3. Convex functions

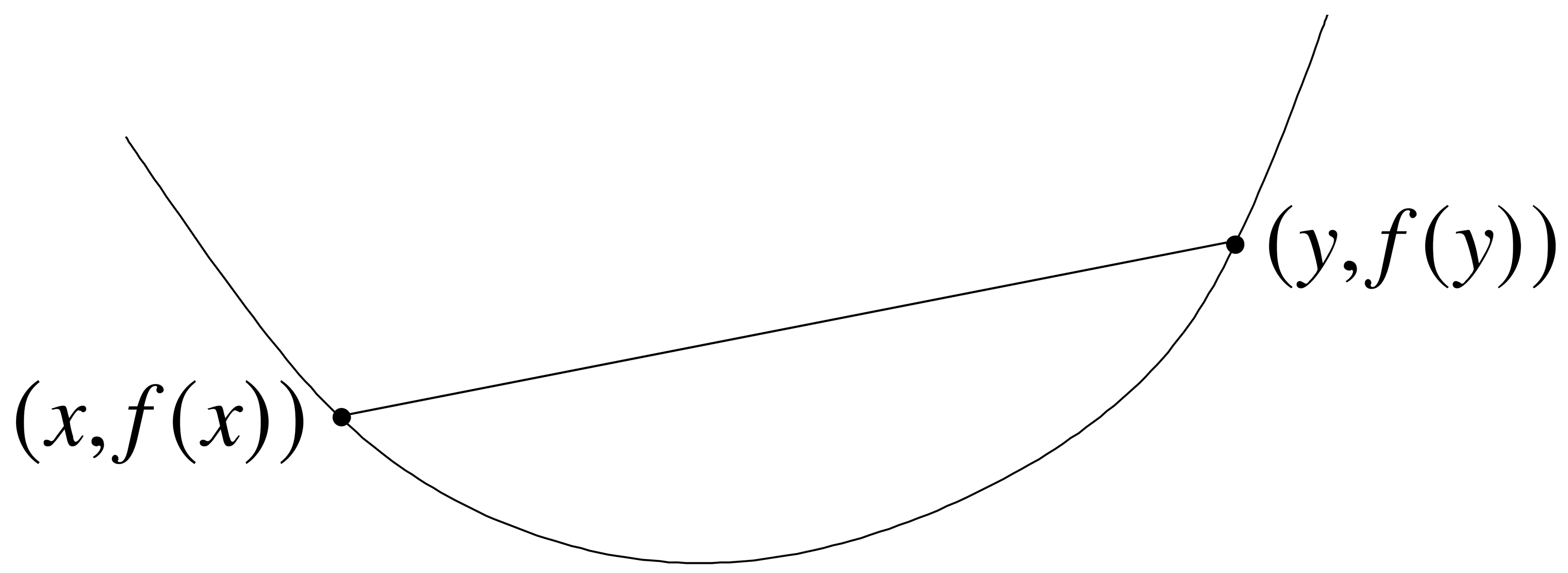
Outline

Definition

Important!

- $f : \mathbf{R}^n \to \mathbf{R}$ is convex if $\mathbf{dom}\, f$ is a convex set and for all $x, y \in \mathbf{dom}\, f$, $0 \le \theta \le 1$,

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



- f is concave if $-f$ is convex
- f is strictly convex if $\mathbf{dom}\, f$ is convex and for $x, y \in \mathbf{dom}\, f$, $x \ne y$, $0 < \theta < 1$,

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

Examples on $\mathbf{R}$

convex functions:

- affine: $ax + b$ on $\mathbf{R}$, for any $a, b \in \mathbf{R}$
- 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 $p \geq 1$
- positive part (relu): $\max\{0, x\}$

concave functions:

- affine: $ax + b$ on $\mathbf{R}$, for any $a, b \in \mathbf{R}$
- powers: x^α on $\mathbf{R}_{++}$, for $0 \leq \alpha \leq 1$
- logarithm: $\log x$ on $\mathbf{R}_{++}$
- entropy: $-x \log x$ on $\mathbf{R}_{++}$
- negative part: $\min\{0, x\}$

Examples on $\mathbf{R}^n$

convex functions:

- affine functions: $f(x) = a^T x + b$
- any norm, *e.g.*, the ℓ_p norms
  - $\|x\|_p = (|x_1|^p + \cdots + |x_n|^p)^{1/p}$ for $p \geq 1$
  - $\|x\|_\infty = \max\{|x_1|, \ldots, |x_n|\}$
- sum of squares: $\|x\|_2^2 = x_1^2 + \cdots + x_n^2$
- max function: $\max(x) = \max\{x_1, x_2, \ldots, x_n\}$
- softmax or log-sum-exp function: $\log(\exp x_1 + \cdots + \exp x_n)$

Triangle ineq: $\|x+y\| \leq \|x\| + \|y\|$.

Scalar mult: $\|\alpha x\| = \alpha \|x\|$.

$\max(x) \leq LSE(x) \leq \max(x) + \log(n)$

$\left(e^{\max(x)} \leq \sum_i e^{x_i} \leq n\, e^{\max(x)}\right)$

$\frac{\partial LSE}{\partial x_i} = \frac{e^{x_i}}{e^{x_1} + \cdots + e^{x_n}}$.

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

- $X \in \mathbf{R}^{m\times n}$ ($m \times n$ matrices) is the variable
- general affine function has form

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

treating A, X like vectors.

  for some $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}$

- spectral norm (maximum singular value) is convex

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

$\sigma_{max}(X) = \max_{\|x\|_2 \le 1, \|y\|_2 \le 1} x^T X y$

- log-determinant: for $X \in \mathbf{S}^n_{++}$, $f(X) = \log\det X$ is concave

Extended-value extension

- suppose f is convex on $\mathbf{R}^n$, with domain $\mathbf{dom}\, f$
- its extended-value extension $\tilde{f}$ is function $\tilde{f} : \mathbf{R}^n \to \mathbf{R} \cup \{\infty\}$

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

- often simplifies notation; for example, the condition

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

  (as an inequality in $\mathbf{R} \cup \{\infty\}$), means the same as the two conditions
  - $\mathbf{dom}\, f$ is convex
  - $x, y \in \mathbf{dom}\, f$, $0 \leq \theta \leq 1 \quad \Longrightarrow \quad f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$

Restriction of a convex function to a line

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

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

  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

Example

- $f : \mathbf{S}^n \to \mathbf{R}$ with $f(X) = \log\det X$, $\mathbf{dom}\, f = \mathbf{S}^n_{++}$
- consider line in $\mathbf{S}^n$ given by $X + tV$, $X \in \mathbf{S}^n_{++}$, $V \in \mathbf{S}^n$, $t \in \mathbf{R}$

direction. *} s.t. $X + tV \in S^n_{++}$*

$$\begin{aligned} g(t) &= \log\det(X + tV) \\ &= \log\det\left(X^{1/2}\left(I + tX^{-1/2}VX^{-1/2}\right)X^{1/2}\right) \\ &= \log\det X + \log\det\left(I + tX^{-1/2}VX^{-1/2}\right) \\ &= \log\det X + \sum_{i=1}^{n} \log(1 + t\lambda_i) \end{aligned}$$

Recall def of char. poly.

$\chi_A(s) = \det(sI - A) = \prod_{i=1}^{n}(s - \lambda_i)$

$\Leftrightarrow \det(I + A) = \prod_{i=1}^{n}(1 + \lambda_i)$

where λ_i are the eigenvalues of $X^{-1/2}VX^{-1/2}$

- g is concave in t (for any choice of $X \in \mathbf{S}^n_{++}$, $V \in \mathbf{S}^n$); hence f is concave

