Lecture 10: The revised simplex method

1 Calculating the reduced costs

Last time, we used the notation $A_{\mathcal{I}}$ and $\mathbf{x}_{\mathcal{I}}$ to pick out columns of a matrix, or entries from a vector, with indices given by a sequence $\mathcal{I}$.

We used this to write down a formula for a basic solution to the system of equations $A\mathbf{x} = \mathbf{b}$: if the basic variables are numbered by $\mathcal{B}$, and the nonbasic variables are numbered by $\mathcal{N}$, then the corresponding basic solution has $\mathbf{x}_{\mathcal{B}} = (A_{\mathcal{B}})^{-1}\mathbf{b}$ and $\mathbf{x}_{\mathcal{N}} = \mathbf{0}$. A general solution has $\mathbf{x}_{\mathcal{B}} = (A_{\mathcal{B}})^{-1}(\mathbf{b} - A_{\mathcal{N}}\mathbf{x}_{\mathcal{N}})$.

Let's continue by doing the same thing for the objective function. In general, this is an expression of the form $\mathbf{c}^{\mathsf{T}}\mathbf{x} = c_1x_1 + c_2x_2 + \cdots + c_nx_n$. This, too, can be split up by basic and nonbasic variables: $\mathbf{c}^{\mathsf{T}}\mathbf{x} = (\mathbf{c}_{\mathcal{B}})^{\mathsf{T}}\mathbf{x}_{\mathcal{B}} + (\mathbf{c}_{\mathcal{N}})^{\mathsf{T}}\mathbf{x}_{\mathcal{N}}$. If we want to know the objective value at a basic solution, we set $\mathbf{x}_{\mathcal{B}} = (A_{\mathcal{B}})^{-1}\mathbf{b}$ and $\mathbf{x}_{\mathcal{N}} = \mathbf{0}$ to get $(\mathbf{c}_{\mathcal{B}})^{\mathsf{T}}(A_{\mathcal{B}})^{-1}\mathbf{b}$.

What about the reduced costs? Well, let's write $(\mathbf{c}_{\mathcal{B}})^{\mathsf{T}}\mathbf{x}_{\mathcal{B}} + (\mathbf{c}_{\mathcal{N}})^{\mathsf{T}}\mathbf{x}_{\mathcal{N}}$ just in terms of $\mathbf{x}_{\mathcal{N}}$. To do this, we use the formula $\mathbf{x}_{\mathcal{B}} = (A_{\mathcal{B}})^{-1}(\mathbf{b} - A_{\mathcal{N}}\mathbf{x}_{\mathcal{N}})$ and get

$$\begin{aligned}\mathbf{c}^{\mathsf{T}}\mathbf{x} &= (\mathbf{c}_{\mathcal{B}})^{\mathsf{T}}(A_{\mathcal{B}})^{-1}(\mathbf{b} - A_{\mathcal{N}}\mathbf{x}_{\mathcal{N}}) + (\mathbf{c}_{\mathcal{N}})^{\mathsf{T}}\mathbf{x}_{\mathcal{N}} \\ &= (\mathbf{c}_{\mathcal{B}})^{\mathsf{T}}(A_{\mathcal{B}})^{-1}\mathbf{b} + \Big((\mathbf{c}_{\mathcal{N}})^{\mathsf{T}} - (\mathbf{c}_{\mathcal{B}})^{\mathsf{T}}(A_{\mathcal{B}})^{-1}A_{\mathcal{N}}\Big)\mathbf{x}_{\mathcal{N}}.\end{aligned}$$

So the row vector of our reduced costs is given by the formula $(\mathbf{c}_{\mathcal{N}})^{\mathsf{T}} - (\mathbf{c}_{\mathcal{B}})^{\mathsf{T}}(A_{\mathcal{B}})^{-1}A_{\mathcal{N}}$.

We're writing the product $(\mathbf{c}_{\mathcal{B}})^{\mathsf{T}}(A_{\mathcal{B}})^{-1}$ a lot, so let's give it a name: let's call it $\mathbf{u}^{\mathsf{T}}$. (It has a transpose because it's a row vector.) We'll learn much more about this vector later; for now, it's just a vector that's handy in our calculations!

