3. Convex functions

Xinfu Liu

Outline

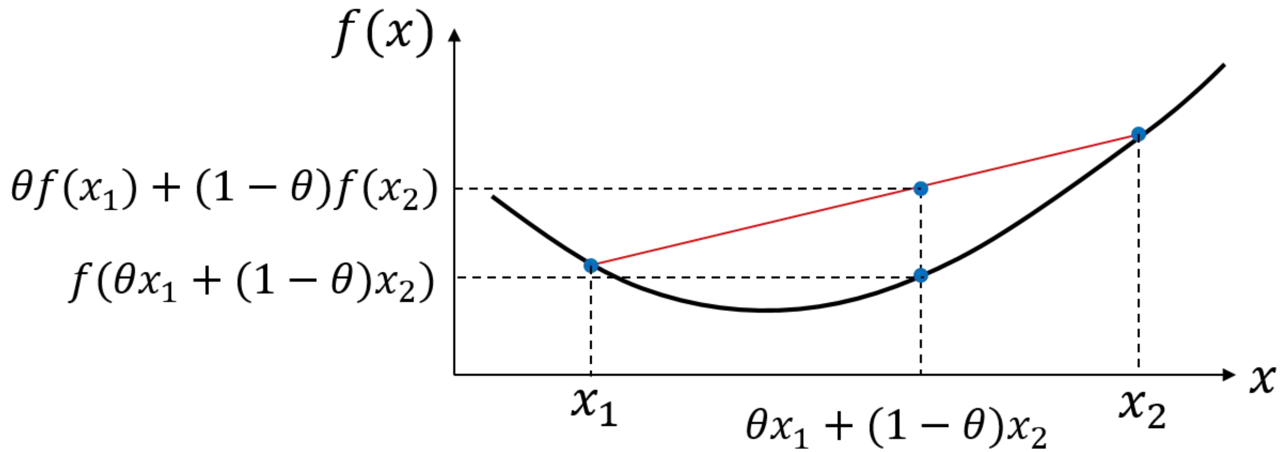
- basic properties and examples
- operations that preserve convexity
- the conjugate function
- quasi-convex functions
- log-concave and log-convex functions
- convexity with respect to generalized inequalities

Definition

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

$$f(\theta x_1 + (1-\theta)x_2) \leq \theta f(x_1) + (1-\theta) f(x_2)$$

for all $x_1, x_2 \in \mathbf{dom}\, f, 0 \leq \theta \leq 1$



- f is concave if $-f$ is convex
- f is **strictly convex** if $\mathbf{dom}\, f$ is convex and

$$f(\theta x_1 + (1-\theta)x_2 < \theta f(x_1) + (1-\theta) f(x_2)$$

for all $x_1, x_2 \in \mathbf{dom}\, f, x_1 \neq x_2,\ 0 < \theta < 1$

Restriction of a convex function to a line

A function is convex if and only if it is convex when restricted to any line that intersects its domain.

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

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

is convex (in t) for any $x \in \mathbf{dom}\, f$, $v \in \boldsymbol{R}^n$

can check convexity of f by checking convexity of functions of one variable

example. $f : \boldsymbol{S}^n \to \boldsymbol{R}$ with $f(X) = \log\det X$, $\mathbf{dom}\, f = \boldsymbol{S}^n_{++}$

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

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

g is concave in t (for any choice of $X \succ 0$, V); hence f is concave

Extended-value extension

If f is convex, its extended-value extension $\tilde{f}$ is defined as

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

It can often **simplify notation** since we do not need to explicitly specify the domain. For example, for $0 \leq \theta \leq 1$, the condition

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

for any x and y means **the same as the following two conditions**

- $\mathbf{dom}\, f$ is convex
- for $x,\ y \in \mathbf{dom}\, f$ and $0 \leq \theta \leq 1$

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

First-order condition

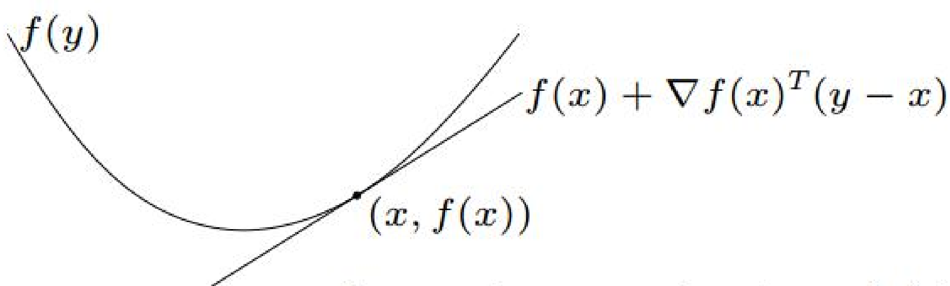
f is **differentiable** if $\mathbf{dom}\ f$ is open and the gradient

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \ldots, \frac{\partial f(x)}{\partial x_n}\right)$$

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

1st-order condition: differentiable f with convex domain is convex iff

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



first-order approximation of f is global underestimator

Q: what if $\nabla f(x)$=0?

