03. Convex functions

By Yang Lin[1] (2024 秋季, @NJU)

Contents

1 Basic properties and examples

2 Operations that preserve convexity

3 Conjugate function

4 Quasiconvex functions

*5 Log-concave and log-convex functions

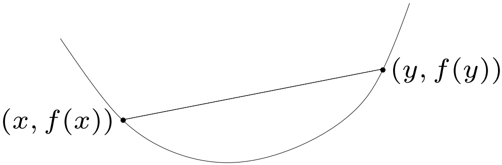
[1]Institute: Nanjing University. Email: linyang@nju.edu.cn.

1 Basic properties and examples

Definition $f : \boldsymbol{R}^n \to \boldsymbol{R}$ is convex if $\mathbf{dom}\, f$ is a convex set and

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

for all $x, y \in \mathbf{dom}\, f$, $0 \le \theta \le 1$



f is concave if $-f$ is convex

f is strictly convex if $\mathbf{dom}\, f$ is convex and

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

for all $x, y \in \mathbf{dom}\, f$, $x \ne y$ and $0 \le \theta \le 1$

1 Basic properties and examples

Extended-value extension $\tilde{f}$ of f is

$$\tilde{f}(x) = f(x), \; x \in \textbf{dom } f, \; \tilde{f}(x) = \infty, \; x \notin \textbf{dom } f$$

often simplifies notation; also satisfy

$$0 \leq \theta \leq 1 \Rightarrow \tilde{f}(\theta x + (1 - \theta y)) \leq \theta \tilde{f}(x) + (1 - \theta)\tilde{f}(y)$$

(as an inequality in $R \bigcup \{\infty\}$)

1 Basic properties and examples

Extended-value extension $\tilde{f}$ of f is

$$\tilde{f}(x) = f(x), \; x \in \textbf{dom } f, \; \tilde{f}(x) = \infty, \; x \notin \textbf{dom } f$$

often simplifies notation; also satisfy

$$0 \leq \theta \leq 1 \Rightarrow \tilde{f}(\theta x + (1-\theta y)) \leq \theta \tilde{f}(x) + (1-\theta)\tilde{f}(y)$$

(as an inequality in $R \bigcup \{\infty\}$)

proof. Apparently holds when $\theta = 0$ or 1.

If x or $y \notin \textbf{dom } f$ and $\theta \neq 0, 1$, the R.H.S. is $+\infty$, and thus the inequality still holds

1 Basic properties and examples

Jensen' s inequality: if f is convex, then for $0 \leq \theta \leq 1$,

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

Extension: if f is convex, then

$$f(\mathbf{E}z) \leq \mathbf{E}f(z)$$

for any random variable z

Basic inequality is special case with discrete distribution $p(x) = \theta, p(y) = 1 - \theta$

1 Basic properties and examples

proof. for $\theta_1 + \theta_2 + \cdots + \theta_k = 1$ and $\theta_i \geq 0$

$$
\begin{aligned}
& f(\theta_1 x_1 + \cdots + \theta_k x_k) \\
=& f\left((1-\theta_k)\left(\frac{\theta_1}{1-\theta_k}x_1 + \cdots + \frac{\theta_{k-1}}{1-\theta_k}x_{k-1}\right) + \theta_k x_k\right) \\
\leq& (1-\theta_k) f\left(\frac{\theta_1}{1-\theta_k}x_1 + \cdots + \frac{\theta_{k-1}}{1-\theta_k}x_{k-1}\right) + \theta_k f(x_k) \\
=& (1-\theta_k) f\left(\left(1-\frac{\theta_{k-1}}{1-\theta_k}\right)\left(\frac{\theta_1}{1-\theta_{k-1}-\theta_k}x_1 + \cdots\right) + \frac{\theta_{k-1}}{1-\theta_k}x_{k-1}\right) + \theta_k f(x_k) \\
\leq& (1-\theta_k)\left(1-\frac{\theta_{k-1}}{1-\theta_k}\right) f\left(\frac{\theta_1}{1-\theta_{k-1}-\theta_k}x_1 + \cdots\right) + \theta_{k-1} f(x_{k-1}) + \theta_k f(x_k) \\
=& \cdots
\end{aligned}
$$

