Đại học thái nguyên
Trường đại học sư phạm

- - - - - - - - - - - - - - - - -

Nông Thị Mai

Dưới vi phân của hàm lồi và một số ứng dụng trong tối ưu

Chuyên ngành: Giải tích
Mã số:60.46.01

Luận văn thạc sĩ toán học

Người hướng dẫn khoa học:
GS -TSKH Lê Dũng Mưu

Thái nguyên - Năm 2008

Mục lục

Danh mục các ký hiệu, các chữ viết tắt

Với n là số nguyên dương, ký hiệu:
R^n: không gian Euclide n-chiều trên trường số thực;
R^n_+: góc không âm của R^n (tập các véc-tơ có mọi toạ độ đều không âm);
R: trục số thực ($R = R^1$);
$\overline{R}$: trục số thực mở rộng ($\overline{R} = R \cup \{-\infty, +\infty\}$);
N: tập hợp số nguyên dương;
2^{R^n}: tập hợp tất cả các tập con của R^n;
Với mọi véc-tơ $x, y \in R^n$, ký hiệu:
x_i: toạ độ thứ i của x;
x^T: véc-tơ hàng (chuyển vị của x);
$\langle x, y\rangle = x^T y = xy := \sum_{j=1}^n x_j y_j$: tích vô hướng của hai véc-tơ x và y;
$||x|| = \sqrt{\sum_{j=1}^n x_j^2}$: chuẩn Euclide của x;
$[x, y]$: đoạn thẳng đóng nối x và y;
(x, y): đoạn thẳng mở nối x và y;
Với tập A, ký hiệu:
$\overline{A}$: bao đóng của A;
coA: bao lồi của A;
aff A: bao a-phin của A;
$intA$: tập hợp các điểm trong của A;
ri A: tập hợp các điểm trong tương đối của A;
Với hàm f của n biến, ký hiệu:
$\overline{f}$: hàm bao đóng của f;
$\mathrm{dom}\, f$: tập hữu dụng của f;
f^*: hàm liên hợp của f;
epi f: trên đồ thị của f;
$\partial f(x)$: dưới vi phân của f tại x;
$\partial_\epsilon f(x)$: ϵ- dưới vi phân của f tại x;
$\nabla f(x)$ hoặc $f'(x)$: đạo hàm của f tại x;
$f'(x, d)$: đạo hàm theo phương d của f tại x;

Lời nói đầu

Giải tích lồi là một bộ môn quan trọng trong giải tích phi tuyến hiện đại. Giải tích lồi nghiên cứu những khía cạnh giải tích của tập lồi và hàm lồi. Dưới vi phân là một khái niệm cơ bản của giải tích lồi. Đây là mở rộng cho đạo hàm khi hàm không khả vi. Điều này cho thấy vai trò của dưới vi phân trong giải tích hiện đại cũng có tầm quan trọng như vai trò của đạo hàm trong giải tích cổ điển. Dưới vi phân của hàm lồi có rất nhiều ứng dụng trong giải tích phi tuyến và đặc biệt trong các bộ môn toán ứng dụng, như tối ưu hoá, bất đẳng thức biến phân, cân bằng v...v.

Mục đích của luận văn là trình bày một cách có hệ thống, các kiến thức cơ bản và quan trọng nhất về dưới vi phân của hàm lồi và xét một số ứng dụng điển hình của dưới vi phân trong tối ưu hoá.

Luận văn gồm 3 chương. Trong chương 1 sẽ trình bày những kiến thức cơ bản về tập lồi và hàm lồi. Đây là các kiến thức bổ trợ cho chương 2 và do đó sẽ không được chứng minh trong luận văn này. Trong chương 2 sẽ đề cập về đạo hàm theo phương, dưới vi phân, dưới vi phân xấp xỉ và một số tính chất cơ bản của chúng. Dựa trên các kết quả đã nghiên cứu trong các chương trước, trong chương 3 sẽ trình bày các điều kiện cực trị cho các bài toán quy hoạch lồi với các rằng buộc khác nhau (không rằng buộc, rằng buộc đẳng thức, rằng buộc bất đẳng thức).

Bản luận văn này được hoàn thành dưới sự hướng dẫn khoa học của GS -TSKH Lê Dũng Mưu. Nhân đây em xin chân thành cảm ơn thầy đã hướng dẫn, động viên, khuyến khích em học tập, nghiên cứu để hoàn thành luận văn này.

Chương 1

Các kiến thức cơ bản về tập lồi và hàm lồi

Trong luận văn này, chúng ta sẽ làm việc với không gian euclid-n chiều trên trường số thực R. Không gian này được kí hiệu là R^n. Chương này nhằm giới thiệu những khái niệm cơ bản nhất của tập lồi và hàm lồi cùng với những tính chất đặc trưng của nó. Các kiến thức ở trong chương này được lấy ở tài liệu :

+ Giáo trình "Nhập môn giải tích lồi ứng dụng" của tác giả Lê Dũng Mưu và Nguyễn Văn Hiền.

+ Cuốn "Convex Analysis" của tác giả T.Rockafellar.

Do chương này chỉ mang tính chất bổ trợ, nên ta không chứng minh các kết quả nêu ở đây.

1.1 Tập lồi

Định nghĩa 1.1. Đoạn thẳng nối hai điểm a và b trong R^n là tập hợp các véc-tơ x có dạng

$$\{x \in R^n \mid x = \alpha a + \beta b\ , \alpha \geqslant 0\ ,\ \beta \geqslant 0\ , \alpha + \beta = 1\}.$$

Định nghĩa 1.2. Một tập $C \subseteq R^n$ được gọi là một tập lồi nếu C chứa mọi đoạn thẳng đi qua hai điểm bất kỳ của nó. Tức là

$$C \text{ lồi khi và chỉ khi } \forall x, y \in C, \lambda \in [0, 1] \Longrightarrow \lambda x + (1 - \lambda) y \in C.$$

Ví dụ 1.1. (Về tập lồi).

a) Tập $C = R^2_+$ là tập lồi.

b) Tập $C = [-2; 3)$ là tập lồi.

c) Tập $C \equiv oxy$ trong R^3 là tập lồi.

d) Các tam giác, hình tròn trong mặt phẳng là các tập lồi.

Ví dụ 1.2. (Về tập không lồi).

a) Tập $C = (-2; 0) \cup (0; 3)$ không là tập lồi.

b) Tập $C = \{(x, y) \in R^2 \mid xy = 0\}$ không là tập lồi.

Định nghĩa 1.3. Ta nói x là tổ hợp lồi của các điểm (véc-tơ) $x^1, ..., x^k$ nếu

$$x = \sum_{j=1}^{k} \lambda_j x^j \ , \lambda_j \geqslant 0 \ , \ \forall j = 1, ..., k \ , \ \sum_{j=1}^{k} \lambda_j = 1.$$

Định nghĩa 1.4. Siêu phẳng trong không gian R^n là một tập hợp các điểm có dạng

$$\{x \in R^n \mid a^T x = \alpha\},$$

trong đó $a \in R^n$ là một véc-tơ khác 0 và $\alpha \in R$.

Véc-tơ a thường được gọi là véc-tơ pháp tuyến của siêu phẳng. Một siêu phẳng sẽ chia không gian ra hai nửa không gian. Nửa không gian được định nghĩa như sau:

Định nghĩa 1.5. Nửa không gian là một tập hợp có dạng

$$\{x \mid a^T x \geqslant \alpha\},$$

trong đó $a \neq 0$ và $\alpha \in R$. Đây là nửa không gian đóng.

Định nghĩa 1.6. Cho $C \subseteq R^n$ là một tập lồi và $x \in C$. Tập

$$N_C(x) := \{\omega \mid \langle \omega, y - x \rangle \leqslant 0 \ , \ \forall y \in C\},$$

được gọi là nón pháp tuyến ngoài của C tại x.

Nhận xét. $N_C(x)$ là một nón lồi đóng.

Ví dụ 1.3. Trong R^2, xét tập $C = R^2_+$.

$$\begin{aligned} N_C(0) &= \{\omega \mid \langle \omega, y - 0\rangle \leqslant 0\ ,\ \forall y \in C\} \\ &= \{\omega \mid \sum_{i=1}^{2} \omega_i y_i \leqslant 0\} \\ &= \{\omega \mid \omega_i \leqslant 0\}. \end{aligned}$$

Định nghĩa 1.7. Một điểm $a \in C$ được gọi là điểm trong tương đối của C nếu nó là điểm trong của C theo tô-pô cảm sinh bởi $\operatorname{aff} C$.

Ta sẽ ký hiệu tập hợp các điểm trong tương đối của C là $\operatorname{ri} C$. Theo định nghĩa trên ta có:

$$\operatorname{ri} C := \{a \in C \mid \exists B : (a + B) \cap \operatorname{aff} C \subset C\},$$

trong đó B là một lân cận mở của gốc. Hiển nhiên

$$\operatorname{ri} C := \{a \in \operatorname{aff} C \mid \exists B : (a + B) \cap \operatorname{aff} C \subset C\}.$$

Như thường lệ, ta ký hiệu $\overline{C}$, là bao đóng của C. Tập hợp $\overline{C} \setminus \operatorname{ri} C$ được gọi là biên tương đối của C.

Mệnh đề 1.1. *Cho $C \subseteq R^n$ là một tập lồi. Giả sử $x \in \operatorname{ri} C$. Khi đó với mọi $y \in \overline{C}$ tất cả các điểm trên đoạn thẳng nối x và y, có thể trừ y, đều thuộc $\operatorname{ri} C$. Nói cách khác, với mọi $0 \leqslant \lambda < 1$, thì $(1 - \lambda) \operatorname{ri} C + \lambda \overline{C} \subset \operatorname{ri} C$.*

Định nghĩa 1.8. Một đường thẳng nối hai điểm (hai véc-tơ) a,b trong R^n là tập hợp tất cả các véc-tơ $x \in R^n$ có dạng

$$\{x \in R^n \mid x = \alpha a + \beta b\ ,\ \alpha\ ,\ \beta \in R\ ,\ \alpha + \beta = 1\}.$$

Định nghĩa 1.9. Một tập C được gọi là tập a-phin nếu nó chứa mọi đường thẳng đi qua hai điểm bất kỳ của nó, tức là

$$\forall x, y \in C\ ,\ \forall \lambda \in R \Longrightarrow \lambda x + (1 - \lambda) y \in C.$$

Ví dụ 1.4. (Về tập a-phin).

Tập $C = R^2$ là tập a-phin, không gian con là một tập affine

Nhận xét. Tập a-phin là một trường hợp riêng của tập lồi.

Định nghĩa 1.10. Bao lồi của một tập E là giao của tất cả các tập lồi chứa E. Bao lồi của một tập E sẽ được ký hiệu là coE.

Bao lồi đóng của một tập E là tập lồi đóng nhỏ nhất chứa E. Ta sẽ ký hiệu bao lồi đóng của một tập E là $\overline{co}E$.

Bao a-phin của E là giao của tất cả các tập a-phin chứa E. Bao a-phin của một tập E sẽ được ký hiệu là $\mathrm{aff}\, E$.

Định nghĩa 1.11. Cho $E \subseteq R^n$.

Điểm a được gọi là điểm trong của E nếu tồn tại một lân cận mở $U(a)$ của a sao cho $U(a) \subset E$.

Ký hiệu tập hợp các điểm trong của tập E là $intE$ và B là quả cầu đơn vị tâm ở gốc. Khi đó theo định nghĩa ta có

$$intE = \{x \mid \exists r > 0 : x + rB \subset E\}.$$

Điểm a được gọi là điểm biên của E nếu mọi lân cận của a đều có điểm thuộc E và điểm không thuộc E.

Tập E được gọi là tập mở nếu mọi điểm của E đều là điểm trong của E.

Tập E được gọi là tập đóng nếu E chứa mọi điểm biên của nó.

Tập E được gọi là bị chặn, nếu tồn tại một hình cầu chứa E.

Trong R^n tập E được gọi là tập compắc nếu E là một tập đóng và bị chặn.

Định nghĩa 1.12. Cho C là một tập lồi.

Một tập $F \subset C$ được gọi là một diện của một tập lồi C nếu

$$F \text{ là tập lồi và } \forall x, y \in C\ ,\ tx + (1-t)y \in F\ ,\ 0 < t < 1 \Longrightarrow [x, y] \subset F.$$

Ví dụ 1.5. Cho $C := \{(x, y, z) \in R^3 \mid x, y, z \in [0, 1]\}$.

Tập $F_1 := \{(x, y, z) \in R^3 \mid x, y \in [0, 1], z = 0\}$ là một diện của tập C.

Tập $F_2 := \{(x, y, z) \in R^3 \mid y \in [0, 1], x = 1, z = 0\}$ là một diện của tập C.

Điểm cực biên là diện có thứ nguyên (chiều) bằng 0.

Định nghĩa 1.13. Cho $x^0 \in C$. Ta nói $a^T x = \alpha$ là siêu phẳng tựa của C tại x^0, nếu

$$a^T x^0 = \alpha \ , \ a^T x \geqslant \alpha \ \forall x \in C.$$

Như vậy siêu phẳng tựa của C tại $x^0 \in C$ là siêu phẳng đi qua x^0 và để tập C về một phía. Nửa không gian $a^T x \geqslant \alpha$ trong định nghĩa trên, được gọi là nửa không gian tựa của C tại x^0.

Định lý 1.1. *(Krein-Milman).*

Mọi tập lồi đóng khác rỗng, không chứa đường thẳng đều có điểm cực biên.

Định lý 1.2. *(Xấp xỉ tuyến tính tập lồi).*

Mọi tập lồi đóng khác rỗng và không trùng với toàn bộ không gian đều là giao của tất cả các nửa không gian tựa của nó.

Định nghĩa 1.14. Cho hai tập C và D khác rỗng.

Ta nói siêu phẳng $a^T x = \alpha$ tách C và D nếu

$$a^T x \leqslant \alpha \leqslant a^T y \ , \ \forall x \in C \ , \ \forall y \in D.$$

Ta nói siêu phẳng $a^T x = \alpha$ tách chặt C và D nếu

$$a^T x < \alpha < a^T y \ , \ \forall x \in C \ , \ \forall y \in D.$$

Ta nói siêu phẳng $a^T x = \alpha$ tách mạnh C và D nếu

$$\mathrm{Sup}_{x \in C}\, a^T x < \alpha < \inf_{y \in D} a^T y.$$

Ví dụ 1.6. (Tách nhưng không tách chặt).

Cho tập

$$C = \{(x, y) \in R^2 \mid x^2 + y^2 \leqslant 1\},$$

và

$$D = \{(x, y) \in R^2 \mid -1 \leqslant x \leqslant 1, 1 \leqslant y \leqslant 3\}.$$

Ta có:

+ C và D khác rỗng.

+ C, D tách được vì tồn tại siêu phẳng $(0,1)(x,y) = 1$ thoả mãn

$$(0,1)(x,y) \leqslant 1 \leqslant (0,1)(x',y') \quad \forall (x,y) \in C, \forall (x',y') \in D.$$

Hay

$$y \leqslant 1 \leqslant y' \quad \forall (x,y) \in C, \forall (x',y') \in D.$$

+ C, D không tách chặt được vì không tồn tại siêu phẳng $(a_1,a_2)(x,y) = \alpha$ nào thoả mãn

$$(a_1,a_2)(x,y) < \alpha < (a_1,a_2)(x',y') \quad \forall (x,y) \in C, \forall (x',y') \in D.$$

Ví dụ 1.7. (Tách nhưng không tách mạnh).

Cho tập

$$C = \{(x,y) \in R^2 \mid x \geqslant 0, y = 0\},$$

và

$$D = \{(x,y) \in R^2 \mid y \geqslant \frac{1}{x}, y > 0, x > 0\}.$$

Ta có:

+ C và D khác rỗng.

+ C, D tách được vì tồn tại siêu phẳng $(0,1)(x,y) = 0$ thoả mãn

$$(0,1)(x,y) = 0 \leqslant (0,1)(x',y') \qquad \forall (x,y) \in C, \forall (x',y') \in D.$$

Hay

$$y = 0 \leqslant y' \qquad \forall (x,y) \in C, \forall (x',y') \in D.$$

+ C, D không tách mạnh được vì

$$\mathrm{Sup}_{(x,y)\in C}(0,1)(x,y) = 0,$$

$$\inf_{(x',y')\in D}(0,1)(x',y') = 0.$$

Định lý 1.3. *(Định lý tách 1).*

Cho C và D là hai tập lồi khác rỗng trong R^n sao cho $C \cap D = \emptyset$. Khi đó có một siêu phẳng tách C và D.

Hệ quả 1.1. *(Bổ đề liên thuộc).*

Cho $C \subset R^n$ là một tập lồi khác rỗng. Giả sử $x^0 \notin C$. Khi đó tồn tại $t \in R^n$, $t \neq 0$ thoả mãn

$$\langle t, x\rangle \geqslant \langle t, x^0\rangle \ \ \forall x \in C.$$

Định lý 1.4. *(Định lý tách 2).*

Cho C và D là hai tập lồi đóng khác rỗng sao cho $C \cap D = \emptyset$. Giả sử có ít nhất một tập là compắc. Khi đó hai tập này có thể tách mạnh được bởi một siêu phẳng.

Hệ quả 1.2. *Cho $C \subset R^n$ là một tập lồi đóng khác rỗng sao cho $0 \notin C$. Khi đó tồn tại một véc-tơ $t \in R^n$, $t \neq 0$ và $\alpha > 0$ sao cho*

$$\langle t, x\rangle \geqslant \alpha > 0 \ , \ \forall x \in C.$$

1.2 Hàm lồi

1.2.1 Hàm lồi

Cho $C \subseteq R^n$ và $f : C \longrightarrow R \cup \{-\infty, +\infty\}$. Ta sẽ kí hiệu:

$\operatorname{dom} f := \{x \in C \mid f(x) < +\infty\}$. Tập $\operatorname{dom} f$ được gọi là miền hữu dụng của f

$\operatorname{epi} f := \{(x, \mu) \in C \times R \mid f(x) \leqslant \mu\}$. Tập $\operatorname{epi} f$ được gọi là trên đồ thị của hàm f.

