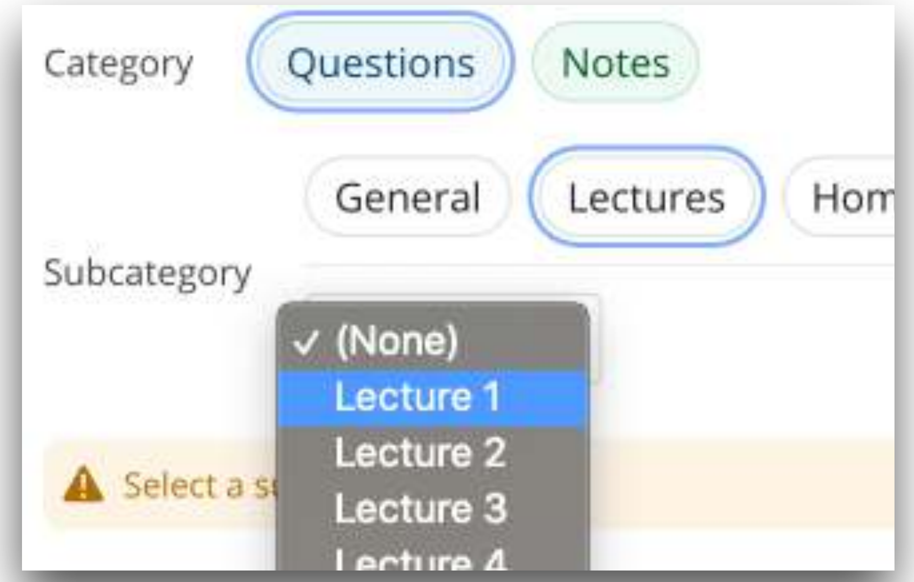
ORF522 – Linear and Nonlinear Optimization

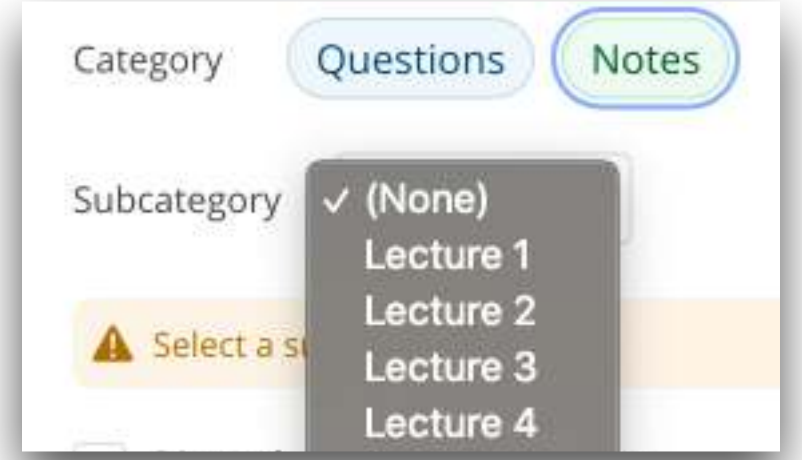
3. Geometry and polyhedra

Bartolomeo Stellato — Fall 2022

Ed Forum



- **General Forum**
  - Please select the relecent Lecture in Notes/Question
  - Just need one question or comment. Not both



- **Questions/Notes**
  - Converting a problem in standard form increases the dimension of the problem, potentially by quite a lot. Is this ever an issue? Are there cases where we may want to not convert to standard form?
  - Why in general, machine learning people would love to use l2 norm in their loss function? Also, what's the intuition behind the fact that l2 norm cannot fully recover the sparse signal but l1 norm can?

Today's agenda
Readings [Chapter 2, Bertsimas and Tsitsiklis]

- Polyhedra and linear algebra
- Corners: extreme points, vertices, basic feasible solutions
- Constructing basic solutions
- Existence and optimality of extreme points

