Convexity in Several Variables

Last modified: 2026-10-31 15:22:31 (PDT)

Convexity in several variables

Source

This section is adapted from the convexity part of Zhou [1], used under the MIT License (see the license text in Linear Algebra). The source states the first-order test (in one direction) and the second-derivative test for convexity without proof; here both directions of the first-order test are proved, and the second-derivative test is proved from the Taylor theorem of Section 2 in Critical Points and Optimality Conditions. The source defines both concave and strictly concave as "$-f$ is strictly convex"; here f is concave when $-f$ is convex. The source states that an intersection of convex sets is convex and that a maximum of convex functions is convex for any finite number of them; here each is stated for two and the general case follows by induction. These parts of the source are not part of this section:

- its catalogs of convex sets and convex functions, including the positive semidefinite matrices and $-\log\det$
- intersections of infinitely many convex sets, and suprema of infinitely many convex functions
- its examples, exercises and figures, including the multivariate normal maximum likelihood example

Definition 0.1 (Convex set)**.** A set $K \subseteq \mathbb{R}^p$ is **convex** if [2] for all $\underset{\sim}{x}, \underset{\sim}{y} \in K$ and all $t \in [0,1]$, the point $t\underset{\sim}{x} + (1-t)\underset{\sim}{y}$ is in K: K contains the whole line segment between any two of its points.

Example 0.1 (Convex sets)**.**

- $\mathbb{R}^p$ is convex, and so is any set with at most one point.

- The closed ball (closed ball) $B = \{\underset{\sim}{x} : \|\underset{\sim}{x}\| \le r\}$ is convex: for $\underset{\sim}{x}, \underset{\sim}{y} \in B$ and $t \in [0,1]$,

$$\begin{aligned} \|t\underset{\sim}{x} + (1-t)\underset{\sim}{y}\| &\le \|t\underset{\sim}{x}\| + \|(1-t)\underset{\sim}{y}\| && \text{(triangle inequality)} \\ &= t\,\|\underset{\sim}{x}\| + (1-t)\,\|\underset{\sim}{y}\| && \text{(norm of a multiple; } t, 1-t \ge 0) \\ &\le t\,r + (1-t)\,r && (\underset{\sim}{x}, \underset{\sim}{y} \in B) \\ &= r. && \text{(arithmetic)} \end{aligned}$$

  The first step is the triangle inequality, and the second uses the norm of a multiple.

Example 0.2 (A set that is not convex)**.** $K = \{x \in \mathbb{R} : |x| \ge 1\}$ is not convex: -1 and 1 are in K, but with $t = \frac{1}{2}$ the point $\frac{1}{2}(-1) + \frac{1}{2}(1) = 0$ is not.

Definition 0.2 (Half-space)**.** Let $\underset{\sim}{a} \in \mathbb{R}^p$ be a nonzero vector and c a real number. The set $\{\underset{\sim}{x} \in \mathbb{R}^p : \underset{\sim}{a} \cdot \underset{\sim}{x} \ge c\}$ is a (closed) **half-space**: the points on one side of the boundary $\{\underset{\sim}{x} : \underset{\sim}{a} \cdot \underset{\sim}{x} = c\}$, together with the boundary. In $\mathbb{R}^2$, a half-space is called a **half-plane**.

Example 0.3 (The right half-plane)**.** With $\underset{\sim}{a} = (1,0)^\top$ and $c = 0$, the half-plane $\{\underset{\sim}{x} \in \mathbb{R}^2 : x_1 \ge 0\}$ is the set of points on or to the right of the x_2-axis. $(1,5)^\top$ is in it, since $1 \ge 0$; $(0,-3)^\top$ is in it, on the boundary; and $(-1,0)^\top$ is not, since $-1 < 0$.

Theorem 0.1 (An intersection of convex sets is convex)**.** *If K_1 and K_2 are convex subsets of $\mathbb{R}^p$ (Definition 0.1), then $K_1 \cap K_2$ is convex.*

i Proof

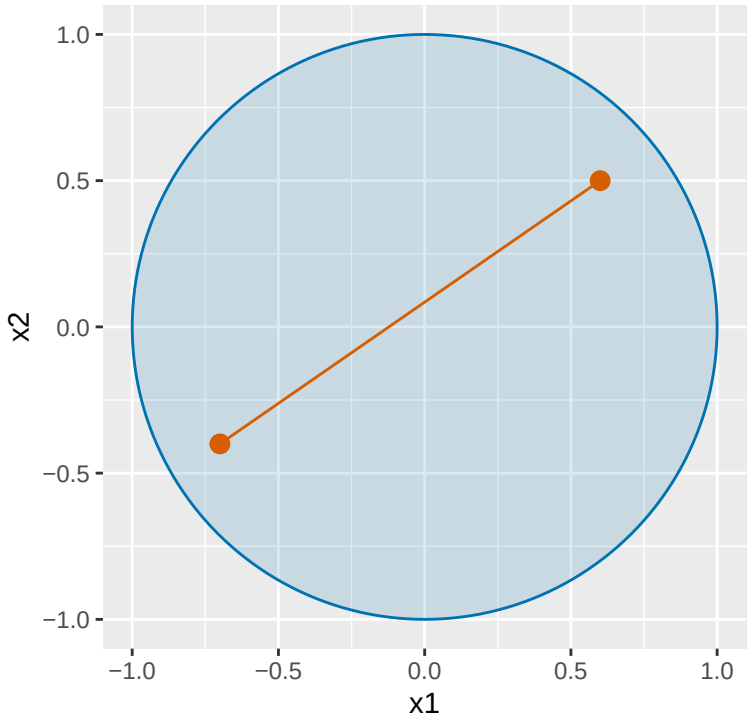
Proof. Let $\underset{\sim}{x}, \underset{\sim}{y} \in K_1 \cap K_2$ and $t \in [0,1]$. Then $\underset{\sim}{x}, \underset{\sim}{y} \in K_1$, so $t\underset{\sim}{x} + (1-t)\underset{\sim}{y} \in K_1$, since K_1 is convex; in the same way $t\underset{\sim}{x} + (1-t)\underset{\sim}{y} \in K_2$. So $t\underset{\sim}{x} + (1-t)\underset{\sim}{y} \in K_1 \cap K_2$. □

Example 0.4 (Intersecting a ball with a half-plane)**.** The closed unit ball (unit ball) in $\mathbb{R}^2$ is convex (Example 0.1), and so is the half-plane (Definition 0.2) $\{\underset{\sim}{x} : x_1 \ge 0\}$: if $x_1 \ge 0$ and $y_1 \ge 0$, then $tx_1 + (1-t)y_1 \ge 0$ for $t \in [0,1]$. By Theorem 0.1, the closed half-disk $\{\underset{\sim}{x} : \|\underset{\sim}{x}\| \le 1,\ x_1 \ge 0\}$ is convex. A union need not be convex: $\{x \in \mathbb{R} : x \le -1\}$ and $\{x \in \mathbb{R} : x \ge 1\}$ are convex, but their union is the set of Example 0.2.

