Convexity, Infimum and Supremum

Last modified: 2026-11-24 09:19:45 (PST)

Convex functions

Exercise 0.1 (A value compared with an average of values)**.** Let $f(x) = (x-2)^2$, and take the points $x = 0$ and $y = 4$ with $t = \frac{1}{2}$.

1. Compute $f(tx + (1-t)y)$.
2. Compute $tf(x) + (1-t)f(y)$.
3. Is $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$?

Solution

Solution 0.1.

1. With $t = \frac{1}{2}$,

$$\begin{aligned} tx + (1-t)y &= \tfrac{1}{2} \cdot 0 + \tfrac{1}{2} \cdot 4 \\ &= 2, \end{aligned}$$

   so

$$\begin{aligned} f(2) &= (2-2)^2 \\ &= 0. \end{aligned}$$

2.

$$\begin{aligned} f(0) &= (0-2)^2 \\ &= 4 \end{aligned}$$

   and

$$f(4) = (4-2)^2$$
$$= 4,$$

so $\frac{1}{2} \cdot 4 + \frac{1}{2} \cdot 4 = 4$.

3. Yes: $0 \leq 4$. The value of f at the midpoint of 0 and 4 is below the average of its values at the two endpoints.

Definition 0.1 (Line segment)**.** Let $x, y \in \mathbb{R}^p$. The **line segment** from x to y [1] is the set of points

$$\{tx + (1-t)y : t \in [0,1]\}.$$

The value $t = 1$ gives x, the value $t = 0$ gives y, and $t = \frac{1}{2}$ gives the midpoint $\frac{1}{2}(x+y)$.

Example 0.1 (A line segment in $\mathbb{R}$)**.** For $p = 1$, $x = 0$ and $y = 4$, the line segment from x to y is the interval $[0, 4]$. The value $t = \frac{1}{4}$ gives the point $\frac{1}{4} \cdot 0 + \frac{3}{4} \cdot 4 = 3$.

Definition 0.2 (Chord)**.** Let $f : \mathbb{R}^p \to \mathbb{R}$ be a function, and let $x, y \in \mathbb{R}^p$. The **chord** of f from x to y [2] is the line segment (Definition 0.1) in $\mathbb{R}^{p+1}$ from the point $(x, f(x))$ to the point $(y, f(y))$ of the graph of f:

$$\{(tx + (1-t)y,\ tf(x) + (1-t)f(y)) : t \in [0,1]\}.$$

Above the point $tx + (1-t)y$, the chord has height $tf(x) + (1-t)f(y)$.

Example 0.2 (A chord of x^2)**.** For $f(x) = x^2$, $x = -1$ and $y = 3$, the chord runs from $(-1, f(-1)) = (-1, 1)$ to $(3, f(3)) = (3, 9)$. With $t = \frac{1}{2}$, it passes through $\left(\frac{1}{2} \cdot (-1) + \frac{1}{2} \cdot 3,\ \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 9\right) = (1, 5)$, so above the point 1 it has height 5.

Definition 0.3 (Convex function)**.** Let $f : \mathbb{R}^p \to \mathbb{R}$ be a function. f is **convex** [3] if

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

for all $x, y \in \mathbb{R}^p$ and all $t \in [0, 1]$.

Remark (Chords lie on or above the graph)

Remark 0.1 (Chords lie on or above the graph). The point $tx + (1-t)y$ lies on the line segment from x to y, and the right-hand side is the height above that point of the chord of f from x to y. So f is convex when every chord lies on or above the graph.
For example, take $f(x) = x^2$, $x = -1$, $y = 3$, and $t = \frac{1}{2}$. The point is $\frac{1}{2} \cdot (-1) + \frac{1}{2} \cdot 3 = 1$, where the graph has height $f(1) = 1$ and the chord has height

$$\begin{aligned} \tfrac{1}{2} f(-1) + \tfrac{1}{2} f(3) &= \tfrac{1}{2} \cdot 1 + \tfrac{1}{2} \cdot 9 \\ &= 5. \end{aligned}$$

The chord is above the graph: $1 \leq 5$.

