Convex Functions

Yuanming Shi

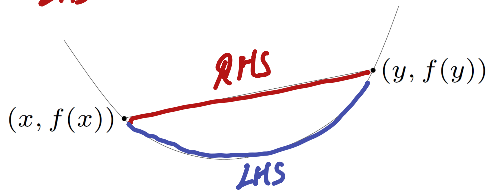
ShanghaiTech University

Outline

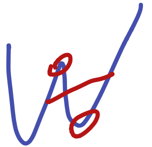
Definition of Convex Function

- A function $f : \mathbb{R}^n \Rightarrow \mathbb{R}$ is said to be **convex** if the domain, $\mathsf{dom} f$, is convex and for any $\boldsymbol{x}, \boldsymbol{y} \in \mathsf{dom} f$ and $0 \le \theta \le 1$,

$$\underbrace{f(\theta \boldsymbol{x} + (1-\theta)\boldsymbol{y})}_{LHS} \le \underbrace{\theta f(\boldsymbol{x}) + (1-\theta) f(\boldsymbol{y})}_{RHS}$$



nonconvex



- f is **strictly convex** if the inequality is strict for $0 < \theta < 1$

- f is **concave** if $-f$ is convex

Strongly convex: if $\exists\, \alpha > 0$, such that

$g(x) = f(x) - \alpha \|x\|^2$ is convex

lemma:

strong convexity $\overset{①}{\Rightarrow}$ strict convexity $\overset{②}{\Rightarrow}$ convexity

Proof: ① ⇒ ②, strong convexity of f implies

$$f(\lambda x + (1-\lambda) y) - \alpha \|\lambda x + (1-\lambda) y\|^2 \leq$$
$$\lambda [f(x) - \alpha \|x\|^2] + (1-\lambda)[f(y) - \alpha \|y\|^2]$$

but

$$\lambda \alpha \|x\|^2 + (1-\lambda) \alpha \|y\|^2 - \alpha \|\lambda x + (1-\lambda) y\|^2 > 0$$

$\forall\, x, y,\quad x \neq y,\quad \forall \lambda \in (0,1)$

$$\Downarrow$$

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

The converse statements are not true, e.g.,

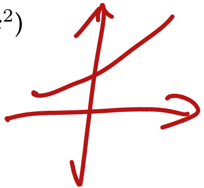
$f(x) = x$ is convex, not strictly convex

Examples on $\mathbb{R}$

Convex functions:

- affine: $ax + b$ on $\mathbb{R}$
- powers of absolute value: $|x|^p$ on $\mathbb{R}$, for $p \geq 1$ (e.g., $|x|$)
- powers: x^p on $\mathbb{R}_{++}$, for $p \geq 1$ or $p \leq 0$ (e.g., x^2)
- exponential: e^{ax} on $\mathbb{R}$
- negative entropy: $x \log x$ on $\mathbb{R}_{++}$

e^x



Concave functions:

- affine: $ax + b$ on $\mathbb{R}$
- powers: x^p on $\mathbb{R}_{++}$, for $0 \leq p \leq 1$
- logarithm: $\log x$ on $\mathbb{R}_{++}$

$\log x$

Examples on $\mathbb{R}^n$

$\|x\|_p$ ($p \geq 1$, convex)

