#### 10-725: Optimization

## Lecture 3: September 3

Lecturer: Barnabas Poczos/Ryan Tibshirani Scribes: Kirstin Early, Willie Neiswanger, Nicole Rafidi

**Note**: LaTeX template courtesy of UC Berkeley EECS dept.

**Disclaimer**: These notes have not been subjected to the usual scrutiny reserved for formal publications. They may be distributed outside this class only with the permission of the Instructor.

# 3.1 Pivot Transformation

Consider the following LP problem

$$\min z \quad \text{s.t.}$$

$$2x_1 + 2x_2 + 2x_3 + x_4 + 4x_5 = z$$

$$4x_1 + 2x_2 + 13x_3 + 3x_4 + x_5 = 17$$

$$x_1 + x_2 + 5x_3 + x_4 + x_5 = 7$$

$$x_1, \dots, x_5 \ge 0$$
(3.1)

We define a **pivot** to be a nonzero element in the above problem, for example  $3x_4$ . We choose a pivot, then use it to eliminate the corresponding variable (in this case,  $x_4$ ) from the remaining equations. After we apply a pivot transformation we arrive at an equivalent system of equations, where the solution set is left the same. This process is equivalent to Gaussian elimination. The pivot transformation proceeds as follows.

We begin with an initial system of equations

$$2x_{1} + 2x_{2} + 2x_{3} + x_{4} + 4x_{5} = z$$
  

$$4x_{1} + 2x_{2} + 13x_{3} + 3x_{4} + x_{5} = 17$$
  

$$x_{1} + x_{2} + 5x_{3} + x_{4} + x_{5} = 7$$
(3.2)

We choose  $3x_4$  as a pivot and aim to eliminate  $x_4$  from the remaining equations. We begin by dividing the second equation by 3 to get the system

$$2x_{1} + 2x_{2} + 2x_{3} + x_{4} + 4x_{5} = z$$

$$\frac{4x_{1}}{3} + \frac{2x_{2}}{3} + \frac{13x_{3}}{3} + x_{4} + \frac{x_{5}}{3} = \frac{17}{3}$$

$$x_{1} + x_{2} + 5x_{3} + x_{4} + x_{5} = 7$$
(3.3)

Note that the coefficients for  $x_4$  in the first and third equations are now the same as the coefficient for  $x_4$  in

Fall 2012

the second equation. We then subtract the the second equation from the first and third equations to get

$$\frac{2x_1}{3} + \frac{4x_2}{3} - \frac{7x_3}{3} + 0x_4 + \frac{11x_5}{3} = z - \frac{17}{3} 
\frac{4x_1}{3} + \frac{2x_2}{3} + \frac{13x_3}{3} + x_4 + \frac{x_5}{3} = \frac{17}{3} 
-\frac{x_1}{3} + \frac{x_2}{3} + \frac{2x_3}{3} + 0x_4 + \frac{2x_5}{3} = \frac{4}{3}$$
(3.4)

Note that now  $x_4$  has been eliminated from equations one and three (i.e. the coefficient for  $x_4$  in these equations is now 0). Next, we choose another pivot, such as  $\frac{x_2}{3}$ . We use this pivot (in the third equation) to eliminate  $x_2$  from the first and second equations by subtracting 4-times the third equation from the first equation, to get

$$2x_{1} + 0x_{2} - 5x_{3} + 0x_{4} + x_{5} = z - 11$$

$$\frac{4x_{1}}{3} + \frac{2x_{2}}{3} + \frac{13x_{3}}{3} + x_{4} + \frac{x_{5}}{3} = \frac{17}{3}$$

$$-\frac{x_{1}}{3} + \frac{x_{2}}{3} + \frac{2x_{3}}{3} + 0x_{4} + \frac{2x_{5}}{3} = \frac{4}{3}$$
(3.5)

And then by subtracting 2-times the third equation from the second equation, to get

$$2x_{1} + 0x_{2} - 5x_{3} + 0x_{4} + x_{5} = z - 11$$
  

$$2x_{1} + 0x_{2} + 3x_{3} + x_{4} - x_{5} = 3$$
  

$$-\frac{x_{1}}{3} + \frac{x_{2}}{3} + \frac{2x_{3}}{3} + 0x_{4} + \frac{2x_{5}}{3} = \frac{4}{3}$$
(3.6)