1 Basic properties and examples

Use definition to check convexity

1. 范数（利用三角不等式和齐次性）

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

2. 最大值函数，对任意 $0 \le \theta \le 1$，函数 $f(x)$ 满足

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

1 Basic properties and examples

Affine functions are convex and concave; all norms are convex

Examples on $\mathbf{R}^n$

Affine function: $f(x) = a^T x + b$

Norms: $\|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$ for $p \geq 1$; $\|x\|_\infty = \max_k |x_k|$

Examples on $\mathbf{R}^n$ and $\mathbf{R}^{m \times n}$

Affine function

$$f(X) = \mathbf{tr}(A^T X) + b = \sum_{i=1}^m \sum_{j=1}^n A_{ij} X_{ij} + b$$

1 Basic properties and examples

Convex:

Affine: $ax + b$ on $\mathbf{R}$, for any $a, b \in R$

Exponential: e^{ax}, for any $a \in \mathbf{R}$

Powers: x^α on $\mathbf{R}_{++}$, for $\alpha \geq 1$ or $\alpha \leq 0$

Powers of absolute value: $|x|^p$ on $\mathbf{R}$, for $p \geq 1$

Negative entropy: $-x \log x$ on $\mathbf{R}_{++}$

Concave:

Affine: $ax + b$ on $\mathbf{R}$, for any $a, b \in R$

Powers: x^α on $\mathbf{R}_{++}$, for $0 \leq \alpha \leq 1$

Logarithmic: $\log x$ on $\mathbf{R}_{++}$

1 Basic properties and examples

判断凸性的其他方法：**Restrict convex function to a line**

$f : \mathbf{R}^n \to \mathbf{R}$ is convex if and only if the function $g : \mathbf{R} \to \mathbf{R}$

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

is convex (in t) for any x in $\mathbf{dom}\ f$, $v \in \mathbf{R}^n$

Can check convexity of f by checking convexity of functions of one variable

1 Basic properties and examples

判断凸性的其他方法：**Restrict convex function to a line**

$f : \mathbf{R}^n \to \mathbf{R}$ is convex if and only if the function $g : \mathbf{R} \to \mathbf{R}$

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

is convex (in t) for any x in $\mathbf{dom}\ f$, $v \in \mathbf{R}^n$

Can check convexity of f by checking convexity of functions of one variable

proof. $\Rightarrow$: take any two points x_1 and x_2, they must lie on a line which is in the form of $x + tv$, ...

$\Leftarrow$: $x + tv$ interacts $\mathbf{dom} f$, forming a new convex set. It is straightforward that the points in the new convex set satisfy the convexity condition

1 Basic properties and examples

Example $f : \mathbf{S}^n \to \mathbf{R}$ with $f(X) = \log\det X$, **dom** $f = \mathbf{S}^n_{++}$

$$g(t) = \log\det(X + tV) = \log\det X + \log\det(I + t(X^{-1/2})^T V X^{-1/2})$$
$$= \log\det X + \sum_{i=1}^{n} \log(1 + t\lambda_i)$$

where λ_i are the eigenvalues of $(X^{-1/2})^T V X^{-1/2}$

g is concave in t (for any choice of $X \succ 0,\ V$); hence f is concave

1 Basic properties and examples

First-order condition

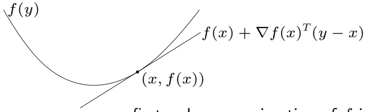
f is differentiable if **dom** f is open and the gradient

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

exists at each $x \in \textbf{dom } f$

1st-order condition: differentiable f with convex domain is convex iff

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



first-order approximation of f is global underestimator

1 Basic properties and examples

proof.

