MATH 4211/6211 – Optimization

Linear Programming

Xiaojing Ye
Department of Mathematics & Statistics
Georgia State University

The standard form of a **Linear Program (LP)**

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

where

- $\boldsymbol{x} \in \mathbb{R}^n$ is the unknown variable;
- $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ is given, where $m < n$ and rank$(\boldsymbol{A}) = m$;
- $\boldsymbol{b} \in \mathbb{R}^m$ is given.

Other forms of LP can be converted to the standard form.

For example, suppose an LP is given as

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

Then we can rewrite it as an equivalent standard form

$$\begin{aligned} \text{minimize} \quad & (-\boldsymbol{c})^\top \boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{b} \\ & \boldsymbol{x}, \boldsymbol{y} \ge \boldsymbol{0} \end{aligned}$$

where $[\boldsymbol{x}; \boldsymbol{y}] \in \mathbb{R}^{n+m}$ is the variable to solve for now, and $\boldsymbol{y}$ is called the **slack** variable.

If $[\boldsymbol{x}^*; \boldsymbol{y}^*]$ is a solution of the new problem, then $\boldsymbol{x}^*$ is a solution of the original problem.

Example. Convert the LP below into the standard form

$$\begin{aligned} \text{maximize} \quad & x_2 - x_1 \\ \text{subject to} \quad & 3x_1 = x_2 - 5 \\ & |x_2| \le 2 \\ & x_1 \le 0 \end{aligned}$$

- $x_1' := -x_1$, then $x_1 \le 0 \Longleftrightarrow x_1' \ge 0$
- $x_2 = u - v$ where $u, v \ge 0$
- $|x_2| \le 2 \Longleftrightarrow -2 \le u - v \le 2$
- Introduce slack variables $x_3, x_4 \ge 0$ such that $u - v + x_3 = 2$ and $u - v - x_4 = -2$

Example (cont). Now we obtain an equivalent problem:

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

Note that there are 5 variables now.

Example. Suppose a factory wants to manufacture 4 products with the cost and budget (availability), as well as their profits, given in the table below. The goal is to maximize total profit.

	P1	P2	P3	P4	availability
man weeks	1	2	1	2	20
kg of material A	6	5	3	2	100
boxes of material B	3	4	9	12	75
profit	6	4	7	5	

Solution. Let x_i be the quantity to manufacture product i, then the LP is

$$\begin{aligned} \text{maximize} \quad & 6x_1 + 4x_2 + 7x_3 + 5x_4 \\ \text{subject to} \quad & x_1 + 2x_2 + x_3 + 2x_4 \leq 20 \\ & 6x_1 + 5x_2 + 3x_3 + 2x_4 \leq 100 \\ & 3x_1 + 4x_2 + 9x_3 + 12x_3 \leq 75 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

Exercise: convert this into the standard form of LP.

Geometric interpretation of the feasible set $\Omega = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}, \boldsymbol{x} \ge \boldsymbol{0}\}$:

Let $\boldsymbol{a}_i^\top$ be the ith row of $\boldsymbol{A}$ for $i = 1, \dots, m$. Then $\boldsymbol{a}_i^\top \boldsymbol{x} \le b_i$ is a half space.

Hence $\boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}$ is the intersection of m half spaces, which is a **polyhedra**. So is Ω.

The objective $\boldsymbol{c}^\top \boldsymbol{x}$ is a plane with slope defined on Ω, and the optimal point $\boldsymbol{x}^*$ is the point in Ω that has the minimum value $\boldsymbol{c}^\top \boldsymbol{x}^*$.

Basic solutions of linear system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$.

Suppose $\boldsymbol{A} = [\boldsymbol{B}\ \boldsymbol{D}]$ where $\boldsymbol{B} \in \mathbb{R}^{m\times m}$ has $\mathrm{rank}(\boldsymbol{B}) = m$. Denote $\boldsymbol{x}_B = \boldsymbol{B}^{-1}\boldsymbol{b} \in \mathbb{R}^m$. Then

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x}_B \\ \boldsymbol{0} \end{bmatrix} \in \mathbb{R}^n$$

is called a **basic solution** of $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$.

Definition. We introduce the following terms:

- $\boldsymbol{x}_B \in \mathbb{R}^m$: basic variables

- If $\boldsymbol{x}_B$ has zero component(s), then $\boldsymbol{x}$ is called a degenerate basic solution

- $\boldsymbol{B}$: basic columns of $\boldsymbol{A}$. Note that in practice there are up to $\binom{n}{m}$ ways to choose the basic columns.

- $\boldsymbol{x}$ is called a feasible point if $\boldsymbol{x} \in \Omega := \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}\}$.

- $\boldsymbol{x}$ is called a basic feasible point if $\boldsymbol{x}$ is a basic solution and also feasible.

- $\boldsymbol{x}^* \in \Omega$ is called an optimal (feasible) solution if $\boldsymbol{c}^\top \boldsymbol{x}^* \leq \boldsymbol{c}^\top \boldsymbol{x}$ for all $\boldsymbol{x} \in \Omega$.

