ĐẠI HỌC ĐÀ NẴNG
TRƯỜNG ĐẠI HỌC SƯ PHẠM
KHOA TOÁN
*****

BÀI GIẢNG

TOÁN TỐI ƯU

Biên soạn : **TS. Hoàng Quang Tuyến**

Đà Nẵng - 2012

Giới thiệu

Tập tài liệu này được biên soạn bởi Thầy giáo TS Hoàng Quang Tuyến, sử dụng cho giảng dạy môn **Toán Tối Ưu** trong chương trình đào tạo thạc sỹ ngành Phương Pháp Toán Sơ Cấp của Đại Học Đà Nẵng. Đã có một số bản đánh máy tài liệu này, nhưng các bản trước đó đều có khá nhiều lỗi chẳng hạn như thiếu một số dòng, sai ký hiệu, sai công thức,... Mình đã mượn thầy Tuyến bản viết tay giáo trình của môn Toán Tối Ưu của thầy và soạn lại trên **Latex**. Hy vọng sẽ giúp ích cho các bạn học viên khóa sau đỡ vất vả hơn khi học môn này.

Đây là bản đầu tiên nên có thể vẫn còn một vài chỗ nhầm lẫn, mong được mọi người cùng góp ý để giáo trình được hoàn thiện một cách chính xác nhất. Mọi ý kiến đóng góp, xin gửi vào địa chỉ email của mình

`hablack18@gmail.com`

Chương 1

CƠ BẢN VỀ GIẢI TÍCH LỒI

1.1 Tập lồi

Các ký hiệu:

- Một vector a luôn hiểu là một vector cột.
- Chuyển vị của vector a là một vector hàng a^T.
- Tích vô hướng của hai vector a, b là $\langle a, b\rangle$ hay $a^T b$.
- Tập các số thực là $\mathbb{R}$.

Định nghĩa 1.1. *Đường thẳng đi qua hai điểm a, b trong không gian Euclid n-chiều $\mathbb{R}^n$ là tập hợp các điểm $x \in \mathbb{R}^n$ có dạng*

$$x = \lambda a + (1-\lambda)b, \lambda \in \mathbb{R}.$$

Định nghĩa 1.2. *Đoạn thẳng nối hai điểm a, b trong $\mathbb{R}^n$ là tập hợp các điểm $x \in \mathbb{R}^n$ có dạng*

$$x = \lambda a + (1-\lambda)b, 0 \leq \lambda \leq 1.$$

Định nghĩa 1.3. *Tập $M \subset \mathbb{R}^n$ gọi là đa tạp affine nếu với hai điểm bất kỳ $x, y \in M$ thì đường thẳng đi qua x, y cũng thuộc M. Tức là*

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

Mỗi đa tạp affine đều có duy nhất một không gian con L song song với nó. Tức là $L = M + a, a \in \mathbb{R}^n$. Thứ nguyên của M là thứ nguyên của L.

Định nghĩa 1.4. *Siêu phẳng trong $\mathbb{R}^n$ là tập*

$$\{x = (x_1, x_2, \ldots x_n) | x_1 a^1 + x_2 a^2 + \ldots + x_n a^n = \alpha, a^i \in \mathbb{R}, \forall i = 1..n, \alpha \in \mathbb{R}\}.$$

Ví dụ 1.1.1. *Siêu phẳng trong không gian 2 chiều là đường thẳng, trong không gian 3 chiều là mặt phẳng.*

Bài tập 1.1. *Siêu phẳng có phải là đa tạp?*

Định nghĩa 1.5. *(Về các nửa không gian)*

- *Nửa không gian đóng trong* $\mathbb{R}^n$ *là tập*

  $$\left\{x = (x_1, x_2, \dots x_n) \,\middle|\, x_1a^1 + x_2a^2 + \dots + x_na^n \leq \alpha, a^i \in \mathbb{R}, \forall i = 1..n, \alpha \in \mathbb{R}\right\}.$$

- *Nửa không gian mở trong* $\mathbb{R}^n$ *là tập*

  $$\{x = (x_1, x_2, \dots x_n) | x_1a^1 + x_2a^2 + \dots + x_na^n < \alpha, a^i \in \mathbb{R}, \forall i = 1..n, \alpha \in \mathbb{R}\}$$

- *Đây là các nửa không gian được xác định bởi siêu phẳng*

  $$x_1a^1 + x_2a^2 + \dots + x_na^n = \alpha$$

- *Hai nửa không gian đóng, mở nằm bên kia siêu phẳng so với hai nửa siêu phẳng trên là*

  $$\{x = (x_1, x_2, \dots x_n) | x_1a^1 + x_2a^2 + \dots + x_na^n \geq \alpha, a^i \in \mathbb{R}, \forall i = 1..n, \alpha \in \mathbb{R}\},$$
  $$\{x = (x_1, x_2, \dots x_n) | x_1a^1 + x_2a^2 + \dots + x_na^n > \alpha, a^i \in \mathbb{R}, \forall i = 1..n, \alpha \in \mathbb{R}\}.$$

Định nghĩa 1.6. *(Tập lồi)*
Tập $D \subset \mathbb{R}^n$ *gọi là tập lồi nếu*

$$\forall a, b \in D \text{ và } \lambda \in [0, 1] \text{ ta có } \lambda a + (1 - \lambda)b \in D.$$

Định nghĩa 1.7. *(Nón lồi)*
Tập $D \subset \mathbb{R}^n$ *gọi là nón lồi nếu*

$$\forall x, y \in D \text{ thì } x + y \in D \text{ và } tx \in D, \forall t \geq 0.$$

Ví dụ 1.1.2. $\mathbb{R}^n_+$ *là nón lồi.*

Bài tập 1.2. *Nón lồi có phải là tập lồi?*

Định nghĩa 1.8. *(Bao lồi)*
Bao lồi của tập A là tập lồi nhỏ nhất chứa A, ký hiệu $CovA$.

Ví dụ 1.1.3. $A = \{x; y\} \Rightarrow CovA = \{\lambda x + (1 - \lambda)y | 0 \leq \lambda \leq 1\}$.

Định nghĩa 1.9. *(Tổ hợp lồi của hai tập).*
Cho $A \subset \mathbb{R}^n, B \subset \mathbb{R}^n$, tổ hợp lồi của A và B là tập hợp các điểm thuộc $\mathbb{R}^n$ có dạng

$$x = \lambda a + (1-\lambda)b, a \in A, b \in B, 0 \leq \lambda \leq 1.$$

Bài tập 1.3. *Tổ hợp lồi là tập lồi?*

Định lý 1.1. *Tập lồi là đóng với phép giao, phép cộng, phép nhân với một số và phép lấy tổ hợp tuyến tính. Tức là, nếu A, B là hai tập lồi trong $\mathbb{R}^n$ thì các tập sau đây cũng lồi :*

i) $A \cap B := \{x | x \in A, x \in B\}$,

ii) $\lambda A + \beta B := \{x = \lambda a + \beta b | a \in A, b \in B, \lambda, \beta \in \mathbb{R}\}$.

Định nghĩa 1.10. *Thứ nguyên của một tập lồi A là thứ nguyên của đa tạp affine nhỏ nhất chứa A, gọi là bao affine của A ký hiệu là $affA$. Thứ nguyên của tập lồi A ký hiệu là $dimA$.*

Nhận xét 1. *Nếu $A \subset \mathbb{R}^n$ thì $dimA \leq n$.*

Định nghĩa 1.11. *Tập hợp các điểm trong tương đối của một tập $A \subset \mathbb{R}^n$ là tập hợp*

$$riA := \{x \in A | \exists U(x), U(x) \cap affA \subset A\},$$

trong đó : $U(x)$ là lân cận mở của x.

Bài tập 1.4. *Nếu $A \neq \emptyset$ và lồi thì $riA \neq \emptyset$.*

Định nghĩa 1.12. *Một tập hợp được gọi là tập lồi đa diện (hay khúc lồi) nếu nó là giao của hữu hạn các nửa không gian đóng.*

Như vậy, khúc lồi là tập hợp thỏa mãn các bất phương trình dạng :

$$\begin{cases} a_{11}x^1 + a_{12}x^2 + & \ldots + a_{1n}x^n \leq b_1 \\ a_{21}x^1 + a_{22}x^2 + & \ldots + a_{2n}x^n \leq b_2 \\ & \ldots \\ a_{m1}x^1 + a_{m2}x^2 + & \ldots + a_{mn}x^n \leq b_m \end{cases}$$

Hệ bất phương trình này có thể viết dưới dạng $Ax \leq b$, trong đó

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}; x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}; b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Nhận xét 2. *Khúc lồi là một tập đóng, có thể không bị chặn.*

Định nghĩa 1.13. *Một khúc lồi bị chặn gọi là đa diện lồi. Một tập con A' của khúc lồi A được gọi là một diện của A nếu:*

$$\forall a, b \in A, x = \lambda a + (1-\lambda)b; 0 < \lambda < 1, x \in A' \Rightarrow a, b \in A'.$$

Nhận xét 3.

- *Mọi diện của một tập lồi đa diện cũng là tập lồi đa diện. (Chứng minh nhận xét này xem như bài tập)*
- *Một diện có thứ nguyên 0 gọi là một đỉnh (điểm cực biên).*
- *Cạnh là diện có thứ nguyên bằng 1.*

Định nghĩa 1.14. *Điểm $x \in C$ gọi là điểm cực biên của tập C (C không nhất thiết lồi) nếu C không có đoạn thẳng nào nhận x làm điểm trong.*

Định nghĩa 1.15. *Một vector $h \neq 0$ được gọi là phương vô hạn của tập C nếu :*

$$x + \lambda h \subset C, \forall x \in C, \forall \lambda > 0.$$

Định lý 1.2.

i) Mọi khúc lồi không chứa trọn một đường thẳng đều có ít nhất một đỉnh.

ii) Mọi khúc lồi A có đỉnh đều là tập

$$A := \left\{ x = \sum_{i\in I} \lambda_i v^i + \sum_{j\in J} \beta_j d^j \,\middle|\, \sum_{i\in I} \lambda_i = 1, \lambda_i, \beta_j \geq 0 \right\}.$$

Trong đó: $v^i \in \{$Tập I đỉnh$\}, d^j \in \{$Tập J phương vô hạn$\}$

Chú ý 1.1.1.

i) Nếu khúc lồi A bị chặn thì A chỉ là tổ hợp lồi của các đỉnh (tập I đỉnh):

$$A := \left\{ x = \sum_{i\in I} \lambda_i v^i \,\middle|\, \sum_{i\in I} \lambda_i = 1, \lambda_i \geq 0, \right\}.$$

ii) Nếu D là tập lồi đa diện (khúc lồi) thì D có thể biểu diễn:

$$D = E + D_0,$$

trong đó: E là không gian con, D_0 là khúc lồi có đỉnh.

