

References for linear programming

- Donald Goldfarb, Michael J. Todd, Linear Programming, in G.L. Nemhauser, A.H.G. Rinnooy Kan, M. J. Todd (eds.), Optimization North-Holland 1989, pp. 73-170

give for free until page 130 google goldfarb todd linear programming

- R.S. Garfinkel, G.L. Nemhauser, Integer Programming, Wiley, 1972.
- C.H. Papadimitriou, K. Steiglitz, Combinatorial Optimization: Algorithms and Complexity, Prentice-Hall, 1982.

available as free ebook on the web.

- V. Chvátal, Linear Programming, Freeman, 1983.
- Norbert Blum, Algorithmen und Datenstrukturen, 2. Auflage, Oldenbourg, 2013

3. Linear programming

For some optimization problems like

- the maximum flow problem
- the maximum matching problem
- the maximum weighted matching problem

⋮

we have developed specialized algorithms. For doing this we have first analyzed the problem for finding combinatorial structures which make it possible to construct an efficient algorithm which solves the considered optimization problem. Normally, such an algorithm depends on the special combinatorial structures and cannot be applied to other optimization problems.

Sometimes one is not able to find such nice combinatorial structures for a given optimization problem such that one can develop an efficient specialized algorithm

for the solution of the problem.



The development of a general method for the solution of multitude of optimization problems is useful and interesting.

Question:

What is an appropriate specification of an optimization problem?

The input of an optimization problem P consists of

- a description of the set F of feasible solutions

and

- an objective function $z: F \rightarrow \mathbb{R}$

The goal is to compute an $x \in F$ such that

- $z(x) \leq z(y) \quad \forall y \in F$ if P is a minimization problem
- $z(x) \geq z(y) \quad \forall y \in F$ if P is a maximization problem.

For the development of an efficient method we have to give a precise description of F and of z . The kind of description has to be general enough such that it is possible to describe a lot of optimization problems. Also, it has to be restricted enough such that the development of an efficient method for the computation of $x \in F$ which optimizes $z(x)$ is possible.

Notations:

Let $n \in \mathbb{N}$ and $b, c_1, c_2, \dots, c_n \in \mathbb{R}$.

- A function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ with

$$f(x_1, x_2, \dots, x_n) = \sum_{j=1}^n c_j x_j$$

is called a linear function.

- $f(x_1, x_2, \dots, x_n) = b$ is a linear equation

- $f(x_1, x_2, \dots, x_n) \leq b$ and $f(x_1, x_2, \dots, x_n) \geq b$ are called linear inequations.

- Linear equations and also linear inequations

are called linear restrictions.

A linear programming problem or shortly a linear program is the problem of the maximization or the minimization of a linear function such that a set of linear restrictions is fulfilled. If in addition, the components of the feasible solution vector have to be integers then we obtain a linear integer programming problem or shortly a linear integer program.

In the following, let always $|V| = n$ and $|E| = m$.

Examples:

a) The weighted matching problem

Given a weighted undirected graph $G = (V, E, w)$ the goal is to compute a maximum weighted matching of G .

Let $V = \{1, 2, \dots, n\}$ and assume that the edges in E are numbered from 1 to m .

For the definition of the corresponding linear program we need the vectors

$$w^T = (w_1, w_2, \dots, w_m) \quad \text{and}$$

$$x^T = (x_1, x_2, \dots, x_m).$$

If (i, j) is the k -th edge in E then we identify x_{ij} and x_k . In the solution vector x^T the variable x_{ij} has the value 1 iff (i, j) is an edge in the corresponding matching and 0 otherwise.

The following linear integer program is equivalent to the maximum weighted matching problem:

$$\text{maximize } z(x) = w^T x$$

$$\sum_{(i,j) \in E} x_{ij} \leq 1 \quad 1 \leq i \leq n$$

$$x \geq 0$$

$$x \text{ integral}$$

The objective function z computes the weight of the matching and the restrictions take care that the feasible solution space contains exactly the vectors which correspond to a matching.

b) The maximum flow problem

Given a flow network $G = (V, E, c, s, t)$ we have to compute a maximum flow from the source s to the sink t .

w. l. o. g. we can assume that the nodes in V are numbered such that $s = 1$ and $t = n$.

Let c_{ij} be the capacity of the edge (i, j) .

For $2 \leq i \leq n-1$ let

$$k(i) := \sum_{j: (j,i) \in E} x_{ji} - \sum_{j: (i,j) \in E} x_{ij}$$



$$\max \quad z(x) = \sum_{(i,n) \in E} x_{in}$$

$$k(i) = 0 \quad 2 \leq i \leq n-1$$

$$x_{ij} \leq c_{ij} \quad \forall (i,j) \in E$$

$$x_{ij} \geq 0 \quad \forall (i,j) \in E$$

The objective function computes the flow which enters the sink. $k(i) = 0$ implies that Kirchhoff's law is fulfilled with respect to the node $i \in V$. The other restrictions take care that the capacity conditions are fulfilled for all edges in E .

c) Vertex cover

Given an undirected graph $G = (V, E)$ our goal is to compute a vertex cover of minimum size. $V' \subseteq V$ is a vertex cover if $v \in V'$ or $w \in V'$ $\forall (v,w) \in E$. Assume that $V = \{1, 2, \dots, n\}$. For each node $i \in V$ we associate the variable x_i . $x_i = 1$ iff i is a node in the corresponding vertex cover and $x_i = 0$ otherwise.



$$\min z(x) = \sum_{i \in V} x_i$$

$$x_i + x_j \geq 1 \quad \forall (i, j) \in E$$

$$x_i \geq 0 \quad 1 \leq i \leq n$$

$$x_i \text{ integral} \quad 1 \leq i \leq n.$$

Exercise:

Prove the equivalence of the defined linear programs and linear integer programs, respectively and the corresponding optimization problems.

3.1 Foundations

First we shall show that given any linear program LP it is easy to transform LP to an equivalent linear program LP' which fulfils certain normal form properties. Hence, for the development of algorithms we can always assume that the given linear program is in a certain normal form.

Replacement of unconstrained variables

An unconstrained variable $x_j \leq 0$ can be replaced by two nonnegative variables x_j' and x_j'' . For doing this, we add the restrictions

$$x_j = x_j' - x_j'', \quad x_j' \geq 0 \text{ and } x_j'' \geq 0.$$

Turning of an inequality

We can turn the relation of an inequality by multiplying both sides of the inequality by -1 .

Replacement of inequalities by equalities

An inequality

$$\sum_{j=1}^n a_{ij} x_j \leq b_i$$

can be converted to an equality by adding a nonnegative slack variable x_{n+i} .

For doing this, we replace this inequality by the equality

$$\sum_{j=1}^n a_{ij} x_j + x_{n+i} = b_i$$

and add the restriction

$$x_{n+i} \geq 0.$$

Replacement of equalities by inequalities

We can replace the equality

$$\sum_{j=1}^n a_{ij} x_j = b_i$$

by the two inequalities

$$\sum_{j=1}^n a_{ij} x_j \leq b_i \quad \text{and} \quad \sum_{j=1}^n a_{ij} x_j \geq b_i.$$

Replacement of the optimization operation

Maximizing the linear function $c^T x$ is equivalent to minimizing the linear function $-c^T x$.

Often, linear programs are written in one of the following two forms:

canonical form

$$\begin{aligned} \min \quad & z(x) = c^T x \\ & Ax \leq b \\ & x \geq 0 \end{aligned}$$

standard form

$$\begin{aligned} \min \quad & z(x) = c^T x \\ & Ax = b \\ & x \geq 0 \end{aligned}$$

where $A \in \mathbb{R}^m \times \mathbb{R}^n$, $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ and x is an n -vector of variables.

Exercise:

Show that any linear program can be transformed into an equivalent linear program in canonical form and in standard form, respectively.

What is the effect to the size of the linear program?

For the understanding of linear programming, the geometric interpretation of linear programming and the algebraic characterization of some geometric concepts are very important.

Notations:

Let $x, y \in \mathbb{R}^n$ be any two points in the n -dimensional vector space $\mathbb{R}^n$. Each point $z \in \mathbb{R}^n$ such that

$$z = \lambda x + (1-\lambda)y, \quad \lambda \in [0, 1]$$

is a convex combination of x and y .

If $\lambda \neq 0$ and $\lambda \neq 1$ then z is called to be a strict convex combination of x and y .

More generally, for $x_1, x_2, \dots, x_t \in \mathbb{R}^n$, each point

$$z = \lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_t x_t \quad \text{with } \lambda_i \in [0, 1], \\ 1 \leq i \leq t \text{ and } \sum_{i=1}^t \lambda_i = 1$$

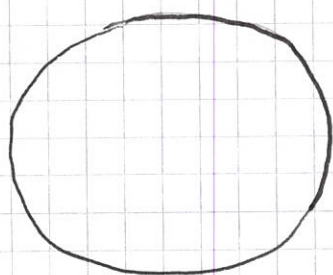
is a convex combination of the points $x_1, x_2, \dots, x_t$.

