Lecture 2
Convex functions

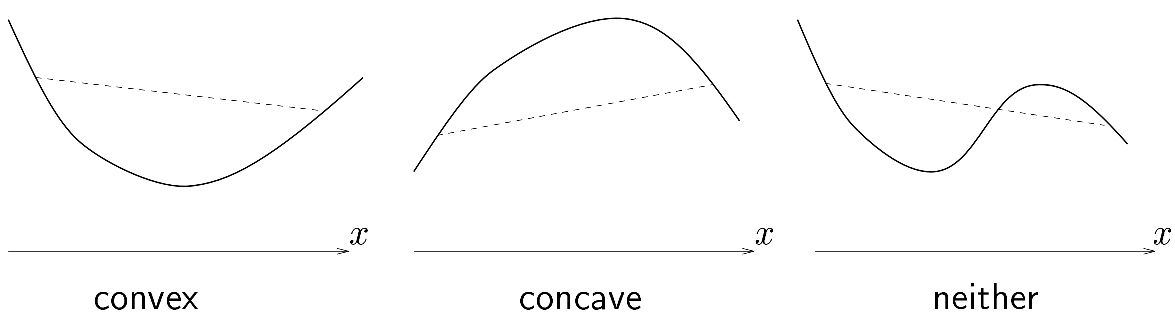
- convex functions, epigraph
- simple examples, elementary properties
- more examples, more properties
- Jensen's inequality
- quasiconvex, quasiconcave functions
- log-convex and log-concave functions
- K-convexity

Convex functions

$f : \mathbf{R}^n \to \mathbf{R}$ is *convex* if $\mathbf{dom}\, f$ is convex and

$$x, y \in \mathbf{dom}\, f, \quad \lambda \in [0,1]$$
$$\Downarrow$$
$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y) \qquad (1)$$

f is concave if $-f$ is convex



'Modern' definition: $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$

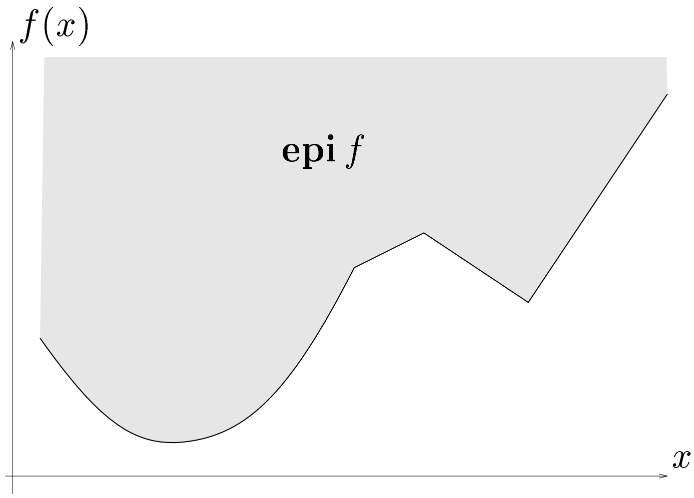
(but not identically $+\infty$)

f is convex if (1) holds as an inequality in $\mathbf{R} \cup \{+\infty\}$

Epigraph & sublevel sets

The *epigraph* of the function f is

$$\mathbf{epi}\, f = \{(x,t) \mid x \in \mathbf{dom}\, f,\ \ f(x) \le t\ \}\,.$$



f convex function $\Leftrightarrow$ $\mathbf{epi}\, f$ convex set

The (α-)*sublevel set* of f is

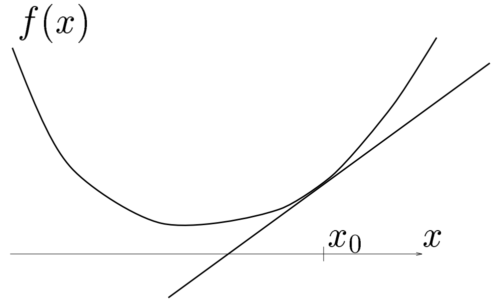
$$C(\alpha) \triangleq \{x \in \mathbf{dom}\, f \mid f(x) \le \alpha\}\,.$$

f convex $\Rightarrow$ sublevel sets are convex (converse false)