$0 \leq p < 1$, non-convex

$\langle a, x \rangle$

- **Affine functions** $f(\boldsymbol{x}) = \boldsymbol{a}^T\boldsymbol{x} + b$ are convex and concave on $\mathbb{R}^n$
- **Norms** $\|\boldsymbol{x}\|$ are convex on $\mathbb{R}^n$ (e.g., $\|\boldsymbol{x}\|_\infty, \|\boldsymbol{x}\|_1, \|\boldsymbol{x}\|_2$)
- **Quadratic functions** $f(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{P}\boldsymbol{x} + 2\boldsymbol{q}^T\boldsymbol{x} + r$ are convex $\mathbb{R}^n$ if and only if $\boldsymbol{P} \succeq \boldsymbol{0}$
- The **geometric mean** $f(\boldsymbol{x}) = (\prod_{i=1}^n x_i)^{1/n}$ is concave on $\mathbb{R}^n_{++}$
- The **log-sum-exp** $f(\boldsymbol{x}) = \log \sum_i e^{x_i}$ is convex on $\mathbb{R}^n$ (it can be used to approximate $\max_{i=1,\cdots,n} x_i$) $\max x_i \leq f(x) \leq \max x_i + \log n$
- **Quadratic over linear:** $f(\boldsymbol{x}, y) = \boldsymbol{x}^T\boldsymbol{x}/y$ is convex on $\mathbb{R}^n \times \mathbb{R}_{++}$

$x_i := x$

Examples on $\mathbb{R}^{n\times n}$

- **Affine functions:** (prove it!)

$$f(\boldsymbol{X}) = \mathrm{Tr}(\boldsymbol{A}\boldsymbol{X}) + b$$

are convex and concave on $\mathbb{R}^{n\times n}$

- **Logarithmic determinant function:** (prove it!)

$$f(\boldsymbol{X}) = \mathrm{logdet}(\boldsymbol{X})$$

$\underline{\log\det(I + H^* Q H)}$

channel capacity

is concave on $\mathbb{S}^n = \{\boldsymbol{X} \in \mathbb{R}^{n\times n} \mid \boldsymbol{X} \succeq \boldsymbol{0}\})$

- **Maximum eigenvalue function**: (prove it!)

$$f(\boldsymbol{x}) = \lambda_{\max}(\boldsymbol{X}) = \underline{\sup_{\boldsymbol{y}\neq\boldsymbol{0}} \frac{\boldsymbol{y}^T\boldsymbol{X}\boldsymbol{y}}{\boldsymbol{y}^T\boldsymbol{y}}}$$

is convex on $\mathbb{S}^n$

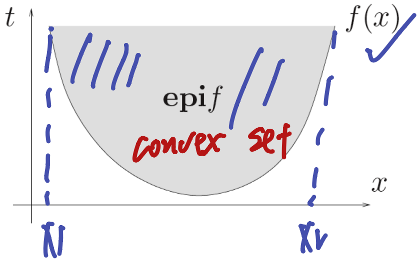
Epigraph

- The **epigraph** of f if the set

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

- Relation between convexity in sets and convexity in functions:

$$f \text{ is convex} \iff \text{epi}\, f \text{ is convex}$$



Outline

Restriction of a Convex Function to a Line

proof is straight from the definition

$x+tv$ $\quad v$ $\quad x$

gradient descent: $x_{k+1} = x_k - t\nabla f(x_k)$

t: learning rate

step size

A

- $f \colon \mathbb{R}^n \longrightarrow \mathbb{R}$ is convex if and only if the function $g \colon \mathbb{R} \longrightarrow \mathbb{R}$

$$g(t) = f(\boldsymbol{x} + t\boldsymbol{v}), \quad \operatorname{dom} g = \{t \mid \boldsymbol{x} + t\boldsymbol{v} \in \operatorname{dom} f\}$$

is convex for any $\boldsymbol{x} \in \operatorname{dom} f$, $\boldsymbol{v} \in \mathbb{R}^n$

$\min_t g(t)$ exact line search

A

- In words: a function is convex if and only if it is convex when restricted to an arbitrary line.

$g(t) = f(x_k - t\nabla f(x_k))$

- Implication: we can check convexity of f by checking convexity of functions of one variable!

