Linear Programming I

Asst. Prof. Dr.-Ing. Sudchai Boonto

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

September 23, 2025

Objective

At the end of this chapter you should be able to:

- Describe, implement, and use Linear Programming
- Understand Linear Programming

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.

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

  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.

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

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

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

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

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

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:

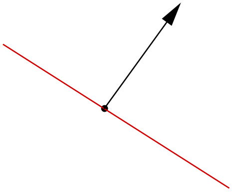
$$\begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n + b_2 \\ a_{21}x_1 + \cdots + a_{2n}x_n + b_2 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n + b_m \end{bmatrix} \implies \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\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

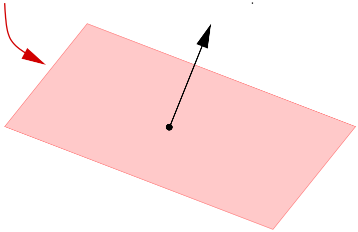
which can be written simply as $\mathbf{A}\mathbf{x} + \mathbf{b}$. Same definitions apply to:

- 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

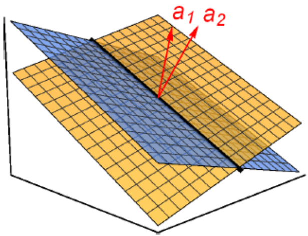
- The set of points $x \in \mathbb{R}^n$ that satisfies a linear equation $a_1x_1 + \cdots + a_nx_n = 0$ (or $\mathbf{a}^T\mathbf{x} = 0$) is called a **hyperplane**. The vector $\mathbf{a}$ is *normal* to the hyperplane.
- If the right=hand side is nonzero: $\mathbf{a}^T\mathbf{x} = \mathbf{b}$, the solution set is called an **affine hyperplane.** (It's a shifted hyperplane.)





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)

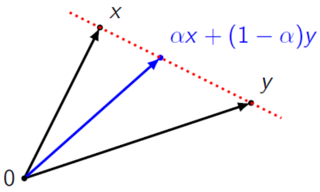
- A hyperplane in $\mathbb{R}^n$ is a subspace of dimension $n-1$.
- The intersection of k hyperplanes has dimension at least $n-k$ ("at least" because of potential redundancy).

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.**



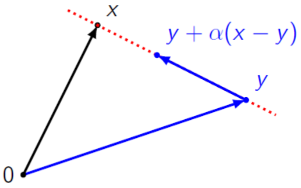
If $\mathbf{Ax} = \mathbf{b}$ and $\mathbf{Ay} = \mathbf{b}$, then $\mathbf{Aw} = \mathbf{b}$. So affine combinations of points in an (affine) subspace also belong to the subspace.

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.**



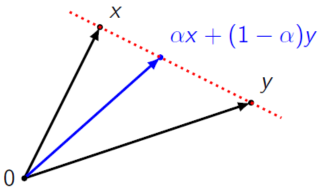
If $\mathbf{Ax} = \mathbf{b}$ and $\mathbf{Ay} = \mathbf{b}$, then $\mathbf{Aw} = \mathbf{b}$. So affine combinations of points in an (affine) subspace also belong to the subspace.

Convex Combinations

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

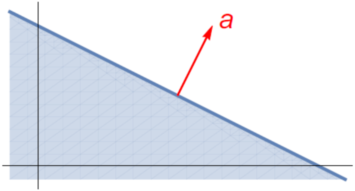
$$\mathbf{w} = \alpha\mathbf{x} + (1-\alpha)\mathbf{y} \quad \text{for some } 0 \leq \alpha \leq 1$$

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



Geometry of affine inequalities

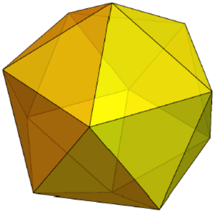
- The set of points $\mathbf{x} \in \mathbb{R}^n$ that satisfies a linear inequality $a_1x_1 + \cdots + a_nx_n \leq b$ (or $a^T \leq \mathbf{b}$) is called a **halfspace**. The vector a is *normal* to the halfspace and $\mathbf{b}$ shifts it.
- Define $\mathbf{w} = \alpha\mathbf{x} + (1-\alpha)\mathbf{y}$ 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 $\mathbf{A}\mathbf{x} \leq \mathbf{b}$) 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 $\mathbf{A}\mathbf{x} \leq \mathbf{b}$ and $\mathbf{A}\mathbf{y} \leq \mathbf{b}$, then $\mathbf{A}\mathbf{w} \leq \mathbf{b}$.



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:

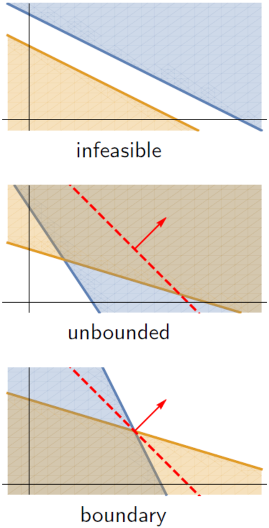
- real-valued variables ($\mathbf{x} \in \mathbb{R}^n$)
- affine objective function ($\mathbf{c}^T\mathbf{x} + \mathbf{d}$), can be minimized or maximized.
- constraints may be:
    - affine equations ($\mathbf{A}\mathbf{x} = \mathbf{b}$)
    - affine inequalities ($\mathbf{A}\mathbf{x} \leq \mathbf{b}$) 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!).



Standard form: Top Brass Data

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: Top Brass Data

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

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

Matrix form

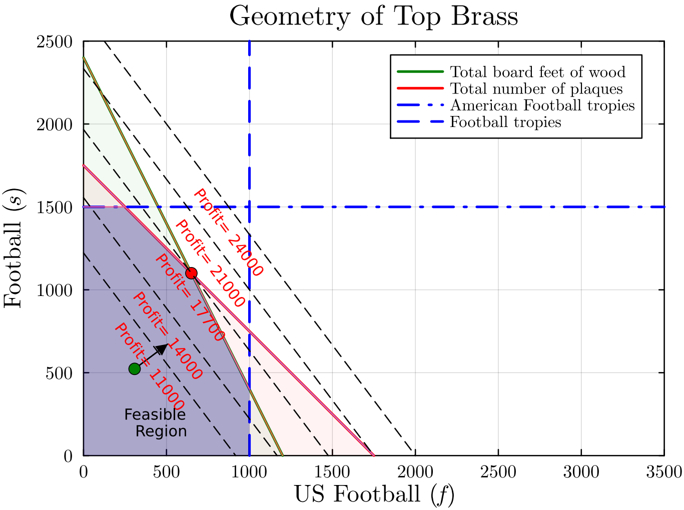
$$
\begin{aligned}
\underset{\mathbf{x}}{\text{maximize}} \quad & \mathbf{c}^T\mathbf{x} \\
\text{subject to} \quad & \mathbf{A}\mathbf{x} \le \mathbf{b} \\
& \mathbf{x} \ge 0
\end{aligned}
$$

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: Example

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: Example

1. Converting minimize to maximize or vice versa

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

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: Example

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

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

Standard Form: Example

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

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

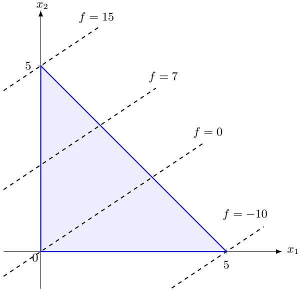
$$\mathbf{x} = \begin{bmatrix} x_1' & x_2 & x_3' & x_3'' & s_2 & e_3 & s_4 \end{bmatrix}^T$$

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.

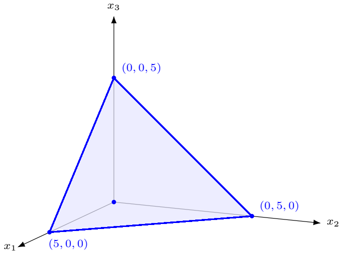


From the figure, it is obvious that the minimum value of the objective function over the feasible region is $f^* = -10$, and the optimal point is $\mathbf{x}^* = \begin{bmatrix} 5 & 0 \end{bmatrix}^T$.

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?

Objective

- Linear programming (LP) problems occur in a diverse range of real-life applications in economic analysis and planning, operations research, computer science, medicine, and engineering.
- These prolems, it is known that nay minima occur at the vertices of the feasible region and can be determined through a “brute-force” or exhaustive approach by evaluating the objective function at all the vertices of the feasible region.
- The number of variables involved in practical LP problem is often vary large and an exhaustive approach would entail a considerable amount of computation.
- In 1947, Dantzig developed a method for solving LP problems known as the *simplex method*. He solved this problem because he came to the class late and thought an unsolved problem on a blackboard was homework.
- Named one of the “Top 10 algorithms of the 20th century” by *Computing in Science & Engineering magazine*. Full list at: `https://www.siam.org/pdf/news/637.pdf`
- The simplex method has been the primary method for solving LP problems since its introduction.

General Constrained Optimization Problem

A general constrained optimization problems:

$$
\begin{aligned}
\underset{\mathbf{x}}{\text{minimize}} \quad & f(\mathbf{x}) \\
\text{subject to} \quad & g_i(\mathbf{x}) \leq 0, & \text{for } i = 1, 2, \ldots, q \\
& h_j(\mathbf{x}) = 0, & \text{for } j = 1, 2, \ldots, p \\
& x_{L_k} \leq x_k \leq x_{U_k}, & \text{for } k = 1, 2, \ldots, n,
\end{aligned}
$$

where x_L and x_U are lower bound and upper bound, respectively.

