

Chương 9

TÍNH CHẤT CỰC TRỊ, BẤT ĐẲNG THỨC LỒI VÀ ĐỊNH LÝ HELLEY

9.1. Cực đại và cực tiểu của hàm lồi	128
9.2. Hạng của hàm lồi	135
9.3. Bất đẳng thức lồi	139
9.4. Định lý Helley	143
9.5. Bài tập	148

Cực trị (cực đại, cực tiểu) của một hàm lồi trên một tập lồi có những tính chất riêng, lý thú. Sự nghiên cứu về tính chất cực trị của một hàm lồi trên một tập lồi là một đề tài quan trọng của lý thuyết tối ưu. Chương này trước hết giới thiệu một số tính chất chung về cực tiểu và cực đại của một hàm lồi trên một tập lồi. Tiếp theo trình bày về hạng của hàm lồi, đại lượng đo "mức độ phi tuyến" của nó. Tính chất cực trị của hàm lồi có liên quan chặt chẽ với các bất đẳng thức lồi sẽ được đề cập trong phần tiếp theo. Cuối chương là các định lý Helley về sự tương giao của các tập lồi.

9.1. Cực đại và cực tiểu của hàm lồi

Trong nhiều vấn đề ứng dụng ta thường gặp bài toán tìm cực tiểu hoặc cực đại của một hàm lồi trên một tập lồi. Hai bài toán này có những tính chất cơ bản rất khác nhau. Tuy nhiên tính chất lồi kéo theo những đặc thù riêng cho mỗi bài toán. Lợi dụng các tính chất này, người ta đã đưa ra được những phương pháp giải quyết khác nhau cho mỗi bài toán kể trên.

Định nghĩa 9.1. Cho $C \subseteq \mathbb{R}^n$ khác rỗng và $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$. Một 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

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

Điểm $x^* \in C$ được gọi là *cực đại địa phương* nếu

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

Nếu

$$f(x) \geq f(x^*) \quad \forall x \in C,$$

thì x^* được gọi là *cực tiểu toàn cục* hay *cực tiểu tuyệt đối* của f trên C , và nếu

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

thì x^* được gọi là *cực đại toàn cục* hay *cực đại tuyệt đối* của f trên C .

Cực tiểu hàm lồi. Mệnh đề dưới đây cho thấy mọi điểm cực tiểu địa phương của một hàm lồi trên một tập lồi cũng chính là điểm cực tiểu tuyệt đối. Điều này rất quan trọng vì nó cho phép sử dụng các công cụ mang tính địa phương như phép tính vi phân trong việc xây dựng lý thuyết và các phương pháp giải cho bài toán này.

Mệnh đề 9.1. Cho $f : \mathbf{R}^n \rightarrow \mathbf{R} \cup \{+\infty\}$ lồi. Khi đó mọi điểm cực tiểu địa phương của f trên một tập lồi đều là cực tiểu toàn cục. Hơn nữa tập hợp các điểm cực tiểu của f là một tập lồi. Nếu f lồi chặt, thì điểm cực tiểu, nếu tồn tại, sẽ duy nhất.

Chứng minh. Cho $C \subseteq \mathbf{R}^n$. Giả sử x^* là điểm cực tiểu địa phương của f trên C . Khi đó tồn tại lân cận U của x^* sao cho

$$f(x) \geq f(x^*) \quad \forall x \in U \cap C.$$

Với mọi $x \in C$, và $0 < \lambda < 1$, do C lồi và U là lân cận của $x^* \in C$, nên điểm $x_\lambda := (1 - \lambda)x^* + \lambda x \in C \cap U$ khi λ đủ nhỏ. Do $f(x^*) \leq f(x_\lambda)$ và f lồi, ta có

$$f(x^*) \leq f(x_\lambda) \leq (1 - \lambda)f(x^*) + \lambda f(x).$$

Từ đây suy ra $f(x^*) \leq f(x)$. Chứng tỏ x^* là cực tiểu toàn cục của f trên C .

Giả sử $x^*, y^* \in C$ là các điểm cực tiểu của f trên C . Vậy $f(x^*) = f(y^*) \leq f(x)$ với mọi $x \in C$. Lấy $z^* := \lambda x^* + (1 - \lambda)y^*$, với $0 < \lambda < 1$. Do C lồi, nên $z^* \in C$ và do f lồi, nên

$$f(z^*) \leq \lambda f(x^*) + (1 - \lambda)f(y^*) \leq f(x).$$

Suy ra z^* cũng là điểm cực tiểu của f trên C . Chứng tỏ tập các điểm cực tiểu của f trên C là lồi. Để thấy rằng tập hợp này chỉ gồm nhiều nhất một điểm khi f lồi chặt. $\square$

Mệnh đề trên mang tính chất định tính. Mệnh đề dưới đây mang nhiều tính chất định lượng. Ta hãy xét bài toán tìm cực tiểu một hàm lồi trên một tập lồi có dạng sau:

$$(OP) \begin{cases} \min f(x) \text{ với các điều kiện} \\ g_i(x) \leq 0, i = 1, \dots, m, \\ h_j(x) = 0, j = 1, \dots, k \\ x \in X \end{cases}$$

trong đó $X \subseteq \mathbf{R}^n$ là một tập lồi đóng khác trống và f, g_i ($i = 1, \dots, m$) là các hàm lồi hữu hạn trên X , còn h_j ($j = 1, \dots, k$) là các hàm a-phân hữu hạn trên tập a-phân của X . Ta sẽ luôn giả sử rằng X có điểm trong và các hàm a-phân h_j ($j = 1, \dots, k$) độc lập tuyến tính trên X , theo nghĩa, nếu $\sum_{j=1}^k \mu_j h_j(x) = 0$ với mọi $x \in X$, thì $\mu_j = 0$ với mọi j .