Definition 0.3 (Strictly convex function)**.** A function $f : \mathbb{R}^p \to \mathbb{R}$ is **strictly convex** if $f(t\underset{\sim}{x} + (1-t)\underset{\sim}{y}) < tf(\underset{\sim}{x}) + (1-t)f(\underset{\sim}{y})$ for all $\underset{\sim}{x} \ne \underset{\sim}{y}$ in $\mathbb{R}^p$ and all $t \in (0,1)$.

R

```r
cs_angle_fc <- seq(0, 2 * pi, length.out = 201)
cs_cos_fc <- cos(cs_angle_fc)
cs_sin_fc <- sin(cs_angle_fc)
cs_disk_fc <- tibble::tibble(x1 = cs_cos_fc, x2 = cs_sin_fc)
cs_points_x1_fc <- c(-0.7, 0.6)
cs_points_x2_fc <- c(-0.4, 0.5)
cs_points_fc <- tibble::tibble(x1 = cs_points_x1_fc, x2 = cs_points_x2_fc)
cs_segment_fc <- tibble::tibble(
  x1 = -0.7, x2 = -0.4, x1_end = 0.6, x2_end = 0.5
)
cs_aes_fc <- ggplot2::aes(x1, x2)
cs_seg_aes_fc <- ggplot2::aes(xend = x1_end, yend = x2_end)
ggplot2::ggplot(cs_disk_fc, cs_aes_fc) +
  ggplot2::geom_polygon(fill = "#0072B2", alpha = 0.15, colour = "#0072B2") +
  ggplot2::geom_segment(
    data = cs_segment_fc, mapping = cs_seg_aes_fc, colour = "#D55E00"
  ) +
  ggplot2::geom_point(data = cs_points_fc, colour = "#D55E00", size = 3) +
  ggplot2::coord_fixed() +
  ggplot2::labs(x = "x1", y = "x2")
```



Python


```python
import matplotlib.pyplot as plt
import numpy as np

cs_angle_fc = np.linspace(0, 2 * np.pi, 201)
cs_x1_fc = np.array([-0.7, 0.6])
cs_x2_fc = np.array([-0.4, 0.5])
```

Definition 0.4 (Concave function)**.** A function $f : \mathbb{R}^p \to \mathbb{R}$ is **concave** if $-f$ is convex.

Definition 0.5 (Strictly concave function)**.** A function $f : \mathbb{R}^p \to \mathbb{R}$ is **strictly concave** if $-f$ is strictly convex (Definition 0.3).

Example 0.5 (Strictly convex, and convex but not strictly)**.**

- $f(x) = x^2$ is strictly convex:

$$\begin{aligned} tx^2 + (1-t)y^2 - (tx + (1-t)y)^2 &= tx^2 + (1-t)y^2 - t^2x^2 - 2t(1-t)xy - (1-t)^2y^2 && \text{(expand the squa} \\ &= t(1-t)x^2 - 2t(1-t)xy + t(1-t)y^2 && (t - t^2 = t(1-t) \\ &= t(1-t)(x-y)^2, && \text{(factor)} \end{aligned}$$

  which is positive when $x \neq y$ and $t \in (0,1)$. At $t \in \{0,1\}$ or $x = y$ it is 0, so x^2 is also convex.

- $f(x) = 2x + 1$ is convex but not strictly convex:

$$\begin{aligned} f(tx + (1-t)y) &= 2tx + 2(1-t)y + 1 \\ &= tf(x) + (1-t)f(y) \end{aligned}$$

  for every x, y, t, so the inequality holds, but never strictly. It is also concave.

Theorem 0.2 (The maximum of convex functions is convex)**.** *If f_1 and f_2 are convex functions on $\mathbb{R}^p$, then $f(\underset{\sim}{x}) \overset{def}{=} \max\{f_1(\underset{\sim}{x}), f_2(\underset{\sim}{x})\}$ is convex.*

Proof

Proof. Let $\underset{\sim}{x}, \underset{\sim}{y} \in \mathbb{R}^p$ and $t \in [0,1]$. For $i = 1, 2$,

$$\begin{aligned} f_i(t\underset{\sim}{x} + (1-t)\underset{\sim}{y}) &\leq tf_i(\underset{\sim}{x}) + (1-t)f_i(\underset{\sim}{y}) && (f_i \text{ is convex}) \\ &\leq tf(\underset{\sim}{x}) + (1-t)f(\underset{\sim}{y}). && (f_i \leq f, \text{ and } t, 1-t \geq 0) \end{aligned}$$

Both f_1 and f_2 at $t\underset{\sim}{x} + (1-t)\underset{\sim}{y}$ are at most $tf(\underset{\sim}{x}) + (1-t)f(\underset{\sim}{y})$, so their maximum $f(t\underset{\sim}{x} + (1-t)\underset{\sim}{y})$ is too. □

Example 0.6 (A maximum is convex; a minimum need not be)**.**

- $f_1(x) = x$ is convex, by the computation for $2x + 1$ in Example 0.5 with 2 and 1 replaced by 1 and 0, and $f_2(x) = x^2$ is convex by the same example, so $\max\{x, x^2\}$ is convex by Theorem 0.2.
- The minimum of two convex functions need not be convex: $g(x) = \min\{x^2, (x-2)^2\}$ has

$$\begin{aligned} g(0) &= g(2) \\ &= 0, \end{aligned}$$

but at the midpoint

$$\begin{aligned} g(1) &= \min\{1, 1\} \\ &= 1 \\ &> \tfrac{1}{2}g(0) + \tfrac{1}{2}g(2) \\ &= 0. \end{aligned}$$

Definition 0.6 (Epigraph)**.** The **epigraph** [3] of $f : \mathbb{R}^p \to \mathbb{R}$ is the set of points on or over its graph,

$$\operatorname{epi} f \stackrel{\text{def}}{=} \{(\underset{\sim}{x}, s) : \underset{\sim}{x} \in \mathbb{R}^p,\ s \in \mathbb{R},\ f(\underset{\sim}{x}) \le s\}.$$

A pair $(\underset{\sim}{x}, s)$ is treated as the $(p+1)$-vector with entries $x_1, \ldots, x_p, s$, so $\operatorname{epi} f \subseteq \mathbb{R}^{p+1}$, and sums and multiples of pairs are taken entry by entry.