Definition: Regular point

A point $\mathbf{x}$ is called a *regular point* of the equality constraints if $\mathbf{x}$ satisfies $h_j(\mathbf{x}) = 0$ and column vector $\nabla h(\mathbf{x})$ are linearly independent.

- $\mathbf{x}$ is a regular point of the equality constraints if it is a solution of $h_j(\mathbf{x}) = 0$ and the Jacobian $J = \left[\nabla_{h_1}(\mathbf{x}) \quad \nabla_{h_2}(\mathbf{x}), \ldots, \nabla_{h_p}(\mathbf{x})\right]^T$ has full row rank.

General Constrained Optimization Problem

Consider the equality constraints

$$
\begin{aligned}
-x_1 + x_3 - 1 &= 0 \\
x_1^2 + x_2^2 - 2x_1 &= 0
\end{aligned}
$$

The Jacobian of the constraints is given by

$$
\mathbf{J} = \begin{bmatrix} -1 & 0 & 1 \\ 2x_1 - 2 & 2x_2 & 0 \end{bmatrix}
$$

- The Jacobian has rank 2 except $\mathbf{x} = \begin{bmatrix} 1 & 0 & x_3 \end{bmatrix}^T$
- $\mathbf{x} = \begin{bmatrix} 1 & 0 & x_3 \end{bmatrix}^T$ does not satisfy the second constraint.
- Any points satisfying both constraints is regular.

General Constrained Optimization Problem: Inequality constraints

Consider the constraints

$$g_1(\mathbf{x}) \leq 0, \quad g_2(\mathbf{x}) \leq 0, \quad \cdots \quad g_q(\mathbf{x}) \leq 0$$

- For the feasible point $\mathbf{x}$, these inequalities can be divided into two classes.
- The set of constraints with $g_i(\mathbf{x}) = 0$ are called ***active constraints***.
- The set of constraints with $g_i(\mathbf{x}) < 0$ is called ***inactive constraints***.
- We can convert inequality constraints into equality constraints by adding *slack variable* $s \geq 0$ as

$$\hat{g}_i(\mathbf{x}) = g_i(\mathbf{x}) + s_i = 0$$

General Properties: Formulation of LP problems

The standard-form LP problem:

$$\underset{\mathbf{x}}{\text{minimize}} \quad \mathbf{c}^T\mathbf{x} \tag{4.1a}$$

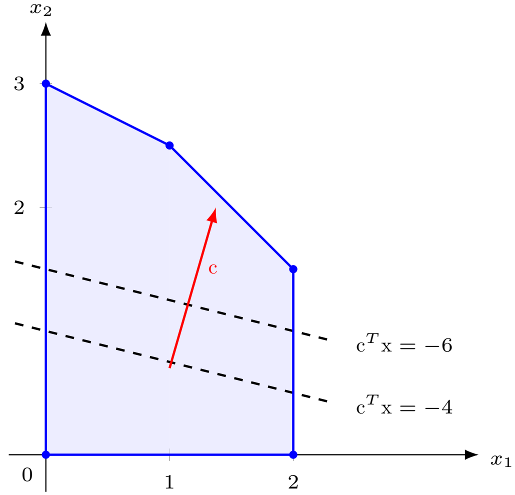
$$\text{subject to} \quad \mathbf{A}\mathbf{x} = \mathbf{b} \tag{4.1b}$$

$$\mathbf{x} \geq 0 \tag{4.1c}$$

where $\mathbf{c} \in \mathbb{R}^{n\times 1}$ with $\mathbf{c} \neq 0$, $\mathbf{A} \in \mathbb{R}^{p\times n}$, and $\mathbf{b} \in \mathbb{R}^{p\times 1}$ are given. We assume that $\mathbf{A}$ is of full row rank, i.e., $\text{rank}(\mathbf{A}) = p$. To be meaningful LP problem, full row rank in $\mathbf{A}$ implies that $p < n$.

- For $n = 2, \mathbf{c}^T\mathbf{x} = \beta$ represents a linea and $\mathbf{c}^T\mathbf{x} = \beta$ for $\beta = \beta_1, \beta_2, \ldots,$ represents a family of parallel lines.
- The normal of these lines is $\mathbf{c}$, and the vector $\mathbf{c}$ is often referred to as the *normal vector* of the objective function.

General Properties: Formulation of LP problem



General Properties: Formulation of LP problems

Another LP problem, which is often encountered in practice, involves minimizing a linear function subject to inequality constraints, i.e.,

$$\underset{\mathbf{x}}{\text{minimize}} \quad \mathbf{c}^T\mathbf{x} \tag{4.2a}$$

$$\text{subject to} \quad \mathbf{A}\mathbf{x} \leq \mathbf{b} \tag{4.2b}$$

where $\mathbf{c} \in \mathbb{R}^{n\times 1}$ with $\mathbf{c} \neq 0$, $\mathbf{A} \in \mathbb{R}^{p\times n}$, and $\mathbf{b} \in \mathbb{R}^{p\times 1}$ are given. This will be referred to as the *alternative-form* LP problem hereafter. If we let

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_p^T \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{bmatrix}, \text{ then } a_i^T\mathbf{x} \leq b_i, \quad \text{for } i = 1, 2, \ldots, p$$

where vector $\mathbf{a}_i$ is the normal of the ith inequality constraint, and $\mathbf{A}$ is usually referred to as the *constraint matrix*.

General Properties: Formulation of LP problems

- by introducing a p-dimensional slack vector variable $\mathbf{s}$, the LP problem can be reformaulated as

$$\mathbf{A}\mathbf{x} + \mathbf{s} = \mathbf{b} \qquad \text{for } \mathbf{s} \geq 0$$

The vector variable $\mathbf{x}$ can be decomposed as

$$\mathbf{x} = \mathbf{x}' - \mathbf{x}'' \quad \text{with} \quad \mathbf{x}' \geq 0 \quad \text{with} \quad \mathbf{x}'' \geq 0$$

Hence if we let

$$\hat{\mathbf{x}} = \begin{bmatrix} \mathbf{x}' & \mathbf{x}'' & \mathbf{s} \end{bmatrix}^T, \quad \hat{\mathbf{c}} = \begin{bmatrix} \mathbf{c} & -\mathbf{c} & 0 \end{bmatrix}^T, \quad \hat{\mathbf{A}} = \begin{bmatrix} \mathbf{A} & -\mathbf{A} & \mathbf{I}_p \end{bmatrix}$$

General Properties: Formulation of LP problems

- The non-standard LP can be reformulated as a standard-form LP problem , the increase in problem size leasd to reduced computational efficiency which can sometimes be a serious problem particularly when the number of inequality constraints is large.
- To solve each form LP will be described separately to enable us to solve each of these problems directly without the need of converting the one form into the other.

General Properties: KKT Conditions

- **Lagrange Multipliers** use to convert a constrained problem into a form such that the derivative test of an unconstrained problem can be applied.
- From the LP:

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

The Lagrangian and the optimal conditions are

$$\begin{aligned} \mathcal{L}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\lambda}) &= \mathbf{c}^T\mathbf{x} + \boldsymbol{\lambda}^T(\mathbf{A}\mathbf{x} - \mathbf{b}) - \boldsymbol{\mu}^T\mathbf{x} \\ \nabla\mathcal{L}_{\mathbf{x}} &= \mathbf{c} + \mathbf{A}^T\boldsymbol{\lambda} - \boldsymbol{\mu} = 0 \\ \nabla\mathcal{L}_{\boldsymbol{\lambda}} &= \mathbf{A}\mathbf{x} - \mathbf{b} = 0 \\ \nabla\mathcal{L}_{\boldsymbol{\mu}} &= \mathbf{x} = 0 \end{aligned}$$

General Properties: KKT Conditions

Theorem: Karush-Kuhn-Tucker (KKT) conditions for standard LP

If $\mathbf{x}^*$ is regular for the constraints that are active at $\mathbf{x}^*$, then it is a global solution of the LP problem in the standard LP if an only if

- $\mathbf{A}\mathbf{x}^* = \mathbf{b}$, (4.3a)
- $\mathbf{x}^* \geq 0$, (4.3b)
- there exist Lagrange multipliers $\boldsymbol{\lambda}^* \in \mathbb{R}^{p\times 1}$ and $\boldsymbol{\mu}^* \in \mathbb{R}^{n\times 1}$ such that $\boldsymbol{\mu}^* \geq 0$ and

$$\mathbf{c} + \mathbf{A}^T\boldsymbol{\lambda}^* - \boldsymbol{\mu}^* = 0 \tag{4.3c}$$

- $\mu_i^* x_i^* = 0$ for $1 \leq i \leq n$ (4.3d)

- The first two condition simply say that solution $\mathbf{x}^*$ must be a feasible point. The constraint matrix $\mathbf{A}$ and vector $\mathbf{c}$ are related through the Lagrange multipliers $\boldsymbol{\lambda}^*$ and $\boldsymbol{\mu}^*$.

General Properties: Formulation of LP problems

- From (4.3a)-(4.3d), in most cases solution $\mathbf{x}^*$ cannot be strictly feasible.
- The term *strictly feasible points* is the points that satisfy the equality constraints with $x_i^* > 0$ for $1 \leq i \leq n$
- From (4.3d), $\boldsymbol{\mu}^*$ must be a zero vector for a strictly feasible point $\mathbf{x}^*$ to be a solution ($x_i^* > 0$). Hence (4.3c) becomes

$$\mathbf{c} + \mathbf{A}^T \boldsymbol{\lambda}^* = 0$$

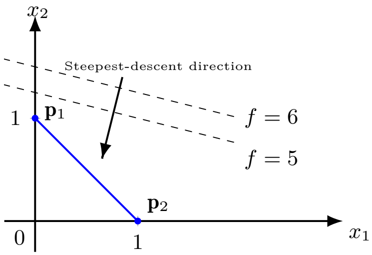
- For **strictly feasible point** to be a minimizer of the standard-form LP problem, the n-dimensional vector $\mathbf{c}$ must lie in the p-dimensional subspace spanned by the p columns of $\mathbf{A}^T$. Since $p < n$, the probability that $\mathbf{c} + \mathbf{A}^T \boldsymbol{\lambda}^* = 0$ is very small.
- Any solutions of the problem are very likely to be located on the **boundary** of the feasible region.

