Linear Programming I

Asst. Prof. Dr.-Ing. Sudchai Boonto

September 1, 2023

Department of Control Systems and Instrumentation Engineering
King Mongkut's Unniversity of Technology Thonburi
Thailand

Objective

Topics:

- Review: linear algebra
- Geometrical Intuition
- Standard form for LPs
- Examples

Most parts of this lecture is taken from Laurent Lessarn , "*Introduction to Optimization*," Lecture Note, University of Wisconsin–Madison.

Matrix basic

Two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ can be multiplied together in two ways. Both are valid matrix multiplications:

- **inner product:** produces a scalar.

$$\mathbf{x}^T\mathbf{y} = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = x_1y_1 + \cdots + x_ny_n$$

  Also called *dot product*. Sometime write $\mathbf{x} \cdot \mathbf{y}$ or $\langle \mathbf{x}, \mathbf{y} \rangle$.

- **Outer product:** produces an $n \times n$ matrix.

$$\mathbf{x}\mathbf{y}^T = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix} = \begin{bmatrix} x_1y_1 & \cdots & x_1y_n \\ \vdots & \ddots & \vdots \\ x_ny_1 & \cdots & x_ny_n \end{bmatrix}$$

Matrix basic

- Matrices and vectors can be stacked and combined to form bigger matrices as long as the dimesions agree, e.g. If $\mathbf{x}_1, \dots, \mathbf{x}_m \in \mathbb{R}^n$, then $\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_m \end{bmatrix} \in \mathbb{R}^{m \times n}$
- Matrices can also be concatenated in blocks. For example

$$\mathbf{Y} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}$$

  if $\mathbf{A}, \mathbf{C}$ have same number of columns, $\mathbf{A}, \mathbf{B}$ have same number of rows, etc.
- Matrix multiplication also works with block matrices:

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{P} \\ \mathbf{Q} \end{bmatrix} = \begin{bmatrix} \mathbf{AP} + \mathbf{BQ} \\ \mathbf{CP} + \mathbf{DQ} \end{bmatrix}$$

  as long as $\mathbf{A}$ has as many columns as $\mathbf{P}$ has rows, etc.

Linear and Affine Functions

- A function $f(x_1, \ldots, x_m)$ is **linear** in the variables $x_1, \ldots, x_m$ if there exist constants $a_1, \ldots, a_m$ such that

$$f(x_1, \ldots, x_m) = a_1 x_1 + \cdots + a_m x_m = a^T x$$

- A function $f(x_1, \ldots, x_m)$ is **affine** in the variables $x_1, \ldots, x_m$ if there exist constants $b, a_1, \ldots, a_m$ such that

$$f(x_1, \ldots, x_m) = a_0 + a_1 x_1 + \cdots + a_m x_m = \mathbf{a}^T \mathbf{x} + \mathbf{b}$$

Example:

- $3x - y$ is linear in (x, y).
- $2xy + 1$ is affine in x and y but not in (x, y).
- $x^2 + y^2$ is not linear or affine.

Linear and Affine Functions

Several linear or affine functions can be combined:

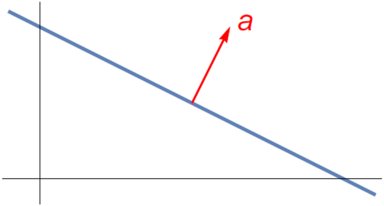
$$\begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n + b_2 \\ a_{21}x_1 + \cdots + a_{2n}x_n + b_2 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n + b_m \end{bmatrix} \implies \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

which can be written simply as $\mathbf{A}\mathbf{x} + \mathbf{b}$. Same definitions apply to:

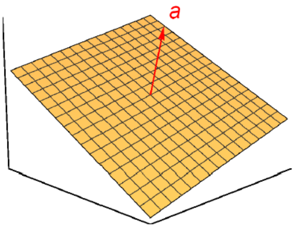
- A vector-valued function $F(\mathbf{x})$ is **linear** in $\mathbf{x}$ if there exists a constant matrix $\mathbf{A}$ such that $F(\mathbf{x}) = \mathbf{A}\mathbf{x}$.
- A vector-valued function $F(\mathbf{x})$ is **affine** in $\mathbf{x}$ if there exists a constant matrix $\mathbf{A}$ and vector $\mathbf{b}$ such that $F(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$.

