

# Mục lục

|                                                         |           |
|---------------------------------------------------------|-----------|
| <b>Mở đầu</b>                                           | <b>2</b>  |
| <b>Chương 1: Dưới vi phân</b>                           | <b>5</b>  |
| 1.1. Định nghĩa và kí hiệu . . . . .                    | 5         |
| 1.2. Một số tính chất cơ bản của dưới vi phân . . . . . | 6         |
| 1.3. Phép toán về dưới vi phân . . . . .                | 12        |
| <b>Chương 2: Điều kiện tồn tại nghiệm tối ưu</b>        | <b>18</b> |
| 2.1. Sự tồn tại nghiệm tối ưu . . . . .                 | 18        |
| 2.2. Các bài toán tối ưu . . . . .                      | 22        |
| 2.3. Bài toán tối ưu không ràng buộc . . . . .          | 23        |
| 2.3.1 Điều kiện tối ưu cấp 1 . . . . .                  | 24        |
| 2.3.2 Điều kiện tối ưu cấp 2 . . . . .                  | 25        |
| 2.4. Bài toán tối ưu có ràng buộc . . . . .             | 40        |
| 2.4.1 Điều kiện tối ưu cấp 1 . . . . .                  | 44        |
| 2.4.2 Điều kiện tối ưu cấp 2 . . . . .                  | 47        |
| <b>Kết luận</b>                                         | <b>62</b> |
| <b>Tài liệu tham khảo</b>                               | <b>63</b> |

# Mở đầu

Trong tự nhiên sự vận hành và phát triển của vạn vật đều có thể qui được về hai vấn đề cơ bản sau:

1) Tồn tại hay không tồn tại? Theo ngôn ngữ toán học: có tồn tại hay không nghiệm của phương trình

$$f(x) = 0, x \in D \quad (1)$$

2) Tồn tại như thế nào? Theo ngôn ngữ toán học: Tìm nghiệm tối ưu của bài toán

$$\min_{x \in D} f(x) \quad (2)$$

Chính vì vậy toán học nói chung luôn là công cụ hữu hiệu giải quyết các bài toán nảy sinh từ thực tế sinh động. Lý thuyết tối ưu nói riêng trong thời đại ngày nay đang được sử dụng một cách khá triệt để trong mọi lĩnh vực của cuộc sống.

Hai bài toán trên cũng có liên quan với nhau. Đôi khi để giải quyết bài toán (1) ta chỉ cần giải bài toán (2) và ngược lại. Bài toán (2) đóng vai trò chính trong lý thuyết tối ưu. Để nghiên cứu, chứng minh sự tồn tại nghiệm và tìm phương pháp giải ra nghiệm của bài toán này, người ta thường phân loại theo cấu trúc của tập hợp  $D$  và tính chất của hàm số  $f$ . Nếu  $D$  là tập mở và  $f$  là hàm số khả vi thì (2) được gọi là bài toán tối ưu trơn. Đối với bài toán này, ta đã có dịp làm quen trong chương trình phổ thông. Sự tồn tại nghiệm của nó được qui về xét các điều kiện của các đạo hàm cấp 1, 2. Nếu  $f$  là hàm số không có đạo hàm, bài toán (2) được gọi là bài toán tối ưu không trơn. Mục đích của luận văn này

là trình bày một số cách tiếp cận để nghiên cứu các điều kiện cần và đủ cho việc tồn tại nghiệm của bài toán (2). Như chúng ta đã biết trong giáo trình giải tích cổ điển, ngay cả trong  $\mathbb{R}^1$  nhiều hàm  $f$  lồi không khả vi tại điểm  $x$  nào đó thuộc  $(a; b)$ , vì vậy rất khó xấp xỉ các hàm số này tại lân cận của  $x$  bởi một hàm tuyến tính. Khi đó ta không có được các điều kiện cần và đủ tối ưu cho bài toán tối ưu như đối với các hàm khả vi. Những năm 60 của thế kỷ XX, Rockafellar đã xây dựng lý thuyết dưới vi phân cho lớp hàm lồi và ý tưởng cơ bản của lý thuyết này là xấp xỉ hàm lồi tại điểm cho trước bằng cả một tập hợp có tính chất khá đẹp được gọi là tập dưới vi phân thay vì chỉ có một hàm tuyến tính như trong trường hợp khả vi. Các tập dưới vi phân chứa các thông tin về các điều kiện cần và đủ tối ưu cho các bài toán tối ưu liên quan đến các hàm này. Đây là một vấn đề khó nhưng có nhiều ứng dụng trong thực tế. Chính vì lẽ đó mà tác giả đã chọn đề tài: "**Dưới vi phân của hàm lồi và ứng dụng trong tối ưu hoá không trơn**".

Luận văn được chia làm 2 chương.

Chương I: Dưới vi phân. Trong chương I, tác giả trình bày các kiến thức cơ bản về dưới vi phân như: định nghĩa, các tính chất và các phép toán về dưới vi phân.

Chương II: Điều kiện tồn tại nghiệm tối ưu. Trong chương II, tác giả trình bày một cách chi tiết các điều kiện tối ưu cấp 1 và cấp 2 đối với hai loại bài toán tối ưu không trơn là bài toán tối ưu không ràng buộc và bài toán tối ưu có ràng buộc và có sự so sánh với bài toán tối ưu trơn.

Bản luận văn được hoàn thành dưới sự hướng dẫn tận tình của GS.TS Trần Vũ Thiệu. Tác giả hi vọng rằng một phần kiến thức nhỏ

trong luận văn sẽ là tài liệu tham khảo cho các bạn sinh viên đại học, cao đẳng, những người làm toán quan tâm và yêu thích đề tài này.

Mặc dù tác giả đã cố gắng hết sức nhưng kết quả đạt được trong luận văn còn rất khiêm tốn và không tránh khỏi những thiếu sót, tác giả mong nhận được nhiều ý kiến đóng góp quý báu của các thầy cô và đồng nghiệp.

Hà Nội, tháng 11 năm 2009

# Chương 1

## Dưới vi phân

### 1.1 Định nghĩa và kí hiệu

**Định nghĩa 1.1.** Cho  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  là một hàm lồi. Một véctơ  $g \in \mathbb{R}^n$  là dưới gradient của  $f$  tại  $x \in \mathbb{R}^n$  nếu

$$f(x + \delta) \geq f(x) + \delta^T g, \quad \forall x + \delta \in \mathbb{R}^n. \quad (1.1)$$

**Định nghĩa 1.2.** Tập tất cả dưới gradient của  $f$  tại  $x$  được gọi là dưới vi phân của hàm  $f$  tại  $x$ , kí hiệu là  $\partial f(x)$ , tức là:

$$\partial f(x) = \{ g : f(x + \delta) \geq f(x) + \delta^T g, \quad \forall x + \delta \in \mathbb{R}^n \} \quad (1.2)$$

Kí hiệu:

$$\partial f^{(k)} = \partial f(x^{(k)}), \quad f^{(k)} = f(x^{(k)}).$$

**Ví dụ 1.1.** Cho  $f(x) = |x|$ . Khi đó

$$\partial f(0) = [-1, 1],$$

$$\partial f(x) = \begin{cases} \{1\} & \text{nếu } x > 0 \\ \{-1\} & \text{nếu } x < 0. \end{cases}$$

**Ví dụ 1.2.** Cho  $f(x) = e^x - 1$ . Khi đó

$$\partial f(0) = [0, 1],$$

$$\partial f(x) = \begin{cases} \{e^x\} & \text{nếu } x > 0 \\ \{0\} & \text{nếu } x < 0. \end{cases}$$

**Định nghĩa 1.3.** Hàm  $f$  được gọi là khả dưới vi phân tại  $x$  nếu tập  $\partial f(x) \neq \emptyset$ .

## 1.2 Một số tính chất cơ bản của dưới vi phân

**Bổ đề 1.1.** Dưới vi phân  $\partial f(x)$  là một tập đóng, tức là: nếu ta có dãy  $x^{(k)} \rightarrow x', g^{(k)} \rightarrow g', g^{(k)} \in \partial f(x^{(k)})$  thì  $g' \in \partial f(x')$ .

*Chứng minh.* Lấy  $y \in K$ , vì  $g^{(k)} \in \partial f(x^{(k)})$  nên với mọi  $k$  ta có

$$f(y) \geq f(x^{(k)}) + (y - x^{(k)})^T g^{(k)}. \quad (1.3)$$

Trong (1.3) cho  $k \rightarrow \infty$  ta được

$$f(y) \geq f(x') + (y - x')^T g', \quad \forall y \in K$$

suy ra  $g' \in \partial f(x')$ . □

**Bổ đề 1.2.**  $\partial f(x)$  là tập bị chặn với mọi  $x \in B \subset \text{Int}(K)$  trong đó  $K \subset \mathbb{R}^n$  và  $B$  là tập compact.

*Chứng minh.* Giả sử ngược lại, khi đó tồn tại dãy  $g^{(k)} \in \partial f(x^{(k)})$  và dãy  $x^{(k)} \in B$  sao cho  $\|g^{(k)}\|_2 \rightarrow \infty$ . Do tính compact nên tồn tại  $x^{(k)} \rightarrow x'$ .

Định nghĩa

$$\delta^{(k)} = \frac{g^{(k)}}{\|g^{(k)}\|_2^2}.$$

Khi đó  $x^{(k)} + \delta^{(k)} \in K$  với  $k$  đủ lớn và theo (1.1) ta có

$$f(x^{(k)} + \delta^{(k)}) \geq f^{(k)} + g^{(k)T} \delta^{(k)} = f^{(k)} + 1.$$

Nhưng qua giới hạn thì

$$f^{(k)} \rightarrow f', \quad \delta^{(k)} \rightarrow 0.$$

Vì vậy  $f(x^{(k)} + \delta^{(k)}) \rightarrow f'$ . Mâu thuẫn. □

### Nhận xét 1.1.

- i) Từ hai bổ đề trên suy ra  $\partial f(x)$  là một tập compact.
- ii) Nếu  $f$  khả vi tại  $x$  thì

$$f(x + \delta) = f(x) + \delta^T \nabla f(x) + o(\|\delta\|)$$

mà

$$f(x + \delta) \geq f(x) + \delta^T g$$

nên

$$\delta^T (g - \nabla f(x)) \leq o(\|\delta\|).$$

Chọn  $\delta = \theta(g - \nabla f(x))$ ,  $\theta \downarrow 0$  sao cho  $g = \nabla f$ . Từ đây ta có  $\partial f(x)$  là vectơ  $\nabla f(x)$ .

### Bổ đề 1.3. Xét hàm đa trị

$$\partial f : K \rightarrow 2^K \quad (K \subset \mathbb{R}^n)$$

$$x \longmapsto \partial f(x)$$

Khi đó hàm đa trị  $\partial f$  đơn điệu, tức là với mọi  $x_1, x_2 \in K$  luôn tồn tại  $g_1 \in \partial f(x_1)$ ,  $g_2 \in \partial f(x_2)$  sao cho

$$(g_2 - g_1)^T (x_2 - x_1) \geq 0.$$

*Chứng minh.* Lấy  $x_1, x_2 \in K$ ,  $g_1 \in \partial f(x_1)$ ,  $g_2 \in \partial f(x_2)$ . Theo định nghĩa của dưới vi phân ta có

$$f(x_2) \geq f(x_1) + g_1^T(x_2 - x_1)$$

$$f(x_1) \geq f(x_2) + g_2^T(x_1 - x_2).$$

Cộng hai bất đẳng thức trên ta được

$$(g_2 - g_1)^T(x_2 - x_1) \geq 0. \quad \square$$

Xét lớp các hàm lồi đa diện  $h : \mathbb{R}^m \rightarrow \mathbb{R}^1$ ,  $h(c)$  được định nghĩa bởi

$$h(c) = \max_i c^T h_i + b_i \quad (1.4)$$

trong đó  $h_i$  là các cột của ma trận hữu hạn  $H$  cho trước. Định nghĩa

$$\mathcal{A} = \mathcal{A}(c) = \{ i : c^T h_i + b_i = h(c) \} \quad (1.5)$$

là tập các siêu phẳng tựa tại  $c$  và do đó đạt giá trị lớn nhất. Khi đó dễ dàng nhận thấy các siêu phẳng này xác định dưới vi phân  $\partial h(c)$ . Điều này được nêu trong bổ đề sau:

**Bổ đề 1.4.**  $\partial h(c) = \text{conv}_{i \in \mathcal{A}} h_i$

*Chứng minh.* Ta có

$$\partial h(c) = \{ \lambda : h(c + \delta) \geq h(c) + \delta^T \lambda, \forall \delta \} \quad (1.6)$$

Lấy  $\lambda \in \text{conv}_{i \in \mathcal{A}} h_i$ , ta có  $\lambda = \sum_{i \in \mathcal{A}} h_i \mu_i$  với  $\mu_i \geq 0$ ,  $\sum_{i \in \mathcal{A}} \mu_i = 1$ . Khi đó với mọi  $\delta$  ta có

$$\begin{aligned} h(c) + \delta^T \lambda &= \max_{i \in \mathcal{A}} (c^T h_i + b_i) + \sum_{i \in \mathcal{A}} \delta^T h_i \mu_i \\ &\leq \max_{i \in \mathcal{A}} (c^T h_i + b_i) + \max_{i \in \mathcal{A}} \delta^T h_i \\ &= \max_{i \in \mathcal{A}} (c + \delta)^T h_i + b_i \leq h(c + \delta). \end{aligned}$$

Do đó  $\lambda \in \partial(h(c))$  nên  $\text{conv}_{i \in \mathcal{A}} h_i \subset \partial h(c)$ .

Ngược lại giả sử  $\lambda \in \partial h(c)$ ,  $\lambda \notin \text{conv}_{i \in \mathcal{A}} h_i$ . Khi đó theo Bổ đề 1.5 ở phần dưới sẽ tồn tại  $s \neq 0$ ,  $s^T \lambda > s^T \mu$ ,  $\forall \mu \in \text{conv}_{i \in \mathcal{A}} h_i$ .

Lấy  $\delta = \alpha s$  và từ  $h_i \in \text{conv}_{i \in \mathcal{A}} h_i$  ta có

$$\begin{aligned} h(c) + \delta^T \lambda &= \max_i (c^T h_i + b_i) + \alpha s^T \lambda \\ &> c^T h_i + b_i + \alpha s^T h_i, \forall i \in \mathcal{A} \\ &= \max_{i \in \mathcal{A}} (c + \alpha s)^T h_i + b_i \\ &\geq \max_i (c + \alpha s)^T h_i + b_i = h(c + \delta) \end{aligned}$$

nên

$$h(c) + \delta^T \lambda > h(c + \delta).$$

Với  $\alpha$  đủ nhỏ, khi đó max đạt được trên một tập con của  $\mathcal{A}$ , mâu thuẫn với  $\lambda \in \partial h(c)$ .

Do đó với  $\lambda \in \text{conv}_{i \in \mathcal{A}} h_i$  thì  $\partial h(c) \subset \text{conv}_{i \in \mathcal{A}} h_i$ .

Vậy  $h(c) = \text{conv}_{i \in \mathcal{A}} h_i$ . □

### **Bổ đề 1.5. (Bổ đề về siêu phẳng tách các tập lồi)**

Nếu  $K$  là một tập lồi, đóng,  $\lambda \notin K$ . Khi đó tồn tại một siêu phẳng tách  $\lambda$  và  $K$ .

