Mục lục

Chương 1 . Bài toán tối ưu 1

Chương 2 . Một số khái niệm và kết quả cơ bản của Giải tích lồi 3

Chương 3 . Qui hoạch tuyến tính
(Linear Programming) 5

Chương 4 . Qui hoạch phi tuyến
(Nonlinear Programming) 7

4.1 Hàm lồi 7
4.1.1 Định nghĩa 7
4.1.2 Các phép toán về hàm lồi 11
4.1.3 Tính liên tục của hàm lồi 12
4.1.4 Đạo hàm theo hướng của hàm lồi 12
4.1.5 Tiêu chuẩn nhận biết hàm lồi khả vi 15
4.1.6 Cực trị của hàm lồi 16
4.2 Bài toán quy hoạch phi tuyến không ràng buộc 19
4.2.1 Điều kiện tối ưu 19
4.2.2 Phương pháp hướng giảm 22
4.2.3 Phương pháp gradient 29
4.2.4 Phương pháp Newton 34
4.2.5 Cực tiểu hàm một biến 42
4.2.6 Phương pháp tìm kiếm trực tiếp 45
4.3 Bài toán quy hoạch phi tuyến có ràng buộc 50
4.3.1 Điều kiện tối ưu 50
4.3.2 Phương pháp nhân tử Lagrange 60
4.3.3 Phương pháp tuyến tính hóa giải quy hoạch lồi 67
4.3.4 Phương pháp hướng có thể giải bài toán cực tiểu hàm trơn với ràng buộc tuyến tính 73
4.3.5 Phương pháp Frank-Wolfe giải bài toán quy hoạch lồi với ràng buộc tuyến tính 76
4.3.6 Phương pháp hàm phạt 79

**Chương 5 . Tính chất của tập nghiệm,
điều kiện tối ưu và đối ngẫu 89**

Chương 6 . Qui hoạch toàn cục 91

Chương 1

Bài toán tối ưu

1.1 Bài toán tối ưu và các khái niệm cơ bản

Xét bài toán tối ưu

$$\min\ f(x) \text{ v.đ.k. } x \in D, \qquad (P_1)$$

trong đó:

$D \subseteq \mathbb{R}^n$ - *tập nghiệm chấp nhận được* hay *tập ràng buộc*

$f : D \to \mathbb{R}$ - *hàm mục tiêu*

$x \in D$ - *nghiệm chấp nhận được* hay *một phương án chấp nhận được* (có thể gọi tắt là *một phương án*)

Định nghĩa 1.

i) Điểm $x^* \in \mathbb{R}^n$ được gọi là *nghiệm tối ưu* của bài toán (P_1) nếu

$x^* \in D$;
$f(x^*) \leq f(x) \ \ \forall x \in D.$

ii) Điểm $x^* \in \mathbb{R}^n$ được gọi là *nghiệm tối ưu chặt* của bài toán (P_1) nếu

$x^* \in D$;
$f(x^*) < f(x) \ \ \forall x \in D.$

iii) Ký hiệu $\mathrm{Argmin}(P_1)$ (hoặc $\mathrm{Argmin}\{f(x) : x \in D\}$ - *tập nghiệm tối ưu*

Định nghĩa 2. Ta gọi $t_0 = \inf f(D)$ là *giá trị tối ưu* của bài toán (P_1), tức:

hoặc

$f(x) \geq t_0 \quad \forall x \in D$;
$\exists \{x^k\} \subset D$ sao cho $\lim_{k\to\infty} f(x^k) = t_0$.

hoặc

$\forall \varepsilon > 0$, tồn tại $\bar{x} \in D$ sao cho $f(\bar{x}) < t_0 + \varepsilon$.

Nếu $\text{Argmin}(P_1) \neq \emptyset$ thì giá trị tối ưu $t_0 = f(x^*)$, trong đó $x^* \in \text{Argmin}(P_1)$. Khi đó, giá trị tối ưu có thể được ký hiệu là

$$\min\{f(x) : x \in D\}.$$

Định nghĩa 3.

i) Điểm $x^* \in \mathbb{R}^n$ được gọi là *nghiệm tối ưu địa phương* của bài toán (P_1) nếu:

$x^* \in D$;
$\exists$ lân cận $U(x^*)$ sao cho $f(x^*) \leq f(x) \ \ \forall x \in U(x^*) \cap D$.

iii) Điểm $x^* \in \mathbb{R}^n$ được gọi là *nghiệm tối ưu địa phương chặt* của bài toán (P_1) nếu:

$x^* \in D$;
$\exists$ lân cận $U(x^*)$ sao cho $f(x^*) \leq f(x) \ \ \forall x \in U(x^*) \cap D$.

Chú ý 1.

$x^* \in \text{Argmin}\{f(x) : x \in D\} \quad \Rightarrow \quad x^*$ là nghiệm tối ưu địa phương

$x^* \in \text{Argmin}\{f(x) : x \in D\} \quad \not\Leftarrow \quad x^*$ là nghiệm tối ưu địa phương

Chú ý 2.

$$\text{Argmin}\{f(x) : x \in D\} = \text{Argmax}\{-f(x) : x \in D\}$$

và

$$\min\{f(x) : x \in D\} = -\max\{-f(x) : x \in D\}.$$

1.2 Điều kiện tồn tại nghiệm

Định lý 1.1. *Cho $D \subset \mathbb{R}^n$ là tập compac khác rỗng. Khi đó,*
i) *Nếu hàm f nửa liên tục dưới trên D thì* $\operatorname{Argmin}\{f(x) : x \in D\} \neq \emptyset$;
ii) *Nếu hàm f nửa liên tục trên trên D thì* $\operatorname{Argmax}\{f(x) : x \in D\} \neq \emptyset$.

Hệ quả 1.1. (Định lý Weierstrass) *Nếu $D \subset \mathbb{R}^n$ là tập compac khác rỗng và hàm f liên tục trên D thì*

$$\operatorname{Argmin}\{f(x) : x \in D\} \neq \emptyset \quad \text{và} \quad \operatorname{Argmax}\{f(x) : x \in D\} \neq \emptyset.$$

Định lý 1.2. *Cho tập đóng khác rỗng $D \subseteq \mathbb{R}^n$. Nếu hàm f nửa liên tục dưới trên D và thỏa mãn điều kiện bức (coercive) trên D,*

$$f(x) \to +\infty \text{ khi } x \in D, \text{ và } \|x\| \to +\infty$$

thì bài toán $\min\{f(x) : x \in D\}$ *có nghiệm tối ưu.*

Bài tập

1. Giải bài toán sau

$$\max\{x_1^2 + x_2^2 \mid 3x_1 + x_2 \leq 15,\ 2x_1 - 3x_2 \leq 6,\ x_2 \leq 10,\ x_1, x_2 \geq 0\}.$$

2. Xét bài toán min $\{f(x) \mid x \in \mathbb{Z}\}$, trong đó $\mathbb{Z}$ là tập các số nguyên. Chứng minh rằng mỗi điểm $x \in \mathbb{Z}$ đều là nghiệm cực tiểu địa phương.

3. Cho $X = [a, b] \subset \mathbb{R}$ và f là hàm thực xác định trên X. Xét bài toán

$$\min\ f(x) \ \text{ v.đ.k. } \ x \in X. \qquad (VD_1)$$

Chứng minh rằng:

i) Nếu $x = a$ là nghiệm tối ưu địa phương của bài toán (VD_1) và tồn tại ${f_+}'(a)$ thì $f'_+(a) \geq 0$.

ii) Nếu $x = b$ là nghiệm tối ưu địa phương của bài toán (VD_1) và tồn tại $f'_-(b)$ thì $f'_-(b) \leq 0$.

4. Điểm $(0,0)^T$ có phải là nghiệm cực tiểu địa phương, nghiệm cực tiểu toàn cục của bài toán $\min\{f(x)|x \in \mathbb{R}^2\}$ không, trong đó:

   i) $f(x_1, x_2) = \sqrt{x_1^2 + x_2^2}$.
   ii) $f(x_1, x_2) = x_1^2 + x_2^2(2 - x_1)^3$.

   Giải thích?

5. Cho $h(x) = x_1^2 + 2x_2^2 + 3x_3^2 + \cdots + nx_n^2$ và siêu phẳng

$$Q = \{x \in \mathbb{R}^n | \langle v, x \rangle = \alpha, v \in \mathbb{R}^n \setminus \{0\}, \alpha \in \mathbb{R}\}.$$

   Bài toán $\min\{h(x)|x \in Q\}$ có nghiệm tối ưu không? Vì sao?

Chương 2

Một số khái niệm và kết quả cơ bản của Giải tích lồi

2.1 Tập afin

Cho x^1, x^2 là hai điểm trong $\mathbb{R}^n$. *Đường thẳng* qua x^1 và x^2 là tập

$$\begin{aligned}\Gamma(a,b) &= \{x \in \mathbb{R}^2 : x = \lambda x^1 + (1-\lambda)x^2, \quad \lambda \in \mathbb{R}\}\\ &= \{x \in \mathbb{R}^2 : x = x^2 + \lambda(x^1 - x^2), \quad \lambda \in \mathbb{R}\}\end{aligned}$$

Tập $M \subseteq \mathbb{R}^n$ được gọi là *tập afin* (affine set) nếu M chứa trọn cả đường thẳng đi qua hai điểm bất kỳ của M, nghĩa là

$$\forall x^1, x^2 \in M,\ \lambda \in \mathbb{R} \ \Rightarrow\ \lambda x^1 + (1-\lambda)x^2 \in M.$$

$\diamond$ Cho k điểm $x^1, x^2, \ldots, x^k \in \mathbb{R}^n$. Điểm

$$x = \sum_{i=1}^{k} \lambda_i x^i \quad \text{với} \ \ \lambda_1, \ldots, \lambda_k \in \mathbb{R} \ \ \text{và} \ \ \sum_{i=1}^{k} \lambda_i = 1$$

được gọi là *tổ hợp afin* của các điểm $x^1, x^2, \ldots, x^k \in \mathbb{R}^n$.

Mệnh đề 2.1. *Tập $M \subseteq \mathbb{R}^n$ là tập afin khi và chỉ khi*

$$M = x^0 + L = \{x^0 + v \mid v \in L\}, \tag{2.1}$$

trong đó $x^0 \in M$ và L là không gian con của $\mathbb{R}^n$.

Chú ý 1. Ta gọi L trong (2.1) là *không gian con song song* với M vì:

i) Không gian con L trong (2.1) không phụ thuộc vào cách chọn $x^0 \in M$.

ii) Không gian con L này xác định duy nhất.

Thứ nguyên (dimension) hay *số chiều* của tập afin M, ký hiệu là $\dim M$, là thứ nguyên của không gian con song song với nó.

Ví dụ 2.1. Cho A là ma trận cấp $m \times n$ và véc tơ $b \in \mathbb{R}^m$ và

$$M = \{x \in \mathbb{R}^n : Ax = b\}.$$

Chứng minh M là tập afin.

Chứng minh. Lấy tùy ý hai điểm x^1, $x^2 \in M$ và số thực bất kỳ $\lambda \in \mathbb{R}$. Ta sẽ chứng minh

$$x^0 = \lambda x^1 + (1-\lambda)x^2 \in M.$$

Thật vậy, do $x^1, x^2 \in M$ nên

$$\begin{aligned} A(\lambda x^1 + (1-\lambda)x^2) &= \lambda A x^1 + (1-\lambda)Ax^2 \\ &= \lambda b + (1-\lambda)b \\ &= b. \end{aligned}$$

Suy ra $x^0 \in M$. Theo định nghĩa, M là một tập afin. Dễ thấy

$$M = x^0 + L, \quad x^0 \in M, \quad L = \mathrm{Ker} A := \{x \in \mathbb{R}^n \mid Ax = 0\}.$$

Vì vậy,

$$\dim M = \dim \mathrm{Ker} A = n - \mathrm{rank} A,$$

trong đó $\mathrm{rank} A$ là hạng của ma trận A.

Bao afin (affine hull) của một tập $E \subseteq \mathbb{R}^n$, ký hiệu là aff$E$, là giao của tất cả các tập afin chứa E. Đó là tập afin nhỏ nhất chứa E.

Ví dụ 2.2.

i) $E_1 = \{x^1, x^2\} \quad \Rightarrow \quad \text{aff } E_1 = \{x \in \mathbb{R}^3 \mid x = \lambda x^1 + (1-\lambda)x^2,\ \lambda \in \mathbb{R}\}.$

ii) $E_2 = \{x \in \mathbb{R}^3 \mid -1 \leq x_1 \leq 1,\ -1 \leq x_2 \leq 1,\ x_3 = 0\}$

$\Rightarrow \quad \text{aff}E_2 = \{x \in \mathbb{R}^3 \mid x_3 = 0\}.$

iii) $E_3 = \{x \in \mathbb{R}^n \mid \|x\| \leq 3\} \quad \Rightarrow \quad \text{aff}E_3 = \mathbb{R}^n.$

iv) $E_4 = \{x \in \mathbb{R}^n \mid \|x\| = 5\} \quad \Rightarrow \quad \text{aff}E_4 = \mathbb{R}^n.$

2.2 Số chiều và điểm trong tương đối

Cho tập tùy ý $E \subset \mathbb{R}^n$. Ta định nghĩa

$$\dim E = \dim(\text{aff}E).$$

Cho tập $M \subset \mathbb{R}^n$ có dim$M < n$. Một điểm $a \in M$ được gọi là *điểm trong tương đối* (relative interior point) của M nếu

$$\exists B(a,\varepsilon)\ :\ \Big(B(a,\varepsilon) \cap \text{aff}M\Big) \subset M.$$

Phần trong tương đối của tập M, kí hiệu là ri$M$, là tập tất cả các điểm trong tương đối của M.

Một tập $M \subseteq \mathbb{R}^n$ được gọi là có *thứ nguyên đầy đủ* nếu dim$M = n$.

Chú ý.

$$\text{int}M \neq \emptyset \quad \Rightarrow \quad \dim M = n.$$
$$\text{int}M \neq \emptyset \quad \not\Leftarrow \quad \dim M = n.$$

Ví dụ 2.3. Cho

$E_1 = \{a, b\} \subset \mathbb{R}^3, \qquad E_2 = \{x \in \mathbb{R}^3 \mid 0 \le x_i \le 1, i = 1, 2\ x_3 = 0\}$
$E_3 = \{x \in \mathbb{R}^3 \mid \|x\| \le 1\}, \qquad E_4 = \{x \in \mathbb{R}^3 \mid \|x\| = 1\}.$

Ta có:

i) $\mathrm{int}E_1 = \emptyset$, $\mathrm{ri}E_1 = \emptyset$, $\dim E_1 = 1$.

ii) $\mathrm{int}E_2 = \emptyset$, $\mathrm{ri}E_2 = \{x \in \mathbb{R}^3 | 0 < x_1 < 1, 0 < x_2 < 1, x_3 = 0\}$ và $\dim E_2 = 2$.

iii) $\mathrm{int}E_3 = \{x \in \mathbb{R}^3 \mid \|x\| < 1\}, \dim E_3 = 3$, tức là E_3 có thứ nguyên đầy đủ.

iv) $\mathrm{int}E_4 = \emptyset$, $\dim E_4 = 3$, tức là E_4 có thứ nguyên đầy đủ.

2.3 Tập lồi

Cho hai điểm x^1 và x^2 thuộc $\mathbb{R}^n$. *Đoạn thẳng* nối x^1 và x^2 là

$$\begin{aligned}[x^1, x^2] &= \{x \in \mathbb{R}^2 : x = \lambda x^1 + (1 - \lambda)x^2, \quad \lambda \in [0, 1]\} \\ &= \{x \in \mathbb{R}^2 : x = x^2 + \lambda(x^1 - x^2), \quad \lambda \in [0, 1]\}\end{aligned}$$

Tập $M \subseteq \mathbb{R}^n$ được gọi là *tập lồi* (convex set) nếu

$$\forall x^1, x^2 \in M, \quad \Rightarrow \quad [x^1, x^2] \subset M$$

tức

$$\forall x^1, x^2 \in M, \forall \lambda \in [0, 1] \ \Rightarrow \ \lambda x^1 + (1 - \lambda)x^2 \in M.$$

Chú ý: Lược đồ chung để chứng minh một tập $Q \subset \mathbb{R}^n$ là tập lồi:

- $\diamond$ Lấy tùy ý $q^1, q^2 \in Q$ và số thực bất kỳ $t \in [0, 1]$.
- $\diamond$ Ta sẽ chứng minh $\bar{q} = tq^1 + (1 - t)q^2 \in K$.
- $\diamond$ Thật vậy, ...

Ví dụ 2.4. Cho A là ma trận cấp $m \times n$ và véc tơ $b \in \mathbb{R}^m$ và

$$M = \{x \in \mathbb{R}^n : Ax \le b,\ x \ge 0\}.$$

Giả sử $M \neq \emptyset$. Chứng minh M là tập lồi.

Chứng minh. Lấy tùy ý hai điểm x^1, $x^2 \in M$ và số thực bất kỳ $\lambda \in [0, 1]$. Ta sẽ chứng minh

$$x^0 = \lambda x^1 + (1 - \lambda)x^2 \in M.$$

Thật vậy, do $x^1, x^2 \in M$ nên

$$Ax^1 \le b,\ Ax^2 \le b,\ x^1 \ge 0 \text{ và } x^2 \ge 0.$$

Ta có

$$A(\lambda x^1 + (1 - \lambda)x^2) = \lambda Ax^1 + (1 - \lambda)Ax^2 \le b \tag{2.2}$$

và

$$x^0 = \lambda x^1 + (1 - \lambda)x^2 \ge 0. \tag{2.3}$$

Kết hợp (2.2) và (2.3) suy ra $x^0 \in M$. Theo định nghĩa, M là một tập lồi. □

Mệnh đề 2.2. *Cho hai tập lồi $M_1,\ M_2 \subset \mathbb{R}^n$ và số thực $\alpha \in \mathbb{R}$. Khi đó, các tập sau đây là tập lồi:*

i) $M_1 \cap M_2$;

ii) $\alpha M_1 = \{y \mid y = \alpha x,\ x \in M_1\}$;

iii) $M_1 + M_2 = \{x \mid x = x^1 + x^2,\ x^1 \in M_1, x^2 \in M_2\}$;

iv) $M_1 - M_2 = \{x \mid x = x^1 - x^2,\ x^1 \in M_1, x^2 \in M_2\}$.

Chú ý. Hợp của các tập lồi chưa chắc là tập lồi.

Tổ hợp lồi

◇ Điểm x được gọi là *tổ hợp lồi* của $x^1, x^2, \ldots, x^k \in \mathbb{R}^n$ nếu

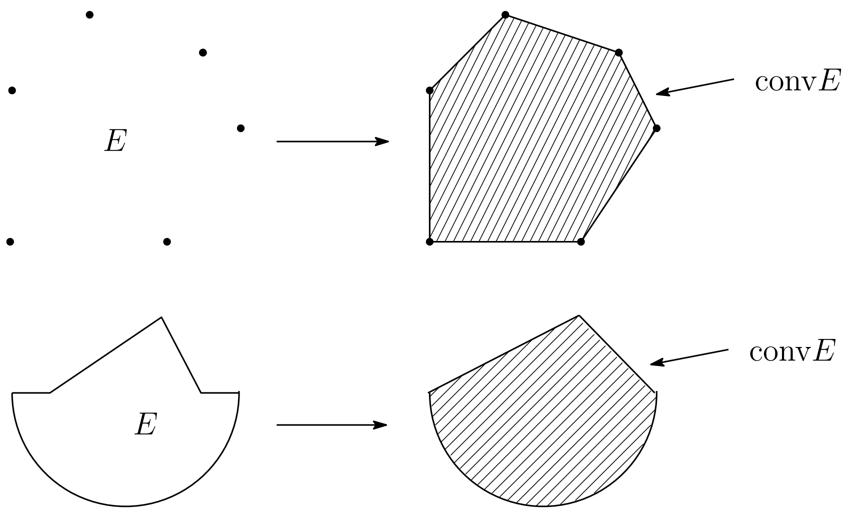
$$x = \sum_{i=1}^{k} \lambda_i x^i \quad \text{với } \lambda_1, \ldots, \lambda_k \geq 0 \text{ và } \sum_{i=1}^{k} \lambda_i = 1.$$

◇ Điểm x được gọi là *tổ hợp lồi chặt* của $x^1, x^2, \ldots, x^k \in \mathbb{R}^n$ nếu

$$x = \sum_{i=1}^{k} \lambda_i x^i \quad \text{với } \lambda_1, \ldots, \lambda_k > 0 \text{ và } \sum_{i=1}^{k} \lambda_i = 1.$$

Mệnh đề 2.3. *Một tập $M \subseteq \mathbb{R}^n$ là lồi khi và chỉ khi nó chứa tất cả các tổ hợp lồi của những phần tử thuộc nó.*

Bao lồi *Bao lồi* (convex hull) của tập $E \subset \mathbb{R}^n$ là giao của tất cả các tập lồi chứa E và được ký hiệu là conv$E$. Đó là tập lồi nhỏ nhất chứa E.



Mệnh đề 2.4. *Bao lồi của tập $E \subset \mathbb{R}^n$ chứa tất cả các tổ hợp lồi của các phần tử thuộc nó.*

Định lý 2.1. (Carathéodory) *Giả sử $E \subset \mathbb{R}^n$ là tập con của một tập afin k chiều với $k \leq n$. Khi đó, mỗi điểm $x \in \text{conv}E$ đều có thể biểu diễn bởi một tổ hợp lồi của không quá $k+1$ phần tử thuộc E.*

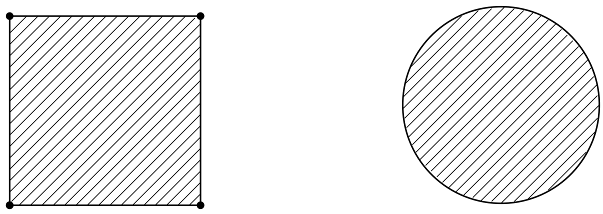
2.4 Điểm cực biên

Cho tập lồi $M \subset \mathbb{R}^n$. Một điểm $x \in M$ được gọi là *điểm cực biên* (extreme point) của M nếu

$$\nexists y, z \in M, y \neq z \text{ sao cho } x = \lambda y + (1-\lambda) z \text{ với } \quad 0 < \lambda < 1,$$

tức

$$x = \lambda y + (1-\lambda) z, \quad y, z \in M \text{ và } 0 < \lambda < 1 \;\Rightarrow\; x = y = z.$$



(a) (b)

(a) – Hình vuông có 4 điểm cực biên; (b) – Hình tròn có vô số điểm cực biên

Chú ý. Khi tập lồi có hữu hạn điểm cực biên thì chúng thường được gọi là *các đỉnh.*

$$x \text{ là điểm cực biên của } M \quad \Rightarrow \quad x \text{ là điểm biên của } M$$

$$x \text{ là điểm cực biên của } M \quad \not\Leftarrow \quad x \text{ là điểm biên của } M$$

Mệnh đề 2.5. *Một tập lồi đóng khác rỗng $M \subset \mathbb{R}^n$ có điểm cực biên khi và chỉ khi nó không chứa trọn một đường thẳng nào.*

Định lý 2.2. (Krein-Milman) *Một tập lồi compac $M \subset \mathbb{R}^n$ là bao lồi của các điểm cực biên của nó.*

2.5 Diện

Cho tập lồi khác rỗng $M \subset \mathbb{R}^n$. Một tập con lồi khác rỗng F của M được gọi là một *diện* của M nếu hễ F chứa một điểm trong tương đối của một đoạn thẳng nào đó thuộc M thì F chứa trọn cả đoạn thẳng đó, nghĩa là

$$y \in P, z \in P,\ x = \lambda y + (1-\lambda)z \in F \text{ với } 0 < \lambda < 1 \ \Rightarrow\ y \in F, z \in F.$$

Mệnh đề 2.6. *Cho tập lồi khác rỗng $D \subset \mathbb{R}^n$ và véc tơ $v \in \mathbb{R}^n \setminus \{0\}$. Giả sử*

$$\operatorname{Argmin}\{g(x) = \langle v, x\rangle : x \in D\} \neq \emptyset. \tag{P_0}$$

Khi đó $\operatorname{Argmin}(P_0)$ *là một diện của D.*

Chứng minh. Lấy $y, z \in D$, $x \in \operatorname{Argmin}(P_0)$ với

$$x = \lambda y + (1-\lambda)z, \quad 0 < \lambda < 1.$$

Ta phải chứng minh $y \in \operatorname{Argmin}(P_0)$ và $z \in \operatorname{Argmin}(P_0)$.

Thật vậy, giả sử $\langle c, y\rangle \geq \langle c, z\rangle$. Khi đó

$$\langle c, x\rangle = \lambda\langle c, y\rangle + (1-\lambda)\langle c, z\rangle \geq \lambda\langle c, z\rangle + (1-\lambda)\langle c, z\rangle = \langle c, z\rangle. \tag{2.4}$$

Vì $z \in D$ và $x \in \operatorname{Argmin}(P_0)$, nên

$$\langle c, x\rangle \leq \langle c, z\rangle. \tag{2.5}$$

Từ (2.4) và (2.5) suy ra

$$\langle c, x\rangle = \langle c, z\rangle \quad \Rightarrow \quad z \in \operatorname{Argmin}(P_0).$$

Hơn nữa ta có

$$\langle c, x\rangle = \lambda\langle c, y\rangle + (1-\lambda)\langle c, z\rangle = \lambda\langle c, y\rangle + (1-\lambda)\langle c, x\rangle.$$

Do đó

$$\langle c, y\rangle = \langle c, x\rangle \quad \Rightarrow \quad y \in \operatorname{Argmin}(P_0).$$

Theo định nghĩa, $\operatorname{Argmin}(P_0)$ là một diện của D. $\square$

2.6 Siêu phẳng, nửa không gian

• **Siêu phẳng** Tập $H \subset \mathbb{R}^n$ là một *siêu phẳng* nếu

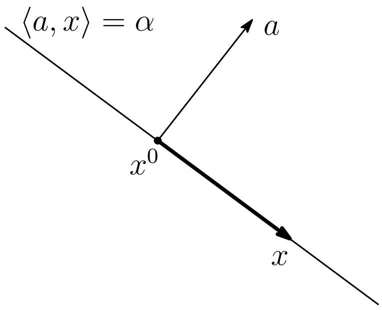
$$H := \{x \in \mathbb{R}^n | \langle a, x\rangle = \alpha\},$$

trong đó $a \in \mathbb{R}^n \setminus \{0\}$ và $\alpha \in \mathbb{R}$. Như vậy

◇ Siêu phẳng H là một tập afin có $\dim H = n - 1$.

◇ Siêu phẳng H là một mặt mức của hàm tuyến tính $\varphi(x) = \langle a, x\rangle$.

Véc tơ a được gọi là *véc tơ pháp tuyến* của siêu phẳng H.



• **Nửa không gian** Cho $a \in \mathbb{R}^n \setminus \{0\}$ và $\alpha \in \mathbb{R}$. Khi đó

◇ Tập

$$\{x \in \mathbb{R}^n \mid \langle a, x\rangle \le \alpha\} \quad (\text{hoặc } \{x \in \mathbb{R}^n \mid \langle a, x\rangle \ge \alpha\})$$

được gọi là *nửa không gian đóng* xác định bởi siêu phẳng

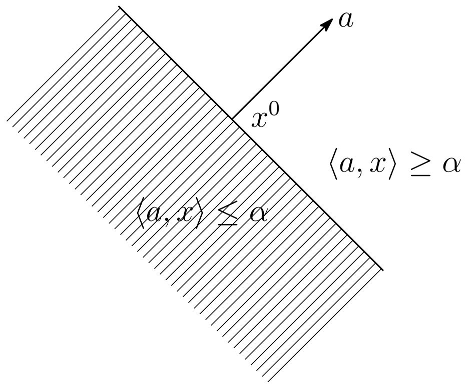
$$\{x \in \mathbb{R}^n \mid \langle a, x\rangle = \alpha\}.$$

◇ Tập

$$\{x \in \mathbb{R}^n \mid \langle a, x\rangle < \alpha\} \quad (\text{hoặc } \{x \in \mathbb{R}^n \mid \langle a, x\rangle > \alpha\})$$

được gọi là *nửa không gian mở* xác định bởi siêu phẳng

$$\{x \in \mathbb{R}^n \mid \langle a, x\rangle = \alpha\}.$$



• **Siêu phẳng tựa**

Cho tập $M \subset \mathbb{R}^n$, véc tơ $a \in \mathbb{R}^n \backslash \{0\}$ và số thực α. Ta gọi siêu phẳng

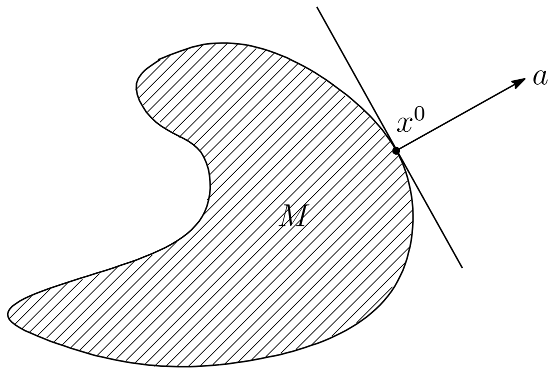
$$H = \{x \in \mathbb{R}^n | \langle a, x\rangle = \alpha\}$$

là *siêu phẳng tựa* (supporting hyperplane) của M tại $x^0 \in M$

$$x^0 \in H$$

$$\langle a, x^0\rangle = \alpha \quad \text{và} \quad \langle a, x\rangle \leq \alpha \;\; \text{với mọi } x \in M.$$

Tập $\{x \in \mathbb{R}^n \mid \langle a, x\rangle \leq \langle a, x^0\rangle\}$ được gọi là *nửa không gian tựa* của M tại x^0.



Định lý 2.3. *Qua mỗi điểm biên x^0 của tập lồi $M \subset \mathbb{R}^n$ tồn tại ít nhất một siêu phẳng tựa của M tại x^0.*

Định lý 2.4. *Một tập lồi đóng khác rỗng $M \subset \mathbb{R}^n$ là giao của họ các nửa không gian tựa của nó.*

2.7 Nón

Tập $M \subset \mathbb{R}^n$ được gọi là *nón* (cone) nếu

$$x \in M \Rightarrow \lambda x \in M \quad \forall \quad \lambda \geq 0.$$

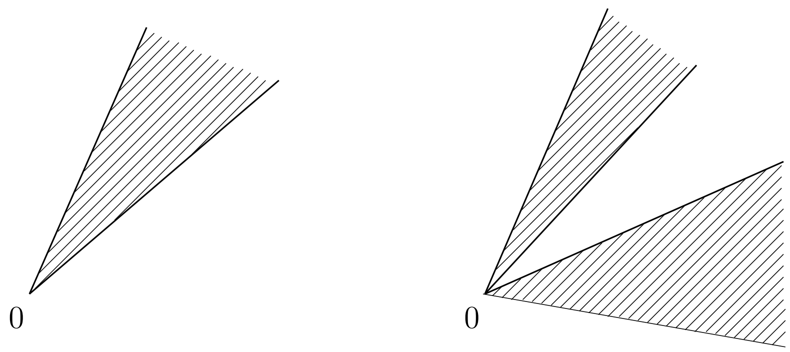
Một nón luôn chứa điểm gốc $0 \in \mathbb{R}^n$.

Tập $M \subset \mathbb{R}^n$ được gọi là *nón lồi* nếu M vừa là nón vừa là tập lồi.

Mệnh đề 2.7. *Tập $M \subseteq \mathbb{R}^n$ là nón lồi khi và chỉ khi*

i) $M + M \subset M$,

ii) $\alpha M \subset M, \quad \forall \alpha \geq 0.$



(a) − Nón lồi; (b) − Nón không lồi

Mệnh đề 2.8. *Tập $M \subset \mathbb{R}^n$ là nón lồi khi và chỉ khi nó chứa tất cả các tổ hợp tuyến tính không âm của các phần tử của nó.*

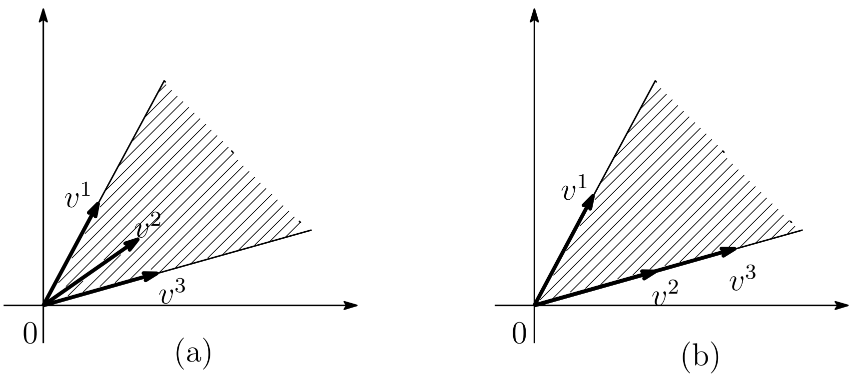
Cho k vec tơ $v^1, \ldots, v^k \in \mathbb{R}^n$. Tập

$$\text{cone}\{v^1, \ldots, v^k\} := \{v \in \mathbb{R}^n \mid v = \sum_{i=1}^{k} \lambda_i v^i, \ \lambda_i \geq 0, \ i = 1, \ldots, k\}$$

được gọi là *nón sinh bởi* các véc tơ $v^1, \ldots, v^k$.

Véc tơ $v^h \in \{v^1, \ldots, v^k\}$ là *không thiết yếu* (non essential) nếu

$$\text{cone}\{v^1, \ldots, v^{h-1}, v^{h+1}, \ldots, v^k\} = \text{cone}\{v^1, \ldots, v^k\}.$$



(a) – Véc tơ v^2 là không thiết yếu vì $\text{cone}\{v^1, v^2, v^3\} = \text{cone}\{v^1, v^3\}$*;*
(b) – Hoặc véc tơ v^2, hoặc véc tơ v^3 là không thiết yếu vì
$\text{cone}\{v^1, v^2, v^3\} = \text{cone}\{v^1, v^3\} = \text{cone}\{v^1, v^2\}$

Cho tập khác rỗng $K \subset \mathbb{R}^n$. Nón sinh bởi K, ký hiệu bởi cone$K$, được định nghĩa bởi

$$\text{cone}K := \{v \in \mathbb{R}^n | v = tx,\ x \in K,\ t \geq 0\}.$$

Dễ thấy rằng, cone$K$ là nón nhỏ nhất chứa K.

2.8 Phương lùi xa, phương cực biên

• Cho tập lồi khác rỗng $D \subseteq \mathbb{R}^n$. Véc tơ $d \neq 0$ được gọi là *phương lùi xa* (recession direction) của D nếu

$$\{x + \lambda d \mid \lambda \geq 0\} \subset D \quad \text{với mỗi } x \in D.$$

$\Rightarrow$ Mọi nửa đường thẳng song song với một phương lùi xa d xuất phát từ một điểm bất kỳ của D đều nằm trọn trong D.

Ký hiệu

$$\text{rec}D = \{\text{Tất cả các phương lùi xa của tập lồi } D\} \cup \{0\}.$$

$\Rightarrow$ rec$D$ là nón lồi và được gọi là *nón lùi xa* của tập D.

Mệnh đề 2.9. *Tập đóng $D \subseteq \mathbb{R}^n$ không bị chặn khi và chỉ khi* rec$D \neq \{0\}$.

• Ta nói hai phương d^1 và d^2 là *khác biệt* (distinct) nếu

$$\nexists \alpha > 0 : \quad d^1 = \alpha d^2.$$

Ví dụ.

◇ $v^1 = (2,5)^T$ và $v^2 = (4,10)^T$ cùng phương;

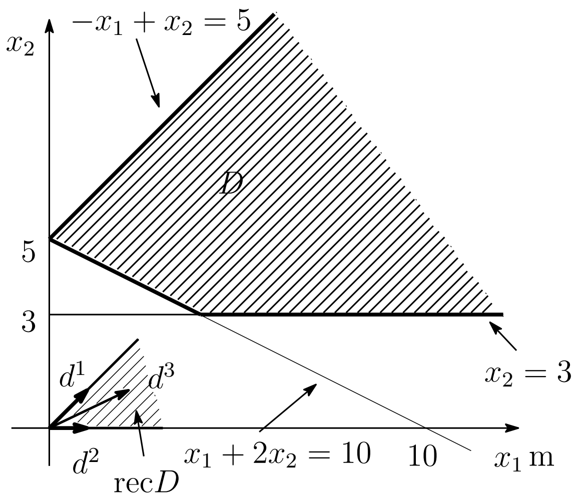
◇ $v^3 = (1,2)^T$ và $v^4 = (2,1)^T$ là khác biệt.

• Phương lùi xa d của tập D được gọi là *phương cực biên* (extreme direction) của D nếu không tồn tại các phương lùi xa khác biệt d^1 và d^2 của D sao cho

$$d = \lambda_1 d^1 + \lambda_2 d^2, \quad \lambda_1,\ \lambda_2 > 0.$$

Ví dụ 2.5. Xét tập

$$D := \{x \in \mathbb{R}^2 \mid -x_1 + x_2 \le 5, x_1 + 2x_2 \ge 10, x_1 \ge 0, x_2 \ge 3\}.$$



Tập D có:

◇ 02 đỉnh là $v^1 = (0,5)^T$ và $v^2 = (4,3)^T$,

◇ 02 phương cực biên là $d^1 = (1,1)^T$ và $d^2 = (1,0)^T$;

◇ $\text{rec}D = \text{cone}\{(1,1)^T, (1,0)^T\}$,

◇ $d^3 = (2,1)^T \in \text{rec}D$ nhưng d^3 không phải là phương cực biên của D.

2.9 Tập lồi đa diện

Tập $P \subset \mathbb{R}^n$ được gọi là *tập lồi đa diện* nếu là giao của một số hữu hạn nửa không gian đóng, tức P là tập nghiệm của hệ

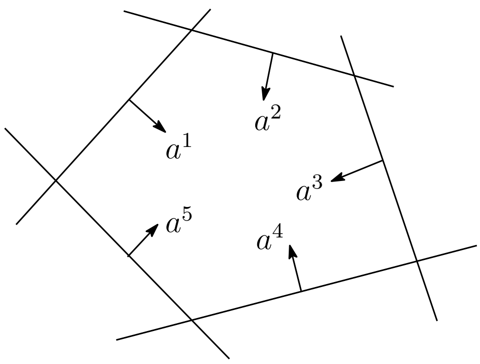
$$\langle a^i, x\rangle \geq b_i, \quad i = 1, \ldots, \ell, \tag{2.6}$$

trong đó $a^i \in \mathbb{R}^n$, $b_i \in \mathbb{R}$ với mọi $i = 1, \ldots, \ell$.

- ◇ Mỗi bất đẳng thức trong hệ (2.6) được gọi là *một ràng buộc* của P.
- ◇ Ràng buộc $k \in \{1, \ldots, \ell\}$ là *thừa* nếu

$$\{x \mid \langle a^i, x\rangle \geq b_i, i = 1, \ldots, \ell\} = \{x \mid \langle a^i, x\rangle \geq b_i, i \in \{1, \ldots, \ell\} \setminus \{k\}\}.$$

- ◇ Tập lồi đa diện là một tập lồi, đóng. Một tập lồi đa diện bị chặn được gọi là *đa diện lồi* hay gọi tắt là *đa diện*.



Ký hiệu

Ma trận $A \in \mathbb{R}^{\{}\ell \times n$ với các hàng là $a^1, \ldots, a^\ell$;

Véc tơ $b = (b_1, \ldots, b_\ell)^T \in \mathbb{R}^\ell$.

Véc tơ $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$.

Khi đó

$$(2.6) \quad \Leftrightarrow \quad Ax \geq b.$$

Mệnh đề 2.10. *Cho tập lồi đa diện P xác định bởi hệ (2.6). Khi đó, tập con lồi khác rỗng $F \subseteq P$ là một diện của P khi và chỉ khi tồn tại một tập chỉ số $I \subseteq \{1, 2, \ldots, \ell\}$ sao cho F là tập nghiệm của hệ*

$$\begin{cases} \langle a^i, x \rangle = b_i, & i \in I \\ \langle a^j, x \rangle \geq b_j, & j \in \{1, 2, \ldots, \ell\} \backslash I. \end{cases}$$

Hơn nữa, ta có $\dim F = n - \operatorname{rank}\{a^i, i \in I\}$.

Chú ý.

- *Đỉnh* của tập lồi đa diện P là một diện có thứ nguyên bằng 0.
- *Cạnh* của P là một diện có thứ nguyên bằng 1.
- Mỗi cạnh vô hạn của tập lồi đa diện P tương ứng với một phương cực biên của nó.

Cho $x^0 \in P$. Tập

$$I(x^0) := \{i \in \{1, 2, \ldots, \ell\} \mid \langle a^i, x^0 \rangle = b_i\}$$

được gọi là *tập hợp các chỉ số của những ràng buộc thỏa mãn chặt* tại $x^0 \in P$.

Hệ quả 2.1. *Cho tập lồi đa diện P xác định bởi hệ (2.6).*

i) *Điểm $x^0 \in P$ là đỉnh của P khi và chỉ x^0 thỏa mãn chặt n ràng buộc độc lập tuyến tính của hệ (2.6).*

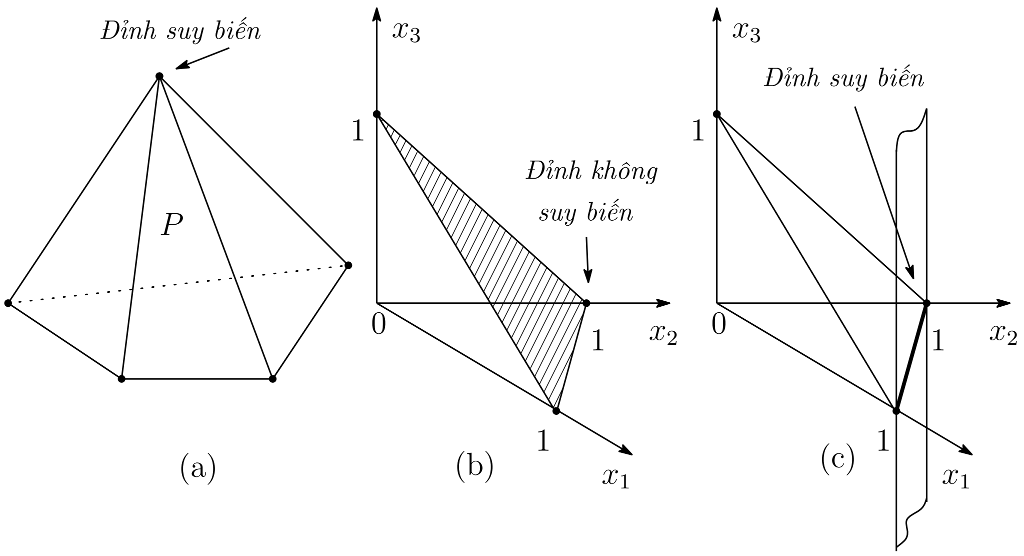
ii) *Một đoạn thẳng (hoặc nửa đường thẳng, hoặc đường thẳng) $\Gamma \subset P$ là một cạnh của P khi và chỉ khi nó là tập các điểm của P thỏa mãn chặt $(n-1)$ ràng buộc độc lập tuyến tính của hệ (2.6).*

Chú ý.

- ◇ Mỗi tập lồi đa diện chỉ có một số hữu hạn đỉnh hoặc cạnh.
- ◇ Một diện của tập lồi đa diện P cũng là một tập lồi đa diện.
- ◇ Mỗi đỉnh của một diện của P cũng là một đỉnh của P.

Định nghĩa. Cho x^0 là một đỉnh của tập lồi đa diện P xác định bởi hệ (2.6).

- ◇ Đỉnh x^0 được gọi là *đỉnh không suy biến* nếu $|I(x^0)| = n$.
- ◇ Đỉnh x^0 là *đỉnh suy biến* nếu $|I(x^0)| > n$.
- ◇ Hai đỉnh x^1 và x^2 được gọi là *kề nhau* nếu đoạn thẳng nối chúng là một cạnh.



Đỉnh suy biến và không suy biến

Định lý 2.5. (Định lý biểu diễn tập lồi đa diện) *Giả sử tập lồi đa diện P có tập đỉnh là*

$$\{v^1, \ldots, v^N\}$$

và tập các phương cực biên là

$$\{d^1, \ldots, d^M\}.$$

Khi đó, mỗi điểm $x \in P$ đều có thể biểu diễn dưới dạng

$$x = \sum_{i=1}^{N} \lambda_i v^i + \sum_{j=1}^{M} \mu_j d^j, \tag{2.7}$$

$$\lambda_i \geq 0,\ i = 1, \ldots, N,\ \mu_j \geq 0,\ j = 1, \ldots, M \ \text{ và } \ \sum_{i=1}^{N} \lambda_i = 1.$$

Nhận xét 2.1. Nếu P là đa diện lồi thì trong (2.7) chỉ còn lại tổng thứ nhất, tức là mỗi điểm của P đều được biểu diễn bằng tổ hợp lồi của các đỉnh của nó.

Ví dụ 2.6. *Xét bài toán*

$$\min f(x) = \langle c, x \rangle \quad v.đ.k. \quad x \in D, \tag{VD_1}$$

trong đó $c \in \mathbb{R}^n \setminus \{0\}$ và $D \subset \mathbb{R}^n$ là đa diện khác rỗng. Khi đó, bài toán (VD_1) đạt nghiệm tối ưu tại ít nhất một đỉnh của X.

Chứng minh. Vì $f(x)$ là hàm liên tục và D là tập compact khác rỗng nên $\text{Argmin}(VD_1) \neq \emptyset$. Giả đa diện có tập đỉnh là $\{v^1, \ldots, v^k\}$.

Lấy tùy ý điểm $x \in D$. Theo Định lý biểu diễn tập lồi đa diện,

$$x = \sum_{i=1}^{k} \lambda_i v^i \quad \lambda_i \geq 0,\ i = 1, \ldots, k, \sum_{i=1}^{k} \lambda_i = 1.$$

Ta có

$$\langle c, x \rangle = \langle c, \sum_{i=1}^{k} \lambda_i v^i \rangle = \sum_{i=1}^{k} \lambda_i \langle c, v^i \rangle \geq \sum_{i=1}^{k} \lambda_i \langle c, v^{i_0} \rangle = \lambda_i \langle c, v^{i_0} \rangle,$$

trong đó

$$v^{i_0} \in \ \{v^1, \ldots, v^k\} \text{ thỏa mãn } \ \langle c, v^{i_0} \rangle = \min\{\langle c, v^i \rangle, i = 1, \ldots, k\}.$$

Vì phương án được chọn tùy ý nên ta có

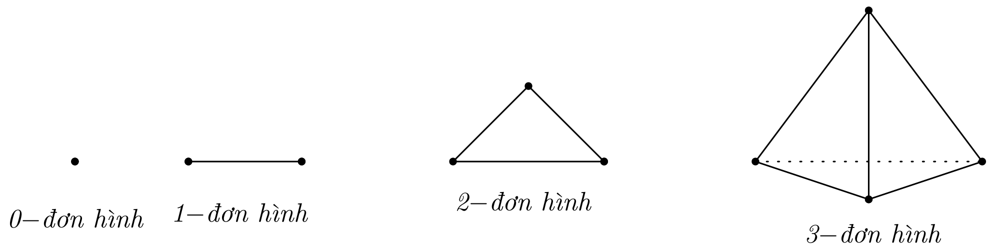
$$\langle c, x \rangle \geq \langle c, v^{i_0} \rangle \ \ \forall\, x \in D.$$

Ta có điều phải chứng minh. $\square$

Ta nói $k+1$ điểm (hay véc tơ) $v^0, v^1, \ldots, v^k \in \mathbb{R}^n$ là *độc lập afin* (affinely independent) nếu k véc tơ $v^1 - v^0, \ldots, v^k - v^0$ là độc lập tuyến tính. Bao lồi của $k+1$ điểm độc lập afin trong $\mathbb{R}^n$ được gọi là đơn hình k chiều hay *$k-$đơn hình*. Tập

$$S = \{x \in \mathbb{R}^n \mid \sum_{j=1}^{n} x_j \leq 1,\ x_j \geq 0,\ j = 1, \ldots, n\}$$

được gọi là *đơn hình chuẩn* trong $\mathbb{R}^n$.



Hình 2.13. *Các đơn hình trong* $\mathbb{R}^3$

Ví dụ 2.7. Một điểm là 0−đơn hình, đoạn thẳng là 1−đơn hình, tam giác là 2−đơn hình và tứ diện là 3−đơn hình (xem Hình 2.13). Đơn hình chuẩn trong $\mathbb{R}^3$ là tứ diện với bốn đỉnh là $(0,0,0)^T$, $(1,0,0)^T$, $(0,1,0)^T$ và $(0,0,1)^T$.

Bài tập Chương 2

1. Cho tập $E \subseteq \mathbb{R}^n$. Chứng minh rằng:

$$x \in \text{aff} E \quad \Leftrightarrow \quad x = \sum_{i=1}^{k} \lambda_i x^i;\ \ x^1, x^2, \ldots, x^k \in E,\ \sum_{i=1}^{k} \lambda_i = 1.$$

2. Chứng minh rằng không gian con song song với tập afin M là xác định duy nhất.

3. Chứng minh rằng:

   i) Giao của một họ hữu hạn các tập lồi là tập lồi.

   ii) Hợp của hai tập lồi chưa chắc đã là tập lồi, cho ví dụ cụ thể.

4. Chứng minh rằng tập $M = \{x \in \mathbb{R}^n | Ax = b,\ x \geq 0\}$ là tập lồi, trong đó A là ma trận cấp $m \times n$ và $b \in R^m$.

5. Cho $\{d^1, \ldots, d^k\}$ là các hướng lùi xa của tập
$$D = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\},$$
với A là ma trận cấp $m \times n$ và $b \in \mathbb{R}^m$. Chứng minh rằng véc tơ
$$0 \neq d = \sum_{i=1}^{k} \alpha_i d^i \text{ với } \alpha_i \geq 0$$
cũng là một hướng lùi xa của D.

6. Cho $P = \{x \in \mathbb{R}^n | Ax = b,\ x \geq 0\}$. Chứng minh rằng nếu $d \neq 0$ thỏa mãn
$$Ad = 0 \quad \text{và} \quad d \geq 0$$
khi và chỉ khi d là một hướng lùi xa của tập P.

7. Cho $p \neq 0$ là một hướng lùi xa của tập
$$D = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\},$$
với A là ma trận cấp $m \times n$ và $b \in \mathbb{R}^m$. Chứng minh rằng $-p$ không thể là một hướng lùi xa của D.

8. Cho tập $M = \{x \in \mathbb{R}^2 \mid x_1 + x_2 \geq 1;\ -x_1 + x_2 \leq 2;\ x_1, x_2 \geq 0\}$. Tìm các phương cực biên của M và xác định nón lùi xa rec$M$.

9. Chứng minh rằng tập $D = \{x \in \mathbb{R}^n | Ax < b\}$ không có một điểm cực biên nào. Cho ví dụ về một tập lồi có hữu hạn điểm cực biên và một tập lồi có vô hạn điểm cực biên.

10. Cho tập lồi đa diện $P = \{x \in \mathbb{R}^6 | Ax = b,\ x \geq 0\}$, trong đó
$$A = \begin{pmatrix} 1 & 4 & 7 & 1 & 0 & 0 \\ 2 & 5 & 8 & 0 & 1 & 0 \\ 3 & 6 & 9 & 0 & 0 & 1 \end{pmatrix} \quad \text{và} \quad b = \begin{pmatrix} 12 \\ 15 \\ 18 \end{pmatrix}.$$
Điểm $x = (1, 1, 1, 0, 0, 0)^T$ có phải là điểm cực biên của P không? Giải thích.

11. Cho tập $P = \{x \in \mathbb{R}^2 | -3x_1 + 2x_2 \leq 30,\ -2x_1 + x_2 \leq 12,\ x_1, x_2 \geq 0\}$.
    i) Vẽ tập P.
    ii) Xác định các điểm cực biên và chỉ ra hai hướng lùi xa độc lập tuyến tính của P.

12. Hãy biểu diễn điểm $(2,2)^T$ như một tổ hợp lồi của ba điểm $(0,0)^T$, $(1,4)^T$ và $(3,1)^T$.

13. Vẽ bao lồi của tập các điểm sau: $(0,0)^T$, $(1,0)^T$, $(-1,2)^T$, $(3,1)^T$, $(2,6)^T$, $(-2,1)^T$, $(-3,-2)^T$, $(3,3)^T$. Chỉ rõ các điểm cực biên và các điểm trong của bao lồi này.

14. Hãy xác định siêu phẳng đi qua các điểm $(1,1,1,1)^T$, $(2,0,1,0)^T$, $(0,2,0,1)^T$ và $(1,1,-1,0)^T$ thuộc $\mathbb{R}^4$.

15. Cho C_1 và C_2 là hai tập lồi trong $\mathbb{R}^n$. Chứng minh rằng tồn tại một siêu phẳng tách chặt C_1 và C_2 khi và chỉ khi

$$\inf\{\|x-y\| \;|x \in C_1,\; y \in C_2\} > 0.$$

16. Cho A là ma trận cấp $m \times n$ và $c \in \mathbb{R}^m \setminus \{0\}$. Chứng minh rằng có một và chỉ một hệ phương trình trong hai hệ phương trình sau có nghiệm:

    i) $Ax = c$.

    ii) $A^T y = 0,\; \langle c, y\rangle = 1$.

17. Cho ma trận A cấp $m \times n$, véc tơ $b \in \mathbb{R}^m$, $b \geq 0$. Chứng minh rằng $x^0 = 0$ là một đỉnh của tập lồi đa diện

$$D = \{x \in \mathbb{R}^n \mid Ax \leq b, x \geq 0\}.$$

Chương 3

Qui hoạch tuyến tính (Linear Programming)

"*Một trong các thành tựu vĩ đại của thế kỷ XX là đã phát minh và phát triển lý thuyết quy hoạch tuyến tính*".

~ Bản tin của Liên đoàn Toán học thế giới 1/2005 ~

- Năm 1939: Leonid Vitaliyevich Kantorovich đã công bố công thức toán học của các vấn đề kinh tế cơ bản, phác thảo về phương pháp giải và thảo luận về ý nghĩa kinh tế của nó. Thực chất, công trình đó chứa đựng những ý tưởng chính về lý thuyết và thuật toán giải quy hoạch tuyến tính, phương Tây đã không biết đến công trình này trong nhiều năm.

