Convex Sets and Functions (part II)

Prof. Dan A. Simovici

UMB

Definition

Let D be a subset of a real linear space L.
A function $f : D \longrightarrow \mathbb{R}$ is convex if for every $x, y \in D$ such that $(1-t)x + ty \in D$ for $t \in [0,1]$ we have $f((1-t)x + ty) \leqslant (1-t)f(x) + tf(y)$.
A function $g : D \longrightarrow \mathbb{R}$ is concave if $-g$ is convex at $\mathbf{x}$, that is, $g((1-t)x + ty) \geqslant (1-t)g(x) + tg(y)$ for $x, y \in D$.
If $\mathbf{x}, \mathbf{y} \in D$ implies the strict inequality

$$f((1-t)x + ty) < (1-t)f(x) + tf(y),$$

then we say that f is strictly convex.
Similarly, if

$$f((1-t)x + ty) > (1-t)f(x) + tf(y),$$

for every $x, y \in D$, then we say that f is *strictly concave*.

It is useful for the study of convex functions to extend the notion of convex function by allowing ∞ as a value. Thus, if a function f is defined on a subset S of a linear space L, $f : S \longrightarrow \mathbb{R}$, the extended-value function of f is the function $\hat{f} : L \longrightarrow \hat{\mathbb{R}}$ defined by

$$\hat{f}(x) = \begin{cases} f(x) & \text{if } x \in S, \\ \infty & \text{otherwise,} \end{cases}$$

If a function $f : S \longrightarrow \mathbb{R}$ is convex, where $S \subseteq L$ is a convex set, then its extended-value function $\hat{f}$ satisfies the inequality that defines convexity $\hat{f}((1-t)x + ty) \leqslant (1-t)\hat{f}(x) + t\hat{f}(y)$ for every $x, y \in L$ and $t \in [0,1]$, if we adopt the convention that $0 \cdot \infty = 0$.

A extended-value convex function $f : S \longrightarrow \hat{\mathbb{R}}$ is properly convex if f is not the constant function defined by $f(x) = \infty$.
The effective domain of a convex function $f : S \longrightarrow \hat{\mathbb{R}}$ is the set $\mathrm{Dom}(f) = \{x \in S \mid f(\mathbf{x}) < \infty\}$.

Example

Let $f : (0, \infty) \longrightarrow \mathbb{R}$ be defined by $f(x) = x^2$. The definition domain of f is clearly convex and we have

$$\begin{aligned} f((1-t)x_1 + tx_2) &= ((1-t)x_1 + tx_2)^2 \\ &= (1-t)^2 x_1^2 + t^2 x_2^2 + 2(1-t)tx_1x_2. \end{aligned}$$

Therefore,

$$\begin{aligned} & f((1-t)x_1 + tx_2) - (1-t)f(x_1) - tf(x_2) \\ &= (1-t)^2 x_1^2 + t^2 x_2^2 + 2(1-t)tx_1x_2 - (1-t)x_1^2 - tx_2^2 \\ &= -t(1-t)(x_1 - x_2)^2 \leqslant 0, \end{aligned}$$

which implies that f is indeed convex.

Example

The function $f : \mathbb{R} \longrightarrow \mathbb{R}$ defined by $f(x) = |a - xb|$ is convex because

$$\begin{aligned} f((1-t)x_1 + tx_2) &= |a - ((1-t)x_1 + tx_2)b| \\ &= |a(1-t) + at - ((1-t)x_1 + tx_2)b| \\ &= |(1-t)(a - x_1 b) + t(a - x_2 b)| \\ &\leqslant |(1-t)(a - x_1 b)| + |t(a - x_2 b)| \\ &= (1-t)f(x_1) + tf(x_2) \end{aligned}$$

for $t \in [0, 1]$.