General Properties: Example

Solve the LP problem

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & f(\mathbf{x}) = x_1 + 4x_2 \\ \text{subject to} \quad & x_1 + x_2 = 1 \\ & x_i \geq 0, \quad i = 1, 2 \end{aligned}$$

- The feasible region of the problem is the segment of the line $x_1 + x_2 = 1$ in the first quadrant.
- The dashed lines are contours of the form $f(\mathbf{x}) =$ constant, and the arrow points to the steepest-descent direction of $f(\mathbf{x})$



We have

$$\mathbf{c} = \begin{bmatrix} 1 \\ 4 \end{bmatrix} \text{ and } \mathbf{A}^T = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Since $\mathbf{c}$ and $\mathbf{A}^T$ are linearly independent, $\mathbf{c} = \mathbf{A}^T \boldsymbol{\lambda}^*$ cannot be satisfied and no interior feasible point can be a solution.

General Properties: Example

From the figure, the unique minimizer is $\mathbf{x}^* = \begin{bmatrix} 1 & 0 \end{bmatrix}^T$. At $\mathbf{x}^*$ the constraint $x_1 + x_2 = 1$ and $x_2 = 0$ are active. The Jacobian of these constraints,

$$\mathbf{J} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

is nonsingular, $\mathbf{x}^*$ is a regular point. From $\mu_i^* x_i^* = 0$ and $x_1^* = 1$, then $\mu_1^* = 0$

$$\begin{aligned} \mathbf{c} + \mathbf{A}^T \boldsymbol{\lambda}^* - \boldsymbol{\mu}^* &= 0 \\ \begin{bmatrix} 1 \\ 4 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \boldsymbol{\lambda}^* - \begin{bmatrix} 0 \\ \mu_2^* \end{bmatrix} &= 0 \\ \boldsymbol{\lambda}^* &= -1 \text{ and } \mu_2^* = 3 \end{aligned}$$

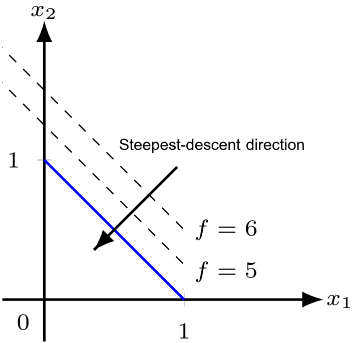
This is confirm that $\mathbf{x}^* = \begin{bmatrix} 1 & 0 \end{bmatrix}^T$ is indeed a global solution (KKT condition).

Note: if the objective function is changed to

$$f(\mathbf{x}) = \mathbf{c}^T\mathbf{x} = 4x_1 + 4x_2$$

We can have

$$\mathbf{c} + \mathbf{A}^T\boldsymbol{\lambda}^* - \boldsymbol{\mu}^* = 0 \quad \Longrightarrow \quad \boldsymbol{\lambda}^* = -4, \boldsymbol{\mu}^* = 0$$



Any feasible point becomes a global solution. The objective function remains constant ($x_1 + x_2 = 1$)in the feasible region, i.e.,

$$f(\mathbf{x}) = 4(x_1 + x_2) = 4, \text{ for } \mathbf{x} \in \mathbb{R}^2$$

General Properties: Alternative form

Consider an alternative LP

$$
\begin{aligned}
\underset{\mathbf{x}}{\text{minimize}} \quad & f(\mathbf{x}) = \mathbf{c}^T\mathbf{x} \\
\text{subject to} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b}
\end{aligned}
$$

Theorem: Necessary and sufficient conditions for a minimum in alternative form LP problem

If $\mathbf{x}^*$ is regular for the constraints in (4.2b) that are active at $\mathbf{x}^*$, then it is a global solution of the problem in (4.2a) if and only if

1. $\mathbf{A}\mathbf{x}^* \leq \mathbf{b}$ (4.4a)
2. there exists a $\boldsymbol{\mu}^* \in \mathbb{R}^{p\times 1}$ such that $\boldsymbol{\mu}^* \geq 0$ and
$$\mathbf{c} + \mathbf{A}^T\boldsymbol{\mu}^* = 0 \tag{4.4b}$$
3. $\boldsymbol{\mu}_i^*(\mathbf{a}_i^T\mathbf{x}^* - b_i) = 0$ for $1 \leq i \leq p$ (4.4c)
where $\mathbf{a}_i^T$ is the ith row of $\mathbf{A}$

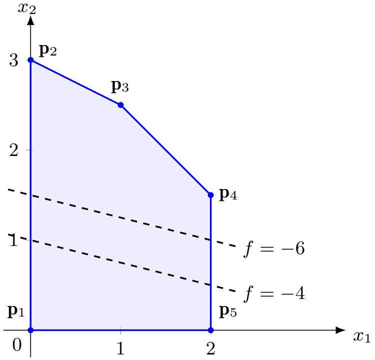
General Properties: Alternative form

- The theorem show that the solutions of the problem must be located on the boundary of the feasible region.
- If $\mathbf{x}^*$ is a strictly feasible point satisfying $\mu_i^*(\mathbf{a}_i^T\mathbf{x} - b_i) = 0$, then $\mathbf{A}\mathbf{x}^* < \mathbf{b}$ and the complementarity condition in (4.4c) implies that $\boldsymbol{\mu}^* = 0$. Hence (4.4b) cannot be satisfied unless $\mathbf{c} = 0$
- If $\mathbf{c} = 0$, it would lead to a meaningless LP problem.
- In other word, any solutions of (4.4a)-(4.4c) can only occur on the boundary of the feasible region.

General Properties: Alternative form Example

Solve the LP problem

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & f(\mathbf{x}) = -x_1 - 4x_2 \\ \text{subject to} \quad & -x_1 \leq 0 \\ & x_1 \leq 2 \\ & -x_2 \leq 0 \\ & x_1 + x_2 - 3.5 \leq 0 \\ & x_1 + 2x_2 - 6 \leq 0 \end{aligned}$$



The five constraints can be expressed as $\mathbf{Ax} \leq \mathbf{b}$ with

$$\mathbf{A} = \begin{bmatrix} -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 0 \\ 2 \\ 0 \\ 3.5 \\ 6 \end{bmatrix}$$

the feasible region is the polygon shown above.

General Properties: Alternative form Example

- The solution cannot be inside the polygon, we consider the five edges of the polygon. At any point of $\mathbf{x}$ on an edge other than the five vertices $\mathbf{p}_i$ only one constraints is active. This mean that only one of the five μ_i's is nonzero.
- At such an x_i, (which is on the edge.), (4.4b) becomes

$$\mathbf{c} = \begin{bmatrix} -1 & -4 \end{bmatrix}^T = -\mu_i \mathbf{a}_i$$

  where $\mathbf{a}_i$ is the transpose of the ith row in $\mathbf{A}$.
- Since each $\mathbf{a}_i$ is linearly independent of $\mathbf{c}$, no μ_i exists that satisfies $\mathbf{c} = -\mu_i \mathbf{a}_i$
- We have five vertices for verification. At $\mathbf{p}_1 = \begin{bmatrix} 0 & 0 \end{bmatrix}^T$, the constraints $-x_1 = 0$, and $-x_2 = 0$ are active. Then $\mathbf{c} = -\mathbf{A}^T \boldsymbol{\mu}$ is

$$\begin{bmatrix} -1 \\ -4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_3 \end{bmatrix}, \Longrightarrow \quad \mu_1 = -1, \mu_3 = -4$$

- Since $\mu_i \leq 0$, then $\mathbf{p}_1$ is not a solution.

General Properties: Alternative form Example

- At the point $\mathbf{p}_2 = \begin{bmatrix} 0 & 3 \end{bmatrix}^T$, the constraints $-x_1 = 0$ and $x_1 + 2x_2 - 6 = 0$ are active. Then $\mathbf{c} = -\mathbf{A}^T \boldsymbol{\mu}$ is

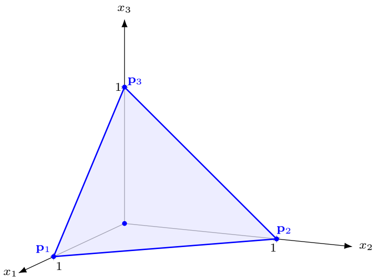
$$\begin{bmatrix} -1 \\ -4 \end{bmatrix} = - \begin{bmatrix} -1 & 0 \\ 1 & 2 \end{bmatrix}^T \begin{bmatrix} \mu_1 \\ \mu_5 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_5 \end{bmatrix}$$

$$\mu_1 = 1, \quad \mu_5 = 2$$

$$\boldsymbol{\mu} = \boldsymbol{\mu}^* = \begin{bmatrix} 1 & 0 & 0 & 0 & 2 \end{bmatrix}^T \geq 0$$

- $\mathbf{p}_2 = \begin{bmatrix} 0 & 3 \end{bmatrix}^T$ is a minimizer, i.e., $\mathbf{x} = \mathbf{x}^* = \mathbf{p}_2$.
- By checking the other vertex point, the point $\mathbf{p}_2$ is the unique solution to the problem.