Geometry of Affine Equation

- The set of points $x \in \mathbb{R}^n$ that satisfies a linear equation $a_1 x_1 + \cdots + a_n x_n = 0$ (or $\mathbf{a}^T \mathbf{x} = 0$) is called a **hyperplane**. The vector $\mathbf{a}$ is *normal* to the hyperplane.
- If the right=hand side is nonzero: $\mathbf{a}^T \mathbf{x} = \mathbf{b}$, the solution set is called an **affine hyperplane.** (It's a shifted hyperplane.)



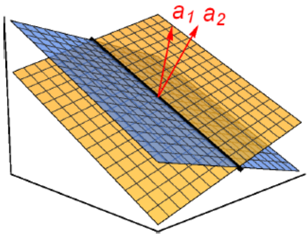
Affine hyperplane in 2D



Affine hyperplane in 3D

Geometry of Affine Equation

- The set of points $\mathbf{x} \in \mathbb{R}^n$ satisfying many linear equations $a_{i1}x_1 + \cdots + a_{im}x_n = 0$ for $i = 1, \ldots, m$ (or $\mathbf{A}\mathbf{x} = 0$) is called a **subspace** (the intersection of many hyperplanes).
- If the right-hand side is nonzero: $\mathbf{A}\mathbf{x} = \mathbf{b}$, the solution set is called an **affine subspace,** (the shifted subspace).



Intersections of affine hyperplanes are affine subspaces.

Geometry of Affine Equation

The **dimension** of a subspace is the number of independent directions it contains. A line has dimension 1, a plane has dimension 2, and so on. (Hyperplanes are subspaces)

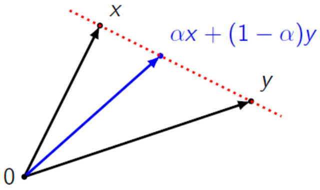
- A hyperplane in $\mathbb{R}^n$ is a subspace of dimension $n - 1$.
- The intersection of k hyperplanes has dimension at least $n - k$ (”at least” because of potential redundancy).

Affine Combinations

If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then the combination

$$\mathbf{w} = \alpha\mathbf{x} + (1-\alpha)\mathbf{y} \qquad \text{for some } \alpha \in \mathbb{R}$$

is called an **affine combination.**



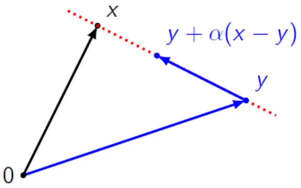
If $\mathbf{Ax} = \mathbf{b}$ and $\mathbf{Ay} = \mathbf{b}$, then $\mathbf{Aw} = \mathbf{b}$. So affine combinations of points in an (affine) subspace also belong to the subspace.

Affine Combinations

If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then the combination

$$\mathbf{w} = \alpha\mathbf{x} + (1 - \alpha)\mathbf{y} \qquad \text{for some } \alpha \in \mathbb{R}$$

is called an **affine combination.**



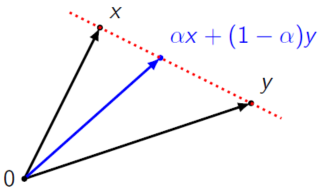
If $\mathbf{Ax} = \mathbf{b}$ and $\mathbf{Ay} = \mathbf{b}$, then $\mathbf{Aw} = \mathbf{b}$. So affine combinations of points in an (affine) subspace also belong to the subspace.