Bài toán (OP) này được gọi là một quy hoạch lồi. Hàm f được gọi là hàm mục tiêu. Các điều kiện $x \in X, g_i(x) \leq 0$ ($i = 1, \dots, m$), $h_j(x) = 0$, ($j = 1, \dots, k$) được gọi là các ràng buộc. Tập

$$D := \{x \in X \mid g_i(x) \leq 0 \ i = 1, \dots, m, h_j(x) = 0, j = 1, \dots, k\}$$

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 ($i = 1, \dots, m$) lồi trên X và h_j ($j = 1, \dots, k$) a-phân, 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):

$$L(x, \lambda, \mu) := \lambda_0 f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{j=1}^k \mu_j h_j(x).$$

Dựa vào hàm Lagrange, ta có kết quả sau:

Định lý 9.1 (Karush-Kuhn-Tucker). Nếu x^* là nghiệm của bài toán quy hoạch lồi (OP), thì tồn tại $\lambda_i^* \geq 0$ ($i = 0, 1, \dots, m$) và μ_j^* ($j = 1, \dots, k$) không đồng thời bằng 0 sao cho

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

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

Hơn nữa nếu $\text{int} X \neq \emptyset$ và điều kiện Slater sau thoả mãn

$$\exists x^0 \in D : g_i(x^0) < 0 \ (i = 1, \dots, m)$$

thì $\lambda_0^* > 0$ và hai điều kiện đạo hàm triệt tiêu và độ lệch bù ở trên, cũng là điều kiện đủ để điểm chấp nhận được x^* là nghiệm tối ưu của bài toán (OP).

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

$$C := \{(\lambda_0, \lambda_1, \dots, \lambda_m, \mu_1, \dots, \mu_k) \mid (\exists x \in X) :$$

$$f(x) - f(x^*) < \lambda_0, g_i(x) \leq \lambda_i, i = 1, \dots, m, h_j(x) = \mu_j, j = 1, \dots, k\}$$

Do $X \neq \emptyset$ lồi, f, g_i lồi và h_j a-phân hữu hạn trên X , nên C là một tập lồi, khác trống trong $\mathbb{R}^{m+k+1}$. Hơn nữa $0 \notin C$. Thật vậy, vì nếu trái lại $0 \in C$, thì tồn tại một điểm chấp nhận được x thỏa mãn $f(x) < f(x^*)$. Điều này mâu thuẫn với việc x^* là nghiệm tối ưu của (OP).

Khi đó theo định lý tách 1, tồn tại λ_i^* ($i = 0, 1, \dots, m$), μ_j^* ($j = 1, \dots, k$) không đồng thời bằng 0 sao cho

$$\sum_{i=0}^m \lambda_i^* \lambda_i + \sum_{j=1}^k \mu_j^* \mu_j \geq 0 \quad \forall (\lambda_0, \dots, \lambda_m, \mu_1, \dots, \mu_k) \in C. \quad (9.1)$$

Chú ý rằng với mọi $\lambda_0, \dots, \lambda_m > 0$, thì $(\lambda_0, \dots, \lambda_m, 0, \dots, 0) \in C$, vì theo định nghĩa của C ta chỉ lấy $x = x^*$. Từ chú ý này, ta suy ra ngay tất cả $\lambda_0^*, \lambda_1^*, \dots, \lambda_m^* \geq 0$. Hơn nữa, với mọi $\epsilon > 0$ và $x \in X$, ta lấy $\lambda_0 = f(x) - f(x^*) + \epsilon$, $\lambda_i = g_i(x)$ ($i = 1, \dots, m$), $\mu_j = h_j(x)$ ($j = 1, \dots, k$) rồi thay vào (9.1) và cho $\epsilon \rightarrow 0$, sẽ được

$$\begin{aligned} \lambda_0^* f(x) + \sum_{i=1}^m \lambda_i^* g_i(x) + \sum_{i=1}^k \mu_i^* h_i(x) &\geq \\ \lambda_0^* f(x^*) + \sum_{i=1}^m \lambda_i^* g_i(x^*) + \sum_{i=1}^k \mu_i^* h_i(x^*) &\quad \forall x \in X. \end{aligned}$$

Hay

$$L(x^*, \lambda^*, \mu^*) \leq L(x, \lambda^*, \mu^*) \quad \forall x \in X.$$

Đây chính là điều kiện đạo hàm triệt tiêu.

Để chứng minh điều kiện độ lệch bù, ta chú ý rằng do x^* chấp nhận được, nên $g_i(x^*) \leq 0$ với mọi i . Nếu như tồn tại một i nào đó mà $g_i(x^*) = \zeta < 0$, thì với mọi $\epsilon > 0$, ta có

$$(\epsilon, \dots, \zeta, \epsilon, \dots, \epsilon, 0, \dots, 0) \in C \quad (\zeta \text{ ở vị trí thứ } i+1)$$

Thay vào (9.1) và cho $\epsilon \rightarrow 0$, ta thấy $\lambda_i^* \zeta \geq 0$. Nhưng $\zeta < 0$, nên $\lambda_i^* \leq 0$. Suy ra $\lambda_i^* = 0$. Điều kiện độ lệch bù do đó cũng được thoả mãn.

