

TỐI ƯU LỖI

GIẢNG VIÊN: TRẦN HÀ SƠN

ĐẠI HỌC KHOA HỌC TỰ NHIÊN-ĐHQG TP HCM

Ngày 31 tháng 3 năm 2024

Overview

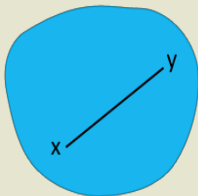
- 1 Tập lồi
- 2 Hàm lồi một biến
- 3 Ma trận xác định dương
- 4 Hàm lồi nhiều biến
- 5 Dạng toàn phương

Khái niệm tập lồi

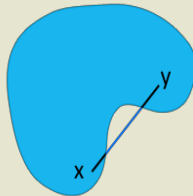
Định nghĩa

Tập $S \subset \mathbb{R}^n$ khác rỗng được gọi là tập lồi nếu với mọi cặp điểm $x, y \in S$ và với mọi số $\lambda \in [0, 1]$, ta có $\lambda x + (1 - \lambda)y \in S$.

Convex set



Non-convex set



Nhận xét

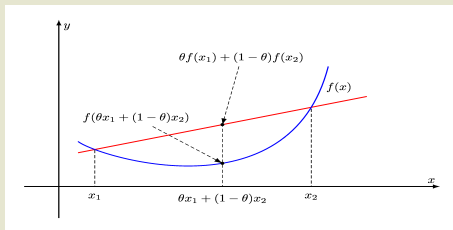
Về mặt hình học, tập S là lồi nếu với mọi điểm $x, y \in S$, toàn bộ đoạn thẳng xy đều nằm trong S .

Hàm lồi một biến

Định nghĩa

Hàm số $f : (a, b) \rightarrow \mathbb{R}$ được gọi là hàm lồi trên (a, b) nếu với mọi $x_1, x_2 \in (a, b)$ và $\theta \in [0, 1]$, ta có

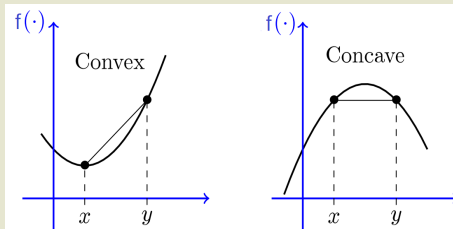
$$f(\theta x_1 + (1 - \theta)x_2) \leq \theta f(x_1) + (1 - \theta)f(x_2).$$



Hàm lồi một biến

Định nghĩa

Cho $f : (a, b) \rightarrow \mathbb{R}$, f được gọi là hàm lõm (concave) nếu $-f$ là hàm lồi.



Định lý

Cho $f : (a, b) \rightarrow \mathbb{R}$, có đạo hàm cấp hai tồn tại và liên tục trên (a, b) .
Các phát biểu sau là tương đương:

- 1 f là hàm lồi.
- 2 $f(y) \geq f(x) + f'(x)(y - x)$ với mọi $x, y \in (a, b)$.
- 3 $f''(x) \geq 0$ với mọi $x \in (a, b)$.

Hàm lồi một biến

Ví dụ

Chứng minh rằng các hàm $f(x) = x^{2k}$, $k \in \mathbb{N}$, $g(x) = e^x$ là các hàm lồi.

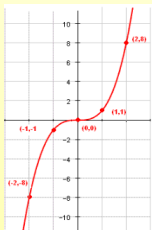
Hàm lồi một biến

Hệ quả

Cho $f : (a, b) \rightarrow \mathbb{R}$, có đạo hàm cấp hai tồn tại và liên tục trên (a, b) . Các phát biểu sau là tương đương:

- 1 f là hàm lõm.
- 2 $f(y) \leq f(x) + f'(x)(y - x)$ với mọi $x, y \in (a, b)$.
- 3 $f''(x) \leq 0$ với mọi $x \in (a, b)$.

Chứng minh hàm số $y = x^3$ lõm trên $(-\infty, 0)$ và lồi trên $(0, +\infty)$.



Hàm lồi một biến

Hệ quả

Cho $f : (a, b) \rightarrow \mathbb{R}$, có đạo hàm bậc nhất tồn tại và liên tục trên (a, b) , $x_0 \in (a, b)$ thỏa $f'(x_0) = 0$.