Second-order condition

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

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

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

2nd-order condition: for twice differentiable f, it is convex if and only if $\mathbf{dom}\, f$ is convex and its Hessian is positive semidefinite:

$$\nabla^2 f(x) \succcurlyeq 0 \text{ for all } x \in \mathbf{dom}\, f$$

- if $\nabla^2 f(x) \succ 0$ for all $x \in \mathbf{dom}\, f$, then f is **strictly convex**

Examples on $\boldsymbol{R}$

convex:

- affine: $ax + b$ on $\boldsymbol{R}$, for any $a, b \in \boldsymbol{R}$
- exponential: e^{ax}, for any $a \in \boldsymbol{R}$
- powers: x^{α} on $\boldsymbol{R}_{++}$ for $\alpha \geq 1$ or $\alpha \leq 0$
- powers of absolute value: $|x|^p$ on $\boldsymbol{R}$, for $p \geq 1$
- negative entropy: $x \log x$ on $\boldsymbol{R}_{++}$

concave:

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

Examples on $\boldsymbol{R}^n$ and $\boldsymbol{R}^{m \times n}$

example on $\boldsymbol{R}^n$

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

  all **norms** are convex

examples on $\boldsymbol{R}^{m \times n}$ ($m \ \times \ n$ matrices)

- affine function

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

- spectral (maximum singular value) norm

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

Example

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

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

convex if $P \succcurlyeq 0$

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

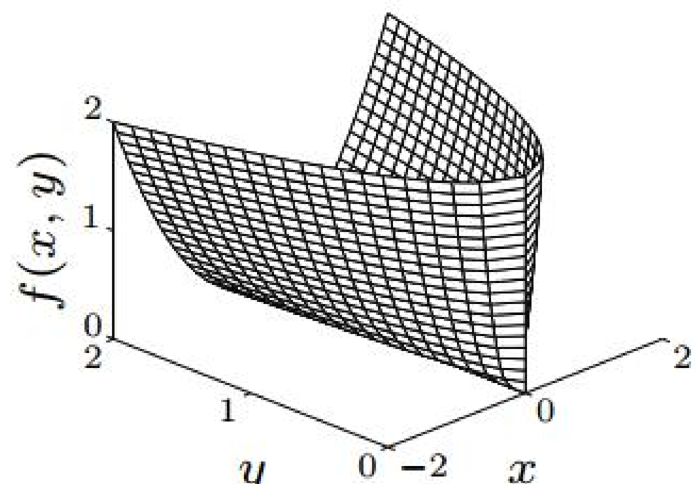
$$\nabla f(x) = 2A^T(Ax - b), \; \nabla^2 f(x) = 2A^T A$$

convex (for any A)

