Linear Algebra

Lecture 12 - Linear Programming

Oskar Kędzierski

13 January 2025

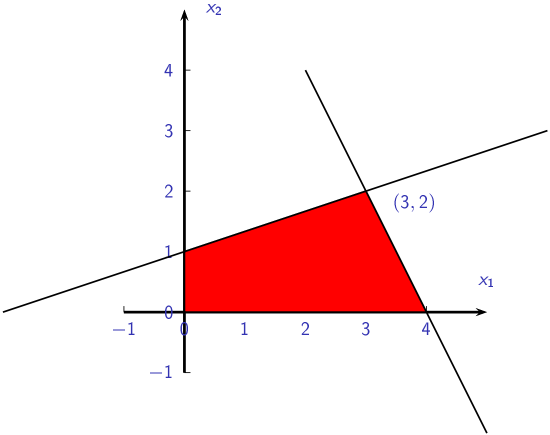
What is Linear Programming?

Example

Maximize the value $x_1 + 2x_2$ under the constraints

$$\begin{cases} x_1 & \geqslant & 0 & & \\ x_2 & \geqslant & 0 & & \\ 2x_1 & + & x_2 & \leqslant & 8 \\ -x_1 & + & 3x_2 & \leqslant & 3 \end{cases}$$

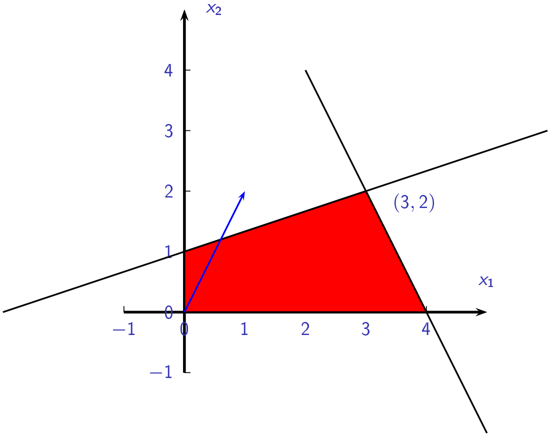
What is Linear Programming?



maximize $x_1 + 2x_2$

$$\begin{cases} x_1 & \geqslant & 0 & & \\ x_2 & \geqslant & 0 & & \\ 2x_1 & + & x_2 & \leqslant & 8 \\ -x_1 & + & 3x_2 & \leqslant & 3 \end{cases}$$

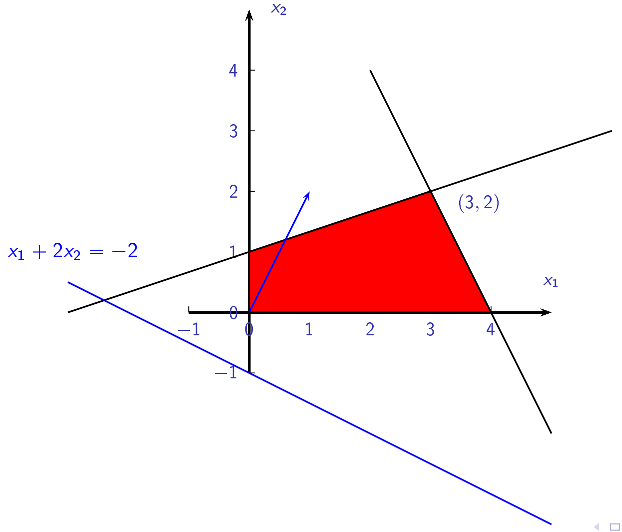
What is Linear Programming?



maximize $x_1 + 2x_2$

$$\begin{cases} x_1 & \geqslant & 0 & & \\ x_2 & \geqslant & 0 & & \\ 2x_1 & + & x_2 & \leqslant & 8 \\ -x_1 & + & 3x_2 & \leqslant & 3 \end{cases}$$

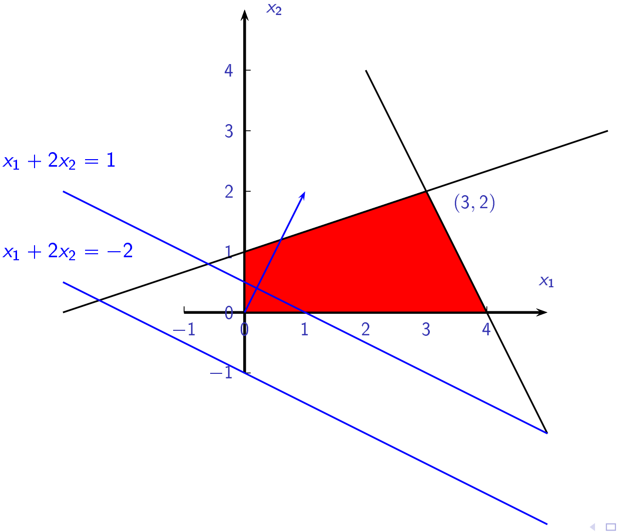
What is Linear Programming?



maximize $x_1 + 2x_2$

$$\begin{cases} x_1 & \geqslant & 0 & & \\ x_2 & \geqslant & 0 & & \\ 2x_1 & + & x_2 & \leqslant & 8 \\ -x_1 & + & 3x_2 & \leqslant & 3 \end{cases}$$

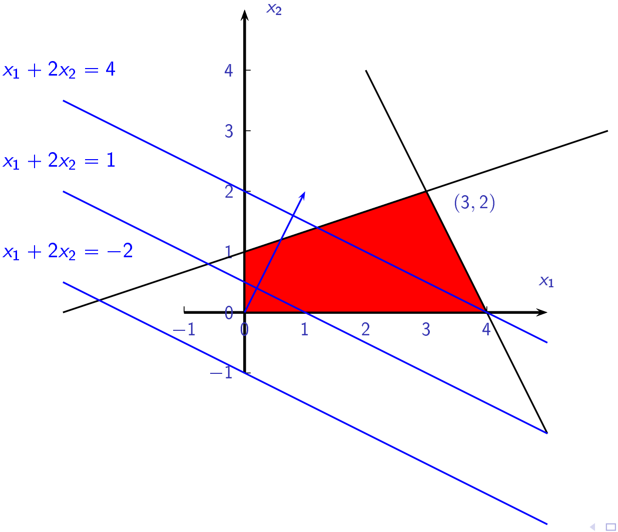
What is Linear Programming?



maximize $x_1 + 2x_2$

$$\begin{cases} x_1 & \geqslant & 0 & & \\ x_2 & \geqslant & 0 & & \\ 2x_1 & + & x_2 & \leqslant & 8 \\ -x_1 & + & 3x_2 & \leqslant & 3 \end{cases}$$

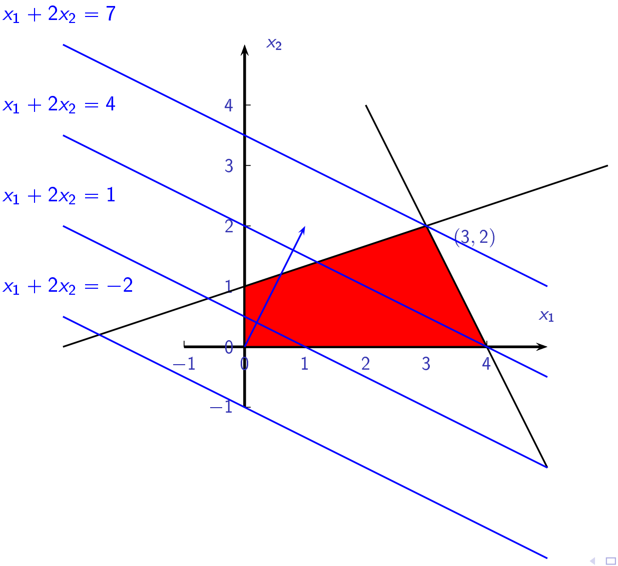
What is Linear Programming?



maximize $x_1 + 2x_2$

$$\begin{cases} x_1 & \geqslant & 0 \\ x_2 & \geqslant & 0 \\ 2x_1 + x_2 & \leqslant & 8 \\ -x_1 + 3x_2 & \leqslant & 3 \end{cases}$$

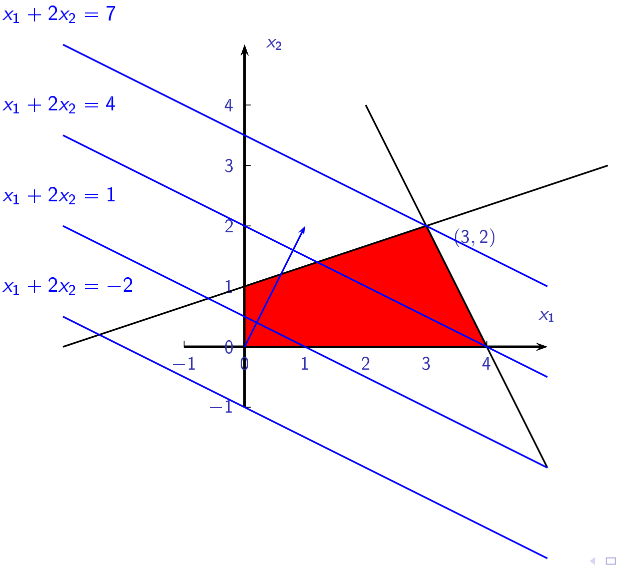
What is Linear Programming?



maximize $x_1 + 2x_2$

$$\begin{cases} x_1 & \geqslant & 0 \\ x_2 & \geqslant & 0 \\ 2x_1 + x_2 & \leqslant & 8 \\ -x_1 + 3x_2 & \leqslant & 3 \end{cases}$$

What is Linear Programming?



maximize $x_1 + 2x_2$

$$\begin{cases} x_1 & \geqslant & 0 \\ x_2 & \geqslant & 0 \\ 2x_1 & + & x_2 & \leqslant & 8 \\ -x_1 & + & 3x_2 & \leqslant & 3 \end{cases}$$

optimal solution is $(3, 2)$

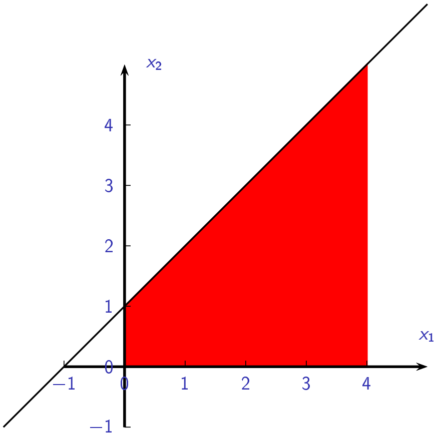
What is Linear Programming?

Example

Maximize the value $x_1 + 2x_2$ under the constraints

$$\begin{cases} x_1 & \geqslant & 0 & & \\ x_2 & \geqslant & 0 & & \\ -x_1 & + & x_2 & \leqslant & 1 \end{cases}$$

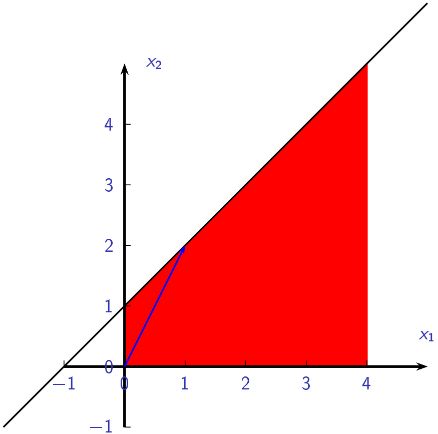
What is Linear Programming?



maximize $x_1 + 2x_2$

$$\begin{cases} x_1 & \geqslant & 0 \\ x_2 & \geqslant & 0 \\ -x_1 & + & x_2 & \leqslant & 1 \end{cases}$$

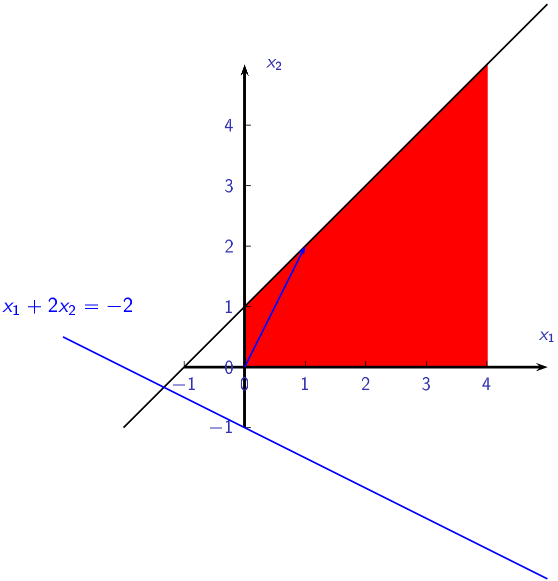
What is Linear Programming?



maximize $x_1 + 2x_2$

$$\begin{cases} x_1 & \geqslant & 0 \\ x_2 & \geqslant & 0 \\ -x_1 & + & x_2 & & \leqslant & 1 \end{cases}$$

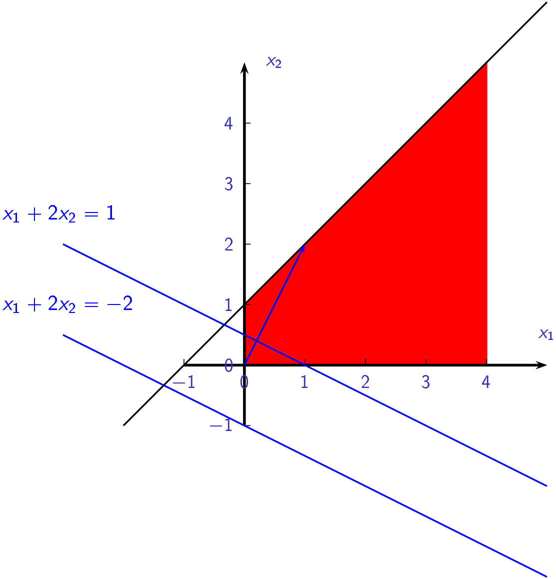
What is Linear Programming?



maximize $x_1 + 2x_2$

$$\begin{cases} x_1 & \geqslant & 0 \\ x_2 & \geqslant & 0 \\ -x_1 + x_2 & \leqslant & 1 \end{cases}$$

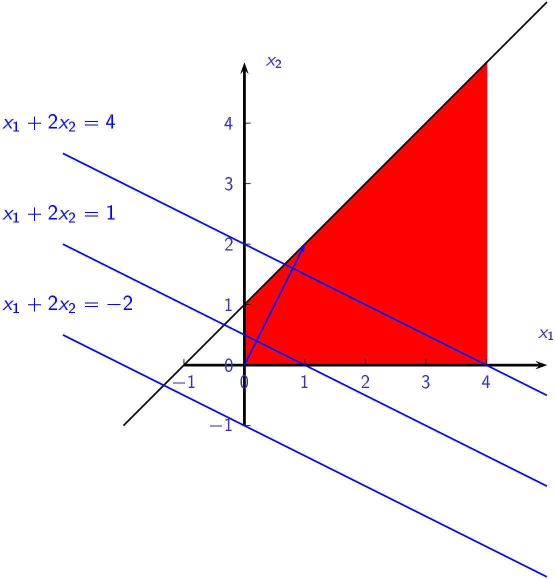
What is Linear Programming?



maximize $x_1 + 2x_2$

$$\begin{cases} x_1 & \geqslant & 0 \\ x_2 & \geqslant & 0 \\ -x_1 & + & x_2 & \leqslant & 1 \end{cases}$$

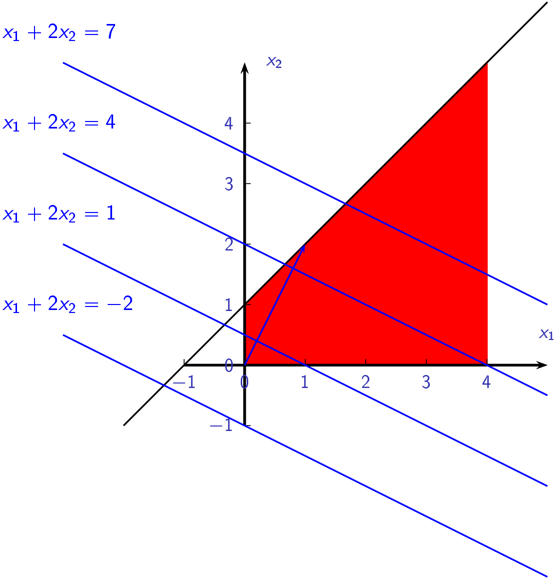
What is Linear Programming?