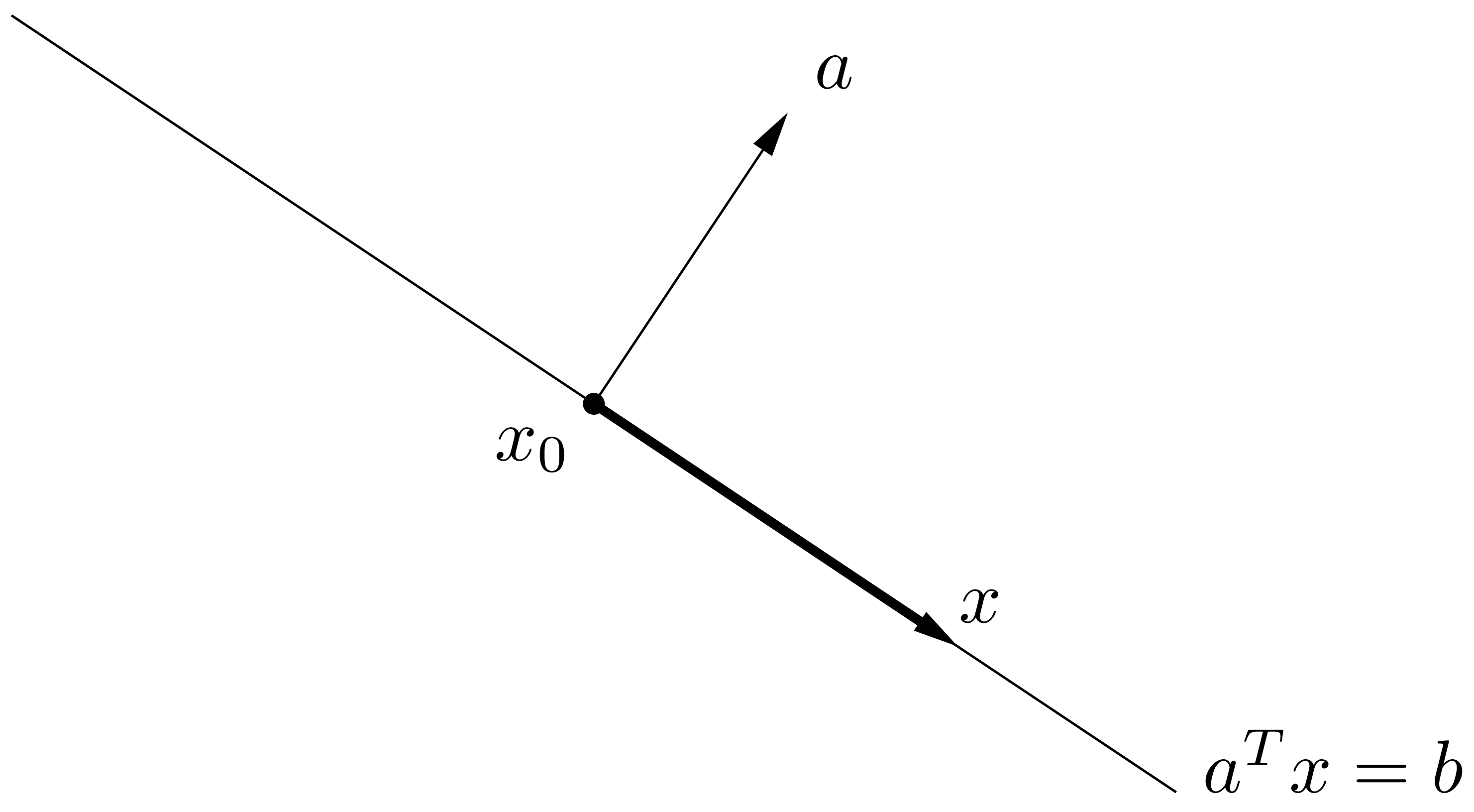
Polyhedra and linear algebra

Hyperplanes and halfspaces

Definitions

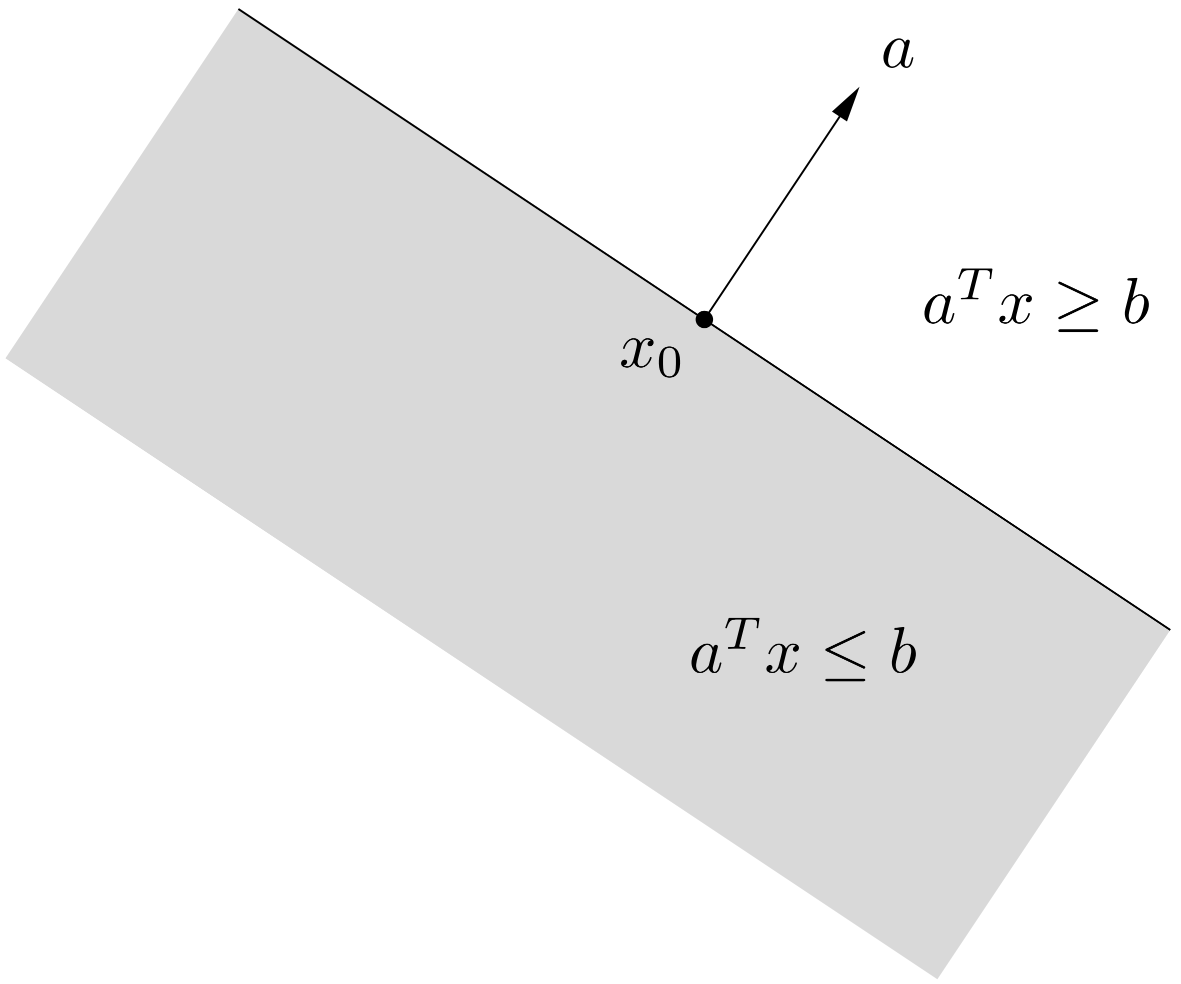
Hyperplane

$\{x \mid a^T x = b\}$



Halfspace

$\{x \mid a^T x \leq b\}$

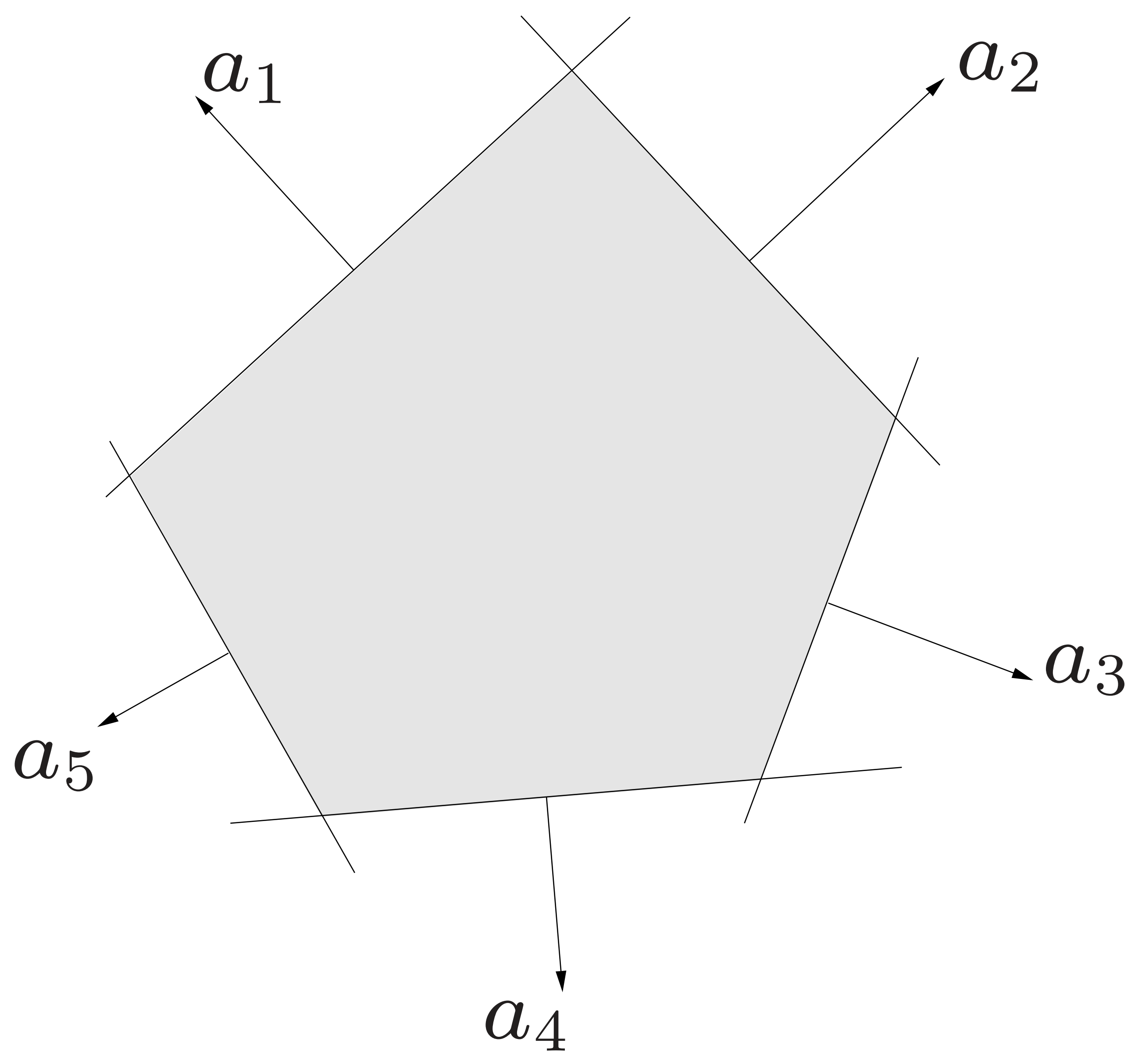


- x_0 is a specific point in the hyperplane
- For any x in the hyperplane defined by $a^T x = b$, $x - x_0 \perp a$
- The halfspace determined by $a^T x \leq b$ extends in the direction of $-a$

Polyhedron

Definition

$$P = \{x \mid a_i^T x \leq b_i, \quad i = 1, \ldots, m\} = \{x \mid Ax \leq b\}$$

