Andrew Shao

Linear Algebra I

Unit 5: Inverses

Matrix Inverses:

A *Matrix Inverse* is a matrix where when taking AB, you yield I. A *Left Inverse* is one where if you take matrix C, shaped mxn, and evaluate CA, you get I_n. A *Right Inverse* is one where if you take matrix B, shaped mxn, and evaluate AB, you get I_m

How can we compute a left or right inverse?

To compute a left inverse, we need to break down the definition of matrix multiplication and solve for the identity matrix. We know that:

$AB = A[b_1, ..., b_n]$

$AB = [Ab_1, Ab_2, ..., Ab_n]$

Thus, we want to solve for when $[Ab_1, ..., Ab_n] = [e_1, ..., e_n]$. Note that this is just solving n linear equations, and thus we can simply use row-reduction and RREF to solve for a left inverse.

Also note that for this system to be consistent, only square matrices can have both a left and right inverse, since if A_{mxn}, $where\ m > n$, then A cannot have a right inverse, and vice-versa.

Theorem 5.1.3: A matrix is invertible when it has an inverse. A square matrix, while having a left and right inverse, has only one unique solution.

Theorem 5.1.3 can be proven very easily:

$I = BA, where\ B = A^{-1}$ and:

$AC = I, where\ C = A^{-1}$

$B = BI = B(AC) = (BA)C = IC = C$

Sample Problem: Solve for the inverse of the matrix: $\begin{bmatrix} 1 & 3 & 2 \\ -2 & 2 & -2 \\ 3 & -1 & 3 \end{bmatrix}$

We simply structure the problem as a row-reduction problem, where the system is [A|I], and the resulting inverse is $[I \mid A^{-1}]$.

$$\left[\begin{array}{ccc|ccc} 1 & 3 & 2 & 1 & 0 & 0 \\ -2 & 2 & -2 & 0 & 1 & 0 \\ 3 & -1 & 3 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 11/4 & 5/2 \\ 0 & 1 & 0 & 0 & 3/4 & 1/2 \\ 0 & 0 & 1 & 1 & -5/2 & -2 \end{array}\right]$$

Solving, we find the inverse to be $\begin{bmatrix} -1 & 11/4 & 5/2 \\ 0 & 3/4 & 1/2 \\ 1 & -5/2 & -2 \end{bmatrix}$. Computing an inverse can be quite tedious, but it is very easy to check if your answer is correct. $AB = I = BA$, should be valid when solving for an inverse in a square matrix.

Also, keep in mind that an inverse is only possible if the RREF turns out to be I, and doesn't result in infinite solutions or impossible solutions.

This process can be rather tedious, and thus we need an easier way to compute the inverse. A specific formula to quickly compute the inverse of a 2x2 matrix is the following:

if $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then:

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

Note that if ad - bc is 0, then the matrix is not invertible.

Left and Right Inverses:

The previous inverse matrices are considered to be *2-sided-inverses* since they can be applied to either the left or the right side of a given matrix and both will result in the identity matrix. However, 2-sided inverses only exist for square matrices. Sometimes, we want an inverse matrix that can still work on rectangular matrices.

Left Inverse: A left inverse is defined as: $(A^T A)^{-1} A^T$, as it can be applied to only the left side of the matrix A. The reasoning is because any matrix multiplied by its inverse will yield a square matrix. Thus, if we can perform this and get a square matrix using A and A^T, we can simply compute an inverse! Keep in mind that the order does matter in matrix multiplication, which is why this only works on the left.

Right Inverse: A right inverse is pretty much the same, but now we consider $A^T (AA^T)^{-1}$ as the inverse.

Properties of Invertible Matrices:

Theorem 5.1.6: If k is some constant and A and B are invertible matrices.

