Mô hình hoá và tối ưu hoá cho học máy

Homework 1

▼ 1. Convex sets

▼ a. Closed sets and convex sets

▼ i. Chứng minh rằng: Một Polyhedron $\{x \in \mathbb{R}^n : Ax \leq b\}$, với $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$, vừa là tập lồi, vừa là tập đóng.

i. Show that a polyhedron $\{x \in \mathbb{R}^n : Ax \leq b\}$, for some $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$, is both convex and closed.

Giải

- Polyhedron $P = \{x \in \mathbb{R}^n : Ax \leq b\}$
- Chứng minh P là tập lồi

  Chọn $x_1, x_2 \in P \Rightarrow \begin{cases} Ax_1 \leq b \\ Ax_2 \leq b \end{cases}$

  Xét điểm $tx_1 + (1-t)x_2$:

  $$\begin{aligned} & A(tx_1 + (1-t)x_2) \\ & = tAx_1 + (1-t)Ax_2 \\ & \leq tb + (1-t)b \\ & = b \\ & \Rightarrow (tx_1 + (1-t)x_2) \in P \end{aligned}$$

  Vậy P là tập lồi.

- Chứng minh P là tập đóng

  $P = \{x \in \mathbb{R}^n : Ax \leq b\}$
  $\Rightarrow P$ là giao của các nửa không gian (halfspace)
  mà các nửa không gian là các tập đóng
  $\Rightarrow P$ là tập đóng.

▼ ii. Chứng minh rằng: Nếu $S_i \subseteq \mathbb{R}^n, i \in I$ là một tập của các tập lồi, thì phần giao của chúng $\bigcap_{i \in I} S_i$ cũng là tập lồi. Chứng minh điều tương tự nếu thay "lồi" bằng "đóng".

ii. Show that if $S_i \subseteq \mathbb{R}^n$, $i \in I$ is a collection of convex sets, then their intersection $\cap_{i \in I} S_i$ is also convex. Show that the same statement holds if we replace "convex" with "closed".

Giải

- Chứng minh: Nếu $S_i \subseteq \mathbb{R}^n, i \in I$ là một tập của các tập lồi, thì phần giao của chúng $\bigcap_{i \in I} S_i$ cũng là tập lồi.

  Chọn $s_1, s_2 \in \bigcap_{i \in I} S_i$
  $\Rightarrow s_1, s_2 \in S_i \quad \forall i \in I$
  mà S_i là tập lồi $\forall i \in I$
  $\Rightarrow ts_1 + (1-t)s_2 \in S_i \quad \forall i \in S_i$
  $\Rightarrow ts_1 + (1-t)s_2 \in \bigcap_{i \in I} S_i$
  $\Rightarrow \bigcap_{i \in I} S_i$ là tập lồi.

- Chứng minh: Nếu $S_i \subseteq \mathbb{R}^n, i \in I$ là một tập của các tập đóng, thì phần giao của chúng $\bigcap_{i \in I} S_i$ cũng là tập đóng.

  Chọn $\{s_n\}$ là chuỗi thuộc $\bigcap_{i \in I} S_i$.

  $\Rightarrow \{s_n\} \in S_i \ \ \forall i \in I$

  mà S_i là tập đóng

  $\Rightarrow \exists i \in I : s_0 = \lim_{n \to \infty} s_n \in S_i$

mà $\{s_n\} \in \bigcap_{i \in I} S_i$

$\Rightarrow s_0 \in \bigcap_{i \in I} S_i$

$\Rightarrow \bigcap_{i \in I} S_i$ là tập đóng.

