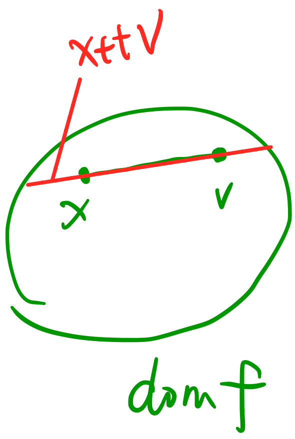
$f: \mathbb{R}^n \to \mathbb{R}$ is convex if and only if

$g: \mathbb{R} \to \mathbb{R}$, defined by



$g(t) = f(x + tv)$,

$\operatorname{dom} g = \{t : x + tv \in \operatorname{dom} f\}$,

is convex in t for $x \in \operatorname{dom} f$; $v \in \mathbb{R}^n$.

(restricted to a line).

We can use this property to check the convexity of a function by resticting it to a line. (with only one scalar variable.).

Ex. $f: S^n \to \mathbb{R}$. $f(X) = \log\det X$.

$\text{dom} f = S^n_{++}$

V has a spectral decomposition

$\Uparrow$

$g(t) = \log\det(Z + tV)$. $Z \in S^n_{++}$, $\boxed{V \in S^n}$

$\text{dom}\, g = \{t : \underline{Z + tV > 0}\}$.

$g(t) = \log\det\left[Z^{1/2}\left(\underbrace{\mathbb{1} + tZ^{-1/2}VZ^{-1/2}}_{>0}\right)Z^{1/2}\right]$

$= \log\det Z + \log\det(\mathbb{1} + tZ^{-1/2}VZ^{-1/2})$

$= \log\det Z + \sum_{i=1}^{n} \log\underbrace{(1 + t\lambda_i)}_{>0}$

where λ_i are the eigenvalues of $Z^{-1/2}VZ^{-1/2}$.

$\log\det AB$
$= \log\det A \cdot \det B$
$= \log\det A + \log\det B$
$A, B > 0$.

If $P > 0$.
$\det P = \prod_i \lambda_i$
where λ_i are eigenvalues of P

Therefore, $f(X)$ is concave.

Extended value extension

If $f: \mathbb{R}^n \to \mathbb{R}$ is convex, its extended-value extension $\tilde{f}: \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is defined by

$$\tilde{f}(x) = \begin{cases} f(x), & \text{if } x \in \mathbf{dom} f; \\ \infty, & \text{otherwise.} \end{cases}$$

Sometimes it can simplify notation without describing the domain.

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

for $0 \le \theta \le 1$, $\forall x, y \in \mathbb{R}^n$.

1. $\mathbf{dom} f$ is convex.
   If $x, y \in \mathbf{dom} f$, the same inequality as previous.
2. If x or $y \notin \mathbf{dom} f$, the right hand side of the inequality is ∞.
   we don't know $\tilde{f}(\theta x + (1-\theta) y)$

△ In this book (from now on) we will simply use $f = \tilde{f}$ if there is no ambiguity.

$f: \mathbb{R}^n \to \mathbb{R}$

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

$$\nabla f(x) = \begin{pmatrix} \partial f/\partial x_1 \\ \partial f/\partial x_2 \\ \vdots \\ \partial f/\partial x_n \end{pmatrix}$$

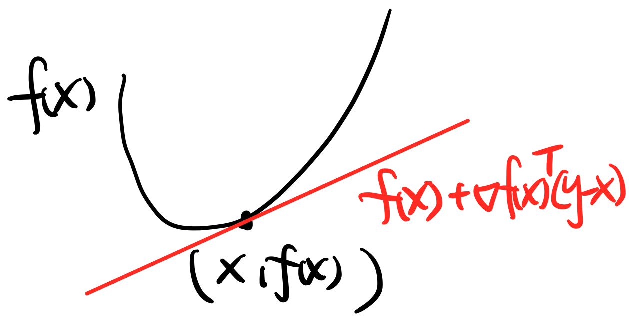
exists $\forall x \in \operatorname{dom} f$.

First-order condition

A differentiable function f is (strictly) convex (concave) if and only if $\operatorname{dom} f$ is convex, and

$$f(y) \underset{(\le)}{\overset{(>)}{\ge}} f(x) + \nabla f(x)^T (y-x)$$

$\forall x, y \in \operatorname{dom} f$.



