Algorithmic Aspects of Communication Networks

Chapter 3
General Optimization Methods for Network Design

Part 1

Thomas Böhme
Technische Universität Ilmenau

WS 2021/22

Optimization Problems

- An *optimization problem* is given by a set M and a function $f : X \to \mathbb{R}$ with $X \supseteq M$.
- The usual terminology is as follows:
  - M is called the *feasible set*.
  - f is called the *objective function*
  - Elements of M are called *feasible solutions*.
  - $x^* \in M$ is an *optimal (maximal) solution* if $f(x^*) \geq f(x)$ for all $x \in M$, i.e.

$$f(x^*) = \max\{f(x) \mid x \in M\}.$$

- An optimization method is an algorithm that computes an optimal solution x^* given the input (M, f) if there is any.
- Note that $\min\{f(x) \mid x \in M\} = -\max\{-f(x) \mid x \in M\}$. Thus, there is no need to deal with minimization problems separately.

Linear Optimization Problems (LOP)

- An optimization problem (M, f) is a *linear optimization problem (LOP)* if $M \subseteq \mathbb{R}^n$ for some $n \in \mathbb{N}$ consists of all $\vec{x} \in \mathbb{R}^n$ satisfying a finite set of linear inequalities and/or linear equations, and $f : \mathbb{R}^n \to \mathbb{R}$ is a linear function.
- Example 1:
  Maximize
  $$f(x, y) = x + 3y$$
  subject to
  $$\begin{array}{rcl} -x + y & \leq & 1 \\ x + y & \leq & 2 \\ x, y & \geq & 0 \end{array}$$

Some Remarks on Notation

- $\mathbb{R}^n$ is the set of all column vectors $\vec{x} = (x_1, \ldots, x_n)^T$ with $x_1, \ldots x_n \in \mathbb{R}$.
- We write $\vec{a} \leq \vec{b}$ if and only if $a_k \leq b_k$ for all $k \in \{1, \ldots, n\}$, and use this notation for row vectors correspondingly. (Caution: This is not a linear order. For example, neither $(1, 2)^T \leq (2, 1)^T$ nor $(1, 2)^T \geq (2, 1)^T$ obtains.)
- $\vec{o}$ denotes a zero vector (with appropriately many entries).
- I denotes an identity matrix (with appropriately many rows and columns).
- $\|\vec{x}\| = \sqrt{x_1^2 + \cdots + x_n^2}$ denotes the euclidean norm of $\vec{x}$.

Canonical Form LOP's

Canonical Form:
Maximize
$$f(\vec{x}) = \vec{c}^T \vec{x}$$
subject to
$$A\vec{x} \leq \vec{b}, \vec{x} \geq \vec{o}$$
Here, $A \in \mathbb{R}^{m \times n}$ is real matrix with m rows and n columns and $\vec{b} \in \mathbb{R}^m$ is a real column vector with m entries.

Transformation into Canonical Form

- If there is a variable x_k subject to $x_k \leq 0$ replace every appearance of x_k with $-x_k$ and replace $x_k \leq 0$ with $x_k \geq 0$.
- If a variable x_k is unrestricted, replace every appearance of x_k by $x_k^+ - x_k^-$ and let $x_k^+ \geq 0$, $x_k^- \geq 0$. This results in an equivalent LOP where each variable is non-negative.
- The inequality $\vec{a}_k^T \vec{x} \leq b_k$ is equivalent to $-\vec{a}_k^T \vec{x} \geq -b_k$, and the equality $\vec{a}_k^T \vec{x} = b_k$ is equivalent to $\vec{a}_k^T \vec{x} \leq b_k$ and $\vec{a}_k^T \vec{x} \geq b_k$. Thus, any LOP can be transformed into an equivalent LOP in canonical form.

Standard Form LOP's

- **Standard Form:**
  Maximize
  $$f(\vec{x}) = \vec{c}^T\vec{x}$$
  subject to
  $$A\vec{x} = \vec{b}$$
  and
  $$\vec{x} \geq \vec{o}$$
  Here, $A \in \mathbb{R}^{m \times n}$ is a matrix with rank m and $\vec{b} \in \mathbb{R}^m$ is a column vector with $\vec{b} \geq \vec{o}$.

Transformation into Standard Form

- Assume the LOP is given in general form, i.e.the feasible set M is written as $M = \{\vec{x} \in \mathbb{R}^n \mid (A\vec{x} \leq \vec{b}) \wedge (\vec{x} \geq \vec{o})\}$ such that $A \in \mathbb{R}^{m \times n}$, $\vec{b} \in \mathbb{R}^m$.
- Clearly,
  $$M = \{\vec{x} \in \mathbb{R}^n | \, \exists \vec{y} \in \mathbb{R}^m : \ (A\vec{x}+\vec{y} = \vec{b}) \wedge (\vec{x} \geq \vec{o}) \wedge (\vec{y} \geq \vec{o}))\}.$$
  The additional variables in $\vec{y}$ are called *slack variables*.
- Note that the equality $A\vec{x} + \vec{y} = \vec{b}$ can be re-written as
  $$(A|I)\begin{pmatrix} \vec{x} \\ \vec{y} \end{pmatrix} = \vec{b}.$$
- If $\vec{b}$ does not satisfy $\vec{b} \geq \vec{o}$, multiply all rows k with $b_k < 0$ by (-1).
- Finally, recall that if $A \in \mathbb{R}^{m \times n}$, then $rk(A) \leq m$, and that if $rk(A) < m$, then the system of linear equations $A\vec{x} = \vec{b}$ either has no solution at all or there are $m - rk(A)$ redundant equations. Consequently, the assumption that $rk(A) = m$ is no restriction.