Example 0.7 (The epigraph of x^2)**.** For $f(x) = x^2$, $\operatorname{epi} f = \{(x, s) : x^2 \le s\}$, the points (x, s) on the parabola $s = x^2$ or on the side of it where s is larger. $(1, 2)$ is in it, since $1^2 \le 2$; $(2, 1)$ is not, since $2^2 > 1$.

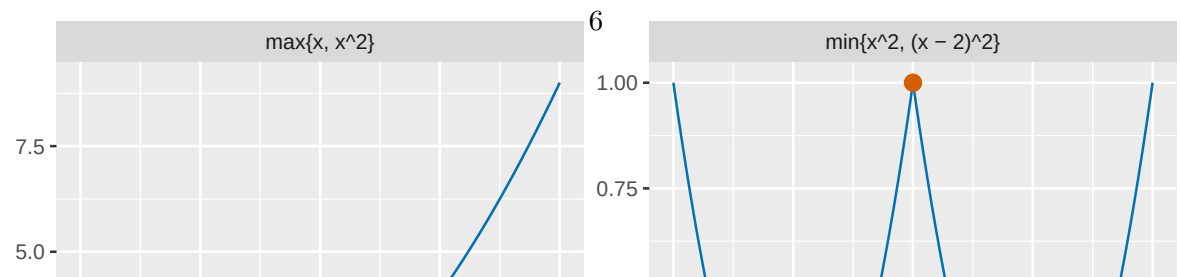
Theorem 0.3 (Convexity through the epigraph)**.** *$f : \mathbb{R}^p \to \mathbb{R}$ is convex if and only if its epigraph (Definition 0.6) is convex (Definition 0.1).*

Proof

Proof. **If.** Let $\underset{\sim}{x}, \underset{\sim}{y} \in \mathbb{R}^p$ and $t \in [0, 1]$. The points $(\underset{\sim}{x}, f(\underset{\sim}{x}))$ and $(\underset{\sim}{y}, f(\underset{\sim}{y}))$ are in $\operatorname{epi} f$, since $f(\underset{\sim}{x}) \le f(\underset{\sim}{x})$ and $f(\underset{\sim}{y}) \le f(\underset{\sim}{y})$. Since $\operatorname{epi} f$ is convex, it contains

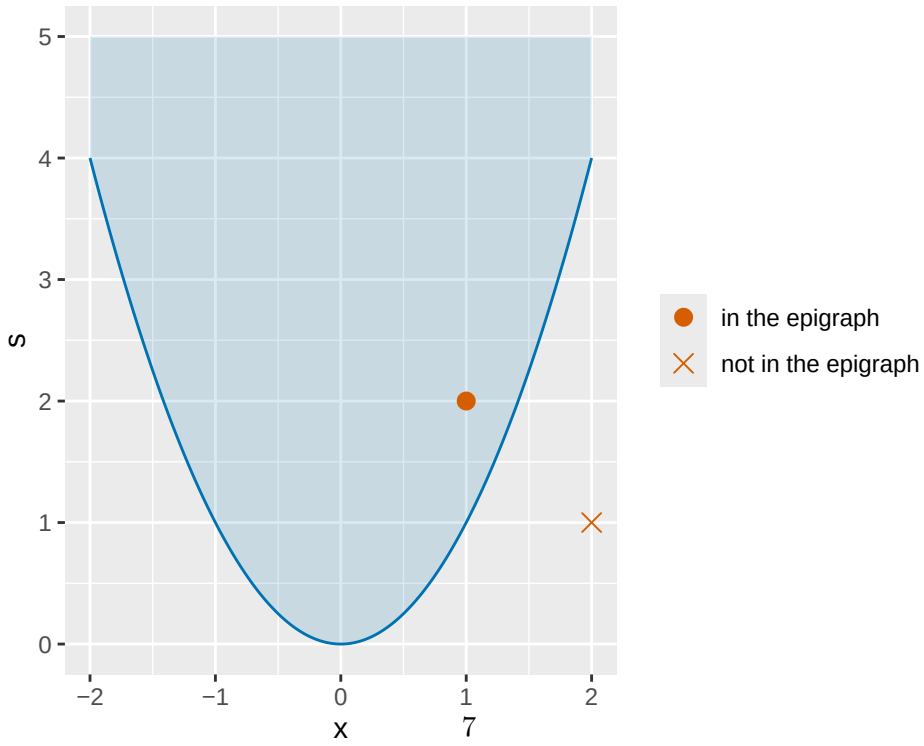
R

```
cm_x_fm <- seq(-1, 3, length.out = 201)
cm_levels_fm <- c("max{x, x^2}", "min{x^2, (x - 2)^2}")
cm_max_value_fm <- pmax(cm_x_fm, cm_x_fm^2)
cm_min_value_fm <- pmin(cm_x_fm^2, (cm_x_fm - 2)^2)
cm_max_fm <- tibble::tibble(
  x = cm_x_fm, value = cm_max_value_fm, panel = cm_levels_fm[1]
)
cm_min_fm <- tibble::tibble(
  x = cm_x_fm, value = cm_min_value_fm, panel = cm_levels_fm[2]
)
cm_main_fm <- dplyr::bind_rows(cm_max_fm, cm_min_fm)
cm_main_fm$panel <- factor(cm_main_fm$panel, levels = cm_levels_fm)
cm_square_value_fm <- cm_x_fm^2
cm_line_fm <- tibble::tibble(
  x = cm_x_fm, value = cm_x_fm, panel = cm_levels_fm[1], part = "x"
)
cm_square_fm <- tibble::tibble(
  x = cm_x_fm, value = cm_square_value_fm, panel = cm_levels_fm[1],
  part = "x^2"
)
cm_parts_fm <- dplyr::bind_rows(cm_line_fm, cm_square_fm)
cm_parts_fm$panel <- factor(cm_parts_fm$panel, levels = cm_levels_fm)
cm_chord_fm <- tibble::tibble(
  x = 0, xend = 2, value = 0, panel = cm_levels_fm[2]
)
cm_chord_fm$panel <- factor(cm_chord_fm$panel, levels = cm_levels_fm)
cm_mid_fm <- tibble::tibble(x = 1, value = 1, panel = cm_levels_fm[2])
cm_mid_fm$panel <- factor(cm_mid_fm$panel, levels = cm_levels_fm)
cm_aes_fm <- ggplot2::aes(x, value)
cm_parts_aes_fm <- ggplot2::aes(group = part)
cm_chord_aes_fm <- ggplot2::aes(xend = xend, yend = value)
ggplot2::ggplot(cm_main_fm, cm_aes_fm) +
  ggplot2::geom_line(
    data = cm_parts_fm, mapping = cm_parts_aes_fm,
    colour = "grey60", linewidth = 0.4
  ) +
  ggplot2::geom_line(colour = "#0072B2") +
  ggplot2::geom_segment(
    data = cm_chord_fm, mapping = cm_chord_aes_fm, colour = "#D55E00"
  ) +
  ggplot2::geom_point(data = cm_mid_fm, colour = "#D55E00", size = 3) +
  ggplot2::facet_wrap(~panel, scales = "free_y") +
  ggplot2::labs(x = "x", y = NULL)
```