*Chứng minh.* Lấy  $x_0 \in K$ . Khi đó tập

$$\{ x : \|x - \lambda\|_2 \leq \|x_0 - \lambda\|_2 \}$$

là một tập bị chặn. Do đó tồn tại điểm cực tiểu  $\bar{x}$  đối với bài toán

$$\min \|x - \lambda\|_2, x \in K.$$

Khi đó với bất kì  $x \in K$  ta có

$$\|(1 - \theta)\bar{x} + \theta x - \lambda\|_2^2 \geq \|\bar{x} - \lambda\|_2^2.$$

Cho  $\theta \rightarrow 0$  ta được

$$(x - \bar{x})^T(\lambda - \bar{x}) \leq 0, \forall x \in K.$$

Từ đó véctor  $s = \lambda - \bar{x} \neq 0$  và thoả mãn đồng thời

$$s^T(\lambda - \bar{x}) > 0, s^T(x - \bar{x}) \leq 0, \forall x \in K$$

nên

$$s^T \lambda > s^T \bar{x} \geq s^T x$$

do đó

$$s^T \lambda > s^T x, \forall x \in K.$$

Vậy siêu phẳng  $s^T(x - \bar{x}) = 0$  tách  $K$  và  $\lambda$ . □

**Bổ đề 1.6.** Cho  $f(x)$  xác định trên một tập lồi  $K \subset \mathbb{R}^n, x' \in \text{int}(K)$ .

Nếu  $x^{(k)} \rightarrow x'$  là dãy định hướng bất kì với  $\delta^{(k)} \downarrow 0$  và  $s^{(k)} \rightarrow s$

(ở đây  $x^{(k)} - x' = \delta^{(k)} s^{(k)}, \forall k$ ) thì

$$\lim_{k \rightarrow \infty} \frac{f^{(k)} - f'}{\delta^{(k)}} = \max_{g \in \partial f'} s^T g. \quad (1.7)$$

*Chứng minh.* Ta có  $x^{(k)} = x' + \delta^{(k)} s^{(k)}$ .

Nếu  $g^{(k)} \in \partial f^{(k)}$  thì với mọi  $k$  đủ lớn ta có

$$\begin{aligned} f' = f(x') &= f(x^k + x' - x^k) \\ &\geq f(x^k) + (x' - x^k)^T g^{(k)} \\ &= f^{(k)} - (x^k - x')^T g^{(k)} \\ &= f^{(k)} - \delta^{(k)} s^{(k)T} g^{(k)} \end{aligned}$$

và

$$f^{(k)} \geq f' + \delta^{(k)} s^{(k)T} g, \forall g \in \partial f'.$$

Từ đó

$$s^{(k)T} g^{(k)} \geq \frac{f^{(k)} - f'}{\delta^{(k)}} \geq \max_{g \in \partial f'} s^T g. \quad (1.8)$$

Vì  $\partial f^{(k)}$  là một tập bị chặn trong một lân cận của  $x'$  ( theo Bổ đề 1.2 ) nên tồn tại dãy  $g^{(k)} \in \partial f^{(k)}$  mà  $g^{(k)} \rightarrow g'$  và  $g' \in \partial f'$  ( theo Bổ đề 1.1 ). Nếu (1.8) không đúng thì (1.9) chỉ ra mâu thuẫn tại giới hạn của dãy con nêu trên.  $\square$

### Nhận xét 1.2.

i) Nếu  $x^*$  là cực tiểu địa phương của hàm  $f(x)$  thì  $f^{(k)} \geq f^*$  với  $k$  đủ lớn và từ (1.8) ta có

$$\max_{g \in \partial f^*} s^T g \geq 0, \quad \forall s : \|s\| = 1. \quad (1.9)$$

Do đó điều kiện cần đối với cực tiểu địa phương là đạo hàm theo mọi hướng phải không âm. Từ đó

$$f'(g; s) \geq 0.$$

Điều này tương đương với

$$0 \in \partial f^*. \quad (1.10)$$

Đó là điều kiện tổng quát của đòi hỏi  $g^* = 0$  đối với các hàm trơn.

ii) Nếu  $0 \notin \partial f'$  thì theo Bổ đề 1.5 ( với  $\lambda = 0$ ,  $K = \partial f'$  ), tồn tại vectơ  $s = -\frac{\bar{g}}{\|\bar{g}\|_2}$  sao cho  $s^T g < 0$ ,  $\forall g \in \partial f'$  với  $\bar{g}$  là vectơ cực tiểu của  $\|g\|_2$ .

Từ kết quả này tại  $x^*$  thì (1.10) và (1.11) là tương đương. Ta thấy rằng cả (1.10) và (1.11) cũng là điều kiện đủ đối với cực tiểu toàn cục tại  $x^*$ .

Thật vậy, nếu  $0 \in \partial f(x^*)$  thì

$$f(x^* + \delta) \geq f(x^*) + \delta^T 0$$

hay

$$f(x^* + \delta) \geq f(x^*), \forall x^* + \delta \in K,$$

do đó  $x^*$  là cực tiểu toàn cục.

Cuối cùng để nghiên cứu sự tồn tại và duy nhất của điểm cực tiểu, chúng ta có kết quả sau đây:

**Mệnh đề 1.1.** *Cho  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  là một hàm lồi và  $C$  là một tập con lồi đóng khác rỗng của  $\mathbb{R}^n$ . Khi đó*

*i) Nếu  $f$  lồi chặt thì  $f$  có nhiều nhất một cực tiểu trên  $C$ .*

*ii) Nếu  $f$  lồi mạnh thì  $f$  có duy nhất điểm cực tiểu trên  $C$ .*

### 1.3 Phép toán về dưới vi phân

**Bổ đề 1.7.** *Cho  $A$  và  $B$  là hai tập con lồi compact khác rỗng của  $\mathbb{R}^n$ . Khi đó*

$$i) A \subseteq B \Leftrightarrow \Gamma_A \leq \Gamma_B$$

$$ii) A = B \Leftrightarrow \Gamma_A = \Gamma_B$$

trong đó  $\Gamma_A$  là hàm tựa của tập lồi  $A$  được định nghĩa bởi

$$\Gamma_A(x) = \sup_{y \in A} \langle y, x \rangle.$$

*Chứng minh.*

i) Theo định nghĩa của hàm tựa ta thấy ngay nếu  $A \subseteq B$  thì  $\Gamma_A \leq \Gamma_B$ . Để chứng minh chiều ngược lại ta giả sử  $A \not\subseteq B$ , tức là tồn tại  $a \in A$  và  $a \notin B$ . Vì  $B$  là tập lồi đóng khác rỗng nên từ định lý tách các tập lồi,  $a$  và  $B$  có thể được tách ngặt bởi một siêu phẳng, nghĩa là tồn tại  $s \in \mathbb{R}^n$

và  $\gamma \in \mathbb{R}$  sao cho

$$\langle s, b \rangle < \gamma < \langle s, a \rangle, \forall b \in B$$

mà

$$\Gamma_B(s) \leq \gamma < \langle s, a \rangle \leq \Gamma_A(s)$$

trái với giả thiết  $\Gamma_A \leq \Gamma_B$ . Vậy  $A \subseteq B$ .

ii) Suy ra từ (i). □

Trước hết ta xét dưới vi phân của một tổ hợp dương các hàm lồi:

**Mệnh đề 1.2.** Cho  $f_1, f_2 : \mathbb{R}^n \rightarrow \mathbb{R}$  là các hàm lồi và  $t_1, t_2 > 0$ . Khi đó

$$\partial(t_1 f_1 + t_2 f_2)(x) = t_1 \partial f_1(x) + t_2 \partial f_2(x) \quad \forall x \in \mathbb{R}^n.$$

*Chứng minh.* Lấy  $x \in \mathbb{R}^n$  và đặt

$$A = \partial(t_1 f_1 + t_2 f_2)(x)$$

và

$$B = t_1 \partial f_1(x) + t_2 \partial f_2(x).$$

Cả hai tập này đều là các tập lồi, khác rỗng và compact. Theo Bổ đề 1.7, nếu  $\Gamma_A = \Gamma_B$  thì  $A = B$ .

Ta có

$$\Gamma_A = (t_1 f_1 + t_2 f_2)'(x, \cdot)$$

$$\Gamma_B = t_1 \Gamma_{\partial f_1(x)} + t_2 \Gamma_{\partial f_2(x)} = t_1 f_1'(x, \cdot) + t_2 f_2'(x, \cdot).$$

Mặt khác, theo tính chất của đạo hàm theo hướng thì

$$(t_1 f_1 + t_2 f_2)'(x, \cdot) = t_1 f_1'(x, \cdot) + t_2 f_2'(x, \cdot)$$

nên  $\Gamma_A = \Gamma_B$ , do đó  $A = B$ . □

Sau đây ta sẽ kiểm tra dưới vi phân của cận trên đúng của các hàm lồi. Cho  $\{f_j\}_{j \in J}$  là tập hợp các hàm lồi từ  $\mathbb{R}^n$  vào  $\mathbb{R}$ . Ta xét hàm  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$  được định nghĩa bởi

$$f(x) = \sup_{j \in J} f_j(x), \quad \forall x$$

ở đây ta giả sử rằng  $f$  nhận giá trị hữu hạn. Dễ thấy  $f$  là hàm lồi và liên tục trên  $\mathbb{R}^n$ . Lấy  $x \in \mathbb{R}^n$ . Một câu hỏi đặt ra là làm thế nào để tính được  $\partial f(x)$  từ các  $\partial f_j(x), j \in J$ . Để giải quyết câu hỏi đó ta đưa ra tập

$$J(x) = \{ j \in J | f_j(x) = f(x) \},$$

$J(x)$  có thể rỗng. Ví dụ nếu  $J = \mathbb{N}_0$  và nếu  $f_j(x) = -\frac{1}{j}$  với mọi  $x$  và  $j$  thì  $f(x) = 0, \forall x$  và  $J(x) = \emptyset$ .

**Mệnh đề 1.3.** Với mọi  $x \in \mathbb{R}^n$  ta có

$$\partial f(x) \supseteq \overline{\text{conv}} \{ \cup \partial f_j(x) | j \in J(x) \}$$

trong đó  $\overline{\text{conv}}$  kí hiệu cho bao lồi đóng.

*Chứng minh.* Nếu  $J(x) = \emptyset$ , mệnh đề luôn đúng.

Vì vậy ta giả sử  $J(x) \neq \emptyset$ . Từ  $\partial f(x)$  lồi, đóng, ta chứng minh

$$\partial f_j(x) \subseteq \partial f(x), \quad \forall j \in J(x).$$

Lấy  $j \in J(x)$  và  $s \in \partial f_j(x)$ . Khi đó

$$f(y) \geq f_j(y) \geq f_j(x) + \langle s, y - x \rangle, \quad \forall y \in \mathbb{R}^n.$$

Từ  $f_j(x) = f(x)$  suy ra  $s \in \partial f(x)$ . Mệnh đề được chứng minh.  $\square$

Để đạt được dấu đẳng thức, ta giả sử  $J$  là tập hữu hạn,  $J(x) \neq \emptyset$ .

Ta có:

**Mệnh đề 1.4.** Nếu  $J = \{1, \dots, m\}$  thì

$$\partial f(x) = \text{conv} \{ \cup \partial f_j(x) | j \in J(x) \}, \forall x \in \mathbb{R}^n.$$

*Chứng minh.* Theo Mệnh đề 1.3 ta chứng minh được bao hàm thức  $\subseteq$ .

Để chứng minh phần còn lại ta xét

$$S = \{ \cup \partial f_j(x) | j \in J(x) \}.$$

Từ  $J$  là hữu hạn và  $\partial f_j(x)$  là compact với mọi  $j \in J(x)$ , ta có  $S$  là compact và do đó có  $\text{conv} S$ . Theo Bổ đề 1.7 ta chứng minh được

$$\Gamma_{\partial f(x)} \leq \Gamma_{\text{conv}(S)}$$

tức là

$$f'(x; d) \leq \sup_{j \in J(x)} f'_j(x; d), \forall d \in \mathbb{R}^n.$$

Thật vậy với mỗi  $d$  thì

$$\Gamma_{\text{conv}(S)}(d) = \Gamma_s(d) = \sup_{s \in S} \langle s, d \rangle = \sup_{j \in J(x)} \sup_{s_j \in \partial f_j(x)} \langle s_j, d \rangle = \sup_{j \in J(x)} f'_j(x; d).$$

Cho  $d \in \mathbb{R}^n$ ,  $d \neq 0$  và xét dãy  $t_k \rightarrow 0$ ,  $t_k > 0$  với mọi  $k$ . Khi đó với mỗi  $k$ , lấy một phần tử  $j_k$  trong tập  $J(x + t_k d)$ . Từ  $\{j_k\}_k$  là một dãy mà các phần tử thuộc vào tập hữu hạn  $\{1, \dots, m\}$ , tồn tại một dãy con của  $\{j_k\}$  mà ta vẫn kí hiệu là  $\{j_k\}$  sao cho  $j_k = j^*$  với mọi  $k$ . Hơn nữa  $j^* \in J(x)$ . Thật vậy với mọi  $k$  ta có

$$f_{j^*}(x + t_k d) = f_{j_k}(x + t_k d) = f(x + t_k d)$$

và cho  $k \rightarrow +\infty$  ta được

$$f_{j^*}(x) = f(x)$$

tức là  $j^* \in J(x)$  (ở đây ta đã sử dụng tính liên tục của  $f_{j^*}$  và  $f$  tại  $x$ ).

Cuối cùng với mỗi  $k$  ta có

$$f'(x; d) \leq \frac{f(x + t_k d) - f(x)}{t_k} = \frac{f_{j^*}(x + t_k d) - f_{j^*}(x)}{t_k}.$$

Cho  $k \rightarrow +\infty$  ta được

$$f'(x; d) \leq f'_{j^*}(x; d) \leq \sup_{j \in J(x)} f'_j(x; d)$$

bởi vì  $j^* \in J(x)$ . Mệnh đề được chứng minh.  $\square$

**Hệ quả 1.1.** Nếu  $f_1, \dots, f_m$  là các hàm lồi khả vi thì

$$\partial f(x) = \text{conv} \{ \nabla f_j(x) | j \in J(x) \}, \forall x \in \mathbb{R}^n.$$

**Ví dụ 1.3.** Xét hàm  $f(x) = \max \{ f_1(x), f_2(x), f_3(x) \}$

với

$$f_1(x) = -x_1 - x_2, \quad f_2(x) = -x_1 + x_2, \quad f_3(x) = x_1$$

và cho điểm  $(4, 8)$ .

Từ  $J((4, 8)) = \{2, 3\}$  và  $\nabla f_2(4, 8) = (-1, 1)^T$ ,  $\nabla f_3(4, 8) = (1, 0)^T$  ta có

$$\partial f(4, 8) = \text{conv} \{ (-1, 1)^T, (1, 0)^T \}.$$

Trong trường hợp khi tập  $J$  vô hạn, ta có thể chứng minh kết quả sau:

**Mệnh đề 1.5.** Giả sử  $J$  là tập compact ( trong không gian metric ) sao cho với mọi  $x \in \mathbb{R}^n$  hàm

$$\begin{aligned} f(x) : \quad J &\rightarrow \mathbb{R} \\ j &\mapsto f_j(x) \end{aligned}$$

là nửa liên tục trên ( tức là  $j_n \rightarrow j^*$  dẫn tới  $\limsup f_{j_n}(x) \leq f_{j^*}(x)$  ).