Example

The function $g : \mathbb{R}^2 \longrightarrow \mathbb{R}$ given by $g(\mathbf{x}) = |a - x_1 x_2|$ is not convex, in general. Consider, for example the special case $g(\mathbf{x}) = |12 - x_1 x_2|$. We have

$$f \begin{pmatrix} 6 \\ 2 \end{pmatrix} = f \begin{pmatrix} 2 \\ 6 \end{pmatrix} = 0.$$

Note that

$$\begin{pmatrix} 4 \\ 4 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 6 \\ 2 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 2 \\ 6 \end{pmatrix} \text{ and } f \begin{pmatrix} 4 \\ 4 \end{pmatrix} = 4 > \frac{1}{2} f \begin{pmatrix} 6 \\ 2 \end{pmatrix} + \frac{1}{2} f \begin{pmatrix} 2 \\ 6 \end{pmatrix}.$$

Example

Any norm ν on a real linear space L is convex. Indeed, for $t \in [0, 1]$ we have

$$\nu(tx + (1 - t)y) \leqslant \nu(tx) + \nu((1 - t)y) = t\nu(x) + (1 - t)\nu(y)$$

for $x, y \in L$.

Example

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix. The function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ given by $f(\mathbf{x}) = \mathbf{x}'A\mathbf{x}$ is convex if and only if A is a positive semidefinite matrix. Indeed, suppose that f is convex. For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ we have

$$((1-t)\mathbf{x} + t\mathbf{y})'A((1-t)\mathbf{x} + t\mathbf{y}) \leqslant (1-t)\mathbf{x}'A\mathbf{x} + t\mathbf{y}'A\mathbf{y},$$

for $t \in (0,1)$, which amounts to

$$(t^2 - t)\mathbf{x}'A\mathbf{x} + (t^2 - t)\mathbf{y}'A\mathbf{y} + (1-t)t\mathbf{y}'A\mathbf{x} + t(1-t)\mathbf{x}'A\mathbf{y} \leqslant 0.$$

Since A is symmetric, we have $(\mathbf{y}'A\mathbf{x})' = \mathbf{x}'A\mathbf{y}$ and because both terms of the last equality are scalars we have $\mathbf{y}'A\mathbf{x} = \mathbf{x}'A\mathbf{y}$. Note that $t^2 - t \leqslant 0$ because $t \in [0,1]$. Consequently,

$$\mathbf{x}'A\mathbf{x} + \mathbf{y}'A\mathbf{y} - \mathbf{y}'A\mathbf{x} - \mathbf{x}'A\mathbf{y} \geqslant 0,$$

which amounts to $(\mathbf{x} - \mathbf{y})'A(\mathbf{x} - \mathbf{y}) \geqslant 0$, so A is positive semidefinite. The reverse implication is an exercise!

Local vs. Global Minima

Definition

Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function. The point $\mathbf{x}_0$ is a global minimum for f if $f(\mathbf{x}) \geqslant f(\mathbf{x}_0)$ for every $\mathbf{x}$.
The point $\mathbf{x}_1$ is a local minimum for f if there exists $\epsilon > 0$ such that $f(\mathbf{x}) \geqslant f(\mathbf{x}_0)$ for every $\mathbf{x} \in B[\mathbf{x}_0, \epsilon]$.

If $\mathbf{x}_1$ is a local minimum for f and $\mathbf{x}_0$ is a global minimum, we have $f(\mathbf{x}_1) \geqslant f(\mathbf{x}_0)$.

Theorem

If $\mathbf{x}_1$ *is a local minimum of a convex function* $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, *then* $\mathbf{x}_1$ *is a global minimum for* f.

Proof

Let $\mathbf{x}_0$ be a global minimum of f and let $\mathbf{x}_1$ be a local minimum. We have $f(\mathbf{x}_0) \leqslant f(\mathbf{x}_1)$. Since $\mathbf{x}_1$ is a local minimum, there exists ϵ such that if $\parallel \mathbf{x}_1 - \mathbf{x} \parallel \geqslant \epsilon$, then $f(\mathbf{x}_1) \leqslant f(\mathbf{x})$.
Let $\mathbf{z} = (1-a)\mathbf{x}_1 + a\mathbf{x}_0$, where $a \in [0,1]$. We have $\mathbf{x}_1 - \mathbf{z} = a(\mathbf{x}_1 - \mathbf{x}_0)$. By choosing a such that

$$a < \frac{\epsilon}{\parallel \mathbf{x}_1 - \mathbf{x}_0 \parallel}$$

we have $\parallel \mathbf{x}_1 - \mathbf{z} \parallel \leqslant \epsilon$, which implies $\mathbf{z} \in B[\mathbf{x}_1, \epsilon]$, so $f(\mathbf{z}) \geqslant f(\mathbf{x}_1)$ because $\mathbf{x}_1$ is a local minimum. By the convexity of f we have

$$f(\mathbf{z}) = f((1-a)\mathbf{x}_1 + a\mathbf{x}_0) \leqslant (1-a)f(\mathbf{x}_1) + af(\mathbf{x}_0) \leqslant f(\mathbf{x}_1),$$