maximize $x_1 + 2x_2$

$$\begin{cases} x_1 \geqslant 0 \\ x_2 \geqslant 0 \\ -x_1 + x_2 \leqslant 1 \end{cases}$$

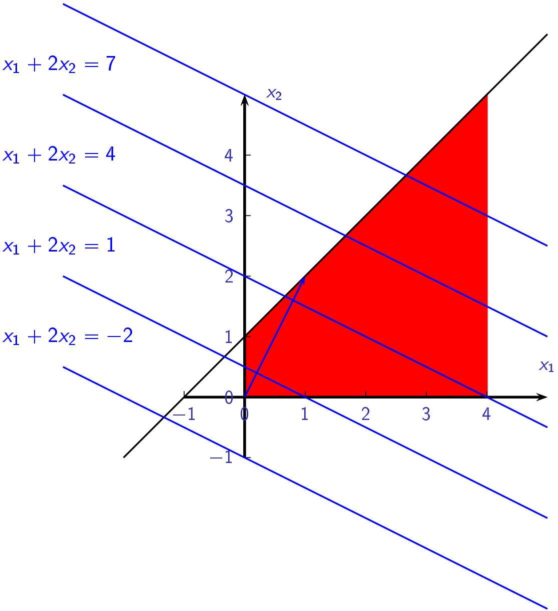
What is Linear Programming?



maximize $x_1 + 2x_2$

$$\begin{cases} x_1 & \geqslant & 0 \\ x_2 & \geqslant & 0 \\ -x_1 + x_2 & \leqslant & 1 \end{cases}$$

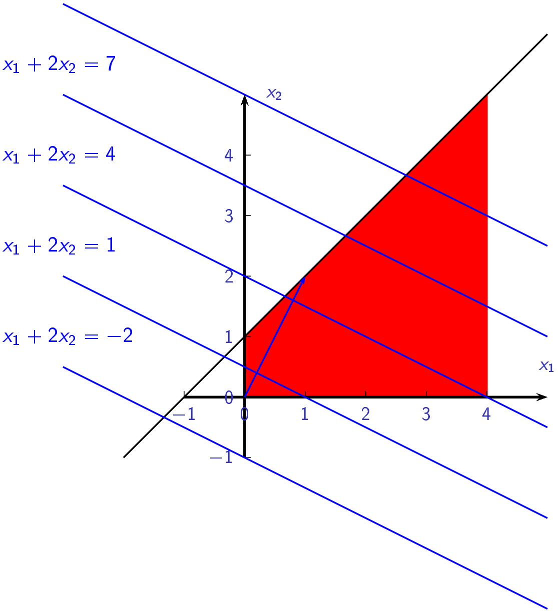
What is Linear Programming?



maximize $x_1 + 2x_2$

$$\begin{cases} x_1 & \geqslant & 0 \\ x_2 & \geqslant & 0 \\ -x_1 & + & x_2 & \leqslant & 1 \end{cases}$$

What is Linear Programming?



maximize $x_1 + 2x_2$

$$\begin{cases} x_1 \geqslant 0 \\ x_2 \geqslant 0 \\ -x_1 + x_2 \leqslant 1 \end{cases}$$

no optimal solution

Economy and Economical

The second meaning of ‘economy’ in the Oxford British and World English Dictionary

Careful management of available resources.

Economy and Economical

The second meaning of 'economy' in the Oxford British and World English Dictionary

Careful management of available resources.

The first meaning of 'economical' in the Oxford British and World English Dictionary

Giving good value or return in relation to the money, time, or effort expended.

Economy and Economical

The second meaning of ‘economy’ in the Oxford British and World English Dictionary

Careful management of available resources.

The first meaning of ‘economical’ in the Oxford British and World English Dictionary

Giving good value or return in relation to the money, time, or effort expended.

from Greek

oikonomia=household management, housekeeping

Linear Programming Problem

Definition

Linear programming problem is a task of maximizing or minimizing a linear function (called an **objective function**) over a set $X \subset \mathbb{R}^n$ described by a finite number of linear inequalities.

Linear Programming Problem

Definition

Linear programming problem is a task of maximizing or minimizing a linear function (called an **objective function**) over a set $X \subset \mathbb{R}^n$ described by a finite number of linear inequalities.

That is, we look for the maximal or the minimal value of the function $f((x_1, x_2, \ldots, x_n)) = c_1x_1 + c_2x + \ldots + c_nx_n$ on the set $X \subset \mathbb{R}^n$ of points satisfying the system of linear inequalities, i.e.

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & \leqslant & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & \leqslant & b_2 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & \leqslant & b_m \end{cases}$$

Linear Programming Problem (continued)

Those conditions (also called **constraints**) can be written in a concise form. Let

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix},$$

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \; c = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}, \; b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

This is an example of a global optimization problem with (inequality) constraints

Linear Programming Problem (continued)

The linear programming problem can be written in the form: maximize (or minimize) the linear function $f(x) = c^{\mathsf{T}}x$ over the set $X \subset \mathbb{R}^n$ given by

$$Ax \leqslant b.$$

Linear Programming Problem (continued)

The linear programming problem can be written in the form: maximize (or minimize) the linear function $f(x) = c^\mathsf{T} x$ over the set $X \subset \mathbb{R}^n$ given by

$$Ax \leqslant b.$$

Equivalently, one can write $f(x) \longrightarrow max$ (resp. $f(x) \longrightarrow min$).

Linear Programming Problem (continued)

The linear programming problem can be written in the form: maximize (or minimize) the linear function $f(x) = c^\intercal x$ over the set $X \subset \mathbb{R}^n$ given by

$$Ax \leqslant b.$$

Equivalently, one can write $f(x) \longrightarrow max$ (resp. $f(x) \longrightarrow min$).

Remark

Since

$$a_1x_1 + \ldots + a_nx_n = b \Longleftrightarrow \begin{cases} a_1x_1 + \ldots + a_nx_n \leqslant & b \\ -a_1x_1 - \ldots - a_nx_n \leqslant & -b \end{cases}$$

a set given by a finite number of linear equations and finite number of inequalities can be expressed by a finite number of inequalities.

Real Life Applications - Transportation Problem

A firm stores some goods at l supply centers and ships those goods to k markets. The cost of transporting a unit of those goods from the i-th supply center to the j-th market is a_{ij}. Each market demands at least of b_j units of those goods. Each supply center produces at most w_i units of goods.

Real Life Applications - Transportation Problem

A firm stores some goods at l supply centers and ships those goods to k markets. The cost of transporting a unit of those goods from the i-th supply center to the j-th market is a_{ij}. Each market demands at least of b_j units of those goods. Each supply center produces at most w_i units of goods.
Introduce $l \times k$ variables x_{ij} for $i = 1, \ldots, l$ and $j = 1, \ldots, k$ denoting the amount of the transport from the i-th supply center to the j-th market. We want to minimize the cost of transport and to satisfy demands of all markets.

Real Life Applications - Transportation Problem

A firm stores some goods at l supply centers and ships those goods to k markets. The cost of transporting a unit of those goods from the i-th supply center to the j-th market is a_{ij}. Each market demands at least of b_j units of those goods. Each supply center produces at most w_i units of goods.

Introduce $l \times k$ variables x_{ij} for $i = 1, \ldots, l$ and $j = 1, \ldots, k$ denoting the amount of the transport from the i-th supply center to the j-th market. We want to minimize the cost of transport and to satisfy demands of all markets. We minimize the linear function $\sum_{i=1}^{l} \sum_{j=1}^{k} a_{ij} x_{ij}$ under the constraints

$$\begin{cases} x_{11} + x_{12} + x_{13} + \ldots + x_{1k} \leqslant w_1 \\ x_{21} + x_{22} + x_{23} + \ldots + x_{2k} \leqslant w_2 \\ \vdots \\ x_{l1} + x_{l2} + x_{l3} + \ldots + x_{lk} \leqslant w_l \end{cases}$$

i.e. no supply center cannot supply more than w_i of goods and

Real Life Applications - Transportation Problem

$$\begin{cases} x_{11} + x_{21} + x_{31} + \ldots + x_{l1} \geqslant b_1 \\ x_{12} + x_{22} + x_{32} + \ldots + x_{l2} \geqslant b_2 \\ \vdots \\ x_{1k} + x_{2k} + x_{3k} + \ldots + x_{lk} \geqslant b_k \end{cases}$$

i.e. the demand of each market is satisfied.

Real Life Applications - Transportation Problem

$$\begin{cases} x_{11} + x_{21} + x_{31} + \ldots + x_{l1} \geqslant b_1 \\ x_{12} + x_{22} + x_{32} + \ldots + x_{l2} \geqslant b_2 \\ \vdots \\ x_{1k} + x_{2k} + x_{3k} + \ldots + x_{lk} \geqslant b_k \end{cases}$$

i.e. the demand of each market is satisfied. We want to transport from a supply center to a market so we assume

$$x_{ij} \geqslant 0 \text{ for } i = 1, \ldots, l \text{ and } j = 1, \ldots, k.$$

Real Life Application - Diet Problem

Suppose there are n foods available. The cost of serving per j-th food is q_j. Assume there are k nutrients and each serving of j-th type of food contains z_{ij} units of the i-th nutrient.

Real Life Application - Diet Problem

Suppose there are n foods available. The cost of serving per j-th food is q_j. Assume there are k nutrients and each serving of j-th type of food contains z_{ij} units of the i-th nutrient. We want to find a healthy diet minimizing its cost. Let N_i denotes the minimal amount of units of the i-th nutrient in a healthy diet. Introduce n variables $x_1, \ldots, x_n$, where x_j stands for the amount of servings of the j-th food.

Real Life Application - Diet Problem

Suppose there are n foods available. The cost of serving per j-th food is q_j. Assume there are k nutrients and each serving of j-th type of food contains z_{ij} units of the i-th nutrient. We want to find a healthy diet minimizing its cost. Let N_i denotes the minimal amount of units of the i-th nutrient in a healthy diet. Introduce n variables $x_1, \ldots, x_n$, where x_j stands for the amount of servings of the j-th food. We minimize the function $q_1x_1 + q_2x_2 + \ldots + q_nx_n$ under the constraints $x_1, \ldots, x_n \geqslant 0$ and

Real Life Application - Diet Problem

Suppose there are n foods available. The cost of serving per j-th food is q_j. Assume there are k nutrients and each serving of j-th type of food contains z_{ij} units of the i-th nutrient. We want to find a healthy diet minimizing its cost. Let N_i denotes the minimal amount of units of the i-th nutrient in a healthy diet. Introduce n variables $x_1, \ldots, x_n$, where x_j stands for the amount of servings of the j-th food. We minimize the function $q_1x_1 + q_2x_2 + \ldots + q_nx_n$ under the constraints $x_1, \ldots, x_n \geqslant 0$ and

$$\begin{cases} z_{11}x_1 + z_{12}x_2 + z_{13}x_3 + \ldots + x_{1n}x_n \geqslant N_1 \\ z_{21}x_1 + z_{22}x_2 + z_{23}x_3 + \ldots + x_{2n}x_n \geqslant N_2 \\ \vdots \\ z_{k1}x_1 + z_{k2}x_2 + z_{k3}x_3 + \ldots + x_{kn}x_n \geqslant N_k \end{cases}$$

Real Life Applications - Diet Problem

If needed one may add another constraints for the minimal or maximal amount of servings of each type of food.

Real Life Applications - Diet Problem

If needed one may add another constraints for the minimal or maximal amount of servings of each type of food. A similar problem was considered in 1930s and 1940s in order to find an optimal diet for the US soldiers.

Real Life Applications

And many more: portfolio optimization, network design, vehicle routing.

Convex Set

Definition

For any $p, q \in \mathbb{R}^n$ the **line segment** joining p and q is the set

$$[p, q] = \{tp + (1 - t)q \in \mathbb{R}^n \mid t \in [0, 1]\}.$$

Convex Set

Definition

For any $p, q \in \mathbb{R}^n$ the **line segment** joining p and q is the set

$$[p, q] = \{tp + (1 - t)q \in \mathbb{R}^n \mid t \in [0, 1]\}.$$

Definition

A set $X \subset \mathbb{R}^n$ is **convex** if

$$[p, q] \subset X \text{ for any } p, q \in X.$$

Open and Closed Ball

Definition

An open ball with center $x \in \mathbb{R}^n$ and the radius $r > 0$ is the set

$$B(x, r) = \{y \in \mathbb{R}^n \mid \|x - y\| < r\}.$$

A closed ball with center $x \in \mathbb{R}^n$ and the radius $r > 0$ is the set

$$\overline{B}(x, r) = \{y \in \mathbb{R}^n \mid \|x - y\| \leqslant r\}.$$

Convex Set (continued)

Proposition

A ball is a convex set.

Convex Set (continued)

Proposition

A ball is a convex set.

Proof.

Let $p, q \in B(x, r)$. Then for any $t \in [0, 1]$

$$\|x - (tp + (1 - t)q)\| = \|t(x - p) + (1 - t)(x - q)\| \leqslant$$

$$\leqslant t\|x - p\| + (1 - t)\|x - q\| < r,$$

that is

$$[p, q] \subset B(x, r).$$

Convex Set (continued)

Proposition

A ball is a convex set.

Proof.

Let $p, q \in B(x, r)$. Then for any $t \in [0, 1]$

$$\|x - (tp + (1-t)q)\| = \|t(x-p) + (1-t)(x-q)\| \leqslant$$

$$\leqslant t\|x-p\| + (1-t)\|x-q\| < r,$$

that is

$$[p, q] \subset B(x, r).$$

The same proof works for a closed ball. $\square$

Convex Set (continued)

Proposition

Intersection of a family of convex sets is a convex set. In particular, if $X_1, \ldots, X_m \subset \mathbb{R}^n$ are convex sets then

$$X_1 \cap \ldots \cap X_m = \{x \in \mathbb{R}^n \mid x \in X_i \text{ for } i = 1, \ldots, m\},$$

is a convex set.

Half-space

Definition

A **half-space** $H_\leqslant \subset \mathbb{R}^n$ is a set given by a single inequality, that is

$$H_\leqslant = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid a_1x_1 + \ldots + a_nx_n \leqslant b\}.$$

Half-space

Definition

A **half-space** $H_{\leqslant} \subset \mathbb{R}^n$ is a set given by a single inequality, that is

$$H_{\leqslant} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid a_1 x_1 + \ldots + a_n x_n \leqslant b\}.$$

Proposition

A half-space is a convex set.

Half-space

Definition

A **half-space** $H_\leqslant \subset \mathbb{R}^n$ is a set given by a single inequality, that is

$$H_\leqslant = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid a_1 x_1 + \ldots + a_n x_n \leqslant b\}.$$

Proposition

A half-space is a convex set.

Proof.

Let $p = (p_1, \ldots, p_n), q = (q_1, \ldots, q_n) \in H_\leqslant$. Then for any $t \in [0, 1]$

$$a_1(tp_1 + (1-t)q_1) + a_2(tp_2 + (1-t)q_2) + \ldots + a_n(tp_n + (1-t)q_n) =$$

$$t(a_1 p_1 + a_2 p_2 + \ldots + a_n p_n) + (1-t)(a_1 q_1 + a_2 q_2 + \ldots + a_n q_n) \leqslant tb + (1-t)b = b,$$

i.e.

$$[p, q] \subset H_\leqslant.$$

$\square$

Open Sets, Closed Sets

Definition

A set $U \subset \mathbb{R}^n$ is **open** if for every $x \in U$ there exists a radius $r > 0$ such that

$$B(x, r) \subset U.$$

Open Sets, Closed Sets

Definition

A set $U \subset \mathbb{R}^n$ is **open** if for every $x \in U$ there exists a radius $r > 0$ such that
$$B(x, r) \subset U.$$
A set $D \subset \mathbb{R}^n$ is **closed** if the set $\mathbb{R}^n \backslash D$ is open.

Open Sets, Closed Sets

Definition

A set $U \subset \mathbb{R}^n$ is **open** if for every $x \in U$ there exists a radius $r > 0$ such that

$$B(x, r) \subset U.$$

A set $D \subset \mathbb{R}^n$ is **closed** if the set $\mathbb{R}^n \backslash D$ is open.

