ELG6108: Introduction to Convex Optimization

Lecture 3: Convex Functions

Dr. Sergey Loyka

EECS, University of Ottawa

February 10, 2021

Convex functions[1]

- Definition
- Examples
- 1st order condition
- 2nd order condition
- Operations that preserve convexity
- Quasiconvex functions
- Log-concave and log-convex functions
- Generalized inequalities

[1] adapted from Boyd & Vandenberghe, Convex Optimization, Lecture slides.

Definition of convex/concave function

- $f(\mathbf{x})$ is convex if **dom** f is a convex set and

$$f(\theta\mathbf{x} + (1-\theta)\mathbf{y}) \leq \theta f(\mathbf{x}) + (1-\theta)f(\mathbf{y}) \tag{1}$$

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

Definition of convex/concave function

- $f(\mathbf{x})$ is convex if **dom** f is a convex set and

$$f(\theta \mathbf{x} + (1-\theta)\mathbf{y}) \leq \theta f(\mathbf{x}) + (1-\theta) f(\mathbf{y}) \tag{1}$$

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

- geometrically: any line segment is above the graph

Definition of convex/concave function

- $f(\mathbf{x})$ is convex if **dom** f is a convex set and

$$f(\theta \mathbf{x} + (1-\theta)\mathbf{y}) \leq \theta f(\mathbf{x}) + (1-\theta) f(\mathbf{y}) \tag{1}$$

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

- geometrically: any line segment is above the graph
- strictly convex: if the inequality is strict for any $\mathbf{x} \neq \mathbf{y}$ and $0 < \theta < 1$

Definition of convex/concave function

- $f(\mathbf{x})$ is convex if **dom** f is a convex set and

$$f(\theta\mathbf{x} + (1-\theta)\mathbf{y}) \leq \theta f(\mathbf{x}) + (1-\theta) f(\mathbf{y}) \quad (1)$$

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

- geometrically: any line segment is above the graph
- strictly convex: if the inequality is strict for any $\mathbf{x} \neq \mathbf{y}$ and $0 < \theta < 1$
- f is concave if $-f$ is convex (equivalently: opposite inequality)

Examples: convex $f(x)$ of scalar x

- **quadratic**: x^2 (most simple, my favorite)

Examples: convex $f(x)$ of scalar x

- **quadratic**: x^2 (most simple, my favorite)
- affine (linear): $ax + b$ for any a, b (convex and concave sim.)
- exponential: e^{ax}, for any a
- powers: x^α for $x > 0$, $\alpha \geq 1$ or $\alpha \leq 0$
- powers of absolute value: $|x|^p$ for $p \geq 1$
- negative entropy: $x \log x$ for $x > 0$

Examples: concave $f(x)$ of scalar x

- **negative quadratic**: $-x^2$ (most simple, my favorite)

Examples: concave $f(x)$ of scalar x

- **negative quadratic**: $-x^2$ (most simple, my favorite)
- affine (linear): $ax + b$ for any a, b
- powers: x^α for $x > 0$ and $0 \leq \alpha \leq 1$
- logarithm: $\log x$ for $x > 0$

Examples: vector argument $\mathbf{x}$

- affine functions are convex and concave; all norms are convex

Examples: vector argument $\mathbf{x}$

- affine functions are convex and concave; all norms are convex

- Euclidean norm = length $|\mathbf{x}| = |\mathbf{x}|_2$
- affine function $f(\mathbf{x}) = \mathbf{a}^T\mathbf{x} + b$
- l_p norms: $|\mathbf{x}|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$ for $p \geq 1$
- $|x|_\infty = \max_k |x_k|$

Examples: convex $f(\mathbf{X})$ of matrix $\mathbf{X}$

- trace: $f(\mathbf{X}) = tr(\mathbf{X})$ for any $\mathbf{X}$ (convex and concave)
- affine function for any $\mathbf{X}$ (convex and concave)

$$f(\mathbf{X}) = \mathrm{tr}(\mathbf{A}^T\mathbf{X}) + b = \sum_{i,j} a_{ij}x_{ij} + b \tag{2}$$

- max. eigenvalue: $f(\mathbf{X}) = \lambda_{max}(\mathbf{X})$ for $\mathbf{X}^T = \mathbf{X}$
- spectral norm (max. singular value) for any $\mathbf{X}$

$$f(\mathbf{X}) = |\mathbf{X}|_2 = \sigma_{max}(\mathbf{X}) = (\lambda_{max}(\mathbf{X}^T\mathbf{X}))^{1/2} \tag{3}$$

Examples: concave $f(\mathbf{X})$ of matrix $\mathbf{X}$

- trace: $f(\mathbf{X}) = tr(\mathbf{X})$ for any $\mathbf{X}$ (convex and concave)
- min. eigenvalue: $f(\mathbf{X}) = \lambda_{min}(\mathbf{X})$ for $\mathbf{X}^T = \mathbf{X}$
- log-det: $f(\mathbf{X}) = \log|\mathbf{X}|$ for $\mathbf{X} > 0$

Restriction to a line

- makes it simple to check convexity in many cases
- $f(\mathbf{x})$ is convex if and only if $g(t)$ is convex:

$$g(t) = f(\mathbf{x} + t\mathbf{y}) \tag{4}$$

  for any $\mathbf{x}, \mathbf{y}, t$ such that $(\mathbf{x} + t\mathbf{y}) \in \mathbf{dom}\ f$
- same applies to $f(\mathbf{X})$
- note that $g(t)$ is simpler than $f(\mathbf{x})$: t - scalar, but $\mathbf{x}$ - vector
- can check convexity of $f(\mathbf{x})$ by checking convexity of $g(t)$

Example: $f(\mathbf{X}) = \log|\mathbf{X}|,\ \mathbf{X} > 0$

$$g(t) = \log|\mathbf{X} + t\mathbf{Y}| \tag{5}$$

$$= \log|\mathbf{X}| + \log|\mathbf{I} + t\mathbf{X}^{-1/2}\mathbf{Y}\mathbf{X}^{-1/2}| \tag{6}$$

$$= \log|\mathbf{X}| + \sum_i \log(1 + t\lambda_i) \tag{7}$$

$\lambda_i = \lambda_i(\mathbf{X}^{-1/2}\mathbf{Y}\mathbf{X}^{-1/2})$ are the eigenvalues