Properties of basic solutions

Theorem (Fundamental Theorem of LP). For an LP with nonempty feasible set Ω, the following statements hold:

(a) A basic feasible point exists.

(b) If LP has solution, then there exists an optimal basic feasible solution.

Proof. Part (a). Let $\boldsymbol{x} \in \Omega$ (since $\Omega \neq \varnothing$), i.e., $\boldsymbol{Ax} = \boldsymbol{b}$ and $\boldsymbol{x} \geq \mathbf{0}$. WLOG, assume $x_1, \dots, x_p > 0$ and $x_{p+1}, \dots, x_n = 0$. Then

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

where $\boldsymbol{A} = [\boldsymbol{a}_1, \dots, \boldsymbol{a}_p, \dots, \boldsymbol{a}_n]$.

If $\boldsymbol{a}_1, \dots, \boldsymbol{a}_p$ are linearly independent, then $p \leq m$ (since $\boldsymbol{A} \in \mathbb{R}^{m \times n}$), and hence $\boldsymbol{x}$ is a basic feasible point.

If $\boldsymbol{a}_1, \dots, \boldsymbol{a}_p$ are linearly dependent, then $\exists\, y_1, \dots, y_p \in \mathbb{R}$ not all zero such that

$$\boldsymbol{Ay} = y_1 \boldsymbol{a}_1 + \cdots + y_p \boldsymbol{a}_p = \mathbf{0}$$

where $\boldsymbol{y} = [y_1, \dots, y_p, 0, \dots, 0]^\top \in \mathbb{R}^n$ (WLOG we assume at least one $y_i > 0$ otherwise take $\boldsymbol{y} = -\boldsymbol{y}$), then we know $\boldsymbol{A}(\boldsymbol{x} - \varepsilon \boldsymbol{y}) = \boldsymbol{b}$ for all ε.

Proof (cont). Choose $\varepsilon = \min\{x_i/y_i : y_i > 0\}$, then $\boldsymbol{x} - \varepsilon\boldsymbol{y} \in \Omega$ and has only $p-1$ nonzero components.

We update $\boldsymbol{x}$ to $\boldsymbol{x} - \varepsilon\boldsymbol{y}$ (then $\boldsymbol{x} \geq 0$ and has $p-1$ positive components).

Repeat this procedure until we find the basic columns of $\boldsymbol{A}$ and its corresponding basic feasible point.

Proof (cont). Part (b). Let $\boldsymbol{x} \in \Omega$ be optimal, and again WLOG assume $x_1, \dots, x_p > 0$.

If $\boldsymbol{a}_1, \dots, \boldsymbol{a}_p$ are linearly independent, then $p \leq m$ and $\boldsymbol{x}$ is also a basic solution. Hence $\boldsymbol{x}$ is optimal, basic, and feasible.

If $\boldsymbol{a}_1, \dots, \boldsymbol{a}_p$ are linearly dependent, then with the same argument as in (a), $\exists\, y_1, \dots, y_p \in \mathbb{R}$ not all zero such that

$$\boldsymbol{A}\boldsymbol{y} = y_1\boldsymbol{a}_1 + \cdots + y_p\boldsymbol{a}_p = \boldsymbol{0}$$

If $\boldsymbol{c}^\top \boldsymbol{y} \neq 0$ (say > 0), then by choosing any $0 < \varepsilon \leq \min\{|x_i/y_i| : y_i \neq 0\}$, we have $\boldsymbol{x} - \varepsilon\boldsymbol{y} \in \Omega$ and $\boldsymbol{c}^\top(\boldsymbol{x} - \varepsilon\boldsymbol{y}) = \boldsymbol{c}^\top\boldsymbol{x} - \varepsilon\boldsymbol{c}^\top\boldsymbol{y} < \boldsymbol{c}^\top\boldsymbol{x}$, which contradicts to the optimality of $\boldsymbol{x}$. Hence $\boldsymbol{c}^\top\boldsymbol{y} = 0$.

So we can choose ε in the same way as in (a) to get $\boldsymbol{x} - \varepsilon\boldsymbol{y}$ which has $p - 1$ nonzeros. Repeat to obtain an optimal, basic, and feasible solution.

Theorem. Consider an LP with nonempty Ω, then x is an extreme point of Ω iff x is a basic feasible point.

In theory, we only need to examine the basic feasible points of Ω and find the optimal basic feasible solution of the LP.

However, there could be a large amount of extreme points to examine.

Now we study the **Simplex Method** designed for LP.

Recall basic row operations to linear system $\boldsymbol{Ax} = \boldsymbol{b}$:

- interchange two rows
- multiply one row by a real nonzero scalar
- adding a scalar multiple of one row to another row

Each of these operations corresponds to an invertible matrix $\boldsymbol{E} \in \mathbb{R}^m$ premultiplying to $\boldsymbol{A}$.

