最优化方法

东南大学
计算机&人工智能学院
飞
songmf@seu.edu.cn

数

- 基本性质和案例
- 保　运算
- 共　　数
- 　　　数
- 对数-　　数和对数-　　数
- 关于广义不等式的　性

定义

定义

- $f : \mathbf{R}^n \to \mathbf{R}$ 为 数 则其满足定义域为 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 数 则其满足定义域为 集 且

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

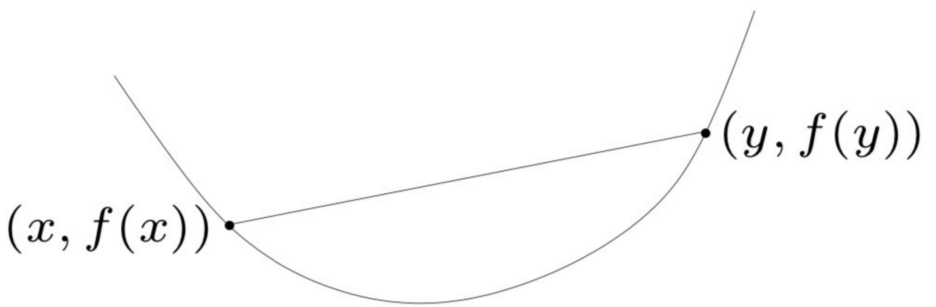
  - 对所有 $x, y \in \mathbf{dom}\, f$, $0 \le \theta \le 1$

定义

- $f : \mathbf{R}^n \to \mathbf{R}$ 为　　　数　则其满足定义域为　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　　　数　则-f为　　　数
- 　射　数

严格　　数

严格　　数

- f为严格　　数　则其满足定义域为　集　且

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

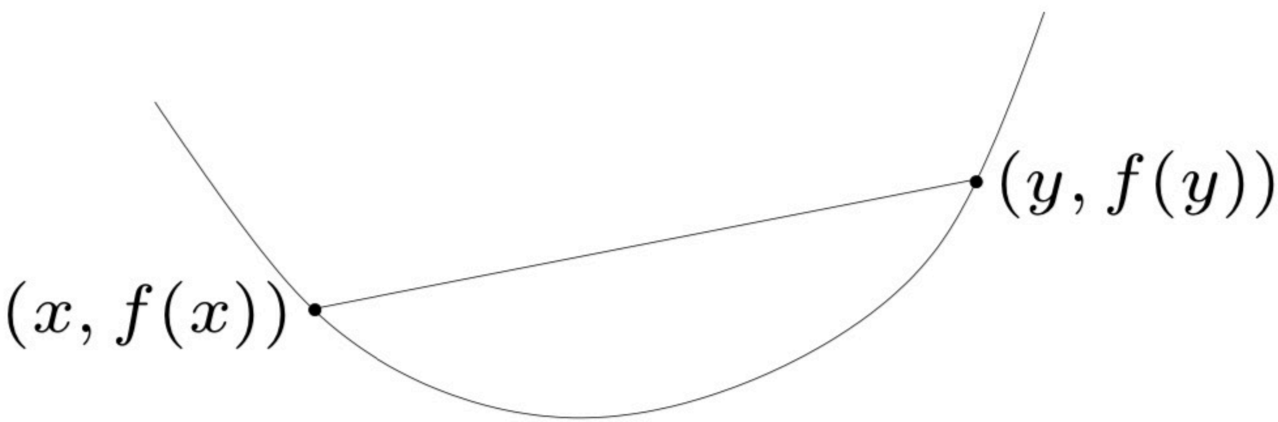
  - 对所有$x, y \in \mathbf{dom}\, f,\ x \neq y,\ 0 < \theta < 1$

严格 数

- f为严格 数 则其满足定义域为 集 且

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

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

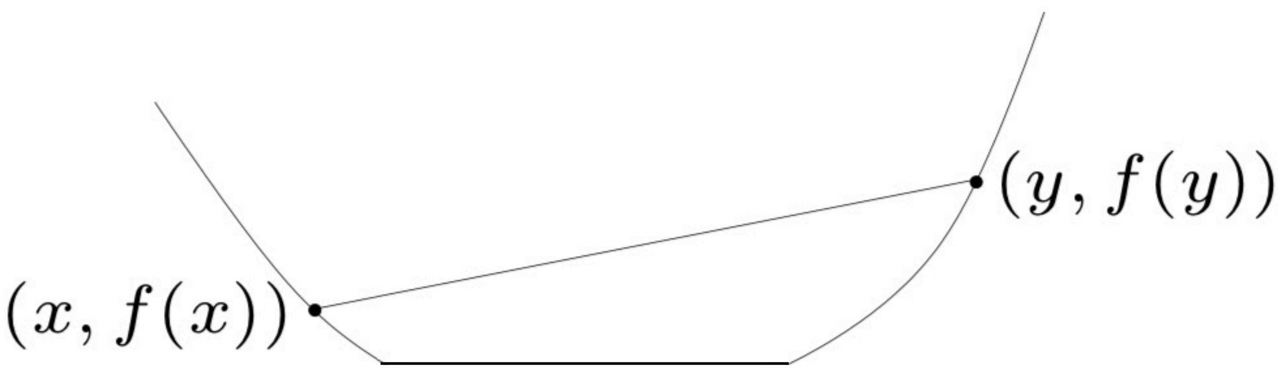


严格　　数

- f为严格　　数　则其满足定义域为　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　　数到直线

约束 数到直线

- $f : \mathbf{R}^n \to \mathbf{R}$为 数 当且仅当 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 数 对任意 $x \in \mathbf{dom}\, f,\ v \in \mathbf{R}^n$

约束 数到直线

- $f:\mathbf{R}^n \to \mathbf{R}$ 为 数 当且仅当 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 数 对任意 $x \in \mathbf{dom}\, f,\ v \in \mathbf{R}^n$

- 可将 数的检测转化为单变量 数 性的检测

约束 数到直线

- $f:\mathbf{R}^n \to \mathbf{R}$为 数 当且仅当 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 数 对任意 $x \in \mathbf{dom}\, f$, $v \in \mathbf{R}^n$
- 可将 数的检测转化为单变量 数 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 数到直线

- $f:\mathbf{R}^n \to \mathbf{R}$ 为 数 当且仅当 数 $g:\mathbf{R}\to\mathbf{R}$

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

  为 数 对任意 $x \in \mathbf{dom}\, f$, $v \in \mathbf{R}^n$

- 可将 数的检测转化为单变量 数 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 数到直线

- $f:\mathbf{R}^n\to\mathbf{R}$为 数 当且仅当 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 数 对任意 $x\in\mathbf{dom}\,f,\ v\in\mathbf{R}^n$

- 可将 数的检测转化为单变量 数 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 数到直线

- $f:\mathbf{R}^n \to \mathbf{R}$ 为 数 当且仅当 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 数 对任意 $x\in\mathbf{dom}\, f$, $v\in\mathbf{R}^n$

- 可将 数的检测转化为单变量 数 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 数g 为 数 因此f也为 数

展值延

展值延

- 数f的　展值延　$\tilde{f}$表示

展值延

- 数f的　展值延　$\tilde{f}$表示

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

展值延

- 数f的 展值延 $\tilde{f}$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 数的表示:

展值延

- 数f的 展值延 $\tilde{f}$ 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 数的表示:

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

展值延

- 数f的 展值延 $\tilde{f}$ 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 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.** 数f的定义域为 集
  - **2.** 对 $x, y \in \mathbf{dom}\, f$,

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

展值延

- 例　　集的示性　数是　　数

展值延

- 例　集的示性　数是　数

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

展值延

- 例 集的示性 数是 数

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

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

一　条件

一　条件

- 数f可微　其满足　数f的定义域是开集　且度

一　条件

- 数f可微　其满足　数f的定义域是开集　且度

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

- 在定义域处处存在

一　条件

- 数f可微　其满足　数f的定义域是开集　且度

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

  - 在定义域处处存在

- 一　条件　定义域为　集的可微　数f为数　当且仅当

一 条件

- 数f可微 其满足 数f的定义域是开集 且 度

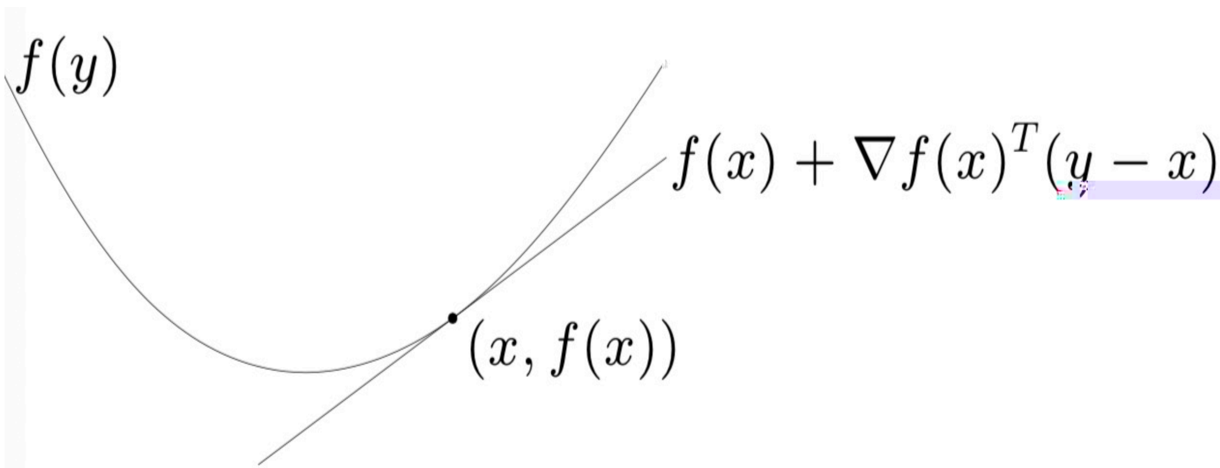
$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \ldots, \frac{\partial f(x)}{\partial x_n} \right)$$

  - 在定义域处处存在

- 一 条件 定义域为 集的可微 数f为 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　条件



数f的一　近似表示全局下估计

一　条件的证明

一　条件的证明

- 考虑一维情况

一　条件的证明

- 考虑一维情况
  - f　$R \to R$为　　　定义域为　　　且

一　条件的证明

- 考虑一维情况
  - f　$R \to R$为　　定义域为　　且

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

一　条件的证明

- 考虑一维情况
  - f　$R \rightarrow R$为　　　定义域为　　　且

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

- 证明

一　条件的证明

- 考虑一维情况
  - f　$R \to R$为　　　定义域为　　　且

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

- 证明
  - f为　　　对定义域内的任意两点 和　　实数

一　条件的证明

- 考虑一维情况
  - *f*　$R \rightarrow R$为　　　定义域为　　　且

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

- 证明
  - *f*为　　　对定义域内的任意两点 和　　实数

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

一　条件的证明

- 考虑一维情况
  - f　$R \to R$为　　　定义域为　　　且

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

- 证明
  - f为　　　对定义域内的任意两点 和　　　实数

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

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

一　条件的证明

- 考虑一维情况
  - f　$R \to R$为　　定义域为　　且
$$f(y) \geq f(x) + f'(x)(y - x)$$
- 证明
  - f为　　对定义域内的任意两点 和　实数
$$0 < t \leq 1,\ x + t(y - x) \in \mathbf{dom}\, f$$
$$f(x + t(y - x)) \leq (1 - t)f(x) + tf(y)$$
$$tf(y) \quad tf(x) + f(x + t(y \quad x)) \quad f(x)$$

一　条件的证明

- 考虑一维情况
  - f　$R \rightarrow R$为　　定义域为　　且

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

- 证明
  - f为　　对定义域内的任意两点 和　　实数

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

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

$$tf(y) \quad tf(x) + f(x + t(y \quad x)) \quad f(x)$$

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

一　条件的证明

- 考虑一维情况
  - f　$R \to R$为　　定义域为　　且

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

- 证明
  - f为　　对定义域内的任意两点 和　　实数

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

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

$$tf(y) \quad tf(x) + f(x+t(y \quad x)) \quad f(x)$$

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

一　条件的证明

- 考虑一维情况
  - f　$R \rightarrow R$为　　定义域为　　且

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

- 证明
  - f为　　对定义域内的任意两点 和　　实数

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

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

$$tf(y) \quad tf(x) + f(x + t(y \quad x)) \quad f(x)$$

lim $_{0+}$ $\quad f(y) \geq f(x) + \boxed{\dfrac{f(x + t(y-x)) - f(x)}{t}}$

一　条件的证明

一　条件的证明

- 证明

一　条件的证明

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

一　条件的证明

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

一　条件的证明

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

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

一　条件的证明

- 证明
  - 设定义域两个点 $x \neq y$ 以及 $0 \leq \theta \leq 1$
  - 通过　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　条件的证明

证明

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

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

一　条件的证明

❑ 证明

- ❖ 设定义域两个点 $x \neq y$ 以及 $0 \leq \theta \leq 1$
- ❖ 通过　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　条件的证明

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

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

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

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

一　条件的证明

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

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

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

$$\theta f(x) + (1-\theta) f(y) \ge f(z) + f'(z)\left(\boxed{\theta(x-z) + (1-\theta)(y-z)}\right)$$

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

一　条件的证明

一　条件的证明

- 高维情况

一　条件的证明

- 高维情况
- 证明
  - 设f为　考虑定义域内两点 和 构造 数g

一　条件的证明

- 高维情况
- 证明
  - 设f为　　　考虑定义域内两点 和　　构造　数g

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

一　条件的证明

- 高维情况
- 证明
  - 设f为　　考虑定义域内两点 和　构造　数g

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

一 条件的证明

- 高维情况
- 证明
  - 设f为 考虑定义域内两点 和 构造 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 条件的证明

- 高维情况
- 证明
  - 设f为　　考虑定义域内两点 和　　构造　　数g

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

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

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

一　条件的证明

- 高维情况
- 证明
  - 设f为　　考虑定义域内两点　和　　构造　　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 条件的证明

- 高维情况
- 证明
  - 设f为 考虑定义域内两点 和 构造 数g

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

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

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

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

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

一　条件的证明

一　条件的证明

- 证明
  - 考虑定义域内两点 和

一　条件的证明

- 证明
  - 考虑定义域内两点 和

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

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

一　条件的证明

- 证明
  - 考虑定义域内两点 和

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

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

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

一　条件的证明

- 证明
  - 考虑定义域内两点 和

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

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

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

一　条件的证明

- 证明
  - 考虑定义域内两点 和

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

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

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

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

一　条件的证明

- 证明
  - 考虑定义域内两点 和

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

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

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

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

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

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

一　条件的证明

- 证明
  - 考虑定义域内两点 和

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

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

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

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

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

一　条件的证明

证明

- 考虑定义域内两点 和

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

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

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

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

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

一　条件的证明

证明

- 考虑定义域内两点 和

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

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

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

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

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

二　条件

二 条件

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

二 条件

- 数 f 为二 可微 其满足 数 f 的定义域是开集 且**Hessian** 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　条件

- 数f为二　可微　其满足　数f的定义域是开集　且Hessian　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　条件
  - 对定义域为　集的二　可微　数f　若其为　数　当且仅当

二　条件

- 数f为二　可微　其满足　数f的定义域是开集　且**Hessian**　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　条件
  - 对定义域为　集的二　可微　数f　若其为　数　当且仅当

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

二　条件

- 数f为二　可微　其满足　数f的定义域是开集　且Hessian　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　条件
  - 对定义域为　集的二　可微　数f　若其为　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　数

例

例

- 二次　数

例

- 二次 数

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

例

- 二次 数

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

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

例

- 二次 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 数

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

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

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

- 若 $P \succeq 0$ 则为 数

例

- 二次　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　　数
- 问题
  - 严格　时　P是否一定为正定　阵

例

例

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

例

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

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

例

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

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

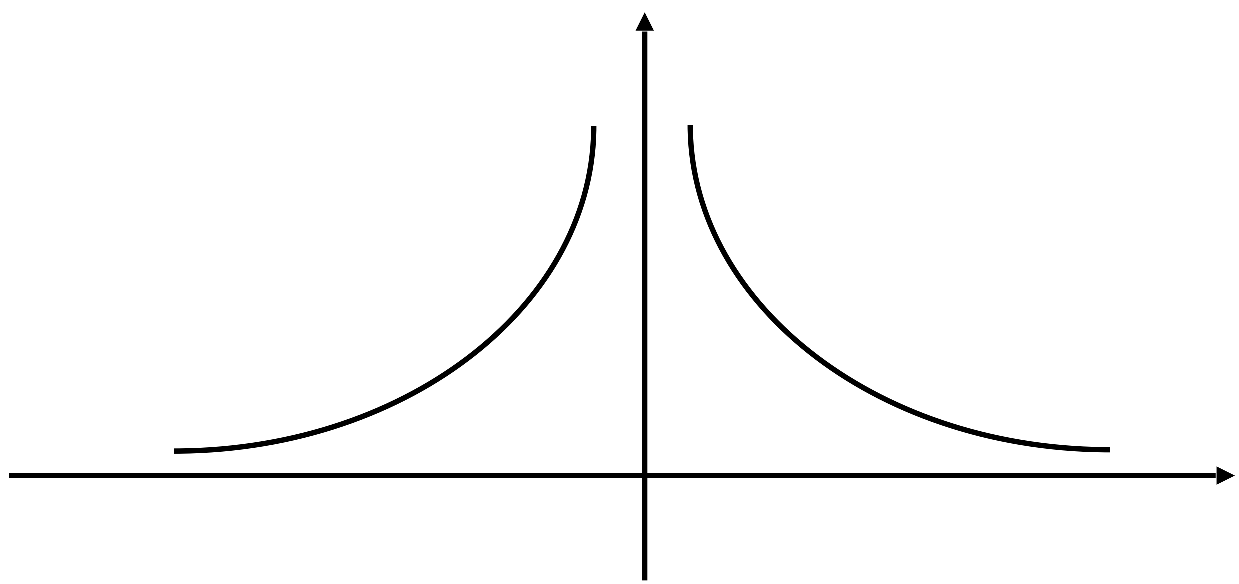
$$f''(x) > 0 \quad \textbf{(6 -4)}$$

例

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

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

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



例

例

- 射　数 $ax+b$ $\nabla^2 f(x)=0$

例

- 射　数 $ax+b$ $\nabla^2 f(x) = 0$
- 指数　数 e^{ax}

例

- 射　数 $ax+b$ $\nabla^2 f(x) = 0$
- 指数　数 e^{ax}
  - $f'(x) = ae^{ax}$

例

- 射　数 $ax+b$ $\nabla^2 f(x) = 0$
- 指数　数 e^{ax}
  - $f'(x) = ae^{ax}$
  - $f''(x) = a^2e^{ax}$

例

- 射 数 $ax+b$ $\nabla^2 f(x)=0$
- 指数 数 e^{ax}
  - $f'(x)=ae^{ax}$
  - $f''(x)=a^2e^{ax}$
- 数 x^a 为正实数

例

- 射　数 $ax+b$ $\nabla^2 f(x) = 0$
- 指数　数 e^{ax}
  - $f'(x) = ae^{ax}$
  - $f''(x) = a^2e^{ax}$
- 数　x^a　为正实数
  - $f'(x) = ax^{a\ \ 1}$

例

- 射　数 $ax+b$　$\nabla^2 f(x) = 0$
- 指数　数 e^{ax}
  - $f'(x) = ae^{ax}$
  - $f''(x) = a^2e^{ax}$
- 　数　x^a　　为正实数
  - $f'(x) = ax^{a \quad 1}$
  - $f''(x) = a(a \quad 1)x^{a \quad 2}$

例

- 射 数 $ax+b$ $\nabla^2 f(x) = 0$
- 指数 数 e^{ax}
  - $f'(x) = ae^{ax}$
  - $f''(x) = a^2 e^{ax}$
- 数 x^a 为正实数
  - $f'(x) = ax^{a\ 1}$
  - $f''(x) = a(a\ \ 1)x^{a\ 2}$
  - $f''(x) = \begin{cases} 0 & \text{if } a\ \ 1 \text{ or } a\ \ 0 \\ 0 & \text{if } 0\ \ a\ \ 1 \end{cases}$

例

例

- 绝对值的　　　数 $|x|^p$

例

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

例

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

例

- 绝对值的　　　数

❖ $$f'(x) = \begin{cases} px^{p\quad 1} & \text{if } x \quad 0 \\ \quad p(\quad x)^{p\quad 1} & \text{if } x < 0 \end{cases}$$

❖ $$f''(x) = \begin{cases} p(p \quad 1)x^{p\quad 2} & \textbf{if } x \quad 0 \\ p(p \quad 1)(\quad x)^{p\quad 2} & \textbf{if } x < 0 \end{cases}$$

❖

例

- 绝对值的　数 $|x|^p$
  - $$f'(x) = \begin{cases} px^{p\quad 1} & \text{if } x \quad 0 \\ \quad p(\quad x)^{p\quad 1} & \text{if } x < 0 \end{cases}$$
  - $$f''(x) = \begin{cases} p(p \quad 1)x^{p\quad 2} & \textbf{if } x \quad 0 \\ p(p \quad 1)(\quad x)^{p\quad 2} & \textbf{if } x < 0 \end{cases}$$
  - $p > 1$ 为
  - $p = 1$ 为

例

- 绝对值的　　数 $|x|^p$
  - $f'(x) = \begin{cases} px^{p\ \ 1} & \text{if } x\ \ \ 0 \\ p(\ \ x)^{p\ \ 1} & \text{if } x < 0 \end{cases}$
  - $f''(x) = \begin{cases} p(p\ \ \ 1)x^{p\ \ 2} & \textbf{if } x\ \ \ 0 \\ p(p\ \ \ 1)(\ \ x)^{p\ \ 2} & \textbf{if } x < 0 \end{cases}$
  - $p > 1$ 为
  - $p = 1$ 为
  - $p < 1$??