Định nghĩa 1.16. *Ta nói siêu phẳng $H = \{x\,|\langle v, x\rangle = \alpha\}$ tách hai tập A và B nếu:*

$$\langle v, a\rangle \le \alpha, \langle v, a\rangle \ge \alpha, \forall a \in A, \forall b \in B, \tag{1.1}$$

ta nói H tách hẳn A và B nếu (1.1) có ít nhất một đẳng thức thực sự.

Định lý 1.3. *Cho A là một tập lồi đóng và $x^0 \notin A$. Lúc đó tồn tại một siêu phẳng tách A và x^0*

Hệ quả 1.3.1. *(Bổ đề Farkas)*
Cho $a \in \mathbb{R}^n$ và A là ma trận cấp m×n. Khi đó:

$$\langle a, x\rangle \ge 0, \forall x \text{ thỏa mãn } Ax \ge 0 \Leftrightarrow \exists y \ge 0 \in \mathbb{R}^m \text{ sao cho } a = A^T y.$$

Nhận xét 4. *Ý nghĩa hình học của bổ đề là siêu phẳng đi qua gốc tọa độ $\langle a, x\rangle = 0$ tách nón $\{x|Ax \ge 0\}$ về một phía khi và chỉ khi vector pháp tuyến a của siêu phẳng thuộc nón sinh bởi các hàng của ma trận A.*

1.2 Hàm lồi

Giáo trình này chỉ xét các hàm số thực và nhận giá trị hữu hạn.

Định nghĩa 1.17. $\forall x, y \in A, 0 \le \lambda \le 1$

- *, Hàm số f xác định trên tập lồi A gọi là hàm lồi trên A nếu :*

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda)f(y).$$

- *Hàm f gọi là lồi chặt nếu :*

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda)f(y), \forall x, y \in A, 0 < \lambda < 1.$$

- *Hàm f gọi là **tựa lồi** (quasi convex) trên A nếu*

$$\forall \lambda \in \mathbb{R},\ \text{tập mức } \{x \in A|f(x) \le \lambda\} \text{ là một tập lồi .}$$

- *Hàm f gọi là **tựa lõm** (quasi convcave) trên A nếu $-f$ tựa lồi.*

Ví dụ 1.2.1. $f(x) = \dfrac{\langle a, x\rangle + \alpha}{\langle b, x\rangle + \beta}$

Định nghĩa 1.18. *Các hàm $\lambda f, f + g$ và $\max(f, g)$ được định nghĩa như sau:*

$$(\lambda f)(x) := \lambda f(x),$$
$$(f + g)(x) := f(x) + g(x),$$
$$\max(f, g)(x) := \max\{f(x), g(x)\}.$$

Định lý 1.4. *Cho f là hàm lồi trên tập lồi A và g là hàm lồi trên tập lồi B. Lúc đó trên $A \cap B$ các hàm sau là lồi:*

i) $\lambda f + \beta g, \forall \lambda, \beta \geq 0,$

ii) $\max(f, g)$.

Chứng minh định lý này như bài tập.

Định lý 1.5. *Một hàm lồi xác định trên tập lồi A thì liên tục tại mọi điểm trong của A.*

- Chú ý: Hàm lồi xác định trên tập lồi thì liên tục tại mọi điểm trong, chưa chắc liên tục trên điểm biên.
- Kí hiệu: $f'(a)$ hoặc $\nabla f(a)$ là **đạo hàm** của f tại a.

Định lý 1.6.

1. *Cho $f : A \to \mathbb{R}$ là một hàm khả vi trên tập lồi mở A. Điều kiện cần và đủ để f lồi trên A là :*

$$f(x) + \langle \nabla f(x), y - x \rangle \leq f(y), \forall x, y \in A.$$

2. *Nếu f khả vi hai lần thì f lồi trên A khi và chỉ khi $\forall x \in A$ ma trận Hessian H(x) của f tại x xác định không âm,tức là :*

$$y^T H(x) y \geq 0, \forall x \in A, y \in \mathbb{R}^n.$$

Chú ý 1.2.1. *Tính khả vi của một hàm lồi giữ vai trò quan trọng bậc nhất trong tối ưu hóa.*

Định nghĩa 1.19. *Ta gọi đạo hàm theo hướng d của một hàm số f (không nhất thiết lồi) tại x là một đại lượng số :*

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

nếu giới hạn này tồn tại.

Định lý 1.7. *Nếu f là một hàm lồi trên tập A thì $\forall x \in A$ và $\forall d \in \mathbb{R}^n$ sao cho $x + d \in A$ đạo hàm theo hướng d của f tại x luôn tồn tại và nghiệm đúng*

$$f'(x, d) \leq f(x + d) - f(x).$$

Ngoài ra, với mỗi x cố định, $f'(x, .)$ là hàm lồi trên tập lồi $\{d : x + d \in A\}$.

Nhận xét 5.

- *Nếu f khả vi thì:* $f'(x,d) = \langle \nabla f(x), d\rangle, \forall d$
- *Hàm lồi chưa chắc khả vi tại mọi điểm*

Định nghĩa 1.20. *Cho f là một hàm trên tập lồi A. Một vector $y^* \in \mathbb{R}^n$ được gọi là dưới vi phân tại $x^* \in A$ nếu*

$$f(x) \geq f(x^*) + \langle y^*, x - x^*\rangle, \forall x \in A.$$

Tập các điểm y^* thỏa mãn bất đẳng thức này được ký hiệu $\partial f(x^*)$.
Trường hợp $\partial f(x^*)$ chỉ có một điểm ta nói f khả vi tại x^*.

Nhận xét 6.

i) *Tương tự trường hợp hàm một biến, bất đẳng thức*

$$f(x) \geq f(x^*) + \langle y^*, x - x^*\rangle, \forall x \in A$$

có nghĩa rằng siêu phẳng đi qua điểm $(x^*, f(x^*))$ nằm dưới đồ thị hàm số.

ii) *Tập $\partial f(x^*)$ có thể rỗng, tuy nhiên với hàm lồi khác $\emptyset$.*

Định lý 1.8. *Cho f là hàm lồi (hữu hạn) trên tập lồi A. Khi đó f có dưới vi phân tại mọi điểm trong tương đối riA.*

Nhận xét 7.
Nếu $A \equiv \mathbb{R}^n$ thì f có dưới vi phân tại mọi điểm vì $ri\mathbb{R}^n \equiv \mathbb{R}^n$.

1.3 Tính chất cực trị

Cho $D \subset \mathbb{R}^n, D \neq \emptyset$ và hàm số $f : D \to \mathbb{R}$ (không nhất thiết lồi).

Định nghĩa 1.21. *Một điểm $x^* \in D$ được gọi là cực tiểu địa phương của f trên D nếu tồn tại một lân cận mở U của x^* sao cho $f(x^*) \leq f(x)$ với mọi $x \in D \cap U$. Điểm x^* được gọi là cực tiểu tuyệt đối (toàn cục) của f trên D nếu :*

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

Dưới đây là hai tính chất cơ bản về cực trị của hàm lồi :

Định lý 1.9.

i) Mọi điểm cực tiểu địa phương của một hàm lồi trên một tập lồi đều là điểm cực tiểu tuyệt đối.

ii) Nếu x^* là điểm cực tiểu của hàm lồi f trên tập lồi D và $x^* \in intD$ thì $0 \in \partial f(x^*)$.

Định lý 1.10. *Cực đại của hàm lồi (nếu có) trên tập lồi có điểm cực biên bao giờ cũng đạt tại một điểm trên biên.*

Chương 2

ĐIỀU KIỆN TỐI ƯU

2.1 Bài toán tối ưu

Nhiều vấn đề thực tế trong các lĩnh vực đều có thể mô tả như một bài toán tối ưu.

Ví dụ 2.1.1. *Một xí nghiệp sản xuất n loại sản phẩm cần sử dụng m loại nguyên liệu khác nhau. Gọi x_j là số lượng sản phẩm thứ $j(j = \overline{1,n})$ và c_j là lãi thu được của một sản phẩm j. Biết rằng để sản xuất một sản phẩm loại j cần một lượng nguyên liệu $a_{ij}(i = \overline{1,m})$. Gọi b_i là số lượng tối đa của nguyên liệu i mà xí nghiệp có.*
Bài toán đặt ra là hãy sản xuất mỗi loại sản phẩm với số lượng bao nhiêu để tổng lợi nhuận thu được là lớn nhất.

Ta có mô hình toán học của bài toán trên như sau:

$$\max \sum_{j=1}^{n} c_j x_j \tag{2.1}$$

với điều kiện :

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i, i = \overline{1,m}, \tag{2.2}$$

$$x_j \geq 0, j = \overline{1,n}, \tag{2.3}$$

Dạng tổng quát của bài toán tối ưu được mô tả như sau:

$$\min f(x) \text{ Với điều kiện } x \in D, \tag{2.4}$$

$$\max f(x) \text{ Với điều kiện } x \in D. \tag{2.5}$$

Trong đó, D là một tập (có thể rỗng) trong không gian nào đó, f là hàm số thực xác định trên một tập chứa D. Thông thường D được môt tả như tập nghiệm của hệ đẳng thức (bất đẳng thức), cũng có thể là tập nghiệm của hệ phương trình vi phân (tích phân). D thường được gọi là tập **phương án chấp nhận được**.

Chú ý 2.1.1.

$$\min\{f(x)|x \in D\} = -\max\{-f(x)|x \in D\},$$

Và tập các lời giải tối ưu cho hai bài toán này trùng nhau. Do đó, ta có thể đưa bài toán tìm cực đại về bài toán tìm cực tiểu và ngược lại.

Xét bài toán (2.4), có bốn khả năng xảy ra đối với nghiệm tối ưu (tuyệt đối):

1. D là một tập rỗng (không có phương án chấp nhận được).
2. Cực tiểu của f trên D bằng $-\infty$.
3. Cực tiểu của f trên D hữu hạn nhưng không đạt trên D.
4. f đạt cực tiểu hữu hạn trên D.

Để tổng quát, nhiều khi người ta thay inf cho min và sup cho max.

Định nghĩa 2.1.

1. inf *của hàm f trên tập D là số t lớn nhất thỏa mãn $t \leq f(x), \forall x \in D$. Ký hiệu:* inf *của f trên D là* $\inf f(D)$
2. sup *của hàm f trên tập D là số t nhỏ nhất thoản mãn $t \geq f(x), \forall x \in D$. Ký hiệu :* sup *của f trên D là* $\sup f(D)$.

Ví dụ 2.1.2. $\min e^x$ *với ràng buộc $x < 0$ không đạt cực tiểu trên tập $x < 0$ nhưng* $\inf\limits_{x<0} e^x = 0$

Từ nay về sau, ta xét bài toán tối ưu trong không gian Euclide -$\mathbb{R}^n$:

$$\min f(x) \text{ với điều kiện } x \in D. \tag{2.6}$$

Trong đó D là tập đóng trong $\mathbb{R}^n$ gọi là miền chấp nhận được hay miền ràng buộc của bài toán (2.6). Một điểm thuộc D gọi là điểm chấp nhận được.
f là hàm số xác định trên một tập nào đó chứa D và được gọi là **hàm mục**

tiêu.

