Matrix Multiplication

Reading: Strang 2.3, 2.4

Learning objective: Understand the motivation for the definition of matrix multiplication and be able to multiply matrices.

Overview

Today we learn how to multiply matrices together.

We motivate matrix multiplication through Gaussian elimination.

We can formulate the row operations of Gaussian elimination in terms of the actions of matrices.

The multiplication of matrices represents the composition of these actions.

Matrix-vector multiplication

We have seen how a matrix multiplies a vector.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = u_1 \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} + u_2 \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} + u_3 \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix}$$

The result is a linear combination of the columns of the matrix, with coefficients of the linear combination given by components of the vector.

Identity Matrix

What does this matrix do to a vector?

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

Nothing! This is called the identity matrix.

In general the identity matrix of size n has ones on the diagonal and zeros everywhere else.

Back to Gaussian elimination

Consider doing Gaussian elimination on the following system.

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ 2x_1 + 4x_2 + 8x_3 &= 5 \\ 3x_1 + 9x_2 + 27x_3 &= 14 \end{aligned}$$

$$\begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 2 & 4 & 8 & \mathbf{5} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix}$$

augmented matrix

The first step is to subtract two times the first equation from the second equation.

Back to Gaussian elimination

The first step is to subtract two times the first equation from the second equation.

$$\begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 2 & 4 & 8 & \mathbf{5} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix} \xrightarrow[R_2' = R_2 - 2R_1]{} \begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 0 & 2 & 6 & \mathbf{3} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix}$$

Let's focus on the "right hand side" vector $\vec{b}$. How does it change?

$$\begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 2 & 4 & 8 & \mathbf{5} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix} \xrightarrow[R_2' = R_2 - 2R_1]{} \begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 0 & 2 & 6 & \mathbf{3} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix}$$

Let's focus on the "right hand side" vector $\vec{b}$. How does it change?

$$\vec{b} = \begin{bmatrix} 1 \\ 5 \\ 14 \end{bmatrix} \qquad \vec{b}_{new} = \begin{bmatrix} 1 \\ 3 \\ 14 \end{bmatrix}$$

initially after row operation

We want to find a matrix E that represents this operation:

$$E\vec{b} = \vec{b}_{new}$$

$$\vec{b} = \begin{bmatrix} 1 \\ 5 \\ 14 \end{bmatrix} \xrightarrow{\quad R_2' = R_2 - 2R_1 \quad} \vec{b}_{new} = \begin{bmatrix} 1 \\ 3 \\ 14 \end{bmatrix}$$

initially

after row operation

The matrix representing this row operation is:

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

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ 14 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} + 5 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 14 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 14 \end{bmatrix}$$

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

No matter what the vector $\vec{b}$ is, E performs the action

$$R_2' = R_2 - 2R_1$$

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = b_1 \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} + b_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + b_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 - 2b_1 \\ b_3 \end{bmatrix}$$

The first and third component of $\vec{b}$ remain unchanged.

Elementary Matrix

A matrix like E that implements a row operation of Gaussian elimination is called an elementary matrix.

How do we form the elementary matrix corresponding to the row operation $R_3' = R_3 - 3R_1$?

Start with the identity matrix: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$

Perform the row operation on the identity matrix.

The result is the corresponding elementary matrix!

Check

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow[R_3' = R_3 - 3R_1]{} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 - 3b_1 \end{bmatrix}$$

The first and second rows remain unchanged.

The third row becomes the original third row minus three times the first row.

Back to example

Let's go back to our Gaussian elimination example.

