最优化方法

东南大学
计算机&人工智能学院
宋沫飞
songmf@seu.edu.cn

凸函数

- 基本性质和案例
- 保凸运算
- 共轭函数
- 拟凸函数
- 对数-凹函数和对数-凸函数
- 关于广义不等式的凸性

定义

定义

- $f : \mathbf{R}^n \to \mathbf{R}$ 为凸函数，则其满足定义域为凸集，且

$$f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta) f(y)$$

  - 对所有 $x, y \in \mathbf{dom}\, f$, $0 \le \theta \le 1$

定义

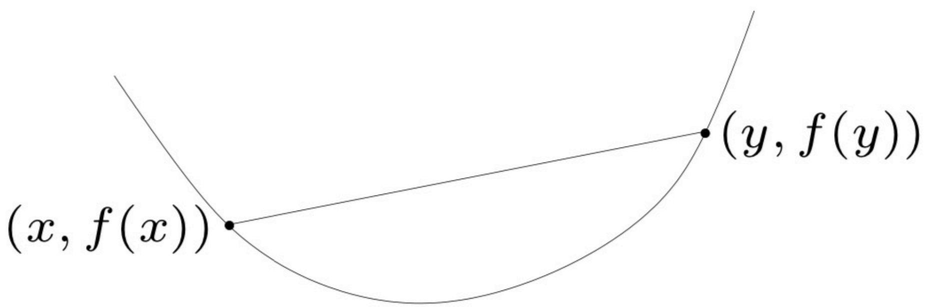
- $f : \mathbf{R}^n \to \mathbf{R}$ 为凸函数，则其满足定义域为凸集，且
$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$$
  - 对所有 $x, y \in \mathbf{dom}\, f$, $0 \leq \theta \leq 1$

定义

- $f : \mathbf{R}^n \to \mathbf{R}$ 为凸函数，则其满足定义域为凸集，且

$$f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta) f(y)$$

  - 对所有 $x, y \in \mathbf{dom}\, f$, $0 \le \theta \le 1$



- 若f为凹函数，则$-f$为凸函数
- 仿射函数

严格凸函数

严格凸函数

- f为严格凸函数，则其满足定义域为凸集，且

$$f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta)f(y)$$

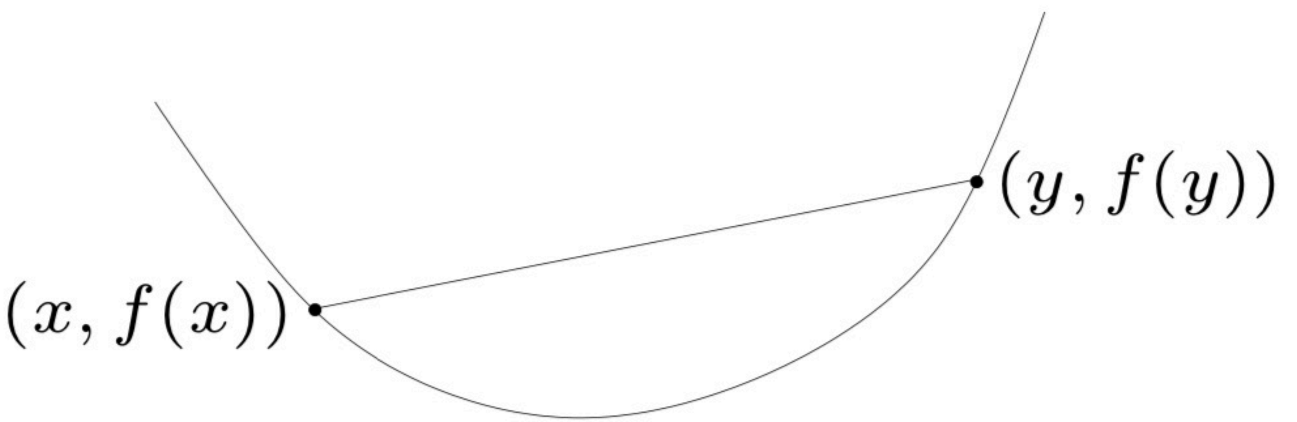
  - 对所有$x, y \in \mathbf{dom}\, f$, $x \neq y$, $0 < \theta < 1$

严格凸函数

- f为严格凸函数，则其满足定义域为凸集，且

$$f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta)f(y)$$

  - 对所有 $x, y \in \mathbf{dom}\, f,\ x \neq y,\ 0 < \theta < 1$

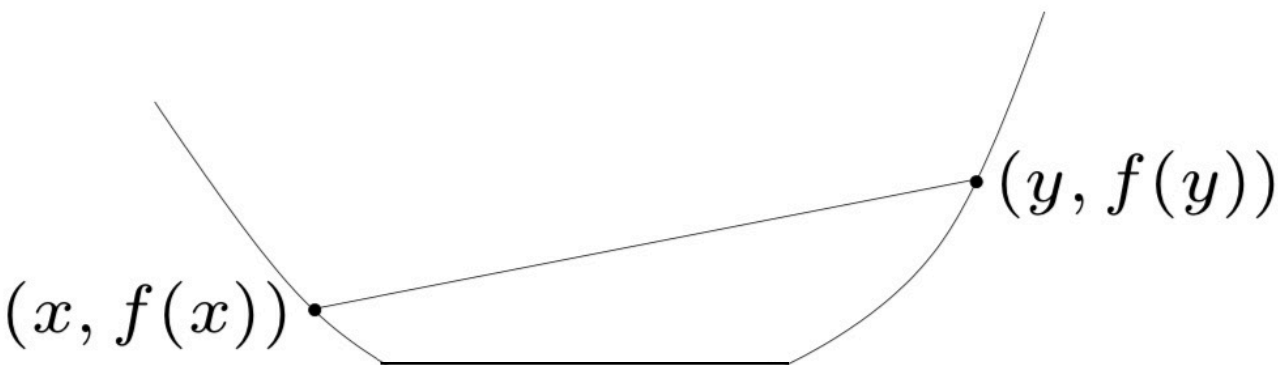


严格凸函数

- f为严格凸函数，则其满足定义域为凸集，且

$$f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta)f(y)$$

  - 对所有 $x, y \in \mathbf{dom}\, f,\ x \neq y,\ 0 < \theta < 1$



约束凸函数到直线

约束凸函数到直线

- $f : \mathbf{R}^n \to \mathbf{R}$ 为凸函数，当且仅当函数 $g : \mathbf{R} \to \mathbf{R}$

$$g(t) = f(x + tv), \quad \mathbf{dom}\, g = \{t \mid x + tv \in \mathbf{dom}\, f\}$$

为凸函数，对任意 $x \in \mathbf{dom}\, f$, $v \in \mathbf{R}^n$

约束凸函数到直线

- $f:\mathbf{R}^n\rightarrow\mathbf{R}$ 为凸函数，当且仅当函数 $g:\mathbf{R}\rightarrow\mathbf{R}$

$$g(t)=f(x+tv),\quad \mathbf{dom}\, g=\{t\mid x+tv\in\mathbf{dom}\, f\}$$

  为凸函数，对任意 $x\in\mathbf{dom}\, f,\ v\in\mathbf{R}^n$
- 可将凸函数的检测转化为单变量函数凸性的检测

约束凸函数到直线

- $f:\mathbf{R}^n \to \mathbf{R}$为凸函数，当且仅当函数 $g:\mathbf{R}\to\mathbf{R}$

$$g(t)=f(x+tv), \quad \mathbf{dom}\, g=\{t \mid x+tv\in\mathbf{dom}\, f\}$$

为凸函数，对任意 $x\in\mathbf{dom}\, f$, $v\in\mathbf{R}^n$

- 可将凸函数的检测转化为单变量函数凸性的检测
- 例：$f:\mathbf{S}^n\to\mathbf{R}$ 此处：$f(X)=\log\det X$, $\mathbf{dom}\, f=\mathbf{S}^n_{++}$

约束凸函数到直线

- $f:\mathbf{R}^n\to\mathbf{R}$ 为凸函数，当且仅当函数 $g:\mathbf{R}\to\mathbf{R}$

$$g(t)=f(x+tv),\quad \mathbf{dom}\,g=\{t\mid x+tv\in\mathbf{dom}\,f\}$$

为凸函数，对任意 $x\in\mathbf{dom}\,f,\ v\in\mathbf{R}^n$

- 可将凸函数的检测转化为单变量函数凸性的检测
- 例：$f:\mathbf{S}^n\to\mathbf{R}$ 此处：$f(X)=\log\det X,\ \mathbf{dom}\,f=\mathbf{S}^n_{++}$

$$\begin{aligned} g(t)=\log\det(X+tV) &= \log\det X+\log\det(I+tX^{-1/2}VX^{-1/2}) \\ &= \log\det X+\sum_{i=1}^{n}\log(1+t\lambda_i) \end{aligned}$$

约束凸函数到直线

- $f: \mathbf{R}^n \to \mathbf{R}$ 为凸函数，当且仅当函数 $g: \mathbf{R} \to \mathbf{R}$

$$g(t) = f(x+tv), \quad \mathbf{dom}\, g = \{t \mid x+tv \in \mathbf{dom}\, f\}$$

为凸函数，对任意 $x \in \mathbf{dom}\, f$, $v \in \mathbf{R}^n$

- 可将凸函数的检测转化为单变量函数凸性的检测
- 例：$f: \mathbf{S}^n \to \mathbf{R}$ 此处：$f(X) = \log\det X$, $\mathbf{dom}\, f = \mathbf{S}^n_{++}$

$$\begin{aligned} g(t) = \log\det(X+tV) &= \log\det X + \log\det(I + tX^{-1/2}VX^{-1/2}) \\ &= \log\det X + \sum_{i=1}^{n} \log(1+t\lambda_i) \end{aligned}$$

  - 此处，λ_i 是 $X^{-1/2}VX^{-1/2}$ 的特征值

约束凸函数到直线

- $f : \mathbf{R}^n \to \mathbf{R}$为凸函数，当且仅当函数 $g : \mathbf{R} \to \mathbf{R}$

$$g(t) = f(x+tv), \quad \mathbf{dom}\, g = \{t \mid x + tv \in \mathbf{dom}\, f\}$$

为凸函数，对任意 $x \in \mathbf{dom}\, f$, $v \in \mathbf{R}^n$

- 可将凸函数的检测转化为单变量函数凸性的检测
- 例：$f : \mathbf{S}^n \to \mathbf{R}$ 此处：$f(X) = \log\det X$, $\mathbf{dom}\, f = \mathbf{S}^n_{++}$

$$\begin{aligned} g(t) = \log\det(X+tV) &= \log\det X + \log\det(I + tX^{-1/2}VX^{-1/2}) \\ &= \log\det X + \sum_{i=1}^{n} \log(1+t\lambda_i) \end{aligned}$$

  - 此处，λ_i 是 $X^{-1/2}VX^{-1/2}$ 的特征值
  - 由于函数g 为凹函数，因此f 也为凹函数。

扩展值延伸

扩展值延伸

- 函数f的扩展值延伸$\tilde{f}$表示：

扩展值延伸

- 函数f的扩展值延伸$\tilde{f}$表示：

$$\tilde{f}(x) = f(x), \quad x \in \mathbf{dom}\, f, \qquad \tilde{f}(x) = \infty, \quad x \notin \mathbf{dom}\, f$$

扩展值延伸

- 函数f的扩展值延伸$\tilde{f}$表示：

$$\tilde{f}(x) = f(x), \quad x \in \mathbf{dom}\, f, \qquad \tilde{f}(x) = \infty, \quad x \notin \mathbf{dom}\, f$$

- 据此，可简化凸函数的表示:

扩展值延伸

- 函数f的扩展值延伸$\tilde{f}$表示：

$$\tilde{f}(x) = f(x), \quad x \in \mathbf{dom}\, f, \qquad \tilde{f}(x) = \infty, \quad x \notin \mathbf{dom}\, f$$

- 据此，可简化凸函数的表示:

$$0 \le \theta \le 1 \quad \Longrightarrow \quad \tilde{f}(\theta x + (1-\theta)y) \le \theta \tilde{f}(x) + (1-\theta)\tilde{f}(y)$$

扩展值延伸

- 函数f的扩展值延伸$\tilde{f}$表示：

$$\tilde{f}(x) = f(x), \quad x \in \mathbf{dom}\, f, \qquad \tilde{f}(x) = \infty, \quad x \notin \mathbf{dom}\, f$$

- 据此，可简化凸函数的表示：

$$0 \le \theta \le 1 \quad \Longrightarrow \quad \tilde{f}(\theta x + (1-\theta)y) \le \theta \tilde{f}(x) + (1-\theta)\tilde{f}(y)$$

- 等价于如下两个表达式：
  - **1.** 函数f的定义域为凸集；
  - **2.** 对 $x, y \in \mathbf{dom}\, f$,

$$0 \le \theta \le 1 \quad \Longrightarrow \quad f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta)f(y)$$

扩展值延伸

扩展值延伸

- 例：凸集的示性函数是凸函数

扩展值延伸

- 例：凸集的示性函数是凸函数

$$I_C(x) = \begin{cases} 0 & x \in C \\ \text{无定义} & x \notin C \end{cases}$$

扩展值延伸

- 例： 凸集的示性函数是凸函数

$$I_C(x) = \begin{cases} 0 & x \in C \\ \text{无定义} & x \notin C \end{cases}$$

$$\tilde{I}_C(x) = \begin{cases} 0 & x \in C \\ \infty & x \notin C \end{cases}$$

扩展值延伸

- 例： 凸集的示性函数是凸函数

$$I_C(x) = \begin{cases} 0 & x \in C \\ \text{无定义} & x \notin C \end{cases}$$

$$\tilde{I}_C(x) = \begin{cases} 0 & x \in C \\ 1 & x \notin C \end{cases}$$

一阶条件

一阶条件

- 函数f可微，其满足函数f的定义域是开集，且梯度

一阶条件

- 函数f可微，其满足函数f的定义域是开集，且梯度

$$\nabla f(x)=\left(\frac{\partial f(x)}{\partial x_1},\frac{\partial f(x)}{\partial x_2},\ldots,\frac{\partial f(x)}{\partial x_n}\right)$$

  - 在定义域处处存在

一阶条件

- 函数f可微，其满足函数f的定义域是开集，且梯度

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \ldots, \frac{\partial f(x)}{\partial x_n}\right)$$

  - 在定义域处处存在
- 一阶条件：定义域为凸集的可微函数f为凸函数，当且仅当

一阶条件

- 函数f可微，其满足函数f的定义域是开集，且梯度

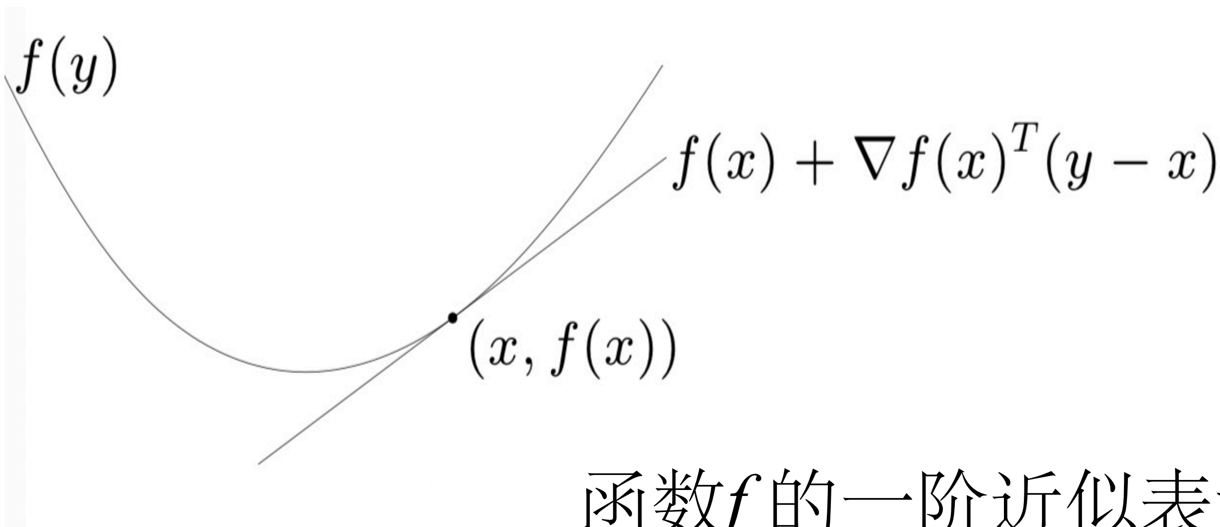
$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \ldots, \frac{\partial f(x)}{\partial x_n} \right)$$

  - 在定义域处处存在

- 一阶条件：定义域为凸集的可微函数f为凸函数，当且仅当

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) \quad \text{for all } x, y \in \mathbf{dom}\, f$$

一阶条件



函数f的一阶近似表示全局下估计

一阶条件的证明

一阶条件的证明

- 考虑一维情况

一阶条件的证明

- 考虑一维情况
  - f: $R \to R$为凸$\Longleftrightarrow$定义域为凸，且

一阶条件的证明

- 考虑一维情况
  - f: $R \to R$为凸$\Leftrightarrow$定义域为凸，且

$$f(y) \geq f(x) + f'(x)(y - x)$$

一阶条件的证明

- 考虑一维情况
  - f: $R \to R$为凸$\Leftrightarrow$定义域为凸，且

$$f(y) \geq f(x) + f'(x)(y-x)$$

- 证明$\Rightarrow$

一阶条件的证明

- 考虑一维情况
  - f: $R \to R$为凸$\Leftrightarrow$定义域为凸，且

$$f(y) \geq f(x) + f'(x)(y - x)$$

- 证明$\Rightarrow$
  - f为凸，对定义域内的任意两点x和y，实数t

一阶条件的证明

- 考虑一维情况
  - f: $R \rightarrow R$为凸$\Longleftrightarrow$定义域为凸，且

$$f(y) \geq f(x) + f'(x)(y - x)$$

- 证明$\Rightarrow$
  - f为凸，对定义域内的任意两点x和y，实数t

$$0 < t \leq 1,\ x + t(y - x) \in \mathbf{dom}\, f$$

一阶条件的证明

- 考虑一维情况
  - f: $R\rightarrow R$为凸$\Longleftrightarrow$定义域为凸，且

$$f(y) \geq f(x) + f'(x)(y-x)$$

- 证明$\Rightarrow$
  - f为凸，对定义域内的任意两点x和y，实数t

$$0 < t \leq 1,\ x + t(y-x) \in \mathbf{dom}\, f$$

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

一阶条件的证明

- 考虑一维情况
  - f: $R\rightarrow R$为凸$\Longleftrightarrow$定义域为凸，且

$$f(y) \geq f(x) + f'(x)(y-x)$$

- 证明$\Rightarrow$
  - f为凸，对定义域内的任意两点x和y，实数t

$$0 < t \leq 1,\ x + t(y-x) \in \mathbf{dom}\, f$$

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

$$tf(y) \geq tf(x) + f(x+t(y-x)) - f(x)$$

一阶条件的证明

- 考虑一维情况
  - f: $R\rightarrow R$为凸$\Longleftrightarrow$定义域为凸，且

$$f(y) \geq f(x) + f'(x)(y-x)$$

- 证明$\Rightarrow$
  - f为凸，对定义域内的任意两点x和y，实数t

$$0 < t \leq 1,\ x + t(y-x) \in \mathbf{dom}\, f$$

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

$$tf(y) \geq tf(x) + f(x+t(y-x)) - f(x)$$

$$f(y) \geq f(x) + \frac{f(x+t(y-x)) - f(x)}{t}$$

一阶条件的证明

- 考虑一维情况
  - f: $R \rightarrow R$为凸$\Longleftrightarrow$定义域为凸，且

$$f(y) \geq f(x) + f'(x)(y-x)$$

- 证明$\Rightarrow$
  - f为凸，对定义域内的任意两点x和y，实数t

$$0 < t \leq 1,\ x + t(y-x) \in \mathbf{dom}\, f$$

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

$$tf(y) \geq tf(x) + f(x + t(y-x)) - f(x)$$

$$\lim_{t \rightarrow 0+} \quad f(y) \geq f(x) + \frac{f(x+t(y-x)) - f(x)}{t}$$

一阶条件的证明

- 考虑一维情况
  - f: $R \rightarrow R$为凸$\Longleftrightarrow$定义域为凸，且

$$f(y) \geq f(x) + \boxed{f'(x)(y-x)}$$

- 证明$\Rightarrow$
  - f为凸，对定义域内的任意两点x和y，实数t

$$0 < t \leq 1,\ x + t(y-x) \in \mathbf{dom}\, f$$

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

$$tf(y) \geq tf(x) + f(x + t(y-x)) - f(x)$$

$$\lim_{t \rightarrow 0+} \quad f(y) \geq f(x) + \boxed{\frac{f(x+t(y-x)) - f(x)}{t}}$$

一阶条件的证明

一阶条件的证明

- 证明$\Leftarrow$

一阶条件的证明

- 证明$\Leftarrow$
  - 设定义域两个点 $x \neq y$以及 $0 \leq \theta \leq 1$

一阶条件的证明

- 证明$\Leftarrow$
  - 设定义域两个点 $x \neq y$以及 $0 \leq \theta \leq 1$
  - 通过凸组合构造 $z = \theta x + (1-\theta) y$

一阶条件的证明

- 证明$\Leftarrow$
  - 设定义域两个点 $x \neq y$以及 $0 \leq \theta \leq 1$
  - 通过凸组合构造 $z = \theta x + (1-\theta)y$

$$f(x) \geq f(z) + f'(z)(x-z)$$

一阶条件的证明

- 证明$\Leftarrow$
  - 设定义域两个点 $x \neq y$以及 $0 \leq \theta \leq 1$
  - 通过凸组合构造 $z = \theta x + (1-\theta)y$

$$f(x) \geq f(z) + f'(z)(x-z)$$

$$f(y) \geq f(z) + f'(z)(y-z)$$

一阶条件的证明

- 证明$\Leftarrow$
  - 设定义域两个点 $x \neq y$以及 $0 \leq \theta \leq 1$
  - 通过凸组合构造 $z = \theta x + (1-\theta) y$

$$\boxed{\begin{array}{c} f(x) \\ f(y) \end{array}} \begin{array}{l} \geq f(z) + f'(z)(x-z) \\ \geq f(z) + f'(z)(y-z) \end{array}$$

一阶条件的证明

- 证明$\Leftarrow$
  - 设定义域两个点$x \neq y$以及$0 \leq \theta \leq 1$
  - 通过凸组合构造$z = \theta x + (1-\theta)y$

$$\boxed{f(x)} \geq f(z) + f'(z)(x-z)$$

$$\boxed{f(y)} \geq f(z) + f'(z)(y-z)$$

$$\theta f(x) + (1-\theta) f(y) \geq f(z) + f'(z)\big(\theta(x-z) + (1-\theta)(y-z)\big)$$

一阶条件的证明

- 证明$\Leftarrow$
  - 设定义域两个点 $x \neq y$以及 $0 \leq \theta \leq 1$
  - 通过凸组合构造 $z = \theta x + (1-\theta) y$

$$\boxed{\begin{array}{c} f(x) \\ f(y) \end{array}} \begin{array}{l} \geq f(z) + f'(z)(x-z) \\ \geq f(z) + f'(z)(y-z) \end{array}$$

$$\theta f(x) + (1-\theta) f(y) \geq f(z) + f'(z)\left(\boxed{\begin{array}{l}\theta(x-z) \\ +(1-\theta)(y-z)\end{array}}\right)$$

一阶条件的证明

- 证明$\Leftarrow$
  - 设定义域两个点 $x \neq y$以及 $0 \leq \theta \leq 1$
  - 通过凸组合构造 $z = \theta x + (1-\theta) y$

$$\boxed{\begin{matrix} f(x) \\ f(y) \end{matrix}} \begin{matrix} \geq f(z) + f'(z)(x-z) \\ \geq f(z) + f'(z)(y-z) \end{matrix}$$

$$\theta f(x) + (1-\theta) f(y) \geq f(z) + f'(z)\left(\boxed{\begin{matrix} \theta(x-z) \\ +(1-\theta)(y-z) \end{matrix}}\right)$$

$$\boxed{\theta x + (1-\theta) y - z}$$

一阶条件的证明

一阶条件的证明

- 高维情况

一阶条件的证明

- 高维情况
- 证明$\Rightarrow$
  - 设f为凸，考虑定义域内两点x和y，构造函数g

一阶条件的证明

- 高维情况
- 证明⇒
  - 设f为凸，考虑定义域内两点x和y，构造函数g

$$g(t) = f(ty + (1-t)x)$$

一阶条件的证明

- 高维情况
- 证明⇒
  - 设f为凸，考虑定义域内两点x和y，构造函数g

$$g(t) = f(\boxed{ty + (1-t)x}) \rightarrow \boxed{x + t(y-x)}$$

一阶条件的证明

- 高维情况
- 证明⇒
  - 设f为凸，考虑定义域内两点x和y，构造函数g

$$g(t) = f(\boxed{ty + (1-t)x}) \rightarrow \boxed{x + t(y - x)}$$

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

一阶条件的证明

- 高维情况
- 证明⇒
  - 设f为凸，考虑定义域内两点x和y，构造函数g

$$g(t) = f(\boxed{ty + (1-t)x}) \rightarrow \boxed{x + t(y-x)}$$

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

$$g(t) \geq g(\tilde{t}) + g'(\tilde{t})(t - \tilde{t})$$

一阶条件的证明

- 高维情况
- 证明⇒
  - 设f为凸，考虑定义域内两点x和y，构造函数g

$$g(t) = f(\boxed{ty + (1-t)x}) \rightarrow \boxed{x + t(y-x)}$$

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

$$g(t) \geq g(\tilde{t}) + g'(\tilde{t})(t-\tilde{t})$$

$$g(1) \geq g(0) + g'(0)$$

一阶条件的证明

- 高维情况
- 证明⇒
  - 设f为凸，考虑定义域内两点x和y，构造函数g

$$g(t) = f(\boxed{ty + (1-t)x}) \rightarrow \boxed{x + t(y-x)}$$

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

$$g(t) \geq g(\tilde{t}) + g'(\tilde{t})(t - \tilde{t})$$

$$g(1) \geq g(0) + g'(0)$$

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

一阶条件的证明

一阶条件的证明

- 证明$\Leftarrow$
  - 考虑定义域内两点x和y

一阶条件的证明

- 证明⇐
  - 考虑定义域内两点x和y

$$ty + (1-t)x \in \mathbf{dom}\, f$$

$$\tilde{t}y + (1-\tilde{t})x \in \mathbf{dom}\, f$$

一阶条件的证明

- 证明⇐
  - 考虑定义域内两点*x*和*y*

$$ty + (1-t)x \in \mathbf{dom}\, f$$

$$\tilde{t}y + (1-\tilde{t})x \in \mathbf{dom}\, f$$

$$f(ty+(1-t)x) \geq f(\tilde{t}y+(1-\tilde{t})x) + \nabla f(\tilde{t}y+(1-\tilde{t})x)^T(y-x)(t-\tilde{t})$$

一阶条件的证明

- 证明⇐
  - 考虑定义域内两点*x*和*y*

$$ty+(1-t)x \in \textbf{dom}\, f$$

$$\tilde{t}y+(1-\tilde{t})x \in \textbf{dom}\, f$$

$$f(ty+(1-t)x) \geq f(\tilde{t}y+(1-\tilde{t})x) + \nabla f(\tilde{t}y+(1-\tilde{t})x)^T \boxed{(y-x)(t-\tilde{t})}$$

一阶条件的证明

- 证明⇐
  - 考虑定义域内两点*x*和*y*

$$ty + (1-t)x \in \mathbf{dom}\, f$$

$$\tilde{t}y + (1-\tilde{t})x \in \mathbf{dom}\, f$$

$$f(ty+(1-t)x) \geq f(\tilde{t}y+(1-\tilde{t})x) + \nabla f(\tilde{t}y+(1-\tilde{t})x)^T(y-x)(t-\tilde{t})$$

一阶条件的证明

- 证明⇐
  - 考虑定义域内两点*x*和*y*

$$ty+(1-t)x \in \mathbf{dom}\, f$$

$$\tilde{t}y+(1-\tilde{t})x \in \mathbf{dom}\, f$$

$$g(t) = f(ty+(1-t)x)$$

$$f(ty+(1-t)x) \ge f(\tilde{t}y+(1-\tilde{t})x) + \nabla f(\tilde{t}y+(1-\tilde{t})x)^T(y-x)(t-\tilde{t})$$

一阶条件的证明

- 证明⇐
  - 考虑定义域内两点*x*和*y*

$$ty + (1-t)x \in \mathbf{dom}\, f$$

$$\tilde{t}y + (1-\tilde{t})x \in \mathbf{dom}\, f$$

$$g(t) = f(ty + (1-t)x)$$

$$f(ty + (1-t)x) \geq f(\tilde{t}y + (1-\tilde{t})x) + \nabla f(\tilde{t}y + (1-\tilde{t})x)^T (y-x)(t-\tilde{t})$$

$$g(t) \geq g(\tilde{t}) + g'(\tilde{t})(t-\tilde{t})$$

一阶条件的证明

- 证明$\Leftarrow$
  - 考虑定义域内两点x和y

$$ty + (1-t)x \in \mathbf{dom}\, f$$

$$\tilde{t}y + (1-\tilde{t})x \in \mathbf{dom}\, f$$

$$g(t) = f(ty + (1-t)x)$$

$$\boxed{f(ty + (1-t)x)} \geq f(\tilde{t}y + (1-\tilde{t})x) + \nabla f(\tilde{t}y + (1-\tilde{t})x)^T (y-x)(t-\tilde{t})$$

$$\boxed{g(t)} \geq g(\tilde{t}) + g'(\tilde{t})(t-\tilde{t})$$

一阶条件的证明

- 证明$\Leftarrow$
  - 考虑定义域内两点x和y

$$ty + (1-t)x \in \mathbf{dom}\, f \qquad g(t) = f(ty + (1-t)x)$$

$$\tilde{t}y + (1-\tilde{t})x \in \mathbf{dom}\, f$$

$$\boxed{f(ty + (1-t)x)} \geq \boxed{f(\tilde{t}y + (1-\tilde{t})x)} + \nabla f(\tilde{t}y + (1-\tilde{t})x)^T (y-x)(t-\tilde{t})$$

$$\boxed{g(t)} \geq \boxed{g(\tilde{t})} + g'(\tilde{t})(t-\tilde{t})$$

一阶条件的证明

- 证明⇐
  - 考虑定义域内两点x和y

$$ty+(1-t)x \in \mathbf{dom}\, f \qquad g(t)=f(ty+(1-t)x)$$

$$\tilde{t}y+(1-\tilde{t})x \in \mathbf{dom}\, f$$

$$\boxed{f(ty+(1-t)x)} \geq \boxed{f(\tilde{t}y+(1-\tilde{t})x)} + \boxed{\nabla f(\tilde{t}y+(1-\tilde{t})x)^T(y-x)}(t-\tilde{t})$$

$$\boxed{g(t)} \geq \boxed{g(\tilde{t})} + \boxed{g'(\tilde{t})}(t-\tilde{t})$$

二阶条件

二阶条件

- 函数f为二阶可微，其满足函数f的定义域是开集，且**Hessian**矩阵处处存在$\nabla^2 f(x) \in \mathbf{S}^n$，

二阶条件

- 函数f为二阶可微，其满足函数f的定义域是开集，且**Hessian**矩阵处处存在$\nabla^2 f(x) \in \mathbf{S}^n$，

$$\nabla^2 f(x)_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}, \quad i, j = 1, \ldots, n,$$

二阶条件

- 函数f为二阶可微，其满足函数f的定义域是开集，且**Hessian**矩阵处处存在$\nabla^2 f(x) \in \mathbf{S}^n$,

$$\nabla^2 f(x)_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}, \quad i, j = 1, \ldots, n,$$

- 二阶条件
  - 对定义域为凸集的二阶可微函数f，若其为凸函数，当且仅当

二阶条件

- 函数f为二阶可微，其满足函数f的定义域是开集，且**Hessian**矩阵处处存在$\nabla^2 f(x) \in \mathbf{S}^n$，

$$\nabla^2 f(x)_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}, \quad i, j = 1, \ldots, n,$$

- 二阶条件
  - 对定义域为凸集的二阶可微函数f，若其为凸函数，当且仅当

$$\nabla^2 f(x) \succeq 0 \quad \text{for all } x \in \textbf{dom}\, f$$

二阶条件

- 函数f为二阶可微，其满足函数f的定义域是开集，且**Hessian**矩阵处处存在$\nabla^2 f(x) \in \mathbf{S}^n$,

$$\nabla^2 f(x)_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}, \quad i, j = 1, \ldots, n,$$

- 二阶条件
  - 对定义域为凸集的二阶可微函数f，若其为凸函数，当且仅当

$$\nabla^2 f(x) \succeq 0 \quad \text{for all } x \in \mathbf{dom}\, f$$

- 若$\nabla^2 f(x) \succ 0$ for all $x \in \mathbf{dom}\, f$，则其为严格凸函数

例

例

- 二次函数

例

- 二次函数

$$f(x) = (1/2)x^T P x + q^T x + r \text{ (with } P \in \mathbf{S}^n\text{)}$$

例

- 二次函数

$$f(x) = (1/2)x^T P x + q^T x + r \quad (\text{with } P \in \mathbf{S}^n)$$

$$\nabla f(x) = Px + q,$$

例

- 二次函数

$$f(x) = (1/2)x^T Px + q^T x + r \quad (\text{with } P \in \mathbf{S}^n)$$

$$\nabla f(x) = Px + q,$$

$$\nabla^2 f(x) = P$$

例

- 二次函数

$$f(x) = (1/2)x^T P x + q^T x + r \quad (\text{with } P \in \mathbf{S}^n)$$

$$\nabla f(x) = Px + q,$$

$$\nabla^2 f(x) = P$$

- 若$P \succeq 0$，则为凸函数

例

- 二次函数

$$f(x) = (1/2)x^T P x + q^T x + r \ \text{(with } P \in \mathbf{S}^n\text{)}$$

$$\nabla f(x) = Px + q,$$

$$\nabla^2 f(x) = P$$

- 若$P \succeq 0$，则为凸函数
- 问题：
  - 严格凸时，P是否一定为正定矩阵?

