ORF522 – Linear and Nonlinear Optimization

9. Introduction to nonlinear optimization

Bartolomeo Stellato — Fall 2024

Today's lecture

[Chapter 2-4 and 6, CO][Chapter 1, LSCOMO][Chapter A and B, FCA]

- Nonlinear optimization
- Convex analysis review: sets and functions
- Convex optimization

What if the problem is no longer linear?

Nonlinear optimization

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g_i(x) \leq 0, \quad i = 1, \dots, m \end{array}$$

$x = (x_1, \dots, x_n)$ Variables

$f : \mathbf{R}^n \to \mathbf{R}$ Nonlinear objective function

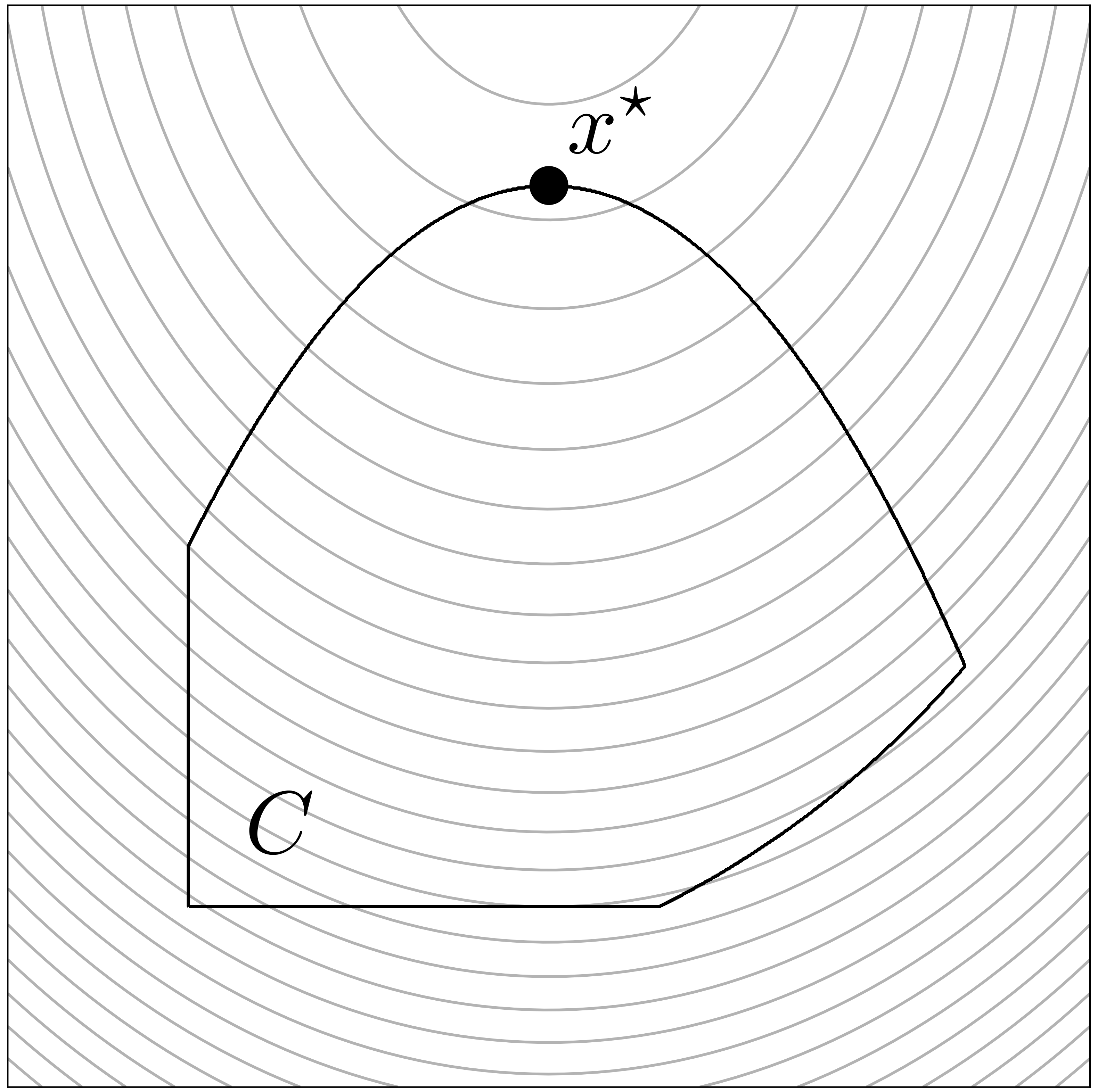
$g_i : \mathbf{R}^n \to \mathbf{R}$ Nonlinear constraints functions

Feasible set

$$C = \{x \mid g_i(x) \leq 0, \quad i = 1, \dots, m\}$$

Small example

$$\begin{aligned} \text{minimize} \quad & 0.5x_1^2 + 0.25x_2^2 \\ \text{subject to} \quad & e^{x_1} - 2 - x_2 \le 0 \\ & (x_1 - 1)^2 + x_2 - 3 \le 0 \\ & x_1 \ge 0 \\ & x_2 \ge 1 \end{aligned}$$



Contour plot has curves (no longer lines)

Feasible set is no longer a polyhedron

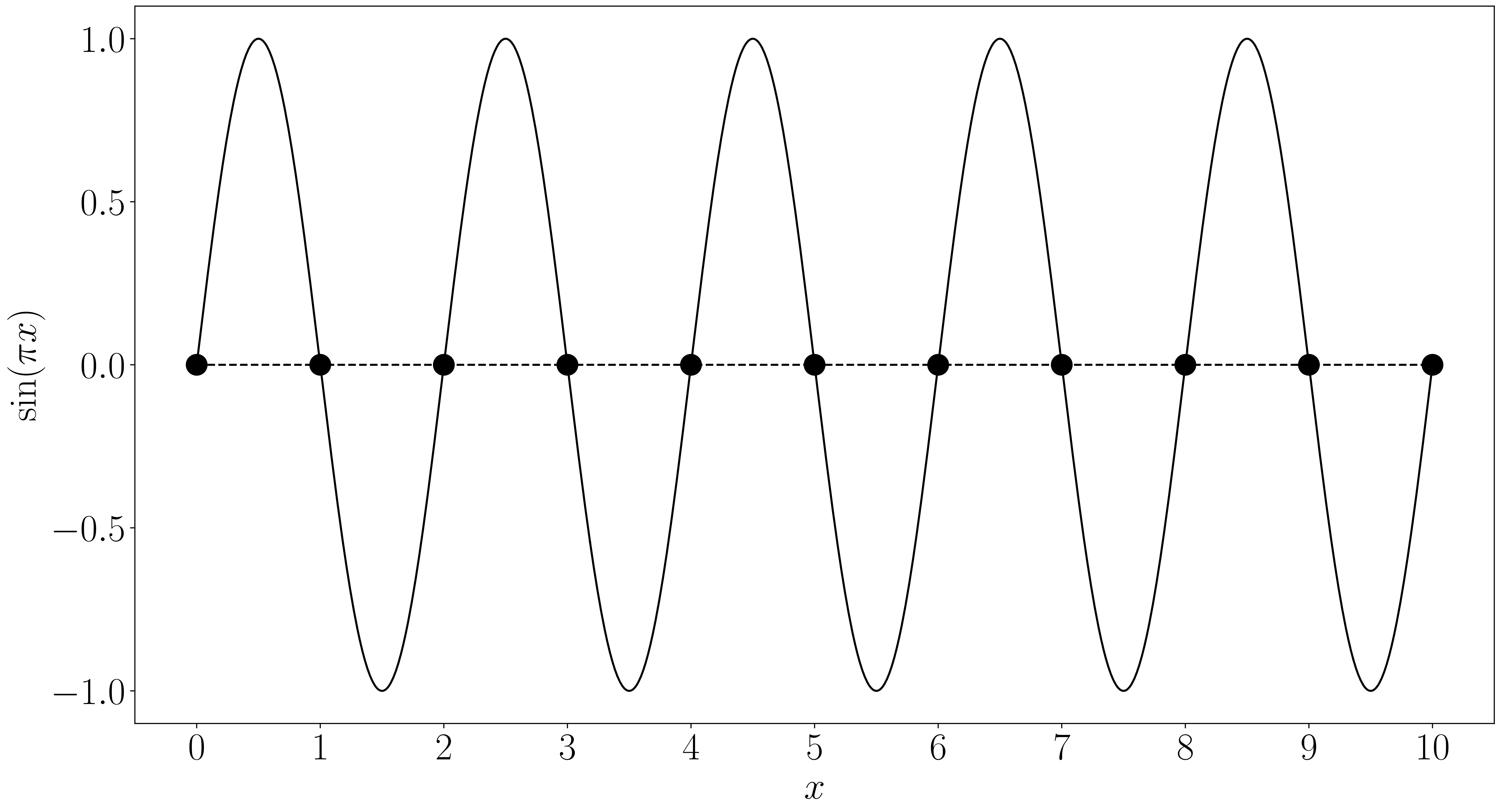
Integer optimization

It's still nonlinear optimization

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in \mathbf{Z} \end{array}$$

$\updownarrow$

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & \sin(\pi x) = 0 \end{array}$$



Portfolio optimization

We have a total of n assets

x_i is fraction of money invested in asset i
p_i is the relative price change of asset i → **Returns** $p^T x$

p random variable: mean μ, covariance Σ

Portfolio optimization

$$\begin{array}{ll} \text{maximize} & \mu^T x - \gamma x^T \Sigma x \\ \text{subject to} & \mathbf{1}^T x = 1 \\ & x \geq 0 \end{array}$$

