Tối ưu lồi

Đại học Khoa học Tự Nhiên, ĐHQG Thành phố Hồ Chí Minh

Ngày 10 tháng 11 năm 2021

1. Hàm số nhiều biến
2. Dạng toàn phương
3. Hàm lồi

Outlines

Định nghĩa

Một ánh xạ $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ gắn mỗi vector $\mathrm{x} = (x_1, x_2, \ldots, x_n) \in D$ một số thực

$$y = f(\mathrm{x}) = f(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$$

được gọi là một hàm số thực (real - value function) với n biến thực (real variable) trên miền xác định D.

- Nếu $n = 1$, f được gọi là hàm số 1 biến.
- Nếu $n \geq 2$ thì f được gọi là hàm số nhiều biến.

Cho hàm $f : D \subset \mathbb{R}^n$ xác định bởi

$$f(x, y) = x^3 + x^2y^3 - 2y^2$$

Ta có $f(0, y) = -2y^2, f(x, 0) = x^3, f(1, 1) = 0$

Định nghĩa

Cho $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ có đạo hàm liên tục tới cấp 2 và $\mathrm{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, vector gradient và ma trận Hess của f tại x, ký hiệu lần lượt là $\nabla f(\mathrm{x})$ và $\nabla^2 f(\mathrm{x})$ được xác định bởi

$$\nabla f(\mathrm{x}) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(\mathrm{x}) \\ \dfrac{\partial f}{\partial x_2}(\mathrm{x}) \\ \vdots \\ \dfrac{\partial f}{\partial x_n}(\mathrm{x}) \end{pmatrix}$$

$$\nabla^2 f(\mathrm{x}) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1}(\mathrm{x}) & \dfrac{\partial^2 f}{\partial x_2 \partial x_1}(\mathrm{x}) & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_1}(\mathrm{x}) \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2}(\mathrm{x}) & \dfrac{\partial^2 f}{\partial x_2 \partial x_2}(\mathrm{x}) & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_2}(\mathrm{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(\mathrm{x}) & \dfrac{\partial^2 f}{\partial x_2 \partial x_n}(\mathrm{x}) & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n}(\mathrm{x}) \end{pmatrix}$$

Mệnh đề

$\nabla^2 f(\mathrm{x})$ là ma trận đối xứng (khi f có đạo hàm liên tục tới cấp 2).

Ví dụ

Với $f(x, y) = x^3 + x^2y^3 - 2y^2$, tìm $\nabla f(\mathrm{x}), \nabla^2 f(\mathrm{x})$ tại $M(1, 1)$.

Ta có

$$\nabla f(\mathrm{x}) = \begin{pmatrix} 3x^2 + 2xy^3 \\ 3x^2y^2 - 4y \end{pmatrix}, \nabla f(1,1) = \begin{pmatrix} 5 \\ -1 \end{pmatrix}$$

$$\nabla^2 f(\mathrm{x}) = \begin{pmatrix} 6x + 2y^3 & 6xy^2 \\ 6xy^2 & 6x^2y - 4 \end{pmatrix}, \nabla f(1,1) = \begin{pmatrix} 8 & 6 \\ 6 & 2 \end{pmatrix}$$

Ví dụ

Cho hàm số $f(x,y,z) = 3x^3 - 2x^2y^3 + y^2x^2 + 2z^2y^2$, tìm $\nabla f(\mathrm{x}), \nabla^2 f(\mathrm{x})$ tại $M(1,2,1)$.

Định nghĩa

Cho a $= (a_1, \ldots, a_n) \in \mathrm{R}^n$, hàm số $f : \mathrm{R}^n \longrightarrow \mathbb{R}$ được xác định bởi

$$f(\mathrm{x}) = f(x_1, x_2, \ldots, x_n) = \mathrm{a}^T\mathrm{x} = \begin{pmatrix} a_1 & a_2 & \ldots & a_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} a_i x_i$$

được gọi là một dạng tuyến tính với a là vector xác định dạng.
Hơn nữa, nếu $b \in \mathbb{R}$, hàm số $f : \mathrm{R}^n \longrightarrow \mathbb{R}$ được xác định bởi

$$f(\mathrm{x}) = \mathrm{a}^T\mathrm{x} + b = b + \sum_{i=1}^{n} a_i x_i$$

được gọi là một hàm affine.