$f(x) + \nabla f(x)^T (y-x)$ is the first order approximation of the convex function f (Appendix A). and is also a global <u>underestimator</u>.

Proof. $\boxed{\text{The case } n=1.}$ $f: \mathbb{R} \to \mathbb{R}$ is convex if and only if $f(y) \geq f(x) + f'(x)(y-x)$.

(Necessity). Assume that f is convex

$$0 < t \leq 1, \ x, y \in \operatorname{dom} f \Rightarrow \underbrace{(1-t)x + ty}_{= x + t(y-x)} \in \operatorname{dom} f$$

(convex set)

By assumption

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

$$\Rightarrow f(y) \geq f(x) + \frac{f(x+t(y-x)) - f(x)}{t}$$

$$= f(x) + \boxed{\frac{f(x+t(y-x)) - f(x)}{t(y-x)}} (y-x)$$

(Take $t \to 0$) $\to f(x) + f'(x)(y-x)$

$$\boxed{f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}}$$

$\therefore f(y) \geq f(x) + f'(x)(y-x)$

(Sufficiency)

Choose $x \neq y \in \operatorname{dom} f$. $0 \leq \theta \leq 1$

Let $z = \theta x + (1-\theta) y \in \operatorname{dom} f$.

$$\begin{cases} f(x) \geq f(z) + f'(z)(x-z) \\ f(y) \geq f(z) + f'(z)(y-z) \end{cases}$$

by assumption.

$\times\theta$, $\times(1-\theta)$

$$\theta f(x) + (1-\theta) f(y) \geq f(z) + f'(z)\underbrace{(\theta x + (1-\theta) y - z)}_{=0}$$

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

$\therefore f$ is convex

The general case $f: \mathbb{R}^n \to \mathbb{R}$

Let $x, y \in \mathbb{R}^n$ Consider f restricted to a line.

$ty + (1-t)x$.

$$\frac{\partial \left(ty + (1-t)x\right)}{\partial t} = y - x.$$

$$g(t) = f\left(ty + (1-t)x\right)$$

$$\Rightarrow g'(t) = \underline{\nabla f\left(ty + (1-t)x\right)^T}(y-x) \quad \text{by } \underline{\text{chain rule}}.$$

(Necessity) Assume that f is convex. $\Rightarrow$ g is convex (restricted to a line)

By the argument of the case $n=1$,

$$\underset{\substack{\| \\ f(y)}}{g(1)} \geq \underset{\substack{\| \\ f(x)}}{g(0)} + \underset{\substack{\| \\ \nabla f(x)^T(y-x)}}{g'(0)}(1-0)$$

$$\boxed{\begin{aligned} y &= 1 \\ x &= 0. \end{aligned}}$$

(Sufficiency)

Consider $v = ty + (1-t)x \in \operatorname{dom} f$.

$w = t'y + (1-t')x \in \operatorname{dom} f$.

By assumption

$$f(v) \geq f(w) + \nabla f(w)^T(v-w).$$

$$\underset{\substack{\| \\ f(ty+(1-t)x) \\ \| \\ g(t)}}{f(v)} \geq \underset{\substack{\| \\ f(t'y+(1-t')x) + \nabla f(t'y+(1-t')x)^T(y-x)(t-t') \\ g(t') + g'(t')(t-t')}}{f(w) + \nabla f(w)^T(v-w)}$$

Thus g is convex

$\Rightarrow f$ is convex.

f is twice differentiable if $\mathrm{dom} f$ is open and its Hessian or second derivative $\nabla^2 f(x) \in S^n$ exists, where $[\nabla^2 f(x)]_{ij} = \dfrac{\partial^2 f(x)}{\partial x_i \partial x_j}$, $i, j = 1, \dots, n$.

Second order conditions

A twice-differentiable function f is convex (concave) if and only if $\mathrm{dom} f$ is convex and

$$\nabla^2 f(x) \succeq 0 \;(\preceq 0) \quad \forall x \in \mathrm{dom} f.$$

(See Exercise 3.8 of BV).

If $\nabla^2 f(x) \succ 0$ $\forall x \in \mathrm{dom} f$, then f is strictly convex.

