

Tối ưu hóa

Faculty of Computer Science
Phenikaa University

Ngày 17 tháng 4 năm 2024

Tối ưu lồi (Convex optimization)

Tập lồi

Đường thẳng và đoạn thẳng

Cho hai điểm x, y trong không gian \mathbb{R}^n . Tập hợp tất cả các điểm có dạng

$$z = \theta x + (1 - \theta)y,$$

ở đó $\theta \in \mathbb{R}$, tạo thành một **đường thẳng** đi qua x và y . Tập hợp các điểm z với $\theta \in [0, 1]$ tạo thành một **đoạn thẳng** nối x và y .

Tối ưu lồi (Convex optimization)

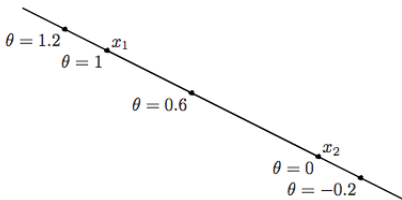
Tập lồi

Đường thẳng và đoạn thẳng

Cho hai điểm x, y trong không gian \mathbb{R}^n . Tập hợp tất cả các điểm có dạng

$$z = \theta x + (1 - \theta)y,$$

ở đó $\theta \in \mathbb{R}$, tạo thành một **đường thẳng** đi qua x và y . Tập hợp các điểm z với $\theta \in [0, 1]$ tạo thành một **đoạn thẳng** nối x và y .



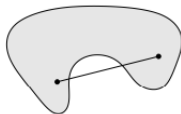
Tập lồi

Một tập hợp khác rỗng $C \subset \mathbb{R}^n$ được gọi là lồi nếu đoạn thẳng nối hai điểm bất kỳ của C nằm trong C .

Tập lồi

Một tập hợp khác rỗng $C \subset \mathbb{R}^n$ được gọi là lồi nếu đoạn thẳng nối hai điểm bất kỳ của C nằm trong C .

Hay nói cách khác, nếu $x, y \in C$ thì với mọi $\theta \in [0, 1]$, ta có $\theta x + (1 - \theta)y \in C$.

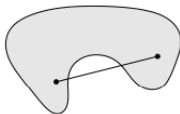


Tối ưu lồi

Tập lồi

Một tập hợp khác rỗng $C \subset \mathbb{R}^n$ được gọi là lồi nếu đoạn thẳng nối hai điểm bất kỳ của C nằm trong C .

Hay nói cách khác, nếu $x, y \in C$ thì với mọi $\theta \in [0, 1]$, ta có $\theta x + (1 - \theta)y \in C$.



Người ta quy ước tập rỗng là tập lồi

Ví dụ về tập lồi

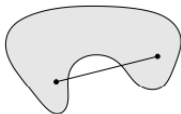
- Tập rỗng, tập chỉ gồm một phần tử (điểm), toàn bộ không gian \mathbb{R}^n

Tối ưu lồi

Tập lồi

Một tập hợp khác rỗng $C \subset \mathbb{R}^n$ được gọi là lồi nếu đoạn thẳng nối hai điểm bất kỳ của C nằm trong C .

Hay nói cách khác, nếu $x, y \in C$ thì với mọi $\theta \in [0, 1]$, ta có $\theta x + (1 - \theta)y \in C$.



Người ta quy ước tập rỗng là tập lồi

Ví dụ về tập lồi

- Tập rỗng, tập chỉ gồm một phần tử (điểm), toàn bộ không gian \mathbb{R}^n
- Một đoạn thẳng bất kỳ trong \mathbb{R}^n

Các tính chất của tập lồi

Nếu S_1 và S_2 là các tập lồi thì

- $S_1 \cap S_2$ là lồi;
- $S_1 + S_2 = \{x_1 + x_2 \mid x_1 \in S_1, x_2 \in S_2\}$ là lồi.

Các tính chất của tập lồi

Nếu S_1 và S_2 là các tập lồi thì

- $S_1 \cap S_2$ là lồi;
- $S_1 + S_2 = \{x_1 + x_2 \mid x_1 \in S_1, x_2 \in S_2\}$ là lồi.

Hợp của hai tập lồi có phải là tập lồi?