Example

An open ball is an open set. A closed ball is a closed set. A half-space is a closed set.

Open Sets, Closed Sets

Definition

A set $U \subset \mathbb{R}^n$ is **open** if for every $x \in U$ there exists a radius $r > 0$ such that

$$B(x, r) \subset U.$$

A set $D \subset \mathbb{R}^n$ is **closed** if the set $\mathbb{R}^n \backslash D$ is open.

Example

An open ball is an open set. A closed ball is a closed set. A half-space is a closed set.

Proof.

It is enough to show that $\mathbb{R}^n \backslash H_{\leqslant}$ is an open set. If $x \in \mathbb{R}^n \backslash H_{\leqslant}$ let $r = d(x, H)$ be the distance of x from H. Then

$$B(x, r) \subset \mathbb{R}^n \backslash H_{\leqslant}.$$

$\square$

Open Sets, Closed Sets (continued)

Proposition

Let $X_1, \ldots, X_m \subset \mathbb{R}^n$ *be open sets. Then the sets*

$$X_1 \cup \ldots \cup X_m \subset \mathbb{R}^n,$$

and

$$X_1 \cap \ldots \cap X_m \subset \mathbb{R}^n,$$

are open.

Open Sets, Closed Sets (continued)

Proposition

Let $X_1, \ldots, X_m \subset \mathbb{R}^n$ be open sets. Then the sets

$$X_1 \cup \ldots \cup X_m \subset \mathbb{R}^n,$$

and

$$X_1 \cap \ldots \cap X_m \subset \mathbb{R}^n,$$

are open.

Proof.

Let $x \in X_1 \cup \ldots \cup X_m \subset \mathbb{R}^n$. Then $x \in X_i$ for some i. Since X_i is open there exists $r > 0$ such that

$$B(x, r) \subset X_i \subset X_1 \cup \ldots \cup X_m.$$

Open Sets, Closed Sets (continued)

Proposition

Let $X_1, \dots, X_m \subset \mathbb{R}^n$ be open sets. Then the sets

$$X_1 \cup \ldots \cup X_m \subset \mathbb{R}^n,$$

and

$$X_1 \cap \ldots \cap X_m \subset \mathbb{R}^n,$$

are open.

Proof.

Let $x \in X_1 \cup \ldots \cup X_m \subset \mathbb{R}^n$. Then $x \in X_i$ for some i. Since X_i is open there exists $r > 0$ such that

$$B(x, r) \subset X_i \subset X_1 \cup \ldots \cup X_m.$$

If $x \in X_i$ for $i = 1, \dots, m$, then there exist $r_i > 0$ such that $B(x, r_i) \subset X_i$ for $i = 1, \dots, m$. Let $r = \min\{r_1, \dots, r_m\}$. Then

$$B(x, r) \subset X_1 \cap \ldots \cap X_m.$$

Open Sets, Closed Sets (continued)

Corollary

Let $X_1, \ldots, X_m \subset \mathbb{R}^n$ *be closed sets. Then the sets*

$$X_1 \cup \ldots \cup X_m \subset \mathbb{R}^n,$$

and

$$X_1 \cap \ldots \cap X_m \subset \mathbb{R}^n,$$

are closed.

Proof.

$$\mathbb{R}^n \backslash (X_1 \cup \ldots \cup X_m) = (\mathbb{R}^n \backslash X_1) \cap \ldots \cap (\mathbb{R}^n \backslash X_m),$$
$$\mathbb{R}^n \backslash (X_1 \cap \ldots \cap X_m) = (\mathbb{R}^n \backslash X_1) \cup \ldots \cup (\mathbb{R}^n \backslash X_m),$$

where $\mathbb{R}^n \backslash X_i$ are open. $\square$

Open Sets, Closed Sets (continued)

Corollary

Let $X_1, \ldots, X_m \subset \mathbb{R}^n$ *be closed sets. Then the sets*

$$X_1 \cup \ldots \cup X_m \subset \mathbb{R}^n,$$

and

$$X_1 \cap \ldots \cap X_m \subset \mathbb{R}^n,$$

are closed.

Proof.

$$\mathbb{R}^n \backslash (X_1 \cup \ldots \cup X_m) = (\mathbb{R}^n \backslash X_1) \cap \ldots \cap (\mathbb{R}^n \backslash X_m),$$
$$\mathbb{R}^n \backslash (X_1 \cap \ldots \cap X_m) = (\mathbb{R}^n \backslash X_1) \cup \ldots \cup (\mathbb{R}^n \backslash X_m),$$

where $\mathbb{R}^n \backslash X_i$ are open. □

Remark

In general, the union of any family of open sets is an open set and the intersection of any family of closed sets is a closed set.

Convex Polytopes

Definition

Convex polytope $X \subset \mathbb{R}^n$ is a non–empty set of solutions of a system of linear inequalities, i.e.

$$X = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid Ax \leqslant b\},$$

where $A \in M(m \times n; \mathbb{R})$. Equivalently, it is a non–empty intersection of finite number of half-spaces.

Convex Polytopes

Definition

Convex polytope $X \subset \mathbb{R}^n$ is a non–empty set of solutions of a system of linear inequalities, i.e.

$$X = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid Ax \leqslant b\},$$

where $A \in M(m \times n; \mathbb{R})$. Equivalently, it is a non–empty intersection of finite number of half-spaces.

Proposition

Convex polytope is a closed convex set.

Convex Polytopes

Definition

Convex polytope $X \subset \mathbb{R}^n$ is a non–empty set of solutions of a system of linear inequalities, i.e.

$$X = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid Ax \leqslant b\},$$

where $A \in M(m \times n; \mathbb{R})$. Equivalently, it is a non–empty intersection of finite number of half-spaces.

Proposition

Convex polytope is a closed convex set.

Proof.

It is an intersection of closed convex sets. $\square$

Polyhedra

Remark

Sometimes a different terminology is used: a **polyhedron** *(or a* **polyhedral set***) is a set of solutions of a system $Ax \leqslant b$ and a* **polytope** *is a bounded polyhedron.*

Compact Set

Definition

Set $X \subset \mathbb{R}^n$ is **bounded** if there exists $x \in \mathbb{R}^n$ and $r > 0$ such that

$$X \subset B(x, r).$$

Definition

Set $X \subset \mathbb{R}^n$ is **compact** if it is **closed** and **bounded**.

Extreme Value Theorem

Theorem

Let $X \subset \mathbb{R}^n$ be a compact set and let

$$f : X \to \mathbb{R},$$

be a continuous function. Let

$$m = \inf_{x \in X} f(x), \quad M = \sup_{x \in X} f(x).$$

Then there exist $x_m, x_M \in X$ such that

$$m = f(x_m), \quad M = f(x_M).$$

Extreme Value Theorem

Theorem

Let $X \subset \mathbb{R}^n$ *be a compact set and let*

$$f : X \to \mathbb{R},$$

be a continuous function. Let

$$m = \inf_{x \in X} f(x), \quad M = \sup_{x \in X} f(x).$$

Then there exist $x_m, x_M \in X$ *such that*

$$m = f(x_m), \quad M = f(x_M).$$

Remark

Linear functions are continuous.

Convex Polytopes (continued)

Suppose we are given a linear programming problem with constraints $Ax \leqslant b$ and the objective function $f(x) = c^{\mathsf{T}}x \longrightarrow min$.

Convex Polytopes (continued)

Suppose we are given a linear programming problem with constraints $Ax \leqslant b$ and the objective function $f(x) = c^{\mathsf{T}}x \longrightarrow min$.

Definition

A **feasible region** (also a **feasible set)** is the set of all points $X \subset \mathbb{R}^n$ satisfying the conditions $Ax \leqslant b$. An optimal solution is any point $\overline{x} \in X$ such that $f(\overline{x}) \leqslant f(x)$ for any $x \in X$.

Convex Polytopes (continued)

Suppose we are given a linear programming problem with constraints $Ax \leqslant b$ and the objective function $f(x) = c^{\mathsf{T}}x \longrightarrow min$.

Definition

A **feasible region** (also a **feasible set)** is the set of all points $X \subset \mathbb{R}^n$ satisfying the conditions $Ax \leqslant b$. An optimal solution is any point $\overline{x} \in X$ such that $f(\overline{x}) \leqslant f(x)$ for any $x \in X$.

Remark

A feasible region is a convex polytope.

Convex Polytopes (continued)

Suppose we are given a linear programming problem with constraints $Ax \leqslant b$ and the objective function $f(x) = c^{\mathsf{T}}x \longrightarrow min$.

Definition

A **feasible region** (also a **feasible set)** is the set of all points $X \subset \mathbb{R}^n$ satisfying the conditions $Ax \leqslant b$. An optimal solution is any point $\overline{x} \in X$ such that $f(\overline{x}) \leqslant f(x)$ for any $x \in X$.

Remark

A feasible region is a convex polytope. If it is bounded (i.e. contained in a ball) then there exists an optimal solution. An optimal solution may not be unique.

Supporting Hyperplane

Definition

Let $X \subset \mathbb{R}^n$ be a convex closed set. A **supporting hyperplane** of X is a hyperplane

$$H = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid a_1 x_1 + \ldots + a_n x_n = b\},$$

such that $(a_1, \ldots, a_n) \neq 0$

$$H \cap X \neq \emptyset \text{ and } X \subset H_{\leqslant}.$$

Supporting Hyperplane

Definition

Let $X \subset \mathbb{R}^n$ be a convex closed set. A **supporting hyperplane** of X is a hyperplane

$$H = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid a_1 x_1 + \ldots + a_n x_n = b\},$$

such that $(a_1, \ldots, a_n) \neq 0$

$$H \cap X \neq \emptyset \text{ and } X \subset H_{\leqslant}.$$

Definition

A **face** of a convex polytope X is the intersection of X with a supporting hyperplane.

Supporting Hyperplane

Definition

Let $X \subset \mathbb{R}^n$ be a convex closed set. A **supporting hyperplane** of X is a hyperplane

$$H = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid a_1 x_1 + \ldots + a_n x_n = b\},$$

such that $(a_1, \ldots, a_n) \neq 0$

$$H \cap X \neq \emptyset \text{ and } X \subset H_{\leqslant}.$$

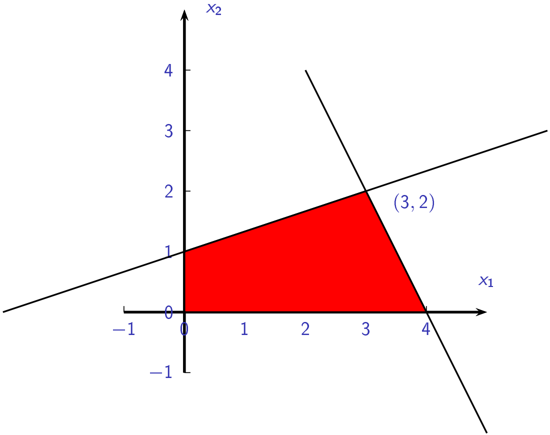
Definition

A **face** of a convex polytope X is the intersection of X with a supporting hyperplane.

Remark

A face of a convex polytope is a convex polytope.

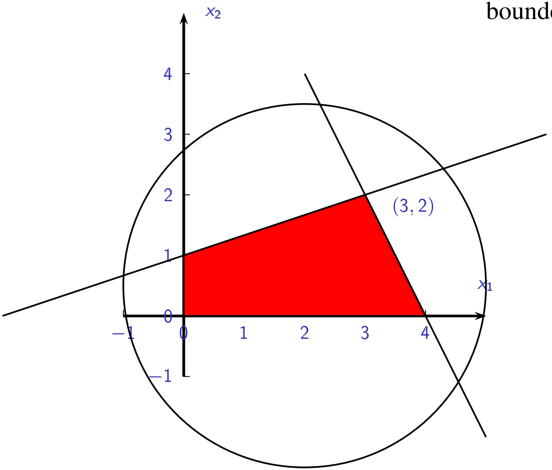
Example



$$\begin{cases} x_1 & \geqslant & 0 & & \\ x_2 & \geqslant & 0 & & \\ 2x_1 & + & x_2 & \leqslant & 8 \\ -x_1 & + & 3x_2 & \leqslant & 3 \end{cases}$$

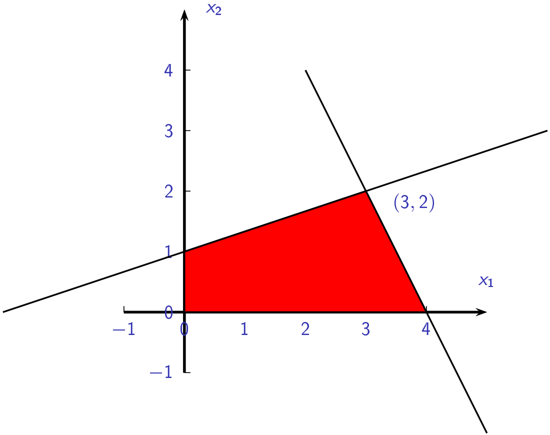
Example

bounded $\Rightarrow$ optimal solution exists



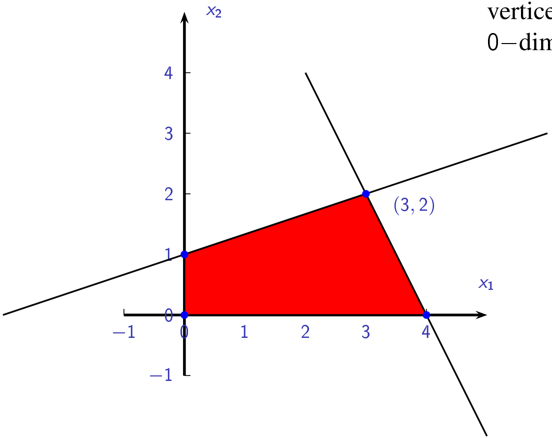
$$\begin{cases} x_1 & \geqslant & 0 \\ x_2 & \geqslant & 0 \\ 2x_1 & + & x_2 & \leqslant & 8 \\ -x_1 & + & 3x_2 & \leqslant & 3 \end{cases}$$

Example



$$\begin{cases} x_1 & \geqslant & 0 & & \\ x_2 & \geqslant & 0 & & \\ 2x_1 & + & x_2 & \leqslant & 8 \\ -x_1 & + & 3x_2 & \leqslant & 3 \end{cases}$$

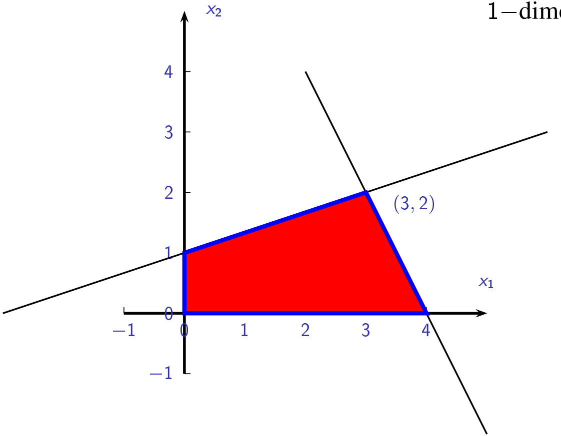
Example



vertices, i.e.
0−dimensional faces

$$\begin{cases} x_1 & \geqslant & 0 & & \\ x_2 & \geqslant & 0 & & \\ 2x_1 & + & x_2 & \leqslant & 8 \\ -x_1 & + & 3x_2 & \leqslant & 3 \end{cases}$$

