Lecture 7
Theory of LP

Solving systems with more variables than equations
Matrix representation of a standart LP
Getting a basic solution
Theory of optimality and feasibility conditions
Some basic theorems and their proofs

Solving Systems with More Variables than Equations

Suppose now that $\mathbf{A} \in \mathbb{R}^{m \times n}$ where $m \leq n$. Let $\mathbf{b} \in \mathbb{R}^m$. Then the equation:

$$(3.46) \quad \mathbf{Ax} = \mathbf{b}$$

has more variables than equations and is underdetermined and if $\mathbf{A}$ has full row rank then the system will have an infinite number of solutions. We can formulate an expression to describe this infinite set of solutions.

Sine $\mathbf{A}$ has full row rank, we may choose any m linearly independent columns of $\mathbf{A}$ corresponding to a subset of the variables, say $x_{i_1}, \dots, x_{i_m}$. We can use these to form the matrix

$$(3.47) \quad \mathbf{B} = [\mathbf{A}_{\cdot i_1} \cdots \mathbf{A}_{\cdot i_m}]$$

from the columns $\mathbf{A}_{\cdot i_1}, \dots, \mathbf{A}_{\cdot i_m}$ of $\mathbf{A}$, so that B is invertible. It should be clear at this point that B will be invertible precisely because we've chosen m linearly independent column vectors. We can then use elementary column operations to write the matrix $\mathbf{A}$ as:

$$(3.48) \quad \mathbf{A} = [\mathbf{B}|\mathbf{N}]$$

The matrix $\mathbf{N}$ is composed of the $n - m$ other columns of $\mathbf{A}$ not in $\mathbf{B}$. We can similarly sub-divide the column vector $\mathbf{x}$ and write:

$$(3.49) \quad [\mathbf{B}|\mathbf{N}] \begin{bmatrix} \mathbf{x_B} \\ \mathbf{x_N} \end{bmatrix} = \mathbf{b}$$

where the vector $\mathbf{x_B}$ are the variables corresponding to the columns in $\mathbf{B}$ and the vector $\mathbf{x_N}$ are the variables corresponding to the columns of the matrix $\mathbf{N}$.

DEFINITION 3.46 (Basic Variables). For historical reasons, the variables in the vector $\mathbf{x_B}$ are called the *basic variables* and the variables in the vector $\mathbf{x_N}$ are called the *non-basic* variables.

We can use matrix multiplication to expand the left hand side of this expression as:

$$\mathbf{B}\mathbf{x_B} + \mathbf{N}\mathbf{x_N} = \mathbf{b} \tag{3.50}$$

The fact that $\mathbf{B}$ is composed of all linearly independent columns implies that applying Gauss-Jordan elimination to it will yield an $m \times m$ identity and thus that $\mathbf{B}$ is invertible. We can solve for basic variables $\mathbf{x_B}$ in terms of the non-basic variables:

$$\mathbf{x_B} = \mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{N}\mathbf{x_N} \tag{3.51}$$

We can find an arbitrary solution to the system of linear equations by choosing values for the variables the non-basic variables and solving for the basic variable values using Equation 3.51.

DEFINITION 3.47. (Basic Solution) When we assign $\mathbf{x_N} = 0$, the resulting solution for $\mathbf{x}$ is called a *basic solution* and

$$\mathbf{x_B} = \mathbf{B}^{-1}\mathbf{b} \tag{3.52}$$

Example 3.48. Consider the problem:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 8 \end{bmatrix} \tag{3.53}$$

Then we can let $x_3 = 0$ and:

$$\mathbf{B} = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix} \tag{3.54}$$

We then solve[1]:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{B}^{-1} \begin{bmatrix} 7 \\ 8 \end{bmatrix} = \begin{bmatrix} \frac{-19}{3} \\ \frac{20}{3} \end{bmatrix} \tag{3.55}$$

Other basic solutions could be formed by creating $\mathbf{B}$ out of columns 1 and 3 or columns 2 and 3.

Exercise 38. Find the two other basic solutions in Example 3.48 corresponding to

$$\mathbf{B} = \begin{bmatrix} 2 & 3 \\ 5 & 6 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 1 & 3 \\ 4 & 6 \end{bmatrix}$$

In each case, determine what the matrix $\mathbf{N}$ is. [Hint: Find the solutions any way you like. Make sure you record exactly which x_i ($i \in \{1, 2, 3\}$) is equal to zero in each case.]

Matrix representation of a standart LP

- Consider the following standart LP problem:

max $z = c^T x$

s.t. $Ax = b$,

$x \geq 0$.

Getting a Basic Feasible Solution

To get a basic solution rewrite the linear program in terms of the basis where B is the basis of A:

$A=[B\ N]$, $c = [c_B\ c_N]^T$, $x = [x_B\ x_N]^T$, $x_B \in R^m$, $x_N \in R^{n-m}$

max $z = [c_B\ c_N]\ [x_B\ x_N]^T$

s.t. $[B\ N]\ [x_B\ x_N]^T = b$

$x_B , x_N \geq 0$.

