2 \geq 0 \end{cases}$$

hay: $z = 8x_1 + 6x_2 + 0x_3 + 0x_4 \rightarrow \text{Max}$ với các ràng buộc

$$(b) \begin{cases} 4x_1 + 2x_2 + x_3 = 60 \\ 2x_1 + 4x_2 - x_4 = 48 \\ x_1, x_2, x_3, x_4 \geq 0 \end{cases}$$

Ta có thể đưa bài toán về dạng chính tắc sau gọi là bài toán M:

$\text{Max } z = 8x_1 + 6x_2 + 0x_3 + 0x_4 - Mx_5$ (trong đó $M \approx +\infty$) với các ràng buộc

$$(c) \begin{cases} 4x_1 + 2x_2 + x_3 = 60 \\ 2x_1 + 4x_2 - x_4 + x_5 = 48 \\ x_1, x_2, x_3, x_4, x_5 \geq 0 \end{cases}$$

Cách 1: Có thể giải BTQH TT với các điều kiện ràng buộc (a) bằng phương pháp đồ thị để nhận được kết quả: phương án tối ưu là $(x_1 = 0, x_2 = 30)$ và $z_{\max} = 180$.

Cách 2: Giải BTQH TT với các điều kiện ràng buộc (c) bằng cách lập bảng đơn hình như thông thường nhưng chú ý hệ số $M \approx +\infty$ (xem bảng I.2).

Bảng I.2. Các bảng đơn hình giải bài toán M

Hệ số hàm mục tiêu	Biến cơ sở	Phương án	8	6	0	0	-M
			x_1	x_2	x_3	x_4	x_5
0	x_3	60	4	2	1	0	0
-M	x_5	48	2	4	0	-1	+1
Hàng z		$z_0 = -48M$	$z_1 = -2M$	$z_2 = -4M$	$z_3 = 0$	$z_4 = M$	$z_5 = -M$
Hàng Δ_j			$\Delta_1 = 8+2M$	$\Delta_2 = 6+4M$	$\Delta_3 = 0$	$\Delta_4 = -M$	$\Delta_5 = 0$
0	x_3	36	3	0	1	1/2	-1/2
6	x_2	12	1/2	1	0	-1/4	1/4
Hàng z		72	3	6	0	-3/2	3/2
Hàng Δ_j			5	0	0	3/2	-M - 3/2
0	x_4	72	6	0	2	1	-1
6	x_2	30	2	1	1/2	0	0
Hàng z		180	12	6	3	0	0
Hàng Δ_j			-4	0	-3	0	-M

Tại bảng đơn hình cuối cùng, ta thấy $\Delta_j \leq 0 \forall j$ nên phương án tối ưu đã đạt được với $x_2 = 30, x_4 = 72$, các x_j khác $= 0$ và $z_{\text{Max}} = 180$.

Lưu ý

– Khi một biến giả đã được đưa ra khỏi cơ sở thì không bao giờ quay lại nữa. Do đó ta có thể xoá cột biến giả đó khỏi bảng đơn hình.

– Nếu dấu hiệu dừng xuất hiện ($\Delta_j \leq 0 \forall j$) nhưng vẫn còn biến giả với giá trị dương trong số các biến cơ sở thì điều này chứng tỏ bài toán ban đầu không thể có phương án khả thi (có thể chứng minh bằng phản chứng).

– Với ví dụ trên (xem bảng I.2) ta thấy quá trình giải chia làm hai pha: pha 1 nhằm giải bài toán M cho tới khi biến giả (x_5) được đưa ra khỏi số biến cơ sở (lúc này có phương án cực biên xuất phát cho bài toán (b)) và pha 2 nhằm tìm phương án tối ưu cho bài toán (b).

– Phần mềm tính toán Lingo có thể giải được tất cả các BTQH TT không đòi hỏi người dùng phải đưa chúng về dạng chính tắc.

3. Mô hình quy hoạch tuyến tính đa mục tiêu

3.1. Các khái niệm cơ bản

Phát biểu mô hình

Trong các bài toán kỹ thuật, công nghệ, quản lý, kinh tế nông nghiệp v.v... nảy sinh từ thực tế, chúng ta thường phải xem xét để tối ưu hoá đồng thời một lúc nhiều mục tiêu. Các mục tiêu này thường là khác về thứ nguyên, tức là chúng được đo bởi các đơn vị khác nhau. Những tình huống như vậy tạo ra các bài toán tối ưu đa mục tiêu. Như vậy, chúng ta cần phải tối ưu hoá (cực đại hoá hoặc cực tiểu hoá tùy theo tình huống thực tế) không phải là chỉ một mục tiêu nào đó, mà là đồng thời tất cả các mục tiêu đã đặt ra.

Bài toán tối ưu đa mục tiêu mà trong đó miền ràng buộc D là tập lồi đa diện và các mục tiêu $z_i = f_i(X)$, với $i = 1, 2, \dots, p$, là các hàm tuyến tính xác định trên D , được gọi là bài toán quy hoạch tuyến tính đa mục tiêu. Khi đó, ta có mô hình toán học sau đây được gọi là mô hình quy hoạch tuyến tính đa mục tiêu :

Max CX với ràng buộc $X \in D$, trong đó:

C là ma trận cấp $p \times n$

$D = \{ X \in R^n : AX \leq B \}$

với A là ma trận cấp $m \times n$ và $B \in R^m$.

Ví dụ: BTQH TT với hai mục tiêu

$f_1(X) = x_1 + 2x_2 \rightarrow \text{Min}$ hay $z_1 = f_1(X) = -x_1 - 2x_2 \rightarrow \text{Max}$

$z_2 = f_2(X) = 2x_2, \rightarrow \text{Max}$

với các ràng buộc

$$\begin{cases} -x_1 + x_2 \leq 3 \\ x_1 + x_2 \geq 3 \\ x_1, x_2 \geq 0 \end{cases}$$

Ta có thể viết bài toán này dưới dạng ma trận như sau: Max CX với ràng buộc $X \in D = \{X \in \mathbb{R}^2 : AX \leq B\}$, trong đó $X = (x_1, x_2)^T$, $B = (3, -3, 0, 0)^T$, còn

$$C = \begin{bmatrix} -1 & -2 \\ 0 & 2 \end{bmatrix}, \quad A = \begin{bmatrix} -1 & 1 \\ -1 & -1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Có thể nói, BTQH TT đa mục tiêu là BTQH TT mà trong đó chúng ta phải tối ưu hoá cùng một lúc nhiều mục tiêu. Tuy nhiên, các mục tiêu này thường đối chọi cạnh tranh với nhau. Việc làm tốt hơn mục tiêu này thường dẫn tới việc làm xấu đi một số mục tiêu khác. Vì vậy việc giải các bài toán tối ưu đa mục tiêu, tức là tìm ra một phương án khả thi tốt nhất theo một nghĩa nào đó, thực chất chính là một bài toán ra quyết định. Có thể thấy lại ở đây một lần nữa khẳng định "Tối ưu hoá chính là công cụ định lượng chủ yếu nhất của quá trình ra quyết định".

Hiện tại các tài liệu, sách chuyên khảo, tạp chí cập nhật về lĩnh vực liên ngành Toán – Tin, Khoa học quản lí, Công nghệ, Kinh tế, Điện, Cơ khí nông nghiệp,... đề cập rất nhiều tới bài toán tối ưu đa mục tiêu. Vấn đề nghiên cứu cơ sở lí thuyết, thuật toán, lập mô hình, xây dựng hệ máy tính trợ giúp quyết định, và áp dụng các mô hình tối ưu đa mục tiêu cho các quá trình công nghệ, quản lí,... là một vấn đề liên ngành được rất nhiều nhà khoa học và kĩ sư thực hành quan tâm.

Phương án tối ưu Pareto

Khái niệm then chốt trong tối ưu hoá đa mục tiêu là khái niệm phương án tối ưu Pareto.

Định nghĩa: Một phương án tối ưu Pareto X^* có tính chất sau đây:

– Trước hết nó phải thuộc vào miền các phương án khả thi của bài toán, tức là phải thoả mãn tất cả các ràng buộc: $X^* \in D$.

– Với mọi phương án khả thi khác $X \in D$ mà có một mục tiêu nào đó tốt hơn ($f_i(X)$ tốt hơn $f_i(X^*)$) thì cũng phải có ít nhất một mục tiêu khác xấu hơn ($f_j(X)$ xấu hơn $f_j(X^*)$, $j \neq i$).

Nói một cách khác, không tồn tại một phương án khả thi nào $X \in D$ có thể trội hơn X^* trên tổng thể.

Để minh hoạ định nghĩa trên, ta xét ví dụ đã cho.

Ví dụ:

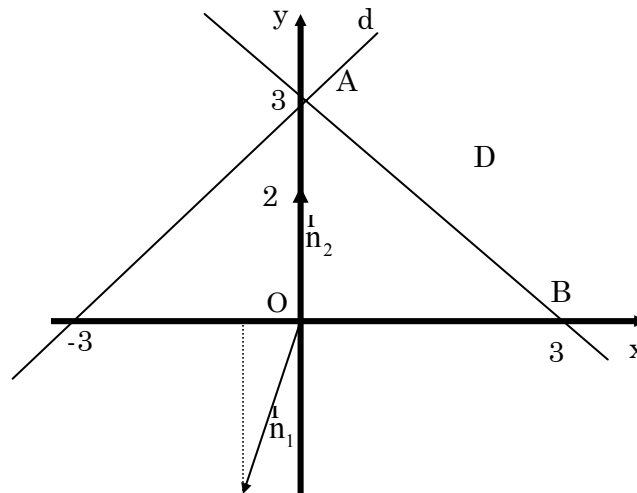
Xét BTQH TT với hai mục tiêu.

$$f_1(X) = x_1 + 2x_2 \rightarrow \text{Min}$$

$$f_2(X) = 2x_2 \rightarrow \text{Max}$$

Với các ràng buộc

$$\begin{cases} -x_1 + x_2 \leq 3 \\ x_1 + x_2 \geq 3 \\ x_1, x_2 \geq 0 \end{cases}$$



Hình I.6. Minh họa đồ thị BTQH TT hai mục tiêu

Miền các phương án khả thi D (miền giới hạn bởi đoạn AB và các tia Ad, Bx) được biểu thị trên hình I.6. $\vec{n}_1(-1, -2)$ là hướng giảm của mục tiêu 1, còn $\vec{n}_2(0, 2)$ là hướng tăng của mục tiêu 2.

Lúc này A(0, 3) cũng như B(3, 0) là hai phương án tối ưu Pareto của bài toán trên. Dễ thấy tập hợp P tất cả các phương án tối ưu Pareto bao gồm các điểm nằm trên đoạn AB và Ad.

3.2. Một số phương pháp giải BTQH TT đa mục tiêu

Định nghĩa 1

Giải bài toán tối ưu toàn cục đa mục tiêu là chọn ra từ tập hợp P các phương án tối ưu Pareto của bài toán một (hoặc một số) phương án tốt nhất (thỏa mãn nhất) theo một nghĩa nào đó dựa trên cơ cấu ưu tiên của người ra quyết định.

Trong ví dụ trên, tùy theo cơ cấu ưu tiên của người ra quyết định, chúng ta có thể chọn ra một hoặc một số điểm tối ưu Pareto nằm trên AB hoặc tia Ad làm phương án tối ưu của bài toán.

Cách 1: Bằng một phương pháp tối ưu toán học thích hợp tìm ra tập hợp P tất cả các phương án tối ưu Pareto. Người ra quyết định sẽ đề ra cơ cấu ưu tiên của mình đối với tập P nhằm tìm ra phương án tối ưu Pareto thoả mãn nhất cho bài toán đa mục tiêu ban đầu.

Cách 2: Việc tìm tập hợp P trong trường hợp các bài toán nhiều biến là khá khó và mất nhiều thời gian. Vì vậy, so với cách 1, cách 2 sẽ tiến hành theo trình tự ngược lại. Trước hết người ra quyết định sẽ đề ra cơ cấu ưu tiên của mình. Dựa vào cơ cấu ưu tiên đó, các mục tiêu sẽ được tổ hợp vào một mục tiêu duy nhất, tiêu biểu cho hàm tổng tiện ích của bài toán. Bài toán tối ưu với hàm mục tiêu tổ hợp này sẽ được giải bằng một phương pháp tối ưu toán học thích hợp, để tìm ra một (hoặc một số) phương án tối ưu Pareto. Lúc này, người ra quyết định sẽ chọn ra trong số các phương án tối ưu Pareto đó một phương án tốt nhất.

Chúng ta sẽ tiếp tục phân tích cách thứ 2. Rõ ràng, người ra quyết định không thể đề ra cơ cấu ưu tiên của mình một cách chính xác ngay từ đầu. Trong quá trình giải bài toán, trong mỗi bước lặp, sau khi xem xét lại cơ cấu ưu tiên đã đề ra, cũng như phương án trung gian vừa tìm được, người ra quyết định có thể dựa vào các thông tin đó để thay đổi lại cơ cấu ưu tiên của mình. Sau đó, quá trình giải lại được tiếp tục, cho tới khi một phương án tối ưu cuối cùng được đưa ra.

Định nghĩa 2

Phương pháp giải bài toán tối ưu đa mục tiêu dựa trên sự trợ giúp của hệ máy tính, nhằm giúp người ra quyết định từng bước thay đổi các quyết định trung gian một cách thích hợp để đi tới một phương án tối ưu Pareto thoả mãn nhất, được gọi là phương pháp tương tác người – máy tính.

Phương pháp tương tác người – máy tính giải bài toán tối ưu đa mục tiêu có các yếu tố cấu thành sau:

- Cơ cấu ưu tiên của người ra quyết định và hàm tổ hợp tương ứng.
- Kiểu tương tác người – máy tính: cho biết các thông tin nào máy tính phải đưa ra lại trong các bước lặp trung gian, và cách thay đổi các thông số của cơ cấu ưu tiên từ phía người ra quyết định.
- Kỹ thuật tối ưu toán học được xây dựng dựa trên lí thuyết tối ưu hoá nhằm tìm ra các phương án tối ưu Pareto cho các bài toán cần giải trong các bước lặp trung gian.

Cho tới thời điểm hiện nay, hàng chục phương pháp giải BTQHTT đa mục tiêu đã được đề cập tới trong các tạp chí chuyên ngành, mà đa số chúng đều có những ứng dụng rất thành công trong nhiều lĩnh vực, như: phương pháp tham số, phương pháp nón pháp tuyến, phương pháp véc tơ cực đại, phương pháp trọng số tương tác của Chebysev, phương pháp thoả dụng mờ tương tác của Nguyễn Hải Thanh.

3.3. Phương pháp thoả dụng mờ tương tác giải BTQHTT đa mục tiêu

Thuật giải

a. Bước khởi tạo

– Nhập số liệu cho các hàm mục tiêu tuyến tính z_i ($i = 1, 2, \dots, p$) và m điều kiện ràng buộc.

– Giải BTQHTT cho từng mục tiêu z_i ($i = 1, 2, \dots, p$) với m ràng buộc ban đầu, thu được các phương án tối ưu X^1, X^2, \dots, X^p (nếu với một mục tiêu nào đó bài toán không cho phương án tối ưu thì cần xem xét để chỉnh sửa lại các điều kiện ràng buộc ban đầu).

– Tính giá trị hàm mục tiêu tại p phương án X^1, X^2, \dots, X^p .

Lập bảng pay-off. Xác định giá trị cận trên z_i^B và giá trị cận dưới z_i^w của mục tiêu z_i ($i=1, 2, \dots, p$).

– Xác định các hàm thoả dụng mờ $\mu_1(z_1), \mu_2(z_2), \dots, \mu_p(z_p)$ cho từng mục tiêu dựa vào thông tin từ bảng pay-off theo công thức:

$$\mu_i(z_i) = \frac{z_i - z_i^w}{z_i^B - z_i^w}, \quad i = 1, 2, \dots, p.$$

– Đặt $k := 1$.

b. Các bước lặp (xét bước lặp thứ k)

Bước 1: Xây dựng hàm mục tiêu tổ hợp từ các hàm thoả dụng trên:

$$w_1\mu_1(z_1) + w_2\mu_2(z_2) + \dots + w_p\mu_p(z_p) \rightarrow \text{Max}$$

Trong đó: w_1, w_2, \dots, w_p là các trọng số phản ánh tầm quan trọng của từng hàm thoả dụng trong thành phần hàm tổ hợp, với

$$w_1 + w_2 + \dots + w_p = 1 \quad \text{và} \quad 0 \leq w_1, w_2, \dots, w_p \leq 1.$$

Bước 2:

– Giải BTQHTT với hàm mục tiêu tổ hợp và m ràng buộc ban đầu để tìm được phương án tối ưu của bước lặp thứ k là $X^{(k)}$ và giá trị của các hàm mục tiêu z_i cũng như của các hàm thoả dụng $\mu_i(z_i)$ (với $i = 1, 2, \dots, p$).

– Nếu người ra quyết định cảm thấy chưa thoả mãn với các giá trị đạt được của các hàm mục tiêu cũng như của các hàm thoả dụng thì phương án thu được $X^{(k)}$ chưa phải là phương án tối ưu thoả mãn nhất. Đặt $k := k + 1$, quay về bước 1.

– Nếu người ra quyết định đã cảm thấy thoả mãn thì phương án thu được là $X^{(k)}$. Chuyển sang bước 3.

Bước 3: Kết thúc.

Ví dụ: Giải BTQHHT hai mục tiêu.

$$z_1 = 8x_1 + 6x_2 \rightarrow \text{Max}$$

$$z_2 = x_1 + 3x_2 \rightarrow \text{Max}$$

với các ràng buộc:

$$(D) \begin{cases} 4x_1 + 3x_2 \leq 60 \\ 2x_1 + 4x_2 \leq 48 \\ x_1, x_2 \geq 0 \end{cases}$$

a. Bước khởi tạo

– Giải BTQHHT cho từng mục tiêu trong ví dụ trên ta có hai bài toán: Max $z_1 = 8x_1 + 6x_2 \rightarrow \text{Max}$ với điều kiện ràng buộc (D) cho phương án tối ưu $X^1(12, 6)$ và Max $z_1 = 132$;

$z_2 = x_1 + 3x_2 \rightarrow \text{Max}$ cho phương án tối ưu $X^2(0, 12)$ và Max $z_2 = 36$.

Như vậy miền phương án tối ưu Pareto chính là mọi phương án thuộc AB (xem hình I.1), với A(0, 12) và B(12, 6). $\frac{1}{n_1}(8, 6)$ là hướng tăng của mục tiêu 1, còn $\frac{1}{n_2}(1, 3)$ là hướng tăng của mục tiêu 2. Do đó, khi chọn phương án tối ưu Pareto dịch dần từ B về A thì z_1 giảm, z_2 tăng. Cần tìm phương án tối ưu Pareto "thoả mãn nhất" thuộc AB bằng cách "thương lượng" giữa z_1 và z_2 .

– Lập bảng pay-off cho các mục tiêu

Phương án X_i	z_1	z_2
$X^1(12, 6)$	132	30
$X^2(0, 12)$	72	36

Dựa trên thông tin của bảng pay-off, ta có $z_1^W = 72$, $z_1^B = 132$; còn $z_2^W = 30$, $z_2^B = 36$. Do đó, đoạn biến thiên cần xét cho z_1 là $[72, 132]$ và cho z_2 là $[30, 36]$. Từ đó chúng ta có thể thiết lập các hàm thoả dụng mờ ứng với hai mục tiêu đã cho như sau:

$$\mu_1(z_1) = \frac{z_1 - z_1^W}{z_1^B - z_1^W} = \frac{z_1 - 72}{132 - 72} = \frac{z_1}{60} - \frac{72}{60} = \frac{z_1}{60} - 1,2$$

Hàm thoả dụng mờ trên đây phụ thuộc vào z_1 , nên phụ thuộc vào (x_1, x_2) . Khi có một phương án khả thi (x_1, x_2) ta tính được độ thoả dụng $\mu_1(z_1)$ đối với mục tiêu z_1 . Tương tự đối với z_2 ta có hàm thoả dụng mờ:

$$\mu_2(z_2) = \frac{z_2 - z_2^W}{z_2^B - z_2^W} = \frac{z_2 - 30}{36 - 30} = \frac{z_2}{6} - 5.$$

Lập hàm thoả dụng tổ hợp $u = w_1 \mu_1(z_1) + w_2 \mu_2(z_2)$, trong đó w_1, w_2 là các trọng

số thoả mãn $0 \leq w_1, w_2 \leq 1$ và $w_1 + w_2 = 1$.

b. Các bước lập

Xét $w_1 = 0,5$ và $w_2 = 0,5$, thì có $u = 0,5 \left(\frac{z_1}{60} - 1,2 \right) + 0,5 \left(\frac{z_2}{6} - 5 \right) = \left(\frac{z_1}{120} + \frac{z_2}{12} \right) +$

1,9. Để cực đại hoá hàm thoả dụng tổ hợp, ta chỉ cần tìm $\text{Max} \left\{ \frac{z_1}{120} + \frac{z_2}{12} \right\}$. Vậy chúng

ta cần giải bài toán: $\text{Max} u = \frac{z_1}{120} + \frac{z_2}{12}$ với các ràng buộc (D), hay bài toán tương

đương:

$z = 120u/18 = x_1 + 2x_2 \rightarrow \text{Max}$ với các ràng buộc (D). Giải BTQHHT này ta sẽ có kết quả (0, 12). Nếu thay đổi giá trị của bộ trọng số (w_1, w_2) thì sẽ thu được các phương án Pareto khác nhau, nhưng đều nằm trên đoạn AB.

Chú ý: Chú ý rằng, với mỗi bộ trọng số (w_1, w_2) ta tìm được một phương án tối ưu Pareto (x_1, x_2) bằng cách làm cực đại hoá hàm thoả dụng tổ hợp u đã thiết lập được. Cứ thế chọn các bộ trọng số w_1, w_2 khác nhau ta có các phương án tối ưu Pareto khác nhau. Từ các phương án đó, chúng ta chọn được phương án tối ưu Pareto tốt nhất.

Phần mềm MULTIOPT phiên bản 1.0 được xây dựng dựa trên phương pháp thoả dụng mờ tương tác nhằm giải BTQHHT đa mục tiêu. Chương trình được xây dựng bằng ngôn ngữ Visual Basic, có giao diện đơn giản và dễ dùng. Các chức năng chính của phần mềm bao gồm:

- Chức năng nhập dữ liệu cho BTQHHT đa mục tiêu một cách trực quan: nhập dữ liệu qua bàn phím hoặc nhập dữ liệu từ tệp.
- Xuất bài toán đã được nhập ra tệp.
- Giải bài toán: theo dõi bảng pay-off và giải theo phương pháp trọng số.
- Xuất kết quả ra tệp: xuất kết quả trung gian ra tệp / xuất kết quả cuối cùng ra tệp. Sau khi đã cài đặt chương trình, một liên kết đến chương trình sẽ được tạo tại menu *Start > Programs* của Windows. Để khởi động chương trình, kích hoạt menu *Start > Programs > Tối ưu đa mục tiêu*.

Sau đó, cần phải chọn bài toán bằng cách vào menu *Chương trình*, để chọn *Bài toán tuyến tính*.

Người sử dụng có thể nhập dữ liệu theo một trong hai cách sau:

Nhập dữ liệu từ bàn phím

Chọn menu *Xử lý bài toán > Nhập dữ liệu > Nhập từ bàn phím*.

Các số liệu đầu vào của chương trình bao gồm:

- n – số biến;
- m – số các ràng buộc;
- m1 – số các ràng buộc \leq ;

- m2 - số các ràng buộc \geq ;
- p - số các mục tiêu;
- ep - số dương đủ nhỏ;
- gz - số dương đủ lớn.

Nhập dữ liệu từ tệp

Thực hiện lệnh *Xử lí bài toán*>*Nhập dữ liệu*>*Nhập từ tệp*, sau đó chọn tệp dữ liệu.

Ví dụ:

Giải bài toán quy hoạch tuyến tính ba mục tiêu.

$$z_1 = x_1 - 2x_2 + 3x_4 \rightarrow \text{Max}$$

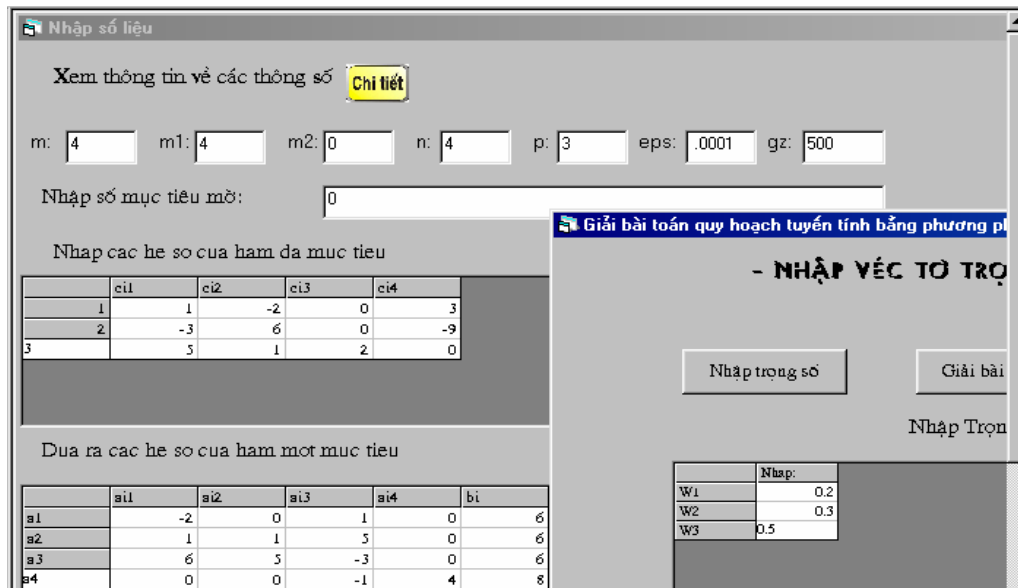
$$z_2 = -3x_1 + 6x_2 - 9x_4 \rightarrow \text{Max}$$

$$z_3 = 5x_1 + x_2 + 2x_3 \rightarrow \text{Max}$$

với các ràng buộc:

$$\begin{cases} -2x_1 + x_3 \leq 6; \\ x_1 + x_2 + 5x_3 \leq 6; \\ 6x_1 + 5x_2 - 3x_3 \leq 6; \\ -x_3 + 4x_4 \leq 8; \\ x_1, x_2, x_3, x_4 \geq 0. \end{cases}$$

Như vậy, ở đây chúng ta muốn cực đại hoá đồng thời ba hàm mục tiêu trên miền D các phương án khả thi thoả mãn đồng thời bốn ràng buộc và điều kiện không âm của các biến.

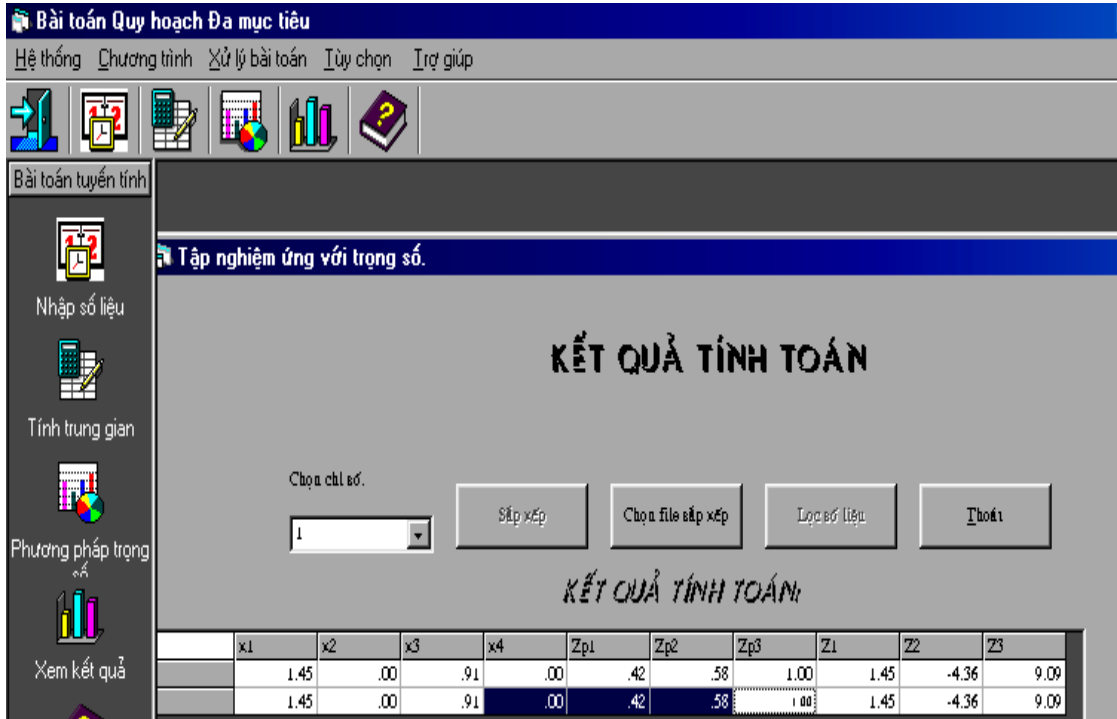


Hình 1.7. Nhập dữ liệu từ bàn phím

Trong bước khởi tạo, ta nhập dữ liệu từ bàn phím như trên hình 1.7.

– Sau khi đã nhập dữ liệu, thực hiện tính các bước trung gian (tính các giá trị trong bảng pay-off, tính các giá trị cho các hàm thoả dụng) bằng cách chọn *Xử lí bài toán*>*Tính toán*>*Tính trung gian*.

– Thực hiện *Xử lí bài toán*>*Tính toán*>*Phương pháp trọng số*. Trên màn hình xuất hiện form nhập các giá trị trọng số (xem hình 1.7: các giá trị của trọng số được nhập sao cho có tổng là 1). Sau khi nhập các giá trị trọng số nhấn vào nút *Giải bài toán* để giải bài toán (xem hình 1.8).



Hình 1.8. Nhập các giá trị trọng số và giải bài toán

Kết quả trên hình 1.8 là $x_1 = 1,45$, $x_2 = 0$, $x_3 = 0,91$, $x_4 = 0$, $\mu_1(z_1) = 0,42$, $\mu_2(z_2) = 0,58$, $\mu_3(z_3) = 1$; $z_1 = 1,45$, $z_2 = -4,36$ và $z_3 = 9,09$ ứng với bộ trọng số $w_1 = 0,2$, $w_2 = 0,3$ và $w_3 = 0,5$.

4. Mô hình tối ưu phi tuyến đơn và đa mục tiêu

4.1. Một số khái niệm cơ bản

Mô hình tối ưu tổng quát

Mô hình tối ưu tổng quát, hay bài toán tối ưu tổng quát, có dạng:

$$F(X) \rightarrow \text{Min (Max)} \text{ với } X \in D \subset R^n.$$

Ở đây $F(X)$ có thể là một hàm vô hướng hay hàm véc tơ, tuyến tính hay phi tuyến. Trong trường hợp $F(X)$ là hàm vô hướng thì ta có mô hình tối ưu đơn mục tiêu, còn nếu F là hàm véc tơ thì có mô hình tối ưu đa mục tiêu. D được gọi là miền ràng buộc hay miền phương án khả thi, thường được biểu diễn bởi các đẳng thức và/hoặc các bất đẳng thức.

Mô hình tối ưu phi tuyến đơn mục tiêu

Dạng chính tắc của bài toán tối ưu một mục tiêu được biểu diễn như sau:

$$f(X) \rightarrow \text{Min (Max)}, \quad X = (x_1, x_2, \dots, x_n) \in R^n,$$

với:

$$(i) \quad g_j(X) \leq 0, \quad j = 1, 2, \dots, k,$$

$$(ii) \quad g_j(X) = 0, \quad j = k+1, k+2, \dots, m,$$

Trong các bài toán thực tế có thể bổ sung các ràng buộc

$$(iii) \quad a_i \leq x_i \leq b_i, \quad i = 1, 2, \dots, n.$$

Trong trường hợp hoặc hàm mục tiêu $f(X)$ hoặc có ít nhất một trong các hàm ràng buộc $g_j(X)$, $j = 1, 2, \dots, m$, là hàm phi tuyến, chúng ta có bài toán tối ưu phi tuyến. Khi tất cả các tọa độ x_i đều bắt buộc nhận các giá trị nguyên, $i = 1, 2, \dots, n$, thì ta có bài toán tối ưu nguyên. Còn nếu chỉ có một số tọa độ (nhưng không phải tất cả các tọa độ) bắt buộc nhận giá trị nguyên thì ta có bài toán tối ưu hỗn hợp nguyên.

Kí hiệu D là miền các phương án (miền ràng buộc) cho bởi các ràng buộc (i), (ii) và/hoặc (iii) thì bài toán tối ưu trên đây có thể viết gọn hơn như sau:

$$f(X) \rightarrow \text{Min (Max) với } X \in D.$$

Lúc này, đối với bài toán cực tiểu hoá, $X^* \in D$ được gọi là phương án tối ưu toàn cục nếu $\forall X \in D$ ta luôn có: $f(X^*) \leq f(X)$. Trong trường hợp $f(X^*) \leq f(X)$ chỉ đúng với $\forall X \in D$ trong một lân cận nào đó của X^* thì X^* được gọi là phương án tối ưu địa phương. Một cách tương tự, ta có thể định nghĩa khái niệm phương án tối ưu toàn cục hoặc địa phương cho bài toán cực đại hoá. Nếu chúng ta chỉ quan tâm tới việc tìm kiếm phương án tối ưu toàn cục thì ta có *bài toán tối ưu toàn cục*.

Trong các bài toán tối ưu phi tuyến ứng dụng nói chung, trong lĩnh vực cơ khí – điện lực nói riêng, phương án tối ưu toàn cục có một ý nghĩa quan trọng. Chẳng hạn trong thiết kế máy nông nghiệp, sau khi dùng phương pháp phân tích hồi quy nhiều chiều, ta thường thu được hàm mục tiêu $f(X)$ có dạng phi tuyến. Bài toán đặt ra là phải tìm được phương án tối ưu toàn cục.

Có rất nhiều phương pháp giải các lớp bài toán tối ưu phi tuyến, nhưng chưa có phương pháp nào tỏ ra hữu hiệu cho mọi bài toán tối ưu phi tuyến, đặc biệt là các bài toán tối ưu nguyên và hỗn hợp nguyên.

Mô hình tối ưu phi tuyến đa mục tiêu

Mô hình tối ưu đa mục tiêu có dạng:

$$z_j = f_j(X) \rightarrow \text{Min (Max)}, \quad X = (x_1, x_2, \dots, x_n), \quad j = 1, 2, \dots, p \quad (p \geq 2)$$

với:

$$(i) \quad g_j(X) \leq 0, \quad j = 1, 2, \dots, k,$$

$$(ii) \quad g_j(X) = 0, \quad j = k+1, k+2, \dots, m,$$

Trong các bài toán thực tế có thể bổ sung các ràng buộc

$$(iii) \quad a_i \leq x_i \leq b_i, \quad i = 1, 2, \dots, n.$$

Trong mô hình này, ta có p mục tiêu cần tối ưu hoá, các hệ số của các hàm mục

tiêu và ràng buộc nói chung được giả sử là các giá trị thực xác định. Trong trường hợp có ít nhất một trong các hàm mục tiêu hay các hàm ràng buộc là hàm phi tuyến, chúng ta có bài toán tối ưu phi tuyến đa mục tiêu.

Đối với bài toán tối ưu phi tuyến đa mục tiêu chúng ta cũng có khái niệm phương án tối ưu Pareto như đã trình bày trong mục 3.1 và 3.2 đối với BTQHTT đa mục tiêu.

Cũng như đối với các BTQHTT đa mục tiêu, phương pháp giải bài toán tối ưu phi tuyến đa mục tiêu dựa trên sự trợ giúp của hệ máy tính, nhằm giúp người ra quyết định từng bước thay đổi các quyết định trung gian một cách thích hợp để đi tới một phương án tối ưu Pareto thoả mãn nhất, được gọi là phương pháp tương tác người–máy tính.

4.2. Một số phương pháp và phần mềm giải bài toán tối ưu phi tuyến đơn mục tiêu

Các phương pháp giải bài toán tối ưu toàn cục

Các phương pháp giải bài toán tối ưu toàn cục phi tuyến đơn mục tiêu được phân ra thành hai lớp: phương pháp tất định (*deterministic methods*) và phương pháp ngẫu nhiên (*stochastic methods*).

Phương pháp tất định sử dụng các tính chất giải tích của hàm mục tiêu và các hàm ràng buộc. Một số dạng bài toán tối ưu toàn cục với những tính chất giải tích nhất định của hàm mục tiêu và các hàm ràng buộc có thể giải được bằng các phương pháp tất định thích hợp, chẳng hạn như phương pháp quy hoạch toàn phương, quy hoạch tách, quy hoạch lồi, quy hoạch d.c... Trong các trường hợp đó phương án tối ưu toàn cục có thể tìm được sau một số hữu hạn bước tính toán với độ chính xác chọn trước. Tuy nhiên, đối với nhiều lớp bài toán tối ưu toàn cục phương pháp tất định tỏ ra không có hiệu quả.

Trong khi đó, các phương pháp ngẫu nhiên như: phương pháp đa khởi tạo (*multistart*), mô phỏng tôi (*simulated annealing*), thuật giải di truyền (*genetic algorithm*),... có thể áp dụng để giải các bài toán tối ưu toàn cục dạng bất kỳ, không đòi hỏi các tính chất đặc biệt của hàm mục tiêu hay các hàm ràng buộc. Các phương pháp ngẫu nhiên đặc biệt tỏ ra có hiệu quả đối với các bài toán tối ưu phi tuyến nguyên và hỗn hợp nguyên. Tuy nhiên, các phương pháp này thường chỉ cho phương án “gần” tối ưu khá tốt sau một số hữu hạn bước mà không kiểm soát được độ chính xác của phương án tìm được.

Phần mềm Lingo giải bài toán quy hoạch toàn phương

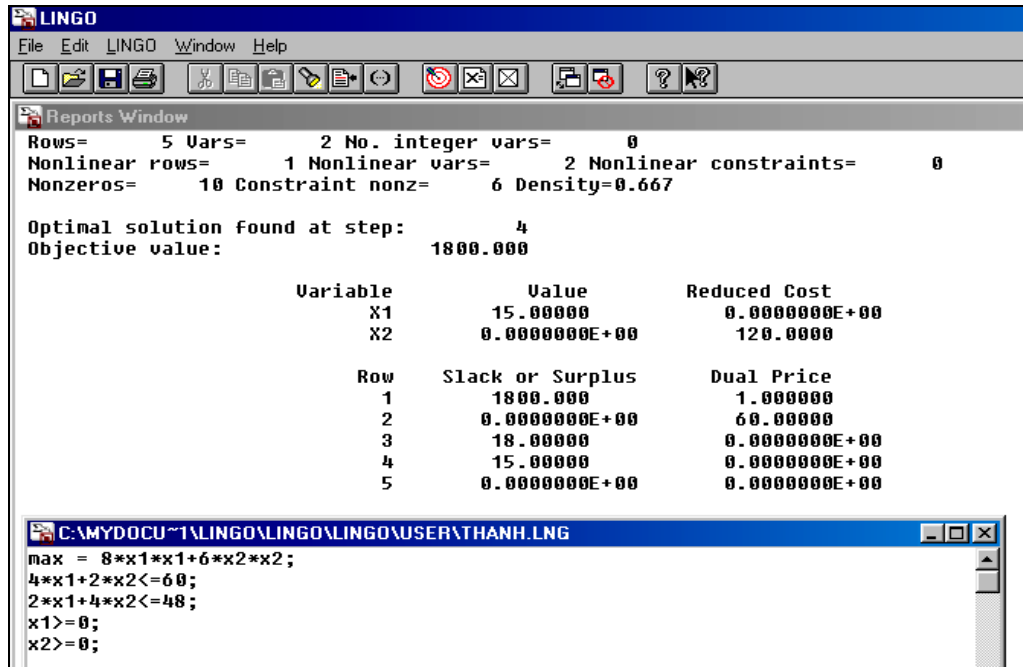
Ví dụ: Giải bài toán tối ưu phi tuyến dạng toàn phương

$$z = 8x_1^2 + 6x_2^2 \rightarrow \text{Max},$$

với các ràng buộc:

$$\begin{cases} 4x_1 + 2x_2 \leq 60 \\ 2x_1 + 4x_2 \leq 48 \\ x_1, x_2 \geq 0. \end{cases}$$

Để giải bài toán trên, chúng ta nhấn vào biểu tượng Lingo trên màn hình để vào cửa sổ Lingo. Sau đó thực hiện các lệnh Lingo: *Menu > New > <Untitled>* và gõ vào các dữ liệu của bài toán (tương tự như khi giải BTQHTT bằng phần mềm Lingo, xem lại mục 1.4, hình I.4).