Để chứng minh điều kiện đủ, ta giả sử điều kiện Slater thoả mãn. Ta có $\lambda_0^* > 0$. Thật vậy, vì nếu $\lambda_0^* = 0$, thì do điều kiện đạo hàm triệt tiêu và điều kiện độ lệch bù, ta có

$$0 = \sum_{i=1}^m \lambda_i^* g_i(x^*) + \sum_{j=1}^k \mu_j^* h_j(x^*) \leq \sum_{i=1}^m \lambda_i^* g_i(x) + \sum_{j=1}^k \mu_j^* h_j(x) \quad \forall x \in X.$$

Thế nhưng do $\lambda_0^* = 0$, nên phải có hoặc $\lambda_i^* > 0$ với một i nào đó, hoặc nếu $\lambda_i^* = 0$ với mọi i , thì sẽ có $\mu_j^* > 0$ với một j nào đó. Trong trường hợp đầu, thay x^0 vào bất đẳng thức trên, sẽ được

$$0 = \sum_{i=1}^m \lambda_i^* g_i(x^*) + \sum_{j=1}^k \mu_j^* h_j(x^*) \leq \sum_{i=1}^m \lambda_i^* g_i(x^0) + \sum_{j=1}^k \mu_j^* h_j(x^0) < 0.$$

Mâu thuẫn.

Trong trường hợp sau, ta có

$$0 = \sum_{j=0}^k \mu_j^* h_j(x^*) \leq \sum_{j=0}^k \mu_j^* h_j(x) \quad \forall x \in X.$$

Do $\text{int}X \neq \emptyset$ và h_j a-phin với mọi j , nên từ đây suy ra $\sum_{j=0}^k \mu_j^* h_j(x) = 0 \quad \forall x \in X$. Từ đây và do các hàm h_j độc lập tuyến tính trên X , ta có $\mu_j^* = 0$ với mọi j . Điều này mâu thuẫn với việc tất cả các nhân tử λ_i^* và μ_j^* không đồng thời bằng 0. Vậy $\lambda_0^* > 0$.

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

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{j=1}^k \mu_j h_j(x)$$

Do điều kiện đạo hàm triệt tiêu và độ lệch bù, nên với mọi x chấp nhận được, ta có:

$$\begin{aligned} f(x^*) &= f_0(x^*) + \sum_{i=1}^m \lambda_i g_i(x^*) + \sum_{j=1}^k \mu_j h_j(x^*) \\ &\leq f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{j=1}^k \mu_j h_j(x) \leq f(x) \end{aligned}$$

Chứng tỏ x^* là lời giải tối ưu của (OP). $\square$

Chú ý. Khi X là tập mở (nói riêng là toàn không gian) và mọi hàm đều khả vi, thì điều kiện đạo hàm triệt tiêu sẽ là

$$0 = \lambda_0^* \nabla f_0(x^*) + \sum_{j=1}^m \lambda_j^* \nabla g_j(x^*) + \sum_{i=1}^k \mu_i^* \nabla h_i(x^*).$$

Cực đại hàm lồi Các tính chất cực đại của một hàm lồi khác hẳn các tính chất về cực tiểu của nó. Cụ thể ta thấy rằng cực đại địa phương của một hàm lồi không nhất thiết là cực đại tuyệt đối. Ví dụ hàm $f(x) = x^2$ có điểm cực đại địa phương trên đoạn $[-1, 2]$ là $x = -1$, nhưng điểm cực đại tuyệt đối lại là $x = 2$. Nếu xét hàm này trên đoạn $[-2, 2]$ ta thấy tập các điểm cực đại tuyệt đối của nó trên đoạn này là không lồi vì nó chỉ gồm hai điểm -2 và 2 . Dưới đây, nếu không nói gì thêm, ta luôn hiểu cực đại là cực đại tuyệt đối.

Từ định nghĩa hàm lồi, ta thấy rằng, nếu f lồi chính thường trên một tập lồi C và $a, b \in C$, thì với mọi x thuộc đoạn (a, b) , tức là $x = \lambda a + (1 - \lambda)b$, $0 < \lambda < 1$, ta có

$$f(x) \leq \lambda f(a) + (1 - \lambda)f(b) \leq \max\{f(a), f(b)\}.$$

Từ đây suy ra rằng cực đại của một hàm lồi f trên một đoạn $[a, b]$ đạt tại đầu mút của đoạn đó. Một cách tổng quát ta có:

- Mệnh đề 9.2.** (i) Giả sử f là một hàm lồi chính thường trên $\mathbb{R}^n$ và $C \subseteq \mathbb{R}^n$ là một tập lồi. Khi đó nếu f đạt cực đại trên C tại một điểm trong tương đối của C , thì f là hằng số trên C .
- (ii) Nếu f là một hàm lồi, chính thường trên $\mathbb{R}^n$ và bị chặn trên trong một tập a -phân, thì nó là hằng số trên tập này.

Chứng minh. (i) Giả sử $a \in \text{ri}C$ là điểm tại đó f đạt cực đại của nó trên C . Theo tính chất của điểm trong tương đối, nên với mọi $x \in C$, đều tồn tại $y \in C$ sao cho $a \in (x, y)$. Do $f(x) \leq f(a)$, $f(y) \leq f(a)$, và f lồi, theo nhận xét ở trên, ta suy ra $f(x) = f(a)$.

