ORF522 – Linear and Nonlinear Optimization

4. The simplex method

Bartolomeo Stellato — Fall 2020

Ed forum

- Basic feasible solutions in geometric vs algebraic form (next slides)
- More efficient transformation methods from geometric to standard form when there is structure? (Pre-processing + do not need to calculate all extreme points)
- Do equality constraints in geometric form correspond to two linearly dependent inequalities?
- Equivalence proofs between corners (next slides)
- Definition of contain a line (Typo!)
- How do we start if initial solution is infeasible?
- Jupyter notebook: only pdf or also ipdb? Only pdf.
- Video/audio not in sync.

Recap

Standard form polyhedra

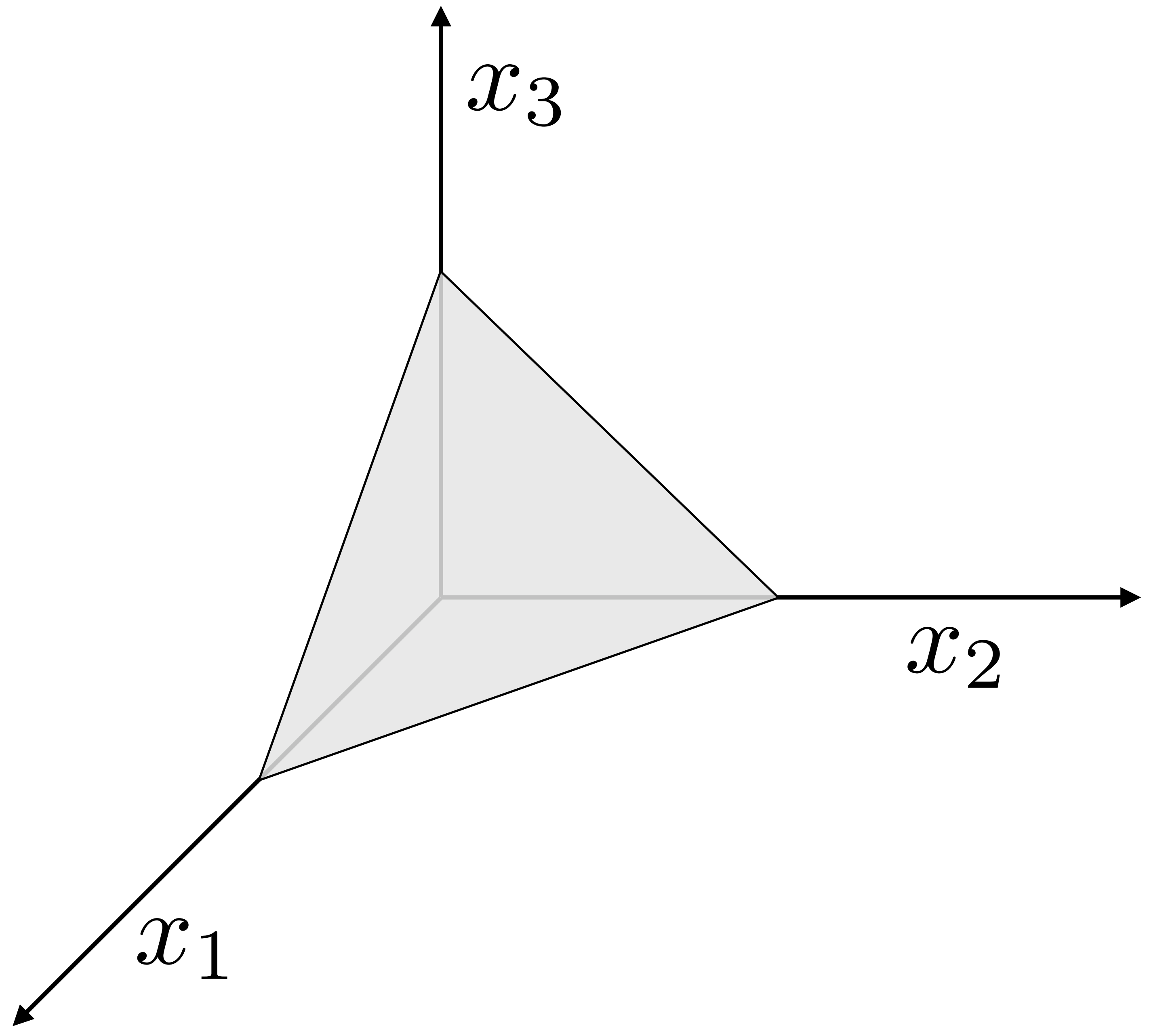
Definition

Standard form LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array}$$

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

Transformation to standard form

minimize $c^T x$
subject to $Ax \leq b$ $\longrightarrow$

minimize $c^T(x^+ - x^-)$
subject to $\begin{bmatrix} A & -A & I \end{bmatrix} \begin{bmatrix} x^+ \\ x^- \\ s \end{bmatrix} = b$
$(x^+, x^-, s) \geq 0$ $\longrightarrow$

minimize $\tilde{c}^T \tilde{x}$
subject to $\tilde{A}\tilde{x} = b$
$\tilde{x} \geq 0$

Variables: $\tilde{n} = 2n + m$
(Equality) constraints: $\tilde{m} = m$

There are $\tilde{m}$ active constraints
We need $\tilde{n} - \tilde{m} = 2n$ inequalities active $\Rightarrow \tilde{x}_i = 0$ (non basic)