R

```r
ep_x_fe <- seq(-2, 2, length.out = 201)
ep_curve_fe <- tibble::tibble(x = ep_x_fe, s = ep_x_fe^2, top = 5)
ep_marks_x_fe <- c(1, 2)
ep_marks_s_fe <- c(2, 1)
ep_marks_kind_fe <- c("in the epigraph", "not in the epigraph")
ep_marks_fe <- tibble::tibble(
  x = ep_marks_x_fe, s = ep_marks_s_fe, kind = ep_marks_kind_fe
)
ep_aes_fe <- ggplot2::aes(x, s)
ep_ribbon_aes_fe <- ggplot2::aes(ymin = s, ymax = top)
ep_marks_aes_fe <- ggplot2::aes(shape = kind)
ep_shapes_fe <- c("in the epigraph" = 16, "not in the epigraph" = 4)
ggplot2::ggplot(ep_curve_fe, ep_aes_fe) +
  ggplot2::geom_ribbon(
    mapping = ep_ribbon_aes_fe, fill = "#0072B2", alpha = 0.15
  ) +
  ggplot2::geom_line(colour = "#0072B2") +
  ggplot2::geom_point(
    data = ep_marks_fe, mapping = ep_marks_aes_fe,
    colour = "#D55E00", size = 3
  ) +
  ggplot2::scale_shape_manual(values = ep_shapes_fe) +
  ggplot2::labs(x = "x", y = "s", shape = NULL)
```



Python

```python
import matplotlib.pyplot as plt
import numpy as np
```

$$t\,(\underset{\sim}{x}, f(\underset{\sim}{x})) + (1-t)\,(\underset{\sim}{y}, f(\underset{\sim}{y})) = (t\underset{\sim}{x} + (1-t)\underset{\sim}{y},\ tf(\underset{\sim}{x}) + (1-t)f(\underset{\sim}{y})),$$

and by Definition 0.6 that membership means $f(t\underset{\sim}{x} + (1-t)\underset{\sim}{y}) \le tf(\underset{\sim}{x}) + (1-t)f(\underset{\sim}{y})$.
Only if. Let $(\underset{\sim}{x}, s)$ and $(\underset{\sim}{y}, u)$ be in epi f, so $f(\underset{\sim}{x}) \le s$ and $f(\underset{\sim}{y}) \le u$, and let $t \in [0,1]$. Then

$$\begin{aligned} f(t\underset{\sim}{x} + (1-t)\underset{\sim}{y}) &\le tf(\underset{\sim}{x}) + (1-t)f(\underset{\sim}{y}) && (f \text{ is convex}) \\ &\le ts + (1-t)u, && (f(\underset{\sim}{x}) \le s,\ f(\underset{\sim}{y}) \le u, \text{ and } t, 1-t \ge 0) \end{aligned}$$

so $t\,(\underset{\sim}{x}, s) + (1-t)\,(\underset{\sim}{y}, u) = (t\underset{\sim}{x} + (1-t)\underset{\sim}{y},\ ts + (1-t)u)$ is in epi f. □

Example 0.8 (Two epigraphs)**.**

- The epigraph of $f(x) = |x|$ is $\{(x,s) : |x| \le s\}$, a wedge. It is the intersection of the half-planes $\{(x,s) : s - x \ge 0\}$ and $\{(x,s) : s + x \ge 0\}$. Each half-plane is convex: if $s_1 - x_1 \ge 0$ and $s_2 - x_2 \ge 0$, then $(ts_1 + (1-t)s_2) - (tx_1 + (1-t)x_2) = t(s_1 - x_1) + (1-t)(s_2 - x_2) \ge 0$, and the same with $+$ for the other. So the wedge is convex (Theorem 0.1), and $|x|$ is convex by Theorem 0.3.
- The epigraph of $g(x) = -x^2$ is not convex: $(-1,-1)$ and $(1,-1)$ are in it, since $g(\pm 1) = -1$, but their midpoint $(0,-1)$ is not, since $g(0) = 0 > -1$. So $-x^2$ is not convex.

Definition 0.7 (Tangent plane)**.** Let $f : \mathbb{R}^p \to \mathbb{R}$ have a gradient at $\underset{\sim}{y}$ (gradient). The **tangent plane** [4] to the graph of f at $\underset{\sim}{y}$ is the graph of the function

$$\underset{\sim}{x} \mapsto f(\underset{\sim}{y}) + \left(\frac{\partial}{\partial \underset{\sim}{x}} f(\underset{\sim}{y})\right)^{\top} (\underset{\sim}{x} - \underset{\sim}{y}),$$

which takes the value $f(\underset{\sim}{y})$ at $\underset{\sim}{x} = \underset{\sim}{y}$ and has the same gradient as f there.

