L02 Convexity and Optimality

CS 169 / CS 268
Introduction to Optimization

Ioannis Panageas

Fall 2026

Local and global minimizers

Let $f : S \to \mathbb{R}$, where $S \subseteq \mathbb{R}$, and let $x^* \in S$.

Global minimizer. No feasible point has a smaller value:

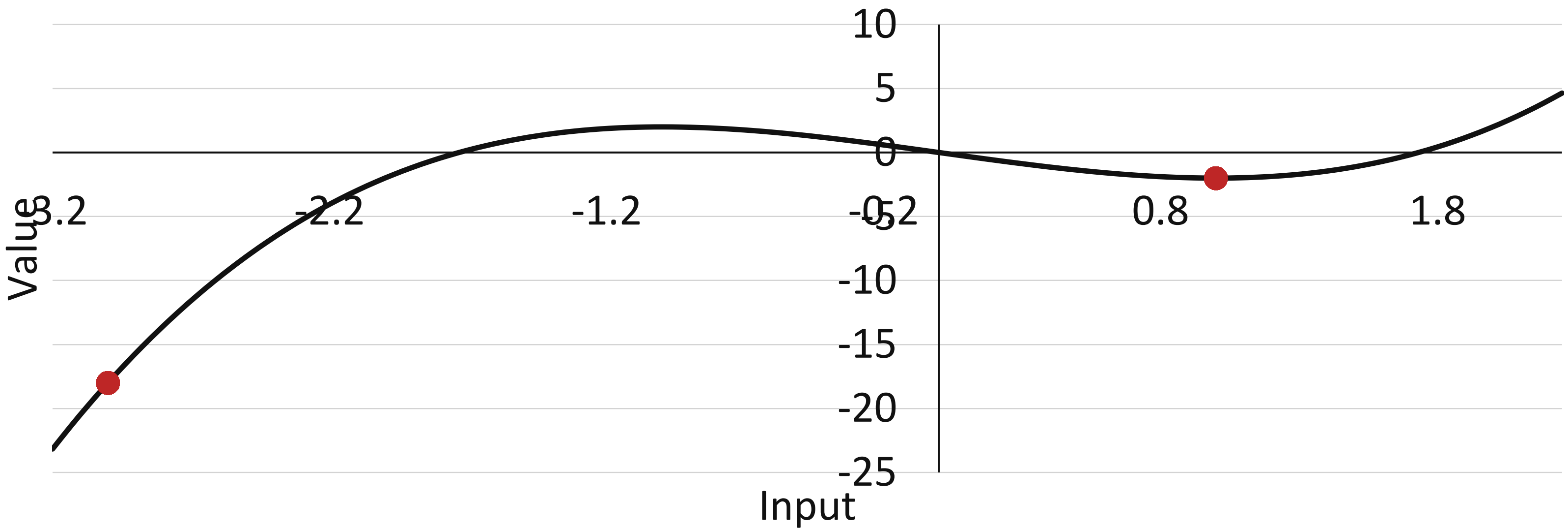
$$f(x^*) \leq f(y) \qquad \text{for every } y \in S.$$

Local minimizer. The same is true in some neighborhood: for some $\varepsilon > 0$,

$$f(x^*) \leq f(y) \qquad \text{if } y \in S \text{ and } |y - x^*| < \varepsilon.$$

A local minimum need not be global

$$f(x) = x^3 - 3x, \qquad x \in \mathbb{R}.$$



$x = 1$ is a local minimizer, but $f(-3) = -18 < f(1) = -2$.

A necessary condition in one variable

Stationarity. If f is differentiable on an open interval and x^* is a local minimizer, then