Hình I.9. Kết quả bài toán quy hoạch toàn phương trong Lingo.

Tiếp theo, cần nhấp chuột vào nút LINGO và giải bài toán để thu được kết quả chi tiết như trên hình I.9. Kết quả trên cho ta biết giá trị cực đại của hàm mục tiêu là 180 với phương án tối ưu là: $x_1 = 15$, $x_2 = 0$. Các giá trị tối ưu của các biến đối ngẫu là $y_1 = 5/3$ và $y_2 = y_3 = y_4 = 0$.

Giải bài toán tối ưu phi tuyến bằng phần mềm RST2ANU

Phần mềm RST2ANU 1.0 được sử dụng để giải các bài toán tối ưu toàn cục phi tuyến dạng tổng quát với các biến liên tục, các biến nguyên và cho các bài toán hỗn hợp nguyên. Quá trình xây dựng phương pháp tính toán tối ưu, thuật giải, cài đặt trên ngôn ngữ C và sau này là ngôn ngữ Visual C++ 6.0 cũng như chạy thử nghiệm kéo dài gần tám năm. Ngoài ưu điểm giải được các bài toán hỗn hợp nguyên, phần mềm có độ tin cậy rất cao trong việc tìm ra các phương án tối ưu toàn cục và có giao diện thân thiện đối với người sử dụng. Phần mềm đã được đóng gói tránh sao chép và có thể dùng để giải các bài toán lớn khi được cài đặt trên hệ máy tính mạnh.

Thuật giải

Thuật giải ngẫu nhiên RST2AN (hay RST2ANU), được đưa ra bởi C. Mohan và Nguyễn Hải Thanh. Thuật giải này là thuật giải tìm kiếm ngẫu nhiên có điều khiển, có kết hợp thuật toán mô phỏng tôi (SA). Thuật giải RST2AN là thuật giải lặp, bao gồm hai pha: pha cục bộ và pha toàn cục. Sau đây là thuật giải RST2AN được phát biểu một

cách ngắn gọn cho bài toán tối ưu chính tắc dạng cực tiểu hoá.

Trong pha toàn cục, một số lượng thích hợp đủ lớn các phương án khả thi được được phát sinh ra một cách ngẫu nhiên và lưu trữ trong mảng có tên A. Đánh dấu hai điểm có giá trị hàm mục tiêu lớn nhất và nhỏ nhất tương ứng là M và L.

Trong pha cục bộ, các phương án được xử lý nhằm thu được giá trị tốt hơn của hàm mục tiêu. Trong pha này, thuật giải xác định X là điểm được nội suy bậc hai dựa trên phương án L và hai phương án khác được chọn ngẫu nhiên trong mảng A. Nếu như X là phương án khả thi thì với $f(X) \leq f(M)$, M sẽ được thay thế bởi X trong mảng A; còn với $f(X) > f(M)$, M sẽ được thay thế bởi X với xác suất $p = \exp(-\beta(f(X)-f(M))/(f(X)-f(L)))$, trong đó $\beta > 0$ là tham số được lựa chọn thích hợp. Nếu X không phải là phương án khả thi, bỏ qua X và chọn hai phương án khác trong A một cách ngẫu nhiên rồi cùng với L tiếp tục sinh ra phương án mới. Quá trình cứ thế tiếp diễn như vậy cho tới khi tập hợp các phương án trong A sẽ có xu hướng co cụm lại xung quanh một phương án tối ưu toàn cục.

Ví dụ: Giải bài toán tối ưu phi tuyến hỗn hợp nguyên.

$$z = x_1^{0,6} + x_2^{0,6} + x_3^{0,4} + 2x_4 + 5x_5 - 4x_3 - x_6 \rightarrow \text{Min}$$

với các ràng buộc:

$$\left\{ \begin{array}{l} x_2 - 3x_1 - 3x_4 = 0; \\ x_3 - 2x_2 - 2x_5 = 0; \\ 4x_4 - x_6 = 0; \\ x_1 + 2x_4 \leq 4; \\ x_2 + x_5 \leq 4; \\ x_3 + x_6 \leq 6; \\ x_1 \leq 3; \quad x_2 \leq 4; \quad x_3 \leq 4; \\ x_4 \leq 1; \quad x_5 \leq 2; \quad x_6 \leq 6; \\ x_1, x_2, x_3, x_4, x_5, x_6 \geq 0; \\ x_4, x_5, x_6 \text{ là các biến nguyên.} \end{array} \right.$$

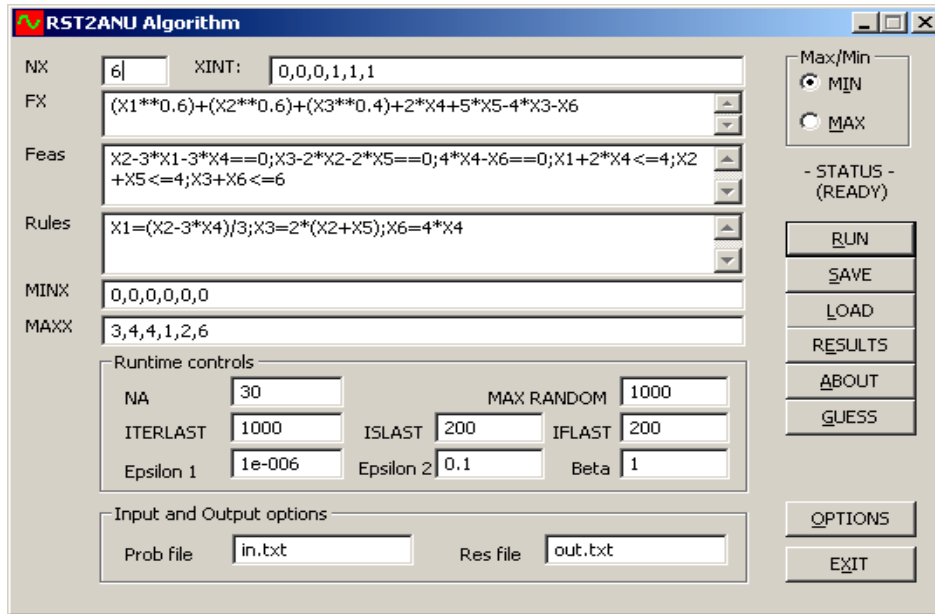
Hướng dẫn sử dụng

Chương trình được gói gọn trong một file chạy duy nhất mang tên rst2anu1.0.exe. Khi bắt đầu khởi động chương trình, người dùng sẽ được hỏi mã đăng kí sử dụng chương trình. Mỗi người dùng sẽ được cấp một mã đăng kí và phải có mã đăng kí mới sử dụng được chương trình, do đó chương trình không thể bị sao chép.

Sau khi nhập mã đăng kí, người dùng có thể nhập bài toán một cách dễ dàng (xem hình I.10) với:

– NX là số biến của bài toán.

– XINT xác định biến nguyên và biến không nguyên. Như trong hình trên, XINT = 0,0,0,1,1,1 cho biết ba biến đầu là biến liên tục, ba biến sau là biến nguyên.



Hình 1.10. Màn hình giao diện sau khi nhập xong dữ liệu

– FX là xâu xác định hàm ràng buộc, được nhập theo cú pháp của EvaluateExpression. Các biến được viết bằng kí hiệu “X” có kèm theo chỉ số. Ví dụ, X1 là biến thứ nhất, X5 là biến thứ 5.

– Nếu bài toán tối ưu là bài toán tìm cực tiểu thì lựa chọn ô MIN và ngược lại chọn ô MAX với bài toán tìm cực đại.

– Feas xâu cho biết các hàm ràng buộc, được nhập cách nhau bởi dấu chấm phẩy hoặc xuống dòng. Các xâu này cũng tuân theo cú pháp của EvaluateExpression.

– Rules là các xâu chỉ ra các luật. Ở đây, một luật có thể coi như là một lệnh gán giá trị của một biến bởi giá trị của một biểu thức các biến khác.

– MINX là mảng xác định cận dưới cho các biến, các giá trị viết cách nhau bởi dấu phẩy (,).

– MAXX là mảng xác định cận trên cho các biến, các giá trị viết cách nhau bởi dấu phẩy (,).

– NA là kích thước của mảng A (có thể chọn tùy ý, tối thiểu là $2(n + 1)$ với n là số biến của bài toán).

– MAX RANDOM là số lần cố gắng tối đa để tìm một phương án chấp nhận được bằng phương pháp ngẫu nhiên.

– ITERLAST, ISLAST, IFLAST là các giới hạn về số vòng lặp, số lần thất bại trong việc cải thiện giá trị hàm mục tiêu, số lần thất bại trong việc nội suy phương án mới chấp nhận được.

– Epsilon1, epsilon2 là các số dương đủ nhỏ nhằm xác định tiêu chuẩn co cụm của mảng A theo thuật giải.

– Beta là hằng số sử dụng trong công thức tính xác suất thay thế một phương án tốt hơn trong mảng A bởi một phương án tồi hơn.

– Prob file và Res file là các tệp đầu vào và tệp kết quả. Có thể soạn sẵn tệp bài toán đầu vào rồi nạp bài toán. Cũng có thể lưu một bài toán đã nhập ra tệp.

Chạy chương trình

Sau khi nhập bài toán hay nạp bài toán từ tệp, có thể chạy chương trình bằng cách kích chuột vào nút RUN. Trong khi chạy chương trình, ô trạng thái ở phía trên nút RUN sẽ xuất hiện dòng chữ SEARCHING. Khi thuật giải chạy xong thì ô trạng thái sẽ trở về READY cho biết đã sẵn sàng cho các bài toán tiếp theo. Mọi thông tin về phần mềm và cách sử dụng sẽ được biết nếu kích chuột vào nút ABOUT.

Sau khi chạy xong chương trình, kết quả chạy sẽ được xem trực tiếp khi kích chuột vào nút RESULTS và có thể lưu ra file văn bản, bao gồm phương án tối ưu, giá trị hàm mục tiêu, mảng A,... có cấu trúc như trên hình I.11.

```
*****RESULTS*****
fx= (X1**0.6)+(X2**0.6)+(X3**0.4)+2*X4+5*X5-4*X3-X6
ITER = 158/1000; ISMAX = 76/200; IFMAX = 1001/1000
X*   = 0.666667 2.000000 4.000000 0.000000 0.000000 0.000000
f(X*) = -11.959130

=====ARRAY A=====
0.666667 2.000000 4.000000 0.000000 0.000000 0.000000 -11.959130
0.666667 2.000000 4.000000 0.000000 0.000000 0.000000 -11.959130
0.666667 2.000000 4.000000 0.000000 0.000000 0.000000 -11.959130
0.666667 2.000000 4.000000 0.000000 0.000000 0.000000 -11.959130
0.666667 2.000000 4.000000 0.000000 0.000000 0.000000 -11.959130
0.666667 2.000000 4.000000 0.000000 0.000000 0.000000 -11.959130
0.666321 1.998962 3.997925 0.000000 0.000000 0.000000 -11.951906
0.663472 1.990417 3.980834 0.000000 0.000000 0.000000 -11.892427
0.662414 1.987243 3.974487 0.000000 0.000000 0.000000 -11.870339
0.659851 1.979553 3.959105 0.000000 0.000000 0.000000 -11.816829
0.640848 1.922544 3.845088 0.000000 0.000000 0.000000 -11.420634
0.640116 1.920347 3.840693 0.000000 0.000000 0.000000 -11.405380
```

Hình I.11. Cấu trúc file kết quả

Như vậy, bài toán đã được giải xong, với kết quả: $x_1 = 2/3$, $x_2 = 2$, $x_3 = 4$, $x_4 = 0$, $x_5 = 0$, $x_6 = 0$, và giá trị tối ưu của hàm mục tiêu là $-11,95913$.

Bài toán tối ưu thông số sàng phân loại

Chúng ta có thể sử dụng phần mềm RST2ANU để tìm nghiệm của hệ phương trình phi tuyến sau phát sinh trong việc tính toán một số thông số hình học và động học của cơ cấu sàng phân loại dao động (cần chú ý rằng nhiều phương pháp tính toán thông dụng khác của giải tích số đã tỏ ra không hiệu quả):

$$\left\{ \begin{array}{l} r \cos \varphi_1 + l \cos \varphi_2 + l''_3 \cos \varphi_3 + l_4 \cos \varphi_4 - x_{C1} = 0; \end{array} \right.$$

$$r \sin \varphi_1 + l \sin \varphi_2 + l'_3 \sin \varphi_3 + l_4 \sin \varphi_4 - y_{C1} = 0;$$

$$r \cos \varphi_1 + l \cos \varphi_2 + l'_3 \cos(\varphi_3 - \alpha) + l_5 \cos \varphi_5 - x_{D1} = 0;$$

$$r \sin \varphi_1 + l \sin \varphi_2 + l'_3 \sin(\varphi_3 - \alpha) + l_5 \sin \varphi_5 - y_{D1} = 0;$$

Trong hệ phi tuyến trên các thông số đã biết là: $r = 0,05\text{m}$; $l = 0,30\text{m}$; $l'_3 = 0,15\text{m}$; $l'_3 = 1,075\text{m}$; $l_3 = 1,025\text{m}$; $l_4 = 0,50\text{m}$; $l_5 = 0,40\text{m}$; $x_{C1} = 0,365\text{m}$; $y_{C1} = 0,635\text{m}$; $x_{D1} = 1,365\text{m}$; $y_{D1} = 0,635\text{m}$; $\alpha = \pi/8$.

Để sử dụng phần mềm RST2ANU giải hệ phương trình phi tuyến cho $\varphi_1 = k\pi/8$ ($k = 0, \dots, 9$), trước hết chúng ta cần thiết lập hàm mục tiêu sau:

$$z = (r \cos \varphi_1 + l \cos \varphi_2 + l'_3 \cos \varphi_3 + l_4 \cos \varphi_4 - x_{C1})^2 + (r \sin \varphi_1 + l \sin \varphi_2 + l'_3 \sin \varphi_3 + l_4 \sin \varphi_4 - y_{C1})^2 + (r \cos \varphi_1 + l \cos \varphi_2 + l'_3 \cos(\varphi_3 - \alpha) + l_5 \cos \varphi_5 - x_{D1})^2 + (r \sin \varphi_1 + l \sin \varphi_2 + l'_3 \sin(\varphi_3 - \alpha) + l_5 \sin \varphi_5 - y_{D1})^2 \rightarrow \text{Min.}$$

Kết quả được cho trong bảng I.3 với $z_{\min} = 0$.

Bảng I.3. Kết quả tính toán giá trị các thông số của sàng phân loại

$\varphi_1 \in [0, 2\pi]$	$\varphi_2 \in [0, \pi]$	$\varphi_3 \in [0, \pi]$	$\varphi_4 \in [0, \pi]$	$\varphi_5 \in [0, \pi]$
0	0,226128	0,551311	1,783873	1,666775
$\pi/18$	0,199269	0,550518	1,784628	1,670250
$2\pi/18$	0,170835	0,550590	1,782751	1,668853
$3\pi/18$	0,143343	0,550490	1,778826	1,663697
$4\pi/18$	0,112669	0,552073	1,770032	1,652171
$5\pi/18$	0,090986	0,551991	1,759350	1,639575
$6\pi/18$	0,066036	0,553576	1,745374	1,622823
$7\pi/18$	0,051284	0,554296	1,730174	1,602970
$8\pi/18$	0,039053	0,555262	1,713242	1,581813
$9\pi/18$	0,033773	0,556277	1,695605	1,560720

4.3. Một số phương pháp giải bài toán tối ưu phi tuyến đa mục tiêu

Phương pháp tương tác người-máy tính

Phương pháp PRELIME (*PREference Level Interactive Method*) hay còn gọi là phương pháp tương tác dựa trên mức ưu tiên do C. Mohan và Nguyễn Hải Thanh đề xuất. Còn phương pháp trọng số quy chuẩn là do Andrezej Osyczka đề xuất. Các phương pháp này đều thuộc lớp phương pháp tương tác người-máy tính giải bài toán tối ưu đa mục tiêu với các yếu tố cấu thành sau:

- Cơ cấu ưu tiên của người ra quyết định và hàm tổ hợp tương ứng.

– Kiểu tương tác người – máy tính: các thông tin nào máy tính phải đưa ra trong các bước lập trung gian, và cách thay đổi các thông số của cơ cấu ưu tiên từ phía người ra quyết định.

– Kỹ thuật tối ưu toán học được xây dựng dựa trên lý thuyết tối ưu hoá nhằm tìm ra các phương án tối ưu Pareto cho các bài toán cần giải trong các bước lập trung gian.

Bài toán thiết kế trục máy

Bài toán có hai mục tiêu sau:

– Mục tiêu 1 là cực tiểu hoá thể tích của trục máy

$$f_1(X) = 0,785[x_1(6400 - x_2^2) + (1000 - x_1)(1000 - x_2^2)] \text{ (mm}^3\text{)}$$

– Mục tiêu 2 là cực tiểu hoá độ nén tĩnh của trục

$$f_2(X) = 3,298 \times 10^{-5} \left[\left(\frac{1}{4,096 \times 10^7 - x_2^4} - \frac{1}{10^8 - x_2^4} \right) x_1^3 + \frac{10^9}{10^8 - x_2^4} \right]$$

(mm/N)

Trong đó, $X = (x_1, x_2)$ là véc tơ quyết định hay véc tơ phương án, với x_1, x_2 là các biến quyết định sau: x_1 – độ dài phần giáp nối trục, x_2 – đường kính trong của trục. Các thông số khác đã được thể hiện trong các hàm mục tiêu $f_1(X)$ và $f_2(X)$. Chúng ta cần chọn các giá trị cho các biến quyết định (còn gọi là các biến thiết kế) x_1, x_2 để tối ưu hoá đồng thời các mục tiêu 1 và 2 trong các điều kiện ràng buộc sau:

$$\begin{cases} g_1(X) = 180 - \frac{9,78 \times 10^6 x_1}{4,096 \times 10^7 - x_2^4} \geq 0, & (1) \\ g_2(X) = 75,2 - x_2 \geq 0, & (2) \\ g_3(X) = x_2 - 40 \geq 0, & (3) \\ g_4(X) = x_1 \geq 0, & (4) \end{cases}$$

trong đó các điều kiện (2), (3), (4) là dễ hiểu (ngoài ra, ta giả thiết $x_1 \leq 1000$), còn điều kiện (1) nảy sinh do yêu cầu sau: Một mặt, trục máy phải chịu đựng được tới mức tối đa lực $F_{\max} = 12000\text{N}$. Mặt khác, độ nén kết nối cho phép là 180N/mm .

Việc phát biểu bài toán tối ưu đa mục tiêu dưới dạng toán học (chính là việc lập mô hình toán học cho vấn đề phát sinh từ thực tế) là một khâu rất quan trọng nhằm mô tả tốt nhất hành vi của hệ thống đang được xem xét, mặt khác nhằm tìm ra được các phương pháp tối ưu hoá có hiệu quả để đi tới một phương án đủ tốt và mang lại lợi ích.

Sau đây, chúng ta hãy phân tích vắn tắt hai phương pháp giải bài toán thiết kế trục máy đã nêu ra ở trên.

Phương pháp trọng số quy chuẩn

Trong yếu tố cấu thành thứ nhất, hàm tổ hợp các mục tiêu cho bởi $f(X) = \omega_1 f_1(X)$

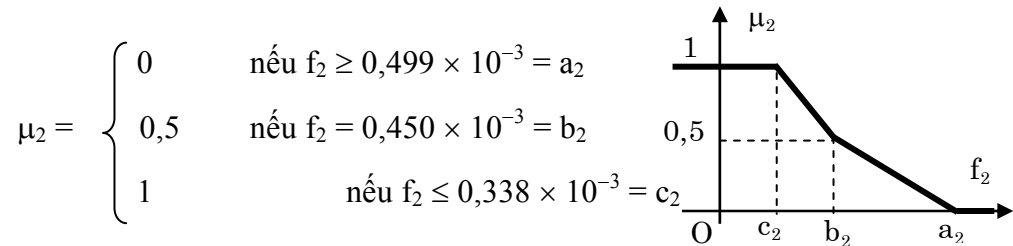
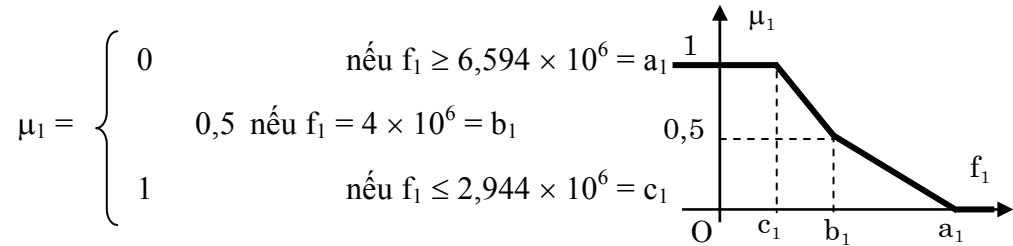
+ $\omega_2 f_2(X)$, trong đó ω_1, ω_2 là các trọng số không âm ứng với các hàm $f_1(X)$ và $f_2(X)$, $\omega_1 + \omega_2 = 1$. Do giá trị của hàm $f_1(X)$ thường lớn gấp rất nhiều lần giá trị của hàm $f_2(X)$, ω_1 và ω_2 được quy chuẩn như sau: $f(X) = \omega_1' f_1(X) + \omega_2' f_2(X)$, với $\omega_1' = \omega_1 \cdot 10^{-6} / 2,961$; $\omega_2' = \omega_2 \cdot 10^{+3} / 0,338$.

Ở yếu tố cấu thành thứ hai, trong các bước lặp trung gian, người ra quyết định thay đổi lần lượt các cặp trọng số (ω_1, ω_2) với các giá trị là $(0,2; 0,8)$, $(0,8; 0,2)$, $(0,6; 0,4)$ và $(0,4; 0,6)$. Cặp trọng số cuối cùng cho phương án tối ưu Pareto thoả mãn nhất là $x_1 = 237,1$ và $x_2 = 68,2$, với $f_1(X) = 3,529 \times 10^6$; $f_2(X) = 0,437 \times 10^{-3}$.

Còn ở yếu tố cấu thành thứ ba, tác giả Andrejz Osyczka đã sử dụng thuật toán tối ưu dò tìm ngẫu nhiên.

Phương pháp tương tác dựa trên mức ưu tiên PRELIME

Trước hết, ở yếu tố cấu thành thứ nhất, hai mục tiêu $f_1(X)$ và $f_2(X)$ được chuyển thành hai hàm (liên) thuộc mờ phản ánh độ thoả mãn của người ra quyết định đối với từng mục tiêu. Các hàm thuộc mờ này là các hàm tuyến tính từng khúc, được viết dưới dạng gián lược như sau cho một số điểm nội suy:



Đồ thị của các hàm thuộc mờ cho ở các hình vẽ trên. Phân tích hàm thuộc mờ μ_1 , ta thấy: người ra quyết định sẽ có độ thoả mãn 0 đối với mọi phương án làm cho $f_1 \geq 6,594 \times 10^6$; độ thoả mãn 1 nếu $f_1 \leq 2,944 \times 10^6$; và độ thoả mãn 0,5 nếu $f_1 = 4 \times 10^6$. Độ thoả mãn 0,5 được coi là độ thoả mãn tối thiểu và mức $f_1 = 4 \times 10^6 = b_1$ được gọi là mức ưu tiên tương ứng đối với mục tiêu f_1 . Tương tự chúng ta có thể phân tích về hàm thuộc μ_2 và mức ưu tiên b_2 . Sau đó, hàm thoả dụng tổ hợp dạng Max-Min được thiết lập cho hai hàm mục tiêu riêng rẽ trên dưới dạng: $\text{Max}\{\text{Min}[\mu_1, \mu_2]\}$ nhằm tìm ra phương án thoả

dụng (x_1, x_2) trong miền ràng buộc của bài toán.

Đối với yếu tố cấu thành thứ hai, người ra quyết định sẽ căn cứ vào thông tin do máy tính đưa ra để điều chỉnh các mức ưu tiên b_1 và b_2 . Thay đổi các cặp mức ưu tiên (b_1, b_2) từ $(4 \times 10^6; 0,45 \times 10^{-3})$ sang $(3,6 \times 10^6; 0,435 \times 10^{-3})$, sẽ nhận được phương án sau $(x_1, x_2) = (235,67; 67,67)$ với $(f_1, f_2) = (3,58 \times 10^6; 0,433 \times 10^{-3})$.

Trong yếu tố cấu thành thứ ba, các tác giả đã dùng thuật toán tìm kiếm ngẫu nhiên có điều khiển RST2ANU kết hợp với thuật toán mô phỏng tôỉ (SA) để tìm ra các phương án tối ưu Pareto cho các bài toán trung gian thông qua việc giải bài toán tối ưu phi tuyến đơn mục tiêu dạng $\text{Max}\{\text{Min}[\mu_1, \mu_2]\}$.

Chương II

CÁC MÔ HÌNH MẠNG

1. Mô hình mạng vận tải

1.1. Phát biểu bài toán vận tải

Bài toán vận tải được áp dụng rất rộng rãi trong lĩnh vực lập kế hoạch phân bổ sản phẩm hàng hoá (dịch vụ) từ một số địa điểm cung / cấp phát tới một số địa điểm cầu / tiêu thụ. Thông thường, tại mỗi địa điểm cung (nơi đi) chỉ có một số lượng giới hạn hàng, mỗi địa điểm cầu (nơi đến) chỉ cần một số lượng nhất định hàng. Với các cung đường vận chuyển hàng đa dạng, với cước phí vận tải khác nhau, mục tiêu đặt ra là xác định phương án vận tải tối ưu. Nói cách khác, vấn đề đặt ra là cần xác định nên vận chuyển từ mỗi địa điểm cung tới mỗi địa điểm cầu bao nhiêu đơn vị hàng nhằm thoả mãn nhu cầu của từng nơi đến đồng thời đạt tổng chi phí vận tải là nhỏ nhất.

Ví dụ: Ta có 3 điểm cung cấp hàng A, B, C và 4 điểm cầu S, T, U và V với lượng hàng cung và cầu tại mỗi điểm cũng như cước phí vận tải trên một đơn vị hàng cho mỗi cung đường như trong bảng II.1.

Bảng II.1. Các dữ liệu của bài toán vận tải

Điểm cung	Lượng hàng	Điểm cầu	Lượng hàng
A	5000	S	6000
B	6000	T	4000
C	2500	U	2000
Tổng	13500	V	1500
		Tổng	13500

Nơi đi	Cước phí vận chuyển/đơn vị hàng c_{ij} (USD) đến			
	S	T	U	V
A	3	2	7	6
B	7	5	2	3
C	2	5	4	5

Từ điểm cung i đến điểm cầu j ta có cước phí vận tải / một đơn vị hàng là c_{ij} đã biết, chẳng hạn như c_{11} là 3USD / một đơn vị hàng. Cần thiết lập phương án vận tải hàng đáp ứng được cung cầu và tổng chi phí vận tải là nhỏ nhất. Chú ý rằng bài toán vận tải đang xét có tổng cung bằng tổng cầu, nên được gọi là bài toán vận tải cân bằng thu phát. Đây là dạng đơn giản nhất trong các dạng bài toán vận tải.

1.2. Tạo phương án vận tải xuất phát

Khái niệm bảng vận tải

Bảng vận tải có m hàng, n cột gồm $m \times n$ ô, m là số điểm cung, n là số điểm cầu với cước phí c_{ij} được ghi trong ô (i, j) cho cung đường (i, j) . Khi $m = 3$, $n = 4$ như trong ví dụ trên, ta có bảng vận tải II.2.

Bảng II.2. Bảng vận tải

3	2	7	6	Cung 1: 5000
7	5	2	3	Cung 2: 6000
2	5	4	5	Cung 3: 2500
Cầu 1: 6000	Cầu 2: 4000	Cầu 3: 2000	Cầu 4: 1500	Tổng: 13500

Ta cần tìm phương án phân hàng vào các ô (i, j) sao cho tổng theo hàng hay cột đều khớp với các lượng cung, cầu và tổng chi phí vận tải là nhỏ nhất. Mỗi ô (i, j) biểu diễn một cung đường vận chuyển hàng từ điểm cung i về điểm cầu j .

Các phương pháp tạo phương án xuất phát

Có một số phương pháp tạo phương án xuất phát. Ta nghiên cứu hai phương pháp sau đây.

a. Phương pháp "góc tây bắc"

Phương pháp này được phát biểu như sau:

- Phân phát hàng tối đa vào góc tây bắc của bảng vận tải.
- Sau khi (hàng) cung hoặc (cột) cầu đã thỏa mãn thì ta thu gọn bảng vận tải bằng cách bỏ bớt hàng cung hoặc cột cầu đó đi (chỉ bỏ một trong hai thứ hoặc hàng hoặc cột, ở đây là toán tử hoặc loại trừ, *OR exclusive*).
- Tiếp tục lặp lại hai bước trên đây cho tới khi hàng được phân phối hết vào các ô (các ô được phân hàng được gọi là ô sử dụng).

Bằng phương pháp "góc tây bắc" ta tạo được phương án A trong bảng II.3 với sáu ô sử dụng $(1, 1)$, $(2, 1)$, $(2, 2)$, $(2, 3)$, $(3, 3)$ và $(3, 4)$.

Bảng II.3. Phương án xuất phát với phương pháp "góc tây bắc"

3		2		7		6	
	5000						
7		5		2		3	
	1000		4000		1000		
2		5		4		5	
					1000		1500

Tổng chi phí vận tải:

$$\Sigma CPVT = (3 \times 5 + 7 \times 1 + 5 \times 4 + 2 \times 1 + 4 \times 1 + 5 \times 1,5) \times 1000 = 55500.$$

b. Phương pháp cước phí tối thiểu

Phương pháp này được phát biểu tương tự phương pháp "góc tây bắc" nhưng ưu tiên phân phát hàng vào ô có cước phí bé nhất (nếu có nhiều ô như vậy thì chọn ô bất kì trong số đó). Lúc này ta có phương án xuất phát là phương án B cho trong bảng II.4.

Bảng II.4. Phương án xuất phát với phương pháp cước phí tối thiểu

3		2		7	6
	1000		4000		
7	2500	5		2	3
				2000	1500
2	2500	5		4	5

Tổng chi phí vận tải:

$$\Sigma CPVT = (3 \times 1 + 2 \times 4 + 7 \times 2,5 + 2 \times 2 + 3 \times 1,5 + 2 \times 2,5) \times 1000 = 42000$$

Nhận xét

– Phương pháp cước phí tối thiểu thường cho phương án xuất phát tốt hơn phương pháp “góc tây bắc”.

– Bảng vận tải có số ô sử dụng là $3 + 4 - 1 = 7 - 1 = 6$. Một cách tổng quát bảng vận tải m hàng, n cột có số ô sử dụng là $m + n - 1$.

– Bài toán vận tải cũng là BTQHHT. Trong ví dụ đang xét, nếu kí hiệu x_{ij} là lượng hàng vận chuyển trên cung đường (i, j) thì chúng ta BTQHHT sau:

$$z = c_{11}x_{11} + c_{12}x_{12} + \dots + c_{34}x_{34} \rightarrow \text{Min}$$

với các ràng buộc:

$$\begin{cases} x_{11} + x_{12} + x_{13} + x_{14} = 5000 \\ x_{21} + x_{22} + x_{23} + x_{24} = 6000 \\ x_{31} + x_{32} + x_{33} + x_{34} = 2500 \\ x_{11} + x_{21} + x_{31} = 6000 \\ x_{12} + x_{22} + x_{32} = 4000 \\ x_{13} + x_{23} + x_{33} = 2000 \\ x_{14} + x_{24} + x_{34} = 1500 \\ x_{ij} \geq 0 \quad \forall i = 1, 2, 3; j = 1, 2, 3, 4 \end{cases}$$

Hệ các ràng buộc có 12 biến với 7 phương trình. Nếu lấy tổng 3 phương trình đầu trừ đi tổng 3 phương trình tiếp theo thì được phương trình cuối. Có thể kiểm nghiệm dễ dàng, số phương trình độc lập tuyến tính của hệ là $7 - 1 = 6$.

– Mỗi phương án xuất phát A hay B tìm được của bài toán vận tải chính là một phương án cực biên xuất phát khi giải BTQHHT. Bài toán vận tải có thể hoàn toàn giải được bằng phương pháp đơn hình. Tuy nhiên do cấu trúc đặc biệt của mình, bài toán vận tải có thể giải bằng phương pháp đặc biệt với thuật toán chuyên dụng.

1.3. Phương pháp phân phối giải bài toán vận tải

Chúng ta áp dụng phương pháp “đá lăn” (tạm dịch từ *Stepping Stone Method*), hay chính thức hơn còn gọi là phương pháp phân phối (*Distribution Method*) để giải bài toán vận tải.

Phương pháp “đá lăn” là một quy trình tính toán nhằm từng bước cải thiện phương án vận tải đã có để cuối cùng tìm được phương án vận tải tối ưu.

Xác định hiệu suất của các ô chưa sử dụng

Quay lại bảng vận tải II.3 với phương án xuất phát tìm được theo phương pháp “góc tây bắc”. Trong bảng đó chỉ có một số ô đã sử dụng, ta coi chúng như các hòn đá nổi lên trong một cái ao. Xét một ô (i, j) bất kì chưa sử dụng trong phương án đã có. Ta cần tính hiệu suất của ô đó, kí hiệu là e_{ij} , (e là viết tắt của từ *effect*) theo các bước sau:

– Đầu tiên ta cần tìm một đường đi có tính chất: đi qua một ô (i, j) chưa sử dụng (ô xuất phát) và một số ô đã sử dụng khác, mỗi bước phải đi theo hàng hoặc theo cột xen kẽ nhau (không được đi liền hai bước trên một hàng hay một cột) để cuối cùng quay về ô (i, j). Điều này giống như đang ở trên thuyền, muốn ra khỏi thuyền mà không ướt ta phải nhảy qua các hòn đá nổi lên trong ao để cuối cùng lại quay về thuyền (vì vậy phương pháp có tên là phương pháp “đá lăn”). Một điều thú vị nữa là con đường nhảy trên các hòn đá như vậy là duy nhất.

Tóm lại xuất phát từ ô (1, 2) chẳng hạn, ta sẽ có đường đi như sau: (1, 2) → (2, 2) → (2, 1) → (1, 1) → (1, 2), trên đường đi này chỉ duy nhất có một ô chưa sử dụng (xem bảng II.5).

Bảng II.5. Tính hiệu suất các ô chưa sử dụng

3	2	7	6	5000
5000				
7	5	2	3	6000
1000	4000	1000		
2	5	4	5	2500
(-7)	(-2)	1000	1500	
6000	4000	2000	1500	

– Đánh dấu cộng trừ xen kẽ tại các đỉnh trên đường đi mà trong đó ô chưa sử dụng được đánh dấu “+”. Giả sử ta cần luân chuyển một đơn vị hàng theo đường đi đã xác định mà vẫn thoả mãn được cung cầu (tức là các ô mang dấu “+”: ô (1, 2) và ô (2, 1) có thêm một đơn vị hàng, các ô mang dấu “-”: ô (2, 2) và ô (1, 1) rút bớt đi một đơn vị hàng). Lúc này tổng chi phí sẽ thay đổi một lượng tiền là: $e_{12} = +c_{12} - c_{22} + c_{21} - c_{11} = 2 - 5 + 7 - 3 = +1$. Nói cách khác, tổng chi phí vận tải sẽ tăng thêm lên 1USD cho mỗi một đơn vị hàng luân chuyển theo đường đi trên. Ta gọi e_{12} là *hiệu suất* của ô(1, 2).

Tương tự:

$$e_{13} = 7 - 2 + 7 - 3 = +9,$$

$$e_{14} = 6 - 5 + 4 - 2 + 7 - 3 = +7,$$

$$e_{24} = 3 - 5 + 4 - 2 = 0,$$

$$e_{31} = 2 - 7 + 2 - 4 = -7,$$

$$e_{32} = 5 - 5 + 2 - 4 = -2.$$

Chỉ có hai ô với hiệu suất âm là ô (3, 1) và ô (3, 2) (xem bảng II.5) có thể lựa chọn để đưa vào sử dụng trong phương án mới. Ta quyết định trong phương án mới sẽ chọn ô (3, 2) để đưa vào sử dụng, mỗi đơn vị hàng đưa vào sử dụng tại ô (3, 2) sẽ làm tổng chi phí giảm 2USD. Kí hiệu $e = e_{32}$.

Chú ý: Có thể chứng minh được $e_{ij} = \Delta_{ij}$ với Δ_{ij} là giá trị trên hàng Δ ứng với cột x_{ij} nếu giải bài toán vận tải bằng phương pháp đơn hình.

Xác định lượng hàng đưa vào ô chọn

Như trên đã phân tích, một đơn vị hàng đưa vào ô (3, 2) làm giảm tổng chi phí vận tải 2 USD. Ta cần tìm q, lượng hàng tối đa có thể đưa vào ô (3, 2). Đường đi qua ô (3, 2) và một số ô đã được sử dụng là: (3, 2) → (2, 2) → (2, 3) → (3, 3) → (3, 2), với các ô được đánh dấu cộng trừ xen kẽ (ô (3, 2) mang dấu +). Lượng hàng q được tính theo quy tắc:

$q =$ giá trị nhỏ nhất của các lượng hàng tại các ô mang dấu (-) = Min {lượng hàng tại ô (2, 2), lượng hàng tại ô (3, 3)} = Min {4000, 1000} = 1000.

Vậy trong phương án mới, lượng hàng tại các ô mang dấu “+” (các ô (3, 2), ô (2, 3)) được tăng thêm 1000 đơn vị, còn tại các ô mang dấu “-” (các ô (2, 2) và ô (3, 3)) lượng hàng giảm đi 1000 đơn vị (xem bảng II.6). Phương án mới gồm 6 ô sử dụng (ô (3, 3) ứng với q=1000 đã bị loại ra).