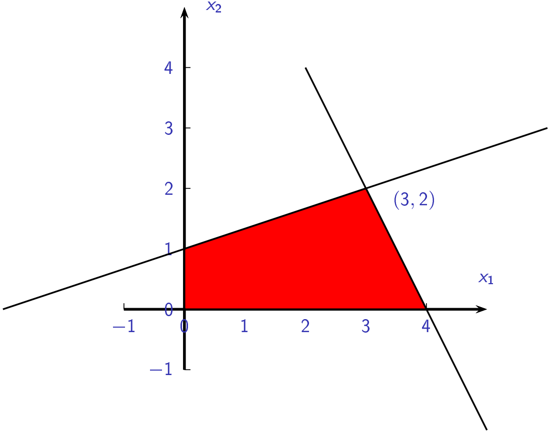
Example



1−dimensional faces

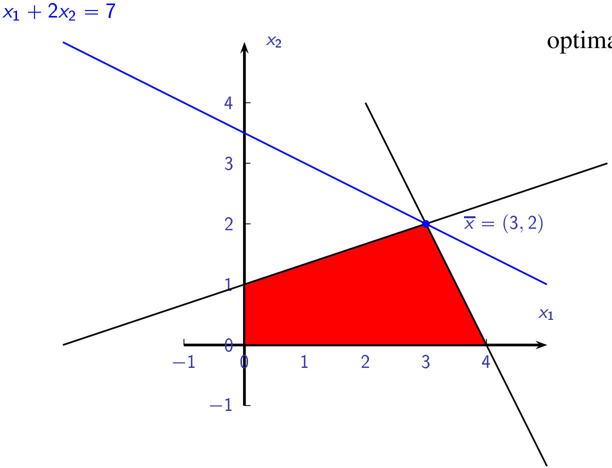
$$\begin{cases} x_1 & \geqslant & 0 \\ x_2 & \geqslant & 0 \\ 2x_1 & + & x_2 & \leqslant & 8 \\ -x_1 & + & 3x_2 & \leqslant & 3 \end{cases}$$

Example



$$\begin{cases} x_1 & \geqslant & 0 & & \\ x_2 & \geqslant & 0 & & \\ 2x_1 & + & x_2 & \leqslant & 8 \\ -x_1 & + & 3x_2 & \leqslant & 3 \end{cases}$$

Example

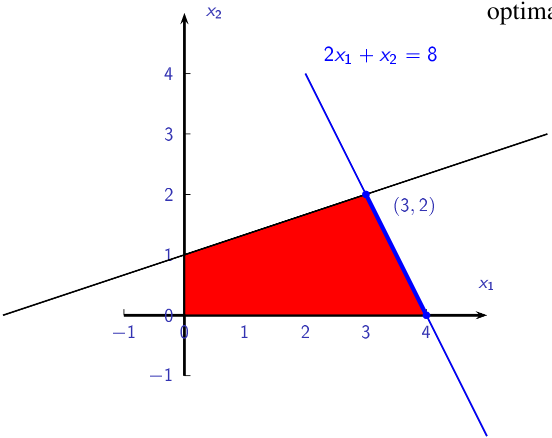


optimal solution is a vertex

maximize $x_1 + 2x_2$

$$\begin{cases} x_1 & \geqslant & 0 & & \\ x_2 & \geqslant & 0 & & \\ 2x_1 & + & x_2 & \leqslant & 8 \\ -x_1 & + & 3x_2 & \leqslant & 3 \end{cases}$$

Example



optimal solution may not be unique

maximize $2x_1 + x_2$

$$\begin{cases} x_1 & \geqslant & 0 \\ x_2 & \geqslant & 0 \\ 2x_1 & + & x_2 & \leqslant & 8 \\ -x_1 & + & 3x_2 & \leqslant & 3 \end{cases}$$

Extreme Points

Definition

Let $X \subset \mathbb{R}^n$ be a convex set. Point $x \in \mathbb{R}^n$ is an **extreme point** of X if for any $p, q \in X$

$$\text{if } x \in [p, q] \text{ then } x = p \text{ or } x = q.$$

Extreme Points of a Convex Polytope

Definition

Let $X \subset \mathbb{R}^n$ be a convex polytope given by $Ax \leqslant b$, where $A \in M(m \times n; \mathbb{R})$, $b \in \mathbb{R}^m$. Let $a_1, \ldots, a_m \in \mathbb{R}^n$ denote the rows of matrix A. For any $p \in X$ denote by

$$J(p) = \{i \in \{1, \ldots, m\} \mid a_i^T p = b_i\}$$

the set of **active constraints**. Let $A_{J(p)}$ denote the submatrix of matrix A consisting of rows of A indexed by the set $J(p)$, the same for $b_{J(p)}$. In particular

$$A_{J(p)} p = b_{J(p)}.$$

Extreme Points of a Convex Polytope

Definition

Let $X \subset \mathbb{R}^n$ be a convex polytope given by $Ax \leqslant b$, where $A \in M(m \times n; \mathbb{R})$, $b \in \mathbb{R}^m$. Let $a_1, \ldots, a_m \in \mathbb{R}^n$ denote the rows of matrix A. For any $p \in X$ denote by

$$J(p) = \{i \in \{1, \ldots, m\} \mid a_i^T p = b_i\}$$

the set of **active constraints**. Let $A_{J(p)}$ denote the submatrix of matrix A consisting of rows of A indexed by the set $J(p)$, the same for $b_{J(p)}$. In particular

$$A_{J(p)} p = b_{J(p)}.$$

Proposition

Let $X \subset \mathbb{R}^n$ be a convex polytope given by $Ax \leqslant b$ and let $p \in X$ be a point. Then p is an extreme point of X if and only if $r(A_{J(p)}) = n$.

Extreme Points of a Convex Polytope (continued)

Proof.

Let $p \in X$ be an extreme point of X. Assume $r(A_{J(p)}) < n$. Then, by the Kronecker–Capelli theorem, there exists a non–zero solution $q \in \mathbb{R}^n$, that is

$$A_{J(p)} q = 0 \text{ and } q \neq 0.$$

Let $a_1, \ldots, a_m \in \mathbb{R}^n$ denote the rows of matrix A. For sufficiently small $t \in \mathbb{R}$, $t \neq 0$

$$\left| t(a_i^\top q) \right| < b_i - a_i^\top p \text{ for any } i \notin J(p),$$

which gives $p \pm tq \in X$ since $a_i^\top (p \pm tq) = b_i$ for $i \in J(p)$. Then $p \neq p \pm tq$ and $p \in [p - tq, p + tq]$ because

$$p = \frac{1}{2}(p - tq) + \frac{1}{2}(p + tq),$$

which leads to a contradiction with $p \in X$ being an extreme point.[1]

[1] Proof based on N. Lauritzen, *Lectures on Convex Sets*.

Extreme Points of a Convex Polytope (continued)

Proof.

Assume that $r(A_{J(p)}) = n$ and let $p = tp_1 + (1-t)p_2$ for some $t \in (0,1), p_1, p_2 \in X$, where $p_1 \neq p_2$. Then

$$b_{J(p)} = A_{J(p)}p = tA_{J(p)}p_1 + (1-t)A_{J(p)}p_2 \leqslant b_{J(p)},$$

which implies

$$A_{J(p)}p = A_{J(p)}p_1 = A_{J(p)}p_2 = b_{J(p)}.$$

Extreme Points of a Convex Polytope (continued)

Proof.

Assume that $r(A_{J(p)}) = n$ and let $p = tp_1 + (1-t)p_2$ for some $t \in (0,1)$, $p_1, p_2 \in X$, where $p_1 \neq p_2$. Then

$$b_{J(p)} = A_{J(p)}p = tA_{J(p)}p_1 + (1-t)A_{J(p)}p_2 \leqslant b_{J(p)},$$

which implies

$$A_{J(p)}p = A_{J(p)}p_1 = A_{J(p)}p_2 = b_{J(p)}.$$

Since $r(A_{J(p)}) = n$ the system of linear equations $A_{J(p)}x = b_{J(p)}$ has a unique solution hence $p = p_1 = p_2$. By contradiction, either $t = 0$ or $t = 1$. $\square$

Extreme Points of a Convex Polytope (continued)

Corollary

Let $X \subset \mathbb{R}^n$ *be a convex polytope given by* $Ax \leqslant b$, *where* $A \in M(m \times n; \mathbb{R}), b \in \mathbb{R}^m$. *Then if* $X \neq \varnothing$

$$X \text{ has no extreme points} \iff r(A) < n.$$

Extreme Points of a Convex Polytope (continued)

Corollary

Let $X \subset \mathbb{R}^n$ be a convex polytope given by $Ax \leqslant b$, where $A \in M(m \times n; \mathbb{R}), b \in \mathbb{R}^m$. Then if $X \neq \varnothing$

$$X \text{ has no extreme points} \iff r(A) < n.$$

Proof.

$(\Leftarrow)$ follows from the previous proposition,

Extreme Points of a Convex Polytope (continued)

Corollary

Let $X \subset \mathbb{R}^n$ be a convex polytope given by $Ax \leqslant b$, where $A \in M(m \times n; \mathbb{R}), b \in \mathbb{R}^m$. Then if $X \neq \varnothing$

$$X \text{ has no extreme points} \iff r(A) < n.$$

Proof.

($\Leftarrow$) follows from the previous proposition,

($\Rightarrow$) let $p \in X$ be some point, if $J(p) = \{1, \ldots, m\}$ then $A = A_{J(p)}$ and $r(A) < n$ since p in not an extreme point. If $J(p) \subsetneq \{1, \ldots, m\}$ then $r(A_{J(p)}) < n$ and there exist $q \in \mathbb{R}^n$, such that

$$A_{J(p)}q = 0 \text{ and } q \neq 0.$$

If $a_i^\intercal q = 0$ for $i \notin J(p)$ then $Aq = 0$ and $r(A) < n$. If $a_i^\intercal q < 0$ for all $i \notin J(p)$ then one can replace q with $-q$.

Extreme Points of a Convex Polytope (continued)

Proof.

Let

$$t = \min \left\{ \frac{b_i - a_i^\mathsf{T} p}{a_i^\mathsf{T} q} \in \mathbb{R} \mid i \notin J(p) \text{ and } a_i^\mathsf{T} q > 0 \right\}.$$

Then $t > 0$, $p + tq \in X$ and $J(p) \subsetneq J(p + tq)$. Eventually, by replacing p with $p + tq$ as above, one can find $p \in X$ such that $J(p) = \{1, \ldots, m\}$.[2] □

Corollary

If $m < n$ then the convex polytope X given by

$$Ax \leqslant b,$$

where $A \in M(m \times n; \mathbb{R})$, has no extreme points.

[2] Proof based on N. Lauritzen, *Lectures on Convex Sets*.

Vertices of Convex Polytopes

Definition

Let $X \subset \mathbb{R}^n$ be a convex polytope. Point $p \in X$ is a **vertex** of X if it is a face of X, i.e. there exists a half–space $H_{\leqslant} \subset \mathbb{R}^n$ such that

$$X \subset H_{\leqslant} \text{ and } X \cap H = \{p\}.$$

Proposition

Let $X \subset \mathbb{R}^n$ be a convex polytope given by the system of inequalities $Ax \leqslant b$. Let $p \in X$. Then

$$p \textit{ is an extreme point of } X \iff p \textit{ is a vertex of } X.$$

Vertices of Convex Polytopes (continued)

Proof.

($\Leftarrow$) Let

$$H = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid a_{m+1}^\mathsf{T} x = b_{m+1}\},$$

be the supporting hyperplane such that $X \cap H = \{p\}$. Since $X \subset H_\leqslant$ the polytope X is given by the system of inequalities $A'x \leqslant b'$ where

$$A' = \left[\frac{A}{a_{m+1}}\right], \quad b' = \left[\frac{b}{b_{m+1}}\right].$$

If $r(A'_{J(p)}) < n$ then, as in the previous proof, there exists $q \neq 0$ such that $A'_{J(p)}q = 0$ and $p + tq \in X$ for small $t \in \mathbb{R}$. Since $m + 1 \in J(p)$, that is $a_{m+1}^\mathsf{T} q = 0$

$$p + tq \in X \cap H,$$

which leads to a contradiction with $X \cap H = \{p\}$.

Vertices of Convex Polytopes (continued)

Proof.

($\Rightarrow$) Let $X \subset \mathbb{R}^n$ be given by $Ax \leqslant b$ and let $p \in X$ be an extreme point of X. Let

$$a_{m+1} = \sum_{i \in J'(p)} a_i, \quad b_{m+1} = \sum_{i \in J'(p)} b_i,$$

where

$$J'(p) = \{i_1, \dots, i_n\} \subset J(p),$$

and the n rows $a_{i_1}, \dots, a_{i_n} \in \mathbb{R}^n$ of $A_{J(p)}$ are linearly independent. Let

$$H = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid a_{m+1}^{\mathsf{T}} x = b_{m+1}\}.$$

By linear independence $a_{m+1} \neq 0$. Moreover, if $q \in X \cap H$, then $a_{i_j}^{\mathsf{T}} q = b_{i_j}$ for $j = 1, \dots, n$ (if sum is active then each summand is active too) which implies $q = p$. Therefore

$$X \subset H_{\leqslant}, \quad X \cap H = \{p\}.$$

$\square$

Standard Form

Definition

A linear programming problem in $\mathbb{R}^n$ is in the **standard form** if the constraints are given by a system of linear equations and all variables are non–negative, i.e

$$Ax = b, \; x_1, \ldots, x_n \geqslant 0,$$

and we look for the **minimum** of the objective function $f(x) = c^{\mathsf{T}} x$.

Standard Form

Definition

A linear programming problem in $\mathbb{R}^n$ is in the **standard form** if the constraints are given by a system of linear equations and all variables are non–negative, i.e

$$Ax = b, \ x_1, \ldots, x_n \geqslant 0,$$

and we look for the **minimum** of the objective function $f(x) = c^{\mathsf{T}} x$.

Moreover, we assume that $A \in M(m \times n; \mathbb{R})$ and

$$r(A) = r([A|b]) = m.$$

Standard Form

Definition

A linear programming problem in $\mathbb{R}^n$ is in the **standard form** if the constraints are given by a system of linear equations and all variables are non–negative, i.e

$$Ax = b, \ x_1, \ldots, x_n \geqslant 0,$$

and we look for the **minimum** of the objective function $f(x) = c^{\mathsf{T}} x$.

Moreover, we assume that $A \in M(m \times n; \mathbb{R})$ and

$$r(A) = r([A|b]) = m.$$

Remark

If $r(A) = r([A|b]) < m$ one can remove redundant equations. If $r(A) \neq r([A|b])$ then $X = \emptyset$.

Standard Form (continued)

Theorem

Any linear programming problem can be brought to the standard form.

Standard Form (continued)

Theorem

Any linear programming problem can be brought to the standard form.

The following operations on the a linear programming data give an equivalent problem:

i) the condition $f(x) \longrightarrow max$ can be replaced by

$$-f(x) \longrightarrow min,$$

Standard Form (continued)

Theorem

Any linear programming problem can be brought to the standard form.

The following operations on the a linear programming data give an equivalent problem:

i) the condition $f(x) \longrightarrow max$ can be replaced by

$$-f(x) \longrightarrow min,$$

ii) the inequality $a_1x_1 + \ldots + a_nx_n \leqslant b$ can replaced by $a_1x_1 + \ldots + a_nx_n + x_{n+1} = b$ and $x_{n+1} \geqslant 0$, the inequality $a_1x_1 + \ldots + a_nx_n \geqslant b$ can replaced by $a_1x_1 + \ldots + a_nx_n - x_{n+1} = b$ and $x_{n+1} \geqslant 0$, the newly introduced variable x_{n+1} is called **slack variable**,

Standard Form (continued)

Theorem

Any linear programming problem can be brought to the standard form.

The following operations on the a linear programming data give an equivalent problem:

i) the condition $f(x) \longrightarrow max$ can be replaced by

$$-f(x) \longrightarrow min,$$

ii) the inequality $a_1x_1 + \ldots + a_nx_n \leqslant b$ can replaced by $a_1x_1 + \ldots + a_nx_n + x_{n+1} = b$ and $x_{n+1} \geqslant 0$, the inequality $a_1x_1 + \ldots + a_nx_n \geqslant b$ can replaced by $a_1x_1 + \ldots + a_nx_n - x_{n+1} = b$ and $x_{n+1} \geqslant 0$, the newly introduced variable x_{n+1} is called **slack variable**,

iii) the condition $x_i \leqslant 0$ can be replaced by $x_i' \geqslant 0$ and $x_i' = -x_i$,

Standard Form (continued)

Theorem

Any linear programming problem can be brought to the standard form.

The following operations on the a linear programming data give an equivalent problem:

i) the condition $f(x) \longrightarrow max$ can be replaced by

$$-f(x) \longrightarrow min,$$

ii) the inequality $a_1x_1 + \ldots + a_nx_n \leqslant b$ can replaced by $a_1x_1 + \ldots + a_nx_n + x_{n+1} = b$ and $x_{n+1} \geqslant 0$, the inequality $a_1x_1 + \ldots + a_nx_n \geqslant b$ can replaced by $a_1x_1 + \ldots + a_nx_n - x_{n+1} = b$ and $x_{n+1} \geqslant 0$, the newly introduced variable x_{n+1} is called **slack variable**,

iii) the condition $x_i \leqslant 0$ can be replaced by $x_i' \geqslant 0$ and $x_i' = -x_i$,

iv) if there are no constraints on the variable x_i, one can introduce two slack variables $x_i^-, x_i^+ \geqslant 0$ and set $x_i = x_i^+ - x_i^-$.