Bài toán (2.6) được gọi là một quy hoạch lồi nếu D lồi và f lồi trên D.

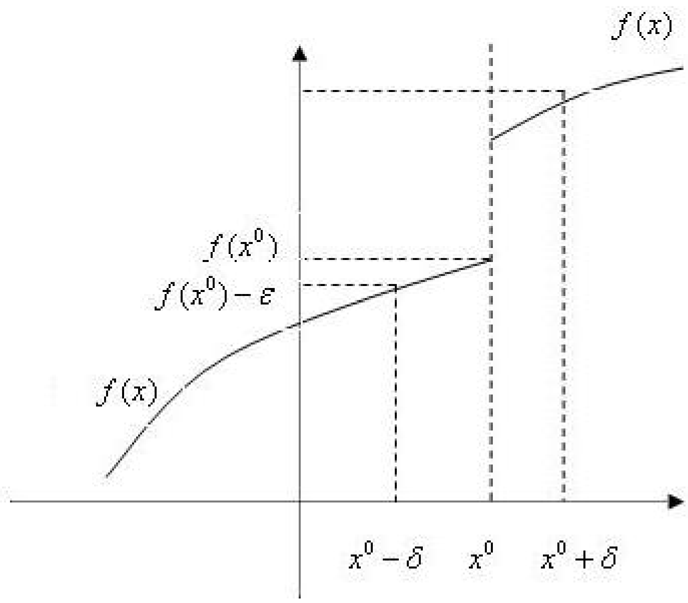
Sau đây, để khảo sát sự tồn tại nghiệm của bài toán tối ưu, ta nhắc lại một số khái niệm giải tích.

Định nghĩa 2.2.

1. *Một hàm f xác định trên X gọi là nửa liên tục dưới tại điểm $x_0 \in X$ nếu $\forall \epsilon > 0, \exists \delta > 0$ sao cho :*

$$f(x) \geq f(x_0) - \epsilon, \forall x \in X \text{ thỏa mãn } : \|x - x_0\| < \delta.$$

2. *Hàm f gọi là nửa liên tục trên tại x_0 nếu $-f$ là nửa liên tục dưới tại x_0.*



Hình 2.1: Minh họa hàm nửa liên tục dưới tại x^0

Chú ý 2.1.2. *Hàm f nửa liên tục dưới trên tập đóng X khi và chỉ khi $\forall t \in \mathbb{R}$ tập mức:*

$$levf(X) = \{x \in X | f(x) \leq t\}$$

là tập đóng.

Định lý 2.1. *Điều kiện cần và đủ để hàm f đạt cực tiểu trên D là tập $F^{\perp}(D) := \{t \in \mathbb{R} | f(x) \leq t, x \in D\}$ đóng và bị chặn dưới.*

Chứng minh. Nếu x^* là điểm cực tiểu của f trên D thì $F^{\perp}(D) := [f(x^*), +\infty)$ là đóng (do phần bù mở), và bị chặn dưới bởi $f(x^*)$.
Ngược lại $F^{\perp}(D)$ bị chặn dưới suy ra: $\inf F^{\perp}(D) = t_* > -\infty$.
Do $F^{\perp}(D)$ đóng $\Rightarrow t_* \in F^{\perp}(D) \Rightarrow \exists x^* \in D : t_* = f(x^*)$
Vậy, x^* là một điểm cực tiểu của f trên D. □

Định lý 2.2. *Nếu D compact và f nửa liên tục dưới trên D thì f đạt cực tiểu trên D.*

Chứng minh. Đặt $t_* = \inf f(D)$. Theo định nghĩa inf tồn tại $\{x^n\} \subset D$: $f(x^n) \to t_*$. Do D compact nên tồn tại dãy con của $\{x^n\}$ hội tụ đến $x^* \in D$.
Do f nửa liên tục dưới: $f(x^{n_k}) \geq f(x^*) - \epsilon$
Qua giới hạn ta được: $t_* = \lim f(x^{n_k}) \geq f(x^*)$
Do t_* là inf $\Rightarrow t_* = f(x^*)$. Suy ra x^* là điểm cực tiểu của f trên D. □

2.2 Điều kiện tối ưu và đối ngẫu Lagrange

2.2.1 Điều kiện tối ưu

Định nghĩa 2.3. *Một vector $d \neq 0$ được gọi là **hướng chấp nhận được** của tập D tại $x^* \in D$ nếu tồn tại số thực $\lambda_* > 0$ sao cho:*

$$x^* + \lambda d \in D \text{ với mọi } 0 < \lambda \leq \lambda_*.$$

Tập các hướng chấp nhận được của D tại x^* được ký hiệu là $D(x^*)$ và bao đóng là $\overline{D}(x^*)$.

Định lý 2.3. *Giả sử f khả vi trong một tập mở chứa D. Nếu x^* là cực tiểu địa phương của f trên D thì $d^T \nabla f(x^*) \geq 0$ với mọi $d \in \overline{D}(x^*)$.*

Chứng minh. Khai triển Taylor tại x^*:

$$f(x^* + \lambda d) = f(x^*) + \lambda \langle \nabla f(x^*), d \rangle + r(\lambda d),$$

$$\text{trong đó: } \frac{r(h)}{h} \to 0 \text{ khi } h \to 0.$$

Từ đây và do x^* cực tiểu địa phương, $d \in D(x^*)$ nên:

$$\begin{aligned} & f(x^* + \lambda d) - f(x^*) \geq 0, \forall \lambda \text{ đủ nhỏ}. \\ & \Rightarrow \lambda \langle \nabla f(x^*), d \rangle + r(\lambda d) \geq 0, \forall \lambda \text{ đủ nhỏ}. \\ & \Rightarrow d^T \nabla f(x^*) = \langle \nabla f(x^*), d \rangle \geq 0, \forall d \in D(x^*). \end{aligned}$$

Nếu $d \in \overline{D}(x^*)$ thì suy ra: $d = \lim_k d^k$ với $d^k \in D(x^*)$.
Vì $\left\langle d^k, \nabla f(x^*) \right\rangle \geq 0$, cho $k \to +\infty$ ta được:

$$d^T \nabla f(x^*) = \langle d, \nabla f(x^*) \rangle \geq 0, \forall d \in \overline{D}(x^*).$$

□

Định nghĩa 2.4. *Một vector $x^* \in D$ được gọi là **điểm dừng** của f trên D nếu $d^T \nabla f(x^*) \geq 0, \forall d \in D(x^*)$.*

Nhận xét 8.

1. *Nếu $x^* \in int(D)$ thì $D(x^*) = \mathbb{R}^n$ và $\nabla f(x^*) = 0$.*
   *(**Chứng minh:** Lấy $d = \nabla f(x^*)$ và $d = -\nabla f(x^*)$ thế vào công thức điểm dừng).*

2. *Từ định lý 2.3, nếu x^* là cực tiểu địa phương thì x^* là điểm dừng, tuy nhiên điều ngược lại không đúng. Chẳng hạn 0 là điểm dừng của $f(x) = x^3$ nhưng nó không phải là cực tiểu trên đoạn $[a, b]$ chứa điểm 0 nào. Tuy vậy, với quy hoạch lồi điểm dừng chính là điểm cực tiểu.*

Định lý 2.4. *Giả sử D là một tập lồi, f là một hàm lồi khả vi trên tập mở chứa D. Lúc đó, điều kiện cần và đủ cho $x^* \in D$ làm hàm cực tiểu f trên D là x^* là điểm dừng của f trên D.*

Chứng minh.
$\Rightarrow$) x^* cực tiểu f trên D nên suy ra x^* là điểm dừng (định lý 2.3).

$\Leftarrow$) Do f và D lồi, lấy bất kỳ $x \in D$ và $0 < t < 1$, ta có:

$$f(x^* + t(x - x^*)) = f(tx + (1-t)x^*) \leq tf(x) + (1-t)f(x^*)$$
$$\Rightarrow \frac{f(x^* + t(x - x^*)) - f(x^*)}{t} \leq f(x) - f(x^*)$$

Qua giới hạn cho $t \to 0^+$:

$$\lim_{t \to 0^+} \frac{f(x^* + t(x - x^*)) - f(x^*)}{t} = (x - x^*)^T \nabla f(x^*) \leq f(x) - f(x^*)$$

Mà $(x - x^*)^T \nabla f(x^*) \geq 0$ *(điểm dừng)* nên suy ra $f(x) \leq f(x^*), \forall x \in D$. □

Trong các bài toán quy hoạch miền D thường gặp là tập nghiệm của hệ bất phương trình và phương trình sau:

$$g_j(x) \leq 0, (j = \overline{1, m}); h_i(x) = 0, (i = \overline{1, k}) \tag{2.7}$$

Trong đó g_j, h_i là hàm xác định trong tập mở chứa D.
Dễ thấy nếu g_j lồi, h_i affine(tuyến tính) thì D lồi, đóng.

Định nghĩa 2.5. *Đối với bài toán tối ưu* (2.6) *với miền* D *được cho như* (2.7) *thì hàm Lagrange được định nghĩa:*

$$L(x,\lambda,\mu) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x) + \sum_{i=1}^{k} \mu_i h_i(x).$$

Sử dụng hàm Lagrange ta thu được điều kiện cần (và đủ nếu là quy hoạch lồi) tối ưu.

Khái niệm điều kiện chính quy:

Cho x^0 là điểm chấp nhận được của (2.7). Giả sử các hàm g_j, h_i của (2.7) khả vi. Ký hiệu $S(x^0)$ là tập các vector d thỏa mãn hệ tuyến tính:

$$\left\langle \nabla h_j(x^0), d \right\rangle = 0, j = \overline{1,k}, \tag{2.8}$$

$$\left\langle \nabla g_i(x^0), d \right\rangle \leq 0, i \in A(x^0), \tag{2.9}$$

trong đó $A(x^0)$ là tập chỉ số i có $g_i(x^0) = 0$. Cho $x^0 \in D$ (D cho bởi (2.7)).
Ta nói rằng : **điều kiện chính quy** được thỏa mãn tại x^0 nếu $\overline{D}(x^0) = S(x^0)$

Bổ đề 2.1. $\forall x^0 \in D$ *có* $\overline{D}(x^0) \subset S(x^0)$

Chứng minh. Cho $d \in D(x^0)$, nếu (phản chứng) $d^T \nabla g_i(x^0) > 0, (i \in A(x^0)$ thì do g_i khả vi: $g_i(x^0 + td) > g_i(x^0) = 0, \forall t$ đủ nhỏ
$\Rightarrow$ trái với giả thiết $d \in D(x^0)$.
Tương tự, $\forall j$ ta phải có $d^T \nabla h_j(x^0) = 0$ vì nếu ngược lại thì $h_j(x^0 + td) \neq 0$
$\Rightarrow d \notin D(x^0)$
$\Rightarrow$ mâu thuẫn.
Vậy $D(x^0) \subset S(x^0)$. Do $S(x^0)$ đóng $\Rightarrow \overline{D}(x^0) \subset S(x^0)$. □

Định lý 2.5. *(Định lý Kuhn-Tucker)*
Giả sử các hàm $f, g_j, h_i(j = \overline{1,m}, i = \overline{1,k}$ *khả vi liên tục trên tập mở chứa* D. *Cho* x^* *là cực tiểu địa phương của bài toán* (2.7) *và tại đó* $\overline{D}(x^0) = S(x^0)$ *(điều kiện chính quy được thỏa mãn).*
Lúc đó tồn tại các vector $\lambda^* = (\lambda_1^*, \lambda_2^*, \ldots, \lambda_m^* \geq 0$ *và* $\mu^* = (\mu_1^*, \mu_2^*, \ldots, \mu_k^*)$ *sao cho:*

$$\nabla f(x^*) + \sum_{j=1}^{m} \lambda_j^* \nabla g_j(x^*) + \sum_{i=1}^{k} \mu_i^* \nabla h_i(x^*) = 0. \tag{2.10}$$

$$\lambda_j^* g_j(x^*) = 0, \forall j = 1, \ldots, m. \tag{2.11}$$

Nếu (2.6) *là quy hoạch lồi, tức là* $f, g_j(j = \overline{1,m})$ *là các hàm lồi và* $h_i(i = \overline{1,k})$ *là hàm affine và thỏa mãn điều kiện chính quy thì* (2.10) *và* (2.11) *cũng là điều kiện đủ để* $x^* \in D$ *là lời giải của* (2.6).