$g(t)$ is concave (why?), for any $t, \mathbf{X}, \mathbf{Y}$ such that $\mathbf{X} + t\mathbf{Y} > 0$

hence, $f(\mathbf{X})$ is also concave

First-order condition

Assume $f(\mathbf{x})$ is differentiable, the gradient $\nabla f(\mathbf{x})$

$$\nabla f(\mathbf{x}) = \left(\frac{\partial f(\mathbf{x})}{\partial x_1}, \frac{\partial f(\mathbf{x})}{\partial x_2}, ..., \frac{\partial f(\mathbf{x})}{\partial x_n} \right)^T \tag{8}$$

exists for each $\mathbf{x} \in \textbf{dom } f$

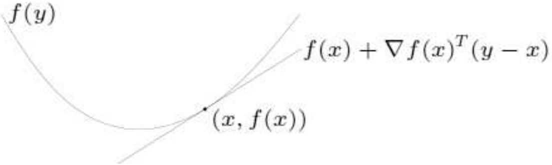
1st-order condition: differentiable $f(\mathbf{x})$ with convex domain is convex iff

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) \quad \text{for all } \mathbf{x}, \mathbf{y} \in \textbf{dom } f \tag{9}$$

First-order condition

Differentiable $f(\mathbf{x})$ with convex domain is convex iff

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



Geometry: first-order approximation of $f(\mathbf{x})$ is its **global underestimator**

Second-order condition

Twice differentiable $f(\mathbf{x})$, Hessian $\mathbf{H} = \nabla^2 f(\mathbf{x})$ exists at each $\mathbf{x} \in \mathbf{dom}\ f$,

$$\mathbf{H} = \nabla^2 f(\mathbf{x}) = \frac{\partial^2 f(\mathbf{x})}{\partial \mathbf{x}^T \partial \mathbf{x}} = \left\{ \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} \right\} \tag{10}$$

2nd-order conditions: for twice differentiable $f(\mathbf{x})$ with convex domain

- $f(\mathbf{x})$ is convex if and only if

$$\mathbf{H} = \nabla^2 f(\mathbf{x}) \geq 0 \text{ for all } \mathbf{x} \in \mathbf{dom} f \tag{11}$$

- strictly convex if $\nabla^2 f(\mathbf{x}) > 0$

Examples

- **quadratic function:** $f(\mathbf{x}) = (1/2)\mathbf{x}^T\mathbf{P}\mathbf{x} + \mathbf{q}^T\mathbf{x} + r$

$$\nabla f(\mathbf{x}) = \mathbf{P}\mathbf{x} + \mathbf{q}, \quad \nabla^2 f(\mathbf{x}) = \mathbf{P} \tag{12}$$

convex if $\mathbf{P} \succeq 0$, concave if $\mathbf{P} \leq 0$

Examples

- **quadratic function:** $f(\mathbf{x}) = (1/2)\mathbf{x}^T\mathbf{P}\mathbf{x} + \mathbf{q}^T\mathbf{x} + r$

$$\nabla f(\mathbf{x}) = \mathbf{P}\mathbf{x} + \mathbf{q}, \quad \nabla^2 f(\mathbf{x}) = \mathbf{P} \tag{12}$$

  convex if $\mathbf{P} \succeq 0$, concave if $\mathbf{P} \leq 0$

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

$$\nabla f(\mathbf{x}) = 2\mathbf{A}^T(\mathbf{A}\mathbf{x} - \mathbf{b}), \quad \nabla^2 f(\mathbf{x}) = 2\mathbf{A}^T\mathbf{A} \geq 0 \tag{13}$$

  convex for any $\mathbf{A}$ (even non-square)

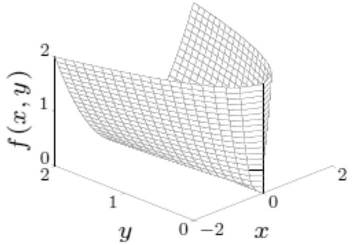
Examples

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

Examples

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

$$\nabla^2 f(x,y) = \frac{2}{y^3} \begin{bmatrix} y \\ -x \end{bmatrix} \begin{bmatrix} y \\ -x \end{bmatrix}^T \succeq 0 \quad (14)$$



Examples

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

Examples

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

$$\nabla^2 f(\mathbf{x}) = \frac{1}{\mathbf{1}^T\mathbf{z}} \operatorname{diag}(\mathbf{z}) - \frac{1}{\left(\mathbf{1}^T\mathbf{z}\right)^2}\mathbf{z}\mathbf{z}^T \quad (z_k = \exp x_k) \tag{15}$$

Proof: show that $\nabla^2 f(\mathbf{x}) \succeq 0$ via $\mathbf{v}^T\nabla^2 f(\mathbf{x})\mathbf{v} \geq 0$ for all $\mathbf{v}$:

$$\mathbf{v}^T\nabla^2 f(\mathbf{x})\mathbf{v} = \frac{\left(\sum_k z_k v_k^2\right)\left(\sum_k z_k\right) - \left(\sum_k v_k z_k\right)^2}{\left(\sum_k z_k\right)^2} \geq 0 \tag{16}$$

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

Examples

- **geometric mean:** $f(\mathbf{x}) = \left(\prod_{k=1}^{n} x_k\right)^{1/n}$ is concave for $\{x_k > 0, \ \forall k\}$ (similar proof as for log-sum-exp)

Sublevel set

- α-sublevel set of f:

$$C_\alpha = \{\mathbf{x} \in \operatorname{dom} f : f(\mathbf{x}) \leq \alpha\} \tag{17}$$

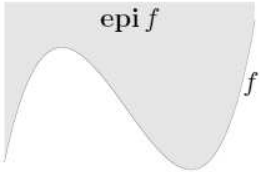
i.e. the set of points where function values do not exceed given level α: $f(\mathbf{x}) \leq \alpha$

- sublevel sets of convex functions are convex (converse is false)

Epigraph

- the set of points above the function's graph
- epigraph of $f(\mathbf{x})$:

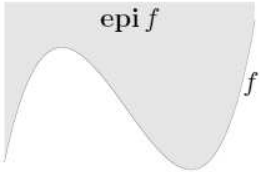
$$\textbf{epi } f = \{(\mathbf{x}, t) : \mathbf{x} \in \text{dom } f, f(\mathbf{x}) \leq t\} \tag{18}$$



Epigraph

- the set of points above the function's graph
- epigraph of $f(\mathbf{x})$:

$$\mathbf{epi}\ f = \{(\mathbf{x}, t) : \mathbf{x} \in \text{dom}\, f, f(\mathbf{x}) \leq t\} \tag{18}$$



- $f(\mathbf{x})$ is convex if and only if **epi** f is a convex set

Jensen's inequality

- the definition of convexity of $f(\mathbf{x})$: for $0 \leq \theta \leq 1$,

$$f(\theta \mathbf{x} + (1-\theta)\mathbf{y}) \leq \theta f(\mathbf{x}) + (1-\theta) f(\mathbf{y}) \tag{19}$$

- extension: if $f(\mathbf{x})$ is convex, then

$$f(\mathbb{E}\mathbf{z}) \leq \mathbb{E} f(\mathbf{z}) \tag{20}$$

for any random vector $\mathbf{z}$; $\mathbb{E}\{\cdot\}$ is statistical expectation

Jensen's inequality

- the definition of convexity of $f(\mathbf{x})$: for $0 \le \theta \le 1$,

$$f(\theta\mathbf{x} + (1-\theta)\mathbf{y}) \le \theta f(\mathbf{x}) + (1-\theta) f(\mathbf{y}) \tag{19}$$

- extension: if $f(\mathbf{x})$ is convex, then

$$f(\mathbb{E}\mathbf{z}) \le \mathbb{E} f(\mathbf{z}) \tag{20}$$

  for any random vector $\mathbf{z}$; $\mathbb{E}\{\cdot\}$ is statistical expectation
- the definition is special case with discrete distributions

$$\Pr(\mathbf{z} = \mathbf{x}) = \theta, \quad \Pr(\mathbf{z} = \mathbf{y}) = 1 - \theta \tag{21}$$

Jensen's inequality

- powerful applications
  - communications
  - information theory
  - signal processing
  - control, etc.
- examples:
  - entropy/mutual information/channel capacity
  - error rate in fading channels
  - mean square error

How to establish convexity ?

How to establish convexity ?

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

How to establish convexity ?

1. Use definition (often simplified by restricting to a line)
2. Use $\nabla^2 f(\mathbf{x}) \succeq 0$ (for twice differentiable functions)

How to establish convexity ?

1. Use definition (often simplified by restricting to a line)
2. Use $\nabla^2 f(\mathbf{x}) \succeq 0$ (for twice differentiable functions)
3. Show that $f(\mathbf{x})$ is obtained from simple convex functions by operations that preserve convexity

How to establish convexity ?

1. Use definition (often simplified by restricting to a line)
2. Use $\nabla^2 f(\mathbf{x}) \succeq 0$ (for twice differentiable functions)
3. Show that $f(\mathbf{x})$ is obtained from simple convex functions by operations that preserve convexity

Some methods may be much simpler than others (example: $f(x) = x^2$)

Convexity-preserving operations

- nonnegative weighted sum
- composition with affine function
- pointwise maximum/supremum
- composition
- minimization
- perspective

Positive weighted sum

- **nonnegative multiple:** $\alpha f(\mathbf{x})$ is convex if $f(\mathbf{x})$ is convex and $\alpha \geq 0$
- **sum:** $f_1(\mathbf{x}) + f_2(\mathbf{x})$ convex if $f_1(\mathbf{x}), f_2(\mathbf{x})$ convex (extends to infinite sums, integrals)
- **positive weighted sum:** convex if $f_i(\mathbf{x})$ are convex and $\alpha_i \geq 0$

$$\sum_i \alpha_i f_i(\mathbf{x}),\ \alpha_i \geq 0 \tag{22}$$

- also extends to infinite sums and integrals

Composition with affine function

- **composition with affine function:** $f(\mathbf{Ax}+\mathbf{b})$ is convex if f is convex
  - proof?

Composition with affine function

- **composition with affine function:** $f(\mathbf{Ax}+\mathbf{b})$ is convex if f is convex
  - proof?
- **examples**
  - log barrier for linear inequalities

$$f(\mathbf{x}) = -\sum_{i=1}^{m} \log\left(b_i - \mathbf{a}_i^T\mathbf{x}\right), \quad \mathbf{dom}\ f = \{\mathbf{x} : \mathbf{a}_i^T\mathbf{x} < b_i\ \forall i\} \tag{23}$$

Composition with affine function

- **composition with affine function:** $f(\mathbf{A}\mathbf{x}+\mathbf{b})$ is convex if f is convex
  - proof?
- **examples**
  - log barrier for linear inequalities

$$f(\mathbf{x}) = -\sum_{i=1}^{m} \log\left(b_i - \mathbf{a}_i^T\mathbf{x}\right), \quad \textbf{dom } f = \{\mathbf{x} : \mathbf{a}_i^T\mathbf{x} < b_i \ \forall i\} \qquad (23)$$

  - any norm of affine function: $f(\mathbf{x}) = \|\mathbf{A}\mathbf{x}+\mathbf{b}\|$
  - proof?