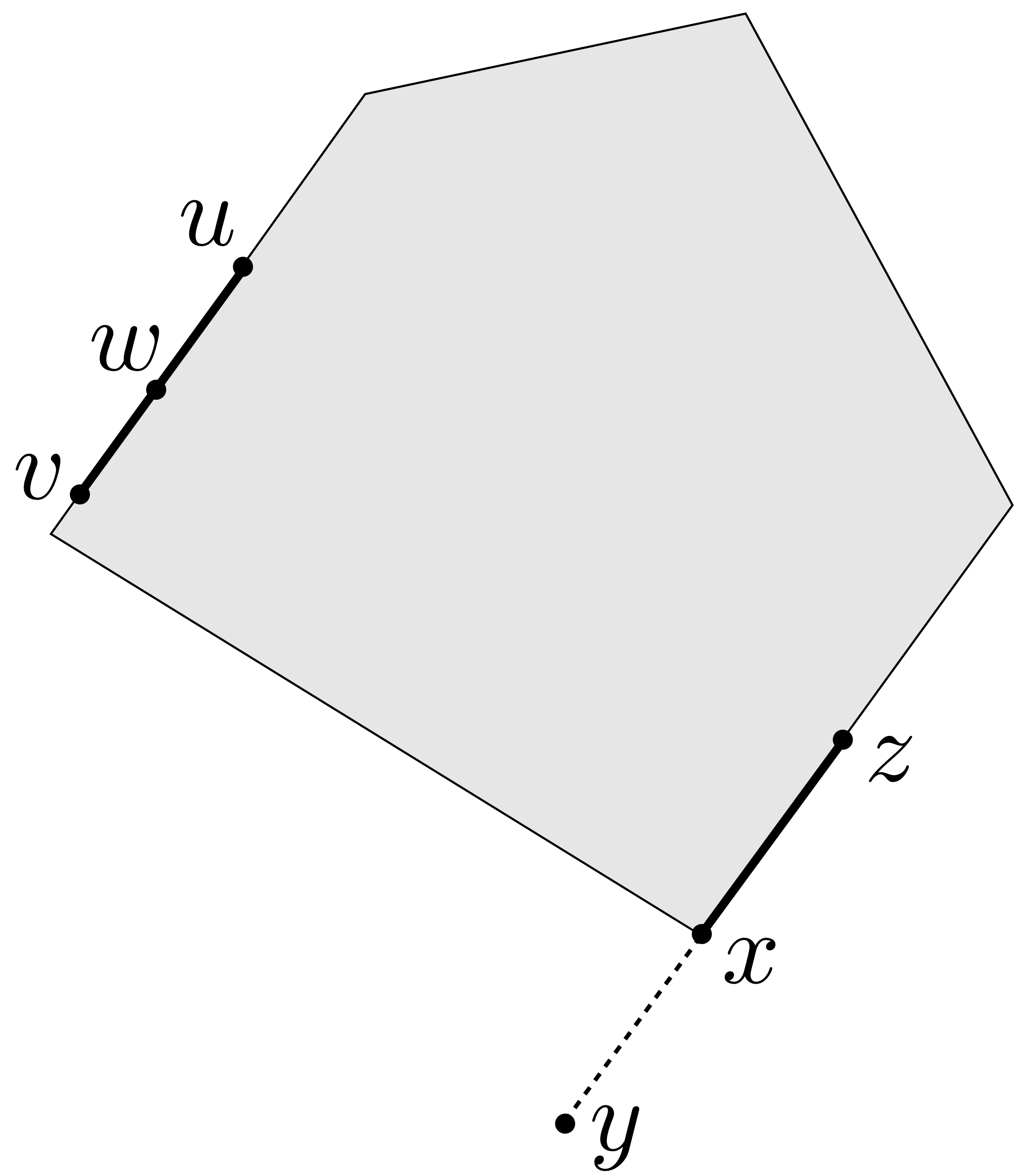
Which corresponds to m inequalities inactive $\Rightarrow \tilde{x}_i > 0$ (basic)

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 $x = \alpha y + (1-\alpha) z$



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

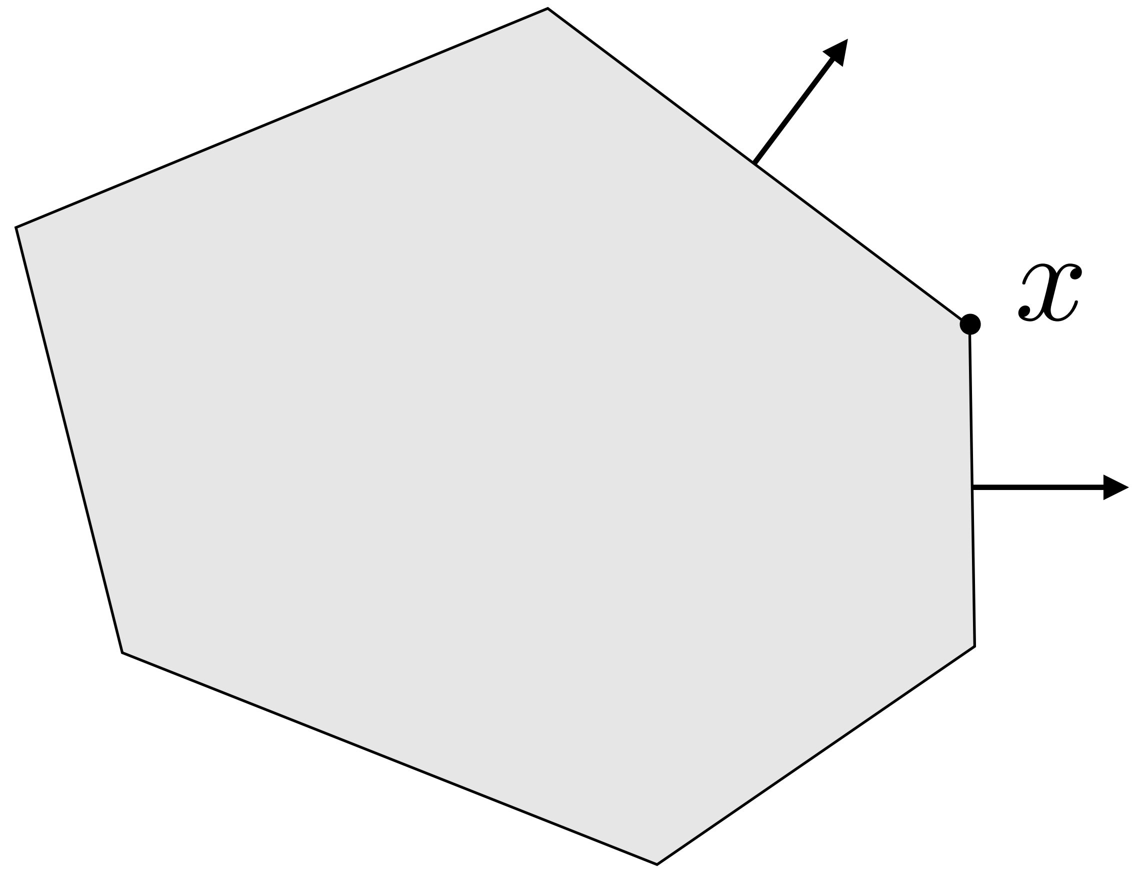
$$B = \begin{bmatrix} | & | & & | \\ A_{B(1)} & A_{B(2)} & \ldots & A_{B(m)} \\ | & | & & | \end{bmatrix}, \quad x_B = \begin{bmatrix} x_{B(1)} \\ \vdots \\ x_{B(m)} \end{bmatrix} \longrightarrow \quad x_B = B^{-1}b$$

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

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 $x = \lambda y + (1 - \lambda) z$

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

Therefore, $c^T x = \lambda c^T y + (1 - \lambda) c^T z > \lambda c^T x + (1 - \lambda) c^T x = c^T x$: **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: $a_i^T d = 0, \quad \forall i \in \mathcal{I}(x)$

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 $\lambda = 0.5$.