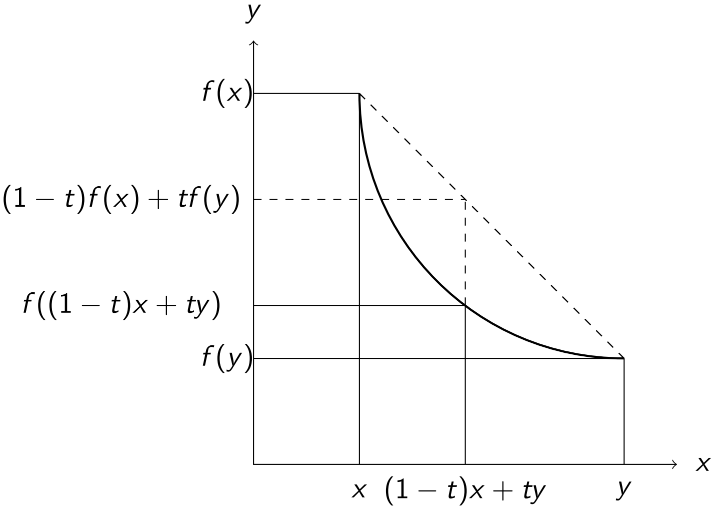
so $f(\mathbf{z}) = f(\mathbf{x}_1)$. This, in turn implies

$$f(\mathbf{x}_1) \leqslant (1-a)f(\mathbf{x}_1) + af(\mathbf{x}_0),$$

which yields $f(\mathbf{x}_1) \leqslant f(\mathbf{x}_0)$, hence $f(\mathbf{x}_1) = f(\mathbf{x}_0)$. Therefore, the local minimum $\mathbf{x}_1$ is also a global minimum.

One-argument convex function

The curve representing $f(x)$ is located under the chord determined by $(u, f(u))$ and $(v, f(v))$.



Lemma

Let $f : [a, b] \longrightarrow \mathbb{R}$ *be a convex function. If* $x \in (a, b)$, *then*

$$\frac{f(x) - f(a)}{x - a} \leqslant \frac{f(b) - f(a)}{b - a} \leqslant \frac{f(b) - f(x)}{b - x}.$$

Proof

It is easy to see that x can be regarded as either of the following convex combinations:

$$\begin{aligned} x &= \left(1-\frac{x-a}{b-a}\right)a+\frac{x-a}{b-a}b, \\ &= \frac{b-x}{b-a}a+\left(1-\frac{b-x}{b-a}\right)b. \end{aligned}$$

The existence of the first convex combination implies

$$\begin{aligned} f(x) = f\left(1-\frac{x-a}{b-a}\right)a+\frac{x-a}{b-a}b, & \\ & \leqslant \left(1-\frac{x-a}{b-a}\right)f(a)+\frac{x-a}{b-a}f(b), \end{aligned}$$

which is equivalent to

$$f(x) \leqslant \frac{b-x}{b-a}f(a)+\frac{x-a}{b-a}f(b).$$

This gives the first inequality of the lemma. The second one can be

Lemma

Let $f : I \longrightarrow \mathbb{R}$ be a function, where I is an open interval. The following statements are equivalent for $a < b < c$, where $a, b, c \in I$:

1. $(c - a)f(b) \leqslant (b - a)f(c) + (c - b)f(a)$;
2. $\frac{f(b)-f(a)}{b-a} \leqslant \frac{f(c)-f(a)}{c-a}$;
3. $\frac{f(c)-f(a)}{c-a} \leqslant \frac{f(c)-f(b)}{c-b}$.

Proof

(i) is equivalent to (ii): Suppose that (i) holds. Then we have

$$(c-a)f(b)-(c-a)f(a) \leqslant (b-a)f(c)+(c-b)f(a)-(c-a)f(a),$$

which is equivalent to

$$(c-a)(f(b)-f(a)) \leqslant (b-a)(f(c)-f(a)). \tag{1}$$

By dividing both sides by $(b-a)(c-a) > 0$ we obtain Inequality (ii).
Conversely, note that (ii) implies Inequality (1). By adding $(c-a)f(a)$ to both sides of this inequality we obtain (i).
In a similar manner it is possible to prove the equivalence between (i) and (iii).