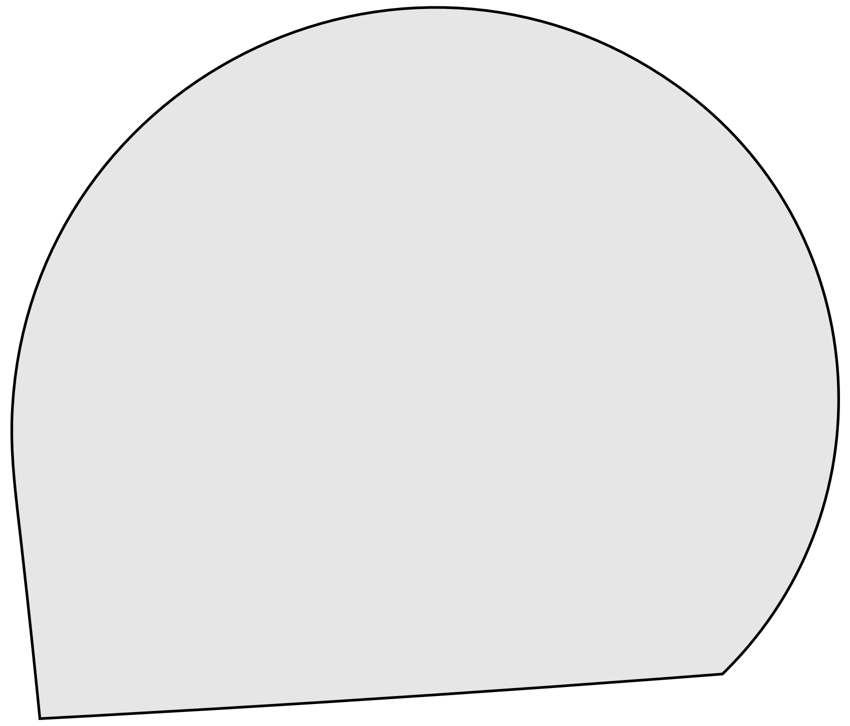


- Intersection of finite number of halfspaces
- Can include equalities

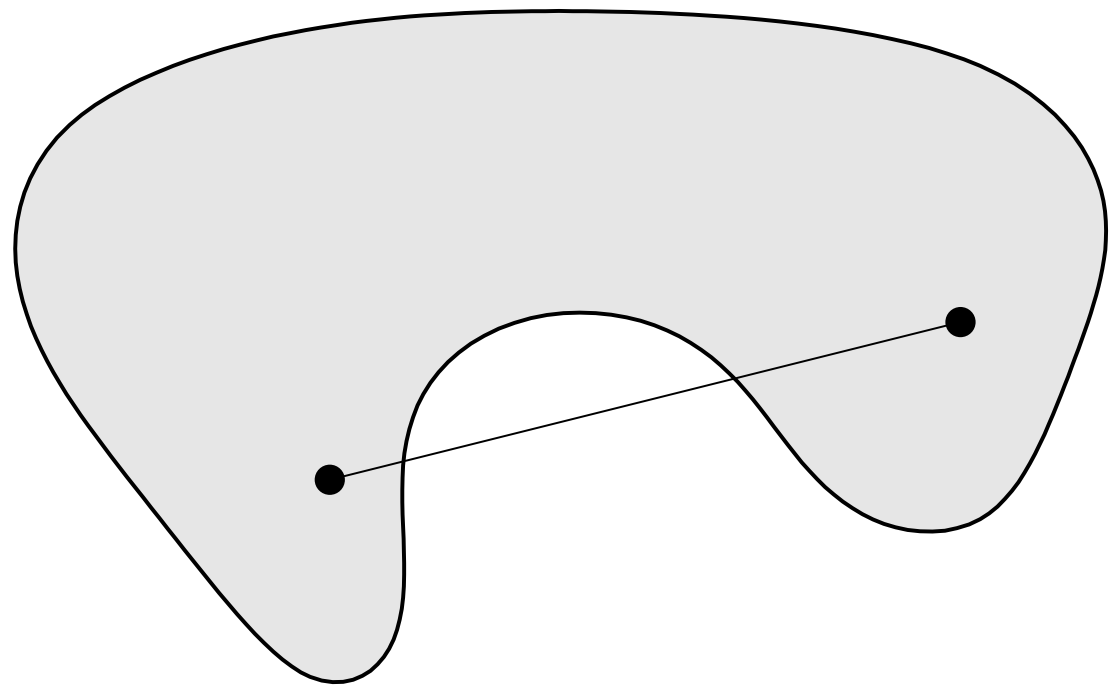
Convex set

Definition

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



Convex



Not convex

Examples

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

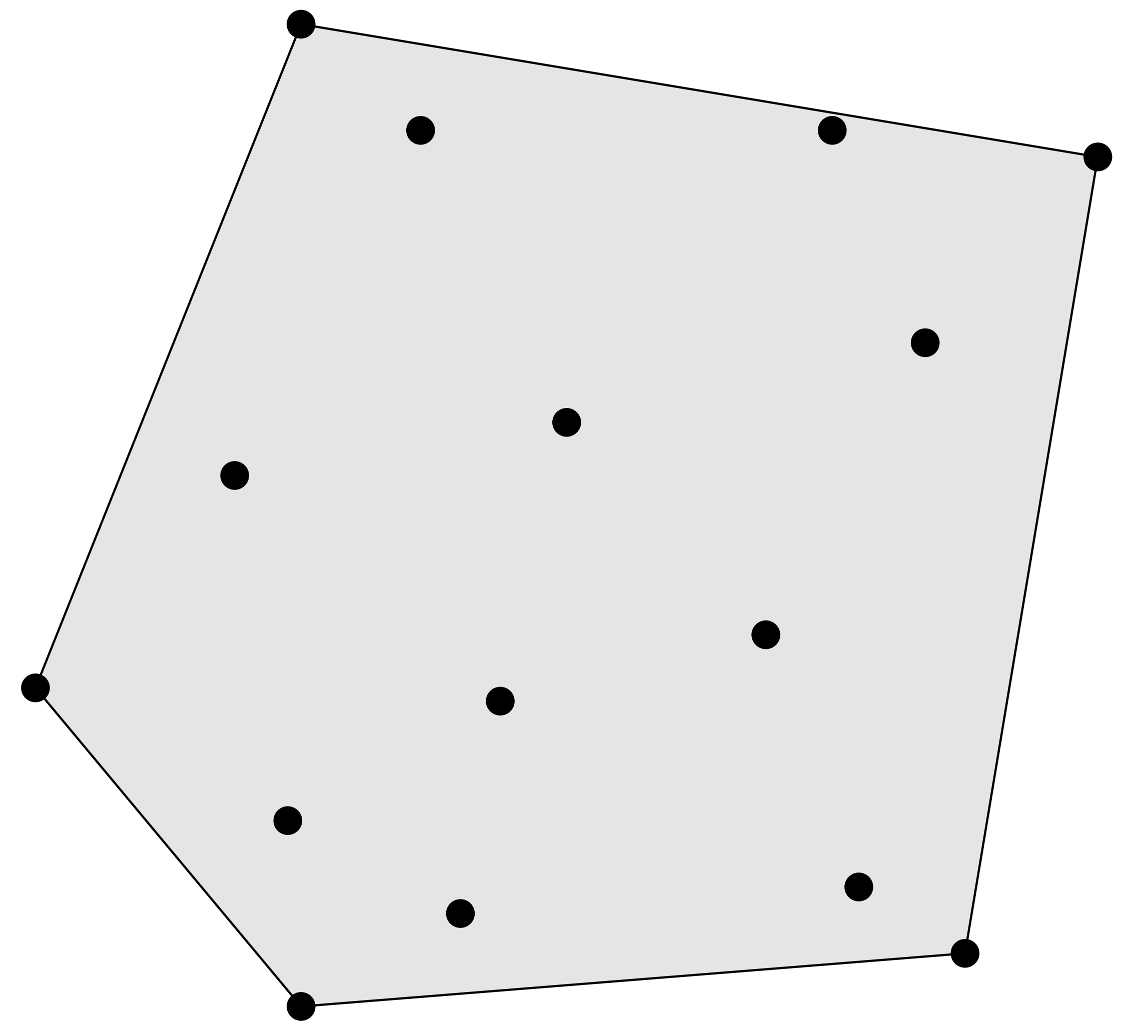
Convex combinations

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

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



Linear independence

a nonempty set of vectors $\{v_1, \ldots, v_k\}$ is **linearly indepdendent** if

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = 0$$

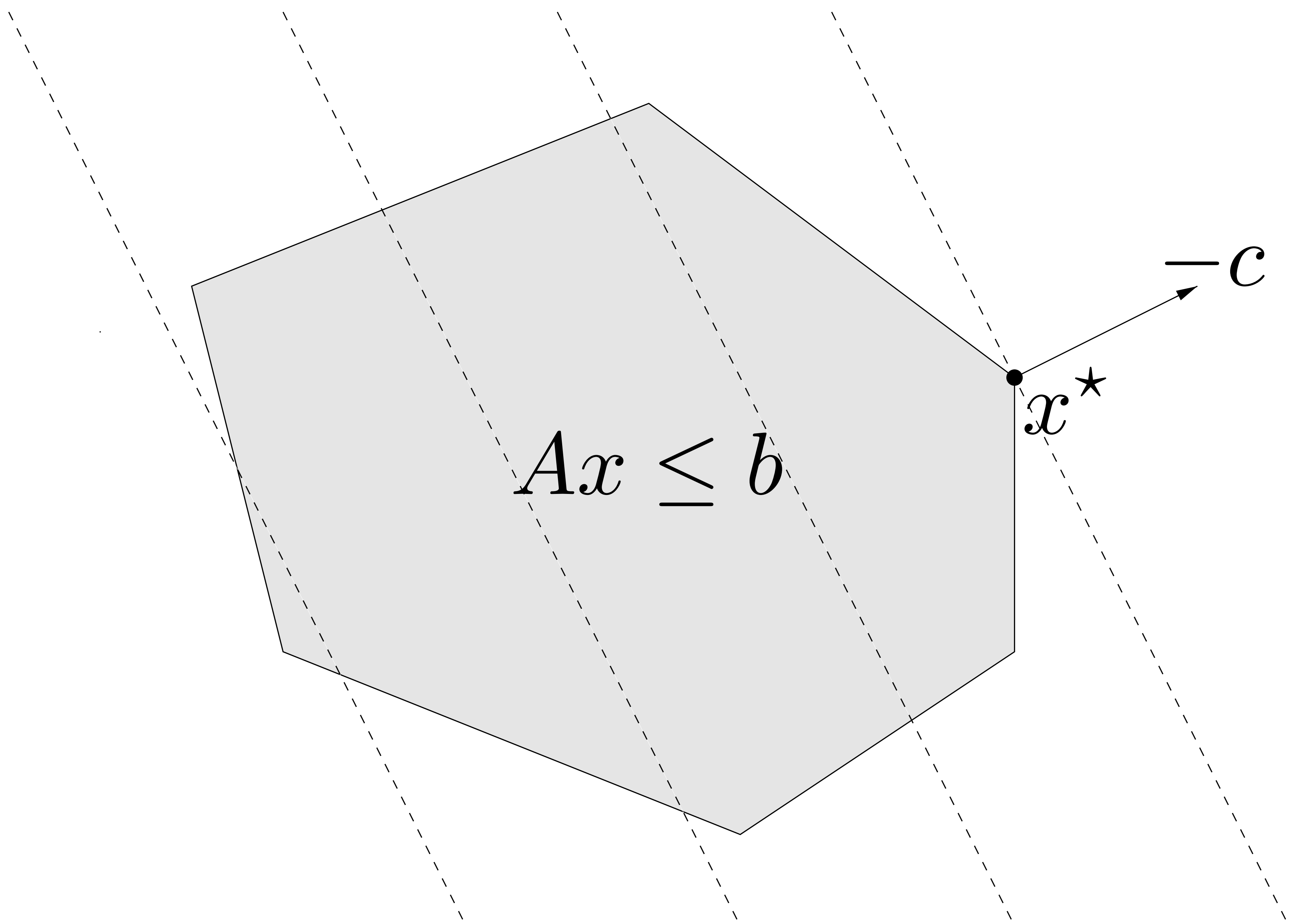
holds only for $\alpha_1 = \cdots = \alpha_k = 0$

Properties

- The coefficients α_k in a linear combination $x = \alpha_1 v_1 + \cdots + \alpha_k v_k$ are unique
- None of the vectors v_i is a linear combination of the other vectors

Geometrical interpretation of linear optimization

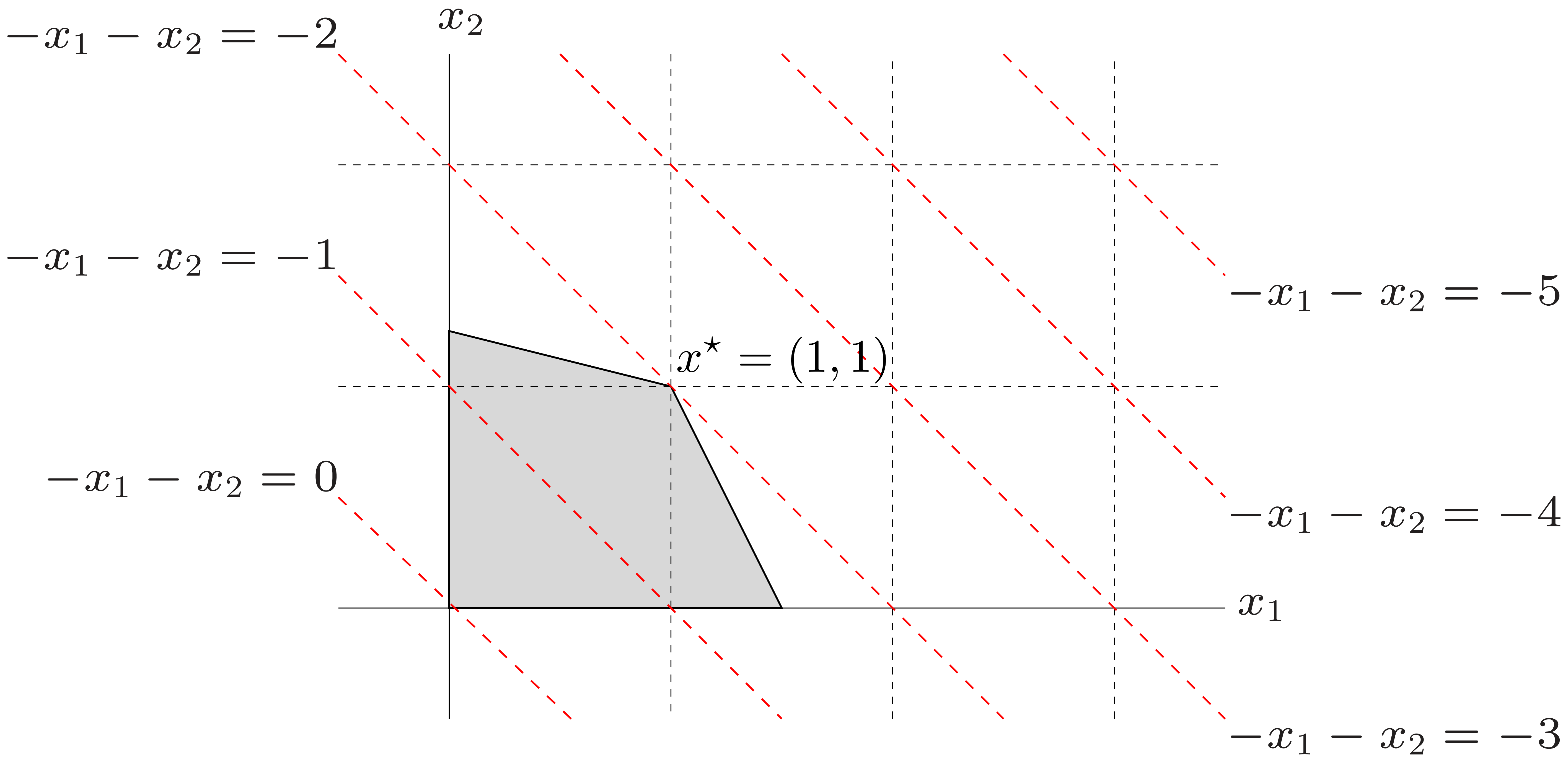
$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \end{array}$$



Dashed lines (hyperplanes) are level sets $c^T x = \alpha$ for different α

Example of linear optimization

$$
\begin{array}{ll}
\text{minimize} & -x_1 - x_2 \\
\text{subject to} & 2x_1 + x_2 \leq 3 \\
& x_1 + 4x_2 \leq 5 \\
& x_1 \geq 0, \; x_2 \geq 0
\end{array}
$$



Optimal solutions tend to be at a “**corner**” of the feasible set

How do we formalize it?