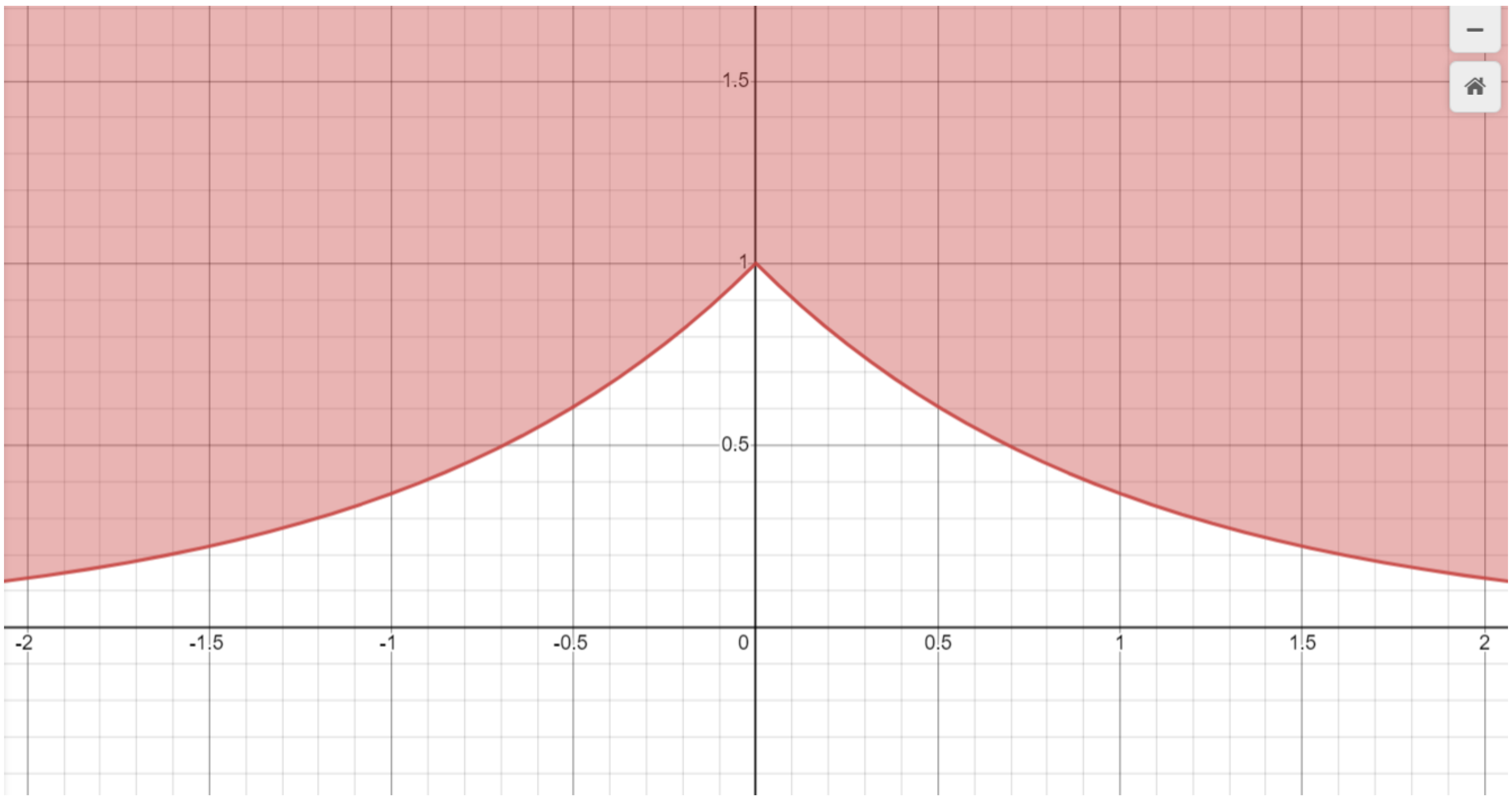
$\lim_{x \to 1} x = 1$

▼ iii. Lấy ví dụ về một tập đóng trong $\mathbb{R}^2$ mà bao lồi (convex hull) của nó không đóng.

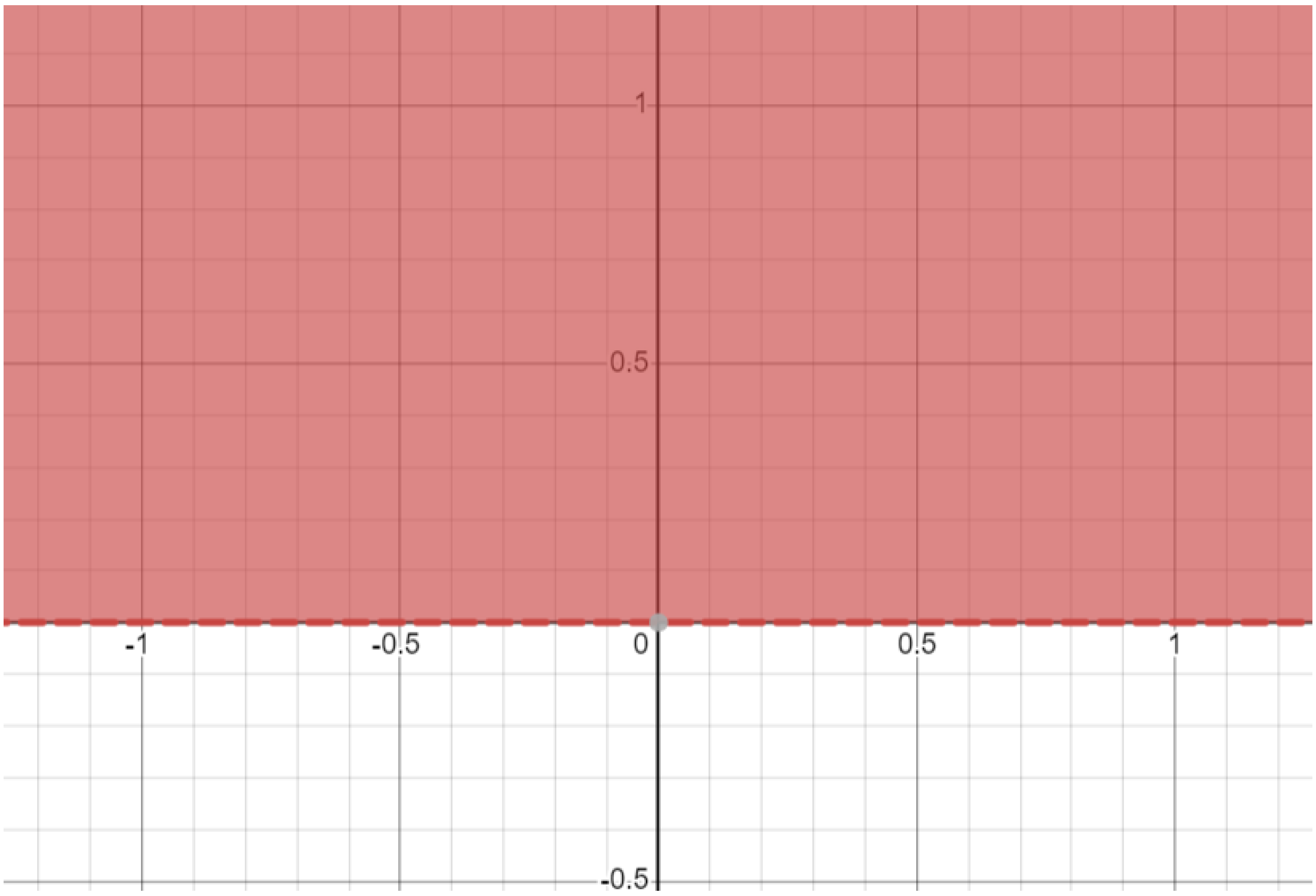
iii. Given an example of a closed set in $\mathbb{R}^2$ whose convex hull is not closed.

Giải

$C = \{(x, y) \in \mathbb{R}^2 : y \geq e^{-|x|}\}$



$\Rightarrow conv(C) = \{(x, y) \in \mathbb{R}^2 : y > 0\}$



▼ iv. Cho $A \in \mathbb{R}^{m \times n}$.
Chứng minh rằng: Nếu $S \subseteq \mathbb{R}^m$ là tập lồi, thì $A^{-1}(S) = \{x \in \mathbb{R}^n : Ax \in S\}$ cũng là tập lồi.
Chứng minh điều tương tự nếu thay lồi" bằng "đóng".

iv. Let $A \in \mathbb{R}^{m \times n}$. Show that if $S \subseteq \mathbb{R}^m$ is convex then so is $A^{-1}(S) = \{x \in \mathbb{R}^n : Ax \in S\}$, which is called the preimage of S under the map $A : \mathbb{R}^n \to \mathbb{R}^m$. Show that the same statement holds if we replace "convex" with "closed".

Giải

- Chứng minh: Nếu $S \subseteq \mathbb{R}^m$ là tập lồi, thì $A^{-1}(S) = \{x \in \mathbb{R}^n : Ax \in S\}$ cũng là tập lồi.