Pointwise maximum

- if $f_1, \ldots, f_m$ are convex, then

$$f(\mathbf{x}) = \max\{f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})\} \quad (24)$$

is convex

  - proof: via $\max\{a_1 + b_1, a_2 + b_2\} \leq \max\{a_1, a_2\} + \max\{b_1, b_2\}$

Pointwise maximum

- if $f_1, \ldots, f_m$ are convex, then

$$f(\mathbf{x}) = \max\{f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})\} \tag{24}$$

  is convex
  - proof: via $\max\{a_1 + b_1, a_2 + b_2\} \leq \max\{a_1, a_2\} + \max\{b_1, b_2\}$
- examples
  - piecewise-linear function: $f(\mathbf{x}) = \max_{i=1,\ldots,m}\left(\mathbf{a}_i^T\mathbf{x} + b_i\right)$

Pointwise maximum: examples

- sum of r largest components of $\mathbf{x}$:

$$f(\mathbf{x}) = x_{[1]} + x_{[2]} + \cdots + x_{[r]} \quad (25)$$

$x_{[i]}$ is i-th largest component of $\mathbf{x}$,

$$x_{[1]} \geq x_{[2]} \geq ... \geq x_{[r]} \quad (26)$$

Pointwise maximum: examples

- sum of r largest components of $\mathbf{x}$:

$$f(\mathbf{x}) = x_{[1]} + x_{[2]} + \cdots + x_{[r]} \tag{25}$$

$x_{[i]}$ is i-th largest component of $\mathbf{x}$,

$$x_{[1]} \geq x_{[2]} \geq ... \geq x_{[r]} \tag{26}$$

- proof: via

$$f(x) = \max\{x_{i_1} + x_{i_2} + \cdots + x_{i_r} \mid 1 \leq i_1 < i_2 < \cdots < i_r \leq n\} \tag{27}$$

Pointwise maximum: examples

- sum of r largest components of $\mathbf{x}$:

$$f(\mathbf{x}) = x_{[1]} + x_{[2]} + \cdots + x_{[r]} \tag{25}$$

$x_{[i]}$ is i-th largest component of $\mathbf{x}$,

$$x_{[1]} \geq x_{[2]} \geq ... \geq x_{[r]} \tag{26}$$

- proof: via

$$f(x) = \max\{x_{i_1} + x_{i_2} + \cdots + x_{i_r} \mid 1 \leq i_1 < i_2 < \cdots < i_r \leq n\} \tag{27}$$

- example: $f(\mathbf{x}) = x_{[1]}$

Pointwise maximum: examples

- sum of r largest components of $\mathbf{x}$:

$$f(\mathbf{x}) = x_{[1]} + x_{[2]} + \cdots + x_{[r]} \tag{25}$$

$x_{[i]}$ is i-th largest component of $\mathbf{x}$,

$$x_{[1]} \geq x_{[2]} \geq ... \geq x_{[r]} \tag{26}$$

- proof: via

$$f(x) = \max\{x_{i_1} + x_{i_2} + \cdots + x_{i_r} \mid 1 \leq i_1 < i_2 < \cdots < i_r \leq n\} \tag{27}$$

- example: $f(\mathbf{x}) = x_{[1]}$
- Q: what about smallest component?

Pointwise maximum (supremum)

- if $f(\mathbf{x}, \mathbf{y})$ is convex in $\mathbf{x}$ for each $\mathbf{y} \in \mathcal{A}$, then $g(\mathbf{x})$ is also convex,

$$g(\mathbf{x}) = \max_{\mathbf{y} \in \mathcal{A}} f(\mathbf{x}, \mathbf{y}) \tag{28}$$

Pointwise maximum (supremum)

- if $f(\mathbf{x},\mathbf{y})$ is convex in $\mathbf{x}$ for each $\mathbf{y} \in \mathcal{A}$, then $g(\mathbf{x})$ is also convex,

$$g(\mathbf{x}) = \max_{\mathbf{y}\in\mathcal{A}} f(\mathbf{x},\mathbf{y}) \tag{28}$$

- **examples**
  - support function of a set C : $S_C(\mathbf{x}) = \max_{\mathbf{y}\in C} \mathbf{y}^T\mathbf{x}$

Pointwise maximum (supremum)

- if $f(\mathbf{x}, \mathbf{y})$ is convex in $\mathbf{x}$ for each $\mathbf{y} \in \mathcal{A}$, then $g(\mathbf{x})$ is also convex,

$$g(\mathbf{x}) = \max_{\mathbf{y} \in \mathcal{A}} f(\mathbf{x}, \mathbf{y}) \quad (28)$$

- **examples**
  - support function of a set C : $S_C(\mathbf{x}) = \max_{\mathbf{y} \in C} \mathbf{y}^T \mathbf{x}$
  - distance to farthest point in a set C :

$$f(\mathbf{x}) = \max_{\mathbf{y} \in C} |\mathbf{x} - \mathbf{y}| \quad (29)$$

Pointwise maximum (supremum)

- if $f(\mathbf{x},\mathbf{y})$ is convex in $\mathbf{x}$ for each $\mathbf{y} \in \mathcal{A}$, then $g(\mathbf{x})$ is also convex,

$$g(\mathbf{x}) = \max_{\mathbf{y}\in\mathcal{A}} f(\mathbf{x},\mathbf{y}) \tag{28}$$

- **examples**
  - support function of a set C : $S_C(\mathbf{x}) = \max_{\mathbf{y}\in C} \mathbf{y}^T\mathbf{x}$
  - distance to farthest point in a set C :

$$f(\mathbf{x}) = \max_{\mathbf{y}\in C} |\mathbf{x}-\mathbf{y}| \tag{29}$$

  - maximum eigenvalue of symmetric matrix $\mathbf{X} = \mathbf{X}^T$:

$$\lambda_1(\mathbf{X}) = \max_{|\mathbf{y}|_2=1} \mathbf{y}^T\mathbf{X}\mathbf{y} \tag{30}$$

Pointwise maximum (supremum)

- if $f(\mathbf{x}, \mathbf{y})$ is convex in $\mathbf{x}$ for each $\mathbf{y} \in \mathcal{A}$, then $g(\mathbf{x})$ is also convex,

$$g(\mathbf{x}) = \max_{\mathbf{y} \in \mathcal{A}} f(\mathbf{x}, \mathbf{y}) \tag{28}$$

- **examples**
  - support function of a set C : $S_C(\mathbf{x}) = \max_{\mathbf{y} \in C} \mathbf{y}^T \mathbf{x}$
  - distance to farthest point in a set C :

$$f(\mathbf{x}) = \max_{\mathbf{y} \in C} |\mathbf{x} - \mathbf{y}| \tag{29}$$

  - maximum eigenvalue of symmetric matrix $\mathbf{X} = \mathbf{X}^T$:

$$\lambda_1(\mathbf{X}) = \max_{|\mathbf{y}|_2 = 1} \mathbf{y}^T \mathbf{X} \mathbf{y} \tag{30}$$

  - Q: what about minimum eigenvalue? 2nd largest?