Bằng cách cho $f(x) = +\infty$ nếu $x \notin C$, ta có thể coi f được xác định trên toàn không gian và hiển nhiên là

$$\operatorname{dom} f := \{x \in R^n \mid f(x) < +\infty\}.$$

$$\operatorname{epi} f := \{(x, \mu) \in R^n \times R \mid f(x) \leqslant \mu\}.$$

Định nghĩa 1.15. Cho $\emptyset \neq C \subseteq R^n$ lồi và $f : C \longrightarrow R \cup \{-\infty, +\infty\}$. Ta nói f là hàm lồi trên C nếu $\operatorname{epi} f$ là một tập lồi trong R^{n+1}.

Sau đây ta sẽ chủ yếu làm việc với hàm $f : R^n \longrightarrow R \cup \{+\infty\}$.Trong trường hợp này, định nghĩa trên tương đương với:

Hàm $f : R^n \longrightarrow R \cup \{+\infty\}$ là hàm lồi trên C nếu

$$f[\lambda x + (1-\lambda)y] \leqslant \lambda f(x) + (1-\lambda)f(y)\ ,\ \forall x, y \in C\ ,\ \forall \lambda \in (0,1)$$

Hàm $f : R^n \longrightarrow R \cup \{+\infty\}$ là hàm lồi chặt trên C nếu

$$f[\lambda x + (1-\lambda)y] < \lambda f(x) + (1-\lambda)f(y)\ ,\ \forall x, y \in C\ ,\ \forall \lambda \in (0,1)$$

Hàm $f : R^n \longrightarrow R \cup \{+\infty\}$ là hàm lồi mạnh trên C với hệ số lồi $\eta > 0$ nếu

$$f[\lambda x + (1-\lambda)y] \leqslant \lambda f(x) + (1-\lambda)f(y) - \frac{1}{2}\eta\lambda(1-\lambda)||x-y||^2\ ,$$
$$\forall x, y \in C\ ,\ \forall \lambda \in (0,1).$$

Hàm f được gọi là một hàm lõm trên C, nếu $-f$ là hàm lồi trên C.

Ví dụ 1.8. Hàm a-phin. $f(x) = a^T x + \alpha, a \in R^n, \alpha \in R$

$\forall x, y \in R^n, \forall \lambda \in (0,1)$, ta có

$$\begin{aligned} f[\lambda x + (1-\lambda)y] &= a^T[\lambda x + (1-\lambda)y] + \alpha \\ &= \lambda a^T x + (1-\lambda)a^T y + \alpha \\ &= \lambda a^T x + \lambda\alpha + (1-\lambda)a^T y + (1-\lambda)\alpha \\ &= \lambda(a^T x + \alpha) + (1-\lambda)(a^T y + \alpha) \\ &= \lambda f(x) + (1-\lambda)f(y). \end{aligned}$$

Vậy f là một hàm lồi trên R^n.

$\forall x, y \in R^n, \forall \lambda \in (0,1)$, lại có

$$\begin{aligned} -f[\lambda x + (1-\lambda)y] &= -a^T[\lambda x + (1-\lambda)y] - \alpha \\ &= -\lambda a^T x - (1-\lambda)a^T y - \alpha \\ &= -\lambda a^T x - \lambda\alpha - (1-\lambda)a^T y - (1-\lambda)\alpha \\ &= -\lambda(a^T x + \alpha) - (1-\lambda)(a^T y + \alpha) \\ &= -\lambda f(x) - (1-\lambda)f(y). \end{aligned}$$

Vậy $-f$ là một hàm lồi trên R^n. Suy ra f là một hàm lõm trên R^n.

Ví dụ 1.9. Hàm chỉ. Cho $C \neq \emptyset$ là một tập lồi .

Đặt $\delta_C(x) := \begin{cases} 0 & \text{nếu} \quad x \in C, \\ +\infty & \text{nếu} \quad x \notin C. \end{cases}$

Ta nói δ_C là hàm chỉ của C.

+ $\forall x, y \in C, \forall \lambda \in (0,1)$, ta có: $\delta_C(x) = 0$, $\delta_C(y) = 0$.

Do C lồi nên $\lambda x + (1-\lambda)y \in C$.

Suy ra $\delta_C[\lambda x + (1-\lambda)y] = 0 = \lambda\delta_C(x) + (1-\lambda)\delta_C(y)$.

+ $\forall x \in C, \forall y \notin C, \forall \lambda \in (0,1)$, ta có :

$\delta_C(x) = 0$, $\delta_C(y) = +\infty$, $\delta_C[\lambda x + (1-\lambda)y] \leqslant +\infty$.

Suy ra $\delta_C[\lambda x + (1-\lambda)y] \leqslant \lambda\delta_C(x) + (1-\lambda)\delta_C(y)$.

+ $\forall x, y \notin C, \forall \lambda \in (0,1)$, ta có :

$\delta_C(x) = +\infty$, $\delta_C(y) = +\infty$, $\delta_C[\lambda x + (1-\lambda)y] \leqslant +\infty$.

Suy ra $\delta_C[\lambda x + (1-\lambda)y] \leqslant \lambda\delta_C(x) + (1-\lambda)\delta_C(y)$.

Vậy δ_C là hàm lồi trên R^n.

Ví dụ 1.10. Hàm tựa.

Đặt $S_C(y) := \mathrm{Sup}_{x\in C}\langle y, x\rangle$.Ta nói S_C là hàm tựa của C.

$\forall x, y \in C, \forall \lambda \in (0,1)$, ta có

$$\begin{aligned} S_C[\lambda x + (1-\lambda)y] &= \mathrm{Sup}_{z\in C}\langle \lambda x + (1-\lambda)y, z\rangle \\ &= \mathrm{Sup}_{z\in C}\{\langle \lambda x, z\rangle + \langle (1-\lambda)y, z\rangle\} \\ &\leqslant \mathrm{Sup}_{z\in C}\langle \lambda x, z\rangle + \mathrm{Sup}_{z\in C}\langle (1-\lambda)y, z\rangle \\ &= \lambda\,\mathrm{Sup}_{z\in C}\langle x, z\rangle + (1-\lambda)\,\mathrm{Sup}_{z\in C}\langle y, z\rangle \\ &= \lambda S_C(x) + (1-\lambda)S_C(y). \end{aligned}$$

Vậy S_C là hàm lồi trên C.

Định nghĩa 1.16. Cho $f : R^n \longrightarrow R \cup \{+\infty\}$ (không nhất thiết lồi), $C \subseteq R^n$ là một tập lồi khác rỗng và η là một số thực .

Ta nói η là hệ số lồi của f trên C, nếu với mọi $\lambda \in (0,1)$, với mọi $x, y \in C$, ta có:

$$f[(1-\lambda)x + \lambda y] \leqslant (1-\lambda)f(x) + \lambda f(y) - \frac{1}{2}\eta\lambda(1-\lambda)||x-y||^2.$$

Nếu $\eta = 0$ thì f lồi trên C.

Nếu f có hệ số lồi trên C là $\eta > 0$, thì f lồi mạnh trên C với hệ số η.

Định nghĩa 1.17. Một hàm $f : R^n \longrightarrow R \cup \{+\infty\}$ được gọi là chính thường nếu $\operatorname{dom} f \neq \emptyset$ và $f(x) > -\infty$ với mọi x.

Định nghĩa 1.18. Hàm $f : R^n \longrightarrow R \cup \{+\infty\}$ được gọi là đóng, nếu $\operatorname{epi} f$ là một tập đóng trong R^{n+1}

Chú ý 1.1. 1. Nếu f là một hàm lồi trên một tập lồi C, thì có thể thác triển f lên toàn không gian bằng cách đặt

$$f_e(x) = \begin{cases} f(x) & \text{nếu} \quad x \in C, \\ +\infty & \text{nếu} \quad x \notin C. \end{cases}$$

Hiển nhiên $f_e(x) = f(x)$ với mọi $x \in C$ và f_e lồi trên R^n. Hơn nữa f_e là chính thường khi và chỉ khi f chính thường. Tương tự f_e đóng khi và chỉ khi f đóng.

2. Nếu f là một hàm lồi trên R^n thì $\operatorname{dom} f$ là một tập lồi vì $\operatorname{dom} f$ chính là hình chiếu trên R^n của $\operatorname{epi} f$, tức là:

$$\operatorname{dom} f = \{x | \exists \mu \in R : (x, \mu) \in \operatorname{epi} f\}.$$

Định nghĩa 1.19. Cho $f : R^n \longrightarrow R \cup \{+\infty\}$.

Hàm f được gọi là thuần nhất dương (bậc 1) trên R^n nếu

$$f(\lambda x) = \lambda f(x) \qquad \forall x \in R^n, \forall \lambda > 0.$$

Hàm f được gọi là dưới cộng tính nếu $f(x+y) \leqslant f(x) + f(y) \qquad \forall x, y$.

Hàm f được gọi là dưới tuyến tính nếu f là thuần nhất dương và dưới cộng tính.

Ví dụ 1.11. Hàm chuẩn $f(x) = \|x\|$ là hàm dưới tuyến tính. Thật vậy,

$\forall x \in R^n, \forall \lambda > 0$, ta có: $f(\lambda x) = \|\lambda x\| = |\lambda|.\|x\| = \lambda\|x\| = \lambda f(x)$.

$\forall x, y \in R^n$, ta có: $f(x+y) = \|x+y\| \leqslant \|x\| + \|y\| = f(x) + f(y)$.

Mệnh đề 1.2. *Cho $f : R^n \longrightarrow R \cup \{+\infty\}$ là một hàm thuần nhất dương trên R^n.*

Khi đó: f lồi khi và chỉ khi f là dưới cộng tính.

1.2.2 Tính liên tục của hàm lồi

Định nghĩa 1.20. Cho hàm $f : E \longrightarrow R \cup \{-\infty, +\infty\}$.

Hàm f được gọi là nửa liên tục dưới tại một điểm $x \in E$ nếu với mọi dãy $\{x^k\} \subset E, x^k \to x$ ta có

$$\liminf f(x^k) \geqslant f(x).$$

Hàm f được gọi là nửa liên tục trên tại $x \in E$ nếu $-f$ nửa liên tục dưới tại $x \in E$. Như vậy f nửa liên tục trên tại $x \in E$ nếu với mọi dãy $\{x^k\} \subset E, x^k \to x$ ta có

$$\limsup f(x^k) \leqslant f(x).$$

Hàm f được gọi là liên tục tại $x \in E$ nếu như nó vừa nửa liên tục trên và nửa liên tục dưới tại $x \in E$.

Hàm f được gọi là nửa liên tục dưới trên E nếu nó nửa liên tục dưới tại mọi điểm thuộc E.

Hàm f được gọi là nửa liên tục trên trên E nếu nó nửa liên tục trên tại mọi điểm thuộc E.

Hàm f được gọi là liên tục trên E nếu nó nửa liên tục trên và nửa liên tục dưới trên E.

Định nghĩa 1.21. Cho hai hàm f và g xác định trên R^n.

Ta nói g là bao đóng của f, nếu $\operatorname{epi} g = \overline{\operatorname{epi} f}$. Bao đóng của f sẽ được kí hiệu là $\overline{f}$. Vậy $\operatorname{epi} \overline{f} = \overline{\operatorname{epi} f}$.

Hàm f được gọi là đóng nếu $\operatorname{epi} f = \overline{\operatorname{epi} f}$.

1.2.3 Các phép toán bảo toàn tính lồi

Định nghĩa 1.22. Giả sử $\{f_\alpha\}_{\alpha \in I}$ là một họ tuỳ ý các hàm số trên R^n và $E \subseteq R^n$. Hàm cận trên của họ hàm này trên coE, ký hiệu là $V_{\alpha \in I} f_\alpha$ là hàm số được định nghĩa như sau:

$$(V_{\alpha \in I} f_\alpha)(x) := \operatorname{Sup}_{\alpha \in I} f_\alpha(x)$$

với mỗi $x \in coE$.

Mệnh đề 1.3. *Giả sử $\{f_\alpha\}_{\alpha \in I}$ là một họ hàm lồi trên R^n và $E \subseteq R^n$. Khi đó hàm cận trên của họ hàm này là một hàm lồi trên coE.*

1.2.4 Bất đẳng thức lồi

Định nghĩa 1.23. Cho $D \subseteq R^n$ là một tập lồi và $f_1, ..., f_m$ là các hàm lồi trên R^n. Hệ bất đẳng thức

$$x \in D, f_i(x) <= 0, i \in I$$

được gọi là hệ bất đẳng thức lồi, trong đó I là tập chỉ số và ký hiệu $<=$ có thể hiểu là $<$ hoặc $\leqslant$.

Mệnh đề 1.4. *Cho $f_1, ..., f_m$ là các hàm lồi hữu hạn trên một tập lồi $D \neq \emptyset$ và A là một ma trận thực cấp $k \times n$. Giả sử $b \in \mathrm{ri}\, A(D)$. Khi đó hệ*

$$x \in D,\ Ax = b,\ f_i(x) < 0\ i = 1, .., m$$

không có nghiệm, khi và chỉ khi tồn tại $t \in R^k$ và $\lambda_i \geqslant 0,\ i = 1, .., m$ sao cho $\sum_{i=1}^m \lambda_i = 1$ và

$$\langle t, Ax - b\rangle + \sum_{i=1}^m \lambda_i f_i(x) \geqslant 0 \qquad \forall x \in D.$$

1.2.5 Hàm liên hợp

Định nghĩa 1.24. Cho $f : R^n \longrightarrow [-\infty, +\infty]$ là một hàm bất kỳ. Hàm

$$f^*(x^*) := \mathrm{Sup}\{\langle x^*, x\rangle - f(x) \mid x \in R^n\}$$

được gọi là hàm liên hợp của f.

Chú ý 1.2. Như thường lệ, trong định nghĩa trên ta qui ước cận trên đúng trên một tập rỗng là $-\infty$. Như vậy nếu $f \equiv +\infty$, thì $f^* \equiv -\infty$, ngoài ra nếu f có nhận giá trị $-\infty$ thì $f^* \equiv +\infty$.

Để khỏi phải làm việc với hàm liên hợp đồng nhất bằng $+\infty$ hoặc đồng nhất bằng $-\infty$, ta sẽ hạn chế việc xét hàm liên hợp trong lớp hàm có tính chất sau:

$$f \not\equiv +\infty \text{ và tồn tại một hàm non a-phin của f.}$$

Ví dụ 1.12. Xét hàm chỉ

$$\delta_C(x) = \begin{cases} 0 & \text{nếu} \quad x \in C, \\ +\infty & \text{nếu} \quad x \notin C. \end{cases}$$

Ta có:

$$\begin{aligned} \delta_C^*(x^*) &:= \mathrm{Sup}_{x\in R^n}\{\langle x^*, x\rangle - \delta_C(x)\} \\ &= \mathrm{Sup}_{x\in C}\{\langle x^*, x\rangle - \delta_C(x)\} \\ &= \mathrm{Sup}_{x\in C}\{\langle x^*, x\rangle - 0\} \\ &= \mathrm{Sup}_{x\in C}\langle x^*, x\rangle \\ &= S_C(x^*). \end{aligned}$$

Mệnh đề 1.5. *Với mọi hàm số f, hàm liên hợp f^* là một hàm lồi đóng thoả mãn bất đẳng thức Fenchel sau:*

$$f^*(x^*) \geqslant \langle x^*, x\rangle - f(x) \qquad \forall x, \forall x^*.$$

Chú ý 1.3. Trong nhiều trường hợp, ta quan tâm đến hàm liên hợp thứ hai. Theo định nghĩa hàm liên hợp thì

$$f^{**}(x) := (f^*)^*(x) = \mathrm{Sup}\{\langle x, s\rangle - f^*(s) \mid s \in R^n\}.$$

Hàm liên hợp thứ hai tất nhiên luôn là một hàm lồi đóng.

Mệnh đề 1.6. *Giả sử $f \not\equiv +\infty$ và tồn tại một hàm non a-phin của f. Khi đó*

$$\mathrm{epi}\, f^{**} = \overline{co}(\mathrm{epi}\, f).$$

Hệ quả 1.3. *$f \equiv f^{**}$ khi và chỉ khi f là hàm lồi, đóng.*

Định nghĩa 1.25. Hàm l là hàm non a-phin của một hàm f trên R^n nếu

$$l \text{ là hàm a-phin trên } R^n \text{ và } l(x) \leqslant f(x) \qquad \forall x \in R^n.$$

Chương 2

Dưới vi phân của hàm lồi

Phép tính vi phân là một trong những đề tài cơ bản nhất của giải tích cổ điển. Trong giải tích lồi, lý thuyết này lại càng trở nên phong phú nhờ những tính chất đặc biệt của tập lồi và hàm lồi. Mục đầu tiên của chương này sẽ xét đến đạo hàm theo phương của một hàm lồi. Tiếp đến ở mục 2, sẽ đưa ra định nghĩa về dưới vi phân và các tính chất của nó như: Xét tính khả vi của hàm lồi, khảo sát tính đơn điệu của dưới vi phân, khảo sát tính liên tục của ánh xạ dưới vi phân và một số phép tính với dưới vi phân. Mục cuối của chương sẽ giới thiệu về dưới vi phân xấp xỉ và một số tính chất của nó.

2.1 Đạo hàm theo phương

Cho một hàm n-biến $f : R^n \longrightarrow R \cup \{+\infty\}$. Khi cố định một phương và xét hàm nhiều biến trên phương đó , thì ta có một hàm một biến. Giả sử $y \neq 0$ là một phương cho trước xuất phát từ điểm x^0. Khi đó mọi điểm x thuộc đường thẳng đi qua x^0 và có phương y đều có dạng $x = x^0 + \lambda y$ với $\lambda \in R$. Nếu đặt $\xi(\lambda) = f(x^0 + \lambda y)$ thì ξ lồi trên R khi và chỉ khi f lồi trên R^n.