$Ax = Bx_B + Nx_N = b$

, $x_N = 0$, $x_B = B^{-1}b$

- Algebraically manipulating the program we can solve for x_B :

Algebraically manipulating the program we can solve for x_B:

$$\begin{aligned} Bx_B + Nx_N &= b \\ x_B &= B^{-1}(b - N_{x_N}) \end{aligned}$$

Then rearranging $c'x$:

$$\begin{aligned} c'x &= c'_B x_B + c'_N x_N \\ &= c'_B B^{-1}(b - Nx_N) + c'_N X_N \\ &= c'_B B^{-1} b - c'_B B^{-1} N x_N + c'_N x_N \\ &= z_0 - \sum_j (z_j - c_j) x_j \end{aligned}$$

where:

$$\begin{aligned} z_0 &= c'_B B^{-1} b \\ z_j &= c'_B B^{-1} A_j \ , \ j \in N \end{aligned}$$

We can now see that:

$$max\ c'x = z_0 - \sum_j (z_j - c_j) x_j$$

So to test optimality we can check:

$$z_j - c_j \geq 0 \ \forall j \in N$$

3. The Simplex Algorithm

Suppose we have a basic feasible solution $\mathbf{x} = (\mathbf{x_B}, \mathbf{x_N})$. We can divide the cost vector $\mathbf{c}$ into its basic and non-basic parts, so we have $\mathbf{c} = [\mathbf{c_B}|\mathbf{c_N}]^T$. Then the objective function becomes:

$$(5.10) \quad \mathbf{c}^T\mathbf{x} = \mathbf{c_B}^T\mathbf{x_B} + \mathbf{c_N}^T\mathbf{x_N}$$

We can substitute Equation 5.8 into Equation 5.10 to obtain:

$$(5.11) \quad \mathbf{c}^T\mathbf{x} = \mathbf{c_B}^T\left(\mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{N}\mathbf{x_N}\right) + \mathbf{c_N}\mathbf{x_N} = \mathbf{c_B}^T\mathbf{B}^{-1}\mathbf{b} + \left(\mathbf{c_N}^T - \mathbf{c_B}^T\mathbf{B}^{-1}\mathbf{N}\right)\mathbf{x}_N$$

Let $\mathcal{J}$ be the set of indices of non-basic variables. Then we can write Equation 5.11 as:

$$(5.12) \quad z(x_1, \dots, x_n) = \mathbf{c_B}^T\mathbf{B}^{-1}\mathbf{b} + \sum_{j \in \mathcal{J}} \left(\mathbf{c}_j - \mathbf{c_B}^T\mathbf{B}^{-1}\mathbf{A}_{\cdot j}\right) x_j$$

Consider now the fact $x_j = 0$ for all $j \in \mathcal{J}$. Further, we can see that:

$$(5.13) \quad \frac{\partial z}{\partial x_j} = \mathbf{c}_j - \mathbf{c_B}^T\mathbf{B}^{-1}\mathbf{A}_{\cdot j}$$

This means that if $\mathbf{c}_j - \mathbf{c}_{\mathbf{B}}^T\mathbf{B}^{-1}\mathbf{A}_{\cdot j} > 0$ and we *increase* x_j from zero to some new value, then we will *increase* the value of the objective function. For historic reasons, we actually consider the value $\mathbf{c}_{\mathbf{B}}^T\mathbf{B}^{-1}\mathbf{A}_{\cdot j} - \mathbf{c}_j$, called the *reduced cost* and denote it as:

$$(5.14) \quad -\frac{\partial z}{\partial x_j} = z_j - c_j = \mathbf{c}_{\mathbf{B}}^T\mathbf{B}^{-1}\mathbf{A}_{\cdot j} - \mathbf{c}_j$$

In a maximization problem, we chose non-basic variables x_j with negative reduced cost to become basic because, in this case, $\partial z/\partial x_j$ is *positive*.

Assume we chose x_j, a non-basic variable to become non-zero (because $z_j - c_j < 0$). We wish to know which of the basic variables will become zero as we *increase* x_j away from zero. We must also be very careful that *none* of the variables become negative as we do this.

By Equation 5.8 we know that the only current basic variables will be affected by increasing x_j. Let us focus explicitly on Equation 5.8 where we include only variable x_j (since all other non-basic variables are kept zero). Then we have:

$$(5.15) \quad \mathbf{x}_{\mathbf{B}} = \mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{A}_{\cdot j}x_j$$

Let $\overline{\mathbf{b}} = \mathbf{B}^{-1}\mathbf{b}$ be an $m \times 1$ column vector and let that $\overline{\mathbf{a}_j} = \mathbf{B}^{-1}\mathbf{A}_{\cdot j}$ be another $m \times 1$ column. Then we can write:

$$(5.16) \quad \mathbf{x}_{\mathbf{B}} = \overline{\mathbf{b}} - \overline{\mathbf{a}}_j x_j$$