Now we want to compose these operations.

$$\begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 2 & 4 & 8 & \mathbf{5} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix}$$

$$\Big\downarrow \; R_2' = R_2 - 2R_1$$

$$\begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 0 & 2 & 6 & \mathbf{3} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix} \xrightarrow[R_3' = R_3 - 3R_1]{} \begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 0 & 2 & 6 & \mathbf{3} \\ 0 & 6 & 24 & \mathbf{11} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 2 & 4 & 8 & \mathbf{5} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix} \xrightarrow{R_2' = R_2 - 2R_1} \begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 0 & 2 & 6 & \mathbf{3} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix} \xrightarrow{R_3' = R_3 - 3R_1} \begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 0 & 2 & 6 & \mathbf{3} \\ 0 & 6 & 24 & \mathbf{11} \end{bmatrix}$$

We obtain the final right hand side vector by composing these operations.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \left(\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ 14 \end{bmatrix} \right)$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 14 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 3 \\ 11 \end{bmatrix}$$

Matrix Multiplication

The composition of operations is given by matrix multiplication.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \left(\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ 14 \end{bmatrix} \right)$$

$$= \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \begin{bmatrix} 1 \\ 5 \\ 14 \end{bmatrix}$$

The product of these two matrices should perform the cumulative action of composing these two row operations.

Deriving Matrix Multiplication

The definition of matrix-vector multiplication tells us what the left hand side to be.

$$A(B\vec{u}) = (AB)\vec{u}$$

In order for matrix multiplication to represent the composition of actions, we want this to equal the right hand side.

We want this to hold for every vector $\vec{u}$.

$$A(B\vec{u}) = (AB)\vec{u}$$

We want this to hold for every vector $\vec{u}$.

This will determine the rule for how to multiply matrices.

Suppose that A, B are 2-by-2 matrices and $\vec{u} = (1, 0)$.

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \left(\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix}$$

More compactly, using python-like notation, this is

$$A(B[:, 1])$$

A times the first column of B.

Note: unlike python we start counting with 1 instead of 0.

$$A(B\vec{u}) = (AB)\vec{u}$$

When $\vec{u} = (1, 0)$ the left hand side is $A(B[:, 1])$.

What about the right hand side?

Say that $C = AB$. Then

$$\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} c_{11} \\ c_{21} \end{bmatrix}$$

This is the first column of AB.

We write this as $(AB)[:, 1]$.

Deriving Matrix Multiplication

$$A(B\vec{u}) = (AB)\vec{u}$$

Demanding that this equation holds with $\vec{u} = (1, 0)$ tells us that

$\boxed{\text{The first column of } AB \text{ is } A(B[:, 1]).}$

Deriving Matrix Multiplication

$$A(B\vec{u}) = (AB)\vec{u}$$

How do we figure out the second column of AB ?

Now let's take $\vec{u} = (0, 1)$.

On the left hand side we have $A(B[:, 2])$.

On the right hand side we have $(AB)[:, 2]$.

Deriving Matrix Multiplication

$$A(B\vec{u}) = (AB)\vec{u}$$

Demanding that this equation holds with $\vec{u} = (0, 1)$ tells us that

The second column of AB is $A(B[:, 2])$.

General Case

Now let's derive the formula for matrix multiplication in the general case.

First question, when is matrix multiplication defined?

Say that B is a m-by-n matrix, and $\vec{u} \in \mathbb{R}^n$.

Then $\vec{v} = B\vec{u}$ is a m-dimensional vector.

$$A(B\vec{u}) = A\vec{v}$$

For this to make sense, A must have m columns.

General Case

Conclusion: The matrix product AB is only defined when the number of columns of A equals the number of rows of B.

Now let B be a m-by-n matrix and A a k-by-m matrix.

We again use the desired property

$$A(B\vec{u}) = (AB)\vec{u} \quad \text{for every} \quad \vec{u} \in \mathbb{R}^n$$

to determine the product AB.

Now let B be a m-by-n matrix and A a k-by-m matrix.

$$A(B\vec{u}) = (AB)\vec{u} \quad \text{for every} \quad \vec{u} \in \mathbb{R}^n$$

Let $\vec{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ be the vector whose i^{th} component is one, and all other components are zero.

Then $B\vec{e}_i = B[:, i]$ is the i^{th} column of B.

Thus the left hand side is $A(B[:, i])$.

This is a k-dimensional vector.