  Chọn $x_1, x_2 \in A^{-1}(S) \Rightarrow Ax_1, Ax_2 \in S$

  mà S là tập lồi

  $\Rightarrow tAx_1 + (1-t)Ax_2 \in S$

  $\Rightarrow A(tx_1 + (1-t)x_2) \in S$

  $\Rightarrow tx1 + (1-t)x_2 \in A^{-1}(S)$

  $\Rightarrow A^{-1}(S)$ là tập lồi.

- Chứng minh: Nếu $S \subseteq \mathbb{R}^m$ là tập đóng, thì $A^{-1}(S) = \{x \in \mathbb{R}^n : Ax \in S\}$ cũng là tập đóng.

  Chọn $T = \mathbb{R}^m \backslash S$. Do S là tập đóng $\Rightarrow T$ là tập mở

  $A^{-1}(T) = \{x \in \mathbb{R}^n : Ax \in T\}$
  $= \{x \in \mathbb{R}^n : Ax \notin S\}$
  $= \mathbb{R}^n \backslash A^{-1}(S)$

  Do T là tập mở $\Rightarrow A^{-1}(T)$ là tập mở $\Rightarrow A^{-1}(S)$ là tập đóng.

▼ v. Cho $A \in \mathbb{R}^{m \times n}$. Chứng minh rằng: Nếu $S \subseteq \mathbb{R}^n$ là tập lồi, thì $A(S) = \{Ax \in \mathbb{R}^m : x \in S\}$ cũng là tập lồi.

v. Let $A \in \mathbb{R}^{m \times n}$. Show that if $S \subseteq \mathbb{R}^n$ is convex then so is $A(S) = \{Ax : x \in S\}$, called the image of S under A.

Giải

Chọn $Ax_1, Ax_2 \in A(S) \Rightarrow x_1, x_2 \in S$

Do S là tập lồi $\Rightarrow tx_1 + (1-t)x_2 \in S$

Xét $tAx_1 + (1-t)Ax_2 = A(tx_1 + (1-t)x_2)$

Do $tx_1 + (1-t)x_2 \in S \Rightarrow A(tx_1 + (1-t)x_2) \in A(S)$

$\Rightarrow A(S)$ là tập lồi.

▼ vi. Lấy ví dụ về một ma trận $A \in \mathbb{R}^{m\times n}$ và một tập vừa đóng vừa lồi $S \subseteq \mathbb{R}^n$ nhưng $A(S)$ lại không đóng.

vi. Give an example of a matrix $A \in \mathbb{R}^{m\times n}$ and a set $S \subseteq \mathbb{R}^n$ that is closed and convex but such that $A(S)$ is not closed.

Giải

$C = \{(x,y) \in \mathbb{R}^2 : y \geq e^x\}$

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

$\Rightarrow A(C) = \{(0,y) \in \mathbb{R}^2 : y > 0\}$ là tập mở.

▼ b. Polyhedra.

▼ i. Chứng minh rằng: Nếu $p \subseteq \mathbb{R}^n$ là một Polydedron, thì với ma trận $A \in \mathbb{R}^{m\times n}$, $A(P)$ cũng là Polyhedron.

Gợi ý: Có thể sử dụng điều sau:

$P \subseteq \mathbb{R}^{m+n}$ là Polyhedron $\Rightarrow \{x \in \mathbb{R}^n : (x,y) \in P \ for \ some \ y \in \mathbb{R}^m\}$ cũng là Polyhedron.

i. Show that if $P \subseteq \mathbb{R}^n$ is a polyhedron, and $A \in \mathbb{R}^{m\times n}$, then $A(P)$ is a polyhedron. Hint: you may use the fact that

$$P \subseteq \mathbb{R}^{m+n} \text{ is a polyhedron } \Rightarrow \{x \in \mathbb{R}^n : (x,y) \in P \text{ for some } y \in \mathbb{R}^m\} \text{ is a polyhedron.}$$

Giải

Chọn $Q = \{(Ax, x) \in \mathbb{R}^{m+n} : x \in P\} \Rightarrow Q \subseteq \mathbb{R}^{m+n}$

Do $x \in P \Rightarrow M_P x \leq b_P$

Đặt $M_Q = [0 \quad M_P] \in \mathbb{R}^{m\times(m+n)}$

$\Rightarrow M_Q \begin{bmatrix} Ax \\ x \end{bmatrix} = M_P x \leq b_P$

$\Rightarrow Q$ là Polyhedron

$\Rightarrow \{Ax \in \mathbb{R}^m : (Ax, x) \in Q \ với \ x \in \mathbb{R}^n\}$ cũng là Polyhedron

$\Rightarrow A(P) = \{Ax \in \mathbb{R}^m : x \in P\}$ là Polyhedron

▼ ii. Chứng minh rằng: Nếu $Q \subseteq \mathbb{R}^m$ là Polyhedron, thì với ma trận $A \in \mathbb{R}^{m\times n}$, $A^{-1}(Q)$ cũng là Polyhedron.

ii. Show that if $Q \subseteq \mathbb{R}^m$ is a polyhedron, and $A \in \mathbb{R}^{m\times n}$, then $A^{-1}(Q)$ is a polyhedron.

Giải

Chọn $P = \{(x, Ax) \in \mathbb{R}^{n+m} : Ax \in Q\} \Rightarrow P \subseteq \mathbb{R}^{n+m}$

Do $Ax \in Q \Rightarrow M_Q(Ax) \leq b_Q$

Đặt $M_P = [0 \quad M_Q] \in \mathbb{R}^{n\times(n+m)}$

$\Rightarrow M_P \begin{bmatrix} x \\ Ax \end{bmatrix} = M_Q(Ax) \leq b_Q$

$\Rightarrow P$ là Polyhedron

$\Rightarrow \{x \in \mathbb{R}^n : (x, Ax) \in P \ với \ Ax \in \mathbb{R}^m\}$ cũng là Polyhedron

$\Rightarrow A^{-1}(Q) = \{x \in \mathbb{R}^n : Ax \in Q\}$ là Polyhedron

▼ 2. Convex functions

▼ a. Chứng minh rằng: Hàm số Entropy là hàm lõm chặt.

(a, 2 pts) Prove that the *entropy function*, defined as

$$f(x) = -\sum_{i=1}^{n} x_i \log(x_i),$$

with $\mathrm{dom}(f) = \{x \in \mathbb{R}^n_{++} : \sum_{i=1}^n x_i = 1\}$, is strictly concave.