Also recall that, to solve a linear system $\boldsymbol{Ax} = \boldsymbol{b}$, we first form the augmented matrix $[\boldsymbol{A}, \boldsymbol{b}]$, then apply a series of row operations, say $\boldsymbol{E}_1, \dots, \boldsymbol{E}_t$, to the matrix to obtain

$$\boldsymbol{E}_t \cdots \boldsymbol{E}_1 [\boldsymbol{A}, \boldsymbol{b}] = [\boldsymbol{I}, \boldsymbol{D}, \boldsymbol{B}^{-1}\boldsymbol{b}]$$

where $\boldsymbol{D}$ is such that $\boldsymbol{E}_t \dots \boldsymbol{E}_1 \boldsymbol{A} = [\boldsymbol{I}, \boldsymbol{D}]$ and $\boldsymbol{B} = (\boldsymbol{E}_t \cdots \boldsymbol{E}_1)^{-1} \in \mathbb{R}^{m \times m}$ is invertible. Note $\boldsymbol{A} = [\boldsymbol{B}, \boldsymbol{BD}]$.

Then a particular solution of $\boldsymbol{Ax} = \boldsymbol{b}$ is $\boldsymbol{x} = [\boldsymbol{B}^{-1}\boldsymbol{b}; \boldsymbol{0}] \in \mathbb{R}^n$.

In addition, note that $[-\boldsymbol{D}\boldsymbol{x}_D; \boldsymbol{x}_D] \in \mathbb{R}^n$ is a solution of $[\boldsymbol{I}, \boldsymbol{D}]\boldsymbol{x} = \boldsymbol{0}$ (hence a solution of $\boldsymbol{Ax} = \boldsymbol{0}$) for any $\boldsymbol{x}_D \in \mathbb{R}^{n-m}$.

Therefore, any solution of $\boldsymbol{Ax} = \boldsymbol{b}$ can be written for some $\boldsymbol{x}_D \in \mathbb{R}^{n-m}$ as

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{B}^{-1}\boldsymbol{b} \\ \boldsymbol{0} \end{bmatrix} + \begin{bmatrix} -\boldsymbol{D}\boldsymbol{x}_D \\ \boldsymbol{x}_D \end{bmatrix}$$

Suppose we have applied basic row operations and converted $\boldsymbol{Ax} = \boldsymbol{b}$ to $[\boldsymbol{I}, \boldsymbol{Y}]\boldsymbol{x} = \boldsymbol{y}_0$ (called the **canonical form**) where

$$\boldsymbol{Y} = \begin{bmatrix} y_{1\,m+1} & \cdots & y_{1\,n} \\ \vdots & \ddots & \vdots \\ y_{m\,m+1} & \cdots & y_{m\,n} \end{bmatrix} \in \mathbb{R}^{m\times(n-m)} \qquad \boldsymbol{y}_0 = \begin{bmatrix} y_{10} \\ \vdots \\ y_{m0} \end{bmatrix} \in \mathbb{R}^m$$

Note that the column $\boldsymbol{y}_q = [y_{1q}; \dots; y_{mq}] \in \mathbb{R}^m$ ($q > m$) gives the coefficients to represent $\boldsymbol{a}_q$ using $\boldsymbol{a}_1, \dots, \boldsymbol{a}_m$, where $\boldsymbol{a}_i$ is the ith column of $\boldsymbol{A}$:

$$\boldsymbol{a}_q = y_{1q}\boldsymbol{a}_1 + \cdots + y_{mq}\boldsymbol{a}_m = [\boldsymbol{a}_1, \dots, \boldsymbol{a}_m]\boldsymbol{y}_q$$

since $\boldsymbol{A} = [\boldsymbol{a}_1, \dots, \boldsymbol{a}_m, \boldsymbol{a}_{m+1}, \dots, \boldsymbol{a}_n] = [\boldsymbol{B}, \boldsymbol{BY}]$.

Now we are using $\boldsymbol{a}_1, \dots, \boldsymbol{a}_m$ as basic columns, and $\boldsymbol{x} = [\boldsymbol{B}^{-1}\boldsymbol{b}; \boldsymbol{0}]$ is the corresponding basic solution to $\boldsymbol{Ax} = \boldsymbol{b}$. If $\boldsymbol{B}^{-1}\boldsymbol{b} \geq \boldsymbol{0}$, then $\boldsymbol{x}$ is a basic feasible point of the LP. We temporarily assume we started with a basic feasible point.

Now we want to move to another basic feasible point. To this end, we need to exchange one of $\boldsymbol{a}_1, \dots, \boldsymbol{a}_m$ with another $\boldsymbol{a}_q$ ($q > m$) to form the new basic columns, find the corresponding basic solution $\boldsymbol{x}$ of $\boldsymbol{Ax} = \boldsymbol{b}$, making sure that $\boldsymbol{x} \geq \boldsymbol{0}$.

Exchanging basic columns p and q where $p \leq m < q$ is simple: just apply basic row operations $[\boldsymbol{I}, \boldsymbol{Y}]$ so that the qth column becomes $\boldsymbol{e}_p$. Here $\boldsymbol{e}_p \in \mathbb{R}^n$ is the vector with 1 as the pth component and 0 elsewhere.