- Năm 1947: Geogre Bernard Dantzig cùng các cộng sự phát hiện lại mô hình quy hoạch tuyến tính khi nghiên cứu bài toán lập kế hoạch cho không quân Mỹ và công bố thuật toán đơn hình (simplex algorithm) nổi tiếng để giải bài toán quy hoạch tuyến tính.

- Năm 1975: L. V. Kantorovich và T. C. Koopmans được trao giải thưởng giải thưởng Nobel dành cho khoa học kinh tế về những đóng góp quan trọng trong việc ứng dụng quy hoạch tuyến tính giải quyết các bài toán trong kinh tế .

3.1 Mô hình toán học

Bài toán quy hoạch tuyến tính tổng quát:

$$\min\{f(x) = \langle c, x\rangle \mid x \in D\}, \tag{LP}$$

trong đó $c = (c_1, c_2, \ldots, c_n)^T \in \mathbb{R}^n$ và $D \subset \mathbb{R}^n$ là tập lồi đa diện.

Khi nghiên cứu quy hoạch tuyến tính cũng như khi áp dụng nó, người ta thường dùng hai dạng đặc thù sau:

3.1.1 Dạng chuẩn tắc

$$\begin{aligned} \min\ f(x) = \langle c, x\rangle = c_1x_1 + \cdots + c_nx_n \\ \text{v.đ.k.}\quad \langle a^i, x\rangle \geq b_i,\ i = 1, \ldots, m \\ x_j \geq 0,\ j = 1, \ldots, n, \end{aligned} \tag{LP^*_{ct}}$$

trong đó $c, a^1, \ldots, a^m \in \mathbb{R}^n$, số thực $b_i \in \mathbb{R}$ với mọi $i = 1, \ldots, m$.

- Với mỗi $i \in \{1, \ldots, m\}$, bất đẳng thức

$$\langle a^i, x\rangle \geq b_i$$

  được gọi là một *ràng buộc chính*.

- Với mỗi $j \in \{1, \ldots, n\}$, ràng buộc $x_j \geq 0$ được gọi là *ràng buộc dấu*.

- Kí hiệu $A \in \mathbb{R}^{m\times n}$ với các hàng $a^i i = 1, \ldots, m$ và véc tơ

$$b = (b_1, \ldots, b_m)^T \in \mathbb{R}^m.$$

  Bài toán LP^*_{ct}) được viết lại dưới dạng ma trận như sau:

$$\begin{aligned} \min\ f(x) = &\langle c, x\rangle \\ \text{v.đ.k.}\quad & Ax \geq b \\ & x \geq 0. \end{aligned}$$

3.1.2 Dạng chính tắc

$$\begin{aligned} \min\ f(x) = \langle c, x\rangle = c_1x_1 + \cdots + c_nx_n \qquad\qquad & (LP_{ct})\\ \text{v.đ.k.}\quad \langle a^i, x\rangle = b_i,\ i = 1,\dots,m &\\ x_j \geq 0,\ j = 1,\dots,n, &\end{aligned}$$

trong đó $c, a^1, \dots, a^m \in \mathbb{R}^n$, số thực $b_i \geq 0$ với mọi $i = 1, \dots, m$.

- Với mỗi $i \in \{1, \dots, m\}$, bất đẳng thức
$$\langle a^i, x\rangle = b_i$$
được gọi là một *ràng buộc chính.*
- Với mỗi $j \in \{1, \dots, n\}$, ràng buộc $x_j \geq 0$ được gọi là *ràng buộc dấu.*
- Kí hiệu $A \in \mathbb{R}^{m\times n}$ với các hàng $a^i i = 1, \dots, m$ và véc tơ
$$b = (b_1, \dots, b_m)^T \in \mathbb{R}^m_+.$$
Bài toán LP_{ct}) được viết lại dưới dạng ma trận như sau:
$$\begin{aligned} \min\ f(x) =& \langle c, x\rangle\\ \text{v.đ.k.}\quad Ax =& b\\ x \geq& 0.\end{aligned}$$

Chú ý. Trong bài toán quy hoạch tuyến tính dạng chính tắc,

$$b_i \geq 0, \quad \forall\ \ i = 1, \dots, m.$$

Ví dụ 3.1. Bài toán

$$\begin{aligned} \min\ f(x) =&\ 7x_1 + 6x_2 - 8x_3\\ \text{v.đ.k}\quad &\ 2x_1 - 5x_2 + 4x_3 = 8\\ &\ 2x_1 + 9x_2 - 5x_3 = 15\\ &\ x_1,\ x_2,\ x_3 \geq 0\end{aligned}$$

là một bài toán quy hoạch tuyến tính chính tắc, trong đó

$$x = \begin{pmatrix} x_1\\ x_2\\ x_3\end{pmatrix},\ \ c = \begin{pmatrix} 7\\ 6\\ -8\end{pmatrix},\ \ A = \begin{pmatrix} 2 & -5 & 4\\ 2 & 9 & -5\end{pmatrix},\ \ b = \begin{pmatrix} 8\\ 15\end{pmatrix}.$$

Bài toán này có $n = 3$ biến và $m = 2$ ràng buộc chính.

Chú ý. Kí hiệu tập nghiệm của bài toán (LP_{ct}^*) là

$$D^* = \{x : Ax \geq b, \quad x \geq 0\}$$

và tập nghiệm của bài toán (LP_{ct}) là

$$D = \{x : Ax = b, \quad x \geq 0\}.$$

Dễ thấy

$$D^* \subset \mathbb{R}^n_+ \quad \text{và} \quad D \subset \mathbb{R}^n_+.$$

Do đó

$$D^* \neq \emptyset \quad \Rightarrow \quad D^* \text{ có đỉnh,}$$
$$D \neq \emptyset \quad \Rightarrow \quad D \text{ có đỉnh.}$$

3.1.3 Chuyển bài toán quy hoạch tuyến tính bất kỳ về dạng chuẩn tắc hay chính tắc

Mọi bài toán quy hoạch tuyến tính đều có thể đưa về bài toán quy hoạch tuyến tính dạng chính tắc hoặc chuẩn tắc bằng các phép biến đổi sơ cấp sau:

- ◇ Nếu x_j không bị ràng buộc dấu thì đặt

$$x_j = \bar{x}_j - \bar{\bar{x}}_j \quad \text{với } \bar{x}_j \geq 0, \quad \bar{\bar{x}}_j \geq 0.$$

- ◇ Nếu $x_j \leq 0$ đổi biến

$$\bar{x}_j = -x_j \geq 0.$$

- ◇ Mỗi ràng buộc bất đẳng thức

$$\sum a_{ij}x_j \leq b_i$$

  có thể chuyển về ràng buộc đẳng thức nhờ đưa thêm vào một *biến phụ* $x_{n+i} \geq 0$,

$$\sum a_{ij}x_j + x_{n+i} = b_i.$$

- ◇ Mỗi ràng buộc bất đẳng thức

$$\langle a^i, x\rangle = \sum a_{ij}x_j \geq b_i$$

  có thể chuyển về ràng buộc đẳng thức nhờ đưa thêm vào một *biến phụ* $x_{n+i} \geq 0$,

$$\sum a_{ij}x_j - x_{n+i} = b_i.$$

⋄ Mỗi ràng buộc

$$\langle a^i, x\rangle = \sum a_{ij}x_j \le b_i$$

có thể viết lại thành

$$\langle -a^i, x\rangle = -\sum a_{ij}x_j \ge -b_i.$$

⋄ Mỗi ràng buộc đẳng thức

$$\langle a^i, x\rangle = \sum a_{ij}x_j = b_i$$

có thể thay bằng hai ràng buộc bất đẳng thức

$$\langle a^i, x\rangle = \sum a_{ij}x_j \ge b_i \text{ và } \langle -a^i, x\rangle = -\sum a_{ij}x_j \ge -b_i.$$

Ví dụ 3.2. Xét bài toán quy hoạch tuyến tính

$$\begin{aligned} \min\ f(x) &= 3x_1 + 5x_2 - 4x_3 \\ \text{v.đ.k.}\quad 3x_1 - 5x_2 + 3x_3 &\le 5 \\ 2x_1 + 4x_2 + 6x_3 &= 8 \\ -4x_1 - 9x_2 + 4x_3 &\le -4 \\ x_1 \ge -2,\ 0 \le x_2 \le 4,\ x_3 &\in \mathbb{R}. \end{aligned}$$

Để chuyển bài toán quy hoạch tuyến tính này về dạng chính tắc, trước hết thực hiện:

- Nhân hai vế của ràng buộc chính thứ ba với -1 ta được

$$4x_1 + 9x_2 - 4x_3 \ge 4;$$

- Đổi biến x_1 thành $\bar{x}_1$ với

$$\bar{x}_1 := x_1 + 2 \ge 0 \quad \Rightarrow \quad x_1 = \bar{x}_1 - 2;$$

- Ràng buộc cận trên của biến thứ hai $x_2 \le 4$ được xem như ràng buộc chính thứ tư;
- Biến thứ ba được đổi thành

$$x_3 = \bar{x}_3 - \bar{\bar{x}}_3 \text{ với } \bar{x}_3 \ge 0,\ \bar{\bar{x}}_3 \ge 0.$$

Sau khi thay thế các biến đổi trên vào bài toán ban đầu ta nhận được bài toán:

$$\begin{array}{lll} \min\ z = & 3\bar{x}_1 + 5x_2 - 4\bar{x}_3 + 4\bar{\bar{x}}_3 - 6 & \\ \text{v.đ.k.} & 3\bar{x}_1 - 5x_2 + 3\bar{x}_3 - 3\bar{\bar{x}}_3 & \le 11 \\ & 2\bar{x}_1 + 4x_2 + 6\bar{x}_3 - 6\bar{\bar{x}}_3 & = 12 \\ & 4\bar{x}_1 + 9x_2 - 4\bar{x}_3 + 4\bar{\bar{x}}_3 & \ge 12 \\ & \qquad x_2 & \le 4 \\ & \qquad \bar{x}_1,\ x_2,\ \bar{x}_3,\ \bar{\bar{x}}_3 & \ge 0. \end{array}$$

Cuối cùng, sau khi bỏ hằng số -6 ở hàm mục tiêu và thêm các biến phụ $x_4,\ x_5,\ x_6 \ge 0$ lần lượt vào các ràng buộc chính thứ nhất, thứ ba, thứ tư của bài toán này, ta nhận được bài toán quy hoạch tuyến tính dạng chính tắc sau:

$$\begin{array}{lll} \min\ f'(x) = & 3\bar{x}_1 + 5x_2 - 4\bar{x}_3 + 4\bar{\bar{x}}_3 & \\ \text{v.đ.k.} & 3\bar{x}_1 - 5x_2 + 3\bar{x}_3 - 3\bar{\bar{x}}_3 + x_4 & = 11 \\ & 2\bar{x}_1 + 4x_2 + 6\bar{x}_3 - 6\bar{\bar{x}}_3 & = 12 \\ & 4\bar{x}_1 + 9x_2 - 4\bar{x}_3 + 4\bar{\bar{x}}_3 \quad - x_5 & = 12 \\ & \qquad x_2 \qquad\qquad\qquad + x_6 & = 4 \\ & \qquad \bar{x}_1,\ x_2,\ \bar{x}_3,\ \bar{\bar{x}}_3,\ x_4,\ x_5,\ x_6 & \ge 0. \end{array}$$

Giả sử $(\bar{x}_1^*,\ x_2^*,\ \bar{x}_3^*,\ \bar{\bar{x}}_3^*,\ x_4^*,\ x_5^*,\ x_6^*)^T$ là nghiệm tối ưu của bài toán quy hoạch tuyến tính chính tắc này. Khi đó nghiệm tối ưu của bài toán ban đầu là $x^{opt} = (x_1^{opt},\ x_2^{opt},\ x_3^{opt})^T$ và giá trị tối ưu là $f(x^{opt}) = 3x_1^{opt} + 5x_2^{opt} - 4x_3^{opt}$, trong đó $x_1^{opt} = \bar{x}_1^* - 2$, $x^{opt} = x_2^*$ và $x_3^{opt} = \bar{x}_3^* - \bar{\bar{x}}_3^*$.

3.2 Điều kiện tồn tại nghiệm và tính chất tập nghiệm của bài toán quy hoạch tuyến tính

Xét bài toán quy hoạch tuyến tính tổng quát

$$\min f(x) = \langle c, x \rangle \quad \text{v.đ.k.} \quad x \in D, \tag{LP}$$

trong đó $c \in \mathbb{R}^n$ và $D \subset \mathbb{R}^n$ là tập lồi đa diện.

3.2.1 Điều kiện tồn tại nghiệm

Trường hợp 1. $D = \emptyset \quad \Rightarrow \quad \operatorname{Arg\,min}\{f(x) : x \in D\} = \emptyset.$

Trường hợp 2.

Định lý 3.1. *Nếu* $D\emptyset$ *bị chặn thì* $\operatorname{Arg\,min}\{f(x) : x \in D\} \neq \emptyset$

Chứng minh. Nhắc lại, tập lồi đa diện là tập đóng. Theo giả thiết, D khác rỗng và bị chặn. Suy ra

$$D \text{ là tập compac, khác rỗng.} \tag{1}$$

Vì hàm mục tiêu f là hàm tuyến tính, nên

$$f \text{ là hàm liên tục trên } \mathbb{R}^n. \tag{2}$$

Do (1) và (2), theo Định lý Weierstrass, ta có

$$\operatorname{Arg\,min}\{f(x) : x \in D\} \neq \emptyset$$

□

Trường hợp 3.

Định lý 3.2. *Nếu D khác rỗng và $f(x) = \langle c, x\rangle$ bị chặn dưới trên D, tức*

$$\exists \alpha \in \mathbb{R} \quad : \ f(x) \geq \alpha \quad \forall\ x \in D$$

thì $\operatorname{Arg\,min}\{f(x) : x \in D\} \neq \emptyset$.

Chứng minh. Vì mọi quy hoạch tuyến tính đều có thể chuyển về dạng chuẩn tắc hoặc chính tắc nên không giảm tổng quát ta giả thiết tập D có đỉnh. Giả sử D có tập đỉnh $\{v^1, \ldots, v^N\}$ và tập các phương cực biên là $\{d^1, \ldots, d^M\}$.

Lấy tùy ý $x \in D$. Theo Định lý biểu diễn tập lồi đa diện,

$$x = \sum_{i=1}^{N} \lambda_i v^i + \sum_{j=1}^{M} \mu_j d^j, \tag{3}$$

$$\lambda_i \geq 0,\ i = 1, \ldots, N,\ \mu_j \geq 0,\ j = 1, \ldots, M, \quad \sum_{i=1}^{N} \lambda_i = 1.$$

Do hàm mục tiêu $f(x) = \langle c, x\rangle$ bị chặn dưới trên D nên

$$\langle c, d^j\rangle \geq 0 \quad \forall d^j, \ \ j = 1, \ldots, M. \tag{4}$$

Thật vậy, giả sử tồn tại phương cực biên $d^{j_0} \in \{d^1, \ldots, d^M\}$ sao cho

$$\langle c, d^{j_0}\rangle < 0.$$

Vì d^{j_0} là một phương cực biên nên

$$x + td^{j_0} \in D \ \ \forall x \ \in D, \quad \forall t \geq 0$$

và

$$\langle c, x + td^{j_0}\rangle = \langle c, x\rangle + t\langle c, d^{j_0}\rangle \longrightarrow -\infty \text{ khi } \ t \to +\infty.$$

Điều này mâu thuẫn với tính bị chặn dưới của hàm $f(x) = \langle c, x\rangle$ và chứng tỏ khẳng định (4) là đúng.

Chọn một đỉnh v^{i_0} của D sao cho

$$\langle c, v^{i_0}\rangle = \min\{\langle c, v^i\rangle \mid i = 1, \ldots, N\}.$$

Theo (3) và (4), với bất kỳ $x \in D$, ta có

$$\langle c, x\rangle \overset{(3)}{=} \sum_{i=1}^{N} \lambda_i \langle c, v^i\rangle + \sum_{j=1}^{M} \mu_j \langle c, d^j\rangle \overset{(4)}{\geq} \sum_{i=1}^{N} \lambda_i \langle c, v^i\rangle \geq \sum_{i=1}^{N} \lambda_i \langle c, v^{i_0}\rangle = \langle c, v^{i_0}\rangle.$$

Suy ra v^{i_0} là nghiệm tối ưu của bài toán (LP). $\square$

Chú ý 3.1. Kết luận của Định lý 3.2 nói chung không còn đúng đối với bài toán quy hoạch phi tuyến. Ví dụ:

i) Bài toán

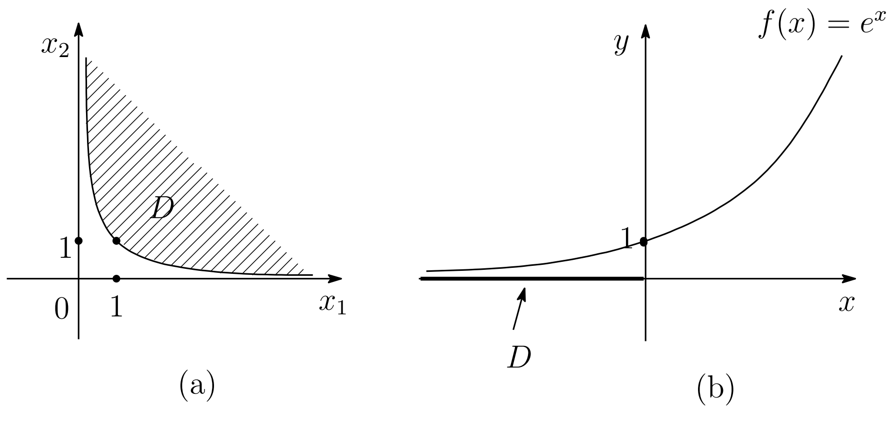
$$\inf\{f(x) = x_2 \mid x \in D\},$$

trong đó $D = \{x \in \mathbb{R}^2 \mid x_1x_2 \geq 1,\ x_1,\ x_2 \geq 0\}$, có hàm mục tiêu là tuyến tính và bị chặn dưới bởi 0. Tập nghiệm chấp nhận được D là tập lồi khác rỗng nhưng không phải tập lồi đa diện. Đây không phải là bài toán quy hoạch tuyến tính và dễ thấy, $x = (x_1, 0)^T \notin D$ với mọi $x_1 \geq 0$. Vì thế bài toán này không có nghiệm tối ưu (xem Hình 3.1(a)) và $\inf f(D) = 0$.

ii) Bài toán

$$\min\{f(x) = e^x \mid x \in D\},$$

trong đó $D = \{x \in \mathbb{R} \mid x \leq 0\}$, có tập chấp nhận được là tập lồi đa diện nhưng hàm mục tiêu là phi tuyến và cũng bị chặn dưới bởi 0. Rõ ràng cũng không tồn tại một điểm $x \in D$ để $e^x = 0$ và bài toán này không có nghiệm tối ưu (xem Hình 3.1(b)), giá trị tối ưu $\inf f(D) = 0$.



Hình 3.1

3.2.2 Tính chất tập nghiệm

Định lý 3.3. *Nếu bài toán quy hoạch tuyến tính* (LP) *có nghiệm tối ưu thì tập nghiệm tối ưu của nó là một diện của tập lồi đa diện chấp nhận được.*

Chứng minh. Đây là trường hợp riêng của Định lý:

Cho tập lồi khác rỗng $D \subset \mathbb{R}^n$ *và véc tơ* $v \in \mathbb{R}^n \setminus \{0\}$. *Giả sử*

$$\mathrm{Argmin}\{g(x) = \langle v, x \rangle : x \in D\} \neq \emptyset. \tag{P_0}$$

Khi đó $\mathrm{Argmin}(P_0)$ *là một diện của* D.

Hệ quả 3.1. *i) Tập nghiệm của bài toán qui hoạch tuyến tính là một tập lồi đa diện.*

ii) Nếu một quy hoạch tuyến tính có nghiệm tối ưu và tập lồi đa diện ràng buộc có đỉnh thì nghiệm tối ưu phải đạt tại ít nhất một đỉnh, tức đạt tại ít nhất một phương án cực biên.

Chứng minh. Theo định nghĩa, phương án cực biên chính là một đỉnh của tập lồi đa diện chấp nhận được của bài toán quy hoạch tuyến tính. Hệ quả được suy trực tiếp từ Định lý 3.3 và sự kiện là đỉnh của một diện của một tập lồi đa diện cũng chính là đỉnh của tập lồi đa diện đó (Hệ quả 2.3). □

Định lý 3.4. *Nếu x^* là nghiệm tối ưu địa phương của bài toán quy hoạch tuyến tính (LP) thì x^* cũng là nghiệm tối ưu toàn cục.*

Chứng minh. Giả sử $x^* \in D$ là nghiệm tối ưu địa phương của bài toán (LP). Theo định nghĩa, tồn tại một hình cầu mở $B(x^*, \varepsilon)$ sao cho

$$\langle c, x^* \rangle \leq \langle c, x \rangle \quad \forall x \in B(x^*, \varepsilon) \cap D.$$

Lấy tùy ý phương án chấp nhận được $x \in D$. Do D là tập lồi đa diện

$$[x^*, x] \subset D.$$

Lấy điểm

$$x^0 \in ([x^*, x] \cap B(x^*, \varepsilon)\Big) \subset \Big(D \cap B(x^*, \varepsilon)\Big)$$

$$\Rightarrow\ x^0 = \lambda x^* + (1-\lambda)x \quad \text{với } 0 < \lambda < 1.$$

Do x^* là nghiệm tối ưu địa phương và $x^0 \in B(x^*, \varepsilon) \cap D$ nên

$$\langle c, x^0 \rangle = \lambda \langle c, x^* \rangle + (1-\lambda)\langle c, x \rangle \geq \langle c, x^* \rangle.$$

$$\Rightarrow (1-\lambda)\langle c, x \rangle \geq (1-\lambda)\langle c, x^* \rangle$$

$$\Rightarrow \langle c, x \rangle \geq \langle c, x^* \rangle.$$

Vì $x \in D$ được lấy tùy ý nên

$$\langle c, x \rangle \geq \langle c, x^* \rangle \ \ \forall x \in D \Rightarrow\ x^* \in \text{Argmin}\{g(x) = \langle v, x \rangle : x \in D\}.$$

□

3.3 Giải bài toán quy hoạch tuyến tính hai biến bằng phương pháp hình học

Xét bài toán quy hoạch tuyến tính hai biến

$$\min f(x) = \langle c, x \rangle \ \text{ v.đ.k. } \ x \in D, \qquad (LP^{2b})$$

hoặc bài toán

$$\max f(x) = \langle c, x \rangle \ \text{ v.đ.k. } \ x \in D, \qquad (LP^{2b}_*)$$

trong đó $c = (c_1, c_2)^T \in \mathbb{R}^2 \setminus \{0\}$ và $D \subset \mathbb{R}^2$ là tập lồi đa diện.

Có sở lý thuyết

i) Qua mỗi điểm $\bar{x} = (\bar{x}_1, \bar{x}_2)^T \in \mathbb{R}^2$ chỉ có duy nhất một đường mức

$$L(\alpha, f) = \{x \in \mathbb{R}^2 \mid \langle c, x \rangle = \alpha\}$$

của hàm $f(x) = \langle c, x \rangle$ với mức $\alpha = \langle c, \bar{x} \rangle$.

ii) $\nabla f(x) = c$ tại mọi điểm $x \in \mathbb{R}^n$. Do đó véc tơ hàm mục tiêu c là véc tơ pháp tuyến của mọi đường mức.

iii) Các đường mức của hàm $f(x) = \langle c, x \rangle$ song song với nhau.

iv) Giá trị hàm $f(x) = \langle c, x \rangle$ tăng theo hướng véc tơ gradient $\nabla f(x) = c$ và giảm theo hướng ngược véc tơ c.

Ví dụ Xét hàm $f(x) = 2x_1 + 3x_2$. Ta có

$$\nabla f(x) = (f'_{x_1}, f'_{x_2})^T = (2, 3)^T$$

là véc tơ pháp tuyến của mọi đường mức

$$L(\alpha, f) = \{x \in \mathbb{R}^2 : 2x_1 + 3x_2 = \alpha\}, \quad \text{với} \quad \alpha \in \mathbb{R}$$

.

Lấy tùy ý $\alpha_1 \neq \alpha_2$, ta có $L(\alpha_1, f)//L(\alpha_2, f)$. Chẳng hạn

$$L(6, f) = \{x \in \mathbb{R}^2 : 2x_1 + 3x_2 = 6\}//L(12, f) = \{x \in \mathbb{R}^2 : 2x_1 + 3x_2 = 12\}.$$

• Phát biểu theo ngôn ngữ hình học:

◇ Bài toán $\min\{f(x) : x \in D\}$: *Trong số các đường mức cắt tập D, hãy tìm đường mức có giá trị mức nhỏ nhất.*

◇ Bài toán $\max\{f(x) : x \in D\}$: *Trong số các đường mức cắt tập D, hãy tìm đường mức có giá trị mức lớn nhất.*

Thuật toán.

Bước 1. Vẽ:

- Tập chấp nhận được D;

- Véc tơ hàm mục tiêu c;

- Đường mức $L(0, f) = \{x \in \mathbb{R}^2 \mid \langle c, x\rangle = 0\}$ đi qua điểm gốc 0 và vuông góc với c.

Bước 2.

- Lấy một điểm bất kỳ $\bar{x} \in D$.

- Vẽ đường thẳng L đi qua $\bar{x}$ và song song với đường mức $L(0, f)$. *Đường thẳng L là đường mức*

$$\{x \in \mathbb{R}^2 \mid \langle c, x\rangle = \langle c, \bar{x}\rangle\}$$

.

Bước 3.

◇ Bài toán $\min\{f(x) : x \in D\}$:

Dịch chuyển song song đường mức L **theo hướng ngược với hướng véc tơ** c đến khi việc dịch chuyển tiếp theo làm cho đường mức không còn cắt D nữa thì dừng. Các điểm của D nằm trên đường mức cuối cùng này là các nghiệm tối ưu của bài toán (LP^{2b}), còn giá trị mức này chính là giá trị tối ưu của bài toán.

◇ Bài toán $\max\{f(x) : x \in D\}$:

Dịch chuyển song song đường mức L **theo hướng hướng véc tơ** c đến khi việc dịch chuyển tiếp theo làm cho đường mức không còn cắt D nữa thì dừng. Các điểm của D nằm trên đường mức cuối cùng này là các nghiệm tối ưu của bài toán (LP^{2b}), còn giá trị mức này chính là giá trị tối ưu của bài toán.

Cách trình bày bài làm giải bài toán qui hoạch tuyến tính bằng thuật toán hình học

Ví dụ Xét bài toán quy hoạch tuyến tính:

$$\begin{aligned} \min\ f(x) = -20x_1 + 10x_2 & \\ \text{v.đ.k.} \quad -x_1 + 2x_2 &\le 4 \\ x_1 + \ \ x_2 &\le 5 \\ x_1 \qquad &\le 4 \\ x_1,\ x_2 &\ge 0. \end{aligned}$$

Lời giải.

- ⋄ Tập chấp nhận được D, hướng của véc tơ hàm mục tiêu $c = (-20, 10)^T$, đường mức
$$L_0 = \{x \in \mathbb{R}^2 \ : \ \langle c, x\rangle = 0\}$$
và đường mức đi qua điểm $\bar{x} \in D$
$$L = \{x \in \mathbb{R}^2 \ : \ \langle v, x\rangle = \langle c, \bar{x}\rangle\}$$
được minh họa như hình vẽ.

- ⋄ $*$ Bài toán $\min\{f(x) : x \in D\}$:
Vì **bài toán tìm cực tiểu hàm** $f(x) = \langle v, x\rangle$ trên D nên tịnh tiến đường mức L **theo hướng ngược hướng véc tơ** c.

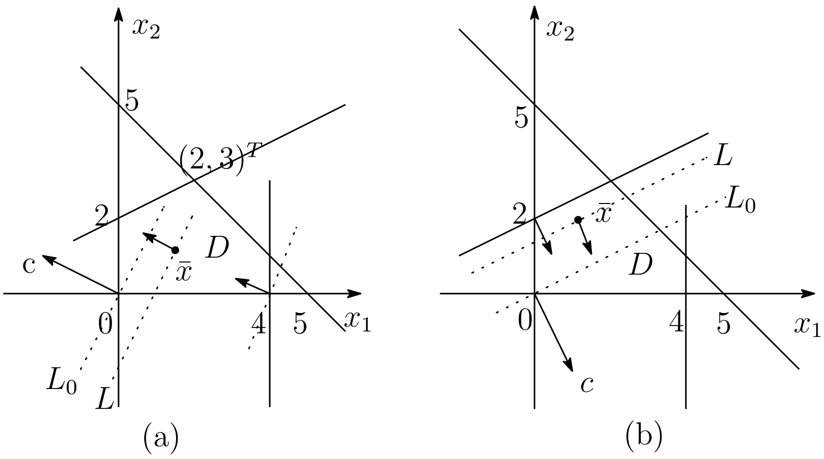
  $*$ Bài toán $\max\{f(x) : x \in D\}$:
Vì **bài toán tìm cực đại hàm** $f(x) = \langle v, x\rangle$ trên D nên tịnh tiến đường mức L **theo hướng véc tơ** c.

- ⋄ Theo thuật toán hình học, bài toán có nghiệm tối ưu là .... và giá trị tối ưu là .... (hoặc bài toán không có nghiêm tối ưu)

Chú ý. Trong trường hợp $0 \in D$ thì có thể bỏ qua câu:
"*đường mức đi qua điểm* $\bar{x} \in D$

$$L = \{x \in \mathbb{R}^2 \ : \ \langle v, x\rangle = \langle c, \bar{x}\rangle\}\text{”}.$$

Nếu thay hàm mục tiêu của bài toán này bởi $f'(x) = 10x_1 - 20x_2$ thì ta nhận được tập nghiệm tối ưu là cả cạnh $[(0,2)^T, (2,3)^T]$. Trường hợp này, bài toán có vô số nghiệm. Ta có thể lấy một nghiệm tối ưu đại diện là đỉnh $(0,2)^T$.



(a) – Bài toán có duy nhất nghiệm; (b) – Bài toán có vô số nghiệm

Ví dụ 3.3. Xét bài toán quy hoạch tuyến tính:

$$\begin{aligned} \min\ f(x) = \ & x_1 + \ \ x_2 \\ \text{v.đ.k.}\quad & x_1 + 2x_2 \ \geq 2 \\ & -x_1 + 2x_2 \ \leq 4 \\ & x_1,\ x_2 \ \geq 0. \end{aligned}$$

Các thông tin về bài toán này như tập chấp nhận được, hướng của véc tơ mục tiêu $c = (1,1)^T$, đường mức L_0 và đường mức L đi qua điểm $\bar{x} \in D$ được minh họa ở Hình 3.3(a). Tập chấp nhận được của bài toán là không bị chặn. Theo Thuật toán 3.1, ta thấy bài toán có nghiệm tối ưu duy nhất là $x^* = (0,1)^T$.

Nếu thay hàm mục tiêu của bài toán này bởi $f'(x) = -x_1 - x_2$ thì bài toán vô nghiệm, hàm mục tiêu giảm vô hạn trên tập phương án (xem Hình 3.3(b)).

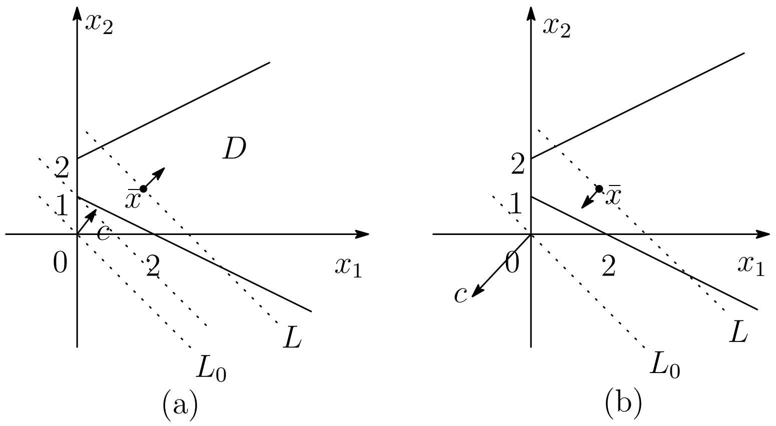
Ví dụ 3.4. Xét bài toán quy hoạch tuyến tính:

$$\begin{aligned} \min\ f(x) = \ & 2x_1 - 3x_2 \\ \text{v.đ.k.}\quad & -2x_1 + 3x_2 \leq 6 \\ & 2x_1 - 3x_2 \leq 6. \end{aligned}$$

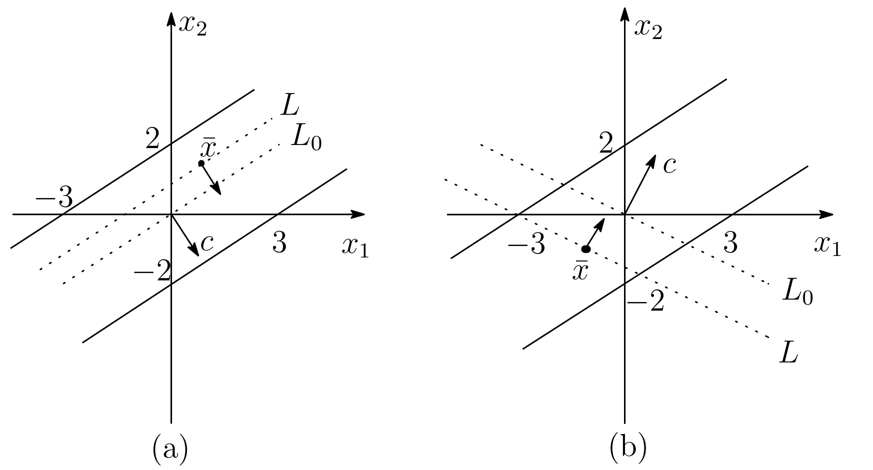
Tập nghiệm tối ưu của bài toán là cả đường thẳng (xem Hình 3.4(a)) xác định bởi phương trình

$$-2x_1 + 3x_2 = 6.$$

Nếu thay hàm mục tiêu của bài toán này bởi $f'(x) = x_1 + x_2$ thì bài toán vô nghiệm, hàm mục tiêu giảm vô hạn trên tập phương án (xem Hình 3.4(b)).



Hình 3.3. *(a) – Bài toán có duy nhất nghiệm; (b) – Bài toán vô nghiệm*



Hình 3.4. *(a) - Tập nghiệm tối ưu là cả đường thẳng; (b) - Bài toán vô nghiệm*

Nhận xét 3.1. Qua các ví dụ trên ta thấy:

i) Nếu tập chấp nhận được khác rỗng và bị chặn thì chắc chắn quy hoạch tuyến tính có nghiệm tối ưu;

ii) Nếu quy hoạch tuyến tính có nghiệm tối ưu và tập chấp nhận được D có đỉnh thì nghiệm tối ưu đạt trên ít nhất một đỉnh của D (Ví dụ 3.3 và 3.4);

iii) Trường hợp bài toán không có nghiệm tối ưu và tập chấp nhận được khác rỗng, ta có hàm mục tiêu không bị chặn dưới trên D (Ví dụ 3.4);

iv) Nếu tập chấp nhận được D không có đỉnh thì quy hoạch tuyến tính có thể không có nghiệm hoặc có nghiệm. Trường hợp có nghiệm thì nghiệm tối ưu không phải là đỉnh (Ví dụ 3.5).

3.4 Phương pháp đơn hình giải bài toán quy hoạch tuyến tính chính tắc

Xét bài toán quy hoạch tuyến tính chính tắc:

$$\min f(x) = \langle c, x \rangle \quad \text{v.đ.k.} \quad x \in D, \qquad (LP^{ct})$$

trong đó $c \in \mathbb{R}^n \setminus \{0\}$ và $D \subset \mathbb{R}^n$ là tập lồi đa diện xác định bởi

$$Ax = b, \quad x \geq 0, \qquad (3.5)$$

với A là ma trận cấp $m \times n, \ m < n$ và $b = (b_1, \dots, b_m)^T \geq 0$.

Giả thiết.

$$\text{rank}A = m \quad \text{và} \quad m < n.$$

Định lý 3.5. *Cho tập lồi đa diện khác rỗng $D = \{x \mid Ax = b, \ \ x \geq 0\}$, trong đó A là ma trận cấp $m \times n$ với các hàng $a^1, \dots, a^m$. Giả sử rằng* rank$A = k < m$ *và các hàng $a^{i_1}, \dots, a^{i_k}$ độc lập tuyến tính. Khi đó $D = D'$, trong đó D' là tập lồi đa diện xác định bởi*

$$D' = \{x \mid \langle a^{i_1}, x \rangle = b_{i_1}, \dots, \langle a^{i_k}, x \rangle = b_{i_k}, x \geq 0\}.$$

Phương pháp đơn hình của Dantzig giải bài toán quy hoạch tuyến tính dựa trên hai tính chất sau:

i) Nếu bài toán (LP^{ct}) có nghiệm tối ưu thì nghiệm tối ưu phải đạt trên ít nhất một đỉnh của tập lồi đa diện chấp nhận được D (Hệ quả 3.1).

ii) Nghiệm tối ưu địa phương của bài toán (LP^{ct}) là nghiệm tối ưu toàn cục (Định lý 3.4).

3.4.1 Mô tả hình học của phương pháp đơn hình

Thuật toán đơn hình xuất phát từ một đỉnh $x^0 \in D$. Tại đỉnh x^0 chỉ có một trong ba trường hợp sau xảy ra:

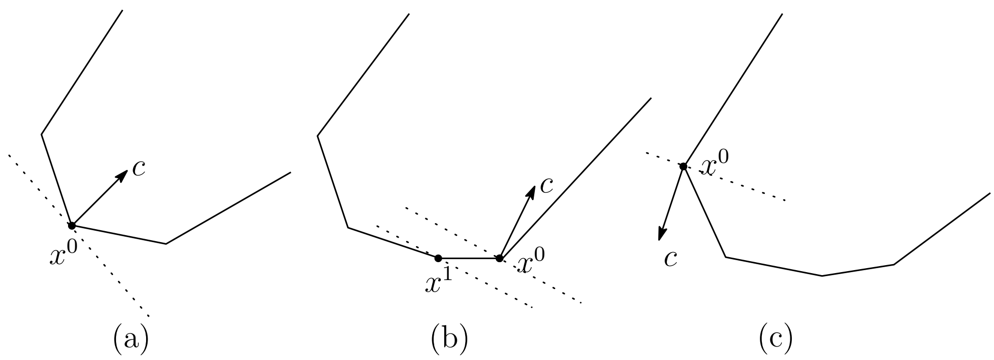
i) Trên mọi cạnh của tập nghiệm chấp nhận được xuất phát từ x^0, giá trị hàm mục tiêu đều không giảm. Khi đó x^0 là nghiệm tối ưu toàn cục của (LP^{ct}) (Bài tập) (xem Hình 3.5(a));

ii) Mọi cạnh xuất phát từ x^0, theo đó giá trị hàm mục tiêu giảm, đều là cạnh hữu hạn. Đi theo một cạnh như thế, ta sẽ đến một đỉnh x^1 kề với x^0 mà

$$\langle c, x^1 \rangle < \langle c, x^0 \rangle.$$

Gán $x^0 := x^1$ và lặp lại quá trình tính toán với đỉnh x^0 mới (xem Hình 3.5(b));

iii) Có một cạnh vô hạn xuất phát từ x^0, theo đó giá trị hàm mục tiêu giảm. Khi đó giá trị hàm mục tiêu sẽ tiến đến $-\infty$ theo cạnh này và bài toán không có nghiệm tối ưu (xem Hình 3.5(c)).



Hình 3.5

3.4.2 Cơ sở lý thuyết của phương pháp đơn hình

a. Phương án cực biên

Xét bài toán quy hoạch tuyến tính chính tắc:

$$\min f(x) = \langle c, x\rangle \text{ v.đ.k. } x \in D, \qquad (LP^{ct})$$

trong đó $c \in \mathbb{R}^n \setminus \{0\}$ và $D \subset \mathbb{R}^n$ là tập lồi đa diện xác định bởi

$$Ax = b, \quad x \geq 0, \qquad (3.5)$$

với A là ma trận cấp $m \times n$, $m < n$ và $b = (b_1, \dots, b_m)^T \geq 0$.

Ký hiệu A_j là cột thứ j của ma trận A, $j = 1, \dots, n$. Khi đó hệ (3.5) được viết dưới dạng véc tơ như sau:

$$A_1x_1 + A_2x_2 + \cdots + A_nx_n = b, \quad x_j \geq 0, j = 1, \dots, n. \qquad (3.6)$$

Xét một phương án chấp nhận được $x^0 = (x_1^0, x_2^0, \dots, x_n^0)^T \in D$, tức x^0 thoả mãn (3.6). Ký hiệu

$$J(x^0) := \{j \in \{1, \dots, n\} \mid x_j^0 > 0\}.$$

Ví dụ.

$$x^1 = (0, 5, 0, 6, 8, 0)^T \quad \Rightarrow \quad J(x^1) = \{2, 4, 5\}.$$

$$x^2 = (4, 0, 0, 3, 0, 1)^T \quad \Rightarrow \quad J(x^2) = \{1, 4, 6\}.$$

Định lý 3.6. *Cho phương án chấp nhận được $x^0 \in D$. Khi đó,*

x^0 là phương án cực biên $\Leftrightarrow$ $\{A_j \mid j \in J(x^0)\}$ độc lập tuyến tính.

Chứng minh. Không giảm tổng quát, có thể giả thiết rằng $J(x^0) = \{1, \dots, k\}$, tức $x^0 = (x_1^0, \dots, x_k^0, 0, \dots, 0)^T$ với $x_j^0 > 0$ với mọi $j = 1, \dots, k$.

($\Rightarrow$) Giả sử $x^0 \in D$ là phương án cực biên. Ta có

$$\sum_{j=1}^{n} x_j^0 A_j = \sum_{j=1}^{k} x_j^0 A_j = b. \tag{3.7}$$

Giả thiết phản chứng rằng $\{A_1, \dots, A_k\}$ phụ thuộc tuyến tính, tức tồn tại các số thực $d_j \in \mathbb{R}$ với $j = 1, \dots, k$ và có ít nhất một hệ số $d_j \neq 0$ sao cho

$$d_1 A_1 + \cdots + d_k A_k = 0 \Leftrightarrow t d_1 A_1 + \cdots + t d_k A_k = 0, \ \forall t > 0. \tag{3.8}$$

Từ (3.7) và (3.8) suy ra

$$(x_1^0 - t d_1) A_1 + \cdots + (x_k^0 - t d_k) A_k = b,$$

$$(x_1^0 + t d_1) A_1 + \cdots + (x_k^0 + t d_k) A_k = b$$

hay

$$Ay = b \ \text{ và } \ Az = b,$$

trong đó

$$y = (x_1^0 - t d_1, \dots, x_k^0 - t d_k, 0, \dots, 0)^T \in \mathbb{R}^n,$$

$$z = (x_1^0 + t d_1, \dots, x_k^0 + t d_k, 0, \dots, 0)^T \in \mathbb{R}^n.$$

Vì $x_j^0 > 0$ với mọi $j = 1, \dots, k$ nên có thể chọn $t > 0$ đủ nhỏ sao cho

$$y_j \geq 0, z_j \geq 0, j = 1, \dots, n \Rightarrow y \geq 0, \ z \geq 0.$$

$$\Rightarrow \quad y, z \in D \text{ y} \neq z.$$

Ta có

$$x^0 = \frac{1}{2} y + \frac{1}{2} z.$$

Điều này mâu thuẫn với giả thiết x^0 là phương án cực biên của D.

Suy ra $A_1, \dots, A_k$ phải độc lập tuyến tính.

($\Leftarrow$) Giả sử các véc tơ $A_1, \ldots, A_k$ độc lập tuyến tính. Giả thiết phản chứng rằng x^0 không phải pacb, tức tồn tại $y, z \in D$, $y \neq z$ sao cho

$$x^0 = \lambda y + (1-\lambda)z,\ 0 < \lambda < 1. \tag{3.9}$$

$$\Updownarrow$$

$$x_j^0 = \lambda y_j + (1-\lambda)z_j,\ 0 < \lambda < 1,\ \forall\ j = 1, \ldots, n. \tag{3.9}$$

Vì

$$\lambda > 0,\ (1-\lambda) > 0,\ \ x_j^0 = 0, , \forall i = k+1, \ldots, n$$

nên

$$y_j = 0,\ z_j = 0,\ \ \forall i = k+1, \ldots, n.$$

Vì $y, z \in D$ nên

$$y_1 A_1 + \cdots + y_k A_k = b,$$

$$z_1 A_1 + \cdots + z_k A_k = b.$$

Do $\{A_1, \ldots, A_k\}$ độc lập tuyến tính nên chỉ có duy nhất một biểu diễn của b qua chúng, tức

$$y_j = z_j\ \forall\ \ j = 1, \ldots, k \ \ \Rightarrow \ \ y = z.$$

Điều này mâu thuẫn với giả thiết phản chứng là $y \neq z$ và chứng tỏ x^0 phải là phương án cực biên. $\square$

Theo Định lý 3.6, một véc tơ $x^0 \in \mathbb{R}^n$ là phương án cực biên của bài toán quy hoạch tuyến tính (LP^{ct}) nếu nó thỏa mãn đồng thời hai điều kiện sau:

i) $x^0 \in D$ (tức x^0 có thoả mãn hệ (3.6));

ii) Các véc tơ $\{A_j \mid j \in J(x^0)\}$ độc lập tuyến tính.

Ví dụ 3.5. Cho bài toán quy hoạch tuyến tính dạng chính tắc với tập ràng buộc được xác định bởi hệ

$$\begin{aligned} x_1 + 2x_2 - x_3 \qquad\qquad &= 2 \\ -x_1 + 2x_2 \qquad + x_4 \qquad &= 4 \\ x_1 \qquad\qquad\qquad + x_5 &= 5 \\ x_1,\ x_2, \ldots, x_5 &\geq 0 \end{aligned}$$

và các véc tơ $x^1 = (0, 2, 2, 0, 5)^T$, $x^2 = (1, 1, 1, 3, 4)^T$ và $x^3 = (2, 0, 0, 6, 5)^T$. Xác định xem các véc tơ đó có phải là phương án cực biên của bài toán đã cho không?

Giải.

- ◇ Xét $x^1 = (0, 2, 2, 0, 5)^T$. Dễ thấy x^1 là một phương án chấp nhận được vì nó thỏa mãn mọi ràng buộc của bài toán. Ta có

$$J(x^1) = \{j \in \{1, \ldots, 5\} \mid x_j^1 > 0\} = \{2, 3, 5\}$$

và ba véc tơ

$$A_2 = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}, \quad A_5 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

độc lập tuyến tính. Vậy x^1 là một phương án cực biên của bài toán đang xét.

- ◇ Xét $x^2 = (1, 1, 1, 3, 4)^T$. Ta cũng có x^2 là một phương án chấp nhận được với $J(x^2) = \{1, 2, 3, 4, 5\}$. Nhưng hiển nhiên là năm véc tơ $A_1, \ldots, A_5 \in \mathbb{R}^3$ phụ thuộc tuyến tính. Vậy x^2 không phải là phương án cực biên của bài toán.

- ◇ Xét $x^3 = (2, 0, 0, 6, 5)^T$. Vì x^3 vi phạm ràng buộc chính thứ ba của bài toán nên nó không là phương án chấp nhận được. Do đó x^3 cũng không phải là phương án cực biên của bài toán.

Sau đây là các hệ quả trực tiếp của Định lý 3.6.

Hệ quả 3.2. *Số thành phần dương trong mỗi phương án cực biên của bài toán quy hoạch tuyến tính chính tắc không vượt quá m.*

Chứng minh. Vì A là ma trận cấp $m \times n$ và $\text{rank}A = m$ nên số véc tơ cột độc lập tuyến tính của A không thể vượt quá m. Kết luận của Hệ quả được suy ra trực tiếp từ Định lý 3.6. □

Hệ quả 3.3. *Số phương án cực biên của bài toán quy hoạch tuyến tính dạng chính tắc là hữu hạn.*

Chứng minh. Theo Định lý 3.6, mỗi phương án cực biên của bài toán tương ứng với $k \leq m$ véc tơ độc lập tuyến tính của A và k véc tơ độc lập tuyến tính của A xác định nhiều nhất một phương án cực biên. Vì ma trận A có n cột nên số hệ gồm k véc tơ cột của A là C_n^k. Vậy số phương án cực biên của bài toán không lớn hơn $\sum_{k=0}^{n} C_n^k = 2^n$. □

Định nghĩa. Cho phương án chấp nhận được $x^0 \in D$.

• Nếu $|J(x^0)| = m$ thì x^0 là *phương án cực biên không suy biến* x^0 của bài toán quy hoạch tuyến tính chính tắc (LP^{ct}).

• Nếu $|J(x^0)| < m$ thì x^0 là *phương án cực biên suy biến* x^0 của bài toán quy hoạch tuyến tính chính tắc (LP^{ct}).

• Nếu tất cả các phương án cực biên của tập chấp nhận được là không suy biến thì Bài toán quy hoạch tuyến tính (LP^{ct})là *bài toán qhtt không suy biến*

• Nếu có ít nhất một phương án cực biên suy biến thì Bài toán quy hoạch tuyến tính (LP^{ct})là *bài toán qhtt suy biến.*

Ví dụ 3.6. Xét quy hoạch tuyến tính có tập chấp nhận được là tập nghiệm của hệ sau

$$\begin{aligned} 3x_1 + 4x_2 + 10x_3 &= 10 \\ x_1 - x_2 + x_3 &= 1 \\ x_j \geq 0,\ j &= 1,2,3. \end{aligned}$$

Ta có $n = 3$ và $m = 2$. Dễ thấy $v^1 = (2, 1, 0)^T$ và $v^2 = (0, 0, 1)^T$ là hai phương án cực biên của bài toán này. Vì

$$J(v^1) = \{j \in \{1, 2, 3\} \mid v^1_j > 0\} = \{1, 2\}$$

có số phần tử $|J(v^1)| = 2 = m$ nên v^1 là phương án cực biên không suy biến. Còn v^2 là phương án cực biên suy biến do

$$J(v^2) = \{j \in \{1, 2, 3\} \mid v^2_j > 0\} = \{3\}$$

có số phần tử $|J(v^2)| = 1 < m$. Vì vậy, bài toán này là bài toán quy hoạch tuyến tính suy biến.

Nhận xét 3.2. Theo Định lý 3.6, ta có thể dễ dàng xác định được tất cả các phương án cực biên của bài toán quy hoạch tuyến tính chính tắc trong trường hợp số ràng buộc m và số biến n đủ nhỏ. Vì nghiệm tối ưu của bài toán quy hoạch tuyến tính đạt tại ít nhất một phương án cực biên (Hệ quả 3.1) nên nếu thêm giả thiết là tập chấp nhận được bị chặn, tức nó là đa diện lồi, thì ta có thể tìm nghiệm tối ưu của bài toán bằng cách so sánh giá trị của hàm mục tiêu tại các phương án cực biên.

Ví dụ 3.7. Xác định tất cả các phương án cực biên của bài toán quy hoạch tuyến tính với tập chấp nhận được xác định bởi hệ:

$$\begin{aligned} x_1 + 4x_2 - x_3 &= 3 \\ 5x_1 + 2x_2 + 3x_3 &= 4 \\ x_j \geq 0,\ j &= 1,2,3. \end{aligned}$$

Giải. Bài toán này có $n = 3$ biến và $m = 2$ ràng buộc chính.

◇ *Xác định các phương án cực biên không suy biến*: Theo định nghĩa, một phương án cực biên không suy biến có đúng $m = 2$ thành phần dương. Vì vậy, từ hệ trên ta có:

  – Nếu $x_1 = 0$ thì $x_2 = \frac{13}{14}$, $x_3 = \frac{5}{7}$ và $x^1 = \left(0, \frac{13}{14}, \frac{5}{7}\right)^T$ là một phương án chấp nhận được;

– Nếu $x_2 = 0$ thì hệ trên vô nghiệm;

– Nếu $x_3 = 0$ thì $x_1 = \frac{5}{9}$, $\quad x_2 = \frac{11}{18}$ và $x^2 = \left(\frac{5}{9}, \frac{11}{18}, 0\right)^T$ là một phương án chấp nhận được.

Vì $J(x^1) = \{2, 3\}$ và hai véc tơ $A_2 = (4, 2)^T$ và $A_3 = (-1, 3)^T$ độc lập tuyến tính nên x^1 là phương án cực biên không suy biến. Tương tự, x^2 cũng là phương án cực biên không suy biến vì $J(x^2) = \{1, 2\}$ và hai véc tơ $A_1 = (1, 5)^T$; $A_2 = (4, 2)^T$ độc lập tuyến tính.

◇ *Xác định các phương án cực biên suy biến*: Một phương án cực biên suy biến có ít hơn $m = 2$ thành phần dương, tức nó phải có ít nhất $n - m + 1 = 2$ thành phần bằng 0. Dễ thấy bài toán này không có phương án cực biên suy biến. Vì vậy, nó là bài toán quy hoạch tuyến tính không suy biến.

b. Điều kiện tối ưu

Xét bài toán quy hoạch tuyến tính chính tắc

$$\min f(x) = \langle c, x \rangle \ \text{ v.đ.k. } \ x \in D, \qquad (LP^{ct})$$

trong đó $c \in \mathbb{R}^n \setminus \{0\}$ và tập chấp nhận được

$$D = \{x \in R^n \mid Ax = b, x \geq 0\},$$

trong đó $b = (b_1, \dots, b_m)^T \geq 0$, A là ma trận cấp $m \times n$ với các cột $A_1, \dots, A_n$, $\text{rank}A = m$ và $m < n$.

Định nghĩa. Một bộ gồm m véc tơ cột độc lập tuyến tính $B = \{A_{j_1}, A_{j_2}, \dots, A_{j_m}\}$ của ma trận A cấp $m \times n$ có $\text{rank}A = m$ được gọi là một *cơ sở* của ma trận A.

Cho $x^0 = (x_1^0, \dots, x_n^0)^T$ là một phương án cực biên của bài toán (LP^{ct}). Theo Định lý 3.6, các véc tơ $\{A_j \mid j \in J(x^0)\}$ độc lập tuyến tính. Vì $\text{rank}A = m$ nên:

- ◇ Nếu x^0 là phương án cực biên không suy biến, tức $|J(x^0)| = m$, thì $B = \{A_j \mid j \in J(x^0)\}$ là cơ sở duy nhất của A tương ứng với x^0.