$\Leftarrow$: for all $0 < t \leq 1$, $f(x+t(y-x)) \leq (1-t)f(x)+tf(y)$, there is

$$\begin{aligned} f(y) \geq & \frac{1}{t}f(x+t(y-x)) - \frac{1-t}{t}f(x) \\ = & f(x) + \frac{f(x+t(y-x))-f(x)}{t} \\ = & f(x) + f'(x)(y-x) \end{aligned}$$

$\Rightarrow$: let $z = \theta x + (1-\theta)y$. Then, $f(x) \geq f(z) + f'(z)(x-z)$ and $f(y) \geq f(z) + f'(z)(y-z)$. Then,

$$\begin{aligned} & \theta f(x) + (1-\theta)f(y) \\ \geq & \theta f(z) + \theta f'(z)(x-z) + (1-\theta)f(z) + (1-\theta)f'(z)(y-z) \\ = & f(z) + f'(z)[\theta x + (1-\theta)y - z] = f(z) \end{aligned}$$

1 Basic properties and examples

Second-order condition

f is twice differentiable if **dom** f is open and the Hassian $\nabla^2 f(x) \in \boldsymbol{S}^n$

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

exists at each $x \in$ **dom** f

2nd-order condition: for twice differentiable f with convex domain

- f is convex iff
$$\nabla^2 f(x) \succeq 0, \ \text{for all } x \in \textbf{dom } f$$
- If $\nabla^2 f(x) \succ 0$, for all $x \in$ **dom** f, then f is stricly convex

1 Basic properties and examples

proof.

observe $f(y) \geq f(x) + \nabla f(x)(y-x)$, that is $\frac{f(y)-f(x)}{y-x} \geq \nabla f(x)$

$\Leftarrow$: there are $f'(x) \leq \frac{f(y)-f(x)}{y-x}$ and $f(x) \geq f(y) + f'(y)(x-y)$

hence, $f'(y) \geq \frac{f(y)-f(x)}{y-x} \geq f'(x)$

$\Rightarrow$: there is a $z \in [x, y]$, such that $\frac{f(y)-f(x)}{y-x} = f'(z)$

hence, $f'(x) \leq f'(z) = \frac{f(y)-f(x)}{y-x}$

as a result, $f(y) \geq f(x) + (y-x)f'(x)$

1 Basic properties and examples

Example

设函数 $f\colon \mathbb{R} \to \mathbb{R}$ 是凸函数，其定义域为 $\mathbb{R}$。函数在 $\mathbb{R}$ 上有上界。证明该函数是常数。

1 Basic properties and examples

Example

设函数 $f: \mathbb{R} \to \mathbb{R}$ 是凸函数，其定义域为 $\mathbb{R}$。函数在 $\mathbb{R}$ 上有上界。证明该函数是常数。

证明：任取一点 x。如果 $f'(x) \neq 0$，不妨设 $f'(x) > 0$。

根据凸函数的性质，我们有 $f'(z) \geq f'(x)$ 对所有的 $z \geq x$ 都成立。

假设 $f'(x) = \delta > 0$。那么我们有

$$f(z) - f(x) \geq \delta(z - x)$$

因此当 $z \to \infty$，$f(z)$ 趋向于正无穷。此时，$f(x)$ 不存在上界。因此 $f'(x) > 0$ 不成立。同样，$f'(x) < 0$ 也不成立。因此，只能有 $f'(x) = 0$。得证。

1 Basic properties and examples

Example

设函数 $f: \mathbb{R} \to \mathbb{R}$ 是凸函数，其定义域为 $\mathbb{R}$。函数在 $\mathbb{R}$ 上有上界。证明该函数是常数。

证明：任取一点 x。如果 $f'(x) \neq 0$，不妨设 $f'(x) > 0$。

根据凸函数的性质，我们有 $f'(z) \geq f'(x)$ 对所有的 $z \geq x$ 都成立。

假设 $f'(x) = \delta > 0$。那么我们有

$$f(z) - f(x) \geq \delta(z - x)$$

因此当 $z \to \infty$，$f(z)$ 趋向于正无穷。此时，$f(x)$ 不存在上界。因此 $f'(x) > 0$ 不成立。同样，$f'(x) < 0$ 也不成立。因此，只能有 $f'(x) = 0$。得证。

Q: How about $x \in \mathbb{R}^n$?

