Linear Program in Standard Form

Throughout this chapter, we consider the following standard form problem:

\[
\begin{align*}
\text{minimize} & \quad c^T \cdot x \\
\text{subject to} & \quad A \cdot x = b \\
& \quad x \geq 0
\end{align*}
\]

with \( A \in \mathbb{R}^{m \times n}, \text{rank}(A) = m, b \in \mathbb{R}^m, \) and \( c \in \mathbb{R}^n. \)

Recall:

- Let \( B = (A_{B(1)}, \ldots, A_{B(m)}) \) be a basis matrix of \( A. \) Then \( B \) corresponds to the basic solution \( x = (x_B, x_N)^T, \) where \( x_B = B^{-1}b \) and \( x_N = 0. \)

- \( x = (x_B, x_N)^T \) is a basic feasible solution if \( x_B \geq 0. \)
Feasible directions

**Definition 3.1 in [BT]**

Let \( x \in P \subseteq \mathbb{R}^n \) be a point in a polyhedron. A vector \( d \in \mathbb{R}^n \) is called a feasible direction at \( x \), if there exists some \( \theta \in \mathbb{R}_{>0} \) with \( x + \theta d \in P \).

The set of all feasible directions at a fixed point \( x \) forms the cone of feasible directions at \( x \).

Picture ...

In standard form, we try to raise the value of a nonbasic variable while keeping the other nonbasic variables at zero. This is called the \( j \)-th basic direction.

**Definition page 83 [BT]**

Let \( x_j \) be a nonbasic variable. The \( j \)-th basic direction \( d \in \mathbb{R}^n \) is defined via

- \( d_j = 1 \)
- \( d_i = 0 \) for all other nonbasic variables \( x_i \)
- \( d_B = -B^{-1}A_j \in \mathbb{R}^m \).  

Prop.: If \( d \) is a \( j \)-th basic direction at \( x \) (and in particular \( Ax = b \)), then \( A(x + \theta d) = b \).  

At a nondegenerate basic solution \( x \) the \( j \)-th basic direction is always a feasible direction. Not true at degenerate basic solutions! Picture ...
Reduced Costs

Change of the cost function moving along a basic direction:
If \( d \) is the \( j \)th basic direction, then the rate of change is \( c^T d \).

\[
c^T d = c_B^T d_B + c_j
\]

where \( c_B = (c_B(1), \ldots, c_B(m)) \).

Since \( d_B = -B^{-1}A_j \):

\[
c^T d = c_B^T (-B^{-1}A_j) + c_j = c_j - c_B^T B^{-1}A_j.
\]

Definition 3.2 in [BT]

Let \( x \) be a basic solution, let \( B \) be an associated basis matrix, let \( c_B \) be the vector of costs of the basic variables. The reduced cost \( \bar{c}_j \) of the variable \( x_j \) is defined as

\[
\bar{c}_j = c_j - c_B^T B^{-1}A_j.
\]

It is the change of the cost function when moving in the \( j \)th basic direction.
Main Idea of the Simplex Method

Idea
Change basis by exchanging one basic column with one non-basic column.

More precisely:

- Start with a basis $B$ whose associated basic solution is a basic feasible solution.
- Then proceed in iterations. In each iteration:
  - select a nonbasic column $j$ such that bringing $j$ into the basis decreases the value of the objective function (i.e., $\bar{c}_j < 0$). Stop, if no such column exists, i.e., $\bar{c} \geq 0$.
  - Go along the $j$-th basic direction. Select a basic column $\ell$ corresponding to a blocking variable $x_\ell$: remove $\ell$ from the basis and bring $j$ into the basis

Iterations are called pivot steps.
An Example

A simple linear programming problem:

\[
\begin{align*}
\text{min} & \quad -10 \, x_1 - 12 \, x_2 - 12 \, x_3 \\
\text{s.t.} & \quad x_1 + 2 \, x_2 + 2 \, x_3 \leq 20 \\
& \quad 2 \, x_1 + x_2 + 2 \, x_3 \leq 20 \\
& \quad 2 \, x_1 + 2 \, x_2 + x_3 \leq 20 \\
& \quad x_1, x_2, x_3 \geq 0
\end{align*}
\]
Set of Feasible Solutions

\[ A = (0, 0, 0)^T \]
\[ B = (0, 0, 10)^T \]
\[ C = (0, 10, 0)^T \]
\[ D = (10, 0, 0)^T \]
\[ E = (4, 4, 4)^T \]
Introducing Slack Variables

$$\begin{align*}
\text{min} & \quad -10x_1 - 12x_2 - 12x_3 \\
\text{s.t.} & \quad x_1 + 2x_2 + 2x_3 \leq 20 \\
& \quad 2x_1 + x_2 + 2x_3 \leq 20 \\
& \quad 2x_1 + 2x_2 + x_3 \leq 20 \\
& \quad x_1, x_2, x_3 \geq 0
\end{align*}$$

