

# Convex Set

# Convex set

**line segment** between  $x_1$  and  $x_2$ : all points

$$x = \theta x_1 + (1 - \theta)x_2$$

with  $0 \leq \theta \leq 1$

**convex set**: contains line segment between any two points in the set

$$x_1, x_2 \in C, \quad 0 \leq \theta \leq 1 \quad \implies \quad \theta x_1 + (1 - \theta)x_2 \in C$$

**examples** (one convex, two nonconvex sets)



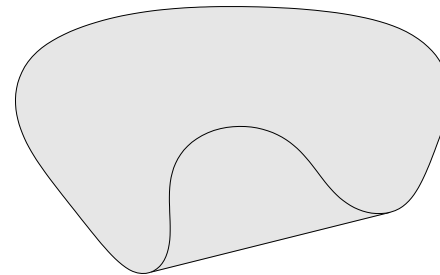
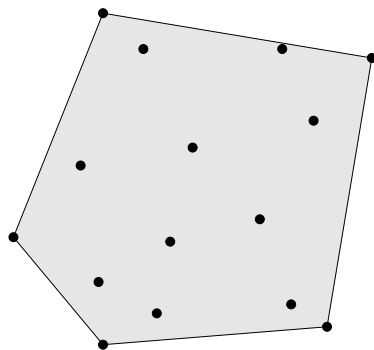
# Convex combination and convex hull

**convex combination** of  $x_1, \dots, x_k$ : any point  $x$  of the form

$$x = \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_k x_k$$

with  $\theta_1 + \dots + \theta_k = 1$ ,  $\theta_i \geq 0$

**convex hull**  $\text{conv } S$ : set of all convex combinations of points in  $S$



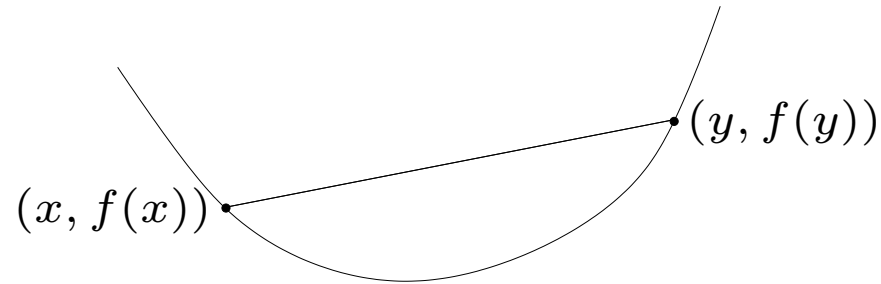
# Convex Function

# Convex function

$f : \mathbf{R}^n \rightarrow \mathbf{R}$  is convex if  $\mathbf{dom} f$  is a convex set and

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

for all  $x, y \in \mathbf{dom} f$ ,  $0 \leq \theta \leq 1$



- $f$  is concave if  $-f$  is convex
- $f$  is strictly convex if  $\mathbf{dom} f$  is convex and

$$f(\theta x + (1 - \theta)y) < \theta f(x) + (1 - \theta)f(y)$$

for  $x, y \in \mathbf{dom} f$ ,  $x \neq y$ ,  $0 < \theta < 1$

## Second-order conditions

$f$  is **twice differentiable** if  $\text{dom } f$  is open and the Hessian  $\nabla^2 f(x) \in \mathbf{S}^n$ ,

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

exists at each  $x \in \text{dom } f$

**2nd-order conditions:** for twice differentiable  $f$  with convex domain

- $f$  is convex if and only if

$$\nabla^2 f(x) \succeq 0 \quad \text{for all } x \in \text{dom } f$$

- if  $\nabla^2 f(x) \succ 0$  for all  $x \in \text{dom } f$ , then  $f$  is strictly convex

# Jensen's inequality

**basic inequality:** if  $f$  is convex, then for  $0 \leq \theta \leq 1$ ,

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

**extension:** if  $f$  is convex, then

$$f(\mathbf{E} z) \leq \mathbf{E} f(z)$$

for any random variable  $z$

basic inequality is special case with discrete distribution

$$\mathbf{prob}(z = x) = \theta, \quad \mathbf{prob}(z = y) = 1 - \theta$$

# Positive weighted sum & composition with affine function

**nonnegative multiple:**  $\alpha f$  is convex if  $f$  is convex,  $\alpha \geq 0$

**sum:**  $f_1 + f_2$  convex if  $f_1, f_2$  convex (extends to infinite sums, integrals)

**composition with affine function:**  $f(Ax + b)$  is convex if  $f$  is convex

## examples

- log barrier for linear inequalities

$$f(x) = - \sum_{i=1}^m \log(b_i - a_i^T x), \quad \text{dom } f = \{x \mid a_i^T x < b_i, i = 1, \dots, m\}$$

- (any) norm of affine function:  $f(x) = \|Ax + b\|$



# Duality

# Lagrangian

**standard form problem** (not necessarily convex)

$$\begin{array}{ll}\text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \\ & h_i(x) = 0, \quad i = 1, \dots, p\end{array}$$

variable  $x \in \mathbf{R}^n$ , domain  $\mathcal{D}$ , optimal value  $p^*$

**Lagrangian:**  $L : \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^p \rightarrow \mathbf{R}$ , with  $\text{dom } L = \mathcal{D} \times \mathbf{R}^m \times \mathbf{R}^p$ ,

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x)$$

- weighted sum of objective and constraint functions
- $\lambda_i$  is Lagrange multiplier associated with  $f_i(x) \leq 0$
- $\nu_i$  is Lagrange multiplier associated with  $h_i(x) = 0$

# Lagrange dual function

**Lagrange dual function:**  $g : \mathbf{R}^m \times \mathbf{R}^p \rightarrow \mathbf{R}$ ,

$$\begin{aligned} g(\lambda, \nu) &= \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) \\ &= \inf_{x \in \mathcal{D}} \left( f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x) \right) \end{aligned}$$

$g$  is concave, can be  $-\infty$  for some  $\lambda, \nu$

**lower bound property:** if  $\lambda \succeq 0$ , then  $g(\lambda, \nu) \leq p^*$

# The dual problem

## Lagrange dual problem

$$\begin{array}{ll}\text{maximize} & g(\lambda, \nu) \\ \text{subject to} & \lambda \succeq 0\end{array}$$

- finds best lower bound on  $p^*$ , obtained from Lagrange dual function
- a convex optimization problem; optimal value denoted  $d^*$
- $\lambda, \nu$  are dual feasible if  $\lambda \succeq 0, (\lambda, \nu) \in \mathbf{dom} g$
- often simplified by making implicit constraint  $(\lambda, \nu) \in \mathbf{dom} g$  explicit

# Karush-Kuhn-Tucker (KKT) conditions

the following four conditions are called KKT conditions (for a problem with differentiable  $f_i, h_i$ ):

1. primal constraints:  $f_i(x) \leq 0, i = 1, \dots, m, h_i(x) = 0, i = 1, \dots, p$
2. dual constraints:  $\lambda \succeq 0$
3. complementary slackness:  $\lambda_i f_i(x) = 0, i = 1, \dots, m$
4. gradient of Lagrangian with respect to  $x$  vanishes:

$$\nabla f_0(x) + \sum_{i=1}^m \lambda_i \nabla f_i(x) + \sum_{i=1}^p \nu_i \nabla h_i(x) = 0$$

if strong duality holds and  $x, \lambda, \nu$  are optimal, then they must satisfy the KKT conditions