IEOR 4004
Formal Simplex Method

Here we outline key steps in the Simplex Method applied to an LP in standard form:

$$\begin{aligned} \max z &= c^T x \\ \text{Subject to:} \quad Ax &= b \\ x &\geq 0. \end{aligned}$$

Here A is an $m \times n$ matrix.

Step I. Find an initial basic feasible solution. This is **not** a trivial step. Among other points, the LP may not have any feasible solution, and so no bfs. We will skip this step for now, and deal with it in greater detail later.

Step II. Assuming we have found a bfs, we now proceed to optimize. Let us assume that the initial basis is made up of the following column set: $\{1, 2, \ldots, m\}$. In general we cannot assume that this is the case, however by renaming variables if necessary it will be.

After performing some elementary row operations, we have a set of equations:

$$\begin{array}{llllllllll} z & & & & & -\bar{c}_{m+1}\,x_{m+1} & -\bar{c}_{m+2}\,x_{m+2} & \cdots & -\bar{c}_n\,x_n & = \bar{z}_0 \\ \hline & x_1 & & & & +\bar{a}_{1,m+1}\,x_{m+1} & +\bar{a}_{1,m+2}\,x_{m+2} & \cdots & +\bar{a}_{1,n}\,x_n & = \bar{b}_1 \\ & & x_2 & & & +\bar{a}_{2,m+1}\,x_{m+1} & +\bar{a}_{2,m+2}\,x_{m+2} & \cdots & +\bar{a}_{2,n}\,x_n & = \bar{b}_2 \\ & & & \cdots & & & & & & \\ & & & & x_{m-1} & +\bar{a}_{m-1,m+1}\,x_{m+1} & +\bar{a}_{m-1,m+2}\,x_{m+2} & \cdots & +\bar{a}_{m-1,n}\,x_n & = \bar{b}_{m-1} \\ & & & & x_m & +\bar{a}_{m,m+1}\,x_{m+1} & +\bar{a}_{m,m+2}\,x_{m+2} & \cdots & +\bar{a}_{m,n}\,x_n & = \bar{b}_m \end{array}$$

Here we are using "bars" over the coefficients to indicate that they were obtained by elementary row operations. This "tableau" indicates the following b.f.s:

$$x_j = \bar{b}_j, \quad 1 \leq j \leq m, \quad \text{(basic variables)} \tag{1a}$$

$$x_j = 0, \quad m+1 \leq j, \quad \text{(nonbasic variables)} \tag{1b}$$

$$z = \bar{z}_0. \tag{1c}$$

Since this is a bfs, it is feasible, so $\bar{b} \geq 0$. Note that the first equation can be written as:

$$z = \bar{z}_0 + \bar{c}_{m+1}x_{m+1} + \bar{c}_{m+2}x_{m+2} \cdots + \bar{c}_n x_n.$$

The fact that we obtained this equation by elementary row operations implies that this equation is valid for **every** feasible x. So we have two cases

1. $\bar{c}_j \leq 0$ for $m+1 \leq j \leq n$. In this case the bfs is **optimal** and we **stop**. Do you see why?

   Here is an example with $m = 3$ and $n = 5$.

$$z = 20 - 3x_4 - x_5.$$

2. There is a nonbasic variable x_j (so $m+1 \leq j$) with $\bar{c}_j > 0$. In this case our bfs may not be optimal (we do not know). Do you see why? In fact there may be many nonbasic variables x_j with $\bar{c}_j > 0$. We will pick any one of them. For now, it does not matter which (later we will discuss policies for making this choice). This variable x_k that we choose will be the **entering** variable. We will move to a new bfs where k is part of the basis, and some currently basic variable becomes nonbasic. That is the **leaving** variable. The choice of this variable is not arbitrary.

To see how we pick the leaving variable, suppose (as we have done before) that we perturb the solution we have by (slowly) increasing x_k from zero, while keeping all other nonbasic variables fixed at zero, and adjusting the basic variables so as to maintain feasibility. Using the tableau, we will obtain the following equations that are similar to equations (1):

$$\begin{aligned} z &= \bar{z}_0 + \bar{c}_k x_k && \text{(2a)} \\ x_j &= \bar{b}_j - \bar{a}_{j,k}\, x_k, \quad 1 \leq j \leq m, \quad \text{(basic variables)} && \text{(2b)} \\ x_j &= 0, \quad m+1 \leq j \neq k, \quad \text{(other nonbasic variables).} && \text{(2c)} \end{aligned}$$

Since $\bar{c}_k > 0$ we can see that if we can make $x_k > 0$ we will indeed increase the value of z. As a result:

if $a_{j,k} \leq 0$ for **every** j then the problem is **unbounded**

Can you see why?