We also multiply the third equation by 3 to get

$$2x_1 + 0x_2 - 5x_3 + 0x_4 + x_5 = z - 11$$
  

$$2x_1 + 0x_2 + 3x_3 + x_4 - x_5 = 3$$
  

$$-x_1 + x_2 + 2x_3 + 0x_4 + 2x_5 = 4$$
(3.7)

We can now rewrite this as

$$\begin{aligned} z &+ 2x_1 &+ 0x_2 &- 5x_3 &+ 0x_4 &+ x_5 &= -11 \\ 2x_1 &+ 0x_2 &+ 3x_3 &+ x_4 &- x_5 &= 3 \\ -x_1 &+ x_2 &+ 2x_3 &+ 0x_4 &+ 2x_5 &= 4 \end{aligned}$$
(3.8)

The system written in this way is said to be in **canonical form**. In particular, we say that this system of equations is in canonical form with respect to (-z),  $x_4$ , and  $x_2$  variables, and that

 $x_1, x_3, x_5$  are independent (nonbasic) variables  $-z, x_4, x_2$  are dependent (basic) variables, which are expressed in terms of other variables.

If we set the nonbasic variables to zero, then we can get values for the basic variables. For example, setting  $x_1 = 0$ ,  $x_3 = 0$ , and  $x_5 = 0$ , i.e. by setting  $X_N = (x_1, x_3, x_5) = (0, 0, 0)$  we get that z = 11,  $X_B = (x_4, x_2) = (3, 4)$ .

Note that if we had initially started with a different pivot, say  $x_1$  and then  $x_4$  (instead of  $x_4$  and then  $x_2$ ), then we would've arrived at z = 3,  $X_N = (x_2, x_3, x_5) = (0, 0, 0)$ , and  $X_B = (x_1, x_4) = (-4, 11)$ .

#### 3.1.1 The Goals of Pivots

The primary goal of the pivoting procedure is to reduce the original LP problem to canonical form. After a reduction to canonical form, it is easy to find a (basic) solution to the system of equations. To do so, we simply set the nonbasic variables to zero. Note that this procedure returns a valid solution, because the pivoting process does not alter the solution set of the system of equations (i.e. after pivots, the systems are equivalent).

However, the basic solution returned by the pivoting process might be:

- 1. Not a feasible solution. This could occur because of the boundary constraint (e.g. we required that  $x_1, \ldots, x_5 > 0$ , in the case of the previous example).
- 2. Not an optimal solution. For example, z might not be minimal.

### 3.1.2 Formal Definition of Canonical Form

A system of *m* equations and *n* nonbasic variables  $(x_j)_{j=1}^n$  is in **canonical form** with respect to *m* basic variables  $(x_{j_i})_{i=1}^m$  if and only if either of the two equivalent conditions hold:

- 1.  $x_{j_i}$  has one coefficient in equation *i*, for all  $j \neq i$ .
- 2. The system  $I_m X_B + A X_N = b$

Using the previous example, the second condition would be equivalent to

	[1	0	0	$\left[-z\right]$		2	-5	1]	$\begin{bmatrix} x_1 \end{bmatrix}$		[-11]	]
M =	0	1	0	$x_4$	+	2	3	-1	$x_3$	=	3	
M =	0	0	1	$x_2$		$\lfloor -1 \rfloor$	2	2	$x_5$		4	

We more-formally define a **basic solution** to be the solution gained by setting  $X_N = 0$  in the canonical form.

For example, if we let

$$X_B = (x_1, \dots, x_m)$$
$$X_N = (x_{m+1}, \dots, x_n)$$

then we'd set  $x_{m+1}, \ldots, x_n = 0$ , and would arrive at the solutions  $x_1 = b_1, \ldots, x_m = b_m$ . This basic solution is feasible if and only if  $b_1 \ge 0, \ldots, b_m \ge 0$ .