Let $\overline{\mathbf{b}} = [\overline{b}_1, \dots \overline{b}_m]^T$ and $\overline{\mathbf{a}}_j = [\overline{a}_{j_1}, \dots, \overline{a}_{j_m}]$, then we have:

$$(5.17) \quad \begin{bmatrix} x_{B_1} \\ x_{B_2} \\ \vdots \\ x_{B_m} \end{bmatrix} = \begin{bmatrix} \overline{b}_1 \\ \overline{b}_2 \\ \vdots \\ \overline{b}_m \end{bmatrix} - \begin{bmatrix} \overline{a}_{j_1} \\ \overline{b}_{j_2} \\ \vdots \\ \overline{b}_{j_m} \end{bmatrix} x_j = \begin{bmatrix} \overline{b}_1 - \overline{a}_{j_1} x_j \\ \overline{b}_2 - \overline{a}_{j_2} x_j \\ \vdots \\ \overline{b}_m - \overline{a}_{j_m} x_j \end{bmatrix}$$

We know (a priori) that $\overline{b}_i \geq 0$ for $i = 1, \dots, m$. If $\overline{a}_{j_i} \leq 0$, then as we increase x_j, $\overline{b}_i - \overline{a}_{j_i} \geq 0$ no matter how large we make x_j. On the other hand, if $\overline{a}_{j_i} > 0$, then as we increase x_j we know that $\overline{b}_i - \overline{a}_{j_i} x_j$ will get smaller and eventually hit zero. In order to ensure that *all* variables remain non-negative, we cannot increase x_j beyond a certain point.

For each i ($i = 1, \dots, m$) such that $\overline{a}_{j_i} > 0$, the value of x_j that will make x_{B_i} goto 0 can be found by observing that:

$$(5.18) \quad x_{B_i} = \overline{b}_i - \overline{a}_{j_i} x_j$$

and if $x_{B_i} = 0$, then we can solve:

$$(5.19) \quad 0 = \overline{b}_i - \overline{a}_{j_i} x_j \implies x_j = \frac{\overline{b}_i}{\overline{a}_{j_i}}$$

Thus, the *largest possible value* we can assign x_j and ensure that all variables remain positive is:

$$(5.20) \quad \min \left\{ \frac{\overline{b}_i}{\overline{a}_{j_i}} : i = 1, \dots, m \text{ and } a_{j_i} > 0 \right\}$$

Expression 5.20 is called the *minimum ratio test*. We are interested in which index i is the minimum ratio.

Suppose that in executing the minimum ratio test, we find that $x_j = \overline{b}_k/\overline{a}_{j_k}$. The variable x_j (which was non-basic) becomes basic and the variable $x_{\mathbf{B}_k}$ becomes non-basic. All other basic variables remain basic (and positive). In executing this procedure (of exchanging one basic variable and one non-basic variable) we have moved from one extreme point of X to another.

THEOREM 5.6. *If $z_j - c_j \geq 0$ for all $j \in \mathcal{J}$, then the current basic feasible solution is optimal.*

EXAMPLE 5.9. Consider the Toy Maker Problem (from Example 2.3). The linear programming problem given in Equation 2.8 is:

$$\begin{cases} \max & z(x_1, x_2) = 7x_1 + 6x_2 \\ s.t. & 3x_1 + x_2 \leq 120 \\ & x_1 + 2x_2 \leq 160 \\ & x_1 \leq 35 \\ & x_1 \geq 0 \\ & x_2 \geq 0 \end{cases}$$

We can convert this problem to standard form by introducing the slack variables s_1, s_2 and s_3:

$$
\begin{cases}
\max \quad z(x_1, x_2) = 7x_1 + 6x_2 \\
s.t. \;\; 3x_1 + x_2 + s_1 = 120 \\
x_1 + 2x_2 + s_2 = 160 \\
x_1 + s_3 = 35 \\
x_1, x_2, s_1, s_2, s_3 \geq 0
\end{cases}
$$

which yields the matrices

$$
\mathbf{c} = \begin{bmatrix} 7 \\ 6 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ s_1 \\ s_2 \\ s_3 \end{bmatrix} \quad \mathbf{A} = \begin{bmatrix} 3 & 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 120 \\ 160 \\ 35 \end{bmatrix}
$$

We can begin with the matrices:

$$
\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \mathbf{N} = \begin{bmatrix} 3 & 1 \\ 1 & 2 \\ 1 & 0 \end{bmatrix}
$$

In this case we have:

$$
\mathbf{x_B} = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} \quad \mathbf{x_N} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \mathbf{c_B} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{c_N} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}
$$

and $\quad \mathbf{B}^{-1}\mathbf{b} = \begin{bmatrix} 120 \\ 160 \\ 35 \end{bmatrix} \quad \mathbf{B}^{-1}\mathbf{N} = \begin{bmatrix} 3 & 1 \\ 1 & 2 \\ 1 & 0 \end{bmatrix}$

Therefore:

$$\mathbf{c}_{\mathbf{B}}^T\mathbf{B}^{-1}\mathbf{b} = 0 \quad \mathbf{c}_{\mathbf{B}}^T\mathbf{B}^{-1}\mathbf{N} = \begin{bmatrix} 0 & 0 \end{bmatrix} \quad \mathbf{c}_{\mathbf{B}}^T\mathbf{B}^{-1}\mathbf{N} - \mathbf{c}_{\mathbf{N}} = \begin{bmatrix} -7 & -6 \end{bmatrix}$$