Example 0.3 (A function that is not convex)**.** The function in Exercise 4 in Equalities, Inequalities and Minimizers, $f(x) = x^3 - 3x$, is not convex. Take $x = -2$, $y = 0$, $t = \frac{1}{2}$. The point is $\frac{1}{2} \cdot (-2) + \frac{1}{2} \cdot 0 = -1$, where the graph has height

$$\begin{aligned} f(-1) &= -1 + 3 \\ &= 2, \end{aligned}$$

but the chord has height

$$\begin{aligned} \tfrac{1}{2} f(-2) + \tfrac{1}{2} f(0) &= \tfrac{1}{2}(-8 + 6) + \tfrac{1}{2} \cdot 0 \\ &= -1, \end{aligned}$$

and $2 \leq -1$ is false.

Example 0.4 (A convex and a non-convex function)**.** $f(x) = (x-2)^2$ is convex (Definition 0.3): by expanding the square, $f(tx+(1-t)y) - tf(x) - (1-t)f(y) = -t(1-t)(x-y)^2 \leq 0$. Its local minimizer $x^* = 2$ is also a global minimizer, since

$$\begin{aligned} f(x) &= (x-2)^2 \\ &\geq 0 \\ &= f(2) \end{aligned}$$

for every x.
Without convexity, a local minimizer need not be global: $f(x) = x^3 - 3x$ is not convex (Example 0.3), and it has a local minimizer at $x^* = 1$ that is not global (Exercise 4 in Equalities, Inequalities and Minimizers).

Theorem 0.1 (Local minimizers of convex functions are global)**.** *Let $f : \mathbb{R}^p \to \mathbb{R}$ be a convex function. Every local minimizer of f is a global minimizer of f over $\mathbb{R}^p$.*

Proof

Proof. Let x^* be a local minimizer of f, so there is a $\delta > 0$ with $f(x^*) \leq f(x)$ whenever $\|x - x^*\| < \delta$ (Definition 19 in Equalities, Inequalities and Minimizers). Suppose x^* is not a global minimizer (Definition 18 in Equalities, Inequalities and Minimizers). Then some $y \in \mathbb{R}^p$ has $f(y) < f(x^*)$, and in particular $y \neq x^*$.
Let $t = \min\left\{\frac{1}{2}, \frac{\delta}{2\|y - x^*\|}\right\}$, so $t \in (0, 1)$, and let $z = ty + (1-t)x^*$. Then $z - x^* = t(y - x^*)$, so $\|z - x^*\| = t\|y - x^*\| \leq \delta/2 < \delta$.
By convexity,

$$f(z) \leq tf(y) + (1-t)f(x^*) < tf(x^*) + (1-t)f(x^*) = f(x^*),$$

where the strict inequality uses $t > 0$ and $f(y) < f(x^*)$. So $f(z) < f(x^*)$ with $\|z - x^*\| < \delta$, which contradicts x^* being a local minimizer. Hence x^* is a global minimizer. □

Theorem 0.2 (Jensen's inequality for a weighted average)**.** *Let $f : \mathbb{R}^p \to \mathbb{R}$ be a convex function [4]. Let $x_1, \ldots, x_n \in \mathbb{R}^p$, and let $w_1, \ldots, w_n \geq 0$ be weights with $\sum_{i=1}^n w_i = 1$. Then*

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

Proof

Proof. The proof is by induction on n.
For $n = 1$ the weight is $w_1 = 1$, and both sides equal $f(x_1)$.
Now let $n \geq 2$, and assume the inequality holds for $n - 1$ points. If $w_n = 1$, the other weights are all 0, and both sides equal $f(x_n)$. Otherwise let

$$\begin{aligned} s &= 1 - w_n \\ &= \sum_{i=1}^{n-1} w_i \\ &> 0 \end{aligned}$$

and

$$y = \sum_{i=1}^{n-1} \frac{w_i}{s} x_i.$$

The weights w_i/s are non-negative and sum to 1, so the induction assumption gives

$$f(y) \leq \sum_{i=1}^{n-1} \frac{w_i}{s} f(x_i).$$

Since $\sum_{i=1}^{n} w_i x_i = sy + (1-s)x_n$ and $s \in (0,1]$, the definition of a convex function (Definition 0.3) gives

$$f\left(\sum_{i=1}^{n} w_i x_i\right) = f(sy + (1-s)x_n) \leq sf(y) + (1-s)f(x_n).$$

Substituting the bound on $f(y)$, and using $1 - s = w_n$,

$$sf(y) + (1-s)f(x_n) \leq \sum_{i=1}^{n-1} w_i f(x_i) + w_n f(x_n) = \sum_{i=1}^{n} w_i f(x_i).$$

