Chapter 15 Introduction to Linear Programming

An Introduction to Optimization

Spring, 2014

Wei-Ta Chu

Brief History of Linear Programming

- The goal of linear programming is to determine the values of decision variables that maximize or minimize a linear objective function, where the decision variables are subject to linear constraints.
- A linear programming problem is a special case of a general constrained optimization problem. The objective function is linear, and the set of feasible points is determined by a set of linear equations and/or inequalities.

Brief History of Linear Programming

- The solution to a linear programming problem can be found by searching through a particular finite number of feasible points, known as ***basic feasible solutions***.
- We can simply compare the basic feasible solutions and find one that minimizes or maximizes the objective function – ***brute-force approach***.
- An alternative approach is to use experienced planners to optimize this problem. Such an approach relies on heuristics. Heuristics come close, but give suboptimal answers.

Brief History of Linear Programming

- Efficient methods became available in the late 1930s.
- In 1939, Kantorovich presented a number of solutions to some problems related to production and transportation planning.
- During World War II, Koopmans contributed significantly to the solution of transportation problems.
- They were awarded a Nobel Prize in Economics in 1975 for their work on the theory of optimal allocation of resources.
- In 1947, Dantzig developed the ***simplex method***.

Brief History of Linear Programming

- The simplex method has the undesirable property that in the worst case, the number of steps required to find a solution grows exponentially with the number of variables. Thus, the simplex method is said to have ***exponential worst-case complexity***.
- Khachiyan, in 1979, devise a polynomial complexity algorithm. In 1984, Karmarkar proposed a new linear programming algorithm that has polynomial complexity and appears to solve some complicated real-world problems of scheduling, routing, and planning more efficiently than the simplex method. His work led to the development of many ***interior-point methods***. This approach is currently still an active research area.

Simple Examples of Linear Programs

- Formally, a linear program is an optimization problem of the form

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

  where $\boldsymbol{c} \in R^n, \boldsymbol{b} \in R^m, \boldsymbol{A} \in R^{m \times n}$. The vector inequality $\boldsymbol{x} \geq \mathbf{0}$ means that each component of $\boldsymbol{x}$ is nonnegative.

- Several variations of this problem are possible. For example, we can maximize, or the constraints may be in the form of inequalities, such as $\boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}$ or $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$. In fact, these variations can all be rewritten into the standard form shown above.

Example

- A manufacturer produces four different products: X_1, X_2, X_3, X_4 there are three inputs to this production process: labor in person-weeks, kilograms of raw material A, and boxes of raw material B. Each product has different input requirements. In determining each week's production schedule, the manufacturer cannot use more than the available amounts of labor and the two raw materials. The relevant information is presented in this table. Every production decision must satisfy the restrictions on the availability of inputs. These constraints can be written using the data in this table.

$$x_1 + 2x_2 + x_3 + 2x_4 \leq 20$$
$$6x_1 + 5x_2 + 3x_3 + 2x_4 \leq 100$$
$$3x_1 + 4x_2 + 9x_3 + 12x_4 \leq 75$$

Inputs	Products				Input Availabilities
	X_1	X_2	X_3	X_4	
man weeks	1	2	1	2	20
kilograms of material A	6	5	3	2	100
boxes of material B	3	4	9	12	75
production levels	x_1	x_2	x_3	x_4	

Example

- Because negative production levels are not meaningful, we must impose the following nonnegativity constraints on the production levels: $x_i \geq 0, i = 1, 2, 3, 4$
- Now, suppose that one unit of product X_1 sells for \$6, and X_2, X_3, X_4 sell for \$4, \$7, \$5, respectively. Then, the total revenue for any production decision (x_1, x_2, x_3, x_4) is

$$f(x_1, x_2, x_3, x_4) = 6x_1 + 4x_2 + 7x_3 + 5x_4$$

- The problem is then to maximize f subject to the given constraints (the three inequalities and four nonnegativity constraints).

Example

- Using vector notation with $\boldsymbol{x} = [x_1, x_2, x_3, x_4]^T$, the problem can be written in the compact form

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

where

$$\boldsymbol{c}^T = [6, 4, 7, 5] \qquad \boldsymbol{A} = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 6 & 5 & 3 & 2 \\ 3 & 4 & 9 & 12 \end{bmatrix} \qquad \boldsymbol{b} = \begin{bmatrix} 20 \\ 100 \\ 75 \end{bmatrix}$$

Example

- ***Diet Problem***. Assume that n different food types are available. The jth food sells at a price c_j per unit. In addition, there are m basic nutrients. To achieve a balanced diet, you must receive at least b_i units of the ith nutrient per day. Assume that each unit of food j contains a_{ij} units of the i th nutrient. Denote by x_j the number of units of food j in the diet. The objective is to select the x_j to minimize the total cost of the diet:

$$\text{minimize} \quad c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

subject to the nutritional constraints, and the nonnegativity constraints

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \geq b_1$$
$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \geq b_2 \qquad x_1 \geq 0, x_2 \geq 0, ..., x_n \geq 0$$

$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \geq b_m$$

Example

- In the more compact vector notation, this problem becomes

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

where $\boldsymbol{x} = [x_1, x_2, ..., x_n]^T$ is an n-dimensional column vector, $\boldsymbol{c}^T$ is an n-dimensional row vector, $\boldsymbol{A}$ is an $m \times n$ matrix, and $\boldsymbol{b}$ is an m-dimensional column vector.

Example

- A manufacturer produces two different products, X_1, X_2, using three machines: M_1, M_2, M_3 . Each machine can be used for only a limited amount of time. Production times of each product on each machine are given in this table. The objective is to maximize the combined time of utilization of all three machines.
- Every production decision must satisfy the constraints on the available time. These restrictions can be written as

$$x_1 + x_2 \leq 8$$
$$x_1 + 3x_2 \leq 18$$
$$2x_1 + x_2 \leq 14$$

x_1 and x_2 denote the production levels

Machine	Production time (hours/unit) X_1	Production time (hours/unit) X_2	Available time (hours)
M_1	1	1	8
M_2	1	3	18
M_3	2	1	14
Total	4	5	

Example

- The combined production time of all three machines is

$$f(x_1, x_2) = 4x_1 + 5x_2$$

- Thus, writing $\boldsymbol{x} = [x_1, x_2]^T$, the problem in compact notation has the form

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

where

$$\boldsymbol{c}^T = [4, 5] \qquad \boldsymbol{A} = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 2 & 1 \end{bmatrix} \qquad \boldsymbol{b} = \begin{bmatrix} 8 \\ 18 \\ 14 \end{bmatrix}$$

Example

- A manufacturing companying has plants in cities A, B, and C. The company produces and distributes its product to dealers in various cities. On a particular day, the company has 30 units of its product in A, 40 in B, and 30 in C. The company plans to ship 20 units to D, 20 to E, 25 to F, and 35 to G, following orders received from dealers. The transportation costs per unit of each product between the cities are given in this table. In the table, the quantities supplied and demand appear at the right and along the bottom of the table. The quantities to be transported from the plants to different destinations are represented by the decision variables.

To / From	D	E	F	G	Supply
A	$7	$10	$14	$8	30
B	$7	$11	$12	$6	40
C	$5	$8	$15	$9	30
Demand	20	20	25	35	100

Example

- The problem can be stated in the form

$$\text{minimize } 7x_{11} + 10x_{12} + 14x_{13} + 8x_{14} + 7x_{21} + 11x_{22} + 12x_{23} + 6x_{24} + 5x_{31} + 8x_{32} + 15x_{33} + 9x_{34}$$

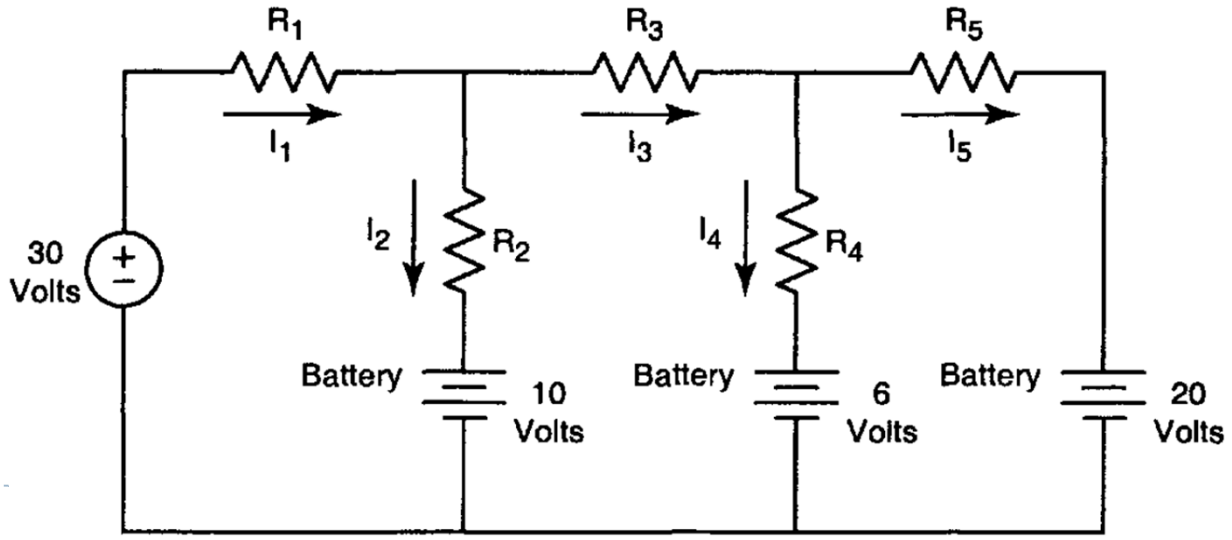
$$\begin{aligned}
\text{subject to } & x_{11} + x_{12} + x_{13} + x_{14} = 30 \\
& x_{21} + x_{22} + x_{23} + x_{24} = 40 \\
& x_{31} + x_{32} + x_{33} + x_{34} = 30 \\
& x_{11} + x_{21} + x_{31} = 20 \\
& x_{12} + x_{22} + x_{32} = 20 \\
& x_{13} + x_{23} + x_{33} = 25 \\
& x_{14} + x_{24} + x_{34} = 35 \\
& x_{11}, x_{12}, ..., x_{34} \geq 0
\end{aligned}$$

Example

- In this problem one of the constraint equations is redundant because it can be derived from the rest of the constraint equations. The mathematical formulation of the transportation problem is then in a linear programming form with twelve (3×4) decision variables and six $(3 + 4 - 1)$ linearly independent constraint equations. Obviously, we also require nonnegativity of the decision variables, since a negative shipment is impossible and does not have a valid interpretation.

