Chương 1.

BÀI TOÁN TỐI ƯU VÀ CÁC KIẾN THỨC CƠ SỞ

Trong cương này, chúng tôi lần lượt trình bày các vấn đề của lý thuyết tối ưu và các khái niệm, kết quả cơ bản nhất được dùng cho các chương sau, cụ thể là trình bày:

- Mục đích, ý nghĩa và quy luật hoạt động của trạng thái (vật thể) trong tự nhiên.
- Bài toán tối ưu và các hướng nghiên cứu của tối ưu hóa.
- Các khái niệm cơ bản như: không gian tuyến tính, tuyến tính định chuẩn, không gian Hibert, không gian Banach, biến phân, đạo hàm, tập lồi, hàm lồi và các định lý cơ bản liên quan đến các khái niệm trên.
- Nhắc lại bài toán Quy hoạch tuyến tính và thuật toán đơn hình và sử dụng thư viện MATHLAB để giải bài toán này.

1.1. NHỮNG BÀI TOÁN KINH ĐIỂN VÀ Ý NGHĨA

1.1.1 Những ví dụ

Ví dụ 1.1.1. *Bài toán đẳng chu (thế kỷ thứ 5 trước công nguyên) Tìm đường cong khép kín trên mặt phẳng có chu vi cho trước sao cho hình nó tạo ta có diện tích lớn nhất.*

Ví dụ 1.1.2. *(Euclid 365 trước công nguyên) Cho tam giác ABC. Hãy tìm điểm E trên cạnh BC sao cho hình bình hành $ADEF$, với D, F nằm trên AB và AC, có diện tích lớn nhất.*

Ví dụ 1.1.3. *(Heron 75 trước công nguyên) Tìm điểm C trên đường thẳng cho trước sao cho tổng khoảng cách từ C đến A và B là lớn nhất.*

Ví dụ 1.1.4. *(Tartaglia 1500-1557) Tìm hai số tự nhiên a, b thỏa mãn $a + b = 8$ sao cho $ab(b - a)$ lớn nhất.*

Ví dụ 1.1.5. *(Kepler 1571-1630) Tìm hình trụ nội tiếp trong hình cầu cho trước sao cho thể tích lớn nhất.*

Ví dụ 1.1.6. *(Fermat 1601-1665) Tìm hai cạnh góc vuông bằng một số cho trước sao cho diện tích lớn nhất.*

Ví dụ 1.1.7. *(Steiner 1796-1863) Một đa giác được gọi là nội tiếp trong một đa giác ngoại tiếp nếu nó nằm trong đó và trên mỗi cạnh của đa giác ngoại tiếp có ít nhất một điểm của đa giác nội tiếp. Hãy tìm đa giác nội tiếp có chu vi nhỏ nhất.*

1.1.2 Ý nghĩa thực tiễn

Các ví dụ trên có tính chất hàn lâm, không mang ý nghĩa thực tế. Do đó, trong mục này chúng ta sẽ chỉ ra rằng mọi trạng thái của các vật thể trong tự nhiên đều hoạt động tuân theo một quy luật tối ưu nào đó và đồng thời đưa ra một số ví dụ minh họa trong các lĩnh vực ứng dụng quan trọng của lý thuyết tối ưu.

1.1.3 Hoạt động của trạng thái trong tự nhiên

Câu hỏi đặt ra ở đây là, các trạng thái (động hay tĩnh) của vật thể trong tự nhiên hoạt động tuân theo quy luật nào?
Ngay từ thế kỷ XVIII L. Euler đã viết:**" Vì thế giới được thiết lập một cách hoàn hảo nhất và vì nó là sản phẩm của đấng sáng tạo tinh thông nhất, nên không thể tìm thấy cái gì mà không mang theo tính chất cực đại hay cực tiểu nào đó".** Như vậy:

- Ngay thế kỷ XVIII các quy luật cơ bản của tự nhiên đã được phát biểu dưới dạng các nguyên lý cực trị.

- Mọi diễn biến trong tự nhiên đều tuân theo một nguyên lý tối ưu nào đó.

Những nguyên lý sau thể hiện khẳng định trên.
1. **(Nguyên lý Fermat)** ***Ánh sáng chọn đường đi mà thời gian đi là ngắn nhất.***
2. (Nguyên lý cực tiểu thế năng Dirichlet) ***Một hệ bảo toàn (năng lượng) có trạng thái cân bằng ổn định khi và chỉ khi thế năng của nó đạt giá trị cực tiểu.*** Nói cách khác: khi không bị tác động từ bên ngoài, một vật nằm lại ở vị trí mà thế năng nhỏ nhất (so với các vị trí lân cận)
3. **(Nguyên lý tác động dừng** (hay **nguyên lý tác động nhỏ nhất**)) ***Chuyển động giữa hai thời điểm*** t_0, t_1 ***sẽ diễn ra sao cho tích phân tác động***

$$W = \int_{t_0}^{t_1} (T - U)dt$$

đạt giá trị thấp nhất ($\to \min$) ***hay trạng thái điểm dừng, trong đó*** T ***là động năng,*** U ***là thế năng,*** $T - U$ ***là thế động lực.***

1.1.4 Các bài toán thực tế

Ví dụ 1.1.8. *(Bài toán thanh uốn) Cho thanh đàn hồi có độ dài l, modul đàn hồi E và mô men quán tính I Khi dựng đứng thanh đàn hồi và tác dụng lên đầu trên một lực P thì nó bị cong đi. Gọi x là góc giữa trục thanh uốn và phương thẳng đứng. Năng lượng tương ứng với công sinh ra biến dạng trong thanh uốn là*

$$\frac{1}{2}\int_0^l EI\dot{x}^{(2)}(s)ds.$$

Thế năng của trọng lực P là

$$P\int_0^l cosx(s)ds.$$

Do đó, tổng thế năng của thanh uốn là:

$$\frac{1}{2}\int_0^l EI\dot{x}^{(2)}(s)ds + P\int_0^l cosx(s)ds.$$

Theo nguyên lý cực tiểu thế năng Dirichlet thì hình dạng ổn định của thanh uốn là trạng thái có thế năng nhỏ nhất. Do đó để tìm trạng thái đó ta phải tìm $x(\cdot)$ sao cho tổng thế năng của thanh uốn là nhỏ nhất.

Ví dụ 1.1.9. *(Bài toán lựa chọn đầu tư) Một trong những ứng dụng nổi trong kinh tế là bài toán lựa chọn đầu tư do H. M. Markowitz đề xuất. Bài toán phát biểu như sau: Phân phối vốn qua n chứng khoán (asset) có sẵn để có thể giảm thiểu rủi ro và tối đa lợi nhuận, tức là tìm véc tơ tỉ lệ $x \in D,\ D := \{x = (x_1, x_2, \ldots, x_n) \mid \sum_{j=1}^n x_j = 1\}$ để $f(x) = \omega x^T Ax - \rho^T x$ đạt giá trị nhỏ nhất, trong đó $x_j, j = 1, \ldots, n$, là tỷ lệ chứng khoán thứ j trong danh mục đầu tư, ω là tham số rủi ro, $A \in I\!R^{n\times n}$ là ma trận hiệp phương sai, $\rho \in I\!R^n$ là véc tơ lợi nhuận kỳ vọng.*

Ví dụ 1.1.10. *(Bài toán tối ưu chi phí phát điện) Một vấn đề thường được nghiên cứu của phát điện tối ưu, tức là bài toán phân bố lượng điện năng cho từng tổ máy phát nhiệt điện sao cho tổng chi phí (giá thành) là cực tiểu, đồng thời vẫn đáp ứng được nhu cầu lượng điện năng và thoả mãn ràng buộc về công suất phát ra của mỗi tổ máy. Người ta thường giả thiết hàm chi phí tổng cộng (bao gồm các chi phí nhiên liệu (fuel cost), chi phí tải sau (load-following cost), chi phí dự phòng quay (sprinning-reserve cost), chi phí dự phòng bổ sung (supplemental-reserve cost), chi phí tổn thất phát và truyền dẫn điện năng) là hàm toàn phương, lồi ngặt và có dạng*

$$F(P) = \sum_{i=1}^{n} F_i(P_i),$$

trong đó n là số tổ máy phát, $P := (P_1, P_2, \ldots, P_n)$, $P_i \in [P_{i\min}, P_{i\max}]$ là lượng điện năng phát ra của tổ máy thứ i, $P_{i\min}, P_{i\max}$ là công suất phát nhỏ nhất và lớn nhất của tổ máy phát thứ i, $F_i(P_i) = a_i + b_iP_i + c_iP_i^2$ là hàm chi phí của tổ máy phát thứ i và a_i, b_i, c_i là các hệ số giá của tổ máy phát thứ $i \in \{1, 2, \ldots, n\}$.

Đặc biệt, nếu hiệu ứng điểm-van được xét đến thì hàm chi phí toàn phương phải được hiệu chỉnh bởi tổng hữu hạn các hàm dạng sin*, tức là*

$$F(P) = \sum_{i=1}^{n} \big(F_i(P_i) + |e_i \sin(f_i(P_{i\min} - P_i))|\big),$$

trongđó e_i, f_i *là các hệ số hiệu ứng điểm-van.*

1.2. LÝ THUYẾT TỐI ƯU

1.2.1 Quá trình hình thành và phát triển

1. Thế kỷ XVIII, một hướng nghiên cứu bài toán cực trị hàm mục tiêu là phiếm hàm tích phân gọi là Phép tính biến phân.

2. Những năm 30-40 của thế kỷ XX xuất hiện Lý thuyết Quy hoạch tuyến tính.

3. Những năm 50- thế kỷ XX xuất hiện Quy hoạch lồi.

4. Từ những những năm 70 của thế kỷ XX hình thành nhiều hướng nghiên cứu khác nhau như Tối ưu không lồi, tối ưu phi tuyến, tối ưu rời rạc, tối ưu tổ hợp và tối ưu đa mục tiêu.

5. Từ những năm 50-60 của thế kỷ XX xuất hiện Lý thuyết điều khiển được và điều khiển tối ưu.

1.2.2 Mô hình toán học

Cho $f : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$, với X là không gian nào đó. Bài toán tối ưu phát biểu như sau:

$$f(x) \to \inf(\sup) \quad \text{với ràng buộc} \quad x \in D \subset X, \tag{1.1}$$

trong đó:

1. Hàm $f(x)$ gọi là hàm mục tiêu,

2. X gọi là không gian chấp nhận được,

3. D là miền chấp nhận được, hay là miền ràng buộc,

4. $x \in D$ gọi là nghiệm chấp nhận được.

5. Điểm x^* tại đó f nhận giá trị tối ưu, tức là:

$$f(x^*) \leq f(x), \forall x \in D \text{ hay } f(x^*) \geq f(x), \forall x \in D$$

   được gọi là nghiệm tối ưu toàn cục.

6. Trong trường hợp X được trang bị topo (không gian tuyến tính định chuẩn là một trường hợp riêng), nếu tồn tại lân cận V của điểm x^* sao cho

$$f(x^*) \leq f(x), \forall x \in D \cap V \text{ hay } f(x^*) \geq f(x), \forall x \in D \cap V$$

   thì x^* gọi là nghiệm tối ưu địa phương.

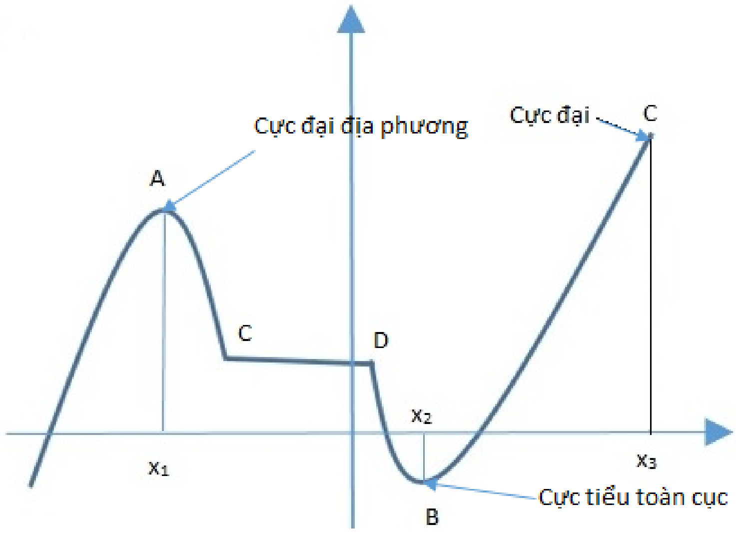
7. Nếu $D = X$ thì bài toán tối ưu trên gọi là bài toán tối ưu không ràng buộc, ngược lại gọi là bài toán tối ưu bị ràng buộc.

8. Điều kiện $x \in D$ thường xuất hiện ở các dạng sau (có thể cùng lúc ở cả 3 dạng):

   - Ràng buộc đẳng thức: $F(x) = 0$ với $F : X \to Y$.
   - Ràng buộc bất đẳng thức: $f_i(x) \leq 0$ với $f_i : X \to I\!R, i = 1, \ldots, m$.
   - Ràng buộc bao hàm thức: $x \in A, A \subset X$ với A cho trước.

1.2.3 Phân loại bài toán tối ưu

1. ***Quy hoạch tuyến tính***: Hàm mục tiêu và các hàm ràng buộc đều là các hàm tuyến tính. Như vậy miền chấp nhận được là một tập lồi đa diện.



Hình 1.1: **Cực đại, cực tiểu, địa phương, toàn cục**

2. ***Quy hoạch phi tuyến (Tối ưu phi tuyến)***: Tối thiểu có hàm mục tiêu hoặc hàm ràng buộc là phi tuyến. Tối ưu phi tuyến bao gồm: Tối ưu trơn (hàm mục tiêu và ràng buộc là trơn), Tối ưu lồi (hàm mục tiêu và ràng buộc là lồi), Tối ưu không lồi (hàm mục tiêu hoặc miền chấp nhận được không lồi).

3. ***Tối ưu rời rạc hay tối ưu tổ hợp***: Miền chấp nhận được là một tập rời rạc. Trường hợp các biến số nhận giá trị nguyên là bài toán ***quy hoạch nguyên***.

4. ***Tối ưu đa mục tiêu***: Mục tiêu gồm nhiều hàm không hòa hợp nhau. Tối ưu đa mục tiêu cũng được phân chia thành nhiều bài toán con khác nhau tùy theo tính chất của hàm mục tiêu và tập ràng buộc.

5. ***Quy hoạch ngẫu nhiên***: Tức là bài toán tối ưu mà các tham số trong đó không có giá trị xác định mà được mô tả bởi tham số xác suất.

6. ***Quy hoạch động***: Tức là bài toán tối ưu mà các đối tượng được xét có thể chia ra nhiều giai đoạn hoặc qua trình phát triển theo thời gian. Ngoài ra còn nhiều bài toán tối ưu hóa khác như: Quy hoạch

Lípshitz, quy hoạch nón, tối ưu không trơn ...

Điều quan trọng ở đây là lúc đầu người ta tưởng như các hướng nghiên cứu trên hoàn toàn riêng, nhưng dần dần người ta phát hiện ra nhiều điểm tương đồng. Do đó thúc đẩy đi tìm những nét đặc trưng chung cho các bài toán cực trị và dẫn đến sự hình thành lý thuyết các bài toán cực trị.

Nhận xét 1.2.1. *Nếu* $f(x)$ *lồi thì* $-f(x)$ *là lõm và* $f(x) \to \sup$ *tương đương với* $-f(x) \to \inf$ *nên bài toán* (1.1) *với hàm mục tiêu là lồi tương đương với việc nghiên cứu* 2 *bài toán quy hoạch lồi và quy hoạch lõm.*

1.2.4 Những vấn đề của lý thuyết tối ưu

Lý thuyết tối ưu quan tâm giải quyết những vấn đề cơ bản sau:

1. Tìm công cụ toán học để nghiên cứu.
2. Tìm điều kiện cần cho bài toán tối ưu.
3. Tìm điều kiện đủ cho bài toán tối ưu.
4. Tìm điều kiện tồn tại nghiệm.
5. Tìm các phương pháp để giải các bài toán tối ưu (phương pháp số và các phương pháp tiến hóa như GEN, PSO).

Mục đích của chuyên đề là đi theo lược đồ trên để trình bày các kết quả trong lý thuyết tối ưu, các thuật toán giải bài toán tối ưu và cài đặt các chương trình tính toán với các thuật toán cụ thể (việc viết chương trình tìm lời giải tối ưu sẽ được giao cho học viên thực hiện như là bài tập lớn).

1.3. CÔNG CỤ GIẢI TÍCH CHO BÀI TOÁN TỐI ƯU

1.3.1 Một số kiến thức cơ sở

Định nghĩa 1.3.1. *Tập X gọi là không gian véc tơ tuyến tính nếu trên đó xác định các phép toán "+" và "*" vô hướng thỏa mãn các tính chất sau:*

1. $\forall x, y \in X, \lambda, \mu \in I\!R \to \lambda x + \mu y \in X$
2. $1 * x = x, 0 * x = 0, 0 + x = x$
3. *với mỗi* $x \in X$, *tồn tại duy nhất* $(-x)$ *sao cho* $: x + (-x) = 0$

Ví dụ 1.3.11. *Ký hiệu* $I\!R^n := \{x = (x_1, x_2, \ldots, x_n \mid x_i \in I\!R, i = 1, 2, \ldots, n\}$ *với các phép toán* $x+y := (x_1+y_1, \ldots, x_n+y_n), y = (y_1, \ldots, y_n)$ *là không gian véc tơ tuyến tính. Một hệ véc tơ* $u_i \in I\!R^n, i = 1, \ldots, n$ *được gọi là cơ sở của không gian* $I\!R^n$ *nếu với mọi* $x \in I\!R^n$ *luôn có duy nhất biểu diễn* $x = \sum_{i=1}^n \alpha_i u_i$. *Từ định nghĩa trên hệ n véc tơ* $\{e_1 = (1, 0, \ldots, 0), e_2 = (0, 1, 0, \ldots, 0), \ldots, e_{n-1} = (0, 0, \ldots, 0, 1, 0), e_n = (0, 0, \ldots, 0, 1)\}$ *là cơ sở trong* $I\!R^n$ *và cơ sở này gọi là cơ sở đơn vị.*

Định nghĩa 1.3.2. *Bộ* $(X, \|\cdot\|)$ *gọi là không gian véc tơ tuyến tính định chuẩn nếu*

1. *X là không gian véc tơ tuyến tính,*
2. $\|\cdot\| : X \to R^+$ $(\|\cdot\|)$ *được gọi là chuẩn nếu thỏa mãn:*
   $\|x\| = 0$ *khi và chỉ khi* $x = 0$, $\|\alpha x\| = |\alpha|\|x\|, \|x + y\| \le \|x\| + \|y\|$.

Một trong những không gian quan trọng

Định nghĩa 1.3.3. *Cho X là không gian véc tơ tuyến tính định chuẩn trên trường số thực,* $\langle \cdot, \cdot \rangle : X \times X \to R$ *gọi là tích vô hướng trong không gian véc tơ tuyến tính định chuẩn nếu với mọi* $x,\ y,\ x \in X,\ \lambda \in I\!R$

1. $\langle x, y \rangle = \langle y, x \rangle$,

2. $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$,

3. $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$,

4. $\langle x, x\rangle \geq 0$; $\langle x, x\rangle = 0$ *khi và chỉ khi* $x = 0.q$

Ví dụ 1.3.12. *Trong không gian tuyến tính với tích vô hướng đặt* $\|\dot{|}\|$ *như sau* $\|x\| := \sqrt{\langle x, x\rangle}$ *sẽ là chuẩn trong* X.

Thật vậy, theo định nghĩa tích vô hướng và chuẩn trên thì

$$\|x+\lambda y\|^2 = \|x\|^2 + 2\lambda\langle x, y\rangle + \lambda^2\|y\|^2 \geq 0,$$

với mọi $\lambda \in I\!R$. Do đó biệt thức Δ của tam thức bậc 2 tham số λ phài nhỏ hơn hoặc bằng 0, tức là: $\langle x, y\rangle^2 - \|x\|^2\|y\|^2 \leq 0$. Vậy nên

$$|\langle x, y\rangle| \leq \|x\|\|y\|$$

Đây chính là bất đẳng thức Schwartz. Ngoài ra từ bất đẳng thức này dễ dàng chỉ ra hàm được định nghĩa như trên thỏa mãn điều kiện bất đẳng thức tam giác của chuẩn

$$\|x+y\|^2 = \|x\|^2+\langle x, y\rangle+\|y\|^2 \leq \|x\|^2+2|\langle x, y\rangle|+\|y\|^2 \leq \|x\|^2+2\|x\|\|y\|+\|y\|^2$$

hay

$$\|x+y\| \leq \|x\| + \|y\|.$$

Định nghĩa 1.3.4. *1.* $(X, \|\cdot\|)$ *gọi là không gian Banach nếu mọi dãy Cosi đều hội tụ trong* X. *Trong đó,* $\{x^n\}$ *được gọi là dãy Cosi nếu:*

$$\forall\epsilon > 0 \ \ \exists N : \quad (n, m > N \implies \|x^n - x^m\| \leq \epsilon).$$

2. Không gian tuyến tính có tích vô hướng $(X, \langle\cdot,\cdot\rangle)$ *với chuẩn* $\|x\| = \sqrt{\langle x, x\rangle}$ *đầy đủ gọi là không gian Hibert.*

Không gian $C[a, b] := \{x(t) \mid x(.)$ liên tục trên $[a, b]$ với chuẩn $\|x\| = \max_{t\in[a,b]} |x(t|\}$ là không gian Banach.

Không gian $I\!R^n$ với chuẩn $\|x\| := \sqrt{\sum_{i=1}^n x_i^2}$ và tích vô hướng $\langle x, y\rangle := \sum_{i=1}^n x_i y_i$ thỏa mãn mọi tính chất của chuẩn và tích vô hướng nên là không gian Hilbert.

Định nghĩa 1.3.5. *Cho X, Y là các không gian tuyến tính định chuẩn $A : X \to Y$ được gọi là toán tử tuyến tính nếu:*
$A(\alpha x + \beta y) = \alpha A(x) + \beta A(y)$ với mọi $x, y \in X, \alpha, \beta \in I\!R$ và gọi là tuyến tính liên tục nếu nó liên tục tại điểm 0 *(khi đó cũng sẽ liên tục tại mọi $x \in X$). Tập các toán tử tuyến tính từ X vào Y ký hiệu là $L(X, Y)$.*

1.3.2 Biến phân bậc nhất và đạo hàm

Như chúng ta đã biết, khi nghiên cứu cực trị của hàm một biến, ta có định lý về điều kiện cần Fermat ($f'(x^*) = 0$) và định lý về điều kiện đủ $f'(x^*) = 0, f''(x^*) < 0$, hoặc $f''(x^*) < 0$. Câu hỏi tự nhiên được đặt ra ở đây là: đối với hàm nhiều biến (tức là đối số nằm trong không gian hữu hạn hoặc vô hạn chiều) khái niệm đạo hàm phải được mở rộng ra như thế nào để Định lý Fermat vẫn còn hiệu lực? Mục sau đây đưa ra một số định nghĩa được hiểu nôm na là đạo hàm nhằm mở rộng các định lý trên về điểm cực trị.

Định nghĩa 1.3.6. *Cho X, Y không gian tô pô tuyến tính (tuyến tính và được trang bị tô pô, không gian tuyến tính định chuẩn là một trường hợp đặc biệt), V là lân cận của $x \in X, F : X \to Y$. Nếu*

$$\delta F(x, h) := \lim_{t\to 0} t^{-1}(F(x + th) - F(x)) \tag{1.2}$$

tồn tại với mọi $h \in X$ thì ánh xạ $h \to \delta F(x, h)$ được gọi là biến phân bậc nhất của F tại x.

Nếu tồn tại toán tử Λ sao cho

$$\Lambda h = \delta F(x, h)\ \forall h \in X$$

thì Λ gọi là đạo hàm Gato và ký hiệu là $F'_G(x)$ hay $F'(x)$ và ta nói F khả *vi Gato tại x. Điều này xẩy ra khi và chỉ khi*

$$F(x+th) = F(x) + t\Lambda h + o(t)\ \forall h \in X.$$

Ví dụ hàm $f(x) = rcos\varphi, r, \varphi$ tọa độ cực của $x \in I\!R^2$. Khi đó $\delta f(0,h) = \delta f(x,h) := \lim_{t\to 0} t^{-1}(f(th) - f(0)) = \lim_{t\to 0} t^{-1}(trcos\varphi - 0) = rcos\varphi = f(h)$. Vì $\delta f(0,h)$ không tuyết tính nên f không khả vi Gato tại $0 \in I\!R^2$

Định nghĩa 1.3.7. *Nếu X, Y là không gian Banach. $F : X \to Y$ gọi là khả vi Frechet tại x nếu tồn tại toán tử tuyến tính liên tục Λ :*

$$F(x+h) := F(x) + \Lambda h + r(h) \text{ với } \lim_{\|h\|\to 0} \frac{\|r(h)\|_Y}{\|h\|_X} = 0$$

và Λ gọi là đạo hàm Frechet ký hiệu là $F'_G(x)$ hoặc $F'(x)$. Ánh xạ F là chính quy tại x nếu nó khả vi Frechet tại x và $F'(x)X = Y$. Λ là đạo hàm Frechet ký hiệu là $F'_G(x)$ hoặc $F'(x)$. F là chính quy tại x nếu khả vi Frechet tại x và $F'(x)X = Y$.

Mệnh đề 1.3.1. (a) *Nếu F khả vi Frechet tại x thì liên tục và khả vi Gato tại đây.*

(b) *Nếu F khả vi Gato tại x thì tồn tại biến phân bậc nhất tại đó và $\delta F(x,h) = F'_G h$.*

Chú ý rằng hàm $f(x) = \begin{cases} 1 & \text{nếu } x_1 = x_2^2 \text{ và } x_2 \neq 0 \\ 0 & \text{trường hợp ngược lại} \end{cases}$ khả vi Gato tại $(0,0) \in I\!R^2$ nhưng không liên tục tại đó nên không tồn tại đạo hàm Frechet được.

Với các khái niệm được mở rộng như trên, các tính chất cơ bản của giải tích cổ điển như như định lý về trị trung bình, định lý về hàm hợp, định lý về hàm ẩn, đạo hàm bậc cao vẫn còn hiệu lực và đóng vai trò quan trọng trong giải tích hàm và lý thuyết tối ưu. Sau đây là những định lý đó.

Định lý 1.3.1. (**Định lý giá trị trung bình**) *X, Y không gian vecto tô pô, U tập mở của X, $F : U \to Y$ khả vi Gato tại mọi điểm trên $[x, x+h] \subset U$.. Khi đó*

(a) *Nếu ánh xạ* $z \mapsto F'_G(z)h$ *là một ánh xạ liên tục của* $[x, x+h]$ *vào* Y *thì*

$$F(x+h) - F(x) = \int_0^1 F'_G(x+th)hdt.$$

(b) *Nếu* X, Y *là k.g Banach và* $z \mapsto F'_G(z)h$ *là một ánh xạ liên tục của* $[x, x+h]$ *vào* Y *thì*

$$\|F(x+h) - F(x)\| \leq \sup_{0 \leq t \leq 1} \|F'_G(x+th)\|.\|h\|$$

và với mỗi $\Lambda \in L(X, Y)$ *thì*

$$\|F(x+h) - F(x) - \Lambda h\| \leq \sup_{0 \leq t \leq 1} \|F'_G(x+th) - \Lambda\|.\|h\|.$$

Đặc biệt, với mọi $z \in [x, x+h]$ *thì*

$$\|F(x+h) - F(x) - F'_G(z)h\| \leq \sup_{0 \leq t \leq 1} \|F'_G(x+th) - F'_G(z)\|.\|h\|.$$

1.3.3 Biến phân và đạo hàm bậc cao (đọc thêm)

Nếu $h \in X$, hàm $\phi_h(t) := F(x+th)$ khả vi ít nhất n lần tại x thì

$$\delta^n(F(x,h) := \frac{d^n}{dt^n}\phi(t)|_{t=0} \tag{1.3}$$

Đạo hàm của đạo hàm Frechet cấp $n-1$ là đạo hàm Frechet cấp n nếu tồn tại.

Định lý 1.3.2. (**Định lí về đạo hàm riêng của Schwartz**) *X, Y và Z là Banach, $U \in X \times Y$ và $F : U \to Z$ có đạo hàm (Frechet) riêng $F_x(x,y)$ và $F_y(x,y)$ tại mọi điểm $(x,y) \in U$. Nếu $(x,y) \mapsto F_x(x,y)$ và $(x,y) \mapsto F_y(x,y)$ liên tục (theo topo đều) tại $(\overline{x}, \overline{y}) \in U$ thì F khả vi Frechet tại đó và*

$$F'(\overline{x}, \overline{y})[(\xi, \eta)] = F_x(\overline{x}, \overline{y})[\xi] + F_y(\overline{x}, \overline{y})[\eta].$$

Định lý 1.3.3. (**Quy tắc dây chuyền**) *Cho X, Y và Z Banach, $U \subset X$, $V \subset Y, F : U \to Y$ và $G : V \to Z$. Cho $x \in U$ với $F(x) \in V$. Nếu F khả vi*

Frechet tại x và G khả vi Frechet tại $F(x)$ thì ánh xạ $H = G \circ F$ cũng khả vi Frechet tại x và

$$H'(x) = G'(F(x)) \circ F'(x).$$

Định lý 1.3.4. (**Định lí hàm ẩn**) *X, Y và Z Banach, $(x_0, y_0) \in U \subset X \times Y$ và $F : U \to Z$ khả vi Frechet liên tục. Giả sử $F(x_0, y_0) = 0$ và $F_y(x_0, y_0)$ là một phép đồng phôi tuyến tính. Khi đó tồn tại $\epsilon > 0, \delta > 0$ và một ánh xạ $x \mapsto y(x)$ từ quả cầu $B(x_0, \delta) \subset X$ vào quả cầu $B(y_0, \epsilon) \subset Y$ sao cho:*

(a) *Hai quan hệ $F(x, y) = 0$ và $y = y(x)$ tương tự trên tập $B(x_0, \delta) \times B(y_0, \epsilon)$.*

(b) *$y(.)$ khả vi liên tục và $y'(x) = -[F_y(x, y(x)]^{-1} \circ F_x(x, y(x))$.*

Định nghĩa 1.3.8. (**Nón**) *Tập $D \subset X$ được gọi là nón nếu với mọi $\lambda \geq 0$ và $x \in D$ thì $\lambda x \in D$.*

Định nghĩa 1.3.9. *Cho $M \subset X, x \in X$ là vecto tiếp tuyến tập M tại $x_0 \in M$ nếu tồn tại $\epsilon > 0$ và ánh xạ $r : [0, \epsilon] \to X$, thỏa mãn $\lim_{t\to 0} ||r(t)||/t = 0$, sao cho*

$$x_0 + tx + r(t) \in M \ \forall t \in [0, \epsilon].$$

Tập các vecto tiếp tuyến của M là một nón đóng và khác rỗng (vì chứa điểm 0), gọi là nón tiếp tuyến tập M tại x_0 kí hiệu là $T_M(x_0)$.

Định lý 1.3.5. (**Định lí Lyusternik**) *X, Y Banach, $x_0 \in V \subset X$, và $F : V \to Y$ khả vi Frechet. Giả sử F chính qui tại x_0 (tức $\mathrm{Im\,F'(x) = Y}$) và khả vi liên tục tại x_0. Khi đó tập $M = \{x \in U : F(x) = F(x_0)\}$ có một không gian tiếp tuyến tại x_0 và $T_M(x_0) = KerF'(x)$.*

Các định lý trên được ứng dụng nhiều trong việc nghiên cứu các bài toán cực trị trong không gian vô hạn chiều, tuy nhiên trong giáo trình này chúng ta chỉ hạn chế nghiên cứu trong không gian hữu hạn chiều, khi đó *các định nghĩa về đạo hàm như trên liên quan mật thiết đến đạo hàm riêng theo các biến, chúng ta sẽ trình bày rõ hơn ở các phần sau.*

1.3.4 Tập lồi

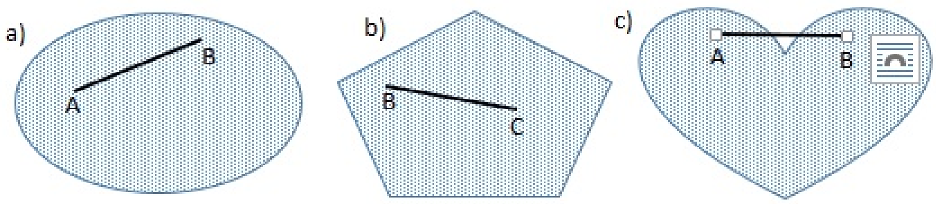
Định nghĩa 1.3.10. (**Tập lồi**) *Tập $D \subset I\!R^n$ gọi là lồi nếu*

$$x, y \in D, \lambda \in [0,1] \Longrightarrow \lambda x + (1-\lambda)y \in D.$$

Định nghĩa 1.3.11. (**Tổ hợp lồi**) *Cho $x^1, \ldots x^m$ là các véc tơ trong $I\!R^n$ gọi*

$$\sum_{i=1}^{m} \lambda_i x^i, \quad \sum_{i=1}^{m} \lambda_i = 1, \quad \lambda_i \geq 0$$

là tổ hợp lồi của các véc tơ $x^1, \ldots x^m$,



Hình 1.2: **Tập lồi a), b), tập không lồi c)**

Mệnh đề 1.3.2. *1. Tổng đại số của hữu hạn tập lồi là lồi.*

2. Giao của họ các tập lồi là lồi.

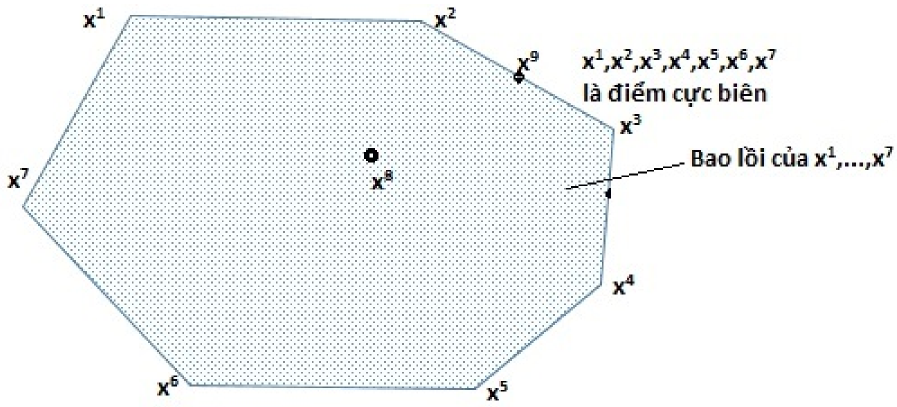
3. Tích Đề các của các tập lồi là lồi.

4. Ảnh và nghịch ảnh của tập lồi qua ánh xạ tuyến tính cũng là lồi.

5. $D = \{x \mid x = \sum_{i=1}^{m} \lambda_i x^i, x^i \in D, \sum_{i=1}^{m} \lambda_i = 1, \lambda_i \geq 0, i = 1, 2, \ldots, m; m \geq 1\}$.

6. Nón là một tập lồi.

Định nghĩa 1.3.12. (**Điểm cực biên**) *Điểm x^* được gọi là điểm cực biên của tập lồi D nếu không tồn tại hai điểm khác nhau $x^1, x^2 \in D$ sao cho $x^* = \frac{1}{2}x_1 + \frac{1}{2}x_2$. Điều này tương đương với nếu $x_1, x_2 \in D$ thỏa mãn $x^* = \frac{1}{2}x_1 + \frac{1}{2}x_2$ thì $x^* = x_1 = x_2$. Tập các điểm cực biên của tập lồi ký hiệu là* $\mathrm{Ext}(D)$.



Hình 1.3: **Điểm cực biên và bao lồi**

Định nghĩa 1.3.13. (**Bao lồi**) *Cho D là một tập hợp, bao lồi của D là giao của mọi tập lồi chứa D hay nói cách khác bao lồi của D là tập lồi nhỏ nhất chứa D. Bao lồi của D ký hiệu là* $\mathrm{co}(D)$, hoặc $\mathrm{conv}(D)$.

Định lý 1.3.6. (**Định lý Minkovski**) *Cho D là tập lồi trong X, khi đó $D = \mathrm{co}(\mathrm{Ext}(D))$.*

Định lý 1.3.7. (**Định lý Hahn - Banach**) *Cho không gian topo tuyến tính X, $A \subset X$ là tập lồi mở, $L \subset X$ là một không gian con, thỏa mãn $A \cap L = \emptyset$. Khi đó tồn tại một phiếm hàm tuyến tính liên tục x^* sao cho:*

$$\langle x^*, x\rangle > 0\ \forall x \in A \text{ và } \langle x^*, x\rangle = 0\ \forall x \in L.$$

Định lý 1.3.8. (**Định lý tách**) *Cho A và B là hai tập lồi trong không gian tuyến tính X, có tính chất $A \cap B = \emptyset$ và $\mathrm{int}A \neq \emptyset$. Khi đó A và B có thể tách được bằng một phiếm hàm tuyến tính khác 0, tức*

$$\exists x^* \in X^* \setminus \{0\}\ \forall x \in A\ \forall y \in B : \langle x^*, x\rangle \geq \langle x^*, y\rangle.$$

Định nghĩa 1.3.14. (**Bao affine**) *Cho $D \subset I\!R^n$ gọi tập $\{x : x = \alpha x^1 + (1-\alpha)x^2, x^1, x^2 \in D, \alpha \in I\!R\}$ bao affine của D ký hiệu là* $\mathrm{aff}(D)$.

Tập Với $x \in D$, $x-\mathrm{aff}(D)$ là một không gian con, số chiều của không gian con này được gọi là thứ nguyên của $\mathrm{aff}(D)$.

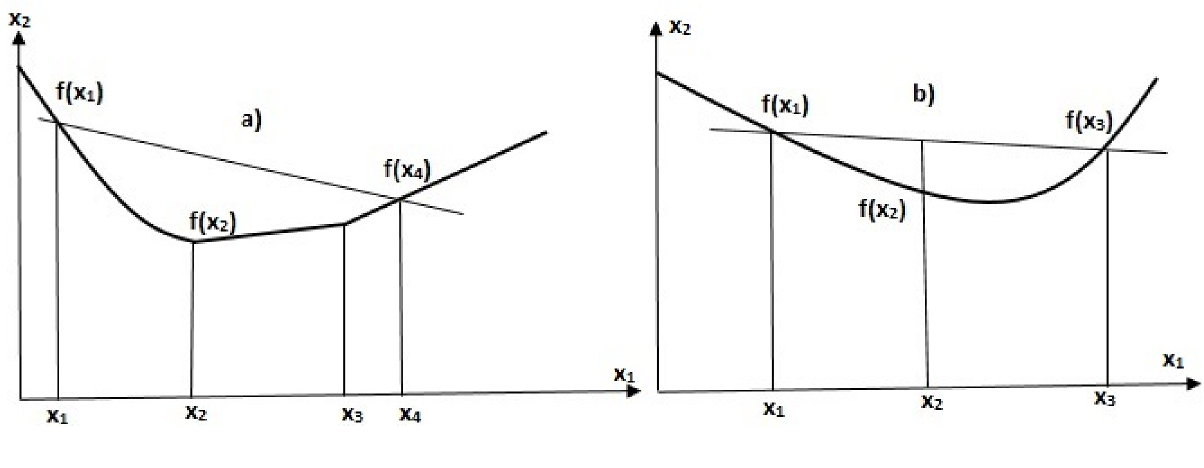
Định lý 1.3.9. (**Caratheodory**) *Nếu D là một tập hợp chứa trong một đa tạp r thứ nguyên thì mọi điểm $x \in \mathrm{co}\,(D)$ đều có thể biểu diễn thành một tổ hợp lồi của không quá $r+1$ điểm của D.*

1.3.5 Hàm lồi

Định nghĩa 1.3.15. (**Hàm lồi**) *Hàm $f(x)$ xác định trên tập lồi D gọi là lồi nếu với mọi*

$$\forall x, y \in D, \lambda \in [0,1] : f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda)f(y), \qquad (1.4)$$

nếu bất đẳng thức trên là thực sự với mọi $\lambda \in (0,1)$ thì hàm f được gọi là lồi ngặt.