The Structure of the Feasible Set 1

Definition: A subset X of $\mathbb{R}^n$ is called

- *convex* if for any two points $x_1, x_2 \in X$ the whole straight line segment $\overline{x_1 x_2} = \{\alpha x_1 + (1-\alpha)x_2 \mid \alpha \in [0,1]\}$ is in X,
- *closed* if the limit of every convergent sequence in X is in X too, and
- *bounded* if there is $K \in \mathbb{R}$ such that $\|\vec{x}\| \leq K$ for all $\vec{x} \in X$.

The Structure of the Feasible Set 2

We recall the following statements.

- The intersection of two convex sets is convex.
- The intersection of two closed sets is closed.
- If $f : X \to \mathbb{R}$ is continuous and $M \subseteq X$ is nonempty, closed and bounded, then there is a $x^* \in M$ such that $f(x) \leq f(x^*)$ for all $x \in M$, i.e. f attains its maximum on M.
- If $\vec{d} \in \mathbb{R}^n$ and $t \in \mathbb{R}$, then the set $\{\vec{x} \in \mathbb{R}^n \mid \vec{d}^T \vec{x} \leq t\}$ is convex and closed.
- Linear functions from $\mathbb{R}^n$ into $\mathbb{R}$ are continuous.

The Structure of the Feasible Set 3

Consider an LOP in standard form.

Maximize $f(\vec{x}) = \vec{c}^T\vec{x}$ subject to $A\vec{x} = \vec{b}$, $\vec{x} \geq \vec{o}$ with $A \in \mathbb{R}^{m\times n}$, $\vec{b} \geq \vec{o}$ and $rk(A) = m$.

By the statements from the previous slide we deduce that the feasible set $M = \{\vec{x} \in \mathbb{R}^{m\times n} \mid (A\vec{x} = \vec{b}) \wedge (\vec{x} \geq \vec{o})\ \}$ is convex and closed. If M is also nonempty and bounded, f attains its maximum on M.

Extreme Points and Basic Solutions 1

Definition: Let M be a convex set. A point $\vec{x}_0 \in M$ is called an *extreme point* of M if $\vec{x}_0$ cannot be expressed as a convex linear combination $\alpha\vec{x}_1 + (1-\alpha)\vec{x}_2$ with $\alpha \in [0,1]$ of two points $\vec{x}_1, \vec{x}_2 \in M$ with $\vec{x}_1 \neq \vec{x}_0 \neq \vec{x}_2$.

We mention the following statements without proof.

(A) If the set $M = \{\vec{x} \in \mathbb{R}^{m\times n} \mid (A\vec{x} = \vec{b}) \wedge (\vec{x} \geq \vec{o})\ \}$ is nonempty, then M has at most finitely many and at least one extreme point.

(B) If M is in addition bounded, then M is the convex hull of its extreme points $\vec{x}_1, \ldots, \vec{x}_k$, i.e. any point $\vec{x}$ in M can be expressed as convex linear combination $\vec{x} = \sum_{i=1}^{k} \alpha_k \vec{x}_i$ with $\alpha_1, \ldots, \alpha_k \in [0,1]$ and $\sum_{i=1}^{k} \alpha_i = 1$.

Extreme Points and Basic Solutions 2

Theorem (1)

If M is nonempty and bounded, then there is an extreme point $\vec{x}^* \in M$ such that $f(\vec{x}^*) \geq f(\vec{x})$ for all $\vec{x} \in M$.

Proof.

- Since f is linear and M is closed and bounded, there is a point $\vec{x}_0 \in M$ such that $f(\vec{x}_0) \geq f(\vec{x})$ for all $\vec{x} \in M$.
- Since M is nonempty and bounded, $\vec{x}_0$ can be expressed as convex linear combination

$$\vec{x}_0 = \sum_{i=1}^{k} \alpha_i \vec{x}_i$$

of extreme points $\vec{x}_1, \ldots, \vec{x}_k$ with $\alpha_1, \ldots, \alpha_k \in [0,1]$ and $\sum_{i=1}^{k} \alpha_i = 1$. By linearity of f it is $f(\vec{x}_0) = \sum_{i=1}^{k} \alpha_i f(\vec{x}_i) \leq \max\{f(\vec{x}_1), \ldots, f(\vec{x}_k)\}$. By maximality of $f(\vec{x}_0)$ it is $f(\vec{x}_0) \geq \max\{f(\vec{x}_1), \ldots, f(\vec{x}_k)\}$. Hence, there is an $i \in \{1, \ldots, k\}$ such that $f(\vec{x}_0) = f(\vec{x}_i)$. □

Extreme Points and Basic Solutions 3

If M is nonempty and unbounded there need not be an optimal solution. However, if there is one, a similar but more elaborate argument can be used to prove the following theorem.