Định nghĩa 2.1. Cho $f : R^n \longrightarrow R \cup \{+\infty\}$ và $x^0 \in R^n$ sao cho $f(x^0) < +\infty$.

Nếu với một véc-tơ $y \in R^n$ mà giới hạn $\lim\limits_{\lambda \to 0} \frac{f(x^0+\lambda y)-f(x^0)}{\lambda}$ tồn tại (hữu hạn hay vô hạn) thì ta nói f có đạo hàm theo phương y tại điểm x^0. Ta sẽ ký hiệu giới hạn này là $f'(x^0, y)$.

Ví dụ 2.1. Giả sử f được cho như sau:

$$f(x) = \begin{cases} 0 & \text{nếu} \quad x < 0, \\ 1 & \text{nếu} \quad x = 0, \\ +\infty & \text{nếu} \quad x > 0. \end{cases}$$

Ta có

$\operatorname{dom} f = (-\infty; 0] \Rightarrow \operatorname{dom} f \neq \emptyset$,

$f(x) > -\infty, \forall x$. Vậy f là hàm chính thường .

Ta có:

$f'(0,-1) = \lim\limits_{\lambda \to 0} \frac{f(0+\lambda(-1))-f(0)}{\lambda} = \lim\limits_{\lambda \to 0} \frac{0-1}{\lambda} = -\infty,$

$f'(0,0) = \lim\limits_{\lambda \to 0} \frac{f(0+\lambda 0)-f(0)}{\lambda} = \lim\limits_{\lambda \to 0} \frac{1-1}{\lambda} = 0,$

$f'(0,1) = \lim\limits_{\lambda \to 0} \frac{f(0+\lambda 1)-f(0)}{\lambda} = \lim\limits_{\lambda \to 0} \frac{\infty-1}{\lambda} = +\infty.$

Suy ra $f'(0,.)$ không là hàm chính thường.

Mệnh đề 2.1. *Cho* $f : R^n \longrightarrow R \cup \{+\infty\}$ *lồi. Khi đó với mọi* $x \in \operatorname{dom} f$ *và mọi* $y \in R^n$ *ta có:*

i) φ *là hàm đơn điệu không giảm trên* $(0; +\infty)$ *, trong đó*

$$\varphi(\lambda) := \frac{f(x+\lambda y) - f(x)}{\lambda},$$

và do đó $f'(x,y)$ *tồn tại với mọi* $y \in R^n$ *và*

$$f'(x,y) := \inf\nolimits_{\lambda > 0} \frac{f(x+\lambda y) - f(x)}{\lambda}.$$

ii) Hàm $f'(x,.)$ *thuần nhất dương bậc 1.*

Ngoài ra nếu $f'(x,.) > -\infty$ *thì hàm* $f'(x,.)$ *là dưới tuyến tính trên* R^n *(do đó nó là hàm lồi chính thường trên* R^n*).*

iii) $-f'(x,-y) \leqslant f'(x,y) \qquad \forall y \in R^n.$

iv) Hàm $f'(x,.)$ *nhận giá trị hữu hạn trên* F *khi và chỉ khi* $x \in \operatorname{ri}(\operatorname{dom} f)$, *trong đó* F *là không gian con của* $\operatorname{dom} f$.

Chứng minh. i) Ta chứng minh hàm φ đơn điệu không giảm trên miền $(0; +\infty)$.

Định nghĩa hàm $h : R \longrightarrow R \cup \{+\infty\}$ xác định bởi

$$h(\lambda) = f(x + \lambda.y) - f(x).$$

Khi đó $h(0) = 0$.

Giả sử $0 < \lambda' \leqslant \lambda$, do f là hàm lồi nên h là hàm lồi , không nhận giá trị $-\infty$.

Ta có

$$\begin{aligned} h(\lambda') &= h[\frac{\lambda'}{\lambda}\lambda + (1 - \frac{\lambda'}{\lambda})0] \\ &\leqslant \frac{\lambda'}{\lambda}h(\lambda) + (1 - \frac{\lambda'}{\lambda})h(0) \\ &= \frac{\lambda'}{\lambda}h(\lambda). \end{aligned}$$

Do $\varphi(\lambda) = \frac{f(x+\lambda y)-f(x)}{\lambda} = \frac{h(\lambda)}{\lambda}$ nên $\varphi(\lambda') \leqslant \varphi(\lambda)$.

Vậy φ là hàm không giảm trên miền $(0; +\infty)$.

Suy ra $f'(x, y) = \lim\limits_{\lambda \to 0} \varphi(\lambda)$ tồn tại và

$$\lim_{\lambda \to 0} \varphi(\lambda) = \inf_{\lambda > 0} \varphi(\lambda) = \inf_{\lambda > 0} \frac{f(x + \lambda.y) - f(x)}{\lambda}.$$

ii) Theo định nghĩa, ta có

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

Chứng minh tính thuần nhất dương .

Với $t > 0$, ta viết

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

Đặt $\lambda' = \lambda t$, ta có tiếp

$$f'(x, ty) = t \lim_{\lambda \to 0} \frac{f(x + \lambda' y) - f(x)}{\lambda'} = tf'(x, y).$$

Vậy $f'(x, .)$ thuần nhất dương.

Chứng minh tính dưới tuyến tính.

Giả sử $f'(x,.) > -\infty$, với mọi u và v ta có:

$$\begin{aligned} f'(x,u+v) &= \inf_{\lambda>0} \frac{f[x+\frac{\lambda}{2}(u+v)]-f(x)}{\frac{\lambda}{2}} \qquad \text{(theo i)} \\ &= \inf_{\lambda>0} \frac{f[(\frac{x}{2}+\frac{\lambda}{2}u)+(\frac{x}{2}+\frac{\lambda}{2}v)]-\frac{1}{2}f(x)-\frac{1}{2}f(x)}{\frac{\lambda}{2}}. \end{aligned}$$

Do f là hàm lồi không nhận giá trị $-\infty$,nên

$$\begin{aligned} &f[(\frac{x}{2}+\frac{\lambda}{2}u)+(\frac{x}{2}+\frac{\lambda}{2}v)]-\frac{1}{2}f(x)-\frac{1}{2}f(x) \\ &\leqslant \frac{1}{2}[f(x+\lambda u)-f(x)]+\frac{1}{2}[f(x+\lambda v)-f(x)]. \end{aligned}$$

Do đó

$$\begin{aligned} f'(x,u+v) &\leqslant \inf_{\lambda>0} \frac{f(x+\lambda u)}{\lambda} + \inf_{\lambda>0} \frac{f(x+\lambda v)}{\lambda} \\ &= f'(x,u)+f'(x,v). \end{aligned}$$

($f'(x,u)+f'(x,v)$ có nghĩa vì $f'(x,.) > -\infty$).

Vậy $f'(x,.)$ là hàm dưới cộng tính. Suy ra $f'(x,.)$ là hàm dưới tuyến tính trên R^n.

Vì $f'(x,.) > -\infty, f'(x,0) = 0$ và $f'(x,.)$ là dưới tuyến tính trên R^n, nên nó là hàm lồi, chính thường trên toàn không gian.

iii) Do $f'(x,0) = 0$ và theo tính chất dưới cộng tính, ta có:

$$0 = f'(x,0) = f'(x,y-y) \leqslant f'(x,y)+f'(x,-y) \qquad \forall y \in R^n.$$

Suy ra $-f'(x,-y) \leqslant f'(x,y)$ với mọi $y \in R^n$.

iv) Giả sử $x \in \operatorname{ri}(\operatorname{dom} f)$. Ta cần chứng tỏ $f'(x,.)$ hữu hạn trên F.

Từ iii) suy ra $f'(x,.) > -\infty$. Vậy cần chỉ ra $f'(x,y) < +\infty$ với mọi $y \in F$.

Do $x \in \operatorname{ri}(\operatorname{dom} f)$, nên $\forall y \in F$, $x+\lambda.y \in \operatorname{dom} f$ $\forall \lambda > 0$ đủ nhỏ.

Do đó $f'(x,y) = \inf_{\lambda>0} \frac{f(x+\lambda.y)-f(x)}{\lambda} < +\infty$.

Ngược lại, giả sử $f'(x,y)$ hữu hạn với mọi $y \in F$. Ta cần chứng tỏ $x \in \operatorname{ri}(\operatorname{dom} f)$.

Thật vậy, nếu trái lại sẽ tồn tại $y \in F$ và một dãy $\{\lambda_k\}$ các số dương hội tụ đến 0 và $x + \lambda_k.y \notin \operatorname{dom} f$ với mọi k đủ lớn. Trong trường hợp này

$$f(x + \lambda_k.y) - f(x) = +\infty \text{ với mọi k đủ lớn.}$$

Do đó $f'(x, y) = +\infty$. Mâu thuẫn với giả thiết. Vậy $x \in \operatorname{ri}(\operatorname{dom} f)$. $\square$

2.2 Dưới vi phân và các tính chất

2.2.1 Dưới vi phân

Định nghĩa 2.2. Cho $f : R^n \longrightarrow R \cup \{+\infty\}$. Ta nói $x^* \in R^n$ là dưới đạo hàm của f tại x nếu

$$\langle x^*, z - x\rangle + f(x) \leqslant f(z) \qquad \forall z.$$

Kí hiệu tập hợp tất cả các dưới đạo hàm của f tại x là $\partial f(x)$. Vậy $\partial f(x)$ là một tập (có thể bằng $\emptyset$) trong R^n. Khi $\partial f(x) \neq \emptyset$, thì ta nói hàm f khả dưới vi phân tại x.

Theo định nghĩa, một điểm $x^* \in \partial f(x)$ khi và chỉ khi nó thoả mãn một hệ vô hạn các bất đẳng thức tuyến tính . Như vậy $\partial f(x)$ là giao của các nửa không gian đóng. Vậy $\partial f(x)$ luôn là một tập lồi đóng (có thể rỗng).

Kí hiệu $\operatorname{dom}(\partial f) := \{x | \partial f(x) \neq \emptyset\}$.

Ví dụ 2.2. 1) Hàm chuẩn $f(x) = \|x\|, x \in R^n$.

Tại điểm $x = 0$, ta có

$$\partial f(0) = \{x^* | \langle x^*, x\rangle \leqslant \|x\|, \forall x\}.$$

Vậy hàm $f(x)$ khả dưới vi phân.

Lại có

$$\lim_{x \to 0} \frac{f(x) - f(0) - \langle x^*, x - 0\rangle}{\|x - 0\|} = \lim_{x \to 0} \frac{\|x\| - \langle x^*, x\rangle}{\|x\|} = 1 \neq 0.$$

Vậy hàm $f(x)$ không khả vi tại $x = 0$.

2) Hàm chỉ

$$f(x) = \delta_C(x) := \begin{cases} 0 & \text{nếu } x \in C, \\ +\infty & \text{nếu } x \notin C. \end{cases}$$

Trong đó C là một tập lồi khác $\emptyset$.

Khi đó với $x^0 \in C$, ta có

$$\partial f(x^0) = \partial \delta_C(x^0) = \{x^* | \langle x^*, x - x^0 \rangle \leqslant \delta_C(x), \forall x\}.$$

Với $x \notin C$ thì $\delta_C(x) = +\infty$, nên bất đẳng thức này luôn đúng.

Vậy $\partial f(x^0) = \partial \delta_C(x^0) = \{x^* | \langle x^*, x - x^0 \rangle \leqslant 0, \forall x \in C\} = N_C(x^0)$.

Vậy dưới vi phân của hàm chỉ của một tập lồi C khác $\emptyset$ tại một điểm $x^0 \in C$ chính là nón pháp tuyến ngoài của C tại x^0.

Mệnh đề 2.2. *i) $x^* \in \partial f(x)$ khi và chỉ khi $f'(x, y) \geqslant \langle x^*, y \rangle$, $\forall y$.*

ii) Nếu f là hàm lồi chính thường trên R^n, thì với mọi $x \in \operatorname{dom}(\partial f)$, ta có $f(x) = \overline{f}(x)$ và $\partial f(x) = \partial \overline{f}(x)$.

Chứng minh. i) Theo định nghĩa

$$x^* \in \partial f(x) \Leftrightarrow \langle x^*, z - x \rangle + f(x) \leqslant f(z) \qquad \forall z.$$

Với bất kì y, lấy $z = x + \lambda.y, \lambda > 0$, ta có

$$\langle x^*, \lambda.y \rangle + f(x) \leqslant f(x + \lambda.y).$$

Từ đây suy ra

$$\langle x^*, y \rangle \leqslant \frac{f(x + \lambda.y) - f(x)}{\lambda} \qquad \forall \lambda > 0. \tag{2.1}$$

Theo định nghĩa của $f'(x, y)$, suy ra ngay $\langle x^*, y \rangle \leqslant f'(x, y) \qquad \forall y$.

Ngược lại, giả sử (2.1) thoả mãn.

Lấy z bất kì và áp dụng (2.1) với $y = z - x$ và $\lambda = 1$, ta có

$$\langle x^*, z - x \rangle \leqslant f(z) - f(x) \qquad \forall z.$$

Vậy $x^* \in \partial f(x)$.

ii) Cho $x \in \operatorname{dom}(\partial f)$, thì $\partial f(x) \neq \emptyset$, tức là tồn tại $x^* \in \partial f(x)$.

Theo định nghĩa của $\overline{f}$, ta có $\operatorname{epi}\overline{f} = \overline{\operatorname{epi} f}$.

Mặt khác, ta lại có $\operatorname{epi} f \subset \overline{\operatorname{epi} f}$, suy ra $\operatorname{epi} f \subset \operatorname{epi}\overline{f}$. Vậy

$$f(x) \geqslant \overline{f}(x). \tag{2.2}$$

Theo giả thiết f là hàm lồi chính thường trên R^n, nên $\overline{f}$ là hàm lồi đóng trên R^n, theo hệ quả 1.1, ta có

$$\overline{f}(x) = f^{**}(x). \tag{2.3}$$

Theo tính chất của hàm liên hợp thứ 2, ta có

$$f^{**}(x) \geqslant < \langle x^*, x\rangle - f^*(x^*) = f(x). \tag{2.4}$$

Từ (2.2),(2.3) và (2.4) ta có $f(x) = \overline{f}(x)$.

Ta lấy $y^* \in \partial\overline{f}(x)$ thì $\forall z$ tacó

$$\langle y^*, z - x\rangle + \overline{f}(x) \leqslant \overline{f}(z).$$

Mặt khác

$$f(z) \geqslant \overline{f}(z) \geqslant \langle y^*, z - x\rangle + \overline{f}(x) = \langle y^*, z - x\rangle + f(x).$$

Suy ra $y^* \in \partial f(x)$. Vậy

$$\partial\overline{f}(x) \subset \partial f(x). \tag{2.5}$$

Ngược lại, lấy $z^0 \in \operatorname{ri}(\operatorname{dom} f)$. Với mọi z ta có

$$\overline{f}(z) = f(z) = \lim_{t\to 0} f[(1-t).z + t.z^0].$$

Vậy theo định nghĩa của dưới vi phân ta có :

$$x^* \in \partial f(x) \Leftrightarrow \langle x^*, (1-t).z + t.z^0 - x\rangle + f(x) \leqslant f[(1-t).z + t.z^0].$$

Cho $t \to 0$ ta được :

$$\langle x^*, z - x\rangle + f(x) \leqslant \overline{f}(z).$$

Hay

$$\langle x^*, z - x\rangle + \overline{f}(x) \leqslant \overline{f}(z)$$

Chứng tỏ $x^* \in \partial\overline{f}(x)$. Vậy

$$\partial f(x) \subset \partial\overline{f}(x). \tag{2.6}$$

Từ (2.5) và (2.6) ta có $\partial f(x) = \partial\overline{f}(x)$. □

Mệnh đề 2.3. *Cho $f : R^n \longrightarrow R \cup \{+\infty\}$ lồi, khi đó :*

i) Nếu $x \notin \operatorname{dom} f$, thì $\partial f(x) = \emptyset$.

ii) $x \in \operatorname{ri}(\operatorname{dom} f)$ khi và chỉ khi $\partial f(x) \neq \emptyset$ và compắc.

Chứng minh. i) Cho $z \in \operatorname{dom} f$, thì $f(z) < +\infty$. Vậy nếu $x \notin \operatorname{dom} f$ thì $f(x) = +\infty$ và do đó không thể tồn tại x^* tho mãn

$$\langle x^*, z - x\rangle + f(x) \leqslant f(z) < +\infty.$$

Vậy $\partial f(x) = \emptyset$.

ii) Giả sử $x \in \operatorname{ri}(\operatorname{dom} f)$. Ta có điểm $(x, f(x))$ nằm trên biên của $\operatorname{epi} f$.

Do f lồi, chính thường, nên tồn tại siêu phẳng tựa của $\operatorname{epi} f$ đi qua $(x, f(x))$.

Tức là tồn tại $p \in R^n, t \in R$ không đồng thời bằng 0 sao cho

$$\langle p, x\rangle + t.f(x) \leqslant \langle p, y\rangle + t.\mu \ , \forall (y, \mu) \in \operatorname{epi} f. \tag{2.7}$$

Ta có $t \neq 0$, vì nếu $t = 0$ thì $\langle p, x\rangle \leqslant \langle p, y\rangle \ , \forall y \in \operatorname{dom} f$.

Hay $\langle p, x - y\rangle \leqslant 0 \ , \forall y \in \operatorname{dom} f$.

Nhưng do $x \in \operatorname{ri}(\operatorname{dom} f)$, nên điều này kéo theo $p = 0$. Mâu thuẫn với p, t không đồng thời bằng 0. Vậy $t \neq 0$.

Hơn nữa $t > 0$, vì nếu $t < 0$ thì trong bất đẳng thức (2.7), khi cho $\mu \to \infty$ ta suy ra mâu thuẫn vì vế trái cố định.