Giải

Hàm số Entropy: $f(x) = -\sum\limits_{i=1}^{n} x_i \log(x_i)$

với $dom(f) = \{x \in \mathbb{R}^n_{++} : \sum\limits_{i=1}^{n} x_i = 1\}$

- Chứng minh $dom(f)$ là tập lồi

  Chọn $x_1, x_2 \in dom(f) \Rightarrow \begin{cases} \sum\limits_{i=1}^{n} x_{1i} = 1 \\ \sum\limits_{i=1}^{n} x_{2i} = 1 \end{cases}$

  Xét $tx_1 + (1-t)x_2 : \sum\limits_{i=1}^{n} tx_{1i} + (1-t)x_{2i} = t + 1 - t = 1$

  $\Rightarrow dom(f)$ là tập lồi.

- Chứng minh $\nabla^2 f(x) \prec 0$

  $\dfrac{\partial f(x)}{\partial x_i} = -\log(x_i) - 1 \Rightarrow \nabla f(x) = [-\log(x_i) - 1] \in \mathbb{R}^n$

  $\dfrac{\partial^2 f(x)}{\partial x_i} = -\dfrac{1}{x_i}$

  $\Rightarrow \nabla^2 f(x) = [a_{ij}] = \begin{cases} -\dfrac{1}{x_i} \;\; nếu \;\; i = j \\ 0 \;\; nếu \;\; i \neq j \end{cases} \in \mathbb{R}^{n \times n}$

  $\Rightarrow \nabla^2 f(x) \prec 0$

$\Rightarrow f(x)$ là hàm lõm chặt.

▼ b. Cho hàm số f có đạo hàm khả vi, với $dom(f)$ là tập lồi. Chứng minh rằng:
f là hàm lồi $\Leftrightarrow (\nabla f(x) - \nabla f(y))^T (x - y) \geq 0 \; \forall \; x, y$

(b, 4 pts) Let f be twice differentiable, with $\mathrm{dom}(f)$ convex. Prove that f is convex if and only if

$$(\nabla f(x) - \nabla f(y))^T (x - y) \geq 0,$$

for all x, y. This property is called *monotonicity* of the gradient ∇f.

Giải

- Chứng minh: f lồi $\Rightarrow (\nabla f(x) - \nabla f(y))^T (x - y) \geq 0 \; \forall \; x, y$

  Do f là hàm lồi$\Rightarrow \begin{cases} f(y) \geq f(x) + \nabla f(x)^T (y - x) & (1) \\ f(x) \geq f(y) + \nabla f(y)^T (x - y) & (2) \end{cases} \;\; \forall x, y$

  Cộng hai vế của (1) và (2)

  $\Rightarrow 0 \geq \nabla f(x)^T (y - x) + \nabla f(y)^T (x - y)$

  $\Leftrightarrow (\nabla f(x) - \nabla f(y))^T (x - y) \geq 0$

- Chứng minh: $(\nabla f(x) - \nabla f(y))^T (x - y) \geq 0 \; \forall \; x, y \Rightarrow f$ lồi

  Đặt $g(t) = f(tx + (1-t)y) \Rightarrow \begin{cases} g(0) = f(y) \\ g(1) = f(x) \end{cases} \;\; \forall t \in [0, 1]$

$g'(t) = \nabla f(tx + (1-t)y)^T(x-y)$

Chọn $0 \leq t_1 < t_2 \leq 1$:

$$g'(t_2) - g'(t_1)$$
$$= (\nabla f(t_2x + (1-t_2)y - \nabla f(t_1x + (1-t_1)y)^T(x-y)$$
$$= \frac{1}{t_2 - t_1}(\nabla f(t_2x + (1-t_2)y - \nabla f(t_1x + (1-t_1)y)^T(t_2x + (1-t_2)y - t_1x - (1-t_1)y) \geq 0$$

$\Rightarrow g'(t)$ đồng biến

$\Rightarrow g(t)$ là hàm lồi

$\Rightarrow g(t) \leq tg(1) + (1-t)g(0)$

$\Rightarrow f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$

$\Rightarrow f(x)$ là hàm lồi.

▼ c. Lấy ví dụ về một hàm lồi chặt nhưng không chứa điểm nhỏ nhất của nó.

(c, 2 pts) Give an example of a strictly convex function that does not attain its infimum.

Giải

Hàm $f(x) = e^x$. $f''(x) = e^x > 0 \Rightarrow f(x)$ là hàm lồi chặt

$\min_x f(x) = \lim_{x \to +\infty} e^x = 0$

▼ d. Một hàm số $f : \mathbb{R}^n \to \mathbb{R}$ được gọi là hàm "ép buộc" nếu $\lim_{\|x\|_2 \to \infty} f(x) = \infty$. Một tính chất của hàm "ép buộc" là nó chứa điểm nhỏ nhất của nó.

Cho f là một hàm lồi mạnh (strongly convex), có đạo hàm khả vi. Chứng minh f là một hàm "ép buộc".

(d, 3 pts) A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *coercive* provided that $f(x) \to \infty$ as $\|x\|_2 \to \infty$. A key fact about coercive functions is that they attain their infimums. Prove that a twice differentiable, strongly convex function is coercive and hence attains its infimum. Hint: use Q3 part (b.iv).

Giải

Một hàm số $f(x)$ được gọi là lồi mạnh nếu $f(x)$ lồi và $f(x) - \frac{m}{2}\|x\|_2^2$ cũng lồi.

Đặt $g(x) = f(x) - \frac{m}{2}\|x\|_2^2$. Do $g(x)$ lồi

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

$\Leftrightarrow f(y) - \frac{m}{2}\|y\|_2^2 \geq f(x) - \frac{m}{2}\|x\|_2^2 + (\nabla f(x) - mx)^T(y-x)$

Đặt $\begin{cases} \nabla f(x) - mx = \alpha \\ f(x) - \frac{m}{2}\|x\|_2^2 - \alpha^T x = c_x \end{cases}$

$\Rightarrow f(y) \geq \frac{m}{2}\|y\|_2^2 + \alpha^T y + c_x$

$\Rightarrow f(y) \geq \frac{m}{2}\|y\|_2^2 - \|\alpha\|_2 \cdot \|y\|_2 + c_x$

Do $\lim_{\|y\|_2 \to \infty} \frac{m}{2}\|y\|_2^2 - \|\alpha\|_2 \cdot \|y\|_2 + c_x = \infty$

$\Rightarrow \lim_{\|y\|_2 \to \infty} f(y) = \infty$

▼ e. "Chứng minh rằng cực đại của hàm lồi trên một khối đa diện bị chặn phải xảy ra tại một trong các đỉnh." -Google dịch-

(e, 3 pts) Prove that the maximum of a convex function over a bounded polyhedron must occur at one of the vertices. Hint: you may use the fact that a bounded polyhedron can be represented as the convex hull of its vertices.