例

例

$$f(x) = 1/x^2$$

例

$$f(x) = 1/x^2$$

$$\mathbf{dom}\, f = \{x \in \mathbf{R} \mid x \neq 0\}$$

例

$$f(x) = 1/x^2$$

$$\mathbf{dom}\, f = \{x \in \mathbf{R} \mid x \neq 0\}$$

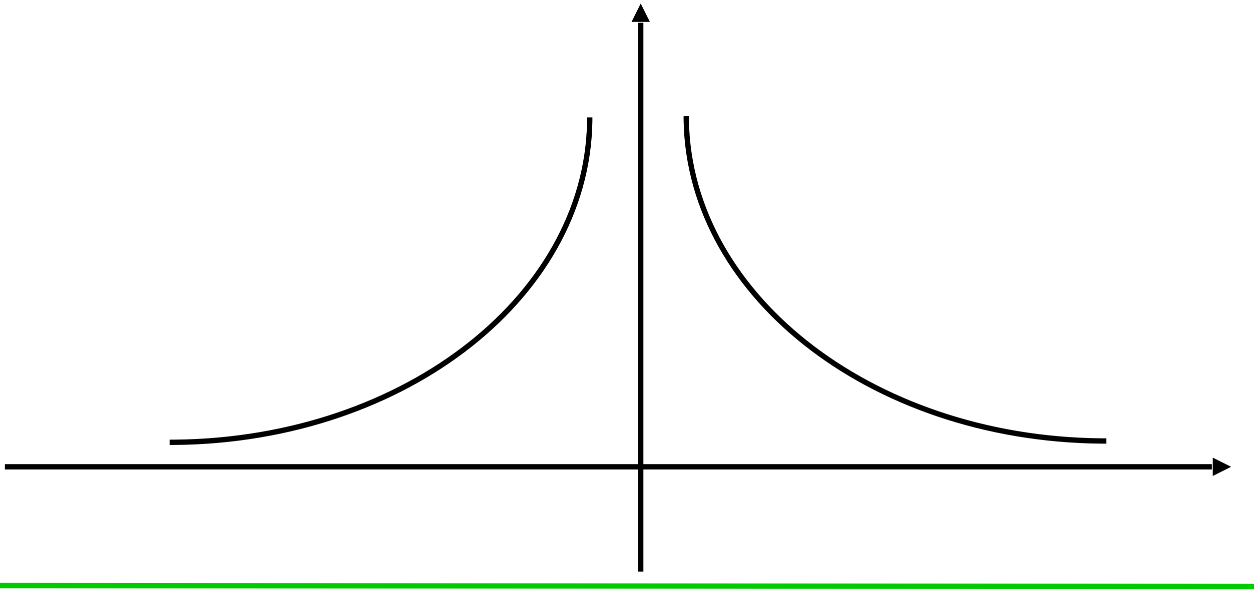
$$f''(x) > 0 \quad (6x^{-4})$$

例

$$f(x) = 1/x^2$$

$$\mathbf{dom}\, f = \{x \in \mathbf{R} \mid x \neq 0\}$$

$$f''(x) > 0 \quad (\mathbf{6}x^{\mathbf{-4}})$$



例

例

- 仿射函数 $ax+b$ $\nabla^2 f(x)=0$

例

- 仿射函数 $ax+b$ $\nabla^2 f(x)=0$
- 指数函数 e^{ax}

例

- 仿射函数 $ax+b$ $\nabla^2 f(x)=0$
- 指数函数 e^{ax}
  - $f'(x)=ae^{ax}$

例

- 仿射函数 $ax+b$ $\nabla^2 f(x) = 0$
- 指数函数 e^{ax}
  - $f'(x) = ae^{ax}$
  - $f''(x) = a^2 e^{ax}$

例

- 仿射函数 $ax+b$ $\nabla^2 f(x)=0$
- 指数函数 e^{ax}
  - $f'(x)=ae^{ax}$
  - $f''(x)=a^2e^{ax}$
- 幂函数：x^a，x为正实数

例

- 仿射函数 $ax+b$ $\nabla^2 f(x)=0$
- 指数函数 e^{ax}
  - $f'(x)=ae^{ax}$
  - $f''(x)=a^2e^{ax}$
- 幂函数：x^a，x为正实数
  - $f'(x)=ax^{a-1}$

例

- 仿射函数 $ax+b$ $\nabla^2 f(x)=0$
- 指数函数 e^{ax}
  - $f'(x)=ae^{ax}$
  - $f''(x)=a^2e^{ax}$
- 幂函数：x^a，x为正实数
  - $f'(x)=ax^{a-1}$
  - $f''(x)=a(a-1)x^{a-2}$

例

- 仿射函数 $ax+b$ $\nabla^2 f(x)=0$
- 指数函数 e^{ax}
  - $f'(x)=ae^{ax}$
  - $f''(x)=a^2e^{ax}$
- 幂函数： x^a ， x为正实数
  - $f'(x)=ax^{a-1}$
  - $f''(x)=a(a-1)x^{a-2}$
  - $f''(x)=\begin{cases}\geq 0 & \textbf{if } a\geq 1 \ or \ a\leq 0\\ \leq 0 & \textbf{if } 0\leq a\leq 1\end{cases}$

例

例

- 绝对值的幂函数 $|x|^p$

例

- 绝对值的幂函数 $|x|^p$
  - $$f'(x)=\begin{cases} px^{p-1} & \text{if } x\geq 0 \\ -p(-x)^{p-1} & \text{if } x<0 \end{cases}$$

例

- 绝对值的幂函数 $|x|^p$
  - $$f'(x) = \begin{cases} px^{p-1} & \text{if } x \geq 0 \\ -p(-x)^{p-1} & \text{if } x < 0 \end{cases}$$
  - $$f''(x) = \begin{cases} p(p-1)x^{p-2} & \textbf{if } x \geq 0 \\ p(p-1)(-x)^{p-2} & \textbf{if } x < 0 \end{cases}$$

例

- 绝对值的幂函数 $|x|^p$
  - $f'(x) = \begin{cases} px^{p-1} & \text{if } x \geq 0 \\ -p(-x)^{p-1} & \text{if } x < 0 \end{cases}$
  - $f''(x) = \begin{cases} p(p-1)x^{p-2} & \textbf{if } x \geq 0 \\ p(p-1)(-x)^{p-2} & \textbf{if } x < 0 \end{cases}$
  - $p > 1$ 为凸

例

- 绝对值的幂函数 $|x|^p$
  - $$f'(x) = \begin{cases} px^{p-1} & \text{if } x \geq 0 \\ -p(-x)^{p-1} & \text{if } x < 0 \end{cases}$$
  - $$f''(x) = \begin{cases} p(p-1)x^{p-2} & \textbf{if } x \geq 0 \\ p(p-1)(-x)^{p-2} & \textbf{if } x < 0 \end{cases}$$
  - $p > 1$ 为凸
  - $p = 1$ 为凸

例

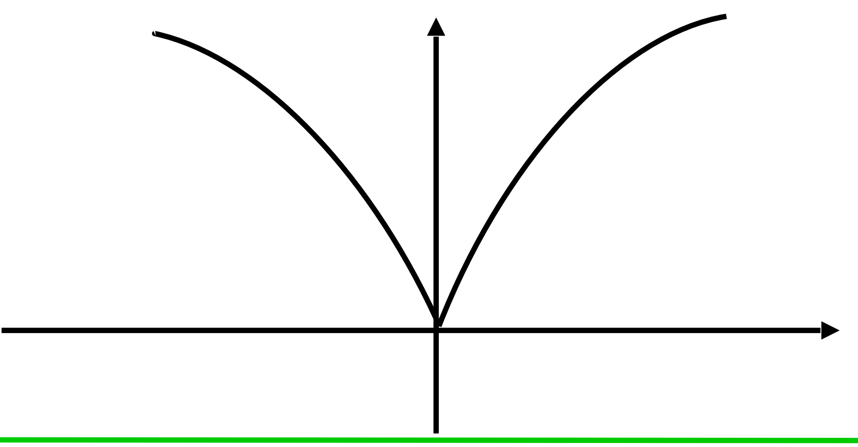
- 绝对值的幂函数$|x|^p$
  - $$f'(x) = \begin{cases} px^{p-1} & \text{if } x \ge 0 \\ -p(-x)^{p-1} & \text{if } x < 0 \end{cases}$$
  - $$f''(x) = \begin{cases} p(p-1)x^{p-2} & \textbf{if } x \ge 0 \\ p(p-1)(-x)^{p-2} & \textbf{if } x < 0 \end{cases}$$
  - $p > 1$ 为凸
  - $p = 1$ 为凸
  - $p < 1$ **??**

例

- 绝对值的幂函数$|x|^p$
  - $$f'(x) = \begin{cases} px^{p-1} & \text{if } x \geq 0 \\ -p(-x)^{p-1} & \text{if } x < 0 \end{cases}$$
  - $$f''(x) = \begin{cases} p(p-1)x^{p-2} & \textbf{if } x \geq 0 \\ p(p-1)(-x)^{p-2} & \textbf{if } x < 0 \end{cases}$$
  - $p > 1$ 为凸
  - $p = 1$ 为凸
  - $p < 1$ **??**
    - $p = 0.5$

例

- 绝对值的幂函数 $|x|^p$
  - $f'(x) = \begin{cases} px^{p-1} & \text{if } x \geq 0 \\ -p(-x)^{p-1} & \text{if } x < 0 \end{cases}$
  - $f''(x) = \begin{cases} p(p-1)x^{p-2} & \textbf{if } x \geq 0 \\ p(p-1)(-x)^{p-2} & \textbf{if } x < 0 \end{cases}$
  - $p > 1$ 为凸
  - $p = 1$ 为凸
  - $p < 1$??
    - $p = 0.5$



例

例

- 对数函数 $\log x$，x为正实数

例

- 对数函数 $\log x$，x为正实数
  - $f'(x) = 1/x$

例

- 对数函数 $\log x$, x为正实数
  - $f'(x) = 1/x$
  - $f''(x) = -1/x^2 < 0$

例

- 对数函数 $\log x$，x为正实数
  - $f'(x) = 1/x$
  - $f''(x) = -1/x^2 < 0$
- 负熵 $x \log x$，x为正实数

例

- 对数函数 $\log x$, x为正实数
  - $f'(x) = 1/x$
  - $f''(x) = -1/x^2 < 0$
- 负熵 $x\log x$, x为正实数
  - $f'(x) = logx + 1$

例

- 对数函数 $\log x$, x为正实数
  - $f'(x) = 1/x$
  - $f''(x) = -1/x^2 < 0$
- 负熵 $x \log x$, x为正实数
  - $f'(x) = log x + 1$
  - $f''(x) = 1/x > 0$

例

例

- 范数：$\mathbf{R}^n$空间的范数$P(x)$

例

- 范数：$\mathbf{R}^n$空间的范数$P(\mathbf{x})$
  - $P(ax) = |a|\ P(x)$
  - $P(x + y) \le P(x) + P(y)$
  - $P(x) = 0 \Longleftrightarrow x = 0$

例

- 范数：$\mathbf{R}^n$空间的范数$P(\mathbf{x})$
  - $P(ax) = |a|\ P(x)$
  - $P(x + y) \le P(x) + P(y)$
  - $P(x) = 0 \Longleftrightarrow x = 0$

$$f(\theta x + (1 - \theta)y) \le f(\theta x) + f((1 - \theta)y)$$

例

- 范数：$\mathbf{R}^n$空间的范数$P(x)$
  - $P(ax) = |a|\ P(x)$
  - $P(x + y) \le P(x) + P(y)$
  - $P(x) = 0 \Longleftrightarrow x = 0$

$$
\begin{aligned}
f(\theta x + (1-\theta)y) &\le f(\theta x) + f((1-\theta)y) \\
&= \theta f(x) + (1-\theta) f(y)
\end{aligned}
$$

例

- 范数：$\mathbf{R}^n$空间的范数$P(x)$
  - $P(ax) = |a|\ P(x)$
  - $P(x + y) \le P(x) + P(y)$
  - $P(x) = 0 \Longleftrightarrow x = 0$

$$
\begin{aligned}
f(\theta x + (1-\theta)y) &\le f(\theta x) + f((1-\theta)y) \\
&= \theta f(x) + (1-\theta)f(y)
\end{aligned}
$$

- 零范数：计算非零元素的数目

例

- 范数：$\mathbf{R}^n$空间的范数$P(x)$
  - $P(ax) = |a|\, P(x)$
  - $P(x + y) \leq P(x) + P(y)$
  - $P(x) = 0 \Longleftrightarrow x = 0$

$$
\begin{aligned}
f(\theta x + (1-\theta)y) &\leq f(\theta x) + f((1-\theta)y) \\
&= \theta f(x) + (1-\theta) f(y)
\end{aligned}
$$

- 零范数：计算非零元素的数目
  - 令n=1

例

- 范数：$\mathbf{R}^n$空间的范数$P(\boldsymbol{x})$
  - $P(ax) = |a|\ P(x)$
  - $P(x + y) \leq P(x) + P(y)$
  - $P(x) = 0 \Longleftrightarrow x = 0$

$$
\begin{aligned}
f(\theta x + (1-\theta)y) &\leq f(\theta x) + f((1-\theta)y) \\
&= \theta f(x) + (1-\theta) f(y)
\end{aligned}
$$

- 零范数：计算非零元素的数目
  - 令n**=1**
  - 不是范数

例

例

- 极大值函数

例

- 极大值函数

$$f(x) = \max\{x_1, \ldots, x_n\}$$

例

- 极大值函数

$$f(x) = \max\{x_1, \ldots, x_n\}$$

$$f(\theta x + (1-\theta)y) \quad = \quad \max_i(\theta x_i + (1-\theta)y_i)$$

例

- 极大值函数

$$f(x) = \max\{x_1, \ldots, x_n\}$$

$$\begin{aligned} f(\theta x + (1-\theta)y) &= \max_i(\theta x_i + (1-\theta)y_i) \\ &\leq \theta \max_i x_i + (1-\theta)\max_i y_i \end{aligned}$$

例

❑ 极大值函数

$$f(x) = \max\{x_1, \ldots, x_n\}$$

$$\begin{aligned}
f(\theta x + (1-\theta)y) &= \max_i(\theta x_i + (1-\theta)y_i) \\
&\leq \theta \max_i x_i + (1-\theta)\max_i y_i \\
&= \theta f(x) + (1-\theta)f(y)
\end{aligned}$$

例

- 极大值函数

$$f(x) = \max\{x_1, \ldots, x_n\}$$

$$\begin{aligned} f(\theta x + (1-\theta)y) &= \max_i(\theta x_i + (1-\theta)y_i) \\ &\leq \theta \max_i x_i + (1-\theta)\max_i y_i \\ &= \theta f(x) + (1-\theta)f(y) \end{aligned}$$

- 现实中不易求解：不可导
  - 解析逼近

例

例

- 指数和的对数

例

- 指数和的对数 $f(x) = \log\left(e^{x_1} + \cdots + e^{x_n}\right)$

例

- 指数和的对数 $f(x) = \log(e^{x_1} + \cdots + e^{x_n})$

  $\max\{x_1, \ldots, x_n\} \leq f(x) \leq \max\{x_1, \ldots, x_n\} + \log n$

例

- 指数和的对数 $f(x) = \log(e^{x_1} + \cdots + e^{x_n})$

  $\max\{x_1, \ldots, x_n\} \leq f(x) \leq \max\{x_1, \ldots, x_n\} + \log n$

例

- 指数和的对数 $f(x) = \log\left(e^{x_1} + \cdots + e^{x_n}\right)$

  $\max\{x_1, \ldots, x_n\} \le f(x) \le \max\{x_1, \ldots, x_n\} + \log n$

  - $\dfrac{\partial f}{\partial x_i} = \dfrac{e^{x_i}}{e^{x_1} + \ldots + e^{x_n}}$

例

- 指数和的对数 $f(x) = \log\left(e^{x_1} + \cdots + e^{x_n}\right)$
  $\max\{x_1, \ldots, x_n\} \leq f(x) \leq \max\{x_1, \ldots, x_n\} + \log n$
  - $\dfrac{\partial f}{\partial x_i} = \dfrac{e^{x_i}}{e^{x_1} + \ldots + e^{x_n}}$
  - **Hessian**矩阵$[H_{ij}]$

例

- 指数和的对数 $f(x) = \log(e^{x_1} + \cdots + e^{x_n})$

$$\max\{x_1,\ldots,x_n\} \le f(x) \le \max\{x_1,\ldots,x_n\} + \log n$$

  - $\dfrac{\partial f}{\partial x_i} = \dfrac{e^{x_i}}{e^{x_1} + \ldots + e^{x_n}}$
  - **Hessian**矩阵$[H_{ij}]$
  - $i \neq j$: $\dfrac{\partial^2 f}{\partial x_i \partial x_j} = \dfrac{-e^{x_i} e^{x_j}}{(e^{x_1} + \ldots + e^{x_n})^2}$

例

- 指数和的对数 $f(x) = \log(e^{x_1} + \cdots + e^{x_n})$
  $\max\{x_1, \ldots, x_n\} \le f(x) \le \max\{x_1, \ldots, x_n\} + \log n$
  - $\dfrac{\partial f}{\partial x_i} = \dfrac{e^{x_i}}{e^{x_1} + \ldots + e^{x_n}}$
  - **Hessian**矩阵$[H_{ij}]$
  - $i \neq j$: $\dfrac{\partial^2 f}{\partial x_i \partial x_j} = \dfrac{-e^{x_i}e^{x_j}}{(e^{x_1} + \ldots + e^{x_n})^2}$
  - $i = j$: $\dfrac{\partial^2 f}{\partial x_i \partial x_i} = \dfrac{-e^{x_i}e^{x_i} + e^{x_i}(e^{x_1} + \ldots + e^{x_n})}{(e^{x_1} + \ldots + e^{x_n})^2}$

例

- 指数和的对数 $f(x) = \log(e^{x_1} + \cdots + e^{x_n})$

  $\max\{x_1, \ldots, x_n\} \leq f(x) \leq \max\{x_1, \ldots, x_n\} + \log n$

  - $\dfrac{\partial f}{\partial x_i} = \dfrac{e^{x_i}}{e^{x_1} + \ldots + e^{x_n}}$
  - **Hessian**矩阵$[H_{ij}]$
  - $i \neq j$: $\dfrac{\partial^2 f}{\partial x_i \partial x_j} = \dfrac{-e^{x_i} e^{x_j}}{(e^{x_1} + \ldots + e^{x_n})^2}$
  - $i = j$: $\dfrac{\partial^2 f}{\partial x_i \partial x_i} = \dfrac{-e^{x_i} e^{x_i} + e^{x_i}(e^{x_1} + \ldots + e^{x_n})}{(e^{x_1} + \ldots + e^{x_n})^2}$

  $z = (e^{x_1}, \ldots, e^{x_n})$

例

- 指数和的对数 $f(x) = \log(e^{x_1} + \cdots + e^{x_n})$
  $\max\{x_1, \ldots, x_n\} \le f(x) \le \max\{x_1, \ldots, x_n\} + \log n$
  - $\dfrac{\partial f}{\partial x_i} = \dfrac{e^{x_i}}{e^{x_1} + \ldots + e^{x_n}}$
  - **Hessian**矩阵$[H_{ij}]$
  - $i \neq j$: $\dfrac{\partial^2 f}{\partial x_i \partial x_j} = \dfrac{-e^{x_i} e^{x_j}}{(e^{x_1} + \ldots + e^{x_n})^2}$
  - $i = j$: $\dfrac{\partial^2 f}{\partial x_i \partial x_i} = \dfrac{-e^{x_i} e^{x_i} + e^{x_i}(e^{x_1} + \ldots + e^{x_n})}{(\boxed{e^{x_1} + \ldots + e^{x_n}})^2}$

  $z = (e^{x_1}, \ldots, e^{x_n})$

例

- 指数和的对数 $f(x) = \log(e^{x_1} + \cdots + e^{x_n})$

  $\max\{x_1, \ldots, x_n\} \le f(x) \le \max\{x_1, \ldots, x_n\} + \log n$

  - $$\frac{\partial f}{\partial x_i} = \frac{e^{x_i}}{e^{x_1} + \ldots + e^{x_n}}$$
  - **Hessian**矩阵$[H_{ij}]$
  - $i \neq j$: $$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{-e^{x_i} e^{x_j}}{(e^{x_1} + \ldots + e^{x_n})^2}$$
  - $i = j$: $$\frac{\partial^2 f}{\partial x_i \partial x_i} = \frac{-e^{x_i} e^{x_i} + e^{x_i}(e^{x_1} + \ldots + e^{x_n})}{(\underbrace{e^{x_1} + \ldots + e^{x_n}}_{=\mathbf{1}^T z})^2}$$

  $z = (e^{x_1}, \ldots, e^{x_n})$

例

- 指数和的对数 $f(x) = \log\left(e^{x_1} + \cdots + e^{x_n}\right)$

$$\nabla^2 f(x) = \frac{1}{(\mathbf{1}^T z)^2}\left((\mathbf{1}^T z)\,\mathbf{diag}(z) - zz^T\right)$$

- **Hessian**矩阵$[H_{ij}]$
- $i \neq j$: $\dfrac{\partial^2 f}{\partial x_i \partial x_j} = \dfrac{-e^{x_i} e^{x_j}}{(e^{x_1} + \ldots + e^{x_n})^2}$
- $i = j$: $\dfrac{\partial^2 f}{\partial x_i \partial x_i} = \dfrac{-e^{x_i} e^{x_i} + e^{x_i}(e^{x_1} + \ldots + e^{x_n})}{(\boxed{e^{x_1} + \ldots + e^{x_n}})^2}$ $= \mathbf{1}^T z$

$z = \left(e^{x_1}, \ldots, e^{x_n}\right)$

例

- 指数和的对数

$$\nabla^2 f(x) = \frac{1}{(\mathbf{1}^T z)^2}\left((\mathbf{1}^T z)\,\mathbf{diag}(z) - zz^T\right)$$

例

- 指数和的对数

$$\nabla^2 f(x) = \frac{1}{(\mathbf{1}^T z)^2}\left(\boxed{(\mathbf{1}^T z)\,\mathbf{diag}(z) - zz^T}\right) \quad \mathbf{K}$$

例

- 指数和的对数

$$\nabla^2 f(x) = \frac{1}{(\mathbf{1}^T z)^2}\left(\boxed{(\mathbf{1}^T z)\,\mathbf{diag}(z) - zz^T}\right) \quad \mathbf{K}$$

- 证明该**Hessian**矩阵半正定

例

- 指数和的对数

$$\nabla^2 f(x) = \frac{1}{(\mathbf{1}^T z)^2}\left(\boxed{(\mathbf{1}^T z)\,\mathbf{diag}(z) - zz^T}\right) \quad \mathbf{K}$$

- 证明该**Hessian**矩阵半正定
  - 对任意*v*，有

例

- 指数和的对数

$$\nabla^2 f(x) = \frac{1}{(\mathbf{1}^T z)^2}\left(\boxed{(\mathbf{1}^T z)\,\mathbf{diag}(z) - zz^T}\right) \quad \mathbf{K}$$

- 证明该**Hessian**矩阵半正定
  - 对任意v，有
  - $v^TKv=$

例

- 指数和的对数

$$\nabla^2 f(x) = \frac{1}{(\mathbf{1}^T z)^2}\left(\boxed{(\mathbf{1}^T z)\,\mathbf{diag}(z) - zz^T}\right) \quad \mathbf{K}$$

- 证明该**Hessian**矩阵半正定
  - 对任意v，有
  - $v^TKv = (\mathbf{1}^T z)\, v^T \mathbf{diag}(z)\, v - v^T z z^T v$

例

- 指数和的对数

$$\nabla^2 f(x) = \frac{1}{(\mathbf{1}^T z)^2}\left(\boxed{(\mathbf{1}^T z)\,\mathbf{diag}(z) - zz^T}\right) \quad \mathbf{K}$$

- 证明该**Hessian**矩阵半正定
  - 对任意v，有
  - $v^TKv = (\mathbf{1}^T z)\, v^T\mathbf{diag}(z)\, v - v^T z z^T v$

$$= \left(\sum_{i=1}^n z_i\right)\left(\sum_{i=1}^n v_i^2 z_i\right) - \left(\sum_{i=1}^n v_i z_i\right)^2$$

例

- 指数和的对数

$$\nabla^2 f(x) = \frac{1}{(\mathbf{1}^T z)^2}\left(\boxed{(\mathbf{1}^T z)\,\mathbf{diag}(z) - zz^T}\right) \quad \mathbf{K}$$

- 证明该**Hessian**矩阵半正定
  - 对任意*v*，有
  - v^TKv= $(\mathbf{1}^T z) v^T \mathbf{diag}(z) v - v^T z z^T v$

$$= \left(\sum_{i=1}^n z_i\right)\left(\sum_{i=1}^n v_i^2 z_i\right) - \left(\sum_{i=1}^n v_i z_i\right)^2$$

$$a_i = v_i\sqrt{z_i}$$

例

- 指数和的对数

$$\nabla^2 f(x) = \frac{1}{(\mathbf{1}^T z)^2}\left(\boxed{(\mathbf{1}^T z)\,\mathbf{diag}(z) - zz^T}\right) \quad \mathbf{K}$$

- 证明该**Hessian**矩阵半正定
  - 对任意*v*，有
  - $v^TKv = (\mathbf{1}^T z)\, v^T \mathbf{diag}(z)\, v - v^T z z^T v$

$$= \left(\sum_{i=1}^{n} z_i\right)\left(\sum_{i=1}^{n} v_i^2 z_i\right) - \left(\sum_{i=1}^{n} v_i z_i\right)^2$$

$a_i = v_i\sqrt{z_i} \quad b_i = \sqrt{z_i}$.

例

- 指数和的对数

$$\nabla^2 f(x) = \frac{1}{(\mathbf{1}^T z)^2}\left(\boxed{(\mathbf{1}^T z)\,\mathbf{diag}(z) - zz^T}\right) \quad \mathbf{K}$$

- 证明该**Hessian**矩阵半正定
  - 对任意v，有
  - v^TKv= $(\mathbf{1}^T z) v^T \mathbf{diag}(z) v - v^T z z^T v$

$$= \left(\sum_{i=1}^n z_i\right)\left(\sum_{i=1}^n v_i^2 z_i\right) - \left(\sum_{i=1}^n v_i z_i\right)^2$$

$$a_i = v_i\sqrt{z_i} \quad b_i = \sqrt{z_i} \qquad (a^T a)$$

例

- 指数和的对数

$$\nabla^2 f(x) = \frac{1}{(\mathbf{1}^T z)^2}\left(\boxed{(\mathbf{1}^T z)\,\mathbf{diag}(z) - zz^T}\right) \quad \mathbf{K}$$

- 证明该**Hessian**矩阵半正定
  - 对任意*v*，有
  - $v^TKv = (\mathbf{1}^T z)\, v^T \mathbf{diag}(z)\, v - v^T z z^T v$

$$= \left(\sum_{i=1}^n z_i\right)\left(\sum_{i=1}^n v_i^2 z_i\right) - \left(\sum_{i=1}^n v_i z_i\right)^2$$

$a_i = v_i\sqrt{z_i} \quad b_i = \sqrt{z_i} \quad (a^T a)(b^T b)$

例

- 指数和的对数

$$\nabla^2 f(x) = \frac{1}{(\mathbf{1}^T z)^2}\left(\boxed{(\mathbf{1}^T z)\,\mathbf{diag}(z) - zz^T}\right)$$

K

- 证明该**Hessian**矩阵半正定
  - 对任意v，有
  - $v^TKv = (\mathbf{1}^T z)\, v^T \mathbf{diag}(z)\, v - v^T z z^T v$

$$= \left(\sum_{i=1}^n z_i\right)\left(\sum_{i=1}^n v_i^2 z_i\right) - \left(\sum_{i=1}^n v_i z_i\right)^2$$

$a_i = v_i\sqrt{z_i}$ $b_i = \sqrt{z_i}$ $(a^T a)(b^T b) -$

例

- 指数和的对数

$$\nabla^2 f(x) = \frac{1}{(\mathbf{1}^T z)^2}\left(\boxed{(\mathbf{1}^T z)\,\mathbf{diag}(z) - zz^T}\right) \quad \mathbf{K}$$

- 证明该**Hessian**矩阵半正定
  - 对任意*v*，有
  - $v^TKv = (\mathbf{1}^T z)\, v^T \mathbf{diag}(z)\, v - v^T z z^T v$

$$= \left(\sum_{i=1}^n z_i\right)\left(\sum_{i=1}^n v_i^2 z_i\right) - \left(\sum_{i=1}^n v_i z_i\right)^2$$

$$a_i = v_i\sqrt{z_i} \quad b_i = \sqrt{z_i} \qquad (a^Ta)(b^Tb) - (a^Tb)^2$$

例

- 指数和的对数

$$\nabla^2 f(x) = \frac{1}{(\mathbf{1}^T z)^2}\left(\boxed{(\mathbf{1}^T z)\,\mathbf{diag}(z) - zz^T}\right) \quad \mathbf{K}$$

- 证明该**Hessian**矩阵半正定
  - 对任意*v*，有
  - $v^T K v = (\mathbf{1}^T z)\, v^T \mathbf{diag}(z)\, v - v^T z z^T v$

$$= \left(\sum_{i=1}^{n} z_i\right)\left(\sum_{i=1}^{n} v_i^2 z_i\right) - \left(\sum_{i=1}^{n} v_i z_i\right)^2$$

$$a_i = v_i\sqrt{z_i} \quad b_i = \sqrt{z_i} \qquad (a^T a)(b^T b) - (a^T b)^2 \geq 0$$

例

例

- 几何平均：凹函数

例

- 几何平均：凹函数

$$f(x) = \left(\prod_{i=1}^{n} x_i\right)^{1/n}$$

例

- 几何平均：凹函数

$$f(x) = \left(\prod_{i=1}^{n} x_i\right)^{1/n}$$

$$\frac{\partial^2 f(x)}{\partial x_k^2} = -(n-1)\frac{\left(\prod_{i=1}^{n} x_i\right)^{1/n}}{n^2 x_k^2}$$

例

- 几何平均：凹函数

$$f(x) = \left(\prod_{i=1}^{n} x_i\right)^{1/n}$$

$$\frac{\partial^2 f(x)}{\partial x_k^2} = -(n-1)\frac{\left(\prod_{i=1}^{n} x_i\right)^{1/n}}{n^2 x_k^2}$$

$$\frac{\partial^2 f(x)}{\partial x_k \partial x_l} = \frac{\left(\prod_{i=1}^{n} x_i\right)^{1/n}}{n^2 x_k x_l} \quad \text{for } k \neq l$$

例

- 几何平均：凹函数

$$f(x) = \left(\prod_{i=1}^n x_i\right)^{1/n}$$

$$\frac{\partial^2 f(x)}{\partial x_k^2} = -(n-1)\frac{\left(\prod_{i=1}^n x_i\right)^{1/n}}{n^2 x_k^2}$$

$$\frac{\partial^2 f(x)}{\partial x_k \partial x_l} = \frac{\left(\prod_{i=1}^n x_i\right)^{1/n}}{n^2 x_k x_l} \quad \text{for } k \neq l$$

$$\nabla^2 f(x) = -\frac{\prod_{i=1}^n x_i^{1/n}}{n^2}\left(n\,\mathbf{diag}(1/x_1^2, \ldots, 1/x_n^2) - qq^T\right) \; q_i = 1/x_i$$

例

- 几何平均：凹函数

$$f(x) = \left(\prod_{i=1}^n x_i\right)^{1/n}$$

$$\frac{\partial^2 f(x)}{\partial x_k^2} = -(n-1)\frac{\left(\prod_{i=1}^n x_i\right)^{1/n}}{n^2 x_k^2}$$

$$\frac{\partial^2 f(x)}{\partial x_k \partial x_l} = \frac{\left(\prod_{i=1}^n x_i\right)^{1/n}}{n^2 x_k x_l} \quad \text{for } k \neq l$$

$$\nabla^2 f(x) = -\frac{\prod_{i=1}^n x_i^{1/n}}{n^2}\left(n\,\mathbf{diag}(1/x_1^2, \ldots, 1/x_n^2) - qq^T\right) \; q_i = 1/x_i$$

$$v^T \nabla^2 f(x) v = -\frac{\prod_{i=1}^n x_i^{1/n}}{n^2}\left(n\sum_{i=1}^n v_i^2/x_i^2 - \left(\sum_{i=1}^n v_i/x_i\right)^2\right) \leq 0$$

例

例

$$f(X) = \log\det X,\ \mathbf{dom}\, f = \mathbf{S}^n_{++}$$

例

$$f(X) = \log\det X, \ \mathbf{dom}\, f = \mathbf{S}^n_{++}$$

$$\begin{aligned} g(t) = \log\det(X + tV) &= \log\det X + \log\det(I + tX^{-1/2}VX^{-1/2}) \\ &= \log\det X + \sum_{i=1}^{n} \log(1 + t\lambda_i) \end{aligned}$$

例

$$f(X) = \log\det X,\ \mathbf{dom}\, f = \mathbf{S}^n_{++}$$

$$\begin{aligned} g(t) = \log\det(X+tV) &= \log\det X + \log\det(I + tX^{-1/2}VX^{-1/2}) \\ &= \log\det X + \sum_{i=1}^{n} \log(1+t\lambda_i) \end{aligned}$$

- 此处，λ_i是$X^{-1/2}VX^{-1/2}$的特征值

例

$$f(X) = \log\det X,\ \mathbf{dom}\, f = \mathbf{S}^n_{++}$$

$$\begin{aligned} g(t) = \log\det(X + tV) &= \log\det X + \log\det(I + tX^{-1/2}VX^{-1/2}) \\ &= \log\det X + \sum_{i=1}^{n} \log(1 + t\lambda_i) \end{aligned}$$

- 此处，λ_i 是 $X^{-1/2}VX^{-1/2}$ 的特征值
- $g'(t) = \sum_{i=1}^{n} \dfrac{\lambda_i}{1 + t\lambda_i}$

例

$$f(X) = \log\det X, \ \mathbf{dom}\, f = \mathbf{S}_{++}^n$$

$$\begin{aligned} g(t) = \log\det(X + tV) &= \log\det X + \log\det(I + tX^{-1/2}VX^{-1/2}) \\ &= \log\det X + \sum_{i=1}^{n} \log(1 + t\lambda_i) \end{aligned}$$

- 此处，λ_i是$X^{-1/2}VX^{-1/2}$的特征值
- $g'(t) = \sum_{i=1}^{n} \frac{\lambda_i}{1 + t\lambda_i}$
- $g''(t) = -\sum_{i=1}^{n} \frac{\lambda_i^2}{(1 + t\lambda_i)^2}$

例

$$f(X) = \log\det X,\ \mathbf{dom}\, f = \mathbf{S}^n_{++}$$

$$\begin{aligned} g(t) = \log\det(X + tV) &= \log\det X + \log\det(I + tX^{-1/2}VX^{-1/2}) \\ &= \log\det X + \sum_{i=1}^{n} \log(1 + t\lambda_i) \end{aligned}$$

- 此处，λ_i 是 $X^{-1/2}VX^{-1/2}$ 的特征值
- $g'(t) = \sum_{i=1}^{n} \frac{\lambda_i}{1 + t\lambda_i}$
- $g''(t) = -\sum_{i=1}^{n} \frac{\lambda_i^2}{(1 + t\lambda_i)^2}$
- $g''(t) \le 0$

保凸运算

保凸运算

- 如何判断某一函数是否为凸函数?