Example

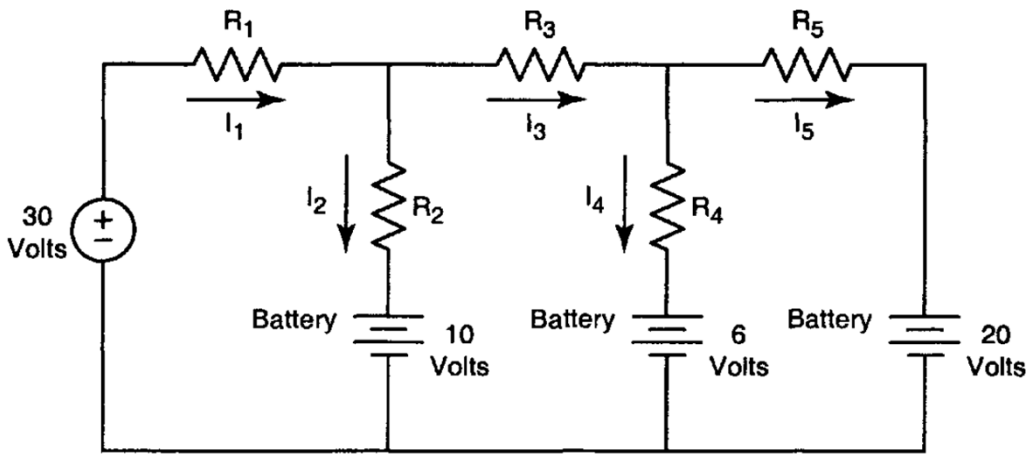
- An electric circuit is designed to use a 30-V source to charge 10-V, 6-V, and 20-V batteries connected in parallel. Physical constraints limit the currents I_1, I_2, I_3, I_4, I_5 to a maximum of 4A, 3A, 3A, 2A, and 2A, respectively. In addition, the batteries must not be discharged; that is, the currents I_1, I_2, I_3, I_4, I_5 must not be negative. We wish to find the values of the currents such that the total power transferred to the batteries is maximized.



Example

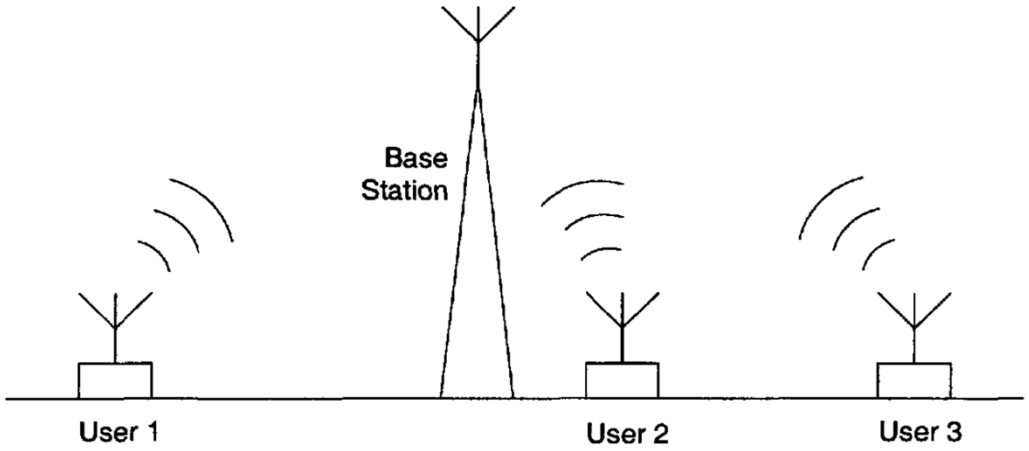
- The total power transferred to the batteries is the sum of the powers transferred to each battery and is given by $10I_2 + 6I_4 + 20I_5$ W. From the circuit, we observe that the currents satisfy the constraints $I_1 = I_2 + I_3$ and $I_3 = I_4 + I_5$. Therefore, the problem can be posed as the following linear program:

$$
\begin{aligned}
&\text{maximize } 10I_2 + 6I_4 + 20I_5 \\
&\text{subject to } I_1 = I_2 + I_3 \\
&\qquad\qquad I_3 = I_4 + I_5 \\
&\qquad\qquad I_1 \leq 4 \\
&\qquad\qquad I_2 \leq 3 \\
&\qquad\qquad I_3 \leq 3 \\
&\qquad\qquad I_4 \leq 2 \\
&\qquad\qquad I_5 \leq 2 \\
&\qquad\qquad I_1, I_2, I_3, I_4, I_5 \geq 0
\end{aligned}
$$



Example

- Consider the wireless communication system. There are n mobile users. For each $i = 1, ..., n$, user i transmits a signal to the base station with power p_i and an attenuation factor of h_i (i.e., the actual signal power received at the base station from user i is $h_i p_i$). When the base station is receiving from user i the total power received from all other users is considered ***interference*** (i.e., $\sum_{j \neq i} h_j p_j$). For the communication with user i to be reliable, the signal-to-interference ratio must exceed a threshold γ_i , where the "signal" is the power received from user i .



Example

- We are interested in minimizing the total power transmitted by all users subject to having reliable communication for all users. We can formulate the problem as a linear programming problem of the form

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

- The total power transmitted is $p_1 + \cdots + p_n$. The signal-to-interference ratio for user i

$$\frac{h_i p_i}{\sum_{j \neq i} h_j p_j}$$

Example

- Hence, the problem can be written as

$$\begin{aligned} &\text{minimize } p_1 + \cdots + p_n \\ &\text{subject to } \frac{h_i p_i}{\sum_{j \neq i} h_j p_j} \geq \gamma_i \qquad i = 1, ..., n \\ &p_1, ..., p_n \geq 0 \end{aligned}$$

- We can write the above as the linear programming problem

$$\begin{aligned} &\text{minimize } p_1 + \cdots + p_n \\ &\text{subject to } h_i p_i - \gamma_i \sum_{j \neq i} h_j p_j \geq 0 \qquad i = 1, ..., n \\ &p_1, ..., p_n \geq 0 \end{aligned}$$

- In matrix form,

$$\boldsymbol{c} = [1, ..., 1]^T$$

$$\boldsymbol{A} = \begin{bmatrix} h_1 & -\gamma_1 h_2 & \cdots & -\gamma_1 h_n \\ -\gamma_2 h_1 & h_2 & \cdots & -\gamma_2 h_n \\ \vdots & & \ddots & \vdots \\ -\gamma_n h_1 & -\gamma_n h_2 & \cdots & h_n \end{bmatrix}$$

Two-Dimensional Linear Programs

- Consider the following linear program

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

  where $\boldsymbol{x} = [x_1, x_2]^T$ and

$$\boldsymbol{c} = [1, 5]^T \qquad \boldsymbol{A} = \begin{bmatrix} 5 & 6 \\ 3 & 2 \end{bmatrix} \qquad \boldsymbol{b} = \begin{bmatrix} 30 \\ 12 \end{bmatrix}$$