Example

Bring to the standard form the following linear programming problem:

$x_1 + 2x_2 \longrightarrow max$

$$\begin{cases} x_1 & \geqslant & 0 & & \\ x_2 & \geqslant & 0 & & \\ 2x_1 & + & x_2 & \leqslant & 8 \\ -x_1 & + & 3x_2 & \leqslant & 3 \end{cases}$$

Example

Bring to the standard form the following linear programming problem:

$x_1 + 2x_2 \longrightarrow max$

$$\begin{cases} x_1 & \geqslant & 0 & & \\ x_2 & \geqslant & 0 & & \\ 2x_1 & + & x_2 & \leqslant & 8 \\ -x_1 & + & 3x_2 & \leqslant & 3 \end{cases}$$

A standard form: $-x_1 - 2x_2 \longrightarrow min$

$$\begin{cases} 2x_1 & + & x_2 & + & x_3 & & & = & 8 \\ -x_1 & + & 3x_2 & & & + & x_4 & = & 3 \end{cases}$$

and $x_1, x_2, x_3, x_4 \geqslant 0$.

Example (continued)

Equivalently, it can be written $c^T x \longrightarrow min$, $Ax = b$, $x \geqslant 0$ where

$$A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ -1 & 3 & 0 & 1 \end{bmatrix}, \; b = \begin{bmatrix} 8 \\ 3 \end{bmatrix}, \; x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \; c = \begin{bmatrix} -1 \\ -2 \\ 0 \\ 0 \end{bmatrix}$$

Example (continued)

Equivalently, it can be written $c^\mathsf{T} x \longrightarrow min, \;\; Ax = b, \;\; x \geqslant 0$ where

$$A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ -1 & 3 & 0 & 1 \end{bmatrix}, \; b = \begin{bmatrix} 8 \\ 3 \end{bmatrix}, \; x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \; c = \begin{bmatrix} -1 \\ -2 \\ 0 \\ 0 \end{bmatrix}$$

The optimal solution is

$$\overline{x} = \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \end{bmatrix} \text{ and } c^\mathsf{T}\overline{x} = -7$$

Optimal Solution and Vertices

Proposition

If the (non–empty) convex polytope X is given by $Ax = b, x \geqslant 0$, where $A \in M(m \times n; \mathbb{R})$, then it has a vertex.

Optimal Solution and Vertices

Proposition

If the (non–empty) convex polytope X is given by $Ax = b, x \geqslant 0$, where $A \in M(m \times n; \mathbb{R})$, then it has a vertex.

Proof.

It can be given by a system of inequalities

$$\left[\begin{array}{c} A \\ \hline -A \\ \hline -I_n \end{array}\right] x \leqslant \left[\begin{array}{c} b \\ \hline -b \\ \hline 0 \end{array}\right]$$

where the matrix of coefficients has rank n. $\square$

Optimal Solution and Vertices (continued)

Proposition

Let the convex polytope X be given by $Ax = b, x \geqslant 0$ where $A \in M(m \times n; \mathbb{R})$. If $\overline{x} \in X$ is an optimal solution for the problem $f(x) = c^{\mathsf{T}}x \longrightarrow min$, $c \neq 0$ then there exists a vertex $\overline{x}'$ of X such that

$$f(\overline{x}') = f(\overline{x}).$$

That is, an optimal solution, if it exists, can be chosen to be a vertex of the feasible set.

Optimal Solution and Vertices (continued)

Proof.

If $\overline{x} \in X$ is an optimal solution then

$$H = \{x \in \mathbb{R}^n \mid c^\mathsf{T} x = c^\mathsf{T} \overline{x}\},$$

is a supporting hyperplane of X such that $Y = X \cap H$ is a face of X and the function f is constant on Y. Therefore Y is a convex polytope which can be described by

$$\left[\begin{array}{r} A \\ \hline -A \\ \hline -I_n \\ \hline c \\ \hline -c \end{array}\right] x \leqslant \left[\begin{array}{r} b \\ \hline -b \\ \hline 0 \\ \hline c^\mathsf{T}\overline{x} \\ \hline -c^\mathsf{T}\overline{x} \end{array}\right].$$

It follows that Y has a vertex $\overline{x}' \in Y$.

Optimal Solution and Vertices (continued)

Proof.

The point $\overline{x}' \in Y$ is also a vertex of X since the convex polytope X can be described by matrix of rank n

$$\left[\begin{array}{r} A \\ \hline -A \\ \hline -I_n \\ \hline -c \end{array}\right] x \leqslant \left[\begin{array}{r} b \\ \hline -b \\ \hline 0 \\ \hline -c^{\mathsf{T}}\overline{x} \end{array}\right].$$

and the rank of the submatrix given by active inequalities for $\overline{x}'$ has rank n (the same as for Y). □

Basic Set, Basic Variables

Consider a linear programming problem in the standard form $c^\mathsf{T} x \longrightarrow min, Ax = b, x \geqslant 0$ where

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

and $r(A) = r([A|b]) = m$.

Definition

A basic set $\mathcal{B} = \{i_1, \ldots, i_m\} \subset \{1, \ldots, n\}$ is a set of m elements such that columns $c_{i_1}, \ldots, c_{i_m}$ of the matrix A are linearly independent (or equivalently, the determinant of square submatrix of A consisting of columns $c_{i_1}, \ldots, c_{i_m}$ is non–zero).

Basic Set, Basic Variables

Consider a linear programming problem in the standard form $c^{\mathsf{T}}x \longrightarrow min, Ax = b, x \geqslant 0$ where

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

and $r(A) = r([A|b]) = m$.

Definition

A basic set $\mathcal{B} = \{i_1, \ldots, i_m\} \subset \{1, \ldots, n\}$ is a set of m elements such that columns $c_{i_1}, \ldots, c_{i_m}$ of the matrix A are linearly independent (or equivalently, the determinant of square submatrix of A consisting of columns $c_{i_1}, \ldots, c_{i_m}$ is non–zero). The variables $x_{i_1}, \ldots, x_{i_m}$ are called **basic variables**. The other variables are called **non-basic**.

Basic Solution and Basic Feasible Solution

Definition

Let $\mathcal{B}$ be a basic set. The unique solution $\overline{x}_{\mathcal{B}} \in \mathbb{R}^n$ of the system of linear equations $Ax = b$ with $x_i = 0$ for $i \notin \mathcal{B}$ is called a **basic solution**.

Basic Solution and Basic Feasible Solution

Definition

Let $\mathcal{B}$ be a basic set. The unique solution $\overline{x}_{\mathcal{B}} \in \mathbb{R}^n$ of the system of linear equations $Ax = b$ with $x_i = 0$ for $i \notin \mathcal{B}$ is called a **basic solution**. The basic set $\mathcal{B}$ such that $\overline{x}_{\mathcal{B}} \geqslant 0$ is called a **feasible basic set** and the solution $\overline{x}_{\mathcal{B}}$ is called a **feasible basic solution**.

Basic Solution and Basic Feasible Solution

Definition

Let $\mathcal{B}$ be a basic set. The unique solution $\overline{x}_{\mathcal{B}} \in \mathbb{R}^n$ of the system of linear equations $Ax = b$ with $x_i = 0$ for $i \notin \mathcal{B}$ is called a **basic solution**. The basic set $\mathcal{B}$ such that $\overline{x}_{\mathcal{B}} \geqslant 0$ is called a **feasible basic set** and the solution $\overline{x}_{\mathcal{B}}$ is called a **feasible basic solution**. Otherwise the basic set $\mathcal{B}$ and the basic solution $\overline{x}_{\mathcal{B}}$ are called **infeasible**.

Basic Solution and Basic Feasible Solution

Definition

Let $\mathcal{B}$ be a basic set. The unique solution $\overline{x}_{\mathcal{B}} \in \mathbb{R}^n$ of the system of linear equations $Ax = b$ with $x_i = 0$ for $i \notin \mathcal{B}$ is called a **basic solution**. The basic set $\mathcal{B}$ such that $\overline{x}_{\mathcal{B}} \geqslant 0$ is called a **feasible basic set** and the solution $\overline{x}_{\mathcal{B}}$ is called a **feasible basic solution**. Otherwise the basic set $\mathcal{B}$ and the basic solution $\overline{x}_{\mathcal{B}}$ are called **infeasible**.

Example

Consider a linear programming problem
$c^\mathsf{T} x \longrightarrow min, \; Ax = b, \; x \geqslant 0$ where

$$A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ -1 & 3 & 0 & 1 \end{bmatrix}, \; b = \begin{bmatrix} 8 \\ 3 \end{bmatrix}$$

There are $\binom{4}{2} = 6$ basic sets, i.e. every set of 2 elements is basic.

$$\begin{aligned}
\mathcal{B}_1 &= \{1,2\}, \; \overline{x}_{\mathcal{B}_1} = (3,2,0,0), \\
\mathcal{B}_2 &= \{1,3\}, \; \overline{x}_{\mathcal{B}_2} = (-3,0,14,0), \\
\mathcal{B}_3 &= \{1,4\}, \; \overline{x}_{\mathcal{B}_3} = (4,0,0,7), \\
\mathcal{B}_4 &= \{2,3\}, \; \overline{x}_{\mathcal{B}_4} = (0,1,7,0), \\
\mathcal{B}_5 &= \{2,4\}, \; \overline{x}_{\mathcal{B}_5} = (0,8,0,-21), \\
\mathcal{B}_6 &= \{3,4\}, \; \overline{x}_{\mathcal{B}_6} = (0,0,8,3),
\end{aligned}$$

The sets $\mathcal{B}_1, \mathcal{B}_3, \mathcal{B}_4, \mathcal{B}_6$ are basic feasible, the sets $\mathcal{B}_2, \mathcal{B}_5$ are basic infeasible.

Example

Consider a linear programming problem
$c^\mathsf{T} x \longrightarrow min,\ \ Ax = b,\ \ x \geqslant 0$ where

$$A = \left[\begin{array}{rrrr} 2 & -6 & 1 & 0 \\ -1 & 3 & 0 & 1 \end{array}\right],\ b = \left[\begin{array}{c} 1 \\ 2 \end{array}\right]$$

The set $\mathcal{B} = \{1, 2\}$ is **not** basic because $\det \begin{bmatrix} 2 & -6 \\ -1 & 3 \end{bmatrix} = 0.$

Vertices and the Standard Form

Proposition

Let $A \in M(m \times n; \mathbb{R})$ *be a matrix such that* $r(A) = m$ *and* $m \leqslant n$*. Let* $X \subset \mathbb{R}^n$ *be a convex polytope*

$$X = \{x \in \mathbb{R}^n \mid Ax = b, x_1, \ldots, x_n \geqslant 0\}.$$

Let $p \in X$*. Then*

p *is a vertex of* $X \iff$ *there exists a basic feasible set* $\mathcal{B}$ *such that* $p = \overline{x}_{\mathcal{B}}$.

Vertices and the Standard Form

Proposition

Let $A \in M(m \times n; \mathbb{R})$ be a matrix such that $r(A) = m$ and $m \leqslant n$. Let $X \subset \mathbb{R}^n$ be a convex polytope

$$X = \{x \in \mathbb{R}^n \mid Ax = b, x_1, \ldots, x_n \geqslant 0\}.$$

Let $p \in X$. Then

p is a vertex of $X \iff$ there exists a basic feasible set $\mathcal{B}$ such that $p = \overline{x}_{\mathcal{B}}$.

Remark

This is not one–to–one correspondence. For example, different basic feasible sets $\mathcal{B} = \{1, 4\}$, $\mathcal{B}' = \{2, 4\}$

$$\begin{bmatrix} 1 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix}^T = \begin{bmatrix} 0 \\ 3 \end{bmatrix}, x_1, \ldots, x_4 \geqslant 0$$

give the same vertex $\overline{x}_{\mathcal{B}} = \overline{x}_{\mathcal{B}'} = (0, 0, 0, 3)$.

Vertices and the Standard Form (continued)

Proof.

($\Leftarrow$) Let

$$H_{\leqslant} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid \sum_{i \notin \mathcal{B}} -x_i \leqslant 0\}.$$

Then

$$X \subset H_{\leqslant},$$

and for $p = (p_1, \ldots, x_n) \in X$

$$p \in X \cap H_{\leqslant} \Leftrightarrow p_i = 0 \text{ for } i \notin \mathcal{B} \Leftrightarrow p = \overline{x}_{\mathcal{B}}.$$

Vertices and the Standard Form (continued)

Proof.

($\Rightarrow$) Just a sketch. Let $p = (p_1, \ldots, p_n) \in X$ be an extreme point. Let

$$I = \{i \in \{1, \ldots, n\} \mid p_i > 0\}.$$

Columns c_i for $i \in I$ are linearly independent. Otherwise there exists a $q \in \mathbb{R}^n$ such that $q_i = 0$ for $i \notin I$ such that $p + tq \in X$ for small $|t| < \varepsilon$. It is now enough to observe that $|I| \leqslant m$ (exercise) and, if necessary, pick additional $m - |I|$ linearly independent columns.

Example

Consider a linear programming problem
$c^\top x \longrightarrow min,\ Ax = b,\ x \geqslant 0$ where

$$A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ -1 & 3 & 0 & 1 \end{bmatrix},\ b = \begin{bmatrix} 8 \\ 3 \end{bmatrix}$$

Example

Consider a linear programming problem
$c^{\mathsf{T}}x \longrightarrow min,\ \ Ax = b,\ \ x \geqslant 0$ where

$$A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ -1 & 3 & 0 & 1 \end{bmatrix},\ \ b = \begin{bmatrix} 8 \\ 3 \end{bmatrix}$$

The set $\mathcal{B} = \{3, 4\}$ is basic. The corresponding basic solution $\overline{x}_{\mathcal{B}} = \begin{bmatrix} 0 & 0 & 8 & 3 \end{bmatrix}^{\mathsf{T}}$ is feasible since $\overline{x}_{\mathcal{B}} \geqslant 0$.

Example

Consider a linear programming problem
$c^\mathsf{T} x \longrightarrow min, \;\; Ax = b, \;\; x \geqslant 0$ where

$$A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ -1 & 3 & 0 & 1 \end{bmatrix}, \;\; b = \begin{bmatrix} 8 \\ 3 \end{bmatrix}$$

The set $\mathcal{B} = \{3, 4\}$ is basic. The corresponding basic solution $\overline{x}_{\mathcal{B}} = \begin{bmatrix} 0 & 0 & 8 & 3 \end{bmatrix}^\mathsf{T}$ is feasible since $\overline{x}_{\mathcal{B}} \geqslant 0$. It corresponds to the vertex $(0, 0)$ of a polytope given by the original problem.

Example

Consider a linear programming problem
$c^{\mathsf{T}}x \longrightarrow min,\ Ax = b,\ x \geqslant 0$ where

$$A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ -1 & 3 & 0 & 1 \end{bmatrix},\ b = \begin{bmatrix} 8 \\ 3 \end{bmatrix}$$

The set $\mathcal{B} = \{3, 4\}$ is basic. The corresponding basic solution $\overline{x}_{\mathcal{B}} = \begin{bmatrix} 0 & 0 & 8 & 3 \end{bmatrix}^{\mathsf{T}}$ is feasible since $\overline{x}_{\mathcal{B}} \geqslant 0$. It corresponds to the vertex $(0, 0)$ of a polytope given by the original problem.

The set $\mathcal{B} = \{2, 4\}$ is basic. The corresponding basic solution $\overline{x}_{\mathcal{B}} = \begin{bmatrix} 0 & 8 & 0 & -21 \end{bmatrix}^{\mathsf{T}}$ is infeasible since $\overline{x}_{\mathcal{B}} \not\geqslant 0$.

