Ch 3. convex functions.

凸函数：$\forall\theta\in[0,1]$. dom f 凸. $f(\theta x+(1-\theta)y)\le\theta f(x)+(1-\theta)f(y)$.

严格凸 $<$.

判断凸函数：First-order condition 凸.

dom f 凸 ⊕ $f(y)\ge f(x)+\nabla f(x)^T(y-x)$ $\forall x,y$.

First-order 严格凸.

dof 凸 ⊕ $\forall x\neq y$. $f(y)>f(x)+\nabla f(x)^T(y-x)$.

second-order 凸.

f 二阶可微 ⊕ $\forall x\in$ dom f, $\nabla^2 f(x)\succeq 0$.

严格凸的条件是~~不必要的~~ 充分不必要的：$\nabla^2 f(x)\succ 0$.

some examples:

凸：e^{ax}, $a\in R$; x^a when $a\ge1$ or $a\le0$; $|x|^p$, $p\ge1$; $x\log x$; $\|\cdot\|$; max ; log sum exp

approximate

二次线性分式 $f(x,y)=\frac{x^2}{y}$.

凹：x^a, $0\le a\le1$; $\log x$; 几何平均 $f(x)=(\prod_{i=1}^{n}x_i)^{\frac{1}{n}}$; $\log\det X$;

technique for proving: prove on a direction. $X+tV$.

α 下水平集 (sublevel set) $C_\alpha=\{x \text{ in dom} f: f(x)\le\alpha\}$.

α 上水平集. $\{x: f(x)\ge\alpha\}$

上镜图. epigraph. epi $f=\{(x,t)\mid x\in\text{dom} f, f(x)\le t\}$.

亚图：hypo $f=\{(x,t): f(x)\ge t\}$.

convex function $\Longleftrightarrow$ epigraph convex.

concave function $\Longleftrightarrow$ hypograph convex.

$\Rightarrow$ Matrix fractional function. $f(x,Y)=x^TY^{-1}x$ convex.

[note: Matrix definiteness & Schur complete]

上镜图在边界点 $(x,f(x))$ 上的支撑超平面：$(\nabla f(x),-1)$

凸函数的保凸运算.

① nonnegative weighted sum. $f=w_1f_1+\cdots+w_mf_m$ / $g(x)=\int w(y)f(x,y)dy$ if $\forall y$, $f(x,y)$ convex for x.

② 复合仿射映射：f convex. $g(x)=f(Ax+b)$. convex.

concave $g(x)$ concave.

③ pointwise maximum & supremum. f_1,f_2 凸. $f(x)=\max\{f_1(x),f_2(x)\}$.

$\forall y$, $f(x,y)$ 关于 x 凸. $\sup_y f(x,y)=g(x)$ 凸.

仿射函数表示定理：every convex function can be expressed as pointwise supremum of a family of affine function.

$f(x)=\sup_g\{g\mid g$ 是仿射, $g(z)\le f(z)$, $\forall z\}$

④

Composition: [1] Scalar composition $f=h\circ g(x)$. 在所有定理中. f 和 h 保持一致.

g 与 h 是否保持一致 ±

h 的单增/减 ±

[2] Vector composition. $f = h \circ g(x) = f(g_1(x), \cdots, g_k(x))$.

是它在每一维上情况的延伸.

⑤ 最小化：(inf). 也可保凸, 但没有比 sup 更严苛的条件.

$f(x,y)$ 关于 (x,y) 凸 且 C 非空凸 且 for $g(x) = \inf_{y \in C} f(x,y)$. $\exists x$ s.t. $g(x) > -\infty$. $\Rightarrow g$ 凸.

⑥ the perspective function of a function.

└ $f: \mathbb{R}^n \to \mathbb{R}$. 透视函数 $g(x,t) = t f(\frac{x}{t})$. 保凸也保凹

└→ relative entropy & KL divergence are convex.

~~共轭~~

共轭函数. conjugate function of f: $f^*(y) = \sup_{x \in \mathrm{dom} f} (y^T x - f(x))$ $\quad y \in \mathrm{dom} f^* \Leftrightarrow y^T x - f(x)$ upper bdd.

important examples: ① $f(x) = \frac{1}{2} x^T Q x$, $Q \in S_{++}^n \Rightarrow f^*(y) = \frac{1}{2} y^T Q^{-1} y$.

② $f(X) = \log\det X^{-1}$, $X \in S_{++}^n \Rightarrow f^*(Y) = \log\det(-Y)^{-1} - n$ $\quad$ 矩阵标准内积 $\mathrm{tr}(Y^T X)$

③ indicator function $I_S(x) = 0$ if $x \in S$. $\quad = \infty$ if $x \notin S$. $\Rightarrow I_S^*(y) = \sup_{x \in S} y^T x$ support function of S.