1) $(kA)^{-1} = \frac{1}{k} A^{-1}$
2) $(AB)^{-1} = B^{-1} A^{-1}$
3) $(A^T)^{-1} = (A^{-1})^T$

Proof:

1) $(kA)(1/k(A^{-1})) = (k/k)(A)(A^{-1}) = I$
2) $(AB)(B^{-1} A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = I$

3) $(A^T)(A^{-1})^T = (AA^{-1})^T = I$

Theorem 5.1.7: The Invertible Matrix Theorem

There are 8 properties that are all equivalent when considering invertible matrices. This means that if one of these properties are true, they must all be true, and if one is false, they all are. Here are they:

1) A is invertible
2) RREF of A = I
3) rankA = n
4) Ax = b is consistent with a unique solution every single time.
5) Null(A) = 0
6) Col(A) forms a basis for $\mathbb{R}^n$
7) Row(A) forms a basis for $\mathbb{R}^n$
8) A^T is invertible

Proof:

1) Iff RREF of A = 1(part 2), then we can find an inverse.
2) An invertible matrix exists only when the system is consistent with a unique solution, this only happens when RREF of A = I
3) For rankA = n, the RREF of A has to be consistent with a unique solution. This is true if RREF of A = I
4) If there are no free variables since RREF of A = I, then it is only possible that a unique solution exists.
5) Since there exists only unique solutions proved by 4, then the only solution for Ax = 0 is the trivial one
6) Since there are only unique solutions, then the rows and columns are linearly independent and thus have to span $\mathbb{R}^n$ if there are n of them and they are all linearly independent.
7) Same proof for 6.
8) If both rows and columns are linearly independent, then the transpose of the matrix will still have the above properties.

Solving Linear Systems with Inverses:

Now knowing how invertible matrices work, we can now use them to solve linear equations with ease. Iff A is invertible and Ax = B, we can solve the equation as follows:

$Ax = B$

$A^{-1}Ax = A^{-1}B$

$x = A^{-1}B$

Ex. Solve the following system of equations:

$2x_1 + 4x_2 = 3$

$-x_1 - 5x_2 = 5$

We first construct the coefficient Matrix, but non-augmented as we will solve for the inverse. $\begin{bmatrix} 2 & 4 \\ -1 & -5 \end{bmatrix}$. Using the formula for a 2x2 matrix inverse, we compute that the a = 2, b = 4, c = -1, and d = -5. Thus:

$$A^{-1} = \frac{1}{2(-5)-4(-1)}\begin{bmatrix} -5 & -(4) \\ -(-1) & 2 \end{bmatrix}$$

Or:

$$A^{-1} = \frac{-1}{6}\begin{bmatrix} -5 & -4 \\ 1 & 2 \end{bmatrix}$$

We plug this in and solve for $A^{-1}b$, yielding: $\begin{bmatrix} 35/6 \\ -13/6 \end{bmatrix}$, which is the solution to the linear equation.

Elementary Matrices:

Elementary Matrix: a matrix that stores an elementary row operation. Matrix multiplication using this elementary matrix will result in the same effect as performing an elementary row operation.

We have already seen elementary matrices when finding inverses, as matrix A^{-1}has the ability to convert the original matrix A to the identity matrix. Thus, the row reduction operations that we did to solve for A^{-1}must be somehow encoded into this matrix.

Let E represent the elementary matrix that encodes how to perform a given elementary row operation, and EA as the matrix A after having the operation applied to it. Since elementary row operations cannot change the rank of a matrix, we immediately get the following:

rank(EA) = rank(A)

Note that we can get an elementary matrix by simply applying the given operation onto the identity matrix, and we can compute the inverse of a given elementary operation to “undo” its effects.

Ex. *Find a 2x2 elementary matrix corresponding to* $4R_2$.

We simply multiply the second row of the 2x2 identity matrix to get it’s elementary matrix, being:

$$\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$$

Matrix Decomposition: Utilizing Elementary matrices, we can “decompose” entire row reductions to make them more understandable(called “factoring” a matrix).

This is because we know that any elementary row reduction operation can be represented as a matrix multiplication with an elementary matrix, and thus we can represent any matrix as a sequence of elementary matrix inverses from the reduced form. In other words:

$A = E^{-1}_1\, matmul \ldots\ matmul\, E^{-1}_k\, matmul\, R$, since all elementary matrices are invertible.