Convex Combinations

If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then the combination

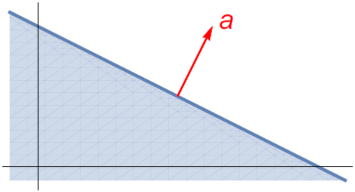
$$\mathbf{w} = \alpha\mathbf{x} + (1-\alpha)\mathbf{y} \quad \text{for some } 0 \leq \alpha \leq 1$$

is called a **convex combination**. It's the line segment that connects $\mathbf{x}$ and $\mathbf{y}$.



Geometry of affine inequalities

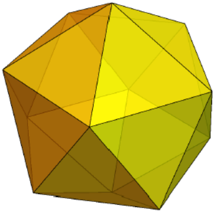
- The set of points $\mathbf{x} \in \mathbb{R}^n$ that satisfies a linear inequality $a_1 x_1 + \cdots + a_n x_n \leq b$ (or $a^T \leq \mathbf{b}$) is called a **halfspace**. The vector a is *normal* to the halfspace and $\mathbf{b}$ shifts it.
- Define $\mathbf{w} = \alpha\mathbf{x} + (1-\alpha)\mathbf{y}$ where $0 \leq \alpha \leq 1$. If $\mathbf{a}^T\mathbf{x} \leq \mathbf{b}$ and $\mathbf{a}^T\mathbf{y} \leq \mathbf{b}$, then $\mathbf{a}^T\mathbf{w} \leq \mathbf{b}$.



Halfspace

Geometry of affine inequalities

- The set of points $\mathbf{x} \in \mathbb{R}^n$ satisfying many linear inequalities $a_{i1}x_1 + \cdots + a_{in}x_n \leq b_i$ for $i = 1, \ldots, m$ (or $\mathbf{Ax} \leq \mathbf{b}$) is called a **polyhedron** (the intersection of many halfspaces). Some sources use the term **polytope** instead.
- As before: let $\mathbf{w} = \alpha\mathbf{x} + (1 - \alpha)\mathbf{y}$ where $0 \leq \alpha \leq 1$. If $\mathbf{Ax} \leq \mathbf{b}$ and $\mathbf{Ay} \leq \mathbf{b}$, then $\mathbf{Aw} \leq \mathbf{b}$.



Intersections of halfspaces are polyhedra.

Linear Programming

- Many engineering optimization problem can be cast as a linear programming (planning or scheduling) problem.
- The Linear Programming (LP) is an optimization problem where the objective function and the constraints are linear functions of the optimization variables.
- Several nonlinear optimization problems can be solved by iteratively solving linearized versions of the original problem.
- In 1947, George Dantzig developed the famous Simplex method.
- Several variations of the Simplex method were introduced after that. Some variations are commercial products, which are secret. They can solve several thousand variables problem in less than one minute.
- The more efficient (most but not always) technique is the interior-point method (IPM).

Linear Programming

- We can put every LP in the form:
  Maximization:

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{R}^n}{\text{maximize}} \quad & \mathbf{c}^T\mathbf{x} \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned}$$

  Minimization:

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{R}^n}{\text{minimize}} \quad & \mathbf{c}^T\mathbf{x} \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned}$$

The linear program

A linear program (LP) is an optimization model with:

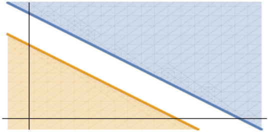
- real-valued variables ($\mathbf{x} \in \mathbb{R}^n$)
- affine objective function ($\mathbf{c}^T\mathbf{x} + \mathbf{d}$), can be minimized or maximized.
- constraints may be:
  - affine equations ($\mathbf{A}\mathbf{x} = \mathbf{b}$)
  - affine inequalities ($\mathbf{A}\mathbf{x} \leq \mathbf{b}$) or ($\mathbf{A}\mathbf{x} \geq \mathbf{b}$)
  - combinations of the above
- individual variables may have:
  - box constraints ($p \leq x_i$, or $x_i \leq q$, or $p \leq x_i \leq q$)
  - no constraints (x_i is unconstrained)

There are many equivalent ways to express the same LP.