Theorem

Let I be an open interval and let $f : I \longrightarrow \mathbb{R}$ is a function. Each of the conditions of the previous Lemma is equivalent to the convexity of f for $a < b < c$ and $a, b, c \in I$.

Proof

Let $f : I \longrightarrow \mathbb{R}$ be a convex function and let $a, b, c \in I$ be such that $a < b < c$. Define $t = \frac{c-b}{c-a}$. Clearly $0 < t < 1$ and by the convexity property,

$$\begin{aligned} f(b) &= f(at + (1-t)c) \leqslant tf(a) + (1-t)f(c) \\ &= \frac{c-b}{c-a} f(a) + \frac{b-a}{c-a} f(c), \end{aligned}$$

which yields the first inequality of Lemma 10.

Conversely, suppose that the first inequality of Lemma 10 is satisfied. Choose $a = x$, $c = y$ and $b = tx + (1-t)y$ for $t \in (0,1)$. We have $(c-a)f(b) = (y-x)f(tx + (1-t)y)$ and $(b-a)f(c) + (c-b)f(a) = (1-t)(y-x)f(y) + t(y-x)f(x)$ Taking into account that $y > x$, we obtain the inequality $f(tx + (1-t)y) \leqslant tf(x) + (1-t)f(y)$, which means that f is convex.

Theorem

Let I be an open interval and let $f : \mathbb{R} \longrightarrow \mathbb{R}$ is a convex function. The function $g(x, h)$ defined for $x \in I$ and $h \in \mathbb{R} - \{0\}$ as

$$g(x, h) = \frac{f(x + h) - f(x)}{h}$$

is increasing with respect to each of its arguments.

Proof

We need to examine three cases: $0 < h_1 < h_2$, $h_1 < h_2 < 0$, and $h_1 < 0 < h_2$.

In the first case choose $a = x$, $b = x + h_1$ and $c = x + h_2$ in the second inequality of Lemma 10, where all three numbers $x, x + h_1$ and $x + h_2$ belong to I. We obtain $\frac{f(x+h_1)-f(x)}{h_1} \leqslant \frac{f(x+h_2)-f(x)}{h_2}$, which shows that $g(x, h_1) \leqslant g(x, h_2)$.

If $h_1 < h_2 < 0$, choose $a = x + h_1$, $b = x + h_2$ and $c = x$ in the last inequality of Lemma 10. This yields: $\frac{f(x)-f(x+h_1)}{-h_1} \leqslant \frac{f(x)-f(x+h_2)}{-h_2}$, that is $g(x, h_1) \leqslant g(x, h_2)$.

In the last case, $h_1 < 0 < h_2$, begin by noting that the last two inequalities of Lemma 10 imply

$$\frac{f(b) - f(a)}{b - a} \leqslant \frac{f(c) - f(b)}{c - b}.$$

By taking $a = x + h_1$, $b = x$, and $c = x + h_2$ in this inequality we obtain

$$\frac{f(x) - f(x + h_1)}{-h_1} \leqslant \frac{f(x + h_2) - f(x)}{h_2},$$

Proof (cont'd)

To prove the monotonicity in the first argument let x_1, x_2 be in I such that $x_1 < x_2$ and let h be a number such that both $x_1 + h$ and $x_2 + h$ belong to I. Since g is monotonic in its second argument we have

$$g(x_1, h) = \frac{f(x_1 + h) - f(x_1)}{h} \leqslant \frac{f(x_2 + h) - f(x_1)}{h + (x_2 - x_1)}$$

and

$$\begin{aligned}
&\frac{f(x_2 + h) - f(x_1)}{h + (x_2 - x_1)} \\
= \quad & \frac{f(x_1) - f(x_2 + h)}{-h - (x_2 - x_1)} \\
= \quad & \frac{f((x_2 + h) - h - (x_2 - x_1)) - f(x_2 + h)}{-h - (x_2 - x_1)} \\
\leqslant \quad & \frac{f((x_2 + h) - h) - f(x_2 + h)}{-h} = \frac{f(x_2 + h) - f(x_2)}{h},
\end{aligned}$$

which proves the monotonicity in its first argument.

Convexity of functions of n arguments can be expressed as convexity of one-argument functions.