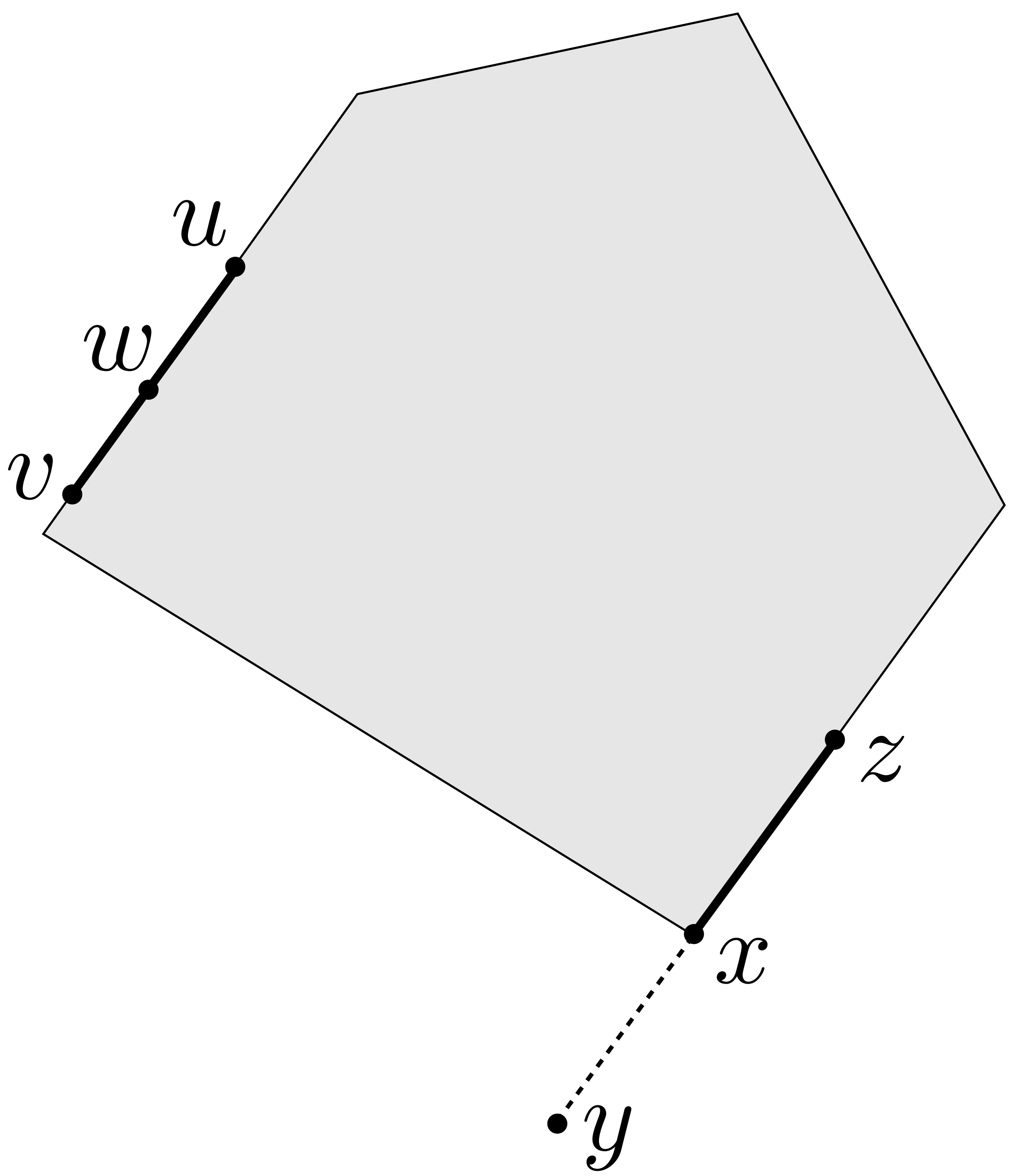
Corners of linear optimization

Extreme points

Definition

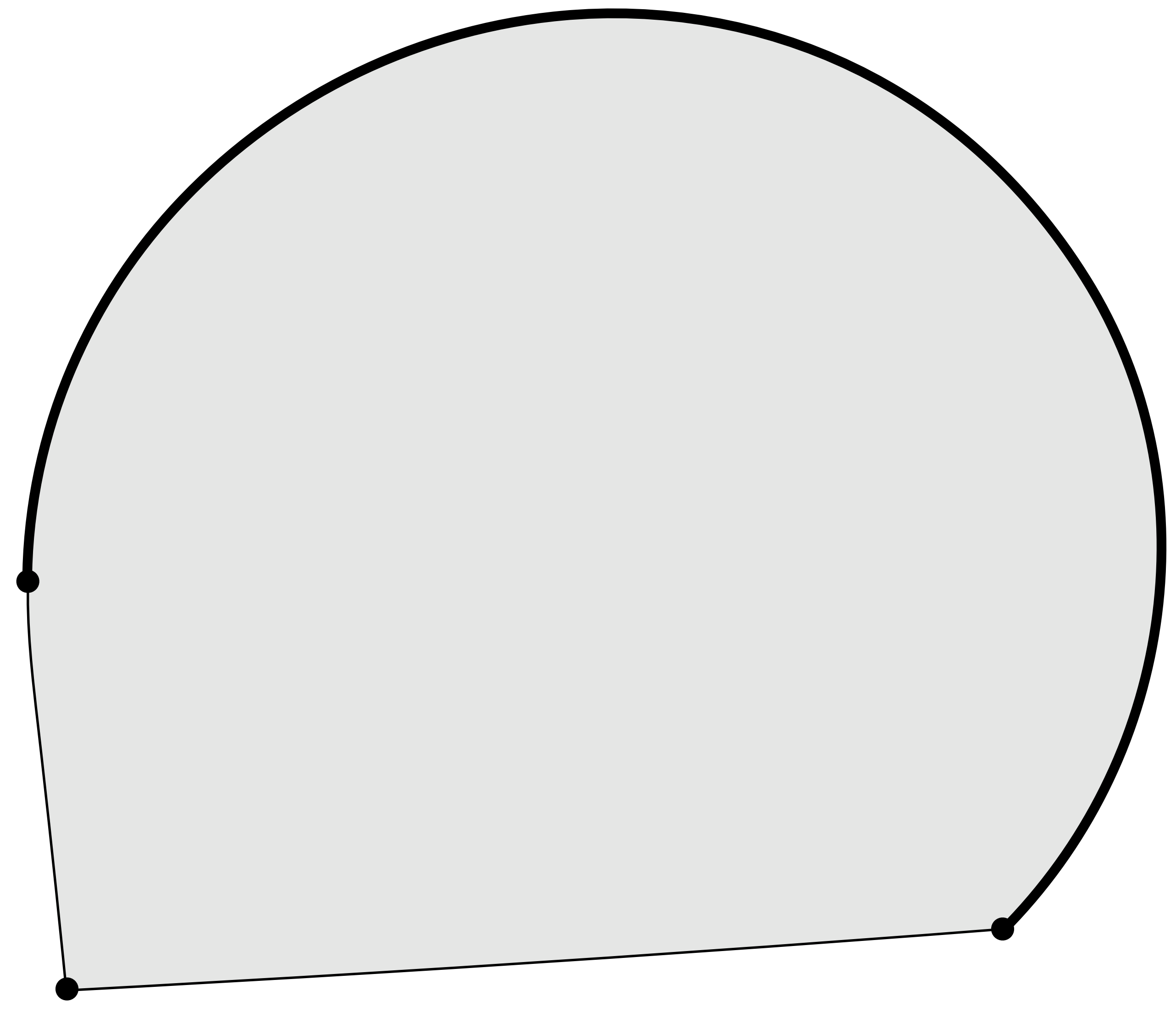
$x \in P$ is said to be an **extreme point** of P if

$\nexists y, z \in P \; (y \neq x, z \neq x)$ and $\alpha \in [0, 1]$ such that $x = \alpha y + (1 - \alpha) z$



Extreme points

Convex sets



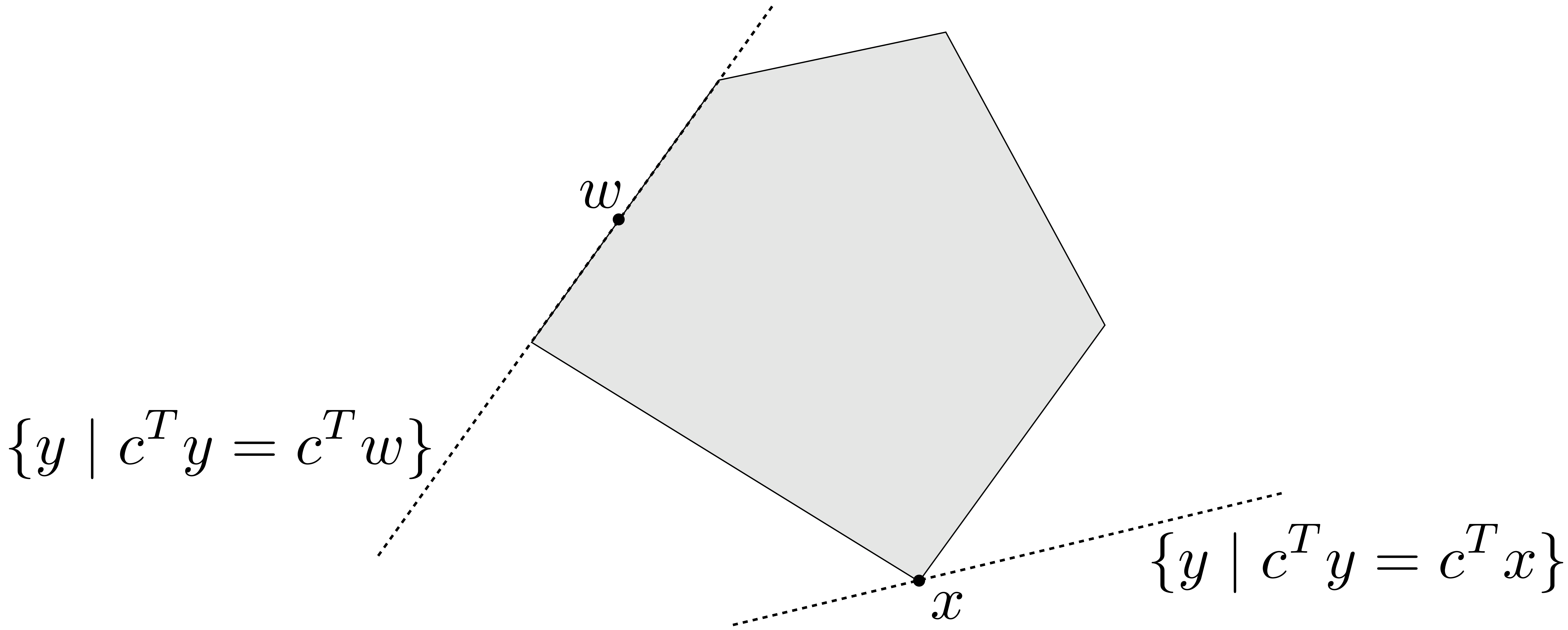
- Convex sets can have an infinite number of extreme points
- Polyhedra are convex sets with a finite number of extreme points

Vertices

Definition

$x \in P$ is a **vertex** if $\exists c$ such that x is the unique optimum of

$$\begin{array}{ll} \text{minimize} & c^T y \\ \text{subject to} & y \in P \end{array}$$



