Lecture 03

Paul Scemama

OCW: *This lecture looks at matrix multiplication from five different points of view. We then learn how to find the inverse of a matrix using elimination, and why the Gauss-Jordan method works.*

1 Outline

- Matrix multiplication (5 ways)
- Inverse of A, AB, and A^T
- Gauss-Jordan / Find A^{-1}

2 Matrix multiplication: row $\times$ column

$$\underbrace{\begin{bmatrix} - & - & - & - \\ - & - & - & - \\ A_{31} & A_{32} & A_{33} & A_{34} \\ - & - & - & - \end{bmatrix}}_{A} \overbrace{\begin{bmatrix} - & - & - & B_{14} \\ - & - & - & B_{24} \\ - & - & - & B_{34} \\ - & - & - & B_{44} \end{bmatrix}}^{B} = \underbrace{\begin{bmatrix} - & - & - & - \\ - & - & - & - \\ - & - & - & C_{34} \\ - & - & - & - \end{bmatrix}}_{C}$$

$$\begin{aligned} C_{34} &= (\text{row 3 of } A) \cdot (\text{column 4 of } B) \\ &= \underbrace{a_{31}b_{14}}_{k=1} + \underbrace{a_{32}b_{24}}_{k=2} \cdots = \sum_{k=1}^{n} a_{3k}b_{k4} \end{aligned}$$

Q: When are we allowed to multiply matrices?

$$A_{m\times n} \cdot B_{n\times p} = C_{m\times p}$$

So the number of columns in A have to match the number of rows in B.

3 Matrix multiplication: whole columns

We've just seen a way to multiply to matrices. By multiplying the rows of the first with the columns of the second to fill the *output* matrix.

Now we will look at matrix multiplication with whole columns in mind. For instance, let's suppose we have an $m \times n$ matrix A and an $n \times p$ matrix B that we want to multiply together:

$$A_{m\times n} \cdot B_{n\times p} = C_{m\times p}$$

But instead look at A times the *columns* of B one at a time.

$$\underbrace{\begin{bmatrix} & & & \\ & & & \\ & & & \\ & & & \end{bmatrix}}_{A} \cdot \overbrace{\begin{bmatrix} | & & & \\ | & & & \\ \underbrace{|}_{Col_1} & & & \end{bmatrix}}^{B} = \begin{bmatrix} | & & & \\ | & & & \\ | & & & \\ \underbrace{|}_{A(Col_1)} & & & \end{bmatrix}$$

The matrix A times the first column of B is a column itself; it is a linear combination of the columns of A. The scaling of each column of A is determined by the elements in the first column of B.

And we can do this for each column in B.

$$\begin{gathered} (\text{Matrix } \cdot Col_1) \text{ is } Col_1 \text{ of answer} \\ (\text{Matrix } \cdot Col_2) \text{ is } Col_2 \text{ of answer} \\ \vdots \end{gathered}$$

To reiterate, we can think of AB as multiplying A to each column of B and getting the columns of the answer. Columns of C are combinations of the columns of A (notice how columns of A have length m as do columns of C).

4 Matrix multiplication: whole rows

As we did with columns, we can also look at whole rows.

$$\underbrace{\begin{bmatrix} & & \\ - & - & - \\ & & \end{bmatrix}}_{A} \cdot \overbrace{\begin{bmatrix} & & \\ & & \\ & & \end{bmatrix}}^{B} = \begin{bmatrix} & & \\ - & - & - \\ & & \end{bmatrix}$$

The rows of C are combinations of the rows of B.

5 Matrix multiplication: sum of (columns × rows)

Consider multiplying a column of A to a row of B.

$$\underbrace{\text{column of } A}_{m\times 1} \times \underbrace{\text{row of } B}_{1\times p} = \underbrace{\text{a big matrix}}_{m\times p}$$

For example, the first column of A and the first row of B may be a and b respectively. And they may be multiplied so that,

$$\underbrace{\begin{bmatrix}2\\3\\4\end{bmatrix}}_{a}\overbrace{\begin{bmatrix}1 & 6\end{bmatrix}}^{b} = \begin{bmatrix}2 & 12\\3 & 18\\4 & 24\end{bmatrix}$$

One way to look at the output is from a column perspective: the first column of the output is a linear combination of the column $a \rightarrow 1\cdot a$. Likewise, the second column of the output is a linear combination of the column $a \rightarrow 6\cdot a$. This is shown below.