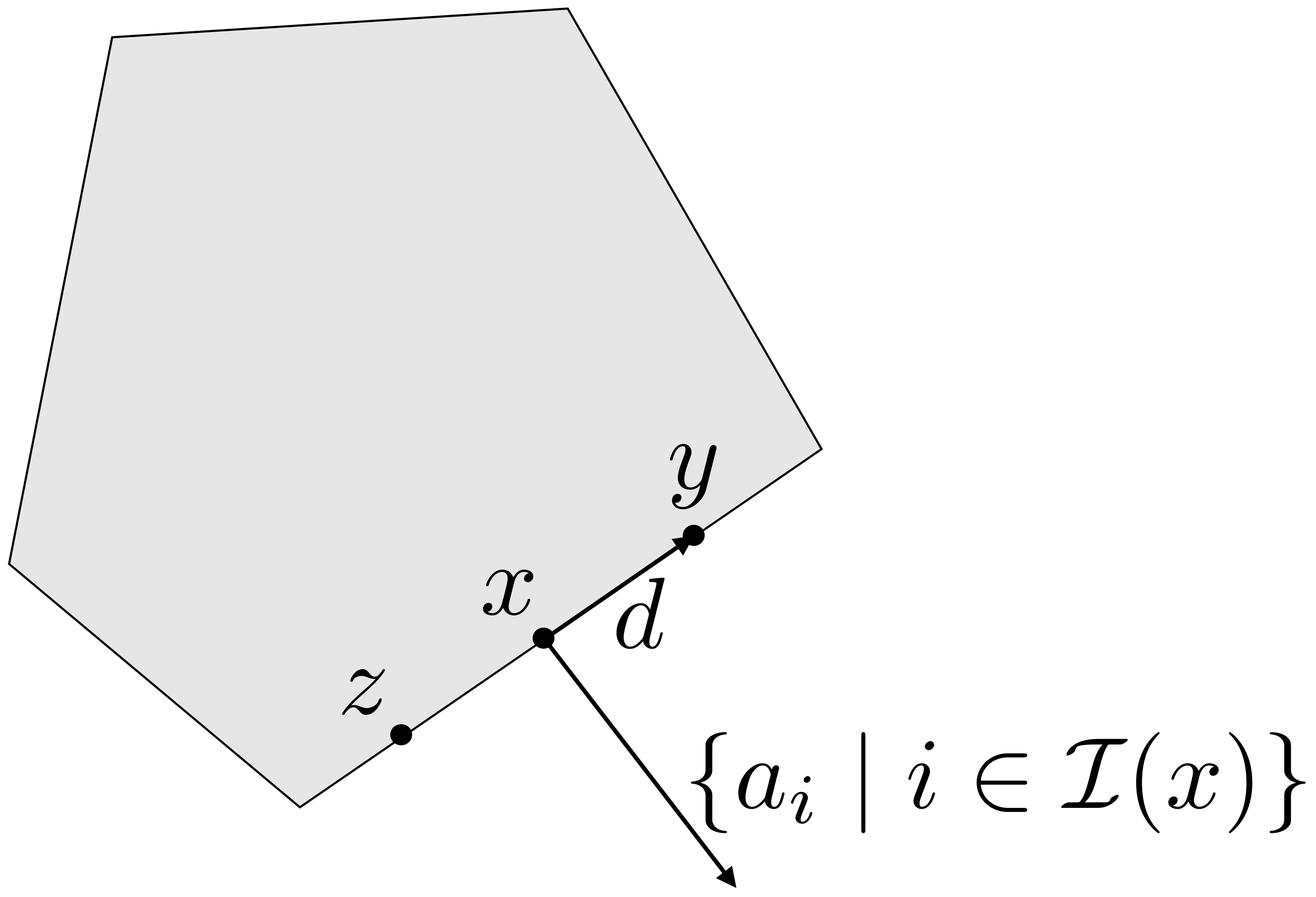
x is **not an extreme point** $\blacksquare$

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 $\lambda = 0.5$.
x is **not an extreme point** ∎