Khi đó với mọi  $x \in \mathbb{R}^n$  ta có

$$\partial f(x) = \text{conv} \{ \nabla f_j(x) | j \in J(x) \}.$$

## Chương 2

# Điều kiện tồn tại nghiệm tối ưu

### 2.1 Sự tồn tại nghiệm tối ưu

Trong thực tế ta thấy, để tìm một thứ gì đó trước hết ta phải xem xét nó có tồn tại hay không đã. Muốn sản xuất ra một loại hàng hoá nào đó, trước hết phải xem có phương án hay cách thức nào đó để sản xuất hay không? Muốn xây dựng một trung tâm thương mại ở khu dân cư sao cho tối ưu, trước hết phải tính toán xem có cách nào để đạt được không?... Nói tóm lại, muốn tìm được lời giải của một bài toán tối ưu, trước hết ta phải có cách nào đó nhận biết được xem nghiệm ấy có tồn tại hay không đã rồi mới đưa ra cách để tìm nó. Ta biết trong bài toán tối ưu có hai đối tượng quan trọng: Tập ràng buộc và hàm mục tiêu xác định trên tập đó. Vì thế khi ta xét đến điều kiện để tồn tại nghiệm tối ưu, ta phải quan tâm đến các điều kiện, tính chất của hai đối tượng ấy. Ví dụ, trong giải tích cổ điển, định lý Weierstrass khẳng định rằng một hàm liên tục trên một tập compact hay mở rộng là một hàm nửa liên tục dưới trên một tập compact khác rỗng bao giờ cũng đạt trên tập compact giá trị lớn nhất và giá trị nhỏ nhất. Nói cách khác, một bài toán tối ưu

có dữ kiện như vậy bao giờ cũng có nghiệm tối ưu. Đối với bài toán tối ưu trơn, nếu một điểm nào đó thuộc phần trong của miền nghiệm tối ưu thì đạo hàm của hàm số tại điểm ấy phải bằng không. Điều kiện như vậy được gọi là điều kiện cần tối ưu. Vậy muốn tìm nghiệm tối ưu của bài toán này, ta chỉ cần tìm trên tập con của miền ràng buộc mà trên đó đạo hàm của hàm số triệt tiêu. Tại những điểm này mà ta sử dụng những điều kiện liên quan tới đạo hàm bậc nhất để suy ra hàm đạt giá trị tối ưu thì những điều kiện đó được gọi là điều kiện đủ tối ưu cấp một. Tiếp theo, nếu hàm số có đạo hàm bậc hai và tại những điểm của tập con này, đạo hàm bậc hai dương chặt (hoặc âm chặt) thì điểm ấy chính là nghiệm tối ưu của bài toán. Điều kiện này được gọi là điều kiện đủ tối ưu cấp hai.

Mục đích của chương này là tìm các điều kiện cần và đủ để bài toán tối ưu không trơn có nghiệm dựa trên các thông tin về các tập dưới vi phân và ma trận Hessian. Trước hết ta nhắc lại khái niệm về các loại nghiệm của bài toán tối ưu.

Cho  $X$  là không gian tôpô tuyến tính lồi địa phương Hausdorff và  $D \subset X$  là tập hợp khác rỗng. Xét hàm  $f : D \rightarrow \mathbb{R}$ , ta có

### **Định nghĩa 2.1.**

*i)  $x_0 \in D$  là điểm cực tiểu (cực tiểu chặt) của  $f$  trên  $D$  nếu*

$$f(x_0) \leq f(x), \forall x \in D \quad (f(x_0) < f(x), \forall x \in D, x \neq x_0).$$

*ii)  $x_0 \in D$  được gọi là điểm cực tiểu địa phương nếu tồn tại lân cận  $U$  chứa  $x_0$  để các bất đẳng thức trên thoả mãn với mọi  $x \in U \cap D$ .*

Nhiều khi ta sử dụng kí hiệu

$$f(x_0) = \min_{x \in D} f(x) \quad (P)$$

chung cho các loại tối ưu trên.

Bài toán tìm cực đại của một hàm trên tập cho trước cũng được phát biểu một cách tương tự. Nhưng để ý

$$\min_{x \in D} f(x) = -\max_{x \in D} (-f(x))$$

ta suy ra bài toán cực đại hoàn toàn có thể quy về bài toán cực tiểu. Do đó trong lý thuyết tối ưu, nói chung ta chỉ cần xây dựng lý thuyết giải cho một trong hai loại: cực tiểu hoặc cực đại. Trong chương này ta chỉ quan tâm tới bài toán cực tiểu.

Chú thích:

*Nếu  $x_0 \in D$  là cực tiểu của  $f$  trên  $D$  thì  $x_0$  còn được gọi là cực tiểu toàn cục của  $f$  trên  $D$ . Khi ấy bài toán (P) được gọi là bài toán tối ưu toàn cục. Trái lại, bài toán (P) được gọi là bài toán tối ưu địa phương.*

Mệnh đề và định lý sau đây cho ta những điều kiện tổng quát nhất về sự tồn tại nghiệm tối ưu của các bài toán dạng trên.

**Mệnh đề 2.1.** *Điều kiện cần và đủ để tồn tại nghiệm cực tiểu của hàm  $f$  là tập hợp*

$$f(D)_+ = \{t \in \mathbb{R} | f(x) \leq t, x \in D\}$$

*đóng và có một cận dưới hữu hạn.*

*Chứng minh.* Giả sử  $x_0$  là nghiệm tối ưu của bài toán (P). Khi đó ta có

$$f(x_0) = \min_{x \in D} f(x), \quad f(D)_+ = [f(x_0), +\infty).$$

Hiển nhiên  $f(D)_+$  là tập đóng và nhận  $f(x_0)$  là một cận dưới.

Ngược lại, nếu tập  $f(D)_+$  có một cận dưới hữu hạn thì cận dưới lớn nhất

( hay infimum ) của tập này là hữu hạn và ta ký hiệu nó là  $t_0$ . Theo định nghĩa của infimum,  $t \geq t_0$  với mọi  $t \in f(D)_+$  và tồn tại  $\{t_n\} \subset f(D)_+$  hội tụ đến  $t_0$ . Vì  $f(D)_+$  là tập đóng nên  $t_0 \in f(D)_+$ . Theo định nghĩa của tập  $f(D)_+$  tồn tại  $x_0 \in D$  sao cho  $t_0 \geq f(x_0)$ . Hiển nhiên  $f(x_0) \in f(D)_+$  và vì  $t_0$  là cận dưới lớn nhất của tập  $f(D)_+$  nên ta có  $f(x_0) \geq t_0$ . Suy ra  $t_0 = f(x_0)$ . Điều đó chứng tỏ  $x_0$  là nghiệm tối ưu của bài toán (P).  $\square$

**Định lý 2.1.** *Cho  $D$  là tập compact khác rỗng. Khi đó nếu  $f$  nửa liên tục dưới trên  $D$  thì  $f$  đạt cực tiểu trên  $D$ .*

*Chứng minh.* Theo Mệnh đề 2.1, ta chỉ cần chỉ ra  $f(D)_+$  là tập đóng và bị chặn dưới. Thật vậy giả sử tập này không bị chặn dưới tức là tồn tại dãy điểm  $\{x_n\} \subset D$  sao cho

$$\lim_{n \rightarrow \infty} f(x_n) = -\infty.$$

Vì  $D$  là tập compact, không mất tính tổng quát ta có thể giả thiết

$$\lim_{n \rightarrow \infty} x_n = x_0 \in D.$$

Ta thấy rằng giá trị của hàm  $f$  tại  $x_0$  là hữu hạn và hàm  $f$  là nửa liên tục dưới, do đó không thể có

$$f(x_0) \leq \lim_{n \rightarrow \infty} f(x_n) = -\infty.$$

Vậy  $f(D)_+$  bị chặn dưới. Đặt  $t$  bằng cận dưới của tập này. Theo định nghĩa của infimum,  $t$  cũng là infimum của hàm  $f$  trên  $D$ . Do vậy tồn tại  $\{x_n\} \subset D$  sao cho

$$\lim_{n \rightarrow \infty} f(x_n) = t.$$

Vì  $D$  compact nên  $\lim_{n \rightarrow \infty} x_n = x_0 \in D$ . Do  $f$  nửa liên tục dưới kéo theo

$$t = \lim_{n \rightarrow \infty} f(x_n) \geq f(x_0).$$

Từ đây ta suy ra  $t = f(x_0)$  và  $x_0$  là cực tiểu của hàm  $f$  trên tập  $D$ .  $\square$

## 2.2 Các bài toán tối ưu

Cho hàm  $f : D \rightarrow \mathbb{R}$ . Bài toán

$$\min_{x \in D} f(x) \quad (P)$$

được gọi là bài toán tối ưu không ràng buộc.

Cho các hàm  $h_1, \dots, h_k : D \rightarrow \mathbb{R}$ ,  $g_1, \dots, g_m : D \rightarrow \mathbb{R}$ . Bài toán

$$\min_{x \in D_0} f(x) \quad (CP)$$

với

$$D_0 = \{ x \in D \mid g_i(x) \leq 0, h_j(x) = 0, \forall i = \overline{1, m}; j = \overline{1, k} \}$$

được gọi là bài toán tối ưu có ràng buộc. Tập  $D_0$  được gọi là tập chấp nhận được.

**Định nghĩa 2.2.** Điểm  $x_0 \in D$  được gọi là nghiệm tối ưu của bài toán  $(CP)$  nếu nó là cực tiểu của hàm  $f$  trên tập chấp nhận được  $D_0$ . Khi đó  $f(x_0)$  được gọi là giá trị tối ưu của bài toán.

**Định lý 2.2.** Cho  $D$  là tập compact trong không gian  $X$ ,  $f, g_1, \dots, g_m$  là các hàm nửa liên tục dưới và  $h_1, \dots, h_k$  là các hàm liên tục trong  $D$ . Khi đó bài toán  $(CP)$  có nghiệm nếu  $D \neq \emptyset$ .

*Chứng minh.* Định lý được chứng minh nhờ tính nửa liên tục dưới của các hàm số  $g_i$ , tính liên tục của các hàm  $h_j$  để đảm bảo tính compact của tập  $D_0$  và Định lý 2.1.  $\square$

## 2.3 Bài toán tối ưu không ràng buộc

Trong phần này, chúng ta sẽ nghiên cứu các điều kiện tối ưu đối với hàm hợp

$$\phi(x) = f(x) + h(c(x)) \quad (2.1)$$

với  $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}^1$ ,  $c(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$  là các hàm trơn thuộc lớp  $\mathbb{C}^1$ , còn  $h(c) : \mathbb{R}^m \rightarrow \mathbb{R}^1$  là các hàm lồi nhưng không trơn thuộc lớp  $\mathbb{C}^0$  và ở đây ta nghiên cứu hàm  $h(c)$  có dạng

$$h(c) = \max_i c^T h_i + b_i \quad (2.2)$$

với các vectơ  $h_i$  và các số  $b_i$  cho trước.

Như vậy  $h(c)$  là một hàm lồi đa diện và đồ thị của nó được tạo bởi một số hữu hạn các siêu phẳng tựa  $c^T h_i + b_i$ . Hầu hết sự quan tâm hướng về ba trường hợp đặc biệt sau đây mà trong đó  $b_i = 0$  với mọi  $i$

Trường hợp 1 :  $h(c) = \max_i c_i$

Trường hợp 2 :  $h(c) = \|c\|_\infty$

Trường hợp 3 :  $h(c) = \|c\|_1$

Khi đó dưới vi phân của các hàm trên lần lượt là:

$$\partial \max_i c_i = \{ \lambda : \sum \lambda_i = 1, \lambda \geq 0, c_i < \max_i c_i \Rightarrow \lambda_i = 0 \}.$$

$$\begin{aligned} \partial \|c\|_\infty = \{ \lambda : \quad c \neq 0 \Rightarrow \sum \lambda_i = 1 \\ |c_i| < \|c\|_\infty \Rightarrow \lambda_i = 0 \\ |c_i| = \|c\|_\infty \Rightarrow \lambda_i c_i \geq 0 \\ c = 0 \Rightarrow \sum \lambda_i \leq 1 \}. \end{aligned}$$

$$\partial \|c\|_1 = \{ \lambda : |\lambda_i| \leq 1, c_i \neq 0 \Rightarrow \lambda_i = \text{sign} c_i \}.$$

### 2.3.1 Điều kiện tối ưu cấp 1

Cho  $x^{(k)} \rightarrow x'$  là dãy định hướng với  $\delta^{(k)} \downarrow 0$  và  $s^{(k)} \rightarrow s$  ( tức là  $x^{(k)} = x' + \delta^{(k)} s^{(k)}$  ).

Theo khai triển Taylor ta có

$$f^{(k)} = f' + \delta^{(k)} g'^T s^{(k)} + o(\delta^{(k)})$$

trong đó  $g' = \nabla f(x')$ . Từ đó

$$\frac{f^{(k)} - f'}{\delta^{(k)}} \rightarrow g'^T s$$

và

$$c^{(k)} = c' + \delta^{(k)} A'^T s^{(k)} + o(\delta^{(k)}),$$

với  $A'$  là ma trận có cột  $i$  là  $\nabla c_i(x')$ . Vì thế  $c^{(k)} \rightarrow c'$  là dãy định hướng trong  $\mathbb{R}^m$  với

$$\frac{c^{(k)} - c'}{\delta^{(k)}} \rightarrow A'^T s.$$

Từ Bổ đề 1.6 với  $h(c)$  ta có

$$\lim_{k \rightarrow \infty} \frac{\phi^{(k)} - \phi'}{\delta^{(k)}} = \max_{\lambda \in \partial h'} s^T (g' + A' \lambda) \quad (2.3)$$

và điều này dẫn tới đạo hàm theo hướng tại  $x'$  là theo hướng  $s$  đối với hàm  $\phi(x)$  trong (2.1).

Từ (2.3) nếu  $x^*$  là cực tiểu địa phương của  $\phi(x)$  thì  $\phi^{(k)} \geq \phi^*$  với mọi  $k$  đủ lớn và do đó

$$\max_{\lambda \in \partial h^*} s^T (g^* + A^* \lambda) \geq 0, \quad \forall s : \|s\| = 1.$$

Đây là điều kiện cần cấp 1 đối với cực tiểu địa phương và cũng giống như (1.10), nó có thể được hiểu là đạo hàm theo mọi hướng là không âm. Một lần nữa ta đưa ra kết quả tương tự là

$$0 \in \partial \phi(x^*) = \{ \gamma : \gamma = g + A \lambda, \quad \forall \lambda \in \partial h \}_{x=x^*}. \quad (2.4)$$

Do đó tập  $\partial\phi^*$  xác định, là tập lồi, compact nhưng không khả dưới vi phân vì  $\phi$  có thể không là hàm lồi.

Để phát biểu điều kiện (2.4) theo một cách khác, ta đưa ra hàm Lagrange

$$L(x, \lambda) = f(x) + \lambda^T c(x).$$

Khi đó một phát biểu tương tự (2.4) là

**Định lý 2.3. (Điều kiện cần cấp một)**

*Nếu  $x^*$  là cực tiểu của  $\phi(x)$  thì tồn tại một vectơ  $\lambda^* \in \partial h^*$  thoả mãn*

$$\nabla L(x^*, \lambda^*) = g^* + A^* \lambda^* = 0. \quad (2.5)$$

*Chứng minh.* Định lý dễ dàng chứng minh dựa vào giả thiết  $\partial\phi^*$  là tập các vectơ  $\nabla L(x^*, \lambda)$  với mọi  $\lambda \in \partial h^*$ .  $\square$

Trong trường hợp tổng quát, hàm  $\phi(x)$  có thể không lồi. Khi đó điều kiện của Định lý 2.3 không phải là điều kiện đủ.

### 2.3.2 Điều kiện tối ưu cấp hai

Cho  $\lambda^*$  là một vectơ bất kì tồn tại trong Định lý 2.3 và xét

$$X = \{ x : h(c(x)) = h(c(x^*)) + (c(x) - c(x^*))^T \lambda^* \}. \quad (2.6)$$

**Định nghĩa 2.3.**  $\mathcal{G}^*$  là tập các phương tiếp xúc chấp nhận được của  $X$  tại  $x^*$  ( tức là nếu lấy  $s \in \mathcal{G}^*$  thì tồn tại một dãy định hướng  $x^{(k)} \rightarrow x^*$  với  $x^{(k)} \in X$  sao cho  $s^{(k)} \rightarrow s$  trong đó  $\|s^{(k)}\|_2 = 1$  ).