Theorem (2)

If M is nonempty and there exists an K such that $f(\vec{x}) \leq K$ for all $\vec{x} \in M$, then there is an extreme point $\vec{x}^* \in M$ such that $f(\vec{x}^*) \geq f(\vec{x})$ for all $\vec{x} \in M$.

Extreme Points and Basic Solutions 4

Definition: Let $A \in \mathbb{R}^{m \times n}$ be a matrix with column vectors $\vec{a}_1, \ldots, \vec{a}_n$, $M = \{\vec{x} \in \mathbb{R}^{m \times n} \, |(A\vec{x} = \vec{b}) \wedge (\vec{x} \geq \vec{o}) \, \}$, $\vec{x} \in M$, and $S(\vec{x}) = \{\vec{a}_i \mid (i \in \{1, \ldots, n\}) \wedge (x_i \neq 0)\}$. Then $\vec{x}$ is called a *basic solution* if $S(\vec{x})$ is linear independent. Note that linear independence of $S(\vec{x})$ is equivalent to the assertion that there is no $\vec{v} \neq \vec{o}$ such that $A\vec{v} = \vec{o}$ with the property that $v_i \neq 0$ implies $x_i \neq 0$.

Theorem (3)

Let $\vec{x} \in M$. Then the following two statements are equivalent.
(a) $\vec{x}$ is an extreme point of M.
(b) $\vec{x}$ is a basic solution.

Extreme Points and Basic Solutions 5

Proof of Theorem (3).
(a) implies (b): Suppose that $\vec{x}$ is not a basic solution, i.e. $S(\vec{x})$ is not linear independent. Then there is a vector $\vec{v} \neq \vec{o}$ such that $A\vec{v} = \vec{o}$ and $v_i \neq 0$ implies $x_i \neq 0$ for all $i \in \{1, \ldots, n\}$. Thus, there is an $\varepsilon > 0$ such that $\vec{u}_1 = \vec{x} + \varepsilon\vec{v} \geq \vec{o}$ and $\vec{u}_2 = \vec{x} - \varepsilon\vec{v} \geq \vec{o}$. Clearly, $\vec{u}_1, \vec{u}_2 \in M$, $\vec{u}_1 \neq \vec{x} \neq \vec{u}_2$, and $\frac{1}{2}\vec{u}_1 + \frac{1}{2}u_2 = \vec{x}$, i.e. $\vec{x}$ is not an extreme point.
(b) implies (a): Suppose $\vec{x}$ is not an extreme point. Then there are vectors $\vec{v}_1, \vec{v}_2 \in M$ such that $\vec{v}_1 \neq \vec{x} \neq \vec{v}_2$ and an $\alpha \in (0, 1)$ such that $\vec{x} = \alpha\vec{v}_1 + (1 - \alpha)\vec{v}_2$. It follows that $\vec{v}_1 \neq \vec{v}_2$, $A(\vec{v}_1 - \vec{v}_2) = \vec{o}$, and that if the i-th entry of $\vec{v}_1 - \vec{v}_2$ is not 0, then $x_i \neq 0$. This implies that $S(\vec{x})$ is not linear independent, and so $\vec{x}$ is not a basic solution. □

The Simplex Algorithm 1

We consider the LOP in standard form (L) as follows.

Maximize

$$f(\vec{x}) = \vec{c}^T \vec{x}$$

subject to

$$A\vec{x} = \vec{b}$$

and

$$\vec{x} \geq \vec{o}$$

with

$$A \in \mathbb{R}^{m \times n}, \vec{b} \geq \vec{o}, \text{ and } rk(A) = m.$$

Let $M = \{\vec{x} \in \mathbb{R}^n \mid (A\vec{x} = \vec{b}) \wedge (\vec{x} \geq \vec{o})\}$.

The Simplex Algorithm 2

We know so far:

- M has at least one extreme point.
- Any extreme point of M is a basic solution of (L) and vice versa.
- (L) has at most $\binom{n}{m}$ basic solutions.
- If there is an optimal solution of (L), then there is an optimal solution of (L) that is a basic solution.
- If M is nonempty and bounded, then there is an optimal solution of (L).

It follows that (at least) if M is nonempty and bounded, there is a finite algorithm that finds an optimal solution of (L).

The Simplex Algorithm 3

- The simplex algorithm consists of two parts, called Phase 1 and Phase 2.
- The input of Phase 2 is a feasible basic solution.
- The simplex algorithm stops when either an optimal basic has been found, or if it has been detected that the objective function is unbounded on M. In the latter case there is no optimal solution of (L).
- Phase 1 is needed only if there is no feasible basic solution known.
- Phase 1 consists in applying Phase 2 to an auxiliary LOP. Phase 1 stops when a feasible basic solution of (L) has been found or if it has been detected that M is empty.