Equivalent theorem proof

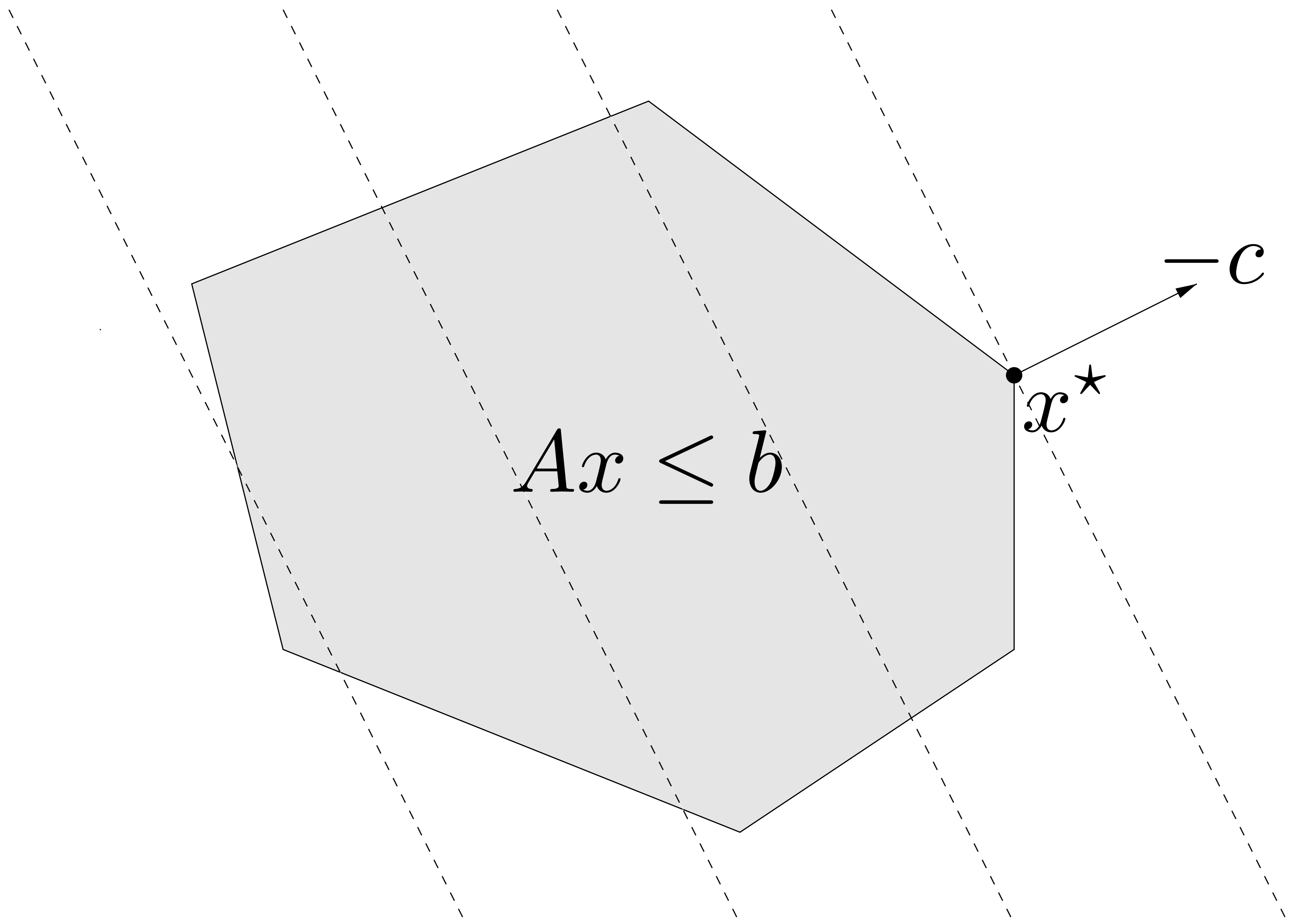
Basic feasible solution —> Vertex

Left as exercise

Hint

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

Optimality of extreme points



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

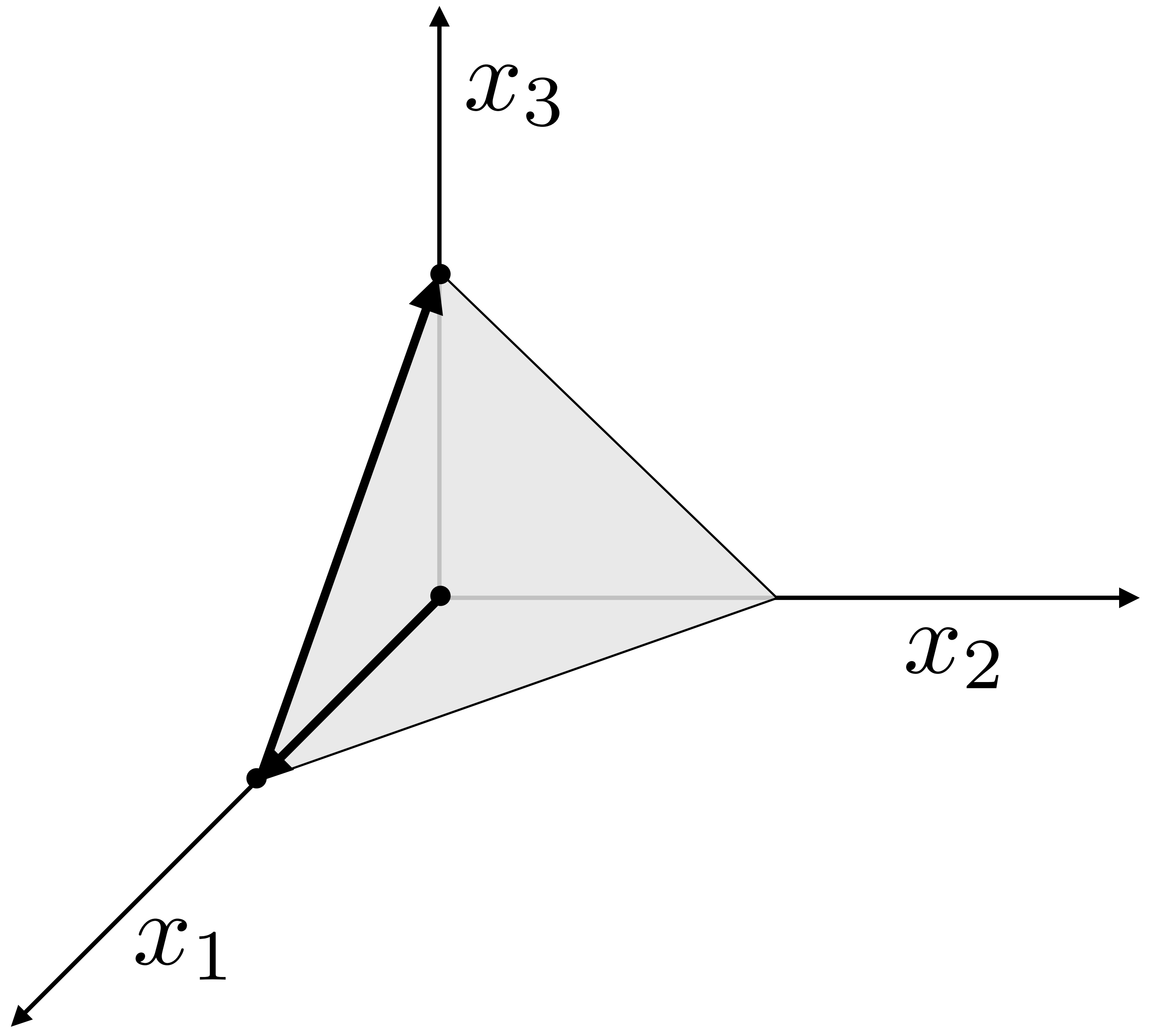
If
- P has at least one extreme point
- There exists an optimal solution $x^\star$

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

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

Conceptual algorithm

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



Today's agenda

Readings: [Chapter 3, Bertsimas and Tsitsiklis]