- ➊ Nếu f là hàm lồi thì f đạt giá trị nhỏ nhất tại x_0 .
- ➋ Nếu f là hàm lõm thì f đạt giá trị lớn nhất tại x_0 .

Ví dụ

Tìm giá trị lớn nhất, nhỏ nhất (nếu có) của các hàm số sau.

- a) $f(x) = x^4 + 6x^2$.
- b) $g(x) = -x \ln(x)$ trên $(0, +\infty)$.

Định nghĩa

Cho ma trận vuông $A \in M_n(\mathbb{R})$.

- A được gọi là xác định dương (ký hiệu $A > 0$) nếu $x^T A x > 0$ với mọi $x \in \mathbb{R}^n \setminus 0$.
- A được gọi là nửa xác định dương (ký hiệu $A \geq 0$) nếu $x^T A x \geq 0$ với mọi $x \in \mathbb{R}^n$.
- A được gọi là xác định âm (ký hiệu $A < 0$) nếu $x^T A x < 0$ với mọi $x \in \mathbb{R}^n \setminus 0$.
- A được gọi là nửa xác định âm (ký hiệu $A \leq 0$) nếu $x^T A x \leq 0$ với mọi $x \in \mathbb{R}^n$.

Ma trận xác định dương

Ví dụ (Ma trận xác định dương)

Cho ma trận vuông $A = \begin{bmatrix} 2 & 4 \\ 4 & 9 \end{bmatrix}$ là ma trận xác định dương vì với mọi $x \in \mathbb{R}^2$,

$$x^T A x = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 4 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2x_1^2 + 8x_1x_2 + 9x_2^2 = 2(x_1 + x_2)^2 + x_2^2$$

(do $2(x_1 + x_2)^2 + x_2^2 > 0, \forall x \in \mathbb{R}^2 \setminus \{0\}$).

Lưu ý: A cũng là ma trận nửa xác định dương.

Ma trận xác định dương

Ví dụ (Ma trận xác định âm)

Cho ma trận vuông $B = \begin{bmatrix} -2 & 0 \\ 0 & -9 \end{bmatrix}$ là ma trận xác định dương vì với mọi $x \in \mathbb{R}^2$,

$$x^T B x = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} -2 & 0 \\ 0 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -2x_1^2 - 9x_2^2$$

(do $-2x_1^2 - 9x_2^2 < 0, \forall x \in \mathbb{R}^2 \setminus 0$).

Lưu ý: B cũng là ma trận nửa xác định âm.

Một số tiêu chuẩn kiểm tra ma trận xác định dương

Định lý (Tiêu chuẩn kiểm tra ma trận xác định dương)

Cho ma trận $A \in M_n(\mathbb{R})$.

- 1 Nếu $x^T A x > 0, \forall x \in \mathbb{R}^n$ thì A xác định dương.
- 2 Nếu tất cả các trị riêng của A đều dương thì A là xác định dương.
- 3 Nếu các định thức con D_1, D_2, \dots, D_n của A dương thì A xác định dương.

Một số tiêu chuẩn kiểm tra ma trận xác định dương

Ví dụ

Chứng minh rằng ma trận $A = \begin{bmatrix} 9 & 6 & -5 \\ 6 & 6 & -1 \\ -5 & -1 & 6 \end{bmatrix}$ là ma trận xác định dương.

Một số tiêu chuẩn kiểm tra ma trận xác định âm

Định lý (Một số tiêu chuẩn kiểm tra ma trận xác định âm)

Cho ma trận $A \in M_n(\mathbb{R})$.

- 1 Nếu $x^T A x < 0, \forall x \in \mathbb{R}^n$ thì A xác định âm.
- 2 Nếu tất cả các trị riêng của A đều âm thì A là xác định âm.
- 3 Nếu các định thức con của A đan dấu, nghĩa là

$$(-1)^k D_k > 0, k = 1, 2, \dots, n$$

thì A xác định âm.

Một số tiêu chuẩn kiểm tra ma trận xác định âm

Ví dụ (Một số tiêu chuẩn kiểm tra ma trận xác định âm)

Cho ma trận $B = \begin{bmatrix} -3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ là ma trận xác định âm vì

1 Kiểm tra các trị riêng

B có ba trị riêng $\lambda_1 = -3 < 0$, $\lambda_2 = -2 < 0$ và $\lambda_3 = -1 < 0$.