Bảng II.6. Phương án vận tải sau hai bước

3	2	7	6	5000
5000				
7	5	2	3	6000
1000	3000	2000		
2	5	4	5	2500
(-5)	1000		1500	
6000	4000	2000	1500	

Tổng chi phí vận tải:

$$\sum CPVT = (3 \times 5 + 7 \times 1 + 5 \times 3 + 2 \times 2 + 5 \times 1 + 5 \times 1,5) \times 1000 = 3500;$$

hoặc

$$\sum CPVT_{\text{mới}} = \sum CPVT_{\text{cũ}} - e \times q = 55500 - 2 \times 1000 = 53500.$$

Điều kiện tối ưu

Thực hiện theo quy trình trên cho tới khi tất cả các hiệu suất $e_{ij} \geq 0 \forall (i, j)$ là các ô chưa sử dụng. Đây chính là điều kiện tối ưu hay điều kiện dừng. Điều kiện này thực chất là điều kiện $\Delta_{ij} \geq 0$ với mọi biến ngoài cơ sở x_{ij} nếu giải bài toán bằng phương pháp đơn hình.

Để giải tiếp bài toán, cần tính các hiệu suất cho các ô chưa sử dụng trong phương án mới:

$$\begin{aligned} e_{12} &= 2 - 5 + 7 - 3 = +1; & e_{13} &= 7 - 2 + 7 - 3 = +19; \\ e_{14} &= 6 - 5 + 5 - 5 + 7 - 3 = +5; & e_{24} &= 3 - 5 + 5 - 5 = -2; \\ e_{31} &= 2 - 7 + 5 - 5 = -5; & e_{33} &= 4 - 5 + 5 - 2 = +2. \end{aligned}$$

Ta quyết định sử dụng ô chọn (3, 1) trong phương án mới vì $e_{31} = -5$. Tìm được $q = 1000$ theo quy tắc đã biết. Có hai ô ứng với q tìm được, chúng ta chỉ bỏ đi ô (2, 1) còn phải giữ lại ô (3, 2) để đưa vào sử dụng. Phương án sau bước thứ ba cho trong bảng II.7.

Bảng II.7. Phương án vận tải sau ba bước

3	2	7	6	5000
5000				
7	5	2	3	6000
	4000	2000	(-2)	
2	5	4	5	2500
1000	0		1500	
6000	4000	2000	1500	

Tổng chi phí vận tải:

$$\Sigma CPVT = 53500 - 5 \times 1000 = 48500.$$

Tiếp tục tính các hiệu suất:

$$e_{12} = +1; \quad e_{13} = 7 - 2 + 5 - 5 + 4 = 9;$$

$$e_{14} = 6 - 5 + 2 - 3 = 0; \quad e_{21} = 7 - 2 + 5 - 5;$$

$$e_{24} = 3 + 5 + 5 - 5 = -2; \quad e_{33} = 4 - 5 + 5 - 2 = 2.$$

Chọn ô (2, 4) đưa vào sử dụng và tính $q = 1500$. Từ đó có phương án mới sau bốn bước như trong bảng II.8.

Bảng II.8. Phương án vận tải sau bốn bước

3	2 (-4)	7	6	5000
5000				
7	5	2	3	6000
	2500	2000	1500	
2	5	4	5	2500
1000	1500			
6000	4000	2000	1500	

Tổng chi phí vận tải:

$$\Sigma CPVT = 48500 - 2 \times 1500 = 45500.$$

Tiếp tục tính các hiệu suất:

$$e_{12} = 2 - 5 + 2 - 3 = -4; \quad e_{13} = 7 - 2 + 5 - 5 + 2 - 3 = 4;$$

$$e_{14} = 6 - 3 + 5 - 5 + 2 - 3 = 2; \quad e_{21} = 7 - 2 + 5 - 5 = 5;$$

$$e_{33} = 4 - 5 + 5 - 2 = 2; \quad e_{34} = 5 - 5 + 5 - 2 = 3;$$

Ta có $e_{12} = -4$ và chọn ô (1, 2) làm ô chọn với $q = 1500$ và chuyển sang phương án mới như trong bảng II.9.

Bảng II.9. Phương án vận tải sau năm bước

3	2	7	6	5000
3500	1500			
7	5	2	3	6000
	2500	2000	1500	
2	5	4	5	2500
2500				
6000	4000	2000	1500	

Tổng chi phí vận tải:

$$\Sigma CPVT = 45500 - 4 \times 1500 = 39500.$$

Lúc này $e_{ij} \geq 0$ với mọi ô (i, j) chưa sử dụng. Điều kiện tối ưu đã được thoả mãn. Phương án vận tải tối ưu cho trong bảng II.9 với tổng chi phí nhỏ nhất là 39500.

Bài toán vận tải không cân bằng thu phát

Trường hợp tổng lượng cung lớn hơn tổng lượng cầu, cần bố trí thêm một điểm cầu giả mà mọi chi phí vận tải đến đó đều được coi bằng 0.

Tương tự, nếu cầu vượt cung thì cần bố trí một điểm cung giả và coi mọi chi phí vận chuyển từ đó đi đều bằng 0.

1.4. Phương pháp phân phối cải biên giải bài toán vận tải

Phương pháp “đá lăn” hay phương pháp phân phối có một nhược điểm là việc tính hiệu suất của các ô khá dài dòng. Vì vậy, ta sẽ nghiên cứu phương pháp phân phối cải biên nhằm tính các hiệu suất e_{ij} ngắn gọn hơn.

Xét phương án xuất phát tìm được bằng phương pháp cực phí cực tiểu cho trong bảng II.10 (với tổng chi phí vận tải là 42000).

Bảng II.10. Phương án vận tải xuất phát

3	2	7	6	5000
1000	4000			
7	5	2	3	6000
2500		2000	1500	
2	5	4	5	2500
2500				
6000	4000	2000	1500	

Ta có $e_{13} = 7 - 2 + 7 - 3 = +9$. Ta tìm cách tính e_{13} bằng cách khác nhanh hơn như trình bày sau đây.

Trước hết cần xây dựng hệ thống số thế vị hàng và cột $\{(u_i, v_j), i = 1, 2, 3; j = 1, 2, 3, 4\}$. Có thể gán cho một thế vị bất kì giá trị 0 (hoặc một giá trị bất kì khác), thế vị này thường được chọn ở hàng hay cột có nhiều ô sử dụng nhất. Chẳng hạn chọn $u_2 = 0$.

Các thế vị khác được tính bởi công thức: $u_i + v_{ij} = c_{ij} \forall \text{ ô } (i, j) \text{ sử dụng.}$

$$u_2 = 0 \Rightarrow v_1 = 7 (= c_{21} - u_2)$$

$$v_3 = 2 (= c_{23} - u_2)$$

$$v_4 = 3 (= c_{24} - u_2)$$

$$u_1 = -4 (= c_{11} - v_1)$$

$$u_3 = -5 (= c_{37} - v_1)$$

$$v_2 = 6 (= c_{12} - u_1)$$

Công thức tổng quát để tính các hiệu suất cho các ô (i, j) chưa sử dụng là:

$$e_{ij} = c_{ij} - (u_i + v_j).$$

Chẳng hạn ta có $e_{13} = c_{13} - (u_1 + v_3) = 7 - (-4 + 2) = 9$. Các hiệu suất khác được tính tương tự (xem bảng II.11).

Bảng II.11. Tính toán các thế vị và các hiệu suất

	$v_1 = 7$	$v_2 = 6$	$v_3 = 2$	$v_4 = 3$	
$u_1 = -4$	3 1000	2 4000	7	6	5000
$u_2 = 0$	7 2500	5 (-1)	2 2000	3 1500	6000
$u_3 = -5$	2 2500	5	4	5	2500
	6000	4000	2000	1500	

Trong bảng II.11 ta thấy $e_{22} = -1 < 0$. Chọn ô (2, 2) để đưa vào sử dụng ứng với $q = 2500$, ta chuyển sang phương án mới và tính lại các hệ thống số thế vị như trong bảng II.12.

Bảng II.12. Tính toán các thế vị và các hiệu suất cho phương án mới

	$v_1 = 6$	$v_2 = 6$	$v_3 = 2$	$v_4 = 3$	
$u_1 = -3$	3 3500	2 1500	7	6	5000
$u_2 = 0$	7	5 2500	2 2000	3 1500	6000
$u_3 = -4$	2 2500	5	4	5	2500
	6000	4000	2000	1500	

Chọn $u_2 = 0 \Rightarrow v_2 = 5 (= 5 - 0); v_3 = 2 (= 2 - 0); v_4 = 3 (= 3 - 0);$
 $u_1 = -3 (= 2 - 5); v_1 = 6 (= 3 - (-3)); u_3 = -4 (= 2 - 6).$

Tổng chi phí vận tải:

$$\begin{aligned} \Sigma \text{CPVT} &= (3 \times 3,5 + 2 \times 1,5 + 5 \times 2,5 + 2 \times 2 + 3 \times 1,5 + 2 \times 2,5) \times 1000 \\ &= 39500 \quad (\text{tính cách khác, } \Sigma \text{CPVT}_{\text{mới}} = 42000 - 1 \times 2500). \end{aligned}$$

Tiếp tục tính toán các hiệu suất:

$$e_{13} = c_{13} - (u_1 + v_3) = 7 - (-3 + 2) = 8;$$

$$e_{14} = c_{14} - (u_1 + v_4) = 6 - (-3 + 3) = 6;$$

$$e_{21} = c_{21} - (u_2 + v_1) = 7 - (0 + 6) = 1;$$

$$e_{32} = c_{32} - (u_3 + v_2) = 5 - (-4 + 5) = 4;$$

$$e_{33} = c_{33} - (u_3 + v_4) = 4 - (-4 + 2) = 6;$$

$$e_{34} = c_{34} - (u_3 + v_4) = 5 - (-4 + 3) = 6.$$

Ta thấy $e_{ij} \geq 0 \forall \hat{o}(i, j)$ chưa sử dụng nên điều kiện tối ưu đã được thoả mãn. Phương án tối ưu cho trong bảng II.12, với tổng chi phí vận tải nhỏ nhất là 39500.

Chú ý:

– Đối với bài toán vận tải cần cực đại hoá hàm mục tiêu thì tiêu chuẩn dừng sẽ là $e_{ij} \leq 0 \forall \hat{o}(i, j)$ chưa sử dụng.

– Đối với bài toán vận tải có ô cấm (cung đường không được sử dụng) thì đặt cước phí $M = +\infty$ cho các ô cấm với bài toán Min hoặc $M = -\infty$ với bài toán Max.

Giải bài toán vận tải bằng phần mềm Lingo

Để giải bài toán vận tải trong Lingo, ta có thể sử dụng các bài toán mẫu bằng cách nhấn vào biểu tượng Lingo và thực hiện các lệnh *File > Open > Tran.lng* để vào bài toán vận tải mẫu. Sau đó nhập các số liệu đầu vào của bài toán cần giải, chẳng hạn, của ví dụ đã xét trong các mục trên thay cho các số liệu của bài toán mẫu (xem hình II.1).

```

LINGO - [D:\THAYTH-1\ATHANH\LINGO\LINGO\LINGO\SAMPLES\TRAN.LNG]
File Edit LINGO Window Help
MODEL:
! A 3 Warehouse, 4 Customer Transportation Problem;
SETS:
    WAREHOUSE / WH1, WH2, WH3/      : CAPACITY;
    CUSTOMER   / C1, C2, C3, C4/     : DEMAND;
    ROUTES( WAREHOUSE, CUSTOMER)    : COST, VOLUME;
ENDSETS

! Here are the parameters;
DATA:
    CAPACITY = 5000, 6000, 2500 ;
    DEMAND = 6000, 4000, 2000, 1500 ;
    COST = 3, 2, 7, 6,
           7, 5, 2, 3,
           2, 5, 4, 5 ;
ENDDATA
END

```

Hình II.1. Nhập số liệu cho bài toán vận tải

Sau đó chúng ta thực hiện *LINGO>Solve*, kết quả tính toán sẽ hiện ra trên màn hình (xem hình II.2).

```

LINGO
File Edit LINGO Window Help
Reports Window
Rows= 8 Uars= 12 No. integer vars= 0 ( all are linear)
Nonzeros= 43 Constraint nonz= 24( 24 are +- 1) Density=0.413
Smallest and largest elements in absolute value= 1.00000 6000.00
No. < : 3 No. =: 0 No. > : 4, Obj=MIN, GUBs <= 4
Single cols= 0

Optimal solution found at step: 8
Objective value: 39500.00

Variable Value Reduced Cost
VOLUME( WH1, C1) 3500.000 0.000000E+00
VOLUME( WH1, C2) 1500.000 0.000000E+00
VOLUME( WH1, C3) 0.000000E+00 8.000000
VOLUME( WH1, C4) 0.000000E+00 6.000000
VOLUME( WH2, C1) 0.000000E+00 1.000000
VOLUME( WH2, C2) 2500.000 0.000000E+00
VOLUME( WH2, C3) 2000.000 0.000000E+00
VOLUME( WH2, C4) 1500.000 0.000000E+00
VOLUME( WH3, C1) 2500.000 0.000000E+00
VOLUME( WH3, C2) 0.000000E+00 4.000000
VOLUME( WH3, C3) 0.000000E+00 6.000000
VOLUME( WH3, C4) 0.000000E+00 6.000000

```

Hình II.2. Kết quả của bài toán vận tải

2. Mô hình mạng PERT

(*Program Evaluation and Review Technique*)

2.1. Các khái niệm cơ bản về PERT

Vai trò của PERT

PERT có thể được hiểu là phương pháp hoặc kỹ thuật theo dõi và đánh giá dự án với mục đích giúp cho bộ máy quản lý trả lời các câu hỏi sau đây:

- Dự án sẽ hoàn thành khi nào?
 - Mỗi hoạt động của dự án nên được bắt đầu vào thời điểm nào và kết thúc vào thời điểm nào?
 - Những hoạt động nào của dự án phải kết thúc đúng thời hạn để tránh cho toàn bộ dự án bị kết thúc chậm hơn so với kế hoạch?
 - Liệu có thể chuyển các nguồn dự trữ (nhân lực, vật lực) từ các hoạt động “không căng” sang các hoạt động “căng” (các hoạt động phải hoàn thành đúng tiến độ) mà không ảnh hưởng tới thời hạn hoàn thành dự án?
 - Những hoạt động nào cần tập trung theo dõi?
- Để bước đầu hình dung về PERT, chúng ta xét ví dụ sau đây.

Ví dụ:

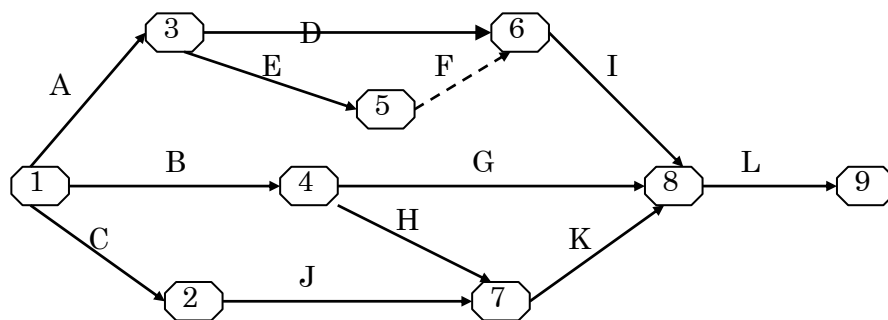
Giả sử cần thực hiện một dự án hoặc chương trình có các hoạt động được liệt kê trong bảng II.13.

Bảng II.13. Các hoạt động của một dự án, thứ tự và thời gian thực hiện

Hoạt động	Hoạt động kế trước	Thời gian thực hiện (tuần)
A	–	2
B	–	2
C	–	2
D	A	3
E	A	4
F	E	0 (hoạt động giả)
G	B	7
H	B	6
I	D, F	4
J	C	10
K	H, J	3
L	G, I, K	4

Ta cần lập kế hoạch thực hiện dự án trên để hoàn thành toàn bộ các hoạt động của dự án trong thời gian ngắn nhất, đồng thời phải xác định được những hoạt động nào cần chú trọng (được hiểu là các hoạt động “căng”).

Vẽ sơ đồ mạng PERT



Hình II.3. Sơ đồ mạng PERT

Trên hình II.3 ta thấy mạng PERT là một mạng các nút có đánh số được nối với nhau bởi các cung có mũi tên. Mỗi cung có mũi tên biểu diễn một hoạt động của dự án, còn mỗi nút biểu diễn thời điểm kết thúc một số hoạt động và / hoặc thời điểm bắt đầu của một số hoạt động khác.

Hoạt động giả F được kí hiệu bởi cung mũi tên với nét rời có thời gian thực hiện bằng 0, nhằm tránh cho hoạt động D và E có cùng nút bắt đầu và nút kết thúc. Như vậy, trong sơ đồ mạng PERT ta buộc phải tuân theo quy ước: hai hoạt động khác nhau thì không được có cùng nút bắt đầu cũng như nút kết thúc.

Xác định thời gian tối thiểu thực hiện dự án

Để xác định thời gian tối thiểu thực hiện dự án, trước hết chúng ta nghiên cứu khái niệm thời điểm bắt đầu sớm nhất và thời điểm kết thúc sớm nhất (EST và EFT – *Earliest start time* và *Earliest finish time*) cho từng hoạt động.

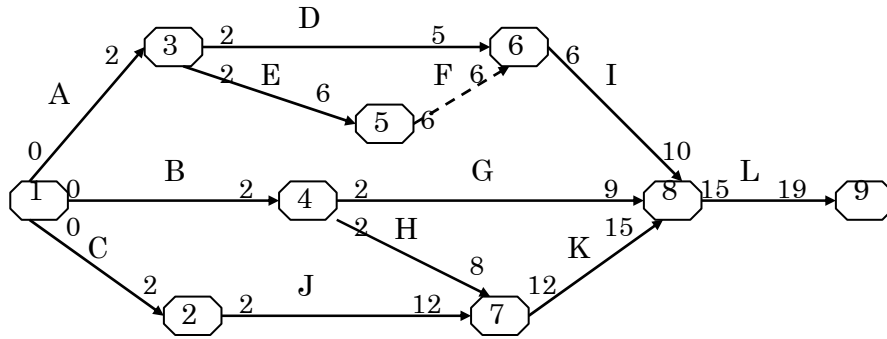
Ví dụ: Hoạt động A có $EST_A = 0$ và $EFT_A = 2$, vì

- Thời điểm bắt đầu sớm nhất là khi bắt đầu khởi động dự án,
- Thời điểm kết thúc sớm nhất là sau 2 tuần.

Mối quan hệ giữa EST và FFT là:

$$EFT = EST + \text{thời gian thực hiện hoạt động.}$$

Một cách tổng quát, để xác định EST chúng ta có quy tắc “*thời điểm bắt đầu sớm nhất*”: thời điểm bắt đầu sớm nhất của một hoạt động rời một nút nào đó là thời điểm muộn nhất trong các thời điểm kết thúc sớm nhất đối với các hoạt động đi vào nút đó. Áp dụng quy tắc trên đây, có thể tính được $EST_K = 12$ (do $EFT_H = 8$, $EFT_J = 12$ và số lớn hơn là 12) và $EFT_K = 15$. Kết quả tìm EST và EFT cho các hoạt động dự án được tính toán tiến từ nút 1 đến nút 9 và được tóm tắt trong bảng II.14 và hình II.4. Vậy thời gian kết thúc sớm nhất dự án là sau 19 tuần.



Hình II.4. Tính EST và EFT cho các hoạt động của dự án

Bảng II.14. Tính EST, LST, EFT, LFT và tìm đường găng

Hoạt động	EST	LST	EFT	LFT	LST-EST (LFT-EFT)	Trên cung găng
A	0	5	2	7	5	
B	0	4	2	6	4	
C	0	0	2	2	0	*
D	2	8	5	11	6	
E	2	7	6	11	5	
F	6	11	6	11	5	
G	2	8	9	15	6	
H	2	6	8	12	4	
I	6	11	10	15	5	
J	2	2	12	12	0	*
K	12	12	15	15	0	*
L	15	15	19	19	0	*

Bước tiếp theo là xác định thời điểm bắt đầu muộn nhất và thời điểm kết thúc muộn nhất (LST và LFT – *Latest start time* và *Latest finish time*) cho từng hoạt động.

Ví dụ: Hoạt động L có $LST_L = 15$ và $LFT_L = 19$, vì

– Thời điểm kết thúc muộn nhất là sau 19 tuần (nếu ta ấn định dự án phải kết thúc sau 19 tuần),

– Thời điểm bắt đầu muộn nhất là tuần 15 (do hoạt động L cần thời gian 4 tuần để thực hiện).

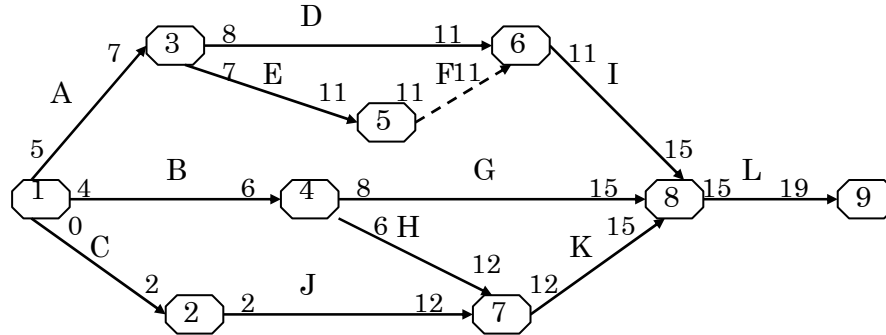
Mối quan hệ giữa LST và LFT là:

$$LST = LFT - \text{thời gian thực hiện hoạt động.}$$

Một cách tổng quát, để xác định LFT chúng ta có quy tắc “*thời điểm kết thúc muộn nhất*”: thời điểm kết thúc muộn nhất của một hoạt động đi vào một nút nào đó là thời điểm sớm nhất trong các thời điểm bắt đầu muộn nhất đối với các hoạt động rời

nút đó.

Áp dụng quy tắc trên đây, có thể tính được $LFT_A = 7$ (do $LST_D = 8$, $LST_E = 7$ và số bé hơn là 7) và $LST_A = 5$. Kết quả tìm LFT và LST cho các hoạt động dự án được tính toán lùi từ nút 9 về nút 1 và được tóm tắt trong bảng II.14 và hình II.5.



Hình II.5. Tính LFT và LST cho các hoạt động của dự án

Chú ý: Mỗi cung có mũi tên là một hoạt động, nhưng có thể bao gồm nhiều hoạt động nhỏ khác. Nói cách khác, bản thân từng hoạt động của dự án có thể lại là một mạng PERT nhỏ.

Xác định hoạt động găng, đường găng

Hoạt động găng là hoạt động mà

$$LST - EST = LFT - EFT = 0, \text{ hay } [EST, EFT] \equiv [LST, LFT]$$

$$\Leftrightarrow \begin{cases} EST = LST \\ EFT = LFT \end{cases} \Leftrightarrow \begin{cases} Slack = LST - EST = 0 \\ Slack = LFT - EFT = 0 \end{cases} \text{ (độ trễ cho phép bằng 0).}$$

Giải thích: Slack \equiv độ nói lỏng (độ trễ).

Trong ví dụ đang xét, các hoạt động găng là: $C \rightarrow J \rightarrow K \rightarrow L$ (xem bảng II.14) và tạo thành đường găng (*Critical Path*). Vì vậy, phương pháp mạng PERT còn có tên là phương pháp đường găng (*CPM - Critical Path Method*).

Xác định đường găng bằng phần mềm Lingo

Để xác định đường găng bằng phần mềm Lingo, ta có thể sử dụng các bài toán mẫu bằng cách nhấn vào biểu tượng Lingo và thực hiện các lệnh *File > Open > Pert.lng* để vào bài toán PERT mẫu. Sau đó nhập các số liệu đầu vào của bài toán cần giải vào thay các số liệu của bài toán mẫu, chẳng hạn như số liệu của ví dụ đã cho (xem hình II.6).

```

LINGO - [D:\THAYTH-1\ATHANH\LINGO\LINGO\LINGO\SAMPLES\PERT1.LNG]
File Edit LINGO Window Help

! For each task there is: Time, Early Start (ES),
  Late Start (LS), and Slack;
TASK/ FIRST, A,B,C,D,E,F,G,H,I,J,K,L/: TIME, ES, LS, SLACK;

! Here are the precedence relations, the first task in the
  precedence relationship needs to be completed before the
  second task can be started;
PRED( TASK, TASK)/  FIRST,A,    FIRST,B, FIRST,C,
  A,D,    A,E,    E,F,    B,G,    B,H,    D,I,
  F,I,    C,J,    H,K,    J,K,    G,L,    I, L, K,L/;
ENDSETS

DATA: ! Here are the task times;
      TIME = 0, 2,2,2,3,4,0,7,6,4,10,3,4;
ENDDATA

```

Hình II.6. Nhập số liệu cho bài toán PERT

Sau đó chúng ta thực hiện *LINGO > Solve*, kết quả tính toán sẽ hiện trên màn hình (xem hình II.7).

```

LINGO
File Edit LINGO Window Help

Reports Window

SLACK( FIRST)      0.0000000E+00
SLACK( A)          5.000000
SLACK( B)          4.000000
SLACK( C)          0.0000000E+00
SLACK( D)          6.000000
SLACK( E)          5.000000
SLACK( F)          5.000000
SLACK( G)          6.000000
SLACK( H)          4.000000
SLACK( I)          5.000000
SLACK( J)          0.0000000E+00
SLACK( K)          0.0000000E+00
SLACK( L)          0.0000000E+00

```

Hình II.7. Kết quả tìm cung găng của bài toán PERT

2.2. Sơ đồ PERT với số liệu ngẫu nhiên

Thời gian thực hiện từng hoạt động của dự án nói chung là một lượng biến động khó dự đoán trước, chúng ta giả thiết chúng là các biến ngẫu nhiên. Giả sử ta có các số liệu ước tính về thời gian thực hiện các hoạt động của dự án (xem bảng II.15) a, m, b. Lúc đó thời gian trung bình và độ lệch chuẩn thời gian thực hiện các hoạt động được

$$t = \frac{a + 4m + b}{6}.$$

Bảng II.15. Số liệu ước tính về thời gian thực hiện các hoạt động

Hoạt động	Hoạt động kế trước	Thời gian ước tính				
		a (sớm nhất)	m (nhiều khả năng xảy ra nhất)	b (muộn nhất)	t (thời gian trung bình)	σ (độ lệch tiêu chuẩn, độ biến thiên)
A	–	1	2	3	2	1/3
B	–	1	2	3	2	1/3
C	–	1	2	3	2	1/3
D	A	1	2	9	3	4/3
E	A	2	3	10	4	4/3
F	E	0	0	0	0	0
G	B	3	6	15	7	2
H	B	2	5	14	6	2
I	D, F	1	4	7	4	1
J	C	4	9	20	10	8/3
K	H, J	1	2	9	3	4/3
L	G, I, K	4	4	4	4	0

Bước tiếp theo là lập sơ đồ mạng cho dự án với các thời gian trung bình t và tìm đường găng. Đường găng là $C \rightarrow J \rightarrow K \rightarrow L$ bao gồm các hoạt động găng C, J, K và L. Các hoạt động này có độ trễ cho phép bằng 0, hay nói cách khác, không cho phép sự chậm trễ nào. Đây là các hoạt động cần hết sức chú trọng, việc chậm thực hiện bất cứ một hoạt động nào trong số này đều kéo theo sự chậm trễ trong tiến độ của cả dự án. Từ *Critical Path* (tiếng Anh) được dịch sang tiếng Việt là đường găng vì lí do đó.

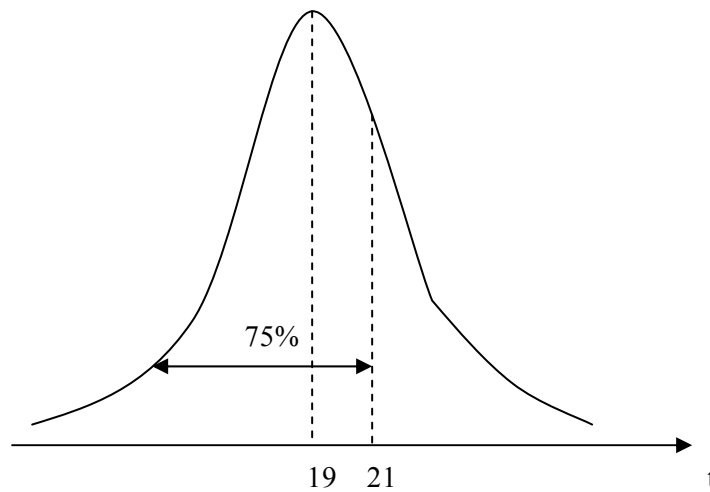
Thời gian thực hiện dự án là một lượng ngẫu nhiên tính theo công thức: $T = T_C + T_J + T_K + T_L$. Ta tìm kì vọng của T (thời gian trung bình thực hiện dự án) theo công thức: $m = m_T = t_C + t_J + t_K + t_L = 2 + 10 + 3 + 4 = 19$ (tuần).

Tính độ lệch chuẩn của thời gian thực hiện dự án:

$$\sigma = \sigma_T = \sqrt{\sigma_C^2 + \sigma_J^2 + \sigma_K^2 + \sigma_L^2} = \sqrt{(1/3)^2 + (8/3)^2 + (4/3)^2 + 0} = 3.$$

Ta coi T (thời gian thực hiện dự án) là biến ngẫu nhiên tuân theo luật chuẩn $N(m = 19; \sigma = 3)$.

Đồ thị hàm mật độ xác suất của T cho trên hình II.8.



Hình II.8. Đường cong mật độ chuẩn

Để tính P, xác suất thực hiện dự án trong vòng (không vượt quá) 19 tuần, ta phải quy T về biến ngẫu nhiên với phân phối chuẩn tắc $N(0, 1)$ như cho trong phụ lục 1. Lúc đó:

$$P(T \leq 19) = P\left(\frac{T - m}{\sigma} \leq \frac{19 - 19}{3}\right) = P(Z \leq 0) = 0,5 \text{ (hay 50\%)},$$

ở đây $Z = (T - m)/\sigma$ là biến ngẫu nhiên tuân theo phân phối $N(0, 1)$.

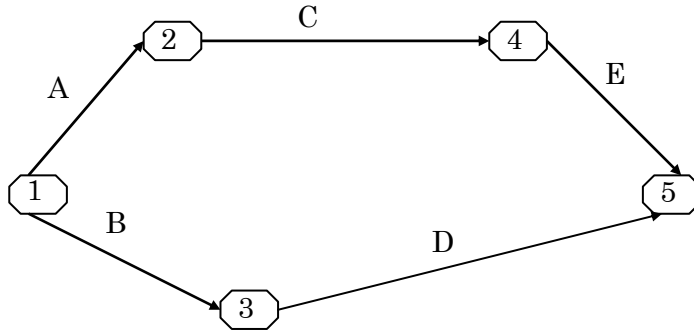
Tương tự, xác suất thực hiện dự án trong vòng (không vượt quá) 21 tuần được tính như sau:

$$P(T \leq 21) = P\left(\frac{T - m}{\sigma} \leq \frac{21 - 19}{3}\right) = P(Z \leq 0,666) = 75\%.$$

Ta chuyển sang xem xét vấn đề về độ tin cậy của thời gian hoàn thành dự án. Chẳng hạn chúng ta muốn trả lời câu hỏi sau: Muốn thời gian thực hiện dự án có độ tin cậy 90% thì thời gian tối thiểu (tính theo số tuần) là bao nhiêu? Đặt $P(T \leq t) = 90\%$. Tra bảng phân phối chuẩn tắc $N(0, 1)$, tìm được $z = 1,28$. Vì $z = (t - 19)/3 = 1,28$ nên $t = 19 + 3 \cdot 1,28 \approx 23$ (tuần). Như vậy, dự án đang xem xét có khả năng hoàn thành với độ tin cậy tới 90% trong vòng (không vượt quá) 23 tuần.

2.3. Điều chỉnh dự án khi kế hoạch một số hoạt động bị phá vỡ

Ví dụ: Đôi khi trong quá trình thực hiện dự án, kế hoạch của một số hoạt động bị phá vỡ. Chính vì vậy, khi phát hiện dự án đang bị chậm so với kế hoạch đề ra ta cần định lại thời gian thực hiện (thời gian rút gọn) một số hoạt động trong giai đoạn tới. Xét các dữ kiện cho trong hình II.9 và bảng II.16.



Hình II.9. Sơ đồ mạng PERT dự án cần điều chỉnh

Bảng II.16. Số liệu điều chỉnh khi kế hoạch bị phá vỡ

Hoạt động	Thời gian định mức	Thời gian rút gọn	Kinh phí bổ sung/ 1 đơn vị thời gian rút gọn (triệu đồng)
A	6	4	2
B	4	3	3
C	3	2	1
D	8	6	1,5
E	7	4	0,5

Sau khi có thời gian định mức cho các hoạt động như trong bảng II.16, dễ dàng tìm được thời gian tối thiểu cần thiết để hoàn thành kế hoạch là 16 (tuần). Tuy nhiên do yêu cầu mới, cần rút gọn thời gian hoàn thành dự án trong vòng (không vượt quá) 10 (tuần). Muốn vậy ta thực hiện các điểm sau:

- Tìm thời gian tối thiểu dự định thực hiện dự án (16 tuần) và tìm đường găng.
- Ước tính thời gian rút gọn tối đa (cột 3, bảng II.16)
- Khi rút gọn thời gian trên đường găng cũng phải chú trọng đồng thời các cung đường khác.

Trên hình II.9, ta thấy cần thực hiện A, C và E với thời gian rút gọn tối đa (4, 2, 4 để tổng các thời gian thực hiện các hoạt động găng là 10 tuần), đồng thời rút gọn các hoạt động B và D ở mức cho phép:

- Phương án 1: rút bớt thời gian thực hiện hoạt động B một tuần và rút bớt D một tuần.
- Phương án 2: không rút bớt B và rút bớt D hai tuần.

Vậy khi cần điều chỉnh thời gian thực hiện dự án ta cần thay đổi kế hoạch của một số hoạt động theo các bước đã nêu trên.

Tuy có nhiều phương án điều chỉnh dự án, nhưng trong việc phá vỡ kế hoạch các hoạt động của dự án để đáp ứng tiến độ mới cần chú ý về khía cạnh chi phí gia tăng để có một phương án tối ưu đảm bảo rút gọn được thời gian thực hiện với chi phí nhỏ nhất. Đối với ví dụ trên ta chọn phương án 2.

Có thể áp dụng phương pháp tổng quát để điều chỉnh dự án theo các mục tiêu ở trên (phương pháp đơn hình cho BTQH TT đơn và đa mục tiêu) như sẽ được trình bày sau đây.

2.4. Tính thời gian rút gọn tối ưu bằng phương pháp đơn hình

Để tính thời gian rút gọn bằng phương pháp đơn hình (có thể sử dụng các phần mềm máy tính thích hợp), ta phải đưa ra được mô hình toán học, hay cách khác, cần phát biểu được BTQH TT (đơn hay đa mục tiêu).

Trước hết, cần xác định các biến quyết định. Gọi x_1, x_2, x_3, x_4, x_5 là các thời điểm mà các hoạt động xảy ra (tại các nút); y_A, y_B, y_C, y_D, y_E là thời gian cần rút bớt cho các hoạt động để yêu cầu mới về đẩy nhanh tiến độ được thoả mãn. Ta có BTQH TT đa mục tiêu sau (cần cực tiểu hóa cả thời gian thực hiện dự án lẫn tổng chi phí gia tăng):

$$\text{Mục tiêu 1: } z_1 = x_5 \rightarrow \text{Min}$$

$$\text{Mục tiêu 2: } z_2 = 2y_A + 3y_B + y_C + 1,5y_D + 0,5y_E \rightarrow \text{Min}$$

với các ràng buộc:

$$\begin{cases} x_2 \geq 6 - y_A + x_1 \\ x_4 \geq 3 - y_C + x_2 \\ x_3 \geq 4 - y_B + x_1 \\ x_5 \geq 7 - y_E + x_4 \\ x_5 \geq 8 - y_D + x_3 \\ x_i \geq 0, i = 1, 2, 3, 4, 5 \\ y_j \geq 0, j = A, B, C, D, E \\ y_A \leq 2, y_B \leq 1, y_C \leq 1, y_D \leq 2, y_E \leq 3 \\ x_5 \leq x_1 + 10 \end{cases} \quad (*)$$

Có 2 cách giải mô hình:

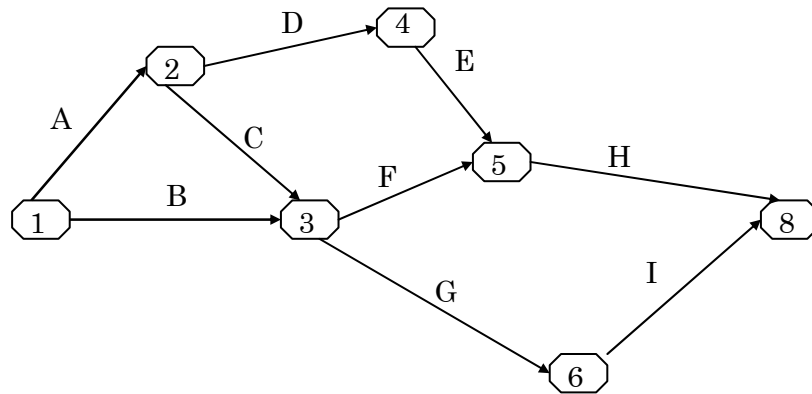
– Chuyển mục tiêu 1 thành ràng buộc (*). Nếu lúc đó BTQH TT không có phương án khả thi thì phải nói lỏng dần (*): chẳng hạn thay (*) bởi $x_5 \leq x_1 + 11$.

– Để nguyên cả hai mục tiêu để giải theo phương pháp BTQH TT đa mục tiêu.

2.5. Áp dụng mạng PERT trong phân tích chi phí và quản lý tài chính dự án

Trong giai đoạn đầu ứng dụng PERT và CPM, các phương pháp này thường được áp dụng cho bài toán tìm thời gian tối thiểu thực hiện dự án, tìm các hoạt động gắng. Chúng ít khi được áp dụng để phân tích chi phí, mặc dù trong các dự án thì việc phân tích chi phí (bao gồm chi phí trực tiếp, gián tiếp và chi phí tiện ích) cũng rất quan trọng. Tuy nhiên ngày nay, PERT và CPM được áp dụng rất rộng rãi cho các bài toán dạng này.

Ví dụ: Chúng ta xem xét dự án với các dữ kiện cho trong bảng II.17 và hình II.10.



Hình II.10. Mạng PERT cho bài toán phân tích chi phí

Bảng II.17. Dữ kiện cho bài toán PERT chi phí

Hoạt động	EST	LST	Thời gian thực hiện (tháng)	Tổng chi phí (triệu đồng)	Chi phí / một tháng (triệu đồng)
A	0	0	3	30	10
B	0	8	2	200	100
C	3	9	1	40	40
D	3	3	4	20	5
E	7	7	5	75	15
F	4	10	2	100	50
G	4	10	1	75	75
H	12	12	3	18	6
I	5	11	4	240	60

Để thấy, thời gian tối thiểu để hoàn thành dự án là 15 (tháng).

Nguyên tắc điều hành tài chính một dự án là:

- Luồng kinh phí phải được đưa vào dần dần sao cho đáp ứng được tiến độ dự án.
- Nếu kinh phí đưa vào thừa hoặc thiếu (theo tiến độ) thì phải kịp thời điều chỉnh.

Cần nắm bắt được: những hoạt động nào không dùng hết kinh phí dự kiến, những hoạt động nào sử dụng kinh phí nhiều hơn dự kiến để có sự điều chỉnh thích hợp.