1 Basic properties and examples

设函数 $f\colon \mathbb{R} \to \mathbb{R}$ 是凸函数。试证明：利用凸函数的定义证明，对于三变量 $x_1 < x_2 < x_3$, 函数 f 满足如下公式（单调性质）：

$$\frac{f(x_2)-f(x_1)}{x_2-x_1} \leq \frac{f(x_3)-f(x_2)}{x_3-x_2}$$

1 Basic properties and examples

设函数 f: $\mathbb{R} \to \mathbb{R}$ 是凸函数。试证明：利用凸函数的定义证明，对于三变量 $x_1 < x_2 < x_3$, 函数 f 满足如下公式（单调性质）:

$$\frac{f(x_2)-f(x_1)}{x_2-x_1} \leq \frac{f(x_3)-f(x_2)}{x_3-x_2}$$

proof.

$$f(x_3) \geq f(x_2) + \nabla f(x_2)(x_3 - x_2)$$
$$f(x_1) \geq f(x_2) + \nabla f(x_2)(x_1 - x_2)$$
$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \nabla f(x_2)$$
$$\frac{f(x_3) - f(x_2)}{x_3 - x_2} \geq \nabla f(x_2) \geq \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

This completes the proof.

1 Basic properties and examples

Examples

Quadratic function: $f(x) = (1/2)x^T Px + q^T x + r$, ($P \in \boldsymbol{S}^n$, convex if $P \succeq 0$)

$$\nabla f(x) = Px + q, \;\; \nabla^2 f(x) = P$$

Least-squares objective: $f(x) = \|Ax - b\|_2^2$ (convex for any A)

$$\nabla f(x) = 2A^T(Ax - b), \;\; \nabla^2 f(x) = 2A^T A$$

Quadratic-over-linear: $f(x, y) = x^2/y$ (convex for $y > 0$)

$$\nabla^2 f(x, y) = \frac{2}{y^3} \begin{bmatrix} y \\ -x \end{bmatrix} \begin{bmatrix} y \\ -x \end{bmatrix}^T \succeq 0$$

An example of nonconvex: $f(x) = 1/x^2$

1 Basic properties and examples

Log-sum-exp: $f(x) = \log \sum_{k=1}^{n} \exp(x_k)$ is convex

$$\nabla^2 f(x) = \frac{1}{\mathbf{1}^T z} \mathbf{diag}(z) - \frac{1}{(\mathbf{1}^T z)^2} z z^T, \quad (z_k = \exp(x_k))$$

to show $\nabla^2 f(x) \succeq 0$, we must verify that $v^T \nabla^2 f(x) v \geq 0$ for all v:

$$v^T \nabla^2 f(x) v = \frac{(\sum_k z_k v_k^2)(\sum_k z_k) - (\sum_k z_k v_k)^2}{(\sum_k z_k)^2} \geq 0$$

since $(\sum_k z_k v_k)^2 \leq (\sum_k z_k v_k^2)(\sum_k z_k)$ (from Cauchy-Schwarz inequality)

Geometric mean: $f(x) = (\prod_{k=1}^{n} x_k)^{1/n}$ on $\mathbf{R}_{++}^n$ is concave

(similar proof as for log-sum-exp)

1 Basic properties and examples

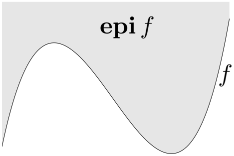
α-sublevel set of f: $\mathbf{R}^n \to \mathbf{R}$

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

Sublevel sets of convex functions are convex (converse is false)

Epigraph of f: $\mathbf{R}^n \to \mathbf{R}$

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



f is convex if and only if **epi** f is a convex set

2 Operations that preserve convexity

Practical methods for establishing convexity of a function:

1. Verify definition (often simplified by restricting to a line)

2. For twice differentiable functions, show $\nabla^2 f(x) \succeq 0$

3. Show that f is obtained from simple convex functions by operations that preserve convexity

- Nonnegative weighted sum
- Composition with affine function
- Pointwise maximum and supremum
- Composition
- Minimization
- Perspective

2 Operations that preserve convexity

Sum: $f_1 + f_2$ convex if f_1, f_2 convex (extends to infinite sums,integrals)

Nonnegative multiple: αf is convex if f is convex, $\alpha \geq 0$

Positive weighted sum

Composition with affine function: $f(Ax + b)$ is convex if f is convex

Example:

- Log barrier for linear inequalities

$$f(x) = -\log(b_i - a_i^T x), \mathbf{dom}\ f = \{x | a_i^T x < b_i,\ \ i = 1, ..., m\}$$

- (Any) norm of affine function: $f(x) = \|Ax + b\|$

2 Operations that preserve convexity

Pointwise maximum:

If $f_1, \ldots, f_m$ are convex, then $f(x) = \max\{f_1(x), \ldots, f_m(x)\}$ is convex

Example:

- Piecewise-linear function: $f(x) = \max_{i=1,\ldots,m}(a_i^T x + b_i)$ is convex
- Sum of r largest components of $x \in \mathbf{R}^n$:

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

is convex ($x_{[i]}$ is i-th largest component of x)

2 Operations that preserve convexity

Pointwise maximum:

If $f_1, \ldots, f_m$ are convex, then $f(x) = \max\{f_1(x), \ldots, f_m(x)\}$ is convex

Example:

- Piecewise-linear function: $f(x) = \max_{i=1,\ldots,m}(a_i^T x + b_i)$ is convex
- Sum of r largest components of $x \in \mathbf{R}^n$:

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

is convex ($x_{[i]}$ is i-th largest component of x)

proof.

$$f(x) = \max\{x_{i_1} + x_{i_2} + \cdots + x_{i_r} | 1 \leq i_1 < i_2 < \cdots < i_r \leq n\}$$

2 Operations that preserve convexity

Pointwise supremum: if $f(x, y)$ is convex in x for each $y \in \mathcal{A}$, then

$$g(x) = \sup_{y \in \mathcal{A}} f(x, y)$$

is convex

Example:

- Support function of a set C: $S_C(x) = \sup_{y \in C} y^T x$ is convex
- Distance to farthest point in a set C:

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

- Maximum eigenvalue of symmetric matrix: for $X \in \mathbf{S}^n$

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

2 Operations that preserve convexity

Composition with scalar functions: composition of $g : \mathbf{R}^n \to \mathbf{R}$ and $h : \mathbf{R} \to \mathbf{R}$:

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

f is convex if: (1) g convex, h convex, $\tilde{h}$ nondecreasing; (2) g concave, h convex, $\tilde{h}$ nonincreasing

proof. (for $n = 1$, differentiable g, h)

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

note: monotonicity must hold for extended-value extension $\tilde{h}$

Example:

- $\exp g(x)$ is convex if g is convex
- $1/g(x)$ is convex if g is concave and positive

2 Operations that preserve convexity

Vector composition of $g : \mathbf{R}^n \to \mathbf{R}^k$ and $h : \mathbf{R}^k \to \mathbf{R}$:

$$f(x) = h(g(x)) = h(g_1(x), g_2(x), ..., g_k(x))$$

f is convex if: (1) g_i convex, h convex, $\tilde{h}$ nondecreasing in each argument; (2) g_i concave, h convex, $\tilde{h}$ nonincreasing in each argument

proof. (for $n = 1$, differentiable g, h)

$$f''(x) = g'(x)^T \nabla^2 h(g(x)) g'(x) + \nabla h(g(x))^T g''(x)$$

note: monotonicity must hold for extended-value extension $\tilde{h}$

Example:

- $\sum_{i=1}^m \log g_i(x)$ is concave if g_i are concave and positive
- $\log \sum_{i=1}^m \exp(g_i(x))$ is convex if g_i are convex

2 Operations that preserve convexity

Minimization: if $f(x, y)$ is convex in (x, y) and C is a convex set, then

$$g(x) = \inf_{y \in C} f(x, y)$$

is convex

Example:

- Distance to a set: $\mathbf{dist}(x, S) = \inf_{y \in S} \|x - y\|$ is convex if S is convex

2 Operations that preserve convexity

proof.

$$(x, t) \in \textbf{epi } g \Leftrightarrow \begin{cases} (1) \;\; t \geq f(x, y) \text{ or } (x, y, t) \in \textbf{epi } f \\ (2) \;\; y \in C \text{ which is convex} \end{cases}$$

condition (1) and (2) result in two sets, respectively

$$S_1 = \{(x, y, t) | y \in C\}$$
$$S_2 = \{(x, y, t) | (x, y, t) \in \textbf{epi } f\}$$

both S_1 and S_2 are convex

by analyzing the conditions, **epi** g is the projection of the interaction of two convex sets. Hence, **epi** g is convex

2 Operations that preserve convexity

Perspective of a function $f : \mathbf{R}^n \to \mathbf{R}$ is the function $g : \mathbf{R}^n \times \mathbf{R} \to \mathbf{R}$

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

g is convex if f is convex

proof.

$$\begin{aligned}(x, t, s) \in \mathbf{epi}\ g \Leftrightarrow & g(x, t) \leq s \\ \Leftrightarrow & tf(x/t) \leq s \\ \Leftrightarrow & f(x/t) \leq s/t \\ \Leftrightarrow & (x/t, s/t) \in \mathbf{epi}\ f\end{aligned}$$

It implies **epi** f is the perspective of **epi** g. Hence, **epi** f convex iff **epi** g convex

2 Operations that preserve convexity

Example:

- $f(x) = x^T x$ is convex; hence $g(x, t) = x^T x/t$ is convex for $t > 0$
- Negative logarithm $f(x) = -\log x$ is convex; hence relative entropy $g(x, t) = t \log t - t \log x$ is convex on $\mathbf{R}^2_{++}$
- If f is convex, then

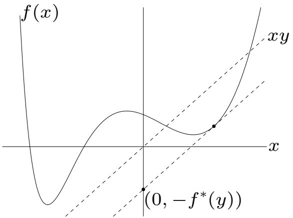
$$g(x) = (c^T x + d) f\Big((Ax + b)/(c^T x + d)\Big)$$

is convex on $\{x | c^T x + d > 0, (Ax + b)/(c^T x + d) \in \mathbf{dom}\ f\}$

3 The conjugate function

The conjugate of a function f is

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



f^* is convex (even if f is not), why?

will be useful in chapter 5

3 The conjugate function

Examples:

- Negative logarithm $f(x) = -\log x$

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

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

$$f^*(y) = \sup_x \left(y^T x - (1/2)x^T Qx\right) = \frac{1}{2} y^T Q^{-1} y$$

Preliminaries: derivative on matrix

$$\frac{\partial(\mathbf{a}^T\mathbf{x})}{\partial\mathbf{x}} = \frac{\partial(\mathbf{x}^T\mathbf{a})}{\partial\mathbf{x}} = \mathbf{a}$$

$$\frac{\partial(\mathbf{a}^T\mathbf{X}\mathbf{b})}{\partial\mathbf{X}} = \mathbf{a}\mathbf{b}^T = \mathbf{a}\otimes\mathbf{b} \in \mathbb{R}^{n\times n}$$

$$\frac{\partial(\mathbf{a}^T\mathbf{X}^T\mathbf{b})}{\partial\mathbf{X}} = \mathbf{b}\mathbf{a}^T = \mathbf{b}\otimes\mathbf{a} \in \mathbb{R}^{n\times n}$$

$$\frac{\partial(\mathbf{x}^T\mathbf{A}\mathbf{x})}{\partial\mathbf{x}} = (\mathbf{A}+\mathbf{A}^T)\mathbf{x}$$

$$\frac{\partial[(\mathbf{X}\mathbf{b}+\mathbf{c})^T\mathbf{A}(\mathbf{X}\mathbf{b}+\mathbf{c})]}{\partial\mathbf{X}} = (\mathbf{A}+\mathbf{A}^T)(\mathbf{X}\mathbf{b}+\mathbf{c})\mathbf{b}^T$$