However, which $\boldsymbol{a}_q$ we should choose to enter the basic columns, and which $\boldsymbol{a}_p$ to leave?

Since $\boldsymbol{x} = [y_{10}; \dots; y_{m0}; 0; \dots; 0] \in \mathbb{R}^n_+$ is a basic solution corresponding to $\boldsymbol{Ax} = \boldsymbol{b}$, we know

$$y_{10}\boldsymbol{a}_1 + \cdots + y_{m0}\boldsymbol{a}_m = \boldsymbol{b}$$

Suppose we decide to let $\boldsymbol{a}_q$ enter the basic columns, then due to

$$\boldsymbol{a}_q = y_{1q}\boldsymbol{a}_1 + \cdots + y_{mq}\boldsymbol{a}_m = [\boldsymbol{a}_1, \dots, \boldsymbol{a}_m]\boldsymbol{y}_q$$

we know for any $\varepsilon \geq 0$ there is

$$(y_{10} - \varepsilon y_{1q})\boldsymbol{a}_1 + \cdots + (y_{m0} - \varepsilon y_{mq})\boldsymbol{a}_m + \varepsilon \boldsymbol{a}_q = \boldsymbol{b}$$

Note that $\boldsymbol{y}_q$ have positive and nonpositive components (if all are nonpositive then the problem is unbounded), then for $\varepsilon > 0$ gradually increasing from 0, one of the coefficients (say p) will become 0 first. We will choose $\boldsymbol{a}_p$ to leave the basic columns.

More precisely, we choose $\varepsilon = \min_i\{y_{i0}/y_{iq} : y_{iq} > 0\}$, and obtain the basic columns

$$\boldsymbol{a}_1, \dots, \widehat{\boldsymbol{a}}_p, \dots, \boldsymbol{a}_m, \boldsymbol{a}_q$$

Now we consider which $\boldsymbol{a}_q$ to enter the basic column.

Ideally, we should choose $\boldsymbol{a}_q$ such that $\boldsymbol{c}^\top \boldsymbol{x}$ decreases the most.

Recall that the current point is $\boldsymbol{x} = [\boldsymbol{y}_0; \boldsymbol{0}] \in \mathbb{R}^n$. If we choose $\boldsymbol{a}_q$ to enter, then the objective function becomes

$$\begin{aligned}\boldsymbol{c}^\top \left(\begin{bmatrix} \boldsymbol{y}_0 - \varepsilon \boldsymbol{y}_q \\ \boldsymbol{0} \end{bmatrix} + \varepsilon \boldsymbol{e}_q \right) &= z_0 + \varepsilon [c_q - (c_1 y_{1q} + \cdots + c_m y_{mq})] \\ &= z_0 + (c_q - z_q)\varepsilon\end{aligned}$$

where $z_i := \boldsymbol{c}_{[1:m]}^\top \boldsymbol{y}_i$ for $i = 0, 1, \ldots, m$.

If $c_q - z_q < 0$ for some q, then choosing $\boldsymbol{a}_q$ to enter can further reduce the objective function since $\varepsilon > 0$. If there are more than one such q, then choose the one with smallest index or the one with smallest $c_q - z_q$.

If $c_q - z_q \geq 0$ for all q, then the optimal basic point is reached.

These answered which $\boldsymbol{a}_q$ to enter, and when the simplex method should stop.

To simplify the process, we consider the matrix form of the simplex method.

Given the standard form of an LP

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

we form the **tableau** of the problem as the matrix

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{b} \\ \boldsymbol{c}^\top & 0 \end{bmatrix} = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{D} & \boldsymbol{b} \\ \boldsymbol{c}_B^\top & \boldsymbol{c}_D^\top & 0 \end{bmatrix} \in \mathbb{R}^{(m+1)\times(n+1)}$$

Note that the tableau contains all information of the LP.

Premultiplying the matrix $\begin{bmatrix} \boldsymbol{B}^{-1} & \boldsymbol{0} \\ \boldsymbol{0}^\top & 1 \end{bmatrix}$ to the tableau, we obtain

$$\begin{bmatrix} \boldsymbol{B}^{-1} & \boldsymbol{0} \\ \boldsymbol{0}^\top & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{B} & \boldsymbol{D} & \boldsymbol{b} \\ \boldsymbol{c}_B^\top & \boldsymbol{c}_D^\top & 0 \end{bmatrix} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{B}^{-1}\boldsymbol{D} & \boldsymbol{B}^{-1}\boldsymbol{b} \\ \boldsymbol{c}_B^\top & \boldsymbol{c}_D^\top & 0 \end{bmatrix}$$