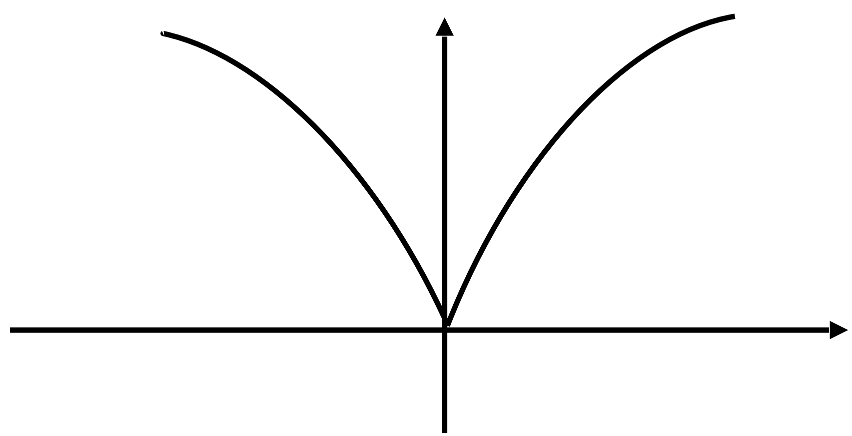
例

- 绝对值的 数

  - $$f'(x) = \begin{cases} p x^{p\quad 1} & \text{if } x \quad 0 \\ p(\quad x)^{p\quad 1} & \text{if } x < 0 \end{cases}$$

  - $$f''(x) = \begin{cases} p(p \quad 1) x^{p\quad 2} & \textbf{if } x \quad 0 \end{cases}$$

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



例

例

- 对数　数 $\log x$，为正实数

例

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

例

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

例

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

例

- 对数　数 $\log x$，为正实数
  - $f'(x) = 1/x$
  - $f''(x) = \quad 1/x^2 < 0$
- 负　$x \log x$，为正实数
  - $f'(x) = logx + 1$

例

- 对数　数 $\log x$，为正实数
  - $f'(x) = 1/x$
  - $f''(x) = \quad 1/x^2 < 0$
- 负　$x \log x$，为正实数
  - $f'(x) = logx + 1$
  - $f''(x) = 1/x > 0$

例

例

- 范数　**R** 空间的范数$P(\ )$

例

- 范数　R 空间的范数$P(\ )$
  - $P(a\ \ ) = |a|\, P(\ )$
  - $P(\ \ +\ \ ) \le P(\ ) + P(\ )$
  - $P(\ ) = 0 \qquad\qquad = 0$

例

- 范数 **R** 空间的范数$P(\ )$
  - ❖ $P(a\ \ ) = |a|\, P(\ )$
  - ❖ $P(\ \ +\ \ ) \le P(\ ) + P(\ )$
  - ❖ $P(\ ) = 0 \qquad\qquad = 0$

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

例

- 范数 **R** 空间的范数*P*( )
  - $P(a \quad) = |a|\ P(\quad)$
  - $P(\quad + \quad) \leq P(\quad) + P(\quad)$
  - $P(\quad) = 0 \qquad\qquad = 0$

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

例

- 范数　R 空间的范数$P(\ )$
  - $P(a\ ) = |a|\ P(\ )$
  - $P(\ +\ ) \le P(\ ) + P(\ )$
  - $P(\ ) = 0\qquad = 0$

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

- 零范数　计算非零元素的数目

例

- 范数　R 空间的范数**P( )**
  - $P(a\ \ ) = |a|\ P(\ )$
  - $P(\ \ +\ \ ) \leq P(\ ) + P(\ )$
  - $P(\ ) = 0 \qquad\quad = 0$

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

- 零范数　计算非零元素的数目
  - 令　**=1**

例

- 范数　**R** 空间的范数$P(\ )$
  - $P(a\ \ ) = |a|\ P(\ )$
  - $P(\ \ +\ \ ) \le P(\ ) + P(\ )$
  - $P(\ ) = 0 \qquad\qquad = 0$

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

- 零范数　计算非零元素的数目
  - 令　**=1**
  - 不是范数

例

例

- 极大值　数

例

- 极大值　数

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

例

- 极大值　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　数

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

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

例

- 极大值　数

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

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

例

- 极大值　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 $f(x) = \log\left(e^{x_1} + \cdots + e^{x_n}\right)$

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

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

例

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

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

  - $\dfrac{f}{x_i} = \dfrac{e^{x_i}}{e^{x_1} + \ldots + e^{x_n}}$
  - **Hessian** 阵$[H_{ij}]$

例

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

例

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

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

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

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

例

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

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

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

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

例

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

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

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

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

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** 阵$[H_{ij}]$

❖ $i \neq j$: $\dfrac{{}^2 f}{x_i \ \ x_j} = \dfrac{e^{x_i}e^{x_j}}{(e^{x_1} + \ldots + e^{x_n})^2}$

❖ $i = j$: $\dfrac{{}^2 f}{x_i \ \ x_i} = \dfrac{e^{x_i}e^{x_i} + e^{x_i}(e^{x_1} + \ldots + e^{x_n})}{(\boxed{e^{x_1} + \ldots + e^{x_n}})^2}$ $= \mathbf{1}^T z$

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

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** 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**　阵半正定
  - 对任意　　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** 阵半正定
  - 对任意 有
  - $^T K =$

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** 阵半正定
  - 对任意 有
  - $^TK = (\mathbf{1}^T z)v^T\mathbf{diag}(z)v - v^Tzz^Tv$

例

- 指数和的对数

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

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

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

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** 阵半正定
  - 对任意 有
  - $^TK = (\mathbf{1}^T z) v^T \mathbf{diag}(z) v - v^T z z^T v$

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

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

例

- 指数和的对数

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

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

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

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

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 阵半正定
  - 对任意 有
  - $^TK = (\mathbf{1}^T z) v^T \mathbf{diag}(z) v - v^T z z^T v$

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

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

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 阵半正定
  - 对任意 有
  - $^TK = (\mathbf{1}^T z) v^T \mathbf{diag}(z) v - v^T zz^T v$

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

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

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 阵半正定
  - 对任意 有
  - $^TK = (\mathbf{1}^T z) v^T \mathbf{diag}(z) v - v^T z z^T v$

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

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

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　阵半正定
  - 对任意　 有
  - $^TK = (\mathbf{1}^T z) v^T \mathbf{diag}(z) v - v^T z z^T v$

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

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

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** 阵半正定
  - 对任意 有
  - $^TK = (\mathbf{1}^T z) v^T \mathbf{diag}(z) v - v^T zz^T v$

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

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

例

例

- 几何平　　　　　　　数

例

- 几何平 数

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

例

- 几何平 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　　　　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 数

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

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

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

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

例

- 几何平 数

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

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

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

$$\nabla^2 f(x) \;\text{[redacted]}\; \frac{\prod_{i=1}^{n} x_i^{1/n}}{n^2} \;\text{[redacted]}\quad q_i = 1/x_i$$

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

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

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

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

- 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
- $g'(t) = \sum_{i=1}^{n} \frac{\lambda_i}{1+t\lambda_i}$
- $g''(t) = \sum_{i=1}^{n} \frac{\lambda_i^2}{(1+t\lambda_i)^2}$
- $g''(t) \quad 0$

保　运算

保 运算

- 如何 断某一 数是否为 数

保 运算

- 如何 断某一 数是否为 数
  - 方式**1:**验证定义 通过将 数限制在直线 断

保　运算

- 如何　断某一　数是否为　　数
  - 方式**1:**验证定义　通过将　数限制在直线　断
  - 方式**2:**对于二　可微　数　验证 $\nabla^2 f(x) \succeq 0$

保　运算

- 如何　断某一　数是否为　　数
  - 方式1:验证定义　通过将　数限制在直线　断
  - 方式2:对于二　可微　数　验证 $\nabla^2 f(x) \succeq 0$
  - 方式3:考察　数是否可由多个简单　　数通过保　运算得到

保　运算

- 如何　断某一　数是否为　　数
  - 方式1:验证定义　通过将　数限制在直线　断
  - 方式2:对于二　可微　数　验证 $\nabla^2 f(x) \succeq 0$
  - 方式3:考察　数是否可由多个简单　　数通过保　运算得到
    - 非负加权求和

保 运算

- 如何 断某一 数是否为 数
  - 方式1:验证定义 通过将 数限制在直线 断
  - 方式2:对于二 可微 数 验证 $\nabla^2 f(x) \succeq 0$
  - 方式3:考察 数是否可由多个简单 数通过保 运算得到
    - 非负加权求和
    - 复合 射 射

保 运算

- 如何 断某一 数是否为 数
  - 方式1:验证定义 通过将 数限制在直线 断
  - 方式2:对于二 可微 数 验证 $\nabla^2 f(x) \succeq 0$
  - 方式3:考察 数是否可由多个简单 数通过保 运算得到
    - 非负加权求和
    - 复合 射 射
    - 逐点最大和逐点上确界

保 运算

- 如何 断某一 数是否为 数
  - 方式**1**:验证定义 通过将 数限制在直线 断
  - 方式**2**:对于二 可微 数 验证 $\nabla^2 f(x) \succeq 0$
  - 方式**3**:考察 数是否可由多个简单 数通过保 运算得到
    - 非负加权求和
    - 复合 射 射
    - 逐点最大和逐点上确界
    - 复合

保　运算

- 如何　断某一　数是否为　　数
  - 方式**1**:验证定义　通过将　数限制在直线　断
  - 方式**2**:对于二　可微　数　验证 $\nabla^2 f(x) \succeq 0$
  - 方式**3**:考察　数是否可由多个简单　　数通过保　运算得到
    - 非负加权求和
    - 复合　射 射
    - 逐点最大和逐点上确界
    - 复合
    - 最小化

保 运算

- 如何 断某一 数是否为 数
  - 方式**1**:验证定义 通过将 数限制在直线 断
  - 方式**2**:对于二 可微 数 验证
  - 方式**3**:考察 数是否可由多个简单 数通过保运算得到
    - 非负加权求和

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　　数　当　数f为　　数　$\alpha \geq 0$

非负加权求和

- 非负乘积
  - αf 为 数 当 数f 为 数 $\alpha \geq 0$
- 求和

非负加权求和

- 非负乘积
  - αf 为　　数　当　数f为　　数　$\alpha \geq 0$
- 求和
  - $f_1 + f_2$为　　数　当 f_1, f_2为　　数

非负加权求和

- 非负乘积
  - αf 为 数 当 数f 为 数 $\alpha \geq 0$
- 求和
  - $f_1 + f_2$为 数 当 f_1, f_2为 数
- 非负加权求和 $f = w_1 f_1 + \cdots + w_m f_m$
  - 是保 运算

非负加权求和

- 非负乘积
  - αf 为 数 当 数f为 数 $\alpha \geq 0$
- 求和
  - $f_1 + f_2$为 数 当 f_1, f_2为 数
- 非负加权求和 $f = w_1 f_1 + \cdots + w_m f_m$
  - 是保 运算
- 可 展到无限项 积分

非负加权求和

- 非负乘积
  - αf 为 数 当 数 f 为 数 $\alpha \geq 0$
- 求和
  - $f_1 + f_2$ 为 数 当 f_1, f_2 为 数
- 非负加权求和 $f = w_1 f_1 + \cdots + w_m f_m$
  - 是保 运算
- 可 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　　数　当　数f为　　数　$\alpha \geq 0$
- 求和
  - $f_1 + f_2$为　　数　当 f_1, f_2为　　数
- 非负加权求和　$f = w_1 f_1 + \cdots + w_m f_m$
  - 是保　运算
- 可　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　$f(\ ,\ )$是　　数

复合 射 射

复合 射 射

- 复合 射 射 $f(Ax+b)$ 为 数 当 数f为 数

复合 射 射

- 复合 射 射 $f(Ax+b)$ 为 数 当 数f为 数
- 证明 令 $g(x)=f(Ax+b)$

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

复合　射　射

- 复合　射　射　　　　　　为　　数　　当　　数f为数
- 证明　令

复合 射 射

- 复合 射 射 $f(Ax+b)$ 为 数 当 数 f 为 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 $\mathbf{dom}\, g$, 0 θ 1
  - $g(\theta x + (1 \quad \theta) y)$

复合　射 射

- 复合　射 射 $f(Ax+b)$ 为　数　当　数f为　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　$\mathbf{dom}\, g$, 0　θ　1
  - $g(\theta x + (1 \quad \theta) y)$

$$= f(\theta A x + (1 \quad \theta) A y + b)$$

复合 射 射

- 复合 射 射 $f(Ax+b)$ 为 数 当 数f为 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 $\mathbf{dom}\, g$, 0 θ 1
  - $g(\theta x+(1 \quad \theta)y)$

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

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

复合 射 射

- 复合 射 射 $f(Ax+b)$ 为 数 当 数 f 为 数
- 证明 令 $g(x)=f(Ax+b)$

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

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

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

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

$$\theta f(Ax+b)+(1 \quad \theta)f(Ay+b)$$

复合 射 射

- 复合 射 射 $f(Ax+b)$ 为 数 当 数f为 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 $\mathbf{dom}\, g$, $0 \quad \theta \quad 1$
  - $g(\theta x + (1 \quad \theta)y)$

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

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

$$\theta f(Ax+b) + (1 \quad \theta)f(Ay+b)$$

$$= \theta g(x) + (1 \quad \theta)g(y)$$

复合 射 射

复合 射 射

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

复合 射 射

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

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

复合 射 射

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

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

- 例 线性不等式的对数 数

复合 射 射

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

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

- 例 线性不等式的对数 数

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

复合 射 射

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

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

- 例 线性不等式的对数 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 射 射

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

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

- 例 线性不等式的对数 数

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

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

- 射 数的任意范数

复合 射 射

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

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

- 例 线性不等式的对数 数

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

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

- 射 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　　数,

逐点最大

- 若 $f_1, \ldots, f_m$ 为 数,
  $f(x) = \max\{f_1(x), \ldots, f_m(x)\}$ 为 数

逐点最大

- 若 $f_1, \ldots, f_m$ 为 数,

  $$f(x) = \max\{f_1(x), \ldots, f_m(x)\} \text{ 为 } \quad \text{数}$$

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

逐点最大

- 若 $f_1, \ldots, f_m$ 为 数,

  $$f(x) = \max\{f_1(x), \ldots, f_m(x)\} \text{ 为 数}$$

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

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

逐点最大

- 若 $f_1, \ldots, f_m$ 为 数,

  $$f(x) = \max\{f_1(x), \ldots, f_m(x)\} \text{ 为 } \quad \text{数}$$

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

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

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

逐点最大

- 若 $f_1, \ldots, f_m$ 为 数，

  $f(x) = \max\{f_1(x), \ldots, f_m(x)\}$ 为 数

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

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

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

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

逐点最大

- 若 $f_1, \ldots, f_m$ 为 数,

  $f(x) = \max\{f_1(x), \ldots, f_m(x)\}$ 为 数

- 取m=2:

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

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

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

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

    $= \theta f(x) + (1 \quad \theta)f(y)$

逐点最大

- 若 为 数,
  为 数
- 取m=**2:**
  -

逐点最大

- 若 $f_1, \ldots, f_m$ 为　　数,

$$f(x) = \max\{f_1(x), \ldots, f_m(x)\} \text{ 为　　数}$$

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

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

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

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

$$= \theta f(x) + (1 \quad \theta) f(y)$$

- 分段线性　数为　　数

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

逐点最大

逐点最大

- 最大 个分量之和($_{[i]}$为 中第i大的分量)

逐点最大

- 最大 个分量之和($_{[i]}$为 中第i大的分量)

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

逐点最大

- 最大 个分量之和($_{[i]}$为 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 个分量之和($_{[i]}$为 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 个分量之和($_{[i]}$为 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 为 数 则

逐点上确界

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

逐点上确界

- 若对任意$y \in \mathcal{A}$ $f(x,y)$关于 为 数 则
- $$g(x) = \sup_{y \in \mathcal{A}} f(x,y)$$ 为 数
- 例

逐点上确界

- 若对任意$y \in \mathcal{A}$ $f(x,y)$ 关于 为 数 则
- $g(x) = \sup_{y \in \mathcal{A}} f(x,y)$ 为 数
- 例
  - 在集合C中 最远点的 离 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 为 数 则
- $g(x) = \sup\limits_{y \in \mathcal{A}} f(x,y)$为 数
- 例
  - 在集合C中 最远点的 离 数

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

  - 对称 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 为 数 则
- $g(x) = \sup_{y \in \mathcal{A}} f(x,y)$为 数

- 例
  - 在集合C中 最远点的 离 数

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

  - 对称 阵的最大特征值

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

数复合

数复合

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

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

数复合

- 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，k=1　定义域　为实数集　g和h二　可微

数复合

- 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，k=**1** 定义域 为实数集 g和h二 可微
- $\nabla^2 f(x) \succeq 0$ f为 数

数复合

- 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，k=1　定义域　为实数集　g和h二　可微
- $\nabla^2 f(x) \succeq 0$　f为　数

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

数复合

- 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，k=1 定义域 为实数集 g和h二 可微
- $\nabla^2 f(x) \succeq 0$ f为 数

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

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

数复合

- 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，k=1 定义域 为实数集 g和h二 可微
- $\nabla^2 f(x) \succeq 0$ f为 数

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

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

数复合

- 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，k=1　定义域　为实数集　g和h二　可微
- $\nabla^2 f(x) \succeq 0$　f为　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　不减　g为　则f为

数复合

- 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，k=1　定义域　为实数集　g和h二　可微
- $\nabla^2 f(x) \succeq 0$　f为　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　不减　g为　则f为
- h为　不增　g为　则f为

数复合

- 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，*k*=1　定义域　为实数集　*g*和*h*二　可微
- $\nabla^2 f(x)\succeq 0$　*f*为　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　不减　*g*为　则*f*为
- *h*为　不增　*g*为　则*f*为
- *h*为　不减　增　*g*为　则*f*为

数复合

数复合

- 假设　展

数复合

- 假设　展
  - 高维　, k>1
  - 定义域不为全空间

数复合

- 假设　展
  - 高维　, k>1
  - 定义域不为全空间
  - h　g　不可微

数复合

- 假设　展
  - 高维　, k>1
  - 定义域不为全空间
  - h　g　不可微
- 　展值延

数复合

- 假设　展
  - 高维　, $k>1$
  - 定义域不为全空间
  - h　g　不可微
- 　展值延
- 　数 f 为　　数

数复合

- 假设 展
  - 高维 , *k*>1
  - 定义域不为全空间
  - h g 不可微
- 展值延
- 数f为 数
  - 当g为 数 h为 数 且h的 展值延 非减

数复合

- 假设　展
  - 高维　, k>1
  - 定义域不为全空间
  - h　g　不可微
- 　展值延
- 　数f为　　数
  - 当g为　　数　h为　　数　且h的　展值延　非减
  - 当g为　　数　h为　　数　且h的　展值延　非增

数复合

- 假设　展
  - 高维　, k>1
  - 定义域不为全空间
  - h　g　不可微
- 　展值延
- 　数 f 为　　数
  - 当 g 为　　数　h 为　　数　且 h 的　展值延　非减
  - 当 g 为　　数　h 为　　数　且 h 的　展值延　非增
  - 注意　　展值延　须保证单调性

数复合

数复合

- 展值延　无单调性

数复合

- 展值延　无单调性

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

数复合

- 展值延　无单调性

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

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

数复合

- 展值延　无单调性

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

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

$$f(x)=0$$

数复合

- 展值延　无单调性

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

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

$$f(x) = 0$$

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

数复合

数复合

- 证明　当g为　　h为　　且h的　展值延　非减　则f为

数复合

- 证明　当g为　　h为　　且h的　展值延　非减　则f为　x, $y \in \mathbf{dom}\, f$ $0 \le \theta \le 1$

数复合

- 证明 当g为 h为 且h的 展值延 非减 则f为 $x,\ y \in \mathbf{dom}\, f$ $0 \le \theta \le 1$

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

数复合

- 证明 当g为 h为 且h的 展值延 非减 则f为 x, $y \in \mathbf{dom}\, f$ $0 \le \theta \le 1$

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

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

数复合

- 证明　当g为　　h为　　且h的　展值延　非减　则f为　$x, y \in \mathbf{dom}\, f$ $0 \le \theta \le 1$

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

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

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

数复合

- 证明　当g为　　h为　　且h的　展值延　非减　则f为　$x, y \in \mathbf{dom}\, f$ $0 \le \theta \le 1$

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

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

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

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

数复合

- 证明 当g为 h为 且h的 展值延 非减 则f为 $x, y \in \mathbf{dom}\, f$ $0 \leq \theta \leq 1$

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

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

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

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

数复合

- 证明 当g为 h为 且h的 展值延 非减 则f为 $x,\ y \in \mathbf{dom}\, f$ $0 \le \theta \le 1$

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

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

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

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

$$f(y)$$

数复合

- 证明 当g为 h为 且h的 展值延 非减 则f为 $x, y \in \mathbf{dom}\, f$ $0 \le \theta \le 1$

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

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

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

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

$$f(y)$$

数复合

- 证明 当g为 h为 且h的 展值延 非减 则f为 $x, \; y \in \textbf{dom}\, f$ $0 \le \theta \le 1$

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

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

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

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

数复合

- 证明 当g为 h为 且h

数复合

- 证明 当g为 h为 且h的 展值延 非减 则f为 x, $y \in \mathbf{dom}\, f$ $0 \le \theta \le 1$

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

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

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

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

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

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

数复合

- 证明 当g为 h为 且h的 展值延 非减 则f为 $x, y \in \mathbf{dom}\, f$ $0 \le \theta \le 1$

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

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

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

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

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

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

数复合

数复合

- 若g为 则 $\exp g(x)$为 数

数复合

- 若g为　　则 $\exp g(x)$ 为　　数
  - $h($　$)=\exp$　　　　　　　展值不减

数复合

- 若g为 则 $\exp g(x)$为 数
  - $h($ $)=\exp$ 展值不减
- 若g为 数且为正 则 $\log g(x)$为 数

数复合

- 若g为　　　则 $\exp g(x)$为　　　数
  - $h(\ \ )$=exp　　　　　　　　　展值不减
- 若g为　　　数且为正　则 $\log g(x)$为　　　数
  - $h(\ \ )$=log　，　　　　　展值不减