Hình 1.4: **Hàm lồi a) và lồi ngặt b)**

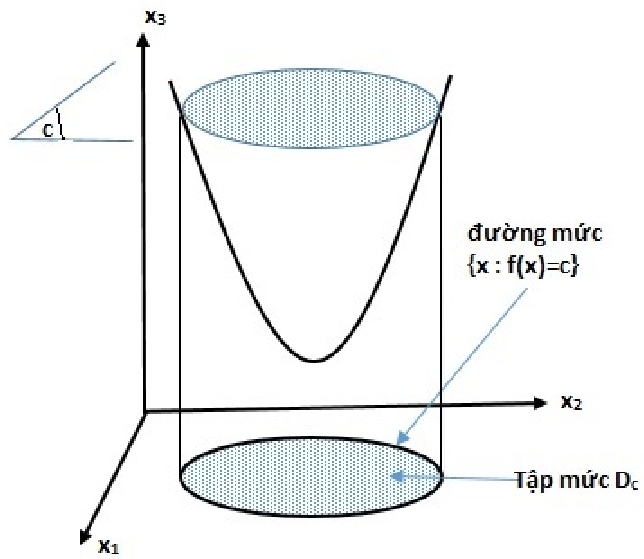
Định nghĩa 1.3.16. (**Tập mức của hàm lồi**) *Hàm $f(x)$ xác định trên tập lồi D gọi tập*

$$D_\alpha := \{x \in D \mid f(x) < c\} \qquad (1.5)$$

là tập mức dưới của hàm lồi f trên D. và tập

$\{x \in D \mid f(x) = \alpha\}$ gọi là đường mức của hàm f trên D là tập mức dưới của hàm lồi f và tập Hàm lồi có rất nhiều tính chất giải tích quan trọng, ta quan tâm đến các tính chất tối ưu hóa sau:

- ***Cực tiểu địa phương là cực tiểu toàn cục.***
- ***Tập mức dưới*** $\mathcal{L}(\alpha, f) := \{x \in D \mid f(x) \leq \alpha\}$ ***là tập lồi.***
- ***Điểm dừng là điểm cực tiểu toàn cục.***
- ***Nếu*** D ***là tập compắc thì hàm đạt cực đại tại ít nhất một điểm cực biên.***



Hình 1.5: **Tập mức và đường mức**

1.3.6 Về bài toán Quy hoạch tuyến tính (đọc thêm)

Cho ma trận $A = \{a_{ij}\} \in I\!R^{m\times n}, c = (c_1, c_2, \ldots, c_n) \in I\!R^n, b = (b_1, b_2, \ldots, b_m)' \in I\!R^m$. Ký hiệu $A_i = (a_{i1}, a_{i2}, \ldots, a_{in}) \in I\!R^n, A^j = (a_{1j}, a_{2j}, \ldots, a_{mj})' \in I\!R^m$ là véc tơ hàng và véc tơ cột tương ứng của ma trận A.

Bài toán quy hoạch tuyến tính dạng tổng quát
Tìm véc tơ $x \in I\!R^n$ sao cho

$$
\begin{aligned}
&\langle c, x\rangle \to \min(\max)\\
&\langle A_i, x\rangle \geq b_i, \quad i \in I \subset M := \{1, 2, \ldots, m\}\\
&\langle A_i, x\rangle = b_i, \quad i \in\subset \{1, 2, \ldots, m\} \setminus I\\
&x_j \geq 0, \quad j \in J \subset N := \{1, 2, \ldots, n\}
\end{aligned}
$$

Bài toán quy hoạch tuyến tính dạng chuẩn tắc
Tìm véc tơ $x \in I\!R^n$ sao cho

$$
\begin{aligned}
&\langle c, x\rangle \to \min(\max)\\
&\langle A_i, x\rangle \geq b_i, \quad i \in M := \{1, 2, \ldots, m\}\\
&x_j \geq 0, \quad j \in N := \{1, 2, \ldots, n\}
\end{aligned}
$$

Bài toán quy hoạch tuyến tính dạng chính tắc
Tìm véc tơ $x \in I\!R^n$ sao cho

$$\langle c, x\rangle \to \min(\max) \tag{1.6}$$
$$\langle A_i, x\rangle = b_i, \quad i \in M := \{1, 2, \ldots, m\} \tag{1.7}$$
$$x_j \geq 0, \quad j \in N := \{1, 2, \ldots, n\} \tag{1.8}$$

Nhận xét

1. Có thể đưa bài toán xét min về xét max.
2. Có thể đổi dấu "$\leq$",thành "$\geq$" và ngược lại.
3. Có thể đổi dấu "$\leq$", "$\geq$" thành dấu "$=$".
4. Có thể thay biến âm x_j thành hai biến không âm $x_j^+, x_j^- \geq 0$ trong đó $x_j = x_j^+ - x_j^-$.
5. ***Từ các nhận xét trên suy ra mọi bài toán quy hoạch tuyến tính đều có thể đưa về bài toán quy hoạch tuyến tính dạng***

chính tắc. Do đó ta chỉ xét bài toán quy hoạch tuyến tính dạng chính tắc. Khi đó miền ràng buộc $D = \{\langle A_i, x\rangle = b_i, \quad i \in M := \{1, 2, \ldots, m\}, x_j \geq 0, \quad j \in N := \{1, 2, \ldots, n\}$.

Từ nhận xét trên nên từ nay ta chỉ nghiên cứu bài toán quy hoạch tuyến tính với bài toán min .

Mệnh đề 1.3.3. *(Về phương án tối ưu của bài toán quy hoạch tuyến tính)*

1. *Nếu D là đa diện lồi trong $I\!R^n$ thì bài toán quy hoạch tuyến tính có phương án tối ưu.*

2. *Nếu bài toán quy hoạch tuyến tính có phương án tối ưu thì có ít nhất một phương án tối ưu là điểm cực biên.*

3. *Nếu hàm mục tiêu của bài toán* min *(bài toán* max*) bị chặn dưới (bị chặn trên) thì tồn tại phương án tối ưu.*

Ký hiệu $J_0 := \{j_1, j_2, \ldots, j_m; j_i \in \{1, 2, \ldots, n\}; i \in \{1, 2 \ldots, m\}\}$.

Định nghĩa 1.3.17. • *Điểm $x^0 = (x_j^0)$ gọi là phương án cực biên của bài toán QHTT chính tắc nếu x^0 là phương án chấp nhận được và là điểm cực biên của D.*

- *Phương án cực biên $x^0 = (x_i^0), i = 1, 2, \ldots, n$ gọi là không suy biến nếu $x_i > 0$ với mọi $i \in J_0$, tức là x^0 có đúng m phần tử lớn hơn 0 và gọi là suy biến nếu có ít hơn m phần tử lớn hơn 0.*

- *Thông thường ta ký hiệu $x^0 = (x_i^0) \in I\!R^m$; $i \in J_0$ mà bỏ qua những phần tử bằng 0 và gọi là phương án cực biên quy gọn thay cho $x^0 = (x_i); i = 1, \ldots, n$, $(x_j > 0$ với $i \in J_0)$.*

Định lý 1.3.10. *Ký hiệu $H(x^0) := \{A^i \mid x_i^0 > 0\}$. Khi đó x^0 là phương án cực biên khi và chỉ khi $H(x^0)$ độc lập tuyến tính.*

Từ định lý trên suy ra, nếu x^0 là phương án cực biên không suy biến thì $H(x^0)$ là cơ sở trong $I\!R^m$. Do đó với mọi $j = 0, 1 \ldots, n$

$$A^j = \sum_{i \in J_0} x_i^j A^i,$$

ở đây $A^0 := b$ và biểu diễn là duy nhất. Ký hiệu $x^j := (x_i^j), c^0 := (c_i) \in I\!R^m$, $i \in J_0$. Gọi $\Delta_j := \langle c^0, x^j \rangle - c_j = \sum_{i \in J_0} c_i^0 x_i^j - c_j$ là ước lượng của véc tơ A^j, ta dễ nhận thấy rằng $\Delta_j = 0$ với $j \in J_0$.

Định lý 1.3.11. *Nếu x^0 là phương án cực biên không suy biến khi đó:*

1. *Nếu $\Delta_j \leq 0, j = 1, 2, \ldots, n$ thì x^0 là phương án tối ưu của bài toán QHTT dạng chính tắc.*

2. *Nếu tồn tại $\Delta_j > 0$ sao cho ứng với j đó $x^j \leq 0$ thì bài toán không có phương án tối ưu.*

3. *Nếu tồn tại $\Delta_j > 0$ sao cho ứng với j đó tồn tại $x_i^j > 0$ khi đó có thể xây dựng phương án cực biên mới tốt hơn phương án đã có.*

Thuật toán đơn hình khi biết phương án cực biên và cơ sở đơn vị
Giả sử $b \geq 0$ và $B := [A^{j_1}, A^{j_2}, \ldots, A^{j_m}] = E$. Khi đó dễ thấy $x^0 = B^{-1}b \geq 0$ là phương án cực biên xuất phát. Thuật toán đơn hình dạng bảng gồm các bước sau:

B1. a/ Tính $x^j = B^{-1}A^j, j = 0, \ldots, n$.
b/ Lập bảng đơn hình.
c/ Tính Δ_j.

B2. a/ Nếu $\Delta_j \leq 0, j = 1, 2, \ldots, n$ thì x^0 là phương án tối ưu của bài toán QHTT dạng chính tắc
b/ Nếu tồn tại $\Delta_j > 0$ sao cho ứng với j đó $x^j \leq 0$ ta dừng và kết luận bài toán không có phương án tối ưu.
c/ Nếu tồn tại $\Delta_j > 0$ sao cho ứng với j đó tồn tại $x_i^j > 0$ ta thực

hiện:
- Đưa véc tơ mới A^k gọi là cột xoay vào cơ sở, k được xác định:

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

- Loại A^s ra khỏi cơ sở cũ gọi là dòng xoay, s được xác định:

$$\min\{x_i^0/x_i^k \mid x_i^k > 0\} = x_s^0/x_s^k,$$

x_s^k gọi là phần tử xoay, chuyển B3.

B3. Tính lại x^j theo cơ sở mới. Công thức chuyển đổi như sau:
a/ $x_i^j(m) := x_i^j(c) - x_s^j x_i^k / x_s^k, i \neq s$,
b/ $x_s^j(m) := x_s^j(c)/x_s^k$, c/ lập bảng đơn hình mới. Quay lại b/, c/ của B1. Thuật toán kết thúc sau hữu hạn bước.

Thuật toán đơn hình khi cơ sở không là cơ sở đơn vị

Trong trường hợp này nếu $x^0 = B^{-1}b \geq 0$ thì x^0 là phương án cực biên, tuy nhiên điều này gặp một số khó khăn khi tính toán, hơn nữa nhiều khi ma trận B có thể không khả ngược, do đó người ta sử dụng một số thuật toán khác nhằm tránh hạn chế đó như phương pháp 2 pha (pha thứ nhất tìm phương án cực biên thông qua bài toán phụ, pha thứ hai tìm phương án tối ưu của bài toán gốc), phương pháp đánh thuế (tìm phương án tối ưu của bài toán phụ với cơ sở đơn vị gồm những thành phần không có trong bài toán gốc sau đó tìm nghiệm tối ưu của bài toán gốc), phương pháp đối ngẫu (tìm phương án tối ưu thông qua giả phương án của bài toán đối ngẫu). Để thuận tiện cho việc lập trình tìm phương án tối ưu của bài toán quy hoạch tuyến tính, chúng tôi trình bày thêm phương pháp đánh thuế. Các phương pháp khác có thể tham khảo chi tiết trong các tài liệu chuyên khảo về quy hoạch tuyến tính (Tối ưu hóa - Nguyễn Đức Nghĩa).

Phương pháp đánh thuế Thay vào việc giải bài toán 1.6-1.8 ta giải bài toán phụ sau:

$$\langle c, x\rangle + M(x_{n+1} + x_{n+1}, \dots + x_{n+m}) \to \min(\max) \tag{1.9}$$

$$\langle A_i, x\rangle + x_{n+i} = b_i, \quad i = \{1, 2, \dots, m\} \tag{1.10}$$

$$x_j \geq 0, \quad j \in \{1, 2, \dots, n, n+1, \dots + n+m\} \tag{1.11}$$

trong đó như thường lệ $x = (x_1, x_2, \dots, x_n), w := (x_{n+1}, x_{n+2}, \dots, x_{n+m})$, M là một số dương cực lớn, lớn hơn bất cứ một số cụ thể nào.

Rõ ràng bài toán trên gồm $n + m$ ẩn và có sơ sở đơn vị là $A^{n+1}, A^{n+2}, \dots, A^{n+m}$ trong $I\!R^m$. Do bài toán có phương án cực biên không suy biến ban đầu $(x^0, w^0) = b \geq 0$, trong đó $x^0 = 0$ và $w^0 = (b_1, b_2, \dots, b_m)$, nên ta có thể áp dụng thuật toán đơn hình dạng bảng với cơ sở đơn vị (và phương án cực biên ban đầu đã biết) cho bài toán (1.9) - (1.11). Lưu ý rằng ước lượng của véc tơ A^j trong bài toán này phụ thuộc tuyến tính vào M.

Định lý 1.3.12. (Nhận biết phương án tối ưu của phương pháp đánh thuế) *Giả sử bài toán* (1.9) - (1.11) *có phương án tối ưu* (x^*, w^*). *Khi đó*

1. *Nếu* $w^* \neq 0$ *thì bài toán gốc* (1.6) - (1.8) *không có phương án tối ưu*
2. *Nếu* $w^* = 0$ *thì bài toán gốc* (1.6) - (1.8) *có phương án tối ưu là* x^*.

Ví dụ 1.3.13.

$$\begin{aligned} 2x_1 + 6x_2 - 5x_3 + x_4 + 4x_5 &\to \min \\ x_1 - 4x_2 + 2x_3 - 5x_4 + 9x_5 &= 3 \\ x_2 - 3x_3 + 4x_4 - 5x_5 &= 6 \\ x_2 - x_3 + x_4 - x_5 &= 1 \\ x_j &\geq 0 \end{aligned}$$

c^0	CS	x^0	2	6	-5	1	4	M	M	M
			x^1	x^2	x^3	x^4	x^5	x^6	x^7	x^8
M	A^6	3	1	-4	2	-5	**9**	1	0	0
M	A^7	6	0	1	-3	4	-5	0	1	0
M	A^8	1	0	1	-1	1	-1	0	0	1
	Δ		1	-2	-2	0	**3**	0	0	0
			-2	-6	5	-1	-4	0	0	0
4	A^5	1/3	1/9	-4/9	2/9	-5/9	1		0	0
M	A^7	23/3	5/9	-11/9	-17/9	11/9	0		1	0
M	A^8	4/3	1/9	5/9	-7/9	**4/9**	0		0	1
	Δ		2/3	-2/3	-24/9	**5/3**	0		0	0
			-14/9	-70/9	53/9	-29/9				
4	A^5	2	1/4	1/4	-3/4	0	1		0	
M	A^7	4	1/4	-11/4	**1/4**	0	0		1	0
1	A^4	3	5/4	5/4	-7/4	1	0		0	
	Δ		1/4	-11/4	**1/4**	0	0		0	
			-3/4	1/4	1/4	0			0	
4	A^5	14	1	-8	0	0	1			
-5	A^3	16	1	-11	1	0	0			
1	A^4	31	2	-18	0	1	0			
	Δ		-1	-1	0	0	0			
			-3/4	1/4	1/4	0	0			

Bảng 1.1: *

Ta xây dựng bài toán phụ như sau:

$$
\begin{aligned}
2x_1 + 6x_2 - 5x_3 + x_4 + 4x_5 + M(x_6 + x_7 + x_8) \to &\ \min \\
x_1 - 4x_2 + 2x_3 - 5x_4 + 9x_5 + x_6 &= 3 \\
x_2 - 3x_3 + 4x_4 - 5x_5 + x_7 &= 6 \\
x_2 - x_3 + x_4 - x_5 + x_8 &= 1 \\
x_j &\geq 0.
\end{aligned}
$$

Rõ ràng A^6, A^7, A^8 là cơ sở đơn vị ứng với phương án cực biên không suy biến $(x^0, w^0) = b = (3, 6, 1)'$. (Xem bảng đơn hình)

Phương án tối ưu là $(x^*, w^*) = (0, 0, 16, 31, 14, 0, 0, 0)$. Vì $w^* = (0, 0, 0)$ nên phương án tối ưu của bài toán gốc là $x^* = (0, 0, 16, 31, 14)$.

Ghi chú: Việc lập trình giải bài toán QHTT không phức tạp trên MATHLAB, học viên có thể tự thực hiện. Tuy nhiên sử dung hàm linprog trong MATHLAB ta có thể viết giải đơn giản như sau:

c=[ ]; (Nhập giá trị của vectơ c) A=[ ]; (Nhập ma trận A của ràng buộc bất đẳng thức)

b=[ ]; (Nhập vectơ b của ràng buộc bất đẳng thức)

Aeq=[ ]; (Nhập ma trận Aeq của ràng buộc đẳng thức)

beq=[ ]; (Nhập vectơ beq của ràng buộc đẳng thức)

ub=[ ]; lb=zeros(3,1); (Nhập cận trên và dưới của nghiệm)

disp('nghiem khong am');

[x, fval,exiflag,ouput]=linprog(c,A,b,Aeq,beq,lb,ub)

maxf=-fval

disp('nghiem khong hoac mot:')

[x, fval,exiflag,ouput]=linprog(c,A,b,Aeq,beq,lb,ub)

maxf=-fval

Bảng sau đây chỉ cho chúng ta cách giải các bài toán tối ưu thông qua thư viện của MATLAB

MATLAB để giải các bài toán tối ưu

Ví dụ 1.3.14.

$$f(x_1, x_2) = 9.82x_1x_2 + 2x_1 \to \min$$

với các ràng buộc sau:

$$g_1(x_1, x_2) = 2500/(\pi x_1 x_2) - 500 \le 0$$

Loại bài toán tối ưu	Mô hình	Tên ct MATLAB
Hàm một biến	Tìm x: $f(x) \to min$ $x1 \leq x \leq x2$	fminbnd
Tối ưu không ràng buộc	Tìm x: $f(x) \to min$	fminbnd fminsearch
Quy hoạch tuyến tính	Tìm x: $f^T x \to min$ $[A]x \leq b, [Aeq]x = beq, l \leq x \leq u$	linnprog
Quy hoạch toàn phương	Tìm x: $x^T[H]x + f^T x \to min$ $[A]x \leq b, [Aeq]x = beq, l \leq x \leq u$	quadprog
Tối ưu với ràng buộc bổ sung	Tìm x: $x^T[H]x + f^T x \to min$ $c(x) \leq 0, ceq = 0, [A]x \leq b$ $[Aeq]x = beq, l \leq x \leq u$	fmincon

Bảng 1.2: *

$$g_2(x_1, x_2) = 2500/(x_1x_2) - \pi(x_1^2 + x_2^2)/0.5882 \leq 0;$$

$$g_3(x_1, x_2) = -x_1 + 2 \leq 0; g_4(x_1, x_2) = x_1 - 14 \leq 0;$$

$$g_5(x_1, x_2) = -x_2 + 0.2 \leq 0; g_6(x_1, x_2) = x_2 - 0.8 \leq 0;$$

Quá trình tìm lời giải tối ưu:

B1: 1 Viết M-file probofminobj.m cho hàm mục tiêu.
function f= probofminobj (x)
f= 9.82*x(1)*x(2)+2*x(1);

B2: Viết M-file conprobformin.m cho các ràng buộc.
function [c, ceq] = conprobformin(x)
Các ràng buộc phi tuyến bất đẳng thức:
$c = [2500/(\pi * x(1) * x(2)) - 500; 2500/(\pi * x(1) * x(2)) - (\pi^2 * (x(1)^2$
$+x(2)^2))/0.5882; -x(1) + 2; x(1) - 14; -x(2) + 0.2; x(2) - 0.8];$
Các ràng buộc phi tuyến đẳng thức:
ceq = [];

B3: Nội dung chương trình (ghi vào file Mathlab mới):
x0 = [7 0.4];

fprintf ('The values of function value and constraints at starting point');

f=probofminobj (x0)

$[c, ceq] = conprobformin(x0)$

options = optimset ('LargeScale', 'off');

$[x, fval] = fmincon$ (@ probofminobj, x0, [], [], [], [], [],[],

fprintf('The values of constraints at optimum solution');

$[c, ceq] = conprobformin(x)$

Kết quả được đưa ra như sau: The values of function value and constraints at starting point

f=41.4960

c = -215.7947 -540.6668 -5.0000 -7.0000 -0.2000

-0.4000

ceq =[]

Optimization terminated: first-order optimality measure lessthan options. TolFun and maximum constraint violation is less than options.TolCon. Active inequalities (to within options.TolCon = 1e-006): lower upper ineqlin ineqnonlin

1

2

x=5.4510 0.2920

fval =26.5310

The values of constraints at optimum solution c=

-0.0000

-0.0000

-3.4510

-8.5490

-0.0920

-0.5080

$ceq = []$

Chương 2.

QUY HOẠCH TRƠN, LỒI

2.1. BÀI TOÁN TRƠN

2.1.1 Bài toán trơn không ràng buộc

Bài toán trơn không ràng buộc là

$$f(x) \to \inf,\ x \in X, \tag{2.12}$$

nếu hàm mục tiêu f là trơn (lồi) thì gọi là bài toán trơn (lồi) không ràng buộc. Hàm lồi có những tính chất tối ưu quan trọng như cực tiểu địa phương là cực tiểu toàn cục, điểm dừng là điểm cực tiểu toàn cục hay tập mức dưới là tập lồi, đối với các bài toán trơn, định lý Fermat và định lý nhân tử Lagrange lại đóng một vai trò rất quan trọng. Các định lý này được ví như định lý nền để xem xét điều kiện cần cho cực trị địa phương tại một điểm nào đó.

Định lý 2.1.13. (**Định lý Fermat**) *a/ Nếu x^* là nghiệm cực tiểu địa phương (là cực tiểu toàn cục khi f) của* (2.12) *và $f(x)$ có biến phân bậc nhất $\delta f(x^*, h)$. Khi đó*

$$\delta f(x^*, h) = 0\ \forall h \in X.$$

b/ Nếu X là không gian Banach và f khả vi Frechet tại x^* thì $f'(x^*) = 0$.

Khi $X = \mathbb{R}^n$ thì tương đương với

$$\nabla f(x^*) := (\frac{\partial f(x^*)}{\partial x_1}, \frac{\partial f(x^*)}{\partial x_2}, \dots, \frac{\partial f(x^*)}{\partial x_n})^T = 0.$$

Chứng minh. Theo định nghĩa biến phân bậc nhất thì $\delta f(x^*, h) \geq 0$ và $\delta f(x^*, h) = -\delta f(x^*, -h)$ với mọi $h \in X$. Do đó $\delta f(x^*, h) = 0$ với mọi $h \in X$.

b/Theo giả thiết thì $f'(x^*)h = f'_G(x^*)h = \delta f(x^*, h) = 0\ \forall h \in X$. □

Định lý 2.1.14. (**Định lý về điều kiện đủ**) *Xét bài toán tối ưu không ràng buộc* (2.12) *với* $X = I\!R^n$.

a) Nếu x^* *là cực trị địa phương của hàm* f *hai lần khả vi trên* $I\!R^n$ *thì* $\nabla^2 f(x^*) \geq 0$.

b) Ngược lại, nếu x^* *tại đó hàm* f *hai lần khả vi và* $\nabla f(x^*) = 0, \nabla^2 f(x^*) > 0$ *thì* x^* *là điểm cực tiểu địa phương của* f *trên* $I\!R^n$.

Trong đó

$$\nabla^2 f(x^*) := \begin{pmatrix} \frac{\partial^2 f(x^*)}{\partial^2 x_1} & \frac{\partial^2 f(x^*)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x^*)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x^*)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x^*)}{\partial^2 x_2} & \cdots & \frac{\partial^2 f(x^*)}{\partial x_2 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f(x^*)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x^*)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(x^*)}{\partial^2 x_n} \end{pmatrix}$$

Ma trận trên gọi là ma trận Heissian.

Chứng minh. Ta chứng minh b). Nếu $\nabla^2 f(x^*) > 0$, thì mọi véc tơ riêng của ma trận trên đều lớn hơn 0. Ta gọi giá trị riêng nhỏ nhất là λ. Do đó

$$\langle \nabla^2 f(x^*) x, x \rangle \geq \lambda \|x\|^2$$

với mọi $x \in I\!R^n$. Mặt khác gọi $\Delta x = x - x^*$, theo công thức Taylo ta có

$$\begin{aligned} f(x^* + \Delta x) - f(x^*) &= \nabla f(x^*) + \langle \nabla^2 f(x^*) \Delta x, \Delta x \rangle + 0(\|\Delta x\|^2) \\ &> 0 \end{aligned}$$

khi $\|\Delta\|^2$ đủ nhỏ. Từ đây ta suy ra x^* là điểm cực tiểu của f.

□

Ví dụ 2.1.15. *Cho hai môi trường đồng chất, nằm ở 2 phía của một mặt phẳng và hai điểm* a, b *nằm trong 2 môi trường đó. Tìm đường đi của ánh sáng từ* a *tới* b.

Trước tiên ta có nhận xét sau:

1 Không giảm tổng quát ta coi $a = (a_1, 0, 0); b = (b_1, b_2, 0), a_>0 > b_1, b_2 > 0$ và mặt phẳng ngăn cách là:$\{x = (x_1, x_2, x_3) \in I\!R^3, x_1 = 0\}$.

2. Ánh sáng truyền theo đường mà thời gian đi ngắn nhất (nguyên lý Fermat) nên trong môi trường đồng chất vận tốc không thay đổi nên ánh sáng phải đi theo đường thẳng.

3. Từ các kết luận trên suy ra cần xác định điểm $z = (0, z_2, z_3)$ nơi mà ánh sáng truyền từ môi trường này sang môi trường kia.

Gọi v_1, v_2 là vận tốc ánh sáng tương ứng trong 2 môi trường. Khi đó thời gian tưng ứng sẽ là: $|z-a|/v_1$ và $|z-b|/v_2$. Do đó bài toán trở thành

$$|z-a|/v_1 + |z-b|/v_2 \to \min.$$

Theo định lý Fermat thì

$$\frac{z_2}{v_1|z-a|} + \frac{z_2 - b_2}{v_2|z-b|} = 0 \text{ và } \frac{z_3}{v_1|z-a|} + \frac{z_3}{v_2|z-b|} = 0$$

Từ phương trình sau suy ra $z_3 = 0$ và z_2 được xác định duy nhất từ phương trình đầu. Gọi α_1, α_2 là góc giữa pháp tuyến và tia sáng. Khi đó

$$\sin\alpha_1 = z_2/|z-a| \text{ và } \sin\alpha_2 = z_2 - b_2/|z-b|.$$

Vậy nên

$$\sin\alpha_1/\sin\alpha_2 = v_1/v_2.$$

Đây là kết luận quen biết trong lý thuyết quang học.

2.1.2 Bài toán trơn với ràng buộc đẳng thức

Trước tiên ta xét ví dụ sau:

$$f_0(x) = x_1^2 + 4x_2^2 \to \min$$

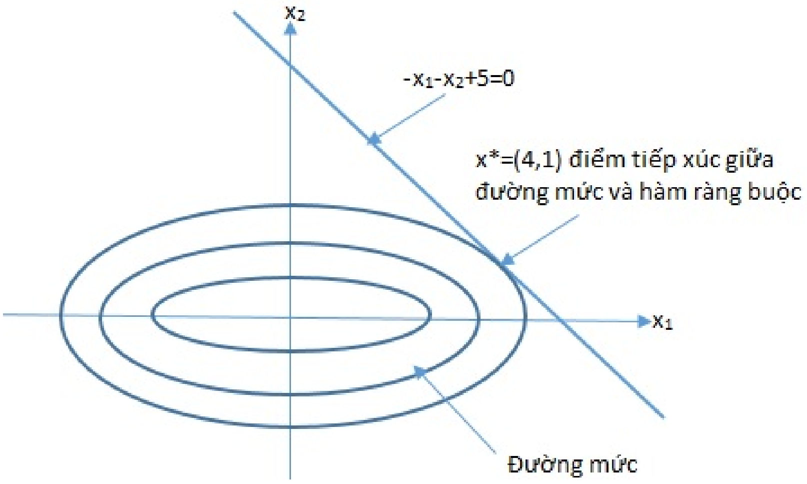
trong đó $x := (x_1, x_2)^T \in I\!R^2$. Áp dụng đinh lý Fermat $f_0'(x) = (2x_1, 8x_2) = (0, 0)$. Từ đây suy ra

Dễ thấy điều kiện cần cực trị này cho một nghiệm duy nhất $x^* = (0,0)^T$ và chính là nghiệm cực tiểu toàn cục của $f_0(x)$. Nếu thêm ràng buộc

$$f_1(x) = -x_1 - x_2 + 5 = 0$$

thì khi đó $(0,0)^T$ không còn là cực tiểu toàn cục nữa. Vì $f_0(0,0) = 0$ và $f(x) > 0$ với mọi x thỏa mãn $f_1(x) = 0$. Do đó $f_0(.)$ sẽ đạt cực tiểu tại điểm tiếp xúc x^* giữa đường thẳng $f_1(x) = 0$ và đường mức của $f_0(x)$.

Vì hai đường này tiếp xúc nên hai véc tơ f_0' và f_1' phải song song với nhau, tức là $f_0'(x^*) + \lambda_1 f_1'(x^*) = 0$.



Hình 2.6: **Nghiệm của bài toán bị ràng buộc**

$$2x_1^* - \lambda_1 = 0, 8x_2^* - \lambda_1 = 0.$$

Do đó $\lambda_1 = 8, x_1^* = 4, x_2^* = 1$. Đây chính là nghiệm tối ưu của của bài toán với ràng buộc trên.

Bây giờ ta xét **bài toán trơn ràng buộc đẳng thức** như sau:

$$f_0(x) \to \inf; \tag{2.13}$$

$$f_i(x) = 0, i = 1, \ldots, m, f_i : \mathbb{R}^n \to \mathbb{R}. \tag{2.14}$$

Gọi

$$\mathcal{L}(x, \lambda_0, \lambda_1, \ldots, \lambda_m) := \sum_{i=0}^{m} \lambda_i f_i(x) \text{ là hàm Lagrange .} \tag{2.15}$$

Ta có định lý sau:

Định lý 2.1.15. (Quy tắc nhân tử Lagrange) *Cho $f_i, i = 0, \ldots, m$, khả vi liên tục trong lân cận $V \subset I\!R^n$ của x^*. Nếu x^* là nghiệm cực tiểu địa phương của bài toán* (2.13) *thì tồn tại các nhân tử Lagrange $\lambda_i \geq 0, i = 0, 1, \ldots, m$, sao cho chúng không cùng triệt tiêu và thỏa mãn*

$$\mathcal{L}_x(x^*, \lambda_0, \lambda_1, \ldots, \lambda_m) := \sum_{i=0}^{m} \lambda_i f_i'(x^*) = 0. \tag{2.16}$$

Nếu $f_1'(x^*), f_2'(x^*), \ldots, f_m'(x^*)$ độc lập tuyến tính thì $\lambda_0 \neq 0$ và có thể chọn bằng 1.

Bây giờ ta xét bài toán thứ hai.

(Đọc thêm) Cho X, Y là các không gian Banach, $f_0 : X \to I\!R$ và $F : X \to Y$. Xét bài toán

$$\begin{aligned} f_0(x) &\to \inf & (2.17)\\ F(x) &= 0. & (2.18) \end{aligned}$$

Định lý 2.1.16. (Quy tắc nhân tử Lagrange (L. A. Lyusternik)) *Giả thiết f, F khả vi Frechet tại $x_*, F(x_*) = 0$ và ảnh của ánh xạ $x \mapsto F'(x_*)x$ là đóng. Nếu x_* là cực tiểu địa phương của* (3.55) *thì tồn tại nhân tử Euler-Lagrange λ_0, y^* : chúng không cùng triệt tiêu, thỏa mãn phương trình Euler-Lagrange*

$$\mathcal{L}_x(x, \lambda_0, y^*) := \lambda_0 f_0(x) + \langle y^*, F(x) \rangle. \tag{2.19}$$

Nếu F khả vi liên tục và chính quy tại x_*, tức là $F'(x_*)X = Y$, thì $\lambda \neq 0$ và có thể chọn $\lambda_0 = 1$.

Ví dụ 2.1.16. *Cực tiểu hóa hàm sau: $f(x) = \prod_{i=1}^{n} x_i \to \sup; \sum_{i=1}^{n} x_i = a, x_i > 0, i = 1, \ldots, n$.*

Hàm Lagrange là

$$\mathcal{L} = \lambda_0 \prod_{i=1}^{n} x_i + \lambda_1 \sum_{i=1}^{n} x_i.$$

Theo định lý thì $\mathcal{L}_x = 0$. Rõ ràng λ_0 không thể bằng 0, do đó suy ra $\prod_{i\neq j} x_i^* = -\lambda_1/\lambda_0, j = 1, \ldots, n$. Vậy

$$x_1^* = x_2^* = \cdots = x_n^* = \frac{a}{n}.$$

Mặt khác tập $\{x = (x_1, \ldots, x_n) \mid f(x) \geq (\frac{a}{n})^n\}$ của hàm liên tục f là khác rỗng và compact nên x^* là điểm hàm đạt cực đại.

2.1.3 Bài toán trơn với ràng buộc tệp

Bài toán trơn với ràng buộc tệp là bài toán sau:

$$f(x) \to \min,\ f : D \to I\!R^n \text{ trong đó } D \subset I\!R^n \text{ là tập tùy ý, } f \text{ khả vi trên } D. \tag{2.20}$$

Ở chương 1 chúng ta đã sơ lược về nón tiếp tuyến khi nói về Định lý Lyusternik, ta sẽ trình bày những định nghĩa tương đương về phương tiếp tuyến và nón tiếp tuyến để phục vụ cho nghiên cứu các điều kiện tồn tại nghiệm tối ưu trong phần này.

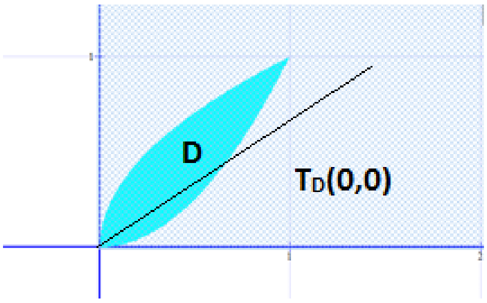
Định nghĩa 2.1.18. (**Phương tiếp tuyến**) *Cho $D \subset I\!R^n$ và $x^0 \in D$ ta nói vec tơ $0 \neq v \in I\!R^n$ là một phương tiếp tuyến của D tại x^0 nếu tồn tại dãy $\{x^k\} \subset D$ và $x^k \neq x^0$ và dãy số dương t_k đơn điệu giảm về 0 sao cho $v^k := (x^k - x^0)/t_k \to v)$ khi $k \to +\infty$.*

Do $x^k = x^0 + t_k v^k \in D$ nên định nghĩa trên có thể hiểu là, gần tùy ý x^0 và theo phương tùy ý sát v đều có những điểm thuộc D.

Định nghĩa 2.1.19. (**Nón tiếp tuyến**) *Tất cả các phương tiếp xúc của D tại $x^0 \in D$ và véc tơ 0 gọi là nón tiếp xúc của D tại x^0 ký hiệu là $T_D(x^0)$.*

Ví dụ 2.1.17. (**Nón tiếp tuyến**)

- *Nếu $x^0 \in \mathrm{int}(D)$ khi đó $T_D(x^0) = I\!R^n$,*
- *nếu $D = \{x^0\}$ khi đó $T_D(x^0) = \{0\}$,*



Hình 2.7: **Nón tiếp tuyến của D tại** $(0,0)$

- *nếu* $D = I\!R^n_+$ *khi đó* $T_D(x^0) = D$,
- *nếu* $D = \{x \in I\!R^2 \mid x_1 \geq x_2^2, x_2 \geq x_1^2\}, x^0 = (0,0)^T$ *thì* $T_D(x^0) = \{v \in I\!R^2 \mid v \geq 0$.

Ta kiểm tra phương $v = (0,1)^T$ là phương tiếp xúc của D tại điểm $(0,0)$. Xét các điểm $x^k := (1/k^2, 1/k)$ và $t_k = 1/k$ và $v^k = (1/k, 1)$ rõ ràng $x^k = x^0 + tv^k \in D$ với mọi k và $v^k \to v = (0,1)^T$ khi $k \to +\infty$

Định nghĩa 2.1.20. (**Tập hướng giảm**) *Cho* $x^* \in D$ *ta gọi tập* $D(x^*) := \{x \subset I\!R^n \mid \langle \nabla f(x^*), x\rangle < 0\}$ *là tập hướng giảm của* f *tại* x^*.

Định lý 2.1.17. (**Định lý về điều kiện cần**) *Giả sử* x^* *là điểm cực tiểu địa phương của* $f(x)$ *trên* D, *khi đó*

$$\langle \nabla f(x^*), x\rangle \geq 0 \text{ với mọi } x \in T_D(x^*),$$

tức là $T_D(x^*) \cap D(x^*) = \emptyset$.

Hệ quả 2.1.1. (**Định lý về điều kiện cần**) *Giả sử* x^* *là điểm cực tiểu địa phương của* $f(x)$ *trên* D *và là điểm trong của* D *khi đó* $\nabla f(x^*) = 0$ *tức là* $T_D(x^*) \cap D(x^*) = \emptyset$.

Định nghĩa 2.1.21. (**Hướng chấp nhận được**) *Cho* $x^0 \in D \subset I\!R^n$ *véc tơ* $0 \neq v \in I\!R^n$ *gọi là hướng chấp nhận được tại* x^0 *nếu tồn tại* $t_0 > 0$ *sao cho* $x^0 + tv \in D$ *với mọi* $t \in [0, t_0]$.

Định lý 2.1.18. (**Định lý về điều kiện cần cấp 2**) *Nếu f khả vi 2 lần và liên tục trên D. x^* là điểm cực tiểu địa phương của $f(x)$ trên D, thì với mỗi hướng chấp nhận được $v \in I\!R^n$ tại x^*ta có:*

a $\langle \nabla f(x^*), v \rangle \geq 0$.

b) Nếu $\langle \nabla f(x^*), v \rangle = 0$ *thì* $\langle \nabla^2 f(x^*)v, v \rangle \geq 0$.

Chứng minh. Theo công thức Taylo với $0 \leq t \leq t_0$ ta có:

$$f(x + tv) - f(x^*) = \langle \nabla f(x^*), v \rangle + \langle \nabla^2 f(x^*)v, v \rangle + o(t^2 \, |v|)$$

nên kết hợp giả thiết ta được $\langle \nabla^2 f(x^*)v, v \rangle \geq 0$. □

Định lý 2.1.19. (**Định lý về điều kiện đủ cấp 1**) *Cho f khả vi liên tục trên D, x^* là điểm cực tiểu địa phương của $f(x)$ trên $D \subset I\!R^n$. Nếu $x^* \in D$ thỏa mãn điều kiện*

$$\langle \nabla f(x^*), v \rangle \geq 0$$

vói mọi $v \in T_D(x^*), v \neq 0$, thì x^* là điểm cực tiểu địa phương chặt của f trên D.

