The Simplex Method

Yinyu Ye
Department of Management Science and Engineering
Stanford University
Stanford, CA 94305, U.S.A.

http://www.stanford.edu/~yyye

(LY, Chapters 2.3-2.5, 3.1-3.4)

Geometry of Linear Programming (LP)

Consider the following LP problem:

$$
\begin{array}{llll}
\text{maximize} & x_1 & +2x_2 & \\
\text{subject to} & x_1 & & \leq 1 \\
 & & x_2 & \leq 1 \\
 & x_1 & +x_2 & \leq 1.5 \\
 & x_1, & x_2 & \geq 0.
\end{array}
$$

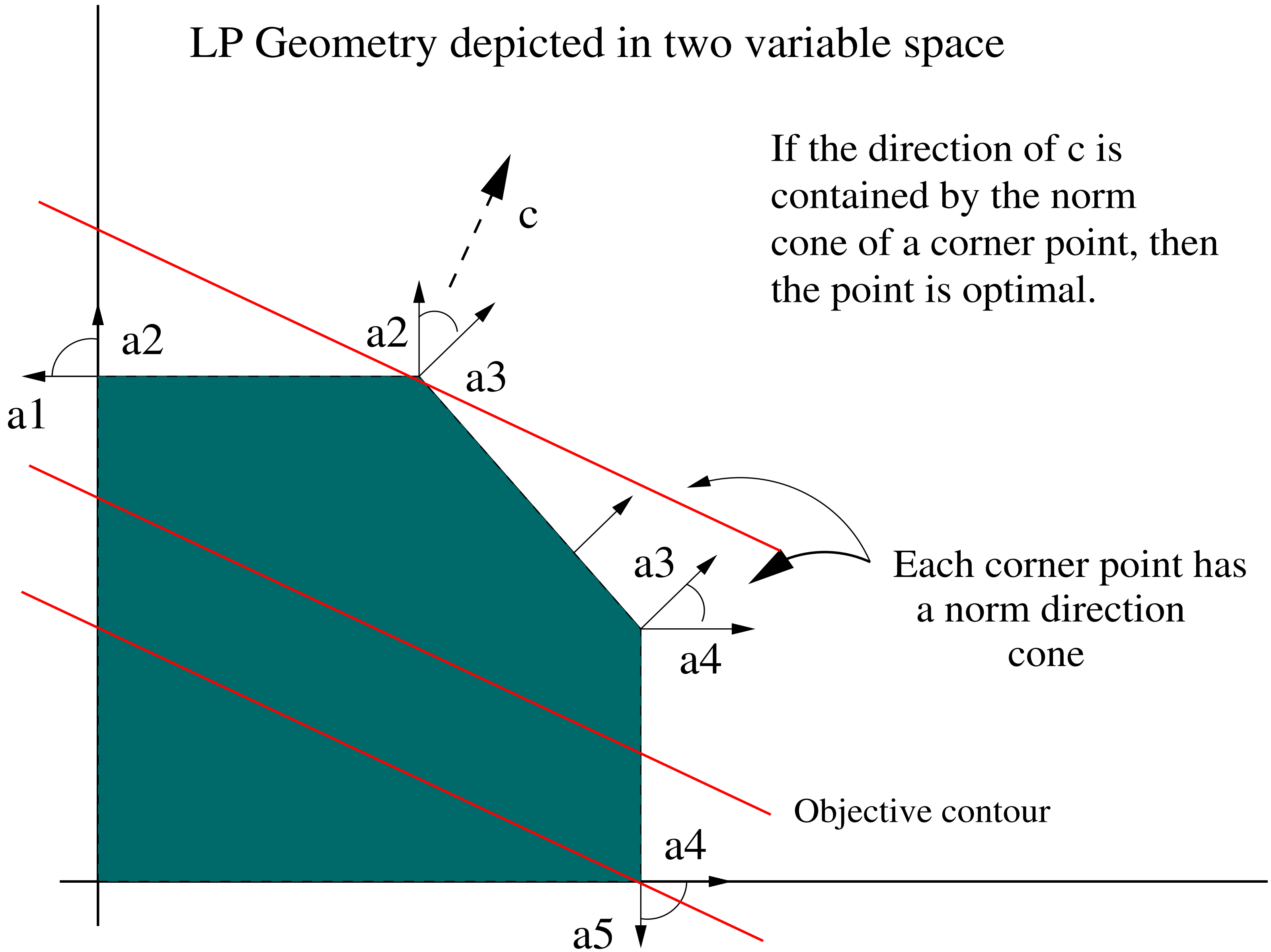

Figure 1: Feasible region with objective contours

- Solution (decision, point): any specification of values for all decision variables, regardless of whether it is a desirable or even allowable choice.
- Feasible Solution: a solution for which all the constraints are satisfied.
- Feasible Region (constraint set, feasible set): the collection of all feasible solution.
- Interior, Boundary, and Face
- Extreme Point or Corner Point or Vertex
- Objective Function Contour: iso-profit (or iso-cost) line.
- Optimal Solution: a feasible solution that has the most favorable objective value.
- Optimal Objective Value: the value of the objective function evaluated at an optimal solution.
- Active Constraint: binding constraint.

Definition of Face and Extreme Points

- Let P be a polyhedron in $\mathcal{R}^n$, then F is a face of P if and only if there is a vector $\mathbf{b}$ for which F is the set of points attaining $\max\{\mathbf{b}^T\mathbf{y} : \mathbf{y} \in P\}$ provided this maximum is finite.
- A polyhedron has only finite many faces; each face is a nonempty polyhedron.
- A vector $\mathbf{y} \in P$ is an extreme point or a vertex of P if $\mathbf{y}$ is not a convex combination of two distinct feasible points.

Theory of LP

All LP problems fall into one of the following three classes:

- Problem is infeasible: feasible region is empty.
- Problem is unbounded: feasible region is unbounded towards the optimizing direction.
- Problem is feasible and bounded:
  - There exists an optimal solution or optimizer.
  - There may be a unique optimizer or multiple optimizers.
  - All optimizers are on a face of the feasible region.
  - There is always at least one corner (extreme) optimizer if the face has a corner.
  - If a corner point is not worse than all its adjacent or neighboring corners, then it is optimal.