Simplex method

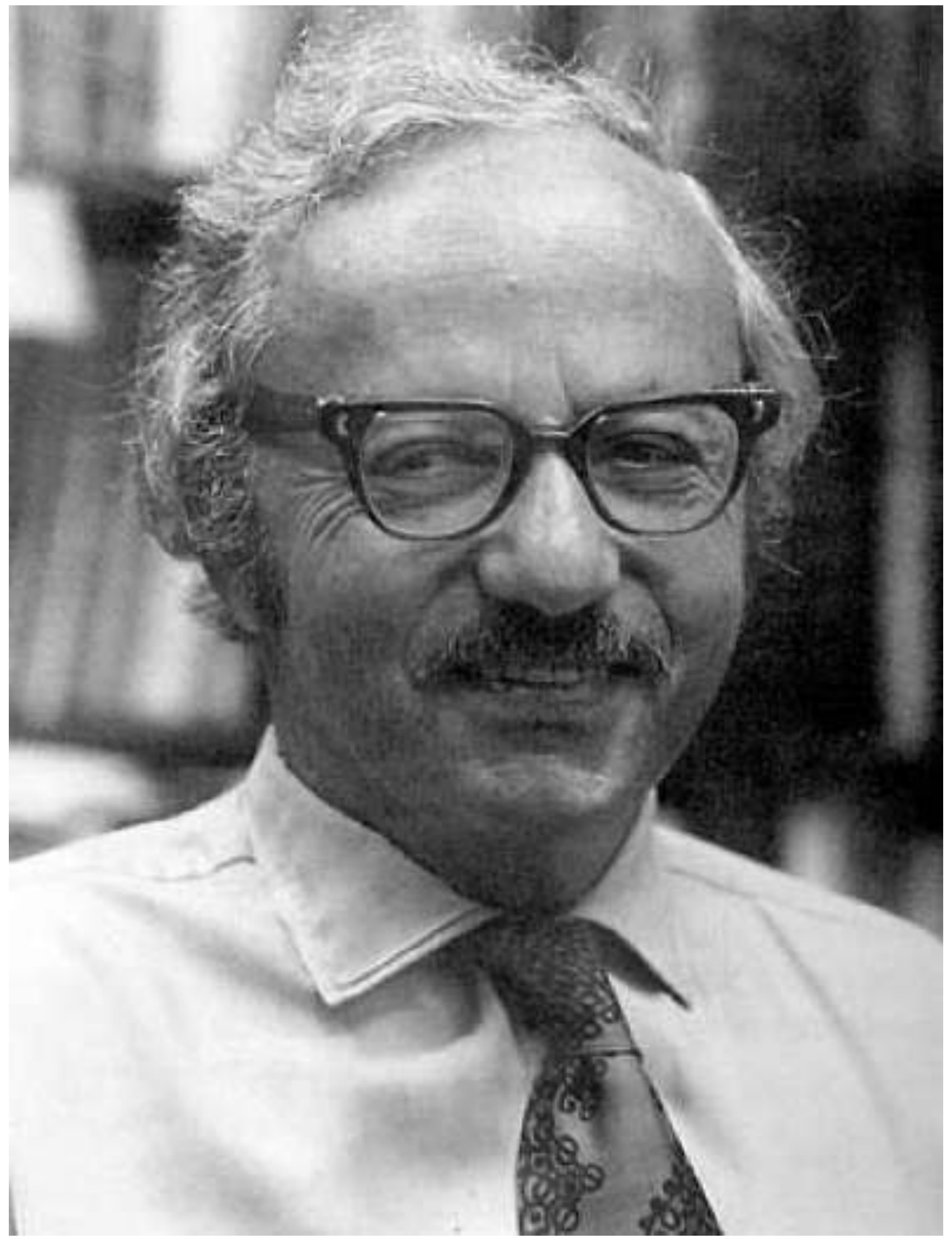
- Iterate between neighboring basic solutions
- Optimality conditions
- Simplex iterations

The simplex method

Top 10 algorithms of the 20th century

1946: Metropolis algorithm
1947: Simplex method
1950: Krylov subspace method
1951: The decompositional approach to matrix computations
1957: The Fortran optimizing compiler
1959: QR algorithm
1962: Quicksort
1965: Fast Fourier transform
1977: Integer relation detection
1987: Fast multipole method

George Dantzig



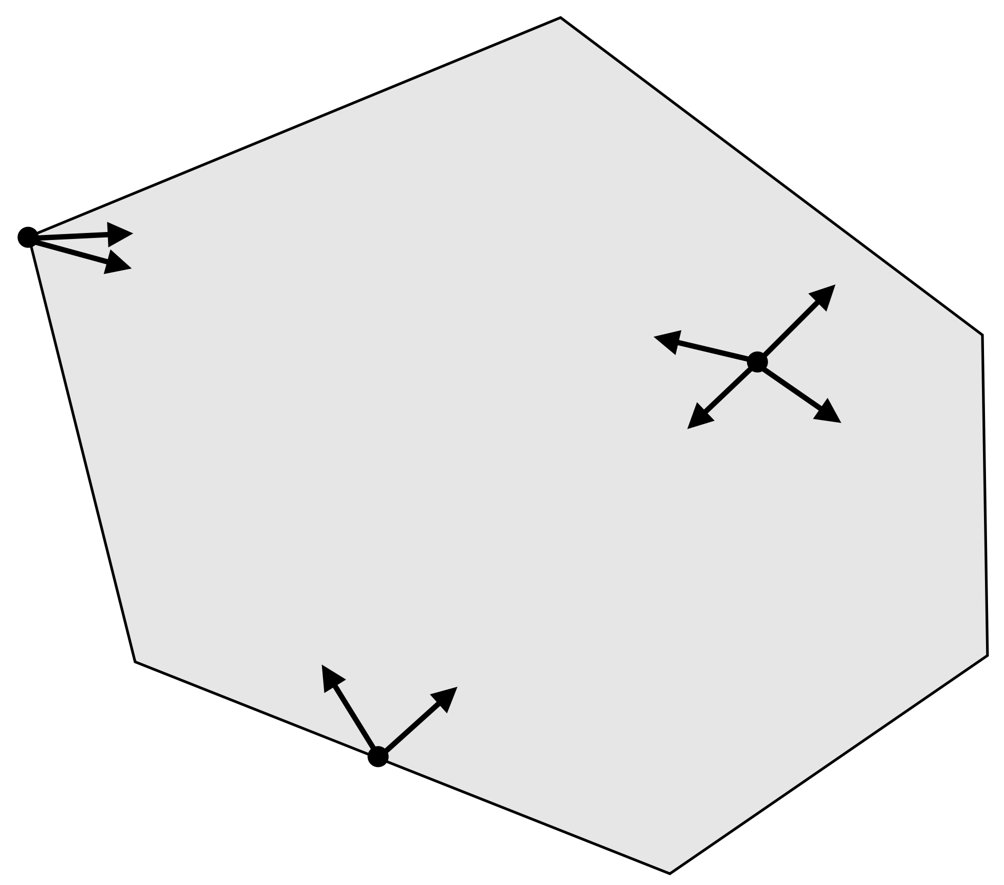
[SIAM News (2000)]