Differentiable convex functions

f differentiable and convex

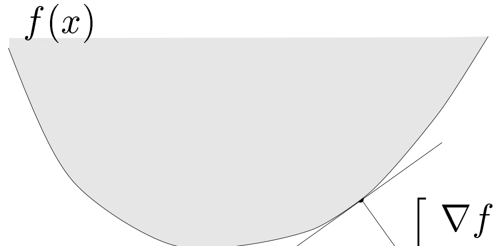
$$\iff \forall x, x_0 : f(x) \geq f(x_0) + \nabla f(x_0)^T(x - x_0)$$



Interpretation

- 1st order Taylor appr. is a *global lower bound* on f
- supporting hyperplane to $\mathbf{epi}\, f$:

$$(x, t) \in \mathbf{epi}\, f \implies \begin{bmatrix} \nabla f(x_0) \\ -1 \end{bmatrix}^T \begin{bmatrix} x - x_0 \\ t - f(x_0) \end{bmatrix} \leq 0$$



f twice differentiable and convex $\iff \nabla^2 f(x) \succeq 0$

Simple examples

- linear and affine functions: $f(x) = a^T x + b$
- convex quadratic functions:
  $f(x) = x^T P x + 2q^T x + r$ with $P = P^T \succeq 0$
- any norm

Examples on $\mathbf{R}$

- x^α is convex on $\mathbf{R}_+$ for $\alpha \geq 1$, $\alpha \leq 0$; concave for $0 \leq \alpha \leq 1$
- $\log x$ is concave, $x \log x$ is convex on $\mathbf{R}_+$
- $e^{\alpha x}$ is convex
- $|x|$, $\max(0, x)$, $\max(0, -x)$ are convex
- $\log \int_{-\infty}^{x} e^{-t^2} dt$ is concave

Elementary properties

- a function is convex iff it is convex on all lines:

$$f \text{ convex } \iff f(x_0 + th) \text{ convex in } t \text{ for all } x_0, h$$

- positive multiple of convex function is convex:

$$f \text{ convex}, \alpha \geq 0 \implies \alpha f \text{ convex}$$

- sum of convex functions is convex:

$$f_1, f_2 \text{ convex } \implies f_1 + f_2 \text{ convex}$$

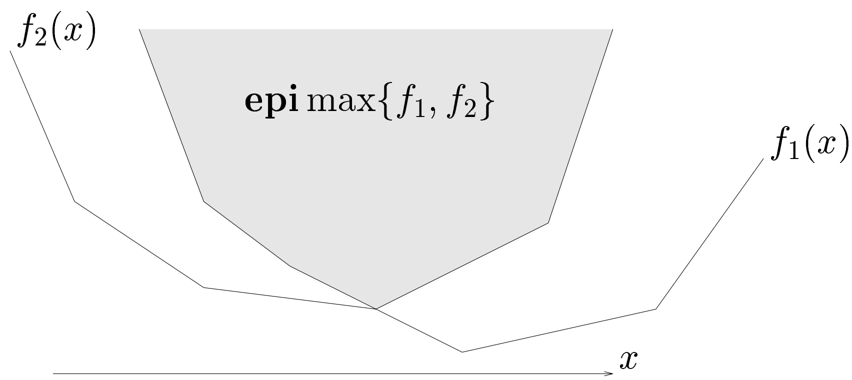
- extends to infinite sums, integrals:

$$g(x, y) \text{ convex in } x \implies \int g(x, y) dy \text{ convex}$$

- pointwise maximum:

$$f_1, f_2 \text{ convex } \implies \max\{f_1(x), f_2(x)\} \text{ convex}$$

(corresponds to intersection of epigraphs)



- pointwise supremum:

$$f_\alpha \text{ convex } \implies \sup_{\alpha \in \mathcal{A}} f_\alpha \text{ convex}$$

- affine transformation of domain

$$f \text{ convex } \Rightarrow f(Ax+b) \text{ convex}$$

More examples

- piecewise-linear functions: $f(x) = \max_i \{a_i^T x + b_i\}$ is convex in x ($\mathbf{epi}\, f$ is polyhedron)

- max distance to any set, $\sup_{s \in S} \|x - s\|$, is convex in x

- $f(x) = x_{[1]} + x_{[2]} + x_{[3]}$ is convex on $\mathbf{R}^n$
  ($x_{[i]}$ is the ith largest x_j)