保凸运算

- 如何判断某一函数是否为凸函数?
  - 方式**1:**验证定义（通过将函数限制在直线判断）

保凸运算

- 如何判断某一函数是否为凸函数?
  - 方式**1:**验证定义（通过将函数限制在直线判断）
  - 方式**2:**对于二阶可微函数，验证 $\nabla^2 f(x) \succeq 0$

保凸运算

- 如何判断某一函数是否为凸函数?
  - 方式**1:**验证定义（通过将函数限制在直线判断）
  - 方式**2:**对于二阶可微函数，验证 $\nabla^2 f(x) \succeq 0$
  - 方式**3:**考察函数是否可由多个简单凸函数通过保凸运算得到

保凸运算

- 如何判断某一函数是否为凸函数?
  - 方式**1:**验证定义（通过将函数限制在直线判断）
  - 方式**2:**对于二阶可微函数，验证 $\nabla^2 f(x) \succeq 0$
  - 方式**3:**考察函数是否可由多个简单凸函数通过保凸运算得到
    - 非负加权求和

保凸运算

- 如何判断某一函数是否为凸函数?
  - 方式**1:**验证定义（通过将函数限制在直线判断）
  - 方式**2:**对于二阶可微函数，验证 $\nabla^2 f(x) \succeq 0$
  - 方式**3:**考察函数是否可由多个简单凸函数通过保凸运算得到
    - 非负加权求和
    - 复合仿射映射

保凸运算

- 如何判断某一函数是否为凸函数?
  - 方式**1:**验证定义（通过将函数限制在直线判断）
  - 方式**2:**对于二阶可微函数，验证 $\nabla^2 f(x) \succeq 0$
  - 方式**3:**考察函数是否可由多个简单凸函数通过保凸运算得到
    - 非负加权求和
    - 复合仿射映射
    - 逐点最大和逐点上确界

保凸运算

- 如何判断某一函数是否为凸函数?
  - 方式**1**:验证定义（通过将函数限制在直线判断）
  - 方式**2**:对于二阶可微函数，验证 $\nabla^2 f(x) \succeq 0$
  - 方式**3**:考察函数是否可由多个简单凸函数通过保凸运算得到
    - 非负加权求和
    - 复合仿射映射
    - 逐点最大和逐点上确界
    - 复合

保凸运算

- 如何判断某一函数是否为凸函数?
  - 方式**1**:验证定义（通过将函数限制在直线判断）
  - 方式**2**:对于二阶可微函数，验证 $\nabla^2 f(x) \succeq 0$
  - 方式**3**:考察函数是否可由多个简单凸函数通过保凸运算得到
    - 非负加权求和
    - 复合仿射映射
    - 逐点最大和逐点上确界
    - 复合
    - 最小化

保凸运算

- 如何判断某一函数是否为凸函数?
  - 方式**1**:验证定义（通过将函数限制在直线判断）
  - 方式**2**:对于二阶可微函数，验证 $\nabla^2 f(x) \succeq 0$
  - 方式**3**:考察函数是否可由多个简单凸函数通过保凸运算得到
    - 非负加权求和
    - 复合仿射映射
    - 逐点最大和逐点上确界
    - 复合
    - 最小化
    - 透视函数

非负加权求和

非负加权求和

- 非负乘积

非负加权求和

- 非负乘积
  - αf 为凸函数，当函数f为凸函数，$\alpha \geq 0$

非负加权求和

- 非负乘积
  - αf 为凸函数，当函数f为凸函数，$\alpha \geq 0$
- 求和

非负加权求和

- 非负乘积
  - αf 为凸函数，当函数f为凸函数，$\alpha \geq 0$
- 求和
  - $f_1 + f_2$为凸函数，当 f_1, f_2为凸函数

非负加权求和

- 非负乘积
  - αf 为凸函数，当函数f为凸函数，$\alpha \geq 0$
- 求和
  - $f_1 + f_2$为凸函数，当 f_1, f_2为凸函数
- 非负加权求和

非负加权求和

- 非负乘积
  - αf 为凸函数，当函数f为凸函数，$\alpha \geq 0$
- 求和
  - $f_1 + f_2$为凸函数，当 f_1, f_2为凸函数
- 非负加权求和 $f = w_1 f_1 + \cdots + w_m f_m$

非负加权求和

- 非负乘积
  - αf 为凸函数，当函数f为凸函数，$\alpha \geq 0$
- 求和
  - $f_1 + f_2$为凸函数，当 f_1, f_2为凸函数
- 非负加权求和 $f = w_1 f_1 + \cdots + w_m f_m$
  - 是保凸运算

非负加权求和

- 非负乘积
  - αf 为凸函数，当函数f为凸函数，$\alpha \geq 0$
- 求和
  - $f_1 + f_2$为凸函数，当 f_1, f_2为凸函数
- 非负加权求和 $f = w_1 f_1 + \cdots + w_m f_m$
  - 是保凸运算
- 可扩展到无限项、积分

非负加权求和

- 非负乘积
  - αf 为凸函数，当函数f为凸函数，$\alpha \geq 0$
- 求和
  - $f_1 + f_2$为凸函数，当 f_1, f_2为凸函数
- 非负加权求和 $f = w_1 f_1 + \cdots + w_m f_m$
  - 是保凸运算
- 可扩展到无限项、积分

$$g(x) = \int_{\mathcal{A}} w(y) f(x, y)\, dy \quad w(y) \geq 0$$

非负加权求和

- 非负乘积
  - αf 为凸函数，当函数f为凸函数，$\alpha \geq 0$
- 求和
  - $f_1 + f_2$为凸函数，当 f_1, f_2为凸函数
- 非负加权求和 $f = w_1 f_1 + \cdots + w_m f_m$
  - 是保凸运算
- 可扩展到无限项、积分

$$g(x) = \int_A w(y) f(x, y)\, dy \quad w(y) \geq 0$$

  - 对集合A中每一个y，$f(x, y)$是凸函数

复合仿射映射

复合仿射映射

- 复合仿射映射：$f(Ax+b)$为凸函数，当函数f为凸函数

复合仿射映射

- 复合仿射映射: $f(Ax+b)$为凸函数，当函数f为凸函数
- 证明：令$g(x)=f(Ax+b)$

$$\mathbf{dom}\, g=\{x \mid Ax+b\in \mathbf{dom}\, f\}$$

复合仿射映射

- 复合仿射映射：$f(Ax+b)$ 为凸函数，当函数f 为凸函数
- 证明：令 $g(x) = f(Ax+b)$

$$\mathbf{dom}\, g = \{x \mid Ax + b \in \mathbf{dom}\, f\}$$

  - 任意取$x, y \in \mathbf{dom}\, g$，$0 \le \theta \le 1$

复合仿射映射

- 复合仿射映射：$f(Ax+b)$ 为凸函数，当函数 f 为凸函数
- 证明：令 $g(x) = f(Ax+b)$

$$\mathbf{dom}\, g = \{x \mid Ax + b \in \mathbf{dom}\, f\}$$

  - 任意取 $x, y \in \mathbf{dom}\, g$，$0 \le \theta \le 1$
  - $g(\theta x + (1-\theta)y)$

复合仿射映射

- 复合仿射映射：$f(Ax+b)$ 为凸函数，当函数 f 为凸函数
- 证明：令 $g(x) = f(Ax+b)$

$$\mathbf{dom}\, g = \{x \mid Ax + b \in \mathbf{dom}\, f\}$$

  - 任意取 $x, y \in \mathbf{dom}\, g$，$0 \le \theta \le 1$
  - $g(\theta x + (1-\theta)y)$

$$= f(\theta Ax + (1-\theta)Ay + b)$$

复合仿射映射

- 复合仿射映射: $f(Ax+b)$ 为凸函数，当函数f 为凸函数
- 证明：令 $g(x)=f(Ax+b)$

$$\mathbf{dom}\, g=\{x \mid Ax+b\in \mathbf{dom}\, f\}$$

  - 任意取$x, y \in \mathbf{dom}\, g$, $0\le\theta\le 1$
  - $g(\theta x+(1-\theta)y)$

$$=f(\theta Ax+(1-\theta)Ay+b)$$

$$=f(\theta(Ax+b)+(1-\theta)(Ay+b))$$

复合仿射映射

- 复合仿射映射: $f(Ax+b)$ 为凸函数，当函数f 为凸函数
- 证明： 令 $g(x)=f(Ax+b)$

$$\mathbf{dom}\, g=\{x \mid Ax+b\in \mathbf{dom}\, f\}$$

  - 任意取$x, y \in \mathbf{dom}\, g$, $0 \le \theta \le 1$
  - $g(\theta x+(1-\theta)y)$

$$\begin{aligned}&= f(\theta Ax+(1-\theta)Ay+b)\\&= f(\theta(Ax+b)+(1-\theta)(Ay+b))\\&\le \theta f(Ax+b)+(1-\theta)f(Ay+b)\end{aligned}$$

复合仿射映射

- 复合仿射映射：$f(Ax+b)$为凸函数，当函数f为凸函数
- 证明：令$g(x)=f(Ax+b)$

$$\mathbf{dom}\, g = \{x \mid Ax+b \in \mathbf{dom}\, f\}$$

  - 任意取$x, y \in \mathbf{dom}\, g$，$0 \le \theta \le 1$
  - $g(\theta x+(1-\theta)y)$

$$\begin{aligned}
&= f(\theta Ax+(1-\theta)Ay+b) \\
&= f(\theta(Ax+b)+(1-\theta)(Ay+b)) \\
&\le \theta f(Ax+b)+(1-\theta)f(Ay+b) \\
&= \theta g(x)+(1-\theta)g(y)
\end{aligned}$$

复合仿射映射

复合仿射映射

- 存在一系列凸函数 $f_i: \mathbf{R}^n \to \mathbf{R}$
  - 则$g(x) = A[f_1(x), \ldots, f_m(x)] + b$ 是否为凸函数?

复合仿射映射

- 存在一系列凸函数$f_i: \mathbf{R}^n \rightarrow \mathbf{R}$
  - 则$g(x) = A[f_1(x), \ldots, f_m(x)] + b$ 是否为凸函数?

$$f = w_1 f_1 + \cdots + w_m f_m$$

复合仿射映射

- 存在一系列凸函数 $f_i: \mathbf{R}^n \to \mathbf{R}$
  - 则$g(x) = A[f_1(x), \ldots, f_m(x)] + b$ 是否为凸函数?

$$f = w_1 f_1 + \cdots + w_m f_m$$

- 例： 线性不等式的对数惩罚函数

复合仿射映射

- 存在一系列凸函数f_i: $\mathbf{R}^n \to \mathbf{R}$
  - 则$g(x) = A[f_1(x), \ldots, f_m(x)] + b$ 是否为凸函数?

$$f = w_1 f_1 + \cdots + w_m f_m$$

- 例：线性不等式的对数惩罚函数

$$f(x) = -\sum_{i=1}^{m} \log(b_i - a_i^T x)$$

复合仿射映射

- 存在一系列凸函数f_i: $\mathbf{R}^n \to \mathbf{R}$
  - 则$g(x) = A[f_1(x), \ldots, f_m(x)] + b$ 是否为凸函数?

$$f = w_1 f_1 + \cdots + w_m f_m$$

- 例：线性不等式的对数惩罚函数

$$f(x) = -\sum_{i=1}^{m} \log(b_i - a_i^T x)$$

$$\mathbf{dom}\, f = \{x \mid a_i^T x < b_i, i = 1, \ldots, m\}$$

复合仿射映射

- 存在一系列凸函数 f_i: $\mathbf{R}^n \to \mathbf{R}$
  - 则 $g(x) = A[f_1(x), \ldots, f_m(x)] + b$ 是否为凸函数?

$$f = w_1 f_1 + \cdots + w_m f_m$$

- 例：线性不等式的对数惩罚函数

$$f(x) = -\sum_{i=1}^{m} \log(b_i - a_i^T x)$$

$$\mathbf{dom}\, f = \{x \mid a_i^T x < b_i, i = 1, \ldots, m\}$$

- 仿射函数的任意范数

复合仿射映射

- 存在一系列凸函数 $f_i: \mathbf{R}^n \to \mathbf{R}$
  - 则 $g(x) = A[f_1(x), \ldots, f_m(x)] + b$ 是否为凸函数?

$$f = w_1 f_1 + \cdots + w_m f_m$$

- 例：线性不等式的对数惩罚函数

$$f(x) = -\sum_{i=1}^{m} \log(b_i - a_i^T x)$$

$$\mathbf{dom}\, f = \{x \mid a_i^T x < b_i, i = 1, \ldots, m\}$$

- 仿射函数的任意范数

$$f(x) = \|Ax + b\|$$

逐点最大

逐点最大

- 若 $f_1, \ldots, f_m$ 为凸函数,

逐点最大

- 若 $f_1, \ldots, f_m$ 为凸函数,

$$f(x) = \max\{f_1(x), \ldots, f_m(x)\}$$ 为凸函数

逐点最大

- 若 $f_1, \ldots, f_m$ 为凸函数,

$$f(x) = \max\{f_1(x), \ldots, f_m(x)\} \text{ 为凸函数}$$

- 取 ***m*****=2:**

逐点最大

- 若 $f_1, \ldots, f_m$ 为凸函数,

  $$f(x) = \max\{f_1(x), \ldots, f_m(x)\} \text{ 为凸函数}$$

- 取m=**2:**
  - $f(\theta x + (1-\theta)y)$

逐点最大

- 若 $f_1, \ldots, f_m$ 为凸函数,

$$f(x) = \max\{f_1(x), \ldots, f_m(x)\}$$ 为凸函数

- 取m=**2:**
  - $f(\theta x + (1-\theta)y)$

$$= \max\{f_1(\theta x + (1-\theta)y), f_2(\theta x + (1-\theta)y)\}$$

逐点最大

- 若 $f_1, \ldots, f_m$ 为凸函数,

  $$f(x) = \max\{f_1(x), \ldots, f_m(x)\}$$ 为凸函数

- 取 m=**2:**
  - $f(\theta x + (1-\theta)y)$

    $= \max\{f_1(\theta x + (1-\theta)y), f_2(\theta x + (1-\theta)y)\}$

    $\leq \max\{\theta f_1(x) + (1-\theta)f_1(y), \theta f_2(x) + (1-\theta)f_2(y)\}$

逐点最大

- 若 $f_1, \ldots, f_m$ 为凸函数,

  $$f(x) = \max\{f_1(x), \ldots, f_m(x)\} \text{ 为凸函数}$$

- 取m=**2:**
  - $f(\theta x + (1-\theta)y)$

  $$= \max\{f_1(\theta x + (1-\theta)y), f_2(\theta x + (1-\theta)y)\}$$

  $$\leq \max\{\theta f_1(x) + (1-\theta)f_1(y), \theta f_2(x) + (1-\theta)f_2(y)\}$$

  $$\leq \max\{\theta f_1(x), \theta f_2(x)\} + \max\{(1-\theta)f_1(y), (1-\theta)f_2(y)\}$$

逐点最大

- 若 $f_1, \ldots, f_m$ 为凸函数,

$$f(x) = \max\{f_1(x), \ldots, f_m(x)\}$$ 为凸函数

- 取*m*=**2:**
  - $f(\theta x + (1-\theta)y)$

$$
\begin{aligned}
&= \max\{f_1(\theta x + (1-\theta)y), f_2(\theta x + (1-\theta)y)\} \\
&\leq \max\{\theta f_1(x) + (1-\theta)f_1(y), \theta f_2(x) + (1-\theta)f_2(y)\} \\
&\leq \max\{\theta f_1(x), \theta f_2(x)\} + \max\{(1-\theta)f_1(y), (1-\theta)f_2(y)\} \\
&= \theta f(x) + (1-\theta)f(y)
\end{aligned}
$$

逐点最大

- 若 $f_1, \ldots, f_m$ 为凸函数,

  $$f(x) = \max\{f_1(x), \ldots, f_m(x)\}$$ 为凸函数

- 取m=2:
  - $$\begin{aligned} & f(\theta x + (1-\theta)y) \\ &= \max\{f_1(\theta x + (1-\theta)y), f_2(\theta x + (1-\theta)y)\} \\ &\le \max\{\theta f_1(x) + (1-\theta)f_1(y), \theta f_2(x) + (1-\theta)f_2(y)\} \\ &\le \max\{\theta f_1(x), \theta f_2(x)\} + \max\{(1-\theta)f_1(y), (1-\theta)f_2(y)\} \\ &= \theta f(x) + (1-\theta)f(y) \end{aligned}$$

- 分段线性函数为凸函数

逐点最大

- 若 $f_1, \ldots, f_m$ 为凸函数,

$$f(x) = \max\{f_1(x), \ldots, f_m(x)\} \text{ 为凸函数}$$

- 取m=2:
  - $f(\theta x + (1-\theta)y)$

$$\begin{aligned}
&= \max\{f_1(\theta x + (1-\theta)y), f_2(\theta x + (1-\theta)y)\} \\
&\le \max\{\theta f_1(x) + (1-\theta)f_1(y), \theta f_2(x) + (1-\theta)f_2(y)\} \\
&\le \max\{\theta f_1(x), \theta f_2(x)\} + \max\{(1-\theta)f_1(y), (1-\theta)f_2(y)\} \\
&= \theta f(x) + (1-\theta)f(y)
\end{aligned}$$

- 分段线性函数为凸函数

$$f(x) = \max_{i=1,\ldots,m}(a_i^T x + b_i)$$

逐点最大

逐点最大

- 最大r个分量之和($x_{[i]}$为x中第i大的分量)

逐点最大

- 最大r个分量之和($x_{[i]}$为x中第i大的分量)

$$x_{[1]} \geq x_{[2]} \geq \cdots \geq x_{[n]}$$

逐点最大

- 最大r个分量之和($x_{[i]}$为x中第i大的分量)

$$x_{[1]} \geq x_{[2]} \geq \cdots \geq x_{[n]}$$

$$f(x) = x_{[1]} + x_{[2]} + \cdots + x_{[r]}$$

逐点最大

- 最大r个分量之和($x_{[i]}$为x中第i大的分量)

$$x_{[1]} \geq x_{[2]} \geq \cdots \geq x_{[n]}$$

$$f(x) = x_{[1]} + x_{[2]} + \cdots + x_{[r]}$$

$$f(x) = \sum_{i=1}^{r} x_{[i]}$$

逐点最大

- 最大r个分量之和($x_{[i]}$为x中第i大的分量)

$$x_{[1]} \geq x_{[2]} \geq \cdots \geq x_{[n]}$$

$$f(x) = x_{[1]} + x_{[2]} + \cdots + x_{[r]}$$

$$f(x) = \sum_{i=1}^{r} x_{[i]}$$

$$= \max\{x_{i_1} + \cdots + x_{i_r} \mid 1 \leq i_1 < i_2 < \cdots < i_r \leq n\}$$

逐点上确界

逐点上确界

- 若对任意$y \in \mathcal{A}$,$f(x, y)$ 关于x为凸函数，则

逐点上确界

- 若对任意$y \in \mathcal{A}$，$f(x,y)$ 关于x为凸函数，则
- $$g(x) = \sup_{y \in \mathcal{A}} f(x,y)$$ 为凸函数

逐点上确界

- 若对任意$y \in \mathcal{A}$，$f(x,y)$关于x为凸函数，则
- $g(x) = \sup_{y \in \mathcal{A}} f(x,y)$为凸函数
- 例：

逐点上确界

- 若对任意$y \in \mathcal{A}$，$f(x, y)$ 关于x为凸函数，则
- $$g(x) = \sup_{y \in \mathcal{A}} f(x, y)$$ 为凸函数
- 例：
  - 在集合C中距x最远点的距离函数：

$$f(x) = \sup_{y \in C} \|x - y\|$$

逐点上确界

- 若对任意$y \in \mathcal{A}$,$f(x,y)$ 关于x为凸函数，则
- $$g(x) = \sup_{y \in \mathcal{A}} f(x,y)$$ 为凸函数
- 例：
  - 在集合C中距x最远点的距离函数：
  $$f(x) = \sup_{y \in C} \|x - y\|$$
  - 对称矩阵的最大特征值：

逐点上确界

- 若对任意$y \in \mathcal{A}$,$f(x,y)$ 关于x为凸函数，则
- $$g(x) = \sup_{y \in \mathcal{A}} f(x,y)$$ 为凸函数
- 例：
  - 在集合C中距x最远点的距离函数：

$$f(x) = \sup_{y \in C} \|x - y\|$$

  - 对称矩阵的最大特征值：

$$\lambda_{\max}(X) = \sup_{\|y\|_2 = 1} y^T X y$$

函数复合

函数复合

- 函数 $g : \mathbf{R}^n \to \mathbf{R}^k$ 和 $h : \mathbf{R}^k \to \mathbf{R}$ 的复合：

$$f(x) = h(g(x))$$

函数复合

- 函数 $g: \mathbf{R}^n \to \mathbf{R}^k$ 和 $h: \mathbf{R}^k \to \mathbf{R}$ 的复合：
$$f(x) = h(g(x))$$
- 假设n, k=**1**，定义域均为实数集，g和h二阶可微

函数复合

- 函数 $g : \mathbf{R}^n \to \mathbf{R}^k$ 和 $h : \mathbf{R}^k \to \mathbf{R}$ 的复合：

$$f(x) = h(g(x))$$

- 假设n, k=**1**，定义域均为实数集，g和h二阶可微
- $\nabla^2 f(x) \succeq 0$ ⇔f为凸函数

函数复合

- 函数 $g:\mathbf{R}^n \to \mathbf{R}^k$ 和 $h:\mathbf{R}^k \to \mathbf{R}$ 的复合：

$$f(x) = h(g(x))$$

- 假设n, k=1，定义域均为实数集，g和h二阶可微
- $\nabla^2 f(x) \succeq 0$ ⇔f为凸函数

$$f''(x) = h'(g(x))g'(x)$$

函数复合

- 函数 $g: \mathbf{R}^n \to \mathbf{R}^k$ 和 $h: \mathbf{R}^k \to \mathbf{R}$ 的复合：

$$f(x) = h(g(x))$$

- 假设n, k=**1**，定义域均为实数集，g和h二阶可微
- $\nabla^2 f(x) \succeq 0 \Leftrightarrow f$为凸函数

$$f'(x) = h'(g(x))g'(x)$$

$$f''(x) = h''(g(x))g'(x)^2 + h'(g(x))g''(x)$$

函数复合

- 函数 $g : \mathbf{R}^n \rightarrow \mathbf{R}^k$ 和 $h : \mathbf{R}^k \rightarrow \mathbf{R}$ 的复合：

$$f(x) = h(g(x))$$

- 假设n, k=1，定义域均为实数集，g和h二阶可微
- $\nabla^2 f(x) \succeq 0$ ⇔f为凸函数

$$f'(x) = h'(g(x))g'(x)$$

$$f''(x) = \boxed{h''(g(x))}g'(x)^2 + h'(g(x))g''(x)$$

函数复合

- 函数 $g : \mathbf{R}^n \to \mathbf{R}^k$ 和 $h : \mathbf{R}^k \to \mathbf{R}$ 的复合：

$$f(x) = h(g(x))$$

- 假设n, k=**1**，定义域均为实数集，g和h二阶可微
- $\nabla^2 f(x) \succeq 0$ ⇔f为凸函数

$$f'(x) = h'(g(x))g'(x)$$

$$f''(x) = \boxed{h''(g(x))}g'(x)^2 + \boxed{h'(g(x))}\boxed{g''(x)}$$

函数复合

- 函数 $g: \mathbf{R}^n \to \mathbf{R}^k$ 和 $h: \mathbf{R}^k \to \mathbf{R}$ 的复合：

$$f(x) = h(g(x))$$

- 假设*n*, *k*=**1**，定义域均为实数集，*g*和*h*二阶可微
- $\nabla^2 f(x) \succeq 0$ ⇔*f*为凸函数

$$f'(x) = h'(g(x))g'(x)$$

$$f''(x) = \boxed{h''(g(x))}g'(x)^2 + \boxed{h'(g(x))}\boxed{g''(x)}$$

  - *h*为凸，不减，*g*为凸，则*f*为凸

函数复合

- 函数 $g : \mathbf{R}^n \to \mathbf{R}^k$ 和 $h : \mathbf{R}^k \to \mathbf{R}$ 的复合：

$$f(x) = h(g(x))$$

- 假设*n*, *k*=**1**，定义域均为实数集，*g*和*h*二阶可微
- $\nabla^2 f(x) \succeq 0$ ⇔*f*为凸函数

$$f'(x) = h'(g(x))g'(x)$$

$$f''(x) = \boxed{h''(g(x))}g'(x)^2 + \boxed{h'(g(x))}\boxed{g''(x)}$$

- *h*为凸，不减，*g*为凸，则*f*为凸
- *h*为凸，不增，*g*为凹，则*f*为凸

函数复合

- 函数 $g:\mathbf{R}^n \to \mathbf{R}^k$ 和 $h:\mathbf{R}^k \to \mathbf{R}$ 的复合：

$$f(x) = h(g(x))$$

- 假设n, k=1，定义域均为实数集，g和h二阶可微
- $\nabla^2 f(x) \succeq 0$ ⇔f为凸函数

$$f'(x) = h'(g(x))g'(x)$$

$$f''(x) = \boxed{h''(g(x))}g'(x)^2 + \boxed{h'(g(x))}\boxed{g''(x)}$$

  - h为凸，不减，g为凸，则f为凸
  - h为凸，不增，g为凹，则f为凸
  - h为凹，不减（增），g为凹（凸），则f为凹

函数复合

函数复合

- 假设扩展

函数复合

- 假设扩展
  - 高维：$n, k>1$

函数复合

- 假设扩展
  - 高维：*n*, *k*>1
  - 定义域不为全空间

函数复合

- 假设扩展
  - 高维：$n, k>1$
  - 定义域不为全空间
  - h，g均不可微

函数复合

- 假设扩展
  - 高维： $n, k>1$
  - 定义域不为全空间
  - h， g均不可微
- 扩展值延伸

函数复合

- 假设扩展
  - 高维： $n, k>1$
  - 定义域不为全空间
  - h， g均不可微
- 扩展值延伸
- 函数f为凸函数

函数复合

- 假设扩展
  - 高维：$n, k>1$
  - 定义域不为全空间
  - h，g均不可微
- 扩展值延伸
- 函数f为凸函数
  - 当g为凸函数，h为凸函数，且h的扩展值延伸非减

函数复合

- 假设扩展
  - 高维：$n, k>1$
  - 定义域不为全空间
  - h，g均不可微
- 扩展值延伸
- 函数f为凸函数
  - 当g为凸函数，h为凸函数，且h的扩展值延伸非减
  - 当g为凹函数，h为凸函数，且h的扩展值延伸非增

函数复合

- 假设扩展
  - 高维：$n, k>1$
  - 定义域不为全空间
  - h，g均不可微
- 扩展值延伸
- 函数f为凸函数
  - 当g为凸函数，h为凸函数，且h的扩展值延伸非减
  - 当g为凹函数，h为凸函数，且h的扩展值延伸非增
  - 注意：扩展值延伸须保证单调性

函数复合

函数复合

- 扩展值延伸无单调性

函数复合

- 扩展值延伸无单调性

$$g(x) = x^2 \quad \mathbf{dom}\, g = \mathbf{R}$$

函数复合

- 扩展值延伸无单调性

$$g(x) = x^2 \quad \mathbf{dom}\, g = \mathbf{R}$$

$$h(x) = 0 \quad \mathbf{dom}\, h = [1, 2]$$

函数复合

- 扩展值延伸无单调性

$$g(x) = x^2 \quad \mathbf{dom}\, g = \mathbf{R}$$

$$h(x) = 0 \quad \mathbf{dom}\, h = [1, 2]$$

$$f(x) = 0$$

函数复合

- 扩展值延伸无单调性

$$g(x) = x^2 \quad \mathbf{dom}\, g = \mathbf{R}$$

$$h(x) = 0 \quad \mathbf{dom}\, h = [1, 2]$$

$$f(x) = 0$$

$$\mathbf{dom}\, f = [-\sqrt{2}, -1] \cup [1, \sqrt{2}]$$

函数复合

函数复合

- 证明：当g为凸，h为凸，且h的扩展值延伸非减，则f为凸

函数复合

- 证明：当g为凸，h为凸，且h的扩展值延伸非减，则f为凸 $x, y \in \mathbf{dom}\, f$ $0 \leq \theta \leq 1$

函数复合

- 证明：当g为凸，h为凸，且h的扩展值延伸非减，则f为凸 x，$y \in \mathbf{dom}\, f$ $0 \leq \theta \leq 1$

$$\theta x + (1-\theta)y \in \mathbf{dom}\, g$$

函数复合

- 证明：当g为凸，h为凸，且h的扩展值延伸非减，则f为凸 x，$y \in \mathbf{dom}\, f$ $0 \le \theta \le 1$

$$\theta x + (1-\theta)y \in \mathbf{dom}\, g$$

$$g(\theta x + (1-\theta)y) \le \theta g(x) + (1-\theta)g(y)$$

函数复合

- 证明：当g为凸，h为凸，且h的扩展值延伸非减，则f为凸 x，$y \in \mathbf{dom}\, f$ $0 \le \theta \le 1$

$$\theta x + (1-\theta) y \in \mathbf{dom}\, g$$

$$g(\theta x + (1-\theta) y) \le \theta g(x) + (1-\theta) g(y)$$

$$\theta g(x) + (1-\theta) g(y) \in \mathbf{dom}\, h$$

函数复合

- 证明：当g为凸，h为凸，且h的扩展值延伸非减，则f为凸 x，$y \in \mathbf{dom}\, f$ $0 \le \theta \le 1$

$$\theta x + (1-\theta)y \in \mathbf{dom}\, g$$

$$g(\theta x + (1-\theta)y) \le \theta g(x) + (1-\theta)g(y)$$

$$\theta g(x) + (1-\theta)g(y) \in \mathbf{dom}\, h$$

$$h(\theta g(x) + (1-\theta)g(y)) \le \theta h(g(x)) + (1-\theta)h(g(y))$$

函数复合

- 证明：当g为凸，h为凸，且h的扩展值延伸非减，则f为凸 $x, \ y \in \textbf{dom}\, f \ \ 0 \le \theta \le 1$

$$\theta x + (1-\theta) y \in \textbf{dom}\, g$$

$$g(\theta x + (1-\theta) y) \le \theta g(x) + (1-\theta) g(y)$$

$$\theta g(x) + (1-\theta) g(y) \in \textbf{dom}\, h$$

$$h(\theta g(x) + (1-\theta) g(y)) \le \theta h(g(x)) + (1-\theta) \boxed{h(g(y))}$$

函数复合

- 证明：当g为凸，h为凸，且h的扩展值延伸非减，则f为凸 $x, y \in \mathbf{dom}\, f$ $0 \le \theta \le 1$

$$\theta x + (1-\theta) y \in \mathbf{dom}\, g$$

$$g(\theta x + (1-\theta)y) \le \theta g(x) + (1-\theta) g(y)$$

$$\theta g(x) + (1-\theta) g(y) \in \mathbf{dom}\, h$$

$$h(\theta g(x) + (1-\theta) g(y)) \le \theta h(g(x)) + (1-\theta)\boxed{h(g(y))}$$

$$f(y)$$

函数复合

- 证明：当g为凸，h为凸，且h的扩展值延伸非减，则f为凸 $x, y \in \mathbf{dom}\, f$ $0 \le \theta \le 1$

$$\theta x + (1-\theta)y \in \mathbf{dom}\, g$$

$$g(\theta x + (1-\theta)y) \le \theta g(x) + (1-\theta)g(y)$$

$$\theta g(x) + (1-\theta)g(y) \in \mathbf{dom}\, h$$

$$h(\theta g(x) + (1-\theta)g(y)) \le \theta \boxed{h(g(x))} + (1-\theta)\boxed{h(g(y))}$$

$$f(y)$$

函数复合

- 证明：当g为凸，h为凸，且h的扩展值延伸非减，则f为凸 $x, \ y \in \mathbf{dom}\, f \ 0 \le \theta \le 1$

$$\theta x + (1-\theta) y \in \mathbf{dom}\, g$$

$$g(\theta x + (1-\theta) y) \le \theta g(x) + (1-\theta) g(y)$$

$$\theta g(x) + (1-\theta) g(y) \in \mathbf{dom}\, h$$

$$h(\theta g(x) + (1-\theta) g(y)) \le \theta \underbrace{\boxed{h(g(x))}}_{f(x)} + (1-\theta) \underbrace{\boxed{h(g(y))}}_{f(y)}$$

