By a cone, we normally mean a cone which is convex as well. And it is not difficult to prove the following proposition.

Proposition 3.2

A set $K \subseteq \mathbb{R}^n$ is a convex cone if it is closed in nonnegative linear combination or conic combination: $\forall\ x$, $y \in K$ and $\forall \lambda$, $\mu \geq 0$, $\lambda x + \mu y \in K$.

The conic combination also can be extended to a finite number of vectors. Similarly we can define conic hull of a set S to be the smallest cone including S as a subset. Also, then we can show the conic hull of S is the set of conic combinations of vectors from S.

Definition 3.3

If a convex cone K is a conic hull of a finite set of vectors $\{y^1, \ldots, y^k\}$

$$K = \text{cone}\{y^1, \ldots, y^k\} = \{\lambda_1 y^1 + \cdots + \lambda_k y^k : \lambda_1 \geq 0, \ldots, \lambda_k \geq 0\},$$

then K is said to be a *finitely generated cone.*

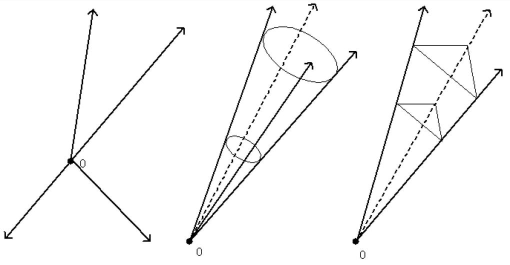


Figure: Cone, convex cone and finitely generated cone.

Definition 3.4

For any matrix $A \in \mathbb{R}^{m \times n}$, the polyhedron $K = \{y : Ay \geq 0\}$ is a convex cone by Proposition 3.2. We call K a polyhedral cone.

Proposition 3.5

Minkowski's theorem: *Every polyhedral cone is finitely generated.*

Proof: Any vector y of $K = \{y : Ay \geq 0\}$ can be scaled down to be contained in a unit hypercube centered at the origin. Hence every $y \in K$ is a positive multiplication of a vector from $\bar{K} = \{y : Ay \geq 0, \; -e \leq y \leq e\}$. Since $\bar{K}$ is bounded, it is the convex hull of its finite number of vertices. It implies $K = \{y : Ay \geq 0\}$ is the conic hull of the vertices of $\bar{K}$, and hence finitely generated. $\square$

If a polyhedron $P = \{x : Ax \geq b\}$ includes a half line $x + \lambda y$, then we should have $Ay \geq 0$. Conversely, if $Ay \geq 0$, then for any $x \in P$, the half line $x + \lambda y$ is included in P. It suggests the following proposition.

Proposition 3.6

Every polyhedron $P = \{x : Ax \geq b\}$ is the sum of a bounded polyhedron and a polyhedral cone. In other words, there is a finite set of vectors $\{x^1$, x^2, ..., $x^p\}$ and $\{y^1$, y^2, ..., $y^q\}$ such that for any $x \in P$, there are λ and μ such that

$$\begin{gathered} x = \lambda_1 x^1 + \lambda_2 x^2 + \cdots + \lambda_p x^p + \mu_1 y^1 + \mu_2 y^2 + \cdots + \mu_q y^q, \\ \lambda_1 + \lambda_2 + \cdots + \lambda_p = 1,\ \lambda_1 \geq 0,\ \lambda_2 \geq 0,\ \cdots, \lambda_p \geq 0, \\ \mu_1 \geq 0,\ \mu_2 \geq 0,\ \cdots, \mu_q \geq 0. \end{gathered} \tag{3.2}$$

Linear Programs and Simplex Method

Optimization Lab.