All this can be summarized by putting our dictionaries in matrix form:

$$\begin{array}{c} \zeta \;=\; \mathbf{u}^{\mathsf{T}}\mathbf{b} \;+\; \Big((\mathbf{c}_{\mathcal{N}})^{\mathsf{T}} - \mathbf{u}^{\mathsf{T}}A_{\mathcal{N}}\Big)\mathbf{x}_{\mathcal{N}} \\ \hline \mathbf{x}_{\mathcal{B}} = (A_{\mathcal{B}})^{-1}\mathbf{b} - (A_{\mathcal{B}})^{-1}A_{\mathcal{N}}\mathbf{x}_{\mathcal{N}} \end{array}$$

When doing the ordinary simplex method, it would be bad to recompute the dictionary at every step using these formulas, because computing $(A_{\mathcal{B}})^{-1}$ at every step is expensive. On the other hand, this can be useful to compute a dictionary if, for some reason, all you know is which variables are basic.

We will also use these formulas in the revised simplex method: an improvement on the simplex method which is more computationally efficient by avoiding unnecessary calculations.

[1] This document comes from the Math 3272 course webpage: https://facultyweb.kennesaw.edu/mlavrov/courses/3272-fall-2022.php

2 The revised simplex method

2.1 Finding a basic solution using matrices

Consider the following problem:

Problem 1. *You are an adventurer who has just slain a dragon. You're standing in the dragon's lair, admiring the hoard of gold, jewels, magic artifacts, and so forth. Unfortunately, you can't take it all. You're limited by volume (whatever will fit in your backpack: say, 100 units) and weight (whatever you can carry: say, 30 kg), and you want to find the most valuable combination of objects possible under those constraints.*

We assume that all kinds of objects are continuous enough that we can say "take x_i kilograms of the i^{th} object" for any plausible x_i, and plentiful enough that there's no constraints other than the total weight and volume. However, there's lots of them: maybe you have a table along the lines of

	Gold	*Silver*	*Rubies*	*Diamonds*	*Magic rings*	*Spell scrolls*	*Stale cookies*
Price/kg	*2*	*1*	*3*	*5*	*2*	*5*	*0*
Volume/kg	*3*	*3*	*1*	*2*	*4*	*5*	*5*

How do you figure out the most efficient combination of precious items?

We just have two constraints here, aside from nonnegativity constraints:

- if $x_1, \dots, x_7$ measure the total amount of the objects in kilograms, then we want $x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 \leq 30$.
- The volume/kg row of the table gives us the constraint on volume: $3x_1 + 3x_2 + x_3 + 2x_4 + 4x_5 + 5x_6 + 5x_7 \leq 100$.

The price/kg row gives us the objective function: we want to maximize $2x_1 + x_2 + 3x_3 + 5x_4 + 2x_5 + 5x_6$.

The challenging part is the number of variables (most of which will not be used in the optimal solution). If 7 variables (9 when we add slack variables) is not bad enough for you, you can imagine a more varied hoard for which the problem would be much worse.

We will do something unusual with the notation today. To make it easier to connect our dictionary to the matrix formulas, we will name our slack variables x_8 and x_9, putting them at the end of our vector $\mathbf{x}$. The variables that describe our linear program are:

$$\mathbf{c}^\mathsf{T} = \begin{bmatrix} 2 & 1 & 3 & 5 & 2 & 5 & 0 & 0 & 0 \end{bmatrix}$$
$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 3 & 3 & 1 & 2 & 4 & 5 & 5 & 0 & 1 \end{bmatrix}$$
$$\mathbf{b} = \begin{bmatrix} 30 \\ 100 \end{bmatrix}.$$

Normally, our first choice of basic variables would be $\mathcal{B} = (8, 9)$: the slack variables. To try out our new formulas, we'll take $\mathcal{B} = (1, 7)$: we'll consider filling up our backpack with gold and stale cookies.

The first thing to compute is $(A_{\mathcal{B}})^{-1}$. We have

$$A_{\mathcal{B}} = A_{(1,7)} = \begin{bmatrix} 1 & 1 \\ 3 & 5 \end{bmatrix} \implies (A_{\mathcal{B}})^{-1} = \begin{bmatrix} \frac{5}{2} & -\frac{1}{2} \\ -\frac{3}{2} & \frac{1}{2} \end{bmatrix}.$$

No matter what we do next, we probably want to know the basic feasible solution (though you have only my word that it's going to turn out feasible) and the associated objective value:

$$(A_{\mathcal{B}})^{-1}\mathbf{b} = \begin{bmatrix} \frac{5}{2} & -\frac{1}{2} \\ -\frac{3}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 30 \\ 100 \end{bmatrix} = \begin{bmatrix} 25 \\ 5 \end{bmatrix} \qquad (\mathbf{c}_{\mathcal{B}})^{\mathsf{T}}(A_{\mathcal{B}})^{-1}\mathbf{b} = \begin{bmatrix} 2 & 0 \end{bmatrix} \begin{bmatrix} 25 \\ 5 \end{bmatrix} = 50.$$

So we are currently taking 25 kg of gold and 5 kg of stale cookies, which will bring us a profit of 50 in whatever currency.