Further premultiplying the matrix $\begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ -\boldsymbol{c}_B^\top & 1 \end{bmatrix}$ yields

$$\begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ -\boldsymbol{c}_B^\top & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & \boldsymbol{B}^{-1}\boldsymbol{D} & \boldsymbol{B}^{-1}\boldsymbol{b} \\ \boldsymbol{c}_B^\top & \boldsymbol{c}_D^\top & 0 \end{bmatrix} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{B}^{-1}\boldsymbol{D} & \boldsymbol{B}^{-1}\boldsymbol{b} \\ \boldsymbol{0}^\top & \boldsymbol{c}_D^\top - \boldsymbol{c}_B^\top\boldsymbol{B}^{-1}\boldsymbol{D} & -\boldsymbol{c}_B^\top\boldsymbol{B}^{-1}\boldsymbol{b} \end{bmatrix}$$

It is easy to check that

- $[\boldsymbol{B}^{-1}\boldsymbol{b}; \mathbf{0}] \in \mathbb{R}^n$ is the basic solution
- $\boldsymbol{c}_D^\top - \boldsymbol{c}_B^\top \boldsymbol{B}^{-1}\boldsymbol{D} \in \mathbb{R}^{n-m}$ contains the reduced cost coefficients, i.e., $z_q - c_q$ for $q = m+1, \ldots, n$
- $\boldsymbol{c}_B^\top \boldsymbol{B}^{-1}\boldsymbol{b} \in \mathbb{R}$ is the objective function

According to the discussion above, the simplex method executes the following actions in order:

- find the index, say q, of the most negative component among $\boldsymbol{c}_D^\top - \boldsymbol{c}_B^\top \boldsymbol{B}^{-1} \boldsymbol{D} \in \mathbb{R}^{n-m}$;

- find $p = \arg\min_i \{y_{i0}/y_{iq} : y_{iq} > 0\}$;

- pivot the tableau about the (p, q) entry by basic row operations so that the column becomes 0 except the (p, q) entry which is 1.

- $[1 : m, n+1]$ of the tableau gives the current basic solution (basic feasible point), the nonzeros in $[m + 1, 1 : n]$ are the reduced cost coefficients, and the $(m + 1, n + 1)$ entry is the negative of objective function.

Example. Consider the following linear programming problem:

$$\begin{aligned} \text{maximize} \quad & 7x_1 + 6x_2 \\ \text{subject to} \quad & 2x_1 + x_2 \le 3 \\ & x_1 + 4x_2 \le 4 \\ & x_1, x_2 \ge 0 \end{aligned}$$

Solve this LP using the simplex method.

Solution. We first convert this LP to the standard form:

$$\begin{aligned} \text{minimize} \quad & -7x_1 - 6x_2 \\ \text{subject to} \quad & 2x_1 + x_2 + x_3 = 3 \\ & x_1 + 4x_2 + x_4 = 4 \\ & x_1, x_2, x_3, x_4 \ge 0 \end{aligned}$$

Solution (cont). The tableau of this LP is

$$\begin{bmatrix} 2 & 1 & 1 & 0 & 3 \\ 1 & 4 & 0 & 1 & 4 \\ -7 & -6 & 0 & 0 & 0 \end{bmatrix}$$

where the upper left 2×4 submatrix corresponds to $\boldsymbol{A} = [\boldsymbol{a}_1, \dots, \boldsymbol{a}_4]$, $\boldsymbol{c} = [-7; -6, 0; 0]$, and $\boldsymbol{b} = [3; 4]$.

Note that $\boldsymbol{A}$ is already in the canonical form: $\boldsymbol{a}_3, \boldsymbol{a}_4$ are the basic columns, and $\boldsymbol{x} = [0; 0; 3; 4]$ is the basic solution of $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$.

Since -7 is the most negative term, we will let $\boldsymbol{a}_1$ enter the basic column. Comparing $3/2$ and $4/1$, we see the former is smaller and hence decide to let $\boldsymbol{a}_3$ leave the basic column.

Now pivoting about the $(1,1)$ entry, we obtain

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} & 0 & \frac{3}{2} \\ 0 & \frac{7}{2} & -\frac{1}{2} & 1 & \frac{5}{2} \\ 0 & -\frac{5}{2} & \frac{7}{2} & 0 & \frac{21}{2} \end{bmatrix}$$

Now we see $-5/2$ is the most negative term, so we let $\boldsymbol{a}_2$ enter the basic columns. Comparing $\frac{3/2}{1/2} = 3$ and $\frac{5/2}{7/2} = 5/7$, we let $\boldsymbol{a}_4$ leave the basic column. Then pivoting about the $(2,2)$ entry, we obtain

$$\begin{bmatrix} 1 & 0 & \frac{4}{7} & -\frac{1}{7} & \frac{8}{7} \\ 0 & 1 & -\frac{1}{7} & \frac{2}{7} & \frac{5}{7} \\ 0 & 0 & \frac{22}{7} & \frac{5}{7} & \frac{86}{7} \end{bmatrix}$$

There are no negative entries in the last row, so we stop. The optimal solution is $\boldsymbol{x}^* = [\frac{8}{7}, \frac{5}{7}, 0, 0]$ and optimal value is $-\frac{86}{7}$ of the LP in its *standard form*. These correspond to the optimal solution $x_1 = 8/7$, $x_2 = 5/7$, and optimal value $86/7$ of the *original problem*.