History of the Simplex Method

George B. Dantzig's Simplex Method for LP stands as one of the most significant algorithmic achievements of the 20th century. It is now over 50 years old and still going strong.

The basic idea of the simplex method to confine the search to corner points of the feasible region (of which there are only finitely many) in a most intelligent way. In contrast, interior-point methods will move in the interior of the feasible region, hoping to by-pass many corner points on the boundary of the region.

The key for the simplex method is to make computers see corner points; and the key for interior-point methods is to stay in the interior of the feasible region.

From Geometry to Algebra

- How to make computer recognize a corner point?
- How to make computer tell that two corners are neighboring?
- How to make computer terminate and declare optimality?
- How to make computer identify a better neighboring corner?

LP Standard Form

$$
\begin{array}{rl}
\text{minimize} & c_1x_1 + c_2x_2 + \ldots + c_nx_n \\
\text{subject to} & a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1, \\
& a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2, \\
& \vdots \\
& a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \\
& x_j \geq 0, \quad j = 1, 2, \ldots, n.
\end{array}
$$

Equivalently,

$$
\begin{array}{lll}
(LP) & \text{minimize} & \mathbf{c}^T\mathbf{x} \\
& \text{subject to} & A\mathbf{x} = \mathbf{b}, \\
& & \mathbf{x} \geq \mathbf{0}.
\end{array}
$$

Reduction to the Standard Form

- Eliminating "free" variables: substitute with the difference of two nonnegative variables

$$x = x^+ - x^- \quad \text{where} \quad x^+, x^- \geq 0.$$

- Eliminating inequalities: add slack variables

$$\mathbf{a}^T\mathbf{x} \leq b \quad \Leftrightarrow \quad \mathbf{a}^T\mathbf{x} + s = b, \quad s \geq 0,$$

$$\mathbf{a}^T\mathbf{x} \geq b \quad \Leftrightarrow \quad \mathbf{a}^T\mathbf{x} - s = b, \quad s \geq 0.$$

- Eliminating upper bounds: move them to constraints

$$x \leq 3 \quad \Leftrightarrow \quad x + s = 3, \quad s \geq 0.$$

- Eliminating nonzezro lower bounds: shift the decision variables

$$x \geq 3 \quad \Rightarrow \quad x := x - 3.$$

- Change $\max\ \mathbf{c}^T\mathbf{x}$ to $\min\ -\mathbf{c}^T\mathbf{x}$.