Expected return → $\mu^T x$

Risk → $x^T \Sigma x$

Risk-aversion parameter → γ

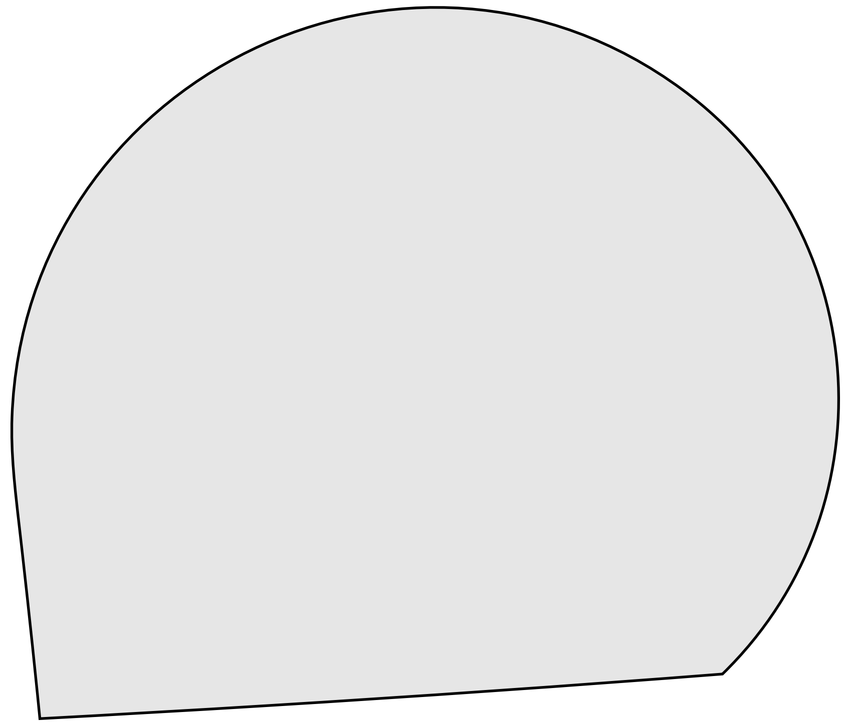
We cannot solve most nonlinear optimization problems

Convex analysis review
Sets

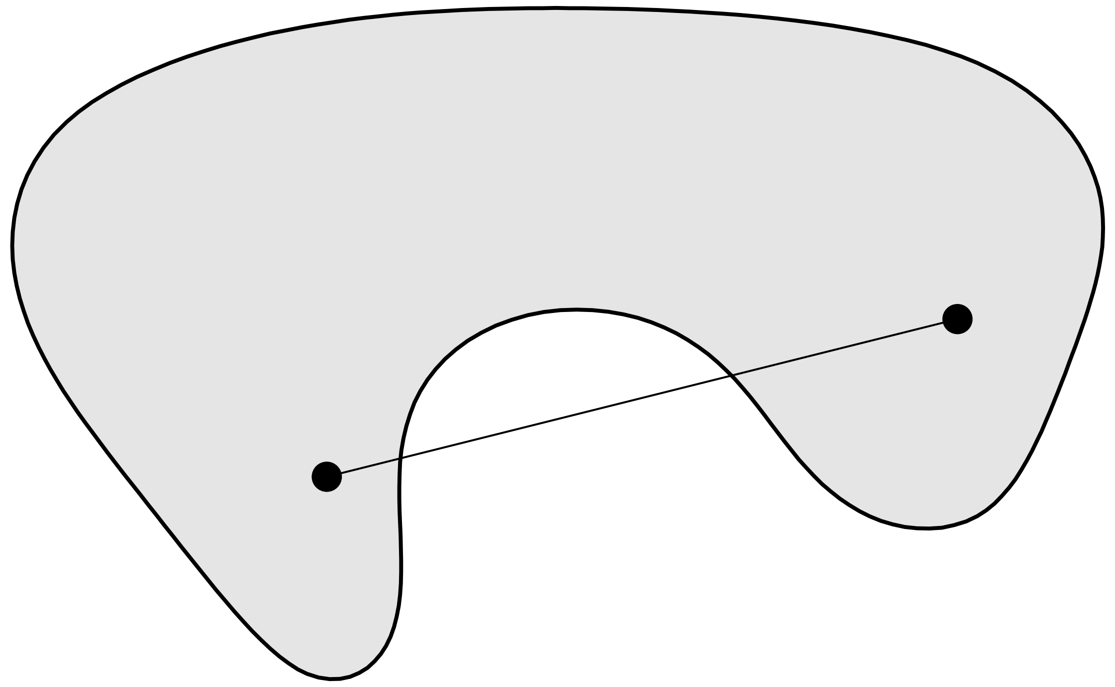
Convex set

Definition

For any $x, y \in C$ and any $\alpha \in [0, 1]$ $\quad \alpha x + (1 - \alpha) y \in C$



Convex



Not convex

Examples

- $\mathbf{R}^n$
- Hyperplanes
- Hyperspheres
- Polyhedra

Examples

- Cardinality constraint $\mathrm{card}(x) \leq k$
- $\mathbf{Z}^n$
- Any disjoint set

intersection of convex sets is convex

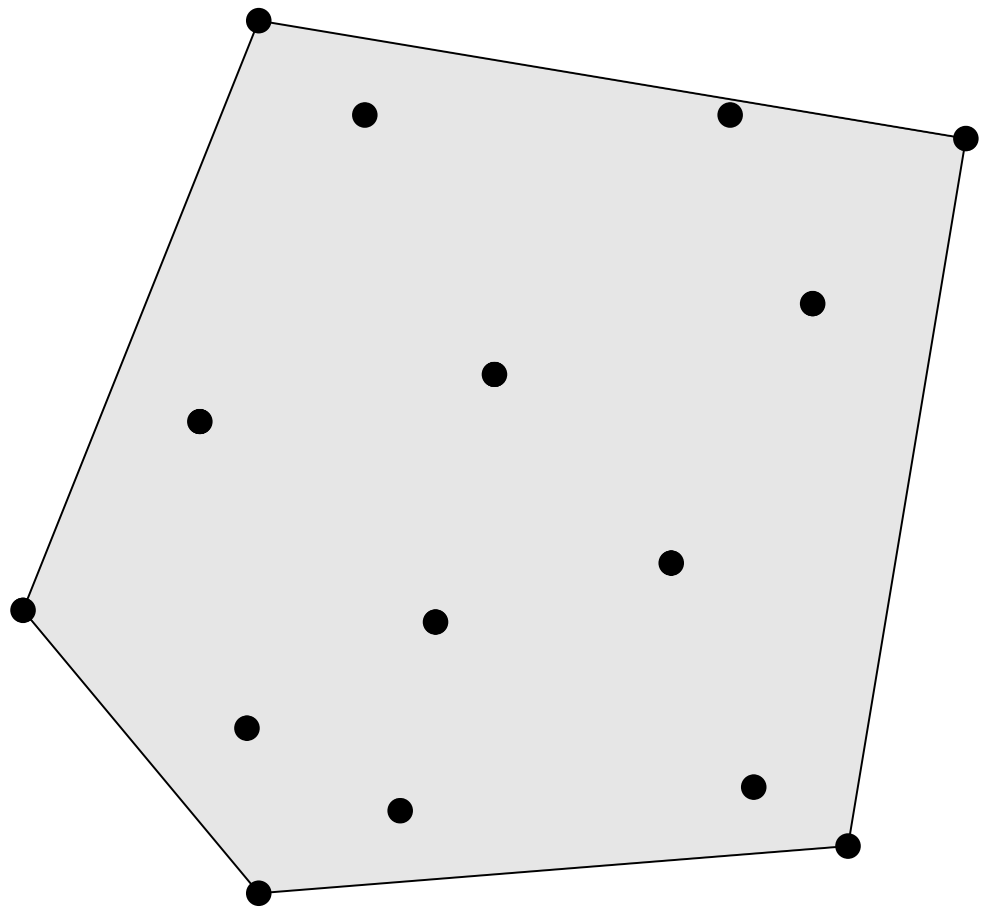
Convex combinations

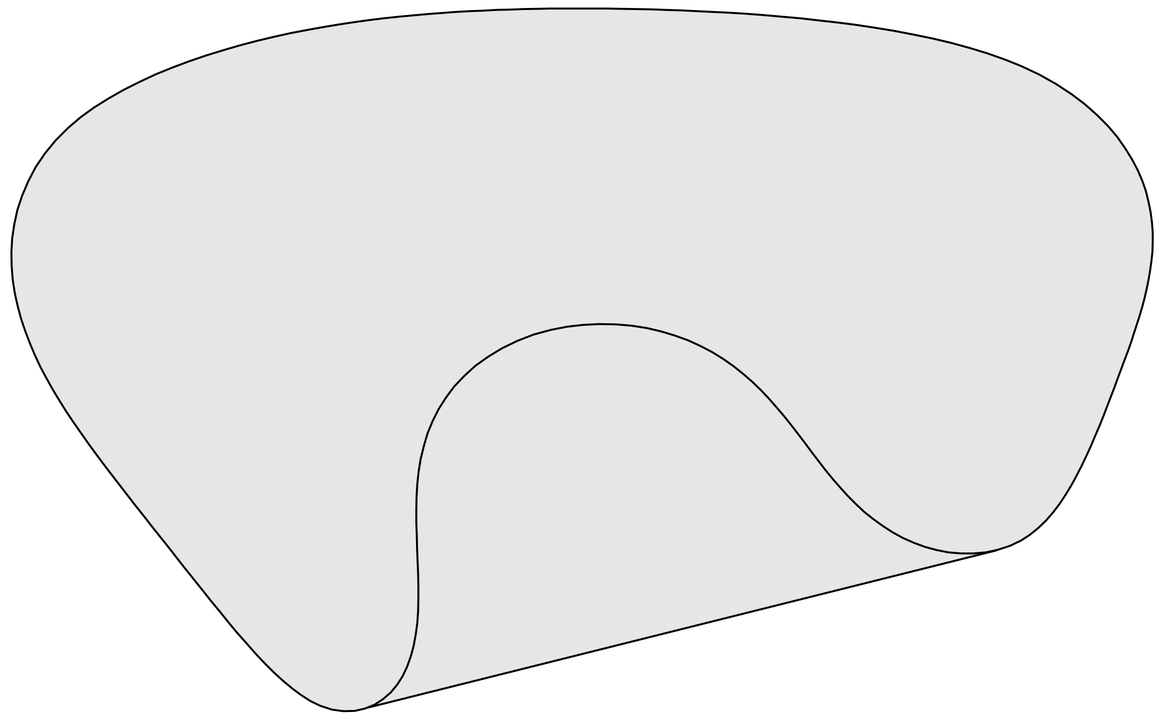
Convex combination

$\alpha_1 x_1 + \cdots + \alpha_k x_k$ for any $x_1, \ldots, x_k$ and $\alpha_1, \ldots, \alpha_k$ such that $\alpha_i \geq 0,\ \sum_{i=1}^k \alpha_i = 1$