(ii) Nếu f không là hằng số trên tập a -phân M , có nghĩa là tồn tại $a, b \in M$ sao cho $f(a) < f(b)$. Mọi điểm x thuộc nửa đường thẳng xuất phát từ a và có hướng $b - a$ đều có dạng $x = a + \lambda(b - a)$ với $\lambda > 0$. Khi đó

$$b = \frac{1}{\lambda}x + \frac{\lambda - 1}{\lambda}a.$$

Với mọi $\lambda > 1$, theo tính lồi của f ta có

$$f(b) \leq \frac{1}{\lambda}f(x) + \frac{\lambda - 1}{\lambda}f(a).$$

Từ đây và do giả thiết $f(x) \leq m < \infty$ với mọi $x \in M$, ta suy ra

$$f(b) - f(a) \leq \frac{1}{\lambda}f(x) - \frac{1}{\lambda}f(a) \leq \frac{1}{\lambda}[m - f(a)].$$

Điều này đúng với mọi $\lambda > 1$, nên khi cho $\lambda \rightarrow +\infty$ ở vế phải, do $f(a)$ hữu hạn, nên vế phải tiến tới 0, trong khi đó theo giả thiết, vế trái $f(b) - f(a) > 0$. Mâu thuẫn. Vậy f phải là hằng số trên tập a -phân M . $\square$

Hệ quả 9.1. Nếu một hàm lồi đạt cực đại trên một tập lồi có điểm cực biên, thì cực đại sẽ đạt tại một điểm cực biên của tập lồi đó.

Chứng minh. Giả sử x^* là điểm cực đại của f trên tập lồi C . Nếu x^* không phải điểm cực biên của C , thì tồn tại $a, b \in C$ và $\lambda \in (0, 1)$ sao cho $x^* = \lambda a + (1 - \lambda)b$. Theo (i) của Mệnh đề 9.2, ta có: $f(x^*) = f(x)$ với mọi $x \in [a, b]$. $\square$

Hệ quả 9.2. Cho $\Gamma := \{\lambda d \mid \lambda \geq 0\}$ và $\Gamma_a := a + \Gamma$ với $a, d \in \mathbb{R}^n$. Giả sử f là một hàm lồi chính thường trên $\mathbb{R}^n$ và f bị chặn trên trên nửa đường thẳng Γ_a . Khi đó f bị chặn trên trên mọi nửa đường thẳng song song với Γ_a . Ngoài ra cực đại của f trên nửa đường thẳng Γ_a đạt tại đầu mút của nó.

Chứng minh. Do f bị chặn trên trên tia Γ_a , nên $\Gamma_a \subset \text{dom} f$. Nếu $\Gamma_b // \Gamma_a$, thì $\Gamma_b \subset \text{dom} f$. $\square$

Theo Hệ quả 9.2 tính bị chặn trên của một hàm lồi trên một nửa đường thẳng không phụ thuộc vào đầu mút của nửa đường thẳng mà chỉ phụ thuộc vào hướng của nó.

9.2. Hạng của hàm lồi

Từ định nghĩa, ta đã thấy rằng một hàm a-phin là một hàm vừa lồi vừa lõm. Trong trường hợp này bài toán cực đại hay cực tiểu hàm tuyến tính đều có tất cả các tính chất về cực tiểu và cực đại hàm lồi, nêu ở mục trên. Điều này gợi ý đến việc xem xét "mức độ tuyến tính" của một hàm lồi.

Một câu hỏi được đặt ra là: nếu một hàm lồi không là hàm a-phin, thì "mức độ phi tuyến" của nó như thế nào. Hạng của một hàm lồi chính là đại lượng đo mức độ phi tuyến của nó.

Giả sử f là một hàm lồi chính thường trên $\mathbb{R}^n$. Khi đó với mỗi

$\alpha \in \mathbb{R}$, tập

$$L_f(\alpha) := \{x : f(x) \leq \alpha\}$$

là lồi. Tập này thường được gọi là *tập mức dưới*. Tập

$$U_f(\alpha) := \{x : f(x) \geq \alpha\}$$

được gọi là *tập mức trên*.

Dĩ nhiên khi α càng lớn, thì tập mức dưới ứng với nó càng rộng. Tuy nhiên tất cả các tập mức dưới đều có chung một nón lồi xa và một không gian thẳng. Cụ thể ta có mệnh đề sau:

Mệnh đề 9.3. Cho f là một hàm lồi, chính thường, nửa liên tục dưới trên $\mathbb{R}^n$. Khi đó tất cả các tập mức dưới $L_f(\alpha)$ (nếu khác rỗng) đều có chung một nón lồi xa và một không gian thẳng. Nón lồi xa là tập hợp hướng của các nửa đường thẳng, theo đó hàm f bị chặn trên, còn không gian thẳng là không gian con song song với không gian a -phân trên đó f là hằng số.

Chứng minh. Giả sử Γ_a là một hướng xuất phát từ a , theo đó f bị chặn trên. Với bất kỳ $\alpha \in \mathbb{R}$ và $a \in L_f(\alpha)$, do f lồi và bị chặn trên trên Γ_a , nên cực đại của f trên Γ_a đạt tại đầu mút a . Tức là $f(x) \leq f(a)$ với mọi $x \in \Gamma_a$. Suy ra $\Gamma_a \subseteq L_f(f(a))$.

Ngược lại, nếu Γ_a là một hướng thuộc nón lồi xa của tập mức dưới $L_f(\alpha)$, thì $\Gamma_a \subset L_f(\alpha)$ với mọi $a \in L_f(\alpha)$. Vậy $f(x) \leq \alpha$ với mọi $x \in \Gamma_a$. Chứng tỏ f bị chặn trên trên nửa đường thẳng Γ_a .