Starting the Simplex Method

The simplex method discussed above is referred to the **Phase II** which requires an initial point to be a basic feasible point.

Now we consider the **Phase I** which finds one of such basic feasible points. To this end, we introduce **artificial variables** $\boldsymbol{y} \in \mathbb{R}^m$ and consider the following **artificial problem**:

$$\begin{aligned} \text{minimize} \quad & \mathbf{1}^\top \boldsymbol{y} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} + \boldsymbol{E}\boldsymbol{y} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \mathbf{0}, \ \boldsymbol{y} \geq \mathbf{0} \end{aligned}$$

where $\mathbf{1} = [1, \dots, 1]^\top \in \mathbb{R}^m$, $\boldsymbol{E}$ is diagonal with $\boldsymbol{E}_{ii} = \text{sign}(b_i)$.

Note that the equality constraint is equivalent to $\boldsymbol{EAx} + \boldsymbol{y} = \boldsymbol{Eb} = |\boldsymbol{b}|$. That is, if $b_i < 0$, then multiply -1 to both sides of the ith equality constraint.

Then the artificial problem

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{1}^\top \boldsymbol{y} \\ \text{subject to} \quad & \boldsymbol{EAx} + \boldsymbol{y} = |\boldsymbol{b}| \\ & \boldsymbol{x} \geq \boldsymbol{0},\ \boldsymbol{y} \geq \boldsymbol{0} \end{aligned}$$

has an obvious basic feasible point $[\boldsymbol{x}; \boldsymbol{y}] = [0; \dots; 0; |b_1|; \dots; |b_m|]$, which we can use as an initial and apply the Phase II simplex method.

Moreover, the artificial problem in Phase I has a solution iff the original problem has nonempty feasible set Ω. In this case, the artificial problem above have an optimal objective value 0 and optimal solution $[\boldsymbol{x}^*; \boldsymbol{y}^*]$, such that $\boldsymbol{y}^* = \boldsymbol{0}$ and $\boldsymbol{x}^*$ is a basic feasible point of the original LP problem:

- If Ω is nonempty, then there exists $\boldsymbol{x} \in \Omega$, i.e., $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ and $\boldsymbol{x} \geq \boldsymbol{0}$. So $[\boldsymbol{x}; \boldsymbol{0}]$ solves the artificial problem in Phase I.

- If the artificial problem has a solution such that $\boldsymbol{1}^\top \boldsymbol{y} = 0$, then $\boldsymbol{y} = \boldsymbol{0}$, and $\boldsymbol{x}$ satisfies $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ and $\boldsymbol{x} \geq \boldsymbol{0}$, which means $\boldsymbol{x} \in \Omega$.

Revised Simplex Method

If $m \gg n$, it is wasteful to apply row operations to all columns since most of them do not give new basic variables. Instead, we can keep tracking $\boldsymbol{B}^{-1}$ and the current basic columns only.

Specifically, we start with only the $m \times (m+1)$ portion of $\boldsymbol{A} \in \mathbb{R}^{m\times(n+1)}$: $[\boldsymbol{I}, \boldsymbol{b}]$. We set $\boldsymbol{B} = \boldsymbol{B}^{-1} = \boldsymbol{I}$, and the rest of $\boldsymbol{A}$ as $\boldsymbol{D}$. Let $\boldsymbol{y}_0 = \boldsymbol{B}^{-1}\boldsymbol{b} = \boldsymbol{b}$, so we have $[\boldsymbol{B}^{-1}, \boldsymbol{y}_0] = [\boldsymbol{I}, \boldsymbol{b}]$ and record the corresponding basic variables.

In each iteration, we will need an updated $[\boldsymbol{B}^{-1}, \boldsymbol{y}_0]$ and the corresponding basic variables to do the followings.

Revised Simplex Method (continued)

Then we can compute $\boldsymbol{r}_D^\top := \boldsymbol{c}_D^\top - (\boldsymbol{c}_B^\top \boldsymbol{B}^{-1})\boldsymbol{D}$, and look for the most negative component of $\boldsymbol{r}_D^\top$ to decide which x_q to become new basic variable.

Note that in the original simplex method we needed to apply row operations to $\boldsymbol{D}$, but here we only compute $(\boldsymbol{c}_B^\top \boldsymbol{B}^{-1})\boldsymbol{D}$ which only require two matrix-vector multiplications.

We then compute $\boldsymbol{y}_q = \boldsymbol{B}^{-1}\boldsymbol{a}_q$ ($\boldsymbol{a}_q$ is the qth column of $\boldsymbol{A}$), and check $\boldsymbol{y}_0/\boldsymbol{y}_q$ componentwisely to find the smallest positive number, say p, then pivot $[\boldsymbol{B}^{-1}, \boldsymbol{y}_0, \boldsymbol{y}_q]$ on the pth component of $\boldsymbol{y}_q$ (so the last column becomes $\boldsymbol{e}_p$ after pivoting). Then delete the last column to obtain the updated $[\boldsymbol{B}^{-1}, \boldsymbol{y}_0]$ (and use x_q to switch x_p out of the basic variables) and repeat from top of this page.

