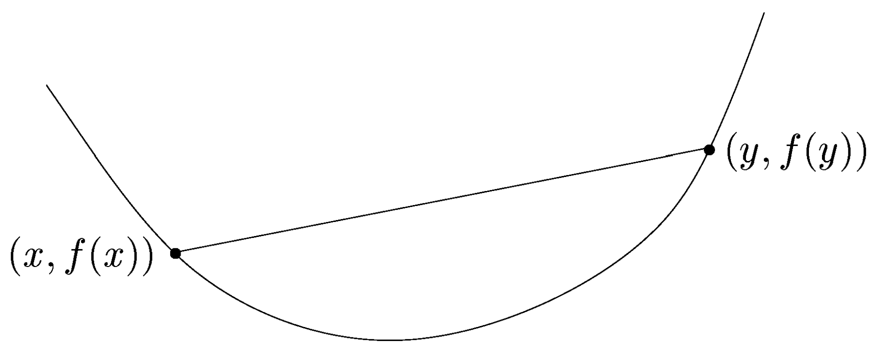
CONVEX FUNCTIONS

Definition

Convex function

A function $f : \mathbf{R}^n \to \mathbf{R}$ is *convex* if $\mathbf{dom}\, f$ is a convex set and if for all x, $y \in \mathbf{dom}\, f$, and θ with $0 \leq \theta \leq 1$, we have

