Convex functions

October 27, 2022

1 Basic properties and examples

1.1 Definition

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

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

A function f is strictly convex if strict inequality above whenever $x \neq y$ and $0 < \theta < 1$. We say f is concave if -f is convex, and strictly concave if -f is strictly convex.

All affine (and therefore also linear) functions are both convex and concave. Conversely, any function that is convex and concave is affine.

A function is convex if and only if it is convex when restricted to any line that intersects its domain. In other words f is convex if and only if for all $x \in \mathbf{dom} f$ and all v, the function $g(t) = f(x + tv)$ is convex (on its domain, $\{t | x + tv \in \mathbf{dom} f\}$).

A convex function is continuous on the relative interior of its domain; it can have discontinuities only on its relative boundary.

1.2 Extended-value extensions

It is often convenient to extend a convex function to all of $\mathbf{R}^n$ by defining its value to be $\boldsymbol{\infty}$ outside its domain. If f is convex we define its extended-value extension $\widetilde{f} : \mathbf{R}^n \to R \cup \{\infty\}$ by

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

The extension $\widetilde{f}$ is defined on all $\mathbf{R}^n$, and takes values in $\mathbf{R} \cup \{\infty\}$. We can recover the domain of the original function f from the extension $\widetilde{f}$ as $\mathbf{dom} f = \{x | \widetilde{f} < \infty\}$.

1.3 First-order conditions

Suppose f is differentiable (i.e., its gradient ∇f exists at each point in $\mathbf{dom} f$, which is open). Then f is convex if and only if $\mathbf{dom} f$ is convex and

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

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

The affine function of y given by $f(x) + \nabla f(x)^T(y - x)$ is, of course, the first-order Taylor approximation of f near x. The inequality states that for a convex function, the first-order Taylor approximation is in fact a global underestimator of the function. Conversely, if the first-order Taylor approximation of a function is always a global underestimator of the functio

the inequality shows that 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 of the function f.

Strict convexity can also be characterized by a first-order condition: f is strictly convex if and only if domf is convex and for $x, y \in \mathbf{dom} f$, $x \neq y$, we have

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

For concave functions we have the corresponding characterization: f is concave if and only if domf is convex and

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

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

1.4 Second-order conditions

We now assume that f is twice differentiable, that is, its Hessian or second derivative $\nabla^2 f$ exists at each point in $\mathbf{dom} f$, which is open. Then f is convex if and only if $\mathbf{dom} f$ is convex and its Hessian is positive semidefinite: for all $x \in \mathbf{dom} f$,

$$\nabla^2 f \succeq 0$$

For a function on $\mathbf{R}$, this reduces to the simple condition $f''(x) \geq 0$ (and $\mathbf{dom} f$ convex, i.e., an interval), which means that the derivative is nondecreasing. The condition $\nabla^2 f \succeq 0$ can be interpreted geometrically as the requirement that the graph of the function have positive (upward) curvature at x.

Similarly, f is concave if and only if domf is convex and $\nabla^2 f \preceq 0$ for all $x \in \mathbf{dom} f$. Strict convexity can be partially characterized by second-order conditions. If $\nabla^2 f \succ 0$ for all $x \in \mathbf{dom} f$, then f is strictly convex. The converse, however, is not true: for example, the function $f : \mathbf{R} \to \mathbf{R}$ given by $f(x) = x^4$ is strictly convex but has zero second derivative at $x = 0$.

1.5 Examples

We start with some functions on $\mathbf{R}$, with variable x.

1. Exponential. e^{ax} is convex on $\mathbf{R}$, for any $a \in \mathbf{R}$.

2. Powers. x^a is convex on $\mathbf{R}_{++}$ when $a \geq 1$ or $a \leq 0$, and concave for $0 \leq a \leq 1$.

3. Powers of absolute value. $|x|^p$, for $p \geq 1$, is convex on textbfR.

4. Logarithm. log x is concave on $\mathbf{R}_{++}$.

5. Negative entropy. x log x (either on $\mathbf{R}_{++}$, or on $\mathbf{R}_+$, defined as 0 for $x = 0$) is convex.

We now give a few interesting examples of functions on $\mathbf{R}^n$.

1. Norms. Every norm on $\mathbf{R}^n$ is convex.
2. Max function. $f(x) = max\{x_1, \cdots, x_n\}$ is convex on $\mathbf{R}^n$.
3. Quadratic-over-linear function. The function $f(x, y) = x^2/y$, with

$$\mathbf{dom}f = \mathbf{R} \times \mathbf{R}_{++} = \{(x, y) \in \mathbf{R}^2 | y > 0\}$$

   is convex
4. Log-sum-exp. The function $f(x) = log(e^{x_1} + \cdots + e^{x_n})$ is convex on $\mathbf{R}^n$. This function can be interpreted as a differentiable (in fact, analytic) approximation of the max function, since

$$max\{x_1, \cdots, x_n\} \leq f(x) \leq max\{x_1, \cdots, x_n\} + logn$$

   for all x.
5. Geometric mean. The geometric mean $f(x) = (\Pi_{i=1}^n x_i)^{1/n}$ is concave on $\mathbf{dom}f = \mathbf{R}_{++}^n$
6. Log-determinant. The function $f(X) = log(detX)$ is concave on $\mathbf{dom}f = \mathbf{S}_{++}^n$.