数复合

- 若g为　　则 $\exp g(x)$为　　数
  - $h(\ )$=exp　　　　　展值不减
- 若g为　　数且为正　则 $\log g(x)$为　　数
  - $h(\ )$=log　，　　　展值不减
- 若g为　　数且为正　则 $1/g(x)$ 为　　数

数复合

- 若g为　　　则 $\exp g(x)$为　　　数
  - $h(\ )=\exp$　　　　　　　展值不减
- 若g为　　　数且为正　则 $\log g(x)$为　　　数
  - $h(\ )=\log$　，　　　　展值不减
- 若g为　　　数且为正　则 $1/g(x)$ 为　　　数
  - $h(\ )=1/$　，　　　　展值不增

数复合

- 若g为　　　则 $\exp g(x)$为　　　数
  - $h(\ )$=exp　　　　　　　　展值不减
- 若g为　　　数且为正　则 $\log g(x)$为　　　数
  - $h(\ )$=log　，　　　　　展值不减
- 若g为　　　数且为正　则 $1/g(x)$ 为　　　数
  - $h(\ )$=1/ ，　　　　　展值不增
- 若g为　　　数且非负　　**1**　则 $g(x)^p$ 为　　　数

数复合

- 若g为 则 $\exp g(x)$ 为 数
  - $h(\ )=\exp$ 展值不减
- 若g为 数且为正 则 $\log g(x)$ 为 数
  - $h(\ )=\log$, 展值不减
- 若g为 数且为正 则 $1/g(x)$ 为 数
  - $h(\ )=1/$, 展值不增
- 若g为 数且非负 **1** 则 $g(x)^p$ 为 数
  - $h(x) = x^p(x \quad 0)$, 展值不减

数复合

- 若g为　　则 $\exp g(x)$ 为　　数
  - $h($　$)=\exp$　　　　展值不减
- 若g为　　数且为正　则 $\log g(x)$ 为　　数
  - $h($　$)=\log$　,　　　展值不减
- 若g为　　数且为正　则 $1/g(x)$ 为　　数
  - $h($　$)=1/$　,　　　展值不增
- 若g为　　数且非负　　**1**　则 $g(x)^p$ 为　　数
  - $h(x) = x^p(x$　　$0)$,　　　展值不减

数复合

- 若g为 则 $\exp g(x)$ 为 数
  - $h($ $)$=exp 展值不减
- 若g为 数且为正 则 $\log g(x)$ 为 数
  - $h($ $)$=log , 展值不减
- 若g为 数且为正 则 $1/g(x)$ 为 数
  - $h($ $)$=1/ , 展值不增
- 若g为 数且非负 **1** 则 $g(x)^p$ 为 数
  - $h(x) = x^p(x \quad 0)$, 展值不减
  - $h(x) = \begin{cases} x^p & \textbf{if } x \quad 0 \\ 0 & \textbf{if } x < 0 \end{cases}$

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
  - $g(x) = \inf\limits_{y\ \ C} f(x,y)$为凸函数
- 例：$f(x,y) = x^T Ax + 2x^T By + y^T Cy$

最小化

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

最小化

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

最小化

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

透视　数

透视　数

- 　数 $f : \mathbf{R}^n \to \mathbf{R}$ 的透视　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 数

- 数 $f:\mathbf{R}^n \to \mathbf{R}$ 的透视 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 数 g为 数

透视 数

- 数 $f:\mathbf{R}^n\to\mathbf{R}$ 的透视 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 数 g为 数

- 例 欧几里得范数的平方 $f(x)=x^Tx$ 为 数

透视　数

- 　数 $f:\mathbf{R}^n\to\mathbf{R}$ 的透视　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　　数　g为　　数

- 例　欧几里得范数的平方 $f(x)=x^Tx$ 为　　数

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

透视　数

- 　数 $f:\mathbf{R}^n \to \mathbf{R}$ 的透视　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　　数　g为　　数

- 例　欧几里得范数的平方 $f(x)=x^Tx$ 为　　数

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

透视　数

- 　数 $f:\mathbf{R}^n\to\mathbf{R}$ 的透视　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　　数　g 为　　数

- 例　欧几里得范数的平方 $f(x)=x^Tx$ 为　　数

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

- 例　负的对数 $f(x)=-\log x$

透视　数

- 　数 $f:\mathbf{R}^n\to\mathbf{R}$ 的透视　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　　　数　g为　　　数

- 例　欧几里得范数的平方 $f(x)=x^Tx$ 为　　　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　数