A subset $C \subseteq \mathbb{R}^n$ is called convex if for all points $x, y \in C$ the set C contains also each convex combination of x and y .

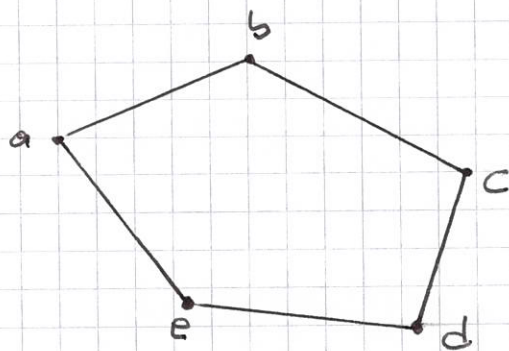
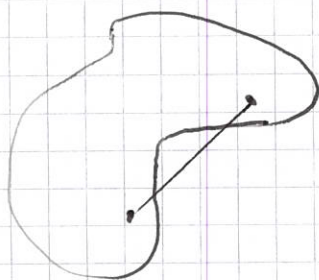
Example:

$\mathbb{R}^n$, $\emptyset$ and $\{x\}$, $x \in \mathbb{R}^n$ are convex.
 In $\mathbb{R}$ each interval is convex and each convex subset is an interval.

$C \subseteq \mathbb{R}^n$ is convex iff for any points $x, y \in C$, all points on the line segment which connects x and y are in C .



convex set

Set of all convex combinations of a, b, c, d and e .

nonconvex set

An extreme point of a convex set C is a point $x \in C$ which is not a convex combination of two other points in C .

Let $a \in \mathbb{R}^n \setminus \{0\}$ and $b \in \mathbb{R}$. Then

$$H = \{x \in \mathbb{R}^n \mid a^T x = b\}$$

is called a hyperplane. The sets

$$\bar{H} = \{x \in \mathbb{R}^n \mid a^T x \leq b\} \quad \text{and}$$

$$\bar{H} = \{x \in \mathbb{R}^n \mid a^T x \geq b\}$$

are closed half-spaces.

The hyperplane H associated with the half-space $\bar{H}$ is the bounding hyperplane of that half-space.

The intersection of a finite number of closed half-spaces is called a convex polyhedron. If it is nonempty and bounded then it is called convex polytope, or simply polytope.

Exercise

Prove that closed half-spaces and that the intersection of convex sets be convex.

Since closed half-spaces and the intersection of convex sets are convex, a convex polyhedron is also convex.

The set of feasible solutions of a linear programming problem

$$F = \{x \in \mathbb{R}^n \mid Ax \leq b, x \geq 0\}$$

is a convex polyhedron since it is the intersection of the half-spaces defined by the

inequalities

$$a_1^T x \leq b_1, a_2^T x \leq b_2, \dots, a_m^T x \leq b_m$$

and

(*)

$$e_1^T x \geq 0, e_2^T x \geq 0, \dots, e_n^T x \geq 0$$

where a_i^T is the i -th row of the matrix A and e_i^T is the i -th row of the $n \times n$ identity matrix.

A (linear) subspace S of $\mathbb{R}^n$ is a subset of $\mathbb{R}^n$ which is closed under vector addition and scalar multiplication. Equivalently, a subspace is the set of those points in $\mathbb{R}^n$ fulfilling a set of homogeneous linear equalities; i.e.,

$$S = \{x \in \mathbb{R}^n \mid Ax = 0\}$$

for a matrix $A \in \mathbb{R}^m \times \mathbb{R}^n$.

The maximal number of linearly independent vectors in S is called the dimension of S and is denoted by dim(S). Note that

$$\dim(S) = n - \text{rank}(A).$$

An affine subspace S_a of $\mathbb{R}^n$ is a subspace S of $\mathbb{R}^n$ which is shifted by a vector $u \in \mathbb{R}^n$; i.e.,

$$S_a = \{u + x \mid x \in S\}.$$

(104)

The dimension $\dim(S_a)$ of S_a is the same as the dimension of S . Equivalently, an affine subspace of $\mathbb{R}^n$ is the set of those points in $\mathbb{R}^n$ fulfilling a set of inhomogeneous equalities; i.e.,

$$S_a = \{x \in \mathbb{R}^n \mid Ax = b\}$$

for a matrix $A \in \mathbb{R}^m \times \mathbb{R}^n$ and or $b \in \mathbb{R}^m$.

Exercise:

Show that $S_a \subseteq \mathbb{R}^n$ is affine iff for any $x, y \in S_a$ and any $-\infty < \lambda < \infty$ always $z = \lambda x + (1-\lambda)y \in S_a$.

A hyperplane in $\mathbb{R}^n$ is an $(n-1)$ -dimensional affine subspace of $\mathbb{R}^n$. We say that S_a is parallel to S if $S_a = \{u+x \mid x \in S\}$ for a vector $u \in \mathbb{R}^n$.

The dimension of any subset $C \subseteq \mathbb{R}^n$ is the minimal dimension of an affine subspace of $\mathbb{R}^n$ which contains C ; i.e.,

$$\dim(C) := \min \{ \dim(S_a) \mid C \subseteq S_a \text{ and } S_a \text{ is an affine subspace of } \mathbb{R}^n \}.$$

A supporting hyperplane of a convex set $C \subseteq \mathbb{R}^n$ is a hyperplane H such that $H \cap C \neq \emptyset$ and $C \subseteq \bar{H}$, one of the two closed half-spaces associated with H .

Let $P \subseteq \mathbb{R}^n$ be a convex polyhedron and H be any supporting hyperplane of P . The intersection $P \cap H$ defines a face of P .

We distinguish three kinds of faces

- vertex face of dimension 0
- edge face of dimension 1
- facet face of dimension $n-1$.

If $P = \{x \in \mathbb{R}^n \mid Ax \leq b, x \geq 0\}$ then every facet of P corresponds to the intersection of P with a half-space defined by one of the linear restrictions in (*).

However, not all such intersections necessarily define facets since some of the inequalities may be redundant; i.e., deleting them from the definition of P does not change P .

Vertices of a convex polyhedron P are obviously extreme points of P . Edges are either line segments which connect neighbouring vertices or are semi-infinite lines emanating from a vertex.

Let $F = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$ be the set of feasible solutions of a linear program in standard form.

Since F contains an infinite number of points we cannot consider each point in F for the computation of a solution of the given linear program.



Question:

Is it always possible to solve a linear program by the consideration of a finite number of points? If the answer is "yes", which points should be considered?

For answering these questions, we characterize the vertices of the polyhedron

$$P = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}.$$

Theorem 3.1

A point $x \in P = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$ is a vertex of P iff the columns corresponding to positive components of x are linearly independent.

Proof:

W. l. o. g., let us assume the first p components of x are positive and the last $n-p$ components of x are zero. Let

$$x = \begin{pmatrix} \bar{x} \\ 0 \end{pmatrix}, \quad \bar{x} > 0 \quad \text{and let } \bar{A} \text{ be the matrix}$$

which consists of the first p columns of A .

Then

$$Ax = \bar{A}\bar{x} = b.$$

" $\Rightarrow$ "

Suppose that the columns of $\bar{A}$ are not linearly independent.

Then there exists a vector $\bar{w} \neq 0$ such that

$$\bar{A}\bar{w} = 0.$$

Hence, for all $\varepsilon > 0$ there hold

$$\bar{A}(\bar{x} + \varepsilon\bar{w}) = \bar{A}(\bar{x} - \varepsilon\bar{w}) = \bar{A}\bar{x} = b.$$

Choose ε small enough such that

$$\bar{x} + \varepsilon\bar{w} \geq 0 \quad \text{and} \quad \bar{x} - \varepsilon\bar{w} \geq 0.$$

Then both points

$$y' = \begin{pmatrix} \bar{x} + \varepsilon\bar{w} \\ 0 \end{pmatrix} \quad \text{and} \quad y'' = \begin{pmatrix} \bar{x} - \varepsilon\bar{w} \\ 0 \end{pmatrix}$$

are points in P .

Since, $x = \frac{1}{2}(y' + y'')$, x cannot be a vertex.

$\Rightarrow$

In the case that x is a vertex of P , the columns of A have to be linearly independent.

" $\Leftarrow$ "

Suppose now that x is not a vertex of P .

Then there exists $y', y'' \in P$, $y' \neq y''$ and $\lambda \in]0, 1[$ such that

$$x = \lambda y' + (1 - \lambda) y''.$$

Since $x, y' \in P$ it holds

$$A(x - y') = Ax - Ay' = b - b = 0.$$

Further, since $\lambda > 0$ and $1 - \lambda > 0$, the last $n - p$ components of y' and hence, also the last $n - p$ components of $x - y'$ have to be 0.

$\Rightarrow$

$A(x - y')$ is a linear combination of the columns in $\bar{A}$.

$\Rightarrow$ The columns in $\bar{A}$ are linear dependent.

Hence, if the columns in $\bar{A}$ are linearly independent then x is a vertex of P ,
a contradiction.