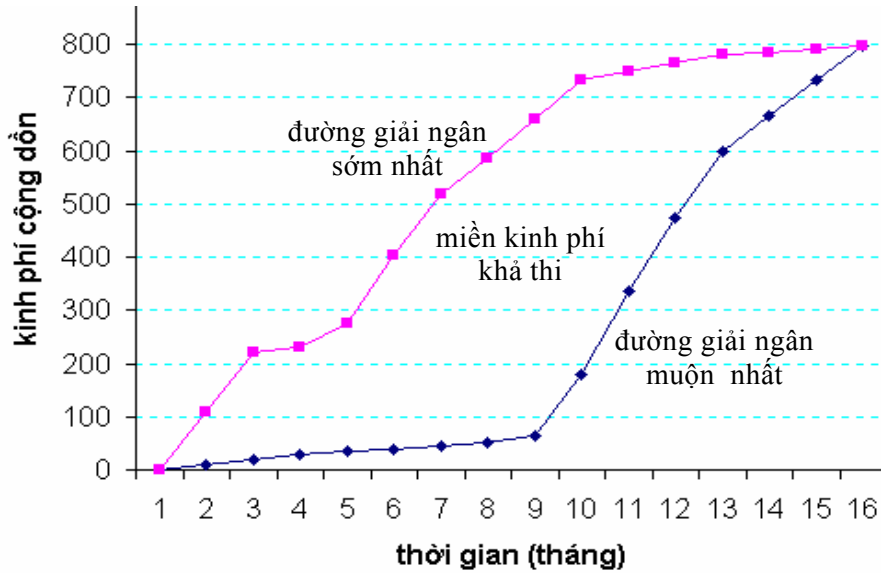
- Các báo cáo định kì cho phép kiểm soát được dự án về tiến độ và luồng kinh phí.

Muốn vậy, trước hết cần lập bảng theo dõi kinh phí cho dự án từ tháng 1 đến tháng 15 (xem bảng II.18). Phần trên của từng ô ứng với các hoạt động giải ngân sớm nhất, phần dưới ứng với giải ngân muộn nhất. Hai hàng cuối bảng dành cho kinh phí

trong từng tháng và tổng kinh phí cộng dồn cho tới tháng đó tương ứng với hoạt động giải ngân sớm nhất và giải ngân muộn nhất.

Bảng II.18. Dữ kiện cho bài toán PERT chi phí

T.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	10 10	10 10	10 10												
B	100	100								100	100				
C				40							40				
D				5 5	5 5	5 5	5 5								
E								15 15	15 15	15 15	15 15	15 15			
F					50	50						50	50		
G					75							75			
H													6 6	6 6	6 6
I						60	60	60	60				60	60	60
Σ	110 10	110 10	10 10	45 5	130 5	115 5	65 5	75 15	75 115	15 155	15 140	15 125	6 66	6 66	6 66
$\Sigma+$	110 10	220 20	230 30	275 35	405 40	520 45	585 50	660 65	735 180	750 335	765 475	780 600	786 666	792 732	798 798



Hình II.11. Đồ thị miền kinh phí khả thi

Dựa vào bảng II.18, có thể vẽ được đồ thị miền kinh phí khả thi như trên hình II.11. Nếu tiến độ giải ngân nằm ngoài miền kinh phí khả thi thì cần gấp rút đưa ra các biện pháp điều chỉnh tiến độ giải ngân. Ngoài ra, cũng có thể điều chỉnh kinh phí các hoạt động của dự án dựa vào bảng II.18.

Lưu ý

Các vấn đề cơ bản cần giải quyết khi áp dụng phương pháp PERT hay CPM trong theo dõi và đánh giá dự án là:

- Xác định được sơ đồ mạng PERT của dự án.
- Tìm được đường găng và các hoạt động găng.
- Tính được độ tin cậy ứng với các mốc thời hạn hoàn thành dự án khi số liệu là ngẫu nhiên.
- Biết cách điều chỉnh thời gian rút gọn khi tiến độ thực hiện dự án là chậm so với kế hoạch.
- Phân tích chi phí và điều hành kinh phí dự án.

3. Một số mô hình mạng khác

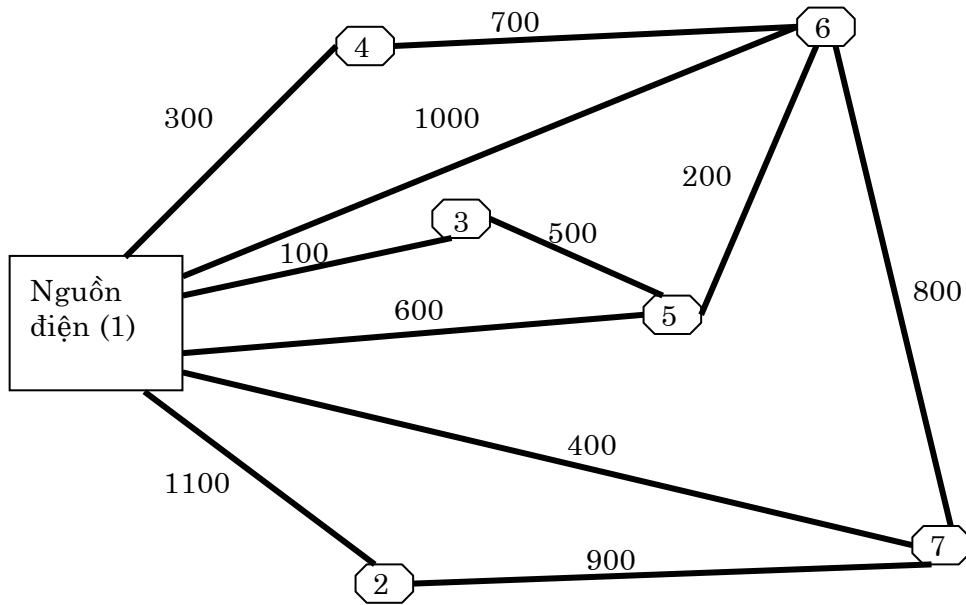
3.1. Bài toán cây khung tối thiểu

Bài toán cây khung tối thiểu được nghiên cứu và ứng dụng trong nhiều lĩnh vực (Công nghệ thông tin, Điện lực, Quy hoạch thủy lợi,...). Vấn đề đặt ra là cần xác định một mạng đường đi tới mọi nút của mạng xuất phát từ một nút nào đó trong mạng, sao cho tổng độ dài các cung đường này là ngắn nhất. Phương pháp tốt nhất giải bài toán cây khung tối thiểu thuộc về R. Prim sẽ được trình bày trong mục này.

Ví dụ: Mắc điện cho 7 xã trong huyện từ nguồn phát điện đến được 7 xã với chi phí đường dây là bé nhất. Sơ đồ khoảng cách từ nguồn cung cấp điện tới trạm điện các xã như trên hình II.12.

Bài toán đặt ra là phải phát triển được cây khung hay đường đi tối thiểu sao cho tổng chiều dài các cung đường là bé nhất.

Để giải ta lập bảng II.19 (chiều dài các cung đường được quy gọn), trong đó M là kí hiệu một số $\approx +\infty$, biểu thị cung đường không thể xảy ra trên thực tế. Mỗi hàng hay mỗi cột của bảng đều biểu thị các nút, chẳng hạn ô nằm trên giao của hàng 2 và cột 7 (cũng giống như ô nằm trên giao của hàng 7 và cột 2) đều chứa số 9, là khoảng cách giữa hai nút 2 và 7. Một hàng và một cột được nói là liên thông với nhau nếu ô nằm trên giao của hàng và cột này chứa giá trị khác M .



Hình II.12. Sơ đồ khoảng cách từ nguồn điện tới các xã

Bảng II.19. Bảng khoảng cách các cung đường

(Nút hàng)	Nút (cột)						
	1√	2	3√	4	5	6	7
√ 1	0	11	1	3	6	10	4
2	11	0	M	M	M	M	9
√ 3	1	M	0	M	5	M	M
√ 4	3	M	M	0	M	7	M
5	6	M	5	M	0	2	M
...	10	M	M	7	2	0	8
√ 7	4	9	M	M	M	8	0

Thuật giải Prim

– Bước khởi tạo: Lập bảng khoảng cách giữa các nút mạng. Trong bảng trên, chọn cột bất kì (ví dụ cột 1, tức là ta chọn nút 1 để bắt đầu), gạch bỏ cột vừa chọn ra khỏi bảng.

– Các bước lặp:

Bước 1: Đánh dấu vào hàng tương ứng (hàng cùng chỉ số) với cột vừa chọn. Trên các hàng đã được đánh dấu tìm ô có giá trị nhỏ nhất.

Bước 2: Chọn cột tương ứng với ô vừa tìm được (cột 3 biểu diễn nút chọn mới, ghi cung đường vừa tìm được $1 \rightarrow 3$), rồi gạch bỏ nó đi (gạch bỏ cột 3). Nếu trong bảng vẫn còn các cột chưa gạch bỏ hết thì quay về bước 1, nếu trái lại chuyển sang bước kết thúc.

– Bước kết thúc: Nếu tất cả các cột đã bị gạch bỏ hết thì dừng với tất cả các cung đường liên thông tìm được tạo nên cây khung tối thiểu.

Chú ý: Những câu in nghiêng minh họa cho bước khởi tạo và bước lặp đầu tiên. Sau 6 bước lặp, quá trình giải kết thúc với các cung đường sau: $1 \rightarrow 3, 1 \rightarrow 4, 1 \rightarrow 7, 3 \rightarrow 5, 5 \rightarrow 6$ và $7 \rightarrow 2$. Tổng độ dài các cung đường của cây khung tối thiểu là $\Sigma = 1 + 3 + 4 + 5 + 2 + 9 = 24$. Ngoài ra, có thể chọn nút khởi tạo là bất cứ nút nào.

Thuật toán Prim còn được ứng dụng trong các bài toán xác định chi phí tối thiểu nhiều dạng khác. Việc chứng minh thuật giải trên xin dành lại cho người đọc quan tâm nghiên cứu các vấn đề về thuật toán.

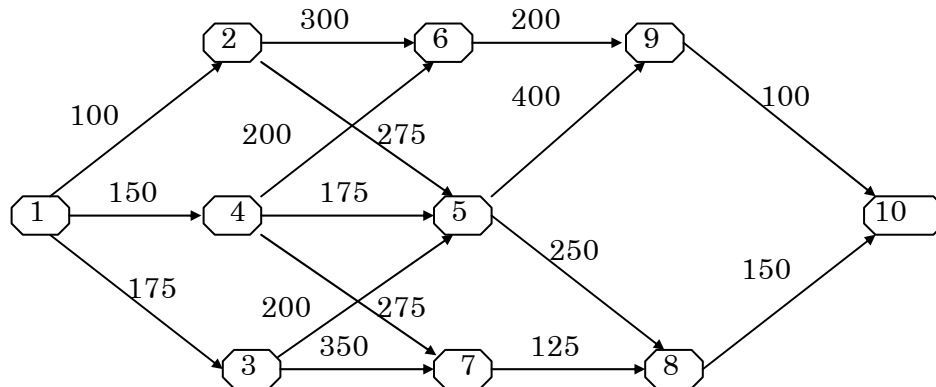
3.2. Bài toán tìm đường đi ngắn nhất và quy hoạch động

Bài toán tìm đường đi ngắn nhất

Trong bài toán tìm đường đi ngắn nhất, chúng ta muốn xác định hành trình ngắn nhất từ một địa điểm xuất phát (điểm gốc) để đi tới điểm cần đến (điểm đích) trên một mạng liên thông. Để cho dễ hiểu, chúng ta xem xét ví dụ sau đây.

Ví dụ: Bài toán người đi du lịch.

Có một người đi du lịch, xuất phát từ nút 1 và kết thúc hành trình ở nút 10 theo hành trình trên hình II.13.



Hình II.12. Sơ đồ hành trình đường đi

Người du lịch xuất phát từ nút 1. Trong giai đoạn đầu anh ta chỉ được quyền (và bắt buộc) chọn một trong ba nút (thành phố) 2, 3, 4 để vào thăm quan. Giai đoạn tiếp theo, anh ta chỉ được chọn một trong ba nút 5, 6, 7 để du lịch. Trong giai đoạn tiếp nối, anh ta có quyền vào một trong hai nút 8 hoặc 9 trước khi kết thúc hành trình tại nút 10.

Như vậy, trong mỗi giai đoạn người đi du lịch chỉ được quyền đi vào một thành phố (mỗi thành phố được coi là một trạng thái của giai đoạn đó). Hãy tìm cách xác định đường đi ngắn nhất từ nút 1 tới nút 10 thỏa mãn các điều kiện đặt ra của bài toán.

Nguyên tắc tối ưu Bellman trong quy hoạch động

Sử dụng nguyên tắc tối ưu Bellman trong quy hoạch động để giải bài toán người du lịch, chúng ta chia bài toán thành nhiều giai đoạn, tức là thành nhiều bài toán nhỏ. Tại mỗi giai đoạn ta cần tìm phương án tối ưu là các phương án tốt nhất của tình trạng hiện có, xét trong mối quan hệ với các phương án tối ưu đã tìm được của các giai đoạn trước.

Ta có thể giải quyết bài toán dần theo từng giai đoạn theo cách *tính toán tiến* hoặc *tính toán lùi*. Để giải bài toán này, ta áp dụng cách tính toán lùi (*backward computing*) với các kí hiệu và dữ kiện cho trong bảng II.20.

Bảng II.20. Các biến trạng thái của bài toán quy hoạch động

Giai đoạn	Đầu vào	Đầu ra	Đường đi tối ưu	Khoảng cách tới đích
Giai đoạn I	8	10	8 → 10	150
	9	10	9 → 10	100
Giai đoạn II	5	8	5 → 8	400
	6	9	6 → 9	300
	7		7 → 8	275
Giai đoạn III	2	5	2 → 6	600
	3	6	3 → 5	600
	4	7	4 → 6	500
Giai đoạn IV	1	2	1 → 2	700
		3	1 → 3	775
		4	1 → 4	650

Giải thích: Sử dụng nguyên tắc tối ưu Bellman, để tìm đường đi ngắn nhất từ nút 4 tới nút 10 chúng ta tìm được phương án tối ưu là đi từ nút 4 tới nút 6 cho giai đoạn III (lúc này $d(4, 10) = d(4, 6) + \text{Min } d(6, 10) = 200 + 300 = 500$). Điều này là do hai lựa chọn khác là đi từ nút 4 tới nút 5 hay 7 thì đều cho khoảng cách từ nút 4 tới đích là nút 10 lớn hơn (chẳng hạn nếu đi qua nút 5 thì $d(4, 10) = d(4, 5) + \text{Min } d(5, 10) = 175 + 400 = 575$).

Trong bảng II.20, tại giai đoạn IV, ta thấy khoảng cách ngắn nhất tới đích là 650. Đi ngược lại, từ điểm gốc tới điểm đích ta xác định được đường đi ngắn nhất là: 1 → 4 → 6 → 9 → 10 với tổng chiều dài là 650.

Quy trình tính toán tổng quát

– Trước hết, cần chọn có các biến trạng thái (*state variables*) như mô tả trong bảng II.21.

Bảng II.21. Các biến trạng thái của bài toán quy hoạch động

Biến	Số trạng thái	Các trạng thái (nút)	Giá trị có thể xảy ra của các biến trạng thái
x_4	1	1	$x_4 \equiv 1$
x_3	3	2, 3, 4	$x_3 = 2$; $x_3 = 3$; $x_3 = 4$
x_2	3	5, 6, 7	$x_2 = 5$; $x_2 = 6$; $x_2 = 7$
x_1	2	8, 9	$x_1 = 8$; $x_1 = 9$
x_0	1	10	$x_0 = 10$

Biến trạng thái mô tả trạng thái của hệ thống trong từng giai đoạn.

– Xác định hàm mục tiêu: Đặt $F_i(x_i)$ là khoảng cách ngắn nhất tới đích tính tại giai đoạn i . Theo bảng II.20, ta thấy:

$$F_1(x_1) = \begin{cases} 150 & \text{với } x_1 = 8 \\ 100 & \text{với } x_1 = 9 \end{cases}$$

$$F_2(x_2) = \begin{cases} 400 & \text{với } x_2 = 5 \\ 300 & \text{với } x_2 = 6 \\ 275 & \text{với } x_2 = 7 \end{cases}$$

Mục đích của bài toán là cần tìm được giá trị $F_4(x_4) = F_4(1)$.

– Lập hàm truy toán: $F_{i+1}(x_{i+1}) = \text{Min} [F_i(x_i) + f_i(u_i)]$, Min tìm theo mọi tổ hợp thích hợp x_i và u_i , trong đó u_i là biến điều khiển để điều khiển chuyển trạng thái từ trạng thái x_i sang x_{i+1} và $f_i(u_i)$ là hiệu ứng của biến điều khiển tác động lên hàm truy toán (và lên hàm mục tiêu, nếu tính đến bài toán cuối cùng). Theo biểu thức của hàm truy toán ta thấy, nếu $F_i(x_i) + f_i(u_i)$ là hàm phi tuyến thì phải dùng kỹ thuật tối ưu thích hợp để tìm ra $F_{i+1}(x_{i+1})$.

Sau đây chúng ta đi tìm các hàm truy toán $F_{i+1}(x_{i+1})$ với quy trình *tính toán lùi* để giải bài toán theo từng giai đoạn, nhằm cuối cùng tìm ra được $F_4(x_4) = F_4(1)$.

Giai đoạn 1: Trong giai đoạn này, muốn chuyển từ nút 10 ($x_0 = 10$) về nút 8 ($x_1 = 8$) chẳng hạn, thì biến điều khiển u_0 phải có giá trị 150 ($u_0 = 150$). Hiệu ứng gây nên bởi u_0 là $f(u_0) = 150$. Điều này có nghĩa là nếu chuyển từ nút 10 ngược về nút 8 thì cần đi quãng đường có chiều dài là 150.

$F_0(x_0) = 0$	$x_0 = 10$	u_0	$f_0(u_0)$	$F_1(x_1)$
$x_1 = 8$	+ $u_0 = 150$	150	150	150
$x_1 = 9$	+ $u_0 = 100$	100	100	100

Chú ý: Không phải bài toán nào u_i cũng trùng với hiệu ứng $f_i(u_i)$ của nó. Nói chung, biến điều khiển u_i có thể gây ra hiệu ứng $f_i(u_i)$ khác với u_i cả về độ lớn cũng như đơn vị đo.

Giai đoạn 2:

x_2	$x_1 = 8$	$x_1 = 9$	$F_1(x_1) + f_1(u_1)$		$F_2(x_2) = \text{Min}[F_1(x_1) + f_1(u_1)]$
			$x_1 = 8$	$x_1 = 9$	
5	$+u_1 = 250$	$+u_1 = 400$	400	500	$400 = 150 + 250$
6	–	$+u_1 = 200$	–	300	$300 = 100 + 200$
7	$+u_1 = 125$	–	275	–	$275 = 150 + 125$

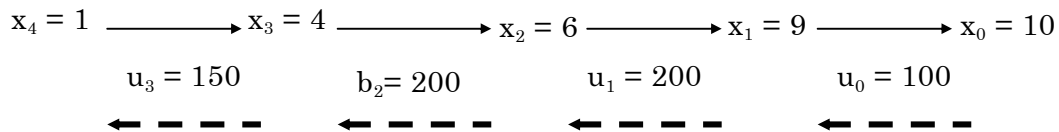
Giai đoạn 3:

x_3	x_2			$F_2(x_2) + f_2(u_2)$			$F_3(x_3) = \text{Min}[F_2(x_2) + f_2(u_2)]$
	5	6	7	$x_2 = 5$	$x_2 = 6$	$x_2 = 7$	
2	$u_2 = 275$	$u_2 = 300$	–	675	600	–	600
3	$u_2 = 200$	–	$u_2 = 350$	600	–	625	600
4	$u_2 = 175$	$u_2 = 200$	$u_2 = 275$	575	500	550	500

Giai đoạn 4:

x_4	$x_3 = 2$	$x_3 = 3$	$x_3 = 4$	$F_3(x_3) + f_3(u_3)$			$F_4(x_4) = \text{Min}[F_3(x_3) + f_3(u_3)]$
				$x_3 = 2$	$x_3 = 3$	$x_3 = 4$	
1	$u_3 = 100$	$u_3 = 175$	$u_3 = 150$	700	775	650	650

Đáp số: $F_4(x_4) = F_4(1) = 650$ với đường đi ngắn nhất trên hình II.14.



Hình II.14. Đường đi ngắn nhất $1 \rightarrow 4 \rightarrow 6 \rightarrow 9 \rightarrow 10$

3.3. Áp dụng quy hoạch động cho một số bài toán ngành điện

Bài toán 1

Cần phân phối công suất tối ưu của n nhà máy điện với phụ tải tổn thất cố định. Biết chi phí của các nhà máy là hàm $f_i(p_i)$ phụ thuộc vào công suất p_i , với $i = 1, 2, \dots, n$. Cần xác định các giá trị của p_i sao cho tổng chi phí là cực tiểu. Vậy ta có bài toán tối ưu sau:

Hàm mục tiêu:

$$z = f_1(p_1) + \dots + f_n(p_n) \rightarrow \text{Min}$$

với các ràng buộc:

$$\begin{cases} p_1 + p_2 + \dots + p_n = P \\ 0 \leq p_i \leq P_{i, \max} \end{cases}$$

trong đó P là tổng phụ tải, $P_{i, \max}$ là công suất tối đa cho phép.

Chẳng hạn, với $n = 3$ ta có BTQHHTT (nguyên) sau đây:

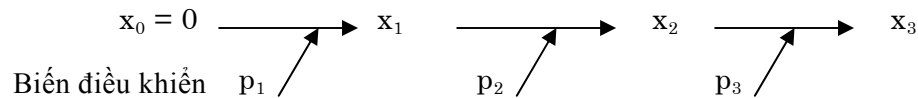
$$z = 3p_1 + 2p_2 + p_3 \rightarrow \text{Min}$$

$$\begin{cases} p_1 + p_2 + p_3 = 15 \\ 0 \leq p_i \leq 6; 0 \leq p_2 \leq 6; 0 \leq p_3 \leq 8 \end{cases}$$

nếu đã biết:

$$\begin{cases} f_1(p_1) = 3p_1 \\ f_2(p_2) = 2p_2 \\ f_3(p_3) = p_3 \end{cases}$$

Chúng ta xét phương pháp giải bài toán này với giả thiết các công suất p_i là nguyên. Đặt các biến trạng thái là x_1, x_2, x_3 ; các biến điều khiển là p_1, p_2, p_3 với quan hệ như sau: $x_1 = p_1, x_2 = p_1 + p_2, x_3 = p_1 + p_2 + p_3 = 15$. Các hiệu ứng gây nên bởi các biến điều khiển là $f_i(p_i)$ với $i = 1, 2, 3$.



Thiết lập hàm truy toán $F_{i+1}(x_{i+1}) = \text{Min} [F_i(x_i) + f_{i+1}(p_{i+1})]$. Đặt $F_0(x_0) = 0$, dễ thấy: $F_1(x_1) = \text{Min} f_1(p_1)$, $F_2(x_2) = \text{Min}[f_1(p_1) + f_2(p_2)]$ và $F_3(x_3) = \text{Min}[f_1(p_1) + f_2(p_2) + f_3(p_3)] = 3p_1 + 2p_2 + p_3$. Mục tiêu cuối cùng là cực tiểu hoá $z = F_3(x_3)$.

Sử dụng nguyên tắc tối ưu Bellman ta chia bài toán ra các giai đoạn sau đây (với quy trình tính toán tiến).

Giai đoạn 1: chỉ xét công suất p_1 ;

Giai đoạn 2: chỉ xét công suất p_1 và p_2 ;

Giai đoạn 3: xét các công suất p_1, p_2 và p_3 .

Giai đoạn 1: (Coi $F_0(x_0) = 0$)

x_1	$x_0 = 0$	$f_1(p_1) = 3p_1$	$F_1(x_1) = \text{Min}$ $[F_0(x_0) + f_1(p_1)]$
0	$p_1 = 0$	0	0
1	$p_1 = 1$	3	3
2	$p_1 = 2$	6	6
3	$p_1 = 3$	9	9
4	$p_1 = 4$	12	12
5	$p_1 = 5$	15	15
6	$p_1 = 6$	18	18

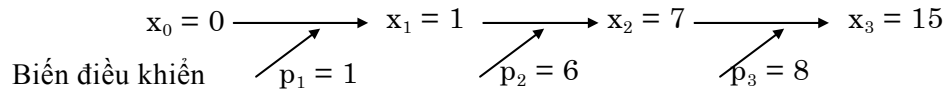
Giai đoạn 2:

x_2	x_1							$F_1(x_1) + f_2(p_2)$							$F_2(x_2) = \text{Min}$ $[F_1(x_1) + f_2(p_2)]$
	0	1	2	3	4	5	6	0	1	2	3	4	5	6	
	p_2							0	1	2	3	4	5	6	
0	0	-	-	-	-	-	-	0	-	-	-	-	-	-	0
1	1	0	-	-	-	-	-	2	3	-	-	-	-	-	2
2	2	1	0	-	-	-	-	4	5	6	-	-	-	-	4
3	3	2	1	0	-	-	-	6	7	8	9	-	-	-	6
4	4	3	2	1	0	-	-	8	9	10	11	12	-	-	8
5	5	4	3	2	1	0	-	10	11	12	13	14	15	-	10
6	6	5	4	3	2	1	0	12	13	14	15	16	17	18	12
7	-	6	5	4	3	2	1	-	15	16	17	18	19	20	15
8	-	-	6	5	4	3	2	-	-	18	19	20	21	22	18
9	-	-	-	6	5	4	3	-	-	-	21	22	23	24	21
10	-	-	-	-	6	5	4	-	-	-	-	24	25	26	24
11	-	-	-	-	-	6	5	-	-	-	-	-	27	28	27
12	-	-	-	-	-	-	6	-	-	-	-	-	-	30	30

Giai đoạn 3:

x_3	x_2								$F_2(x_2) + f_3(p_3)$						$F_3(x_3) = \text{Min}$ $[F_2(x_2) + f_3(p_3)]$
	0	6	7	8	9	10	11	12	7	8	9	10	11	12	
	p_3								7	8	9	10	11	12	
15	-	-	8	7	6	5	4	3	23	25	27	29	31	33	23

Đáp số: Tổng chi phí đạt giá trị cực tiểu là 23, với $p_1 = 1$, $p_2 = 6$, $p_3 = 8$.



Lưu ý

Các vấn đề cơ bản cần giải quyết khi áp dụng phương pháp quy hoạch động theo nguyên tắc Bellman là:

- Chia bài toán thành nhiều giai đoạn nhỏ để giải bài toán tối ưu cho từng giai đoạn. Các yếu tố của bài toán quy hoạch động là biến trạng thái, biến điều khiển, hàm truy toán và hàm mục tiêu.

- Khi chuyển từ một trạng thái nào đó (trong một giai đoạn) sang trạng thái khác (giai đoạn khác) cần có biến điều khiển.

- Mỗi giá trị của biến điều khiển gây ra một hiệu ứng lên hàm mục tiêu.

- Tùy theo các bài toán tối ưu phát sinh trong các giai đoạn mà lựa chọn phương pháp tối ưu thích hợp.

Trong ví dụ đang xét, khi các hiệu ứng $f_i(p_i)$ cho dưới dạng hàm tuyến tính với các biến p_i nhận các giá trị rời rạc/nguyên thì hàm truy toán $F_{i+1}(x_{i+1}) = \text{Min} [F_i(x_i) + f_{i+1}(p_{i+1})]$ sẽ tính được bằng thuật giải dựa trên bảng liệt kê (như phương pháp giải đã trình bày). Nếu $f_i(p_i)$ phi tuyến với các biến p_i nhận các giá trị liên tục thì để tìm $F_{i+1}(x_{i+1}) = \text{Min}[F_i(x_i) + f_{i+1}(p_{i+1})]$ ta có hai cách:

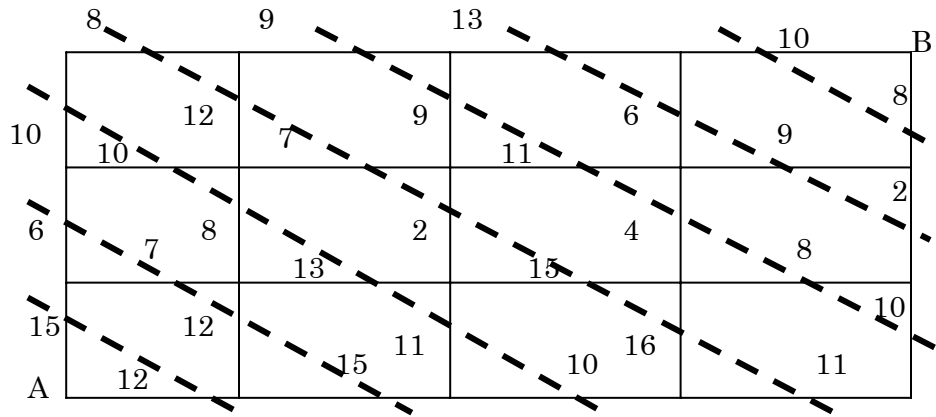
- Cách 1: rời rạc hoá theo từng mức. Chẳng hạn với $p_1 \in [0, 6]$, thì coi $p_1 \in \{0, 1, 2, 3, 4, 5, 6\}$.

- Cách 2: áp dụng phương pháp tối ưu thích hợp với biến liên tục (xem chương I) cho hàm mục tiêu. Chẳng hạn, trong ví dụ trên khi cần tìm $F_2(x_2) = \text{Min} [F_1(x_1) + f_2(p_2)] = \text{Min}[f_1(p_1) + f_2(p_2)] = \text{Min} [3p_1 + 2p_2]$ với điều kiện ràng buộc: $p_1 + p_2 \leq 15$ và $0 \leq p_1 \leq 6, 0 \leq p_2 \leq 6$, có thể áp dụng phương pháp đơn hình.

Bài toán 2

Xác định tuyến đường đi của đường dây truyền tải điện từ điểm A đến điểm B, với các chướng ngại vật khác nhau, sao cho tổng chi phí là nhỏ nhất. Các dữ kiện của bài toán cho trên hình II.15.

Như vậy để thiết lập sơ đồ đường truyền tải điện thì xuất phát từ A ta có thể định tuyến đi của đường truyền tải điện trước hết phải qua một trong hai điểm sát gần, theo hướng bắc hay hướng đông, với các chi phí là 15 và 12. Từ một trong hai điểm này, chúng ta lại tiếp tục xác định tuyến đi cho đường truyền tải điện, với các chi phí đã biết... Vậy ta có bài toán tìm đường đi với chi phí nhỏ nhất.



Hình II.15. Sơ đồ tuyến đi cho dây truyền tải điện

Bài toán này hoàn toàn tương tự với bài toán người du lịch đã xét và có thể giải bằng phương pháp quy hoạch động (*Hướng dẫn*: Chia bài toán thành nhiều giai đoạn nhỏ theo các đường với nét đứt nối trên hình II.15).

GIỚI THIỆU LÝ THUYẾT MÔ PHÒNG VÀ MÔ HÌNH HÀNG CHỜ

1. Mục đích và các công cụ của mô phỏng

1.1. Khái niệm về mô phỏng ngẫu nhiên

Mô phỏng (*Simulation*) được ứng dụng rộng rãi trong kinh tế, kỹ thuật và nhiều lĩnh vực khác. Theo *Từ điển chính xác Oxford*, bản 1976, "mô phỏng có nghĩa là giả cách, ..., làm ra vẻ như, hành động như, bắt chước giống với, mang hình thức của, giả bộ như..., làm giả các điều kiện của tình huống nào đó thông qua một mô hình với mục đích huấn luyện hoặc tiện lợi".

Về mặt ý nghĩa kỹ thuật, mô phỏng (hay nói đúng hơn, phương pháp mô phỏng) hàm chứa việc áp dụng một mô hình nào đó để tạo ra kết quả, chứ không có nghĩa là thử nghiệm một hệ thống thực tế nào đó đang cần nghiên cứu hay khảo sát. Nếu mô hình có chứa các thành phần hay yếu tố ngẫu nhiên thì chúng ta có mô phỏng ngẫu nhiên.

Thuật ngữ "phương pháp Monte-Carlo" xuất hiện từ thế chiến thứ hai khi tiến hành các mô phỏng ngẫu nhiên trong quá trình phát kiến bom nguyên tử. Ngày nay, thuật ngữ này đôi khi cũng được dùng đồng nghĩa với thuật ngữ phương pháp mô phỏng ngẫu nhiên, như khi ta nói phương pháp Monte-Carlo tính tích phân chẳng hạn, tuy nhiên, nó không được sử dụng một cách rộng rãi.

Chúng ta xét mô phỏng trên hai quan điểm: nghệ thuật và kỹ thuật (với tư cách một công cụ), mà trong một số trường hợp rất khó phân định ranh giới rạch ròi. Trong chương này chúng ta nghiên cứu mô phỏng ngẫu nhiên về phương diện một số kỹ thuật, công cụ thường được sử dụng.

1.2. Các công cụ chủ yếu của mô phỏng

Nguồn ngẫu nhiên (Source of randomness)

Để áp dụng mô phỏng ngẫu nhiên trước hết cần phải có được một nguồn các *số ngẫu nhiên*. Các số ngẫu nhiên như vậy có thể được tạo ra bởi các hàm sinh số ngẫu nhiên.

Trong nhiều ngôn ngữ lập trình (như Visual C++ 6.0, hay Builder C++ 5.0,...), ta sẽ thấy có một cặp hàm dạng SRAND (*seed*) và RANDOM để phát sinh các số (được coi là) ngẫu nhiên. Hàm SRAND, có tham số là *seed* được gọi là *hạt mầm ngẫu nhiên*, đóng vai trò khởi tạo dãy số ngẫu nhiên. Còn hàm RANDOM là hàm sinh các số ngẫu nhiên sau khi có giá trị khởi tạo.

Thông thường, các nguồn này được coi như tồn tại một cách đương nhiên. Câu hỏi

đặt ra là chúng đã "đủ tốt" hay chưa? Trong giáo trình này chúng ta không đi sâu vào phân tích vấn đề trên. Một cách khái quát có thể nói rằng, các số được gọi là số ngẫu nhiên được tạo ra như vậy còn xa mới thực sự là ngẫu nhiên. Một cách chính xác hơn, chúng chỉ có thể gọi là các số *giả ngẫu nhiên* mà thôi. Chất lượng của nguồn ngẫu nhiên có thể ảnh hưởng rất lớn tới kết quả nghiên cứu khi sử dụng phương pháp mô phỏng ngẫu nhiên.

Xét về thực chất, các số giả ngẫu nhiên là các số có tính chất tất định (*deterministic*), nhưng chúng có tính chất giống với một dãy các giá trị thể hiện của các biến ngẫu nhiên độc lập, có phân phối đều. Ví dụ, xét dãy số: 13, 8, 1, 2, 11, 14, 7, 12, 13, 12, 17, 2, 11, 10, 3,... Dãy số này trông thì có vẻ ngẫu nhiên, nhưng thực chất là tuân theo một quy tắc (hãy phát hiện ra quy tắc này). Việc tìm kiếm các thuật giải (hay các quy tắc tất định) để phát sinh ra các số giả ngẫu nhiên đủ tốt là một lĩnh vực nghiên cứu chuyên sâu của Toán học và Tin học. Mặc dù trong thực tế, khi áp dụng mô phỏng ngẫu nhiên, người ta ít khi dùng các số ngẫu nhiên tuân theo luật phân phối xác suất đều $U[0, 1)$ trên $[0, 1)$, nhưng nguồn số ngẫu nhiên loại này chính là cơ sở để mô phỏng các phân phối xác suất khác (xem mục 1.3).

Mô hình ngẫu nhiên

Hai lí do chính cho việc áp dụng mô phỏng ngẫu nhiên là:

- Tổng hợp dữ liệu theo sự phân loại nhất định.
- Đưa ra các dự báo.

Muốn áp dụng mô phỏng ngẫu nhiên cần phải có mô hình. Như vậy, mục đích của mô phỏng ngẫu nhiên cũng gần với mục đích của mô hình hoá (*modelling*). Có hai loại mô hình thường được áp dụng, đó là: mô hình cơ chế (*mechanistic model*) và mô hình tiện dụng (*convenient model*). Cả hai loại này đều có thể được sử dụng để trợ giúp các công việc nghiên cứu, khảo sát nhằm gia tăng sự nhận biết và tìm kiếm tri thức, dự báo và hỗ trợ việc ra quyết định.

Để ứng dụng một mô hình, ta có hai sự lựa chọn sau:

– Tiến hành các phân tích về mặt toán học để tìm hiểu hành vi của mô hình. Vấn đề này nhiều khi trở nên rất phức tạp với các hệ phi tuyến nhiều biến, do đó chúng ta cần đặt ra thêm các giả thiết. Tuy nhiên những giả thiết "chặt chẽ quá" của toán học đôi khi trở nên "đáng nghi ngờ" trong thực tế.

– Thí nghiệm với mô hình đang xem xét. Đối với các mô hình ngẫu nhiên các giá trị phản hồi (đầu ra) sẽ biến thiên, vì vậy chúng ta cần tạo ra hàng loạt các thể hiện (dữ liệu nhân tạo) với những bộ tham số khác nhau của mô hình.

Đôi khi cũng cần xem xét tới sự lựa chọn thứ ba, đó là tiếp cận lai (*hybrid approach*) của hai lựa chọn trên.

1.3. Mô phỏng một số phân phối xác suất

Một số phân phối xác suất thường gặp

Để áp dụng mô phỏng ngẫu nhiên cần biết một số kiến thức cơ bản mà chúng ta sẽ nhắc lại ngay sau đây. Biến ngẫu nhiên là một khái niệm quan trọng trong lý thuyết xác suất thống kê. Một cách giản lược, biến ngẫu nhiên (*random variable*), còn gọi là đại lượng ngẫu nhiên, được hiểu là biến nhận giá trị tùy thuộc vào kết quả của phép thử (phép đo, quan sát, thí nghiệm) mà không thể đoán trước được.

Biến ngẫu nhiên chia làm hai loại chính: rời rạc và liên tục. Biến rời rạc có thể nhận các giá trị từ một tập hợp (có lực lượng) hữu hạn hoặc đếm được. Biến liên tục là một khái niệm toán học về loại biến ngẫu nhiên có thể nhận các giá trị dày sát nhau trên một hoặc một số khoảng / đoạn số thực nào đó (để trình bày vấn đề đơn giản, ở đây chúng ta chỉ nói tới biến ngẫu nhiên nhận các giá trị là số thực). Trong thực tế, không có một đại lượng ngẫu nhiên nào là liên tục theo nghĩa tuyệt đối, chẳng qua là chúng ta không nhận biết được (một cách cố ý hay không cố ý) khoảng cách giữa các giá trị rất sát nhau của nó mà thôi.