Composition with scalar functions

- composition of $g(\mathbf{x})$ and $h(y)$:

$$f(\mathbf{x}) = h(g(\mathbf{x})) \tag{31}$$

- $f(\mathbf{x})$ is convex if:
  - $g(\mathbf{x})$ convex, $h(y)$ convex and nondecreasing
  - $g(\mathbf{x})$ concave, $h(y)$ convex and nonincreasing

Composition with scalar functions

- composition of $g(\mathbf{x})$ and $h(y)$:

$$f(\mathbf{x}) = h(g(\mathbf{x})) \tag{31}$$

- $f(\mathbf{x})$ is convex if:
  - $g(\mathbf{x})$ convex, $h(y)$ convex and nondecreasing
  - $g(\mathbf{x})$ concave, $h(y)$ convex and nonincreasing
- proof for scalar x, differentiable g, h:

$$f''(x) = h''(g(x))g'(x)^2 + h'(g(x))g''(x) \tag{32}$$

Composition with scalar functions

- composition of $g(\mathbf{x})$ and $h(y)$:

$$f(\mathbf{x}) = h(g(\mathbf{x})) \tag{31}$$

- $f(\mathbf{x})$ is convex if:
  - $g(\mathbf{x})$ convex, $h(y)$ convex and nondecreasing
  - $g(\mathbf{x})$ concave, $h(y)$ convex and nonincreasing
- proof for scalar x, differentiable g, h:

$$f''(x) = h''(g(x))g'(x)^2 + h'(g(x))g''(x) \tag{32}$$

examples

- $e^{g(\mathbf{x})}$ is convex if $g(\mathbf{x})$ is convex
- $1/g(\mathbf{x})$ is convex if $g(\mathbf{x})$ is concave and positive

Vector composition

composition of $\mathbf{g}(\mathbf{x})$ and $h(\mathbf{y})$:

$$f(\mathbf{x}) = h(\mathbf{g}(\mathbf{x})) = h(g_1(\mathbf{x}), g_2(\mathbf{x}), \dots, g_k(\mathbf{x})) \tag{33}$$

$f(\mathbf{x})$ is convex if:

- $g_i(\mathbf{x})$ convex, $h(\mathbf{y})$ convex, h nondecreasing in each argument
- $g_i(\mathbf{x})$ concave, $h(\mathbf{y})$ convex, h nonincreasing in each argument

Vector composition

composition of $\mathbf{g}(\mathbf{x})$ and $h(\mathbf{y})$:

$$f(\mathbf{x}) = h(\mathbf{g}(\mathbf{x})) = h(g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_k(\mathbf{x})) \tag{33}$$

$f(\mathbf{x})$ is convex if:

- $g_i(\mathbf{x})$ convex, $h(\mathbf{y})$ convex, h nondecreasing in each argument
- $g_i(\mathbf{x})$ concave, $h(\mathbf{y})$ convex, h nonincreasing in each argument

proof for scalar x, differentiable g, h

$$f''(x) = \mathbf{g}'(x)^T \nabla^2 h(g(x)) \mathbf{g}'(x) + \nabla h(\mathbf{g}(x))^T \mathbf{g}''(x) \tag{34}$$

Vector composition

composition of $\mathbf{g}(\mathbf{x})$ and $h(\mathbf{y})$:

$$f(\mathbf{x}) = h(\mathbf{g}(\mathbf{x})) = h(g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_k(\mathbf{x})) \tag{33}$$

$f(\mathbf{x})$ is convex if:

- $g_i(\mathbf{x})$ convex, $h(\mathbf{y})$ convex, h nondecreasing in each argument
- $g_i(\mathbf{x})$ concave, $h(\mathbf{y})$ convex, h nonincreasing in each argument

proof for scalar x, differentiable g, h

$$f''(x) = \mathbf{g}'(x)^T \nabla^2 h(g(x))\mathbf{g}'(x) + \nabla h(\mathbf{g}(x))^T \mathbf{g}''(x) \tag{34}$$

examples:

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

Minimization

If $f(\mathbf{x}, \mathbf{y})$ is (jointly) convex in $(\mathbf{x}, \mathbf{y})$ and C is a convex set, then $g(\mathbf{x})$ is convex,

$$g(\mathbf{x}) = \min_{\mathbf{y} \in C} f(\mathbf{x}, \mathbf{y}) \tag{35}$$

Minimization

If $f(\mathbf{x},\mathbf{y})$ is (jointly) convex in $(\mathbf{x},\mathbf{y})$ and C is a convex set, then $g(\mathbf{x})$ is convex,

$$g(\mathbf{x}) = \min_{\mathbf{y}\in C} f(\mathbf{x},\mathbf{y}) \tag{35}$$

examples

- distance to a set S: convex if S is convex,

$$\text{dist}(\mathbf{x},S) = \min_{\mathbf{y}\in S} |\mathbf{x}-\mathbf{y}| \tag{36}$$

Minimization

examples

- distance to a set S: convex if S is convex,

$$\text{dist}(\mathbf{x}, S) = \min_{\mathbf{y}\in S} |\mathbf{x} - \mathbf{y}| \tag{37}$$

- $f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T\mathbf{A}\mathbf{x} + 2\mathbf{x}^T\mathbf{B}\mathbf{y} + \mathbf{y}^T\mathbf{C}\mathbf{y}$ with

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{bmatrix} \succeq 0, \quad \mathbf{C} \succ 0 \tag{38}$$

minimizing over $\mathbf{y}$ gives

$$g(\mathbf{x}) = \min_{\mathbf{y}} f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \left(\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T\right) \mathbf{x} \tag{39}$$

Since $g(\mathbf{x})$ is convex, Schur complement $\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^T \succeq 0$

Perspective

the **perspective** of a function f is the function g,

$$g(\mathbf{x}, t) = tf(\mathbf{x}/t), \quad \operatorname{dom} g = \{(\mathbf{x}, t) \mid \mathbf{x}/t \in \operatorname{dom} f, t > 0\} \tag{40}$$