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

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

convex for $y > 0$



Example

$f(x,\ y) =\ x^2y^2$

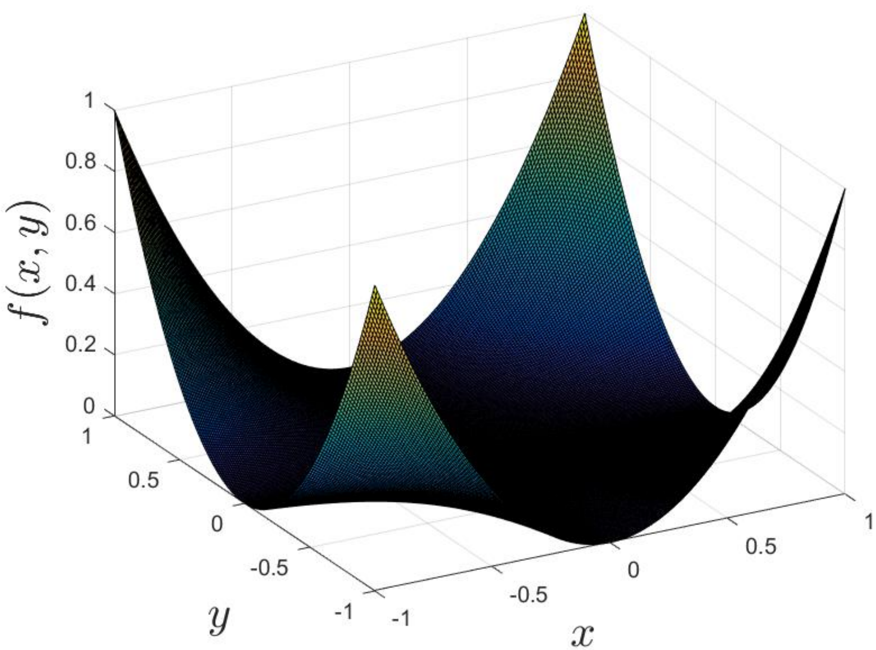
dom $f = \{(x,y) | x > 0, y > 0\}$

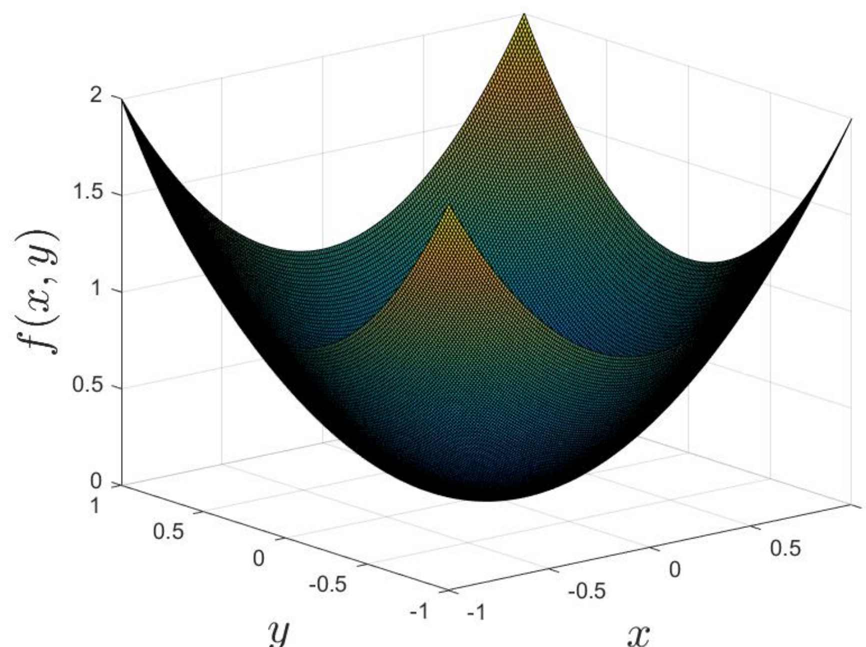
$\nabla^2 f(x,\ y) = \begin{bmatrix} 2y^2 & 4xy \\ 4xy & 2x^2 \end{bmatrix}$