Bao lồi

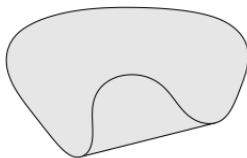
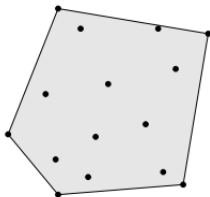
- Điểm $\theta_1 x_1 + \dots + \theta_k x_k$, ở đó $\theta_1 + \dots + \theta_k = 1$ và $\theta_i \geq 0$ với $i = 1, \dots, k$, được gọi là **một tổ hợp lồi** của các điểm x_1, \dots, x_k .
- **Bao lồi** của một tập hợp C , ký hiệu là $\text{conv } C$, là tập hợp tất cả các tổ hợp lồi của các điểm thuộc C .

$$\text{conv } C = \{\theta_1 x_1 + \dots + \theta_k x_k \mid x_i \in C, \theta_1, \dots, \theta_k \geq 0, \theta_1 + \dots + \theta_k = 1\}$$

Bao lồi

- Điểm $\theta_1 x_1 + \dots + \theta_k x_k$, ở đó $\theta_1 + \dots + \theta_k = 1$ và $\theta_i \geq 0$ với $i = 1, \dots, k$, được gọi là **một tổ hợp lồi** của các điểm x_1, \dots, x_k .
- **Bao lồi** của một tập hợp C , ký hiệu là $\text{conv } C$, là tập hợp tất cả các tổ hợp lồi của các điểm thuộc C .

$$\text{conv } C = \{\theta_1 x_1 + \dots + \theta_k x_k \mid x_i \in C, \theta_1, \dots, \theta_k \geq 0, \theta_1 + \dots + \theta_k = 1\}$$



Không gian Euclide n chiều trên trường số thực

Tích vô hướng

Tích vô hướng của 2 vectơ $u = (u_1, u_2, \dots, u_n)$ và $v = (v_1, v_2, \dots, v_n)$ trên Không gian Euclide \mathbb{R}^n

$$\langle u, v \rangle = u_1 v_1 + u_2 v_2 + \dots + u_n v_n.$$

Không gian Euclide n chiều trên trường số thực

Tích vô hướng

Tích vô hướng của 2 vectơ $u = (u_1, u_2, \dots, u_n)$ và $v = (v_1, v_2, \dots, v_n)$ trên Không gian Euclide \mathbb{R}^n

$$\langle u, v \rangle = u_1 v_1 + u_2 v_2 + \dots + u_n v_n.$$

Ta có thể viết $u^T v$ hoặc $v^T u$ thay cho $\langle u, v \rangle$

Siêu phẳng và nửa không gian

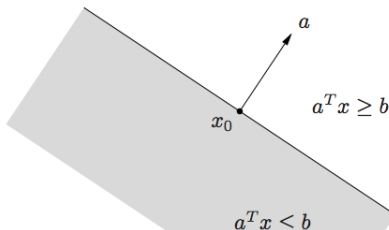
- Một **siêu phẳng** là một tập hợp có dạng

$$\{x \in \mathbb{R}^n \mid a^T x = b\} \quad \text{hoặc} \quad \{x \in \mathbb{R}^n \mid a_1 x_1 + \cdots + a_n x_n = b\},$$

ở đó, $a \in \mathbb{R}^n$, $a \neq 0$, và $b \in \mathbb{R}$.

- Một siêu phẳng chia không gian \mathbb{R}^n thành hai nửa không gian. Một **nửa không gian** (đóng) là một tập hợp có dạng

$$\{x \in \mathbb{R}^n \mid a^T x \leq b, a \neq 0\}.$$



Chuẩn trong không gian Euclide

Cho $v = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$

Chuẩn trong không gian Euclide

Cho $v = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$

$\|v\|$ ký hiệu **chuẩn** (độ dài) của véc tơ v , được cho bởi công thức

Chuẩn trong không gian Euclide

Cho $v = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$

$\|v\|$ ký hiệu **chuẩn** (độ dài) của véc tơ v , được cho bởi công thức

$$\|v\| = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2}.$$

Hình cầu & Ellipsoid

- **Hình cầu** tâm x_c bán kính $r > 0$ trong \mathbb{R}^n có dạng

$$\begin{aligned} B(x_c, r) &= \{x \in \mathbb{R}^n \mid \|x - x_c\| \leq r\} \\ &= \{x \in \mathbb{R}^n \mid (x - x_c)^T (x - x_c) \leq r^2\}. \end{aligned}$$

- Một **ellipsoid** là một tập hợp có dạng

$$E = \{x \in \mathbb{R}^n \mid (x - x_c)^T P (x - x_c) \leq 1\},$$

ở đó P là một ma trận đối xứng và nửa xác định dương cỡ n .

