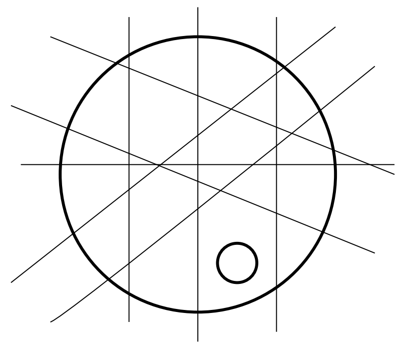
Math 1b TA: Padraic Bartlett

Recitation 1: Matrices!

Week 1 *Caltech 2011*

1 Random Question

Suppose you have a circle S in the plane with diameter 100, and 32 lines in the plane that intersect your circle. Show that there is always a circle of radius 3 contained within the interior of S that doesn't intersect any of these lines.



2 Administrivia

Here are most of the random administrative details for the course:

- My email: padraic@caltech.edu
- My office: 360 Sloan.
- My office hours: 10-11pm on Sunday night/ by appointment.
- My website: www.its.caltech.edu/∼padraic. Course notes for every recitation will be posted here, ideally within a few days of the recitation.
- HW policy: The course-wide policy is that every student is allowed at most 1 late HW without a note from the deans or health center, with an extension of at most one week. Homeworks after this one will require a note from the health center or the deans: it bears noting that both entities are remarkably kind, and as long as your reason for needing more time is not something like "all-night SC2 marathon," they'll grant an extension.
- Random questions: I post a random question at the start of every recitation! If you've seen the material in rec before, and get distracted, they're meant to offer something mathematically interesting to focus on until the lecture returns to a place you haven't seen. Because we're at Caltech, and pretty much anything we talk about in Math

1 *some* of you have seen before, it struck me as a decent way to avoid boring some students without losing others. If you solve any of them, tell me! I am always interested to see solutions.

3 Matrices and Solving Systems of Equations

Why do we study matrices? Later on in mathematics, you'll learn that there are dozens of subtle mathematical reasons to examine matrices. However, from a beginning linear algebra student's perspective, there's one main reason we study matrices: to solve systems of linear equations.

For example, suppose that we wanted to find all of the values of w, x, y, z that satisfy the following three equations simultaneously:

$$\begin{bmatrix} w & + & 2x & + & 3y & + & 0z & = & 0 \\ -w & + & 2x & + & 0y & + & z & = & 2 \\ 0w & + & 4x & + & y & + & z & = & 4. \\ 2w & + & 4x & + & 6y & + & 0z & = & 0 \end{bmatrix}$$

How can we do this? Well, as it turns out, matrices provide an excellent way to systematically find solutions to systems like the above. Specifically, we have the following three-step process:

1. First, we turn this system of equations into a matrix, made out of the coefficients of the above equations:

$$\begin{bmatrix} w & + & 2x & + & 3y & + & 0z & = & 0 \\ -w & + & 2x & + & 0y & + & z & = & 2 \\ 0w & + & 4x & + & y & + & z & = & 4 \\ 2w & + & 4x & + & 6y & + & 0z & = & 0 \end{bmatrix} \longrightarrow \left(\begin{array}{cccc|c} 1 & 2 & 3 & 0 & 0 \\ -1 & 2 & 0 & 1 & 2 \\ 0 & 4 & 1 & 1 & 4 \\ 2 & 4 & 6 & 0 & 0 \end{array}\right).$$

   But wait – what's that vertical bar mean? That bar is a bit of shorthand that allows us to distinguish between the matrix entries that correspond to our variables (w, x, y, z), and those that correspond to the constants our equations above are trying to be equal to. In essence, the matrix entries to the left of the bar above are what we're really trying to study and deal with: the entries to the right are in some sense just placeholders, which we have in our matrix so we can keep track of the constants more easily.

2. Now, what we want to do is reduce the coefficient-part of our matrix (again, the part to the left of the vertical bar) into its reduced row-echelon form[1]. How do we do this?

[1] A matrix is in **reduced row-echelon form** if

- the **leading coefficient** (first nonzero entry) of every row is 1,
- the leading coefficient of the k-th row is strictly to the right of the leading coefficient of the $k-1$-th row,
- the leading coefficient of any row is the only nonzero entry in its respective column, and
- all of the zero rows are at the bottom of the matrix.