Neighboring basic solutions

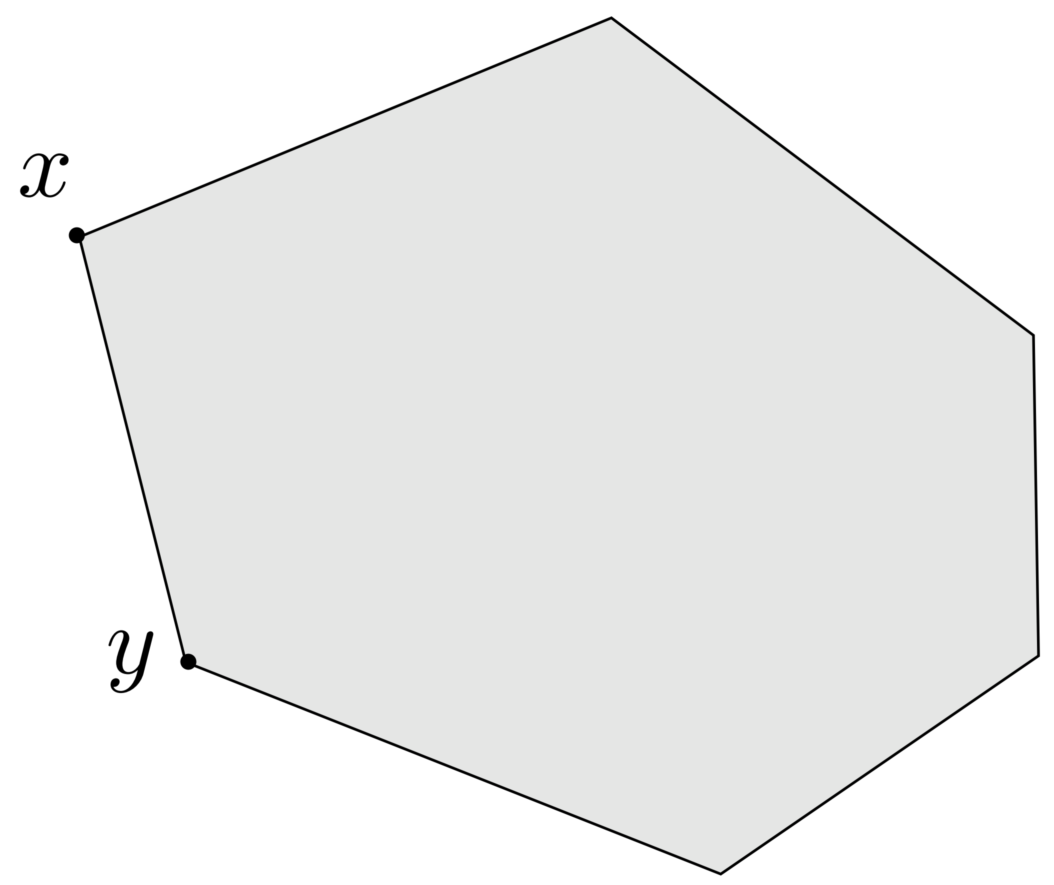
Feasible directions and neighboring solutions

Definition

Let $x \in P$, a vector d is a **feasible direction** at x if $\exists \theta > 0$ for which $x + \theta d \in P$



Two basic solutions are **neighboring** if their basic indices differ by exactly one variable



Feasible directions

Conditions

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

Given a basis matrix $B = \begin{bmatrix} A_{B(1)} & \dots & A_{B(m)} \end{bmatrix}$

we have basic feasible solution x:

- $x_B = B^{-1}b$
- $x_i = 0,\ \forall i \neq B(1), \dots, B(m)$

Feasible direction d

- $A(x + \theta d) = b \implies Ad = 0$
- $x + \theta d \geq 0$

Feasible directions

Computation

Nonbasic indices

- $d_j = 1 \longrightarrow$ **Basic direction**
- $d_k = 0, \ \forall k \notin \{j, B(1), \ldots, B(m)\}$

Basic indices

$$Ad = 0 = \sum_{i=1}^{n} A_i d_i = B d_B + A_j = 0 \Longrightarrow d_B = -B^{-1} A_j$$

Non-negativity (non-degenerate assumption)