The right hand side is the i^{th} column of AB.

Matrix Multiplication

If A is a k-by-m matrix and B is a m-by-n matrix, then the product AB is defined and is a k-by-n matrix.

The i^{th} column of AB is $A(B[:, i])$.

$$\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{km} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ A(B[:, 1]) & A(B[:, 2]) & \cdots & A(B[:, n]) \\ \vdots & \vdots & & \vdots \end{bmatrix}$$

Matrix Multiplication

If A is a k-by-m matrix and B is a m-by-n matrix, then the product AB is defined and is a k-by-n matrix.

The i^{th} column of AB is $A(B[:,i])$.

Every column of AB is a linear combination of the columns of A.

Example

Let's go back to our example from Gaussian elimination.

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

$R_3' = R_3 - 3R_1$ $\qquad$ $R_2' = R_2 - 2R_1$

Action on augmented matrix

So far we have only been talking about the evolution of the right hand side vector.

$$\begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 2 & 4 & 8 & \mathbf{5} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix} \xrightarrow[R_2' = R_2 - 2R_1]{} \begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 0 & 2 & 6 & \mathbf{3} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ 14 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 14 \end{bmatrix}$$

But each column of the augmented matrix changes in the same way!

$$\begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 2 & 4 & 8 & \mathbf{5} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix} \xrightarrow[R_2' = R_2 - 2R_1]{} \begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 0 & 2 & 6 & \mathbf{3} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix}$$

Each column of the augmented matrix changes in the same way!

For example the first column:

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

$$\begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 2 & 4 & 8 & \mathbf{5} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix} \xrightarrow[R_2' = R_2 - 2R_1]{} \begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 0 & 2 & 6 & \mathbf{3} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix}$$

Each column of the augmented matrix changes in the same way!

The second column:

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 9 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 9 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 2 & 4 & 8 & \mathbf{5} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix} \xrightarrow[R_2' = R_2 - 2R_1]{} \begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 0 & 2 & 6 & \mathbf{3} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix}$$

Each column of the augmented matrix changes in the same way!

The third column:

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 8 \\ 27 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \\ 27 \end{bmatrix}$$

Action on augmented matrix

$$\begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 2 & 4 & 8 & \mathbf{5} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix} \xrightarrow[R_2' = R_2 - 2R_1]{} \begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 0 & 2 & 6 & \mathbf{3} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix}$$

The transformation of the augmented matrix is given by multiplying on the left by the corresponding elementary matrix.

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 2 & 4 & 8 & \mathbf{5} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 0 & 2 & 6 & \mathbf{3} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{km} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ A(B[:,1]) & A(B[:,2]) & \cdots & A(B[:,n]) \\ \vdots & \vdots & & \vdots \end{bmatrix}$$

The transformation of the augmented matrix is given by multiplying on the left by the corresponding elementary matrix.

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 2 & 4 & 8 & \mathbf{5} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 0 & 2 & 6 & \mathbf{3} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix}$$

Finishing Gaussian elimination

$$\begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 2 & 4 & 8 & \mathbf{5} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix} \xrightarrow{R_2' = R_2 - 2R_1} \begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 0 & 2 & 6 & \mathbf{3} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix} \xrightarrow{R_3' = R_3 - 3R_1} \begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 0 & 2 & 6 & \mathbf{3} \\ 0 & 6 & 24 & \mathbf{11} \end{bmatrix}$$

$$\xrightarrow{R_3' = R_3 - 3R_2} \begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 0 & 2 & 6 & \mathbf{3} \\ 0 & 0 & 6 & \mathbf{2} \end{bmatrix}$$

The cumulative operation is given by:

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

Check

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 3 & -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 2 & 4 & 8 & \mathbf{5} \\ 3 & 9 & 27 & \mathbf{14} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \mathbf{1} \\ 0 & 2 & 6 & \mathbf{3} \\ 0 & 0 & 6 & \mathbf{2} \end{bmatrix}$$