Facets, Edges, and Vertices



$$x_1 + x_2 + x_3 \leq 1$$

$$x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0$$

$$\mathbf{A}\mathbf{x} \leq \mathbf{b}$$

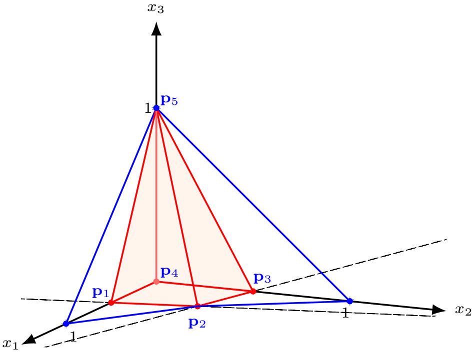
$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix},\ \mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

The polyhedron is three-dimension face, which has four facets, six edges, and four vertices.

A vertex is a feasible point $\mathbf{p}$ at which there exist at least n active constraints which contain n linearly independent constraints where n is the dimension of $\mathbf{x}$. Vertex $\mathbf{p}$ is said to be ***nondegenerate*** if exactly n constraints are active at $\mathbf{p}$ or ***degenerative*** if more than n constraints are active at $\mathbf{p}$.

$\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4$ are nondegenerate vertices.

Facets, Edges, and Vertices



$$x_1 + x_2 + x_3 \le 1$$

$$0.5x_1 + 2x_2 + x_3 \le 1$$

$$x_1 \ge 0, x_2 \ge 0, x_3 \ge 0$$

$$\mathbf{A}\mathbf{x} \le \mathbf{b}$$

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 0.5 & 2 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

- The convex polyhedron has five facets, eight edges, and five vertices.
- At vertex $\mathbf{p}_5$ four constraints are active but since $n = 3$, $\mathbf{p}_5$ is degenerate.
- The other four vertices, namely, $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$, and $\mathbf{p}_4$, are nondegenerate.

Feasible Descent Directions

Start from an initial point, we need to find a better new point:

Theorem: Feasible direction

Let $\boldsymbol{\delta} = \alpha \mathbf{d}$ be a change in $\mathbf{x}$ where α is a positive constant and $\mathbf{d}$ is a direction vector. If Ω is the feasible region and a constant $\hat{\alpha} > 0$ exists such that

$$\mathbf{x} + \alpha \mathbf{d} \in \Omega$$

for all α in the range $0 \leq \alpha \leq \hat{\alpha}$, then $\mathbf{d}$ is said to be a *feasible direction* at point $\mathbf{x}$.

- A vector $\mathbf{d} \in \mathbb{R}^{n \times 1}$ is said to be a ***feasible descent direction*** at a feasible point $\mathbf{x} \in \mathbb{R}^{n \times 1}$ if $\mathbf{d}$ is a feasible direction and the linear objective function strictly decreases along $\mathbf{d}$, i.e., $f(\mathbf{x} + \alpha \mathbf{d}) < f(\mathbf{x})$ for $\alpha > 0$, where $f(\mathbf{x}) = \mathbf{c}^T \mathbf{x}$.
- This implies that

$$f(\mathbf{x} + \alpha \mathbf{d}) = \mathbf{c}^T (\mathbf{x} + \alpha \mathbf{d}) = \mathbf{c}^T \mathbf{x} + \alpha \mathbf{c}^T \mathbf{d}$$

$$\frac{1}{\alpha} \left[f(\mathbf{x} + \alpha \mathbf{d}) - f(\mathbf{x}) \right] = \frac{1}{\alpha} \left[\mathbf{c}^T \mathbf{x} + \alpha \mathbf{c}^T \mathbf{d} - \mathbf{c}^T \mathbf{x} \right] = \mathbf{c}^T \mathbf{d} < 0$$

Feasible Descent Directions

- For LP, we denote $\mathbf{A}_a$ as the matrix whose rows are the rows of $\mathbf{A}$ associated with the constraints active at $\mathbf{x}$.
- We can construct a matrix $\mathbf{A}_a$ the ***active constraint matrix*** at $\mathbf{x}$. If $\mathcal{J} = [j_1, j_2, \ldots, j_k]$ is the set of indices that identify active constraints at $\mathbf{x}$ then

$$\mathbf{A}_a = \begin{bmatrix} \mathbf{a}_{j_1}^T \\ \mathbf{a}_{j_2}^T \\ \vdots \\ \mathbf{a}_{j_k}^T \end{bmatrix}, \quad \mathbf{a}_j^T \mathbf{x} = b_j \qquad \text{for } j \in \mathcal{J}$$

- If $\mathbf{d}$ is a feasible direction, we must have

$$\mathbf{A}_a(\mathbf{x} + \alpha \mathbf{d}) \leq \mathbf{b}_a$$

where $\mathbf{b}_a = \begin{bmatrix} b_{j_1} & b_{j_2} & \cdots & b_{j_k} \end{bmatrix}^T$

Feasible Descent Directions

- $\mathbf{A}_a(\mathbf{x} + \alpha\mathbf{d}) \leq \mathbf{b}_a \Rightarrow \mathbf{A}_a\mathbf{x} + \alpha\mathbf{A}_a\mathbf{d} \leq \mathbf{b}_a$ since $\mathbf{A}_a\mathbf{x} = \mathbf{b}_a$, we must have $\mathbf{A}_a\mathbf{d} \leq 0$
- So the characterizes of a feasible descent direction $\mathbf{d}$ is

$$\mathbf{A}_a\mathbf{d} \leq 0 \quad \text{and} \quad \mathbf{c}^T\mathbf{d} < 0$$

The point $\mathbf{x}^*$ is a solution of the problem in (4.2a) and (4.2b) if and only if there is **no feasible descent directions exist at $\mathbf{x}^*$**.

Theorem: Necessary and sufficient conditions for a minimum in alternative form LP problem

Point $\mathbf{x}^*$ is a solution of the problem in (4.2a) and (4.2b) if and only if it is feasible and

$$\mathbf{c}^T\mathbf{d} \geq 0 \quad \text{for all } \mathbf{d} \text{ with } \mathbf{A}_{a^*}d \leq 0$$

where $\mathbf{A}_{a^*}$ is the active constraint matrix at $\mathbf{x}^*$.

The theorem shows that we could not find the feasible descent directions.

Feasible Descent Directions

- For the standard form LP problem in (4.1a)-(4.1c), a feasible descent direction $\mathbf{d}$ at a feasible point $\mathbf{x}^*$ satisfies the constraints $\mathbf{Ad} = 0$ and $d_j \geq 0$ for $j \in \mathcal{J}_*$ and $\mathbf{c}^T\mathbf{d} \leq 0$, where $\mathcal{J}_* = [j_1, j_2, \ldots, j_k]$ is the set of indices for the constraints in (4.1c) that are active at $\mathbf{x}^*$.

Theorem: Necessary and sufficient conditions for a minimum in standard form LP problem

Point $\mathbf{x}^*$ is a solution of the problem in (4.1a)-(4.1c) if and only if it is feasible and

$$\mathbf{c}^T\mathbf{d} \geq 0 \quad \text{for all } \mathbf{d} \text{ with } \mathbf{d} \in \mathcal{N}(\mathbf{A}) \text{ and } d_j \geq 0 \text{ for } j \in \mathcal{J}_*$$

where $\mathcal{N}(\mathbf{A})$ denotes the null space of $\mathbf{A}$.

$\mathbf{d} \in \mathcal{N}(\mathbf{A})$ means the set of $\mathbf{d}$ such that $\mathbf{Ad} = 0$.

Finding a Vertex (not the LP solution)

- We know that the solution of the LP problems can occur at vertex points. Under some conditions a vertex minimizer always exists.
- We need to have a strategy that can be used to find a minimizer vertex for the LP problem starting with a feasible point $\mathbf{x}_0$.
- In the kth iteration, if the active constraint matrix at $\mathbf{x}_k$, $\mathbf{A}_{a_k}$, has rank n, then $\mathbf{x}_k$ itself is already a vertex.
- Assume that $\mathrm{rank}(\mathbf{A}_{a_k}) < n$. We will generate a feasible point $\mathbf{x}_{k+1}$ such that the active constraint matrix at $\mathbf{x}_{k+1}$, $\mathbf{A}_{a_{k+1}}$, is an *augmented* version of $\mathbf{A}_{a_k}$ with $\mathrm{rank}(\mathbf{A}_{a_{k+1}})$ increased by one.
- $\mathbf{x}_{k+1}$ is a point such that (a) it is feasible, (b) all the constraints that are active at $\mathbf{x}_k$ remain active at $\mathbf{x}_{k+1}$, and (c) there is a new active constraint at $\mathbf{x}_{k+1}$, which was inactive at $\mathbf{x}_k$. A vertex can be identified in a finite number of steps.
- Let $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k$. To make sure that all active constraints at $\mathbf{x}_k$ remain active at $\mathbf{x}_{k+1}$, we must have

$$\mathbf{A}_{a_k}\mathbf{x}_{k+1} = \mathbf{b}_{a_k}$$

Finding a Vertex

- Since $\mathbf{A}_{a_k}\mathbf{x}_k = \mathbf{b}_{a_k}$,

$$\mathbf{A}_{a_k}\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \alpha_k \mathbf{A}_{a_k}\mathbf{d}_k = \mathbf{b}_{a_k}$$

  it follows that $\mathbf{A}_{a_k}\mathbf{d}_k = 0$ (no more feasible direction).