Phần khẳng định về không gian thẳng của tập mức dưới nêu trong mệnh đề được suy ra từ định nghĩa của không gian thẳng và tính chất là một hàm lồi bị chặn trên trong một tập a -phân, sẽ là hằng số trên tập đó. $\square$

Theo Mệnh đề 9.3 mọi tập mức dưới của một hàm lồi, chính thường, nửa liên tục dưới đều có chung một nón lồi xa và một không gian thẳng. Nón này được gọi là *nón lồi xa*, còn không gian này được gọi là *không gian hằng* của hàm số.

Với tư cách là một tập lồi trong không gian $\mathbf{R}^{n+1}$, trên đồ thị của một hàm lồi f cũng có một không gian thẳng. Hình chiếu của không gian thẳng của $\text{epi} f$ trên $\mathbf{R}^n$ được gọi là *không gian thẳng* của f . Ta sẽ ký hiệu $L(f)$ là không gian thẳng của f . Như vậy

$$L(f) = \{y \in \mathbf{R}^n \mid \exists t \in \mathbf{R} : (t, y) \in L(\text{epi} f)\}.$$

Mệnh đề dưới đây cho thấy cấu trúc của không gian thẳng của f .

Mệnh đề 9.4. *Giả sử f là một hàm lồi chính thường trên $\mathbf{R}^n$. Khi đó $y \in L(f)$ khi và chỉ khi tồn tại $t \in \mathbf{R}$ thỏa mãn*

$$f(x + \lambda y) = f(x) + \lambda t \quad \forall x \in \text{dom} f, \quad \forall \lambda \in \mathbf{R}.$$

Nếu f đóng, thì trong công thức trên có thể thay với mọi $x \in \text{dom} f$ bởi với một $x \in \text{dom} f$.

Chứng minh. Theo định nghĩa, (y, t) thuộc không gian thẳng của $\text{epi} f$ khi và chỉ khi

$$(x, f(x)) + \lambda(y, t) \in \text{epi} f, \quad \forall x \in \text{dom} f, \quad \forall \lambda \in \mathbf{R}.$$

Tức là

$$f(x + \lambda y) - \lambda t \leq f(x) \quad \forall \lambda \in \mathbf{R}.$$

Vậy hàm lồi một biến $\varphi(\lambda) := f(x + \lambda y) - \lambda t$ bị chặn trên trong $\mathbf{R}$. Theo Mệnh đề 9.2, φ là hằng số trên $\mathbf{R}$, và do đó

$$f(x + \lambda y) - \lambda t = f(x) \quad \forall \lambda.$$

Trong trường hợp f đóng, tức là $\text{epi} f$ đóng, thì điều khẳng định sau của mệnh đề được suy ra từ sự kiện (Mệnh đề 1.7) là một véc-tơ y thuộc nón các phương vô hạn của một tập lồi đóng C khi và chỉ khi

$$x + \lambda y \in C, \quad \forall \lambda \geq 0$$

với một $x \in C$. □

Theo công thức (9.1) hàm $\phi(\lambda) := f(x + \lambda y)$ là một hàm a-phin khi $y \in L(f)$. Như vậy $L(f)$ là tập hợp các hướng y trên đó hàm f a-phin. Ký hiệu $\text{lin} f$ là thứ nguyên (số chiều) của không gian thẳng $L(f)$. Như vậy $\text{lin} f$ chính là đại lượng đo "mức độ tuyến tính" của f và do đó

$$\text{rank} f := \dim f - \text{lin} f$$

sẽ là đại lượng đo "mức độ phi tuyến" của f và nó được gọi là *hạng* của f . Nhớ rằng $\dim f$ là số chiều của tập hữu dụng của f .

Khái niệm về hạng của một hàm lồi có vai trò rất quan trọng trong vấn đề phân rã các bài toán phi tuyến. Nói chung trong hầu hết các lĩnh vực của toán học tính toán, các bài toán phi tuyến thường khó giải hơn nhiều so với các bài toán tuyến tính cùng loại. Ngay với các thể hệ máy tính hiện nay, khi đã có thể giải các bài toán tuyến tính với số biến khá lớn thì việc giải các bài toán phi tuyến cùng loại chỉ mới thực hiện được trên các bài toán có số biến khá nhỏ so với các bài toán tuyến tính. Trong tình huống này một câu hỏi tự nhiên được đặt ra là phải phân loại các "biến phi tuyến" tức là các biến làm cho bài toán trở nên phi tuyến và các "biến tuyến tính" để xử lý một cách riêng biệt. Từ đó có thể phân rã bài toán đang xét thành các bài toán phi tuyến và tuyến tính, theo đó trong các bài toán phi tuyến chỉ có sự tham gia của các biến phi tuyến hoặc ít ra là các biến phi tuyến chỉ tham gia trong các phép toán gây nên tính khó giải của bài toán. Để minh hoạ ta hãy xét bài toán tối ưu

$$\max\{f(x) : x \in D \subseteq \mathbf{R}^n\}, \quad (9.2)$$

trong đó f là một hàm toàn phương lồi cho bởi

$$f(x) := f(x_1, \dots, x_n) = \sum_{j=1}^k \lambda_j x_j^2 + \sum_{j=k+1}^n \lambda_j x_j \quad (9.3)$$

với $\lambda_j > 0$ ($j = 1, \dots, k$) và D là một tập lồi đa diện, ví dụ D được cho bởi

$$D := \{x = (x_1, \dots, x_n) \geq 0 : Ax = b\},$$

với A là một ma trận thực cấp $m \times n$, $b \in \mathbb{R}^m$. Dễ dàng kiểm tra rằng $\dim f = n$ và $\text{rank} f = k$.