Using this information, we can compute:

$$\mathbf{c}_{\mathbf{B}}^T\mathbf{B}^{-1}\mathbf{A}_{\cdot 1} - \mathbf{c}_1 = -7$$
$$\mathbf{c}_{\mathbf{B}}^T\mathbf{B}^{-1}\mathbf{A}_{\cdot 2} - \mathbf{c}_2 = -6$$

and therefore:

$$\frac{\partial z}{\partial x_1} = 7 \text{ and } \frac{\partial z}{\partial x_2} = 6$$

Based on this information, we could chose either x_1 or x_2 to enter the basis and the value of the objective function would increase. If we chose x_1 to enter the basis, then we must determine which variable will leave the basis. To do this, we must investigate the elements of $\mathbf{B}^{-1}\mathbf{A}_{\cdot 1}$ and the current basic feasible solution $\mathbf{B}^{-1}\mathbf{b}$. Since each element of $\mathbf{B}^{-1}\mathbf{A}_{\cdot 1}$ is positive, we must perform the minimum ratio test on each element of $\mathbf{B}^{-1}\mathbf{A}_{\cdot 1}$. We know that $\mathbf{B}^{-1}\mathbf{A}_{\cdot 1}$ is just the first column of $\mathbf{B}^{-1}\mathbf{N}$ which is:

$$\mathbf{B}^{-1}\mathbf{A}_{\cdot 1} = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$$

Performing the minimum ratio test, we see have:

$$\min \left\{ \frac{120}{3}, \frac{160}{1}, \frac{35}{1} \right\}$$

In this case, we see that index 3 (35/1) is the minimum ratio. Therefore, variable x_1 will enter the basis and variable s_3 will leave the basis. The new basic and non-basic variables will be:

$$\mathbf{x_B} = \begin{bmatrix} s_1 \\ s_2 \\ x_1 \end{bmatrix} \quad \mathbf{x_N} = \begin{bmatrix} s_3 \\ x_2 \end{bmatrix} \quad \mathbf{c_B} = \begin{bmatrix} 0 \\ 0 \\ 7 \end{bmatrix} \quad \mathbf{c_N} = \begin{bmatrix} 0 \\ 6 \end{bmatrix}$$

and the matrices become:

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad \mathbf{N} = \begin{bmatrix} 0 & 1 \\ 0 & 2 \\ 1 & 0 \end{bmatrix}$$

Note we have simply swapped the column corresponding to x_1 with the column corresponding to s_3 in the basis matrix $\mathbf{B}$ and the non-basic matrix $\mathbf{N}$. We will do this repeatedly in the example and we recommend the reader keep track of which variables are being exchanged and why certain columns in $\mathbf{B}$ are being swapped with those in $\mathbf{N}$.

Using the new $\mathbf{B}$ and $\mathbf{N}$ matrices, the derived matrices are then:

$$\mathbf{B}^{-1}\mathbf{b} = \begin{bmatrix} 15 \\ 125 \\ 35 \end{bmatrix} \quad \mathbf{B}^{-1}\mathbf{N} = \begin{bmatrix} -3 & 1 \\ -1 & 2 \\ 1 & 0 \end{bmatrix}$$

The cost information becomes:

$$\mathbf{c}_{\mathbf{B}}^T\mathbf{B}^{-1}\mathbf{b} = 245 \quad \mathbf{c}_{\mathbf{B}}^T\mathbf{B}^{-1}\mathbf{N} = \begin{bmatrix} 7 & 0 \end{bmatrix} \quad \mathbf{c}_{\mathbf{B}}^T\mathbf{B}^{-1}\mathbf{N} - \mathbf{c}_{\mathbf{N}} = \begin{bmatrix} 7 & -6 \end{bmatrix}$$

Using the new $\mathbf{B}$ and $\mathbf{N}$ matrices, the derived matrices are then:

$$\mathbf{B}^{-1}\mathbf{b} = \begin{bmatrix} 15 \\ 125 \\ 35 \end{bmatrix} \quad \mathbf{B}^{-1}\mathbf{N} = \begin{bmatrix} -3 & 1 \\ -1 & 2 \\ 1 & 0 \end{bmatrix}$$

The cost information becomes:

$$\mathbf{c}_{\mathbf{B}}^T\mathbf{B}^{-1}\mathbf{b} = 245 \quad \mathbf{c}_{\mathbf{B}}^T\mathbf{B}^{-1}\mathbf{N} = \begin{bmatrix} 7 & 0 \end{bmatrix} \quad \mathbf{c}_{\mathbf{B}}^T\mathbf{B}^{-1}\mathbf{N} - \mathbf{c}_{\mathbf{N}} = \begin{bmatrix} 7 & -6 \end{bmatrix}$$

using this information, we can compute:

$$\mathbf{c}_{\mathbf{B}}^T\mathbf{B}^{-1}\mathbf{A}_{\cdot 5} - \mathbf{c}_5 = 7$$
$$\mathbf{c}_{\mathbf{B}}^T\mathbf{B}^{-1}\mathbf{A}_{\cdot 2} - \mathbf{c}_2 = -6$$

