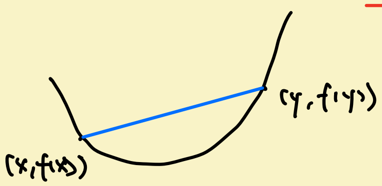
梯度：$\exists g\in\mathbb{R}^n \quad \lim\limits_{h\to 0}\dfrac{f(x+h)-f(x)-g^T h}{\|h\|}=0,$
则 g 是 f 在 x 点处的梯度，记 $\nabla f(x)$

Fréchet 可微 $\quad \lim\limits_{V\to 0}\dfrac{f(X+V)-f(X)-\langle V,G\rangle}{\|V\|}=0,$

Gâteaux 可微 $\quad \lim\limits_{t\to 0}\dfrac{f(X+tV)-f(X)-t\langle V,G\rangle}{t}=0$

① 用定义求梯度 ② 用微分 $df=\langle\nabla f(x),dx\rangle$



Defn 凸函数：$f:\mathbb{R}^n\to\mathbb{R}$ 为适当函数，如果 $\mathrm{dom}\, f$ 是凸集，且
$$f(\theta x+(1-\theta)y)\le\theta f(x)+(1-\theta)f(y)$$
$\forall x,y\in\mathrm{dom}\, f$，$0\le\theta\le 1$ 成立，称 f 是凸函数。

Defn 凹函数：若 f 是凸函数，则 $-f$ 是凹函数。

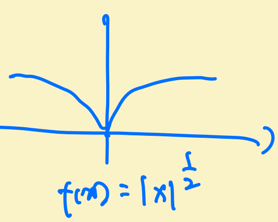
E.g. $\quad f(x)=Ax+b \quad \forall x\in\mathbb{R}^n \to$ 既凸又凹

Defn 严格凸函数

$f:\mathbb{R}^n\to\mathbb{R}$ 为适当函数，如果 $\mathrm{dom}\, f$ 是凸集，且
$$f(\theta x+(1-\theta)y)<\theta f(x)+(1-\theta)f(y)$$
$\forall x,y\in\mathrm{dom}\, f$，$0\le\theta\le 1$ 成立，称 f 是严格凸函数。

E.g. $\quad f(x)=ax+b \qquad f(x)=e^{ax} \qquad f(x)=|x|^p \quad p\ge 1$

$f(x)=x\log x \quad x\in\mathbb{R}_{++}$



E.g. $\quad f(x)=a^Tx+b \qquad f(x)=\|x\|_p \quad p\ge 1 \qquad f(x)=\|x\|_\infty=\max\limits_k |x_k|$

$f(x)=\langle A,X\rangle+B \qquad f(X)=\|X\|_2 \qquad f(X)=\|X\|_F$

凸函数 ⊇ 严格凸函数 ⊇ 强凸函数

Defn (强凸函数) 若存在常数 $m \geq 0$，使得

$$g(x) = f(x) - \frac{m}{2}\|x\|^2$$

若 g 为凸函数，则称 f 为强凸函数，m 称为强凸参数。

$$f(x) \longrightarrow g(x) = f(x) + \frac{m}{2}\|x\|^2$$

Thm (凸函数的判定定理)

$f: \mathbb{R}^n \to \mathbb{R}$ 是凸函数，当且仅当 $\forall x \in \mathrm{dom} f$，$g: \mathbb{R} \to \mathbb{R}$ 是关于 t 的凸函数：

$$g(t) = f(x + tv) \qquad \mathrm{dom}\, g: \{t \mid x + tv \in \mathrm{dom} f\}$$

E.g. $f(X) = -\log\det X \qquad \mathrm{dom} f = S^n_{++}$

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

Proof: 设 f 是凸函数，$\forall t_1, t_2 \in \mathrm{dom}\, g \quad \forall \theta \in (0,1)$

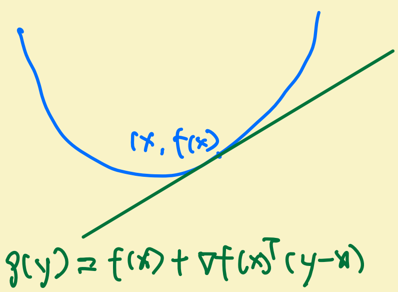
$$x + t_1 v \in \mathrm{dom} f, \quad x + t_2 v \in \mathrm{dom} f.$$

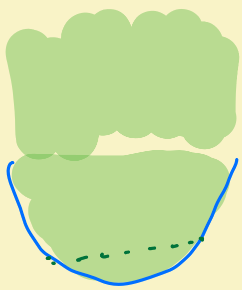
$$\Rightarrow \theta(x + t_1 v) + (1-\theta)(x + t_2 v) = x + (\theta t_1 + (1-\theta)t_2)v \in \mathrm{dom} f$$

因为 $\mathrm{dom}\, g = \{t \mid x + tv \in \mathrm{dom} f\}$，$\theta t_1 + (1-\theta)t_2 \in \mathrm{dom}\, g$.