Example 0.9 (A tangent line and a tangent plane)**.**

- For $f(x) = x^2$ at $y = 1$, $f(1) = 1$ and $f'(1) = 2$, so the tangent line is

$$\begin{aligned} s &= 1 + 2\,(x-1) && (\text{substitute } f(1) \text{ and } f'(1)) \\ &= 1 + 2x - 2 && (\text{distribute}) \\ &= 2x - 1. && (1 - 2 = -1) \end{aligned}$$

- For $f(\underset{\sim}{x}) = x_1^2 + x_2^2$ at $\underset{\sim}{y} = (1,1)^{\top}$, $f(\underset{\sim}{y}) = 2$ and the gradient is $(2y_1,\ 2y_2)^{\top} = (2,2)^{\top}$, so the tangent plane is

$$
\begin{aligned}
s &= 2 + 2\,(x_1 - 1) + 2\,(x_2 - 1) && \text{(substitute } f(\underset{\sim}{y}) \text{ and the gradient)} \\
&= 2 + 2x_1 - 2 + 2x_2 - 2 && \text{(distribute)} \\
&= 2x_1 + 2x_2 - 2. && (2 - 2 - 2 = -2)
\end{aligned}
$$

ℹ Remark (The tangent plane in one variable)

Remark 0.1 (The tangent plane in one variable). For $p = 1$ the tangent plane is the tangent line $s = f(y) + f'(y)\,(x - y)$, the line through $(y, f(y))$ with slope $f'(y)$.

Theorem 0.4 (First-order characterization of convexity)**.** *Let $f : \mathbb{R}^p \to \mathbb{R}$ have first partial derivatives that are continuous on $\mathbb{R}^p$ (continuity). Then f is convex if and only if*

$$
f(\underset{\sim}{x}) \geq f(\underset{\sim}{y}) + \left(\frac{\partial}{\partial \underset{\sim}{x}} f(\underset{\sim}{y})\right)^{\top} (\underset{\sim}{x} - \underset{\sim}{y}) \qquad \text{for all } \underset{\sim}{x}, \underset{\sim}{y} \in \mathbb{R}^p : \tag{1}
$$

every tangent plane (Definition 0.7) lies on or under the graph. If the inequality is strict whenever $\underset{\sim}{x} \neq \underset{\sim}{y}$, then f is strictly convex (Definition 0.3).

ℹ Proof

Proof. **Only if.** Let $\underset{\sim}{x}, \underset{\sim}{y} \in \mathbb{R}^p$, let $\underset{\sim}{d} \overset{\text{def}}{=} \underset{\sim}{x} - \underset{\sim}{y}$, and let $g(t) \overset{\text{def}}{=} f(\underset{\sim}{y} + t\underset{\sim}{d})$. The inner function $t \mapsto \underset{\sim}{y} + t\underset{\sim}{d}$ has derivative $\underset{\sim}{d}^{\top}$ (vector-valued derivative), so the vector chain rule (vector chain rule) gives

$$
\begin{aligned}
g'(0) &= \underset{\sim}{d}^{\top} \frac{\partial}{\partial \underset{\sim}{x}} f(\underset{\sim}{y}) && \text{(vector chain rule)} \\
&= \left(\frac{\partial}{\partial \underset{\sim}{x}} f(\underset{\sim}{y})\right)^{\top} \underset{\sim}{d}. && \text{(a } 1 \times 1 \text{ matrix equals its transpose)}
\end{aligned}
$$

For $t \in (0, 1]$, $\underset{\sim}{y} + t\underset{\sim}{d} = t\underset{\sim}{x} + (1 - t)\underset{\sim}{y}$, so

$$
\begin{aligned}
g(t) - g(0) &= f(t\underset{\sim}{x} + (1 - t)\underset{\sim}{y}) - f(\underset{\sim}{y}) && \text{(definition of } g) \\
&\leq t f(\underset{\sim}{x}) + (1 - t) f(\underset{\sim}{y}) - f(\underset{\sim}{y}) && (f \text{ is convex}) \\
&= t\,(f(\underset{\sim}{x}) - f(\underset{\sim}{y})), && \text{(collect terms)}
\end{aligned}
$$

and dividing by $t > 0$, $\dfrac{g(t) - g(0)}{t} \leq f(\underset{\sim}{x}) - f(\underset{\sim}{y})$. As $t \to 0$ from the right, the left side tends to $g'(0)$ (g is differentiable at 0, so the one-sided limit equals the two-sided derivative), and a limit of numbers that are all at most $f(\underset{\sim}{x}) - f(\underset{\sim}{y})$ is at most $f(\underset{\sim}{x}) - f(\underset{\sim}{y})$. So $\left(\frac{\partial}{\partial \underset{\sim}{x}} f(\underset{\sim}{y})\right)^{\top} (\underset{\sim}{x} - \underset{\sim}{y}) \leq f(\underset{\sim}{x}) - f(\underset{\sim}{y})$, which rearranges to Equation 1.

If. Let $\underset{\sim}{x}, \underset{\sim}{y} \in \mathbb{R}^p$, $t \in [0, 1]$, $\underset{\sim}{z} \overset{\text{def}}{=} t\underset{\sim}{x} + (1 - t)\underset{\sim}{y}$, and $\underset{\sim}{g} \overset{\text{def}}{=} \frac{\partial}{\partial \underset{\sim}{x}} f(\underset{\sim}{z})$. Applying Equation 1 at

the point $\underset{\sim}{z}$, once toward $\underset{\sim}{x}$ and once toward $\underset{\sim}{y}$,

$$f(\underset{\sim}{x}) \geq f(\underset{\sim}{z}) + \underset{\sim}{g}^{\top}(\underset{\sim}{x} - \underset{\sim}{z}), \qquad f(\underset{\sim}{y}) \geq f(\underset{\sim}{z}) + \underset{\sim}{g}^{\top}(\underset{\sim}{y} - \underset{\sim}{z}).$$