$g(\mathbf{x}, t)$ is convex if $f(\mathbf{x})$ is convex

Perspective

the **perspective** of a function f is the function g,

$$g(\mathbf{x}, t) = tf(\mathbf{x}/t), \quad \operatorname{dom} g = \{(\mathbf{x}, t) \mid \mathbf{x}/t \in \operatorname{dom} f, t > 0\} \tag{40}$$

$g(\mathbf{x}, t)$ is convex if $f(\mathbf{x})$ is convex

examples

- $f(\mathbf{x}) = \mathbf{x}^T\mathbf{x}$ is convex; hence $g(\mathbf{x}, t) = \mathbf{x}^T\mathbf{x}/t$ is convex for $t > 0$

Perspective

the **perspective** of a function f is the function g,

$$g(\mathbf{x}, t) = tf(\mathbf{x}/t), \quad \operatorname{dom} g = \{(\mathbf{x}, t) \mid \mathbf{x}/t \in \operatorname{dom} f, t > 0\} \tag{40}$$

$g(\mathbf{x}, t)$ is convex if $f(\mathbf{x})$ is convex

examples

- $f(\mathbf{x}) = \mathbf{x}^T\mathbf{x}$ is convex; hence $g(\mathbf{x}, t) = \mathbf{x}^T\mathbf{x}/t$ 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 for $x, t > 0$

Perspective

the **perspective** of a function f is the function g,

$$g(\mathbf{x}, t) = tf(\mathbf{x}/t), \quad \mathrm{dom}\, g = \{(\mathbf{x}, t) \mid \mathbf{x}/t \in \mathrm{dom}\, f, t > 0\} \tag{40}$$

$g(\mathbf{x}, t)$ is convex if $f(\mathbf{x})$ is convex

examples

- $f(\mathbf{x}) = \mathbf{x}^T\mathbf{x}$ is convex; hence $g(\mathbf{x}, t) = \mathbf{x}^T\mathbf{x}/t$ 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 for $x, t > 0$
- if f is convex, then

$$g(\mathbf{x}) = \left(\mathbf{c}^T\mathbf{x} + d\right) f\left(\frac{\mathbf{A}\mathbf{x} + \mathbf{b}}{\mathbf{c}^T\mathbf{x} + d}\right) \tag{41}$$

is convex on $\left\{\mathbf{x} : \mathbf{c}^T\mathbf{x} + d > 0, (\mathbf{A}\mathbf{x} + \mathbf{b})/\left(\mathbf{c}^T\mathbf{x} + d\right) \in \mathrm{dom}\, f\right\}$

The conjugate function

the conjugate of a function $f(\mathbf{x})$ is

$$f^*(\mathbf{y}) = \max_{\mathbf{x} \in \text{dom } f} \left(\mathbf{y}^T \mathbf{x} - f(\mathbf{x}) \right) \tag{42}$$

- $f^*(\mathbf{y})$ is convex, *even if* $f(\mathbf{x})$ is not
- Q: why?
- is used in duality theory

The conjugate function: examples

- negative logarithm $f(x) = -\ln x$

$$f^*(y) = \max_{x>0}(xy + \ln x) \quad (43)$$

$$= \begin{cases} -1 - \ln(-y) & y < 0 \\ \infty & \text{otherwise} \end{cases} \quad (44)$$

The conjugate function: examples

- negative logarithm $f(x) = -\ln x$

$$f^*(y) = \max_{x>0}(xy + \ln x) \tag{43}$$

$$= \begin{cases} -1 - \ln(-y) & y < 0 \\ \infty & \text{otherwise} \end{cases} \tag{44}$$

- strictly convex quadratic $f(\mathbf{x}) = (1/2)\mathbf{x}^T\mathbf{Q}\mathbf{x}$, $\mathbf{Q} > 0$,

$$f^*(\mathbf{y}) = \max_{\mathbf{x}} \left(\mathbf{y}^T\mathbf{x} - (1/2)\mathbf{x}^T\mathbf{Q}\mathbf{x}\right) \tag{45}$$

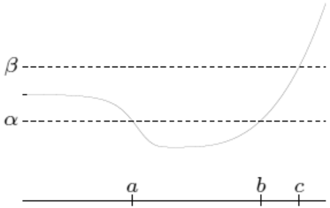
$$= \frac{1}{2}\mathbf{y}^T\mathbf{Q}^{-1}\mathbf{y} \tag{46}$$

Quasiconvex functions

$f(\mathbf{x})$ is quasiconvex if dom f is convex and the sublevel sets

$$S_\alpha = \{\mathbf{x} : f(\mathbf{x}) \le \alpha\} \quad (47)$$

are convex for all α

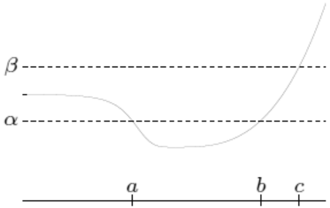


Quasiconvex functions

$f(\mathbf{x})$ is quasiconvex if dom f is convex and the sublevel sets

$$S_\alpha = \{\mathbf{x} : f(\mathbf{x}) \leq \alpha\} \tag{47}$$

are convex for all α



- f is quasiconcave if $-f$ is quasiconvex
- f is quasilinear if it is quasiconvex and quasiconcave

Quasiconvex functions: examples

- $\sqrt{|x|}$ is quasiconvex

Quasiconvex functions: examples

- $\sqrt{|x|}$ is quasiconvex
- $\text{ceil}(x) = \min\{z \in \mathbf{Z} : z \geq x\}$ is quasilinear

Quasiconvex functions: examples

- $\sqrt{|x|}$ is quasiconvex
- $\text{ceil}(x) = \min\{z \in \mathbf{Z} : z \geq x\}$ is quasilinear
- $\log x$ is quasilinear on $x > 0$

Quasiconvex functions: examples

- $\sqrt{|x|}$ is quasiconvex
- $\text{ceil}(x) = \min\{z \in \mathbf{Z} : z \geq x\}$ is quasilinear
- $\log x$ is quasilinear on $x > 0$
- $f(x_1, x_2) = x_1 x_2$ is quasiconcave on $x_1, x_2 > 0$