$\therefore$ dom g 是凸集.

$$
\begin{aligned}
g(\theta t_1 + (1-\theta)t_2) &= f(x + (\theta t_1 + (1-\theta)t_2)v) \\
&\leq \theta f(x + t_1 v) + (1-\theta)f(x + t_2 v) \\
&= \theta g(t_1) + (1-\theta)g(t_2)
\end{aligned}
$$





Thm (一阶条件) $f: \mathbb{R}^n \to \mathbb{R}$ 连续函数, $\mathrm{dom} f$ 是凸集. f 是凸函数当且仅当, f 可微

$$f(y) \geq f(x) + \nabla f(x)^T (y-x) \quad \forall x, y \in \mathrm{dom} f$$

Thm (梯度的单调性) $f: \mathbb{R}^n \to \mathbb{R}$ 连续函数, $\mathrm{dom} f$ 是凸集, f 可微, 则 f 是凸函数当且仅当 ∇f 为单调映射.

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

Proof:

$$\left.\begin{aligned} f(y) &\geq f(x) + \nabla f(x)^T(y-x) \\ f(x) &\geq f(y) + \nabla f(y)^T(x-y) \end{aligned}\right\}$$

Thm (上图) 函数 f 为凸函数, 当且仅当 其上图 $\mathrm{epi} f$ 为凸集

Thm (二阶条件) 函数 f 是凸函数当且仅当

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

如果 $\nabla^2 f(x) \succ 0$, $\forall x \in \mathrm{dom} f$, 则 f 是严格凸函数

Ex. $f(x) = \frac{1}{2} x^T P x + q^T x + r \quad P \in S^n \quad x \in \mathbb{R}^n$

$\Rightarrow$ 什么情况下, $f(x)$ 为凸函数

$\nabla f(x) = Px + q \qquad \nabla^2 f(x) = P \qquad P \succeq 0 \Rightarrow f$ 为凸

Ex. 利用二阶条件判断 $f(x) = \|Ax - b\|_2^2$ 是否为凸函数.

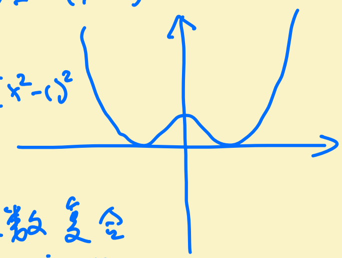
$f(x) = (Ax-b)^T (Ax-b) = x^T A^T A x - 2b^T A x + b^T b$

$\nabla f(x) = 2A^T A x - 2A^T b \qquad \nabla^2 f(x) = 2A^T A \qquad f$ 是凸函数

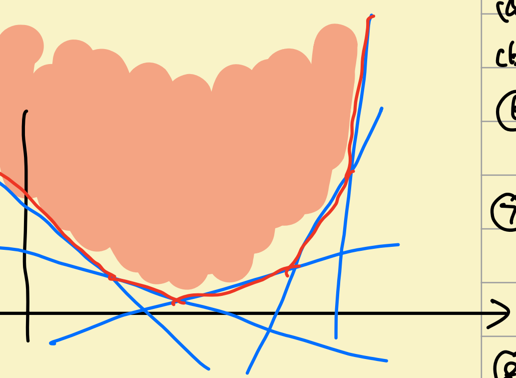
$f(x) = x^2$

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

$g(x) = (x^2-1)^2$



凸函数复合并不一定凸



保凸的运算:

① f 是凸函数 $\Rightarrow$ $\forall \alpha \geq 0$, αf 是凸函数

② f_1, f_2 是凸函数 $\Rightarrow$ $f_1 + f_2$ 是凸函数

③ f 是凸函数 $\Rightarrow$ $f(Ax+b)$ 是凸函数

④ $g: \mathbb{R}^n \to \mathbb{R}$ $h: \mathbb{R} \to \mathbb{R}$. 令 $f(x) = h(g(x))$

(a). g 是凸函数, h 是凸函数且单调不减 ↗

(b) g 是凹函数, h 是凸函数且单调不增 ↘

$\Rightarrow$ f 是凸函数

⑤ $g: \mathbb{R}^n \to \mathbb{R}^k$, $h: \mathbb{R}^k \to \mathbb{R}$, 令 $f(x) = h(g(x))$

(a). g 是凸函数, h 是凸函数且对每个分量单调不减

(b) g 是凹函数, h 是凸函数且对每个分量单调不增

$\Rightarrow$ f 是凸函数

⑥ 若 $f_1, f_2 \cdots f_k$ 是凸函数, 则 $f(x) = \max\{f_1(x), f_2(x), \cdots f_m(x)\}$ 是凸函数

⑦ 若对于每个 $y \in A$, $f(x,y)$ 是关于 x 的凸函数, 则

$$g(x) = \sup_{y \in A} f(x,y) \quad \text{是凸函数}$$

⑧ 若 $f(x,y)$ 关于 (x,y) 整体是凸函数, C 是凸集, 则

$$g(x) = \inf_{y \in C} f(x,y) \quad \text{是凸函数}$$

⑨ 若 $f: \mathbb{R}^n \to \mathbb{R}$ 是凸函数, $g(x,t) = t f(\frac{x}{t})$, $\operatorname{dom} g = \{(x,t) \mid \frac{x}{t} \in \operatorname{dom} f, t>0\}$