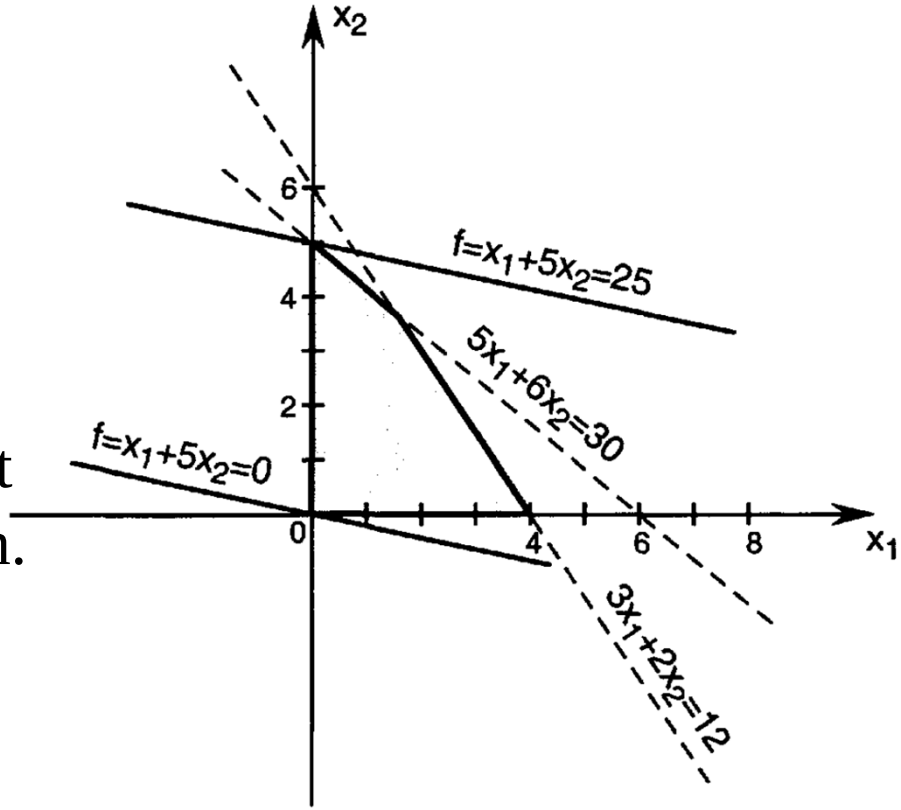
- First, we note that the set of equations $\{\boldsymbol{c}^T\boldsymbol{x} = x_1 + 5x_2 = f, f \in R\}$ specifies a family of straight lines in R^2. Each member of this family can be obtained by setting f equal to some real number. Thus, for example, $x_1 + 5x_2 = -5$, $x_1 + 5x_2 = 0$, and $x_1 + 5x_2 = 3$ are three parallel lines belonging to the family.

Two-Dimensional Linear Programs

- Now, suppose that we try to choose several values for x_1 and x_2 and observe how large we can make f while still satisfying the constraints on x_1 and x_2. We first try $x_1 = 1$ and $x_2 = 3$. This point satisfies the constraints. For this point, $f = 16$. If we now select $x_1 = 0$ and $x_2 = 5$, then $f = 25$ and this point yields a larger value for f than does $\boldsymbol{x} = [1, 3]^T$. There are infinitely many points $[x_1, x_2]^T$ satisfying the constraints. Therefore, we need a better method than trial and error to solve the problem.

Two-Dimensional Linear Programs

- For the example above we can easily solve the problem using geometric arguments. First let us sketch the constraints in R^2 . The region of feasible points (the set of points $\boldsymbol{x}$ satisfying the constraints $\boldsymbol{Ax} \leq \boldsymbol{b},\ \boldsymbol{x} \geq \mathbf{0}$) is depicted by the shaded region in this figure.
- Geometrically, maximizing $\boldsymbol{c}^T\boldsymbol{x} = x_1 + 5x_2$ subject to the constraints can be thought of as finding the straight line $f = x_1 + 5x_2$ that intersects the shaded region and has the largest f. The point $[0, 5]^T$ is the solution.



Example

- Suppose that you are given two different types of concrete. The first type contains 30% cement, 40% gravel, and 30% sand (all percentages of weight). The second type contains 10% cement, 20% gravel, and 70% sand. The first type of concrete costs $5 per pound and the second type costs $1 per pound. How many pounds of each type of concrete should you buy and mix together so that your cost is minimized but you get a concrete mixture that has at least a total of 5 pounds of cement, 3 pounds of gravel, and 4 pounds of sand?

Example

- The problem can be represented as

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

$$\boldsymbol{c} = [5, 1]^T \quad \boldsymbol{A} = \begin{bmatrix} 0.3 & 0.1 \\ 0.4 & 0.2 \\ 0.3 & 0.7 \end{bmatrix} \quad \boldsymbol{b} = \begin{bmatrix} 5 \\ 3 \\ 4 \end{bmatrix}$$

- Using the graphical method described above, we get a solution of $[0, 50]^T$, which means that we would purchase 50 pounds of the second type of concrete.

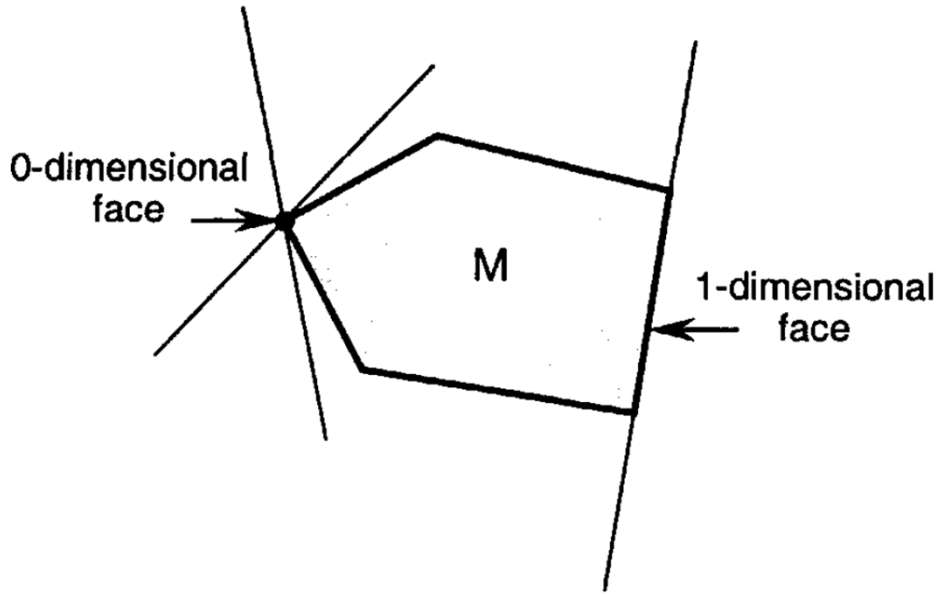
- In some case, there may be more than one point of intersection, and therefore any one of them is a solution.

Convex Polyhedra and Linear Programming

- We discuss linear programs from a geometric point of view. The set of points satisfying these constraints can be represented as the intersection of a finite number of closed half-spaces. Thus, the constraints define a convex polytope.
- We assume, for simplicity, that this polytope is nonempty and bounded. In other words, the equations of constraints define a polyhedron M in R^n. Let H be a hyperplane of support of this polyhedron. If the dimension of M is less than n, then the set of all points common to the hyperplane H and the polyhedron M coincides with M.

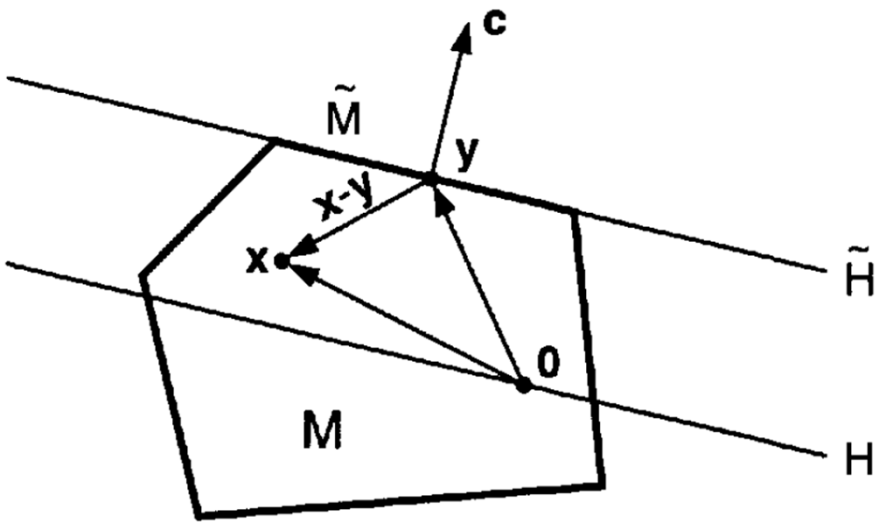
Convex Polyhedra and Linear Programming

- If the dimension of M is equal to n, then the set of all points common to the hyperplane H and the polyhedron M is a face of the polyhedron. If this face is $(n-1)$-dimensional, then there exists only one hyperplane of support, namely, the carrier of this face. If the dimension of the face is less than $n-1$, then there exist an infinite number of hyperplanes of support whose intersection with this polyhedron yields this face.



Convex Polyhedra and Linear Programming

- The goal of our linear programming problem is to maximize a linear objective function $f(\boldsymbol{x}) = \boldsymbol{c}^T\boldsymbol{x} = c_1x_1 + \cdots + c_nx_n$ on the convex polyhedron M. Next, let H be the hyperplane defined by the equation $\boldsymbol{c}^T\boldsymbol{x} = 0$. Draw a hyperplane of support $\tilde{H}$ to the polyhedron M, which is parallel to H and positioned such that the vector $\boldsymbol{c}$ points in the direction of the half-space that does not contain M.