$$\underbrace{\begin{bmatrix}2\\3\\4\end{bmatrix}}_{a}\overbrace{\begin{bmatrix}1 & 6\end{bmatrix}}^{b} = \begin{bmatrix}\underbrace{\boxed{\begin{matrix}2\\3\\4\end{matrix}}}_{1\cdot a} & \begin{matrix}12\\18\\24\end{matrix}\end{bmatrix}$$

$$\underbrace{\begin{bmatrix}2\\3\\4\end{bmatrix}}_{a}\overbrace{\begin{bmatrix}1 & 6\end{bmatrix}}^{b} = \begin{bmatrix}\begin{matrix}2\\3\\4\end{matrix} & \underbrace{\boxed{\begin{matrix}12\\18\\24\end{matrix}}}_{6\cdot a}\end{bmatrix}$$

Another way to look at the output is from a row perspective: the first row is a linear combination of the row $b \rightarrow 2\cdot b$. The second row is a linear combination of the row $b \rightarrow 3\cdot b$. And the third row is a linear combination of the row $b \rightarrow 4\cdot b$.

$$\begin{aligned}
\underbrace{\begin{bmatrix}2\\3\\4\end{bmatrix}}_{a}\overbrace{\begin{bmatrix}1 & 6\end{bmatrix}}^{b} &= \begin{bmatrix}\boxed{2 \quad 12}\\3 \quad 18\\4 \quad 24\end{bmatrix} \leftarrow 2\cdot b\\
&= \begin{bmatrix}2 \quad 12\\\boxed{3 \quad 18}\\4 \quad 24\end{bmatrix} \leftarrow 3\cdot b\\
&= \begin{bmatrix}2 \quad 12\\3 \quad 18\\\boxed{4 \quad 24}\end{bmatrix} \leftarrow 4\cdot b
\end{aligned}$$

And so what will happen if we continue with the second column of A and the second row of B? We will end up with a matrix of the same size, and constructed in the same way, but with different numbers. And so another way we can look at matrix multiplication is,

$$AB = \text{Sum of (Columns of } A) \times \text{(Rows of } B)$$

For example,

$$\begin{bmatrix} 2 & 7 \\ 3 & 8 \\ 4 & 9 \end{bmatrix} \begin{bmatrix} 1 & 6 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & 6 \end{bmatrix} + \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} \begin{bmatrix} 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 12 \\ 3 & 18 \\ 4 & 24 \end{bmatrix} + 0 = \begin{bmatrix} 2 & 12 \\ 3 & 18 \\ 4 & 24 \end{bmatrix}$$

Strang takes a moment here to note something special about the resultant matrix above. We can first inspect the first two columns,

$$\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} ; \begin{bmatrix} 12 \\ 18 \\ 24 \end{bmatrix}$$

They are pointing in the same direction! Now let's look at the rows,

$$\begin{bmatrix} 2 & 12 \end{bmatrix} ; \begin{bmatrix} 3 & 18 \end{bmatrix} ; \begin{bmatrix} 4 & 24 \end{bmatrix}$$

We notice that these too are pointing in the same direction! And so what we will later call the *row space* is just a line. And similarly, the *column space* is just a line.

- *row space*: all combinations of the rows.
- *column space*: all combinations of the columns.

6 Matrix multiplication: by block

Finally, you can also cut the matrix into blocks and do the multiplication by block.

$$\left[\begin{array}{c|c} A_1 & A_2 \\ \hline A_3 & A_4 \end{array}\right] \left[\begin{array}{c|c} B_1 & B_2 \\ \hline B_3 & B_4 \end{array}\right] = \left[\begin{array}{c|c} A_1B_1 + A_2B_3 & ... \\ \hline ... & ... \end{array}\right]$$

Part 1 of lecture done; now onto Inverses.

7 Inverses (square matrices)

A square matrix A is said to be *invertible* (singular) if there exists a matrix A^{-1} such that

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

We call A^{-1} the *inverse* of A; and we do so because it *inverts* (undoes) what A does when multiplied to another vector/matrix.

7.1 Cases with *no* inverse

We will look at cases where there is *no* inverse to reveal an insight that will be important later on. Take an example,

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$$

Why does A not have an inverse? There are various ways to answer this question. First, to frame the question in a different way, *why can't $A^{-1}A$ (or equivalently AA^{-1}) equal* $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$*?* Well, when thinking about columns – if we multiply another matrix by A, the result has columns which are combinations of the columns of A. So for A to have an inverse,

$$a \begin{bmatrix} 1 \\ 2 \end{bmatrix} + b \begin{bmatrix} 3 \\ 6 \end{bmatrix} =? \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{1}$$

For any combination of (a, b), (1) cannot be satisfied. This is because $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 6 \end{bmatrix}$ lie on the same line. And so every combination of these two vectors will be on that line. Unfortunately, $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is not on this line, and so it is impossible to "arrive at" $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ by taking combinations of A's columns, and hence it is impossible for (1) to be satisfied. A visualization is provided in Figure 1.