Based on this information, we can only choose x_2 to enter the basis to ensure that the value of the objective function increases. We can perform the minimum ration test to figure out which basic variable will leave the basis. We know that $\mathbf{B}^{-1}\mathbf{A}_{\cdot 2}$ is just the second column of $\mathbf{B}^{-1}\mathbf{N}$ which is:

$$\mathbf{B}^{-1}\mathbf{A}_{\cdot 2} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

Performing the minimum ratio test, we see have:

$$\min \left\{ \frac{15}{1}, \frac{125}{2} \right\}$$

In this case, we see that index 1 (15/1) is the minimum ratio. Therefore, variable x_2 will enter the basis and variable s_1 will leave the basis. The new basic and non-basic variables will be: The new basic and non-basic variables will be:

$$\mathbf{x_B} = \begin{bmatrix} x_2 \\ s_2 \\ x_1 \end{bmatrix} \quad \mathbf{x_N} = \begin{bmatrix} s_3 \\ s_1 \end{bmatrix} \quad \mathbf{c_B} = \begin{bmatrix} 6 \\ 0 \\ 7 \end{bmatrix} \quad \mathbf{c_N} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

and the matrices become:

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad \mathbf{N} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}$$

The derived matrices are then:

$$\mathbf{B}^{-1}\mathbf{b} = \begin{bmatrix} 15 \\ 95 \\ 35 \end{bmatrix} \quad \mathbf{B}^{-1}\mathbf{N} = \begin{bmatrix} -3 & 1 \\ 5 & -2 \\ 1 & 0 \end{bmatrix}$$

The cost information becomes:

$$\mathbf{c}_{\mathbf{B}}^T\mathbf{B}^{-1}\mathbf{b} = 335 \quad \mathbf{c}_{\mathbf{B}}^T\mathbf{B}^{-1}\mathbf{N} = \begin{bmatrix} -11 & 6 \end{bmatrix} \quad \mathbf{c}_{\mathbf{B}}^T\mathbf{B}^{-1}\mathbf{N} - \mathbf{c}_{\mathbf{N}} = \begin{bmatrix} -11 & 6 \end{bmatrix}$$

Based on this information, we can only choose s_3 to (re-enter) the basis to ensure that the value of the objective function increases. We can perform the minimum ration test to figure out which basic variable will leave the basis. We know that $\mathbf{B}^{-1}\mathbf{A}_{\cdot 5}$ is just the fifth column of $\mathbf{B}^{-1}\mathbf{N}$ which is:

$$\mathbf{B}^{-1}\mathbf{A}_{\cdot 5} = \begin{bmatrix} -3 \\ 5 \\ 1 \end{bmatrix}$$

Performing the minimum ratio test, we see have:

$$\min \left\{ \frac{95}{5}, \frac{35}{1} \right\}$$

In this case, we see that index 2 (95/5) is the minimum ratio. Therefore, variable s_3 will enter the basis and variable s_2 will leave the basis. The new basic and non-basic variables will be:

$$\mathbf{x}_{\mathbf{B}} = \begin{bmatrix} x_2 \\ s_3 \\ x_1 \end{bmatrix} \quad \mathbf{x}_{\mathbf{N}} = \begin{bmatrix} s_2 \\ s_1 \end{bmatrix} \quad \mathbf{c}_{\mathbf{B}} = \begin{bmatrix} 6 \\ 0 \\ 7 \end{bmatrix} \quad \mathbf{c}_{\mathbf{N}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

and the matrices become:

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \quad \mathbf{N} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}$$

The derived matrices are then:

$$\mathbf{B}^{-1}\mathbf{b} = \begin{bmatrix} 72 \\ 19 \\ 16 \end{bmatrix} \quad \mathbf{B}^{-1}\mathbf{N} = \begin{bmatrix} 6/10 & -1/5 \\ 1/5 & -2/5 \\ -1/5 & 2/5 \end{bmatrix}$$

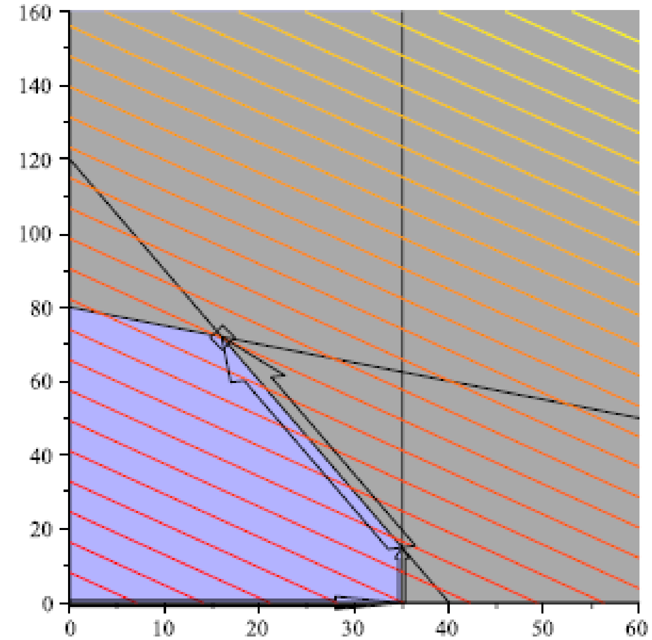

Figure 5.1.