$f(x,\ y) =\ x^2 + y^2$

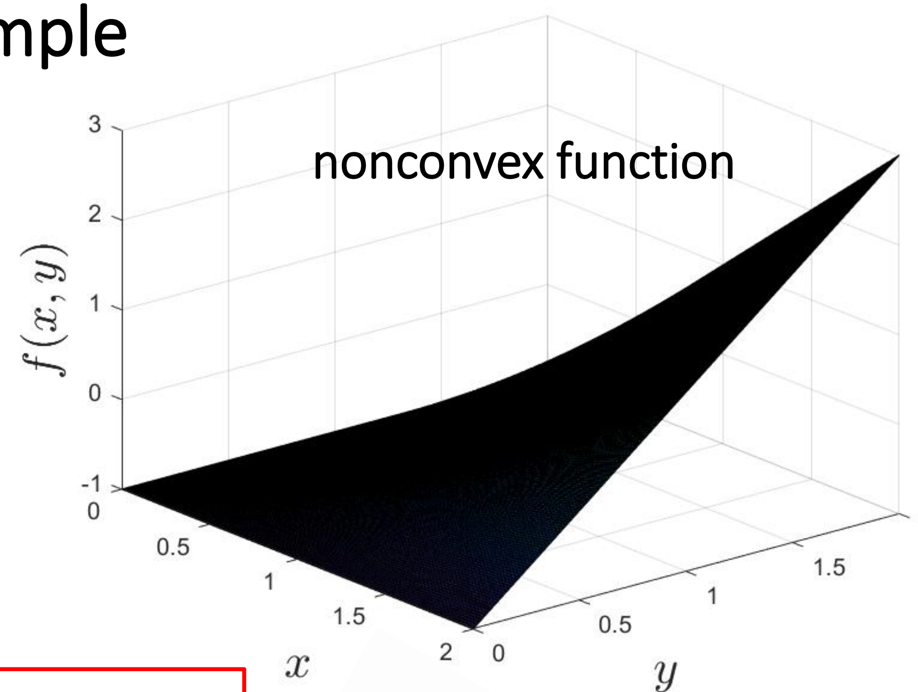
dom $f = \{(x,y) | x > 0, y > 0\}$

$\nabla^2 f(x,\ y) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$





Example

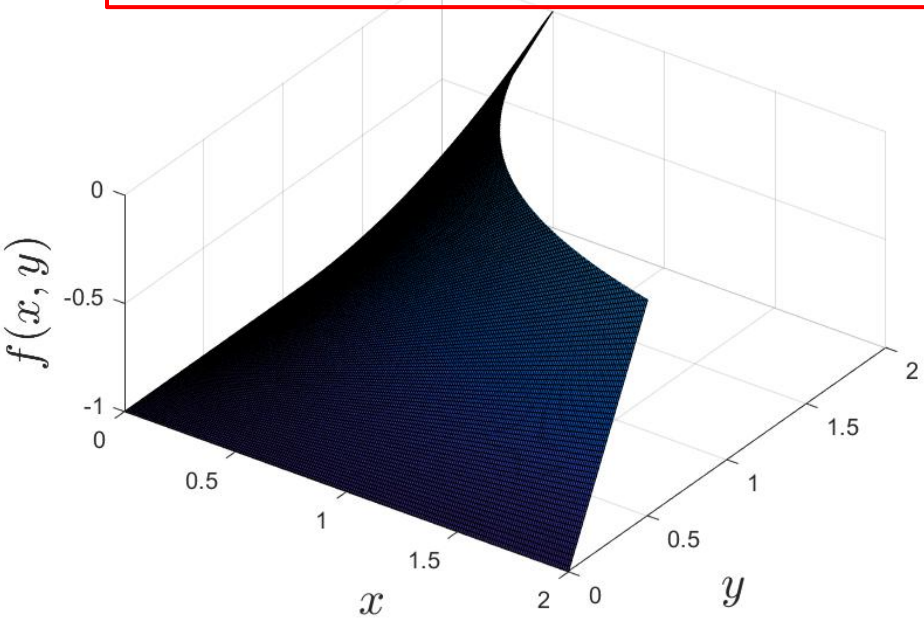


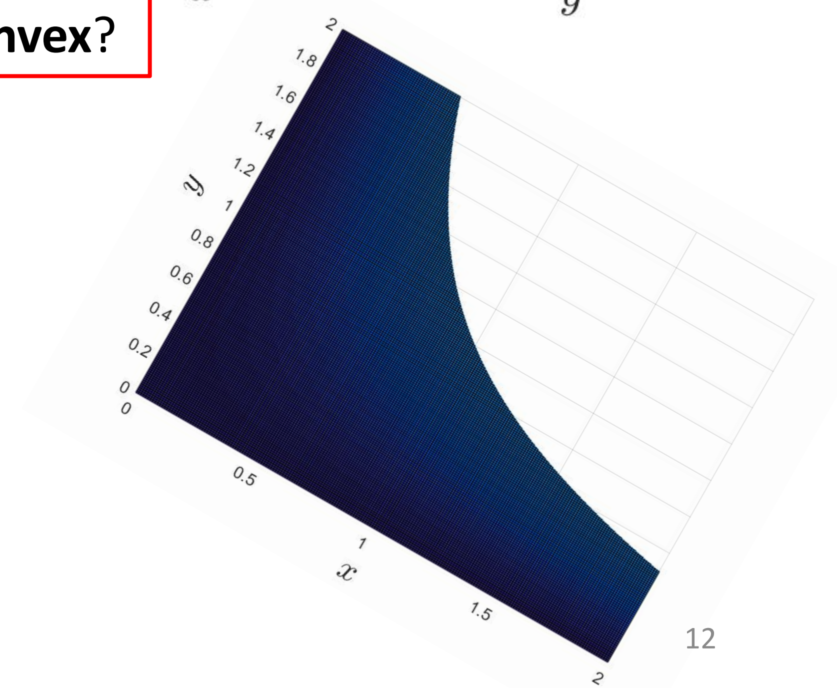
$$f(x,\ y) = xy - 1$$

$$\textbf{dom}\ f = \{(x, y) | x > 0, y > 0\}$$

$$\nabla^2 f(x,\ y) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

constraint: $xy - 1 \leq 0$ **convex or nonconvex**?