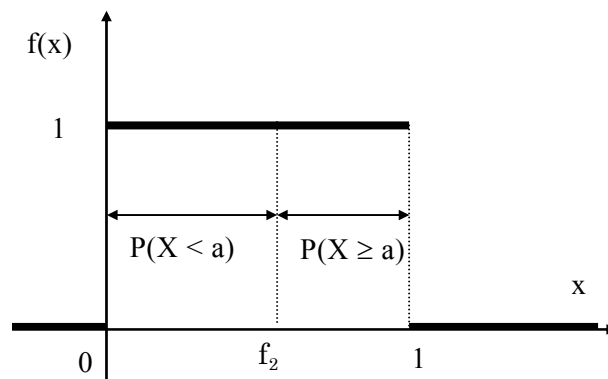
Phân phối xác suất của biến ngẫu nhiên rời rạc được minh họa qua ví dụ sau: Xét biến X có thể rơi vào một trong ba trạng thái được định lượng bởi các giá trị 6, 9, 12 với các xác suất tương ứng của các trạng thái là 0,3, 0,4 và 0,3. Chú ý rằng tổng các xác suất bằng 1 (100%) được phân phối vào các giá trị biến ngẫu nhiên X có thể lấy như trình bày trong bảng sau đây, được gọi là bảng phân phối xác suất.

Các giá trị của $X: x_i$	6	9	12
Xác suất tương ứng: p_i	0,3	0,4	0,3

(Chú ý: $\sum p_i = 1$)

Một số phân phối xác suất thường dùng của biến ngẫu nhiên liên tục và rời rạc được liệt kê dưới đây.

Phân phối đều trong $[0,1)$: X nhận các giá trị thuộc nửa khoảng $[0,1)$ với khả năng “như nhau”. Hàm mật độ xác suất $f(x)$ của nó được biểu diễn trên hình III.1.



Hình III.1. Đồ thị hàm mật độ phân phối đều

Phân phối Poát-xông: Với một hệ thống hàng chờ một kênh (xem mục 3), số lượng X tín hiệu đến trong một khoảng thời gian là một biến ngẫu nhiên.

Giả sử số tín hiệu đến trung bình trong một khoảng thời gian đã biết được (kí hiệu λ), thì với một số điều kiện nhất định có thể coi X tuân theo luật phân phối xác suất Poát-xông (*Poisson*) như sau:

Các giá trị của X: x_i	0 1x..... $+\infty$
Xác suất p_i tương ứng	$p(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$

$$\sum p_i = e^{-\lambda} \left[\frac{\lambda^0}{0!} + \frac{\lambda^1}{1!} + \frac{\lambda^2}{2!} + \dots + \frac{\lambda^x}{x!} + \dots \right] = e^{-\lambda} \times e^{\lambda} = 1 \quad .$$

Chú ý rằng số đặc trưng cho giá trị trung bình của biến ngẫu nhiên X được gọi là kì vọng. Trong phân phối Poát-xông, kì vọng của X là λ . Số đặc trưng cho độ phân tán các giá trị của X xung quanh giá trị kì vọng của nó được gọi là độ lệch chuẩn σ . Với phân phối Poát-xông thì $\sigma^2 = \lambda$.

Phân phối mũ: Trên đây ta đã xét phân phối Poát-xông của số các tín hiệu đến trong một đơn vị thời gian. Một kiểu biến ngẫu nhiên thường xét là khoảng thời gian giữa hai tín hiệu liên tiếp sẽ tuân theo phân phối mũ. Đây là biến ngẫu nhiên liên tục chỉ nhận các giá trị không âm với hàm mật độ xác suất là $f(\tau) = \lambda e^{-\lambda\tau}$. Kí hiệu biến

ngẫu nhiên đang xét là τ thì xác suất $P(\tau \leq t) = \int_0^t \lambda e^{-\lambda\tau} d\tau$ có thể hiểu là xác suất cộng dồn cho tới t. Do đó hàm phân phối xác suất của τ là: $F(t) = \int_0^t f(\tau) d\tau = \int_0^t \lambda e^{-\lambda\tau} d\tau = -e^{-\lambda\tau} \Big|_0^t = 1 - e^{-\lambda t}$.

Phân phối chuẩn tắc N(0, 1): Giả sử X là biến ngẫu nhiên có phân phối chuẩn tắc N(0,1). Lúc đó nó có kì vọng $m = 0$ và độ lệch chuẩn $\sigma = 1$. Hàm phân phối xác suất của X có dạng:

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(x) dx = \int_{-\infty}^x (1/\sqrt{2\pi}) \exp(-x^2/2) dx .$$

Cho X là biến ngẫu nhiên tuân theo luật phân phối chuẩn $N(m, \sigma^2)$ có kì vọng m, độ lệch chuẩn σ . Lúc đó, thực hiện phép đổi biến $Z = \frac{X - m}{\sigma}$ thì Z là một biến ngẫu nhiên tuân theo luật phân phối chuẩn tắc N(0,1).

Mô phỏng các phân phối xác suất

Ví dụ 1: Mô phỏng phân phối đều trên $[0, 1)$

Cách 1: Dùng bảng số ngẫu nhiên (xem phụ lục 2A và 2B). Đây là các bảng số ghi lại các số (giá) ngẫu nhiên được phát sinh nhờ các hàm sinh số ngẫu nhiên trong máy tính. Chẳng hạn, sử dụng phụ lục 2B chúng ta nhận được một dãy số ngẫu nhiên: 0,10; 0,09; 0,73; 0,25 ...

Cách 2: Sử dụng các hàm sinh số ngẫu nhiên (*Random number generator*) đã được cài đặt trên máy tính.

Dù dùng bảng số ngẫu nhiên hay sử dụng các hàm sinh số ngẫu nhiên trong máy tính, ta cũng lấy ra hoặc tính được liên tiếp các số ngẫu nhiên x_i trong $[0, 1)$ với $i = 1, 2, \dots, n$. Tần số các giá trị này rơi vào k khoảng nhỏ với độ dài bằng nhau $1/k$ được chia ra từ $[0, 1)$ là gần như nhau ($\approx n/k$). Với n lớn thì các tần số đó càng sát gần n/k . Vì vậy ta coi các giá trị phát sinh được là các thể hiện của biến ngẫu nhiên X tuân theo phân phối đều trên $[0, 1)$.

Trong trường hợp cần mô phỏng biến Y phân phối đều trên $[a, b)$, ta chỉ việc tính $y_i = a + (b - a)x_i$. Chú ý rằng để phát sinh các số ngẫu nhiên nhận giá trị nguyên $0, 1, 2, \dots, N$, chỉ cần áp dụng công thức $y_i = [(N + 1)x_i]$, trong đó vế phải là phần nguyên của $(N + 1)x_i$. Một số bảng số ngẫu nhiên nguyên hay hàm sinh số ngẫu nhiên nguyên cài đặt sẵn trong các hệ máy tính cũng giúp giải quyết vấn đề này.

Ví dụ 2: Mô phỏng phân phối rời rạc với luật phân phối xác suất sau

Các giá trị của $X: x_i$	6	9	12
Xác suất p_i	0,3	0,4	0,3

Muốn mô phỏng phân phối trên, trước hết cần tạo ra một dãy các chữ số ngẫu nhiên bằng cách tra bảng số ngẫu nhiên hay dùng hàm sinh số ngẫu nhiên đã được cài đặt trong máy tính. Chẳng hạn ta có thể chọn dãy sau 1009732533 7652013586 3467354876 ... lấy từ hàng đầu bảng số ngẫu nhiên trong phụ lục 2B. Ta quy định nếu các chữ số 0, 1, 2 xuất hiện thì coi $X = 6$, nếu 3, 4, 5, 6 xuất hiện thì coi $X = 9$, còn nếu có 7, 8, 9 xuất hiện thì coi $X = 12$. Lúc đó ứng với 10 chữ số đầu tiên của dãy trên $a_1 a_2 \dots a_{10} = 1009732533$ ta có bảng sau đây cho biết các giá trị của X có thể lấy:

a_i	1	0	0	9	7	3	2	5	3	3
Các giá trị của $X: x_i$	6	6	6	12	12	9	6	9	9	9

Như vậy, đã có 10 giá trị (thể hiện) của X được tạo ra. Tương tự, có thể tạo ra các thể hiện khác của X . Do tần suất (hay xác suất thực nghiệm) của mỗi chữ số ngẫu nhiên từ 0 tới 9 trong bảng số ngẫu nhiên là khoảng 10% nên tần suất (xác suất thực nghiệm) X nhận giá trị 6, 9 và 12 theo thứ tự là 30%, 40% và 30%. Do đó có thể coi $P(X = 6) = 30\%$, $P(X = 9) = 40\%$, $P(X = 12) = 30\%$.

Vậy muốn mô phỏng phân phối của X phải phát sinh ra một loạt các giá trị (các thể hiện) x_i của biến ngẫu nhiên X tuân theo quy luật phân phối đã cho.

Ví dụ 3: Mô phỏng phân phối mũ.

Giả sử biến ngẫu nhiên τ tuân theo phân phối mũ với hàm phân phối xác suất là $F(t) = P(\tau \leq t) = 1 - e^{-\lambda t}$. Đây chính là xác suất để τ nhận giá trị không lớn hơn một số t cho trước; λ là tham số đã cho của phân phối mũ.

Nếu r là biến ngẫu nhiên có phân phối đều trên $[0, 1)$ thì $P(r \geq e^{-\lambda t}) = 1 - e^{-\lambda t} = P(\tau \leq t)$ (xem hình III.1). Do đó, $P(\ln r \geq -\lambda t) = P(-\frac{1}{\lambda} \ln r \leq t) = P(\tau \leq t)$. Vậy để phát sinh ra các giá trị ngẫu nhiên (các thể hiện) của τ thì trước hết cần phát sinh ra các giá trị ngẫu nhiên r và tính $\tau = -\frac{1}{\lambda} \ln r$. Chẳng hạn, từ bảng số ngẫu nhiên (phụ lục 2B), nếu lấy

$r = 0,10$ và $\lambda = 5$ thì $\tau = -0,2 \times \ln r = -0,2 \times \ln 0,1 = 0,46$. Tiếp theo, nếu lấy $r = 0,09$ thì $\tau = -0,2 \times \ln 0,09 = 0,482$. Cứ như vậy ta thu được một dãy các thể hiện của τ .

2. Áp dụng mô phỏng ngẫu nhiên

2.1. Vai trò của phương pháp mô phỏng

Nhiều bài toán thực tế chứa các yếu tố ngẫu nhiên, bất ổn định không giải được bằng các phương pháp giải tích. Nếu chúng ta áp dụng các phương pháp giải tích, thì trong nhiều trường hợp buộc phải công nhận những giả thiết chặt chẽ không được thoả mãn trên thực tế, và do đó lời giải tìm được cũng ít có giá trị thực tiễn. Phương pháp mô phỏng được dùng rộng rãi để giải các bài toán loại đó, nhất là những bài toán liên quan đến hệ thống lớn, bất ổn định, hàm chứa nhiều yếu tố ngẫu nhiên.

Chúng ta cần áp dụng phương pháp mô phỏng trong các tình huống sau đây:

- Khi không tìm được mô hình giải tích nào thích hợp.
 - Các hoạt động của hệ thống thường bị ngắt quãng, đứt đoạn không theo quy luật nào cả.
 - Mô phỏng là phương pháp duy nhất cho chi phí tiết kiệm và tốn ít thời gian.
- Tuy nhiên phương pháp mô phỏng có một số điểm hạn chế sau:
- Không đưa ra được lời giải chính xác.
 - Khó xác định được sai số.
 - Mô phỏng chỉ sử dụng khi môi trường có tính bất ổn định.
 - Mô phỏng chỉ tạo ra các phương án đánh giá chứ không đưa ra được kỹ thuật tìm lời giải tối ưu.
 - Mô phỏng đôi khi rất đắt tiền.

2.2. Các bước cần tiến hành khi áp dụng mô phỏng

- Xác định vấn đề hay hệ thống cần mô phỏng.
- Xác định mô hình mô phỏng.
- Đo và thu thập số liệu cần thiết cho mô hình.
- Chạy mô phỏng.
- Phân tích kết quả mô phỏng, nếu cần thì phải sửa lại phương án đã được đánh giá qua chạy mô phỏng.
- Chạy mô phỏng để kiểm chứng phương án mới.
- Kiểm tra tính đúng đắn của mọi kết luận về hệ thống thực tế được rút ra sau khi chạy mô phỏng.

Trên đây là các bước cần làm khi áp dụng mô phỏng ngẫu nhiên để tìm ra các phương án hợp lý cho các bài toán thực tế. Ngoài ra, mô phỏng còn được áp dụng để giải quyết nhiều vấn đề khác.

2.3. Một số ví dụ về áp dụng phương pháp mô phỏng

Ví dụ 1: Cần lựa chọn một trong hai chiến lược để phát triển sản phẩm, với các số liệu thu thập được cho trong ba bảng III.1, III.2 và III.3.

Bảng III.1. Xác suất thời gian phát triển sản phẩm

Thời gian phát triển sản phẩm	Xác suất	
	Chiến lược I	Chiến lược II
6	0,2	0,4
9	0,3	0,4
12	0,5	0,2

Bảng III.2. Chi phí lợi nhuận

Chi phí/giá bán	Chiến lược I	Chiến lược II
Chi phí cố định	600.000	1.500.000
Chi phí biến thiên/đơn vị	7,5	6,75
Giá bán/đơn vị sản phẩm	10	10

Bảng III.3. Doanh số phụ thuộc thời gian phát triển sản phẩm

Doanh số	Xác suất		
	6 tháng	9 tháng	12 tháng
1.000.000	0,2	0,4	0,5
1.500.000	0,8	0,6	0,5

Vấn đề đặt ra là áp dụng phương pháp mô phỏng để tính lợi nhuận trung bình của từng chiến lược, sau đó kiểm tra kết quả (so sánh với kết quả lí thuyết).

Như vậy có năm phân phối xác suất cần mô phỏng ứng với năm biến ngẫu nhiên: X_1 – thời gian phát triển sản phẩm (theo chiến lược) I, X_2 – thời gian phát triển sản phẩm II, X_3 – doanh số cho thời gian 6 tháng, X_4 – doanh số cho thời gian 9 tháng và X_5 – doanh số cho thời gian 12 tháng. Trong ví dụ này, để trình bày đơn giản về vấn đề mô phỏng các phân phối xác suất của các biến trên, ta dùng mười số ngẫu nhiên, mỗi số gồm mười chữ số ngẫu nhiên rút ra từ bảng số ngẫu nhiên – phụ lục 2A (vì vậy các chữ số 0, 1, 2,..., 9 mỗi số chiếm khoảng 10%).

a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
1	5	8	1	9	2	2	3	9	6
2	0	6	8	5	7	7	9	8	4
8	2	6	2	1	3	0	8	9	2
8	3	7	4	8	3	6	0	4	9
4	6	3	7	5	6	7	4	8	8
0	9	2	8	1	0	5	5	8	2
7	2	9	5	0	8	8	5	7	9
9	5	8	6	1	1	1	6	3	2
7	0	5	5	5	0	8	7	6	7
6	4	7	2	3	8	2	9	3	4

Ta quy định a_1 ứng với X_1 , a_2 ứng với X_2 , a_6 ứng với X_3 , a_8 ứng với X_4 và a_{10} ứng với X_5 . Ngoài ra cũng quy định:

$$a_1 = \begin{cases} 0,1 & \text{thì } X_1 = 6 \text{ tháng (thời gian phát triển sản phẩm I)} \\ 2,3,4 & \text{thì } X_1 = 9 \text{ tháng} \\ 5,6,7,8,9 & \text{thì } X_1 = 12 \text{ tháng} \end{cases}$$

$$a_2 = \begin{cases} 0,1,2,3 & \text{thì } X_2 = 6 \text{ tháng (thời gian phát triển sản phẩm II)} \\ 4,5,6,7 & \text{thì } X_2 = 9 \text{ tháng} \\ 8,9 & \text{thì } X_2 = 12 \text{ tháng} \end{cases}$$

$$a_6 = \begin{cases} 0,1 & \text{thì } X_3 = 10^6 \text{ (doanh số 6 tháng phát triển sản phẩm)} \\ 2,3,\dots, 9 & \text{thì } X_3 = 1,5 \cdot 10^6 \end{cases}$$

$$a_8 = \begin{cases} 0,1,2,3 & \text{thì } X_4 = 10^6 \text{ (doanh số 9 tháng phát triển sản phẩm)} \\ 4,5,\dots, 9 & \text{thì } X_4 = 1,5 \cdot 10^6 \end{cases}$$

$$a_{10} = \begin{cases} 0,1,2,3,4 & \text{thì } X_5 = 10^6 \text{ (doanh số 12 tháng phát triển sản phẩm)} \\ 5,6,\dots, 9 & \text{thì } X_5 = 1,5 \cdot 10^6 \end{cases}$$

Cần nhắc lại một số công thức trong lĩnh vực quản trị kinh doanh như sau:

$$+ \text{ Lợi nhuận} = (\text{Doanh số} - \text{Điểm hoà vốn}) \times (\text{Lợi nhuận} / \text{đơn vị sản phẩm})$$

$$+ \text{ Điểm hoà vốn} = (\text{Chi phí cố định}) / (\text{Lợi nhuận} / \text{đơn vị sản phẩm})$$

$$+ \text{ Lợi nhuận} / \text{đơn vị sản phẩm} = (\text{Giá bán/đơn vị sản phẩm}) - (\text{chi phí} / \text{đơn vị sản phẩm})$$

Các tính toán mô phỏng được tổng hợp trong bảng III.4.

Bảng III.4. Kết quả tính toán mô phỏng

Số ngẫu nhiên										Thời gian		Doanh số		Lợi nhuận	
a ₁	a ₂				a ₆		a ₈		a ₁₀	I	II	I	II	I	II
1	5	8	1	9	2	2	3	9	6	6	9	1,5.10 ⁶	10 ⁶	3,15.10 ⁶	1,75.10 ⁶
2	0	6	8	5	7	7	9	8	4	9	6	1,5.10 ⁶	1,5.10 ⁶	3,15.10 ⁶	3,38.10 ⁶
8	2	6	2	1	3	0	8	9	2	12	6	10 ⁶	1,5.10 ⁶	1,9.10 ⁶	3,38.10 ⁶
8	3	7	4	8	3	6	0	4	9	12	6	1,5.10 ⁶	1,5.10 ⁶	3,15.10 ⁶	3,38.10 ⁶
4	6	3	7	5	6	7	4	8	8	9	9	1,5.10 ⁶	1,5.10 ⁶	3,15.10 ⁶	3,38.10 ⁶
0	9	2	8	1	0	5	5	8	2	6	12	10 ⁶	10 ⁶	1,9.10 ⁶	1,75.10 ⁶
7	2	9	5	0	8	8	5	7	9	12	6	1,5.10 ⁶	1,5.10 ⁶	3,15.10 ⁶	3,38.10 ⁶
9	5	8	6	1	1	1	6	3	2	12	9	10 ⁶	1,5.10 ⁶	1,9.10 ⁶	3,38.10 ⁶
7	0	5	5	5	0	8	7	6	7	12	6	1,5.10 ⁶	10 ⁶	3,15.10 ⁶	1,75.10 ⁶
6	4	7	2	3	8	2	9	3	4	12	9	10 ⁶	1,5.10 ⁶	1,9.10 ⁶	3,38.10 ⁶

$$\text{Điểm hoà vốn của chiến lược I} = \frac{600.000}{10 - 7,5} = 240.000$$

$$\text{Điểm hoà vốn của chiến lược II} = \frac{1.500.000}{10 - 6,75} = 461538$$

Bảng III.5. So sánh lợi nhuận giữa chiến lược I và II

Tổng lợi nhuận	Chiến lược I	Chiến lược II
	26,5 × 10 ⁶	28,91 × 10 ⁶
Lợi nhuận trung bình (Σ lợi nhuận/ 10)	2,65 × 10 ⁶	2,891 × 10 ⁶

Cần chú ý rằng trong bảng III.5 là kết quả tính toán khi chạy mô phỏng 10 lượt ứng với 10 số đã chọn ra. Nếu ta lấy càng nhiều số ngẫu nhiên thì độ chính xác đạt được càng cao. Vì vậy, nếu việc tính toán trên đây được lập trình và chạy trên máy tính với hàng trăm, hàng ngàn lượt thì độ chính xác sẽ rất cao.

Qua các phân tích trên ta thấy

Để tiến hành mô phỏng cần phải có:

– Cơ sở dữ liệu (*DataBase*)

– Cơ sở tri thức (*KnowledgeBase*)

Kiểm tra kết quả mô phỏng trên bằng cách so sánh với kết quả lí thuyết được thực hiện như sau:

Doanh số chiến lược I = $0,2 \times (0,2 \times 10^6 + 0,8 \times 1,5 \times 10^6) + 0,3 \times (0,4 \times 10^6 + 0,6 \times 1,5 \times 10^6) + 0,5 \times (0,5 \times 10^6 + 0,5 \times 1,5 \times 10^6) = 1,295 \times 10^6$. Lợi nhuận trung bình chiến lược I = $(1,295 - 0,24) \times 2,5 \times 10^6 = 2,637 \times 10^6$. Kết quả tính toán mô phỏng là $2,65 \times 10^6$ rất sát với kết quả này.

Tương tự ta tính được doanh số và lợi nhuận trung bình cho chiến lược II ($2,84 \times 10^6$), và rút ra được kết luận về độ chính xác của tính toán mô phỏng.

Ví dụ 2: Tìm xác suất p để bao lồi của 4 điểm lấy bất kì trong vòng tròn đơn vị là một hình tam giác (*bài toán Sylvester*).

Có lẽ cách đơn giản nhất để giải bài toán này là áp dụng mô phỏng ngẫu nhiên theo các bước sau đây:

i) Gán cho biến đếm *Counter* giá trị ban đầu bằng 0.

ii) Tiến hành một đợt gieo ngẫu nhiên tám số thực $0 \leq r_i \leq 1$ và $0 \leq \varphi_i \leq 2\pi$ (để gieo φ_i ta lấy số ngẫu nhiên thuộc $[0, 1)$ gieo được nhân thêm với 2π), $i = 1, 2, 3, 4$. Đặt $x_i = r_i \sin \varphi_i$, $y_i = r_i \cos \varphi_i$, ta có 4 điểm nằm trong hình tròn đơn vị. Đặt $A = (x_1, y_1)$, $B = (x_2, y_2)$, $C = (x_3, y_3)$, $D = (x_4, y_4)$.

iii) Ta tính diện tích 4 tam giác ABC, ABD, ACD và BCD. Nếu ta có diện tích của một tam giác bằng tổng diện tích ba tam giác còn lại thì ta được bao lồi của bốn điểm A, B, C và D là một tam giác. Ta tăng giá trị của biến đếm *Counter* lên thêm 1, nếu trái lại biến đếm giữ nguyên giá trị cũ và quay về bước ii).

Quá trình cứ thế tiếp diễn cho tới khi số đợt gieo đạt tới một giá trị khá lớn được chọn từ trước (chẳng hạn 10000 đợt hoặc 20000 đợt, hoặc 100000 đợt). Mỗi lần biến đếm *Counter* sẽ có giá trị kết thúc khác nhau. Lấy tỉ số của số đó và số đợt, ta có tần suất xuất hiện của sự kiện "*bao lồi của 4 điểm là tam giác*". Số tần suất này theo luật số lớn là giá trị gần đúng của xác suất cần tính.

Theo các tài liệu chuyên khảo, lời giải đúng của bài toán là: $p = 35/(12\pi^2) \approx 0,29552$. Rõ ràng, trong trường hợp này, ta nên áp dụng mô phỏng ngẫu nhiên để tính ra tần suất (việc dễ thực hiện), thay thế cho việc tính xác suất theo lí thuyết (việc khó thực hiện).

Sau đây là văn bản chương trình máy tính với ngôn ngữ lập trình C giải bài toán Sylvester.

```
#include <stdio.h>
#include <math.h>
#include <stdlib.h>
#include <conio.h>
```

```
#define PI 3.14159265358979
```

```

const double esp =4.5e-12;
struct diem {double x,y;};

/* Tao bo so ngau nhien bang cach tron – shuffling */
int rd(){return rand()%10;}
double radm()
{return rand()%100+0.1*rd()+0.001*rd()+0.0001*rd();}

/* Chuong trinh chinh */
void main()
{
clrscr(); long int count = 0, reps; diem d[4]; double r, goc;
printf("\n Provide number of repetitions:");
scanf("%ld",&reps); printf("\n reps= %ld",reps); srand(19587);

/* Gieo ngau nhien 4 diem va tinh dien tich bon tam giac */
for(long int i=0;i<reps;i++)
{
for (int j=0;j<4;j++)
{
r=radm(); goc=radm()/100;
d[j].x=r*cos(goc); d[j].y=r*sin(goc);
}
double d12=sqrt(pow(d[0].x-d[1].x,2)+pow(d[0].y-d[1].y,2));
double d13=sqrt(pow(d[0].x-d[2].x,2)+pow(d[0].y-d[2].y,2));
double d14=sqrt(pow(d[0].x-d[3].x,2)+pow(d[0].y-d[3].y,2));
double d23=sqrt(pow(d[1].x-d[2].x,2)+pow(d[1].y-d[2].y,2));
double d24=sqrt(pow(d[1].x-d[3].x,2)+pow(d[1].y-d[3].y,2));
double d34=sqrt(pow(d[2].x-d[3].x,2)+pow(d[2].y-d[3].y,2));
double p123=(d12+d23+d13)/2;
double p124=(d12+d24+d14)/2;
double p134=(d13+d34+d14)/2;
double p234=(d23+d24+d34)/2;
double s123=p123*(p123-d12)*(p123-d13)*(p123-d23);
double s124=p124*(p124-d12)*(p124-d14)*(p124-d24);
double s134=p134*(p134-d13)*(p134-d14)*(p134-d34);
double s234=p234*(p234-d23)*(p234-d24)*(p234-d34);

```



```

/* Cac truong hop bao loi cua 4 diem la tam giac */
if(s123>0&& s123>0&& s134>0&& s234>0)
{
s123=sqrt(s123);s124=sqrt(s124);
s134=sqrt(s134);s234=sqrt(s234);
if(fabs(s123-(s124+s134+s234))<esp)
count++;
else if(fabs(s124-(s123+s134+s234))<esp)
count++;
else if(fabs(s134-(s123+s124+s234))<esp)
count++;
else if(fabs(s234-(s123+s124+s134))<esp)
count++;
}
else count++;
}
printf("\n Number of repetitions = %ld",reps);
printf("\n Number of successes = %ld",count);
printf("\n Probability to compute= %0.9f", count*1.0/reps);
getch();
}

```

Các kết quả chạy chương trình trong bốn lần như sau:

esp = 4.5e-12	esp = 4.5e-12	esp = 4.5e-12	esp = 4.5e-12
số lần lặp 10000	số lần lặp 20000	số lần lặp 100000	số lần lặp 200000
số lần thành công 3050	số lần thành công 5941	số lần thành công 29594	số lần thành công 58993
xác suất: 0.30500	xác suất: 0.29705	xác suất: 0.29594	xác suất: 0.294965

Ví dụ 3: Giải bài toán tối ưu toàn cục sau (xem lại mục 4.2 chương I về bài toán tối ưu toàn cục phi tuyến).

$$f(x) = 4x_1^2 - 2,1x_1^4 + x_1^6 / 3 + x_1x_2 - 4x_2^2 + 4x_2^4 \rightarrow \text{Min}$$

với điều kiện ràng buộc (miền ràng buộc D):

$$\begin{cases} -2,5 \leq x_1 \leq 2,5 \\ -1,5 \leq x_2 \leq 1,5 \end{cases}$$

Bài toán trên đây có 4 cực tiểu địa phương và 2 cực tiểu toàn cục, có tên gọi là

“Bài toán lưng lạc đà”.

Ta tìm phương án tối ưu toàn cục bằng phương pháp mô phỏng tôi SA (*Simulated Annealing*). Phương pháp (SA) mô phỏng quá trình một vật thể rắn sau khi bị nung nóng ở nhiệt độ rất cao được để nguội từ từ về một nhiệt độ rất thấp. Lúc đó hàm năng lượng của vật thể sẽ đạt mức thấp nhất. Thuật giải SA áp dụng mô phỏng ngẫu nhiên (bằng lí thuyết xích Markov, như sẽ trình bày trong chương IV, có thể chứng minh được quá trình SA sẽ hội tụ theo xác suất về lời giải tối ưu toàn cục) như sau:

Bước khởi tạo

Ta xuất phát từ một phương án X bất kì ban đầu thoả điều kiện ràng buộc. Lấy nhiệt độ $T = T_{\text{ban đầu}}$ khá cao ($T_{\text{ban đầu}} = 10000$, chẳng hạn).

Các bước lặp

Tại mỗi mức nhiệt độ T thực hiện các bước sau:

i) Chọn $X' \in D$ và thuộc một lân cận đủ nhỏ của X .

ii) Xét $\Delta f = f(X') - f(X)$. Nếu $\Delta f < 0$ thì đặt $X := X'$. Nếu trái lại khi $\Delta f > 0$ thì chấp nhận $X := X'$ với xác suất $p = \exp(-\Delta f / (K_b \times T))$, trong đó K_b là hằng số Boltzmann ($K_b = 1,38 \cdot 10^{23}$), T là nhiệt độ hiện thời trong quá trình nguội.

Quy trình i) và ii) lặp lại một số lần L đủ lớn (chẳng hạn $L = 200, 300, \dots$).

Sau đó tính mức nhiệt độ mới theo công thức $T := \alpha T$ ($\alpha \approx 1$, chẳng hạn như $\alpha = 0,95$ hay $0,99 \dots$). Thuật toán dừng khi $T \leq T_{\text{cuối}}$ ($T_{\text{cuối}}$ là giá trị đã chọn trước ≈ 0).

Sau đây là văn bản chương trình *annealing.cpp*:

```
/* Su dung ky thuat annealing – mo phong toi
giai bai toan toi uu toan cuc co rang buoc */
```

```
#include <stdio.h>
#include <stdlib.h>
#include <conio.h>
#include <math.h>
```

```
/* Tinh gia tri ham so can cuc tieu hoa */
float f(float x, float y)
{
float fg = 4*pow(x,2)-2.1*pow(x,4)+pow(x,6)/3;
fg = fg + x*y -4*pow(y,2)+4*pow(y,4);
return fg;
}
```

```

/* Kiem tra cac dieu kien rang buoc */
int constraint(float x,float y)
{
    float fg;
    fg = x + 2.5;
    if (fg<0) goto ZERO;
    fg = 2.5 - x;
    if (fg<0) goto ZERO;
    fg = y + 1.5;
    if (fg<0) goto ZERO;
    fg = 1.5 - y;
    if (fg<0) goto ZERO;
    return 1;
    ZERO: return 0;
}

/* Thu tuc tim diem thay the mo phong qua trinh annealing */
void sa(float x0,float y0,int k,float T,float Tlast,float alfa,float delta)
{
    float x1,y1,deltaf,u,p,ux,uy;
    int l=1; srand(27556);
    printf("\n starting value of the function=%f",f(x0,y0)); getch();
    printf("\n obtained at x=%f,y=%f",x0,y0);getch();
    do
    {
        do
        {
            ux=float(random(10000))/10000-0.5;
            x1=x0-delta*ux;
            uy=float(random(10000))/10000-0.5;
            y1=y0-delta*uy;
            if(constraint(x1,y1)==1)
            {
                deltaf=f(x1,y1)-f(x0,y0);
                if(deltaf<=0) {x0=x1;y0=y1;}
            }
            else
            {
                u=float(random(10000))/10000;

```

```

        p=exp(-deltaf/T);
        if(u<=p) {x0=x1;y0=y1;}
    }
}
l=l+1;
}while(l<=k);
T=T*alfa;
}while(T>Tlast);
printf("\n Optimal value fMin=%f",f(x0,y0)); getch();
printf("\n obtained at x=%f,y=%f",x0,y0);getch();
}
/* Chuong trinh chinh tim diem toi uu */
void main()
{
    clrscr();
    float x0,y0,T,Tlast,k,alfa,delta;
    printf("\n Within the range -2.5 to 2.5 provide value x0=");
    scanf("%f",&x0);
    printf("\n Within the range -1.5 to 1.5 provide value y0=");
    scanf("%f",&y0);
    printf("\n Specify number of iterations at each T level=");
    scanf("%d",&k);
    printf("\n Specify a very high value for Tstart=");
    scanf("%f",&T);
    printf("\n Specify a very low value for Tlast=");
    scanf("%f",&Tlast);
    printf("\n Specify reduction coefficient alfa=");
    scanf("%f",&alfa);
    printf("\n Specify neighbourhood radius delta=");
    scanf("%f",&delta);
    sa(x0,y0,k,T,Tlast,alfa,delta);
    getch();
}

```

Kết quả chạy chương trình máy tính với thuật giải SA là:

Hạt mầm (Seed)	Phương án ban đầu	Phương án tối ưu	Giá trị f_{Min}
27556	(0, 0)	(-0.0898613, 0.7124848)	-1.0316283
19587	(0.1, 0.1)	(0.0898837, -0.7125957)	-1.0316284

(với $\alpha = 0.997$, $\delta = 0.01$, $T_{\text{ban đầu}} = 10000$, $L = 500$, $T_{\text{cuối}} = 0.0001$).

3. Một số vấn đề về mô hình hàng chờ

3.1. Một số yếu tố cơ bản của hệ thống hàng chờ

Như đã biết, trong nhiều hoạt động sản xuất kinh doanh cũng như trong đời sống chúng ta áp dụng các hệ dịch vụ đám đông hay hệ phục vụ công cộng. Chúng có tên gọi chung là hệ thống hàng chờ (*Waiting Line System*). Chẳng hạn các xí nghiệp sửa chữa máy móc, các cửa hàng, các bến xe, bến cảng, trạm tổng đài, các hệ thống điện tử viễn thông, dịch vụ Internet,... là các ví dụ về hệ thống hàng chờ.

Mô hình hàng chờ

Trong các hệ thống hàng chờ thường xuyên diễn ra hai quá trình: quá trình nảy sinh các yêu cầu (một yêu cầu còn được coi là một tín hiệu cần được phục vụ) và quá trình phục vụ các yêu cầu ấy. Song trong quá trình phục vụ của các hệ thống, do nhiều nguyên nhân khác nhau, thường xảy ra các tình trạng sau: Trong nhiều trường hợp, quá trình phục vụ không đáp ứng các yêu cầu và do đó dẫn đến kết quả là nhiều yêu cầu phải chờ để được phục vụ. Ngược lại, trong một số tình huống khác, khả năng phục vụ của hệ thống vượt quá số yêu cầu cần được phục vụ, với kết quả là hệ thống không sử dụng hết phương tiện phục vụ. Vì vậy bài toán đặt ra là:

– Phân tích bản chất của quá trình diễn ra trong các hệ thống hàng chờ và thiết lập các mối liên hệ về lượng giữa các đặc trưng của các quá trình ấy. Điều đó có nghĩa là cần thiết lập hay lựa chọn một mô hình hàng chờ (*Waiting Line Model*) phản ánh được bản chất của hệ thống.

– Trên cơ sở các mối liên hệ đã được xây dựng và các số liệu thu được từ hệ thống, cần tính toán, phân tích và đưa ra các quyết định nhằm tìm ra các giá trị thích hợp cho các tham số điều khiển / thiết kế của hệ thống để thiết kế hay điều khiển các hoạt động của hệ thống hoạt động một cách có hiệu quả hơn.

Các phương pháp giải bài toán mô hình hàng chờ

Để tìm lời giải cho một mô hình hàng chờ người ta thường sử dụng hai phương pháp: phương pháp giải tích và phương pháp mô phỏng trên máy tính. Phương pháp giải tích để giải mô hình hàng chờ gồm các bước sau:

Bước 1: Phân tích hệ thống, chủ yếu là phân tích bản chất của dòng yêu cầu / tín hiệu đến và các trạng thái của hệ thống.

Bước 2: Thiết lập hệ phương trình trạng thái cho các xác suất trạng thái (xác suất để hệ thống ở một trạng thái nào đó tại thời điểm t).

Bước 3: Giải hệ phương trình để tìm các xác suất trạng thái. Từ đó thiết lập các mối quan hệ giữa các chỉ tiêu cần phân tích.

Bước 4: Tính toán, phân tích các chỉ tiêu, trên cơ sở đó đưa ra các nhận xét và các quyết định.

Phương pháp giải tích thường sử dụng các giả thiết rất chặt chẽ của Toán học về các đặc trưng của hệ thống, vì vậy nó có một số hạn chế nhất định khi giải các bài toán thực tế.

Trong khi đó, phương pháp mô phỏng / mô phỏng ngẫu nhiên để giải mô hình hàng chờ được áp dụng cho các bài toán dịch vụ đám đông không giải được bằng công cụ giải tích, nhất là những bài toán liên quan đến hệ thống lớn, bất ổn định, hàm chứa nhiều yếu tố ngẫu nhiên, không tuân theo các giả thiết quá chặt chẽ của Toán học. Trong nhiều trường hợp phương pháp mô phỏng cho ta tiết kiệm được thời gian và chi phí nghiên cứu. Tuy phương pháp mô phỏng chỉ tạo ra các phương án đủ tốt để đánh giá hoạt động của hệ thống chứ không đưa ra được kĩ thuật tìm lời giải tốt nhất, nó tỏ ra rất thành công khi giải quyết nhiều bài toán hàng chờ nảy sinh từ thực tiễn. Các bước cần tiến hành khi áp dụng phương pháp mô phỏng bao gồm:

Bước 1: Xác định bài toán hay hệ thống hàng chờ cần mô phỏng và mô hình mô phỏng.

Bước 2: Đo và thu thập số liệu cần thiết để khảo sát thống kê các số đặc trưng / các yếu tố cơ bản của mô hình.

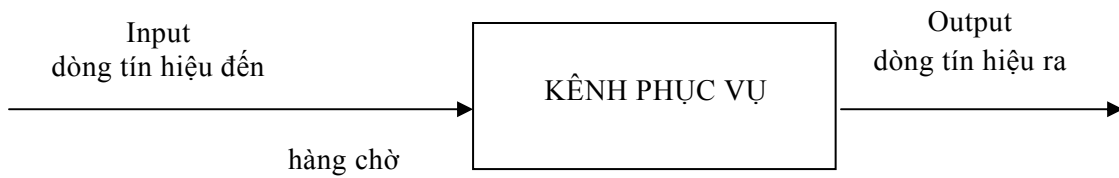
Bước 3: Chạy mô phỏng kiểm chứng (*test simulation*) mô hình và so sánh kết quả kiểm chứng với các kết quả đã biết được trong thực tế. Phân tích kết quả chạy mô phỏng kiểm chứng, nếu cần thì phải sửa lại phương án đã được đánh giá qua chạy mô phỏng.

Bước 4: Chạy mô phỏng để kiểm chứng phương án cuối cùng và kiểm tra tính đúng đắn của mọi kết luận về hệ thống thực tế được rút ra sau khi chạy mô phỏng. Triển khai hoạt động của hệ thống hàng chờ dựa trên phương án tìm được.

Từ những phân tích trên đây có thể thấy Lí thuyết hàng chờ (*Waiting Line Theory*) còn gọi là Lí thuyết hệ phục vụ công cộng hay Lí thuyết hệ dịch vụ đám đông là lĩnh vực rất quan trọng của Toán ứng dụng / Vận trù học. Nhiều bài toán thực tế trong các lĩnh vực hệ thống dịch vụ, kĩ thuật, ... đã được giải quyết thành công nhờ áp dụng phương pháp mô phỏng mô hình hàng chờ.

Các yếu tố cơ bản của hệ thống hàng chờ

Hệ thống hàng chờ tổng quát được minh hoạ như trên hình III.2.