Note: You can only use *elementary* row operations, combination row operations will lead to incorrect answers.

Also notice that since R is the identity matrix if matrix A is invertible, which also tells us that $A^{-1} = E_k\, matmul \ldots\ matmul\, E_1$, and $A = E_1\, matmul \ldots\ matmul\, E_k$

Determinants:

The *determinant* is an indicator of whether or not a given matrix is invertible or not. It is denoted either with a vertical bar matrix or as det(matrix). Ex:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

Computing the determinant of larger and larger matrices quickly becomes tedious, but it can be done recursively, with the base case being when the matrix is 1x1(The determinant is just the value inside).

To understand the recursive calculation of the determinant, you can use what is called cofactors, which simply determine whether to use a positive or negative coefficient in front(You will see what this means later).

$$C_{ij} = (-1)^{i+j} det(A(i, j))$$

What this calculation represents is to use the indices of the matrix(row and column) to find out whether or not the coefficient is positive or negative, and then delete the i'th row and j'th column before calculating the determinant again(smaller problem now).

Sample Problem: Compute the Cofactor $a_{2,3}$ of matrix A: $\begin{bmatrix} 3 & 2 & -1 \\ 1 & 4 & 3 \\ 5 & 1 & 2 \end{bmatrix}$

We first find the value inside of (2, 3), being 3. Then, we eliminate the second row and third column, yielding:

$$\begin{bmatrix} 3 & 2 \\ 5 & 1 \end{bmatrix}$$

We then continue on computing the cofactor:

$$C_{(2,3)} = (-1)^{2+3} \begin{vmatrix} 3 & 2 \\ 5 & 1 \end{vmatrix}$$

We know already how to compute a 2x2 determinant(ad - bc), and do so to get: 3(1) - 2(5) = -7

$$C_{(2,3)} = -1 * -7 = 7$$

Using the cofactors, we can compute a determinant by taking the dot product between the cofactors in the first row with the coefficients in the first row.

Sample Problem: Compute: $\begin{vmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{vmatrix}$

We start by computing the cofactors of the first row. Note we only need to compute the first one, as the first row has many zeros(which is convenient)

$C_{(1,1)} = (-1)^{1+1}\begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix}.$

Recursively computing cofactors again:

$C_{(1,1)} = (-1)^{1+1}\begin{vmatrix} 1 \end{vmatrix}$

$C_{(1,1)} = 1 * 1$

$\begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} = 1 * 1 = 1$

Moving up a recursion layer.

$C_{(1,1)} = 1 * 1 = 1$

$\begin{vmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{vmatrix} = 2 * 1 = 2$

Thus, the determinant is 2.

Cofactor Expansions:

We are not limited to just using the top row to compute the determinant. You can perform a cofactor expansion by using the cofactors of any arbitrary row or column, and all will have the final result. Keep in mind that the coefficients of the cofactor expansion may change depending on the location(due to the (-1) ^ (i + j)), so don't get lazy.

Cofactor expansions are really useful in the sense that we can simply perform a cofactor expansion over the row or column that has the most zeros inside, which can heavily reduce the computation needed.

Calculating the Determinant using Triangular Matrices:

Since we now know what cofactor expansions are, we can utilize triangular matrices to quickly compute the determinant of large matrices.

Upper Triangular Matrices: $u_{ij} = 0$ whenever i > j. This makes the diagonal row full, and everything below it to be 0. We know that the determinant will simply be all of the diagonal entries multiplied. Why? Well, by cofactor expansion, we only ever use the top left entry, which results in it just being multiplied across the diagonal.

Lower Triangular Matrices: $u_{ij} = 0$ when i < j. This makes everything above the diagonal to be 0.

But, we cannot just hope that we always get diagonal entries. We need a way to reduce the matrix down to a triangular matrix. We can perform this with basic row operations. However, these row operations can cause a change in the determinant, so we have to be sure to revert the changes later.