Chứng minh. Dùng khai triển Taylor

$$f(x^* + \lambda d) = f(x^*) + \langle \nabla f(x^*), \lambda d\rangle + r(\lambda d)$$
$$\Rightarrow \langle \nabla f(x^*), \lambda d\rangle \geq 0, \forall d \in D(x^*).$$

Do x^* chính quy ($\overline{D}(x^*) = S(x^*)$) nên $\langle \nabla f(x^*), \lambda d\rangle \geq 0, \forall d \in S(x^*)$.
Áp dụng bổ đề Farkas với ma trận A có các dòng:

$$-\nabla g_j(x^*), j \in A(x^*), \nabla h_i(x^*), -\nabla h_i(x^*), i = \overline{1, k}.$$

chú ý điều kiện (2.8) $\langle \nabla h_i(x^*), \lambda d\rangle = 0$ *tương đương với* $\langle \nabla h_i(x^*), \lambda d\rangle \geq 0$ và $\langle \nabla h_i(x^*), \lambda d\rangle \leq 0$.

Ta có các số thực $\lambda_j \geq 0, j \in A(x^*)$ và $\alpha_i \geq 0, \beta_i \geq 0, i = \overline{1, k}$ sao cho:

$$\nabla f(x^*) + \sum_{j \in A(x^*)} \lambda_j \nabla g_j(x^*) + \sum_{i=1}^{k} (\alpha_i - \beta_i) \nabla h_i(x^*) = 0.$$

Đặt: $\lambda_j^* = \begin{cases} \lambda_j, \forall j \in A(x^*) \\ \lambda_j = 0, \forall j \notin A(x^*) \end{cases}$, và $\mu_i^* = \alpha_i - \beta_i, \forall i = \overline{1, k}$, ta được (2.10) và (2.11).
Học viên kiểm tra kết quả (2.11)

Ngược lại, nếu (2.6) là quy hoạch lồi.
Giả sử x^* không tối ưu (phản chứng). Khi đó $\exists x \in D : f(x) < f(x^*)$.
Đặt $d = x - x^* \Rightarrow d \in D(x^*)$ ta có:

$$\langle \nabla f(x^*), d\rangle = \lim_{t \to 0} \frac{f(x^* + td) - f(x^*)}{t} < 0 \tag{2.12}$$

Mặt khác $\lambda_j g_j(x^*) = 0, \forall j$ nên $\lambda_j = 0$ nếu $j \notin A(x^*)$ (*chú ý rằng* $\overline{D}(x^*) = S(x^*)$)

Vậy:

$$\langle \lambda_j \nabla g_j(x^*), d\rangle \leq 0, \forall j \tag{2.13}$$

Với h_i, dễ dàng suy ra:

$$\langle \nabla h_i(x^*), d\rangle = 0, \forall i \;(\textit{chú ý tính chất lồi của hàm affine}) \tag{2.14}$$

Kết hợp (2.12),(2.13) và (2.14), ta được:

$$\langle \nabla f(x^*), d\rangle + \sum_{j=1}^{m} \lambda_j \langle \nabla g_j(x^*), d\rangle + \sum_{i=1}^{k} \langle \mu_i \nabla h_i(x^*), d\rangle < 0$$

Mâu thuẫn với (2.10). Vậy x^* phải là nghiệm tối ưu. □

Chú ý 2.2.1. *Các điều kiện (2.10) và (2.11) gọi là **điều kiện Kuhn-Tucker**. Các số λ^*, μ^* gọi là các nhân tử Lagrange.*

2.2.2 Đối ngẫu Lagrange

Cho bài toán ban đầu:

$$(P) \qquad \min f(x)$$

Với ràng buộc:

$$x \in X, g_j(x) \leq 0, j = \overline{1, m}$$

Từ bài toán ban đầu (P) xây dựng một bài toán tối ưu khác có dạng:

$$(D) \quad \max d(y)$$

Với ràng buộc:

$$y \in Y$$

Định nghĩa 2.6. *Ta nói D là bài toán đối ngẫu của (P) nếu với mọi điểm chấp nhận x của (P) và y của (D) ta có:*

$$f(x) \geq d(y)$$

Cặp đối ngẫu (P) và (D) giọi là chính xác nếu $\exists x^* \in X, y^* \in Y$ sao cho $f(x^*) = d(y^*)$. Tức là x^* là nghiệm của bài toán (P) và y^* là nghiệm của bài toán (D).

Đối ngẫu Lagrange của (P) được xây dựng như sau:

1. Xây dựng hàm Lagrange

$$L(x, y) = f(x) + \sum_{j=1}^{m} y_j g_j(x).$$

2. Xây dựng hàm mục tiêu của bài toán đối ngẫu (D):

$$d(y) := \inf_{x \in X} L(x, y).$$

3. Miền ràng buộc của

$$(D) : Y := \mathbb{R}^m_+$$

4. Bài toán đối ngẫu (D):

$$(D) \qquad \sup_{y \geq 0} d(y) = \sup_{y \geq 0} \inf_{x \in X} L(x, y).$$

Định lý 2.6. *Bài toán (D) là đối ngẫu của bài toán (P).*

Chứng minh. Cho x là chấp nhận của (P) và y là chấp nhận của (D). Do

$$g_j(x) \leq 0 \quad \forall j, y_j \geq 0 \Rightarrow \sum_{j=1}^{m} y_j g_j(x) \leq 0.$$

Do đó:

$$f(x) \geq L(x, y) \geq \inf_{x \in X} L(x, y) = d(y).$$

□

Định lý 2.7. *Giả sử:*

i) Bài toán (P) có nghiệm.

ii) f và $g_j(j = \overline{1, m})$ là các hàm lồi, liên tục trên tập đóng lồi X.

iii) Điều kiện Slater thỏa mãn, tức $\exists x^0 : g_j(x^0) < 0, \forall j$.

Khi đó (P) và (D) là cặp đối ngẫu chính xác.

Chứng minh. Ký hiệu: $g(x) = (g_1(x), g_2(x), \ldots, g_m(x))$. Ta nói g là lồi khi mọi tọa độ của g lồi. Xét tập

$$A := \{(t, z) \in \mathbb{R} \times \mathbb{R}^m | t > f(x), z \geq g(x), x \in X\}$$

Do f, g là tập lồi. Giả sử x^* là nghiệm của (P), khi đó $(f(x^*), 0) \notin A$. Theo định lý tách ta có: $\exists(\alpha, y) \neq 0, (\alpha, y) \in \mathbb{R} \times \mathbb{R}^m$ sao cho

$$\alpha t + \langle y, z \rangle \geq \alpha f(x^*), \forall (t, z) \in A \tag{2.15}$$

Do f, g liên tục nên (2.15) cũng đúng $\forall (t, z) \in \overline{A}$ (*bao đóng của* A), thế vào (2.15) ta có:

$$\alpha f(x) + \langle y, z \rangle \geq \alpha f(x^*), \forall x \in X \tag{2.16}$$

Chú ý rằng trong (2.16), *bên trái là hàm Lagrange của bài toán* (P), *nếu ta chỉ được* $y \geq 0$ *và* $\alpha > 0$ *(để chia 2 vế).*
Trước hết ta chứng tỏ $y \geq 0$
Giả sử tồn tại một tọa độ $y_j < 0$. Khi đó ta lấy $(t_0, z_0) \in A$, ta xây dựng điểm $(t_0, z) = (t_0, z_0 + \xi c^j)$.

Trong đó $c^j = (\overbrace{0, 0, \ldots, \underbrace{1}_{\text{thứ } j}, 0, \ldots, 0}^{m})^T$ và $\xi c^j \geq 0$.

Ta thấy ngay $(t_0, z^0) = (t_0, z_0 + \xi c^j) \in A$. Thế (t_0, z^0) vào (2.15) và cho $\xi \to +\infty$ ta được:
$-\infty \geq \alpha f(x^*)$ (hữu hạn) $\Rightarrow$ Vô lý. Vậy $y \geq 0$.

Tương tự, $\alpha \geq 0$. Hơn nữa $\alpha > 0$ vì nếu ngược lại ($\alpha = 0$), thế vào (2.16):

$$\langle y, g(x) \rangle \geq 0, \forall x \in X$$

Điều này mâu thuẫn với giả thiết Slater ($\exists x^0 \in X : \langle y, g(x^0) \rangle < 0$).
Chia 2 vế của (2.16) cho α ta có:

$$f(x) + \left\langle \frac{y}{\alpha}, g(x) \right\rangle \geq f(x^*), \forall x \in X \tag{2.17}$$

$$\Rightarrow d(\frac{y}{\alpha}) = \min_{x \in X}(f(x) + \langle y, g(x) \rangle) \geq f(x^*). \tag{2.18}$$

Áp dụng định lý (2.6):

$$f(x^*) \geq \max_{y \geq 0} d(y). \tag{2.19}$$

Kết hợp (2.18) với (2.19) ta được:

$$d(y^*) = d(\frac{y}{\alpha}) = f(x^*). \tag{2.20}$$

Tức (P) và (D) là cặp đối ngẫu chính xác. □

2.2.3 Điểm yên ngựa

Điểm yên ngựa là phương án rất hiệu quả khi nghiên cứu các điều kiện tối ưu và đối ngẫu.