Định lý 2.1.20. (**Định lý về điều kiện đủ cấp 2**) *Cho f 2 lần khả vi liên tục tại điểm x^*. Nếu $x^* \in D$ thỏa mãn điều kiện $\nabla f(x^*) = 0$ và $\langle \nabla^2 f(x^*)v, v \rangle > 0$ với mọi $v \in T_D(x^*), v \neq 0$ thì x^* là điểm cực tiểu địa phương chặt của f trên D.*

Hệ quả 2.1.2. *Nếu $x^* \in D$ thì điều kiện đủ để x^* là điểm cực tiểu địa phương chặt của $f(x)$ trên D là $\nabla^2 f(x^*)$ xác định dương trên $I\!R^n$.*

Ví dụ 2.1.18. *Cho*

$$f(x) = x_1^2 + 2x_1x_2 - 2x_1 + 2x_2, (x_1, x_2) \in D = I\!R^2_+.$$

Dễ thấy rằng tại hàm đạt cực tiểu tại $x^* = (1, 0)^T$ và $f_{\min} = -1$. Mặt khác

$$\frac{\partial f(x)}{\partial x_1} = 2x_1 + 2x_2 - 2, \frac{\partial f(x)}{\partial x_2} = 2x_1 + 2$$

và

$$\nabla^2 f(x) = \begin{pmatrix} 2 & 2 \\ 2 & 0 \end{pmatrix}$$

Các đạo hàm riêng

$$\frac{\partial f(x^*)}{\partial x_1} = 0, \frac{\partial f(x^*)}{\partial x_2} = 4 \neq 0.$$

Mặt khác dễ chứng minh được hướng chấp nhận được $v = (v_1, v_2)$ của tập D thỏa mãn $v_2 \geq 0$. Do đó $\langle \nabla^2 f(x^*)v, v\rangle = 4v_2 > 0$, vậy điều kiện a) thỏa mãn. Ta cũng có thể tính được $T_D(x^*) = \{v = (v_1, 0)^T \mid v_1 \geq 0\}$ nên $\langle v\nabla^2 f(x^*), v\rangle = 2v_1^2 \geq 0$ vì thế b) được kiểm tra.

2.1.4 Bài toán trơn với ràng buộc bất đẳng thức và đẳng thức

Trong mục này chúng ta xét bài toán

$$f_0(x) \to \min; \tag{2.21}$$

$$f_i(x) \leq 0, i = 1, \ldots, m, h_j(x) = 0, j = 1, 2, \ldots, p, \tag{2.22}$$

trong đó $f_i, h_j I\!R^n \to I\!R$ là các hàm khả vi. Ký hiệu

$$D := \{x \in I\!R^n \mid f_i(x) \leq 0, i = 1, \ldots, m, g_j(x) = 0, j = 1, 2, \ldots, p.$$

Định nghĩa 2.1.22. (**Nón chấp nhận được tuyến tính hóa**) *Cho $x^0 \in D$ gọi $S(x^0)$ là tập tất cả các véc tơ v nghiệm đúng của hệ phương trình và bất phương trình tuyến tính*

$$\begin{cases} \langle \nabla f_i(x^0), v\rangle \leq 0, i \in I(x^0) \\ \langle \nabla g_i(x^0), v\rangle = 0, j = 1, 2, \ldots, p \end{cases}$$

trong đó $I(x^0) := \{i \mid g_i(x^0) = 0\}$. $S(x^0)$ được gọi là nón chấp nhận được tuyến tính hóa.

Định nghĩa 2.1.23. (**Điểm chính quy**) *Điểm x^0 gọi là điểm chính quy nếu $T_D(x^0) = S(x^0)$.*

Nhận xét 2.1.2. *Điểm x^0 là chính quy nếu nó thỏa mãn một trong các tính chất sau:*

- *Các hàm $f_i, i \in I(x^0)$ và g_j là các hàm afine,*
- *Các véc tơ $\nabla f_i, i \in I(x^0), \nabla g_j$ là độc lập tuyến tính,*
- *g_j là hàm afine, f_i là các hàm lồi và tồn tại $u^0 \in D$ sao cho $g_i(u^0) < 0$ với mọi i mà g_i không phải là hàm afine (điều kiện chính quy Slater). Điều kiện này còn đản bảo mọi điểm chấp nhận được đều là điểm chính quy.*

Ví dụ 2.1.19. *Cho $D = \{x = (x_1, x_2) \mid f_1(x) = x_1^2 + x_2^2 - 2 \leq 0, g_1 = -2x_1 + x_2^2 + 1 = 0\}$*

Tại điểm chấp nhận được $x^0 = (1,1)^T$ ta có $f_1(x^0) = g_1(x^0) = 0$ và $\nabla f_1 = (2,2)^T, \nabla g_1 = (-2,2)^T$. Hai véc tơ này độc lập tuyến tính nên x^0 là điểm chính quy.

Định lý 2.1.21. (**Định lý về điều kiện cần cấp 1- Karush-Kuhn-Tucker**) *Giả sử các hàm $f, f_i(x) \leq 0, i = 1, \dots, m, g_j(x) = 0, j = 1, 2, \dots, p$ khả vi trên tập mở chứa D, $x^* \in D$ là điểm cực tiểu địa phương của bài toán 2.22 và 2.22 và x^* là điểm chính quy. Khi đó, tồn tại véc tơ $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m) \geq 0$ và $\mu = (\mu_1, \dots, \mu_p)$ thỏa mãn*

$$\begin{cases} \nabla f_i(x^*) + \sum_{i=1}^m \lambda_i \nabla(x^*) + \sum j = 1^p, \mu \nabla g_j(x^*) = 0 \\ \lambda_i f_i(x^0) = 0, i = 1, 2, \dots, m, \\ g_j(x^*) \leq 0, i = 1, 2, \dots, p. \end{cases}$$

Ví dụ 2.1.20. *Tìm phương án tối ưu của bài toán sau:*

$$\min\{-\log x_1 - \log x_2 \mid x_1 + x_2 \leq 0, x_1 \geq 0, x_2 \geq 0\}.$$

Hàm Lagrange có dạng:

$$\mathcal{L}(x, \lambda) := -\log x_1 - \log x_2 + \lambda(x_1 + x_2 - 2).$$

Điều kiện Kuhn-Tucker và ràng buộc là hệ phương trình:

$$\begin{cases} -\frac{1}{x_1} + \lambda \geq 0 \\ -\frac{1}{x_2} + \lambda \geq 0 \\ x_1(\lambda - \frac{1}{x_1} = 0 \\ x_2(\lambda - \frac{1}{x_2} = 0 \\ \lambda x_1 + x_2 - 2 = 0 \\ x_1 \geq 0, x_2 \geq 0, \lambda \geq 0. \end{cases}$$

- Nếu $\lambda = 0$, thì thay vào phương trình thứ 2, 3 ta nhận được -1=0. Do đó $\lambda \neq 0$.
- Từ phương trình cuối cùng suy ra $x_1 + x_2 - 2 = 0$ và từ phương trình 3, 4 suy ra $x_1, x_2 \neq 0$. Vậy $x_1 = x_2 = \frac{1}{\lambda}$.
- Thay vào phương trình $x_1 + x_2 = 2$ ta suy ra $x^* = (1,1)^T$ và $\lambda = 1$. Điểm x^* chính là điểm cực tiểu toàn cục. Lưu ý rằng với x^* nhận được ta chưa thể kết luận x^* là điểm cực tiểu, điều này sẽ trình bày sau.

2.2. BÀI TOÁN LỒI

2.2.1 Bài toán lồi không có ràng buộc

Bài toán được xét ở đây là

$$f(x) \to \inf,\ \ f : X \to I\!R \text{ là hàm lồi.} \tag{2.23}$$

Khi nghiên cứu Bài toán (2.12) nếu bỏ giả thiết trơn thì khi đó không thể dùng các định lý liên quan đến đạo hàm để tìm điểm cực trị. Để khảo sát các điểm cực trị cần phải mở rộng tiếp các khái niệm liên quan đến đạo hàm. Do đó khái niệm dưới vi phân được xét đến.

Định nghĩa 2.2.24. (**Dưới vi phân**) *Cho $f : D \subset X \to I\!R$ gọi gọi dưới vi phân của hàm tại điểm $x^* \in D$ ký hiệu là $\partial f(x^*)$ xác định như sau:*

$$\partial f(x^*) := \{y \in X^* \mid f(x) - f(x^*) \geq \langle y, x - x^* \rangle, \forall x \in D\} \tag{2.24}$$

Định lý 2.2.22. (**Định lý về điều kiện cần và đủ**) *Hàm lồi $f(x)$ nhận giá trị cực tiểu tại x^* khi và chỉ khi $0 \in \partial f(x^*)$.*

2.2.2 Một số nhận xét và ví dụ

Để nghiên cứu tiếp về điểm cực trị của hàm lồi, ta có định nghĩa sau: Cho $z \in X$ cố định. Ta gọi

Định nghĩa 2.2.25. (**Đạo hàm theo hướng**)
Nếu tồn tại $\lim_{\lambda \to 0} \frac{f(x+\lambda z)-f(x)}{\lambda}$ thì giới hạn đó gọi là đạo hàm theo hướng z của f tại x và ký hiệu là $f'(x,z)$. tức là:

$$f'(x,z) := \lim_{\lambda \to 0} \frac{f(x+\lambda z)-f(x)}{\lambda}$$

Với định nghĩa trên ta dễ suy ra:

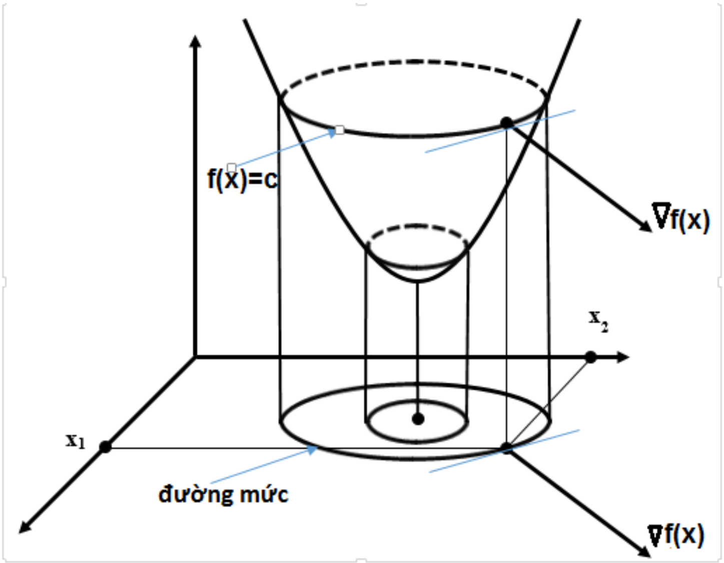
- $f'(x,z) \leq f(x+z) - f(x)$.
- Nếu f là khả vi thì
$$f'(x,z) = \langle \bigtriangledown f(x), z \rangle,$$
- $f'(x,z) = \langle \bigtriangledown f(x), z \rangle = | \bigtriangledown f(x) || z | \cos(\bigtriangledown f(x), z)$,

trong đó
$$\bigtriangledown f(x) := \Big(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \dots, \frac{\partial f(x)}{\partial x_n}\Big)^T$$
là hình chiếu của véc tơ $\bigtriangledown f(x)$ lên hướng z. (xem hình vẽ)
Cũng cần chú ý rằng đạo hàm có hướng thì véc tơ z là cố định, đó là điều khác với biến phân bậc nhất dù rằng vế phải na ná giống nhau.

Ví dụ 2.2.21. *(Dưới vi phân của hàm chuẩn $f(x) = \|x\|$)*

Từ định nghĩa suy ra $x^* \in \partial\|0\|$ khi và chỉ khi $\|z\| \geq |\langle x^*, z\rangle|$ $\forall z \in X$ tương đương với $\|x^*\| \leq 1$. Điều này có nghĩa là

$$\partial\|0\| = \{x^* \in X^* \mid \|x^*\| \leq 1\}.$$

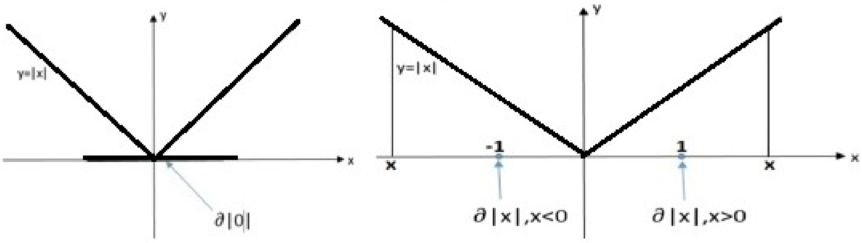


Hình 2.8: **Véc tơ pháp và đường mức**

(Nếu $X = I\!R^n$ thì X^* cũng là $I\!R^n$). Ta có thể chứng minh, với $x \neq 0$ thì

$$\partial \|x\| = \{x^* \in X^* \mid \|x^*\| = 1 \text{ và } \langle x^*, x\rangle = \|x\|\}.$$

Như vậy nếu $y = |x|$ trong $I\!R$ thì $\partial|0| = [-1, 1]$ còn $\partial|x| = \{1\}$ với $x > 0$ và $\partial|x| = \{-1\}$ với $x < 0$.



Hình 2.9: **Dưới vi phân của hàm** $y = |x|$

Ví dụ 2.2.22. *(Dưới vi phân của hàm chỉ định)*

$$\delta(x \mid A) := \begin{cases} 0 & khi\ x \in A \\ +\infty & khi\ x \notin A. \end{cases}$$

Với mọi $x \in A$ thì $\partial\delta(a \mid A) \neq \emptyset$ vì nó đều chứa 0. Ta có thể chứng minh được rằng

$$\partial\delta(x \mid A) = N(x|A|) = \{x^* \in X \mid \langle x^*, z - x\rangle \leq 0\}.$$

2.2.3 Bài toán quy hoạch lồi có ràng buộc bao hàm thức

Xét bài toán

$$f_0(x) \to \inf,\ x \in A. \tag{2.25}$$

Định lý 2.2.23. *Cho X là một không gian tô pô tuyến tính lồi địa phương, f_0 là một hàm lồi trên X và liên tục tại x^*, $A \subset X$ là tập lồi. Khi đó x^* là nghiệm cực tiểu toàn cục của bài toán* (2.25) *khi và chỉ khi*

$$0 \in \partial f_0(x^*) + \partial(\delta(x^*|A)),$$

điều này tương đương với

$$\exists y^* \in \partial f_0(x^*) : -y^* \in N(x^*|A).$$

Hệ quả 2.2.3. *Cho $y^* \in X^*$ và $A \subset X$ lồi. Điều kiện cần và đủ để x^* là nghiệm cực tiểu của bài toán*

$$\langle y^*, x\rangle \to \inf, x \in A$$

là $-y^* \in N(x^*|A)$.

2.2.4 Bài toán lồi với ràng buộc bất đẳng thức,

Cho X là một không gian tuyến tính, $f_i : X \to I\!R$ (với $i = 0, 1, \dots m$ là các hàm lồi, $A \subset X$ là tập lồi. Xét bài toán lồi

$$f_0(x) \to \inf, \tag{2.26}$$

$$D = \{x \in A \mid f_1(x) \le 0, f_2(x) \le 0, \dots, f_m(x) \le 0\}. \tag{2.27}$$

Đối với bài toán này, vì các hàm f_i chỉ cho là lồi nên định lý về nhân tử Lagrange không còn làm việc được. Do đó cần phải có những biến đổi phù hợp để sao cho có thể có được một kết quả gần như định lý này. Định lý Kuhn-Tucker chính là một cải biên hợp lý định lý trên theo hướng này.

Hàm Lagrange của bài toán trên:

$$\mathcal{L}(x, \lambda_0, \ldots, \lambda_m) := \sum_{i=0}^{m} \lambda_i f_i(x), \lambda_i \in I\!R\, i = 0, 1, \ldots, m$$

Tập phương án chấp nhận được của bài toán là:

$$D = \cap_{i=1}^{m} \{x \in A \mid f_i(x) \leq 0\}.$$

Định lý 2.2.24. (Định lý Kuhn-Tucker),(xem [6], trang 76)

(a) *Nếu x^* là nghiệm cực tiểu của bài toán* (2.26) *thì tồn tại các nhân tử Lagrange $\lambda_i \geq 0, i = 0, \ldots, m$, sao cho chúng không cùng triệt tiêu, thỏa mãn điều kiện Kuhn-Tucker*

$$\mathcal{L}(x^*, \lambda_0, \ldots, \lambda_m) = \min_{x \in A} \mathcal{L}(x, \lambda_0, \ldots, \lambda_m) \tag{2.28}$$

và điều kiện bù

$$\lambda_i f_i(x^*) = 0 \text{ với mọi } i = 1, \ldots, m. \tag{2.29}$$

Nếu thêm điều kiện Slater

$$\exists z \in A : f_i(z) < 0 \text{ với mọi } i = 1, \ldots, m,$$

thoả mãn thì $\lambda_0 \neq 0$ và có thể coi $\lambda_0 = 1$.

(b) *Nếu tồn tại x^* thỏa mãn* (2.28), (2.29) *với $\lambda_0 = 1$ thì x^* là nghiệm cực tiểu của bài toán* (2.26)–(2.27).

Định lý này được W. H. Kuhn và A. W. Tucker chứng minh năm 1951 được coi là công trình khai phá Quy hoạch lồi. Điều kiện Slater được M. Slater đưa ra vào năm 1950.

(a) Xét tập

$$\begin{aligned} C := \{(\mu_0, \ldots \mu_m) \in I\!R^{m+1} \quad | \quad & \exists x \in A : f_0(x) - f_0(x^*) < \mu_0 \\ & f_1(x) \leq \mu_1, \ldots f_m(x) \leq \mu_m\} \end{aligned}$$

Ta thấy, nếu $\mu_i > 0$ với mọi $i = 1, \ldots m$ thì $(\mu_0, \ldots, \mu_m) \in C$ (vì $\mu_0 > f_0(x^*) - f_0(x^*), \mu_i > f_i(x^*)$). Do đó $\mathrm{int}(C) \neq \emptyset$. Vì f_i là lồi nên C là lồi. Thêm vào đó dễ thấy rằng $0 \notin C$. Theo định lý tách ta có thể tách C và véc tơ 0 bằng một phiếm hàm tuyến tính khác 0, tức là tồn tại $m+1$ số $\lambda_0, \ldots \lambda_m$ không đồng thời triệt tiêu sao cho

$$\sum_{i=0}^{m} \lambda_i \mu_i \geq 0 \ \forall (\mu_0, \ldots, \mu_m) \in C. \tag{2.30}$$

Mặt khác, do $\mathrm{int}\mathbb{R}_+^{m+1} \subset C$ nên với mọi $i \in \{0, 1 \ldots, m\}$ ta rút ra từ biểu thức trên

$$\lambda_i \geq \lim_{\mu_j \to 0, j \neq i} \frac{1}{\mu_i}\big(- \sum_{j \neq i} \lambda_j \mu_j \big),$$

nên $\lambda_i \geq 0$ với mọi $i = 0, 1, 2, \ldots, m$.

Cho $\mu_0 \to f_0(x) - f_0(x^*)$ và $\mu_i = f_i(x), 1 \leq i \leq m$, ta suy ra từ (2.30)

$$\sum_{i=0}^{m} \lambda_i f_i(x) \geq \lambda_0 f_0(x^*) \quad \forall x \in A \tag{2.31}$$

Nếu $f_i(x^*) = -\alpha < 0$ cho một chỉ số i nào đó và với mọi $\epsilon > 0$. $\mu_i := -\alpha$ $\mu_j := \epsilon$ ($j \neq i$) thì suy ra $(\mu_0, \ldots, \mu_m) \in C$. Thay vào (2.30) và cho $\epsilon \to 0$ ta được $-\lambda_i \alpha \geq 0$. Suy ra $\lambda_i = 0$ nếu $f_i(x^*) < 0$. Vì vậy

$$\lambda_i f_i(x^*) = 0 \text{ với mọi } i = 1, \ldots, m.$$

Kết hợp biểu thức cuối với (2.31) ta suy ra

$$\sum_{i=0}^{m} \lambda_i f_i(x) \geq \sum_{i=0}^{m} \lambda_i f(x^*) \tag{2.32}$$

Giả sử điều kiện Slater thỏa mãn. Nếu $\lambda_0 = 0$ thì tồn tại ít nhất một $\lambda_i > 0$ và do đó

$$\sum_{i=0}^{m} \lambda_i f_i(x) < 0 = \sum_{i=0}^{m} \lambda_i f_i(x^*),$$

điều này mâu thuẫn với (2.32). Vậy điều kiện Slater kéo theo $\lambda_0 \neq 0$.

(b) Nếu (2.28)–(2.29) thỏa mãn với $\lambda_0 = 1$ thì với mọi phương án chấp nhận được x của bài toán (2.26) ta có

$$f_0(x) \geq f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) \geq \sum_{i=0}^{m} \lambda_i f_i(x^*) = f_0(x^*),$$

tức x^* là nghiệm tối ưu của bài toán (2.26)–(2.27).

Định lý 2.2.25. (Dạng dưới vi phân của Định lý Kuhn-Tucker) *Giả thiết rằng X là một không gian Hausdorff lồi địa phương và $f_i\, i = 1, \ldots, m$, là các hàm lồi, cùng liên tục ít nhất tại một điểm của tập lồi $A \subset I\!R^n$. Cho x^* là một nghiệm chấp nhận được của bài toán* (2.26)–(2.27).

(a) Nếu x^* là nghiệm cực tiểu của bài toán thì tồn tại các nhân tử Lagrange $\lambda_i \geq 0, i = 0, \ldots, m$, sao cho chúng không cùng triệt tiêu, thỏa mãn phương trình

$$0 \in \sum_{i=0}^{m} \lambda_i \partial f_i(x^*) + N(x^*|A) \tag{2.33}$$

và

$$\lambda_i f_i(x^*) = 0 \text{ với mọi } i = 1, \ldots, m, \tag{2.34}$$

trong đó $N(x^*|A) := \{y \in X^* \mid \langle y, x - x^* \rangle \leq 0 \;\; \forall x \in A$ là nón pháp tuyến của A tại x^*. Nếu điều kiện Slater

$$\exists z \in A : f_i(z) < 0 \text{ với mọi } i = 1, \ldots, m,$$

thỏa mãn thì $\lambda_0 \neq 0$ và có thể coi $\lambda_0 = 1$.

(b) Nếu tồn tại x^* thỏa mãn (2.33), (2.34) với mọi $\lambda_0 = 1$ thì x^* là nghiệm cực tiểu của bài toán (2.26)–(2.27).

Nhận xét 2.2.1. *Nếu $A = I\!R^n$ thì khi đó $N(x^*|A) = \{0\}$, nên biểu thức* (2.33) *được thay bởi*

$$0 \in \sum_{i=0}^{m} \lambda_i \partial f_i(x^*). \tag{2.35}$$

Ví dụ 2.2.23. *Cho hai điểm ngoài hình tròn đơn vị. Tìm một điểm thuộc hình tròn đó sao cho khoảng cách đến hai điểm ấy là nhỏ nhất.*

Bài toán có dạng giải tích là:

$$f_0(x) = \|x - y\| + \|x - z\| \to \inf, f_1(x) = \|x\| - 1,$$

trong đó $|y| > 1, |z| > 1$ là hai điểm cho trước. Đây là bài toán cực tiểu của hàm liên tục trên tập compact nên luôn tồn tại nghiệm.

Vì điều kiện Slater thỏa mãn ($f_1(0) < -1 < 0$) nên sử dụng Định lý 2.2.25 với $\lambda_0 = 1$. Theo đó tồn tại $\lambda_1 \geq 0$ sao cho

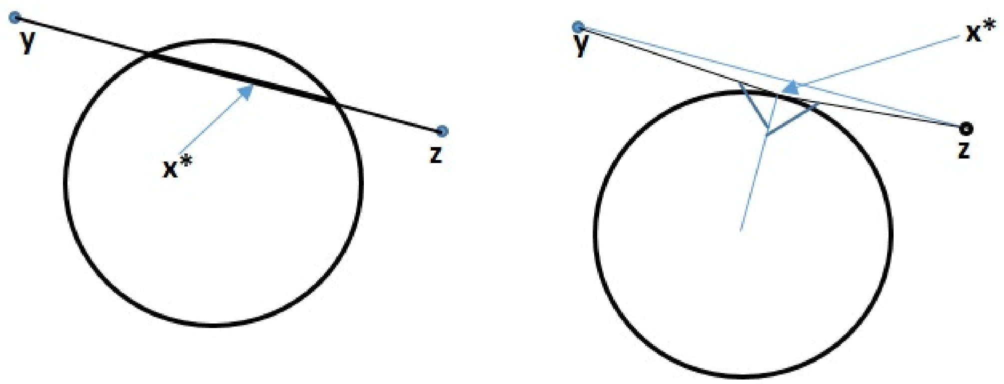
$$0 \in \partial\|x^* - y\| + \partial\|x^* - z\| + \lambda_1\partial\|x^*\|$$

$$\lambda_1(\|x^*\| - 1) = 0.$$

Ta biết rằng dưới vi phân của chuẩn trong không gian Euclid tại điểm khác 0 là véc tơ đơn vị từ 0 hướng tới điểm ấy. Vì vậy, khi $\|x^*\| = 1$ ta có

$$0 = \frac{x^* - y}{\|x^* - y\|} + \frac{x^* - z}{\|x^* - z\|} + \lambda_1 x^*.$$

Điều này có nghĩa là góc giữa hai đoạn $[0, x^*]$ và $[y, x^*]$ bằng góc giữa hai đoạn $[0, x^*]$ và $[z, x^*]$. Như vậy ta được ba phương trình với ba ẩn số, do đó có thể giải để tìm x^*.



Hình 2.10: **Vị trí của nghiệm ở ví dụ trên**

Trong trường hợp $\|x^*\| < 1$, điều kiện bù kéo theo $\lambda_1 = 0$, khi đó

$$0 = \frac{x^* - y}{|x^* - y|} + \frac{x^* - z}{|x^* - z|},$$

nghĩa là x^* nằm trên đoạn $[y, z]$. Điều đó chỉ xẩy ra khi đoạn $[y, z]$ cắt đường tròn đơn vị. Hiển nhiên, lúc đó giao giữa $[y, z]$ và hình tròn đơn vị là tập nghiệm tối ưu.

Định lý 2.2.26. (xem [8], trang 47). *Xét bài toán sau*

$$\begin{aligned} &\langle Mx, x\rangle + \langle b, x\rangle \to \inf; \\ &D = \{x \in I\!R^n \mid \langle C_i, x\rangle \le d_i,\, i = 1, \ldots, m\}, \end{aligned}$$

trong đó $M \in I\!R^{n\times n}$ *là ma trận đối xứng,* $C_i \in I\!R^n$, $i = 1, \ldots, m$. *Khi đó, nếu* x^* *là điểm cực tiểu địa phương thì tồn tại các nhân tử Lagrange* $\lambda_i \ge 0,\ i = 1, \ldots, m$, *sao cho chúng thỏa mãn các điều kiện*

$$(2Mx^* + b) + \sum_{i=1}^{m} \lambda_i C_i = 0,$$

và

$$\lambda_i(\langle C_i, x^*\rangle - d_i) = 0 \text{ với mọi } i = 1, \ldots, m.$$

Định lý 2.2.27. (Xem [8], trang 79). *Xét bài toán (P) với D là tập lồi đa diện, khi đó*

(a) *Nếu M là ma trận đối xứng xác định dương và* $D \neq \emptyset$ *thì bài toán có điểm cực tiểu toàn cục duy nhất.*

(b) *Nếu M là ma trận đối xứng xác âm thì điểm cực tiểu địa phương của bài toán là một điểm cực biên của D.*

Nhận xét 2.2.2. *Kết luận* (b) *của định lý trên tương đương với phát biểu sau "Nếu M đối xứng xác định dương nên điểm cực đại địa phương của bài toán (P) là điểm cực biên của D."*

Ghi chú Các kết quả được trình bày ở các mục trên được trích từ các tài liệu: [6], [7], [8],[9], [10],. . . và [11].

2.2.5 Điểm yên ngựa và định lý Kuhn-Tucker

Từ đây ta xét bài toán quy hoạch lồi với ràng buộc bất đẳng thức trong không gian Euclid và hàm mục tiêu $f_0(x)$ thay bởi $f(x)$., tức là bài toán quy hoạch lồi có dạng:

$$f(x) \;\to\; \inf, \tag{2.36}$$

$$D = \{x \in A \mid f_1(x) \;\le\; 0, f_2(x) \le 0, \ldots, f_m(x) \le 0, \}. \tag{2.37}$$

Định nghĩa 2.2.26. (**Hướng chấp nhận được**) *Cho phương án* $\overline{x}$*, ta nói véc tơ* $z \in I\!R^n$ *là hướng chấp nhận được tại* $\overline{x}$ *nếu tồn tại* $\epsilon > 0$ *sao cho* $\overline{x} + \epsilon z$ *cũng là phương án (do đó* $\forall \lambda : 0 \le \lambda \ge \epsilon$ *thì* $\overline{x} + \lambda z$ *cũng là phương án của bài toán* (2.36)–(2.37).

Ta có nhận xét sau: Với D là tập lồi, gọi $D_{\overline{x}}$ là tập tất cả các hướng chấp nhận được tại $\overline{x}$. Khi đó, $z = \epsilon'(x - \overline{x})$ với $x \in D$ cũng là hướng chấp nhận được. Thật vậy:

$$\overline{x} + \epsilon z = \overline{x} + \epsilon\epsilon'(x - \overline{x}) = \epsilon\epsilon' x + (1 - \epsilon\epsilon')\overline{x} \in D(\text{do } D \text{ lồi }).$$

Định lý 2.2.28. (**Điều kiện cần và đủ**) *Phương án* x^* *là phương án tối ưu khi và chỉ khi*

$$f'(x^*; z) \ge 0 \quad \forall z \in D_{x^*}. \tag{2.38}$$

Chứng minh. Điều kiện cần: Cho x^* là phương án tối ưu và $z \in D_{x^*}$. Khi đó tồn tại $\epsilon > 0$ đủ nhỏ để $x^* + \epsilon z \in D_{x^*}$ do đó

$$\frac{f(x^* + \epsilon z) - f(x^*)}{\epsilon} \ge 0.$$

Cho $\epsilon \to 0^+$ ta suy ra điều kiện cần.

Điều kiện đủ: Giả sử biểu thức (2.38) thỏa mãn, tức là $\delta f(x^*; z) \ge 0 \forall z \in D_{x^*}$. Theo tính chất của hàm lồi ta có $f(x^* + z) - f(x^*) \ge \delta f(x^*; z)$, nên thay $z := x - x^*$ ta được

$$\begin{aligned}
\forall x \in D: \;\; & f(x^* + (x - x^*)) - f(x^*) \ge \delta f(x^*, x - x^*) \\
& \Rightarrow \;\; f(x) - f(x^*) \ge \delta f(x^*, z) \\
& \Rightarrow \;\; f(x) \ge f(x^*).
\end{aligned}$$

Do đó x^* là phương án tối ưu của bài toán (2.36)–(2.37). $\square$

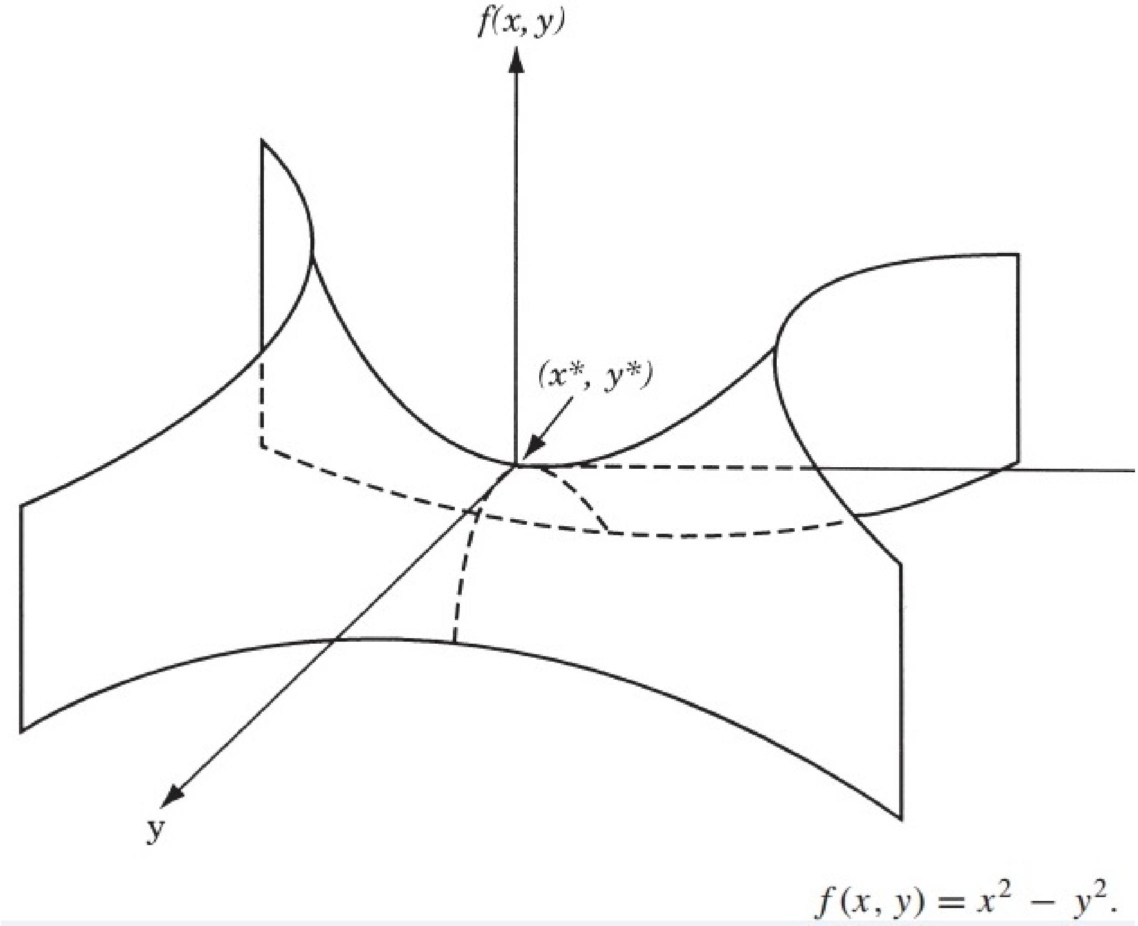
Hệ quả 2.2.4. *điểm yên ngựa Nếu f khả vi và $\frac{\partial f(x)}{\partial x_j} = 0$ $(j = 1, \ldots m)$ tại $x \in D_{x^*}$ thì x^* là phương án tối ưu.*

Định nghĩa 2.2.27. (Điểm yên ngựa) *Ta nói một điểm $(\overline{x}, \overline{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^m$ là điểm yên ngựa (hay điểm đèo) của hàm Lagrange*

$$\mathcal{L}(x, \lambda) := f(x) + \sum_{i=1}^{m} \lambda_i f_i(x),$$

trong đó $\lambda := (\lambda_1, \lambda_2, \ldots, \lambda_m)$, nhận các giá trị thực, nếu

$$\overline{x} \in A, \overline{\lambda} \geq 0$$
$$\forall x \in A\ \forall \lambda \geq 0 : \mathcal{L}(\overline{x}, \lambda) \leq \mathcal{L}(\overline{x}, \overline{\lambda}) \leq \mathcal{L}(x, \overline{\lambda}). \tag{2.39}$$



Hình 2.11: **Điểm yên ngựa**

Từ định nghĩa và biểu thức (2.39) ta thấy, khi cố định $x = \overline{x}$, thì $(\overline{x}, \overline{\lambda})$ là điểm "cao" nhất của $\mathcal{L}(x, \lambda)$. Khi cố định $\lambda = \overline{\lambda}$ thì $(\overline{x}, \overline{\lambda})$ lại là điểm " thấp " nhất.

Định lý 2.2.29. (Phát biểu khác của Định lý Kuhn – Tucker) *Giả sử bài toán quy hoạch lồi thỏa điều kiện Slater*

$$\exists z \in A : f_i(z) < 0 \quad \forall i = 1, \ldots, m.$$

Khi đó $x^* \in D$ là phương án tối ưu khi và chỉ khi $\exists \lambda^* \in I\!R^m, \lambda^* \geq 0$ sao cho (x^*, λ^*) là điểm yên ngựa của hàm Lagrange $\mathcal{L}(x, \lambda)$ trong miền $x \in D$.

Nhận xét 2.2.3. *Phần phải của* (2.39) *có nghĩa là :*

$$\mathcal{L}(x^*, \lambda^*) = \min_{x \in A} \mathcal{L}(x, \lambda^*).$$

Phần trái của (2.39) *có nghĩa là :*

$$\mathcal{L}(x^*, \lambda^*) = \max_{x \in A} \mathcal{L}(x^*, \lambda).$$

Do đó định lý có thể phát biểu tương đương như sau: x^* là phương án tối ưu của bài toán khi và chỉ khi tồn tại λ^* sao cho

(a) *x^* là lời giải của bài toán*

$$\min \mathcal{L}(x, \lambda^*).$$

(b) *x^* là lời giải của bài toán*

$$\max \mathcal{L}(x^*, \lambda).$$

Ta dễ dàng thấy rằng điều kiện (b) tương đương với

$$\sum_{i=1}^{m} \lambda_i^* f_i(x^*) = 0, \;\; f_i(x^* \leq 0, \;\; i = 1, 2, \ldots, m.$$

Thực vậy, vì $\lambda_i \geq 0, f_i(x^*) \leq 0$ nên $\sum_{i=1}^{m} \lambda_i f_i(x^*) = 0$. Ngược lại cũng tương tự. Từ đây, ta nhận được phát biểu ban đầu của Định lý Kuhn-Tucker.

Nhận xét 2.2.4. *Khi $A = I\!R^n$ và các hàm $f, f_i, i = 1, \ldots, m$ là khả vi, khi đó theo Hệ quả 2.2.4 thì điều kiện Kuhn-Tucker có dạng*

$$\frac{\partial f(x^*)}{\partial x_j} + \sum_{i=1}^{m} \frac{\partial f_i(x^*)\lambda_i^*}{\partial x_j} = 0, j = 1, \dot{;} n$$

hoặc

$$\bigtriangledown f(x^*) + \sum_{i=1}^{m} \frac{\bigtriangledown f_i(x^*)\lambda_i^*}{\partial x_j} = 0.$$

Nhận xét 2.2.5. Ý nghĩa và ứng dụng

- *Coi các nhân tử Lagrange λ_i như là tiền phạt (hay giá) phải trả nếu $f_i(x)$ vượt quá mức cho phép một đơn vị...(mức tối đa cho phép của $f_i(x)$ là 0.)*

- *Coi f(x) là chi phí phải trả nếu chọn điểm $x \in D$.*

Khi đó tổng số tiền phải trả là cho $x \in D$ là $\mathcal{L}(x, \lambda_i)$. Như vậy theo các định lý đã phát biểu ở các dạng khác nhau ở trên ta thấy rằng nếu chọn giá trị $\lambda^* = (\lambda_1^*, \ldots, \lambda_m^*)$ thích hợp thì lời giải x^* cũng sẽ là lời giải của bài toán đặt ra. Giá trị λ^* là thích hợp nếu số tiền phạt là lớn nhất đối với phương án x^* đã tìm.

Trong nhiều trường hợp thực tế, việc tìm cực tiểu của một hàm lồi bất kỳ $f(x)$ trên tập D có thể tiến hành tương đối đơn giản nhờ một thuật toán tốt, hoặc một cơ chế tự động nào đó có thể tin cậy được. Khi ấy cách làm thứ hai trên đây có thể thực hiện được bằng cách lấy một giá trị $\lambda > 0$ tìm $x \in D$ đạt cực tiểu hàm lồi $\mathcal{L}(x, \lambda)$ rồi điều chỉnh dần giá λ cho đến khi nó trở thành thích hợp. Như thế bài toán sẽ quy về điều chỉnh véc tơ m chiều λ và nếu m rất nhỏ so với n (số chiều của x) thì phương pháp này có thể hiệu lực hơn là tìm trực tiếp x. Đó là cơ sở khoa học của phương pháp xử lý hiện đại đối với nhiều bài toán quản lý kinh tế và điều khiển các hệ thống phức tạp nói chung.

Ghi chú Các kết quả được trình bày ở các mục trên có thể tìm thấy trong các tài liệu sau: [1],[2],. . .

2.3. PHƯƠNG PHÁP SỐ GIẢI BÀI TOÁN QUY HOẠCH LỒI

2.3.1 Phương pháp giải Frank và Wolfe

Phương pháp này cho phép tìm kiếm lời giải gần đúng của Bài toán QHL với ràng buộc tuyến tính sau:

$$f(x) \rightarrow \min \tag{2.40}$$

$$Ax \geq b, x \geq 0, \tag{2.41}$$

trong đó $A \in \mathbb{R}^{m\times n}$ và x là véc tơ cột trong $\mathbb{R}^n$. Đây là trường hợp đặc biệt của của bài toán QHL đã xét, tập phương án D là tập lồi đa diện.