Multiplying the first by $t \geq 0$ and the second by $1 - t \geq 0$ and adding,

$$\begin{aligned}
t f(\underset{\sim}{x}) + (1-t) f(\underset{\sim}{y}) &\geq t(f(\underset{\sim}{z}) + \underset{\sim}{g}^{\top}(\underset{\sim}{x} - \underset{\sim}{z})) + (1-t)(f(\underset{\sim}{z}) + \underset{\sim}{g}^{\top}(\underset{\sim}{y} - \underset{\sim}{z})) && \text{(multiply by } t, 1-t \geq 0 \text{ and add)} \\
&= (t + (1-t))\, f(\underset{\sim}{z}) + t\, \underset{\sim}{g}^{\top}(\underset{\sim}{x} - \underset{\sim}{z}) + (1-t)\, \underset{\sim}{g}^{\top}(\underset{\sim}{y} - \underset{\sim}{z}) && \text{(distribute } t \text{ and } 1-t) \\
&= f(\underset{\sim}{z}) + t\, \underset{\sim}{g}^{\top}(\underset{\sim}{x} - \underset{\sim}{z}) + (1-t)\, \underset{\sim}{g}^{\top}(\underset{\sim}{y} - \underset{\sim}{z}) && (t + (1-t) = 1) \\
&= f(\underset{\sim}{z}) + \underset{\sim}{g}^{\top}(t(\underset{\sim}{x} - \underset{\sim}{z}) + (1-t)(\underset{\sim}{y} - \underset{\sim}{z})) && \text{(factor out } \underset{\sim}{g}^{\top}\text{; matrix products dis} \\
&= f(\underset{\sim}{z}) + \underset{\sim}{g}^{\top}(t\underset{\sim}{x} + (1-t)\underset{\sim}{y} - \underset{\sim}{z}) && \text{(collect terms)} \\
&= f(\underset{\sim}{z}). && \text{(definition of } \underset{\sim}{z})
\end{aligned}$$

That inequality is the defining inequality of a convex function. If Equation 1 is strict for distinct points, and $\underset{\sim}{x} \neq \underset{\sim}{y}$ and $t \in (0,1)$, then:

- $\underset{\sim}{z}$ differs from both $\underset{\sim}{x}$ and y, since $\underset{\sim}{x} - \underset{\sim}{z} = (1-t)(\underset{\sim}{x} - \underset{\sim}{y}) \neq \underset{\sim}{0}$ and $\underset{\sim}{y} - \underset{\sim}{z} = t(\underset{\sim}{y} - \underset{\sim}{x}) \neq \underset{\sim}{0}$;
- so both inequalities at $\underset{\sim}{z}$ are strict;
- both multipliers t and $1-t$ are positive, so the first step of the display is strict, and f is strictly convex.

□

Example 0.10 (Tangent lines under a parabola)**.** For $f(x) = x^2$, the right side of Equation 1 is $y^2 + 2y(x-y) = 2xy - y^2$, and

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

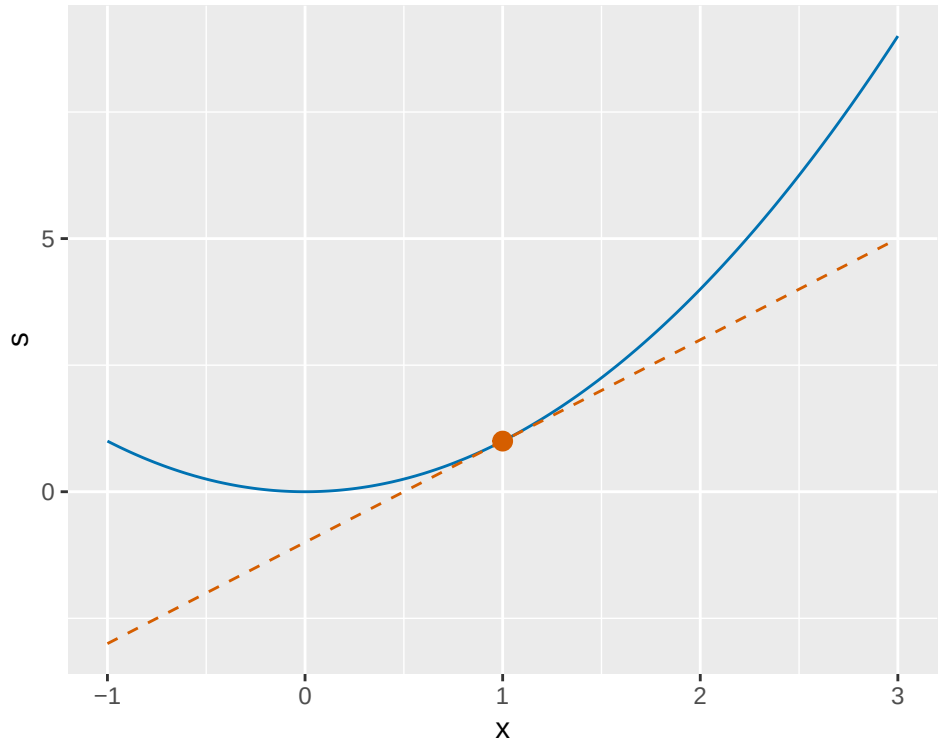
with equality only at $x = y$, confirming that x^2 is strictly convex. For $g(x) = -x^2$ the inequality fails: at $y = 0$ the tangent line is $s = 0$, and $g(1) = -1 < 0$.