Theorem

Let $f : \mathbb{R}^n \longrightarrow \hat{\mathbb{R}}$ be a function. The function f is convex if and only if the function $\phi_{\mathbf{x},\mathbf{h}} : \mathbb{R} \longrightarrow \hat{\mathbb{R}}$ given by $\phi_{\mathbf{x},\mathbf{h}}(t) = f(\mathbf{x} + t\mathbf{h})$ is a convex function for every $\mathbf{x}$ and $\mathbf{h}$ in $\mathbb{R}^n$.

Proof

Suppose that f is convex. We have

$$\begin{aligned}
\phi_{\mathbf{x},\mathbf{h}}(ta+(1-t)b) &= f(\mathbf{x}+(ta+(1-t)b)\mathbf{h}) \\
&= f(t(\mathbf{x}+a\mathbf{h})+(1-t)(\mathbf{x}+b\mathbf{h})) \\
&\leqslant tf(\mathbf{x}+a\mathbf{h})+(1-t)f(\mathbf{x}+b\mathbf{h}) \\
&= t\phi_{\mathbf{x},\mathbf{h}}(a)+(1-t)\phi_{\mathbf{x},\mathbf{h}}(b),
\end{aligned}$$

which shows that $\phi_{\mathbf{x},\mathbf{h}}$ is indeed convex. The converse implication follows in a similar manner.

Lemma

Let $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ *be a convex function such that* $g(\mathbf{0}_n) = 0$. *We have* $-g(\mathbf{x}) \leqslant g(\mathbf{x})$ *for* $\mathbf{x} \in \mathbb{R}^n$.

Proof: Note that if $\mathbf{x} \neq \mathbf{0}_n$, $\mathbf{0}_n \in [-\mathbf{x}, \mathbf{x}]$. Since g is clearly mid-point convex, we have

$$0 = g(\mathbf{0}_n) = g\left(\frac{-\mathbf{x} + \mathbf{x}}{2}\right) \leqslant \frac{1}{2}g(-\mathbf{x}) + \frac{1}{2}g(\mathbf{x}),$$

which implies the desired inequality.

Theorem

Let $f : [a, b] \longrightarrow \mathbb{R}$ be a convex function. The function f is continuous at every $x_0 \in (a, b)$.

Proof

Let $g : (a - x_0, b - x_0) \longrightarrow \mathbb{R}$ be defined as $g(x) = f(x + x_0) - f(x_0)$. It is clear that g is convex on $(a - x_0, b - x_0)$, $0 \in (a - x_0, b - x_0)$, and $g(0) = 0$; also, g is continuous in 0 if and only if f is continuous at x_0. For $x \in (a - x_0, b - x_0)$ let

$$s(x) = \begin{cases} 1 & \text{if } x \leqslant 0, \\ -1 & \text{if } x < 0. \end{cases}$$

If $|x| < \delta$, then the convexity of g implies

$$\begin{aligned} g(x) &= g\left(\frac{|x|}{\delta} s(x)\delta + \left(1 - \frac{|x|}{\delta}\right) 0\right) \\ &\leqslant \frac{|x|}{\delta} g(s(x)\delta) + \left(1 - \frac{|x|}{\delta}\right) g(0) \\ &= \frac{|x|}{\delta} g(s(x)\delta). \end{aligned}$$

Proof (cont'd)

Therefore, $g(x) \leqslant \frac{1}{\delta}\max\{g(-\delta), g(\delta)\}|x|$. The convexity of g implies that $-g(-x) \leqslant g(x)$ by the previous lemma, so
$|g(x)| \leqslant \frac{1}{\delta}\max\{g(-\delta), g(\delta)\}|x|$.
If $\lim_{n\to snn} x_n = 0$, where (x_n) is a sequence in $(a - x_0, b - x_0)$, then $\lim_{n\to\infty} g(x_n) = 0 = g(0)$, so g is continuous in 0. This implies that f is continuous in x_0.

Definition

Let $f : I \longrightarrow \mathbb{R}$ be a convex function where I is an interval of $\mathbb{R}$ and let $a \in I$. The function f is left-differentiable if the limit

$$f'(a-) = \lim_{x \to a, x < a} \frac{f(x) - f(a)}{x - a}$$

exists. In this case, the value of the limit is known as left derivative of f in a.