Polyhedron & Polytope

- Một **Polyhedron** (tập lồi đa diện) là tập nghiệm của một hệ gồm hữu hạn các đẳng thức và bất đẳng thức tuyến tính

$$P = \{x \in \mathbb{R}^n \mid a_j^T x \leq b_j, j = 1, \dots, m, c_j^T x = d_j, j = 1, \dots, p\}.$$

Polyhedron & Polytope

- Một **Polyhedron** (tập lồi đa diện) là tập nghiệm của một hệ gồm hữu hạn các đẳng thức và bất đẳng thức tuyến tính

$$P = \{x \in \mathbb{R}^n \mid a_j^T x \leq b_j, j = 1, \dots, m, c_j^T x = d_j, j = 1, \dots, p\}.$$

- Một tập lồi đa diện là giao hữu hạn các nửa không gian đóng.

Polyhedron & Polytope

- Một **Polyhedron** (tập lồi đa diện) là tập nghiệm của một hệ gồm hữu hạn các đẳng thức và bất đẳng thức tuyến tính

$$P = \{x \in \mathbb{R}^n \mid a_j^T x \leq b_j, j = 1, \dots, m, c_j^T x = d_j, j = 1, \dots, p\}.$$

- Một tập lồi đa diện là giao hữu hạn các nửa không gian đóng.
- Một tập lồi đa diện bị chặn đôi khi được gọi là một **polytope**.



regular
tetrahedron



cube



regular
octahedron



regular
dodecahedron



regular
icosahedron

Polyhedron & Polytope

- Một **Polyhedron** (tập lồi đa diện) là tập nghiệm của một hệ gồm hữu hạn các đẳng thức và bất đẳng thức tuyến tính

$$P = \{x \in \mathbb{R}^n \mid a_j^T x \leq b_j, j = 1, \dots, m, c_j^T x = d_j, j = 1, \dots, p\}.$$

- Một tập lồi đa diện là giao hữu hạn các nửa không gian đóng.
- Một tập lồi đa diện bị chặn đôi khi được gọi là một **polytope**.



regular
tetrahedron



cube



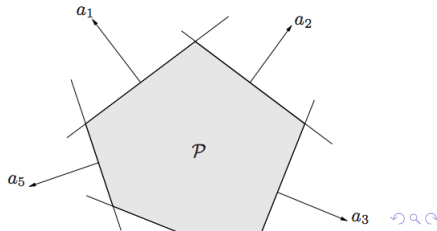
regular
octahedron



regular
dodecahedron



regular
icosahedron



Cho hàm số $f : \mathbb{R}^n \rightarrow \mathbb{R}$

- Miền hữu hiệu $\text{dom } f := \{x \in \mathbb{R}^n \mid f(x) < +\infty\}$

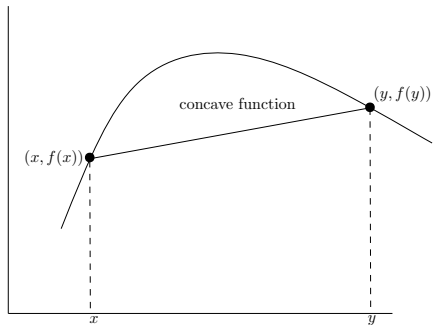
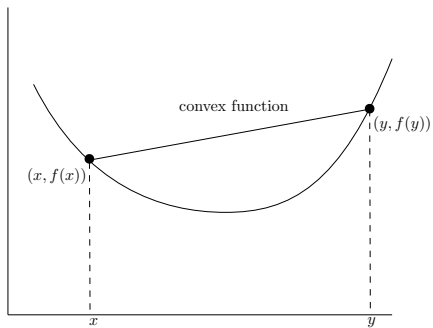
Định nghĩa

Hàm số $f : \mathbb{R}^n \rightarrow \mathbb{R}$ được gọi là hàm lồi nếu với mọi $x, y \in \text{dom } f$, $\theta \in [0, 1]$, bất đẳng thức Jensen thỏa mãn:

$$f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y).$$

f được gọi là

- **lồi chặt** nếu bất đẳng thức Jensen là chặt
- **lõm** nếu $-f$ là lồi
- **lõm chặt** nếu $-f$ lồi chặt