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



Definition

Strictly convex function

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

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

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

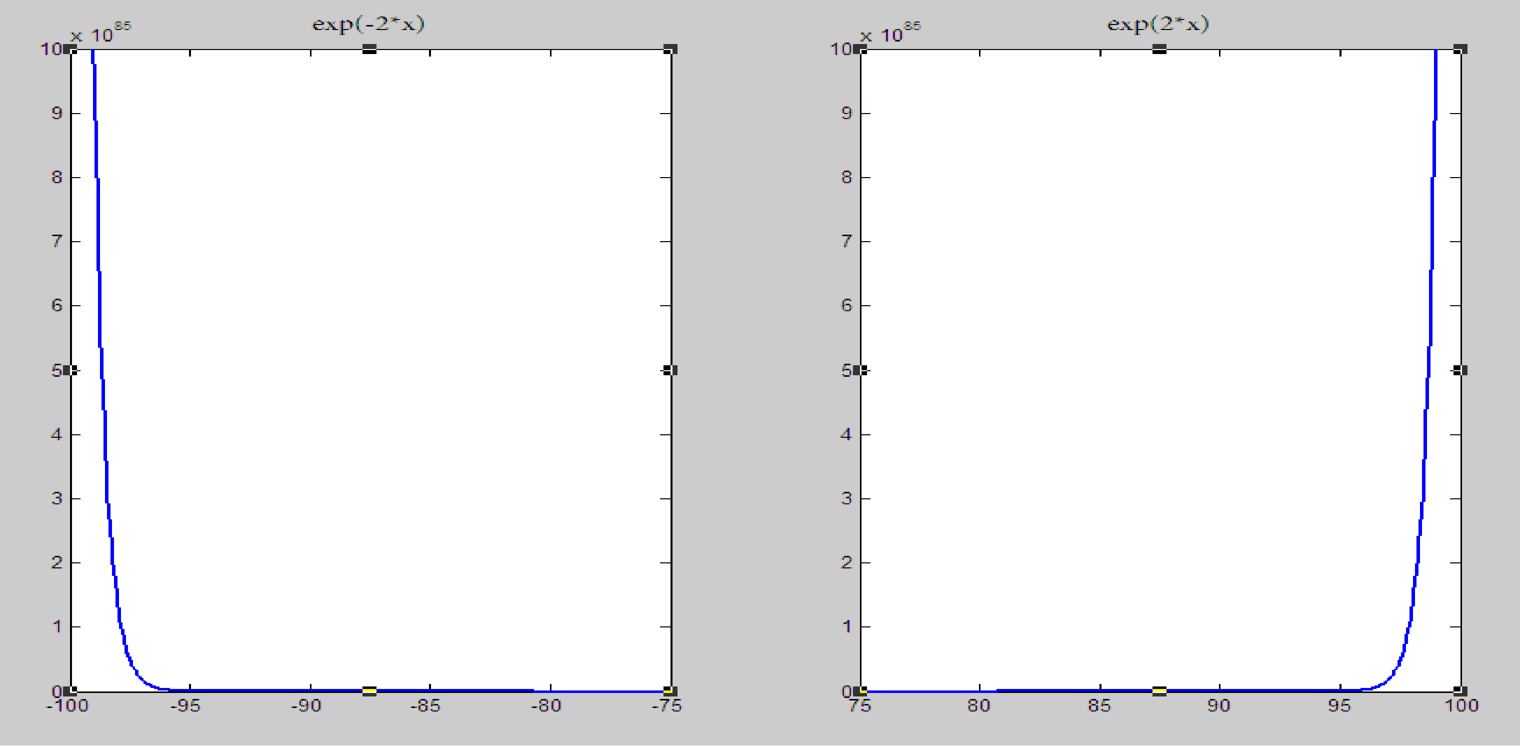
Concave function

f is concave if $-f$ is convex

Some examples

affine: $ax + b$ on $\mathbf{R}$, for any $a, b \in \mathbf{R}$ convex concave

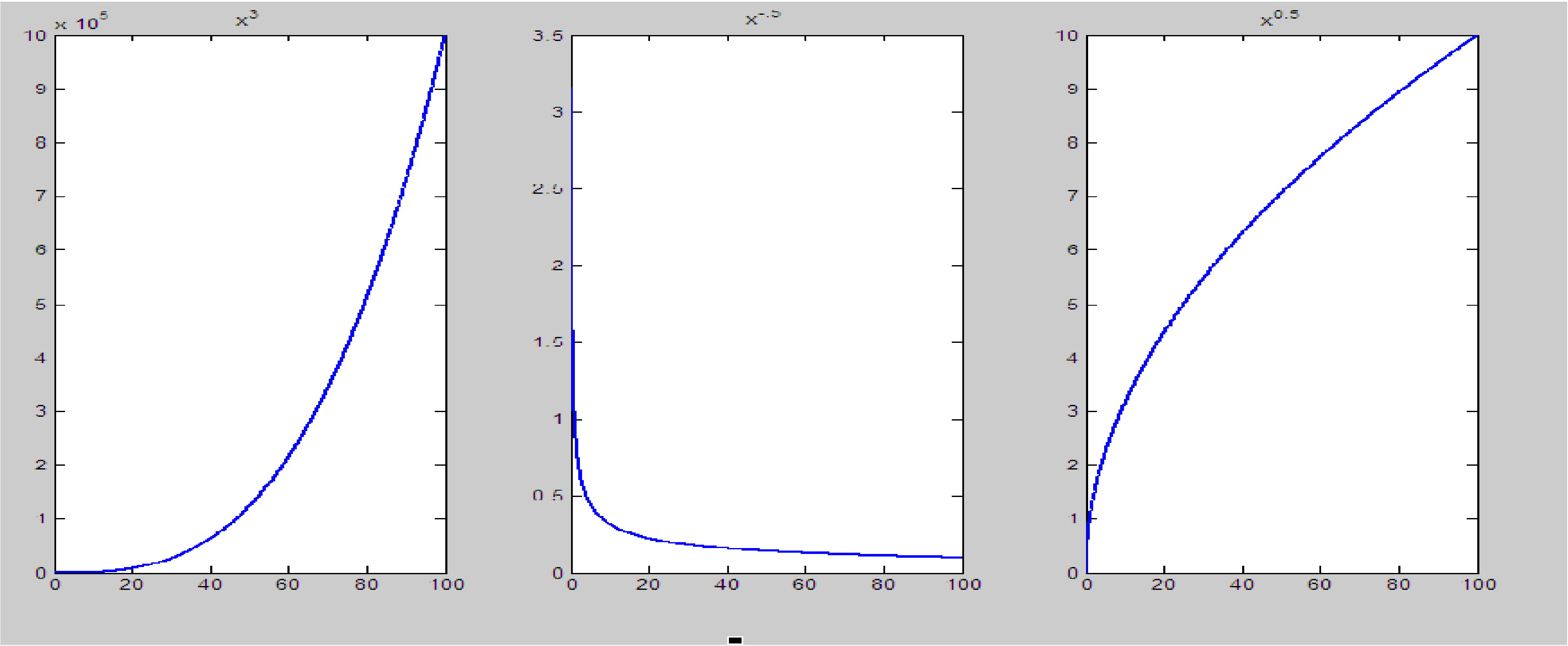
exponential: e^{ax}, for any $a \in \mathbf{R}$ convex



Some examples

powers: x^α on $\mathbf{R}_{++}$, for $\alpha \geq 1$ or $\alpha \leq 0$ **convex**

