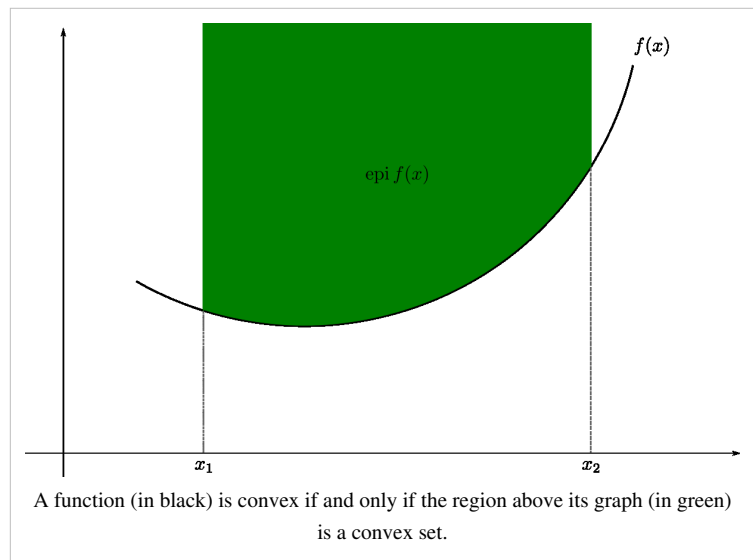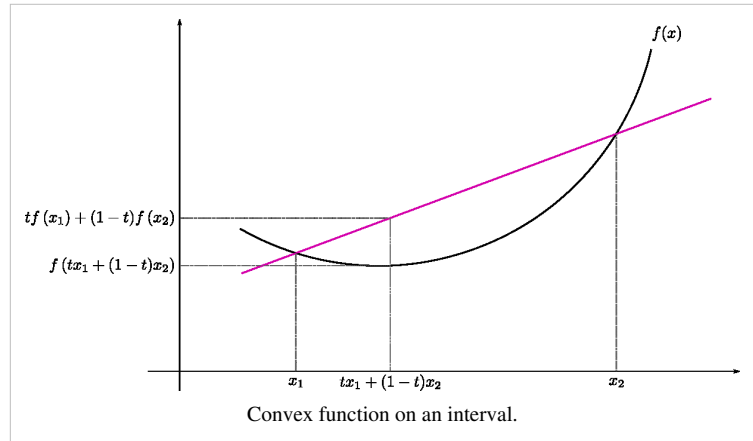# Convex function

In mathematics, a real-valued function $f(x)$ defined on an interval is called **convex** (or **convex downward** or **concave upward**) if the line segment between any two points on the graph of the function lies above the graph, in a Euclidean space (or more generally a vector space) of at least two dimensions. Equivalently, a function is convex if its epigraph (the set of points on or above the graph of the function) is a convex set. Well-known examples of convex functions are the quadratic function $f(x) = x^2$ and the exponential function $f(x) = e^x$ for any real number x.

Convex functions play an important role in many areas of mathematics. They are especially important in the study of optimization problems where they are distinguished by a number of convenient properties. For instance, a (strictly) convex function on an open set has no more than one minimum. Even in infinite-dimensional spaces, under suitable additional hypotheses, convex functions continue to satisfy such properties and, as a result, they are the most well-understood functionals in the calculus of variations. In probability theory, a convex function applied to the expected value of a random variable is always less or equal to the expected value of the convex function of the random variable. This result, known as Jensen's inequality underlies many important inequalities (including, for instance, the arithmetic-geometric mean inequality and Hölder's inequality).

Exponential growth is a special case of convexity. Exponential growth narrowly means "increasing at a rate *proportional* to the current value", while convex growth generally means "increasing at an increasing rate (but not necessarily proportionally to current value)".

Convex function on an interval.

A function (in black) is convex if and only if the region above its graph (in green) is a convex set.

Definition

Let X be a convex set in a real vector space and let $f : X \to \mathbb{R}$ be a function.

- f is called **convex** if:

$$\forall x_1, x_2 \in X, \forall t \in [0, 1] : \qquad f(tx_1 + (1 - t)x_2) \leq tf(x_1) + (1 - t)f(x_2).$$

- f is called **strictly convex** if:

$$\forall x_1 \neq x_2 \in X, \forall t \in (0, 1) : \qquad f(tx_1 + (1 - t)x_2) < tf(x_1) + (1 - t)f(x_2).$$

- A function f is said to be (strictly) concave if $-f$ is (strictly) convex.