□

Example 0.5 (The average of the squares is at least the square of the average)**.** Take $f(x) = x^2$. It is convex by Definition 0.3: for $t \in [0,1]$,

$$tx^2 + (1-t)y^2 - (tx + (1-t)y)^2 = t(1-t)(x-y)^2 \geq 0.$$

Take the three points $x_1 = 1$, $x_2 = 2$, $x_3 = 6$, and the equal weights

$$\begin{aligned} w_1 &= w_2 \\ &= w_3 \\ &= \tfrac{1}{3}. \end{aligned}$$

The weighted average of the points is their ordinary average, $\frac{1}{3}(1+2+6) = 3$, so the left side of Theorem 0.2 is $f(3) = 9$. The right side is the average of the squares, $\frac{1}{3}(1+4+36) = \frac{41}{3}$. The inequality holds: $9 \leq \frac{41}{3}$.
The gap is $\frac{41}{3} - 9 = \frac{14}{3}$. That is the average squared distance of the points from their average, $\frac{1}{3}((1-3)^2 + (2-3)^2 + (6-3)^2) = \frac{14}{3}$, which statistics calls the variance of the three points (about their mean, defined later). For $f(x) = x^2$, Jensen's inequality says this average squared distance is never negative.

i Remark (The inequality reverses for concave functions)

Remark 0.2 (The inequality reverses for concave functions). If f is concave, then $-f$ is convex, and applying Theorem 0.2 to $-f$ reverses the inequality: $f\left(\sum_i w_i x_i\right) \geq \sum_i w_i f(x_i)$. For $f(x) = -x^2$ and the points in Example 0.5, $f(3) = -9 \geq -\frac{41}{3}$.

Infimum and supremum

Definition 0.1 (Upper bound)**.** Let $A \subseteq \mathbb{R}$. A real number u is an **upper bound** for A [5] if $a \leq u$ for all $a \in A$.

Definition 0.2 (Lower bound)**.** Let $A \subseteq \mathbb{R}$. A real number ℓ is a **lower bound** for A [5] if $\ell \leq a$ for all $a \in A$.

Definition 0.3 (Bounded above)**.** A set $A \subseteq \mathbb{R}$ is **bounded above** if it has an upper bound (Definition 0.1). A real-valued function is bounded above if its image is.

Definition 0.4 (Bounded below)**.** A set $A \subseteq \mathbb{R}$ is **bounded below** if it has a lower bound (Definition 0.2). A real-valued function is bounded below if its image is.

Definition 0.5 (Bounded)**.** A set $A \subseteq \mathbb{R}$ is **bounded** if it is both bounded above (Definition 0.3) and bounded below (Definition 0.4). A real-valued function is bounded if its image is.

Example 0.1 (Bounded and unbounded sets)**.**

- For $A = (0, 1]$, 1, 2, and 100 are upper bounds, and 0 and -5 are lower bounds, so A is bounded. 0.5 is not an upper bound, because $0.6 \in A$ and $0.6 > 0.5$.
- The natural numbers $\mathbb{N} = \{1, 2, 3, ...\}$ are bounded below, by 1, but not bounded above: for any real number u, some natural number n satisfies $n > u$.

Theorem 0.1 (Completeness of the real numbers)**.** *Let $A \subseteq \mathbb{R}$ be nonempty [6].*

- *If A is bounded above, the set of upper bounds of A has a minimum.*
- *If A is bounded below, the set of lower bounds of A has a maximum.*

Proof

Proof. The first statement is the least-upper-bound property of $\mathbb{R}$ [7], Definition 1.8, p. 4, and Theorem 1.19, p. 8. For the second, let $-A = \{-a : a \in A\}$. Negating both sides of an inequality reverses it (Theorem 6 in Equalities, Inequalities and Minimizers), so ℓ is a lower bound of A exactly when $-\ell$ is an upper bound of $-A$. The set $-A$ is nonempty and bounded above, so by the first statement its upper bounds have a minimum u, and

then $-u$ is the largest lower bound of A. □

Example 0.2 (Completeness for an interval)**.** For $A = (1, 2]$, the lower bounds of A are the numbers $\ell \leq 1$, and the largest of them is 1. The upper bounds of A are the numbers $u \geq 2$, and the smallest of them is 2.