$$f'(x^*) = 0.$$

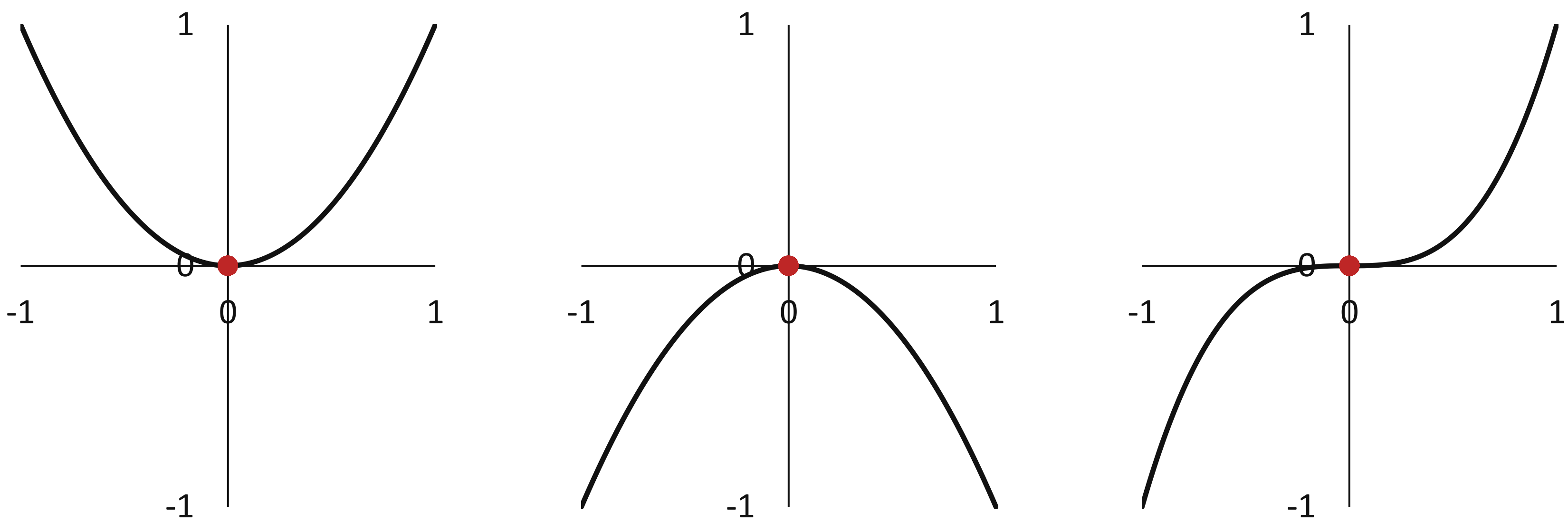
For small $h > 0$, the difference quotient is nonnegative.
For small $h < 0$, it is nonpositive.

$$\frac{f(x^* + h) - f(x^*)}{h}$$

Both one-sided limits equal $f'(x^*)$, so it must be 0.

Stationary does not mean minimum

$$f_1(x) = x^2, \qquad f_2(x) = -x^2, \qquad f_3(x) = x^3.$$



Minimum Maximum Neither

All three derivatives vanish at 0.
Only x^2 has a local minimum there.

Vectors and distance

A point in $\mathbb{R}^n$ has n real coordinates:

$$x = (x_1, \ldots, x_n), \qquad y = (y_1, \ldots, y_n).$$

Euclidean distance. The distance between two points is

$$\|x - y\|_2 = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

For $n = 2$, this is the usual distance in the plane.
Local neighborhoods use this distance.

Our running example

$$f(x_1, x_2) = (x_1 - 1)^2 + 2(x_2 + 1)^2.$$

We now choose two coordinates.
Each squared term is nonnegative.