Let A be an $(m \times n)$ -matrix. If $\text{rank}(A) = m$ then we obtain an equivalent characterization of the vertices of P which leads to an answer to the question posed above.

$$\text{rank}(A) = m \Rightarrow m \leq n.$$

The case $m > n$ will be reduced to the case $m \leq n$ later on.

Let B be any nonsingular $m \times m$ matrix composed of m linearly independent columns of A ; i.e., B is a basis of A . The components of x corresponding to the columns of B are called basic variables; the other components of x are called nonbasic variables with respect to the basis B .

A point $x \in \mathbb{R}^n$ with $Ax = b$ and the property that all nonbasic variables with respect to B are equal to zero is said to be a basic solution with respect to the basis B .

Given a basis B , we obtain after setting the corresponding nonbasic variables to zero the following system of m equations in m unknowns:

$$B x_B = b$$

which is uniquely solvable for the basic variables x_B . If a basic solution x with respect to a basis B is nonnegative then it is called a basic feasible solution.

The following corollary is a direct consequence of Theorem 2.1.

Corollary 3.1

A point $x \in P$ is a vertex of P iff x is a basic feasible solution with respect to some basis B .

There are $\binom{n}{m}$ possibilities to choose

m columns of an $m \times n$ matrix A . Hence, we obtain the following corollary.

Corollary 2.2

The polyhedron P has only a finite number of vertices.

Exercise:

Let P be a polytope. Prove that each point $x \in P$ is a convex combination of the vertices of P .

(Hint: Prove first the assertion for all vertices, then for each point on an edge, then for all points on a facet and finally for each point in the interior of P .)

A vector $d \in \mathbb{R}^n \setminus \{0\}$ is called a direction of a polyhedron P if for each point $x_0 \in P$, the ray $\{x \in \mathbb{R}^n \mid x = x_0 + \lambda d, \lambda \geq 0\}$ lies entirely in P . Obviously P is unbounded iff P has a direction.

The following lemma characterizes the directions of a polyhedron.

Lemma 3.1

Let $d \neq 0$. Then d is a direction of

$P = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$ iff $Ad = 0$ and $d \geq 0$.

Proof:

" $\Rightarrow$ "

Suppose that d is a direction of P . Then

$$\{x \in \mathbb{R}^n \mid x = x_0 + \lambda d, \lambda \geq 0\} \subseteq P$$

for all $x_0 \in P$.

Assume that $d \neq 0$.

Then there exists a component d_i of d with $d_i < 0$.

For λ' large enough, for any $x_0 \in P$ there holds

$$x_0^i + \lambda' d_i < 0$$

where x_0^i is the i -th component of x_0 .

$\Rightarrow$

$x_0 + \lambda' d \neq 0$ and hence, $x_0 + \lambda' d \notin P$.

This contradicts that d is a direction of P .

Therefore

$$d \geq 0.$$

Assume that $Ad \neq 0$.

$\Rightarrow$

$$A\lambda d \neq 0 \quad \text{for } \lambda > 0.$$

This implies for any point $x_0 \in P$

$$A(x_0 + \lambda d) = Ax_0 + A\lambda d = b + A\lambda d \neq b.$$

$$\Rightarrow x_0 + \lambda d \notin P$$

This contradicts that d is a direction of P .

Therefore

$$Ad = 0.$$

$\Leftarrow^4$

Suppose that $Ad = 0$ and $d \geq 0$.

$\Rightarrow$

$$A(x_0 + \lambda d) = Ax_0 + A\lambda d = b + \lambda Ad = b$$

for all $x_0 \in P$ and all $\lambda \geq 0$.

Definition $\Rightarrow$

d is a direction of P .



Lemma 3.1 implies directly that for each direction $d \in \mathbb{R}^n$ of a polyhedron P and each $\lambda > 0$, the vector λd is also a direction of P .

If $x \in P$ is not a convex combination of the vertices of P then x can be described as the sum of a point which is a convex combination of the vertices and a direction of P . This observation gives us the following simple representation theorem:

Theorem 3.2

Let $P \subseteq \mathbb{R}^n$ be any polyhedron and let

$\{v_i \mid i \in I\}$ be the set of vertices of P . Then every point $x \in P$ can be represented as

$$x = \sum_{i \in I} \lambda_i v_i + d$$

where $\sum_{i \in I} \lambda_i = 1$, $\lambda_i \geq 0$ for all $i \in I$ and either $d = 0$ or d is a direction of P .

Proof:

exercise

■

Theorem 3.2 implies directly the following corollary.

Corollary 3.3

A nonempty polyhedron $P = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$ has at least one vertex.

Now we can prove the Fundamental theorem of linear programming.

Theorem 3.3

Let P be a nonempty polyhedron. Then the minimum value of $z(x) = c^T x$ for $x \in P$ is attained at a vertex of P or z has no lower bound on P .

Proof:

If P has a direction d with $c^T d < 0$ then

P is unbounded and the value of z converges on the direction d to $-\infty$.

Otherwise, the minimum is attained at points which can be expressed as convex combinations of the vertices v_i of P . Let

$$\hat{x} = \sum_{i \in I} \lambda_i v_i$$

be any such a point where $\{v_i \mid i \in I\}$ is the set of vertices of P , $\sum_{i \in I} \lambda_i = 1$ and $\lambda_i \geq 0$ for all $i \in I$. Then

$$c^T \hat{x} = c^T \sum_{i \in I} \lambda_i v_i = \sum_{i \in I} \lambda_i c^T v_i$$

$$\geq \min \{ c^T v_i \mid i \in I \}$$

Hence, the minimum of z is attained at a vertex of P . ■

Theorem 3.3 implies that for getting an optimal solution of a linear programming problem it suffices to consider basic feasible solutions and to investigate if there is a direction along which $z \rightarrow -\infty$.

Assume that ^{for} a given linear program

$$\begin{aligned}\min \quad z(x) &= c^T x \\ Ax &= b \\ x &\geq 0\end{aligned}$$

there is a finite optimal solution.

Since the number of basic solutions can be $\binom{n}{m}$ where m is the number of rows and n is the number of columns of A , for large m and n we cannot consider all basic solutions for the computation of an optimal feasible basic solution. Hence, we need

1. a strategy for the consideration of the basic solutions and
2. a criterion which decides if the current considered feasible basic solution is optimal

2.2 The simplex method

Goal:

Development of a method to solve the linear program

$$\begin{aligned}\min \quad z(x) &= c^T x \\ Ax &= b \\ x &\geq 0\end{aligned}$$

where A is an $(m \times n)$ -matrix of row rank m . The case $\text{rank}(A) < m$ will be discussed later on.

Geometric motivation

- Start in any vertex x_0 of the polyhedron

$$P = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$$

Go from vertex to vertex along edges of P that are "downhill" with respect to the objective function $z(x) = c^T x$, generating a sequence of vertices with strictly decreasing objective value.

$\Rightarrow$

Once the method leaves a vertex the method can never return to that vertex.

$\Rightarrow$

In a finite number of steps, a vertex will be reached which is optimal, or an edge will be chosen which goes off to infinity and along which z goes to $-\infty$.

Question:

How to convert the above geometric description of the simplex method into an algebraic and, hence, computational form?

The vertex x_0 corresponds to a basic feasible solution

$$x_0 = \begin{pmatrix} x_B \\ x_N \end{pmatrix} = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}.$$

The corresponding objective value $z(x_0)$ is obtained by

117

$$z(x_0) = c_B^T B^{-1} b$$

where c_B contains exactly the components of c which correspond to the basic variables x_B .

Goal:

The computation of a so-called downhill edge which starts in x_0 and on which the objective value $z(x)$ strictly decreases.

Two cases are possible:

1. The downhill edge ends in a neighboured vertex x' with $z(x') < z(x_0)$.
2. The edge has an infinite length such that the objective value $z(x)$ converges on the edge to $-\infty$.

In the second case, the algorithm knows that no finite optimum exists and terminates. In the first case, the algorithm looks for a downhill edge which starts in x' . If no such an edge exists, the algorithm terminates.

To prove the correctness of the method we have to show that a basic feasible solution for which no downhill edge exists is always an optimum solution.

To transform the geometric consideration into an algebraic method, we have to solve the following problems:

1. If $P \neq \emptyset$ then find a vertex of P .
2. Given a vertex of P compute a downhill edge which starts in this vertex if such an edge exists. If no such an edge exists this has to be established.

First we shall investigate the second problem.

Let B be the basis corresponding to a given vertex of the polyhedron P . Let

$$x_0 = \begin{pmatrix} x_B \\ x_N \end{pmatrix} = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$$

be the corresponding basic feasible solution where $A = [B, N]$ and $c^T = [c_B^T, c_N^T]$ are partitioned with respect to basic and nonbasic variables. $Ax = b$ can be described as

$$Bx_B + Nx_N = b$$

Since B is nonsingular the inverse matrix B^{-1} exists.