Nhận thấy các phương này liên quan chặt chẽ tới tập  $G^*$  là tập các hướng tiếp xúc có độ dốc 0, tức là

$$G^* = \{ s : \max_{\lambda \in \partial h^*} s^T (g^* + A^* \lambda) = 0, \|s\|_2 = 1 \}. \quad (2.7)$$

**Bổ đề 2.1.**  $\mathcal{G}^* \subset G^*$ .

*Chứng minh.* Lấy  $s \in \mathcal{G}^*$ , vì vậy tồn tại một dãy định hướng trong  $X$  là

$$s^{(k)} \rightarrow s, \quad \|s\|_2 = 1.$$

Từ (2.1), (2.3) và (2.6) ta có

$$\begin{aligned} \max_{\lambda \in \partial h^*} s^T(g^* + A^*\lambda) &= \lim_{k \rightarrow \infty} \frac{\phi^{(k)} - \phi^*}{\delta^{(k)}} \\ &= \lim_{k \rightarrow \infty} \frac{f^{(k)} - f^* + h(c^{(k)}) - h(c^*)}{\delta^{(k)}} \\ &= \lim_{k \rightarrow \infty} \frac{f^{(k)} - f^* + (c^{(k)} - c^*)^T \lambda^*}{\delta^{(k)}}. \end{aligned}$$

Khai triển Taylor ta có

$$\begin{aligned} f^{(k)} &= f^* + \delta^{(k)} g^{*T} s^{(k)} + o(\delta^{(k)}) \\ c^{(k)} &= c^* + \delta^{(k)} A^{*T} s^{(k)} + o(\delta^{(k)}). \end{aligned}$$

Do đó

$$\begin{aligned} &\frac{f^{(k)} - f^* + (c^{(k)} - c^*)^T \lambda^*}{\delta^{(k)}} \\ &= (g^* + A^* \lambda^*) s^{(k)T} + \frac{o(\delta^{(k)})}{\delta^{(k)}} + \frac{o(\delta^{(k)})}{\delta^{(k)}} \lambda^*. \end{aligned} \quad (2.8)$$

Trong (2.8) chuyển qua giới hạn khi  $k \rightarrow \infty$  ta được

$$\max_{\lambda \in \partial h^*} s^T(g^* + A^*\lambda) = s^T(g^* + A^*\lambda^*)$$

mà  $x^*$  là cực tiểu địa phương nên  $g^* + A^*\lambda^* = 0$ , do đó

$$\max_{\lambda \in \partial h^*} s^T(g^* + A^*\lambda) = 0,$$

suy ra  $s \in G^*$ . Vậy  $\mathcal{G}^* \subset G^*$ . □

Nếu  $\mathcal{G}^* = G^*$  thì điều kiện này được gọi là điều kiện chính quy.

**Định lý 2.4. (Điều kiện cần cấp hai)**

Nếu  $x^*$  là cực tiểu của  $\phi(x)$  thỏa mãn Định lý 2.3 và tồn tại  $\lambda^*$  để  $\mathcal{G}^* = G^*$  xảy ra thì

$$s^T \nabla^2 L(x^*, \lambda^*) s \geq 0, \forall s \in G^*.$$

*Chứng minh.* Với  $s \in G^*$  thì  $s \in \mathcal{G}^*$ . Do đó tồn tại một dãy định hướng chấp nhận được  $x^{(k)} \rightarrow x^*$ , tức là  $x^{(k)} = x^* + s^{(k)} \delta^{(k)}$  với  $s^{(k)} \rightarrow s$ .

Sử dụng khai triển Taylor đối với  $L(x, \lambda^*)$  tại  $x^*$  ta có

$$\begin{aligned} f(x^{(k)}) + \lambda^{*T} c(x^{(k)}) &= L(x^{(k)}, \lambda^*) \\ &= f^* + c^{*T} \lambda^* + e^{(k)T} \nabla L(x^*, \lambda^*) + \frac{1}{2} e^{(k)T} \nabla^2 L(x^*, \lambda^*) e^{(k)} + o(\|e^{(k)}\|^2) \\ &= L(x^*, \lambda^*) + \frac{1}{2} (\delta^{(k)})^2 s^{(k)T} \nabla^2 L(x^*, \lambda^*) s^{(k)} + o((\delta^{(k)})^2) \end{aligned} \quad (2.9)$$

với  $e^{(k)} = x^{(k)} - x^*$  và vì  $\|s^{(k)}\| = 1$ ,  $\delta^{(k)}$  là một vô hướng.

Từ giả thiết  $x^{(k)}$  là chấp nhận được ( $x^{(k)} \in X$ ) và từ (2.9) ta có

$$\phi^{(k)} = f^{(k)} + h(c^{(k)})$$

$$\phi^* = f^* + h(c^*)$$

nên

$$\begin{aligned} \phi^{(k)} - \phi^* &= f^{(k)} - f^* + h(c^{(k)}) - h(c^*) \\ &= f^{(k)} - f^* + (c^{(k)} - c^*)^T \lambda^* \\ &= L(x^{(k)}, \lambda^*) - c^{(k)T} \lambda^* - f^* + c^{(k)T} \lambda^* - c^{*T} \lambda^* \\ &= L(x^{(k)}, \lambda^*) - f^* - c^{*T} \lambda^* \\ &= \frac{1}{2} \delta^{(k)2} s^{(k)T} \nabla^2 L(x^*, \lambda^*) s^{(k)} + o((\delta^{(k)})^2). \end{aligned}$$

Vì  $x^*$  là cực tiểu địa phương nên  $\phi^{(k)} \geq \phi^*$  với  $k$  đủ lớn. Do đó

$$0 \leq \frac{\phi^{(k)} - \phi^*}{\frac{1}{2} \delta^{(k)2}} = s^{(k)T} \nabla^2 L(x^*, \lambda^*) s^{(k)} + \frac{o(\delta^{(k)2})}{\frac{1}{2} \delta^{(k)2}}$$

Cho  $k \rightarrow +\infty$  ta được

$$s^T \nabla^2 L(x^*, \lambda^*) s \geq 0, \forall s \in G^*. \quad \square$$

**Định lý 2.5. (Điều kiện đủ cấp hai)**

Nếu tồn tại  $\lambda^* \in \partial h^*$  thỏa mãn  $\nabla L(x^*, \lambda^*) = g^* + A^* \lambda^* = 0$  và nếu  $s^T \nabla^2 L(x^*, \lambda^*) s > 0, \forall s \in G^*$  thì  $x^*$  là cực tiểu địa phương chặt của  $\phi(x)$ .

*Chứng minh.* Giả sử ngược lại, khi đó sẽ tồn tại một dãy định hướng  $x^{(k)} \rightarrow x^*$  mà  $\phi^{(k)} \leq \phi^*$ .

Ta có

$$0 \leq \max_{\lambda \in \partial h^*} s^T (g^* + A^* \lambda^*) = \mu.$$

Nếu  $\mu > 0$  thì từ giới hạn (2.3)

$$\lim_{k \rightarrow +\infty} \frac{\phi^{(k)} - \phi^*}{\delta^{(k)}} = \mu$$

dẫn tới mâu thuẫn với  $\phi^{(k)} \leq \phi^*$ , suy ra  $\mu = 0$ . Vậy  $s \in G^*$ .

Xét

$$\begin{aligned} & L(x^{(k)}, \lambda^*) - L(x^*, \lambda^*) \\ &= f^{(k)} + c^{(k)T} \lambda^* - f^* - c^{*T} \lambda^* \\ &= \phi^{(k)} - \phi^* - [h(c^{(k)}) - h(c^*) - (c^{(k)} - c^*)^T \lambda^*] \\ &\leq \phi^{(k)} - \phi^*. \end{aligned}$$

Sử dụng bất đẳng thức dưới gradient, do  $\lambda^* \in \partial h^*$  nên

$$h(c^* + \delta) \geq h(c^*) + \delta^T \lambda^*.$$

Chọn  $\delta = c^{(k)} - c^*$ , suy ra

$$h(c^{(k)}) \geq h(c^*) + (c^{(k)} - c^*)^T \lambda^*.$$

Từ (2.9) ta có

$$0 \geq \phi^{(k)} - \phi^* \geq \frac{1}{2} \delta^{(k)^2} s^{(k)^T} \nabla^2 L(x^*, \lambda^*) s^{(k)} + o(\delta^{(k)^2})$$

nên

$$0 \geq \frac{\phi^{(k)} - \phi^*}{\frac{1}{2} \delta^{(k)^2}} \geq s^{(k)^T} \nabla^2 L(x^*, \lambda^*) s^{(k)} + \frac{o(\delta^{(k)^2})}{\frac{1}{2} \delta^{(k)^2}}.$$

Chuyển qua giới hạn khi  $k \rightarrow \infty$  ta được

$$0 \geq s^T \nabla^2 L(x^*, \lambda^*) s,$$

mâu thuẫn với giả thiết là

$$s^T \nabla^2 L(x^*, \lambda^*) s > 0 \quad \forall s \in G^*.$$

□

Ta biết rằng trong điều kiện cần cấp hai thì một giả thiết rất quan trọng đặt ra đó là điều kiện chính quy phải được thoả mãn, tức là  $\mathcal{G}^* = G^*$ . Vậy khi nào thì điều kiện chính quy xảy ra? Bổ đề sau đây sẽ làm sáng tỏ điều đó.

## **Bổ đề 2.2. (Điều kiện chính quy)**

Cho  $h(c) = \max_i (c^T h_i + b_i)$  và  $\mu_i^*$  là các nhân tử thoả mãn

$$\begin{cases} \sum_i \mu_i^* = 1 \\ \sum_i \mu_i^* (g^* + A^* h_i) = 0 \\ \mu_i^* \geq 0 \end{cases}$$

nếu tồn tại chỉ số  $p$  mà  $\mu_p^* > 0$  và các vectơ

$$A^*(h_p - h_i), \quad i \in \mathcal{A}^* \setminus p$$

là độc lập tuyến tính thì  $\mathcal{G}^* = G^*$ , trong đó

$$\mathcal{A}^* = \{ i : h(c^*) = c^{*T} h_i + b_i \}.$$

*Chứng minh.* Cho  $\lambda^* = \sum_i \mu_i^* h_i$  là các véctơ tương ứng trong  $\partial h^*$  và xét

$$\mathcal{A}_+^* = \{ i : \mu_i^* > 0 \}.$$

Lấy  $s \in G^*$  sao cho

$$\max_{\lambda \in \partial h^*} s^T(g^* + A^*\lambda) = 0.$$

Đặt

$$\mathcal{A}_s^* = \{ i : s^T(g^* + A^*\lambda) = 0, i \in \mathcal{A}^* \}$$

là kí hiệu tập hợp các véctơ  $h_i$  ( phụ thuộc vào  $s$  ) mà đạt max. Theo (2.5) thì  $\lambda^*$  thoả mãn

$$g^* + A^*\lambda^* = 0$$

nên

$$g^* + A^*\left(\sum_i \mu_i^* h_i\right) = g^* + \sum_i \mu_i^* A^* h_i = 0$$

Do đó:

$$\sum_i \mu_i^* g^* + \sum_i \mu_i^* A^* h_i = \sum_i \mu_i^* (g^* + A^* h_i) = 0.$$

Nếu  $\mu_i^* > 0$  ( tức là  $i \in \mathcal{A}_+^*$  ) thì  $g^* + A^* h_i = 0$ , suy ra  $i \in \mathcal{A}_s^*$ . Do đó  $\mathcal{A}_+^* \subset \mathcal{A}_s^* \subset \mathcal{A}^*$ . Vậy nếu  $p \in \mathcal{A}_+^*$  thì

$$s^T A^* (h_p - h_i) = 0, \quad i \in \mathcal{A}_s^* \setminus p$$

$$s^T A^* (h_p - h_i) > 0, \quad i \in \mathcal{A}^* \setminus \mathcal{A}_s^*.$$

Thật vậy: ta có đẳng thức

$$\begin{cases} \sum_i \mu_i^* (g^* + A^* h_i) = 0 \\ \mu_i^* \geq 0 \end{cases} \quad (2.10)$$

Với  $i \in \mathcal{A}_s^* \setminus p$ , suy ra  $\mu_p^* > 0$ ,  $\mu_i^* = 0$  với mọi  $i \in \mathcal{A}^* \setminus p$ . Vậy từ (2.10) ta có  $g^* + A^* h_p = 0$  hay  $A^* h_p = -g^*$ .

Với  $i \in \mathcal{A}_s^*$  thì

$$s^T(g^* + A^*h_i) = 0$$

hay

$$s^T A^*h_i = -s^T g^* = s^T A^*h_p$$

Vậy

$$s^T A^*(h_p - h_i) = 0 \quad \forall i \in \mathcal{A}_s^* \setminus p.$$

Nếu  $i \in \mathcal{A}^* \setminus \mathcal{A}_s^*$  thì  $i \notin \mathcal{A}_s^*$ , do đó

$$s^T(g^* + A^*h_i) < 0$$

hay

$$s^T A^*h_i < s^T(-g^*).$$

Vì  $p \in \mathcal{A}_+^*$  nên  $\mu_p^* > 0$  dẫn tới  $g^* + A^*h_p = 0$

hay

$$-g^* = A^*h_p$$

do đó

$$s^T A^*h_i < s^T A^*h_p$$

Vậy

$$s^T A^*(h_p - h_i) > 0, \quad i \in \mathcal{A}^* \setminus \mathcal{A}_s^*.$$

Nếu  $A^*(h_p - h_i)$ ,  $i \in \mathcal{A}^* \setminus p$  độc lập tuyến tính thì  $\mathcal{A}_s^* \setminus p$  chứa ít hơn  $n$  phần tử ( vì nếu chứa đúng  $n$  phần tử thì  $s^T = s^{(i)^T}$  với mọi  $i = \overline{1, n}$ , mâu thuẫn với  $\|s\| = 1$  ). Vì vậy tồn tại  $x(\theta)$  với  $x^* = x(0)$ ,  $\dot{x}(0) = s$  xác định với  $\theta \geq 0$  đủ nhỏ sao cho với  $\theta > 0$  thì

$$h_p^T c(x(\theta)) + b_p = h_i^T c(x(\theta)) + b_i, \quad i \in \mathcal{A}_s^* \setminus p$$

$$h_p^T c(x(\theta)) + b_p > h_i^T c(x(\theta)) + b_i, \quad i \in \mathcal{A}^* \setminus \mathcal{A}_s^*.$$

Do đó với  $\theta$  đủ nhỏ thì

$$h(c(x(\theta))) = h_i^T c(x(\theta)) + b_i$$

điều này dẫn tới

$$h(c(x(\theta))) - h^* = h_i^T (c(x(\theta)) - c^*), \forall i \in \mathcal{A}_s^*.$$

□

**Hệ quả 2.1.** Nếu trong tập  $\mathcal{A}^*$  chứa  $n + 1$  phần tử và  $\mathcal{A}_+^* = \mathcal{A}^*$  thì  $G^*$  là  $\emptyset$ .

*Chứng minh.* Từ giả thiết ta có  $\mathcal{A}_s^* \setminus p$  phải có  $n$  phần tử. Khi đó nếu  $G^* \neq \emptyset$  thì tồn tại  $s$  sao cho  $s^T = s^{(i)^T} = 0, \forall i = \overline{1, n}$ , mâu thuẫn với  $\|s\| = 1$ . Do đó  $G^* = \emptyset$  □

**Ví dụ 2.1.** Xét bài toán

$$\min \|c(x)\|_1$$

trong đó

$$c_1(x) = x_1 - x_2^3 + 5x_2^2 - 2x_2 - 12$$

$$c_2(x) = x_1 + x_2^3 + x_2^2 - 14x_2 - 29$$

Chứng minh:  $x^* = (6.4638, -0.8968)^T$  là cực tiểu địa phương của bài toán.

*Giải.* Từ giả thiết ta có:  $f(x) = 0$  và  $h(x) = \|c(x)\|_1$

Vì  $f(x) = 0$  nên  $g^* = 0$

Mặt khác:

$$c_1^* = c_1(x^*) = 0.9999 > 0$$

$$c_2^* = c_2(x^*) = -9.898 < 0$$

Nên  $\partial h^* = \{ (1, -1)^T \}$  dẫn tới  $\lambda^* = (1, -1)^T$ .