Chia hai vế của (2.7) cho $t > 0$, ta được:

$$\langle \frac{p}{t}, x\rangle + f(x) \leqslant \langle \frac{p}{t}, y\rangle + \mu \qquad \forall y \in \operatorname{dom} f.$$

Thay $\mu = f(y)$, ta được

$$\langle \frac{p}{t}, x\rangle + f(x) \leqslant \langle \frac{p}{t}, y\rangle + f(y) \qquad \forall y \in \operatorname{dom} f.$$

Đặt $x^* = -\frac{p}{t}$, ta được

$$-\langle x^*, x\rangle + f(x) \leqslant -\langle x^*, y\rangle + f(y) \qquad \forall y \in \operatorname{dom} f.$$

Hay

$$\langle x^*, y - x\rangle + f(x) \leqslant f(y) \qquad \forall y \in \operatorname{dom} f.$$

Nếu $y \notin \operatorname{dom} f$ thì $f(y) = \infty$, do đó

$$\langle x^*, y - x\rangle + f(x) \leqslant f(y) \qquad \forall y.$$

Chứng tỏ $x^* \in \partial f(x)$. Vậy $\partial f(x) \neq \emptyset$.

Bây giờ ta chỉ ra tập $\partial f(x)$ compắc.

Do $x \in \operatorname{ri}(\operatorname{dom} f)$, theo mệnh đề (2.2)

$$x^* \in \partial f(x) \Longleftrightarrow f'(x, d) \geqslant \langle x^*, d\rangle \qquad \forall d. \tag{2.8}$$

Gọi F là không gian tuyến tính của $\operatorname{dom} f$. Lấy e^i là véc-tơ đơn vị thứ i (i=1,...,n) của R^n (toạ độ thứ i của e^i bằng 1 và mọi toạ độ khác là 0). Không giảm tổng quát, ta giả sử rằng các véc-tơ đơn vị $e^1, ...e^k \in F$, áp dụng (2.8) lần lượt với $d = e^i$ với i=1,...k, ta có $x_i^* \leqslant f'(x, e^i)$.

Tương tự , áp dụng với $d = -e^i$ với i=1,...k, ta có $-x_i^* \leqslant f'(x, -e^i)$. Hay $x_i^* \geqslant -f'(x, -e^i)$.

Tóm lại $-f'(x, -e^i) \leqslant x_i^* \leqslant f'(x, e^i)$,với mọi i=1,...k.

Theo (iv) mệnh đề (2.1), do $x \in \operatorname{ri}(\operatorname{dom} f)$ và F là không gian con của $\operatorname{dom} f$, nên $f'(x, y)$ hữu hạn với mọi $y \in F$. Nói riêng $f'(x, -e^i)$ và $f'(x, e^i)$ hữu hạn với mọi i=1,...k. Vậy $\partial f(x)$ bị chặn , và do tính đóng nên nó là compắc.

Ngược lại, giả sử rằng $\partial f(x) \neq \emptyset$ và $\partial f(x)$ compắc. Ta chỉ ra rằng $x \in \operatorname{ri}(\operatorname{dom} f)$.

Do $\partial f(x) \neq \emptyset$ nên $x \in \operatorname{dom} f$. Nếu trái lại $x \notin \operatorname{ri}(\operatorname{dom} f)$, thì x ở trên biên tương đối của $\operatorname{dom} f$.

Do $\operatorname{dom} f$ lồi, theo mệnh đề về siêu phẳng tựa, tồn tại một siêu phẳng tựa của $\operatorname{dom} f$ tại x, tức là tồn tại vectơ $p \in R^n, p \neq 0$ sao cho

$$\langle p, x\rangle \geqslant \langle p, z\rangle \qquad \forall z \in \operatorname{dom} f.$$

Lấy $x^* \in \partial f(x)$. Từ đây và theo định nghĩa dưới vi phân ta có:

$$\begin{aligned} f(z) - f(x) &\geqslant \langle x^*, z - x\rangle \\ &\geqslant \langle x^*, z - x\rangle + \lambda.\langle p, z - x\rangle \\ &= \langle x^* + \lambda.p, z - x\rangle \qquad \forall \lambda > 0, \forall z. \end{aligned}$$

Chứng tỏ $x^* + \lambda.p \in \partial f(x) \qquad \forall \lambda > 0$.

Điều này mâu thuẫn với tính bị chặn của $\partial f(x)$. Vậy $x \in \operatorname{ri}(\operatorname{dom} f)$. $\square$

Ví dụ 2.3. Cho hàm một biến

$$f(x) = \begin{cases} -2x^{\frac{1}{2}} & \text{nếu} \quad x \geqslant 0, \\ +\infty & \text{nếu} \quad x < 0. \end{cases}$$

Ta có

$$\operatorname{dom} f = [0; +\infty)\,, 0 \notin int(\operatorname{dom} f).$$

$$\begin{aligned} x^* \in \partial f(0) &\Leftrightarrow \langle x^*, x\rangle + f(0) \leqslant f(x)\,, \forall x \\ &\Leftrightarrow x^*.x \leqslant -2x^{\frac{1}{2}}, \forall x > 0. \end{aligned} \tag{2.9}$$

Nếu $x^* < 0$, ta chọn $x = 0.01$ thì (2.9) không thoả mãn.

Nếu $x^* \leqslant 0$ thì (2.9) không thoả mãn.

Vậy $\partial f(0) = \emptyset$.

Ví dụ trên cho thấy nếu $x \notin int(\operatorname{dom} f)$ thì tập $\partial f(x)$ có thể bằng rỗng.

Mệnh đề 2.4. *Cho $f : R^n \longrightarrow R \cup \{+\infty\}$ và $x \in \operatorname{dom} f$. Khi đó*

i) Nếu $x \in \operatorname{ri}(\operatorname{dom} f)$, thì $f'(x, y) = \max_{x^* \in \partial f(x)} \langle x^*, y\rangle\,, \forall y$.

ii) Với mọi tập bị chặn $C \subset int(\operatorname{dom} f)$, tập $\cup_{x \in C} \partial f(x)$ bị chặn .

iii) Nếu có thêm f đóng, thì

$$f^*(x^*) + f(x) = \langle x^*, x\rangle \Longleftrightarrow x^* \in \partial f(x), x \in \partial f(x^*).$$

Chứng minh. i) Do $f'(x,.)$ là hàm lồi, thuần nhất dương, nên mọi hàm non a-phin của $f'(x,.)$ đều tuyến tính, tức là có dạng $\langle p,.\rangle$. Vậy nếu $\langle p,.\rangle$ là hàm non a-phin của $f'(x,.)$ trên R^n, thì

$$\langle p, y\rangle \leqslant f'(x,y)\ ,\forall y.$$

Theo mệnh đề 2.2 ta có $p \in \partial f(x)$.

Hơn nữa, do $f'(x,.)$ là một hàm lồi đóng, nên theo định lý xấp xỉ tập lồi nó là bao trên của các hàm non a-phin của nó. Vậy

$$f'(x,y) = \mathrm{Sup}_{p\in\partial f(x)} \langle p, y\rangle.$$

ii) Giả sử $C \subseteq int(\mathrm{dom}\, f)$.

Đặt

$$\xi = \mathrm{Sup}_{x^*\in\partial f(C)} \|x^*\| = \mathrm{Sup}_{x\in C}\,\mathrm{Sup}_{x^*\in\partial f(C)} \|x^*\| \tag{2.10}$$

Xét ánh xạ tuyến tính $\langle x^*, z\rangle$. Chuẩn của ánh xạ tuyến tính này là

$$\|x^*\| = \mathrm{Sup}_{\|z\|=1} \langle x^*, z\rangle.$$

Thay vào (2.10) ta có :

$$\xi = \mathrm{Sup}_{x\in C}\,\mathrm{Sup}_{x^*\in\partial f(C)}\,\mathrm{Sup}_{\|z\|=1} \langle x^*, z\rangle.$$

Do

$$f'(x,z) = \mathrm{Sup}_{x^*\in\partial f(x)} \langle x^*, z\rangle$$

nên ta có tiếp

$$\xi = \mathrm{Sup}_{\|z\|=1}\,\mathrm{Sup}_{x\in C}\, f'(x,z).$$

Đặt $g(z) = \mathrm{Sup}_{x\in C}\, f'(x,z)$.

Do $x \in C \subseteq int(\mathrm{dom}\, f)$, nên hàm $f'(x,.)$ lồi trên R^n (do đó liên tục). Suy ra hàm g liên tục vì là bao trên của một họ hàm lồi liên tục trên R^n.

Vậy

$$\xi = \mathrm{Sup}_{\|z\|=1}\, g(z) = \max_{\|z\|=1} g(z) < +\infty.$$

Chứng tỏ $\partial f(C)$ bị chặn.

iii) Theo định nghĩa hàm liên hợp, ta có

$$f^*(x^*) = \mathrm{Sup}_x \{\langle x^*, x\rangle - f(x)\}.$$

Điều này tương đương với

$$f^*(x^*) \geqslant \langle x^*, y\rangle - f(y)\ , \forall y.$$

Do đó

$$\begin{aligned} f^*(x^*) + f(x) &= \langle x^*, x\rangle \\ \Leftrightarrow \langle x^*, y\rangle - f(y) + f(x) &\leqslant \langle x^*, x\rangle\ , \forall y. \end{aligned}$$

Hay

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

Vậy $x^* \in \partial f(x)$.
Do f đóng, nên theo hệ quả 1.1, ta có $f = f^{**}$.
Theo định nghĩa của hàm liên hợp, ta có:

$$f^{**} = \mathrm{Sup}_{x^*} \{\langle x, x^*\rangle - f^*(x^*)\}.$$

Điều này tương đương với

$$f^{**}(x) \geqslant \langle x, y\rangle - f^*(y)\ , \forall y.$$

Hay

$$f(x) \geqslant \langle x, y\rangle - f^*(y)\ , \forall y.$$

Do đó

$$\begin{aligned} f^*(x^*) + f(x) &= \langle x^*, x\rangle \\ \Leftrightarrow \langle x, y\rangle - f^*(y) + f^*(x^*) &\leqslant \langle x^*, x\rangle\ , \forall y. \end{aligned}$$

Hay

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

Vậy $x \in \partial f^*(x^*)$. $\square$

2.2.2 Tính khả vi của hàm lồi

Định nghĩa 2.3. Cho một hàm f xác định trên một lân cận của $x \in R^n$. Hàm f được gọi là khả vi tại x, nếu tồn tại x^* sao cho

$$\lim_{z \to x} \frac{f(z) - f(x) - \langle x^*, z - x \rangle}{\|z - x\|} = 0.$$

Một điểm x^* như thế nếu tồn tại sẽ duy nhất và được gọi là đạo hàm của f tại x. Thông thường đạo hàm này được kí hiệu là $\triangledown f(x)$ hoặc $f'(x)$.

Giả sử $f : R^n \longrightarrow R \cup \{+\infty\}$ lồi, chính thường và $x \in \operatorname{dom} f$. Nếu f khả vi tại x, thì với mọi $y \neq 0$, ta có:

$$\lim_{\lambda \to 0} \frac{f(x + \lambda.y) - f(x) - \langle \triangledown f(x), \lambda.y \rangle}{\lambda.\|y\|} = 0.$$

Hay là

$$\frac{f'(x, y) - \langle \triangledown f(x), y \rangle}{\|y\|} = 0.$$

Suy ra $f'(x, y) = \langle \triangledown f(x), y \rangle, \forall y$.

Lấy $y = e^i$ (i=1,...,n) là véc-tơ đơn vị thứ i của R^n, ta có :

$$\langle \triangledown f(x), e^i \rangle = (\frac{\partial f}{\partial x_i})(x)(i = 1, .., n).$$

Vậy

$$f'(x, y) = \sum_{i=1}^{n} y_i (\frac{\partial f}{\partial x_i})(x).$$

Từ đây ta có mệnh đề sau:

Mệnh đề 2.5. *Giả sử* $f : R^n \longrightarrow R \cup \{+\infty\}$ *lồi, chính thường và* $x \in \operatorname{dom} f$. *Khi đó* f *khả vi tại* x *khi và chỉ khi tồn tại* $x^* \in R^n$ *sao cho*

$$f'(x, y) = \langle x^*, y \rangle, \forall y.$$

Ngoài ra $x \in int(\operatorname{dom} f)$ *và* $\triangledown f(x) = x^*$.

Chứng minh. Nếu f khả vi tại x thì như ở trên, ta đã chỉ ra rằng

$$f'(x, y) = \langle \triangledown f(x), y \rangle, \forall y.$$

Vậy $f'(x,y)$ hữu hạn trên toàn R^n, nên $x \in int(\operatorname{dom} f)$.

Ngược lại $f'(x,y) = \langle \nabla f(x), y\rangle\ , \forall y$. Trước hết ta có $x \in int(\operatorname{dom} f)$ vì $f'(x,.)$ hữu hạn trên toàn R^n. Để chứng minh tính khả vi của f tại x, ta lấy

$$g(y) := f(x+y) - f(x) - \langle x^*, y\rangle.$$

Do f lồi, chính thường và f hữu hạn, nên g cũng là một hàm lồi, chính thường trên R^n. Ta cần chứng tỏ

$$\lim_{y\to 0} \frac{g(y)}{\|y\|} = 0.$$

Trước hết từ $f'(x,y) = \langle x^*, y\rangle$, theo định nghĩa của $f'(x,y)$, ta có

$$g(y) \geqslant 0\ , \forall y \text{ và } g(0) = 0.$$

Nếu $y \neq 0$ thì véc-tơ $\frac{y}{\|y\|}$ thuộc siêu hộp $H := [-1,1]^n$. Vậy theo định lý Krein-Milman điểm $\frac{y}{\|y\|}$ biểu diễn được bởi một tổ hợp lồi của các đỉnh của H, tức là tồn tại các số thực β_i (phụ thuộc y) sao cho

$$\beta_i \geqslant 0, \sum_{i\in I}\beta_i = 1 \text{ và } \frac{y}{\|y\|} = \sum_{i\in I}\beta_i v^i,$$

trong đó $v^i (i \in I)$ là các đỉnh của H.

Ta có

$$g(y) := g(\|y\|\frac{y}{\|y\|}) = g(\|y\|\sum_{i\in I}\beta_i v^i) = g(\sum_{i\in I}\beta_i\|y\|v^i).$$

Theo tính lồi của g thì

$$g(\sum_{i\in I}\beta_i\|y\|v^i) \leqslant \sum_{i\in I}\beta_i g(\|y\|v^i).$$

Tóm lại

$$0 \leqslant \frac{g(y)}{\|y\|} \leqslant \sum_{i\in I}\beta_i\frac{g(\|y\|v^i)}{\|y\|}.$$

Theo định nghĩa của g ta lại có

$$\frac{g(\|y\|v^i)}{\|y\|} \to 0 \text{ khi } y \to 0.$$

Vậy $\frac{g(y)}{\|y\|} \to 0$.

Chứng tỏ f khả vi tại x và do đó $\nabla f(x) = x^*$. □

Mệnh đề 2.6. *Cho* $f : R^n \longrightarrow R \cup \{+\infty\}$ *khả vi và* $C \subset R^n$*. Ba điều kiện sau tương đương.*

a) η *là hệ số lồi của* f *trên* C.

b) $f(y) \geqslant f(x) + \langle f'(x), y - x\rangle + \frac{\eta}{2}\|x - y\|^2 \qquad \forall x, y \in C$

c)$\langle f'(y) - f'(x), y - x\rangle \geqslant \eta\|x - y\|^2 \qquad \forall x, y \in C$

Chứng minh. (a)$\rightarrow$ (b):

Do η là hệ số lồi của trên C, nên với $t \in (0; 1)$ và mọi x,y thuộc C ta có:

$$f[t.y + (1-t)x] \leqslant tf(y) + (1-t)f(x) - \frac{\eta}{2}t(1-t)\|x-y\|^2.$$

Suy ra

$$\begin{aligned} f(y) - f(x) &\geqslant \frac{f[ty + (1-t)x] - f(x) + \frac{\eta}{2}t(1-t)\|x-y\|^2}{t} \\ &= \frac{f[x + t(y-x)] - f(x)}{t} + \frac{\eta}{2}(1-t)\|x-y\|^2. \end{aligned}$$

Cho $t \rightarrow 0$, do f khả vi, ta được:

$$f(y) - f(x) \geqslant f'(x, y - x) + \frac{\eta}{2}\|x - y\|^2.$$

(b)$\rightarrow$(a):

Cho $t \in (0; 1)$ và $\omega = (1-t)x + ty$. Khi đó

$$\begin{aligned} y &= \omega + (1-t)(y-x), \\ x &= \omega + (-t)(y-x). \end{aligned}$$

Ap dụng (b), ta được:

$$\begin{aligned} f(y) &\geqslant f(\omega) + \langle f'(\omega), y - \omega\rangle + \frac{\eta}{2}\|\omega - y\|^2, \\ f(x) &\geqslant f(\omega) + \langle f'(\omega), x - \omega\rangle + \frac{\eta}{2}\|\omega - x\|^2. \end{aligned}$$

Hay

$$\begin{aligned} f(y) &\geqslant f(\omega) + \langle f'(\omega), (1-t)(y-x)\rangle + \frac{\eta}{2}(1-t)^2\|y-x\|^2, \\ f(x) &\geqslant f(\omega) + \langle f'(\omega), (-t)(y-x)\rangle + \frac{\eta}{2}t^2\|y-x\|^2. \end{aligned}$$

Nhân bất đẳng thức trên với $t > 0$ và bất đẳng thức dưới với $1 - t > 0$ ta được :

$$\begin{aligned} tf(y) &\geqslant tf(\omega) + \langle f'(\omega), t(1-t)(y-x)\rangle \\ &+ \frac{\eta}{2}t(1-t)^2\|y-x\|^2, \\ (1-t)f(x) &\geqslant (1-t)f(\omega) + \langle f'(\omega), -t(1-t)(y-x)\rangle \\ &+ \frac{\eta}{2}t^2(1-t)\|y-x\|^2. \end{aligned}$$

Cộng hai bất đẳng thức trên và chuyển vế, ta có:

$$tf(y) + (1-t)f(x) \geqslant f(\omega) + \frac{\eta}{2}t(1-t)\|y-x\|^2.$$

Hay

$$f[(1-t)x + ty] \leqslant (1-t)f(x) + tf(y) - \frac{\eta}{2}t(1-t)\|y-x\|^2.$$

Chứng tỏ η là hệ số lồi của f trên C.