Example (Revised simplex method). Consider the following linear program:

$$\begin{aligned} \text{maximize} \quad & 3x_1 + 5x_2 \\ \text{subject to} \quad & x_1 + x_2 \leq 4 \\ & 5x_1 + 3x_2 \geq 8 \\ & x_1, x_2 \geq 0 \end{aligned}$$

Solve this LP using the revised simplex method.

Solution. We first convert this LP to the standard form:

$$\begin{aligned} \text{minimize} \quad & -3x_1 - 5x_2 \\ \text{subject to} \quad & x_1 + x_2 + x_3 = 4 \\ & 5x_1 + 3x_2 - x_4 = 8 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

Solution (cont.) We first form the artificial problem $x_5 = y_1$:

$$\begin{aligned} \text{minimize} \quad & x_5 \\ \text{subject to} \quad & x_1 + x_2 + x_3 = 4 \\ & 5x_1 + 3x_2 - x_4 + x_5 = 8 \\ & x_1, x_2, x_3, x_4, x_5 \geq 0 \end{aligned}$$

with a basic feasible point $[x_1, x_2, x_3, x_4, x_5] = [0, 0, 4, 0, 8]$ to start.

The complete tableau for this artificial problem is

$\boldsymbol{a}_1$	$\boldsymbol{a}_2$	$\boldsymbol{a}_3$	$\boldsymbol{a}_4$	$\boldsymbol{a}_5$	$\boldsymbol{b}$
1	1	1	0	0	4
5	3	0	-1	1	8
0	0	0	0	1	0

where the last row contains $\boldsymbol{c}^\top = [0, 0, 0, 0, 1]$.

Solution (cont.) The initial $\boldsymbol{B}^{-1} = \boldsymbol{I}$, $\boldsymbol{y}_0 = [4; 8]$, and the basic variables are x_3, x_5.

Then we check

$$\begin{aligned} \boldsymbol{r}_D^\top &= \boldsymbol{c}_D^\top - (\boldsymbol{c}_B^\top \boldsymbol{B}^{-1})\boldsymbol{D} \\ &= [0, 0, 0] - \left([0, 1] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} 1 & 1 & 0 \\ 5 & 3 & -1 \end{bmatrix} \\ &= [-5, -3, 1] = [r_1, r_2, r_4] \end{aligned}$$

So we decide to let x_1 become a new basic variable since r_1 is most negative.

Compute $\boldsymbol{y}_1 = \boldsymbol{B}^{-1}\boldsymbol{a}_1 = [1; 5]$, then $\boldsymbol{y}_0/\boldsymbol{y}_1 = [4, 8/5]$ so we let $\boldsymbol{x}_5$ not be basic variable anymore since $8/5$ is smaller.

Solution (cont.) Now we do pivoting on the 2nd component of $\boldsymbol{y}_1$:

$$[\boldsymbol{B}^{-1}, \boldsymbol{y}_0, \boldsymbol{y}_1] = \begin{bmatrix} 1 & 0 & 4 & 1 \\ 0 & 1 & 8 & 5 \end{bmatrix} \to \begin{bmatrix} 1 & -\frac{1}{5} & \frac{12}{5} & 0 \\ 0 & \frac{1}{5} & \frac{8}{5} & 1 \end{bmatrix}$$

Then the updated $[\boldsymbol{B}^{-1}, \boldsymbol{y}_0]$ is $\begin{bmatrix} 1 & -\frac{1}{5} & \frac{12}{5} \\ 0 & \frac{1}{5} & \frac{8}{5} \end{bmatrix}$ and the corresponding basic variables are x_3, x_1.

Then we check

$$\begin{aligned} \boldsymbol{r}_D^\top &= \boldsymbol{c}_D^\top - (\boldsymbol{c}_B^\top \boldsymbol{B}^{-1})\boldsymbol{D} \\ &= [0, 0, 1] - \left([0, 0] \begin{bmatrix} 1 & -\frac{1}{5} \\ 0 & \frac{1}{5} \end{bmatrix} \right) \begin{bmatrix} 1 & 0 & 0 \\ 3 & -1 & 1 \end{bmatrix} \\ &= [0, 0, 1] = [r_2, r_4, r_5] \geq \boldsymbol{0} \end{aligned}$$

So the Phase I is completed, and we get solution $[x_1, \ldots, x_5] = [\frac{8}{5}, 0, \frac{12}{5}, 0, 0]$.

Solution (cont.) Now we start Phase II. The complete tableau is

$\boldsymbol{a}_1$	$\boldsymbol{a}_2$	$\boldsymbol{a}_3$	$\boldsymbol{a}_4$	$\boldsymbol{b}$
1	1	1	0	4
5	3	0	−1	8
−3	−5	0	0	0

where the last row contains $\boldsymbol{c}^\top = [-3, -5, 0, 0]$.