Quasiconvex functions: examples

- $\sqrt{|x|}$ is quasiconvex
- $\text{ceil}(x) = \min\{z \in \mathbf{Z} : z \geq x\}$ is quasilinear
- $\log x$ is quasilinear on $x > 0$
- $f(x_1, x_2) = x_1 x_2$ is quasiconcave on $x_1, x_2 > 0$
- linear-fractional function is quasilinear,

$$f(\mathbf{x}) = \frac{\mathbf{a}^T\mathbf{x} + b}{\mathbf{c}^T\mathbf{x} + d}, \quad \text{dom } f = \left\{\mathbf{x} : \mathbf{c}^T\mathbf{x} + d > 0\right\} \tag{48}$$

Quasiconvex functions: examples

- $\sqrt{|x|}$ is quasiconvex
- $\text{ceil}(x) = \min\{z \in \mathbf{Z} : z \geq x\}$ is quasilinear
- $\log x$ is quasilinear on $x > 0$
- $f(x_1, x_2) = x_1 x_2$ is quasiconcave on $x_1, x_2 > 0$
- linear-fractional function is quasilinear,

$$f(\mathbf{x}) = \frac{\mathbf{a}^T\mathbf{x} + b}{\mathbf{c}^T\mathbf{x} + d}, \quad \text{dom } f = \left\{\mathbf{x} : \mathbf{c}^T\mathbf{x} + d > 0\right\} \tag{48}$$

- distance ratio is quasiconvex,

$$f(\mathbf{x}) = \frac{|\mathbf{x} - \mathbf{a}|_2}{|\mathbf{x} - \mathbf{b}|_2}, \quad \text{dom } f = \{\mathbf{x} : |\mathbf{x} - \mathbf{a}|_2 \leq |\mathbf{x} - \mathbf{b}|_2\} \tag{49}$$

Quasiconvex functions: properties

- **modified Jensen inequality:** for quasiconvex $f(\mathbf{x})$

$$0 \leq \theta \leq 1 \quad \Longrightarrow \quad f(\theta \mathbf{x} + (1-\theta)\mathbf{y}) \leq \max\{f(\mathbf{x}), f(\mathbf{y})\} \qquad (50)$$

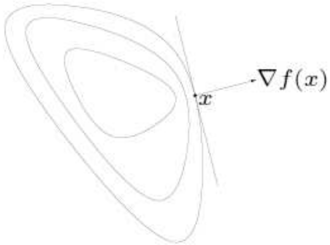
Quasiconvex functions: properties

- **modified Jensen inequality:** for quasiconvex $f(\mathbf{x})$

$$0 \leq \theta \leq 1 \quad \Longrightarrow \quad f(\theta \mathbf{x} + (1-\theta)\mathbf{y}) \leq \max\{f(\mathbf{x}), f(\mathbf{y})\} \qquad (50)$$

- **first-order condition:** differentiable $f(\mathbf{x})$ with convex domain is quasiconvex iff

$$f(\mathbf{y}) \leq f(\mathbf{x}) \quad \Longrightarrow \quad \nabla f(\mathbf{x})^T(\mathbf{y}-\mathbf{x}) \leq 0 \qquad (51)$$



sums of quasiconvex functions are not necessarily quasiconvex

Log-concave and log-convex functions

a positive function $f(\mathbf{x})$ is log-concave if $\log f(\mathbf{x})$ is concave:

$$f(\theta\mathbf{x} + (1-\theta)\mathbf{y}) \geq f(\mathbf{x})^{\theta} f(\mathbf{y})^{1-\theta} \quad \text{for } 0 \leq \theta \leq 1 \tag{52}$$

$f(\mathbf{x})$ is log-convex iff $\log f(\mathbf{x})$ is convex

Log-concave and log-convex functions

a positive function $f(\mathbf{x})$ is log-concave if $\log f(\mathbf{x})$ is concave:

$$f(\theta\mathbf{x} + (1-\theta)\mathbf{y}) \geq f(\mathbf{x})^{\theta} f(\mathbf{y})^{1-\theta} \quad \text{for } 0 \leq \theta \leq 1 \tag{52}$$

$f(\mathbf{x})$ is log-convex iff $\log f(\mathbf{x})$ is convex

examples

- powers: x^p on $x > 0$ is log-convex for $p \leq 0$, log-concave for $p \geq 0$

Log-concave and log-convex functions

a positive function $f(\mathbf{x})$ is log-concave if $\log f(\mathbf{x})$ is concave:

$$f(\theta \mathbf{x} + (1-\theta)\mathbf{y}) \geq f(\mathbf{x})^{\theta} f(\mathbf{y})^{1-\theta} \quad \text{for } 0 \leq \theta \leq 1 \tag{52}$$

$f(\mathbf{x})$ is log-convex iff $\log f(\mathbf{x})$ is convex

examples

- powers: x^p on $x > 0$ is log-convex for $p \leq 0$, log-concave for $p \geq 0$
- many common probability densities are log-concave, e.g. normal:

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{R}|}} e^{-\frac{1}{2}(\mathbf{x}-\bar{\mathbf{x}})^T \mathbf{R}^{-1}(\mathbf{x}-\bar{\mathbf{x}})} \tag{53}$$

- cumulative Gaussian distribution function Φ is log-concave

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} du \tag{54}$$

Properties of log-concave functions

- twice differentiable $f(\mathbf{x})$ with convex domain is log-concave iff

$$f(\mathbf{x})\nabla^2 f(\mathbf{x}) \preceq \nabla f(\mathbf{x})\nabla f(\mathbf{x})^T \ \forall\ \mathbf{x} \in \operatorname{dom} f \tag{55}$$

Properties of log-concave functions

- twice differentiable $f(\mathbf{x})$ with convex domain is log-concave iff

$$f(\mathbf{x})\nabla^2 f(\mathbf{x}) \preceq \nabla f(\mathbf{x})\nabla f(\mathbf{x})^T \ \forall \ \mathbf{x} \in \operatorname{dom} f \tag{55}$$

- product of log-concave functions is log-concave
- sum of log-concave functions is not always log-concave
- integration: if $f(\mathbf{x}, \mathbf{y})$ is log-concave, then

$$g(\mathbf{x}) = \int f(\mathbf{x}, \mathbf{y}) d\mathbf{y} \tag{56}$$

is log-concave (not easy to show)