- Example: concavity of $\operatorname{logdet}(\boldsymbol{X})$ follows from concavity of $\log(x)$

S^n_{++}

$A, B \in S_{++}^{n}$

① $\det(AB) = \underline{\det(A)\det(B)}$

② $\det(A) = \prod_{i=1}^{n} \lambda_i$

$\underline{M = I + tQ}$, $Q = U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^T$

$UU^T = I$

$M = U \begin{pmatrix} \underline{1+t\lambda_1} & & \\ & \ddots & \\ & & \underline{1+t\lambda_n} \end{pmatrix} U^T$

$\det(M) = \prod_{i=1}^{n} \underline{(1+t\lambda_i)}$

$\log\det(M) = \log \prod_{i=1}^{n} (1+t\lambda_i)$

$= \sum_{i=1}^{n} \log(1+t\lambda_i)$

Example

$X+tV = X^{\frac{1}{2}}\underbrace{(I + t X^{-\frac{1}{2}} V X^{-\frac{1}{2}})}X^{\frac{1}{2}}$

Example: concavity of $\operatorname{logdet}(\boldsymbol{X})$:

$$\begin{aligned} g(t) = \operatorname{logdet}(\boldsymbol{X} + t\boldsymbol{V}) = \quad & \operatorname{logdet}(\boldsymbol{X}) + \operatorname{logdet}(\boldsymbol{I} + t\boldsymbol{X}^{-1/2}\boldsymbol{V}\boldsymbol{X}^{-1/2}) \\ & = \operatorname{logdet}(\boldsymbol{X}) + \sum_{i=1}^{n} \log(1 + t\lambda_i) \end{aligned}$$

where λ_i's are the eigenvalues of $\boldsymbol{X}^{-1/2}\boldsymbol{V}\boldsymbol{X}^{-1/2}$. $: Q$

The function g is concave in t for any choice of $\boldsymbol{X} \succ \boldsymbol{0}$ and $\boldsymbol{V}$; therefore, f is concave.

$\{t \mid X \succ 0,\ X+tV \succ 0\}$

Outline

First and Second Order Conditions I

- **Gradient** (for differentiable f):

$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial f(\boldsymbol{x})}{\partial x_1} & \cdots & \frac{\partial f(\boldsymbol{x})}{\partial x_n} \end{bmatrix}^T \in \mathbb{R}^n$$

- **Hessian** (for twice differentiable f):

$$\nabla^2 f(\boldsymbol{x}) = \left(\frac{\partial^2 f(\boldsymbol{x})}{\partial x_i \partial x_j} \right)_{ij} \in \mathbb{R}^{n \times n} \quad S^n$$

- **Taylor series**:

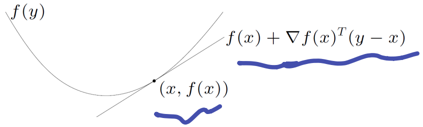
$$f(\boldsymbol{x} + \boldsymbol{\delta}) = f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^T \boldsymbol{\delta} + \frac{1}{2} \boldsymbol{\delta}^T \nabla^2 f(\boldsymbol{x}) \boldsymbol{\delta} + o\left(\|\boldsymbol{\delta}\|^2\right)$$

$$\lim_{\delta \to 0} o(\|\delta\|^2) = 0$$

First and Second Order Conditions II

- **First-order condition:** a differentiable f with convex domain is convex if and only if

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



- Interpretation: first-order approximation is a global under estimator

- **Second-order condition:** a twice differentiable f with convex domain is convex if and only if

$$\nabla^2 f(\boldsymbol{x}) \succeq \mathbf{0} \quad \forall \boldsymbol{x} \in \operatorname{dom} f$$

① f is convex

② $f(y) \geq f(x) + \nabla f(x)^T (y-x)$, $\forall x, y \in \text{dom} f$

③ $\nabla^2 f(x) \succeq 0$, $\forall x \in \text{dom} f$

Goal: ①, ②, ③ are equivalent

① ⟺ ② then ② ⟺ ③ ⟹ ① ⟺ ② ⟺ ③

Proof: ① ⟺ ②

① ⟹ ②, if f is convex, by definition

$f(\lambda y + (1-\lambda)x) \leq \lambda f(y) + (1-\lambda) f(x)$, $\forall \lambda \in [0,1]$
$x, y \in \text{dom}(f)$