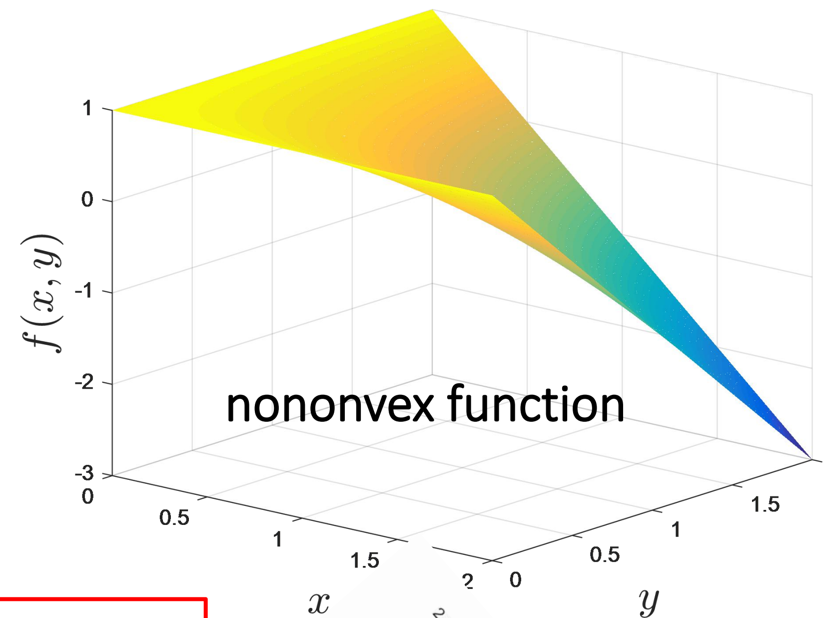




Example

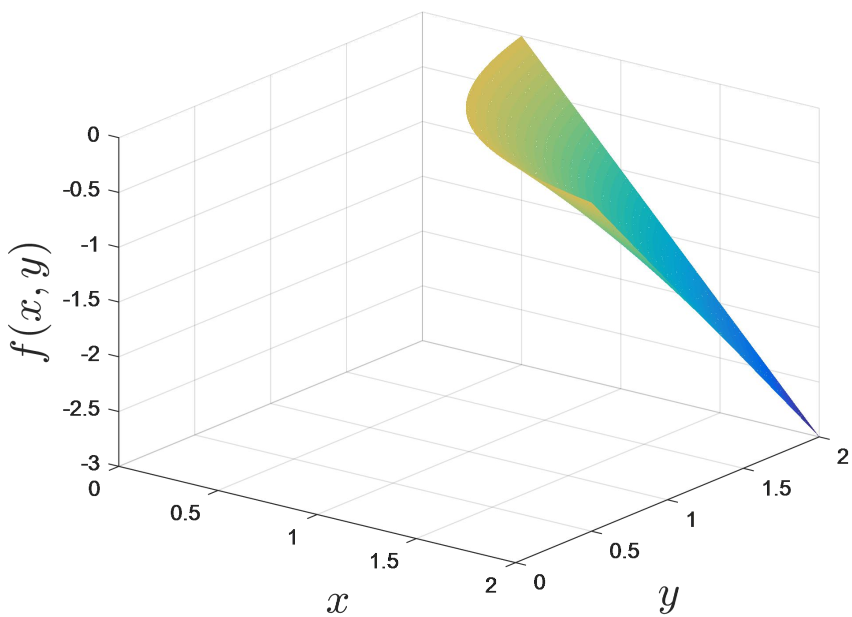
$f(x,\ y) = 1 - xy$

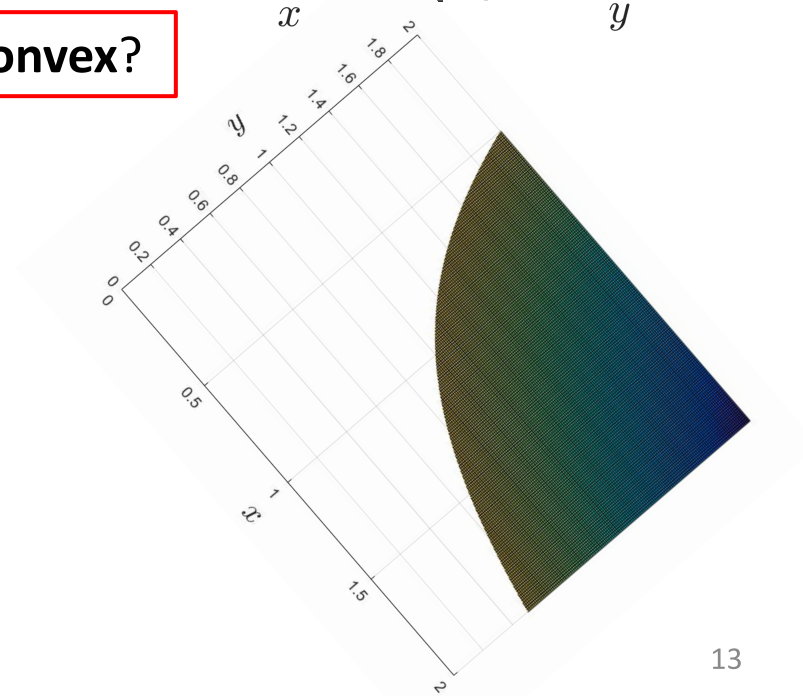
dom $f\ = \{(x, y)|x > 0, y > 0\}$



$\nabla^2 f(x,\ y) = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$

constraint: $1 - xy \leq 0$ **convex or nonconvex**?





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

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

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

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

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)

geometric mean: $f(x) = \left(\prod_{k=1}^{n} x_k\right)^{1/n}$ on $\boldsymbol{R}_{++}^n$ is concave

(similar proof as for log-sum-exp)

Sublevel set and epigraph

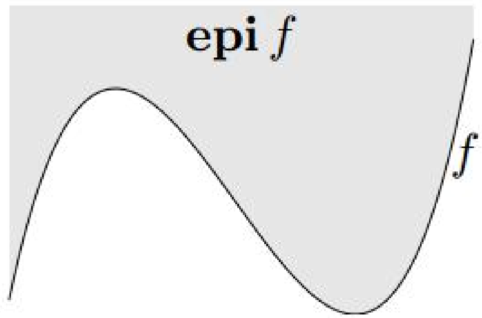
α-**sublevel set** of $f : \boldsymbol{R}^n \to \boldsymbol{R}$:

$$C_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \leq \alpha\}$$

sublevel sets of convex functions are convex (converse is false, examples?)

epigraph of $f : \boldsymbol{R}^n \to \boldsymbol{R}$:

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



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

Operations that preserve convexity

practical methods for establishing convexity of a function

1. verify definition (often simplified by restricting to a line)
2. for twice differentiable functions, show $\nabla^2 f(x) \succcurlyeq 0$
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

Positive weighted sum & composition with affine function

nonnegative multiple: αf is convex if f is convex, $\alpha \geq 0$

sum: $f_1 + f_2$ convex if $f_1,\ f_2$ convex (extends to infinite sums, integrals)