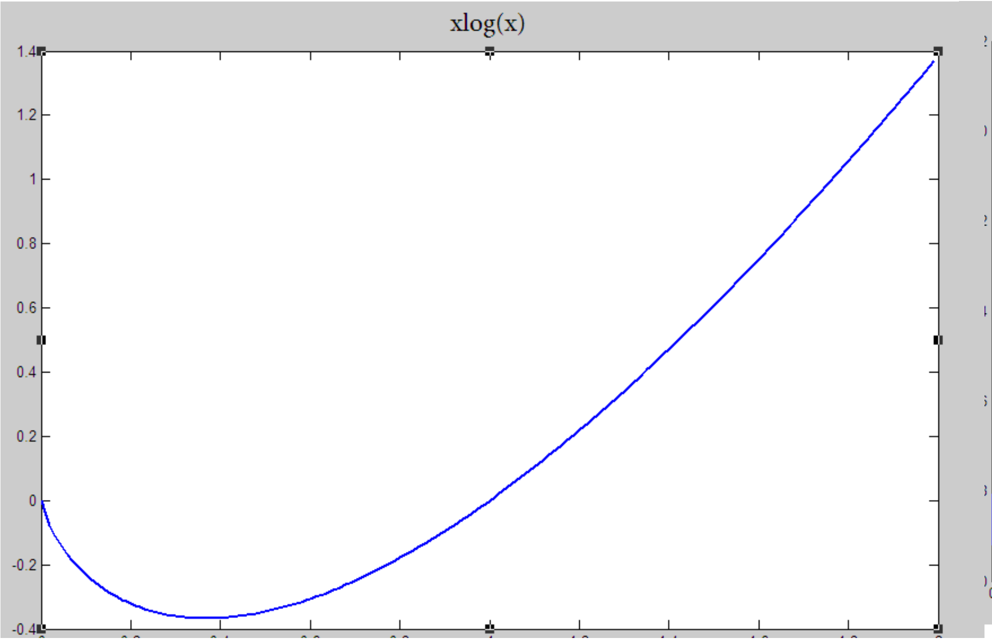
powers: x^α on $\mathbf{R}_{++}$, for $0 \leq \alpha \leq 1$ **concave**

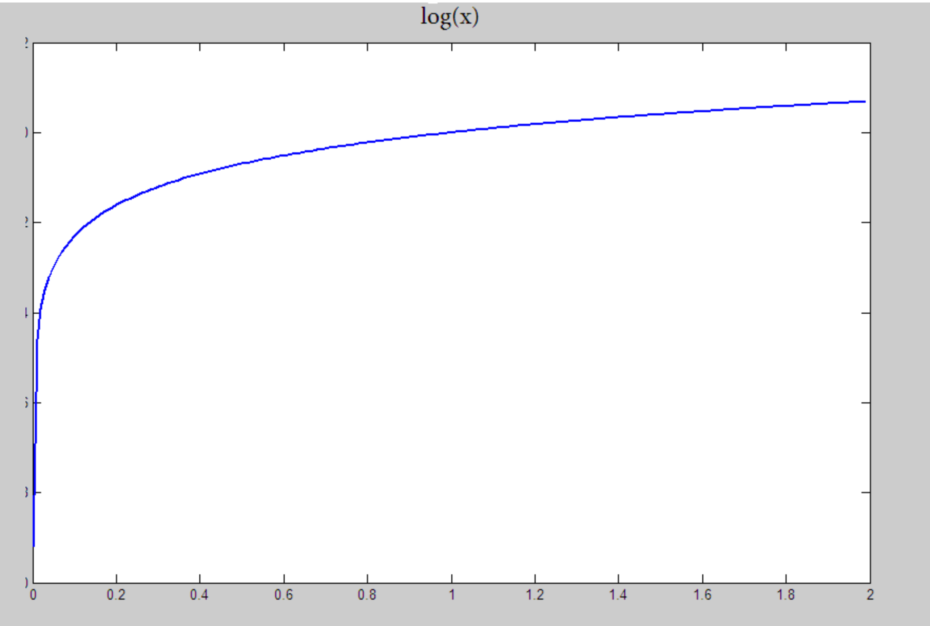


Some examples

negative entropy: $x \log x$ on $\mathbf{R}_{++}$ convex

logarithm: $\log x$ on $\mathbf{R}_{++}$ concave





Some examples

examples on $\mathbf{R}^n$

- affine function $f(x) = a^T x + b$
- norms

$$\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p} \text{ for } p \geq 1;$$