After rewritting, we have

$$f(x + \lambda(y-x)) \leq f(x) + \lambda(f(y) - f(x))$$

$$\Rightarrow f(y) - f(x) \geq \frac{f(x+\lambda(y-x)) - f(x)}{\lambda(y-x)} \cdot (y-x)$$

$\lambda \downarrow 0$ $\Downarrow$ $\nabla f(x)^T$

As $\lambda \downarrow 0$, we get

$$f(y) - f(x) \geq \nabla f(x)^T (y-x) \quad \text{--- 1)}$$

② ⟹ ①. suppose 1) holds, $\forall x, y \in dom(f)$

Take any $x, y \in dom(f)$, and let

$$z = \lambda x + (1-\lambda) y$$

$$f(x) \geq f(z) + \nabla f(z)^T (x - z) \quad \leftarrow 2)$$

$$f(y) \geq f(z) + \nabla f(z)^T (y - z) \quad \leftarrow 3)$$

Multiplying 2) by λ, 3) by $(1-\lambda)$, adding, we get

$$\lambda f(x) + (1-\lambda) f(y) \geq f(z) + \nabla f(z)^T \big[\underbrace{\lambda x + (1-\lambda) y - z}_{=0}\big]$$

$$= f(z)$$

$$= f(\lambda x + (1-\lambda) y)$$

---

② ⟺ ③, $n=1$ case

② ⟹ ③ $(n=1)$. let $x, y \in dom(f)$, $y > x$.

we have

$$f(y) \geq f(x) + f'(x)(y-x) \quad \text{— } 4)$$

$$f(x) \geq f(y) + f'(y)(x-y) \quad \text{— } 5)$$

$\Longrightarrow$

$$\underbrace{f'(x)(y-x)}_{LHS} \overset{4)}{\leq} f(y) - f(x) \overset{5)}{\leq} \underbrace{f'(y)(y-x)}_{RHS}$$

Dividing LHS and RHS by $(y-x)^2$

$$\frac{f'(y) - f'(x)}{y-x} \geq 0, \quad \forall x \neq y$$

Let $x \to y$ $\Longrightarrow$ $f''(x) \geq 0, \quad \forall x \in dom(f)$

$3) \Rightarrow 2)$ $(n=1)$ suppose $f''(x) \geq 0, \ \forall x \in dom(f)$

By the mean value version of Taylor's theorem

$$f(y) = f(x) + f'(x)(y-x) + \frac{1}{2}\underbrace{f''(z)}_{\geq 0}(y-x)^2$$

for some $z \in [x,y]$

$$\Downarrow$$

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

② $\Leftrightarrow$ ③ , $n > 1$

Recall : $f: \mathbb{R}^n \to \mathbb{R}$ is convex iff

$g(\alpha) = f(x_0 + \alpha v)$ is convex

$\forall x_0 \in dom(f)$, $dom(g) = \{\alpha \mid x_0 + t\alpha \in dom(f)\}$

we just prove this happens iff

$$g''(\alpha) = v^T \underbrace{D^2 f(x_0 + \alpha v)}_{\succeq 0} v \geq 0$$

$\forall x_0 \in dom(f)$, $\forall v \in \mathbb{R}^n$, α, s.t. $x_0 + \alpha v \in dom(f)$

Examples

"The matrix workbook" by Koone Brandt

❧ **Quadratic function:** $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{P}\boldsymbol{x} + \boldsymbol{q}^T\boldsymbol{x} + r$ (with $\boldsymbol{P} \in \mathbb{S}^n$)

$$\nabla f(\boldsymbol{x}) = \boldsymbol{P}\boldsymbol{x} + \boldsymbol{q}, \qquad \nabla^2 f(\boldsymbol{x}) = \boldsymbol{P}$$

is convex if $\boldsymbol{P} \succeq \boldsymbol{0}$.