2.2 Being careful about what we compute

The fundamental idea of the revised simplex method is that now we are going to be very careful not to do too much work. In particular, to fill in the entire dictionary at this point, we'd need to compute $(A_{\mathcal{B}})^{-1}A_{\mathcal{N}}$, and that's a really annoying matrix multiplication. Can we avoid it?

Well, our first step in the simplex method is to choose an entering variable. This only involves looking at the reduced costs. We have two choices:

- We could go ahead and compute the entire row of reduced costs. This has the formula $(\mathbf{c}_{\mathcal{N}})^{\mathsf{T}} - (\mathbf{c}_{\mathcal{B}})^{\mathsf{T}}(A_{\mathcal{B}})^{-1}A_{\mathcal{N}}$. To compute this as efficiently as possible, we'd begin by finding $\mathbf{u}^{\mathsf{T}} = (\mathbf{c}_{\mathcal{B}})^{\mathsf{T}}(A_{\mathcal{B}})^{-1}$, then calculating $(\mathbf{c}_{\mathcal{N}})^{\mathsf{T}} - \mathbf{u}^{\mathsf{T}}A_{\mathcal{N}}$. This avoids having to deal with the product $(A_{\mathcal{B}})^{-1}A_{\mathcal{N}}$.
- If we use Bland's rule for pivoting, then we get to save some work. After computing $\mathbf{u}^{\mathsf{T}}$, we can find the reduced cost of variable x_i by calculating $c_i - \mathbf{u}^{\mathsf{T}}A_i$: x_i's component of $(\mathbf{c}_{\mathcal{N}})^{\mathsf{T}} - \mathbf{u}^{\mathsf{T}}A_{\mathcal{N}}$. Bland's rule says that we can stop once we find the first positive reduced cost.

  This helps counteract the disadvantage of Bland's rule: its slowness. We don't mind doing more pivot steps if each pivot step becomes faster!

Either way, we begin by computing

$$\mathbf{u}^{\mathsf{T}} = (\mathbf{c}_{\mathcal{B}})^{\mathsf{T}}(A_{\mathcal{B}})^{-1} = \begin{bmatrix} 2 & 0 \end{bmatrix} \begin{bmatrix} \frac{5}{2} & -\frac{1}{2} \\ -\frac{3}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 5 & -1 \end{bmatrix}.$$

Let's try computing the reduced costs one at a time. Silver (x_2) gives us

$$c_2 - \mathbf{u}^{\mathsf{T}}A_2 = 1 - \begin{bmatrix} 5 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = 1 - (5 \cdot 1 - 1 \cdot 3) = -1.$$

Doing the same calculation for rubies (x_3) gives us $c_3 - \mathbf{u}^{\mathsf{T}}A_3 = 3 - (5 \cdot 1 - 1 \cdot 1) = -1$, but for diamonds ($x_4$) we finally get $c_4 - \mathbf{u}^{\mathsf{T}}A_4 = 5 - (5 \cdot 1 - 1 \cdot 2) = 2$, which is positive.

Now that we know x_4 is our entering variable, we want to find our leaving variable. The trick is that we don't need all of $(A_\mathcal{B})^{-1}A_\mathcal{N}$ to do this! We only care about x_4's column of that matrix, which is given by

$$(A_\mathcal{B})^{-1}A_4 = \begin{bmatrix} \frac{5}{2} & -\frac{1}{2} \\ -\frac{3}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{3}{2} \\ -\frac{1}{2} \end{bmatrix}.$$

Remember that our dictionary has $\mathbf{x}_\mathcal{B} = (A_\mathcal{B})^{-1}(\mathbf{b} - A_\mathcal{N}\mathbf{x}_\mathcal{N})$ in it, so we are *subtracting* $(A_\mathcal{B})^{-1}A_4x_4$. Our shortlist of leaving variables comes from negative coefficients, which means we're looking for positive values in $(A_\mathcal{B})^{-1}A_4$. The $\frac{3}{2}$ is positive, which puts our first variable in $\mathcal{B} = (1, 7)$ on our shortlist: x_1.

If we had more than one variable on our shortlist, we'd continue by computing the ratios between the column $(A_\mathcal{B})^{-1}A_4$ we just found, and the column $(A_\mathcal{B})^{-1}\mathbf{b}$ that we computed earlier. But in this case, we can skip that step: x_1 is the only candidate.

So now we know x_4 is our entering variable and x_1 is our leaving variable. We're done, right? We can just go to the next step with $\mathcal{B} = (4, 7)$.