函数复合

- 证明：当g为凸，h为凸，且h的扩展值延伸非减，则f为凸 $x, \ y \in \mathbf{dom}\, f$ $0 \leq \theta \leq 1$

$$\theta x + (1-\theta) y \in \mathbf{dom}\, g$$

$$g(\theta x + (1-\theta) y) \leq \theta g(x) + (1-\theta) g(y)$$

$$\theta g(x) + (1-\theta) g(y) \in \mathbf{dom}\, h$$

$$\boxed{h(\theta g(x) + (1-\theta) g(y))} \leq \theta \underbrace{\boxed{h(g(x))}}_{f(x)} + (1-\theta) \underbrace{\boxed{h(g(y))}}_{f(y)}$$

函数复合

- 证明：当g为凸，h为凸，且h的扩展值延伸非减，则f为凸 $x, \; y \in \mathbf{dom}\, f \;\; 0 \le \theta \le 1$

$$\theta x + (1-\theta) y \in \mathbf{dom}\, g$$

$$g(\theta x + (1-\theta)y) \le \theta g(x) + (1-\theta) g(y)$$

$$\theta g(x) + (1-\theta) g(y) \in \mathbf{dom}\, h$$

$$\boxed{h(\theta g(x) + (1-\theta)g(y))} \le \theta \boxed{h(g(x))} + (1-\theta)\boxed{h(g(y))}$$

$$f(x) \qquad\qquad f(y)$$

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

函数复合

- 证明：当g为凸，h为凸，且h的扩展值延伸非减，则f为凸 $x,\ y \in \textbf{dom}\, f\ \ 0 \le \theta \le 1$

$$\theta x + (1-\theta)y \in \textbf{dom}\, g$$

$$g(\theta x + (1-\theta)y) \le \theta g(x) + (1-\theta)g(y)$$

$$\theta g(x) + (1-\theta)g(y) \in \textbf{dom}\, h$$

$$\boxed{h(\theta g(x) + (1-\theta)g(y))} \le \theta \underbrace{\boxed{h(g(x))}}_{f(x)} + (1-\theta)\underbrace{\boxed{h(g(y))}}_{f(y)}$$

$\le$

$$f(\theta x + (1-\theta)y) = h(g(\theta x + (1-\theta)y))$$

函数复合

函数复合

- 若g为凸，则 $\exp g(x)$为凸函数

函数复合

- 若g为凸，则 $\exp g(x)$为凸函数
  - $h(x)=\exp x$，凸，扩展值不减

函数复合

- 若g为凸，则 $\exp g(x)$为凸函数
  - $h(x)=\exp x$，凸，扩展值不减
- 若g为凹函数且为正，则 $\log g(x)$为凹函数

函数复合

- 若g为凸，则 $\exp g(x)$为凸函数
  - $h(x)=\exp x$，凸，扩展值不减
- 若g为凹函数且为正，则 $\log g(x)$为凹函数
  - $h(x)=\log x$, 凹，扩展值不减

函数复合

- 若g为凸，则 $\exp g(x)$为凸函数
  - $h(x)=\exp x$，凸，扩展值不减
- 若g为凹函数且为正，则 $\log g(x)$为凹函数
  - $h(x)=\log x$, 凹，扩展值不减
- 若g为凹函数且为正，则 $1/g(x)$ 为凸函数

函数复合

- 若g为凸，则 $\exp g(x)$为凸函数
  - $h(x)=\exp x$，凸，扩展值不减
- 若g为凹函数且为正，则 $\log g(x)$为凹函数
  - $h(x)=\log x$, 凹，扩展值不减
- 若g为凹函数且为正，则 $1/g(x)$ 为凸函数
  - $h(x)=1/x$, 凸，扩展值不增

函数复合

- 若g为凸，则 $\exp g(x)$为凸函数
  - $h(x)=\exp x$，凸，扩展值不减
- 若g为凹函数且为正，则 $\log g(x)$为凹函数
  - $h(x)=\log x$, 凹，扩展值不减
- 若g为凹函数且为正，则 $1/g(x)$ 为凸函数
  - $h(x)=1/x$, 凸，扩展值不增
- 若g为凸函数且非负，$p\geq\mathbf{1}$，则 $g(x)^p$ 为凸函数

函数复合

- 若g为凸，则 $\exp g(x)$为凸函数
  - $h(x)=\exp x$，凸，扩展值不减
- 若g为凹函数且为正，则 $\log g(x)$为凹函数
  - $h(x)=\log x$, 凹，扩展值不减
- 若g为凹函数且为正，则 $1/g(x)$ 为凸函数
  - $h(x)=1/x$, 凸，扩展值不增
- 若g为凸函数且非负，$p\geq\mathbf{1}$，则 $g(x)^p$ 为凸函数
  - $h(x) = x^p(x \geq 0)$, 凸，扩展值不减

函数复合

- 若g为凸，则 $\exp g(x)$为凸函数
  - $h(x)=\exp x$，凸，扩展值不减
- 若g为凹函数且为正，则 $\log g(x)$为凹函数
  - $h(x)=\log x$, 凹，扩展值不减
- 若g为凹函数且为正，则 $1/g(x)$ 为凸函数
  - $h(x)=1/x$, 凸，扩展值不增
- 若g为凸函数且非负，$p\geq\mathbf{1}$，则 $g(x)^p$ 为凸函数
  - $h(x) = x^p(x \geq 0)$, 凸，扩展值不减

函数复合

- 若g为凸，则 $\exp g(x)$为凸函数
  - $h(x)=\exp x$，凸，扩展值不减
- 若g为凹函数且为正，则 $\log g(x)$为凹函数
  - $h(x)=\log x$, 凹，扩展值不减
- 若g为凹函数且为正，则 $1/g(x)$ 为凸函数
  - $h(x)=1/x$, 凸，扩展值不增
- 若g为凸函数且非负，$p\geq\mathbf{1}$，则 $g(x)^p$ 为凸函数
  - $h(x) = x^p(x \geq 0)$, 凸，扩展值不减
  - $h(x) = \begin{cases} x^p & \textbf{if } x \geq 0 \\ 0 & \textbf{if } x < 0 \end{cases}$

最小化

最小化

- 若$f(x, y)$是关于(x, y)的凸函数，且集合C为凸集

最小化

- 若$f(x,y)$是关于(x,y)的凸函数，且集合C为凸集
  - $g(x)=\inf_{y\in C} f(x,y)$为凸函数

最小化

- 若$f(x,y)$是关于(x,y)的凸函数，且集合C为凸集
  - $g(x)=\inf_{y\in C} f(x,y)$为凸函数
- 例：$f(x,y)=x^TAx+2x^TBy+y^TCy$

最小化

- 若$f(x,y)$是关于(x,y)的凸函数，且集合C为凸集
  - $g(x)=\inf\limits_{y\in C} f(x,y)$为凸函数
- 例：$f(x,y)=x^TAx+2x^TBy+y^TCy$
  - 其中，$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix}\geq 0$，$C>0$

最小化

- 若$f(x,y)$是关于(x,y)的凸函数，且集合C为凸集
  - $g(x)=\inf\limits_{y\in C} f(x,y)$为凸函数
- 例：$f(x,y)=x^TAx+2x^TBy+y^TCy$
  - 其中，$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \geq 0$，$C>0$
  - $g(x)=\inf\limits_{y} f(x,y)=x^T(A-BC^{-1}B^T)x$

最小化

- 若$f(x,y)$是关于(x,y)的凸函数，且集合C为凸集
  - $g(x)=\inf\limits_{y\in C} f(x,y)$为凸函数
- 例：$f(x,y)=x^TAx+2x^TBy+y^TCy$
  - 其中，$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \geq 0$，$C>0$
  - $g(x)=\inf\limits_{y} f(x,y)=x^T(A-BC^{-1}B^T)x$
  - $g(x)$为凸函数

最小化

- 若$f(x,y)$是关于(x,y)的凸函数，且集合C为凸集
  - $g(x)=\inf_{y\in C} f(x,y)$为凸函数
- 例：$f(x,y)=x^T A x+2x^T B y+y^T C y$
  - 其中，$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \geq 0$，$C>0$
  - $g(x)=\inf_{y} f(x,y)=x^T(A-BC^{-1}B^T)x$
  - $g(x)$为凸函数
  - $A-BC^{-1}B^T \geq 0$

透视函数

透视函数

- 函数 $f : \mathbf{R}^n \to \mathbf{R}$ 的透视函数形如 $g : \mathbf{R}^n \times \mathbf{R} \to \mathbf{R}$

$$g(x,t) = tf(x/t), \qquad \mathbf{dom}\, g = \{(x,t) \mid x/t \in \mathbf{dom}\, f,\ t > 0\}$$

透视函数

- 函数 $f : \mathbf{R}^n \to \mathbf{R}$ 的透视函数形如 $g : \mathbf{R}^n \times \mathbf{R} \to \mathbf{R}$

$$g(x,t) = tf(x/t), \qquad \mathbf{dom}\, g = \{(x,t) \mid x/t \in \mathbf{dom}\, f,\ t > 0\}$$

  - 若f为凸函数，g为凸函数。

透视函数

- 函数 $f : \mathbf{R}^n \to \mathbf{R}$ 的透视函数形如 $g : \mathbf{R}^n \times \mathbf{R} \to \mathbf{R}$

$$g(x,t) = tf(x/t), \qquad \mathbf{dom}\, g = \{(x,t) \mid x/t \in \mathbf{dom}\, f,\ t > 0\}$$

  - 若f为凸函数，g为凸函数。
- 例：欧几里得范数的平方 $f(x) = x^T x$ 为凸函数

透视函数

- 函数 $f : \mathbf{R}^n \to \mathbf{R}$ 的透视函数形如 $g : \mathbf{R}^n \times \mathbf{R} \to \mathbf{R}$

$$g(x,t) = tf(x/t), \qquad \mathbf{dom}\, g = \{(x,t) \mid x/t \in \mathbf{dom}\, f,\ t > 0\}$$

  - 若f为凸函数，g为凸函数。

- 例： 欧几里得范数的平方 $f(x) = x^T x$ 为凸函数

$$g(x,t) = t(x/t)^T (x/t)$$

透视函数

- 函数 $f: \mathbf{R}^n \to \mathbf{R}$ 的透视函数形如 $g: \mathbf{R}^n \times \mathbf{R} \to \mathbf{R}$

$$g(x,t) = tf(x/t), \qquad \mathbf{dom}\, g = \{(x,t) \mid x/t \in \mathbf{dom}\, f,\ t>0\}$$

  - 若f为凸函数，g为凸函数。
- 例：欧几里得范数的平方 $f(x) = x^T x$ 为凸函数

$$g(x,t) = t(x/t)^T(x/t) = \frac{x^T x}{t}$$

透视函数

- 函数 $f:\mathbf{R}^n \to \mathbf{R}$ 的透视函数形如 $g:\mathbf{R}^n \times \mathbf{R} \to \mathbf{R}$

$$g(x,t) = tf(x/t), \qquad \mathbf{dom}\, g = \{(x,t) \mid x/t \in \mathbf{dom}\, f,\ t > 0\}$$

  - 若f为凸函数，g为凸函数。

- 例：欧几里得范数的平方 $f(x) = x^T x$ 为凸函数

$$g(x,t) = t(x/t)^T(x/t) = \frac{x^T x}{t}$$

- 例：负的对数 $f(x) = -\log x$

透视函数

- 函数 $f:\mathbf{R}^n \to \mathbf{R}$ 的透视函数形如 $g:\mathbf{R}^n \times \mathbf{R} \to \mathbf{R}$

$$g(x,t) = tf(x/t), \qquad \mathbf{dom}\, g = \{(x,t) \mid x/t \in \mathbf{dom}\, f,\ t>0\}$$

  - 若f为凸函数，g为凸函数。

- 例：欧几里得范数的平方 $f(x) = x^T x$ 为凸函数

$$g(x,t) = t(x/t)^T(x/t) = \frac{x^T x}{t}$$

- 例：负的对数 $f(x) = -\log x$

$$g(x,t) = -t\log(x/t) = t\log(t/x)$$

透视函数

- 函数 $f:\mathbf{R}^n\to\mathbf{R}$ 的透视函数形如 $g:\mathbf{R}^n\times\mathbf{R}\to\mathbf{R}$

  $g(x,t)=tf(x/t),\qquad \mathbf{dom}\,g=\{(x,t)\mid x/t\in\mathbf{dom}\,f,\ t>0\}$

  - 若f为凸函数，g为凸函数。

- 例：欧几里得范数的平方 $f(x)=x^Tx$ 为凸函数

$$g(x,t)=t(x/t)^T(x/t)=\frac{x^Tx}{t}$$

- 例：负的对数 $f(x)=-\log x$

$$g(x,t)=-t\log(x/t)=t\log(t/x)$$
$$=t\log t-t\log x$$

透视函数

透视函数

- 向量的相对熵

透视函数

- 向量的相对熵

$$\sum_{i=1}^{n} u_i \log(u_i/v_i)$$

透视函数

- 向量的相对熵

$$\sum_{i=1}^{n} u_i \log(u_i/v_i)$$

  - 注意：不是非负加权和

透视函数

- 向量的相对熵

$$\sum_{i=1}^{n} u_i \log(u_i/v_i)$$

  - 注意：不是非负加权和

- **KL**散度

透视函数

- 向量的相对熵

$$\sum_{i=1}^{n} u_i \log(u_i/v_i)$$

  - 注意：不是非负加权和

- **KL**散度

$$D_{\mathrm{kl}}(u,v) = \sum_{i=1}^{n} \left(u_i \log(u_i/v_i) - u_i + v_i\right)$$

透视函数

- 向量的相对熵

$$\sum_{i=1}^{n} u_i \log(u_i/v_i)$$

  - 注意：不是非负加权和

- **KL**散度

$$D_{\mathrm{kl}}(u,v)=\sum_{i=1}^{n}\left(u_i \log(u_i/v_i)-u_i+v_i\right)$$

  - 凸函数

下水平集

下水平集

- 函数 $f: \mathbf{R}^n \to \mathbf{R}$ 的**α-**下水平集：

下水平集

- 函数 $f: \mathbf{R}^n \to \mathbf{R}$ 的**α-**下水平集：

$$C_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \le \alpha\}$$

下水平集

- 函数 $f: \mathbf{R}^n \to \mathbf{R}$ 的**α-**下水平集：

$$C_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \le \alpha\}$$

- 凸函数的所有下水平集为凸集

下水平集

- 函数 $f : \mathbf{R}^n \to \mathbf{R}$ 的**α-**下水平集：

$$C_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \le \alpha\}$$

- 凸函数的所有下水平集为凸集
  - 对于 $x,\ y \in C_\alpha \quad 0 \le \theta \le 1$

下水平集

- 函数 $f:\mathbf{R}^n \to \mathbf{R}$ 的α-下水平集：

$$C_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \le \alpha\}$$

- 凸函数的所有下水平集为凸集
  - 对于 $x,\ y \in C_\alpha \quad 0 \le \theta \le 1$

$$f(x) \le \alpha \quad f(y) \le \alpha$$

下水平集

- 函数 $f:\mathbf{R}^n \to \mathbf{R}$ 的**α-**下水平集：

$$C_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \le \alpha\}$$

- 凸函数的所有下水平集为凸集
  - 对于 $x,\ y \in C_\alpha \quad 0 \le \theta \le 1$

$$f(x) \le \alpha \quad f(y) \le \alpha$$

$$f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta)f(y)$$

下水平集

- 函数 $f:\mathbf{R}^n \rightarrow \mathbf{R}$ 的**α-**下水平集：

$$C_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \le \alpha\}$$

- 凸函数的所有下水平集为凸集
  - 对于 $x,\ y \in C_\alpha \quad 0 \le \theta \le 1$

$$f(x) \le \alpha \quad f(y) \le \alpha$$

$$f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta) f(y)$$
$$\le \alpha$$

下水平集

- 函数 $f : \mathbf{R}^n \to \mathbf{R}$ 的α-下水平集：

$$C_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \le \alpha\}$$

- 凸函数的所有下水平集为凸集
  - 对于 $x,\ y \in C_\alpha \quad 0 \le \theta \le 1$

$$f(x) \le \alpha \quad f(y) \le \alpha$$

$$\begin{aligned} f(\theta x + (1-\theta)y) &\le \theta f(x) + (1-\theta) f(y) \\ &\le \alpha \end{aligned}$$

$$\theta x + (1-\theta) y \in C_\alpha$$

下水平集

- 函数 $f : \mathbf{R}^n \to \mathbf{R}$ 的**α-**下水平集：

$$C_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \leq \alpha\}$$

- 凸函数的所有下水平集为凸集
  - 对于 $x,\ y \in C_\alpha \quad 0 \leq \theta \leq 1$

$$f(x) \leq \alpha \quad f(y) \leq \alpha$$

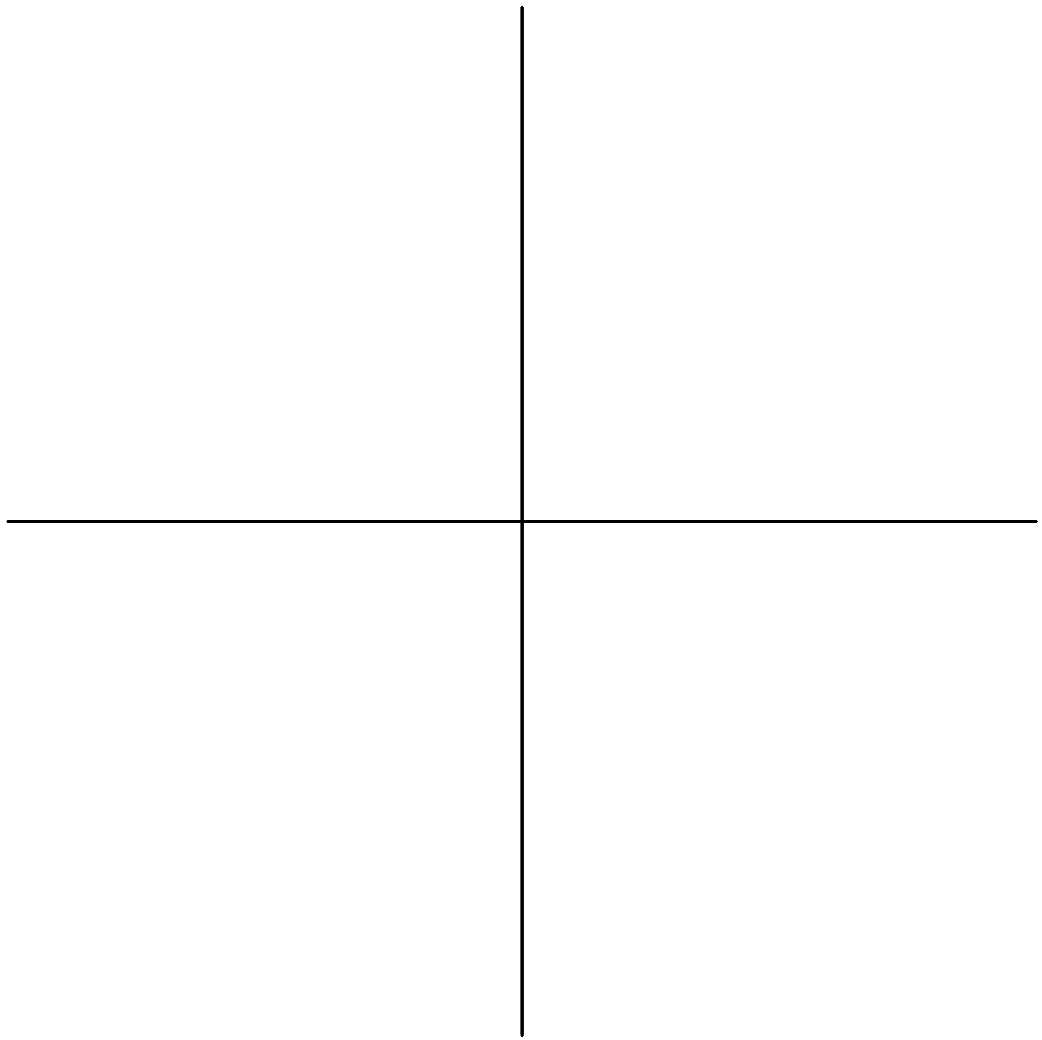
$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$$
$$\leq \alpha$$