Convex hull

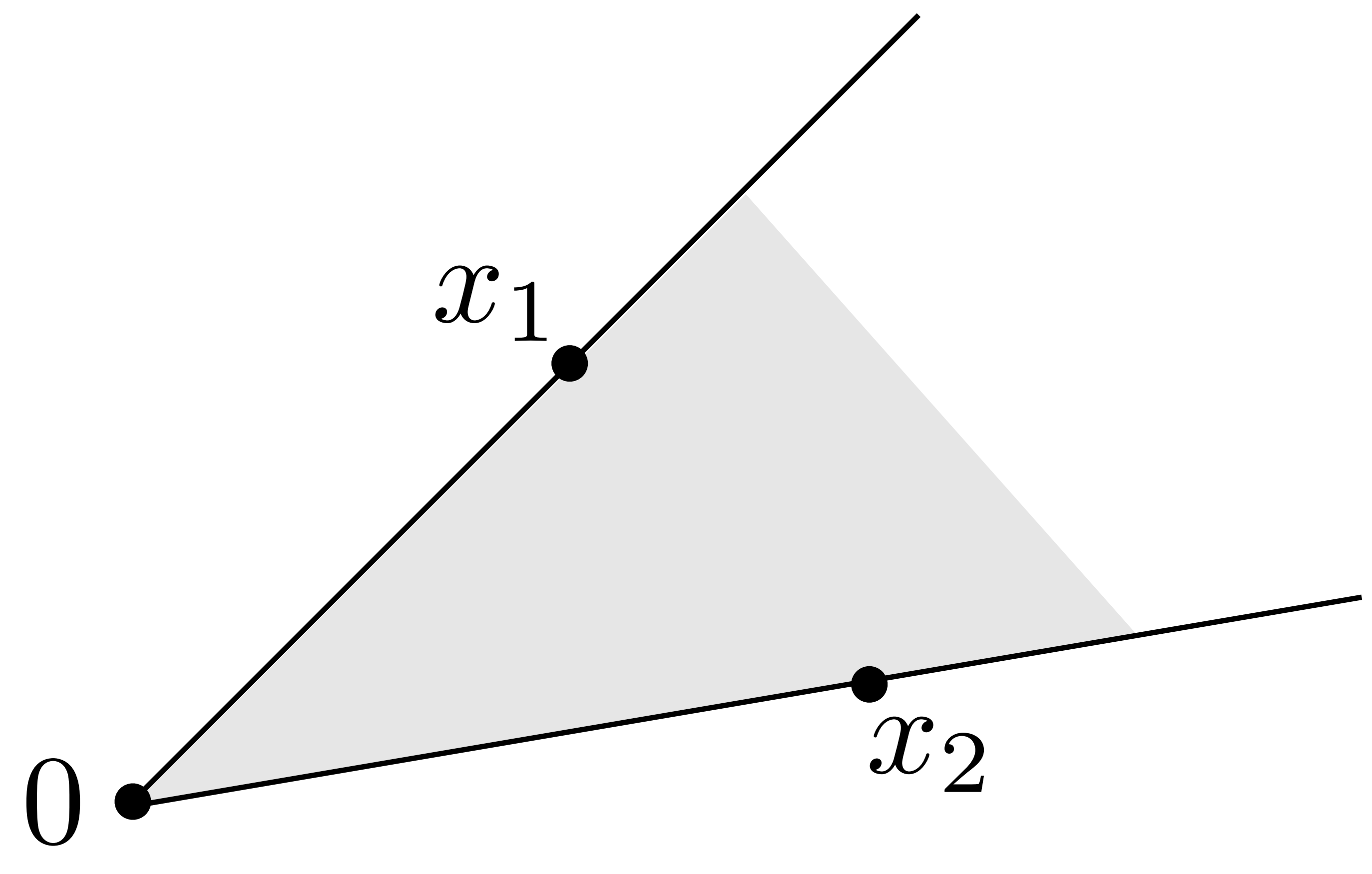
$$\mathbf{conv}\, C = \left\{ \sum_{i=1}^k \alpha_i x_i \mid x_i \in C, \quad \alpha_i \geq 0, \quad i = 1, \ldots, k, \quad \mathbf{1}^T \alpha = 1 \right\}$$





Cones

Cone $x \in C \implies tx \in C$ for all $t \geq 0$



Convex cone (a cone that is also convex)

$x_1, x_2 \in C \implies t_1 x_1 + t_2 x_2 \in C$ for all $t_1, t_2 \geq 0$

Examples

Nonnegative orthant $\mathbf{R}^n_+ = \{x \in \mathbf{R}^n \mid x \geq 0\}$

Norm-cone $\{(x, t) \mid \|x\| \leq t\}$ (if 2-norm, second-order cone)

Positive semidefinite cone $\mathbf{S}^n_+ = \{X \in \mathbf{S}^n \mid z^T X z \geq 0, \quad \text{for all} \quad z \in \mathbf{R}^n\}$

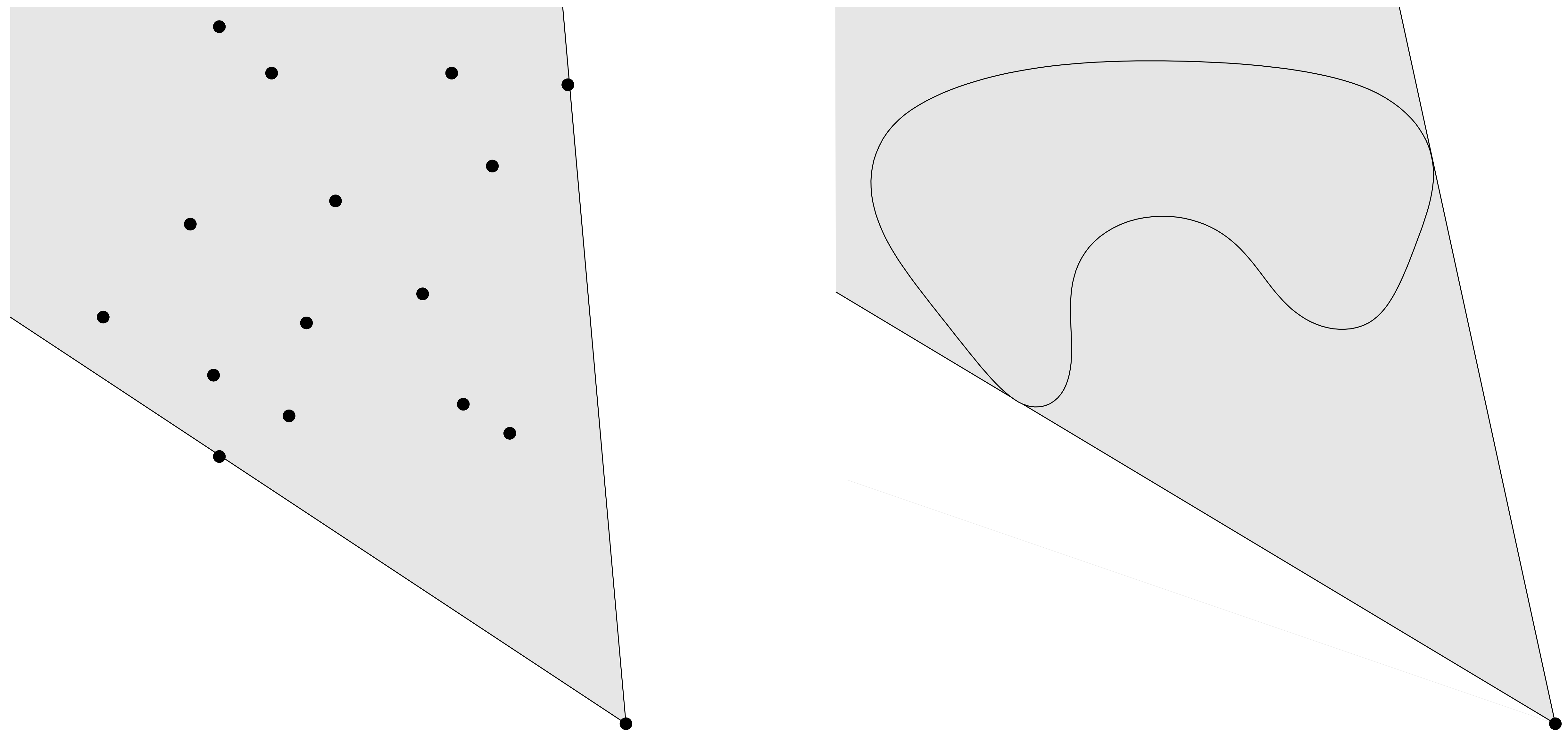
Conic combinations

Conic combination

$\alpha_1 x_1 + \cdots + \alpha_k x_k$ for any $x_1, \ldots, x_k$ and $\alpha_1, \ldots, \alpha_k$ such that $\alpha_i \geq 0$