First-order condition

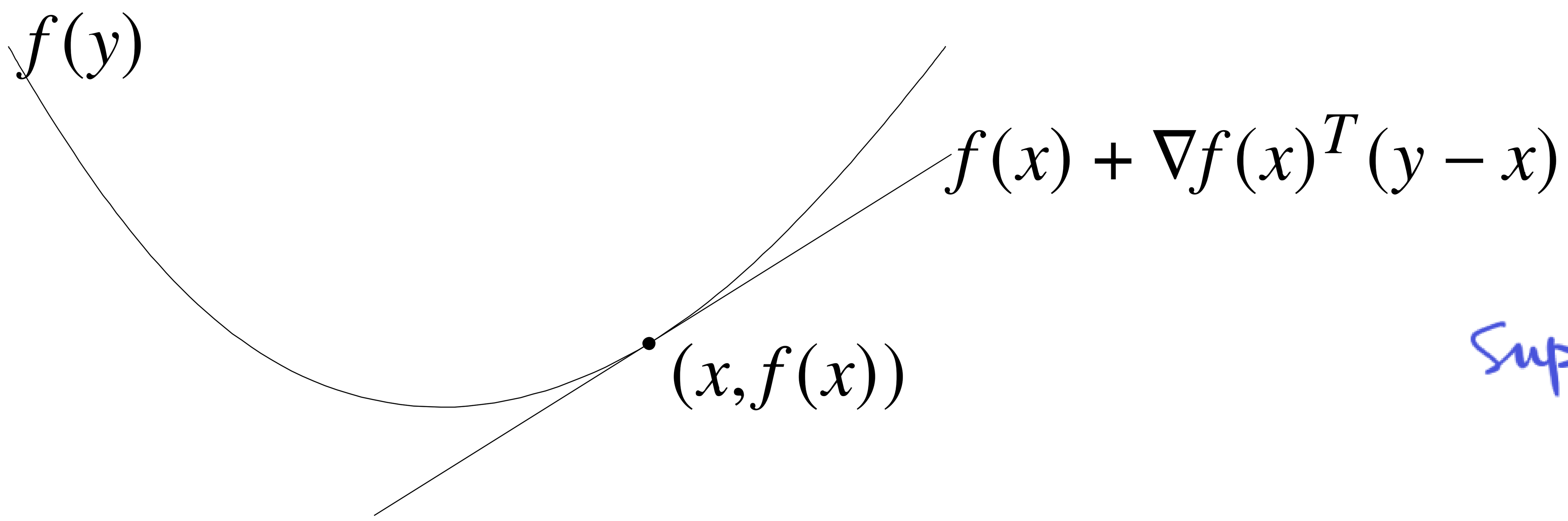
- f is **differentiable** if $\mathbf{dom}\, f$ is open and the gradient

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \ldots, \frac{\partial f(x)}{\partial x_n} \right) \in \mathbf{R}^n$$

  exists at each $x \in \mathbf{dom}\, f$
- **1st-order condition:** differentiable f with convex domain is convex if and only if

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

- first order Taylor approximation of convex f is a **global underestimator** of f



Supporting hyperplane!

Second-order conditions

- f is **twice differentiable** if $\mathbf{dom}\, f$ is open and the Hessian $\nabla^2 f(x) \in \mathbf{S}^n$,

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

  exists at each $x \in \mathbf{dom}\, f$

- **2nd-order conditions:** for twice differentiable f with convex domain
  - f is convex if and only if $\nabla^2 f(x) \succeq 0$ for all $x \in \mathbf{dom}\, f$
  - if $\nabla^2 f(x) \succ 0$ for all $x \in \mathbf{dom}\, f$, then f is strictly convex

Examples

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

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

  convex if $P \succeq 0$ (concave if $P \preceq 0$)

- **least-squares objective**: $f(x) = \|Ax - b\|_2^2$

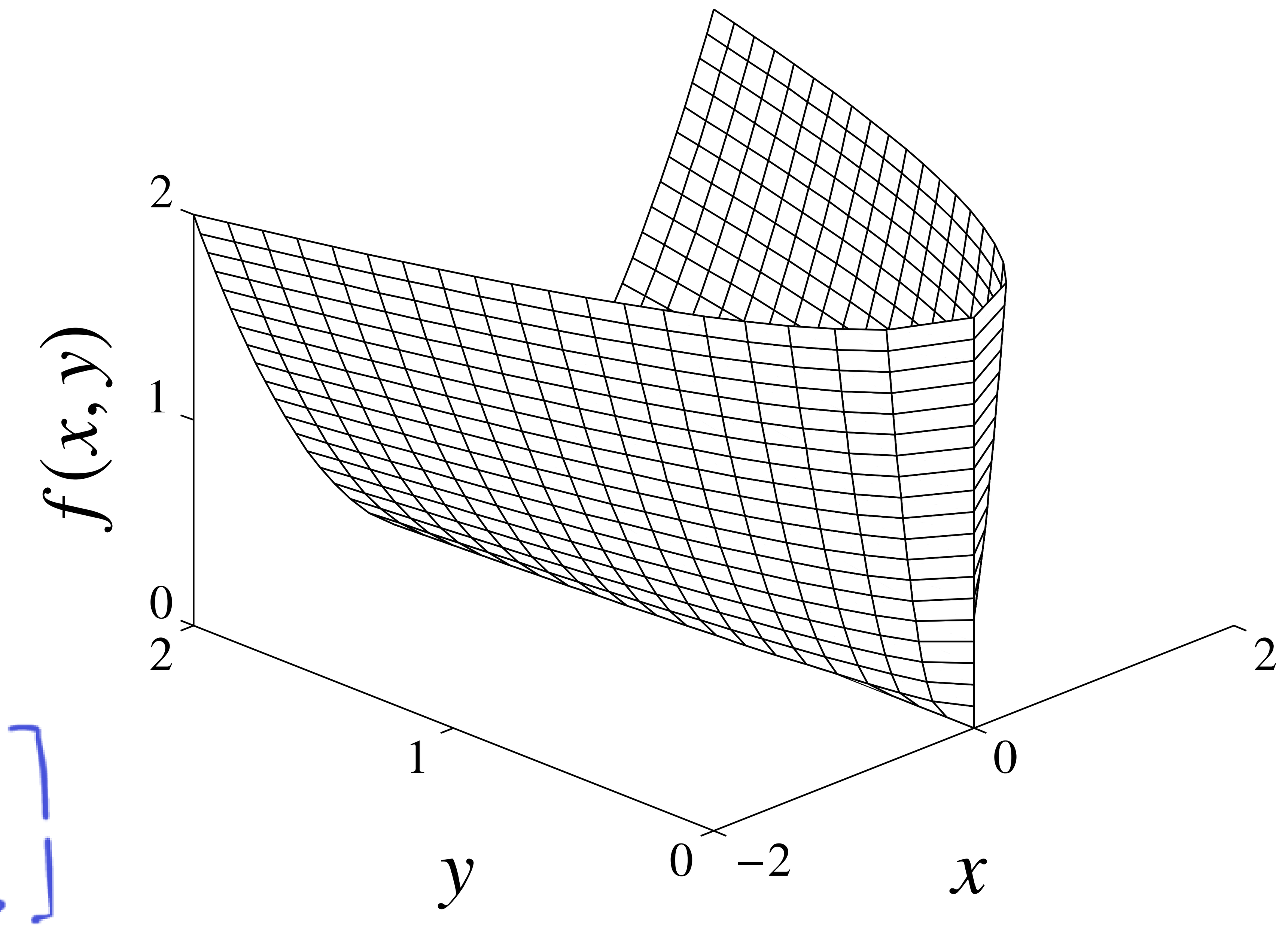
$$\nabla f(x) = 2A^T(Ax - b), \qquad \nabla^2 f(x) = 2A^T A$$

  convex (for any A)

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

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

  convex for $y > 0$



$$\nabla f = \begin{bmatrix} 2x/y \\ -x^2/y^2 \end{bmatrix}. \quad \nabla^2 f = \begin{bmatrix} \frac{2}{y} & -\frac{2x}{y^2} \\ -\frac{2x}{y^2} & -2\frac{x^2}{y^3} \end{bmatrix}$$

More examples

$\frac{\partial f}{\partial x_k} = \frac{\partial f}{\partial z_k} \frac{\partial z_k}{\partial x_k} = z_k \frac{\partial f}{\partial z_k} = \frac{z_k}{\sum_k z_k}$

$\frac{\partial^2 f}{\partial x_{k'} \partial x_k} = \frac{\partial}{\partial x_{k'}} \left(\frac{\partial f}{\partial x_k} \right) = z_{k'} \frac{\partial}{\partial z_{k'}} \left(\frac{z_k}{\sum z_k} \right) = z_{k'} \left(\delta_{kk'} \frac{1}{\sum z_k} - \frac{z_k}{(\sum z_k)^2} \right).$