The cost information becomes:

$$\mathbf{c_B}^T\mathbf{B}^{-1}\mathbf{b} = 544 \quad \mathbf{c_B}^T\mathbf{B}^{-1}\mathbf{N} = \begin{bmatrix} 11/5 & 8/5 \end{bmatrix} \quad \mathbf{c_B}^T\mathbf{B}^{-1}\mathbf{N} - \mathbf{c_N} = \begin{bmatrix} 11/5 & 8/5 \end{bmatrix}$$

Since the reduced costs are now positive, we can conclude that we've obtained an optimal solution because no improvement is possible. The final solution then is:

$$\mathbf{x_B}^* = \begin{bmatrix} x_2 \\ s_3 \\ x_1 \end{bmatrix} = \mathbf{B}^{-1}\mathbf{b} = \begin{bmatrix} 72 \\ 19 \\ 16 \end{bmatrix}$$

Simply, we have $x_1 = 16$ and $x_2 = 72$ as we obtained in Example 2.3. The path of extreme points we actually took in traversing the boundary of the polyhedral feasible region is shown in Figure 5.1.

4. Simplex Method–Tableau Form

No one executes the simplex algorithm in algebraic form. Instead, several representations (tableau representations) have been developed to lesson the amount of writing that needs to be done and to collect all pertinent information into a single table.

To see how a *Simplex Tableau* is derived, consider Problem P in standard form:

$$P \begin{cases} \max & \mathbf{c}^T\mathbf{x} \\ s.t. & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{cases}$$

We can re-write P in an unusual way by introducing a new variable z and separating $\mathbf{A}$ into its basic and non-basic parts to obtain:

$$\begin{aligned} &\max \quad z \\ (5.22) \qquad &s.t. \quad z - \mathbf{c}_\mathbf{B}^T\mathbf{x}_\mathbf{B} - \mathbf{c}_\mathbf{N}^T\mathbf{x}_\mathbf{N} = 0 \\ &\qquad\quad \mathbf{B}\mathbf{x}_\mathbf{B} + \mathbf{N}\mathbf{x}_\mathbf{N} = \mathbf{b} \\ &\qquad\quad \mathbf{x}_\mathbf{B}, \mathbf{x}_\mathbf{N} \geq 0 \end{aligned}$$

From the second equation, it's clear

$$(5.23) \quad \mathbf{x}_\mathbf{B} + \mathbf{B}^{-1}\mathbf{N}\mathbf{x}_\mathbf{N} = \mathbf{B}^{-1}\mathbf{b}$$

We can multiply this equation by $\mathbf{c}_\mathbf{B}^\mathbf{T}$ to obtain:

$$(5.24) \quad \mathbf{c}_\mathbf{B}^T\mathbf{x}_\mathbf{B} + \mathbf{c}_\mathbf{B}^T\mathbf{B}^{-1}\mathbf{N}\mathbf{x}_\mathbf{N} = \mathbf{c}_\mathbf{B}^T\mathbf{B}^{-1}\mathbf{b}$$

If we add this equation to the equation $z - \mathbf{c}_\mathbf{B}^T\mathbf{x}_\mathbf{B} - \mathbf{c}_\mathbf{N}^T\mathbf{x}_\mathbf{N} = 0$ we obtain:

$$(5.25) \quad z + 0^T\mathbf{x}_\mathbf{B} + \mathbf{c}_\mathbf{B}^T\mathbf{B}^{-1}\mathbf{N}\mathbf{x}_\mathbf{N} - \mathbf{c}_\mathbf{N}^T\mathbf{x}_\mathbf{N} = \mathbf{c}_\mathbf{B}^T\mathbf{B}^{-1}\mathbf{b}$$

Here $\mathbf{0}$ is the vector of zeros of appropriate size. This equation can be written as:

$$(5.26) \quad z + \mathbf{0}^T \mathbf{x_B} + \left(\mathbf{c}_\mathbf{B}^T \mathbf{B}^{-1}\mathbf{N} - \mathbf{c}_\mathbf{N}^T\right) \mathbf{x_N} = \mathbf{c}_\mathbf{B}^T \mathbf{B}^{-1}\mathbf{b}$$

We can now represent this set of equations as a large matrix (or tableau):

	z	$\mathbf{x_B}$	$\mathbf{x_N}$	RHS	
z	1	0	$\mathbf{c}_\mathbf{B}^T \mathbf{B}^{-1}\mathbf{N} - \mathbf{c}_\mathbf{N}^T$	$\mathbf{c}_\mathbf{B}^T \mathbf{B}^{-1}\mathbf{b}$	Row 0
$\mathbf{x_B}$	0	1	$\mathbf{B}^{-1}\mathbf{N}$	$\mathbf{B}^{-1}\mathbf{b}$	Rows 1 through m