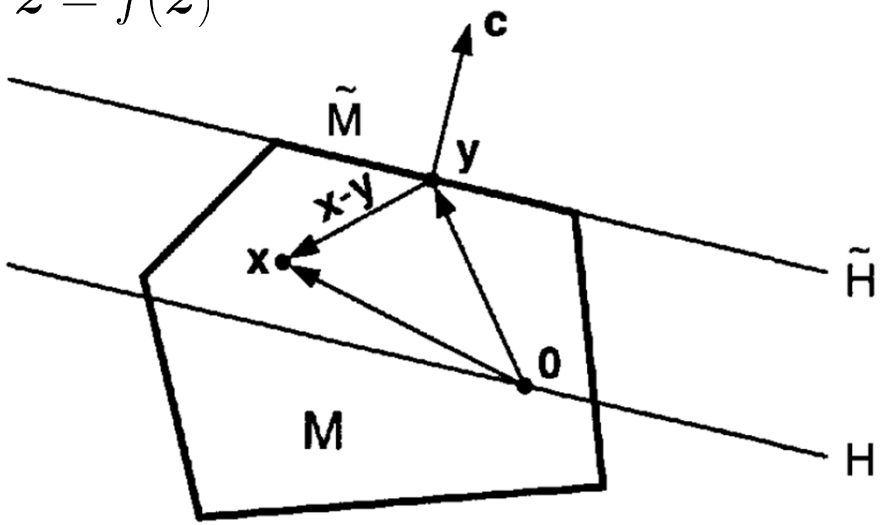


Convex Polyhedra and Linear Programming

- The equation of the hyperplane $\tilde{H}$ has the form $\boldsymbol{c}^T\boldsymbol{x} = \beta$, and for all $\boldsymbol{x} \in M$, we have $\boldsymbol{c}^T\boldsymbol{x} \leq \beta$. Denote by $\tilde{M}$ the convex polyhedron that is the intersection of the hyperplane of support $\tilde{H}$ with the polyhedron M. We now show that f is constant on $\tilde{M}$ and that $\tilde{M}$ is the set of all points in M for which f attains its maximum value. To this end, let $\boldsymbol{y}$ and $\boldsymbol{z}$ be two arbitrary points in $\tilde{M}$. This implies that both $\boldsymbol{y}$ and $\boldsymbol{z}$ belong to $\tilde{H}$. Hence,

$$f(\boldsymbol{y}) = \boldsymbol{c}^T\boldsymbol{y} = \beta = \boldsymbol{c}^T\boldsymbol{z} = f(\boldsymbol{z})$$

which means that f is constant on $\tilde{M}$.



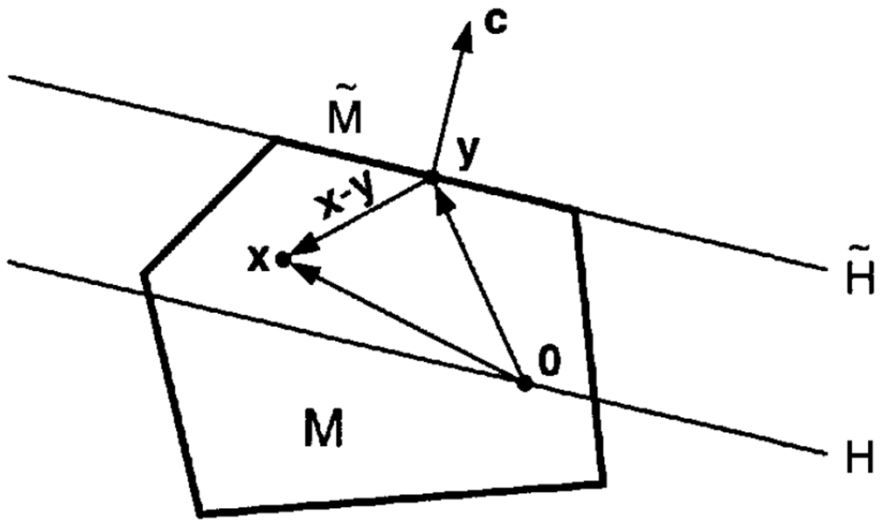
Convex Polyhedra and Linear Programming

- Let $\boldsymbol{y}$ be a point of $\tilde{M}$, and let $\boldsymbol{x}$ be a point of $M \setminus \tilde{M}$; that is, $\boldsymbol{x}$ is a point of M that does not belong to $\tilde{M}$. Then,

$$\boldsymbol{c}^T\boldsymbol{x} < \beta = \boldsymbol{c}^T\boldsymbol{y}$$

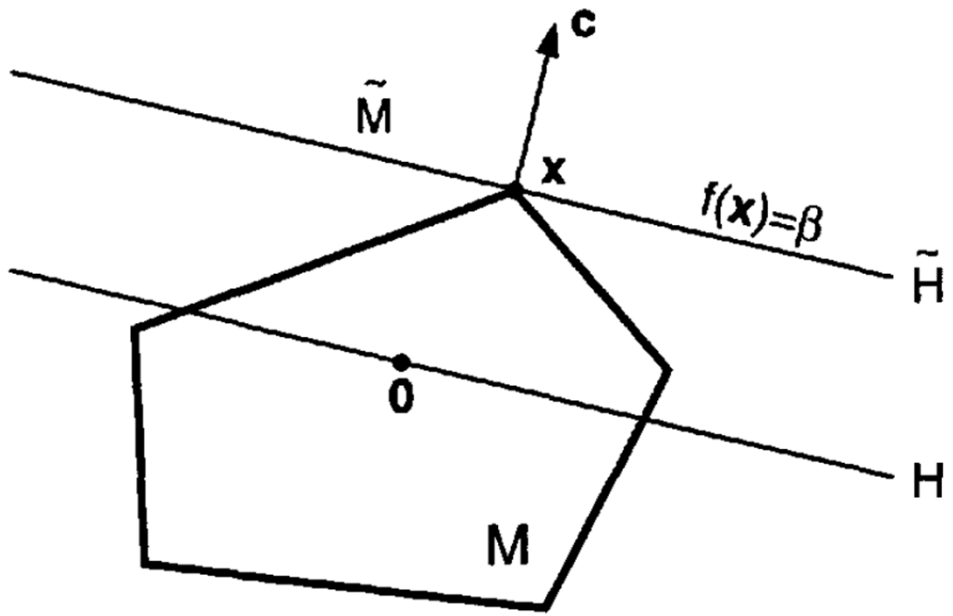
  which implies that $f(\boldsymbol{x}) < f(\boldsymbol{y})$

- Thus, the values of f at the points of M that do not belong to $\tilde{M}$ are smaller than the values at points of $\tilde{M}$. Hence, f achieves its maximum on M at points in $\tilde{M}$.



Convex Polyhedra and Linear Programming

- It may happen that $\tilde{M}$ contains only a single point, in which case f achieves its maximum at a unique point.
- This occurs when the hyperplane of support passes through an extreme point of M.



Standard Form Linear Programs

- We refer to a linear program of the form

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

  as a linear program in ***standard form***. Here $\boldsymbol{A}$ is an $m \times n$ matrix composed of real entries, $m < n,\ rank(\boldsymbol{A}) = m$

- Without loss of generality, we assume that $\boldsymbol{b} \geq \mathbf{0}$. If a component of $\boldsymbol{b}$ is negative, say the i th component, we multiply the ith constraint by -1 to obtain a positive right-hand side.

Standard Form Linear Programs

- Theorems and solution techniques for linear programs are usually stated for problems in standard form. Other forms of linear programs can be converted to the standard form.
- If a linear program is in the form

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

  then by introducing ***surplus variables*** y_i, we can convert the original problem into the standard form

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^T\boldsymbol{x} \\ \text{subject to} \quad & a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n - y_i = b_i \qquad i = 1, ..., m \\ & x_1 \geq 0, x_2 \geq 0, ..., x_n \geq 0 \\ & y_1 \geq 0, y_2 \geq 0, ..., y_m \geq 0 \end{aligned}$$

Standard Form Linear Programs

- In more compact notation, the formulation above can be represented as

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^T \boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} - \boldsymbol{I}_m \boldsymbol{y} = [\boldsymbol{A}, -\boldsymbol{I}_m] \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \qquad \boldsymbol{y} \geq \boldsymbol{0} \end{aligned}$$

  where $\boldsymbol{I}_m$ is the $m \times m$ identity matrix.

Standard Form Linear Programs

- If, on the other hand, the constraints have the form

$$\boldsymbol{Ax} \leq \boldsymbol{b} \quad \boldsymbol{x} \geq \mathbf{0}$$

then we introduce ***slack variables*** y_i to convert the constraints into the form

$$\boldsymbol{Ax} + \boldsymbol{I}_m\boldsymbol{y} = [\boldsymbol{A}, \boldsymbol{I}_m] \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} = \boldsymbol{b}$$

$$\boldsymbol{x} \geq \mathbf{0} \quad \boldsymbol{y} \geq \mathbf{0}$$

where $\boldsymbol{y}$ is the vector of slack variables. Note that neither surplus nor slack variables contribute to the objective function $\boldsymbol{c}^T\boldsymbol{x}$

Standard Form Linear Programs

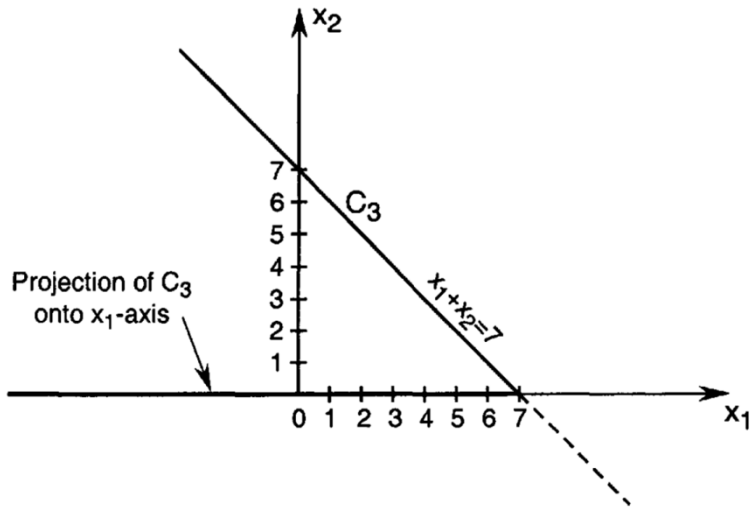
- At first glace, it may appear that the two problems

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^T\boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{aligned} \qquad \begin{aligned} \text{minimize} \quad & \boldsymbol{c}^T\boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} - \boldsymbol{I}_m\boldsymbol{y} = [\boldsymbol{A}, -\boldsymbol{I}_m] \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \\ & \boldsymbol{y} \geq \boldsymbol{0} \end{aligned}$$

  are different, in that the first problem refers to the intersection of half-spaces in the n-dimensional space, whereas the second problem refers to an intersection of half-spaces and hyperplanes in the $(n+m)$-dimensional space. It turns out that both formulations are algebraically equivalent in the sense that a solution to one of the problems implies a solution to the other.