# 3.2 Warming up to the simplex algorithm

#### 3.2.1 Starting from Canonical Form

Let's assume that we have a system of equations in canonical form, as well as a feasible basic solution:

Recall that we have the additional constraints that  $x_i \ge 0 \forall i = 1, ..., n$ , and that the coefficients  $c_i$  are called relative cost factors (these may be different in a different canonical form basis). This system can be written in matrix notation as well:

$$\begin{bmatrix} 1 & 0 & C \\ 0 & I_m & A \end{bmatrix} \begin{bmatrix} -z \\ x_B \\ x_N \end{bmatrix} = \begin{bmatrix} -z_0 \\ b \end{bmatrix}$$
(3.9)

In this form, the basic solution is  $z = z_0$ , and the assignments are  $x_B = b$  and  $x_N = 0$ . The solution is feasible if and only if  $b \ge 0$ .

#### 3.2.2 Improving a Nonoptimal Basic Solution

Let's continue the example from the previous section:

$$\begin{aligned} -z + 2x_1 + 0x_2 - 5x_3 + 0x_4 + 1x_5 &= -11 \\ 2x_1 + 0x_2 + 3x_3 + 1x_4 - 1x_5 &= 3 \\ - x_1 + 1x_2 + 2x_3 + 0x_4 + 2x_5 &= 4 \end{aligned}$$
 (3.10)

The basic feasible solution here is that z = 11, with assignments  $x_B = (x_4, x_2) = (3, 4)$  and  $x_N = (x_1, x_3, x_5) = (0, 0, 0)$ .

Recall that our goal is to minimize z such that  $x_i \ge 0$ . How can we improve our current solution?

Examine the relative cost factors of the variables in  $x_N$ , and notice that  $x_3$  has a negative cost factor. If we make  $x_3 > 0$ , we have a chance to reduce z.

To see how much we can increase  $x_3$  while maintaining feasibility, we keep  $x_3, -z$ , and  $x_B = (x_2, x_4)$  as parameters while setting  $x_N = (x_1, x_5) = (0, 0)$ . Our system now becomes:

$$z = 11 - 5x_3$$
  

$$x_4 = 3 - 3x_3$$
  

$$x_2 = 4 - 2x_3$$
  
(3.11)

So we can decrease  $x_3$  as much as we want as long as  $x_4 \ge 0$  and  $x_2 \ge 0$ . The limiting factor in this case is  $x_4$  which is 0 when  $x_3 = 1$ . If we think of this in matrix form as Ax = b:

$$\begin{bmatrix} 2 & 0 & 3 & 1 & -1 \\ -1 & 1 & 2 & 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ x_2 \\ x_3 \\ x_4 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$
(3.12)

We are comparing the ratios  $\frac{b_1}{A_{13}} = \frac{3}{3} = 1$  and  $\frac{b_2}{A_{23}} = \frac{4}{2} = 2$ , so in this case  $\frac{b_1}{A_{13}} < \frac{b_2}{A_{23}}$ . The conclusion is the same: we can only increase  $x_3$  enough to make  $x_4 = 0$ .

Thus, we move  $x_4$  from  $x_B$  to  $x_N$ , and  $x_3$  moves into  $x_B$ .

This is equivalent to taking equation 3.10 and using Gaussian elimination with  $x_3$  as a pivot variable:

$$\begin{aligned} -z + 2x_1 + 0x_2 - 5x_3 + 0x_4 + 1x_5 &= -11 \\ 2x_1 + 0x_2 + 3x_3 + 1x_4 - 1x_5 &= 3 \\ - x_1 + 1x_2 + 2x_3 + 0x_4 + 2x_5 &= 4 \end{aligned}$$
 (3.13)

We perform three changes: multiply the second equation by  $\frac{5}{3}$  and add it to the first equation, multiply the second equation by  $\frac{-2}{3}$  and add it to the third equation, and finally divide the second equation by 3. This gets us:

$$-z + (\frac{10}{3} + 2)x_1 + 0x_2 + 0x_3 + \frac{5}{3}x_4 + (1 - \frac{5}{3})x_5 = -11 + 5$$
  
$$\frac{2}{3}x_1 + 0x_2 + 1x_3 + \frac{1}{3}x_4 - \frac{1}{3}x_5 = 1$$
  
$$(-\frac{4}{3} - 1)x_1 + 1x_2 + 0x_3 - \frac{2}{3}x_4 + (2 + \frac{2}{3})x_5 = 4 - 2$$
  