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

examples

- log barrier for linear inequalities

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

- (any) norm of affine function: $f(x) = \|Ax + b\|$

Pointwise maximum

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

examples

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

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

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

proof:

$$f(x) = \max\left\{x_{i_1} + x_{i_2} + \ldots + x_{i_r} \,\middle|\, 1 \leq i_1 < i_2 < \ldots < i_r \leq n\right\}$$

Pointwise supremum

if $f(x,\ y)$ is convex in x for each $y \in A$, then

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

is convex

examples

- support function of a set C: $S_C(x) = \sup_{y \in C} y^T x$ is convex
- distance to farthest point in a set C:

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

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

$$\lambda_{max}(X) = \sup_{\|y\|_2 = 1} y^T X y$$

Composition with scalar functions

composition of $g: \boldsymbol{R}^n \to \boldsymbol{R}$ and $h: \boldsymbol{R} \to \boldsymbol{R}$:

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

- for $n = 1$, and differentiable g, h, we can get

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

f is convex if
- h convex, h nondecreasing, and g convex
- h convex, h nonincreasing, and g concave

f is concave if
- h concave, h nondecreasing, and g concave
- h cocave, h nonincreasing, and g convex

examples

- $\exp g(x)$ is convex if g is convex
- $1/g(x)$ is convex if g is concave and positive