An LP Example in Standard Form

$$
\begin{array}{llllllll}
\text{minimize} & -x_1 & -2x_2 & & & & \\
\text{subject to} & x_1 & & +x_3 & & & = 1 \\
 & & x_2 & & +x_4 & & = 1 \\
 & x_1 & +x_2 & & & +x_5 & = 1.5 \\
 & x_1, & x_2, & x_3, & x_4, & x_5 & \geq 0.
\end{array}
$$

Basic and Basic Feasible Solution (BFS)

In the LP standard form, let's assume that we selected m linearly independent columns, denoted by the index set B from A and solve

$$A_B\mathbf{x}_B = \mathbf{b}$$

for the m-vector $\mathbf{x}_B$. By setting the variables $\mathbf{x}_N$ of $\mathbf{x}$ corresponding to the remaining columns of A equal to zero, we obtain a solution $\mathbf{x}$ of $A\mathbf{x} = \mathbf{b}.$

Then $\mathbf{x}$ is said to be a basic solution to (LP) with respect to basis A_B. The components of $\mathbf{x}_B$ are called basic variables and those of $\mathbf{x}_N$ are called nonbasic variables.

Two basic solutions are adjacent if they differ by exactly one basic (or nonbasic) variable. If a basic solution satisfies $\mathbf{x}_B \geq \mathbf{0}$, then $\mathbf{x}$ is called a basic feasible solution (BFS), and it is an extreme point of the feasible region.

If one or more components in $\mathbf{x}_B$ has value zero, $\mathbf{x}$ is said to be degenerate.

Geometry vs Algebra

Theorem 1 *Consider the polyhedron in the standard LP form. Then a basic feasible solution and a corner point are equivalent; the former is algebraic and the latter is geometric.*

In the LP example:

Basis	3,4,5	1,4,5	3,4,1	3,2,5	3,4,2	1,2,3	1,2,4	1,2,5
Feasible?	$\surd$	$\surd$		$\surd$		$\surd$	$\surd$	
x_1, x_2	$0, 0$	$1, 0$	$1.5, 0$	$0, 1$	$0, 1.5$	$.5, 1$	$1, .5$	$1, 1$

Theorem 2 *(The Fundamental Theorem of LP in Algebraic form)*[a] *Given (LP) and (LD) where A has full row rank m,*

i) *if there is a feasible solution, there is a basic feasible solution;*

ii) *if there is an optimal solution, there is an optimal basic solution.*

The simplex method is to proceed from one BFS (a corner point of the feasible region) to an adjacent or neighboring one, in such a way as to continuously improve the value of the objective function.

---

[a] Text p.38

Rounding to a Basic Feasible Solution for LP

Step 1: Start with any feasible solution x^0 and without loss of generality, assume $x^0 > 0$. Let $k = 0$ and $A^0 = A$.

Step 2: Find any $A^k d = 0, \; d \neq 0$, and let $x^{k+1} := x^k + \alpha d$ where α is chosen so that $x^{k+1} \geq 0$ and at least one of x^{k+1} equals 0.

Step 3: Eliminate the the variable(s) in x^{k+1} and column(s) in A^k corresponding to $x_j^{k+1} = 0$ and let the new matrix be A^{k+1}.

Step 4: Return to Step 2.

Neighboring Basic Solutions

- Two basic solutions are neighboring or adjacent if they differ by exactly one basic (or nonbasic) variable.
- A basic feasible solution is optimal if no ”better” neighboring basic feasible solution exists.