- Since $\operatorname{rank}(\mathbf{A}_{a_k}) < n$, the solutions of $\mathbf{A}_{a_k}\mathbf{d}_k = 0$ form the null space of $\mathbf{A}_{a_k}$ of dimension $n - \operatorname{rank}(\mathbf{A}_{a_k})$. For a fixed $\mathbf{x}_k$ and $\mathbf{d}_k \in \mathcal{N}(\mathbf{A}_{a_k})$. We call an inactive constraint $\mathbf{a}_i^T\mathbf{x}_k - b_i < 0$ *increasing* with respect to $\mathbf{d}_k$ if $\mathbf{a}_i^T\mathbf{d}_k > 0$. (not in the null space.)

- If the ith constraint is an increasing constraint with respect to $\mathbf{d}_k$, then moving from $\mathbf{x}_k$ to $\mathbf{x}_{k+1}$ along $\mathbf{d}_k$, the constraint becomes

$$\begin{aligned} \mathbf{a}_i^T\mathbf{x}_{k+1} - b_i &= \mathbf{a}_i^T(\mathbf{x}_k + \alpha_k\mathbf{d}_k) - b_i \\ &= (\mathbf{a}_i^T\mathbf{x}_k - b_i) + \alpha_k\mathbf{a}_i^T\mathbf{d}_k = 0 \end{aligned}$$

  with $\mathbf{a}_i^T\mathbf{x}_k - b_i < 0$ and $\mathbf{a}_i^T\mathbf{d}_k > 0$.

Finding a Vertex

- A positive α_k that makes the ith constraint active at point $\mathbf{x}_{k+1}$ can be identified as (= 0)

$$\alpha_k = -\frac{\mathbf{a}_i^T \mathbf{x}_k - b_i}{\mathbf{a}_i^T \mathbf{d}_k}$$

- The moving point along $\mathbf{d}_k$ also affects other inactive constraints and care must be taken to ensure that the value of α_k used does not lead to an infeasible $\mathbf{x}_{k+1}$.
- Two problems need to be addressed. (1) how to find a direction $\mathbf{d}_k$ in the null space $\mathcal{N}(\mathbf{A}_{a_k})$ such that there is at least one decreasing inactive constraint with respect to $\mathbf{d}_k$. (2) if $\mathbf{d}_k$ is found, how to determine the step size α_k.
- Given $\mathbf{x}_k$ and $\mathbf{A}_{a_k}$, we can find an inactive constraint whose normal $\mathbf{a}_i^T$ is linearly independent of the rows of $\mathbf{A}_{a_k}$. It follows that the system of equations

$$\begin{bmatrix} \mathbf{A}_{a_k} \\ \mathbf{a}_i^T \end{bmatrix} \mathbf{d}_k = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

has a solution $\mathbf{d}_k$ with $\mathbf{d}_k \in \mathcal{N}(\mathbf{A}_{a_k})$ and $\mathbf{a}_i^T \mathbf{d}_k > 0$.

Finding a Vertex

- The set of indices corresponding to increasing active constraints with respect to $\mathbf{d}_k$ can be defined as

$$\mathcal{I}_k = \left\{ i : \mathbf{a}_i^T \mathbf{x}_k - b_i < 0, \mathbf{a}_i^T \mathbf{d}_k > 0 \right\}$$

- The value of α_k can be determined as the value for which $\mathbf{x}_k + \alpha_k \mathbf{d}_k$ intersects the nearest new constraint. Hence

$$\alpha_k = \min_{i \in \mathcal{I}_k} \left(-\frac{(\mathbf{a}_i^T \mathbf{x}_k - b_i)}{\mathbf{a}_i^T \mathbf{d}_k} \right)$$

- If $i = i^*$ is an index in $\mathcal{I}_k$ that yields the α_k , then it is quite clear that at point $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k$ the active constraint matrix becomes

$$\mathbf{A}_{a_{k+1}} = \begin{bmatrix} \mathbf{A}_{a_k} \\ \mathbf{a}_{i^*}^T \end{bmatrix}$$

where $\mathrm{rank}(\mathbf{A}_{a_{k+1}}) = \mathrm{rank}(\mathbf{A}_{a_k}) + 1$.

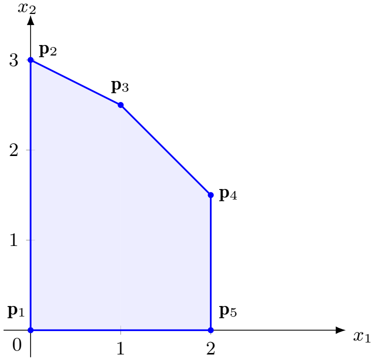
Finding a Vertex

- By repeating the above steps, a feasible point $\mathbf{x}_k$ with $\mathrm{rank}(\mathbf{A}_{a_k}) = n$ will eventually be reached, and point $\mathbf{x}_k$ is then deemed to be a vertex.

Finding a Vertex: Example

Solve the LP problem

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & f(\mathbf{x}) = -x_1 - 4x_2 \\ \text{subject to} \quad & -x_1 \le 0 \\ & x_1 \le 2 \\ & -x_2 \le 0 \\ & x_1 + x_2 - 3.5 \le 0 \\ & x_1 + 2x_2 - 6 \le 0 \end{aligned}$$



The five constraints can be expressed as $\mathbf{A}\mathbf{x} \le \mathbf{b}$ with

$$\mathbf{A} = \begin{bmatrix} -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 0 \\ 2 \\ 0 \\ 3.5 \\ 6 \end{bmatrix}$$

the feasible region is the polygon shown above.

Finding a Vertex: Example

- Starting from point $\mathbf{x}_0 = [1\ 1]^T$, apply the iterative procedure to find a vertex for the LP problem. Since the components of the residual vector at $\mathbf{x}_0$ is

$$\mathbf{r}_0 = \mathbf{A}\mathbf{x}_0 - \mathbf{b} = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} 0 \\ 2 \\ 0 \\ 3.5 \\ 6 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1.5 \\ -3 \end{bmatrix} \text{ are all negative .}$$

- There are no active constraints at $\mathbf{x}_0$. If the first constraint (whose residual is the smallest) is chosen to form

$$\begin{bmatrix} \mathbf{A}_{a_k} \\ \mathbf{a}_i^T \end{bmatrix} \mathbf{d}_k = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \implies \mathbf{a}_1^T \mathbf{d}_0 = \begin{bmatrix} -1 & 0 \end{bmatrix} \mathbf{d}_0 = -d_{01} + (0)d_{02} = 1$$

$$\mathbf{d}_0 = \begin{bmatrix} -1 & 0 \end{bmatrix}^T .$$

Finding a Vertex: Example

- The set $\mathcal{I}_0$ in this case contains only one index, i.e., $\mathcal{I}_0 = \{1\}$.

$$\alpha_0 = -\frac{(\mathbf{a}_1^T \mathbf{x}_0 - b_1)}{\mathbf{a}_1^T \mathbf{d}_0} = -\left(\frac{-1-0}{1}\right) = 1, \quad i^* = 1$$

- Hence

$$\mathbf{x}_1 = \mathbf{x}_0 + \alpha_0 \mathbf{d}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ with } \mathbf{A}_{a_1} = \begin{bmatrix} -1 & 0 \end{bmatrix}.$$

- At point $\mathbf{x}_1$

$$\mathbf{r}_1 = \mathbf{A}\mathbf{x}_1 - \mathbf{b} = \begin{bmatrix} 0 \\ -2 \\ -1 \\ -2.5 \\ -4 \end{bmatrix} \text{ Only } -x_1 \leq 1 \text{ is active.}$$

Finding a Vertex: Example

- The third constraint (whose residual is the smallest and inactive) is chosen to be active:

$$\begin{bmatrix} \mathbf{A}_{a_1} \\ \mathbf{a}_3^T \end{bmatrix} \mathbf{d}_1 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} d_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

  we obtain $\mathbf{d}_1 = \begin{bmatrix} 0 & -1 \end{bmatrix}^T$. It follows that $\mathcal{I}_1 = \{3\}$.

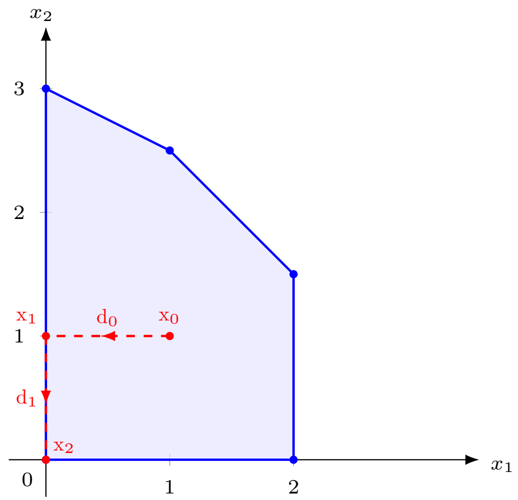
-

$$\alpha_1 = -\frac{\mathbf{a}_3^T \mathbf{x}_1 - b_3}{\mathbf{a}_3^T \mathbf{d}_1} = -\left(\frac{-1-0}{1}\right) = 1 \text{ with } i^* = 3$$

$$\mathbf{x}_2 = \mathbf{x}_1 + \alpha_1 \mathbf{d}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \ \mathbf{A}_{a_2} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

- Since $\operatorname{rank} \mathbf{A}_{a_2} = 2 = n$, $\mathbf{x}_2$ is a vertex.