LP in standard form

$$\begin{align*}
\text{min} & \quad -10x_1 - 12x_2 - 12x_3 \\
\text{s.t.} & \quad x_1 + 2x_2 + 2x_3 + x_4 = 20 \\
& \quad 2x_1 + x_2 + 2x_3 + x_5 = 20 \\
& \quad 2x_1 + 2x_2 + x_3 + x_6 = 20 \\
& \quad x_1, \ldots, x_6 \geq 0
\end{align*}$$

Observation

The right hand side of the system is non-negative. Therefore the point 
$$(0, 0, 0, 20, 20, 20)^T$$ is a basic feasible solution and we can start the simplex method with basis $B(1) = 4, B(2) = 5, B(3) = 6.$

Example of the simplex algorithm ...
From one basis to an adjacent one: An iteration of the simplex method

Goal: An iteration of the simplex method moves from a basic feasible solution to a basic feasible solution. It either decreases the objective function or stays at its basic feasible solution.

To prove this, we first formalize an iteration of the simplex algorithm:

- Let \( x \in P \) basic feasible solution. \( B \) an associated basis, \( j \notin B \)
- Let \( d \) a \( j \)-th basic direction at \( x \) with \( \bar{c}_j < 0 \). (if \( \bar{c} \geq 0 \), the simplex alg. terminates)
- \( \theta^* := \max\{\theta \geq 0 \mid x + \theta d \in P\} \) (if \( \theta^* = \infty \), then the LP is unbounded)
  - \( \theta^* \) finite \( \Rightarrow \) some \( d_i < 0 \)
  - For each \( d_i < 0 \): \( x_i + \theta d_i \geq 0 \) becomes \( \theta \leq -x_i/d_i \)
  - Thus \( \theta^* = \min\{-x_i/d_i \mid d_i < 0\} = \min\{-x_{B(j)}/d_{B(j)} \mid d_{B(j)} < 0\} \)
- Let \( \ell \) be a minimizing index, in particular \( \bullet x_{B(\ell)} + \theta^* d_{B(\ell)} = 0 \) and \( \bullet d_{B(\ell)} < 0 \)
- \( y := x + \theta^* d \) is feasible. \( y_{B(\ell)} = 0 \) and \( y_j = \theta^* \)
- \( \overline{B}(i) = \begin{cases} B(i) & \text{if } i \neq \ell \\ j & \text{if } i = \ell \end{cases} \)

**Theorem 3.2 in [BT]**

- \( A_{\overline{B}} \) has full rank, i.e., \( \overline{B} \) is a basis.  
- \( y \) is a basic feasible solution associated to \( \overline{B} \).
Optimality

Theorem 3.1 in [BT]

Consider a basic feasible solution $x$ associated with a basis matrix $B$, and let $\bar{c}$ be the corresponding vector of reduced costs.

(a) If $\bar{c} \geq 0$, then $x$ is optimal.

(b) If $x$ is optimal and nondegenerate, then $\bar{c} \geq 0$.

Proof: ...

“Careful” Corollary: If the optimum is nondegenerate and the simplex algorithm reaches it, then the simplex algorithm terminates.
Cycling (Ch. 3.4 in [BT])

Problem: If an LP is degenerate, the simplex method might end up in an infinite loop (cycling).

Example:

\[
\begin{align*}
\text{min} & \quad -\frac{3}{4}x_1 + 20x_2 - \frac{1}{2}x_3 + 6x_6 \\
\text{s.t.} & \quad \frac{1}{4}x_1 - 8x_2 - x_3 + 9x_4 + x_5 = 0 \\
& \quad \frac{1}{2}x_1 - 12x_2 - \frac{1}{2}x_3 + 3x_4 + x_6 = 0 \\
& \quad x_3 + x_7 = 1
\end{align*}
\]

We observe this with these pivoting rules:

- **Column selection**: let nonbasic variable with most negative reduced cost $\bar{c}_j$ enter the basis, i.e., steepest descent rule.
- **Row selection**: among basic variables that are eligible to exit the basis, select the one with smallest subscript.