Basic feasible solution

$$P = \{x \mid a_i^T x \leq b_i, \quad i = 1, \ldots, m\}$$

Active constraints at $\bar{x}$

$$\mathcal{I}(\bar{x}) = \{i \in \{1, \ldots, m\} \mid a_i^T \bar{x} = b_i\}$$

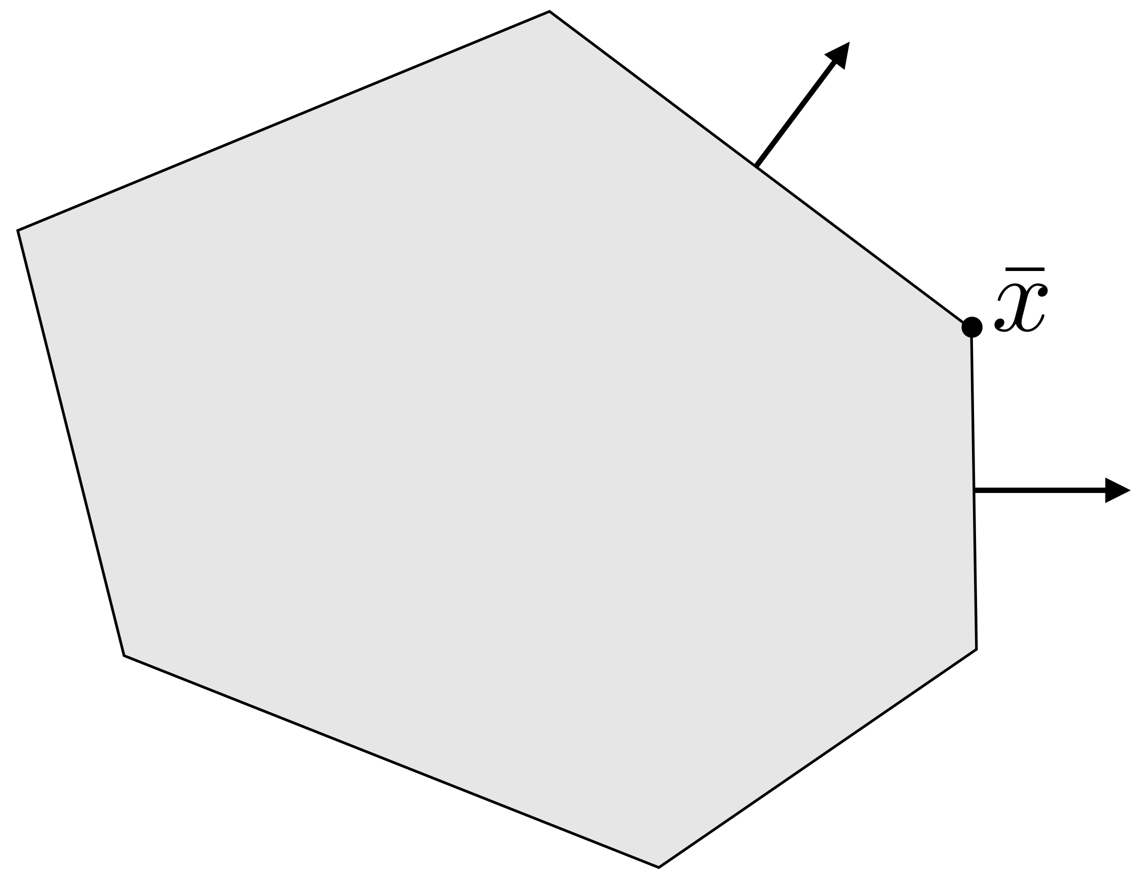
Index of all the constraints satisfied as **equality**

Basic solution $\bar{x}$

- $\{a_i \mid i \in \mathcal{I}(\bar{x})\}$ has n linearly independent vectors

Basic feasible solution $\bar{x}$

- $\bar{x} \in P$
- $\{a_i \mid i \in \mathcal{I}(\bar{x})\}$ has n linearly independent vectors

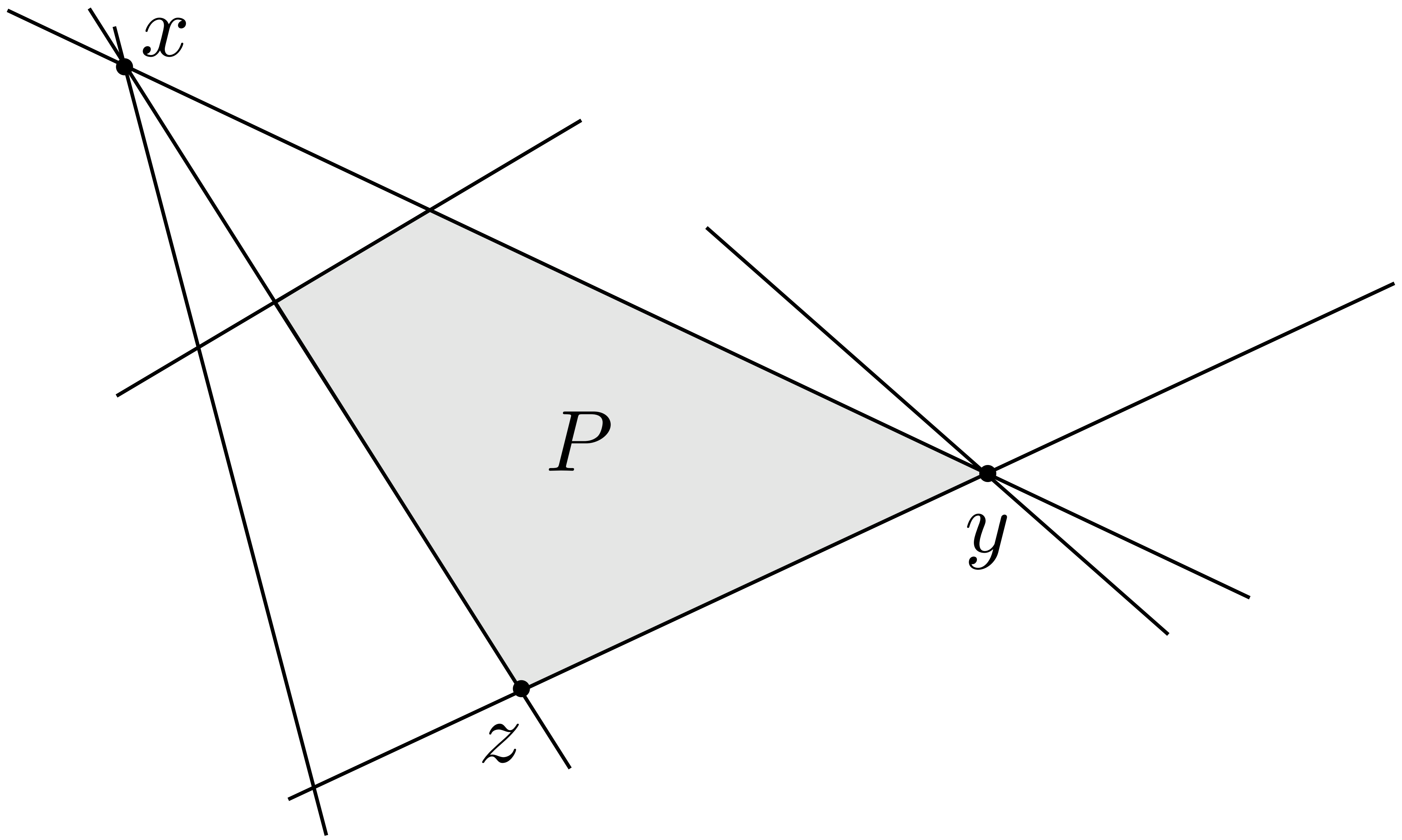


Degenerate basic feasible solutions

A solution $\bar{x}$ is degenerate if $|\mathcal{I}(\bar{x})| > n$

True or False?

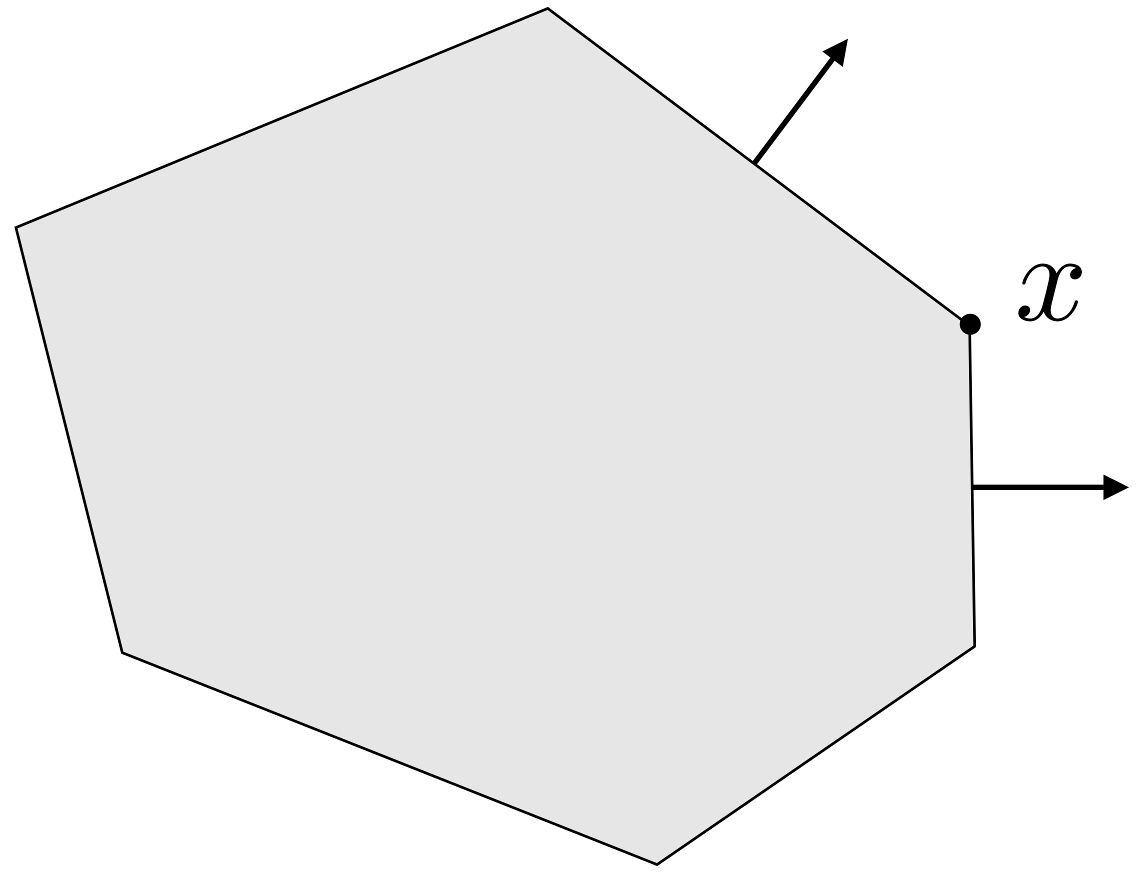
	Basic	Feasible	Degenerate
x			
y			
z			



Equivalence

Theorem

Given a nonempty polyhedron $P = \{x \mid Ax \leq b\}$



Let $x \in P$

x is a **vertex** $\iff$ x is an **extreme point** $\iff$ x is a **basic feasible solution**