Optimality test

Consider a BFS $(x_1, x_2, x_3, x_4, x_5) = (0, 0, 1, 1, 1.5)$. It's NOT optimal to the problem

$$
\begin{array}{llllllll}
\text{minimize} & -x_1 & -2x_2 & & & & \\
\text{subject to} & x_1 & & +x_3 & & & = 1 \\
& & x_2 & & +x_4 & & = 1 \\
& x_1 & +x_2 & & & +x_5 & = 1.5 \\
& x_1, & x_2, & x_3, & x_4, & x_5 & \geq 0;
\end{array}
$$

but it's optimal to the below problem

$$
\begin{array}{llllllll}
\text{minimize} & x_1 & +2x_2 & & & & \\
\text{subject to} & x_1 & & +x_3 & & & = 1 \\
& & x_2 & & +x_4 & & = 1 \\
& x_1 & +x_2 & & & +x_5 & = 1.5 \\
& x_1, & x_2, & x_3, & x_4, & x_5 & \geq 0.
\end{array}
$$

At a BFS, when the objective coefficients to all the basic variables are zero and those to all the non-basic variables are non-negative, then the BFS is optimal.

LP Canonical Form

An LP in the standard form is said to be in canonical form at a BFS if

- The objective coefficients to all the basic variables are zero, that is, the objective function contains only non-basic variables and
- The constraint matrix for the basic variables form an identity matrix (with some permutation if necessary).

It's easy to see whether or not the current BFS is optimal or not, if the LP is in canonical form.

Is there a way to transform the original LP problem to an equivalent LP in the canonical form? That is, is there an algorithm that the feasible set and optimal solution set remain the same when compared to the original LP?

Transforming to canonical form

Suppose the basis of a BFS is A_B and the rest is A_N. One can transform the equality constraint to

$$A_B^{-1}A\mathbf{x} = A_B^{-1}\mathbf{b}, \quad (\text{so that } \mathbf{x}_B = A_B^{-1}\mathbf{b} - A_B^{-1}A_N\mathbf{x}_N)$$

That is, we express $\mathbf{x}_B$ in terms of $\mathbf{x}_N$, the degree of freedom variables. Then the objective function becomes

$$\begin{aligned} \mathbf{c}^T\mathbf{x} = \mathbf{c}_B^T\mathbf{x}_B + \mathbf{c}_N^T\mathbf{x}_N &= \mathbf{c}_B^T A_B^{-1}\mathbf{b} - \mathbf{c}_B^T A_B^{-1}A_N\mathbf{x}_N + \mathbf{c}_N^T\mathbf{x}_N \\ &= \mathbf{c}_B^T A_B^{-1}\mathbf{b} + (\mathbf{c}_N^T - \mathbf{c}_B^T A_B^{-1}A_N)\mathbf{x}_N. \end{aligned}$$

The transformed problem is then in canonical form with respect to the BFS A_B; and it's equivalent to the original LP. That is, its feasible set and optimal solution set don't change.

Equivalent LP

By ignoring the constant term in the objective function, we have an equivalent intermediate LP problem in canonical form:

$$
\begin{array}{ll}
\text{minimize} & \mathbf{r}^T\mathbf{x} \\
\text{subject to} & \bar{A}\mathbf{x} = \bar{\mathbf{b}}, \\
 & \mathbf{x} \geq \mathbf{0},
\end{array}
$$

where

$$
\mathbf{r}_B := \mathbf{0}, \quad \mathbf{r}_N := \mathbf{c}_N - A_N^T(A_B^{-1})^T\mathbf{c}_B, \quad \bar{A} := A_B^{-1}A, \quad \bar{\mathbf{b}} := A_B^{-1}\mathbf{b}.
$$

Optimality Test Revisited

The vector $\mathbf{r} \in R^n$ defined by

$$\mathbf{r} = \mathbf{c} - \bar{A}^T \mathbf{c}_B = \mathbf{c} - A^T (A_B^{-1})^T \mathbf{c}_B$$

is called the reduced cost coefficient vector.