2 Kiểm tra các định thức con

B có ba định thức con:

$$D_1 = |-3| = -3 < 0, D_2 = \begin{vmatrix} -3 & 0 \\ 0 & -2 \end{vmatrix} = 6 > 0,$$

$$D_3 = \begin{vmatrix} -3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -1 \end{vmatrix} = -6 < 0.$$

Một số tiêu chuẩn kiểm tra ma trận nửa xác định âm, dương

Định lý

Cho ma trận đối xứng $A \in M_n(\mathbb{R})$.

- 1 A là ma trận nửa xác định dương khi và chỉ khi các trị riêng của nó không âm.
- 2 A là ma trận nửa xác định âm khi và chỉ khi các trị riêng của nó không dương.

Định nghĩa

Cho Ω là một tập mở lồi không rỗng của \mathbb{R}^n .

- ① Hàm số $f : \Omega \rightarrow \mathbb{R}$ được gọi là **hàm lồi** trên Ω nếu với mọi $x_1, x_2 \in \Omega$ và $\theta \in [0, 1]$, ta có

$$f(\theta x_1 + (1 - \theta)x_2) \leq \theta f(x_1) + (1 - \theta)f(x_2).$$

- ② Hàm số $f : \Omega \rightarrow \mathbb{R}$ được gọi là **hàm lõm** trên Ω nếu với mọi $x_1, x_2 \in \Omega$ và $\theta \in [0, 1]$, ta có

$$f(\theta x_1 + (1 - \theta)x_2) \geq \theta f(x_1) + (1 - \theta)f(x_2).$$

Định nghĩa

Cho $f : \mathbb{R}^n \rightarrow \mathbb{R}$ là hàm thuộc lớp C^2 . Vector gradient của f tại $x_0 = (x_1, x_2, \dots, x_n)$, ký hiệu là $\nabla f(x_0)$ được định nghĩa như sau:

$$\nabla f(x_0) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x_0) \\ \frac{\partial f}{\partial x_2}(x_0) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x_0) \end{bmatrix}$$

Hàm lồi nhiều biến

Ví dụ

Cho $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ được xác định bởi

$$f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - x_1x_2 + x_1x_3$$

Khi đó

$$\nabla f(x_1, x_2, x_3) = \begin{bmatrix} 2x_1 - x_2 + x_3 \\ 2x_2 - x_1 \\ 2x_3 + x_1 \end{bmatrix}$$

Ví dụ

Cho $f : \mathbb{R}^n \rightarrow \mathbb{R}$ là hàm thuộc lớp C^2 . Ma trận Hesse của f tại $x_0 = (x_1, x_2, \dots, x_n)$, ký hiệu là $\nabla^2 f(x_0)$ được định nghĩa như sau:

$$\nabla^2 f(x_0) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(x_0) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_0) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}(x_0) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_0) & \frac{\partial^2 f}{\partial x_2^2}(x_0) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(x_0) & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x_0) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x_0) \end{bmatrix}$$

Nhận xét: $\nabla^2 f(x)$ là một ma trận đối xứng vì $\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x)$ với mọi $1 \leq i, j \leq n$.

Ví dụ

Cho $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ được xác định bởi

$f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - x_1x_2 + x_1x_3$. Khi đó

$$\nabla f(x_1, x_2, x_3) = \begin{bmatrix} 2x_1 - x_2 + x_3 \\ 2x_2 - x_1 \\ 2x_3 + x_1 \end{bmatrix} \quad \text{và} \quad \nabla^2 f(x_1, x_2, x_3) = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}$$

Định lý

Cho $f : \Omega \rightarrow \mathbb{R}$ thuộc lớp C^2 xác định trên một tập mở lồi Ω của \mathbb{R}^n . Các phát biểu sau là tương đương.

- 1 f là hàm lồi.
- 2 $f(y) \geq f(x) + \nabla f(x)^T (y - x)$.
- 3 $\nabla^2 f(x) \geq 0$ với $\forall x \in \Omega$.