Not so fast! We really don't want to compute $(A_\mathcal{B})^{-1}$ again at each step. (In this example, it's only a 2×2 matrix inverse, but for larger systems, the inverse is much harder to compute.) Let's try to compute the inverse of $A_{(4,7)}$ (the new inverse we want) from the inverse of $A_{(1,7)}$ (the old inverse we have).

Here's the idea. Using our old $\mathcal{B}$, we already know all the entries of

$$(A_{(1,7)})^{-1}A_{(1,4,7)} = \begin{bmatrix} \frac{5}{2} & -\frac{1}{2} \\ -\frac{3}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 3 & 2 & 5 \end{bmatrix} = \begin{bmatrix} 1 & \frac{3}{2} & 0 \\ 0 & -\frac{1}{2} & 1 \end{bmatrix}.$$

The column corresponding to x_4, we just computed. The columns corresponding to x_1 and x_7 must form an identity matrix by the definition of a matrix inverse.

What we want to see is a result of the form

$$(A_{(4,7)})^{-1}A_{(1,4,7)} = \begin{bmatrix} ? & ? \\ ? & ? \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 3 & 2 & 5 \end{bmatrix} = \begin{bmatrix} ? & 1 & 0 \\ ? & 0 & 1 \end{bmatrix}$$

because whatever $(A_{(4,7)})^{-1}$ is, multiplying by it must turn the x_4 and x_7 columns of A into the identity matrix.

We can figure out what row operations turn $(A_{(1,7)})^{-1}A_{(1,4,7)}$ (the first 2×3 matrix above) into $(A_{(4,7)})^{-1}A_{(1,4,7)}$ (the second 2×3 matrix above). To do this, we multiply the first row by $\frac{2}{3}$ (to turn $\frac{3}{2}$ into 1) and then add half the result to the second row (to turn $-\frac{1}{2}$ into 0).

But row operations are just matrix multiplication from the left. So those same row operations will turn $(A_{(1,7)})^{-1}$ into $(A_{(4,7)})^{-1}$, which is what we want! We take $(A_{(1,7)})^{-1}$, multiply the first row by $\frac{2}{3}$, and then add half the result to the second row:

$$(A_{(1,7)})^{-1} = \begin{bmatrix} \frac{5}{2} & -\frac{1}{2} \\ -\frac{3}{2} & \frac{1}{2} \end{bmatrix} \qquad \rightsquigarrow \qquad A_{(4,7)}^{-1} = \begin{bmatrix} \frac{5}{3} & -\frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} \end{bmatrix}.$$

2.3 A summary of the revised simplex method

Let's summarize what we did in a set of instructions, so that we can do it again.

0. At the beginning of each pivot step, we should already know $\mathcal{B}$ (the sequence of basic variables) and $(A_{\mathcal{B}})^{-1}$ (the inverse of the corresponding matrix).

1. We calculate $(A_{\mathcal{B}})^{-1}\mathbf{b}$ (which tells us the current basic feasible solution) and $\mathbf{u}^{\mathsf{T}} = (\mathbf{c}_{\mathcal{B}})^{\mathsf{T}}(A_{\mathcal{B}})^{-1}$ (which will be useful for calculations).

2. To determine the entering variable, we compute reduced costs: as before, we want a positive reduced cost for maximizing and a negative one for minimizing.

   The reduced cost of x_i is given by $c_i - \mathbf{u}^{\mathsf{T}}A_i$. We can compute this for all the variables, but if we're using Bland's rule, we can find them one at a time until we get one with the correct sign.

3. Let x_j be the entering variable. We compute $(A_{\mathcal{B}})^{-1}A_j$ to find the coefficients of x_j in our dictionary. The rules are slightly different due to a negative sign in our formulas:

   - The leaving variables on our shortlist correspond to the *positive* components of $(A_{\mathcal{B}})^{-1}A_j$.
   - If multiple variables are on our shortlist, choose the one with the smallest ratio, dividing a component of $(A_{\mathcal{B}})^{-1}\mathbf{b}$ by the corresponding component of $(A_{\mathcal{B}})^{-1}A_j$.

4. Let x_k be the leaving variable, and suppose that it's the i^{th} variable in the list $\mathcal{B}$. Our new sequence of basic variables will be $\mathcal{B}'$ where x_k is replaced by x_j.

   Before we begin the next pivot step, we must compute $(A_{\mathcal{B}'})^{-1}$. Here, let $\mathcal{I}$ be the combination of $\mathcal{B}$ and $\mathcal{B}'$: all the previously basic variables, together with j.