IE department
Seoul National University

9th May 2018

A linear program is an optimization problem of minimizing a real-valued linear function over a polyhedron:

$$
\begin{array}{llccccccl}
\text{min/max} & z= & c_1x_1 & +c_2x_2 & \cdots & +c_nx_n & & & \text{Objective} \\
\text{sub.to} & & a_{11}x_1 & +a_{12}x_2 & \cdots & +a_{1n}x_n & =,\leq,\geq & b_1 & \text{Constraints} \\
& & a_{21}x_1 & +a_{22}x_2 & \cdots & +a_{2n}x_n & =,\leq,\geq & b_2 & \\
& & & & \vdots & \vdots & & & \\
& & a_{m1}x_1 & +a_{m2}x_2 & \cdots & +a_{mn}x_n & =,\leq,\geq & b_m & \\
& & & & & & & & \text{Nonnegativity} \\
& & x_1 \geq 0 & x_2 \geq 0 & \cdots & x_n \geq 0, & & & \text{Restrictions}
\end{array}
$$

Assumption 1.1

1. We assume $b_i \geq 0$ for each i (or we can multiply the constraint by -1).

2. We do not lose generality by the sign restriction since any real variable can be represented as the difference of two nonnegative variables.

Also notice that any LP can be transformed into $\min\{c^T x | Ax \geq b\}$.

By introducing an additional nonnegative variable, each inequality can be transformed into an equality constraint.

Example 1.2

$$\begin{array}{llrrrll} \min & z = & 3x_1 & -x_2 & +2x_3 & & \\ \text{sub.to} & & -x_1 & +5x_2 & +2x_3 & = & 5 \\ & & 2x_1 & -2x_2 & -x_3 & \leq & 3 \\ & & x_1 & & +2x_3 & \geq & 1 \\ & & x_1 \geq 0 & x_2 \geq 0 & x_3 \geq 0. & & \end{array}$$

The second constraint is equivalent to $2x_1 - 2x_2 - x_3 + x_4 = 3$, $x_4 \geq 0$. I.e. a $\leq$-inequality constraint amounts to an equality constraint and a nonnegativity restriction. Similarly, using a nonnegative variable, say x_5, we can transform the third constraint into $x_1 + 2x_3 - x_5 = 1$, $x_5 \geq 0$.

Example 1.3

And we get the following standard linear program.

$$
\begin{array}{lllllllll}
\min & z = & 3x_1 & -x_2 & +2x_3 & & & & \\
\text{sub.to} & & -x_1 & +5x_2 & +2x_3 & & & = & 5 \\
& & 2x_1 & -2x_2 & -x_3 & +x_4 & & = & 3 \\
& & x_1 & & +2x_3 & & -x_5 & = & 1 \\
& & x_1 \geq 0 & x_2 \geq 0 & x_3 \geq 0 & x_4 \geq 0 & x_5 \geq 0. & &
\end{array}
$$

Remark 1.4

Although of no importance, x_4 is called a slack variable, and x_5 a surplus variable.

Conventionally, LP algorithms assume an LP is given in standard form.

Problem 1.5

$$\begin{aligned} \min \quad & c^T x & \\ \text{sub to} \quad & Ax &= b \\ & x &\geq 0 \end{aligned} \tag{Standard LP}$$

As usual $A \in \mathbb{R}^{m \times n}$ is assumed to have a full row rank m (hence the equality system has a solution). If a solution also satisfies nonnegativity restriction, it is called feasible. If LP has no feasible solution, it is said to be infeasible.

$$x_1 + x_2 = -1$$
$$x_1 \geq 0 \; x_2 \geq 0$$

Relying on the primal-dual pair ((3.7) and (3.8) in the previous chapter) and their weak and strong duality, we can derive the standard dual LP,

Problem 1.6

$$\begin{array}{rll} \max & b^T y & \\ \text{sub to} & A^T y & \leq c \end{array} \qquad \text{(Standard dual linear program)}$$

and the duality theorems on standard linear programs.

Theorem 1.7

(Weak duality) Every feasible pair (x, y) of (1.5) and (1.6), satisfies $c^T x \geq y^T b$.

Theorem 1.8

(Strong duality) If either (1.5) or (1.6) has an optimal solution, then so does the other problem and their objective values are the same.