We also have basic variables $x_3 = \frac{12}{5}, x_1 = \frac{8}{5}$ and

$$[\boldsymbol{B}^{-1}, \boldsymbol{y}_0] = \begin{bmatrix} 1 & -\frac{1}{5} & \frac{12}{5} \\ 0 & \frac{1}{5} & \frac{8}{5} \end{bmatrix}$$

Solution (cont.) Then we check

$$\begin{aligned} \boldsymbol{r}_D^\top &= \boldsymbol{c}_D^\top - (\boldsymbol{c}_B^\top \boldsymbol{B}^{-1})\boldsymbol{D} \\ &= [-5, 0] - \left([0, -3] \begin{bmatrix} 1 & -\frac{1}{5} \\ 0 & \frac{1}{5} \end{bmatrix} \right) \begin{bmatrix} 1 & 0 \\ 3 & -1 \end{bmatrix} \\ &= [-\frac{16}{5}, -\frac{3}{5}] = [r_2, r_4] \end{aligned}$$

So we let x_2 become a new basic variable.

Compute $\boldsymbol{y}_2 = \boldsymbol{B}^{-1}\boldsymbol{a}_2 = [\frac{2}{5}; \frac{3}{5}]$, then $\boldsymbol{y}_0/\boldsymbol{y}_1 = [6, \frac{8}{3}]$ so we let x_1 not be basic variable anymore since $\frac{8}{3}$ is smaller.

Solution (cont.) Now we do pivoting on the 2nd component of $\boldsymbol{y}_2$:

$$[\boldsymbol{B}^{-1}, \boldsymbol{y}_0, \boldsymbol{y}_2] = \begin{bmatrix} 1 & -\frac{1}{5} & \frac{12}{5} & \frac{2}{5} \\ 0 & \frac{1}{5} & \frac{8}{5} & \frac{3}{5} \end{bmatrix} \to \begin{bmatrix} 1 & -\frac{1}{3} & \frac{4}{3} & 0 \\ 0 & \frac{1}{3} & \frac{8}{3} & 1 \end{bmatrix}$$

Then the updated $[\boldsymbol{B}^{-1}, \boldsymbol{y}_0]$ is $\begin{bmatrix} 1 & -\frac{1}{3} & \frac{4}{3} \\ 0 & \frac{1}{3} & \frac{8}{3} \end{bmatrix}$ and the corresponding basic variables are $x_3 = \frac{4}{3}, x_2 = \frac{8}{3}$.

Then we check

$$\begin{aligned} \boldsymbol{r}_D^\top &= \boldsymbol{c}_D^\top - (\boldsymbol{c}_B^\top \boldsymbol{B}^{-1})\boldsymbol{D} \\ &= [-3, 0] - \left([0, -5] \begin{bmatrix} 1 & -\frac{1}{3} \\ 0 & \frac{1}{3} \end{bmatrix}\right) \begin{bmatrix} 1 & 0 \\ 5 & -1 \end{bmatrix} \\ &= [\frac{16}{3}, -\frac{5}{3}] = [r_1, r_4] \end{aligned}$$

So we let x_4 be a new basic variable.

Solution (cont.) Compute $\boldsymbol{y}_4 = \boldsymbol{B}^{-1}\boldsymbol{a}_4 = [\frac{1}{3}; -\frac{1}{3}]$, then $\boldsymbol{y}_0/\boldsymbol{y}_4 = [4, -8]$ so we let x_3 not be basic variable anymore.

Now we do pivoting on the first component of $\boldsymbol{y}_4$:

$$[\boldsymbol{B}^{-1}, \boldsymbol{y}_0, \boldsymbol{y}_4] = \begin{bmatrix} 1 & -\frac{1}{3} & \frac{4}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{8}{3} & -\frac{1}{3} \end{bmatrix} \to \begin{bmatrix} 3 & -1 & 4 & 1 \\ 1 & 0 & 4 & 0 \end{bmatrix}$$

Then the updated $[\boldsymbol{B}^{-1}, \boldsymbol{y}_0]$ is $\begin{bmatrix} 3 & -1 & 4 \\ 1 & 0 & 4 \end{bmatrix}$ and the corresponding basic variables are $x_4 = 4, x_2 = 4$. Then we check

$$\begin{aligned} \boldsymbol{r}_D^\top &= \boldsymbol{c}_D^\top - (\boldsymbol{c}_B^\top \boldsymbol{B}^{-1})\boldsymbol{D} \\ &= [-3, 0] - \left([0, -5] \begin{bmatrix} 3 & -1 \\ 1 & 0 \end{bmatrix}\right) \begin{bmatrix} 1 & 1 \\ 5 & 0 \end{bmatrix} \\ &= [2, 5] = [r_1, r_3] \geq \boldsymbol{0} \end{aligned}$$

So Phase II is finished, and solution is $\boldsymbol{x} = [0; 4; 0; 4]$. Hence the solution for the original problem is then $[x_1, x_2] = [0, 4]$.