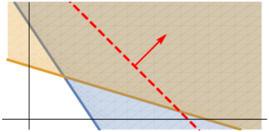
Solutions of an LP

There are exactly three possible cases:

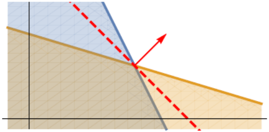
- Model is **infeasible:** there is no $\mathbf{x}$ that satisfies all the constraints. (is the model correct?)
- Model is feasible, but **unbounded:** the cost function can be arbitrarily improved. (forgot a constraint?)
- Model has a solution which occurs **on the boundary** of the set. (there may be many solutions!).



infeasible



unbounded



boundary

Standard form

Top Brass Trophy Company makes large championship trophies for youth athletic leagues. At the moment, they are planning production for fall sports: US football and football. Each US football trophy has a wood base, an engraved plaque, a large brass US football on top, and returns $12 in profit. Football trophies are similar except that a brass football ball is on top, and the unit profit is only $9. Since the US football has an asymmetric shape, its base requires 4 board feet of wood; the football base requires only 2 board feet. At the moment there are 1000 brass US footballs in stock, 1500 football balls, 1750 plaques, and 4800 board feet of wood. What trophies should be produced from these supplies to maximize total profit assuming that all that are made can be sold?

US football | football | both

Standard form

Recipe for building each trophy

	wood	plaques	US footballs	soccer balls	profit
US football	4 ft	1	1	0	$ 12
football	2 ft	1	0	1	$ 9

Quantity of each ingredient in stock

	wood	plaques	US football balls	Football balls
in stock	4800 ft	1750	1000	1500

Linear Programming

$$\begin{aligned} \underset{f,s}{\text{maximize}} \quad & 12f + 9s \\ \text{subject to} \quad & 4f + 2s \leq 4800 \\ & f + s \leq 1750 \\ & 0 \leq f \leq 1000 \\ & 0 \leq s \leq 1500 \end{aligned}$$

Matrix form

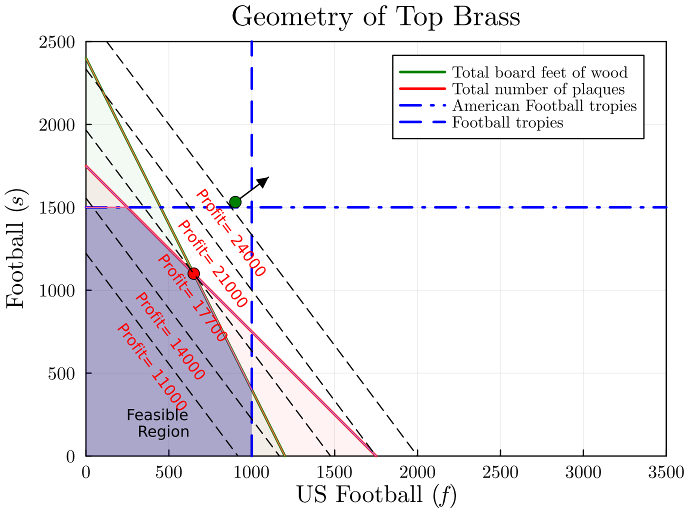
$$\begin{aligned} \underset{\mathbf{x}}{\text{maximize}} \quad & \mathbf{c}^T \mathbf{x} \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned}$$

This is in matrix form, with:

$$\mathbf{A} = \begin{bmatrix} 4 & 2 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 4800 \\ 1750 \\ 1000 \\ 1500 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 12 \\ 9 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} f \\ s \end{bmatrix}$$

Graphical Method

Define $z = 12f + 9s$, where $z =$ profit. Here $s = -\frac{12}{9}f + \frac{z}{9}$



Standard Form

The standard form of the linear programming problem is :

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}^T\mathbf{x} \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned}$$

Example:

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) = 4x_1 - 5x_2 + 3x_3 \\ \text{subject to} \quad & 3x_1 - 2x_2 + 7x_3 = 7 \\ & 8x_1 + 6x_2 + 6x_3 = 5 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 4 \\ -5 \\ 3 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 3 & -2 & 7 \\ 8 & 6 & 6 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 7 \\ 5 \end{bmatrix}$$