Nội dung của phương pháp này là: Giải bài toán qua nhiều bước mà ở mỗi bước ta tuyến tính hóa bài toán và xây dựng dần một dãy phương án

$$x^{(0)}, x^{(1)}, \ldots \text{ sao cho } f(x^{(k)}) \text{ giảm dần và } f(x^{(k)}) \rightarrow \min_{x\in D} f(x).$$

Giả thiết của bài toán là:

(a) Hàm mục tiêu $f(x)$ khả vi liên tục (có đạo hàm riêng theo từng biến và các đạo hàm ấy liên tục).

(b) Với bất kì $x^{(k)} \in D$ hàm tuyến tính $\langle \bigtriangledown f(x^{(k)}), x\rangle$ luôn luôn vị chặn dưới trong miền D (điều này chắc chắn xảy ra nếu D là một đa diện lồi).

Thuật toán gồm các bước sau:

B1: Lấy một điểm bất kì $x^{(0)} \in D$ (tìm phương án bằng QHTT)

B2: Khi đã có $x^{(k)}$ ta tìm $x^{(k+1)}$ như sau:

$$\min_{x\in D}\langle \bigtriangledown f(x^{(k)}), x - x^{(k)}\rangle, x^{(k)} \text{ đã biết}$$

Đây là bài toán QHTT dạng chính tắc và tương đương với

$$\min_{x\in D}\sum_{i=1}^{n}\frac{\partial f(x^{(k)})}{\partial x_j}(x_j - x_j^k).$$

Do giả thiết (b) bài toán trên có lời giải $\overline{x}^{(k)}$.

Hai khả năng có thể xảy ra :

1. $\langle \bigtriangledown f(x^{(k)}), \overline{x}^{(k)} - x^{(k)}\rangle \geq 0$

   Khi đó $\Rightarrow \forall x \in D : \langle \bigtriangledown f(x^{(k)}), x - x^{(k)}\rangle \geq 0$, theo định lý 2.2.28 $x^{(k)}$ là lời giải cần tìm. Quá trình dừng lại.

2. $\langle \bigtriangledown f(x^{(k)}), \overline{x}^{(k)} - x^{(k)}\rangle < 0$
   Khi ấy đạo hàm riêng theo hướng $\overline{x}^{(k)} - x^{(k)}$ âm nên $f(x)$ giảm nếu ta đi từ $x^{(k)}$ tới $\overline{x}^{(k)}$. Ta chọn điểm $x^{(k+1)}$ đạt cực tiểu của $f(x)$ trên đọan $[x^{(k)}, \overline{x}^{(k)}]$ tức là giải bài toán một biến số λ :

   $$\min_{0\leq\lambda\leq 1}\{f(x^{(k)} + \lambda(\overline{x}^{(k)} - x^{(k)})\} = \min \phi(\lambda)$$

   (Cho $\phi'(\lambda) = 0$, tìm điểm dừng λ^*, so sánh các giá trị $\phi(\lambda^*), \phi(0), \phi(1)$))

Giả sử λ_k là giá trị đã tìm được ta lấy

$$x^{(k+1)} = x^{(k)} + \lambda_k(\overline{x}^{(k)} - x^{(k)}).$$

Khi đó theo Định lý 2.2.28 ta có:

1. $f(x^{(k)}$ giảm dần tới $\min f(x)$

2. Với mọi k thì $f(x^{(k)}) - \min f(x) \leq \langle \bigtriangledown f(x^{(k)}), \overline{x}^{(k)} - x^{(k)}\rangle$.

Kết luận:

1. Thuật toán trên đảm bảo sẽ hội tụ về nghiệm.

2. Nếu dừng ở bước thứ k thì được phương án $x^{(k)}$ xấp xỉ phương án tối ưu với sai số không vượt quá

   $$\langle \bigtriangledown f(x^{(k)}), x^{(k)} - \overline{x}^{(k)}\rangle.$$

Chứng minh. Thuật toán được chứng minh dựa trên hai bước sau:

1. Gọi E là tập hợp các đỉnh D có mặt trong dãy $\{\overline{x}^{(k)}\}$. Dãy $x^{(0)}, \ldots, x^{(k)}, \ldots$ nằm trong bao lồi của $\{x^{(0)}\} \cup E$.

- Tập $\{x^{(0)}\} \cup E$ gồm một số hữu hạn điểm nên là tập compact. Do đó dãy $x^{(0)}, \ldots, x^{(k)}, \ldots$ có một điểm tụ x^* và $f(x^{(0)}) > f(x^{(1)}) > \cdots > f(x^{(k)}) \ldots$

- Hàm $f(x)$ liên tục nên $\lim f(x^{(k)}) = f(x^*) = \mu$.

Ta phải chứng minh x^* là điểm cực tiểu.

- Tập hợp các đỉnh của D là hữu hạn nên trong dãy $\{x^{(k)}\}$ có những đỉnh được lặp lại nhiều lần.

- Có thể có dãy con của dãy $\{x^{(k)}\}$ hội tụ tới x^*. Vì vậy trong dãy vô hạn $\{x^{(k)}\}$ phải tìm được dãy con $\{x^{(kr)}\}$ sao cho

$$\lim_{r\to\infty} x^{(kr)} = x^* \text{ và những đỉnh } x^{(kr)} \text{ đều trùng nhau.}$$

Gọi đỉnh trùng nhau đó là $\overline{x}$, với λ cố định $(0 < \lambda < 1)$ ta có

$$f(x^{(kr)} + \lambda(\overline{x} - x^*)) \geq f(x^{(kr)}) \forall \ \ \lambda \in (0, 1),$$

nên

$$f(\overline{x} + \lambda(\overline{x} - x^*)) \geq f(x^*) \ \ \forall \lambda \subset (0, 1).$$

Do đó

$$\lim_{\lambda\to 0} f(x^* + \lambda(\overline{x} - x^*)) = \langle \nabla f(x^*), \overline{x} - x^* \rangle.$$

Mặt khác theo cách xây dựng $x^{(kr)}$ (là nghiêm của bài toán min) ta có :

$$\forall x \in D \ \ \langle \bigtriangledown f(x^{(kr)}), x - x^{(kr)} \rangle \geq \langle \bigtriangledown f(x), \overline{x} - x^{(kr)} \rangle \geq 0.$$

Vậy theo Định lý 2.2.28 x^* là điểm cực tiểu của $f(x)$ trong D.
2. Đặt $x = x^{(k)} + (x - x^{(k)})$, khi đó

$$\forall x \in D f(x) - f(x^{(k)}) = f(x^{(k)} + (x - x^{(k)}) - f(x^{(k)}) \geq \langle \bigtriangledown f(x^{(k)}), \overline{x}^{(k)} - x^{(k)} \rangle.$$

Do đó với $x = \overline{x}$ sao cho

$$f(x) - f(x^{(k)}) \geq \langle \bigtriangledown f(x^{(k)}), \overline{x}^{(k)} - x^{(k)} \rangle$$

thì suy ra

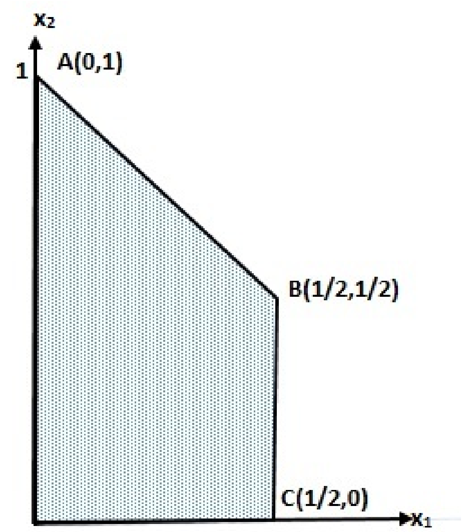
$$f(x^{(k)}) - \mu \leq \langle \bigtriangledown f(x^{(k)}), x^{(k)} - \overline{x}^{(k)} \rangle.$$

Ví dụ 2.3.24. *Giải QHL với ràng buộc tuyến tính*

$$f(x) = 4x_1^2 + 5x_2^2 + 6x_1x_2 + 25x_1 - 40x_2 \to \min.$$

Với miền D giới hạn bởi:

$$D = \begin{cases} f_1(x) = x_1 + x_2 - 1 \leq 0 \\ f_2(x) = x_1 - 1/2 \leq 0 \\ x_1, x_2 \geq 0. \end{cases}$$



Hình 2.12: **Miền D**

Thực hiện thuật toán.

B1. Lấy $x^{(0)} = (0,0)^T$,

$$\frac{\partial f}{\partial x_1} = 8x_1 + 6x_2 + 25$$
$$\frac{\partial f}{\partial x_2} = 10x_2 + 6x_1 - 40,$$

$$\nabla f(x^{(0)}) = (25, -40)^T, \langle \nabla f(x^{(0)}), x - x^{(0)} \rangle = 25x_1 - 40x_2.$$

Giải bài toán QHTT phụ:

$$\Phi(x) = 25x_1 - 40x_2 \to \min,\ x \in D.$$

Ta có

$$A(0,1), B(1/2,1/2), C(1/2,0),$$

$$\Phi(0) = 0, \phi(A) = -80, \Phi(C) = 25,$$

vậy $\overline{x}^{(0)} = (0,1)^T$; và $\langle \nabla f(x^{(0)}), \overline{x}^{(0)} - x^{(0)} \rangle = -80 < 0$ nên $x^{(0)}$ chưa phải là phương án tối ưu của quy hoạch lồi. Ta lại có

$$x^{(0)} + \lambda(\overline{x}^{(0)} - x^{(0)}) = (0,0) + \lambda(0,1) = (0,\lambda)^T, 0 \leq \lambda \leq 1$$

$$f(x^{(0)} + \lambda(\overline{x}^{(0)} - x^{(0)})) = 5\lambda^2 - 40\lambda := \phi(\lambda)$$

nên $\phi'(\lambda) = 10\lambda - 40 = 0$ khi $\lambda = 4$ vượt quá giới hạn cho phép nên ta lấy $\lambda = 1$. Suy ra chọn $x^{(1)} = (0,1)^T$.
B2. Ta có

$$\nabla f(x^{(1)}) = (31, -30), \langle \nabla f(x^{(1)}), (x_1, x_2 - 1) \rangle = 31x_1 - 30x_2 + 30.$$

- Giải bài toán QHTT phụ:

$$\Phi(x) = 31x_1 - 30x_2 \to \min,\ x \in D.$$

Ta có $\Phi(0) = 0, \Phi(A) = -30, \Phi(B) = 60, \Phi(C) = 1$ nên $\overline{x}^{(1)} = (0,1)^T$. Vì

$$\langle \nabla f(x^{(1)}), \overline{x}^{(1)} - x^{(1)} \rangle = 0$$

nên là phương án tối ưu của quy hoạch lồi ràng buộc tuyến tính.

Bài toán thực tế. Xét bài toán phân phối tối ưu công suất giữa thủy điện và nhiệt điện. Cho trước biểu đồ phụ tải trong một ngày đêm (24 giờ) tức là công suất phụ tải $P_{pt}(k), k = 1, \ldots, 24$, tính bằng MW. Giả sử năng lượng thủy điện có thể khai thác trong một ngày đêm là A(MWh). Hãy xác định công suất của các nhà máy nhiệt điện P_k sao cho đường biểu diễn công suất là bằng phẳng nhất có thể được và sao cho sử dụng hết năng lực của thủy điện.

Từ yêu cầu ta có thể lập mô hình như sau: xác định các công suất P_k sao cho:

$$\sum_{k=1}^{24} \left(P_k - \frac{\sum_{k=1}^{24} P_k}{24} \right)^2 \to \min,$$

điều kiện

$$\sum_{k=1}^{24} (P_{pt}(k) - P_k) = A$$

và

$$P_{\min} \leq P_k \leq P_{\max}, k = 1, \ldots, 24.$$

Biểu thức hàm mục tiêu có thể giải thích là tổng các bình phương có độ lệch giữa công suất nhiệt điện cần tìm tại mỗi giờ và công suất trung bình. Ràng buộc thứ nhất thể hiện sử dụng hết năng lượng thủy điện. Ràng buộc thứ 2 là hạn chế các nhà máy nhiệt điện. Bài toán tổng quát có dạng như trên, để đơn giản ta xét biểu đồ phụ tải cho 3 giờ với năng lượng thủy điện là $A = 3MW$. Từ yêu cầu phụ tải ta có: $P_{pt}(1) = 8MW, P_{pt}(2) = 6MW, P_{pt}(3) = 9MW$. Lấy cận trên của P_k là $P_{pt}(k)$, cận dưới là 0.

Ta có bài toán sau:

$$\sum_{k=1}^{3} \left(P_k - \frac{\sum_{k=1}^{3} P_k}{3} \right)^2 \to \min,$$

điều kiện

$$(8 - P_1) + (6 - P_2) + (9 - P_3) = 3$$

và

$$0 \leq P_1 \leq 8, 0 \leq P_2 \leq 6, 0 \leq P_3 \leq 9.$$

Từ ràng buộc thứ nhất ta suy ra :

$$P_1 + P_2 + P_3 = 20.$$

Do đó $P_3 = 20 - P_1 - P2$. Vậy bài toán có thể biến đổi về dang:

$$2P_1^2 + 2P_2^2 + 2P_1P_2 - 40P_1 - 40P_2 + 800/3 \to \min$$

điều kiện

$$0 \leq P_1 \leq 8, 0 \leq P_2 \leq 6, P_1 + P_2 \geq 11.$$

Miền ràng buộc là tam giác ABC trong đó $A(5,6), C(8,3)$

1. Chọn phương án xuất phát $P^{(0)} = (5,6)^T$

2. Tính Gradien của các hàm mục tiêu:

$$\begin{aligned} \frac{\partial f}{\partial P_1} &= -40 + 4P_1 + 2P_2 \\ \frac{\partial f}{\partial P_2} &= -40 + 2P_1 + 4P_2. \end{aligned}$$

Do đó

$$\bigtriangledown f(P^{(0)} = (-8,-6)^T$$

3. Ta có bài toán QHTT phụ thứ nhất :

$$\langle \bigtriangledown f(P^{(0)}), P - P^{(0)} \rangle = \langle -8, -6), (P_1 - 5, P_2 - 6\rangle = -8P_1 - 6P_2 + 76 \to \min$$

Phương án tối ưu của bài toán này là $\overline{P}^{(0)} = (8,6)^T$ Vì

$$\langle \bigtriangledown f(P^{(0)}, \overline{P}^{(0)} - P^0 = -24 < 0$$

nên $\overline{P}^{(0)}$ chưa là phương án tối ưu của bài toán xuất phát thao hướng từ $P^{(0)}$ đến $\overline{P}^{(0)}$ (tức là từ A đến B) hàm mục tiêu giảm. Lại có

$$P^{(0)} + \lambda(\overline{P}^{(0)} - P^{(0)}) = (5 + 3\lambda, 6)^T$$

$$\phi(\lambda) = 2(5+3\lambda)^2 + 2*6^2 + 12(5+3\lambda) - 40(5+3\lambda) - 40*6 + 800/3$$

tức là

$$\phi(\lambda) = 50 + 60\lambda + 18\lambda^2 + 72 + 60 + 36\lambda - 200 - 120\lambda) - 240 + 800/3$$

$$\phi(\lambda) = 18\lambda^2 - 24\lambda - 141 - 1/3$$

$\phi'(2/3) = 0$ nên $\lambda_0 = 2/3$.
4. Chọn

$$P^{(1)} = P^{(0)} + 2/3(\overline{P}^{(0)} - P^{(0)}) = (7,6)^T \text{ suy ra } \bigtriangledown f(P^{(1)}) = (0,-2)^T.$$

5. Ta có bài toán QHTT phụ thứ 2:

$$\langle \bigtriangledown f(P^{(1)}), P - P^{(1)} \rangle = -2P_2 - 4 \to \min$$

Suy ra $\overline{P}^{(1)} = (8,6) = B$

$$\langle \bigtriangledown f(P^{(1)}), \overline{P}^{(1)} - P^{(1)} \rangle = \langle (0,-2)^T, (1,0)^T \rangle = 0$$

Vậy $\overline{P}^{(1)}$ là phương án tối ưu của bài toán xuất phát.

$$P_1 = 7, P_2 = 6, P_3 = 7.$$

2.3.2 Bài toán quy hoạch lồi với ràng buộc phi tuyến (lồi).

Xét bài toán quy hoạch lồi tổng quát

$$f(x) \to \min, x \in D \quad (2.42)$$
$$D = \{x \in A, f_i(x) \leq 0, i = 1, \ldots, m,\} \quad (2.43)$$

trong đó f, f_i là các hàm lồi. Vì bài toán quy hoạch lồi là một trường hợp riêng cả quy hoạch phi tuyến nên các phương pháp đã xét cho quy hoạch lồi tối ưu cục bộ cũng là tối ưu toàn cục.

Các giả thiết của bài toán:

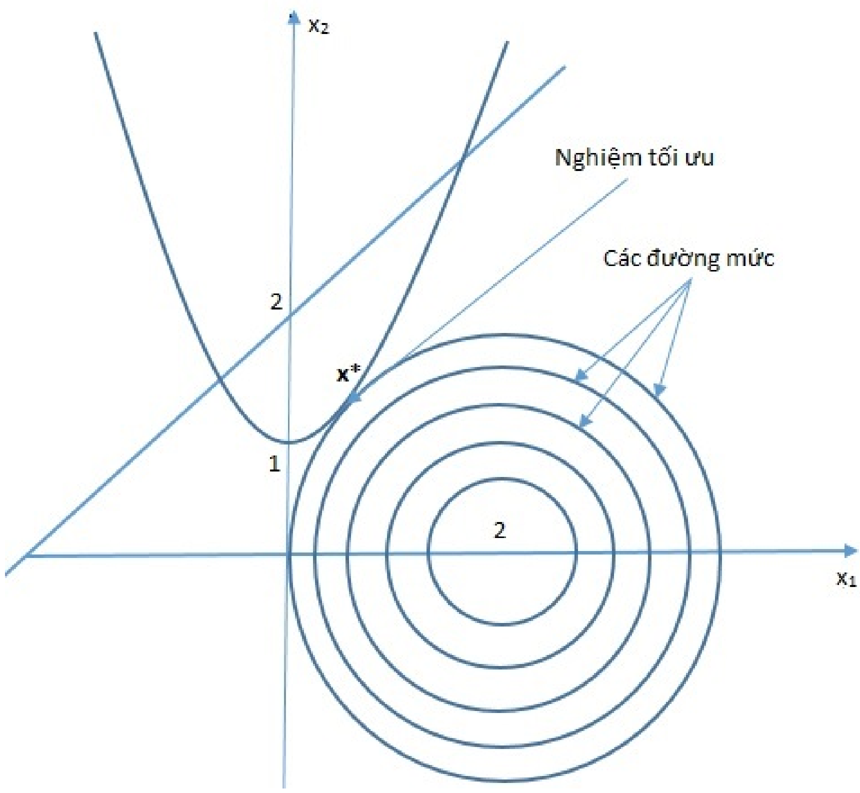
1. Các hàm f, f_i đều khả vi liên tục.
2. Tồn tại một phương án $x^{(0)} \in D$ sao cho $f_i(x^{(0)}) < 0$ với mọi ràng buộc f_i phi tuyến (giả thiết chính quy).
3. Tập D giới nội. Khi ấy tồn tại đa diện $Q : D \subset Q$.

Trước tiên ta xây dựng hình ảnh bài toán qua ví dụ sau:

$$f(x) = x_1^2 + x_2^2 - 4x_1 + 4 \to \min$$

Với các ràng buộc sau:

$$\begin{aligned} f_1(x_1, x_2) &= x_1 - x_2 + 2 \geq 0 \\ f_2(x_1, x_2) &= x_1^2 - x_2 + 1 \geq 0 \\ f_3(x_1, x_2) &= x_1 \geq 0 \\ f_4(x_1, x_2) &= x_2 \leq 0 \end{aligned}$$



Hình 2.13: **Điểm cực tiểu là điểm tiếp xúc giữa đường mức và miền**

Hàm f có thể biểu diễn dưới dạng $f(x) = (x_1 - 2)^2 + x_2^2$. Tại điểm $(2, 0)$ hàm đạt giá trị nhỏ nhất trong $I\!R^2$ nhưng trong miền D hàm đạt giá trị nhỏ nhất tại (x_1^*, x_2^*) sao cho đường mức của hàm $f = 4$. Ví dụ trên ccho ta thấy phương án tối ưu nằm ở điểm tiếp xúc của dường mức và miền ràng buộc của bài toán.

Mục tiêu: Xây dựng dãy phương án $x^{(1)}, x^{(2)}, \ldots$ sao $f(x^{(k)}) \to \mu = \min_{x \in D} f(x)$.

Nội dung phương pháp Gradient:

B1. Xây dựng một phương án xuất phát $x^{(0)}$.

B2. Giả thử đã có phương án $x^{(k)}$ theo một quy tắc (R) nào đó, ta chọn hướng $z^{(k)}$ chấp nhận được tại $x^{(k)}$ sao cho

$$\langle \bigtriangledown f(x^{(k)}), z^{(k)} \rangle < 0. \tag{2.44}$$

- Nếu không có hướng nào như thế, thì $\langle \bigtriangledown f(x^{(k)}), z \rangle \geq 0$ với mọi hướng z chấp nhận được tại $x^{(k)}$ thì $x^{(k)}$ là phương án tối ưu.
- Bất đẳng thức (2.44) có nghĩa là khi đi từ $x^{(k)}$ theo hướng $z^{(k)}$ thì một trong lân cận nào đó của x giá trị $f(x)$ giảm dần. ta tìm điểm $x^{(k+1)}$ "thấp" nhất của D theo B3

B3. Xây dựng $x^{(k+1)} = x^{(k)} + \lambda_k z^{(k)}$ với

$$f(x^{(k+1)}) = \min\{f(x^{(k)} + \lambda z^{(k)}) \mid \lambda > 0, x^{(k)} + \lambda z^{(k)} \in D\}.$$

Rõ ràng $f(x^{(k+1)}) < f(x^{(k)})$.

Phương pháp xác định quy tắc (R) (thực chất là xây dựng hướng chấp nhận được để $f(x^{(k)}) \to \mu = \min_{x \in D} f(x)$)

Giả sử ta đã có một phương án $x^{(0)}$ sao cho điều kiện chính quy của $f_i(x^{(0)})$ thỏa mãn. Với mỗi $a \in D$ đặt

$$P(a) := \{x \mid \langle \bigtriangledown f_i(a), x - a \rangle \leq -f_i(a), i = 1, \ldots, m\}.$$

Dễ thấy từ định nghĩa:

1. $D \subset P(a)$
2. Nếu f_i là afin thì bất đẳng thức $\langle \bigtriangledown f_i(a), x - a \rangle \leq -f_i(a)$ trùng với $f_i(x) \leq 0$.

Phương pháp gồm các bước sau:

Nếu $a \in D$ và $x \in P(a)$ thì $\forall \epsilon > 0$ hướng chấp nhận được tại a là:

$$z := (x - a) + \epsilon(x^{(0)} - a) \in D.$$

Phương án được xây dựng qua các bước sau:
B_k. Giả sử bước k (k = 0,1,2,3,4...) ta có $x^{(k)} \in D$
B_{k+1}. Giải QHTT phụ:

$$\min \langle \bigtriangledown f(x^{(k)}), x - x^{(k)} \rangle \mid x \in P(x^{(k)}) \cap Q,$$

trong đó Q là đa diện lồi chứa D. Chính vì Q là đa diện lồi nên quy hoạch này bao giờ cũng có lời giải $\overline{x}^{(k)}$.
Đặt $\sigma_k := \langle \bigtriangledown f(x^{(k)}), \overline{x}^{(k)} - x^{(k)} \rangle$ khi đó có hai khả năng:

1. $\forall x \in [P(x^{(k)}) \cap Q]$ $f(x) - f(x^{(k)}) \geq \langle \bigtriangledown f(x^{(k)}), x - x^{(k)} \rangle \geq \sigma_k \geq 0.$ Trong trường hợp này $x^{(k)}$ là phương án tối ưu.

2. $\sigma_k < 0$. Khi đó hướng

$$z^{(k)} := (\overline{x}^{(k)} - x^{(k)}) + \epsilon_k(x^{(0)} - x^{(k)})$$

là chấp nhận tại $x^{(k)}$, trong đó

$$\epsilon_k := \min\{1, -\sigma_k/(2|\delta_k|), \delta_k := \langle \bigtriangledown f(x^{(k)}), x^{(0)} - x^{(k)} \rangle.$$

Chọn $x^{(k+1)} := x^{(k)} + \lambda_k z^{(k)}$, trong đó λ_k chọn lớn nhất sao cho:

$$f(x^{(k+1)}) = \min\{f(x^{(k)} + \lambda z^{(k)}) \mid \lambda > 0, x^{(k)} + \lambda z^{(k)} \in D, \}$$

tức là nếu đặt $\phi(\lambda) := f(x^{(k)} + \lambda z^{(k)})$ thì

$$\lambda_k = \max\{\lambda > 0 \mid x^{(k)} - \lambda z^{(k)} \in D, \phi'(\lambda) \leq 0\}.$$

Chuyển qua bước $(k+1)$ sau khi đã xác định được $x^{(k+1)}$.

Tóm lại ta có định lý sau:

Định lý 2.3.30. *Với các giả thiết đã cho của bài toán (2.42) thì*

1. $f(x^{(k)}) \to \mu = \min_{x\in D} f(x)$.
2. $f(x^{(k)}) - \mu = \langle \triangledown f(x^{(k)}), x^{(k)} - \overline{x}^{(k)} \rangle$.

Ví dụ minh họa thuật toán

$$\begin{aligned} f(x) &= 4x_1^2 + 5x_2^2 + 6x_1x_2 + 25x_1 - 40x_2 \to \min \\ f_1(x) &= x_1 + x_2 - 1 \leq 0, \\ f_2(x) &= 8x_1^2 + x_2^2 - 2 \leq 0 \\ & x_1, x_2 \geq 0. \end{aligned}$$

Ta có

$$\begin{aligned} \frac{\partial f(x)}{\partial x_1} &= 8x_1 + 6x_2 + 25 \\ \frac{\partial f(x)}{\partial x_2} &= 6x_1 + 10x_2 - 40. \end{aligned}$$

Ta chọn miền Q như sau:

$$\begin{cases} x_1 + x_2 - 1 \leq 0 \\ x_1 - 1/2 \leq 0 \\ x_1, x_2 \geq 0. \end{cases}$$

$$\frac{\partial f_1(x)}{\partial x_1} = 1, \frac{\partial f_1(x)}{\partial x_2} = 1, \frac{\partial f_2(x)}{\partial x_1} = 16x_1, \frac{\partial f_2(x)}{\partial x_2} = 2x_2$$

Lần lặp 1:

B1: Chọn phương án xuất phát $x^{(0)} = (0,0)^T$ thỏa mãn điều kiện chính quy theo f_1, f_2.

B2: Xây dựng

$$P(x^{(0)}) := \{x \mid \langle \triangledown f_i(x^{(0)}), x - x^{(0)} \rangle \leq -f_i(x^{(0)})\}.$$

Ta có

$$\triangledown f_1(x^{(0)}) = (1,1)^T, \triangledown f_2(x^{(0)}) = (0,0)^T,$$

$$\langle \bigtriangledown f_1(x^{(0)}), (x_1, x_2)\rangle = x_1 + x_2 \leq -f_1(x^{(0)}) = 1,$$

tức là $x_1 + x_2 - 1 \leq 0$.

$$\langle \bigtriangledown f_2(x^{(0)}), (x_1, x_2)\rangle = 0 \leq -f_2(x^{(0)}) = 2.$$

Do đó

$$P(x^{(0)}) = \{x \mid x_1 + x_2 - 1 \leq 0 \text{ và } P(x^{(0)}) \cap Q = Q.$$

B3: Giải bài toán phụ:

$$\langle \bigtriangledown f(x^{(0)}), x - x^{(0)})\rangle = 25x_1 - 40x_2 \to \min,$$

$$\begin{cases} x_1 + x_2 \leq 1 \\ x_1 - 1/2 \leq 0 \\ x_1, x_2 \geq 0. \end{cases}$$

Phương án tối ưu của quy hoạch tuyến tính phụ là $\overline{x}^0 = (0,1)^T$. Mặt khác

$$\langle \bigtriangledown f(x^{(0)}), \overline{x}^{(0)} - x^{(0)})\rangle = -40$$

nên $x^{(0)}$ chưa phải là phương án tối ưu.

B4: Xác định hướng chấp nhận được:

$$\begin{aligned} z^{(0)} &= (\overline{x}^{(0)} - x^{(0)}) + \epsilon_0(x^{(0)} - x^{(0)}) = (\overline{x}^{(0)} - x^{(0)}) = (0,1)^T \\ \phi(\lambda) &= f(x^{(0)} + \lambda z^{(0)}) = 5\lambda^2 - 40\lambda \\ \phi'(\lambda) &= 20\lambda - 80 \leq 0 \end{aligned}$$

suy ra $\lambda \leq 4$.

B5: Tìm

$$\lambda_0 = \max\{\lambda > 0 \mid x^{(0)} + \lambda z^0 \in D, \phi(\lambda_0) \leq 0\}.$$

Vì $x^{(0)} + \lambda z^0 = (0,0)^T + \lambda(0,1)^T = (0,\lambda)^T$ và $x_1 + x_2 - 1 \leq 0$ nên $\lambda \leq 1$. Do đó suy ra $\lambda_0 = 1$.

6. Đặt $x^{(1)} := x^{(0)} + \lambda_0 z^{(0)} = (0,1)^T$.

Lần lặp 2:

B1: Tìm

$$P(x^{(1)}) := \{x \mid \langle \triangledown f_i(x^{(1)}), x - x^{(1)}\rangle \leq -f_i(x^{(1)})\}.$$

Ta có

$$\langle \triangledown f_1(x^{(1)}), (x_1, x_2 - 1)\rangle = x_1 + x_2 - 1 \leq -f_1(x^{(1)}) = 0,$$

tức là $x_1 + x_2 - 1 \leq 0$.

$$\langle \triangledown f_2(x^{(1)}), (x_1, x_2 - 1)\rangle = 2x_2 - 2 \leq -f_2(x^{(1)}) = 1,$$

tức là $2x_2 - 3 \leq 0$ hay $x_2 \leq 3/2$. Do đó

$$P \cap Q = \{x \mid \left\{\begin{array}{l} x_1 + x_2 \leq 1 \\ x_1 \leq 1/2 \\ x_2 \leq 3/2 \\ x_1, x_2 \geq 0. \end{array}\right\}.$$

Từ trên ta suy ra $P \cap Q = Q$.

B2: Giải bài toán phụ:

$$\langle \triangledown f(x^{(1)}), x - x^{(1)})\rangle = \langle (31, -30), (x_1, x_2 - 1)\rangle = 31x_1 - 30x_2 + 30 \rightarrow \min,$$

$$\left\{\begin{array}{l} x_1 + x_2 - 1 \leq 0 \\ x_1 - 1/2 \leq 0 \\ x_1, x_2 \geq 0. \end{array}\right.$$

Phương án tối ưu của quy hoạch tuyến tính phụ là $\overline{x}^{(1)} = (0,1)'$. Mặt khác

$$\langle \triangledown f(x^{(1)}), \overline{x}^1 - x^{(1)})\rangle = 0$$

nên $x^{(1)}$ là phương án tối ưu của bài toán xuất phát và $f_{\min} = -35$.

2.3.3 Phương pháp xấp xỉ tuyến tính

Xét bài toán quy hoạch lồi tổng quát

$$f(x) = \langle c, x\rangle \to \min, x \in D \tag{2.45}$$

$$D = \{x \in I\!R^n \mid f_i(x) \leq 0, i = 1, \dots, m, \} \tag{2.46}$$

trong đó f_i là các hàm lồi khả vi trên $I\!R^n$ và D là lồi, compact. Thuật toán tuyến tính hóa ràng buộc giải bài toán (2.45)–(2.46 do Kelley đề xuất năm 1960 (còn gọi là phương pháp siêu phẳng cắt"). Nội dung phương pháp như sau:

Ý tưởng thuật toán.
Ở bước lặp $k = 1, 2, \dots$ thì

- Tập compact D được thay bằng đa diện lồi $D_1 \supseteq D_2 \supseteq \cdots \supseteq D$.
- Bài toán (2.45)–(2.46) được thay bằng dãy bài toán quy hoạch tuyến tính

$$\min\langle c, x\rangle : \ x \in D_k. \tag{2.47}$$

$\Rightarrow$ Dãy nghiệm tối ưu $x^{(k)}$ của bài toán (2.47) hội tụ đến nghiệm tối ưu của bài toán (2.45)-(2.46).

Thuật toán tuyến tính hóa ràng buộc
B1: (Bước xây dựng đa diện lồi ban đầu $D_1 \supseteq D$). Chọn tùy ý p điểm $x^{(1)}, x^{(2)}, \dots, x^{(p)}$ và dựng tại mỗi điểm đã chọn siêu phẳng tiếp xúc với mặt cong $y = f_i(x)$, tức là lập các hàm tuyến tính

$$h_{i,k}(x) = f_i(x^{(k)}) + \langle \bigtriangledown f_i(x^{(k)}), x - x^{(k)})\rangle, k = 1, \dots, p; i = 1, \dots, m.$$

Do hàm $f_i(x)$ lồi nên

$$h_{i,k}(x) = f_i(x^{(k)}) + \langle \bigtriangledown f_i(x^{(k)}), x - x^{(k)})\rangle + f_i(x^{(k)}) \leq f_i(x) \tag{2.48}$$

$$\forall x \in I\!R^n, k = 1, \dots, p; i = 1, \dots, m. \tag{2.49}$$

Đặt

$$D_1 = \{x \in I\!R^n : h_{k,i} \leq 0, k = 1, \dots, p; i = 1, \dots, m.\}$$

Với mỗi $x \in D$ thì $f_i(x) \leq 0$ với mọi i, do đó theo (2.49) thì $h_{k,i} \leq 0$ nên $D_1 \supseteq D$.

B2: (Bước lặp k=1,2,...)

(a) Ở bước lặp k, tìm điểm $x^{(k+p)} \in D_k$ nghiệm tối ưu của bài toán quy hoạch tuyến tính

$$\langle c, x^{(k+p)} \rangle = \min\{\langle c, x \rangle : x \in D_k\}$$

(b) Nếu $x^{(p+k)} \in D$ thì $x^{(p+k)}$ là nghiệm tối ưu của bài toán (2.45)–(2.46) và dừng thuật toán, vì $D \subset D_k$ và

$$\begin{aligned} \langle c, x^{(p+k)} \rangle = \min\{\langle c, x \rangle : x \in D_k\} &\leq \min\{\langle c, x \rangle : x \in D\} \\ &\leq \min\{\langle c, x^{(p+k)} \rangle\}. \end{aligned}$$

(c) Nếu $x^{(p+k)} \notin D$ thì $x^{(p+k)}$ vi phạm ít nhất một ràng buộc $f_i(x) \leq 0$ nào đó. Ký hiệu

$$I_k := \{i \in \{1, 2, \ldots, m\} : f_i(x^{(p+k)}) > 0\}.$$

Đặt

$$D_{k+1} := D_k \cap \{x \in I\!R^n : \langle \nabla f_i(x^{(p+k)}, x - x^{(p+k)}) \rangle + f_i(x^{(p+k)}) \leq 0, i \in I_k\}. \tag{2.50}$$

(d) Đặt $k \leftarrow k + 1$, quay lại bước k.

Nhận xét.

1. Bài toán quy hoạch lồi tổng quát có thể đưa về bài toán với hàm mục tuyến tính nhờ thêm biến mới x_{n+1} và xét bài toán

$$\min\{x_{n+1} \mid f_i(x) \leq 0, i = 1, \ldots, m, f_{m+1} := f(x) - x_{n+1} \leq 0\}.$$

Rõ ràng bài toán này và bài toán quy hoạch lồi là tương đương nhau.

2. Để xây dựng đa diện ban đầu D_1 ta nên chọn p điểm ít nhất sao cho $D_1 \supseteq D$. Thường chọn $p = n+1$ và $x^{(1)}, \ldots, x^{(p)}$ là các đỉnh của một đơn hình n-chiều, như trong ví dụ dưới đây.

3. Thay cho (2.50) ta có thể đặt

$$D_{k+1} := D_k \cap \{x \in I\!R^n : \langle \bigtriangledown f_i^*(x^{(p+k)}), x - x^{(k)}) \rangle + f_i^*(x^{(p+k)}) \leq 0, i \in I_k\},$$

trong đó $f_i^*(x^{(p+k)}) = \max_{1 \leq i \leq m}\{f_i(x^{(p+k)})\}$, nghĩa là đưa vào D_k bất đẳng thức tương ứng với ràng buộc bị vi phạm nhiều nhất.

4. Bài toán (2.45) ở bước thứ $k+1$ nhận được từ bước thứ k bằng cách thêm một hay một số ràng buộc bất đẳng thức tuyến tính, vì thế giải bài toán ở bước thứ $k+1$ ta có thể dùng thuật toán đơn hình đối ngẫu, xuất phát từ nghiệm tối ưu đã có $x^{(p+k)}$.

5. Các bất đẳng thức thêm vào D_k sẽ cắt bỏ một phần chứa điểm $x^{(p+k)}$ của D_k, bởi vì

$$\langle \bigtriangledown f_i(x^{(p+k)}), x^{(p+k)} - x^{(p+k)} \rangle + f_i(x^{(p+k)}) = f_i(x^{(p+k)}) > 0 \ \forall i \in I_k.$$

Vì thế, các siêu phẳng

$$h_{ki}(x) = \{x \in I\!R^n \mid \langle \bigtriangledown f_i(x^{(p+k)}), x - x^{(p+k)} \rangle + f_i(x^{(p+k)}) = 0\}, i \in I_k.$$

được gọi là siêu phẳng cắt và chúng tách tập lồi D_{k+1} với điểm $x^{(p+k)} \notin D_{k+1}$

6. Ở mọi bước k ta luôn có $\bigtriangledown f_i(x^{(p+k)}) \neq 0$ với mọi $i \in I_k$, bởi vì nếu tồn tại $i \in I_k$ sao cho $\bigtriangledown f_i(x^{(p+k)}) = 0$ thì do f_i là hàm lồi nên với i đó ta có

$$\begin{aligned} f_i(x) &\geq f_i(x^{(p+k)} + \langle \bigtriangledown f_i(x^{(p+k)}), x^{(p+k)} - x^{(p+k)} \rangle + f_i(x^{(p+k)}) \\ &= f_i(x^{(p+k)}) \ \forall x \in I\!R^n. \end{aligned}$$

Điều này mâu thuẫn với giả thiết tập $D \neq \emptyset$.