$$\theta x + (1-\theta) y \in C_\alpha$$

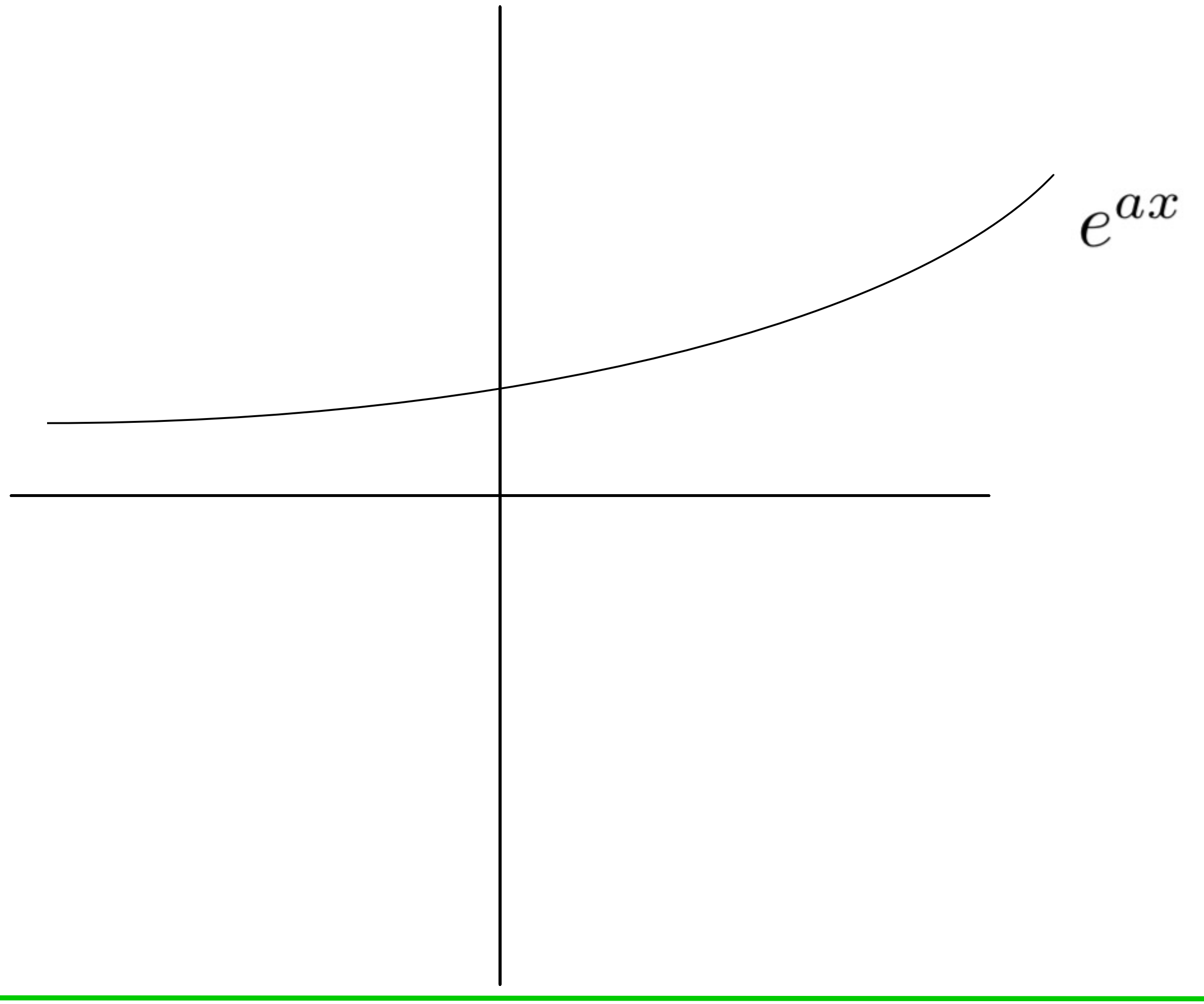
  - 反之不一定正确

下水平集

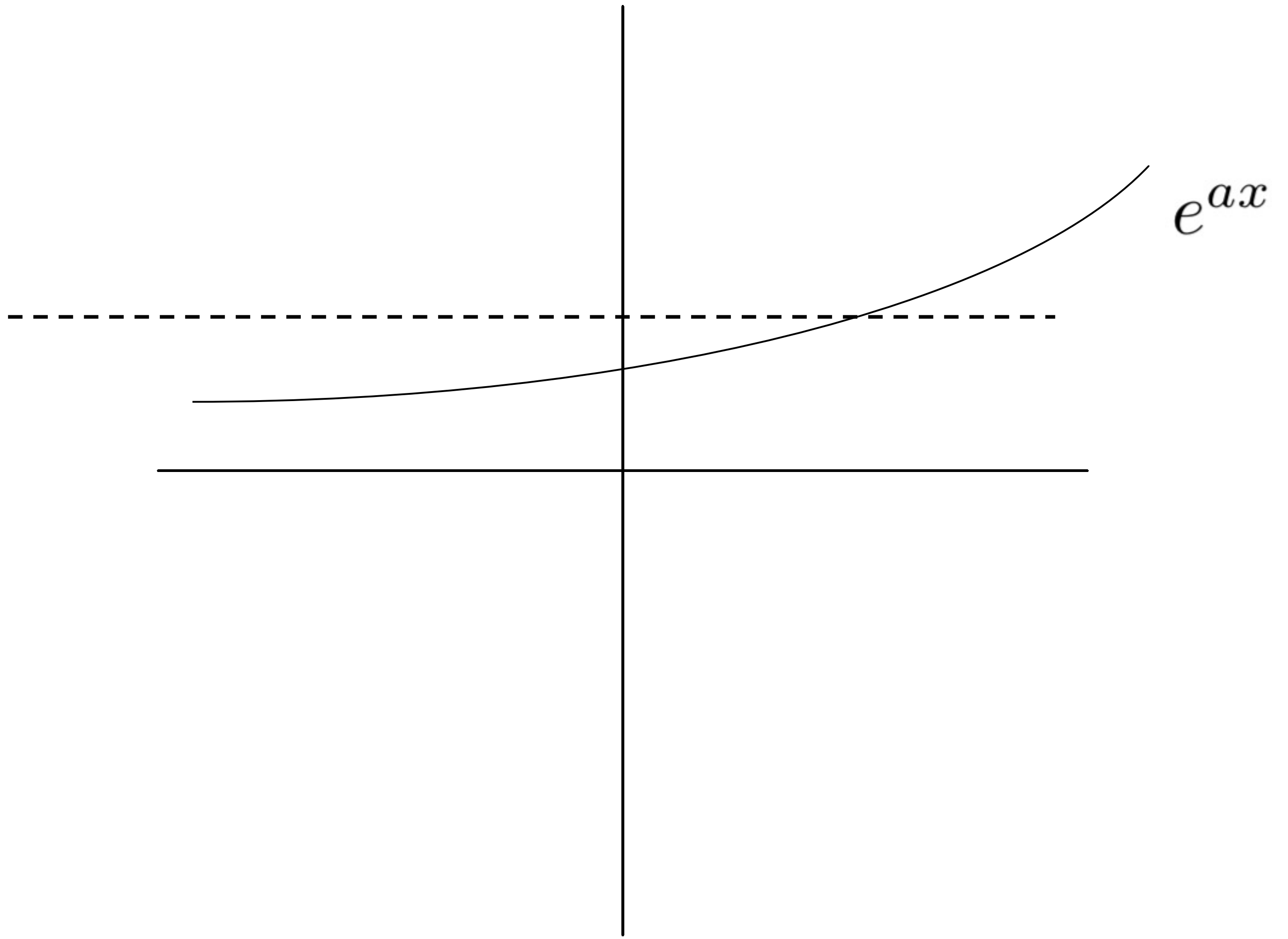
下水平集



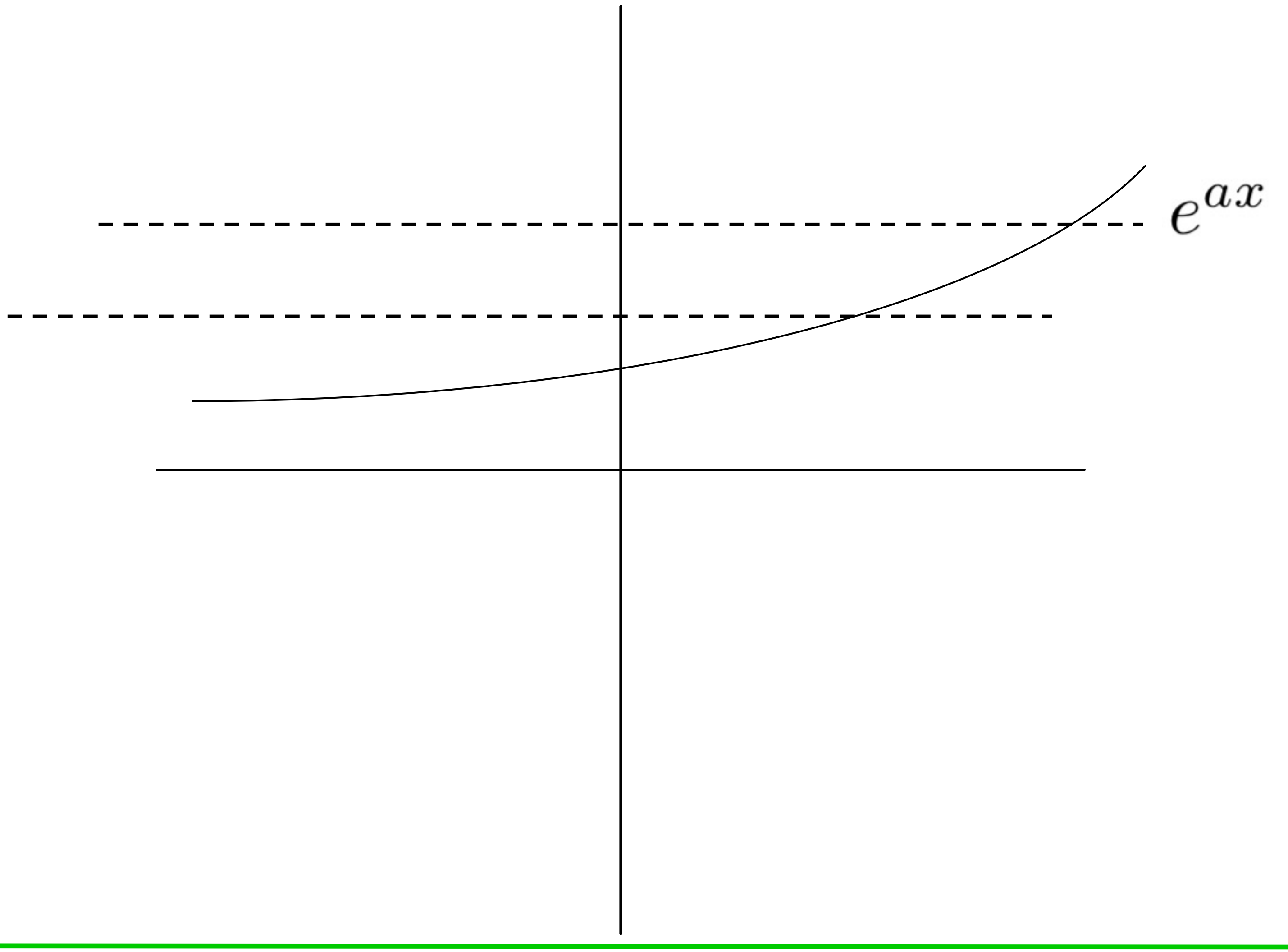
下水平集



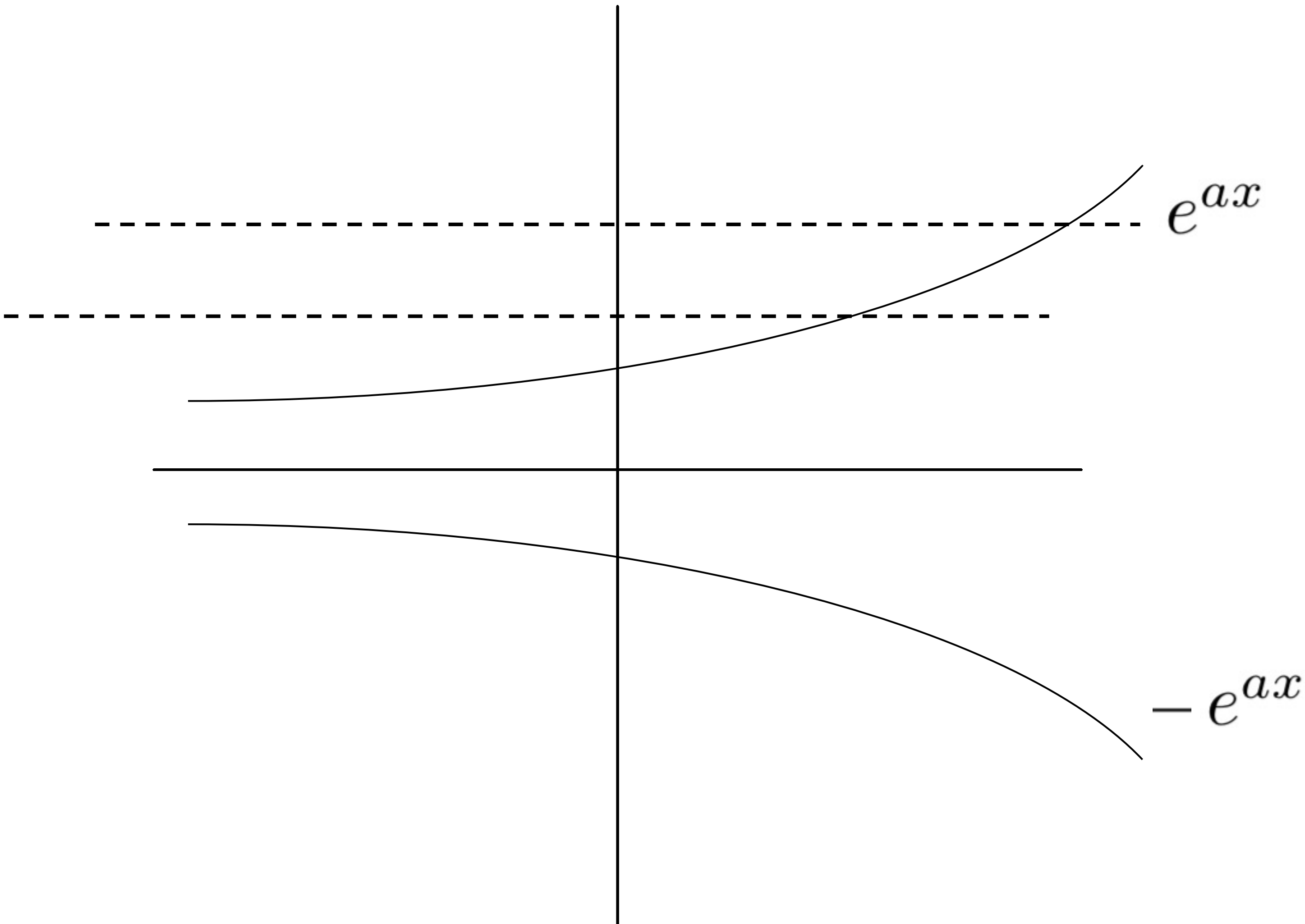
下水平集



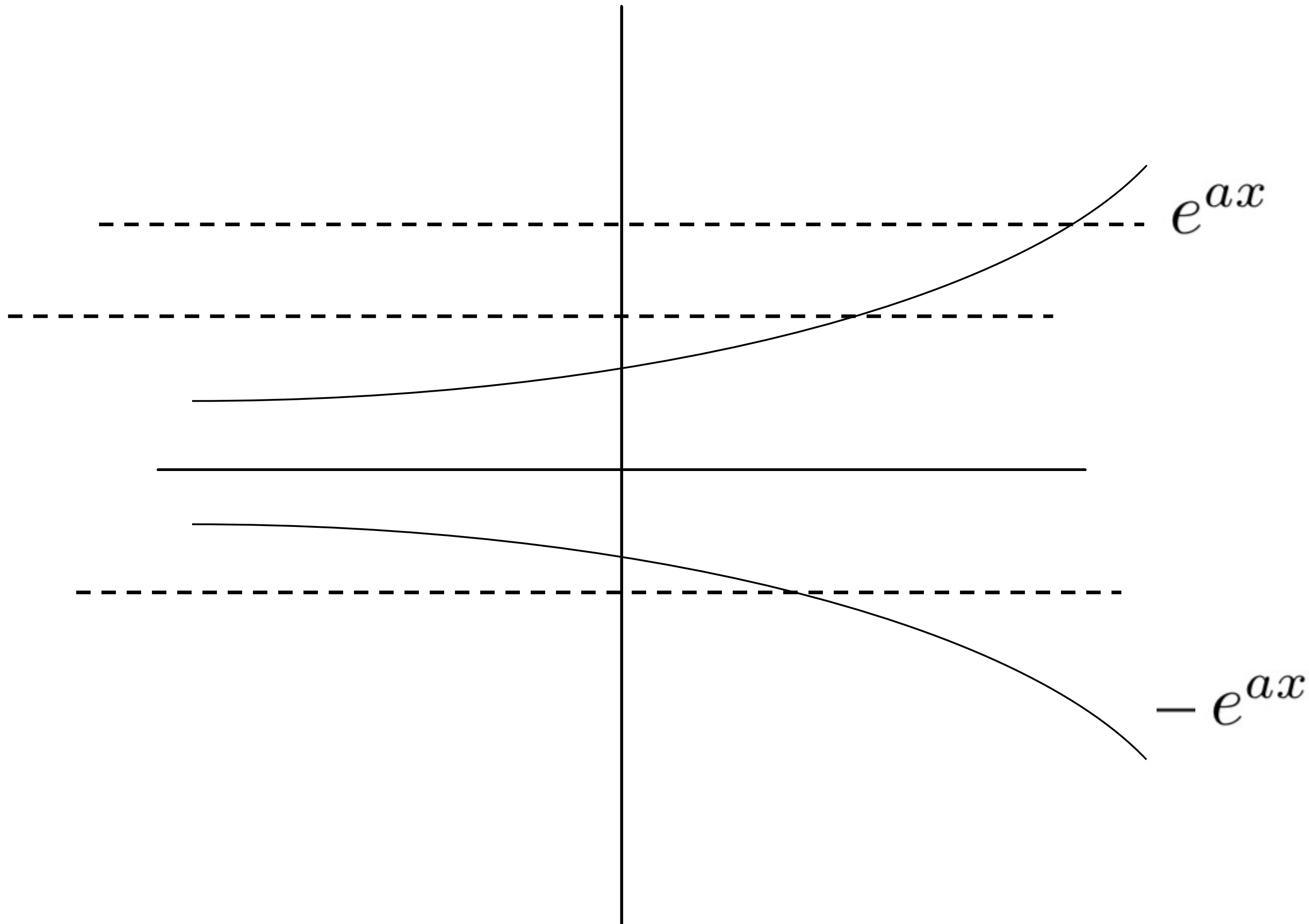
下水平集



下水平集



下水平集



上镜图

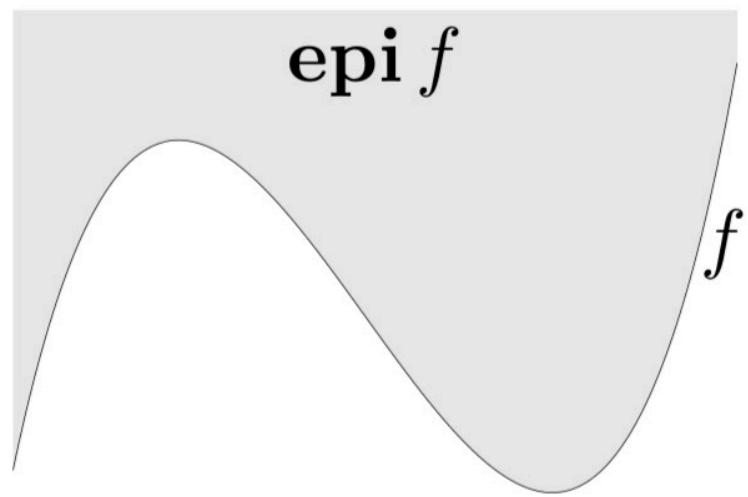
上镜图

- 函数 $f : \mathbf{R}^n \to \mathbf{R}$ 的上镜图:

上镜图

- 函数 $f : \mathbf{R}^n \to \mathbf{R}$ 的上镜图：

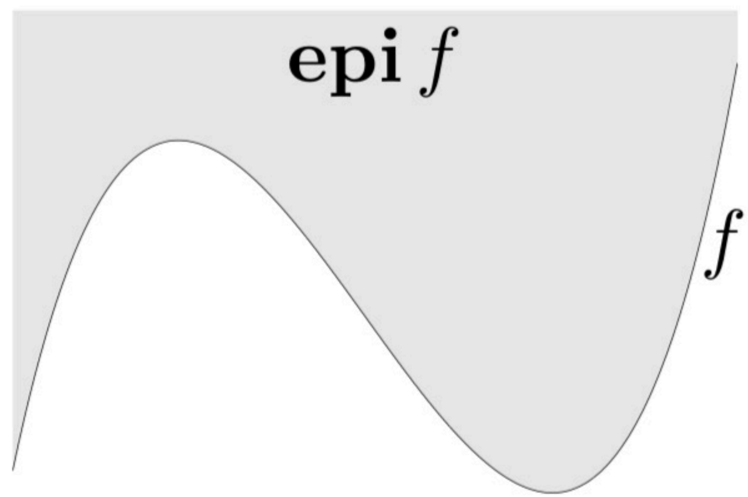
$$\mathbf{epi}\, f = \{(x,t) \in \mathbf{R}^{n+1} \mid x \in \mathbf{dom}\, f,\ \ f(x) \le t\}$$



上镜图

- 函数 $f : \mathbf{R}^n \to \mathbf{R}$ 的上镜图：

$$\textbf{epi}\, f = \{(x, t) \in \mathbf{R}^{n+1} \mid x \in \textbf{dom}\, f,\ f(x) \le t\}$$



- 函数为凸函数，当且仅当其上镜图是凸集

共轭函数

共轭函数

- 函数f的共轭函数为

共轭函数

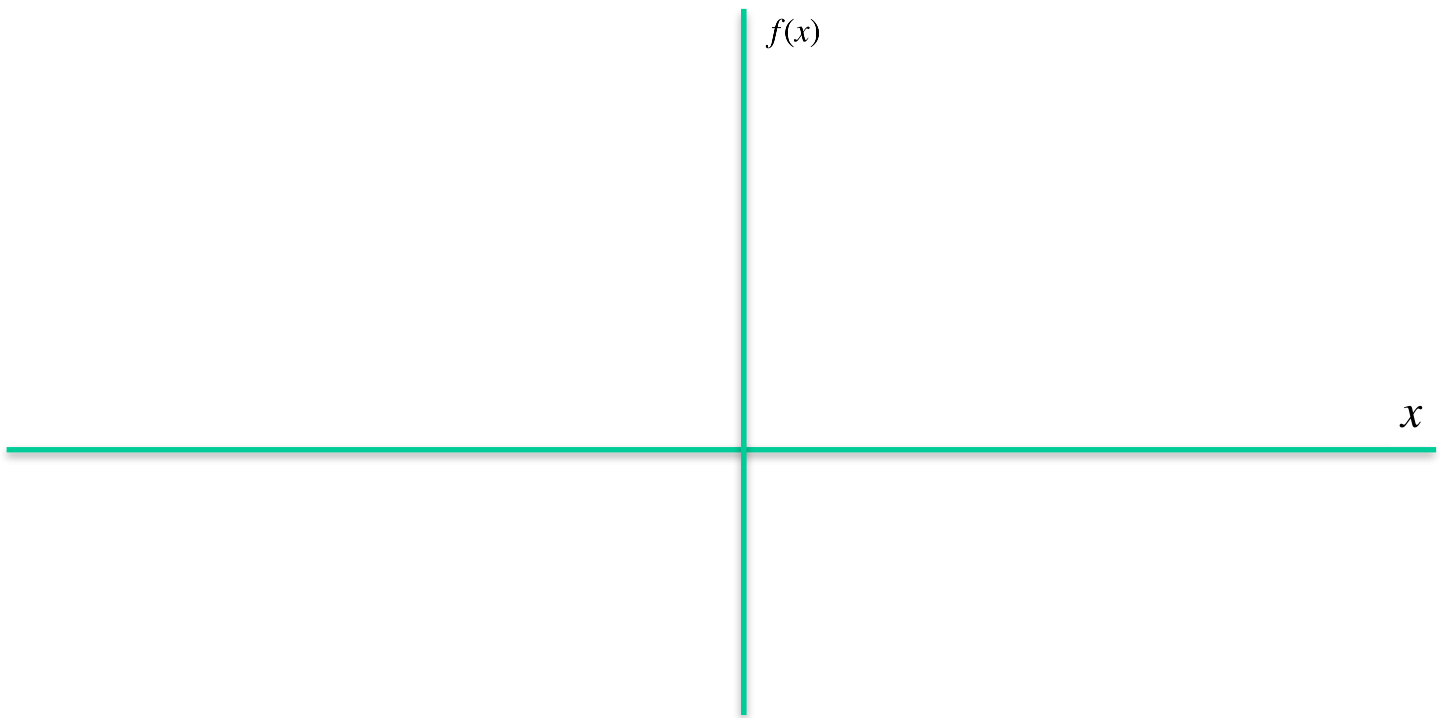
- 函数f的共轭函数为

$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} \left(y^T x - f(x) \right)$$

共轭函数

- 函数f的共轭函数为

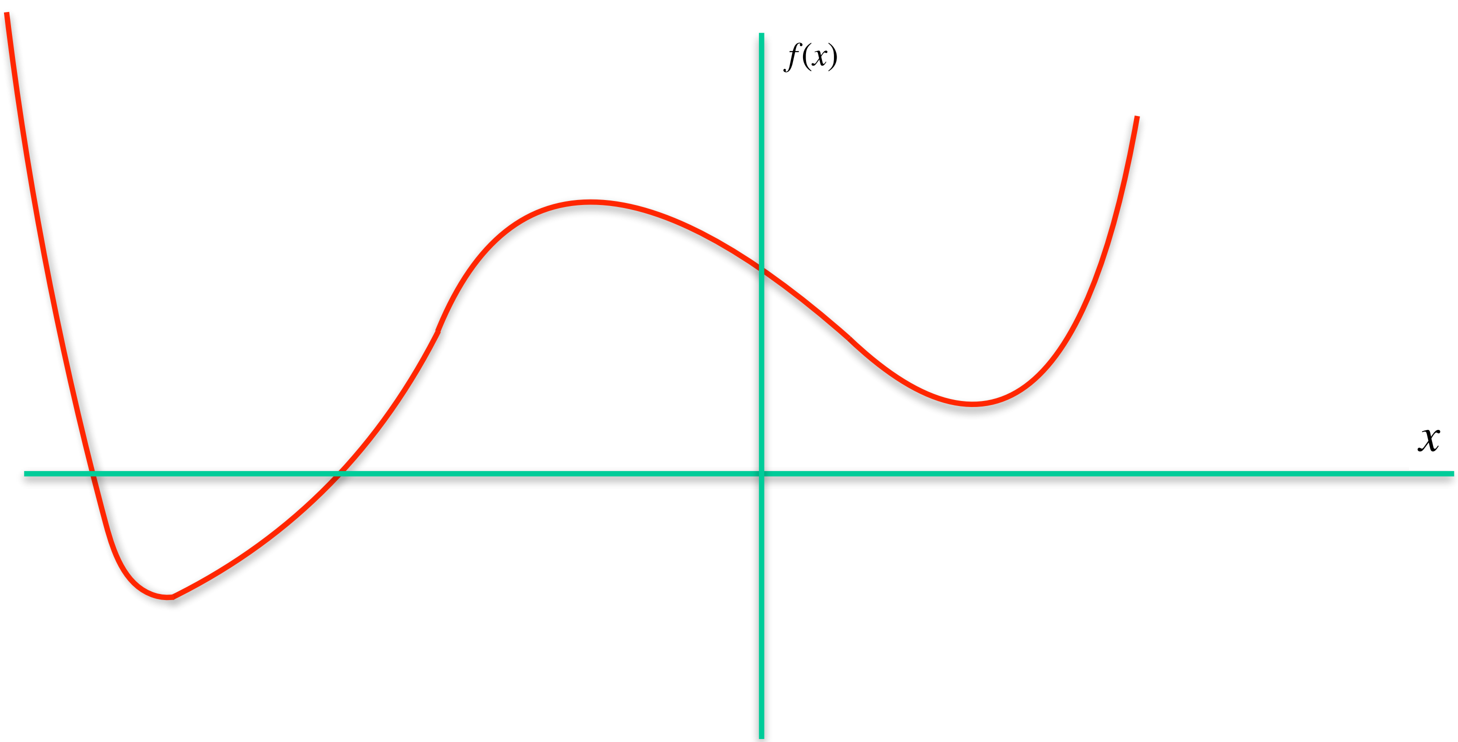
$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} \left(y^T x - f(x)\right)$$



共轭函数

- 函数*f*的共轭函数为

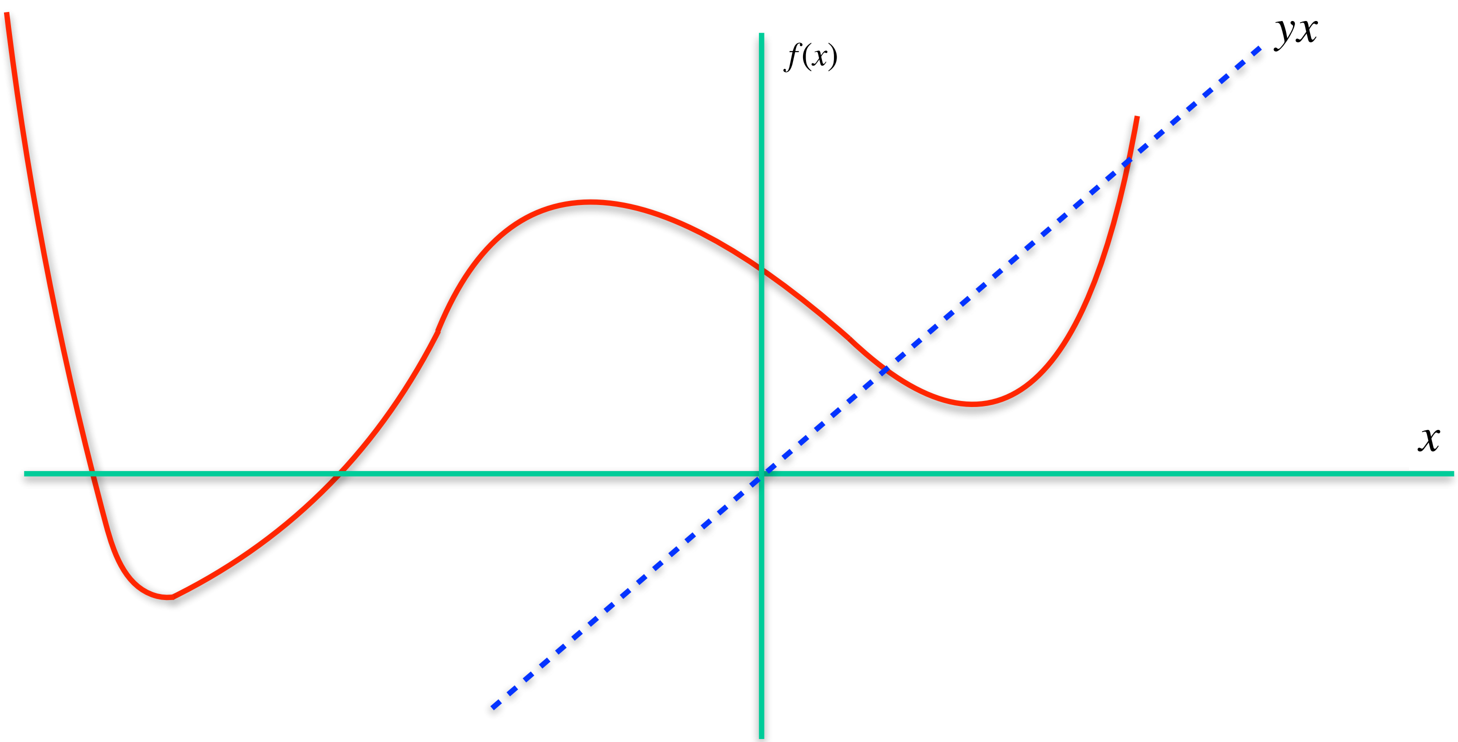
$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} (y^T x - f(x))$$



共轭函数

- 函数f的共轭函数为

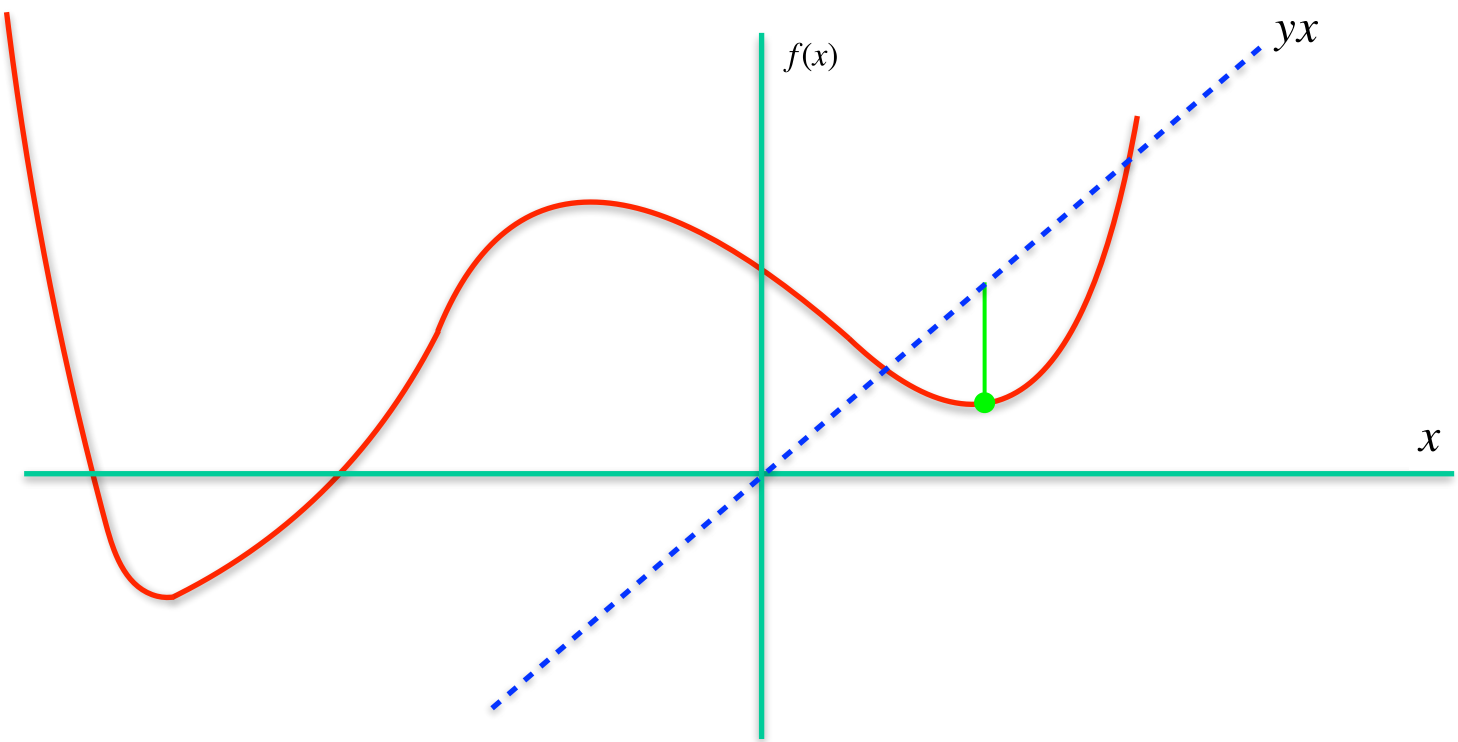
$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} \left(y^T x - f(x) \right)$$



共轭函数

- 函数f的共轭函数为

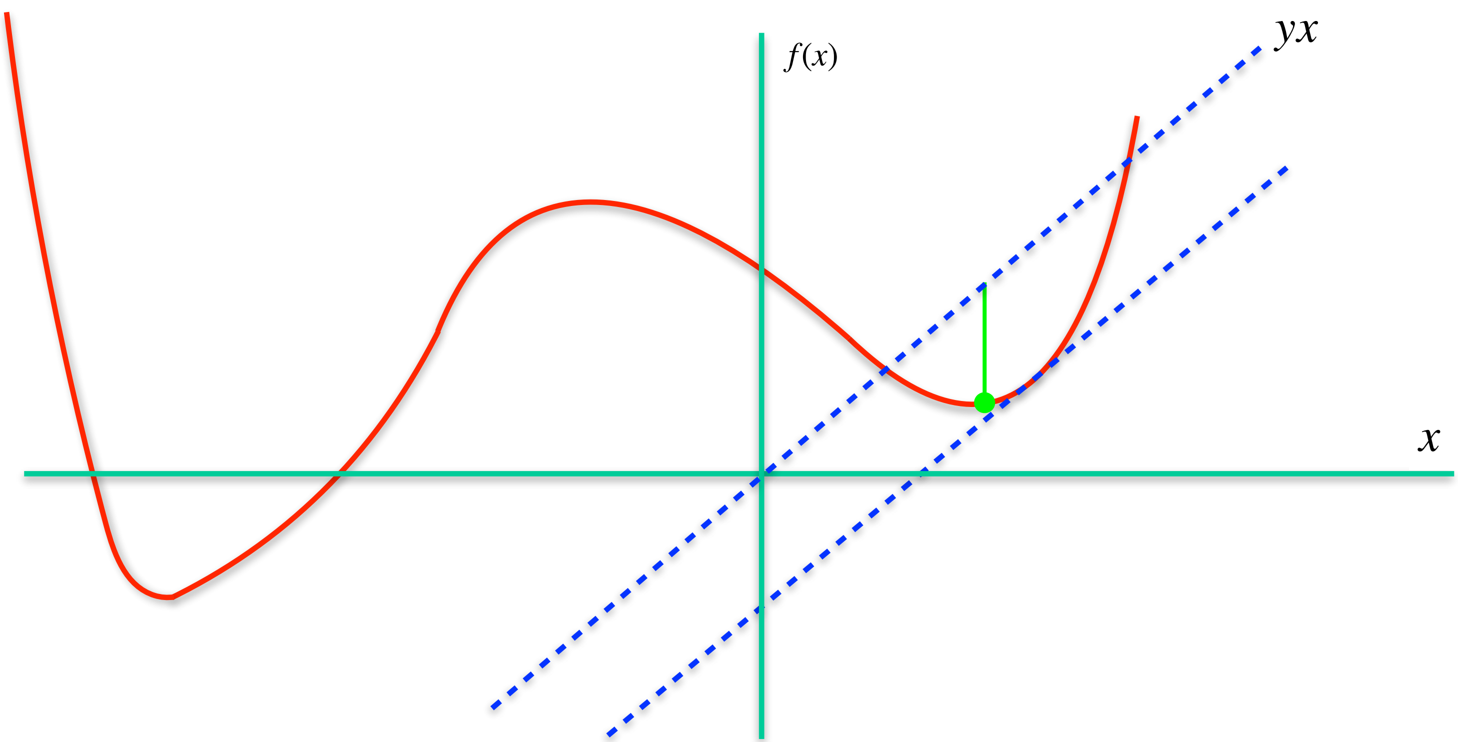
$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} \left(y^T x - f(x) \right)$$



共轭函数

- 函数f的共轭函数为

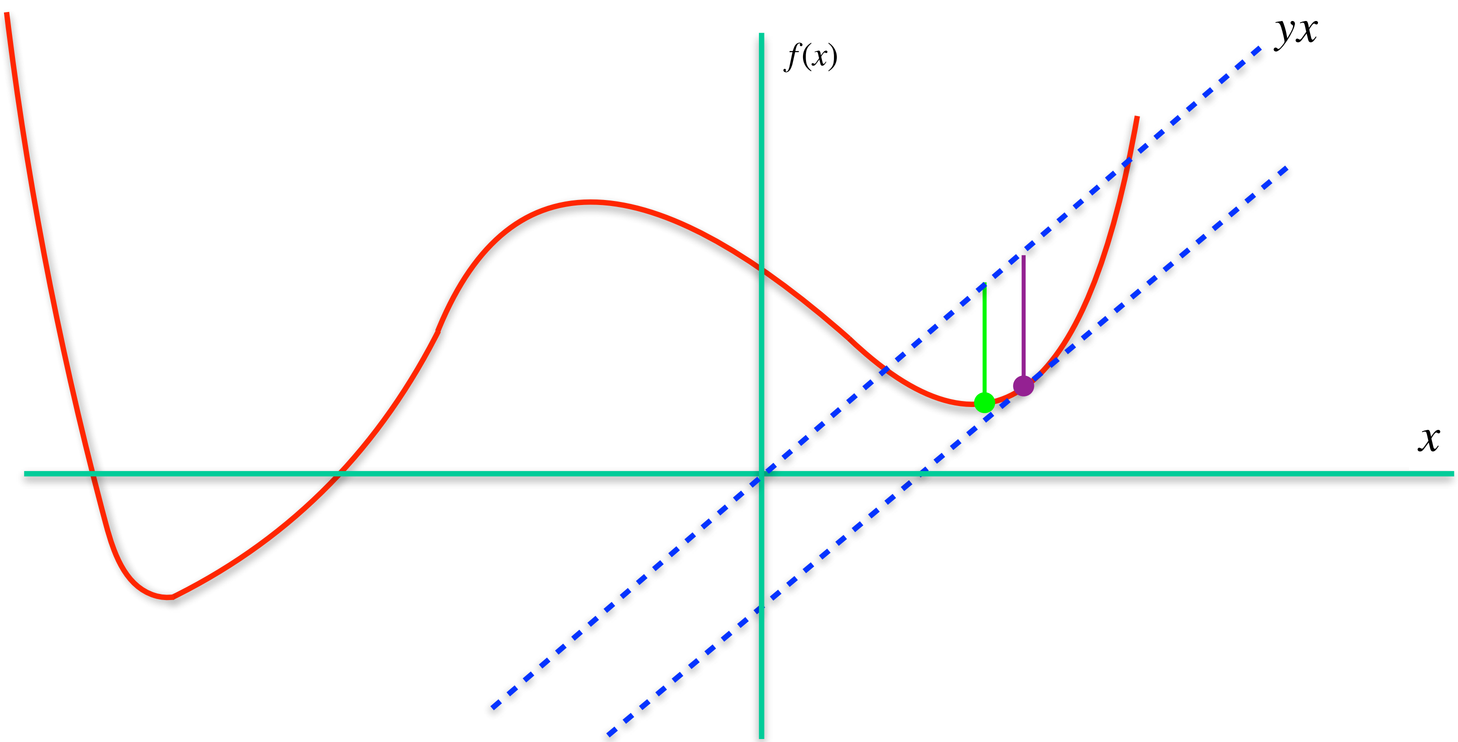
$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} \left(y^T x - f(x) \right)$$



共轭函数

- 函数f的共轭函数为

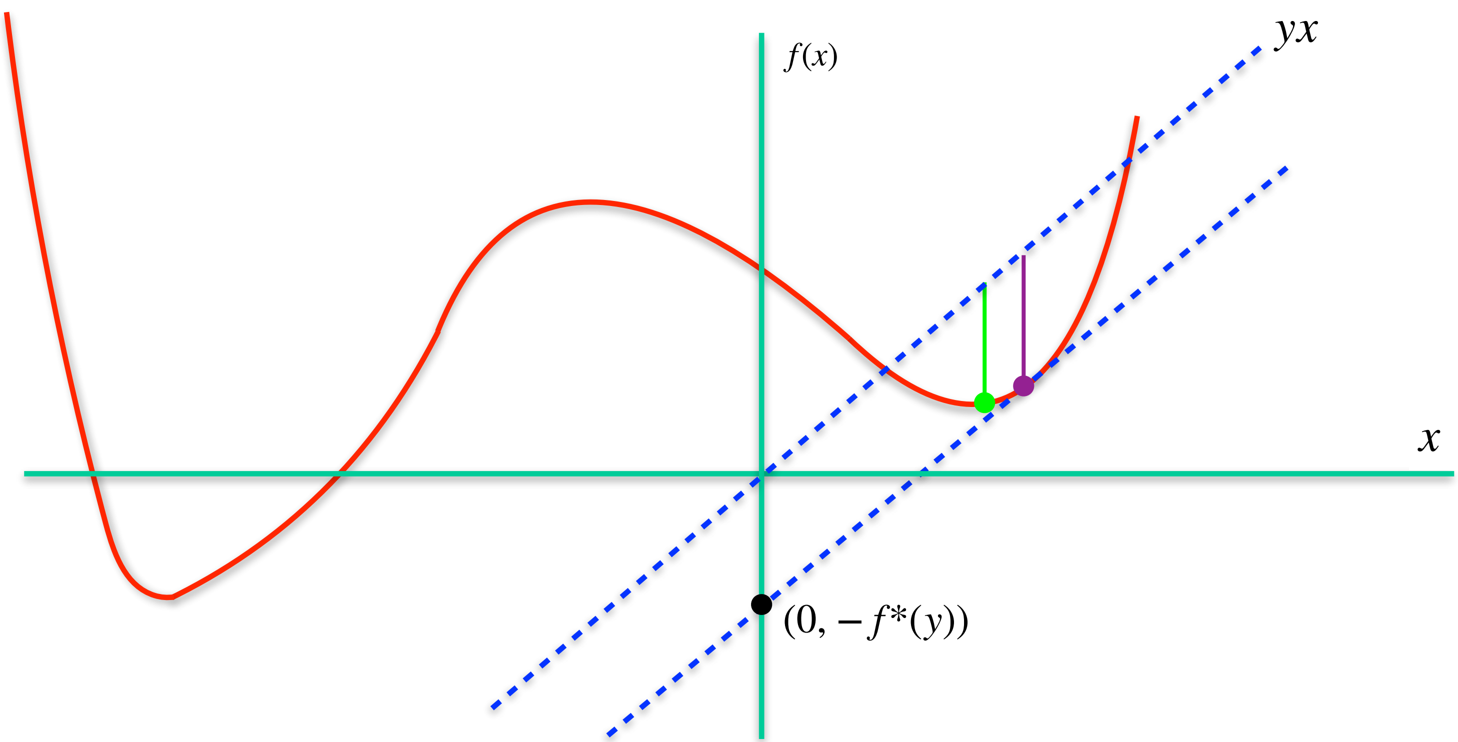
$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} \left(y^T x - f(x) \right)$$



共轭函数

- 函数f的共轭函数为

$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} \left(y^T x - f(x)\right)$$



