Lecture 4
Convex functions

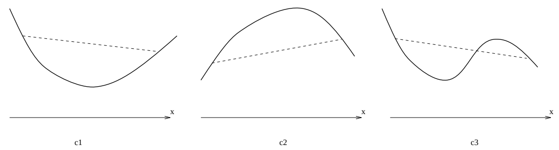
- convex functions, epigraph
- examples, properties
- Jensen's inequality
- conjugate functions
- quasiconvex, quasiconcave functions
- log-convex and log-concave functions
- K-convexity

Convex functions

$f : \mathbf{R}^n \to \mathbf{R}$ is **convex** if **dom** f is convex and for all $x, y \in$ **dom** f, $\theta \in [0, 1]$

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

f is concave if $-f$ is convex



examples (on **R**)

- $f(x) = x^2$ is convex
- $f(x) = \log x$ is concave (**dom** $f = \{x | x > 0\}$)
- $f(x) = 1/x$ is convex (**dom** $f = \{x | x > 0\}$)

Extended-valued extensions

for f convex, it's convenient to define the *extension*

$$\tilde{f}(x) = \begin{cases} f(x) & x \in \textbf{dom}\, f \\ +\infty & x \notin \textbf{dom}\, f \end{cases}$$

inequality

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

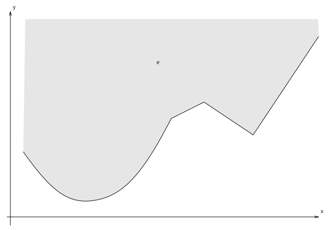
holds for all $x, y \in \mathbf{R}^n$, $0 \le \theta \le 1$
(as an inequality in $\mathbf{R} \cup \{+\infty\}$)

we'll use same symbol for f and its extension, *i.e.*, we'll implicitly assume convex functions are extended

Epigraph & sublevel sets

epigraph of a function f is

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



f convex function $\Leftrightarrow$ **epi** f convex set

the (α-)**sublevel set** of f is

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

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

Differentiable convex functions

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

$$\nabla f(x) = \left[\frac{\partial f}{\partial x_1} \; \frac{\partial f}{\partial x_2} \; \cdots \; \frac{\partial f}{\partial x_n} \right]^T \quad \text{(evaluated at } x\text{)}$$

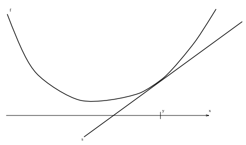
first order Taylor approximation at x_0:

$$f(x) \simeq f(x_0) + \nabla f(x_0)^T (x - x_0)$$

first-order condition: for f differentiable,
f is convex $\Longleftrightarrow$ for all $x, x_0 \in \textbf{dom}\, f$,

$$f(x) \geq f(x_0) + \nabla f(x_0)^T (x - x_0)$$

i.e., 1st order approx. is a *global underestimator*

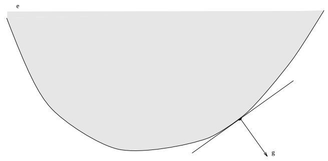


epigraph interpretation

for all $(x, t) \in \textbf{epi}\, f$,

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

i.e., $(\nabla f(x_0), -1)$ defines supporting hyperplane to **epi** f at $(x_0, f(x_0))$



Hessian of a twice differentiable function:

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

(evaluated at x)

2nd order Taylor series expansion around x_0:

$$f(x) \simeq f(x_0) + \nabla f(x_0)^T (x - x_0) + \frac{1}{2} (x - x_0)^T \nabla^2 f(x_0)(x - x_0)$$

second order condition: for f twice differentiable,
f is convex $\Longleftrightarrow$ for all $x \in \textbf{dom}\, f$, $\nabla^2 f(x) \succeq 0$

Simple examples

- linear and affine functions are convex and concave
- quadratic function $f(x) = x^T P x + 2q^T x + r$
  convex $\Longleftrightarrow P \succeq 0$; concave $\Longleftrightarrow P \preceq 0$
  $(P = P^T)$
- any norm is convex

examples on 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:

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

- sum of convex functions is convex:

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

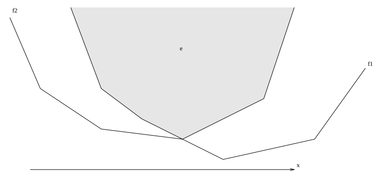
- extends to infinite sums, integrals:

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

- pointwise maximum:

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

  (corresponds to intersection of epigraphs)



- pointwise supremum:

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

- affine transformation of domain

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

More examples

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

- max distance to any set, $\mathbf{sup}_{s \in S} \|x - s\|$, 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)

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

- $f(x) = \Sigma_{i=1}^m \log(b_i - a_i^T x)^{-1}$ is convex (**dom** $f = \{x \mid a_i^T x < b_i, i = 1, \ldots, m\}$)