Finding a Vertex: Example



Finding a Vertex: Example II

Find the vertex for the convex polygon $x_1 + x_2 + x_3 = 1$ such that $x \geq 0$ starting with $\mathbf{x}_0 = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}^T$. (For LP, we need $\mathbf{x} \geq 0$ to be $-\mathbf{x} \leq 0$.)

- We have

$$\mathbf{r}_0 = \mathbf{A}\mathbf{x}_0 - \mathbf{b} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{1}{3} \\ -\frac{1}{3} \\ -\frac{1}{3} \end{bmatrix}$$

  The problem is standard form.

- We select the first inequality constraint (they are equal) so

$$\begin{bmatrix} \mathbf{A}_{a0} \\ \mathbf{a}_2^T \end{bmatrix} \mathbf{d}_0 = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 0 \end{bmatrix} \mathbf{d}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

  Since $d_{01} + d_{02} + d_{03} = 0$ and $-d_{01} = 1$, we have $d_{02} + d_{03} = 1$. Here we select $d_{02} = 1$ and $d_{03} = 0$. Then $\mathbf{d}_0 = \begin{bmatrix} -1 & 1 & 0 \end{bmatrix}^T, \quad \mathcal{I}_0 = \{2\}$.

Finding a Vertex: Example II

- We have

$$\alpha_0 = -\frac{\mathbf{a}_2^T \mathbf{x}_0 - b_1}{\mathbf{a}_2^T \mathbf{d}_0} = -\frac{-\frac{1}{3} - 0}{1} = \frac{1}{3}, \text{ with } i^* = 2$$

$$\mathbf{x}_1 = \mathbf{x}_0 + \alpha_0 \mathbf{d}_0 = \begin{bmatrix} 0 & \frac{2}{3} & \frac{1}{3} \end{bmatrix}^T$$

- $\mathbf{r}_1 = \begin{bmatrix} 0 & 0 & -\frac{2}{3} & -\frac{1}{3} \end{bmatrix}^T$. Choosing the fourth inequality constraint , we have

$$\begin{bmatrix} \mathbf{A}_a \\ \mathbf{a}_4^T \end{bmatrix} \mathbf{d}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \Longrightarrow \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \mathbf{d}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\mathbf{d}_1 = \begin{bmatrix} 0 & 1 & -1 \end{bmatrix}^T, \text{ with } \mathcal{I}_1 = \{4\}$$

$$\alpha_1 = -\frac{\mathbf{a}_4^T \mathbf{x}_1 - b_3}{\mathbf{a}_4^T \mathbf{d}_1} = -\frac{-\frac{1}{3} - 0}{1} = \frac{1}{3}, \text{ with } i^* = 4$$

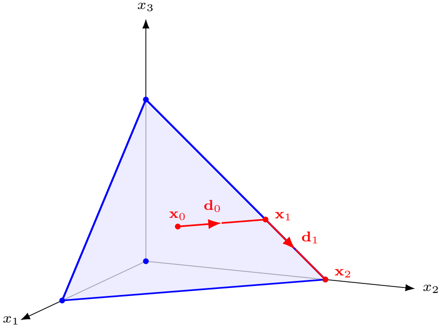
$$\mathbf{x}_2 = \mathbf{x}_1 + \alpha_1 \mathbf{d}_1 = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T$$

Finding a Vertex: Example II

We have

$$\mathbf{A}\mathbf{x}_2 - \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \end{bmatrix} \Longrightarrow \mathbf{A}_{a_2} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \text{ with } \operatorname{rank}(\mathbf{A}_{a_2}) = 3.$$

The point $\mathbf{x}_2$ is a vertex.



Vertex Minimizers

The iterative method for finding a vertex described in the previous section does not involve the objective function $f(\mathbf{x}) = \mathbf{c}^T\mathbf{x}$. The vertex obtained may not be a minimizer. If we start the iterative step at a minimizer, a vertex would eventually be reached without increasing the objective function, which is a vertex minimizer.

Theorem: Existence of a vertex minimizer in alternative-form LP problem

If the minimum of $f(\mathbf{x})$ in the alternative-form LP problem is finite, then there is a vertex minimizer.

Proof: If $\mathbf{x}_0$ is a minimizer, then $\mathbf{x}_0$ is finite and satisfies the condition $\mathbf{A}\mathbf{x}_0 \le \mathbf{b}$ and there exists a $\boldsymbol{\mu}^* \ge 0$ such that

$$\mathbf{c} + \mathbf{A}^T\boldsymbol{\mu}^* = 0 \quad \Rightarrow \quad \mathbf{c} + \mathbf{A}_{a_0}^T\boldsymbol{\mu}_a^* = 0,$$

where $\mathbf{A}_{a_0}$ is the active constraint matrix at $\mathbf{x}_0$ and $\boldsymbol{\mu}_a^*$ is composed of the entries of $\boldsymbol{\mu}^*$ that correspond to the active constraints.

Vertex Minimizers

- If $\mathbf{x}_0$ is not a vertex, the method described in the previous section can be applied to yield a point

$$\mathbf{x}_1 = \mathbf{x}_0 + \alpha_0 \mathbf{d}_0$$

which is closer to a vertex, where $\mathbf{d}_0$ is a feasible direction that satisfies the condition $\mathbf{A}_{a_0}\mathbf{d}_0 = 0$.

- It follows that at $\mathbf{x}_1$ the objective function remains the same as at $\mathbf{x}_0$, i.e.,

$$\begin{aligned} f(\mathbf{x}_1) = \mathbf{c}^T\mathbf{x}_1 = \mathbf{c}^T(x_0 + \alpha_0\mathbf{d}_0) = \mathbf{c}^T\mathbf{x}_0 - \alpha_0\mathbf{c}^T\mathbf{d}_0 \\ = \mathbf{c}^T\mathbf{x}_0 - \alpha_0(\boldsymbol{\mu}_a^*)^T\mathbf{A}_{a_0}\mathbf{d}_0 = \mathbf{c}^T\mathbf{x}_0 = f(\mathbf{x}_0) \end{aligned}$$

- If $\mathbf{x}_1$ is not yet a vertex, then the process is continued to generate minimizers $\mathbf{x}_2, \mathbf{x}_3, \ldots$ until a vertex minimizer is reached.

Vertex Minimizers

- To apply the theorem to the standard form, it follow that

$$\mathbf{c} = -\mathbf{A}^T\boldsymbol{\lambda}^* + \boldsymbol{\mu}^* = -\mathbf{A}^T\boldsymbol{\lambda}^* + \mathbf{I}_0^T\boldsymbol{\mu}_a^*$$

  where $\mathbf{I}_0$ consists of the rows of the $n \times n$ identity matrix that are associated with the inequality constraints $\mathbf{x} \geq 0$ that are active at $\mathbf{x}_0$, and $\boldsymbol{\mu}_a^*$ is composed of the entries of $\boldsymbol{\mu}^*$ that correspond to the active (inequality) constraints.

- At $\mathbf{x}_0$, the active constraint matrix $\mathbf{A}_{a_0}$ is given by

$$\mathbf{A}_{a_0} = \begin{bmatrix} -\mathbf{A} \\ \mathbf{I}_0 \end{bmatrix} \quad \Rightarrow \quad \mathbf{c} = \mathbf{A}_{a_0}^T \begin{bmatrix} \boldsymbol{\lambda}^* \\ \boldsymbol{\mu}_a^* \end{bmatrix}$$

  It can show that the objective function is not change.

Theorem: Existence of a vertex minimizer in standard-form LP problem

If the minimum of $f(\mathbf{x})$ in the standard LP problem is finite, then a vertex minimizer exists.

Simplex Method: Nondegenerate

- If the minimum value of the objective function in the feasible region is finite, then a vertex minimizer exists.
- Let $\mathbf{x}_0$ be a vertex and assume that it is not a minimizer. The simplex method generates an adjacent vertex $\mathbf{x}_1$ with $f(\mathbf{x}_1) < f(\mathbf{x}_0)$ and continues doing so until a vertex minimizer is reached.
- Given a vertex $\mathbf{x}_k$, a vertex $\mathbf{x}_{k+1}$ is *adjacent* to $\mathbf{x}_k$ if $\mathbf{A}_{a_{k+1}}$ is different from $\mathbf{A}_{a_k}$ by only one row.

$$\mathbf{A}_{a_k} = \begin{bmatrix} \mathbf{a}_{j_1}^T \\ \mathbf{a}_{j_2}^T \\ \vdots \\ \mathbf{a}_{j_n}^T \end{bmatrix}, \quad \mathcal{J}_k = \{j_1, j_2, \ldots, j_n\}$$

- If $\mathcal{J}_k$ and $\mathcal{J}_{k+1}$ have exactly $(n-1)$ members, vertices $\mathbf{x}_k$ and $\mathbf{x}_{k+1}$ are adjacent.

Simplex Method: Nondegenerate

- At vertex $\mathbf{x}_k$, the simplex method verifies whether $\mathbf{x}_k$ is a vertex minimizer, and if it is not, it finds an adjacent vertex $\mathbf{x}_{k+1}$ that yields a *reduced* value of the objective function.
- Since a vertex minimizer exists and there is only a finite number of vertices, the method will find the solution using a finite number of iterations.
- Under the nondegeneracy assumption, $\mathbf{A}_{ak}$ is square and nonsingular. There exists a $\boldsymbol{\mu}_k \in \mathbb{R}^{n\times 1}$ such that

$$\mathbf{c} + \mathbf{A}_{a_k}^T \boldsymbol{\mu}_k = 0$$

  Since $\mathbf{x}_k$ is a feasible point, we conclude that $\mathbf{x}_k$ is a vertex minimizer if and only if

$$\boldsymbol{\mu}_k \geq 0$$

  $\mathbf{x}_k$ is not a vertex minimizer if and only if at least one component of $\boldsymbol{\mu}_k$ or $(\boldsymbol{\mu}_k)_l$ is negative.