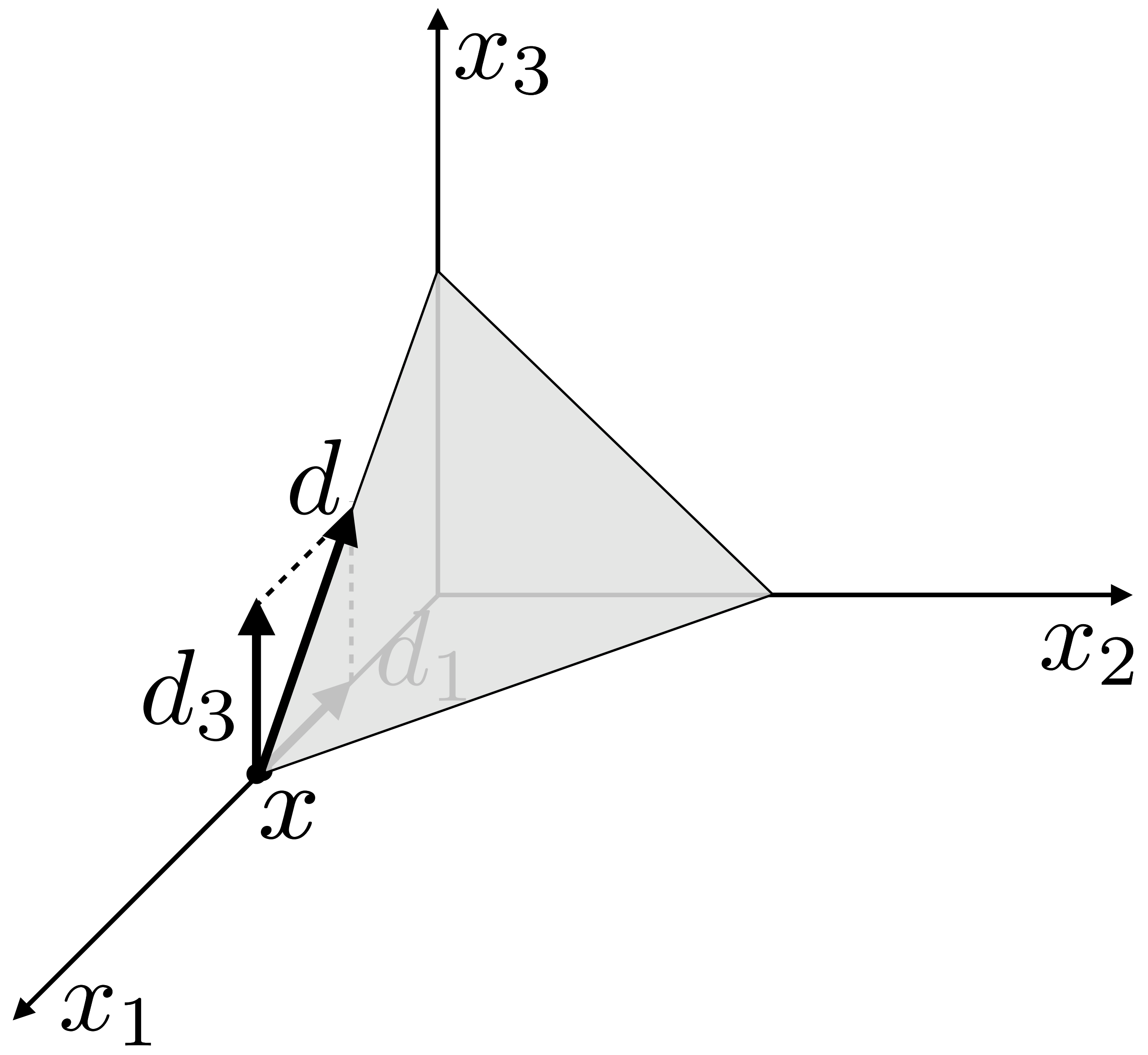
- Non-basic variables: $x_i = 0$. Nonnegative direction $d_i \geq 0$
- Basic variables: $x_B > 0$. Therefore $\exists \theta > 0$ such that $x_B + \theta d_B \geq 0$

Feasible directions

Example

$$P = \{x \mid x_1 + x_2 + x_3 = 2, \quad x \geq 0\}$$

$$x = (2, 0, 0) \qquad\qquad B = \{1\}$$



Basic index $j = 3 \longrightarrow d = (-1, 0, 1)$

$$d_j = 1$$
$$d_B = -B^{-1}A_j$$

How does the cost change?

The **new cost** is $c^T(x + \theta d)$

The **cost improvement** is $c^T(x + \theta d) - c^T x = \theta c^T d$

We call $\bar{c}_j$ the **reduced cost** of (introducing) variable x_j

$$\bar{c}_j = c^T d = \sum_{i=1}^{n} c_j d_j = c_j + c_B^T d_B = c_j - c_B^T B^{-1} A_j$$

Reduced costs

Meaning

Change in objective/marginal cost of adding x_j to the basis

$$\bar{c}_j = \boxed{c_j} \boxed{- c_B^T B^{-1} A_j}$$

Cost per-unit increase of variable x_j

Cost to change other variables compensating for x_j to enforce $Ax = b$

- $c_j > 0$: adding x_j will increase the objective (bad)
- $c_j < 0$: adding x_j will decrease the objective (good)

Reduced costs for basic variables is 0

$$\bar{c}_{B(i)} = c_{B(i)} - c_B^T B^{-1} A_{B(i)} = c_{B(i)} - c_B^T e_i = c_{B(i)} - c_{B(i)} = 0$$

Optimality conditions

Optimality conditions

Theorem

Let x be a basic feasible solution associated with basis matrix B
Let $\bar{c}$ be the vector of reduced costs.

If $\bar{c} \geq 0$, then x is **optimal**

Remark

This is a **stopping criterion** for the simplex algorithm.

If the **neighboring solutions** do not improve the cost, we are done (because of convexity).

Optimality conditions

Proof

For a basic feasible solution x with basis matrix B the reduced costs are $\bar{c} \geq 0$.

Consider any feasible solution y and define $d = y - x$

Since x and y are feasible, then $Ax = Ay = b$ and $Ad = 0$

$$Ad = Bd_B + \sum_{i \in N} A_i d_i = 0 \quad \Rightarrow \quad d_B = -\sum_{i \in N} B^{-1} A_i d_i$$

N are the nonbasic indices

The change in objective is

$$c^T d = c_B^T d_B + \sum_{i \in N} c_i d_i = \sum_{i \in N} (c_i - c_B^T B^{-1} A_i) d_i = \sum_{i \in N} \bar{c}_i d_i$$

Since $y \geq 0$ and $x_i = 0,\ i \in N$, then $d_i = y_i - x_i \geq 0, i \in N$

$$c^T d = c^T (y - x) \geq 0 \quad \Rightarrow \quad c^T y \geq c^T x.$$

■

Simplex iterations

Stepsize

What happens if some $\bar{c}_j$ < 0?

We can decrease the cost by bringing x_j into the basis

How far can we go?

$\theta^\star = \max\{\theta \mid \theta \geq 0$ and $x + \theta d \geq 0\}$

d is the j-th basic direction

Unbounded

If $d \geq 0$, then $\theta^\star = \infty$. The LP is unbounded.