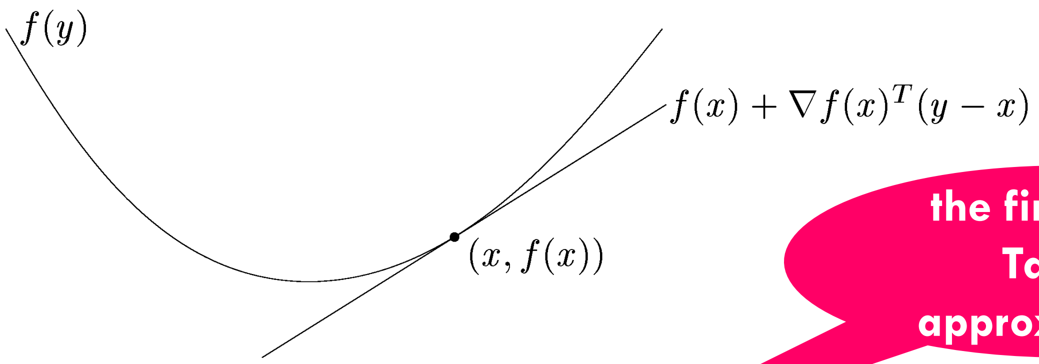
$$\|x\|_\infty = \max_k |x_k|$$

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

- affine function $$f(X) = \mathbf{tr}(A^T X) + b = \sum_{i=1}^m \sum_{j=1}^n A_{ij} X_{ij} + b$$
- spectral (maximum singular value) norm

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

First-order condition



the first-order Taylor approximation

Proof: page 70

$$f(y) \geq \underbrace{f(x) + \nabla f(x)^T (y - x)}$$

A global under-estimator

from local information about a convex function derive global information

First-order condition

For example

if $\nabla f(x) = 0$, then for all $y \in \mathbf{dom}\, f$, $f(y) \geq f(x)$, *i.e.*, x is a global minimizer

Strict convexity

$$f(y) > f(x) + \nabla f(x)^T (y - x).$$

concave

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

Second-order conditions

f is **twice differentiable** if 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 $x \in \mathbf{dom}\, f$

Strictly convex

f is convex if and only if $\mathbf{dom}\, f$ is convex and its Hessian is positive semidefinite:

$$\nabla^2 f(x) \succeq 0.$$

graph of the function have positive curvature

$\nabla^2 f(x) \preceq 0$ concave

Examples

$\boxed{\text{quadratic function}}$

$$f(x) = (1/2)x^T P x + q^T x + r$$

$P \in \mathbf{S}^n$, $q \in \mathbf{R}^n$, and $r \in \mathbf{R}$

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

convex if $P \succeq 0$

Examples

$\boxed{\textbf{least-squares objective}}$

$$f(x) = ||Ax - b||_2^2$$

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

convex (for any A)

examples

quadratic-over-linear

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

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

(for $y > 0$)

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

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

example

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

$$\frac{\partial^2 f(x)}{\partial x_k \partial x_l} = \frac{\left(\prod_{i=1}^{n} x_i\right)^{1/n}}{n^2 x_k x_l} \quad \text{for } k \neq l, \qquad \frac{\partial^2 f(x)}{\partial x_k^2} = -(n-1)\frac{\left(\prod_{i=1}^{n} x_i\right)^{1/n}}{n^2 x_k^2}$$

$$q_i = 1/x_i \qquad \nabla^2 f(x) = -\frac{\prod_{i=1}^{n} x_i^{1/n}}{n^2}\left(n\,\mathbf{diag}(1/x_1^2, \ldots, 1/x_n^2) - qq^T\right)$$

$$v^T \nabla^2 f(x) v = -\frac{\prod_{i=1}^{n} x_i^{1/n}}{n^2}\left(n\sum_{i=1}^{n} v_i^2/x_i^2 - \left(\sum_{i=1}^{n} v_i/x_i\right)^2\right) \leq 0$$