Kết luận: Với các giả thiết của bài toán (2.45) và D là tập lồi, compact thì dãy $x^{(k)}$ nhận được trong thuật toán trên có một dãy con hội tụ tới nghiệm tối ưu của bài toán, tức là

$$\lim\langle c, x^{(p+k)}\rangle = \min\{\langle c, x\rangle : x \in D\}.$$

Ví dụ minh họa thuật toán. Tìm cực tiểu của hàm

$$f(x) = x_1 + 2x_2,$$

với các điều kiện $x \in D = \{x \in I\!R^2 \mid f_1(x) := (x_1-3)^2 + (x_2-2)^2 - 9 \leq 0,\ f_2(x) := -x_1 + (x_2-2)^2 + 1 \leq 0\}$. Các bước giải thực hiện như sau:

1. Tính

$$\bigtriangledown f_1(x) = \begin{pmatrix} 2x_1 - 6 \\ 2x_2 - 4 \end{pmatrix},\ \bigtriangledown f_2(x) = \begin{pmatrix} -1 \\ 2x_2 - 4 \end{pmatrix}$$

2. Xây dựng đa diện lồi ban đầu D_1. Chọn 3 điểm ($p = 3$): $x^{(1)} = (2,0)^T; x^{(2)} = (2,4)^T; x^{(3)} = (6,2)^T$.

- Với $x^{(1)}$:

$$\bigtriangledown f_1(x^{(1)}) = \begin{pmatrix} -2 \\ -4 \end{pmatrix},\ \bigtriangledown f_2(x^{(1)}) = \begin{pmatrix} -1 \\ -4 \end{pmatrix}, f_1(x^{(1)}) = -4; f_2(x^{(1)}) = 3.$$

$$h_{11} = \langle \bigtriangledown f_1(x^{(1)}), x - x^{(1)}\rangle + f_1(x^{(1)}) = -x_1 - 2x_2 \leq 0$$

$$h_{12} = \langle \bigtriangledown f_2(x^{(1)}), x - x^{1}\rangle + f_2(x^{(1)}) = -x_1 - 4x_2 + 5 \leq 0$$

- Với $x^{(2)}$:

$$\bigtriangledown f_1(x^{(2)}) = \begin{pmatrix} -2 \\ 4 \end{pmatrix},\ \bigtriangledown f_2(x^{(2)}) = \begin{pmatrix} -1 \\ 4 \end{pmatrix}, f_1(x^{(2)}) = -4; f_2(x^{(2)}) = 3.$$

$$h_{21} = \langle \bigtriangledown f_1(x^{(2)}), x - x^{(2)}\rangle + f_1(x^{(2)}) = -x_1 + 2x_2 - 8 \leq 0$$

$$h_{22} = \langle \bigtriangledown f_2(x^{(2)}), x - x^{(2)}\rangle + f_2(x^{(2)}) = -x_1 + 4x_2 - 11 \leq 0$$

- Với $x^{(3)}$:

$$\bigtriangledown f_1(x^{(3)}) = \begin{pmatrix} 6 \\ 0 \end{pmatrix},\ \bigtriangledown f_2(x^{(3)}) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, f_1(x^{(3)}) = 0; f_2(x^{(3)}) = -5.$$

$$h_{31} = \langle \bigtriangledown f_1(x^{(3)}), x - x^{(3)} \rangle + f_1(x^{(3)}) = -x_1 - 6 \leq 0$$

$$h_{32} = \langle \bigtriangledown f_2(x^{(3)}), x - x^{(3} \rangle + f_2(x^{(3)}) = -x_1 + 1 \leq 0$$

Đa diện

$$\begin{aligned} D_1 &= \{x \in I\!R^2 \mid h_{ki}(x) \leq 0\} \\ &= \{x \in I\!R^2 \mid x_1 + 4x_2 \geq 5, -x_1 + 4x_2 \leq 11, x_1 \leq 6, x_1 \geq 1\}. \end{aligned}$$

3. **Bước lặp 1.**

(a) Giải bài toán

$$f(x) = x_1 + 2x_2 \to \min,$$

với các ràng buộc

$$\begin{cases} x_1 + 4x_2 \geq 5 \\ -x_1 + 4x_2 \leq 11 \\ x_1 \leq 6 \\ x_1 \geq 1, \end{cases}$$

Nghiệm tối ưu của bài toán này là $x^{(4)} = (1, 1), f(x^{(4)}) = 3$.

(b) Do $f_1(x^{(4)}) = -4 < 0$ và $f_2(x^{(4)}) = 1 > 0$ nên $I_1 = \{2\}$ ta thực hiện bước (c)

(c) Ta có $\bigtriangledown f_1(x^{(4)}) = (-1, -2)^T, \langle \bigtriangledown f_2(x^{(4)}), x - x^{(4)} \rangle + f_2(x^{(4)}) = -x_1 - 2x_2 + 4$ và đặt

$$\begin{aligned} D_2 &= D_1 \cap \{x \in I\!R^2 \mid \langle \bigtriangledown f_2(x^{(4)}), x - x^{(4)} \rangle + f_2(x^{(4)}) \leq 0\} \\ &= D_1 \cap \{x \in I\!R^2 \mid x_1 + 2x_2 \geq 4\} \end{aligned}$$

Đặt $k \leftarrow k + 1 = 2$ chuyển qua bước lặp $k = 2$.

Bước lặp $k = 2$.

(a) Giải bài toán

$$f(x) = x_1 + 2x_2 \to \min, x \in D_2.$$

Nghiệm tối ưu của bài toán này là $x^{(5)} = (1, 1.5)^T, f(x^{(5)}) = 4$.

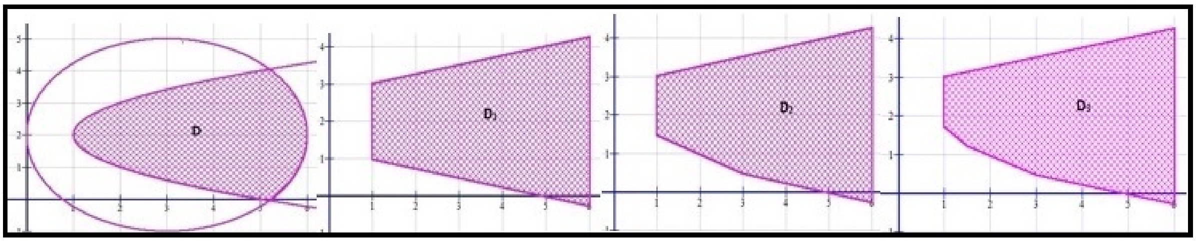
(b) Do $f_1(x^{(5)}) = -4.75 < 0$ và $f_2(x^{(5)}) = 0.25 > 0$ nên $I_1 = \{2\}$ ta thực hiện bước (c)

(c) Ta có $\bigtriangledown f_1(x^{(5)}) = (-1, -1)^T, \langle \bigtriangledown f_2(x^{(5)}), x - x^{(5)}\rangle + f_2(x^{(5)}) = -x_1 - x_2 + 2.75$ và đặt

$$\begin{aligned} D_3 &= D_2 \cap \{x \in I\!R^2 \mid \langle \bigtriangledown f_2(x^{(5)}), x - x^{(5)}\rangle + f_2(x^{(5)}) \leq 0\} \\ &= D_2 \cap \{x \in I\!R^2 \mid x_1 + x_2 \geq 2.75\}. \end{aligned}$$

Đặt $k \leftarrow k + 1 = 3$ chuyển qua bước lặp $k = 3$.

Bước lặp 3.



Hình 2.14: **Miền** D, D_1, D_2, D_3

(a) Giải bài toán

$$f(x) = x_1 + 2x_2 \to \min, x \in D_3.$$

Nghiệm tối ưu của bài toán này là $x^{(6)} = (1.5, 1.25)^T, f(x^{(6)}) = 4.$ và $x^{(6)} = (3, 0.5)^T, f(x^{(6)}) = 4.$

(b) Do $f_1(x^{(6)}) = -6.1875 < 0$ và $f_2(x^{(6)}) = 0.0625 > 0$ nên $I_1 = \{2\}$ ta thực hiện bước (c)

(c) Ta có $\bigtriangledown f_2(x^{(6)}) = (-1, -1.5)^T, \langle \bigtriangledown f_2(x^{(6)}), x - x^{(6)}\rangle + f_2(x^{(6)}) = -x_1 - 1.5x_2 + 1.9375$ và đặt

$$\begin{aligned} D_4 &= D_3 \cap \{x \in I\!R^2 \mid \langle \bigtriangledown f_2(x^{(6)}), x - x^{(5)}\rangle + f_2(x^{(6)}) \leq 0\} \\ &= D_3 \cap \{x \in I\!R^2 \mid x_1 + 1.5x_2 \geq 1.9375\} \end{aligned}$$

Tuy nhiên trong trường hợp này D_4 trùng với D_3 cho nên ta phải quay lại chọn nghiệm khác để cho $D_4 \subset D_3$, nghĩa là $x^{(6)} = (3, 0.5)^T$. Tính toán lại ta có:

(b) Do $f_1(x^{(6)}) = -6.75 < 0$ và $f_2(x^{(6)}) = 0.25 > 0$ nên $I_1 = \{2\}$ ta thực hiện bước (c)

(c) Ta có $\bigtriangledown f_2(x^{(6)}) = (-1, -3)^T$, $\langle \bigtriangledown f_2(x^{(6)}), x - x^{(6)} \rangle + f_2(x^{(6)}) = -x_1 - 3x_2 + 4.75 \leq 0$ và đặt

$$\begin{aligned} D_4 &= D_3 \cap \{x \in I\!R^2 \mid \langle \bigtriangledown f_2(x^{(6)}), x - x^{(5)} \rangle + f_2(x^{(6)}) \leq 0\} \\ &= D_3 \cap \{x \in I\!R^2 \mid x_1 + 3x_2 \geq 4.75\} \end{aligned}$$

Nghiệm tối ưu của bài toán trên D_4 là $x^{(7)} = (1.75, 1.125)^T$ và $f_{\min} = 4$. Đặt $k \leftarrow k + 1 = 5$ chuyển qua bước lặp $k = 5$.

Kết quả tiếp theo được thể hiện qua bảng tính sau:

Bước k	Xấp xỉ mới x^{3+k}	Tập I_k	$f(x^{3+k})$
1	$x^{(4)} = (1, 1)$	$\{1, 2\}$	3
2	$x^{(5)} = (1, 1.5)$	$\{2\}$	4
3	$x^{(6)} = (1.5, 1.25)$	$\{2\}$	4
4	$x^{(7)} = (1.75, 1.125)$	$\{2\}$	4
5	$x^{(8)} = (1.875, 1.0625)$	$\{2\}$	4
6	$x^{(9)} = (1.9375, 1.03125)$	$\{2\}$	4
7	$x^{(10)} = (1.96875, 1.015625)$	$\{2\}$	4
8	$x^{(11)} = (1.984375, 1.00390625)$	$\{2\}$	4
9	$x^{(12)} = (1.9921875, 1.00390625)$	$\{2\}$	4
10	$x^{(13)} = (1.99609375, 1.00193125)$	$\{2\}$	4

Bảng 2.3: *

Phương án tối ưu là $x^{(13)} \approx (2, 1)^T$ và giá trị hàm mục tiêu là $f_{\min} = 4$.

$$\min\{f(x) + p(x) : x \in I\!R^n\} \tag{2.51}$$

2.3.4 Quy hoạch toàn phương

Quy hoạch toàn phương là bài toán tìm cực tiểu hàm lồi bậc hai với các ràng buộc tuyến tính. Có ba dạng bài toán cơ bản :

1. Dạng chính tắc

$$f(x) = \langle Cx, x\rangle + 2\langle p, x\rangle + p_0 \to \min$$

   với các ràng buộc

$$\begin{cases} Ax = b \\ x \geq 0, \end{cases}$$

2. Dạng chuẩn

$$f(x) = \langle Cx, x\rangle + 2\langle p, x\rangle + p_0 \to \min$$

   với các ràng buộc

$$\begin{cases} Ax \leq b \\ x \geq 0, \end{cases}$$

3. Hàm mục tiêu giống như bài toán trên, còn các phương trình ràng buộc có dạng hỗn hợp như sau

$$\begin{aligned} \langle A_i, x\rangle &\geq b_i,\ i \in I \subseteq M := \{1, 2, \ldots, m.\} \\ \langle A_i, x\rangle &= b_i,\ i \in M \setminus I \\ x_j &\geq 0\ j \in J \subseteq \{1, 2, \ldots, n\}, \end{aligned}$$

   trong đó A_i là véc tơ hàng của A, b_i là thành phần thứ i của véc tơ b. Trong bài toán trên $x = (x_1, x_2, \ldots, x_n)^T$ là véc tơ cần tìm. $C \in \mathbb{R}^{n\times n}$ là ma trận đối xứng không âm, $p \in \mathbb{R}^n, b \in \mathbb{R}^m, p_0$ là hằng số. Ta luôn coi $b \geq 0$.

2.3.5 Phương pháp Hildreth – D'Esopo

Điều kiện để áp dụng phương pháp này là :

1. Hàm mục tiêu lồi chặt.
2. Ràng buộc có dạng bất đẳng thức.
3. C là ma trận vuông đối xứng xác định dương và không giả thiết $b \geq 0$.

Với các giả thiết trên bài toán sẽ có một nghiệm tối ưu duy nhất. Để tìm nghiệm này ta xét bài toán đối ngẫu, với các tiêu chuẩn tối ưu sau:

$$Ax + y = b, 2Cx + A^T u = -p,$$

$$u \geq 0, y \geq 0, u^T y = 0.$$

Vì ma trận C xác định dương nên có ma trận ngược C^{-1} và ta có:

$$x(u) = (-1/2) \times C^{-1}(A^T u + p).$$

Ta viết lại tiêu chuẩn tối ưu

$$2Gu - y = -h, u \geq 0, u^T y = 0.$$

Với các ký hiệu mới:

$$h = (1/2)AC^{-1}p + b, G = (1/4)AC^{-1}A^T.$$

Véc tơ h có m thành phần , ma trận G vuông cấp m, nửa xác định dương và các phần tử trên đường chéo luôn dương. Điều kiện tối ưu trên cũng là điều kiện tối ưu của bài toán sau:

$$\min(\phi(u) = h^T u + u^T G u : u \geq 0)$$

là bài toán đối ngẫu của bài toán ban đầu. Bài toán này cũng quy hoạch toàn phương nhưng có ràng buộc đơn giản. Giải bài toán này thì ta cũng có được lời giải của bài toán ban đầu. Bây giờ ta đi tìm u. Xuất phát từ một xấp xỉ ban đầu, tùy ý $u^{(0)} \geq 0$ (thường chọn $u^{(0)} = 0$). Các thành phần xấp xỉ theo u^1 được tìm bằng cách làm cực tiểu $\phi(u)$ lần lượt theo từng tọa độ. Nói chung $u^{(k+1)}$ với $k = 0, 1, 2, \ldots$ được tính theo công thức:

$$u_i^{(k+1)} = \max\{0, w_i^{(k+1)}\}$$

với

$$w_i^{(k+1)} = (\frac{-1}{f_{ii}}\Big(\sum_{j=1}^{i-1} f_{ij}u_j^{(k+1)} + \frac{h_i}{2} + \sum_{j=i+1}^{m} f_{ij}u_j^{(k)}\Big). \tag{2.52}$$

Ví dụ minh họa. Giải bài toán quy hoạch toàn phương:

$$0.5x_1^2 + 0.5x_2^2 - x_1 - 2x_2 \to \min,$$

Với các ràng buộc

$$\begin{cases} 2x_1 + 3x_2 \le 6 \\ x_1 + 4x_2 \le 5 \\ x_1 \ge 0 \\ x_2 \ge 0, \end{cases}$$

Ta có

$$x = (x_1, x_2)^T, p = (-1, -2)^T, b = (6, 5, 0, 0)^T$$

$$C = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}.$$

$$A = \begin{pmatrix} 2 & 3 \\ 1 & 4 \\ -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

$$C^{-1} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Ta tính được

$$h = (1/2)AC^{-1}p + b = (-2, -4, 1, 2)^T$$

và

$$G = (1/4)AC^{-1}A^T == \begin{pmatrix} 13/2 & 7 & -1 & -3/2 \\ 7 & 17/2 & -1/2 & -2 \\ -1 & -1/2 & 1/2 & 0 \\ -3/2 & -2 & 0 & 1/2 \end{pmatrix}.$$

Áp dụng công thức (2.52), tính các điểm xấp xỉ :

$$u^{(k)} = (u_1^{(k)}, u_2^{(k)}, u_3^{(k)}, u_4^{(k)})'$$

thỏa

$$\begin{aligned}
u_1^{(k)} &= \max\{0, (-2/13)(-1 + 7u_2^{(k-1)} - u_3^{(k-1)} - 3/2u_4^{(k-1)} + \frac{h_1}{2}\} \\
u_2^{(k)} &= \max\{0, (-2/17)(7u_1^{(k)} - 2 - 1/2u_3^{(k-1)} - 2u_{(4)}^{(k-1)} + \frac{h_2}{2}\} \\
u_3^{(k)} &= \max\{0, -2(-u_1^{(k)} - 1/2u_2^{(k)} + 0 * u_4^{(k-1)} + \frac{h_3}{2}\} \\
u_4^{(k)} &= \max\{0, -2(-3/2u_1^{(k)} - 2u_2^{(k)} + 0 * u_3^{(k)} + 1 + \frac{h_4}{2}\}.
\end{aligned}$$

Chọn $u^{(0)} = (0, 0, 0, 0)$ ta tính được $u^{(1)} = (u_1^{(1)}, u_2^{(1)}, u_3^{(1)}, u_4^{(1)})$ theo

$$\begin{aligned}
u_1^{(1)} &= \max\{0, (-2/13)(-1) = 0.1538\} \\
u_2^{(1)} &= \max\{0, (-2/17)(7 * 0.15 - 2 = 0.11\} \\
u_3^{(1)} &= \max\{0, -2(-0.15 - 1/2 * 0.11 + 1/2) = 0\} \\
u_4^{(1)} &= \max\{0, -2(-3/2 * 0.15 - 2 * 0.11 + 0 + 1) = 0\}.
\end{aligned}$$

Vậy ta có $u^{(1)} = (0.15, 0.11, 0, 0)$ Ta tiếp tục tính $u^{(2)}$

$$\begin{aligned}
u_1^{(2)} &= \max\{0, (-2/13)(-1 + 7 * 0.11 - 0 - 0)\} = 0.04 \\
u_2^{(2)} &= \max\{0, (-2/17)(7 * 0.04 - 2 - 0 - 0\} = 0.2 \\
u_3^{(2)} &= \max\{0, -2(-0.04 - 1/2 * 0.2 - 1/2)\} = 0 \\
u_4^{(2)} &= \max\{0, -2(-3/2 * 0.04 - 2 * 0.2 + 0 + 1)\} = 0.
\end{aligned}$$

Do đó $u^{(2)} = (0.04, 0.2, 0, 0)$ Ta tiếp tục tính $u^{(3)}$

$$\begin{aligned} u_1^{(3)} &= \max\{0, (-2/13)(-1 + 7 * 0.2 - 0 - 0)\} = 0 \\ u_2^{(3)} &= \max\{0, (-2/17)(7 * 0 - 2 - 0 - 0\} = 0.24 \\ u_3^{(3)} &= \max\{0, -2(-0 - 1/2 * 0.24 + 1/2)\} = 0 \\ u_4^{(3)} &= \max\{0, -2(0 - 2 * 0.24 + 0 + 1)\} = 0. \end{aligned}$$

Do đó $u^{(3)} = (0, 0.24, 0, 0)$. Ta tiếp tục tính $u^{(4)}$

$$\begin{aligned} u_1^{(4)} &= \max\{0, (-2/13)(-1 + 7 * 0.24 - 0 - 0)\} = 0 \\ u_2^{(4)} &= \max\{0, (-2/17)(7 * 0 - 2 - 0 - 0\} = 0.24 \\ u_3^{(4)} &= \max\{0, -2(-0 - 1/2 * 0.24 + 1/2)\} = 0 \\ u_4^{(4)} &= \max\{0, -2(0 - 2 * 0.24 + 0 + 1)\} = 0. \end{aligned}$$

Vậy $u^{(4)} = (0, 0.24, 0, 0)$. Ở các bước trên được tính toán trên MathLap, kết quả cuối cùng của thuật toán này là:

$$x^* = -1/2C^{-1}(A^T u + p) = \begin{pmatrix} 0.5837 \\ 1.1041 \end{pmatrix}.$$

Giá trị tối ưu của hàm mục tiêu: $f_{\min} = -2.0120$.

Có thể dùng thư viện MATHLAB để giải bài toán quy hoạch toàn phương. Ví dụ sau đây chỉ ra điều đó.

$$f(x) = x_1^2 + 2x_2^2 - 4x_1 - 2x_1x_2 \to \min,$$

Với các ràng buộc

$$\begin{cases} 2x_1 + x_2 \le 6 \\ x_1 - 4x_2 \le 0 \\ x_1 \ge 0 \\ x_2 \ge 0, \end{cases}$$

B1: Đưa về dạng $f(x) = 1/2x'Cx + g'x$, ta được:

$$H = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix} \quad g = \begin{pmatrix} -4 \\ 0 \end{pmatrix}$$

B2: Chuẩn hóa dạng ràng buộc dạng $Ax \leq b$ trong đó $b = (6, 0)'$
$A = \begin{pmatrix} 2 & 1 \\ 1 & -4 \end{pmatrix}$

B3: Sử dụng lệnh giải bài toán quy hoạch toàn phương

$$[x, fval] = \text{quadprog}(C, g, A, b)$$

Nghiệm tính toán nhận được với số liệu được nhập vào trong quá trình sau:

$$C = [2 \quad -2; -2 \quad 4]$$

$$g = [-4 \quad 0]$$

$$A = [2 \quad 1; 1 \quad -4]$$

$$b = [6; 0]$$

$$[x, fval] = \text{quadprog}(C, g, A, b).$$

$$x =$$

2.4.615

1.0769

2.4. THUẬT TOÁN XẤP XỈ NGOÀI

Xét bài toán

$$f(x) \to \min, x \in D, \tag{P}$$

trong đó D là tập lồi trong $I\!R^n$ và $f : I\!R^n \to I\!R$ là hàm lõm. Phương pháp xấp xỉ ngoài của tập chấp nhận được bởi một dãy các tập thay thế đơn giản hơn là một phương pháp cơ bản trong lý thuyết tối ưu. Nội dung cơ bản của phương pháp là: thay vì giải bài toán (P) người ta giải thay thế bằng một loạt bài toán đơn giản hơn

$$f(x) \to \min, x \in D_k, \tag{P_k},$$

trong đó $I\!R^n \supset D_1 \supset D_2 \supset \ldots D$ và

$$\min f(D_k) \to \min f(D) \text{ khi } k \to \infty.$$

Thông thường tập D_k phải nằm trong họ F, thỏa mãn một số tính chất sau:

(a) Dãy D_k là đóng, và bài toán (P_k) với $D_k \subset F$ có nghiệm và có thể giải được theo một thuật toán có sẵn nào đó.

(b) Với mỗi $D_k \subset F$ nào đó chứa D và điểm $x^{(k)} \in D_k \setminus D$ có thể tìm được hàm $l_k : I\!R^n \to I\!R$ thỏa mãn:

- $l_k(x) \leq 0 \ \forall x \in D$,
- $l_k(x^{(k)}) > 0$,
- $\{x \in D_k \mid l_k(x) \leq 0\} \in F$.

Với các điều kiện trên lời giải của phương pháp được xác định như sau:

2.4.1 Phương pháp xấp xỉ ngoài tổng quát

Phương pháp xấp xỉ ngoài tổng quát như sau (dùng cho hàm mục tiêu liên tục)

B1: Chọn $D_1 \in F$ sao cho $D_1 \supset D$. Đặt $k := 1$.

B2: Thực hiện với $k = 1, 2, \ldots$.

B3: Giải bài toán (P_k) nhận được $x^{(k)} \in \text{argmin} f(D_k)$ thỏa mãn các điều kiện của (b). Nếu $x^{(k)} \in D$, dừng và $x^{(k)}$ nghiệm của (P). Nếu không chuyển B4.

B4: Cấu trúc hàm l_k thỏa mãn (b) và $D_{k+1} := D_k \cap \{x \mid l_k(x) \leq 0\}, k := k+1$, chuyển B2.

Định lý 2.4.31. *Giả sử rằng*

(i) l_k *nửa liên tục dưới với mọi* $k = 1, 2, \dots,$

(ii) Mỗi dãy con hội tụ $\{x^q\} \subset \{x^{(k)}\}$ *thỏa mãn* $\lim_{q\to\infty} x^q = \overline{x}$ *chứa* $\{x^r\} \subset \{x^q\}$ *sao cho* $\lim_{r\to\infty} l_r(x^r) = \lim_r(\overline{x})$,

(iii) $\lim_{r\to\infty} l_r(\overline{x}) = 0$ *suy ra* $\overline{x} \in D$.

Khi đó mọi điểm tích lũy của dãy $\{x^{(k)}\}$ *nằm trong* D *và kéo theo là nghiệm của (P).*

Ghi chú Mục trên được trích dẫn từ các tài liệu sau: [3], [4], [**?**],...

Chương 3.

CÁC PHƯƠNG PHÁP TÌM NGHIỆM CỦA BÀI TOÁN QUY HOẠCH PHI TUYẾN

Trong chương này chúng tôi tập trung trình bày các thuật toán tìm kiếm phương án tối ưu (cực trị địa phương và toàn cục). Những vấn đề lý thuyết của chương này đã được trình bày ở các chương trước, nên chúng tôi chỉ nhắc những định lý và định nghĩa khi thật cần thiết. Lưu ý rằng đối với tối ưu phi tuyến cực trị địa phương và cực trị toàn cục là hoàn toàn khác biệt. Bài toán sau gọi là bài toán quy hoạch phi tuyến

$$f(x) \rightarrow \min, x \in I\!R^n \tag{3.53}$$

$$f_i(x) \leq 0, i = 1, \dots m, \tag{3.54}$$

$$h_j(x) = 0, j = 1, \dots, p \tag{3.55}$$

trong đó ít nhất một trong các hàm f, f_i, h_j là phi tuyến. Nếu bài toán chỉ có dạng (3.53) thì gọi là bài toán quy hoạch phi tuyến không ràng buộc. Để cho gọn ta có thể viết lại bài toán trên như sau

$$\begin{aligned} f(x) &\rightarrow \min, x \in D \qquad (3.56)\\ D &= \{x \in I\!R^n \mid f_i(x) \leq 0, h_j(x) = 0, i = 1, \dots m, j = 1, \dots, p\} \end{aligned}$$

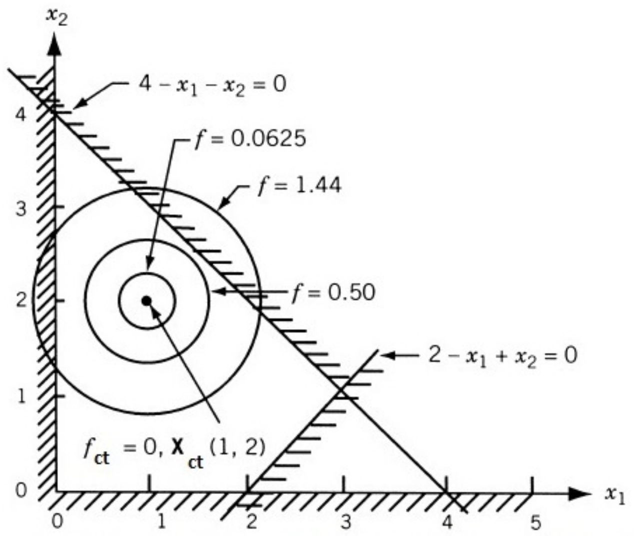
trong đó ít nhất một trong các hàm f, f_i, h_j là phi tuyến. Nhắc lại rằng: Điểm $x^* \in D$ gọi là điểm cực tiểu toàn cục của bài toán (3.56) nếu

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

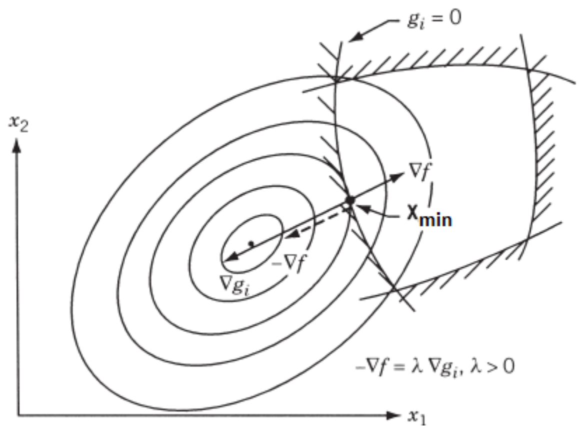
và gọi là cực tiểu địa phương nếu tồn tại lân cận U chứa x^* sao cho

$$f(x^*) \leq f(x) \quad \forall x \in U.$$

Điểm cực trị đạt được nói chung phụ thuộc nhiều vào các ràng buộc và các hình dưới đây sẽ mô tả vị trí của các điểm đạt cực trị tùy theo các ràng buộc của bài toán.

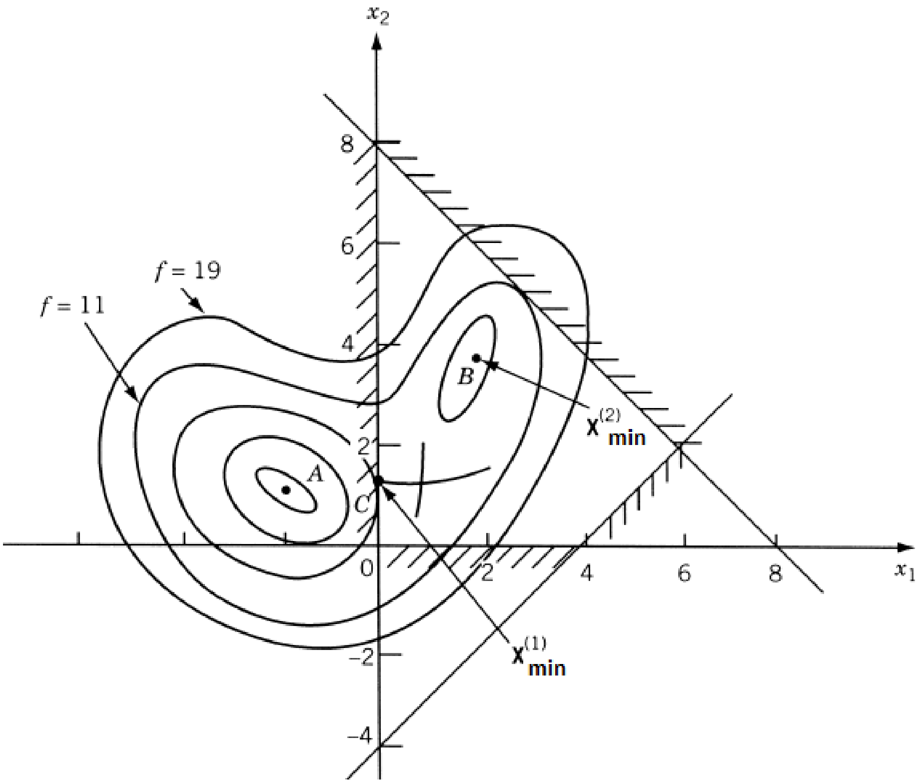


Hình 3.15: Cực trị không ràng buộc và ràng buộc là như nhau(ràng buộc tuyến tính)

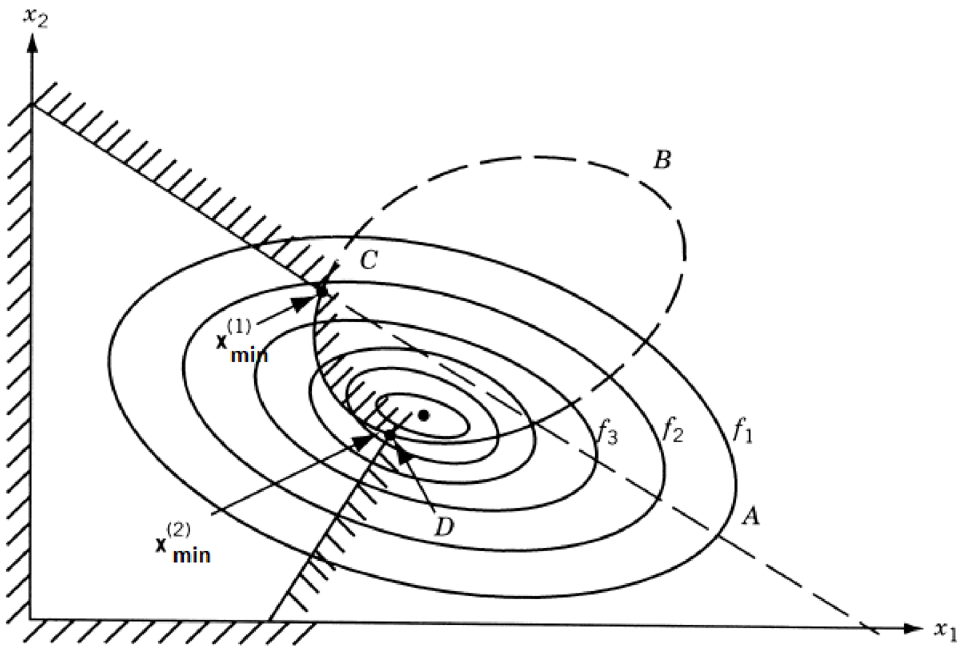


Hình 3.16: Cực trị với ràng buộc phi tuyến

Lưu ý rằng, trong chương này ta chỉ xét các bài toán tối ưu trong không gian Euclid $I\!R^n$ và để tránh nhầm lẫn với ký hiệu x mũ k ta ký hiệu véc tơ nghiệm ở bước lặp thứ k là $x^{(k)}$. Đồng thời cũng lưu ý rằng véc tơ gradient chúng ta cũng để ở dạng véc tơ cột thay vì như trước chúng ta để ở dạng véc tơ hàng, tức là: $\nabla f(x) := (\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \ldots, \frac{\partial f(x)}{\partial x_n})^T$.



Hình 3.17: Cực tri tương đối nhận được theo hàm mục tiêu



Hình 3.18: Cực trị nhận được theo ràng buộc

3.1.2 Phân loại các phương pháp

Có nhiều phương pháp giải bài toán quy hoạch phi tuyến. Ta có thể chia chúng ra làm các nhóm sau:

1. Các phương pháp gradient (khi f khả vi), phương pháp này dựa theo tính chất sau: gradient của hàm mục tiêu tại phương án bất kỳ là véc tơ nằm trong hướng tăng cục bộ của $f(x)$ do đó ta đi theo hướng ngược lại chừng nào hàm mục tiêu chưa tăng. Sau khi xác định được điểm mới ta lại tìm theo hướng mới, ta lặp lại quá trình trên.

2. Phương pháp hướng chấp nhận được, thực chất của phương pháp là: đối với mỗi điểm x thuộc miền ràng buộc, chọn một hướng mà theo đó hàm mục tiêu giảm và bước di chuyển hợp lý để không ra khỏi miền ràng buộc.

3. Phương pháp hàm phạt là thay cho hàm f, ta giải bài toán tối ưu không ràng buộc với hàm mục tiêu mới $f(x) + p(x)$, trong đó $p(x)$ là lượng phạt khi x vi phạm các ràng buộc.

4. Phương pháp tổ hợp và tìm kiếm ngẫu nhiên là hoặc nêu tất cả các phương án, hoặc tìm tiêu chuẩn bỏ bớt một số phương án mà chắc chắn không cho nghiệm bằng cách dùng quá trình ngẫu nhiên kiểu xích Markov hoặc phương pháp Monte-Carlo.

5. Phương pháp cực tiểu hàm lõm là dùng phương pháp cắt và chia nón.

3.2. CÁC PHƯƠNG PHÁP CỰC TIỂU HÀM MỘT BIẾN

Trong mục này ta xét bài toán sau:

$$f(x) \to \min,\ x \in [a, b].$$

3.2.1 Phương pháp phân đôi

Trong trường hợp f là liên tục và lồi, thuật toán phân đôi gồm các bước sau:

B1: Tính $\overline{x} = (a+b)/2, f(\overline{x}), f(a), f(b)$ chia đều thành 4 đoạn

$$[a, x_1], [x_1, \overline{x}], [\overline{x}, x_3], [x_3, b]. \text{Tính } f(x_1), f(x_2).$$

B2: Nếu $f(\overline{x}) \le f(x_1)$ và $f(\overline{x}) \le f(x_2)$ thì $\overline{a} := x_1, \overline{b} := x_2$, qua lại B1 với a, b là $\overline{a}, \overline{b}$ tương ứng.

B3: Nếu $f(x_2) \le f(\overline{x})$ thì gán $\overline{a} := \overline{x}$ quay lại B1.

B3: Nếu $f(x_1) \le f(\overline{x})$ thì gán $\overline{b} := \overline{x}$ quay lại B1. Quá trình tiếp tục cho đến khi $|\overline{a} - \overline{b}| \le \epsilon$ cho trước thì dừng.

Nhận xét 3.2.6. - *Nghiệm gần đúng là* $\overline{x}$ *và* $|x_1 - \overline{x}| = |x_2 - \overline{x}| \le (\overline{b} - \overline{a})/2$.
- *Khi* f *là hàm liên tục, để áp dụng phương pháp trên ta thực hiện như sau: Chia trục* x *thành* m *khoảng đủ nhỏ để có thể bắt được những khoảng chứa điểm cực trị (thường chia tăng theo cấp số nhân* $x_{i+1} = qx_i, q > 1$ *để mở rộng khoảng được xét), sau đó tính* $f(x_i)$ *với giả thiết các khoảng chia đủ nhỏ thì điểm cực tiểu sẽ nằm trong đoạn* $[x_{k-1}, x_k]$ *nếu thỏa mãn điều kiện* $f(x_k) \le f(x_{k+1})$ *và* $f(x_k) \le f(x_{k-1})$*. Ký hiệu* $a = x_{k-1}, b = x_{k+1}$ *và coi* f *trong khoảng đó là lồi.*

3.2.2 Phương pháp lát cắt vàng

Gọi $\tau^* = \frac{\sqrt{5}-1}{2} = 0.61804, 1 - \tau^* = 0.382$. Ta thấy $1 - \tau^* = \tau^{*2}]$.

B1: Chia ba đoạn $[a.b]$ bởi các điểm x_1, x_2 thỏa mãn $x_1 := a + (1-\tau^*)(b - a), x_2 := a + \tau^*(b-a)$, tính $f(a), f(b), f_1 = f(x_1), f_2 = f(x_2)$.

B2: Nếu $f_1 < f_2$ chuyển B3, nếu không chuyển B4.

B3: Lấy đoạn $[\overline{a}, \overline{b}]$ mới với $\overline{a} := a, \overline{b} := x_2$.
- Nếu $|\overline{b} - \overline{a}| \leq \epsilon$ thì cực tiểu là $x^* = x_1$, dừng tính toán.
- Nếu $|\overline{b} - \overline{a}| > \epsilon$, chia đoạn $[\overline{a}, \overline{b}]$ thành ba đoạn bởi các điểm chia $\overline{x}_1, \overline{x}_2$ với $\overline{x}_2 = x_1$, tính $\overline{f}_2 = f(\overline{x}_2) = f_1$ và $\overline{x}_1 = \overline{a} + \tau^*(\overline{b} - \overline{a})$, tính $\overline{f}_1 = f(\overline{x}_1)$.
- Chuyển B2 với các giá trị mới $\overline{a}, \overline{b}, \overline{x}_1, \overline{x}_2, \overline{f}_1, \overline{f}_2$.

B4: Lấy đoạn $[\overline{a}, \overline{b}]$ mới với $\overline{a} = x_1, \overline{b} = b]$. - Nếu $|\overline{b} - \overline{a}| \leq \epsilon$ thì cực tiểu là $x^* = x_2$, dừng tính toán.
- Nếu $|\overline{b} - \overline{a}| > \epsilon$, chia đoạn $[\overline{a}, \overline{b}]$ thành ba đoạn bởi các điểm chia $\overline{x}_1, \overline{x}_2$ với $\overline{x}_1 = x_2$, do đó $\overline{f}_1 = f(\overline{x}_1) = f_2$ và $\overline{x}_2 = \overline{a} + (1 - \tau^*)(\overline{b} - \overline{a})$, tính $\overline{f}_2 = f(\overline{x}_2)$.
- Chuyển B2 với các giá trị mới $\overline{a}, \overline{b}, \overline{x}_1, \overline{x}_2, \overline{f}_1, \overline{f}_2$.
- Sau hữu hạn bước ta xác định được phương án tối ưu.

Nhận xét 3.2.7. - *Phương án tối ưu* $x^* \in [\overline{x}_1, \overline{x}_2]$.
- *Ưu điểm của phương pháp là một trong hai điểm chia trùng với điểm chia cũ, do đó ở mỗi bước lặp chỉ cần tính thêm một giá trị hàm ứng với điểm chia mới.*

3.2.3 Phương pháp nội suy

Định nghĩa 3.2.28. *Giả sử biết giá trị của hàm tại 3 điểm* α, β, γ *và giá trị hàm tương ứng* $f_\alpha, f_\beta, f_\gamma$. *Tam thức bậc 2*

$$\phi(x) := Ax^2 + Bx + C,$$

trong đó A, B, C *được xác lập từ hệ phương trình:*

$$\begin{cases} A\alpha^2 + B\alpha + C = f_\alpha \\ A\beta^2 + B\beta + C = f_\beta \\ A\gamma^2 + B\gamma + C = f_\gamma \end{cases}$$

gọi là hàm nội suy bậc 2 tại 3 điểm α, β, γ.