The Simplex Algorithm 4 - Phase 2

- Let $\vec{x}$ be a feasible basic solution of (L).
- Let $T(\vec{x}) \subseteq \{\vec{a}_1, \dots, \vec{a}_n\}$ be a maximal linear independent set of column vectors of A such that

$$x_k \neq 0 \implies \vec{a}_k \in T(\vec{x}).$$

  (Note: Since $rk(A) = m$ the set $T(\vec{x})$ has m elements.)
- Let $B = \{k \in \{1, \dots, n\} \mid \vec{a}_k \in T(\vec{x})\}$ and $N = \{1, \dots, n\} \setminus B$.
- If $T(\vec{x})$ contains column vectors $\vec{a}_k$ such that $x_k = 0$, the basic solution $\vec{x}$ is called *degenerate*, otherwise it is said to be *non-degenerate*.

The Simplex Algorithm 5 - Phase 2

- $T(\vec{x})$ is a basis of the linear subspace generated by all column vectors of A. Thus, each column vector $\vec{a_j}$ can represented as a linear combination of the elements of $T(\vec{x})$, and this representation is unique:

$$\vec{a_j} = \sum_{k \in B} t_{k,j} \vec{a}_k.$$

- Clearly, if $j \in B$, then $t_{j,j} = 1$ and $t_{k,j} = 0$ if $k \neq j$.
- For $j \in N$ let $u_j = \sum_{k \in B} t_{k,j} c_k$ and $d_j = u_j - c_j$. (Recall that the objective function f is given by $f(\vec{x}) = \vec{c}^T \vec{x}$ with $\vec{c} = (c_1, \ldots, c_n)^T$.)

The Simplex Algorithm 6 - Phase 2

Let $\vec{y} \in \mathbb{R}^n$ be an arbitrary feasible solution of (L), i.e. $\sum_{i=1}^n y_i \vec{a}_i = \vec{b}$ and $\vec{y} \geq \vec{o}$. Then

$$\begin{aligned}
\vec{b} &= \sum_{k \in B} x_k \vec{a}_k \\
&= \sum_{i=1}^{n} y_i \vec{a}_i \\
&= \sum_{i=1}^{n} y_i (\sum_{k \in B} t_{k,i} \vec{a}_k) \\
&= \sum_{k \in B} (\sum_{i=1}^{n} t_{k,i} y_i) \vec{a}_k.
\end{aligned}$$

Since $T(\vec{x})$ is linear independent we know that for all $k \in B$

$$x_k = \sum_{i=1}^{n} t_{k,i} y_i = y_k + \sum_{i \in N} t_{k,i} y_i.$$

The Simplex Algorithm 7 - Phase 2

Consequently, $y_k = x_k - \sum_{i \in N} t_{k,i} y_i$ for all $k \in B$. We can now express $f(\vec{y}) = \vec{c}^T \vec{y}$ as follows

$$\begin{aligned}
\vec{c}^T \vec{y} &= \sum_{i=1}^{n} c_i y_i \\
&= \sum_{k \in B} c_k y_k + \sum_{i \in N} c_i y_i \\
&= \sum_{k \in B} c_k (x_k - \sum_{i \in N} t_{k,i} y_i) + \sum_{i \in N} c_i y_i \\
&= \sum_{k \in B} c_k x_k + \sum_{i \in N} (c_i - \sum_{k \in B} c_k t_{k,i}) y_i \\
&= \vec{c}^T \vec{x} - \sum_{i \in N} d_i y_i
\end{aligned}$$

The Simplex Algorithm 8 - Phase 2

We distinguish three cases.

Case 1 It is $d_i \geq 0$ for all $i \in N$. Then $f(\vec{x}) \geq f(\vec{y})$ for all feasible solutions $\vec{y}$, i.e. $\vec{x}$ is an optimal solution. In this case the algorithm stops.

Case 2 There is an index $j \in N$ such that $d_j < 0$ and $t_{k,j} \leq 0$ for all $k \in B$. Let $\alpha > 0$. Then the vector $\vec{z} \in \mathbb{R}^n$ with

$$z_k = \begin{cases} x_k - \alpha t_{k,j} & \text{for } k \in B \\ \alpha & \text{for } k = j \\ 0 & \text{for } k \in N \text{ and } k \neq j \end{cases}$$

is nonnegative, and

$$A\vec{z} = \sum_{k=1}^{n} z_k \vec{a}_k = \sum_{k \in B} x_k \vec{a}_k - \alpha \sum_{k \in B} t_{k,j} \vec{a}_k + \alpha \vec{a}_j =$$

$$= \sum_{k \in B} x_k \vec{a}_k - \alpha \vec{a}_j + \alpha \vec{a}_j = \sum_{k \in B} x_k \vec{a}_k = A\vec{x} = \vec{b}.$$

The Simplex Algorithm 9 - Phase 2

Consequently, $\vec{z}$ is a feasible solution for all $\alpha > 0$. Since $d_j < 0$,

$$\lim_{\alpha\to\infty} f(\vec{z}) = \lim_{\alpha\to\infty} f(\vec{x}) - d_j\alpha = \infty,$$

i.e. the objective function f is unbounded from above on M. In this case the algorithm stops.