(b)→(c):

Do (b), nên $\forall x, y \in C$, ta có:

$$\begin{aligned} f(y) - f(x) &\geqslant \langle f'(x), y-x\rangle + \frac{\eta}{2}\|x-y\|^2, \\ f(x) - f(y) &\geqslant \langle f'(y), x-y\rangle + \frac{\eta}{2}\|x-y\|^2. \end{aligned}$$

Cộng hai bất đẳng thức lại ta được:

$$0 \geqslant \langle f'(x) - f'(y), y-x\rangle + \eta\|x-y\|^2.$$

Hay

$$\langle f'(y) - f'(x), y-x\rangle \geqslant \eta\|x-y\|^2.$$

(c)→(b):

Đặt

$$\gamma(t) = f[(1-t)x + ty] = f[x + t(y-x)].$$

Khi đó

$$\gamma'(t) = \langle f'[x + t(y-x)], y-x\rangle,$$

và

$$f(y) - f(x) = \gamma(1) - \gamma(0) = \int_0^1 \gamma'(t)dt.$$

Bây giờ giả sử có (c). Với $x, y \in C$, đặt $h := y - x$. Khi đó

$$\begin{aligned} f(y) - f(x) &= \int_0^1 \gamma'(t)dt \\ &= \int_0^1 \langle f'[x + t(y - x)], y - x\rangle dt \\ &= \int_0^1 f'(x + th).hdt \\ &= \int_0^1 [f'(x) + f'(x + th) - f'(x)]hdt \\ &= \int_0^1 (f'(x)h + [f'(x + th) - f'(x)]h)dt \\ &= f'(x)ht|_0^1 + \int_0^1 [f'(x + th) - f'(x)]hdt \\ &= f'(x)h + \int_0^1 [f'(x + th) - f'(x)]hdt. \end{aligned}$$

Theo (c) ta có :

$$\langle f'(x + th) - f'(x), th\rangle \geqslant \eta\|th\|^2,$$

hay

$$[f'(x + th) - f'(x)]h \geqslant \eta t\|h\|^2.$$

Suy ra

$$\begin{aligned} \int_0^1 [f'(x + th) - f'(x)]hdt &\geqslant \int_0^1 \eta t\|h\|^2 dt \\ &= \eta\frac{t^2}{2}\|h\|^2|_0^1 \\ &= \frac{\eta}{2}\|h\|^2. \end{aligned}$$

Vậy

$$f(y) - f(x) \geqslant f'(x)h + \frac{\eta}{2}\|h\|^2,$$

hay

$$f(y) - f(x) \geqslant \langle f'(x), y - x\rangle + \frac{\eta}{2}\|y - x\|^2.$$

$\square$

2.2.3 Tính đơn điệu của dưới vi phân

Cho T là một toán tử đa trị trên R^n, tức là với mỗi $x \in R^n$, thì $T(x)$ là một tập (có thể bằng rỗng). Như thường lệ ta ký hiệu tập hợp tất cả các tập con của R^n là 2^{R^n}.

Kí hiệu miền xác định của T là

$$\operatorname{dom} T := \{x \in R^n \mid T(x) \neq \emptyset\},$$

và đồ thị của T là

$$G(T) := \{(x, y) \in R^n \times R^n \mid y \in T(x)\}.$$

Định nghĩa 2.4. Cho $T : R^n \longrightarrow 2^{R^n}$ và $C \subseteq \operatorname{dom} T$.

Ta nói T là đơn điệu tuần hoàn trên C, nếu với mọi số nguyên dương m và mọi cặp $(x^i, y^i) \in G(T), x^i \in C$ (i=0,...,m) ta có:

$$\langle x^1 - x^0, y^0\rangle + \langle x^2 - x^1, y^1\rangle + ... + \langle x^0 - x^m, y^m\rangle \leqslant 0. \tag{2.11}$$

Nếu (2.11) chỉ đúng với $m = 1$, thì ta nói T đơn điệu trên C, tức là

$$\langle y - y', x - x'\rangle \geqslant 0\ , \forall x, x' \in C\ , \forall y \in T(x)\ , \forall y' \in T(x').$$

Nếu T đơn điệu (hoặc đơn điệu tuần hoàn) trên toàn $\operatorname{dom} T$, thì ta nói ngắn gọn là T đơn điệu (đơn điệu tuần hoàn).

Nếu $T \equiv \partial f$ thì T đơn điệu tuần hoàn trên $\operatorname{dom}(\partial f)$. Thật vậy:
$\forall m \in N\ , \forall (x^i, y^i) \in G(\partial f)\ , x^i \in \operatorname{dom}(\partial f)\ (i = 0, ...n)$ ta có:

$$G(\partial f) = \{(x^i, y^i) \mid y^i \in \partial f(x^i)\}.$$

Suy ra

$$\langle y^0, x^1 - x^0\rangle + f(x^0) \leqslant f(x^1)$$
$$\langle y^1, x^2 - x^1\rangle + f(x^1) \leqslant f(x^2)$$
$$...$$
$$...$$
$$...$$
$$\langle y^m, x^0 - x^m\rangle + f(x^m) \leqslant f(x^0).$$

Cộng vế với vế của các bất đẳng thức trên, ta được:

$$\langle y^0, x^1 - x^0\rangle + \langle y^1, x^2 - x^1\rangle + ... + \langle y^m, x^0 - x^m\rangle \leqslant 0.$$

Theo định nghĩa, $T \equiv \partial f$ đơn điệu tuần hoàn trên $\mathrm{dom}(\partial f)$.

Một câu hỏi được đặt ra là điều ngược lại có đúng không? Trả lời câu hỏi này ta có mệnh đề sau:

Mệnh đề 2.7. *Giả sử S là một toán tử đa trị từ $R^n \longrightarrow R^n$.*

Điều kiện cần và đủ để tồn tại một hàm lồi, đóng, chính thường f trên R^n sao cho $S(x) \subseteq \partial f(x)$,$\forall x$ là toán tử S đơn điệu tuần hoàn.

Chứng minh. Điều kiện cần: Nếu tồn tại một hàm f lồi, đóng, chính thường trên R^n sao cho $S(x) \subseteq \partial f(x)$,$\forall x$ thì S là toán tử đơn điệu tuần hoàn.

$\forall m \in N$,$\forall (x^i, y^i) \in G(S), x^i \in \mathrm{dom}\, S(i = 0, ...m)$.

Từ $(x^i, y^i) \in G(S) \Rightarrow y^i \in S(x^i) \subseteq \partial f(x^i), (\forall i = 0, ..m)$, do ∂f là đơn điệu tuần hoàn trên $\mathrm{dom}(\partial f)$ nên

$$\langle y^0, x^1 - x^0\rangle + \langle y^1, x^2 - x^1\rangle + ... + \langle y^m, x^0 - x^m\rangle \leqslant 0.$$

Suy ra S là đơn điệu tuần hoàn trên $\mathrm{dom}\, S$.

Điều kiện đủ: Nếu S là toán tử đơn điệu tuần hoàn thì tồn tại một hàm f lồi, đóng, chính thường trên R^n sao cho $S(x) \subseteq \partial f(x)$,$\forall x$

Giả sử $(x^0, y^0) \in G(S)$, định nghĩa hàm f bằng cách lấy

$$f(x) := \mathrm{Sup}\{\langle x - x^m, y^m\rangle + ... + \langle x^1 - x^0, y^0\rangle\},$$

trong đó cận trên đúng được lấy trên tất cả các cặp $(x^i, y^i) \in G(S)$ và các số nguyên dương m.

Ta chứng minh: f lồi, đóng, chính thường và $S(x) \subseteq \partial f(x)$, $\forall x$.

Do f là bao trên của một họ các hàm a-phin, nên f là một hàm lồi đóng.

Do tính đơn điệu tuần hoàn của S, nên

$$f(x^0) := \text{Sup}\{\langle x^o - x^m, y^m\rangle + \langle x^m - x^{m-1}, y^{m-1}\rangle + ... + \langle x^1 - x^0, y^0\rangle\} := 0,$$

suy ra $\text{dom} f \neq \emptyset$. Vậy f là chính thường.

Với bất kì cặp $(x, x^*) \in G(S)$, ta có $x^* \in S(x)$, ta sẽ chứng minh $x^* \in \partial f(x)$. Muốn thế ta sẽ chứng minh rằng

$$\forall \alpha < f(x) \text{ và } y \in R^n \text{ , ta có } \alpha + \langle x^*, y - x\rangle < f(y).$$

Thật vậy, do $\alpha < f(x)$, nên theo tính chất của cận trên đúng, sẽ tồn tại các cặp $(x^i, y^i) \in G(S)$ và số nguyên dương m (i=1,...m) thỏa mãn

$$\alpha < \langle x - x^m, y^m\rangle + \langle x^m - x^{m-1}, y^{m-1}\rangle + ... + \langle x^1 - x^0, y^0\rangle.$$

Theo định nghĩa của $f(y)$, ta được:

$$\begin{aligned} f(y) &\geqslant \langle y - x^m, y^m\rangle + ... + \langle x^1 - x^0, y^0\rangle \\ &= \langle y - x^{m+1}, y^m\rangle + \langle x^{m+1} - x^m, y^m\rangle + ... + \langle x^1 - x^0, y^0\rangle. \end{aligned}$$

Thay $x^{m+1} = x$, $y^m = x^*$, ta có

$$\begin{aligned} f(y) &\geqslant \langle y - x, x^*\rangle + \langle x - x^m, y^m\rangle + \langle x^m - x^{m-1}, y^{m-1}\rangle \\ &\quad + ... + \langle x^1 - x^0, y^0\rangle \\ &> \langle y - x, x^*\rangle + \alpha. \end{aligned}$$

Điều này đúng với mọi $(x, x^*) \in G(S)$ nên $S(x) \subset \partial f(x)$, $\forall x$. □

Định nghĩa 2.5. Ta nói một toán tử $T : R^n \longrightarrow 2^{R^n}$ là đơn điệu cực đại nếu nó là đơn điệu và đồ thị của nó không phải là tập con thực sự của đồ thị của một toán tử đơn điệu nào khác.

Toán tử T được gọi là đơn điệu tuần hoàn cực đại, nếu nó là đơn điệu tuần hoàn và đồ thị của nó không phải là tập con thực sự của đồ thị của một toán tử đơn điệu tuần hoàn nào khác.

Ví dụ 2.4. (Toán tử đơn điệu)

Xét $N_C(x) := \{\omega \mid \langle \omega, y - x\rangle \leqslant 0\ ,\ \forall y \in C\}$.

Ta chứng tỏ rằng nón pháp tuyến có tính chất đơn điệu theo nghĩa

$$\langle \omega - \omega', x - x'\rangle \geqslant 0 \quad \forall x, x' \in C\ ,\ \forall \omega \in N_C(x)\ ,\ \forall \omega' \in N_C(x').$$

Thật vậy: $\forall x, x' \in C$ ta có:

+ $\omega \in N_C(x) \Leftrightarrow \langle \omega, y - x\rangle \leqslant 0 \quad \forall y \in C$. Với $y = x'$,
ta có $\langle \omega, x' - x\rangle \leqslant 0$.

+ $\omega' \in N_C(x') \Leftrightarrow \langle \omega', y - x'\rangle \leqslant 0 \quad \forall y \in C$. Với $y = x$,
ta có $\langle \omega', x - x'\rangle \leqslant 0$.

$\Rightarrow \langle \omega, x' - x\rangle + \langle \omega', x - x'\rangle \leqslant 0$.

$\Rightarrow \langle \omega - \omega', x - x'\rangle \geqslant 0$.

Ví dụ 2.5. (Toán tử đơn điệu)

Xét ánh xạ

$$\begin{aligned} f : R^2 &\longrightarrow R^2 \\ x &\longmapsto f(x) = Qx = (x_2, -x_1). \end{aligned}$$

Với $x = (x_1, x_2)$, $Q = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Với mọi $x = (x_1, x_2), y = (y_1, y_2) \in R^2$ ta có:

+ $x - y = (x_1 - y_1, x_2 - y_2)$.

+ $f(x) - f(y) = (x_2 - y_2, -x_1 + y_1)$.

Suy ra

$$\begin{aligned} \langle f(x) - f(y), x - y\rangle &= (x_2 - y_2)(x_1 - y_1) + (-x_1 + y_1)(x_2 - y_2) \\ &= 0 \quad \forall x, y \in R^2. \end{aligned}$$

Vậy f là ánh xạ đơn điệu trên R^2.

Hệ quả 2.1. *Mọi toán tử đơn điệu tuần hoàn cực đại trong R^n đều là dưới vi phân của một hàm lồi, đóng, chính thường trên R^n.*

Chứng minh. Giả sử S là toán tử đơn điệu tuần hoàn cực đại trong R^n. Theo định nghĩa ta có :

+ S là toán tử đơn điệu tuần hoàn .

+ G(S) không là tập con thực sự của đồ thị của một toán tử đơn điệu tuần hoàn nào khác.

Do S là toán tử đơn điệu tuần hoàn nên theo mệnh đề 2.7, tồn tại một hàm lồi, đóng, chính thường f trên R^n sao cho $S(x) \subseteq \partial f(x)$, $\forall x$.

Ta có :$\forall (x, y) \in G(S) \Rightarrow y \in S(x)$ do $S(x) \subseteq \partial f(x), \forall x$ nên $y \in \partial f(x)$ $\Rightarrow (x, y) \in G(\partial f)$. Vậy $G(S) \subset G(\partial f)$.

Do G(S) không là tập con thực sự của đồ thị của một toán tử đơn điệu tuần hoàn nào khác và ∂f là toán tử đơn điệu tuần hoàn nên $G(S) = G(\partial f)$. Suy ra $S \equiv \partial f$. □

2.2.4 Tính liên tục của dưới vi phân

Định nghĩa 2.6. Một ánh xạ $T : R^n \longrightarrow 2^{R^n}$ được gọi là đóng tại x, nếu với mọi dãy $x^k \to x$, mọi $y^k \in T(x^k)$ và $y^k \to y$ thì $y \in T(x)$

Định nghĩa 2.7. Một ánh xạ $T : R^n \longrightarrow 2^{R^n}$ được gọi là nửa liên tục trên tại x, nếu với mọi tập mở G chứa T(x), tồn tại một lân cận mở U của x sao cho

$$T(z) \subset G, \forall z \in U.$$

Ta nói ánh xạ T là đóng (nửa liên tục trên) trên tập C, nếu nó đóng (nửa liên tục trên) tại mọi điểm thuộc C.

Một ánh xạ T được gọi là đóng nếu đồ thị của T là một tập đóng.

Nói một cách khái quát, bổ đề dưới đây chỉ ra rằng: Một dãy hàm lồi nếu bị chặn trên bởi một hàm lồi theo từng điểm ở trên một tập lồi, mở, thì sẽ bị chặn trên đều bởi chính hàm lồi đó trên mọi tập compắc thuộc tập mở này.

Bổ đề 2.1. *Cho một tập lồi, mở $G \subseteq R^n$ và f là một hàm lồi nhận giá trị hữu hạn trên G. Giả sử $\{f_i\}_i \in I$ là một dãy các hàm lồi hữu hạn trên G và hội tụ theo từng điểm trên G đến f. Giả sử* $\limsup f_i(x) \leqslant f(x)$, $\forall x \in G$.

Khi đó với mọi tập compắc $K \subseteq G$, với mọi $\epsilon > 0$, tồn tại chỉ số i_ϵ sao cho

$$f_i(x) \leqslant f(x) + \epsilon\,, \forall i \geqslant i_\epsilon\,, \forall x \in K.$$

Chứng minh. Với mọi $x \in G$, mọi $i \in N$, định nghĩa

$$g_i(x) := \max\{f_i(x), f(x)\}.$$

Hàm g_i lồi, hữu hạn trên G vì nó là hàm bao trên của hai hàm lồi, hữu hạn trên G.

Do G mở, nên g_i liên tục trên G.

Do $K \subseteq G$ compắc, nên dãy $\{g_i(x)\}_i \in I$ bị chặn. Không giảm tổng quát, bằng cách qua dãy con, ta có thể coi

$$g_i(x) \to l(x) \text{ khi } i \to +\infty.$$

Theo định nghĩa của $g_i(x)$ và do $\limsup f_i(x) \leqslant f(x)\ \forall x \in K$, ta suy ra $l(x) = f(x)$.

Vậy dãy $g_i(x)$ hội tụ theo từng điểm đến f trên tập compắc K, nên nó hội tụ đều đến f trên K.

Nhưng do $g_i := \max\{f_i(x), f(x)\}$, nên với mọi tập compắc $K \subseteq G$, với mọi $\epsilon > 0, \exists i_\epsilon : \forall i \geqslant i_\epsilon$ ta có

$$f_i(x) \leqslant f(x) + \epsilon\,, \forall x \in K.$$

□

Từ mệnh đề sau, suy ra tính nửa liên tục trên của ánh xạ ∂f. Cụ thể là:

Mệnh đề 2.8. *Cho một tập lồi, mở $U \subseteq R^n$ và f là một hàm lồi nhận giá trị hữu hạn trên U. Giả sử $\{f_i\}_i \in I$ là một dãy các hàm lồi hữu hạn trên U và hội tụ theo từng điểm trên U đến f.*

Khi đó, nếu dãy $\{x^i\} \subset U$ hội tụ đến $x \in U$, thì với mọi $\epsilon > 0$, tồn tại chỉ số i_ϵ sao cho

$$\partial f_i(x^i) \subset \partial f(x) + \epsilon.\overline{B(0,1)}\,, \forall i \geqslant i_\epsilon,$$

trong đó $\overline{B(0,1)}$ là hình cầu đơn vị đóng tâm ở O

Chứng minh. Cho $\alpha > 0$, $y \in R^n$. Với mỗi $x \in U$ đặt

$$\mu := f'(x, y) + \alpha.$$

Do f lồi, hữu hạn trên tập U mở và $x \in U$, nên $x \in int(\mathrm{dom}\, f)$, do đó μ hữu hạn.