Theorem 0.5 (Second-derivative test for convexity)**.** *Let $f : \mathbb{R}^p \to \mathbb{R}$ have first and second partial derivatives that are continuous on $\mathbb{R}^p$ (continuity).*

1. *If $\mathbf{H}_f(\underset{\sim}{x})$ is positive semidefinite for every $\underset{\sim}{x}$, then f is convex.*
2. *If $\mathbf{H}_f(\underset{\sim}{x})$ is positive definite for every $\underset{\sim}{x}$, then f is strictly convex (Definition 0.3).*

R

```r
tl_x_ft <- seq(-1, 3, length.out = 201)
tl_curve_ft <- tibble::tibble(x = tl_x_ft, s = tl_x_ft^2)
tl_line_ft <- tibble::tibble(x = tl_x_ft, s = 2 * tl_x_ft - 1)
tl_point_ft <- tibble::tibble(x = 1, s = 1)
tl_aes_ft <- ggplot2::aes(x, s)
ggplot2::ggplot(tl_curve_ft, tl_aes_ft) +
  ggplot2::geom_line(colour = "#0072B2") +
  ggplot2::geom_line(
    data = tl_line_ft, colour = "#D55E00", linetype = "dashed"
  ) +
  ggplot2::geom_point(data = tl_point_ft, colour = "#D55E00", size = 3) +
  ggplot2::labs(x = "x", y = "s")
```



Python

```python
import matplotlib.pyplot as plt
import numpy as np

tl_x_ft = np.linspace(-1, 3, 201)
fig, ax = plt.subplots(figsize=(5, 4), layout="constrained")
ax.plot(tl_x_ft, tl_x_ft**2, color="#0072B2", label="f(x) = x^2")
ax.plot(tl_x_ft, 2 * tl_x_ft - 1, "--", color="#D55E00",
        label="tangent line at 1")
ax.plot(1, 1, "o", color="#D55E00", markersize=8)
ax.set_xlabel("x")
ax.set_ylabel("s")
ax.legend(loc="upper center")
plt.show()
```

Proof

Proof. Let $\underset{\sim}{x}, \underset{\sim}{y} \in \mathbb{R}^p$. By Theorem 6 in Critical Points and Optimality Conditions with $\underset{\sim}{z} = \underset{\sim}{y}$ and $\underset{\sim}{h} = \underset{\sim}{x} - \underset{\sim}{y}$, there is a $\tau \in (0,1)$ such that, with $\underset{\sim}{w} \stackrel{\text{def}}{=} \underset{\sim}{y} + \tau(\underset{\sim}{x} - \underset{\sim}{y})$,

$$
\begin{aligned}
f(\underset{\sim}{x}) &= f(\underset{\sim}{y}) + \left(\frac{\partial}{\partial \underset{\sim}{x}} f(\underset{\sim}{y})\right)^\top (\underset{\sim}{x} - \underset{\sim}{y}) + \frac{1}{2} (\underset{\sim}{x} - \underset{\sim}{y})^\top \mathbf{H}_f(\underset{\sim}{w}) (\underset{\sim}{x} - \underset{\sim}{y}) && \text{(Theorem 6 in Critical Points and Optimality} \\
&\geq f(\underset{\sim}{y}) + \left(\frac{\partial}{\partial \underset{\sim}{x}} f(\underset{\sim}{y})\right)^\top (\underset{\sim}{x} - \underset{\sim}{y}). && (\mathbf{H}_f(\underset{\sim}{w}) \text{ is positive semidefinite})
\end{aligned}
$$

This inequality is Equation 1, so f is convex by Theorem 0.4. Under the hypothesis of part 2 and with $\underset{\sim}{x} \neq \underset{\sim}{y}$, the quadratic-form term is positive, so the inequality is strict, and f is strictly convex by the last sentence of Theorem 0.4. □

Example 0.11 (Using the Hessian to show convexity)**.**

- For a symmetric positive semidefinite $\mathbf{A}$, $f(\underset{\sim}{x}) = \underset{\sim}{x}^\top \mathbf{A} \underset{\sim}{x}$ has Hessian $2\mathbf{A}$ (Hessian of a quadratic form), which is positive semidefinite because $\underset{\sim}{h}^\top (2\mathbf{A}) \underset{\sim}{h} = 2\, \underset{\sim}{h}^\top \mathbf{A} \underset{\sim}{h} \geq 0$; so f is convex.
- $f(\underset{\sim}{x}) = x_1^2 + x_2^2 - 2x_1$ has Hessian $2\mathbf{I}_2$ (Example 15 in Critical Points and Optimality Conditions), which is positive definite, so f is strictly convex.
- $f(\underset{\sim}{x}) = e^{2x_1 + x_2} - x_1$ has Hessian $e^u \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$ with $u = 2x_1 + x_2$ (Example), and

$$
\begin{aligned}
\underset{\sim}{h}^\top \mathbf{H}_f(\underset{\sim}{x})\, \underset{\sim}{h} &= e^u (4h_1^2 + 4h_1 h_2 + h_2^2) \\
&= e^u (2h_1 + h_2)^2 \\
&\geq 0,
\end{aligned}
$$

  so f is convex. The Hessian is not positive definite ($\underset{\sim}{h} = (1, -2)^\top$ gives 0), so part 2 does not apply, and in fact f is not strictly convex: along $\underset{\sim}{x} = t\,(1, -2)^\top$,

$$
\begin{aligned}
u &= 2t - 2t \\
&= 0
\end{aligned}
$$

  and $f = 1 - t$ is linear in t, so the convexity inequality holds with equality there.

Example 0.12 (Strictly convex with a singular Hessian)**.** Part 2 is sufficient but not necessary. $f(x) = x^4$ has

$$\begin{aligned} f''(0) &= 12 \cdot 0^2 \\ &= 0, \end{aligned}$$

so $\mathbf{H}_f(0) = [0]$ is singular and not positive definite. Yet f is strictly convex, by the strict form of Equation 1: for $x \neq y$,

$$\begin{aligned} x^4 - y^4 - 4y^3(x-y) &= (x-y)(x^3 + x^2y + xy^2 + y^3) - 4y^3(x-y) && \text{(factor } x^4 - y^4\text{)} \\ &= (x-y)(x^3 + x^2y + xy^2 - 3y^3) && \text{(collect the } y^3 \text{ terms)} \\ &= (x-y)^2(x^2 + 2xy + 3y^2) && ((x-y)(x^2 + 2xy + 3y^2) = x^3 + x^2y + x \\ &= (x-y)^2((x+y)^2 + 2y^2), && \text{(complete the square)} \end{aligned}$$

which is positive: $(x-y)^2 > 0$, and $(x+y)^2 + 2y^2 = 0$ would need $y = 0$ and then

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

Theorem 0.6 (A convex function has a positive semidefinite Hessian)**.** *Let $f : \mathbb{R}^p \to \mathbb{R}$ have first and second partial derivatives that are continuous on $\mathbb{R}^p$. If f is convex, then the Hessian $\mathbf{H}_f(\underset{\sim}{x})$ is positive semidefinite for every $\underset{\sim}{x} \in \mathbb{R}^p$.*