❧ **Least-squares objective:** $f(\boldsymbol{x}) = \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2$

$$\nabla f(\boldsymbol{x}) = 2\boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}), \qquad \nabla^2 f(\boldsymbol{x}) = 2\boldsymbol{A}^T\boldsymbol{A}$$

is convex.

❧ **Quadratic-over-linear:** $f(x, y) = x^2/y$

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

is convex for $y > 0$.

$$\langle a, x\rangle = a^T x = \sum_i a_i x_i$$

$$\frac{\partial a^T x}{\partial x} = \frac{\partial \left(\sum_i a_i x_i\right)}{\partial x} = \begin{pmatrix} \frac{\partial\left(\sum_i a_i x_i\right)}{\partial x_i} \\ \vdots \\ \frac{\partial\left(\sum_i a_i x_i\right)}{\partial x_i} \end{pmatrix}$$

$$= \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = a \quad \text{— ①}$$

$$\frac{\partial x^T a}{\partial x} = a \quad \text{— ②}$$

$$f(x) = \|Ax - b\|_2^2 = (Ax-b)^T(Ax-b)$$

$$= x^T A^T A x - x^T A^T b - b^T A x + \|b\|_2^2$$

$$b^T A x \rightarrow (A^T b)^T x$$

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

$$= A^T A x + A^T A x$$

$$= 2A^T A x$$

Outline

Operations that Preserve Convexity I

How to establish the convexity of a given function?

- Applying the definition
- With first- or second-order conditions
- By restricting to a line
- Showing that the functions can be obtained from simple functions by operations that preserve convexity:
  - nonnegative weighted sum
  - composition with affine function (and other compositions)
  - pointwise maximum and supremum, minimization
  - perspective

Operations that Preserve Convexity II

- **Nonnegative weighted sum:** if f_1, f_2 are convex, then $\alpha_1 f_1 + \alpha_2 f_2$ is convex, with $\alpha_1, \alpha_2 \geq 0$.

- **Composition with affine functions:** if f is convex, then $f(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b})$ is convex (e.g., $\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|$ is convex, $\log\det(\boldsymbol{I} + \boldsymbol{H}\boldsymbol{X}\boldsymbol{H}^T)$ is concave).

$\|z\|$ channel capacity

- **Pointwise maximum:** $f := \max\{f_1, \cdots, f_m\}$ is convex, if $f_1, \cdots, f_m$ are convex

A

Example: sum of r largest components of $\boldsymbol{x} \in \mathbb{R}^n$:

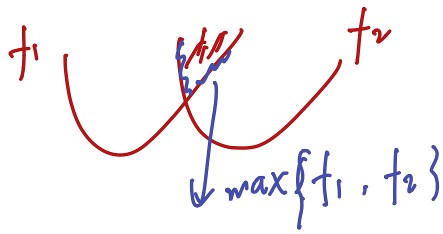
$$f(\boldsymbol{x}) = x_{[1]} + x_{[2]} + \cdots + x_{[r]}$$

where $x_{[i]}$ is the ith largest component of $\boldsymbol{x}$.

Proof: $f(\boldsymbol{x}) = \max\{\underbrace{x_{i_1} + x_{i_2} + \cdots + x_{i_r}}_{f_i(x):=}|1 \leq i_1 < i_2 < \cdots < i_r \leq n\}$.