Via **pivots**!

First, we should say what a pivot is:

Definition. If A is a matrix such that its (i, j)-th entry a_{ij} is nonzero, **pivoting** at (i, j) is simply performing the following two matrix operations:

- First, multiply the i-th row by $1/a_{ij}$. This makes the (i, j)-th entry 1.
- Then, subtract a_{kj}-many copies of the i-th row from the k-th row, for every $k \neq i$. This makes every other entry in the j-th column equal to 0.

To reduce our matrix of coefficients to reduced row-echelon form, then, we just have to do the following:

- Find the furthest-left column j_1 that's not made entirely of zeroes.
- By swapping rows, make it so that a_{1,j_1} is nonzero.
- Pivot at $(1, j_1)$.
- Now, find the furthest-left column j_2 that's not equal to j_1 and not made entirely of zeroes.
- By swapping rows, make it so that a_{2,j_2} is nonzero.
- Pivot at $(2, j_2)$.
- *Now,* find the furthest-left column j_3 that's not equal to j_1 or j_2 and not made entirely of zeroes ... and basically keep repeating this process until your matrix is in reduced row-echelon form!

In practice, what usually winds up happening is that you will first pivot at (1,1), then at (2,2), then at (3,3) ... so on and so forth until your matrix is in reduced row-echelon form. However, you sometimes might wind up having to switch rows around to make this possible; also, you might have all-zero columns which mean that you have to skip over columns occasionally! This is why we write up the above process: it will always work, and without fail transform your matrix into reduced row-echelon form.

These descriptions are kind of dry without an example: so let's calculate one. Specifically, let's return to our matrix

$$\left(\begin{array}{rrrr|r} 1 & 2 & 3 & 0 & 0 \\ -1 & 2 & 0 & 1 & 2 \\ 0 & 4 & 1 & 1 & 4 \\ 2 & 4 & 6 & 0 & 0 \end{array}\right).$$

If we follow the above process, we want to pivot on the furthest-left entry we can get into row 1. In this case, we don't have to switch any rows to do this, as $(1, 1)$ is already nonzero: so we can just pivot there, at (1,1).

To do this, first multiply row 1 by $1/a_{11} = 1$, and then

- subtract $a_{21} = -1$ copies of row 1 from row 2,
- subtract $a_{31} = 0$ copies of row 1 from row 3, and finally
- subtract $a_{41} = 2$ copies of row 1 from row 4.

Doing this yields the matrix

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 0 & 0 \\ 0 & 4 & 3 & 1 & 2 \\ 0 & 4 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Now, we look to pivot in the furthest-left column $\neq 1$ in row 2. Again, we don't have to swap any rows to do this, as (2,2) is already nonzero.

So, let's pivot there, at (2,2). To do this, we first multiply row 2 by $1/a_{22} = 1/4$, and then

- subtract $a_{12} = 2$ copies of row 2 from row 1,
- subtract $a_{32} = 4$ copies of row 2 from row 3, and finally
- subtract $a_{42} = 0$ copies of row 2 from row 4.

Doing this yields the matrix

$$\left(\begin{array}{cccr|r} 1 & 0 & 3/2 & -1/2 & -1 \\ 0 & 1 & 3/4 & 1/4 & 1/2 \\ 0 & 0 & -2 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Finally, we want to find the furthest-left nonzero entry we can get in row 3 that doesn't live in columns 1 or 2: again, as (3,3) is already nonzero, we don't have to move any rows around and can just pivot there.

Pivoting at (3,3) just requires us to multiply row 3 by $1/(-2) = -1/2$, and then

- subtract $a_{13} = 3/2$ copies of row 3 from row 1,
- subtract $a_{23} = 3/4$ copies of row 3 from row 2, and finally
- subtract $a_{43} = 0$ copies of row 3 from row 4.

Doing this yields the matrix

$$\left(\begin{array}{cccr|r} 1 & 0 & 0 & -1/2 & 1/2 \\ 0 & 1 & 0 & 1/4 & 5/4 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right),$$

which is now in reduced row-echelon form.

