Lecture 11
The simplex method

- extreme points
- adjacent extreme points
- one iteration of the simplex method
- degeneracy
- initialization
- numerical implementation

Idea of the simplex method

move from one extreme point to an adjacent extreme point with lower cost until an optimal extreme point is reached

- invented in 1947 (George Dantzig)
- usually developed and implemented for LPs in standard form

questions

1. how do we characterize extreme points? (answered in lecture 3)
2. how do we move from an extreme point to an adjacent one?
3. how do we select an adjacent extreme point with a lower cost?
4. how do we find an initial extreme point?

Extreme points

to check whether x is an extreme point of a polyhedron defined by

$$a_i^T x \leq b_i, \quad i = 1, \ldots, m$$

- check that $Ax \leq b$
- define

$$A_I = \begin{bmatrix} a_{i_1}^T \\ a_{i_2}^T \\ \vdots \\ a_{i_K}^T \end{bmatrix}, \qquad I = \{i_1, \ldots, i_K\}$$

where I is the set of active constraints at x:

$$a_k^T x = b_k, \quad k \in I, \qquad a_k^T x < b_k, \quad k \notin I$$

- x is an extreme point if and only if $\mathbf{rank}(A_I) = n$

Degeneracy

an extreme point is **nondegenerate** if exactly n constraints are active at x

- A_I is square and nonsingular ($K = n$)
- $x = A_I^{-1} b_I$, where $b_I = (b_{i_1}, b_{i_2}, \ldots, b_{i_n})$

an extreme point is **degenerate** if more than n constraints are active at x

- extremality is a geometric property (depends on $\mathcal{P}$)
- degeneracy/nondegeneracy depend on the representation of $\mathcal{P}$ (*i.e.*, A and b)

Assumptions

we will develop the simplex algorithm for an LP in inequality form

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

with $A \in \mathbf{R}^{m \times n}$

we assume throughout the lecture that $\mathbf{rank}(A) = n$

- if $\mathbf{rank}(A) < n$, we can reduce the number of varables
- implies that the polyhedron has at least one extreme point (page 3–24)
- implies that if the LP is solvable, it has an optimal extreme point (page 3–27)

until page 11–20 we assume that all the extreme points are nondegenerate

Adjacent extreme points

extreme points are **adjacent** if they have $n-1$ common active constraints

moving to an adjacent extreme point
given extreme point x with active index set I and an index $k \in I$, find an extreme point $\hat{x}$ that has the active constraints $I \setminus \{k\}$ in common with x

1. solve the n equations

$$a_i^T \Delta x = 0, \quad i \in I \setminus \{k\}, \qquad a_k^T \Delta x = -1$$

2. if $A\Delta x \leq 0$, then $\{\hat{x} + \alpha \Delta x \mid \alpha \geq 0\}$ is a feasible half-line:

$$A(x + \alpha \Delta x) \leq b \quad \forall \alpha \geq 0$$

3. otherwise, $\hat{x} = x + \alpha \Delta x$, where

$$\alpha = \min_{i : a_i^T \Delta x > 0} \frac{b_i - a_i^T x}{a_i^T \Delta x}$$

comments

- step 1: equations are solvable because A_I is nonsingular
- step 3: $\alpha > 0$ because $a_i^T \Delta x > 0$ means $i \notin I$, hence $a_i^T x < b_i$ (for nondegenerate x)
- new active set is $\hat{I} = I \setminus \{k\} \cup \{j\}$ where

$$j = \operatorname*{argmin}_{i : a_i^T \Delta x > 0} \frac{b_i - a_i^T x}{a_i^T \Delta x}$$

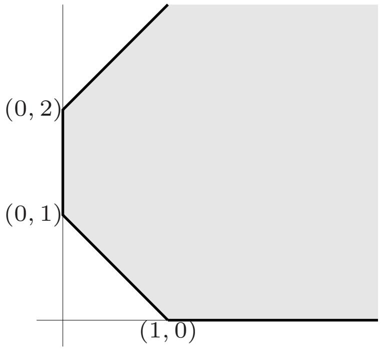
- $A_{\hat{I}}$ is nonsingular because

$$a_i^T \Delta x = 0, \quad i \in I \setminus \{k\}, \qquad a_j^T \Delta x > 0$$

  implies that a_j is linearly independent of the vectors a_i, $i \in I \setminus \{k\}$