Changes to the Determinant:

1) By multiplying a row of matrix A by a constant C, det(A) = cdet(cA). The determinant will change by the same amount.
2) By adding one row of matrix A to a multiple of another, the determinant doesn't change.
3) By swapping two rows, the determinant becomes negative.
4) Transposing a Matrix doesn't change the determinant.

We know the changes to the determinant for every elementary row operation, and thus with our linear algebra techniques, we can reduce any matrix down to this triangular form. For small matrices, this doesn't have a huge benefit, but for very large ones, the computation without this method becomes impossible.

Furthermore, notice that since the transpose of a matrix has the same determinant as the original, this also means we can perform column operations and the same properties.

Notice the relationship between elementary row operations and these determinant changes, and how det(EA) = det(E) det(A). Why? Well each elementary row matrix is just an elementary row operation on the identity matrix, and thus the determinant will be changed by the same amount.

Theorem 5.3.8: A given matrix A is invertible iff det(A) != 0. This is because for a determinant to be 0, the reduced form of the matrix using elementary row operations will not be the identity matrix.

Theorem 5.3.9: Det(AB) = Det(A) * Det(B). This is due to how we can reduce both A and B separately down to the identity matrix before multiplying them, yielding a determinant only composed of the elementary matrices used to row reduce A and B. $Det(E_1 E_{...} I\ matmul\ E_1 E_{...} I)$

Theorem 5.3.10: $Det(A^{-1}) = 1 / Det(A)$.This can be easily shown from Theorem 5.3.9 and breaking down the matrix A into it's elementary matrices.

Cofactor Method and Cramer's Rule:

We already know from Theorem 5.3.8 that a matrix is only invertible if it's determinant is not 0. There indeed exists a relationship between the cofactors of a matrix and its inverse. The method of solving for the inverse is the Cofactor method, which states that:

$A^{-1} = \frac{1}{det(A)} adj(A)$. Where adj(A), or the adjugate matrix is the transpose of the cofactor method.

Ex. solve for the inverse of $\begin{bmatrix} 2 & 1 & -1 \\ 1 & 3 & -1 \\ 2 & -1 & 2 \end{bmatrix}$

We first solve for the cofactors of the matrix, and we get:

$$\begin{bmatrix} 5 & -4 & -7 \\ -1 & 6 & 4 \\ 2 & 1 & 5 \end{bmatrix}$$

Using the cofactors, we compute the determinant and get det(A) as 13. This is good because it shows that there actually exists an inverse.

Then, we transpose the cofactor matrix to get the adjugate matrix $\begin{bmatrix} 5 & -1 & 2 \\ -4 & 6 & 1 \\ -7 & 4 & 5 \end{bmatrix}$.

Finally, we multiply the reciprocal of the determinant with the adjugate matrix to get the inverse.

You may ask, why do all this work? Well, by using the adjugate matrix, we can compute for any given index in the matrix its inverse. This is useful when matrices are filled with variables, making it super tough to reduce via RREF.

Similarly, we can create Cramer's rule by using this definition of the inverse.

We know that $A^{-1} = \frac{1}{detA}adj(A)$, and thus in a system with $Ax = b$, we solve for x using $x = A^{-1}b$

Substituting in, we get $x = \frac{1}{detA}adj(A)b$, which means that for any x_i, $x_i = \frac{b_1A_{1i} + ... + b_nA_{ni}}{detA}$

Note that this is exactly the same as substituting the i'th column with the b vector and solving for the determinant, then dividing by det(A). Thus, for any x_i, $x_i = \frac{detA_i}{detA}$.

Once again, this is super slow, but can be helpful when a matrix is full of variables.

Geometric Interpretation of the Determinant:

The determinant also has a geometric interpretation of computing the area of a parallelogram. Note that the area of a parallelogram is equivalent to the absolute value of the determinant of a matrix of it's base vectors(The 2 sides not parallel, we ignore the parallel sides).

This definition also extends to the definition of a parallelepiped, where the determinant of this 3x3 matrix is equivalent to its volume.

Final Unit: Diagonalization