Định nghĩa 2.7. *Cho $X \subset \mathbb{R}^n, Y \subset \mathbb{R}^m$ và $F : X \times Y \to \mathbb{R}$.*
Một điểm $(x^*, y^*) \in X \times Y$ được gọi là điểm yên ngựa của hàm F trên $X \times Y$ nếu:

$$F(x^*, y) \leq F(x^*, y^*) \leq F(x, y^*), \forall x \in X, \forall y \in Y.$$

Nhận xét 9. *Nếu (x^*, y^*) là điểm yên ngựa thì x^* là điểm cực tiểu của $F(., y^*)$ trên X và y^* là cực đại của hàm $F(x^*, .)$ trên Y.*

Hãy xét điểm yên ngựa của hàm Lagrange của bài toán (P):

$$L(x, y) = f(x) + \sum_{j=1}^{m} y_j g_j(x).$$

Định lý 2.8. *Điểm $(x^*, y^*) \in X \times \mathbb{R}^m_+$ là điểm yên ngựa của $L(x, y)$ trên $X \times \mathbb{R}^m_+$ khi và chỉ khi* *i)* x^* *làm cực tiểu của hàm* $L(x, y^*)$ *trên* X.

ii) $g_j(x^*) \leq 0, (j = \overline{1, m})$.

iii) $y_j^* g_j(x^*) = 0, (j = \overline{1, m})$.

Chứng minh.

i) (x^*, y^*) là điểm yên ngựa $\Rightarrow i)$ là hiển nhiên.

ii) Nếu tồn tại một $g_j(x^*) > 0$, lấy $y = \xi e^1$ với $e^1 = (1, 0, 0, \dots, 0)$.

$$\xi \geq 0 \Rightarrow \xi e^1 \in \mathbb{R}_+^m \text{ cho } \xi \to +\infty \Rightarrow L(x^*, \xi e^1) \to +\infty$$

$$\Rightarrow \text{mâu thuẫn (vì } L(x^*, \xi e^1) \leq L(x^*, y^*), \forall \xi e^1).$$

iii) $\Rightarrow$) Để ý rằng:

$$L(x^*, 0) = f(x^*) \geq f(x^*) + \left\langle y^*, \underbrace{g(x^*)}_{\leq 0} \right\rangle = L(x^*, y^*).$$

Mặt khác $L(x^*, y^*) \geq L(x^*, 0)$ nên

$$\begin{aligned} f(x^*) &\leq f(x^*) + \langle y^*, g(x^*) \rangle \\ &\Rightarrow \langle y^*, g(x^*) \rangle = 0. \end{aligned}$$

Nhưng do: $g(x^*) \leq 0, y^* \geq 0$ nên suy ra $y_j^* g_j(x^*) = 0, \forall j \in [1, m]$
$\Leftarrow$) Ngược lại, do $i)$ nên ta chỉ cần chứng minh $L(x^*, y^*) = \max_{y \geq 0} L(x^*, y)$.
Thật vậy, từ $ii), iii)$ có ngay

$$L(x^*, y^*) = f(x^*) \geq f(x^*) + \langle y, g(x^*) \rangle = L(x^*, y), \forall y \geq 0.$$

Suy ra y^* làm cực đại hàm $L(x^*, .)$ trên $\mathbb{R}_+^m$.

□

Định lý 2.9. *Nếu* (x^*, y^*) *là điểm yên ngựa của* $L(x, y)$ *trên* $X \times \mathbb{R}_+^m$ *thì* x^* *là nghiệm của* (P) *và* y^* *là nghiệm của* (D).

Chứng minh. Do (x^*, y^*) là điểm yên ngựa, theo định lý (2.8) $i), ii), iii)$ suy ra:

$$L(x^*, y^*) = f(x^*) \geq f(x) + \langle y^*, g(x) \rangle, \forall x \in X.$$

Vì $g(x) \leq 0$ (điều kiện ràng buộc của bài toán (P) nên

$$f(x^*) \geq f(x) \forall x \in X \Rightarrow x^* \in argmin(P).$$

Hơn nữa, $\forall y \geq 0$ ta có:

$$d(y) = \inf_{x\in X}[f(x) + \langle y, g(x)\rangle] \leq f(x^*) + \langle y, g(x^*)\rangle \leq f(x^*).$$

Mặt khác:

$$d(y^*) := \inf_{x\in X} L(x, y^*) = \min_{x\in X} L(x, y^*) = L(x^*, y^*) = f(x^*).$$

Suy ra y^* là nghiệm của bài toán đối ngẫu (D). □

Định lý 2.10. *Giả sử (P) là một quy hoạch lồi (X, f, g lồi) thỏa mãn điều kiện Slater. Lúc đó x^* là lời giải của (P) khi và chỉ khi tồn tại $y^* \geq 0$ để (x^*, y^*) là điểm yên ngựa của L(hàm Lagrange) trên $X \times \mathbb{R}^m_+$ và y^* là nghiệm của bài toán đối ngẫu (D).*

Chứng minh. $\Leftarrow$) Suy ra từ định lý (2.9)

$\Rightarrow$) Giả sử $x^* \in argmin(P)$. Do (P) lồi, thỏa mãn điều kiện Slater (tức $\exists x^0 \in X : g_j(x^0) < 0, \forall j = \overline{1, m}$). Do đó, theo định lý (2.7) cặp $(P), (D)$ là đối ngẫu chính xác:

$$f(x^*) = d(y^*) \text{ với } y^* \geq 0$$

Theo định nghĩa của $d(y^*)$:

$$f(x^*) = d(y^*) = \inf_{x\in X} L(x, y^*)$$

Nghĩa là x^* là điểm cực tiểu của $L(., y^*)$ (thỏa mãn điều kiện i) của định lý (2.8))
Ngoài ra

$$f(x^*) \leq f(x) + \langle y^*, g(x)\rangle, \forall x \in X$$
$$\Rightarrow g(x^*) = 0, \langle y^*, g(x^*)\rangle = 0$$

$\Rightarrow ii), iii)$ của định lý (2.8) thỏa mãn. Thế thì (x^*, y^*) là điểm yên ngựa của $L(x, y)$ trên $X \times \mathbb{R}^m_+$.
Theo định lý (2.9) y^* là nghiệm của (D). □

2.3 Bài tập chương 2

Bài tập 2.1. *Cho f tựa lồi trên tập lồi A, chứng minh rằng $\forall x, y \in A$ và $z = tx + (1-t)y, 0 \leq t \leq 1$ ta có:*

$$f(z) \leq \max\{f(x), f(y)\}.$$

Bài tập 2.2. *Chứng minh rằng f đạt cực tiểu tuyệt đối trên D nếu*

1. *D đóng, $D \neq \emptyset$.*

2. *f nửa liên tục dưới trên D.*

3. *$\exists t \in \mathbb{R}$ để tập mức $\{x \in D | f(x) \leq t\} \neq \emptyset$ và bị chặn.*

Bài tập 2.3. *Viết điều kiện Kuhn-Tucker tại $x^* = 1$ của bài toán tối ưu .*

$$\min\{-x^2 : -1 \leq x \leq 1\}.$$

Bài tập 2.4. *Tìm điểm yên ngựa của hàm số*

$$K(x, y) = x.y \quad \text{trên} \quad \mathbb{R} \times \mathbb{R}.$$

Bài tập 2.5. *Chứng minh rằng hàm Lagrange của bài toán tối ưu*

$$\min\{-x^2 : -1 \leq x \leq 1\},$$

không có điểm yên ngựa trên $\mathbb{R}_+ \times \mathbb{R}_+^2$.

Bài tập 2.6. *Viết bài toán đối ngẫu của quy hoạch toàn phương*

$$\min\{\frac{1}{2}x^T Q x + q^T x\},$$

với ràng buộc:

$$Ax \leq b, x \geq 0$$

Trong đó Q và A là các ma trận $n \times n, m \times n$.

Chương 3

PHƯƠNG PHÁP CÓ THỂ VÀ PHƯƠNG PHÁP TUYẾN TÍNH HÓA

Phương pháp có thể là phương pháp rất hiệu quả để giải các bài toán tối ưu phi tuyến.

3.1 Hướng chấp nhận tụt

Định nghĩa 3.1. *Xét bài toán*

$$\min\{f(x)|x \in D\}. \tag{3.1}$$

*Ta nói rằng một vector $d \neq 0$ là **hướng chấp nhận tụt** của bài toán* (3.1) *tại điểm $x \in D$ nếu:*
$d \in D(x)$ và theo hướng d, hàm $f(x)$ giảm.
Xét bài toán

$$\min f(x) \tag{3.2}$$

với ràng buộc:

$$x \in D := \{x \in \mathbb{R}^n | g(x) \leq 0\},$$

trong đó $g(x)$ lồi, xác định trên $\mathbb{R}^n$.

Chú ý 3.1.1. *g lồi $\Rightarrow$ g liên tục tại mọi $x \in int(D)$ và luôn giả thiết D là tập compact.*

Mệnh đề 3.1. *Giả sử D được cho bởi* (3.2)*, với g lồi, liên tục trên $\mathbb{R}^n$. Giả sử $g(x^k) = 0$. Khi đó d là hướng chấp nhận được tại x^k khi và chỉ khi bài toán* *tối ưu một biến*

$$\min_{0\le t\le\lambda} g(x^k+td) \tag{3.3}$$

có nghiệm $t^k > 0$.

Chứng minh.
$\Leftarrow$) Nếu t^k là nghiệm của (3.3) $\Rightarrow g(x^k+t^kd) \le g(x^k) = 0$.
Do g lồi suy ra:

$$g(x^k+t^kd) \le 0, \forall t \in [0,t^k] \Rightarrow d \in D(x^k).$$

$\Rightarrow$) Vì $d \in D(x^k) \Rightarrow \exists \lambda > 0 : g(x^k+t^kd) \le 0, \forall t \in [0,\lambda]$
Và do $g(x^k) = 0$ nên suy ra bài toán (3.3) có ít nhất một nghiệm $t^k > 0$. □

3.2 Phương pháp FRANK-WOLFE (phương pháp hướng có thể)

Xét bài toán quy hoạch tuyến tính

$$\min f(x) \tag{3.4}$$

với ràng buộc:

$$x \in D := \{x|Ax \le b, x \ge 0\}$$

trong đó f khả vi liên tục trên D, A là ma trận $(a_{ij})_{m\times n}, b \in \mathbb{R}^m$ và D bị chặn.
*** Thuật toán hướng chấp nhận tụt (Frank-Wolfe)**

1. **Bước 1:** Dùng quy hoạch tuyến tính (nếu cần) tìm điểm xuất phát $x^k \in D, k = 0$.