Example

Consider a linear programming problem
$c^{\mathsf{T}}x \longrightarrow min,\ Ax = b,\ x \geqslant 0$ where

$$A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ -1 & 3 & 0 & 1 \end{bmatrix},\ b = \begin{bmatrix} 8 \\ 3 \end{bmatrix}$$

The set $\mathcal{B} = \{3, 4\}$ is basic. The corresponding basic solution $\overline{x}_{\mathcal{B}} = \begin{bmatrix} 0 & 0 & 8 & 3 \end{bmatrix}^{\mathsf{T}}$ is feasible since $\overline{x}_{\mathcal{B}} \geqslant 0$. It corresponds to the vertex $(0, 0)$ of a polytope given by the original problem.

The set $\mathcal{B} = \{2, 4\}$ is basic. The corresponding basic solution $\overline{x}_{\mathcal{B}} = \begin{bmatrix} 0 & 8 & 0 & -21 \end{bmatrix}^{\mathsf{T}}$ is infeasible since $\overline{x}_{\mathcal{B}} \ngeqslant 0$. The basic set $\mathcal{B} = \{2, 4\}$ is infeasible.

Basic Feasible Solution

Remark

Let $\mathcal{B} = \{i_1, \ldots, i_m\}$ be a basic set. Let

$$[A|b] \xrightarrow[\text{row operations}]{\text{elementary}} [A'|b'],$$

where the columns $i_1, \ldots, i_m$ of A' are equal to

$$\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix},$$

respectively. Let $\overline{x}_{\mathcal{B}} = (x_1, \ldots, x_n)$. Then

$$x_i = 0 \text{ for } i \notin \mathcal{B},$$

$$x_{i_j} = b'_j \text{ for } j = 1, \ldots, m,$$

and $\mathcal{B}$ is feasible if and only if $b' \geqslant 0$.

Example

Consider a linear programming problem
$c^{\mathsf{T}}x \longrightarrow min, \;\; Ax = b, \;\; x \geqslant 0$ where

$$A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ -1 & 3 & 0 & 1 \end{bmatrix}, \;\; b = \begin{bmatrix} 8 \\ 3 \end{bmatrix}$$

Example

Consider a linear programming problem
$c^\intercal x \longrightarrow min,\ \ Ax = b,\ \ x \geqslant 0$ where

$$A = \left[\begin{array}{rrrr} 2 & 1 & 1 & 0 \\ -1 & 3 & 0 & 1 \end{array}\right],\ \ b = \left[\begin{array}{c} 8 \\ 3 \end{array}\right]$$

The set $\mathcal{B} = \{2, 4\}$ is basic.

Example

Consider a linear programming problem
$c^T x \longrightarrow min, \;\; Ax = b, \;\; x \geqslant 0$ where

$$A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ -1 & 3 & 0 & 1 \end{bmatrix}, \;\; b = \begin{bmatrix} 8 \\ 3 \end{bmatrix}$$

The set $\mathcal{B} = \{2, 4\}$ is basic. We compute the basic solution by using elementary row operations on $[A|b]$ to get the 2−nd column equal to $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and the 4−th column equal to $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Example

Consider a linear programming problem
$c^{\mathsf{T}}x \longrightarrow min,\ \ Ax = b,\ \ x \geqslant 0$ where

$$A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ -1 & 3 & 0 & 1 \end{bmatrix},\ \ b = \begin{bmatrix} 8 \\ 3 \end{bmatrix}$$

The set $\mathcal{B} = \{2, 4\}$ is basic. We compute the basic solution by using elementary row operations on $[A|b]$ to get the 2−nd column equal to $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and the 4−th column equal to $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

$$\left[\begin{array}{cccc|c} 2 & 1 & 1 & 0 & 8 \\ -1 & 3 & 0 & 1 & 3 \end{array}\right] \xrightarrow{r_2 - 3r_1} \left[\begin{array}{cccc|c} 2 & 1 & 1 & 0 & 8 \\ -7 & 0 & -3 & 1 & -21 \end{array}\right]$$

Example

Consider a linear programming problem
$c^\top x \longrightarrow min, \;\; Ax = b, \;\; x \geqslant 0$ where

$$A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ -1 & 3 & 0 & 1 \end{bmatrix}, \;\; b = \begin{bmatrix} 8 \\ 3 \end{bmatrix}$$

The set $\mathcal{B} = \{2, 4\}$ is basic. We compute the basic solution by using elementary row operations on $[A|b]$ to get the 2−nd column equal to $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and the 4−th column equal to $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

$$\left[\begin{array}{cccc|c} 2 & 1 & 1 & 0 & 8 \\ -1 & 3 & 0 & 1 & 3 \end{array}\right] \xrightarrow{r_2 - 3r_1} \left[\begin{array}{cccc|c} 2 & 1 & 1 & 0 & 8 \\ -7 & 0 & -3 & 1 & -21 \end{array}\right]$$

Therefore if $x_1 = x_3 = 0$ (non-basic variables) then
$x_2 = 8, \;\; x_4 = -21$ (basic variables).

Example

Consider a linear programming problem
$c^T x \longrightarrow min, \; Ax = b, \; x \geqslant 0$ where

$$A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ -1 & 3 & 0 & 1 \end{bmatrix}, \; b = \begin{bmatrix} 8 \\ 3 \end{bmatrix}$$

The set $\mathcal{B} = \{2, 4\}$ is basic. We compute the basic solution by using elementary row operations on $[A|b]$ to get the 2−nd column equal to $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and the 4−th column equal to $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

$$\left[\begin{array}{cccc|c} 2 & 1 & 1 & 0 & 8 \\ -1 & 3 & 0 & 1 & 3 \end{array}\right] \xrightarrow{r_2 - 3r_1} \left[\begin{array}{cccc|c} 2 & 1 & 1 & 0 & 8 \\ -7 & 0 & -3 & 1 & -21 \end{array}\right]$$

Therefore if $x_1 = x_3 = 0$ (non-basic variables) then
$x_2 = 8, \; x_4 = -21$ (basic variables). Since $x_4 < 0$ the basic solution
$\overline{x}_{\mathcal{B}} = \begin{bmatrix} 0 & 8 & 0 & -21 \end{bmatrix}^T$ is infeasible.

Next Lecture - Simplex Method

We will learn an algorithm, called simplex method, for finding an optimal solution.

Next Lecture - Simplex Method

We will learn an algorithm, called simplex method, for finding an optimal solution. Simplex method starts from a basic feasible set and with each turn moves to another basic feasible set decreasing (possibly) the objective function.

Dual Linear Program

Definition

For given linear programming problem $c^{\mathsf{T}}x \to max, \ Ax \leqslant b$ the **dual linear program** is

$$b^{\mathsf{T}}y \to min, \ A^{\mathsf{T}}y = c, \ y \geqslant 0.$$

The original problem is called **primal** and the latter **dual**.

Dual Linear Program

Definition

For given linear programming problem $c^{\intercal}x \to max, \ Ax \leqslant b$ the **dual linear program** is

$$b^{\intercal}y \to min, \ A^{\intercal}y = c, \ y \geqslant 0.$$

The original problem is called **primal** and the latter **dual**.

Example

The linear programming problem dual to

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix}^{\intercal} x \to max, \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} x \leqslant \begin{bmatrix} 2 \\ 2 \\ 3 \\ 0 \\ 0 \end{bmatrix},$$

is

Dual Linear Program (continued)

Example

$$\begin{bmatrix} 2 \\ 2 \\ 3 \\ 0 \\ 0 \end{bmatrix}^{\mathsf{T}} y \to min, \quad \begin{bmatrix} 1 & 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 & -1 \end{bmatrix} y = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad y \geqslant 0.$$

Dual Linear Program (continued)

Remark

Some authors give a different definition.

For given linear programming problem $c^\intercal x \to max, \ Ax \leqslant b, \ x \geqslant 0$ *the* **dual linear program** *is*

$$b^\intercal y \to min, \ A^\intercal y \geqslant c, \ y \geqslant 0.$$

It is easy to see the definitions are equivalent. For example, in the above setting the primal is equivalent to
$c^\intercal x \to max, \ \left[\begin{array}{c} A \\ \hline -I \end{array}\right] x \leqslant \left[\begin{array}{c} b \\ \hline 0 \end{array}\right]$ *which is dual to*
$\left[\begin{array}{c} b \\ \hline 0 \end{array}\right]^\intercal \left[\begin{array}{c} y \\ z \end{array}\right] \to min, \ \left[\begin{array}{c|c} A^\intercal & -I \end{array}\right] \left[\begin{array}{c} y \\ z \end{array}\right] = c, \ y, z \geqslant 0$ *which in turn is equivalent to (z describes slack variables)*
$b^\intercal y \to min, \ Ay \geqslant c, \ y \geqslant 0.$

Weak Duality Theorem

Proposition

For any feasible (not necessarily basic) solution x of the **primal** *problem and for any feasible (not necessarily basic) solution y of the* **dual** *problem*

$$c^\mathsf{T} x \leqslant b^\mathsf{T} y.$$

Proof.

Note that since $y \geqslant 0$ and $b - Ax \geqslant 0$ then

$$0 \leqslant y^\mathsf{T}(b - Ax) = y^\mathsf{T} b - (A^\mathsf{T} y)^\mathsf{T} x = y^\mathsf{T} b - c^\mathsf{T} x.$$

$\square$

Weak Duality Theorem (continued)

Corollary

$$\begin{matrix}\textit{the } \textbf{primal} \textit{ problem is feasible but}\\ \textit{the objective function attains no maximum}\end{matrix} \implies \textit{the } \textbf{dual} \textit{ problem is infeasible}$$

$$\begin{matrix}\textit{the } \textbf{dual} \textit{ problem is feasible but}\\ \textit{the objective function attains no minimum}\end{matrix} \implies \textit{the } \textbf{primal} \textit{ problem is infeasible}$$

Remark

The converse does not hold in general. For example when

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}, \quad b = \begin{bmatrix} -1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \quad c = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

both primal and dual problems, i.e. $Ax \leqslant b$ *and* $A^{\mathsf{T}} y = c, \; y \geqslant 0$*, are infeasible.*

Strong Duality Theorem

Theorem

$$\begin{array}{c}\textit{x^* is an optimal solution}\\ \textit{of the \textbf{primal} problem}\end{array} \Longrightarrow \begin{array}{c}\textit{there exists y^* an optimal solution}\\ \textit{of the \textbf{dual} problem}\end{array}$$

Moreover

$$c^\mathsf{T} x^* = b^\mathsf{T} y^*.$$

Proof.

Omitted. □

Strong Duality Theorem (continued)

The linear programming problem

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix}^{\mathsf{T}} x \to max, \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} x \leqslant \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}, \quad x \geqslant 0,$$

has the optimal solution $x^* = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, and $c^{\mathsf{T}} x^* = 5$. The dual problem yields the tableaux

$$\left[\begin{array}{ccccc|c} 2 & 2 & 3 & 0 & 0 & 0 \\ \hline 1 & 0 & 1 & -1 & 0 & 2 \\ 0 & 1 & 1 & 0 & -1 & 1 \end{array}\right] \xrightarrow{r_0 - 2r_1 - 2r_2} \left[\begin{array}{ccccc|c} 0 & 0 & -1 & 2 & 2 & -6 \\ \hline 1 & 0 & 1 & -1 & 0 & 2 \\ 0 & 1 & 1 & 0 & -1 & 1 \end{array}\right]$$

For $\mathcal{B} = \{1, 2\}$ we have $s = 3$ and $r = 2$ therefore

$$\xrightarrow[r_1 - r_2]{r_0 + r_2} \left[\begin{array}{ccccc|c} 0 & 0 & 0 & 2 & 1 & -5 \\ \hline 1 & -1 & 0 & -1 & 1 & 1 \\ 0 & 1 & 1 & 0 & -1 & 1 \end{array}\right]$$

Strong Duality Theorem (continued)

The optimal solution of the dual problem is $y^* = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ and

$b^\intercal y^* = 5.$

Hyperplane Separation Theorem (for cones)

It is relatively easy to prove the Strong Duality Theorem using the Hyperplane Separation Theorem for a cone.

Theorem

For some $v_1, \ldots, v_k \in \mathbb{R}^n$ *let*

$$V = \operatorname{cone}(v_1, \ldots, v_k) = \left\{ \sum_{i=1}^{k} t_i v_i \mid t_i \geqslant 0 \right\}.$$

Then $v \notin V$ *is and only if there exits* $d \in \mathbb{R}^n$ *such that* $d^\mathsf{T} v > 0$ *and*

$$d^\mathsf{T} v_i \leqslant 0 \quad \text{for} \quad i = 1, \ldots, k,$$

that is the hyperplane $d^\mathsf{T} x = 0$ *separates* V *(in particular the vectors* v_i*) from vector* v*.*

Proof.

If $v \in V$ and such d exists then $0 < d^\mathsf{T} v = \sum_{i=1}^{k} t_i \left(d^\mathsf{T} v_i\right) \leqslant 0$. We omit the converse. $\square$

Strong Duality Theorem – Proof

Let $x^* \in \mathbb{R}^n$ be an optimal solution of the primal problem, in particular $Ax^* \leqslant b$. Let $I = J(x^*)$ be the set of all active inequalities in $Ax^* \leqslant b$. Denote the rows of $A \in M(m \times n; \mathbb{R})$ by $a_1, \ldots, a_m \in \mathbb{R}^n$. Let $V = \operatorname{cone}(a_i)_{i \in I}$. Then $c \in V$. Otherwise, by the hyperplane separation theorem, there exists $d \in \mathbb{R}^n$ such that $d^\mathsf{T} c > 0$ and $d^\mathsf{T} a_i \leqslant 0$ for $i \in I$. Then for sufficiently small $\varepsilon > 0$ (active constraints are weakened and inactive still hold for sufficiently small $\varepsilon > 0$)

$$A(x^* + \varepsilon d) \leqslant b,$$

$$c^\mathsf{T}(x^* + \varepsilon d) > c^\mathsf{T} x^*.$$

This contradicts optimality of x^* hence $c \in V$, i.e. for $i \in I$ there exists $y_i \geqslant 0$ such that

$$c = \sum_{i \in I} y_i a_i.$$

Let $y^* \in \mathbb{R}^m$ be given by the above y_i's where $y_i = 0$ for $i \notin I$. It turns out that y^* is an optimal solution of the dual problem.

Strong Duality Theorem – Proof (continued)

Obviously $y^* \geqslant 0$. Moreover

$$A^\mathsf{T} y^* = \sum_{i \in I} y_i a_i = c.$$

That is y^* is feasible. The set I indexes all active constraints of the primal problem hence

$$b^\mathsf{T} y^* = \sum_{i \in I} b_i y_i = \sum_{i \in I} (a_i^\mathsf{T} x^*) y_i = \left(\sum_{i \in I} y_i a_i \right)^\mathsf{T} x^* = c^\mathsf{T} x^*.$$

By the Weak Duality, for any feasible y

$$c^\mathsf{T} x^* \leqslant b^\mathsf{T} y,$$

i.e., y^* is an optimal solution for the dual problem.

Complementary Slackness

Proposition

Let x, y be a feasible solutions of the primal and the dual problem respectively, i.e.

$$\begin{cases} c^\intercal x \to max, \\ Ax \leqslant b \end{cases} \quad \text{and} \quad \begin{cases} b^\intercal y \to min, \\ A^\intercal y = c, \\ y \geqslant 0 \end{cases}$$

Then

$$\underset{\text{are optimal solutions}}{x=x^*, y=y*} \iff y_i = 0 \text{ or } a_i x = b_i \text{ for } i = 1, \ldots, m,$$

where $A \in M(m \times n; \mathbb{R})$ and a_i denotes the $i-th$ row of A.

Proof.

By the Weak Duality, for feasible x, y

$$c^\intercal x = y^\intercal Ax \leqslant y^\intercal b.$$

Both solutions are optimal if and only if $y^\intercal Ax = y^\intercal b$. If $y_i > 0$ then $a_i x = b_i$. $\square$

Primal–Dual Method

There exists a method for solving a linear programming problem using any feasible solution of a primal to solve a smaller, related to dual problem and use it to improve the original solution. This is called the **Primal–Dual Method**.

Carathéodory's Theorem

Theorem

For any $v \in \operatorname{cone}(v_1, \dots, v_k)$ *there exist* $1 \leqslant i_1 < i_2 < \dots i_l \leqslant k$ *such that,*

i) *vectors* $v_{i_1}, \dots, v_{i_l}$ *are linearly independent,*

ii) $v \in \operatorname{cone}(v_{i_1}, \dots, v_{i_l})$.