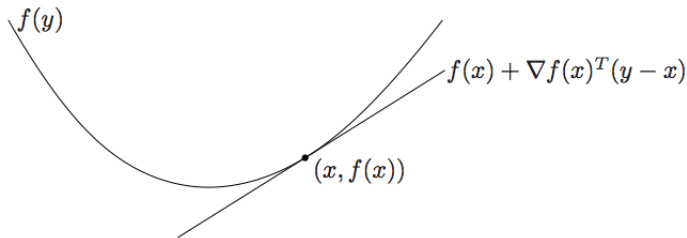
Đặc trưng tính lồi

Điều kiện cấp 1

Giả sử hàm số $f : \mathbb{R}^n \rightarrow \mathbb{R}$ khả vi. Khi đó, f là lồi khi và chỉ khi $\text{dom } f$ là lồi và

$$f(y) \geq f(x) + \nabla f(x)^T (y - x)$$

đúng với mọi $x, y \in \text{dom } f$.



Đạo hàm cấp hai (nhắc lại)

Ma trận Hessian

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

Ma trận con chính, định thức con chính

Ký hiệu H_k là ma trận con chính (ma trận con phía trên, bên trái) của ma trận Hessian H .

$$H_1 = \left[\frac{\partial^2 f}{\partial x_1^2} \right]; H_2 = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix}; H_3 = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \frac{\partial^2 f}{\partial x_3^2} \end{bmatrix} \dots$$

Ký hiệu $\Delta_1 = \det(H_1)$, $\Delta_2 = \det(H_2)$, $\Delta_3 = \det(H_3) \dots$

Ma trận con chính, định thức con chính

Ký hiệu H_k là ma trận con chính (ma trận con phía trên, bên trái) của ma trận Hessian H .

$$H_1 = \left[\frac{\partial^2 f}{\partial x_1^2} \right]; \quad H_2 = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix}; \quad H_3 = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \frac{\partial^2 f}{\partial x_3^2} \end{bmatrix} \dots$$

Ký hiệu $\Delta_1 = \det(H_1)$, $\Delta_2 = \det(H_2)$, $\Delta_3 = \det(H_3) \dots$

- Δ_k được gọi là định thức con chính cấp k của H

Điều kiện cấp 2

Giả sử hàm số $f : \mathbb{R}^n \rightarrow \mathbb{R}$ khả vi đến cấp 2. Khi đó, f là lồi *khi và chỉ khi* ma trận Hessian $\nabla^2 f(x)$ là nửa xác định dương, tức là

$$\nabla^2 f(x) \succeq 0 \quad \forall x \in \text{dom } f.$$

Trong trường hợp ma trận Hessian matrix là **xác định dương**, hàm số **lồi chặt**

Trong trường hợp ma trận Hessian matrix là **xác định dương**, hàm số **lồi chặt**

Trường hợp không gian một chiều

Giả sử hàm số $f : \mathbb{R} \rightarrow \mathbb{R}$ khả vi đến cấp 2. Khi đó, f là lồi khi và chỉ khi $\text{dom } f$ là lồi và $f''(x) \geq 0$ với mọi $x \in \text{dom } f$.

Ma trận xác định dương & nửa xác định dương

- Ma trận A được gọi là **xác định dương** (*positive definite*) nếu $x^T A x > 0$ với mọi vectơ $x \neq 0$. Ta ký hiệu $A \succ 0$.

Ma trận xác định dương & nửa xác định dương

- Ma trận A được gọi là **xác định dương** (*positive definite*) nếu $x^T A x > 0$ với mọi vectơ $x \neq 0$. Ta ký hiệu $A \succ 0$.
- Nếu $x^T A x \geq 0$ với mọi vectơ $x \in \mathbb{R}^n$ thì A được gọi là **nửa xác định dương** (*positive semidefinite*). Ta ký hiệu $A \succeq 0$.

Ma trận xác định dương & nửa xác định dương

- Ma trận A được gọi là **xác định dương** (*positive definite*) nếu $x^T A x > 0$ với mọi vectơ $x \neq 0$. Ta ký hiệu $A \succ 0$.
- Nếu $x^T A x \geq 0$ với mọi vectơ $x \in \mathbb{R}^n$ thì A được gọi là **nửa xác định dương** (*positive semidefinite*). Ta ký hiệu $A \succeq 0$.