$\Rightarrow$

In dependence to the variables x_N corresponding to the nonbasic variables, the

values x_B corresponding to the basic variables can be expressed as follows:

$$(1) \quad x_B = B^{-1}b - B^{-1}N x_N$$

After the elimination of x_B in the equation

$$z(x_0) = c_B^T x_B + c_N^T x_N$$

we obtain

$$(2) \quad z(x_0) = c_B^T B^{-1}b - (c_B^T B^{-1}N - c_N^T) x_N$$

Two bases are called neighbourhood if they differ only in one column. A basic solution where at least one basic variable has the value zero is called degenerate. Otherwise, the basic solution is nondegenerate.

Neighbourhood vertices of the polyhedron correspond to neighbourhood bases.

$\Rightarrow$

Going from one vertex to a neighbourhood vertex is equivalent to the exchange of one column of the corresponding basis by another where each other column of the basis remains to be unchanged.

For simplicity assume that the basic feasible solution x_0 is nondegenerate.

going from x_0 to a neighbored vertex corresponds to increasing the value of a nonbasic variable x_j where all other nonbasic variables remain to be zero.

Question:

- How to find an appropriate nonbasic variable x_j ?

In other words

- How to find a downhill edge which starts in x_0 ?

For getting an answer to this question let us consider equation (2) again.

Increasing the nonbasic variable x_j and all other nonbasic variables remain to be zero implies that the second summand has the value

$$-(c_B^T B^{-1} a_j - c_j) x_j$$

where a_j is the j -th column of A .

$$\bar{c}_j := -(c_B^T B^{-1} a_j - c_j)$$

is called reduced cost for x_j .

$\Rightarrow$

x_j corresponds to a downhill edge iff $\bar{c}_j < 0$.

It is useful to combine the equations (1) and (2). (12)

~>

$$\begin{bmatrix} z(x) \\ x_B \end{bmatrix} = \begin{bmatrix} c_B^T B^{-1} b \\ B^{-1} b \end{bmatrix} - \begin{bmatrix} c_B^T B^{-1} N - c_N^T \\ B^{-1} N \end{bmatrix} x_N$$

Let R be the set of the indices of the columns in N and let

$$z(x) = x_{B_0} \quad \text{and} \quad x_B = (x_{B_1}, x_{B_2}, \dots, x_{B_m})^T.$$

For simplification of the notation we define

$$y_0 := \begin{bmatrix} y_{00} \\ y_{10} \\ \vdots \\ y_{m0} \end{bmatrix} := \begin{bmatrix} c_B^T B^{-1} b \\ B^{-1} b \end{bmatrix}$$

and for $j \in R$; i.e., a_j is a column in N ,

$$y_j := \begin{bmatrix} y_{0j} \\ y_{1j} \\ \vdots \\ y_{mj} \end{bmatrix} := \begin{bmatrix} c_B^T B^{-1} a_j - c_j \\ B^{-1} a_j \end{bmatrix}.$$

Then we can write (1) and (2) for $i = 0, 1, \dots, m$ as follows

$$(3) \quad x_{B_i} = y_{i0} - \sum_{j \in R} y_{ij} x_j.$$

If we set in (3)

$$x_j = 0 \quad \text{for all } j \in R$$

then we obtain the basic solution which corresponds to the basis B .

Definition $\Rightarrow$

$$y_{0j} = -\bar{c}_j \quad \text{for all } j \in R.$$

Suppose that x_B is nondegenerate and that

$$y_{0q} > 0 \quad \text{for any } q \in R.$$

Increasing x_q and simultaneous fixing of the other nonbasic variables to zero decreases x_{B_0} proportional to y_{0q} .

Moreover, each x_{B_i} is a linear function of x_q and decreased proportional to y_{iq} .

If $y_{iq} > 0$ then

$$x_{B_i} \geq 0 \quad \text{as long as} \quad x_q < \frac{y_{i0}}{y_{iq}}.$$

At the moment when

$$x_q = \frac{y_{i0}}{y_{iq}}$$

there holds $x_{B_i} = 0$.

If $y_{iq} \leq 0$ for $1 \leq i \leq m$ then x_q can be increased arbitrary without a basic variable

(123)

x_{B_i} , $1 \leq i \leq m$ becomes to be negative such that the solution remains to be feasible.

$\Rightarrow$

We can improve the value of the objective function within the feasible polyhedron arbitrary.

$\Rightarrow$

The given linear program is unbounded.

Exercise:

Show that every unbounded linear program has such a basic feasible solution.

$\leadsto$

Theorem 3.4

A linear program LP is unbounded iff there is a basic feasible solution x_B and a nonbasic variable x_q such that for the vector y_q corresponding to x_q there hold:

$$y_{0q} > 0 \text{ and } y_{iq} \leq 0 \text{ for } 1 \leq i \leq m.$$

For the development of the simplex method we assume that the given linear program is bounded.

Let x_{B_r} be any basic variable with

$$0 < \frac{y_{p0}}{y_{pq}} = \min_{1 \leq i \leq m} \left\{ \frac{y_{i0}}{y_{iq}} \mid y_{iq} > 0 \right\}$$

If we increase x_q to $\frac{y_{p0}}{y_{pq}}$ and fix the other nonbasic variables to zero then we obtain

$$x_q = \frac{y_{p0}}{y_{pq}}$$

$$x_{B_i} = y_{i0} - y_{iq} \frac{y_{p0}}{y_{pq}} \quad \text{for } i = 0, 1, \dots, m.$$

We obtain a new basic feasible solution with

$$x_q > 0, \quad x_{B_p} = 0 \quad \text{and} \quad x_{B_0} = y_{00} - y_{0q} \frac{y_{p0}}{y_{pq}}.$$

Since $y_{0q} > 0$ and $\frac{y_{p0}}{y_{pq}} > 0$ the value x_{B_0} decreases strictly.

Next we shall investigate the computation of the values corresponding to the new basis the new y_{ij} 's in the equations (3).

We obtain the value corresponding to the new basic variable x_q if we solve the equation in (3) which corresponds to x_{B_p} for x_q . After doing this we obtain

$$(4) \quad x_q = \frac{y_{p0}}{y_{pq}} - \sum_{j \in R \setminus \{q\}} \frac{y_{pj}}{y_{pq}} x_j - \frac{1}{y_{pq}} x_{B_p}.$$

The value which corresponds to the basic variable x_{B_i} , $i \neq p$ is obtained if we eliminate x_q in (3) using (4).

~>

$$x_{B_i} = y_{i0} - \frac{y_{iq} y_{p0}}{y_{pq}} - \sum_{j \in R \setminus \{q\}} \left(y_{ij} - \frac{y_{iq} y_{pj}}{y_{pq}} \right) x_j + \frac{y_{iq}}{y_{pq}} x_{B_p}.$$

Let R' denotes the set of the indexes of the nonbasic variables after the exchange of the basic variable x_{B_p} and the nonbasic variable x_q .

Then

$$R' = \{B_p\} \cup R \setminus \{q\}.$$

Let y'_{ij} be the new value for y_{ij} with respect to the equations (3). Then there hold for $0 \leq i \leq m$, $i \neq p$:

$$y'_{i0} = y_{i0} - \frac{y_{iq} y_{p0}}{y_{pq}},$$

$$y'_{ij} = y_{ij} - \frac{y_{iq} y_{pj}}{y_{pq}} \quad \text{for } j \in R' \setminus \{B_p\} \text{ and}$$

$$y'_{iB_p} = - \frac{y_{iq}}{y_{pq}}.$$

The old p -th basic variable x_{B_p} has been replaced by the new p -th basic variable x_q .
Hence, we obtain from equation (4)

$$y'_{p0} = \frac{y_{p0}}{y_{pq}},$$

$$y'_{pj} = \frac{y_{pj}}{y_{pq}} \quad \text{for } j \in R' \setminus \{B_p\} \text{ and}$$

$$y'_{pB_p} = \frac{1}{y_{pq}}.$$

Goal:

Correctness proof for the Simplex method.

We have to show that a basic feasible solution is optimum if the corresponding vertex of the polyhedron has no downhill edge. This means that

$$y_{0j} \leq 0 \quad \text{for all } j \in R.$$

For doing this let B be any basis with basic feasible solution x' and $y_{0j} \leq 0 \quad \forall j \in R$.

(2) and (3) $\Rightarrow$

$$(5) \quad z(x) = c_B^T B^{-1} b - \sum_{j \in R} y_{0j} x_j \quad \forall x \in P.$$

Since $y_{0j} \leq 0$ and $x_j \geq 0 \quad \forall j \in R$, the value $c_B^T B^{-1} b$ is a lower bound for $z(x)$.

Because of

$$x'_B = B^{-1}b \quad \text{and} \quad x'_N = 0$$

we obtain

$$z(x') = c_B^T B^{-1}b.$$

Hence, x' is an optimum solution.

Altogether, we have proved the following theorem.

Theorem 3.5

The basic solution described by the equations (3) is an optimal solution of the given linear program if the following properties are fulfilled:

1. $y_{i0} \geq 0$ for $i=1,2,\dots,m$ (feasibility)
2. $y_{0j} \leq 0$ for all $j \in R$ (optimality)

Now we can present an algorithm for the simplex method. We assume the the computed basic solutions are nondegenerate. Later on we shall extend the algorithm such that the computed basic solution can be degenerate.