Definition 2.1

If a feasible solution x of (1.5) is a basic solution of $Ax = b$ as well, it is said to be a *basic feasible solution* (BFS).

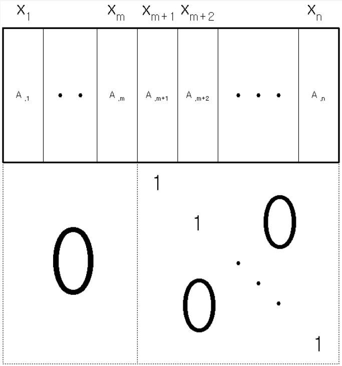
Let B be the basis of a BFS x and x_B the the sub-vector of basic variables (which we assume are ordered to the columns of B). So we have $x_B = B^{-1}b \geq 0$. Unlike a basic solution which is guaranteed by Gauss-Jordan elimination, BFS due to nonnegativity restriction requires a more elaborated algorithm.

We now see the basic feasible solutions are exactly the vertices of the polyhedron.

Theorem 2.2

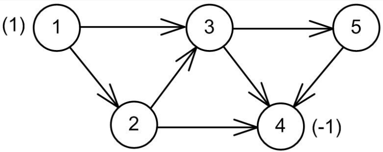
Of a standard LP, its BFSs and vertices are the same thing.

Proof: Let $\bar{x}$ be a vertex. Since its face subsystem has rank is n, it includes, besides $Ax = b$ (whose rank is assumed to be m), $n - m$ nonnegativity restrictions which are, we may assume, the last $n - m$ ones. Then the corresponding $n - m$ variables should be all 0. And $B = [A_{\cdot 1}, \ldots, A_{\cdot m}]$ is a basis of the column space. Since $\bar{x}_B \equiv (\bar{x}_1, \ldots, \bar{x}_m)^T = B^{-1}b$, $\bar{x}$ is basic. Since $\bar{x}$ is feasible. it is a BFS.



Conversely, suppose $\bar{x}$ is a BFS. We can reorder, if necessary, the columns of its basis and add $n - m$ nonnegative restrictions from nonbasic variables to get the subsystem of active inequalities of $\bar{x}$ as in the figure. Its rank is clearly n and hence $\bar{x}$ is a vertex. $\square$

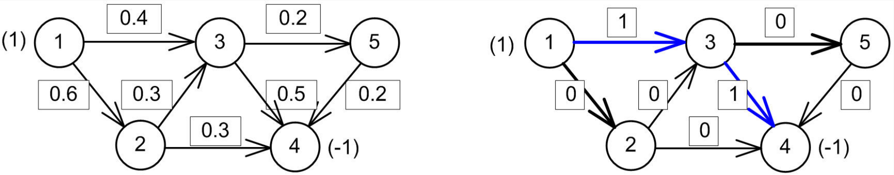
Consider the following "pipe" network on which the total throughput of 1 flows from node 1 to node 4.



Let x_{ij} be the flow on arc (i, j). Then x is a feasible solution of the following standard LP. If we remove the last row which is redundant, the equality system has full-row rank, $n - 1$.

$$
\begin{array}{rrrrrrrl}
x_{12} & +x_{13} & & & & & & = 1 \\
-x_{12} & & +x_{23} & +x_{23} & & & & = 0 \\
& -x_{13} & -x_{23} & & +x_{34} & +x_{35} & & = 0 \\
& & & -x_{24} & -x_{34} & & -x_{54} & = 1 \\
\hline
& & & & & -x_{35} & +x_{54} & = 0 \\
& & & x_{ij} & \geq 0, & \forall i, j & &
\end{array}
\tag{2.1}
$$

The figure indicates two feasible network flows. The left one is not a BFS (why?) whereas the right one is a BFS of (2.1).