Example

$$\begin{bmatrix} 0 & -1 \\ -1 & -1 \\ -1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq \begin{bmatrix} 0 \\ -1 \\ 0 \\ 2 \end{bmatrix}$$



extreme points

x	$b - Ax$	I
$(1, 0)$	$(0, 0, 1, 3)$	$\{1, 2\}$
$(0, 1)$	$(1, 0, 0, 1)$	$\{2, 3\}$
$(0, 2)$	$(2, 1, 0, 0)$	$\{3, 4\}$

compute extreme points adjacent to $x = (1, 0)$

1. try to remove $k = 1$ from active set $I = \{1, 2\}$
   - compute Δx

$$\begin{bmatrix} 0 & -1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \qquad \Longrightarrow \qquad \Delta x = (-1, 1)$$

   - minimum ratio test: $A\Delta x = (-1, 0, 1, 2)$

$$\alpha = \min\{\frac{b_3 - a_3^T x}{a_3^T \Delta x}, \frac{b_4 - a_4^T x}{a_4^T \Delta x}\} = \min\{1/1, 3/2\} = 1$$

new extreme point: $\hat{x} = (0, 1)$ with active set $\hat{I} = \{2, 3\}$

2. try to remove $k = 2$ from active set $I = \{1, 2\}$
   - compute Δx

$$\begin{bmatrix} 0 & -1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \qquad \Longrightarrow \qquad \Delta x = (1, 0)$$

   - $A\Delta x = (0, -1, -1, -1)$:

$$\{(1, 0) + \alpha(1, 0) \mid \alpha \geq 0\}$$

is an unbounded edge of the feasible set

Finding an adjacent extreme point with lower cost

given extreme point x with active index set I

1. define $z \in \mathbf{R}^m$ with

$$A_I^T z_I + c = 0, \qquad z_j = 0, \quad j \notin I$$

2. if $z \geq 0$, then x, z are primal and dual optimal

3. otherwise select k with $z_k < 0$ and determine Δx as on page 11–6:

$$\begin{aligned} c^T(x + \alpha\Delta x) &= c^T x - \alpha z_I^T A_I \Delta x \\ &= c^T x + \alpha z_k \end{aligned}$$

cost decreases in the direction Δx

One iteration of the simplex method

given an extreme point x with active set I

1. compute $z \in \mathbf{R}^m$ with

$$A_I^T z_I + c = 0, \qquad z_j = 0, \quad j \notin I$$

if $z \geq 0$, terminate (x is optimal)

2. choose k such that $z_k < 0$, compute $\Delta x \in \mathbf{R}^n$ with

$$a_i^T \Delta x = 0, \quad i \in I \setminus \{k\}, \qquad a_k^T \Delta x = -1$$

if $A\Delta x \leq 0$, terminate ($p^\star = -\infty$)

3. set $I := I \setminus \{k\} \cup \{j\}$, $x := x + \alpha\Delta x$ where

$$j = \underset{i:a_i^T\Delta x > 0}{\operatorname{argmin}} \frac{b_i - a_i^T x}{a_i^T \Delta x}, \qquad \alpha = \frac{b_j - a_j^T x}{a_j^T \Delta x}$$

Pivot selection and convergence

step 2: which k do we choose if z_k has several negative components?

many variants, *e.g.*,

- choose most negative z_k
- choose maximum decrease in cost αz_k
- choose smallest k

all three variants work (if all extreme points are nondegenerate)

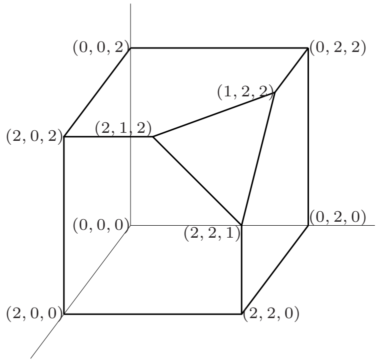
step 3: j is unique and $\alpha > 0$ (if all extreme points are nondegenerate)

convergence follows from:

- number of extreme points is finite
- cost strictly decreases at each step

Example

$$\text{min. } x_1 + x_2 - x_3 \qquad \text{s.t.} \quad \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \leq \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2 \\ 2 \\ 2 \\ 5 \end{bmatrix}$$