We assume that at the beginning, an initial basic feasible solution is known.

Algorithm SIMPLEX

Input: linear program

$$LP: \min z(x) = c^T x$$

$$Ax = b$$

$$x \geq 0.$$

Output: An optimum solution for LP if the optimum is bounded and the information "optimum unbounded" otherwise.

Method:

(1) Initialisation:

Start with a basic feasible solution $(x_B, 0)$.

(2) Optimality test:

If $y_{0j} \leq 0$ for all $j \in R$ then

- the current basic solution $(x_B, 0)$ is an optimum solution.

Output $(x_B, 0)$ and STOP.

(3) Determination of a downhill edge:

- Choose an appropriate variable x_q , $q \in R$ to enter the basis.
- Choose a variable x_{B_p} with

$$\frac{y_{p0}}{y_{pq}} = \min_{1 \leq i \leq m} \left\{ \frac{y_{i0}}{y_{iq}} \mid y_{iq} > 0 \right\}$$

to leave the basis if $\frac{y_{p0}}{y_{pq}}$ is defined.

- If $\frac{y_{p0}}{y_{pq}}$ is not defined; i.e., $y_{iq} \leq 0$ for $1 \leq i \leq m$ then LP is unbounded.

Output "unbounded" and STOP.

(4) Pivot step:

Solve the equations (3) for x_q and x_{B_i} , $i \neq p$ as described above.

$$x_j := 0 \quad \text{for } j \in \{B_p\} \cup R \setminus \{q\}$$

goto (2).

It is possible to implement the simplex method using so-called tables. The following table represent the equations (3).

basic variables

	nonbasic variables			
	...	$-x_j$	...	$-x_q$...
x_{B_0}	y_{00}	y_{0j}		y_{0q}
$\vdots$	$\vdots$	$\vdots$		$\vdots$
x_{B_i}	y_{i0}	y_{ij}		y_{iq}
$\vdots$	$\vdots$	$\vdots$		$\vdots$
x_{B_p}	y_{p0}	y_{pj}		y_{pq}
$\vdots$	$\vdots$	$\vdots$		$\vdots$
x_{B_m}	y_{m0}	y_{mj}		y_{mq}

table of type 1.

The i -th row of the table corresponds to the i -th equation. We say that the table is of Type 1.

Next we shall investigate how to perform the change of the basis using tables of Type 1.

Suppose that the basic variable x_{B_p} has to be replaced by the nonbasic variable x_q .

This can be done as follows:

(1) Divide the p -th row by y_{pq} pivot element

(2) For $0 \leq i \leq m$, $i \neq p$

multiply the new p -th row by y_{iq} and subtract the resulting row from the i -th row.

(3) Divide the old q -th column by y_{pq} and multiply this column by -1 . In the resulting column replace the component in the p -th row by $\frac{1}{y_{pq}}$. Associate with the new q -th column the new nonbasic variable x_{B_p} .



	...	$-x_j$	...	$-x_{B_p}$	...
x_{B_0}	$y_{00} - \left(\frac{y_{0q}y_{p0}}{y_{pq}}\right)$	$y_{0j} - \left(\frac{y_{0q}y_{pj}}{y_{pq}}\right)$		$-\frac{y_{0q}}{y_{pq}}$	
$\vdots$					
x_{B_i}	$y_{i0} - \left(\frac{y_{iq}y_{p0}}{y_{pq}}\right)$	$y_{ij} - \left(\frac{y_{iq}y_{pj}}{y_{pq}}\right)$		$-\frac{y_{iq}}{y_{pq}}$	
$\vdots$					
x_q	$\frac{y_{p0}}{y_{pq}}$	$\frac{y_{pj}}{y_{pq}}$		$\frac{1}{y_{pq}}$	
$\vdots$					

Next we shall investigate how to get the initial feasible basic solution for the given linear program

$$\begin{aligned} \text{LP: } \min \quad z(x) &= C^T x \\ Ax &= b \\ x &\geq 0. \end{aligned}$$

W.l.o.g. we can assume $b \geq 0$. If for an equation

$$a_j^T x = b_j$$

$b_j < 0$ then we can multiply the equation by -1 .

Idea:

In dependence of LP, define a linear program LP' such that:

1. LP' has a trivial feasible basic solution such that we can start the algorithm SIMPLEX with LP' and the trivial feasible basic solution.
2. With help of the solution of LP' it is easy to compute a feasible basic solution of LP.



Consider

$$x = (x_1, x_2, \dots, x_{n+m})^T,$$

$$\tilde{x} = (x_1, x_2, \dots, x_n)^T, \text{ and}$$

$$\bar{x} = (x_{n+1}, x_{n+2}, \dots, x_{n+m})^T$$

and the following linear program:

$$LP': \min z(x) = \sum_{i=n+1}^{n+m} x_i$$

$$A\tilde{x} + \bar{x} = b$$

$$x \geq 0.$$

feasible basic solution of LP' :

$$x_i = 0 \quad \text{for } 1 \leq i \leq n$$

$$x_{n+i} = b_i \quad \text{for } 1 \leq i \leq m.$$

basis: unit matrix I .

The variables $x_{n+1}, x_{n+2}, \dots, x_{n+m}$ are called artificial variables.

Question:

What is the output of SIMPLEX with input

LP' and the feasible basic solution above?

- If LP has no feasible solution then SIMPLEX terminates with a basic feasible solution which has positive values for some artificial variables; i.e. $\bar{x} \neq 0$.
- If LP has a feasible solution then SIMPLEX terminates with a basic feasible solution where all artificial variables are zero; i.e., $\bar{x} = 0$.
- If no artificial variable is a basic variable then we obtain a basic feasible solution of LP by omitting the artificial variables
- Otherwise, the solution is degenerate.

Goal:

Exchange of artificial basic variables by nonbasic variables which are not artificial

or

deletion redundant equations

such that all remaining basic variables are not artificial. Then we obtain a basic feasible solution by omitting the artificial variables.



Suppose that the p th variable of B is artificial. Let e_p denote the p th column of the unit matrix. We distinguish two cases.

Case 1: There is a nonbasic variable x_q , $q \leq n$ with $e_p^T B^{-1} a_q \neq 0$, i.e., $y_{pq} \neq 0$

Apply a pivot step to x_{B_p} and x_q ; i.e., x_{B_p} is replaced by x_q .

$\Rightarrow$

We obtain a basic feasible solution of the same cost but with one artificial variable less.

Case 2: $e_p^T B^{-1} a_j = 0$; i.e. $y_{pj} = 0 \quad \forall$ nonbasic variables x_j , $j \leq n$.

$\Rightarrow$

By applying elementary row operations a zero-row has been constructed from the original matrix

$\Rightarrow$

$Ax = b$ is redundant such that the p -th column of B and the p -th row of A can be deleted. an artificial variable has been deleted from the basis.

Repeat this procedure until no basic variable is artificial.

- The computation of the initial feasible basic solution is called Phase 1 of the simplex method.
- The subsequent computation of an optimal solution is called Phase 2.
- During the development of Phase 1 we have seen that in the case of a redundant system

$$Ax = b$$

of equations one column and one row can be deleted without changing the polyhedron of feasible solutions.

$\Rightarrow$

We know what to do in the case that $\text{rank}(A) < m$.

Exercise:

Develop an algorithm for the solution of linear programs

$$\min z(x) = C^T x$$

$$Ax = b$$

$$x \geq 0$$

with A is an $(m \times n)$ -matrix and $\text{rank}(A) < m$.

It remains the discussion what to do in the case of degenerate basic feasible solutions.

Assumption:

$\exists q \in R, p \in \{1, 2, \dots, m\}$ with $y_{pq} = 0$, $y_{0q} > 0$ and $y_{pq} > 0$.

If the nonbasic variable x_q enters the basis then x_{B_p} or another basic variable with value 0 has to leave the basis. Although we get a new basis, the value of the objective function does not change. The new basic solution is also degenerate (note that $y_{q0} = 0$ after the basis exchange).

Maybe, in the case of degenerate basic solutions, the algorithm SIMPLEX could be in an endless loop. This would be the case if the algorithm repeats a cycle of basic feasible solutions corresponding to the same vertex of the polyhedron infinitely often.

This is called cycling.

If we have no restriction on the choice of the column which leaves the basis there are examples where cycling occurs.

Goal:

The development of selection rules for the column of the basis such that no cycling is possible.

For the development of such selection rules we shall use tables of Type 1. First we need some notations.

A vector $v \neq 0$ is lexicographical positive if its first $\neq 0$ -component is positive. Then we write $v \geq_e 0$. A vector v is called lexico = graphical larger than a vector u if $v - u \geq_e 0$. A sequence $v^1, v^2, \dots, v^s$ of vectors with $v^{i+1} - v^i \geq_e 0$ for $1 \leq i < s$ is called lexicographically increasing. $v \geq_e 0$ means $v = 0$ or $v \geq_e \vec{0}$. Analogously, we define lexicographical smaller and lexicographically decreasing.

To get the selection rules, we extend the table of Type 1 corresponding to the initial feasible basic solution in the following way:

Extended table of
Type 1.