- ◇ Nếu x^0 là phương án cực biên suy biến, tức $|J(x^0)| < m$, thì ta bổ sung thêm các véc tơ cột của A thuộc tập $\{A_j \mid j \notin J(x^0)\}$ vào bộ véc tơ $\{A_j \mid j \in J(x^0)\}$ để nhận được bộ m véc tơ độc lập tuyến tính $\{A_j \mid j \in J\}$ với $J \supset J(x^0)$ và $|J| = m$. Khi đó, $\bar{B} = \{A_j \mid j \in J\}$ là một cơ sở của ma trận A.

Ví dụ 3.8. Xét quy hoạch tuyến tính như ở Ví dụ 3.8. Ta đã biết $v^1 = (2, 1, 0)^T$, với $J(v^1) = \{1, 2\}$, là phương án cực biên không suy biến nên có duy nhất một cơ sở tương ứng với nó là $\{A_1, A_2\} = \Big\{(3, 1)^T, (4, -1)^T\Big\}$.

Phương án cực biên suy biến $v^2 = (0, 0, 1)^T$ có tập $J(v^2) = \{3\}$ chỉ có một phần tử $(1 < m = 2)$ nên ta bổ sung thêm một véc tơ cột của A để nhận được cơ sở của A tương ứng với v^2. Dễ thấy tương ứng với v^2 có hai cơ sở là:

i) Cơ sở $\bar{B}^1 = \{A_1, A_3\} = \Big\{(3, 1)^T, (10, 1)^T\Big\}$;

ii) Cơ sở $\bar{B}^2 = \{A_2, A_3\} = \Big\{(4, -1)^T, (10, 1)^T\Big\}$.

Từ đây, để đơn giản, ta xét bài toán quy hoạch tuyến tính chính tắc (LP^{ct}) với giả thiết thêm rằng: *Bài toán* (LP^{ct}) *là không suy biến và biết trước một phương án cực biên của bài toán này.*

Giả sử ta đã biết phương án cực biên không suy biến $x^0 = (x_1^0,\ x_2^0,\ \ldots,\ x_n^0)^T$ có $|J(x^0)| = m$. Hệ các véc tơ độc lập tuyến tính $\{A_j \mid j \in J(x^0)\}$ là cơ sở duy nhất của ma trận A tương ứng với x^0. Ta gọi:

$J(x^0) = \{j \mid x_j^0 > 0\}$	là *tập chỉ số cơ sở*,
$\{1, \ldots, n\} \setminus J(x^0)$	là *tập chỉ số phi cơ sở*,
$\{x_j \mid j \in J(x^0)\}$	là các *biến cơ sở*,
$\{A_j \mid j \in J(x^0)\}$	là các *véc tơ cơ sở*,
$\{x_j \mid j \notin J(x^0)\}$	là các *biến phi cơ sở*,
$\{A_j \mid j \notin J(x^0)\}$	là các *véc tơ phi cơ sở*.

Do $\{A_j | j \in J(x^0)\}$ là cơ sở của ma trận A nên mỗi véc tơ cột A_k, $k \in \{1, 2, \ldots, n\}$, được biểu diễn dưới dạng

$$A_k = \sum_{j \in J(x^0)} z_{jk} A_j, \tag{3.10}$$

tức

$$\sum_{j \in J(x^0)} a_{ij} z_{jk} = a_{ik},\ i = 1, \ldots, m$$

và bộ số thực z_{jk}, $j \in J(x^0)$ được xác định duy nhất. Vì $x_j^0 = 0$ với mọi $j \notin J(x^0)$ nên

$$\sum_{j \in J(x^0)} x_j^0 A_j = b. \tag{3.11}$$

Giá trị hàm mục tiêu tại x^0 là

$$f(x^0) = \langle c, x^0 \rangle = \sum_{j=1}^{n} x_j^0 c_j = \sum_{j \in J(x^0)} x_j^0 c_j. \tag{3.12}$$

Đại lượng

$$\Delta_k = \sum_{j \in J(x^0)} z_{jk} c_j - c_k, \quad k \in \{1, 2, \ldots, n\}, \tag{3.13}$$

được gọi là *ước lượng* của biến x_k.

Sau đây là điều kiện đủ để phương án cực biên x^0 là phương án tối ưu của bài toán (LP^{ct}).

Định lý 3.7. *Nếu phương án cực biên x^0 thỏa mãn*

$$\Delta_k \leq 0 \ \text{ với mọi } \ k \notin J(x^0)$$

thì x^0 là một phương án tối ưu của bài toán quy hoạch tuyến tính chính tắc (LP^{ct}).

Chứng minh. Vì phương án cực biên $x^0 = (x_1^0, \ldots, x_n^0)^T$ có $x_k^0 = 0$ với mọi $k \notin J(x^0)$ nên

$$f(x^0) = \langle c, x^0 \rangle = \sum_{j \in J(x^0)} c_j x_j^0$$

và

$$b = \sum_{j \in J(x^0)} x_j^0 A_j. \tag{3.14}$$

Để chứng tỏ x^0 là phương án tối ưu, với phương án chấp nhận được bất kỳ $y \in D$, ta sẽ chứng minh rằng

$$f(x) = \langle c, x^0 \rangle \leq f(y) = \langle c, y \rangle.$$

Thật vậy, vì $y \in D$ nên

$$\begin{aligned} b = \sum_{j=1}^{n} y_j A_j &= \sum_{j \in J(x^0)} y_j A_j + \sum_{k \notin J(x^0)} y_k A_k \\ &\overset{(3.10)}{=} \sum_{j \in J(x^0)} y_j A_j + \sum_{k \notin J(x^0)} y_k \Big(\sum_{j \in J(x^0)} z_{jk} A_j \Big) \\ &= \sum_{j \in J(x^0)} \Big(y_j + \sum_{k \notin J(x^0)} y_k z_{jk} \Big) A_j. \end{aligned} \tag{3.15}$$

Do phép biểu diễn b qua cơ sở là duy nhất nên từ (3.14) và (3.15) ta có

$$x_j^0 = y_j + \sum_{k \notin J(x^0)} y_k z_{jk}, \ \forall j \in J(x^0).$$

Suy ra

$$y_j = x_j^0 - \sum_{k \notin J(x^0)} y_k z_{jk}, \ \forall j \in J(x^0) \tag{3.16}$$

và

$$\begin{aligned}
f(y) = \langle c, y\rangle &= \sum_{j\in J(x^0)} c_j y_j + \sum_{k\notin J(x^0)} c_k y_k \\
&\overset{(3.16)}{=} \sum_{j\in J(x^0)} c_j\Big(x_j^0 - \sum_{k\notin J(x^0)} y_k z_{jk}\Big) + \sum_{k\notin J(x^0)} c_k y_k \\
&= \sum_{j\in J(x^0)} c_j x_j^0 - \sum_{k\notin J(x^0)} \Big(\sum_{j\in J(x^0)} c_j z_{jk} - c_k\Big) y_k \\
&\overset{(3.13)}{=} f(x^0) - \sum_{k\notin J(x^0)} \Delta_k y_k. \qquad (3.17)
\end{aligned}$$

Theo giả thiết, $\Delta_k \le 0$ với mọi $k \notin J(x^0)$ và $y_k \ge 0$ với mọi $k = 1, \ldots, n$ nên $\sum_{k\notin J(x^0)} \Delta_k y_k \le 0$. Do đó $f(y) \ge f(x^0)$. Ta có điều phải chứng minh. $\square$

Hệ quả 3.4. *Giả sử x^0 là phương án cực biên tối ưu của bài toán quy hoạch tuyến tính (LP^{ct}). Nếu $\Delta_k < 0$ với mọi $k \notin J(x^0)$ thì x^0 là phương án tối ưu duy nhất. Ngược lại, nếu tồn tại $k \notin J(x^0)$ sao cho $\Delta_k = 0$ thì x^0 không phải là phương án tối ưu duy nhất.*

Chứng minh. Giả sử x^0 là phương án tối ưu thỏa mãn $\Delta_k < 0$ với mọi $k \notin J(x^0)$. Khi đó x^0 là phương án tối ưu duy nhất của bài toán (LP^{ct}). Thật vậy, giả thiết phản chứng rằng x^0 không phải phương án tối ưu duy nhất, tức bài toán (LP^{ct}) có thêm ít nhất một phương án tối ưu $x^1 \neq x^0$. Ta có

$$f(x^1) = f(x^0). \qquad (3.18)$$

Theo (3.17) trong chứng minh Định lý 3.7, vì x^1 là phương án chấp nhận được nên

$$f(x^1) = f(x^0) - \sum_{k\notin J(x^0)} \Delta_k x_k^1. \qquad (3.19)$$

Kết hợp (3.18) và (3.19) suy ra

$$\sum_{k\notin J(x^0)} \Delta_k x_k^1 = 0.$$

Do $\Delta_k < 0$ và $x_k^1 \ge 0$ với mọi $k \notin J(x^0)$ nên phải có $x_k^1 = 0$ với mọi $k \notin J(x^0)$. Vì x^0 và x^1 đều là các phương án chấp nhận được nên

$$\sum_{j=1}^n x_j^0 A_j = \sum_{j=1}^n x_j^1 A_j \quad \text{hay} \quad \sum_{j\in J(x^0)} x_j^0 A_j = \sum_{j\in J(x^0)} x_j^1 A_j.$$

Do $\{A_j, j \in J(x^0)\}$ là cơ sở nên phải có $x_j^0 = x_j^1$ với mọi $j \in J(x^0)$. Vậy, $x_j^1 = x_j^0$ với mọi $j = 1, \ldots, n$ hay $x^0 = x^1$. Điều ngược lại là hiển nhiên. $\square$

Định lý 3.8. *Cho x^0 là một phương án cực biên của bài toán quy hoạch tuyến tính chính tắc (LP^{ct}). Khi đó,*

i) *Nếu tồn tại $k \notin J(x^0)$ sao cho*

$$\Delta_k > 0 \ \ \text{và } z_{jk} \leq 0 \ \ \forall j \in J(x^0)$$

thì hàm mục tiêu giảm vô hạn trên tập chấp nhận được và bài toán không có nghiệm tối ưu;

ii) *Nếu tồn tại $k \notin J(x^0)$ sao cho*

$$\Delta_k > 0 \ \ \text{và } \ \exists j \in J(x^0) \text{ sao cho } z_{jk} > 0 \tag{3.20}$$

thì ta có thể chuyển được tới phương án cực biên x^1 tốt hơn phương án cực biên x^0, nghĩa là $\langle c, x^1\rangle < \langle c, x^0\rangle$.

Chứng minh. Giả sử tồn tại chỉ số $k \notin J(x^0)$ sao cho $\Delta_k > 0$. Ta có

$$\sum_{j=1}^{n} x_j^0 A_j = \sum_{j\in J(x^0)} x_j^0 A_j = b \tag{3.21}$$

và

$$\sum_{j\in J(x^0)} z_{jk} A_j = A_k \quad \Rightarrow \quad \sum_{j\in J(x^0)} \theta z_{jk} A_j = \theta A_k \text{ với } \theta > 0. \tag{3.22}$$

Trừ (3.21) cho (3.22) và chuyển vế θA_k, ta được

$$\sum_{j\in J(x^0)} (x_j^0 - \theta z_{jk}) A_j + \theta A_k = b. \tag{3.23}$$

Đặt $z^k = (z_1^k, \ldots, z_n^k)^T$, với

$$z_j^k = \begin{cases} -z_{jk}, & \forall j \in J(x^0) \\ 0, & j \notin J(x^0) \text{ và } j \neq k \\ 1, & j = k \end{cases} \tag{3.24}$$

và

$$\bar{x}(\theta) = x^0 + \theta z^k.$$

Khi đó (3.23) có thể viết được dưới dạng

$$A\bar{x}(\theta) = b. \tag{3.25}$$

Ta xét hai trường hợp:

- *Trường hợp 1: Có $z_{jk} \le 0$ với mọi $j \in J(x^0)$.* Trong trường hợp này:

$$\bar{x}(\theta) = x^0 + \theta z^k \ge 0, \ \forall \theta \ge 0. \tag{3.26}$$

Hệ thức (3.25) và (3.26) chứng tỏ tất cả các điểm $\bar{x}(\theta)$, với $\theta \ge 0$, nằm trên tia xuất phát từ x^0 theo hướng z^k đều thuộc tập chấp nhận được. Do $\Delta_k > 0$, tính toán trực tiếp ta có

$$\begin{aligned} f(\bar{x}(\theta)) &= \sum_{j \in J(x^0)} (x_j^0 - \theta z_{jk}) c_j + \theta c_k \\ &= \sum_{j \in J(x^0)} x_j^0 c_j - \theta \Big(\sum_{j \in J(x^0)} z_{jk} c_j - c_k \Big) \\ &= \langle c, x^0 \rangle - \theta \Delta_k = f(x^0) - \theta \Delta_k \to -\infty \text{ khi } \theta \to +\infty, \end{aligned}$$

tức giá trị hàm mục tiêu giảm vô hạn trên tập phương án. Bài toán không có nghiệm tối ưu.

- *Trường hợp 2: Tồn tại $z_{jk} > 0$ với ít nhất một $j \in J(x^0)$.* Rõ ràng rằng, trong trường hợp này ta không có $\bar{x}(\theta) \ge 0$ với mọi $\theta > 0$. Tuy nhiên, ta vẫn có hệ thức (3.25) và $\Delta_k > 0$ nên

$$f(\bar{x}(\theta)) = f(x^0) - \theta \Delta_k < f(x^0) \ \text{ với mỗi } \theta > 0. \tag{3.27}$$

Vấn đề là cần chọn số $\theta_0 > 0$ để $\bar{x}(\theta_0) \ge 0$, tức $\bar{x}(\theta_0)$ là phương án chấp nhận được, và hơn nữa, nó còn phải là phương án cực biên kề với x^0 theo hướng z^k.

Đặt

$$\theta_0 = \min \left\{ \frac{x_j^0}{z_{jk}} \ \Big| \ z_{jk} > 0, \ j \in J(x^0) \right\} = \frac{x_r^0}{z_{rk}} \ \text{ với } r \in J(x^0). \tag{3.28}$$

Vì $\bar{x}(\theta) = x^0 + \theta z^k$ nên $\bar{x}(\theta) \ge 0$ với $0 \le \theta \le \theta_0$. Kết hợp sự kiện này với (3.25) ta có các véc tơ $\bar{x}(\theta)$ với $0 \le \theta \le \theta_0$ là các phương án chấp nhận được của bài toán (LP^{ct}). Cụ thể, chúng thuộc cạnh hữu hạn của tập chấp nhận được xuất phát từ x^0 theo hướng z^k. Đặt $x^1 := \bar{x}(\theta_0)$. Các tọa độ của x^1 được xác định bởi

$$x_j^1 = \begin{cases} x_j^0 - \theta_0 z_{jk}, & \forall j \in J(x^0) \\ 0, & \forall j \notin J(x^0) \text{ và } j \ne k \\ \theta_0, & j = k. \end{cases}$$

Kết hợp công thức trên và (3.28) ta có

$$x_r^1 = x_r^0 - \theta_0 z_{rk} = 0 \ \text{ và } \ x_k^1 = \theta_0 > 0.$$

Ký hiệu $J(x^1) = \{j \in \{1, \ldots, n\} \mid x_j^1 > 0\}$. Dễ thấy $J(x^1) = (J(x^0) \setminus \{r\}) \cup \{k\}$ và $|J(x^1)| = |J(x^0)| = m$. Có thể kiểm tra được hệ $\{A_j \mid j \in J(x^1)\}$ độc lập tuyến tính. Theo Định lý 3.6, ta có x^1 là phương án cực biên với bộ chỉ số cơ sở $J(x^1)$. □

Chú ý 3.2.

i) Nếu tồn tại chỉ số $k \notin J(x^0)$ sao cho

$$\Delta_k > 0 \ \text{ và } \ \exists j \in J(x^0) \text{ sao cho } z_{jk} > 0$$

thì ta có thể chuyển được đến phương án cực biên mới x^1 tốt hơn x^0. Ta có

$$\langle c, x^1 \rangle = \langle c, x^0 \rangle - \theta_0 \Delta_k < \langle c, x^0 \rangle$$

và

$$J(x^1) = (J(x^0) \setminus \{r\}) \cup \{k\},$$

trong đó, chỉ số r được xác định bởi

$$\theta_0 = \min \left\{ \frac{x_j^0}{z_{js}} \;\middle|\; z_{js} > 0, \ j \in J(x^0) \right\} = \frac{x_r^0}{z_{rs}} \ \text{ với } r \in J(x^0).$$

Khi đó:

- ⋄ Biến x_k (*biến phi cơ sở đối với* x^0) trở thành biến cơ sở đối với x^1;
- ⋄ Biến x_r (*biến cơ sở đối với* x^0) trở thành biến phi cơ sở đối với x^1.
- ⋄ Hai phương án cực biên x^0 và x^1 *kề nhau* nếu *cơ sở của chúng chỉ sai khác nhau đúng một véc tơ.*

ii) Nếu có nhiều chỉ số $k \notin J(x^0)$ cùng thỏa mãn

$$\Delta_k > 0 \ \text{ và } \ \exists j \in J(x^0) \text{ sao cho } z_{jk} > 0$$

thì ta chọn chỉ số $k = s$ với s là chỉ số thỏa mãn

$$\Delta_s = \max\{\Delta_k \mid \Delta_k > 0\}.$$

Ta gọi s là *chỉ số của véc tơ đưa vào cơ sở mới.*

Khi đó, đưa véc tơ A_s vào cơ sở và đưa A_r ra khỏi cơ sở, với chỉ số r (*chỉ số của véc tơ đưa ra khỏi cơ sở cũ*) được xác định bởi

$$\theta_0 = \min \left\{ \frac{x_j^0}{z_{js}} \;\middle|\; z_{js} > 0, \ j \in J(x^0) \right\} = \frac{x_r^0}{z_{rs}} \ \text{ với } r \in J(x^0) \tag{3.29}$$

và nhận được phương án cực biên mới $x^1 = (x_1^1, x_2^1, \ldots, x_n^1)^T$ với

$$x_j^1 = \begin{cases} x_j^0 - \dfrac{x_r^0}{z_{rs}} z_{js}, & \forall j \in J(x^0) \text{ (trong đó } x_r^1 = 0) \\ 0, & \forall j \notin J(x^0) \text{ và } j \neq s \\ \dfrac{x_r^0}{z_{rs}}, & j = s. \end{cases} \tag{3.30}$$

3.4.3 Thuật toán đơn hình giải bài toán qui hoạch tuyến tính chính tắc

Xét bài toán quy hoạch tuyến tính dạng chính tắc

$$\min \quad f(x) = \langle c, x \rangle \} \quad \text{v.đ.k.} \quad x \in D, \tag{LP^{ct}}$$

trong đó $c = (c_1, c_2, \ldots, c_n)^T \in \mathbb{R}^n \setminus \{0\}$,

$$D = \{xin\mathbb{R}^n \mid Ax = b,\ x \geq 0\},$$

A là ma trận cấp $m \times n$ và véc tơ $b = (b_1, \ldots, b_m)^T \geq 0$ với các giả thiết:

i) $\text{rank}A = m$,

ii) $m < n$;

iii) Bài toán quy hoạch tuyến tính (LP^{ct}) không suy biến;

iv) Biết trước một phương án cực biên x^0.

Cho phương án cực biên $x^0 \in D$. Khi đó:

- Nếu $\Delta_k \leq 0\ \forall\ \ k \notin J(x^0)$ thì $x^0 \in \text{Argmin}(LP^{ct})$.
- Nếu tồn tại $k \notin J(x^0)$ sao cho

$$\Delta_k > 0 \ \text{ và } z_{jk} \leq 0 \ \ \forall j \in J(x^0)$$

thì $\text{Argmin}(LP^{ct}) = \emptyset$ và giá trị tối ưu là $-\infty$.

• Nếu tồn tại chỉ số $k \notin J(x^0)$ sao cho

$$\Delta_k > 0 \text{ và } \exists j \in J(x^0) \text{ sao cho } z_{jk} > 0$$

thì ta có thể chuyển được đến phương án cực biên mới x^1 tốt hơn x^0. Ta có

$$\langle c, x^1 \rangle = \langle c, x^0 \rangle - \theta_0 \Delta_k < \langle c, x^0 \rangle$$

và

$$J(x^1) = (J(x^0) \setminus \{r\}) \cup \{k\},$$

trong đó, chỉ số r được xác định bởi

$$\theta_0 = \min\left\{ \frac{x_j^0}{z_{js}} \;\middle|\; z_{js} > 0,\ j \in J(x^0) \right\} = \frac{x_r^0}{z_{rs}} \text{ với } r \in J(x^0).$$

Khi đó:

◇ Biến x_k (*biến phi cơ sở đối với* x^0) trở thành biến cơ sở đối với x^1;

◇ Biến x_r (*biến cơ sở đối với* x^0) trở thành biến phi cơ sở đối với x^1.

◇ Hai phương án cực biên x^0 và x^1 *kề nhau* nếu *cơ sở của chúng chỉ sai khác nhau đúng một véc tơ*.

• Nếu có nhiều chỉ số $k \notin J(x^0)$ cùng thỏa mãn

$$\Delta_k > 0 \text{ và } \exists j \in J(x^0) \text{ sao cho } z_{jk} > 0$$

thì ta chọn chỉ số $k = s$ với s là chỉ số thỏa mãn

$$\Delta_s = \max\{\Delta_k \mid \Delta_k > 0\}.$$

Ta gọi s là *chỉ số của véc tơ đưa vào cơ sở mới*.

Khi đó, đưa véc tơ A_s vào cơ sở và đưa A_r ra khỏi cơ sở, với chỉ số r (*chỉ số của véc tơ đưa ra khỏi cơ sở cũ*) được xác định bởi

$$\theta_0 = \min\left\{ \frac{x_j^0}{z_{js}} \;\middle|\; z_{js} > 0,\ j \in J(x^0) \right\} = \frac{x_r^0}{z_{rs}} \text{ với } r \in J(x^0) \tag{3.29}$$

và nhận được phương án cực biên mới $x^1 = (x_1^1, x_2^1, \ldots, x_n^1)^T$ với

$$x_j^1 = \begin{cases} x_j^0 - \dfrac{x_r^0}{z_{rs}} z_{js}, & \forall j \in J(x^0) \text{ (trong đó } x_r^1 = 0) \\ 0, & \forall j \notin J(x^0) \text{ và } j \neq s \\ \dfrac{x_r^0}{z_{rs}}, & j = s. \end{cases} \tag{3.30}$$

Thuật toán 3.2.

Bước khởi tạo. Xuất phát từ một pacbn x^0 và cơ sở $\{A_j,\ j \in J(x^0)\}$.

Bước 1. Tính giá trị hàm mục tiêu

$$f(x^0) = \sum_{j \in J(x^0)} c_j x_j^0.$$

Bước 2. Với mỗi $k \notin J(x^0)$, xác định các số z_{jk} bằng việc giải hệ

$$\sum_{j \in J(x^0)} z_{jk} A_j = A_k$$

và tính các ước lượng

$$\Delta_k = \sum_{j \in J(x^0)} z_{jk} c_j - c_k.$$

Bước 3. (*Kiểm tra điều kiện tối ưu*).
  If $\Delta_k \leq 0\ \forall k \notin J(x^0)$ **Then** STOP (x^0 *là nghiệm tối ưu*)

Bước 4. (*Kiểm tra bài toán không có lời giải*)
  If $\exists k \notin J(x^0)$ sao cho $\Delta_k > 0$ và $z_{jk} \leq 0 \forall j \in J(x^0)$

  Then STOP (Argmin$(LP^{ct}) = \emptyset$ và $f_{opt} = -\infty$)

Bước 5. (*Xây dựng phương án cực biên mới*)

- ⋄ Tìm véc tơ A_s để đưa vào cơ sở mới, trong đó chỉ số s thỏa mãn

$$\Delta_s = \max\{\Delta_k \mid \Delta_k > 0\}.$$

- ⋄ Tìm véc tơ A_r để đưa ra khỏi cơ sở cũ với chỉ số r thỏa mãn

$$\theta_0 = \min\left\{ \frac{x_j^0}{z_{js}} \;\middle|\; z_{js} > 0 \right\} = \frac{x_r^0}{z_{rs}}.$$

- ⋄ Xây dựng phương án cực biên mới là x^1 theo công thức (3.30) với tập chỉ số cơ sở mới là $J(x^1) = (J(x^0) \setminus \{r\}) \cup \{s\}$.

- ⋄ Đặt $x^0 := x^1$ và quay lại Bước 1.

3.4.4 Thuật toán đơn hình dạng bảng

a. Bảng đơn hình

Giả sử $x^0 = (x_1^0, x_2^0, \ldots, x_n^0)^T$ là phương án cực biên tương ứng với bộ chỉ số cơ sở $J_B = J(x^0) = \{j_1, \ldots, j_m\}$ và cơ sở đơn vị $B = \{A_{j_1}, \ldots, A_{j_m}\}$.

Bảng 3.1

Hệ số C_B	Cơ sở B	Phương án	c_1 A_1	$\ldots$	c_k A_k	$\ldots$	c_n A_n	θ
c_{j_1}	A_{j_1}	$x_{j_1}^0$	$z_{j_1 1}$	$\ldots$	$z_{j_1 k}$	$\ldots$	$z_{j_1 n}$	θ_{j_1}
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
c_j	A_j	x_j^0	z_{j1}	$\ldots$	z_{jk}	$\ldots$	z_{jn}	θ_j
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
c_{j_m}	A_{j_m}	$x_{j_m}^0$	$z_{j_m 1}$	$\ldots$	$z_{j_m k}$	$\ldots$	$z_{j_m n}$	θ_{j_m}
		$f(x^0)$	Δ_1	$\ldots$	Δ_k	$\ldots$	Δ_n	

Bảng đơn hình gồm $n+4$ cột và $m+2$ dòng.

Cột 1. (*Hệ số* C_B) Ghi giá trị hệ số hàm mục tiêu tương ứng với biến cơ sở.

Cột 2. (*Cơ sở* B) Ghi tên các véc tơ cơ sở. **Chú ý**: tên các véc tơ này phải được ghi theo thứ tự $A_{j_1}, \ldots, A_{j_m}$ sao cho ma trận $B = (A_{j_1}, \ldots, A_{j_m}) \equiv I_m$.

Cột 3. (*Pacb*) Ghi giá trị của các biến cơ sở của phương án cực biên đang xét.

n cột tiếp theo. Cột thứ $3+k$ ứng với tên véc tơ A_k, $k = 1, \ldots, n$. Phía trên tên mỗi cột A_k ghi giá trị c_k tương ứng. Trong cột A_k, ghi giá trị các hệ số $z_{jk}, j \in J_B$ là nghiệm của hệ

$$\sum_{j \in J_B} z_{jk} A_j = A_k, \quad k = 1, \ldots, n.$$

Cột cuối cùng. Dành để ghi tỷ số $\theta_j, j \in J_B$ (xem Bước 3 − Thuật toán 3.3).

Dòng cuối cùng. Tại vị trí dưới cột thứ ba, ghi giá trị

$$f(x^0) = \sum_{j \in J_B} c_j x_j^0.$$

Tại vị trí dưới cột ứng với véc tơ A_k, $k \in \{1, \ldots, n\}$, ghi ước lượng

$$\Delta_k = \sum_{j \in J_B} z_{jk} c_j - c_k.$$

Ta có $\Delta_j = 0$ với mọi $j \in J_B$.

b. Thuật toán đơn hình dạng bảng

Thuật toán 3.3.
Bước 0. (*Chuẩn bị*) Xây dựng bảng đơn hình xuất phát tương ứng với pacb x^0.

Bước 1. (*Kiểm tra điều kiện tối ưu*) Xét dòng cuối của bảng.
If $\Delta_k \leq 0$ với mọi $k = 1, \ldots, n$ **Then** Dừng thuật toán
(*Nghiệm tối ưu là phương án cực biên tương ứng với bảng này*)

Bước 2. (*Kiểm tra bài toán không có lời giải*)
If Tồn tại $k \notin J_B$ sao cho $\Delta_k > 0$ và $z_{jk} \leq 0$ với mọi $j \in J_B$
Then Dừng thuật toán (*Bài toán không có lời giải, giá trị tối ưu là* $-\infty$)

Bước 3. Thực hiện:

- ◇ *Tìm cột quay.* Xác định véc tơ A_s để đưa vào cơ sở mới với chỉ số s thỏa mãn $\Delta_s = \max\{\Delta_k \mid \Delta_k > 0\}$. Cột tương ứng với véc tơ A_s được gọi là *cột quay.*

- ◇ *Tìm dòng quay.* Tính các $\theta_j, j \in J_B$, như sau:

$$\theta_j = \begin{cases} \dfrac{x_j^0}{z_{js}} & \text{nếu } z_{js} > 0,\ j \in J_B \\ +\infty & \text{nếu } z_{js} \leq 0,\ j \in J_B \end{cases}$$

và xác định

$$\theta_r = \min\{\theta_j \mid j \in J_B\}.$$

Dòng r được gọi là *dòng quay.* Phần tử z_{rs} nằm trên giao của dòng quay và cột quay được gọi là *phần tử chính* của phép quay. Các phần tử z_{js} $(j \neq r)$ được gọi là các *phần tử quay.*

Bước 4. (*Chuyển bảng mới tương ứng với phương án cực biên mới*) Thực hiện:

- $\diamond$ Trong cột thứ nhất (cột hệ số C_B) thay giá trị c_r bởi c_s. Trong cột thứ hai (cột tên các véc tơ cơ sở), thay tên A_r bởi A_s.

- $\diamond$ Chia các phần tử của dòng quay cho phần tử chính, ta được dòng mới (có số 1 tại vị trí của z_{rs} cũ) gọi là *dòng chính*, tức ta có quy tắc là

$$\textbf{Dòng chính (mới)} := \frac{\textbf{Dòng quay (cũ)}}{\textbf{phần tử chính}}.$$

- $\diamond$ Biến đổi mỗi dòng còn lại theo quy tắc

  Dòng mới := (- Phần tử quay t. ứng) $\times$ Dòng chính + Dòng cũ t. ứng.

  Ta được số 0 ở mọi vị trí còn lại của cột quay cũ (*sau phép quay thì trong bảng mới ta có $\Delta_s = 0$ vì lúc này s là chỉ số cơ sở của phương án cực biên mới và A_s trở thành véc tơ đơn vị cơ sở*).

- $\diamond$ Quay lại Bước 1 với bảng mới.

3.5 Tìm phương án cực biên xuất phát và cơ sở xuất phát

3.5.1 Trường hợp bài toán có dạng chuẩn tắc

Xét bài toán quy hoạch tuyến tính dạng chuẩn tắc

$$\min f(x) = \langle c, x\rangle \quad \text{v.đ.k} \quad x \in D, \qquad (P_0)$$

trong đó $c \in \mathbb{R}^n \setminus \{0\}$ và $D = \{x \in \mathbb{R}^n \mid Ax \le b,\ x \ge 0\}$, véc tơ $b \in \mathbb{R}^m_+$ và A là ma trận cấp $m \times n$.

Bài toán quy hoạch tuyến tính chính tắc tương ứng với bài toán (P_0) là

$$\min \bar{f}(x) = \langle d, x\rangle \quad \text{v.đ.k} \quad x \in \bar{D}, \qquad (P_1)$$

trong đó $d = (c, 0) = (c_1, \ldots, c_n, 0, ..., 0)^T \in \mathbb{R}^{n+m}$ và

$$\bar{D} = \{v = (x, u) \in \mathbb{R}^{n+m} \mid Ax + u = b,\ x, u \ge 0\}.$$

Dễ thấy, $v^0 = (0, b)$ là phương án cực biên không suy biến của bài toán quy hoạch tuyến tính dạng chính tắc (P_1) tương ứng với cơ sở đơn vị.

3.5.2 Trường hợp bài toán có dạng chính tắc

Xét bài toán quy hoạch tuyến tính dạng chính tắc

$$\min \{\langle c, x\rangle \mid x \in D\}, \qquad (LP^{ct})$$

trong đó tập lồi đa diện $D \subset \mathbb{R}^n$ xác định bởi $D = \{x \in \mathbb{R}^n \mid Ax = b,\ x \ge 0\}$, với A là ma trận cấp $m \times n$, $b \in \mathbb{R}^m$ và $b \ge 0$.

a. Trường hợp đặc biệt

Nếu mỗi ràng buộc chính đều có một biến số với hệ số bằng 1 đồng thời biến số này không có mặt trong các phương trình khác (gọi là biến số cô lập với hệ số bằng 1) thì ta có ngay một phương án cực biên với cơ sở đơn vị. Chẳng hạn, tập chấp nhận được của quy hoạch tuyến tính đang xét là tập nghiệm của hệ

$$\begin{array}{l} x_1 \qquad\qquad\qquad + a_{1\,m+1}x_{m+1} + a_{1\,m+2}x_{m+2} + \cdots + a_{1n}x_n = b_1 \\ \qquad x_2 \qquad\qquad + a_{2\,m+1}x_{m+1} + a_{2\,m+2}x_{m+2} + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad\qquad \cdots\cdots\cdots \\ \qquad\qquad x_m + a_{m\,m+1}x_{m+1} + a_{m\,m+2}x_{m+2} + \cdots + a_{mn}x_n = b_m \\ \qquad\qquad\qquad\qquad\qquad\qquad x_1, x_2, \ldots, x_n \ge 0, \end{array}$$

trong đó $b_i \ge 0$ với mọi $i = 1, \ldots, m$. Trong trường hợp may mắn này, ta có ngay một phương án cực biên với cơ sở đơn vị, cụ thể $x^0 = (b_1, \ldots, b_m, 0, 0, \ldots, 0)^T$ ứng với cơ sở $\{A_1, A_2, \ldots, A_m\} = I_m$..

b. Trường hợp tổng quát
Đưa thêm $u = (u_1, \dots, u_m)^T \geq 0$ và xét tập lồi đa diện

$$D' = \{(x, u) \in \mathbb{R}^{n+m} \mid Ax + u = b,\ (x, u) \geq 0\}.$$

Mệnh đề 3.1. *Điểm x^0 là phương án cực biên của D khi và chỉ khi $(x^0, 0)$ là phương án cực biên của D'.*

• Xét bài toán quy hoạch tuyến tính phụ

$$\begin{array}{ll} \min\ g(x,u) = u_1 + u_2 + \cdots + u_m & \\ \text{v.đ.k.} \qquad\quad (x,u) \in D'. & \end{array} \qquad (LP_*^{ct})$$

- Các biến $u_1, \dots, u_m$ được gọi là các *biến giả*;

- Các véc tơ cột của ma trận ràng buộc tương ứng với các biến giả được gọi là các *véc tơ giả*.

Nhận xét 3.3.

i) Dễ thấy $(x, u) = (0, b) \in D'$ nên $D' \neq \emptyset$. Hàm mục tiêu $g(x, u) \geq 0$ với mọi $(x, u) \in D' \subset \mathbb{R}^{n+m}$. Vì vậy Argmin$(LP_*^{ct}) \neq \emptyset$;

ii) Dễ thấy, $(0, b)$ là phương án cực biên không suy biến tương ứng với cơ sở đơn vị $\{e^1, \dots, e^m\}$. Vì vậy có thể áp dụng ngay Thuật toán đơn hình để giải bài toán (LP_*^{ct}).

Định lý 3.9. *Giả sử (x^0, u^0) là phương án cực biên tối ưu của bài toán phụ (LP_*^{ct}). Khi đó,*
i) *Nếu $g(x^0, u^0) > 0$ thì $D = \emptyset$.*
ii) *Nếu $g(x^0, u^0) = 0$ thì x^0 là phương án cực biên của bài toán ban đầu (LP^{ct}).*

Chứng minh.

i) Giả sử phản chứng rằng giá trị tối ưu của bài toán phụ là $g(x^0, u^0) > 0$ nhưng $D \neq \emptyset$, tức

$$\exists\ x' \in D.$$

Theo định nghĩa, $(x', 0) \in D'$. Giá trị hàm mục tiêu của bài toán phụ (LP_*^{ct}) tại $(x', 0)$ là $g(x', 0) = 0$ nhỏ hơn giá trị tối ưu. Điều mâu thuẫn này chứng tỏ $D = \emptyset$.

ii) Do $u^0 \geq 0$ nên $g(x^0, u^0) = 0$ khi và chỉ khi $u^0 = 0$. Vậy trong trường hợp này, phương án cực biên tối ưu của bài toán phụ có dạng $(x^0, 0)$. Theo Mệnh đề 3.1, ta có ngay điều phải chứng minh. $\square$

Thuật toán đơn hình hai pha

Pha 1. (*Tìm phương án cực biên xuất phát cho thuật toán đơn hình giải* (LP^{ct}))
Giải bài toán quy hoạch tuyến tính phụ (LP^{ct}_*)

$$\min\,\{g(x,u) \mid Ax+u=b,\ (x,u)\geq 0\},$$

nhận được phương án cực biên tối ưu (x^0,u^0).

If $g(x^0,u^0)>0$ **Then** $D=\emptyset$, Dừng thuật toán
Else (x^0 *là phương án cực biên của bài toán* (LP^{ct})). Chuyển sang Pha 2.

Pha 2. Giải bài toán (LP^{ct}) đang xét bằng phương pháp đơn hình xuất phát từ phương án cực biên x^0 với chú ý rằng bảng đơn hình đầu tiên của Pha 2 là bảng đơn hình cuối cùng ở Pha 1 nhưng với một số sửa đổi như sau:

- ⋄ Xóa tất cả các cột tương ứng với các biến giả.
- ⋄ Thay cột C_B bởi hệ số mục tiêu cơ sở tương ứng của bài toán gốc.
- ⋄ Thay các hệ số mục tiêu của bài toán phụ ở dòng 1 bằng hệ số mục tiêu của bài toán gốc.
- ⋄ Nếu trong cơ sở tương ứng với phương án tối ưu của bài toán phụ (LP^{ct}_*) không có véc tơ giả thì cơ sở ứng với phương án này cũng chính là cơ sở tương ứng với x^0. Để có bảng đơn hình xuất phát cho bài toán ban đầu, ta chỉ cần tính lại giá trị hàm mục tiêu tại x^0 và tính lại các ước lượng theo công thức

$$f(x^0)=\sum_{j\in J(x^0)} c_j x^0_j \quad \text{và} \quad \Delta_k=\sum_{j\in J(x^0)} z_{jk}c_j - c_k.$$

  Xem Ví dụ 3.17.

- ⋄ Nếu trong cơ sở của phương án cực biên tối ưu của bài toán phụ (LP^{ct}_*) có ít nhất một véc tơ giả thì ta nhận được một phương án cực biên suy biến của bài toán ban đầu, tức $|J(x^0)|=\left|\left\{j\in\{1,\dots,n\} \mid x^0_j>0\right\}\right|<m$. Khi đó, để đẩy hết biến giả ra khỏi cơ sở, ta thực hiện thêm một vài bước lặp đơn hình nữa như sau: chọn cột quay là cột ứng với véc tơ phi cơ sở A_k với $k\leq n$ (tức A_k không phải véc tơ giả) mà nó có phần tử khác 0 ở dòng tương ứng với véc tơ giả và dòng này được chọn là dòng quay. Chú ý, phần tử chính lúc này có thể dương hoặc âm (khác 0). Do giả thiết $\text{rank}A=m$ nên nếu có biến giả trong cơ sở thì chắc chắn sẽ tìm được cột quay như thế. Sau khi bổ sung véc tơ để nhận được cơ sở $B=\{A_j,\ j\in J\}$ với $J\supset J(x^0)$, ta bỏ các cột tương ứng với biến giả rồi tiếp tục thuật toán (xem Ví dụ 3.18).

Chú ý 3.3.

i) Khi xây dựng bài toán phụ, ta chỉ cộng thêm biến giả vào những phương trình cần thiết để tạo ra ma trận m cột là các véc tơ đơn vị khác nhau. Hàm mục tiêu là tổng các biến giả.

ii) Khi một véc tơ giả bị loại ra khỏi cơ sở thì cột tương ứng với nó không cần tính ở các bước tiếp theo.

iii) Để tiện trong việc tính toán, ta đặt tên biến giả thứ i là x^g_{n+i} thay vì u_i và gọi véc tơ tương ứng với biến giả này là *véc tơ giả* A^g_{n+i}.

Ví dụ 3.9. Giải bài toán sau bằng thuật toán đơn hình hai pha:

$$\begin{aligned} \min \quad f(x) = {} & x_1 - 2x_2 + x_3 \\ \text{v.đ.k.} \quad & x_1 + x_2 + x_3 = 1 \\ & x_1 + 2x_2 + 3x_3 = 6 \\ & x_1,\ x_2,\ x_3 \geq 0. \end{aligned}$$

Giải. Quá trình tính toán như sau:

Pha 1. Giải bài toán quy hoạch tuyến tính phụ:

$$\begin{aligned} \min \quad f^p(x) \ = {} & \qquad\qquad\qquad\quad x^g_4 + x^g_5 \\ \text{v.đ.k.} \quad & x_1 + x_2 + x_3 + x^g_4 \qquad = \ 1 \\ & x_1 + 2x_2 + 3x_3 \qquad + x^g_5 \ = \ 6 \\ & x_1,\ x_2,\ x_3,\ x^g_4,\ x^g_5 \ \geq \ 0. \end{aligned}$$

Dễ thấy $x^0 = (0,\ 0,\ 0,\ 1,\ 6)^T$ là một phương án cực biên của bài toán phụ tương ứng với cơ sở đơn vị $\{A^g_4, A^g_5\} = \{e^1, e^2\}$. Xuất phát từ x^0, kết quả tính toán giải bài toán quy hoạch tuyến tính phụ theo thuật toán đơn hình được ghi ở Bảng 3.9.

Bài toán phụ có giá trị tối ưu là $f^p_{min} = 3 > 0$. Vì vậy, theo Định lý 3.9, bài toán quy hoạch tuyến tính ban đầu có tập chấp nhận được bằng rỗng.

Bảng 3.9

Hệ số C_B	Cơ sở B	Phương án	0 A_1	0 A_2	0 A_3	1 A_4^g	1 A_5^g	θ
1	A_4^g	1	1	1	[1]	1	0	1
1	A_5^g	6	1	2	3	0	1	2
	Bảng 1	7	2	3	4	0	0	
0	A_3	1	1	1	1	1	0	
1	A_5^g	3	-2	-1	0	-3	1	
	Bảng 2	3	-2	-1	0	-4	0	

Ví dụ 3.10. Giải bài toán sau bằng thuật toán đơn hình hai pha:

$$\begin{aligned} \min\ \ & f(x) = 2x_1 + 3x_2 \\ \text{v.đ.k.}\ \ & 2x_1 - 4x_2 \geq 2 \\ & 4x_1 + 3x_2 \leq 19 \\ & 3x_1 + 2x_2 = 14 \\ & x_1, x_2 \geq 0. \end{aligned}$$

Giải. Trước hết, ta đưa bài toán này về dạng chính tắc sau:

$$\begin{aligned} \min\ \ & f(x) = 2x_1 + 3x_2 \\ \text{v.đ.k.}\ \ & 2x_1 - 4x_2 - x_3 = 2 \\ & 4x_1 + 3x_2 + x_4 = 19 \\ & 3x_1 + 2x_2 = 14 \\ & x_1, x_2, x_3, x_4 \geq 0. \end{aligned}$$

Pha 1. Để ý rằng, ràng buộc chính thứ hai của bài toán dạng chính tắc có biến x_4 với hệ số bằng 1 và biến x_4 không xuất hiện trong hai ràng buộc chính còn lại nên ta chỉ cần thêm hai biến giả x_5^g và x_6^g trong bài toán phụ. Sau đây là bài toán phụ tương ứng với bài toán đang xét

$$\begin{aligned} \min\ \ & f^p(x) = x_5^g + x_6^g \\ \text{v.đ.k.}\ \ & 2x_1 - 4x_2 - x_3 + x_5^g = 2 \\ & 4x_1 + 3x_2 + x_4 = 19 \\ & 3x_1 + 2x_2 + x_6^g = 14 \\ & x_1,\ x_2,\ x_3,\ x_4,\ x_5^g,\ x_6^g \geq 0. \end{aligned}$$

Bảng 3.10

Hệ số C_B	Cơ sở B	Phương án	0 A_1	0 A_2	0 A_3	0 A_4	1 A_5^g	1 A_6^g	θ
1	A_5^g	2	[2]	-4	-1	0	1	0	1
0	A_4	19	4	3	0	1	0	0	$\frac{19}{4}$
1	A_6^g	14	3	2	0	0	0	1	$\frac{14}{3}$
	Bảng 1	16	5	-2	-1	0	0	0	
0	A_1	1	1	-2	$-\frac{1}{2}$	0		0	∞
0	A_4	15	0	[11]	2	1		0	$\frac{15}{11}$
1	A_6^g	11	0	8	$\frac{3}{2}$	0		1	$\frac{11}{8}$
	Bảng 2	11	0	8	$\frac{3}{2}$	0		0	
0	A_1	$\frac{41}{11}$	1	0	$-\frac{3}{22}$	$\frac{2}{11}$		0	∞
0	A_2	$\frac{15}{11}$	0	1	$\frac{2}{11}$	$\frac{1}{11}$		0	$\frac{15}{2}$
1	A_6^g	$\frac{1}{11}$	0	0	$[\frac{1}{22}]$	$-\frac{8}{11}$		1	2
	Bảng 3	$\frac{1}{11}$	0	0	$\frac{1}{22}$	$-\frac{8}{11}$		0	
0	A_1	4	1	0	0	-2			
0	A_2	1	0	1	0	3			
0	A_3	2	0	0	1	-16			
	Bảng 4	0	0	0	0	0			

Quá trình tính toán giải bài toán phụ theo Pha 1, xuất phát từ phương án cực biên $x^0 = (0,\ 0,\ 0,\ 19,\ 2,\ 14)^T$ tương ứng với cơ sở $\{A_5^g, A_4, A_6^g\}$, được ghi ở Bảng 3.10.

Bài toán phụ có giá trị tối ưu $f_{min}^p = 0$. Không có các biến giả trong cơ sở của phương án cực biên tối ưu $x^0 = (4,\ 1,\ 2,\ 0)^T$. Ta chuyển sang Pha 2.

Pha 2. Xuất phát từ phương án cực biên không suy biến x^0 tương ứng với cơ sở đơn vị $\{A_1, A_2, A_3\}$, quá trình tính được trình bày ở Bảng 3.11.

Chú ý rằng bảng xuất phát ở Pha 2 (bảng đơn hình đầu tiên của Bảng 3.11) chỉ khác bảng cuối cùng của Pha 1 (bảng đơn hình cuối cùng của Bảng 3.10) ở cột hệ số hàm mục tiêu, dòng 1 (được thay bằng hệ số hàm mục tiêu của bài toán ban đầu) và dòng ước lượng (dòng cuối cùng).

Bảng đơn hình cuối cùng của Bảng 3.11 tương ứng với phương án cực biên tối ưu duy nhất $x^* = (\frac{14}{3},\ 0,\ \frac{22}{3},\ \frac{1}{3})^T$ của bài toán ban đầu vì ước lượng tương ứng với biến phi cơ sở duy nhất $\Delta_2 = -\frac{5}{3} < 0$.

Bảng 3.11

Hệ số C_B	Cơ sở B	Phương án	2 A_1	3 A_2	0 A_3	0 A_4	θ
2	A_1	4	1	0	0	-2	∞
3	A_2	1	0	1	0	[3]	$\frac{1}{3}$
0	A_3	2	0	0	1	-16	∞
	Bảng 1	11	0	0	0	5	
2	A_1	$\frac{14}{3}$	1	$\frac{2}{3}$	0	0	
0	A_4	$\frac{1}{3}$	0	$\frac{1}{3}$	0	1	
0	A_3	$\frac{22}{3}$	0	$\frac{16}{3}$	1	0	
	Bảng 2	$\frac{28}{3}$	0	$-\frac{5}{3}$	0	0	

Ví dụ 3.11. *(Bảng cuối của Pha 1 còn véc tơ giả trong cơ sở)* Giải bài toán quy hoạch tuyến tính sau bằng thuật toán đơn hình hai pha:

$$\begin{aligned}
\min\ \ & f(x) = 2x_1 + 4x_2 + \frac{1}{2}x_3 - 3x_4 \\
\text{v.đ.k.}\ \ & 2x_1 + 2x_2 + 3x_3 + 3x_4 + x_5 = 50 \\
& 4x_1 + 8x_2 + 2x_3 + 3x_4 = 80 \\
& 4x_1 + 4x_2 + x_3 + 2x_4 = 40 \\
& x_1,\ x_2,\ x_3,\ x_4,\ x_5 \geq 0.
\end{aligned}$$

Giải.
Pha 1. Bài toán phụ tương ứng với bài toán này là:

$$\begin{aligned}
\min\ \ & f^p(x) = x_6^g + x_7^g \\
\text{v.đ.k.}\ \ & 2x_1 + 2x_2 + 3x_3 + 3x_4 + x_5 = 50 \\
& 4x_1 + 8x_2 + 2x_3 + 3x_4 + x_6^g = 80 \\
& 4x_1 + 4x_2 + x_3 + 2x_4 + x_7^g = 40 \\
& x_1,\ x_2,\ x_3,\ x_4,\ x_5,\ x_6^g,\ x_7^g \geq 0.
\end{aligned}$$

Quá trình tính toán giải bài toán phụ xuất phát từ phương án cực biên

$$(0,\ 0,\ 0,\ 0, 50,\ 80,\ 40)^T$$

được trình bày ở Bảng 3.12. Sau một bước biến đổi đơn hình, ta nhận được phương án tối ưu $(0,\ 10,\ 0,\ 0, 30,\ 0,\ 0)^T$ nhưng trong cơ sở của phương án này còn véc tơ giả A_6^g (bảng thứ hai). Để loại véc tơ giả A_6^g ra khỏi cơ sở, ta đưa véc tơ phi cơ sở A_1 vào cơ sở (vì $z_{61} = -4 \neq 0$) và thực hiện một bước biến đổi đơn hình chuyển sang bảng cuối cùng của Bảng 3.12. Bảng cuối cùng này không còn véc tơ giả trong cơ sở của phương án cực biên tương ứng vì vậy ta chuyển sang Pha 2.

Pha 2. Trong bảng cuối cùng của Pha 1, ta loại bỏ cột ứng với véc tơ giả A_6^g. Thay cột hệ số C_B, thay dòng đầu tiên bởi các hệ số mục tiêu của bài toán gốc, tính lại các ước lượng, ta nhận được bảng xuất phát của Pha 2. Sau hai bước biến đổi đơn hình, ta nhận được phương án cực biên suy biến tối ưu $x^* = (0,\ 7,\ 12,\ 0,\ 0)^T$ và giá trị tối ưu $f_{min} = 34$ (xem bảng cuối cùng của Bảng 3.13).

Ta để ý rằng, bảng đơn hình thứ nhất của Bảng 3.13 tương ứng với phương án cực biên suy biến $(0,\ 10,\ 0,\ 0, 30)^T$ có cơ sở là $\{A_5, A_1, A_2\}$ và giá trị hàm mục tiêu tại đây bằng 40. Ta chọn véc tơ A_1 đưa ra cơ sở tương ứng với $\theta_1 = \min\{15, 0, 40\} = 0$ nên bảng đơn hình thứ hai vẫn tương ứng với phương án này nhưng với cơ sở thứ hai của nó là $\{A_5, A_4, A_2\}$, còn giá trị hàm mục tiêu hiển nhiên là không đổi.

Bảng 3.12

Hệ số C_B	Cơ sở B	Phương án	0 A_1	0 A_2	0 A_3	0 A_4	0 A_5	1 A_6^g	1 A_7^g	θ
0	A_5	50	2	2	3	3	1	0	0	25
1	A_6^g	80	4	8	2	3	0	1	0	10
1	A_7^g	40	4	[4]	1	2	0	0	1	10
	Bảng 1	120	8	12	3	5	0	0	0	
0	A_5	30	0	0	$\frac{5}{2}$	2	1	0		
1	A_6^g	0	$[-4]$	0	0	- 1	0	1		
0	A_2	10	1	1	$\frac{1}{4}$	$\frac{1}{2}$	0	0		
	Bảng 2	0	- 4	0	0	-1	0	0		
0	A_5	30	0	0	$\frac{5}{2}$	2	1			
0	A_1	0	1	0	0	$\frac{1}{4}$	0			
0	A_2	10	0	1	$\frac{1}{4}$	$\frac{1}{4}$	0			
	Bảng 3	0	0	0	0	0	0			

Bảng 3.13

Hệ số C_B	Cơ sở B	Phương án	2 A_1	4 A_2	1/2 A_3	- 3 A_4	0 A_5	θ
0	A_5	30	0	0	$\frac{5}{2}$	2	1	15
2	A_1	0	1	0	0	$[\frac{1}{4}]$	0	0
4	A_2	10	0	1	$\frac{1}{4}$	$\frac{1}{4}$	0	40
	Bảng 1	40	0	0	$\frac{1}{2}$	$\frac{9}{2}$	0	
0	A_5	30	-8	0	$[\frac{5}{2}]$	0	1	12
- 3	A_4	0	4	0	0	1	0	∞
4	A_2	10	-1	1	$\frac{1}{4}$	0	0	40
	Bảng 2	40	-18	0	$\frac{1}{2}$	0	0	
1/2	A_3	12	$-\frac{16}{5}$	0	1	0	$\frac{2}{5}$	
-3	A_4	0	4	0	0	1	0	
4	A_2	7	$-\frac{1}{5}$	1	0	0	$-\frac{1}{10}$	
	Bảng 3	34	$-\frac{82}{5}$	0	0	0	$-\frac{1}{5}$	

3.5.3 Phương pháp đánh thuế hay phương pháp bài toán (M)

Trong thuật toán đơn hình hai pha, ta phải thực hiện thuật toán đơn hình hai lần: một lần để giải bài toán quy hoạch tuyến tính phụ (LP_*^{ct}) và một lần để giải bài toán quy hoạch tuyến tính gốc (nếu tập nghiệm chấp nhận được của nó khác rỗng). Phương pháp đánh thuế cho phép kết hợp cả hai pha này lại.

Xét bài toán quy hoạch tuyến tính dạng chính tắc:

$$\min\{\langle c, x\rangle \mid Ax = b,\ x \geq 0\}, \qquad (LP^{ct})$$

với ma trận A cấp $m \times n$, véc tơ $b \in \mathbb{R}^m$ và $b \geq 0$.