- $f(x) = \left(\prod_i x_i \right)^{1/n}$ is concave on $\mathbf{R}^n_+$

- $f(x) = \sum_{i=1}^m \log(b_i - a_i^T x)^{-1}$ is convex on
  $\mathcal{P} = \{x \mid a_i^T x < b_i, i = 1, \ldots, m\}$

- least-squares cost as functions of weights,

  $$f(w) = \inf_x \sum_i w_i (a_i^T x - b_i)^2,$$

  is concave in w

Convex functions of matrices

- $\mathbf{Tr}\, X$ is linear in X; more generally,
  $\mathbf{Tr}\, A^T X = \sum_{i,j} A_{ij} X_{ij} = \mathbf{vec}(A)^T\, \mathbf{vec}(X)$

- $\log\det X^{-1}$ is convex on $X = X^T \succ 0$
  Proof: let λ_i be the eigenvalues of $X_0^{-1/2} H X_0^{-1/2}$

$$\begin{aligned} f(t) &\triangleq \log\det(X_0 + tH)^{-1} \\ &= \log\det X_0^{-1} + \log\det(I + tX_0^{-1/2} H X_0^{-1/2})^{-1} \\ &= \log\det X_0^{-1} - \sum_i \log(1 + t\lambda_i) \end{aligned}$$

  is a convex function of t

- $(\det X)^{1/n}$ is concave on $X = X^T \succ 0$, $X \in \mathbf{R}^{n \times n}$

- $\lambda_{\mathsf{max}}(X)$ is convex on $X = X^T$

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

- $\|X\| = \left(\lambda_{\mathsf{max}}(X^T X)\right)^{1/2}$ is convex on $\mathbf{R}^{n \times m}$

  Proof: $\|X\| = \sup_{\|u\|=1,\, \|v\|=1} u^T X v$

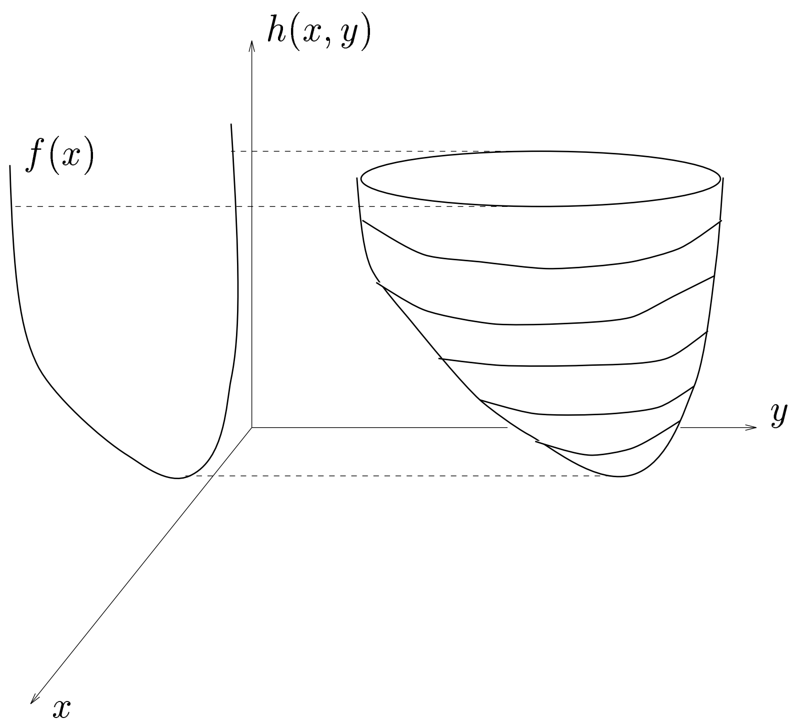
Minimizing over some variables

If $h(x, y)$ is convex in x and y, then

$$f(x) = \inf_y h(x, y)$$

is convex in x

corresponds to projection of epigraph, $(x, y, t) \to (x, t)$



Example. If $S \subseteq \mathbf{R}^n$ is convex then (min) distance to S,

$$\mathbf{dist}(x, S) = \inf_{s \in S} \|x - s\|$$

is convex in x

Example. If $g(x)$ is convex, then

$$f(y) = \inf\{g(x) \mid Ax = y\}$$

is convex in y.