$$\boxed{|x^T y| \leq \|x\|_2 \|y\|_2 \text{ for any } x, y \in \mathbf{R}^n}$$

Sublevel sets and epigraphs

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

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

sublevel sets of convex functions are convex

Inverse is not true

The graph of a function $f : \mathbf{R}^n \to \mathbf{R}$ is defined as

$$\{(x, f(x)) \mid x \in \mathbf{dom}\, f\},$$

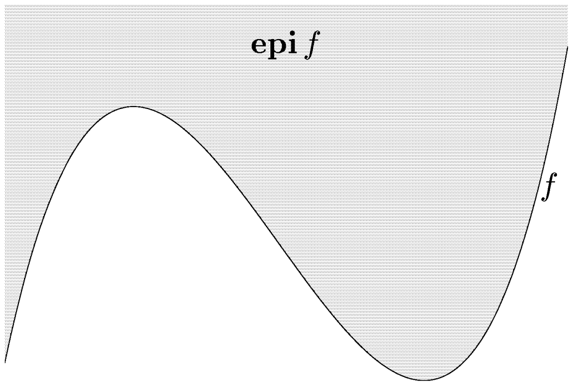
which is a subset of $\mathbf{R}^{n+1}$

The *epigraph* of a function $f : \mathbf{R}^n \to \mathbf{R}$ is defined as

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

which is a subset of $\mathbf{R}^{n+1}$.

Sublevel sets and epigraphs



✓The link between convex sets and convex functions is via the epigraph
✓A function is convex if and only if its epigraph is a convex set.

Establishing convexity

practical methods for establishing convexity of a function

verify definition

for twice differentiable functions, show PS of Hessian

show that f is obtained from simple convex functions by operations that preserve convexity

Operations that preserve convexity

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)

nonnegative weighted sum of convex functions:

$$f = w_1 f_1 + \cdots + w_m f_m$$

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

Operations that preserve convexity

Examples:

- log barrier for linear inequalities

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

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

Operations that preserve convexity

Pointwise maximum

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

examples

the maximum of all possible sums of r different components of x.

piecewise-linear function

$$f(x) = \max\{a_1^T x + b_1, \dots, a_L^T x + b_L\}$$

sum of r largest components of $x \in \mathbf{R}^n$ $\quad x_{[1]} \geq x_{[2]} \geq \cdots \geq x_{[n]} \quad f(x) = \sum_{i=1}^{r} x_{[i]},$

$$f(x) = \sum_{i=1}^{r} x_{[i]} = \max\{x_{i_1} + \cdots + x_{i_r} \mid 1 \leq i_1 < i_2 < \cdots < i_r \leq n\},$$

Operations that preserve convexity

If for each $y \in \mathcal{A}$, $f(x, y)$ is convex in x, then the function g, defined as

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

is convex in x.

examples

distance to farthest point in a set C

The distance (in any norm) to the farthest point of C $f(x) = \sup_{y \in C} \|x - y\|,$ is convex.

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

$$f(X) = \sup\{y^T X y \mid \|y\|_2 = 1\} = \sup_{\|y\|_2 = 1} y^T X y$$

linear functions of X

Operations that preserve convexity

Composition with scalar functions

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

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

$$\boxed{f \text{ is convex if } \begin{array}{l} g \text{ convex, } h \text{ convex, } \tilde{h} \text{ nondecreasing} \\ g \text{ concave, } h \text{ convex, } \tilde{h} \text{ nonincreasing} \end{array}}$$

- proof (for $n = 1$, differentiable g, h)

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

Operations that preserve convexity

$$f \text{ is convex if } \begin{array}{l} g \text{ convex, } h \text{ convex, } \tilde{h} \text{ nondecreasing} \\ g \text{ concave, } h \text{ convex, } \tilde{h} \text{ nonincreasing} \end{array}$$

examples

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

Operations that preserve convexity

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

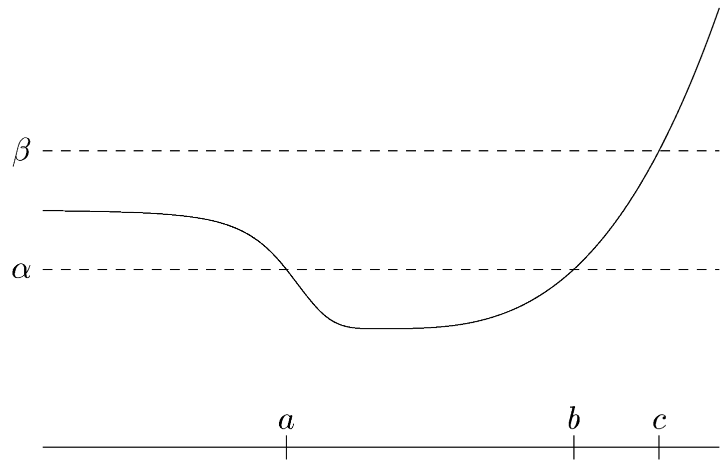
Quasiconvex functions

A function $f : \mathbf{R}^n \to \mathbf{R}$ is called *quasiconvex* (or *unimodal*) if its domain and all its sublevel sets

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

for $\alpha \in \mathbf{R}$, are convex.