Carathéodory's Theorem

Theorem

For any $v \in \operatorname{cone}(v_1, \ldots, v_k)$ *there exist* $1 \leqslant i_1 < i_2 < \ldots i_l \leqslant k$ *such that,*

i) *vectors* $v_{i_1}, \ldots, v_{i_l}$ *are linearly independent,*

ii) $v \in \operatorname{cone}(v_{i_1}, \ldots, v_{i_l})$.

Corollary

Finitely generated cone is a union of finite number of **symplicial** *cones.*

Carathéodory's Theorem – Proof

Let k be the smallest positive number such that[3] (by changing the indices if necessary)

$$v = t_1 v_1 + \ldots + t_k v_k, \text{ for some } t_1, \ldots, t_k > 0.$$

Assume that $v_1, \ldots, v_k$ are linearly dependent. Then there exist $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$, not all equal to 0, such that

$$\alpha_1 v_1 + \ldots + \alpha_k v_k = 0,$$

where $\alpha_i > 0$ for some i (multiply sidewise by -1 if necessary). Let

$$C = \min\left\{ \frac{t_i}{\alpha_i} \mid \alpha_i > 0 \right\}.$$

Then

$$v = (t_1 - C\alpha_1)v_1 + \ldots + (t_k - C\alpha_k)v_k,$$

where

$$t_i - C\alpha_i \begin{cases} \geqslant t_i & \text{if} \quad C < \frac{t_i}{\alpha_i}, \alpha_i \neq 0 \text{ or } \alpha_i = 0, \\ = 0 & \text{if} \quad C = \frac{t_i}{\alpha_i}, \alpha_i > 0. \end{cases}.$$

[3] This proof and the following ones based on N. Lauritzen, *Lectures on*

Carathéodory's Theorem (continued)

Proposition

Let $V = \operatorname{cone}(v_1, \dots, v_k) \subset \mathbb{R}^n$ *be a finitely generated cone. Then* V *is convex and closed.*

Proof.

By Carathéodory's Theorem it is enough to assume that V is symplicial. Complete $v_1, \dots, v_k$ to a basis $v_1, \dots, v_n$ of $\mathbb{R}^n$ and define linear homeomorphism $\varphi \colon \mathbb{R}^n \to \mathbb{R}^n$ by the condition

$$\varphi(\varepsilon_i) = v_i.$$

Then $V = \varphi(\mathbb{R}^k_{\geqslant 0} \times 0)$ is the image of a closed set.

Convexity is left as an exercise. $\square$

Point Separation

Proposition

Let $X \subset \mathbb{R}^n$ be a non–empty, convex, closed set such that $0 \notin X$. Then there exists a unique $x_0 \in X$ such that

$$\|x_0\| = \inf_{x \in X} \|x\|.$$

Moreover $x_0 \neq 0$.

Point Separation

Proposition

Let $X \subset \mathbb{R}^n$ be a non–empty, convex, closed set such that $0 \notin X$. Then there exists a unique $x_0 \in X$ such that

$$\|x_0\| = \inf_{x \in X} \|x\|.$$

Moreover $x_0 \neq 0$.

Proof.

Without loss of generality one can assume that X bounded (exercise) hence compact. If a continuous function $\|\cdot\|$ attains on X minima at points $x_0, y_0 \in X$ then

$$\left\| \frac{1}{2} x_0 + \frac{1}{2} y_0 \right\| \leqslant \frac{1}{2}\|x_0\| + \frac{1}{2}\|y_0\| = \|x_0\|,$$

and $\frac{1}{2}x_0 + \frac{1}{2}y_0 \in X$ by convexity of X. Triangle inequality becomes equality if and only if $x_0 = ty_0$. As $\|x_0\| = \|y_0\|$ it follows that $t = \pm 1$. Since $0 \notin X$ we have $t = 1$.

Point Separation (continued)

Corollary

Let $X \subset \mathbb{R}^n$ be a non–empty, convex, closed set such that $0 \notin X$. Then there exists an affine hyperplane $H \subset \mathbb{R}^n$ separating (strictly) 0 from X, i.e. if H is given by the equation $d^\mathsf{T} x = c$ then

$$0 = d^\mathsf{T} 0 < c \text{ and } d^\mathsf{T} x > c \text{ for any } x \in X.$$

Point Separation (continued)

Corollary

Let $X \subset \mathbb{R}^n$ be a non–empty, convex, closed set such that $0 \notin X$. Then there exists an affine hyperplane $H \subset \mathbb{R}^n$ separating (strictly) 0 from X, i.e. if H is given by the equation $d^\mathsf{T} x = c$ then

$$0 = d^\mathsf{T} 0 < c \text{ and } d^\mathsf{T} x > c \text{ for any } x \in X.$$

Proof.

Let $x_0 \in X$ be a point as above. It is enough to take $d = x_0$ and $c = \frac{x_0^\mathsf{T} x_0}{2}$, i.e. hyperplane H is given by the formula

$$x_0^\mathsf{T} x = \frac{x_0^\mathsf{T} x_0}{2}.$$

Obviously $0 < \frac{\|x_0\|^2}{2}$ and for $x = x_0$ we have $d^\mathsf{T} x > c$. Assume there exists $x \in H \cap X$, i.e. $x_0^\mathsf{T} x = \frac{x_0^\mathsf{T} x_0}{2}$. Then the segment joining x and x_0 is contained in X.

Point Separation (continued)

Proof.

For $t \in [0,1]$

$$\|x_0\|^2 \leqslant \|(1-t)x_0 + tx\|^2 = (1-t)^2\|x_0\|^2 + 2t(1-t)x_0^\top x + t^2\|x\|^2 =$$

$$= (1-t)^2\|x_0\|^2 + t(1-t)\|x_0\|^2 + t^2\|x\|^2.$$

This is equivalent to

$$0 \leqslant -t\|x_0\|^2 + t^2\|x\|.$$

For $t \in (0,1]$

$$\|x_0\| \leqslant t\|x\|,$$

which contradicts that $0 \notin X$ (as 0 is not an accumulation point of X). $\square$

Point Separation (continued)

Corollary

For any non–empty, convex, closed set such that $X \subset \mathbb{R}^n$ *and* $v \notin X$ *there exists an affine hyperplane* $H \subset \mathbb{R}^n$ *separating (strictly)* v *from* X*, i.e. if* H *is given by the equation* $d^\intercal x = c$ *then*

$$d^\intercal v < c \text{ and } d^\intercal x > c \text{ for any } x \in X.$$

Proof.

Exercise. Consider the set $0 \notin X - v$ which is closed and convex too. $\square$

Hyperplane Separation Theorem (for cones) – Proof

Let $V = \mathrm{cone}(v_1, \ldots, v_k) \subset \mathbb{R}^n$ and let $v \in \mathbb{R}^n$ be a vector such that $v \notin V$. The set V is closed and convex hence there exists a hyperplane

$$H\colon d^\mathsf{T} x = c,$$

such that for any $x \in V$

$$d^\mathsf{T} x < c,$$

and (if necessary replace d, c with $-d, -c$)

$$d^\mathsf{T} v > c.$$

Since $0 \in V$ we have $0 < c$. Since for any $t \geqslant 0$

$$d^\mathsf{T}(tx) = t\,(d^\mathsf{T} x) < c,$$

it follows that $d^\mathsf{T} x \leqslant 0$, in particular, for $i = 1, \ldots, k$

$$d^\mathsf{T} x \leqslant 0.$$

Moreover

$$d^\mathsf{T} v > c > 0.$$

Farkas' Lemma

Corollary (Farkas' Lemma)

For $A \in M(m \times n; \mathbb{R})$, $b \in M(n \times 1; \mathbb{R})$ *exactly one of the following sentences is true*

i) *there exists* $x \in \mathbb{R}^n$ *such that* $Ax = b$, $x \geqslant 0$,

ii) *there exists* $y \in \mathbb{R}^m$ *such that* $A^\intercal y \leqslant 0$ *and* $y^\intercal b > 0$.

Remark

This is essentially reformulation of the Hyperplane Separation Theorem. Point i) says b *lies in the cone* V *generated by columns of* A *and point ii) says the hyperplane* $y^\intercal x = 0$ *separates the cone* V *from point* b. *There exist several equivalent variants of this lemma, for example with inequalities reversed in point ii).*

Remarks

The duality can be used in proofs of some results from combinatorial optimization and other theoretical considerations.

Lagrange Duality

Consider the problem $c^\mathsf{T} x \to max$, $Ax \leqslant b$ where $A \in M(m \times n; \mathbb{R})$ with an optimal solution x^*. For any $\lambda \in \mathbb{R}^m$, $\lambda \geqslant 0$ define the Lagrangian function

$$g(x, \lambda) = c^\mathsf{T} x + \lambda^\mathsf{T}(b - Ax).$$

By definition, for any feasible x

$$g(x, \lambda) \geqslant c^\mathsf{T} x.$$

In particular $g(x^*, \lambda) \geqslant c^\mathsf{T} x^*$. Set (a function possibly attaining infinity as a value)

$$g(\lambda) = \sup_{x \in \mathbb{R}^n} g(x, \lambda).$$

Then

$$g(\lambda) \geqslant c^\mathsf{T} x^*,$$

is an upper bound for the optimal value. Moreover, the lowest upper bound is

$$g^* = \min_{\lambda \geqslant 0} g(\lambda) \geqslant c^\mathsf{T} x^*.$$

Lagrange Duality (continued)

This is equivalent to

$$g^* = \min_{\lambda \geqslant 0} g(\lambda) = \min_{\lambda \geqslant 0} \sup_{x \in \mathbb{R}^n} \left(c^\mathsf{T} x + \lambda^\mathsf{T}(b - Ax)\right) =$$

$$= \min_{\lambda \geqslant 0} \left(\lambda^\mathsf{T} b + \sup_{x \in \mathbb{R}^n} \left(c^\mathsf{T} - \lambda^\mathsf{T} A\right) x \right).$$

If at least one entry of $c^\mathsf{T} - \lambda^\mathsf{T} A$ is non–zero then $g(\lambda) = +\infty$ which gives no finite upper bound. Hence one may restrict the domain of $g(\lambda)$ (as it does not change the minimum) to $\lambda' s$ such that $\lambda \geqslant 0$ and $A^\mathsf{T}\lambda - c = 0$, i.e.

$$g^* = \min_{\substack{\lambda \geqslant 0 \\ A^\mathsf{T}\lambda = c}} b^\mathsf{T}\lambda.$$

This is exactly the dual problem and the Strong Duality Theorem implies that $g^* = c^\mathsf{T} x^*$.

Maximum Matching/Minimum Cover

Let $G = (V, E)$ be an undirected graph.

Definition

A set $M \subset E$ is a **matching** in graph G if for any $e_1, e_2 \in M$ edges e_1, e_2 are not adjacent. A set $M \subset E$ is a **maximum (cardinality) matching** if it is a matching in G and for any other matching E' in G

$$|M'| \leqslant |M|.$$

Definition

A set $C \subset V$ is a **(vertex) cover** in graph G if any edge $e \in E$ has at least one of its vertices in C. A set $C \subset V$ is a **minimum (cardinality) cover** if it is a cover in G and for any other cover C' in G

$$|C| \leqslant |C'|.$$

Maximum Matching/Minimum Cover (continued)

Let $G = (V, E)$ be an undirected graph, where $V = \{v_1, \ldots, v_n\}, E = \{e_1, \ldots, e_m\}$. Let $B = B_G \in M(n \times m; \mathbb{R})$ be the incidence matrix of G. For any subset $C \in V$ let $v_C \in \mathbb{R}^n$ denote a vector with $i-$th coordinate equal to 1 if $v_i \in C$ and equal to 0 otherwise. For any subset $M \in E$ let $e_M \in \mathbb{R}^m$ denote a vector with $i-$th coordinate equal to 1 if $e_i \in M$ and equal to 0 otherwise.

Proposition

Set $M \subset E$ is a matching if and only if $e = e_M \in \{0, 1\}^m$ and

$$Be \leqslant \mathbb{1}_n.$$

Proof.

Components of Be are degrees of vertices $v_1, \ldots, v_n$ in a subgraph formed by edges from M. No two edges in a matching share a vertex. $\square$

Maximum Matching/Minimum Cover (continued)

Proposition

Set $C \subset V$ is a cover if and only if $v = v_C \in \{0,1\}^n$ and

$$B^\intercal v \geqslant \mathbb{1}_m.$$

Proof.

Components of $B^\intercal v$ are equal to either $0, 1$ or 2 (each row of $B^\intercal$ contains exactly two $1's$), which counts how many times the corresponding edge is covered by vertices from C. In a cover each edge should be covered by at least one vertex. $\square$

Maximum Matching/Minimum Cover (continued)

Proposition

An optimal solution of the following problem

$$e = e_M \in \{0,1\}^m,$$

$$\mathbb{1}_n^\top e \to max,$$

$$Be \leqslant \mathbb{1}_n,$$

is a maximum matching.

Proposition

An optimal solution of the following problem

$$v = v_C \in \{0,1\}^n,$$

$$\mathbb{1}_n^\top v \to min,$$

$$B^\top v \geqslant \mathbb{1}_m,$$

is a minimum cover.

Fractional Maximum Matching

Proposition

For any graph G both problems

$$\begin{cases} e = e_M \in \{0,1\}^m, \\ \mathbb{1}_m^\intercal e \to max \\ Be \leqslant \mathbb{1}_n. \end{cases} \quad \text{and} \quad \begin{cases} e \geqslant 0, \\ \mathbb{1}_m^\intercal e \to max \\ Be \leqslant \mathbb{1}_n. \end{cases}$$

have the same optimal **value**, *i.e. the cardinality of maximum matching.*

Proof.

The second problem possibly attains a bigger optimal value as $A \subset B \implies \sup_A f \leqslant \sup_B f$. Optimum value is attained at a vertex (of a feasible set/polytope) of the second problem. That vertex has integral components as it is a (unique) solution of a system of active inequalities in $Be \leqslant \mathbb{1}$ and B is a totally unimodular matrix. For any feasible solution $e = (e_1, \dots, e_n)$ of the second problem $e_1, \dots, e_n \leqslant 1$ and hence $e \in \{0,1\}^n$. An optimal solution of the second problem corresponds to a matching and therefore is also a solution of the first problem. $\square$

Fractional Minimum Cover

Proposition

If graph G has no isolated vertices then both problems

$$\begin{cases} v = v_C \in \{0,1\}^n, \\ \mathbb{1}_n^\intercal v \to min \\ B^\intercal v \geqslant \mathbb{1}_m. \end{cases} \quad and \quad \begin{cases} v \geqslant 0, \\ \mathbb{1}_n^\intercal v \to min \\ B^\intercal v \geqslant \mathbb{1}_m. \end{cases}$$

have the same optimal **value**, *i.e. the cardinality of minimum cover.*

Proof.

The second problem possibly attains a smaller optimal value as $A \subset B \implies \inf_B f \leqslant inf_A f$. As above, components of an optimal solution of the second problem are nonnegative integers. Assume that $v^* = (v_1^*, \dots, v_n^*)$ is an optimal solution of the second problem. If say $v_1^* \geqslant 2$ then $v' = (v_1^* - 1, \dots, v_n^*) \geqslant 0$ and $B^\intercal v' \geqslant \mathbb{1}$ but $\mathbb{1}^\intercal v' < \mathbb{1}^\intercal v^*$ (double vertex is wasteful). Therefore optimal solution of the second problem corresponds to a cover and therefore is also a solution of the first problem. $\square$

König's Theorem

Theorem

Let G be a bipartite (undirected) graph. Then the size of maximum matching is equal to the size of minimum cover.

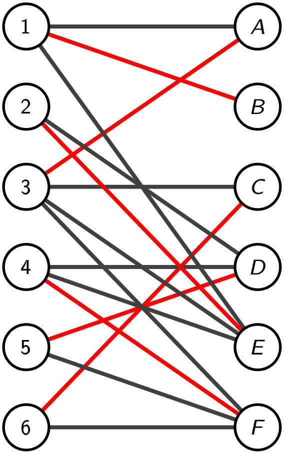
Proof.