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

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

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

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

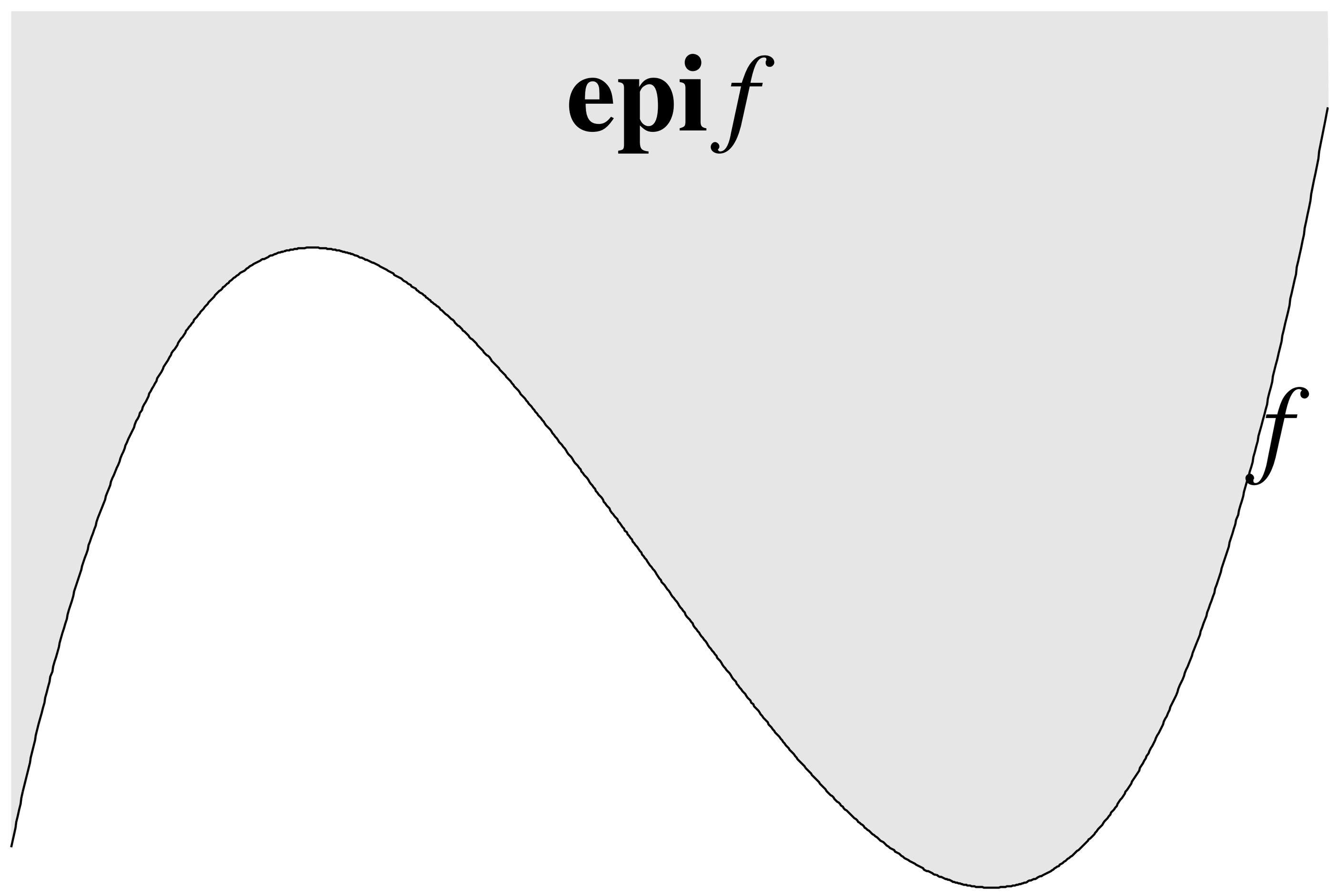
since $(\sum_k v_k z_k)^2 \leq (\sum_k z_k v_k^2)(\sum_k z_k)$ (from Cauchy-Schwarz inequality)

$z_k > 0.$ $\lambda_k = \frac{z_k}{\sum z_k}.$ $(\sum_k \lambda_k v_k)^2 \leq \sum \lambda_k v_k^2$

- **geometric mean**: $f(x) = (\prod_{k=1}^n x_k)^{1/n}$ on $\mathbf{R}_{++}^n$ is concave (similar proof as above)

Epigraph and sublevel set

- α-**sublevel set** of $f : \mathbf{R}^n \to \mathbf{R}$ is $C_\alpha = \{x \in \mathbf{dom} f \mid f(x) \le \alpha\}$
- sublevel sets of convex functions are convex sets (but converse is false)
- **epigraph** of $f : \mathbf{R}^n \to \mathbf{R}$ is $\mathbf{epi} f = \{(x, t) \in \mathbf{R}^{n+1} \mid x \in \mathbf{dom} f,\ f(x) \le t\}$



- f is convex if and only if $\mathbf{epi} f$ is a convex set

Jensen's inequality

- **basic inequality:** if f is convex, then for $x, y \in \textbf{dom} f$, $0 \le \theta \le 1$,

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

- **extension:** if f is convex and z is a random variable on $\textbf{dom} f$,

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

- basic inequality is special case with discrete distribution

$$\textbf{prob}(z = x) = \theta, \qquad \textbf{prob}(z = y) = 1 - \theta$$

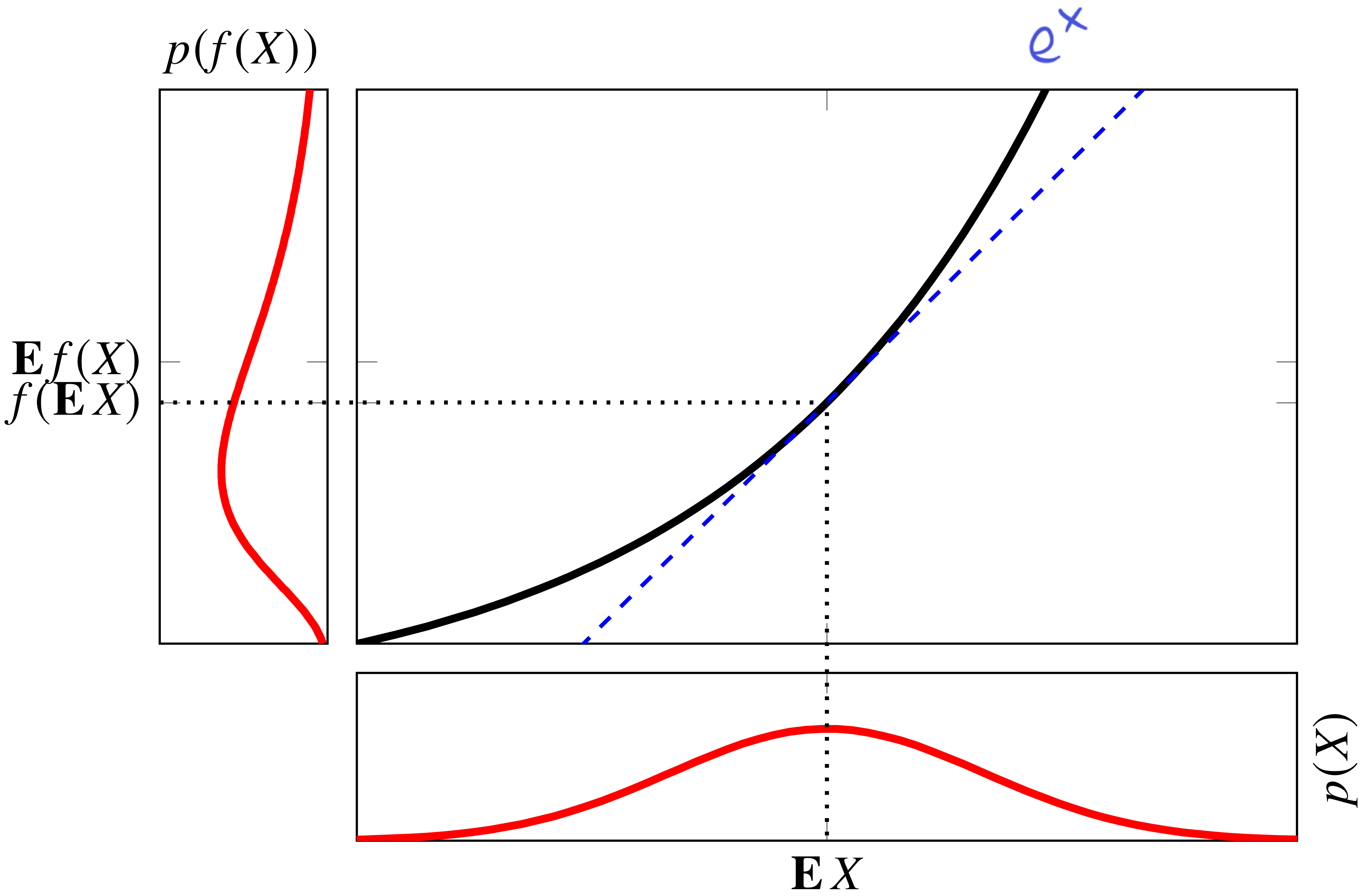
Example: log-normal random variable

- suppose $X \sim \mathcal{N}(\mu, \sigma^2)$
- with $f(u) = \exp u$, $Y = f(X)$ is log-normal
- we have $\mathbf{E} f(X) = \exp(\mu + \sigma^2/2)$
- Jensen's inequality is

$$f(\mathbf{E}\,X) = \exp \mu \leq \mathbf{E} f(X) = \exp(\mu + \sigma^2/2)$$

which indeed holds since $\exp \sigma^2/2 > 1$

Example: log-normal random variable



Outline

Showing a function is convex

methods for establishing convexity of a function f

1. verify definition (often simplified by restricting to a line)
2. for twice differentiable functions, show $\nabla^2 f(x) \succeq 0$
   - recommended only for **very simple** functions
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