Conic hull

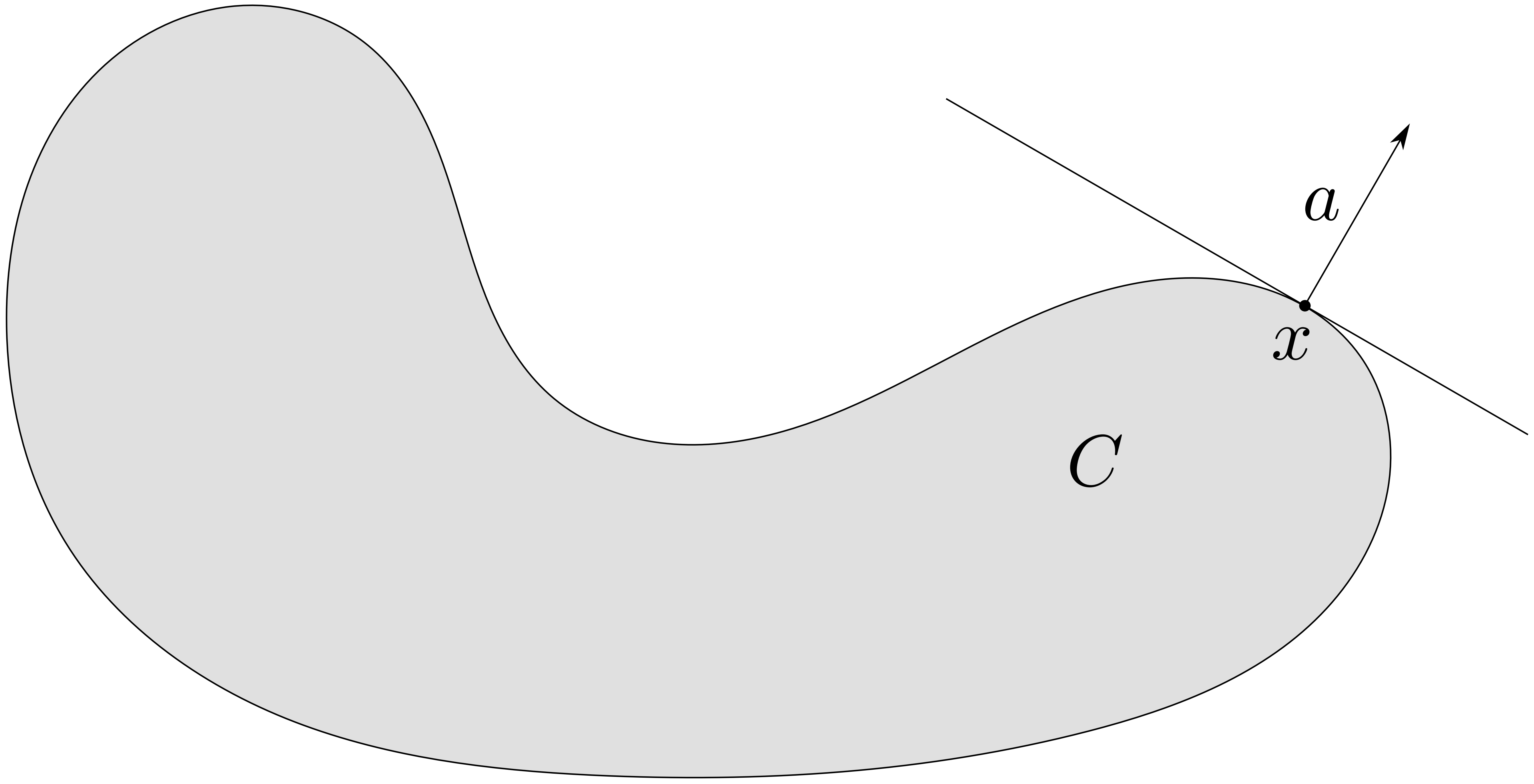
$$\left\{ \sum_{i=1}^{k} \alpha_i x_i \mid x_i \in C, \quad \alpha_i \geq 0, \quad i = 1, \ldots, k \right\}$$



Supporting hyperplanes

Given a set C point x at the boundary of C
a hyperplane $\{z \mid a^T z = a^T x\}$ is a **supporting hyperplane** if

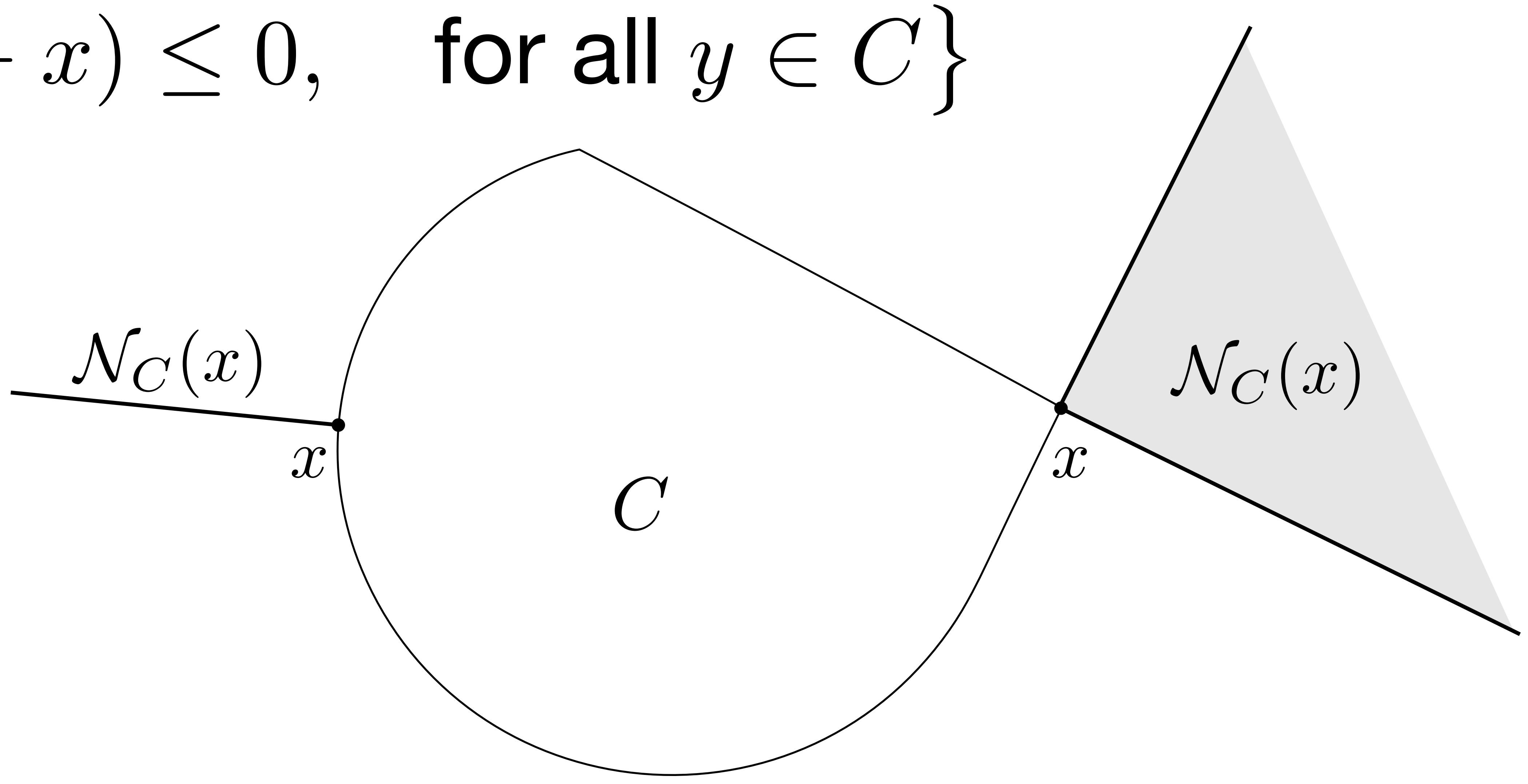
$$a^T (y - x) \leq 0, \quad \forall y \in C$$



Normal cone

For any set C and point $x \in C$, we define

$$\mathcal{N}_C(x) = \{g \mid g^T(y - x) \leq 0, \quad \text{for all } y \in C\}$$



$\mathcal{N}_C(x)$ is **always convex**

Proof For $g_1, g_2 \in \mathcal{N}_C(x)$,

$$(t_1 g_1 + t_2 g_2)^T(y - x) = t_1 g_1^T(y - x) + t_2 g_2^T(y - x) \leq 0 \qquad \text{for all } t_1, t_2 \geq 0$$

How does it relate to supporting hyperplanes?