Giải

- Trước hết, giả sử x^* không nằm trên cạnh của Polyhedron.

Chọn hai điểm x_1, x_2 nằm trên cạnh của Polyhedron sao cho x_1, x^*, x_2 thẳng hàng

$$\Rightarrow \begin{cases} \exists\, t \in [0,1] : x^* = tx_1 + (1-t)x_2 \\ f(x_1) < f(x^*) \\ f(x_2) < f(x^*) \end{cases}$$

$f(x^*)$
$= f(tx_1 + (1-t)x_2)$
$\le tf(x_1) + (1-t)f(x_2)$
$< tf(x^*) + (1-t)f(x^*)$
$= f(x^*)$
$\Rightarrow f(x^*) < f(x^*)$ (vô lí)
$\Rightarrow x^*$ nằm trên cạnh của Polyhedron.

- Giả sử x^* không nằm trên đỉnh của Polyhedron.

  Chọn hai điểm x_1, x_2 là hai đỉnh liền kề của Polyhedron sao cho x_1, x^*, x_2 thẳng hàng

$$\Rightarrow \begin{cases} \exists\, t \in [0,1] : x^* = tx_1 + (1-t)x_2 \\ f(x_1) < f(x^*) \\ f(x_2) < f(x^*) \end{cases}$$

$f(x^*)$
$= f(tx_1 + (1-t)x_2)$
$\le tf(x_1) + (1-t)f(x_2)$
$< tf(x^*) + (1-t)f(x^*)$
$= f(x^*)$
$\Rightarrow f(x^*) < f(x^*)$ (vô lí)
$\Rightarrow x^*$ nằm trên ít nhất một đỉnh của Polyhedron.

▼ 3. Partial optimization with l_2 penalties

Cho bài toán sau: $\min\limits_{\beta,\sigma\ge 0} f(\beta) + \frac{\lambda}{2}\sum\limits_{i=1}^{n} g(\beta_i, \sigma_i)$ (1)

với f là hàm lồi, $dom(f) = \mathbb{R}^n$, $\lambda \ge 0$ và

$$g(x,y) = \begin{cases} \dfrac{x^2}{y} + y & if\ y > 0 \\ 0 & if\ x = y = 0 \\ \infty & else. \end{cases}$$

Consider the problem

$$\min_{\beta,\,\sigma\ge 0}\ f(\beta) + \frac{\lambda}{2}\sum_{i=1}^{n} g(\beta_i, \sigma_i),$$

for some convex f with domain $\mathbb{R}^n$, $\lambda \ge 0$, and

$$g(x,y) = \begin{cases} x^2/y + y & \text{if } y > 0 \\ 0 & \text{if } x = 0,\ y = 0 \\ \infty & \text{else.} \end{cases}$$

In other words, the problem (1) is just the weighted ℓ_2 penalized problem

$$\min_{\beta,\,\sigma\ge 0}\ f(\beta) + \frac{\lambda}{2}\sum_{i=1}^{n}\left(\frac{\beta_i^2}{\sigma_i} + \sigma_i\right),$$

but being careful to treat the ith term in the sum as zero when $\beta_i = \sigma_i = 0$.

▼ a. Chứng minh g là hàm lồi, sau đó suy ra (1) là bài toán tối ưu lồi.

(a, 5 pts) Prove that g is convex. Hence argue that (1) is a convex problem. Note that this means we can perform partial optimization in (1) and expect it to return another convex problem. Hint: use the definition of convexity.

Giải

- Chứng minh g là hàm lồi.

$$\nabla g(x,y) = \begin{cases} \begin{bmatrix} \frac{2x}{y} \\ -\frac{x^2}{y^2} + 1 \end{bmatrix} & if\ y > 0 \\ 0 & if\ x = y = 0 \\ \infty & else. \end{cases}$$

$$\nabla^2 g(x,y) = \begin{cases} \begin{bmatrix} \frac{2}{y} & \frac{-2x}{y^2} \\ \frac{-2x}{y^2} & \frac{2x^2}{y^3} \end{bmatrix} & if\ y > 0 \\ 0 & if\ x = y = 0 \\ 0 & else. \end{cases}$$

Do $\nabla^2 g(x,y) \succeq 0 \Rightarrow g(x,y)$ làm hàm lồi.

- Chứng minh (1) là bài toán tối ưu lồi

  Do $g(x,y)$ là hàm lồi

  $\Rightarrow \frac{\lambda}{2} \sum_{i=1}^{n} g(\beta_i, \sigma_i)$ là hàm lồi

  $\Rightarrow f(\beta) + \frac{\lambda}{2} \sum_{i=1}^{n} g(\beta_i, \sigma_i)$ là hàm lồi

  $\Rightarrow$ (1) là bài toán tối ưu lồi.

▼ b. Chứng minh rằng $\min\limits_{y \geq 0} g(x,y) = |2x|$.

(b, 2 pts) Argue that $\min_{y \geq 0} g(x,y) = 2|x|$.

Giải

- TH1: $y > 0$:

  $g'_y(x,y) = -\frac{x^2}{y^2} + 1$

  $g''_y(x,y) = \frac{2x^2}{y^3}$

  Xét $g'_y(x,y) = 0 \Rightarrow y = |x|$

  Xét $g''_y(x,y) = g''_y(x,|x|) = \frac{2x^2}{|x|^3} = \frac{2}{|x|} > 0\ \forall x \neq 0$

  $\Rightarrow g(x)$ đạt cực tiểu tại $(x, |x|)$

  $\Rightarrow \min\limits_{y \geq 0} g(x,y) = g(x,|x|) = \frac{x^2}{|x|} + |x| = 2|x|$

  $\Rightarrow \min\limits_{y \geq 0} g(x,y) = 2|x|$

- TH2: $y = x = 0 \Rightarrow g(x,y) = 0 = 2|x|$
- TH3: Các trường hợp còn lại của $(x,y) \Rightarrow g(x,y) = \infty$

▼ c. Chứng minh rằng khi tối ưu (1) theo $\sigma \geq 0$ thì sẽ thu được bài toán tối ưu dạng: $\min\limits_{\beta} f(\beta) + \lambda \|\beta\|_1$.