Bounded

If $d_i < 0$ for some i, then

$$\theta^\star = \min_{\{i \mid d_i < 0\}} \left(-\frac{x_i}{d_i}\right) = \min_{\{i \in B \mid d_i < 0\}} \left(-\frac{x_i}{d_i}\right)$$

(Since $d_i \geq 0,\ i \in N$)

Moving to a new basis

Next feasible solution

$$x + \theta^\star d$$

Let $B(\ell) \in \{B(1), \ldots, B(m)\}$ be the index such that $\theta^\star = -\dfrac{x_{B(\ell)}}{d_{B(\ell)}}$. Then,

$$x_{B(\ell)} + \theta^\star d_{B(\ell)} = 0$$

New solution

- $x_{B(\ell)}$ becomes 0 (exits)
- x_j becomes $\theta^\star$ (enters)

New basis

$$\bar{B} = \begin{bmatrix} A_{B(1)} & \ldots & A_{B(\ell-1)} & A_j & A_{B(\ell+1)} & \ldots & A_{B(m)} \end{bmatrix}$$

Example

$$P = \{x \mid x_1 + x_2 + x_3 = 2, \quad x \geq 0\}$$

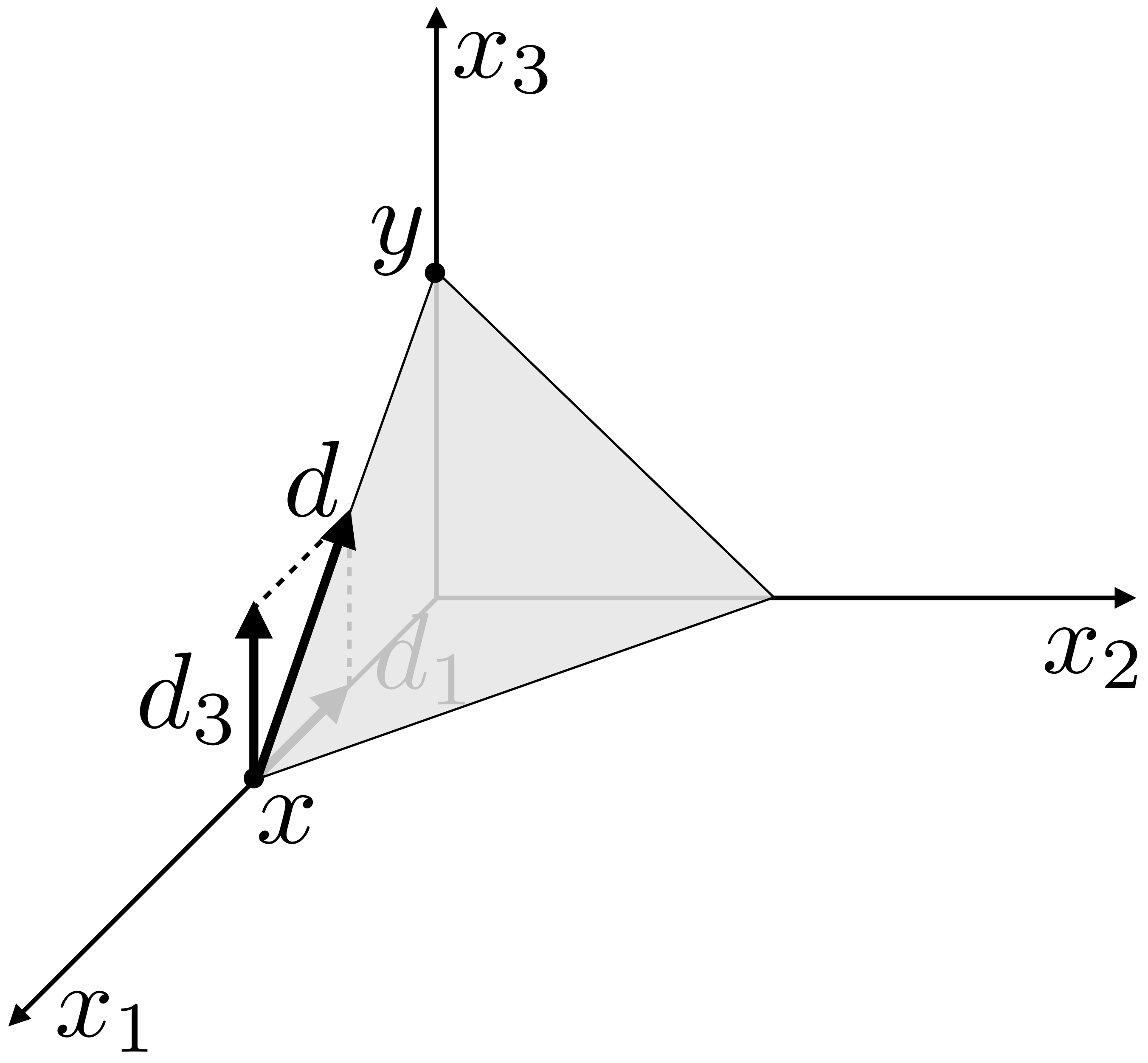
$$x = (2, 0, 0) \qquad\qquad B = \{1\}$$

Basic index $j = 3 \longrightarrow d = (-1, 0, 1)$

$$d_j = 1$$
$$d_B = -B^{-1}A_j$$

Stepsize $\theta^\star = -\dfrac{x_1}{d_1} = 2$

New solution $y = x + \theta^\star d = (0, 0, 2) \qquad \bar{B} = \{3\}$



An iteration of the simplex method

First part

We start with a basic feasible solution x and a basis matrix $B = \begin{bmatrix} A_{B(1)} & \ldots, A_{B(m)} \end{bmatrix}$

1. Compute the reduced costs $\bar{c}_j = c_j - c_B^T B^{-1} A_j$ for $j \in N$
2. If $\bar{c_j} \geq 0$, x **optimal**. **break**
3. Choose j such that $\bar{c}_j < 0$

An iteration of the simplex method

Second part

4. Compute search direction components $d_B = -B^{-1}A_j$

5. If $d_B \geq 0$, the problem is **unbounded** and the optimal value is $-\infty$. **break**

6. Compute step length $\theta^\star = \min_{\{i \in B | d_i < 0\}} \left(-\frac{x_i}{d_i} \right)$

7. Define y such that $y = x + \theta^\star d$

Finite convergence

Assume that

- $P = \{x \mid Ax = b, x \geq 0\}$ not empty
- Every basic feasible solution **non degenerate**

Then

- The simplex method **terminates after a finite number of iterations**
- At termination we either have one of the following
  - an **optimal basis** B
  - a **direction** d such that $Ad = 0,\ d \geq 0,\ c^T d < 0$ and the optimal cost is $-\infty$

Finite convergence

Proof sketch

At each iteration the algorithm improves

- by a **positive** amount $\theta^\star$
- along the **direction** d such that $c^T d < 0$

Therefore

- The cost strictly decreases
- No basic feasible solution can be visited twice

Since there is a **finite number of basic feasible solutions**

The algorithm **must eventually terminate**

$\blacksquare$

The simplex method

Today, we learned to:

- **Iterate** between basic feasible solutions
- **Verify** optimality and unboundedness conditions
- **Apply** a single iteration of the simplex method
- **Prove** finite convergence of the simplex method in the non-degenerate case

Next lecture

- Finding initial basic feasible solution
- Degeneracy
- Complexity