The augmented matrix shown within the table:

$$(5.27) \quad \left[\begin{array}{ccc|c} 1 & 0 & \mathbf{c}_\mathbf{B}^T \mathbf{B}^{-1}\mathbf{N} - \mathbf{c}_\mathbf{N}^T & \mathbf{c}_\mathbf{B}^T \mathbf{B}^{-1}\mathbf{b} \\ 0 & 1 & \mathbf{B}^{-1}\mathbf{N} & \mathbf{B}^{-1}\mathbf{b} \end{array}\right]$$

is simply the matrix representation of the simultaneous equations described by Equations 5.23 and 5.26. We can see that the first row consists of a row of the first row of the $(m+1) \times (m+1)$ identity matrix, the reduced costs of the non-basic variables and the current objective function values. The remainder of the rows consist of the rest of the $(m+1) \times (m+1)$ identity matrix, the matrix $\mathbf{B}^{-1}\mathbf{N}$ and $\mathbf{B}^{-1}\mathbf{b}$ the current non-zero part of the basic feasible solution.

This matrix representation (or tableau representation) contains all of the information we need to execute the simplex algorithm. An entering variable is chosen from among the columns containing the reduced costs and matrix $\mathbf{B}^{-1}\mathbf{N}$. Naturally, a column with a negative reduced cost is chosen. We then chose a leaving variable by performing the minimum ratio test on the chosen column and the right-hand-side (RHS) column. We pivot on the element at the entering column and leaving row and this transforms the tableau into a new tableau that represents the new basic feasible solution.

Theorem: The set of all feasible solutions to an LP problem is a convex set.

Proof:

Theorem: An LP problem assumes its optimum at an extreme point. If it assumes its optimum at more than one extreme point, then it takes on the same value for every convex combination of these particular points.

Proof:

Appendix

1. Convex Sets

DEFINITION 4.1 (Convex Set). Let $X \subseteq \mathbb{R}^n$. Then the set X is convex if and only if for all pairs $\mathbf{x}_1, \mathbf{x}_2 \in X$ we have $\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in X$ for all $\lambda \in [0,1]$.

The definition of convexity seems complex, but it is easy to understand. First recall that if $\lambda \in [0,1]$, then the point $\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$ is on the line segment connecting $\mathbf{x}_1$ and $\mathbf{x}_2$ in $\mathbb{R}^n$. For example, when $\lambda = 1/2$, then the point $\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$ is the midpoint between $\mathbf{x}_1$ and $\mathbf{x}_2$. In fact, for every point $\mathbf{x}$ on the line connecting $\mathbf{x}_1$ and $\mathbf{x}_2$ we can find a value $\lambda \in [0,1]$ so that $\mathbf{x} = \lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$. Then we can see that, convexity asserts that if $\mathbf{x}_1, \mathbf{x}_2 \in X$, then every point on the line connecting $\mathbf{x}_1$ and $\mathbf{x}_2$ is also in the set X.

6. Extreme Points

DEFINITION 4.27 (Extreme Point of a Convex Set). Let C be a convex set. A point $\mathbf{x}_0 \in C$ is a *extreme point* of C if there are *no points* $\mathbf{x}_1$ and $\mathbf{x}_2$ ($\mathbf{x}_1 \neq \mathbf{x}_0$ or $\mathbf{x}_2 \neq \mathbf{x}_0$) so that $\mathbf{x} = \lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$ for some $\lambda \in (0,1)$.[2]

An extreme point is simply a point in a convex set C that cannot be expressed as a strict convex combination of any other pair of points in C. We will see that extreme points must be located in specific locations in convex sets.

Definition 4.3 (Convex Combination). Let $\mathbf{x}_1, \dots, \mathbf{x}_m \in \mathbb{R}^n$. If $\lambda_1, \dots, \lambda_m \in [0, 1]$ and

$$\sum_{i=1}^{m} \lambda_i = 1 \quad \text{then}$$

$$\mathbf{x} = \sum_{i=1}^{m} \lambda_i \mathbf{x}_i \tag{4.2}$$

is called a *convex combination* of $\mathbf{x}_1, \dots, \mathbf{x}_m$. If $\lambda_i < 1$ for all $i = 1, \dots, m$, then Equation 4.2 is called a *strict convex combination*.

7. Linear Combinations, Span, Linear Independence

Definition 3.27. Let $\mathbf{x}_1, \dots, \mathbf{x}_m$ be vectors in $\in \mathbb{R}^n$ and let $\alpha_1, \dots, \alpha_m \in \mathbb{R}$ be scalars. Then

$$\alpha_1 \mathbf{x}_1 + \cdots + \alpha_m \mathbf{x}_m \tag{3.34}$$

is a *linear combination* of the vectors $\mathbf{x}_1, \dots, \mathbf{x}_m$.

Clearly, any linear combination of vectors in $\mathbb{R}^n$ is also a vector in $\mathbb{R}^n$.