<u>Case 3</u> There is an index $j \in N$ such that $d_j < 0$ and there is an index $\ell \in B$ such that $t_{\ell,j} > 0$. Let $\varepsilon = \min\{x_k/t_{k,j} \mid (k \in B) \wedge (t_{k,j} > 0)\}$. Clearly, $\varepsilon \geq 0$. Suppose that $\varepsilon = x_\ell/t_{\ell,j}$ and define $\vec{z} \in \mathbb{R}^n$ by

$$z_k = \begin{cases} x_k - \varepsilon t_{k,j} & \text{for } k \in B \\ \varepsilon & \text{for } k = j \\ 0 & \text{for } k \in N \text{ and } k \neq j \end{cases} .$$

Similarly as in Case 2 it is not hard to see that $\vec{z}$ is a feasible solution. Furthermore, $z_\ell = 0$ and since $t_{\ell,j} \neq 0$, the j-th column vector $\vec{a}_j$ of A cannot be expressed as a linear combination of the column vectors in $S(\vec{x}) \setminus \{\vec{a}_\ell\}$. Thus, $\vec{z}$ is a feasible basic solution.

The Simplex Algorithm 10 - Phase 2

By $d_j < 0$ and $\varepsilon > 0$, it is $f(\vec{z}) = f(\vec{x}) - d_j\varepsilon \geq f(\vec{x})$. The algorithm starts over with the new feasible basic solution $\vec{z}$.

Special care is needed to avoid 'cycling'. One way to avoid cycling is to apply Bland's rule (in Case 3 above)[1]

<u>Bland's Rule.</u>

- Let $j = \min\{i \in N \mid d_i < 0\}$.
- Let $\ell = \min\{s \in B \mid x_s/t_{s,j} = \varepsilon\}$, with ε as above.

[1] A proof that Bland's rule indeed prevents cycling can be found in: J.Matousek and B. Grtner, *Understanding and Using Linear Programming*, Springer, 2006.

The Simplex Algorithm 11 - Phase 1

If the original LOP is given in standard from (L), and no feasible basic solution is known, an auxiliary LOP can be used to find one.

The Auxiliary LOP
Minimize $g(\vec{x}, \vec{y}) = \sum_{i=1}^{m} y_i$

subject to

$$A\vec{x} + \vec{y} = \vec{b}, \vec{x} \geq \vec{o}, \vec{y} \geq \vec{o}.$$

A feasible basic solution for the auxiliary problem is $\vec{y} = \vec{b}, \vec{x} = \vec{o}$. Thus, the feasible set of the auxiliary LOP nonempty.

The Simplex Algorithm 12 - Phase 1

Theorem (4)

(a) The auxiliary LOP has an optimal solution $(\vec{y}^*, \vec{x}^*)$.
(b) The original LOP has a feasible solution if and only if $\vec{y}^* = \vec{o}$.

Proof. (a) Since $\vec{y} \geq \vec{o}$, we have $g(\vec{x}, \vec{y}) = \sum_{i=1}^{m} y_i \geq 0$. Furthermore, the feasible set of the auxiliary LOP is nonempty. It follows that the auxiliary LOP has an optimal solution.
(b) It is not hard to see that $\vec{x} = \vec{x}_0$ is a feasible solution of the original LOP if and only if $\vec{x} = \vec{x}_0, \vec{y} = \vec{o}$ is a feasible solution of the auxiliary one. A feasible solution of the auxiliary LOP with $\vec{y} = \vec{o}$ is optimal because $g(\vec{x}, \vec{o}) = 0$ for all $\vec{x} \in \mathbb{R}^n$. □

The Simplex Algorithm 13 - Phase 1

Remark.

If the original LOP is given in general form with $\vec{b} \geq \vec{o}$, then a feasible solution of the corresponding LOP in standard form

$$\text{Maximize } \vec{c}^T \vec{x} \text{ subject to } A\vec{x} + I\vec{y} = \vec{b},\ \vec{x} \geq \vec{o}, \text{ and } \vec{y} \geq \vec{o}$$

can be obtained as follows: $\vec{y} = \vec{b}$, $\vec{x} = \vec{o}$.

Simplex Tableaus 1

Consider a system of linear equations $A\vec{x} = \vec{b}$. Recall from linear algebra that $A\vec{x} = \vec{b}$ can be transformed into an equivalent system of linear equations $\hat{A}\vec{x} = \hat{\vec{b}}$ using the following 'row operations'.

(1) Switching two equations.

(2) Multiplying one equation with a **nonzero** factor.

(3) Adding a multiple of an equation to **another** equation.

Here, 'equivalent' means that a vector $\vec{x}_0 \in \mathbb{R}^n$ is a solution of $A\vec{x} = \vec{b}$ if and only if it is a solution of $\hat{A}\vec{x} = \hat{\vec{b}}$.