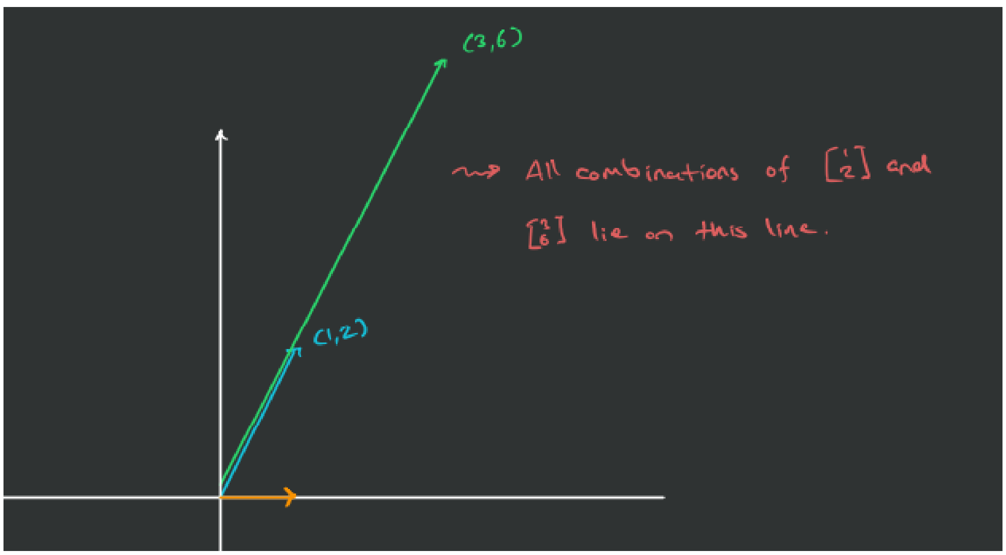

Figure 1: *The columns of A reside on the same line, and cannot reach* $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$

What if the columns of A were on the same line, ***but on this line was*** ?? $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ We again run into an issue. Recall that it is sufficient to find an A^{-1} such that

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

Expanded out,

$$\underbrace{\begin{bmatrix} 5 & 15 \\ 0 & 0 \end{bmatrix}}_{A} \overbrace{\begin{bmatrix} a & c \\ b & d \end{bmatrix}}^{A^{-1}} = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{I}$$

Which is the same as

$$a \begin{bmatrix} 5 \\ 0 \end{bmatrix} + b \begin{bmatrix} 15 \\ 0 \end{bmatrix} =? \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{1}$$

$$c \begin{bmatrix} 5 \\ 0 \end{bmatrix} + d \begin{bmatrix} 15 \\ 0 \end{bmatrix} =? \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{2}$$

Indeed we can find $(a, b) = (\frac{1}{5}, 0)$ such that (1) is satisfied since $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is on the same line as $\begin{bmatrix} 5 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 15 \\ 0 \end{bmatrix}$. *However*, we need to simultaneously be able to satisfy (2). We are incapable of this because we are bound by the assumption that the columns of A are on the same line, but the columns of I are *not* on the same line, and so we cannot possibly reach both. A visualization is shown in Figure 2.