Simplex Method: Nondegenerate

- Assume that $\mathbf{x}_k$ is not a vertex minimizer and let $(\boldsymbol{\mu}_k)_l < 0$.
- The simplex method finds an edge as a feasible descent direction $\mathbf{d}_k$ that points from $\mathbf{x}_k$ to an adjacent vertex $\mathbf{x}_{k+1}$ given by $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k$.
- A feasible descent direction $\mathbf{d}_k$ is characterized by

$$\mathbf{A}_{a_k}\mathbf{d}_k \le 0 \quad \text{and} \quad \mathbf{c}^T\mathbf{d}_k < 0 \qquad (*)$$

- To find the edge that satisfies $(*)$, we denote the lth coordinate vector (the lth column of the $n \times n$ identify matrix as $\mathbf{e}_l$ and examine vector $\mathbf{d}_k$ that solves the equation $\mathbf{A}_{a_k}\mathbf{d}_k = -\mathbf{e}_l$
- We note that $\mathbf{A}_{a_k}\mathbf{d}_k \le 0$. We have

$$\begin{aligned} \mathbf{c}^T + \boldsymbol{\mu}_k^T\mathbf{A}_{a_k} = 0 \quad \Rightarrow \quad \mathbf{c}^T\mathbf{d}_k + \boldsymbol{\mu}_k^T\mathbf{A}_{ak}\mathbf{d}_k &= 0 \\ \mathbf{c}^T\mathbf{d}_k = -\boldsymbol{\mu}_k^T\mathbf{A}_{ak}\mathbf{d}_k = \boldsymbol{\mu}_k^T\mathbf{e}_l &= (\boldsymbol{\mu}_k)_l < 0 \end{aligned}$$

Hence $\mathbf{d}_k$ satisfies $(*)$ and it is a feasible descent direction.

Simplex Method: Nondegenerate

- For $i \neq l$, $\mathbf{A}_{a_k}\mathbf{d}_k = -\mathbf{e}_l$ implies that

$$\mathbf{a}_{j_i}^T(\mathbf{x}_k + \alpha\mathbf{d}_k) = \mathbf{a}_{j_i}^T\mathbf{x}_k + \alpha\mathbf{a}_{j_i}^T\mathbf{d}_k = b_{j_i}$$

- There are exactly $n-1$ constraints that are active at $\mathbf{x}_k$ and remain active at $\mathbf{x}_k + \alpha\mathbf{d}_k$. This means that $\mathbf{x}_k + \alpha\mathbf{d}_k$ with $\alpha > 0$ is an edge that connects $\mathbf{x}_k$ to an adjacent vertex $\mathbf{x}_{k+1}$ with $f(\mathbf{x}_k + 1) < f(\mathbf{x}_k)$. The right step size α_k can be identified as

$$\alpha_k = \min_{i \in \mathcal{I}_k} \left(\frac{-(\mathbf{a}_i^T\mathbf{x}_k - b_i)}{\mathbf{a}_i^T\mathbf{d}_k} \right) = \frac{-(\mathbf{a}_{i*}^T\mathbf{x}_k - b_{i*})}{\mathbf{a}_{i*}^T\mathbf{d}_k}$$

  where $\mathcal{I}_k$ contains the indices of the constraints that are inactive at $\mathbf{x}_k$ with $\mathbf{a}_i^T\mathbf{d}_k > 0$.
- The vertex $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k\mathbf{d}_k$. Then at $\mathbf{x}_{k+1}$ the i^*th constraint becomes active.

Simplex Method: Nondegenerate

- Substituting j_lth constraint with the active i^*th constraint in $\mathbf{A}_{a_{k+1}}$, there are exactly n active constraints at $\mathbf{x}_{k+1}$ and $\mathbf{A}_{a_{k+1}}$ is given by

$$\mathbf{A}_{a_{k+1}} = \begin{bmatrix} \mathbf{a}_{j_1}^T \\ \vdots \\ \mathbf{a}_{j_{l-1}}^T \\ \mathbf{a}_{i^*}^T \\ \mathbf{a}_{j_{l+1}}^T \\ \vdots \\ \mathbf{a}_{j_n}^T \end{bmatrix}$$

- The index set is given by

$$\mathcal{J}_{k+1} = \{j_1, \ldots, j_{l-1}, i^*, j_{l+1}, \ldots, j_n\}$$

Simplex Method: Nondegenerate

Algorithm: Simplex algorithm for the alternative-form LP problem, nondegenerate vertices

1. Input vertex $\mathbf{x}_0$, and form $\mathbf{A}_{a_0}$ and $\mathcal{J}_0$. Set $k = 0$
2. Solve $\mathbf{A}_{a_k}^T \boldsymbol{\mu}_k = -\mathbf{c}$ for $\boldsymbol{\mu}_k$. If $\boldsymbol{\mu}_k \geq 0$, stop ($\mathbf{x}_k$ is a vertex minimizer): otherwise, select the index l that corresponds to the most negative component in $\boldsymbol{\mu}_k$.
3. Solve $\mathbf{A}_{a_k} \mathbf{d}_k = -\mathbf{e}_l$, where $\mathbf{e}_l$ is a unit vector at l index for $\mathbf{d}_k$.
4. Compute the residual vector $\mathbf{r}_k = \mathbf{A}\mathbf{x}_k - \mathbf{b} = (r_i)_{i=1}^p$ If the index set

$$\mathcal{I}_k = \{i : r_i < 0 \text{ and } \mathbf{a}_i^T \mathbf{d}_k > 0\} \text{ is empty, stop}$$

(The objective function tends to $-\infty$ in the feasible region);

Algorithm: Simplex algorithm for the alternative-form LP problem, nondegenerate vertices cont.

4. (cont.) otherwise, compute

$$\alpha_k = \min_{i \in \mathcal{I}_k} \left(\frac{-r_i}{\mathbf{a}_i^T \mathbf{d}_k} \right)$$

and record the index i^* with $\alpha_k = -r_{i^*}/(\mathbf{a}_{i^*}^T \mathbf{d}_k)$. **Note:** $\mathbf{a}_i^T$ is the row of $\mathbf{A}_{a_k}$ such that $r_i < 0$.

5. Set $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{d}_k$. Update $\mathbf{A}_{a_{k+1}}$ and $\mathcal{J}_{k+1}$ using

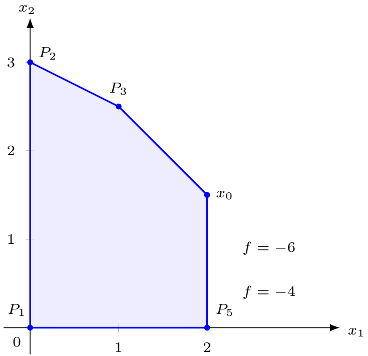
$$\mathbf{A}_{a_{k+1}} = \begin{bmatrix} \mathbf{a}_{j1} & \cdots & \mathbf{a}_{j_{l-1}} & \mathbf{a}_{i^*} & \mathbf{a}_{j_{l+1}} & \cdots & \mathbf{a}_{j_n} \end{bmatrix}^T$$
$$\mathcal{J}_{k+1} = \{j_1, \ldots, j_{l-1}, i^*, j_{l+1}, \ldots, j_n\}$$

Set $k = k + 1$ and repeat from Step 2.

Simplex Method: Nondegenerate Example

Solve the LP problem with initial vertex $x_0 = \begin{bmatrix} 2 & 1.5 \end{bmatrix}^T$

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f(\mathbf{x}) = -x_1 - 4x_2 \\ \text{subject to} \quad & -x_1 \le 0 \\ & x_1 \le 2 \\ & -x_2 \le 0 \\ & x_1 + x_2 - 3.5 \le 0 \\ & x_1 + 2x_2 - 6 \le 0 \end{aligned}$$



The five constraints can be expressed as $\mathbf{A}\mathbf{x} \le \mathbf{b}$ with

$$\mathbf{A} = \begin{bmatrix} -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 0 \\ 2 \\ 0 \\ 3.5 \\ 6 \end{bmatrix}$$ the feasible region is the polygon shown above.

Simplex Method: Nondegenerate Example

- With $\mathbf{x}_0$ the second and fourth constraints are active and hence

$$\mathbf{A}_{a_0} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad \mathcal{J} = \{2, 4\}$$

- Solving $\mathbf{A}_{a_0}^T \boldsymbol{\mu}_0 = -\mathbf{c}$ for $\boldsymbol{\mu}_0$ where $\mathbf{c} = \begin{bmatrix} -1 & -4 \end{bmatrix}^T$, we obtain $\boldsymbol{\mu}_0 = \begin{bmatrix} -3 & 4 \end{bmatrix}^T$. Since $\boldsymbol{\mu}_{0_1}$ is negative, $\mathbf{x}_0$ is not a minimizer and $l = 1$. Next we solve

$$\mathbf{A}_{a_0} \mathbf{d}_0 = -\mathbf{e}_1 \quad \Rightarrow \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \mathbf{d}_0 = -\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\mathbf{d}_0 = \begin{bmatrix} -1 & 1 \end{bmatrix}^T$$

- The residual vector at $\mathbf{x}_0$ is given by

$$\mathbf{r}_0 = \mathbf{A}\mathbf{x}_0 - \mathbf{b} = \begin{bmatrix} -2 & 0 & -1.5 & 0 & -1 \end{bmatrix}^T$$

Simplex Method: Nondegenerate Example

- $\mathbf{r}_0$ shows that the first, third, and fifth constrains are inactive at $\mathbf{x}_0$.

$$\begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_3^T \\ \mathbf{a}_5^T \end{bmatrix} \mathbf{d}_0 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \qquad \mathcal{I}_0 = \{1, 5\}$$

$$\alpha_0 = \min \left(\frac{-r_{01}}{\mathbf{a}_1^T \mathbf{d}_0}, \frac{-r_{05}}{\mathbf{a}_5^T \mathbf{d}_0} \right) = 1, \qquad i^* = 5$$

$$\mathbf{x}_1 = \mathbf{x}_0 + \alpha_0 \mathbf{d}_0 = \begin{bmatrix} 2 \\ 1.5 \end{bmatrix} + 1 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2.5 \end{bmatrix}$$

- Since $l = 1$, we have (by swapping $\mathbf{a}_1^T$ and $\mathbf{a}_5^T$),

$$\mathbf{A}_{a1} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \text{ and } \mathcal{J}_1 = \{5, 4\}$$

- End of the first iteration.