Simplex Tableaus 2

Now suppose, as above, that $A \in \mathbb{R}^{m \times n}$ with rank m, and $\vec{b} \in \mathbb{R}^m$. Let $\vec{x}_0 \in \mathbb{R}^n$ be a feasible basic solution, i.e. $A\vec{x}_0 = \vec{b}$, $\vec{x}_0 \geq \vec{o}$, and the set $S(\vec{x}_0) = \{\vec{a}_i \mid x_{0i} \neq 0\}$ of column vectors of A is linear independent. Let $T(\vec{x}_0) \supseteq S(\vec{x}_0)$ be a maximal linear independent set of column vectors containing $S(\vec{x}_0)$, $B = \{i \mid \vec{a}_i \in T(\vec{x}_0)\}$, and $N = \{1, \dots, n\} \setminus B$. Note that $|B| = m$ because $rank(A) = m$. In order to simplify our notation we assume henceforth that $B = \{1, \dots, m\}$.

Using the row operations (1),(2), and (3), the system of linear equations $A\vec{x} = \vec{b}$ can be transformed into an equivalent systems of linear equations $\hat{A}\vec{x} = \hat{\vec{b}}$ such that the first m columns of $\hat{A}$ form an $m \times m$ identity matrix I. Thus $\hat{A}$ can be written as $(\hat{A}_B, \hat{A}_N) = (I, \hat{A}_N)$. Note that this implies that $\hat{\vec{b}} = (\vec{x}_0)_B$.

Simplex Tableaus 3

Let $j \in N$ and express the column vector $\vec{a}_j$ as a linear combination of the column vectors in $T(\vec{x}_0)$

$$\vec{a}_j = \sum_{k \in B} t_{k,j} \vec{a}_k.$$

Let $\vec{u} \in \mathbb{R}^n$ with

$$u_i = \begin{cases} t_{k,j} & \text{for } i = k \\ -1 & \text{for } i = j \\ 0 & \text{else} \end{cases}.$$

Then $A\vec{u} = \vec{o}$. It follows that $\hat{A}\vec{u} = \vec{o}$. From the structure of $\hat{A}$ we see that

$$(\hat{A}_N)_{k,j} = t_{k,j}$$

for all $k \in B$ and $j \in N$.

Simplex Tableaus 4

Since

$$t_{k,i} = \begin{cases} 1 & \text{for } k = i \\ 0 & \text{else} \end{cases}$$

for all $i, k \in B$, it finally follows that $(\hat{A})_{k,j} = t_{k,j}$ for all $k, j \in \{1, \ldots, n\}$. Now we extend the system of linear equations $A\vec{x} = \vec{b}$ by appending the equation $\vec{c}^T \vec{x} = z$

$$\begin{pmatrix} A \\ \vec{c}^T \end{pmatrix} \vec{x} = \begin{pmatrix} \vec{b} \\ z \end{pmatrix}.$$

This system of linear equation is equivalent to

$$\begin{pmatrix} \hat{A} \\ \vec{c}^T \end{pmatrix} \vec{x} = \begin{pmatrix} \hat{\vec{b}} \\ z \end{pmatrix}.$$

Simplex Tableaus 5

In this system of linear equations we add for each $k \in B$ the k-th equation multiplied by $-c_k$ to the last newly appended equation $\vec{c}^T \vec{x} = z$. We get

$$\begin{pmatrix} \hat{A} \\ \hat{\vec{c}}^T \end{pmatrix} \vec{x} = \begin{pmatrix} \hat{\vec{b}} \\ z \end{pmatrix},$$

where

$$\hat{c}_i = \begin{cases} 0 & \text{for } i \in B \\ c_i - \sum_{k \in B} c_k t_{k,i} & \text{for } i \in N \end{cases}.$$

Note that $c_i - \sum_{k \in B} c_k t_{k,i} = -d_i$ (see 'The Simplex Algorithm 7'.)

Simplex Tableaus 6

Producing a new basic solution (Case 3 on 'The Simplex Algorithm 9'.) can now be described as follows.

Choose $j \in N$ such that $d_j < 0$. Choose $\ell \in B$ such that $t_{\ell,j} > 0$ and $x_\ell / t_{\ell,j}$ is minimal. If there are more than one candidate for j and ℓ, respectively, apply Bland's rule. Multiply the ℓ-th equation by $1/t_{\ell,j}$ and then, add the resulting ℓ-th equation multiplied by $-t_{k,j}$ to the k-th equation for $k \in \{1, \ldots, m\}$ with $k \neq \ell$. Add the ℓ-th equation multiplied by $-\hat{c}_j = -d_j$ to the equation $\hat{\vec{c}}\vec{x} = z$.

Duality 1

Let $A \in \mathbb{R}^{m \times n}$, $\vec{c} \in \mathbb{R}^n$, and $\vec{b} \in \mathbb{R}^m$. Assume there are $\vec{x} \in \mathbb{R}^n$ and $\vec{y} \in \mathbb{R}^m$ such that $\vec{x} \geq \vec{o}$ and $\vec{y} \geq \vec{o}$, $\vec{y}^T A \geq \vec{c}^T$, and $A\vec{x} \leq \vec{b}$.

Observation *

$$\vec{c}^T \vec{x} \leq \vec{y}^T A \vec{x} \leq \vec{y}^T \vec{b}.$$

We rephrase this observation in terms of two LOP's (P) and (D).