(3.14)

Simplifying:

$$-z + \frac{16}{3}x_1 + 0x_2 + 0x_3 + \frac{5}{3}x_4 - \frac{2}{3}x_5 = -6$$
  
$$-\frac{2}{3}x_1 + 0x_2 + 1x_3 + \frac{1}{3}x_4 - \frac{1}{3}x_5 = 1$$
  
$$-\frac{7}{3}x_1 + 1x_2 + 0x_3 - \frac{2}{3}x_4 + \frac{8}{3}x_5 = 2$$
(3.15)

Now the current solution is z = 6, with assignments  $x_N = (x_1, x_4, x_5) = (0, 0, 0)$  and  $x_B = (x_3, x_2) = (1, 2)$ . Examining the first equation, we now see that  $c_5 < 0$ , so we can improve our current solution by increasing  $x_5$  the same way we just increased  $x_3$ . Just as before, we keep  $z, x_B, x_5$  as parameters and set  $x_N$  to 0 so that our system becomes:

$$z = 6 - \frac{2}{3}x_{5}$$

$$x_{3} = 1 + \frac{1}{3}x_{5}$$

$$x_{2} = 2 - \frac{8}{3}x_{5}$$
(3.16)

In this case how much we can increase  $x_5$  is limited only by  $x_2$ . In matrix form we have:

$$\begin{bmatrix} \frac{2}{3} & 0 & 1 & \frac{1}{3} & -\frac{1}{3} \\ -\frac{7}{3} & 1 & 0 & -\frac{2}{3} & \frac{8}{3} \end{bmatrix} \begin{bmatrix} 0 \\ x_2 \\ x_3 \\ 0 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$
(3.17)

Performing the same comparison as before, between  $\frac{b_1}{A_{15}}$  and  $\frac{b_2}{A_{25}}$ , we get that  $-3 < \frac{3}{4}$ , however, we want to keep every  $x_i \ge 0$ , so again, the limiting factor is  $x_2$ . Note that if both of these ratios were less than zero, we could minimize z to  $-\infty$  by increasing  $x_5$ .

We can thus pivot on  $x_5$  by setting it to  $\frac{3}{4}$ . While we could pivot on either equation containing  $x_5$  in 3.15, the only pivot that will get us a feasible solution is the one that moves  $x_2$  into  $x_N$ . Applying Gaussian elimination by pivoting on  $x_5$  in the third equation, our system is now:

$$-z + \frac{57}{12}x_1 + \frac{1}{4}x_2 + 0x_3 + \frac{3}{2}x_4 + 0x_5 = -\frac{11}{2}$$

$$\frac{3}{8}x_1 + \frac{1}{8}x_2 + 1x_3 + \frac{1}{4}x_4 - 0x_5 = \frac{5}{4}$$

$$- \frac{7}{8}x_1 + \frac{3}{8}x_2 + 0x_3 - \frac{2}{8}x_4 + 1x_5 = \frac{3}{4}$$
(3.18)

The solution is that  $z = \frac{11}{2}$ , with  $x_B = (x_3, x_5) = (\frac{5}{4}, \frac{3}{4})$  and  $x_N = (x_1, x_2, x_4) = (0, 0, 0)$ . Looking at the first equation, we see that every cost factor is positive, meaning that we cannot improve z any further - we're done!

## 3.3 The Simplex Algorithm

#### 3.3.1 Key Components of the Algorithm

In this example, we traced the simplex algorithm by hand. The key takeaways are:

- 1. We check for optimality by making sure that each  $c_j \ge 0 \forall j$ . If this holds, then our basic feasible solution is optimal.
- 2. If the optimality condition is not met, we then must bring one variable into  $x_B$  (the variable with the most negative  $c_i$ ), and send another variable to  $x_N$  in exchange.
- 3. We choose the variable to send to  $x_N$  by ensuring that our non-negativity constraints  $(x_j \ge 0 \forall j)$  are met.