By the Strong Duality Theorem both problems attain the same optimal **value**

$$\begin{cases} e \geqslant 0, \\ \mathbb{1}_n e \to max, \\ Be \leqslant \mathbb{1}_n. \end{cases} \qquad \begin{cases} v \geqslant 0, \\ \mathbb{1}_m v \to min \\ B^\intercal v \geqslant \mathbb{1}_m. \end{cases}$$

$\square$

Sample Maximal Matching



6 candidates applied for 6 jobs, first candidate applied for A, B, second candidate for D, E etc. How to hire maximum number of candidates?

Scheduling

Say we have n activities, each activity starts at time p_i, it finishes at time q_i and it brings profit c_i when completed. How to pick non–overlapping activities with maximal profit? Consider the following problem

$$c^\mathsf{T} x \to max,$$

$$x_i + x_j \leqslant 1, \quad \text{for each overlapping activities } i, j,$$

$$x \in \{0, 1\}^n.$$

It has the same optimal solutions as the problem

$$c^\mathsf{T} x \to max,$$

$$x_i + x_j \leqslant 1, \quad \text{for each overlapping activities } i, j,$$

$$x \geqslant 0,$$

as the matrix is an incidence matrix of a bipartite graph (activities i, j are joined by an edge if they overlap) hence totally unimodular.

Fourier–Motzkin Elimination

Theorem

Let $P\colon \mathbb{R}^n \to \mathbb{R}^n$ be the orthogonal projection onto the subspace spanned by the first $n-1$ standard unit vectors, i.e.

$$P(x_1, \ldots, x_{n-1}, x_n) = (x_1, \ldots, x_{n-1}).$$

Let $X \subset \mathbb{R}^n$ be a convex polyhedron. Then $P(X) \subset \mathbb{R}^{n-1}$ is a convex polyhedron.

Fourier–Motzkin Elimination

Theorem

Let $P\colon \mathbb{R}^n \to \mathbb{R}^n$ be the orthogonal projection onto the subspace spanned by the first $n-1$ standard unit vectors, i.e.

$$P(x_1, \ldots, x_{n-1}, x_n) = (x_1, \ldots, x_{n-1}).$$

Let $X \subset \mathbb{R}^n$ be a convex polyhedron. Then $P(X) \subset \mathbb{R}^{n-1}$ is a convex polyhedron.

Proof.

Assume $X \neq \mathbb{R}^n$ is given by the system of inequalities

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & \leqslant b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & \leqslant b_2 \\ \vdots & & \vdots & & \ddots & & \vdots & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & \leqslant b_m \end{cases}$$

Fourier–Motzkin Elimination (continued)

Proof.

Let $N_<, N_0, N_>$ be a partition of the set $\{1, \ldots, m\}$ given by the conditions

$$N_< = \{1 \leqslant i \leqslant m \mid a_{in} < 0\}, N_0 = \{1 \leqslant i \leqslant m \mid a_{in} = 0\},$$

$$N_> = \{1 \leqslant i \leqslant m \mid a_{in} > 0\}.$$

Any $(x_1, \ldots, x_{n-1}) \in P(X)$ satisfies inequality $a_i^\mathsf{T} x \leqslant b_i$ for $i \in N_0$ and a linear combination (with non–negative coefficients) of inequalities $i \in N_<, j \in N_>$

$$a_{jn} \left(\sum_{k=1}^{n} a_{ik} x_k \right) - a_{in} \left(\sum_{k=1}^{n} a_{jk} x_k \right) \leqslant a_{jn} b_i - a_{in} b_j,$$

where x_n is eliminated, i.e.,

$$a_{jn} \left(\sum_{k=1}^{n-1} a_{ik} x_k \right) - a_{in} \left(\sum_{k=1}^{n-1} a_{jk} x_k \right) \leqslant a_{jn} b_i - a_{in} b_j.$$

Fourier–Motzkin Elimination (continued)

Proof.

After dividing by $-1/a_{in}a_{jn}$ this can be rewritten as

$$-\frac{1}{a_{in}}\left(\sum_{k=1}^{n-1} a_{ik}x_k\right) + \frac{1}{a_{jn}}\left(\sum_{k=1}^{n-1} a_{jk}x_k\right) \leqslant -\frac{1}{a_{in}}b_i + \frac{1}{a_{jn}}b_j,$$

that is

$$-\frac{1}{a_{in}}\left(\sum_{k=1}^{n-1} a_{ik}x_k - b_i\right) \leqslant -\frac{1}{a_{jn}}\left(\sum_{k=1}^{n-1} a_{jk}x_k - b_j\right).$$

This implies that

$$\max_{i\in N_<} -\frac{1}{a_{in}}\left(\sum_{k=1}^{n-1} a_{ik}x_k - b_i\right) \leqslant \min_{j\in N_>} -\frac{1}{a_{jn}}\left(\sum_{k=1}^{n-1} a_{jk}x_k - b_j\right).$$

Choosing x_n between those numbers one can see that $(x_1, \ldots, x_n) \in X$. $\square$

Gale's Theorem

Theorem

Let $A \in M(m \times m; \mathbb{R})$, $b \in M(m \times 1; \mathbb{R})$. Then the following conditions are equivalent

i) *the inequality $Ax \leqslant b$ has no solutions,*

ii) *there exists $y \in \mathbb{R}^m$, $y \geqslant 0$ such that $A^\intercal y = 0$, $b^\intercal y < 0$.*

Proof.

Use Fourier–Motzkin elimination to project convex polyhedron X give by $Ax \leqslant b$ onto $0-$dimensional subspace. The image of projection is non–empty is and only if X is non–empty. Each projection amount to multiplying the inequality $Ax \leqslant b$ by some matrix $y \in M(r \times m; \mathbb{R})$, $y \geqslant 0$. The product of such $y's$ gives inequality $y^\intercal A0 \leqslant y^\intercal b$. If X is empty one of the inequalities is $0 \leqslant c$ where $c < 0$. $\square$

Farkas' Lemma Revisited

Corollary (Farkas' Lemma)

For $A \in M(m \times n; \mathbb{R})$, $b \in M(n \times 1; \mathbb{R})$ exactly one of the following sentences is true

i) *there exists $x \in \mathbb{R}^n$ such that $Ax = b$, $x \geqslant 0$,*

ii) *there exists $y \in \mathbb{R}^m$ such that $A^\intercal y \leqslant 0$ and $y^\intercal b > 0$.*

Proof.

As in the previous proof, both conditions cannot be satisfied. If $Ax = b, x \geqslant 0$ has a solution, then $\begin{bmatrix} A \\ -A \\ -I \end{bmatrix} x \leqslant \begin{bmatrix} b \\ -b \\ 0 \end{bmatrix}$ has a solution.

Farkas' Lemma Revisited (continued)

Proof.

By Gale's Theorem for all $\overline{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \geqslant 0$

$$A^\mathsf{T} y_1 - A^\mathsf{T} y_2 - y_3 \neq 0, \quad \text{or} \quad b^\mathsf{T} y_1 - b^\mathsf{T} y_2 \geqslant 0.$$

With $y = y_2 - y_1$ this can be rewritten as

$$A^\mathsf{T} y \neq -y_3, \quad \text{or} \quad b^\mathsf{T} y \leqslant 0,$$

for all $y_3 \geqslant 0$, i.e., for any $y \in \mathbb{R}^m$

$$A^\mathsf{T} y \nleqslant 0, \quad \text{or} \quad y^\mathsf{T} b \leqslant 0,$$

which is exactly the opposite of the condition ii) of Farkas' Lemma.

The converse can be proven in a similar way (exercise). $\square$

Certificate of Infeasibility

Remark

To prove that the problem $Ax = b, x \geqslant 0$ *is infeasible it is enough to find* $y \in \mathbb{R}^m$ *such that* $A^\intercal y \leqslant 0$ *and* $y^\intercal b > 0$. *Therefore any such* y *is called a* **certificate of infesibility**.

Extremal Set Theory

Let S be a finite set and let $\mathcal{A} \subset P(S)$ be a family of subsets of the set S. Let A be a matrix which rows are indicator vectors of subsets in $\mathcal{A}$. Then optimal solutions of the first problems correspond to subsets of $X \subset S$ of maximal cardinality such that $|X \cap A| \leqslant 1$ and the and optimal solutions of the second problem to a subfamilies $\mathcal{Y} \subset \mathcal{A}$ of minimal cardinality such that $\bigcup \mathcal{Y} = S$.

$$\begin{cases} x \in \mathbb{Z}, \\ x \geqslant 0, \\ \mathbb{1}^\intercal x \to max, \\ Ax \leqslant \mathbb{1}. \end{cases} \qquad \begin{cases} y \in \mathbb{Z}, \\ y \geqslant 0, \\ \mathbb{1}^\intercal v \to min \\ A^\intercal y \geqslant \mathbb{1}. \end{cases}$$

Extremal Set Theory (continued)

Optimal solutions of the first problems correspond to subsets of $X \subset S$ of minimal cardinality such that $|X \cap A| \geqslant 1$ (that is X intersects all subsets in the family $\mathcal{A}$) and the and optimal solutions of the second problem to a subfamilies $\mathcal{Y} \subset \mathcal{A}$ of maximal cardinality, containing pairwise disjoint sets.

$$\begin{cases} x \in \mathbb{Z}, \\ x \geqslant 0, \\ \mathbb{1}^\mathsf{T} x \to min, \\ Ax \geqslant \mathbb{1}. \end{cases} \qquad \begin{cases} y \in \mathbb{Z}, \\ y \geqslant 0, \\ \mathbb{1}^\mathsf{T} v \to max \\ A^\mathsf{T} y \leqslant \mathbb{1}. \end{cases}$$

However, for some families $\mathcal{A}$ optimal values of these **integral linear programming problems** may differ. For example let $\mathcal{A} = \{\{1,2\},\{1,3\},\{2,3\}$ and $S = \{1,2,3\}$.

Modeling in Linear Programming[4]

Sometimes it is desirable to impose additional constraints on the optimal solution. This can be achieved by introducing auxiliary variables $t, y_1, \dots, y_n$ (or $t \in \mathbb{R}$ if needed)

$$t \geqslant \max\{x_1, \dots, x_n\} \iff t \geqslant x_i \text{ for } i = 1, \dots, n,$$

$$t \leqslant \min\{x_1, \dots, x_n\} \iff t \leqslant x_i \text{ for } i = 1, \dots, n,$$

$$t \geqslant \max\{a_i^\top x + b_i \mid i = 1, \dots, m\} \iff t \geqslant a_i^\top x + b_i \text{ for } i = 1, \dots, m,$$

in particular

$$t \geqslant |x_i| \iff -t \leqslant x_i \leqslant t,$$

as $|x| = \max\{-x, x\}$.

[4] based on https://docs.mosek.com/modeling-cookbook/index.html

Modeling in Linear Programming (continued)

$$|x_1| + \ldots + |x_n| \leqslant t \iff |x_i| \geqslant y_i \text{ for } i = 1, \ldots, n, \ \sum_{i=1}^{n} y_i = t \iff$$

$$\iff -y_i \leqslant x_i \leqslant y_i \text{ for } i = 1, \ldots, n, \ \sum_{i=1}^{n} y_i = t.$$

The above observation may be used to look (by a heuristic rule) for a sparse solution of the system $Ax = b$ by solving a linear programming problem

$$y_1 + \ldots + y_n \to min,$$

with constraints

$$Ax = b, \quad y \geqslant 0, \quad -y_i \leqslant x_i \leqslant y_i, \quad i = 1, \ldots, n.$$

Sparse Solution of a System of System of Linear Equations – Example

Remark

The previous method quintuples the number of variables. Note that in the generic case an overdetermined system of linear equations $Ax = b$ will have rank A*- sparse solution. Note that in practice large matrices will have low rank which is computationally expensive to estimate.*

Consider the sytem $Ax = b$ where

$$A = \begin{bmatrix} 1 & 0 & 3 & 6 \\ 0 & 1 & 1 & 2 \end{bmatrix} \quad b = \begin{bmatrix} -3 \\ 2 \end{bmatrix}.$$

It does not have 1-sparse solution (why?), i.e. with at most one non–zero entry. To find a solution with the least 1-norm we should add variables $y_1, y_2, y_3, y_4 \geqslant 0$ with inequalities $-y_i \leqslant x_i \leqslant y_i$ for $i = 1, \ldots, 4$ and minimze $y_1 + \ldots + y_4$.

Sparse Solution of a System of System of Linear Equations – Example

The standard form of the above problem is

$$A' = \left[\begin{array}{r|r|c|c|r} A & -A & 0 & 0 & 0 \\ \hline I & -I & I & 0 & -I \\ \hline -I & I & 0 & I & -I \end{array}\right],$$

with $b' = (b, 0, \ldots, 0)$ and $c' = (0, \ldots, 0, 1, 1, 1, 1)$. The first 8 columns of A' correspond to $x_i = x_i^+ - x_i^-$. The second line corresponds to inequalites $x_i + x_i' \leqslant y_i$ i.e., $x_i^+ - x_i^- + x_i' - y_i = 0$. The third line corresponds to inequalites $-y_i \leqslant x_i$ i.e., $-x_i^+ + x_i^- + x_i'' - y_i = 0$. Moreover $x_i^+, x_i^-, x_i', x_i'', y_i \geqslant 0$ for $i = 1, \ldots, 4$.

Sparse Solution of a System of System of Linear Equations – Example

The problem has an optimal solution

$$\bar{x} =$$

$$= \begin{bmatrix} 0 & 3 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 1 & 0 & 6 & 0 & 0 & 0 & 3 & 0 & \frac{1}{2} \end{bmatrix}^T$$

which leads to a 2-sparse solution $x = \left(0, 3, 0, -\frac{1}{2}\right)$.

Modeling in Linear Programming – Sum of m Maximal Components

Proposition

Let $X \subset\in \mathbb{R}^n$ be a section of an n-dimensional cube with a hyperplane $\sum_{i=1}^{n} = m$ where $m \in \{0, 1, \ldots, n\}$, i.e.,

$$X = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid 0 \leqslant x_i \leqslant 1,\ x_1 + \ldots + x_n = m\}.$$

Then vertices of polytope X are of the form

$$(x_1, \ldots, x_n) \quad \textit{where} \quad x_i \in \{0, 1\}, x_1 + \ldots + x_n = m,$$

i.e., sums of m different vectors of the standard basis of $\mathbb{R}^n$.

Modeling in Linear Programming – Sum of m Maximal Components (continued)

Proof.

The constrainst can be rewritten as
$\sum x_i \leqslant m, -\sum x_i \leqslant -m, x_1 \leqslant 1, -x_1 \leqslant 0, \ldots, x_n \leqslant 1, -x_n \leqslant 0$ It is enough to consider submatrices of matrix,

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & \cdots & 1 & 1 & 1 \\ -1 & -1 & -1 & \cdots & \cdots & -1 & -1 & -1 \\ 1 & 0 & 0 & \cdots & \cdots & 0 & 0 & 0 \\ -1 & 0 & 0 & \cdots & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & \cdots & 0 & 0 & 0 \\ 0 & -1 & 0 & \cdots & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \cdots & 0 & 0 & 1 \\ 0 & 0 & 0 & \cdots & \cdots & 0 & 0 & -1 \end{bmatrix}$$

consisting of rows corresponding to active inequalities of rank A. The unique solution is exactly of the required form. Both first rows are always active. $\square$

Modeling in Linear Programming – Sum of m Maximal Components (continued)

Corollary

A solution of the linear programming program $c^\intercal x \to \max$ over X is the sum of m maximal components of vector c.

Modeling in Linear Programming – Sum of m Maximal Components (continued)

If you want to optimize the sum of m maximal components of a point in polyhedron the objective function becomes quadratic. This can by avoided by passing to a dual problem und using the strong duality.

$$d^\mathsf{T} x \to min, Ax = b, x \geqslant 0 \iff b^\mathsf{T} y \to max, A^\mathsf{T} y \leqslant d.$$

Take

$$b = \begin{bmatrix} -m \\ -1 \\ \vdots \\ -1 \end{bmatrix}, \quad A = \left[\begin{array}{c|c} -1 & \\ \vdots & -I \\ -1 & \\ \hline 0 & -I \end{array}\right], \quad d = \begin{bmatrix} -c_1 \\ \vdots \\ -c_n \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Modeling in Linear Programming – Sum of m Maximal Components (continued)

The dual problem becomes

$$mt + \sum_{i=1}^{n} y_i \to min,$$

under the constraints

$$y_i + t \geqslant c_i,$$

$$y_i \geqslant 0,$$

for $i = 1, \ldots, n$.