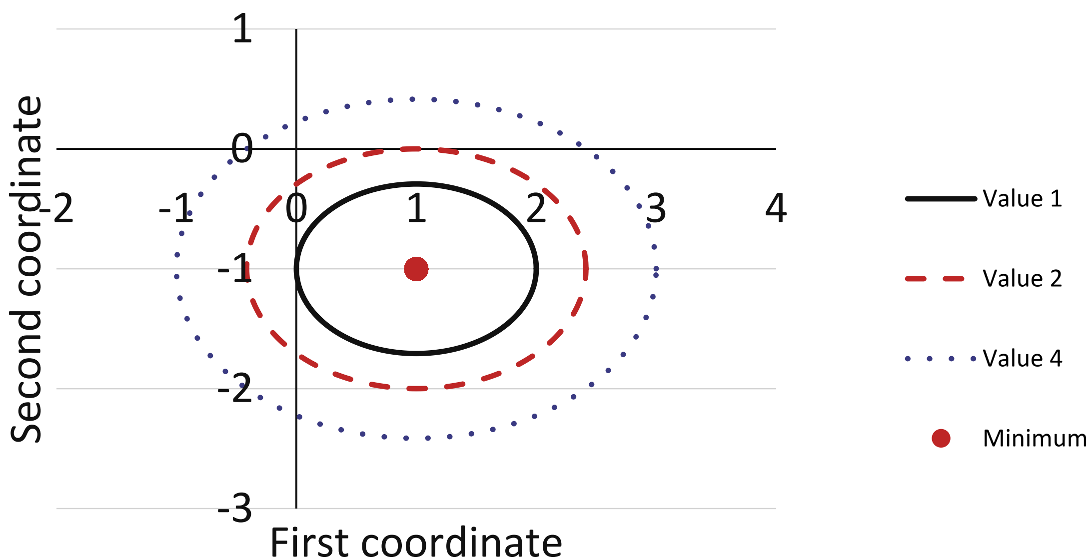
$$f(x_1, x_2) \geq 0, \qquad f(1, -1) = 0.$$

The unique global minimizer is $(1, -1)$.
Both squared terms must vanish to reach 0.

Level sets

Level set. All points with the same function value c:

$$(x_1 - 1)^2 + 2(x_2 + 1)^2 = c.$$

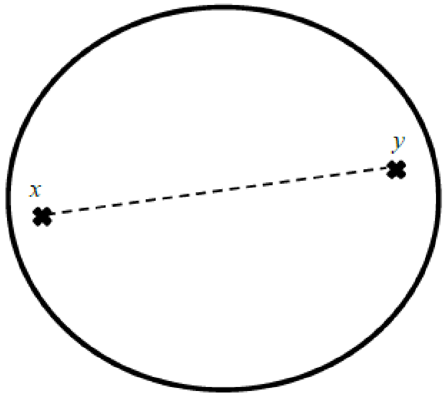


Level 0 consists only of the center $(1, -1)$.

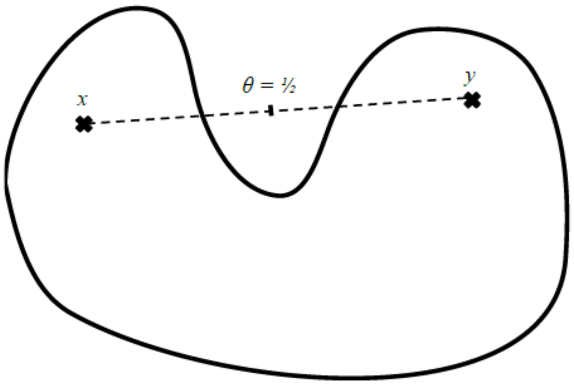
Convex sets in several variables

Definition (Convex set).
A set $S \subseteq \mathbb{R}^n$ is convex if, for every $x, y \in S$ and $t \in [0,1]$,

$$(1-t)x + ty \in S.$$



Convex



Nonconvex

A single segment leaving the set disproves convexity.

Convex functions in several variables

Definition (Convex function).
Let $C \subseteq \mathbb{R}^n$ be convex. A function $f : C \to \mathbb{R}$ is convex if, for every $x, y \in C$ and $t \in [0,1]$,

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

Along every segment, this is ordinary one-variable convexity:

$$g(t) = f((1-t)x + ty), \qquad 0 \leq t \leq 1.$$

Partial derivatives

Partial derivative. Vary one coordinate and keep all other coordinates fixed.

$$f(x_1, x_2) = (x_1 - 1)^2 + 2(x_2 + 1)^2.$$

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = 2(x_1 - 1).$$

$$\frac{\partial f}{\partial x_2}(x_1, x_2) = 4(x_2 + 1).$$

The gradient

Gradient. Collect the partial derivatives into one column vector. Assume f is differentiable on an open domain.

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix}. \qquad \nabla f(2,0) = \begin{pmatrix} 2 \\ 4 \end{pmatrix}.$$

The gradient has the same number of coordinates as x.