Consequences of integration property

- convolution $f * g$ of log-concave functions f, g is log-concave

$$(f * g)(x) = \int f(x - y)g(y)dy \tag{57}$$

Consequences of integration property

- convolution $f * g$ of log-concave functions f, g is log-concave

$$(f * g)(x) = \int f(x - y)g(y)dy \tag{57}$$

- if set C convex and $\mathbf{y}$ is a random variable with log-concave pdf, then $f(\mathbf{x})$ is log-concave,

$$f(\mathbf{x}) = \Pr(\mathbf{x} + \mathbf{y} \in C) \tag{58}$$

Consequences of integration property

- convolution $f * g$ of log-concave functions f, g is log-concave

$$(f * g)(x) = \int f(x-y)g(y)dy \tag{57}$$

- if set C convex and $\mathbf{y}$ is a random variable with log-concave pdf, then $f(\mathbf{x})$ is log-concave,

$$f(\mathbf{x}) = \Pr(\mathbf{x}+\mathbf{y} \in C) \tag{58}$$

- proof: write $f(\mathbf{x})$ as integral of product of log-concave functions,

$$f(\mathbf{x}) = \int g(\mathbf{x}+\mathbf{y})p(\mathbf{y})d\mathbf{y}, \quad g(\mathbf{u}) = \begin{cases} 1 & \mathbf{u} \in C \\ 0 & \mathbf{u} \notin C, \end{cases} \tag{59}$$

  $p(\mathbf{y})$ is the pdf of $\mathbf{y}$

Example: yield function

$$Y(\mathbf{x}) = \Pr(\mathbf{x} + \mathbf{w} \in S) \tag{60}$$

- $\mathbf{x}$: nominal parameter values for product
- $\mathbf{w}$: random variations of parameters in manufactured product
- S : set of acceptable values

Example: yield function

$$Y(\mathbf{x}) = \Pr(\mathbf{x} + \mathbf{w} \in S) \tag{60}$$

- $\mathbf{x}$: nominal parameter values for product
- $\mathbf{w}$: random variations of parameters in manufactured product
- S : set of acceptable values

if S is convex and $\mathbf{w}$ has a log-concave pdf, then

- $Y(\mathbf{x})$ is log-concave
- yield regions $\{\mathbf{x} : Y(\mathbf{x}) \geq \alpha\}$ are convex

Convexity with respect to generalized inequalities

- matrix inequality (positive semi-definite):

$$\mathbf{A} \geq \mathbf{B} \Leftrightarrow \mathbf{A} - \mathbf{B} \geq 0 \Leftrightarrow \mathbf{z}^T(\mathbf{A} - \mathbf{B})\mathbf{z} \geq 0 \ \forall \ \mathbf{z} \tag{61}$$

Convexity with respect to generalized inequalities

- matrix inequality (positive semi-definite):

$$\mathbf{A} \geq \mathbf{B} \Leftrightarrow \mathbf{A} - \mathbf{B} \geq 0 \Leftrightarrow \mathbf{z}^T(\mathbf{A} - \mathbf{B})\mathbf{z} \geq 0 \ \forall \ \mathbf{z} \tag{61}$$

- $f(\mathbf{X})$ is matrix-convex if $\mathrm{dom}\, f$ is convex and

$$f(\theta\mathbf{X} + (1-\theta)\mathbf{Y}) \preceq \theta f(\mathbf{X}) + (1-\theta) f(\mathbf{Y}) \tag{62}$$

for $\mathbf{X}, \mathbf{Y} \in \mathrm{dom}\, f, 0 \leq \theta \leq 1$

Convexity with respect to generalized inequalities

- matrix inequality (positive semi-definite):

$$\mathbf{A} \geq \mathbf{B} \Leftrightarrow \mathbf{A} - \mathbf{B} \geq 0 \Leftrightarrow \mathbf{z}^T(\mathbf{A} - \mathbf{B})\mathbf{z} \geq 0 \ \forall \ \mathbf{z} \tag{61}$$

- $f(\mathbf{X})$ is matrix-convex if dom f is convex and

$$f(\theta\mathbf{X} + (1-\theta)\mathbf{Y}) \preceq \theta f(\mathbf{X}) + (1-\theta)f(\mathbf{Y}) \tag{62}$$

for $\mathbf{X}, \mathbf{Y} \in \text{dom}\, f, 0 \leq \theta \leq 1$

- **example:** $f(\mathbf{X}) = \mathbf{X}^2$ is matrix-convex on $\mathbf{X} = \mathbf{X}^T$

Convexity with respect to generalized inequalities

- matrix inequality (positive semi-definite):

$$\mathbf{A} \geq \mathbf{B} \Leftrightarrow \mathbf{A} - \mathbf{B} \geq 0 \Leftrightarrow \mathbf{z}^T(\mathbf{A} - \mathbf{B})\mathbf{z} \geq 0 \ \forall \ \mathbf{z} \tag{61}$$

- $f(\mathbf{X})$ is matrix-convex if $\text{dom}\, f$ is convex and

$$f(\theta\mathbf{X} + (1-\theta)\mathbf{Y}) \preceq \theta f(\mathbf{X}) + (1-\theta) f(\mathbf{Y}) \tag{62}$$

for $\mathbf{X}, \mathbf{Y} \in \text{dom}\, f, 0 \leq \theta \leq 1$

- **example:** $f(\mathbf{X}) = \mathbf{X}^2$ is matrix-convex on $\mathbf{X} = \mathbf{X}^T$

proof: for fixed $\mathbf{z}$, $g(\mathbf{X}) = \mathbf{z}^T\mathbf{X}^2\mathbf{z} = |\mathbf{X}\mathbf{z}|_2^2$ is convex in $\mathbf{X}$, i.e.

$$\mathbf{z}^T(\theta\mathbf{X} + (1-\theta)\mathbf{Y})^2\mathbf{z} \leq \theta\mathbf{z}^T\mathbf{X}^2\mathbf{z} + (1-\theta)\mathbf{z}^T\mathbf{Y}^2\mathbf{z} \tag{63}$$

therefore, $(\theta\mathbf{X} + (1-\theta)\mathbf{Y})^2 \preceq \theta\mathbf{X}^2 + (1-\theta)\mathbf{Y}^2$, as required