Chú ý: xác định dương \rightarrow nửa xác định dương

Ma trận

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

là xác định dương vì $x^T A x = x_1^2 + x_2^2 > 0$ với mọi $x = (x_1, x_2) \neq (0, 0)$.

Hàm số $f(x_1, x_2) = x_1^2 + x_2^2$ là lồi

Đặc trưng ma trận đối xứng xác định dương

Cho A là ma trận đối xứng cỡ $n \times n$. Khi đó, A là **xác định dương khi và chỉ khi** một trong các điều kiện sau thỏa mãn

- Tất cả các giá trị riêng dương
- Tất cả các định thức con chính dương

Đặc trưng ma trận đối xứng nửa xác định dương

Cho A là ma trận đối xứng cỡ $n \times n$. Khi đó, A là **nửa xác định dương** khi và chỉ khi

- Tất cả các giá trị riêng **không âm**

Đặc trưng ma trận đối xứng nửa xác định dương

Cho A là ma trận đối xứng cỡ $n \times n$. Khi đó, A là **nửa xác định dương** khi và chỉ khi

- Tất cả các giá trị riêng **không âm**



Chú ý: Không thể dùng tiêu chuẩn qua dấu các định thức con chính để kiểm tra một ma trận đối xứng nửa xác định dương

Đặc trưng ma trận đối xứng nửa xác định dương

Cho A là ma trận đối xứng cỡ $n \times n$. Khi đó, A là **nửa xác định dương** khi và chỉ khi

- Tất cả các giá trị riêng **không âm**



Chú ý: Không thể dùng tiêu chuẩn qua dấu các định thức con chính để kiểm tra một ma trận đối xứng nửa xác định dương

Ma trận A nửa xác định dương có tính chất sau đây

- $\det(A) = 0$, tất cả các định thức con chính cấp nhỏ hơn cỡ của A không âm

Ví dụ: Hàm số $f(x_1, x_2) = x_1^2 + x_2^2$ là lồi (chặt) vì ma trận Hessian

$$H = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

là xác định dương

Các ví dụ trong không gian \mathbb{R}

- $f(x) = e^{ax}$ là hàm lồi trên \mathbb{R} với mọi $a \in \mathbb{R}$

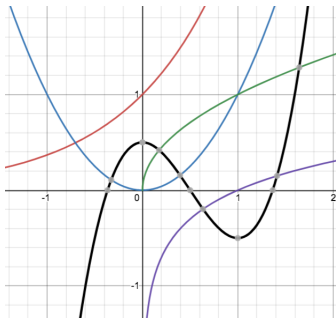
Các ví dụ trong không gian \mathbb{R}

- $f(x) = e^{ax}$ là hàm lồi trên \mathbb{R} với mọi $a \in \mathbb{R}$
- $f(x) = x^a$ là hàm lồi trên \mathbb{R}_+ nếu $a \geq 1$ hoặc $a \leq 0$, và lõm nếu $a \in [0, 1]$

Tối ưu lồi

Các ví dụ trong không gian \mathbb{R}

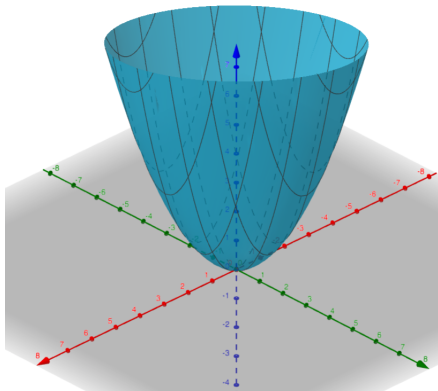
- $f(x) = e^{ax}$ là hàm lồi trên \mathbb{R} với mọi $a \in \mathbb{R}$
- $f(x) = x^a$ là hàm lồi trên \mathbb{R}_+ nếu $a \geq 1$ hoặc $a \leq 0$, và lõm nếu $a \in [0, 1]$
- $f(x) = \log x$ là hàm lõm trên \mathbb{R}_+



Ví dụ trong không gian 2 chiều

$f(x, y) = x^2 + xy + y^2$ là lồi trên \mathbb{R}^2 bởi vì

$$\nabla^2 f(x, y) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \text{ và } \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 2(x^2 + xy + y^2) \geq 0$$



Hàm toàn phương trong \mathbb{R}^n

Một hàm toàn phương (quadratic function) $f : \mathbb{R}^n \rightarrow \mathbb{R}$, với $\text{dom } f = \mathbb{R}^n$, được cho bởi

$$f(x) = (1/2)x^T P x + q^T x + r$$

ở đó P là ma trận đối xứng, $q \in \mathbb{R}^n$, và $r \in \mathbb{R}$. Do $\nabla^2 f(x) = P$, hàm số f là lồi **khi và chỉ khi** $P \succeq 0$.