Equivalent theorem proof

Vertex —> Extreme point

If x is a vertex, $\exists c$ such that $c^T x < c^T y, \quad \forall y \in P, y \neq x$

Let's assume x is not an extreme point:

$\exists y, z \neq x$ such that $x = \lambda y + (1 - \lambda) z$

Since x is a vertex, $c^T x < c^T y$ and $c^T x < c^T z$

Therefore, $c^T x = \lambda c^T y + (1 - \lambda) c^T z > \lambda c^T x + (1 - \lambda) c^T x = c^T x$

$\implies$ **contradiction** ■

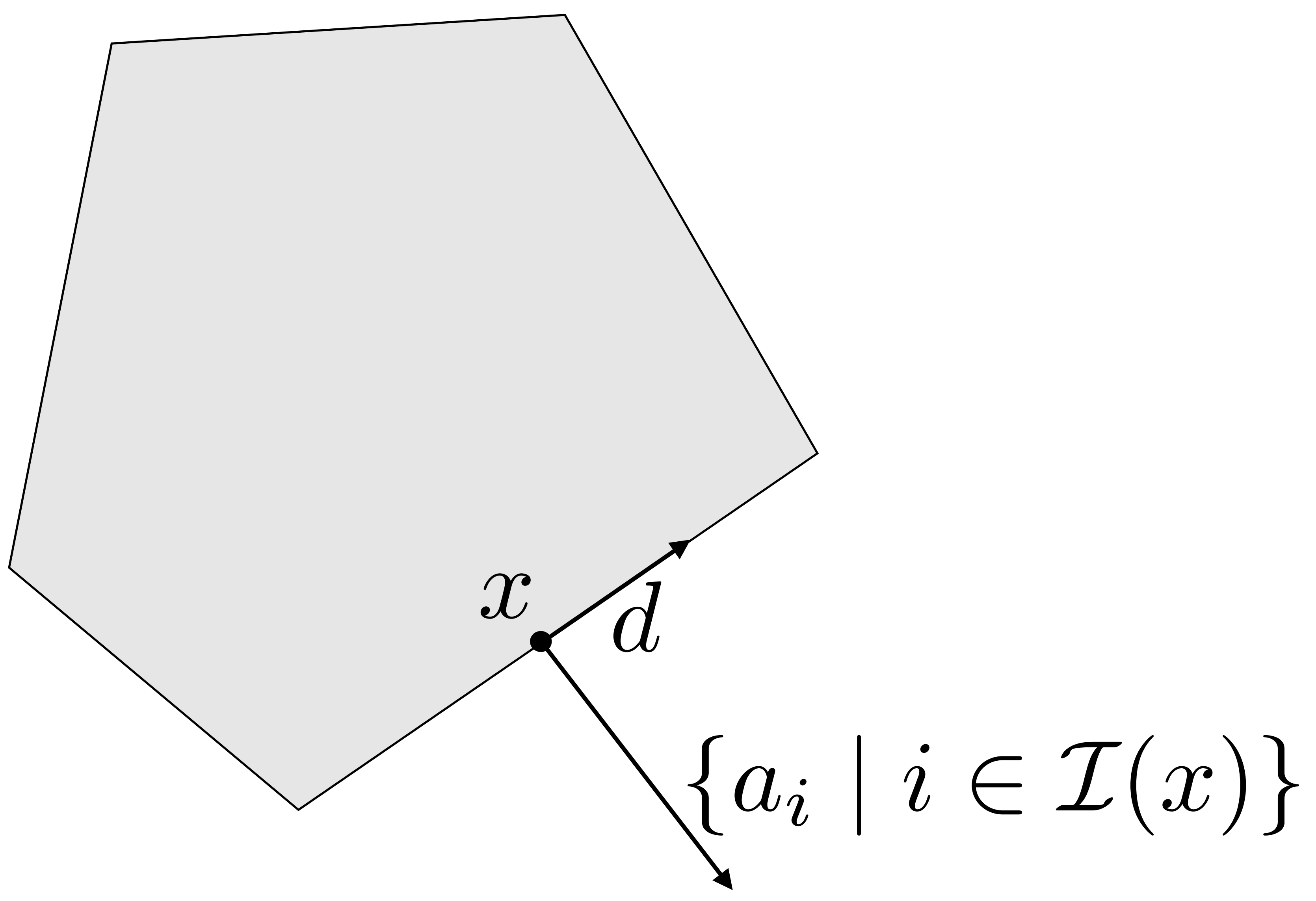
Equivalent theorem proof

Extreme point —> Basic feasible solution (proof by **contraposition**)

Suppose $x \in P$ is **not basic feasible solution**

$\{a_i \mid i \in \mathcal{I}(x)\}$ does not span $\mathbf{R}^n$

$\exists d \in \mathbf{R}^n$ perpendicular to all of them: $a_i^T d = 0, \quad \forall i \in \mathcal{I}(x)$



Equivalent theorem proof

Extreme point —> Basic feasible solution (proof by **contraposition**)

Suppose $x \in P$ is **not basic feasible solution**

$\{a_i \mid i \in \mathcal{I}(x)\}$ does not span $\mathbf{R}^n$

$\exists d \in \mathbf{R}^n$ perpendicular to all of them: $a_i^T d = 0, \quad \forall i \in \mathcal{I}(x)$

Let $\epsilon > 0$ and define $y = x + \epsilon d$ and $z = x - \epsilon d$

For $i \in \mathcal{I}(x)$ we have $a_i^T y = b_i$ and $a_i^T z = b_i$

For $i \notin \mathcal{I}(x)$ we have $a_i^T x < b_i \quad \Rightarrow \quad a_i^T(x + \epsilon d) < b_i$ and $a_i^T(x - \epsilon d) < b_i$

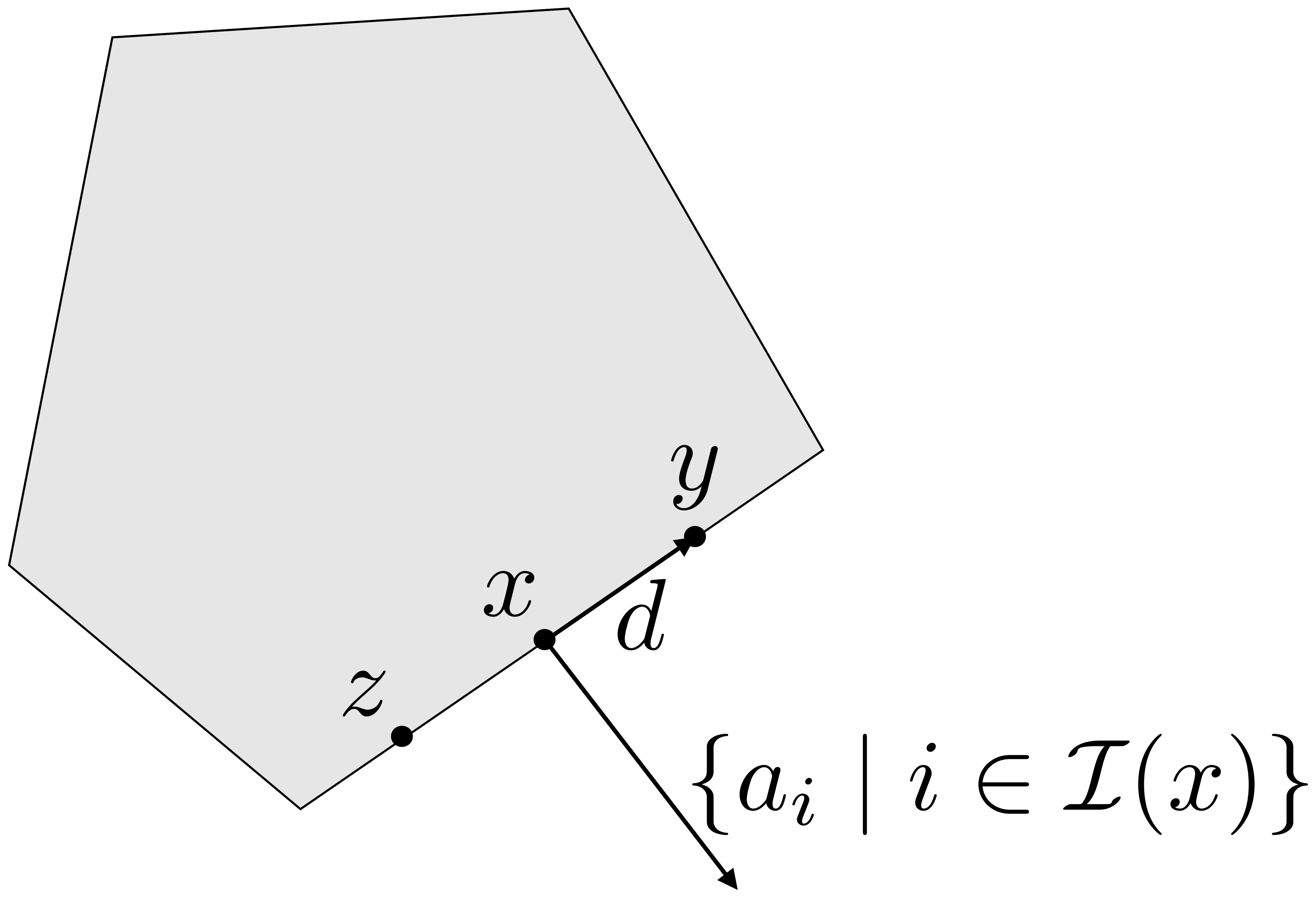
Hence, $y, z \in P$ and $x = \lambda y + (1 - \lambda) z$ with $\lambda = 0.5$.

$\Longrightarrow x$ is **not an extreme point** ■

Equivalent theorem proof

Extreme point —> Basic feasible solution (proof by **contraposition**)

Suppose $x \in P$ is **not basic feasible solution**



Hence, $y, z \in P$ and $x = \lambda y + (1 - \lambda)z$ with $\lambda = 0.5$.