Bài toán (9.2) là một bài toán quan trọng của tối ưu toàn cục vì rất hay gặp nó trong nhiều ứng dụng như kinh tế, sinh học v.v... Chú ý rằng khi $k = 0$, thì f là hàm tuyến tính, và do đó (9.2) trở thành một quy hoạch tuyến tính. Các thuật toán đã có hiện nay có thể giải bài toán quy hoạch tuyến tính này với số biến lên đến hàng triệu. Tuy nhiên vấn đề trở nên phức tạp khi $k > 0$. Người ta đã chứng minh rằng khi k đủ chỉ bằng 1, thì (9.2) đã thuộc lớp bài toán khó giải, gọi là lớp *NP-khó*, trong khi $k = 0$ bài toán (9.2) là một quy hoạch tuyến tính, có thể giải được bằng các thuật toán có thời gian đa thức. Về mặt lý thuyết, đã có nhiều phương pháp giải bài toán (9.2). Tuy nhiên các thuật toán không sử dụng đến $\text{rank} f$ chỉ có thể giải được bài toán trên với số chiều n khá nhỏ (khoảng 20) ngay cả với các thể hệ máy tính hiện nay. Thế nhưng nếu để ý rằng do $\text{rank} f = k$, nên có một không gian với số chiều là $n - k$ mà trên đó hàm f tuyến tính và vì vậy tính phi tuyến của f chỉ còn hạn chế trên một không gian k chiều. Điều này rất có ý nghĩa khi k rất nhỏ so với n . Dựa vào nhận xét này người ta đã xây dựng được các thuật toán giải bài toán (9.2) với số chiều n khá lớn, miễn là $\text{rank} f$ vừa nhỏ.

9.3. Bất đẳng thức lồi

Cho $D \subseteq \mathbb{R}^n$ là một tập lồi và $f_1, \dots, f_m$ là các hàm lồi trên $\mathbb{R}^n$. Hệ bất đẳng thức

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

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

Hệ bất đẳng thức lồi xuất hiện trong nhiều ứng dụng. Do tính chất của hàm lồi, tập nghiệm của một hệ bất đẳng thức lồi luôn là một tập lồi. Dựa vào các định lý tách, ta có thể đưa ra những điều kiện cần hoặc/và đủ về sự tồn tại nghiệm của một hệ bất đẳng thức lồi.

Mệnh đề 9.5. Giả sử $f_0, f_1, \dots, f_m$ là các hàm lồi hữu hạn trên tập lồi $D \neq \emptyset$. Khi đó hệ

$$x \in D, f_i(x) < 0, i = 0, 1, \dots, m \quad (9.4)$$

không có nghiệm, khi và chỉ khi tồn tại các số $\lambda_i \geq 0$ ($i = 0, 1, \dots, m$) không đồng thời bằng 0 sao cho

$$\sum_{i=0}^m \lambda_i f_i(x) \geq 0 \quad \forall x \in D. \quad (9.5)$$

Ngoài ra nếu có điều kiện chính quy Slater: tồn tại $x^0 \in D, f_i(x^0) < 0$ với mọi $i = 1, \dots, m$, thì $\lambda_0 > 0$.

Chứng minh. Điều kiện đủ là hiển nhiên, nên ta chỉ cần chứng minh điều kiện cần. Giả sử hệ (9.4) không có nghiệm. Đặt

$$C := \{y = (y_0, y_1, \dots, y_m) \in \mathbf{R}^{m+1} \mid \exists x \in D, f_i(x) < y_i, i = 0, \dots, m\}.$$

Do D và f_i lồi ($i = 0, \dots, m$), nên C lồi. Theo giả thiết hệ (9.4) không có nghiệm, nên $0 \notin C$. Áp dụng định lý tách, tồn tại $\lambda = (\lambda_0, \lambda_1, \dots, \lambda_m) \neq 0$ sao cho

$$\sum_{i=0}^m \lambda_i y_i \geq 0 \quad \forall y \in C.$$

Chú ý rằng theo định nghĩa của C , với mọi $x \in D$ và $\epsilon > 0$, ta có

$$(f_0(x) + \epsilon, \dots, f_m(x) + \epsilon) \in C.$$

Vậy

$$\sum_{i=0}^m \lambda_i (f_i(x) + \epsilon) \geq 0 \quad \forall x \in D. \quad (9.5)$$

Điều này đúng với mọi $\epsilon > 0$, nên suy ra

$$\sum_{i=0}^m \lambda_i f_i(x) \geq 0 \quad \forall x \in D. \quad (9.6)$$

Ta sẽ chứng tỏ $\lambda_i \geq 0$ với mọi i . Thật vậy, nếu trái lại có $\lambda_j < 0$ với một j nào đó, thì do với mọi $x \in D$, mọi $y_i > f_i(x)$, ta có $(y_0, \dots, y_m) \in C$, nên

$$\sum_{i=0}^m \lambda_i y_i \geq 0.$$

Cho $y_j \rightarrow +\infty$ còn mọi y_i khác cố định, ta thấy vế trái của bất đẳng thức trên tiến đến $-\infty$. Mâu thuẫn vì vế phải bằng 0. Vậy $\lambda_i \geq 0$ với mọi i .