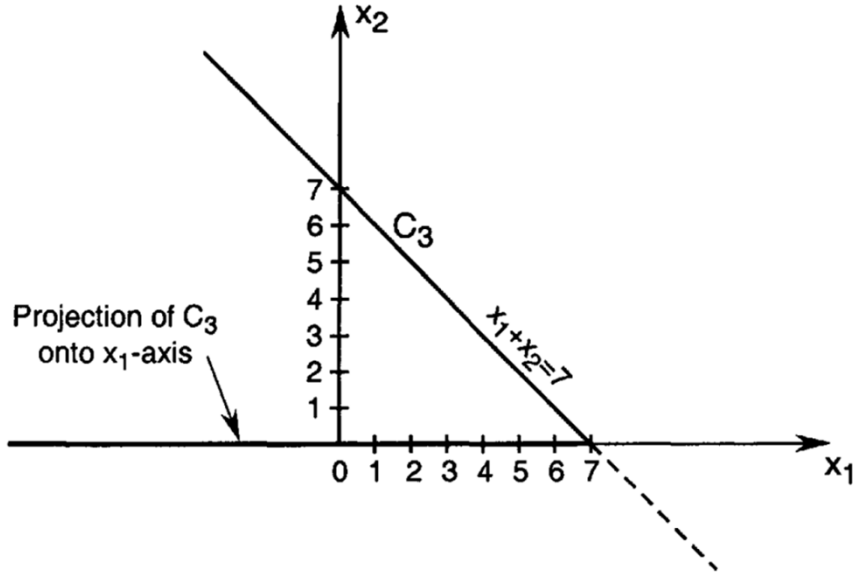
Example

- Suppose that we are given the inequality constraint $x_1 \leq 7$. We convert this to an equality constraint by introducing a slack variable $x_2 \geq 0$ to obtain $x_1 + x_2 = 7, x_2 \geq 0$
- Consider the sets $C_1 = \{x_1 : x_1 \leq 7\}$ and $C_2 = \{x_1 : x_1 + x_2 = 7 : x_2 \geq 0\}$ Are the sets C_1 and C_2 equal? It is clear that indeed they are; we give a geometric interpretation for their equality.
- Consider a third set $C_3 = \{[x_1, x_2]^T : x_1 + x_2 = 7, x_2 \geq 0\}$. From this figure we can see that the set C_3 consists of all points on the line to the left and above the point of intersection of the line with the x_1-axis.



Example

- This set, being a subset of R^2, is of course not the same set as the set C_1 (a subset of R). However, we can project the set C_3 onto the x_1-axis. We can associate with each point $x_1 \in C_1$ a point $[x_1, 0]^T$ on the orthogonal projection of C_3 onto the x_1-axis, and vice versa. Note that $C_2 = \{x_1 : [x_1, x_2]^T \in C_3\} = C_1$



Example

- Consider the inequality constraints

$$a_1x_1 + a_2x_2 \leq b \qquad x_1, x_2 \geq 0$$

where a_1, a_2, b are positive numbers. Again, we introduce a slack variable $x_3 \geq 0$ to get

$$a_1x_1 + a_2x_2 + x_3 = b \qquad x_1, x_2, x_3 \geq 0$$

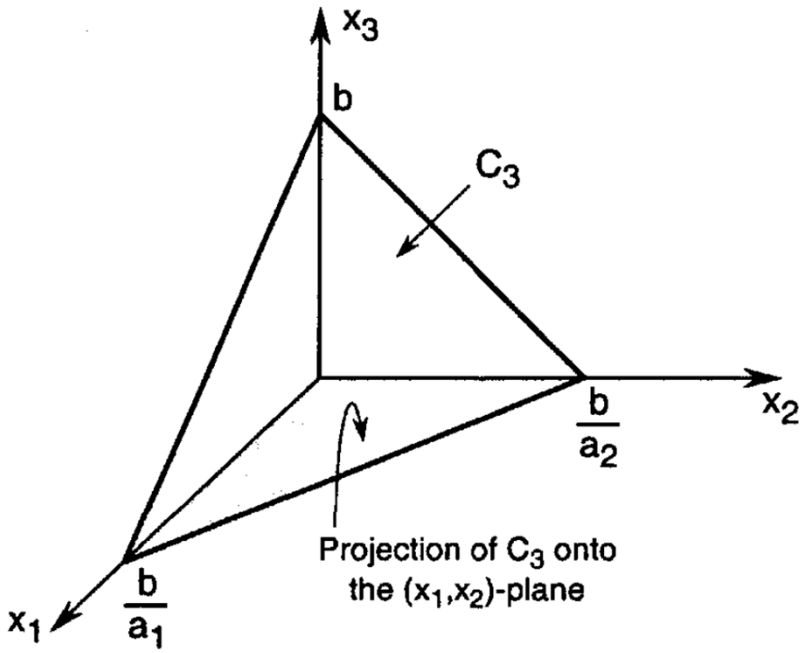
Define the sets

$$C_1 = \{[x_1, x_2]^T : a_1x_1 + a_2x_2 \leq b, x_1, x_2 \geq 0\}$$
$$C_2 = \{[x_1, x_2]^T : a_1x_1 + a_2x_2 + x_3 = b, x_1, x_2, x_3 \geq 0\}$$
$$C_3 = \{[x_1, x_2, x_3]^T : a_1x_1 + a_2x_2 + x_3 = b, x_1, x_2, x_3 \geq 0\}$$

We again see that C_3 is not the same as C_1. However, the orthogonal projection of C_3 onto the (x_1, x_2)-plane allows us to associate the resulting set with the set C_1.

Example

- We associate the points $[x_1, x_2, 0]^T$ resulting from the orthogonal projection of C_3 onto the (x_1, x_2)-plane with the points in C_1 . Note that $C_2 = \{[x_1, x_2]^T : [x_1, x_2, x_3]^T \in C_3\} = C_1$



Example

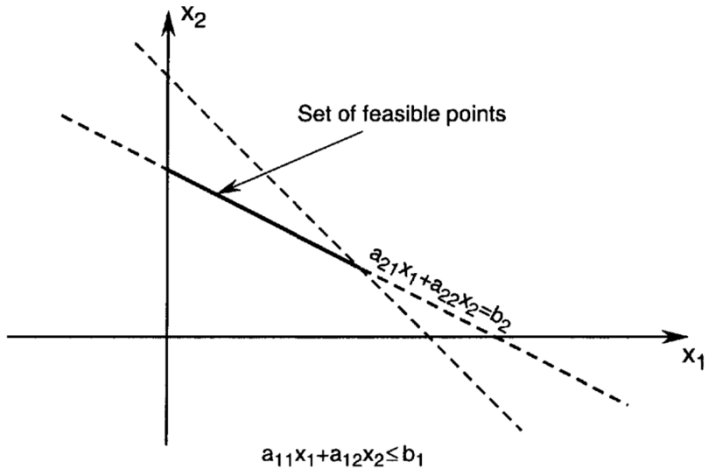
- Suppose that we wish to maximize

$$f(x_1, x_2) = c_1 x_1 + c_2 x_2$$

subject to the constraints

$$a_{11} x_1 + a_{12} x_2 \leq b_1$$
$$a_{21} x_1 + a_{22} x_2 = b_2$$
$$x_1, x_2 \geq 0$$

where, for simplicity, we assume that each $a_{ij} > 0$ and $b_1, b_2 \geq 0$ The set of feasible points is depicted in this figure. Let $C_1 \subset R^2$ be the set of points satisfying the constraints.