Do $x \in int(\mathrm{dom}\, f)$, nên với mọi y, tồn tại $\delta > 0$ sao cho

$$x + \lambda.y \in int(\mathrm{dom}\, f) \text{ với mọi } 0 < \lambda < \delta.$$

Do $\alpha > 0$ và định nghĩa của $f'(x, y)$, ta có :

$$\frac{f(x + \lambda.y) - f(x)}{\lambda} < \mu\,, \forall \lambda \in (0, \delta). \tag{2.12}$$

Do $f_i(x + \lambda.y) \to f(x + \lambda.y)$ và $f_i(x^i) \to f(x)$, nên từ (2.12), tồn tại i_1 sao cho

$$\frac{f_i(x^i + \lambda.y) - f_i(x^i)}{\lambda} < \mu\,, \forall i \geqslant i_1\,, \forall \lambda \in (0, \delta).$$

Do

$$f_i'(x^i, y) \leqslant \frac{f_i(x^i + \lambda.y) - f_i(x^i)}{\lambda}$$

nên $f_i'(x^i, y) \leqslant \mu$ với mọi $i \geqslant i_1$.

Vậy

$$\limsup f_i'(x^i, y) \leqslant \mu = f'(x, y) + \alpha.$$

Do điều này đúng với mọi $\alpha > 0$, ta suy ra

$$\limsup f_i'(x^i, y) < f'(x, y).$$

Vì $f_i'(x^i, .)$ và $f'(x, .)$ lồi, hữu hạn trên U (do $x \in U \subset int(\mathrm{dom}\, f)$) nên áp dụng bổ đề 2.1 cho các hàm lồi $f_i'(x^i, .)$ và $f'(x, .)$ với $G = R^n$, $K = \overline{B(0, 1)}$ ta có :

Với mọi tập compắc $\overline{B(0, 1)} \subseteq R^n$, $\forall \epsilon > 0$, $\exists i_\epsilon$ sao cho $\forall i \geqslant i_\epsilon$ ta có

$$f_i'(x^i, y) \leqslant f'(x, y) + \epsilon\,, \forall y \in \overline{B(0, 1)}.$$

Từ đây, với mọi $y \neq 0$, theo tính chất thuần nhất dương của $f'(x, .)$, ta có:

$$\frac{1}{\|y\|} f_i'(x^i, y) = f_i'(x^i, \frac{y}{\|y\|}) \leqslant f'(x, \frac{y}{\|y\|}) + \epsilon.$$

Hay

$$f_i'(x^i, y) \leqslant f'(x, y) + \epsilon.\|y\|\ , \forall i \geqslant i_\epsilon\ , \forall y.$$

Do $f_i'(x^i, y)$ là hàm tựa của $\partial f_i(x^i)$ và $f'(x, y)$ là hàm tựa của $\partial f(x)$ nên từ đây suy ra

$$\partial f_i(x^i) \subseteq \partial f(x) + \epsilon.\overline{B(0,1)}.$$

□

Mệnh đề 2.9. *Cho f là một hàm lồi chính thường trên R^n. Khi đó ánh xạ dưới vi phân $x \longrightarrow \partial f(x)$ nửa liên tục trên tại mọi điểm $x \in int(\mathrm{dom}\, f)$*

Chứng minh. Ta có f lồi nên nó hữu hạn trên tập $int(\mathrm{dom}\, f)$. Vậy áp dụng mệnh đề 2.8 với $U = int(\mathrm{dom}\, f)$ và $f_i = f, \forall i$, ta có:

Nếu $x \in int(\mathrm{dom}\, f)$ và $\{x^i\} \subset int(\mathrm{dom}\, f)$ hội tụ đến x thì với mọi $\epsilon > 0$, tồn tại i_ϵ sao cho $\forall i \geqslant i_\epsilon$ ta có

$$\partial f(x^i) \subset \partial f(x) + \epsilon.\overline{B(0,1)}.$$

Suy ra ∂f nửa liên tục trên tại mọi điểm $x \in int(\mathrm{dom}\, f)$. □

Mệnh đề 2.10. *Cho f là một hàm lồi chính thường trên R^n.*

Nếu f khả vi trên tập $int(\mathrm{dom}\, f)$ thì nó khả vi liên tục trên tập này.

Chứng minh. Cho $x \in int(\mathrm{dom}\, f)$ và dãy$\{x^i\} \subset int(\mathrm{dom}\, f)$. Ta áp dụng mệnh đề 2.8 với $U = int(\mathrm{dom}\, f)$ và $\nabla f_i = \nabla f\ , \forall i$, ta có:

$\forall \epsilon > 0$ tồn tại i_ϵ sao cho

$$\forall i \geqslant i_\epsilon, \|\nabla f(x^i) - \nabla f(x)\| \leqslant \epsilon.$$

Chứng tỏ $\nabla f(.)$ liên tục tại x. □

2.2.5 Phép tính với dưới đạo hàm

Tương tự giải tích cổ điển, nếu f là một hàm lồi chính thường trong R^n thì

$$\partial(\lambda f)(x) = \lambda.\partial f(x)\ , \forall \lambda > 0\ , \forall x.$$

Thật vậy

$$\begin{aligned} x^* \in \partial(\lambda f)(x) &\Leftrightarrow \langle x^*, z - x\rangle + (\lambda f)(x) \leqslant (\lambda f)(z)\,, \forall z \\ &\Leftrightarrow \langle x^*, z - x\rangle + \lambda f(x) \leqslant \lambda f(z)\,, \forall z \\ &\Leftrightarrow \langle \frac{x^*}{\lambda}, z - x\rangle + f(x) \leqslant f(z)\,, \forall z \\ &\Leftrightarrow \frac{x^*}{\lambda} \in \partial f(x) \\ &\Leftrightarrow x^* \in \lambda \partial f(x). \end{aligned}$$

Đối với dưới vi phân của tổng các hàm lồi, ta có định lý sau:

Mệnh đề 2.11. *(Định lý Moreau-Rockafellar). Cho f_i, i=1,...,m là các hàm lồi chính thường trên R^n. Khi đó*

$$\sum_{i=1}^{m} \partial f_i(x) \subseteq \partial(\sum_{i=1}^{m} f_i(x))\,, \forall x.$$

Nếu $\cap\, \mathrm{ri}(\mathrm{dom}\, f_i) \neq \emptyset$ *thì*

$$\sum_{i=1}^{m} \partial f_i(x) = \partial(\sum_{i=1}^{m} f_i(x))\,, \forall x$$

Chứng minh. Ta chứng minh

$$\sum_{i=1}^{m} \partial f_i(x) \subseteq \partial(\sum_{i=1}^{m} f_i(x)), \forall x.$$

Nếu $x^* \in \sum_{i=1}^{m} \partial f_i(x)$ thì $x^* = \sum_{i=1}^{m} x_i^*\,, x_i^* \in \partial f_i(x), i = 1, .., m$. Ta có

$$\begin{aligned} x_i^* \in \partial f_i(x), i = 1, ..., m &\Leftrightarrow \langle x_i^*, z - x\rangle + f_i(x) \leqslant f_i(z)\,, \forall z\,,\ i = 1, ..., m \\ &\Rightarrow \langle \sum_{i=1}^{m} x_i^*, z - x\rangle + \sum_{i=1}^{m} f_i(x) \leqslant \sum_{i=1}^{m} f_i(z), \forall z \\ &\Leftrightarrow \langle x^*, z - x\rangle + \sum_{i=1}^{m} f_i(x) \leqslant \sum_{i=1}^{m} f_i(z), \forall z \\ &\Leftrightarrow x^* \in \partial(\sum_{i=1}^{m} f_i(x))\,, \forall x. \end{aligned}$$

Vậy $\sum_{i=1}^{m} \partial f_i(x) \subseteq \partial(\sum_{i=1}^{m} f_i(x)), \forall x.$

Ta chứng minh

$$\sum_{i=1}^{m} \partial f_i(x) \supseteq \partial(\sum_{i=1}^{m} f_i(x)) \,, \forall x.$$

Chỉ cần chứng minh với m = 2, với $m > 2$ dùng quy nạp.

Lấy $x^0 \in R^n$ và

$$x^* \in \partial(\sum_{i=1}^{2} f_i(x^0)) = \partial(f_1(x^0) + f_2(x^0)) = \partial(f_1 + f_2)(x^0).$$

Theo định nghĩa của dưới vi phân ta có:

$$\begin{aligned}
&\langle x^*, x - x^0\rangle + f_1(x^0) + f_2(x^0) \leqslant f_1(x) + f_2(x) \qquad \forall x\\
\Leftrightarrow &f_1(x) + f_2(x) - f_1(x^0) - f_2(x^0) - \langle x^*, x - x^0\rangle \geqslant 0 \qquad \forall x\\
\Leftrightarrow &\begin{cases} f_1(x) + f_2(y) - f_1(x^0) - f_2(x^0) - \langle x^*, x - x^0\rangle < 0\\ x = y\end{cases}
\end{aligned}$$

không có nghiệm.

Lấy $D = \mathrm{dom}\, f_1 \times \mathrm{dom}\, f_2$ và $A(x, y) = x - y$. Theo giả thiết f_1 liên tục tại một điểm $a \in \mathrm{dom}\, f_1 \cap \mathrm{dom}\, f_2$, nên tồn tại một lân cận U của gốc sao cho

$$U = (a + U) - a \subset \mathrm{dom}\, f_1 - \mathrm{dom}\, f_2 = A(D).$$

Vậy $0 \in intA(D)$, áp dụng mệnh đề 1.4 với

$$f(x, y) = f_1(x) + f_2(y) - f_1(x^0) - f_2(x^0) - \langle x^*, x - x^0\rangle$$

$$A(x, y) = x - y.$$

Ta có

$$\langle t, x - y\rangle + [f_1(x) + f_2(y) - f_1(x^0) - f_2(x^0) - \langle x^*, x - x^0\rangle] \geqslant 0,$$

$\forall x \in \mathrm{dom}\, f_1 \,, y \in \mathrm{dom}\, f_2$.

Đối với $x \notin \mathrm{dom}\, f_1$ và $y \notin \mathrm{dom}\, f_2$, thì bất đẳng thức trên là hiển nhiên.

Vậy

$$\langle t, x - y\rangle + [f_1(x) + f_2(y) - f_1(x^0) - f_2(x^0) - \langle x^*, x - x^0\rangle] \geqslant 0 \,, \forall x, y.$$

Lấy $x = x^0$ ta có

$$\langle t, y - x^0 \rangle + f_2(x^0) \leqslant f_2(y)\,, \forall y.$$

Chứng tỏ $t \in \partial f_2(x^0)$.
Lấy $y = x^0$ ta có

$$\langle x^* - t, x - x^0 \rangle + f_1(x^0) \leqslant f_1(x)\,, \forall x.$$

Chứng tỏ $x^* - t \in \partial f_1(x^0)$.
Do đó $x^* = (x^* - t) + t \in \partial f_1(x^0) + \partial f_2(x^0)$. □

2.3 Dưới vi phân xấp xỉ

Dưới vi phân xấp xỉ, hay còn được gọi là ϵ-dưới vi phân, thường được sử dụng trong thực tế bởi hai lý do chính sau đây. Một là hàm lồi có thể không khả dưới vi phân tại những điểm thuộc biên của miền hữu dụng của nó, trong khi đó, như sẽ thấy dưới đây, trong miền này, dưới vi phân xấp xỉ luôn tồn tại. Lý do thứ hai quan trọng hơn là trong ứng dụng, thường người ta chỉ cần, và nhiều khi chỉ tính được dưới vi phân một cách xấp xỉ.

Định nghĩa 2.8. Cho $\epsilon \geqslant 0$. Một véctơ x^* được gọi là ϵ-dưới đạo hàm của f tại x, nếu

$$\langle x^*, y - x \rangle + f(x) \leqslant f(y) + \epsilon\,, \forall y.$$

Kí hiệu tập hợp tất cả ϵ-dưới đạo hàm của f tại x là $\partial_\epsilon f(x)$. Tập hợp này được gọi là ϵ-dưới vi phân.

Hiển nhiên $\partial_0 f(x) = \partial f(x)$. Vậy dưới đạo hàm xấp xỉ là một khái niệm tổng quát hoá của dưới đạo hàm chính xác.

Nhận xét:
Theo định nghĩa

$$x^* \in \partial_\epsilon f(x) \Leftrightarrow \langle x^*, z - x \rangle + f(x) \leqslant f(z) + \epsilon\,, \forall z.$$

Thay $z = y + x$, ta được

$$\langle x^*, y\rangle + f(x) \leqslant f(x+y) + \epsilon\,, \forall y.$$

Từ đây, nếu đặt $h(y) = f(x+y) - f(x)$, ta có

$$\langle x^*, y\rangle - h(y) \leqslant \epsilon\,, \forall y.$$

Theo định nghĩa hàm liên hợp, ta có

$$\partial_\epsilon f(x) = \{x^* | h^*(x^*) \leqslant \epsilon\}.$$

Do h^* là một hàm lồi, đóng nên từ đây thấy rằng $\partial_\epsilon f(x)$ luôn là một tập lồi, đóng .

Hiển nhiên $\partial_\epsilon f(x) \subseteq \partial_{\epsilon'} f(x)$ nếu $\epsilon \leqslant \epsilon'$. Hơn nữa $\cap_{\epsilon>0}\partial_\epsilon f(x)$ nếu khác rỗng thì sẽ bằng $\partial f(x)$.

Ví dụ 2.6. Cho hàm một biến

$$f(x) = \begin{cases} -2x^{\frac{1}{2}} & \text{nếu} \quad x \geqslant 0, \\ +\infty & \text{nếu} \quad x < 0. \end{cases}$$

Ta có $\partial_\epsilon f(0) \neq \emptyset$ với $\epsilon > 0$. Thật vậy, lấy $x^* \in \partial_\epsilon f(0)$, ta có:

$$\begin{aligned} x^* \in \partial_\epsilon f(0) &\Leftrightarrow \langle x^*, y - 0\rangle + f(0) \leqslant f(y) + \epsilon\,, \forall y \\ &\Leftrightarrow x^* y \leqslant -2\sqrt{y} + \epsilon\,, \forall y \geqslant 0. \end{aligned}$$

Nếu $y = 0$ thì $0 \leqslant \epsilon$. Vậy bất đẳng thức trên được thoả mãn với mọi x^*.
Nếu $y > 0$ thì $x^* \leqslant \frac{\epsilon - 2\sqrt{y}}{y}$.
Vậy $\partial_\epsilon f(0) \neq \emptyset$ với $\epsilon > 0$.
Mặt khác $\partial f(0) = \emptyset$ (đã chứng minh phần trước).
Ví dụ trên cho thấy rằng $\partial_\epsilon f(x) \neq \emptyset$ với mọi $\epsilon > 0$, tuy nhiên $\partial f(x) = \emptyset$.

Mệnh đề sau đây nói rằng mọi hàm lồi đóng đều khả ϵ-dưới vi phân với mọi $\epsilon > 0$ tại mọi điểm thuộc miền hữu dụng của nó .

Mệnh đề 2.12. *Cho f là một hàm lồi đóng chính thường trên R^n.*

Khi đó với mọi $\epsilon > 0$ và mọi $x^0 \in \operatorname{dom} f$, tập $\partial_\epsilon f(x^0) \neq \emptyset$. Hơn nữa với mọi tập bị chặn $C \subset int(\operatorname{dom} f)$ tập $\partial_\epsilon f(C)$ bị chặn.

Chứng minh. Do f lồi, đóng, chính thường nên $\operatorname{epi} f$ lồi, đóng, khác rỗng.

Do $\operatorname{epi} f$ đóng nên điểm $(x^0, f(x^0) - \epsilon) \notin \operatorname{epi} f$ với mọi $\epsilon > 0$. Ta áp dụng định lý tách chặt cho hai tập $C := \operatorname{epi} f$ và $D := \{(x^0, f(x^0) - \epsilon)\}$, sẽ tồn tại $(p, t) \neq 0, p \in R^n, t \in R$ và số thực η (phụ thuộc ϵ) sao cho:

$$\langle p, t\rangle - t\nu < \eta < \langle p, x^0\rangle - t[f(x^0) - \epsilon]\,, \forall (x, \nu) \in \operatorname{epi} f. \qquad (2.13)$$

Từ đây ta có $t \neq 0$, vì nếu t=0 thì

$$\langle p, x\rangle < \eta < \langle p, x^0\rangle\,, \forall x \in \operatorname{dom} f$$

và do đó ta có mâu thuẫn nếu lấy $x = x^0$.

Hơn nữa $t > 0$, vì nếu $t < 0$ ta sẽ có mâu thuẫn từ (2.13) khi cho ν đủ lớn .

Thay $\nu = f(x)$ và chia hai vế của (2.13) cho $t > 0$, ta có

$$\langle \frac{p}{t}, x\rangle - f(x) < \frac{\eta}{t} < \langle \frac{p}{t}, x^0\rangle - f(x^0) + \epsilon.$$

Suy ra

$$\langle \frac{p}{t}, x - x^0\rangle + f(x^0) < f(x) + \epsilon.$$

Vậy $\frac{p}{t} \in \partial_\epsilon f(x^0)$ hay $\partial_\epsilon f(x^0) \neq \emptyset$.

Giả sử $C \subset int(\operatorname{dom} f)$, C bị chặn.