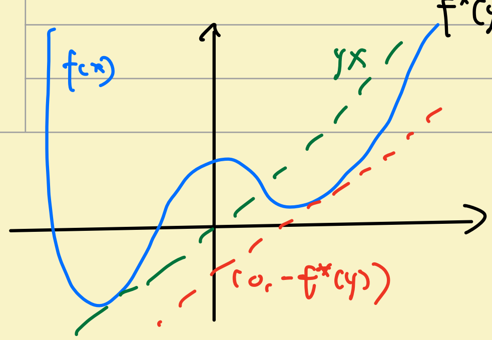
g: f 的透视函数 是凸函数.

E.g. $f(x) = x^T x = \|x\|_2^2$ $\quad g(x,t) = \frac{x^T x}{t}$ $\quad g(x,t)$ 在 $\{(x,t), t>0\}$ 的凸函数.

共轭函数

Defn (共轭函数)

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



Prop (Fenchel 不等式)

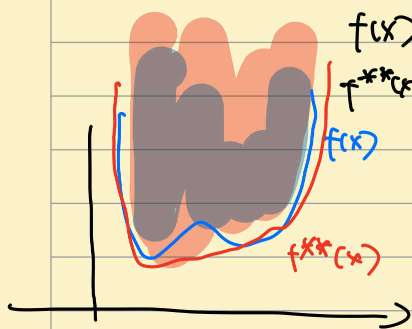
$$f(x) + f^*(y) \geq x^T y$$

Proof:

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

Defn (二次共轭函数)

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



$$\left.\begin{array}{l} f(x) + f^*(y) \geq x^T y \\ f^{**}(x) + f^*(y) \geq x^T y \end{array}\right\} \Rightarrow f^{**}(x) \leq f(x)$$

$$\mathrm{epi} f \leq \mathrm{epi} f^{**}$$

Thm: 若f为闭凸函数，则

$$f^{**}(x) = f(x), \ \forall x.$$

E.g. $f(x) = \frac{1}{2} x^T A x + b^T x + c$

$$f^*(y) = \sup_{x \in \mathrm{dom} f} (y^T x - f(x)) \qquad A \succ 0$$

$$= \sup_{x \in \mathrm{dom} f} (y^T x - \tfrac{1}{2} x^T A x - b^T x - c)$$

$$g(x) = x^T y - \tfrac{1}{2} x^T A x - b^T x - c \qquad \nabla g(x) = y - Ax - b = 0$$

$$x^* = A^{-1}(y - b)$$

$$f^*(y) = (x^*)^T y - \tfrac{1}{2}(x^*)^T A x^* - b^T x^* - c$$

$$= (y-b)^T A^{-1} y - \left(\tfrac{1}{2}(y-b)^T A^{-1} A A^{-1}(y-b) + b^T A^{-1}(y-b) + c\right)$$

$$= \tfrac{1}{2}(y-b)^T A^{-1}(y-b) - c$$

$f(x) = \|x\|_2 \qquad f^*(y) = \sup_{x\in \mathrm{dom} f} \{y^T x - f(x)\}$

$g(x) = y^T x - \|x\|_2$

范数 $\|\cdot\|$，对偶范数：$\|y\|_* = \sup_{\|x\|\le 1} x^T y$

$\|y\|_* \sup_{\|x\|_2 \le 1} \qquad \|x\| \|y\| \ge |x^T y|$

$\|y\|_2 \ge \dfrac{|x^T y|}{\|x\|_2}$

$\|y\|_2 \le 1 \qquad y^T x \le \|y\|_2 \|x\|_2 \le \|x\|_2$，$x=0$ 等号成立

$\Rightarrow \sup_{x\in \mathrm{dom} f} \{y^T x - \|x\|_2\} = 0$

$\|y\|_2 > 1 \qquad P_y x = t \cdot \dfrac{y}{\|y\|_2}$

$\sup_{x\in \mathrm{dom} f} y^T x - \|x\|_2 = t(\|y\|_2 - 1)$

$= \sup_{t\in \mathbb{R}} t(\|y\|_2 - 1) = +\infty$

指示函数 $\rightarrow$ $f^*(y) = \begin{cases} 0 & \|y\|_2 \le 1 \\ +\infty & \|y\|_2 > 1 \end{cases}$

结论：范数的共轭函数为其对偶范数单位球的指示函数。

常见范数 vs 对偶范数

$\|x\|_1 \longleftrightarrow \|x\|_\infty$

$\|x\|_\infty \longleftrightarrow \|x\|_1$

$\|x\|_2 \longleftrightarrow \|x\|_2$

$\|x\|_p \longleftrightarrow \|x\|_q \qquad \frac{1}{p} + \frac{1}{q} = 1$