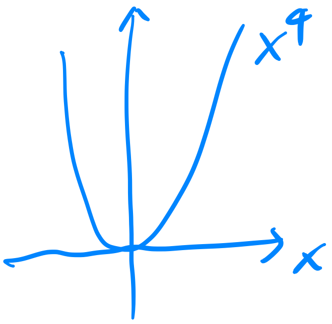
(The converse is not true.)

Ex. $f(x) = x^4$

$f'(x) = 4x^3$

$f''(x) = 12x^2 \geq 0.$

$\Rightarrow f$ is convex



Ex. The quadratic function $f: \mathbb{R}^n \to \mathbb{R}$.

$f(x) = \frac{1}{2} x^T P x + q^T x + r$, $P \in S^n$, $q \in \mathbb{R}^n$, $r \in \mathbb{R}$.

$\nabla^2 f(x) = P$ $\forall x$. f is convex if and only if $P \succeq 0$.

Ex. Least-square objective

$f(x) = \|Ax-b\|_2^2$ is convex for any $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$,

$f(x) = (Ax-b)^T(Ax-b)$

$= x^T A^T A x - x^T A^T b - b^T A x + \|b\|_2^2$

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

$\nabla^2 f(x) = \underline{2A^T A \succeq 0}$

$\left(x^T \nabla^2 f(x)\, x = 2x^T A^T A x = 2\|Ax\|_2^2 \geq 0 \quad \forall x \in \mathbb{R}^n \right)$

Ex: quardatic-over-line.

$f(x,y) = \frac{x^2}{y}$ is convex for $y > 0$.

$\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial^2 f}{\partial x \partial y}, \frac{\partial^2 f}{\partial x^2} \ldots$

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

$\underline{vv^T \succeq 0}$

WolframAlpha computational intelligence.

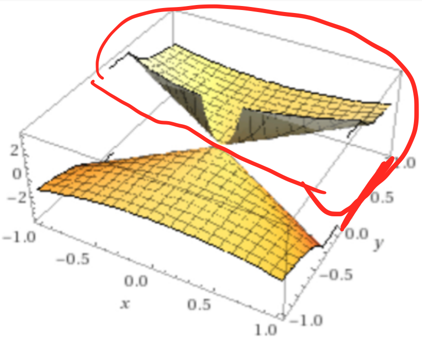
x^2/y

Extended Keyboard Upload Examples Random

Input:

$$\frac{x^2}{y}$$

3D plot:



$y > 0$

Show contour lines

Root:

$y \neq 0, \quad x = 0$

Partial derivatives:

$$\frac{\partial}{\partial x}\left(\frac{x^2}{y}\right) = \frac{2x}{y}$$

$$\frac{\partial}{\partial y}\left(\frac{x^2}{y}\right) = -\frac{x^2}{y^2}$$

Step-by-step solution

Ex. $f(X) = \log\det X$. $\operatorname{dom} f = S_{++}^n$.
(revisited)

For $Z \succ 0$, $Z + tV \succ 0$.

$$g(t) = f(Z + tV)$$
$$= \log\det Z + \sum_i \log(1 + t\lambda_i)$$

where λ_i are the eigenvalues of $Z^{-1/2} V Z^{-1/2}$

$$g'(t) = \sum_{i=1}^n \frac{\lambda_i}{1 + t\lambda_i}$$
$$g''(t) = -\sum_{i=1}^n \frac{\lambda_i^2}{(1+t\lambda_i)^2} \le 0$$

$\therefore f$ is concave.

For $\alpha \in \mathbb{R}$, the α-sublevel set of a function $f: \mathbb{R}^n \to \mathbb{R}$ is defined as