Transformation tricks

1. Converting $\operatorname{minimize}$ to $\operatorname{maximize}$ or vice versa

$$\underset{\mathbf{x}}{\operatorname{minimize}}\ \mathbf{c}^T\mathbf{x} = -\underset{\mathbf{x}}{\operatorname{maximize}}\ -\mathbf{c}^T\mathbf{x}$$

2. Reversing inequalities (flip the sign if $\mathbf{b}$ is negative):

$$\mathbf{A}\mathbf{x} \leq \mathbf{b} \iff (-\mathbf{A})\mathbf{x} \geq (-\mathbf{b})$$

3. If a variable has a lower bound other than zeros

$$x \geq 5, \quad \rightarrow \quad x' = x - 5, \quad \rightarrow \quad x' \geq 0$$

4. Inequalities to equalities (add slack variable):

$$f(\mathbf{x}) \leq 0 \iff f(\mathbf{x}) + s = 0 \text{ and } s \geq 0$$

Transformation tricks

5. Unbounded to bounded (add difference):

$$x \in \mathbb{R} \iff u \geq 0,\ v \geq 0,\ \text{ and } \ x = u - v$$

6. Bounded to unbounded (convert to inequality):

$$p \leq x \leq q \iff \begin{bmatrix} 1 \\ -1 \end{bmatrix} x \leq \begin{bmatrix} q \\ -p \end{bmatrix}$$

Standard Form: Example

Consider a linear programming problem:

$$\begin{aligned} \text{maximize} \quad & f(\mathbf{x}) = -5x_1 - 3x_2 + 7x_3 \\ \text{subject to} \quad & 2x_1 + 4x_2 + 6x_3 = 7 \\ & 3x_1 - 5x_2 + 3x_3 \leq 5 \\ & -4x_1 - 9x_2 + 4x_3 \leq -4 \\ & x_1 \geq -2, 0 \leq x_2 \leq 4 \end{aligned}$$

Convert to a minimization problem and make the third constraint to be nonnegative:

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) = 5x_1 + 3x_2 - 7x_3 \\ \text{subject to} \quad & 2x_1 + 4x_2 + 6x_3 = 7 \\ & 3x_1 - 5x_2 + 3x_3 \leq 5 \\ & 4x_1 + 9x_2 - 4x_3 \geq 4 \\ & x_1 \geq -2, 0 \leq x_2 \leq 4 \end{aligned}$$

Standard Form: Example

Transform x_1 to $x_1' = x_1 + 2$, make bound for $x_3 = x_3' - x_3''$, and change $0 \leq x_2 \leq 4$ to be $x_2 \geq 0$ and $x_2 \leq 4$

$$\begin{aligned}
\text{minimize} \quad & f(\mathbf{x}) = 5x_1 + 3x_2 - 7x_3 \\
\text{subject to} \quad & 2x_1 + 4x_2 + 6x_3 = 7 \\
& 3x_1 - 5x_2 + 3x_3 \leq 5 \\
& 4x_1 + 9x_2 - 4x_3 \geq 4 \\
& x_2 \leq 4 \\
& x_1', x_2, x_3', x_3'' \geq 0
\end{aligned}$$

Substitute all things

$$\begin{aligned}
\text{minimize} \quad & f(\mathbf{x}) = 5x_1' + 3x_2 - 7x_3' + 7x_3'' - 10 \\
\text{subject to} \quad & 2x_1' + 4x_2 + 6x_3' - 6x_3'' = 11 \\
& 3x_1' - 5x_2 + 3x_3' - 3x_3'' \leq 11 \\
& 4x_1' + 9x_2 - 4x_3' + 4x_3'' \geq 12 \\
& x_2 \leq 4 \\
& x_1', x_2, x_3', x_3'' \geq 0
\end{aligned}$$