④ norm: $f(x) = \|x\|$. $f^*(y) = \begin{cases} 0 & \|y\|_* \le 1 \\ \infty & \text{otherwise} \end{cases} = I_{\|y\|_* \le 1}$

two notes: • dual norm: $\|y\|_* = \sup_{\|x\| \le 1} x^T y$

• $x^T y \le \|x\| \|y\|_*$.

$f(x) = \frac{1}{2}\|x\|^2 \quad f^*(y) = \frac{1}{2}\|y\|_*^2$

important properties: ① Fenchel/Young inequality $x^T y \le f(x) + f^*(y)$.

② f convex 且 closed. $\Rightarrow f^{**} = f$ $\quad$ define f closed: epi f closed.

③ when f is differentiable: $f^*(y) = [(\nabla f)^{-1}(y)]^T y - f((\nabla f)^{-1}(y))$

④ 伸缩和复合仿射变换: $g(x) = a f(x) + b$, $a > 0$. $\Rightarrow \cancel{g^*(x) =}\ g^*(y) = a f^*(\frac{y}{a}) - b$.

$g(x) = f(Ax+b)$, A nonsingular. $\Rightarrow g^*(y) = \cancel{f^*(A^{-1}}\ f((A^{-1})^T y) - b^T (A^{-1})^T y$.

$\mathrm{dom}\, g^* = A^T \mathrm{dom}\, f^*$.

⑤ $f_1(u)$, $f_2(v)$ convex. $\quad \cancel{f_1 + f}\ [f_1(u) + f_2(v)]^*(w, z) = f_1^*(w) + f_2^*(z)$.

拟凸函数 Quasiconvex function.

[1] f quasiconvex $\Leftrightarrow$ dom f convex 且 all sublevel sets of f convex.

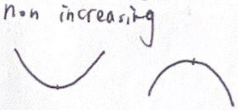
$\Leftrightarrow$ dom f convex 且 $f(\theta x + (1-\theta) y) \le \max\{f(x), f(y)\}$ $\quad 0 \le \theta \le 1$. $x, y \in \mathrm{dom} f$.

[2] quasiconcave: $-f$ quasiconvex.

[3] under the condition f continuous.

f quasiconvex $\Leftrightarrow$ one of these holds: ① f nondecreasing

② f non increasing

③ f  有一点给单增单减分界.

[4] operations preserve quasiconvexity (refer to book P101).

log-concave & log-convex.

└ defined through concave [unlike defining convex].

└→ $f(x) \geq 0$, $\log f$ concave $\Leftrightarrow$ $f(\theta x + (1-\theta)y) \geq \cancel{f(\theta} f(x)^{\theta} f(y)^{1-\theta}$

log-convex : $\frac{1}{f(x)}$ log-concave.

important examples:

log-concave

affine function $f(x) = a^T x + b$.

$f(x) = x^a \quad a \geq 0$

$\Phi(x)$. Gaussian cumulative.

$\det X$.

$\frac{\det X}{\cancel{\text{tr}}\ \text{tr}(X)}$.

density functions : ① normal

② exponential

③ uniform (on convex set)

④ Wishart :

$X_1 \cdots X_p \overset{iid}{\sim} N(0, \Sigma)$.

$X = \sum_{i=1}^{p} x_i x_i^T$

$f(X) = a \det(X)^{\frac{p-n-1}{2}} e^{-\frac{1}{2}\text{tr}(\Sigma^{-1}X)}$

~~Moment generating function~~

log-convex.

$f(x) = x^a \quad a \leq 0.$

$\Gamma(x) = \int_0^{\infty} u^{x-1} e^{-u} du \quad x \geq 1.$ gamma function.

Moment generating function of - X (any distribution)

↓

$E e^{-Xt}$

Laplace transform 对 density function. ~~保~~凸.

under the condition $f(x)$ differentiable:

$$\nabla^2 \log f(x) = \frac{1}{f(x)} \nabla^2 f(x) - \frac{1}{f^2(x)} \nabla f(x) \nabla f(x)^T$$

$$\text{log-convex} \Longleftrightarrow f(x) \nabla^2 f(x) \succeq \nabla f(x) \nabla f(x)^T$$

$$\text{log-concave} \Longleftrightarrow f(x) \nabla^2 f(x) \preceq \nabla f(x) \nabla f(x)^T$$

log 保凹/保凸 : ① Multiplication preserve both.

② sum. 不保凹, 只保凸

③ integration 保凸. $\forall y \in C$. $f(x,y)$ 关于 x 对数凸. $g(x) = \int_C f(x,y)\,dy$ 对数凸.

④ integration 保凹, 但要求 $f(x,y)$ 对 x, y 凹.

⑤ convolution 保凹. require both f, g log-concave.

⑥ marginal dist of log-concave density.