Properties

Suppose f is a function of one real variable defined on an interval, and let

$$R(x_1, x_2) = \frac{f(x_1) - f(x_2)}{x_1 - x_2}$$

(note that $R(x_1, x_2)$ is the slope of the purple line in the above drawing; note also that the function R is symmetric in x_1, x_2). f is convex if and only if $R(x_1, x_2)$ is monotonically non-decreasing in x_1, for x_2 fixed (or viceversa). This characterization of convexity is quite useful to prove the following results.

A convex function f defined on some open interval C is continuous on C and Lipschitz continuous on any closed subinterval. f admits left and right derivatives, and these are monotonically non-decreasing. As a consequence, f is differentiable at all but at most countably many points. If C is closed, then f may fail to be continuous at the endpoints of C (an example is shown in the examples' section).

A function is midpoint convex on an interval C if

$$f\left(\frac{x_1 + x_2}{2}\right) \leq \frac{f(x_1) + f(x_2)}{2}$$

for all x_1 and x_2 in C. This condition is only slightly weaker than convexity. For example, a real valued Lebesgue measurable function that is midpoint convex will be convex by Sierpinski Theorem. In particular, a continuous function that is midpoint convex will be convex.

A differentiable function of one variable is convex on an interval if and only if its derivative is monotonically non-decreasing on that interval. If a function is differentiable and convex then it is also continuously differentiable. For the basic case of a differentiable function from (a subset of) the real numbers to the real numbers, "convex" is equivalent to "increasing at an increasing rate".

A continuously differentiable function of one variable is convex on an interval if and only if the function lies above all of its tangents:[69]

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

for all x and y in the interval. In particular, if $f'(c) = 0$, then c is a global minimum of $f(x)$.

A twice differentiable function of one variable is convex on an interval if and only if its second derivative is non-negative there; this gives a practical test for convexity. Visually, a twice differentiable convex function "curves up", without any bends the other way (inflection points). If its second derivative is positive at all points then the function is strictly convex, but the converse does not hold. For example, the second derivative of $f(x) = x^4$ is $f''(x) = 12 x^2$, which is zero for $x = 0$, but x^4 is strictly convex.

More generally, a continuous, twice differentiable function of several variables is convex on a convex set if and only if its Hessian matrix is positive semidefinite on the interior of the convex set.

Any local minimum of a convex function is also a global minimum. A *strictly* convex function will have at most one global minimum.

For a convex function *f*, the sublevel sets {*x* | *f*(*x*) < *a*} and {*x* | *f*(*x*) ≤ *a*} with *a* ∈ **R** are convex sets. However, a function whose sublevel sets are convex sets may fail to be a convex function. A function whose sublevel sets are convex is called a *quasiconvex function*.

Jensen's inequality applies to every convex function *f*. If *X* is a random variable taking values in the domain of *f*, then

$$\mathrm{E}(f(X)) \geq f(\mathrm{E}(X)).$$ (Here E denotes the mathematical expectation.)

If a function *f* is convex, and *f*(0) ≤ 0, then *f* is superadditive on the positive half-axis.

Proof: Since *f* is convex, let *y* = 0, then:

$$f(tx) = f(tx + (1-t) \cdot 0) \leq tf(x) + (1-t)f(0) \leq tf(x), \quad \forall t \in [0,1].$$

From this we have:

$$f(a)+f(b) = f\left((a+b)\frac{a}{a+b}\right)+f\left((a+b)\frac{b}{a+b}\right) \leq \frac{a}{a+b}f(a+b)+\frac{b}{a+b}f(a+b) = f(a+b)$$

Convex function calculus

- If *f* and *g* are convex functions, then so are $m(x) = \max\{f(x), g(x)\}$ and $h(x) = f(x) + g(x)$.
- If *f* and *g* are convex functions and *g* is non-decreasing, then $h(x) = g(f(x))$ is convex. As an example, if *f(x)* is convex, then so is $e^{f(x)}$, because e^x is convex and monotonically increasing.
- If *f* is concave and *g* is convex and non-increasing, then $h(x) = g(f(x))$ is convex.
- Convexity is invariant under affine maps: that is, if $f(x)$ is convex with $x \in \mathbf{R}^n$, then so is $g(x) = f(Ax + b)$, where $A \in \mathbf{R}^{n \times n}, b \in \mathbf{R}^n$.
- If $f(x, y)$ is convex in *x* then $g(x) = \sup_{y \in C} f(x, y)$ is convex in *x* provided $g(x) > -\infty$ for some *x*, even if *C* is not convex.
- If $f(x, y)$ is convex in *(x,y)* and *C* is a nonempty convex set then $g(x) = \inf_{y \in C} f(x, y)$ is convex in *x* provided $g(x) > -\infty$ for some *x*.
- If *f(x)* is convex, then its perspective $g(x, t) = tf(x/t)$ (whose domain is $\left\{(x,t) | \frac{x}{t} \in \mathrm{Dom}(f), t > 0\right\}$) is convex.
- The additive inverse of a convex function is a concave function.