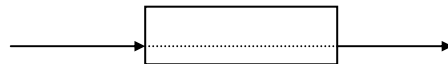
Hình III.2. Hệ thống hàng chờ

Các yếu tố cơ bản của hệ thống hàng chờ bao gồm:

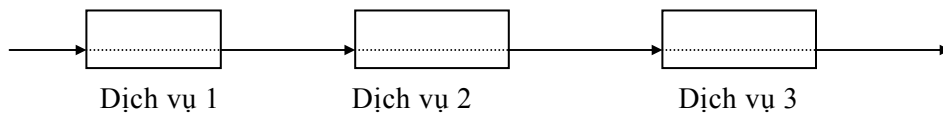
a. Bố trí vật lí của hệ thống

Hệ thống hàng chờ có một số dạng bố trí vật lí (*physical layout*) như minh hoạ trên hình III.3.

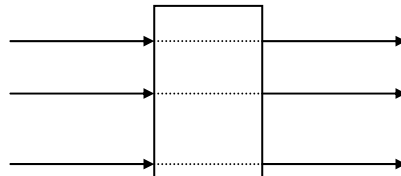
Single Channel – Single Server (Một kênh phục vụ, một loại dịch vụ)



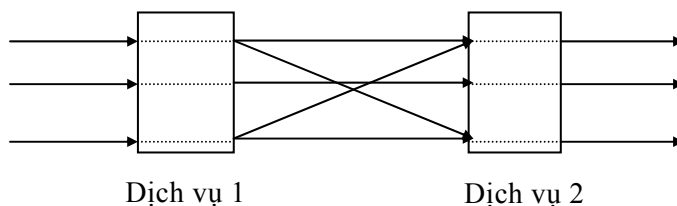
Single Channel – Multi Server (Một kênh phục vụ, nhiều loại dịch vụ)



Multi Channel – Single Server (Nhiều kênh phục vụ, một loại dịch vụ)



Multi Channel – Multi Server (Nhiều kênh phục vụ, nhiều loại dịch vụ)



Hình III.3. Các dạng hệ thống hàng chờ

Trên hình III.3, các kênh phục vụ được hiểu là những thiết bị kĩ thuật hoặc con người hoặc những tổ hợp các thiết bị kĩ thuật và con người được tổ chức quản lí một

cách thích hợp nhằm phục vụ các yêu cầu / các tín hiệu đến hệ thống. Chẳng hạn, ở các trạm điện thoại tự động, kênh phục vụ là các đường dây liên lạc cùng các thiết bị kỹ thuật khác phục vụ cho việc đàm thoại.

b. Nguyên tắc phục vụ

Nguyên tắc phục vụ (hay nội quy) của hệ thống là cách thức nhận các yêu cầu vào các kênh phục vụ. Nguyên tắc phục vụ cho biết trường hợp nào thì yêu cầu được nhận vào phục vụ và cách thức phân bố các yêu cầu vào các kênh như thế nào. Đồng thời nguyên tắc phục vụ cũng cho biết trong trường hợp nào yêu cầu bị từ chối hoặc phải chờ và giới hạn của thời gian chờ.

Một số nguyên tắc phục vụ thường được áp dụng trong các hệ thống hàng chờ là FIFO (*First in first out*), LIFO (*Last in first out*), FCFS (*First come first serve*), có ưu tiên, không ưu tiên,...

c. Các phân phối xác suất của các dòng tín hiệu, dòng phục vụ

Số tín hiệu đến trong một khoảng thời gian cũng như thời gian phục vụ từng tín hiệu nói chung là những biến ngẫu nhiên, và do đó, chúng tuân theo các quy luật phân phối xác suất. Các quy luật phân phối xác suất này được thiết lập căn cứ các số liệu thực nghiệm thu thập từ các quan sát, thí nghiệm, hay từ cơ sở dữ liệu sẵn có.

Đối với dòng tín hiệu đầu vào, thông thường chúng ta giả sử rằng số tín hiệu đến trong vòng một khoảng thời gian nào đó được ấn định trước (1 phút, 3 phút, 5 phút, 30 phút,...) tuân theo luật phân phối Poát-xông $\mathcal{P}(\lambda)$. Ở đây, tham số λ đặc trưng cho số tín hiệu đến (trung bình) trong khoảng thời gian trên. Ví dụ, số khách vào siêu thị (trung bình) là 100 người trong 1 giờ. Có nghĩa là, số khách vào siêu thị là biến ngẫu nhiên X có phân phối Poát-xông với $\lambda = 100$. Hoặc, với số cuộc gọi (trung bình) đến tổng đài trong vòng 1 phút là 3 (tín hiệu) thì có $X \sim \mathcal{P}(3)$.

Một cách chính xác hơn, trong những trường hợp trên, ta có dòng tín hiệu đến là dòng Poát-xông dừng (còn gọi là *dòng tới gián*) với các tính chất trên sau:

– *Tính không hậu quả*: Một dòng tín hiệu có tính không hậu quả nếu xác suất xuất hiện một số tín hiệu nào đó trong một khoảng thời gian nhất định không phụ thuộc vào việc đã có bao nhiêu tín hiệu đã xuất hiện và xuất hiện như thế nào trước khoảng thời gian đó.

– *Tính đơn nhất*: Dòng tín hiệu có tính đơn nhất nếu xét trong khoảng thời gian khá bé thì sự kiện “có nhiều hơn một tín hiệu xuất hiện” hầu như không xảy ra. Về mặt thời gian ta có thể xem dòng tín hiệu có tính đơn nhất nếu thời điểm xuất hiện các tín hiệu không trùng nhau.

– *Tính dừng*: Dòng tín hiệu có tính dừng nếu xác suất xuất hiện một số tín hiệu nào đó trong khoảng thời gian τ chỉ phụ thuộc vào độ dài của τ chứ không phụ thuộc vào điểm khởi đầu của τ .

3.2. Các chỉ số cần khảo sát

Đối với một hệ thống hàng chờ, cần tìm cách để đánh giá được các chỉ số sau:

– A (*Arrival rate*): cường độ dòng tín hiệu đến hay số tín hiệu đến trung bình trong một khoảng thời gian. Ví dụ: A = 6 (6 khách hàng đến trong 2 tiếng); A = 20 (20 cú điện thoại đến tổng đài trong 1 phút).

– S (*Service rate*): cường độ phục vụ hay số tín hiệu trung bình được phục vụ trên một đơn vị thời gian. Ví dụ: S = 7 (hệ thống có thể phục vụ 7 khách trong 1 giờ); S = 25 (tổng đài phục vụ được 25 cú điện thoại trong 2 phút).

– L_q (*Number in queue* hay *Length of queue*): số tín hiệu trung bình trong hàng chờ.

– L_s (*Number in system* hay *Length of system*): số tín hiệu trung bình trong toàn hệ thống (như vậy L_s ≥ L_q).

– W_q (*Waiting time in queue*): thời gian chờ trung bình trong hàng chờ của một tín hiệu.

– W_s (*Waiting time in system*): thời gian chờ trung bình trong hệ thống của một tín hiệu.

– P_w (*Probability the system is busy*): xác suất hệ thống bận (đang hoạt động) hay còn gọi là hệ số (chỉ số) sử dụng của toàn hệ thống (*Utilization factor*).

3.3. Tính toán các chỉ số

Với mục đích tìm hiểu bước đầu, sau đây chúng ta chỉ xét các hệ thống hàng chờ với một loại dịch vụ. Bằng phương pháp giải tích (xem mục 3.1), có thể tìm được công thức tính toán các chỉ số với điều kiện: các giả thiết của mô hình được thỏa mãn.

Mô hình một kênh phục vụ thỏa mãn: số tín hiệu đến có phân phối Poát-xông, thời gian phục vụ có phân phối mũ

Các công thức (I) sau đây đã được chứng minh (bằng phương pháp giải tích):

$$L_q = \frac{A^2}{S(S-A)}; L_s = \frac{A}{S-A}; L_q = \frac{A}{S} \cdot L_s;$$

$$W_q = \frac{A}{S(S-A)}; W_s = \frac{1}{S-A}; P_w = \frac{A}{S}.$$

Chẳng hạn với A = 3, S = 4 thì:

$$L_q = \frac{9}{4(4-3)} = 2,25; L_s = \frac{3}{1} = 3;$$

$$W_q = \frac{3}{4} = 0,75; W_s = 1; P_w = \frac{3}{4} = 0,75.$$

Mô hình một kênh phục vụ thỏa mãn: số tín hiệu đến là phân phối Poát-xông,

thời gian phục vụ có phân phối bất kì.

Các công thức (II) sau đây đã được chứng minh:

$$Lq = \frac{A^2\sigma^2 + (A/S)^2}{2(1-A/S)}; \quad Ls = Lq + \frac{A}{S};$$

$$Wq = \frac{Lq}{A}; \quad Ws = Wq + \frac{1}{S}; \quad Pw = \frac{A}{S}.$$

Trong đó σ độ lệch chuẩn thời gian phục vụ một tín hiệu. Chú ý rằng, nếu thời gian phục vụ tuân theo phân phối mũ thì $\sigma = \frac{1}{S}$ và cũng là thời gian trung bình phục vụ một tín hiệu. Có thể nhắc lại rằng:

$$m = \int_0^{\infty} tf(t)dt = \frac{1}{S} \quad \text{và} \quad \sigma = \sqrt{\int_0^{\infty} (t-m)^2 f(t)dt} = \frac{1}{S}.$$

Lúc này các công thức (II) trở về (I).

Mô hình một kênh phục vụ thoả mãn: số tín hiệu đến có phân phối Poát-xông, thời gian phục vụ có phân phối mũ, hàng chờ có giới hạn số tín hiệu tối đa M

Các công thức (III) sau đây đã được chứng minh:

$$P_0 = \frac{1-A/S}{1-(A/S)^{M+1}}; \quad Pw = 1 - P_0;$$

với P_0 là xác suất không có tín hiệu nào trong hệ thống (hệ số không sử dụng);

$P_M = (A/S)^M P_0$ là tỉ lệ % tín hiệu không được phục vụ do hệ thống “đầy”;

$$Ls = \frac{Pw - M(A/S)P_M}{1-(A/S)}; \quad Lq = Ls - \frac{A(1-P_M)}{S};$$

$$Ws = \frac{Ls}{A(1-P_M)}; \quad Wq = Ws - \frac{1}{S}.$$

Chú ý rằng nếu $M = +\infty$ thì (III) trở về (I).

Mô hình nhiều kênh phục vụ thoả mãn: tín hiệu đến có phân phối Poát-xông, thời gian phục vụ là phân phối mũ.

P_0 – xác suất tất cả các kênh phục vụ đều không có tín hiệu, tìm được bằng cách tra phụ lục 3 dựa trên tỉ số A/kS (k số kênh phục vụ) hoặc tính trực tiếp từ công thức sau:

$$P_0 = \frac{1}{\sum_{n=0}^{k-1} \frac{1}{n!} \left(\frac{A}{S}\right)^n + \frac{1}{k!} \left(\frac{A}{S}\right)^k \frac{kA}{kA-S}} \text{ với } kS > A;$$

$$P_w = \frac{1}{k!} \left(\frac{A}{S}\right)^k \frac{kS}{kS-A} P_0;$$

$$L_s = \frac{AS(A/S)^k}{(k-1)!(kS-A)^2} P_0 + \frac{A}{S}; \quad L_q = L_s - \frac{A}{S}; \quad W_s = \frac{L_s}{A}; \quad W_q = \frac{L_q}{A}.$$

Một số điểm hạn chế của các mô hình hàng chờ

Các mô hình hàng chờ giới thiệu ở trên là những mô hình tiện lợi nhất được áp dụng khá rộng rãi. Tuy nhiên, do các mô hình này công nhận các giả thiết “quá chặt chẽ” ít xảy ra trên thực tế, nên các chuyên gia trong lĩnh vực Toán ứng dụng/Vận trù học/Khoa học quản lý cũng đã đề xuất xem xét nhiều mô hình khác. Đó là các mô hình với các giả thiết như: số tín hiệu cần phục vụ là hữu hạn, dòng tín hiệu đến không phải kiểu Poát-xông, cường độ phục vụ phụ thuộc vào số tín hiệu trong hàng chờ ... và việc giải quyết những mô hình như vậy cần tới sự trợ giúp của phương pháp mô phỏng ngẫu nhiên.

Ngay cả khi các giả thiết khá chặt chẽ của bốn mô hình đã nêu trong mục này (cũng như một số mô hình tương tự khác) là hợp lí, thì việc các mô hình hàng chờ đưa ra các lời giải với trạng thái vững (*steady state solutions*) cũng ít có ý nghĩa thực tế. Trong nhiều ứng dụng thực tiễn, các hệ thống hàng chờ không bao giờ đạt tới các trạng thái vững. Chẳng hạn, trong một hệ thống hàng chờ, cường độ tín hiệu đến trung bình thay đổi nhiều lần trong ngày không cho phép hệ thống đạt được trạng thái vững.

Do đó, để giải quyết nhiều bài toán hàng chờ trong lĩnh vực dịch vụ đám đông và các lĩnh vực khác, cần áp dụng phương pháp mô phỏng để tìm ra các lời giải có tính thực tiễn cho các mô hình hàng chờ khi hệ thống không thể đạt tới trạng thái vững hoặc khi không có các mô hình lí thuyết thích hợp.

3.4. Áp dụng mô phỏng cho một số hệ thống hàng chờ

Ví dụ 1: Bài toán hệ dịch vụ hàng chờ ba kênh với dòng tới gián có từ chối.

Cho biết: dòng tín hiệu đến là dòng Poát-xông dừng (còn gọi là *dòng tới gián*). Giãn cách thời gian giữa thời điểm đến của hai nhu cầu (tín hiệu) liên tiếp có phân phối mũ với tham số $\mu = 5$, tức là có hàm mật độ $f(t) = 5e^{-5t}$. Nếu tín hiệu xuất hiện mà có ít nhất một trong ba kênh không bận (kênh số 1 hoặc kênh số 2 hoặc kênh số 3 không bận) thì tín hiệu được phục vụ tại kênh không bận với số thứ tự nhỏ nhất; nếu trái lại (khi cả ba kênh đều bận) thì tín hiệu bị từ chối. Biết thời gian phục vụ mỗi nhu cầu là 0,5 phút, hãy xác định kì vọng toán số nhu cầu được phục vụ trong khoảng thời gian 4 phút.

Như vậy, cần áp dụng mô hình hàng chờ *MultiChannel – SingleServer System* (Hệ thống nhiều kênh phục vụ – một loại dịch vụ) theo quy tắc *First in first out* (FIFO: Tín hiệu đến trước được phục vụ xong trước). Thời gian giữa hai tín hiệu liên tiếp có phân phối mũ với hàm mật độ xác suất $f(t) = 5e^{-5t}$. Trong bài toán này (nhằm đơn giản các bước tính toán) thời gian phục vụ mỗi tín hiệu được coi là không đổi và bằng 0,5 phút.

Chúng ta sẽ áp dụng mô phỏng để xác định số nhu cầu trung bình cần được phục vụ trong khoảng thời gian 4 phút như trình bày sau đây.

Kí hiệu T_i là thời điểm đến của tín hiệu thứ i , T_{ki} là thời điểm kết thúc dịch vụ của tín hiệu thứ i (nếu có), tại kênh thứ k ($k = 1, 2, 3$). Thời điểm đến của nhu cầu tiếp theo là $T_i = T_{i-1} + \tau_i$ với τ tuân theo luật phân phối mũ có hàm mật độ $f(t) = 5e^{-5t}$ và hàm phân phối là $F(t) = 1 - e^{-5t} = P(\tau \leq t)$.

Lúc đó $T_1 = 0$, $T_{11} = T_1 + 0,5$. Kết quả này cho biết thời điểm đến của tín hiệu thứ nhất là $T_1 = 0$ và được kênh 1 phục vụ. Kết thúc phục vụ tín hiệu 1 là thời điểm $T_{11} = T_1 + 0,5 = 0,5$. Máy đếm ghi nhận 1 đơn vị là số tín hiệu đã được phục vụ.

Để tìm T_2 theo công thức $T_2 = T_1 + \tau_2$, ta phát sinh τ_2 theo cách đã biết ở mục 3.1.3: Trước hết, phát sinh số ngẫu nhiên r_2 có 2 chữ số sau dấu phẩy $0 \leq r_i \leq 1$ (theo bảng số ngẫu nhiên – phụ lục 2B) ta có $r_2 = 0,10$. Sau đó tính $\tau_2 = -\frac{1}{5} \ln r_2$ và $T_2 = T_1$

$-\frac{1}{5} \ln r_2 = 0 - 0,2 \ln 0,1 = 0,46$. Vậy tín hiệu tiếp theo phải vào kênh 2 vì kênh 1 còn đang bận. Máy đếm ghi thêm 1 đơn vị thời điểm kết thúc phục vụ tín hiệu 2 là $T_{22} = T_2 + 0,5 = 0,46 + 0,5 = 0,96$.

Tiếp tục phát sinh $r_3 = 0,09$, ta có $\tau_3 = -0,2 \ln 0,09 = 0,482$. Do đó thời điểm đến của tín hiệu 3 là $T_3 = T_2 + \tau_3 = 0,46 + 0,482 = 0,942$. Lúc này kênh 1 đã được giải phóng do đã phục vụ xong tín hiệu 1, nên tín hiệu 3 được tiếp nhận vào kênh 1. Tại thời điểm kết thúc phục vụ tín hiệu 3 là $T_{13} = T_3 + 0,5 = 0,942 + 0,5 = 1,442$ máy đếm lại ghi tiếp 1 đơn vị. Thực hiện tính toán tương tự, kết quả tổng hợp được ghi trong bảng III.6.

Bảng III.6. Tính toán mô phỏng tìm số nhu cầu được phục vụ

Thứ tự tín hiệu	Số ngẫu nhiên r_i	$-\ln r_i$	$\tau_i =$ $-1/5 \ln r_i$	Thời điểm đến T_i	Thời điểm T_{ki} kết thúc phục vụ tại kênh k			Đếm số tín hiệu	
					1	2	3	nhận	bỏ
1				0	0,5			1	
2	0,10	2,30	0,46	0,46		0,96		1	
3	0,09	2,44	0,482	0,942	1,442			1	
4	0,73	0,32	0,064	1,006		1,506		1	
5	0,25	1,39	0,278	1,284			1,784	1	
6	0,33	1,11	0,222	1,506	2,006			1	
7	0,76	0,27	0,054	1,560		2,060		1	
8	0,52	0,65	0,13	1,690					1
9	0,01	4,6	0,92	2,61	3,11			1	
10	0,35	1,05	0,21	2,82		3,32		1	
11	0,86	0,15	0,03	2,85			3,35	1	
12	0,34	1,08	0,216	3,066					1
13	0,67	0,40	0,08	3,146	3,646			1	
14	0,35	1,05	0,21	3,356		3,856		1	
15	0,48	0,73	0,146	3,502			4,022	1	
16	0,76	0,27	0,054	3,556					1
17	0,80	0,22	0,044	3,600					1
18	0,95	0,05	0,01	3,61					1
19	0,9	0,10	0,02	3,63					1
20	0,91	0,09	0,018	3,648	4,148			1	
21	0,17	1,77	0,354	4,002				1	
								14	6

Phân tích kết quả tính toán ta thấy trong 20 nhu cầu đến thì chỉ có 14 nhu cầu được phục vụ. Tính toán tương tự 6 lần nữa ta có kết quả:

x_1	x_2	x_3	x_4	x_5	x_6
14	15	14	12	13	15

Vậy số nhu cầu trung bình được hệ phục vụ trong vòng 4 phút vào khoảng

$$\bar{x} = (14 + 15 + 14 + 12 + 13 + 15)/6 = 13,83.$$

Giải bài toán dịch vụ ba kênh có từ chối trên máy tính

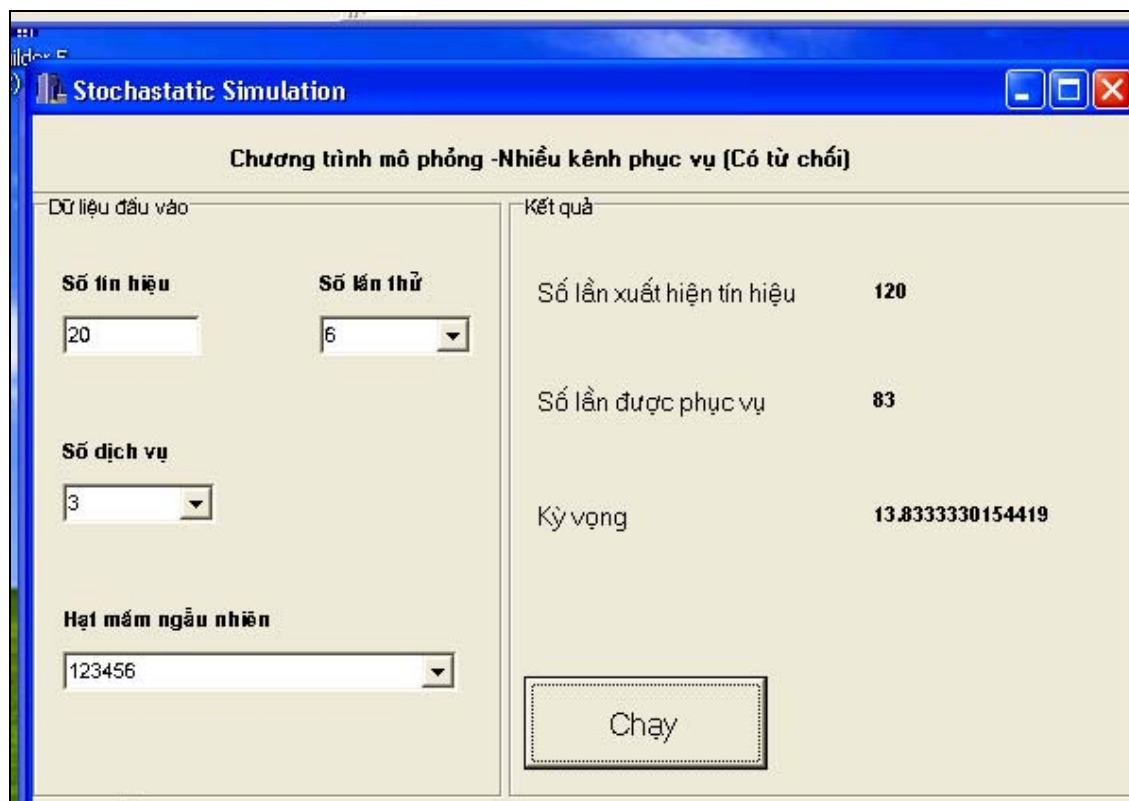
Một phần mềm máy tính với tên gọi MOPHONG1 phiên bản 1.0 đã được thiết kế dựa trên ngôn ngữ Builder C++ 5.0 để giải bài toán một dịch vụ với nhiều kênh phục vụ

có từ chối như đã trình bày trong ví dụ 1 với các điều kiện sau:

- Số kênh phục vụ n có thể lấy giá trị từ 2 tới 10.
- Dòng tín hiệu đến là dòng Poát-xông, thời gian dẫn cách giữa hai tín hiệu liên tiếp tuân theo phân phối mũ với hàm mật độ xác suất là $f(t) = 5e^{-5t}$.
- Thời gian trung bình phục vụ (xử lí) một tín hiệu là 0,5 phút.
- Tính số tín hiệu được phục vụ trong số m tín hiệu đến (m có thể lấy giá trị từ 10 tới 100).
- Thực hiện k lần mô phỏng (k lần thử, k có thể lấy giá trị từ 4 tới 10).

Mục đích của việc xây dựng phần mềm này nhằm vào: phục vụ dạy và học môn Mô phỏng ngẫu nhiên, cũng như tiếp tục nâng cấp phiên bản 1.0 để mô phỏng được hệ nhiều kênh phục vụ – nhiều loại dịch vụ có từ chối trong trường hợp tổng quát khi dòng tín hiệu đến và thời gian phục vụ tín hiệu có phân phối bất kì.

Để chạy phần mềm MOPHONG1 chúng ta cần cài đặt Builder C++ 5.0 vào máy tính. Sau khi kích chuột vào biểu tượng của phần mềm, chọn *File > Open Project > Look in > Mophong1 (Thư mục lưu trữ phần mềm) > Mp1.bpr*. Sau đó chọn *Run* trên thanh công cụ để chạy phần mềm (xem hình III.4).



Hình III.4. Chạy phần mềm MOPHONG 1

Chúng ta nhập số kênh phục vụ $n = 3$ vào ô số (kênh) dịch vụ, số tín hiệu phát

sinh $m = 20$ vào ô số tín hiệu, số lần mô phỏng $k = 6$ vào ô số lần thử. Ngoài ra ta phải chọn hạt mầm là một số nguyên (đủ lớn) nhằm khởi tạo hàm sinh số ngẫu nhiên có phân phối đều trong $[0, 1)$, chẳng hạn số 123456 để ghi vào ô hạt mầm ngẫu nhiên, nhằm từ đó mô phỏng dòng Poát-xông các tín hiệu đến. Sau đó chúng ta kích chuột vào nút *Chạy* để chạy chương trình.

Kết quả ta thấy tổng cộng số tín hiệu đến là 120, trong đó có 83 tín hiệu được phục vụ (do đó có 37 tín hiệu bị từ chối). Vậy trung bình trong 20 tín hiệu đến có 13,833 tín hiệu được phục vụ. Kết quả này so với kết quả tính toán trên giấy cho ví dụ trên là khá sát nhau.

Chú ý: Việc tính toán dựa vào bảng số ngẫu nhiên cho trong phụ lục hoặc dựa vào hàm sinh số ngẫu nhiên trong máy tính không cho kết quả hoàn toàn giống nhau trong các lần chạy mô phỏng khác nhau. Điều này xảy ra là vì các bộ số ngẫu nhiên tạo được không giống nhau. Với các hạt mầm khác nhau thì hàm sinh số ngẫu nhiên cũng cho các bộ số ngẫu nhiên khác nhau và do đó các kết quả cuối cùng cũng không trùng nhau. Muốn kết quả mô phỏng ổn định hơn cần chọn số tín hiệu đến m đủ lớn.

Ngoài ra, chúng ta có thể chọn số kênh phục vụ tùy ý, chẳng hạn $n = 5$. Kết quả chạy phần mềm với hạt mầm 123456 cho biết trong số 20 tín hiệu đến trung bình có 18,167 tín hiệu được phục vụ.

Ví dụ 2: Một công ti điều hành một kho nguyên liệu để cấp phát cho các đốc công của 10 phân xưởng. Hiện tại, hai nhân viên phục vụ đang được công ti giao cho nhiệm vụ cấp phát nguyên liệu. Bộ phận quản lí công ti muốn cân nhắc liệu có nên thêm một nhân viên phục vụ nữa hay không.

Rõ ràng rằng số đốc công (tín hiệu cần phục vụ) là hữu hạn, phân phối của số tín hiệu đến trong một đơn vị thời gian cũng không theo kiểu Poát-xông và thời gian phục vụ tín hiệu cũng không tuân theo luật phân phối mũ. Do đó, không thể tìm ra được lời giải giải tích thông qua một mô hình nhiều kênh phục vụ với các giả thiết như vậy. Phương pháp “duy nhất” là tìm cách áp dụng mô phỏng.

Số liệu thu thập được

– Quan sát trong vòng một tháng vào các ngày làm việc, mỗi ngày một giờ vào các thời điểm ngẫu nhiên, các số liệu về thời gian phục vụ một tín hiệu và tần suất tương ứng đã được thu thập (bảng III.7).

– Tính thời gian phục vụ trung bình cho một tín hiệu (tính kì vọng):

$$8 \times 0,1 + 9 \times 0,2 + 10 \times 0,3 + 11 \times 0,4 = 10 \text{ (phút).}$$

– Ngoài ra cũng đã khảo sát được: giãn cách thời gian trung bình giữa hai tín hiệu liên tiếp là 5 phút và số lượng tín hiệu trung bình đến trong một khoảng 5 phút là một tín hiệu.

Bảng III.7. Thời gian phục vụ một tín hiệu và tần suất

Thời gian phục vụ một tín hiệu (phút)	Số lượng	Tần suất/ Xác suất thực nghiệm
8	15	0,1
9	30	0,2
10	45	0,3
11	60	0,4
	$\Sigma = 150$	$\Sigma = 1$

Cần chú ý rằng, dòng tín hiệu đến chưa chắc tuân theo phân phối Poát-xông và thời gian phục vụ một tín hiệu không nhất thiết tuân theo phân phối mũ. Do đó, không áp dụng được công thức của mục 3.3 mà phải dùng mô phỏng để giải quyết vấn đề: cần bố trí bao nhiêu kênh phục vụ (nhân viên phục vụ) trong kho cấp phát nguyên liệu là hợp lý nhất? Như vậy mô phỏng có khả năng xử lý các tình huống, sự kiện như trong thực tế xảy ra chứ không bắt chúng tuân theo các phân phối xác suất nhất định hay theo các hành vi gò ép.

Mô phỏng hệ thống hàng chờ

– Mô phỏng tín hiệu đến: trung bình 5 phút có 1 tín hiệu đến. Chúng ta dùng 24 số ngẫu nhiên sau lấy ra từ bảng số ngẫu nhiên (phụ lục 2A), mỗi số gồm 10 chữ số để mô phỏng 24 khoảng 5 phút (như vậy tổng cộng là 120 phút, mỗi số dùng để mô phỏng một khoảng 5 phút từ 9h – 11h): 1581922396, 2068577984, 8262130892, 8374856049, 4637567488, 0928105582, 7295088579, 9586111652, 7055508767, 6472382984, 4112077556, 3440672486,..., 5973470495. Nếu chữ số 7 xuất hiện trong số 10 chữ số đã chọn, ta coi như 1 tín hiệu đến trong khoảng thời gian tương ứng (vì trung bình trong 5 phút có 1 tín hiệu đến cũng giống như trung bình trong một số có 10 chữ số có một chữ số 7). Chẳng hạn trong khoảng 5 phút đầu không có tín hiệu nào, khoảng 5 phút thứ hai có 2 tín hiệu đến.

Ta thấy số tín hiệu đến chỉ có thể là 0, 1, 2, 3, 4 tín hiệu. Thời điểm đến của tín hiệu trong mỗi khoảng 5 phút được quy định như sau tùy theo số tín hiệu đến trong khoảng đó (bảng III.8). Chẳng hạn, nếu có hai tín hiệu đến, thì thời điểm đến là vào đầu phút thứ nhất và đầu phút thứ ba.

Bảng III.8. Quy định thời điểm đến của tín hiệu

Phút	1	2	3	4	5
Nếu 1 tín hiệu đến	*				
Nếu 2 tín hiệu	*		*		
Nếu 3 tín hiệu	*		*		*
Nếu 4 tín hiệu		*	*	*	*

– Mô phỏng thời gian phục vụ X một tín hiệu: Ta mô phỏng phân phối xác suất ở bảng III.7 theo cách đã biết. Trước hết lấy ba số ngẫu nhiên có 10 chữ số: 9846413446, 8306646692, 0661684251 (hàng 4 từ dưới lên, phụ lục 2A). X = 8 phút nếu xuất hiện chữ số 0; X = 9 phút nếu xuất hiện chữ số 1, 2; X = 10 phút nếu xuất hiện chữ số 3, 4, 5; X = 11 phút nếu xuất hiện chữ số 6, 7, 8 hoặc 9. Bảng III.9 tổng hợp kết quả mô phỏng số tín hiệu đến và thời gian phục vụ các tín hiệu.

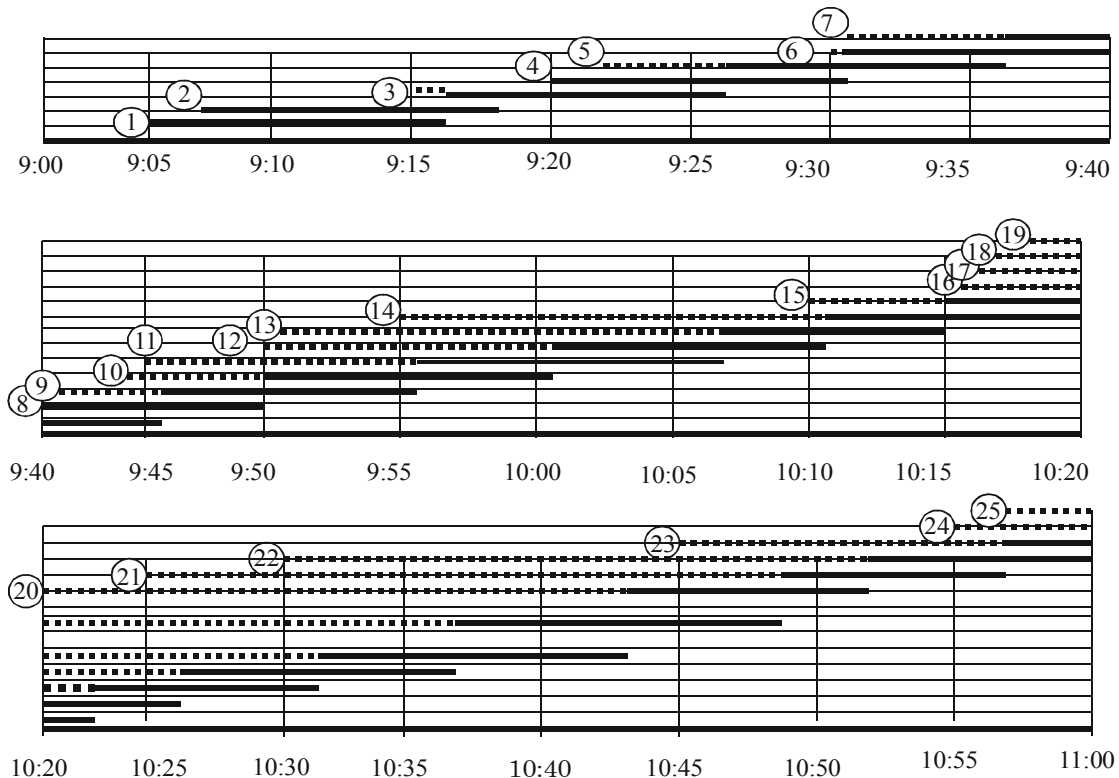
Bảng III. 9. Kết quả mô phỏng số tín hiệu đến và thời gian phục vụ tín hiệu

Chu kì	Số ngẫu nhiên 10 chữ số										Số tín hiệu đến	Thời gian phục vụ
1	1	5	8	1	9	2	2	3	9	6	0	
2	2	0	6	8	5	7	7	9	8	4	2	11, 11
3	8	2	6	2	1	3	0	8	9	2	0	
4	8	3	7	4	8	5	6	0	4	9	1	10
5	4	6	3	7	5	6	7	4	8	8	2	11, 10
6	0	9	2	8	1	0	5	5	8	2	0	
7	7	2	9	5	0	8	8	5	7	9	2	9, 10
8	9	5	8	6	1	1	1	6	5	2	0	
9	7	0	5	5	5	0	8	7	6	7	3	10, 10, 11
10	6	4	7	2	3	8	2	9	3	4	1	11
11	4	1	1	2	0	7	7	5	5	6	2	10, 8
12	3	4	4	0	6	7	2	4	8	6	1	11
13	1	8	8	2	4	1	2	9	6	3	0	
14	0	6	8	4	0	1	2	0	0	6	0	
15	0	9	3	3	1	4	7	9	1	4	1	11
16	7	4	5	7	4	7	7	4	6	8	4	10, 11, 11, 11
17	5	4	3	5	8	8	0	7	8	8	1	9
18	9	6	7	0	8	5	2	9	1	3	1	8
19	1	2	9	1	2	6	5	7	3	0	1	11
20	4	8	9	0	0	3	1	3	0	5	0	
21	0	0	9	9	5	2	0	8	5	8	0	
22	3	0	9	0	9	0	8	8	7	2	1	11
23	2	0	3	9	5	9	3	1	8	1	0	
24	5	9	7	3	4	7	0	4	9	5	2	9,11

Để tiến hành minh họa quá trình tính toán mô phỏng cho số tín hiệu đến trong từng khoảng thời gian 5 phút và thời gian phục vụ mỗi tín hiệu, chúng ta quy ước các kí hiệu sau:

○ Tín hiệu đến Chờ ————— Được phục vụ

Hình III.5 tổng hợp kết quả tính toán mô phỏng cho hệ thống chờ hai kênh phục vụ – một dịch vụ.



Hình III.5. Tổng hợp kết quả mô phỏng hệ thống chờ

Theo hình III.5 có thể thấy, trong thời gian 120 phút có 25 tín hiệu đến (25 số 7 xuất hiện ở các chu kì 2, 4, 5, 7, 9, 10, 11, 12, 15, 16, 17, 18, 19, 22, 24). Tổng thời gian đợi 213 phút, vì vậy thời gian đợi trung bình = $\frac{213}{25} = 8,52$ phút.

Chúng ta phân tích ý nghĩa kinh tế của tính toán mô phỏng như sau: Mô hình hàng chờ trên có hai kênh phục vụ (hai nhân viên phục vụ). Giả sử rằng lương 7\$/1 giờ/1 nhân viên phục vụ, lương của mỗi đốc công là 12\$/1 giờ/1 đốc công. Theo dữ kiện của bài toán, thời gian trung bình giữa hai lần tín hiệu đến liên tiếp là 5 phút. Như vậy, trong trong 8 giờ có 96 lần đi, và phải đợi 8,52 phút mỗi lần. Do đó, tổng hao hụt / mỗi ngày làm việc do thời gian đốc công lãng phí do phải đợi là: $(8,52 \times 12\$ \times 96) / 60 = 163,56\$$. Ngoài ra, tổng chi lương / ngày cho hai nhân viên phục vụ là: $7\$ \times 8 \times 2$

người = 112\$. Từ đó, tổng chi phí / ngày cho hệ thống hàng chờ trên là \$ 275,56.

Tương tự, nếu bố trí ba kênh phục vụ (ba nhân viên phục vụ) thì có thể tính được thời gian trung bình chờ đợi là 1,88 phút và tổng chi phí cho hệ thống ba kênh là 204\$. Nếu dùng bốn kênh phục vụ thì thời gian chờ đợi trung bình là 0 phút, tổng chi phí là 224\$. Vậy hệ thống hàng chờ trên nên dùng ba kênh phục vụ là tốt nhất. Nói cách khác, ban điều hành công ti nên điều thêm một nhân viên phục vụ tới làm việc tại kho cấp phát nguyên liệu nhằm giảm bớt chi phí “cơ hội”.

So sánh với lời giải lí thuyết

So sánh lời giải dùng mô phỏng với lời giải dựa trên lí thuyết mô hình hàng chờ (lúc đó cần giả sử rằng: số tín hiệu đến tuân theo luật Poát-xông, thời gian phục vụ một tín hiệu tuân theo phân phối mũ).

Khi tra phụ lục 3 tìm P_0 dựa vào tỉ số $A/(kS)$ với $k = 3$ là số kênh phục vụ, ta được

$$P_0 = 0,11435 \quad (A = 12, S = 6; A/(kS) = 0,66);$$

$$L_s = \frac{AS(A/S)^k}{(k-1)!(kS-A)^2} P_0 + \frac{A}{S} = 2,915; \quad L_q = L_s - \frac{A}{S} = 0,915;$$

$$W_q = \frac{L_q}{A} = 0,076 \text{ (giờ)} = 4,5 \text{ phút.}$$

Trong khi đó theo kết quả mô phỏng thì thời gian chờ trung bình là 1,88 phút. Tuy nhiên, nếu chạy mô phỏng cho thời gian dài hơn (10 tới 15 giờ) thông qua chương trình trên máy tính thì có kết quả thời gian chờ trung bình là 2,63 phút tương đối ổn định.

Từ các phân tích trên có thể thấy, trong các tình huống chúng ta không xử lí được các điều kiện toán học phức tạp ẩn chứa trong bài toán hàng chờ, lựa chọn duy nhất là áp dụng mô phỏng để giải bài toán.