$\Longrightarrow x$ is **not an extreme point** ∎

Equivalent theorem proof

Basic feasible solution —> Vertex

Left as exercise

Hint

Define $c = -\sum_{i \in \mathcal{I}(x)} a_i$

Constructing basic solutions

Standard form polyhedra

Definition

Standard form LP

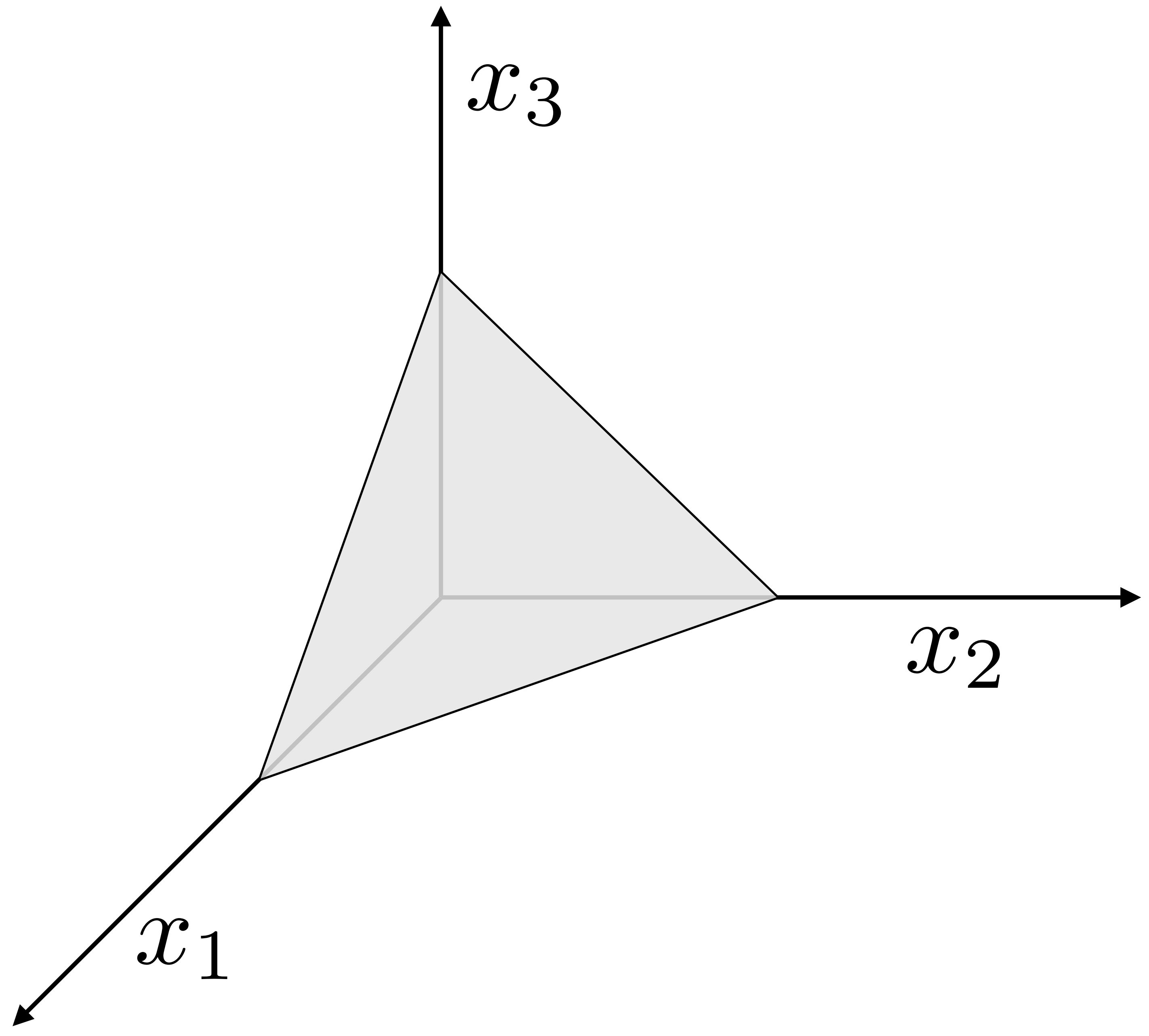
minimize $\quad c^T x$

subject to $\quad Ax = b$

$\quad x \geq 0$

Standard form polyhedron

$$P = \{x \mid Ax = b,\ x \geq 0\}$$



Assumption

$A \in \mathbf{R}^{m \times n}$ has full row rank $m \leq n$

Interpretation

P lives in $(n - m)$-dimensional subspace

Basic solutions

Standard form polyhedra

$P = \{x \mid Ax = b,\ x \geq 0\}$ with $A \in \mathbf{R}^{m \times n}$ has full row rank $m \leq n$

x is a **basic solution** if and only if

- $Ax = b$
- There exist indices $B(1), \dots, B(m)$ such that
  - columns $A_{B(1)}, \dots, A_{B(m)}$ are linearly independent
  - $x_i = 0$ for $i \neq B(1), \dots, B(m)$

x is a **basic feasible solution** if x is a **basic solution** and $x \geq 0$

Constructing basic solution

1. Choose any m independent columns of A: $A_{B(1)}, \ldots, A_{B(m)}$
2. Let $x_i = 0$ for all $i \neq B(1), \ldots, B(m)$
3. Solve $Ax = b$ for the remaining $x_{B(1)}, \ldots, x_{B(m)}$

Basis matrix

Basis columns

Basic variables

$$A_B = \begin{bmatrix} | & | & & | \\ A_{B(1)} & A_{B(2)} & \ldots & A_{B(m)} \\ | & | & & | \end{bmatrix}, \quad x_B = \begin{bmatrix} x_{B(1)} \\ \vdots \\ x_{B(m)} \end{bmatrix} \longrightarrow \text{Solve } A_B x_B = b$$

If $x_B \geq 0$, then x is a **basic feasible solution**

Finding a basic solution

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 2 & -1 & -3 & 0 & 0 \\ 0 & 2 & 8 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 6 \end{bmatrix}$$

$A_{B(1)}$ $A_{B(2)}$ $A_{B(3)}$

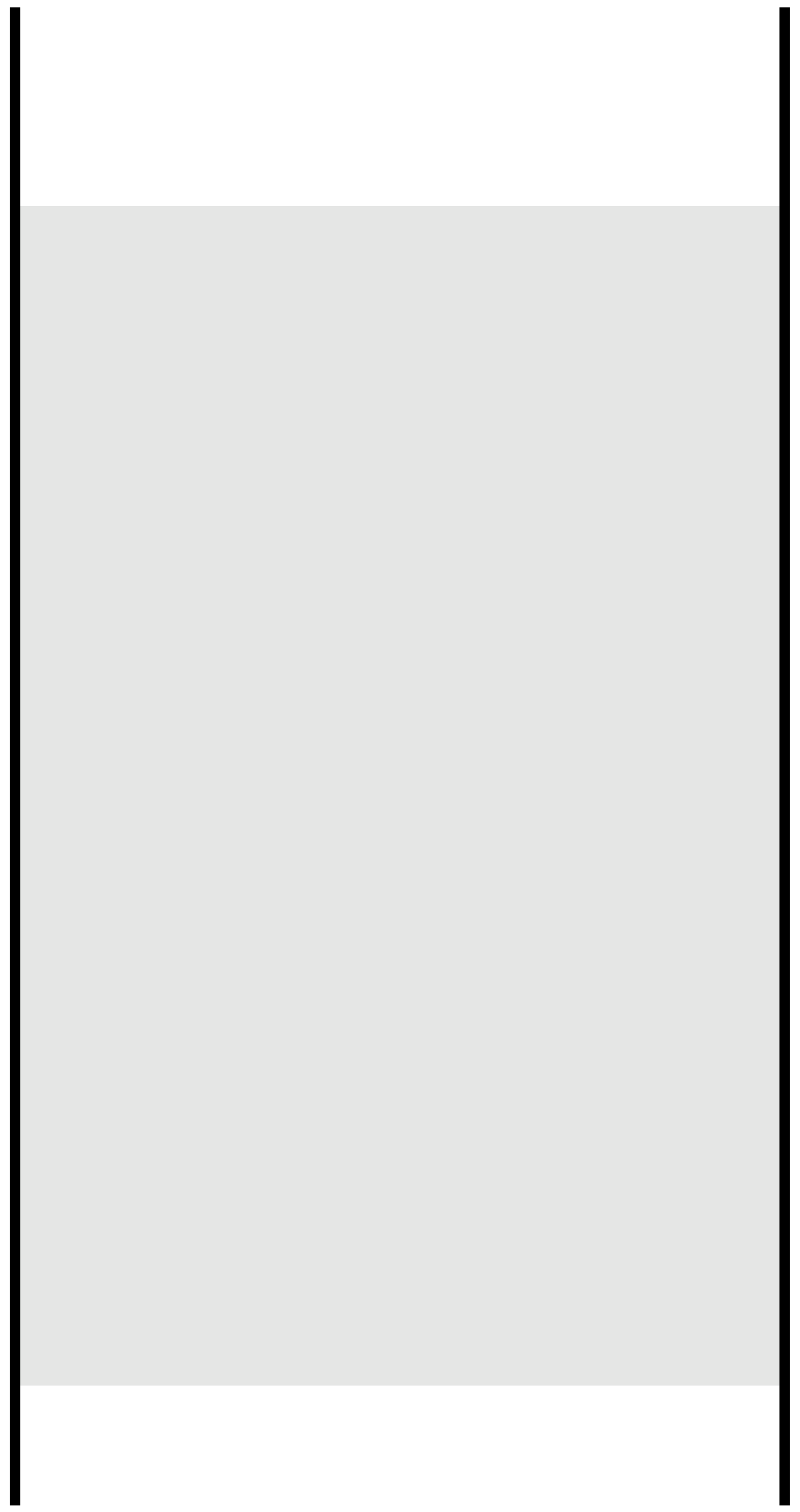
Solve $$\begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 2 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_2 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 6 \end{bmatrix}$$

$$x_B = \begin{bmatrix} x_2 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \geq 0$$