1.6 Sublevel sets

The $\alpha - sublevel$ set of a function $f : \mathbf{R}^n \to \mathbf{R}$ is defined as

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

Sublevel sets of a convex function are convex, for any value of α.

The converse is not true: a function can have all its sublevel sets convex, but not be a convex function. For example, $f(x) = -e^x$ is not convex on R (indeed, it is strictly concave) but all its sublevel sets are convex.

If f is concave, then its $\alpha - superlevel$ set, given by $\{x \in \mathbf{dom}f | f(x) \geq \alpha\}$, is a convex set.

1.7 Epigraph

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

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

which is a subset of $\mathbf{R}^{n+1}$. The epigraph of a function $f : \mathbf{R}^n \to \mathbf{R}$ is defined as

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

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

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. A function is concave if and only if its hypograph, defined as

$$\mathbf{hypo}f = \{(x, t) | t \leq f(x)\}$$

is a convex set.

1.8 Jensen's inequality and extensions

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

is sometimes called Jensen's inequality. It is easily extended to convex combinations of more than two points: If f is convex, $x_1, \cdots, x_k \in \mathbf{dom} f$, and $\theta_1, \cdots, \theta_k \geq 0$ with $\theta_1 + \cdots + \theta_k = 1$, then

$$f(\theta_1 x_1 + \cdots + \theta_k x_k) \leq \theta_1 f(x_1) + \cdots + \theta_k f(x_k)$$

As in the case of convex sets, the inequality extends to infinite sums, integrals, and expected values. For example, if $p(x) \geq 0$ on $S \subseteq \mathbf{dom} f, \int_S p(x)dx = 1$, then

$$f(\int_S p(x)xdx) \leq \int_S f(x)p(x)dx$$

provided the integrals exist. In the most general case we can take any probability measure with support in $\mathbf{dom} f$. If x is a random variable such that $x \in \mathbf{dom} f$ with probability one, and f is convex, then we have

$$f(\mathbf{E}x) \leq \mathbf{E} f(x)$$

1.9 Inequalities

$$\sqrt{ab} \leq \frac{a+b}{2}$$

$$-log(\frac{a+b}{2}) \leq \frac{-loga - logb}{2}$$

$$\sum_{i=1}^n x_i y_i \leq (\sum_{i=1}^n |x_i|^p)^{1/p} (\sum_{i=1}^n |y_i|^q)^{1/q} \qquad p > 1, \frac{1}{p} + \frac{1}{q} = 1$$

2 Operations that preserve convexity

2.1 Nonnegative weighted sums

Combining nonnegative scaling and addition, we see that the set of convex functions is itself a convex cone: a nonnegative weighted sum of convex functions,

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

is convex. Similarly, a nonnegative weighted sum of concave functions is concave. A nonnegative, nonzero weighted sum of strictly convex (concave) functions is strictly convex (concave).

These properties extend to infinite sums and integrals. For example if f(x, y) is convex in x for each $y \in \mathcal{A}$, and $w(y) \geq 0$ for each $y \in \mathcal{A}$, then the function g defined as

$$g(x) = \int_{\mathcal{A}} w(y) f(x, y) dy$$

is convex in x (provided the integral exists).

2.2 Composition with an affine mapping

Suppose $f : \mathbf{R}^n \to \mathbf{R}$, $A \in \mathbf{R}^{n\times m}$, and $b \in \mathbf{R}^n$. Define $g : \mathbf{R}^m \to \mathbf{R}$ by

$$g(x) = f(Ax + b)$$

,

with $\mathbf{dom} g = \{x|Ax + b \in \mathbf{dom} f\}$. Then if f is convex, so is g; if f is concave, so is g.

2.3 Pointwise maximum and supremum

If f_1 and f_2 are convex functions then their pointwise maximum f, defined by

$$f(x) = max\{f_1(x), f_2(x)\}$$

with $\mathbf{dom} f = \mathbf{dom} f_1 \cap \mathbf{dom} f_2$, is also convex.

It is easily shown that if $f_1, \cdots, f_m$ are convex, then their pointwise maximum

$$f(x) = max\{f_1(x), \cdots, f_m(x)\}$$

is also convex.

The pointwise maximum property extends to the pointwise supremum over an infinite set of convex functions. 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.

Here the domain of g is

$$\mathbf{dom} g = \{x|(x, y) \in \mathbf{dom} \text{ for all } y \in \mathcal{A}, \sup_{y\in\mathcal{A}} f(x, y) < \infty\}$$

Similarly, the pointwise infimum of a set of concave functions is a concave function.

$$\mathbf{epi} g = \cap_{y\in\mathcal{A}} \mathbf{epi} f(\cdot, y)$$

Thus, the result follows from the fact that the intersection of a family of convex sets is convex.

2.4 Composition

2.5 Minimization

2.6 Perspective of a function

3

4

5

6