Matrix Inversion

We saw in the last set of notes that one can rewrite a system of linear equations

$$\begin{aligned} A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n &= b_1 \\ \vdots &= \vdots \\ A_{m1}x_1 + A_{m2}x_2 + \cdots + A_{mn}x_n &= b_m \end{aligned}$$

as a matrix equation

$$Ax = b,$$

where A is the $m \times n$ matrix with entries A_{ij}, and x is the vector $(x_1, \ldots, x_n) \in \mathbb{R}^n$, and b is the vector $(b_1, \ldots, b_m) \in \mathbb{R}^m$. Here we specialize to the case where $m = n$, so that we have the same number of equations as unknowns and A is a square matrix.

In this case, we saw that (most of the time) we can apply row reductions to turn this into a diagonal one:

$$Ax = \begin{bmatrix} A_{11} & 0 & \cdots & 0 \\ 0 & A_{22}0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \hat{b}_1 \\ \hat{b}_2 \\ \vdots \\ \hat{b}_n \end{bmatrix},$$

the the $\hat{b}_j$ coeffcients are the ones you get by applying the row operations to the vector b (multiplying on the left as before). Now it's not too hard to turn the matrix on the left hand side into the identity matrix: you divide the jth row by A_{jj} to get

$$x = Ix = \begin{bmatrix} \frac{1}{A_{11}} & 0 & \cdots & 0 \\ 0 & \frac{1}{A_{22}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \frac{1}{A_{nn}} \end{bmatrix} \begin{bmatrix} A_{11} & 0 & \cdots & 0 \\ 0 & A_{22}0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \frac{\hat{b}_1}{A_{11}} \\ \frac{\hat{b}_2}{A_{22}} \\ \vdots \\ \frac{\hat{b}_n}{A_{nn}} \end{bmatrix} = \begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \\ \vdots \\ \tilde{b}_n \end{bmatrix}$$

Now, let's pause for a moment and think about what we're doing in terms of the matrix equation. We're starting with

$$Ax = b,$$

and we want to solve for the unknown vector x. Remember that A is an $n \times n$ matrix, so it could very well have a multiplicative inverse A^{-1}. In fact, most of the time A will have an inverse. If it does, then we can multiply both sides of the equation by A^{-1} on the left to get

$$x = A^{-1}Ax = A^{-1}b.$$

So, in the last set of notes we saw one way to solve a system of linear equations, by performing row operations. Now we see a second way to solve the system of equations, by finding the inverse of a matrix. At this point, we pause to list some properties of matrix inversion.

1. Inverses are unique. That is, if $AB = I$ and $AC = I$ then $B = C$. Again, we write the common inverse as A^{-1}.

2. Let B be an $n \times n$ matrix such that $AB = I$. Then $BA = I$ as well, and so $B = A^{-1}$. In particular inverses for multiplication on the left are the same as inverses for multiplication on the right.

3. The $n \times n$ matrix A is invertible if and only if the matrix equation $Ax = b$ has a unique solution for every choice of right hand side b. Moreover, the solution is $x = A^{-1}b$.

4. If A and B are both invertible then $(AB)^{-1} = B^{-1}A^{-1}$.

Alright, so how do you find the inverse of a matrix? One method is to do row reduction. We see this as follows. Start with

$$Ax = b,$$

and start to multiply on the left by the matrices we discussed last time which give you elementary row operations. After some number of steps, let's say k steps, we have

$$\begin{bmatrix} \hat{A}_{11} & 0 & \cdots & 0 \\ 0 & \hat{A}_{22} & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & 0 & \hat{A}_{nn} \end{bmatrix} = E_k E_{k-1} \cdots E_2 E_1 Ax = E_k E_{k-1} \cdots E_1 b.$$

Next we multiply by the diagonal matrix (that is, matrix whose only nonzero entries are on the diagonal from top left to bottom right) with $1/A_{jj}$ down the diagonal. Then we have

$$Ix = \begin{bmatrix} \frac{1}{\hat{A}_{11}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\hat{A}_{22}} & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & 0 & \frac{1}{\hat{A}_{nn}} \end{bmatrix} E_k E_{k-1} \cdots E_1 x = A^{-1}b.$$