透视　数

- 向量的相对

透视　　数

- 向量的相对

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

透视　数

- 向量的相对

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

  - 注意　不是非负加权和

透视　数

- 向量的相对

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

  - 注意　不是非负加权和

- **KL**散度

透视 数

- 向量的相对

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

  - 注意 不是非负加权和

- **KL**散度

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

透视　数

- 向量的相对

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

  - 注意　不是非负加权和

- **KL**散度

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

  - 　　数

下水平集

下水平集

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

下水平集

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

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

下水平集

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

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

- 数的所有下水平集为 集

下水平集

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

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

- 数的所有下水平集为 集
  - 对于 $x, y \in C_\alpha$ $0 \le \theta \le 1$

下水平集

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

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

- 数的所有下水平集为 集
  - 对于 $x,\ y \in C_\alpha$ $0 \le \theta \le 1$

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

下水平集

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

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

- 数的所有下水平集为 集
  - 对于 $x,\ y \in C_\alpha$ $0 \le \theta \le 1$

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

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

下水平集

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

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

- 数的所有下水平集为 集
  - 对于 $x,\ y \in C_\alpha$ $0 \le \theta \le 1$

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

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

下水平集

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

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

- 数的所有下水平集为 集
  - 对于 $x,\ y \in C_\alpha$ $0 \le \theta \le 1$

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

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

$$\le \alpha$$

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

下水平集

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

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

- 数的所有下水平集为 集
  - 对于 $x,\ y \in C_\alpha$ $0 \le \theta \le 1$

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

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

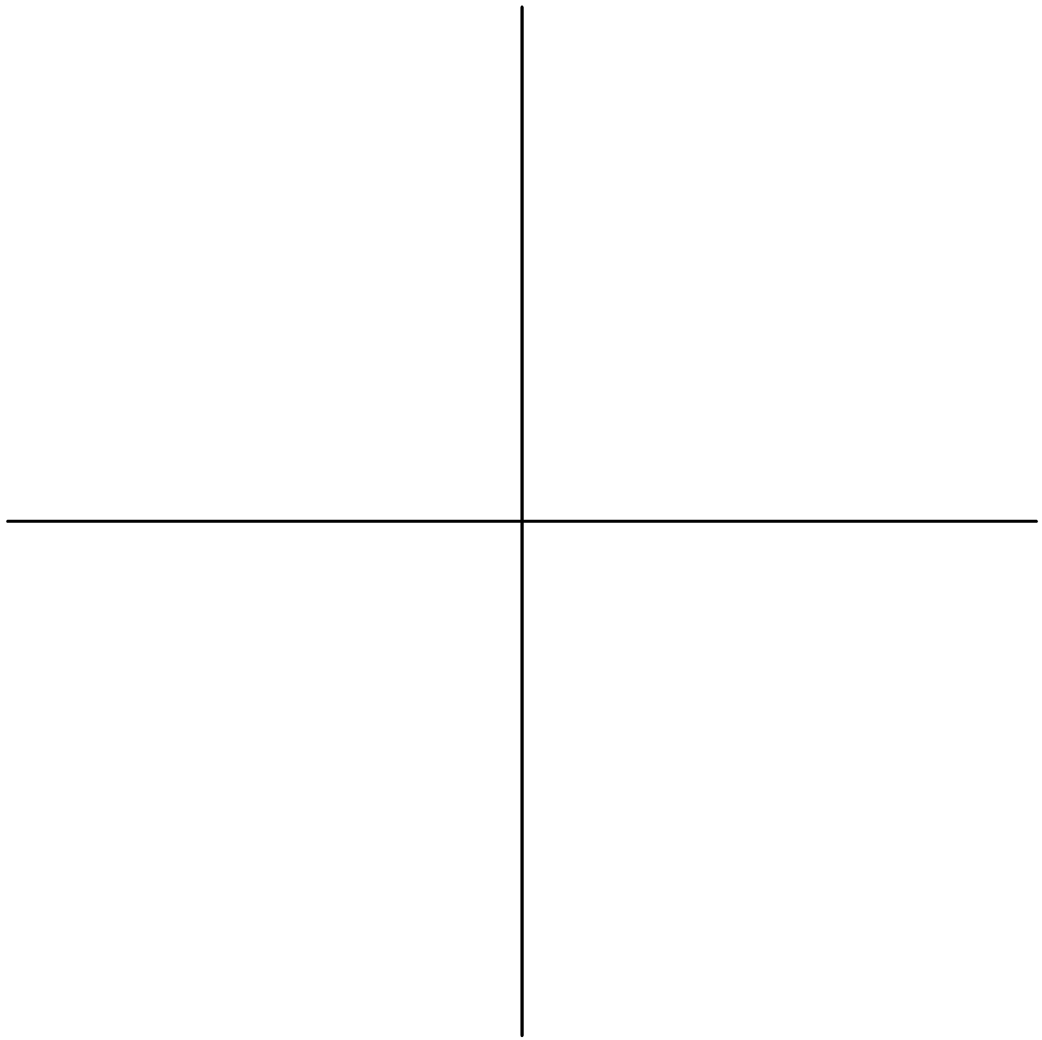
$$\le \alpha$$

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

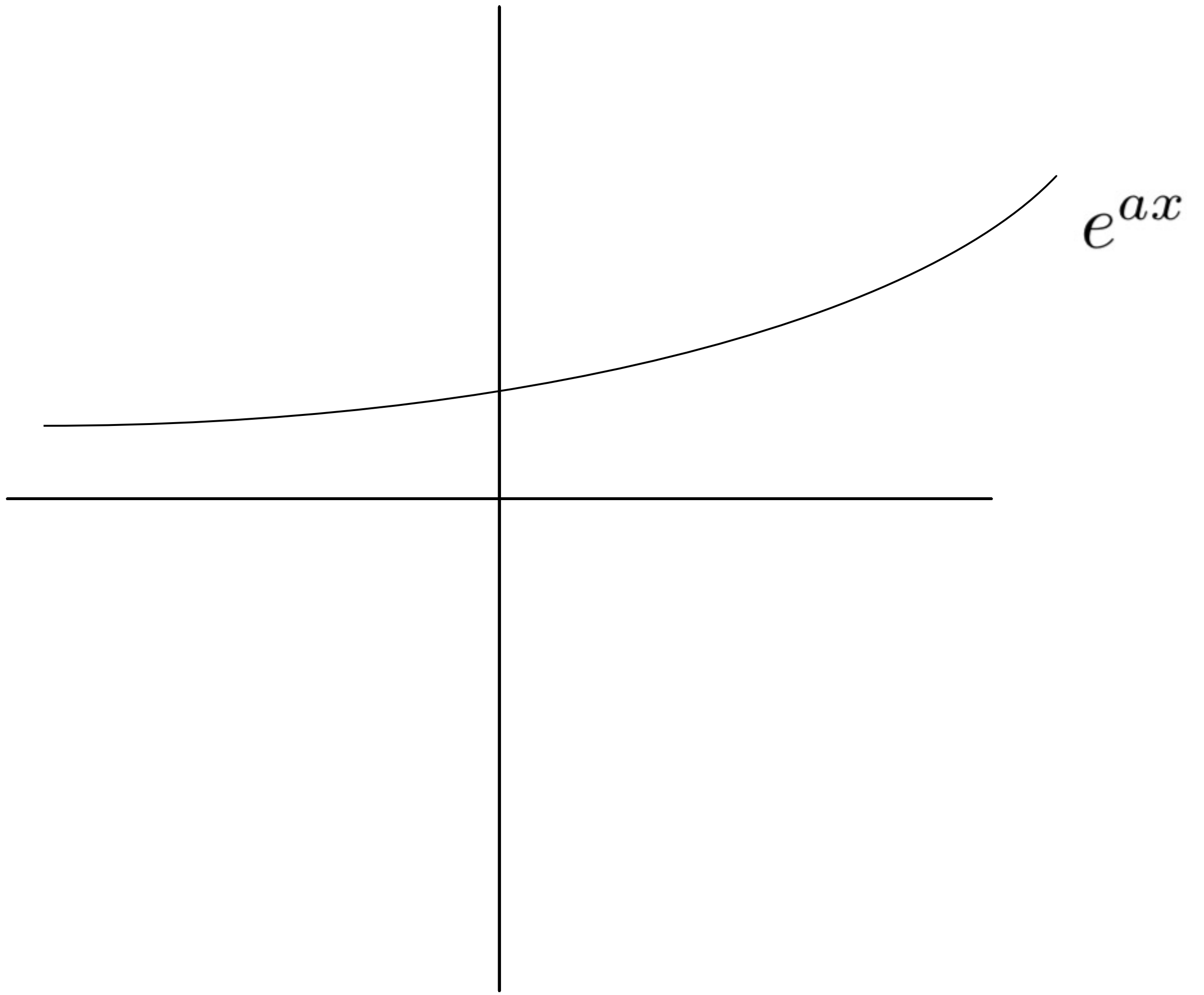
  - 反之不一定正确

下水平集

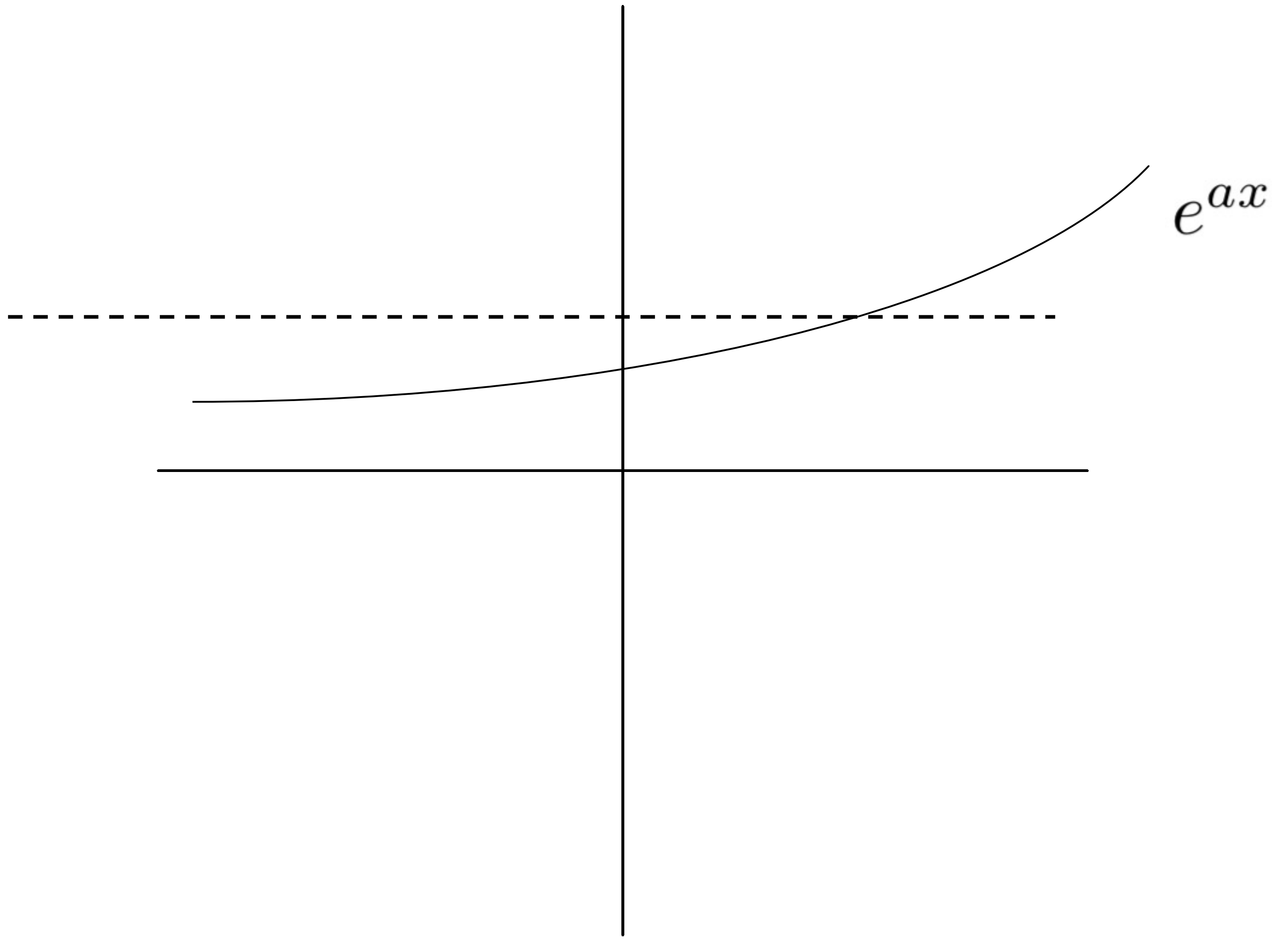
下水平集



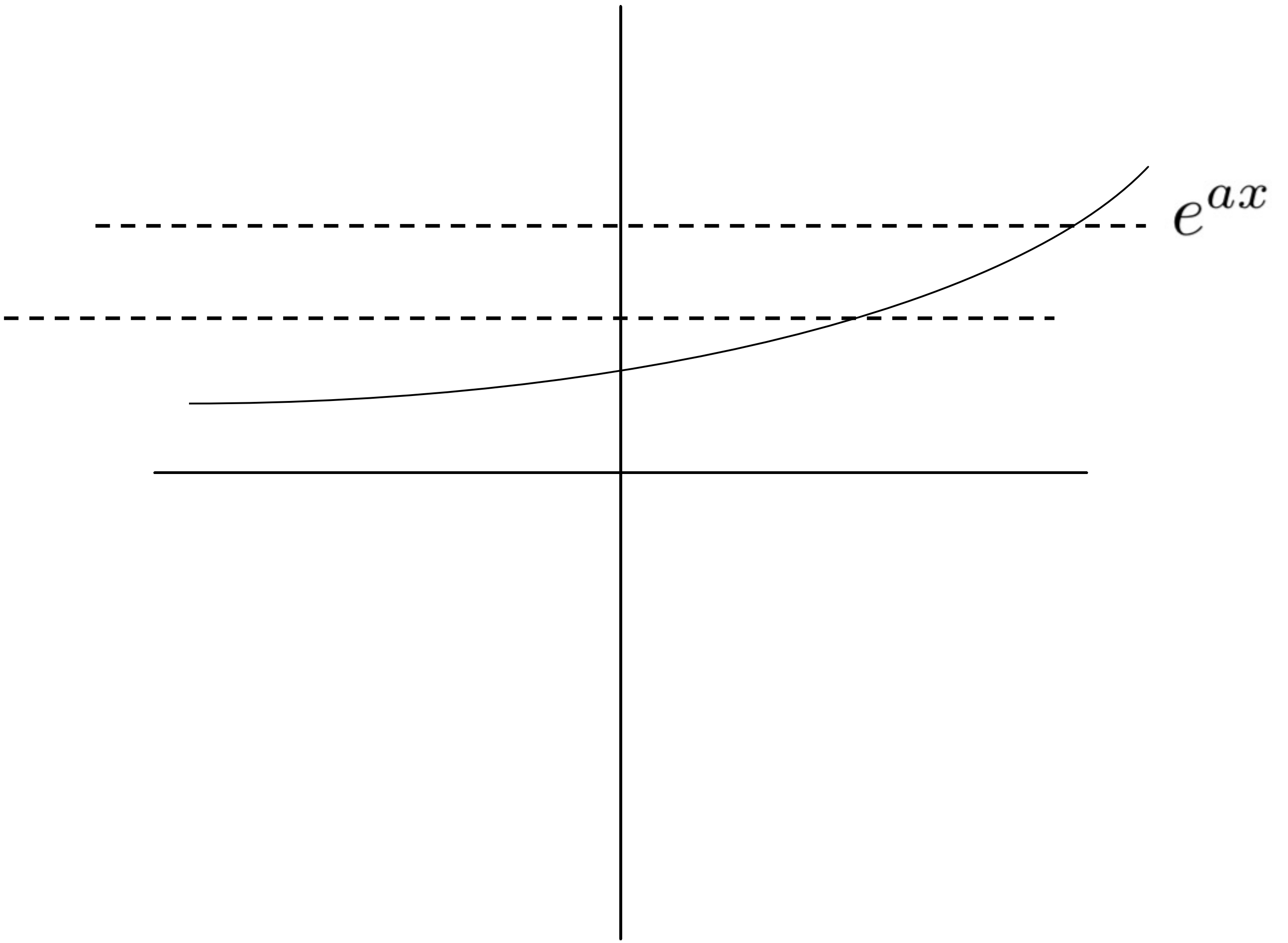
下水平集



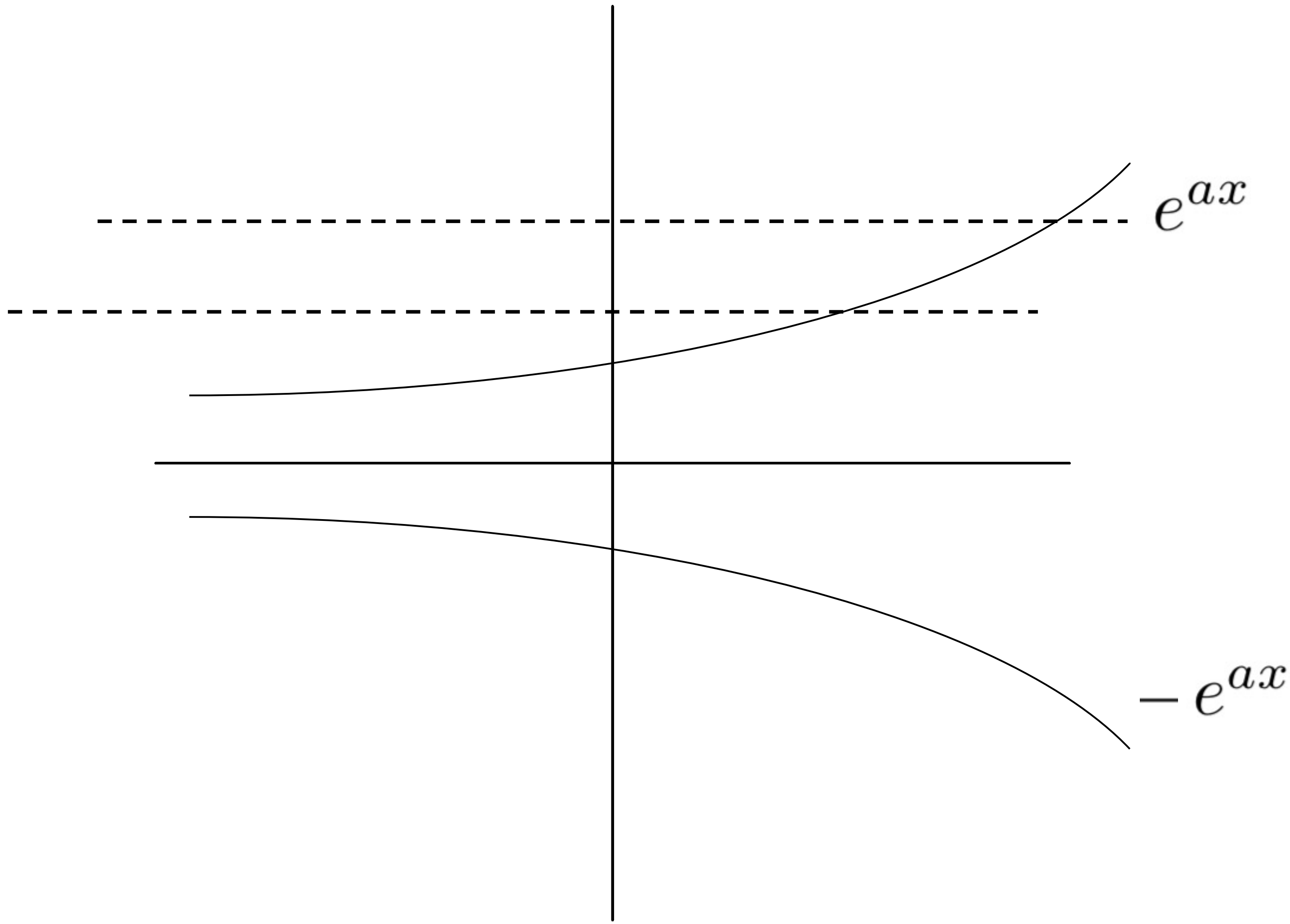
下水平集



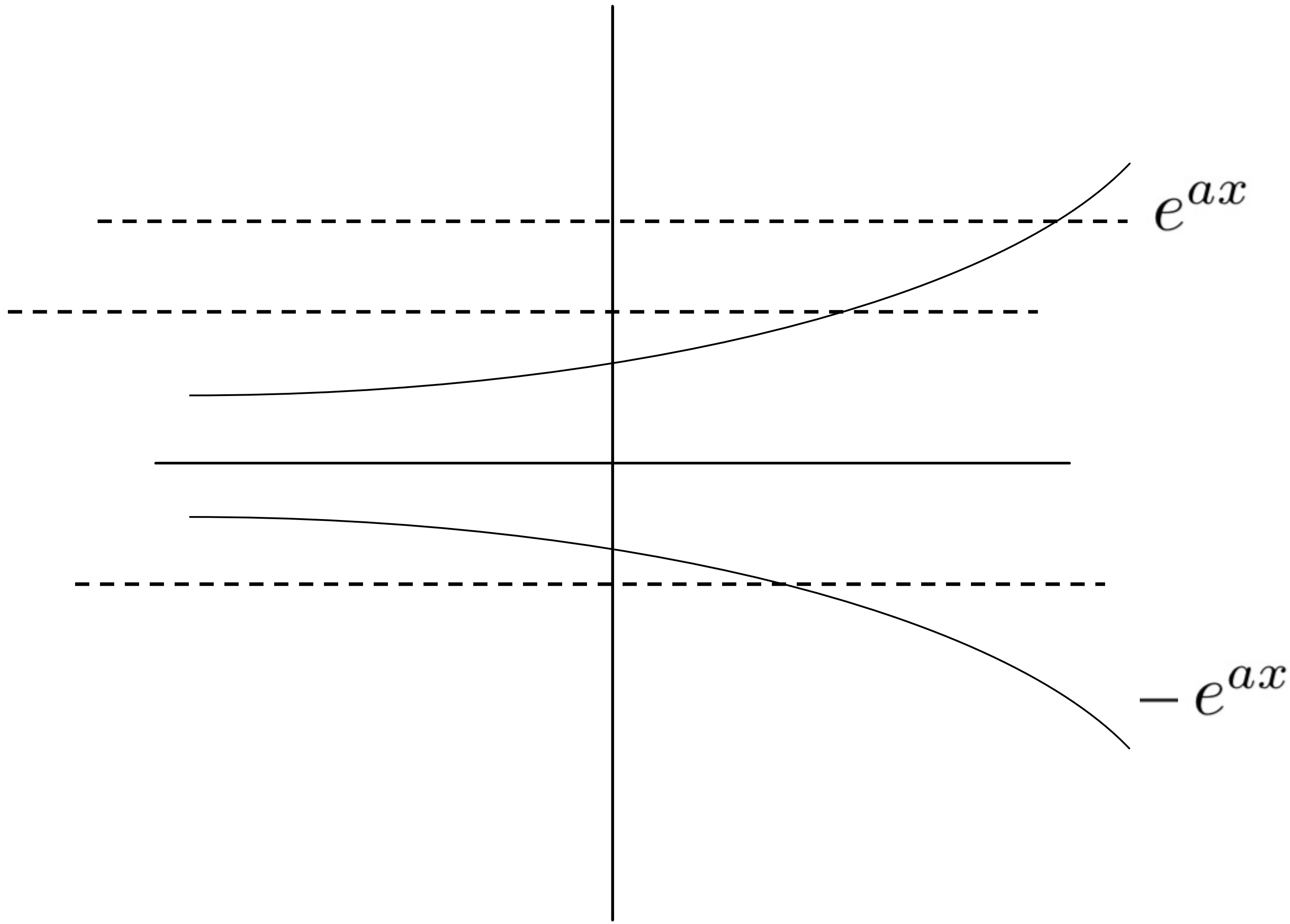
下水平集



下水平集



下水平集



上　图

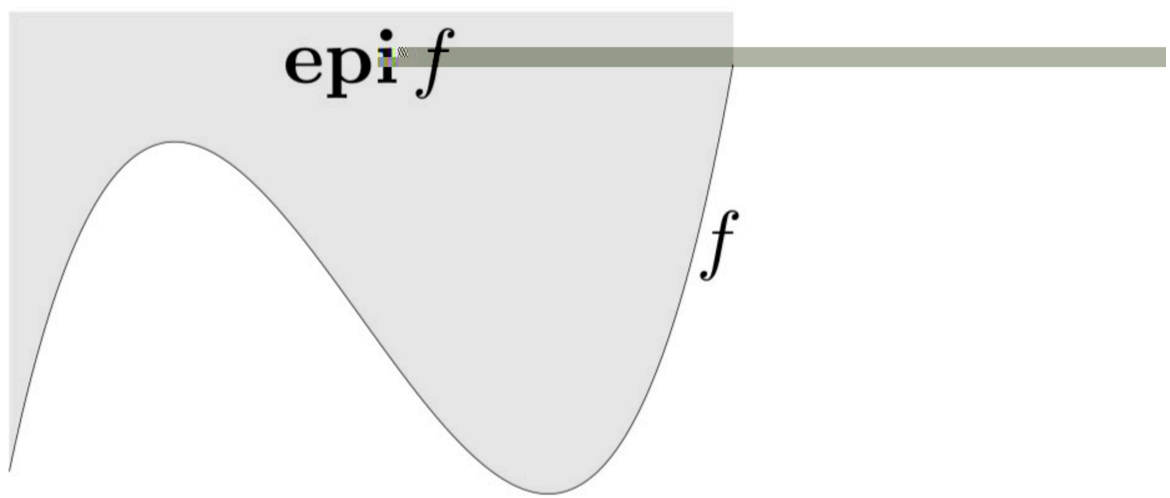
上　图

- 数 $f:\mathbf{R}^n\to\mathbf{R}$的上　图

上 图

- 数 $f : \mathbf{R}^n \to \mathbf{R}$ 的上 图

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



共　　数

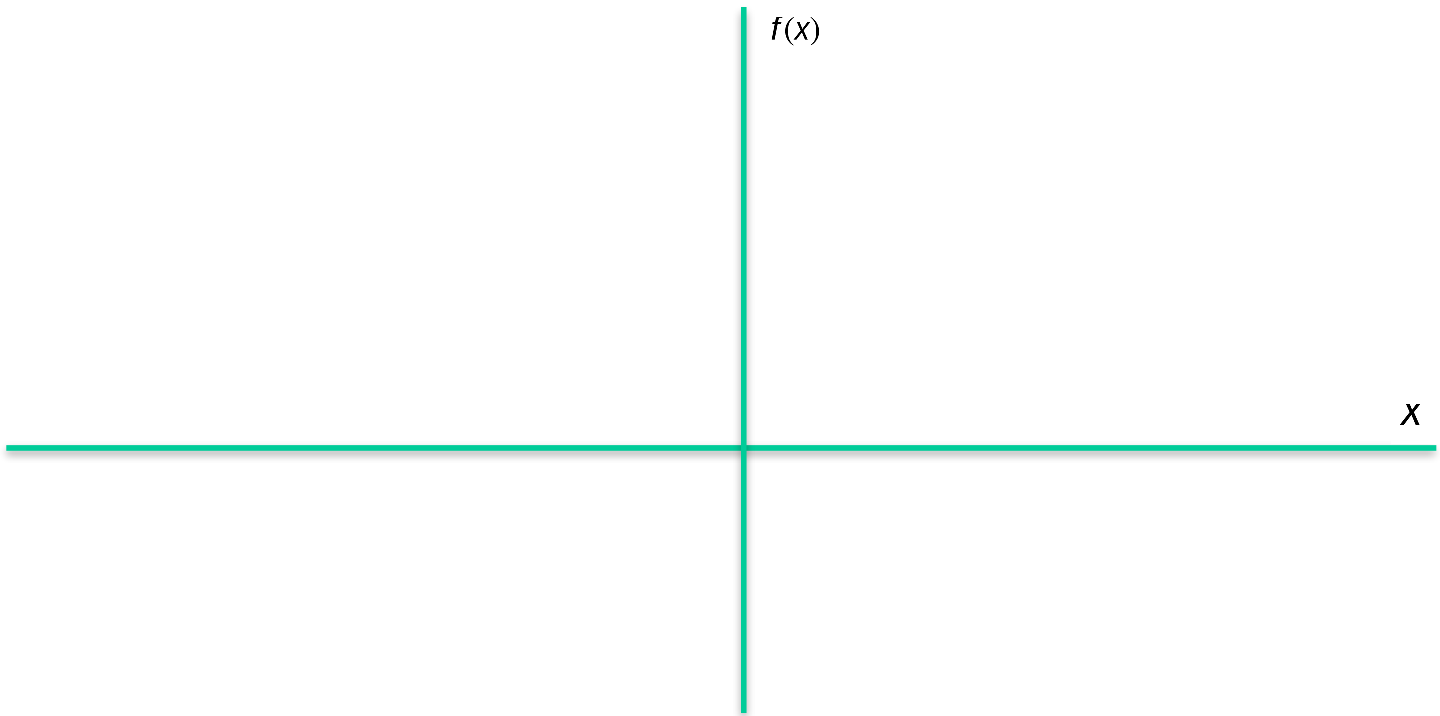
共　　数

- 数f的共　　数为

共　　数

- 数f的共　　数为

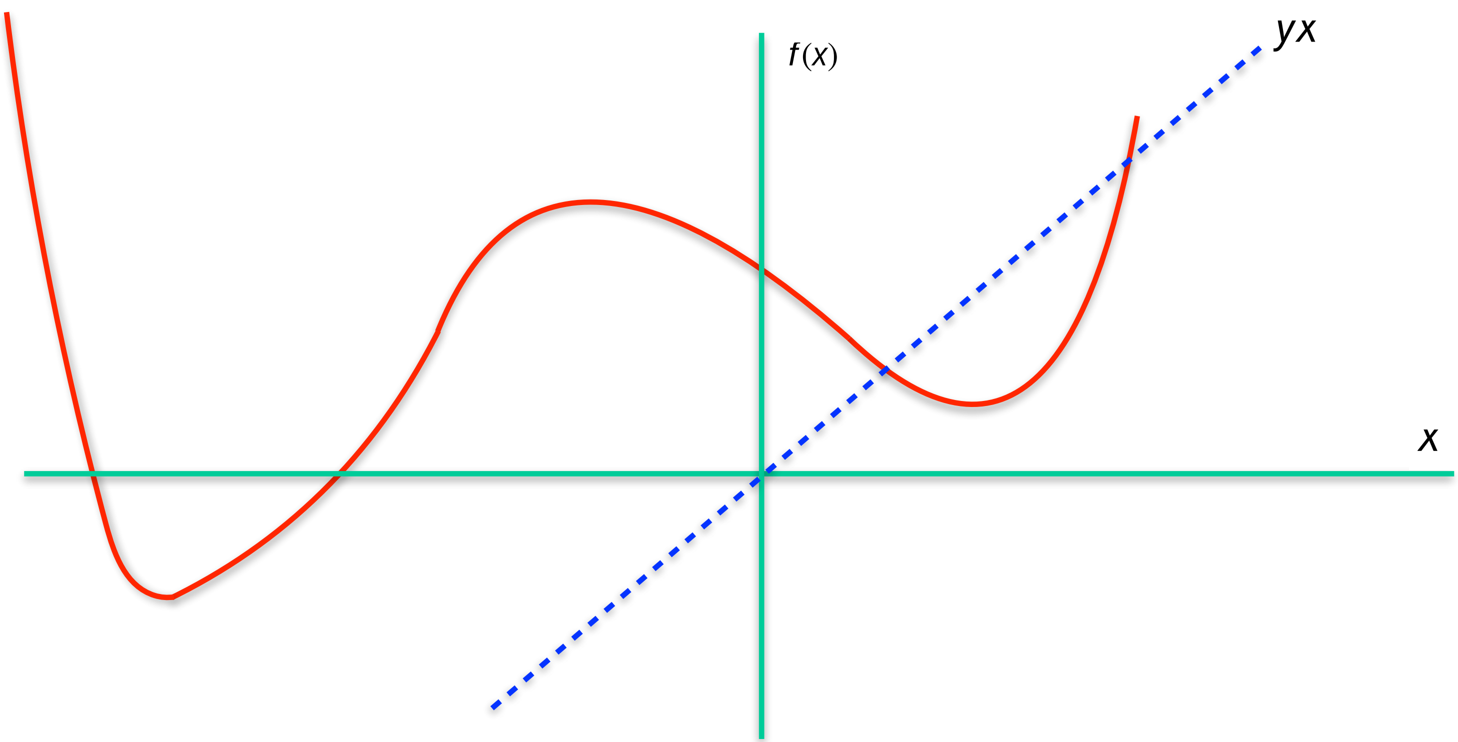
$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} \left(y^T x - f(x) \right)$$



共　　数

- 数f的共　　数为

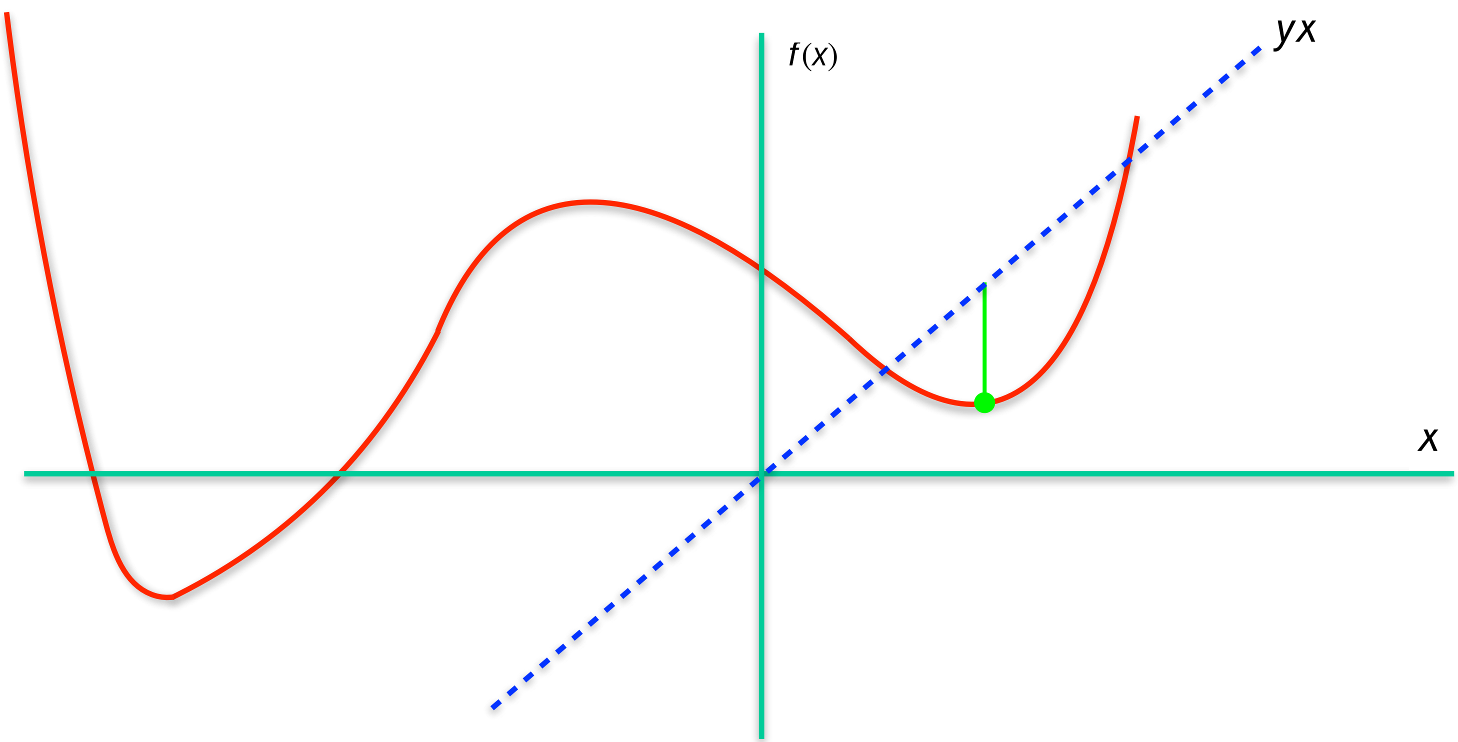
$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} \left(y^T x - f(x)\right)$$