共轭函数

共轭函数

- 若f可微，则$f^*(y)$对应的x必满足$f'(x) = y$

共轭函数

- 若f可微，则$f^*(y)$对应的x必满足$f'(x) = y$
- 函数f^*为凸函数
  - 即使函数f不是凸函数

共轭函数

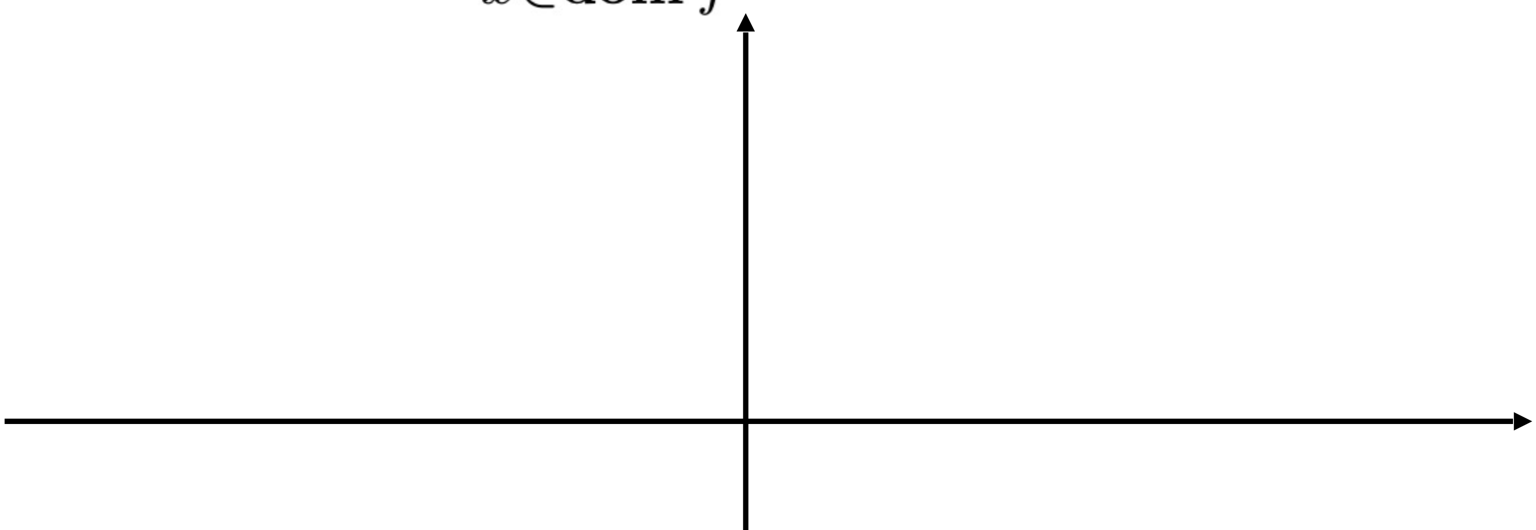
- 若f可微，则$f^*(y)$对应的x必满足$f'(x) = y$
- 函数f^*为凸函数
  - 即使函数f不是凸函数

$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} (y^T x - f(x))$$

共轭函数

- 若f可微，则$f^*(y)$对应的x必满足$f'(x) = y$
- 函数f^*为凸函数
  - 即使函数f不是凸函数

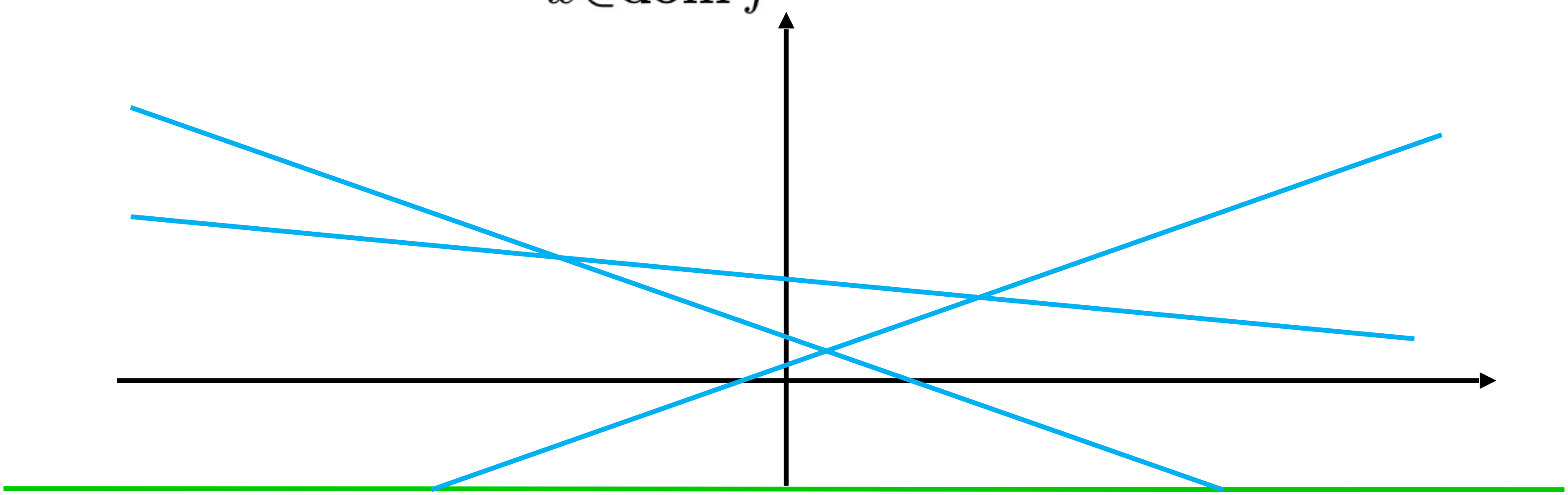
$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} \left(y^T x - f(x)\right)$$



共轭函数

- 若f可微，则$f^*(y)$对应的x必满足$f'(x)=y$
- 函数f^*为凸函数
  - 即使函数f不是凸函数

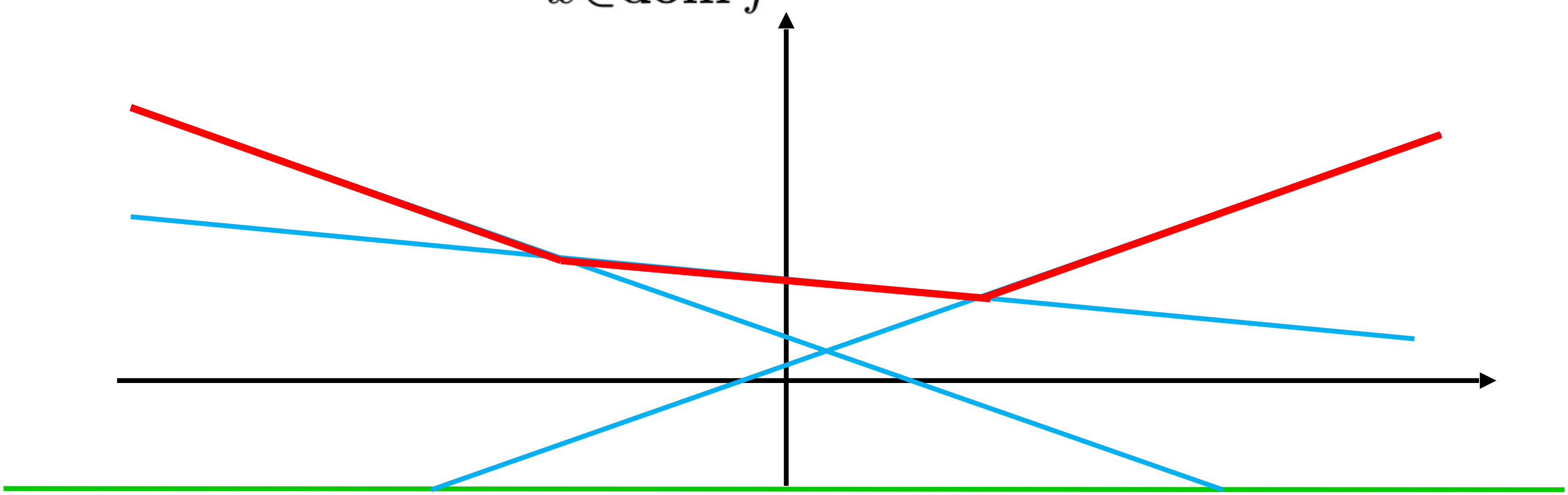
$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} \left(y^T x - f(x) \right)$$



共轭函数

- 若f可微，则$f^*(y)$对应的x必满足$f'(x) = y$
- 函数f^*为凸函数
  - 即使函数f不是凸函数

$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} \left(y^T x - f(x) \right)$$



共轭函数

共轭函数

- 例：$f(x) = ax + b$ 的共轭函数

共轭函数

- 例： $f(x) = ax + b$ 的共轭函数
  - $f^*(y) = \sup(yx - (ax + b))$

共轭函数

- 例：$f(x) = ax + b$的共轭函数
  - $$\begin{aligned} f^*(y) &= \sup(yx - (ax + b)) \\ &= \sup((y - a)x - b) \end{aligned}$$

共轭函数

- 例： $f(x) = ax + b$ 的共轭函数
  - $$f^*(y) = \sup(yx - (ax + b)) = \sup((y - a)x - b) = \begin{cases} -b & \textbf{if } y = a \\ +\infty & \textbf{if } y \neq a \end{cases}$$

共轭函数

- 例： $f(x) = ax + b$ 的共轭函数
  - $$f^*(y) = \sup(yx - (ax+b))$$
    $$= \sup((y-a)x - b)$$
    $$= \begin{cases} -b & \textbf{if } y = a \\ +\infty & \textbf{if } y \neq a \end{cases}$$
- 例： 负对数 $f(x) = -\log x$

共轭函数

- 例： $f(x) = ax + b$ 的共轭函数
  - $f^*(y) = \sup(yx - (ax+b))$

$$= \sup((y-a)x - b)$$

$$= \begin{cases} -b & \textbf{if } y = a \\ +\infty & \textbf{if } y \neq a \end{cases}$$

- 例： 负对数 $f(x) = -\log x$

$$f^*(y) = \sup_{x>0}(xy + \log x)$$

共轭函数

- 例： $f(x) = ax + b$ 的共轭函数
  - $$\begin{aligned} f^*(y) &= \sup(yx - (ax+b)) \\ &= \sup((y-a)x - b) \\ &= \begin{cases} -b & \textbf{if } y = a \\ +\infty & \textbf{if } y \neq a \end{cases} \end{aligned}$$
- 例： 负对数 $f(x) = -\log x$

$$f^*(y) = \sup_{x>0}(xy + \log x)$$

$$y + 1/x = 0$$

$$x = -1/y$$

共轭函数

- 例：$f(x) = ax + b$ 的共轭函数
  - $f^*(y) = \sup(yx - (ax+b))$
    $$= \sup((y-a)x - b)$$
    $$= \begin{cases} -b & \textbf{if } y = a \\ +\infty & \textbf{if } y \neq a \end{cases}$$

- 例：负对数 $f(x) = -\log x$

$$y + 1/x = 0$$
$$x = -1/y$$

$$\begin{aligned} f^*(y) &= \sup_{x>0}(xy + \log x) \\ &= \begin{cases} -1 - \log(-y) & y < 0 \\ \infty & \text{otherwise} \end{cases} \end{aligned}$$

共轭函数

共轭函数

- 二次函数：$f(x) = (1/2)x^T Q x$ with $Q \in \mathbf{S}^n_{++}$

共轭函数

- 二次函数：$f(x) = (1/2)x^TQx$ with $Q \in \mathbf{S}^n_{++}$

$$f^*(y) = \sup_x (y^Tx - (1/2)x^TQx)$$

共轭函数

- 二次函数：$f(x) = (1/2)x^T Qx$ with $Q \in \mathbf{S}^n_{++}$

$$f^*(y) = \sup_x (y^T x - (1/2)x^T Qx)$$

- $$\frac{\partial(y^T x - (1/2)x^T Qx)}{\partial x} = y - Qx = 0$$

共轭函数

- 二次函数：$f(x) = (1/2)x^TQx$ with $Q \in \mathbf{S}^n_{++}$

$$f^*(y) = \sup_x (y^Tx - (1/2)x^TQx)$$

- $\dfrac{\partial(y^Tx - (1/2)x^TQx)}{\partial x} = y - Qx = 0$
- $x = Q^{-1}y$

共轭函数

- 二次函数：$f(x) = (1/2)x^T Qx$ with $Q \in \mathbf{S}^n_{++}$

$$f^*(y) = \sup_x (y^T x - (1/2)x^T Qx)$$

- $\dfrac{\partial(y^T x - (1/2)x^T Qx)}{\partial x} = y - Qx = 0$
- $x = Q^{-1}y$
- $f^*(y) = y^T Q^{-1} y - (1/2) y^T (Q^{-1})^T Q Q^{-1} y$

共轭函数

- 二次函数：$f(x) = (1/2)x^T Qx$ with $Q \in \mathbf{S}^n_{++}$

$$f^*(y) = \sup_x (y^T x - (1/2)x^T Qx)$$

- $\dfrac{\partial(y^T x - (1/2)x^T Qx)}{\partial x} = y - Qx = 0$
- $x = Q^{-1}y$
- $f^*(y) = y^T Q^{-1} y - (1/2) y^T (Q^{-1})^T Q Q^{-1} y$

$$= \frac{1}{2} y^T Q^{-1} y$$

共轭函数的性质

共轭函数的性质

- 复数的共轭

共轭函数的性质

- 复数的共轭
  - $(a+bi)^* = a-bi$

共轭函数的性质

- 复数的共轭
  - $(a+bi)^* = a-bi$
  - $(a-bi)^* = a+bi$

共轭函数的性质

- 复数的共轭
  - $(a+bi)^* = a-bi$
  - $(a-bi)^* = a+bi$
- 函数的共轭

共轭函数的性质

- 复数的共轭
  - $(a+bi)^* = a-bi$
  - $(a-bi)^* = a+bi$
- 函数的共轭
  - $f^{**}=f$ ？

共轭函数的性质

- 复数的共轭
  - $(a+bi)^* = a-bi$
  - $(a-bi)^* = a+bi$
- 函数的共轭
  - $f^{**}=f$ ？
  - 若f为凸，且为闭函数，则$f^{**}=f$

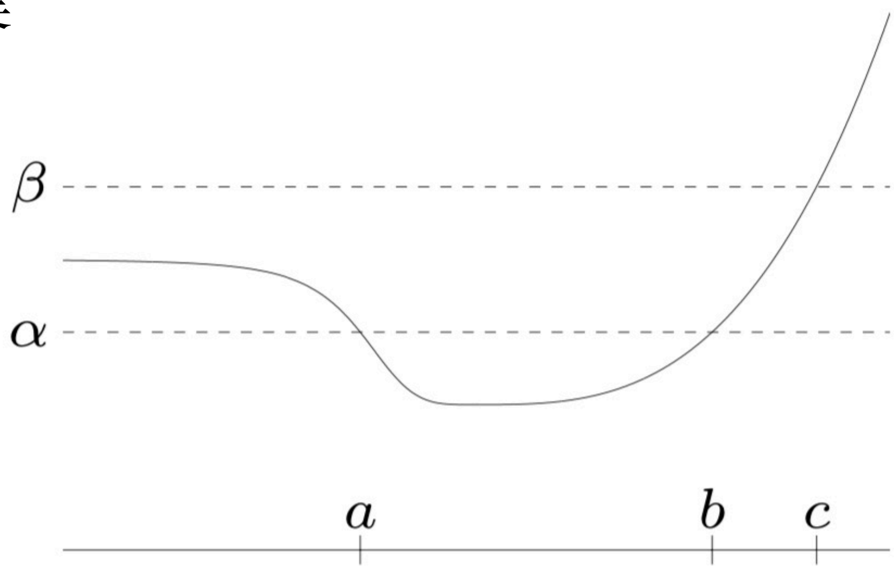
拟凸函数

拟凸函数

- 函数 $f : \mathbf{R}^n \to \mathbf{R}$ 为拟凸函数，当其定义域为凸集，且其下水平集 $S_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \le \alpha\}$ 对所有 α 均为凸集

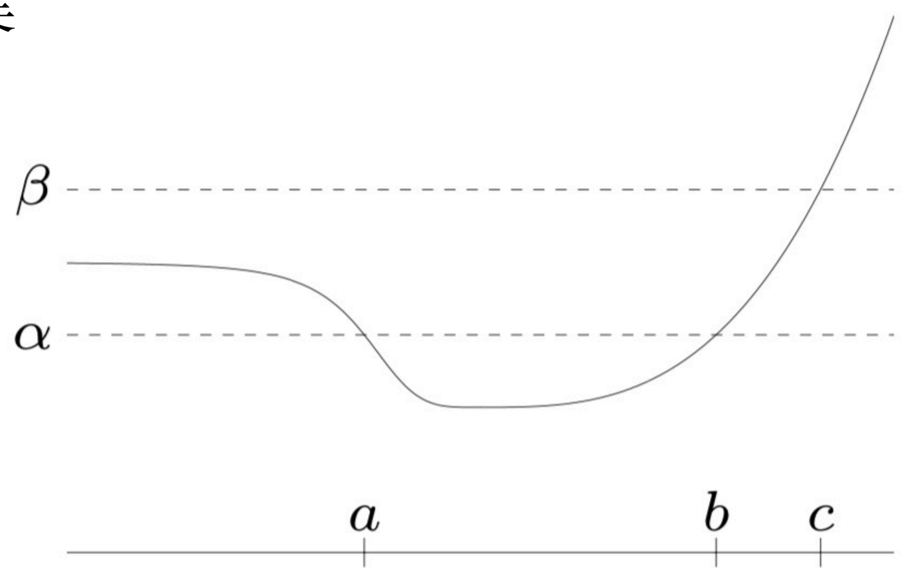
拟凸函数

- 函数 $f: \mathbf{R}^n \to \mathbf{R}$ 为拟凸函数，当其定义域为凸集，且其下水平集 $S_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \le \alpha\}$ 对所有 α 均为凸集



拟凸函数

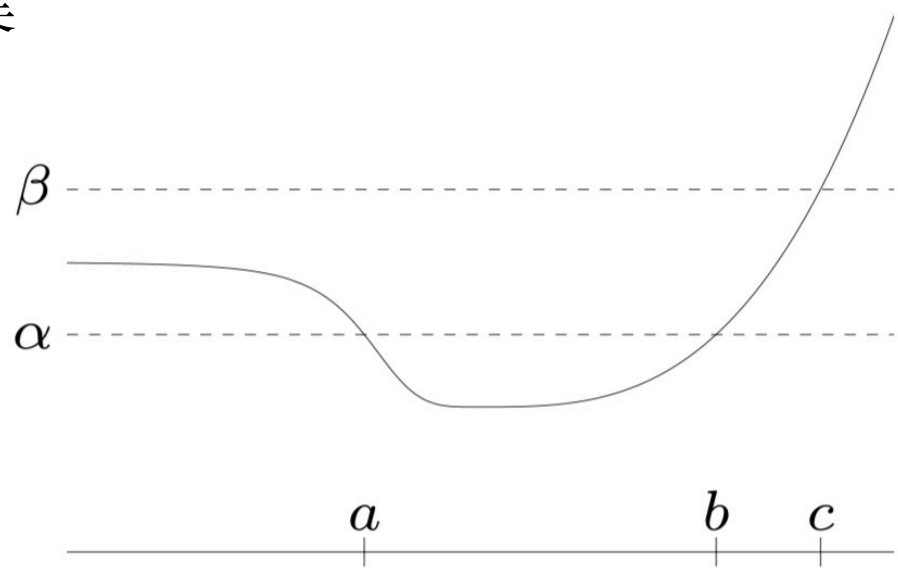
- 函数 $f : \mathbf{R}^n \to \mathbf{R}$ 为拟凸函数，当其定义域为凸集，且其下水平集 $S_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \le \alpha\}$ 对所有 α 均为凸集



- 函数f 为拟凹函数，若-f 为拟凸函数

拟凸函数

- 函数 $f:\mathbf{R}^n\to\mathbf{R}$ 为拟凸函数，当其定义域为凸集，且其下水平集 $S_\alpha=\{x\in\mathbf{dom}\,f\mid f(x)\le\alpha\}$ 对所有 α 均为凸集



- 函数f为拟凹函数，若$-f$为拟凸函数
- 函数f为拟线性函数，若同时为拟凹和拟凸函数

拟凸函数

拟凸函数

- 凸函数为拟凸函数

拟凸函数

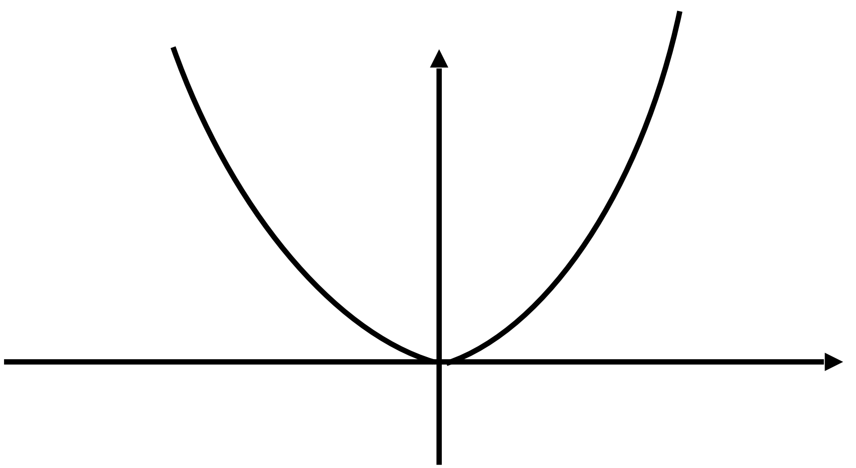
- 凸函数为拟凸函数
- 拟凸函数不一定为凸函数

拟凸函数

- 凸函数为拟凸函数
- 拟凸函数不一定为凸函数
- 拟凸函数为单模态函数

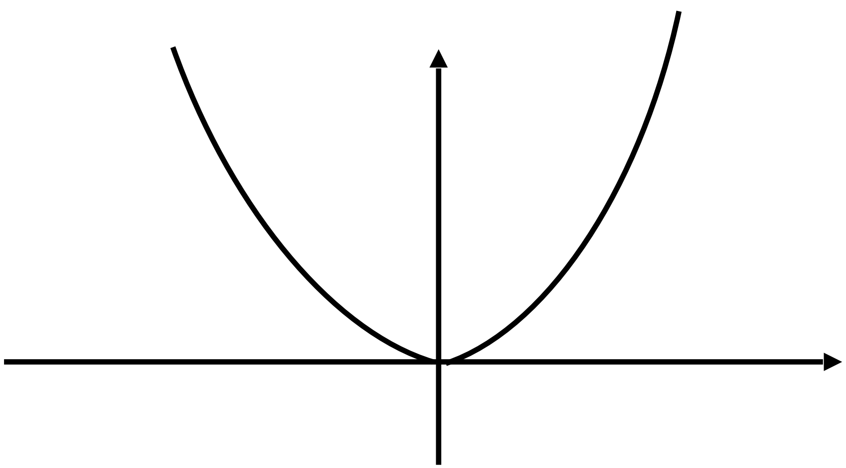
拟凸函数

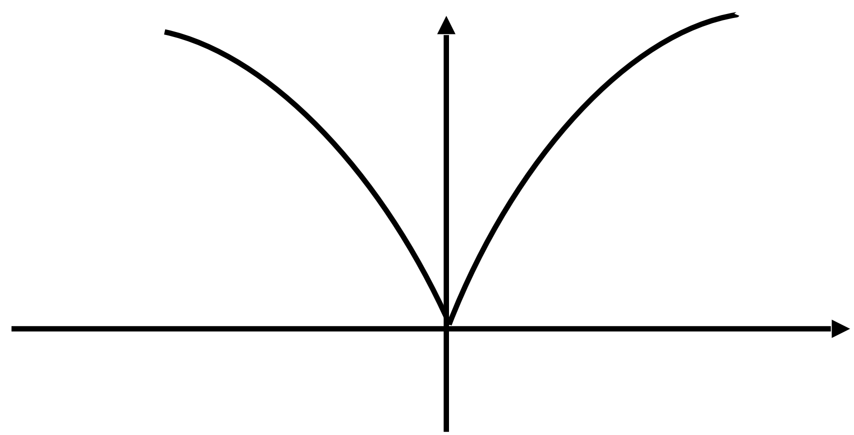
- 凸函数为拟凸函数
- 拟凸函数不一定为凸函数
- 拟凸函数为单模态函数



拟凸函数

- 凸函数为拟凸函数
- 拟凸函数不一定为凸函数
- 拟凸函数为单模态函数





拟凸函数

拟凸函数

- 凸函数：定义域为凸集，且

拟凸函数

- 凸函数：定义域为凸集，且
$$f(\theta x+(1-\theta)y)\le \theta f(x)+(1-\theta)f(y)$$

拟凸函数

- 凸函数：定义域为凸集，且

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

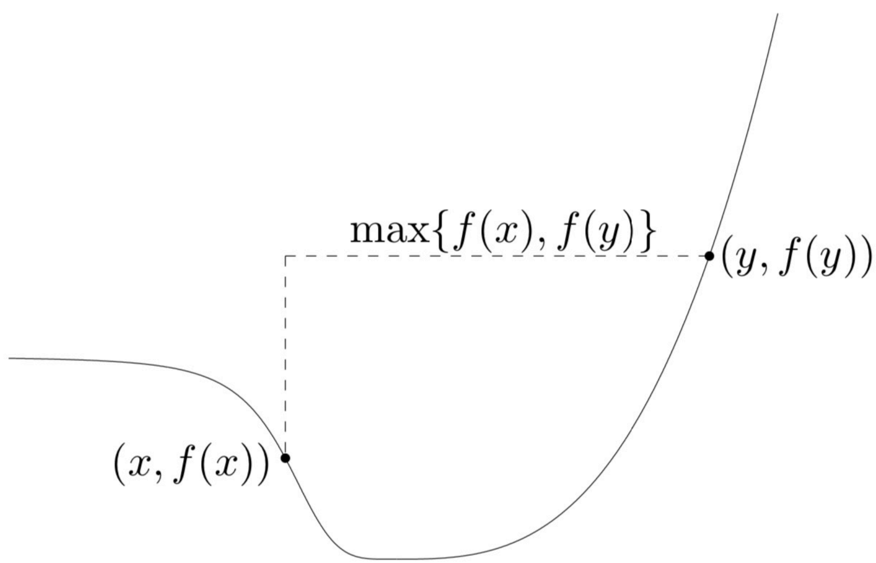
- 拟凸函数：定义域为凸集，且

拟凸函数

- 凸函数：定义域为凸集，且
  $$f(\theta x+(1-\theta)y)\le \theta f(x)+(1-\theta)f(y)$$
- 拟凸函数：定义域为凸集，且
  $$f(\theta x+(1-\theta)y)\le \max\{f(x),f(y)\}$$

拟凸函数

- 凸函数：定义域为凸集，且
  $$f(\theta x+(1-\theta)y)\le \theta f(x)+(1-\theta)f(y)$$
- 拟凸函数：定义域为凸集，且
  $$f(\theta x+(1-\theta)y)\le \max\{f(x),f(y)\}$$



拟凸函数

拟凸函数

- 向量的长度：向量中最后一个非零元素的位置

拟凸函数

- 向量的长度：向量中最后一个非零元素的位置

$$f(x) = \max\{i \mid x_i \neq 0\}$$

拟凸函数

- 向量的长度：向量中最后一个非零元素的位置

$$f(x) = \max\{i \mid x_i \neq 0\}$$

$$f(x) \leq \alpha \iff x_i = 0 \text{ for } i = \lfloor \alpha \rfloor + 1, \ldots, n$$

拟凸函数

- 向量的长度：向量中最后一个非零元素的位置

$$f(x) = \max\{i \mid x_i \neq 0\}$$

$$f(x) \leq \alpha \iff x_i = 0 \text{ for } i = \lfloor \alpha \rfloor + 1, \ldots, n$$

- 线性分数函数

拟凸函数

- 向量的长度：向量中最后一个非零元素的位置

$$f(x) = \max\{i \mid x_i \neq 0\}$$

$$f(x) \leq \alpha \iff x_i = 0 \text{ for } i = \lfloor \alpha \rfloor + 1, \ldots, n$$

- 线性分数函数

$$f(x) = \frac{a^T x + b}{c^T x + d}, \qquad \mathbf{dom}\, f = \{x \mid c^T x + d > 0\}$$

拟凸函数

- 向量的长度：向量中最后一个非零元素的位置

$$f(x) = \max\{i \mid x_i \neq 0\}$$

$$f(x) \leq \alpha \iff x_i = 0 \text{ for } i = \lfloor \alpha \rfloor + 1, \ldots, n$$

- 线性分数函数

$$f(x) = \frac{a^T x + b}{c^T x + d}, \qquad \textbf{dom}\, f = \{x \mid c^T x + d > 0\}$$

$$S_\alpha = \{x \mid c^T x + d > 0,\ (a^T x + b)/(c^T x + d) \leq \alpha\}$$

拟凸函数

- 向量的长度：向量中最后一个非零元素的位置

$$f(x) = \max\{i \mid x_i \neq 0\}$$

$$f(x) \leq \alpha \iff x_i = 0 \text{ for } i = \lfloor \alpha \rfloor + 1, \ldots, n$$

- 线性分数函数

$$f(x) = \frac{a^T x + b}{c^T x + d}, \qquad \mathbf{dom}\, f = \{x \mid c^T x + d > 0\}$$

$$\begin{aligned} S_\alpha &= \{x \mid c^T x + d > 0,\ (a^T x + b)/(c^T x + d) \leq \alpha\} \\ &= \{x \mid c^T x + d > 0,\ a^T x + b \leq \alpha(c^T x + d)\} \end{aligned}$$