Strongly convex functions

The concept of strong convexity extends and parametrizes the notion of strict convexity. A strongly convex function is also strictly convex, but not vice-versa.

A differentiable function *f* is called strongly convex with parameter *m* > 0 if the following inequality holds for all points *x,y* in its domain:

$$(\nabla f(x) - \nabla f(y))^T (x - y) \geq m\|x - y\|_2^2$$

or, more generally,

$$\langle \nabla f(x) - \nabla f(y), (x - y) \rangle \geq m\|x - y\|^2$$

where $\| \cdot \|$ is any norm. Some authors, such as refer to functions satisfying this inequality as elliptic functions.

An equivalent condition is the following:

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) + \frac{m}{2}\|y - x\|_2^2$$

It is not necessary for a function to be differentiable in order to be strongly convex. A third definition for a strongly convex function, with parameter *m*, is that, for all *x,y* in the domain and $t \in [0,1]$,

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \frac{1}{2}mt(1-t)\|x - y\|_2^2$$

Notice that this definition approaches the definition for strict convexity as $m \to 0$, and is identical to the definition of a convex function when $m = 0$. Despite this, functions exist that are strictly convex but are not strongly convex for any $m > 0$ (see example below).

If the function f is twice continuously differentiable, then f is strongly convex with parameter m if and only if $\nabla^2 f(x) \succeq mI$ for all x in the domain, where I is the identity and $\nabla^2 f$ is the Hessian matrix, and the inequality $\succeq$ means that $\nabla^2 f(x) - mI$ is positive semi-definite. This is equivalent to requiring that the minimum eigenvalue of $\nabla^2 f(x)$ be at least m for all x. If the domain is just the real line, then $\nabla^2 f(x)$ is just the second derivative $f''(x)$, so the condition becomes $f''(x) \geq m$. If $m = 0$, then this means the Hessian is positive semidefinite (or if the domain is the real line, it means that $f''(x) \geq 0$), which implies the function is convex, and perhaps strictly convex, but not strongly convex.

Assuming still that the function is twice continuously differentiable, one can show that the lower bound of $\nabla^2 f(x)$ implies that it is strongly convex. Start by using Taylor's Theorem:

$$f(y) = f(x) + \nabla f(x)^T (y - x) + \frac{1}{2}(y - x)^T \nabla^2 f(z)(y - x)$$

for some (unknown) $z \in [x, y]$. Then

$$(y - x)^T \nabla^2 f(z)(y - x) \geq m(y - x)^T (y - x)$$

by the assumption about the eigenvalues, and hence we recover the second strong convexity equation above.

A function f is strongly convex with parameter m if and only if the function $x \mapsto f(x) - \frac{m}{2}\|x\|^2$ is convex.

The distinction between convex, strictly convex, and strongly convex can be subtle at first glimpse. If f is twice continuously differentiable and the domain is the real line, then we can characterize it as follows:

f convex if and only if $f''(x) \geq 0$ for all x.

f strictly convex if $f''(x) > 0$ for all x (note: this is sufficient, but not necessary).

f strongly convex if and only if $f''(x) \geq m > 0$ for all x.

For example, consider a function f that is strictly convex, and suppose there is a sequence of points (x_n) such that $f''(x_n) = \frac{1}{n}$. Even though $f''(x_n) > 0$, the function is not strongly convex because $f''(x)$ will become arbitrarily small.

Strongly convex functions are in general easier to work with than convex or strictly convex functions, since they are a smaller class. Like strictly convex functions, strongly convex functions have unique minima.

Uniformly convex functions

A uniformly convex function, with modulus ϕ, is a function f that, for all x,y in the domain and $t \in [0, 1]$, satisfies

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y) - t(1 - t)\phi(\|x - y\|)$$

where ϕ is a function that is increasing and vanishes only at 0. This is a generalization of the concept of strongly convex function; by taking $\phi(\alpha) = \frac{m}{2}\alpha^2$ we recover the definition of strong convexity.