Cuối cùng giả sử điều kiện Slater thoả mãn. Nếu như $\lambda_0 = 0$, thì theo (9.6), có

$$\sum_{i=1}^m \lambda_i f_i(x) \geq 0 \quad \forall x \in D.$$

Lấy $x = x^0 \in D$, theo điều kiện chính qui Slater thì

$$\sum_{i=1}^m \lambda_i f_i(x^0) < 0.$$

Ta có mâu thuẫn và do đó mệnh đề được chứng minh. $\square$

Trong nhiều ứng dụng, trong một hệ bất đẳng thức lồi thường có sự tham gia của các đẳng thức tuyến tính. Khi đó ta có mệnh đề sau.

Mệnh đề 9.6. Cho $f_1, \dots, f_m$ 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 $b \in \text{int } A(D)$. Khi đó hệ

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

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

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

Chứng minh. Cũng như trên, dễ thấy rằng điều kiện đủ là hiển nhiên. Để chứng minh điều kiện cần, lấy tập

$$E := \{x \in D \mid Ax = b\}.$$

Do D lồi, nên E lồi và theo giả thiết hệ

$$x \in E, f_i(x) < 0, i = 1, \dots, m$$

không có nghiệm. Vậy áp dụng mệnh đề trên, sẽ tồn tại $\lambda_i, i = 1, \dots, m$ thoả mãn $\sum_{i=1}^m \lambda_i > 0$ và

$$\sum_{i=1}^m \lambda_i f_i(x) \geq 0 \quad \forall x \in E.$$

Bằng cách chia cho $\sum_{i=1}^m \lambda_i > 0$, ta có thể coi $\sum_{i=1}^m \lambda_i = 1$.

Với mỗi $x \in D$, lấy $f(x) := \sum_{i=1}^m \lambda_i f_i(x)$. Khi đó f lồi và hữu hạn trên D . Lấy

$$C := \{(y, y_0) \in \mathbf{R}^k \times \mathbf{R} \mid \exists x \in D, Ax - b = y, f(x) < y_0\}.$$

Do D và f lồi, nên C lồi và, theo giả thiết, suy ra $0 \notin C$. Theo Mệnh đề 6.1 có thể tách đúng C và 0 . Tức là tồn tại $(t, t_0) \in \mathbf{R}^k \times \mathbf{R}$ và $(\bar{y}, \bar{y}_0) \in C$ sao cho

$$\langle t, y \rangle + t_0 y_0 \geq 0 \quad \forall (y, y_0) \in C,$$

và

$$\langle t, \bar{y} \rangle + t_0 \bar{y}_0 > 0. \quad (9.7)$$

Dựa vào định nghĩa của C , lập luận tương tự như mệnh đề trước, ta có $t_0 \geq 0$. Nhưng t_0 không thể bằng 0, vì nếu $t_0 = 0$, thì

$$\langle t, y \rangle \geq 0 \quad \forall y \in A(D) - b. \quad (9.8)$$

Thế nhưng theo giả thiết $b \in \text{int}A(D)$, tức là $0 \in \text{int}(A(D) - b)$. Từ đây và (9.8) suy ra

$$\langle t, y \rangle = 0 \quad \forall y \in A(D).$$

Do $t_0 = 0$, nên

$$\langle t, y \rangle + t_0 y_0 = 0 \quad \forall (y, y_0) \in C.$$

Mâu thuẫn với (9.7), vì $\bar{y} \in A(D)$. Vậy $t_0 > 0$.

Nhớ lại định nghĩa của C ta có $(Ax - b, f(x) + \epsilon) \in C \quad \forall \epsilon > 0, \forall x \in D$. Vậy

$$\langle t, Ax - b \rangle + t_0(f(x) + \epsilon) \geq 0 \quad \forall x \in D.$$

Điều này đúng với mọi $\epsilon > 0$, nên

$$\langle t, Ax - b \rangle + t_0 f(x) \geq 0 \quad \forall x \in D.$$

Thay $f(x) = \sum_{i=1}^m \lambda_i f_i(x)$ và chia hai vế cho $t_0 > 0$, ta có điều cần chứng minh. $\square$

Chú ý. Để thấy rằng mệnh đề vẫn còn đúng khi thay hệ $Ax = b$ bởi hệ $Ax \leq b$.

9.4. Định lý Helley

Định lý Helly, đôi khi còn gọi là định lý về sự tương giao của các tập lồi. Đây là một trong các định lý cơ bản trong lý thuyết tổ hợp các tập lồi. Định lý này do G. Helley, một nhà toán học người

Uc phát hiện ra từ giữa thế kỷ 19. Phạm vi ứng dụng của định lý này khá rộng rãi, đặc biệt là trong việc chứng minh về sự tồn tại (điểm bất động, nghiệm của hệ phương trình, bất đẳng thức biến phân, bài toán tối ưu, điểm cân bằng v.v...). Định lý đã được mở rộng theo nhiều hướng khác nhau. Dưới đây sẽ trình bày một dạng cơ bản của định lý này. Ta chọn một chứng minh quy nạp sơ cấp dựa vào mệnh đề sau.