共　　数

- 数f的共　　数为

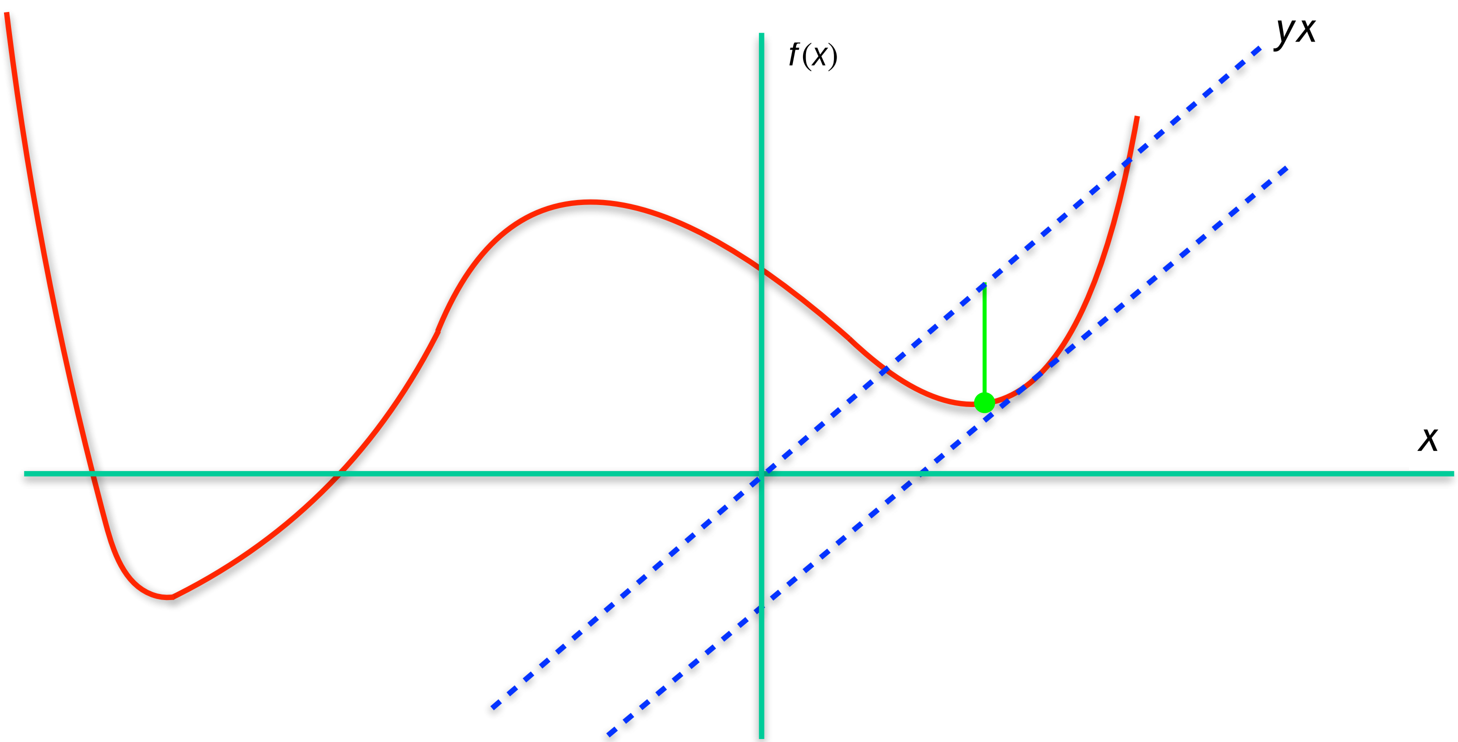
$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} \left(y^T x - f(x) \right)$$



共　　数

- 数f的共　　数为

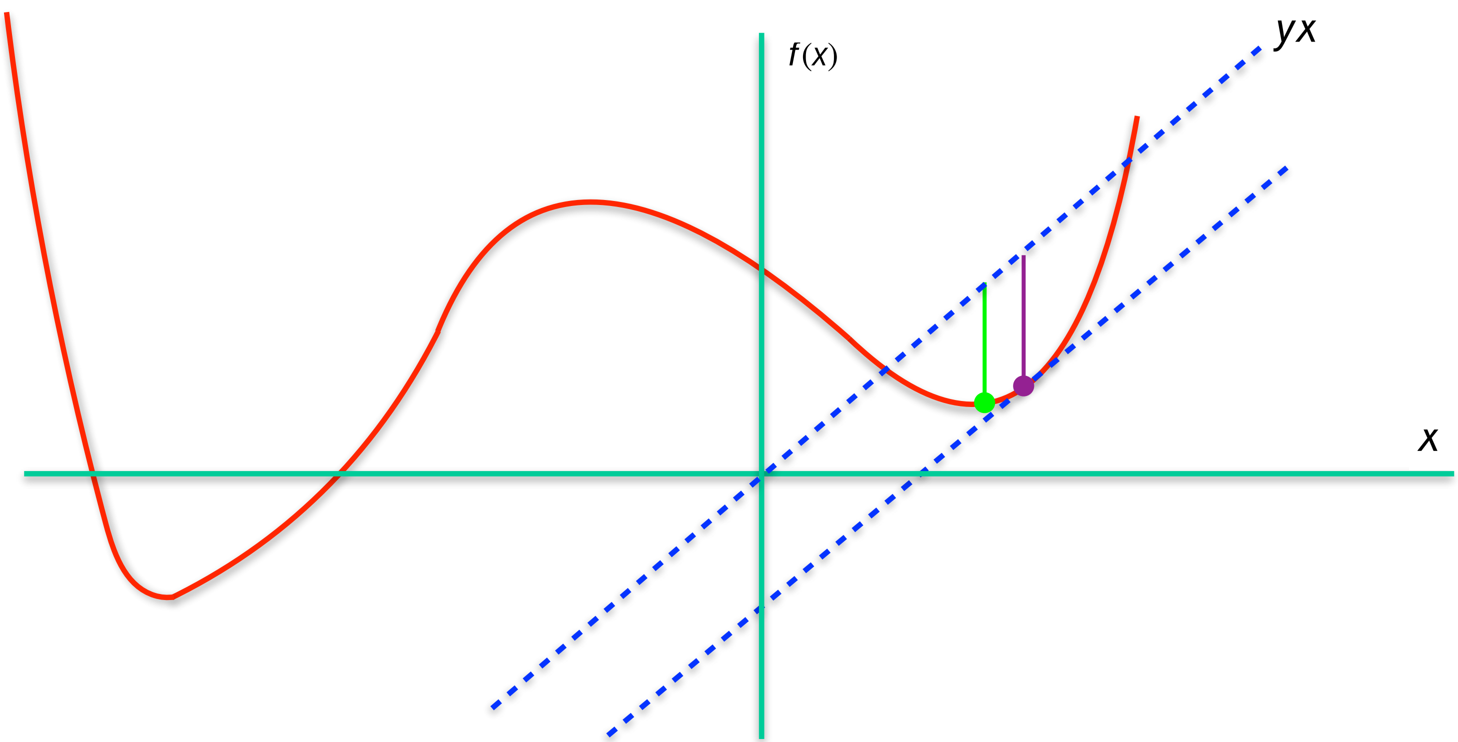
$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} \left(y^T x - f(x) \right)$$



共　　数

- 数f的共　　数为

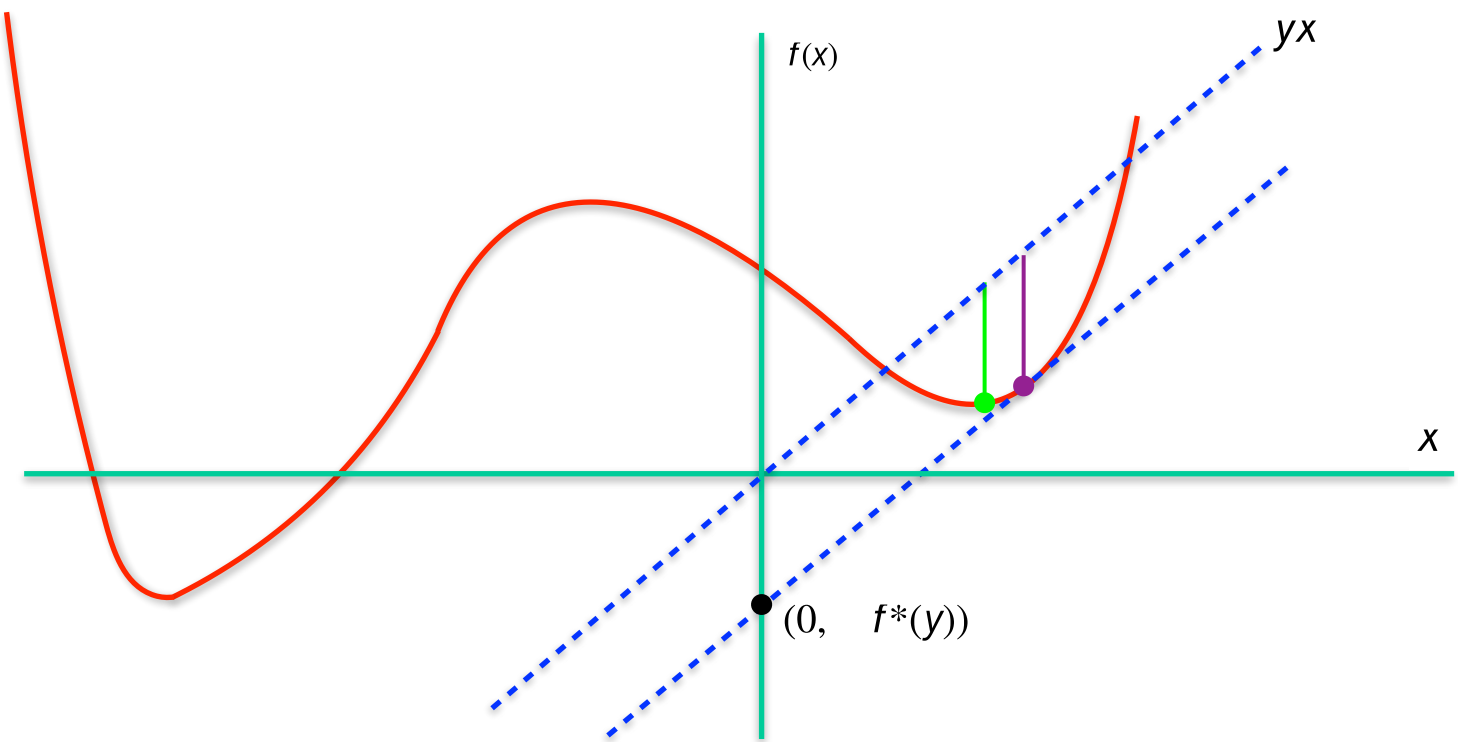
$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} \left(y^T x - f(x)\right)$$



共　　数

- 数f的共　　数为

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



共 数

共　　数

- 若f可微　则$f^*(y)$对应的 必满足$f'(x) = y$

共　　数

- 若f可微　则$f^*(y)$对应的　必满足$f'(x) = y$
- 　数f^*为　　数
  - 即使　数f不是　　数

共　　数

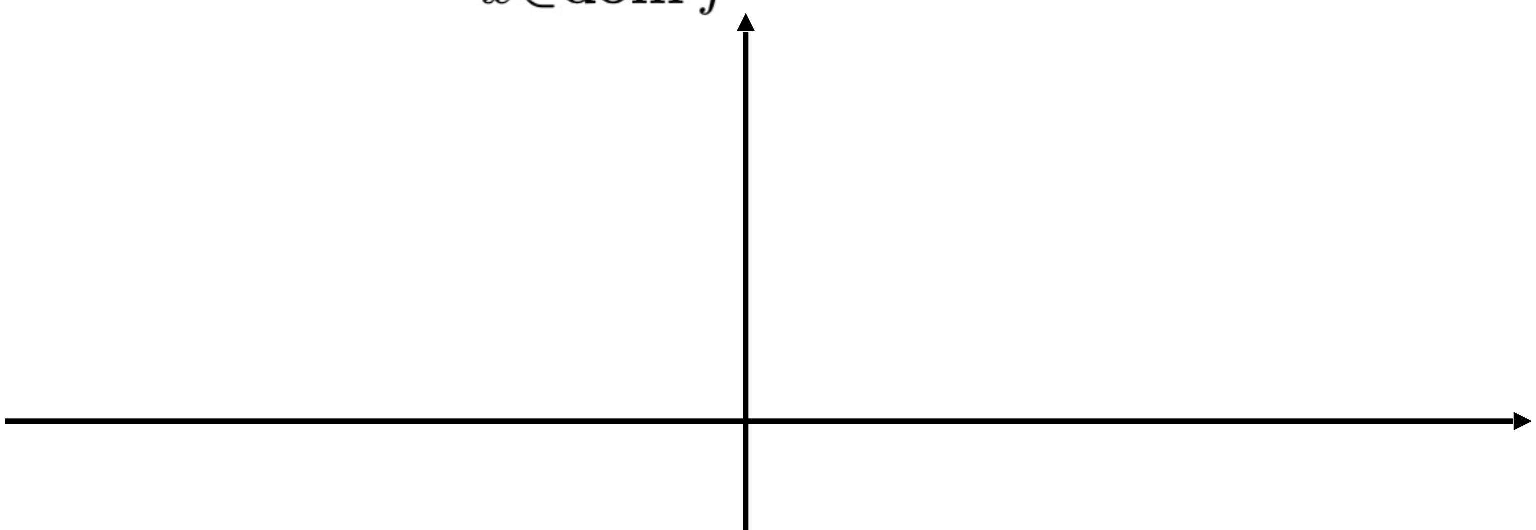
- 若f可微　则$f^*(y)$对应的 必满足$f'(x) = y$
- 　数f^*为　　　数
  - 即使　数f不是　　　数

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

共 数

- 若f可微 则$f^*(y)$对应的 必满足$f'(x) = y$
- 数f^*为 数
  - 即使 数f不是 数

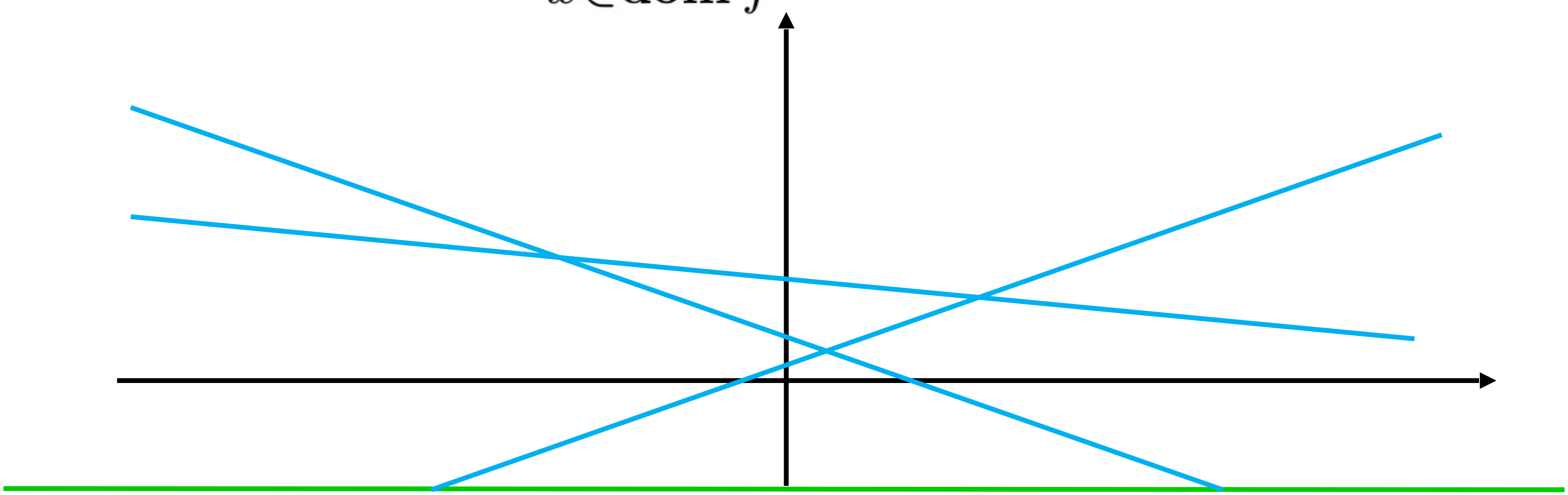
$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} (y^T x - f(x))$$



共　　数

- 若f可微　则$f^*(y)$对应的 必满足$f'(x) = y$
- 数f^*为 　数
  - 即使 数f不是 　数

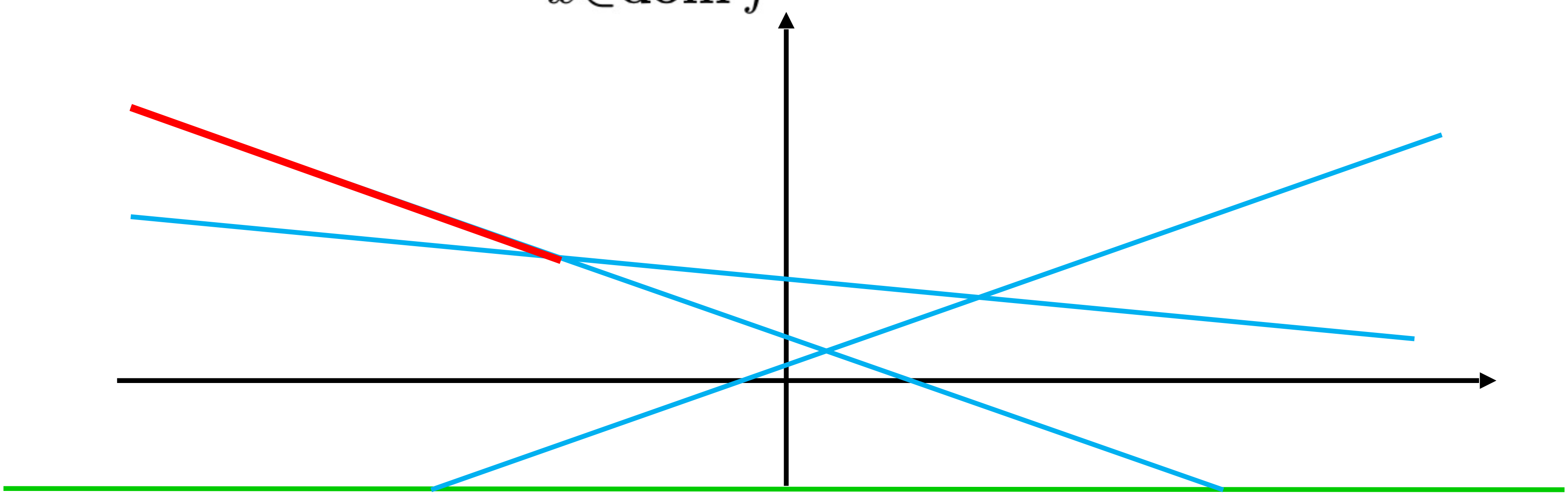
$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} \left(y^T x - f(x) \right)$$



共　　数

- 若f可微　则$f^*(y)$对应的　必满足$f'(x) = y$
- 　数f^*为　　　数
  - 即使　数f不是　　　数

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



共　　数

- 例　$f(x) = ax + b$的共　　数

共　　数

- 例　$f(x) = ax + b$的共　　数
  - $f^*(y) = \sup(yx \quad (ax + b))$

共　　数

- 例　$f(x) = ax + b$ 的共　　数
  - $f^*(y) = \sup(yx \quad (ax + b))$
    $= \sup((y \quad a)x \quad b)$

共　　数

- 例　$f(x) = ax + b$ 的共　　数
  - $f^*(y) = \sup(yx \quad (ax+b))$

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

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

共 数

- 例 的共 数
  - $f^*(y) = \sup(yx \quad (ax + b))$

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

    $=$

共　　数

- 例　$f(x) = ax + b$ 的共　　数
  - $f^*(y) = \sup(yx \quad (ax + b))$

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

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

- 例　负对数 $f(x) = -\log x$

  $f^*(y) = \sup\limits_{x>0}$ [redacted]

共　　数

- 例　$f(x) = ax + b$ 的共　　数
  - $f^*(y) = \sup(yx \quad (ax+b))$

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

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

- 例　负对数 $f(x) = -\log x$

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

$y + 1/x = 0$

$x = \quad 1/y$

共　　数

- 例　$f(x) = ax + b$ 的共　　数
  - $f^*(y) = \sup(yx \quad (ax + b))$

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

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

- 例　负对数 $f(x) = -\log x$

$y + 1/x = 0$

$$f^*(y) = \sup_{x>0}(xy + \log x) \qquad x = \quad 1/y$$

$$= \begin{cases} -1 - \log(-y) & y < 0 \\ \infty & \text{otherwise} \end{cases}$$

共　　数

共 数

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

共　　数

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

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

共 数

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

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

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

共　　数

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

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

- $\dfrac{(y^T x \quad (1/2)x^T Q x)}{x} = y \quad Qx = 0$
- $x = Q^{\;\;1} y$

共 数

- 二次 数

$(y$

共　　数

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

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

- $\dfrac{(y^Tx \quad (1/2)x^TQx)}{x} = y \quad Qx = 0$
- $x = Q^{\ 1}y$
- $f^*(y) = y^TQ^{\ 1}y \quad (1/2)y^T(Q^{\ 1})^TQQ^{\ 1}y$

$$= \quad \frac{1}{2}y^TQ^{-1}y$$

共　　数的性质

共　　数的性质

- 复数的共

共　　数的性质

- 复数的共
  - $(a+bi)^* = a-bi$

共　　数的性质

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

共　数的性质

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

共　数的性质

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

共　　数的性质

- 复数的共
  - $(a+bi)^* = a-bi$
  - $(a-bi)^* = a+bi$
- 　数的共
  - $f^{**}=f$　？
  - 若f为　　且为闭　数　则$f^{**}=f$

