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 <u>feasible solutions</u>

and

- an <u>objective function</u> $z : F \to \mathbb{R}$

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 $b, c_1, c_2, \ldots, c_n \in \mathbb{R}$.

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

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

is called a linear function.

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

- $f(x_1, x_2, \ldots, x_n) \leq b$ and $f(x_1, x_2, \ldots, x_n) \geq b$ 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 $V = \{1, 2, \ldots, n\}$ 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

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

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

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:

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

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

$$x \geq 0$$

$$x \text{ integral}$$

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) <u>The maximum flow problem</u>

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

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

$\rightsquigarrow$

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

$$k(i) = 0 \qquad 2 \le i \le n-1$$
$$x_{ij} \le c_{ij} \qquad \forall (i,j) \in E$$
$$x_{ij} \ge 0 \qquad \forall (i,j) \in E$$

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 $V = \{1, 2, \ldots, n\}$. 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.

$\rightsquigarrow$

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

$$\begin{aligned} x_i + x_j &\geq 1 && \forall (i,j) \in E \\ x_i &\geq 0 && 1 \leq i \leq n \\ x_i &\text{ integral} && 1 \leq i \leq n. \end{aligned}$$

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 fullfies 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 $x_j \lesseqgtr 0$ can be replaced by two nonnegative variables x_j' and x_j''. For doing this, we add the restrictions

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

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 <u>slack variable</u> 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

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

is a <u>convex combination</u> of x and y.

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

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

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

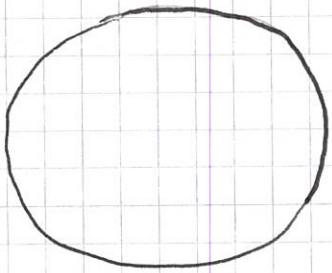
is a <u>convex combination</u> of the points $x_1, x_2, \dots, x_t$.

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

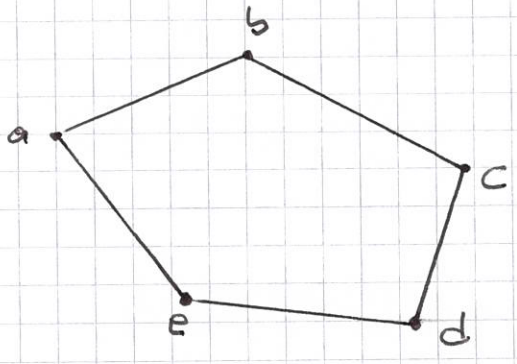
Example:

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

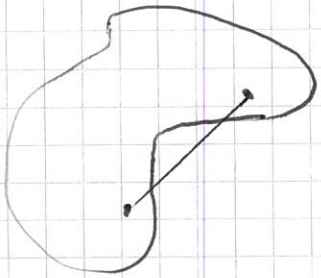
$C \subseteq \mathbb{R}^n$ is convex iff for any points $x, y \in C$, all points on the line segment which connects x and y are in C.



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 <u>hyperplane</u>. The sets

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

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

are <u>closed half-spaces</u>.

The hyperplane H associated with the half-space $\overline{H}$ is the <u>bounding hyperplane</u> of that half-space.

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

<u>Exercise</u>

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

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

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

inequalities

$$a_1^T x \leq b_1, \; a_2^T x \leq b_2, \; \dots, \; a_m^T x \leq b_m \qquad (*)$$

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 <u>dimension</u> of S. and is denoted by <u>dim(S)</u>. Note that

$$\dim(S) = n - \operatorname{rank}[A].$$

An <u>affine subspace</u> 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.,

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

The dimension $\dim(S_a)$ of S_a is the same as the dimension of S. Equivalently, a 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 a $b \in \mathbb{R}^m$.

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 $C \subseteq H$, 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 characterizes 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

$$x = \begin{pmatrix} \bar{x} \\ 0 \end{pmatrix}, \quad \bar{x} > 0$$ and let $\bar{A}$ 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 <u>not</u> 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 $\operatorname{rank}(A) = m$ then we obtain an equivalent characterization of the vertices of P which leads to an answer to the question posed above.

$$\operatorname{rank}(A) = m \implies m \leq n.$$

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 $d \not\geq 0$.

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.

$\Rightarrow$

$x_0 + \lambda' d \not\geq 0$ 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

$$Ad = 0.$$

$\Leftarrow$ a

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 $z \to -\infty$.

Assume that for a given linear program

$$\min \; z(x) = c^T x$$
$$Ax = b$$
$$x \geq 0$$

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 optim

3.2 The simplex method

Goal:
Development of a method to solve the linear program

$$\min \; z(x) = c^T x$$
$$Ax = b$$
$$x \geq 0$$

where A is an $(m \times n)$-matrix of row rank m. The case $\operatorname{rank}(A) < m$ 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

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

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

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) $$z(x) = c_B^T B^{-1} b - \left(c_B^T B^{-1} N - c_N^T \right) x_N$$

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

Neighboured vertices of the polyhedron correspond to neighboured bases.

$\Rightarrow$

Going from one vertex to a neighboured 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 neighboured 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 <u>reduced cost</u> 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).

$$\rightsquigarrow$$

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

Let R be the set of the indizes 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

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

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

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

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

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.

$\rightsquigarrow$

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:

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

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

Let x_{B_p} be any basic variable with

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

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

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

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

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

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

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).

$\rightsquigarrow$

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

Let R' denotes the set of the indizes 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 $0 \leq i \leq m$, $i \neq p$:

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

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

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)

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

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

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

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

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

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

(2) and (3) $\Rightarrow$

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

Since $y_{Oj} \leq 0$ and $x_j \geq 0 \ \forall j \in R$, 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 $i = 1, 2, \ldots, m$ (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

$$LP: \quad \min z(x) = c^T x$$
$$Ax = b$$
$$x \geq 0.$$

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

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

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

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

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

goto (2).

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

	nonbasic variables			
	$\cdots$	$-x_j$	$\cdots$	$-x_q$ $\cdots$
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}

basic variables

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 $y_{pq} \neq 0$ 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}.

	$\cdots$	$-x_j$	$\cdots$	$-x_{B_p}$	$\cdots$
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

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

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:

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

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

$$x \geq 0.$$

<u>feasible basic solution of LP'</u>:

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

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

basis: unit matrix I.

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

<u>Question:</u>

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 pth variable of B is artificial. Let e_p denote the pth 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. $y_{pj} = 0$ $\forall$ 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.