**Bases visited**

\[(5, 6, 7) \rightarrow (1, 6, 7) \rightarrow (1, 2, 7) \rightarrow (3, 2, 7) \rightarrow (3, 4, 7) \rightarrow (5, 4, 7) \rightarrow (5, 6, 7)\]

This is the same basis that we started with! Continuing with these pivoting rules, the simplex method will never terminate in this example.
Anticycling – Perturbation

Perturbation:
- Replace each $x_i$ by $\tilde{x}_i = x_i + \varepsilon^i$.

Observation: $x_i \geq 0$ iff ($\tilde{x}_i \geq 0$ for all $\varepsilon \geq 0$)
- This gives a relaxed LP.

- Observation: For small enough $\varepsilon$, the optimal basis for the perturbed LP equals an optimal basis for the original LP.

Proposition

In the perturbed LP, every basis gives a different basic solution.

Proof:
- Recall that the basic solution to $B$ is $x_B = (A_B)^{-1}b$
- Let $\vec{\varepsilon} = (\varepsilon, \varepsilon^2, \varepsilon^3, \ldots, \varepsilon^n)$, so $\tilde{x} = x + \vec{\varepsilon}$
- Rephrase $Ax = b$ in terms of $\tilde{x}$: $Ax = b \iff A(\tilde{x} - \vec{\varepsilon}) = b \iff A\tilde{x} = b + A\vec{\varepsilon}$
- Every entry of the basic solutions $\tilde{x}_B = (A_B)^{-1}(b + A\vec{\varepsilon})$ is a nonzero polynomial in $\varepsilon$, because $(A_B)^{-1}$ does not have a zero row. Therefore no basic variable is zero.

Conclusion

For a perturbed LP, during the simplex algorithm, whenever $\overline{c} \not\geq 0$, then we have $\theta^* > 0$. Hence on the perturbed LP the simplex algorithm makes nonzero progress on the objective function during each step, and hence switches the vertex (and never visits a vertex twice), and therefore terminates eventually.
Anticycling – Bland’s Rule (much more efficient)

This pivoting rule that is guaranteed to avoid cycling:

Smallest subscript pivoting rule (Bland’s rule) [BT] page 111

1. Choose the column $A_j$ with $\bar{c}_j < 0$ and $j$ minimal to enter the basis.
2. Among all basic variables $x_i$ that could exit the basis, select the one with smallest $i$.

Theorem (without proof)
The simplex algorithm with Bland’s rule does not cycle.

Using Bland’s rule, the simplex algorithm always terminates:

- $\bar{c} \geq 0$ implies optimality and makes the simplex algorithm terminate.
- If $\bar{c}_i < 0$, then we
  - either decrease the value of the objective function (finite $\theta^* > 0$)
  - or we output “unbounded” ($\theta^* = \infty$)
  - or we make one of finitely many steps at the same basic feasible solution ($\theta^* = 0$).

Hence after finitely many steps the algorithm terminates or the value of the objective function decreases. All basic feasible solutions that have been visited thus cannot be visited again.
Finding an Initial Basic Feasible Solution ([BT] Sec. 3.5)

So far we always assumed that the simplex algorithm starts with a basic feasible solution. We now discuss how such a solution can be obtained.

- The two-phase simplex method
- The big-$M$ method
Driving artificial variables out of the basis

Example:

\[
\begin{align*}
\text{min} & \quad x_1 + x_2 + x_3 \\
\text{s.t.} & \quad x_1 + 2x_2 + 3x_3 = 3 \\
& \quad -x_1 + 2x_2 + 6x_3 = 2 \\
& \quad 4x_2 + 9x_3 = 5 \\
& \quad 3x_3 + x_4 = 1 \\
& \quad x_1, \ldots, x_4 \geq 0
\end{align*}
\]

(Without loss of generality we always have \( b \geq 0 \))

Auxiliary problem with artificial variables:

\[
\begin{align*}
\text{min} & \quad x_5 + x_6 + x_7 + x_8 \\
\text{s.t.} & \quad x_1 + 2x_2 + 3x_3 + x_5 = 3 \\
& \quad -x_1 + 2x_2 + 6x_3 + x_6 = 2 \\
& \quad 4x_2 + 9x_3 + x_7 = 5 \\
& \quad 3x_3 + x_4 + x_8 = 1 \\
& \quad x_1, \ldots, x_4, x_5, \ldots, x_8 \geq 0
\end{align*}
\]
Auxiliary Problem