Vector composition

composition of $g: \boldsymbol{R}^n \to \boldsymbol{R}^k$ and $h: \boldsymbol{R}^k \to \boldsymbol{R}$:

$$f(x) = h\big(g(x)\big) = h(g_1(x), g_2(x), \ldots, g_k(x)$$

- for $n = 1$, and differentiable $g,\ h$, we can get

$$f''(x) = g'(x)^T \nabla^2 h(g(x)) g'(x) + \nabla h\big(g(x)\big)^T g''(x)$$

f is convex if
- h convex, h nondecreasing in each argument, and g_i convex
- h convex, h nonincreasing in each argument, and g_i concave

f is concave if
- h concave, h nondecreasing in each argument, and g_i concave
- h concave, h nonincreasing in each argument, and g_i convex

examples

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

Minimization

if $f(x,\ y)$ is convex in $(x,\ y)$ and C is a convex set, then

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

is convex

examples (Schur complement)

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

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \succcurlyeq 0,\ \ C\ \succ\ 0 \quad \text{(means f(x,y) is a convex function)}$$

minimizing over y gives $g(x) = \inf_{y} f(x, y)\ \ = x^T(A\ - BC^{-1}B^T\ )x$

$\boxed{\text{since } g \text{ is convex, we have } \textbf{Schur complement } A - BC^{-1}B^T \succcurlyeq 0}$

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

Perspective

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

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

$\boxed{g \text{ is convex if } f \text{ is convex}}$

examples

- $f(x)\ = x^T x$ is convex; hence $g(x,\ t)\ = x^T x/t$ is convex for $t\ >\ 0$
- negative logarithm $f(x) =- \log x$ is convex; hence relative entropy

  $g(x,\ t) = t\,\mathrm{log}t - t\,\log x$ is convex on $\boldsymbol{R}^2_{++}$
- if f is convex, then

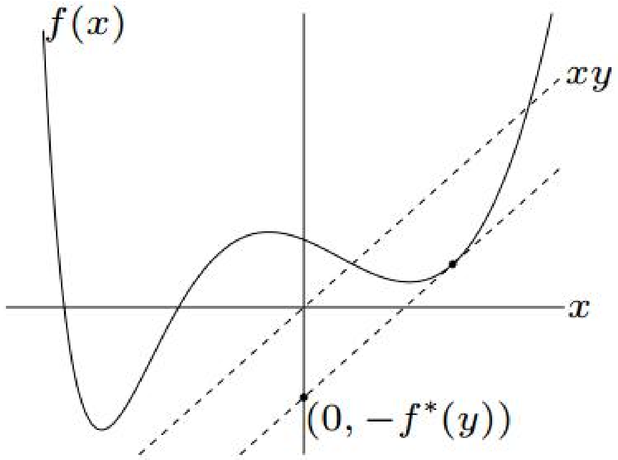
$$g(x) = (c^T x\ +\ d) f(\frac{Ax\ +\ b}{c^T x\ +\ d})$$

  is convex on $\{x \mid c^T x + d > 0, (Ax\ +\ b)/(c^T x + d) \in \mathbf{dom}\, f\}$

The conjugate function

the **conjugate** of a function $f(x)$ is

$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} \left(y^T x - f(x)\right)$$



the domain of the conjugate function: consists of y for which the supremum is **finite**, i.e., $y^T x - f(x)$ is bounded above on **dom** f

- f^* is **convex** (even if f is not). Why?
- will be **useful** in later lectures for "Duality"

examples

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

$$f^*(y) = \sup_{x>0} (xy + \log x)$$

$$= \begin{cases} -1 - \log(-y) & y < 0 \\ \infty & otherwise \end{cases}$$

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

$$f^*(y) = \sup_x (y^T x - (1/2)x^T Q x)$$

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

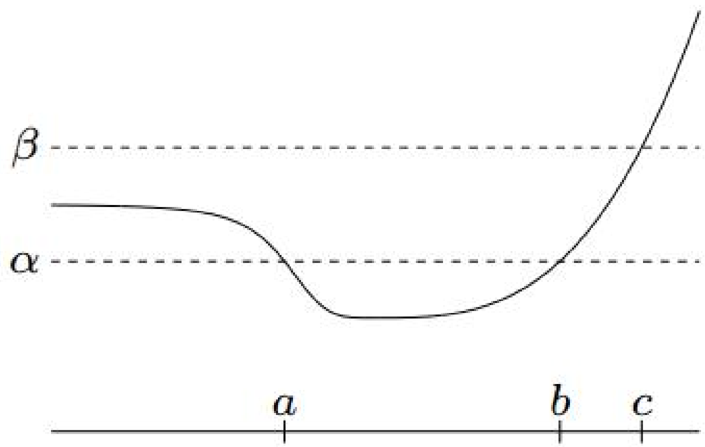
Q: for function f(x)=1/x, x>0, what is its conjugate function?