Often we represent $(A_B^{-1})^T \mathbf{c}_B$ by shadow price vector $\mathbf{y}$. That is,

$$A_B^T \mathbf{y} = \mathbf{c}_B, \quad (\text{so that } \mathbf{r} = \mathbf{c} - A^T \mathbf{y}).$$

Theorem 3 *If $\mathbf{r}_N \geq \mathbf{0}$ (equivalently $\mathbf{r} \geq \mathbf{0}$) at a BFS with basic variable set B, then the BFS $\mathbf{x}$ is an optimal basic solution and A_B is an optimal basis.*

In the LP Example, let the basic variable set $B = \{1, 2, 3\}$ so that

$$A_B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$$

and

$$A_B^{-1} = \begin{pmatrix} 0 & -1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & -1 \end{pmatrix}$$

$$\mathbf{y}^T = (0,\ -1,\ -1) \quad \text{and} \quad \mathbf{r}^T = (0,\ 0,\ 0,\ 1,\ 1)$$

The corresponding optimal corner solution is $(x_1, x_2) = (0.5, 1)$ in the original problem.

Canonical Tableau

The intermediate canonical form data can be organized in a tableau:

B	$\mathbf{r}^T$	$-\mathbf{c}_B^T\bar{\mathbf{b}}$
Basis Indices	$\bar{A}$	$\bar{\mathbf{b}}$

The up-right corner quantity represents negative of the objective value at the current basic solution.

With the initial basic variable set $B = \{3, 4, 5\}$ for the example, the canonical tableau is

B	-1	-2	0	0	0	0
3	1	0	1	0	0	1
4	0	1	0	1	0	1
5	1	1	0	0	1	$\frac{3}{2}$

Moving to a ”Better” Neighboring Corner

As we saw in the previous slides, the BFS given by the canonical form is optimal if the reduced-cost vector $\mathbf{r} \geq \mathbf{0}$.

What to do if it is not optimal?

An effort can be made to find a better neighboring basic feasible solution. That is, a new BFS that differs from the current one by exactly one new basic variable, as long as the reduced cost coefficient of the entering variable is negative.

Changing Basis for the LP Example

With the initial basic variable set $B = \{3, 4, 5\}$:

B	-1	-2	0	0	0	0
3	1	0	1	0	0	1
4	0	1	0	1	0	1
5	1	1	0	0	1	$\frac{3}{2}$

Select the entering basic variable x_1 with $r_1 = -1 < 0$ and consider

$$\begin{pmatrix} x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \frac{3}{2} \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} x_1.$$

How much can we increase x_1 while the current basic variables remain feasible (non-negative)? This connects to the definition of minimum ratio test (MRT).

Minimum Ratio Test (MRT)

- Select the entering variable x_e with its reduced cost $r_e < 0$;
- If column $\bar{A}_{.e} \leq 0$, then Unbounded
- Minimum Ratio Test (MRT):

$$\theta := \min\left\{ \frac{\bar{b}_i}{\bar{A}_{ie}} : \bar{A}_{ie} > 0 \right\}.$$

How Much Can the New Basic Variable be Increased?

What is θ? It is the largest amount that x_e can be increased before one (or more) of the current basic variables x_i decreases to zero.

For the moment, assume that the minimum ratio in the MRT is attained by exactly one basic variable index, o. Then x_o is called the out-going basic variable:

$$\begin{aligned} x_o &= \bar{b}_o - \bar{a}_{oe}\theta = 0 \\ x_i &= \bar{b}_i - \bar{a}_{ie}\theta > 0 \quad \forall i \neq o. \end{aligned}$$

That is,

$$x_o \longleftarrow x_e$$

in the basic variable set.

Tie Breaking

If the MRT does not result in a single index, but rather in a set of two or more indices for which the minimum ratio is attained, we select one of these indices as out-going arbitrarily.

Again, when x_e reaches θ we have generated a corner point of the feasible region that is adjacent to the one associated with the index set B. The new basic variable set associated with this "new" corner point is

$$o \longleftarrow e.$$

But the new BFS is degenerate, one of its basic variable has value 0. Pretending that $\epsilon > 0$ but arbitrarily small and continue the transformation process to the new canonical form.

The Simplex Algorithm

Step 0. Initialize with a minimization problem in feasible canonical form with respect to a basic index set B. Let N denote the complementary index set.

Step 1. Test for termination: first find

$$r_e = \min_{j \in N}\{r_j\}.$$

If $r_e \geq 0$, stop. The solution is optimal. Otherwise determine whether the column of $\bar{A}_{.e}$ contains a positive entry. If not, the objective function is unbounded below. Terminate. Let x_e be the entering basic variable.