Mệnh đề 9.7. *Mọi tập hợp của các điểm khác nhau, phụ thuộc a-phin trong $\mathbb{R}^n$, đều có thể biểu diễn như là hợp của hai tập rời nhau nhưng bao lồi của chúng lại có điểm chung.*

Chứng minh. Giả sử các điểm $x^1, \dots, x^k$ phụ thuộc a-phin. Khi đó tồn tại các số $\alpha_1, \dots, \alpha_k$ không đồng thời bằng 0, sao cho

$$\sum_{j=1}^k \alpha_j x^j = 0, \quad \sum_{j=1}^k \alpha_j = 0.$$

Bằng cách đánh lại chỉ số, nếu cần, ta có thể giả sử, rằng $\alpha_j > 0$ với $j = 1, \dots, p$. Vậy $1 \leq p < k$. Khi đó do

$$\sum_{j=1}^k \alpha_j = 0,$$

nên

$$\alpha := \sum_{j=1}^p \alpha_j = - \sum_{j=p+1}^k \alpha_j > 0.$$

Từ đây và do

$$\sum_{j=1}^k \alpha_j x^j = 0,$$

nên

$$\sum_{j=1}^p \frac{\alpha_j}{\alpha} x^j = \sum_{j=p+1}^k -\frac{\alpha_j}{\alpha} x^j.$$

Đặt

$$x := \sum_{j=1}^p \frac{\alpha_j}{\alpha} x^j$$

Do $\frac{\alpha_j}{\alpha} > 0$ với mọi $j = 1, \dots, p$ và $-\frac{\alpha_j}{\alpha} > 0$ với mọi $j = p+1, \dots, k$ và

$$\sum_{j=1}^p \frac{\alpha_j}{\alpha} = \sum_{j=p+1}^k -\frac{\alpha_j}{\alpha} = 1,$$

nên

$$x \in \text{co}\{x^1, \dots, x^p\} \cap \text{co}\{x^{p+1}, \dots, x^k\}.$$

□

Mệnh đề 9.8 (Định lý Helley). Cho $C_1, \dots, C_k$ là các tập lồi trong $\mathbb{R}^n$. Giả sử rằng bất cứ $(n+1)$ tập trong các tập này đều có điểm chung. Khi đó toàn bộ tập này cũng có điểm chung.

Chứng minh. Ta chứng minh bằng qui nạp theo k . Nếu $k \leq n+1$ thì không có gì cần phải chứng minh. Cho $k > n+1$ và giả sử rằng định lý đúng với $k-1$. Khi đó với mọi $j \in \{1, \dots, k\}$, theo giả thiết quy nạp có

$$x^j \in C_1 \cap \dots \cap C_{j-1} \cap C_{j+1} \cap \dots \cap C_k. \quad (9.9)$$

Như vậy ta có k điểm $x^1, \dots, x^k$ trong $\mathbb{R}^n$ phụ thuộc a-phin (vì $k > n+1$). Theo định lý Radon, bằng cách đánh số lại các chỉ số, nếu cần thiết, ta có một điểm x thoả mãn

$$x \in \text{co}\{x^1, \dots, x^p\} \cap \text{co}\{x^{p+1}, \dots, x^k\} \quad (9.10)$$

với một chỉ số p nào đó thuộc $\{1, \dots, k-1\}$. Thế nhưng từ (9.9) ta có

$$x^i \in \cap_{j=p+1}^k C_j \quad \forall i = 1, \dots, p.$$

Từ đây và từ (9.10) và do C_j lồi, ta suy ra

$$x \in \text{co}\{x^1, \dots, x^p\} \subset \cap_{j=p+1}^k C_j.$$

Một cách tương tự, ta có

$$x \in \text{co}\{x^{p+1}, \dots, x^k\} \subset \bigcap_{j=1}^p C_j.$$

Vậy $x \in \bigcap_{j=1}^k C_j$. □

Chú ý. Trong trường hợp có vô hạn tập, thì Định lý Helley nêu trên không còn đúng. Ví dụ trong trục số thực lấy họ tập $\{C_j\}$ ($j = 0, 1, 2, \dots$) với $C_j = [j, +\infty)$. Trong ví dụ này các tập không compact. Trong trường hợp có ít nhất một tập compact, thì định lý Helley vẫn còn đúng trong trường hợp có vô số tập. Cụ thể ta có mệnh đề sau:

Mệnh đề 9.9 (Định lý Helley khi có vô hạn tập). *Cho Γ là một họ vô hạn các tập lồi, đóng, khác rỗng trong không gian $\mathbb{R}^n$. Giả sử rằng trong họ này có một tập compact và bất cứ $n + 1$ tập của nó đều có giao khác rỗng. Khi đó toàn bộ các tập của họ này cũng có giao khác rỗng.*

Có thể chứng minh định lý này theo nhiều cách khác nhau. Cách chứng minh của Radon dựa vào tính toán đại số, như đã trình bày trong chứng minh ở Mệnh đề 9.8. Còn có thể chứng minh định lý này bằng cách dùng nón đối ngẫu, rất ngắn gọn, và cũng có thể chứng minh dựa vào định lý tách các tập lồi. Dưới đây ta sẽ giới thiệu vắn tắt hai cách chứng minh sau này.

Chứng minh (dùng định lý tách) Ta sẽ chứng minh theo qui nạp. Giả sử mọi s phần tử, với $s \geq n + 1$ của họ Γ đều có điểm chung. Ta cần chứng tỏ điều này đúng với $(s + 1)$ phần tử của họ này. Giả sử trái lại, ta có $C_1 \cap C_2 \cap \dots \cap C_{s+1} = \emptyset$. Do tập $C := C_1 \cap C_2 \cap \dots \cap C_s$ và C_{s+1} là hai tập lồi, đóng, khác rỗng và không giao nhau, nên tồn tại một siêu phẳng H tách đúng hai tập này thoả mãn $C \cap H = \emptyset$. Theo giả thiết, với mọi bộ gồm $(s - 1)$ tập $C_{i_1}, \dots, C_{i_{s-1}}$ ta đều có

$$C_{i_1} \cap \dots \cap C_{i_{s-1}} \cap C_{s+1} \neq \emptyset.$$