- f is quasiconcave if $-f$ is quasiconvex

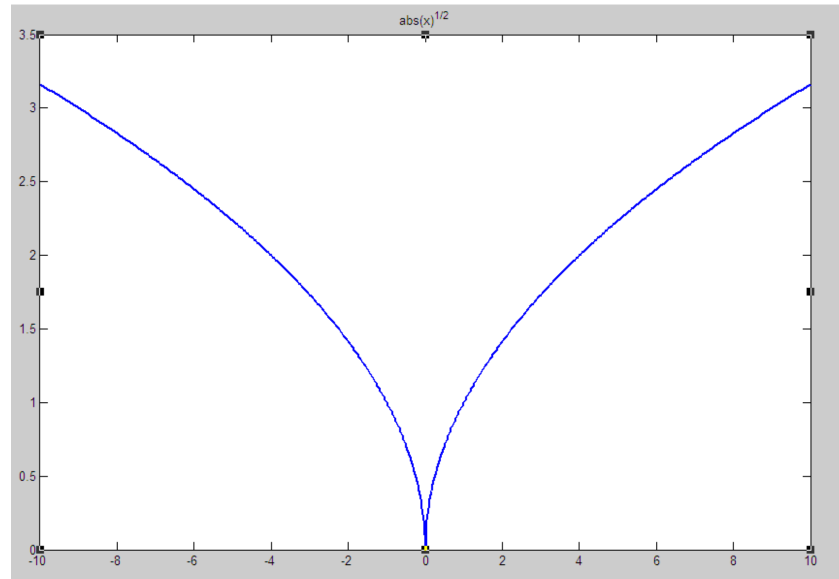


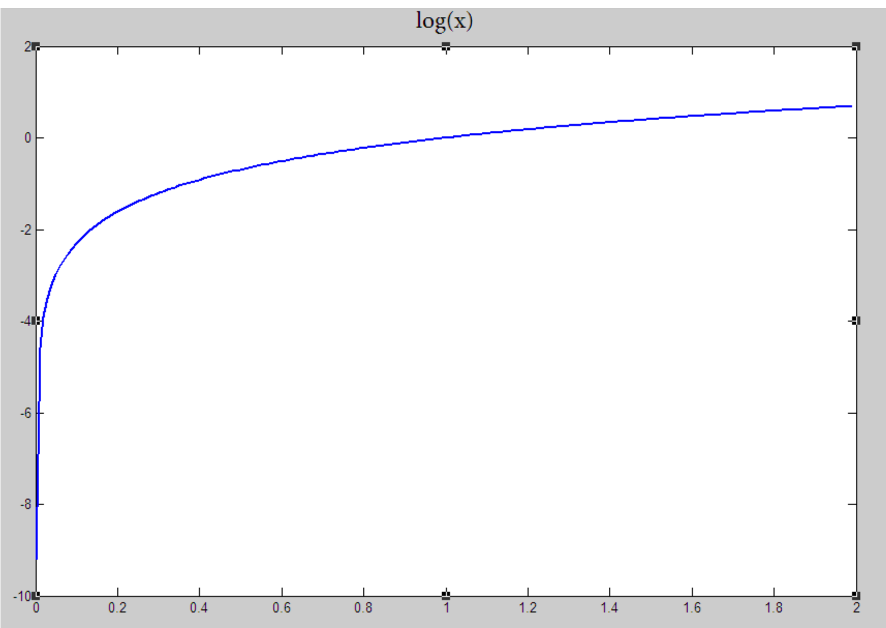
- f is quasilinear if it is quasiconvex and quasiconcave

Quasiconvex functions

examples

- $\sqrt{|x|}$ is quasiconvex on $\mathbf{R}$
- $\log x$ is quasilinear on $\mathbf{R}_{++}$





Quasiconvex functions

Consider $f : \mathbf{R}^2 \to \mathbf{R}$, with $\mathbf{dom}\, f = \mathbf{R}^2_+$ and $f(x_1, x_2) = x_1 x_2$.

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

This function is neither convex nor concave

the superlevel sets

$$\{x \in \mathbf{R}^2_+ \mid x_1 x_2 \geq \alpha\}$$

are convex sets

The function f is quasiconcave