Proof: find B, C s.t.

$$\{x \mid Ax = y\} = \{By + Cz \mid z \in \mathbf{R}^k\}$$

so $f(y) = \inf_z g(By + Cz)$

'Modern' proof: $f(y) = \inf_z g(x) + h(Ax - y)$ where

$$h(z) = \begin{cases} 0 & \text{if } z = 0 \\ +\infty & \text{otherwise} \end{cases}$$

is convex

Composition — one-dimensional case

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

is convex if

- g convex; h convex, nondecreasing
- g concave; h convex, nonincreasing

Examples

- $f(x) = \exp g(x)$ is convex if g is convex
- $f(x) = 1/g(x)$ is convex if g is concave, positive
- $f(x) = g(x)^p$, $p \geq 1$, is convex if $g(x)$ convex, positive
- $f_1, \ldots, f_n$ convex, then $f(x) = -\sum_i \log(-f_i(x))$ is convex on $\{x \mid f_i(x) < 0,\ i = 1, \ldots, n\}$

Proof: (differentiable functions, $x \in \mathbf{R}$)

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

Composition — k-dimensional case

$$f(x) = h(g_1(x), \ldots, g_k(x))$$

with $h : \mathbf{R}^k \to \mathbf{R}$, $g_i : \mathbf{R}^n \to \mathbf{R}$ is convex if

- h convex, nondecreasing in each arg.; g_i convex
- h convex, nonincreasing in each arg.; g_i concave
- etc.

Examples

- $f(x) = \max_i g_i(x)$ is convex if each g_i is
- $f(x) = \log \sum_i \exp g_i(x)$ is convex if each g_i is

Proof: (differentiable functions, $n = 1$)

$$f'' = \nabla h^T \begin{bmatrix} g_1'' \\ \vdots \\ g_k'' \end{bmatrix} + \begin{bmatrix} g_1' \\ \vdots \\ g_k' \end{bmatrix}^T \nabla^2 h \begin{bmatrix} g_1' \\ \vdots \\ g_k' \end{bmatrix}$$

Jensen's inequality

$$f : \mathbf{R}^n \to \mathbf{R} \text{ convex}$$

- two points

$$\lambda \in [0,1]$$
$$\Downarrow$$
$$f(\lambda x_1 + (1-\lambda)x_2) \le \lambda f(x_1) + (1-\lambda) f(x_2)$$

- more than two points

$$\lambda_i \ge 0, \quad \sum_i \lambda_i = 1$$
$$\Downarrow$$
$$f(\sum_i \lambda_i x_i) \le \sum_i \lambda_i f(x_i)$$

- continuous version

$$p(x) \ge 0, \quad \int p(x)dx = 1$$
$$\Downarrow$$
$$f(\int x p(x) dx) \le \int f(x) p(x) dx$$

- most general form:

$$f(\mathbf{E}\, x) \le \mathbf{E}\, f(x)$$

Interpretation: (zero mean) randomization, dithering increases average value of a convex function

Applications

Many (some people claim most) inequalities can be derived from Jensen's inequality

Example. Arithmetic-geometric mean inequality

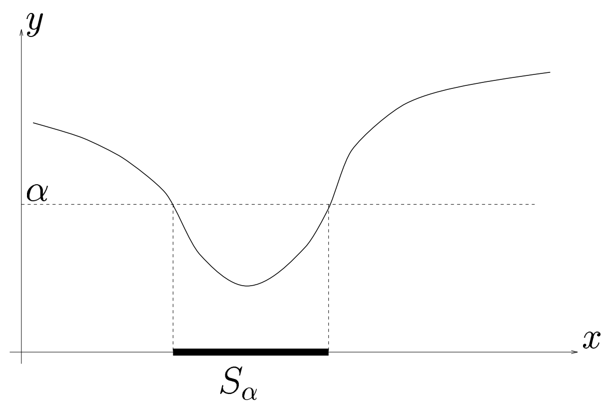
$$a, b \geq 0 \Rightarrow \sqrt{ab} \leq (a+b)/2$$

Proof. $f(x) = \log x$ is concave on $\mathbf{R}_+$:

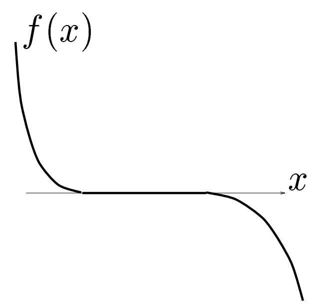
$$\frac{1}{2}(\log a + \log b) \leq \log\left(\frac{a+b}{2}\right)$$

Quasiconvex functions

$f : C \to \mathbf{R}$, C a convex set, is *quasiconvex* if every sublevel set $S_\alpha = \{x \mid f(x) \leq \alpha\}$ is convex.



can have 'locally flat' regions



f is *quasiconcave* if $-f$ is quasiconvex, $i.e.$, superlevel sets $\{x \mid f(x) \geq \alpha\}$ are convex.

A function which is both quasiconvex and quasiconcave is called *quasilinear*.

f convex (concave) $\Rightarrow$ f quasiconvex (quasiconcave)

Examples

- $f(x) = \sqrt{|x|}$ is quasiconvex on $\mathbf{R}$

- $f(x) = \log x$ is quasilinear on $\mathbf{R}_+$

- linear fractional function,
$$f(x) = \frac{a^T x + b}{c^T x + d}$$
is quasilinear on the halfspace $c^T x + d > 0$

- $f(x) = \dfrac{\|x - a\|}{\|x - b\|}$ is quasiconvex on the halfspace $\{x \mid \|x - a\| \le \|x - b\|\}$

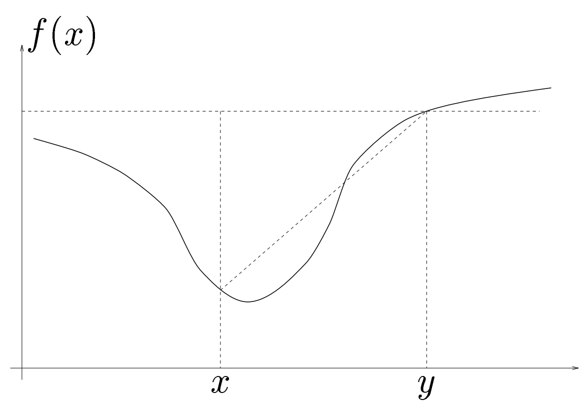
- $f(a) = \text{degree}(a_0 + a_1 t + \cdots + a_k t^k)$ on $\mathbf{R}^{k+1}$

Properties

- f is quasiconvex if and only if it is quasiconvex on lines, *i.e.*, $f(x_0 + th)$ quasiconvex in t for all x_0, h.

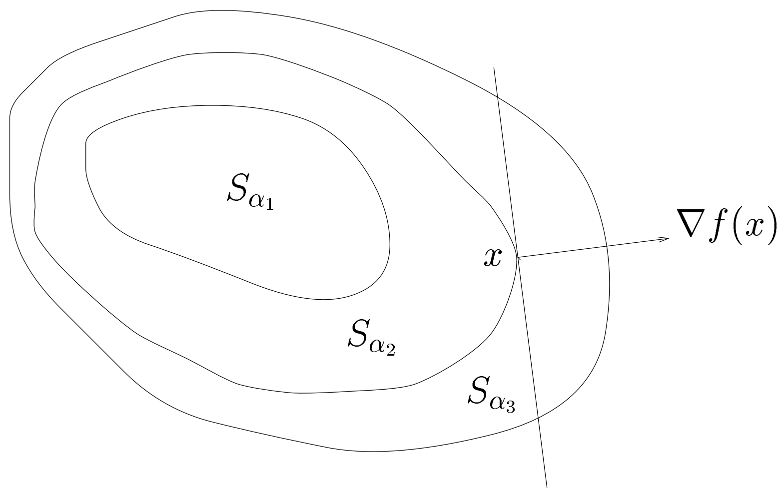
- modified Jensen's inequality: $f : C \to \mathbf{R}$ quasiconvex if and only if

$$x, y \in C, \lambda \in [0, 1]$$
$$\Downarrow$$
$$f(\lambda x + (1 - \lambda)y) \leq \max\{f(x), f(y)\}$$



- for f differentiable, f quasiconvex if and only if for all x, y

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



$$\alpha_1 < \alpha_2 < \alpha_3$$

- positive multiples

$$f \text{ quasiconvex}, \alpha \geq 0 \Longrightarrow \alpha f \text{ quasiconvex}$$