2. **Bước 2:** Tính $\nabla f(x^k)$

   a. Nếu $\nabla f(x^k) = 0$ dừng.

   b. Nếu $\nabla f(x^k) \ne 0$, giải quy hoạch tuyến tính:

   $$(L(x^k)) \quad \min\{\left\langle \nabla f(x^k), x - x^k\right\rangle |x \in D\}$$

   thu được lời giải là u^k (một đỉnh của D).

      i. Nếu $\left\langle \nabla f(x^k), u^k - x^k\right\rangle \ge 0 \Rightarrow$ dừng.

ii. Nếu $\langle \nabla f(x^k), u^k - x^k \rangle < 0 \Rightarrow d^k := u^k - x^k \neq 0$ là hướng chấp nhận tụt (Bài tập: Áp dụng khai triển Taylor cho f).
Theo hướng d^k tìm $x^{k+1} \in D$ sao cho $f(x^{k+1}) < f(x^k)$ thông qua giải bài toán một biến

$$\min\{f(x^k + td^k), 0 \leq t \leq 1\}$$

được nghiệm $t_k > 0$.
Khi đó lấy $x^{k+1} = x^k + t_k d^k$ thế $x^k := x^{k+1}$, quay lại bước 2.

Thuật toán này hội tụ theo định lý sau:

Định lý 3.1. *Thuật toán giải bài toán* (3.4) *cho kết quả:*

i) $f(x^{k+1}) < f(x^k), \forall k.$

ii) Nếu thuật toán dừng ở x^k thì x^k là điểm dừng của f trên D. Nếu thuật toán vô hạn thì mọi điểm tụ của dãy $\{x^k\}$ đều là điểm dừng.

iii) Nếu f lồi thì mọi điểm dừng đều là nghiệm của (3.4). *Ngoài ra dãy $\{f(x^k)\}$ hội tụ đến giá trị tối ưu của f^* và*

$$0 \leq f(x^k) - f^* \leq \langle \nabla f(x^k), x^k - u^k \rangle, \forall k$$

Chứng minh.
i) Dễ thấy vì theo cách xây dựng, thì d^k là hướng chấp nhận tụt.
ii) Giả sử thuật toán kết thúc tại bước k, nghĩa là

$$\langle \nabla f(x^k), u^k - x^k \rangle \geq 0$$

Do $u^k \in argminL(x^k)$ nên

$$\begin{aligned} \langle \nabla f(x^k), x - x^k \rangle &\geq \min\{\langle \nabla f(x^k), x - x^k \rangle : x \in D\} \\ &= \langle \nabla f(x^k), u^k - x^k \rangle \geq 0, \forall x \in D \end{aligned}$$

Từ đó suy ra x^k là điểm dừng.

Giả sử thuật toán vô hạn.
Gọi x^* là điểm tụ của dãy $\{x^k\}$.
Do D compact nên tồn tại dãy con $\{x^{k_j}\} \to x^*$.
Gọi $\{u^{k_j}\}$ là nghiệm của $L(x^{k_j})$.
Do tập đỉnh của D hữu hạn $\Rightarrow$ để cho gọn, ta coi $u^{k_j} = u^*, \forall j$ (nếu không ta sẽ làm phép chứng minh dưới đây với nhiều nhất số lần bằng số đỉnh của D).

Theo tính đơn điệu của dãy $\{f(x^k)\}$ và cách xác định x^{k_j}, u^* thì $\forall t \in (0,1)$ ta có:

$$f(x^{k_{j+1}}) \leq f(x^{k_j+1}) \leq f(x^{k_j} + t(\underbrace{u^* - x^{k_j}}_{h.ch.nh\ tụt})).$$

Cho $j \to +\infty$, do f liên tục:

$$f(x^*) \leq f(x^* + t(u^* - x^*)), \forall t \in (0,1).$$

Qua giới hạn (để ý $f(x^* + t(u^* - x^*)) - f(x^*) \geq 0$) ta có:

$$\lim_{t\to 0^+} \frac{f(x^* + t(u^* - x^*)) - f(x^*)}{t} = \left\langle \nabla f(x^{k_j}), x - x^* \right\rangle, \forall x \in D.$$

Vậy x^* là điểm dừng.
iii) Nếu f lồi suy ra:

$$f(x) - f(x^*) \geq \langle \nabla f(x^*), x - x^*)) \rangle \geq 0, \forall x \in D$$
$$\Rightarrow x^* \in argmin(3.4)$$

□

3.3 Phương pháp cắt KELLEY (phương pháp tuyến tính hóa)

Xét quy hoạch lồi:

$$\min\{c^T x : \ \ x \in D\}, \tag{3.5}$$

trong đó $c \in \mathbb{R}^n, D$ là tập lồi compact cho dưới dạng

$$D := \{x \in \mathbb{R}^n | g(x) \leq 0\},$$

với g là hàm lồi trên $\mathbb{R}^n$.

Chú ý 3.3.1. *Trước đây D thường có dạng $D := \{x \in \mathbb{R}^n | g_j(x) \leq 0, j = \overline{1,m}\}$ với g_j lồi $\forall j \in [1,m]$. Bằng cách đặt:*

$$g(x) = \max_j \{g_j(x) | j = \overline{1,m}\} \Rightarrow g \text{ lồi }.$$

Và tập:

$$\{x \in \mathbb{R}^n | g_j(x) \leq 0, j = \overline{1,m}\} = \{x \in \mathbb{R}^n | g(x) \leq 0\}$$

Ta đều có thể đưa về biểu diễn D của (3.5) *(Bài tập)*

*** Thuật toán cắt Kelley**

1. **Bước chuẩn bị:** Xây dựng đa diện lồi $D_0 \supset D$. Cho $k = 0$

2. **Bước** $k(k = 0, 1, \ldots)$**:** Giải quy hoạch tuyến tính

$$(L_k) \qquad \min\{c^T x | x \in D_k\}$$

Tìm được nghiệm x^k.

a. Nếu $x^k \in D \Rightarrow x^k \in argmin(3.5) \Rightarrow$ dừng.

b. Nếu $x^k \notin D$, xây dựng siêu phẳng (hàm affine)

$$l_k(x) := \left\langle a^{k+1}, x \right\rangle - b_{k+1} \qquad \textit{sao cho thỏa mãn :}$$

$$\left\langle a^{k+1}, x^k \right\rangle - b_{k+1} \leq 0, \forall x \in D$$
$$\left\langle a^{k+1}, x^k \right\rangle - b_{k+1} = g(x^k) \geq 0 \qquad (3.6)$$

Đặt $D_{k+1} := D_k \bigcap \{x | l_k(x) \leq 0\}$, gán $k := k + 1$ rồi quay về bước k.

Chú ý 3.3.2. *Để xây dựng được siêu phẳng $l_k(x)$ thỏa mãn* (3.6) *thông thường lấy $a^{k+1} \in \partial g(x^k)$ và đặt*

$$l_k(x) = \left\langle a^{k+1}, x - x^k \right\rangle + g(x^k).$$

Định nghĩa 3.2. *Một dãy phiếm hàm affine $\{l_k(x) := \left\langle a^{k+1}, x \right\rangle - b_{k+1}\}$ được gọi là bị chặn đều nếu tồn tại một hằng số c sao cho :*

$$\left\| a^k \right\| \leq c, \forall k.$$

Định lý 3.2.

i) Nếu thuật toán Kelley dừng lại ở bước thứ k thì x^k là nghiệm tối ưu của bài toán (3.5)*.*

ii) Nếu thuật toán vô hạn, các siêu phẳng cắt bị chặn đều thì mọi điểm tụ của dãy $\{x^k\}$ đều là nghiệm tối ưu của bài toán (3.5) *và dãy $\{f(x^k)\}$ đơn điệu tăng đến trị tối ưu f_* của bài toán* (3.5)*.*

Chứng minh.

i) Theo cách xây dựng siêu phẳng ta có

$$D \subset D_k, D_{k+1} \subset D_k, \forall k.$$

Giả sử thuật toán dừng tại bước thứ k ta suy ra

$$x^k \in D, D \subset D_k \text{ và } x^k \text{ là nghiệm của bài toán } (L_k)$$

Do đó $x^k \in argmin(3.5)$.

ii) Giả sử thuật toán vô hạn. Do $x^k \in D_0, \forall k$, và D_0 compact $\Rightarrow \{x^k\}$ có dãy con $\{x^{k_j}\}$ hội tụ đến x^*. Để dễ ký hiệu, ta gọi dãy con hội tụ đó là $\{x^k\}$.
Với $k > j$, theo (3.6) ta có:

$$\begin{aligned}
\langle a^{j+1}, x^k \rangle - b_{j+1} &\le 0 \\
\Rightarrow \langle a^{j+1}, x^k - x^j \rangle + g(x^j) &\le 0 \\
\Rightarrow g(x^j) \le \left\| \langle a^{j+1}, x^k - x^j \rangle \right\| &\le \left\| a^{j+1} \right\| . \left\| x^k - x^j \right\| \\
&\le c. \left\| x^k - x^j \right\| .
\end{aligned}$$

Do $g(x^j) > 0, \forall j$ và $\left\| x^k - x^j \right\| \to 0$ khi $k, j \to +\infty$
$\Rightarrow g(x^j) \to 0$.
Tức là $g(x^*) = 0 \Rightarrow x^* \in D$. (*)
Mặt khác : $f(x^k) \le f_*, \forall k$,
do f liên tục suy ra $f(x^*) \le f_*$. Kết hợp điều này với (*) ta đi đến kết quả $x^* \in argmin(3.5)$.
Hơn nữa $f(x^{k+1}) \ge f(x^k), \forall k \Rightarrow f(x^k) \nearrow f_*$

□

3.4 Bài tập chương 3

Bài tập 3.1. *Cho $D := \{x : g(x) \le 0\}$ với g lồi trên $\mathbb{R}^n$. Giả sử $g(u^0) < 0$ và $g(x^k) > 0$. Cho u^k là điểm trên đoạn (u^O, x^k) sao cho $g(u^k) = 0$. Lấy a^{k+1} là dưới vi phân của g tại u^k và xây dựng siêu phẳng:*

$$l_k(x) := \langle a^{k+1}, x - u^k \rangle .$$

Chứng tỏ rằng siêu phẳng $l_k(x)$ thỏa mãn các tính chất trong thuật toán Kelley:

1. *$l_k(x) \le 0, \forall x \in D$,*

2. *$l_k(x^k) > 0$,*

3. *$l_k(x^k) \le g(x^k)$,*

4. *Dãy $\{l_k\}$ bị chặn đều nếu D bị chặn.*

Biết rằng: Ảnh của tập compact qua ánh xạ dưới vi phân của một hàm lồi trên $\mathbb{R}^n$ cũng là một tập compact.

Bài tập 3.2. *Cho $x^0 = (0, 0, 4)^T$. Tính 2 bước đầu tiên của thuật toán Frank-Wolfe cho bài toán*

$$\min\{0,5x_1^2 + 2x_1x_2 + 2,5x_2^2 + x_2x_3 + x_3^2 - 2x_1 - 4x_2 - 6x_3\},$$

với ràng buộc

$$\begin{cases} x_1 + x_2 + x_3 \leq 5 \\ 2x_1 + x_3 \leq 6 \\ x_1, x_2, x_3 \geq 0 \end{cases}$$

Chương 4

PHƯƠNG PHÁP HÀM PHẠT ĐIỂM TRONG, ĐIỂM NGOÀI

4.1 Phương pháp hàm phạt điểm trong

Xét bài toán

$$\min f(x), \tag{4.1}$$

$$\textbf{s.t}$$
$$D := \{x | g_j(x) \leq 0, j = \overline{1, m}, \quad g_j \textit{ liên tục trên } \mathbb{R}^n\}$$
Giả sử D compact.