Đặt

$$\xi = \operatorname{Sup}_{x^* \in \partial_\epsilon f(C)} \|x^*\| = \operatorname{Sup}_{x \in C} \operatorname{Sup}_{x^* \in \partial_\epsilon f(C)} \|x^*\|. \qquad (2.14)$$

Xét ánh xạ tuyến tính $\langle x^*, z\rangle$. Chuẩn của ánh xạ tuyến tính này là

$$\|x^*\| = \operatorname{Sup}_{\|z\|=1} \langle x^*, z\rangle.$$

Thay vào (2.14) ta có

$$\xi = \operatorname{Sup}_{x \in C} \operatorname{Sup}_{x^* \in \partial_\epsilon f(x)} \operatorname{Sup}_{\|z\|=1} \langle x^*, z\rangle.$$

Mặt khác

$$x^* \in \partial_\epsilon f(x) \Leftrightarrow f'(x, y) + \beta \geqslant \langle x^*, y\rangle\,, \forall \beta > 0\,, \forall y.$$

Do đó

$$f'(x,y) + \beta = \mathrm{Sup}_{x^* \in \partial_\epsilon f(x)} \langle x^*, y\rangle \ , \forall \beta > 0 \ , \forall y.$$

Hay

$$f'(x,z) + \beta = \mathrm{Sup}_{x^* \in \partial_\epsilon f(x)} \langle x^*, z\rangle \ , \forall \beta > 0 \ , \forall z.$$

Ta có tiếp

$$\begin{aligned} \xi &= \mathrm{Sup}_{\|z\|=1} \mathrm{Sup}_{x\in C} (f'(x,z) + \beta) \\ &= \mathrm{Sup}_{\|z\|=1} \mathrm{Sup}_{x\in C} f'(x,z) + \beta. \end{aligned}$$

Đặt $g(z) := \mathrm{Sup}_{x\in C} f'(x,z)$.

Do $x \in C \subseteq int(\mathrm{dom}\, f)$, nên hàm $f'(x,z)$ lồi trên R^n (do đó liên tục). Suy ra hàm g liên tục vì là bao trên của một họ hàm lồi liên tục trên R^n.

Vậy

$$\xi = \mathrm{Sup}_{\|z\|=1} g(z) + \beta = \max_{\|z\|=1} g(z) + \beta < +\infty.$$

Chứng tỏ $\partial_\epsilon f(C)$ bị chặn. □

Mệnh đề 2.13. *Cho $f : R^n \longrightarrow R$ là hàm lồi đóng.*

Khi đó một véc-tơ x^* là ϵ- dưới vi phân của f tại $x \in \mathrm{dom}\, f$ khi và chỉ khi

$$f^*(x^*) + f(x) - \langle x^*, x\rangle \leqslant \epsilon.$$

Chứng minh. Dùng định nghĩa ϵ-dưới đạo hàm ta có:

$$\begin{aligned} x^* \in \partial_\epsilon f(x) &\Leftrightarrow \langle x^*, y - x\rangle + f(x) \leqslant f(y) + \epsilon \ , \forall y \in \mathrm{dom}\, f \\ &\Leftrightarrow [\langle x^*, y\rangle - f(y)] + f(x) - \langle x^*, x\rangle \leqslant \epsilon \ , \forall y \in \mathrm{dom}\, f. \end{aligned}$$

Theo định nghĩa của $f^*(x^*)$ ta có:

$$f^*(x^*) = \mathrm{Sup}_y \{\langle x^*, y\rangle - f(y)\}.$$

Vậy $f^*(x^*) + f(x) - \langle x^*, x\rangle \leqslant \epsilon$. □

Định nghĩa 2.9. Cho $C \subseteq R^n$ là một tập lồi, đóng và $x \in C$. Tập ϵ-nón pháp tuyến ngoài của C tại x là ϵ-dưới vi phân của hàm chỉ của C tại x.

Tức là:

$$N_{C,\epsilon}(x) := \partial_\epsilon \delta_C(x) = \{x^* \in R^n \mid \langle x^*, y - x\rangle \leqslant \epsilon \,, \forall y \in C\}$$

Định lý 2.1. *Cho f là hàm lồi, đóng, chính thường trên R^n.*

$$0 \in \partial_\epsilon f(x) \Leftrightarrow f(x) \leqslant f(y) + \epsilon \,, \forall y \in R^n.$$

Chứng minh. Theo định nghĩa dưới vi phân xấp xỉ ta có:

$$\begin{aligned} 0 \in \partial_\epsilon f(x) &\Leftrightarrow \langle 0, y - x\rangle + f(x) \leqslant f(y) + \epsilon \,, \forall y \in R^n \\ &\Leftrightarrow f(x) \leqslant f(y) + \epsilon \,, \forall y \in R^n. \end{aligned}$$

□

Mệnh đề 2.14. *Cho f_i, i=1,...,m là các hàm lồi chính thường trên R^n. Giả sử $\cap ri(\text{dom}\, f_i) \neq \emptyset$. Khi đó*

$$\partial_\epsilon\big(\sum_{i=1}^m f_i(x)\big) \subseteq \sum_{i=1}^m \partial_\epsilon f_i(x) \qquad \forall x.$$

Đặc biệt

$$\partial_\epsilon\big(\sum_{i=1}^m f_i(x)\big) = \sum_{i=1}^m \partial_\epsilon f_i(x), \; \forall x \Leftrightarrow \epsilon = 0.$$

Chứng minh. Ta chứng minh cho $m = 2$. Với $m > 2$ dùng quy nạp.

Ta lấy $x^0 \in R^n$ và

$$x^* \in \partial_\epsilon(\sum_{i=1}^2 f_i(x^0)) = \partial_\epsilon(f_1(x^0) + f_2(x^0)) = \partial_\epsilon(f_1 + f_2)(x^0).$$

Theo định nghĩa của dưới vi phân xấp xỉ, ta có:

$$
\begin{aligned}
& x^* \in \partial_\epsilon(f_1+f_2)(x^0) \\
\Leftrightarrow & \langle x^*, x-x^0\rangle + (f_1+f_2)(x^0) \leqslant (f_1+f_2)(x) + \epsilon \quad \forall x \\
\Leftrightarrow & \langle x^*, x-x^0\rangle + f_1(x^0) + f_2(x^0) \leqslant f_1(x) + f_2(x) + \epsilon \quad \forall x \\
\Leftrightarrow & f_1(x) + f_2(x) - f_1(x^0) - f_2(x^0) - \langle x^*, x-x^0\rangle + \epsilon \geqslant 0 \quad \forall x \\
\Rightarrow & \text{hệ} \begin{cases} f_1(x) + f_2(y) - f_1(x^0) - f_2(x^0) - \langle x^*, x-x^0\rangle + \epsilon < 0 \\ x = y \end{cases} \\
& \text{không có nghiệm.}
\end{aligned} \tag{2.15}
$$

Lấy

$$
\begin{aligned}
D &= \operatorname{dom} f_1 \times \operatorname{dom} f_2, \\
A(x,y) &= x - y, \\
f(x,y) &= f_1(x) + f_2(y) - f_1(x^0) - f_2(x^0) - \langle x^*, x-x^0\rangle + \epsilon.
\end{aligned}
$$

Theo giả thiết f_1 liên tục tại một điểm $a \in \operatorname{dom} f_1 \cap \operatorname{dom} f_2$, nên tồn tại một lân cận U của gốc sao cho

$$U = (a+U) - a \subset \operatorname{dom} f_1 - \operatorname{dom} f_2 = A(D).$$

Vậy $0 \in intA(D)$.

Lúc này (2.15) có dạng:

$$\text{hệ} \quad \begin{cases} f(x,y) < 0 \\ A(x,y) = 0 \\ (x,y) \in D \end{cases} \quad \text{không có mghiệm}$$

Ap dụng mệnh đề 1.4 ta có:

$$\langle t, A(x,y) - 0\rangle + f(x,y) \geqslant 0 \quad \forall (x,y) \in D.$$

$$\Leftrightarrow \langle t, x-y\rangle + f_1(x) + f_2(y) - f_1(x^0) - f_2(x^0) + \langle x^*, x-x^0\rangle + \epsilon \geqslant 0,$$

$\forall x \in \operatorname{dom} f_1$, $\forall y \in \operatorname{dom} f_2$.

Đối với $x \notin \operatorname{dom} f_1$ và $y \notin \operatorname{dom} f_2$ thì bất đẳng thức trên là hiển nhiên. Vậy

$$\langle t, x-y\rangle + f_1(x) + f_2(y) - f_1(x^0) - f_2(x^0) + \langle x^*, x - x^0\rangle + \epsilon \geqslant 0 \qquad \forall x, y.$$

Lấy $x = x^0$ ta có :

$$\begin{aligned} &\langle t, x^0 - y\rangle + f_2(y) - f_2(x^0) + \epsilon \geqslant 0 \qquad \forall y \\ \Leftrightarrow\ &\langle t, y - x^0\rangle + f_2(x^0) \leqslant f_2(y) + \epsilon \qquad \forall y \\ \Leftrightarrow\ &t \in \partial_\epsilon f_2(x^0). \end{aligned}$$

Lấy $y = x^0$ ta có:

$$\begin{aligned} &\langle t, x - x^0\rangle + f_1(x) - f_1(x^0) - \langle x^*, x - x^0\rangle + \epsilon \geqslant 0 \qquad \forall x \\ \Leftrightarrow\ &\langle x^* - t, x - x^0\rangle + f_1(x^0) \leqslant f_1(x) + \epsilon \qquad \forall x \\ \Leftrightarrow\ &x^* - t \in \partial_\epsilon f_1(x^0). \end{aligned}$$

Do đó

$$x^* = (x^* - t) + t \subseteq \partial_\epsilon f_1(x^0) + \partial_\epsilon f_2(x^0).$$

Vậy

$$\partial_\epsilon (f_1(x^0) + f_2(x^0)) \subseteq \partial_\epsilon f_1(x^0) + \partial_\epsilon f_2(x^0).$$

□

Chương 3

Một số ứng dụng của dưới vi phân trong tối ưu hoá

Chương này trước hết giới thiệu một số khái niệm chung về cực tiểu, ϵ- cực tiểu của một hàm lồi. Tiếp theo trình bày điều kiện cần và đủ của nghiệm tối ưu của bài toán lồi với các ràng buộc khác nhau (Không ràng buộc, ràng buộc đẳng thức, ràng buộc bất đẳng thức). Cuối chương trình bày điều kiện cần và đủ của nghiệm tối ưu xấp xỉ của bài toán lồi với các ràng buộc khác nhau.

3.1 Các khái niệm

Định nghĩa 3.1. Cho $C \subseteq R^n$ khác rỗng và $f : R^n \longrightarrow R \cup \{+\infty\}$.

a) Điểm $x^* \in C$ được gọi là cực tiểu địa phương của f trên C nếu tồn tại một lân cận U của x^* sao cho

$$f(x^*) \leqslant f(x), \forall x \in U \cap C.$$

b) Điểm $x^* \in C$ được gọi là cực tiểu toàn cục (hay cực tiểu tuyệt đối) của f trên C nếu

$$f(x^*) \leqslant f(x), \forall x \in C.$$

c) Điểm $x \in C$ được gọi là điểm chấp nhận được của bài toán.

Định nghĩa 3.2. Cho $\epsilon > 0$. Một điểm $x_\epsilon \in C$ được gọi là điểm ϵ-cực tiểu toàn cục của f trên C nếu

$$f(x_\epsilon) \leqslant f(x) + \epsilon, \forall x \in C.$$

3.2 Bài toán lồi không có ràng buộc

Xét bài toán

$$(P1) \qquad \{\min h(x) \mid x \in R^n\}.$$

Trong đó h là một hàm lồi chính thường trên R^n.

Mệnh đề 3.1. *x^* là nghiệm của bài toán (P1) $\Leftrightarrow 0 \in \partial h(x^*)$.*

Chứng minh. Ta có:

$$\begin{aligned} x^* \text{là nghiệm của bài toán (P1)} &\Leftrightarrow x^* \text{là điểm cực tiểu của h trên } R^n \\ &\Leftrightarrow h(x^*) \leqslant h(x)\ , \forall x \in R^n \\ &\Leftrightarrow \langle 0, x - x^* \rangle + h(x^*) \leqslant h(x)\ , \forall x \in R^n \\ &\Leftrightarrow 0 \in \partial h(x^*). \end{aligned}$$

□

3.3 Bài toán lồi với ràng buộc đẳng thức

Xét bài toán

$$(P2) \qquad \{\min f(x) \mid x \in C\}.$$

Trong đó $C \subseteq R^n$ là một tập lồi khác rỗng và f là một hàm lồi trên C.

Mệnh đề 3.2. *Giả sử* $\mathrm{ri}(\mathrm{dom}\, f) \cap \mathrm{ri}\, C \neq \emptyset$.

$x^* \in C$ là nghiệm của bài toán (P2)$\Leftrightarrow 0 \in \partial f(x^*) + N_C(x^*)$, trong đó $N_C(x^*) := \{\omega \mid \langle \omega, x - x^* \rangle \leqslant 0\, , \forall x \in C\}$ là nón pháp tuyến ngoài của C tại x^*.

Chứng minh. Gọi $\delta_C(.)$ là hàm chỉ của tập C , tức là

$$\delta_C(x) := \begin{cases} 0 & \text{nếu} \quad x \in C, \\ +\infty & \text{nếu} \quad x \notin C. \end{cases}$$

Khi đó

$$\begin{aligned}
&x^* \text{ là nghiệm của bài toán (P2)}\\
\Leftrightarrow &x^* \text{ là điểm cực tiểu của f trên C}\\
\Leftrightarrow &x^* \text{ là điểm cực tiểu của } h(x) := f(x) + \delta_C(x) \text{ trên } R^n\\
\Leftrightarrow &0 \in \partial h(x^*) \qquad \text{(theo mệnh đề 3.1).}
\end{aligned}$$

Do $\mathrm{ri}(\mathrm{dom}\, f) \cap \mathrm{ri}\, C \neq \emptyset$, theo định lý Moreau-Rockafellar ta có:

$$\begin{aligned}
\partial h(x^*) &= \partial[f(x^*) + \delta_C(x^*)]\\
&= \partial f(x^*) + \partial \delta_C(x^*).
\end{aligned}$$

Vì $x^* \in C$ nên $\partial \delta_C(x^*) = N_C(x^*)$.
Vậy $\partial h(x^*) = \partial f(x^*) + N_C(x^*)$.
Suy ra $0 \in \partial f(x^*) + N_C(x^*)$. □

3.4 Bài toán lồi với ràng buộc bất đẳng thức

Xét bài toán tìm cực tiểu của một hàm lồi trên một tập lồi có dạng sau:

$$(OP) \qquad \{\min f(x) \mid g_i(x) \leqslant 0 \; (i = 1, ...m), x \in X\}.$$

Trong đó $X \subseteq R^n$ là một tập lồi đóng khác rỗng và f, g_i (i=1,..m) là các hàm lồi hữu hạn trên X. Ta sẽ luôn giả sủ rằng X có điểm trong.

Bài toán (OP) này được gọi là một quy hoạch lồi. Hàm f được gọi là hàm mục tiêu.

Các điều kiện $x \in X, g_i(x) \leqslant 0 \, (i = 1, ...m)$ được gọi là các ràng buộc.

Tập $D := \{x \in X \mid g_i(x) \leqslant 0 \, i = 1, ...m\}$ được gọi là miền chấp nhận được.

Một điểm $x \in D$ được gọi là điểm chấp nhận được của bài toán (OP).

Do X là tập lồi, các hàm g_i (i=1,..,m) lồi trên X nên D là một tập lồi. Điểm cực tiểu của f trên D cũng được gọi là nghiệm tối ưu của bài toán (OP).

Ta xây dựng hàm sau, được gọi là hàm Lagrange, cho bài toán (OP):

$$L(x, \lambda) := \lambda_0 f(x) + \sum_{i=1}^{m} \lambda_i g_i(x),$$

với $\lambda = (\lambda_0, ..., \lambda_m)$

Dựa vào hàm Lagrange ta có kết qủa sau:

Định lý 3.1. *(Karush- Kuhn- Tucker)*

Giả sử $\mathrm{ri}(\mathrm{dom}\, f) \cap \mathrm{ri}(\mathrm{dom}\, g_i) \cap \mathrm{ri}\, X \neq \emptyset$

a) Nếu x^* là nghiệm của bài toán (OP) thì tồn tại $\lambda_i^* \geqslant 0$ (i=0,...,m) không đồng thời bằng 0 sao cho:

1)

$$L(x^*, \lambda^*) = \min_{x \in X} L(x, \lambda^*) \qquad \textit{(điều kiện đạo hàm triệt tiêu)}$$

$$(\Leftrightarrow 0 \in \lambda_0^* \partial f(x^*) + \sum_{i=1}^{m} \lambda_i^* \partial g_i(x^*) + N_X(x^*)\,).$$

Trong đó $N_X(x^*)$ là nón pháp tuyến ngoài của X tại x^* .

2)

$$\lambda_i^* g_i(x^*) = 0\, (i = 1, ..., m) \qquad \textit{(điều kiện độ lệch bù).}$$

Hơn nữa nếu điều kiện Slater sau thoả mãn:

$$\exists x^0 \in X : g_i(x^0) < 0\, (i = 1, ...m) \textit{ thì } \lambda_0^* > 0.$$

b) Nếu hai điều kiện đạo hàm triệt tiêu và độ lệch bù ở trên được thoả mãn với $\lambda_0^* > 0$ thì điểm chấp nhận được x^* là nghiệm tối ưu của bài toán (OP)

Chứng minh. a) Giả sử x^* là nghiệm của bài toán (OP). Đặt

$$\begin{aligned} C :=& \{(\lambda_0, \lambda_1, ..., \lambda_m) \in R^{m+1} | \exists x \in X : \\ & f(x) - f(x^*) < \lambda_0,\ g_i(x) \leqslant \lambda_i, i = 1, ..., m\}. \end{aligned}$$

Do $X \neq \emptyset$ lồi, f, g_i lồi trên X, nên C là một tập lồi .