- optimal point is $x = (0, 0, 2)$
- start simplex method at $x = (2, 2, 0)$

---

iteration 1: $x = (2, 2, 0), \quad b - Ax = (2, 2, 0, 0, 0, 2, 1), \quad I = \{3, 4, 5\}$

1. compute z:

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} z_3 \\ z_4 \\ z_5 \end{bmatrix} = - \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \quad \Longrightarrow \quad z = (0, 0, -1, -1, -1, 0, 0)$$

   not optimal; remove $k = 3$ from active set

2. compute Δx

$$\begin{bmatrix} 0 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} \quad \Longrightarrow \quad \Delta x = (0, 0, 1)$$

3. minimum ratio test: $A\Delta x = (0, 0, -1, 0, 0, 1, 1)$

$$\alpha = \operatorname{argmin}\{2/1, 1/1\} = 1, \qquad j = 7$$

---

iteration 2: $x = (2, 2, 1), \quad b - Ax = (2, 2, 1, 0, 0, 1, 0), \quad I = \{4, 5, 7\}$

1. compute z:

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_4 \\ z_5 \\ z_7 \end{bmatrix} = - \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \quad \Longrightarrow \quad z = (0, 0, 0, -2, -2, 0, 1)$$

   not optimal; remove $k = 5$ from active set

2. compute Δx

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} \quad \Longrightarrow \quad \Delta x = (0, -1, 1)$$

3. minimum ratio test: $A\Delta x = (0, 1, -1, 0, -1, 1, 0)$

$$\alpha = \operatorname{argmin}\{2/1, 1/1\} = 1, \qquad j = 6$$

iteration 3: $x = (2, 1, 2)$, $\quad b - Ax = (2, 1, 2, 0, 1, 0, 0)$, $\quad I = \{4, 6, 7\}$

1. compute z:

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} z_4 \\ z_6 \\ z_7 \end{bmatrix} = - \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \quad \Longrightarrow \quad z = (0, 0, 0, 0, 0, 2, -1)$$

   not optimal; remove $k = 7$ from active set

2. compute Δx

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \quad \Longrightarrow \quad \Delta x = (0, -1, 0)$$

3. minimum ratio test: $A\Delta x = (0, 1, 0, 0, -1, 0, -1)$

$$\alpha = \operatorname{argmin}\{1/1\} = 1, \qquad j = 2$$

iteration 4: $x = (2, 0, 2)$, $\quad b - Ax = (2, 0, 2, 0, 2, 0, 1)$, $\quad I = \{2, 4, 6\}$

1. compute z:

$$\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_2 \\ z_4 \\ z_6 \end{bmatrix} = - \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \quad \Longrightarrow \quad z = (0, 1, 0, -1, 0, 1, 0)$$

   not optimal; remove $k = 4$ from active set

2. compute Δx

$$\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \Delta x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} \quad \Longrightarrow \quad \Delta x = (-1, 0, 0)$$

3. minimum ratio test: $A\Delta x = (1, 0, 0, -1, 0, 0, -1)$

$$\alpha = \operatorname{argmin}\{2/1\} = 2, \qquad j = 1$$

iteration 5: $x = (0, 0, 2)$, $\quad b - Ax = (0, 0, 2, 2, 2, 0, 3)$, $\quad I = \{1, 2, 6\}$

1. compute z:

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_6 \end{bmatrix} = - \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \quad \Longrightarrow \quad z = (1, 1, 0, 0, 0, 1, 0)$$

   optimal

Degeneracy

- if x is degenerate, A_I has rank n but is not square
- if next point is degenerate, we have a tie in the argmin of step 3

solution

- define I to be a subset of n linearly independent active constraints
- A_I is square; steps 1 and 2 work as in the nondegenerate case
- in step 3, break ties arbitrarily

does it work?