- pointwise maximum

$$f_1, f_2 \text{ quasiconvex } \Longrightarrow \max\{f_1, f_2\} \text{ quasiconvex}$$

(extends to supremum over arbitrary set)

- affine transformation of domain

$$f \text{ quasiconvex} \Longrightarrow f(Ax+b) \text{ quasiconvex}$$

- projective transformation of domain

$$f \text{ quasiconvex} \Longrightarrow f\left(\frac{Ax+b}{c^Tx+d}\right) \text{ quasiconvex}$$

on $c^Tx+d>0$

- composition with monotone increasing function

$$f \text{ quasiconvex}, g \text{ monotone increasing}$$
$$\Longrightarrow g(f(x)) \text{ quasiconvex}$$

- sums of quasiconvex functions are **not** quasiconvex in general

- f quasiconvex in x, $y \Longrightarrow g(x) = \inf_y f(x,y)$ quasiconvex in x

Nested sets characterization

f quasiconvex $\Rightarrow$ sublevel sets S_α convex, nested, *i.e.*,

$$\alpha_1 \leq \alpha_2 \Rightarrow S_{\alpha_1} \subseteq S_{\alpha_2}$$

converse: if T_α is a nested family of convex sets, then

$$f(x) = \inf\{\alpha \mid x \in T_\alpha\}$$

is quasiconvex.

Engineering interpretation: T_α are specs, tighter for smaller α

Log-concave functions

$f : \mathbf{R}^n \to \mathbf{R}_+$ is log-concave (log-convex) if $\log f$ is concave (convex)

Log-convex $\Rightarrow$ convex; concave $\Rightarrow$ log-concave

'Modern' definition allows log-concave f to take on value zero, so $\log f$ takes on value $-\infty$

Examples

- normal density, $f(x) = e^{-(1/2)(x-x_0)^T\Sigma^{-1}(x-x_0)}$
- erfc, $f(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty e^{-t^2}dt$
- indicator function of convex set C:

$$I_C(x) = \begin{cases} 1 & x \in C \\ 0 & x \notin C \end{cases}$$

Properties

- sum of log-concave functions not always log-concave (but sum of log-convex functions **is** log-convex)

- products

$$f, g \text{ log-concave } \Longrightarrow fg \text{ log-concave}$$

(immediate)

- integrals

$f(x, y)$ log-concave in $x, y \Longrightarrow \int f(x, y)dy$ log-concave

- convolutions

$$f, g \text{ log-concave} \Longrightarrow \int f(x - y)g(y)dy \text{ log-concave}$$

(immediate from the properties above)

Log-concave probability densities

Many common probability density functions are log-concave.

Examples

- normal ($\Sigma \succ 0$)

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

- exponential($\lambda_i > 0$)

$$f(x) = \left(\prod_{i=1}^{n} \lambda_i \right) e^{-(\lambda_1 x_1 + \cdots + \lambda_n x_n)}$$

  on $\mathbf{R}^n_+$

- uniform distribution on convex (bounded) set C

$$f(x) = \begin{cases} 1/\alpha & x \in C \\ 0 & x \notin C \end{cases}$$

  where α is Lebesgue measure of C ($i.e.$, length, area, volume ...)

K-convexity

convex cone $K \subseteq \mathbf{R}^m$ induces generalized inequality $\preceq_K$

$f : \mathbf{R}^n \to \mathbf{R}^m$ is K-convex if $0 \le \lambda \le 1 \quad \Longrightarrow$

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

Example. K is PSD cone (called *matrix convexity*)
let's show that $f(X) = X^2$ is K-convex on
$\{X | X = X^T\}$, $i.e.$, for $\lambda \in [0,1]$,

$$(\lambda X + (1-\lambda)Y)^2 \preceq \lambda X^2 + (1-\lambda)Y^2 \qquad (1)$$

for any $u \in \mathbf{R}^m$, $u^T X^2 u = \|Xu\|^2$ is a (quadratic) convex fct of X, so

$$u^T(\lambda X + (1-\lambda)Y)^2 u \le \lambda u^T X^2 u + (1-\lambda) u^T Y^2 u$$

which implies (1)