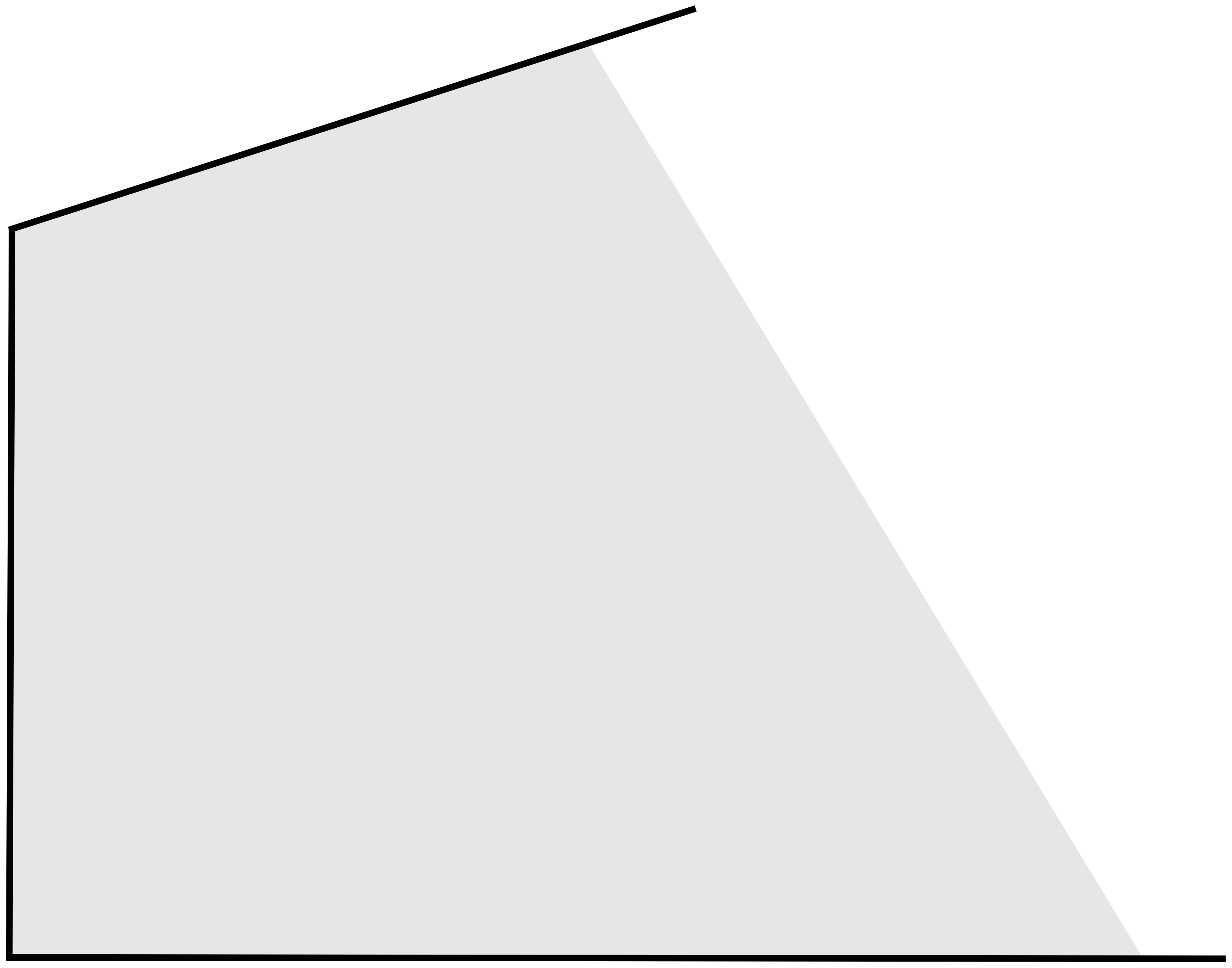
Existence and optimality of extreme points

Existence of extreme points

Example



No extreme points



Extreme points

Existence of extreme points

Characterization

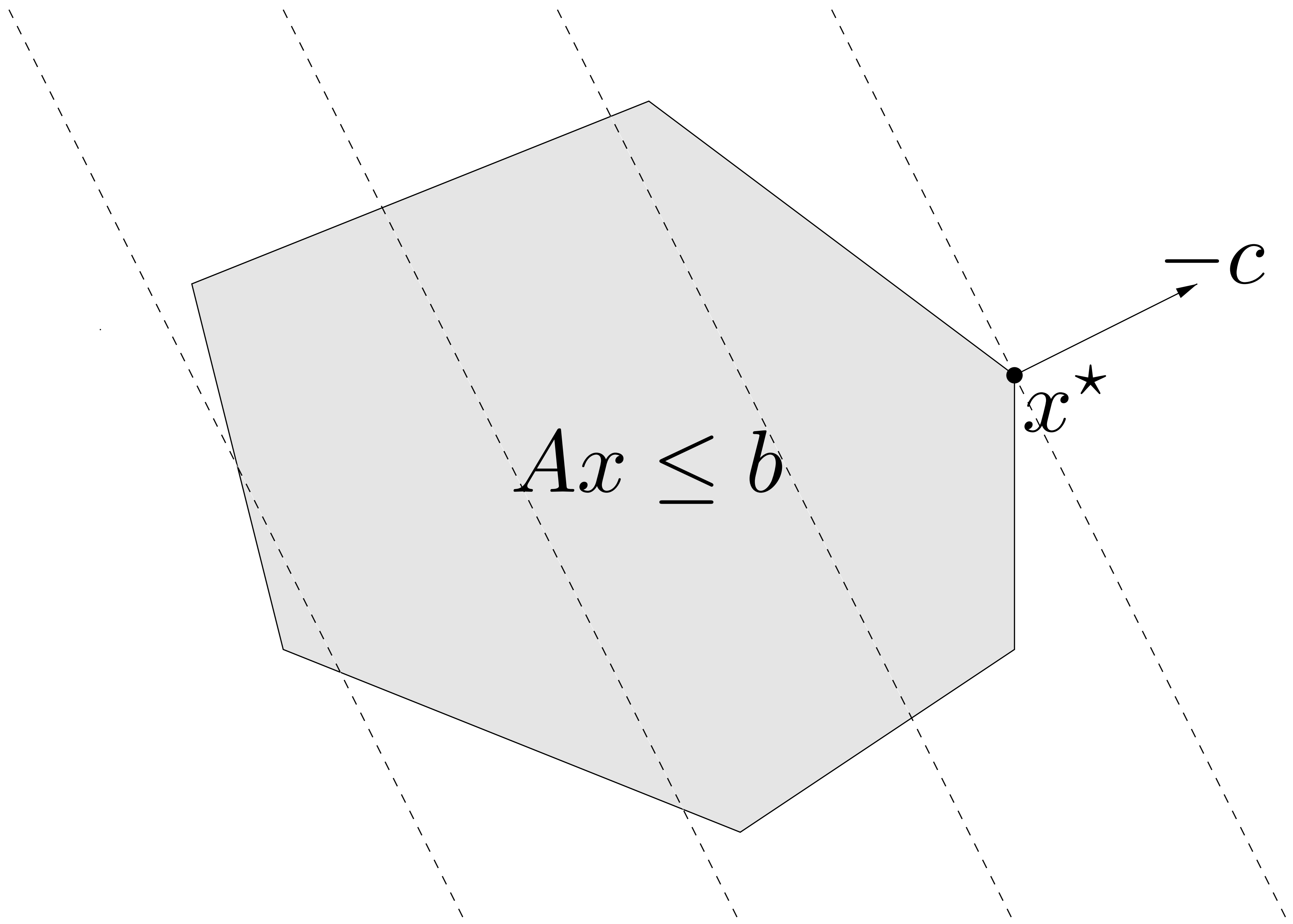
A polyhedron P **contains a line** if
$\exists x \in P$ and a nonzero vector d such that $x + \lambda d \in P, \forall \lambda \in \mathbf{R}$.

Given a polyhedron $P = \{x \mid a_i^T x \leq b_i, \quad i = 1, \dots, m\}$, the following are **equivalent**

- P does not contain a line
- P has at least one extreme point
- n of the a_i vectors are linearly independent

Corollary
Every nonempty **bounded polyhedron** has
at least one basic feasible solution

Optimality of extreme points



$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & Ax \leq b
\end{array}
$$

If
- P has at least one extreme point
- There exists an optimal solution $x^\star$

Then, there exists an optimal solution which is an **extreme point** of P

We only need to search between **extreme points**

How to search among basic feasible solutions?

Idea

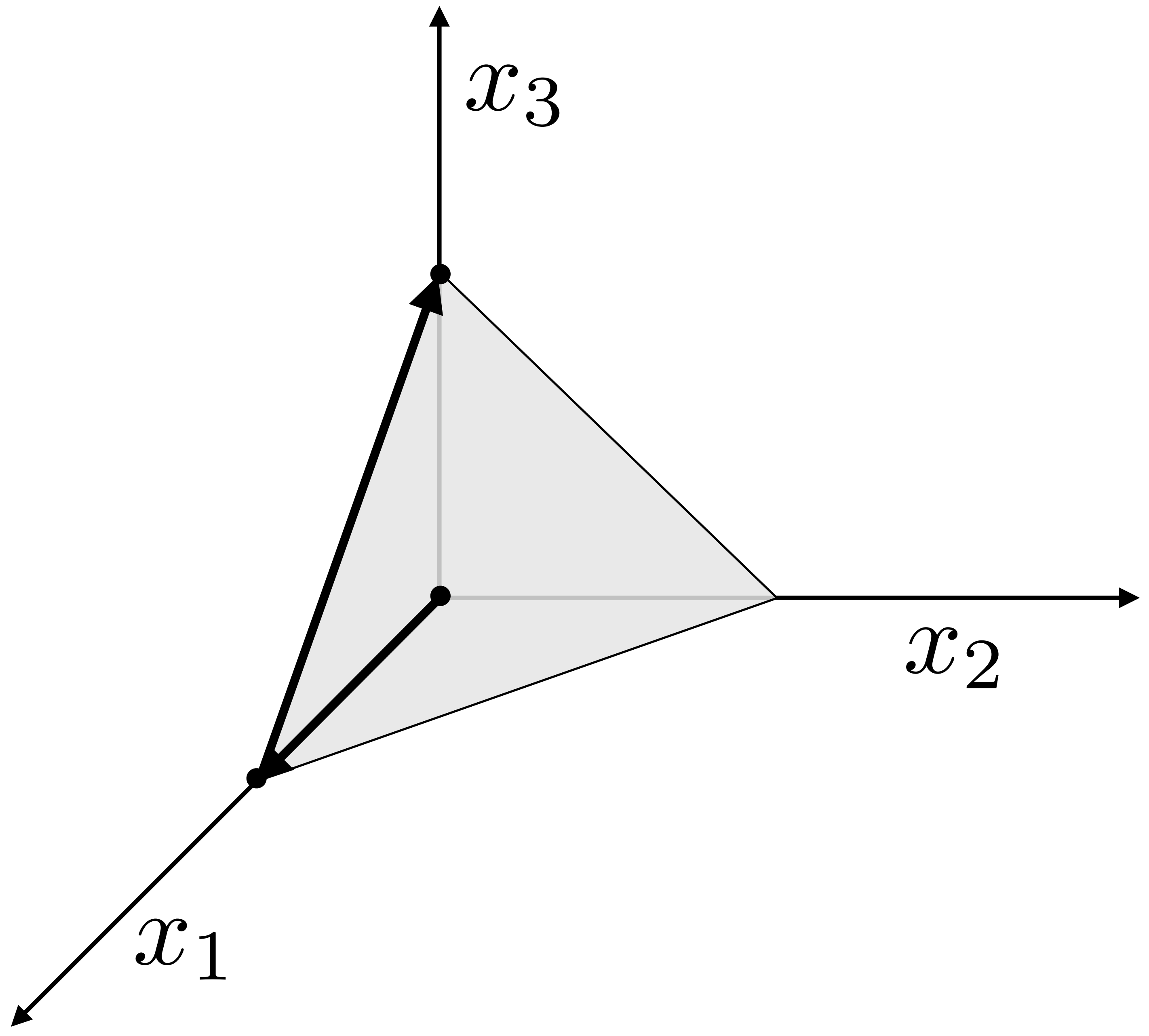
List all the basic feasible solutions, compare objective values and pick the best one.

Intractable!

If $n = 1000$ and $m = 100$, we have 10^{143} combinations!

Conceptual algorithm

- Start at corner
- Visit neighboring corner that improves the objective



Geometry of linear optimization

Today, we learned to:

- **Apply geometric and algebraic properties** of polyhedra to characterize the “corners” of the feasible region.
- **Construct basic feasible solutions** by solving a linear system.
- **Recognize existence and optimality** of extreme points.

Next lecture

The simplex method

- Iterations
- Convergence
- Complexity