Hệ quả

Cho $f : \Omega \rightarrow \mathbb{R}$ thuộc lớp C^2 xác định trên một tập mở lồi Ω của \mathbb{R}^n . Các phát biểu sau là tương đương.

- 1 f là hàm lõm.
- 2 $f(y) \leq f(x) + \nabla f(x)^T (y - x)$.
- 3 $\nabla^2 f(x) \leq 0$ với $\forall x \in \Omega$.

Hệ quả

Cho Ω là tập mở lồi trên \mathbb{R}^n , hàm $f : \Omega \rightarrow \mathbb{R}$ thuộc lớp C^1 , xét điểm x_0 thỏa $\nabla f(x_0) = 0$. Khi đó

- 1 Nếu f lồi thì f đạt giá trị nhỏ nhất tại x_0 .
- 2 Nếu f lõm thì f đạt giá trị lớn nhất tại x_0 .

Ví dụ

Cho $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ được xác định bởi

$$f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - x_1x_2 + x_1x_3.$$

Hãy kiểm tra tính lồi, lõm và xác định cực trị của f (nếu có).

Định nghĩa

Cho vector $q \in \mathbb{R}^n$ và ma trận $A \in M_n(\mathbb{R})$.

- ① Hàm $f : \mathbb{R}^n \rightarrow \mathbb{R}$ xác định bởi

$$f = (q, x) = q^T x = \sum_{i=1}^n q_i x_i$$

được gọi là một dạng tuyến tính.

- ② Hàm $f : \mathbb{R}^n \rightarrow \mathbb{R}$ xác định bởi

$$f(x) = x^T A x = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$$

được gọi là dạng toàn phương.

Định lý

Cho $f_q(x)$ là một dạng tuyến tính với $q \in \mathbb{R}^n$ và $f_A(x) = x^T A x$ là một dạng toàn phương xác định bởi ma trận đối xứng $A \in M_n(\mathbb{R})$. Khi đó

- ① $\nabla f_q(x) = q$ và $\nabla^2 f_q(x) = 0$ với mọi $x \in \mathbb{R}^n$.
- ② $\nabla f_A(x) = 2Ax$ và $\nabla^2 f(x) = 2A$, với mọi $x \in \mathbb{R}^n$.

Định lý

Cho ma trận $A \in M_n(\mathbb{R})$ có hạng bằng n , $b \in \mathbb{R}^n$. Hàm số $f : \mathbb{R}^n \rightarrow \mathbb{R}$ được xác định bởi $f(x) = \|Ax - b\|^2$. Khi đó

- 1 $\nabla^2 f(x) = 2A^T A$ là ma trận nửa xác định dương.
- 2 Nghiệm $x = (A^T A)^{-1} A^T b$ của phương trình $\nabla f(x) = 0$ là điểm cực tiểu toàn cục của hàm f .

Ví dụ

Cho $b = (2, 2, 5, 8) \in \mathbb{R}^4$. Hãy tìm giá trị nhỏ nhất của $\|Ax - b\|$ trong các trường hợp

a)
$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}$$

b)
$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 3 & 3 \\ 1 & 4 & 4 \end{bmatrix}$$

Ví dụ

Khảo sát tính lồi lõm của hàm số xác định bởi

a) $f(x_1, x_2) = (x_1 - 3x_2)^2 + (x_1 - 2x_2)^2.$

b) $f(x_1, x_2) = (x_1 - 3x_2)^2 + (x_1 - 2x_2)^2 + x_1^2.$

Ví dụ

Khảo sát tính lồi lõm của hàm số xác định bởi

$$f(x) = \frac{1}{x_1 x_2} \text{ trên } \Omega = (0, +\infty) \times (0, +\infty).$$

Ví dụ

Khảo sát tính lồi lõm và tìm cực trị toàn cục (nếu có) của của hàm số $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ xác định bởi

$$\frac{3}{4}x_1^2 + \frac{3}{4}x_2^2 - \frac{1}{2}x_1x_2 - 2x_1 + x_2 + 3.$$

Ví dụ

Khảo sát tính lồi lõm và tìm cực trị toàn cục (nếu có) của của hàm số $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ xác định bởi

$$f(x) = x^T A x + q^T x + 1$$

với $A = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 1 & -2 \\ 1 & -2 & 1 \end{bmatrix}$ và $q = (1, 2, -1)$.