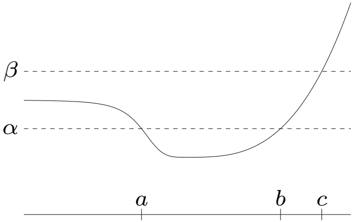
$$\frac{\partial(\mathbf{a}^T\mathbf{X}^T\mathbf{X}\mathbf{b})}{\partial\mathbf{X}} = \mathbf{X}(\mathbf{a}\otimes\mathbf{b}+\mathbf{b}\otimes\mathbf{a})$$

4 Quasiconvex functions

$f : \mathbf{R}^n \to \mathbf{R}$ is quasiconvex if $\mathbf{dom}\, f$ is convex and the sublevel sets

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

are convex for all α



f is quasiconcave if $-f$ is quasiconvex

f is quasilinear if it is quasiconvex and quasiconcave

4 Quasiconvex functions

Examples:

1. $\sqrt{|x|}$ is quasiconvex on $\mathbf{R}$
2. $\mathrm{ceil}(x) = \inf\{z \in \mathbf{Z} | z \geq x\}$ is quasilinear
3. $\log x$ is quasilinear on $\mathbf{R}_{++}$
4. $f(x_1, x_2) = x_1 x_2$ is quasiconcave on $\mathbf{R}^2_{++}$
5. Linear-fractional function

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

is quasilinear

6. Distance ratio

$$f(x) = \frac{\|x - a\|_2}{\|x - b\|_2}, \quad \mathbf{dom}\, f = \{x | \|x - a\|_2 \leq \|x - b\|_2\}$$

is quasiconvex

4 Quasiconvex functions

Properties:

1. Modified Jensen inequality: for quasiconvex f

$$0 \leq \theta \leq 1 \Longrightarrow f(\theta x + (1-\theta)y) \leq \max\{f(x), f(y)\}$$

4 Quasiconvex functions

Properties:

1. Modified Jensen inequality: for quasiconvex f

$$0 \leq \theta \leq 1 \Longrightarrow f(\theta x + (1-\theta)y) \leq \max\{f(x), f(y)\}$$

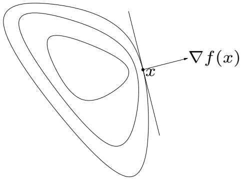
proof. Obviously, x and y are in the sublevel set of $f(x)$ and $f(y)$, respectively. x and y are in the sublevel set of the larger one between $f(x)$ and $f(y)$. Since the sublevel set is convex, $\theta x + (1-\theta)y$ must be in the sublevel set of $f(x)$ or $f(y)$, that is $f(\theta x + (1-\theta)y) \leq \max\{f(x), f(y)\}$

4 Quasiconvex functions

Properties:

2. First-order condition: differentiable f with cvx domain is quasiconvex iff

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



proof. (hint) $f(y) \le f(x)$ means y lies in the interior of a convex sublevel set $C_{f(x)}$, while x lies on the boundary, $\nabla f(x)$ and x determine the half space that contains the sublevel set

3. Sums of quasiconvex functions are not necessarily quasiconvex

5 Log-concave and log-convex functions

A positive function f is log-concave if $\log f$ is concave

$$f(\theta x + (1-\theta) y) \geq f(x)^{\theta} f(y)^{1-\theta} \text{ for } 0 \leq \theta \leq 1$$

f is log-convex if $\log f$ is convex

Example:

- Powers: x^a on $\mathbf{R}_{++}$ is log-convex for $a \leq 0$, log-concave for $a \geq 0$
- Many common probability densities are log-concave, e.g., normal:

$$f(x) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} e^{-\frac{1}{2}(x-\bar{x})^T \Sigma^{-1} (x-\bar{x})}$$

- Cumulative Gaussian distribution function Φ is log-concave

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