(c, 3 pts) Argue that minimizing over $\sigma \geq 0$ in (1) gives the ℓ_1 penalized problem

$$\min_{\beta}\ f(\beta) + \lambda \|\beta\|_1.$$

Giải

$$\min_{\beta, \sigma \geq 0} f(\beta) + \frac{\lambda}{2} \sum_{i=1}^{n} g(\beta_i, \sigma_i)$$

$$= \min_{\beta} \left(f(\beta) + \frac{\lambda}{2} \sum_{i=1}^{n} \min_{\sigma \geq 0} g(\beta_i, \sigma_i)) \right)$$

$$= \min_{\beta} \left(f(\beta) + \frac{\lambda}{2} \sum_{i=1}^{n} 2|\beta_i| \right)$$
$$= \min_{\beta} \left(f(\beta) + \lambda \|\beta\|_1 \right).$$

▼ 4. Lipschitz gradients and strong convexity

Cho hàm lồi f có đạo hàm liên tục và khả vi.

Let f be convex and twice continuously differentiable.

▼ a. Chứng minh các mệnh đề sau là tương đương:

i. ∇f là Lipschitz với hằng số L.
ii. $(\nabla f(x) - \nabla f(y))^T(x-y) \leq L\|x-y\|_2^2 \; \forall x, y$.
iii. $\nabla^2 f(x) \preceq LI \; \forall x$.
iv. $f(y) \leq f(x) + \nabla f(x)^T(y-x) + \dfrac{L}{2}\|y-x\|_2^2 \; \forall x, y$.

(a, 10 pts) Show that the following statements are equivalent.

i. ∇f is Lipschitz with constant L;
ii. $(\nabla f(x) - \nabla f(y))^T(x-y) \leq L\|x-y\|_2^2$ for all x, y;
iii. $\nabla^2 f(x) \preceq LI$ for all x;
iv. $f(y) \leq f(x) + \nabla f(x)^T(y-x) + \frac{L}{2}\|y-x\|_2^2$ for all x, y.

Your solution should have 5 parts, where you prove i $\Rightarrow$ ii, ii $\Rightarrow$ iii, iii $\Rightarrow$ iv, iv $\Rightarrow$ ii, and iii $\Rightarrow$ i.

Giải

- Chứng minh $i \Rightarrow ii$.

  ∇f là Lipschitz với hằng số L

  $\Rightarrow \|\nabla f(x) - \nabla f(y)\|_2 \leq L\|x-y\|_2 \; \forall x, y$

  $\Rightarrow \|\nabla f(x) - \nabla f(y)\|_2 \cdot \|x-y\|_2 \leq L\|x-y\|_2^2 \; \forall x, y$

  mà $(\nabla f(x) - \nabla f(y))^T(x-y) \leq \|\nabla f(x) - \nabla f(y)\|_2 \cdot \|x-y\|_2$

  $\Rightarrow (\nabla f(x) - \nabla f(y))^T(x-y) \leq L\|x-y\|_2^2 \; \forall x, y$

- Chứng minh $ii \Rightarrow iii$.

  Đặt $g(t) = f(tx + (1-t)y) \Rightarrow \begin{cases} g(0) = f(y) \\ g(1) = f(x) \end{cases} \quad \forall t \in [0, 1]$

  $g'(t) = \nabla f(tx + (1-t)y)^T(x-y)$

  Chọn $0 \leq t_1 < t_2 \leq 1$:

  $g'(t_2) - g'(t_1)$
  $= (\nabla f(t_2 x + (1-t_2)y - \nabla f(t_1 x + (1-t_1)y)^T(x-y)$
  $= \dfrac{1}{t_2 - t_1}(\nabla f(t_2 x + (1-t_2)y) - \nabla f(t_1 x + (1-t_1)y)^T(t_2 x + (1-t_2)y - t_1 x - (1-t_1)y)$
  $\leq L(t_2 - t_1)\|x-y\|_2^2$

  $\Rightarrow \dfrac{g'(t_2) - g'(t_1)}{t_2 - t_1} \leq L\|x-y\|_2^2$

  Đặt $t_2 = t_1 + \Delta t$

  $\Rightarrow \lim\limits_{\Delta t \to 0} \dfrac{g'(t_1 + \Delta t) - g'(t_1)}{\Delta t} \leq \lim\limits_{\Delta t \to 0} L\|x-y\|_2^2$

  $\Rightarrow g''(t_1) \leq L\|x-y\|_2^2$

  Chọn $t_1 = 1$

  $\Rightarrow g''(1) = (x-y)^T \nabla^2 f(x)(x-y) \leq L\|x-y\|_2^2$

  $\Rightarrow \nabla^2 f(x) \preceq LI$

- Chứng minh $iii \Rightarrow iv$.

Đặt $g(x) = f(x) - \frac{L}{2}\|x\|_2^2$

$\Rightarrow \nabla^2 g(x) = \nabla^2 f(x) - LI \preceq 0$ (do $\nabla^2 f(x) \preceq LI$)

$\Rightarrow g(x)$ là hàm lõm

$\Rightarrow g(y) \leq g(x) + \nabla g(x)^T(y-x)$

$\Leftrightarrow f(y) - \frac{m}{2}\|y\|_2^2 \leq f(x) - \frac{L}{2}\|x\|_2^2 + (\nabla f(x) - Lx)^T(y-x)$

$\Rightarrow f(y) \leq f(x) + \nabla f(x)^T(y-x) + \frac{L}{2}\|y-x\|_2^2 \ \forall x, y$

- Chứng minh $iv \Rightarrow ii$.

  $f(y) \leq f(x) + \nabla f(x)^T(y-x) + \frac{L}{2}\|y-x\|_2^2 \ \forall x, y$

  $f(x) \leq f(y) + \nabla f(y)^T(x-y) + \frac{L}{2}\|x-y\|_2^2 \ \forall x, y$

  Cộng hai vế lại:

  $\Rightarrow (\nabla f(x) - \nabla f(y))^T(x-y) \leq L\|x-y\|_2^2 \ \forall x, y.$

- Chứng minh $iii \Rightarrow i$.