Convex analysis review

Functions

Convex functions

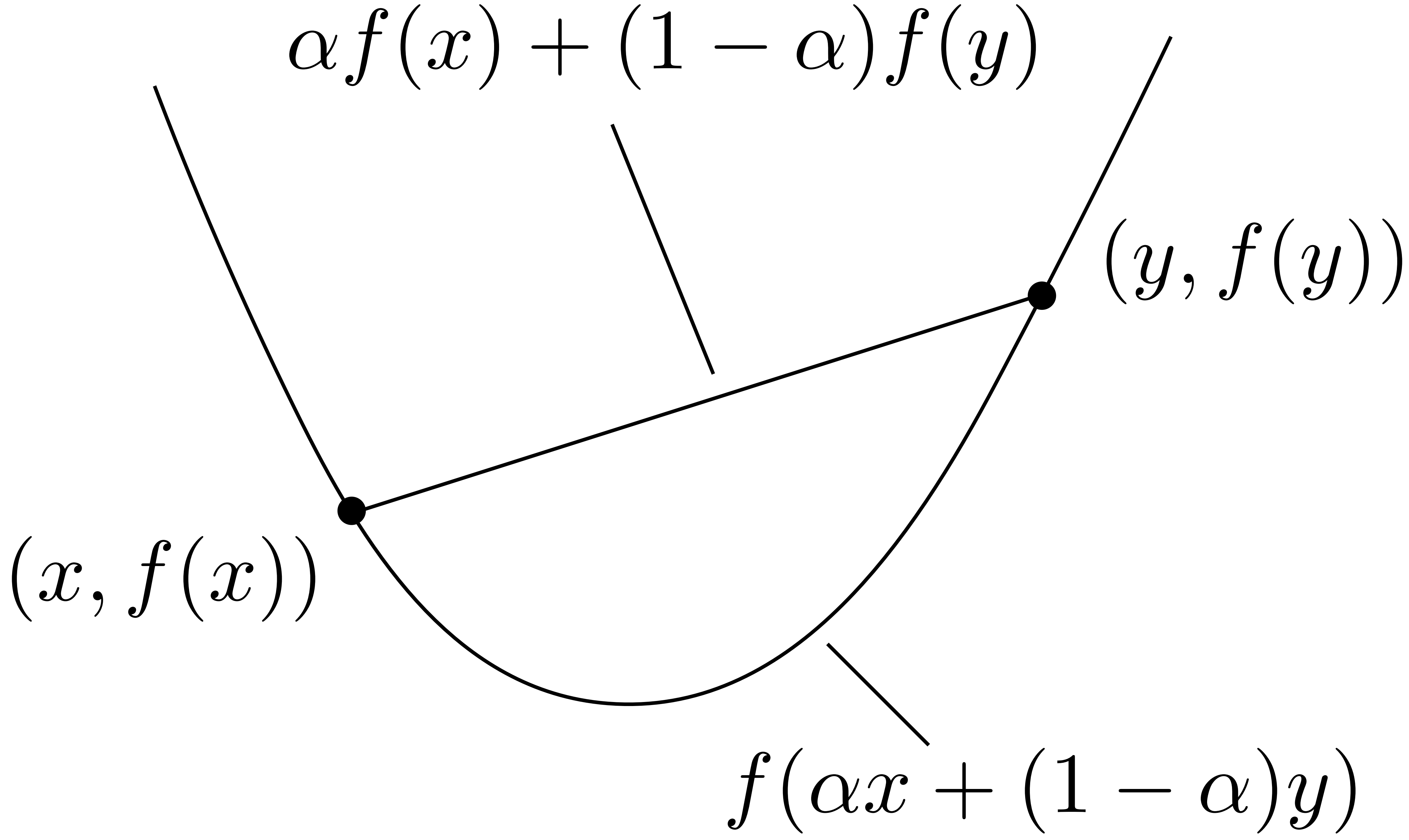
Extended-value functions map $\mathbf{R}^n$ to the extended real line $\mathbf{R} \cup \{\pm\infty\}$

Effective domain of f: $\quad \mathbf{dom}\, f = \{x \in \mathbf{R}^n \mid f(x) < \infty\}$

Convex function

For every $x, y \in \mathbf{R}^n,\ \alpha \in [0, 1]$

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



Concave function

f is concave if and only if $-f$ is convex

Gradient

Derivative

If $f(x) : \mathbf{R}^n \to \mathbf{R}^m$ continuously differentiable, we define

$$Df(x)_{ij} = \frac{\partial f_i(x)}{\partial x_j}, \quad i = 1, \dots, m, \quad j = 1, \dots, n$$

Gradient

If $f : \mathbf{R}^n \to \mathbf{R}$, we define

$$\nabla f(x) = Df(x)^T$$

Example

$$f(x) = (1/2)x^T P x + q^T x$$

$$\nabla f(x) = Px + q$$

First-order approximation

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

(affine function of y)

Hessian

Hessian matrix (second derivative)

If $f(x) : \mathbf{R}^n \to \mathbf{R}$ second-order differentiable, we define

$$\nabla^2 f(x)_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}, \quad i, j = 1, \dots, n$$

Example

$$f(x) = (1/2)x^T P x + q^T x$$

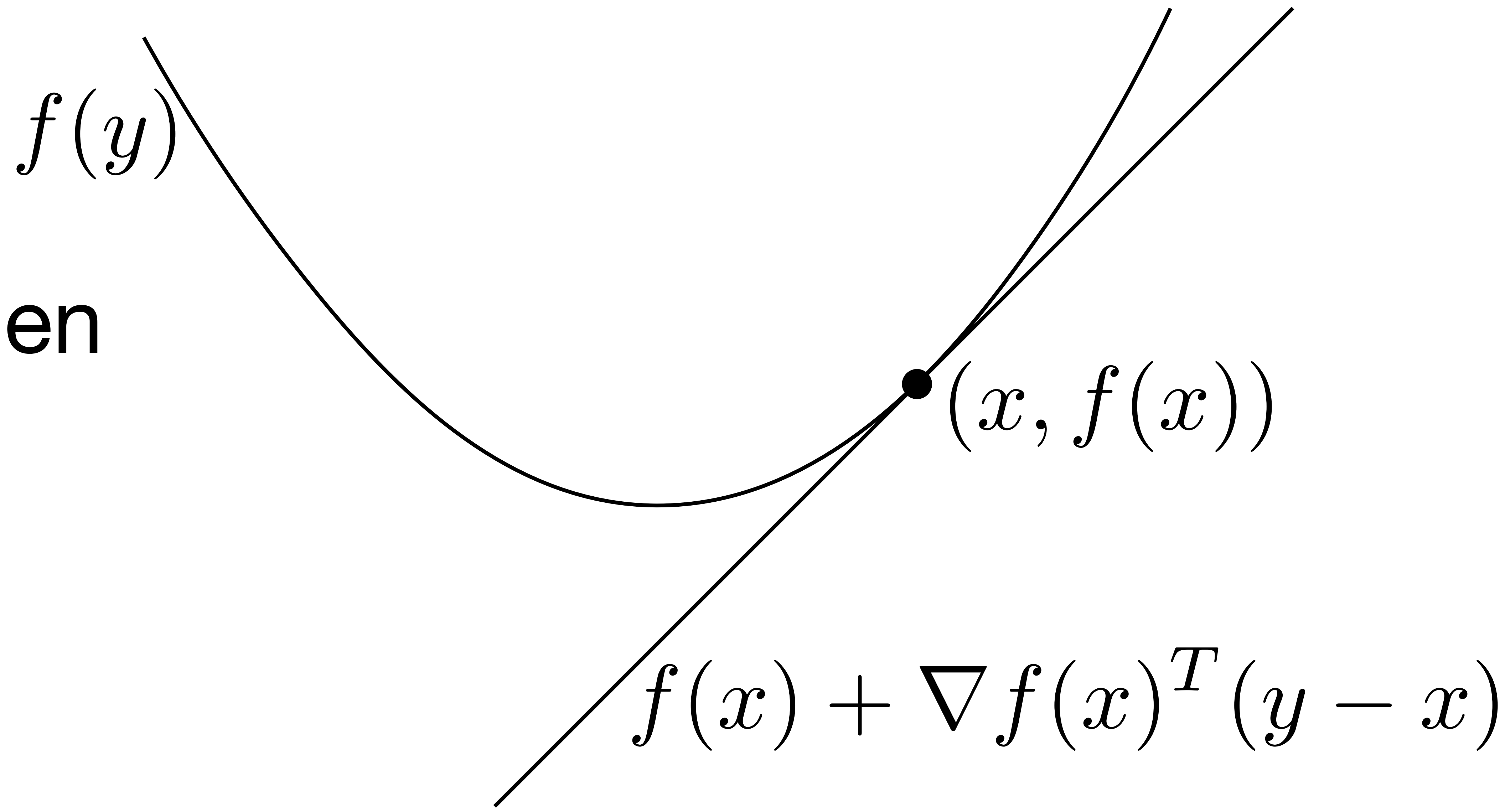
$$\nabla^2 f(x) = P$$

Second-order approximation

$$f(y) \approx f(x) + \nabla f(x)^T (y - x) + (1/2)(y - x)^T \nabla^2 f(x)(y - x)$$

(quadratic function of y)

Convex conditions



First-order

Let f be a continuous differentiable function, then it is convex if and only if $\mathbf{dom} f$ is convex and

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

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

Second-order

If f is twice differentiable, then f is convex if and only if $\mathbf{dom} f$ is convex and