Mệnh đề

Nếu $f(\mathrm{x}) = \mathrm{a}^T\mathrm{x} + b$ là một hàm affine thì

$$\nabla f(\mathrm{x}) = \mathrm{a} \text{ và } \nabla^2 f(\mathrm{x}) = 0 \in \mathbb{R}^n$$

Ví dụ

Cho $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$ với

$$f(x_1, x_2, x_3) = 3x_3 - x_1$$

thì f là một dạng tuyến tính với vector xác định dạng $\mathrm{a} = (-1, 0, 3) \in \mathbb{R}^3$, cụ thể

$$f(x_1, x_2, x_3) = 3x_3 - x_1 = -x_1 + 0x_2 + 3x_3 = \begin{pmatrix} -1 & 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathrm{a}^T\mathrm{x}$$

Ví dụ

Ngoài ra,

$$\nabla f(x_1, x_2, x_3) = \begin{pmatrix} -1 \\ 0 \\ 3 \end{pmatrix} = \mathrm{a} \text{ và } \nabla^2 f(x_1, x_2, x_3) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Định nghĩa

Cho $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ có đạo hàm và $\mathrm{z} = (z_1, z_2, \ldots, z_n) \in \mathbb{R}^n$, xấp xỉ Taylor bậc nhất của f tại z là hàm affine $\hat{f} : \mathbb{R}^n \longrightarrow \mathbb{R}$ được xác định bởi.

$$\hat{f}(\mathrm{x}) = f(\mathrm{z}) + \nabla f(\mathrm{z})^T(\mathrm{x} - \mathrm{z}) = \nabla f(\mathrm{z})^T\mathrm{x} + (f(\mathrm{z}) - \nabla f(\mathrm{z})^T\mathrm{z})$$

Ví dụ

Cho $f(x, y) = x^3 + x^2y^3 - 2y^2$ và $\mathrm{z} = (1, 1)$, ta có $f(\mathrm{z} = 0), \nabla f(\mathrm{z}) = \begin{pmatrix} 5 \\ -1 \end{pmatrix}$, xấp xỉ Taylor bậc nhất của f tại z là hàm affine

$$\hat{f}(\mathrm{x}) = \nabla f(\mathrm{z})^T\mathrm{x} + (f(\mathrm{z}) - \nabla f(\mathrm{z})^T\mathrm{z})$$

$$\begin{pmatrix}5 & -1\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}+\left(0-\begin{pmatrix}5 & -1\end{pmatrix}\begin{pmatrix}1\\ 1\end{pmatrix}\right)=5x-y-4$$

Với hàm $\hat{f}$ đã có, ta tìm được giá trị xấp xỉ cho f tại $(0.9, 1.1)$ là

$$\hat{f}(0.9, 1.1) = -0.6 \approx f(0.9, 1.1) = -0.61289$$

Ví dụ - SV tự giải

Cho hàm số $f(x, y) = x^4 - 2x^2y^3 + y^3$ và z $= (1, 2)$. Dùng xấp xỉ Taylor bậc nhất của f tại z để tính giá trị xấp xỉ cho $f(0.8, 1.2)$

Outlines

Dạng toàn phương

Định nghĩa

Cho $A \in \mathbb{R}^{n\times n}$ là ma trận đối xứng ($A^T = T$), hàm số $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ được xác định bởi

$$f(\mathrm{x}) = f(x_1, x_2, \dots, x_n) = \mathrm{x}^T A\mathrm{x}$$
$$= \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}$$

được gọi là một dạng toàn phương (quadratic form) với A là ma trận xác định dạng.