Để giải bài toán (LP^{ct}), ta xét bài toán mở rộng sau đây (thường gọi là *bài toán (M)*):

$$\begin{array}{ll} \min\ F(x,u) = c_1x_1 + \cdots + c_nx_n + M(u_1 + \cdots + u_m) & (M) \\ \text{v.đ.k.} \qquad Ax + u = b & \\ \qquad\qquad (x,u) \geq 0, & \end{array}$$

trong đó M là một hằng số dương đủ lớn (lớn hơn bất cứ số cụ thể nào mà ta cần so sánh với nó). Về trực quan, có thể thấy rằng cách làm này giống như "đánh thuế" rất nặng vào các biến giả $u_1, \ldots, u_m$ khiến cho nếu bài toán chính (LP^{ct}) có phương án chấp nhận được thì phương án tối ưu (x^*, u^*) của bài toán (M) phải có $u^* = 0$ và x^* là phương án tối ưu của bài toán chính.

Bài toán (M) có m ràng buộc chính và $n + m$ biến. Dễ thấy $(0, b)$ là một phương án cực biên của bài toán này với cơ sở tương ứng là cơ sở đơn vị. Do đó ta có thể xây dựng được ngay bảng đơn hình xuất phát và giải bài toán này bằng Thuật toán 3.3. Tuy nhiên, để ý rằng, do hàm mục tiêu F phụ thuộc tuyến tính vào M nên các ước lượng Δ_k của các biến x_k cũng phụ thuộc tuyến tính vào M. Do đó, ta có thể viết

$$F = \alpha_0 + \beta_0 M \ \text{ và } \ \Delta_k = \alpha_k + \beta_k M,\ k = 1, \ldots, n + m.$$

Khi giải bài toán (M) theo Thuật toán 3.3, ta không cần xác định giá trị cụ thể của số M mà chỉ cần tách dòng ước lượng (dòng cuối cùng của mỗi bảng đơn hình) thành hai dòng: dòng đầu ghi các hệ số α_k, dòng sau ghi các hệ số β_k, $k = 0, 1, \ldots, n + m$. Vì M là hằng số dương đủ lớn nên:

⋄ Với mỗi $k \in \{1, \ldots, n\}$ ta có:

$\Delta_k > 0$ nếu $\beta_k > 0$ và α_k bất kỳ (*hoặc* $\beta_k = 0$ *và* $\alpha_k > 0$);

⋄ Với cặp $i, j \in \{1, \ldots, n\}$, $i \neq j$, ta có:

$\Delta_i > \Delta_j$ nếu $\beta_i > \beta_j$ và α_i, α_j bất kỳ (*hoặc* $\beta_i = \beta_j$ *và* $\alpha_i > \alpha_j$).

Khi thuật toán kết thúc:

i) Nếu bài toán (M) có phương án tối ưu là (x^*, u^*) với $u^* \neq 0$ thì bài toán gốc (LP^{ct}) không có phương án chấp nhận được;

ii) Nếu bài toán (LP^M) có phương án tối ưu là $(x^*, 0)$ thì x^* là phương án tối ưu của bài toán gốc (LP^{ct});

iii) Nếu bài toán (LP^M) không có phương án tối ưu thì bài toán gốc (LP^{ct}) cũng không có phương án tối ưu (giá trị hàm mục tiêu giảm vô hạn).

Ví dụ 3.12. Xét lại bài toán quy hoạch tuyến tính chính tắc như ở Ví dụ 3.17:

$$\begin{aligned}
\min\ \ & f(x) = 2x_1 + 3x_2 \\
\text{v.đ.k.}\ \ & 2x_1 - 4x_2 - x_3 = 2 \\
& 4x_1 + 3x_2 + x_4 = 19 \\
& 3x_1 + 2x_2 = 14 \\
& x_1, x_2, x_3, x_4 \geq 0.
\end{aligned}$$

Cũng lập luận tương tự như khi giải Ví dụ 3.17, ta chỉ cần thêm hai biến giả x_5^g và x_6^g trong bài toán mở rộng. Sau đây là bài toán (M) tương ứng:

$$\begin{aligned}
\min\ \ F(x, x^g) &= 2x_1 + 3x_2 + (x_5^g + x_6^g)M \\
\text{v.đ.k.}\ \ & 2x_1 - 4x_2 - x_3 + x_5^g = 2 \\
& 4x_1 + 3x_2 + x_4 = 19 \\
& 3x_1 + 2x_2 + x_6^g = 14 \\
& x_1,\ x_2,\ x_3,\ x_4,\ x_5^g,\ x_6^g \geq 0.
\end{aligned}$$

Quá trình tính toán giải bài toán này theo Thuật toán 3.3, xuất phát từ phương án cực biên ban đầu $x^0 = (0,\ 0,\ 0,\ 19,\ 2,\ 14)^T$ tương ứng với cơ sở đơn vị $\{A_5^g, A_4, A_6^g\}$ được trình bày ở Bảng 3.14. Như sẽ thấy, bảng đơn hình cuối cùng của Bảng 3.14 cho ta phương án cực biên tối ưu $(\frac{14}{3},\ 0,\ \frac{22}{3},\ \frac{1}{3},\ 0,\ 0)^T$. Do đó, phương án cực biên tối ưu của bài toán ban đầu là $x^* = (\frac{14}{3},\ 0,\ \frac{22}{3},\ \frac{1}{3})^T$. Ta có kết quả trùng hợp với kết quả nhận được khi giải bài toán này bằng thuật toán đơn hình hai pha ở Ví dụ 3.17.

Bảng 3.14

Hệ số C_B	Cơ sở B	Phương án	2 A_1	3 A_2	0 A_3	0 A_4	M A_5	M A_6	θ
M	A_5^g	2	[2]	-4	-1	0	1	0	1
0	A_4	19	4	3	0	1	0	0	$\frac{19}{4}$
M	A_6^g	14	3	2	0	0	0	1	$\frac{14}{3}$
	α	0	-2	-3	0	0	0	0	
Bảng 1	β	16	5	-2	-1	0	0	0	

Hệ số C_B	Cơ sở B	Phương án	2 A_1	3 A_2	0 A_3	0 A_4	M A_5	M A_6	θ
2	A_1	1	1	-2	$-\frac{1}{2}$	0	$\frac{1}{2}$	0	∞
0	A_4	15	0	[11]	2	1	-2	0	$\frac{15}{11}$
M	A_6^g	11	0	8	$\frac{3}{2}$	0	$-\frac{3}{2}$	1	$\frac{11}{8}$
	α	2	0	-7	-1	0	1	0	
Bảng 2	β	11	0	8	$\frac{3}{2}$	0	$-\frac{5}{2}$	0	
2	A_1	$\frac{41}{11}$	1	0	$-\frac{3}{22}$	$\frac{2}{11}$	$\frac{3}{22}$	0	∞
3	A_2	$\frac{15}{11}$	0	1	$\frac{2}{11}$	$\frac{1}{11}$	$-\frac{2}{11}$	0	$\frac{15}{2}$
M	A_6^g	$\frac{1}{11}$	0	0	$[\frac{1}{22}]$	$-\frac{8}{11}$	$-\frac{1}{22}$	1	2
	α	$\frac{127}{11}$	0	0	$\frac{3}{11}$	$\frac{7}{11}$	$-\frac{3}{11}$	0	
Bảng 3	β	$\frac{1}{11}$	0	0	$\frac{1}{22}$	$-\frac{8}{11}$	$-\frac{23}{22}$	0	
2	A_1	4	1	0	0	- 2	0	3	∞
3	A_2	1	0	1	0	[3]	0	- 4	$\frac{1}{3}$
0	A_3	2	0	0	1	- 16	- 1	22	∞
	α	11	0	0	0	5	0	-6	
Bảng 4	β	0	0	0	0	0	- 1	-1	
2	A_1	$\frac{14}{3}$	1	$\frac{2}{3}$	0	0	0	$\frac{1}{3}$	
0	A_4	$\frac{1}{3}$	0	$\frac{1}{3}$	0	1	0	$-\frac{4}{3}$	
0	A_3	$\frac{22}{3}$	0	$\frac{16}{3}$	1	0	- 1	$\frac{2}{3}$	
	α	$\frac{28}{3}$	0	$-\frac{5}{3}$	0	0	0	$\frac{2}{3}$	
Bảng 5	β	0	0	0	0	0	- 1	-1	

3.6 Tính hữu hạn của thuật toán đơn hình

Định nghĩa. Thuật toán giải một bài toán tối ưu được gọi là *hữu hạn* nếu như nó cho phép tìm được phương án tối ưu của bài toán sau một số hữu hạn phép tính.

Định lý 3.10. *Giả sử một bài toán quy hoạch tuyến tính không thoái hóa và có phương án tối ưu. Khi đó, với mọi phương án cực biên xuất phát, thuật toán đơn hình là hữu hạn.*

Chứng minh. Giả sử bài toán quy hoạch tuyến tính chính tắc (LP^{ct}) có phương án cực biên x^0 tương ứng với cơ sở B_0. Xuất phát từ x^0, quá trình thực hiện thuật toán đơn hình để giải bài toán này sẽ sinh ra một dãy các phương án cực biên x^k và dãy cơ sở tương ứng B_k, $k = 1, 2, \dots$ Do bài toán không suy biến nên tất cả các phương án cực biên đều không suy biến và khi chuyển từ x^k sang x^{k+1} thì giá trị hàm mục tiêu giảm thực sự (xem công thức (3.27)). Ta có:

$$f(x^k) \overset{(3.12')}{=} C_{B_k} X_{B_k} \overset{(3.11')}{=} C_{B_k} B_k^{-1} b.$$

Do đó trong quá trình thực hiện thuật toán, không có một ma trận cơ sở nào bị lặp lại. Mặt khác, số cơ sở của ma trận A không vượt quá C_n^m nên sau một số hữu hạn bước lặp, ta sẽ nhận được một phương án cực biên thỏa mãn điều kiện dừng của thuật toán. □

Nhận xét 3.4. Trong trường hợp một bài toán thoái hóa thì tính hữu hạn của thuật toán đơn hình không đảm bảo nữa. Người ta gọi đó là *hiện tượng xoay vòng.*

3.7 Hiện tượng xoay vòng

Hiện tượng xoay vòng là hiện tượng sau một số phép biến đổi đơn hình, ta lại quay lại phương án cực biên suy biến trước đó với cơ sở tương ứng đã gặp. Trong trường hợp này, thuật toán đơn hình không thể kết thúc. Ví dụ đầu tiên về hiện tượng xoay vòng do Hoffman [12] đưa ra năm 1953. Năm 1969, Marshall và Suurballe [30] chứng minh rằng một bài toán quy hoạch tuyến tính dạng chính tắc xoay vòng thì phải có ít nhất sáu biến và ba ràng buộc. Sau đây là ví dụ bài toán quy hoạch tuyến tính xoay vòng do GS. Trần Vũ Thiệu [39] (Viện

Toán học) đưa ra năm 1964:

$$
\begin{aligned}
\min\ f(x) = 4x_1 & \qquad\qquad\qquad\quad - 6x_5 - 5x_6 + 64x_7 \\
\text{v.đ.k.} \quad x_1 & \qquad\quad + \frac{1}{3}x_4 - 2x_5 - \ \ x_6 + 12x_7 = 0 \\
& x_2 + \quad + \frac{1}{2}x_4 - \ \ x_5 - \frac{1}{6}x_6 + \frac{2}{3}x_7 = 0 \\
& \quad x_3 \qquad\quad + \ \ x_5 + \ \ x_6 - \ \ 9x_7 = 2 \\
& x_1,\ x_2,\ x_3, x_4,\ x_5,\ x_6,\ x_7 \geq 0.
\end{aligned}
$$

Giải bài toán này theo thuật toán đơn hình xuất phát từ phương án cực biên suy biến $(0,\ 0,\ 2,\ 0,\ 0,\ 0, 0)^T$ với quy tắc là: tại một bảng đơn hình nào đó, nếu có nhiều véc tơ cùng đạt tiêu chuẩn ra khỏi cơ sở đang xét thì ta chọn để đưa ra khỏi cơ sở véc tơ có chỉ số nhỏ nhất. Kết quả là bảng đơn hình ở bước lặp thứ bảy trùng với bảng đơn hình ở bước đầu tiên. Như vậy, nếu cứ tiếp tục tính ta sẽ lâm vào tình trạng xoay vòng.

Sau hơn nửa thế kỷ áp dụng phương pháp đơn hình giải quy hoạch tuyến tính, người ta thấy rằng hiện tượng xoay vòng rất ít xảy ra đến mức các chương trình phần mềm giải quy hoạch tuyến tính đều không chú ý đến phòng ngừa xoay vòng. Tuy nhiên, về mặt lý thuyết, người ta vẫn phải đưa ra các biện pháp để tránh hiện tượng xoay vòng như: quy tắc R.G. Bland, phương pháp từ điển, phương pháp nhiễu loạn. Bạn đọc quan tâm đến các phương pháp này có thể tham khảo trong [2], [8], [19], [34], [40] hoặc [42].

3.8 Đối ngẫu

3.8.1 Cặp bài toán quy hoạch tuyến tính đối ngẫu

Cho A là ma trận cấp $m \times n$ và véc tơ $b = (b_1, \dots, b_m)^T \in \mathbb{R}^m$. Ký hiệu
a^i, $i = 1, \dots, m$, là các hàng của ma trận A;
A_j, $j = 1, \dots, n$, là các cột của ma trận A.

- *Đối ngẫu của bài toán quy hoạch tuyến tính tổng quát*

Bảng 3.15

Gốc (P)	min $\langle c, x\rangle$	max $\langle b, y\rangle$	Đối ngẫu (D)
Ràng buộc chính	$\langle a^i, x\rangle = b_i, \quad i \in M_1$ $\langle a^i, x\rangle \le b_i, \quad i \in M_2$ $\langle a^i, x\rangle \ge b_i, \quad i \in M_3$	$y_i \in \mathbb{R}, \quad i \in M_1$ $y_i \le 0, \quad i \in M_2$ $y_i \ge 0, \quad i \in M_3$	Ràng buộc dấu
Ràng buộc dấu	$x_j \ge 0, \quad j \in N_1$ $x_j \le 0, \quad j \in N_2$ $x_j \in \mathbb{R}, \quad j \in N_3$	$\langle A_j, y\rangle \le c_j, \quad j \in N_1$ $\langle A_j, y\rangle \ge c_j, \quad j \in N_2$ $\langle A_j, y\rangle = c_j, \quad j \in N_3$	Ràng buộc chính

$(|M_1| + |M_2| + |M_3| = m, \qquad |N_1| + |N_2| + |N_3| = n)$

Nhận xét 3.5.

- Số ràng buộc chính của bài toán quy hoạch tuyến tính gốc (P) bằng số biến (hay số ràng buộc dấu) của bài toán quy hoạch tuyến tính đối ngẫu (D).
- Số biến của bài toán (P) bằng số ràng buộc chính của bài toán (D).
- Các biến của bài toán (D) được gọi là các *biến đối ngẫu*. Các biến đối ngẫu tương ứng 1-1 với các ràng buộc chính của bài toán gốc (P).
- Các ràng buộc chính của bài toán (D) tương ứng 1-1 với các biến của bài toán (P). Tính chất này được sử dụng rất nhiều khi nghiên cứu các bài toán quy hoạch tuyến tính thông qua các bài toán đối ngẫu của chúng.

- *Đối ngẫu của bài toán quy hoạch tuyến tính chuẩn tắc*

$$\begin{aligned} \min z &= \langle c, x\rangle \\ \text{v.đ.k.}\quad Ax &\geq b \\ x &\geq 0 \end{aligned} \qquad\qquad \begin{aligned} \max w &= \langle b, y\rangle \\ \text{v.đ.k.}\quad A^T y &\leq c \\ y &\geq 0. \end{aligned}$$

- *Đối ngẫu của bài toán quy hoạch tuyến tính chính tắc*

$$\begin{aligned} \min z &= \langle c, x\rangle \\ \text{v.đ.k.}\quad Ax &= b \\ x &\geq 0 \end{aligned} \qquad\qquad \begin{aligned} \max w &= \langle b, y\rangle \\ \text{v.đ.k.}\quad A^T y &\leq c \\ y &\in \mathbb{R}. \end{aligned}$$

Mệnh đề 3.2. *Bài toán đối ngẫu của bài toán đối ngẫu là bài toán gốc.*

Chứng minh. Không giảm tổng quát, ta xét bài toán gốc (P) là bài toán quy hoạch tuyến tính dạng chuẩn tắc. Ta có cặp bài toán đối ngẫu sau:

$$\begin{aligned} \min z &= \langle c, x\rangle \qquad (P) \\ \text{v.đ.k.}\quad Ax &\geq b \\ x &\geq 0 \end{aligned} \qquad\qquad \begin{aligned} \max w &= \langle b, y\rangle \qquad (D) \\ \text{v.đ.k.}\quad A^T y &\leq c \\ y &\geq 0. \end{aligned}$$

Bài toán (D) tương đương với bài toán:

$$\begin{aligned} \min w' &= -\langle b, y\rangle \qquad (D') \\ \text{v.đ.k.}\quad -A^T y &\geq -c \\ y &\geq 0. \end{aligned}$$

Bài toán đối ngẫu của (D') là:

$$\begin{aligned} \max z' &= -\langle c, x\rangle \qquad (P') \\ \text{v.đ.k.}\quad -Ax &\leq -b \\ x &\geq 0 \end{aligned} \qquad \Leftrightarrow \qquad \begin{aligned} \min z &= \langle c, x\rangle \qquad (P) \\ \text{v.đ.k.}\quad Ax &\geq b \\ x &\geq 0. \end{aligned}$$

Mệnh đề đã được chứng minh. $\square$

3.8.2 Các định lý về đối ngẫu

Không giảm tổng quát, xét cặp bài toán đối ngẫu dạng chuẩn tắc

$$\begin{array}{ll} \min z(x) = \langle c, x\rangle \quad (P) \\ \text{v.đ.k.} \;\; Ax \geq b \\ \qquad\quad x \geq 0 \end{array} \qquad\qquad \begin{array}{ll} \max w(y) = \langle b, y\rangle \quad (D) \\ \text{v.đ.k.} \;\; A^T y \leq c \\ \qquad\quad y \geq 0. \end{array}$$

Ký hiệu:

X là tập chấp nhận được của bài toán gốc P;

Y là tập chấp nhận được của bài toán đối ngẫu (D).

Định lý 3.11. (Đối ngẫu yếu) *Nếu x^* là phương án chấp nhận được bất kỳ của bài toán gốc (P) và y^* là phương án chấp nhận được bất kỳ của bài toán đối ngẫu (D) thì*

$$\langle b, y^*\rangle \leq \langle c, x^*\rangle.$$

Chứng minh. Lấy tùy ý $x^* \in X$ và $y^* \in Y$. Theo định nghĩa,

$$Ax^* \geq b, \quad x^* \geq 0$$

và

$$A^T y^* \leq c, \quad y^* \geq 0.$$

Ta có:

$$\langle b, y^*\rangle \leq \langle Ax^*, y\rangle = \langle x^*, A^T y^*\rangle \leq \langle x^*, c\rangle = \langle c, x\rangle.$$

□

Hệ quả 3.5.

i) *Nếu hàm $z(x) = \langle c, x\rangle$ không bị chặn dưới trên X thì bài toán (D) không chấp nhận được, tức $Y = \emptyset$;*

ii) *Nếu hàm $\max w(y) = \langle b, y\rangle$ bị chặn trên trên Y thì bài toán (P) không chấp nhận được, tức $X = \emptyset$.*

Chứng minh. i) Giả sử phản chứng là tồn tại $y^0 \in Y$. Theo Định lý 3.11 ta có

$$\langle b, y^0\rangle \leq \langle c, x\rangle \;\; \forall \;\; x \in X,$$

tức $\langle c, x\rangle$ bị chặn dưới trên $X \Rightarrow$ Mâu thuẫn với giả thiết $\Rightarrow \;\; Y = \emptyset$. □

Hệ quả 3.6. *Giả sử* $\bar{x} \in X$, $\bar{y} \in Y$ *và*

$$\langle b, \bar{y} \rangle = \langle c, \bar{x} \rangle.$$

Khi đó

$$\bar{x} \in \text{Argmin}(P) \quad và \bar{y} \text{Argmax}(D).$$

Chứng minh. Lấy tùy ý $x \in X$ và $y \in Y$. Theo Định lý 3.11 ta có

$$\langle b, y \rangle \leq \langle c, \bar{x} \rangle = \langle b, \bar{y} \rangle \leq \langle c, x \rangle.$$

Ta có điều phải chứng minh. □

Định lý 3.12. (Đối ngẫu mạnh) *Nếu một bài toán quy hoạch tuyến tính có phương án tối ưu thì bài toán đối ngẫu của nó cũng có phương án tối ưu và giá trị tối ưu của hai bài toán này bằng nhau.*

Hệ quả 3.7. *Điều kiện cần và đủ để cặp phương án chấp nhận được* $\{\bar{x}, \bar{y}\}$ *của bài toán (P) và (D) là phương án tối ưu là*

$$\langle c, \bar{x} \rangle = \langle b, \bar{y} \rangle.$$

Định lý 3.13. (Định lý tồn tại) *Xét một cặp quy hoạch tuyến tính đối ngẫu (P) và (D). Khi đó chỉ xảy ra một trong bốn trường hợp sau:*

i) $X = \emptyset$ *và* $Y = \emptyset$.

ii) $X \neq \emptyset$ *và* $Y \neq \emptyset$. *Khi đó*

$$\text{Argmin}(P) \neq \emptyset, \quad \text{Argmax}(D) \neq \emptyset$$

và

$$\min\{\langle c, x \rangle \mid x \in X\} = \max\{\langle b, y \rangle \mid y \in Y\}.$$

iii) $X \neq \emptyset$ *và* $Y = \emptyset$. *Khi đó:*

$f(x) = \langle c, x \rangle$ *không bị chặn dưới trên* X.

iv) $X = \emptyset$ *và* $Y \neq \emptyset$. *Khi đó*

$g(y) = \langle b, y \rangle$ *không bị chặn trên trên* Y.

Sau đây là một số ví dụ minh họa cho bốn trường hợp đã nêu trong Định lý 3.13.

Ví dụ 3.13. *(Bài toán gốc và bài toán đối ngẫu đều không có phương án chấp nhận được)*

$$\begin{array}{ll} \min\ f(x) = 2x_1 + x_2 & \\ \quad \text{v.đ.k.} \quad x_1 + x_2 = 5 & \\ \qquad\qquad -x_1 - x_2 = 3 & \\ \qquad\qquad x_1, x_2 \in \mathbb{R} & \end{array} \qquad \begin{array}{ll} \max\ g(y) = 5y_1 + 3y_2 & \\ \quad \text{v.đ.k.} \quad y_1 - y_2 = 2 & \\ \qquad\qquad y_1 - y_2 = 1 & \\ \qquad\qquad y_1, y_2 \in \mathbb{R}. & \end{array}$$

Ví dụ 3.14. *(Bài toán gốc và bài toán đối ngẫu đều có phương án chấp nhận được)*

$$\begin{array}{ll} \min\ f(x) = 6x_1 + 5x_2 & \\ \quad \text{v.đ.k.} \quad x_1 + 2x_2 \geq 4 & \\ \qquad\qquad 2x_1 + x_2 \leq 10 & \\ \qquad\qquad x_1 \geq 0, x_2 \geq 0 & \end{array} \qquad \begin{array}{ll} \max\ g(y) = 4y_1 + 10y_2 & \\ \quad \text{v.đ.k.} \quad y_1 + 2y_2 \leq 6 & \\ \qquad\qquad 2y_1 + y_2 \leq 5 & \\ \qquad\qquad y_1 \geq 0, y_2 \leq 0. & \end{array}$$

Ví dụ 3.15. *(Bài toán gốc không có phương án chấp nhận được và giá trị hàm mục tiêu của bài toán đối ngẫu tăng vô hạn trong miền ràng buộc)*

$$\begin{array}{ll} \min\ f(x) = 2x_1 + x_2 & \\ \quad \text{v.đ.k.} \quad x_1 - x_2 = 5 & \\ \qquad\qquad x_1 + x_2 = 2 & \\ \qquad\qquad x_1, x_2 \geq 0 & \end{array} \qquad \begin{array}{ll} \max\ g(y) = 5y_1 + 2y_2 & \\ \quad \text{v.đ.k.} \quad y_1 + y_2 \leq 2 & \\ \qquad\qquad -y_1 + y_2 \leq 1 & \\ \qquad\qquad y_1, y_2 \in \mathbb{R}. & \end{array}$$

Ví dụ 3.16. *(Giá trị hàm mục tiêu của bài toán gốc giảm vô hạn trong miền ràng buộc và bài toán đối ngẫu không có phương án chấp nhận được)*

$$\begin{array}{ll} \min\ f(x) = -x_1 - x_2 & \\ \quad \text{v.đ.k.} \quad -2x_1 + x_2 \leq 2 & \\ \qquad\qquad x_1 - 3x_2 \leq 3 & \\ \qquad\qquad x_1, x_2 \geq 0 & \end{array} \qquad \begin{array}{ll} \max\ g(y) = 2y_1 + 3y_2 & \\ \quad \text{v.đ.k.} \quad -2y_1 + y_2 \leq -1 & \\ \qquad\qquad y_1 - 3y_2 \leq -1 & \\ \qquad\qquad y_1, y_2 \leq 0\ . & \end{array}$$

3.8.3 Định lý về độ lệch bù (complementary slackness)

Xét cặp bài toán qui hoạch tuyến tính đối ngẫu

$$\begin{array}{ll} \min z(x) = \langle c, x\rangle \quad (P) & \qquad \max w(y) = \langle b, y\rangle \quad (D) \\ \text{v.đ.k. } Ax \geq b & \qquad \text{v.đ.k. } A^T y \leq c \\ \qquad\quad x \geq 0 & \qquad\qquad y \geq 0. \end{array}$$

Ký hiệu:

X là tập chấp nhận được của bài toán gốc P;

Y là tập chấp nhận được của bài toán đối ngẫu (D);

$a^1, \ldots, a^m$ là các véc tơ hàng của ma trận A cấp $m \times n$;

$A_1, \ldots, A_n$ là các véc tơ cột của ma trận A cấp $m \times n$;

Véc tơ $b = (b_1, \ldots, b_m)^T \in \mathbb{R}^m$.

Định lý 3.14. (Định lý về độ lệch bù) *Giả sử $x^* \in X$ và $y^* \in Y$. Khi đó*

$$x^* \in \text{Argmin}(P), \quad \textit{và} \quad y^* \in \text{Argmax}(D)$$

khi và chỉ khi

$$\begin{cases} \langle Ax^* - b, y^*\rangle = 0 & (3.36) \\ \langle c - A^T y^*, x^*\rangle = 0. & (3.37) \end{cases}$$

Chứng minh. Từ (3.36) và (3.37) ta viết được

$$\langle b, y^*\rangle = \langle Ax^*, y^*\rangle \quad \text{và} \quad \langle c, x^*\rangle = \langle A^T y^*, x^*\rangle = \langle y^*, Ax^*\rangle.$$

Do đó

$$\langle b, y^*\rangle = \langle c, x^*\rangle.$$

Theo Hệ quả 3.7, ta có điều phải chứng minh. □

Định lý 3.15. (Định lý về độ lệch bù) *Giả sử* $x^* \in X$ *và* $y^* \in Y$. *Khi đó*

$$x^* \in \operatorname{Argmin}(P), \quad \textit{và} \quad y^* \in \operatorname{Argmax}(D)$$

khi và chỉ khi

$$\begin{cases} \langle Ax^* - b, y^* \rangle = 0 & (3.36) \\ \langle c - A^T y^*, x^* \rangle = 0. & (3.37) \end{cases}$$

Chú ý.

i) Vì $x^* \in X$ nên

$$Ax^* - b \geq 0 \text{ và } x^* \geq 0.$$

ii) Vì $y^* \in Y$ nên

$$c - A^T y^* \geq 0 \text{ và } y^* \geq 0.$$

iii) Vì

$$\langle Ax^* - b, y^* \rangle = \sum_{i=1}^{m} \Big(\langle a^i, x^* \rangle - b_i \Big) y_i^*$$

nên

$$(3.36) \Leftrightarrow \big(\langle a^i, x^* \rangle - b_i \big)\, y_i^* = 0 \ \ \forall i \Leftrightarrow \left[\begin{array}{ll} \text{Nếu } \ y_i^* > 0 & \text{thì } \ \langle a^i, x^* \rangle = b_i, \\ \text{Nếu } \ \langle a^i, x^* \rangle > b_i & \text{thì } \ y_i^* = 0. \end{array} \right.$$

iv) Vì

$$\langle c - A^T y^*, x^* \rangle = \sum_{j=1}^{n} \Big(c_j - \langle A_j, y^* \rangle \Big) x_j^*$$

$$(3.37) \Leftrightarrow (c_j - \langle A_j, y^* \rangle)\, x_j^* = 0 \ \forall j \Leftrightarrow \left[\begin{array}{ll} \text{Nếu } \ x_j^* > 0 & \text{thì } \ \langle A_j, y^* \rangle = c_j, \\ \text{Nếu } \ \langle y^*, A_j \rangle < c_j & \text{thì } \ x_j^* = 0. \end{array} \right.$$

Điều đó có nghĩa là nếu ràng buộc "lệch" khỏi thỏa mãn chặt thì biến phải bằng 0 (tức "bù lại") để "không lệch" khỏi thỏa mãn chặt và ngược lại.

Nhận xét 3.6.

i) Nếu bài toán gốc có dạng chính tắc

$$\min\{\langle c, x\rangle \mid Ax = b,\ x \geq 0\}, \tag{P}$$

thì (3.36) luôn thỏa mãn với mọi phương án chấp nhận được nên điều kiện độ lệch bù cho bài toán quy hoạch tuyến tính chính tắc là điều kiện (3.37).

ii) Điều kiện độ lệch bù có thể cho ta ngay nghiệm tối ưu của quy hoạch tuyến tính khi biết nghiệm tối ưu của bài toán đối ngẫu của nó.

Nhận xét 3.7. Trường hợp bài toán suy biến thì điều kiện độ lệch bù giúp rất ít cho việc xác định nghiệm tối ưu của bài toán quy hoạch tuyến tính.

3.8.4 Một số ứng dụng của lý thuyết đối ngẫu

a. Tìm nghiệm không âm của hệ phương trình tuyến tính

Ta có thể sử dụng Pha 1 của thuật toán đơn hình hai pha để tìm nghiệm không âm của hệ phương trình tuyến tính $Ax = b$, tức tìm nghiệm của hệ

$$Ax = b,\ x \geq 0, \tag{3.38}$$

trong đó A là ma trận cấp $m \times n$ và véc tơ $b \in \mathbb{R}^m,\ \ b \geq 0$. Cụ thể giải bài toán quy hoạch tuyến tính (xem Mục 3.5.2(b))

$$\begin{array}{rr} \min\ g(x,u) = & u_1 + \cdots + u_m \\ \text{v.đ.k.} & Ax + u = b \\ & (x,u) \geq 0, \end{array}$$

được nghiệm tối ưu (x^*, u^*). Nếu $u^* = 0$ thì x^* là nghiệm của hệ (3.38), còn nếu $u^* \neq 0$ thì hệ (3.38) vô nghiệm.

Ngược lại, ta cũng có thể đưa việc giải bài toán quy hoạch tuyến tính về việc tìm nghiệm không âm của một hệ phương trình tuyến tính. Thật vậy, không giảm tổng quát, ta xét bài toán quy hoạch tuyến tính dạng chuẩn tắc và bài toán đối ngẫu của nó

$$\begin{array}{rl} \min z = \langle c, x\rangle & \\ \text{v.đ.k.}\ \ Ax \geq b & \\ x \geq 0 & \end{array} \qquad\qquad \begin{array}{rl} \max w = \langle b, y\rangle & \\ \text{v.đ.k.}\ \ A^T y \leq c & \\ y \geq 0. & \end{array}$$

Theo Hệ quả 3.7, phương án chấp nhận được x là phương án tối ưu của bài toán gốc và phương án chấp nhận được y là phương án tối ưu của bài toán đối ngẫu khi và chỉ khi

$$\langle c, x\rangle = \langle b, y\rangle.$$

Điều đó có nghĩa là cặp nghiệm tối ưu x^* và y^* nghiệm đúng hệ phương trình, bất phương trình sau

$$\begin{cases} \langle c,x\rangle = \langle b,y\rangle \\ Ax \geq b, \quad x \geq 0 \\ A^T y \leq c, \quad y \geq 0\ . \end{cases}$$

Bằng cách đưa thêm biến phụ $u \geq 0, v \geq 0$, hệ trên được viết lại dưới dạng

$$\begin{cases} \langle c,x\rangle = \langle b,y\rangle \\ Ax - u = b, \quad A^T y + v = c, \\ x \geq 0, \quad y \geq 0, \quad u \geq 0, \quad v \geq 0\ . \end{cases}$$

Do đó mọi thuật toán tìm nghiệm không âm của một hệ phương trình tuyến tính đều có thể áp dụng trực tiếp vào việc giải bài toán quy hoạch tuyến tính.

b. Chứng minh Bổ đề Farkas

Bổ đề Farkas là một kết quả cơ bản của giải tích lồi và có nhiều ứng dụng trong quy hoạch toán học.

Bổ đề Farkas. *Cho véc tơ $a \in \mathbb{R}^n$ và ma trận A cấp $m \times n$. Khi đó $\langle a,x\rangle \geq 0$ với mọi x thỏa mãn $Ax \geq 0$ khi và chỉ khi tồn tại véc tơ $y \geq 0$ thuộc $\mathbb{R}^m$ sao cho $a = A^T y$.*

Chứng minh.
($\Leftarrow$) Điều kiện đủ là hiển nhiên. Thật vậy, giả sử tồn tại véc tơ $y \in \mathbb{R}^m$, $y \geq 0$ và $a = A^T y$. Suy ra $\langle a,x\rangle = \langle A^T y,x\rangle = \langle y,Ax\rangle \geq 0$ với mọi x thỏa mãn $Ax \geq 0$.

($\Rightarrow$) Xét quy hoạch tuyến tính

$$\min \varphi(x)\langle a,x\rangle \text{ v.đ.k.} x \in M, \qquad (P)$$

trong đó

$$M = \{x \in \mathbb{R}^n \mid Ax \geq 0\}.$$

Dễ thấy $0 \in M$, tức $M \neq \emptyset$ và

$$\varphi(x) = \langle a,x\rangle \geq 0 \quad \forall x \in M.$$

Suy ra bài toán (P) có nghiệm tối ưu. Vậy bài toán đối ngẫu của nó là

$$\max\{\langle 0,y\rangle \mid A^T y = a,\ y \geq 0\}$$

cũng có nghiệm tối ưu (Định lý 3.13) và tập chấp nhận được của bài toán đối ngẫu khác rỗng, tức

$$\exists\ y \geq 0 \text{ sao cho } A^T y = a.$$

$\square$

c. Ý nghĩa kinh tế của cặp bài toán đối ngẫu
Ta trình bày vấn đề này thông qua một bài toán thực tế: Người ta dự định sản xuất m loại sản phẩm bằng n phương pháp khác nhau. Biết rằng nhu cầu của xã hội về từng loại sản phẩm i là b_i, $i = 1, \dots, m$. Nếu sử dụng một đơn vị thời gian theo phương pháp j thì thu được a_{ij} đơn vị sản phẩm loại i và phải trả một chi phí bằng c_j. Bài toán đặt ra là xác định thời gian t_j sử dụng mỗi phương pháp j, $j = 1, \dots, n$, sao cho số sản phẩm loại i sản xuất được không ít hơn nhu cầu b_i, $i = 1, \dots, m$ đồng thời tổng chi phí sản xuất là thấp nhất.

Sau đây là mô hình toán học của bài toán:

$$\begin{aligned} \min\ f(x) &= c_1 t_1 + \cdots + c_n t_n \\ \text{v.đ.k.}\quad \sum_{j=1}^{n} a_{ij} t_j &\geq b_i, \quad i = 1, \dots, m, \\ t_j &\geq 0, \quad j = 1, \dots, n. \end{aligned}$$

Bài toán đối ngẫu của bài toán này là:

$$\begin{aligned} \max\ g(z) &= b_1 z_1 + \cdots + b_m z_m \\ \text{v.đ.k.}\quad \sum_{i=1}^{m} a_{ij} z_i &\leq c_j, \quad j = 1, \dots, n, \\ z_i &\geq 0, \quad i = 1, \dots, m. \end{aligned}$$

Nếu gọi z_i, $i = 1, \dots, m$ là giá trị (quy ước) của một đơn vị sản phẩm loại i thì hàm mục tiêu của bài toán đối ngẫu là hàm tổng giá trị toàn bộ sản phẩm theo yêu cầu của xã hội. Ràng buộc chính thứ j của bài toán đối ngẫu có nghĩa là: tổng giá trị m loại sản phẩm được sản xuất theo phương pháp thứ j trong một đơn vị thời gian không vượt quá chi phí sản xuất c_j, $j = 1, \dots, n$. Mục đích của bài toán đối ngẫu là xác định giá trị z_i cho mỗi đơn vị sản phẩm loại i sao cho tổng giá trị toàn bộ sản phẩm theo yêu cầu của xã hội là lớn nhất.

Ta có thể phân tích một số ý nghĩa kinh tế của cặp bài toán đối ngẫu trên như sau: Gọi t^* là phương án sản xuất tối ưu của bài toán gốc và z^* là phương án định giá trị tối ưu của bài toán đối ngẫu. Từ Định lý về độ lệch bù (Định lý 3.14), ta suy ra:

i) Nếu phương pháp thứ j được áp dụng (tức $t_j^* > 0$) thì tổng giá trị m loại sản phẩm được sản xuất theo phương pháp này phải vừa đúng bằng chi phí

$$\sum_{i=1}^{m} a_{ij} z_i^* = c_j$$

hoặc phương pháp thứ j sẽ không được áp dụng (tức $t_j^* = 0$) nếu tổng giá trị m loại sản phẩm được sản xuất theo phương pháp đó thấp hơn chi phí

$$\sum_{i=1}^{m} a_{ij} z_i^* < c_j.$$

ii) Nếu sản phẩm loại i có giá trị $z_i^* > 0$ thì tổng số sản phẩm loại đó sản xuất được phải vừa đúng bằng nhu cầu xã hội

$$\sum_{j=1}^{n} a_{ij} t_j^* = b_i$$

hoặc sản phẩm loại i không có giá trị (tức $z_i^* = 0$) nếu tổng số sản phẩm loại đó sản xuất được lại vượt quá nhu cầu xã hội

$$\sum_{j=1}^{n} a_{ij} t_j^* > b_i.$$

Bài tập Chương 3

1. Chuyển bài toán quy hoạch tuyến tính sau về dạng chính tắc:

$$\begin{aligned}
\max\quad & z = 3x_1 + 5x_2 - 4x_3 \\
\text{v.đ.k.}\quad & 7x_1 - 2x_2 - 3x_3 \geq 4 \\
& 2x_1 - 4x_2 - 8x_3 = 3 \\
& 5x_1 - 3x_2 - 2x_3 \leq 9 \\
& x_1 \geq 1, x_2 \leq 7, x_3 \geq 0.
\end{aligned}$$

2. Chuyển bài toán quy hoạch tuyến tính sau về dạng chính tắc:

$$\begin{aligned}
\max\quad & z = 5x_1 - x_2 - 7x_3 \\
\text{v.đ.k}\quad & 7x_1 + 4x_2 - 11x_3 \geq 12 \\
& x_1 - 3x_2 - 9x_3 = -5 \\
& 6x_1 + 5x_2 + 2x_3 \leq 10 \\
& x_1 \geq -2, x_2 \text{ tự do}, \ x_3 \geq 0.
\end{aligned}$$

3. Khi nào bài toán quy hoạch tuyến tính $\min\{\langle c, x\rangle | x \in M\}$ có nghiệm? Cấu trúc của tập nghiệm của bài toán này?

4. Giả sử rằng một quy hoạch tuyến tính với tập chấp nhận được bị chặn có ℓ phương án cực biên tối ưu $v^1, \dots, v^\ell$. Chứng minh rằng một phương án là tối ưu khi và chỉ khi nó là tổ hợp lồi của $v^1, \dots, v^\ell$.
5. Cho x^1, x^2 là hai phương án tối ưu của bài toán $\min\{\langle c, x\rangle \mid Ax = b, x \geq 0\}$ và x^0 là tổ hợp lồi của x^1, x^2. Chứng minh rằng x^0 cũng là nghiệm của bài toán này.
6. Giải bài toán quy hoạch tuyến tính sau bằng phương pháp hình học:

   i)

$$
\begin{aligned}
\min \quad & z = 3x_1 + x_2 \\
\text{v.đ.k.} \quad & x_1 - x_2 \leq 1 \\
& 3x_1 + 2x_2 \leq 12 \\
& 2x_1 + 3x_2 \leq 3 \\
& -2x_1 + 3x_2 \geq 9 \\
& x_1, \ x_2 \geq 0.
\end{aligned}
$$

   ii)

$$
\begin{aligned}
\min \quad & z = x_1 + 2x_2 \\
\text{v.đ.k.} \quad & 2x_1 + x_2 \geq 12 \\
& x_1 + x_2 \geq 5 \\
& -x_1 + 3x_2 \leq 3 \\
& 6x_1 - x_2 \geq 12 \\
& x_1, \ x_2 \geq 0.
\end{aligned}
$$

7. Tìm các phương án cực biên không suy biến của bài toán quy hoạch tuyến tính sau đây và trả lời xem các bài toán này có nghiệm không? Vì sao?

   i)

$$
\begin{aligned}
\min f(x) = & \ 2x_1 + 4x_2 - 3x_3 \\
\text{v.đ.k.} \quad & x_1 - x_2 - x_3 = 1 \\
& x_1 + x_2 + x_3 = 3 \\
& x_1, \ x_2, \ x_3 \geq 0.
\end{aligned}
$$

   ii)

$$
\begin{aligned}
\max f(x) = & \ 3x_1 - 6x_2 + x_3 - 3x_4 \\
\text{v.đ.k.} \quad & x_1 + x_2 + x_3 + x_4 = 10 \\
& 2x_2 + x_3 - x_4 = 6 \\
& x_1, \ x_2, \ x_3, \ x_4 \geq 0.
\end{aligned}
$$

8. Xét bài toán:

$$\begin{aligned}
\min\ & z = 2x_1 + 9x_2 + 3x_3 \\
\text{v.đ.k.}\ & -2x_1 + 2x_2 + x_3 \geq 1 \\
& x_1 + 4x_2 - 3x_3 \geq 1 \\
& x_1,\ x_2,\ x_3 \geq 0.
\end{aligned}$$

   i) Viết bài toán đối ngẫu của bài toán này và giải nó bằng phương pháp hình học.

   ii) Sử dụng điều kiện độ lệch bù để tìm nghiệm tối ưu của bài toán ban đầu.

9. Xét bài toán $\min\{z = \langle c, x\rangle |\ Ax = b,\ x \geq 0\}$. Giả sử bài toán này và bài toán đối ngẫu của nó đều chấp nhận được. Cho x^* là một nghiệm tối ưu của bài toán gốc với giá trị tối ưu tương ứng là z_* và y^* là một nghiệm tối ưu của bài toán đối ngẫu. Chứng minh rằng:

$$z_* = y^{*T} A x^*.$$

10. Cho bài toán quy hoạch tuyến tính

$$\begin{aligned}
\max\ & z = 6x_1 + 2x_2 + 4x_3 \\
\text{v.đ.k.}\ & 2x_1 + 4x_2 + 2x_3 \leq 8 \\
& x_1 + 2x_2 + 3x_3 \leq 7 \\
& x_1,\ x_2,\ x_3 \geq 0.
\end{aligned}$$

   i) Bài toán này có nghiệm không? Vì sao? Tìm một bài toán thực tế có thể mô tả bởi quy hoạch tuyến tính này.

   ii) Chứng minh rằng tập nghiệm của bài toán quy hoạch tuyến tính là tập lồi.

11. Cho bài toán quy hoạch tuyến tính:

$$\begin{aligned}
\min\ & f(x) = -2x_1 - x_2 + 3x_3 + x_4 - 4x_5 \\
\text{v.đ.k}\ & x_1 - x_2 + 4x_3 - 2x_5 - x_6 = -4 \\
& 3x_1 + 2x_2 - x_3 + x_4 + x_7 = 24 \\
& 5x_1 + 3x_2 + x_3 + 2x_4 - x_5 = 46 \\
& x_1, \ldots, x_7 \geq 0.
\end{aligned}$$

   i) Chứng minh rằng $x^* = (0, 2, 0, 20, 0, 2, 0)^T$ là phương án cực biên không suy biến.

ii) Xuất phát từ x^*, giải bài toán trên bằng phương pháp đơn hình.

iii) Tìm một phương án chấp nhận được $\bar{x}$ có $f(\bar{x}) = -87$.

12. Cho bài toán quy hoạch tuyến tính:

$$\begin{array}{rl} \min & z = x_1 - 4x_2 + \alpha.x_3 \quad\quad + 3x_5 \\ & 2x_1 + x_2 + x_3 \quad\quad\quad = 3 \\ & -x_1 \quad + 2x_3 + x_4 \quad = 2 \\ & 3x_1 \quad + 2x_3 \quad + x_5 = 4 \\ & x_1,\ x_2,\ x_3,\ x_4,\ x_5 \geq 0. \end{array}$$

i) Lập bảng đơn hình tương ứng với phương án cực biên $x^0 = (0, 3, 0, 2, 4)^T$.

ii) Tìm điều kiện của tham số α để phương án trên là tối ưu.

13. Chứng minh rằng nếu bài toán $\min\{\langle c, x\rangle \mid Ax = b,\ x \geq 0\}$ có phương án tối ưu thì các bài toán $\min\{\langle c, x\rangle \mid Ax = \bar{b},\ x \geq 0\}$ cũng có phương án tối ưu, trong đó $\bar{b}$ là véc tơ vế phải bất kỳ miễn là tập chấp nhận được $\{x \in \mathbb{R}^n \mid Ax = \bar{b},\ x \geq 0\}$ khác rỗng.

14. Cho bài toán quy hoạch tuyến tính:

$$\begin{array}{rl} \max & z = -4x_1 + 3x_2 - x_3 - 2x_4 \\ \text{v.đ.k.} & x_1 + 2x_2 + x_3 \quad\quad = 1 \\ & 4x_2 - 2x_3 \quad\quad \leq 4 \\ & -3x_2 \quad\quad + x_4 = 5 \\ & x_1,\ x_2,\ x_3,\ x_4 \geq 0. \end{array}$$

i) Viết bài toán đối ngẫu của bài toán trên.

ii) Dùng đối ngẫu kiểm tra xem $x^* = (0,\ 0,\ 1,\ 5)^T$ có phải là phương án tối ưu của bài toán trên không?

15. Giải bài toán quy hoạch tuyến tính sau bằng thuật toán đơn hình hai pha:

i)

$$\begin{array}{rl} \min & z = -4x_1 - 2x_2 \\ \text{v.đ.k.} & 3x_1 - 2x_2 \geq 4 \\ & -2x_1 + x_2 = 2 \\ & x_1,\ x_2 \geq 0. \end{array}$$

ii)

$$\begin{array}{rrl} \min & z = & -\,x_1 - 3x_2 \\ \text{v.đ.k.} & & x_1 + \ x_2 \geq 3 \\ & & -x_1 + \ x_2 \leq -1 \\ & & x_1 + 2x_2 \leq 4 \\ & & x_1, \ \ x_2 \geq 0. \end{array}$$

16. Một xưởng sản xuất may gia công hai kiểu mũ. Thời gian lao động để làm ra một mũ kiểu thứ nhất nhiều gấp hai lần thời gian để làm xong một mũ kiểu thứ hai. Nếu sản xuất toàn mũ kiểu thứ hai thì họ làm được 500 mũ một ngày. Thị trường tiêu thụ được trong mỗi ngày không quá 200 mũ kiểu thứ nhất và 300 mũ kiểu thứ hai. Tiền lãi một mũ kiểu thứ nhất là 50000 *đồng* và một mũ kiểu thứ hai là 30000 *đồng*. Xưởng sản xuất này cần gia công mỗi kiểu mũ bao nhiêu chiếc trong một ngày để tổng tiền lãi là lớn nhất?

17. Hãng hàng không Universal Aviation quyết định đầu tư 4000000 \$ để mua ba loại máy bay A, B, C để thực hiện hợp đồng chuyên chở hàng cho tập đoàn RIO. Biết rằng:

    Loại A giá 80000 \$, có tải trọng 10 *tấn*, tốc độ 350 km/h và có thể bay 18 h trong một ngày.

    Loại B giá 130000 \$, có tải trọng 20 *tấn* và tốc độ 300 km/h và có thể bay 18 h trong một ngày .

    Loại C giá 150000 \$, có tải trọng 18 *tấn* và tốc độ 300 km/h và có thể bay 21 h trong một ngày .

    Máy bay loại A và loại C có thể bay 3 ca một ngày. Máy bay loại B có thể bay 6 ca một ngày.

    Hãng hàng không chỉ có thể thực hiện bay nhiều nhất là 150 ca trong một ngày, và hãng chỉ mua tối đa 30 máy bay.

    Số tiền thu được từ mỗi máy bay loại A, B, C khi chúng chuyên chở 1 *tấn* hàng và bay quãng đường 1 km (gọi tắt là 1 *tấn* - km) lần lượt là 5 \$, 8 \$ và 12 \$.

    Mỗi ngày, hãng hàng không phải thực hiện hợp đồng chuyên chở cho tập đoàn RIO là 3500000 *tấn* - km. Nếu máy bay của hãng Universal Aviation không hoàn thành hợp đồng với tập đoàn RIO thì hãng hàng không này có thể thuê một hãng hàng không khác chở hàng mà vẫn lãi 0.2 \$ trên 1 *tấn* - km. Bài toán đặt ra là cần mua bao nhiêu máy bay mỗi loại sao cho hãng hàng không có thể thu được số tiền lớn nhất mỗi ngày? Hãy thiết lập mô hình quy hoạch tuyến tính cho bài toán này.

18. Mỗi tháng công ty Hoàng Anh cần 90000 m^3 gỗ xẻ đã sấy khô để đóng đồ gia dụng. Công ty mua gỗ xẻ theo hai cách sau:

    Cách 1: Mua hai loại gỗ đã được xẻ sẵn (chưa sấy) của các công ty khác rồi đem sấy. Biết rằng gỗ đã xẻ loại thứ nhất giá 300000 *đồng*/m^3 và sau khi sấy khô 1 m^3 chỉ còn 0.7 m^3, loại thứ hai giá 700000 *đồng*/m^3 và sau khi sấy khô 1 m^3 chỉ còn 0.9 m^3. Mỗi tháng, công ty có thể mua nhiều nhất 40000 m^3 gỗ xẻ loại thứ nhất và 60000 m^3 gỗ xẻ loại thứ hai.

    Cách 2: Mua gỗ hộp (chưa xẻ) sau đó xẻ tại xưởng cưa của công ty và đem sấy, chi phí vận chuyển 1 m^3 gỗ hộp đến nhà máy cưa mất 25000 *đồng*. Tiền mua và xẻ 1 m^3 gỗ hộp là 60000 *đồng* và 1 m^3 gỗ này chỉ còn 0.8 m^3 sau khi sấy khô. Phân xưởng cưa có thể cưa 35000 m^3 gỗ trong 1 tháng.

    Chi phí để sấy gỗ xẻ là 40000 *đồng*/m^3. Tổng thời gian làm khô gỗ tối đa là 40 giờ, trong đó thời gian làm khô 1 m^3 gỗ loại thứ nhất (gỗ xẻ sẵn) là 2 phút, loại thứ hai (gỗ xẻ sẵn) là 0.8 phút và loại gỗ hộp là 1.3 phút.

    Hãy xây dựng mô hình bài toán quy hoạch tuyến tính để giúp công ty có thể cực tiểu chi phí mua nguyên liệu hàng tháng.

Chương 4

Qui hoạch phi tuyến (Nonlinear Programming)

4.1 Hàm lồi

4.1.1 Định nghĩa

Cho tập lồi khác rỗng $X \subseteq \mathbb{R}^n$ và hàm số $f : X \to \mathbb{R}$. Khi đó,

◇ f được gọi là *hàm lồi* (convex function) nếu

$$f(\lambda x^1 + (1-\lambda)x^2) \leq \lambda f(x^1) + (1-\lambda)f(x^2)$$

với bất kỳ $x^1, x^2 \in X$ và số thực $\lambda \in [0, 1]$.

◇ f được gọi là *hàm lồi chặt* (strictly convex function) nếu

$$f(\lambda x^1 + (1-\lambda)x^2) < \lambda f(x^1) + (1-\lambda)f(x^2)$$

với bất kỳ $x^1, x^2 \in X$, $x^1 \neq x^2$ và $0 < \lambda < 1$.

◇ *Miền xác định hữu hiệu* của hàm f là

$$\mathrm{dom} f := \{x \in X |\ f(x) < +\infty\}.$$

◇ *Epigraph* của hàm f, ký hiệu là epi(f) được định nghĩa như sau:

$$\mathrm{epi}(f) := \{(x, \xi) \in X \times \mathbb{R} \mid \xi \geq f(x)\} \subset \mathbb{R}^{n+1}$$

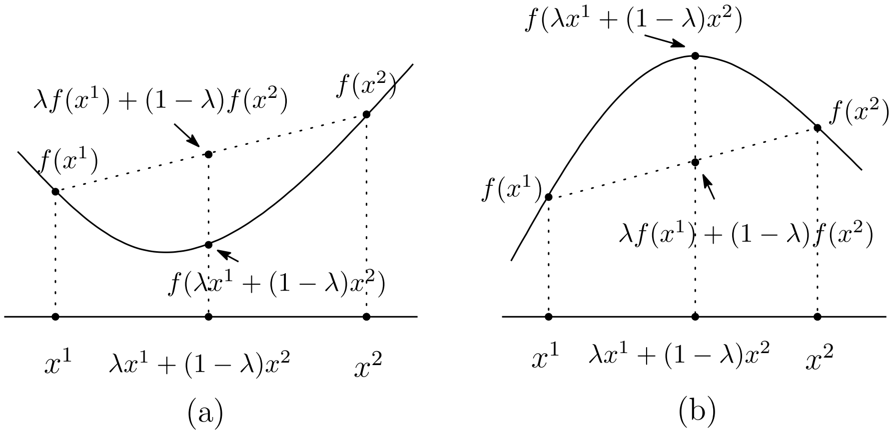
◇ *Hypograph* của hàm f, lý hiệu là hypo(f)), xác định bởi

$$\mathrm{hypo}(f) := \{(x, \zeta) \in X \times \mathbb{R} \mid \zeta \leq f(x)\} \subset \mathbb{R}^{n+1}.$$

◇ Hàm lồi $f : X \to \mathbb{R} \cup \{+\infty\}$ có thể được mở rộng thành một hàm lồi trên toàn không gian $\mathbb{R}^n$ bằng cách đặt $f(x) = +\infty$ nếu $x \notin \text{dom}f$. Vì vậy, để đơn giản, ta thường xét f là hàm lồi trên $\mathbb{R}^n$.