   To do this, find the row reduction steps that take $(A_{\mathcal{B}})^{-1}A_{\mathcal{I}}$ (which should have pivots in $\mathcal{B}$'s columns) to $(A_{\mathcal{B}'})^{-1}A_{\mathcal{I}}$ (which should have pivots in $\mathcal{B}'$'s columns).

   Then, apply those steps to $(A_{\mathcal{B}})^{-1}$ to get $(A_{\mathcal{B}'})^{-1}$.

2.4 One more pivot step

Let's do another pivot step for this problem. Everything will now be in terms of the basis $(4, 7)$.

1. We calculate $(A_{(4,7)})^{-1}\mathbf{b} = \begin{bmatrix} \frac{5}{3} & -\frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 30 \\ 100 \end{bmatrix} = \begin{bmatrix} \frac{50}{3} \\ \frac{40}{3} \end{bmatrix}$ and

$$\mathbf{u}^{\mathsf{T}} = (\mathbf{c}_{(4,7)})^{\mathsf{T}}(A_{(4,7)})^{-1} = \begin{bmatrix} 5 & 0 \end{bmatrix} \begin{bmatrix} \frac{5}{3} & -\frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{25}{3} & -\frac{5}{3} \end{bmatrix}.$$

2. To determine the entering variable, we compute the reduced costs, one at a time. We can skip x_1, since it was just the leaving variable, so we start with x_2:

$$c_2 - \mathbf{u}^{\mathsf{T}}A_2 = 1 - \left(\frac{25}{3} \cdot 1 - \frac{5}{3} \cdot 3\right) = -\frac{7}{3}$$

$$c_3 - \mathbf{u}^\mathsf{T} A_3 = 3 - \left(\frac{25}{3} \cdot 1 - \frac{5}{3} \cdot 1\right) = -\frac{11}{3}$$
$$c_5 - \mathbf{u}^\mathsf{T} A_5 = 2 - \left(\frac{25}{3} \cdot 1 - \frac{5}{3} \cdot 4\right) = \frac{1}{3}.$$

Since c_5 has a positive reduced cost, it will be our entering variable.

3. To find the leaving variable, we compute $(A_{(4,7)})^{-1} A_5 = \begin{bmatrix} \frac{5}{3} & -\frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}$.

   Both rows are positive, so both x_4 and x_7 are on our shortlist of leaving variables. Their current values in the basic solution are given by $(A_{(4,7)})^{-1}\mathbf{b}$: they're $\frac{50}{3}$ and $\frac{40}{3}$. The ratio is $\frac{50/3}{1/3} = 50$ for x_4 and $\frac{40/3}{2/3} = 20$ for x_7, so x_7 leaves the basis.

4. We are turning the basis $(4,7)$ into $(4,5)$. To compute the new inverse matrix $(A_{(4,5)})^{-1}$, we want to find the row reduction that takes

$$(A_{(4,7)})^{-1} A_{(4,5,7)} = \begin{bmatrix} 1 & \frac{1}{3} & 0 \\ 0 & \frac{2}{3} & 1 \end{bmatrix} \qquad \rightsquigarrow \qquad (A_{(4,5)})^{-1} A_{(4,5,7)} = \begin{bmatrix} 1 & 0 & ? \\ 0 & 1 & ? \end{bmatrix}.$$

   To get there, we must multiply the second row by $\frac{3}{2}$ (to turn the $\frac{2}{3}$ into 1) and then subtract $\frac{1}{3}$ of that from the first row (to turn $\frac{1}{3}$ into 0). So let's do the same things to $(A_{(4,7)})^{-1}$:

$$(A_{(4,7)})^{-1} = \begin{bmatrix} \frac{5}{3} & -\frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} \end{bmatrix} \qquad \rightsquigarrow \qquad (A_{(5,7)})^{-1} = \begin{bmatrix} 2 & -\frac{1}{2} \\ -1 & \frac{1}{2} \end{bmatrix}$$

   We are ready for our next pivot step.

3 Lessons learned

The revised simplex method might be obnoxious to do by hand (and I don't encourage it, except to make sure that it makes sense). But there's a few reasons to do what we've done today:

- We can carefully think about the number of operations required for the simplex method, and how it scales with the number of variables and number of equations.
- Our concerns when designing the revised simplex method—we were kind of worried about multiplying $(A_\mathcal{B})^{-1} A_\mathcal{N}$, and we were *really* worried about computing $(A_\mathcal{B})^{-1}$—are common to many algorithms. You are unlikely to have to write computer code to implement the revised simplex method; however, it is much more important to understand what operations are cheap, what operations are expensive, and how we can avoid the expensive ones.

Later in the semester, the dictionary formulas from the beginning of the lecture will also be put to new, unexpected uses.