Định nghĩa 4.1. *Hàm $p : \mathbb{R}^n \to \mathbb{R}$ được gọi là hàm phạt điểm trong trên miền D nếu thỏa mãn:*

a) *p liên tục trên tập hợp*

$$D^0 := \{x \in \mathbb{R}^n | g_j(x) < 0, \forall j = \overline{1, m}\}.$$

b) *$\forall \{x^k\} \subset D^0, x^k \xrightarrow{k \to +\infty} x \notin D^0$ ta có*

$$\lim_{k \to +\infty} p(x^k) = +\infty.$$

Hai hàm phạt điểm trong nổi tiếng (Fiacco, McCormick):

$$p(x) = -\sum_{j=1}^{m} \log(-g_j(x) \quad \text{hoặc} \quad p(x) = -\sum_{j=1}^{m} \frac{1}{g_j(x)}.$$

Xây dựng bài toán phụ (B_t)

1. Xây dựng hàm tham số một biến $s(t) : \mathbb{R}_+ \setminus \{0\} \to \mathbb{R}_+ \setminus \{0\}$ sao cho :

i) $s(t)$ liên tục $\forall t > 0$ và $s(t) \to 0$ khi $t \to +\infty$.

ii) $s(t)$ đơn điệu giảm $(s(t) > s(t'), \forall t' > t > 0)$.

Ví dụ: $s(t) = 1/t, \quad s(t) = 1/t^2$.

2. Xây dựng bài toán phụ

Đặt hàm

$$F(x,t) := f(x) + s(t)p(x), \tag{4.2}$$

với miền $t > 0$ xây dựng bài toán phụ không ràng buộc

$$(B_t) \quad \min_x \{F(x,t) \mid x \in \mathbb{R}^n\}.$$

Mệnh đề 4.1. *Giả sử các điều kiện a, b, i, ii thỏa mãn và bài toán (B_t) có nghiệm $\forall t > 0$. Khi đó, nếu $0 < t_1 < t_2$ và x^i là nghiệm của $B_{t_i} (i = 1, 2)$. Ta có :*

1. $p(x^1) \leq p(x^2)$,

2. $f(x^1) \geq f(x^2)$.

Chứng minh. Để đơn giản ký hiệu, đặt $s(t_i) = s_i, p(x^i) = p_i, f(x^i) = f_i, i = \overline{1,2}$. Khi đó :

$$f_1 + s_1p_1 \leq f_2 + s_1p_2 \quad \text{vì } x^1 \text{ là argmin}(Bt_1), \tag{4.3}$$

$$f_2 + s_2p_2 \leq f_1 + s_2p_1 \quad \text{vì } x^2 \text{ là argmin}(Bt_2). \tag{4.4}$$

Cộng 2 vế và ước lược:

$$s_1p_1 + s_2p_2 \leq s_1p_2 + s_2p_1$$
$$\Rightarrow (s_1 - s_2)(p_1 - p_2) \leq 0.$$

Do $t_1 < t_2 \Rightarrow s_1 > s_2$ (đđ giảm) $\Rightarrow p(x^1) \leq p(x^2)$
$\Rightarrow$ (theo (4.4)) $f(x^1) \geq f(x^2)$. □

Định lý 4.1. *Nếu dãy $\{t_k\}$ đơn điệu tăng đến $+\infty$ và x^k là nghiệm của (B_{t_k}) thì dãy $\{f(x^k)\}$ hội tụ giảm đến trị tối ưu f_* của bài toán (4.1). Ngoài ra mọi điểm tụ của dãy x^k đều là nghiệm của bài toán (4.1).*

Chứng minh. Theo mệnh đề (4.1) thì dãy $\{f(x^k)\}$ đơn điệu giảm, do đó dãy $\{f(x^k)\}$ hội tụ. (vì sao ?)

1) Xét trường hợp nghiệm của (4.1) $x^* \in D^0$:

Do x^k là nghiệm của (B_{t_k}) (để ý $x^k \in D^0$, nếu $x^k \notin D^0$ thì $F = +\infty$) nên:

$$f(x^k) + s(t_k)p(x^k) \leq f(x^*) + s(t_k)p(x^*). \tag{4.5}$$

Do D compact nên $\{x^k\}$ hội tụ đến u^* nào đó.

1.a) Nếu $u^* \in D^0$ qua giới hạn $(s(t_k) \to 0, p(x^*), p(u^*)$ hữu hạn) suy ra:

$$f(u^*) \leq f(x^*) \Rightarrow u^* \in argmin(4.1).$$

1.b) Nếu $u^* \notin D^0$ thì $p(x^k) \xrightarrow{k \to +\infty} +\infty$ nên $\exists K_1$ sao cho $s(t_k)p(x^k) \geq 0, \forall k \geq K_1$. Khi đó, từ 4.5 ta có:

$$f(x^k) \leq f(x^*) + s(t_k)p(x^*), \forall k \geq K_1.$$

Qua giới hạn ta được

$$\lim_{k \to +\infty} f(x^k) \leq f(x^*).$$

Mặt khác do $x^k \in D \Rightarrow f(x^k) \geq f(x^*), \forall k \geq K_1$ nên ta có

$$\lim_{k \to +\infty} f(x^k) \geq f(x^*).$$

Từ đó suy ra

$$\lim_{k \to +\infty} f(x^k) = f(x^*).$$

2) Trường hợp nghiệm tối ưu của bài toán (4.1)$x^* \notin D^0$

Đặt $\beta = \lim_{k \to +\infty} f(x^k)$. Ta suy ra $f^* \leq \beta$ (do $f(x^k) \geq f(x^*), \forall k$)

Nếu thật sự $f^* < \beta$ thì do f liên tục trên D,

nên suy ra $\exists u \in D^0 : f^* < f(u) < \beta$. Từ bất đẳng thức:

$$f(x^k) + s(t_k)p(x^k) \leq f(u)s(t_k)p(u),$$

tương tự **1.b)** ta suy ra: $f(x^k) \leq f(u), \forall k \geq K_1$. Điều này mâu thuẫn với

$$f(x^k) \geq \beta > f(u), \forall k.$$

Vậy $\beta = f(x^*)$.

$\square$

4.2 Phương pháp hàm phạt điểm ngoài

Định nghĩa 4.2. *Hàm số $p : R^n \to R$ được gọi là hàm phạt điểm ngoài miền D của bài toán* (4.1) *nếu thỏa mãn:*

a) *p liên tục trên $\mathbb{R}^n$.*

b) *$p(x) = 0, \forall x \in D, p(x) > 0, \forall x \notin D$.*

Ví dụ 4.2.1. *Với g_j của bài toán* (4.1)*, ta có 2 hàm phạt điểm ngoài*

$$p(x) = \sum_{j=1}^{m} \max(0, g_j(x)), \tag{4.6}$$

$$p(x) = \sum_{j=1}^{m} [\max(0, g_j(x))]^2. \tag{4.7}$$

*** Xây dựng bài toán phụ** (P_t)

1) Xây dựng hàm $r : \mathbb{R}_+ \setminus \{0\} \to \mathbb{R}_+ \setminus \{0\}$.

i) $r(t)$ liên tục tại $\forall t > 0, r(t) \to 0$ khi $t \to +\infty$.

ii) $r(t)$ đơn điệu tăng: $t > t' > 0 \to r(t) > r(t') > 0$.

2) Với mỗi t cố định xây dựng bài toán phụ (P_t):

$$(P_t) \qquad \min\{F(x,t) := f(x) + r(t)p(x) \mid x \in \mathbb{R}^n\}.$$

Ta có mệnh đề và cách chứng minh tương tự mệnh đề (4.1).

Mệnh đề 4.2. *Giả sử (P_t) có nghiệm $\forall t > 0$. Khi đó nếu x là nghiệm của $(P_{t_i}), (i = 1, 2)$ và $0 < t_1 < t_2$ thì*

1) *$p(x^1) \geq p(x^2)$.*

2) *$f(x^1) \leq f(x^2)$.*

Định lý 4.2. *Nếu dãy $\{t_k\}$ đơn điệu tăng đến $+\infty$ và x^k là nghiệm của (P_{t_k}), thì dãy số $\{f(x^k)\}$ hội tụ tăng đến giá trị tối ưu f_* của* (4.1)*. Ngoài ra mọi điểm tụ của dãy $\{x^k\}$ đều là nghiệm của* (4.1)*.*

Chứng minh. Gọi $x^* \in argmin(4.1), x^k \in argmin(P_{t_k})$ suy ra $p(x^*) = 0$ và

$$f(x^k) + r(t_k)p(x^k) \leq f(x^*) + r(t_k)p(x^*) = f_*. \tag{4.8}$$

Ta sẽ chứng minh mọi điểm tụ u^* của dãy $\{x^k\}$ phải thuộc D.
Thật vậy, giả sử $u^* \notin D \Rightarrow p(u^*) > 0$. Cho $r(t_k) \to +\infty$ khi đó, $\exists k$ đủ lớn để:

$$f(u^*) + r(t_k)p(u^*) > f_* \quad \Rightarrow \quad \text{mâu thuẫn với (4.8).}$$

Do đó, $u^* \in D$.

Bây giờ ta chứng minh $u^* \in argmin(4.1)$:
Từ (4.8) ta suy ra $f(x^k) \leq f_*, \forall k$.
Vì f liên tục, qua giới hạn ta được: $f(u^*) \leq f_*$.
Vì $u^* \in D \Rightarrow f(u^*) = f_* \Rightarrow u^* \in argmin(4.1)$

Cuối cùng, ta chứng minh $f(x^k) \nearrow f_*$
Ta có: $f(x^k) \leq f_*, \forall k$ và theo mệnh đề (4.2) $\Rightarrow f(x^k) \nearrow f_*$ khi $t_k \to +\infty$ □

Chú ý 4.2.1.

1. *Nếu tập $\{x \mid f(x) + r(t^0)p(x) \leq c\}$ compact với mọi hằng số C thì bài toán không ràng buộc (P_t) có nghiệm.*

2. *Nếu $g_j, j = \overline{1,m}$ lồi thì (4.6), (4.7) lồi.*

4.3 Bài tập chương 4

Bài tập 4.1. *Cho bài toán*

$$(P) \qquad \min\{f(x) \mid x \in S, g(x) \leq 0\},$$

trong đó, f, g hàm lõm trên S, S là đa diện lồi bị chặn và $g(x) \geq 0, \forall x \notin S$. Chứng minh rằng $\exists t^* \in (0, +\infty)$ sao cho mọi nghiệm của bài toán

$$(P_t) \qquad \min\{f(x) + tg(x) \mid x \in S\},$$

đều là nghiệm của $(P), \forall t \geq t_*$.
(Gợi ý: Cực tiểu của hàm lõm đạt trên đỉnh của S.)