Simplex Method: Nondegenerate Example

- The second iteration starts by solving $\mathbf{A}_{a_1}^T \boldsymbol{\mu}_1 = -\mathbf{c}$ for $\boldsymbol{\mu}_1$

$$\begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \boldsymbol{\mu}_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix} \Longrightarrow \boldsymbol{\mu}_1 = (-1) \begin{bmatrix} 1 & -1 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$$

Then the point $\mathbf{x}_1$ is not a minimizer and $l = 2$.

- By solving $\mathbf{A}_{a_1} \mathbf{d}_1 = -\mathbf{e}_2$

$$\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \mathbf{d}_1 = - \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \Longrightarrow \quad \mathbf{d}_1 = (-1) \begin{bmatrix} 1 & -2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

- We compute the residual vector at $\mathbf{x}_1$ as

$$\mathbf{r}_1 = \mathbf{A}\mathbf{x}_1 - \mathbf{b} = \begin{bmatrix} -1 & -1 & -2.5 & 0 & 0 \end{bmatrix}^T$$

It indicates that the first three constraints are inactive at $\mathbf{x}_1$.

Simplex Method: Nondegenerate Example

- By evaluating

$$\begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \mathbf{a}_3^T \end{bmatrix} \mathbf{d}_1 = \begin{bmatrix} -1 & 0 \\ 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ -1 \end{bmatrix}, \quad \mathcal{I}_1 = \{1\}, \quad \alpha_1 = \frac{-r_{11}}{\mathbf{a}_1^T \mathbf{d}_1} = \frac{1}{2}, i^* = 1$$

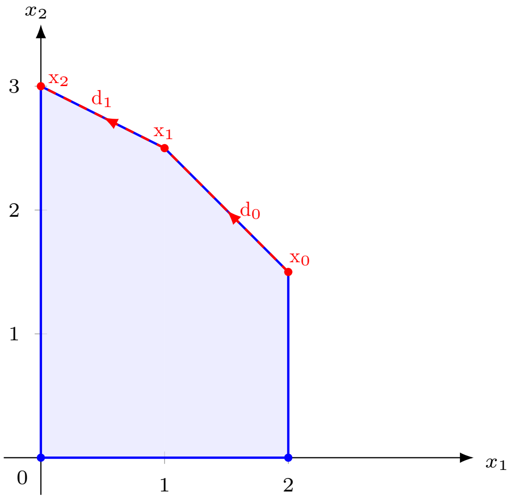
- This leads to $\mathbf{x}_2 = \mathbf{x}_1 + \alpha_1 \mathbf{d}_1 = \begin{bmatrix} 0 & 3 \end{bmatrix}^T$ with $l = 2$ (by swapping $\mathbf{a}_2^T$ and $\mathbf{a}_1^T$)

$$\mathbf{A}_{a_2} = \begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix} \text{ and } \mathcal{J}_2 = \{5, 1\}$$

  Which is complete the second iteration.

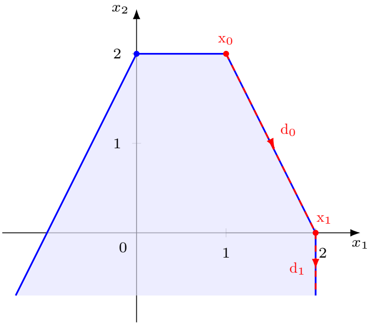
- Vertex x_2 is confirmed to be a minimizer at the beginning of the third iteration since the equation $\mathbf{A}_{a_2}^T \boldsymbol{\mu}_2 = -\mathbf{c}$ and yields nonnegative Lagrange multipliers $\boldsymbol{\mu}_2 = \begin{bmatrix} 2 & 1 \end{bmatrix}^T$.

Simplex Method: Nondegenerate Example



Simplex Method: Nondegenerate Example

$$
\begin{aligned}
\text{minimize} \quad & f(\mathbf{x}) = x_1 + x_2 \\
\text{subject to} \quad & x_1 \leq 2 \\
& x_2 \leq 2 \\
& -2x_1 + x_2 \leq 2 \\
& 2x_1 + x_2 \leq 4
\end{aligned}
$$



the constraints can be written as $\mathbf{A}\mathbf{x} \leq \mathbf{b}$ with

$$
\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -2 & 1 \\ 2 & 1 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 2 \\ 2 \\ 2 \\ 4 \end{bmatrix}
$$

Note: The feasible region is unbounded.

Simplex Method: Nondegenerate Example

- We start with the vertex $\mathbf{x}_0 = \begin{bmatrix} 1 & 2 \end{bmatrix}^T$. At $\mathbf{x}_0$

$$\mathbf{r}_0 = \mathbf{A}\mathbf{x}_0 - \mathbf{b} = \begin{bmatrix} -1 & 0 & -2 & 0 \end{bmatrix}^T, \text{ second and fourth constraints are active.}$$

$$\mathbf{A}_{a_0} = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix} \text{ and } \mathcal{J}_0 = \{2, 4\}$$

- From $\mathbf{A}_{a_0}^T \boldsymbol{\mu}_0 = -\mathbf{c}$, we have

$$\begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix} \mu_0 = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \Longrightarrow \boldsymbol{\mu}_0 = -\frac{1}{2} \begin{bmatrix} 1 & -2 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$$

$\mathbf{x}_0$ is not a minimizer.

- Since both components of $\boldsymbol{\mu}_0$ are negative, we can choose index l to be either 1 or 2.

- Choosing $l = 1$,

$$\mathbf{A}_{a_0}\mathbf{d}_0 = -\mathbf{e}_1 \implies \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}\mathbf{d}_0 = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \quad \implies \mathbf{d}_0 = -\frac{1}{2}\begin{bmatrix} 1 & -1 \\ -2 & 0 \end{bmatrix}\begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ -1 \end{bmatrix}$$

- The residual vector at $\mathbf{x}_0$ is given by

$$\mathbf{r}_0 = \mathbf{A}\mathbf{x}_0 - \mathbf{b} = \begin{bmatrix} -1 & 0 & -2 & 0 \end{bmatrix}^T$$ the first and third constraints are inactive at $\mathbf{x}_0$.

- We compute

$$\begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_3^T \end{bmatrix}\mathbf{d}_0 = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} -\frac{1}{2} \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ -2 \end{bmatrix}, \quad \mathcal{I}_0 = \{1\}$$

$$\alpha_0 = \frac{-r_{01}}{\mathbf{a}_1^T\mathbf{d}_0} = \frac{1}{\frac{1}{2}} = 2, \quad i^* = 1$$

- the next vertex is $\mathbf{x}_1 = \mathbf{x}_0 + \alpha_0 \mathbf{d}_0 = \begin{bmatrix} 2 & 0 \end{bmatrix}^T$, with

$$\mathbf{A}_{a_1} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}, \text{ and } \mathcal{J}_1 = \{1, 4\}$$

- Check whether x_1 is a minimizer by solving

$$\mathbf{A}_{a_1}^T \boldsymbol{\mu}_1 = -\mathbf{c} \quad \Longrightarrow \quad \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \boldsymbol{\mu}_1 = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \quad \Longrightarrow \quad \boldsymbol{\mu}_1 = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

  indicating that $\mathbf{x}_1$ is not a minimizer and $l = 2$
- Solving

$$\mathbf{A}_{a_1} \mathbf{d}_1 = -\mathbf{e}_2 \quad \Longrightarrow \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \mathbf{d}_1 = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \quad \Longrightarrow \mathbf{d}_1 = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

Simplex Method: Nondegenerate Example

- The residual vector at x_1 is

$$\mathbf{r}_1 = \mathbf{A}\mathbf{x}_1 - \mathbf{b} = \begin{bmatrix} 0 & -2 & -6 & 0 \end{bmatrix}^T$$

- The second and third constraints are inactive. We evaluate

$$\begin{bmatrix} \mathbf{a}_2^T \\ \mathbf{a}_3^T \end{bmatrix} \mathbf{d}_1 = \begin{bmatrix} 0 & 1 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

- Since $\mathcal{I}_1$ is empty, we conclude that the solution of this LP problem is un-bounded.

Reference

1. Joaquim R. R. A. Martins, and Andrew Ning, ”*Engineering Design Optimization*,” Cambridge University Press, 2021.
2. Jorge Nocedal, and Stephen J. Wright, ”*Numerical Optimization*,” 2nd, Springer, 2026
3. Mykel J. Kochenderfer, and Tim A. Wheeler, ”*Algorithms for Optimization*,” The MIT Press, 2019