$i \in A$, $i \in \{1, \ldots, C_n^r\} = A$

$f = \max_{i\in\{1,\dots,m\}} f_i$



Proof: $\forall\, x, y \in \mathrm{dom}(f)$, $\lambda \in (0,1]$. Then

$$f(\lambda x + (1-\lambda)y) = f_j(\lambda x + (1-\lambda)y), \text{ for some } j \in \{1,\dots,m\}$$
$$\le \lambda f_j(x) + (1-\lambda) f_j(y)$$
$$\le \lambda \underbrace{\max\{f_1(x),\dots,f_m(x)\}}_{f(x)} + (1-\lambda)\underbrace{\max\{f_1(y),\dots,f_m(y)\}}_{f(y)}$$

Proof via epigraphs

Recall: f is convex $\Leftrightarrow$ $epi(f)$ is a convex set

$$epi(f) = \bigcap_{i=1}^{m} epi(f_i) \Rightarrow \text{convex set}$$

$f = \max_i f_i$

fact: the intersection of convex set is convex

Operations that Preserve Convexity III

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

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

$$\mathrm{epi}(g) = \bigcap_{y \in A} \mathrm{epi}(f(y))$$

is convex.

Example: distance to farthest point in a set C:

$$f(\boldsymbol{x}) = \sup_{\boldsymbol{y} \in C} \|\boldsymbol{x} - \boldsymbol{y}\| \quad 4$$

Example: maximum eigenvalue of symmetric matrix: for $\boldsymbol{X} \in \mathbb{S}^n$,

$$\lambda_{\max}(\boldsymbol{X}) = \sup_{\boldsymbol{y} \neq \boldsymbol{0}} \frac{\boldsymbol{y}^T \boldsymbol{X} \boldsymbol{y}}{\boldsymbol{y}^T \boldsymbol{y}}$$

$f(X, y) = \frac{y^T X y}{y^T y}$ is convex in X given y

Proof: $n=1$, g. h differentiable

$$f''(x) = \underbrace{h''(g(x))}_{\geq 0}\underbrace{g'(x)^2}_{\geq 0}$$

$$+ \underbrace{h'(g(x))}_{\geq 0}\underbrace{g''(x)}_{\geq 0}$$

nondecreasing

Operations that Preserve Convexity IV

- **Composition with scalar functions:** let $g : \mathbb{R}^n \longrightarrow \mathbb{R}$, $h : \mathbb{R} \longrightarrow \mathbb{R}$, then the function $f(\boldsymbol{x}) = h(g(\boldsymbol{x}))$ satisfies:

$$f(\boldsymbol{x}) \text{ is convex if } \begin{array}{l} g \text{ convex, } h \text{ convex nondecreasing} \\ g \text{ concave, } h \text{ convex nonincreasing} \end{array}$$

- **Minimization:** if $f(\boldsymbol{x}, \boldsymbol{y})$ is convex in $(\boldsymbol{x}, \boldsymbol{y})$ and C is a convex set, then

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

is convex (e.g., distance to a convex set).

Example: distance to a set C:

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

A

is convex if C is convex.

Assume the infimum over $y \in C$ is attained for each x, we have

$$\operatorname{epi} g = \{(x, t) \mid (x, y, t) \in \operatorname{epi} f, \text{ for some } y \in C\}$$

the projection of a convex set on some of its components.

Operations that Preserve Convexity V

- **Perspective:** if $f(\boldsymbol{x})$ is convex, then its perspective

$$g(\boldsymbol{x}, t) = tf(\boldsymbol{x}/t), \quad \operatorname{dom} g = \{(\boldsymbol{x}, t) \in \mathbb{R}^{n+1} | \boldsymbol{x}/t \in \operatorname{dom} f, t > 0\}$$

  is convex.

$$t \cdot \left(\frac{x}{t}\right)^T \left(\frac{x}{t}\right) = \frac{x^T x}{t}$$

Example: $f(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{x}$ is convex; hence $g(\boldsymbol{x}, t) = \boldsymbol{x}^T\boldsymbol{x}/t$ is convex for $t > 0$.