Chương 5

TỐI ƯU ĐA MỤC TIÊU

5.1 Điểm hữu hiệu và bài toán tối ưu đa mục tiêu

5.1.1 Điểm hữu hiệu

Định nghĩa 5.1. *Cho nón lồi* $R^p := \{x \in R^p \mid x \leq 0\}, x, y \in R^p$, *ta nói :*

$$x \text{ nhỏ hơn hoặc bằng } y \ (x \leq y) \text{ khi và chỉ khi } x - y \in R^p_-.$$

(Tức là : $x_i \leq y_i \ \ \forall i = \overline{1, p}$*)*

Định nghĩa 5.2. *Cho* $Y \subseteq R^p$, *ta nói :* $y^* \in Y$ *là điểm hữu hiệu hay điểm Pareto của* Y *nếu :*

$$\nexists\, y \in Y \text{ để } y \leq y^* \text{ và } y \neq y^*$$

Nhận xét 10. *Về mặt hình học nếu* y^* *là điểm Pareto của* Y *thì nón có đỉnh tại* y^* *và có phương của các cạnh trùng với phương của các cạnh của nón* R^p_- *không chứa một điểm* $y \in Y, y \neq y^*$ *nào.*

5.1.2 Bài toán tối ưu đa mục tiêu

Cho $R^n \supset D \neq \emptyset, f : R^n \to R^n, Y := f(D)$ là ảnh của D qua f.
Bài toán tối ưu đa mục tiêu được viết như sau:

$$\text{Min}\{f(x) \mid x \in D\}. \tag{5.1}$$

Bài toán này được hiểu là : hãy tìm một tập (có thể là một điểm) các điểm $x^* \in D$ sao cho $y^* := f(x^*)$ là điểm Pareto của Y. Khi đó x^* được gọi là nghiệm tối ưu của bài toán (5.1) hay điểm hữu hiệu của f trên D.

Chú ý 5.1.1.

1. *Khi* $p = 1$ *thì* x^* *là điểm làm cực tiểu tuyệt đối* f *trên* D.

2. *Nếu* D *là khúc lồi,* f *affine trên* D *(mỗi* f_i *là affine) thì* (5.1) *được gọi là bài toán tuyến tính đa mục tiêu.*

5.2 Sự tồn tại và tính chất của nghiệm tối ưu của bài toán tối ưu đa mục tiêu

Mệnh đề 5.1. *Cho* $\lambda \in R^p$ *là vector dương* $\lambda > 0$. *Khi đó mọi nghiệm tối ưu của bài toán một mục tiêu*

$$\min\{\lambda^T f(x) \mid x \in D\} \tag{5.2}$$

đều là điểm hữu hiệu của f *trên* D.

Chứng minh. Gọi $x^* \in \text{argmin}(5.2)$. Giả sử $x^* \notin \text{argmin}(5.1)$ (tức x^* không phải là điểm hữu hiệu của f trên D). Suy ra :

$$\exists x' \in D : \quad f(x') \leq f(x^*) \quad \text{và } f(x') \neq f(x).$$

Kết hợp với $\lambda > 0$ (tức là $\lambda_i > 0, \forall i = \overline{1,p}$), ta có:

$$\lambda^T f(x') < \lambda^T f(x^*).$$

Điều này mâu thuẫn với $x^* \in \text{argmin}(5.2)$.
Vậy, $x^* \in \text{argmin}(5.1)$ □

Hệ quả 5.1.1. *Nếu* D *compact và* $f(x)$ *nửa liên tục dưới thì bài toán tối ưu đa mục tiêu* (5.1) *có nghiệm tối ưu.*

Định lý 5.1. *Giả sử* (5.1) *là bài toán quy hoạch lồi. Khi đó mọi* $u \in \text{argmin}$ (5.1) *đều tồn tại* $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_p) \geq 0$ *sao cho* $u \in \text{argmin}$ *của bài toán* $\min\{\lambda^T f(x) \mid x \in D\}$.

Chứng minh.
Đặt $C := cov\,(K := \{y \in R^p \mid y = f(x) - f(u), x \in D\})$.

1) Ta chứng minh $C \cap R^p = \{0\}$:
Trước hết, $C \neq \emptyset$ vì $\{0\} \in K$ (lấy $x \equiv u$).
Lấy bất kỳ $y \in C$, vì C là bao lồi của K nên suy ra :

$$\exists y^1, y^2 \in K : \quad y = ty^1 + (1-t)y^2, \quad 0 \leq t \leq 1.$$

Và $\exists x^1, x^2 \in D$:

$$y^i = f(x^i) - f(u), \quad i = \overline{1,2}. \tag{5.3}$$

Lấy $x = tx^1 + (1-t)x^2 \Rightarrow x \in D$ (D lồi).

Do f lồi và kết hợp với (5.3) có:

$$\begin{aligned} f(x) - f(u) &\leq tf(x^1) + (1-t)f(x^2) - f(u) \\ &= t(f(x^1) - f(u)) + (1-t)(f(x^2) - f(u)) \\ &= ty^1 + (1-t)y^2 = y \end{aligned}$$

Tức là $f(x) - f(u) \leq y$

Giả sử $y \in C \cap R^p_-$ và $y < 0$, ta suy ra

$$f(x) - f(u) \leq 0 \quad \text{và} \quad f(x) \neq f(u)$$

Do đó u không phải là điểm hữu hiệu của f trên D. Điều này mâu thuẫn với giả thiết u là điểm hữu hiệu của f trên D.

Vậy: $C \cap R^p_- = \{0\}$.

2) Bây giờ ta sẽ chứng minh $u \in \operatorname{argmin}\{\lambda^T f(x) \mid x \in D\}$:

Do C lồi, R^p_- lồi và $C \cap R^p_- = \{0\}$ nên theo định lý tách $\exists \lambda \neq 0 \quad (\lambda = \lambda_1, \lambda_2, \ldots, \lambda_p)$:

$$\lambda^T y \leq 0 \quad \forall y \in R^p_- \tag{5.4}$$

$$\lambda^T y \geq 0 \quad \forall y \in K \tag{5.5}$$

Bằng cách chia cho $\sum_{i=1}^p \lambda_i$ ta coi $\sum_{i=1}^p \lambda_i = 1$.

Từ (5.4) $\Rightarrow \lambda^T \geq 0$.

Từ (5.5) và định nghĩa của K suy ra $\lambda^T y = \lambda^T(f(x) - f(u)) \geq 0, \forall x \in D$.

Do đó, u là nghiệm tối ưu của bài toán (5.2):

$$\min\{\lambda^T f(x) \mid x \in D\}.$$

□

Bây giờ, ta xét bài toán sau:

Bài Toán 5.2.1. *Cho quy hoạch lồi*

$$\operatorname{Min} f(x) \tag{5.6}$$

$$x \in D = \{x \in \mathbb{R}^n \mid g(x) \leq 0\}$$

trong đó $f : \mathbb{R}^n \to \mathbb{R}^p, \quad g : \mathbb{R}^n \to \mathbb{R}^m$ *lồi.*

Định lý 5.2. *Giả sử quy hoạch lồi (5.6) thỏa mãn điều kiện Slater. Khi đó x^0 là điểm hữu hiệu của (5.6) khi và chỉ khi tồn tại (x^0, v^0) sao cho (x^0, v^0) là điểm yên ngựa của hàm*

$$F(u^0, x, v) := \left\langle u^0, f(x) \right\rangle + \langle v, g(x) \rangle \quad \textit{trên} \ \mathbb{R}^n \times \mathbb{R}^m_+$$

Chứng minh.

$\Rightarrow$) Giả sử x^0 là điểm hữu hiệu của quy hoạch lồi (5.6).
Theo định lý (5.1), $\exists u^0 \geq 0 : x^0 \in \operatorname{argmin}\{\min_x \langle u^0, f(x) \rangle \mid g(x) \leq 0\}$
Áp dụng định lý điểm yên ngựa cho hàm Lagrange của bài toán này, ta suy ra tồn tại $v^0 \geq 0$ sao cho (x^0, v^0) là điểm yên ngựa của hàm

$$F(u^0, x, v) = \left\langle u^0, f(x) \right\rangle + \langle v, g(x) \rangle \quad \text{trên} \quad \mathbb{R}^n \times \mathbb{R}^m_+$$

Tức là:

$$F(u^0, x^0, v) \leq F(u^0, x^0, v^0) \leq F(u^0, x, v^0) \quad \forall (x, v) \in \mathbb{R}^n \times \mathbb{R}^m_+ \tag{5.7}$$

$\Leftarrow$) (x^0, u^0) là điểm yên ngựa của $F(u^0, x, v)$ trên $\mathbb{R}^n \times \mathbb{R}^m_+$.
Từ điều kiện Slater, lặp lại chứng minh của định lý điểm yên ngựa (ở chương 2), ta có $u^0 > 0$. Áp dụng (5.7) với $v = 0$ ta có

$$\begin{aligned} \left\langle u^0, f(x^0) \right\rangle \leq \left\langle u^0, f(x) \right\rangle + \langle v, g(x) \rangle = \left\langle x^0, f(x) \right\rangle \quad \forall x \in D \\ \Rightarrow x^0 \in \operatorname{argmin}\{\left\langle u^0, f(x) \right\rangle \mid x \in D\} \end{aligned}$$

Theo mệnh đề (5.1), x^0 là điểm hữu hiệu của f trên D.

□

5.3 Bài tập chương 5

Bài tập 5.1. *Điểm $x^* \in D$ gọi là hữu hiệu yếu của hàm vector $f(x)$ trên D nếu $\nexists x \in D : f(x) < f(x^*)$.*
Chứng minh rằng nếu $\lambda > 0$ thì mọi nghiệm của bài toán

$$\min \lambda^T f(x) | x \in D$$

đều là điểm hữu hiệu yếu của f trên D.

Bài tập 5.2. *Cho bài toán tối ưu đa mục tiêu*

$$\mathrm{Min}_{x\in D} f(x), f : \mathbb{R}^n \to \mathbb{R}^p.$$

Chứng minh x^0 *là điểm hữu hiệu (điểm Pareto) khi và chỉ khi* x^0 *là nghiệm tối ưu của bài toán một mục tiêu*

$$\min h(x) := e^T f(x) | x \in D, f(x) \leq f(x^0),$$

trong đó : $e = (1, 1, \ldots, 1) \in \mathbb{R}^p$.

Bài tập 5.3. *Tìm một điểm hữu hiệu của bài toán đa mục tiêu*

$$\mathrm{Min}\{f(x) = (f_1(x), f_2(x)) \,| x \in D\},$$

trong đó:

$$f_1(x) = x_1 + x_2; f_2(x) = x_1 - 2x_2; D := \{x \,| x_1 \geq x_2, 0 \leq x_1 \leq 2, x_2 \geq 0\}$$

Mục lục