Auxiliary problem with artificial variables:

\[
\begin{align*}
\text{min} & \quad x_5 + x_6 + x_7 + x_8 \\
\text{s.t.} & \quad x_1 + 2x_2 + 3x_3 + x_5 = 3 \\
& \quad -x_1 + 2x_2 + 6x_3 + x_6 = 2 \\
& \quad 4x_2 + 9x_3 + x_7 = 5 \\
& \quad 3x_3 + x_4 + x_8 = 1 \\
& \quad x_1, \ldots, x_4, x_5, \ldots, x_8 \geq 0
\end{align*}
\]

Observation

\( x = (0, 0, 0, 0, 3, 2, 5, 1) \) is a basic feasible solution for this problem with basic variables \((x_5, x_6, x_7, x_8)\). We can start the simplex algorithm from here.

- If the simplex algorithm cannot find a solution with objective function value 0: The original LP was infeasible.
- If the simplex algorithm finds a feasible solution \( x \) with objective function value 0: Remove the artificial variables and start the simplex from the truncated \( x \) with the original objective function.

(If some of the removed variables were basic variables, choose a proper replacements among the remaining variables: The basic columns shall be linearly independent)
Omitting Artificial Variables

**Auxiliary problem**

\[
\begin{array}{l}
\text{min} & x_5 + x_6 + x_7 + x_8 \\
\text{s.t.} & x_1 + 2x_2 + 3x_3 + x_5 = 3 \\
& -x_1 + 2x_2 + 6x_3 + x_6 = 2 \\
& 4x_2 + 9x_3 + x_7 = 5 \\
& 3x_3 + x_4 + x_8 = 1 \\
& x_1, \ldots, x_8 \geq 0
\end{array}
\]

Artificial variable \(x_8\) could have been omitted by setting \(x_4\) to 1 in the initial basis. This is possible as \(x_4\) does only appear in one constraint.

Generally, this can be done, e.g., with all slack variables that have nonnegative right hand sides.
Phase I of the Simplex Method

Given: LP in standard form, linearly independent rows: \( \min \{ c^T \cdot x \mid A \cdot x = b, \ x \geq 0 \} \)

1. Transform problem such that \( b \geq 0 \) (multiply constraints by \(-1\)).
2. Introduce artificial variables \( y_1, \ldots, y_m \) and solve auxiliary problem

\[
\min \sum_{i=1}^{m} y_i \quad \text{s.t. } A \cdot x + I_m \cdot y = b, \ x, y \geq 0.
\]

3. If optimal cost is positive, then STOP (original LP is infeasible).
4. If no artificial variable is in final basis, eliminate artificial variables and STOP (feasible basis for original LP has been found).
5. If artificial variables are in final basis, remove them and fill up with non-artificial basis (keep the columns linearly independent!). Such columns exists, because the rows of \( A \) are linearly independent.
The Two-phase Simplex Method

### Two-phase simplex method

1. Given an LP in standard form, first run phase I.
2. If phase I yields a basic feasible solution for the original LP, enter "phase II" (see above).

### Possible outcomes of the two-phase simplex method

1. Problem is infeasible (detected in phase I).
2. Optimal cost is $-\infty$ (detected in phase II).
3. Problem has optimal basic feasible solution (found in phase II).
Big-$M$ Method

**Alternative idea:** Combine the two phases into one by introducing sufficiently large penalty costs for artificial variables.

This way, the LP

\[
\begin{align*}
\min \quad & \sum_{i=1}^{n} c_i x_i \\
\text{s.t.} \quad & A \cdot x = b \\
& x \geq 0
\end{align*}
\]

becomes:

\[
\begin{align*}
\min \quad & \sum_{i=1}^{n} c_i x_i + M \cdot \sum_{j=1}^{m} y_j \\
\text{s.t.} \quad & A \cdot x + I_m \cdot y = b \\
& x, y \geq 0
\end{align*}
\]

**Remark:** If $M$ is sufficiently large and the original program has a feasible solution, all artificial variables will be driven to zero by the simplex method.