		$-x_j$						
x_{B_0}	y_{00}	0	...	0	...	0	...	y_{0j} ...
x_{B_1}	y_{10}	1	...	0	...	0	...	y_{1j} ...
		$\vdots$		$\vdots$		$\vdots$		$\vdots$
x_{B_i}	y_{i0}	0	...	1	...	0	...	y_{ij} ...
		$\vdots$		$\vdots$		$\vdots$		$\vdots$
x_{B_m}	y_{m0}	0	...	0	...	1	...	y_{mj} ...

- Between the solution column and the columns corresponding to the nonbasic variables, we add the $(m \times m)$ -unit matrix using the rows $1, 2, \dots, m$. Row 0 obtains in the corresponding columns the value 0.

We shall use the additional columns only for the selection of the variables which leave the basis.

Let

$$v_0 := (y_{00}, 0, 0, \dots, 0) \quad \text{and} \\ v_i := (y_{i0}, 0, \dots, 0, 1, 0, \dots, 0), \quad 1 \leq i \leq m$$

be the coefficients of the first $m+1$ columns of the rows $0, 1, 2, \dots, m$ of the extended table.

Note that for $1 \leq i \leq m$

$$y_{i0} \geq 0 \quad \Rightarrow \quad v_i \geq_e 0.$$

Furthermore, these vectors are linear independent.

The selection rule will take care that the algorithm **SIMPLEX** maintains the following invariant:

- The vectors $v_1, v_2, \dots, v_m$ are lexicographical positive and linear independent.

Now we shall investigate the performance of one iteration of the algorithm **SIMPLEX**.

Assumption:

$y_{0q} > 0$ and x_q is chosen to enter the basis.

Let

$$S_q := \{ i \mid i \geq 1, y_{iq} > 0 \} \quad \text{and}$$

$$u_i = \frac{v_i}{y_{iq}} \quad \text{for all } i \in S_q.$$

Let u_p be the lexicographical smallest such a vector; i.e., $u_i \geq_e u_p$ for all $i \in S_q$.

We write then

$$u_p = \text{lexmin} \{ u_i \mid i \in S_q \}.$$

$v_i, 1 \leq i \leq m$ linear independent $\Rightarrow u_p$ is uniquely determined

- Choose the basic variable x_{B_p} to leave the basis.

Since

$$u_p = \text{lexmin} \{ u_i \mid i \in S_q \} \Rightarrow$$

$$\frac{y_{p0}}{y_{pq}} = \min \left\{ \frac{y_{i0}}{y_{iq}} \mid i \in S_q \right\}$$

there holds,

x_{B_p} fulfills the necessary condition to leave the basis.

~~13.12.~~

Let $\hat{v}_i, 1 \leq i \leq m$ be the vectors after the performance of the transformation.

To prove the maintenance of the invariant, we shall prove that

- $\hat{v}_i, 1 \leq i \leq m$ are lexicographical positive
- $\hat{v}_i, 1 \leq i \leq m$ are linear independent.

Note that

$$\hat{v}_p = \frac{v_p}{y_{pq}} \quad \text{and}$$

$$\hat{v}_i = v_i - \frac{y_{iq}}{y_{pq}} v_p = v_i - y_{iq} \cdot \hat{v}_p \quad \text{for } i \neq p.$$

Since $v_p \geq_e 0$ and $y_{pq} > 0$ there hold $\hat{v}_p \geq_e 0$.

- If $y_{iq} \leq 0$ then $-y_{iq} \hat{v}_p \geq_e 0$ and therefore $\hat{v}_i \geq_e 0$.
- If $y_{iq} > 0$ then because of the choice of u_p

$$(*) \quad u_i - u_p = \frac{v_i}{y_{iq}} - \hat{v}_p \geq_e 0.$$

After multiplication of equation (*) with y_{iq} we obtain

$$\hat{v}_i \geq_e 0.$$



$\hat{v}_i, 1 \leq i \leq m$ are lexicographical positive.

Note that

- the addition of a multiple of a vector and another vector maintains the linearly independence of a set of vectors



$\hat{v}_i, 1 \leq i \leq m$ are linear independent.

It remains to show that no cycling occurs.

Since the values in the table are completely determined by the corresponding basis,

until permutation of the nonbasic columns and permutation of the basic rows, the table of the q -th iteration depends only on the current basis B_q .

Hence, for $q < t$ there holds

$$v_0^q \neq v_0^t \Rightarrow B_q \neq B_t.$$

If the sequence $\{v_0^t\}$ is lexicographically decreasing then

$$t_1 \neq t_2 \Rightarrow v_0^{t_1} \neq v_0^{t_2}$$

such that no cycling occurs.

Note that

$$\begin{aligned} v_0^{t+1} &= v_0^t - \frac{y_{0q}^t}{y_{pq}^t} v_p^t \\ &= v_0^t - y_{0q}^t v_p^{t+1} \end{aligned}$$

Since $v_p^{t+1} \geq_e 0$ and $y_{0q}^t > 0$ there holds

$$v_0^{t+1} \leq_e v_0^t.$$

$\Rightarrow$

The sequence $\{v_0^t\}$ is lexicographically decreasing.

3.3 Duality

Duality is a very important and helpful concept in combinatorial optimization. We have used this concept for the development of an efficient algorithm for the maximum weighted matching problem in bipartite graphs. With respect to linear programs, we obtain with help of duality for each linear program a corresponding other linear program which have some interesting properties. Let

$$\begin{aligned} \text{LP: } \min z(x) &= C^T x \\ Ax &= b \\ x &\geq 0 \end{aligned}$$

be a linear program with $\text{rank}(A) = m$ where m is the number of rows of A . Let B be any basis of A and $x = (x_B, x_N)$ be any feasible solution of LP. We have shown above that

$$z(x) = C_B^T B^{-1} b - (C_B^T B^{-1} N - C_N^T) x_N.$$

A sufficient condition for the optimality of the basic solution with respect to the basis B is the following:

- 1) The basic solution which corresponds to the basis B is feasible.

2) $c_B^T B^{-1} N - c_N^T \leq 0.$

We shall investigate the situation that both properties hold at the same time.

Let $\hat{B}$ and $\bar{B}$ be two bases of A such that

$$c_{\hat{B}}^T \hat{B}^{-1} b < c_{\bar{B}}^T \bar{B}^{-1} b.$$

Assume that the basic solution $(x_{\hat{B}}, x_{\hat{N}})$ is feasible for LP; i.e., property 1 holds with respect to the basis $\hat{B}$.

$\Rightarrow$

Property 2 does not hold with respect to the basis $\bar{B}$.

To see this consider the feasible basic solution $x = (x_{\hat{B}}, x_{\hat{N}})$. Then there holds

$$c_{\hat{B}}^T \hat{B}^{-1} b = z(x) = c_{\bar{B}}^T \bar{B}^{-1} b - (c_{\bar{B}}^T \bar{B}^{-1} \bar{N} - c_{\bar{N}}^T) x_{\bar{N}}$$

Since x is a feasible solution for LP there holds

$$x_{\bar{N}} \geq 0$$

Hence

$$\left[c_{\hat{B}}^T \hat{B}^{-1} b < c_{\bar{B}}^T \bar{B}^{-1} b \right.$$

$\Rightarrow$

$$\left. (c_{\bar{B}}^T \bar{B}^{-1} \bar{N} - c_{\bar{N}}^T) \neq 0 \right]$$

Therefore, the basis $\bar{B}$ cannot fulfill Property 2. (144)

Hence, only a basis B which maximizes

$$c_B^T B^{-1} b$$

under fulfilment of

$$c_B^T B^{-1} N \leq c_N^T$$

can solve the linear program LP.

Question:

Can we express this as an optimization problem?

Let

$$\pi^T := c_B^T B^{-1}$$

Then the following linear program defines such an optimization problem:

$$\begin{aligned} \text{LP'}: \quad \max \quad w(\pi) &= \pi^T b \\ \pi^T A &\leq c^T \\ \pi &\geq 0. \end{aligned}$$

$\pi \geq 0$ means that we have no restrictions on the variables in π .

LP is called the primal linear program and LP' is the dual linear program of LP.

Question:

What are the relations of the feasible solutions of a primal linear program and its dual linear program?

(145)

Let x and π be feasible solutions of the primal and the dual linear program. Then

$$(*) \quad c^T x \geq \pi^T A x = \pi^T b.$$

The relation $(*)$ is called weak duality.

Weak duality says that the cost of the primal linear program is always at least as large as the cost of its dual linear program.

Hence the existence of a feasible solution of the primal linear program implies that the dual linear program cannot be unbounded.

Therefore, the existence of an optimal solution for the primal linear program implies the existence of an optimal solution for the dual linear program.

Note that with respect to an optimal basic solution (x_B, x_N) of the primal linear program there holds

$$c_B^T B^{-1} N - c_N^T \leq 0$$

The reverse consideration holds as well.

We wish to prove that the cost of the optimal solutions of the primal and the dual linear programs are always equal.

Because of the weak duality it suffices to construct an explicit dual feasible solution

π_0 such that $\pi_0^T b = c^T x_0$ for an optimal solution x_0 of the primal linear program.