Quasiconvex functions

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

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

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

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

  is quasi-linear

- distance ratio

$$f(x) = \frac{\|x - a\|_2}{\|x - b\|_2}, \ \ \mathbf{dom}\, f = \{x \mid \|x - a\|_2 \leq \|x - b\|_2\}$$

  is quasi-convex

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 $x_0 < 0$ and $x_0 + x_1 + \ldots + x_n > 0$
- present value of cash flow x, for interest rate r:

$$PV(x, r) = \sum_{i=0}^{n} (1+r)^{-i} x_i$$

- internal rate of return is smallest interest rate for which $PV(x, r) = 0$:

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

IRR is quasi-concave: superlevel set is intersection of open halfspaces

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

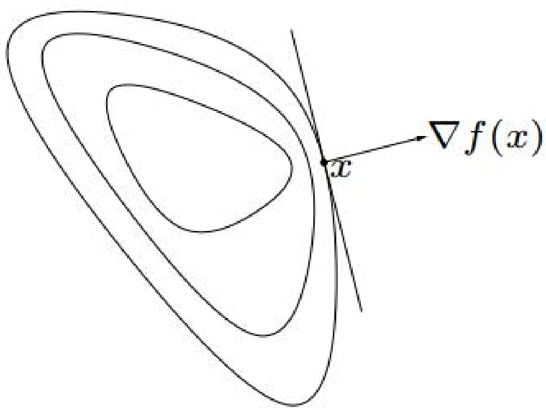
Properties

$f : \boldsymbol{R}^n \rightarrow \boldsymbol{R}$ is **quasiconvex** if $\mathbf{dom}\ f$ is convex and any $x, y \in \mathbf{dom}\ f$ and all $0 \leq \theta \leq 1$,

$$f(\theta x + (1-\theta)y) \leq \max\{f(x),\ f(y)\} \longleftarrow$$ **modified Jensen inequality**

first-order condition: differentiable $f: \boldsymbol{R}^n \rightarrow \boldsymbol{R}$ is **quasiconvex** if and only if $\mathbf{dom}\ f$ is convex and for all $x, y \in \mathbf{dom}\ f$

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



sums of quasiconvex functions are not necessarily quasiconvex

Log-concave and log-convex function

a positive function f is log-concave if $\log f$ is concave:

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

f is log-convex if $\log f$ is convex

- powers: x^a on $\boldsymbol{R}_{++}$ is log-convex for $a \; \leq \; 0$, log-concave for $a \; \geq \; 0$
- many common probability densities are log-concave, e.g. normal:

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

- cumulative Gaussian distribution function is log-concave

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

Properties of log-concave functions

- twice differentiable f with convex domain is log-concave if and only if

$$f(x)\nabla^2 f(x) \preccurlyeq \nabla f(x)\nabla f(x)^T$$

  for all $x \in \mathbf{dom}\, f$

- product of log-concave functions is log-concave
- sum of log-concave functions is not always log-concave
- integration: if $f : \boldsymbol{R}^n \times \boldsymbol{R}^m \to \boldsymbol{R}$ is log-concave, then

$$g(x) = \int f(x,\, y)\, dy$$

  is log-concave (not easy to show)

consequences of integration property

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

$$(f * g)(x) = \int f(x - y)g(y)dy$$

- if $C \subseteq \boldsymbol{R}^n$ convex and y is a random variable with log-concave pdf then

$$f(x) = \mathbf{prob}(x + y \in C)$$

  is log-concave

  proof: write $f(x)$ as integral of product of log-concave functions

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

  p is pdf of y

example: yield function

$$Y(x) = \mathbf{prob}\,(x + w \in S)$$

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

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

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

Convexity with respect to generalized inequalities

$f : \boldsymbol{R}^n \rightarrow \boldsymbol{R}^m$ is K-convex if $\mathbf{dom}\, f$ is convex and

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

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

example $f : S^m \rightarrow S^m, f(X) = X^2\ is\ S_+^m$ -convex

proof: for fixed $z \in R^m,\ z^T X^2 z = \|Xz\|_2^2$ is convex in X, *i.e.*

$$z^T (\theta X + (1 - \theta)Y)^2 z \leq \theta z^T X^2 z + (1 - \theta) z^T Y^2 z$$

for $X, Y \in S^m,\ 0 \leq \theta \leq 1$

therefore $(\theta X + (1 - \theta)Y)^2 \preccurlyeq \theta X^2 + (1 - \theta)Y^2$