Lại có:

$$\nabla c_1(x) = \begin{pmatrix} 1 \\ -3x_2^2 + 10x_2 - 2 \end{pmatrix}, \quad \nabla c_2(x) = \begin{pmatrix} 1 \\ 3x_2^2 + 2x_2 - 14 \end{pmatrix}$$
$$A^* = \nabla c^{*T} = \begin{pmatrix} 1 & 1 \\ -13.38 & -13.38 \end{pmatrix}$$

$$\nabla L(x^*, \lambda^*) = g^* + A^* \lambda^* = 0.$$

Vậy  $x^*$  thoả mãn điều kiện cần cấp 1.

Ta có:

$$\begin{aligned} G^* &= \{ s : \max_{\lambda \in \partial h^*} s^T (g^* + A^* \lambda) = 0, \|s\|_2 = 1 \} \\ &= \{ s : \|s\|_2 = 1 \}. \end{aligned}$$

Xét hàm Lagrange:

$$\begin{aligned} L(x, \lambda) &= \lambda_1 c_1(x) + \lambda_2 c_2(x) \\ &= \lambda_1 (x_1 - x_2^3 + 5x_2^2 - 2x_2 - 12) + \lambda_2 (x_1 + x_2^3 + x_2^2 - 14x_2 - 29). \end{aligned}$$

Ta có:

$$\nabla L(x, \lambda) = \begin{pmatrix} \lambda_1 + \lambda_2 \\ -3\lambda_1 x_2^2 + 10\lambda_1 x_2 - 2\lambda_1 + 3\lambda_2 x_2^2 + 2x_2 \lambda_2 - 14\lambda_2 \end{pmatrix}$$

và

$$\nabla^2 L(x, \lambda) = \begin{pmatrix} 0 & 0 \\ 0 & \lambda_1(-6x_2 + 10) + \lambda_2(6x_2 + 2) \end{pmatrix}$$

Do đó

$$\nabla^2 L(x^*, \lambda^*) = \begin{pmatrix} 0 & 0 \\ 0 & 18.76 \end{pmatrix}$$

Xét

$$s^T \nabla^2 L(x^*, \lambda^*) s = (s_1, s_2) \begin{pmatrix} 0 & 0 \\ 0 & 18.76 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = 18.76 s_2^2 \geq 0, \forall s \in G^*.$$

Vậy  $x^*$  thoả mãn điều kiện đủ cấp 2 nên  $x^*$  là cực tiểu địa phương của bài toán.

**Ví dụ 2.2.** Xét bài toán

$$\min \|c\|_\infty$$

trong đó

$$c_1(x) = x_1 - x_2^3 + 5x_2^2 - 2x_2 - 13$$

$$c_2(x) = x_1 + x_2^3 + x_2^2 - 14x_2 - 29$$

$$c_3(x) = 0.4336x_1$$

Chứng minh:  $x^* = (11.4128, -0.8968)^T$  là cực tiểu địa phương của bài toán.

*Giải.* Từ giả thiết ta có:  $f(x) = 0$  và  $h(x) = \|c(x)\|_\infty$

Vì  $f(x) = 0$  nên  $g^* = 0$

Mặt khác:

$$c^* = c(x^*) = (4.949, -4.949, 4.949)$$

Do đó  $\|c^*\|_\infty = 4.949$ .

Vậy  $\partial h^* = \{ (\lambda_1, \lambda_2, \lambda_3)^T : \lambda_1, \lambda_3 \geq 0, \lambda_2 \leq 0, |\lambda_1| + |\lambda_2| + |\lambda_3| = 1 \}$ .

Với  $\lambda^* \in \partial h^*$  thì  $\lambda^* = (\lambda_1, \lambda_2, \lambda_3)^T$  trong đó  $\lambda_1, \lambda_3 \geq 0, \lambda_2 \leq 0$  và  $|\lambda_1| + |\lambda_2| + |\lambda_3| = 1$ .

Ta có

$$\nabla c_1(x) = \begin{pmatrix} 1 \\ -3x_2^2 + 10x_2 - 2 \end{pmatrix}, \quad \nabla c_2(x) = \begin{pmatrix} 1 \\ 3x_2^2 + 2x_2 - 19 \end{pmatrix}$$

$$\nabla c_3(x) = \begin{pmatrix} 0.4336 \\ 0 \end{pmatrix}$$

Vậy

$$A^* = \nabla c^{*T} = \begin{pmatrix} 0.4336 & 1 & 1 \\ 0 & -13.38 & -13.38 \end{pmatrix}.$$

Hàm Lagrange cho bài toán:

$$\begin{aligned} L(x, \lambda) &= \lambda_1(x_1 - x_2^3 + 5x_2^2 - 2x_2 - 13) \\ &\quad + \lambda_2(x_1 + x_2^3 + x_2^2 - 14x_2 - 29) + 0.4336\lambda_3x_3. \end{aligned}$$

Khi đó

$$\begin{aligned} \nabla L(x^*, \lambda^*) &= g^* + A^*\lambda^* = 0 \\ \Leftrightarrow \begin{pmatrix} 0.4336 & 1 & 1 \\ 0 & -13.38 & -13.38 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} &= 0 \\ \Leftrightarrow \begin{cases} 0.4336\lambda_1 + \lambda_2 + \lambda_3 = 0 \\ \lambda_2 + \lambda_3 = 0 \end{cases} \end{aligned}$$

Kết hợp với điều kiện  $\lambda_1, \lambda_3 \geq 0, \lambda_2 \leq 0, |\lambda_1| + |\lambda_2| + |\lambda_3| = 1$  ta được

$$\lambda_1 = 0, \lambda_2 = -0.5, \lambda_3 = 0.5$$

$$\Rightarrow \lambda^* = (0, -0.5, 0.5)^T.$$

Do đó với  $\lambda^* = (0, -0.5, 0.5)^T$  thì  $x^*$  thoả mãn điều kiện cần cấp 1.

Tiếp theo ta sẽ xác định tập

$$G^* = \{ s : \|s\|_2 = 1, \max_{\lambda \in \partial h^*} s^T(g^* + A^*\lambda) = 0 \}$$

Ta có

$$\max_{\lambda \in \partial h^*} s^T(g^* + A^*\lambda) = 0$$

Hay

$$\max_{\lambda \in \partial h^*} s_1(0.5664\lambda_1 + \lambda_2 + 1) - 13.38s_2(1 + \lambda_2 - \lambda_1) = 0$$

Điều này chỉ xảy ra khi

$$s_1 \leq 0, \quad s_2 \geq 0$$

Từ đó suy ra

$$G^* = \{ s : \|s\|_2 = 1, s_1 \leq 0, s_2 \geq 0 \}.$$

Lại có

$$\nabla L(x, \lambda) = \begin{pmatrix} \lambda_1 + \lambda_2 + 0.4336\lambda_3 \\ 3(\lambda_2 - \lambda_1)x_2^2 + 2(\lambda_2 + 5\lambda_1)x_2 - 2(\lambda_1 + 7\lambda_2) \end{pmatrix}$$

và

$$\nabla^2 L(x, \lambda) = \begin{pmatrix} 0 & 0 \\ 0 & 6(\lambda_2 - \lambda_1)x_2 + 2(\lambda_2 + 5\lambda_1) \end{pmatrix}$$

Do đó

$$\nabla^2 L(x^*, \lambda^*) = \begin{pmatrix} 0 & 0 \\ 0 & 1.6904 \end{pmatrix}.$$

Xét

$$s^T \nabla^2 L(x^*, \lambda^*) s = 1.6904s_2^2 \geq 0, \quad \forall s \in G^*.$$

Vậy  $x^*$  thỏa mãn điều kiện đủ cấp 2 nên  $x^*$  là cực tiểu địa phương của bài toán.

Từ kết quả trên ta có thể đưa ra một sự so sánh giữa bài toán tối ưu của hàm hợp không trơn với bài toán quy hoạch phi tuyến.

Rõ ràng hàm  $\phi(x)$  xác định theo công thức (2.1) có thể viết dưới dạng

$$\phi(x) = \max_i (f(x) + c(x)^T h_i + b_i) \quad (2.11)$$

và tương đương với

$$\phi(x) = \min v : v \geq f(x) + c(x)^T h_i + b_i, \forall i. \quad (2.12)$$

Do đó  $x^*$  là cực tiểu địa phương của  $\phi(x)$  khi và chỉ khi  $x^*, v^*$  là nghiệm địa phương của bài toán quy hoạch phi tuyến

$$\min_{x,v} v$$

với điều kiện

$$v - f(x) - c(x)^T h_i \geq b_i, \forall i. \quad (2.13)$$

Vì vậy điều kiện tối ưu cấp một và cấp hai có thể áp dụng kết quả tương tự trong phần quy hoạch phi tuyến cho bài toán (2.13).

Tiếp theo ta nhắc lại các điều kiện tối ưu của bài toán quy hoạch phi tuyến. Xét bài toán quy hoạch phi tuyến có ràng buộc

$$\min f(x), x \in \mathbb{R}^n$$

với

$$\begin{cases} c_i(x) = 0, & i \in E \\ c_i(x) \geq 0, & i \in I. \end{cases} \quad (2.14)$$

Xét hàm số Lagrange

$$L(x, \lambda) = f(x) - \sum_i \lambda_i c_i(x).$$

Kí hiệu

$$W^* = \nabla_x^2 L(x^*, \lambda^*) = \nabla^2 f(x^*) - \sum_i \lambda_i^* \nabla^2 c_i(x^*)$$

là ma trận Hessian của hàm Lagrange theo  $x$  và

$$a_i^* = \nabla c_i(x^*)$$

$$\mathcal{A}^* = \mathcal{A}(x^*) = \{ i : c_i(x^*) = 0 \}$$

$$\mathcal{A}_+^* = \{ i \mid i \in E \text{ hoặc } \lambda_i^* > 0 \}$$

$$G^* = \{ s \mid s \neq 0, (a_i^*)^T s = 0, i \in \mathcal{A}_+^*, (a_i^*)^T s \geq 0, i \in \mathcal{A}^* \setminus \mathcal{A}_+^* \}$$

$\mathcal{G}^*$  là tập các hướng chấp nhận được tại  $x^*$ .

Khi đó ta có các điều kiện tối ưu sau của bài toán quy hoạch phi tuyến có ràng buộc

### **Định lý 1 (Điều kiện cấp một)**

*Nếu  $x^*$  là cực tiểu địa phương của bài toán (2.14) và nếu  $G^* = \mathcal{G}^*$  được thoả mãn tại  $x^*$  thì tồn tại nhân tử Lagrange  $\lambda^*$  sao cho  $x^*$  và  $\lambda^*$  thoả mãn hệ thức sau*

$$\nabla_x L(x, \lambda) = 0$$

$$c_i(x) = 0, \quad i \in E$$

$$c_i(x) \geq 0, \quad i \in I \tag{2.15}$$

$$\lambda_i \geq 0, \quad i \in I$$

$$\lambda_i c_i(x) = 0, \quad \forall i.$$

### **Định lý 2 (Điều kiện cấp hai)**

*Nếu tại  $x^*$  tồn tại nhân tử  $\lambda^*$  thoả mãn (2.15) và nếu*

$$s^T W^* s > 0, \quad \forall s \in G^*$$

*thì  $x^*$  là cực tiểu địa phương chặt của bài toán (2.14).*

Bây giờ trở lại bài toán (2.13). Hàm Lagrange thích hợp cho Định lý 1 là

$$L(x, v, \mu) = v - \sum_i \mu_i (v - f(x) - c(x)^T h_i - b_i)$$

và nó thoả mãn tại  $x^*, v^*$  nếu tồn tại  $\mu^*$  sao cho

$$\begin{cases} \frac{\partial}{\partial v} L(x^*, v^*, \mu^*) = 0 & \text{hay } \sum_i \mu_i^* = 1; \\ \nabla_x L(x^*, v^*, \mu^*) = 0 & \text{hay } \sum_i \mu_i^* (g^* + A^* h_i) = 0 \\ \mu^* \geq 0 \\ \mu_i^* > 0 \Rightarrow v^* = f^* + c^{*T} h_i + b_i \end{cases} \quad (2.16)$$

Khi đó mối liên hệ giữa bài toán tối ưu không trơn và bài toán quy hoạch phi tuyến được đưa ra trong bổ đề sau

**Bổ đề 2.3.** Với  $x^*$  cho trước, cho  $\mu^*$  thoả mãn (2.16) và  $\lambda^* = H \cdot \mu^*$  ( $H$  là ma trận có các cột là  $h_i$ ). Khi đó điều kiện cấp hai của Định lý 2 đối với bài toán (2.13) là tương đương với các điều kiện trong Định lý 2.5. *Chứng minh.* Trong cả hai trường hợp này, các điều kiện cấp một thoả mãn.

Lấy  $\bar{s}^T = (s^T, s_{n+1})$ . Điều kiện cấp hai của Định lý 2 đối với bài toán (2.13) liên quan đến tập

$$M = \{ \bar{s} : \bar{s} \neq 0, \quad s_{n+1} = s^T (g^* + A^* h_i), \quad i \in \mathcal{A}_+^*, \\ s_{n+1} \geq s^T (g^* + A^* h_i), \quad i \in \mathcal{A}^* \setminus \mathcal{A}_+^* \}.$$

Lấy  $\bar{s} \in M$ . Khi đó một tích trong với  $\mu_i^*$  dẫn đến  $s_{n+1} = 0$ . Thật vậy, ta có

$$\sum_i s_{n+1} \mu_i^* = \sum_i s^T (g^* + A^* h_i) \mu_i^*$$

hay

$$s_{n+1} \sum_i \mu_i^* = s^T \sum_i (g^* + A^* h_i) \mu_i^*$$

Vậy  $s_{n+1} = 0$  vì  $\sum_i \mu_i^* = 1$  và  $\sum_i (g^* + A^* h_i) \mu_i^* = 0$ .

Với  $\lambda \in \partial h^*$  tùy ý, cho  $\lambda = H \mu$ .

Khi đó một tích trong với  $\mu_i$  dẫn tới  $s^T(g^* + A^*\lambda) \leq 0$ , do đó

$$\frac{s}{\|s\|_2} \in G^*.$$

Hơn nữa  $\|\frac{s}{\|s\|_2}\| = 1$  và  $\left(\frac{s}{\|s\|_2}\right)^T (g^* + A^*\lambda) \leq 0$  nên

$$\max \left( \left( \frac{s}{\|s\|_2} \right)^T (g^* + A^*\lambda) \right) = 0.$$

Bây giờ cho  $s \in G^*, p \in \mathcal{A}_+^*$  và định nghĩa

$$s_{n+1} = s^T(g^* + A^*h_p).$$

Theo Bổ đề 2.2 ta có

$$s_{n+1} = s^T(g^* + A^*h_i), \quad i \in \mathcal{A}_s^*$$

$$s_{n+1} \geq s^T(g^* + A^*h_i), \quad i \in \mathcal{A} \setminus \mathcal{A}_s^*.$$

Từ  $\mathcal{A}_s^* \supset \mathcal{A}_+^*$  ta có  $\bar{s} \in M$ .

Ngoài ra việc chuẩn hoá các vectơ trong tập  $M$  và trong  $G^*$  là tương đương. Do đó các điều kiện cấp hai  $s^T W^* s > 0, \forall s \in M$  hoặc  $s \in G^*$  là tương đương. Vậy bổ đề được chứng minh.  $\square$

## 2.4 Bài toán tối ưu có ràng buộc

Xét bài toán tối ưu không trơn có ràng buộc của hàm hợp sau đây

$$\min_{x \in \mathbb{R}^n} \phi(x) = f(x) + h(c(x)) \quad (2.17)$$

$$\text{với điều kiện } t(r(x)) \leq 0$$

trong đó hàm mục tiêu là hàm hợp đã được đề cập tới trong mục (2.3),  $r(x) : \mathbb{R}^n \rightarrow \mathbb{R}^p$  là hàm trơn thuộc lớp  $\mathbb{C}^1$ ,  $t(r) : \mathbb{R}^p \rightarrow \mathbb{R}$  là hàm lồi

nhưng không tron thuộc lớp  $\mathbb{C}^0$ .