you'll mostly use methods 2 and 3

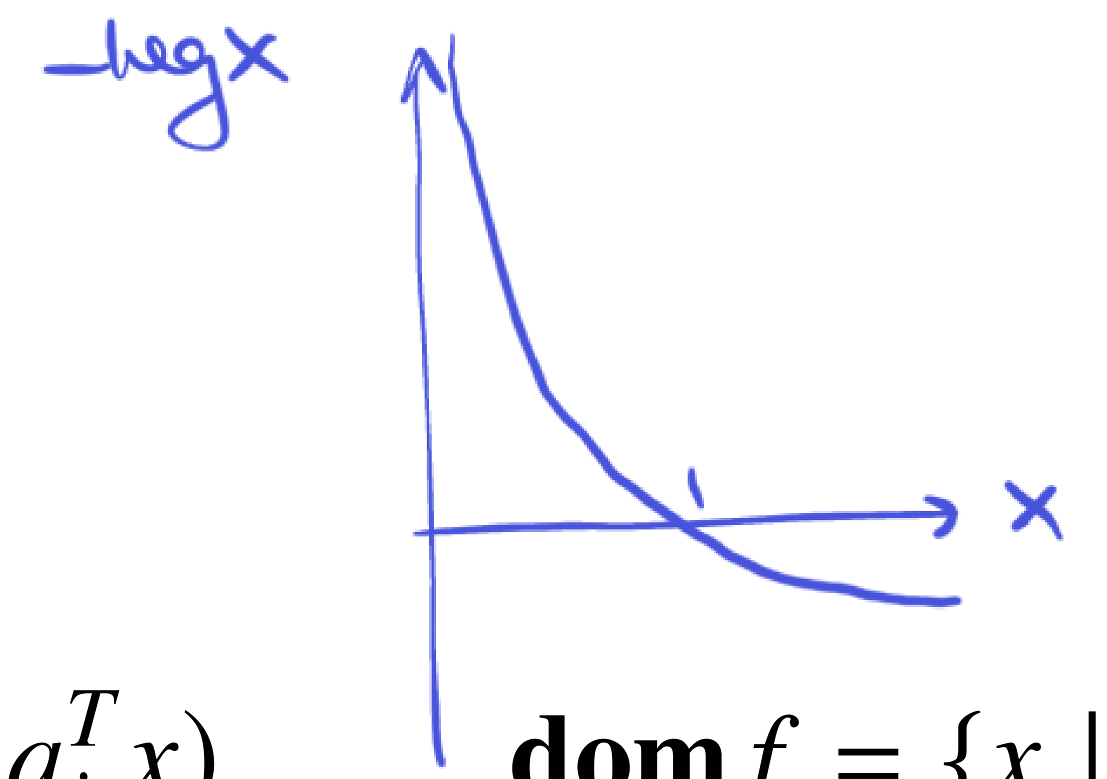
Nonnegative scaling, sum, and integral

- **nonnegative multiple:** αf is convex if f is convex, $\alpha \geq 0$
- **sum:** $f_1 + f_2$ convex if f_1, f_2 convex
- **infinite sum:** if $f_1, f_2, \ldots$ are convex functions, infinite sum $\sum_{i=1}^{\infty} f_i$ is convex
- **integral:** if $f(x, \alpha)$ is convex in x for each $\alpha \in \mathcal{A}$, then $\int_{\alpha \in \mathcal{A}} f(x, \alpha)\, d\alpha$ is convex

- there are analogous rules for concave functions

Composition with affine function

(pre-)composition with affine function: $f(Ax + b)$ is convex if f is convex



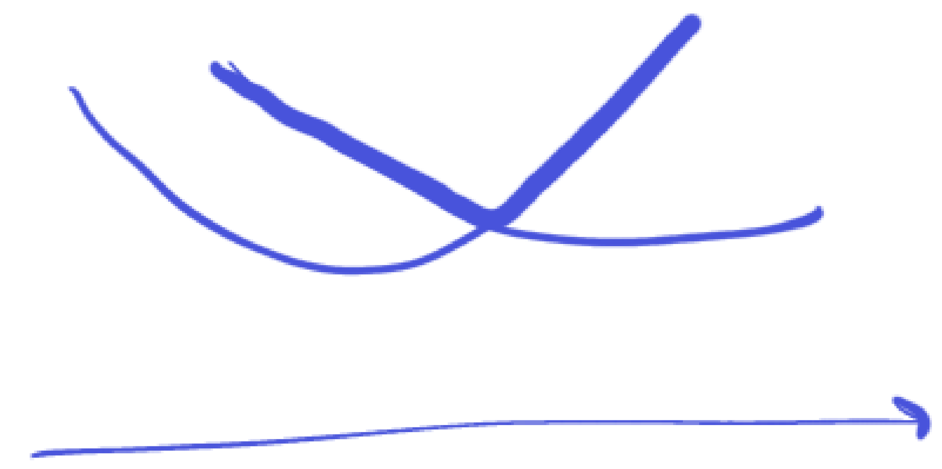
examples

- log barrier for linear inequalities

$$f(x) = -\sum_{i=1}^{m} \log(b_i - a_i^T x), \qquad \mathbf{dom}\, f = \{x \mid a_i^T x < b_i, i = 1, \ldots, m\}$$

- norm approximation error: $f(x) = \|Ax - b\|$ (any norm)

Pointwise maximum



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

examples

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

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

  ($x_{[i]}$ is ith largest component of x)

  proof: $f(x) = \max\{x_{i_1} + x_{i_2} + \cdots + x_{i_r} \mid 1 \le i_1 < i_2 < \cdots < i_r \le n\}$

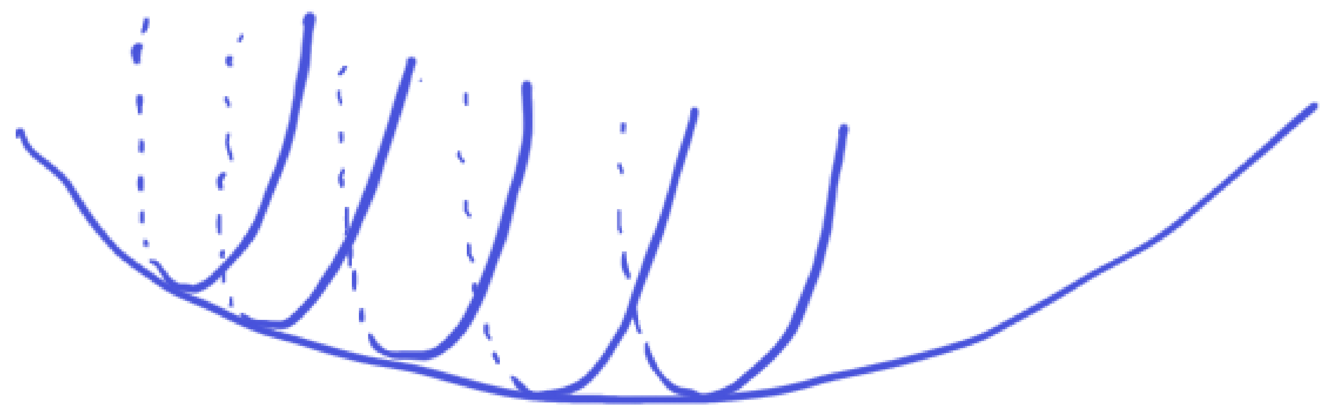
Pointwise supremum

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

examples

- distance to farthest point in a set C: $f(x) = \sup_{y \in C} \|x - y\|$
- maximum eigenvalue of symmetric matrix: for $X \in \mathbf{S}^n$, $\lambda_{\max}(X) = \sup_{\|y\|_2 = 1} y^T X y$ is convex
- support function of a set C: $S_C(x) = \sup_{y \in C} y^T x$ is convex

Partial minimization



- the function $g(x) = \inf_{y \in C} f(x, y)$ is called the **partial minimization** of f (w.r.t. y)
- if $f(x, y)$ is convex in (x, y) and C is a convex set, then partial minimization g is convex

examples

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

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \succeq 0, \qquad C \succ 0$$

  minimizing over y gives $g(x) = \inf_y f(x, y) = x^T (A - BC^{-1}B^T)x$
  g is convex, hence Schur complement $A - BC^{-1}B^T \succeq 0$

$ax^2 + 2bxy + cy^2$, $c > 0$.
$y = -\frac{bx}{c}$ is minimization
$ax^2 - 2\frac{b^2x^2}{c} + \frac{b^2x^2}{c}$
$= (a - b^2/c)x^2$.

