MATH-UA 140 - Linear Algebra

Lecture 6: Inverse matrices

I] Review of matrix operations, and a useful trick

1) Review

In the past lectures, we have covered matrix addition: Two matrices with the same dimensions can be added, and the sum C of the matrices A and B, $C = A + B$, has entries $c_{ij} = a_{ij} + b_{ij}$

We have also seen multiplication by a scalar c: the entries of the matrix cA are ca_{ij}. In other words, all the entries are multiplied by the scalar

Lastly, we also learned how to multiply a matrix with another matrix. The rules were as follows:

If the matrix A has m rows and n columns, then the quantity AB only makes sense if the matrix B has n rows.

If B has p columns, then the matrix $C = AB$ has m rows and p columns. Schematically, we may write (m-by-n)(n-by-p) $\Rightarrow$ (m-by-p)

To multiply two matrices A and B, we take the dot product of each row of A with each column of B:

The entry c_{ij} of $C = AB$ is (row i of A)$\cdot$(column j of B)

Example: $\begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 3 & 3 \\ -2 & -1 & -1 \\ 4 & 2 & 2 \end{bmatrix}$

2) Properties of matrix operations

In the following, we assume that the matrices A, B, and C have dimensions such that the operations shown below are allowed. c is a scalar. The following properties hold:

1. $A + B = B + A$ (matrix addition commutes)
2. $c(A+B) = cA + cB = (A+B)c$ (distributive law)
3. $A + (B+C) = (A+B)+C$ (associative law)
4. $C(A+B) = CA + CB$ (distributive law from the left)
5. $(A+B)C = AC + BC$ (distributive law from the right)
6. $A(BC) = (AB)C$
7. $AI = IA = A$, where I is the identity matrix.

Remember that in general, $AB \neq BA$

Matrix multiplication does not commute

3) Matrix powers

Let A be a square matrix. Powers of the matrix A can be defined in the same way as one defines integer powers of scalars:

$$A^2 = AA, \quad A^3 = AAA, \quad A^4 = AAAA, \ldots, \quad A^p = \underbrace{AA\ldots AA}_{p \text{ factors}}$$

If p and q are integers, $(A^p)(A^q) = A^{p+q}$ and $(A^p)^q = A^{pq}$ just like the case of powers of scalars.

For the special case $p=0$, we have $A^0 = I$, where I is the identity matrix with the same dimensions as A.

Lastly, we will soon learn about A^{-1}, the matrix inverse of A.

4) Block matrices and block operations

It will often be useful to cut matrices into blocks and operate on the matrices block by block. What we mean by blocks are smaller matrices within a larger block.

Example:

$$\left[\begin{array}{cc|cc} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ \hline 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{array}\right]$$

This 4-by-4 matrix can be seen as being made of 4 identical 2-by-2 matrices which are the identity I_2

* Blocks and matrix addition

Since matrix addition is defined by the addition of the entries with the same row and column numbers, it is clear that matrix addition can be calculated a block at a time.

Example:
$$\left[\begin{array}{cc|cc} 2 & 1 & 3 & 4 \\ -1 & -1 & 2 & 2 \\ \hline \pi & \sqrt{2} & 1 & e \\ 0 & 1 & -3 & -4 \end{array}\right] + \left[\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ \hline 2 & -2 & 3 & 7 \\ -5 & 4 & 2 & 1 \end{array}\right]$$

$$= \left[\begin{array}{cc|cc} 3 & 1 & 3 & 5 \\ -1 & 0 & 3 & 3 \\ \hline 2+\pi & \sqrt{2}-2 & 4 & e+7 \\ -5 & 5 & -1 & -3 \end{array}\right]$$

* Blocks and matrix multiplication

If we cut two matrices A and B and if the cuts between columns of A match the cuts between rows of B, then AB can be evaluated by block multiplication.

Consider the block matrices

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

where $A_{11}, B_{11}, A_{12}, B_{12}, A_{21}, B_{21}, A_{22}, B_{22}$ are matrices with dimensions that are compatible with the block matrix products. We can compute the product of A with B by blocks:

$$AB = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}$$

Example

$$\left[\begin{array}{cc|cc} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ \hline 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{array}\right] \left[\begin{array}{cc|cc} 1 & 2 & 1 & 2 \\ 3 & 4 & 3 & 4 \\ \hline 1 & 2 & 1 & 2 \\ 3 & 4 & 3 & 4 \end{array}\right] = \begin{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} & \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \\ \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} & \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 4 & 2 & 4 \\ 6 & 8 & 6 & 8 \\ 2 & 4 & 2 & 4 \\ 6 & 8 & 6 & 8 \end{bmatrix}$$

The example above illustrates a common motivation for splitting a matrix into blocks: it is sometimes easier to visualize how given blocks act than the whole matrix. It is certainly the case here, since we know very well how the identity blocks $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ act.

II] Inverse matrices

For any scalar $c \neq 0$, c^{-1} is called its inverse. It is such that $cc^{-1} = c^{-1}c = 1$. The purpose of this section is to learn about the idea of inverses for matrices. As we will see, the situation is slightly more complex for matrices than it is for scalars, and therefore more interesting!