Since the primal linear program has an optimal solution Theorem 3.3 implies the existence of a basis $\hat{B}$ of the matrix A such that the corresponding basic solution $(x_{\hat{B}}, x_{\hat{N}})$ is also optimal; i.e., $z(x_0) = c_{\hat{B}}^T \hat{B}^{-1} b$.

Let $\pi_0 := c_{\hat{B}}^T \hat{B}^{-1}$. Then

$$w(\pi_0) = \pi_0^T b = c_{\hat{B}}^T \hat{B}^{-1} b = c^T x_0 = z(x_0).$$

Altogether, we have proved the following Theorem.

Theorem 3.6

If a linear program has an optimal solution then its dual linear program has also an optimal solution. The costs of both optimal solutions are equal.

An important property of duality is its symmetry. This means that the dual of the dual linear program is the primal linear program again. To prove this property it is useful to consider a linear program in general form and to determine its corresponding dual linear program. For doing this, we transform the general linear program to standard form such that we can

apply the above construction. Consider

$$\begin{aligned}
 \text{LP: } \min \quad z(x) &= c^T x \\
 A_i x &= b_i \quad i \in M \\
 A_i x &\geq b_i \quad i \in \bar{M} \\
 x_j &\geq 0 \quad j \in N \\
 x_j &\leq 0 \quad j \in \bar{N}
 \end{aligned}$$

where A_i denotes the i -th row of the matrix A .

Goal: The transformation of LP to standard form.

We create

- 1) for each inequality $A_i x \geq b_i$, $i \in \bar{M}$
a slack variable x_i^s

and

- 2) for each unrestricted variable x_j , $j \in \bar{N}$
two new nonnegative variables x_j^+ and x_j^-
with $x_j = x_j^+ - x_j^-$ and replace the
column a_j by two columns a_j and $-a_j$.

Then we obtain the following equivalent linear program

$$\begin{aligned}
 \min \quad \hat{z}(\hat{x}) &= \hat{c}^T \hat{x} \\
 \hat{A} \hat{x} &= b \\
 \hat{x} &\geq 0
 \end{aligned}$$

where

$$\hat{A} = \left[a_j, j \in N \mid (a_j, -a_j), j \in \bar{N} \mid \begin{matrix} 0, i \in M \\ -I, i \in \bar{M} \end{matrix} \right]$$

$$\hat{x} = (x_j, j \in N \mid (x_j^+, x_j^-), j \in \bar{N} \mid x_i^s, i \in \bar{M})^T$$

$$\hat{c} = (c_j, j \in N \mid (c_j, -c_j), j \in \bar{N} \mid 0)^T$$

Then we obtain

$$\hat{z}(\hat{x}) = \hat{c}_B^T \hat{B}^{-1} b - (\hat{c}_B^T \hat{B}^{-1} \hat{N} - \hat{c}_N^T) \hat{x}_N$$

Let
$$\pi^T := \hat{c}_B^T \hat{B}^{-1}$$

Then we obtain the following dual linear program:

$$\begin{aligned} \max \quad & w(\pi) = \pi^T b \\ & \pi^T \hat{A} \leq \hat{c}^T \\ & \pi \geq 0. \end{aligned}$$

$\hat{A}$ and A have the same number of rows.

Each inequality in $\pi^T \hat{A} \leq \hat{c}^T$ corresponds to the product of π^T with a column in $\hat{A}$.

The columns in $\hat{A}$ are separated into three sets. This separates the inequalities in $\pi^T \hat{A} \leq \hat{c}^T$ as follows:

1) $\pi^T a_j \leq c_j$ for $j \in N$

2) $\pi^T a_j \leq c_j$ and $-\pi^T a_j \leq -c_j$

$\Leftrightarrow$

$\pi^T a_j = c_j$ for $j \in \bar{N}$ and

$$3. -\pi_i \leq 0 \Leftrightarrow \pi_i \geq 0 \quad \text{for } i \in M$$

Therefore, the dual linear program can be written in the following way:

$$LP': \max w(\pi) = \pi^T b$$

$$\pi^T a_j \leq c_j \quad j \in N$$

$$\pi^T a_j = c_j \quad j \in \bar{N}$$

$$\pi_i \leq 0 \quad i \in M$$

$$\pi_i \geq 0 \quad i \in \bar{M}$$

Let the linear program LP in general form be the primal linear program. Then the dual linear program of LP is the linear program LP'.

primal	dual
$\min z(x) = c^T x$	$\max w(\pi) = \pi^T b$
$Ax_i = b_i \quad i \in M$	$\pi_i \leq 0$
$Ax_i \geq b_i \quad i \in \bar{M}$	$\pi_i \geq 0$
$x_j \geq 0 \quad j \in N$	$\pi^T a_j \leq c_j$
$x_j \leq 0 \quad j \in \bar{N}$	$\pi^T a_j = c_j$

Now we can prove the symmetry property of the duality.

Theorem 3.7

The dual linear program of the dual is the primal linear program again.

Proof:

First we transform the dual linear program such that

- we have a minimizing problem,
- $\geq$ is the only kind of inequalities in it and
- the rows of the matrix are multiplied with the column vector of the variables.

Then we shall use the resulting linear program as input for the transformation described above.

$\leadsto$

$$\max w(\pi) = \pi^T b$$

$$\pi^T a_j \leq c_j \quad j \in N$$

$$\pi^T a_j = c_j \quad j \in \bar{N}$$

$$\pi_i \geq 0 \quad i \in \bar{M}$$

$$\pi_i \leq 0 \quad i \in M$$

$$\min w(\pi) = \pi^T (-b)$$

$$(-a_j^T) \pi \geq -c_j$$

$$(-a_j^T) \pi = -c_j$$

$$\pi_i \geq 0$$

$$\pi_i \geq 0$$

After the application of the construction above, we obtain the following dual linear program.

$$\max z(x) = x^T (-c)$$

$$x_j \geq 0 \quad j \in N$$

$$x_j \geq 0 \quad j \in \bar{N}$$

$$-A_i x \leq -b_i \quad i \in \bar{M}$$

$$-A_i x = -b_i \quad i \in M$$

$$j \in N$$

$$j \in \bar{N}$$

$$i \in \bar{M}$$

$$i \in M$$

This is equivalent to the initial primal linear program

$$\min z(x) = cx^T$$

$$A_i x = b_i \quad i \in M$$

$$A_i x \geq b_i \quad i \in \bar{M}$$

$$x_j \geq 0 \quad j \in N$$

$$x_j \leq 0 \quad j \in \bar{N}$$



For each linear program LP exactly one of the following three cases is fulfilled:

- 1) LP has a finite optimum.
- 2) LP is unbounded.
- 3) LP has no feasible solution.

The following table describes which combinations of the three cases above are possible for a primal and its dual linear program.

primal/dual	finite optimum	unbounded	infeasible
finite optimum	1	x	x
unbounded	x	x	3
infeasible	x	3	2

Theorems 3.6 and 3.7 exclude in the first row and the first column of the table all cases up to combination 1. If the primal or the dual linear program has unbounded cost, the other linear program cannot have a feasible solution.

To show that the combinations 2 and 3 occur, we give simple examples. Consider

$$\begin{aligned} \text{LP:} \quad \min \quad z(x) &= x_1 \\ x_1 + x_2 &\geq 1 \\ -x_1 - x_2 &\geq 1 \\ x_1 &\leq 0 \\ x_2 &\leq 0 \end{aligned}$$

Obviously, LP has no feasible solution. Its dual linear program is

$$\begin{aligned} \text{LP'} \quad \max \quad w(\pi) &= \pi_1 + \pi_2 \\ \pi_1 - \pi_2 &= 1 \\ \pi_1 - \pi_2 &= 0 \\ \pi_1 &\geq 0 \\ \pi_2 &\geq 0 \end{aligned}$$

LP' has also no feasible solution.

$\Rightarrow$

We have an example for combination 2.

If we replace in LP the restrictions of the variables x_1 and x_2 by $x_1 \geq 0$ and $x_2 \geq 0$ then the primal linear program remains to be infeasible but its dual linear program gets to be unbounded.

Altogether, we have proved the following Theorem.

Theorem 3.8

A primal - dual pair of linear programs always fulfills exactly one of the combinations 1-3 as described in the above table.

Let us consider the primal linear program in general form and its dual linear program again.

The stronger restrictions of the primal linear program (dual linear program) correspond to the sets M and N ($\bar{M}$ and $\bar{N}$) of indices.

A necessary and sufficient condition of a pair x, π of feasible solutions for a primal-dual pair is the so-called complementary slackness condition.

Theorem 3.9 (complementary slackness)

Let x, π be feasible solutions for a primal-dual pair of linear programs. Then x and π are optimal iff

- 1) $u_i = \pi_i (A_i x - b_i) = 0$ for all $i \in M \cup \bar{M}$ and
- 2) $v_j = (c_j - \pi^T a_j) x_j = 0$ for all $j \in N \cup \bar{N}$.