  Đặt $g(x) = f(x) - \frac{L}{2}\|x\|_2^2$

  $\Rightarrow g(x)$ là hàm lõm (do $\nabla^2 f(x) \preceq LI$)

  $\Rightarrow (\nabla g(x) - \nabla g(y))^T(x-y) \leq 0$

  $\Rightarrow (\nabla f(x) - Lx - \nabla f(y) + Ly)^T(x-y) \leq 0$

  $\Rightarrow (\nabla f(x) - \nabla f(y))^T(x-y) \leq L\|x-y\|_2^2$

  $\Rightarrow \nabla f(x)$ là hàm Lipschitz với hằng số L.

▼ b. Chứng minh các mệnh đề sau là tương đương:

i. f là hàm lồi mạnh (strongly convex) với hằng số m.

ii. $(\nabla f(x) - \nabla f(y))^T(x-y) \geq m\|x-y\|_2^2 \ \forall x, y.$

iii. $\nabla^2 f(x) \succeq mI \ \forall x.$

iv. $f(y) \geq f(x) + \nabla f(x)^T(y-x) + \frac{m}{2}\|y-x\|_2^2 \ \forall x, y.$

(b, 8 pts) Show that the following statements are equivalent.

i. f is strongly convex with constant m;

ii. $(\nabla f(x) - \nabla f(y))^T(x-y) \geq m\|x-y\|_2^2$ for all x, y;

iii. $\nabla^2 f(x) \succeq mI$ for all x;

iv. $f(y) \geq f(x) + \nabla f(x)^T(y-x) + \frac{m}{2}\|y-x\|_2^2$ for all x, y.

Your solution should have 4 parts, where you prove i $\Rightarrow$ ii, ii $\Rightarrow$ iii, iii $\Rightarrow$ iv, and iv $\Rightarrow$ i.

- Chứng minh $i \Rightarrow ii$.

  f là hàm lồi mạnh $\Rightarrow g = f - \frac{m}{2}\|x\|_2^2$ là hàm lồi

  $\Rightarrow (\nabla g(x) - \nabla g(y))^T(x-y) \geq 0 \ \forall x, y$

  $\Leftrightarrow (\nabla f(x) - mx - \nabla f(y) + my)^T(x-y) \geq 0 \ \forall x, y$

  $\Leftrightarrow (\nabla f(x) - \nabla f(y))^T(x-y) - m(x-y)^T(x-y) \geq 0 \ \forall x, y$

  $\Leftrightarrow (\nabla f(x) - \nabla f(y))^T(x-y) \geq m\|x-y\|_2^2 \ \forall x, y$

- Chứng minh $ii \Rightarrow iii$.

  Đặt $g(t) = f(tx + (1-t)y) \Rightarrow \begin{cases} g(0) = f(y) \\ g(1) = f(x) \end{cases} \ \forall t \in [0,1]$

  $g'(t) = \nabla f(tx + (1-t)y)^T(x-y)$

  Chọn $0 \leq t_1 < t_2 \leq 1$:

$g'(t_2) - g'(t_1)$
$= (\nabla f(t_2x + (1-t_2)y - \nabla f(t_1x + (1-t_1)y)^T(x-y)$
$= \dfrac{1}{t_2 - t_1}(\nabla f(t_2x + (1-t_2)y - \nabla f(t_1x + (1-t_1)y)^T(t_2x + (1-t_2)y - t_1x - (1-t_1)y)$
$\leq L(t_2 - t_1)\|x-y\|_2^2$

$\Rightarrow \dfrac{g'(t_2) - g'(t_1)}{t_2 - t_1} \geq m\|x-y\|_2^2$

Đặt $t_2 = t_1 + \Delta t$

$\Rightarrow \lim\limits_{\Delta t \to 0} \dfrac{g'(t_1 + \Delta t) - g'(t_1)}{\Delta t} \geq \lim\limits_{\Delta t \to 0} m\|x-y\|_2^2$

$\Rightarrow g''(t_1) \geq m\|x-y\|_2^2$

Chọn $t_1 = 0$

$\Rightarrow g''(1) = (x-y)^T\nabla^2 f(x)(x-y) \geq m\|x-y\|_2^2$

$\Rightarrow \nabla^2 f(x) \succeq mI \; \forall x$

- Chứng minh $iii \Rightarrow iv$.

  Đặt $g(x) = f(x) - \dfrac{m}{2}\|x\|_2^2$

  $\Rightarrow \nabla^2 g(x) = \nabla^2 f(x) - mI \succeq 0$ (do $\nabla^2 f(x) \succeq mI$)

  $\Rightarrow g(x)$ là hàm lồi

  $\Rightarrow f(x)$ là hàm lồi mạnh (strongly convex function)

  $\Rightarrow f(y) \geq f(x) + \nabla f(x)^T(y-x) + \dfrac{m}{2}\|y-x\|_2^2 \; \forall x, y$

- Chứng minh $iv \Rightarrow i$.

  - $f(y) \geq f(x) + \nabla f(x)^T(y-x) + \dfrac{m}{2}\|y-x\|_2^2 \; \forall x, y$ (1)
  - $\dfrac{m}{2}\|y-x\|_2^2 = \dfrac{m}{2}\|y\|_2^2 - mx^Ty + m\|x\|_2^2 - \dfrac{m}{2}\|x\|_2^2$
    $= \dfrac{m}{2}\|y\|_2^2 - mx^T(y-x) - \dfrac{m}{2}\|x\|_2^2$
  - (1) $\Rightarrow f(y) - \dfrac{m}{2}\|y\|_2^2 \geq f(x) - \dfrac{m}{2}\|x\|_2^2 + (\nabla f(x) - mx)^T(y-x)$ (2)
  - Đặt $g(x) = f(x) - \dfrac{m}{2}\|x\|_2^2 \Rightarrow \nabla g(x) = \nabla f(x) - mx$
  - (2) $\Rightarrow g(y) \geq g(x) + \nabla g(x)^T(y-x)$
    $\Rightarrow g(x)$ là hàm lồi
    $\Rightarrow f(x)$ là hàm lồi mạnh (strongly convex function)

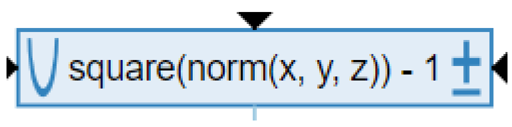
▼ 5. Solving optimization problems with CVX

▼ a. Bài làm tại file 5a.rar

▼ b. Chứng minh lồi và chuyển biểu thức về dạng DCP dựa theo link sau.