Similarly, f is right-differentiable if the limit

$$f'(a+) = \lim_{x \to a, x > a} \frac{f(x) - f(a)}{x - a}$$

exists. In this case, the value of the limit is known as the right derivative of f in a.

Definition

A function $f : I \longrightarrow \mathbb{R}$ is differentiable in a, where $a \in I$, if $f'(a-) = f'(a+)$.
In this case we write $f'(a) = f'(a+) = f'(a-)$.

If f is convex and differentiable on I if f' is increasing.

Example

Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be the function defined by $f(x) = |x|$. We have:

$$f'(0-) = \lim_{x \to 0, x < 0} \frac{|x|}{x} = -1$$

and

$$f'(0+) = \lim_{x \to 0, x > 0} \frac{|x|}{x} = 1.$$

Note that if both $f'(a-)$ and $f'(a+)$ exist and are finite, then f is continuous in a.

Theorem

Let $f : (a, b) \longrightarrow \mathbb{R}$ *be convex function on* $[a, b]$. *If* $x, y \in (a, b)$ *and* $x < y$, *then*

$$f'(x-) \leqslant f'(x+) \leqslant f'(y-) \leqslant f'(y+).$$

If $a < x < b$ we have

$$\frac{f(x) - f(a)}{x - a} \leqslant \frac{f(b) - f(a)}{b - a} \leqslant \frac{f(b) - f(x)}{b - x}.$$

Since $\lim_{a \to x, a < x} \frac{f(x) - f(a)}{x - a} \leqslant \lim_{b \to x, b > x} \frac{f(b) - f(x)}{b - x}$, it follows that $f'(x-) \leqslant f'(x^{+})$; similarly, $f'(y-) \leqslant f'(y+)$.
Let $t \in (x, y)$. By the same previous result we have:

$$\frac{f(t) - f(x)}{t - x} \leqslant \frac{f(y) - f(x)}{y - x} \leqslant \frac{f(y) - f(t)}{y - t}.$$

The first inequality implies

$$f'(x+) = \lim_{t \to x, t > x} \frac{f(t) - f(x)}{t - x} \leqslant \frac{f(y) - f(x)}{y - x},$$

while the second yields

$$\frac{f(y) - f(x)}{y - x} \leqslant \lim_{t \to y, t < y} \frac{f(y) - f(t)}{y - t} = f'(y-),$$

so $f'(x+) \leqslant f'(y-)$.

Theorem

If $f : I \longrightarrow \mathbb{R}$ is a function and f' is increasing, then f is convex.

Proof

For $x, y \in I$ the inequality

$$f((1-a)x + ay) \leqslant (1-a)f(x) + af(y)$$

is immediate for $a = 0$ and $a = 1$. So suppose that $0 < a < 1$. By the Mean Value theorem we have

$$\begin{aligned} f((1-a)x + ay) &= f'(\xi_1)a(y-x), \\ f(y) - f((1-a)x + ay) &= f'(\xi_2)(1-a)(y-x), \end{aligned}$$

where $x < \xi_1 < (1-a)x + ay < xi_2 < y$.
Since $f'(xi_1) \leqslant f'(xi_2)$, the last equalities yield the convexity property (multiply first by $(1-a)$, second by a and subtract).

Corollary

If $f : I \longrightarrow \mathbb{R}$ is twice differentiable function, then f is convex if and only if $f''(x) \geqslant 0$ for every $x \in (a, b)$. If $f''(x) > 0$ for $x \in I$, then f is strictly convex.

Table: Examples of convex or concave functions.

Function	Second Derivative	Convexity Property
x^r for $r > 0$	$r(r-1)x^{r-2}$	concave for $r < 1$ convex for $r \geqslant 1$
$\ln x$	$-\frac{1}{x^2}$	concave
$x \ln x$	$\frac{1}{x}$	convex
e^x	e^x	convex

Theorem (Jensen's Theorem)

Let f be a function that is convex on an interval I. If $t_1, \ldots, t_n \in [0, 1]$ are n numbers such that $\sum_{i=1}^{n} t_i = 1$, then

$$f\left(\sum_{i=1}^{n} t_i x_i\right) \leqslant \sum_{i=1}^{n} t_i f(x_i)$$

for every $x_1, \ldots, x_n \in I$.