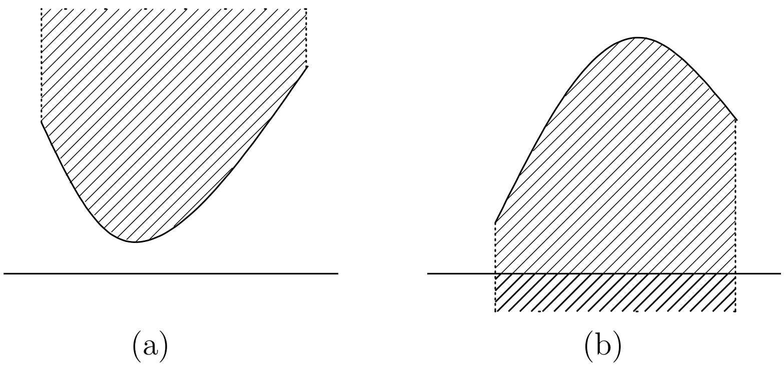
◇ f được gọi là *hàm lõm* (concave function) trên tập lồi X nếu $-f$ là hàm lồi.

◇ f được gọi là *hàm lõm chặt* (strictly concave function)) trên tập lồi X nếu $-f$ là hàm lồi chặt.



Hình 6.1 *(a) – Hàm lồi; (b) – Hàm lõm*

Chú ý. Cho tập lồi khác rỗng $X \subseteq \mathbb{R}^n$ và hàm số $f : X \to \mathbb{R}$. Khi đó f là hàm lồi khi và chỉ khi với mỗi điểm $x^0 \in X$ và mỗi hướng $v \in \mathbb{R}^n$, hàm một biến $\varphi(t) := f(x^0 + tv)$ là hàm lồi trên tập lồi $\{t \in \mathbb{R} \mid x^0 + tv \in X\} \subseteq \mathbb{R}$.



Hình 6.2. *(a) – Epigraph của một hàm lồi* *(b) – Hypograph của một hàm lõm*

Lược đồ chứng minh hàm lồi Cho hàm số $f : X \to \mathbb{R}$, trong đó $X \subseteq \mathbb{R}^n$. Câu hỏi: f có phải là hàm lồi (*trên* X) không ?

Bước 1 Kiểm tra X là tập lồi?

If X là tập lồi **Then** Chuyển Bước 2

Else Dừng. Kết luận: f không phải là hàm lồi trên X.

Bước 2

Cách 1. Lấy tùy ý $x^1, x^2 \in X$ và lấy tùy ý số thực $\lambda \in [0, 1]$. Kiểm tra

$$f(\lambda x^1 + (1-\lambda)x^2) \leq \lambda f(x^1) + (1-\lambda)f(x^2). \tag{1}$$

If (1) đúng **Then** Dừng. Kết luận: f là hàm lồi trên X

Else Dừng. Kết luận: f không phải là hàm lồi trên X.

Cách 2. . . .

Cách 3. . . .

Mệnh đề 4.1. *Cho hàm f xác định trên tập lồi khác rỗng $X \subseteq \mathbb{R}^n$. Khi đó,*

i) *Hàm f là hàm lồi khi và chỉ khi* epi*(f) là tập lồi.*

ii) *Hàm f là hàm lõm khi và chỉ khi* hypo*(f) là tập lồi.*

Chứng minh: Do tính tương tự, ta chỉ cần chứng minh khẳng định i).

($\Rightarrow$) Giả sử f là hàm lồi trên tập lồi X. Lấy hai điểm bất kỳ (x^1, ξ_1) và (x^2, ξ_2) thuộc epi(f) và số thực $\lambda \in [0, 1]$. Ta sẽ chứng minh

$$\lambda(x^1, \xi_1) + (1 - \lambda)(x^2, \xi_2) \in \text{epi}(f).$$

Thật vậy, theo định nghĩa

$$\xi_1 \geq f(x^1) \text{ và } \xi_2 \geq f(x^2). \tag{6.1}$$

Ta có

$$\begin{aligned} f(\lambda x^1 + (1-\lambda)x^2) &\leq \lambda f(x^1) + (1-\lambda) f(x^2) && \text{(do } f \text{ là hàm lồi)} \\ &\leq \lambda \xi_1 + (1-\lambda)\xi_2 && \text{(do (6.1))}. \end{aligned}$$

Suy ra

$$\Big(\lambda x^1 + (1-\lambda)x^2, \lambda\xi_1 + (1-\lambda)\xi_2\Big) = \lambda(x^1, \xi_1) + (1-\lambda)(x^2, \xi_2) \in \text{epi}(f).$$

($\Leftarrow$) Ngược lại, giả sử epi(f) là tập lồi. Lấy tùy ý $x^1, x^2 \in X$ và số thực $\lambda \in [0, 1]$. Ta sẽ chứng minh

$$f(\lambda x^1 + (1-\lambda)x^2) \leq \lambda f(x^1) + (1-\lambda) f(x^2).$$

Thật vậy, vì

$$(x^1, f(x^1)),\ (x^2, f(x^2)) \ \in\ \text{epi}(f)$$

nên

$$\lambda\Big(x^1, f(x^1)\Big) + (1-\lambda)\Big(x^2, f(x^2)\Big) =$$

$$= \Big(\lambda x^1 + (1-\lambda)x^2, \lambda f(x^1) + (1-\lambda) f(x^2)\Big) \in \text{epi}(f).$$

Theo định nghĩa ta có

$$f(\lambda x^1 + (1-\lambda)x^2) \leq \lambda f(x^1) + (1-\lambda) f(x^2) \ \ \forall\, 0 \leq \lambda \leq 1.$$

Điều đó chứng tỏ f là hàm lồi. $\square$

Mệnh đề 4.2.

i) *Nếu hàm số f xác định trên tập lồi $X \subseteq \mathbb{R}^n$ là hàm lồi thì tập mức dưới $L_\alpha(f) := \{x \in X \mid f(x) \leq \alpha\}$ là tập lồi với mọi $\alpha \in \mathbb{R}$.*

ii) *Nếu g là hàm lõm xác định trên tập lồi khác rỗng $X \subseteq \mathbb{R}^n$ thì tập mức trên $L^\alpha(g) := \{x \in X \mid g(x) \geq \alpha\}$ là tập lồi với mọi $\alpha \in \mathbb{R}$.*

Chứng minh: Ta chỉ cần chứng minh i). Lấy bất kỳ $x^1, x^2 \in L_\alpha(f)$ và $0 \leq \lambda \leq 1$. Khi đó

$$\begin{aligned} f(\lambda x^1 + (1-\lambda)x^2) &\leq \lambda f(x^1) + (1-\lambda)f(x^2) \quad &&(\text{do } f \text{ là hàm lồi}) \\ &\leq \lambda\alpha + (1-\lambda)\alpha &&(\text{do } x^1, x^2 \in L_\alpha(f)) \\ &= \alpha. \end{aligned}$$

Suy ra $\lambda x^1 + (1-\lambda)x^2 \in L_\alpha(f)$. Vậy $L_\alpha(f)$ là một tập lồi. □

Chú ý 4.1. Mệnh đề 6.2 chỉ là điều kiện cần, không phải là điều kiện đủ. Cụ thể, xét hàm số $f : X \to \mathbb{R}$, trong đó $X \subset \mathbb{R}^n$ là tập lồi khác rỗng.

- ⋄ Nếu tập mức dưới $L_\alpha(f)$ của hàm f là tập lồi với mọi $\alpha \in \mathbb{R}$ thì hàm f chưa chắc đã là hàm lồi.
- ⋄ Nếu tập mức trên $L^\alpha(f)$ của hàm f là tập lồi với mọi $\alpha \in \mathbb{R}$ thì hàm f chưa chắc đã là hàm lõm.
- ⋄ Một hàm mà mọi tập mức dưới là tập lồi được gọi là một hàm *tựa lồi* (quasiconvex function).
- ⋄ Tương tự, một hàm mà mọi tập mức trên là tập lồi được gọi là một hàm *tựa lõm* (quasiconcave function).

4.1.2 Các phép toán về hàm lồi

Cho hàm lồi f_1 xác định trên tập lồi $X_1 \subseteq \mathbb{R}^n$, hàm lồi f_2 xác định trên tập lồi $X_2 \subseteq \mathbb{R}^n$ và số thực $\lambda \geq 0$. Các phép toán $\lambda f_1,\ f_1 + f_2,\ \max\{f_1, f_2\}$ được định nghĩa như sau:

$$\begin{aligned} (\lambda f_1)(x) &:= \lambda f_1(x), \quad x \in X_1; \\ (f_1 + f_2)(x) &:= f_1(x) + f_2(x), \quad x \in X_1 \cap X_2; \\ \max\{f_1, f_2\}(x) &:= \max\{f_1(x), f_2(x)\}, \quad x \in X_1 \cap X_2. \end{aligned}$$

Kết quả sau dễ dàng suy ra từ định nghĩa.

Mệnh đề 4.3. *Cho f_1 là hàm lồi trên tập lồi X_1, f_2 là hàm lồi trên tập lồi X_2 và các số thực $\alpha > 0$, $\beta > 0$. Khi đó các hàm $\alpha f_1 + \beta f_2$ và $\max\{f_1, f_2\}$ là lồi trên $X_1 \cap X_2$.*

4.1.3 Tính liên tục của hàm lồi

Định lý 4.1. *Nếu f là hàm lồi xác định trên tập lồi mở $X \subseteq \mathbb{R}^n$ thì f liên tục trên tập X.*

Nhận xét 4.1. Sự gián đoạn của hàm lồi chỉ có thể xảy ra tại biên của tập xác định.

4.1.4 Đạo hàm theo hướng của hàm lồi

Định nghĩa. Cho hàm $f : \mathbb{R}^n \to \mathbb{R}$, véc tơ $d \in \mathbb{R}^n \setminus \{0\}$. Nếu tồn tại giới hạn

$$\lim_{t \to 0^+} \frac{f(x^0 + td) - f(x^0)}{t},$$

(hữu hạn hoặc vô cùng), thì giới hạn đó được gọi là *đạo hàm theo hướng* d của hàm f tại điểm $x^0 \in \mathbb{R}^n$ và ký hiệu là $f'(x^0, d)$,tức

$$f'(x^0, d) = \lim_{t \to 0^+} \frac{f(x^0 + td) - f(x^0)}{t}.$$

Ví dụ.

◇ Cho hàm $f : \mathbb{R}^n \to \mathbb{R}$, véc tơ $d = e^i$, trong đó $e^i = (0, \ldots, \underbrace{1}_{i}, \ldots, 0)^T$.

Khi đó
$$\frac{\partial f(x^0)}{\partial x_i} = \lim_{t \to 0^+} \frac{f(x^0 + te^i) - f(x^0)}{t} = f'(x^0, e^i).$$

◇ Cho hàm $f : \mathbb{R} \to \mathbb{R}$, điểm $x^0 \in \mathbb{R}$, $d = 1$. Khi đó

$$f'(x^0, 1) = \lim_{t \to 0^+} \frac{f(x^0 + t) - f(x^0)}{t} = f'_+(x^0),$$

trong đó $f'_+(x^0)$ là đạo hàm phải của f tại x^0

◇ Cho hàm $f : \mathbb{R} \to \mathbb{R}$, điểm $x^0 \in \mathbb{R}$, $d = -1$. Khi đó

$$f'(x^0, -1) = \lim_{t \to 0^+} \frac{f(x^0 + t.(-1)) - f(x^0)}{t} = \lim_{t \to 0^-} \frac{f(x^0 + t) - f(x^0)}{-t} = -f'_-(x^0),$$

với $f'_-(x^0)$ là đạo hàm trái của f tại x^0.

Mệnh đề 4.4. *Cho hàm f xác định trên $\mathbb{R}^n$ và điểm $x^0 \in \mathbb{R}^n$. Nếu f khả vi tại x^0 thì*

$$f'(x^0, d) = \langle \nabla f(x^0), d \rangle \quad \forall\, d \in \mathbb{R}^n \setminus \{0\}.$$

Chứng minh. Giả sử f khả vi tại x^0. Theo định nghĩa, với mọi $v \in \mathbb{R}^n \setminus \{0\}$ và $\|v\|$ đủ nhỏ, ta có

$$f(x^0 + v) = f(x^0) + \langle \nabla f(x^0), v \rangle + o(\|v\|).$$

Vì vậy, với mọi $d \in \mathbb{R}^n \setminus \{0\}$ ta có

$$\lim_{t \to 0^+} \frac{f(x^0 + td) - f(x^0) - \langle \nabla f(x^0), td \rangle}{t\|d\|} = 0.$$

Do đó

$$\frac{f'(x^0, d) - \langle \nabla f(x^0), d \rangle}{\|d\|} = 0$$

và ta nhận được điều phải chứng minh. □

Nhận xét 4.2. Đặt $\varphi(t) := f(x^0 + td)$. Khi đó, theo định nghĩa ta có

$$\varphi'(0) = \frac{d\varphi(t)}{dt}|_{t=0} = \lim_{t\to 0} \frac{\varphi(t) - \varphi(0)}{t} = \lim_{t\to 0^+} \frac{f(x^0 + td) - f(x^0)}{t} = f'(x^0, d).$$

◇ Đạo hàm theo hướng của hàm f tại x^0 phản ánh tốc độ biến thiên của hàm f tại x^0 theo hướng đó.

◇ Theo Mệnh đề 4.4, ta có

$$f'(x^0, d) = \langle \nabla f(x^0), d \rangle.$$

. Theo bất đẳng thức Cauchy-Bunjakowski-Schwarz, trong tất cả các hướng $d \in \mathbb{R}^n$ có $\|d\| = 1$, ta có

$$|\langle \nabla f(x^0), d \rangle| \leq \|\nabla f(x^0)\| \|d\| = \|\nabla f(x^0)\|$$

$$\Rightarrow -\|\nabla f(x^0)\| \leq \langle \nabla f(x^0), d \rangle \leq \|\nabla f(x^0)\|.$$

Suy ra:

- Đạo hàm theo hướng của hàm f tại điểm x^0 đã cho là lớn nhất khi hướng

$$d = \frac{\nabla f(x^0)}{\|\nabla f(x^0)\|},$$

tức giá trị hàm tăng nhanh nhất theo hướng gradient.

- Đạo hàm theo hướng của hàm f tại điểm x^0 đã cho là nhỏ nhất khi

$$d = -\frac{\nabla f(x^0)}{\|\nabla f(x^0)\|},$$

giá trị hàm giảm nhanh nhất theo hướng ngược với gradient.

Tính đặc thù của đạo hàm theo hướng của hàm lồi được mô tả ở định lý sau.

Định lý 4.2. *Nếu $f : X \to \mathbb{R} \cup \{+\infty\}$ là một hàm lồi xác định trên tập lồi $X \subseteq \mathbb{R}^n$ thì nó có đạo hàm theo mọi hướng $d \in \mathbb{R}^n \setminus \{0\}$ tại mọi điểm $x^0 \in \mathrm{dom} f$ và*

$$f'(x^0, d) \leq f(x^0 + d) - f(x^0).$$

Hệ quả 4.1. *Nếu f là hàm lồi khả vi xác định trên tập lồi mở X thì f có đạo hàm theo mọi hướng $d \in \mathbb{R}^n \setminus \{0\}$ tại mọi điểm $x^0 \in \mathrm{dom} f$ và*

$$\langle \nabla f(x^0), d \rangle = f'(x^0, d) \leq f(x^0 + d) - f(x^0).$$

4.1.5 Tiêu chuẩn nhận biết hàm lồi khả vi

Định lý 4.3. *Cho f là hàm khả vi trên tập lồi mở $X \subseteq \mathbb{R}^n$. Khi đó:*

i) Hàm f là hàm lồi trên X khi và chỉ khi

$$f(y) - f(x) \geq \Big\langle \nabla f(x), y - x \Big\rangle \quad \forall\, x, y \in X;$$

ii) Hàm f là hàm lõm trên X khi và chỉ khi

$$f(y) - f(x) \leq \Big\langle \nabla f(x), y - x \Big\rangle \quad \forall\, x, y \in X.$$

Định lý 4.4. *Cho f là hàm khả vi hai lần trên tập lồi mở $X \subseteq \mathbb{R}^n$. Khi đó,*

i) *Hàm f là hàm lồi trên X khi và chỉ khi ma trận Hesse $\nabla^2 f(x)$ là nửa xác định dương trên X, tức với mỗi $x \in X$,*

$$y^T \nabla^2 f(x) y \geq 0 \quad \forall y \in \mathbb{R}^n.$$

Hàm f là hàm lồi chặt trên X nếu $\nabla^2 f(x)$ xác định dương trên X, tức với mỗi $x \in X$,

$$y^T \nabla^2 f(x) y > 0 \quad \forall y \in \mathbb{R}^n \setminus \{0\}.$$

ii) *Hàm f là hàm lõm trên X khi và chỉ khi ma trận Hesse $\nabla^2 f(x)$ là nửa xác định âm trên X, tức với mỗi $x \in X$,*

$$y^T \nabla^2 f(x) y \leq 0 \quad \forall y \in \mathbb{R}^n.$$

Hàm f là hàm lõm chặt trên X nếu $\nabla^2 f(x)$ xác định âm trên X, tức với mỗi $x \in X$,

$$y^T \nabla^2 f(x) y < 0 \quad \forall y \in \mathbb{R}^n \setminus \{0\}.$$

Hệ quả 4.2. *Cho hàm toàn phương*

$$f(x) = \frac{1}{2}\langle x, Qx\rangle + \langle x, a\rangle + \alpha,$$

trong đó Q là ma trận đối xứng cấp $n \times n$. Khi đó:

i) *f là hàm lồi (t.ư., lồi chặt) trên $\mathbb{R}^n$ nếu Q là ma trận nửa xác định dương (t.ư., xác định dương);*

ii) *f là lõm (t.ư., lõm chặt) trên $\mathbb{R}^n$ nếu Q là ma trận nửa xác định âm (t.ư., xác định âm).*

4.1.6 Cực trị của hàm lồi

Mệnh đề 4.5. *Cho hàm lồi $f : \mathbb{R}^n \to \mathbb{R}$ và tập lồi khác rỗng $D \subseteq \mathbb{R}^n$. Giả sử x^* là một nghiệm tối ưu địa phương của bài toán*

$$\min\{f(x) \mid x \in D\}. \tag{P}$$

Khi đó:

i) *Ta có x^* cũng là nghiệm tối ưu toàn cục.*

ii) *Nếu x^* là nghiệm tối ưu địa phương chặt hoặc f là hàm lồi chặt thì x^* là nghiệm tối ưu toàn cục duy nhất của bài toán.*

Chứng minh.
i) Giả sử $x^* \in D$ là một nghiệm tối ưu địa phương của bài toán (P). Theo định nghĩa, tồn tại một ϵ-lân cận $B(x^*, \epsilon)$ của điểm $x^* \in D$ sao cho

$$f(x^*) \leq f(x) \quad \forall x \in B(x^*, \epsilon) \cap D.$$

Giả sử phản chứng rằng $x^* \notin \operatorname{Argmin}(P)$, tức

$$\exists \bar{x} \in D \;\text{ sao cho }\; f(\bar{x}) < f(x^*). \tag{4.2}$$

Chọn số thực $\lambda \in (0, 1)$ và λ đủ nhỏ sao cho

$$x^0 = \lambda\bar{x} + (1 - \lambda)x^* = x^* + \lambda(\bar{x} - x^*) \in B(x^*, \epsilon) \cap D.$$

Do x^* là nghiệm cực tiểu địa phương, hàm f là hàm lồi và (4.2) nên

$$f(x^*) \leq f(x^0) \leq \lambda f(\bar{x}) + (1 - \lambda)f(x^*) < f(x^*).$$

Điều mâu thuẫn này chứng tỏ $x^* \in \operatorname{Argmin}(P)$.

ii) Giả sử x^* là nghiệm tối ưu địa phương chặt. Theo i), $x^* \in \mathrm{Argmin}(P)$.

Bây giờ, ta giả thiết phản chứng rằng

$$\bar{x} \in D,\ \bar{x} \neq x^* \ \text{ và } f(\bar{x}) = f(x^*).$$

Ký hiệu

$$x_\lambda = \lambda\bar{x} + (1-\lambda)x^*, \quad 0 \le \lambda \le 1.$$

Vì D là tập lồi và f là hàm lồi trên D nên

$$x_\lambda \in D \ \text{ và } \ f(x_\lambda) \le \lambda f(\bar{x}) + (1-\lambda)f(x^*) = f(x^*) \ \text{ với mọi } \ 0 \le \lambda \le 1. \quad (4.3)$$

Cho $\lambda \to 0_+$. Ta có thể chọn được $x_\lambda \in B(x^*, \varepsilon) \cap D$ với một $\varepsilon > 0$. Điều này và (4.3) mâu thuẫn với giả thiết x^* là nghiệm tối ưu địa phương chặt. Vì vậy x^* phải là nghiệm tối ưu toàn cục duy nhất.

• Giả sử rằng x^* là nghiệm tối ưu địa phương và hàm mục tiêu f là lồi chặt. Từ i) ta có $x^*\mathrm{Argmin}(P)$.

Giả thiết phản chứng rằng

$$\bar{x} \in D,\ \bar{x} \neq x^* \text{ và } f(\bar{x}) = f(x^*).$$

Do D là tập lồi nên

$$x^0 = (\frac{1}{2}\bar{x} + \frac{1}{2}x^*) \in D$$

Do f là hàm lồi chặt nên

$$f(x^0) = f\Big(\frac{1}{2}\bar{x} + \frac{1}{2}x^*\Big) < \frac{1}{2}f(\bar{x}) + \frac{1}{2}f(x^*) = f(x^*).$$

Điều này mâu thuẫn với giả thiết $x^* \in \mathrm{Argmin}(P)$, chứng tỏ x^* là nghiệm tối ưu toàn cục duy nhất. $\square$

Mệnh đề 4.6. *Cho hàm lồi f xác định trên $\mathbb{R}^n$ và tập lồi khác rỗng $D \subseteq \mathbb{R}^n$. Nếu*

$$x^* \in \text{Argmax}\{f(x) \mid x \in D\} \quad \textit{và} x^* \in \text{ri}D$$

thì $f(x) = f(x^*)$ với mọi $x \in D$.

Chứng minh. Lấy tùy ý $x \in D$. Vì $x^* \in \text{ri}D$ nên tồn tại $y \in D$ sao cho

$$x^* = \lambda x + (1-\lambda)y, \quad 0 < \lambda < 1.$$

Giả sử rằng $f(x) \leq f(y)$. Do f là hàm lồi nên

$$f(x^*) \leq \lambda f(x) + (1-\lambda)f(y) \leq \max\{f(x), f(y)\} = f(y).$$

Kết hợp điều này với giả thiết $x^* \in \text{Argmax}\{f(x) \,|\, x \in D\}$ suy ra $f(x^*) = f(y)$. Do đó,

$$f(x^*) \leq \lambda f(x) + (1-\lambda)f(y) = \lambda f(x) + (1-\lambda)f(x^*) \Rightarrow f(x^*) \leq f(x).$$

Lập luận tương tự như trên, ta có $f(x) = f(x^*)$. Mệnh đề đã được chứng minh.
□

Nhận xét

Cho tập lồi khác rỗng $D \subseteq \mathbb{R}^n$ và hàm số $f : D \to \mathbb{R}$.

- ◇ Nếu f là hàm lõm và

  $$x^* \in \text{Argmin}\{f(x) \mid x \in D\} \quad \text{và} x^* \in \text{ri}D$$

  thì $f(x) = f(x^*)$ với mọi $x \in D$.

- ◇ Không cần xét bài toán $\max\{f(x) \mid x \in \mathbb{R}^n\}$ nếu f là hàm lồi.

- ◇ Không cần xét bài toán $\min\{f(x) \mid x \in \mathbb{R}^n\}$ nếu f là hàm lõm.

4.2 Bài toán quy hoạch phi tuyến không ràng buộc

$$\min\ f(x) \ \text{ v.đ.k. } \ x \in \mathbb{R}^n, \qquad (P^{krb})$$

trong đó $f : \mathbb{R}^n \to \mathbb{R}$ là hàm phi tuyến.

Chú ý: Không cần xét bài toán (P^{krb}) trong trường hợp f là hàm lõm.

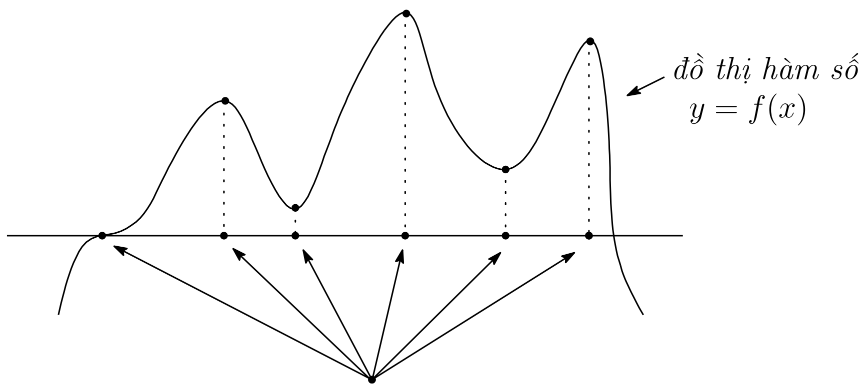
4.2.1 Điều kiện tối ưu

Định lý 4.5. (Điều kiện bậc nhất) *Cho hàm f xác định, khả vi trên $\mathbb{R}^n$. Nếu $x^* \in \mathbb{R}^n$ là nghiệm tối ưu địa phương của bài toán (P^{krb}) thì $\nabla f(x^*) = 0$.*

Chứng minh. Do f khả vi trên $\mathbb{R}^n$ và x^* là nghiệm cực tiểu địa phương nên

$$f'(x^*, d) = \lim_{t\to 0^+} \frac{f(x^* + td) - f(x^*)}{t} = \langle \nabla f(x^*), d\rangle \geq 0 \ \ \forall d \in \mathbb{R}^n.$$

Suy ra $\nabla f(x^*) = 0$. □



Chú ý.

- ◇ Nếu $\nabla f(x^*) = 0$ thì $x^* \in \mathbb{R}^n$ được gọi *điểm dừng* của hàm f.
- ◇ x^* là nghiệm tối ưu địa phương của bài toán $(P^{krb}) \quad \Rightarrow \quad \nabla f(x^*) = 0$
  x^* là nghiệm tối ưu địa phương của bài toán $(P^{krb}) \quad \not\Leftarrow \quad \nabla f(x^*) = 0$
- ◇ Nghiệm cực tiểu địa phương (t.ư., nghiệm cực tiểu hoặc nghiệm cực tiểu toàn cục) của bài toán (P^{krb}) còn được gọi là *điểm cực tiểu địa phương (t.ư., điểm cực tiểu hoặc điểm cực tiểu toàn cục) của hàm f trên* $\mathbb{R}^n$.

Định lý 4.6. *Giả sử f là hàm lồi khả vi trên $\mathbb{R}^n$. Khi đó*

$$x^* \in \text{Argmin}(P^{krb}) \quad \Leftrightarrow \quad \nabla f(x^*) = 0.$$

Chứng minh.
($\Rightarrow$) Theo Định lý 4.5, ta nhận được điều kiện cần của Định lý.

($\Leftarrow$) Giả sử $\nabla f(x^*) = 0$. Ta sẽ chứng minh x^* là nghiệm cực tiểu toàn cục của bài toán (P^{krb}). Thật vậy, do f là hàm lồi khả vi trên $\mathbb{R}^n$ nên nó có đạo hàm theo mọi hướng $d \in \mathbb{R}^n \setminus \{0\}$ tại x^* và

$$\langle \nabla f(x^*), d\rangle = f'(x^*, d) \leq f(x^* + d) - f(x^*), \quad \forall d \in \mathbb{R}^n \setminus \{0\}.$$

Lấy tùy ý $x \in \mathbb{R}^n$, $x \neq x^*$. Đặt , $d = x - x^*$. Ta có

$$0 = \langle \nabla f(x^*), x - x^*\rangle \leq f(x^* + (x - x^*)) - f(x^*) = f(x) - f(x^*).$$

Suy ra

$$f(x^*) \leq f(x) \;\; \forall x \in \mathbb{R}^n,$$

tức x^* là điểm cực tiểu của f trên $\mathbb{R}^n$. □

Nhận xét 4.3. Cho hàm toàn phương

$$f(x) = \frac{1}{2}x^T A x - b^T x + c,$$

trong đó A là ma trận cấp $n \times n$ đối xứng, nửa xác định dương, $b \in \mathbb{R}^n$ và $c \in \mathbb{R}$. Khi đó, việc giải bài toán

$$\min\{f(x) \mid x \in \mathbb{R}^n\}$$

tương đương với việc tìm nghiệm của một hệ phương trình tuyến tính

$$Ax = b.$$

Định lý 4.7. (Điều kiện bậc hai) *Giả sử hàm số f khả vi liên tục hai lần trên $\mathbb{R}^n$. Khi đó:*

i) *Nếu $x^* \in \mathbb{R}^n$ là điểm cực tiểu địa phương của f trên $\mathbb{R}^n$ thì*

$$\begin{aligned} &\nabla f(x^*) = 0 \ \ \textit{và} \\ &\nabla^2 f(x^*) \ \ \textit{nửa xác định dương}; \end{aligned}$$

ii) *Ngược lại, nếu*

$$\begin{aligned} &\nabla f(x^*) = 0 \ \ \textit{và} \\ &\nabla^2 f(x^*) \ \ \textit{xác định dương} \end{aligned}$$

thì x^* là điểm cực tiểu địa phương chặt của f trên $\mathbb{R}^n$.

Chứng minh. Tự đọc:
- Trang 210, Giáo trình in năm 2014
- Trang 223, Giáo trình in năm 2008.

Nhận xét 4.4. Vì hai bài toán $\min\{f(x) \mid x \in \mathbb{R}^n\}$ và $\max\{-f(x) \mid x \in \mathbb{R}^n\}$ tương đương với nhau theo nghĩa tập nghiệm của hai bài toán trùng nhau và giá trị tối ưu ngược dấu nên

i) Nếu $x^* \in \mathbb{R}^n$ là điểm cực đại địa phương của f trên $\mathbb{R}^n$ thì

$$\begin{aligned} &\nabla f(x^*) = 0 \ \ \text{và} \\ &\nabla^2 f(x^*) \ \ \text{nửa xác định âm}; \end{aligned}$$

ii) Ngược lại, nếu

$$\begin{aligned} &\nabla f(x^*) = 0 \ \ \text{và} \\ &\nabla^2 f(x^*) \ \ \text{xác định âm} \end{aligned}$$

thì x^* là điểm cực đại địa phương chặt của f trên $\mathbb{R}^n$.

4.2.2 Phương pháp hướng giảm

Xét bài toán quy hoạch không ràng buộc

$$\min \ f(x) \ \text{ v.đ.k. } \ x \in \mathbb{R}^n, \qquad (P^{krb})$$

trong đó $f : \mathbb{R}^n \to \mathbb{R}$ là hàm phi tuyến, khả vi trên $\mathbb{R}^n$.

Mục đích. Tìm điểm dừng của hàm f trên $\mathbb{R}^n$, tức tìm điểm $x^* \in \mathbb{R}^n$ thỏa mãn $\nabla f(x^*) = 0$.

Y tưởng: Xuất phát một điểm bất kỳ $x^0 \in \mathbb{R}^n$, ta xây dựng một dãy điểm $x^1, x^2, \ldots, x^k, \ldots$ sao cho

$$f(x^0) > f(x^1) > f(x^2) > \cdots$$

và dãy $\{x^k\}$ hội tụ đến điểm dừng $x^* \in \mathbb{R}^n$ của hàm f, tức $\nabla f(x^*) = 0$.

Chú ý. Nếu f là hàm lồi và $\nabla f(x^*) = 0$ thì $x^* \in \text{Argmin}(P^{krb})$.

a. Lược đồ chung

Bước khởi đầu. Xuất phát từ một điểm tùy ý $x^0 \in \mathbb{R}^n$, $\nabla f(x^0) \neq 0$.
Gán $k := 0$.

Bước lặp k. $(k = 0, 1, 2, \ldots)$

(k_1) Xác định $x^{k+1} := x^k + t_k d^k$ sao cho $f(x^{k+1}) < f(x^k)$.

(k_2) **If** x^{k+1} thỏa mãn điều kiện dừng **Then** STOP ($x^* = x^{k+1}$ *là điểm dừng*)
Else Gán $k := k + 1$; Quay lại Bước lặp k.

Chú ý.

◇ Thông thường, điều kiện dừng của thuật toán là

$$\nabla f(x^k) \approx 0 \ \text{ hoặc } \ \|x^k - x^{k-1}\| \text{ đủ nhỏ.}$$

◇ Trong công thức xác định x^{k+1},

$$x^{k+1} := x^k + t_k d^k \ \text{ sao cho } \ f(x^{k+1}) < f(x^k),$$

Véc tơ $d^k \in \mathbb{R}^n$ là *hướng giảm* của f tại x^k;
Số thực $t_k > 0$ là *độ dài bước.*

- ⋄ Tại điểm $x^k \in \mathbb{R}^n$, mỗi cách lựa chọn hướng dịch chuyển d^k và độ dài bước t_k khác nhau cho ta các thuật toán cụ thể tương ứng với các phương pháp hướng giảm khác nhau.

b. Hướng giảm

Định nghĩa. Cho $x^0 \in \mathbb{R}^n$. Ta gọi $d \in \mathbb{R}^n \setminus \{0\}$ là *hướng giảm* của hàm f tại x^0 nếu

$$\exists\, \varepsilon > 0 \quad \text{sao cho} \quad f(x^0 + td) < f(x^0) \ \ \forall \ \ t \in (0, \varepsilon).$$

Mệnh đề 4.7. *(Điều kiện đủ) Cho hàm $f \in C(\mathbb{R}^n)$, điểm $x^0 \in \mathbb{R}^n$ và hướng $d \in \mathbb{R}^n \setminus \{0\}$. Nếu $\langle \nabla f(x^0), d\rangle < 0$ thì d là hướng giảm của f tại x^0.*

Chứng minh. Theo giả thiết ta có

$$f'(x^0, d) = \lim_{t \to 0^+} \frac{f(x^0 + td) - f(x^0)}{t} = \langle \nabla f(x^0), d\rangle < 0.$$

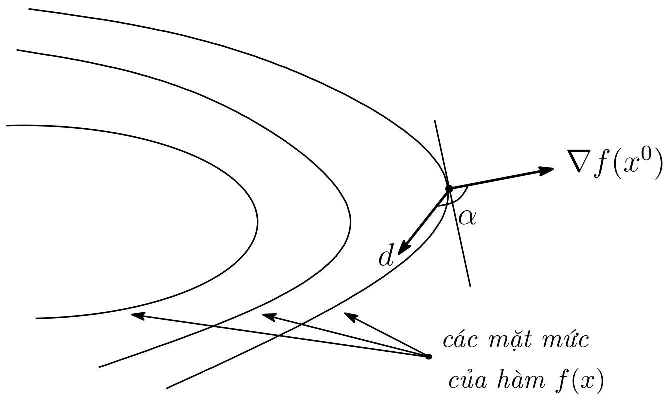
Do đó, $f(x^0 + td) - f(x^0) < 0$ với t đủ nhỏ. Mệnh đề đã được chứng minh. □

Chú ý. Nếu d là hướng giảm của hàm f tại x^0 thì chưa chắc $\langle \nabla f(x^0), d\rangle < 0$.

Ý nghĩa hình học: Cho véc tơ $d \in \mathbb{R}^n \setminus \{0\}$. Ký hiệu α là góc giữa hai véc tơ d và $\nabla f(x^0)$. Ta có

$$\langle \nabla f(x^0), d\rangle = \|\nabla f(x^0)\| \, \|d\| \cos \alpha.$$

Như vậy, nếu α là góc tù thì d là hướng giảm của hàm f tại x^0.



Hướng giảm của hàm f tại x^0

Mệnh đề 4.8. *Cho hàm lồi f khả vi trên $\mathbb{R}^n$, điểm $x^0 \in \mathbb{R}^n$ và hướng $d \in \mathbb{R}^n \setminus \{0\}$. Khi đó, $\langle \nabla f(x^0), d\rangle < 0$ khi và chỉ khi d là hướng giảm của f tại x^0.*

Chứng minh. Do đã có Mệnh đề 4.7, ta chỉ cần chứng minh điều kiện cần.

Giả sử $d \in \mathbb{R}^n \setminus \{0\}$ là hướng giảm của hàm f tại x^0, tức

$$\exists \varepsilon > 0 \text{ sao cho } f(x^0 + td) < f(x^0) \;\; \forall\, t: \quad 0 < t < \varepsilon. \tag{4.5}$$

Vì hàm f lồi khả vi trên $\mathbb{R}^n$ nên hàm f có đạo hàm theo mọi hướng d tại mọi điểm $x^0 \in \mathbb{R}^n$ và

$$f(x^0 + td) - f(x^0) \geq f'(x^0, td) = \langle \nabla f(x^0), td\rangle = t\langle \nabla f(x^0), d\rangle.$$

Kết hợp điều này và (4.5), với $0 < t < \varepsilon$ ta có

$$\langle \nabla f(x^0), d\rangle \leq \frac{f(x^0 + td) - f(x^0)}{t} < 0. \qquad \square$$

Sau đây là hệ quả trực tiếp của Mệnh đề 4.7.

Hệ quả 4.3. *Cho hàm f khả vi trên $\mathbb{R}^n$ và điểm $x^0 \in \mathbb{R}^n$. Nếu $\nabla f(x^0) \neq 0$ thì $d = -\nabla f(x^0)$ là một hướng giảm của f tại x^0.*

Nhắc lại

Mệnh đề 4.9. *Giả sử hàm f khả vi trên $\mathbb{R}^n$ và $\nabla f(x^0) \neq 0$. Trong các hướng giảm d của hàm f tại x^0 có $\|d\| = 1$ thì hàm f giảm nhanh nhất theo hướng $d = -\frac{\nabla f(x^0)}{\|\nabla f(x^0)\|}$.*

c. Xác định độ dài bước

Giả sử ta đã biết hướng giảm d^k của hàm f tại điểm x^k. Cần xác định

$$x^{k+1} := x^k + t_k d^k,$$

với t_k là một số thực dương sao cho

$$f(x^{k+1}) < f(x^k).$$

Thông thường, có hai cách lựa chọn độ dài bước t_k tương ứng với hai thủ tục khác nhau là thủ tục tìm chính xác theo tia (Exact line search) và thủ tục quay lui (Backtracking).

• Thủ tục tìm chính xác theo tia

Cho điểm $x^k \in \mathbb{R}^n$ và hướng giảm d^k của hàm f tại x^k.

Đặt

$$\varphi_k(t) := f(x^k + td^k)$$

và t_k là nghiệm tối ưu của bài toán

$$\min \quad \varphi_k(t) \quad \text{v.đ.k.} \quad t \geq 0. \tag{P_1}$$

Thủ tục này chọn *độ dài bước chính xác* $t_k > 0$ là nghiệm cực tiểu của hàm f theo tia $\{x^k + td^k,\ t \geq 0\}$.

Chú ý.

- ◇ Một số thuật toán giải bài toán (P_1) được giới thiệu ở Mục 6.2.5 (tự đọc).
- ◇ Trường hợp đơn giản, có thể dùng đánh giá $\varphi_k'(t)$ và $\varphi_k''(t)$ như đã biết để tìm cực tiểu hàm một biến.
- ◇ Đặc biệt, nếu hàm mục tiêu $f(x)$ là hàm toàn phương lồi chặt thì với mọi hướng giảm d^k của hàm f tại x^k, ta có công thức tường minh để xác định t_k.

Mệnh đề 4.10. *Cho hàm toàn phương lồi chặt*

$$f(x) = \frac{1}{2}x^T Ax - b^T x + c,$$

trong đó A *là ma trận cấp* $n \times n$*, đối xứng, xác định dương, véc tơ* $b \in \mathbb{R}^n$ *và* $c \in \mathbb{R}$*. Cho* $x^k \in \mathbb{R}^n$ *và hướng giảm* d^k *của hàm* f *tại* x^k*. Khi đó, độ dài bước chính xác* t_k *được xác định bởi*

$$t_k = -\frac{(Ax^k - b)^T d^k}{(d^k)^T Ad^k} > 0.$$

Chứng minh. Vì $f : \mathbb{R}^n \to$ là hàm lồi nên $\varphi_k(t) = f(x^k + td^k)$ là hàm lồi một biến. Nếu t_k là điểm cực tiểu của hàm $\varphi_k(t)$ thì

$$\begin{aligned}
\varphi_k^{'}(t)|_{t=t_k} = \frac{d\varphi_k(t)}{dt}|_{t=t_k} &= \lim_{t\to 0} \frac{\varphi_k(t_k + t) - \varphi_k(t_k)}{t} \\
&= \lim_{t\to 0^+} \frac{f(x^k + (t_k + t)d^k) - f(x^k + t_k d^k)}{t} \\
&= \lim_{t\to 0^+} \frac{f((x^k + t_k d^k) + td^k) - f(x^k + t_k d^k)}{t} \\
&= f'(\underbrace{x^k + t_k d^k}_{x^{k+1}}, d^k) = \langle \nabla f(x^{k+1}), d^k \rangle = 0. \qquad (4.6)
\end{aligned}$$

Vì $\nabla f(x) = Ax - b$ nên

$$\begin{aligned}
\langle \nabla f(x^{k+1}), d^k \rangle &= \langle A(x^k + t_k d^k) - b, d^k \rangle \\
&= \langle Ax^k - b, d^k \rangle + t_k \langle Ad^k, d^k \rangle = 0.
\end{aligned}$$

Do d^k là hướng giảm của hàm f tại x^k và $f(x)$ là hàm lồi nên

$$\langle \nabla f(x^k), d^k \rangle = \langle Ax^k - b, d^k \rangle < 0.$$

Hơn nữa, vì ma trận A xác định dương nên

$$\langle Ad^k, d^k \rangle = (d^k)^T Ad^k > 0 \quad \Rightarrow \quad t_k = -\frac{(Ax^k - b)^T d^k}{(d^k)^T Ad^k} > 0. \qquad \square$$

- **Thủ tục quay lui**

Trong nhiều trường hợp, việc giải bài toán cực tiểu hàm một biến để xác định độ dài bước theo thủ tục tìm chính xác theo tia cũng không dễ dàng và chi phí tính toán cao. Trong thực tế tính toán, người ta thường sử dụng thủ tục quay lui.

Mệnh đề sau là cơ sở của thủ tục quay lui xác định điểm x^{k+1} khi đã biết hướng giảm d^k của hàm f tại x^k.

Mệnh đề 4.11. *Cho hàm f khả vi trên $\mathbb{R}^n$, điểm $x^k \in \mathbb{R}^n$ và véc tơ $d^k \in \mathbb{R}^n \setminus \{0\}$ thỏa mãn $\langle \nabla f(x^k), d^k \rangle < 0$. Cho số thực $m_1 \in (0,1)$. Khi đó*

$$\exists t_0 > 0 \text{ sao cho } f(x^k + td^k) \le f(x^k) + m_1 t \langle \nabla f(x^k), d^k \rangle \ \ \forall t \in (0, t_0].$$

Chứng minh. Vì $\langle \nabla f(x^k), d^k \rangle < 0$ nên d^k là hướng giảm của f tại x^k.

Theo định nghĩa và vì f khả vi nên

$$f'(x^k, d^k) = \lim_{t \to 0^+} \frac{f(x^k + td^k) - f(x^k)}{t} = \langle \nabla f(x^k), d^k \rangle.$$

Do đó

$$\lim_{t \to 0^+} \frac{f(x^k + td^k) - f(x^k)}{t \langle \nabla f(x^k), d^k \rangle} = 1 > m_1.$$

Suy ra, tồn tại số thực $t_0 > 0$ đủ nhỏ sao cho với mọi $t \in (0, t_0]$ ta có

$$\frac{f(x^k + td^k) - f(x^k)}{t \langle \nabla f(x^k), d^k \rangle} \ge m_1.$$

Kết hợp điều này với giả thiết $\langle \nabla f(x^k), d^k \rangle < 0$, ta có điều phải chứng minh. □

Điều kiện

$$f(x^k + t_k d^k) \le f(x^k) + m_1 t_k \langle \nabla f(x^k), d^k \rangle \text{ với } m_1 \in (0,1) \text{ và } t_k > 0$$

được gọi là *điều kiện Armijo*. Mệnh đề 6.11 đã chỉ ra rằng điều kiện này được thỏa mãn với t_k đủ nhỏ.

Thủ tục quay lui sau đây cho phép xác định điểm lặp tiếp theo x^{k+1} sao cho điều kiện Armijo thỏa mãn nhưng độ dài bước t_k không quá bé.

Thủ tục quay lui (*Quy tắc Armijo*)

- ◇ *Đầu vào:* điểm $x^k \in \mathbb{R}^n$ và hướng giảm d^k của hàm f tại x^k.
- ◇ *Đầu ra:* điểm x^{k+1} trên tia $\{x^k + td^k \mid t > 0\}$ thỏa mãn $f(x^{k+1}) < f(x^k)$.

Bước 1. Tùy chọn $m_1 \in (0,1)$ và $\alpha \in (0,1)$ (chẳng hạn, $\alpha = \frac{1}{2}$). Đặt $t_k := 1$.

Bước 2. Tính $x^{k+1} := x^k + t_k d^k$ và $f(x^{k+1})$.

Bước 3. **If** $f(x^{k+1}) \le f(x^k) + m_1 t_k \langle \nabla f(x^k), d^k \rangle$
Then Dừng thủ tục (*ta có* x^{k+1})
Else $t_k := \alpha t_k$ và quay về Bước 2.

Nhận xét rằng, trong thủ tục quay lui, để xác định điểm tiếp theo, ta không cần phải giải bài toán cực tiểu hàm một biến.

d. Tốc độ hội tụ

Với các thuật toán sử dụng chiến lược xây dựng một dãy điểm tiến dần đến nghiệm, người ta thường phải trả lời hai câu hỏi:

– Thuật toán có hội tụ không? (tức dãy điểm do thuật toán sinh ra có hội tụ đến nghiệm cần tìm không?)

– Nếu thuật toán hội tụ thì hội tụ nhanh hay chậm thế nào?

Định nghĩa. Cho dãy $\{x^k\} \subset \mathbb{R}^n$ hội tụ đến $x^* \in \mathbb{R}^n$. Dãy $\{x^k\}$ được gọi là:

– Hội tụ đến x^* với *tốc độ tuyến tính* (linear) nếu

$$\exists 0 \leq \gamma < 1,\ \exists k_0 \ \text{ sao cho } \ \forall k > k_0 :\ \|x^{k+1} - x^*\| \leq \gamma \|x^k - x^*\|;$$

– Hội tụ đến x^* với *tốc độ trên tuyến tính* (super linear) nếu

$$\forall k :\ \ \|x^{k+1} - x^*\| \leq c_k \|x^k - x^*\| \ \text{ và } \ c_k \to 0;$$

– Hội tụ đến x^* với *tốc độ hội tụ bậc hai* (quadratic) nếu

$$\exists \gamma > 0,\ \exists k_0 \text{ sao cho } \ \forall k > k_0 : \|x^{k+1} - x^*\| \leq \gamma \|x^k - x^*\|^2.$$

Nhận xét 4.5. Dễ thấy rằng, nếu dãy $\{x^k\}$ hội tụ đến x^* với tốc độ trên tuyến tính thì nó hội tụ đến x^* với tốc độ tuyến tính nhưng điều ngược lại không đúng (Bài tập). Tương tự, nếu dãy $\{x^k\}$ hội tụ đến x^* với tốc độ bậc hai thì nó hội tụ đến x^* với tốc độ trên tuyến tính vì

$$\|x^{k+1} - x^*\| \leq \gamma \|x^k - x^*\| \|x^k - x^*\| = c_k \|x^k - x^*\|,$$

trong đó $c_k = \gamma \|x^k - x^*\| \to 0$. Tuy nhiên, điều ngược lại không đúng. Ví dụ, cho dãy $\{x^k\} \to x^*$ và

$$\|x^{k+1} - x^*\| \leq \gamma \|x^k - x^*\|^{\frac{3}{2}} = \gamma \|x^k - x^*\|^{\frac{1}{2}} \|x^k - x^*\|.$$

Vì $c_k = \gamma \|x^k - x^*\|^{\frac{1}{2}} \to 0$ nên dãy $\{x^k\}$ hội tụ đến x^* với tốc độ trên tuyến tính nhưng rõ ràng đó không phải là hội tụ với tốc độ bậc hai.

4.2.3 Phương pháp gradient

a. Thuật toán gradient với thủ tục tìm chính xác theo tia

Trong thuật toán này, tại mỗi bước lặp k, điểm lặp tiếp theo được xác định bởi

$$x^{k+1} := x^k - t_k \nabla f(x^k),$$

trong đó t_k là nghiệm cực tiểu của hàm một biến $\varphi_k(t) := f(x^k - t\nabla f(x^k))$ với $t > 0$.

Thuật toán 4.1. (*Thuật toán gradient với thủ tục tìm chính xác theo tia*)

Bước khởi đầu. Chọn trước số $\varepsilon > 0$ đủ nhỏ. Xuất phát từ một điểm tùy ý $x^0 \in \mathbb{R}^n$ có $\nabla f(x^0) \neq 0$. Đặt $k := 0$.
Bước lặp k. $(k = 0, 1, 2, \dots)$

(k_1) Tính $x^{k+1} := x^k - t_k \nabla f(x^k)$, trong đó

$$t_k = \operatorname{argmin}\{\varphi_k(t),\ t > 0\}.$$

(k_2) Tính $\nabla f(x^{k+1})$.

(k_3) **If** $\|\nabla f(x^{k+1})\| < \varepsilon$ **Then** Dừng thuật toán (*lấy điểm dừng* $x^* \approx x^{k+1}$)

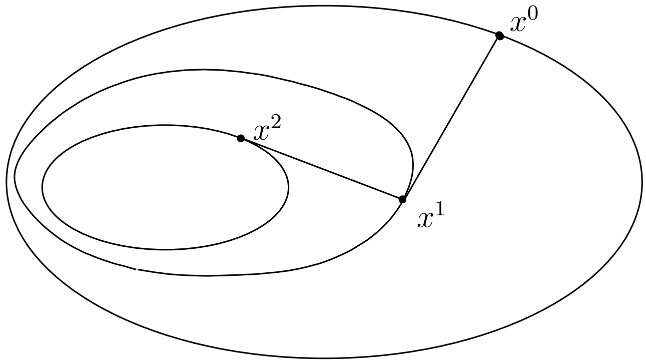
Else $k := k + 1$ và quay lại Bước lặp k.

Chú ý 4.2. Theo Mệnh đề 6.10, nếu hàm mục tiêu của bài toán (P^{krb}) là hàm toàn phương lồi chặt

$$f(x) = \frac{1}{2} x^T A x - b^T x + c$$

thì ta có công thức tính độ dài bước chính xác t_k tại mỗi Bước lặp k là

$$t_k = \frac{(Ax^k - b)^T \nabla f(x^k)}{(\nabla f(x^k))^T A \nabla f(x^k)} > 0.$$



Minh họa hình học Thuật toán 4.1

Ý nghĩa hình học. Trong mỗi bước lặp k, độ dài bước t_k là nghiệm cực tiểu của hàm một biến

$$\begin{aligned} \varphi_k : \mathbb{R}_+ &\to \mathbb{R} \\ t &\mapsto \varphi_k(t) := f(x^k - t\nabla f(x^k)). \end{aligned}$$

ta có

$$\varphi_k^{'}(t)|_{t=t_k} \overset{(6.6)}{=} f'(\underbrace{x^k - t_k\nabla f(x^k)}_{x^{k+1}}, -\nabla f(x^k)) = \langle \nabla f(x^{k+1}), -\nabla f(x^k)\rangle = 0,$$

tức véc tơ $\nabla f(x^{k+1})$ và véc tơ $d^k = -\nabla f(x^k)$ là trực giao với nhau.

Vì $f : \mathbb{R}^n \to$ là hàm lồi nên $\varphi_k(t) = f(x^k + td^k)$ là hàm lồi một biến, trong đó $d^k = -\nabla f(x^k)$. Theo điều kiện cần tối ưu, nếu t_k là điểm cực tiểu của hàm $\varphi_k(t)$ thì

$$\begin{aligned} \varphi_k^{'}(t)|_{t=t_k} = \frac{d\varphi_k(t)}{dt}|_{t=t_k} &= \lim_{t\to 0} \frac{\varphi_k(t_k+t) - \varphi_k(t_k)}{t} \\ &= \lim_{t\to 0^+} \frac{f(x^k + (t_k+t)d^k) - f(x^k + t_kd^k)}{t} \\ &= \lim_{t\to 0^+} \frac{f((x^k + t_kd^k) + td^k) - f(x^k + t_kd^k)}{t} \\ &= f'(\underbrace{x^k + t_kd^k}_{x^{k+1}}, d^k) = \langle \nabla f(x^{k+1}), d^k\rangle = 0. \end{aligned}$$

Thay $d^k = -\nabla f(x^k)$, ta có

$$\langle \nabla f(x^{k+1}), -\nabla f(x^k)\rangle = 0 \quad \Rightarrow \quad \nabla f(x^{k+1}) \perp \nabla f(x^k).$$

Nói cách khác, xuất phát từ x^k, đi theo hướng $d^k = -\nabla f(x^k)$ đến một điểm nằm trên một đường mức nào đó của hàm f mà nhận d^k là tiếp tuyến thì dừng. Đó chính là điểm x^{k+1} cần xác định trong Bước lặp k.