Example: the negative logarithm $f(\boldsymbol{x}) = -\log \boldsymbol{x}$ is convex; hence the relative entropy function $g(\boldsymbol{x}, t) = t \log t - t \log \boldsymbol{x}$ is convex on $\mathbb{R}^2_{++}$.

$$(x, t, s) \in \operatorname{epi} g \iff g(x,t) = \underbrace{t f\left(\frac{x}{t}\right) \le s}$$

$$\iff \underbrace{f\left(\frac{x}{t}\right) \le \frac{s}{t}}$$

$$\iff \underline{\left(\frac{x}{t}, \frac{s}{t}\right)} \in \operatorname{epi} f$$

$\Rightarrow (u, v, w)$ to $(u, w)/v$

Outline

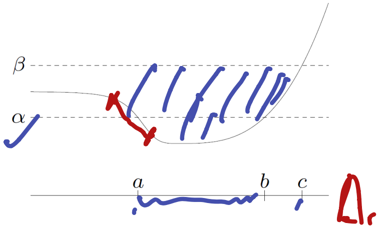
Quasi-Convexity Functions

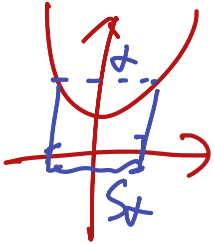
- A function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is quasi-convex if $\mathrm{dom}\, f$ is convex and the sublevel sets

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

are convex for all α.

given





- f is quasiconcave if $-f$ is quasiconvex.

Proof: $x, y \in S_\alpha$, $\lambda \in (0,1)$

$x \in S_\alpha \Rightarrow f(x) \leq \alpha$ — ①

$y \in S_\alpha \Rightarrow f(y) \leq \alpha$ — ②

f convex $\Rightarrow f(\lambda x + (1-\lambda) y)$

$$\leq \lambda \underbrace{f(x)}_{\alpha} + (1-\lambda) \underbrace{f(y)}_{\alpha}$$

$$= \alpha$$

$\lambda x + (1-\lambda) y \in S_\alpha$

convexity $\Rightarrow$ quasi-convexity

$\not\Leftarrow$

Examples

- $\sqrt{|x|}$ is quasiconvex on $\mathbb{R}$
- $\text{ceil}(x) = \inf\{z \in \mathbb{Z} \mid z \geq x\}$ is quasilinear
- $\log x$ is quasilinear on $\mathbb{R}_{++}$
- $f(x_1, x_2) = x_1 x_2$ is quasiconcave on $\mathbb{R}^2_{++}$

bi-section

- the linear-fractional function

[l, u]

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

is quasilinear

$S_\alpha = \{x \mid f(x) \leq \alpha\}$

Log-Convexity

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

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

- f is log-convex if $\log f$ is convex.
- Example: x^a on $\mathbb{R}_{++}$ is log-convex for $a \leq 0$ and log-concave for $a \geq 0$
- Example: many common probability densities are log-concave

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

Convexity w.r.t. Generalized Inequalities

- $\boldsymbol{f} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is K-convex if $\mathrm{dom}\, \boldsymbol{f}$ is convex and for any $\boldsymbol{x}, \boldsymbol{y} \in \mathrm{dom}\, \boldsymbol{f}$ and $0 \leq \theta \leq 1$,

$$\underbrace{\boldsymbol{f}(\theta \boldsymbol{x} + (1-\theta)\boldsymbol{y})}_{b} \preceq_K \underbrace{\theta \boldsymbol{f}(\boldsymbol{x}) + (1-\theta)\boldsymbol{f}(\boldsymbol{y})}_{a}$$

- Example: $\boldsymbol{f} : \mathbb{S}^m \longrightarrow \mathbb{S}^m$, $\boldsymbol{f}(\boldsymbol{X}) = \boldsymbol{X}^2$ is $\mathbb{S}^m_+$-convex

$a - b \in K$

Reference

Chapter 3 of:

- Stephen Boyd and Lieven Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge University Press, 2004.

Book:

- Petersen, Kaare Brandt, and Michael Syskind Pedersen. “The matrix cookbook." Technical University of Denmark 7 (2008): 15.