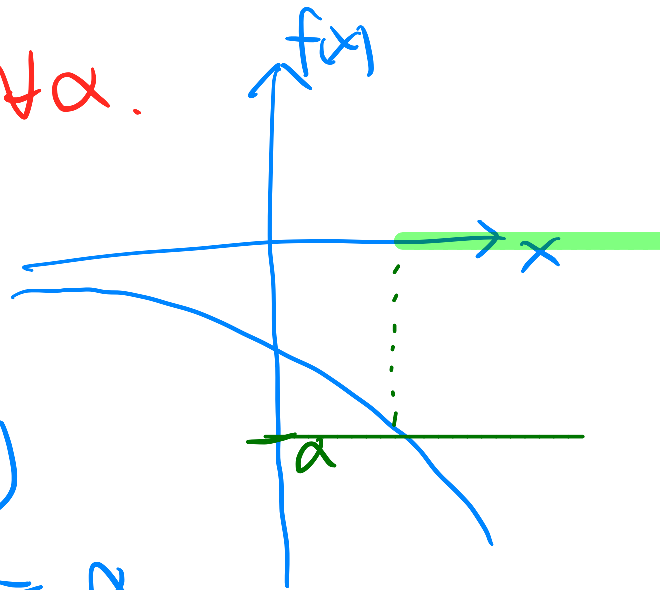
$$C_\alpha = \{x \in \mathrm{dom} f : f(x) \le \alpha\}.$$

▷ Sublevel sets of convex functions are convex. (verify it).

(But the converse is not true.).

Ex: $f(x) = -e^x$, which is concave.

C_α is convex $\forall \alpha$.



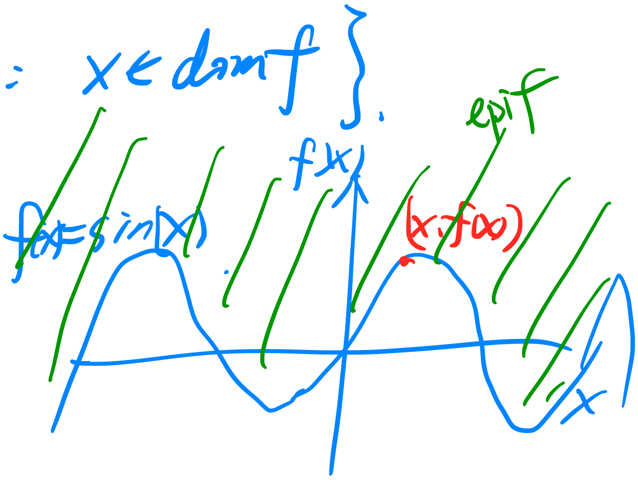
If $-e^x \le \alpha$,

$-e^y \le \alpha$,

$$\Rightarrow -e^{\theta x + (1-\theta) y} \le (-e^{\theta x})(-e^{(1-\theta) y})$$

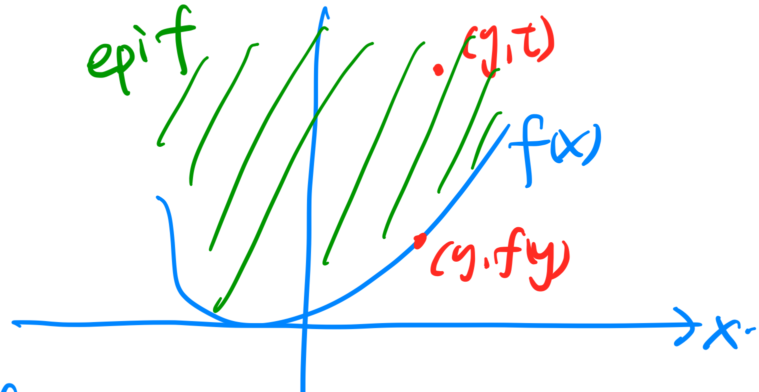
$$\le \alpha^\theta \cdot \alpha^{1-\theta} = \alpha.$$

The graph of a function $f: \mathbb{R}^n \to \mathbb{R}$ is defined as $\{(x, f(x)) \in \mathbb{R}^{n+1} : x \in \operatorname{dom} f\}$.



△ The epigraph (upper) of a function $f: \mathbb{R}^n \to \mathbb{R}$ is defined as

$$\operatorname{epi} f = \{(x, t) : x \in \operatorname{dom} f,\ f(x) \le t\}$$



△ f is convex if and only if $\operatorname{epi} f$ is a convex set.

(Necessity)

(Sufficiency)

Assume $\operatorname{epi} f$ is convex

$$(x, t), (y, s) \in \operatorname{epi} f$$

$$\Rightarrow \theta(x, t) + (1-\theta)(y, s) \in \operatorname{epi} f$$

$$\| $$

$$(\theta x + (1-\theta) y,\ \theta t + (1-\theta) s)$$