Ví dụ 4.1. Giải bài toán $\min\{f(x) \mid x \in \mathbb{R}^3\}$, trong đó

$$f(x) = \frac{1}{2}x^T Ax - b^T x$$

với

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 25 \end{pmatrix} \quad \text{và} \quad b = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}.$$

Do A là ma trận đối xứng xác định dương nên $f(x)$ là hàm lồi chặt. Vì vậy, theo Nhận xét 6.3 và Mệnh đề 6.5, nghiệm cực tiểu duy nhất của bài toán này là

$$x^* = A^{-1}b = \begin{pmatrix} -1.00 \\ -0.20 \\ -0.04 \end{pmatrix}.$$

Bây giờ, ta sẽ tiến hành giải bài toán này theo Thuật toán 6.1, xuất phát từ điểm $x^0 = (0,\ 0,\ 0)^T$ và chọn $\varepsilon = 10^{-8}$. Tại mỗi điểm x^k, ta có $\nabla f(x^k) = Ax^k - b$. Theo Chú ý 6.2, độ dài bước chính xác t_k tại mỗi Bước lặp k được xác định bởi

$$t_k = \frac{(Ax^k - b)^T \nabla f(x^k)}{(\nabla f(x^k))^T A \nabla f(x^k)} > 0.$$

Tại x^0 ta có

$$f(x^0) = 0, \quad \nabla f(x^0) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \|\nabla f(x^0)\| = 1.732051.$$

Do đó $t_0 = 0.096774$ và

$$x^1 = \begin{pmatrix} -0.096774 \\ -0.096774 \\ -0.096774 \end{pmatrix}.$$

Tại x^1 ta có

$$f(x^1) = -0.145161, \quad \nabla f(x^1) = \begin{pmatrix} 0.903226 \\ 0.516129 \\ -1.419355 \end{pmatrix}, \quad \|\nabla f(x^1)\| = 1.759765.$$

Do đó $t_1 = 0.058973$ và

$$x^2 = \begin{pmatrix} -0.150040 \\ -0.127212 \\ -0.013071 \end{pmatrix}.$$

Tại x^2 ta có

$$f(x^2) = -0.236474, \quad \nabla f(x^2) = \begin{pmatrix} 0.849960 \\ 0.363941 \\ 0.673226 \end{pmatrix}, \quad \|\nabla f(x^2)\| = 1.143730.$$

Sau 217 bước lặp, ta sẽ nhận được

$$x^{217} = \begin{pmatrix} -0.999999994 \\ -0.200000000 \\ -0.040000000 \end{pmatrix} \quad \text{với} \quad \|\nabla f(x^{217})\| < 10^{-8}.$$

Dừng thuật toán với điểm dừng $x^* \approx x^{217}$.

Định lý 4.8. (Định lý hội tụ) ([27], trang 41) *Cho $x^0 \in \mathbb{R}^n$ và hàm f khả vi liên tục trên $\mathbb{R}^n$ và có tập mức dưới $\{x \in \mathbb{R}^n \mid f(x) \leq f(x^0)\}$ bị chặn. Khi đó mỗi điểm tụ x^* của dãy $\{x^k\}$ được chọn như trong Thuật toán 6.1 thỏa mãn $\nabla f(x^*) = 0$.*

b. Thuật toán gradient với thủ tục quay lui

Trong thuật toán này, tại mỗi bước lặp k, chọn hướng giảm $d^k = -\nabla f(x^k)$ và độ dài bước t_k được xác định theo thủ tục quay lui.

Thuật toán 6.2. (*Thuật toán gradient với thủ tục quay lui*)

Bước khởi đầu. Tùy chọn $m_1 \in (0, 1)$ và $\alpha \in (0, 1)$. Chọn số thực $\varepsilon > 0$ đủ nhỏ. Xuất phát từ một điểm tùy ý $x^0 \in \mathbb{R}^n$ có $\nabla f(x^0) \neq 0$. Đặt $k := 0$.

Bước lặp k. $(k = 0, 1, 2, \dots)$

(k_1) Đặt $t_k := 1$.

(k_2) Tính $x^{k+1} := x^k - t_k \nabla f(x^k)$ và $f(x^{k+1})$.

(k_3) **If** $f(x^{k+1}) - f(x^k) \leq m_1 t_k \langle \nabla f(x^k), -\nabla f(x^k) \rangle = -m_1 t_k \|\nabla f(x^k)\|^2$
Then Chuyển Bước k_4
Else $t_k := \alpha t_k$ và quay về Bước k_2.

(k_4) Tính $\nabla f(x^{k+1})$.

(k_5) **If** $\|\nabla f(x^{k+1})\| < \varepsilon$ **Then** Dừng thuật toán (*lấy điểm dừng* $x^* \approx x^{k+1}$)
Else $k := k + 1$, quay về Bước lặp k.

Ví dụ 4.2. Xét bài toán (P^{krb}) như ở Ví dụ 6.5, tức có

$$f(x_1, x_2) = x_1^3 + x_2^2 - 3x_1 - 2x_2 + 12.$$

Quá trình giải bài toán này theo Thuật toán 6.2 như sau: Chọn $m_1 = 0.01$, $\alpha = 0.5$. Cũng xuất phát từ điểm $x^0 = (1, 2)^T$ và ta có

$$\nabla f(x^0) = (0, 2)^T \neq (0, 0)^T.$$

Đặt $t_0 = 1$. Suy ra

$$x^1 = x^0 - t_0 \nabla f(x^0) = \begin{pmatrix} 1 \\ 2 \end{pmatrix} - \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad f(x^0) = f(x^1) = 10.$$

Vì

$$f(x^1) - f(x^0) = 10 - 10 = 0 > -m_1 t_0 \|\nabla f(x^0)\|^2 = -0.04$$

nên đặt

$$t_0 = \alpha . t_0 = 0.5.$$

Tính lại

$$x^1 = x^0 - t_0 \nabla f(x^0) = \begin{pmatrix} 1 \\ 2 \end{pmatrix} - 0.5 \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

và

$$f(x^1) - f(x^0) = 9 - 10 = -1 < -m_1 t_0 \|\nabla f(x^0)\|^2 = -0.02.$$

Do đó ta chọn $x^1 = (1,\ 1)^T$. Vì $\nabla f(x^1) = (0, 0)^T$ nên dừng thuật toán với điểm dừng $x^1 = (1, 1)^T$.

Định lý 4.9. (Định lý hội tụ) *Giả sử hàm $f(x)$ bị chặn dưới, gradient $\nabla f(x)$ thỏa mãn điều kiện Lipschitz*[1]*, tức tồn tại $L > 0$ sao cho*

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|, \ \ \forall x, y \in \mathbb{R}^n.$$

Khi đó, với bất kỳ điểm xuất phát x^0, dãy $\{x^k\}$ được chọn như trong Thuật toán 6.2 có tính chất $\|\nabla f(x^k)\| \to 0$ khi $k \to \infty$.

Chứng minh. Xem [37], trang 375−376. □

[1]Rudolph Otto Sigismund LIPSCHITZ (1832 − 1903): Nhà toán học Đức. Chuyên ngành nghiên cứu của ông là Giải tích và Hình học.

4.2.4 Phương pháp Newton

Xét bài toán

$$\min\ f(x)\ \text{ v.đ.k. }\ x \in \mathbb{R}^n, \tag{P_{krb}}$$

trong đó $f : \mathbb{R}^n \to \mathbb{R}$ là hàm khả vi cấp 2. Giả sử, tại mỗi $x \in \mathbb{R}^n$, ta xác định được $f(x)$, $\nabla f(x)$ và $\nabla^2 f(x)$.

Xấp xỉ Taylor bậc hai tại lân cận $x^k \in \mathbb{R}^n$ là

$$f(x^k + p) \approx f(x^k) + \nabla f(x^k), p\rangle + \frac{1}{2} p^T \nabla^2 f(x^k) p.$$

Ký hiệu

$$\varphi(p) = \frac{1}{2} p^T \nabla^2 f(x^k) p + \nabla f(x^k), p\rangle + f(x^k).$$

Bài toán: Tìm hướng $p^k \in \mathbb{R}^n$ là nghiệm tối ưu của bài toán sau:

$$\min\ \varphi(p)\ \text{ v.đ.k. }\ p \in \mathbb{R}^n, \tag{P_0}$$

Giả thiết: Ma trận $\nabla^2 f(x^k)$ xác định dương.

Khi đó, $\varphi(p)$ là hàm lồi chặt trên $\mathbb{R}^n$ và bài toán (P_0) có nghiệm tối ưu duy nhất là p^k thỏa mãn

$$\nabla \varphi(p^k) = \nabla^2 f(x^k) p^k + \nabla f(x^k) = 0 \quad \Rightarrow \quad p^k = -[\nabla^2 f(x^k)]^{-1} \nabla f(x^k).$$

Định nghĩa. Véc tơ $p^k = -[\nabla^2 f(x^k)]^{-1} \nabla f(x^k)$ được gọi là *hướng Newton của hàm f tại điểm x^k.*

Mệnh đề 4.12. *Nếu ma trận Hesse $\nabla^2 f(x^k)$ xác định dương thì hướng Newton p^k của hàm f tại x^k cũng là hướng giảm của hàm f tại x^k.*

Chứng minh. Thật vậy, vì $\nabla^2 f(x^k)$ xác định dương nên ma trận $[\nabla^2 f(x^k)]^{-1}$ cũng xác định dương và

$$\begin{aligned}\langle \nabla f(x^k), p^k\rangle &= \Big[\nabla f(x^k)\Big]^T . p^k \\ &= -[\nabla f(x^k)]^T [\nabla^2 f(x^k)]^{-1} \nabla f(x^k) \quad < 0.\end{aligned}$$

Theo Mệnh đề 6.7, p^k là một hướng giảm của hàm f tại x^k. □

Nội dung Mệnh đề 6.12 cho ta biết lý do của giả thiết $\nabla^2 f(x)$ xác định dương với mỗi $x \in \mathbb{R}^n$ của bài toán (P^{krb}) đang xét. Trong thuật toán Newton thuần túy giải bài toán này, tại mỗi bước lặp k, điểm x^{k+1} được xác định theo công thức

$$x^{k+1} := x^k + p^k,$$

trong đó p^k là hướng Newton của hàm f tại x^k. Như vậy, độ dài bước $t_k = 1$ tại mọi bước lặp k.

Thuật toán 6.3. (*Thuật toán Newton thuần túy giải bài toán* (P^{krb}))

Bước khởi đầu. Xuất phát từ một điểm tùy ý $x^0 \in \mathbb{R}^n$ đủ gần điểm dừng x^* và $\nabla f(x^0) \neq 0$; Chọn trước số $\varepsilon > 0$ đủ nhỏ. Đặt $k := 0$.

Bước lặp k. $(k = 0, 1, 2, \dots)$

(k_1) Tính hướng Newton p^k của hàm f tại x^k bằng việc giải hệ phương trình tuyến tính
$$[\nabla^2 f(x^k)].p^k = -\nabla f(x^k).$$

(k_2) Xác định $x^{k+1} := x^k + p^k$ và $\nabla f(x^{k+1})$.

(k_3) **If** $\|\nabla f(x^{k+1})\| < \varepsilon$ **Then** Dừng thuật toán (*lấy điểm dừng* $x^* \approx x^{k+1}$)
Else $k := k + 1$ và quay lại Bước lặp k.

Chú ý 4.3. Điểm xuất phát x^0 trong Thuật toán 6.3 phải đảm bảo "**đủ gần**" điểm dừng x^* của hàm f (xem Ví dụ 6.11).

Định lý 4.10. (Định lý hội tụ) *Giả sử:*

i) *Hàm f khả vi hai lần trên $\mathbb{R}^n$;*

ii) *Hàm $\nabla^2 f(x)$ là liên tục và Lipschitz trong lân cận của điểm dừng x^* của hàm f, tức tồn tại lân cận $B(x^*, \varepsilon)$ và số $L > 0$ sao cho*
$$\|\nabla^2 f(x) - \nabla^2 f(y)\| \leq L\|x - y\| \quad \forall\, x, y \in B(x^*, \varepsilon);$$

iii) *Ma trận $\nabla^2 f(x)$ xác định dương tại mọi $x \in \mathbb{R}^n$.*

Khi đó, nếu xuất phát từ một điểm đủ gần x^* thì dãy $\{x^k\}$ sinh ra bởi Thuật toán Newton thuần túy hội tụ tới x^* theo tốc độ cấp hai.

Chứng minh. Xem [36], trang 38 − 39. □

Ví dụ 4.3. Sử dụng thuật toán Newton thuần túy giải bài toán
$$\min\{f(x) = x^4 - 1 \mid x \in \mathbb{R}\}$$
với điểm xuất phát $x^0 = 4$. Tính ba điểm tiếp theo x^1, x^2 và x^3. Chứng minh rằng, dãy $\{x^k\}$ sinh bởi thuật toán này sẽ hội tụ đến nghiệm cực tiểu của bài toán đang xét.

Giải. Dễ thấy bài toán có nghiệm cực tiểu duy nhất là $x^* = 0$. Ta có $f'(x) = 4x^3$ và $f''(x) = 12x^2 > 0$ với mọi $x \in \mathbb{R} \setminus \{0\}$. Do đó, hướng Newton p^k của hàm f tại một điểm x^k cũng là hướng giảm của hàm f tại x^k và

$$p^k = -\frac{f'(x^k)}{f''(x^k)} = -\frac{1}{3}x^k.$$

Giả sử biết x^k, theo thuật toán Newton thuần túy, điểm tiếp theo x^{k+1} được xác định bởi

$$x^{k+1} := x^k + p^k = x^k - \frac{1}{3}x^k = \frac{2}{3}x^k.$$

Xuất phát từ $x^0 = 4$, ta có

$$x^1 = \frac{2}{3}x^0 = \frac{8}{3}; \quad x^2 = \frac{2}{3}x^1 = \left(\frac{2}{3}\right)^2 x^0 = \frac{16}{9}; \quad x^3 = \frac{2}{3}x^2 = \left(\frac{2}{3}\right)^3 x^0 = \frac{32}{27}.$$

Do đó

$$x^k = \left(\frac{2}{3}\right)^k x^0 = 4\left(\frac{2}{3}\right)^k \quad \Rightarrow \quad \{x^k\} \overset{k\to\infty}{\longrightarrow} 0 = x^*.$$

- **Các ưu, nhược điểm của Thuật toán Newton thuần túy**

◇ **Ưu điểm**:

i) Nếu hàm mục tiêu f là hàm toàn phương

$$f(x) = \frac{1}{2}x^T Ax - b^T x + c,$$

trong đó A là ma trận cấp $n \times n$, đối xứng, xác định dương, không suy biến, thì Thuật toán 6.3 cho ta ngay một nghiệm cực tiểu toàn cục của bài toán (P^{krb}) chỉ sau một vòng lặp, không phụ thuộc vào điểm xuất phát ban đầu. Thật vậy, trong trường hợp này ta có $\nabla^2 f(x) = A$. Theo Nhận xét 6.3, điểm cực tiểu toàn cục x^* của f trên $\mathbb{R}^n$ là nghiệm của hệ phương trình $Ax = b$. Do đó $x^* = A^{-1}b$. Bây giờ, xuất phát từ một điểm bất kỳ x^0, theo Thuật toán 6.3, ta đến được $x^1 = x^0 - A^{-1}(Ax^0 - b) = A^{-1}b = x^*$.

ii) Nếu xuất phát từ một điểm x^0 đủ gần điểm dừng và ma trận $\nabla^2 f(x^k)$ không suy biến tại mọi bước lặp k thì thuật toán hội tụ rất nhanh (hội tụ cấp hai).

◇ **Nhược điểm**:

i) Tại điểm x^k, nếu ma trận $\nabla^2 f(x^k)$ suy biến thì điểm lặp tiếp theo không xác định được;

ii) Tại điểm x^k, nếu ma trận $\nabla^2 f(x^k)$ không suy biến nhưng $\nabla^2 f(x^k)$ không xác định dương thì hướng Newton p^k tương ứng không phải hướng giảm của hàm f tại x^k. Thuật toán có thể hội tụ đến điểm dừng $x^* \in \mathbb{R}^n$ nhưng đó không phải là điểm cực tiểu địa phương;

iii) Tính toán và lưu trữ các ma trận cấp $n \times n$ rất tốn kém khi n lớn.

c. Phương pháp Newton suy rộng

Trong thuật toán Newton thuần túy (Thuật toán 6.3), việc phải đảm bảo yêu cầu xuất phát từ một điểm x^0 đủ gần điểm dừng x^* gây khó khăn cho việc thực hiện thuật toán. Khắc phục hạn chế này, phương pháp Newton suy rộng cho phép xuất phát từ một điểm tùy ý $x^0 \in \mathbb{R}^n$. Tại mỗi Bước lặp k, điểm x^{k+1} được xác định bởi

$$x^{k+1} := x^k + t_k d^k,$$

trong đó d^k vẫn là hướng Newton của hàm f tại x^k, tức $d^k = -[\nabla^2 f(x^k)]^{-1}\nabla f(x^k)$ và độ dài bước t_k được tính theo thủ tục quay lui.

Thuật toán 6.4. (*Phương pháp Newton suy rộng*)

Bước khởi đầu. Xuất phát từ một điểm $x^0 \in \mathbb{R}^n$ có $\nabla f(x^0) \neq 0$; Chọn trước số $\varepsilon > 0$ đủ nhỏ. Đặt $k := 0$.

Bước lặp k. $(k = 0, 1, 2 \dots)$

(k_1) Tính hướng Newton d^k bằng việc giải hệ phương trình tuyến tính

$$[\nabla^2 f(x^k)]d^k = -\nabla f(x^k).$$

(k_2) Tính $x^{k+1} := x^k + t_k d^k$ theo thủ tục quay lui. Tính $\nabla f(x^{k+1})$.

(k_3) **If** $\|\nabla f(x^{k+1})\| < \varepsilon$ **Then** Dừng thuật toán (*lấy* $x^* \approx x^{k+1}$)
Else gán $k := k + 1$ và quay về Bước lặp k.

Định lý 4.11. (Định lý hội tụ) *Giả sử hàm $f(x)$ khả vi hai lần trên $\mathbb{R}^n$ và f là hàm lồi mạnh với hệ số lồi mạnh $m > 0$ trên tập mức dưới*

$$L\Big(f(x^0), f\Big) = \{x \in \mathbb{R}^n \mid f(x) \leq f(x^0)\},$$

tức là với mọi $x, y \in L\Big(f(x^0), f\Big)$, $\lambda \in (0, 1)$, ta có

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) - \tfrac{1}{2}m\lambda(1-\lambda)\|y - x\|^2,$$

trong đó x^0 là điểm xuất phát của dãy $\{x^k\}$ sinh bởi Thuật toán Newton suy rộng (Thuật toán 6.4). Gọi x^* là điểm cực tiểu duy nhất của f và giả sử rằng

$\nabla^2 f(x)$ liên tục Lipschitz trên $L\Big(f(x^0), f\Big)$ với hằng số Lipschitz $L > 0$. Khi đó:

i) Tồn tại k sao cho với mọi $\ell \geq k$ ta có

$$\|\nabla f(x^\ell)\| < \eta_1 \quad \text{với} \quad \eta_1 \leq \tfrac{m^2}{L};$$

ii) Dãy $\{f(x^\ell)\}$ hội tụ đến $f(x^*)$ với tốc độ bậc hai.

Chứng minh. Xem [36], trang 42−43. □

d. Phương pháp tựa Newton

Xét bài toán (P^{krb}) với giả thiết hàm mục tiêu f khả vi hai lần trên $\mathbb{R}^n$. Theo phương pháp Newton, ở mỗi bước lặp k, ta cần tính hướng Newton

$$p^k = -[\nabla^2 f(x^k)]^{-1}\nabla f(x^k).$$

Đây là công việc khó khăn vì ta phải tính ma trận nghịch đảo $[\nabla^2 f(x)]^{-1}$. Hơn nữa, nếu ma trận $\nabla^2 f(x^k)$ không suy biến nhưng không xác định dương thì hướng Newton d^k tương ứng không phải hướng giảm của hàm f tại x^k. Vì vậy phương pháp Newton ít được sử dụng trong thực tế khi $n > 1$ mặc dù phương pháp này có tốc độ hội tụ bậc hai.

Chiến lược của phương pháp tựa Newton (Quasi Newton Methods) là thay vì hướng Newton p^k tại mỗi bước lặp, ta tính hướng

$$d^k = -H_k\nabla f(x^k),$$

trong đó H_k là ma trận không suy biến, đối xứng, xác định dương. Ma trận H_{k+1} sẽ được tính truy hồi theo H_k, $\nabla f(x^k)$ và $\nabla f(x^{k+1})$ (xem công thức (6.11) dưới đây). Ma trận H_{k+1} cũng là ma trận đối xứng, xác định dương (xem [27], trang 47). Tuy phương pháp tựa Newton có tốc độ hội tụ không còn là bậc hai nữa và phải thực hiện nhiều bước lặp hơn phương pháp Newton nhưng tính toán tại mỗi bước lặp đơn giản hơn. Thuật toán 6.5 mang tên Davidon-Fletcher-Powell[2] trình bày sau đây là một thuật toán tiêu biểu cho phương pháp này, có tốc độ hội tụ trên tuyến tính.

Thuật toán 6.5. (Thuật toán D.F.P.)

Bước khởi đầu.

- Xuất phát tùy ý từ một điểm $x^1 \in \mathbb{R}^n$ có $\nabla f(x^1) \neq 0$.
- Chọn tùy ý ma trận đối xứng xác định dương H_1 (người ta thường chọn H_1 là ma trận đơn vị I).
- Đặt $k := 1$.

Bước lặp k. ($k = 1, 2, \dots$)

(k_1) Đặt $d^k := -H_k\nabla f(x^k)$.

(k_2) Tính $t_k := \text{argmin}\{\varphi(t) = f(x^k + td^k),\ t \geq 0\}$.

[2] Thuật toán này do Davidon đề xuất đầu tiên vào năm 1959, sau đó được Fletcher và Powell phát triển (1963).

(k_3) Tính $x^{k+1} := x^k + t_k d^k$; $\quad v^k := x^{k+1} - x^k$ và $\nabla f(x^{k+1})$.

(k_4) **If** $\|\nabla f(x^{k+1})\| \approx 0$ **Then** Dừng thuật toán $\quad$ (*lấy* $x^* \approx x^{k+1}$)
Else Chuyển Bước k_5

(k_5) Tính $u^k := \nabla f(x^{k+1}) - \nabla f(x^k)$ và

$$H_{k+1} := H_k + \frac{v^k (v^k)^T}{\langle u^k, v^k \rangle} - \frac{(H_k u^k)(H_k u^k)^T}{\langle u^k, H_k u^k \rangle}. \tag{6.11}$$

(k_6) Đặt $k := k + 1$ và quay về Bước lặp k.

Nhắc lại rằng, $v^k \in \mathbb{R}^n$ là véc tơ cột. Do đó $v^k (v^k)^T$ là ma trận cấp $n \times n$. Cụ thể

$$\begin{pmatrix} v_1^k \\ v_2^k \\ \vdots \\ v_n^k \end{pmatrix} \begin{pmatrix} v_1^k & v_2^k & \dots & v_n^k \end{pmatrix} = \begin{pmatrix} v_1^k v_1^k & v_1^k v_2^k & \dots & v_1^k v_n^k \\ v_2^k v_1^k & v_2^k v_2^k & \dots & v_2^k v_n^k \\ \vdots & \vdots & & \vdots \\ v_n^k v_1^k & v_n^k v_2^k & \dots & v_n^k v_n^k \end{pmatrix}.$$

Tương tự, $(H_k u^k)(H_k u^k)^T$ là ma trận cấp $n \times n$.

Ví dụ 4.4. Giải lại bài toán đã xét ở Ví dụ 6.9, $\min\{f(x) \mid x \in \mathbb{R}^3\}$, trong đó $f(x) = \frac{1}{2} x^T A x - b^T x$ và

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 25 \end{pmatrix} \quad \text{và} \quad b = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}.$$

Ta có $\nabla f(x) = Ax - b$ và kết quả giải bài toán này theo Thuật toán 6.5 như sau:

Bước lặp $k = 1$.

$$x^1 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \nabla f(x^1) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad H_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$d^1 = -H_1 \nabla f(x^1) = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} \quad t_1 = 0.096774$$

$$x^2 = x^1 + t_1 d^1 = \begin{pmatrix} -0.096774 \\ -0.096774 \\ -0.096774 \end{pmatrix}$$

$$v^1 = x^2 - x^1 = \begin{pmatrix} -0.096774 \\ -0.096774 \\ -0.096774 \end{pmatrix} \quad \nabla f(x^2) = \begin{pmatrix} 0.903226 \\ 0.516129 \\ -1.419355 \end{pmatrix}$$

$$u^1 = \nabla f(x^2) - \nabla f(x^1) = \begin{pmatrix} -0.096774 \\ -0.483871 \\ -2.419355 \end{pmatrix}$$

$$H_2 = \begin{pmatrix} 1.030722 & 0.024578 & -0.006144 \\ 0.024578 & 0.993856 & -0.159754 \\ -0.006144 & -0.159754 & 0.072197 \end{pmatrix}$$

Bước lặp $k = 2$.

$$d^2 = -H_2 \nabla f(x^2) = \begin{pmatrix} -0.952381 \\ -0.761905 \\ 0.190476 \end{pmatrix}$$

$$t_2 = 0.323077 \qquad x^3 = x^2 + t_2 d^2 = \begin{pmatrix} -0.040467 \\ -0.342928 \\ -0.035236 \end{pmatrix}$$

$$v^2 = x^3 - x^2 = \begin{pmatrix} -0.307692 \\ -0.246154 \\ 0.061538 \end{pmatrix}$$

$$\nabla f(x^3) = \begin{pmatrix} 0.595533 \\ -0.714640 \\ 0.119107 \end{pmatrix} \quad u^2 = \nabla f(x^3) - \nabla f(x^2) = \begin{pmatrix} -0.307692 \\ -1.230769 \\ 1.538462 \end{pmatrix}$$

$$H_3 = \begin{pmatrix} 1.131584 & -0.071580 & 0.009053 \\ -0.040811 & 0.209795 & -0.000326 \\ 0.193668 & 0.153520 & 0.001549 \end{pmatrix}$$

Bước lặp $k = 3$.

$$d^3 = -H_3 \nabla f(x^3) = \begin{pmatrix} -0.726129 \\ 0.174271 \\ -0.005809 \end{pmatrix}$$

$$t_3 = 0.820149 \qquad x^4 = x^3 + t_3 d^3 = \begin{pmatrix} -1.000000 \\ -0.200000 \\ -0.040000 \end{pmatrix}$$

Vì $\nabla f(x^4) = (0,\ 0,\ 0)^T$ nên dừng thuật toán. Ta nhận được nghiệm

$$x^* = (-1\ ,\ -0.2,\ -0.04)^T$$

đúng như mong muốn (xem Ví dụ 6.9).

4.2.5 Cực tiểu hàm một biến

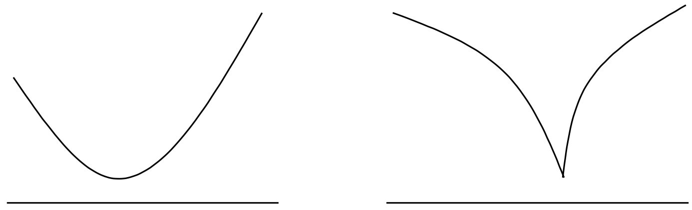
Trong nhiều bước tính toán của bài toán quy hoạch nhiều biến, ta thường phải tìm cực tiểu của hàm theo một hướng nào đó, tức tìm cực tiểu hàm số một biến. Mục này trình bày hai phương pháp đơn giản để giải bài toán

$$\min\ f(x) \ \text{ v.đ.k. } \ x \in [a,b] \subset \mathbb{R}, \tag{P^{1b}}$$

trong đó $f : [a,b] \to \mathbb{R}$ là hàm đơn mốt (unimodal function) và giả sử bài toán (P^{1b}) có nghiệm cực tiểu $x^* \in (a,b)$.

Định nghĩa. Ta nói một hàm số thực f xác định trên đoạn $[a,b] \subset \mathbb{R}$ là hàm *đơn mốt* nếu: i) Hàm f đạt cực tiểu tại $x^* \in (a,b)$; ii) Trong khoảng (a,b), giá trị hàm $f(x)$ hội tụ giảm dần đến $f(x^*)$ khi x hội tụ đơn điệu đến x^*.

Dễ thấy hàm lồi là hàm đơn mốt. Tuy nhiên, hàm đơn mốt có thể không phải hàm lồi, cũng không phải hàm lõm. Xem minh họa ở Hình 6.8.



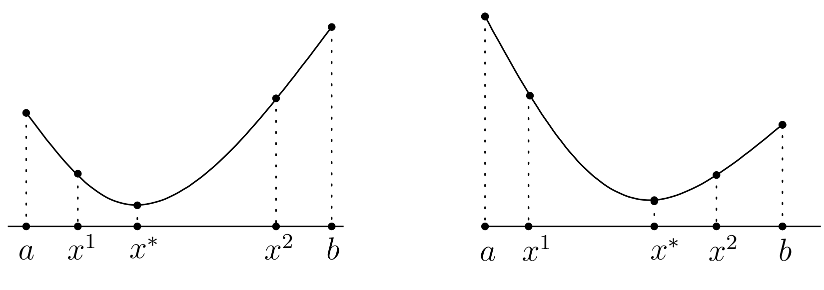
Hình 6.8. *Một số dáng điệu của đồ thị hàm đơn mốt*

Một tính chất quan trọng của hàm đơn mốt trên đoạn $[a,b]$ là: *Giả sử nghiệm cực tiểu x^* thuộc khoảng (a,b) và $a < x^1 < x^2 < b$. Khi đó:*

i) Nếu $f(x^1) < f(x^2)$ thì $x^* \in [a, x^2]$;

ii) Nếu $f(x^1) > f(x^2)$ thì $x^* \in [x^1, b]$.

Xem minh họa ở Hình 6.9.



Hình 6.9

Một khoảng chứa nghiệm cực tiểu x^* được gọi là *khoảng bất định* (uncertainty). Ký hiệu giá trị tối ưu của bài toán cần giải là f_*. Ta có $f_* = f(x^*)$. Hai phương pháp trình bày sau đây cũng như hầu hết các phương pháp khác để tìm cực tiểu của hàm một biến có tính chất của hàm đơn mốt đều dựa trên ý tưởng thu hẹp dần miền chứa nghiệm.

a. Phương pháp chia đôi

Thuật toán 6.6. (*Thuật toán chia đôi*)

Bước khởi đầu. Lấy $\varepsilon > 0$ đủ nhỏ; Đặt $a^1 := a$; $b^1 := b$; $k := 1$.

Bước lặp k. $(k = 1, 2, \dots)$

(k_1) Đặt

$$c := \frac{a^k + b^k}{2}; \quad x^k := c - \frac{\varepsilon}{2}; \quad y^k := c + \frac{\varepsilon}{2}.$$

(*có* $a^k < x^k < y^k < b^k$)

(k_2) Tính $z_1 = f(x^k)$ và $z_2 = f(y^k)$.

(k_3) **If** $z_1 \le z_2$ **Then** Chuyển tới Bước k_4 **Else** Chuyển tới Bước k_5.

(k_4) (*Có* $x^* \in [a^k, y^k]$)

If $y^k - a^k \le \varepsilon$ **Then** Dừng thuật toán (*lấy* $x^* := x^k$ *và* $f_* := z_1$)

Else Đặt $a^{k+1} := a^k$; $b^{k+1} := y^k$; $k := k + 1$; Chuyển về Bước lặp k.

(k_5) (*Có* $x^* \in [x^k, b^k]$)

If $b^k - x^k \le \varepsilon$ **Then** Dừng thuật toán (*lấy* $x^* := y^k$ *và* $f_* := z_2$)

Else Đặt $a^{k+1} := x^k$; $b^{k+1} := b^k$; $k := k + 1$; Chuyển về Bước lặp k.

Nhận xét: Sau mỗi bước lặp, khoảng chứa nghiệm giảm xấp xỉ một nửa so với bước trước nó.

b. Phương pháp lát cắt vàng

Phương pháp này sử dụng tính chất của dãy số Fibonacci $\{F_n\}$. Nhắc lại, dãy $\{F_n\}$ là dãy số: $1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13, \dots$, tức là:

$F_1 = F_2 = 1$;

$F_n = F_{n-1} + F_{n-2}$ với $n = 3, 4, \dots$

(Mỗi số hạng, kể từ số hạng thứ ba, bằng tổng của hai số hạng đứng ngay trước nó).

Dãy số Fibonacci[3] $\{F_n\}$ có tính chất đặc biệt đáng chú ý là: *tỷ số giữa hai số kế tiếp nhau của dãy số này tiến tới tỷ số vàng*

$$\lim_{n\to\infty} \frac{F_{n-1}}{F_n} = \frac{\sqrt{5}-1}{2}.$$

Thuật toán lát cắt vàng giải bài toán (P^{1b}) được mô tả như sau:

Thuật toán 6.7. (*Thuật toán lát cắt vàng*)

Bước khởi đầu. Lấy $\varepsilon > 0$ đủ nhỏ; Đặt $a^1 := a$; $b^1 := b$; $k := 1$ và

$$\alpha := \frac{\sqrt{5}-1}{2}.$$

Bước lặp k. $(k = 1, 2, \dots)$

(k_1) Chia $[a^k, b^k]$ bởi các điểm chia x^k và y^k, trong đó

$$x^k := a^k + (1-\alpha)(b^k - a^k)$$
$$y^k := a^k + \alpha(b^k - a^k).$$

(k_2) Tính $z_1 := f(x^k)$ và $z_2 := f(y^k)$.

(k_3) **If** $z_1 \leq z_2$ **Then** Chuyển tới Bước k_4 **Else** Chuyển tới Bước k_5;

(k_4) (*Có* $x^* \in [a^k, y^k]$)

If $y^k - a^k \leq \varepsilon$ **Then** Dừng thuật toán (*lấy* $x^* := x^k$ *và* $f_* := z_1$)

Else Đặt $a^{k+1} := a^k;\ b^{k+1} := y^k;\ k := k+1$; Chuyển về Bước lặp k.

(k_5) (*Có* $x^* \in [x^k, b^k]$)

If $b^k - x^k \leq \varepsilon$ **Then** Dừng thuật toán (*lấy* $x^* := y^k$ *và* $f_* := z_2$)

Else Đặt $a^{k+1} := x^k;\ b^{k+1} := b^k;\ k := k+1$; Chuyển về Bước lặp k.

---

[3]Leonardo FIBONACCI (1180−1250): Nhà toán học người Italy. Tên thật của ông là Leonardo De Pisa, có nghĩa là Leonard ở thành Pise. Fibonacci là người có công trong việc truyền bá những tính ưu việt của toán học A rập vào châu Âu thời Phục hưng. Dãy số mang tên ông được ông đưa ra năm 1202, bắt nguồn từ bài toán "thỏ đẻ" như sau: *Xuất phát từ một cặp thỏ. Hỏi sau một số tháng ta có bao nhiêu cặp thỏ, biết rằng mỗi cặp thỏ mỗi tháng đẻ ra một cặp và cặp mới này sau hai tháng lại bắt đầu đẻ?*

Nhận xét 4.6.

i) Trong cả hai trường hợp, ta đều có

$$b^{k+1} - a^{k+1} = \alpha(b^k - a^k).$$

ii) Một trong hai điểm chia ở bước sau trùng với điểm chia ở bước trước. Do đó thuật toán này cho phép giảm số phép tính.

Ví dụ 4.5. Tìm $\min\{f(x) = x^2 + 2x + 2 \mid x \in [-5, 3]\}$ với $\varepsilon = 0.001$.

Kết quả giải bài toán này theo Thuật toán 6.7 với $\varepsilon = 0.001$ được trình bày ở Bảng 6.1. Thuật toán dừng ở Bước lặp 19 và ta lấy nghiệm tối ưu $x^* = -1.000163$ và giá trị tối ưu $f_* = 1.000000027$.

Bảng 6.1

Bước k	a^k	b^k	x^k	y^k
1	-5.00000	3.00000	-1.944272	-0.055728
2	-5.00000	-0.055728	-3.111456	-1.944272
3	-3.111456	-0.055728	-1.944272	-1.222912
4	-1.944272	- 0.055728	-1.222912	-0.777088
5	-1.944272	-0.777088	-1.498447	-1.222912
6	-1.498447	-0.777088	-1.222912	-1.052622
7	-1.222912	-0.777088	-1.052622	-0.947377
8	-1.222912	-0.947377	-1.117667	-1.052622
9	-1.117667	-0.947377	-1.052622	-1.012422
10	-1.052622	-0.947377	-1.012422	-0.987577
11	-1.052622	-0.987577	-1.027777	-1.012422
12	-1.027777	-0.987577	-1.012422	-1.002932
13	-1.012422	-0.987577	-1.002932	-0.997067
14	-1.012422	-0.997067	-1.006557	-1.002932
15	-1.006557	-0.997067	-1.002932	-1.000692
16	-1.002932	-0.997067	-1.000692	-0.999308
17	-1.002932	-0.999308	-1.001548	-1.000692
18	-1.001548	-0.999308	-1.000692	-1.000163
19	-1.000692	-0.999308	-1.000163	-0.999836

4.2.6 Phương pháp tìm kiếm trực tiếp

Mục này dành để trình bày hai thuật toán giải bài toán (P^{krb})

$$\min\ f(x)\ \text{ v.đ.k. }\ x \in \mathbb{R}^n,$$

theo phương pháp tìm kiếm trực tiếp là: Thuật toán của Hooke và Jeeves[4] và Thuật toán tìm kiếm theo đơn hình[5] (Sequential simplex search algorithm).

Các thuật toán này được sử dụng để giải bài toán (P^{krb}) khi hàm mục tiêu $f(x)$ không khả vi hoặc có khả vi nhưng việc lấy các đạo hàm riêng là khó khăn do $f(x)$ cấu trúc phức tạp hoặc khi có ít thông tin về $f(x)$.

a. Thuật toán của Hooke và Jeeves

Thuật toán 6.8. (*Thuật toán giảm theo tọa độ*)

Bước 0. Xuất phát từ điểm tùy ý $x^1 = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$. Chọn các véc tơ

$$v^i = (0, \ldots, 0, \underbrace{\delta}_{i}, 0, \ldots, 0)^T, \quad i = 1, \ldots, n,$$

trong đó $\delta > 0$ là số cho trước. Chọn trước số $\varepsilon > 0$ đủ nhỏ.

Bước 1. Đặt $x^{1_0} = x^1$. Trong ba điểm x^{1_0} và $x^{1_0} \pm v^1$, chọn một điểm mà tại đó giá trị hàm mục tiêu bé nhất, ký hiệu điểm đó là x^{1_1}. Để đơn giản ta viết: Tìm

$x^{1_1} = \text{argmin}\{f(x^{1_0}), f(x^{1_0} + v^1), f(x^{1_0} - v^1)\}$ *(theo tọa độ thứ nhất)*

$x^{1_2} = \text{argmin}\{f(x^{1_1}), f(x^{1_1} + v^2), f(x^{1_1} - v^2)\}$ *(theo tọa độ thứ hai)*

$$\vdots$$

$x^{1_n} = \text{argmin}\{f(x^{1_{n-1}}), f(x^{1_{n-1}} + v^n), f(x^{1_{n-1}} - v^n)\}$ *(theo tọa độ thứ n)*

Gán $y^1 := x^{1_n}$.

Bước 2. **If** $x^1 = y^1$ **Then** Chuyển Bước 3 **Else** Chuyển tới Bước 4.

Bước 3. **If** $\delta \le \varepsilon$ **Then** Dừng thuật toán (*lấy* $x^* \approx x^1$)
Else Đặt $v^i = \frac{v^i}{2}$, $i = 1, \ldots, n$ và quay lại Bước 1.

Bước 4. Đặt $x^2 := x^1 + 2(y^1 - x^1)$, $\quad x^1 := x^2$ và quay lại Bước 1.

Ví dụ 4.6. Xét bài toán:

$$\min\{f(x) = x_1^2 + x_2^2 \mid x \in \mathbb{R}^2\}.$$

Bài toán này có $n = 2$. Xuất phát từ $x^1 = (1.0\ ,\ 1.0)^T$. Lấy $\delta = 0.5$, ta có

$$v^1 = (0.5\ ,\ 0)^T, \quad v^2 = (0\ ,\ 0.5)^T$$

Chọn $\varepsilon = 0.5$. Quá trình tính toán giải theo Thuật toán 6.8 như sau:

[4] Thuật toán này do Hooke và Jeeves đề xuất năm 1961.

[5] Năm 1962, thuật toán này được đưa ra bởi Spendly, Hext và Himsworth và được Nelder và Mead cải tiến vào năm 1965.

Bước 1. (lần thứ nhất) Gán $x^{1_0} = x^1 = (1.0\ ,\ 1.0)^T$. Tính trực tiếp ta có:

$$\begin{aligned} f(x^{1_0}) &= f(1.0\ ,\ 1.0) = 2.0 \\ f(x^{1_0} + v^1) &= f(1.5\ ,\ 1.0) = 3.25 \\ f(x^{1_0} - v^1) &= f(0.5\ ,\ 1.0) = 1.25. \end{aligned}$$

Do đó $\min\{f(x^{1_0}),\ f(x^{1_0} + v^1),\ f(x^{1_0} - v^1)\} = f(x^{1_0} - v^1) = 1.25$. Suy ra $x^{1_1} = (0.5\ ,\ 1)^T$.

◇ Lặp lại với x^{1_1},

$$\begin{aligned} f(x^{1_1}) &= f(0.5\ ,\ 1.0) = 1.25 \\ f(x^{1_1} + v^2) &= f(0.5\ ,\ 1.5) = 2.5 \\ f(x^{1_1} - v^2) &= f(0.5\ ,\ 0.5) = 0.5. \end{aligned}$$

Do đó $\min\{f(x^{1_1}),\ f(x^{1_1} + v^2),\ f(x^{1_1} - v^2)\} = f(x^{1_1} - v^2) = 0.5$. Suy ra $x^{1_2} = (0.5\ ,\ 0.5)^T$.

Đặt $y^1 := x^{1_2} = (0.5\ ,\ 0.5)^T$. Vì $y^1 \neq x^1$ nên đặt

$$x^2 = x^1 + 2(y^1 - x^1) = (0\ ,\ 0)^T; \qquad x^1 := x^2.$$

Quay lại Bước 1.

Bước 1. (lần thứ hai) Gán $x^{1_0} = x^1 = (0\ ,\ 0)^T$. Ta có

$$\min\{f(x^{1_0}),\ f(x^{1_0} + v^1),\ f(x^{1_0} - v^1)\} = f(x^{1_0}) = 0.0$$

và

$$x^{1_1} = x^{1_0} = (0.0\ ,\ 0.0)^T.$$

Tiếp tục

$$\min\{f(x^{1_1}),\ f(x^{1_1} + v^2),\ f(x^{1_1} - v^2)\} = f(x^{1_1}) = 0.0$$

và

$$x^{1_2} = x^{1_1} = (0.0\ ,\ 0.0)^T.$$

Gán $y^1 := x^{1_2}$. Vì $y^1 = x^1$ và $\delta = \varepsilon$ nên dừng thuật toán và coi $x^1 = (0, 0)^T$ là nghiệm cực tiểu địa phương.

Dễ thấy rằng, trong ví dụ đơn giản này, điểm $(0\ ,\ 0)^T$ chính là cực tiểu toàn cục của hàm lồi $f(x_1, x_2) = x_1^2 + x_2^2$.

b. Thuật toán tìm kiếm theo đơn hình

Thuật toán 6.9.

Bước 1. ◇ Tạo một đơn hình có $(n + 1)$ đỉnh $x^1, \ldots, x^{n+1} \in \mathbb{R}^n$.

$\diamond$ Tính $f(x^i),\ i = 1, \ldots, n+1$.

Bước 2. Tính:

$f^{max} := f(x^{i_M}) = \max\{f(x^i) \mid i = 1, \ldots, n+1\},\ i_M \in \{1, \ldots, n+1\}$

$f^{min} := f(x^{i_m}) = \min\{f(x^i) \mid i = 1, \ldots, n+1\},\ i_m \in \{1, \ldots, n+1\}$.

Đặt $x^{max} := x^{i_M}$ và $x^{min} = x^{i_m}$.

Bước 3. (Tiêu chuẩn tối ưu)

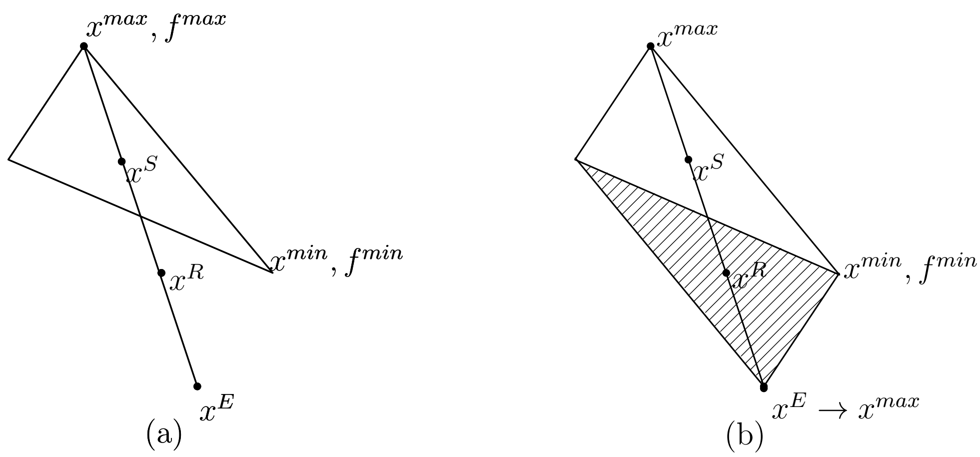
If $|f^{max} - f^{min}| \le \varepsilon$ **Then** Dừng thuật toán

(x^{min} là nghiệm tối ưu địa phương và f^{min} là giá trị tối ưu tương ứng)

Else Chuyển tới Bước 4.

Bước 4. (Tính điểm trọng tâm)

$$x^S = \sum_{i=1}^{n+1} \frac{|f(x^i)|}{\sum_{i=1}^{n+1} |f(x^i)|} x^i \ \text{ (tổ hợp lồi của } x^i,\ i = 1, \ldots, n+1).$$



Hình 6.10

Bước 5. Chiếu đối xứng x^{max} qua x^S được

$$x^R = x^{max} + 2(x^S - x^{max}) \ \text{ (tức } x^S = \frac{1}{2}x^{max} + \frac{1}{2}x^R).$$

Đặt $f^R := f(x^R)$.

Bước 6. **If** $f^R \le f^{min}$ **Then** Chuyển tới Bước 7 **Else** Chuyển tới Bước 8.

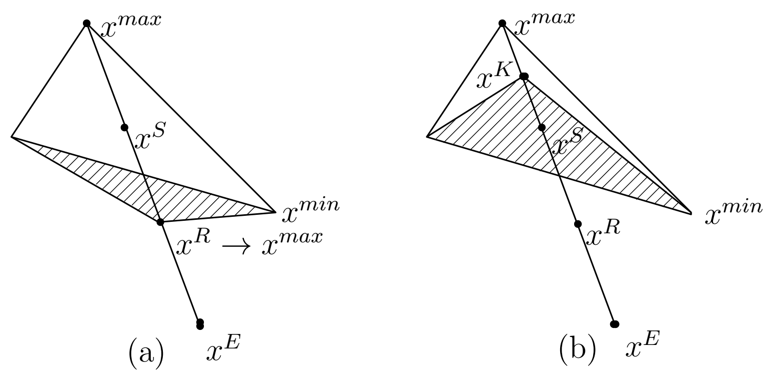
Bước 7. Chiếu đối xứng x^S qua x^R được

$$x^E = x^S + 2(x^R - x^S) \quad \text{(Xem Hình 6.10(a))}.$$

Đặt $f^E := f(x^E)$.

If $f^E < f^{min}$ **Then** Đặt $x^{i_M} := x^E$ (tức thay $x^{max} := x^E$),
được đơn hình mới (Hình 6.10(b)) và quay về Bước 2
Else Đặt $x^{i_M} := x^R$ (tức thay $x^{max} := x^R$),
được đơn hình mới (Hình 6.11(a)) và quay về Bước 2.

Bước 8. (Đã có $f^R > f^{min}$)
If $f^R < f^{max}$ **Then** Đặt $x^{i_M} := x^R$ (tức thay $x^{max} := x^R$),
được đơn hình mới và quay về Bước 2
Else Chuyển Bước 9.



Hình 6.11

Bước 9. (Đã có $f^R \geq f^{max}$) Tính:

$$x^K := \frac{1}{2}x^{max} + \frac{1}{2}x^S \ \text{ và } \ f^K := f(x^K).$$

If $f^K < f^{max}$ **Then** Đặt $x^{i_M} := x^K$ (tức thay $x^{max} := x^K$),
được đơn hình mới (Hình 6.11(b)) và quay về Bước 2
Else Thu hẹp đơn hình theo công thức

$$x^i := \frac{1}{2}x^i + \frac{1}{2}x^{min}, \ \ \forall i \in \{1, \ldots, n+1\} \setminus \{i_m\}.$$

Quay lại Bước 2.

Nhận xét 4.7. Trong thực hành tính theo Thuật toán 6.8 (t.ư., Thuật toán 6.9), ta chọn điểm xuất phát x^0 ngẫu nhiên (t.ư., chọn đơn hình xuất phát có $n+1$ đỉnh ngẫu nhiên). Với các dữ liệu đầu vào khác nhau, có thể dẫn đến các nghiệm tối ưu địa phương khác nhau. Cuối cùng, ta chọn nghiệm tốt nhất có thể.

4.3 Bài toán quy hoạch phi tuyến có ràng buộc

Bài toán quy hoạch phi tuyến có ràng buộc tổng quát được phát biểu như sau

$$\min f(x) \text{ v.đ.k. } x \in X, \qquad (P^{rb})$$

trong đó $X \subset \mathbb{R}^n$ và hàm số f xác định trên X.

4.3.1 Điều kiện tối ưu

Trước hết, ta làm quen với khái niệm nón tiếp xúc (tangent cone).

Định nghĩa. Cho dãy $\{x^q\} \subset \mathbb{R}^n$ hội tụ đến $x^0 \in \mathbb{R}^n$. Ta nói dãy $\{x^q\}$ *hội tụ đến* x^0 *theo hướng* $v \in \mathbb{R}^n$, ký hiệu là $\{x^q\} \xrightarrow{v} x^0$, nếu tồn tại dãy số dương $\{t_q\}$, $\lim_{q\to\infty} t_q = 0$ sao cho

$$x^q = x^0 + t_q v + o(t_q).$$

Nói cách khác, $\{x^q\} \xrightarrow{v} x^0$ nếu tồn tại dãy số dương $\{t_q\}$, $\lim_{q\to\infty} t_q = 0$ sao cho

$$\lim_{q\to\infty} \frac{x^q - x^0}{t_q} = v.$$

Nhận xét 4.8.

i) Nếu $\{x^q\} \xrightarrow{v} x^0$ thì $\{x^q\} \xrightarrow{\lambda v} x^0$ với mọi $\lambda > 0$ (Bài tập);

ii) Không phải mọi dãy hội tụ đều hội tụ theo hướng nhưng với mỗi dãy hội tụ luôn có một dãy con hội tụ theo một hướng nào đó;

iii) Hiển nhiên là $x^0 \xrightarrow{0} x^0$.

Định nghĩa. Cho $X \subset \mathbb{R}^n$. Tập tất cả các hướng $v \in \mathbb{R}^n$ sao cho có một dãy $\{x^q\} \subset X$ hội tụ đến x^0 theo hướng v tạo thành một nón (Bài tập). Ta gọi đó là *nón tiếp xúc* với X tại $x^0 \in X$, ký hiệu là $T(X, x^0)$, cụ thể

$$T(X, x^0) := \{v \in \mathbb{R}^n \mid \exists \{x^q\} \subset X \text{ sao cho } \{x^q\} \xrightarrow{v} x^0\}.$$

Nhận xét 4.9.

i) Nếu $x^0 \in \text{int} X$ thì $T(X, x^0) = \mathbb{R}^n$;

ii) Nếu $X \subset \mathbb{R}^n$ là tập lồi đóng thì $T(X, x^0) = \text{cone}\{(x - x^0) \mid x \in X\}$.

Bổ đề 4.1. *Giả sử $\{x^q\}$ là một dãy thuộc $X \subset \mathbb{R}^n$ hội tụ đến $x^0 \in X$ theo hướng v và f là hàm khả vi liên tục cấp một trên X. Khi đó*

$$\langle \nabla f(x^0), v \rangle = \lim_{t_q \to 0^+} \frac{f(x^q) - f(x^0)}{t_q}.$$

Chứng minh. Giả sử $\{x^q\} \xrightarrow{v} x^0$. Theo định nghĩa, tồn tại dãy số dương $\{t_q\}$, $\lim_{q\to\infty} t_q = 0$, sao cho

$$\lim_{q\to\infty} \frac{x^q - x^0}{t_q} = v.$$

Khai triển Taylor của hàm f tại x^0 (với q đủ lớn) là

$$f(x^q) = f(x^0) + \langle \nabla f(x^0), x^q - x^0 \rangle + o(\|x^q - x^0\|).$$

Suy ra

$$\frac{f(x^q) - f(x^0)}{t_q} = \frac{\langle \nabla f(x^0), x^q - x^0 \rangle}{t_q} + \frac{o(\|x^q - x^0\|)}{t_q}.$$

Do đó, khi $t_q \to 0^+$ ta có

$$\langle \nabla f(x^0), v \rangle = \lim_{q} \frac{f(x^q) - f(x^0)}{t_q}$$

vì

$$\lim_{t_q\to 0^+} \frac{o(\|x^q - x^0\|)}{t_q} = \lim_{t_q\to 0^+} \frac{o(\|x^q - x^0\|)}{\|x^q - x^0\|}.\frac{\|x^q - x^0\|}{t_q} = 0.$$

$\square$

Định lý 4.12.

i) *Giả sử f khả vi trên một tập mở chứa X. Nếu $x^* \in X$ là nghiệm cực tiểu địa phương của bài toán (P^{rb}) thì*

$$\langle \nabla f(x^*), v \rangle \geq 0 \ \ \forall v \in T(X, x^*). \tag{6.12}$$

ii) *Ngược lại, nếu $x^* \in X$ thỏa mãn điều kiện*

$$\langle \nabla f(x^*), v \rangle > 0 \ \ \forall v \in T(X, x^*) \setminus \{0\}, \tag{6.13}$$

thì x^* là một nghiệm tối ưu địa phương chặt của bài toán (P^{rb}).