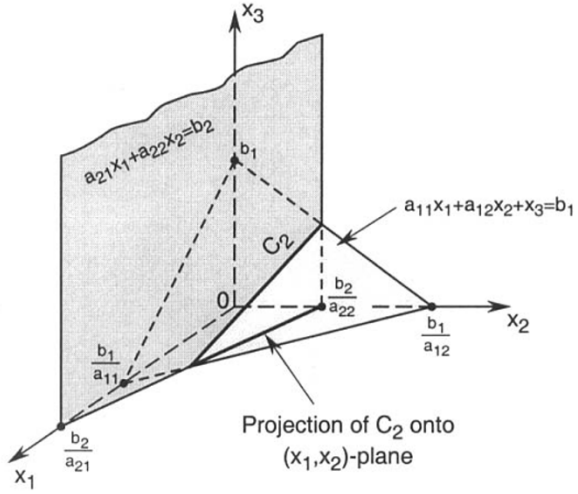


Example

- Introducing a slack variable, we convert the constraints into standard form:

$$a_{11}x_1 + a_{12}x_2 + x_3 = b_1$$
$$a_{21}x_1 + a_{22}x_2 = b_2$$
$$x_i \geq 0, i = 1, 2, 3$$

- Let $C_2 \subset R^3$ be the set of points satisfying the constraints. As illustrated in this figure, this set is a line segment (in R^3). We now project C_2 onto the (x_1, x_2)-plane. The projected set consists of the points $[x_1, x_2, 0]^T$, with $[x_1, x_2, x_3]^T \in C_2$ for some $x_3 \geq 0$. In this figure this set is marked by a heavy line in the (x_1, x_2)-plane. We can associate the points on the projection with the corresponding points in the set C_1



Example

- Consider the following optimization problem

$$\begin{array}{lllll} \text{maximize} & x_2 - x_1 & & & \\ \text{subject to} & 3x_1 = x_2 - 5 & |x_2| \leq 2 & x_1 \leq 0 \end{array}$$

- To convert the problem into a standard form linear programming problem, we perform the following steps:
  - 1. Change the objective function to: minimize $x_1 - x_2$
  - 2. Substitute $x_1 = -x_1'$
  - 3. Write $|x_2| \leq 2$ as $x_2 \leq 2$ and $-x_2 \leq 2$
  - 4. Introduce slack variables x_3, x_4 , and convert the inequalities above to $x_2 + x_3 = 2$ and $-x_2 + x_4 = 2$
  - 5. Write $x_2 = u - v, u, v \geq 0$

Example

- Hence, we obtain

$$\begin{aligned} \text{minimize} \quad & -x_1' - u + v \\ \text{subject to} \quad & 3x_1' + u - v = 5 \\ & u - v + x_3 = 2 \\ & v - u + x_4 = 2 \\ & x_1', u, v, x_3, x_4 \geq 0 \end{aligned}$$

Basic Solutions

- In the following discussion we only consider linear programming problems in standard form.
- Consider the system of equalities $\boldsymbol{Ax} = \boldsymbol{b}$, where $rank(\boldsymbol{A}) = m$ In dealing with this system of equations, we frequently need to consider a subset of columns of the matrix $\boldsymbol{A}$. For convenience, we often reorder the columns of $\boldsymbol{A}$ so that the columns we are interested in appear first.
- Specifically, let $\boldsymbol{B}$ be a square matrix whose columns are m linearly independent columns of $\boldsymbol{A}$.

Basic Solutions

- If necessary, we reorder the columns of $\boldsymbol{A}$ so that the columns in $\boldsymbol{B}$ appear first: $\boldsymbol{A}$ has the form $\boldsymbol{A} = [\boldsymbol{B}, \boldsymbol{D}]$, where $\boldsymbol{D}$ is an $m \times (n - m)$ matrix whose columns are the remaining columns of $\boldsymbol{A}$. The matrix $\boldsymbol{B}$ is nonsingular, and thus we can solve the equation $\boldsymbol{B}\boldsymbol{x}_B = \boldsymbol{b}$ for the m-vector $\boldsymbol{x}_B$. The solution is $\boldsymbol{x}_B = \boldsymbol{B}^{-1}\boldsymbol{b}$. Let $\boldsymbol{x}$ be the n-vector whose first m components are equal to $\boldsymbol{x}_B$ and the remaining components are equal to zero; that is, $\boldsymbol{x} = [\boldsymbol{x}_B^T, 0^T]^T$. Then, $\boldsymbol{x}$ is a solution to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$

Basic Solutions

- Definition 15.1: We call $\boldsymbol{x} = [\boldsymbol{x}_B^T, 0^T]^T$ a ***basic solution*** to $\boldsymbol{Ax} = \boldsymbol{b}$ with respect to the ***basis*** $\boldsymbol{B}$. We refer to the components of the vector $\boldsymbol{x}_B$ as ***basic variables*** and the columns of $\boldsymbol{B}$ as ***basic columns***.
- If some of the basic variables of a basic solution are zero, then the basic solution is said to be a ***degenerate basic solution***.
- A vector $\boldsymbol{x}$ satisfying $\boldsymbol{Ax} = \boldsymbol{b}$, $\boldsymbol{x} \geq 0$, is said to be a ***feasible solution***.
- A feasible solution that is also basic is called a ***basic feasible solution***.
- If the basic feasible solution is a degenerate basic solution, then it is called a ***degenerate basic feasible solution***.
- Note that in any basic feasible solution, $\boldsymbol{x}_B \geq 0$

Example

- Consider the equation $\boldsymbol{Ax} = \boldsymbol{b}$

  $$\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_4] = \begin{bmatrix} 1 & 1 & -1 & 4 \\ 1 & -2 & -1 & 1 \end{bmatrix} \qquad \boldsymbol{b} = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$$

  where $\boldsymbol{a}_i$ denotes the i th column of the matrix $\boldsymbol{A}$

- Then, $\boldsymbol{x} = [6, 2, 0, 0]^T$ is a basic feasible solution with respect to the basis $\boldsymbol{B} = [\boldsymbol{a}_1, \boldsymbol{a}_2]$, $\boldsymbol{x} = [0, 0, 0, 2]^T$ is a degenerate basic feasible solution with respect to the basis $\boldsymbol{B} = [\boldsymbol{a}_3, \boldsymbol{a}_4]$ (as well as $[\boldsymbol{a}_1, \boldsymbol{a}_4]$ and $[\boldsymbol{a}_2, \boldsymbol{a}_4]$), $\boldsymbol{x} = [3, 1, 0, 1]^T$ is a feasible solution that is not basic, and $\boldsymbol{x} = [0, 2, -6, 0]^T$ is a basic solution with respect to the basis $\boldsymbol{B} = [\boldsymbol{a}_2, \boldsymbol{a}_3]$, but is not feasible.

Example

- Consider the system of linear equations $\boldsymbol{Ax} = \boldsymbol{b}$, where

$$\boldsymbol{A} = \begin{bmatrix} 2 & 3 & -1 & -1 \\ 4 & 1 & 1 & -2 \end{bmatrix} \qquad \boldsymbol{b} = \begin{bmatrix} -1 \\ 9 \end{bmatrix}$$

We now find all solutions of this system. Note that every solution $\boldsymbol{x}$ of $\boldsymbol{Ax} = \boldsymbol{b}$ has the form $\boldsymbol{x} = \boldsymbol{v} + \boldsymbol{h}$, where $\boldsymbol{v}$ is a particular solution of $\boldsymbol{Ax} = \boldsymbol{b}$ and $\boldsymbol{h}$ is a solution to $\boldsymbol{Ax} = \boldsymbol{0}$

- We form the augmented matrix $[\boldsymbol{A}, \boldsymbol{b}]$ of the system

$$[\boldsymbol{A}, \boldsymbol{b}] = \begin{bmatrix} 2 & 3 & -1 & -1 & -1 \\ 4 & 1 & 1 & -2 & 9 \end{bmatrix}$$

Using elementary row operations, we transform this matrix into the form given by

$$\begin{bmatrix} 1 & 0 & \frac{2}{5} & -\frac{1}{2} & \frac{14}{5} \\ 0 & 1 & -\frac{3}{5} & 0 & -\frac{11}{5} \end{bmatrix}$$

Example

- The corresponding system of equations is given by

$$x_1 + \frac{2}{5}x_3 - \frac{1}{2}x_4 = \frac{14}{5}$$
$$x_2 - \frac{3}{5}x_3 = -\frac{11}{5}$$

- Solving for the leading unknowns x_1 and x_2, we obtain

$$x_1 = \frac{14}{15} - \frac{2}{5}x_3 + \frac{1}{2}x_4$$
$$x_2 = -\frac{11}{5} + \frac{3}{5}x_3$$

where x_3 and x_4 are arbitrary real numbers. If $[x_1, x_2, x_3, x_4]^T$ is a solution, then we have

$$x_1 = \frac{14}{15} - \frac{2}{5}s + \frac{1}{2}t$$
$$x_2 = -\frac{11}{5} + \frac{3}{5}s$$
$$x_3 = s$$
$$x_4 = t$$

where we have substituted s and t for x_3 and x_4, respectively.

Example

- Using vector notation, we may write the system of equations above as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \frac{14}{5} \\ -\frac{11}{5} \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -\frac{2}{5} \\ \frac{3}{5} \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} \frac{1}{2} \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

- Note that we have infinitely many solutions, parameterized by $s, t \in R$. For the choice $s = t = 0$ we obtain a particular solution to $\boldsymbol{Ax} = \boldsymbol{b}$, given by

$$\boldsymbol{v} = \begin{bmatrix} \frac{14}{5} \\ -\frac{11}{5} \\ 0 \\ 0 \end{bmatrix}$$

Example

- Any other solution has the form $\boldsymbol{v} + \boldsymbol{h}$, where

$$\boldsymbol{h} = s \begin{bmatrix} -\frac{2}{5} \\ \frac{3}{5} \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} \frac{1}{2} \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

- The total number of possible basic solution is at most

$$\binom{n}{m} = \frac{n!}{m!(n-m)!} = \frac{4!}{2!(4-2)!} = 6$$

  to find basic solutions that are feasible, we check each of the basic solutions for feasibility.