Giải hệ trên ta nhận được:

$$\begin{cases} A = (\gamma - \beta)f_\alpha + (\alpha - \gamma)f_\beta + (\beta - \gamma)f_\gamma/\Delta \\ B = (\beta^2 - \gamma^2)f_\alpha + (\gamma^2 - \alpha^2)f_\beta + (\alpha^2 - \beta^2)f_\gamma/\Delta \\ C = \beta\gamma(\gamma - \beta)f_\alpha + \gamma\alpha(\alpha - \gamma)f_\beta + \alpha\beta(\beta - \alpha)f_\gamma/\Delta \end{cases}$$

Khi đó hàm $\phi(x)$ đạt cực tiểu tại $x^* = -B/(2A)$ nếu $A > 0$. Vậy có thể xấp xỉ điểm cực tiểu của hàm $f(x)$ bởi giá trị

$$x^* = \frac{1}{2}\frac{(\beta^2 - \gamma^2)f_\alpha + (\gamma^2 - \alpha^2)f_\beta + (\alpha^2 - \beta^2)f_\gamma}{\beta\gamma(\gamma - \beta)f_\alpha + \gamma\alpha(\alpha - \gamma)f_\beta + \alpha\beta(\beta - \alpha)f_\gamma} \tag{3.57}$$

Nội dung thuật toán:

B1: Khởi tạo giá trị cực tiểu ban đầu $x^{(0)}$ bước dịch chuyển h, tính $f(x(0)), f(x^{(0)} + h)$

B2: Nếu $f(x(0)) < f(x^{(0)} + h)$ chon điểm thứ 3 là $x^{(0)} - h$ tinh $f(x^{(0)} - h)$. Ngược lại chọn điểm thứ ba $x^{(0)} + 2h$, tinh $f(x^{(0)} + 2h)$.

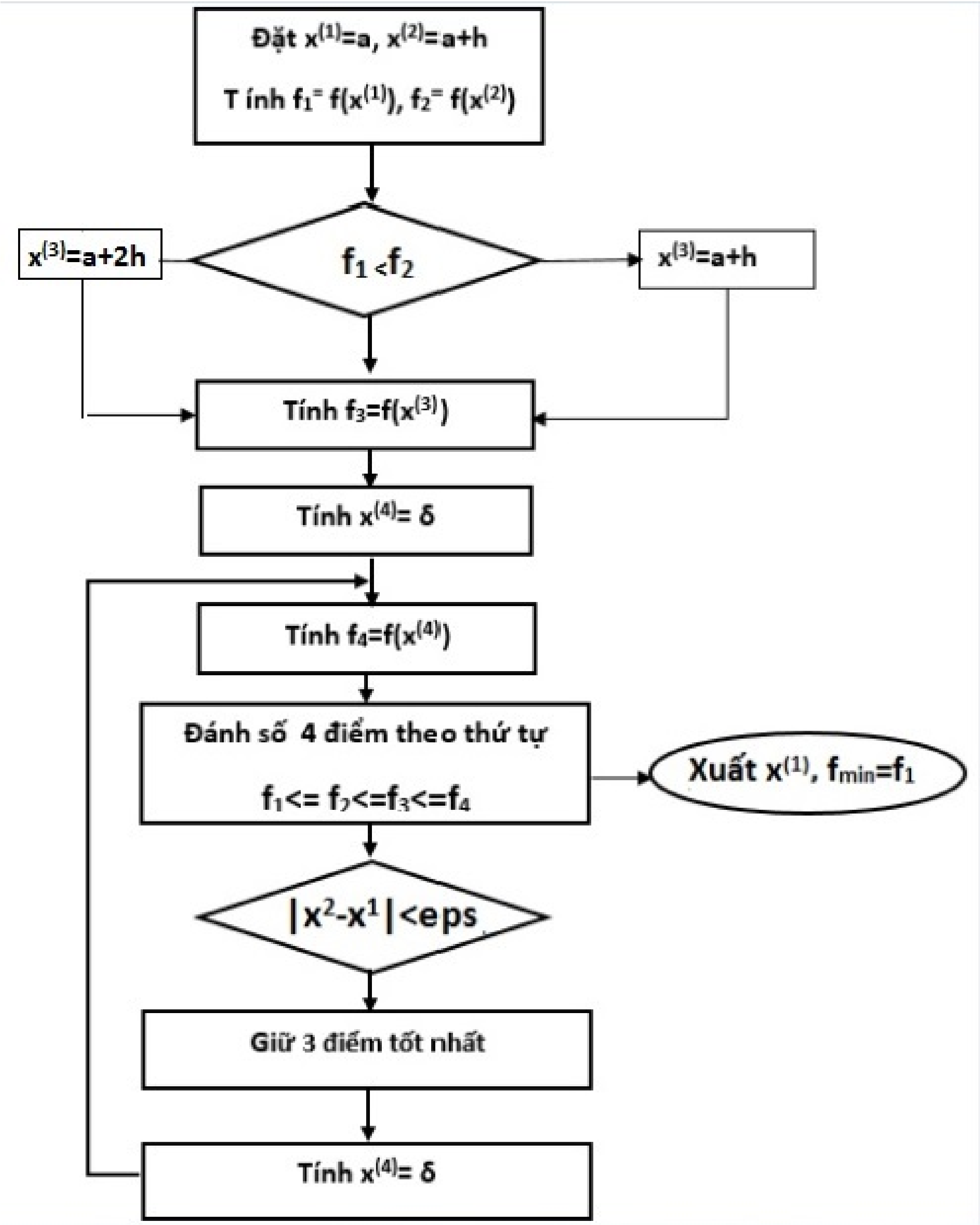
B3: Tính nghiệm theo công thức 3.61 và lặp lại B1. Thuật toán dừng khi trị tuyệt đối của hiệu của nghiệm xấp xỉ hai bước gần nhau nhỏ hơn một số ϵ cho trước.

Nhận xét 3.2.8. *Khi ϵ chọn đủ nhỏ thì $\alpha, \beta, \gamma, f_\alpha, f_\beta, f_\gamma$ rất gần nhau nên (3.61) không tính được x^* do đó từ bước nội suy thứ 2 trở đi người ta thay bởi:*

$$x^* = \frac{1}{2}(\alpha + \beta) + \frac{1}{2}\frac{(f_\alpha - f_\beta)(\beta - \gamma)(\gamma - \alpha)}{(\beta - \gamma)f_\alpha + (\gamma - \alpha)f_\beta + (\alpha - \beta)f_\gamma} \tag{3.58}$$

Ví dụ 3.2.25. *Dùng phép nội suy bậc 2 tìm cực tiểu của hàm số $f(x) = 2x^2 - e^x$ với độ chính xác 0.001. Các giá trị ban đầu được cho là: $a = 0.5, h = 0.5$*

Sử dụng thuật toán trên ta tính được $f_{\min} = -1.17413808$.



Hình 3.19: **Sơ đồ khối phương pháp nội suy**

Sử dụng Optimset trong MATHLAB để tìm điểm cực tiểu của hàm một biến thông qua ví dụ sau:

$$f(x) = 0.65 - \frac{0.75}{1+x^2} - 0.65x\tan^{-1}(\frac{1}{x})$$

B1: Viết hàm mục tiêu vào M-file.

$$\text{function f} = \text{objecfun}(x)$$

$$f = 0.65 - \frac{0.75}{1+x^2} - 0.65x\tan^{-1}(\frac{1}{x})$$

B2: options=optimset('LargeScale','off');

$[x, fval]$=fminbnd(@ objjfun,0,0.5,options)

3.3. QUY HOẠCH PHI TUYẾN KHÔNG RÀNG BUỘC - CÁC PHƯƠNG PHÁP DÙNG ĐẠO HÀM

Trong mục này ta xét bài toán (3.53), tức là bài toán sau:

$$f(x) \to \min, \quad x \in I\!R^n.$$

3.3.1 Các định lý cơ bản của cực trị địa phương

Ta nhắc lại 2 định lý quan trong đã trình bày ở Chương 1 về điều kiện cần và định lý về điều kiện đủ để tồn tại điểm tối ưu sau:

Định lý về điều kiện cần Nếu $f(x)$ khả vi tại x^* và điểm $x*$ là điểm trị địa phương thì $\nabla f(x^*) = 0$, tức là x^* là điểm dừng.

Định lý về điều kiện đủ về tồn tại điểm cực tiểu địa phương Nếu

1. $f(x)$ khả vi tại x^*
2. $\nabla f(x^*) = 0$, tức là x^* là điểm dừng
3. $H(x^*) := \nabla^2 f(x^*) > 0$, tức là ma trận Hessian xác định dương.

thì điểm x^* là điểm cực tiểu địa phương.

Nhớ rằng nếu hàm là lồi thì cực tiểu địa phương là cực tiểu toàn cục.

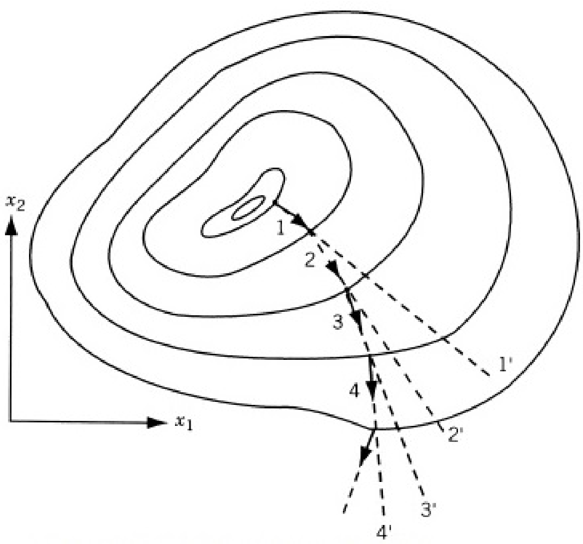
3.3.2 Các phương pháp dùng đạo hàm - Phương pháp gradient

Nội dung của phương pháp như sau:

B1: Chọn $x^{(0)} \in I\!R^n$

B2: Chọn dãy lặp $x^{(k+1)} = x^{(k)} - \lambda_k \bigtriangledown f(x^{(k)})$, trong đó λ_k là hệ số có thể lấy cố định hoăc lấy giá trị cực tiểu theo tham số λ.

B3: Có thể dãy không hội tụ, trong trường hợp này ta chọn lại λ_k nhỏ hơn. Khi λ đủ nhỏ $x^{(k)}$ sẽ hội tụ về nghiệm tối ưu.



Hình 3.20: Mô tả phương pháp gradient

Ví dụ 3.3.26. *Tìm* min *của hàm* $f(x) = x^2 + 3$
Khi đó $\bigtriangledown f(x) = 2x$. *Chọn* $x^{(0)} = 1, \bigtriangledown f(x^{(0)}) = 2$ $x^{(1)} = x^{(0)} - \lambda \bigtriangledown f(x^{(0)}) = 1 - 2\lambda; \lambda > 0$. $\bigtriangledown f(x^{(1)} = 2x^{(1)} = 2(1-2\lambda) \neq 0$ *khi* $\lambda <> 1/2$ $x^{(2)} = (1-2\lambda) - 2\lambda(1-2\lambda) = (1-2\lambda)^2$. *Tương tự* $x^{(k)} = (1-2\lambda)^k$.

Với $0 < \lambda < 1$ thì $x^{(k)} \to 0, k \to \infty$. Do đó $x^* = 0$ là điểm cực tiểu toàn cục.

3.3.3 Các phương pháp dùng đạo hàm - Phương pháp đường dốc nhất

Nội dung phương pháp:

B1: Chọn $x^{(0)}$, tính $\| \bigtriangledown f(x^{(k)})\|; s(k) = \bigtriangledown f(x^{(k)})/\| \bigtriangledown f(x^{(k)})\|$

B2: $x^{(k+1)} := x^{(k)} - \lambda_k s^{(k)}$

B3: Tối ưu hóa hàm $f(x^{(k+1)})$ theo tham số λ_k. Tìm được min λ_k^*.

B4: Chọn điểm mới $x^{(k+1)} := x^{(k)} - \lambda_k^* s^{(k)}$

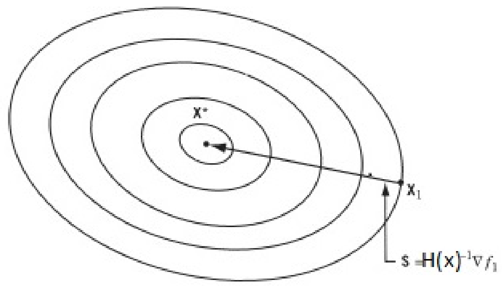
Tìm min của hàm Rosenbrock: $f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$.
- Chọn $x^{(0)} = (-0.5, 0.5)^T, f(x^{(0)}) = 8.5$
- Tính $\bigtriangledown f(x^{(k)})$ và $\| \bigtriangledown f(x^{(k)})\|$
Mặt khác $\frac{\partial f(x^{(k)})}{\partial x_1} = 47, \frac{\partial f(x^{(k)})}{\partial x_2} = 50,$
$\| \bigtriangledown f(x^{(k)})\| = \sqrt{47^2 + 50^2} = 68.6,, \ s^{(k)} = 1/68(47, 50)^T = (0.685, 0.729)^T$.
Véc tơ $s^{(k)}$ vuông góc với đường mức của $f(x)$ tại $x^{(k)} = (-0.5, 0.5)^T$.

3.3.4 Các phương pháp dùng đạo hàm - Phương pháp Newton

Đối với phương pháp này ta phải giả thiết hàm mục tiêu có đạo hàm cấp 2 liên tục. Nhắc lại rằng ma trận Hessian có dạng $H(x) := \nabla^2(x) = \Big(\partial^2 f(x)/\partial x_i x_j\Big)$ Ta chọn các bước lặp như sau:

$$x^{(k+1)} = x^{(k)} + \lambda_k^* H^{-1}(x^{(k)})\nabla f(x^{(k)}) = x^{(k)} + \lambda_k^* s^{(k)}$$

trong đó $s^{(k)} := H^{-1}(x^{(k)})\nabla f(x^{(k)})$.



Hình 3.21: Véc tơ hướng chấp nhận được trong Phương pháp Newton

Ví dụ 3.3.27. *Tìm x^* sao cho*

$$f(x) = 4x_1 + x_2 + \frac{1}{x_1} + \frac{1}{x_2} \to \min$$

-

$$\nabla f(x) = \begin{pmatrix} 4 - \frac{1}{x_1^2} \\ 1 - \frac{1}{x_2^2} \end{pmatrix}$$

-

$$H(x) = \begin{bmatrix} \frac{2}{x_1^3} & 0 \\ 0 & \frac{2}{x_2^3} \end{bmatrix}$$

- Vécto ban đầu $x^{(0)} = (1.13, 3.56)^T$
- $\nabla f(x^{(0)}) = (3.2, 0.92)^T$
-

$$H(x^{(0)}) = \begin{bmatrix} 1.41 & 0 \\ 0 & 0.04 \end{bmatrix}$$

-

$$H^{-1}(x^{(0)}) = \begin{bmatrix} 0.71 & 0 \\ 0 & 2.5 \end{bmatrix}$$

- Xác định λ_k^* từ cực tiểu theo λ của hàm$f(x^{(k)} + \lambda s^{(k)}) =$ và tính được $\lambda_0^* = 0.112$
- Tính ra ta được: $x^{(1)} = x^{(0)} - \lambda_0 H^{-1}(x^{(0)}) \nabla f(x^{(0)}) = (1.13, 3.36)^T - 0.112(2.28, 2.3) = (0.88, 0.98)^T$
- Tiếp tục lược đồ trên ta nhận được nghiệm gần đúng của bài toán.

3.3.5 Các phương pháp dùng đạo hàm - Phương pháp Gradien liên hợp

Định nghĩa 3.3.29. *(**Hướng liên hợp**) Hai hướng s_i, s_j gọi là liên hợp nhau nếu:*

1. $\langle \boldsymbol{H}s_i, s_j\rangle = 0, i \neq j$
2. $\langle \boldsymbol{H}s_i, s_j\rangle \geq= 0, i = j$ *trong đó* $\boldsymbol{H} = \nabla^2 f(x)$

Ví dụ $f(x) = 2x_1^2 + x_2^2 - x_1 + x_2, s_1 = (0,1)^T, s_2 = (1,1)^T$ *không liên hợp vì* $\langle \boldsymbol{H}(s_i), s_j\rangle = \langle \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \rangle = 2 \neq 0.$

Với hàm $f(x) = 2x_1^2 + 16x_2^2 - 2x_1x_2 - x_1 - 6x_2 - 5$ *thì*

$$\langle Hs_1, s_2\rangle = \langle \begin{bmatrix} 4 & -2 \\ -2 & 32 \end{bmatrix} \begin{bmatrix} 15 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \rangle = 0.$$

Do đó với $s_1 = (15,-1)^T, s_2 = (1,1)^T$ *sẽ liên hợp nhau vì* $\langle \boldsymbol{H}s_1, s_2\rangle = 0$

Phương pháp Fletcher - Reeves được thực hiện thông qua các bước sau:

B1: $k = 1$, Chọn điểm xuất phát $x^{(k)}$

B2: Xác định hướng tìm thứ nhất $s_k = -\nabla f(x^{(k)})$

B3: Tìm điểm $x^{(k+1)}$ theo quan hệ $x^{(k+1)} = x^{(k)} + \lambda_k^* s_k$, trong đó λ_k^* là độ dài tối ưu theo hướng s_k. Nếu điều kiện tối ưu thỏa mãn dừng, nếu không chuyển sang B4.

B4: Tính

$$\nabla f(x^{(k+1)}) \text{ và } s_{k+1} = -\nabla f(x^{(k+1)}) + \frac{|\nabla f(x^{(k+1)})|^2}{|\nabla f(x^k)|^2} s_k$$

B5. Tìm điểm $\lambda_{(k+1)}^*$ theo hướng s_{k+1} và quay lại B3.

Ví dụ 3.3.28. *Giải bài toán*

$$f(x) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2 \to \min$$

1. - Chọn $x^{(1)} = (0,0)$

- Tính

$$\nabla f(x^{(1)}) = \begin{bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{bmatrix}$$

Hướng tìm $s_1 = -\nabla f(x^{(1)}) = (-1,1)^T$

- Tối ưu hóa hàm $f(x^{(1)} + \lambda_1 s_1) = f(-\lambda_1, \lambda_1) = \lambda_1^2 - 2\lambda_1$ *nhận được* $\lambda_1^* = 1$.

- Xác định $x^{(2)} = x^{(1)} + \lambda_1^* s_1 = (-1,1)^T$

2. - Tính $\nabla f(x^{(2)}) = (-1,-1)^T, s_2 = -\nabla f(x^{(2)}) + |\nabla f(x^{(2)})|^2 / |\nabla f(x^{(1)})|^2 s_1 = (0,2)^T$ *- Xác định* λ_2^* :

$$f(x^{(2)} + \lambda_2 s_2) = 4\lambda_2^2 - 2\lambda_2 - 1, \lambda_2^* = 1/4$$

-$x^{(3)} = (-1, 1.5)^T$

3.

$$\nabla f(x^{(3)}) = (0,0)^T, |\nabla f(x^{(2)})|^2 = 2, s_3 = (0,0)^T$$

Do đó $x^{(3)}$ *là giá trị tối ưu.*

3.4. QUY HOẠCH PHI TUYẾN KHÔNG RÀNG BUỘC - CÁC PHƯƠNG PHÁP KHÔNG DÙNG ĐẠO HÀM

Xét bài toán

$$f(x) \to \min, \ x \in \mathbb{R}^n.$$

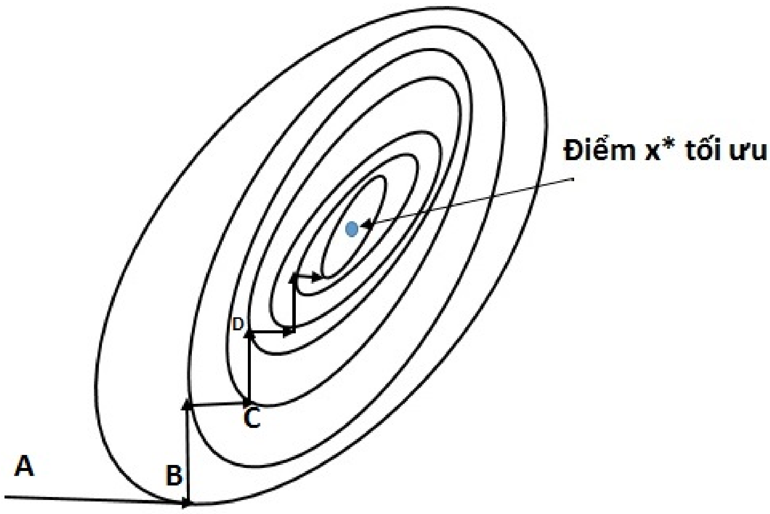
3.4.1 Phương pháp tìm theo tọa độ

Nội dung phương pháp: Giả sử trong $e_1, e_2, \ldots, e_n \in \mathbb{R}^n|$ là cơ sở đơn vị.

B1 Xuất phát từ 1 điểm A_1 bất kỳ tìm giá trị cực tiểu của hàm theo trên tia qua điểm đó và song song với e_1 ta được điểm A_2. (Điểm cực tiểu theo tia sẽ là điểm tia đã cho tiếp xúc với đường mức của hàm).

Bi Lặp lại B_1 từ điểm A_i. với e_i

Khi $i = n + 1$ gán cho $i = 1$ và quay lại B_1.



Hình 3.22: **Tìm theo tọa độ**

Phương pháp có ưu điểm là đơn giản, nhưng không hiệu quả khi n lớn.

3.4.2 Phương pháp tìm kiếm ngẫu nhiên

Phương pháp gồm các bước sau:

B1: Tạo $\delta_i :=$ random $(-0.5, 0.5)\Delta x_i$, trong đó Δx_i là các số gia ứng với thành phần x_i do ta chọn.

B2: Tính $x^{(k+1)} := x^{(k)} + \delta_i e^i$ nếu $f(x^{(k+1)} < f(x^{(k)}),\ i = 1, \ldots, n$.

B3: $x^{(k+1)} := x^{(k)}$ nếu $f(x^{(k+1)}) \geq f(x^{(k)})$.

Ví dụ $f(x) = 4(x_1 - 5)^2 + (x_2 - 6)^2$, chọn $\Delta x = (0.5, 0.5), x^{(0)} = (6, 7)^T$.

- Bước lặp đầu tiên

$x^{(1)} = (6, 7) + (-0.5 \times 0.5 \times 1, 0.5 \times 0.5 \times 0)^T = (5.75, 7), f(x^{(1)})^T = 1.5625$

$< f(x^{(0)})$ nên $x^{(1)} = (5.75, 7)$.

- Bước lặp 2: $x^{(2)} := x^{(1)} + (-0.35 \times 0.5 \times 0, -0.1 \times 0.5 \times 1)^T = (5.75, 6.95)$,

$f(x^{(2)})^T = 1.465 < f(x^{(1)})$ nên $x^{(2)} = (5.75, 6.95)$. Tiếp tục các bước ta nhận được $x^* = (5, 6)^T$.

Nếu miền $D = x \in I\!R^n : a \leq x \leq b$ ta có thể thực hiện như sau: Tạo ngẫu nhiên trong miền k (đủ lớn véc tơ) lấy giá trị nhỏ sao cho $f(x)$ nhỏ nhất.

Nhận xét 3.4.9.

Phương pháp tìm kiếm ngẫu nhiên cũng như các phương pháp lập trình tiến hóa nói chung không đánh giá được tốc độ hội tụ và lời giải tối ưu.

3.4.3 Phương pháp tìm kiếm trực tiếp (Hooke - Jeeves - Wood)

Phương pháp này rất độc đáo và hiệu quả. Ý tưởng của phương pháp là: xuất phát từ một điểm tùy ý gọi là điểm cơ sở, việc tìm kiếm được thực hiện quanh điểm cơ sở nhằm tìm được điểm tốt hơn. Nếu thành công thực hiện bước chuyển dời điểm cơ sở theo hướng giảm của hàm $f(x)$ tới điểm mới. Trái lại quay lại điểm cơ sở trước đó hoặc giảm độ dài bước dò tìm. Thuật toán gồm các bước sau:

A: ***Bước khởi sự:*** Chọn $\mathbf{b}^{(1)} = (x_1, \ldots, x_n)^T$ làm điểm cơ sở, $\Delta x_j > 0$ làm độ dài bước cho biến x_j.

B: ***Dò tìm quanh điểm cơ sở:***

1. Tính $f_b = f(\mathbf{b}^{(1)})$, giá trị của biến bắt đầu từ x_1 bằng cách thêm vào độ dài bước Δx_1. Tính $f(\mathbf{b}^{(1)} + \Delta x_1 e^1)$. Nếu $f(\mathbf{b}^{(1)} + \Delta x_1 e^1) < f(\mathbf{b}^{(1)})$ thì thay $\mathbf{b}^{(1)} := \mathbf{x}^{(1)} + \Delta x_1 e^1, f_b = f(b^{(1)})$, chuyển sang bước tiếp theo

2. Tính $f(b^{(1)} - \Delta x_1 e^1)$. Nếu $f(\mathbf{b}^{(1)} - \Delta x_1 e^1) < f(\mathbf{b}^{(1)})$ thì thay $b^{(1)}) = b^{(1)} - \Delta x_1 e^1, f_b = f(b^{(1)})$, nếu hai bước trên không làm giảm giá trị của hàm thì $\mathbf{b}^{(1)}$ không thay đổi. Trở lại bước 1. thực

hiện với $x_2, \dots, x_n$. Khi tất cả các biến đã được xét ta tìm được $\mathbf{b}^{(2)}$. Kết thúc quá trình dò tìm.

Có hai khả năng xảy ra:

(a) $\mathbf{b}^{(2)} = \mathbf{b}^{(1)}$, nghĩa là việc dò tìm không làm giảm giá trị hàm, ta tiếp tục dò tìm quanh điểm $\mathbf{b}^{(1)}$ với độ dài bước bé hơn (thường giảm 10 lần).

(b) $\mathbf{b}^{(2)} \neq \mathbf{b}^{(1)}$, nghĩa là bước dò tìm có kết quả, ta chuyển sang thực hiện bước C.

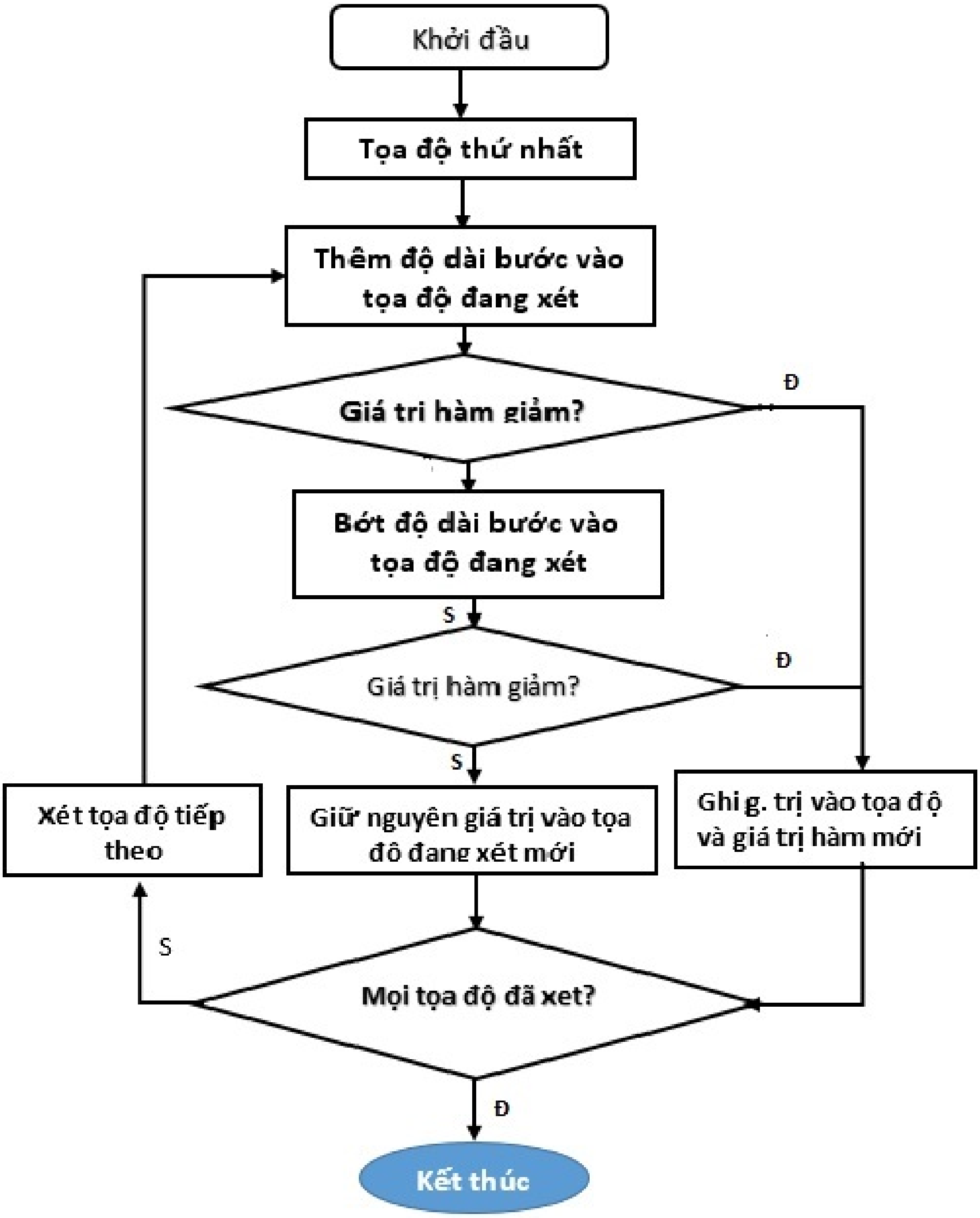
C: ***Dò tìm theo mẫu:***

1. Dịch chuyển từ điểm cơ sở $\mathbf{b}^{(2)}$ theo hướng $v := \mathbf{b}^{(2)} - \mathbf{b}^{(1)}$ vì dò theo hướng này giá trị hàm sẽ giảm ta tính giá trị hàm tại điểm mẫu mới $v := 2\mathbf{b}^{(2)} - \mathbf{b}^{(1)}$.
2. Tiến hành dò tìm quanh điểm mẫu $\mathbf{v}$ như ở bước B, nếu giá trị hàm thu được xung quanh điểm v nhỏ hơn giá trị đã thu được tại điểm cơ sở $\mathbf{b}^{(2)}$ thì ta nhận được điểm cơ sở mới $\mathbf{b}^{(3)}$.
3. Lặp lại bước dò tìm với $\mathbf{b}^{(2)}$ thay cho $\mathbf{b}^{(1)}, \mathbf{b}^{(3)}$ thay cho $\mathbf{b}^{(2)}$. Nếu không tìm được điểm có giá trị hàm tốt hơn thì dừng việc dò tìm theo mẫu từ điểm $\mathbf{b}^{(2)}$ và dò tìm quanh điểm cơ sở cũ $\mathbf{b}^{(2)}$ bằng bước B.

D: ***Dừng qua trình tìm kiếm:*** Khi độ dài bước $\Delta := \max\{\Delta x_1, \dots, \Delta x_n\}^T$ nhỏ hơn một giá trị quy định trước.

Ví dụ 3.4.29. *Dùng phương pháp trên tìm cực tiểu của hàm sau:*

$$f(x_1, x_2) = 3x_1 + 4x_1x_2 + 5x_2^2.$$



Hình 3.23: **Sơ đồ dò tìm quanh điểm cơ sở**

Giải. A. Khởi sự: - Chọn điểm cơ sở ban đầu $b^{(1)} = (4;3)$, độ dài bước $h = 1$ và $\epsilon = 10^{-3}$. Giá trị ban đầu $f_b = f(b^{(1)}) = 141$.

B. Dò tìm quanh điểm cơ sở:

- Đặt $v = b^{(1)}$.

- Xét tọa độ x_1

$+x_1 : v + he_1 = (5,3)^T, f(v + he_1) = 180 > f_b = 141.$

$+x_1 : v - he_1 = (3,3)^T, f(v - he_1) = 108 < f_b = 141.$

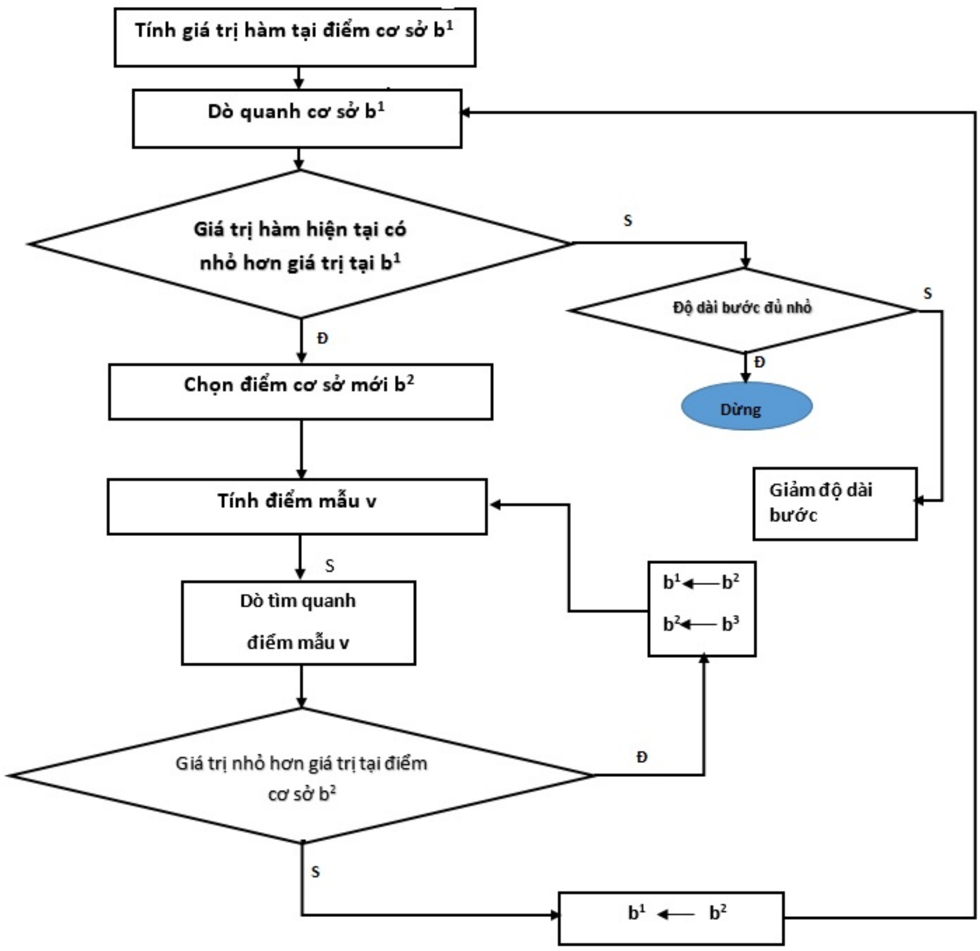
+ Đặt $v := v - he_1 = (3,3)^T$ đây là giá trị tốt nhất mới, $f_b = f(v) = 108$.

- Xét tọa độ x_2

$+x_2 : v + he_2 = (3,4)^T, f(v + he_2) = 155 > f_b = 108.$

$+x_2 : v - he_1 = (3,2)^T, f(v - he_2) = 71 < f_b = 108.$

+ Đặt $b^{(2)} := v - he_2 = (3,2)^T \neq b^{(1)} = (4,3)^T$, giá trị ban đầu mới $f_b = f(b^{(2)}) = 71$.



Hình 3.24: **Sơ đồ dò tìm theo mẫu**

C. Dò tìm theo mẫu từ $b^{(2)}$:

- Điểm mẫu $v = 2b^{(2)} - b^{(1)} = (2,1)^T, f(v) = 25$. Giá trị ban đầu mới là $f_b = f(v) = 25$.

- Xét tọa độ x_1

$+x_1 : v + he_1 = (3,1)^T, f(v + he_1) = 44 > f_b = 25$.

$+x_1 : v - he_1 = (1,1)^T, f(v - he_1) = 12 < f_b = 25$.

+ Đặt $v := v - he_1 = (1,1)^T$ đây là giá trị tốt nhất mới, $f_b = f(v) = 12$.

- Xét tọa độ x_2

$+x_2 : v + he_2 = (1,2)^T, f(v + he_2) = 31 > f_b = 12$.

$+x_2 : v - he_1 = (1,0)^T, f(v - he_2) = 3 < f_b = 12$.

(Các tọa độ đã xét xong.)

+ Đặt $b^{(3)} := v - he_2 = (1,0)^T \neq b^{(2)} = (3,2)^T$, giá trị ban đầu mới $f_b = f(b^{(3)}) = 3$.

C. Dò tìm theo mẫu (tiếp)

- Gán $b^{(1)} = b^{(2)} = (3,2)^T, b^{(3)} = b^{(3)} = (1,0)^T$ và giá trị ban đầu $f_b = 3$.

- Điểm mẫu $v = 2b^{(2)} - b^{(1)} = (-1,-2)^T, f(v) = 31 > f_b = 3$.

- Xét tọa độ x_1

$+x_1 : v + he_1 = (0,-2)^T, f(v + he_1) = 20 > f_b = 3$.

$+x_1 : v - he_1 = (-2,-2)^T, f(v - he_1) = 48 > f_b = 3$.

- Xét tọa độ x_2

$+x_2 : v + he_2 = (-1,-1)^T, f(v + he_2) = 12 > f_b = 3$.

$+x_2 : v - he_2 = (-1,-3)^T, f(v - he_2) = 60 > f_b = 3$.

Dừng tìm kiếm theo mẫu từ $b^{(2)} = (1,0)^T$. Đặt $b^{(1)} = b^{(2)} = (1,0)^T$ và thực hiện dò tìm quanh điểm cơ sở $b^{(1)}$ (tức thực hiện bước B)

B. Dò tìm quanh điểm cơ sở $b^{(1)} = (1,0)^T$ **và** $f_b = 3$. Đặt $v = b^{(1)} = (0,1)^T$ - Xét tọa độ x_1

$+x_1 : v + he_1 = (2,0)^T, f(v + he_1) = 12 > f_b = 3$.

$+x_1 : v - he_1 = (0,0)^T, f(v - he_1) = 0 < f_b = 3$.

+ Đặt $v = v - he_1 = (0,0)^T$. Đạt giá trị ban đầu mới, đặt $f_b = 0$.

- Xét tọa độ x_2

$+x_2 : v + he_2 = (0,1)^T, f(v + he_2) = 5 > f_b = 0$.

$+x_2 : v - he_2 = (0, -1)^T, f(v - he_2) = 5 > f_b = 0.$

+ Đặt $b^{(2)} := v = (0, 0)^T \neq b^{(1)} = (1, 0)^T$, giá trị ban đầu $f_b = 3$.

C. Dò tìm theo mẫu với $b^{(1)} = (1, 0)^T, b^{(2)} = (0, 0)^T, f_b = 0$

- Điểm mẫu $v = 2b^{(2)} - b^{(1)} = (-1, 0)^T, f(v) = 3 > f_b = 0$.

- Xét tọa độ x_1

$+x_1 : v + he_1 = (0, 0)^T, f(v + he_1) = 0 = f_b = 0.$

$+x_1 : v - he_1 = (-2, -0)^T, f(v - he_1) = 12 > f_b = 0.$

- Xét tọa độ x_2

$+x_2 : v + he_2 = (-1, 1)^T, f(v + he_2) = 4 > f_b = 0.$

$+x_2 : v - he_2 = (-1, -1)^T, f(v - he_2) = 12 > f_b = 0.$

Dừng tìm kiếm theo mẫu từ $b^{(2)} = (0, 0)^T$ và dò tìm quanh điểm cơ sở cũ $b^{(2)}$, vì thế ta đặt $b^{(1)} := b^{(2)} = (0, 0)^T$.

+ Do không tìm được điểm tốt hơn nên tiếp tục dò tìm quanh điểm $b^{(1)}$ với độ dài bước h bé hơn 10 lần.

+Lặp lại một số lần dò tìm quanh $b^{(1)}$ vói h nhỏ hơn 10 lần sau mỗi lần do tìm ta sẽ nhận được giá trị cực tiểu của hàm là $f_b = 0$ và điểm cực tiểu là $(0, 0)^T$.