Chứng minh.

i) Xét bất kỳ $v \in T(X, x^*)$. Theo định nghĩa, tồn tại dãy $\{x^q\} \subset X$ và $\{x^q\} \xrightarrow{v} x^*$. Theo Bổ đề 6.1, ta có

$$\langle \nabla f(x^*), v \rangle = \lim_{t_q \to 0+} \frac{f(x^q) - f(x^*)}{t_q}.$$

Vì x^* là nghiệm cực tiểu địa phương nên $f(x^q) - f(x^*) \geq 0$ với q đủ lớn. Do đó $\langle \nabla f(x^*), v \rangle \geq 0$.

ii) Giả thiết phản chứng rằng x^* không phải nghiệm cực tiểu địa phương chặt của bài toán (P^{rb}). Khi đó, tồn tại một dãy $\{x^q\} \subset X$, $x^q \neq x^*$ và $\{x^q\} \to x^*$ sao cho $f(x^q) \leq f(x^*)$. Trích một dãy con (nếu cần), ta có thể giả sử rằng dãy $\{x^q\}$ hội tụ đến x^* theo một hướng $v \in \mathbb{R}^n$. Khi đó, $v \in T(X, x^*)$. Theo Bổ đề 6.1, ta có

$$\langle \nabla f(x^*), v \rangle = \lim_{t_q \to 0+} \frac{f(x^q) - f(x^*)}{t_q} \leq 0.$$

Điều này mâu thuẫn với giả thiết (6.13) và chứng tỏ rằng x^* là nghiệm cực tiểu địa phương chặt của bài toán (P^{rb}). □

Một điểm $x^* \in X$ thỏa mãn

$$\langle \nabla f(x^*), v \rangle \geq 0 \ \ \forall v \in T(X, x^*)$$

được gọi là một *điểm dừng* hay *điểm tới hạn* của hàm f trên X. Chú ý rằng điều kiện (6.12) chỉ là điều kiện cần nhưng không đủ.

Ví dụ 4.7. Xét bài toán

$$\min\{f(x_1, x_2) = x_1^2 - x_2^2 \mid x_1^2 + x_2^2 \leq 1\}.$$

Tập chấp nhận được của bài toán là hình tròn đóng $\bar{B}(0, 1)$ có tâm 0 và bán kính 1. Ta có $\nabla f(x) = (2x_1, \ -2x_2)^T$ và $\nabla f(x^0) = 0$ với $x^0 = (0, \ 0)^T$. Vì $x^0 \in \text{int}\bar{B}(0, 1)$ nên $T(\bar{B}(0, 1), x^0) = \mathbb{R}^2$ và

$$\langle \nabla f(x^0), v \rangle = 0, \quad v \in T\Big(\bar{B}(0, 1), x^0\Big).$$

Tuy nhiên $x^0 = (0, \ 0)^T$ không phải nghiệm cực tiểu địa phương của bài toán đang xét vì $f(0, \pm\varepsilon) = -\varepsilon^2 < 0$ với mọi ε.

Hệ quả 4.4. *Giả sử $x^* \in \text{int}X$ và x^* là điểm cực tiểu địa phương của bài toán (P^{rb}). Khi đó $\nabla f(x^*) = 0$.*

Chứng minh. Do $x^* \in \text{int}X$ nên $T(X, x^*) = \mathbb{R}^n$. Vì x^* là điểm cực tiểu địa phương nên theo Định lý 6.13,

$$\langle \nabla f(x^*), v\rangle \geq 0 \ \ \forall v \in \mathbb{R}^n.$$

Suy ra $\nabla f(x^*) = 0$. □

Định lý 4.13. *Cho f là hàm lồi khả vi trên một tập mở chứa tập lồi $X \subset \mathbb{R}^n$. Điều kiện cần và đủ để $x^* \in X$ là điểm cực tiểu toàn cục của bài toán quy hoạch lồi* $\min\{f(x) \mid x \in X\}$ *là*

$$\langle \nabla f(x^*), v\rangle \geq 0 \ \ \forall v \in T(X, x^*). \tag{6.14}$$

Chứng minh. Biểu thức (6.14) chính là biểu thức (6.12). Do đã có Định lý 6.13, ta chỉ cần chứng minh điều kiện đủ, tức nếu $x^* \in X$ thỏa mãn (6.14) thì nó là điểm cực tiểu của hàm f trên X. Giả sử phản chứng rằng $x^* \in X$ thỏa mãn (6.14) nhưng nó không phải là điểm cực tiểu của hàm f trên X, tức tồn tại $\bar{x} \in X$ sao cho $f(\bar{x}) < f(x^*)$. Đặt $v = \bar{x} - x^*$. Vì X là tập lồi nên $x^* + tv \in X$ với mọi $0 \leq t \leq 1$ và $v \in T(X, x^*)$. Theo Hệ quả 6.1,

$$\langle \nabla f(x^*), v\rangle \leq f(x^* + v) - f(x^*) = f(\bar{x}) - f(x^*) < 0,$$

mâu thuẫn với (6.14). Điều đó chứng tỏ x^* là nghiệm cực tiểu của f trên X. □

Ví dụ 4.8. Xét bài toán

$$\min\{f(x_1, x_2) = x_1^2 + x_2^2 \mid x_1^2 + x_2^2 \leq 1\}$$

Tập chấp nhận được của bài toán là hình tròn đóng $\bar{B}(0, 1)$. Ta có $\nabla f(x) = (2x_1,\ 2x_2)^T$ và $\nabla f(x^0) = 0$ với $x^0 = (0,\ 0)^T$. Dễ thấy ma trận Hesse $\nabla^2 f(x)$ là xác định dương nên hàm mục tiêu $f(x)$ là hàm lồi chặt (Định lý 6.4(i)). Tương tự Ví dụ 6.16, nón tiếp xúc với $\bar{B}(0, 1)$ tại x^0 là $T(\bar{B}(0, 1), x^0) = \mathbb{R}^2$ và

$$\langle \nabla f(x^0), v\rangle = 0, \quad v \in T\Big(\bar{B}(0, 1), x^0\Big).$$

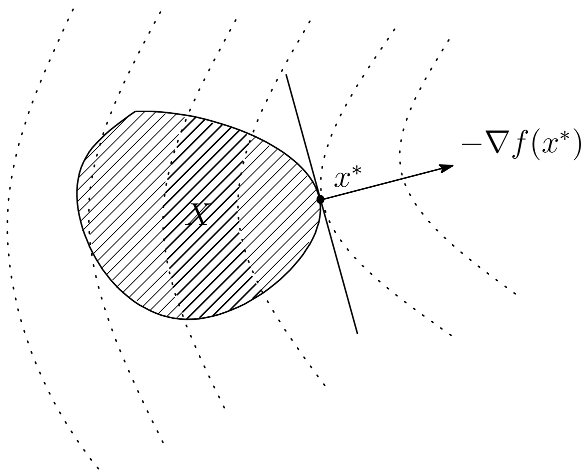
Theo Định lý 6.14, $x^0 = (0,\ 0)^T$ là nghiệm cực tiểu của bài toán đang xét. Hơn nữa, do hàm mục tiêu là lồi chặt nên x^0 là nghiệm cực tiểu duy nhất (Mệnh đề 6.5(ii)).

Hệ quả 4.5. *Cho f là hàm lồi khả vi trên một tập mở chứa tập lồi $X \subset \mathbb{R}^n$. Điểm $x^* \in X$ là điểm cực tiểu toàn cục của bài toán quy hoạch lồi* $\min\{f(x) \mid x \in X\}$ *khi và chỉ khi*

$$\langle \nabla f(x^*), x - x^*\rangle \geq 0 \ \ \forall x \in X.$$

Chứng minh. Suy trực tiếp từ Định lý 6.14 và Nhận xét 6.9(ii). □

Ý nghĩa hình học. Theo Hệ quả 6.5, điểm $x^* \in X$ (có $\nabla f(x^*) \neq 0$) là nghiệm cực tiểu toàn cục của bài toán $\min\{f(x) \mid x \in X\}$ khi và chỉ khi véc tơ $-\nabla f(x^*)$ xác định một siêu phẳng tựa của tập lồi chấp nhận được X tại x^*. Xem minh họa ở Hình 6.12.



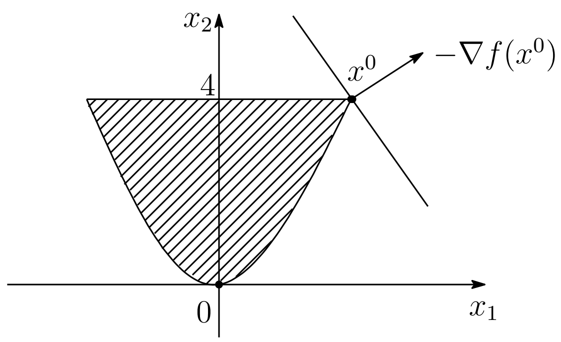
Hình 6.12. *$x^* \in X$ là nghiệm cực tiểu của quy hoạch lồi* $\min\{f(x) \mid x \in X\}$.

Ví dụ 4.9. Xét bài toán $\min\{f(x) = (x_1 - 4)^2 + (x_2 - 6)^2 \mid x_2 \geq x_1^2,\ \ x_2 \leq 4\}$. Điểm $x^0 = (2,\ 4)^T$ có phải nghiệm cực tiểu của bài toán này không? Vì sao?

Giải. Dễ thấy $f(x)$ là hàm lồi và

$$\nabla f(x) = \begin{pmatrix} 2(x_1 - 4) \\ 2(x_2 - 6) \end{pmatrix} \quad \Rightarrow \quad \nabla f(x^0) = \begin{pmatrix} -4 \\ -4 \end{pmatrix}.$$

Tập chấp nhận được X của bài toán này là tập lồi (phần gạch chéo) và véc tơ $-\nabla f(x^0)$ được minh họa ở Hình 6.13. Ta thấy $-\nabla f(x^0)$ xác định một siêu phẳng tựa với X tại x^0. Do đó x^0 là nghiệm cực tiểu của bài toán đang xét.



Hình 6.13

c. Định lý Karush-Kuhn-Tucker

Xét bài toán quy hoạch phi tuyến:

$$\min\ f(x)\ \text{ v.đ.k. }\ x \in X, \qquad (P_1^{rb})$$

trong đó $X \subset \mathbb{R}^n$ là tập nghiệm của hệ

$$\begin{cases} g_i(x) \le 0, & i = 1, \dots, m, \\ h_j(x) = 0, & j = 1, \dots, k, \end{cases}$$

với f, g_i, $i = 1, \dots, m$ và h_j, $j = 1, \dots, k$ là các hàm số khả vi bất kỳ xác định trên $\mathbb{R}^n$, có thể không lồi. Như thường lệ, mỗi hệ thức $g_{i_0}(x) \le 0$, $i_0 \in \{1, \dots, m\}$ hoặc $h_{j_0}(x) = 0$, $j_0 \in \{1, \dots, k\}$, được gọi là một *ràng buộc*.

Nhận xét 4.10.

i) Nếu $g_i(x),\ i = 1, \dots, m$ là các hàm lồi và $h_j(x),\ j = 1, \dots, k$ là các hàm afin thì X là tập lồi, đóng. Nếu thêm điều kiện f là hàm lồi thì bài toán (P_1^{rb}) là một quy hoạch lồi.

ii) Do mọi hàm afin đều là hàm lồi nên quy hoạch tuyến tính là một trường hợp riêng của quy hoạch lồi.

Cho $x^0 \in X$ là một nghiệm chấp nhận được của bài toán (P_1^{rb}). Đặt

$$I(x^0) := \{i \in \{1, \dots, m\} \mid g_i(x^0) = 0\}$$

là *tập các chỉ số của các ràng buộc $g_i(x) \le 0$, $i = 1, \dots, m$, thỏa mãn chặt tại x^0.*

Ký hiệu $S(x^0)$ là tập hợp các véc tơ v thỏa mãn hệ tuyến tính sau:

$$\begin{cases} \langle \nabla h_j(x^0), v\rangle = 0, & j = 1, \dots, k \\ \langle \nabla g_i(x^0), v\rangle \le 0, & i \in I(x^0). \end{cases}$$

Bổ đề 4.2. *Với mọi $x^0 \in X$ ta có $T(X, x^0) \subseteq S(x^0)$.*

Chứng minh. Cho $v \in T(X, x^0)$. Theo định nghĩa, tồn tại dãy $\{x^q\} \subset X$, tồn tại dãy số dương $\{t_q\}$, $\lim_{q\to\infty} t_q = 0$ sao cho

$$\lim_{q\to\infty} \frac{x^q - x^0}{t_q} = v.$$

Vì $x^0 \in X$ và $x^q \in X$ nên $h_j(x^0) = h_j(x^q) = 0$ với mọi $j = 1, \dots, k$ và $g_i(x^q) \le 0$ với mọi $i = 1, \dots, m$. Với mỗi $i \in I(x^0)$, ta có $g_i(x^0) = 0$. Theo Bổ đề 6.1, ta có

$$\langle \nabla h_j(x^0), v\rangle = \lim_{t_q \to 0+} \frac{h_j(x^q) - h_j(x^0)}{t_q} = 0,\ \ \forall j = 1, \dots, k$$

$$\langle \nabla g_i(x^0), v\rangle = \lim_{t_q \to 0+} \frac{g_i(x^q) - g_i(x^0)}{t_q} \le 0, \ \ \forall i \in I(x^0).$$

Do đó $v \in S(x^0)$, hay $T(X, x^0) \subseteq S(x^0)$. □

Định nghĩa. Ta nói *điều kiện chính quy* (regular condition) được thỏa mãn tại x^0 nếu

$$T(X, x^0) = S(x^0).$$

Định lý 4.14. ([27] – trang 23) *Điều kiện chính quy được thỏa mãn tại x^0 nếu có một trong các điều kiện sau:*

i) *Các hàm $h_j,\ j = 1, \ldots, k$ và $g_i,\ i = 1, \ldots, m$ đều là các hàm afin;*

ii) *Các hàm $h_j,\ j = 1, \ldots, k$ là afin; các hàm $g_i,\ i = 1, \ldots, m$ là lồi và điều kiện Slater*[6] *sau đây thỏa mãn*

$$\exists\ \bar{x} \in \mathbb{R}^n\ :\ g_i(\bar{x}) < 0,\ i = 1, \ldots, m \text{ và } h_j(\bar{x}) = 0,\ j = 1, \ldots, k;$$

iii) *Các véc tơ $\{\nabla g_i(x^0),\ i \in I(x^0)\}$ và $\{\nabla h_j(x^0),\ j = 1, \ldots, k\}$ độc lập tuyến tính.*

Điều kiện thứ nhất hoặc thứ hai của Định lý 6.15 đảm bảo điều kiện chính quy được thỏa mãn tại mọi điểm chấp nhận được của bài toán đang xét mà không cần chỉ rõ điểm x^0 nào, còn điều kiện thứ ba đòi hỏi phải biết điểm x^0.

Định lý 4.15. (Định lý Karush-Kuhn-Tucker[7]) *Cho các hàm $f,\ g_i,\ i = 1, \ldots, m$ và $h_j,\ j = 1, \ldots, k$ là các hàm khả vi liên tục trên một tập mở chứa X. Giả sử x^* là nghiệm cực tiểu địa phương của bài toán (P_1^{rb}) và điều kiện chính quy được thỏa mãn tại x^*. Khi đó điều kiện KKT (điều kiện (6.15)–(6.17)) sau đúng:*

i) $g_i(x^*) \le 0,\ i = 1, \ldots, m$ *và* $h_j(x^*) = 0,\ j = 1, \ldots, k;$ (6.15)

ii) *Tồn tại các số $\lambda_i \ge 0,\ i = 1, \ldots, m$ và các số $\mu_j,\ \ j = 1, \ldots, k$ sao cho*

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{j=1}^{k} \mu_j \nabla h_j(x^*) = 0 \ \text{ và} \tag{6.16}$$

iii) $\lambda_i g_i(x^*) = 0,\ \forall i = 1, \ldots, m.$ *(Điều kiện bù)* (6.17)

[6] Điều kiện này được M. Slater đưa ra năm 1950.

[7] Định lý này được Karush đưa ra đầu tiên vào năm 1939 nhưng thiếu điều kiện không âm của các số thực $\lambda_i, \ldots, \lambda_m$. Sau đó được Kuhn và Tucker hoàn thiện năm 1951.

Chứng minh. Giả sử x^* là điểm cực tiểu địa phương của bài toán (P_1^{rb}) và điều kiện chính quy thỏa mãn tại x^*. Do x^* phải là nghiệm chấp nhận được của bài toán nên nó phải thỏa mãn (6.15). Theo Định lý 6.13(i), ta có

$$\langle \nabla f(x^*), v\rangle \geq 0 \quad \forall v \in T(X, x^*).$$

Do $T(X,x^*) \subset S(x^*)$ (Bổ đề 6.2) nên mọi véc tơ $v \in T(X,x^*)$ đều thỏa mãn

$$\begin{cases} \langle \nabla h_j(x^*), v\rangle = 0, & j = 1, \dots, k \\ \langle \nabla g_i(x^*), v\rangle \leq 0, & i \in I(x^*). \end{cases}$$

Đặt A là ma trận cấp $(2k + |I(x^*)|) \times n$ với các hàng là $\nabla h_j(x^*)$, $j = 1, \dots, k$; $-\nabla h_j(x^*)$, $j = 1, \dots, k$ và $-\nabla g_i(x^*)$, $i \in I(x^*)$, trong đó $|I(x^*)|$ là ký hiệu số phần tử của tập $I(x^*)$. Vì điều kiện chính quy thỏa mãn tại x^* nên $T(X,x^*) = S(x^*)$ và ta có thể viết

$$T(X, x^*) = \{v \in \mathbb{R}^n \mid Av \geq 0\}.$$

Theo Bổ đề Farkas (Hệ quả 2.1) ta có

$$\langle \nabla f(x^*), v\rangle \geq 0, \quad \forall v \in T(X, x^*)$$

khi và chỉ khi tồn tại các số không âm $\alpha_j \geq 0$, $\beta_j \geq 0$, $j = 1, \dots, k$, $\lambda_i \geq 0$, $i \in I(x^*)$, sao cho

$$\nabla f(x^*) = \sum_{j=1}^{k} (\alpha_j - \beta_j) \nabla h_j(x^*) - \sum_{i \in I(x^*)} \lambda_i \nabla g_i(x^*)$$

Đặt $\lambda_i = 0, \forall i \notin I(x^*)$ và $-\mu_j = \alpha_j - \beta_j$, $\forall j = 1, \dots, k$ ta được biểu thức (6.16) và (6.17). $\square$

Chú ý 4.4.

i) Điểm $x^* \in \mathbb{R}^n$ thỏa mãn điều kiện KKT được gọi là một *điểm KKT* của bài toán (P_1^{rb}) tương ứng.

ii) Nếu x^* là một nghiệm cực tiểu địa phương của bài toán (P_1^{rb}) và tại đó thỏa mãn điều kiện chính quy thì x^* là điểm KKT, nhưng điều ngược lại chưa chắc đúng.

Ví dụ 4.10. Xét bài toán

$$\min \; -(x_1^2 + x_2^2) \quad \text{v.đ.k.} \;\; x_1 \leq 1.$$

Bài toán này có $m = 1$, $k = 0$ và hàm mục tiêu $f(x) = -(x_1^2 + x_2^2)$ là hàm lõm chặt. Vì $g_1 = x_1 - 1$ là hàm afin nên điều kiện chính quy thỏa mãn tại mọi điểm chấp nhận được của bài toán. Ta có

$$\nabla f(x) = \begin{pmatrix} -2x_1 \\ -2x_2 \end{pmatrix} \quad \text{và} \quad \nabla g_1(x) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Điều kiện KKT tương ứng của bài toán này là:

$$\begin{cases} g_1(x) = x_1 - 1 \le 0 \\ \nabla f(x) + \lambda_1 \nabla g_1(x) = 0 \\ \lambda_1 g_1(x) = 0 \\ \lambda_1 \ge 0, \end{cases} \Leftrightarrow \begin{cases} \quad x_1 - 1 & \le 0 \\ -2x_1 + \lambda_1 & = 0 \\ -2x_2 & = 0 \\ \lambda_1(x_1 - 1) & = 0 \\ \quad \lambda_1 & \ge 0. \end{cases}$$

Giải hệ này ta nhận được hai điểm KKT là điểm $x^1 = (0,0)^T$ ứng với $\lambda_1 = 0$ và điểm $x^2 = (1,0)^T$ ứng với $\lambda_1 = 2$. Dễ thấy $x^1 = (0,0)^T$ là nghiệm cực đại toàn cục của hàm $f(x)$ trên tập chấp nhận được của bài toán đang xét còn điểm $x^2 = (1,0)^T$ không phải nghiệm cực tiểu địa phương cũng không phải nghiệm cực đại địa phương.

- **Định lý Karush-Kuhn-Tucker cho quy hoạch lồi**

Xét bài toán quy hoạch lồi

$$\min \ f(x) \ \text{ v.đ.k. } \ x \in X, \qquad (P_1^{conv})$$

trong đó

$$X = \{x \mid g_i(x) \le 0, \ i = 1, \dots, m\},$$

và f và g_i, $i = 1, 2, \dots, m$, là các hàm lồi, khả vi liên tục trên một tập mở chứa X. Trong bài toán quy hoạch lồi, nếu xuất hiện ràng buộc $h_j(x) = 0$ thì $h_j(x)$ phải là hàm afin. Hơn nữa, vì

$$h_j(x) = 0 \quad \text{tương đương với} \quad h_j(x) \le 0, \ -h_j(x) \le 0,$$

nên ta chỉ cần xét ràng buộc dạng bất đẳng thức.

Định lý 4.16. (Định lý Karush-Kuhn-Tucker cho quy hoạch lồi) *Giả sử các hàm f, g_i, $i = 1, \dots, m$ là các hàm lồi khả vi trên một tập mở chứa X và điều kiện Slater được thỏa mãn. Khi đó $x^* \in \mathbb{R}^n$ là nghiệm cực tiểu của bài toán (P_1^{conv}) khi và chỉ khi x^* thỏa mãn điều kiện KKT (điều kiện (6.18)–(6.20)) sau:*

i) $g_i(x^*) \le 0, \ i = 1, \dots, m;$ (6.18)

ii) *Tồn tại các số* $\lambda_i \geq 0,\ i = 1, \ldots, m$ *sao cho*

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) = 0 \ \ \textit{và} \tag{6.19}$$

iii) $\lambda_i g_i(x^*) = 0,\ \forall i = 1, \ldots, m.$ (6.20)

Chứng minh. Do đã có Định lý 6.16 nên ta chỉ cần chứng minh điều kiện đủ, tức nếu x^* thỏa mãn (6.18)−(6.20) thì x^* là nghiệm cực tiểu của bài toán (P_1^{conv}). Thật vậy, vì f và g_i, $i = 1, \ldots, m$, là các hàm lồi nên $f + \sum_{i=1}^{m} \lambda_i g_i$, với $\lambda_i \geq 0$, $i = 1, \ldots, m$, cũng là hàm lồi (Mệnh đề 6.3). Từ (6.18) suy ra $x^* \in X$. Lấy bất kỳ $x \in X$. Theo Hệ quả 6.1, ta có

$$(x-x^*)^T \underbrace{\Big[\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) \Big]}_{=0 \ do \ (6.19)} \leq \Big[f(x) + \underbrace{\sum_{i=1}^{m} \lambda_i g_i(x)}_{\leq 0 \ do \ x \in X} \Big] - \Big[f(x^*) + \underbrace{\sum_{i=1}^{m} \lambda_i g_i(x^*)}_{=0 \ do \ (6.20)} \Big].$$

Do đó

$$f(x^*) \leq f(x),$$

chứng tỏ x^* là nghiệm cực tiểu của bài toán (P_1^{conv}). □

Chú ý 4.5. Theo Định lý 6.17, việc giải bài toán quy hoạch lồi (P_1^{conv}) tương đương với việc tìm nghiệm của hệ KKT (6.18)−(6.20) tương ứng.

Ví dụ 4.11. Xét bài toán quy hoạch lồi toàn phương:

$$\min \ f(x) = \frac{1}{2} x^T Q x + c^T x$$

$$\text{với điều kiện } \ Ax \leq b,$$

trong đó Q là ma trận cấp $n \times n$ đối xứng và nửa xác định dương, A là ma trận cấp $m \times n$, $b \in \mathbb{R}^m$, $c, x \in \mathbb{R}^n$.

Theo Định lý 6.17, thay vì giải bài toán này, ta có thể giải hệ KTT tương ứng với nó là

$$\begin{cases} Ax \leq b \\ Qx + c + A^T\lambda = 0 \\ \lambda \geq 0 \\ \lambda^T(Ax - b) = 0. \end{cases}$$

Chú ý 4.6. Điều kiện chính quy trong Định lý 6.16 và Định lý 6.17 **không bỏ qua được**. Nếu x^* là nghiệm cực tiểu địa phương nhưng không thỏa mãn điều kiện chính quy thì điều kiện KKT chưa chắc đã được thỏa mãn tại x^*.

Ví dụ 4.12. Xét bài toán[8]:

$$\min \ x_1 \quad \text{v.đ.k.} \quad x_1^3 - x_2^2 = 0.$$

Ta có hàm mục tiêu $f(x) = x_1,\ m = 0,\ k = 1$ và $h_1(x) = x_1^3 - x_2^2 = 0$. Vì mọi nghiệm chấp nhận được $(x_1,\ x_2)^T$ của bài toán này phải thỏa mãn $x_1^3 = x_2^2$ nên $x_1 \geq 0$. Do đó $x^0 = (0,\ 0)^T$ là nghiệm cực tiểu. Ta có

$$\nabla f(x) = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ \ \nabla h_1(x) = \begin{pmatrix} 3x_1^2 \\ -2x_2 \end{pmatrix} \Rightarrow \nabla f(x^0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ \ \nabla h_1(x^0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Nếu x^0 thỏa mãn điều kiện KKT thì phải tồn tại số $\mu_1 \in \mathbb{R}$ sao cho

$$\nabla f(x^0) + \mu_1 \nabla h_1(x^0) = 0 \ \text{ tức } \ \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Mâu thuẫn này chứng tỏ $x^0 = (0,\ 0)^T$ không là điểm KKT. Lý do là véc tơ $\nabla h_1(x^0) = (0,\ 0)^T$ không độc lập tuyến tính, tức điều kiện chính quy bị vi phạm tại $x^0 = (0, 0)^T$.

4.3.2 Phương pháp nhân tử Lagrange

Hàm số

$$\mathbb{L}(x, \lambda_1, \ldots, \lambda_m, \mu_1, \ldots, \mu_k) := f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{k} \mu_j h_j(x),$$

với các số thực $\lambda_1 \geq 0, \ldots, \lambda_m \geq 0, \mu_1, \ldots, \mu_k$ được gọi là *hàm Lagrange*[9], tương ứng với bài toán quy hoạch phi tuyến

$$\min \ \ f(x) \ \ \text{v.đ.k.} \quad x \in X, \qquad (P_1^{rb})$$

trong đó $X = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0,\ i = 1, \ldots, m,\ \ h_j(x) = 0,\ j = 1, \ldots, k\}$. Các số không âm $\lambda_1, \ldots, \lambda_m$ và các số $\mu_1, \ldots, \mu_k$ được gọi là *các nhân tử Lagrange*.

Ký hiệu $\nabla_x \mathbb{L}$ là gradient của hàm $\mathbb{L}$ theo x. Khi đó, Định lý 6.16 được phát biểu lại theo ngôn ngữ hàm Lagrange như sau:

[8] theo [34], trang 34.

[9] Joseph Louis LAGRANGE (1736 − 1813): Nhà toán học người Italy gốc Pháp. Ông là người sáng lập ra Viện Hàn lâm Turin ở Italy. Sau đó, ông được vua Phổ Friedrich II vời đến làm việc tại Berlin với lời mời khiêm tốn: "Nhà toán học lớn nhất châu Âu cần sống bên cạnh ông vua lớn nhất". Năm 1788, ông đã xây dựng phương pháp này để giải bài toán tối ưu với các ràng buộc đều là đẳng thức.

Định lý 4.17. *Cho f, g_i, $i = 1, \ldots, m$ và h_j, $j = 1, \ldots, k$ là các hàm khả vi liên tục trên $\mathbb{R}^n$. Giả sử điểm $x^* \in \mathbb{R}^n$ là nghiệm cực tiểu địa phương của bài toán (P_1^{rb}) và tại đó điều kiện chính quy được thỏa mãn. Khi đó điều kiện KKT sau là đúng:*

i) $g_i(x^*) \le 0$, $i = 1, \ldots, m$ *và* $h_j(x^*) = 0$, $j = 1, \ldots, k$;

ii) *Tồn tại các số thực* $\lambda_i \ge 0$, $i = 1, \ldots, m$ *và* μ_j, $j = 1, \ldots, k$ *sao cho*

$$\nabla_x \mathbb{L}(x^*, \lambda_1, \ldots, \lambda_m, \mu_1, \ldots, \mu_k) = 0 \text{ và}$$

iii) $\lambda_i g_i(x^*) = 0$, $i = 1, \ldots, m$.

Chú ý 4.7.

i) Dễ thấy

$$\nabla_x \mathbb{L}(x^*, \lambda_1, \ldots, \lambda_m, \mu_1, \ldots, \mu_k) = \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{j=1}^{k} \mu_j \nabla h_j(x^*);$$

ii) Giả sử f, g_i, $i = 1, \ldots, m$ và h_j, $j = 1, \ldots, k$ là các hàm khả vi liên tục cấp hai trong lân cận của điểm x^* và x^* là một điểm KKT của bài toán (P_1^{rb}). Khi đó, với $d \in \mathbb{R}^n$ có $\|d\|$ đủ nhỏ, khai triển Taylor của hàm Lagrange tại x^* là

$$\begin{aligned}\mathbb{L}(x^* + d, \lambda, \mu) &= \mathbb{L}(x^*, \lambda, \mu) + d^T \nabla_x \mathbb{L}(x^*, \lambda, \mu) + \frac{1}{2} d^T \nabla^2 \mathbb{L}(\xi, \lambda, \mu) d \\ &= \mathbb{L}(x^*, \lambda, \mu) + \frac{1}{2} d^T \nabla^2 \mathbb{L}(\xi, \lambda, \mu) d,\end{aligned}$$

trong đó ξ nằm giữa x^* và $x^* + d$, $\lambda = (\lambda_1, \ldots, \lambda_m)^T$, $\mu = (\mu_1, \ldots, \mu_k)^T$. Như đã biết, nếu ma trận $\nabla^2 \mathbb{L}(x^*, \lambda, \mu)$ xác định dương thì $\nabla^2 \mathbb{L}(\xi, \lambda, \mu)$ cũng xác định dương và x^* là một nghiệm cực tiểu địa phương chặt của (P_1^{rb}).

Sau đây là phát biểu lại của Định lý 6.17 cho quy hoạch lồi

$$\min \{f(x) \mid g_i(x) \le 0,\ i = 1, \ldots, m\}. \qquad (P_1^{conv})$$

Định lý 4.18. *Giả sử f, g_i, $i = 1, \ldots, m$ là các hàm lồi khả vi trên một tập mở chứa X và điều kiện Slater được thỏa mãn. Khi đó $x^* \in \mathbb{R}^n$ là nghiệm cực tiểu của bài toán (P_1^{conv}) khi và chỉ khi x^* thỏa mãn điều kiện KKT sau:*

i) $g_i(x^*) \le 0$, $i = 1, \ldots, m$;

ii) *Tồn tại các số thực* $\lambda_1 \ge 0, \ldots, \lambda_m \ge 0$ *sao cho*

$$\nabla_x \mathbb{L}(x^*, \lambda_1, \ldots, \lambda_m) = 0 \text{ và}$$

iii) $\lambda_i g_i(x^*) = 0$, $i = 1, \ldots, m$.

Trong trường hợp điều kiện chính quy thỏa mãn, theo Định lý 6.18, thuật toán sau cho phép xác định được các điểm KKT của bài toán (P_1^{rb}).

Thuật toán 6.10.

Bước 1. Lập hàm Lagrange:

$$\mathbb{L}(x, \lambda_1, \dots, \lambda_m, \mu_1, \dots, \mu_k) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{k} \mu_j h_j(x).$$

Bước 2. Giải hệ KKT sau:

$$\begin{cases} \nabla_x \mathbb{L}(x, \lambda_1, \dots, \lambda_m, \mu_1, \dots, \mu_k) = 0 \\ \lambda_1 \geq 0, \dots, \lambda_m \geq 0 \\ \lambda_i g_i(x) = 0, \ i = 1, \dots, m \\ g_i(x) \leq 0, \ i = 1, \dots, m \\ h_j(x) = 0, \ j = 1, \dots, k. \end{cases} \tag{6.21}$$

Mỗi một nghiệm $\bar{x}$ của hệ này tương ứng với một bộ tham số $\bar{\lambda}_1, \dots, \bar{\lambda}_m$, $\bar{\mu}_1, \dots, \bar{\mu}_k$ là một điểm KKT của bài toán đang xét.

Chú ý 4.8.

i) Nếu bài toán (P_1^{rb}) chỉ có các ràng buộc đẳng thức, tức $m = 0$, thì hệ (6.21) trở thành hệ $n + k$ ẩn, $n + k$ phương trình sau:

$$\begin{cases} \nabla_x \mathbb{L}(x, \mu_1, \dots, \mu_k) = \nabla f(x) + \sum_{j=1}^{k} \mu_j \nabla h_j(x) = 0 \\ h_1(x) = 0, \dots, h_k(x) = 0. \end{cases} \tag{6.22}$$

Trong trường hợp này bài toán (P_1^{rb}) thường được gọi là *bài toán trơn với ràng buộc đẳng thức*.

ii) Nếu $k = 0$, tức bài toán chỉ có các ràng buộc bất đẳng thức, thì hệ (6.21) trở thành:

$$\begin{cases} \nabla_x \mathbb{L}(x, \lambda_1, \dots, \lambda_m) = \nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x) = 0 \\ \lambda_i g_i(x) = 0, \quad \forall\, i = 1, \dots, m \\ \lambda_1 \geq 0, \dots, \lambda_m \geq 0 \\ g_1(x) \leq 0, \dots, g_m(x) \leq 0. \end{cases} \tag{6.23}$$

Thêm vào đó, nếu các hàm f, $g_i, i = 1, \dots, m$ là lồi thì bài toán (P_1^{rb}) trở thành bài toán quy hoạch lồi (P_1^{conv}). Khi đó người ta thường gọi bài toán này là *bài toán lồi với ràng buộc bất đẳng thức*. Trong trường hợp này, mỗi nghiệm x^* của hệ (6.23) tương ứng với một bộ $\lambda_1^*, \dots, \lambda_m^*$ là một nghiệm cực tiểu của bài toán (Định lý 6.19).

iii) Nếu bài toán (P_1^{rb}) chỉ có một ràng buộc $m = 1$ (hoặc $k = 1$) thì điều kiện chính quy thỏa mãn với mọi x mà tại đó $\nabla g_1(x) \neq 0$ (hoặc $\nabla h_1(x) \neq 0$).

Ví dụ 4.13. Xét bài toán:

$$\min\{f(x) = x_1^3 + x_2^3 + x_3^3 \mid x_1^2 + x_2^2 + x_3^2 = 4\}.$$

Tập chấp nhận được của bài toán này là một mặt cầu trong $\mathbb{R}^3$ nên nó là tập compac. Hàm mục tiêu là hàm liên tục. Do đó, theo Định lý Weierstrass (Hệ quả 1.1), bài toán này có nghiệm. Ta có $n = 3$, $m = 0$, $k = 1$, $h_1(x) = x_1^2 + x_2^2 + x_3^2 - 4 = 0$,

$$\nabla f(x) = \begin{pmatrix} 3x_1^2 \\ 3x_2^2 \\ 3x_3^2 \end{pmatrix} \quad \text{và} \quad \nabla h_1(x) = \begin{pmatrix} 2x_1 \\ 2x_2 \\ 2x_3 \end{pmatrix} \neq 0 \;\; \forall x \neq 0.$$

Vì vậy điều kiện chính quy được thỏa mãn tại mọi điểm chấp nhận được của bài toán. Quá trình tính toán giải bài toán này theo Thuật toán 6.10 như sau:

- Lập hàm Lagrange:

$$\mathbb{L}(x, \mu_1) = f(x) + \mu_1 h_1(x) = x_1^3 + x_2^3 + x_3^3 + \mu_1(x_1^2 + x_2^2 + x_3^2 - 4).$$

- Giải hệ phương trình $n + k = 4$ ẩn, 4 phương trình:

$$\begin{cases} \nabla_x \mathbb{L}(x, \mu_1) = 0 \\ h_1(x) = 0 \end{cases} \Leftrightarrow \begin{cases} \mathbb{L}'_{x_1} = 3x_1^2 + 2\mu_1 x_1 = 0 \\ \mathbb{L}'_{x_2} = 3x_2^2 + 2\mu_1 x_2 = 0 \\ \mathbb{L}'_{x_3} = 3x_3^2 + 2\mu_1 x_3 = 0 \\ x_1^2 + x_2^2 + x_3^2 = 4 \end{cases} \Leftrightarrow \begin{cases} (3x_1 + 2\mu_1)x_1 = 0 \\ (3x_2 + 2\mu_1)x_2 = 0 \\ (3x_3 + 2\mu_1)x_3 = 0 \\ x_1^2 + x_2^2 + x_3^2 = 4. \end{cases}$$

Ta xét các trường hợp có thể xảy ra:

1. Nghiệm $x = (x_1, x_2, x_3)^T$ có một tọa độ khác 0, hai tọa độ bằng 0. Chẳng hạn,

$$x_1 \neq 0, \; x_2 = x_3 = 0 \implies x_1 = \pm 2 \quad \text{tương ứng với } \mu_1 = \mp 3.$$

Như vậy, trong trường hợp này ta có các điểm KKT sau:

$$(\pm 2, 0, 0)^T, \quad (0, \pm 2, 0)^T, \quad \text{và} \quad (0, 0, \pm 2)^T.$$

2. Nghiệm $x = (x_1, x_2, x_3)^T$ có hai tọa độ khác 0, một tọa độ bằng 0. Chẳng hạn, $x_1 \neq 0,\ x_2 \neq 0,\ x_3 = 0$. Từ hai phương trình đầu suy ra $x_1 = x_2 = -\frac{2}{3}\mu_1$. Kết hợp điều này với phương trình cuối, ta có $x_1 = x_2 = \pm\sqrt{2}$ tương ứng với $\mu_1 = \mp\frac{3}{\sqrt{2}}$. Vậy, trong trường hợp này, ta có các điểm KKT sau:
$$(\sqrt{2}, \sqrt{2}, 0)^T,\ (\sqrt{2}, 0, \sqrt{2})^T,\ (0, \sqrt{2}, \sqrt{2})^T,$$
$$(-\sqrt{2}, -\sqrt{2}, 0)^T,\ (-\sqrt{2}, 0, -\sqrt{2})^T,\ (0, -\sqrt{2}, -\sqrt{2})^T.$$

3. Nghiệm $x = (x_1, x_2, x_3)^T$ có cả ba tọa độ khác 0. Từ ba phương trình đầu suy ra $x_1 = x_2 = x_3 = -\frac{2}{3}\mu_1$. Kết hợp điều này với phương trình cuối, ta được
$$x_1 = x_2 = x_3 = \pm\frac{2}{\sqrt{3}}.$$
Trường hợp này có hai điểm KKT là
$$\Big(\frac{2}{\sqrt{3}}, \frac{2}{\sqrt{3}}, \frac{2}{\sqrt{3}}\Big)^T, \quad \Big(-\frac{2}{\sqrt{3}}, -\frac{2}{\sqrt{3}}, -\frac{2}{\sqrt{3}}\Big)^T$$
tương ứng với $\mu_1 = \mp\sqrt{3}$.

Vì tập chấp nhận được của bài toán đang xét là tập compac nên hàm mục tiêu f phải đạt giá trị cực tiểu và giá trị cực đại tại các điểm nào đó thuộc tập các điểm KKT vừa tính được ở trên. Tính toán trực tiếp và so sánh giá trị của hàm f tại các điểm này, ta có giá trị cực tiểu của f là -8 tương ứng với ba nghiệm cực tiểu của bài toán là
$$(-2, 0, 0)^T,\ (0, -2, 0)^T,\ (0, 0, -2)^T.$$

Ví dụ 4.14. Xét bài toán tối ưu có ràng buộc:
$$\min\ f(x)\ \text{ v.đ.k. }\ x \in D,$$
trong đó
$$D := \{x \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1\}$$
và hàm mục tiêu lồi chặt $f(x) = x_1^2 + x_1x_2 + x_2^2 + 3(x_1 + x_2 - 2)$ (như hàm mục tiêu trong Ví dụ 6.6).

Bài toán này có $n = 2,\ m = 1,\ k = 0,\ g_1(x) = x_1^2 + x_2^2 - 1 \leq 0$. Dễ thấy $g_1(x)$ là hàm lồi và điều kiện Slater được thỏa mãn, tức tồn tại $\bar{x} \in D$ có $g_1(\bar{x}) < 0$. Do đó, điều kiện chính quy được thỏa mãn tại mọi điểm chấp nhận được của bài toán. Bài toán này là một bài toán quy hoạch lồi với hàm mục tiêu lồi chặt. Theo Mệnh đề 6.5 (ii), nó có duy nhất nghiệm.

- Hàm Lagrange tương ứng với bài toán này là:

$$\mathbb{L}(x, \lambda_1) = x_1^2 + x_1x_2 + x_2^2 + 3(x_1 + x_2 - 2) + \lambda_1(x_1^2 + x_2^2 - 1).$$

- Giải hệ:

$$\begin{cases} \nabla_x \mathbb{L}(x, \lambda_1) = 0 \\ \lambda_1 g_1(x) = 0 \\ \lambda_1 \geq 0 \\ g_1(x) \leq 0 \end{cases} \Leftrightarrow \begin{cases} \mathbb{L}'_{x_1} = 2x_1 + x_2 + 3 + 2\lambda_1 x_1 = 0 \\ \mathbb{L}'_{x_2} = x_1 + 2x_2 + 3 + 2\lambda_1 x_2 = 0 \\ \lambda_1(x_1^2 + x_2^2 - 1) = 0 \\ \lambda_1 \geq 0 \\ x_1^2 + x_2^2 - 1 \leq 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} 2(1 + \lambda_1)x_1 + x_2 = -3 \\ x_1 + 2(1 + \lambda_1)x_2 = -3 \\ \lambda_1(x_1^2 + x_2^2 - 1) = 0 \\ \lambda_1 \geq 0 \\ x_1^2 + x_2^2 - 1 \leq 0. \end{cases}$$

Từ hai phương trình đầu tiên của hệ này suy ra $x_1 = x_2$.

Nếu $\lambda_1 = 0$ thì ta có $x_1 = x_2 = -1$. Vì điểm $(-1,\ -1)^T$ không phải là điểm chấp nhận được của bài toán đang xét nên kết luận rằng $\lambda_1 \neq 0$.

Do $\lambda_1 \neq 0$ nên từ phương trình thứ ba của hệ này ta có $x_1^2 + x_2^2 = 1$. Kết hợp với sự kiện $x_1 = x_2$ suy ra $x_1 = x_2 = \pm\frac{\sqrt{2}}{2}$.

◇ Nếu $x_1 = x_2 = \frac{\sqrt{2}}{2}$ thì $\lambda_1 < 0$. Điều kiện $\lambda_1 \geq 0$ không được thỏa mãn nên loại trường hợp này.

◇ Nếu $x_1 = x_2 = -\frac{\sqrt{2}}{2}$ thì $\lambda_1 > 0$. Vậy $x^* = (-\frac{\sqrt{2}}{2},\ -\frac{\sqrt{2}}{2})^T$ là điểm KKT của bài toán này. Vì đây là bài toán quy hoạch lồi nên x^* cũng là nghiệm cực tiểu toàn cục (Định lý 6.19).

Nhận xét 4.11. Qua Ví dụ 6.6 và Ví dụ 6.23, ta thấy rằng với cùng hàm mục tiêu $f(x)$ nhưng nghiệm cực tiểu của bài toán tối ưu không ràng buộc và bài toán tối ưu có ràng buộc khác nhau.

Ví dụ 4.15. Xét bài toán:

$$\min\ f(x) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 \quad \text{v.đ.k.} \quad x \in D,$$

trong đó $D := \{x \in \mathbb{R}^2 \mid -x_1 + x_2 \leq 1,\ x_1 + 2x_2 \geq 2,\ x_1 \leq 3\}$.

Đây là bài toán quy hoạch lồi với hàm mục tiêu lồi chặt nên nó có một nghiệm cực tiểu (cũng là điểm KKT) duy nhất. Trong bài toán này, $n = 2$, $m = 3$, $g_1(x) = -x_1 + x_2 - 1 \leq 0$, $g_2(x) = -x_1 - 2x_2 + 2 \leq 0$, $g_3(x) = x_1 - 3 \leq 0$.

- Lập hàm Lagrange:

$$\mathbb{L}(x, \lambda_1, \lambda_2, \lambda_3) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 + \lambda_1(-x_1 + x_2 - 1) + \lambda_2(-x_1 - 2x_2 + 2) + \lambda_3(x_1 - 3).$$

- Hệ KKT tương ứng với bài toán này là:

$$\begin{cases} \nabla_x \mathbb{L}(x, \lambda_1, \lambda_2, \lambda_3) = 0 \\ \lambda_i g_i(x) = 0, \quad i = 1, 2, 3 \\ \lambda_1 \geq 0, \quad \lambda_2 \geq 0, \quad \lambda_3 \geq 0 \\ g_i(x) \leq 0, \quad i = 1, 2, 3 \end{cases} \Leftrightarrow \begin{cases} \mathbb{L}'_{x_1} = x_1 - \lambda_1 - \lambda_2 + \lambda_3 = 0 \\ \mathbb{L}'_{x_2} = x_2 + \lambda_1 - 2\lambda_2 = 0 \\ \lambda_1(-x_1 + x_2 - 1) = 0 \\ \lambda_2(-x_1 - 2x_2 + 2) = 0 \\ \lambda_3(x_1 - 3) = 0 \\ \lambda_1 \geq 0, \quad \lambda_2 \geq 0, \quad \lambda_3 \geq 0 \\ -x_1 + x_2 - 1 \leq 0 \\ -x_1 - 2x_2 + 2 \leq 0 \\ x_1 - 3 \leq 0. \end{cases} \tag{6.24}$$

Ta xét các trường hợp có thể xảy ra:

1. Nếu $\lambda_1 = \lambda_2 = \lambda_3 = 0$ thì từ hai phương trình đầu ta có $x_1 = x_2 = 0$. Khi đó, bất phương trình thứ tám của hệ (6.24) bị vi phạm $\Rightarrow$ Loại trường hợp này.

2. Có $\lambda_1 \neq 0$, $\lambda_2 = \lambda_3 = 0$. Khi đó, hệ (6.24) trở thành:

$$\begin{cases} x_1 - \lambda_1 = 0 \\ x_2 + \lambda_1 = 0 \\ -x_1 + x_2 - 1 = 0 \\ \lambda_1 > 0 \\ -x_1 + x_2 - 1 \leq 0 \\ -x_1 - 2x_2 + 2 \leq 0 \\ x_1 - 3 \leq 0. \end{cases}$$

   Suy ra $x_1 = -x_2 = \lambda_1$, cụ thể $x_2 = 0.5$, $x_1 = -0.5$ và $\lambda_1 = -0.5 < 0$. Do đó loại trường hợp này.

3. Có $\lambda_1 = 0$, $\lambda_2 \neq 0$, $\lambda_3 = 0$. Khi đó, hệ (6.24) trở thành:

$$\begin{cases} x_1 - \lambda_2 = 0 \\ x_2 - 2\lambda_2 = 0 \\ -x_1 - 2x_2 + 2 = 0 \\ \lambda_2 > 0 \\ -x_1 + x_2 - 1 \leq 0 \\ -x_1 - 2x_2 + 2 \leq 0 \\ x_1 - 3 \leq 0. \end{cases}$$

Dễ thấy, $x_2 = 2x_1$ và $x_1 = \frac{2}{5}$, $x_2 = \frac{4}{5}$, $\lambda_2 = \frac{2}{5} > 0$. Các bất phương trình và phương trình của hệ (6.24) đều thỏa mãn với $x = (\frac{2}{5}, \frac{4}{5})^T$. Ta nhận được điểm KKT $x^* = (\frac{2}{5}, \frac{4}{5})^T$. Vậy, x^* cũng là nghiệm cực tiểu duy nhất của bài toán.

4.3.3 Phương pháp tuyến tính hóa giải quy hoạch lồi

Xét bài toán quy hoạch lồi:

$$\min\ f(x) \ \text{ v.đ.k. } \ x \in X, \qquad (P_1^{conv})$$

trong đó $X = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, i = 1, \ldots, m\}$ là tập lồi compac và f, g_i, $i = 1, \ldots, m$ là các hàm lồi khả vi trên $\mathbb{R}^n$.

Bằng việc thêm biến mới t và ràng buộc $t \geq f(x)$, có thể chứng minh rằng (Bài tập) bài toán (P_1^{conv}) tương đương với bài toán sau:

$$\begin{aligned} &\min\ t \\ \text{v.đ.k. } & g_i(x) \leq 0, \quad i = 1, \ldots, m \\ & g_{m+1}(x) \leq 0, \end{aligned}$$

trong đó $g_{m+1}(x) = f(x) - t$. Vì vậy, ta chỉ cần xét bài toán có dạng:

$$\min\ \langle c, x\rangle \ \text{ v.đ.k. } \ x \in X, \qquad (P_2^{conv})$$

trong đó tập chấp nhận được $X := \{x \in \mathbb{R}^n \mid g_i(x) \leq 0,\ i = 1, \ldots, m\}$ là tập lồi compac, g_i, $i = 1, \ldots, m$, là các hàm lồi khả vi trên $\mathbb{R}^n$ và $c \in \mathbb{R}^n$.

Ý tưởng của phương pháp xấp xỉ tuyến tính là đưa việc giải bài toán quy hoạch lồi (P_2^{conv}) về việc giải một dãy các bài toán quy hoạch tuyến tính:

$$\min\{\langle c, x\rangle \mid x \in X_k\},$$

với X_k, $k = 1, 2, 3, \ldots$, là các đa diện lồi thỏa mãn

$$X_1 \supset X_2 \supset \cdots \supset X,$$

mà dãy nghiệm tối ưu $\{x^k\}$ tương ứng của chúng hội tụ đến nghiệm của bài toán này.

Sau đây là thuật toán xấp xỉ tuyến tính giải bài toán (P_2^{conv}) do Kelley đề xuất năm 1960:

Thuật toán 6.11.

Bước chuẩn bị. Xây dựng đa diện ban đầu X_1 sao cho $X_1 \supseteq X$. (*Cách xây dựng X_1 sẽ được trình bày ngay sau khi mô tả thuật toán*). Đặt $k := 1$.
Bước lặp k. $(k = 1, 2 \ldots)$

(k_1) Giải quy hoạch tuyến tính $\min\{\langle c, x\rangle | x \in X_k\}$ được nghiệm tối ưu x^k.

(k_2) **If** $g_i(x^k) \le 0 \ \ \forall\, i = 1, \ldots, m$ **Then** Dừng thuật toán
(*lấy x^k là nghiệm tối ưu của (P_2^{conv})*)
Else ($\exists j \in \{1, \ldots, m\}$: $g_j(x^k) > 0$) Chuyển tới Bước k_3.

(k_3) Xác định $g_{i_k}(x^k) = \max\{g_i(x^k) \mid 1 \le i \le m\}$ và đặt

$$X_{k+1} := X_k \cap \{x \mid \left\langle \nabla g_{i_k}(x^k), x - x^k \right\rangle + g_{i_k}(x^k) \le 0\}.$$

Đặt $k := k + 1$ và quay lại Bước lặp k.

Lưu ý rằng, tại Bước (k_3) của thuật toán trên:

i) Ta luôn có $\nabla g_{i_k}(x^k) \ne 0$. Lý do là: nếu ngược lại thì do g_{i_k} là hàm lồi nên theo Hệ quả 6.1, ta có

$$g_{i_k}(x) \ge g_{i_k}(x^k) + \left\langle \nabla g_{i_k}(x^k), x - x^k \right\rangle > 0 \quad \forall x$$

và điều này có nghĩa là tập chấp nhận được X của bài toán đang xét bằng rỗng;

ii) Siêu phẳng

$$\{x \in \mathbb{R}^n \mid \left\langle \nabla g_{i_k}(x^k), x - x^k \right\rangle + g_{i_k}(x^k) = 0\}$$

để xây dựng đa diện X_{k+1} được gọi là *siêu phẳng cắt*. Đây chính là siêu phẳng tách tập lồi đóng X và điểm $x^k \notin X$.

Sau đây ta sẽ trình bày **cách xây dựng đa diện lồi xuất phát** $X_1 \supset X$.

Cho trước điểm $\bar{x} \in \mathbb{R}^n$. Vì mỗi hàm g_i, $i \in \{1, \ldots, m\}$, là hàm lồi khả vi nên ta có (Hệ quả 6.1)

$$\langle \nabla g_i(\bar{x}), x - \bar{x} \rangle \leq g_i(x) - g_i(\bar{x}) \ \ \forall x \in \mathbb{R}^n.$$

Đặt

$$h_i(x) = \langle \nabla g_i(\bar{x}), x - \bar{x} \rangle + g_i(\bar{x}).$$

Rõ ràng, với mỗi $i \in \{1, \ldots, m\}$ ta có $h_i(x)$ là hàm afin thỏa mãn:

$$\begin{aligned} &h_i(x) \leq g_i(x) \ \ \forall x \in \mathbb{R}^n \\ &h_i(\bar{x}) = g_i(\bar{x}). \end{aligned}$$

Đặt $\bar{X} = \{x \in \mathbb{R}^n \mid h_i(x) \leq 0, \ i = 1, \ldots, m\}$. Dễ thấy $\bar{X}$ là tập lồi đa diện và $X \subset \bar{X}$ (Bài tập).

Về nguyên tắc, ta có thể lấy p $(p \geq 1)$ điểm $\bar{x}^1, \ldots, \bar{x}^p \in \mathbb{R}^n$ sao cho

$$X_1 = \{x \in \mathbb{R}^n \mid h_{ij}(x) \leq 0, \ i = 1, \ldots, m; \ j = 1, \ldots, p\},$$

là một đa diện lồi chứa X, trong đó

$$h_{ij}(x) = \langle \nabla g_i(\bar{x}^j), x - \bar{x}^j \rangle + g_i(\bar{x}^j).$$

Ví dụ 4.16. Xét bài toán:

$$\min \ x_1 \ \ \text{v.đ.k.} \ \ x \in D,$$

trong đó $D := \{x \in \mathbb{R}^2 \mid (x_1 - 1)^2 + x_2^2 \leq 4; \ \ x_1^2 + x_2^2 \leq 4\}$.