Change along a direction

Dot product. For vectors $u, v \in \mathbb{R}^n$,

$$u^\mathsf{T} v = \sum_{i=1}^{n} u_i v_i.$$

For differentiable f, the rate along a direction d is

$$\frac{d}{dt} f(x + td)\Big|_{t=0} = \nabla f(x)^\mathsf{T} d.$$

$$x = (2, 0), \quad d = (-1, 0): \qquad \nabla f(x)^\mathsf{T} d = -2.$$

A sufficiently small positive step in this direction decreases f.

Interior minima are stationary

If f is differentiable on an open domain and x^* is a local minimizer, then

$$\nabla f(x^*) = 0.$$

The converse still fails. Consider $h(x_1, x_2) = x_1^2 - x_2^2$.

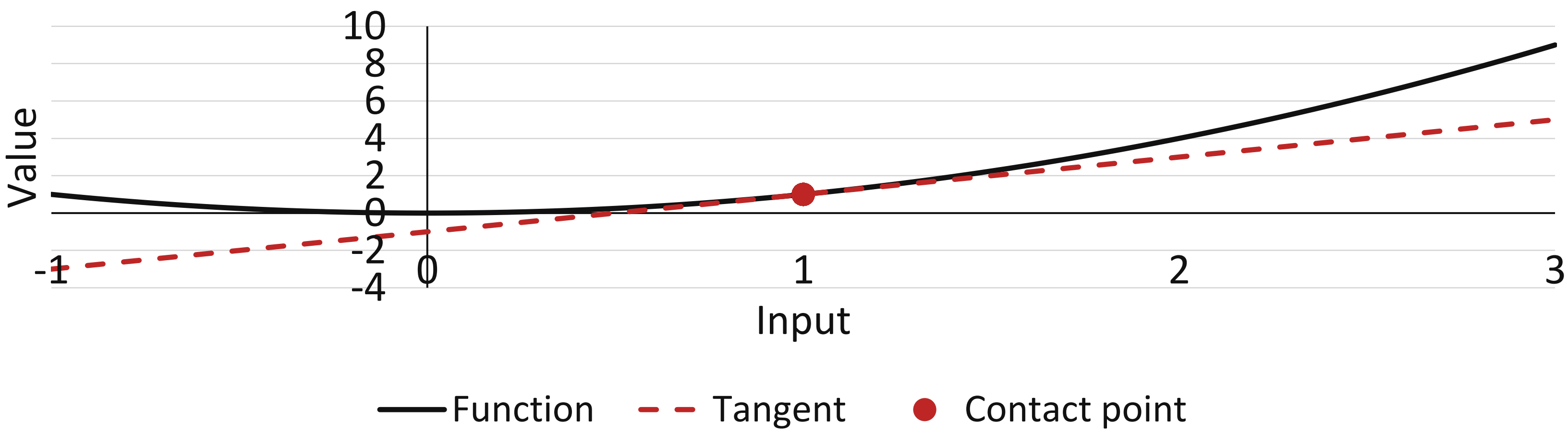
$$\nabla h(0,0) = 0, \qquad h(t,0) = t^2, \quad h(0,t) = -t^2.$$

Arbitrarily close to the origin, values are both larger and smaller.

Convex functions and tangents

For differentiable convex f on an open interval, and any x, y in that interval,

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



The first-order inequality

Let $f : C \to \mathbb{R}$ be differentiable and convex, where $C \subseteq \mathbb{R}^n$ is open and convex. For $x, y \in C$,

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

Proof. Put $g(t) = f(x + t(y - x))$. Convexity gives

$$\frac{g(t) - g(0)}{t} \leq g(1) - g(0), \qquad 0 < t \leq 1.$$

Let $t \downarrow 0$. The chain rule gives $g'(0) = \nabla f(x)^\top (y - x)$.

A certificate of global optimality

Convexity plus stationarity. Let f be differentiable and convex on an open convex domain C. If $x^* \in C$ and $\nabla f(x^*) = 0$, then

$$f(y) \geq f(x^*) + 0^\mathsf{T}(y - x^*) = f(x^*) \qquad (y \in C).$$

For our running example,

$$\nabla f(1, -1) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Every convex local minimum is global

Let f be convex on a convex set C. No differentiability is needed. Suppose x^* is a local minimizer.