Ta có $C \neq \emptyset$. Thật vậy:

+ Lấy $(\lambda_0, ..., \lambda_m) \in intR_+^{m+1}$. Khi đó $\lambda_i > 0$ $(i = 1, ..., m)$.

+ Với $x = x^*$, ta có

$$f(x^*) - f(x^*) = 0 < \lambda_0$$
$$g_i(x^*) \leqslant 0 < \lambda_i \ (i = 1, ..., m).$$

$\Rightarrow (\lambda_0, ..., \lambda_m) \in C.$

$\Rightarrow intR_+^{m+1} \subset C.$

$\Rightarrow C \neq \emptyset$ trong R^{m+1}.

Hơn nữa $0 \notin C$. Thật vậy, nếu $0 \in C$ thì

$$\exists x \in X : f(x) - f(x^*) < 0$$
$$g_i(x) \leqslant 0 \ (i = 1, ..., m).$$

Do đó x^* không là nghiệm của bài toán (OP)$\Rightarrow$ mâu thuẫn. Vậy $0 \notin C$.

Theo định lý tách thứ 1, có thể tách các tập C và $\{0\}$, tức là $\exists \lambda_i^*$ (i=0,...m) không đồng thời bằng 0 sao cho

$$\sum_{i=0}^{m} \lambda_i^* \lambda_i \geqslant 0 \qquad \forall (\lambda_0, ..., \lambda_m) \in C. \tag{3.1}$$

Do $intR_+^{m+1} \subset C$, ta suy ra $\lambda_i^* \geqslant 0$.

Với $\epsilon > 0$ và $x \in X$, lấy

$$\lambda_0 = f(x) - f(x^*) + \epsilon$$
$$\lambda_i = g_i(x) \ (i = 1, ..., m).$$

Thay vào (3.1) ta có

$$\lambda_0^*[f(x) - f(x^*) + \epsilon] + \sum_{i=1}^{m} \lambda_i^* g_i(x) \geqslant 0 \qquad \forall x \in X.$$

Cho $\epsilon \to 0$ ta được

$$\lambda_0^* f(x) + \sum_{i=1}^{m} \lambda_i^* g_i(x) \geqslant \lambda_0^* f(x^*) \qquad \forall x \in X. \tag{3.2}$$

Do x^* là điểm chấp nhận được nên ta có $g_i(x^*) \leqslant 0$ $(i = 1, ..., m)$. Vậy

$$\lambda_0^* f(x^*) \geqslant \lambda_0^* f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*). \tag{3.3}$$

Từ (3.2) và (3.3) ta có

$$\lambda_0^* f(x) + \sum_{i=1}^{m} \lambda_i^* g_i(x) \geqslant \lambda_0^* f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*) \qquad \forall x \in X$$

$\Leftrightarrow L(x, \lambda^*) \geqslant L(x^*, \lambda^*) \qquad \forall x \in X$

$\Leftrightarrow L(x^*, \lambda^*) = \min_{x \in X} L(x, \lambda^*)$ (điều kiện đạo hàm triệt tiêu).

Ta chú ý rằng x^* là nghiệm của bài toán $\{\min L(x, \lambda^*), x \in X\}$ khi và chỉ khi x^* là điểm cực tiểu của hàm $L(x, \lambda^*)$ trên X

$\Leftrightarrow x^*$ là điểm cực tiểu của hàm $L_1(x, \lambda^*) := L(x, \lambda^*) + \delta_X(x)$ trên R^n.

$\Leftrightarrow 0 \in \partial L_1(x^*, \lambda^*)$ (theo mệnh đề 3.1).

Do $\mathrm{ri}(\mathrm{dom}\, f) \cap \mathrm{ri}(\mathrm{dom}\, g_i) \cap \mathrm{ri}\, X \neq \emptyset$ và f, g_i (i:=1,...,m) là các hàm lồi, hữu hạn trên X nên theo định lý Moreau-Rockafellar ta có

$$\begin{aligned}
\partial L_1(x^*, \lambda^*) &= \partial[L(x^*, \lambda^*) + \delta_X(x^*)] \\
&= \partial[\lambda_0^* f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*)] + \partial \delta_X(x^*) \\
&= \lambda_0^* \partial f(x^*) + \sum_{i=1}^{m} \lambda_i^* \partial g_i(x^*) + N_X(x^*) \\
&\qquad \text{(vì } \partial \delta_X(x^*) = N_X(x^*) \text{).}
\end{aligned}$$

Vậy

$$0 \in \lambda_0^* \partial f(x^*) + \sum_{i=1}^{m} \lambda_i^* \partial g_i(x^*) + N_X(x^*).$$

Do x^* là điểm chấp nhận được nên $g_i(x^*) \leqslant 0\, (i = 1, ..., m)$. Nếu $\exists i \in \{1, ..., m\} : g_i(x^*) = \xi < 0$ thì

$$\forall \epsilon > 0\, ,\, f(x^*) - f(x^*) = 0 < \epsilon$$

$$g_j(x^*) \leqslant 0 < \epsilon (j = 1, ..., i-1, i+1, ..., m).$$

$\Rightarrow (\epsilon, ..., \epsilon, \xi, \epsilon, ..., \epsilon) \in C$ (ξ ở vị trí thứ i).

$\Rightarrow \lambda_i^* \xi \geqslant 0$ (thay vào (3.1) và cho $\epsilon \to 0$)

$\Rightarrow \lambda_i^* \leqslant 0.$

Theo chứng minh trên ta có $\lambda_i^* \geqslant 0$. Vậy $\lambda_i^* = 0$.

Như vậy là, nếu $g_i(x^*) < 0$ thì $\lambda_i^* = 0$.

Do đó $\lambda_i^* g_i(x^*) = 0\ (i = 1, .., m)$ (điều kiện độ lệch bù).

Giả sử điều kiện Slater được thoả mãn: $\exists x^0 \in X : g_i(x^0) < 0$.

Khi đó nếu $\lambda_0^* = 0$ thì do điều kiện đạo hàm triệt tiêu và độ lệch bù ta có

$$0 = \lambda_0^* f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*) \leqslant \lambda_0^* f(x) + \sum_{i=1}^{m} \lambda_i^* g_i(x)\ , \forall x \in X.$$

Do $\lambda_0^* = 0$ nên phải có ít nhất một $\lambda_i^* > 0$.

Thay x^0 vào bất đẳng thức trên, sẽ được

$$0 = \lambda_0^* f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*) \leqslant \lambda_0^* f(x^0) + \sum_{i=1}^{m} \lambda_i^* g_i(x^0) < 0.$$

Suy ra mâu thuẫn. Vậy $\lambda_0^* \neq 0$ tức là $\lambda_0^* > 0$.

b) Giả sử x^* là điểm chấp nhận được thoả mãn hai điều kiện đạo hàm triệt tiêu và độ lệch bù ở trên với $\lambda_0^* > 0, \lambda_i^* \geqslant 0\ (i = 1, ..., m)$.

Do $\lambda_0^* > 0$, nên bằng cách chia cho λ_0^*, ta có thể coi hàm Lagrange là

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x).$$

Từ điều kiện đạo hàm triệt tiêu và độ lệch bù, ta có:

$$f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*) \leqslant f(x) + \sum_{i=1}^{m} \lambda_i^* g_i(x) \qquad \forall x \in X$$
$$\lambda_i^* g_i(x^*) = 0\ (i = 1, ..., m).$$

Suy ra

$$f(x^*) \leqslant f(x) + \sum_{i=1}^{m} \lambda_i^* g_i(x)\ , \forall x \in X. \tag{3.4}$$

Với mọi x là điểm chấp nhận được, tức là:

$$x \in X : g_i(x) < 0\,, i = 1, ..., m,$$

ta có

$$f(x) + \sum_{i=1}^{m} \lambda_i^* g_i(x) \leqslant f(x). \tag{3.5}$$

Từ (3.4) và (3.5) suy ra $f(x^*) \leqslant f(x)\,, \forall x \in X$ .Chứng tỏ x^* là nghiệm tối ưu của bài toán (OP). □

Ví dụ 3.1. Ap dụng định lý cho bài toán sau:

$$\{\min f(x) \mid g_i(x) \leqslant 0\ (i = 1, 2)\,, x \in X\}, \tag{OP}$$

trong đó $f(x) = x^2, g_1(x) = x^2 - x, g_2(x) = -x, X = [-\frac{1}{2}, \frac{1}{2}]$.

Giải:

Ta có miền chấp nhận được

$$D = \{x \in X \mid g_i(x) \leqslant 0\ (i = 1, 2)\} = [0, \frac{1}{2}].$$

Giả sử tồn tại $\lambda_i^* \geqslant 0\ (i = 0, ..., 2)$ không đồng thời bằng 0 sao cho:

1) $L(x^*, \lambda^*) = \min_{x \in X} L(x, \lambda^*)$

$(\Leftrightarrow 0 \in \lambda_0^* \partial f(x^*) + \sum_{i=1}^{2} \lambda_i^* \partial g_i(x^*) + N_X(x^*)\,)$.

2) $\lambda_i^* g_i(x^*) = 0,\ i = 1, 2.$

3) $\lambda_0^* > 0.$

Từ định lí 3.1, suy ra x^* là nghiệm tối ưu của bài toán (OP)

$$\begin{aligned}
&\Leftrightarrow f(x^*) \leqslant f(x),\ \forall x \in D\\
&\Leftrightarrow {x^*}^2 \leqslant x^2,\ \forall x \in D\\
&\Leftrightarrow {x^*}^2 \leqslant 0\\
&\Leftrightarrow x^* = 0.
\end{aligned}$$

Ngược lại, nếu $x^* = 0$ là nghiệm của bài toán (OP) thì từ định lí 3.1, suy ra tồn tại $\lambda_i^* \geqslant 0\ (i = 0, ..., 2)$ không đồng thời bằng 0 sao cho :

1) $L(x^*, \lambda^*) = \min_{x\in X} L(x, \lambda^*)$

($\Leftrightarrow 0 \in \lambda_0^* \partial f(x^*) + \sum_{i=1}^2 \lambda_i^* \partial g_i(x^*) + N_X(x^*)$).

2) $\lambda_i^* g_i(x^*) = 0,\ i = 1, 2.$

Ta có

$$\begin{aligned} & L(x^*, \lambda^*) = \min_{x\in X} L(x, \lambda^*) \\ \Leftrightarrow & L(x, \lambda^*) \geqslant L(x^*, \lambda^*),\ \forall x \in X \\ \Leftrightarrow & \lambda_0^* f(x) + \sum_{i=1}^{2} \lambda_i^* g_i(x) \geqslant \lambda_0^* f(x^*) + \sum_{i=1}^{2} \lambda_i^* g_i(x^*),\ \forall x \in X \\ \Leftrightarrow & \lambda_0^* x^2 + \lambda_1^*(x^2 - x) - \lambda_2^* x \geqslant 0,\ \forall x \in X. \end{aligned}$$

Ta có $\lambda_i^* g_i(x^*) = 0,\ i = 1, 2 \Leftrightarrow \lambda_i^*.0 = 0,\ i = 1, 2 \Leftrightarrow \lambda_i^* \geqslant 0,\ i = 1, 2.$

Do $\lambda_i^* \geqslant 0$ ($i = 0, ..., 2$) không đồng thời bằng 0 nên:

+ Chọn $\lambda_1^* = \lambda_2^* = 0.$

Ta có

$$\begin{aligned} & \lambda_0^* x^2 + \lambda_1^*(x^2 - x) - \lambda_2^* x \geqslant 0,\ \forall x \in X \\ \Leftrightarrow & \lambda_0^* x^2 \geqslant 0,\ \forall x \in X \\ \Leftrightarrow & \lambda_0^* > 0 \\ \Rightarrow & \text{Chọn} \lambda_0^* = 1. \end{aligned}$$

+ Chọn $\lambda_1^* = \lambda_2^* = 1.$

Ta có

$$\begin{aligned} & \lambda_0^* x^2 + \lambda_1^*(x^2 - x) - \lambda_2^* x \geqslant 0,\ \forall x \in X \\ \Leftrightarrow & (\lambda_0^* + 1)x^2 - 2x \geqslant 0,\ \forall x \in X \\ \Rightarrow & \text{Không tồn tại } \lambda_0^*. \end{aligned}$$

+ Chọn $\lambda_1^* = 0, \lambda_2^* = 1.$

Ta có

$$\begin{aligned} & \lambda_0^* x^2 + \lambda_1^*(x^2 - x) - \lambda_2^* x \geqslant 0,\ \forall x \in X \\ \Leftrightarrow & \lambda_0^* x^2 - x \geqslant 0,\ \forall x \in X \\ \Rightarrow & \text{Không tồn tại } \lambda_0^*. \end{aligned}$$

+ Chọn $\lambda_1^* = 1, \lambda_2^* = 0.$

Ta có

$$\begin{aligned} &\lambda_0^* x^2 + \lambda_1^*(x^2 - x) - \lambda_2^* x \geqslant 0,\ \forall x \in X \\ \Leftrightarrow &(\lambda_0^* + 1)x^2 - x \geqslant 0,\ \forall x \in X \\ \Rightarrow &\text{Không tồn tại } \lambda_0^*. \end{aligned}$$

Vậy $x^* = 0$ là nghiệm tối ưu của bài toán (OP) và $\lambda_0^* = 1, \lambda_1^* = \lambda_2^* = 0$ là các nhân tử Lagrang tương ứng.

Chú ý 3.1. Trong nhiều trường hợp bài toán (P1), (P2) và (OP) có thể không có lời giải tối ưu chính xác. Hơn nữa trong thực tế thường người ta không tính được lời giải tối ưu (chính xác), mà chỉ tính được lời giải xấp xỉ. Khi đó ta dùng khái niệm lời giải tối ưu xấp xỉ hay còn gọi là ϵ- tối ưu.

Mệnh đề 3.3. *x_ϵ là ϵ -nghiệm của bài toán (P1)* $\Leftrightarrow 0 \in \partial_\epsilon h(x_\epsilon)$

Chứng minh. Ta có:

$$\begin{aligned} &x_\epsilon \text{ là } \epsilon - \text{nghiệm của bài toán (P1)} \\ \Leftrightarrow &x_\epsilon \text{ là điểm } \epsilon - \text{cực tiểu của h trên } R^n \\ \Leftrightarrow &h(x_\epsilon) \leqslant h(x) + \epsilon\,, \forall x \in R^n \qquad \text{(theo định nghĩa3.2)} \\ \Leftrightarrow &0 \in \partial_\epsilon h(x_\epsilon) \qquad \text{(theo định lý2.1).} \end{aligned}$$

□

Mệnh đề 3.4. *Giả sử* $\mathrm{ri}(\mathrm{dom}\, f) \cap \mathrm{ri}\, C \neq \emptyset$. *Khi đó*

$x_\epsilon \in C$ là ϵ -nghiệm của bài toán (P2) $\Longrightarrow 0 \in \partial_\epsilon f(x_\epsilon) + N_{C,\epsilon}(x_\epsilon)$,

trong đó $N_{C,\epsilon}(x_\epsilon) := \{\omega \mid \langle \omega, x - x_\epsilon \rangle \leqslant \epsilon\,, \forall x \in C\}$ là ϵ -nón pháp tuyến ngoài của C tại x_ϵ.

Chứng minh. Gọi $\delta_C(.)$ là hàm chỉ của tập C , tức là

$$\delta_C(x) := \begin{cases} 0 & \text{nếu } x \in C, \\ +\infty & \text{nếu } x \notin C. \end{cases}$$

Khi đó

$$\begin{aligned}
&x_\epsilon \text{ là } \epsilon - \text{nghiệm của bài toán (P2)}\\
\Leftrightarrow &x_\epsilon \text{ là điểm } \epsilon - \text{cực tiểu của f trên C}\\
\Leftrightarrow &x_\epsilon \text{ là điểm } \epsilon - \text{cực tiểu của } h(x) := f(x) + \delta_C(x) \text{ trên } R^n\\
\Leftrightarrow &0 \in \partial_\epsilon h(x_\epsilon) \qquad \text{(theo mệnh đề 3.3).}
\end{aligned}$$

Do $\mathrm{ri}(\mathrm{dom}\, f) \cap \mathrm{ri}\, C \neq \emptyset$, ta có:

$$\begin{aligned}
\partial_\epsilon h(x_\epsilon) &= \partial_\epsilon [f(x_\epsilon) + \delta_C(x_\epsilon)]\\
&\subseteq \partial_\epsilon f(x_\epsilon) + \partial_\epsilon \delta_C(x_\epsilon) \quad \text{(theo mệnh đề 2.14)}\\
&= \partial_\epsilon f(x_\epsilon) + N_{C,\epsilon}(x_\epsilon)
\end{aligned}$$

(Vì $x_\epsilon \in C$ nên theo định nghĩa 2.9 $\partial_\epsilon \delta_C(x_\epsilon) = N_{C,\epsilon}(x_\epsilon)$).

Vậy $0 \in \partial_\epsilon f(x_\epsilon) + N_{C,\epsilon}(x_\epsilon)$. $\square$

Kết luận

Như vậy, luận văn này đã trình bày một cách hệ thống các khái niệm, tính chất cơ bản của tập lồi và hàm lồi. Sau đó lại đề cập về đạo hàm theo phương, dưới vi phân, dưới vi phân xấp xỉ và chứng minh một cách cụ thể một số tính chất của chúng. Cuối cùng luận văn trình bày các điều kiện cực trị cho các bài toán quy hoạch lồi với các rằng buộc khác nhau.

Tài liệu tham khảo

[1] Lê Dũng Mưu và Nguyễn Văn Hiền (2003), *Nhập môn giải tích lồi ứng dụng*, Giáo trình.

[2] . Tạ Quang Sơn (2008), *Some Qualitative Problems In Optimization*, Luận án tiến sĩ.

[3] T. Rockafellar (1970), *Convex Analysis*, Princeton University Press, Princeton, New Jersey.

[4] J. Hiriart-Urruty and C. Lemarechal, *Convex Analysis and Minimization Algorithms*.