This equation above has to be true regardless of which b we choose, so we must have

$$A^{-1} = \begin{bmatrix} \frac{1}{\hat{A}_{11}} & 0 & \cdots & 0 \\ 0 & \frac{1}{\hat{A}_{22}} & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & 0 & \frac{1}{\hat{A}_{nn}} \end{bmatrix} E_k E_{k-1} \cdots E_1.$$

Now let's see an example with a 3×3 matrix. We start with

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 2 & 0 & 2 \end{bmatrix}.$$

We'll just apply successive row operations to A and the identity matrix simultaneously. Each time we apply a row operation, it's the same as multiplying on the left by one of those row operation matrixes. So by applying the row operatorings to I as well, e're keeping track of the product $E_k E_{k-1} \cdots E_1$. **Remember that you always have to apply the same row operations t both matriecs, in the same order.** We write the two matrices side by side so we can see what we're doing:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 2 & 0 & 2 \end{bmatrix}, \qquad I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Start by subtracting three times the first row from the second row, and using the result to replace the second row, so that we have

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & -4 & -8 \\ 2 & 0 & 2 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Next we replace the third row with -2 times the first row plus the third row, to get

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & -4 & -8 \\ 0 & -4 & -4 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}.$$

Next we replace the third row with itself minus the second row, which gives us

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & -4 & -8 \\ 0 & 0 & 4 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix}.$$

Next we divide the second row by -4 and the first row by 4, so that we're left with

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 3/4 & -1/4 & 0 \\ 1/4 & -1/4 & 1/4 \end{bmatrix}.$$

Now, on the left hand side, we have the correct last row, so we leave it alone. To get the correct third row we subtract twice the last row from the second row to get

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 1/4 & 1/4 & -1/2 \\ 1/4 & -1/4 & 1/4 \end{bmatrix}.$$

Now we have the correct second and third rows, so we only need to worry about fixing up the first row. First we subtract 2 times the second row from the first row, which gives us

$$\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1/2 & -1/2 & 1 \\ 1/4 & 1/4 & -1/2 \\ 1/4 & -1/4 & 1/4 \end{bmatrix}.$$

Finally, we subtract 3 times the last row from the first row, so that we're left with

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} -1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & -1/2 \\ 1/4 & -1/4 & 1/4 \end{bmatrix}.$$

We conclude that

$$A^{-1} = \frac{1}{4} \begin{bmatrix} -1 & 1 & 1 \\ 1 & 1 & -2 \\ 1 & -1 & 1 \end{bmatrix},$$

which we can verify by multiplying:

$$AA^{-1} = I.$$

(You might want to check this last part yourself!)

Some people would write the computations above as a series of 3×6 matrices, rather than a series of side-by-side 3×3 matrices. This is a good way to remember that you have to apply row operations to **both** matrices, not just one.

You can use row operations to find the inverse of any invertible square matrix, just as we did above. If A is not invertible, then somewhere in the process above you'd encounter a row of 0's. If this happens, you can stop and say that A is not invertible.

The computation we just did was fairly involved, and a lot of work. Later on we'll learn one or two other techniques for computing the inverse of a matrix, but they're also complicated. It turns out that inverting a matrix is usually a lengthy computation, and one we don't want to do. Instead, we look for ways to get the same information without computing the inverse of the matrix. For instance, we might only want to know whether or not a matrix is invertible, and there's a quick computation we can do to determine this, which we'll learn later.

It turns out that there's a nice formula for the inverse of a 2×2 matrix, which is worth remembering. If A is a 2×2 matrix, write it as

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Then, so long as $\det(A) = ad - bc \neq 0$, we have

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Again, we can verify this formula by direct computation, checking that

$$AA^{-1} = A^{-1}A = I.$$

The quantity $ad - bc$ is the determinant $\det(A)$ of the matrix A. Later on we'll learn how to compute the determinant of any square matrix. It turns out that in general A is invertible if and only if $\det(A) \neq 0$.