Proof

The argument is by induction on n, where $n \geqslant 2$. The basis step, $n = 2$, follows immediately from the definition of convex functions. Suppose that the statement holds for n, and let $u_1, \ldots, u_n, u_{n+1}$ be $n+1$ numbers such that $\sum_{i=1}^{n+1} u_i = 1$. We have

$$\begin{aligned} & f(u_1 x_1 + \cdots + u_{n-1}x_{n-1} + u_n x_n + u_{n+1}x_{n+1}) \\ = \; & f\left(u_1 x_1 + \cdots + u_{n-1}x_{n-1} + (u_n + u_{n+1}) \frac{u_n x_n + u_{n+1}x_{n+1}}{u_n + u_{n+1}} \right). \end{aligned}$$

By the inductive hypothesis, we can write

$$\begin{aligned} & f(u_1 x_1 + \cdots + u_{n-1}x_{n-1} + u_n x_n + u_{n+1}x_{n+1}) \\ \leqslant \; & u_1 f(x_1) + \cdots + u_{n-1} f(x_{n-1}) + (u_n + u_{n+1}) f\left(\frac{u_n x_n + u_{n+1}x_{n+1}}{u_n + u_{n+1}} \right). \end{aligned}$$

Next, by the convexity of f, we have

$$f\left(\frac{u_n x_n + u_{n+1}x_{n+1}}{u_n + u_{n+1}} \right) \leqslant \frac{u_n}{u_n + u_{n+1}} f(x_n) + \frac{u_{n+1}}{u_n + u_{n+1}} f(x_{n+1}).$$

Combining these inequalities gives desired conclusion.

If f is a concave function and $t_1, \ldots, t_n \in [0,1]$ are n numbers such that $\sum_{i=1}^{n} t_i = 1$, then

$$f\left(\sum_{i=1}^{n} t_i x_i\right) \geqslant \sum_{i=1}^{n} t_i f(x_i).$$

Example

We saw that the function $f(x) = \ln x$ is concave. Therefore, if $t_1, \ldots, t_n \in [0,1]$ are n numbers such that $\sum_{i=1}^{n} t_i = 1$, then

$$\ln\left(\sum_{i=1}^{n} t_i x_i\right) \geqslant \sum_{i=1}^{n} t_i \ln x_i .$$

This inequality can be written as

$$\ln\left(\sum_{i=1}^{n} t_i x_i\right) \geqslant \ln \prod_{i=1}^{n} x_i^{t_i},$$

or equivalently

$$\sum_{i=1}^{n} t_i x_i \geqslant \prod_{i=1}^{n} x_i^{t_i},$$

for $x_1, \ldots, x_n \in (0, \infty)$.

In the special case where $t_1 = \cdots = t_n = \frac{1}{n}$, we have the inequality that relates the arithmetic to the geometric average on n positive numbers:

$$\frac{x_1 + \cdots + x_n}{n} \geqslant \left(\prod_{i=1}^{n} x_i \right)^{\frac{1}{n}}. \tag{2}$$

Weighted Means

Let $\mathbf{w} = (w_1, \ldots, w_n) \in \mathbb{R}^n$ be such that $\sum_{i=1}^n w_i = 1$. For $r \neq 0$, the **w**-*weighted mean of order r* of a sequence of n positive numbers $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n_{>0}$ is the number

$$\mu^r_{\mathbf{w}}(\mathbf{x}) = \left(\sum_{i=1}^n w_i x_i^r \right)^{\frac{1}{r}}.$$

Of course, $\mu^r_{\mathbf{w}}(\mathbf{x})$ is not defined for $r = 0$; we will give as special definition

$$\mu^0_{\mathbf{w}}(\mathbf{x}) = \lim_{r \to 0} \mu^r_{\mathbf{w}}(\mathbf{x}).$$

We have

$$\begin{aligned} \lim_{r \to 0} \ln \mu^r_{\mathbf{w}}(\mathbf{x}) &= \lim_{r \to 0} \frac{\ln \sum_{i=1}^n w_i x_i^r}{r} = \lim_{r \to 0} \frac{\sum_{i=1}^n w_i x_i^r \ln x_i}{\sum_{i=1}^n w_i x_i^r} \\ &= \sum_{i=1}^n w_i \ln x_i = \ln \prod_{i=1}^n x_i^{w_i}. \end{aligned}$$