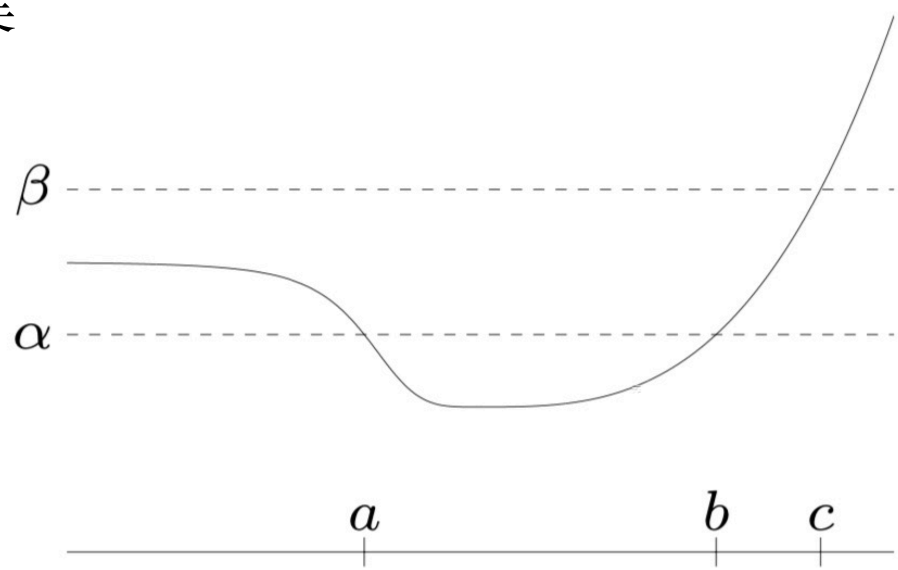
数

数

- 数. $f : \mathbf{R}^n \to \mathbf{R}$ 为　　　数　当其定义域为　集　且其下水平集 $S_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \le \alpha\}$ 对所有 α 为　集

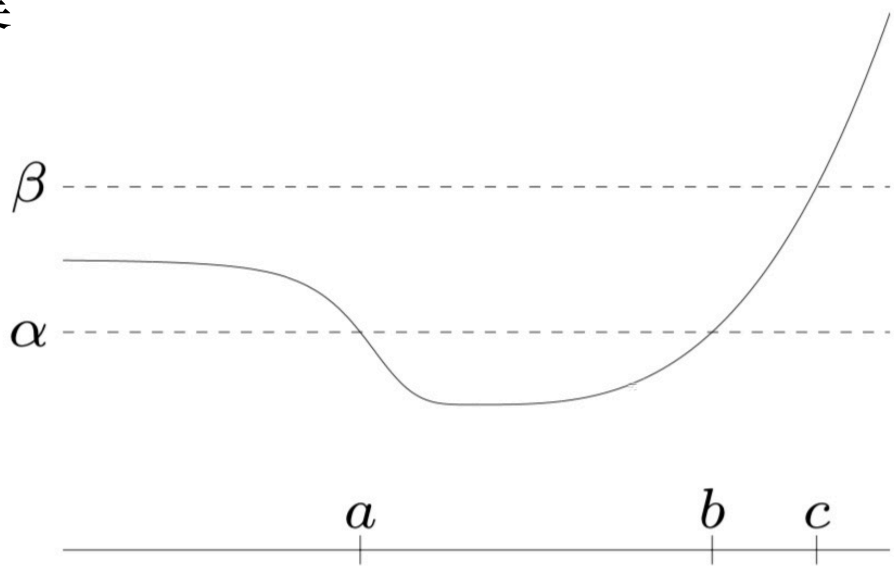
数

- 数. $f:\mathbf{R}^n \to \mathbf{R}$ 为 数 当其定义域为 集 且其下水平集 $S_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \le \alpha\}$ 对所有 α 为 集



数

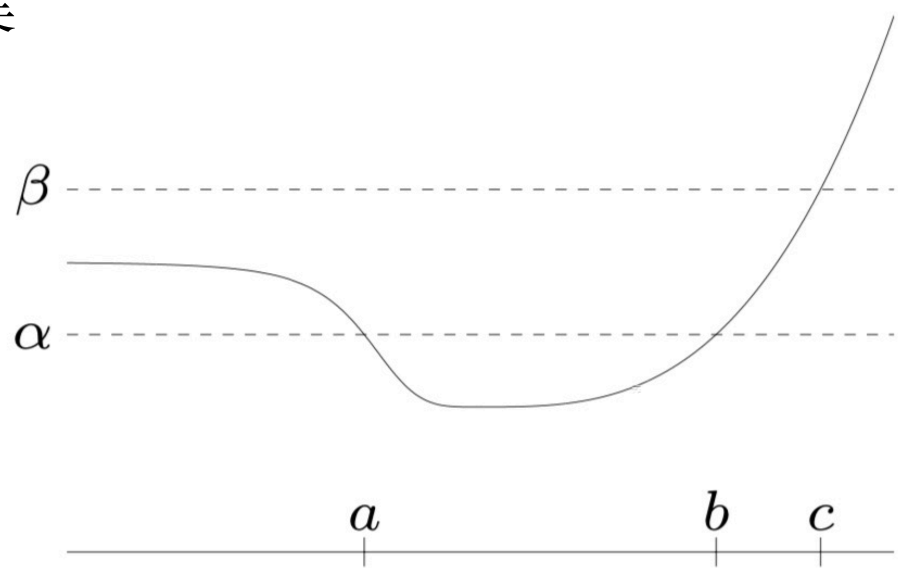
- 数. $f:\mathbf{R}^n \to \mathbf{R}$ 为 数 当其定义域为 集 且其下水平集 $S_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \le \alpha\}$ 对所有 α 为 集



- 数f为 数 若$-f$为 数

数

- 数. $f : \mathbf{R}^n \to \mathbf{R}$ 为 数 当其定义域为 集 且其下水平集 $S_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \le \alpha\}$ 对所有 α 为 集



- 数 f 为 数 若 $-f$ 为 数
- 数 f 为 线性 数 若同时为 和 数

数

数

- 数为 数

数

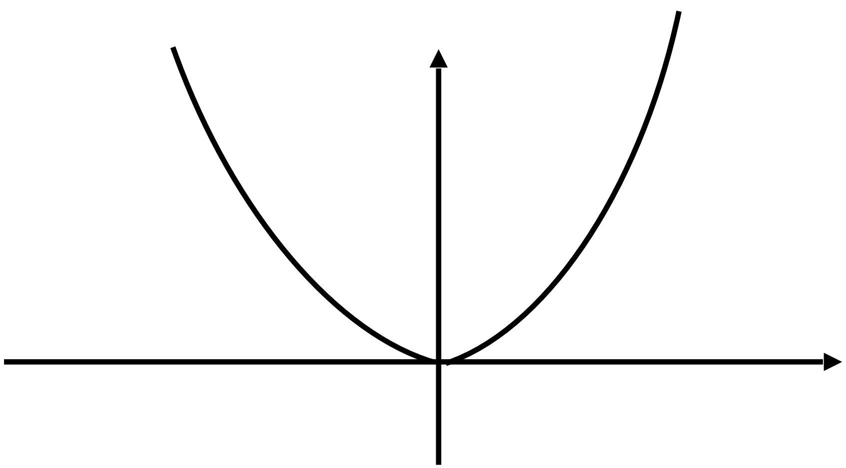
- 数为 数
- 数不一定为 数

数

- 数为 数
- 数不一定为 数
- 数为单模态 数

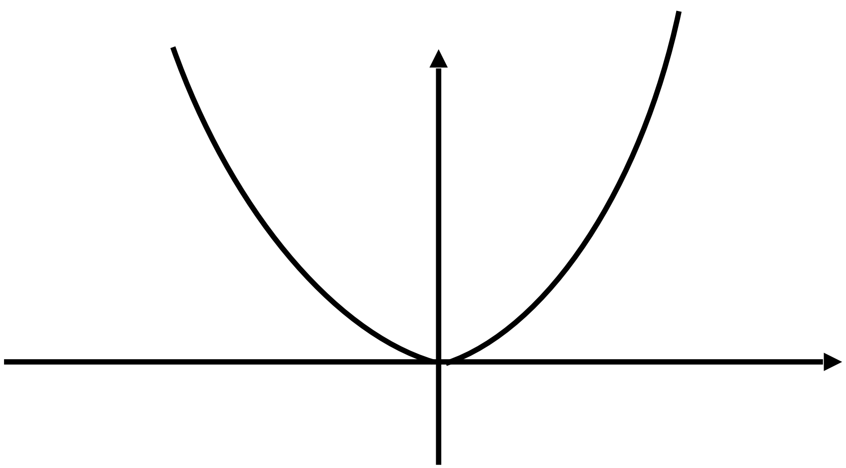
数

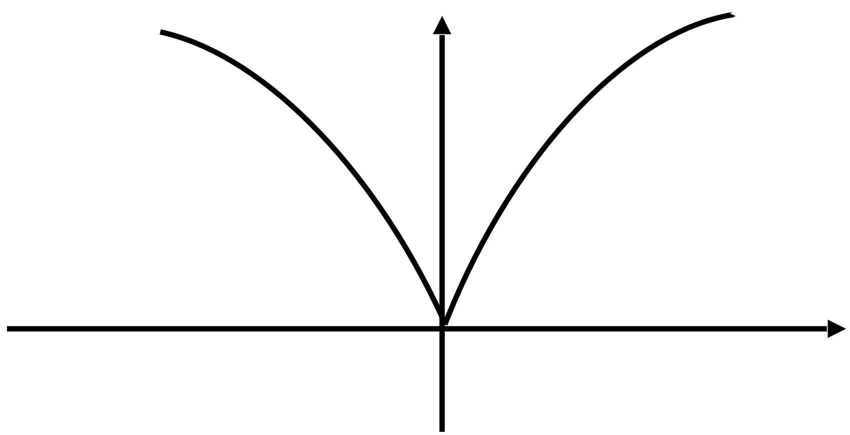
- 数为 数
- 数不一定为 数
- 数为单模态 数



数

- 数为 数
- 数不一定为 数
- 数为单模态 数





数

数

- 数 定义域为 集 且

数

- 数 定义域为 集 且
$$f(\theta x+(1-\theta)y)\le \theta f(x)+(1-\theta)f(y)$$

数

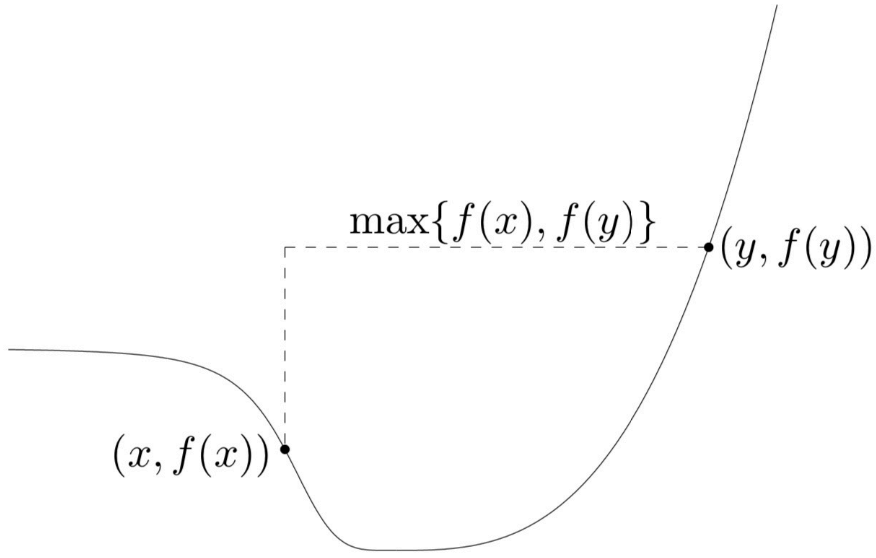
- 数 定义域为 集 且
$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$$
- 数 定义域为 集 且

数

- 数 定义域为 集 且
$$f(\theta x+(1-\theta)y)\le \theta f(x)+(1-\theta)f(y)$$
- 数 定义域为 集 且
$$f(\theta x+(1-\theta)y)\le \max\{f(x),f(y)\}$$

数

- 数 定义域为 集 且
$$f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta)f(y)$$
- 数 定义域为 集 且
$$f(\theta x + (1-\theta)y) \le \max\{f(x), f(y)\}$$