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

Composition with scalar functions

- composition of $g : \mathbf{R}^n \to \mathbf{R}$ and $h : \mathbf{R} \to \mathbf{R}$ is $f(x) = h(g(x))$ (written as $f = h \circ g$)
- composition f is convex if
  - g convex, h convex, $\tilde{h}$ nondecreasing
  - or g concave, h convex, $\tilde{h}$ nonincreasing

  (monotonicity must hold for extended-value extension $\tilde{h}$)
- proof (for $n = 1$, differentiable g, h)

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

$f'(x) = h'(g(x))\, g'(x)$

Nonconvex example: e^{-x^2}

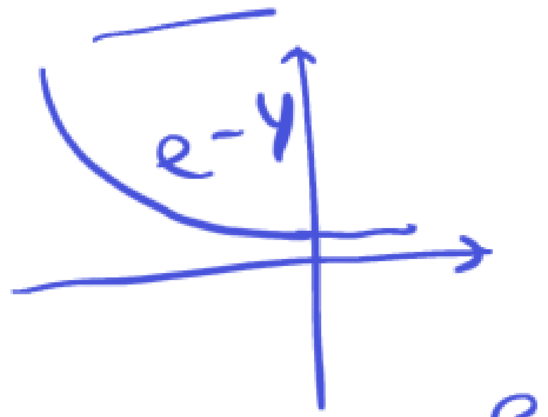
examples

- $f(x) = \exp g(x)$ is convex if g is convex
- $f(x) = 1/g(x)$ is convex if g is concave and positive   e.g. $-\frac{1}{\|x\|}$

General composition rule

- composition of $g : \mathbf{R}^n \to \mathbf{R}^k$ and $h : \mathbf{R}^k \to \mathbf{R}$ is $f(x) = h(g(x)) = h(g_1(x), g_2(x), \ldots, g_k(x))$
- f is convex if h is convex and for each i one of the following holds
  - g_i convex, $\tilde{h}$ nondecreasing in its ith argument
  - g_i concave, $\tilde{h}$ nonincreasing in its ith argument
  - g_i affine

e^{-x^2} vs e^{x^2}



convex, nondecreasing. $\exp(y) \circ x^2$

concave nonincreasing $\exp(-y) \circ (-x^2)$.

- you will use this composition rule **constantly** throughout this course
- you need to commit this rule to memory

Examples

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

(Recall. LogSumExp 's derivatives are softmax-arg) $\frac{e^{x_i}}{\sum_j e^{x_j}} > 0$

- $f(x) = p(x)^2/q(x)$ is convex if
  - p is nonnegative and convex
  - q is positive and concave

$\frac{x^2}{y}$ ← nondecreasing if $x \geq 0$, $y > 0$

- composition rule subsumes others, *e.g.*,
  - αf is convex if f is, and $\alpha \geq 0$
  - sum of convex (concave) functions is convex (concave)
  - max of convex functions is convex
  - min of concave functions is concave

Outline

Convex functions

Operations that preserve convexity

Constructive convex analysis

Perspective and conjugate

Quasiconvexity

Constructive convexity verification

- start with function f given as **expression**
- build parse tree for expression
  - leaves are variables or constants
  - nodes are functions of child expressions
- use composition rule to tag subexpressions as convex, concave, affine, or none
- if root node is labeled convex (concave), then f is convex (concave)
- extension: tag sign of each expression, and use sign-dependent monotonicity

- this is sufficient to show f is convex (concave), but not necessary
- this method for checking convexity (concavity) is readily automated

Example

the function

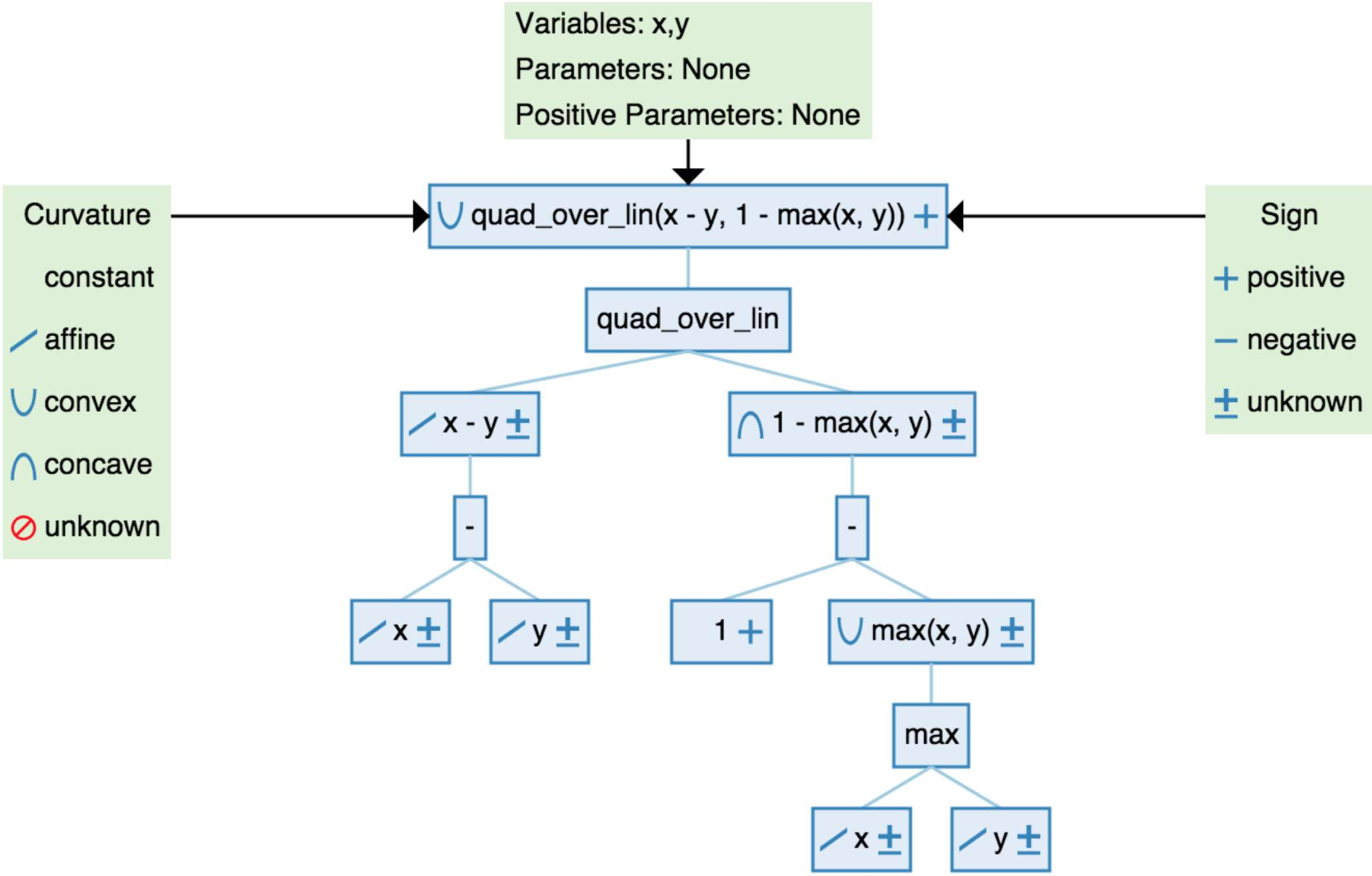
$$f(x, y) = \frac{(x - y)^2}{1 - \max(x, y)}, \qquad x < 1, \quad y < 1$$

is convex

constructive analysis:

- (leaves) x, y, and 1 are affine
- $\max(x, y)$ is convex; $x - y$ is affine
- $1 - \max(x, y)$ is concave
- function u^2/v is convex, monotone decreasing in v for $v > 0$
- f is composition of u^2/v with $u = x - y$, $v = 1 - \max(x, y)$, hence convex

Example (from dcp.stanford.edu)



Disciplined convex programming

in **disciplined convex programming** (DCP) users construct convex and concave functions as expressions using constructive convex analysis

- expressions formed from
  - **variables**,
  - **constants**,
  - and **atomic functions** from a library
- atomic functions have known convexity, monotonicity, and sign properties
- all subexpressions match general composition rule
- a valid DCP function is
  - convex-by-construction
  - 'syntactically' convex (can be checked 'locally')
- convexity depends only on attributes of atomic functions, not their meanings
  - *e.g.*, could swap $\sqrt{\cdot}$ and $\sqrt[4]{\cdot}$, or $\exp \cdot$ and $(\cdot)_+$, since their attributes match

CVXPY example

$$\frac{(x-y)^2}{1-\max(x,y)}, \qquad x < 1, \quad y < 1$$

```
import cvxpy as cp
x = cp.Variable()
y = cp.Variable()
expr = cp.quad_over_lin(x - y, 1 - cp.maximum(x, y))
expr.curvature # Convex
expr.sign # Positive
expr.is_dcp() # True
```

(atom `quad_over_lin(u,v)` includes domain constraint `v>0`)