Here is an example with $m = 3$.

$$\begin{aligned} z &= 10 + 2x_k && \text{(3a)} \\ x_1 &= 15 + 2x_k && \text{(3b)} \\ x_2 &= 4 + 3x_k && \text{(3c)} \\ x_3 &= 0 + 5x_k. && \text{(3d)} \end{aligned}$$

If the problem is unbounded as we have just discussed, again we **STOP**.

Suppose, on the other hand, some $\bar{a}_{j,k} > 0$. **Example:**

$$x_1 = 10 - 2x_k.$$

In this example $j = 1$, $\bar{b}_1 = 10$ and $\bar{a}_{1,k} = 2$. You can see that if $x_k > 5$ then the solution becomes infeasible because x_j goes negative. And so we immediately know that $x_k \leq 5$.
As the example shows, every index j with $\bar{a}_{j,k} > 0$ gives us an upper bound on how large x_k could become. We want to make x_k as large as possible so as to increase z as much as we can given our choice of entering variable.

Here is an example with $m = 4$.

$$\begin{aligned} z &= 10 + 2x_k && \text{(4a)} \\ x_1 &= 15 - 2x_k && \text{(4b)} \\ x_2 &= 4 - 3x_k && \text{(4c)} \\ x_3 &= 0 + 5x_k && \text{(4d)} \\ x_4 &= 20 - 8x_k. && \text{(4e)} \end{aligned}$$

In this example we have $\bar{a}_{1,k} = 2$, $\bar{a}_{2,k} = 3$, $\bar{a}_{3,k} - 5$ and $\bar{a}_{4,k} = -8$. As a result,

$$\begin{aligned} x_k &\leq 15/2, \quad \text{or else } x_1 \text{ becomes negative} \\ x_k &\leq 4/3, \quad \text{or else } x_2 \text{ becomes negative} \\ x_k &\leq 20/8, \quad \text{or else } x_4 \text{ becomes negative} \end{aligned}$$

Since we are trying to make x_k as large as possible, we set to the minimum of the three above limiting conditions: $x_k = 4/3$, which means that x_2 goes to zero. Thus, $\boldsymbol{x_2}$ **is the variable that leaves the basis**. In other words, the basis is now made up of variables $\{1, 3, 4, k\}$.

The general formula is, therefore:

$$x_k = \min\left\{ \frac{\bar{b}_i}{\bar{a}_{i,k}} : \bar{a}_{i,k} > 0 \right\}.$$

This is the so-called **ratio test**. Any row that attains the minimum in this ratio will correspond to a basic variable that will go to zero once we set x_k to the minimum. There could be several such variables, and we pick one arbitrarily to leave the basis. If, say, variable x_i leaves the basis, then x_i and x_k have traded places.

Of course, we have to *show* that it is a basis, i.e. that the set of columns of A is linearly independent. Let's see why that is the case in the previous example.

Back to previous example. Let's write out the tableau before x_k enters the basis. We only write the part of the tableau involving the basic columns and x_k.

$$\begin{array}{ccccccc} z & & & & -2\,x_k & = & 10 \\ \hline & x_1 & & & +2\,x_k & = & 15 \\ & & x_2 & & +3\,x_k & = & 4 \\ & & & x_3 & -5\,x_k & = & 0 \\ & & & & x_4 \;\; +8\,x_k & = & 20 \end{array}$$

As we discussed above, it is the *second constraint* row of this system that provides the limiting value for x_k. Dividing by 3 this row we get:

$$\begin{array}{ccccccc} z & & & & -2\,x_k & = & 10 \\ \hline & x_1 & & & +2\,x_k & = & 15 \\ & & x_2/3 & & +\;\;x_k & = & 4/3 \\ & & & x_3 & -5\,x_k & = & 0 \\ & & & & x_4 \;\; +8\,x_k & = & 20 \end{array}$$

And now we zero out all other entries of x_k in its column:

$$\begin{array}{cccccccr} z & +2/3\,x_2 & & & & = & 10 + 8/3 \\ \hline x_1 & -2/3\,x_2 & & & & = & 15 - 8/3 \\ & 1/3\,x_2 & & & +x_k & = & 4/3 \\ & +5/3\,x_2 & +x_3 & & & = & 20/3 \\ & -8/3\,x_2 & & +x_4 & & = & 20 - 32/3 \end{array}$$

Note that writing the equations this way accomplishes two tasks:

- First, it tells us that $\{1, 2, 4, k\}$ forms a basis, because the corresponding columns are linearly independent. Can you see why?
- Second, from this new tableau we can read the new bfs: $x_1 = 15 - 8/3$, $x_3 = 20/3$, $x_4 = 20 - 32/3$, $x_k = 4/3$. Further, $z = 10 + 8/3$.

At this point we have completed one iteration of the Simplex Method. We have a new bfs, and the objective value has increased from 10 to $10 + 8/3$. The process can now repeat, from the tableau we just wrote, above.

Using matrix notation.