#### 3.3.2 Overview of Algorithm Steps

In the simplex algorithm, we assume that we start with a system in feasible canonical form:

$$\begin{array}{rcl} -z &+ & 0x_B &= & C^T x_N &= & -z_0 \\ & & & Ix_B &+ & Ax_N &= & b \end{array}$$
(3.19)

The starting solution is  $x_B = b$ ,  $x_N = 0$ , and  $z = z_0$ .

Then we complete the following steps:

1. Find the smallest reduced cost, i.e. find  $c_s = min_jc_j$  and  $s = argmin_jc_j$ 

- 2. Test for optimality: if  $c_s \ge 0$  return the current solution and stop.
- 3. If  $c_s < 0$ , then we move  $x_s$  to  $x_B$
- 4. Test for unbounded z: if every entry in the  $s^{th}$  column of A, then  $z^* = -\infty$  by setting  $x_s \to \infty$
- 5. If z is bounded, then we choose the variable to move to  $x_N$  in exchange for  $x_s$ . We find  $r = argmin_{i|A_{is}>0} \frac{b_i}{A_{is}}$  Thus we pivot on the basic variable that is in the  $r^t h$  row of A.
- 6. Pivot on  $A_{rs}$  to get a new basic feasible solution  $\tilde{z}$ , regardless of whether z changes (if  $b_r = 0$ , it will not change z).  $\tilde{z} = z_0 b_r \frac{-c_s}{A_{rs}}$

How do we know that our solution (following the method in step 5) will still be feasible?

**Lemma 3.1** The new basic solution remains feasible,  $\tilde{b_j} \ge 0 \forall j$ 

#### **Proof:**

$$\tilde{b_j} = b_j - b_r \frac{A_{js}}{A_{rs}} \tag{3.20}$$

Assume  $b_j \ge 0$  (i.e. that the previous solution was feasible). If  $A_{js} \le 0$ , then  $\tilde{b_j} \ge 00$ . If  $A_{js} > 0$ , then recall that we chose r so that  $\frac{b_r}{A_{rs}} \le \frac{b_j}{A_{js}} \forall j$ . Thus,  $\tilde{b_j} \ge 0$ .

Notice that in step 6 we specifically stated that z does not have to change. This could potentially lead to an infinite cycle between values that do not change z. We can prevent this by using Bland's Rule:

Whenever the pivot in the simplex method would result in not change of the objective z, do the following:

- 1. Choose from among the incoming column choices  $j \in s$  s.t.  $c_j < 0$  the column with the smallest index j. (Step 3)
- 2. Choose from among the multiple outgoing column choices the eligible column with the smallest index. (Step 5)

## 3.4 Summary of the simplex algorithm

The simplex algorithm can be used for linear programs in standard form:

$$\begin{array}{ll} \underset{x \ge 0}{\text{minimize}} & c^T x\\ \text{subject to} & Ax = b, \end{array}$$
(3.21)

where  $c, x \in \mathbb{R}^n$ ,  $A \in \mathbb{R}^{m \times n}$ , and  $b \in \mathbb{R}^m$ .

**Theorem 3.2** A basic feasible solution is optimal with total cost  $z_0$  if all relative cost factors are nonnegative; *i.e.*, if  $c \ge 0 \in \mathbb{R}^n$ .

**Proof:** 

Since  $c \ge 0$ , increasing any  $x_j : j \in \{1, ..., m\}$  (and adjusting the other  $x_j$ 's to obey the equality constraints) cannot decrease the objective z.

**Theorem 3.3** A basic feasible solution is the unique optimal solution with total cost  $z_0$  if  $c_j > 0$  for all non-basic variables.

**Theorem 3.4** Assuming "non-degeneracy" at each iteration (i.e.,  $b > 0 \in \mathbb{R}^m$ ), the simplex algorithm converges in finitely many steps.

**Proof:** If a linear program is nondegenerate, then it does not have any redundant constraints. Therefore, it is not possible to cycle through any of the finite number of bases (i.e., selections of basic variables) when executing the simplex algorithm.