$$A = \begin{array}{c} \begin{matrix} x_{12} & x_{13} & x_{23} & x_{24} & x_{34} & x_{35} & x_{54} \end{matrix} \\ \left[\begin{array}{rrrrrrr} \mathbf{+1} & \mathbf{+1} & 0 & 0 & \mathbf{0} & \mathbf{0} & 0 \\ \mathbf{-1} & \mathbf{0} & +1 & +1 & \mathbf{0} & \mathbf{0} & 0 \\ \mathbf{0} & \mathbf{-1} & -1 & 0 & \mathbf{+1} & \mathbf{+1} & 0 \\ \mathbf{0} & \mathbf{0} & 0 & -1 & \mathbf{-1} & \mathbf{0} & -1 \\ \hline 0 & 0 & 0 & 0 & 0 & -1 & +1 \end{array}\right] \end{array} \tag{2.2}$$

Its basis consists of A_{12}, A_{13}, A_{34}, A_{35}.

Recall that the optimal solutions constitutes a face F, the intersection of the polyhedron and the supporting hyperplane determined by the objective coefficient vector. Since a standard system has rank n, F should contain a vertex.

Theorem 3.1

An optimal solution of a standard LP is attained at a vertex.

The simplex method searches an optimal BFS by moving from a vertex to an adjacent vertex of a smaller objective value.

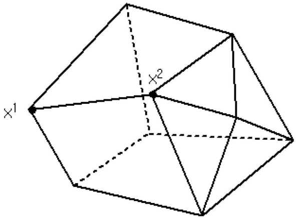
Definition 3.2

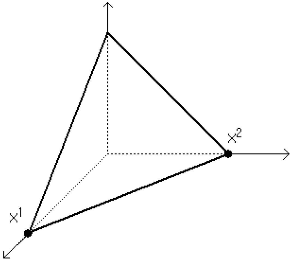
We call two BFSs are *adjacent* if they are adjacent as basic solutions, namely if the bases of the BFSs have exactly $m - 1$ columns in common.

Definition 3.3

It two vertices are on the same one-dimensional face, they are said to be *geometrically adjacent* (이웃 하다).

A one-dimensional face is called an *edge*. Thus two vertices are adjacent iff they are the endpoints of an edge.





Theorem 3.4

Two BFSs are adjacent if and only if geometrically adjacent.

Proof: ($\Rightarrow$) Suppose two BFSs, u and v, have the bases B and B' resp. with exactly $m-1$ columns in common. Then they have exactly $n-m-1$ common nonbasic variables which, along with $Ax=b$, is a face subsystem of a face containing both u and v. Since the rank is $n-1$, the face has a dimension at most 1. Since it contains two distinct points u and v, it is one-dimensional. Therefore u and v are geometrically adjacent.

($\Leftarrow$) Let x° and $x^{\circ\circ}$ be the two end points of an edge. Since the maximum face subsystem of the edge $[x^{\circ}, x^{\circ\circ}]$ has the rank $n-1$, it includes $n-m-1$ nonnegativity restrictions $x_{i_1} \geq 0$, $x_{i_2} \geq 0$, …, $x_{i_{n-m-1}} \geq 0$ which increase the rank of $Ax=b$ by $n-1$.

$$\text{Edge } [x^{\circ}, x^{\circ\circ}] = \left\{x \in P : Ax = b,\ x_{i_1} = \cdots = x_{i_{n-m-1}} = 0\right\}. \qquad (3.3)$$

Proof(*cont'd*): The line from x° to $x^{\circ\circ}$ satisfies the subsystem of (3.3). Hence there should a blocking inequality $x_r \geq 0$ not from (3.3) which is satisfied by equality by $x^{\circ\circ}$ and increases the rank of subsystem into n.

$$\{x^{\circ\circ}\} \;=\; \left\{x : Ax = b, x_r = 0, x_{i_1} = \cdots = x_{i_{n-m-1}} = 0\right\}. \tag{3.4}$$

Therefore, the columns of A corresponding to the remaining variables constitutes a basis of the BFS $x^{\circ\circ}$. (See Figure 2.2.)