Examples

- The function $f(x) = x^2$ has $f''(x) = 2 > 0$ at all points, so f is a convex function. It is also strongly convex (and hence strictly convex too), with strong convexity constant 2.
- The function $f(x) = x^4$ has $f''(x) = 12x^2 \geq 0$, so f is a convex function. It is strictly convex, even though the second derivative is not strictly positive at all points. It is not strongly convex.
- The absolute value function $f(x) = |x|$ is convex (as reflected in the triangle inequality), even though it does not have a derivative at the point $x = 0$. It is not strictly convex.

- The function $f(x) = |x|^p$ for $1 \le p$ is convex.
- The exponential function $f(x) = e^x$ is convex. It is also strictly convex, since $f''(x) = e^x > 0$, but it is not strongly convex since the second derivative can be arbitrarily close to zero. More generally, the function $g(x) = e^{f(x)}$ is logarithmically convex if f is a convex function. The term "superconvex" is sometimes used instead.
- The function f with domain [0,1] defined by $f(0) = f(1) = 1$, $f(x) = 0$ for $0 < x < 1$ is convex; it is continuous on the open interval (0, 1), but not continuous at 0 and 1.
- The function x^3 has second derivative $6x$; thus it is convex on the set where $x \ge 0$ and concave on the set where $x \le 0$.
- The function $-\log\det(X)$ on the domain of positive-definite matrices is convex.[74]
- Every real-valued linear transformation is convex but not strictly convex, since if f is linear, then $f(a + b) = f(a) + f(b)$. This statement also holds if we replace "convex" by "concave".
- Every real-valued affine function, i.e., each function of the form $f(x) = a^T x + b$, is simultaneously convex and concave.
- Every norm is a convex function, by the triangle inequality and positive homogeneity.
- Examples of functions that are monotonically increasing but not convex include $f(x) = \sqrt{x}$ and $g(x) = \log(x)$.
- Examples of functions that are convex but not monotonically increasing include $h(x) = x^2$ and $k(x) = -x$.
- The function $f(x) = 1/x$ has $f''(x) = \dfrac{2}{x^3}$ which is greater than 0 if $x > 0$, so $f(x)$ is convex on the interval $(0, +\infty)$. It is concave on the interval $(-\infty, 0)$.
- The function $f(x) = 1/x^2$, with $f(0) = +\infty$, is convex on the interval $(0, +\infty)$ and convex on the interval $(-\infty, 0)$, but not convex on the interval $(-\infty, +\infty)$, because of the singularity at $x = 0$.

Notes

References

- Bertsekas, Dimitri (2003). *Convex Analysis and Optimization*. Athena Scientific.
- Borwein, Jonathan, and Lewis, Adrian. (2000). Convex Analysis and Nonlinear Optimization. Springer.
- Donoghue, William F. (1969). *Distributions and Fourier Transforms*. Academic Press.
- Hiriart-Urruty, Jean-Baptiste, and Lemaréchal, Claude. (2004). Fundamentals of Convex analysis. Berlin: Springer.
- Krasnosel'skii M.A., Rutickii Ya.B. (1961). *Convex Functions and Orlicz Spaces*. Groningen: P.Noordhoff Ltd.
- Lauritzen, Niels (2013). *Undergraduate Convexity*. World Scientific Publishing.
- Luenberger, David (1984). *Linear and Nonlinear Programming*. Addison-Wesley.
- Luenberger, David (1969). *Optimization by Vector Space Methods*. Wiley & Sons.
- Rockafellar, R. T. (1970). *Convex analysis*. Princeton: Princeton University Press.
- Thomson, Brian (1994). *Symmetric Properties of Real Functions*. CRC Press.
- Zălinescu, C. (2002). *Convex analysis in general vector spaces*. River Edge, NJ: World Scientific Publishing Co., Inc. pp. xx+367. ISBN 981-238-067-1. MR 1921556 (http://www.ams.org/mathscinet-getitem?mr=1921556).

External links

- Stephen Boyd and Lieven Vandenberghe, Convex Optimization (http://www.stanford.edu/~boyd/cvxbook/) (PDF)
- Hazewinkel, Michiel, ed. (2001), "Convex function (of a real variable)" (http://www.encyclopediaofmath.org/index.php?title=p/c026240), *Encyclopedia of Mathematics*, Springer, ISBN 978-1-55608-010-4
- Hazewinkel, Michiel, ed. (2001), "Convex function (of a complex variable)" (http://www.encyclopediaofmath.org/index.php?title=p/c026230), *Encyclopedia of Mathematics*, Springer, ISBN 978-1-55608-010-4

Article Sources and Contributors

Image Sources, Licenses and Contributors

License