DCP is only sufficient

- consider convex function $f(x) = \sqrt{1 + x^2}$
- expression `f1 = cp.sqrt(1+cp.square(x))` is **not** DCP

sqrt is concave.
So can't check by comp. rule

- expression `f2 = cp.norm2([1,x])` **is** DCP
- CVXPY will not recognize `f1` as convex, even though it represents a convex function

Outline

Convex functions

Operations that preserve convexity

Constructive convex analysis

Perspective and conjugate

Quasiconvexity

Perspective (transform)

- the **perspective** of a function $f : \mathbf{R}^n \to \mathbf{R}$ is the function $g : \mathbf{R}^n \times \mathbf{R} \to \mathbf{R}$,

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

- g is convex if f is convex

$(x,t,s) \in \text{epi } g \iff t f(\frac{x}{t}) \le s$

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

$\iff (\frac{x}{t}, \frac{s}{t}) \in \text{epi } f$

$\text{epi } g \xrightarrow{\text{perspective function}} \text{epi } f.$

examples

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

$= -t \log \frac{x}{t}$

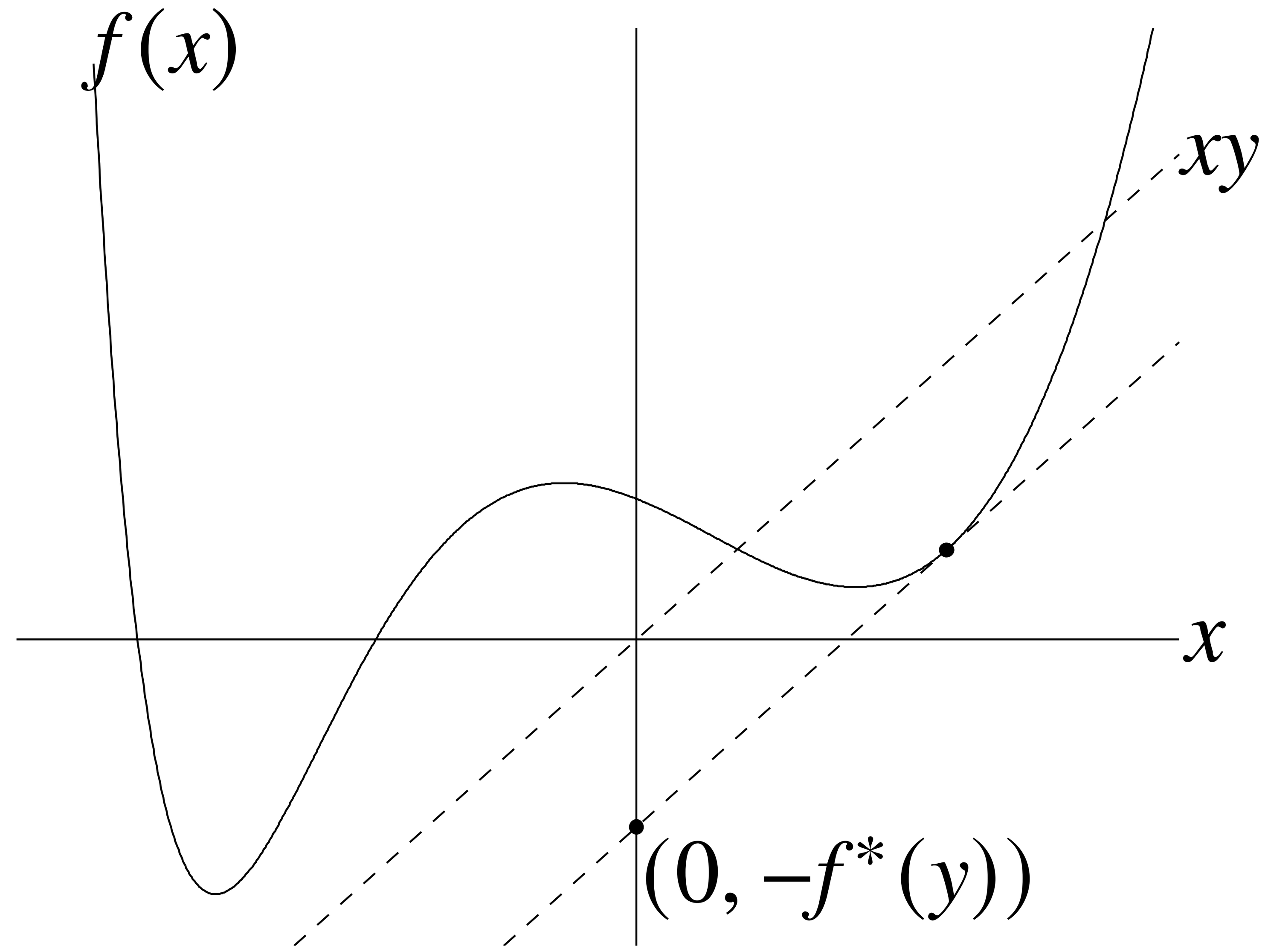
Maybe Confusion: The perspective map itself is not convex. (only defined for $x \in \mathbb{R}$)

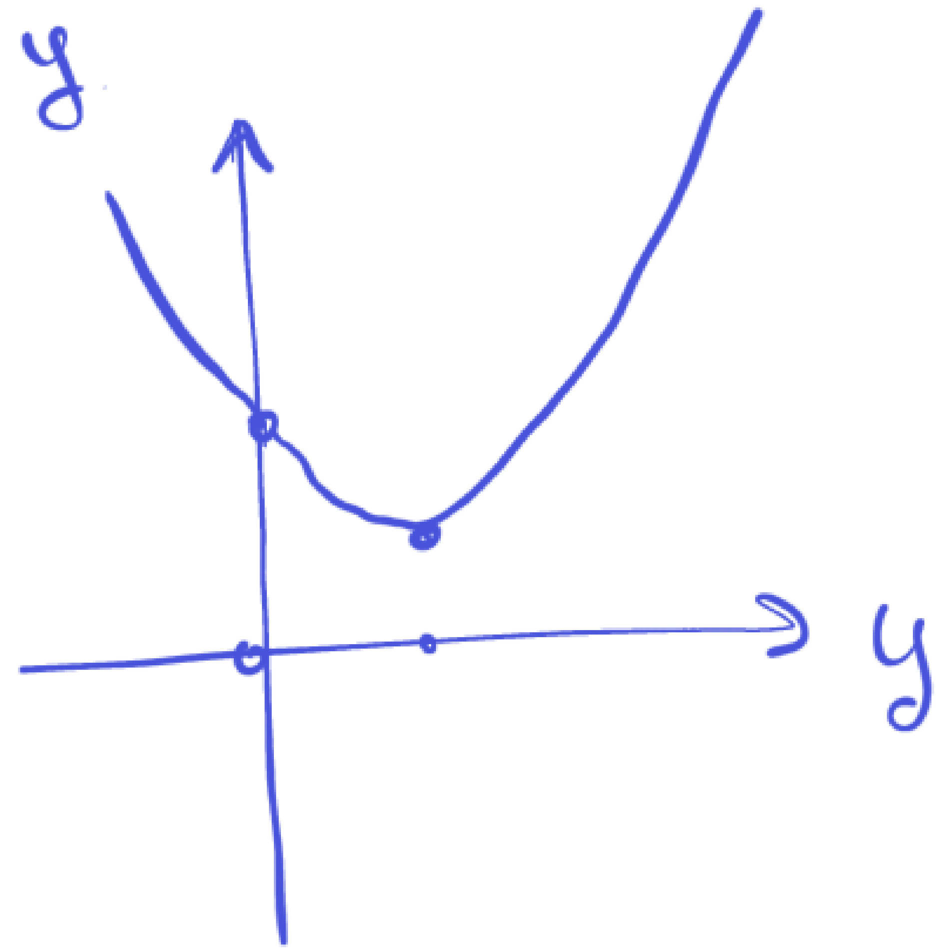
$P(x,t) = \frac{x}{t}$ $\quad \nabla P = \begin{bmatrix} \frac{1}{t} \\ -\frac{x}{t^2} \end{bmatrix}$ $\quad \nabla^2 P = \begin{bmatrix} 0 & -\frac{1}{t^2} \\ -\frac{1}{t^2} & 2\frac{x}{t^3} \end{bmatrix}$ $\quad \det = -\frac{1}{t^4} < 0.$