Phần mềm mô phỏng hệ thống hàng chờ

Phần mềm máy tính với tên gọi MOPHONG2 phiên bản 1.0 đã được thiết kế dựa trên ngôn ngữ Builder C++ 5.0 để mô phỏng hệ thống hàng chờ như đã trình bày trong ví dụ 2 với các điều kiện sau:

– Số kênh phục vụ n có thể lấy giá trị từ 2 tới 10.

– Dòng tín hiệu đến theo quy luật: cứ trung bình 5 phút có một tín hiệu đến, nhưng chưa biết đây có phải là phân phối Poát-xông hay không. Ngoài ra, thời điểm đến của các tín hiệu trong vòng 5 phút cũng chưa biết rõ, nên chúng được mô phỏng căn cứ bảng III.8.

– Thời gian phục vụ (xử lý) một tín hiệu tuân theo phân phối rời rạc đã biết:

Thời gian xử lý một tín hiệu (phút)	8	9	10	11
Xác suất	0,1	0,2	0,3	0,4

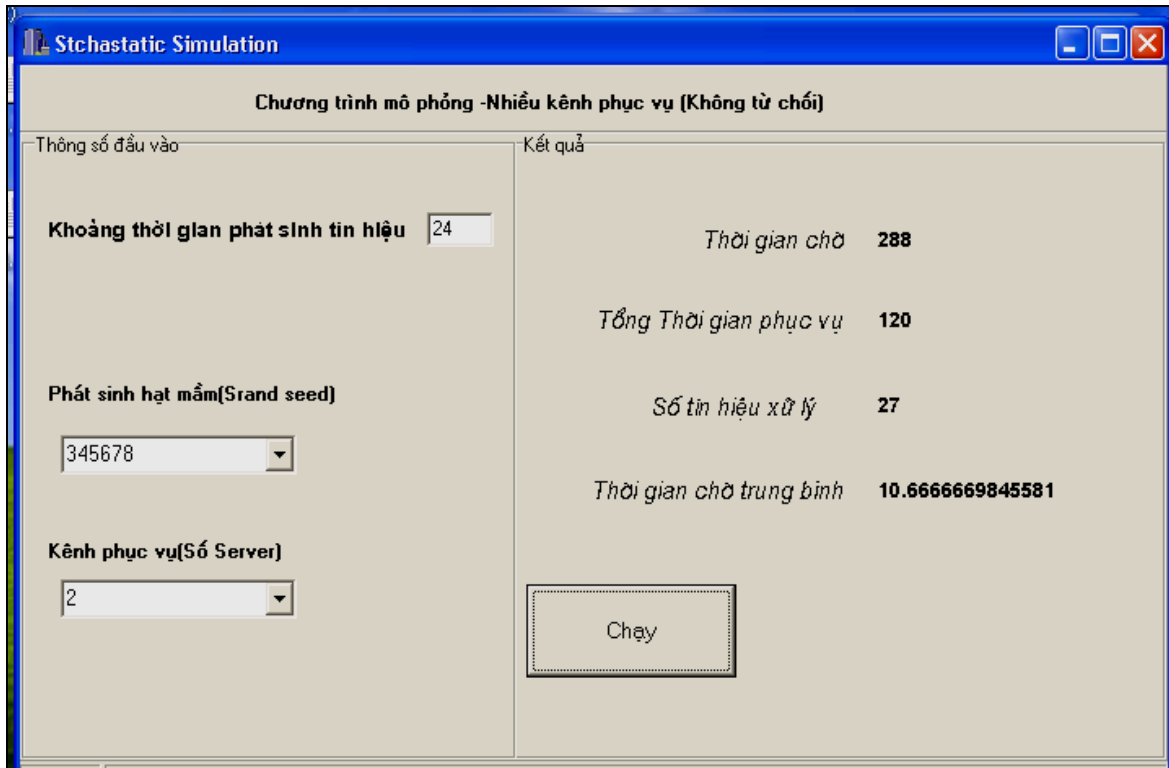
– Khoảng thời gian mô phỏng có thể chọn k chu kì, mỗi chu kì 5 phút (k có thể chọn giá trị 12, 24, 32, 48, ...)

– Cần tính thời gian chờ trung bình (phút) cho mỗi tín hiệu đi đến hệ thống dịch vụ trên đây.

Mục đích của việc xây dựng phần mềm này là phục vụ dạy và học môn Mô phỏng ngẫu nhiên, cũng như tiếp tục nâng cấp phiên bản 1.0 để mô phỏng được hệ nhiều kênh phục vụ – nhiều loại dịch vụ không có từ chối trong trường hợp tổng quát khi dòng tín hiệu đến và thời gian phục vụ tín hiệu có phân phối bất kì.

Để chạy phần mềm MOPHONG2, cũng tương tự như chạy phần mềm MOPHONG1 chúng ta cần cài đặt Builder C++ 5.0 vào máy tính. Sau khi kích chuột vào biểu tượng của phần mềm, chọn *File > Open Project > Look in > Mophong2 (Thư mục lưu trữ phần mềm) > Mp2.bpr*. Sau đó chọn *Run* trên thanh công cụ để chạy phần mềm (xem hình III.6).

Chúng ta nhập số kênh phục vụ $n = 3$ vào ô số (kênh) dịch vụ, khoảng thời gian mô phỏng với $m = 24$ chu kì 5 phút vào ô khoảng thời gian phát sinh tín hiệu. Ngoài ra, ta phải chọn hạt mầm là một số nguyên (đủ lớn) nhằm khởi tạo hàm sinh số ngẫu nhiên có phân phối đều trong $[0, 1)$, chẳng hạn số 345678 để ghi vào ô hạt mầm ngẫu nhiên. Sau đó chúng ta kích chuột vào nút *Chạy* để chạy chương trình.



Hình III.6. Chạy phần mềm MOPHONG2

Kết quả ta thấy tổng cộng có 27 tín hiệu đi tới hệ thống dịch vụ, tổng thời gian chờ là 288 phút, nên trung bình một tín hiệu phải chờ là 10,667 phút. Kết quả này so với kết quả tính toán trên giấy sử dụng bảng số ngẫu nhiên (phụ lục 2A) là khá sát nhau.

Cần nhắc lại rằng việc tính toán dựa vào bảng số ngẫu nhiên cho trong phụ lục 2A và dựa vào hàm sinh số ngẫu nhiên trong máy tính không cho kết quả hoàn toàn giống nhau do các bộ số ngẫu nhiên tạo được không giống nhau. Với các hạt mầm khác nhau thì hàm sinh số ngẫu nhiên cũng cho các bộ số ngẫu nhiên khác nhau và do đó các kết quả cuối cùng cũng không trùng nhau. Chẳng hạn, nếu thay hạt mầm 345678 (cho kết quả thời gian chờ trung bình 10,667 phút) bằng hạt mầm 456789 thì thời gian chờ trung bình sẽ chỉ là 4,158 phút. Nếu lấy trung bình của hai kết quả này thì được con số sát hơn với kết quả tính toán đã biết dựa trên phụ lục 2A.

Từ các phân tích trên ta thấy, muốn kết quả mô phỏng ổn định hơn cần chọn số chu kỳ mô phỏng m đủ lớn và chạy cho nhiều hạt mầm khác nhau để tính giá trị trung bình của kết quả các lần chạy đó. Vấn đề này thực chất là vấn đề *bổ trí thí nghiệm mô phỏng* trên máy tính, cũng là một hướng trong nghiên cứu về mô phỏng ngẫu nhiên hiện nay đang được tiến hành trong nhiều lĩnh vực.

Chương IV

PHÂN TÍCH MARKOV VÀ ỨNG DỤNG

1. Các khái niệm cơ bản về xích Markov

1.1. Một số định nghĩa

Nhiều mô hình ngẫu nhiên trong Vận trù học, Kinh tế, Kỹ thuật, Dân số học, Di truyền học,... dựa trên cơ sở là quá trình Markov. Đặc biệt, hiện tại một lĩnh vực mới về Tin – Sinh học (*Bioinformatics*) chuyên nghiên cứu về gene ứng dụng rất mạnh các vấn đề của lý thuyết các quá trình Markov. Trong ngành Cơ điện hiện nay nhiều chuyên gia lý thuyết và thực hành cũng rất quan tâm tới quá trình Markov nói chung, cũng như các quá trình sinh–tử hay quá trình hồi phục nói riêng.

Ví dụ: Xét một hệ thống vật lý tiến triển theo thời gian. Tại thời điểm $t = 0$, hệ thống có thể rơi vào một trong ba trạng thái (hay vị trí) 1, 2 hoặc 3 một cách ngẫu nhiên. Kí hiệu $X(0)$ là vị trí của hệ thống tại thời điểm $t = 0$, thì $X(0)$ là một biến ngẫu nhiên, có thể nhận các giá trị 1 hoặc 2 hoặc 3 với các xác suất nhất định. Giả sử rằng căn cứ vào các kết quả quan sát hay nghiên cứu, chúng ta có bảng phân phối xác suất sau cho $X(0)$:

Các giá trị của $X(0)$	1	2	3
Xác suất tương ứng	0,2	0,5	0,3

Tại các thời điểm tiếp theo, chẳng hạn, $t = 1, 2, 3, \dots$ vị trí của hệ thống sẽ được mô tả bởi các biến ngẫu nhiên $X(1), X(2), X(3), \dots$ với các bảng phân phối xác suất tương ứng. Dựa trên ví dụ này, chúng ta xét định nghĩa sau về quá trình ngẫu nhiên.

Định nghĩa 1

Xét một hệ thống vật lý (hay một hệ thống sinh thái, hệ thống dịch vụ,...) tiến triển theo thời gian. Gọi $X(t)$ là vị trí (tình trạng) của hệ tại thời điểm t . Như vậy ứng với mỗi thời điểm t , $X(t)$ chính là một biến ngẫu nhiên mô tả vị trí (tình trạng) của hệ thống. Quá trình $\{X(t)\}_{t \geq 0}$ được gọi là một *quá trình ngẫu nhiên*.

Tập hợp các vị trí có thể có của hệ gọi là *không gian trạng thái*. Không gian trạng thái được kí hiệu là S . Trong ví dụ trên, nếu giả sử rằng $X(t)$ chỉ có thể nhận một trong ba giá trị 1, 2, 3 $\forall t$, thì $S = \{1, 2, 3\}$.

Giả sử trước thời điểm s , hệ đã ở trạng thái nào đó, còn tại thời điểm s , hệ ở trạng thái i . Chúng ta muốn đánh giá xác suất để tại thời điểm t ($t > s$), hệ sẽ ở trạng thái j . Nếu xác suất này chỉ phụ thuộc vào bộ bốn (s, i, t, j) , tức là $P[X(t) = j/X(s) = i] = p(s, i, t, j)$.

$t, j)$ là đúng $\forall i, \forall j, \forall s, \forall t$ thì điều này có nghĩa là, sự tiến triển của hệ trong tương lai chỉ phụ thuộc vào hiện tại (trạng thái của hệ tại thời điểm s), và hoàn toàn độc lập với quá khứ (*tính không nhớ*). Đó chính là *tính Markov*. Lúc này quá trình ngẫu nhiên $X(t)$ được gọi là *quá trình Markov*.

Trong ví dụ trên $P[X(1) = 2/X(0) = 1]$ là xác suất có điều kiện của sự kiện $X(1) = 2$ (tại thời điểm $t=1$, hệ thống nằm tại vị trí 2) với điều kiện $X(0) = 1$ (tại thời điểm $t = 0$, hệ thống nằm tại vị trí 1). Nếu quá trình ngẫu nhiên có tính Markov thì xác suất này chỉ phụ thuộc vào tình trạng của hệ tại thời điểm $s = 0$ và hoàn toàn độc lập với các tình trạng của hệ trong quá khứ (trước thời điểm $s = 0$).

Định nghĩa 2

Nếu không gian trạng thái S gồm một số hữu hạn hoặc vô hạn đếm được các trạng thái thì quá trình Markov $X(t)$ được gọi là *xích Markov*. Lúc này, có thể kí hiệu $S = \{1, 2, 3, \dots\}$, tức là các trạng thái được đánh số. Hơn nữa, nếu tập các giá trị t không quá đếm được (chẳng hạn, $t = 0, 1, 2, \dots$) thì ta có xích Markov với *thời gian rời rạc*, hay xích Markov rời rạc. Nếu $t \in [0, \infty)$ thì ta có xích Markov với *thời gian liên tục*, hay xích Markov liên tục.

Định nghĩa 3

Xét một xích Markov. Nếu xác suất chuyển trạng thái $p(s, i, t, j) = p(s+h, i, t+h, j), \forall i, \forall j, \forall s, \forall t$ và $\forall h > 0$, thì ta nói rằng xích Markov *thuần nhất theo thời gian*.

Đây là một khái niệm mới và sẽ được giải thích ngay sau đây trong mục 1.2. Ngoài ra với mục đích tìm hiểu bước đầu, trong các mục 1.2 và 1.3 chúng ta sẽ chỉ xét xích Markov rời rạc và thuần nhất theo thời gian. Ví dụ về xích Markov liên tục sẽ được xem xét trong mục 2.4 và 2.5.

1.2. Ma trận xác suất chuyển trạng thái và phân phối dừng

Trong mục này chúng ta đưa ra khái niệm về ma trận xác suất chuyển trạng thái của một xích Markov rời rạc và thuần nhất theo thời gian với không gian trạng thái gồm N phần tử. Trong trường hợp xích Markov rời rạc và thuần nhất có không gian trạng thái với số phần tử vô hạn đếm được, khái niệm về ma trận xác suất chuyển trạng thái sẽ được xây dựng một cách tương tự.

Ví dụ: Xét lại ví dụ đã có trong mục 1.1, nhưng với một cách minh họa khác trong lĩnh vực dịch vụ. Trong một khu phố 1000 dân (khách hàng) có 3 siêu thị là A, B, và C (A, B, C được coi là các vị trí 1, 2, 3 của hệ thống siêu thị này). Giả sử rằng, trong từng tháng mỗi khách hàng luôn trung thành với một siêu thị. Ngoài ra, cũng giả sử rằng trong tháng đầu số khách vào các siêu thị lần lượt là 200, 500 và 300; tức là có 20% khách hàng vào siêu thị A, 50% vào B và 30% vào C. Như vậy, có thể dự đoán rằng một khách hàng vào A với xác suất 0,2; vào B với xác suất 0,5 và vào C với xác suất 0,3. Để mô tả tình trạng phân chia thị phần trong tháng đầu (tháng 0) của hệ thống siêu thị trên, chúng ta thiết lập biến ngẫu nhiên $X(0)$ với quy tắc: nếu khách hàng mua hàng ở siêu thị A thì đặt $X(0)=1$, ở siêu thị B thì đặt $X(0) = 2$, còn ở siêu thị C thì $X(0) = 3$. Lúc đó, $X(0)$ có bảng phân phối xác suất sau:

Các giá trị của X(0)	1	2	3
Xác suất tương ứng	0,2	0,5	0,3

Kí hiệu $P[X(0) = 1] = \pi_1^{(0)}$, $P[X(0) = 2] = \pi_2^{(0)}$, $P[X(0) = 3] = \pi_3^{(0)}$, thì véc tơ $\Pi^{(0)} = [\pi_1^{(0)}, \pi_2^{(0)}, \pi_3^{(0)}] = [0,2 \ 0,5 \ 0,3]$ được gọi là *véc tơ phân phối xác suất tại thời điểm t = 0* hay *véc tơ phân phối ban đầu*. Các thành phần của $\Pi^{(0)}$ cho biết tỉ lệ phần trăm (%) khách hàng vào các siêu thị A, B và C.

Những tháng sau, ta giả sử xác suất để một người khách, đã vào mua hàng ở siêu thị A tháng trước, vào lại A trong tháng sau luôn là 0,8; chuyển sang mua hàng ở B luôn là 0,1 và chuyển sang C luôn là 0,1. Xác suất để một người khách, đã vào mua hàng ở siêu thị B tháng trước chuyển sang A luôn là 0,07; vào lại B luôn là 0,9 và chuyển sang C luôn là 0,03. Còn xác suất để một người khách, đã vào siêu thị C tháng trước chuyển sang A luôn là 0,083; chuyển sang B luôn là 0,067 và vào lại C luôn là 0,85. Lúc đó các xác suất chuyển của khách hàng được cho thông qua *ma trận xác suất chuyển trạng thái P* (còn gọi là *ma trận chuyển* sau một bước).

$$P = \begin{bmatrix} 0,8 & 0,1 & 0,1 \\ 0,07 & 0,9 & 0,03 \\ 0,083 & 0,067 & 0,85 \end{bmatrix} = [p_{ij}]_{3 \times 3}.$$

Để mô tả tình trạng phân chia thị phần trong tháng t (t = 1, 2, 3, ...) của hệ thống siêu thị trên, có thể thiết lập biến ngẫu nhiên X(t) với quy tắc tương tự như khi thiết lập X(0): nếu khách hàng mua hàng ở siêu thị A thì đặt X(t) = 1, ở siêu thị B thì đặt X(t) = 2, còn ở siêu thị C thì X(t) = 3. Vấn đề đặt ra là X(t) có bảng phân phối xác suất như thế nào.

Trước hết ta đi tìm bảng phân phối xác suất cho X(1). Xét $p_{12} = P[(X(1) = 2 | X(0) = 1)] = 0,1$ là xác suất để một người khách, đã vào mua hàng ở siêu thị A tháng 0 chuyển sang mua hàng ở siêu thị B trong tháng 1. Ngoài ra, $P[X(t+1) = 2 | X(t) = 1] = 0,1 \ \forall t$ là số tự nhiên, vì theo giả thiết của bài toán thì xác suất để một người khách, đã vào mua hàng ở siêu thị A tháng trước chuyển sang mua hàng ở B luôn là 0,1. Vậy p_{12} được gọi là *xác suất chuyển sau một bước* từ vị trí 1 sang vị trí 2, bởi vậy có thể dùng kí hiệu $p_{12}^{(1)}$ để chỉ rõ đây là xác suất chuyển sau một bước. Các phần tử $p_{ij} \ \forall i = 1, 2, 3$ và $\forall j = 1, 2, 3$ của ma trận P có ý nghĩa tương tự.

Dễ thấy rằng trong tháng 1 số khách hàng mua hàng tại siêu thị A là $200 \times 0,8 + 500 \times 0,07 + 300 \times 0,083 = 219,9 (\approx 220)$; số khách hàng mua hàng tại siêu thị B là $200 \times 0,1 + 500 \times 0,9 + 300 \times 0,067 = 490,1 (\approx 490)$; còn số khách hàng mua hàng tại siêu thị C sẽ là $200 \times 0,1 + 500 \times 0,03 + 300 \times 0,85 = 290$. Do tổng số khách hàng là 1000, nên X(1) có bảng phân phối xác suất sau:

Các giá trị của X(1)	1	2	3
Xác suất tương ứng	0,2199	0,4901	0,2900

Vậy véc tơ phân phối xác suất tại thời điểm $t = 1$ là $\Pi^{(1)} = [\pi_1^{(1)}, \pi_2^{(1)}, \pi_3^{(1)}]$ cho biết tỉ lệ phần trăm khách hàng vào các siêu thị A, B và C trong tháng 1. Bằng phép tính ma trận cũng tìm được $\Pi^{(1)}$ như sau:

$$\begin{aligned}\Pi^{(1)} &= \Pi^{(0)} \times P = [0,2 \ 0,5 \ 0,3] \times \begin{bmatrix} 0,8 & 0,1 & 0,1 \\ 0,07 & 0,9 & 0,03 \\ 0,083 & 0,067 & 0,85 \end{bmatrix} \\ &= [0,2199 \ 0,4901 \ 0,2900].\end{aligned}$$

Tương tự có thể tìm được $\Pi^{(2)}$:

$$\begin{aligned}\Pi^{(2)} &= \Pi^{(1)} \times P = [0,2199 \ 0,4901 \ 0,2900] \times \begin{bmatrix} 0,8 & 0,1 & 0,1 \\ 0,07 & 0,9 & 0,03 \\ 0,083 & 0,067 & 0,85 \end{bmatrix} \\ &= [0,234297 \ 0,48251 \ 0,283193].\end{aligned}$$

Sau đây ta đi tìm ma trận xác suất chuyển trạng thái sau hai bước. Kí hiệu $p_{12}^{(2)}$ là xác suất chuyển từ vị trí 1 sang vị trí 2 sau hai bước. Theo công thức xác suất toàn phần ta có:

$$\begin{aligned}p_{12}^{(2)} &= P[X(2) = 2 / X(0) = 1] = P[X(1) = 1 / X(0) = 1] \times P[X(2) = 2 / X(1) = 1] \\ &\quad + P[X(1) = 2 / X(0) = 1] \times P[X(2) = 2 / X(1) = 2] \\ &\quad + P[X(1) = 3 / X(0) = 1] \times P[X(2) = 2 / X(1) = 3] \\ &= p_{11}^{(1)} p_{12}^{(1)} + p_{12}^{(1)} p_{22}^{(1)} + p_{13}^{(1)} p_{32}^{(1)} \\ &= \sum_{k=1}^3 p_{1k}^{(1)} p_{k2}^{(1)} = 0,8 \times 0,1 + 0,1 \times 0,9 + 0,1 \times 0,067 = 0,1767.\end{aligned}$$

Một cách hoàn toàn tương tự, ta có xác suất chuyển từ vị trí i sang vị trí j sau hai bước là $p_{ij}^{(2)} = p_{i1}^{(1)} p_{1j}^{(1)} + p_{i2}^{(1)} p_{2j}^{(1)} + p_{i3}^{(1)} p_{3j}^{(1)} = \sum_{k=1}^3 p_{ik}^{(1)} p_{kj}^{(1)}$. Vậy ta có ma trận chuyển sau hai bước là:

$$P^{(2)} = [p_{ij}^{(2)}]_{3 \times 3} = P^{(1)} \times P^{(1)} = P \times P = P^2$$

$$= \begin{bmatrix} 0,8 & 0,1 & 0,1 \\ 0,07 & 0,9 & 0,03 \\ 0,083 & 0,067 & 0,85 \end{bmatrix} \times \begin{bmatrix} 0,8 & 0,1 & 0,1 \\ 0,07 & 0,9 & 0,03 \\ 0,083 & 0,067 & 0,85 \end{bmatrix}.$$

Để thấy $\Pi^{(2)} = \Pi^{(1)} \times P = \Pi^{(0)} \times P^2$. Tương tự, có thể chứng minh được $\Pi^{(n+m)} = \Pi^{(n)} \times P^{(m)}$, trong đó $\Pi^{(n+m)}$ và $\Pi^{(n)}$ là các véc tơ phân phối tại các thời điểm $t = m + n$ và $t = n$, còn $P^{(m)}$ là ma trận xác suất chuyển trạng thái sau m bước.

Do $P^{(m)} = [p_{ij}^{(m)}]_{3 \times 3}$ nên $P[X(m) = j / X(0) = i] = P[X(n + m) = j / X(n) = i] = P[X(n' + m) = j / X(n') = i] = p_{ij}^{(m)}$, là xác suất chuyển từ vị trí i sang vị trí j sau m bước. Đặt $n = s$,

$t = n+m$ và $h = n' - n$ thì có ngay $P[X(t) = j / X(s) = i] = P[X(t + h) = j / X(s + h) = i]$, hay $p(s, i, t, j) = p(s + h, i, t + h, j)$ luôn đúng $\forall s, \forall t, \forall h$. Từ các phân tích trên đây và đối chiếu với các định nghĩa 1, 2 và 3 mục 1.1, ta thấy quá trình ngẫu nhiên $X(t)$ với $t = 0, 1, 2, \dots$ trong ví dụ này chính là *một xích Markov rời rạc và thuần nhất theo thời gian*.

Để khái quát hóa các khái niệm đã trình bày, chúng ta xét xích Markov rời rạc và thuần nhất theo thời gian $X(t)$, $t = 0, 1, 2, \dots$ với không gian trạng thái gồm N phần tử mà ta kí hiệu là $S = \{1, 2, \dots, N\}$.

Định nghĩa 1

Giả sử tại thời điểm $t = n$, $X(n)$ cũng có thể nhận một trong N giá trị $1, 2, \dots, N$ với các xác suất tương ứng là $\pi_1^{(n)}, \pi_2^{(n)}, \dots, \pi_N^{(n)}$ (với $\pi_1^{(n)} + \pi_2^{(n)} + \dots + \pi_N^{(n)} = 1$) thì véc tơ $\Pi^{(n)} = [\pi_1^{(n)}, \pi_2^{(n)}, \dots, \pi_N^{(n)}]$ được gọi là *véc tơ phân phối tại thời điểm $t = n$* . Với $t = 0$ ta có *véc tơ phân phối ban đầu* $\Pi^{(0)} = [\pi_1^{(0)}, \pi_2^{(0)}, \dots, \pi_N^{(0)}]$.

Ma trận $P = [p_{ij}]_{N \times N}$, trong đó $p_{ij} = p(t, i, t + 1, j) = P[X(t + 1) = j / X(t) = i] \forall t$ là *xác suất chuyển trạng thái từ vị trí i sang vị trí j sau một bước*, $\forall i = 1, 2, \dots, N$ và $\forall j = 1, 2, \dots, N$, được gọi là *ma trận xác suất chuyển trạng thái* hay *ma trận chuyển* sau một bước.

Ví dụ: Tiếp tục xét ví dụ trên, trong đó đã tìm được $\Pi^{(1)} = [0,2199 \quad 0,4901 \quad 0,2900]$, $\Pi^{(2)} = [0,234297 \quad 0,482510 \quad 0,283193]$. Để thấy, các véc tơ phân phối xác suất $\Pi^{(1)}, \Pi^{(2)}, \Pi^{(3)}, \dots$ tại các thời điểm $t = 1, 2, 3, \dots$ được tính theo công thức: $\Pi^{(1)} = \Pi^{(0)} \times P$, $\Pi^{(2)} = \Pi^{(1)} \times P = \Pi^{(0)} \times P^2 \dots$ và $\Pi^{(n+1)} = \Pi^{(n)} \times P = \Pi^{(0)} \times P^{n+1}$, $\forall n$. Sau 21 bước (21 tháng), ta có $\Pi^{(21)} = [0,272257 \quad 0,455523 \quad 0,272220]$.

Các véc tơ phân phối (hay tỉ lệ phần trăm khách hàng vào các siêu thị A, B, C) sau 1, 2, 3, ..., 21 tháng được cho trong bảng IV.1.

Vấn đề đặt ra là liệu $\Pi = \lim_{n \rightarrow \infty} \Pi^{(n)}$ có tồn tại không và nếu tồn tại thì được tìm bằng cách nào. Trong ví dụ này, chúng ta sẽ tìm được $\Pi = [0,273 \quad 0,454 \quad 0,273]$, biểu thị cho tỉ lệ phần trăm cân bằng dừng (*stationary equilibrium*) số khách hàng vào các siêu thị A, B, C sau một thời gian đủ dài.

Cách tính Π

Xuất phát từ $\Pi^{(n+1)} = \Pi^{(n)} \times P$, cho qua giới hạn cả hai vế khi $n \rightarrow \infty$ ta có: $\Pi = \Pi \times P$, hay $\Pi \times (I - P) = 0$.

Do P là ma trận đặc biệt (ma trận chuyển) nên nó là ma trận suy biến. Khi viết lại dưới dạng hệ phương trình (3 ẩn, 3 phương trình) ta phải loại bớt một phương trình đi, và thêm vào hệ thức $\pi_1 + \pi_2 + \pi_3 = 1$ và ràng buộc $\pi_k \geq 0$ ($k = 1, 2, 3$). Kí hiệu $x = \pi_1$, $y = \pi_2$ và $z = \pi_3$, ta sẽ có hệ:

$$\begin{cases} 0,2x - 0,07y - 0,083z = 0 \\ -0,1x + 0,1y - 0,067z = 0 \\ x + y + z = 1 \end{cases} \Leftrightarrow \begin{cases} x = 0,273 \\ y = 0,454 \\ z = 0,273 \end{cases}$$

Vậy $\Pi = [0,273 \ 0,454 \ 0,273]$.

Bảng IV.1. Tỷ lệ phần trăm khách hàng vào các siêu thị

Tháng	A	B	C
1	0.2199	0.4901	0.29
2	0.234297	0.48251	0.283193
3	0.2447183	0.476662631	0.27861905
4	0.2522664	0.472135676	0.2755979
5	0.2577373	0.46861381	0.27364893
6	0.2617056	0.465860633	0.27243373
7	0.2645868	0.463698194	0.27171505
8	0.2666806	0.461991958	0.27132742
9	0.2682041	0.460639762	0.27115613
10	0.269314	0.459563657	0.27112231
11	0.2701238	0.45870389	0.27117228
12	0.2707156	0.458014426	0.27126994
13	0.2711489	0.457459633	0.27139144
14	0.2714668	0.457011789	0.27152141
15	0.2717005	0.456649225	0.27165023
16	0.2718729	0.456354922	0.27177223
17	0.2720002	0.456115454	0.27188433
18	0.2720947	0.455920181	0.27198516
19	0.2721649	0.455760634	0.27207446
20	0.2722173	0.45563005	0.2721526
21	0.2722566	0.455523004	0.27222035

Định nghĩa 2

Xét xích Markov rời rạc và thuần nhất với ma trận chuyển $P = [p_{ij}]_{N \times N}$. Lúc đó, véc tơ phân phối xác suất $\Pi = [\pi_1, \pi_2, \dots, \pi_N]$ thỏa mãn điều kiện $\Pi \times (I - P) = 0$ được gọi là *phân phối dừng* của xích Markov đã cho.

Có thể thấy ngay, phân phối dừng Π không phụ thuộc vào $\Pi^{(0)}$ mà chỉ phụ thuộc vào ma trận P .

Một cách toán học, ta nói *mô hình xích Markov rời rạc thuần nhất* chính là bộ ba $(X(t_n), S/\Pi, P)$. Áp dụng mô hình xích Markov để phân tích một vấn đề nào đó trong Kinh tế, Kỹ thuật, Sinh học,... được coi là việc *ứng dụng phân tích Markov*.

1.3. Các tính chất và định lý

Xét xích Markov rời rạc và thuần nhất với ma trận chuyển $P = [p_{ij}]_{N \times N}$. Có thể chứng minh được các tính chất và định lý sau:

Các tính chất

$$1/ p^{(n+m)}_{ij} = \sum_{k=1}^N p^{(n)}_{ik} p^{(m)}_{kj} \text{ (đây là phương trình Chapman-Kolmogorov).}$$

$$2/ P^{(2)} = P \times P = P^2, P^{(n)} = P^n \text{ và } P^{(n+m)} = P^{(n)} \times P^{(m)}.$$

$$3/ \Pi^{(n+m)} = \Pi^{(n)} \times P^{(m)}.$$

Định lý

1/ Giả sử P là ma trận xác suất chuyển chính quy, tức là tồn tại chỉ số n_0 , sao cho $\forall i, j$, thì xác suất chuyển từ i đến j sau n_0 bước là một số dương: $p^{(n_0)}_{ij} > 0$. Khi đó tồn tại $\pi_1, \pi_2, \dots, \pi_N > 0$, và $\pi_1 + \pi_2 + \dots + \pi_N = 1$ để cho $\lim_{n \rightarrow \infty} p^{(n)}_{ij} = \pi_j$, không phụ thuộc vào i .

Các số $\pi_1, \pi_2, \dots, \pi_N$ được tìm từ hệ phương trình

$$x_j = \sum_{k=1}^N x_k p_{kj}, j=1, 2, \dots, N; x_j \geq 0 \forall j \text{ và } \sum_{j=1}^N x_j = 1.$$

2/ Nếu có các số $\pi_1, \pi_2, \dots, \pi_N$ thỏa mãn điều kiện $\pi_1 + \pi_2 + \dots + \pi_N = 1$ và $\lim_{n \rightarrow \infty} p^{(n)}_{ij} = \pi_j$, không phụ thuộc vào i , thì ma trận P là ma trận chính quy.

Lưu ý

Phân phối $(\pi_1, \pi_2, \dots, \pi_N)$ thỏa mãn điều kiện $\pi_1 + \pi_2 + \dots + \pi_N = 1$ và $\lim_{n \rightarrow \infty} p^{(n)}_{ij} = \pi_j$, không phụ thuộc vào i , được gọi là phân phối giới hạn. Ngoài ra, nếu điều kiện $\pi_j > 0, \forall j$ được thỏa mãn thì phân phối này được gọi là *phân phối Ergodic*. Có thể chứng minh được rằng, nếu phân phối giới hạn tồn tại thì đó là phân phối dừng (duy nhất). Tuy nhiên, điều ngược lại không luôn đúng.

2. Một số ứng dụng của phân tích Markov

Phân tích Markov có nhiều ứng dụng trong Kinh tế, Kỹ thuật, Sinh học, Xã hội học, Công nghệ thông tin,... Trong mục này, chúng ta sẽ xem xét các ứng dụng như tìm cân bằng thị phần, xác định chính sách thay thế vật tư thiết bị, dự báo thất thu cho các hợp đồng thực hiện trước, tìm phân phối giới hạn của một hệ thống kỹ thuật và một ứng dụng của quá trình sinh – tử cho hệ thống hàng chờ.

2.1. Tìm cân bằng thị phần

Ta nhắc lại một cách vắn tắt bài toán cho ở mục 1.2: Trong một khu phố 1000 dân (khách hàng) có 3 siêu thị là A, B, và C. Giả sử, trong tháng đầu, số khách vào các siêu thị lần lượt là 200, 500 và 300. Những tháng sau đó, ta giả sử xác suất để một khách hàng (đã vào siêu thị A lúc trước) vào lại A luôn là 0,8; chuyển sang B luôn là 0,1 và chuyển sang C luôn là 0,1... Các xác suất chuyển khác của khách hàng ("trụ lại" B, chuyển sang A, chuyển sang C...) được cho thông qua ma trận chuyển P

$$P = \begin{bmatrix} 0,8 & 0,1 & 0,1 \\ 0,07 & 0,9 & 0,03 \\ 0,083 & 0,067 & 0,85 \end{bmatrix}$$

Lúc đó, theo kết quả đã biết, tỉ lệ phần trăm cân bằng dừng (khi thời gian đủ dài) số khách hàng vào các siêu thị A, B, C là 27,3%, 45,4% và 27,3% có thể tìm được từ hệ $\Pi \times (I - P) = 0$.

2.2. Chính sách thay thế vật tư thiết bị

Trong một hệ thống điện kỹ thuật, các thiết bị cùng một loại được phân ra các tình trạng sau đây: vừa mới thay, còn tốt, vẫn dùng được và đã bị hỏng. Theo số liệu thống kê được, ta có ma trận xác suất chuyển trạng thái như sau:

$$P = \begin{bmatrix} 0 & 0,8 & 0,2 & 0 \\ 0 & 0,6 & 0,4 & 0 \\ 0 & 0 & 0,5 & 0,5 \\ 1,0 & 0 & 0 & 0 \end{bmatrix},$$

trong đó, sau mỗi tuần (xem hàng đầu của ma trận P) có 0%, 80%, 20% và 0% số các thiết bị mới thay chuyển sang tình trạng mới thay, còn tốt, vẫn dùng được và đã bị hỏng. Các hàng khác của ma trận P được giải thích một cách tương tự. Ta đi tìm phân phối dừng Π bằng phương pháp đã biết.

Xuất phát từ $\Pi^{(n+1)} = \Pi^{(n)} \times P$, cho qua giới hạn cả hai vế khi $n \rightarrow \infty$ ta có: $\Pi = \Pi \times P$, hay $\Pi \times (I - P) = 0$.

Do P là ma trận đặc biệt (ma trận chuyển xác suất) nên nó là ma trận suy biến.

Khi viết lại dưới dạng hệ phương trình (4 ẩn, 4 phương trình) ta phải loại bớt một phương trình đi, và thêm vào hệ thức $\pi_1 + \pi_2 + \pi_3 + \pi_4 = 1$ và ràng buộc $\pi_k \geq 0$ ($k = 1, 2, 3, 4$). Kí hiệu $x_1 = \pi_1, x_2 = \pi_2, x_3 = \pi_3$ và $x_4 = \pi_4$ ta sẽ có hệ:

$$\begin{cases} x_1 - x_4 = 0 \\ -0,8x_1 + 0,4x_2 = 0 \\ -0,2x_1 - 0,4x_2 + 0,5x_3 = 0 \\ -0,5x_3 + x_4 = 0 \\ x_1 + x_2 + x_3 + x_4 = 1 \end{cases} \Leftrightarrow \begin{cases} x_1 = x_4 = \frac{1}{6} \\ x_2 = x_3 = \frac{1}{3} \end{cases}$$

Vậy phân phối dừng $\Pi = [1/6 \ 1/3 \ 1/3 \ 1/6]$.

Giả sử rằng chi phí thay mới một thiết bị là 25 nghìn (đồng) và thất thu khi mỗi một thiết bị hỏng là 18,5 nghìn, thì mỗi tuần hệ thống trên phải chi trung bình trên một thiết bị số tiền là: $(1/6) \times 25 + (1/6) \times 18,5 = 7,25$ nghìn / thiết bị / tuần.

Ta xét phương án thứ hai cho việc thay thế vật tư thiết bị với ma trận xác suất chuyển trạng thái sau đây:

$$P = \begin{bmatrix} 0 & 0,8 & 0,2 \\ 0 & 0,6 & 0,4 \\ 1,0 & 0 & 0 \end{bmatrix}$$

Ma trận này tương ứng với chính sách mới về thay thế vật tư thiết bị là: thay thế mỗi thiết bị một khi kiểm tra và phát hiện thiết bị ở tình trạng vẫn dùng được. Điều này có thể dẫn tới việc giảm thiểu thất thu do thiết bị hỏng gây nên. Thật vậy, ứng với ma trận P trên đây, phân phối dừng $\Pi = [1/4 \ 1/2 \ 1/4]$. Lúc này, mỗi tuần hệ thống trên phải chi trung bình trên một thiết bị số tiền là: $(1/4) \times 25 + (0) \times 18,5 = 6,25$ nghìn / thiết bị / tuần. Như vậy hệ thống sẽ tiết kiệm được 1 nghìn / thiết bị / một tuần. Nếu hệ thống có 2000 thiết bị, thì nhờ chính sách thay thế vật tư mới, mỗi tuần hệ thống sẽ tiết kiệm được 2 triệu (đồng).

2.3. Phân tích Markov trong dự báo thất thu cho các hợp đồng thực hiện trước

Một công ti kinh doanh trong ngành điện chuyên về sửa chữa và thay thế phụ tùng đề ra chính sách tín dụng: đáp ứng yêu cầu của khách hàng trước, thanh toán sau. Phần nhiều hợp đồng sẽ được thanh toán đúng thời hạn, một tỉ lệ nhất định sẽ được công ti cho thanh toán chậm, còn một số ít không thanh toán được. Theo kinh nghiệm, sau hai hay ba hợp đồng thanh toán chậm của một khách hàng nào đó là hợp đồng không thanh toán được sau một thời gian dài, công ti coi đây là hợp đồng “xấu” và sẽ cắt bỏ chính sách tín dụng với khách hàng đó. Như vậy tại từng thời điểm các hợp đồng có thể rơi vào một trong các tình trạng (trạng thái) sau:

- S_0 : hợp đồng được thanh toán,
- S_1 : hợp đồng không được thanh toán,

- S_2 : hợp đồng sẽ được thanh toán đúng thời hạn,
- S_3 : hợp đồng sẽ được thanh toán chậm.