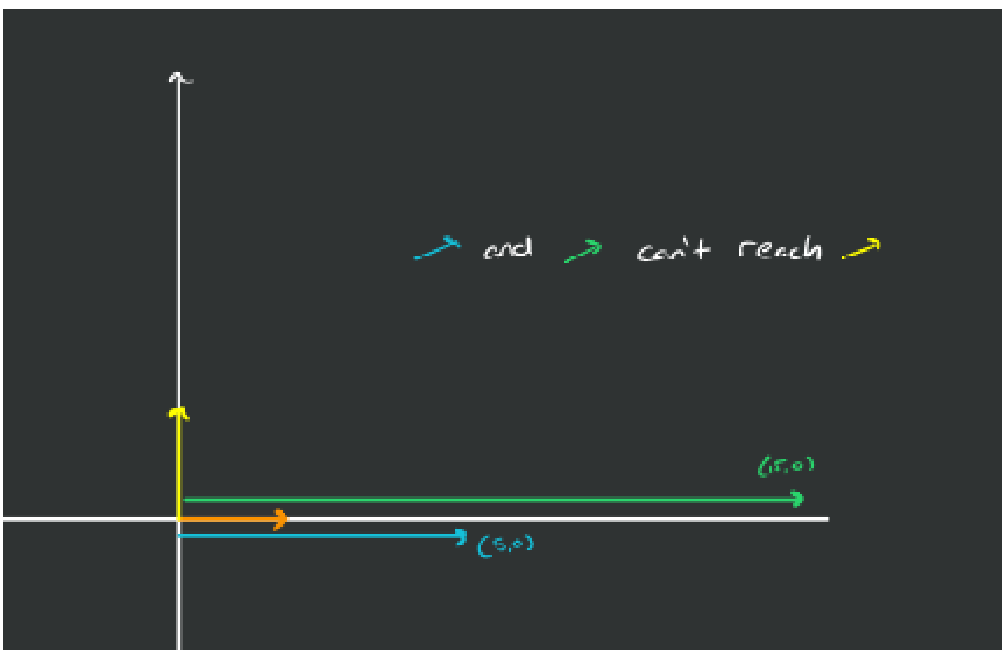

Figure 2: *The columns of A reside on the same line, and can reach $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ but cannot simultaneously reach $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$*

Another way to answer the question – *Why does A not have an inverse?* – is to say: *a square matrix won't have an inverse because we can find a vector $\boldsymbol{x}$ with $A\boldsymbol{x} = 0$* with the condition that $\boldsymbol{x} \neq \mathbf{0}$. To see why, consider the A we previously had,

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$$

We can find a "nontrivial" $\boldsymbol{x} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$ so that

$$A\boldsymbol{x} = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} 3 \\ -1 \end{bmatrix} = 0$$

That is, scaling the first column of A by 3 and the second column by -1 and taking their combination results in 0. Why does this spell disaster for an inverse? Namely, why does $A\boldsymbol{x} = 0$ (for nontrival $\boldsymbol{x}$) spell disaster for an inverse?

Suppose A^{-1} existed,

$$\begin{aligned} A^{-1}A\boldsymbol{x} &= 0 \\ I\boldsymbol{x} &= 0 \\ \boldsymbol{x} &= 0 \end{aligned}$$

a contradiction. And so as a conclusion: *non-invertible (singular matrices) are those where some non-zero combination of their columns gives a zero-column.* In other words, there exists an $\boldsymbol{x}$ such that A "takes it" to 0; and getting back to $\boldsymbol{x}$ once it has been taken to 0 is impossible – hence an inverse does not exist.

8 Gauss-Jordan

For a 2×2 matrix, Gauss-Jordan can be seen as solving two systems of equations at once (this can be extended to $n \times n$ matrices). This relates back to the idea that the columns of A need to *simultaneously* reach the columns of I. And indeed, solving a single system of equations amounts to the columns of A being able to "reach" $\boldsymbol{b}$ – where $\boldsymbol{b}$ is a column of I in our scenario here.

Let's consider an example. We want to find the inverse of A, where

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}$$

This is equivalent to finding A^{-1} such that

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

Which can be expanded out as,

$$\underbrace{\begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}}_{A} \overbrace{\begin{bmatrix} a & c \\ b & d \end{bmatrix}}^{A^{-1}} = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{I}$$

From here, we can break this up by considering the two systems of equations:

$$\begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

And in a different notation,

$$\left[\begin{array}{cc|c} 1 & 3 & 1 \\ 2 & 7 & 0 \end{array}\right]$$
$$\left[\begin{array}{cc|c} 1 & 3 & 0 \\ 2 & 7 & 1 \end{array}\right]$$

Instead of solving these separately via *elimination*, we will *combine* them and then solve via *elimination*:

$$\left[\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 2 & 7 & 0 & 1 \end{array}\right]$$

After doing elimination we get,

$$\left[\begin{array}{c|c} \underbrace{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}}_{I} & \underbrace{\begin{matrix} 7 & -3 \\ -2 & 1 \end{matrix}}_{A^{-1}} \end{array}\right]$$

Why does the RHS of | turn into A^{-1}? Recall that elementary operations do not change the *equality* of the system; and denote E as the elimination matrix culminating all the elimination steps into one. We can summarize what we did above like so,

$$E \left[\ A \mid I\ \right] = \left[\ EA \mid EI\ \right] = \left[\ I \mid E\ \right]$$

Because $EA = I$, E *must be* A^{-1}. Because we were *able* to use E to get $EA = I$, it follows that $E = A^{-1}$.