- in step 3 we can have $\alpha = 0$ ($i.e.$, x does not change)
- maybe this does not hurt, as long as I keeps changing

Example

$$\begin{array}{ll} \text{minimize} & -3x_1 + 5x_2 - x_3 + 2x_4 \\ \text{subject to} & \begin{bmatrix} 1 & -2 & -2 & 3 \\ 2 & -3 & -1 & 1 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \leq \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \end{array}$$

- $x = (0,0,0,0)$ is a degenerate extreme point with

$$b - Ax = (0,0,1,0,0,0,0)$$

- start simplex with $I = \{4,5,6,7\}$

iteration 1: $I = \{4,5,6,7\}$

1. $z = (0,0,0,-3,5,-1,2)$: remove 4 from active set
2. $\Delta x = (1,0,0,0)$
3. $A\Delta x = (1,2,0,-1,0,0,0)$: $\alpha = 0$, add 1 to active set

iteration 2: $I = \{1,5,6,7\}$

1. $z = (3,0,0,0,-1,-7,11)$: remove 5 from active set
2. $\Delta x = (2,1,0,0)$
3. $A\Delta x = (0,1,0,-2,-1,0,0)$: $\alpha = 0$, add 2 to active set

iteration 3: $I = \{1,2,6,7\}$

1. $z = (1,1,0,0,0,-4,6)$: remove 6 from active set
2. $\Delta x = (-4,-3,1,0)$
3. $A\Delta x = (0,0,1,4,3,-1,0)$: $\alpha = 0$, add 4 to active set

iteration 4: $I = \{1,2,4,7\}$

1. $z = (-2,3,0,1,0,0,-1)$: remove 7 from active set
2. $\Delta x = (0,-1/4,7/4,1)$
3. $A\Delta x = (0,0,7/4,0,1/4,-7/4,-1)$: $\alpha = 0$, add 5 to active set

iteration 5: $I = \{1,2,4,5\}$

1. $z = (-1,1,0,-2,4,0,0)$: remove 1 from active set
2. $\Delta x = (0,0,-1,-1)$
3. $A\Delta x = (-1,0,-1,0,0,1,1)$: $\alpha = 0$, add 6 to active set

iteration 6: $I = \{2,4,5,6\}$

1. $z = (0,-2,0,-7,11,1,0)$: remove 2 from active set
2. $\Delta x = (0,0,0,-1)$
3. $A\Delta x = (-3,-1,0,0,0,0,1)$: $\alpha = 0$, add 7 to active set

iteration 7: $I = \{4,5,6,7\}$, the initial active set

Bland's pivoting rule

no cycling will occur if we follow the following rule

- in step 2, always choose the smallest k for which $z_k < 0$
- if there is a tie in step 3, always choose the smallest j

proof (by contradiction) suppose there is a cycle *i.e.*, for some $q > p$

$$x^{(p)} = x^{(p+1)} = \cdots = x^{(q)}, \qquad I^{(p)} \neq I^{(p+1)} \neq \cdots \neq I^{(q)} = I^{(p)}$$

where $x^{(s)}$ ($I^{(s)}$, $z^{(s)}$, $\Delta x^{(s)}$) is the value of x (I, z, Δx) at iteration s

we also define

- k_s: index removed from I in iteration s; j_s: index added in iteration s
- $\bar{k} = \max_{p \leq s \leq q-1} k_s$
- r: the iteration ($p \leq r \leq q-1$) in which $\bar{k}$ is removed ($\bar{k} = k_r$)
- t: the iteration ($r < t \leq q$) in which $\bar{k}$ is added back again ($\bar{k} = j_t$)

at iteration r we remove index $\bar{k}$ from $I^{(r)}$; therefore

- $z_{\bar{k}}^{(r)} < 0$
- $z_i^{(r)} \geq 0$ for $i \in I^{(r)}$, $i < \bar{k}$ (otherwise we should have removed i)
- $z_i^{(r)} = 0$ for $i \notin I^{(r)}$ (by definition of $z^{(r)}$)

at iteration t we add index $\bar{k}$ to $I^{(t)}$; therefore

- $a_{\bar{k}}^T \Delta x^{(t)} > 0$
- $a_i^T \Delta x^{(t)} \leq 0$ for $i \in I^{(r)}$, $i < \bar{k}$

  (otherwise we should have added i, since $b_i - a_i^T x = 0$ for all $i \in I^{(r)}$)
- $a_i^T \Delta x^{(t)} = 0$, for $i \in I^{(r)}$, $i > \bar{k}$

  (if $i > \bar{k}$ and $i \in I^{(r)}$ then it is never removed, so $i \in I^{(t)} \setminus \{k_t\}$)

conclusion: ${z^{(r)}}^T A \Delta x^{(t)} < 0$

a contradiction, because $-{z^{(r)}}^T A \Delta x^{(t)} = c^T \Delta x^{(t)} \leq 0$