**Remark:** There is no need to give $M$ a numerical value.
Example:

\[\begin{align*}
\text{min} & \quad x_1 + x_2 + x_3 \\
\text{s.t.} & \quad x_1 + 2x_2 + 3x_3 = 3 \\
& \quad -x_1 + 2x_2 + 6x_3 = 2 \\
& \quad 4x_2 + 9x_3 = 5 \\
& \quad 3x_3 + x_4 = 1 \\
& \quad x_1, \ldots, x_4 \geq 0
\end{align*}\]
Introducing Artificial Variables and $M$

Auxiliary problem:

$$\begin{align*}
\min & \quad x_1 + x_2 + x_3 + M x_5 + M x_6 + M x_7 \\
\text{s.t.} & \quad x_1 + 2x_2 + 3x_3 + x_5 = 3 \\
& \quad -x_1 + 2x_2 + 6x_3 + x_6 = 2 \\
& \quad 4x_2 + 9x_3 + x_7 = 5 \\
& \quad 3x_3 + x_4 = 1 \\
& \quad x_1, \ldots, x_4, x_5, x_6, x_7 \geq 0
\end{align*}$$

Note that this time the unnecessary artificial variable $x_8$ has been omitted.

We start off with $(x_5, x_6, x_7, x_4) = (3, 2, 5, 1)$. 
Computational Efficiency of the Simplex Method ([BT] Sec. 3.7)

Observation
The computational efficiency of the simplex method is determined by

1. the computational effort of each iteration;
2. the number of iterations.

Question: How many iterations are needed in the worst case?

Idea for negative answer (lower bound)
Describe

- a polyhedron with an exponential number of vertices;
- a path that visits all vertices and always moves from a vertex to an adjacent one that has lower costs.
Consider the unit cube in $\mathbb{R}^n$, defined by the constraints

$$0 \leq x_i \leq 1, \quad i = 1, \ldots, n$$

The unit cube has

- $2^n$ vertices;
- a *spanning path*, i.e., a path traveling the edges of the cube visiting each vertex exactly once.
Klee-Minty cube

Consider a perturbation of the unit cube in $\mathbb{R}^n$, defined by the constraints

\[
0 \leq x_1 \leq 1, \\
\epsilon x_{i-1} \leq x_i \leq 1 - \epsilon x_{i-1}, \quad i = 2, \ldots, n
\]

for some $\epsilon \in (0, 1/2)$.
Klee-Minty cube

\[
0 \leq x_1 \leq 1, \\
\epsilon x_{i-1} \leq x_i \leq 1 - \epsilon x_{i-1}, \quad i = 2, \ldots, n, \quad \epsilon \in (0, 1/2)
\]

Theorem 3.5 in [BT]

Consider the linear programming problem of minimizing \(-x_n\) subject to the constraints above. Then,

- the feasible set has \(2^n\) vertices;
- the vertices can be ordered so that each one is adjacent to and has lower cost than the previous one;
- there exists a pivoting rule under which the simplex method requires \(2^n - 1\) changes of basis before it terminates.
Diameter of Polyhedra

Definition (page 126 in [BT])

- The distance $d(x, y)$ between two vertices $x, y$ is the minimum number of edges required to reach $y$ starting from $x$.

- The diameter $D(P)$ of polyhedron $P$ is the maximum $d(x, y)$ over all pairs of vertices $(x, y)$.

- $\Delta(n, m)$ is the maximum $D(P)$ over all polytopes in $\mathbb{R}^n$ that are represented in terms of $m$ inequality constraints.

- $\Delta_u(n, m)$ is the maximum $D(P)$ over all polyhedra in $\mathbb{R}^n$ that are represented in terms of $m$ inequality constraints.

$\Delta(2, 8) = \left\lfloor \frac{8}{2} \right\rfloor = 4$

$\Delta_u(2, 8) = 8 - 2 = 6$
Hirsch Conjecture

Observation: The diameter of the feasible set in a linear programming problem is a lower bound on the number of steps required by the simplex method, no matter which pivoting rule is being used.

Polynomial Hirsch Conjecture

\[ \Delta(n, m) \leq \text{poly}(m, n) \]

Remarks

- Known lower bounds: \( \Delta_u(n, m) \geq m - n + \left\lfloor \frac{n}{5} \right\rfloor \)
- Known upper bounds:
  \[ \Delta(n, m) \leq \Delta_u(n, m) < m^{1+\log_2 n} = (2n)^{\log_2 m} \]
- The Strong Hirsch Conjecture
  \[ \Delta(n, m) \leq m - n \]

was disproven in 2010 by Paco Santos for \( n = 43, m = 86 \).
Average Case Behavior of the Simplex Method

- Despite the exponential lower bounds on the worst case behavior of the simplex method (Klee-Minty cubes etc.), the simplex method usually behaves well in practice.
- The number of iterations is “typically” $O(m)$.
- There have been several attempts to explain this phenomenon from a more theoretical point of view.
- These results give upper bounds on the running time “on average”.
- One main difficulty is to come up with a meaningful and, at the same time, manageable definition of the term “on average”.