AMES Class Notes – Week 10, Monday: Linear Algebra, cont

Andie Creel

4$^{\text{th}}$ November, 2024

1 Inverting a matrix

We will see how to invert at 2×2 matrix and a 3×3 matrix. In both examples, we will see that the equation equals

$$A^{-1} = \frac{1}{|A|} adj(A) \tag{1}$$

where $|A|$ is the **determinant** and $adj(A)$ is the **adjoint matrix**.

The reason we need to be able to invert matrix is because we cannot divide by a matrix. So this is an important operation we need for solving systems of linear equations.

1.1 Inverting a 2×2 matrix

Remember rules with 1. They have equivalents with the identity matrix.

$$\frac{1}{a}a = 1 \tag{2}$$

$$a^{-1}a = 1 \tag{3}$$

$$\underset{1\times 1}{a}{}^{-1}\underset{1\times 1}{a} = \underset{1\times 1}{I} = 1 \tag{4}$$

$$\underset{n\times n}{A}{}^{-1}\underset{n\times n}{A} = \underset{n\times n}{I} \tag{5}$$

Consider the matrix A and a matrix B

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \tag{6}$$

$$A^{-1} = B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \tag{7}$$

What is $BA = A^{-1}A = I$?

$$\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} \\ ... & ... \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{8}$$

The first element in the first row of AB implies because we're noticing the the element in position $[1,1]$ of the AB matrix is equal to the element in position $[1,1]$ of the identity matrix. Next, we're just doing some algebra to solve for b_{11},

$$b_{11}a_{11} + b_{12}a_{21} = 1 \tag{9}$$

$$\implies b_{11} = \frac{1 - b_{12}a_{21}}{a_{11}} \tag{10}$$

The second element in the first row (position $[1,2]$) of AB implies

$$b_{11}a_{12} + b_{12}a_{22} = 0 \tag{11}$$

we can plug in what we got for b_{11} earlier in equation 10 to get

$$\frac{1 - b_{12}a_{21}}{a_{11}}a_{12} + b_{12}a_{22} = 0 \tag{12}$$

$$\tag{13}$$

which we can solve algebraically to for b_{12} to get

$$\implies b_{12} = \frac{a_{22}}{a_{11}a_{22} - a_{12}a_{21}} \tag{14}$$

which we can plug into eqn 10 to get rid of the b_{12} and have a solution for b_{11} that's only a function of elements from the A matrix.

$$b_{11} = \frac{1 - \frac{a_{22}}{a_{11}a_{22} - a_{12}a_{21}}a_{21}}{a_{11}} \tag{15}$$

$$\tag{16}$$

which simplifies to

$$b_{11} = \frac{a_{22}}{a_{11}a_{22} - a_{12}a_{21}}. \tag{17}$$

The solutions to the other elements are (just giving them so we don't have to solve for them algebraically over and over again):

$$b_{12} = \frac{-a_{12}}{a_{11}a_{22} - a_{12}a_{21}} \tag{18}$$

$$b_{21} = \frac{-a_{21}}{a_{11}a_{22} - a_{12}a_{21}} \tag{19}$$

$$b_{22} = \frac{a_{11}}{a_{11}a_{22} - a_{12}a_{21}} \tag{20}$$

Remember that we're trying to solve for matrix B right now, which is the inverse of A. Notice that there is a pattern here where each element of the inverse matrix is multiplied by the fraction $\frac{1}{a_{11}a_{22} - a_{12}a_{21}}$. So we can write the solution to our inverse matrix as

$$B = A^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \tag{21}$$

The the first term is the inverse of the **determinant**, where the determinant is

$$|A| = a_{11}a_{22} - a_{12}a_{21}$$

and the second term is the **adjoint matrix**

$$adj(A) = \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.$$

So our solution in equation 21 is in the form we saw in equation 1.

Notice that if the determinant is 0, which will happen if you have a row of zeros *aka* if your matrix is not full rank, then your inverse matrix will not be defined because you'll have $\frac{1}{0}$. If this is the case, the inverse of the matrix does not exist.

1.2   Determinants

The *determinant* will *determine* if you can invert a matrix.