Example

LP of page 11–21, same starting point but applying Bland's rule

iteration 1: $I = \{4, 5, 6, 7\}$

1. $z = (0, 0, 0, -3, 5, -1, 2)$: remove 4 from active set
2. $\Delta x = (1, 0, 0, 0)$
3. $A\Delta x = (1, 2, 0, -1, 0, 0, 0)$: $\alpha = 0$, add 1 to active set

iteration 2: $I = \{1, 5, 6, 7\}$

1. $z = (3, 0, 0, 0, -1, -7, 11)$: remove 5 from active set
2. $\Delta x = (2, 1, 0, 0)$
3. $A\Delta x = (0, 1, 0, -2, -1, 0, 0)$: $\alpha = 0$, add 2 to active set

iteration 3: $I = \{1, 2, 6, 7\}$

1. $z = (1, 1, 0, 0, 0, -4, 6)$: remove 6 from active set
2. $\Delta x = (-4, -3, 1, 0)$
3. $A\Delta x = (0, 0, 1, 4, 3, -1, 0)$: $\alpha = 0$, add 4 to active set

iteration 4: $I = \{1, 2, 4, 7\}$

1. $z = (-2, 3, 0, 1, 0, 0, -1)$: remove 1 from active set
2. $\Delta x = (0, -1/4, 3/4, 1)$
3. $A\Delta x = (-1, 0, 3/4, 0, 1/4, -3/4, 0)$: $\alpha = 0$, add 5 to active set

iteration 5: $I = \{2, 4, 5, 7\}$

1. $z = (0, -1, 0, -5, 8, 0, 1)$: remove 2 from active set
2. $\Delta x = (0, 0, 1, 0)$
3. $A\Delta x = (-2, -1, 1, 0, 0, -1, 0)$: $\alpha = 1$, add 3 to active set

new $x = (0, 0, 1, 0)$, $b - Ax = (2, 1, 0, 0, 0, 1, 0)$

iteration 6: $I = \{3, 4, 5, 7\}$

1. $z = (0, 0, 1, -3, 5, 0, 2)$: remove 4 from active set
2. $\Delta x = (1, 0, 0, 0)$
3. $A\Delta x = (1, 2, 0, -1, 0, 0, 0)$: $\alpha = 1/2$, add 2 to active set

new $x = (1/2, 0, 1, 0)$, $b - Ax = (3/2, 0, 0, 1/2, 0, 1, 0)$

iteration 7: $I = \{1, 3, 5, 7\}$

1. $z = (3, 0, 7, 0, -1, 0, 11)$: remove 5 from active set
2. $\Delta x = (2, 1, 0, 0)$
3. $A\Delta x = (0, 1, 0, -2, -1, 0, 0)$: $\alpha = 0$, add 2 to active set

iteration 8: $I = \{1, 2, 3, 7\}$

1. $z = (1, 1, 4, 0, 0, 0, 6)$: optimal

Initialization via phase I

linear program with variable bounds

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \le b, \quad x \ge 0 \end{array}$$

general; can split free x_k as $x_k = x_k^+ - x_k^-$, $x_k \ge 0$, $x_k^- \ge 0$

phase I problem

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & Ax \le (1-t)b, \quad x \ge 0, \quad 0 \le t \le 1 \end{array}$$

- $x = 0$, $t = 1$ is an extreme point of phase I LP
- can compute an optimal extreme point $x^\star$, $t^\star$ of phase I LP via simplex
- if $t^\star > 0$, original problem is infeasible
- if $t^\star = 0$, then $x^\star$ is an extreme point of original problem

Numerical implementation

- most expensive step: solution of two sets of linear equations

$$A_I^T z_I = -c, \qquad A_I \Delta x = (e_k)_I$$

  where e_k is kth unit vector
- one row of A_I changes at each iteration

efficient implementation: propagate LU factorization of A_I

- given the factorization, can solve the equations in $O(n^2)$ operations
- updating LU factorization after changing a row costs $O(n^2)$ operations

total cost is $O(n^2)$ per iteration ($\ll O(n^2)$ if A is sparse)

Complexity of the simplex method

in practice: very efficient (#iterations grows linearly with m, n)

worst-case:

- for most pivoting rules, there exist examples where the number of iterations grows exponentially with n and m
- it is an open question whether there exists a pivoting rule for which the number of iterations is bounded by a polynomial of n and m