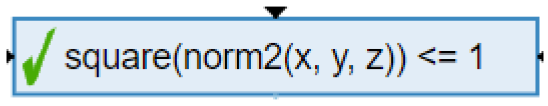
▼ i. $\|(x, y, z)\|_2^2 \leq 1$

- Chứng minh lồi

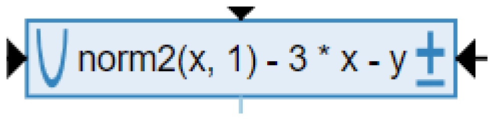


- Đưa về dạng DCP

  square(norm2(x, y, z)) ≤ 1

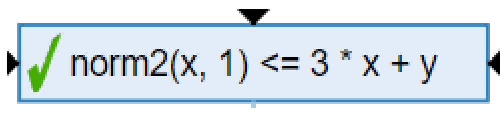


- ii. $\sqrt{x^2+1} \leq 3x+y$
  - Chứng minh lồi

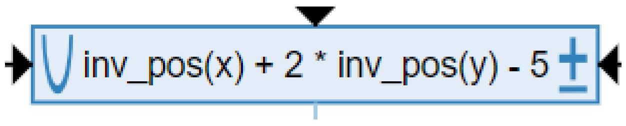


  - Đưa về dạng DCP

    norm2(x, 1) ≤ 3*x + y

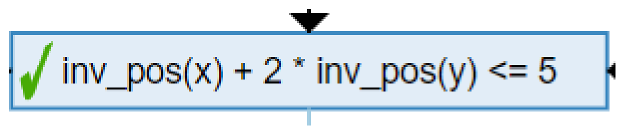


- iii. $\frac{1}{x}+\frac{2}{y} \leq 5, x>0, y>0$
  - Chứng minh lồi

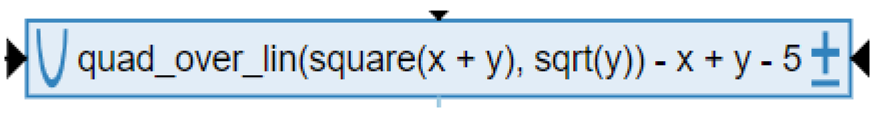


  - Đưa về dạng DCP

    inv_pos(x) + 2*inv_pos(y) ≤ 5

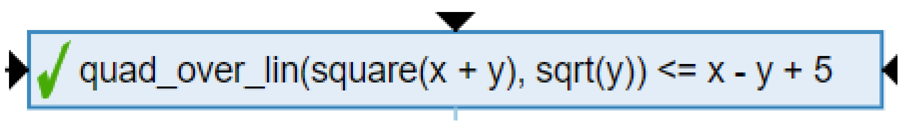


- iv. $\frac{(x+y)^2}{\sqrt{y}} \leq x-y+5, y>0$
  - Chứng minh lồi

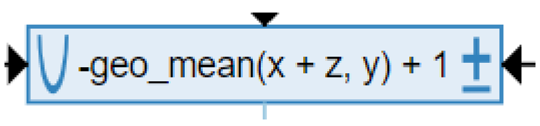


  - Đưa về dạng DCP

    quad_over_lin(square(x+y), sqrt(y)) ≤ x - y + 5

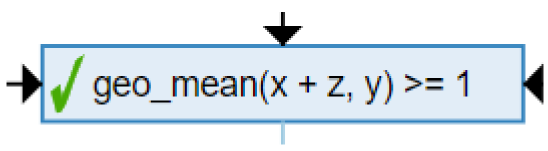


- v. $(x+z)y \geq 1, x+z \geq 0, y \geq 0$
  - Chứng minh lồi



  - Đưa về dạng DCP

    geo_mean(x+z, y) ≥ 1



▼ vi. $\|(x+2y, x-y)\|_2 = 0$

- Chứng minh lồi

  Xét $A = \{(x,y) : \|(x+2y, x-y)\|_2 = 0\}$

  $\|(x+2y, x-y)\|_2 = 0$

  $\Leftrightarrow (x+2y)^2 + (x-y)^2 = 0$

  $\Leftrightarrow 2x^2 + 5y^2 + 2xy = 0$

  Xét hàm số $f(x,y) = 2x^2 + 5y^2 + 2xy$

  Cực trị của $f(x,y)$ đạt tại (x^*, y^*) sao cho:

  $\begin{cases} f'_x(x^*, y^*) = 4x^* + 2y^* = 0 \\ f'_y(x^*, y^*) = 10y^* + 2x^* = 0 \end{cases}$

  $\Leftrightarrow \begin{cases} x = 0 \\ y = 0 \end{cases}$

  $\Rightarrow \min\limits_{x,y} f(x,y) = f(x^*, y^*) = f(0,0) = 0$

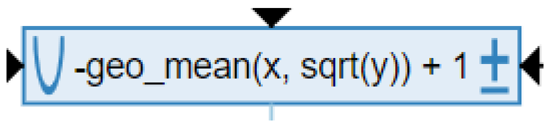
  $\Rightarrow A = \{(0,0)\} \Rightarrow A$ là tập lồi.

- Đưa về dạng DCP

  Bó tay.

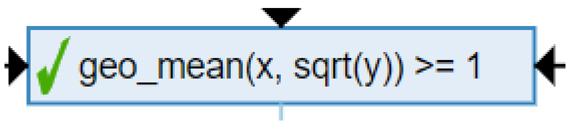
▼ vii. $x\sqrt{y} \geq 1, x \geq 0, y \geq 0$

- Chứng minh lồi



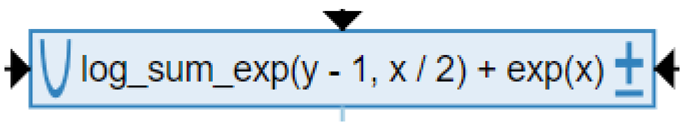
- Đưa về dạng DCP

  geo_mean(x, sqrt(y)) ≥ 1



▼ viii. $\log\left(e^{y-1} + e^{\frac{x}{2}}\right) \leq -e^x$

- Chứng minh lồi



- Đưa về dạng DCP

  log_sum_exp(y-1, x/2) ≤ -exp(x)