- 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

- **Tr** $A^T X = \Sigma_{i,j} A_{ij} X_{ij}$ is linear in X on $\mathbf{R}^{n \times n}$
- $\log\det X^{-1}$ is convex on $X = X^T \succ 0$, $X \in \mathbf{R}^{n \times n}$
  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 $X = X^T \succeq 0$, $X \in \mathbf{R}^{n \times n}$
- $\lambda_{\max}(X)$ is convex on $X = X^T$

  proof: $\lambda_{\max}(X) = \mathbf{sup}_{\|y\|=1}\, y^T X y$
- $\|X\| = (\lambda_{\max}(X^T X))^{1/2}$ is convex on $\mathbf{R}^{m \times n}$

  proof: $\|X\| = \mathbf{sup}_{\|u\|=1} \|Xu\|$

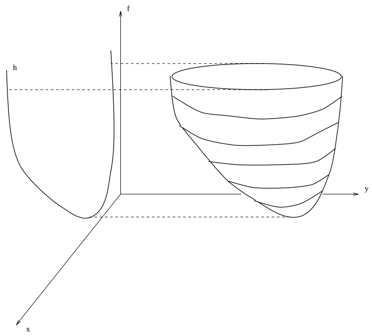
Minimizing over some variables

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, $(x, y, t) \to (x, t)$



examples

- 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
- if g is convex, then

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

  is convex in y

  proof: (assume $A \in \mathbf{R}^{m \times n}$ has rank m)
  find B s.t. $\mathbf{Range}(B) = \mathbf{Nullspace}(A)$; then
  $Ax = y$ iff

  $$x = A^T(AA^T)^{-1}y + Bz$$

  for some z, and hence

  $$f(y) = \inf_z g(A^T(AA^T)^{-1}y + Bz)$$

Composition — one-dimensional case

$f(x) = h(g(x))$ ($g : \mathbf{R}^n \to \mathbf{R}$, $h : \mathbf{R} \to \mathbf{R}$) is convex if

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

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

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

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
- $f(x) = -\Sigma_i \log(-f_i(x))$ is convex on $\{x \mid f_i(x) < 0\}$ if f_i are convex

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.

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

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

examples

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

Jensen's inequality

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

- two points: $\theta_1 + \theta_2 = 1$, $\theta_i \geq 0 \implies$
$$f(\theta_1 x_1 + \theta_2 x_2) \leq \theta_1 f(x_1) + \theta_2 f(x_2)$$
- more than two points: $\sum_i \theta_i = 1$, $\theta_i \geq 0 \implies$
$$f(\sum_i \theta_i x_i) \leq \sum_i \theta_i f(x_i)$$
- continuous version: $p(x) \geq 0$, $\int p(x)\ dx = 1 \implies$
$$f(\int xp(x)\ dx) \leq \int f(x)p(x)\ dx$$
- most general form: for any prob. distr. on x,
$$f(\mathbf{E}x) \leq \mathbf{E}f(x)$$

these are all called *Jensen's inequality*

interpretation of Jensen's inequality:
(zero mean) randomization, dithering increases average value of a convex function

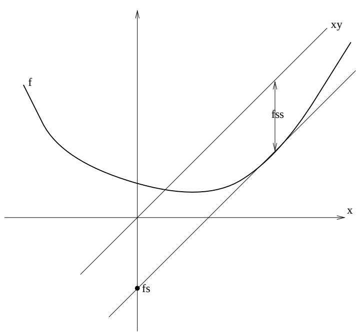
many (some people claim most) inequalities can be derived from Jensen's inequality

example: arithmetic-geometric mean inequality
$$a, b \geq 0 \Rightarrow \sqrt{ab} \leq (a+b)/2$$

proof: $f(x) = \log x$ is concave on $\{x | x > 0\}$, so for $a, b > 0$,
$$\frac{1}{2}(\log a + \log b) \leq \log\left(\frac{a+b}{2}\right)$$

Conjugate functions

the **conjugate** function of $f : \mathbf{R}^n \to \mathbf{R}$ is
$$f^\star(y) = \sup_{x \in \mathbf{dom}\, f} \left(y^T x - f(x)\right)$$



- $f^\star$ is convex (even if f isn't)
- will be useful later

Examples

$f(x) = -\log x$ ($\mathbf{dom}\, f = \{x \mid x > 0\}$):
$$\begin{aligned} f^\star(y) &= \sup_{x>0}(xy + \log x) \\ &= \begin{cases} -1 - \log(-y) & \text{if } y < 0 \\ +\infty & \text{otherwise} \end{cases} \end{aligned}$$

$f(x) = x^T P x$ ($P \succ 0$):
$$f^\star(y) = \sup_x (y^T x - x^T P x) = \frac{1}{4} y^T P^{-1} y$$