数

数

- 向量的长度　向量中最后一个非零元素的位置

数

- 向量的长度　向量中最后一个非零元素的位置

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

数

- 向量的长度　向量中最后一个非零元素的位置

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

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

数

- 向量的长度　向量中最后一个非零元素的位置

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

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

- 线性分数　数

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　数

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

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　数

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

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

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　数

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

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

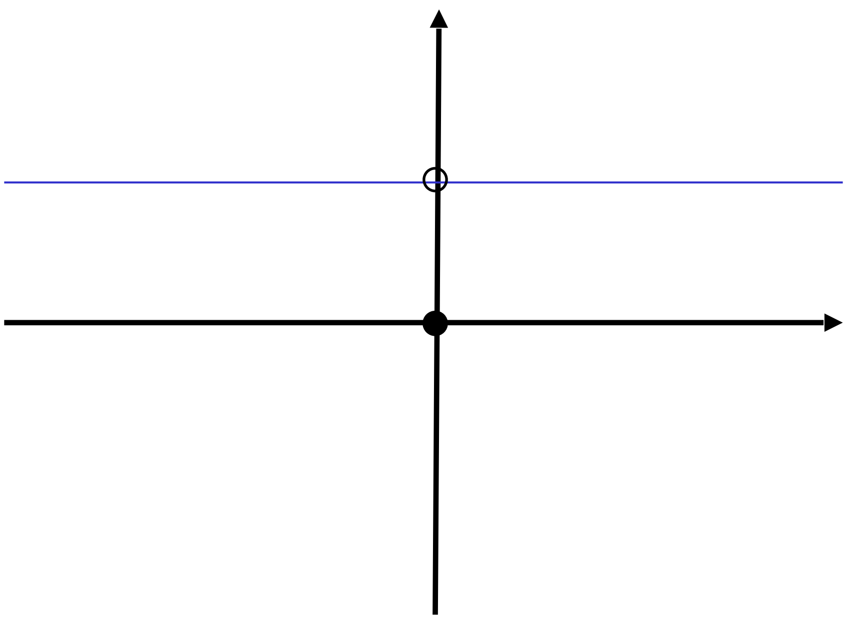
向量的零范数

向量的零范数

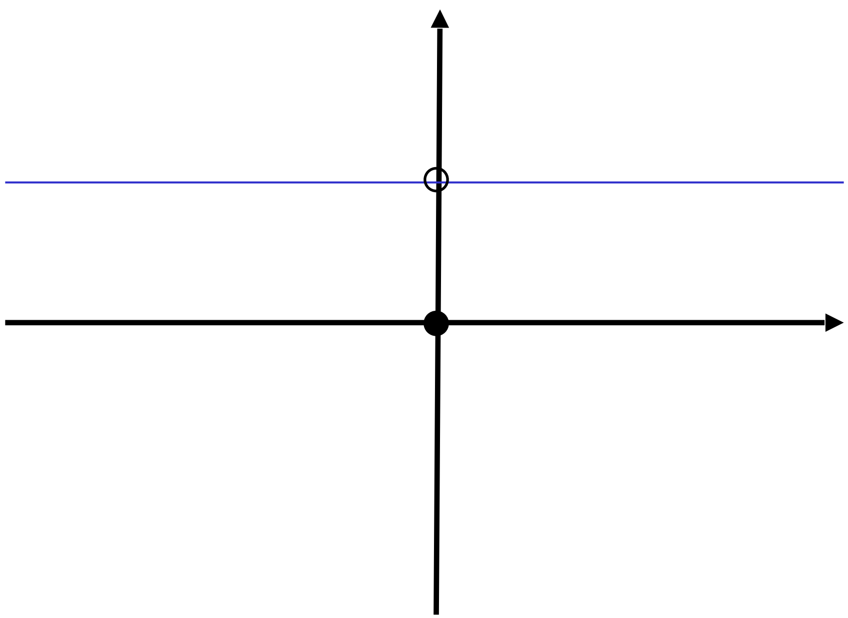
- 非零元素的个数
  - $f(x) = x_{0}$

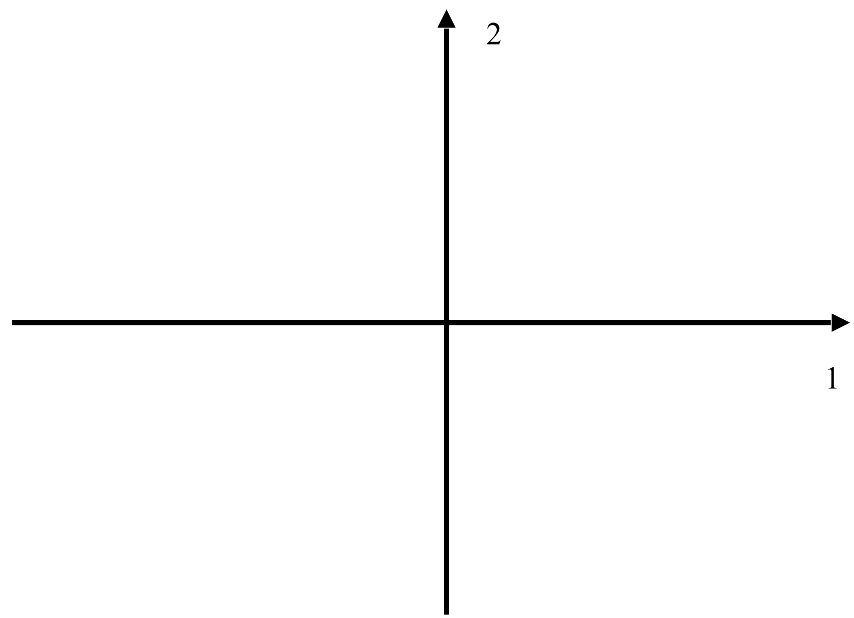
向量的零范数

- 非零元素的个数
  - $f(x) = x_{0}$



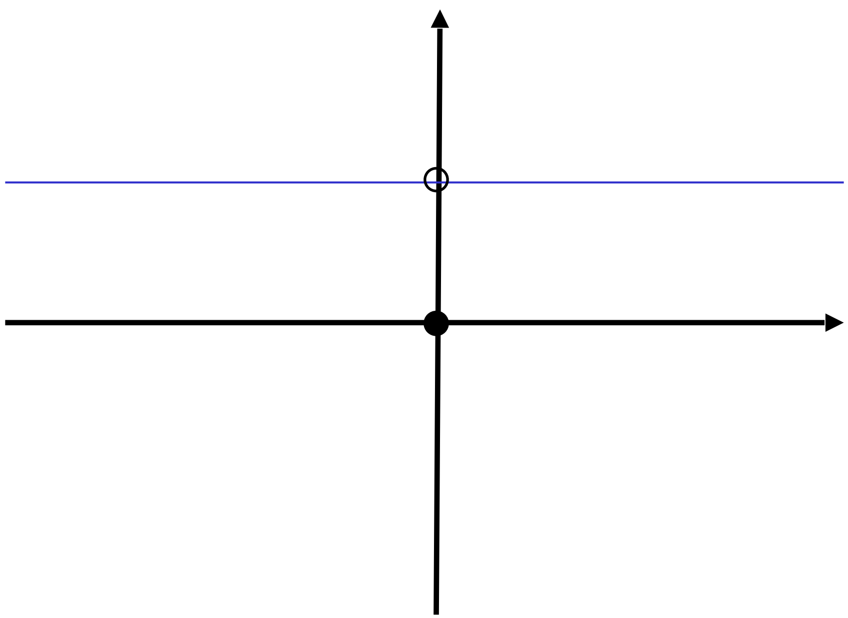
- 非零元素的个数
  - $f(x) = x\ {}_{0}$

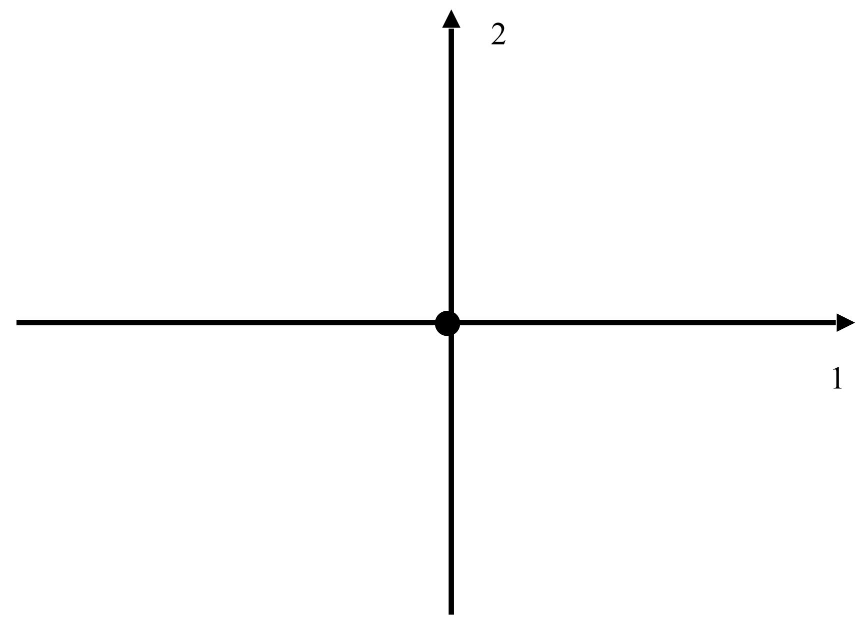




向量的零范数

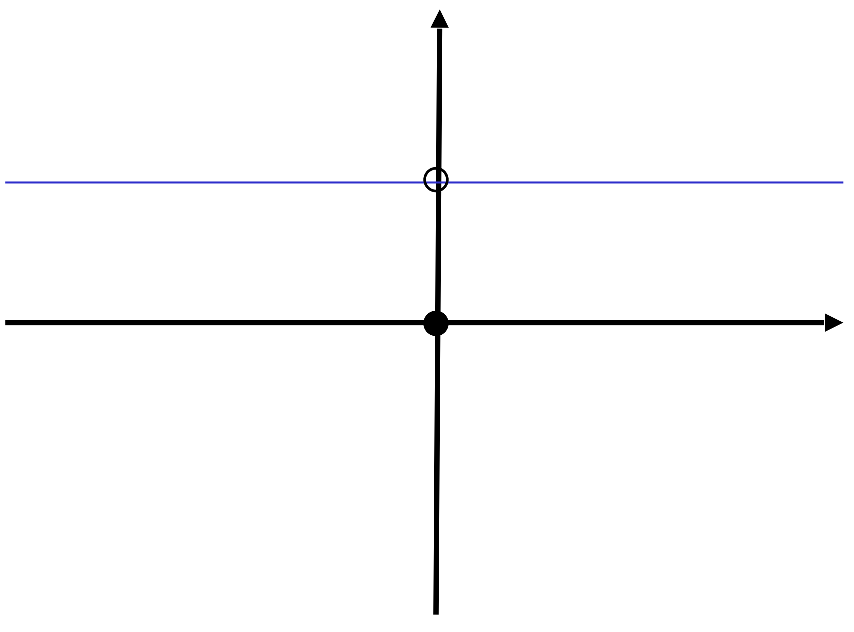
- 非零元素的个数
  - $f(x) = x \;\; _{0}$

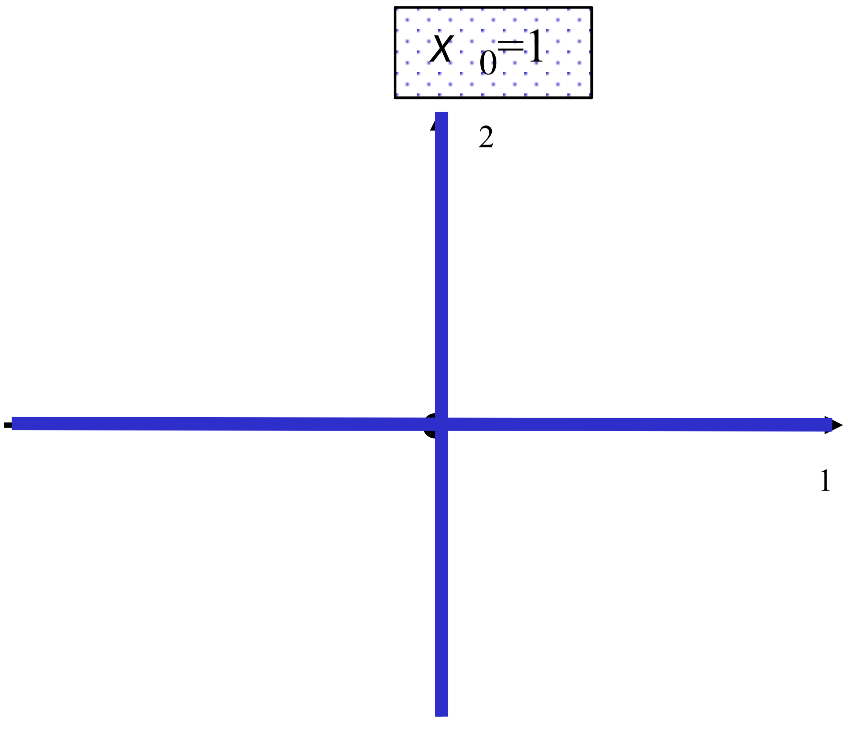




向量的零范数

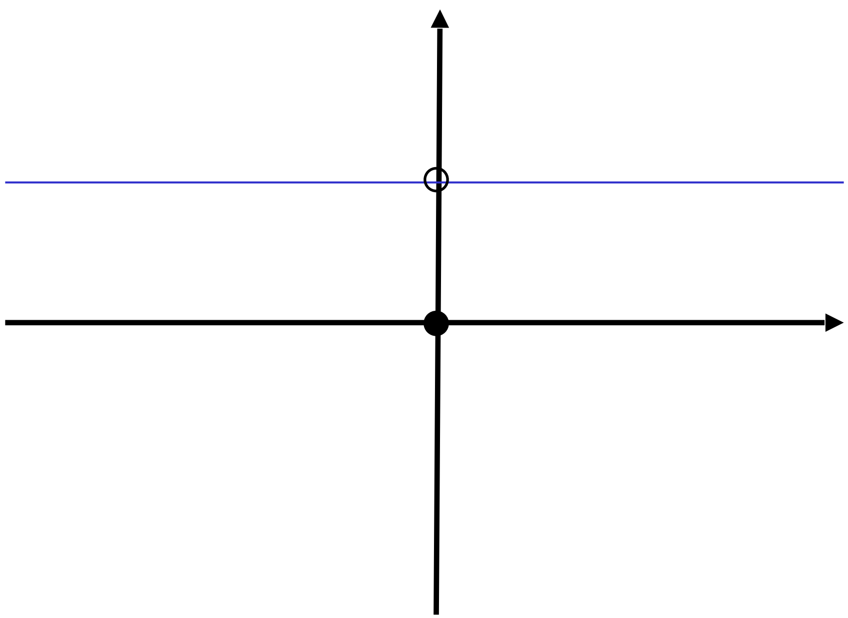
- 非零元素的个数
  - $f(x) = \ x_{0}$

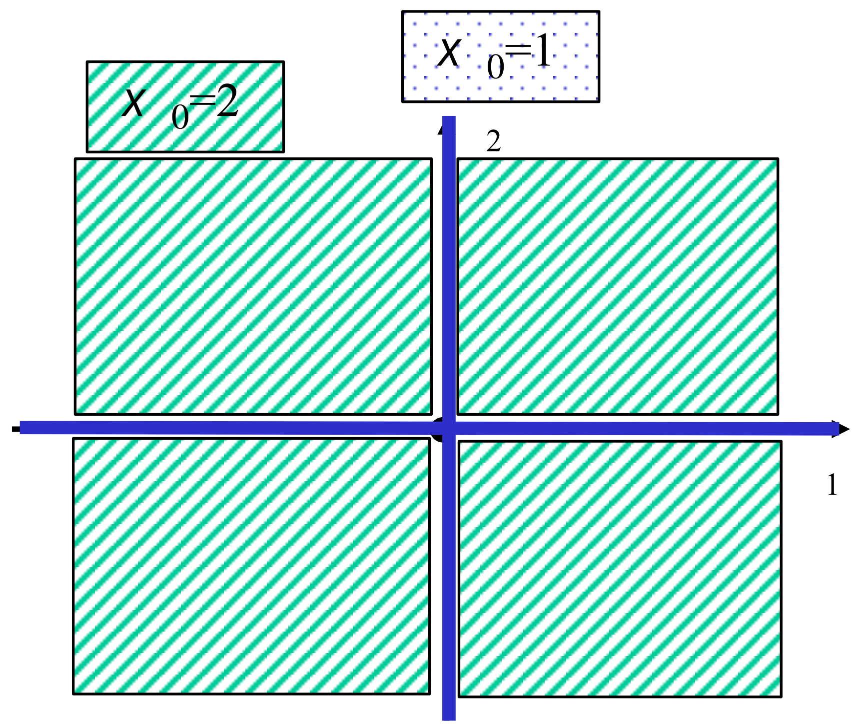




向量的零范数

- 非零元素的个数
  - $f(x) = \|x\|_0$





向量的零范数

向量的零范数

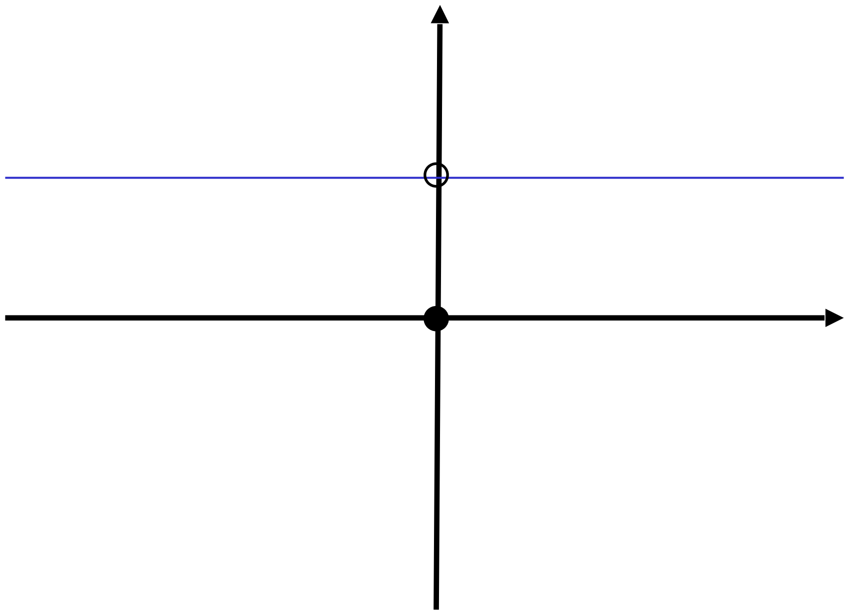
- $\min_{x \in C} \|x\|_0$

向量的零范数

- $\min_{x\ \ C}\quad x\quad {}_{0}$
- $\min_{x\ \ C}\quad x\quad {}_{1}$

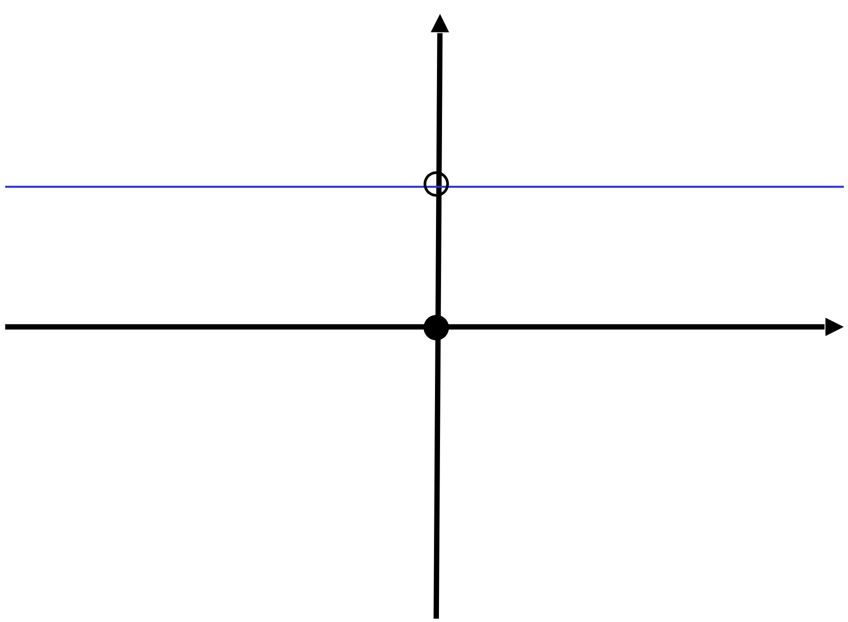
向量的零范数

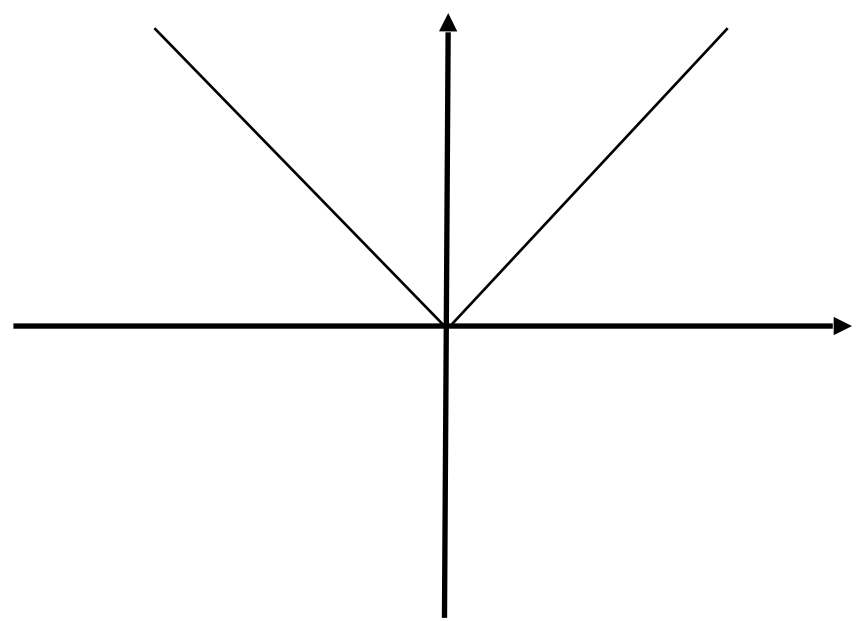
- $\min_{x \quad C} \quad x \quad {}_{0}$
- $\min_{x \quad C} \quad x \quad {}_{1}$



向量的零范数

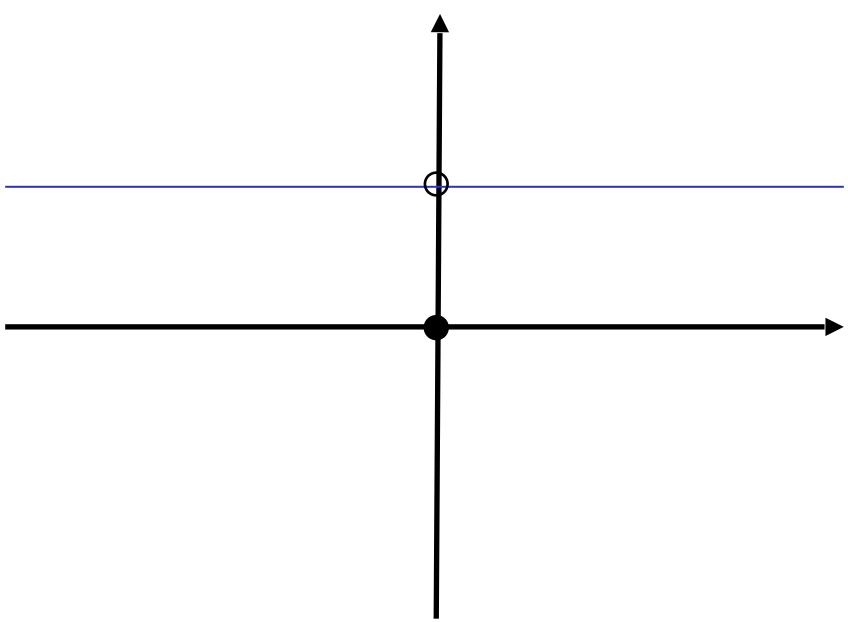
- $\min_{x \quad C} \quad x \quad {}_{0}$
- $\min_{x \quad C} \quad x \quad {}_{1}$





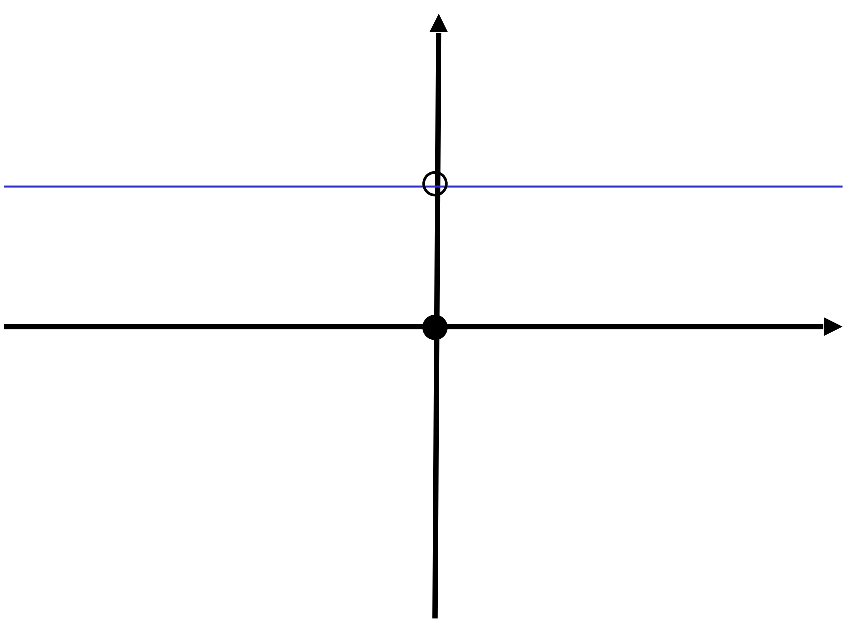
向量的零范数

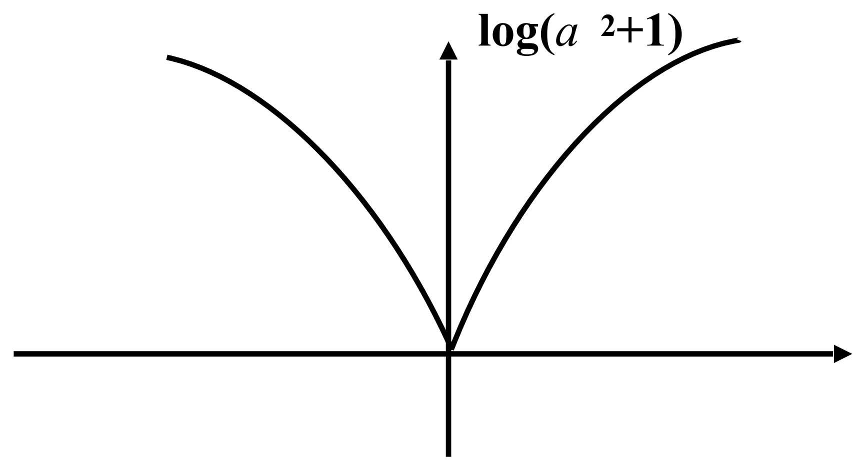
- $\min_{x \quad C} \quad x \quad {}_0$
- $\min_{x \quad C} \quad x \quad {}_1$



向量的零范数

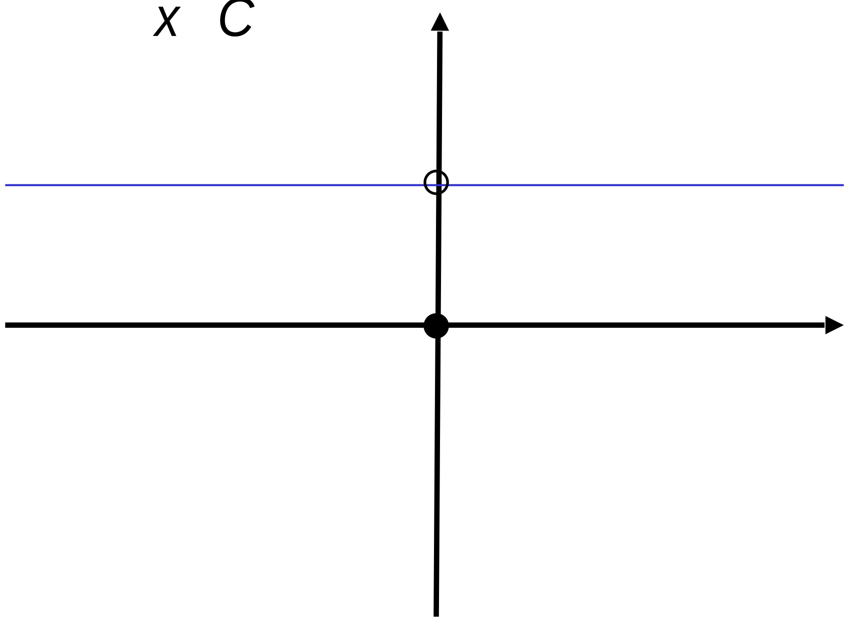
- $\min_{x \quad C} \quad x \quad _0$
- $\min_{x \quad C} \quad x \quad _1$

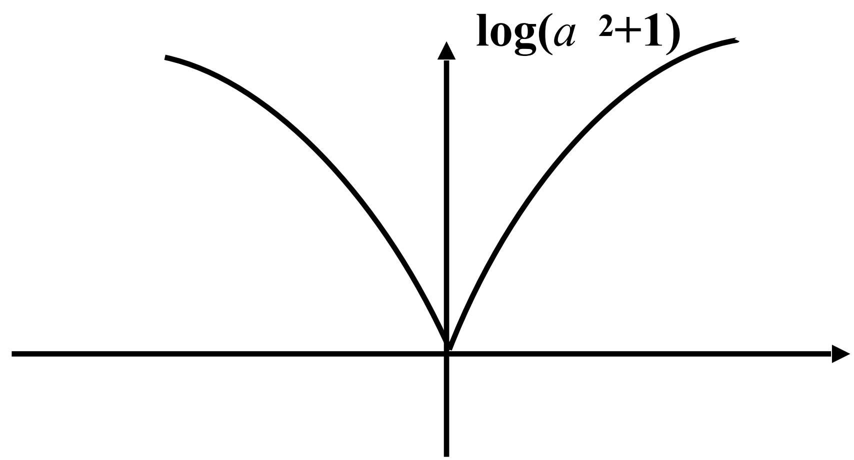




向量的零范数

- $\min_{x \quad C} \quad x \quad {}_0$
- $\min_{x \quad C} \quad x \quad {}_1$
- $\min_{x \quad C} \log(xx^T + 1)$





可微　　　数

可微 数

- 数的一 条件

可微　　数

- 数的一　条件

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

可微 数

- 数的一 条件

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

- 数的一 条件

可微 数

- 数的一 条件

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

- 数的一 条件

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

可微　　　数

- 　　数的一　条件

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

- 　　　数的一　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　　　数

- 　　数的一　条件

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

- 　　数的一　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
  - $0 \leq \theta < 1$

可微　　　数

- 数的一　条件

- 数的一　条件

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

- 证明　仅考虑一维情况
  - $0 \le \theta < 1$

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

可微　　数

- 　　数的一　条件

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

- 　　数的一　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
  - 　$0 \leq \theta < 1$