Proof

Proof. Fix $\underset{\sim}{x}, \underset{\sim}{v} \in \mathbb{R}^p$. Let $h > 0$. By Theorem 6 in Critical Points and Optimality Conditions applied with $\underset{\sim}{z} = \underset{\sim}{x}$ and $\underset{\sim}{h} = h\underset{\sim}{v}$, and again with $\underset{\sim}{h} = -h\underset{\sim}{v}$, there are $\tau_+, \tau_- \in (0,1)$ with

$$f(\underset{\sim}{x} + h\underset{\sim}{v}) = f(\underset{\sim}{x}) + h\left(\frac{\partial}{\partial \underset{\sim}{x}} f(\underset{\sim}{x})\right)^\top \underset{\sim}{v} + \frac{h^2}{2}\, \underset{\sim}{v}^\top \mathbf{H}_f(\underset{\sim}{w}_+)\, \underset{\sim}{v} \quad \text{(Theorem 6 in Critical Points and Optimality Condi}$$

$$f(\underset{\sim}{x} - h\underset{\sim}{v}) = f(\underset{\sim}{x}) - h\left(\frac{\partial}{\partial \underset{\sim}{x}} f(\underset{\sim}{x})\right)^\top \underset{\sim}{v} + \frac{h^2}{2}\, \underset{\sim}{v}^\top \mathbf{H}_f(\underset{\sim}{w}_-)\, \underset{\sim}{v}, \quad \text{(Theorem 6 in Critical Points and Optimality Condi}$$

where $\underset{\sim}{w}_\pm = \underset{\sim}{x} \pm \tau_\pm h\underset{\sim}{v}$. Adding the two equations cancels the first-order summands:

$$f(\underset{\sim}{x} + h\underset{\sim}{v}) + f(\underset{\sim}{x} - h\underset{\sim}{v}) - 2f(\underset{\sim}{x}) = \frac{h^2}{2}\, \underset{\sim}{v}^\top \big(\mathbf{H}_f(\underset{\sim}{w}_+) + \mathbf{H}_f(\underset{\sim}{w}_-)\big)\underset{\sim}{v}.$$

Since f is convex and $\underset{\sim}{x} = \frac{1}{2}(\underset{\sim}{x} + h\underset{\sim}{v}) + \frac{1}{2}(\underset{\sim}{x} - h\underset{\sim}{v})$,

$$f(\underset{\sim}{x}) \leq \frac{1}{2} f(\underset{\sim}{x} + h\underset{\sim}{v}) + \frac{1}{2} f(\underset{\sim}{x} - h\underset{\sim}{v}),$$

so the left side of the previous display is nonnegative. Dividing by $h^2/2$ gives

$$\underset{\sim}{v}^\top \big(\mathbf{H}_f(\underset{\sim}{w}_+) + \mathbf{H}_f(\underset{\sim}{w}_-)\big)\underset{\sim}{v} \geq 0 \qquad \text{for every } h > 0.$$

As $h \to 0$, both $\underset{\sim}{w}_+$ and $\underset{\sim}{w}_-$ approach $\underset{\sim}{x}$, and the entries of $\mathbf{H}_f$ are continuous, so the left side approaches $2\,\underset{\sim}{v}^\top \mathbf{H}_f(\underset{\sim}{x})\,\underset{\sim}{v}$. Therefore $\underset{\sim}{v}^\top \mathbf{H}_f(\underset{\sim}{x})\,\underset{\sim}{v} \geq 0$ for every $\underset{\sim}{v}$. □

Corollary 0.1 (Stationary points of a convex function are global minimizers)**.** *If f is as in Theorem 0.4 and convex, then every stationary point $\underset{\sim}{x}^*$ of f (Definition 4 in Critical Points and Optimality Conditions) is a global minimizer.*

Proof

Proof. For every $\underset{\sim}{x}$,

$$\begin{aligned} f(\underset{\sim}{x}) &\geq f(\underset{\sim}{x}^*) + \left(\frac{\partial}{\partial \underset{\sim}{x}} f(\underset{\sim}{x}^*)\right)^\top (\underset{\sim}{x} - \underset{\sim}{x}^*) && \text{(Equation 1 with } \underset{\sim}{y} = \underset{\sim}{x}^*) \\ &= f(\underset{\sim}{x}^*) + \underset{\sim}{0}^\top (\underset{\sim}{x} - \underset{\sim}{x}^*) && (\underset{\sim}{x}^* \text{ is stationary}) \\ &= f(\underset{\sim}{x}^*). && (\underset{\sim}{0} \cdot \underset{\sim}{v} = 0) \end{aligned}$$

□

Example 0.13 (A convex function, and a non-convex one)**.**

- $f(\underset{\sim}{x}) = x_1^2 + x_2^2 - 2x_1$ is convex by Theorem 0.5, since its Hessian is $2\mathbf{I}_2$ (Example 15 in Critical Points and Optimality Conditions), so its stationary point $(1, 0)^\top$ (Example 7 in Critical Points and Optimality Conditions) is a global minimizer, as completing the square showed in Example 8 in Critical Points and Optimality Conditions.
- Without convexity, a stationary point need not be a minimizer at all: $x_1^2 - x_2^2$ has the stationary point $\underset{\sim}{0}$, which is not a minimizer (Example 8 in Critical Points and Optimality Conditions).
- The corollary says nothing about whether a minimizer exists: $e^{2x_1 + x_2} - x_1$ is convex (Example 0.11) but has no stationary point, and so no local minimizer (Example 8 in Critical Points and Optimality Conditions).

[1] H. Zhou, "Optimization and multivariate calculus." Lecture notes for Biostat 216, Mathematical Methods for Biostatistics, University of California, Los Angeles, 2024. Accessed: Oct. 05, 2026. [Online]. Available: https://ucla-biostat-216.github.io/2024fall/slides/13-optim/13-optim.html

[2] Wikipedia contributors, "Convex set — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Convex_set

[3] Wikipedia contributors, "Epigraph (mathematics) — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Epigraph_(mathematics)

[4] Wikipedia contributors, "Tangent space — Wikipedia, the free encyclopedia." 2026. Available: https://en.wikipedia.org/wiki/Tangent_space