Previously in class we would solve a system of equations for where two lines crossed. If the lines were different (and not parallel), then they would cross somewhere and we could find a point where the two lines are equal at that point. However, if you had two lines that were equal, aka they are the same line, then there is no single point where the two are equal. A **determinant** will let you know if your system of equations has a solution. If the determinant is 0, then two lines are the same or they are parallel (not unique). The reason we're talking about rows and columns of matrices as if they are lines is because the rows and columns of a system of matrices gives us a system of equations.

This issue shows up if some of your variables are perfectly co-linear aka parallel. If you have variables that are co-linear, you won't be able to do a multi-variate regression and the determinant will be zero (unless you have tons and tons of data).

1.3   Inverting 3×3 matrix

Consider the matrix A.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \tag{22}$$

The goal is to solve for the inverse of A, A^{-1}. Remember equation 1. We need to get the adjoint matrix and the determinant.

Step 1: Take the all **minors** (9 total), M_{ij}:

$$M_{11} = \left| \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} \right| = a_{22}a_{33} - a_{23}a_{32} \tag{23}$$

$$M_{23} = \left| \begin{bmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{bmatrix} \right| = a_{11}a_{32} - a_{12}a_{31} \tag{24}$$

and so on....

You could make a (3×3) matrix of minors, where each minor you take goes in the (i, j) position (note that minors are scalars). But, before we we make this, we need to edit the signs of each of these elements. To do that, we use a co-factor equation.

Step 2: Get the co-factors and construct the **co-factor matrix** (3×3). Each element of the co-factor matrix is a function of the minors, which is why we solved for them first:

$$C_{ij} = (-1)^{i+j} * M_{ij} \tag{25}$$

$$= \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \tag{26}$$

Step 3: Then transpose the co-factor matrix (the columns become rows). This is the **adjoint matrix** of A when A is 3×3

$$adj(A) = C^T = \begin{bmatrix} c_{11} & c_{21} & c_{31} \\ c_{12} & c_{22} & c_{32} \\ c_{13} & c_{23} & c_{33} \end{bmatrix} \tag{27}$$

Step 4: Find the **determinant of A**, which can conveniently be written as

$$|A| = a_{11}c_{11} + a_{12}c_{12} + a_{13}c_{13} \tag{28}$$
$$= a_{12}c_{12} + a_{22}c_{22} + a_{32}c_{32} \tag{29}$$
$$= ... \tag{30}$$

so you have a bunch of ways to solve for the determinant. You can choose which of these equations you want to use. Choose any row or column, try to find one that has a lot of 0s to make it easier to calculate.

Step 5: Finally, you can solve for **inverse of A**, A^{-1}, using all these components,

$$A^{-1} = \frac{1}{|A|}adj(A) \tag{31}$$

and equation 1

2 Cramer's rule for solving $Ax = b$ without taking an inverse

Cramer's Rule is a mathematical theorem and a method for solving a system of linear equations with as many equations as unknowns (a square system). It provides an explicit formula for each of the unknown variables in terms of determinants. This is nice because determinants are way easier to find than adjoin matrices. Cramer's Rule is named after the Swiss mathematician Gabriel Cramer.

Here's how Cramer's Rule works for a system of linear equations with n variables $(x_1, x_2, \ldots, x_n)$ and n equations:

1. Write the system of equations in matrix form as $Ax = b$, where A is the coefficient matrix, x is the vector of unknowns, and b is the vector of constants on the right-hand side.

2. Compute the determinant of the coefficient matrix A, denoted as $|A|$.

3. Create n new matrices by replacing each column of A with the column vector b. These matrices are denoted as $A_1, A_2, \ldots, A_n$.

4. Compute the determinants of these matrices: $|A_1|, |A_2|, \ldots, |A_n|$.

5. The solution to the system of equations is given by $x_i = \frac{|A_i|}{|A|}$ for each variable x_i.

It's important to note that Cramer's Rule only works when the determinant of the coefficient matrix A ($|A|$) is non-zero, meaning the system of equations has a unique solution. If $|A|$ is zero, it means that the system of equations is either inconsistent (no solution) or has infinitely many solutions, and Cramer's Rule cannot be applied.

Cramer's Rule can be computationally intensive for larger systems because it involves calculating multiple determinants. In practice, other methods like Gaussian elimination or matrix factorization techniques are often preferred for solving systems of linear equations.