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

可微　　数

- 数的一　条件

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

- 数的一　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
  - $0 \leq \theta < 1$

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

可微　　　数

- 　　数的一　条件

- 　　　数的一　条件

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

- 证明　仅考虑一维情况
  - 　$0 \le \theta \le 1$

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

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

可微　　数

- 　　数的一　条件

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

- 　　数的一　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
  - 　$0 \leq \theta < 1$

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

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

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

可微 数

- 数的一 条件

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

- 数的一 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
  - $0 \leq \theta < 1$

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

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

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

$$f'(x)(y \quad x) \quad 0$$

可微　　　数

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

可微函数

- 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 数

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

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

- 设$x < y$

可微　　　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　　$[x, y]$　满足$f(z) > \max\{f(x), f(y)\}$

可微 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 $[x, y]$ 满足$f(z) > \max\{f(x), f(y)\}$
- 令$\bar{z} = \inf\{\tilde{z}$ $[x, y] : f(\tilde{z}) = f(z)\}$

可微　　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$z \quad [x,y] \quad$满足$f(z) > \max\{f(x), f(y)\}$
- 令$\bar{z} = \inf\{\tilde{z} \quad [x,y] : f(\tilde{z}) = f(z)\}$
- 存在$\epsilon > 0$

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

可微 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 $z \quad [x, y] \quad$ 满足 $f(z) > \max\{f(x), f(y)\}$
- 令 $\bar{z} = \inf\{\tilde{z} \quad [x, y] : f(\tilde{z}) = f(z)\}$
- 存在 $\epsilon > 0$

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

可微 数

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

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

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

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

可微　数

- 证明

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

数

数

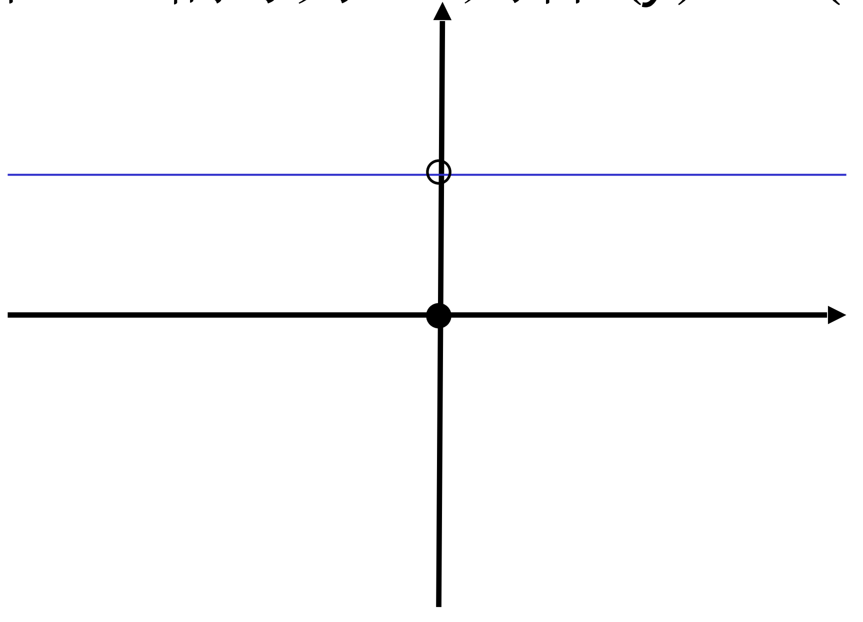
- 数
  - 若一　偏导为**0**　则$f(y)$　$f(x)$

数

- 数
  - 若一 偏导为**0** 则$f(y)$ $f(x)$
- 数
  - 若一 偏导为**0** 则若$f(y)$ $f(x)$ 则 0 0

数

- 数
  - 若一 偏导为**0** 则$f(y)$ $f(x)$
- 数
  - 若一 偏导为**0** 则若$f(y)$ $f(x)$ 则 0 0



数的二　条件

数的二　条件

- 数的二　条件

- 数的二　条件

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

数的二　条件

- 数的二　条件

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

- 数的二　条件

数的二　条件

- 数的二　条件

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

- 数的二　条件

  $$y^T \nabla f(x) = 0 \Longrightarrow y^T \nabla^2 f(x) y \geq 0$$

数的二　条件

- 数的二　条件

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

- 数的二　条件

$$y^T \nabla f(x) = 0 \Longrightarrow y^T \nabla^2 f(x) y \geq 0$$

- 一维　数下
  - $yf'(x) = 0 \rightarrow y^2 f''(x) \quad 0$

数的二　条件

- 　　　数的二　条件

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

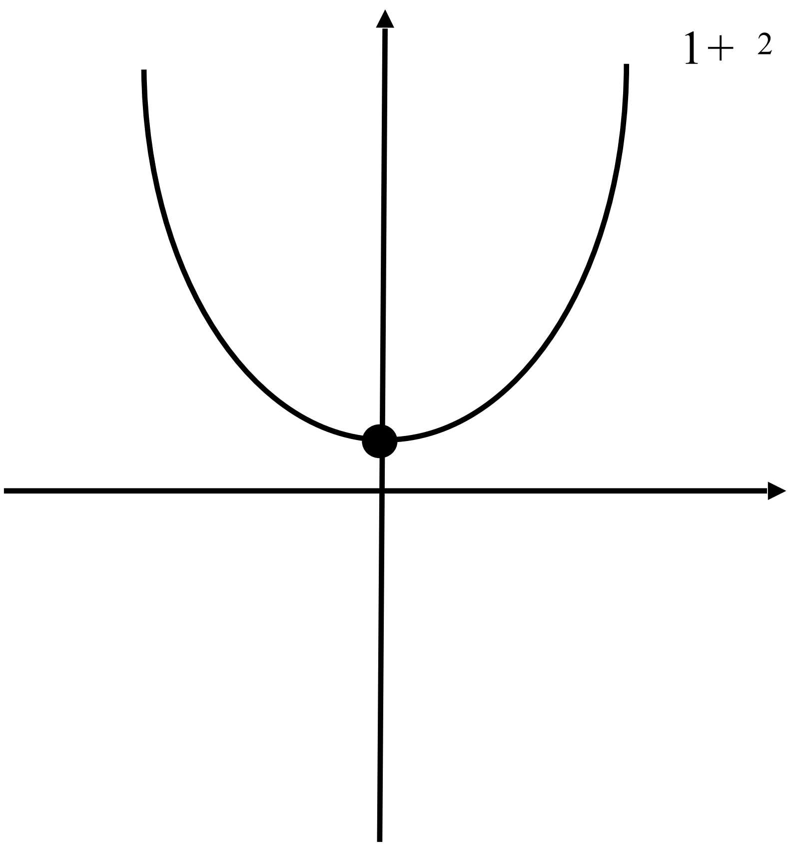
- 　　　　数的二　条件

$$y^T \nabla f(x) = 0 \Longrightarrow y^T \nabla^2 f(x) y \geq 0$$

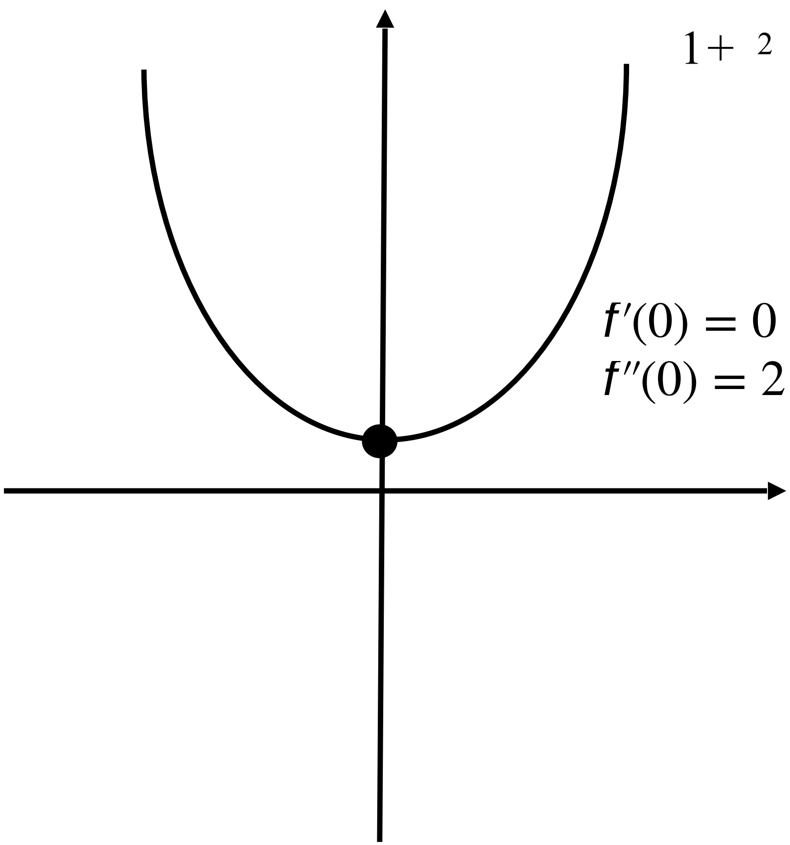
- 一维　数下
  - $yf'(x) = 0 \to y^2 f''(x) \quad 0$
  - $\begin{cases} y = 0 & \quad 0 \quad 0 \\ f'(x) = 0, y \neq 0 & f''(x) \quad 0 \end{cases}$

数的二　条件

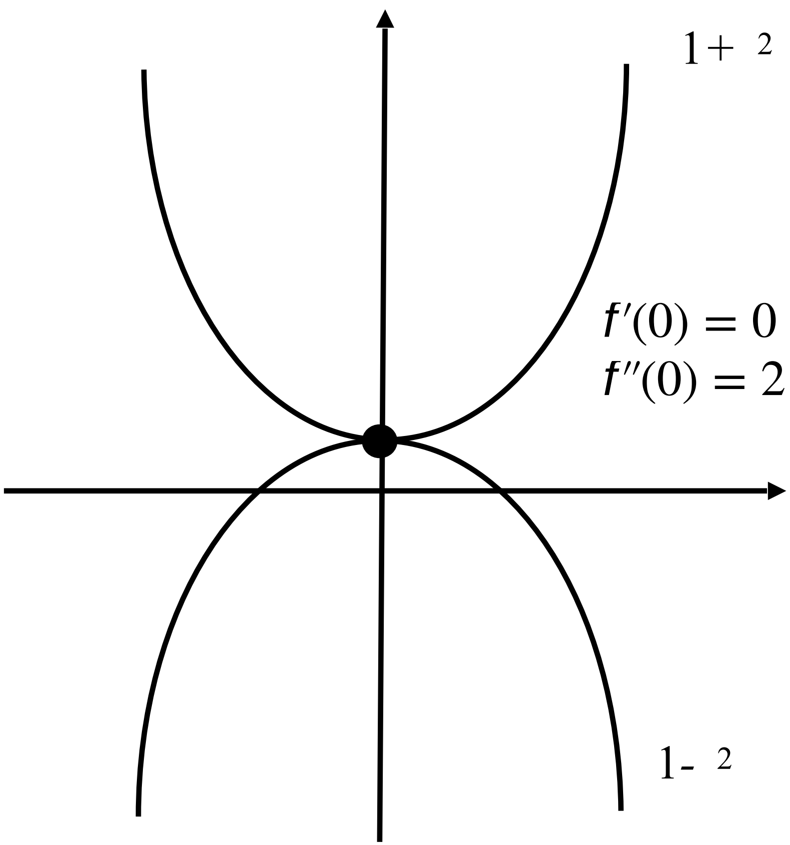
数的二　条件



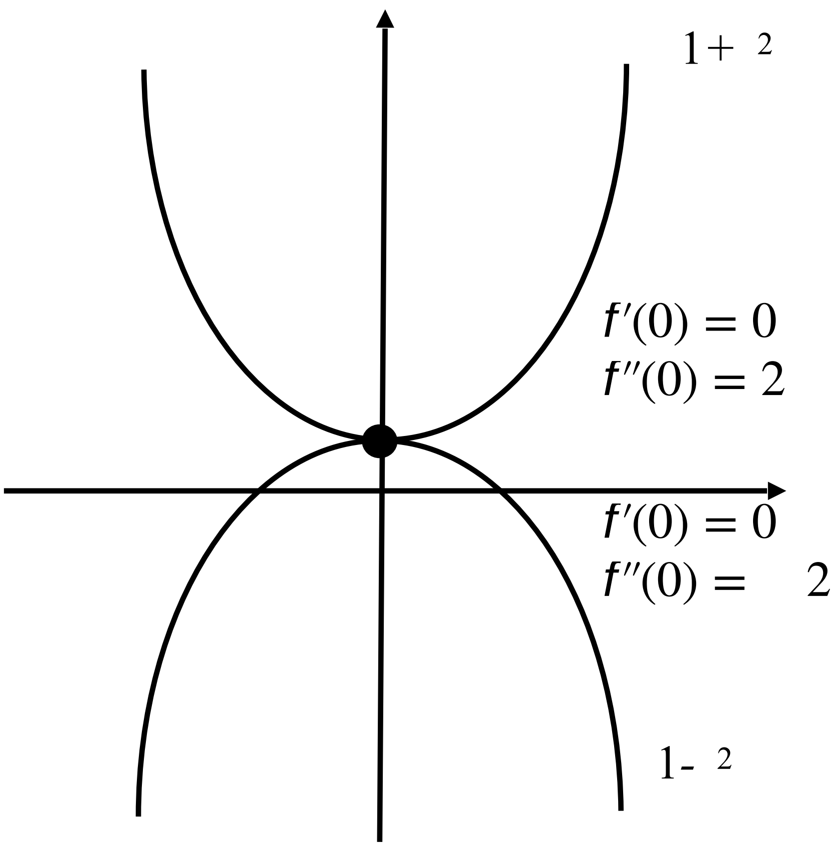
数的二　条件



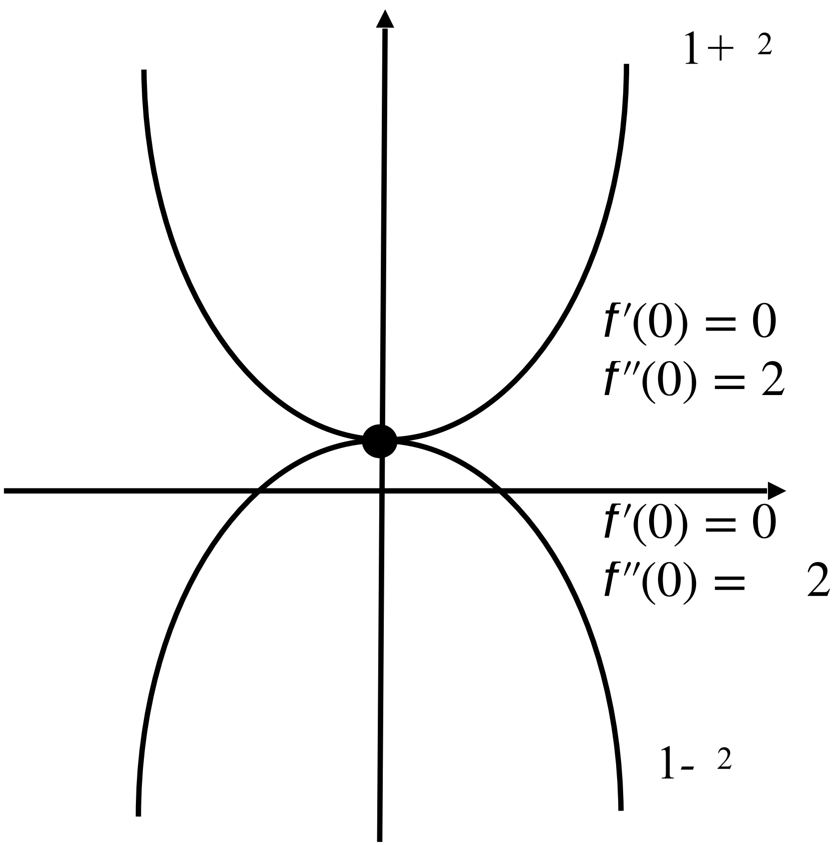
数的二　条件



数的二　条件



数的二　条件



对数-　　数和对数-　　数

- 正值　数f为对数-　　数　需满足$\log f$为　　数

对数-　　数和对数-　　数

- 正值　数f为对数-　　数　需满足$\log f$为　　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-　　数和对数-　　数

- 正值　数f为对数-　　数　需满足$\log f$为　　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　数f为对数-　　数　需满足$\log f$为　　数

对数- 数和对数- 数

- 正值 数f为对数- 数 需满足 为 数

- 正值 数f为对数

对数-　　数和对数-　　数

- 正值　数f为对数-　　数　需满足$\log f$为　数

$$f(\theta x + (1-\theta)y) \ge f(x)^{\theta} f(y)^{1-\theta} \quad \text{for } 0 \le \theta \le 1$$

- 正值　数f为对数-　　数　需满足$\log f$为　数
- 若　数f为对数-　　数　则　数f为　　数
  - $f = e^{\log(f)}$

对数-　数和对数-　数

- 正值　数f为对数-　数　需满足$\log f$为　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　数f为对数-　数　需满足$\log f$为　数
- 若　数f为对数-　数　则　数f为　数
  - $f = e^{\log(f)}$
- 若　数f为　数　则　数f为对数　数

对数-　　数和对数-　　数

- 正值　数f为对数-　　数　需满足$\log f$为　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　数f为对数-　　数　需满足$\log f$为　数
- 若　数f为对数-　　数　则　数f为　　数
  - $f = e^{\log(f)}$
- 若　数f为　　数　则　数f为对数　　数
  - $\log(f)$

对数　　　数

对数　　数

- 数　x^a为若 $a \leq 0$ 为正实数集上的对数-数　若 $a \geq 0$ 为对数-　　数

对数　　数

- 　　数 x^a 为若 $a \leq 0$ 为正实数集上的对数-数　若 $a \geq 0$ 为对数-　　数
- 许多通用的概率密度　数为对数-　　数　如正态分布

对数　　数

- 　数 x^a为若 $a \leq 0$ 为正实数集上的对数-数　若 $a \geq 0$ 为对数-　　数
- 许多通用的概率密度　数为对数-　　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 数

- 数 x^a 为若 $a \leq 0$ 为正实数集上的对数-数 若 $a \geq 0$ 为对数- 数
- 许多通用的概率密度 数为对数- 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 数为对数- 数

对数　　数

- 　　数 x^a 为若 $a \le 0$ 为正实数集上的对数-数　若 $a \ge 0$ 为对数-　　数
- 许多通用的概率密度　数为对数-　　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　数为对数-　　数

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} du$$

对数-　　数的性质

对数-　　数的性质

- 定义域为　集的二　可微　数f为对数-　　　数当且仅当对所有 x

对数-　　数的性质

- 定义域为　集的二　可微　数f为对数-　　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- 数的性质

- 定义域为 集的二 可微 数f为对数- 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- 数的乘积为对数- 数

对数-　　数的性质

- 定义域为　集的二　可微　数f为对数-　　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-　　数的乘积为对数-　　数
- 对数-　　数的和不总是为对数-　　数

对数-　　数的性质

- 定义域为　集的二　可微　数f为对数-　　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-　　数的乘积为对数-　　数
- 对数-　　数的和不总是为对数-　　数
- 积分　若　数 $f : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}$ 为对数-　　数　则

$$g(x) = \int f(x, y)\, dy$$

对数-　　数的性质

- 定义域为　集的二　可微　数f为对数-　　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-　　数的乘积为对数-　　数
- 对数-　　数的和不总是为对数-　　数
- 积分　若　数 $f:\mathbf{R}^n \times \mathbf{R}^m \rightarrow \mathbf{R}$ 为对数-　　数　则

$$g(x) = \int f(x,y)\,dy$$ 为对数-　　数

关于广义不等式的　性

关于广义不等式的 性

- 数 $f : \mathbf{R}^n \to \mathbf{R}^m$ 为K- 数 则定义域为 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$\mathbf{S}^m_+$- 数

关于广义不等式的　性

- 　数 $f:\mathbf{R}^n \to \mathbf{R}^m$ 为*K*-　　　数　则定义域为　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 $\mathbf{S}_+^m$-　　　数
- 证明　给定 $z \in \mathbf{R}^m,\ z^T X^2 z = \|Xz\|_2^2$ 为关于*X*的　　　数

关于广义不等式的 性

- 数 $f: \mathbf{R}^n \to \mathbf{R}^m$ 为K- 数 则定义域为 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 $\mathbf{S}^m_+$- 数

- 证明 给定 $z \in \mathbf{R}^m,\ z^T X^2 z = \|Xz\|_2^2$ 为关于X的 数

$$z^T(\theta X + (1-\theta) Y)^2 z \le \theta z^T X^2 z + (1-\theta) z^T Y^2 z$$

关于广义不等式的 性

- 数 $f:\mathbf{R}^n \to \mathbf{R}^m$ 为 K- 数 则定义域为 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 $\mathbf{S}^m_+$- 数

- 证明 给定 $z \in \mathbf{R}^m,\ z^T X^2 z = \|Xz\|_2^2$ 为关于 X 的 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 性

- 数 $f:\mathbf{R}^n \to \mathbf{R}^m$ 为K- 数 则定义域为 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$\mathbf{S}^m_+$- 数

- 证明 给定 $z \in \mathbf{R}^m,\, z^T X^2 z = \|Xz\|_2^2$ 为关于X的 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 $X, Y \in \mathbf{S}^m,\, 0 \le \theta \le 1$

$$(\theta X + (1-\theta)Y)^2 \preceq \theta X^2 + (1-\theta) Y^2$$