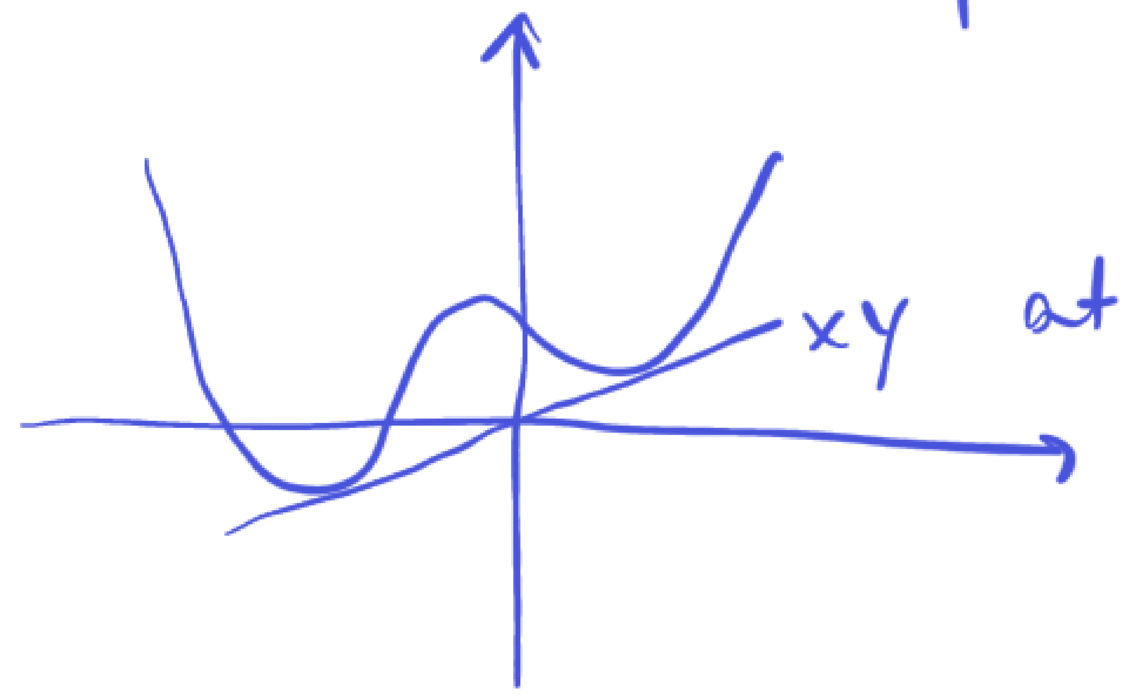
Conjugate function

= "Legendre transform": thermodynamics . . .

- the **conjugate** of a function f is $f^*(y) = \sup_{x \in \mathbf{dom} f} (y^T x - f(x))$ } each is affine in y.







- f^* is convex (even if f is not)
- will be useful in chapter 5

Examples

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

$z = -xy > 0. \quad xy + \log x = \log \frac{z}{|y|} - z = \log z - z - \log |y|.$

$\frac{d}{dz} = \frac{1}{z} - 1.$

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

- strictly convex quadratic, $f(x) = (1/2)x^TQx$ with $Q \in \mathbf{S}^n_{++}$

$$f^*(y) = \sup_x (y^Tx - (1/2)x^TQx) = \frac{1}{2}y^TQ^{-1}y$$

$\nabla_x = y - Qx$

$\Rightarrow x^* = Q^{-1}y$

- exponential. $f(x) = e^x$, $x \in \mathbb{R}$. $\Rightarrow f^*(y) = y \log y - y. \quad y > 0.$

= "Energy $\leftrightarrow$ Entropy"

$xy - e^x$:
- $y < 0$: $+\infty$.
- $y > 0$: $\frac{d}{dx}: y - e^x \Rightarrow x^* = \log y$
- $y = 0$: $\sup_x -e^x = 0.$

Example: Revenue and profit.

- resources consumed $r = (r_1, \dots, r_n)$
- sales revenue $S(r)$
- price per resource $p = (p_1, \dots, p_n)$

Cost. $p^T r$

Profit. $S(r) - p^T r$

Max profit $M(p)$. $\sup_r S(r) - p^T r$

$\Rightarrow$ Max profit is conjugate of sales.

$$M(p) = \sup_r (-p)^T r - (-S(r))$$
$$= (-S)^*(-p)$$

Properties

- Fenchel inequality: $f(x) + f^*(y) \geq x^T y$.

  e.g. $\frac{1}{2} x^T Q x + \frac{1}{2} y^T Q^{-1} y \geq x^T y$.

- Conjugate of conjugate: $f^{**} = f$ if f is convex and closed.
- Legendre transform (when f is differentiable and convex)

$$f^*(y) = x^T \nabla f(x) - f(x) \quad (y = \nabla f(x))$$

$\nabla_x (x^T y - f(x))$
$= y - \nabla f(x)$
$\Rightarrow x^*$ satisfy $y = \nabla f(x^*)$.

Conversely. x^* s.t. $\nabla f(x^*) = y$ maximizes $x^T y - f(x)$

- Thermodynamics

min — internal energy $U(S, V, N)$. $T = \frac{\partial U}{\partial S}$. $-p = \frac{\partial U}{\partial V}$.

max — enthalpy $\Rightarrow$. $H(S, p, N) = pV + U(S, V, N)$

min — Gibbs free energy $G(T, p, N) = H(S, p, N) - TS$

Outline

Convex functions

Operations that preserve convexity

Constructive convex analysis

Perspective and conjugate

Quasiconvexity

Quasiconvex functions

Generalization. Representable by a family of convex functions.

$f(x) \le t \iff \phi_t(x) \le 0$. convex.

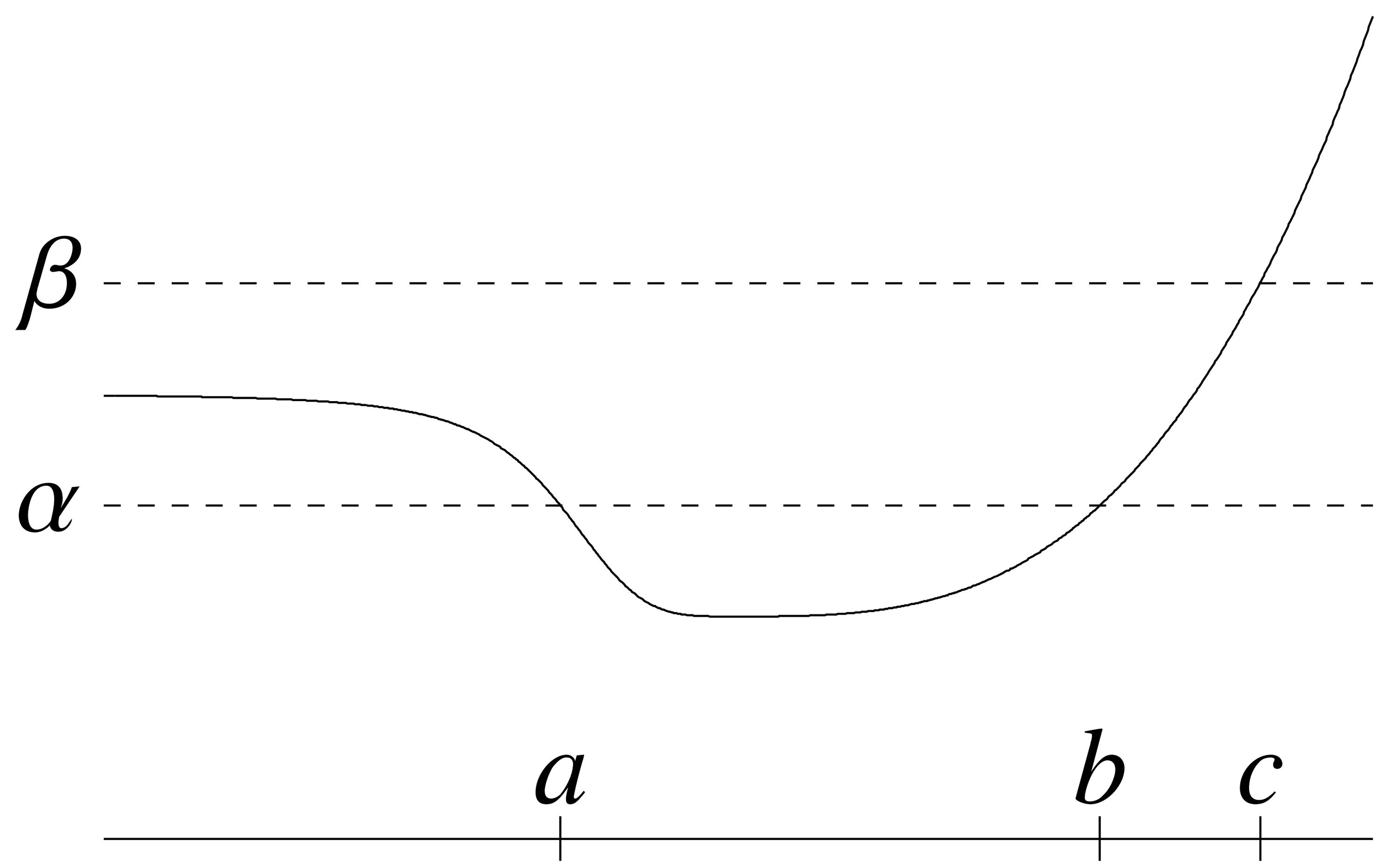
e.g. $\phi_t(x) = \begin{cases} 0, & f(x) \le t \\ \infty, & \text{else} \end{cases}$

$\phi_t(x) = \text{dist}(x, \{z : f(z) \le t\})$.