3. With this matrix in hand, finding the solutions to our original system is actually pretty simple. How? Well, just do the following two things:

   (a) To determine if your system is consistent – i.e. if any solutions exist at all! – simply look at all of the zero-rows in the coefficient part of our matrix. If the constant entry in this row is not also zero, your system is inconsistent – i.e. no solution exists. That's because this row of the matrix is corresponding to the equation $0x_1 + 0x_2 + \ldots 0x_n =$(something not zero), which we know is impossible. Conversely, if this never happens, your system is consistent!

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & -1/2 & 1/2 \\ 0 & 1 & 0 & 1/4 & 5/4 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

   The teal row above is the only zero row in our matrix: because its corresponding entry on the side of our matrix that's keeping track of the constants is 0, we know our matrix is consistent.

   (b) If your system *is* consistent, it's also really easy to express the solutions to your system! To do this, divide the variables you're solving for into two kinds:

      - the "fixed" variables: those variables whose column in the coefficient matrix is one of the leading-coefficient columns we turned into a unit vector when reducing our matrix.
      - the "free" variables: the variables whose column in the coefficient matrix was not one of the leading-coefficient columns.

   For example, in the matrix we've been working with, the variables w, x, and y are all fixed, as their columns (highlighted in pink) all correspond to the unit vectors we made while pivoting. However, the variable z (column highlighted in yellow) is free, as we didn't pivot there.

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & -1/2 & 1/2 \\ 0 & 1 & 0 & 1/4 & 5/4 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

   Once you've divided up your variables this way, look at the rows of our reduced matrix. Specifically, notice the following: in every nonzero row, there is exactly *1* fixed variable. Therefore, we can always write the fixed variables in terms of the free variables! Therefore, we can always express the solutions to our system of equations in the form

$$\Big(\text{(fixed var. 1 expressed via free var's)}, \ldots \text{(fixed var. m expressed via free var's)}, \\ \text{(free variable 1)}, \ldots \text{(free variable n)}\Big).$$

For a concrete example of what we're talking about above, return to our matrix

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & -1/2 & 1/2 \\ 0 & 1 & 0 & 1/4 & 5/4 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

It corresponds to the three linear equations

$$\left[\begin{array}{cccccccc} w & + & & & & -z/2 & = & 1/2 \\ & & x & + & & z/4 & = & 5/4 \\ & & & & y & & = & -1 \end{array}\right],$$

which we can rewrite as

$$\begin{aligned} w &= 1/2 + z/2 \\ x &= 5/4 - z/4 \\ y &= -1. \end{aligned}$$

In other words, we can express the solutions to our system as the collection of vectors of the form

$$(1/2 + z/2, 5/4 - z - 4, -1, z),$$

where z is allowed to range over all of $\mathbb{R}$.

4 Matrices and Operations

The above process shows how we can use matrices to systematically find solutions to systems of linear equations, via this system of "reducing" a matrix to its reduced row-echelon form. When we did this reduction, we used a number of different ways of "manipulating" the rows of the matrix. Specifically, we used the following three transforms repeatedly:

- Switching two rows.
- Multiplying a row by a constant.
- Adding a constant multiple of one row to another.

One question we can ask, then, is whether we can perform such operations via matrices! In other words: can we find a matrix U such that for any $n \times n$ matrix A,

$$UA = A \text{ with row 1 and row 5 of } A \text{ switched?}$$

Or

$$UA = A \text{ with row 1 multiplied by } \pi?$$

As it turns out, yes! Such matrices are called **elementary matrices**, and have the following forms:

$$E = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \lambda & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

If the λ is in the (i, j)-th spot, multiplying A on the left by this matrix adds λ times row j of A to row i of A.

$$E = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \lambda & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

If the λ is in the (i, i)-th spot, multiplying A on the left by this matrix multiplies A's i-th row by λ.

$$E = \begin{pmatrix} 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & \dots & \vdots & \dots & 1 & \dots & 0 \\ 0 & 0 & \dots & 1 & \dots & \vdots & \dots & 0 \\ 0 & 0 & \ddots & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix}$$

The matrix above is the standard identity matrix with its i-th and j-th columns (highlighted) switched. Multiplying A on the left by this matrix switches A's i-th and j-th rows.

Prove these properties to yourself if you don't believe them!