Step 2. Determine the outgoing: execute the MRT to determine the outgoing variable x_o.

Step 3. Update basis: update B and A_B and transform the problem to canonical form and return to Step 1.

Transform to a New Canonical Form: Pivoting

The passage from one canonical form to the next can be carried out by an algebraic process called pivoting. In a nutshell, a pivot on the nonzero (in our case, positive) pivot element $\bar{a}_{oe}$ amounts to

(a) expressing x_e in terms of all the non-basic variables;

(b) replacing x_e in all other formula by using the expression.

In terms of $\bar{A}$, this has the effect of making the column of x_e look like the column of an identity matrix with 1 in the o^{th} row and zero elsewhere.

Pivoting: Gauss-Jordan Elimination Process

Here is what happens to the current table:

1. All the entries in row o are divided by the pivot element $\bar{a}_{oe}$ (This produces a 1 (one) in the column of x_s).

2. For all $i \neq o$, all other entries are modified according to the rule

$$\bar{a}_{ij} \longleftarrow \bar{a}_{ij} - \frac{\bar{a}_{oj}}{\bar{a}_{oe}} \bar{a}_{ie}.$$

   (When $j = e$, the new entry is just 0 (zero).)

The right-hand side and the objective function row are modified in the same way.

The Initial Canonical Form of the Example

Choose $e = 2$ and MRT would decide $o = 4$ with $\theta = 1$:

B	-1	$-\mathbf{2}$	0	0	0		MRT
3	1	0	1	0	0	1	∞
4	0	1	0	1	0	1	$\mathbf{1}$
5	1	1	0	0	1	$\frac{3}{2}$	$\frac{3}{2}$

If necessary, deviding the pivoting row by the pivot value to make the pivot element equal to 1.

Transform to a New Canonical Form by Pivoting

“Objective row := Objective Row + 2$*$Pivot Row”:

B	-1	0	0	2	0	2
3	1	0	1	0	0	1
2	0	1	0	1	0	1
5	1	1	0	0	1	$\frac{3}{2}$

“Bottom row := Bottom Row - Pivot Row”:

B	-1	0	0	2	0	2
3	1	0	1	0	0	1
2	0	1	0	1	0	1
5	1	0	0	-1	1	$\frac{1}{2}$

This is a new canonical form!

Let's Keep Going

Choose $e = 1$ and MRT would decide $o = 5$ with $\theta = 1/2$:

B	$-\mathbf{1}$	0	0	2	0	2	MRT
3	1	0	1	0	0	1	1
2	0	1	0	1	0	1	∞
5	1	0	0	-1	1	$\frac{1}{2}$	$\mathbf{\frac{1}{2}}$

Transform to a New Canonical Form by Pivoting

"Objective Row := Objective Row + Pivot Row":

B	0	0	0	1	1	$2\frac{1}{2}$
3	1	0	1	0	0	1
2	0	1	0	1	0	1
1	1	0	0	-1	1	$\frac{1}{2}$

"Second Row := Second Row - Pivot Row":

B	0	0	0	1	1	$2\frac{1}{2}$
3	0	0	1	1	-1	$\frac{1}{2}$
2	0	1	0	1	0	1
1	1	0	0	-1	1	$\frac{1}{2}$

Stop! It's optimal:)

Detailed Canonical Tableau for Production

If the original LP is the production problem:

$$\begin{array}{ll} \text{maximize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & A\mathbf{x} \le \mathbf{b}(> \mathbf{0}), \\ & \mathbf{x} \ge \mathbf{0}. \end{array}$$

The initial canonical tableau for minimization would be

B	$-\mathbf{c}^T$	$\mathbf{0}$	0
Basis Indices	A	I	$\mathbf{b}$

The intermediate canonical tableau would be

B	$\mathbf{r}^T$	$-\mathbf{y}^T$	$\mathbf{c}_B^T\bar{\mathbf{b}}$
Basis Indices	$\bar{A}$	A_B^{-1}	$\bar{\mathbf{b}}$