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

$$S_\alpha = \{x \in \textbf{dom} f \mid f(x) \le \alpha\}$$

are convex for all α



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

Examples

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

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

is quasilinear

$x_1 x_2 \geq \alpha.$

AM-GM inequality.

$\theta x_1^{(1)} x_2^{(1)} + (1-\theta) x_1^{(2)} x_2^{(2)} \geq (x_1^{(1)} x_2^{(1)})^{\theta} (x_1^{(2)} x_2^{(2)})^{(1-\theta)} \geq \alpha.$

Powerful!

- Length of a vector $f(x) = \max\{i \mid x_i \neq 0\}$. on $\mathbb{R}^n$. is quasiconvex.
- Cardinality of a vector. $\mathrm{card}(x) = \text{"}\|x\|_0\text{"} = \#$ nonzeros in x. is quasiconcave.
- Convex-over-concave: $f(x) = \frac{p(x)}{q(x)}$ $p(x) \geq 0$. $q(x) > 0$. $x \in C$ convex. is quasiconvex.

  $(f(x) \leq t \Leftrightarrow p(x) - q(x)t \leq 0)$

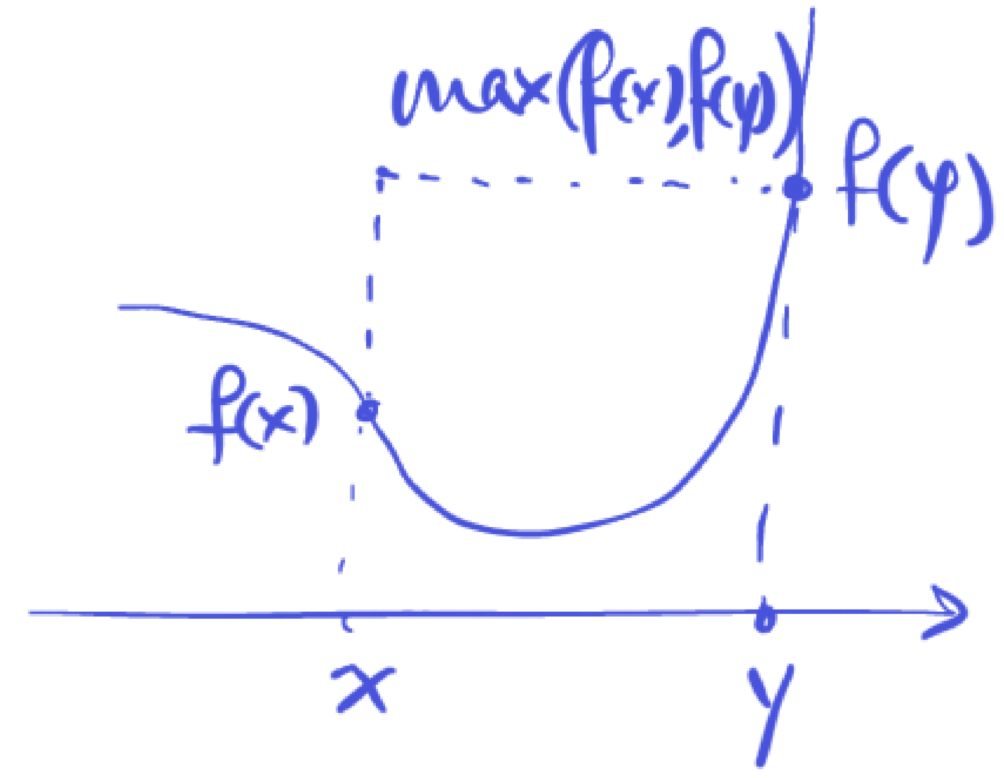
Example: Internal rate of return

- cash flow $x = (x_0, \ldots, x_n)$; x_i is payment in period i (to us if $x_i > 0$)
- we assume $x_0 < 0$ (*i.e.*, an initial investment) and $x_0 + x_1 + \cdots + x_n > 0$
- **net present value** (NPV) of cash flow x, for interest rate r, is $\mathrm{PV}(x, r) = \sum_{i=0}^{n} (1+r)^{-i} x_i$
- **internal rate of return** (IRR) is smallest interest rate for which $\mathrm{PV}(x, r) = 0$:

$$\mathrm{IRR}(x) = \inf\{r \geq 0 \mid \mathrm{PV}(x, r) = 0\}$$

- IRR is quasiconcave: superlevel set is intersection of open halfspaces

$$\mathrm{IRR}(x) \geq R \quad \Longleftrightarrow \quad \sum_{i=0}^{n} (1+r)^{-i} x_i > 0 \text{ for } 0 \leq r < R$$



Properties of quasiconvex functions

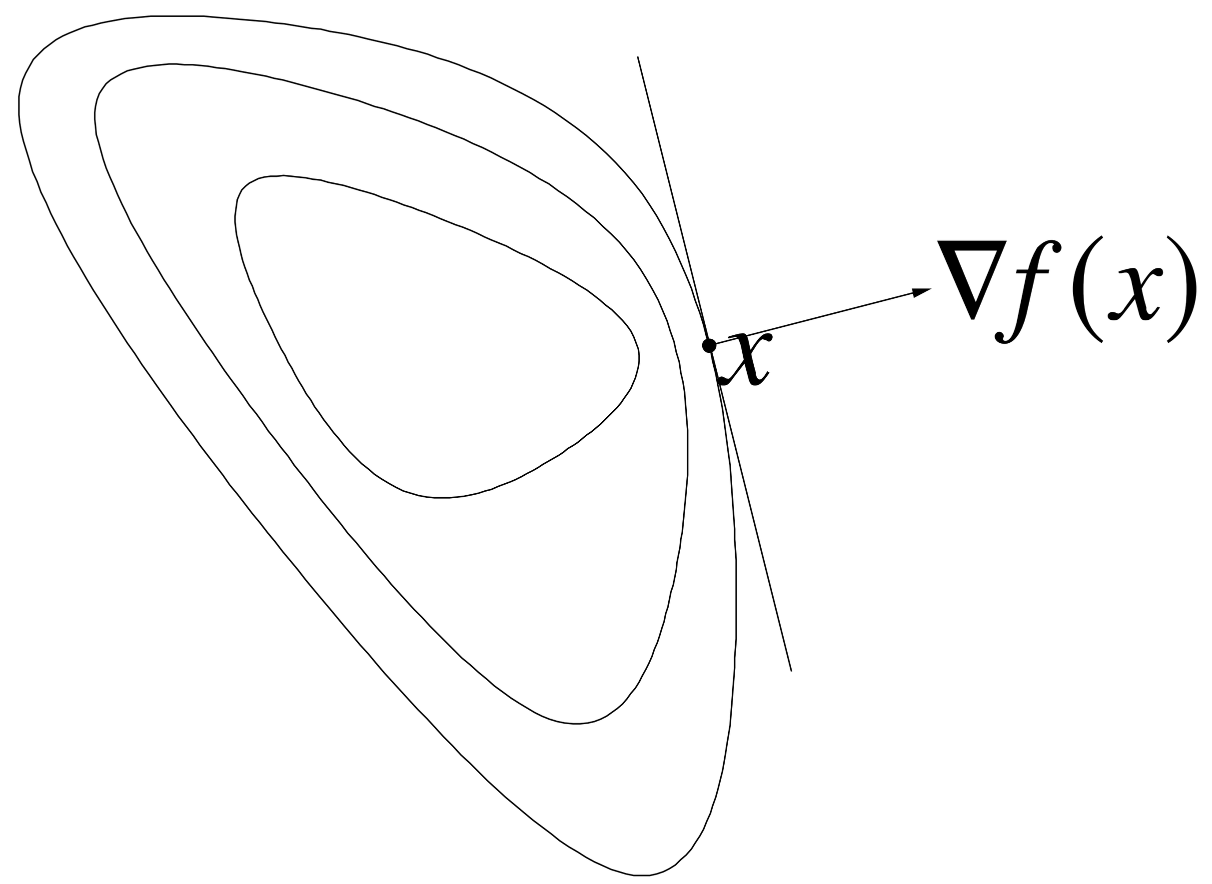
- **modified Jensen inequality:** for quasiconvex f

(equivalent def.)

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

- **first-order condition:** differentiable f with convex domain is quasiconvex if and only if

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



- **sum** of quasiconvex functions is not necessarily quasiconvex