Proof:

The restrictions of both linear programs imply

$$\begin{aligned} u_i &\geq 0 && \text{for all } i \in M \cup \bar{M} \quad \text{and} \\ v_j &\geq 0 && \text{for all } j \in N \cup \bar{N}. \end{aligned}$$

Let

$$u := \sum_{i \in M \cup \bar{M}} u_i \quad \text{and} \quad v := \sum_{j \in N \cup \bar{N}} v_j.$$

Then $u \geq 0$ and $v \geq 0$.

Furthermore

$$\begin{aligned} u = 0 &\Leftrightarrow \pi_i (A_i x - b_i) = 0 \quad \text{for all } i \in M \cup \bar{M} \\ v = 0 &\Leftrightarrow (c_j - \pi^T a_j) x_j = 0 \quad \text{for all } j \in N \cup \bar{N}. \end{aligned}$$

Consider

$$\begin{aligned} u + v &= \sum_{i \in M \cup \bar{M}} \pi_i (A_i x - b_i) + \sum_{j \in N \cup \bar{N}} (c_j - \pi^T a_j) x_j \\ &= c^T x - \pi^T b. \end{aligned}$$

$\Rightarrow$

$$[(u_i = 0 \quad \forall i \in M \cup \bar{M}) \text{ and } (v_j = 0 \quad \forall j \in N \cup \bar{N})]$$

$\Leftrightarrow$

$$u + v = 0$$

$\Leftrightarrow$

$$c^T x = \pi^T b.$$

Since for feasible solutions x and π always

$$c^T x \geq \pi^T b$$

Theorem 3.6 imply the assertion.



We have developed the Simplex algorithm for a given primal linear program. This algorithm is often called primal simplex algorithm.

The primal simplex algorithm solves the linear program by going from a feasible basic solution of the primal linear program (or shortly a primal feasible basis) to another neighbored feasible basic solution.

Goal:

The development of a simplex algorithm which solves the linear program by going from a dual feasible basic solution to another neighbored dual feasible basis.

Such an algorithm is called dual simplex algorithm.

Let us consider a primal-dual pair where the primal linear program is in standard form again.
Let

$$z_0 := \min \{ c^T x \mid Ax = b, x \geq 0 \} \quad \text{and} \\ w_0 := \max \{ \pi^T b \mid \pi^T A \leq c^T, \pi \leq 0 \}.$$

be the cost of an optimal solution of the primal and dual linear program, respectively.

Let B be any basis of A .

The corresponding basic solution (x_B, x_N) with $x_B = B^{-1}b$ and x_N is feasible iff $B^{-1}b \geq 0$.

Then (x_B, x_N) is a primal feasible basic solution.

$\pi^T = c_B^T B^{-1}$ is the solution corresponding to the dual linear program with respect to the basis B .

We shall investigate the situations when $\pi^T = c_B^T B^{-1}$ is feasible for the dual linear program.

Consider $\pi^T A - c^T$. Then we obtain

$$\begin{aligned}\pi^T A - c^T &= c_B^T B^{-1} (B, N) - (c_B, c_N)^T \\ &= (0, c_B^T B^{-1} N - c_N^T)^T\end{aligned}$$

Therefore, $\pi^T = c_B^T B^{-1}$ is feasible for the dual linear program iff

$$c_B^T B^{-1} N - c_N^T \leq 0.$$

A basis B of the matrix A is called dual feasible if $c_B^T B^{-1} N \leq c_N^T$.

We have seen that a basis B defines the point

$$x = (x_B, x_N) = (B^{-1}b, 0) \in \mathbb{R}^n$$

and the point

$$\pi^T = c_B^T B^{-1} \in \mathbb{R}^m.$$

B can be primal feasible, dual feasible, both and none of both.

Lemma 3.2

If a basis B of A is primal and dual feasible then $x = (x_B, x_N) = (B^{-1}b, 0)$ and $\pi^T = c_B^T B^{-1}$ are optimal solutions of the primal and the dual linear programs.

Proof:

$$\text{Since } c^T x = c_B^T x_B = c_B^T B^{-1}b \quad \text{and} \\ \pi^T b = c_B^T B^{-1}b$$

and hence

$$c^T x = \pi^T b,$$

the assertion follows. □

~>

Interpretation: The

- primal simplex algorithm starts with a primal feasible basis and tries to get dual feasibility while maintaining primal feasibility.
- dual simplex algorithm starts with a dual feasible basis and tries to get primal feasibility while maintaining dual feasibility.

We could solve the dual linear program in the following way:

- (1) Transform the linear program to standard form such that the primal simplex can be applied.
- (2) Use the primal simplex algorithm to solve the dual linear program.

Instead of doing this, we shall develop a direct algorithm for the solution of the dual linear program.

Exercise:

Given the dual of a primal linear program. transform the dual linear program such that the primal simplex algorithm could be applied.

Let B be a dual feasible basis of A . Since the dual linear program is feasible, the following combinations for the primal - dual pair of linear programs are possible:

Primal/dual	finite optimum	unbounded
finite optimum	1	x
infeasible	x	3

$\Rightarrow$

The dual simplex algorithm can terminate with the following results:

- 1) A finite optimum is found.
- 2) It is established that the primal linear program is infeasible.

Hence it is useful to develop a sufficient criterion of the infeasibility of the primal linear program.

To get such a criterion let

$$x = (x_B, x_N)$$

be a feasible ^{basic} solution of the primal linear program. Then

$$Bx_B + Nx_N = b \quad \text{implies}$$

$$(*) \quad x_B + B^{-1}Nx_N = B^{-1}b.$$

Let R be the set of indices of the nonbasic variables. Then we can write (*) as follows.

$$x_B + \sum_{j \in R} B^{-1}a_j x_j = B^{-1}b.$$

Let

$$B^{-1}b = (y_{10}, y_{20}, \dots, y_{m0})^T.$$

5. If the basis B is also primal feasible; i.e., $y_{i0} \geq 0$ for all $1 \leq i \leq m$, ^{then} Lemma 3.2 implies that

$(B^{-1}b, 0)$ and $c_B^T B^{-1}$ are optimal solutions for the primal and dual

linear program

Assume that the basis B is not primal feasible and let $B'_1, B'_2, \dots, B'_m$ be the rows of B^{-1} . Then $\exists p \in \{1, 2, \dots, m\}$ with $y_{p0} < 0$. The following lemma gives us the desired criterion.

Lemma 3.3

Let B be a dual feasible basis and let $B^{-1}b = (y_{10}, y_{20}, \dots, y_{m0})^T$. Let $y_{p0} < 0$. If $B'_p a_j \geq 0$ for all $j \in R$ then the primal linear program is infeasible.

Proof:

Note that

$$y_{p0} = x_{B_p} + \sum_{j \in R} B'_p a_j x_j.$$

By definition, for each feasible primal solution there hold $x_j \geq 0$ for all $j \in R$.

Since

$$y_{p0} < 0 \text{ and } B'_p a_j \geq 0 \text{ for all } j \in R$$

there has to be

$$x_{B_p} < 0.$$

$\Rightarrow$

The primal linear program has to be infeasible.

If the criterion of Lemma 3.3 is not fulfilled then there is $q \in R$ with $B_q' a_q < 0$.

Idea

Determine such a q and perform the basis exchange

$$a_{B_p} \leftrightarrow a_q.$$

We have to take care that the dual feasibility is not destroyed by this exchange.

$\leadsto$

For $j \in R$ there holds $y_{0j} = -\bar{c}_j$.

Dual feasibility of $B \Rightarrow$

$$y_{0j} \leq 0 \quad \text{for all } j \in R.$$

After the basis exchange $a_{B_p} \leftrightarrow a_q$, we obtain the following value y'_{0j} for $-\bar{c}_j$:

$$y'_{0j} = y_{0j} - \frac{y_{0q}}{y_{pq}} y_{pj}$$

Since the new basis B' has to be dual feasible, we have to take care that for all $j \in R' = \{B_p\} \cup R \setminus \{q\}$

$$y'_{0j} \leq 0$$

$$\Leftrightarrow y_{0j} - \frac{y_{0q}}{y_{pq}} y_{pj} \leq 0. \quad (+)$$

We distinguish three cases.

a) $y_{pj} = 0$:

Then $y_{oj} \leq 0$ implies that the inequality (+) is fulfilled.

b) $y_{pj} > 0$:

Since $y_{pq} = B_p^T a_q < 0$ and $y_{oq} \leq 0$ the inequality (+) is fulfilled as well.

c) $y_{pj} < 0$:

Then we have to take care. We obtain

$$y_{oj} - \frac{y_{oq}}{y_{pq}} y_{pj} \leq 0$$

$$\Leftrightarrow \frac{y_{oj}}{y_{pj}} \leq \frac{y_{oq}}{y_{pq}}$$

If we choose q such that

$$\frac{y_{oq}}{y_{pq}} = \max \left\{ \frac{y_{oj}}{y_{pj}} \mid y_{pj} < 0 \right\}$$

then the inequality (+) will be always fulfilled

Exercise:

Work out in detail the dual simplex algorithm.
Could you use the same tables as the primal simplex algorithm?