Example

- Our first candidate for a basic feasible solution is obtained by setting $x_3 = x_4 = 0$, which corresponds to the basis $\boldsymbol{B} = [\boldsymbol{a}_1, \boldsymbol{a}_2]$ Solving $\boldsymbol{B}\boldsymbol{x}_B = \boldsymbol{b}$, we obtain $\boldsymbol{x}_B = [14/5, -11/5]^T$, and hence $\boldsymbol{x} = [14/5, -11/5, 0, 0]^T$ is a basic solution that is not feasible.
- For our second candidate basic feasible solution, we set $x_2 = x_4 = 0$. We have the basis $\boldsymbol{B} = [\boldsymbol{a}_1, \boldsymbol{a}_3]$. Solving $\boldsymbol{B}\boldsymbol{x}_B = \boldsymbol{b}$ yields $\boldsymbol{x}_B = [4/3, 11/3]^T$. Hence, $\boldsymbol{x} = [4/3, 0, 11/3, 0]^T$ is a basic feasible solution.
- A third candidate basic feasible solution is obtained by setting $x_2 = x_3 = 0$. However, the matrix

$$\boldsymbol{B} = [\boldsymbol{a}_1, \boldsymbol{a}_4] = \begin{bmatrix} 2 & -1 \\ 4 & -2 \end{bmatrix}$$

is singular. Therefore, $\boldsymbol{B}$ cannot be a basis, and we do not have a basic solution corresponding to $\boldsymbol{B} = [\boldsymbol{a}_1, \boldsymbol{a}_4]$

Example

- We get our fourth candidate for a basic solution by setting $x_1 = x_4 = 0$. We have a basis $\boldsymbol{B} = [\boldsymbol{a}_2, \boldsymbol{a}_3]$, resulting in $\boldsymbol{x} = [0, 2, 7, 0]^T$, which is a basic feasible solution.
- Our fifth candidate for a basic feasible solution corresponds to setting $x_1 = x_3 = 0$, with the basis $\boldsymbol{B} = [\boldsymbol{a}_2, \boldsymbol{a}_4]$
  This results in $\boldsymbol{x} = [0, -11/5, 0, -28/5]^T$, which is a basic solution that is not feasible.
- Finally, the sixth candidate for a basic feasible solution is obtained by setting $x_1 = x_2 = 0$. This results in the basis $\boldsymbol{B} = [\boldsymbol{a}_3, \boldsymbol{a}_4]$, and $\boldsymbol{x} = [0, 0, 11/3, -8/3]^T$, which is a basic solution but is not feasible.

Properties of Basic Solutions

- Definition 15.2: Any vector $\boldsymbol{x}$ that yields the minimum value of the objective function $\boldsymbol{c}^T\boldsymbol{x}$ over the set of vectors satisfying the constraints $\boldsymbol{Ax} = \boldsymbol{b}, \boldsymbol{x} \geq 0$, is said to be an ***optimal feasible solution***. An optimal feasible solution that is basic is said to be an ***optimal basic feasible solution***.
- Theorem 15.1: ***Fundamental Theorem of LP***. Consider a linear program in standard form
  - 1. If there exists a feasible solution, then there exists a basic feasible solution
  - 2. If there exists an optimal feasible solution, then there exists an optimal basic feasible solution.

Proof of Theorem 15.1

- Suppose that $\boldsymbol{x} = [x_1, ..., x_n]^T$ is a feasible solution and it has p positive components. Without loss of generality, we can assume that the first p components are positive, whereas the remaining components are zero. Then, in terms of the columns of $\boldsymbol{A} = [\boldsymbol{a}_1, ..., \boldsymbol{a}_p, ..., \boldsymbol{a}_n]$, this solution satisfies

$$x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_p\boldsymbol{a}_p = \boldsymbol{b}$$

  There are now two cases to consider.

- Case 1: If $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_p$ are linearly independent, then $p \leq m$. If $p = m$, then the solution $\boldsymbol{x}$ is basic and the proof is done. If $p < m$ then, since $rank(\boldsymbol{A}) = m$, we can find $m - p$ columns of $\boldsymbol{A}$ from the remaining $n - p$ columns so that the resulting set of m columns forms a basis. Hence, the solution $\boldsymbol{x}$ is a (degenerate) basic feasible solution corresponding to the basis above.

Proof of Theorem 15.1

- Case 2: Assume that $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_p$ are linearly dependent. Then, there exist numbers $y_i, i = 1, ..., p$, not all zero, such that

$$y_1\boldsymbol{a}_1 + y_2\boldsymbol{a}_2 + \cdots + y_p\boldsymbol{a}_p = 0$$

We can assume that there exists at least one y_i that is positive, for if all the y_i are nonpositive, we can multiply the equation above by -1. Multiply the equation by a scalar ϵ and subtract the resulting equation from $x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_p\boldsymbol{a}_p = \boldsymbol{b}$ to obtain

$$(x_1 - \epsilon y_1)\boldsymbol{a}_1 + (x_2 - \epsilon y_2)\boldsymbol{a}_2 + \cdots + (x_p - \epsilon y_p)\boldsymbol{a}_p = \boldsymbol{b}$$

Let $\boldsymbol{y} = [y_1, ..., y_p, 0, ..., 0]^T$. Then, for any ϵ we can write $\boldsymbol{A}[\boldsymbol{x} - \epsilon\boldsymbol{y}] = \boldsymbol{b}$

Proof of Theorem 15.1

- Let $\epsilon = min\{x_i/y_i : i = 1, ..., p, y_i > 0\}$. Then, the first p components of $\boldsymbol{x} - \epsilon \boldsymbol{y}$ are nonnegative, and at least one of these components is zero. We then have a feasible solution with at most $p - 1$ positive components. We can repeat this process until we get linearly independent columns of $\boldsymbol{A}$, after which we are back to case 1. Therefore, part 1 is proved.

Proof of Theorem 15.1

- We now prove part 2. Suppose that $\boldsymbol{x} = [x_1, ..., x_n]^T$ is an optimal feasible solution and only the first p variables are nonzero. Then, we have two cases to consider.
- The first case is exactly the same as part 1.
- The second case follows the same arguments as in part 1, but in addition we must show that $\boldsymbol{x} - \epsilon \boldsymbol{y}$ is optimal for any ϵ. We do this by showing that $\boldsymbol{c}^T \boldsymbol{y} = 0$. To this end, assume that $\boldsymbol{c}^T \boldsymbol{y} \neq 0$ Note that for ϵ of sufficiently small magnitude $(\epsilon \leq min(|x_i/y_i| : i = 1, ..., p, y_i \neq 0))$, the vector $\boldsymbol{x} - \epsilon \boldsymbol{y}$ is feasible. We can choose ϵ such that $\boldsymbol{c}^T \boldsymbol{x} > \boldsymbol{c}^T \boldsymbol{x} - \epsilon \boldsymbol{c}^T \boldsymbol{y} = \boldsymbol{c}^T (\boldsymbol{x} - \epsilon \boldsymbol{y})$. This contradicts the optimality of $\boldsymbol{x}$. We can now use the procedure from part 1 to obtain an optimal basic feasible solution from a given optimal feasible solution.

Example

- Consider the system of equations

$$\boldsymbol{A} = \begin{bmatrix} 2 & 3 & -1 & -1 \\ 4 & 1 & 1 & -2 \end{bmatrix} \qquad \boldsymbol{b} = \begin{bmatrix} -1 \\ 9 \end{bmatrix}$$

  Find a nonbasic feasible solution to this system and use the method in the proof of the fundamental theorem of LP to find a basic feasible solution.

- Recall that solutions for the system have the form

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \frac{14}{5} \\ -\frac{11}{5} \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -\frac{2}{5} \\ \frac{3}{5} \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} \frac{1}{2} \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

  where $s, t \in R$. Note that if $s = 4, t = 0$, then $\boldsymbol{x}_0 = \begin{bmatrix} \frac{6}{5} \\ \frac{1}{5} \\ 4 \\ 0 \end{bmatrix}$

Example

- There are constants $y_i, i = 1, 2, 3$, such that $y_1\boldsymbol{a}_1 + y_2\boldsymbol{a}_2 + y_3\boldsymbol{a}_3 = \mathbf{0}$
- For example, let $y_1 = -\frac{2}{5}, y_2 = \frac{3}{5}, y_3 = 1$. Note that $\boldsymbol{A}(\boldsymbol{x}_0 - \epsilon\boldsymbol{y}) = \boldsymbol{b}$ where $\boldsymbol{y} = [-\frac{2}{5}, \frac{3}{5}, 1, 0]^T$
  If $\epsilon = 1/3$, then

$$\boldsymbol{x}_1 = \boldsymbol{x}_0 - \epsilon\boldsymbol{y} = \begin{bmatrix} \frac{4}{3} \\ 0 \\ \frac{11}{3} \\ 0 \end{bmatrix}$$

  is a basic feasible solution.