Thus, if we define $\mu^0_{\mathbf{w}}(\mathbf{x}) = \prod_{i=1}^n x_i^{w_i}$, the weighted mean of order r

For $w_1 = \cdots = w_n = \frac{1}{n}$, we have

$$\mu_{\mathbf{w}}^{-1}(\mathbf{x}) = \frac{n x_1 \cdots x_n}{x_2 \cdots x_n + \cdots + x_1 \cdots x_{n-1}}$$

(the harmonic average of $\mathbf{x}$),

$$\mu_{\mathbf{w}}^{0}(\mathbf{x}) = (x_1 \ldots x_n)^{\frac{1}{n}}$$

(the geometric average of $\mathbf{x}$),

$$\mu_{\mathbf{w}}^{1}(\mathbf{x}) = \frac{x_1 + \cdots + x_n}{n}$$

(the arithmetic average of $\mathbf{x}$).

Theorem

If $p < r$, we have $\mu_{\mathbf{w}}^{p}(\mathbf{x}) \leqslant \mu_{\mathbf{w}}^{r}(\mathbf{x})$.

Proof

There are three cases depending on the position of 0 relative to p and r. In the first case, suppose that $r > p > 0$. The function $f(x) = x^{\frac{r}{p}}$ is convex, so by Jensen's inequality applied to $x_1^p, \ldots, x_n^p$, we have

$$\left(\sum_{i=1}^{n} w_i x_i^p\right)^{\frac{r}{p}} \leqslant \sum_{i=1}^{n} w_i x_i^r,$$

which implies

$$\left(\sum_{i=1}^{n} w_i x_i^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{i=1}^{n} w_i x_i^r\right)^{\frac{1}{r}},$$

which is the inequality of the theorem.

If $r > 0 > p$, the function $f(x) = x^{\frac{r}{p}}$ is again convex because $f''(x) = \frac{r}{p}\left(\frac{r}{p} - 1\right) x^{\frac{r}{p}-2} \geq 0$. Thus, the same argument works as in the previous case.

Proof (cont'd)

Finally, suppose that $0 > r > p$. Since $0 < \frac{r}{p} < 1$, the function $f(x) = x^{\frac{r}{p}}$ is concave. Thus, by Jensen's inequality,

$$\left(\sum_{i=1}^{n} w_i x_i^p\right)^{\frac{r}{p}} \geq \sum_{i=1}^{n} w_i x_i^r.$$

Since $\frac{1}{r} < 0$, we obtain again

$$\left(\sum_{i=1}^{n} w_i x_i^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{i=1}^{n} w_i x_i^r\right)^{\frac{1}{r}}.$$

The indicator function of a subset S of a set Z is the function $I_S : Z \longrightarrow \hat{\mathbb{R}}$ defined by

$$I_S(z) = \begin{cases} 0 & \text{if } z \in S, \\ \infty & \text{if } z \not\in\in S. \end{cases}$$

Theorem

A set $S \subseteq \mathbb{R}^n$ is convex if and only if its indicator function I_S is convex.

Proof

If I_S is convex, we have $I_S(t\mathbf{x} + (1-t)\mathbf{y}) \leqslant tI_S(\mathbf{x}) + (1-t)I_S(\mathbf{y})$ for every $\mathbf{x}, \mathbf{y} \in S$. Therefore, if $\mathbf{x}, \mathbf{y} \in S$ we have $I_S(\mathbf{x}) = I_S(\mathbf{y}) = 0$, which implies $I_S(t\mathbf{x} + (1-t)\mathbf{y}) = 0$, so $t\mathbf{x} + (1-t)\mathbf{y} \in S$. Thus, S is convex.
Conversely, suppose that S is convex. We need to prove that

$$I_S(t\mathbf{x} + (1-t)\mathbf{y}) \leqslant tI_S(\mathbf{x}) + (1-t)I_S(\mathbf{y}). \tag{3}$$

If at least one of $\mathbf{x}$ or $\mathbf{y}$ does not belong to S, Inequality (3) is satisfied. The remaining case occurs when we have both $\mathbf{x} \in S$ and $\mathbf{y} \in S$, in which case, $t\mathbf{x} + (1-t)\mathbf{y} \in S$ and $I_S(\mathbf{x}) = I_S(\mathbf{y}) = I_S(t\mathbf{x} + (1-t)\mathbf{y}) = 0$, and, again, Inequality (3) is satisfied.