Trước hết ta đưa ra khái niệm về tập các hướng chấp nhận được tại điểm chấp nhận được  $x'$  như sau:

$$\mathcal{F}' = \{ s : s \neq 0 : \exists \{x^{(k)}\}, t(r(x^{(k)})) \leq 0, \\ x^{(k)} \rightarrow x', s^{(k)} \rightarrow s, \delta^{(k)} \downarrow 0 \} \quad (2.18)$$

trong đó  $x^{(k)} = x' + \delta^{(k)} s^{(k)}$ .

Tập này liên quan đến tập sau đây

$$F' = \{ s : s \neq 0, t' = 0 \Rightarrow \max_{u \in \partial t'} s^T R' u \leq 0 \} \quad (2.19)$$

với  $R = \nabla r^T, t' = t(r(x'))$ .

Khi đó tập  $F'$  có thể được xem như tập các hướng chấp nhận được đối với các ràng buộc tuyến tính tại  $x'$  và sẽ rất thuận lợi nếu các tập  $\mathcal{F}'$  và  $F'$  trùng nhau. Vì thế việc xem xét đánh giá này rất quan trọng. Trước hết ta có bổ đề sau nói lên mối quan hệ giữa  $\mathcal{F}'$  và  $F'$ .

**Bổ đề 2.4.**  $\mathcal{F}' \subseteq F'$

*Chứng minh.* Lấy  $s \in \mathcal{F}'$ , khi đó sẽ tồn tại một dãy định hướng  $x^{(k)} \rightarrow x'$  sao cho  $s^{(k)} \rightarrow s$ . Sử dụng khai triển Taylor tại  $x'$  ta có

$$r^{(k)} = r' + \delta^{(k)} R'^T s^{(k)} + o(\delta^{(k)}).$$

Vì thế  $r^{(k)} - r'$  là một dãy định hướng trong  $\mathbb{R}^p$  với

$$\frac{r^{(k)} - r'}{\delta^{(k)}} \rightarrow R'^T s.$$

Do đó áp dụng Bổ đề 1.6 với hàm  $t(r)$  thì

$$\max_{u \in \partial t'} s^T R' u = \lim_{\delta^{(k)} \downarrow 0} \frac{t^{(k)} - t'}{\delta^{(k)}} \leq 0$$

nếu  $t' = 0$ . Vậy  $s \in F'$ .

Bây giờ ta xem xét điều kiện để  $\mathcal{F}' = F'$  tại điểm chấp nhận được  $x'$ . Giả sử  $\partial t'$  có số chiều là  $q'$  và  $u' \in \partial t'$  tùy ý. Gọi  $H' \in \mathbb{R}^{p \times q'}$  là ma trận có các cột là  $f'_i, i = 1, 2, \dots, q'$  trong đó  $f'_i$  là cơ sở của  $\partial t' - u'$ .

Khi đó  $\partial t'$  có thể biểu diễn dưới dạng

$$\partial t' = \{ u : u = u' + H'.v, v \in V' \subset \mathbb{R}^{q'} \} \quad (2.20)$$

Điều kiện để  $\mathcal{F}' = F'$  tại điểm chấp nhận được  $x'$  được nêu trong bổ đề dưới đây:

**Bổ đề 2.5.** *Điều kiện đủ để  $\mathcal{F}' = F'$  tại điểm chấp nhận được  $x'$  là*

i)  $t' < 0$

ii) *Nếu  $t' = 0$  và hàm  $t(r)$  là tuyến tính địa phương theo  $r'$  ( tức là tồn tại một lân cận mở  $\Omega$  của  $r'$  sao cho  $t(r) = t(r') + \max_{\lambda \in \partial t'} (r - r')^T \lambda$  ) thì  $\text{rank}( R'[ u' : H' ] ) = q' + 1$  hoặc hàm  $r(x)$  là affine.*

*Chứng minh.*

i) Nếu  $t' < 0$  thì  $\mathcal{F}' = \mathbb{R}^n \setminus \{0\}$ , do đó  $\mathcal{F}' = F'$ .

ii) Giả thiết  $t' = 0$ . Vì  $\mathcal{F}' \subseteq F'$  nên ta chỉ cần chứng minh  $F' \subseteq \mathcal{F}'$ .

Lấy  $s \in F'$ .

Nếu  $\max_{u \in \partial t'} s^T R'u < 0$  thì lấy  $x^{(k)} = x' + \delta^{(k)}s$  với dãy  $\delta^{(k)} \downarrow 0$  bất kì, điều này dẫn tới  $t^{(k)} \leq 0$  với  $k$  đủ lớn và do đó  $s \in \mathcal{F}'$  ( vì nếu không sẽ tồn tại một dãy con  $t^{(k)} > 0$ , sử dụng khai triển Taylor và Bổ đề 1.6 sẽ dẫn tới  $\max_{u \in \partial t'} s^T R'u \geq 0$ , mâu thuẫn ).

Nếu  $\max_{u \in \partial t'} s^T R'u = 0$ , lấy  $u' \in \partial t'$  là vectơ bất kì sao cho  $s^T R'u' = 0$ . Không mất tính tổng quát  $u'$  có thể xem như là một vectơ tùy ý trong (2.20). Ta cũng định nghĩa

$$\partial t'_s = \{ u \in \partial t' : s^T R'u = 0 \}.$$

Cho số chiều của  $\partial t'_s$  là  $q'_s$  ( $q'_s < q'$ ) và không mất tổng quát cho  $f'_i$  với  $i = 1, 2, \dots, q'_s$  là một cơ sở của  $\partial t'_s - u'$ . Từ đó

$$s^T R' f'_i \begin{cases} = 0, & i = 1, 2, \dots, q'_s \\ < 0, & i = q'_s + 1, \dots, q'. \end{cases}$$

Nếu  $q'_s + 1 = n$  thì  $s^T R' [u' : H'] = 0^T$  và do đó  $s = 0$  vì theo giả thiết  $\text{rank}(R' [u' : H']) = q' + 1$ , mâu thuẫn với  $s \in F'$ .

Nếu  $q'_s + 1 < n$  thì sẽ tồn tại một hàm trơn  $\text{arc } x(\theta)$ ,  $\theta \in [0, \bar{\theta}]$  với  $x(0) = x'$ ,  $\dot{x}(0) = s$  và

$$(r(x(\theta)) - r')^T u' = \theta s^T R' u' = 0,$$

$$(r(x(\theta)) - r')^T f'_i = \theta s^T R' f'_i \begin{cases} = 0, & i = 1, 2, \dots, q'_s \\ < 0, & i = q'_s + 1, \dots, q'. \end{cases}$$

Điều này dẫn tới

$$\max_{u \in \partial t'} (r(x(\theta)) - r')^T u = 0. \quad (2.21)$$

Sử dụng điều kiện  $t(r)$  là tuyến tính địa phương với  $r'$  thì sẽ tồn tại một lân cận của  $r'$  sao cho  $t(r(x(\theta))) = 0$  và lấy bất kì dãy  $\theta^{(k)} \downarrow 0$  sẽ cho được một dãy định hướng với  $s \in \mathcal{F}'$ .

Cuối cùng nếu  $r(x)$  là hàm affine thì ta

$$x(\theta) = x' + \theta s$$

có

$$r(x(\theta)) = r' + \theta R'^T s.$$

Do đó có thể suy ra hệ thức (2.21) ở trên trực tiếp từ  $\max_{u \in \partial t'} s^T R' u = 0$ .

Điều này một lần nữa dẫn tới  $s \in \mathcal{F}'$ .

Vậy  $F' \subseteq \mathcal{F}'$ , do đó  $F' = \mathcal{F}'$ . □

### 2.4.1 Điều kiện tối ưu cấp một

Xét tập các hướng giảm tại  $x^*$

$$\mathcal{D}(x^*) = \mathcal{D}^* = \{s : \max_{\lambda \in \partial h^*} s^T(g^* + A^*\lambda) < 0\}.$$

Khi đó ta có bổ đề sau

**Bổ đề 2.6.** *Nếu  $x^*$  là cực tiểu địa phương thì  $\mathcal{F}^* \cap \mathcal{D}^* = \emptyset$ .*

*Chứng minh.* Lấy  $s \in \mathcal{F}^*$ , khi đó tồn tại một dãy định hướng chấp nhận được  $x^{(k)} \rightarrow x^*$  với  $s^{(k)} \rightarrow s$ .

Sử dụng khai triển Taylor tại  $x^*$

$$\begin{aligned} f^{(k)} &= f^* + \delta^{(k)} g^{*T} s^{(k)} + o(\delta^{(k)}) \\ c^{(k)} &= c^* + \delta^{(k)} A^{*T} s^{(k)} + o(\delta^{(k)}). \end{aligned}$$

Theo tính chất tối ưu địa phương nên  $\phi^{(k)} \geq \phi^*$  với  $k$  đủ lớn và do đó

$$0 \leq \frac{\phi^{(k)} - \phi^*}{\delta^{(k)}} = \frac{f^{(k)} - f^*}{\delta^{(k)}} + \frac{h(c^{(k)}) - h(c^*)}{\delta^{(k)}}.$$

Chuyển qua giới hạn khi  $k \rightarrow \infty$  và sử dụng Bổ đề 1.6 cùng với thực tế là  $c^{(k)} \rightarrow c^*$  là một dãy định hướng với hướng là  $A^{*T}s$  sẽ dẫn tới

$$0 \leq \max_{\lambda \in \partial h'} s^T(g^* + A^*\lambda).$$

Điều này mâu thuẫn với  $s \in \mathcal{D}^*$ . Vậy bổ đề được chứng minh.  $\square$

**Bổ đề 2.7.** *Nếu trong  $\mathbb{R}^n$ ,  $C$  là một nón lồi đóng,  $B$  là một tập lồi compact khác rỗng và  $B \cap C = \emptyset$ , khi đó tồn tại một siêu phẳng  $s^T x = 0$  tách  $B$  và  $C$ .*

*Chứng minh.* Vì  $B$  là tập compact nên tồn tại các điểm  $\hat{b} \in B, \hat{g} \in C$  làm cực tiểu hàm  $\|b - g\|_L, \forall b \in B, g \in C$ . Từ đó với bất kì  $b \in B$ , do

tính lồi của  $B$  mà

$$(1 - \theta)\hat{b} + \theta b \in B, \theta \in [0, 1].$$

Do đó

$$\|\hat{b} - \hat{g} + \theta(b - \hat{b})\|_2^2 = \|(1 - \theta)\hat{b} + \theta b - \hat{g}\|_2^2 \geq \|\hat{b} - \hat{g}\|_2^2.$$

Chuyển qua giới hạn khi  $\theta \downarrow 0$  ta được

$$(b - \hat{b})^T(\hat{g} - \hat{b}) \leq 0.$$

Nếu  $s = \hat{g} - \hat{b}$  thì  $s^T x \geq 0, \forall x \in C$  và  $s^T \hat{b} < 0$ .

Từ đó  $s^T \hat{b} < 0, \forall b \in B$ . Bổ đề được chứng minh.  $\square$

Bây giờ ta thiết lập hàm Lagrange đối với bài toán (2.17)

$$L(x, \lambda, u, \pi) = f(x) + \lambda^T c(x) + \pi u^T r(x).$$

### **Định lý 2.6. (Điều kiện cần cấp một)**

*Nếu  $x^*$  là cực tiểu địa phương của bài toán (2.17) và điều kiện chính quy*

*$\mathcal{F}^* \cap \mathcal{D}^* = F^* \cap \mathcal{D}^*$  được thoả mãn thì tồn tại các nhân tử  $\lambda^* \in \partial h^*, u^* \in \partial t^*, \pi^* \geq 0$  sao cho*

$$t^* \leq 0$$

$$\pi^* t^* = 0$$

và

$$0 = \nabla L(x^*, \lambda^*, u^*, \pi^*) = g^* + A^* \lambda^* + \pi^* R^* u^*. \quad (2.22)$$

*Chứng minh.* Giả sử  $x^*$  là cực tiểu địa phương của bài toán (2.17). Để chứng minh định lý, ta sẽ chứng minh rằng  $F^* \cap \mathcal{D}^* = \emptyset$  khi và chỉ khi

các điều kiện của định lý được thoả mãn.

Nếu các điều kiện của định lý được thoả mãn thì khi đó lấy  $s \in F^*$  ta có

Nếu  $t^* < 0$  thì  $\pi^* = 0$ , suy ra  $g^* + A^*\lambda^* = 0$  ( theo (2.22) ).

Nếu  $t^* = 0$  thì  $s^T R^* u^* \leq 0$  ( theo (2.19) ), suy ra

$$s^T \frac{-g^* - A^*\lambda^*}{\pi^*} \leq 0$$

( theo (2.22) ), do đó

$$s^T (g^* + A^*\lambda^*) \geq 0.$$

Trong hai trường hợp trên nếu  $s \in \mathcal{D}^*$  thì sẽ mâu thuẫn. Vậy  $F^* \cap \mathcal{D}^* = \emptyset$

Ngược lại nếu các điều kiện của định lý không được thoả mãn thì ta sẽ chỉ ra có một hướng  $s \in F^* \cap \mathcal{D}^*$ . Các điều kiện này của định lý tương đương với phát biểu: Nón lồi đóng

$$C = \{ y : t^* < 0 \Rightarrow y = 0; t^* = 0 \Rightarrow y = -\pi R^* u, \forall \pi \geq 0, u \in \partial t^* \}$$

và tập lồi compac

$$B = \{ b : b = g^* + A^*\lambda^*, \forall \lambda \in \partial h^* \}$$

có một điểm chung.

Do đó nếu các điều kiện của định lý không được thoả mãn tức là không có điểm chung giữa hai tập hợp trên thì theo Bổ đề 2.6, tồn tại một hướng  $s$  sao cho

$$\max_{\lambda \in \partial h^*} s^T (g^* + A^*\lambda) < 0.$$

Điều này dẫn đến  $s \in \mathcal{D}^*$  và  $t^* = 0$ . Do đó

$$\max_{u \in \partial t^*} s^T R^* u \leq 0$$

tức là  $s \in F^*$ . Như vậy  $s \in F^* \cap \mathcal{D}^*$ .

Khi đó định lý chính là kết quả của Bổ đề 2.6 và giả thiết  $\mathcal{F}^* \cap \mathcal{D}^* = F^* \cap \mathcal{D}^*$ .

### 2.4.2 Điều kiện tối ưu cấp hai

Để nghiên cứu điều kiện tối ưu cấp hai, ta phải giả thiết thêm rằng  $c(x)$  và  $r(x)$  là các hàm thuộc lớp  $\mathbb{C}^2$  nhưng  $h(c)$  và  $t(r)$  vẫn là các hàm lồi thuộc lớp  $\mathbb{C}^0$ .

Cho  $x^*, \lambda^*, u^*, \pi^*$  thoả mãn các điều kiện của Định lý 2.6 và xét tập

$$X = \{ x : h(c(x)) = h^* + (c(x) - c^*)^T \lambda^*, \\ t(r(x)) \leq 0, \pi^*(r(x) - r^*)^T u^* = 0 \} \quad (2.23)$$

Định nghĩa  $\mathcal{G}^*$  là tập các định hướng chấp nhận được đã được chuẩn hoá lấy trên tập  $X$  và xét tại  $x^*$

$$\mathcal{G}^* = \{ s : \|s\|_2 = 1, \exists \{x^{(k)}\}, x^{(k)} \in X, \\ x^{(k)} \rightarrow x^*, s^{(k)} \rightarrow s, \delta^{(k)} \downarrow 0 \} \quad (2.24)$$

Tập này liên quan mật thiết với tập

$$G^* = \{ s : \|s\|_2 = 1, s \in F^*, \max_{\lambda \in \partial h^*} s^T (g^* + A^* \lambda) = 0, \\ \pi^* \max_{u \in \partial t^*} s^T R^* u = 0 \} \quad (2.25)$$

$G^*$  có thể được hiểu như là tập các hướng chấp nhận được đối với các ràng buộc tuyến tính tại  $x^*$  có độ dốc 0 và liên quan tới cả hàm  $\phi(x)$  và  $t(r(x))$  (nếu  $\pi^* > 0$ ).