$$\nabla^2 f(x) \succeq 0$$

for all $x \in \mathbf{dom} f$

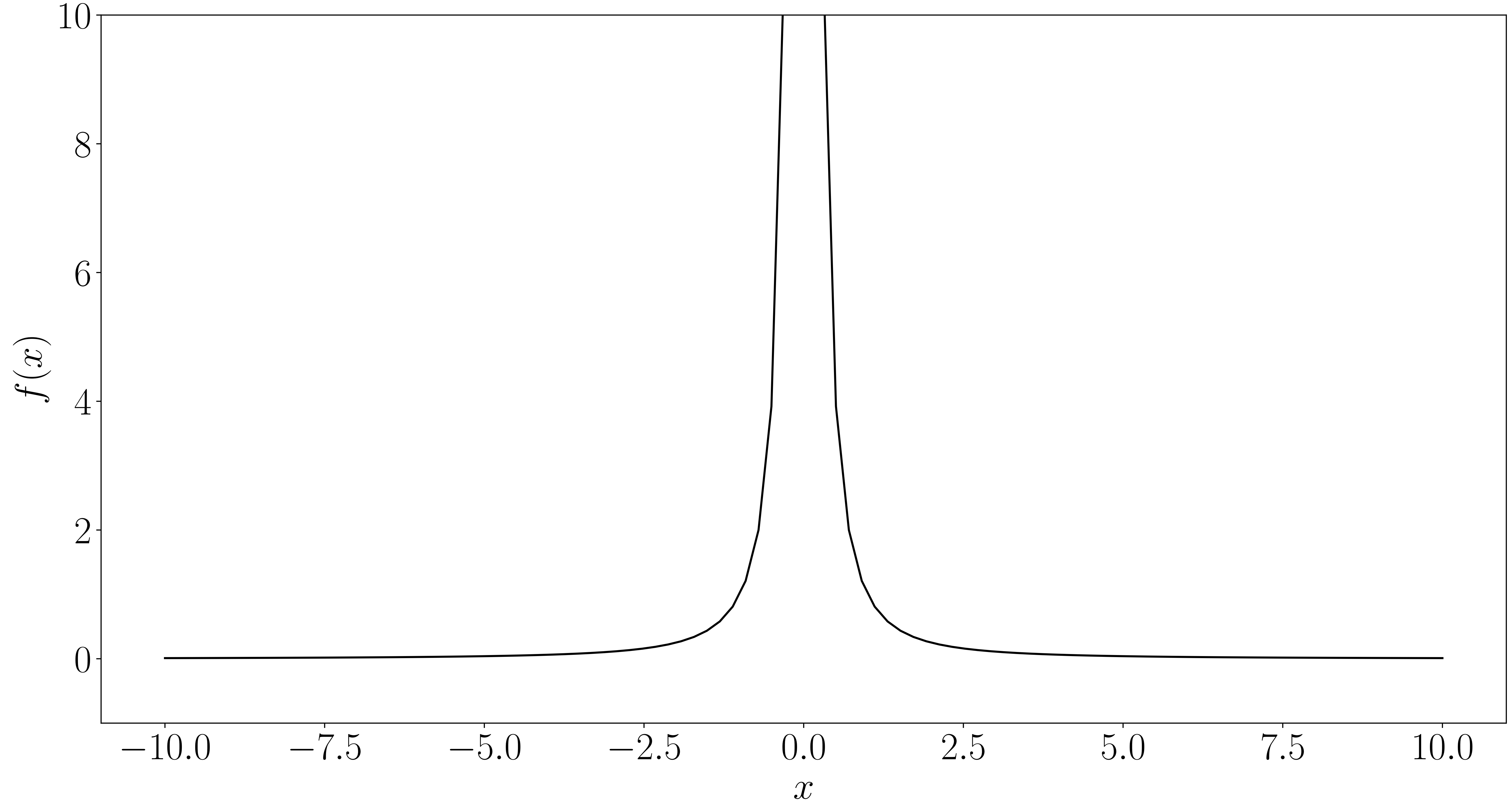
Convex domain requirement

$$f(x) = 1/x^2$$

$$\mathbf{dom} f = \{x \in \mathbf{R} \mid x \neq 0\}$$

$\nabla^2 f(x) > 0$ for all $x \in \mathbf{dom} f$

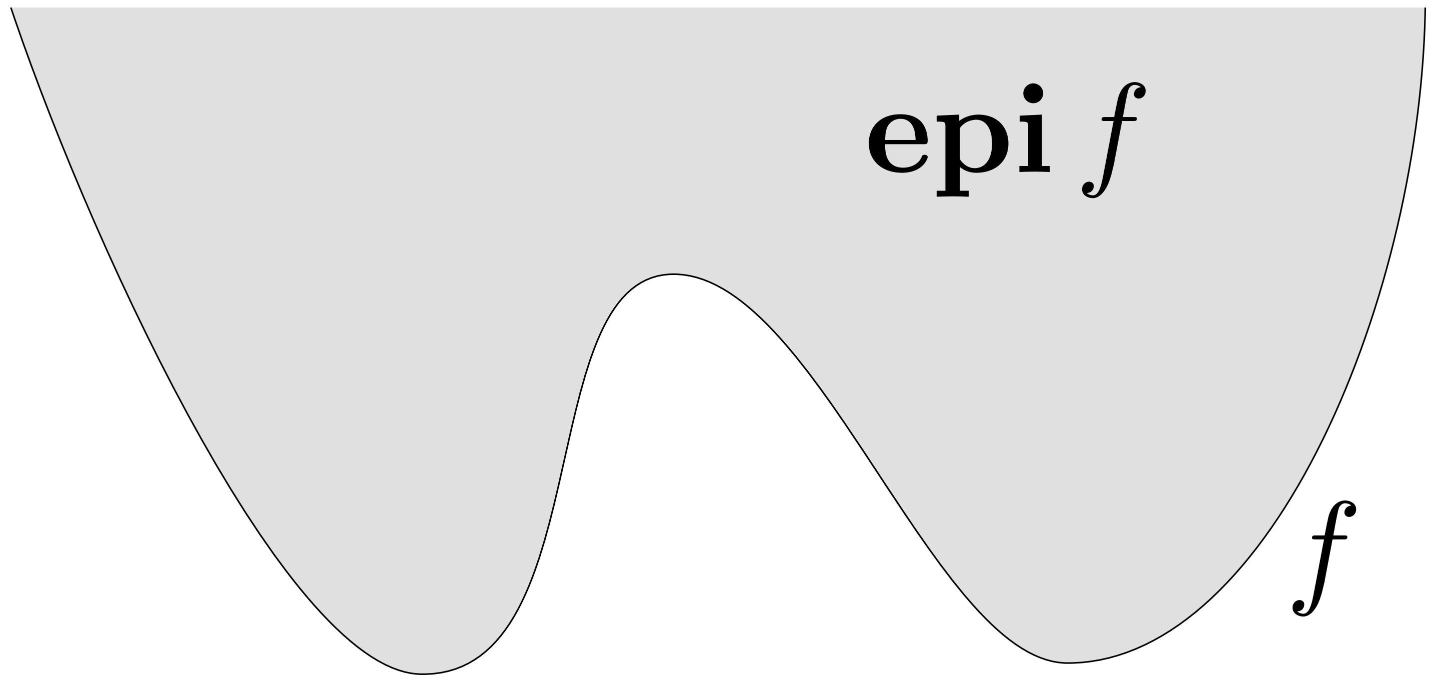
Non convex!

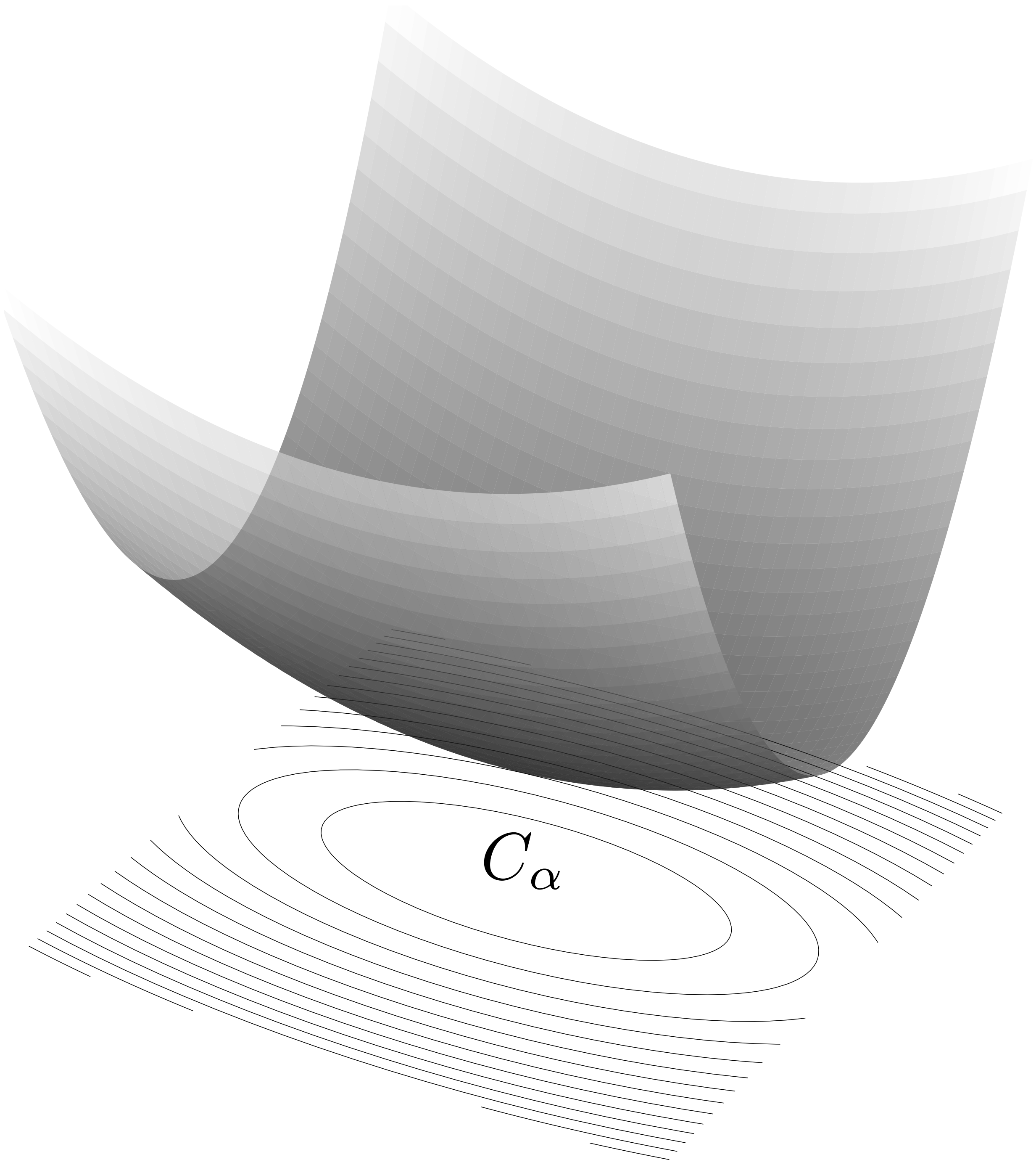


Function epigraph and sublevel sets

function epigraph

$$\mathbf{epi}\, f = \{(x,t) \mid x \in \mathbf{dom}\, f,\ f(x) \le t\}$$





sublevel sets

$$C_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \le \alpha\}$$

Closed convex proper functions

A function f is called **CCP** if it is

closed $\quad \mathbf{epi}\, f$ is a closed set

convex $\quad f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y), \quad \alpha \in [0,1]$

proper $\quad \mathbf{dom}\, f$ is nonempty and $f = -\infty$ never

If not otherwise stated, we assume functions to be CCP

Properties

f is convex $\iff$ $\mathbf{epi}\, f$ is convex

f is convex $\Rightarrow C_\alpha$ is convex $\forall \alpha$ (converse not true, e.g., $f(x) = -e^x$)