Definition 0.6 (Infimum, greatest lower bound)**.** Let $A \subseteq \mathbb{R}$ be nonempty and bounded below (Definition 0.4). The **infimum** of A [8], written $\inf A$, is the greatest real number t satisfying $t \leq a$ for all $a \in A$:

$$\inf A \stackrel{\text{def}}{=} \max\{t \in \mathbb{R} : \forall a \in A,\ t \leq a\}$$

If A is nonempty but not bounded below, we write $\inf A = -\infty$ by convention.

ℹ Remark (Existence of the infimum, and when it is a minimum)

Remark 0.1 (Existence of the infimum, and when it is a minimum). The maximum in Definition 0.6 always exists, by the completeness of the real numbers (Theorem 0.1). For example, for $A = (1, 2]$, the numbers t with $t \leq a$ for all $a \in A$ are those with $t \leq 1$, and the largest of them is 1, so $\inf A = 1$.
If the infimum belongs to A, it equals the minimum: $\inf A = \min A$. For example,

$$\begin{aligned} \inf[1, 2] &= 1 \\ &= \min[1, 2]. \end{aligned}$$

For $A = (1, 2]$, the infimum 1 is not in A, and A has no minimum.

Example 0.3 (Numerical examples of infimum)**.**

- $\inf\{1, 2, 3\} = 1$, since 1 is the smallest element.
- $$\begin{aligned} \inf(0.5, 1] &= 0.5 \\ &= \min[0.5, 1] : \end{aligned}$$

  for intervals open below, the infimum equals the minimum of the corresponding closed-below interval, even though $0.5 \notin (0.5, 1]$. More generally,

  $$\begin{aligned} \inf(c, b] &= \min[c, b] \\ &= c \end{aligned}$$

  for any $c < b$.

- $\inf\{t \geq 0 : t > 0.5\} = 0.5$, even though 0.5 itself is not in the set.
- $\inf\{-1, -2, -3, ...\} = -\infty$, because no real number is less than or equal to every element of that set.

Definition 0.7 (Supremum, least upper bound)**.** Let $A \subseteq \mathbb{R}$ be nonempty and bounded above (Definition 0.3). The **supremum** of A [8], written $\sup A$, is the smallest real number t satisfying $a \leq t$ for all $a \in A$:

$$\sup A \stackrel{\text{def}}{=} \min\{t \in \mathbb{R} : \forall a \in A,\ a \leq t\}$$

If A is nonempty but not bounded above, we write $\sup A = +\infty$ by convention.

Remark (Existence of the supremum, and when it is a maximum)

Remark 0.2 (Existence of the supremum, and when it is a maximum). The minimum in Definition 0.7 always exists, by the completeness of the real numbers (Theorem 0.1). For example, for $A = [1, 2)$, the numbers t with $a \leq t$ for all $a \in A$ are those with $t \geq 2$, and the smallest of them is 2, so $\sup A = 2$.
If the supremum belongs to A, it equals the maximum: $\sup A = \max A$. For example,

$$\begin{aligned} \sup[1, 2] &= 2 \\ &= \max[1, 2]. \end{aligned}$$

For $A = [1, 2)$, the supremum 2 is not in A, and A has no maximum.

Example 0.4 (Numerical examples of supremum)**.**

- $\sup\{1, 2, 3\} = 3$, since 3 is the largest element.
- $\sup\{t \geq 0 : t < 0.5\} = 0.5$, even though 0.5 itself is not in the set.
- $\sup\{1, 2, 3, ...\} = +\infty$, because no real number is greater than or equal to every element of that set.

[1] Wikipedia contributors, "Line segment — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Line_segment

[2] Wikipedia contributors, "Chord (geometry) — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Chord_(geometry)

[3] Wikipedia contributors, "Convex function — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Convex_function

[4] Wikipedia contributors, "Jensen's inequality — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Jensen%27s_inequality

[5] Wikipedia contributors, "Upper and lower bounds — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Upper_and_lower_bounds

[6] Wikipedia contributors, "Least-upper-bound property — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Least-upper-bound_property

[7] W. Rudin, *Principles of mathematical analysis*, 3rd ed. in International series in pure and applied mathematics. New York: McGraw-Hill, 1976.

[8] Wikipedia contributors, "Infimum and supremum — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Infimum_and_supremum