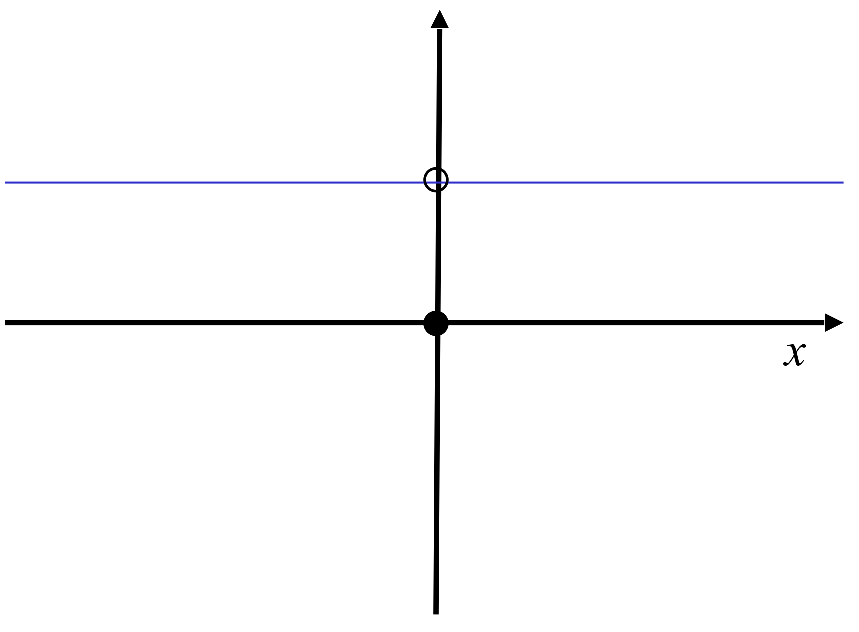
向量的零范数

向量的零范数

- 非零元素的个数
  - $f(x) = \|x\|_0$

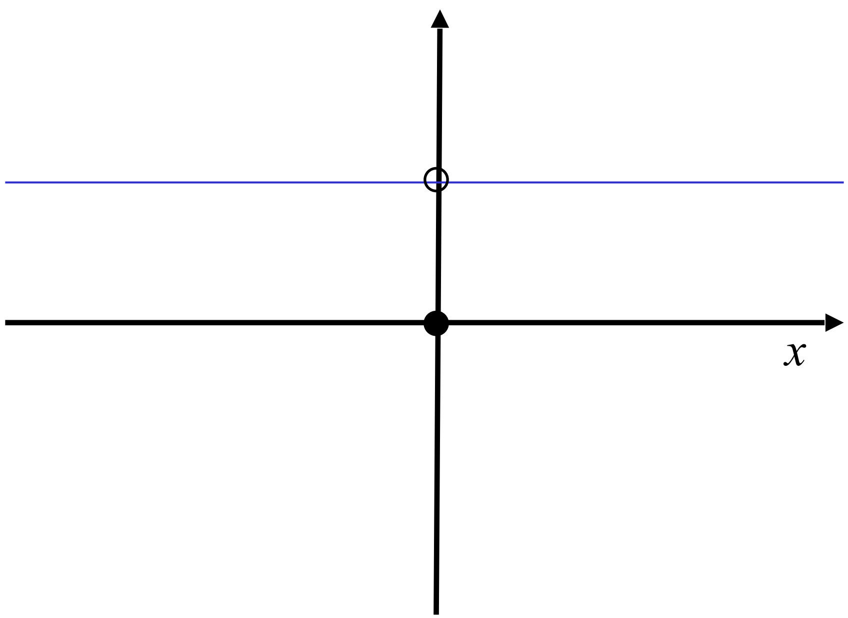
向量的零范数

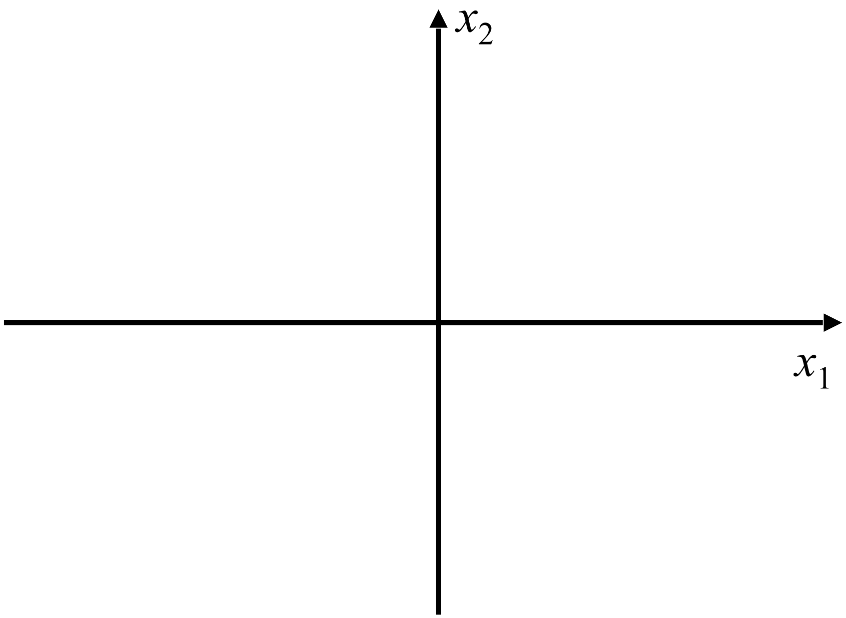
- 非零元素的个数
  - $f(x) = \|x\|_0$



向量的零范数

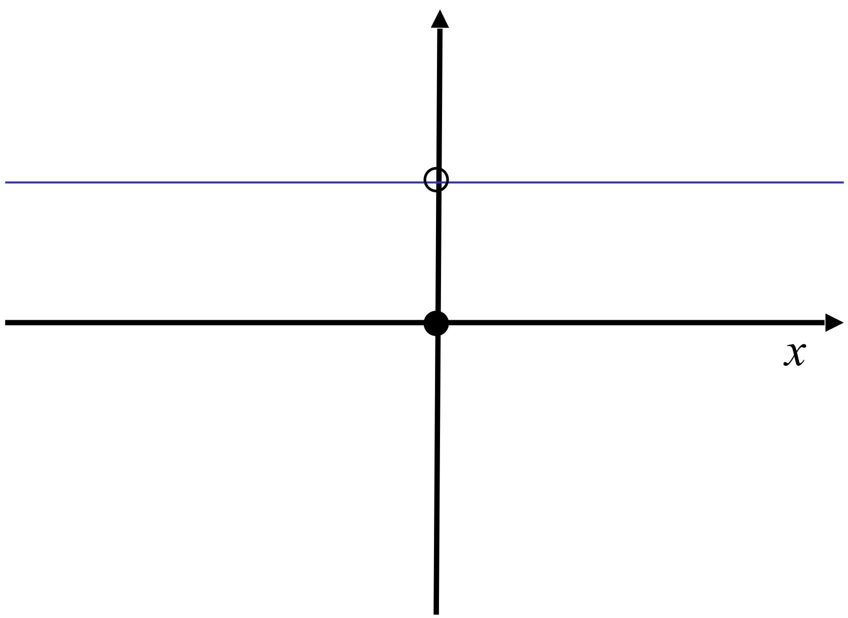
- 非零元素的个数
  - $f(x) = \|x\|_0$

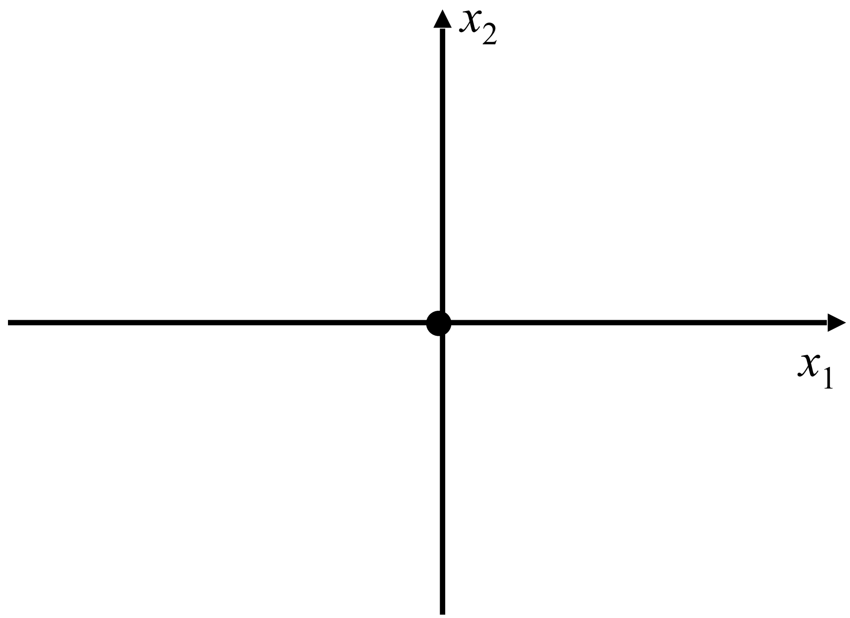




向量的零范数

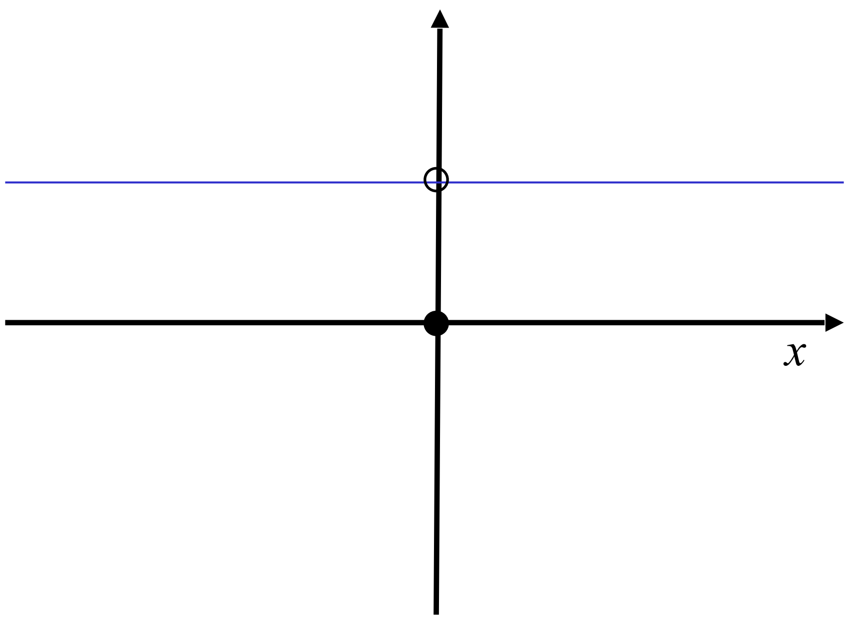
- 非零元素的个数
  - $f(x) = \|x\|_0$

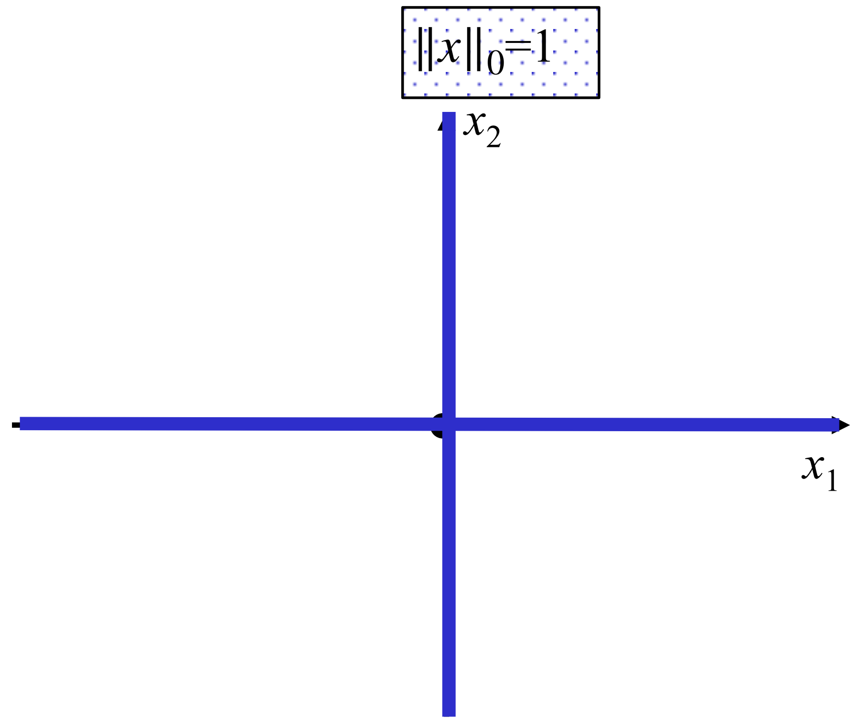




向量的零范数

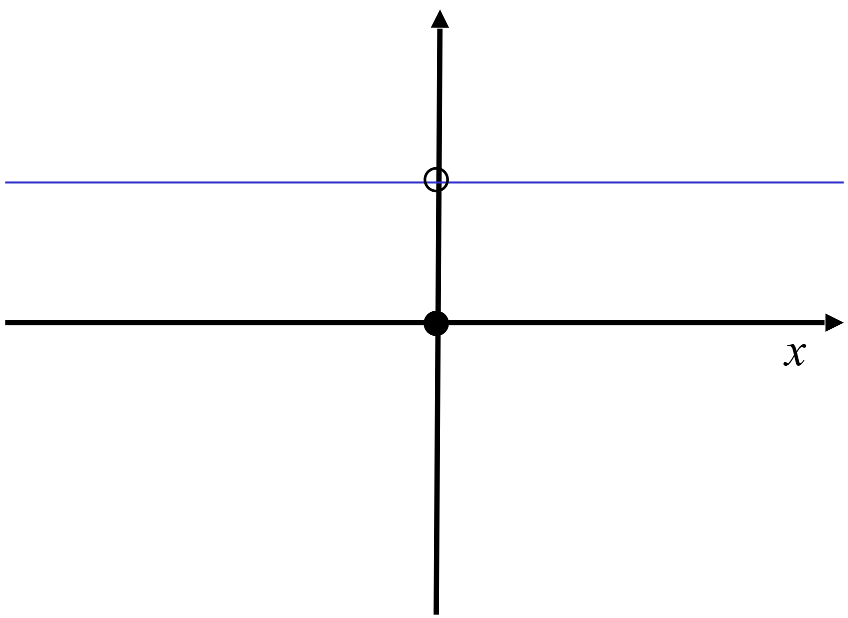
- 非零元素的个数
  - $f(x) = \|x\|_0$

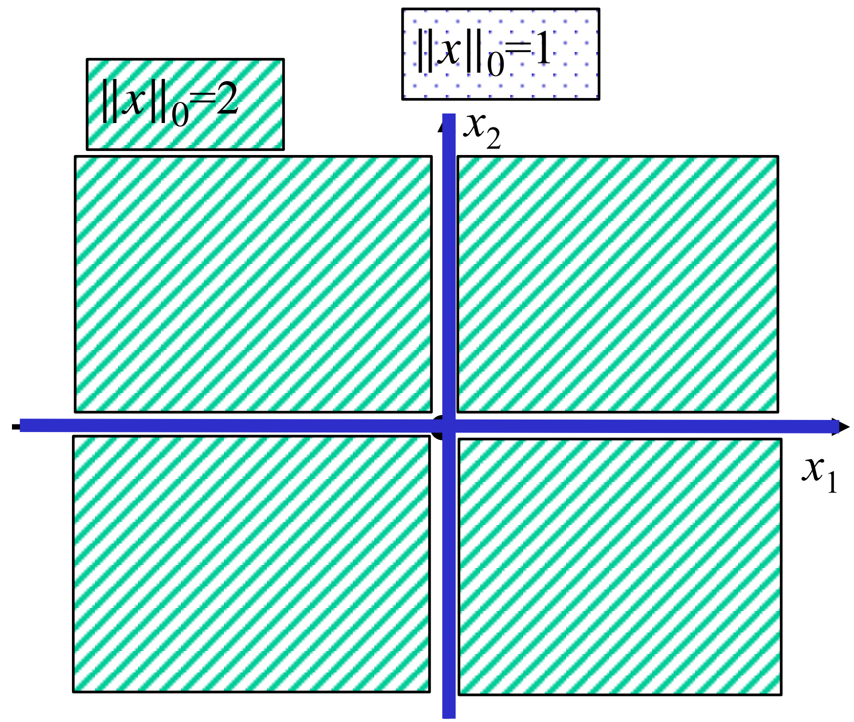




向量的零范数

- 非零元素的个数
  - $f(x) = \|x\|_0$





向量的零范数

向量的零范数

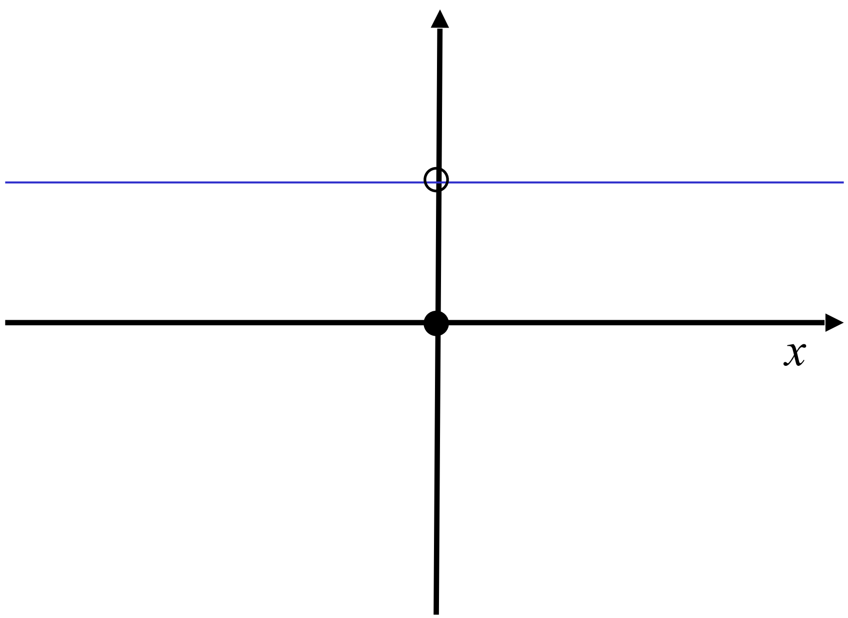
- $\min\limits_{x \in C} \|x\|_0$

向量的零范数

- $\min_{x \in C} \|x\|_0$
- $\min_{x \in C} \|x\|_1$

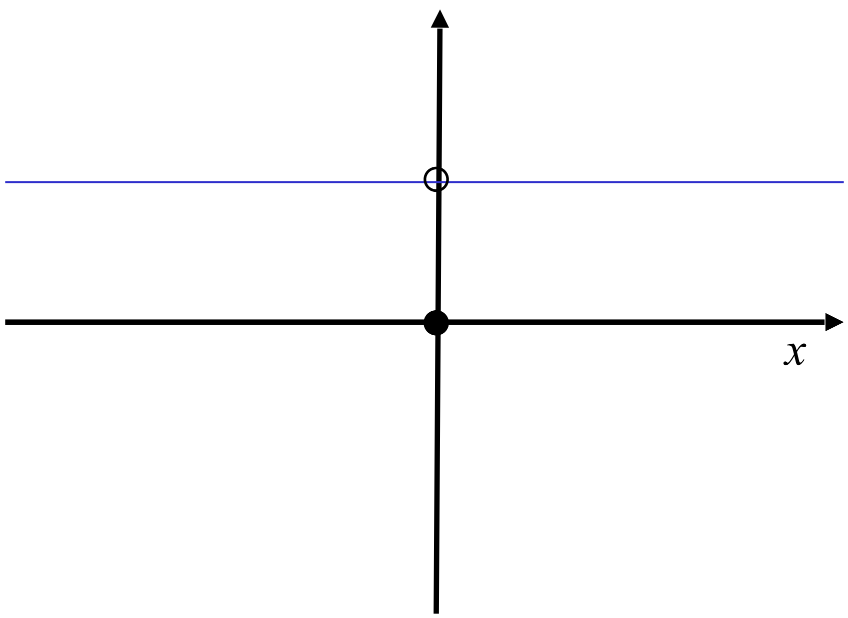
向量的零范数

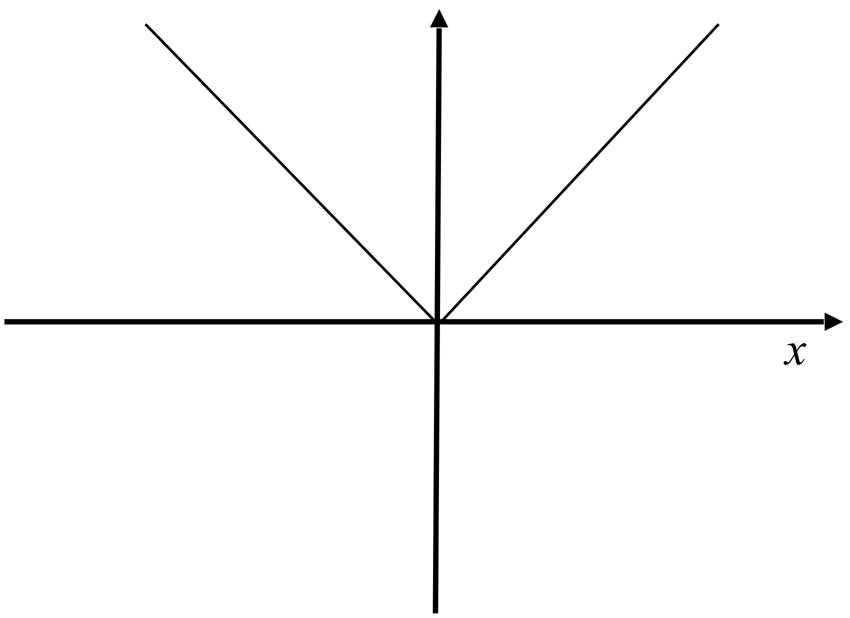
- $\min_{x \in C} \|x\|_0$
- $\min_{x \in C} \|x\|_1$



向量的零范数

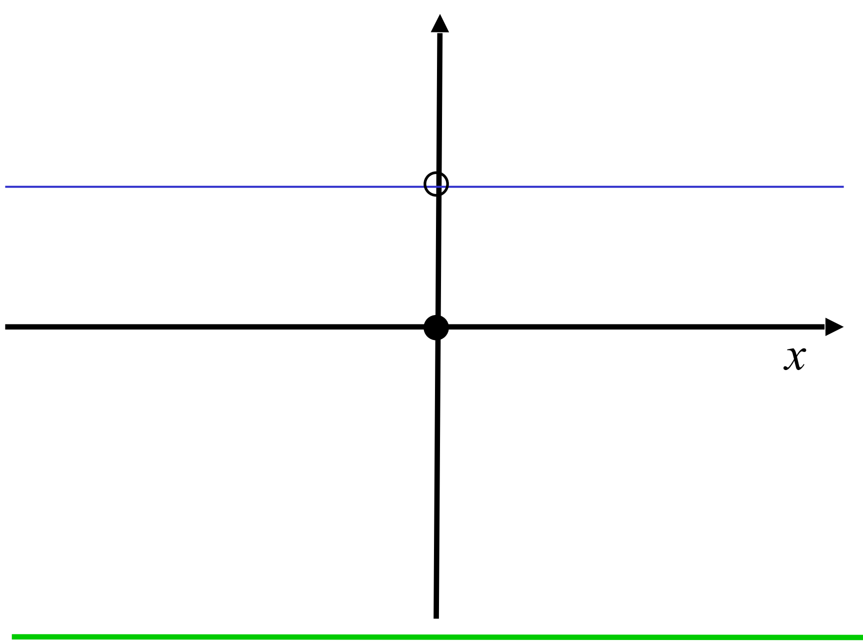
- $\min_{x \in C} \|x\|_0$
- $\min_{x \in C} \|x\|_1$





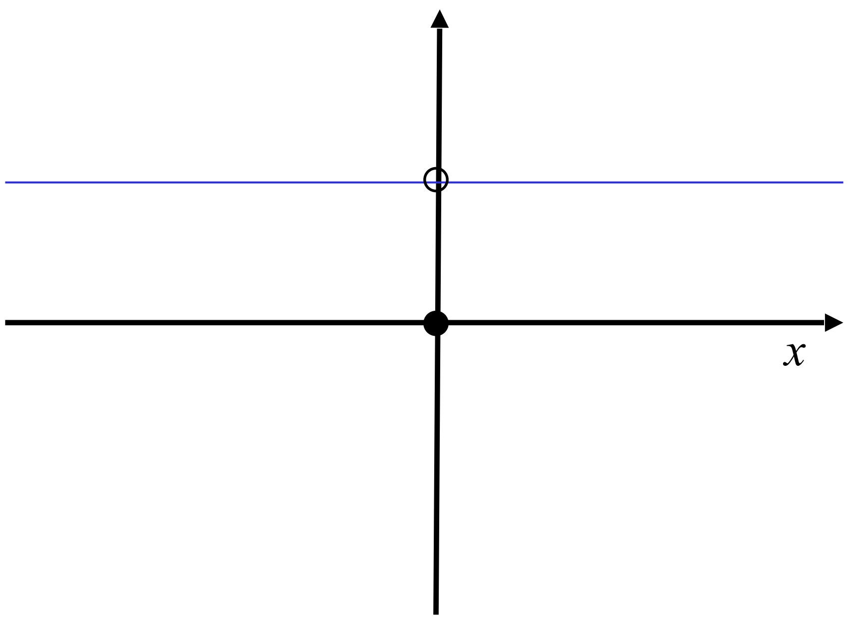
向量的零范数

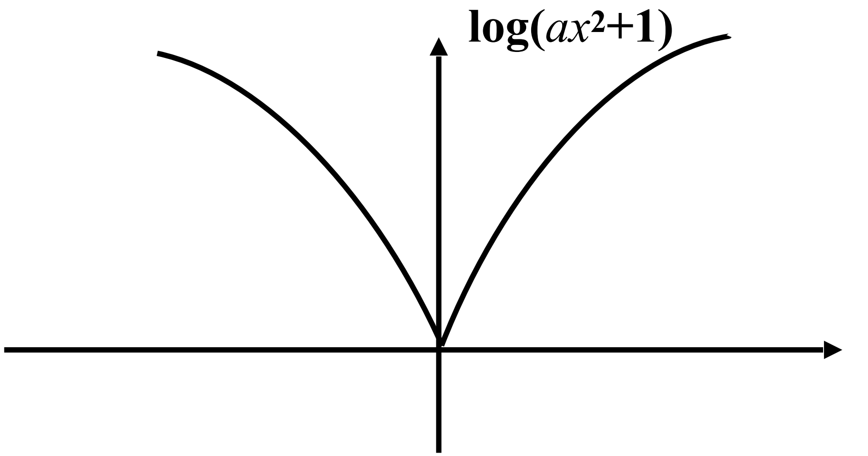
- $\min_{x \in C} \|x\|_0$
- $\min_{x \in C} \|x\|_1$



向量的零范数

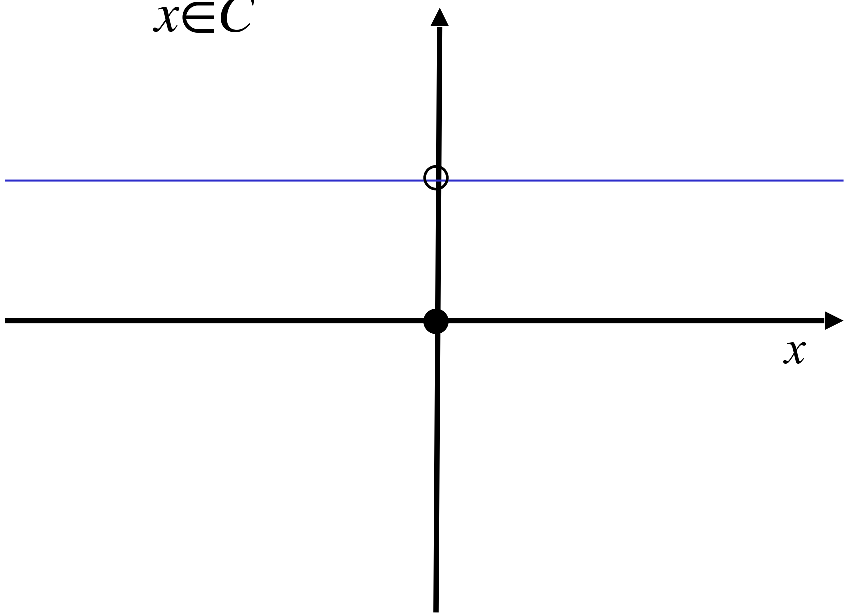
- $\min\limits_{x \in C} \|x\|_0$
- $\min\limits_{x \in C} \|x\|_1$

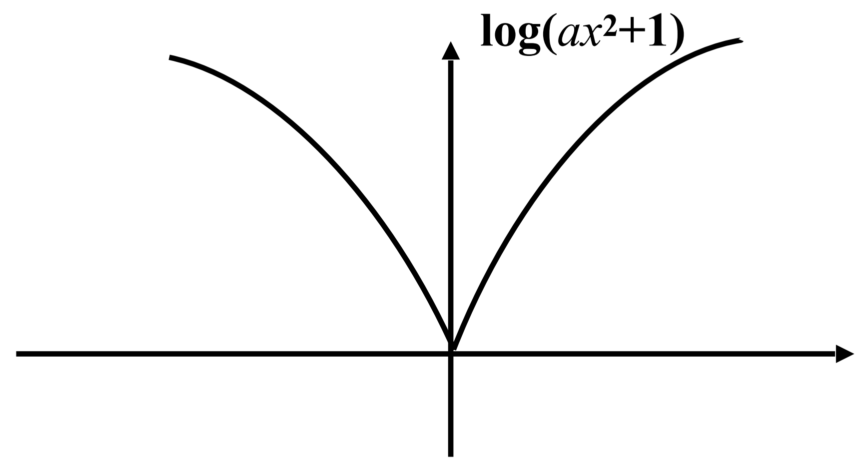




向量的零范数

- $\min\limits_{x \in C} \|x\|_0$
- $\min\limits_{x \in C} \|x\|_1$
- $\min\limits_{x \in C} \log(xx^T + 1)$





可微拟凸函数

可微拟凸函数

- 凸函数的一阶条件

可微拟凸函数

- 凸函数的一阶条件

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

可微拟凸函数

- 凸函数的一阶条件

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

- 拟凸函数的一阶条件

可微拟凸函数

- 凸函数的一阶条件

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

- 拟凸函数的一阶条件

$$f(y) \leq f(x) \Longrightarrow \nabla f(x)^T (y - x) \leq 0$$

可微拟凸函数

- 凸函数的一阶条件

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

- 拟凸函数的一阶条件

$$f(y) \leq f(x) \Longrightarrow \nabla f(x)^T (y - x) \leq 0$$

- 证明：仅考虑一维情况

可微拟凸函数

- 凸函数的一阶条件

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

- 拟凸函数的一阶条件

$$f(y) \leq f(x) \Longrightarrow \nabla f(x)^T (y - x) \leq 0$$

- 证明：仅考虑一维情况
  - $\Rightarrow \quad 0 \leq \theta < 1$

可微拟凸函数

- 凸函数的一阶条件

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

- 拟凸函数的一阶条件

$$f(y) \leq f(x) \Longrightarrow \nabla f(x)^T (y-x) \leq 0$$

- 证明：仅考虑一维情况
  - $\Rightarrow \quad 0 \leq \theta < 1$

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

可微拟凸函数

- 凸函数的一阶条件

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

- 拟凸函数的一阶条件

$$f(y) \leq f(x) \Longrightarrow \nabla f(x)^T (y - x) \leq 0$$

- 证明：仅考虑一维情况
  - $\Rightarrow \quad 0 \leq \theta < 1$

$$f(\theta x + (1 - \theta) y) \leq \max\{f(x), f(y)\} = f(x)$$

可微拟凸函数

- 凸函数的一阶条件

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

- 拟凸函数的一阶条件

$$f(y) \leq f(x) \Longrightarrow \nabla f(x)^T (y - x) \leq 0$$

- 证明：仅考虑一维情况
  - $\Rightarrow \quad 0 \leq \theta < 1$

$$f(\theta x + (1-\theta) y) \leq \max\{f(x), f(y)\} = f(x)$$

可微拟凸函数

- 凸函数的一阶条件

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

- 拟凸函数的一阶条件

$$f(y) \leq f(x) \Longrightarrow \nabla f(x)^T (y - x) \leq 0$$

- 证明：仅考虑一维情况
  - $\Rightarrow \quad 0 \leq \theta < 1$

$$f(\theta x + (1-\theta) y) \leq \max\{f(x), f(y)\} = f(x)$$

$$f(\theta x + (1-\theta) y) - f(\theta x + (1-\theta) x) \leq 0$$

可微拟凸函数

- 凸函数的一阶条件

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

- 拟凸函数的一阶条件

$$f(y) \leq f(x) \Longrightarrow \nabla f(x)^T (y - x) \leq 0$$

- 证明：仅考虑一维情况
  - $\Rightarrow \quad 0 \leq \theta < 1$

$$f(\theta x + (1-\theta) y) \leq \max\{f(x), f(y)\} = f(x)$$

$$f(\theta x + (1-\theta) y) - f(\theta x + (1-\theta) x) \leq 0$$

$$f'(x)(1-\theta)(y-x) + o[((1-\theta)(y-x))^2] \leq 0$$

可微拟凸函数

- 凸函数的一阶条件

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

- 拟凸函数的一阶条件

$$f(y) \leq f(x) \Longrightarrow \nabla f(x)^T (y - x) \leq 0$$

- 证明：仅考虑一维情况
  - $\Rightarrow \quad 0 \leq \theta < 1$

$$f(\theta x + (1-\theta)y) \leq \max\{f(x), f(y)\} = f(x)$$

$$f(\theta x + (1-\theta)y) - f(\theta x + (1-\theta)x) \leq 0$$

$$f'(x)(1-\theta)(y-x) + o[((1-\theta)(y-x))^2] \leq 0$$

$$f'(x)(y-x) \leq 0$$

可微拟凸函数

可微拟凸函数

- 证明$\Leftarrow$ $f(y) \le f(x) \Longrightarrow \nabla f(x)^T(y-x) \le 0$

可微拟凸函数

- 证明⇐ $f(y) \le f(x) \Longrightarrow \nabla f(x)^T (y-x) \le 0$

$$f(\theta x + (1-\theta)y) \le \max\{f(x), f(y)\}$$

可微拟凸函数

- 证明$\Leftarrow$ $\quad f(y) \leq f(x) \Longrightarrow \nabla f(x)^T(y-x) \leq 0$

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

- 设$x < y$

可微拟凸函数

- 证明⇐ $f(y) \le f(x) \Longrightarrow \nabla f(x)^T(y-x) \le 0$

$$f(\theta x + (1-\theta)y) \le \max\{f(x), f(y)\}$$

- 设$x < y$
- 反设存在$z \in [x, y]$，满足$f(z) > \max\{f(x), f(y)\}$

可微拟凸函数

- 证明⇐ $f(y) \le f(x) \Longrightarrow \nabla f(x)^T(y-x) \le 0$

$$f(\theta x+(1-\theta)y) \le \max\{f(x), f(y)\}$$

- 设$x < y$
- 反设存在$z \in [x, y]$，满足$f(z) > \max\{f(x), f(y)\}$
- 令$\bar{z} = \inf\{\tilde{z} \in [x, y] : f(\tilde{z}) = f(z)\}$

可微拟凸函数

- 证明$\Leftarrow$ $\quad f(y) \le f(x) \Longrightarrow \nabla f(x)^T(y-x) \le 0$

$$f(\theta x + (1-\theta)y) \le \max\{f(x), f(y)\}$$

- 设$x < y$
- 反设存在$z \in [x, y]$，满足$f(z) > \max\{f(x), f(y)\}$
- 令$\bar{z} = \inf\{\tilde{z} \in [x, y] : f(\tilde{z}) = f(z)\}$
- 存在$\epsilon > 0$

$$\max\{f(x), f(y)\} < f(\tilde{z}) < f(\bar{z}), \forall \tilde{z} \in (\bar{z} - \epsilon, \bar{z})$$

可微拟凸函数

- 证明$\Leftarrow$ $\quad f(y) \le f(x) \Longrightarrow \nabla f(x)^T(y-x) \le 0$
  $$f(\theta x + (1-\theta)y) \le \max\{f(x), f(y)\}$$
- 设$x < y$
- 反设存在$z \in [x, y]$，满足$f(z) > \max\{f(x), f(y)\}$
- 令$\bar{z} = \inf\{\tilde{z} \in [x, y] : f(\tilde{z}) = f(z)\}$
- 存在$\epsilon > 0$
  $\max\{f(x), f(y)\} < f(\tilde{z}) < f(\bar{z}), \forall \tilde{z} \in (\bar{z} - \epsilon, \bar{z})$
- 存在$z^* \in (\bar{z} - \epsilon, \bar{z}), f'(z^*) > 0$

可微拟凸函数

- 证明$\Leftarrow$ $f(y) \le f(x) \Longrightarrow \nabla f(x)^T(y-x) \le 0$

$$f(\theta x + (1-\theta)y) \le \max\{f(x), f(y)\}$$

- 设$x < y$
- 反设存在$z \in [x, y]$，满足$f(z) > \max\{f(x), f(y)\}$
- 令$\bar{z} = \inf\{\tilde{z} \in [x, y] : f(\tilde{z}) = f(z)\}$
- 存在$\epsilon > 0$

  $\max\{f(x), f(y)\} < f(\tilde{z}) < f(\bar{z}), \forall \tilde{z} \in (\bar{z} - \epsilon, \bar{z})$
- 存在$z^* \in (\bar{z} - \epsilon, \bar{z}), f'(z^*) > 0$
- $f(y) < f(z^*) \Longrightarrow f'(z^*)(y - z^*) > 0$