Duality 2

The LOP (P):

Maximize

$$\vec{c}^T\vec{x}$$

subject to

$$A\vec{x} \leq \vec{b} \text{ and } \vec{x} \geq \vec{o}.$$

The LOP (D):

Minimize

$$\vec{y}^T\vec{b}$$

subject to

$$\vec{y}^T A \geq \vec{c}^T \text{ and } \vec{y} \geq \vec{o}.$$

Duality 3

Our initial Observation * now reads as follows.

Theorem (Weak Duality Theorem (WDT))

(i) *If $\vec{x}_0$ and $\vec{y}_0$ are feasible solutions of (P) and (D), then $\vec{c}^T\vec{x} \leq \vec{y}^T\vec{b}$.*

(ii) *If the objective function $\vec{c}^T\vec{x}$ of (P) is unbounded from above, then (D) has no feasible solution.*

(iii) *If the objective function $\vec{y}^T\vec{b}$ of (D) is unbounded from below, then (P) has no feasible solution.*

Remarks:

- (P) is called the *primal LOP*, (D) the *dual LOP*.
- Note that (i) obtains especially for optimal feasible solutions of (P) and (D), respectively.
- It is possible that both LOP's, (P) and (D), don't have feasible solutions.

Duality 4

Theorem (Strong Duality Theorem (SDT))

(1) *If there are feasible solutions of (P) and its objective function $\vec{c}^T\vec{x}$ is bounded from above, then (P) and (D) have optimal feasible solutions $\vec{x}^*$ and $\vec{y}^*$, respectively, and*

$$\vec{c}^T\vec{x}^* = (\vec{y}^*)^T\vec{b}.$$

(2) *If there are feasible solutions of (D) and its objective function $\vec{y}^T\vec{b}$ is bounded from below, then (P) and (D) have optimal feasible solutions $\vec{x}^*$ and $\vec{y}^*$, respectively, and*

$$\vec{c}^T\vec{x}^* = (\vec{y}^*)^T\vec{b}.$$

Duality 5

We want to give a proof of (1) using ideas from the simplex algorithm. The proof is a modification of the one given in the book by Grtner and Matousek[2]. To this end we recall some of the facts we already know.

(a) If the objective function of (P) $\vec{c}^T\vec{x}$ is bounded on its feasible set $M = \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} \leq \vec{b}, \vec{x} \geq \vec{o}\}$ and $M \neq \emptyset$, then (P) has an optimal solution.

(b) (P) has an optimal solution $\vec{x}^*$ if and only if the standard form LOP ($\hat{\text{P}}$) (see next slide) has one.

(c) If ($\hat{\text{P}}$) has an optimal solution, then it has an optimal basic solution too.

---

[2] J.Matousek and B. Grtner, *Understanding and Using Linear Programming*, Springer, 2006.

Duality 6

The LOP ($\hat{\mathrm{P}}$)

Maximize

$$\hat{\vec{c}}^T \hat{\vec{x}}$$

subject to

$$\hat{A}\hat{\vec{x}} = \vec{b}, \ \hat{\vec{x}} \geq \vec{o}$$

with

$$\hat{A} = (A, I), \quad \hat{\vec{x}} = (\vec{x}, \vec{z})^T, \quad \hat{\vec{c}} = (\vec{c}, \vec{o})^T.$$

Note that we do not assume that $\vec{b} \geq \vec{o}$.

Duality 7

Next, we need to reformulate the optimality criterion for basic solutions of ($\hat{\mathrm{P}}$). Let $\hat{\vec{x}}$ be a feasible basic solution of ($\hat{\mathrm{P}}$). We let $\hat{A}_B$ be the matrix consisting of the columns of A forming the basis corresponding to $\hat{\vec{x}}$, and $\hat{A}_N$ the matrix consisting of the remaining columns. Likewise, we write $\vec{h}_B$ and $\vec{h}_N$ for the vectors consisting of the entries of a vector $\vec{h} \in \mathbb{R}^{n+m}$ corresponding to $\hat{A}_B$ and $\hat{A}_N$, respectively. Thus, $\hat{A}\vec{h} = \hat{A}_B\vec{h}_B + \hat{A}_N\vec{h}_N$. Let $\hat{\vec{y}}$ be a feasible solution ($\hat{\mathrm{P}}$). Note that $\hat{A}_B$ is an invertible $m \times m$-matrix.