Dạng toàn phương

Mệnh đề

Nếu $f(\mathrm{x}) = \mathrm{x}A\mathrm{x}$ là dạng toàn phương thì

$$\nabla f(\mathrm{x}) = 2A\mathrm{x} \text{ và } \nabla^2 f(\mathrm{x}) = 2A$$

Ví dụ

Cho $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$, với

$$f(x_1, x_2, x_3) = x_1^2 - x_3^2 + 2x_1x_3 + 4x_2x_3$$

Ta có f là một dạng toàn phương với

$$f(x_1, x_2, x_3) = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 2 \\ 1 & 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathrm{x}^T A\mathrm{x}$$

Dạng toàn phương

Ví dụ - SV tự giải

Cho $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$, với

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

Tìm ma trận xác định dạng A của dạng toàn phương trên.

Dạng toàn phương

Với dạng toàn phương trên, ta có

$$\nabla f(x_1, x_2, x_3) = \begin{pmatrix} 2x_1 + 2x_3 \\ 4x_3 \\ 2x_1 + 4x_2 - 2x_3 \end{pmatrix} = 2\begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 2 \\ 1 & 2 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 2A\mathrm{x}$$

$$\nabla^2 f(x_1, x_2, x_3) = \begin{pmatrix} 2 & 0 & 2 \\ 0 & 0 & 4 \\ 2 & 4 & -2 \end{pmatrix} = 2A$$

Dạng toàn phương

Định nghĩa

Cho $A \in \mathbb{R}^{n\times n}$ là ma trận đối xứng ($A^T = A$), ta có

- A xác định dương, ký hiệu $A > 0$ nếu $x^T A x > 0$ với mọi $x \neq 0$,
- A nửa xác định dương, ký hiệu $A \geq 0$ nếu $x^T A x \geq 0$ với mọi $x \neq 0$,
- A xác định âm, ký hiệu $A < 0$ nếu $x^T A x < 0$ với mọi $x \neq 0$,
- A nửa xác định âm, ký hiệu $A \leq 0$ nếu $x^T A x \leq 0$ với mọi $x \neq 0$,
- A không xác định nếu có x_1, x_2 sao cho $x_1{}^T A x_1 > 0, x_2{}^T A x_2 < 0$

Dạng toàn phương

Mệnh đề

Cho $A \in \mathbb{R}^{n\times n}$ là ma trận đối xứng ($A^T = A$).

- $A > 0$ khi và chỉ khi tất cả các trị riêng của A đều dương,
- $A \geq 0$ khi và chỉ khi tất cả các trị riêng của A đều không âm,
- $A < 0$ khi và chỉ khi tất cả các trị riêng của A đều âm,
- $A \leq 0$ khi và chỉ khi tất cả các trị riêng của A đều không dương,
- A không xác định khi và chỉ khi A có trị riêng dương và có trị riêng âm.

Dạng toàn phương

Ví dụ

Cho ma trận đối xứng $\begin{pmatrix} 3 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{pmatrix}$.

Ta có ma trận A là ma trận xác định của dạng toàn phương

$$f(x_1, x_2, x_3) = x^T A x = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 3 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$= 3x_1^2 + 2x_1x_2 + 2x_1x_3 + 4x_2x_3$$

Ta có $f(0,1,1) = 4 > 0$ và $f(0,1,-1) = -4 < 0$ nên A không xác định. Ngoài ra, ta có thể tìm các trị riêng của A để kết luận về dạng toàn phương.

Dạng toàn phương

Ví dụ - SV tự giải

Xác định tính chất của dạng toàn phương sau

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

Outlines

Hàm lồi

Hàm số $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ được gọi là lồi(convex) nếu với mọi $x, y \in \mathbb{R}^n$ và $\theta \in \mathbb{R}$ sao cho $0 \leq \theta \leq 1$, ta có

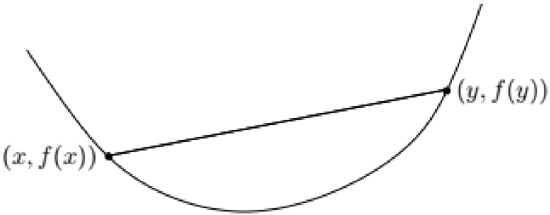
$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y) \tag{1}$$