We can improve on the efficiency of the algorithm and also require less space to describe it by using matrix notation. Thus, at a given iteration of the simplex method, let us denote by $\boldsymbol{B}$ the submatrix of A corresponding to the given basis. We will *also* call B the basis, because it is a basis in the linear algebra sense (a largest submatrix that is linearly independent).

Similarly, we will denote by $\boldsymbol{N}$ the submatrix of A corresponding to the nonbasic variables. The vectors of basic and nonbasic variables will be denoted x_B and x_N, respectively, and the corresponding subvectors of c will be denoted c_B and c_N. The LP can therefore be written as:

$$\begin{aligned} & \max\ c_B^T x_B + c_N^T x_N && \text{(5a)} \\ \text{Subject to}\quad & B x_B + N x_N = b && \text{(5b)} \\ & x_B \geq 0,\ x_N \geq 0. && \text{(5c)} \\ & && \text{(5d)} \end{aligned}$$

To get this system into tableau format we write it as

$$\begin{array}{cccc} z & -c_B^T x_B & -c_N^T x_N & = & 0 \\ \hline & B x_B & +N x_N & = & b \end{array}$$

We simplify this by performing row operations so that "B" becomes "I". Hence the tableau becomes:

$$\begin{array}{cccc} z & -c_B^T x_B & -c_N^T x_N & = & 0 \\ \hline & x_B & +B^{-1} N x_N & = & B^{-1} b \end{array}$$

And then we zero-out the coefficients of B in the objective:

$$\begin{array}{cccc} z & & -(c_N^T - c_B^T B^{-1} N) x_N & = & c_B^T B^{-1} b \\ \hline & x_B & +B^{-1} N x_N & = & B^{-1} b \end{array}$$

Using the tableau in this format we can revisit the basic steps of the simplex method.

1. We see that if $c_N - c_B^T B^{-1} N \leq 0$ then the current bfs is optimal. For a given nonbasic variable x_j, the quantity $c_j - c_B^T B^{-1} N_j$ is what we called $\bar{c}_j$ above. Here, N_j is the column of N corresponding to x_j.

2. We pick any nonbasic variable x_k with $c_k - c_B^T B^{-1} N_k > 0$ as the entering variable. This is the so-called **pricing step.**

3. In terms of the above notation, for any row i, the quantity we called $\bar{a}_{i,k}$ equals $[B^{-1}N_k]_i$. In this expression,

$$B^{-1}N_k$$

is an n-vector, and $[B^{-1}N_k]_i$ is its i^{th} entry. Similarly, $\bar{b}_i = [B^{-1}b]_i$. So the ratio test can be restated as

$$x_k = \min \left\{ \frac{[B^{-1}b]_i}{[B^{-1}N_k]_k} \; : \; [B^{-1}N_k]_i > 0 \right\}.$$

Further efficiencies.

In many practical applications the matrix A will be very sparse. So B and N are sparse and often B^{-1} is sparse, as well. In fact, even c may be sparse. We want to make use of this sparsity so as to obtain rapid computations.

First, recall that we pick, as entering variable, some nonbasic variable x_k such that

$$\bar{c}_k = c_k - c_B^T B^{-1} N_k > 0.$$

The above formulas are efficiently evaluated by **not** ever computing the product $B^{-1}N$ and storing it as a matrix. Instead, we just carry out the product $c_B^T B^{-1}N_k$ each time we need it. We compute this product by paying attention to the sparsity structure, which typically results in far fewer computations than we would need if we treated the matrices as dense. Also note that we may discover a good candidate for entering variable without enumerating all nonbasic variables. That leads to further efficiences.

And, second and most important, we do not explicitly compute the matrix $B^{-1}N$. We only need to compute $B^{-1}N_k$, where x_k is the entering variable, in order to perform thee ratio test. As in the previous paragraph, $B^{-1}N_k$ is computed by taking advantage of the sparsity structure.

A technical detail.

In a typical pivot we perform steps such as "multiply row 1 by 1/5" or "add, to row 4, -3 times row 2". To make an example easy, suppose we have a case with four rows. Then the first example (multiply row 1 by 1/5) is achieved by multiplying the (current) constraints by the matrix

$$\begin{pmatrix} 1/5 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \tag{6}$$

which is clearly invertible. And the second example is attained by multiplying the current constraints (again, from the left) by the matrix

$$\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & -3 & & 1 \end{pmatrix} \tag{7}$$

Either matrix is called an "eta" matrix. It follows that the expression we have for the constraints at any step of the simplex method can be obtained by multiplying, from the left, the original matrix, by the sequence of eta matrices. If you pause to think about it, the sequence must be written in *reverse* order, with the last eta matrix in leftmost position, and the first eta matrix in the rightmost position.

In particular, if B is the basis at the end of a pivot, the corresponding product of eta matrices must equal B^{-1}. Do you see why?

Tue.Jul.19.211715.2016@rabbitchaser