1) Definition and immediate properties

A square matrix A is invertible if there exists a matrix A^{-1}, called "A inverse", such that

$$A^{-1}A = I \quad \text{and} \quad AA^{-1} = I$$

- Note 1: Not all matrices have inverses

In this course, we will learn several methods to find out if a matrix is invertible or not. Many methods will not even require calculating A^{-1}

- Note 2: An n-by-n matrix has an inverse if and only if elimination produces n pivots (i.e. nonzero pivots) (Row exchanges are allowed as necessary)

If A is invertible, the one and only solution $\vec{x}$ to the linear system $A\vec{x} = \vec{b}$ is $\vec{x} = A^{-1}\vec{b}$

To see this, multiply $A\vec{x} = \vec{b}$ by A^{-1} on both sides:

$$A^{-1} A \vec{x} = A^{-1} \vec{b} \iff \vec{x} = A^{-1} \vec{b}$$

Note 3: Suppose there is a nonzero vector $\vec{x}$ such that $A\vec{x} = \vec{0}$. Then A does not have an inverse

If A is invertible, then the linear system $A\vec{x} = \vec{0}$ can only have the zero solution $\vec{x} = A^{-1}\vec{0} = \vec{0}$

Example: $\begin{bmatrix} 4 & 4 \\ 1 & 1 \end{bmatrix}$ is not invertible because the linear system

$\begin{bmatrix} 4 & 4 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ admits the solution $(x_1, x_2) = (1, -1)$

(and many other solutions)

Note 4: If A is invertible, its inverse A^{-1} is unique.

Proof: Let us assume that there exist two matrices B and C such that $BA = I = AB$ and $CA = I = AC$

Now, for any three matrices A, B, and C, we can always write the obvious equality

$$BAC = BAC$$

Using property 6. given at the beginning of this lecture, we rewrite this as

$$B(AC) = (BA)C \iff BI = IC \iff B = C \quad \blacksquare$$

QUESTION: Consider the elimination matrix

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}$$

By interpreting E in terms of operations on rows of a system of equations, find E^{-1}. Verify that you obtained the correct answer by direct computation of $E^{-1}E$.

2) Inverse of the product of matrices

Theorem: If A and B are two square matrices with the same dimensions which are invertible, then AB is also invertible. Furthermore, the inverse of AB is

$$(AB)^{-1} = B^{-1}A^{-1}$$

Note the reversed order.

Proof: $(AB)B^{-1}A^{-1} = A\underbrace{(BB^{-1})}_{I}A^{-1} = AA^{-1} = I$

Hence $B^{-1}A^{-1}$ is indeed the inverse of AB.

QUESTION: In lecture 5, we saw how sequentially applying

$E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $E_{31} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{bmatrix}$

to the system was equivalent to the first two steps of Gaussian elimination. What is the combined matrix which is equivalent to these first two steps? What is its inverse?

- Inverse of a product of more than two matrices

The rule of reversing the order of multiplication for the inverse applies to the product of any number ≥ 2 of matrices. For example, for three matrices A, B, and C,

$$\underline{(ABC)^{-1} = C^{-1} B^{-1} A^{-1}}$$

3) Calculating the matrix inverse by Gauss-Jordan elimination

In the previous two lectures, we have learned how to use Gaussian elimination to solve a system of the form $A\vec{x} = \vec{b}$. As we saw in this lecture, the solution $\vec{x}$, when it exists and is unique, is such that $\vec{x} = A^{-1}\vec{b}$. A beauty of Gaussian elimination is that we did not need to compute A^{-1} to find $\vec{x}$.

Remarkably, however, Gaussian elimination can be used to calculate A^{-1}, and we now show how it works with the example of a 3-by-3 matrix.

Consider the vectors $\vec{i} = (1,0,0)$, $\vec{j} = (0,1,0)$, and $\vec{k} = (0,0,1)$, and a 3-by-3 matrix A. We are looking for A^{-1} such that

$$A A^{-1} = I = [\vec{i} \ \vec{j} \ \vec{k}]$$

Calling $\vec{x}_1, \vec{x}_2$, and $\vec{x}_3$ the unknown columns of A^{-1}, we thus want to solve

$$A \left[\vec{x}_1 \ \vec{x}_2 \ \vec{x}_3 \right] = \left[\vec{i} \ \vec{j} \ \vec{k} \right]$$

In other words, we want to solve 3 systems of equations: $A\vec{x}_1 = \vec{i}$, $A\vec{x}_2 = \vec{j}$, $A\vec{x}_3 = \vec{k}$

Gaussian elimination certainly tells us how to do that! Let us be more specific with an example, namely the matrix