**Bổ đề 2.8.**  $\mathcal{G}^* \subseteq G^*$

*Chứng minh.* Lấy  $s \in \mathcal{G}^*$ , suy ra  $s \in \mathcal{F}^*$  và do đó  $s \in F^*$ . Theo Bổ đề

2.1 thì  $\max_{\lambda \in \partial h^*} s^T(g^* + A^*\lambda) = 0$ .

Bằng lập luận tương tự nếu  $\pi^* > 0$  ( và  $t^* = 0$  ) thì

$$0 = \lim \frac{(r^{(k)} - r^*)^T u^*}{\delta^{(k)}} = s^T R^* u^*.$$

Vì  $s \in F^*$  nên từ (2.19) ta có  $\pi^* \max_{u \in \partial t^*} s^T R^* u = 0$ . Vậy  $s \in G^*$ .  $\square$

**Nhận xét 2.1.** Một cách tổng quát thì không phải bao giờ cũng có hệ thức ngược lại mà nó chỉ xảy ra trong một số trường hợp đặc biệt liên quan tới các hàm tuyến tính địa phương. Điều kiện  $\mathcal{G}^* = G^*$  được gọi là điều kiện chính quy.

Bổ đề sau đây sẽ cho biết khi nào điều kiện chính quy xảy ra:

**Bổ đề 2.9. (Điều kiện chính quy)**

*Nếu  $x^*$  thoả mãn các điều kiện cấp một của Định lý 2.6, nếu  $h(c), t(r)$  tuyến tính địa phương tại  $c^*$  và  $r^*$ , và*

$$\text{rank}([A^*D^* : R^*u^* : R^*H^*]) = l^* + q^* + 1 \quad (2.26)$$

*thì  $\mathcal{G}^* = G^*$ . Giả thiết về hạng có thể được thay thế bởi giả thiết rằng các hàm  $c(x)$  và  $r(x)$  là các hàm affine, ở đây  $D^*$  là ma trận có các cột là  $d_i^*, i = \overline{1, l^*}$  với  $d_i^*$  là cơ sở của  $\partial h(c^*) - \lambda^*$ , còn  $l^*$  là số chiều của  $\partial h(c^*)$ .*

*Chứng minh.* Theo Bổ đề 2.8 thì  $\mathcal{G}^* \subseteq G^*$ . Ta sẽ chứng minh chiều ngược lại.

Lấy  $s \in G^*$ , ta định nghĩa tập

$$\partial h_s^* = \{ \lambda \in \partial h^* : s^T(g^* + A^*\lambda) = 0 \}$$

và giả sử số chiều của  $\partial h_s^*$  là  $l_s^*$  ( $l_s^* < l$ ).

Nếu  $t^* < 0$  hoặc  $t^* = 0, \pi^* = 0$  và  $\max_{u \in \partial t^*} s^T R^* u < 0$  thì sẽ xây

dựng được một hàm trơn  $arc\ x(\theta)$ ,  $\theta \in [0, \bar{\theta})$  sao cho  $x(0) = x^*$  và  $\dot{x}(0) = s$ . Do đó ta có

$$(c(x(\theta)) - c^*)^T d_i^* = \theta s^T A^* d_i \begin{cases} = 0, & i = 1, 2, \dots, l_s^* \\ < 0, & i = l_s^* + 1, \dots, l^* \end{cases}$$

và do đó

$$(c(x(\theta)) - c^*)^T (\lambda - \lambda^*) \begin{cases} = 0, & \lambda \in \partial h_s^* \\ < 0, & \lambda \in \partial h^* \setminus \partial h_s^* \end{cases}$$

hay

$$\max_{\lambda \in \partial h^*} (c(x(\theta)) - c^*)^T (\lambda - \lambda^*) = 0.$$

Sử dụng giả thiết  $h(c)$  là tuyến tính địa phương tại  $c^*$  nên tồn tại một lân cận của  $c^*$  sao cho

$$h(c(x(\theta))) = h(c^*) + (c(x(\theta)) - c^*)^T \lambda^*$$

và lấy bất kì dãy  $\theta^{(k)} \downarrow 0$  sẽ cho ta một dãy định hướng chấp nhận được và do đó  $s \in \mathcal{G}^*$ .

Nếu  $\pi^* = 0$  và  $\max_{u \in \partial t^*} s^T R^* u = 0$  thì không mất tính tổng quát giả sử  $u^*$  là phần tử đạt max, do đó  $s^T R^* u^* = 0$ .

Định nghĩa:

$$\partial h_s^* = \{ \lambda \in \partial h^* : s^T (g^* + A^* \lambda) = 0 \}$$

$$\partial t_s^* = \{ u \in \partial t^* : s^T R^* u = 0 \}$$

phụ thuộc vào  $s$ . Bây giờ  $s \in G^*$  và sử dụng điều kiện cấp một dẫn tới

$$s^T (g^* + A^* \lambda^*) = \pi^* s^T R^* u^* = 0.$$

Do đó

$$s^T A^* (\lambda - \lambda^*) = 0 \quad \forall \lambda \in \partial h_s^*$$

$$s^T R^* u = 0 \quad \forall u \in \partial t_s^*$$

và từ (2.25) ta có

$$s^T A^*(\lambda - \lambda^*) < 0 \quad \forall \lambda \in \partial h^* \setminus \partial h_s^*$$

$$s^T R^* u < 0 \quad \forall u \in \partial t^* \setminus \partial t_s^*.$$

Số chiều của  $\partial h_s^*$  và  $\partial t_s^*$  lần lượt là  $l_s^*$ ,  $q_s^*$  và lấy  $u^*$  là véctơ tùy ý giống như  $u'$  trong (2.20). Không mất tính tổng quát giả sử rằng các véctơ  $d_i^*, i = \overline{1, l_s^*}$  và  $f_i^*, i = \overline{1, q_s^*}$  lập thành cơ sở của  $\partial h_s^* - \lambda^*$  và  $\partial t_s^* - u^*$ . Khi đó

$$\begin{aligned} s^T A^* d_i^* & \begin{cases} = 0, & i = 1, 2, \dots, l_s^* \\ < 0, & i = l_s^* + 1, \dots, l^*, \end{cases} \\ s^T R^* u^* & = 0, \\ s^T R^* f_i^* & \begin{cases} = 0, & i = 1, 2, \dots, q_s^* \\ < 0, & i = q_s^* + 1, \dots, q^*. \end{cases} \end{aligned}$$

Nếu  $l_s^* + q_s^* + 1 = n$  thì từ giả thiết về hạng của ma trận dẫn tới  $s = 0$  mâu thuẫn với  $s \in G^*$ . Với  $l_s^* + q_s^* + 1 < n$  thì ta sẽ xây dựng được một hàm tron

$$arc \, x(\theta), \theta \in [0, \bar{\theta}) \text{ với } x(0) = x^*, \dot{x}(0) = s$$

và

$$\begin{aligned} (c(x(\theta)) - c^*)^T d_i^* & = \theta s^T A^* d_i^*, i = 1, 2, \dots, l_s^* \\ (r(x(\theta)) - r^*)^T u^* & = \theta s^T R^* u^* \\ (r(x(\theta)) - r^*)^T f_i^* & = \theta s^T R^* f_i^*, i = 1, 2, \dots, q_s^*. \end{aligned} \tag{2.27}$$

Từ đó dẫn tới

$$\max_{\lambda \in \partial h^*} (c(x(\theta)) - c^*)^T (\lambda - \lambda^*) = 0 \tag{2.28}$$

và

$$\max_{u \in \partial t^*} (r(x(\theta)) - r^*)^T u = 0, \tag{2.29}$$

từ giả thiết về tính tuyến tính địa phương ta có

$$\begin{aligned} h(c(x(\theta))) &= h^* + (c(x(\theta)) - c^*)^T \lambda^* \\ t(r(x(\theta))) &= 0. \end{aligned}$$

Cộng vào (2.27) phương trình

$$\pi^*(r(x(\theta)) - r^*)^T u^* = 0 \quad (2.30)$$

dẫn tới  $x(\theta) \in X$  trong (2.28) và do đó bằng cách lấy một dãy bất kì  $\theta^{(k)} \downarrow 0$ ,  $s$  là một hướng chấp nhận được trong  $\mathcal{G}^*$ .

Cuối cùng nếu giả thiết  $c(x)$  và  $r(x)$  là affine thì  $\text{arc } x(\theta) = x^* + \theta s$  có

$$\begin{aligned} c(x(\theta)) &= c^* + A^{*T} s \\ r(x(\theta)) &= r^* + R^{*T} s \end{aligned}$$

và dễ dàng suy ra (2.28), (2.29), (2.30) trực tiếp từ phương trình

$$\begin{aligned} \max_{\lambda \in \partial h^*} s^T A^*(\lambda - \lambda^*) &= 0 \\ \max_{u \in \partial t^*} s^T R^* u &= 0 \\ s^T R^* u^* &= 0. \end{aligned}$$

Bổ đề được chứng minh xong. □

Bây giờ ta sẽ nghiên cứu các điều kiện cần và đủ cấp hai cho bài toán tối ưu không trơn có ràng buộc (2.17) ở trên

### **Định lý 2.7. (Điều kiện cần cấp hai)**

*Nếu  $x^*$  là cực tiểu địa phương của bài toán (2.17) và  $\mathcal{F}^* \cap \mathcal{D}^* = F^* \cap \mathcal{D}^*$  thì Định lý 2.6 được thoả mãn. Khi đó với mỗi bộ ba  $\lambda^*, u^*, \pi^*$ , nếu  $\mathcal{G}^* = G^*$  thì*

$$s^T \nabla^2 L(x^*, \lambda^*, u^*, \pi^*) s \geq 0, \quad \forall s \in G^*. \quad (2.31)$$

*Chứng minh.* Lấy  $s \in G^*$ . Khi đó  $s \in \mathcal{G}^*$  và tồn tại một dãy định hướng chấp nhận được trong tập  $X$  được xác định ở (2.23). Khai triển Taylor hàm  $L(x, \lambda^*, u^*, \pi^*)$  tại  $x^*$  ta được

$$\begin{aligned} & L(x^{(k)}, \lambda^*, u^*, \pi^*) \\ &= L(x^*, \lambda^*, u^*, \pi^*) + e^{(k)T} \nabla L(x^*, \lambda^*, u^*, \pi^*) \\ & \quad + \frac{1}{2} e^{(k)T} \nabla^2 L(x^*, \lambda^*, u^*, \pi^*) e^{(k)} + o(\|e^{(k)}\|^2) \end{aligned} \quad (2.32)$$

với  $e^{(k)} = x^{(k)} - x^*$ . Theo định nghĩa của  $L$  thì

$$\begin{aligned} & L(x^{(k)}, \lambda^*, u^*, \pi^*) - L(x^*, \lambda^*, u^*, \pi^*) \\ &= f^{(k)} - f^* + (c^{(k)} - c^*)^T \lambda^* + \pi^* (r^{(k)} - r^*)^T u^* \\ &= f^{(k)} - f^* + h(c^{(k)}) - h^* \\ &= \phi^{(k)} - \phi^*. \end{aligned} \quad (2.33)$$

Vì  $x^*$  là cực tiểu địa phương của hàm  $\phi$  và  $e^{(k)} = x^{(k)} - x^* = \delta^{(k)} s^{(k)}$  nên ta có

$$0 \leq \phi^{(k)} - \phi^* = \frac{1}{2} \delta^{(k)2} s^{(k)T} \nabla^2 L(x^*, \lambda^*, u^*, \pi^*) s^{(k)} + o(\delta^{(k)2}).$$

Chia cả hai vế của bất đẳng thức trên cho  $\frac{1}{2} \delta^{(k)2}$  và lấy giới hạn khi  $k \rightarrow \infty$  ta được (2.31).  $\square$

### **Định lý 2.8. (Điều kiện đủ cấp hai)**

*Nếu tại  $x^*$  mà  $t^* \leq 0$  và tồn tại  $\lambda^* \in \partial h^*, u^* \in \partial t^*, \pi^* \geq 0$  sao cho  $\pi^* t^* = 0$  đồng thời (2.22) thoả mãn. Khi đó nếu*

$$s^T \nabla^2 L(x^*, \lambda^*, u^*, \pi^*) s > 0, \quad \forall s \in G^* \quad (2.34)$$

*thì  $x^*$  là cực tiểu địa phương chặt của bài toán (2.17).*

*Chứng minh.* Giả sử ngược lại  $x^*$  không là cực tiểu địa phương chặt, khi

đó tồn tại một dãy chấp nhận được và do đó tồn tại một dãy định hướng chấp nhận được  $x^{(k)} \rightarrow x^*, s^{(k)} \rightarrow s, \|s\|_2 = 1$  sao cho  $\phi^{(k)} \leq \phi^*$ . Lấy  $s \in \mathcal{F}^*$ , suy ra  $s \in F^*$ . Sử dụng (2.22) ta có

$$t^* < 0 \Rightarrow \pi^* = 0 \Rightarrow g^* + A^* \lambda^* = 0$$

và từ  $s \in F^*$  có

$$t^* = 0 \Rightarrow \max_{u \in \partial t^*} s^T R^* u \leq 0.$$

Do đó trong cả hai trường hợp đều dẫn tới

$$0 \leq \max_{\lambda \in \partial h^*} s^T (g^* + A^* \lambda) = \mu.$$

Nếu  $\mu > 0$  thì theo Bổ đề 1.6 có

$$\lim \frac{\phi^{(k)} - \phi^*}{\delta^{(k)}} = \mu > 0.$$

Do đó  $\phi^{(k)} > \phi^*$ , mâu thuẫn với  $\phi^{(k)} \leq \phi^*$ . Vậy  $\mu = 0$ .

Lấy  $s^T (g^* + A^* \lambda^*) < 0$ . Khi đó điều kiện cấp một dẫn tới  $\pi^* > 0, t^* = 0$  và  $s^T R^* u^* > 0$  mâu thuẫn với  $s \in F^*$ . Do đó  $s^T (g^* + A^* \lambda^*) = 0$  và  $\pi^* s^T R^* u^* = 0$ .

Từ  $s \in F^*$  dẫn tới  $\pi^* \max_{u \in \partial t^*} s^T R^* u^* = 0$  và do đó  $s \in G^*$ .

Bây giờ từ (2.23) ta có

$$\begin{aligned} & L(x^{(k)}, \lambda^*, u^*, \pi^*) - L(x^*, \lambda^*, u^*, \pi^*) \\ &= \phi^{(k)} - \phi^* - (h^{(k)} - h^* - (c^{(k)} - c^*)^T \lambda^*) \\ & \quad + \pi^* (t^{(k)} - t^* - (t^{(k)} - t^* - (r^{(k)} - r^*)^T u^*)) \\ & \leq \phi^{(k)} - \phi^* + \pi^* (t^{(k)} - t^*) \leq \phi^{(k)} - \phi^* \end{aligned}$$

(Sử dụng bất đẳng thức dưới vi phân và tính chấp nhận được)

Từ (2.32) và (2.22) ta được

$$0 \geq \phi^{(k)} - \phi^* \geq \frac{1}{2} \delta^{(k)^2} s^{(k)T} \nabla^2 L(x^*, \lambda^*, u^*, \pi^*) s^{(k)} + o(\delta^{(k)^2}).$$

Chia cả hai vế cho  $\frac{1}{2}\delta^{(k)^2}$  và chuyển qua giới hạn thì

$$s^T \nabla^2 L(x^*, \lambda^*, u^*, \pi^*) s \leq 0$$

mâu thuẫn với (2.34). Vậy điều giả sử là sai.