Sau đây là ma trận xác suất chuyển trạng thái (sau từng tháng):

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0,5 & 0 & 0,3 & 0,2 \\ 0,4 & 0,3 & 0,2 & 0,1 \end{bmatrix}$$

Hiện tại công ty có các hợp đồng phải thanh toán đúng hạn với tổng số 500 triệu, và các hợp đồng cho thanh toán chậm với tổng số 100 triệu. Hãy xác định trong tổng trên có bao nhiêu sẽ được thanh toán, còn bao nhiêu sẽ là nợ “xấu” không đòi được.

Đây là bài toán khá phức tạp liên quan tới phân loại các trạng thái của xích Markov là vấn đề chúng ta sẽ không trình bày trong giáo trình này. Tuy nhiên, có thể thấy ngay rằng các trạng thái S_0 và S_1 là các trạng thái “hấp thụ” (*absorbing state*), tức là mọi hợp đồng dù hiện đang ở trạng thái nào thì cuối cùng sau một thời gian nhất định cũng sẽ rơi vào một trong hai trạng thái trên. Trong khi đó các trạng thái S_2 và S_3 được gọi là các trạng thái truyền ứng (hay các trạng thái di chuyển).

Để tìm câu trả lời cho vấn đề đặt ra, chúng ta cần thực hiện các bước sau: Trước hết, ta chia ma trận P theo khối.

$$P = \begin{bmatrix} J & O \\ K & M \end{bmatrix} \text{ với } J = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, K = \begin{bmatrix} 0,5 & 0 \\ 0,4 & 0,3 \end{bmatrix}, O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, M = \begin{bmatrix} 0,3 & 0,2 \\ 0,2 & 0,1 \end{bmatrix}.$$

Sau đó, ta tìm ma trận $R = I - M$ và ma trận nghịch đảo của nó R^{-1} , ở đây I là ma trận đơn vị cùng cỡ với ma trận M . Ta có:

$$R^{-1} = \begin{bmatrix} 1,5254 & 0,3390 \\ 0,3390 & 1,1864 \end{bmatrix},$$

và tính được:

$$R^{-1} \times K = \begin{bmatrix} 0,8983 & 0,1017 \\ 0,6441 & 0,3559 \end{bmatrix}.$$

Các phần tử trong ma trận trên có ý nghĩa đặc biệt. Trong số các hợp đồng hiện tại ở trạng thái S_2 (phải thanh toán đúng kì hạn) cuối cùng sau một thời gian nhất định có 89,83% sẽ rơi vào trạng thái S_0 (được thanh toán) và 10,17% sẽ rơi vào trạng thái S_1 (không được thanh toán). Còn trong số các hợp đồng hiện tại ở trạng thái S_3 (thanh toán

chậm) cuối cùng sau một thời gian nhất định có 64,41% sẽ rơi vào trạng thái S_0 (được thanh toán) và 35,59% sẽ rơi vào trạng thái S_1 (không được thanh toán).

Thực hiện phép tính:

$$[500 \ 100] \times \begin{bmatrix} 0,8983 & 0,1017 \\ 0,6441 & 0,3559 \end{bmatrix} = [459,32 \ 140,68],$$

ta thấy trong 500 triệu phải thanh toán đúng kì hạn và 100 triệu thanh toán chậm cuối cùng sẽ có 459,32 triệu được thanh toán và 140,68 triệu là nợ “xấu” không đòi được. Để cải thiện tình trạng này, công ti cần nghiên cứu tìm ra một chính sách tín dụng hợp lí hơn.

Ngoài ra, ma trận R^{-1} còn cho biết các thông tin sau:

– Tổng của các phần tử trên hàng thứ nhất là 1,8644 là thời gian trung bình (tháng) mà một hợp đồng dạng phải thanh toán đúng kì hạn sẽ trải qua trước khi rơi vào một trong các trạng thái hấp thụ, tức là trở thành hợp đồng thanh toán được hoặc hợp đồng “xấu”.

– Tổng các phần tử trên hàng thứ hai của R^{-1} cũng có ý nghĩa tương tự đối với các hợp đồng dạng thanh toán chậm.

– Phần tử nằm trên hàng 1 và cột 1 của R^{-1} cho biết thời gian trung bình (tháng) mà một hợp đồng dạng phải thanh toán đúng hạn sẽ ở trong trạng thái S_2 trước khi nó rơi vào một trong các trạng thái hấp thụ là 1,5254 tháng. Phần tử nằm trên hàng 1 và cột 2 cho biết thời gian trung bình (tháng) mà một hợp đồng dạng phải thanh toán đúng hạn sẽ ở trong trạng thái S_3 trước khi nó rơi vào một trong các trạng thái hấp thụ là 0,3390 tháng.

– Các phần tử nằm trên hàng 2 của ma trận R^{-1} có ý nghĩa tương tự đối với một hợp đồng dạng được thanh toán chậm.

Sau đây, chúng ta sẽ đưa ra *một số công thức giải thích* các phân tích trên đây cho trường hợp ma trận xác suất chuyển trạng thái của xích Markov có dạng sau:

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ p_{20} & p_{21} & p_{22} & p_{23} \\ p_{30} & p_{31} & p_{32} & p_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0,5 & 0 & 0,3 & 0,2 \\ 0,4 & 0,3 & 0,2 & 0,1 \end{bmatrix} \quad (\text{nghĩa trong bài toán trên}),$$

$$= \begin{bmatrix} J & O \\ K & M \end{bmatrix} \quad \text{với } J = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, K = \begin{bmatrix} p_{20} & p_{21} \\ p_{30} & p_{31} \end{bmatrix}, O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, M = \begin{bmatrix} p_{22} & p_{23} \\ p_{32} & p_{33} \end{bmatrix},$$

trong đó p_{ij} là xác suất chuyển từ trạng thái S_i sang trạng thái S_j sau một bước. Không gian trạng thái gồm bốn trạng thái S_0, S_1, S_2 và S_3 ; các trạng thái S_0 và S_1 là các trạng thái hấp thụ, còn S_2 và S_3 là các trạng thái truyền ứng. Chúng ta dùng các kí hiệu:

$$U = \begin{bmatrix} u_{20} & u_{21} \\ u_{30} & u_{31} \end{bmatrix},$$

với u_{ik} là xác suất hấp thụ vào trạng thái S_k khi trạng thái ban đầu là S_i , $k = 0, 1$, còn $i = 2, 3$.

$$V = \begin{bmatrix} v_2 \\ v_3 \end{bmatrix},$$

với v_i là thời gian trung bình cho tới khi rơi vào một trong các trạng thái hấp thụ nếu trạng thái ban đầu là S_i , $i = 2, 3$.

$$W = \begin{bmatrix} w_{22} & w_{23} \\ w_{32} & w_{33} \end{bmatrix},$$

với w_{ij} là thời gian trung bình xích Markov ở trong trạng thái S_j trước khi nó rơi vào một trong các trạng thái hấp thụ nếu trạng thái ban đầu là S_i , $i = 2, 3$. Lúc đó:

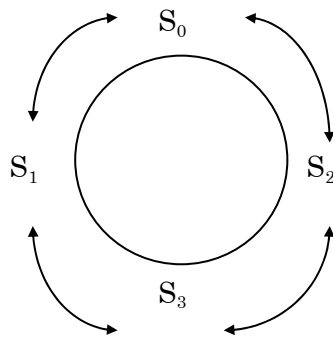
$$U = (I - M)^{-1}, \quad V = (I - M)^{-1} \times \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{và} \quad W = (I - M)^{-1} \times \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Chú ý: Việc chứng minh các công thức trên cho trường hợp tổng quát thực ra cũng không quá khó, có thể tìm thấy trong các sách tham khảo về quá trình Markov. Cách ứng dụng phân tích Markov như trong mục này còn có thể được áp dụng rộng rãi trong nhiều lĩnh vực khác như Sinh học, Xã hội học, Lí thuyết nhận dạng và Thiết kế các hệ thống kĩ thuật, trong đó có Kĩ thuật điện.

2.4. Tìm phân phối giới hạn cho một hệ thống kĩ thuật

Một hệ thống kĩ thuật có hai chi tiết có thể bị hỏng ở bất kì thời điểm nào. Tại mỗi thời điểm hệ thống có thể rơi vào một trong những trạng thái sau (xem hình IV.1):

- S_0 : cả 2 chi tiết tốt;
- S_1 : chi tiết 1 hỏng, chi tiết 2 bình thường;
- S_2 : chi tiết 1 bình thường, chi tiết 2 hỏng;
- S_3 : cả hai chi tiết đều hỏng.



Hình IV.1. Sơ đồ các trạng thái

Nói cách khác, tại mỗi thời điểm t , biến $X(t)$ có thể rơi vào một trong các vị trí / trạng thái S_0, S_1, S_2 và S_3 . Chú ý rằng lúc này ta có xích Markov (thời gian) liên tục với không gian trạng thái $S = \{S_0, S_1, S_2, S_3\}$. Sau đây, chúng ta sẽ tìm cách xác định phân phối giới hạn (*long run distribution*) của $\{X(t)\}_{t \geq 0}$. Đây là một vấn đề khá phức tạp nên chúng ta chỉ có thể trình bày vấn đề một cách vắn tắt.

Trước hết ta nhắc lại về phân phối Poát-xông và phân phối mũ. Giả sử dòng tín hiệu đến (hay xảy ra) tuân theo phân phối Poát-xông $\mathcal{P}(\lambda)$ với λ là số tín hiệu đến trung bình trong một khoảng thời gian nhất định (coi là một đơn vị thời gian), λ còn được gọi là cường độ của dòng tín hiệu đến. Lúc đó, trong khoảng thời gian như trên thì số tín hiệu xảy ra sẽ nhận giá trị k với xác suất $\frac{\lambda^k e^{-\lambda}}{k!}$. Ta gọi phần tử xác suất P là xác suất xuất hiện (ít nhất) một tín hiệu trong khoảng thời gian Δt . Thế thì, do tính “đơn nhất” của quá trình Poát-xông, P cũng là xác suất xuất hiện đúng một tín hiệu trong khoảng thời gian Δt . Theo công thức đã biết thì $P = \lambda \Delta t$ (chính xác tới vô cùng bé $o(\Delta t)$). Chẳng hạn, nếu $\lambda = 6$ tín hiệu/ 1 phút và $\Delta t = 2$ giây, ta sẽ có $P = \lambda \Delta t = 6 \times (1/30) = 1/5 = 0,2$. Từ đó, ta thấy xác suất để có 1 tín hiệu đến trong khoảng thời gian 2 giây là 0,2.

Xét biến ngẫu nhiên T (chẳng hạn thời gian phục vụ một tín hiệu trong một hệ dịch vụ), có phân phối mũ $\varepsilon(\mu)$ với hàm mật độ là $f(\tau) = \mu e^{-\mu\tau}$. μ cũng được gọi là cường độ phục vụ hay cường độ của “*dòng phục vụ*”. Hàm phân phối xác suất của T sẽ là

$$F(\tau) = P(T \leq \tau) = \int_0^{\tau} f(t) dt = \int_0^{\tau} \mu e^{-\mu t} dt = 1 - e^{-\mu\tau}.$$

Còn kì vọng toán và độ lệch chuẩn của biến ngẫu nhiên T là

$$m_T = \int_0^{+\infty} t f(t) dt = \int_0^{+\infty} \mu t e^{-\mu t} dt = \frac{1}{\mu}; \quad \delta_T = \frac{1}{\mu}.$$

Ta nhận thấy ngay rằng:

$$\begin{aligned} P(0 \leq T \leq \Delta t) &= F(\Delta t) - F(0) = 1 - e^{-\mu\Delta t} - [1 - e^0] \\ &= 1 - e^{-\mu\Delta t} = \mu \Delta t \quad (\text{chính xác tới vô cùng bé } o(\Delta t)). \end{aligned}$$

Chú ý: Nếu dòng tín hiệu đến có phân phối Poát-xông $\mathcal{P}(\lambda)$ thì thời gian giữa hai tín hiệu liên tiếp có phân phối mũ $\varepsilon(\lambda)$.

Chúng ta quay lại bài toán đang xét. Gọi λ_1 số lần chi tiết 1 hỏng và λ_2 số lần chi tiết 2 hỏng (tính trung bình) trên 1 đơn vị thời gian. Lúc đó, ta có thể coi dòng tín

hiệu chi tiết 1 và 2 hỏng là dòng Poát-xông với các tham số λ_1 và λ_2 . Gọi T_1 và T_2 là thời gian sửa chữa chi tiết 1 và 2, có phân phối mũ với các kì vọng tsc_1 và tsc_2 là thời gian sửa chữa (trung bình) chi tiết 1 và chi tiết 2. Vậy T_1 và T_2 có phân phối mũ $\varepsilon(\mu_1)$ và $\varepsilon(\mu_2)$, với $\mu_1 = 1/tsc_1$ và $\mu_2 = 1/tsc_2$.

Tại thời điểm t ta có biến ngẫu nhiên $X(t) = X_t$ với phân phối xác suất sau đây:

X_t	S_0	S_1	S_2	S_3
P	$\pi_0(t)$	$\pi_1(t)$	$\pi_2(t)$	$\pi_3(t)$

Ta tính $\pi_0(t + \Delta t)$ tại thời điểm tiếp theo $(t + \Delta t)$ trong hai trường hợp sau đây:

– Trường hợp 1: Tại thời điểm t , hệ thống ở trạng thái S_0 và tại thời điểm $t + \Delta t$, hệ thống vẫn ở trạng thái S_0 (không hỏng).

– Trường hợp 2: Tại thời điểm t , hệ thống ở trạng thái S_1 hoặc S_2 , còn tại thời điểm $t + \Delta t$ hệ thống ở trạng thái S_0 .

Do đó, $\pi_0(t + \Delta t) = \pi_0(t) [1 - (\lambda_1 + \lambda_2)\Delta t] + \pi_1(t) \mu_1 \Delta t + \pi_2(t) \mu_2 \Delta t$ (*). Thật vậy, xác suất do trường hợp 2 gây nên là $\pi_1(t)\mu_1 \Delta t + \pi_2(t)\mu_2 \Delta t$, với $\mu_1\Delta t = P(0 \leq T_1 \leq \Delta t)$ là xác suất sửa chữa xong chi tiết 1 trong khoảng thời gian Δt và $\mu_2 \Delta t = P(0 \leq T_2 \leq \Delta t)$ là xác suất sửa chữa xong chi tiết 2 trong khoảng thời gian Δt . Trong khi đó, xác suất do trường hợp 1 gây nên là $\pi_0(t)[1 - (\lambda_1 + \lambda_2)\Delta t]$, với $\lambda_1\Delta t$: xác suất hỏng chi tiết 1 trong khoảng Δt , còn $\lambda_2\Delta t$: xác suất hỏng chi tiết 2 trong khoảng Δt .

Nói cách khác, chúng ta đã thực hiện công thức xác suất đầy đủ $\pi_0(t + \Delta t) = \pi_0(t)p_{00} + \pi_1(t)p_{10} + \pi_2(t)p_{20}$, trong đó: p_{i0} là xác suất hệ ở trạng thái S_i tại thời điểm t và chuyển sang trạng thái S_0 tại thời điểm $(t + \Delta t)$.

$$\text{Từ (*) ta có: } \frac{\pi_0(t + \Delta t) - \pi_0(t)}{\Delta t} = \pi_1(t)\mu_1 + \pi_2(t)\mu_2 - \pi_0(t)\lambda_1 - \pi_0(t)\lambda_2.$$

Cho $\Delta t \rightarrow 0$ (về phải không liên quan với Δt) thì

$$\frac{d\pi_0(t)}{dt} = \pi_1(t)\mu_1 + \pi_2(t)\mu_2 - \pi_0(t)\lambda_1 - \pi_0(t)\lambda_2$$

Khi t rất lớn (hệ thống hoạt động trong một khoảng thời gian đủ dài) thì hệ thống dần ổn định với phân phối giới hạn có thể tìm được, tức là: $[\pi_0(t), \pi_1(t), \pi_2(t), \pi_3(t)] \rightarrow [\pi_0, \pi_1, \pi_2, \pi_3]$. Vậy ta có:

$$\pi_1\mu_1 + \pi_2\mu_2 - \pi_0\lambda_1 - \pi_0\lambda_2 = 0 \text{ (vì } \frac{d\pi_0(t)}{dt} = 0 \text{ khi } t \text{ đủ lớn).}$$

Một cách tương tự, ta đi đến hệ phương trình:

$$\begin{cases} \frac{d\pi_0}{dt} = \mu_1\pi_1 + \mu_2\pi_2 - (\lambda_1 + \lambda_2)\pi_0 = 0 \\ \frac{d\pi_1}{dt} = \lambda_1\pi_0 + \mu_2\pi_3 - (\lambda_2 + \mu_1)\pi_1 = 0 \\ \frac{d\pi_2}{dt} = \lambda_2\pi_0 + \mu_1\pi_3 - (\lambda_1 + \mu_2)\pi_2 = 0 \\ \frac{d\pi_3}{dt} = \lambda_2\pi_1 + \lambda_1\pi_2 - (\mu_1 + \mu_2)\pi_3 = 0 \end{cases}$$

Một cách tổng quát, phân phối giới hạn được tìm từ hệ phương trình: $-\pi_j q_{jj} = \sum_{i \neq j} \pi_i q_{ij}$ hay $\sum_{i \in S} \pi_i q_{ij} = 0, \forall j \in S$ (**), và $\sum_{i \in S} \pi_i = 1$, trong đó $-q_{ii}$ là cường độ chuyển từ trạng thái i sang các trạng thái khác (không kể i), còn q_{ij} là cường độ chuyển từ trạng thái i sang trạng thái j , được định nghĩa như sau:

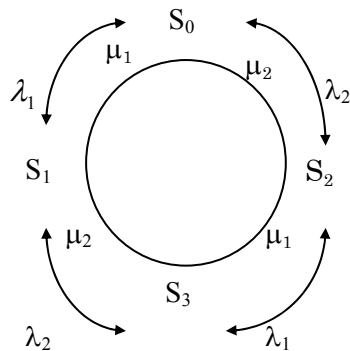
$$q_{ii} = -\lim_{\Delta t \rightarrow 0} (P[X(t + \Delta t) \neq i / X(t) = i] / \Delta t),$$

$$q_{ij} = \lim_{\Delta t \rightarrow 0} (P[X(t + \Delta t) = j / X(t) = i] / \Delta t),$$

Lúc đó, $Q = [q_{ij}]$ được gọi là ma trận cường độ. Từ điều kiện (**) ta thấy, để tìm phân phối giới hạn cần phải giải hệ $[\pi_0 \ \pi_1 \ \pi_2 \ \pi_3]Q = 0$ hay $Q^T[\pi_0 \ \pi_1 \ \pi_2 \ \pi_3]^T = 0$.

Ví dụ: Cho $\lambda_1 = 1, \lambda_2 = 2, \mu_1 = 2, \mu_2 = 3$. Từ sơ đồ cường độ chuyển trạng thái cho trên hình IV.2, có thể tìm được ma trận cường độ Q , với Q^T có dạng sau:

$$Q^T = \begin{bmatrix} -3 & 2 & 3 & 0 \\ 1 & -4 & 0 & 3 \\ 2 & 0 & -4 & 2 \\ 0 & 2 & 1 & -5 \end{bmatrix}$$



Hình IV.2. Sơ đồ cường độ chuyển trạng thái

Giải thích: $q_{00} = -3$ vì cường độ chuyển từ trạng thái S_0 sang các trạng thái khác

là $\lambda_1 + \lambda_2 = 3$, còn $q_{10} = 2$ là cường độ chuyển từ trạng thái S_1 vào trạng thái S_0 .

Giải hệ $[\pi_0 \ \pi_1 \ \pi_2 \ \pi_3]Q = 0$ hay $Q^T[\pi_0 \ \pi_1 \ \pi_2 \ \pi_3]^T = 0$ (với điều kiện bổ trợ $\pi_0 + \pi_1 + \pi_2 + \pi_3 = 1$) có kết quả: $\pi_0 = 6/15 = 0,4$; $\pi_1 = 3/15 = 0,2$; $\pi_2 = 4/15 = 0,27$; $\pi_3 = 2/15 = 0,13$. Cần chú ý rằng, hệ $[\pi_0 \ \pi_1 \ \pi_2 \ \pi_3]Q = 0$ theo một nghĩa nhất định là tương tự với hệ $\Pi \times (I - P) = 0$, như đã trình bày trong các mục 1.2 và 2.1.

Giả sử lợi nhuận / 1 đơn vị thời gian hệ thống mang lại trong các trường hợp có thể xảy ra như sau: nếu hệ thống trong trạng thái S_0 thì lợi nhuận là 8 USD, tại S_1 là 3 USD, tại S_2 là 5 USD, tại S_3 là 0 USD. Vậy lợi nhuận trung bình / 1 đơn vị thời gian là

$$8 \times 0,4 + 3 \times 0,2 + 5 \times 0,27 = 5,15 \text{ (USD)}.$$

Qua ví dụ ta thấy $\pi_0, \pi_1, \pi_2, \pi_3$ được xác định căn cứ vào các giá trị đã biết $\lambda_1, \lambda_2, \mu_1, \mu_2$.

λ_1, λ_2 : số lần chi tiết hỏng (tùy thuộc hệ thống cụ thể),

μ_1, μ_2 : các tham số sửa chữa cần đưa vào.

Lợi nhuận cuối cùng của hệ thống phụ thuộc vào $\lambda_1, \lambda_2, \mu_1, \mu_2$ và được xác định bằng cách giải bài toán tối ưu sau:

$$\text{Lợi nhuận } L = c_0\pi_0 + c_1\pi_1 + c_2\pi_2 \rightarrow \text{Max}$$

(c_0, c_1, c_2 : lợi nhuận từng trạng thái)

với các ràng buộc:

$$\begin{cases} \mu_1\pi_1 + \mu_2\pi_2 - (\lambda_1 + \lambda_2)\pi_0 = 0 \\ \lambda_1\pi_0 + \mu_2\pi_3 - (\lambda_2 + \mu_1)\pi_1 = 0 \\ \lambda_2\pi_0 + \mu_1\pi_3 - (\lambda_1 + \mu_2)\pi_2 = 0 \\ \lambda_2\pi_1 + \lambda_1\pi_2 - (\mu_1 + \mu_2)\pi_3 = 0 \\ \pi_0 + \pi_1 + \pi_2 + \pi_3 = 1 \\ \pi_0, \pi_1, \pi_2, \pi_3 \geq 0; \mu_1, \mu_2 \geq 0 \end{cases}$$

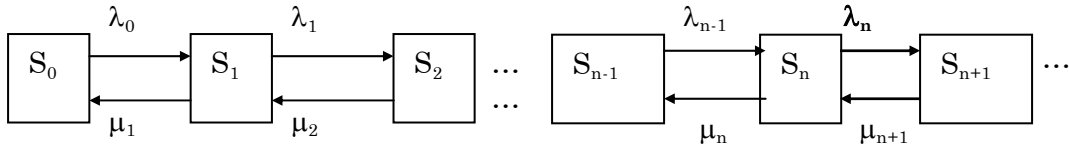
Lưu ý: Bài toán trên đây có 6 biến (λ_1, λ_2 đã biết). Ta phải tìm được μ_1, μ_2 từ bài toán để có phương hướng xây dựng hệ thống với lợi nhuận lớn nhất.

2.5. Một ứng dụng của quá trình sinh-tử cho hệ thống hàng chờ

Quá trình sinh-tử được ứng dụng khá rộng rãi trong Lí thuyết độ tin cậy, là một môn học của ngành Điện / Điện tử, và một số ngành khoa học kĩ thuật khác cũng như trong Quản trị kinh doanh và Vận trù học.

Quá trình sinh - tử là trường hợp riêng của xích Markov thuần nhất thời gian liên tục, với không gian trạng thái S không quá đếm được $S = \{S_0, S_1, S_2, \dots, S_n, \dots\}$ và ma trận cường độ $Q = [q_{ij}]$ có tính chất $q_{ij} = 0$ với $|i - j| \geq 2$. Điều này có nghĩa là việc

chuyển trạng thái trong quá trình sinh–tử chỉ có thể tới “1 đơn vị lên hoặc xuống” (xem hình IV.3).



Hình IV.3. Sơ đồ chuyển trạng thái trong quá trình sinh - tử

Từ trạng thái S_n tại thời điểm t hệ $X(t)$ chỉ có thể chuyển tới một trong các trạng thái S_{n+1} , S_n hoặc S_{n-1} . Vì vậy chúng ta có các cường độ chuyển:

$$\mu_0 = \lambda_{-1} = 0 = q_{00}, q_{n, n+1} = \lambda_n, q_{n, n-1} = \mu_n \text{ và } q_{n, n} = -(\lambda_n + \mu_n) \forall n.$$

Trong trường hợp $\lambda_n, \mu_n > 0, \forall n > 0$, theo định lí đã được chứng minh, phân phối giới hạn có thể tìm được bằng cách giải hệ: $[\pi_0 \pi_1 \pi_2 \pi_3 \dots]Q = 0$, với ma trận cường độ Q đã biết.

Ma trận chuyển vị của Q có dạng:

$$Q^T = \begin{bmatrix} q_{00} & q_{10} & \dots & q_{n0} & q_{n+1,0} & \dots \\ q_{01} & q_{11} & \dots & q_{n1} & q_{n+1,1} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ q_{0n} & q_{1n} & \dots & q_{nn} & q_{n+1,n} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

Ta có $[\pi_0 \pi_1 \pi_2 \pi_3 \dots]Q = 0 \Leftrightarrow Q^T[\pi_0 \pi_1 \pi_2 \pi_3 \dots]^T = 0 \Leftrightarrow$

$$\begin{bmatrix} q_{00} & q_{10} & q_{20} & \dots \\ q_{01} & q_{11} & q_{21} & \dots \\ q_{02} & q_{12} & q_{22} & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix} \times \begin{bmatrix} \pi_0 \\ \pi_1 \\ \pi_2 \\ \dots \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \dots \end{bmatrix}.$$

hay:

$$\begin{cases} q_{00}\pi_0 + q_{10}\pi_1 + q_{20}\pi_2 + \dots = 0, \\ q_{01}\pi_0 + q_{11}\pi_1 + q_{21}\pi_2 + \dots = 0, \\ q_{02}\pi_0 + q_{12}\pi_1 + q_{22}\pi_2 + \dots = 0, \\ \dots \end{cases}$$

Do tính chất đặc biệt, như đã phân tích ở trên, của ma trận cường độ Q của quá trình sinh–tử, hệ trên được viết một cách tường minh hơn như sau:

$$\begin{cases} -\lambda_0\pi_0 + \mu_1\pi_1 + \dots = 0, \\ \lambda_0\pi_0 - (\lambda_1 + \mu_1)\pi_1 + \mu_2\pi_2 + \dots = 0, \\ \lambda_1\pi_1 - (\lambda_2 + \mu_2)\pi_2 + \mu_3\pi_3 + \dots = 0, \\ \dots \end{cases}$$

Từ đây dễ dàng tìm được $\pi_{n+1} = (\lambda_n/\mu_{n+1})\pi_n, \forall n = 1, 2, 3, \dots$ để đi tới công thức tính $\pi_i, \forall i$.

$$\begin{cases} \pi_1 = (\lambda_0/\mu_1)\pi_0, \\ \pi_2 = (\lambda_1/\mu_2)\pi_1 = (\lambda_1\lambda_0/\mu_2\mu_1)\pi_0, \\ \pi_3 = (\lambda_2/\mu_3)\pi_2 = (\lambda_2\lambda_1/\mu_3\mu_2)\pi_1 = (\lambda_2\lambda_1\lambda_0/\mu_3\mu_2\mu_1)\pi_0, \\ \dots \\ \pi_{n+1} = (\lambda_n/\mu_{n+1})\pi_n = \dots = (\lambda_n\lambda_{n-1} \dots \lambda_0/\mu_{n+1}\mu_n \dots \mu_1)\pi_0, \\ \dots \end{cases}$$

Với điều kiện $\sum_{i=0}^{\infty} \pi_i = 1$, cuối cùng ta có:

$$\pi_0 = 1 / \left(1 + \sum_{k=0}^{\infty} (\lambda_k\lambda_{k-1} \dots \lambda_0/\mu_{k+1}\mu_k \dots \mu_1) \right).$$

Đặc biệt khi $\mu_n = 0, \forall n$, thì quá trình sinh–tử trở thành quá trình sinh thuần khiết (*pure birth process*). Quá trình sinh thuần khiết với $\lambda_n = \lambda$ là quá trình Poát–xông với tham số λ .

Ví dụ: Giả sử dòng khách hàng đến mua vé ở một văn phòng bán vé với M quầy phục vụ là dòng Poát–xông với tham số $\lambda = 6$ khách hàng / 1 phút (điều này cũng có nghĩa là khách hàng đến phòng bán vé với các thời điểm đến tuân theo luật phân phối mũ với tham số $\lambda = 6$).

Ngoài ra, còn biết nguyên tắc phục vụ là FCFS (*First come first serve*) và thời gian phục vụ tại mỗi quầy có luật phân phối mũ với kì vọng 1/3 (phút).

Cần trả lời hai câu hỏi sau đây:

– Số quầy hàng tối thiểu là bao nhiêu để hàng chờ không trở nên dài vô hạn?

– Giả sử N_t là số khách hàng đang chờ hay đang được phục vụ tại thời điểm t.

Chọn $M = 4$ và một khách hàng sẽ chờ để được phục vụ nếu $N_t \leq 4$, chờ với xác suất 0,5 nếu

$N_t = 5$ và sẽ bỏ đi nếu $N_t = 6$. Hãy xác định phân phối dừng của quá trình này?

Trước hết, trong ví dụ này chúng ta có một quá trình sinh–tử với không gian trạng

thái $S = \{S_0, S_1, S_2, \dots, S_n, \dots\}$, trong đó S_n là trạng thái trong văn phòng có n khách hàng. Các cường độ chuyển là $\lambda_k = 6$ với $k = 0, 1, 2, \dots$ còn $\mu_k = 3k$ với $k \leq M$ và $\mu_k = 3M$ với $k > M$ (điều này là do biến cực tiểu của các biến ngẫu nhiên với phân phối mũ độc lập cũng có phân phối mũ với tham số bằng tổng các tham số của các phân phối mũ tương ứng).

Do $\lambda_k / \mu_{k+1} = 6/3M < 1$ (khi $k \geq M$) nên với $M \geq 3$ thì:

$$\sum_{k=0}^{\infty} (\lambda_k \lambda_{k-1} \dots \lambda_0 / \mu_{k+1} \mu_k \dots \mu_1) < \infty.$$

Bởi vậy hàng đợi sẽ không dài vô hạn (nếu trái lại, khi chuỗi phân kì, thì $\pi_0 = \pi_1 = \pi_2 = \dots = 0$, nên số khách trong hàng đợi sẽ dần tới một số hữu hạn khi $t \rightarrow \infty$ với xác suất bằng 0).

Trong câu hỏi thứ hai, ta có $\lambda_0 = \lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 6, \lambda_5 = 3$. Theo công thức tính $\pi_0 = 1 / (1 + \sum_{k=0}^5 (\lambda_k \lambda_{k-1} \dots \lambda_0 / \mu_{k+1} \mu_k \dots \mu_1))$ ta có ngay $\pi_0 = 12/89$. Từ đó tính ra $\pi_1 = 24/89, \pi_2 = 24/89, \pi_3 = 16/89, \pi_4 = 8/89, \pi_5 = 4/89$ và $\pi_6 = 1/89$.

3. Mô phỏng xích Markov

3.1. Mô phỏng xích Markov thời gian rời rạc

Phương pháp 1

Xích Markov rời rạc và thuần nhất còn có thể được kí hiệu là X_0, X_1, X_2, \dots . Giả sử không gian trạng thái là S gồm hữu hạn trạng thái: $S = \{0, 1, 2, \dots, N\}$ và ma trận xác suất chuyển trạng thái đã được biết là $P = [p_{ij}]_{N \times N}$. Chúng ta sẽ mô phỏng xích Markov rời rạc và thuần nhất thông qua ví dụ đã trình bày ở các mục 1.2 và 2.1 của chương này.

Ta có phân phối ban đầu là:

X_0	1	2	3
$\Pi^{(0)}$	$\pi_1^{(0)} = 0,2$	$\pi_2^{(0)} = 0,5$	$\pi_3^{(0)} = 0,3$

Để mô phỏng X_0 ta áp dụng phương pháp mô phỏng phân phối rời rạc đã học ở chương III. Trên máy tính, ta phát sinh ra một số ngẫu nhiên $r = \text{RANDOM}[0,1)$ theo luật phân phối đều $U[0,1)$ trong $[0,1)$. Nếu $r \leq 0,2$ ta lấy $X_0 = 1$; nếu $0,2 < r \leq 0,7$ thì ta lấy $X_0 = 2$; còn nếu $r > 0,7$ thì đặt $X_0 = 3$. Căn cứ kết quả mô phỏng X_0 , ta mô phỏng X_1 dựa trên ma trận xác suất chuyển trạng thái:

$$P = \begin{bmatrix} 0,8 & 0,1 & 0,1 \\ 0,07 & 0,9 & 0,03 \\ 0,083 & 0,067 & 0,85 \end{bmatrix}.$$

Giả sử đã biết $X_0 = 2$, lúc đó ta cần mô phỏng biến ngẫu nhiên X_1 căn cứ phân phối sau:

X_1	1	2	3
Xác suất tương ứng	$p_{21} = 0,07$	$p_{22} = 0,9$	$p_{23} = 0,03$

Điều này có thể được thực hiện tương tự như khi mô phỏng X_0 . Cần chú ý rằng, trong hàng thứ hai của bảng trên ta có phân phối xác suất có điều kiện của X_1 với điều kiện $X_0 = 2$. Các bước tiếp theo mô phỏng X_2, X_3, \dots được tiến hành tương tự (cho tới X_{500} chẳng hạn).

Lặp lại quy trình này bắt đầu từ X_0 cho một số bước lặp L đủ lớn (chẳng hạn 1000 lần), ta sẽ có một bộ 1000 số liệu cho X_{500} . Từ đó, có thể tìm được bảng phân phối tần suất (còn gọi là xác suất thực nghiệm) của X_{500} qua thí nghiệm mô phỏng trên đây đối với X_{500} . Như vậy, ta tìm được véc tơ phân phối (xác suất thực nghiệm) $\Pi^{(500)}$. Cuối cùng, chúng ta có kết quả tìm gần đúng phân phối đúng là: $\Pi \approx \Pi^{(500)}$.

Chú ý:

– Trong ví dụ trên đây, ta thấy có thể dùng mô phỏng để tìm phân phối đúng. Tuy nhiên, mục đích chủ yếu của phương pháp 1 là nhằm mô phỏng các xích Markov rời rạc thuần nhất, là các quá trình có thể xảy ra trong các hệ thống phức tạp.

– Khi không gian trạng thái S gồm một số lớn các trạng thái thì phương pháp mô phỏng trên yêu cầu thời gian chạy máy tính khá lớn. Để khắc phục điều này, chúng ta xem xét phương pháp 2 sau đây.

Phương pháp 2

Xét một hệ thống kỹ thuật được biểu diễn bởi xích Markov rời rạc thuần nhất $\{X_t\}$, $t = 0, 1, 2, \dots$ với không gian trạng thái S có N trạng thái (N khá lớn) và ma trận chuyển trạng thái $P = [p_{ij}]_{N \times N}$. Xét thời điểm n , tại thời điểm này giả sử đã mô phỏng được $X_n = s$. Ta sẽ mô phỏng thời gian T_n là thời gian tới lần nhảy tiếp theo sớm nhất mà $X_{n+T_n} \neq s$. Do xích Markov là rời rạc nên T_n chỉ có thể nhận các giá trị $1, 2, \dots$. Đặt $p = p_{ss}$, dễ thấy T_n có phân phối hình học như sau:

T_n	1	2	...	k	...
Xác suất tương ứng	$1-p$	$(1-p)p$...	$(1-p)p^{k-1}$...

Mô phỏng phân phối này ta tìm được giá trị T_n . Còn X_{n+T_n} có phân phối xác suất như sau:

X_{n+T_n}	1	2	...	s	...	N
Xác suất tương ứng	$p_{s1}/(1-p_{ss})$	$p_{s2}/(1-p_{ss})$...	0	...	$p_{sN}/(1-p_{ss})$

Cách mô phỏng này sẽ tiết kiệm hơn thời gian chạy máy tính (khi N khá lớn), nhưng việc lập trình sẽ phức tạp hơn ít nhiều.

Xét ví dụ như đã trình bày trên, nếu dùng phương pháp 2, một cách hoàn toàn tương tự, chúng ta cũng tìm được phân phối dừng $\Pi^{(*)} \approx \Pi^{(500)}$.

3.2. Mô phỏng xích Markov thời gian liên tục

Xét xích Markov thời gian liên tục $\{X(t)\}_{t \in [0, \infty)}$. Giả sử rằng xích đi vào trạng thái i tại thời điểm nào đó, chẳng hạn thời điểm 0, và không rời khỏi trạng thái này cho đến thời điểm s . Lúc đó, do tính “không nhớ” của quá trình Markov, xác suất để xích vẫn tiếp tục ở nguyên trạng thái đó cho tới thời điểm $(t + s)$ sẽ là:

$$P\{(T_i > s + t)/(T_i > s)\} = P\{T_i > t\}$$

trong đó T_i là thời gian quá trình dừng lại ở trạng thái i . Dễ thấy, nếu T_i có phân phối mũ với hàm phân phối $F(T_i < \tau) = 1 - e^{-\lambda\tau}$ thì đẳng thức trên được thoả mãn. Điều ngược lại cũng có thể chứng minh được. Vậy T_i có phân phối mũ.

Từ nhận xét trên, ta có thể đưa ra một *định nghĩa khác cho xích Markov thời gian liên tục*. Xích Markov thời gian liên tục là một quá trình ngẫu nhiên có các tính chất sau mỗi khi nó đi vào trạng thái i :

– Lượng thời gian T_i xích dừng lại tại trạng thái i trước khi nó chuyển sang trạng thái khác là một biến ngẫu nhiên với phân phối mũ có tham số v_i (hay có kì vọng $1/v_i$).

– Một khi quá trình rời khỏi trạng thái i , nó sẽ đi vào trạng thái j nào đó (độc lập với T_i) với các xác suất p_{ij} thoả mãn $\sum_j p_{ij} = 1, p_{ii} = 0, \forall i$.

Vậy để mô phỏng xích Markov thời gian liên tục, chúng ta cần mô phỏng dãy $\tau_0, \tau_1, \tau_2, \dots$ (các lượng thời gian τ_r xích dừng lại tại trạng thái J_r trước khi nó chuyển sang trạng thái khác) và dãy J_0, J_1, J_2, \dots (các trạng thái mà xích chuyển đến). Để phát sinh τ_r , như trên đã nói, ta cần biết tham số v_{J_r} của phân phối mũ tương ứng. Còn để phát sinh trạng thái xích Markov chuyển đến $J_r \forall r$, chúng ta có bảng phân phối xác suất sau:

Trạng thái đến	1	2	...	i	...	N
Xác suất tương ứng	p_{i1}	p_{i2}	...	0	...	p_{iN}

Trong bảng trên, $i = J_{r-1}$ là trạng thái của xích tại bước $r - 1$ (với các xác suất p_{ij} thoả mãn $\sum_j p_{ij} = 1, p_{ii} = 0, \forall i$).

Để thực hiện mô phỏng xích Markov thời gian liên tục, có thể sử dụng số liệu của ví dụ đã xét trong mục 2.4 hay 2.5.