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

we have encountered on multiple occasions by now.
The idea to solve the 3 systems at once is to form the augmented matrix $[A \ \vec{i} \ \vec{j} \ \vec{k}]$ and then apply the elimination steps:

$$\begin{bmatrix} 2 & 1 & 3 & 1 & 0 & 0 \\ 1 & 2 & -1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 2 & 1 & 3 & 1 & 0 & 0 \\ 0 & \frac{3}{2} & -\frac{5}{2} & -\frac{1}{2} & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{bmatrix} \qquad (\text{Row } 2 - \tfrac{1}{2} \text{ Row } 1)$$

$$\rightarrow \left[\begin{array}{cccccc} 2 & 1 & 3 & 1 & 0 & 0 \\ 0 & \frac{3}{2} & -\frac{5}{2} & -\frac{1}{2} & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 & 1 \end{array}\right] \qquad \left(\text{Row } 3 - \frac{1}{2} \text{ Row } 1\right)$$

$$\rightarrow \left[\begin{array}{cccccc} 2 & 1 & 3 & 1 & 0 & 0 \\ 0 & \frac{3}{2} & -\frac{5}{2} & -\frac{1}{2} & 1 & 0 \\ 0 & 0 & \frac{4}{3} & \frac{1}{3} & \frac{1}{3} & 1 \end{array}\right] \qquad \left(\text{Row } 3 - \frac{1}{3} \text{ Row } 2\right)$$

At that point, we are done: back substitution using the 4^{th} column would give us $\vec{x_1}$, back substitution using the 5^{th} column would give $\vec{x_2}$, and back substitution using the 6^{th} column would give $\vec{x_3}$. This is Gauss' idea to get A^{-1}. However, Jordan came up with the idea of continuing with elimination so as to make A^{-1} appear in the augmented matrix. The idea is to use elimination to turn the nonzeros entries above the diagonal in the first three columns into zero entries.

$$\rightarrow \left[\begin{array}{cccccc} 2 & 1 & 3 & 1 & 0 & 0 \\ 0 & \frac{3}{2} & 0 & -\frac{9}{8} & \frac{3}{8} & \frac{15}{8} \\ 0 & 0 & \frac{4}{3} & \frac{1}{3} & \frac{1}{3} & 1 \end{array}\right] \qquad \left(\text{row } 2 + \frac{15}{8} \text{ row } 3\right)$$

$$\rightarrow \left[\begin{array}{cccccc} 2 & 1 & 0 & \frac{7}{4} & \frac{3}{4} & -\frac{9}{4} \\ 0 & \frac{3}{2} & 0 & -\frac{9}{8} & \frac{3}{8} & \frac{15}{8} \\ 0 & 0 & \frac{4}{3} & \frac{1}{3} & \frac{1}{3} & 1 \end{array}\right] \qquad \left(\text{row } 1 - \frac{9}{4} \text{ row } 3\right)$$

$$\rightarrow \left[\begin{array}{cccccc} 2 & 0 & 0 & \frac{5}{2} & \frac{1}{2} & -\frac{7}{2} \\ 0 & \frac{3}{2} & 0 & -\frac{9}{8} & \frac{3}{8} & \frac{15}{8} \\ 0 & 0 & \frac{4}{3} & -\frac{1}{3} & -\frac{1}{3} & 1 \end{array}\right] \qquad \left(\text{row } 1 - \frac{2}{3} \text{ row } 2\right)$$

The last step of Gauss-Jordan elimination is to divide each row by its pivot

$$\rightarrow \begin{bmatrix} 1 & 0 & 0 & \frac{5}{4} & \frac{1}{4} & -\frac{7}{4} \\ 0 & 1 & 0 & -\frac{3}{4} & \frac{1}{4} & \frac{5}{4} \\ 0 & 0 & 1 & -\frac{1}{4} & -\frac{1}{4} & \frac{3}{4} \end{bmatrix}$$

$$= \begin{bmatrix} \vec{i} & \vec{j} & \vec{k} & \vec{x}_1 & \vec{x}_2 & \vec{x}_3 \end{bmatrix} = \begin{bmatrix} I & A^{-1} \end{bmatrix}$$

So $A^{-1} = \begin{bmatrix} \frac{5}{4} & \frac{1}{4} & -\frac{7}{4} \\ -\frac{3}{4} & \frac{1}{4} & \frac{5}{4} \\ -\frac{1}{4} & -\frac{1}{4} & \frac{3}{4} \end{bmatrix}$. You can convince yourself that this is the correct result by multiplying A with A^{-1} and verify that you find I.

Through the elimination algorithm, the Gauss-Jordan process essentially does the following:

<u>Multiply $[A \;\; I]$ by A^{-1} in order to get $[I \;\; A^{-1}]$</u>

QUESTION: Use Gauss-Jordan elimination to compute the inverse A^{-1} of

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

The Gauss-Jordan process always works if A has n pivots. So this proves that if A has n pivots, it is invertible.
One can ask oneself the reverse question: is it true that if A does not have n pivots, then A is not invertible? The answer is yes, and the Gauss-Jordan process gives a hint for why this is true, because when a pivot is zero, we cannot proceed with elimination, and cannot divide the rows by the pivots. We will soon provide a crisper, more rigorous proof. For the moment, it is sufficient to remember the important statement:

<u>A square matrix A is invertible if and only if it has n pivots</u>