Duality 8

Since $\hat{\vec{x}}$ and $\hat{\vec{y}}$ are solutions of $(\hat{\mathrm{P}})$ we know

$$\vec{b} = \hat{A}\hat{\vec{x}} = \hat{A}_B\hat{\vec{x}}_b = \hat{A}_B\hat{\vec{y}}_B + \hat{A}_N\hat{\vec{y}}_N.$$

Hence,

$$\hat{\vec{y}}_B = \hat{\vec{x}}_B - \hat{A}_B^{-1}\hat{A}_N\hat{\vec{y}}_N$$

and

$$\hat{\vec{c}}^{\,T}\hat{\vec{y}} = \hat{\vec{c}}_B^{\,T}(\hat{\vec{x}}_B - \hat{A}_B^{-1}\hat{A}_N\hat{\vec{y}}_N) + \hat{\vec{c}}_N^{\,T}\hat{\vec{y}}_N = \hat{\vec{c}}_B^{\,T}\hat{\vec{x}}_B + (\hat{\vec{c}}_N^{\,T} - \hat{\vec{c}}_B^{\,T}\hat{A}_B^{-1}\hat{A}_N)\hat{\vec{y}}_N.$$

Since $\hat{\vec{c}}^{\,T}\hat{\vec{x}} = \hat{\vec{c}}_B^{\,T}\hat{\vec{x}}_B$ it follows that $\hat{\vec{x}}$ is optimal if and only if

$$(\hat{\vec{c}}_N^{\,T} - \hat{\vec{c}}_B^{\,T}\hat{A}_B^{-1}\hat{A}_N)\hat{\vec{y}}_N \leq 0$$

for all feasible solutions $\hat{\vec{y}}$ of $(\hat{\mathrm{P}})$. Because $\hat{\vec{y}} \geq \vec{o}$, we finally get that $\hat{\vec{x}}$ is optimal if and only if

$$\hat{\vec{c}}_B^{\,T}\hat{A}_B^{-1}\hat{A}_N \geq \hat{\vec{c}}_N^{\,T}.$$

Duality 9

Since $\hat{\vec{c}}_B^{\,T}\hat{A}_B^{-1}\hat{A}_B = \hat{\vec{c}}_B^{\,T}$ the inequality $\hat{\vec{c}}_B^{\,T}\hat{A}_B^{-1}\hat{A}_N \geq \hat{\vec{c}}_N^{\,T}$ is equivalent to

$$\hat{\vec{c}}_B^{\,T}\hat{A}_B^{-1}\hat{A} = \hat{\vec{c}}_B^{\,T}\hat{A}_B^{-1}(\hat{A}_B, \hat{A}_N) = (\hat{\vec{c}}_B^{\,T}, \hat{\vec{c}}_B^{\,T}\hat{A}_B^{-1}\hat{A}_N) \geq (\hat{\vec{c}}_B^{\,T}, \hat{\vec{c}}_N^{\,T}) = \hat{\vec{c}}^{\,T}.$$

We recall that $\hat{\vec{c}}^{\,T} = (\vec{c}^{\,T}, \vec{o}^{\,T})$ and $\hat{A} = (A, I)$. Hence, $\hat{\vec{c}}_B^{\,T}\hat{A}_B^{-1}\hat{A} \geq \hat{\vec{c}}^{\,T}$ is equivalent to $\hat{\vec{c}}_B^{\,T}\hat{A}_B^{-1}A \geq \vec{c}^{\,T}$ and $\hat{\vec{c}}_B^{-1}\hat{A}_B \geq \vec{o}^{\,T}$. Thus, we have proved the following proposition.

Proposition

A feasible basic solution $\hat{\vec{x}}$ of $(\hat{P})$ is optimal if and only if $\hat{\vec{c}}_B^{\,T}\hat{A}_B^{-1}A \geq \vec{c}^{\,T}$ and $\hat{\vec{c}}_B^{-1}\hat{A}_B \geq \vec{o}^{\,T}$.

Duality 10

Now we are ready to prove assertion (1) of SDT.

Proof. If (P) is feasible and bounded, it follows from (a) that (P) has an optimal solution. By (b) and (c) it follows that ($\hat{\text{P}}$)has an optimal basic solution $\hat{\vec{x}}$. We let $\vec{u}^T = \hat{\vec{c}}_B^T \hat{A}_B^{-1}$. By the above proposition $\vec{u}$ is a feasible solution of (D). Recall that $\hat{\vec{x}}_B = \hat{A}_B^{-1}\vec{b}$. Hence,

$$\vec{u}^T \vec{b} = \hat{\vec{c}}_B^T \hat{A}_B^{-1} \vec{b} = \hat{\vec{c}}_B^T \hat{\vec{x}}_B = \hat{\vec{c}}^T \hat{\vec{x}} = \vec{c}^T \vec{x}.$$

By the WDT it follows that $\vec{u}$ is an optimal solution of ($\hat{\text{P}}$). $\square$

Duality 11

The following 'Dualization Recipe' is taken from the book by Grtner and Matousek[3].

	Primal LOP	Dual LOP
Variables	$x_1, x_2, \dots, x_n$	$y_1, y_2, \dots, y_m$
Matrix	A	A^T
Right hand side	$\vec{b}$	$\vec{c}$
Objective function	$\max \vec{c}^T \vec{x}$	$\min \vec{b}^T \vec{y}$
Constraints	ith constraint has $\leq$	$y_i \geq 0$
	ith constraint has $\geq$	$y_i \leq 0$
	ith constraint has $=$	$y_i \in \mathbb{R}$
	$x_j \geq 0$	jth constraint has $\geq$
	$x_j \leq 0$	jth constraint has $\leq$
	$x_j \in \mathbb{R}$	jth constraint has $=$

[3] J.Matousek and B. Grtner, *Understanding and Using Linear Programming*, Springer, 2006.