Bài toán có $n = 2$, $m = 2$,

$$g_1(x) = (x_1 - 1)^2 + x_2^2 - 4 \leq 0, \quad g_2(x) = x_1^2 + x_2^2 - 4 \leq 0,$$

$$\nabla g_1(x) = \begin{pmatrix} 2(x_1 - 1) \\ 2x_2 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}.$$

- *Bước chuẩn bị (Xây dựng đa diện xuất phát D_1)*: Chọn ba điểm

$$\bar{x}^1 = (0, \ 1)^T, \quad \bar{x}^2 = (0, \ -1)^T \ \ \text{và} \ \ \bar{x}^3 = (2, \ 0)^T.$$

◇ Với $\bar{x}^1 = (0, \ 1)^T$:

$$\nabla g_1(\bar{x}^1) = \begin{pmatrix} -2 \\ 2 \end{pmatrix}, \quad \nabla g_2(\bar{x}^1) = \begin{pmatrix} 0 \\ 2 \end{pmatrix}; \ \ g_1(\bar{x}^1) = -2; \ \ g_2(\bar{x}^1) = -3.$$

$$\begin{aligned} h_{11}(x) &= \langle \nabla g_1(\bar{x}^1), x - \bar{x}^1 \rangle + g_1(\bar{x}^1) = -2x_1 + 2x_2 - 4 \le 0; \\ h_{21}(x) &= \langle \nabla g_2(\bar{x}^1), x - \bar{x}^1 \rangle + g_2(\bar{x}^1) = \qquad\qquad 2x_2 - 5 \le 0. \end{aligned}$$

◇ Với $\bar{x}^2 = (0,\ -1)^T$:

$$\nabla g_1(\bar{x}^2) = \begin{pmatrix} -2 \\ -2 \end{pmatrix}, \quad \nabla g_2(\bar{x}^2) = \begin{pmatrix} 0 \\ -2 \end{pmatrix}; \quad g_1(\bar{x}^2) = -2; \quad g_2(\bar{x}^2) = -3.$$

$$\begin{aligned} h_{12}(x) &= \langle \nabla g_1(\bar{x}^2), x - \bar{x}^2 \rangle + g_1(\bar{x}^2) = -2x_1 - 2x_2 - 4 \le 0; \\ h_{22}(x) &= \langle \nabla g_2(\bar{x}^2), x - \bar{x}^2 \rangle + g_2(\bar{x}^2) = \qquad -2x_2 - 5 \le 0. \end{aligned}$$

◇ Với $\bar{x}^3 = (2,\ 0)^T$:

$$\nabla g_1(\bar{x}^3) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad \nabla g_2(\bar{x}^3) = \begin{pmatrix} 4 \\ 0 \end{pmatrix}; \quad g_1(\bar{x}^3) = -3; \quad g_2(\bar{x}^3) = 0.$$

$$\begin{aligned} h_{13}(x) &= \langle \nabla g_1(\bar{x}^3), x - \bar{x}^3 \rangle + g_1(\bar{x}^3) = 2x_1 - 7 \le 0; \\ h_{23}(x) &= \langle \nabla g_2(\bar{x}^3), x - \bar{x}^3 \rangle + g_2(\bar{x}^3) = 4x_1 - 8 \le 0. \end{aligned}$$

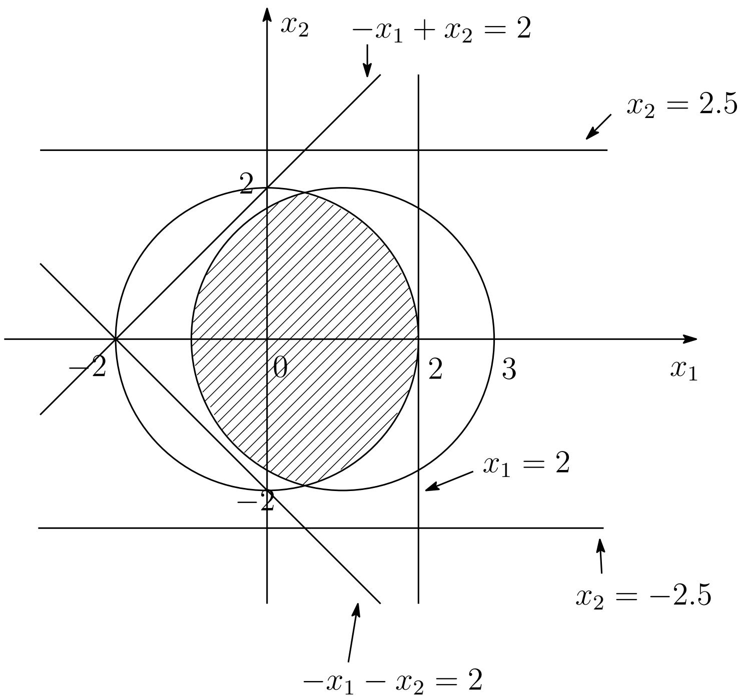
Ta được tập đa diện $D_1 \supset D$ xác định bởi

$$D_1 := \{x \in \mathbb{R}^2 \mid h_{ij}(x) \le 0,\ i = 1, 2,\ j = 1, 2, 3\}.$$

Dễ thấy ràng buộc $h_{22} = 4x_1 - 8 \le 0$ là thừa. Vì vậy ta có

$$D_1 = \{x \in \mathbb{R}^2 \mid -x_1 + x_2 \le 2; \quad x_2 \le 2.5; \quad -x_1 - x_2 \le 2; \quad x_2 \ge -2.5; \quad x_1 \le 2\}$$

(xem Minh họa ở Hình 6.14). Đặt $k := 1$ và sang Bước lặp $k = 1$.



Hình 6.14

- *Bước lặp 1.*

  Bước 1_1. Giải quy hoạch tuyến tính $\min\{f(x) = x_1 \mid x \in D_1\}$ được nghiệm tối ưu là $x^1 = (-2,\ 0)^T$.

  Bước 1_2. Vì $g_1(x^1) = 5 > 0$ nên chuyển Bước 1_3.

  Bước 1_3. Ta có $g_1(x^1) = 5 > g_2(x^1) = 0$ nên chọn $i_1 = 1$ và đặt

$$\begin{aligned} D_2 &:= D_1 \cap \{x \in \mathbb{R}^2 \mid \left\langle \nabla g_1(x^1), x - x^1 \right\rangle + g_1(x^1) \le 0\} \\ &= D_1 \cap \left\{x \in \mathbb{R}^2 \mid x_1 \ge -\frac{7}{6}\right\}. \end{aligned}$$

  Đặt $k := k + 1 = 2$ và chuyển sang Bước lặp $k = 2$.

- *Bước lặp 2.*

  Bước 2_1. Giải quy hoạch tuyến tính $\min\{f(x) = x_1 \mid x \in D_2\}$ được tập nghiệm tối ưu là đoạn $\left[(-\frac{7}{6},\ \frac{5}{6})^T,\ (-\frac{7}{6},\ -\frac{5}{6})^T\right]$. Ta lấy một nghiệm tối ưu đại diện là $x^2 = (-\frac{7}{6},\ 0)^T$.

  Bước 2_2. Vì $g_1(x^2) = \frac{25}{36} > 0$ nên chuyển Bước 2_3.

Bước 2_3. Vì $g_1(x^2) = \frac{25}{36} > g_2(x^2) = -\frac{95}{36}$ nên chọn $i_2 = 1$ và đặt

$$\begin{aligned} D_3 &:= D_2 \cap \{x \in \mathbb{R}^2 \mid \langle \nabla g_1(x^2), x - x^2 \rangle + g_1(x^2) \leq 0\} \\ &= D_2 \cap \{x \in \mathbb{R}^2 \mid x_1 \geq -\frac{157}{156}\}. \end{aligned}$$

Đặt $k := k + 1 = 3$ và chuyển sang Bước lặp $k = 3$.

- *Bước lặp 3.*

Bước 3_1. Giải quy hoạch tuyến tính $\min\{f(x) = x_1 \mid x \in D_3\}$ và lấy một nghiệm tối ưu đại diện là $x^3 = (-\frac{157}{156},\ 0)^T$.

Bước 3_2. Vì $g_1(x^3) \approx 0.025682 > 0$ nên chuyển Bước 2_3.

Bước 3_3. Vì $g_1(x^3) \approx 0.025682 > g_2(x^3) \approx -2.98714$ nên chọn $i_3 = 1$ và đặt

$$\begin{aligned} D_4 &:= D_3 \cap \{x \in \mathbb{R}^2 \mid \langle \nabla g_1(x^3), x - x^3 \rangle + g_1(x^3) \leq 0\} \\ &= D_3 \cap \{x \in \mathbb{R}^2 \mid x_1 \geq -0.9978886\}. \end{aligned}$$

Đặt $k := k + 1 = 4$ và chuyển sang Bước lặp $k = 4$.

- *Bước lặp 4.*

Bước 4_1. Giải quy hoạch tuyến tính $\min\{f(x) = x_1 \mid x \in D_3\}$ và lấy một nghiệm tối ưu đại diện là $x^4 = (-0.9978886,\ 0)^T$.

Bước 4_2. Vì $g_1(x^4) \approx 0.008441142 \approx 0$ và $g_2(x^4) \approx -3.0042183 < 0$ nên dừng thuật toán và lấy nghiệm xấp xỉ là $x^* = (-0.9978886,\ 0)^T$.

Bằng biểu diễn hình học, dễ thấy rằng nghiệm tối ưu của bài toán này là $(-1, 0)^T$.

Định lý 4.19. (Định lý hội tụ) ([27], trang 53) *Giả sử rằng $g_1, \ldots, g_m$ là các hàm lồi khả vi và tập chấp nhận được X của bài toán (P_2^{conv}) là tập compac. Khi đó, dãy $\{x^k\}$ sinh ra bởi Thuật toán 6.11 có một dãy con hội tụ đến nghiệm cực tiểu của bài toán.*

4.3.4 Phương pháp hướng có thể giải bài toán cực tiểu hàm trơn với ràng buộc tuyến tính

Xét bài toán:

$$\min\ f(x) \text{ v.đ.k. } x \in X, \qquad (P_2^{rb})$$

trong đó f là hàm phi tuyến khả vi trên $\mathbb{R}^n$ và $X \subset \mathbb{R}^n$ là tập lồi đa diện khác rỗng xác định bởi

$$X := \{x \in \mathbb{R}^n \mid Ax \leq b,\ Ex = e\},$$

với A là ma trận cấp $m \times n$ với các hàng $a^i \in \mathbb{R}^n$, $i = 1, \ldots, m$, E là ma trận cấp $k \times n$, $b \in \mathbb{R}^m$ và $e \in \mathbb{R}^k$.

Ý tưởng của *phương pháp hướng có thể* (Method of feasible direction) do Zoutendijk đề xuất năm 1960 để giải bài toán (P_2^{rb}) là: Xuất phát từ một điểm chấp nhận được bất kỳ $x^0 \in X$, ta xây dựng một dãy điểm $x^1, x^2, x^3, \ldots$ sao cho:

$$x^{k+1} := x^k + t_k d^k \in X \text{ với } t_k > 0, \qquad f(x^{k+1}) < f(x^k) \qquad (6.25)$$

và dãy $\{x^k\}$ hội tụ đến x^* là một điểm KKT của bài toán đang xét. Véc tơ d^k thỏa mãn (6.25) là một hướng giảm chấp nhận được của bài toán (P_2^{rb}) và t_k là độ dài bước. Do tập chấp nhận được là tập lồi đa diện nên nếu hàm mục tiêu $f(x)$ là hàm lồi thì điểm x^* này cũng chính là nghiệm cực tiểu toàn cục của bài toán đang xét (Định lý 6.17).

a. Cơ sở lý thuyết

Định nghĩa. Cho điểm $x^0 \in X$. Véc tơ $d \in \mathbb{R}^n \setminus \{0\}$ được gọi là một *hướng có thể* (hay *hướng chấp nhận được*) (feasible direction) của tập X tại x^0 nếu tồn tại một số $t_* > 0$ sao cho

$$x^0 + td \in X \text{ với mọi } t \text{ thỏa mãn } 0 < t \leq t_*,$$

nghĩa là khi di chuyển từ điểm x^0 theo hướng d một đoạn đủ nhỏ, ta sẽ không đi vượt ra ngoài miền chấp nhận được X. Ký hiệu $F(X, x^0)$ là tập các hướng chấp nhận được của tập X tại x^0.

Nhận xét 4.12.

i) Bao đóng của tập các hướng chấp nhận được của X tại $x^0 \in X$ bằng nón tiếp xúc với tập X tại x^0, tức

$$\text{cl}F(X, x^0) = T(X, x^0).$$

ii) Nếu X là tập lồi thì $x - x^0$ là hướng chấp nhận được của tập X tại $x^0 \in X$ với mọi $x \in X$ (Bài tập).

Định nghĩa. Véc tơ $d \in \mathbb{R}^n \setminus \{0\}$ được gọi là *hướng dùng được* (usable feasible direction) hoặc *hướng giảm chấp nhận được* (improving feasible direction) của bài toán (P_2^{rb}) tại x^0 nếu nó vừa là hướng giảm của hàm f tại x^0, vừa là hướng chấp nhận được của tập X tại x^0, tức tồn tại một số $t_* > 0$ sao cho

$$x^0 + td \in X \text{ và } f(x^0 + td) < f(x^0) \text{ với mọi } t \text{ thỏa mãn } 0 \le t \le t_*.$$

Cho điểm chấp nhận được x^k. Ký hiệu

$$I(x^k) := \Big\{ i \in \{1, \ldots, m\} \;\Big|\; \langle a^i, x^k \rangle = b_i \Big\}.$$

Mệnh đề sau đây chỉ ra điều kiện cần và đủ để một véc tơ $d \in \mathbb{R}^n$ là hướng chấp nhận được của tập X.

Mệnh đề 4.13. *Cho điểm $x^k \in X$. Khi đó vec tơ d là hướng chấp nhận được của X tại x^k khi và chỉ khi $\langle a^i, d \rangle \le 0$ với mọi $i \in I(x^k)$ và $Ed = 0$.*

Chứng minh.

($\Rightarrow$) Giả sử $d \in \mathbb{R}^n$ là hướng chấp nhận được của X tại x^k. Theo định nghĩa, tồn tại số thực $t_* > 0$ sao cho với mọi $t \in (0, t_*]$ ta có $x^k + td \in X$, tức

$$\langle a^i, x^k + td \rangle = \langle a^i, x^k \rangle + t\langle a^i, d \rangle \le b_i, \ \forall\, i = 1, \ldots, m, \tag{6.26}$$

và

$$E(x^k + td) = \underbrace{Ex^k}_{=e} + tEd = e. \tag{6.27}$$

Từ (6.26) suy ra $\langle a^i, d \rangle \le 0$ với mọi $i \in I(x^k)$ vì $\langle a^i, x^k \rangle = b_i$ với mọi $i \in I(x^k) \subset \{1, \ldots, m\}$ và $t > 0$. Còn biểu thức $Ed = 0$ dễ dàng nhận được từ (6.27).

($\Leftarrow$) Ngược lại, giả sử $\langle a^i, d \rangle \le 0$ với mọi $i \in I(x^k)$ và $Ed = 0$. Khi đó ta có

$$E(x^k + td) = Ex^k + tEd = e \quad \forall \ t \in \mathbb{R}.$$

Hơn nữa,

- $\diamond$ Với mỗi $i \in I(x^k)$ ta có

$$\langle a^i, x^k + td \rangle = \langle a^i, x^k \rangle + t\langle a^i, d \rangle \le \langle a^i, x^k \rangle = b_i, \ \forall\, t > 0.$$

- $\diamond$ Với mỗi $i \notin I(x^k)$ ta có $\langle a^i, x^k \rangle < b_i$ và

$$\langle a^i, x^k + td \rangle = \langle a^i, x^k \rangle + t\langle a^i, d \rangle < b_i + t\langle a^i, d \rangle.$$

Do đó, nếu $\langle a^i, d\rangle \leq 0$ thì $\langle a^i, x^k + td\rangle < b_i$ với mọi $t > 0$. Còn nếu $\langle a^i, d\rangle > 0$ thì

$$\langle a^i, x^k + td\rangle < b_i \text{ với mọi } t \leq \frac{b_i - \langle a^i, x^k\rangle}{\langle a^i, d\rangle}.$$

Bây giờ, đặt

$$t_* = \min\left\{\frac{b_i - \langle a^i, x^k\rangle}{\langle a^i, d\rangle} \;\middle|\; i \notin I(x^k) \text{ và } \langle a^i, d\rangle > 0\right\}.$$

Ta có $x^k + td \in X$ với mọi $t \in (0, t_*]$, chứng tỏ d là hướng chấp nhận được của X tại x^k. $\square$

Nhận xét 4.13. Cho điểm chấp nhận được $x^k \in X$ và véc tơ $d^k \in \mathbb{R}^n$. Ký hiệu

$$I^* = \left\{i \in \{1, \ldots, m\}\backslash I(x^k) \;\middle|\; \langle a^i, d^k\rangle > 0\right\}.$$

Khi đó $x^k + td^k \in X$ với mọi $t \in (0, t_*]$, trong đó

$$t_* = \begin{cases} \min\left\{\dfrac{b_i - \langle a^i, x^k\rangle}{\langle a^i, d\rangle} \;\middle|\; i \in I^*\right\}, & \text{nếu } I^* \neq \emptyset \\ +\infty, & \text{nếu } I^* = \emptyset. \end{cases} \tag{6.28}$$

Mệnh đề 4.14. *Nếu véc tơ $d \in \mathbb{R}^n \setminus \{0\}$ thỏa mãn*

$$\langle a^i, d\rangle \leq 0, \;\; \forall \;\; i \in I(x^k); \quad Ed = 0 \text{ và } \langle \nabla f(x^k), d\rangle < 0$$

thì d là hướng giảm chấp nhận được của bài toán (P_2^{rb}).

Chứng minh. Suy trực tiếp từ Mệnh đề 6.7 và Mệnh đề 6.13. $\square$

Mệnh đề 4.15. *Điểm x^k là điểm KKT của bài toán (P_2^{rb}) khi và chỉ khi bài toán quy hoạch tuyến tính sau có giá trị tối ưu bằng 0:*

$$\begin{aligned} \min \;\; & \langle \nabla f(x^k), d\rangle, \\ \text{v.đ.k.} \;\; & \langle a^i, d\rangle \leq 0, \;\; \forall i \in I(x^k) \\ & Ed = 0, \\ & -1 \leq d_i \leq 1, \;\; \forall i = 1, \ldots, n. \end{aligned} \tag{6.29}$$

Chứng minh. Xem [1], trang $411-412$. □

b. Thuật toán 6.12. (*Giải bài toán* (P_2^{rb}))

Bước khởi đầu. Xác định một đỉnh $x^0 \in X$ (có thể dùng Pha 1 của thuật toán đơn hình). Đặt $k=0$.
Bước lặp k. $(k=0,1,2,\dots)$

(k_1) Giải quy hoạch tuyến tính (6.29) được nghiệm tối ưu d^k.
If $\langle \nabla f(x^k), d^k\rangle = 0$ **Then** Dừng thuật toán (x^k *là điểm KKT*)
Else Chuyển tới Bước k_2.

(k_2) Tính t_* theo công thức (6.28).

(k_3) Xác định $x^{k+1} := x^k + t_k d^k$, trong đó t_k là nghiệm của bài toán cực tiểu hàm một biến $\min \left\{ f(x^k + td^k) \mid t \in [0, t_*] \right\}$.

(k_4) Đặt $k := k+1$ và quay về Bước lặp k.

Chú ý 4.9. Như đã thấy, Thuật toán 6.12 giải bài toán (P_2^{rb}) rất đơn giản, nhưng rất tiếc, nói chung, tính hội tụ của thuật toán không được đảm bảo. Năm 1967, Topkis và Veinott đã cải tiến nó và nhận được thuật toán hội tụ giải bài toán này (xem [1], trang $423-432$). Mục tiếp theo đây sẽ trình bày một thuật toán hội tụ thường được sử dụng để giải bài toán (P_2^{rb}) trong trường hợp hàm mục tiêu f là hàm lồi.

4.3.5 Phương pháp Frank-Wolfe giải bài toán quy hoạch lồi với ràng buộc tuyến tính

Xét quy hoạch lồi

$$\min\ f(x)\ \text{ v.đ.k. }\ x \in X, \qquad (P_3^{conv})$$

trong đó f là hàm lồi trên $\mathbb{R}^n$ và $X \subset \mathbb{R}^n$ là tập lồi đa diện xác định bởi

$$X = \{x \in \mathbb{R}^n \mid Ax \le b\},$$

với A là ma trận cấp $m \times n$ và véc tơ $b \in \mathbb{R}^m$.

Mục này trình bày thuật toán Frank và Wolfe đề xuất năm 1956 để giải bài toán (P_3^{conv}) với giả thiết hàm tuyến tính $\langle \nabla f(\hat{x}), x\rangle$ bị chặn dưới trên X với mỗi $\hat{x} \in \mathbb{R}^n$. Hiển nhiên rằng nếu X là đa diện thì giả thiết này luôn được thỏa mãn.

Cho điểm chấp nhận được $x^k \in X$. Vì X là tập lồi đa diện nên $x - x^k$ là hướng chấp nhận được của X tại x^k với mọi $x \in X$ (Nhận xét 6.12 (ii)). Xét bài toán quy hoạch tuyến tính

$$\min\{\langle \nabla f(x^k), x - x^k\rangle \mid x \in X\}.$$

Giả sử $u^k \in X$ là nghiệm tối ưu của bài toán này. Khi đó:

◇ Nếu giá trị tối ưu $\langle \nabla f(x^k), u^k - x^k\rangle \geq 0$ thì $\langle \nabla f(x^k), x - x^k\rangle \geq 0$ với mọi $x \in X$. Theo Hệ quả 6.5, $x^{opt} = x^k$ là nghiệm cực tiểu của bài toán đang xét.

◇ Ngược lại, nếu giá trị tối ưu $\langle \nabla f(x^k), u^k - x^k\rangle < 0$ thì $u^k - x^k$ là một hướng giảm chấp nhận được của bài toán (P_3^{conv}) (Mệnh đề 6.8).

Sau đây là thuật toán Frank-Wolfe giải bài toán quy hoạch lồi (P_3^{conv}).

Thuật toán 6.13. (*Thuật toán Frank-Wolfe*)

Bước khởi đầu. Tìm một điểm bất kỳ $x^0 \in X$. Đặt $k := 0$.

Bước lặp k. $(k = 0, 1, 2, \dots)$

(k_1) Giải bài toán quy hoạch tuyến tính

$$\min \left\langle \nabla f(x^k), x - x^k \right\rangle \quad \text{v.đ.k.} \;\; x \in X$$

được phương án tối ưu $u^k \in X$.

(k_2) (*kiểm tra điều kiện tối ưu*)
If $\left\langle \nabla f(x^k), u^k - x^k \right\rangle \geq 0$ **Then** Dừng thuật toán (*lấy* $x^{opt} := x^k$)
Else Đặt $d^k := u^k - x^k$ và chuyển Bước (k_3).

(k_3) Xác định điểm $x^{k+1} := x^k + t_k d^k$, trong đó

$$t_k = \operatorname{argmin}\Big\{\varphi(t) = f(x^k + td^k) \mid t \in [0, 1]\Big\}.$$

(k_4) **If** $\nabla f(x^{k+1}) \approx 0$ **Then** Dừng thuật toán $(x^{opt} := x^{k+1})$
Else Đặt $k := k + 1$ và quay lại Bước lặp k.

Định lý 4.20. *Giả sử f là hàm lồi chặt trên $\mathbb{R}^n$ và $X \subset \mathbb{R}^n$ là đa diện khác rỗng. Khi đó, xuất phát từ một điểm bất kỳ $x^0 \in X$, dãy $\{x^k\}$ sinh ra bởi Thuật toán Frank-Wolfe có tính chất: dãy $\{f(x^k)\}$ là dãy số giảm, hội tụ đến giá trị tối ưu f_{opt} của bài toán (P_3^{conv}) và*

$$0 \leq f(x^k) - f_{opt} \leq \langle \nabla f(x^k), x^k - u^k\rangle \quad \forall k.$$

Chứng minh. Xem [32] (trang 116−117) hoặc [37] (trang 356−358) . □

Ví dụ 4.17. Xét lại bài toán đã xét ở Ví dụ 6.24:

$$\min\ f(x) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 \quad \text{v.đ.k.} \quad x \in D,$$

trong đó $D = \{x \in \mathbb{R}^2 \mid -x_1 + x_2 \le 1,\ x_1 + 2x_2 \ge 2,\ x_1 \le 3\}$.

Ta có

$$\nabla f(x) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \text{ và } \nabla^2 f(x) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \Rightarrow \quad f(x) \text{ là hàm lồi chặt.}$$

Xuất phát từ $x^0 = (0,\ 1)^T$ với $\nabla f(x^0) = (0,\ 1)^T$. Đặt $k = 0$.

Bước lặp $k = 0$.

0.1. Giải bài toán $\min\Big\{\langle \nabla f(x^0), x\rangle = x_2 \mid x \in D\Big\}$ được phương án cực biên tối ưu là $u^0 = (3,\ -\frac{1}{2})^T$.

0.2. Vì

$$\langle \nabla f(x^0), u^0 - x^0\rangle = (\nabla f(x^0))^T(u^0 - x^0) = (0 \quad 1)\begin{pmatrix} 3 \\ -\dfrac{3}{2} \end{pmatrix} = -\frac{3}{2} < 0$$

nên x^0 chưa phải nghiệm tối ưu. Đặt $d^0 := u^0 - x^0 = (3,\ -\frac{3}{2})^T$. Ta có d^0 là hướng giảm chấp nhận được của bài toán đang xét tại x^0.

0.3. Tìm điểm chấp nhận được x^1: Ta có

$$x^0 + td^0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + t\begin{pmatrix} 3 \\ -\dfrac{3}{2} \end{pmatrix} = \begin{pmatrix} 3t \\ 1 - \dfrac{3}{2}t \end{pmatrix}.$$

Xét hàm lồi một biến

$$\varphi(t) := f(x^0+td^0) = f\Big(3t,\ 1-\frac{3}{2}t\Big) = \frac{1}{2}\Big(3t\Big)^2+\frac{1}{2}\Big(1-\frac{3}{2}t\Big)^2 \quad \text{với } 0 \le t \le 1.$$

Vì

$$\varphi'(t) = \frac{45}{4}t - \frac{3}{2}.$$

nên

$$\varphi'(t) = 0 \Leftrightarrow t = \frac{2}{15}.$$

Vậy

$$t_0 = \frac{2}{15} \text{ và } x^1 = x^0 + t_0 d^0 = \Big(\frac{2}{5}\ ,\ \frac{4}{5}\Big)^T.$$

0.4. Vì $\nabla f(x^1) = \left(\frac{2}{5},\ \frac{4}{5}\right)^T \neq (0,\ 0)^T$ nên đặt $k := k+1 = 1$ và chuyển Bước lặp $k = 1$.

Bước lặp $k = 1$.

1.1. Giải:

$$\min\Big\{\ \langle \nabla f(x^1), x\rangle = \frac{2}{5}x_1 + \frac{4}{5}x_2\ \Big|\ x \in D\Big\}.$$

Bài toán này có hai phương án cực biên tối ưu là $(0\ ,\ 1)^T$ và $(3\ ,\ -\frac{1}{2})^T$. Giả sử ta chọn $u^1 = (0\ ,\ 1)^T$.

1.2. Khi đó

$$\Big\langle \nabla f(x^1), u^1 - x^1\Big\rangle = \Big(\nabla f(x^1)\Big)^T (u^1 - x^1) = \begin{pmatrix} \frac{2}{5} & \frac{4}{5}\end{pmatrix}\begin{pmatrix} -\frac{2}{5} \\ \frac{1}{5}\end{pmatrix} = 0.$$

Vậy $x^{opt} = x^1 = (\frac{2}{5}\ ,\ \frac{4}{5})^T$.

Nếu trong Bước 1.1 ta chọn $u^1 = (3\ ,\ -\frac{1}{2})^T$ thì

$$\Big\langle \nabla f(x^1), u^1 - x^1\Big\rangle = \Big(\nabla f(x^1)\Big)^T (u^1 - x^1) = \begin{pmatrix} \frac{2}{5} & \frac{4}{5}\end{pmatrix}\begin{pmatrix} \frac{13}{5} \\ -\frac{13}{10}\end{pmatrix} = 0.$$

Do đó ta cũng kết luận được nghiệm tối ưu $x^{opt} = x^1 = (\frac{2}{5}\ ,\ \frac{4}{5})^T$.

4.3.6 Phương pháp hàm phạt

Xét bài toán

$$\min\ f(x) \text{ v.đ.k } x \in D, \tag{P_3^{rb}}$$

trong đó D là tập compac xác định bởi

$$D = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0,\ i = 1, 2, \ldots, m\},$$

f và $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, là các hàm khả vi liên tục.

Ý tưởng cơ bản của phương pháp hàm phạt là thay vì giải bài toán tối ưu có ràng buộc (P_3^{rb}), ta giải một dãy bài toán tối ưu không ràng buộc với hàm mục tiêu phụ thuộc tham số α

$$\min\ f(x) + \alpha p(x)\ \text{ v.đ.k. }\ x \in \mathbb{R}^n,$$

trong đó $p(x)$ là *hàm phạt*, sao cho dãy nghiệm tương ứng của các bài toán không ràng buộc này hội tụ đến nghiệm tối ưu của bài toán (P_3^{rb}).

Mục này trình bày hai cách tiếp cận của phương pháp hàm phạt: phương pháp hàm phạt điểm ngoài (Exterior penalty function method) và phương pháp hàm phạt điểm trong (Interior penalty function method).

a. Phương pháp hàm phạt điểm ngoài

Trong phương pháp này, hàm phạt $p(x)$ được định nghĩa bởi

$$p(x) = \sum_{i=1}^{m} \Phi(g_i(x)),$$

trong đó Φ là hàm một biến liên tục và thỏa mãn

$$\Phi(y) = 0 \text{ nếu } y \leq 0 \;\text{ và }\; \Phi(y) > 0 \text{ nếu } y > 0.$$

Thông thường, hàm $p(x)$ có dạng

$$p(x) := \sum_{i=1}^{m} \max\Big\{0,\ g_i(x)\Big\} \quad \text{hoặc} \quad p(x) := \sum_{i=1}^{m} \Big[\max\{0,\ g_i(x)\}\Big]^2.$$

Hàm mục tiêu của dãy bài toán tối ưu không ràng buộc tương ứng với bài toán (P_3^{rb}) là

$$\varphi(x, \alpha_k) = f(x) + \alpha_k p(x),$$

trong đó dãy tham số $\{\alpha_k\}$ là dãy số dương, đơn điệu tăng đến ∞.

Đại lượng $\alpha_k p(x)$ là *lượng phạt*. Dễ thấy rằng khi x là phương án chấp nhận được của bài toán (P_3^{rb}) thì nó không bị phạt (lượng phạt bằng 0), nhưng khi x không phải là phương án chấp nhận được thì nó phải chịu một lượng là $\alpha_k p(x)$.

Ví dụ 4.18. Xét bài toán:

$$\min\{x \mid -x + 5 \leq 0\}.$$

Bài toán này có $f(x) = x$, $m = 1$ và $g_1(x) = -x + 5 \leq 0$. Đặt

$$p(x) = \Big[\max\{0,\ g_1(x)\}\Big]^2.$$

Khi đó

$$p(x) = \begin{cases} 0 & \text{nếu } x \geq 5 \\ (-x+5)^2 & \text{nếu } x < 5. \end{cases}$$

Thuật toán 6.14.
Bước chuẩn bị. Cho số $\varepsilon > 0$ đủ bé (để kiểm tra điều kiện dừng của thuật toán). Chọn một điểm $x^1 \in \mathbb{R}^n$. Chọn một tham số phạt $\alpha_1 > 0$ và một số $\mu > 1$. Đặt $k = 1$.

Bước lặp k. $(k = 1, 2, \dots)$
Bước k_1. Xuất phát từ x^k, giải bài toán tối ưu không ràng buộc:

$$\min\ \varphi(x, \alpha_k) \ \text{ v.đ.k } \ x \in \mathbb{R}^n.$$

Gọi nghiệm của bài toán này là x^{k+1}.

Bước k_2. **If** $\alpha_k p(x^{k+1}) < \varepsilon$ **Then** Dừng thuật toán
(*lấy* x^{k+1} *là nghiệm tối ưu của bài toán* (P_3^{rb}))
Else Chuyển Bước k_3.

Bước k_3. Đặt tham số phạt mới $\alpha_{k+1} := \mu\alpha_k$.
Đặt $k := k + 1$ và chuyển về Bước lặp k.

Định lý 4.21. *Thuật toán 6.14 có các tính chất sau:*

i) $\varphi(x, \alpha) \geq f(x)$ *với mọi* $\alpha > 0$*;*

ii) *Dãy* $\{\varphi(x^k, \alpha_k)\}$ *là đơn điệu tăng;*

iii) $\{\varphi(x^k, \alpha_k)\}$ *hội tụ đến giá trị tối ưu của bài toán* (P_3^{rb}) *khi* $\{\alpha_k\} \to \infty$.

Chứng minh. Xem [27], trang 59. □

b. Phương pháp hàm phạt điểm trong

Phương pháp này được sử dụng khi biết trước một điểm trong x^1 của tập chấp nhận được, tức $g_i(x^1) < 0$ với mọi $i = 1, \dots, m$. Hàm phạt $p(x)$ phải thỏa mãn tính chất:

i) Không âm và liên tục trên tập $\text{int}D = \{x \in \mathbb{R}^n | g_i(x) < 0, \quad i = 1, \dots, m\}$;

ii) $p(x) \to +\infty$ khi $g_i(x) \to 0^-$.

Vì vậy, người ta còn gọi hàm phạt $p(x)$ này là *hàm chắn* (barrier function). Hai hàm chắn điển hình thường được sử dụng là

$$p(x) = -\sum_{i=1}^{m} \ln[-g_i(x)] \ \text{ và } \ p(x) = -\sum_{i=1}^{m} \frac{1}{g_i(x)}.$$

Phương pháp hàm phạt điểm trong xuất phát từ một điểm trong x^1 của tập chấp nhận được D, giải một dãy các bài toán tối ưu không ràng buộc:

$$\min\ \psi(x, \alpha_k) = f(x) + \alpha_k p(x) \quad \text{v.đ.k. } \ x \in \mathbb{R}^n,$$

trong đó $\{\alpha_k\}$ là dãy tham số dương, giảm đơn điệu về 0.

Thuật toán 6.15.
Bước chuẩn bị. Cho số $\varepsilon > 0$ đủ bé (để kiểm tra điều kiện dừng của thuật toán). Xác định một điểm $x^1 \in D$ thỏa mãn $g_i(x^1) < 0$, $i = 1, \ldots, m$. Chọn một tham số phạt $\alpha_1 > 0$ và một số $\mu \in (0, 1)$. Đặt $k = 1$.

Bước lặp k. $(k = 1, 2, \ldots)$
Bước k_1. Xuất phát từ x^k, giải bài toán tối ưu không ràng buộc:

$$\min\ \psi(x, \alpha_k) \ \text{ v.đ.k. } \ x \in \mathbb{R}^n$$

nhận được nghiệm x^{k+1}.

Bước k_2. **If** $\alpha_k p(x^{k+1}) < \varepsilon$ **Then** Dừng thuật toán
(lấy x^{k+1} là nghiệm tối ưu của bài toán (P_3^{rb}))
Else Chuyển Bước k_3.

Bước k_3. Đặt tham số phạt mới $\alpha_{k+1} := \mu\alpha_k$.
Đặt $k := k + 1$ và chuyển về Bước lặp k.

Định lý 4.22. *Thuật toán 6.15 có các tính chất sau:*

i) $\psi(x, \alpha) \geq f(x)$ *với mọi* $\alpha > 0$*;*

ii) $g_i(x^k) < 0, \ \ i = 1, \ldots, m$*;*

iii) $\{\psi(x^k, \alpha_k)\}$ *hội tụ đến giá trị tối ưu của bài toán* (P_3^{rb}) *khi* $\{\alpha_k\} \to 0$.

Chứng minh. Xem [27], trang $57 - 58$. □

Ví dụ 4.19. Xét bài toán:

$$\begin{aligned} \min\ & f(x) = x_1 - 2x_2 \\ \text{v.đ.k. } & g_1(x) = -1 - x_1 + x_2^2 \leq 0 \\ & g_2(x) = -x_2 \leq 0 \end{aligned}$$

Ta chọn hàm chắn có dạng

$$p(x) = -\ln[-g_1(x)] - \ln[-g_2(x)] = -\ln(1 + x_1 - x_2^2) - \ln(x_2).$$

Khi đó, hàm mục tiêu của các bài toán không ràng buộc tương ứng là

$$\psi(x, \alpha) = x_1 - 2x_2 - \alpha\ln(1 + x_1 - x_2^2) - \alpha\ln(x_2),$$

trong đó α là tham số chắn, dương và giảm đơn điệu về 0.

Với mỗi tham số α cụ thể, bài toán không ràng buộc

$$\min\{\psi(x,\alpha) \mid x \in \mathbb{R}^n\}$$

có hàm mục tiêu $\psi(x,\alpha)$ là hàm lồi chặt theo biến x nên, theo Mệnh đề 6.5 và Định lý 6.6, nó có nghiệm cực tiểu duy nhất cũng chính là điểm dừng. Ta có:

$$\begin{aligned}
\psi'_{x_1}(x,\alpha) &= 1 - \frac{\alpha}{1+x_1-x_2^2},\\
\psi'_{x_2}(x,\alpha) &= -2 + \frac{2\alpha x_2}{1+x_1-x_2^2} - \frac{\alpha}{x_2}.
\end{aligned}$$

Các điểm $x = (x_1, x_2)^T$ thuộc phần trong của tập chấp nhận được có

$$1 + x_1 - x_2^2 > 0 \quad \text{và} \quad x_2 > 0.$$

Giải hệ phương trình

$$\nabla_x \psi(x,\alpha) = 0 \quad \Leftrightarrow \quad \begin{cases} \displaystyle 1 - \frac{\alpha}{1+x_1-x_2^2} = 0 \\ \displaystyle -2 + \frac{2\alpha x_2}{1+x_1-x_2^2} - \frac{\alpha}{x_2} = 0 \end{cases}$$

ta nhận được điểm dừng $x^*(\alpha) = \Big(x_1^*(\alpha), x_2^*(\alpha)\Big)^T$ với

$$x_1^*(\alpha) = \frac{\sqrt{1+2\alpha}+3\alpha-1}{2}, \qquad x_2^*(\alpha) = \frac{1+\sqrt{1+2\alpha}}{2}.$$

Vì tham số $\alpha > 0$ và giảm dần về 0 nên

$$\begin{aligned}
\lim_{\alpha\to 0^+} x_1^*(\alpha) &= \frac{\sqrt{1+2(0)}+3(0)-1}{2} = 0,\\
\lim_{\alpha\to 0^+} x_2^*(\alpha) &= \frac{1+\sqrt{1+2(0)}}{2} = 1.
\end{aligned}$$

Ta có $(0,1)^T$ thuộc tập chấp nhận được và là nghiệm tối ưu của bài toán cần giải.

Bảng 6.2 trình bày giá trị của $x_1^*(\alpha)$ và $x_2^*(\alpha)$ với các tham số $\alpha_1 = 1$, $\mu = 10^{-1}$

$$\alpha_{k+1} = \mu\alpha_k, \quad k = 1, 2, \ldots, 6.$$

Rõ ràng $\left\{x^*(\alpha_k) = \Big(x_1^*(\alpha_k)\ ,\ x_1^*(\alpha_k)\Big)^T\right\}$ hội tụ đến nghiệm tối ưu của bài toán đang xét.

Bảng 6.2

k	α_k	$x_1^*(\alpha_k)$	$x_2^*(\alpha_k)$
1	10^0	1.8660254	1.3660254
2	10^{-1}	0.1977226	1.0477226
3	10^{-2}	0.0199752	1.0049752
4	10^{-3}	0.0019998	1.0004998
5	10^{-4}	0.0002000	1.0000500
6	10^{-5}	0.0000200	1.0000050
7	10^{-6}	0.0000020	1.0000005

Bài tập Chương 6

1. Cho $g_i, i = 1, \dots, m$ là các hàm lồi và $h_j,\ j = 1, \dots, k$ là các hàm tuyến tính afin xác định trên $\mathbb{R}^n$. Chứng minh rằng tập hợp sau đây là lồi:

$$M = \{x \in \mathbb{R}^n \mid g_i(x) \le 0, i = 1, \dots, m;\ h_j(x) = 0,\ j = 1, \dots, k\}.$$

2. Các hàm số sau có phải hàm lồi không? Hãy giải thích bằng ít nhất hai cách:

   i) $f(x) = |x|$ với $x \in \mathbb{R}$;

   ii) $f(x) = e^x + 1 + 5x$ với $x \in \mathbb{R}$;

   iii) $f(x) = -\ln x + 3x^3$ với $0 < x < +\infty$;

   iv) $f(x) = \arctan x$ với $0 < x < +\infty$;

   v) $f(x) = \begin{cases} |x| & \text{nếu } x \le 1 \\ (x-2)^2 - 1 & \text{nếu } x > 1 \end{cases}$ với $x \in \mathbb{R}$;

   vi) $f(x) = \begin{cases} e^x & \text{nếu } x < 0 \\ 3 & \text{nếu } x = 0 \end{cases}$ với $-\infty < x \le 0$.

3. Trong các hàm số sau, hàm nào là hàm lồi, hàm lõm hay không lồi, không lõm. Vì sao?

   i) $f(x_1, x_2) = x_1 e^{-(x_1+x_2)}$;

   ii) $f(x_1, x_2) = x_1^2 + 2x_1x_2 - 10x_1 + 5x_2$;

   iii) $f(x_1, x_2) = -x_1^2 - 5x_2^2 + 2x_1x_2 + 9x_1 - 8x_2$;

   iv) $f(x_1, x_2, x_3) = x_1x_2 + 2x_1^2 + x_2^2 + 2x_3^2 - 6x_1x_3 + 3x_2x_3$;

v) $f(x_1, x_2, x_3) = -x_1^2 - 3x_2^2 - 2x_3^2 + 4x_1x_2 + 2x_1x_3 + 4x_2x_3$;

iv) $f(x_1, x_2, x_3) = x_1^2 + 2x_1x_2 + 4x_1x_2 + 3x_2^2 + x_2x_3 + 6x_3^2$;

vii) $f(x_1, x_2, x_3) = x_1^2 + 4x_1x_2 + 4x_2^2 + 2x_1x_3 + 4x_2x_3$.

4. Chứng minh rằng hàm tuyến tính afin bị chặn trên (hay bị chặn dưới) trên một tập afin thì chỉ có thể đồng nhất bằng hằng số.

5. Chứng minh rằng hàm $f : \mathbb{R}^n \to \mathbb{R}$ là hàm afin khi và chỉ khi f vừa là hàm lồi, vừa là hàm lõm.

6. Cho tập lồi khác rỗng $D \subset \mathbb{R}^n$ và hàm số $f : D \to \mathbb{R}$. Chứng minh rằng:

   i) f là hàm lõm khi và chỉ khi $\text{hypo}(f)$ là tập lồi.

   ii) Nếu f là hàm lõm thì tập $L^\alpha(f) = \{x \in D \mid f(x) \geq \alpha\}$, với $\alpha \in \mathbb{R}$, là tập lồi.

7. Hãy xác định siêu phẳng đi qua các điểm $(1, 1, 1, 1)^T$, $(2, 0, 1, 0)^T$, $(0, 2, 0, 1)^T$ và $(1, 1, -1, 0)^T$ thuộc $\mathbb{R}^4$.

8. Cho C là tập lồi đóng và $y \notin C$. Chứng minh rằng $x^* \in C$ là điểm gần y nhất khi và chỉ khi $\langle x - y, x^* - y\rangle \geq \|x^* - y\|^2$ với mọi x thuộc C.

9. Cho C_1 và C_2 là hai tập lồi trong $\mathbb{R}^n$. Chứng minh rằng tồn tại một siêu phẳng tách chặt C_1 và C_2 khi và chỉ khi

$$\inf\{\|x - y\| \; |x \in C_1, \; y \in C_2\} > 0.$$

10. Cho A là ma trận cấp $m \times n$ và $c \in \mathbb{R}^m \setminus \{0\}$. Chứng minh rằng có một và chỉ một hệ phương trình trong hai hệ phương trình sau có nghiệm:

    i) $Ax = c$;

    ii) $A^T y = 0, \; \langle c, y\rangle = 1$.

11. Cho S là tập lồi khác rỗng trong $\mathbb{R}^n$. Chứng minh rằng hàm f được định nghĩa như sau

$$f(y) := \inf\{\|y - x\| \; |x \in S\}$$

    là hàm lồi.

12. Cho điểm $x^0 = (2, 1)^T \in \mathbb{R}^2$, véc tơ $d = (3, -1)^T$ và hàm số

$$f(x) = x_2 e^{-(x_1 + x_2)}.$$

    i) Vẽ tập các điểm $\{x = x^0 + td \mid t \geq 0\}$.

    ii) Véc tơ d có phải là hướng giảm của hàm f tại x^0 không? Vì sao?

13. Cho hàm toàn phương $f(x) = \frac{1}{2}x^T Ax + b^T x + c$, với A là ma trận cấp $n \times n$, đối xứng, xác định dương, véc tơ $b \in \mathbb{R}^n$ và số thực $c \in \mathbb{R}$. Nói rằng việc giải bài toán $\min\{f(x) \mid x \in \mathbb{R}^n\}$ được đưa về việc tìm nghiệm của hệ phương trình tuyến tính
$$Ax + b = 0$$
có đúng không? Vì sao? Cho ví dụ cụ thể.

14. Cho hàm toàn phương $f(x) = \frac{1}{2}x^T Ax - b^T x$, với A là ma trận cấp $n \times n$, đối xứng, véc tơ $b \in \mathbb{R}^n$. Xét bài toán $\min\{f(x)|x \in \mathbb{R}^n\}$.

    i) Viết điều kiện cần bậc một và bậc hai của nghiệm tối ưu địa phương. Với điều kiện gì thì hàm $f(x)$ có một điểm dừng?
    ii) Khi nào thì hàm $f(x)$ có cực tiểu địa phương?
    iii) Với điều kiện nào thì hàm $f(x)$ có một điểm dừng nhưng đó không phải cực tiểu địa phương, cũng không phải cực đại địa phương?

15. Cho hàm số $f(x) = \alpha x_1^2 + x_2^2 - 2x_1x_2 - 2x_2$, trong đó α là một số thực.

    i) Xác định điểm dừng của $f(x)$ với mỗi giá trị của α?
    ii) Xác định điểm cực đại địa phương và cực tiểu địa phương tương ứng với một giá trị cụ thể nào đó của α. Hàm $f(x)$ đó có điểm cực đại toàn cục hoặc cực tiểu toàn cục không? Vì sao?

16. Giải $\min\{f(x),\ x \in \mathbb{R}^n\}$ với

    i) $f(x) = 15 - 12x - 25x^2 + 2x^3$ (tức $n = 1$);
    ii) $f(x) = 3x^3 + 7x^2 - 15x - 3$ (tức $n = 1$);
    iii) $f(x_1, x_2) = 8x_1^2 + 3x_1x_2 + 7x_2^2 - 25x_1 + 31x_2 - 29$;
    iv) $f(x_1, x_2) = 2x_1^2 + x_2^2 - 2x_1x_2 + 2x_1^3 + x_1^4$;

17. Cho hàm số $f(x_1, x_2) = x_1^2 - x_1x_2 + 2x_2^2 - 2x_1 + e^{x_1+x_2}$.
Điểm $x = (0, 0)^T$ có phải điểm cực tiểu địa phương của hàm f không? Nếu không, tìm một hướng d sao cho hàm f giảm theo d?

18. Cho $f(x)$ là hàm lồi khả vi, xác định trên tập lồi đa diện $X \subset \mathbb{R}^n$. Chứng minh rằng $x^* \in X$ là nghiệm cực tiểu của bài toán $\min\{f(x), x \in X\}$ khi và chỉ khi
$$0 = \min_{x \in X} \langle \nabla f(x^*), x - x^* \rangle.$$
Kết luận này có ý nghĩa gì?

19. Trình bày phương pháp hướng giảm nhanh nhất.

i) Xét bài toán:

$$\min\{f(u) = u_1^2 + 2u_2^2 : u \in \mathbb{R}^2\}.$$

Lấy $u^0 = (2,1)^T$. Tính ba điểm tiếp theo u^1, u^2, u^3 theo phương pháp hướng giảm nhanh nhất. Chứng minh rằng dãy $\{u^k\}$ sinh ra bởi thuật toán hướng giảm nhanh nhất với thủ tục tìm chính xác theo tia có

$$u^k = \left(\frac{1}{3}\right)^k \begin{pmatrix} 2 \\ (-1)^k \end{pmatrix} \quad \text{và} \quad f(u^{k+1}) = \frac{f(u^k)}{9}.$$

ii) Xét bài toán:

$$\min\{f(x) = 4x_1^2 + 2x_2^2 + 4x_1x_2 - 3x_1 \mid x \in \mathbb{R}^2\}.$$

Lấy $x^0 = (2,2)^T$. Tính ba điểm tiếp theo x^1, x^2, x^3 theo phương pháp giảm nhanh nhất.

iii) Cho

$$f(x) = \frac{1}{2}x^T Ax - b^T x$$

với

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \gamma & 0 \\ 0 & 0 & \gamma^2 \end{pmatrix} \quad \text{và} \quad b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Xuất phát từ $x^0 = (0,\ 0,\ 0)^T$, áp dụng Thuật toán 6.1, trình bày ba bước lặp giải bài toán $\min\{f(x)|x \in \mathbb{R}^3\}$ với $\gamma = 1,\ 10,\ 100,\ 1000$.

20. Nêu định nghĩa hướng chấp nhận được tại $x^* \in D \subset \mathbb{R}^n$. Chứng minh rằng nếu D là tập lồi thì $x - x^*$ là hướng chấp nhận được tại x^* với mọi $x \in D$.

21. Phát biểu điều kiện cần của điểm cực tiểu địa phương của bài toán

$$\min\{f(x), x \in \mathbb{R}^n\}$$

và bài toán

$$\min\{f(x)\ x \in D \subset \mathbb{R}^n\}.$$

Khi f là hàm lồi và D là tập lồi thì có kết luận gì? Lấy ví dụ để cùng hàm mục tiêu f nhưng hai bài toán này có nghiệm tối ưu khác nhau.

22. Hàm một biến $f(x) = -\ln(x)$ có phải hàm lồi không? Hãy giải thích bằng ít nhất hai cách.

23. Xét bài toán $\min\{f(x) = \frac{1}{2}x^TQx + c^Tx \mid Ax \le b\}$, trong đó Q là ma trận đối xứng xác định dương, không suy biến cấp $(n \times n)$, ma trận A cấp $(m \times n)$, véc tơ $c, x \in \mathbb{R}^n$ và $b \in \mathbb{R}^m$. Viết điều kiện Karush-Kuhn-Tucker cho bài toán này.

24. Sử dụng Định lý Karush-Kuhn-Tucker tìm nghiệm tối ưu và giá trị tối ưu của bài toán của các bài toán sau. Giải thích chi tiết từng điều kiện áp dụng và cách lấy nghiệm.

    i) $\min\{f(x) = x_1 \mid (x_1 - 1)^2 + x_2^2 \le 1,\ x_1^2 + x_2^2 \le 2\}$.
    ii) $\min\{f(x) = x_1^2 + x_2^2 - 8x_1 - 4x_2 \mid x_1 + x_2 \le 2,\ x_1 \ge 0,\ \ x_2 \ge 0\}$.

25. Trình bày phương pháp nhân tử Langrange giải bài toán trơn với ràng buộc đẳng thức. Giải các bài toán:

    i) $\min\{f(x) = x_1^2 + x_2^2 \mid x_1 + x_2 = 10\}$.
    ii) $\min\{f(x) = x_1^2 - 2x_1 + x_2^2 - x_3^2 + 4x_3 \mid x_1 - x_2 + 2x_3 = 2\}$.
    iii) $\min\{f(x) = 3x_1 + 4x_2 \mid (x_1 + 1)^2 + x_2^2 = 1,\ (x_1 - 1)^2 + x_2^2 = 1\}$.

26. Trình bày thuật toán Frank-Wolfe. Xét bài toán:

$$\min\{f(u) = \frac{1}{2}u_1^2 + u_2^2 + u_1 \mid u_1 \ge 0, u_2 \ge 0, u_1 + u_2 \le 10\}.$$

    Cho $u^0 = (1, 1)^T$. Xây dựng một vài phần tử của dãy lặp $\{u^k\}$ theo thuật toán Frank-Wolfe.

27. Hãy lập chương trình thể hiện thuật toán tìm kiếm theo đơn hình (Thuật toán 6.9) và giải các bài tập sau để đánh giá hiệu quả của thuật toán.

    i) $f(x_1, x_2) = (x_1 - 3.8)^2 + (x_2 - 1.8)^2 \ \longrightarrow\ \min$.
    ii) $f(x) = \sum_{j=1}^n j(x_j - 10)^2 \ \longrightarrow\ \min$ với $n = 10, 20, 30$.
    iii) $f(x_1, x_2) = x_1^4 + x_2^4 + 2x_1^2x_2^2 - 4x_1 + 3 \ \longrightarrow\ \min$.

28. Xét bài toán:

$$\min\{\ x_1 \mid -(x_1 - 2)^2 - 4(x_2 - 2)^2 + 4 \le 0\}.$$

    Điểm $x^0 = (4, 2)^T$ có phải là điểm KKT không? Điểm x^0 có phải nghiệm tối ưu địa phương của bài toán đang xét không? Giải thích?

29. Chứng minh rằng điểm $x^0 = (2, 4, 0)^T$ là nghiệm tối ưu địa phương của bài toán:

$$\min\{f(x) = x_1^2 + x_2^2 + x_3^2 + x_1x_2 + x_2x_3 \mid x_1 \ge 2,\ x_2 \ge 4,\ x_3 \ge 0\}.$$

    Điểm x^0 có phải là nghiệm tối ưu toàn cục không? Vì sao?

Chương 5

Tính chất của tập nghiệm, điều kiện tối ưu và đối ngẫu

Chương 6

Qui hoạch toàn cục

Phụ lục

Tài liệu tham khảo