Tương tự, f ta nói f là

- Lồi ngặt nếu dấu $\leq$ trong (1) được thay bằng dấu $<$,
- Lõm nếu dấu $\leq$ trong (1) được thay bằng dấu $\geq$,
- Lõm ngặt nếu dấu $\leq$ trong (1) được thay bằng dấu $>$

- Nếu f là hàm lồi (lõm) thì $-f$ là hàm lõm (lồi),
- Về mặt hình học, (1) có nghĩa là đoạn thẳng nối $(x, f(x))$ và $(y, f(y))$ luôn nằm trên phần đồ thị nối 2 điểm đó.

Hàm lồi



Hàm lồi

Mệnh đề

Cho $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ có đạo hàm liên tục đến cấp 2, các phát biểu sau là tương đương

i. f là hàm lồi,

ii. $f(y) \geq f(x) + \nabla f(x)^T (y - x)$, với mọi $x, y \in \mathbb{R}^n$,

iii. $\nabla^2 f(x) \geq 0$ hay $\nabla^2 f(x)$ là ma trận nửa xác định dương, với mọi $x \in \mathbb{R}^n$

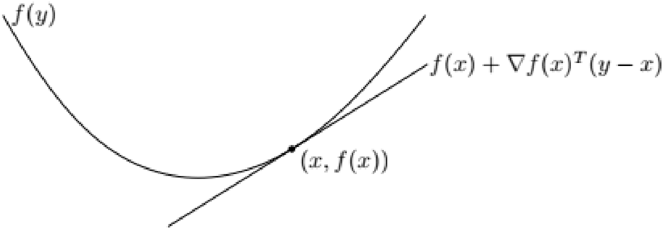
i. f là hàm lõm,

ii. $f(y) \leq f(x) + \nabla f(x)^T (y - x)$, với mọi $x, y \in \mathbb{R}^n$,

iii. $\nabla^2 f(x) \leq 0$ hay $\nabla^2 f(x)$ là ma trận nửa xác định âm, với mọi $x \in \mathbb{R}^n$

Hàm lồi

Từ ii., nếu $\nabla f(x)^T = 0$ thì $f(y) \geq f(x)$ với mọi $y \in \mathbb{R}^n$ nên x là một điểm cực tiểu toàn cục (global minimum point) mà tại đó hàm f nhận giá trị nhỏ nhất (minimum value).
Từ iii., ta có thể kiểm tra tính lồi (lõm) của hàm số bằng kĩ thuật của đại số tuyến tính.



Hàm lồi

Ví dụ

Với dạng toàn phương

$$f(x_1,x_2,x_3) = x_1^2 + x_2^2 + x_3^3 - x_1x_2 + x_1x_3, (x_1,x_2,x_3) \in \mathbb{R}^3$$

Ta có

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

$$\nabla^2 f(x_1,x_2,x_3) = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix}$$

Hàm lồi

Tìm trị riêng của ma trận trên bằng cách giải nghiệm của đa thức đặc trưng, ta có các tri riêng là $2, 2+\sqrt{2}, 2-\sqrt{2}$ nên $\nabla^2 f(x_1, x_2, x_3) \geq 0$.
Do đó nghiệm của phương trình $\nabla f(x_1, x_2, x_3) = (0, 0, 0)$ cho ta cực tiểu toàn cục của f trên R^3.

Ví dụ - SV tự giải

Khảo sát tính lồi(lõm) và tìm các điểm cực trị toàn cục (nếu có) của hàm số 3 biến sau

$$f(x_1, x_2, x_3) = 4x_1^2 + 3x_2^2 + 6x_3^2 - 4x_1x_2 - 4x_2x_3$$

Hàm lồi

Ví dụ - SV tự giải

Khảo sát tính lồi(lõm) của hàm số 3 biến sau

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