Định lý được chứng minh.  $\square$

**Hệ quả 2.2.** *Nếu đạo hàm theo hướng*

$$\max_{\lambda \in \partial h^*} s^T (g^* + A^* \lambda)$$

*đương với mọi hướng chấp nhận được trong  $F^*$  hay một cách tương đương là nếu  $G^* = \emptyset$  thì điều kiện cấp một là đủ để dẫn tới  $x^*$  là cực tiểu địa phương chặt của bài toán (14.6.1)*

*Chứng minh.* Hệ quả này được suy ra trực tiếp từ Định lý 2.8.  $\square$

Sau đây ta sẽ đưa ra một số ví dụ minh họa cho bài toán tối ưu không trơn có ràng buộc (2.17)

**Ví dụ 2.3.** Xét bài toán

$$\min \|x\|_\infty$$

với điều kiện

$$\|r(x)\|_1 \leq 0.5 \tag{2.35}$$

trong đó  $r : \mathbb{R}^3 \rightarrow \mathbb{R}^4$  được định nghĩa bởi

$$\begin{aligned} r_1 &= x_1^2 + x_2^2 - 1 \\ r_2 &= x_1 x_2 - 0.5 \\ r_3 &= x_1 + \frac{1}{2}(x_2^2 - 1) \\ r_4 &= -x_1 + x_3^2 + 0.54 \end{aligned} \tag{2.36}$$

Bài toán này là một ví dụ của bài toán (2.17) với

$$f(x) = 0, \quad c(x) = x, \quad h(c) = \|c\|_\infty, \quad t(r) = \|r\|_1 - 0.5$$

Tại  $x^* = (0.6, 0.8, 0)^T$ ,  $r^* = (0, -0.02, 0.42, -0.06)^T$ , ta có

$$\|x^*\|_\infty = 0.8 \Rightarrow \partial h^* = \{(0, 1, 0)^T\}$$

do đó nếu  $\lambda^* \in \partial h^*$  thì  $\lambda^* = (0, 1, 0)^T$ .

Ta có

$$\nabla r_1(x) = \begin{pmatrix} 2x_1 \\ 2x_2 \\ 0 \end{pmatrix}, \quad \nabla r_2(x) = \begin{pmatrix} x_2 \\ x_1 \\ 0 \end{pmatrix}$$

$$\nabla r_3(x) = \begin{pmatrix} 1 \\ x_2 \\ 0 \end{pmatrix}, \quad \nabla r_4(x) = \begin{pmatrix} -1 \\ 0 \\ 2x_3 \end{pmatrix}$$

Điều này dẫn tới

$$R^* = \nabla r^{*T} = \begin{pmatrix} 1.2 & 0.8 & 1 & -1 \\ 1.6 & 0.6 & 0.8 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$A^* = I.$$

Khi đó

$$t^* = t(x^*) = |r_1^*| + |r_2^*| + |r_3^*| + |r_4^*| - 0.5 = 0$$

$$\partial t^* = \{ (\alpha, -1, 1, -1) : |\alpha| \leq 1 \}.$$

Bây giờ ta phải tìm  $u^* \in \partial t^*$ ,  $\pi^* \geq 0$  thoả mãn

$$g^* + A^* \lambda^* + \pi^* R^* u^* = 0$$

tức là

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \pi^* \begin{pmatrix} 1.2 & 0.8 & 1 & -1 \\ 1.6 & 0.6 & 0.8 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ -1 \\ 1 \\ -1 \end{pmatrix} = 0$$

$$\Leftrightarrow \begin{cases} (1.2\alpha + 1.2)\pi^* = 0 \\ (1.6\alpha + 0.2)\pi^* + 1 = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} \alpha = -1 \\ \pi^* = \frac{5}{7} \end{cases}$$

Khi đó

$$u^* = (-1, -1, 1, -1)^T.$$

Vậy với

$$\lambda^* = (0, 1, 0)^T, \quad u^* = (-1, -1, 1, -1)^T, \quad \pi^* = \frac{5}{7}$$

thì  $x^*$  thoả mãn điều kiện cần cấp 1.

Tiếp theo ta sẽ xác định tập

$$G^* = \left\{ s : \|s\|_2 = 1, s \in F^*, \max_{\lambda \in \partial h^*} s^T (g^* + A^* \lambda) = 0 \right. \\ \left. \pi^* \max_{u \in \partial t^*} s^T R^* u = 0 \right\}$$

Với  $\lambda \in \partial h^*, u \in \partial t^*$  thì

$$s^T (g^* + A^* \lambda) = (s_1, s_2, s_3)(0, 1, 0)^T = s_2$$

$$s^T R^* u = s_1(1.2\alpha + 1.2) + s_2(1.6\alpha + 0.2).$$

Do đó

$$G^* = \{s : \|s\|_2 = 1, s_2 = 0, s_1 \leq 0\}.$$

Số chiều của  $\partial h^*$  và  $\partial t^*$  tương ứng là  $l^* = 0, q^* = 1$ , cơ sở của  $\partial t^* - u^* = (\alpha - 1, 0, 0, 0)^T$  là  $H^* = (1, 0, 0, 0)^T$ . Do đó ma trận

$$[R^* u^* : R^* H^*] = \begin{pmatrix} 0 & 1.2 \\ -1.4 & 1.6 \\ 0 & 0 \end{pmatrix}$$

có hạng là  $2 = l^* + q^* + 1$ , vì vậy điều kiện chính quy  $\mathcal{G}^* = G^*$  được thoả mãn.

Xét hàm Lagrange cho bài toán:

$$\begin{aligned} L(x, \lambda, u, \pi) &= f(x) + \sum \lambda_i c_i(x) + \pi \sum u_j r_j(x) \\ &= \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 + \pi [u_1(x_1^2 + x_2^2 - 1) \\ &\quad + u_2(x_1 x_2 - 0.5) + u_3(x_1 + \frac{1}{2}(x_2^2 - 1)) \\ &\quad + u_4(-x_1 + x_3^2 + 0.54)] \end{aligned}$$

Ta có

$$\nabla L(x, \lambda, u, \pi) = \begin{pmatrix} \lambda_1 + 2\pi u_1 x_1 + \pi u_2 x_2 + \pi u_3 - \pi u_4 \\ \lambda_2 + 2\pi u_1 x_2 + \pi u_2 x_1 + \pi u_3 x_2 \\ \lambda_3 + 2\pi u_4 x_3 \end{pmatrix}$$

và

$$\nabla^2 L(x, \lambda, u, \pi) = \begin{pmatrix} 2\pi u_1 & \pi u_2 & 0 \\ \pi u_2 & 2\pi u_1 + \pi u_3 & 0 \\ 0 & 0 & 2\pi u_4 \end{pmatrix}$$

Điều này dẫn tới

$$\nabla^2 L(x^*, \lambda^*, u^*, \pi^*) = \begin{pmatrix} -2 & -1 & 0 \\ -1 & -1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

Do đó

$$\begin{aligned} s^T \nabla^2 L^* s &= \frac{-5}{7} s^T \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} s \\ &= \frac{-5}{7} (2s_1^2 + 2s_1s_2 + s_2^2 + s_3^2) < 0, \quad \forall s \in G^*. \end{aligned}$$

Điều này chứng tỏ  $x^*$  không thoả mãn điều kiện cần cấp hai, do đó  $x^*$  không phải là nghiệm.

Trên thực tế, lời giải của bài này là

$$x_1^* = x_2^* = \frac{1 + \sqrt{6}}{5}, \quad x_3^* = \sqrt{x_1^* - 0.54} = 0.387167.$$

Nhân tử  $\lambda^* = (0.4367, 0.563299)^T$ ,  $u^* = (-1, -1, 1, 0)^T$ ,  $\pi^* = 0.408247$  thoả mãn điều kiện cấp một. Vì  $\lambda^*$  và  $u^*$  liên quan tới phần trong của  $\partial h^*$  và  $\partial t^*$ ,  $l^* + q^* + 1 = 3$  và vì điều kiện về hạng trong (2.26) được thoả mãn. Do đó  $\mathcal{G}^* = G^* = \emptyset$  và điều kiện cấp một đủ để chỉ ra  $x^*$  là nghiệm.

**Ví dụ 2.4.** Xét bài toán

$$\min \|c(x)\|_1$$

với điều kiện  $\|x\|_\infty \leq 0.8$  trong đó  $c : \mathbb{R}^3 \longrightarrow \mathbb{R}^4$  được định nghĩa bởi

$$c_1 = x_1^2 + x_2^2 - 1$$

$$c_2 = x_1x_2 - 0.5$$

$$c_3 = x_1 + \frac{1}{2}(x_2^2 - 1)$$

$$c_4 = -x_1 + x_3^2 + 0.54$$

Chúng minh  $x^* = (0.6, 0.8, 0)^T$  thoả mãn điều kiện cấp 1 nhưng không thoả mãn điều kiện cần cấp 2.

*Giải.* Từ giả thiết ta có

$$\begin{aligned} f(x) &= 0, \quad h(c) = \|c\|_1, \\ r(x) &= r, \quad t(r) = \|r\|_\infty - 0.8 \end{aligned}$$

Sử dụng kết quả tính toán từ ví dụ (2.3) ta được

$$\begin{aligned} A^* &= \nabla c^{*T} = \begin{pmatrix} 1.2 & 0.8 & 1 & -1 \\ 1.6 & 0.6 & 0.8 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\ R^* &= I. \end{aligned}$$

Từ

$$c^* = c(x^*) = (0, -0.02, 0.42, -0.06)$$

dẫn tới

$$\partial h^* = \partial \|c^*\|_1 = \{ (\lambda, -1, 1, -1) : |\lambda| \leq 1 \}.$$

Mặt khác  $\|x^*\|_\infty = 0.8$  nên

$$\partial t^* = \partial \|x^*\|_\infty = \{(0, 1, 0)^T\} \text{ và } t^* = 0$$

Do đó nếu  $u^* \in \partial t^*$  thì  $u^* = (0, 1, 0)^T$  và nếu  $\lambda^* \in \partial h^*$  thì  $\lambda^* = (\lambda, -1, 1, -1)^T$  trong đó  $|\lambda| \leq 1$ .

Xét điều kiện cấp 1:

$$g^* + A^* \lambda^* + \pi^* R^* u^* = 0$$

Hay

$$\begin{pmatrix} 1.2 & 0.8 & 1 & -1 \\ 1.6 & 0.6 & 0.8 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \lambda \\ -1 \\ 1 \\ -1 \end{pmatrix} + \pi^* \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 0$$

Điều này dẫn tới

$$\begin{cases} 1.2(\lambda + 1) = 0 \\ 1.6\lambda + 0.2 + \pi^* = 0 \end{cases}$$

Do đó  $\lambda = -1$ ,  $\pi^* = \frac{7}{5}$ .

Vậy với  $u^* = (0, 1, 0)^T$ ,  $\lambda^* = (-1, -1, 1, -1)^T$  và  $\pi^* = \frac{7}{5}$  thì  $x^*$  thoả mãn điều kiện cấp 1.

Bây giờ ta kiểm tra điều kiện cần cấp hai đối với  $x^*$ . Tương tự như ví dụ 2.3 thì

$$G^* = \{ s : \|s\|_2 = 1, s_2 = 0, s_1 \leq 0 \}.$$

Số chiều của  $\partial h^*$  và  $\partial t^*$  tương ứng là  $l^* = 1, q^* = 0$ , cơ sở của  $\partial h^* - \lambda^*$  là  $D^* = (1, 0, 0, 0)^T$ . Do đó ma trận

$$[A^* D^* : R^* u^*] = \begin{pmatrix} 1.2 & 0 \\ 1.6 & -1.4 \\ 0 & 0 \end{pmatrix}$$

có hạng là  $2 = l^* + q^* + 1$ , vì vậy điều kiện chính quy  $\mathcal{G}^* = G^*$  được thoả mãn.

Xét hàm Lagrange cho bài toán:

$$\begin{aligned} L(x, \lambda, u, \pi) \\ = \lambda_1(x_1^2 + x_2^2 - 1) + \lambda_2(x_1 x_2 - 0.5) + \lambda_3(x_1 + \frac{1}{2}(x_2^2 - 1)) \\ + \lambda_4(-x_1 + x_3^2 + 0.54) + \pi(u_1 x_1 + u_2 x_2 + u_3 x_3) \end{aligned}$$

Ta có

$$\nabla L(x, \lambda, u, \pi) = \begin{pmatrix} 2\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 - \lambda_4 + \pi u_1 \\ 2\lambda_1 x_2 + \lambda_2 x_1 + \lambda_3 x_2 + \pi u_2 \\ 2\lambda_4 x_3 + \pi u_3 \end{pmatrix}$$

và

$$\nabla^2 L(x, \lambda, u, \pi) = \begin{pmatrix} 2\lambda_1 & \lambda_2 & 0 \\ \lambda_2 & 2\lambda_1 + \lambda_3 & 0 \\ 0 & 0 & 2\lambda_4 \end{pmatrix}$$

Do đó

$$\nabla^2 L(x^*, \lambda^*, u^*, \pi^*) = \begin{pmatrix} -2 & -1 & 0 \\ -1 & -1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

Vậy

$$\begin{aligned} s^T \nabla^2 L^* s &= \frac{-7}{5} s^T \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} s \\ &= \frac{-7}{5} (2s_1^2 + 2s_1s_2 + s_2^2 + s_3^2) < 0, \quad \forall s \in G^*. \end{aligned}$$

Điều này chứng tỏ  $x^*$  không thoả mãn điều kiện cần cấp 2 nên  $x^*$  không là nghiệm của bài toán.

# Kết luận

Luận văn "Dưới vi phân của hàm lồi và ứng dụng trong tối ưu hóa không trơn" đã tập trung nghiên cứu một số vấn đề sau đây:

1. Trình bày khái niệm dưới vi phân cùng với một số tính chất cơ bản và các phép toán của nó.
2. Phát biểu hai bài toán tối ưu không trơn: bài toán không ràng buộc và có ràng buộc.
3. Chứng minh chi tiết các điều kiện tối ưu cấp 1 và cấp 2 cho hai loại bài toán trên và đưa ra mối liên hệ với bài toán tối ưu trơn.
4. Xây dựng các ví dụ minh họa cho các bài toán tối ưu không trơn.

Do khuôn khổ luận văn thạc sĩ có hạn và điều kiện thời gian không cho phép nên còn nhiều ý tưởng hay mà tác giả chưa thể thực hiện được một cách đầy đủ. Chắc chắn rằng, trong thời gian tới, tác giả sẽ tập trung tìm hiểu kỹ lưỡng hơn những vấn đề còn chưa được thực hiện.

Luận văn này được hoàn thành dưới sự hướng dẫn của GS. TS. Trần Vũ Thiệu và sự nghiên cứu, làm việc nghiêm túc của bản thân. Tác giả xin bày tỏ lòng biết ơn sâu sắc đến sự quan tâm hướng dẫn nhiệt tình nhưng hết sức nghiêm khắc của thầy.

Tác giả xin chân thành cảm ơn Ban giám hiệu, Phòng NCKH-ĐTSDH, Khoa Toán - Cơ - Tin Trường Đại học Khoa học Tự nhiên - Đại học Quốc gia Hà Nội đã giúp đỡ, động viên, tạo mọi điều kiện thuận lợi cho tác giả trong quá trình học tập và hoàn thành bản luận văn này.

# Tài liệu tham khảo

- [1] Nguyễn Thị Bạch Kim (2008), *Giáo trình các phương pháp tối ưu lý thuyết và thuật toán*, NXB Bách khoa Hà Nội.
- [2] Đỗ Văn Lưu, Phan Huy Khải (2000), *Giải tích lồi*, NXB Khoa học và kỹ thuật Hà Nội.
- [3] Nguyễn Xuân Tấn, Nguyễn Bá Minh (2007), *Lý thuyết tối ưu không trơn*, NXB Đại học Quốc gia Hà Nội.
- [4] R.Fletcher (2006), *Practical methods of optimization*, University of Dundee, Scotland, UK .
- [5] Manio Gandioso.(2002), “ Nonsmooth optimization in Handbook of applied optimization”, *Oxford univ*, pp.299-309.
- [6] Jean-Jacques strodriot (2003), *Introduction to nonsmooth optimization*, Belgium.
- [7] G.G.Magaril-II'yaev and V.M.Tikhomirov (2000), *Convex analysis: Theory and Applications*, America.
- [8] N.Z.Shor (1985), *Minimization Methods for non-differentiable functions*, Springer - Verlag.