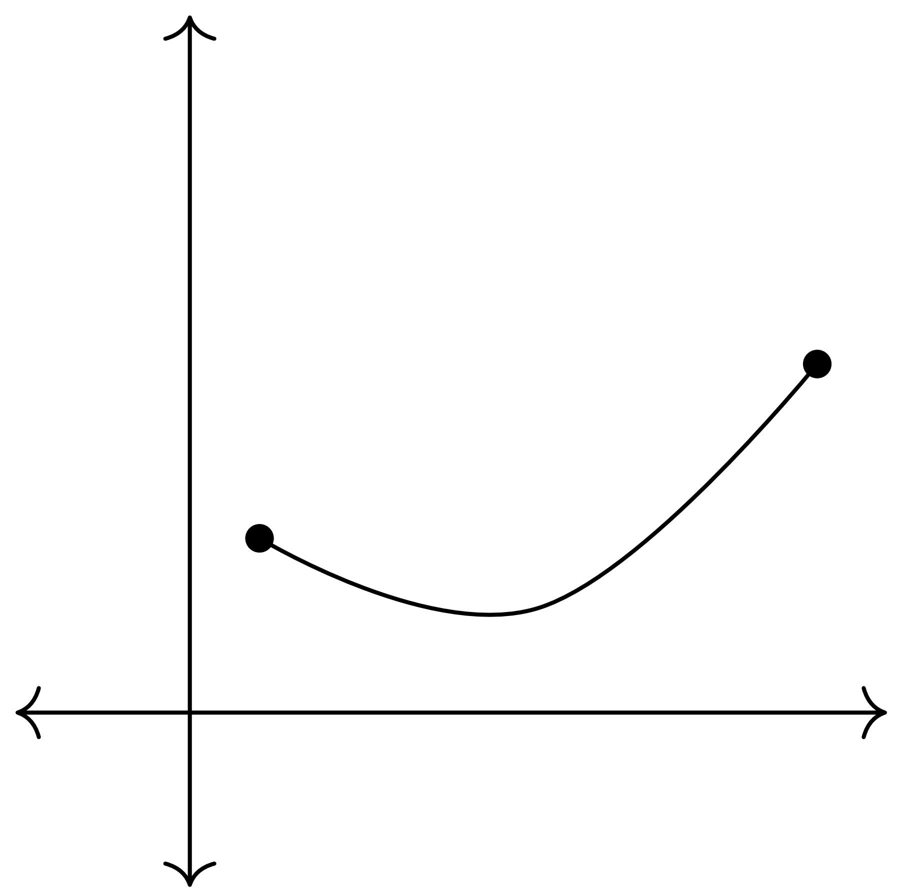
For proper f, f is closed $\iff$ all sublevel sets closed (f lower semicontinuous)

f is CCP $\iff$ $\mathbf{epi}\, f$ is nonempty, closed, convex, without vertical line

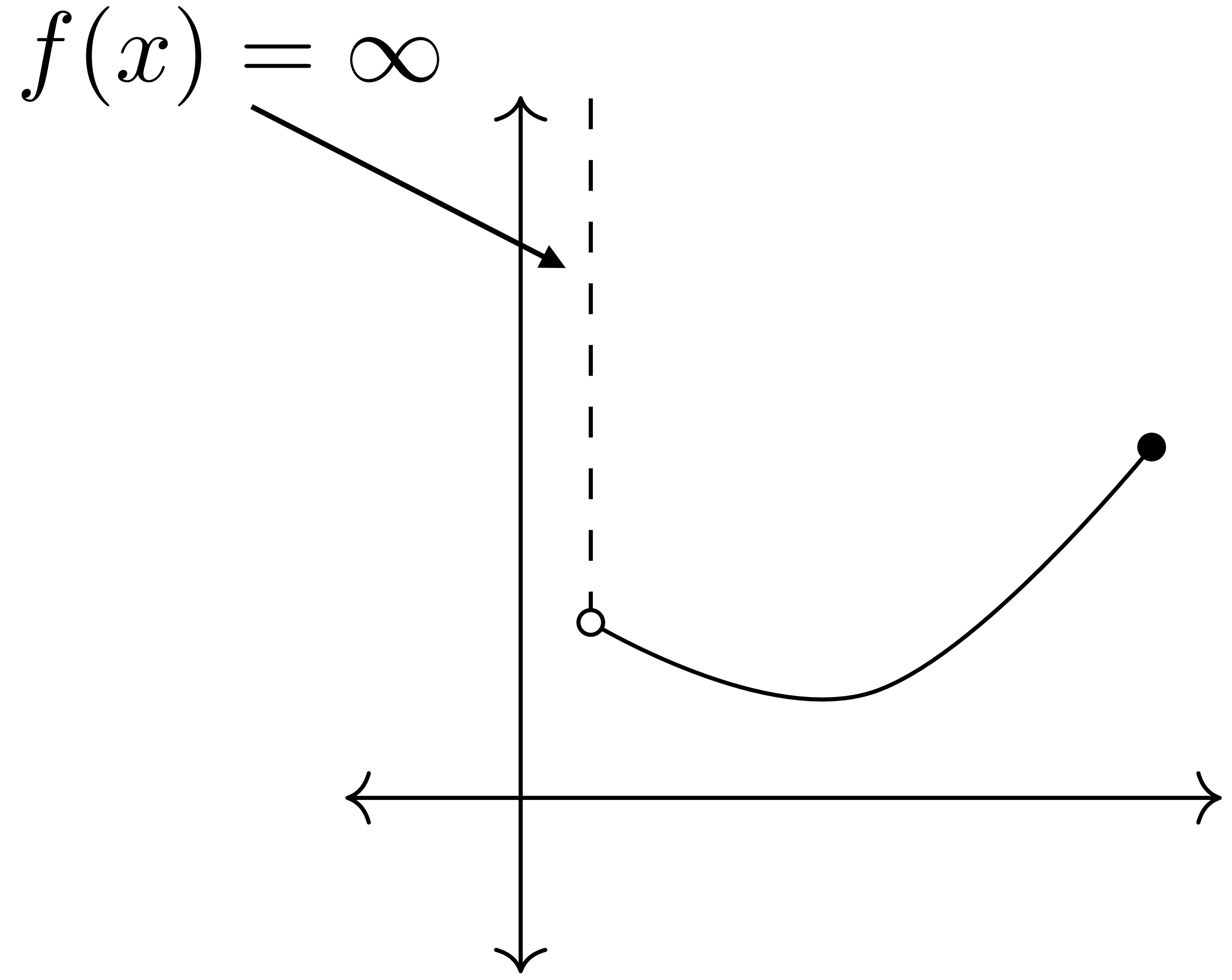
CCP function example

whether convex function f is closed depends on its behavior

on the boundary of $\mathbf{dom}\, f$



closed convex function



convex but not closed

Indicator functions

For $C \subseteq \mathbf{R}^n$, define the *indicator function* as

$$\mathcal{I}_C(x) = \begin{cases} 0 & x \in C \\ \infty & x \notin C \end{cases}$$

If C is convex, closed and nonempty, then $\mathcal{I}_C$ is CCP

Constrained form		**Unconstrained form**
minimize $f(x)$ subject to $x \in C$	$\longleftrightarrow$	minimize $f(x) + \mathcal{I}_C(x)$

Convex optimization

Convex optimization problems

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g_i(x) \leq 0, \quad i = 1, \ldots, m \end{array}$$

$f : \mathbf{R}^n \to \mathbf{R}$ Convex objective function

$g_i : \mathbf{R}^n \to \mathbf{R}$ Convex constraints functions

Convex feasible set

$$C = \{x \mid g_i(x) \leq 0, \quad i = 1, \ldots, m\}$$

Verifying convexity

Basic definition (inequality)

First and second order conditions (gradient, hessian) → **Hard!**

Convex calculus (directly construct convex functions)

- Library of basic functions that are convex/concave
- Calculus rules or transformations that preserve convexity

→ **Easy!**

Disciplined Convex Programming

Convexity by construction

General composition rule

$h(f_1(x), f_2(x), \ldots, f_k(x))$ is convex when h is convex and for each i

- h is nondecreasing in argument i and f_i is convex, or
- h is nonincreasing in argument i and f_i is concave, or
- f_i is affine

Only sufficient condition

Check your functions at `https://dcp.stanford.edu/`

More details and examples in ORF523

Modelling software for convex optimization

Modelling tools simplify the formulation of convex optimization problems

- **Construct problems** using library of basic functions
- **Verify convexity** by general composition rule
- Express the problem in input format required by a specific solver

Examples

- CVX, YALMIP (Matlab)
- CVXPY (Python)
- Convex.jl (Julia)

Local vs global minima (optimizers)