Standard Form: Example

The constant term in the objective function could be remove via a transformation $f'(\mathbf{x}) = f(\mathbf{x}) + 10$. The final step is to add slack and excess variables to convert the general constraints to the equality constraints:

$$
\begin{aligned}
\text{minimize} \quad & f'(\mathbf{x}) = 5x_1' + 3x_2 - 7x_3' + 7x_3'' \\
\text{subject to} \quad & 2x_1' + 4x_2 + 6x_3' - 6x_3'' = 11 \\
& 3x_1' - 5x_2 + 3x_3' - 3x_3'' + s_2 = 11 \\
& 4x_1' + 9x_2 - 4x_3' + 4x_3'' - e_3 = 12 \\
& x_2 + s_4 = 4 \\
& x_1', x_2, x_3', x_3'', s_2, e_3, s_4 \geq 0
\end{aligned}
$$

In matrix form

$$
\begin{aligned}
\text{minimize} \quad & \mathbf{c}^T\mathbf{x} \\
\text{subject to} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \\
& \mathbf{x} \geq 0
\end{aligned}
$$

$$\mathbf{c} = \begin{bmatrix} 5 & 3 & -7 & 7 & 0 & 0 & 0 \end{bmatrix}^T, \quad \mathbf{b} = \begin{bmatrix} 11 & 11 & 12 & 4 \end{bmatrix}^T$$

$$\mathbf{A} = \begin{bmatrix} 2 & 4 & 6 & -6 & 0 & 0 & 0 \\ 3 & -5 & 3 & -3 & 1 & 0 & 0 \\ 4 & 9 & -4 & 4 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

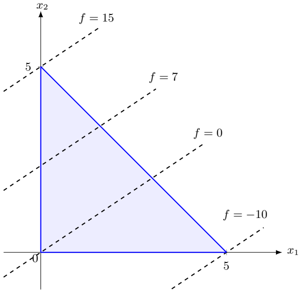
$$\mathbf{x} = \begin{bmatrix} x_1' & x_2 & x_3' & x_3'' & s_2 & e_3 & s_4 \end{bmatrix}^T$$

Standard Form: Example

Put the problem

$$\begin{aligned} \text{minimize} \quad & -2x_1 + 3x_2 \\ \text{subject to} \quad & x_1 + x_2 \leq 5 \\ & \mathbf{x} \geq 0 \end{aligned}$$

in the standard form. Obtain a graphical solution for the original problem and the standard problem.

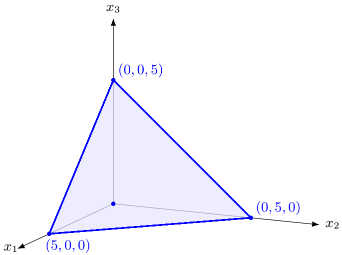


From the figure, it is obvious that the minimum value of the objective function over the feasible region is $f^* = -10$, and the optimal point is $\mathbf{x}^* = \begin{bmatrix} 5 & 0 \end{bmatrix}^T$.

Standard Form: Example

Change it into a standard form by adding a slack variable x_3:

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) = -2x_1 + 3x_2 \\ \text{subject to} \quad & x_1 + x_2 + x_3 = 5 \\ & \mathbf{x} \geq 0 \end{aligned}$$



The minimum of the objective function is at $\mathbf{x}^* = \begin{bmatrix} 5 & 0 & 0 \end{bmatrix}^T$. The optimal solution is the same like the original problem as the slack variable x_3 is set to zero.

How can we find the optimal vertex?

Reference

1. Joaquim R. R. A. Martins, Andrew Ning, "*Engineering Design Optimization*," Cambridge University Press, 2021.
2. Mykel J. kochenderfer, and Tim A. Wheeler, "*Algorithms for Optimization*," The MIT Press, 2019.
3. Ashok D. Belegundu, Tirupathi R. Chandrupatla, "Optimization Concepts and Applications in Engineering," Cambridge University Press, 2019.
4. Laurent Lessarn , "*Introduction to Optimization*," Lecture Note, University of Wisconsin–Madison.