Similarly, there is s such that $\{x^\circ\} = \{x : Ax = b,\ x_s = 0,\ x_{i_1} = \cdots = x_{i_{n-m-1}} = 0\}$. We have $r \neq s$. (Why?) Thus x° and $x^{\circ\circ}$ have exactly $m-1$ columns in common in their bases and thus algebraically adjacent.
□

From the proof, x° and $x^{\circ\circ}$ have the basic variable sets $\{1, \ldots, n\} \setminus \{s, i_1, \ldots, i_{n-m-1}\}$ and $\{1, \ldots, n\} \setminus \{r, i_1, \ldots, i_{n-m-1}\}$, respectively. If we drop x_r and enter x_s into the basic variables of x°, we get the adjacent BFS $x^{\circ\circ}$.

Exercise 3.5

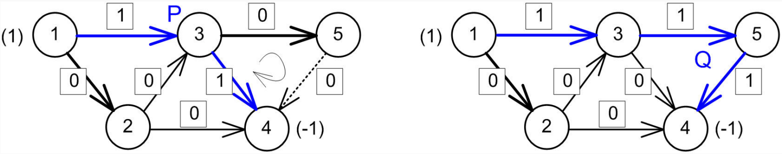
Consider the following standard linear system and the basic feasible solution x^1 *:* $[4,1,0,0,0]^T$*.*

$$\begin{array}{cccccc} x_1 & -2x_2 & +x_3 & -x_4 & +3x_5 & =2 \\ & +x_2 & +2x_3 & +x_4 & -x_5 & =1 \\ x_1, & x_2, & x_3, & x_4, & x_5 & \geq 0 \end{array}$$

(1) Find the basis of x^1*. Compute the adjacent BFS* x^2 *obtained by entering* x_3 *and dropping* x_2 *from the basis.*
(2) Are x^1 *and* x^2 *BFSs? Why?*
(3) Are x^1 *and* x^2 *adjacent? If so, identify the edge.*

Exercise 3.6

Consider the network flow example and the BFS.



We can observe a similar geometrical adjacency of general polyhedron.

Exercise 3.7

$$\begin{array}{cccl} x_1 & & -x_3 & \leq 1 \\ & & x_3 & \leq 4 \\ x_1 & +x_2 & +x_3 & \geq 6 \\ x_1 & & & \geq 0 \\ & x_2 & & \geq 0 \\ & & x_3 & \geq 0 \end{array}$$

(1) Check $x^\circ = (1,5,0)^T$ *is a vertex of the polyhedron. Relax the last inequality of the maximum face subsystem of* x° *to get the face subsystem of an edge.*
(2) Using the edge from (1), find an adjacent vertex x^1 *to* $(1,5,0)^T$*.*
(3) For the objective function $3x_1 + 2x_2 - x_3$*, is* x^1 *-* x° *an improving direction? Why?*

Let $\bar{x}$ be a basic solution with basis $B = [A_{\cdot 1} \ \cdots \ A_{\cdot m}]$. Suppose $\hat{x}$ is an adjacent basic solution obtained by entering x_s and dropping x_r. Recall, from page 64, Chapter 1, any point on the line segment $[\bar{x}, \hat{x}]$ is

$$\bar{x}(\delta) := \bar{x} + \delta \begin{array}{r} \\ \\ \\ s \text{ 번째} \rightarrow \\ \\ \\ \end{array} \left[\begin{array}{c} -B^{-1}A_{\cdot s} \\ \hline 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{array}\right] = \left[\begin{array}{c} B^{-1}b \\ \hline 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{array}\right] + \delta \left[\begin{array}{c} -B^{-1}A_{\cdot s} \\ \hline 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{array}\right], \tag{4.5}$$

for $0 \leq \delta \leq \Delta$, where $\Delta = (B^{-1}b)_r / (B^{-1}A_{\cdot s})_r$. Also we have $\hat{x} = \bar{x}(\Delta)$.