Properties of Basic Solutions

- Observe that the fundamental theorem of LP reduces the task of solving a linear programming problem to that of searching over a finite number of basic feasible solution. That is, we need only check basic feasible solutions for optimality.
- As mentioned before, the total number of basic solutions is at most

$$\binom{n}{m} = \frac{n!}{m!(n-m)!}$$

Although this number is finite, it may be quite large. For example: $\binom{50}{5} = 2118760$

Therefore, a more efficient method of solving linear programs is needed. To this end, we next analyze a geometric interpretation of the fundamental theorem of LP.

Geometric View of Linear Programs

- Recall that a set $\Theta \subset R^n$ is said to be ***convex*** if, for every $\boldsymbol{x}, \boldsymbol{y} \in \Theta$ and every real number $\alpha, 0 < \alpha < 1$, the point $\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y} \in \Theta$
  In other words, a set is convex if given two points in the set, every point on the linear segment joining these two points is also a member of the set.
- Note that the set of points satisfying the constraints $\boldsymbol{Ax} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}$ is convex. To see this, let $\boldsymbol{x}_1, \boldsymbol{x}_2$ satisfy the constraints. Then, for all $\alpha \in (0,1)$, $\boldsymbol{A}(\alpha\boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) = \alpha\boldsymbol{Ax}_1 + (1-\alpha)\boldsymbol{Ax}_2 = \boldsymbol{b}$
  Also, for $\alpha \in (0,1)$, we have $\alpha\boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2 \geq \boldsymbol{0}$

Geometric View of Linear Programs

- Recall that a point $\boldsymbol{x}$ in a convex set Θ is said to be an ***extreme point*** of Θ if there are no two distinct points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ in Θ such that $\boldsymbol{x} = \alpha\boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2$ for some $\alpha \in (0,1)$. In other words, an extreme point is a point that does not lie strictly within the line segment connecting two other points of the set. Therefore, if $\boldsymbol{x}$ is an extreme point, and $\boldsymbol{x} = \alpha\boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2$ for some $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \Theta$ and $\alpha \in (0,1)$, then $\boldsymbol{x}_1 = \boldsymbol{x}_2$. In the following theorem we show that extreme points of the constraint set are equivalent to basic feasible solutions.

Geometric View of Linear Programs

- Theorem 15.2: Let Ω be the convex set consisting of all feasible solutions, that is, all n-vector $\boldsymbol{x}$ satisfying $\boldsymbol{Ax} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}$ where $\boldsymbol{A} \in R^{m \times n}$, $m < n$. Then, $\boldsymbol{x}$ is an extreme point of Ω if and only if $\boldsymbol{x}$ is a basic feasible solution to $\boldsymbol{Ax} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}$
- Form this theorem, it follows that the set of extreme points of the constraint set is equal to the set of basic feasible solutions.
- Combining this observation with the fundamental theorem of LP (Theorem 15.1), we can see that in solving linear programming problems we need only examine the extreme points of the constraint set.

Example

- Consider the following LP problem

$$\begin{aligned} \text{maximize } & 3x_1 + 5x_2 \\ \text{subject to } & x_1 + 5x_2 \leq 40 \\ & 2x_1 + x_2 \leq 20 \\ & x_1 + x_2 \leq 12 \\ & x_1, x_2 \geq 0 \end{aligned}$$

We introduce slack variables x_3, x_4, x_5 to convert this LP problem into standard form

$$\begin{aligned} \text{minimize } & -3x_1 - 5x_2 \\ \text{subject to } & x_1 + 5x_2 + x_3 = 40 \\ & 2x_1 + x_2 + x_4 = 20 \\ & x_1 + x_2 + x_5 = 12 \\ & x_1, x_2, ..., x_5 \geq 0 \end{aligned}$$

Example

- In the remainder of the example we consider only the problem in standard form. We can represent the constraints above as

$$x_1 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 5 \\ 1 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 40 \\ 20 \\ 12 \end{bmatrix}$$

  that is, $x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3 + x_4\boldsymbol{a}_4 + x_5\boldsymbol{a}_5 = \boldsymbol{b}, \boldsymbol{x} \geq \mathbf{0}$. Note that $\boldsymbol{x} = [0, 0, 40, 20, 12]^T$ is a feasible solution. But for this $\boldsymbol{x}$, the value of the objective function is zero. We already know that the minimum of the objective function (if it exists) is achieved at an extreme point of the constraint set Ω defined by the constraints. The point $[0, 0, 40, 20, 12]^T$ is an extreme point of the set of feasible solutions, but it turns out that it does not minimize the objective function.

Example

- Therefore, we need to seek the solution among the other extreme points. To do this we move from one extreme point to an adjacent extreme point such that the value of the objective function decreases. Here, we define two extreme points to be adjacent if the corresponding basic columns differ by only one vector.
- We begin with $\boldsymbol{x} = [0, 0, 40, 20, 12]^T$. We have $0\boldsymbol{a}_1 + 0\boldsymbol{a}_2 + 40\boldsymbol{a}_3 + 20\boldsymbol{a}_4 + 12\boldsymbol{a}_5 = \boldsymbol{b}$. To select an adjacent extreme point, let us choose to include $\boldsymbol{a}_1$ as a basic column in the new basis. We need to remove either $\boldsymbol{a}_3$, $\boldsymbol{a}_4$, or $\boldsymbol{a}_5$ from the old basis.

Example

- We first express $\boldsymbol{a}_1$ as a linear combination of the old basic columns: $\boldsymbol{a}_1 = 1\boldsymbol{a}_3 + 2\boldsymbol{a}_4 + 1\boldsymbol{a}_5$. Multiplying both sides of this equation by $\epsilon_1 > 0$, we get $\epsilon_1\boldsymbol{a}_1 = \epsilon_1\boldsymbol{a}_3 + 2\epsilon_1\boldsymbol{a}_4 + \epsilon_1\boldsymbol{a}_5$
- We now add this equation to the equation $0\boldsymbol{a}_1 + 0\boldsymbol{a}_2 + 40\boldsymbol{a}_3 + 20\boldsymbol{a}_4 + 12\boldsymbol{a}_5 = \boldsymbol{b}$. Collecting terms yields

  $$\epsilon_1\boldsymbol{a}_1 + 0\boldsymbol{a}_2 + (40 - \epsilon_1)\boldsymbol{a}_3 + (20 - 2\epsilon_1)\boldsymbol{a}_4 + (12 - \epsilon_1)\boldsymbol{a}_5 = \boldsymbol{b}$$

  We want to choose ϵ_1 in such a way that each of the coefficients above is nonnegative and at the same time, one of the coefficients $\boldsymbol{a}_3$, $\boldsymbol{a}_4$, or $\boldsymbol{a}_5$ becomes zero. Clearly, $\epsilon_1 = 10$ does the job. The result is $10\boldsymbol{a}_1 + 30\boldsymbol{a}_3 + 2\boldsymbol{a}_5 = \boldsymbol{b}$. The corresponding basic feasible solution (extreme point) is $[10, 0, 30, 0, 2]^T$. For this solution, the objective function value is -30, which is an improvement relative to the objective function value at the old extreme point.

Example

- We now apply the same procedure as above to move to another adjacent extreme point, which hopefully further decreases the value of the objective function. This time, we choose $\boldsymbol{a}_2$ to enter the new basis. We have $\boldsymbol{a}_2 = \frac{1}{2}\boldsymbol{a}_1 + \frac{9}{2}\boldsymbol{a}_3 + \frac{1}{2}\boldsymbol{a}_5$ and $\left(10 - \frac{1}{2}\epsilon_2\right)\boldsymbol{a}_1 + \epsilon_2\boldsymbol{a}_2 + \left(30 - \frac{9}{2}\epsilon_2\right)\boldsymbol{a}_3 + \left(2 - \frac{1}{2}\epsilon_2\right)\boldsymbol{a}_5 = \boldsymbol{b}$

  Substituting $\epsilon_2 = 4$, we obtain $8\boldsymbol{a}_1 + 4\boldsymbol{a}_2 + 12\boldsymbol{a}_3 = \boldsymbol{b}$
  The solution is $[8, 4, 12, 0, 0]^T$ and the corresponding value of the objective function is -44, which is smaller than the value at the previous extreme point.

Example

- To complete the example we repeat the procedure once more. This time, we select $\boldsymbol{a}_4$ and express it as a combination of the vectors in the previous basis, $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, and $\boldsymbol{a}_3$: $\boldsymbol{a}_4 = \boldsymbol{a}_1 - \boldsymbol{a}_2 + 4\boldsymbol{a}_3$ and hence

$$(8 - \epsilon_3)\boldsymbol{a}_1 + (4 + \epsilon_3)\boldsymbol{a}_2 + (12 - 4\epsilon_3)\boldsymbol{a}_3 + \epsilon_3\boldsymbol{a}_4 = \boldsymbol{b}$$

  The largest permissible value for ϵ_3 is 3. The corresponding basic feasible solution is $[5, 7, 0, 3, 0]^T$, with an objective function of -50. The solution $[5, 7, 0, 3, 0]^T$ turns out to be an optimal solution to our problem in standard form. Hence, the solution to the original problem is $[5, 7]^T$, which we can easily obtain graphically.

Example

- The technique used in this example for moving from one extreme point to an adjacent extreme point is also used in the simplex method for solving LP problems. The simplex method is essentially a refined method of performing these manipulations.