可微拟凸函数

- 证明$\Leftarrow$ $\quad f(y) \leq f(x) \Longrightarrow \nabla f(x)^T(y-x) \leq 0$

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

- 设$x < y$
- 反设存在$z \in [x, y]$，满足$f(z) > \max\{f(x), f(y)\}$
- 令$\bar{z} = \inf\{\tilde{z} \in [x, y] : f(\tilde{z}) = f(z)\}$
- 存在$\epsilon > 0$

  $\max\{f(x), f(y)\} < f(\tilde{z}) < f(\bar{z}), \forall \tilde{z} \in (\bar{z} - \epsilon, \bar{z})$
- 存在$z^* \in (\bar{z} - \epsilon, \bar{z}), f'(z^*) > 0$
- $f(y) < f(z^*) \Longrightarrow f'(z^*)(y - z^*) > 0$ ✗

拟凸函数

拟凸函数

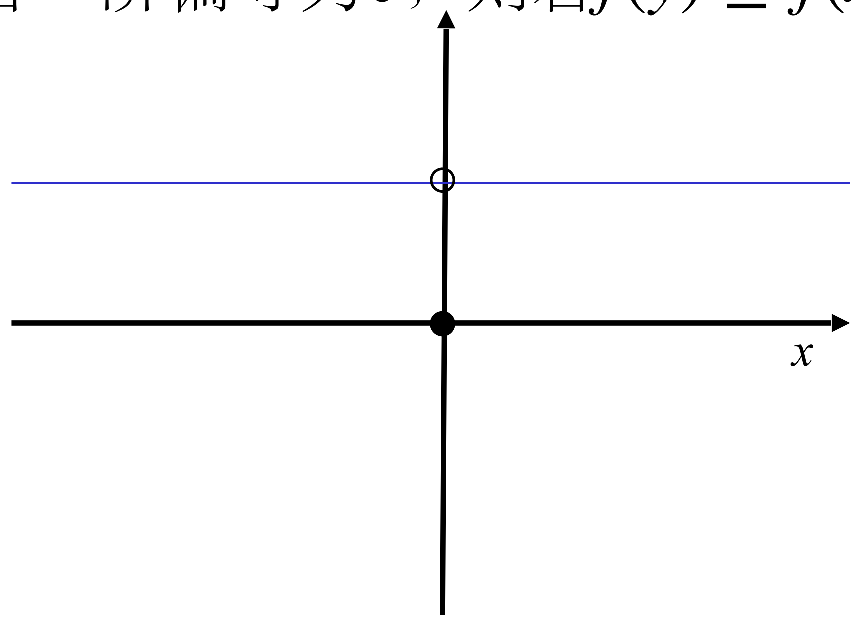
- 凸函数
  - 若一阶偏导为**0**，则$f(y) \geq f(x)$

拟凸函数

- 凸函数
  - 若一阶偏导为**0**，则$f(y) \geq f(x)$
- 拟凸函数
  - 若一阶偏导为**0**，则若$f(y) \leq f(x)$，则 $0 \leq 0$

拟凸函数

- 凸函数
  - 若一阶偏导为**0**，则$f(y) \geq f(x)$
- 拟凸函数
  - 若一阶偏导为**0**，则若$f(y) \leq f(x)$，则 $0 \leq 0$



拟凸函数的二阶条件

拟凸函数的二阶条件

- 凸函数的二阶条件

拟凸函数的二阶条件

- 凸函数的二阶条件

$$\nabla^2 f(x) \succeq 0 \quad \text{for all } x \in \mathbf{dom}\, f$$

拟凸函数的二阶条件

- 凸函数的二阶条件

$$\nabla^2 f(x) \succeq 0 \quad \text{for all } x \in \textbf{dom}\, f$$

- 拟凸函数的二阶条件

拟凸函数的二阶条件

- 凸函数的二阶条件

$$\nabla^2 f(x) \succeq 0 \quad \text{for all } x \in \mathbf{dom}\, f$$

- 拟凸函数的二阶条件

$$y^T \nabla f(x) = 0 \Longrightarrow y^T \nabla^2 f(x) y \geq 0$$

拟凸函数的二阶条件

- 凸函数的二阶条件

$$\nabla^2 f(x) \succeq 0 \quad \text{for all } x \in \mathbf{dom}\, f$$

- 拟凸函数的二阶条件

$$y^T \nabla f(x) = 0 \Longrightarrow y^T \nabla^2 f(x) y \geq 0$$

- 一维函数下

拟凸函数的二阶条件

- 凸函数的二阶条件

$$\nabla^2 f(x) \succeq 0 \quad \text{for all } x \in \mathbf{dom}\, f$$

- 拟凸函数的二阶条件

$$y^T \nabla f(x) = 0 \Longrightarrow y^T \nabla^2 f(x) y \geq 0$$

- 一维函数下
  - $yf'(x) = 0 \rightarrow y^2 f''(x) \geq 0$

拟凸函数的二阶条件

- 凸函数的二阶条件

$$\nabla^2 f(x) \succeq 0 \quad \text{for all } x \in \mathbf{dom}\, f$$

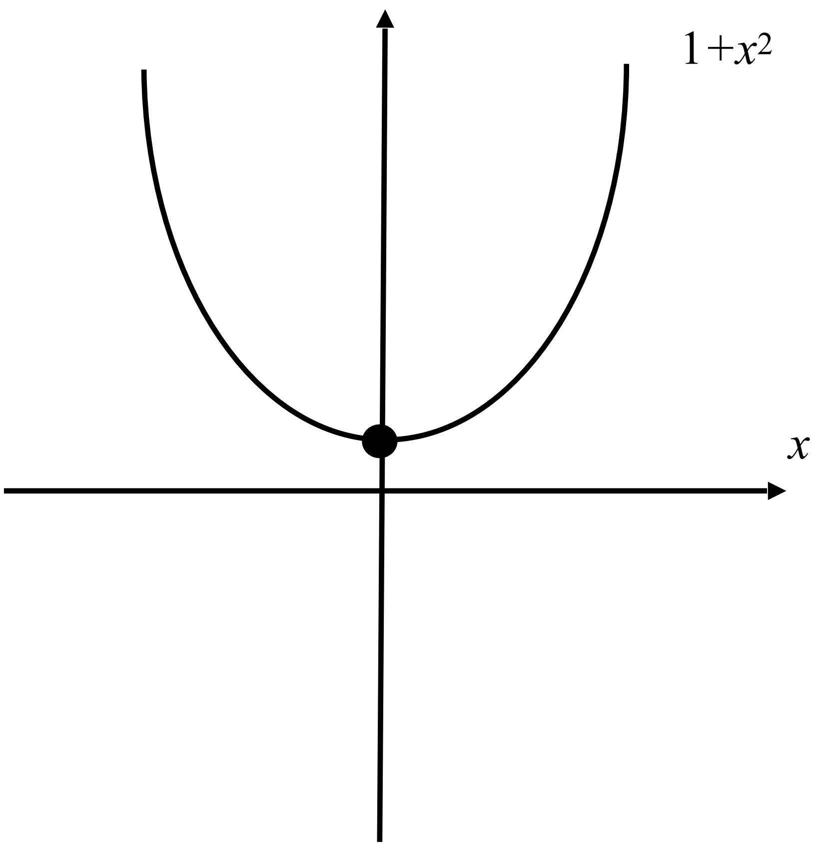
- 拟凸函数的二阶条件

$$y^T \nabla f(x) = 0 \Longrightarrow y^T \nabla^2 f(x) y \geq 0$$

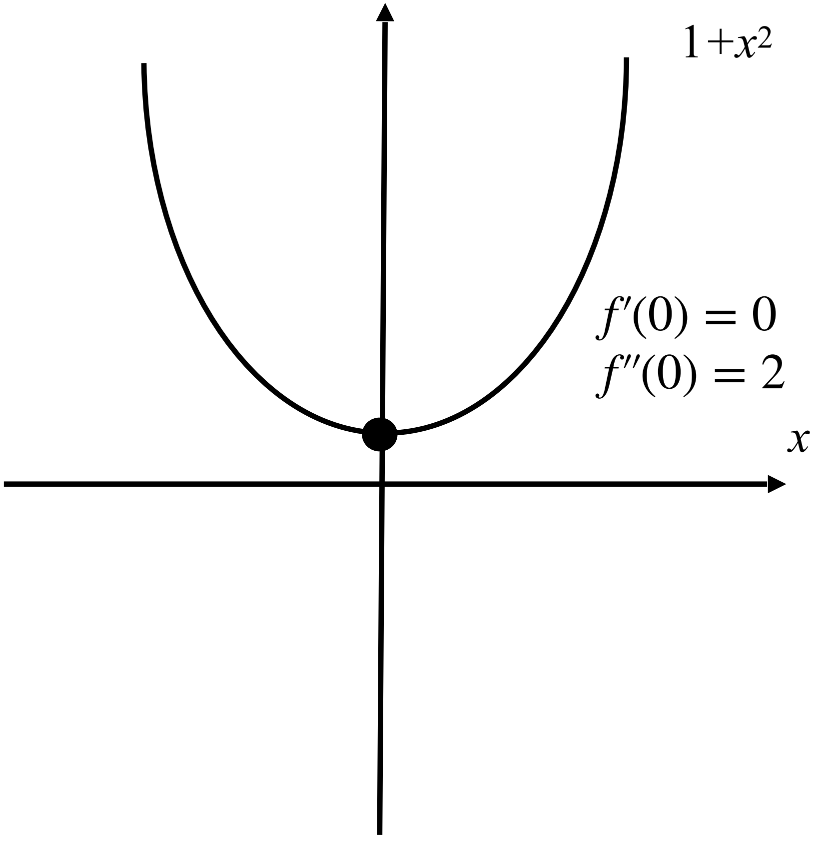
- 一维函数下
  - $yf'(x) = 0 \rightarrow y^2 f''(x) \geq 0$
  - $\begin{cases} y = 0 & 0 \geq 0 \\ f'(x) = 0, y \neq 0 & f''(x) \geq 0 \end{cases}$

拟凸函数的二阶条件

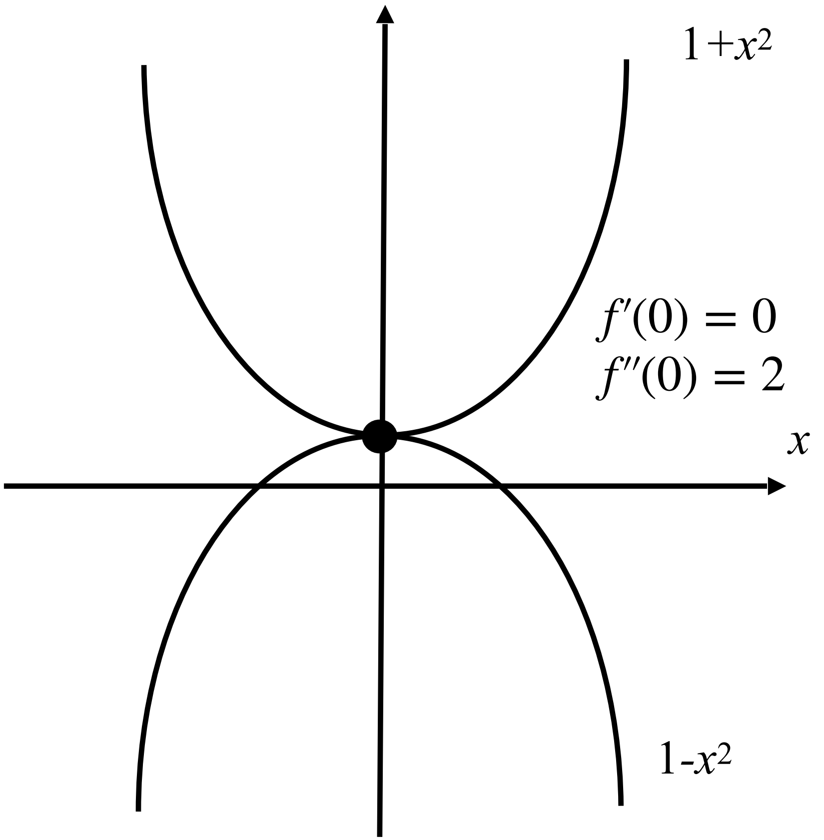
拟凸函数的二阶条件



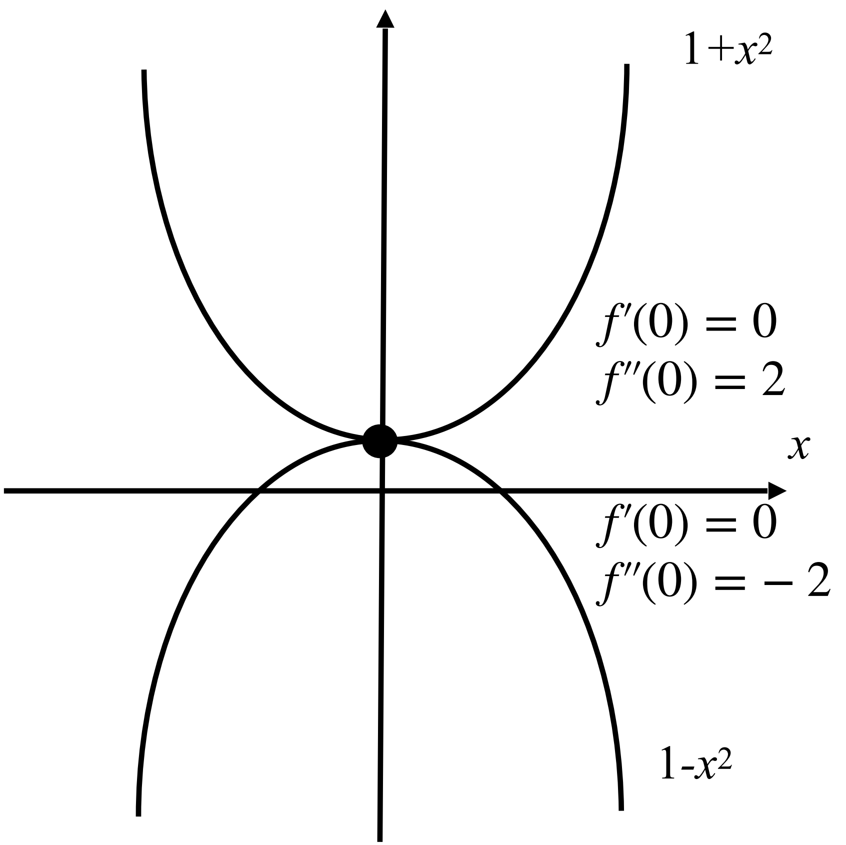
拟凸函数的二阶条件



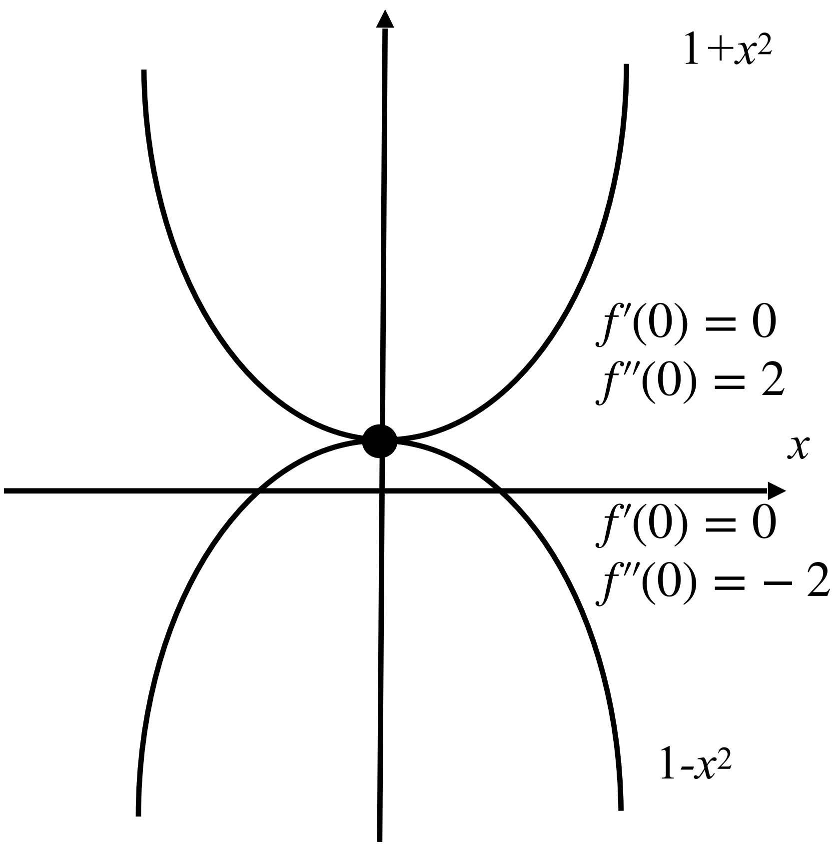
拟凸函数的二阶条件

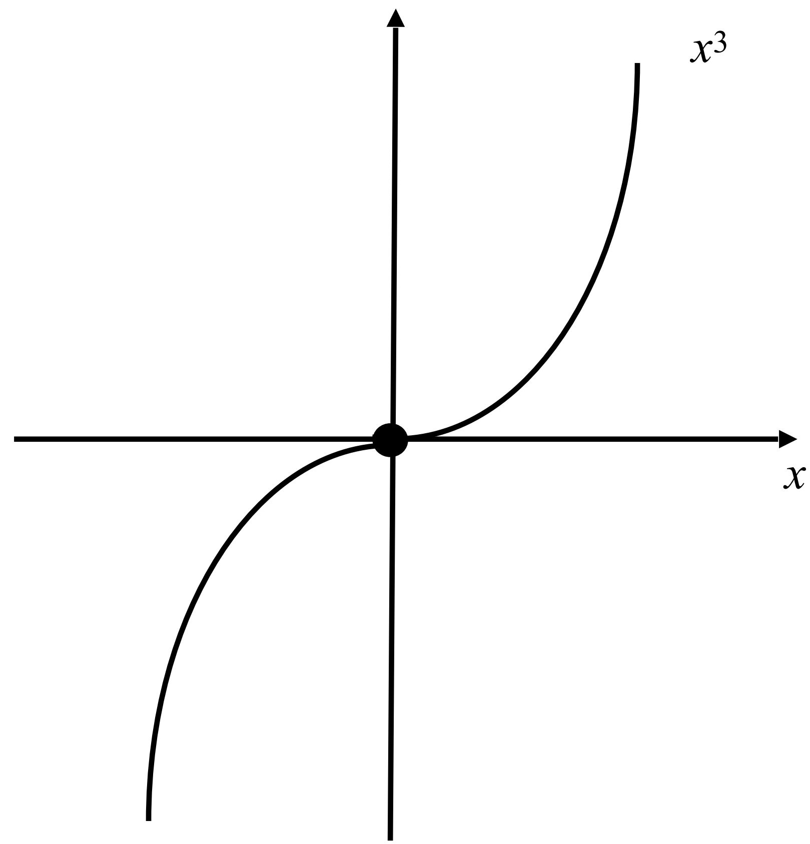


拟凸函数的二阶条件

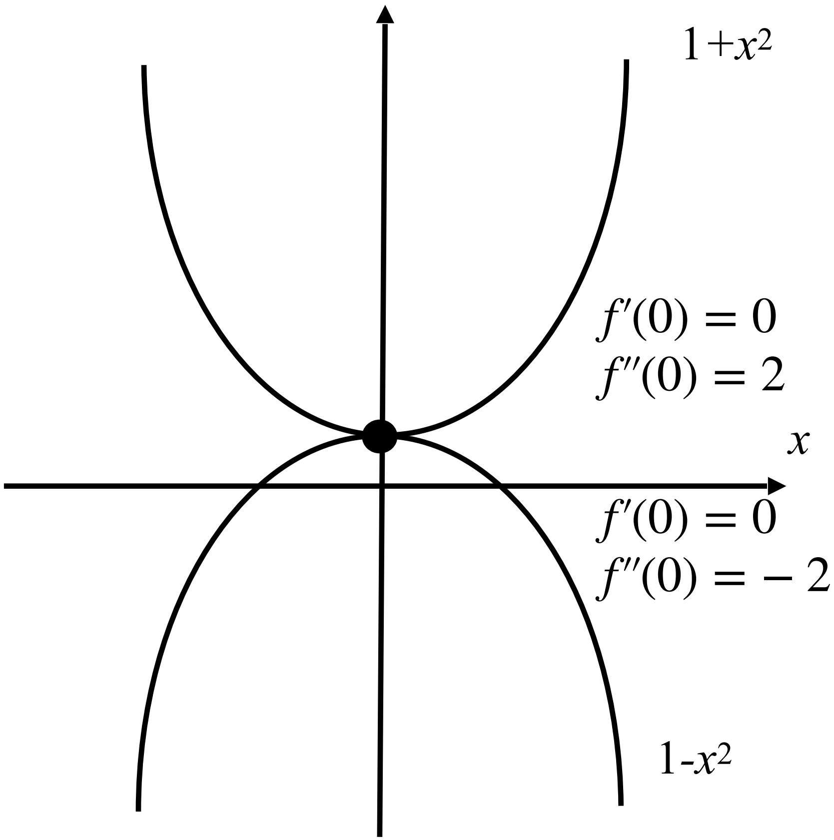


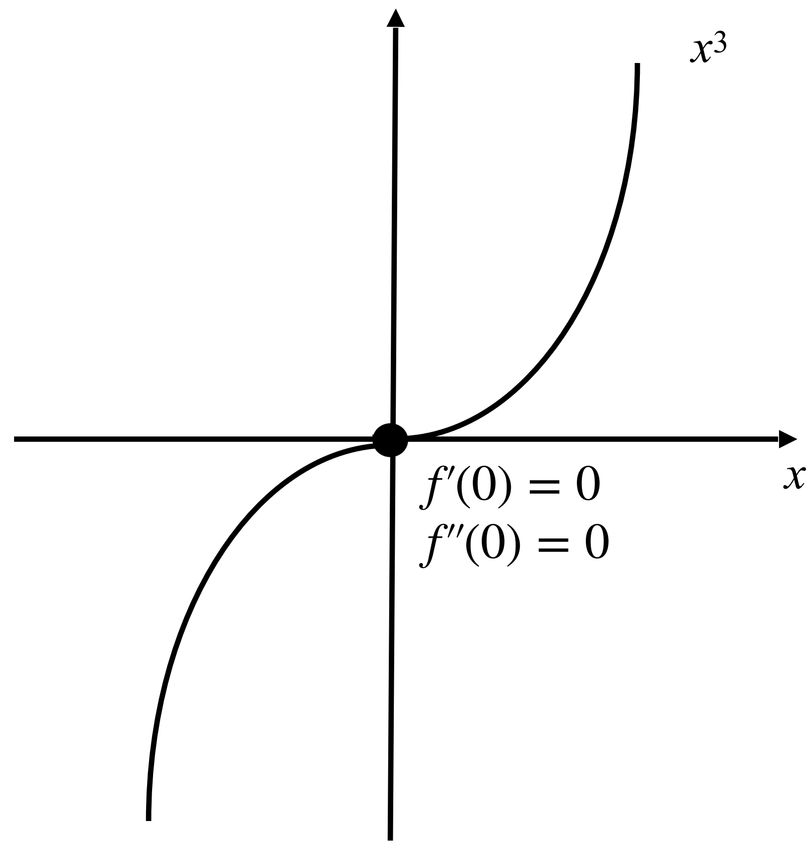
拟凸函数的二阶条件





拟凸函数的二阶条件





对数-凸函数和对数-凹函数

对数-凸函数和对数-凹函数

- 正值函数f为对数-凹函数，需满足$\log f$为凹函数：

对数-凸函数和对数-凹函数

- 正值函数f为对数-凹函数，需满足$\log f$为凹函数：

$$f(\theta x + (1-\theta) y) \geq f(x)^\theta f(y)^{1-\theta} \quad \text{for } 0 \leq \theta \leq 1$$

对数-凸函数和对数-凹函数

- 正值函数f为对数-凹函数，需满足$\log f$为凹函数：

$$f(\theta x+(1-\theta)y) \geq f(x)^{\theta} f(y)^{1-\theta} \quad \text{for } 0 \leq \theta \leq 1$$

- 正值函数f为对数-凸函数，需满足$\log f$为凸函数

对数-凸函数和对数-凹函数

- 正值函数f 为对数-凹函数，需满足$\log f$ 为凹函数：

$$f(\theta x + (1-\theta)y) \geq f(x)^{\theta} f(y)^{1-\theta} \quad \text{for } 0 \leq \theta \leq 1$$

- 正值函数f 为对数-凸函数，需满足$\log f$ 为凸函数
- 若函数f为对数-凸函数，则函数f为凸函数

对数-凸函数和对数-凹函数

- 正值函数f为对数-凹函数，需满足$\log f$为凹函数：

$$f(\theta x + (1-\theta)y) \geq f(x)^{\theta} f(y)^{1-\theta} \quad \text{for } 0 \leq \theta \leq 1$$

- 正值函数f为对数-凸函数，需满足$\log f$为凸函数
- 若函数f为对数-凸函数，则函数f为凸函数
  - $f = e^{\log(f)}$

对数-凸函数和对数-凹函数

- 正值函数f为对数-凹函数，需满足$\log f$为凹函数：

$$f(\theta x + (1-\theta)y) \geq f(x)^{\theta} f(y)^{1-\theta} \quad \text{for } 0 \leq \theta \leq 1$$

- 正值函数f为对数-凸函数，需满足$\log f$为凸函数
- 若函数f为对数-凸函数，则函数f为凸函数
  - $f = e^{\log(f)}$
- 若函数f为凹函数，则函数f为对数凹函数

对数-凸函数和对数-凹函数

- 正值函数f为对数-凹函数，需满足$\log f$为凹函数：
  $$f(\theta x + (1-\theta)y) \geq f(x)^{\theta} f(y)^{1-\theta} \quad \text{for } 0 \leq \theta \leq 1$$
- 正值函数f为对数-凸函数，需满足$\log f$为凸函数
- 若函数f为对数-凸函数，则函数f为凸函数
  - $f = e^{\log(f)}$
- 若函数f为凹函数，则函数f为对数凹函数
  - $\log(f)$

对数凸函数

对数凸函数

- 幂函数: x^a 为若 $a \le 0$ 为正实数集上的对数-凸函数，若 $a \ge 0$ 为对数-凹函数

对数凸函数

- 幂函数: x^a 为若 $a \le 0$ 为正实数集上的对数-凸函数，若 $a \ge 0$ 为对数-凹函数
- 许多通用的概率密度函数为对数-凹函数，如正态分布:

对数凸函数

- 幂函数: x^a 为若 $a \le 0$ 为正实数集上的对数-凸函数，若 $a \ge 0$ 为对数-凹函数
- 许多通用的概率密度函数为对数-凹函数，如正态分布:

$$f(x) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} e^{-\frac{1}{2}(x-\bar{x})^T \Sigma^{-1}(x-\bar{x})}$$

对数凸函数

- 幂函数: x^a为若 $a \le 0$ 为正实数集上的对数-凸函数，若 $a \ge 0$为对数-凹函数
- 许多通用的概率密度函数为对数-凹函数，如正态分布：

$$f(x) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} e^{-\frac{1}{2}(x-\bar{x})^T \Sigma^{-1} (x-\bar{x})}$$

- 累积高斯分布函数为对数-凹函数

对数凸函数

- 幂函数：x^a 为若 $a \le 0$ 为正实数集上的对数-凸函数，若 $a \ge 0$ 为对数-凹函数
- 许多通用的概率密度函数为对数-凹函数，如正态分布：

$$f(x) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} e^{-\frac{1}{2}(x-\bar{x})^T \Sigma^{-1} (x-\bar{x})}$$

- 累积高斯分布函数为对数-凹函数

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2}\, du$$

对数-凹函数的性质

对数-凹函数的性质

- 定义域为凸集的二阶可微函数f为对数-凹函数当且仅当对所有 x

对数-凹函数的性质

- 定义域为凸集的二阶可微函数f为对数-凹函数当且仅当对所有 x

$$f(x)\nabla^2 f(x) \preceq \nabla f(x)\nabla f(x)^T$$

对数-凹函数的性质

- 定义域为凸集的二阶可微函数f为对数-凹函数当且仅当对所有 x

$$f(x)\nabla^2 f(x) \preceq \nabla f(x)\nabla f(x)^T$$

- 对数-凹函数的乘积为对数-凹函数

对数-凹函数的性质

- 定义域为凸集的二阶可微函数f为对数-凹函数当且仅当对所有 x

$$f(x)\nabla^2 f(x) \preceq \nabla f(x)\nabla f(x)^T$$

- 对数-凹函数的乘积为对数-凹函数
- 对数-凹函数的和不总是为对数-凹函数

对数-凹函数的性质

- 定义域为凸集的二阶可微函数f为对数-凹函数当且仅当对所有 x

$$f(x)\nabla^2 f(x) \preceq \nabla f(x)\nabla f(x)^T$$

- 对数-凹函数的乘积为对数-凹函数
- 对数-凹函数的和不总是为对数-凹函数
- 积分：若函数 $f : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}$ 为对数-凹函数，则

$$g(x) = \int f(x, y)\, dy$$

对数-凹函数的性质

- 定义域为凸集的二阶可微函数f为对数-凹函数当且仅当对所有 x

$$f(x)\nabla^2 f(x) \preceq \nabla f(x)\nabla f(x)^T$$

- 对数-凹函数的乘积为对数-凹函数
- 对数-凹函数的和不总是为对数-凹函数
- 积分：若函数 $f : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}$ 为对数-凹函数，则

$$g(x) = \int f(x, y)\, dy \text{ 为对数-凹函数}$$

关于广义不等式的凸性

关于广义不等式的凸性

- 函数 $f : \mathbf{R}^n \to \mathbf{R}^m$为$K$-凸函数，则定义域为凸集，且

$$f(\theta x + (1-\theta)y) \preceq_K \theta f(x) + (1-\theta) f(y)$$

  - 此处：$x,\ y \in \mathbf{dom}\, f,\ 0 \le \theta \le 1$

关于广义不等式的凸性

- 函数$f:\mathbf{R}^n\to\mathbf{R}^m$为$K$-凸函数，则定义域为凸集，且

$$f(\theta x+(1-\theta)y)\preceq_K \theta f(x)+(1-\theta)f(y)$$

  - 此处：$x,\ y\in\mathbf{dom}\,f,\ 0\le\theta\le 1$

- 例：$f:\mathbf{S}^m\to\mathbf{S}^m,\ f(X)=X^2$ 为$\mathbf{S}_+^m$-凸函数

关于广义不等式的凸性

- 函数 $f:\mathbf{R}^n \to \mathbf{R}^m$为$K$-凸函数，则定义域为凸集，且

$$f(\theta x+(1-\theta)y) \preceq_K \theta f(x)+(1-\theta)f(y)$$

  - 此处：$x,\, y \in \mathbf{dom}\, f,\, 0 \le \theta \le 1$
- 例：$f:\mathbf{S}^m \to \mathbf{S}^m,\, f(X)=X^2$ 为$\mathbf{S}^m_+$-凸函数
- 证明：给定 $z \in \mathbf{R}^m,\, z^T X^2 z = \|Xz\|_2^2$ 为关于X的凸函数

关于广义不等式的凸性

- 函数 $f:\mathbf{R}^n\to\mathbf{R}^m$为$K$-凸函数，则定义域为凸集，且

$$f(\theta x+(1-\theta)y)\preceq_K \theta f(x)+(1-\theta)f(y)$$

  - 此处：$x,\, y\in\mathbf{dom}\, f,\ 0\le\theta\le 1$

- 例：$f:\mathbf{S}^m\to\mathbf{S}^m,\ f(X)=X^2$ 为$\mathbf{S}_+^m$-凸函数
- 证明：给定 $z\in\mathbf{R}^m,\ z^TX^2z=\|Xz\|_2^2$为关于$X$的凸函数

$$z^T(\theta X+(1-\theta)Y)^2z\le\theta z^TX^2z+(1-\theta)z^TY^2z$$

关于广义不等式的凸性

- 函数 $f:\mathbf{R}^n \to \mathbf{R}^m$ 为K-凸函数，则定义域为凸集，且

$$f(\theta x+(1-\theta)y) \preceq_K \theta f(x)+(1-\theta)f(y)$$

  - 此处：$x,\ y \in \mathbf{dom}\, f,\ 0 \le \theta \le 1$

- 例：$f:\mathbf{S}^m \to \mathbf{S}^m,\ f(X)=X^2$ 为 $\mathbf{S}_+^m$-凸函数
- 证明：给定 $z \in \mathbf{R}^m,\ z^T X^2 z = \|Xz\|_2^2$ 为关于X的凸函数

$$z^T(\theta X+(1-\theta)Y)^2 z \le \theta z^T X^2 z+(1-\theta)z^T Y^2 z$$

此处：$X,Y \in \mathbf{S}^m,\ 0 \le \theta \le 1$

关于广义不等式的凸性

- 函数 $f:\mathbf{R}^n\to\mathbf{R}^m$为$K$-凸函数，则定义域为凸集，且

$$f(\theta x+(1-\theta)y)\preceq_K \theta f(x)+(1-\theta)f(y)$$

  - 此处：$x,\ y\in\mathbf{dom}\, f,\ 0\le\theta\le 1$

- 例：$f:\mathbf{S}^m\to\mathbf{S}^m,\ f(X)=X^2$ 为$\mathbf{S}_+^m$-凸函数
- 证明：给定 $z\in\mathbf{R}^m,\ z^TX^2z=\|Xz\|_2^2$ 为关于X的凸函数

$$z^T(\theta X+(1-\theta)Y)^2z\le\theta z^TX^2z+(1-\theta)z^TY^2z$$

此处：$X,Y\in\mathbf{S}^m,\ 0\le\theta\le 1$

$$(\theta X+(1-\theta)Y)^2\preceq\theta X^2+(1-\theta)Y^2$$