If a point $y \in C$ had $f(y) < f(x^*)$, move a small distance toward it:

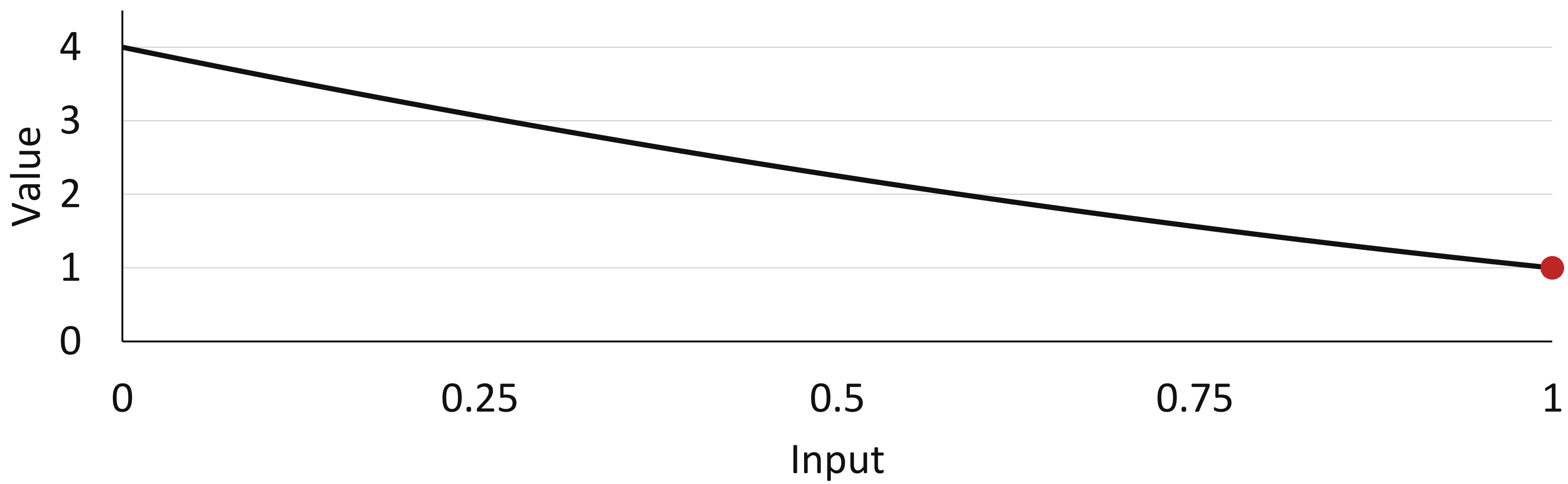
$$x_t = (1-t)x^* + ty \in C, \qquad 0 < t < 1.$$

$$f(x_t) \leq (1-t)f(x^*) + tf(y) < f(x^*).$$

For small t, this contradicts local minimality.

A boundary minimum can have a slope

$$\min_{0 \le x \le 1} (x-2)^2.$$



$$x^* = 1, \qquad f(x^*) = 1, \qquad f'(x^*) = -2 \neq 0.$$

An optimum on a box

$$\min_{(x_1,x_2)\in[0,2]^2} (x_1 - 1)^2 + 2(x_2 + 1)^2.$$

For every feasible point, $x_2 + 1 \geq 1$. Hence

$$f(x_1, x_2) \geq 0 + 2 = 2, \qquad f(1, 0) = 2.$$

$$x^* = (1, 0), \qquad \nabla f(x^*) = \begin{pmatrix} 0 \\ 4 \end{pmatrix}.$$

The negative gradient points outside the feasible box.

Optimality conditions

Interior local minimum. For differentiable f, the gradient must be zero.

Zero gradient. With convexity and differentiability on an open convex domain, this certifies a global minimum.

Convex local minimum. It is global, even without differentiability.

Boundary minimum. The gradient need not vanish.

Questions to check understanding

1. Compute the gradient and find a global minimizer:

$$q(x_1, x_2) = (x_1 - 2)^2 + 3(x_2 - 1)^2.$$

2. Does $\nabla f(x^*) = 0$ imply a local minimum? Give a two-variable counterexample.

3. A convex function has global minimizers u and v. Prove that every point on the segment between them is also a global minimizer.