Nhận xét 3.4.10. *Cũng như phương pháp tìm kiếm ngẫu nhiên, phương pháp này không phải lúc nào cũng cho ta lời giải địa phương và cũng không xác định được tốc độ hội tụ của thuật toán.*

3.4.4 Phương pháp tìm kiếm trực tiếp cho bài toán tìm max

Gần giống như phương pháp trên thực chất là ở mỗi bước chỉ biến đổi một thành phần x_i của x còn các thành phần khác giữ nguyên cho đến khi nhận được gía trị cực đại. Nội dung phưng pháp:

A. Thăm dò bước 1

1. Chọn giá trị đầu $x^{(0)} = (x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)})^T$ có số gia $\Delta x = (\Delta x_1, \ldots, \Delta x_n)^T$

2. Tính $f(x^{(0)})$,

3. Cho x_1 biến đổi còn các thành phần khác giữ nguyên $x_1^{(1)} = x_1^{(0)} + \Delta x_1$. Tính giá trị $\alpha = f(x_1^{(1)}, x_1^{(0)}, \ldots, x_n^{(0)})$.
   - Nếu $\alpha > f(x^{(0)})$ lấy $x_1^{(1)} = x_1^{(0)} - \Delta x_1^{(0)}$. Tính α
   - Nếu không cải tiến được ở cả 2 phía thì cố định $x_1^{(0)}$ không cho biến đổi nữa
4. Tiếp tục tiến hành với $x_2^{(1)} = x_2^{(0)}(+/-)\Delta x_2^{(0)}$ cho đến khi tất cả được biến đổi. Hàm mục tiêu sẽ được thay lại nếu hàm mục tiêu tốt hơn và giữ nguyên nếu không cải tiến được.

B. Tìm theo mẫu Kết thúc thăm dò bước A. ta tìm được $x^{(k)}$. Bước này ta thực hiện như sau:

1. Ta lấy $x^{(k+1)} = tx^{(k)} - x^{(b)}$. Ở lần gặp đầu $x^{(b)} = x^{(0)}$ gọi là điểm cơ sở. t là số biến cần thăm dò 2 biến thì $t = 2$.

A. Thăm dò bước 2

1. Xuất phát từ $x^{(k+1)}$, tính $f(x^{(k+1)})$ và so sánh với $f(x)$ ở bước tìm theo mẫu để thăm dò bước 2 có kết qủa không.
2. Nếu thăm dò bước 2 có kết quả $x^{(k+2)}$ thì tìm theo mẫu được coi là có kết quả nếu $f(x^{(k+2)}) \geq f(x^{(k)})$ và quả trình lặp lại với điểm xuất phát $x^{(b)} := x^{(k)}$.
3. Nếu thăm dò bước 1 liên tiếp không cho hướng mới hiệu quả thì liên tiếp giảm Δx cho tới khi hoặc đã xác định được hướng mới hiệu quả hoặc Δx quá nhỏ.

Việc không có khả năng tăng $f(x)$ khi Δx khá bé có nghĩa là tối ưu địa phương đã đạt được

Cũng như phương pháp tìm kiếm ngẫu nhiên, phương pháp này không phải lúc nào cũng cho ta lời giải địa phương và cũng không xác định được tốc độ hội tụ của thuật toán.

Ví dụ 3.4.30. *Tìm*

$$f(x) := \frac{1}{(x_1+1)^2 + x_2^2} \to \max$$

1: Thăm dò bước 1:

- Chọn

$$x^{(0)} = (2.0, 2.8)^T, \Delta x := (0.6, 0.84)^T$$

- Tính $f(x^{(0)}) = 0.059$
  $x_1^{(1)} = 2.0 + 0.6 = 2.6; f(2.6, 2.8) = 0.048$ (không khả thi)
  $x_1^{(1)} = 2.0 - 0.6 = 1.4; f(1.4, 2.8) = 0.073$ (khả thi)
  $x_2^{(1)} = 2.8 + 0.84 = 3.64; f(1.4, 3.64) = 0.052$ (TB)
  $x_2^{(1)} = 2.8 - 0.84 = 1.96; f(1.4, 1.96) = 0.104$ (KQ)

Chọn thăm dò bước 1 kết quả:

$$x^{(1)} = (1.4, 1.96)^T$$

2: Tìm theo mẫu:

- $x_i^{(k+1)} = 2x_i^{(k)} - x_i^{(b)}$, với $x^{(b)} = x^{(0)} = (2.0, 2.8)^T$ nên

- $x_1^{(2)} = 2 * 1.4 - 2 = 0.8; x_2^{(2)} = 2 * 1.96 - 2.8 = 1.12$

- $x^{(2)} = (0.8, 1.12)^T; f(x^{(2)}) = 0.22$

3: Thăm dò bước 2:

- Tính
  $x_1^{(3)} = 0.8 + 0.6 = 1.4; f(1.4, 1.12) = 0.14$ (TB)
  $x_1^{(3)} = 0.8 - 0.6 = 0.2; f(0.2, 1.12) = 0.38$ (KQ)
  $x_2^{(3)} = 1.12 + 0.84 = 1.96; 3f(0.2, 1.96) = 0.19$ (TB)
  $x_2^{(3)} = 1.12 - 0.84 = 0.28; f(0.2, 0.28) = 0.67$ (KQT)

- Chọn $x^{(3)} = (0.2, 0.28)^T$

- So sánh $f(x^{(3)}) = f(0.2, 0.28) = 0.67 > 0.104 = f(x^{(1)})$ nên tìm theo mẫu có kết quả

Suy ra, điểm cơ sở là

$$x^{(b)} = x^{(1)} = (1.4, 1.96)^T.$$

Điểm $x^{(3)}$ là kết quả thăm dò bước 2, nên chỉ việc tìm theo mẫu xuất phát từ $x^{(3)}$.

4: Tìm theo mẫu xuất phát từ $x^{(3)}$:

- $x_1^{(4)} = 2 * 0.2 - 1.4 = -1.00; x_2^{(4)} = 2 * 0.28 - 1.96 = -1.4$
- Chọn $f(x^{(4)}) = 0.51$

5: Thăm dò bước 2: So sánh với $f(x^{(4)})$ Tính

- $x_1^{(5)} = -1.0 + 0.6 = -0.4; f(-0.4, -1.4) = 0.43$ (TB)
  $x_1^{(5)} = -1.0 - -0.6 = -1.6; f(-1.6, 1.4) = 0.43$ (TB)
  $x_2^{(5)} = -1.4 + 0.84 = -0.56; f(-1.0, -0.56) = 3.18$ (KQT)

- Chọn $x^{(5)} = (-1.0, -0.56)^T$
- So sánh $f(x^{(5)}) = 3.18 > 0.60 = f(x^{(3)})$ nên $x^{(5)}$ được coi là kết quả thăm dò bước 2.

6: Tìm theo mẫu xuất phát từ $x^{(5)}; x^{(b)} := x^{(3)}$

- $x_1^{(6)} = 2 * (-1.0) - 0.2 = -2.2;$
- $x_2^{(6)} = 2 * (-0.56) - 0.28 = -1.4$
- Do đó $x^{(6)} = (-2.2, -1.4)^T; f(x^{(6)}) = 0.29$

7: Thăm dò bước 2 với $f(x^{(6)})$

- $x_1^{(7)} = -2.2 + 0.6 = -1.6; f(-1.6, -1.4) = 0.43(KQ)$

- $x_2^{(7)} = -1.4 + 0.84 = -0.56; f(- - 1.0, -0.56) = 1.49(KQT)$

Do $f(x^{(7)}) < f(x^{(5)})$ nên dù dò tìm có kết quả nhưng tìm theo mẫu thất bại. Ta phải xuất phát từ $x^{(5)}$ thực hiện thăm dò bước 1 cho đến khi không có tìm kiếm nào có hiệu quả. Trong thí dụ này bằng tính toán ta có $x^* = (-1.0, 0.0)^T$ và $f(x^*) = \infty$.

3.4.5 Phương pháp Powell

Xét bài toán

$$f(x) \to \min, \ x \in I\!R^n.$$

khi f được cho là khả vi liên tục.

Kết hợp giữa tối ưu một tham số và tìm theo hướng gradient liên hợp ta có phương pháp Powell sau: (Chú ý rằng chỉ số dưới trong các ký hiệu dưới đây của thuật toán Powell là các vec tơ chứ không phải là các thành phần của một véc tơ. Ví dụ $x_i^{(k)}$ là véc tơ thứ i ở bước lặp thứ k.)

B1: Đặt $k := 0$, chọn điểm xuất phát $x_{\mathbf{0}}^{(k)}$ (véc tơ đầu tiên ở bước lặp $k = 0$), $s_i^{(k)}$ chọn song song với các véc tơ cơ sở e^i. Tính các véc tơ theo $x_i^{(k)}$ theo $s_i^{(k)}, i = 1, \dots, n$.

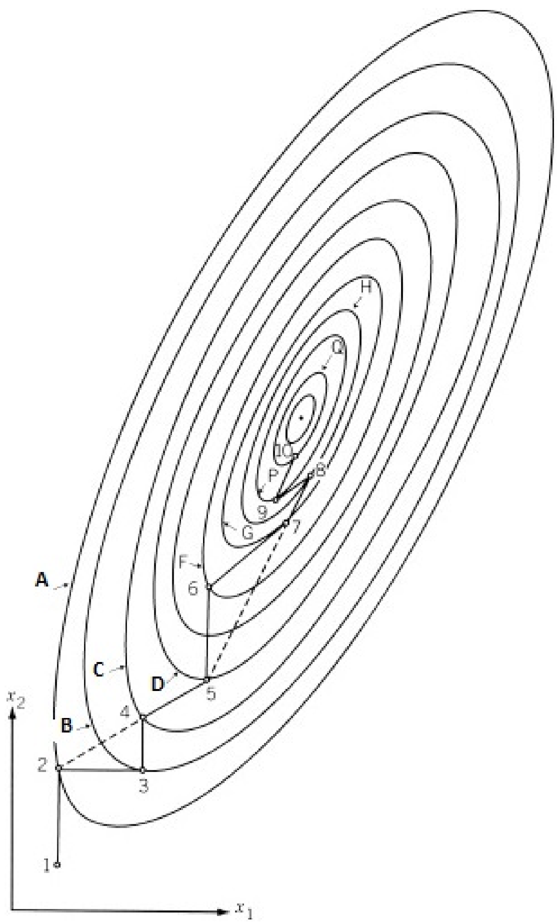
B2:
- Tìm $\min f(x_{i-1}^{(k)} + \lambda_i^{(k)} s_i^{(k)})$ theo một tham số xác định λ_i.
- Đặt $x_{i+1}^{(k)} := x_i^{(k)} + \lambda_i^{(k)} s_i^{(k)}$, nếu $i < n$ quay lại B2, nếu không chuyển B3.

B3: Tính $x_{n+1}^{(k)} := 2x_n^{(k)} + x_0^{(k)}$, chuyển B4.

B4: Tính $\max |f(x_{n+1}^{(k)})| = \Delta_m^{(k)}, i = 1, \dots, n$. Ký hiệu $s_m^{(k)}$ ứng với $\Delta_m^{(k)}$.

B5: Kiểm tra nếu

1. $f(x_{n+1}^{(k)}) \geq f(x_0^{(k)})$ hoặc



Hình 3.25: Minh họa thực hiện quá trình Powell cho hàm hai biến

2. $[f(x_0^{(k)}) - 2f(x_n^{(k)}) + f(x_{n+1}^{(k)})][f(x_n^{(k)}) - \Delta_m^{(k)}]^2 > 0.5\Delta_m^{(k)}[f(x_0^{(k)}) - f(x_{n+1}^{(k)})]^2$,

nếu đúng chuyển qua B6 sai chuyển qua B7.

B6: Đặt $s_i^{(k+1)} := s_i^{(k)}$, kiểm tra $f(x_n^{(k)}) \geq f(x_{n+1}^{(k)})$?, nếu đúng đặt $x_0^{(k+1)} := x_{n+1}^{(k)}$ nếu sai đặt $x_0^{(k+1)} := x_n^{(k)}$, chuyển qua B9.

B7: Xác định min $f(x^{(k)})$ theo hướng $s^{(k)}$ từ $x_{n+1}^{(k)}$ để x_0 ứng với x. Chuyển qua B8.

B8: Thay $s_m^{(k)}$ bởi $s^{(k)}$ và $s_i^{(k+1)} := s_i^{(k)}, i = 1, \ldots, n, i \neq m$. Chuyển qua B9.

B9: Nếu $f(x^{(k+1)}) < 10^{-10}$ và $(x_i^{(k+1)} - x_i^{(k)})/x_i^{(k)} < \epsilon, i = 1, \ldots, n$ chuyển về B10, nếu không tăng k, chuyển về B2.

B10: Dừng thủ tục.

Để thuận tiện cho việc tính toán **phương pháp Powell** được sửa đổi lại như sau:

Cho $x^{(0)}$ là véc tơ lựa chọn ban đầu của hàm

$$z = f(x_1, x_2, \ldots, x_n).$$

Giả sử $e_k = e_k = (0, \ldots, 1_k, \ldots)$ for $k = 1, 2, \ldots, n$ là cơ sở đơn vị trong $I\!R^n$. Đặt $S_k = E_k^T$ for $k = 1, 2, \ldots, n$ là véc tơ chuyển vị của e_k.

Đặt biến đếm $i = 0$.

(i) Cho $P_0 = x^{(i)}$

(ii) Với $k = 1, 2, \ldots, n$ tìm đại lượng $\lambda = \lambda_k$ sao cho $f(P_{k-1} + \lambda S_k)$ đạt min và gọi and set $P_k = P_{k-1} + \lambda_k S_k$.

(iii) Đặt $S_j := S_{j+1}$ for $j = 1, 2, \ldots, n-1$ và đặt $S_n = P_n - P_0$.

(iv) Tăng biến đếm $i := i + 1$.

(v) Tìm đại lượng $\lambda = \lambda_{\min}$ sao cho $f(P_0 + \lambda S_n)$ đạt min, và đặt $x^{(i)} = P_0 + \lambda_{\min} S_n$

(vi) Lặp lại bước từ (i) đến (v) cho đến khi tính hội tụ đạt được.

Ví dụ Tính toán bằng tay phương pháp Powell:

$$f(x) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2 \to \min .$$

1: Dò tìm bước 1 (thăm dò theo tọa độ): - Điểm xuất phát $x^{(1)} = (0, 0)^T$.

- Chọn hướng từ điểm xuất phát $s^{(2)} = (0, 1)^T$ và hệ số $\epsilon = 0.01$.

- Xác định hướng đúng $f_1 = f(x^{(1)}) = 0.0, f^+ = f(x^{(1)} + \epsilon s^{(2)}) = f(0, 0.01) = -0.0099 < f_1$ suy ra hàm giảm theo hướng $s^{(2)}$.

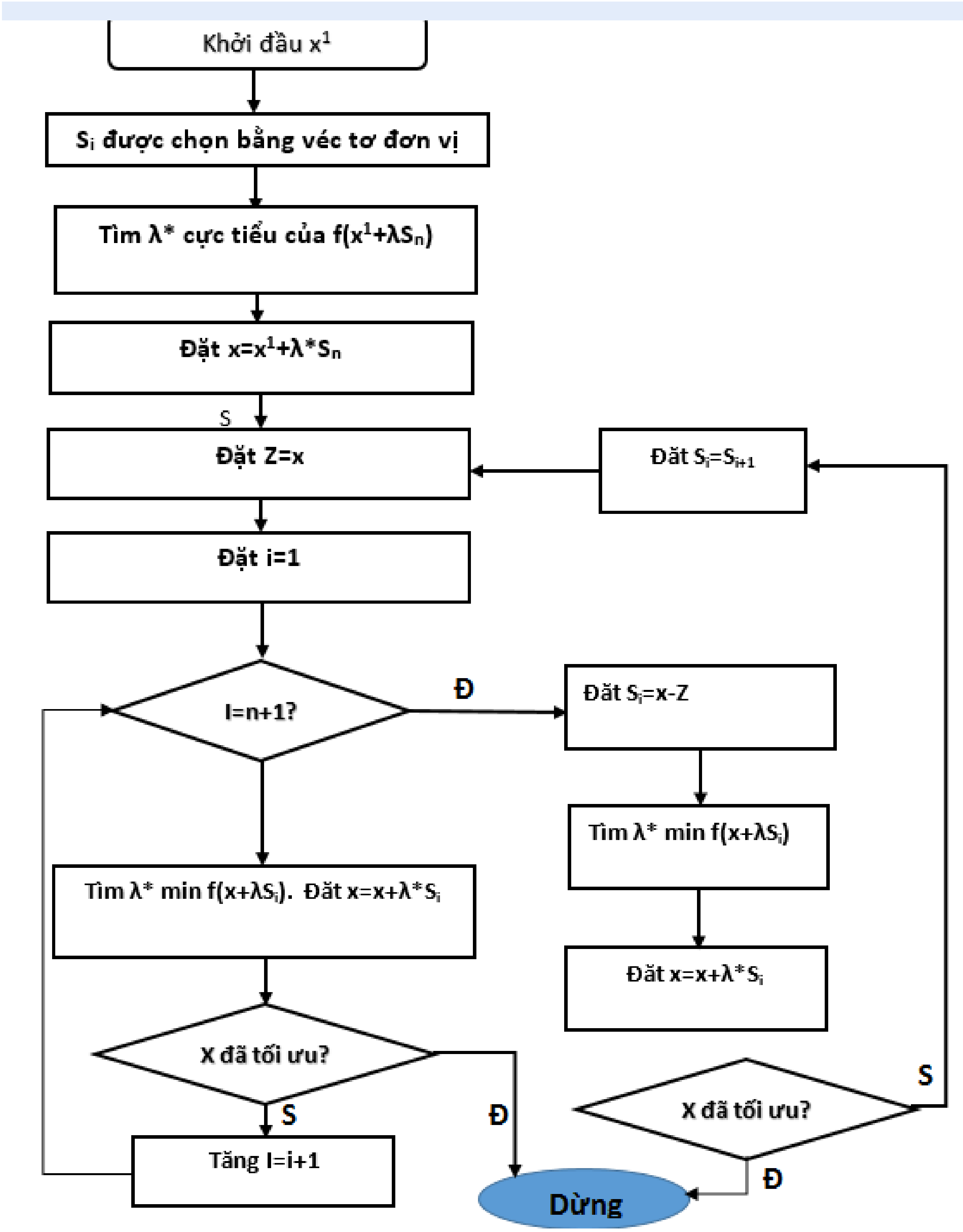
- Tìm độ dài tối ưu theo hướng $s^{(2)}$:

$$f(x^{(1)} + \lambda s^{(2)}) = f(0, 0.\lambda) = \lambda^2 - \lambda,$$

suy ra $\lambda^* = 1/2$.

- Chọn $x^{(2)} = x^{(1)} + \lambda^* s^{(2)} = (0, 0.5)^T$.

- Từ $x^{(2)}$ tìm min của f dọc theo hướng $s^{(1)} = (1, 0)^T$, từ $x^{(2)}$, ta



Hình 3.26: **Sơ đồ khối thuật toán Powell**

được $f_2 = f(x^{(2)}) = -0.25, f^+ = f(x^{(2)} + \epsilon s^{(1)}) = -0.2298 > f_2$ và $f^- = f(x^{(2)} - \epsilon s^{(1)}) = -0.2698 > f_2$. Do đó hàm giảm theo hướng $-s^{(1)}$.
- $f(x^{(2)} - \lambda s^{(1)}) = 2\lambda^2 - 2\lambda - 0.25$ nên $\lambda^* = 1/2$ do đó $x^{(3)} := x^{(2)} - \lambda^* s_1 = (-0.5, 0.5)^T$.
- Tiếp tục tìm min của hàm dọc theo hướng $s^{(2)}$ từ $x^{(3)} = (-0.5, 0.5)^T$, $f_3 = f(x^{(3)}) = -0.75, f^+ = f(x^{(3)} + \epsilon s^{(2)}) = -0.7599 < f_3$. Do đó f giảm theo hướng $s^{(2)}$. Từ đây

$$f(x^{(3)} + \lambda s^{(2)}) = \lambda^2 - \lambda - 0.75$$

dạt min tại $\lambda^* = 1/2$.
- Đặt $x^{(4)} = x^{(3)} + \lambda^* s^{(2)} = (-0.5, 1)^T$
2: Dò tìm bước 2 (dò tìm theo mẫu): - Đặt $s^{(3)} = x^{(4)} - x^{(2)} = (-0.5, 0.5)^T$, tìm min của f dọc $s^{(4)}$ từ $x^{(4)}$ ta thấy $f(x^{(4)}) = -1.00, f^+ = f(x^{(4)} + \epsilon s^{(3)}) = -1.004975$. Do đó hàm giảm theo hướng $s^{(3)}$. - Giá trị min của hàm $f(x^{(4)} + \lambda s^{(3)})$ ứng với $\lambda^* = 1$.
- $x^{(5)} = x^{(4)} + \lambda^* s^{(3)} = (-1, 1.5)^T$.
-$f_5 = f(x^{(5)}) = -1.25, f_+ = f(x^{(5)} + \epsilon s^2) > f_5$ và $f_- > f_5$. Do đó hàm không thể giảm theo cả 2 hướng, nên $x^{(5)}$ là phương án tối ưu.

3.4.6 Phương pháp Nelder - Mead

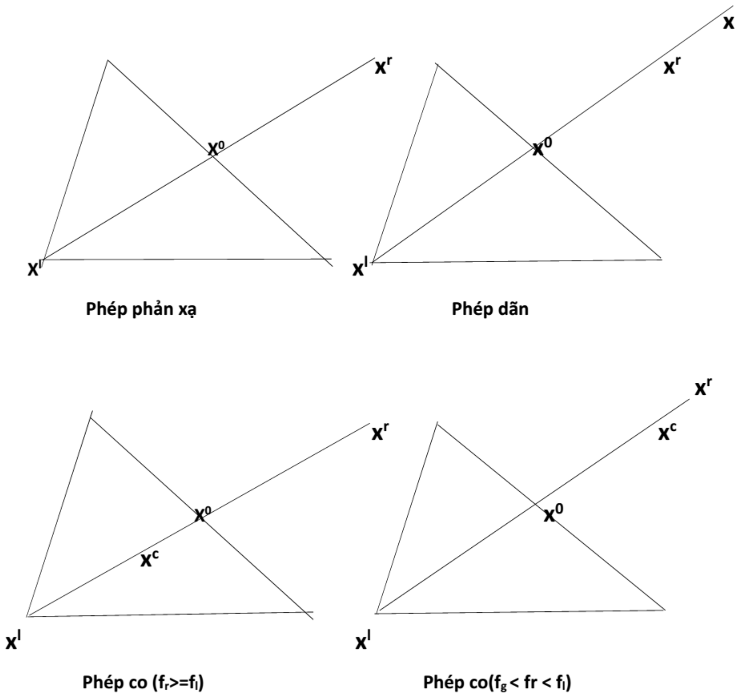
Nelder và Mead đã dùng đơn hình là những đa diện biến dạng nhờ 3 phép biến đổi ***ánh xạ gương, phép co và phép dãn***. Thuật toán bao gồm các bước sau:

B1: Tính giá trị hàm mục tiêu ở các đỉnh của đơn hình $f(x^1), f(x^2), \ldots, f(x^{n+1})$

B2: Xác định giá trị lớn nhất f_l giá trị tiếp theo f_g và giá trị nhỏ nhất f_b các đỉnh tương ứng sẽ là x^l, x^g, x^b.

B3: Xác định trọng tâm hình không kể đến đỉnh x_l theo công thức

$$x_0 = \frac{1}{n} \sum_{i \neq l} x_i \text{ và } f(x_0)$$



Hình 3.27: **Các phép ánh xạ gương-phép co-phép dãn**

B4: Thực hiện ánh xạ gương x^l qua x^0 nhận được x^r, tính $f_r = f(x^r)$. Nếu hệ số ánh xạ $\alpha > 0$ thì vị trí x^r được xác định theo công thức:

$$x^r = x^0 + \alpha(x^0 - x^l), x^r = (1+\alpha)x^0 - \alpha x^l \text{ và } \alpha = |x^r - x^0|/|x^0 - x^l|$$

B5: So sánh f_r và f_b xảy ra:

1. $f_r < f_b$ thì ta nhận được gí trị nhỏ nhất của hàm. Hướng từ f_0 đến f_r là hướng dịch chuyển tốt nhất cho dịch chuyển. Kéo dãn theo hướng này để tìm điểm x^p (hình b)). Hệ số dãn $\gamma > 1$ được tính từ:

   $$x^p - x_0 = \gamma(x^r - x^0), x^p = \gamma x^r + (1-\gamma)x^0, \gamma = |x^p - x^0|/|x^r - x^0|.$$

   Tính $f_p := f(x^p)$.

   - Nếu $f_p < f_b$ ta thay x^l bởi x^p và kiểm tra điểm thứ $(n+1)$ của đơn hình về tính hội tụ đến cực tiểu. Nếu hội tụ thì dừng, nếu không quay lại B2:

- Nếu $f_p > f_b$ thì bỏ điểm x^p vì đã đi quá xa. Thay x^l bởi x^r và kiểm tra tính hội tụ, nếu vẫn chưa thì quay lại B2:

2. Nếu $f_r > f_b$ nhưng $f_r < f_g$ thì x^r tốt hơn so với 2 điểm còn lại của đơn hình. Thay x^l bằng x^r, kiểm tra nếu chưa hội tụ thì quay lại B2:

3. Nếu $f_r > f_b$ và $f_r > f_g$ thì chuyển qua B6:

B6: So sánh các giá trị f_r và f_l

1. - Nếu $f_r > f_l$ thì chuyển về 2 của bước B6:
   - Nếu $f_r < f_l$ thay x^l bởi x^r và thay f_l bởi f_r
   - Nếu $f_r > f_g$ thực hiện bước co.

2. Thực hiện phép co:
   - Nếu $f_r > f_l$ ta đã dịch quá xa theo hướng từ x^l đến x^0 nên cần sửa lại phép co để tìm x^c (hình c): $x^c = x^0 + \beta(x^l) - x^0$ với $0 < \beta < 1$ là hệ số co
   - Nếu $f_r < f_l$ thay f_l bằng f_r va $x^l := x^r$ sau đó thực hiện phép co và tìm x^c (hình d) theo công thức sau: $x^c = x^0 + \beta(x^r - x^0) = \beta x^r + (1-\beta)x^0)$.

B7: Tính f_c. So sánh các giá trị f_c và f_l.

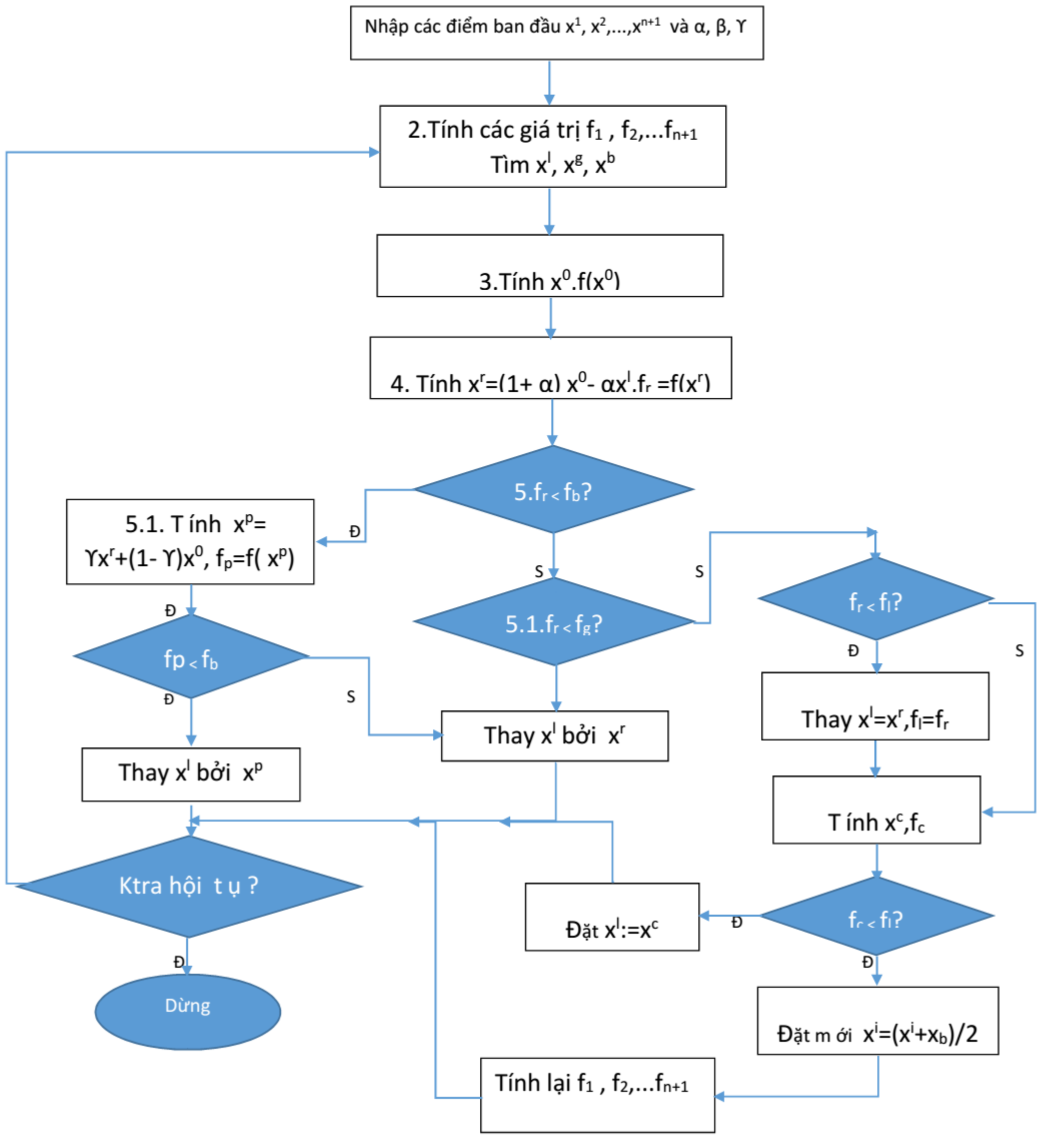
- Nếu $f_c < f_l$ thì thay x^l bởi x^c, nếu chưa hội tụ thì quay về B2:
- Nếu $f_c > f_l$ thì tìm giá trị nhỏ hơn f_l không được, phải chuyển qua B8:

B8: - Thu nhỏ kích thước đơn hình còn một nửa, lấy x^b làm chuẩn. Tức là x^i được thay bằng

$$x^i := x^i - \frac{1}{2}(x^i - x^b) = \frac{1}{2}(x^i - x^b).$$

- Tính các $f_i, i = 1, \ldots, n+1$. Kiểm tra tính hội tụ, nếu không quay lại B2:

B9: Tính $\sigma^2 = \sum_{i=1}^{n+1}(f_i - f)^2/(n+1), f = f_i/(n+1)$ Nếu σ nhỏ hơn độ chính xác ϵ nào đó thì giá trị của các hàm số rất gần nhau vì vậy điểm min rất gần với x^b.



Sơ đồ khối phương pháp Nelder - Mead

Ví dụ 3.4.31.

$$f(x) = 4x_1^2 + x_2^2 - 40x_1 - 12x_2 + 136 = 4(x_1 - 5)^2 + (x_2 - 6)^2 \to \min$$

1. Chọn đơn hình gồm 3 đỉnh $x^1 = (8,9)^T; x^2 = (10,11)^T; x^3 = (8,11)^T$.
2. Bước lặp:

k=0: -

$$f(8,9) = 45 \text{ nhỏ nhất } ; f(10,11) = 125 \text{ lớn nhất } ; f(8,11) = 65.$$

Do đó

$$x^b = (8,9)^T; x^g = (8,11)^T; x^l = (10,11)^T.$$

- Xác định trọng tâm: $x^0 = ((8+8)/2, (9+11)/2) = (8,10)^T$ - Ánh xạ qua $x^0 : x^r = (8+1(8-10), 10+1(10-11))^T = (6,9)^T$
Do $x^r = (6,9)^T, f(x^r) = 13$. Vì $f(x^r) = 13 < 45 = f(x^b)$ nên ta thực hiện phép dãn:
-(Phép dãn:) $x^p = (8+2(6-8), 10+2(9-10))^T = (4,8)^T \dots$

Ví dụ 3.4.32. *Cực tiểu hàm 2 biến sau:*

$$f(x_1, x_2) = x_1^2 + 2x_1x_2 + 3x_2^2$$

Lấy $\alpha = 1, \beta = 0.5, \gamma = 2$.

Vòng 1: - Các đỉnh đơn hình $x^b = (2,1)^T; x^g = (3,1)^T; x^l = (2,2)^T$
- Giá trị hàm $f_b = 11; f_g = 18; f_l = 24$.
- Điểm trọng tâm: $x^0 = (x^b + x^g)/2 = 2.5, 1)^T$
-Phép phản xạ:

$$x^r = 2x^0 - x^l = (3,0)^T; f_r = 9 < f_b$$

- Phép dãn:

$$x^c = 2x^r - x^0 = (3.5, -1)^T; f(x^c) = 8.25 < f_b = 11$$

thay x^h bởi x^e
-$\sigma = 16.8472222 > \epsilon$. Thực hiện bước sau.

Vòng 2: - Các đỉnh đơn hình $x^b = (3.5, -1)^T; x^g = (2,1)^T; x^l = (3,1)^T$
- Giá trị hàm $f_b = 8.25; f_g = 11; f_l = 18$.
- Điểm trọng tâm: $x^0 = (x^b + x^g)/2 = 2.75, 0)^T$
-Phép phản xạ:

$$x^r = 2x^0 - x^l = (2.5, -1)^T; f_r = 4.25 < f_b = 8.25$$

- Phép dãn:

$$x^c = 2x^r - x^0 = (2.25, -2)^T; f_c = f(x^c) = 8.0625 < f_b = 8.25$$

thay x^l bởi x^c

-$\sigma = 1.8 = 29514 > \epsilon$. Thực hiện bước sau.

Vòng 3: - Các đỉnh đơn hình $x^b = (2.25, -2)^T; x^g = (3.5, -1)^T; x^l = (2, 1)^T$

- Giá trị hàm $f_b = 8.0625; f_g = 8.25; f_l = 11$.
- Điểm trọng tâm: $x^0 = (x^b + x^g)/2 = 2.875, -1.5)^T$
- Phép phản xạ:

$$x^r = 2x^0 - x^l = (3.75, -4)^T; f_r = 32.0625 > f_l = 11.$$

- Phép co:

$$x^c = (x^l + x^0)/2 = (2.4375, -0.25)^T; f_c = f(x^c) = 4.91016 < f_l = 11$$

thay x^l bởi x^c

-$\sigma = 2.3474426 > \epsilon$. Thực hiện bước sau.

Vòng 4: - Các đỉnh đơn hình

$$x^b = (2.4375, -0.25)^T; x^g = (2.25, -2)^T; x^l = (3.5, -1)^T$$

- Giá trị hàm $f_b = 4.91; f_g = 8.0625; f_l = 8.25$.
- Điểm trọng tâm: $x^0 = (x^b + x^g)/2 = (2.34375, -1.125)^T$

-Phép phản xạ:

$$x^r = 2x^0 - x^h = (1.1875, -1.25)^T; f_r = 3.1289. < f_b = 4.91.$$

- Phép dãn:

$$x^c = (2x^r - x^0) = (0.03125, -1.375)^T; f_c = f(x^c) = 5.5869 > f_l = 4.91$$

thay x^l bởi x^r

-$\sigma = 4.1611663 > \epsilon$. Thực hiện bước sau.

Vòng 5: - Các đỉnh đơn hình

$$x^b = (1.1875, -1.25)^T; x^g = (2.4375, -0.25)^T; x^l = (2.25, -2)^T.$$

- Giá trị hàm $f_b = 3.128906; f_g = 4.910156; f_l = 8.0625$.
- Điểm trọng tâm: $x^0 = (x^b + x^g)/2 = (1.8125, -0.75)^T$
-Phép phản xạ:

$$x^r = 2x^0 - x^l = (1.1875, -1.25)^T; f_b \leq f_r = 4.015625 < f_g$$

thay x^l bởi x^r
-$\sigma = 0.5288120 > \epsilon$. Thực hiện bước sau.

...

Vòng 17: - Các đỉnh đơn hình $x^b = (0.018651, -0.10235)^T; x^g = (-0.048686, 0.024738)^T; x^l = (0.043274, -0.010235)^T$
- Giá trị hàm $f_b = 0.00171813; f_g = 0.00179747; f_l = 0.00305140$.
-$\sigma = 4.10^{-7} < \epsilon$. Dừng.

3.4.7 Giải bài toán tối ưu phi tuyến không ràng buộc bằng MATHLAB

Tìm cực tiểu hàm sau:

$$f(x) = 100 * (x(2) - x(1) * x(1))^2 + (1 - x(1))^2$$

B1: Tạo M-file ứng với hàm cần tìm phương án tối ưu:
function f=objfunc(x)

$$100 * (x(2) - x(1) * x(1))^2 + (1 - x(1))^2$$

B2: Thực hiện:

- Nhập giá trị ban xuất phat $x0$. Tiến hành tính toán theo các lệnh sau:
$x0 = [-1.2; 1.0]$;

– Tinh giá trị $f(x0)$.
  f=objfunc(x0)
  options=optimset('Large Scale', 'off');
  $[x, fval]$ = fminc(@objfunc,x0,options);

Kết quả sẽ cho:
The values of function value at starting point
f=24.20
Optimization termined: relative infinity-norm of gradient less than options. TolFun
$x = 1.000\ 1.000$
fval=2.8336e-011.

3.5. CÁC PHƯƠNG PHÁP GIẢI BÀI TOÁN QUY HOẠCH PHI TUYẾN CÓ RÀNG BUỘC

Trong mục này ta nghiên cứu bài toán (3.56), tức là bài toán quy hoạch phi tuyến bị ràng buộc sau:

$$\begin{aligned} f(x) &\to \min, x \in D, \\ D &= \{x \in I\!\!R^n \mid f_i(x) \leq 0, h_j(x) = 0, i = 1, \ldots m, j = 1, \ldots, p\}, \end{aligned}$$

trong đó ít nhất một trong các hàm f, f_i, h_j là phi tuyến.
Một trong những phương pháp giải các bài toán quy hoạch phi tuyến có ràng buộc là đưa về bài toán quy hoạch phi tuyến không có ràng buộc, do đó hàm Lagrange vẫn được sử dụng ở đây. Hàm Lagrange trong trường hợp này có dạng:

$$L(x, \lambda, \beta) = f(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{p} \beta_j h_j(x),$$

trong đó $\lambda := (\lambda_1, \lambda_2, \ldots, \lambda_p), \beta := (\beta_1, \ldots, \lambda_p)$.