### 3.5 Phase I

Phase I finds a starting basic feasible solution in canonical form, which Phase II then improves to the optimal solution.

Introduce *m* new optimization variables to the original linear program in (3.21), one for each row of the constraint matrix:  $x_{\text{new}} = [x_1, ..., x_n, x_{n+1}, ..., x_{n+m}]^T$  and augment the constraint matrix  $A_{\text{new}} = [A \ I_m]$ . Solve the new linear program

$$\begin{array}{ll} \underset{x_{\text{new}} \ge 0}{\text{minimize}} & w = \sum_{i=n+1}^{n+m} x_{\text{new},i} \\ \text{subject to} & A_{\text{new}} x_{\text{new}} = b. \end{array}$$
(3.23)

**Theorem 3.5** (3.23) has a feasible optimal solution such that  $x_{n+1} = \dots = x_{n+m} = 0$  iff (3.21) has a feasible solution.

# 3.6 Example of the simplex algorithm

The minimization problem:

 $\begin{array}{rcl} \underset{x \ge 0}{\text{minimize}} & 2x_1 + & x_2 + 2x_3 + & x_4 + 4x_5 & =z \\ \text{subject to} & 4x_1 + & 2x_2 + 13x_3 + & 3x_4 + x_5 & =17 \\ & x_1 + & x_2 + 5x_3 + & x_4 + x_5 & =7. \end{array}$ (3.24)

### 3.6.1 Phase I

Introduce two new optimization variables,  $x_6, x_7 \ge 0$ , disregard the original optimization objective, and solve the new minimization problem

The following stages of the algorithm are shown in tableaux form.

First, convert the problem to canonical form by subtracting the two constraint equations from the objective to give the basic variables  $(x_6, x_7)$  zero coefficients in the objective:

Basic variable	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	RHS
-w	-5	-3	-18	-4	-2	0	0	-24
$x_6$	4	2	13	3	1	1	0	17
$x_7$	1	1	5	1	1	0	1	7

Since  $c_3 = -18 = \min c_j < 0$ , pivot on  $x_3$ .  $x_3$  will replace  $x_6$  as a basic variable since  $r_6 = \frac{17}{13} < \frac{7}{5} = r_7$ :

Basic variable	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	RHS
-w	7/13	-3/13	0	2/13	-8/13	18/13	0	-6/13
$x_3$	4/13	2/13	1	3/13	1/13	1/13	0	17/13
$x_7$	-7/13	3/13	0	-2/13	8/13	-5/13	1	6/13

Now  $c_5 = -8/13 = \min c_j < 0$ , so pivot on  $x_5$ . Because  $r_7 = \frac{6/13}{8/13} < \frac{17/13}{1/13} = r_3$ ,  $x_7$  will go out when  $x_5$  comes in:

Basic variable	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	RHS
-w	0	0	0	0	0	1	1	0
$x_3$	3/8	1/8	1	1/4	0	1/8	-1/8	5/4
$x_5$	-7/8	3/8	0	-1/4	1	-5/8	13/8	3/4

Since  $c \ge 0$ , we have reached an optimal feasible solution to (3.25). Set the nonbasic variables to zero:  $x_1 = x_2 = x_4 = x_6 = x_7 = 0$ ; set the basic variables to the RHS:  $x_3 = 5/4$  and  $x_5 = 3/4$ .

### 3.6.2 Phase II

By Theorem 3.5, the optimal feasible solution found in Phase I will also be a feasible solution to the original linear program, with the original minimization objective,  $z = 2x_1 + x_2 + 2x_3 + x_4 + 4x_5$ . We can include this cost function in the previous table and continue with Phase II to improve this feasible solution:

Basic variable	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	x <sub>7</sub>	RHS
-z	2	1	2	1	4	0	0	0
$x_3$	3/8	1/8	1	1/4	0	1/8	-1/8	5/4
$x_5$	-7/8	3/8	0	-1/4	1	-5/8	13/8	3/4

The final solution to this linear program is  $x = [0, 2, 1, 0, 0]^T$ , which achieves z = 4.