Hàm toàn phương trong \mathbb{R}^n

Một hàm toàn phương (quadratic function) $f : \mathbb{R}^n \rightarrow \mathbb{R}$, với $\text{dom } f = \mathbb{R}^n$, được cho bởi

$$f(x) = (1/2)x^T P x + q^T x + r$$

ở đó P là ma trận đối xứng, $q \in \mathbb{R}^n$, và $r \in \mathbb{R}$. Do $\nabla^2 f(x) = P$, hàm số f là lồi **khi và chỉ khi** $P \succeq 0$.

Ví dụ

$$x^2 + xy + y^2 = (1/2) \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Hàm số $f(x, y) = x^2 + xy + y^2$ là lồi (chặt)

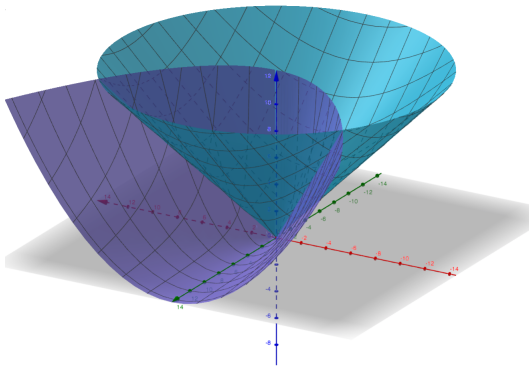
Hàm tuyến tính trong \mathbb{R}^n

$$f(x) = a_1x_1 + \cdots + a_nx_n + b$$

- f vừa là hàm lồi, vừa là hàm lõm

Một số ví dụ trong \mathbb{R}^2

- $f(x, y) = \sqrt{x^2 + y^2}$ là lồi trên \mathbb{R}^2
- $f(x, y) = x^2/y$ là lồi trên $\text{dom } f = \{(x, y) \in \mathbb{R}^2 \mid y > 0\}$



Graph & Epigraph

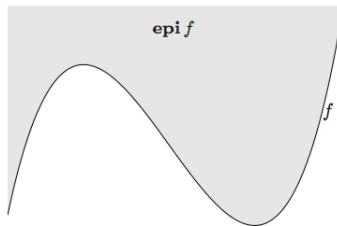
Cho $f : \mathbb{R}^n \rightarrow \mathbb{R}$

- Đồ thị (graph) của f là tập hợp được cho bởi

$$\{(x, f(x)) \mid x \in \text{dom } f\} \subseteq \mathbb{R}^{n+1}$$

- Trên đồ thị (epigraph) của f là tập hợp

$$\text{epi } f = \{(x, t) \mid x \in \text{dom } f, f(x) \leq t\}$$



Hàm số $f : \mathbb{R}^n \rightarrow \mathbb{R}$ là lồi khi và chỉ khi tập hợp $\text{epi } f$ là lồi.

Các phép toán bảo toàn tính lỗi

Các phép toán bảo toàn tính lỗi

- Nếu $f(x)$ là lỗi và $\alpha > 0$, thì $\alpha f(x)$ là lỗi
- Nếu $f(x), g(x)$ là lỗi thì $f(x) + g(x)$ là lỗi
- Nếu $f(x), g(x)$ là lỗi thì $h(x) = \max\{f(x), g(x)\}$ là lỗi

Bài tập

Xác định xem các hàm số sau là lồi hay lõm (giải thích)

- $f(x) = e^x - 1$ trên \mathbb{R} .
- $f(x, y) = xy$ trên $\mathbb{R} \times \mathbb{R}$.
- $f(x, y) = \frac{1}{xy}$ trên $\mathbb{R}_+ \times \mathbb{R}_+$, ở đó $\mathbb{R}_+ = (0, +\infty)$.
- $f(x, y) = \frac{x}{y}$ trên $\mathbb{R}_+ \times \mathbb{R}_+$, ở đó $\mathbb{R}_+ = (0, +\infty)$.