Definition 3.28 (Span). Let $\mathcal{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_m\}$ be a set of vectors in $\in \mathbb{R}^n$, then the span of $\mathcal{X}$ is the set:

$$\text{span}(\mathcal{X}) = \{\mathbf{y} \in \mathbb{R}^n | \mathbf{y} \text{ is a linear combination of vectors in } \mathcal{X}\} \tag{3.35}$$

Definition 3.29 (Linear Independence). Let $\mathbf{x}_1, \dots, \mathbf{x}_m$ be vectors in $\in \mathbb{R}^n$. The vectors $\mathbf{x}_1, \dots, \mathbf{x}_m$ are *linearly dependent* if there exists $\alpha_1, \dots, \alpha_m \in \mathbb{R}$, not all zero, such that

$$\alpha_1 \mathbf{x}_1 + \cdots + \alpha_m \mathbf{x}_m = 0 \tag{3.36}$$

If the set of vectors $\mathbf{x}_1, \dots, \mathbf{x}_m$ is not linearly dependent, then they are *linearly independent* and Equation 3.36 holds just in case $\alpha_i = 0$ for all $i = 1, \dots, n$.

8. Basis

DEFINITION 3.35 (Basis). Let $\mathcal{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ be a set of vectors in $\mathbb{R}^n$. The set $\mathcal{X}$ is called a *basis* of $\mathbb{R}^n$ if $\mathcal{X}$ is a linearly independent set of vectors and every vector in $\mathbb{R}^n$ is in the span of $\mathcal{X}$. That is, for any vector $\mathbf{w} \in \mathbb{R}^n$ we can find scalar values $\alpha_1, \ldots, \alpha_m$ such that

$$\mathbf{w} = \sum_{i=1}^{m} \alpha_i \mathbf{x}_i \tag{3.37}$$

THEOREM 3.37. *If $\mathcal{X}$ is a basis of $\mathbb{R}^n$, then $\mathcal{X}$ contains precisely n vectors.*

LEMMA 3.38. *Let $\{\mathbf{x}_1, \ldots, \mathbf{x}_{m+1}\}$ be a linearly dependent set of vectors in $\mathbb{R}^n$ and let $\mathcal{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ be a linearly independent set. Further assume that $\mathbf{x}_{m+1} \neq \mathbf{0}$. Assume $\alpha_1, \ldots, \alpha_{m+1}$ are a set of scalars, not all zero, so that*

$$\sum_{i=1}^{m+1} \alpha_i \mathbf{x}_i = \mathbf{0} \tag{3.41}$$

For any $j \in \{1, \ldots, m\}$ such that $\alpha_j \neq 0$, if we replace $\mathbf{x}_j$ in the set $\mathcal{X}$ with $\mathbf{x}_{m+1}$, then this new set of vectors is linearly independent.

REMARK 3.39. This lemma proves an interesting result. If $\mathcal{X}$ is a basis of $\mathbb{R}^m$ and $\mathbf{x}_{m+1}$ is another, non-zero, vector in $\mathbb{R}^m$, we can swap $\mathbf{x}_{m+1}$ for any vector $\mathbf{x}_j$ in $\mathcal{X}$ as long as when we express $\mathbf{x}_{m+1}$ as a linear combination of vectors in $\mathcal{X}$ the coefficient of $\mathbf{x}_j$ is not zero. That is, since $\mathcal{X}$ is a basis of $\mathbb{R}^m$ we can express:

$$\mathbf{x}_{m+1} = \sum_{i=1}^{m} \alpha_i \mathbf{x}_i$$

As long as $\alpha_j \neq 0$, then we can replace $\mathbf{x}_j$ with $\mathbf{x}_{m+1}$ and still have a basis of $\mathbb{R}^m$.

9. Rank

DEFINITION 3.40 (Row Rank). Let $\mathbf{A} \in \mathbb{R}^{m \times n}$. The *row rank* of $\mathbf{A}$ is the size of the largest set of row (vectors) from $\mathbf{A}$ that are linearly independent.

EXERCISE 36. By analogy define the *column rank* of a matrix. [Hint: You don't need a hint.]

THEOREM 3.41. *If* $\mathbf{A} \in \mathbb{R}^{m \times n}$ *is a matrix, then elementary row operations on* A *do not change the row rank.*

THEOREM 3.42. *If* $\mathbf{A} \in \mathbb{R}^{m \times n}$ *is a matrix, then the row rank of* $\mathbf{A}$ *is equal to the column rank of* $\mathbf{A}$. *Further,* $\text{rank}(\mathbf{A}) \leq \min\{m, n\}$.

THEOREM 3.43. *If* $\mathbf{A} \in \mathbb{R}^{m \times m}$ *(i.e.,* $\mathbf{A}$ *is a square matrix) and* $\text{rank}(\mathbf{A}) = m$, *then* $\mathbf{A}$ *is invertible.*

DEFINITION 3.44. Suppose that $\mathbf{A} \in \mathbb{R}^{m \times n}$ and let $m \leq n$. Then $\mathbf{A}$ has *full row rank* if $\text{rank}(\mathbf{A}) = m$.

Reference: Kevin G Ross, ISM206 Optimization theory and applications, Lecture notes,
https://courses.soe.ucsc.edu/courses/ism206 , Date accessed: 16.07.2017
Christopher Grin, Linear Programming: Penn State Math 484 Lecture Notes, Version 1.8.2.1
http://www.personal.psu.edu/cxg286/Math484_V1.pdf , Date accessed: 16.07.2017