Nhắc lại và bổ sung điều kiện cần: (Định lý Kuhn-Tucker)

1. f, f_i, h_j khả vi

2. x^* là **điểm cực tiểu địa phương**

Khi đó tồn tại λ^*, β^* thỏa mãn:

1. $\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x^*) + \sum_{j=1}^{p} \beta_j h_j(x^*) = 0$
2. $\lambda_i f_i(x^*) = 0, i = 1, \ldots, m$
3. $\lambda_i \geq 0, i = 1, \ldots, m$

Ví dụ 3.5.33.

$$f(x) = x_1^2 + x_2^2 + x_3^2 + 40x_1 + 20x_2 \to \min$$

Các ràng buộc:

$$\begin{aligned} x_1 - 50 &\geq 0 \\ x_1 + x_2 - 100 &\geq 0 \\ x_1 + x_2 + x_3 - 150 &\geq 0. \end{aligned}$$

Đây là ví dụ về bài toán phi tuyến song do các hàm f, f_i, h_j tồn tại đạo hàm liên tục nên ta có thể sử dụng định lý Tuhn-Tuker. Chỉ khác ở đây do $f_i(c) \geq 0$ nên $\lambda_i \leq 0$. Ta có hệ các phương trình sau:

$$\begin{aligned} 2x_1 + 40 + \lambda_1 + \lambda_2 + \lambda_3 &= 0 \\ 2x_2 + 20 + \lambda_2 + \lambda_3 &= 0 \\ 2x_3 + \lambda_3 &= 0 \end{aligned}$$

$$\begin{aligned} \lambda_1(x_1 - 50) &= 0 \\ \lambda_2(x_1 + x_2 - 100) &= 0 \\ \lambda_3(x_1 + x_2 + x_3 - 150) &= 0 \end{aligned}$$

$$\begin{aligned} f_1(x) = (x_1 - 50) &\geq 0 \\ f_2(x) = (x_1 + x_2 - 100) &\geq 0 \\ f_3(x) = (x_1 + x_2 + x_3 - 150) &\geq 0 \end{aligned}$$

$$\begin{aligned}\lambda_1 &\leq 0\\ \lambda_2 &\leq 0\\ \lambda_3 &\leq 0\end{aligned}$$

Hệ phương trình trên có thể giải bằng phương pháp số để tìm điều kiện cần tối ưu. Ta cũng có thể biện luận để tìm lời giải bằng suy luận toán học. Cụ thể: 1. Nếu $\lambda_1 = 0$ trong trường hợp này không tìm được các giá trị còn lại $\lambda_2, \lambda_3, x_1, x_2, x_3$ thỏa mãn hệ phương trình.
2. Khi $x_1 = 50$ ta làn lượt giải các trường hợp còn lại và tìm được giá trị thỏa mãn là:

$$\lambda_1 = -20; \lambda_2 = -20; \lambda_3 = -100, x_1^* = x_2^* = x_3^* = 50.$$

3.5.1 Phương pháp nhân tử Lagrange tăng cường

Xét bài toán:

$$f(x) \to \min$$

Với các ràng buộc:

$$h_j(x) = 0; j = 1, 2 \ldots, m; m < n.$$

Hàm Lagrange sẽ là:

$$L(x, \lambda) := f(x) + \sum_{j=1}^{m} \lambda_j h_j(x),$$

với $\lambda_j, j = 1, 2, \ldots, m$ là hằng số Lagrange. Xây dựng hàm mới: $\overline{L}(x, \lambda_k, r_k) = f(x) + \sum_{j=1}^{m} \lambda_j h_j(x) + r_k \sum_{j=1}^{m} \lambda_j h_j^2(x)$. Thuật toán tìm phương án tối ưu của bài toán trên gồm các bước sau:

B1: Lấy giá trị đầu: $x^{(1)}, \lambda^{(1)}, r_1, c > 1, r_{\max}$,

B2: Đặt $k = 1$,

B3. Cực tiểu hàm : $\overline{L}(x, \lambda^{(k)}, r_k)$ với điều kiện đầu $x^{(k)}$ và tìm $x^{*(k)}$, B4: Kiểm tra tính hội tụ của $\lambda^{(k)}$ và $x^{*(k)}$, nếu đúng dừng $x* = x^{*(k)}$, nếu không chuyển bước 5,

B5. Đặt $\lambda^{(k+1)} = \lambda^{(k)} + 2r_k h_j(x^{*(k)}), j = 1, 2, \ldots, m,$

B6: Đặt $r_{k+1} = cr_k$,

B7: Nếu $r_{k+1} > r_{\max}$, đặt $r_{k+1} = r_{\max}$,

B8: Đặt $k = k + 1$, quay lại bước 2.

3.5.2 Phương pháp nhân tử Lagrange tăng cường cho bài toán ràng buộc bất đẳng thức

Bài toán được xét đến là:

$$f(x) \to \min$$

Với các ràng buộc:

$$g_j(x) \leq 0; j = 1, 2 \ldots, m.$$

Để sử dụng phương pháp nhân tử Lagrange tăng cường ta đưa thêm biến phụ để đưa về dạng đẳng thức:

$$g_j(x) + y_j^2 = 0, j = 1, 2, \ldots, m.$$

Hàm Lagrange tăng cường sẽ là: $\overline{L}(x, \lambda_k, r_k) = f(x) + \sum_{j=1}^{m} \lambda_j [g_j(x) + y_j^2] + r_k \sum_{j=1}^{m} [g_j(x) + y_j^2]^2(x)$. Thuật toán tìm phương án tối ưu của bài toán trên gồm các bước sau:

B1: Lấy giá trị đầu: $x^{(1)}, \lambda^{(1)}, r_1, c > 1, r_{\max}$,

B2: Đặt $k = 1$,

B3: Cực tiểu hàm : $\overline{L}(x, \lambda^{(k)}, r_k)$ với điều kiện đầu $x^{(k)}$ và tìm $x^{*(k)}$,

B4: Kiểm tra tính hội tụ của $\lambda^{(k)}$ và $x^{*(k)}$, nếu đúng dừng $x* = x^{*(k)}$, nếu không chuyển bước 5, B5: Đặt $\lambda^{(k+1)} = \lambda^{(k)} + 2r_k h_j(x^{*(k)}), j = 1, 2, \ldots, m,$

B6: Đặt $r_{k+1} = cr_k,$

B7: Nếu $r_{k+1} > r_{\max}$, đặt $r_{k+1} = r_{\max},$

B8: Đặt $k = k + 1$, quay lại bước 2.

3.5.3 Phương pháp gradient (chiếu)

Xét bài toán trên khi các hàm f, f_i, h_j là khả vi liên tục. Phương pháp gradient là dò tìm dọc theo hướng gradient hoặc đối gradient phương án tối ưu. Phương pháp gồm các bước sau:

B1: Chọn một điểm bất kỳ $x^{(0)}$ thuộc miền ràng buộc.

B2: Đi theo hướng đối gradient đối với bài toán min và đi theo hướng gradient nếu bài toán max .

B3: Chọn bước đi λ tối ưu

$$x^{(k+1)} = x^{(k)} - \lambda \bigtriangledown f(x^{(k)}),$$

trong đó λ là giá trị để $f(x^{(k+1)}) = f(x^{(k)} - \lambda \langle \bigtriangledown f(x^{(k)}))$ đạt giá trị bé nhất theo λ tức là theo Fermat $\langle \nabla f(x^{(k+1)}), \nabla f(x^{(k)}) \rangle = 0$.

B4: Kiểm tra x có nằm trong miền ràng buộc không?

- Nếu chạm vào biên phải kiểm tra điều kiện $\bigtriangledown f // \bigtriangledown f_i$.
- Quá trình kết thúc khi các véc tơ $\bigtriangledown f$ và $\sum_{j=1}^m \bigtriangledown f_j(x)$ song song, tức là $\langle \bigtriangledown f, \sum_{j=1}^m \bigtriangledown f_j(x) \rangle = 0$.

Ví dụ Xét bài toán

$$f(x) = -6x_1 - 4x_2 + x_1^2 + x_2^2 + 18 \to \min .$$

Với miền D giới hạn bởi:

$$D = \begin{cases} f_1(x) = 4x_1 + 3x_2 - 1 \leq 24 \\ f_2(x) = x_1^2 - 5x_1 + x_2^2 < 6 \\ x_1, x_2 \geq 0. \end{cases}$$

- Lấy bất kỳ $x^{(0)} = (1.5, 3)^T \in D$, tính $f(x^{(0)}) = 8.25$.
- Đối gradient $- \bigtriangledown f(x^{(0)}) = (3, -2) \neq 0$ nên $x^{(0)}$ không phải điểm cực trị.
- Lấy $x^{(1)} = x^{(0)} - \lambda \bigtriangledown f(x^{(0)}) = (1.5 + 3\lambda, 3 - 2\lambda)$ trong đó λ thỏa mãn

$$\frac{d\Delta f}{d\lambda} = \langle - \bigtriangledown f(x^{(1)}), - \bigtriangledown f(x^{(0)}) \rangle = 0 = 13 - 26\lambda = 0,$$

suy ra $\lambda = 0.5$.
- Vì $\frac{d^2\Delta f}{d\lambda^2} = -26 < 0$ nên với $\lambda = 0.5$ thì Δf đạt giá trị lớn nhất.
-Chọn $x^{(1)} = (1.5 + 3 * 0.5, 3 - 2 * 0.5)^T = (3, 2)^T$ và thử lại thấy $x^{(1)} \in D$.
- Tính $\bigtriangledown f(x^{(1)}) = (6 - 3 * 3, 4 - 2 * 2) = (0, 0)$ nghĩa là không còn cách di chuyển $x^{(1)}$ để f giảm được nữa, do đó $x^{(1)}$ là phương án tối ưu.

Ví dụ Xét bài toán

$$f(x) = 2x_1 + 3x_2 - 0.1x_1^2 - 0.1x_2^2 \to \max.$$

Với miền D giới hạn bởi:

$$D = \begin{cases} f_1(x) = 100 - 10x_1 - x_1^2 - 20x_2 - x_2^2 \geq 0 \\ f_2(x) = 120 - 20x_1 - x_1^2 - 10x_2 - x_2^2 \geq 0 \\ f_3(x) = 150 - 20x_1 - x_1^2 - 20x_2 - x_2^2 \geq 0. \end{cases}$$

Giải: 1. Chọn giá trị xuất phát: $x^{(0)} = (1, 3)^T; f(x^{(0)}) = 10$.
2. Tính gradient $\nabla f(x^{(0)}) = (1.8, 2.4)^T$.
3. Đặt $x^{(1)} = x^{(0)} + \lambda_1 \nabla f(x^{(0)}) = (1 + 1.8\lambda_1; 3 + 2 + 2.4\lambda_1)$.
4. Tính $df(x^{(1)})/d\lambda = \langle \nabla f(x^{(k)}), \nabla f(x^{(0)}) = 0$, tính ra $\lambda_1 = 0.5$.
5. $x^{(1)} = (1.9, 4.2)^T$. Do $f_1(x^{(1)}) = -24.5 < 0, f_2(x^{(1)}) > 0, f_3(x^{(1)}) > 0$, tức là ràng buộc đầu tiên bị phá vỡ tại $x^{(1)}$ nên ta tìm $\nabla f_1(x^{(1)})$.
6. Tính $\nabla f_1(x^{(1)}) = (-13.8, -24.8)^T$.
7. Đặt $x^{(2)} = x^{(1)} + \lambda_2(x^{(1)})$ nên $x^{(2)} = (1.9 - 13.8\lambda_2, 4.2 - 24.8\lambda_2)^T$. Lấy

$\lambda_2 = 0.05$ ta được $x^{(2)} = (1.21, 2.96)^T$.

8. Thay vào ta thấy $x^{(2)}$ nằm trong miền ràng buộc và $f(x^{(2)}) = 10.2774$.

9. So sánh $f(x^{(0)}) < f(x^{(2)})$ hai giá trị này gần nhau nên $x^{(2)}$ chuyển dịch gần đường mức của hàm $f(x)$.

10. Để tăng nhanh tốc độ hội tụ ta dịch chuyển theo hướng z từ $x^{(0)}$ qua $x^{(2)}$ tức là $z = x^{(2)} - x^{(0)} = (0.21, -0.04)^T$.

11. Lấy $x^{(3)} = x^{(2)} + \lambda_3 z = (1.21 + 0.21\lambda_3, 2.96 - 0.04\lambda_3)^T$. Tính cực đại của hàm $f(x^{(3)})$ theo λ_3 ta chọn được $\lambda_3 = 6.5$.

12 Vậy $x^{(3)} = (2.5750, 2.700)^T, f(x^{(3)}) = 11.8579$.

13. Kiểm tra ta thấy $f_1(x^{(3)}) > 0, f_2(x^{(3)}) > 0, f_3(x^{(3)}) > 0$ nên $x^{(3)}$ nằm trong miền ràng buộc.

14. Tính $\nabla f(x^{(3)}) = (1.4850, 2.4600), \nabla f_1(x^{(3)}) = (-15.15, -25.40)^T$. Hai véc tơ này gần song song. Do đó việc chuyển dọc theo chúng sẽ đi theo đường dích dắc gần biên và chạm biên.

15. Do vậy ta có thể lấy nghiệm $x^* = x^{(3)}$.

3.6. CÁC PHƯƠNG PHÁP HÀM PHẠT

3.6.1 Phương pháp hàm phạt điểm trong và ngoài

Xét bài toán tối ưu có ràng buộc

$$\begin{aligned} f(x) &\rightarrow \min, x \in D && (3.59)\\ D &= \{f_i(x) \leq 0, i = 1, \ldots, m,\} && (3.60)\end{aligned}$$

Mục đích của phương pháp hàm phạt là thay cho hàm nục tiêu f, ta sẽ giải bài toán tối ưu không ràng buộc với hàm mục tiêu $f + p$, trong đó $p(x)$ là lượng phạt khi x vi phạm các ràng buộc. Các phương pháp hàm phạt được chia thành hai nhóm sau:

(a) *Phương pháp hàm phạt điểm trong.*

(b) *Phương pháp hàm phạt điểm ngoài.*

- Với phương pháp hàm phạt điểm trong, hàm mục tiêu mới thường được xây dựng có dạng:

$$\phi(x, \alpha) := f(x) + \alpha[p(f_1(x)) + \cdots + p(f_m(x))],$$

trong đó $\alpha > 0$ là tham số phạt, p(y) là hàm một biến, đơn điệu thỏa mãn $p(y) \to +\infty$ khi $y \to -0$ và $p(y) > 0$ với mọi $y < 0$. Chẳng hạn $p = -1/f_i(x)$ hay $p = -ln(-f_i(x))$. Nội dung phương pháp được thể hiện qua các bước sau

B1: Xuất phát từ điểm chấp nhận được $x^{(1)} \in D$. Chọn tham số $\alpha_i > 0$. Đặt chỉ số bước lặp $k = 1$.

B2: Tìm cực tiểu hàm không ràng buộc: $\phi(x, \alpha_k) = f(x) - \alpha \sum_{i=1}^{m} \frac{1}{f_i(x)}$ (dùng kỹ thuật tìm cực tiểu không ràng buộc) và nhận được nghiệm cực tiểu $x^{(k)}$.

B3: Kiểm tra $x^{(k)}$ có phải là nghiệm tối ưu của bài toán gốc hay không, nếu đúng dừng, nếu không chuyển bước sau.

B4: Đặt tham số phạt mới $\alpha_{k+1} = \mu\alpha_k$ với $\mu < 1$ (chẳng hạn, $\mu = 10^{-1}$). Đặt k:=k+1 và chuyển về B2.

Định lý 3.6.32. *Phương pháp giải nêu trên có tính chất sau:*
(a) $\phi(x, \alpha) \geq f(x)$ *với mọi* $\alpha > 0,\ x \in D$.
(b) $f_i(x^{(k)}) < 0, i = 1, \ldots, m$.
(c) $\{\phi(x^{(k)}, \alpha_k)\}$ *hội tụ về giá trị tối ưu của bài toán* (2.2.26) *khi* $\{\alpha_k\}$ *dần tới* 0.

- Với phương pháp hàm phạt điểm ngoài hàm phạt được chọn là: $p = \max\{0, f_i(x)\}$ hoặc $p = \max\{0, -f_i(x)\}$.

Ví dụ 3.6.34.

$$f(x) = \frac{1}{3}(x_1 + 1)^3 + x_2 \to \min$$

Với các ràng buộc:

$$f_1(x) = -x_1 + 1 \leq 0$$
$$f_2(x) = -x_2 \leq 0$$

$$\phi(x,\alpha) := \frac{1}{3}(x_1+1)^3 + x_2 - \alpha\left(\frac{1}{-x_1+1} - \frac{1}{x_2}\right) \to \min$$

Áp dụng điều kiện cần ta có:

$$\frac{\partial\phi}{\partial x_1} = (x_1+1)^2 - \frac{\alpha}{(1-x_1)^2} = 0$$

$$\frac{\partial\phi}{\partial x_2} = \left(1 - \frac{\alpha}{x_2^2}\right) = 0$$

Từ các phương trình này ta nhận được:

$$x_1(\alpha) = (\alpha^{1/2}+1)^{\frac{1}{2}}; x_2(\alpha) = \alpha^{1/2}$$

$$\Phi_{\min}(\alpha) = \frac{1}{3}[(\alpha^{1/2}+1)^{1/2}+1]^3 + 2\alpha^{1/2} - \frac{1}{(1/\alpha) - (1/\alpha^{3/2} + 1/\alpha^2)^{1/2}}$$

Khi $\alpha \to 0$ thì $\Phi_{\min}(\alpha) \to f_{\min}; x_1(\alpha) = x_1^*; x_2(\alpha) = x_2^*$. Chọn dãy α dần đến 0 ta tính được các giá trị tương ứng. $x_1^* = 1, x_2^* = 0, f_{\min} = 8/3$.

3.6.2 Phương pháp Caroll

Bài toán chúng ta sẽ xét là (3.59) với ràng buộc là

$$f_j(x) \geq 0, j = 1, 2\ldots, m,$$

phương pháp Caroll là sử dụng hàm phạt trong để xét bài toán với hàm mục tiêu

$$\Phi(x,\alpha_k) := f(x) \pm \alpha_k \sum_{j=1}^{m} \frac{w_j}{f_j(x)}$$

với dấu + khi tìm min và dấu − khi tìmmax, α_k là nhân tử ở bước lặp thứ k, w_j là trọng số (thường chọn bằng 1). Nội dung thuật toán như sau:

B1: Chọn trọng số w_j, bước t_0, thông số α_0, điểm xuất phát $x^{(0)}$, sai số ϵ.

B2: $i := 0; \alpha_i := 0$.

B3: $k := 0$, thiết lập $\nabla\Phi(x^{(k)}, \alpha_k)$, tính $f(x^{(k)}), g(x(k)), \Phi(x^{(k)}, \alpha_k)$

B4: $k := k + 1; t_k := t_0$.

B5: $x^{(k)} := x^{(k-1)} + t_k \nabla\Phi(x^{(k-1)}, \alpha_{k-1})$

B6: Tính $f(x^{(k)}), g(x^{(k)}), \Phi(x^{(k)}, \alpha_k)$

B7: Kiểm tra $\Phi(x^{(k)}, \alpha_k) < \Phi(x^{(k-1)}, \alpha_{k-1})$?, nếu đúng chuyển B8, nếu sai chuyển B10:

B8: Kiểm tra $k - 1 = 0$? nếu đúng chuyển B9:, sai chuyển về B3:

B9: $i := i + 1; \alpha_i := \alpha_{i-1}/2$. Kiểm tra $\alpha_i < \epsilon$?, nếu đúng $x^*, f(x^*)$ dừng nếu không chuyển B3:

B10: Kiểm tra $f_j(x^{(k)}) > 0$?, nếu đúng $k := k + 1$ chuyển B5:, nếu sai đặt $t_k := t_k/2$, tăng $k := k + 1$, chuyển B5:

Thực ra phương pháp do Carroll đưa ra được phát biểu tóm lược như sau: Thay vào tìm min của hàm xuất phát ta dùng phương pháp hàm chắn tìm tối ưu hàm

$$\Phi(x, r_k) = f(x) + r_k \sum_{j=1}^{m} \frac{1}{g_j(x)}$$

1. Chọn giá trị xuất phát $x^{(1)}$ sao cho $g_j(x^{(1)}) > 0$, giá tri $r_1 > 0$ và cho $k = 1$.
2. Tìm min của hàm $\Phi(x, r_k)$ theo một phương pháp không ràng buộc nào đó để có x_k^*.
3. Kiểm tra xem nghiêm x_k^* đã gần là nghiệm của bài toán đã cho chưa. Nếu tìm được thì dừng, nếu không chuyển bước sau.
4. Xác định tham số phạt mới $r_{k+1} = cr_k, c < 1$.
5. Đặt k=k+1 và điểm xuất phát $x^{(1)} = x_k^*$ quay lại bướ 2.

3.6.3 Sử dụng MATHLAB để giải bài toán tối ưu phi tuyến có ràng buộc

Chương trình gôm những bước sau:
B1: Viết M-file chứa hàm mục tiêu (objfun.m, ví dụ hàm $f = x(1)^3 - 6 * x(1)^2 + 11 * x(1) + x(3)$):
function f= objfun (x)

$$f = x(1)^3 - 6 * x(1)^2 + 11 * x(1) + x(3);$$

B2: Viết M-file chứa các hàm ràng buộc M-file constraints.m (c là véc tơ cột các ràng buộc):
function [c, ceq] = constraints (x)

$c = [x(1)^2+x(2)^2-x(3)^2; 4-x(1)^2-x(2)^2-x(3)^2; x(3)-5; -x(1); -x(2); -x(3)];$

ceq = [];
B3: Gọi chương trình tìm phương án tối ưu (Viết nó trong MATLAB file).
clc
clear all
warning off
$x0 = [.1, .1, 3.0]$; fprintf ('The values of function value and constraints at starting pointn');
$f = objfun(x0)$
$[c, ceq] = constraints(x0)$
options = optimset ('LargeScale', 'off');
$[x, fval] = fmincon(@objfun, x0, [], [], [], [], [], [],$
$@constraints, options)$
fprintf ('The values of constraints at optimum solutionn');
$[c, ceq] = constraints(x)$
Kết quả tính toán được liệt kê như sau:
This Produces the Solution or Ouput as follows: The values of function value and constraints at starting point

f=

4.0410

c=

-8.9800

-5.0200

-2.0000

-0.1000

-0.1000

-3.0000

ceq =

[]

Optimization terminated: first-order optimality measure less than options. TolFun and maximum constraint violation is less than options.TolCon.

Active inequalities (to within options.TolCon = 1e-006): lower upper ineqlin ineqnonlin

1

2

4

x=

0 1.4142 1.4142

fval =

1.4142

The values of constraints at optimum solution

c=

476 Nonlinear Programming III: Constrained Optimization Techniques

-0.0000

-0.0000

-3.5858

0

-1.4142

-1.4142

ceq =

[]

3.7. CÁC PHƯƠNG PHÁP TỐI ƯU HIỆN ĐẠI

3.7.1 Phương pháp GEN

3.7.2 Phương pháp PSO

Phương pháp tối ưu bầy đàn là một dạng của các thuật toán tiến hóa quần thể đã được biết đến trước đây như thuật giải di truyền (Genetic algorithm (GA)), Thuật toán đàn kiến(Ant colony algorithm). Tuy vậy PSO khác với GA ở chỗ nó thiên về sử dụng sự tương tác giữa các cá thể trong một quần thể để khám phá không gian tìm kiếm. Tối ưu hóa bầy đàn, viết tắt là PSO , dựa trên hành vi của một nhóm hoặc bầy đàn của loài côn trùng, chẳng hạn như kiến, mối, ong, ong bắp cày và một đàn chim. Một cá thể, một chim trong một đàn gọi là hạt trong một quần thể, mỗi một hạt sử dụng trí thông minh riêng của mình và trí tuệ quần thể hoặc nhóm của bầy đàn để thực hiện mục tiêu tìm kiếm. Thuật toán ra đời vào năm 1995 tại một hội nghị của IEEE bởi James Kennedy và Russell Eberhart. Thuật toán có nhiều ứng dụng quan trọng trong tất cả các lĩnh vực mà ở đó đòi hỏi phải giải các bài toán về tối ưu hóa.

Trong bối cảnh tối ưu hóa đa biến, PSO được giả định là các quy định hoặc kích thước cố định với mỗi hạt (cá thể) đặt ban đầu tại các địa điểm ngẫu nhiên trong không gian nhiều chiều. Mỗi hạt có hai đặc điểm: vị trí và vận tốc. Mỗi hạt lấy ngẫu nhiên không gian đang xét và nhớ vị trí tốt nhất (về nguồn thức ăn hoặc giá trị hàm mục tiêu) mà nó đã phát hiện ra. Các hạt truyền đạt thông tin hoặc các vị trí tốt với nhau và điều chỉnh vị trí cá nhân và vận tốc của nó dựa trên các thông tin nhận được tại vị trí tốt nhất. Ví dụ, xem xét các hành vi của các loài chim trong một đàn.

Mặc dù mỗi con chim có một trí thông minh hạn chế nhưng chúng kiểm soát được hành vi theo các quy tắc sau:

1. Nó sẽ cố gắng không đi quá gần với các con chim khác. (Nguyên tắc gắn kết)
2. Nó sẽ điều tiết theo hướng trung bình của các con chim khác. (Nguyên tắc tách)
3. Nó sẽ cố gắng để phù hợp với vị trí "trung bình" giữa các con chim khác để không quá xa với các con chim trong bầy. (Nguyên tắc liên kết)

PSO được phát triển dựa trên mô hình sau đây:

1. Khi một con chim nằm một mục tiêu hoặc thực phẩm (hoặc tối đa của hàm mục tiêu), nó ngay lập tức truyền thông tin cho tất cả các con chim khác.
2. Tất cả các loài chim khác đổ về các mục tiêu hoặc thực phẩm (hoặc tối đa của mục tiêu chức năng), nhưng không trực tiếp.
3. Có một thành phần của tư duy độc lập của mỗi con chim cũng như bộ nhớ của nó trong quá khứ.

Vì vậy, các mô hình mô phỏng một tìm kiếm ngẫu nhiên trong không gian thiết kế cho các giá trị tối đa của hàm mục tiêu . Như vậy, dần dần qua nhiều lần lặp lại , những con chim đi đến mục tiêu.

Bài toán về tìm kiếm thức ăn của đàn chim có thể được mô hình hóa như sau:
Xét bài toán tối ưu của hàm f trong không gian n chiều. Mỗi vị trí trong không gian là một điểm tọa độ n chiều. Hàm f gọi là hàm mục tiêu xác định trong không gian n chiều và nhận giá trị thực. Mục đích là tìm ra điểm cực tiểu của hàm f trong miền xác định nào đó. Ta bắt đầu xem xét sự liên hệ giữa bài toán tìm thức ăn với bài toán tìm cực tiểu của hàm theo

cách như sau. Giả sử rằng số lượng thức ăn tại một vị trí tỉ lệ nghịch với giá trị của hàm f tại vị trí đó. Có nghĩa là ở một vị trí mà giá trị hàm f càng nhỏ thì số lượng thức ăn càng lớn. Việc tìm vùng chứa thức ăn nhiều nhất tương tự như việc tìm ra vùng chứa điểm cực tiểu của hàm f trên không gian tìm kiếm.

3.7.3 Thuật toán PSO cho bài toán tối ưu ràng buộc hình hộp

Bài toán được xét ở đây là

$$f(x) \to \max, \quad x \in [X^{(l)}, X^{(u)}], \tag{3.61}$$

ở đây $X^(l), X^(u)$ là cận trên và cận dưới tương ứng của x. Thuật toán PSO gồm các bước sau:

B1: **Khởi tạo:**

- **Kích thuộc quần thể,** N.
- **Trọng số quán tính** w.

(Giả sử kích thước của bầy đàn (số hạt) là N. Để giảm tổng số số đánh giá hàm mục tiêu cần thiết để tìm một nghiệm, chúng ta phải giả định một kích thước nhỏ hơn của bầy đàn. Nhưng với một kích thước bầy đàn quá nhỏ có khả năng chúng ta có thể không thể tìm thấy một lời giải nào cả. Thường có kích thước từ 20 đến 30 hạt được giả định cho bầy đàn là một giả thiết chấp nhận được. Đồng thời tạo hệ số quán tính.)

B2: **Khởi tạo các cá thể với vị trí và vận tốc ngẫu nhiên ban đầu.** (Tạo ra quần thể ban đầu của X trong khoảng $X^{(l)}$ và $X^{(u)}$ ngẫu nhiên $X_1, X_2, ..., X_N$. và vận tốc ban đầu. Để thuận tiện, các hạt (vị trí) j và vận tốc của nó trong lần lặp i được ký hiệu là $X_j^{(i)}$ và $V_j^{(i)}$, tương ứng. Do đó, các hạt được tạo ra ban đầu được là $X_1^{(0)}, X_2^{(0)}, ..., X_N^{(0)}$. Các vectơ $X_j^{(0)}, j = 1, 2, ..., N$ được gọi là các hạt hoặc véc tơ tọa độ của các hạt (tương tự như nhiễm sắc thể trong thuật toán di truyền). Vận tốc ban đầu thường chọn là 0.)

B3: **Tính giá trị của hàm mục tiêu trên các hạt**

$$f(X_1^{(0)}), f(X_2^{(0)}), ..., f(X_N^{(0)}).$$

B4: **Tính** $p_{best,j}$ **và** g_{best}.
(Giá trị lịch sử tốt nhất của $X_j^{(i)}$ (hạt thứ j tại bước lặp i), $P_{best,j}$, với giá trị cao nhất của hàm mục tiêu, $f(X_j^{(i)})$, gặp phải hạt j trong tất cả các bước lặp trước. Giá trị lịch sử tốt nhất của $X_j^{(i)}$ (tọa độ của tất cả các hạt cho đến lần lặp hiện tại) , G_{best}, với giá trị cao nhất của hàm mục tiêu $f(X_j^{(i)})$, với tất cả các bước lặp trước đó của N cá thể.)

B5: **Cập nhật,** c_1, c_2, **vận tốc, vị trí,** $p_{best,j}$ **và** g_{best} **của các cá thể.**
(Tính vận tốc của hạt j trong bước lặp i như sau:

$$V_j^{(i)} = V_j^{(i-1)} + c_1 r_1 [P_{best,j} - X_j^{(i-1)}] + c2 r_2 [Gbest - X_j^{(i-1)}], j = 1, 2, ..., N$$

trong đó c_1, c_2 là các giá trị lựa chọn theo một tiêu chí nào đó r_1, r_2 được phân bố đều các số ngẫu nhiên trong khoảng $[0, 1]$. Các thông số c_1 và c_2 biểu thị tầm quan trọng tương đối của bộ nhớ (vị trí) của hạt chính nó vào bộ nhớ (vị trí) của bầy đàn.)
Xác định quần thể mới của các thứ j hạt trong lần lặp thứ i như sau:

$$X_j^{(i)} = X_j^{(i-1)} + V_j^{(i)}, j = 1, 2, ..., N. \tag{3.62}$$

(Giá trị lịch sử tốt nhất của $X_j^{(i)}$ (hạt thứ j tại bước lặp i), $P_{best,j}$, với giá trị cao nhất của hàm mục tiêu, $f(X_j^{(i)})$, gặp phải hạt j trong tất cả các bước lặp trước. Giá trị lịch sử tốt nhất của $X_j^{(i)}$ (tọa độ của tất cả các hạt cho đến lần lặp hiện tại) , G_{best}, với giá trị cao nhất của hàm mục tiêu $f(X_j^{(i)})$, với tất cả các bước lặp trước đó của N cá thể.)

B6: **Kiểm tra điều kiện tiêu chuẩn, nếu thỏa chuyển B7**, nếu không chuyển B1.

B7: **Dừng**.

3.7.4 Cải tiến phương pháp PSO

Thuật toán cho thấy, thường thì vận tốc hạt xây dựng quá nhanh và cực đại của hàm mục tiêu được bỏ qua. Do đó một thuật ngữ quán tính, θ được thêm vào để giảm tốc độ. Thông thường, giá trị của θ được giả định thay đổi tuyến tính $0.9-0.4$ là quá trình lặp đi lặp lại tiến triển. Vận tốc của hạt thứ j, với hạn quán tính, được giả định là

$$V_j^{(i)} = \theta(i)V_j^{(i-1)} + c_1r_1[P_{best,j} - X_j^{(i-1)}] + c_2r_2[G_{best} - X_j^{(i-1)}], j = 1, 2, ..., N. \tag{3.63}$$

Quán tính trọng lượng θ ban đầu được giới thiệu bởi Shi và Eberhart vào năm 1999 để làm giảm vận tốc theo thời gian (hoặc các bước lặp), cho phép quần thể hội tụ chính xác hơn và hiệu quả so với PSO với công thức (3.62) vận tốc trước. Phương trình (3.63) biểu thị một công thức vận tốc thích nghi, cải thiện tốt khả năng điều chỉnh trong việc tìm kiếm giải pháp . Phương trình (3.63) chỉ ra rằng, giá trị lớn hơn của θ thúc đẩy thăm dò toàn cục và một giá trị nhỏ hơn thúc đẩy tìm kiếm địa phương. Vì vậy, giá trị lớn của θ làm cho các thuật toán liên tục khám phá những lĩnh vực mới mà không tìm kiếm địa phương và do đó không tìm thấy sự tối ưu thực sự. Để đạt được một sự cân bằng giữa thăm dò toàn cầu và địa phương để tăng tốc độ hội tụ để tìm tối ưu sự thật, trọng lượng quán tính được sử dụng là:

$$\theta(i) = \theta_{\max} - \frac{\theta_{\max} - \theta_{\min}}{i_{\min}},$$

ở đây $\theta_{\max} = 0.9, \theta_{\min} = 0.4$ là những giá trị ban đầu và cuối cùng của trọng lượng quán tính , tương ứng của các phép lặp được sử dụng trong PSO.

Ví dụ 3.7.35. *Tìm giá trị tối ưu của hàm*

$$f(x) = -x2 + 2x + 11$$

trong khoảng $-2 \leq x \leq 2$ *sử dụng phương pháp PSO.*

Ta thực hiện lần lượt như sau:

1. Khởi tạo số hạt $N = 4$.

2. - Khởi tạo quần thể(vị trí) ban đầu $x_1 = -1.5, x_2 = 0.0, x_3 = 0.5, vx_4 = 1.25$.
   - Tính giá trị hàm mục tiêu tại x_j hiện tại (bước 0), $j = 1, 2, 3, 4 f_1 = f(x_1(0)) = f_{(-1.5)} = 5.75, f_2 = f(x_2(0)) = f(0.0) = 11.0, f_3 = f(x_3(0)) = f(0,5) = 11,75$, và $f_4 = f(x4(0)) = f(1.25) = 11,9375$.

3. Thiết lập vận tốc ban đầu của mỗi hạt bằng không:

   $$v_1(0) = v_2(0) = v_3(0) = v_4(0) = 0,$$

   Thiết lập số lần lặp như $i = 1$ và đi đến bước 4.

4. Tìm $P_{best,1} = -1.5, P_{best,2} = 0.0, P_{best,3} = 0,5, P_{best,4} = 1.25$ và $G_{best} = 1.25$.

5. Tìm vận tốc của các hạt như (bằng cách giả sử $c1 = c2 = 1$ và sử dụng các số ngẫu nhiên trong khoảng $(0, 1)$ là $r_1 = 0,3294$ và $r_2 = 0,9542$)

   $$v_j^{(i)} = v_j^{(i-1)} + r_1[P_{best,j} - x_j^{(i-1)}] + r_2[Gbest - x_j^{(i-1)}]; j = 1, 2, 3, 4 \quad (3.64)$$

   nhận được

   $$\begin{aligned} v_1^{(1)} &= 0 + 0.3294(-1.5 + 1.5 + 0.9542(1.25 + 1.5) = 2.6241 \\ v_2^{(1)} &= 0 + 0.3294(0.0 - 0.0) + 0.9542(1.25 - 0.0) = 1.1927 \\ v_3^{(1)} &= 0 + 0.3294(0.5 - 0.5) + 0.9542(1.25 - 0.5) = 0.7156 \\ v_4^{(1)} &= 0 + 0.3294(1.25 - 1.25) + 0.9542(1.25 - 1.25) = 0.0 \end{aligned}$$

6. Tìm các giá trị mới của $x_j^{(1)}, j = 1, 2, 3, 4,$ theo công thức

$$x_j^{(i)} = x_j^(i - 1) + v_j^{(i)} :$$

$$\begin{aligned}
x_1^{(1)} &= -1.5 + 2.6241 = 1.1241 \\
x_2^{(1)} &= 0.0 + 1.1927 = 1.1927 \\
x_3^{(1)} &= 0.5 + 0.7156 = 1.2156 \\
x_4^{(1)} &= 1.25 + 0.0 = 1.25
\end{aligned}$$

7. Đánh giá các giá trị hàm mục tiêu tại $x_j^{(i)}$: $f(x_1^{(1)}) = 11.9846, f(x_2^{(1)}) = 11.9629, f(x_3^{(1)}) = 11.9535, f(x_4^{(1)}) = 11.9375$ Kiểm tra sự hội tụ của các giải pháp hiện tại . Từ các giá trị của $x_j^{(i)}$ đã làm không hội tụ , chúng ta tăng số lần lặp là $i = 2$ và quay lại bước 4.

8. Tìm $P_{best,j}, G_{best}$ $P_{best,1} = 1.1241, P_{best,2} = 1.1927, P_{best,3} = 1.2156, P_{best,4} = 1.25$, và $G_{best} = 1.1241$.

9. Tính vận tốc mới của các hạt theo công thức (3.64) với $c1 = c2 = 1$ và sử dụng các số ngẫu nhiên trong khoảng $(0, 1)$ là $r_1 = 0.1482, r_2 = 0.4867$:

$$\begin{aligned}
v_1(2) &= 2.6240 + 0.1482(1.1241 - 1.1241) + 0.4867(1.1241 - 1.1241) \\
&= 2.6240 \\
v_2(2) &= 1.1927 + 0.1482(1.1927 - 1.1927) + 0.4867(1.1241 - 1.1927) \\
&= 1.1593 \\
v_3(2) &= 0.7156 + 0.1482(1.2156 - 1.2156) + 0.4867(1.1241 - 1.2156) \\
&= 0.6711 \\
v_4(2) &= 0.0 + 0.1482(1.25 - 1.25) + 0.4867(1.1241 - 1.25) \\
&= -0.0613
\end{aligned}$$

9. Tính giá trị hiện tại của $x_j^{(i)}$ theo $x_j^{(i)} = x_j^{(i-1)} + v_j^{(i)}(i), j = 1, 2, 3, 4$:

$$\begin{aligned} x_1(2) &= 1.1241 + 2.6240 = 3.7481 \\ x_2(2) &= 1.1927 + 1.1593 = 2.3520 \\ x_3(2) &= 1.2156 + 0.6711 = 1.8867 \\ v_4(2) &= 1.25 - 0.0613 = 1.1887 \end{aligned}$$

10. Tìm các giá trị hàm mục tiêu tại $x_j^{(i)}$:
$f(x_1^{(2)}) = 4.4480, f(x_2^{(2)} = 10.1721, f(x_3^{(2)}) = 11.2138, f(x_4^{(2)}) = 11.9644$. Tiếp tục tăng i và quay lại bước 4 nếu điều kiện tối ưu chưa thỏa mãn.

Kết luận: Những thuật toán đã trình bày ở các phần trên chỉ là một phần cơ bản và rất nhỏ trong kho tàng các thuật toán giải các bài toán ***tối ưu hóa***. Hơn nữa, do kiến thức còn hạn chế và thời gian quá gấp nên khi trình bày chuyên đề này thiếu sót là điều không thể tránh khỏi, đặc biệt là hình vẽ minh họa còn chưa thỏa mãn được những yêu cầu của người biên soạn giáo trình này. Với tôi đây là một dịp tốt để tổng hợp lại những kiến thức về môn học này và sẽ cố gắng hoàn thiện trong thời gian tới những khiếm khuyết của bài giảng. Cuối cùng một lần nữa tôi xin cám ơn những góp ý của bạn bè và đồng nghiệp về bản thảo của chuyên đề.

Tài liệu tham khảo

[1] Bùi Minh Trí, *Tối ưu hóa,* Nhà xuất bản KHKT, Hà nội, 2006.

[2] J. W. Daniel, *Stability of the solution of definite quadratic programs,* Math Programming, **5**, 1973, 41–53.

[3] M. Frank and P. Wolfe, *An algorithm for quadratic programming,* Naval Research Logistics Quarterly, **3**, 1956, 95-110.

[4] M. A. Hanson, *On sufficiency of the Kuhn-Tucker conditions,* J. Math. Annal. Appl., **80**, 1981, 545-550.

[5] Nguyễn Nhật Lệ, *Các bài toán cơ bản của tối ưu hóa và điều khiển tối ưu,* Nhà xuất bản KHKT, Hà nội, 2009.

[6] A. D. Ioffe and V. M. Tikhomirov, *Theory of Extremal Problems,* (Tiếng Nga), Nauka, Moscow, 1974.

[7] H.W. Kuhn and A. W. Tucker, *Nonlinear Programming,* Proceedings of the Second Berkeley Symposium on the Mathematical Statistics and Probability, University of California Press, Berkeley, 418-492, 1951.

[8] G. M. Lee, N. N. Tam and N. D. Yen, *Quadratic Programming and Affine Variational Inequalities,* Springer, 2005.

[9] D. G. Luenberg and Yinyu Ye, *Linear and Nonlinear Programming,* Springer, 2007.

[10] H. X. Phu, *Lý thuyết các bài toán cực trị,* Tài liệu giảng dạy cao học -Viện toán hoc, 1998.

[11] B. N. Pshenhishnui, *Convex Analysis and Extremal Problems,* Nauka, Moscow, 1980.

[12] C. H. Paradimitriou and K. Steglitz, *Conbinatorial Optimization: Algorithm and Complexity,* Dover, Nework, 1998.

[13] Singiresu S. Rao, *Engeneering Optimization, Theory and Practice* John Wiley and Inc, 2009.