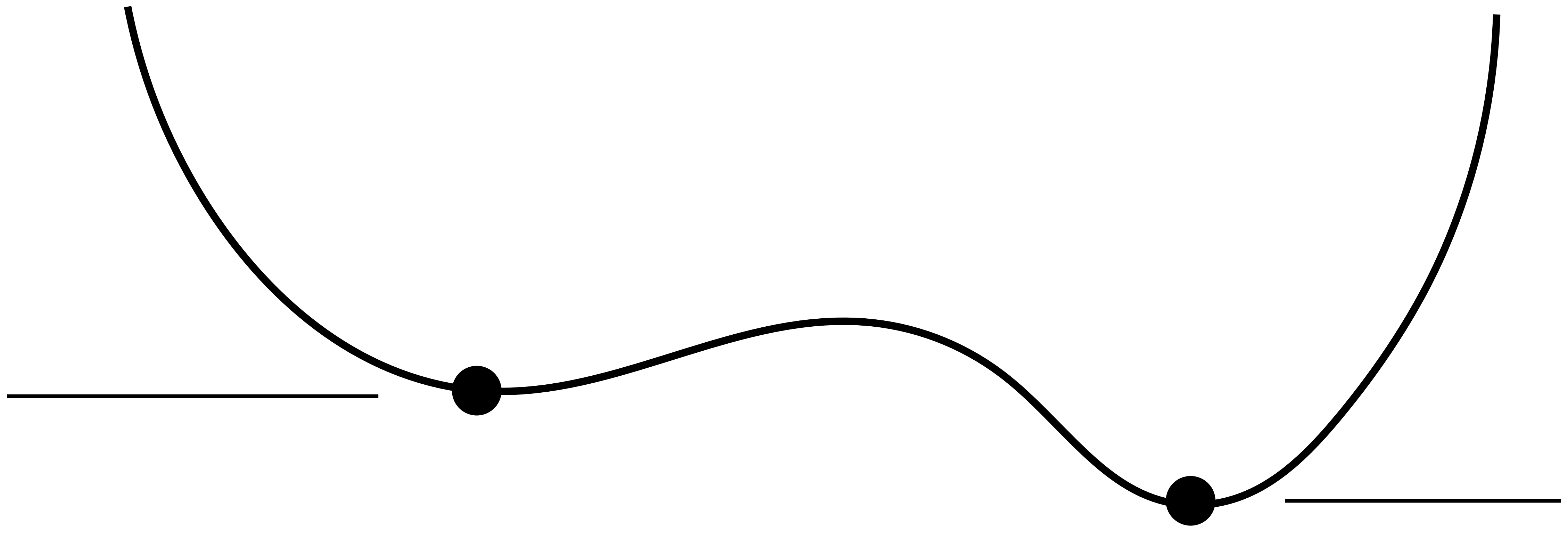
minimize $\quad f(x)$

subject to $\quad x \in C$

Local optimizer x

$f(y) \geq f(x), \quad \forall y$

such that $\|x - y\|_2 \leq R$



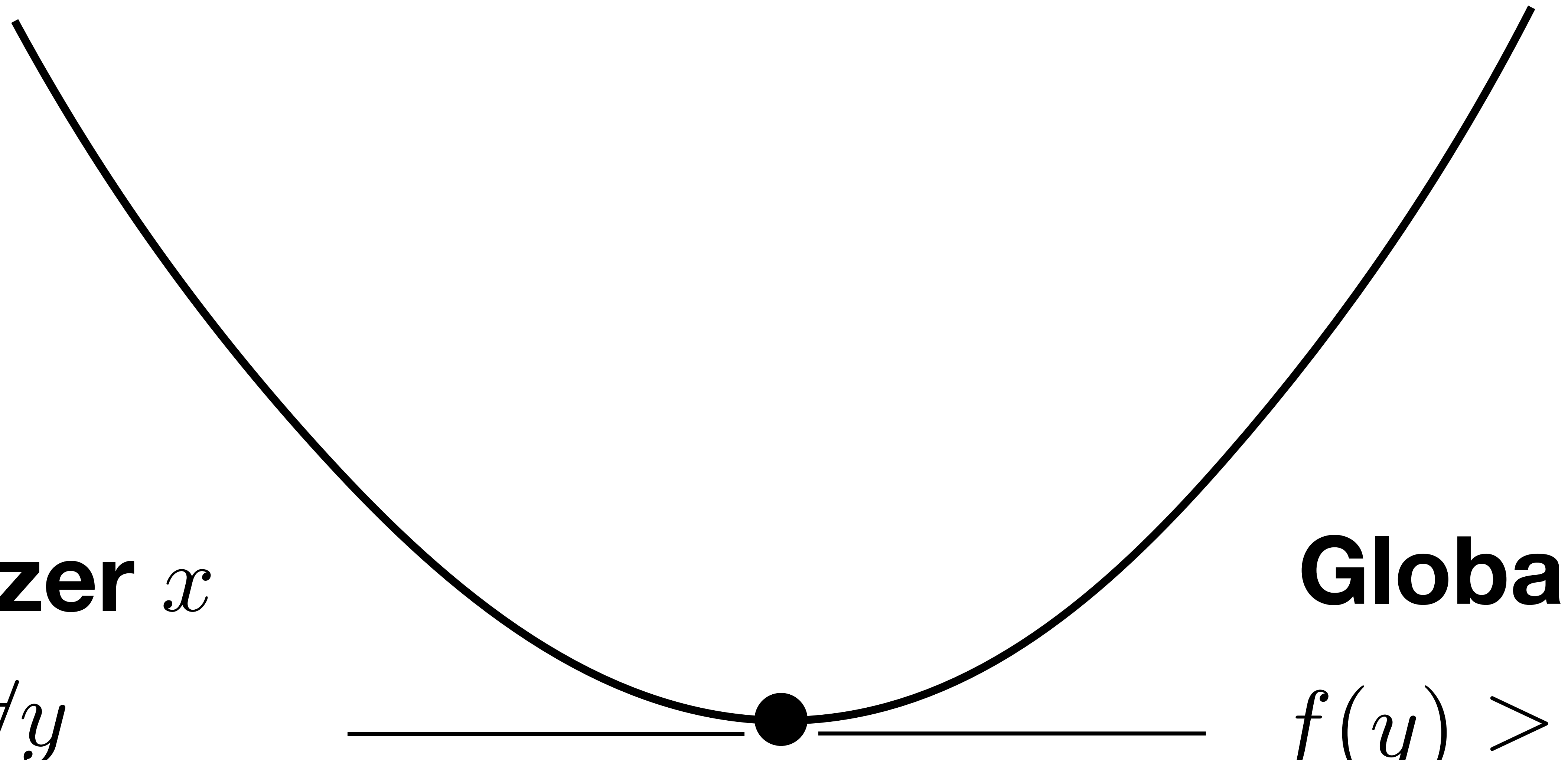
Global optimizer x

$f(y) \geq f(x), \quad \forall y \in C$

Optimality and convexity

Theorem

For a convex optimization problem, any **local minimum** is a **global minimum**



Local optimizer x

$f(y) \geq f(x), \quad \forall y$
such that $\|x - y\|_2 \leq R$

Global optimizer x

$f(y) \geq f(x), \quad \forall y \in C$

Optimality and convexity

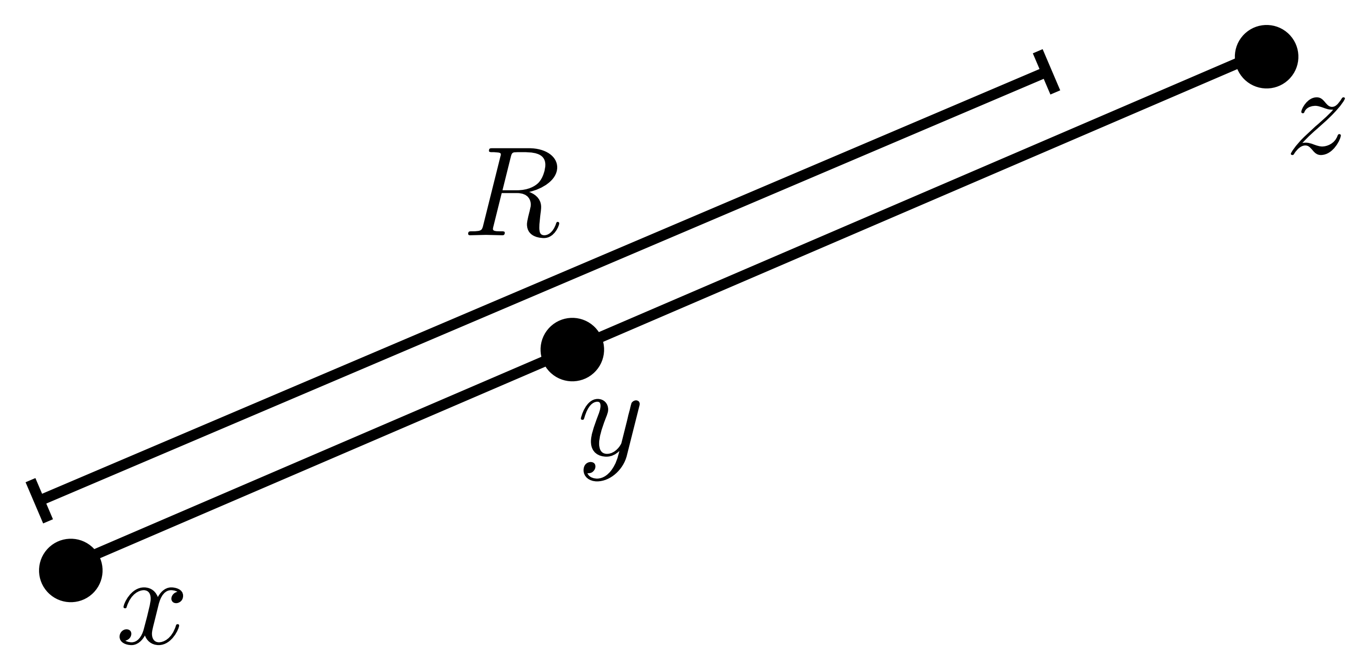
Proof (contradiction)

Suppose that f is convex and x is a local (not global) minimum for f, i.e.,

$$f(y) \geq f(x), \quad \forall y \text{ such that } \|x - y\|_2 \leq R.$$

Therefore, there exists a feasible z such that $\|z - x\| > R$ and $f(z) < f(x)$.

Consider $y = (1 - \alpha)x + \alpha z$ with $\alpha = \dfrac{R}{2\|z - x\|_2}$.



Then, $\|y - x\|_2 = \alpha\|z - x\|_2 = R/2 < R$, and by convexity of the feasible set, y is feasible.

By convexity of f we have $f(y) \leq (1 - \alpha)f(x) + \alpha f(z) < f(x)$, which contradicts the local optimum definition.

Therefore, x is globally optimal. ∎

"...in fact, the great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity."

R. Tyrrell Rockafellar, in SIAM Review, 1993

Introduction to nonlinear optimization

Today, we learned to:

- **Define** nonlinear optimization problems
- **Understand** convex analysis fundamentals (sets, cones, functions, and gradients)
- **Verify** convexity and **construct** convex optimization problems
- **Understand** the importance of *convexity vs nonconvexity* in optimization

Next lecture

- Optimality conditions in nonlinear optimization