Chapter 2

Matrix Algebra

Section 2.1

Matrix Operations

Motivation

Recall: we can turn any system of linear equations into a matrix equation

$$Ax = b.$$

This notation is suggestive. Can we solve the equation by "dividing by A"?

$$x \overset{??}{=} \frac{b}{A}$$

Answer: Sometimes, but you have to know what you're doing.

Today we'll study *matrix algebra*: adding and multiplying matrices.

More Notation for Matrices

Let A be an $m \times n$ matrix.

We write a_{ij} for the entry in the ith row and the jth column. It is called the ij**th entry** of the matrix.

$$\begin{array}{cc} \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix} & i\text{th row} \\ j\text{th column} & \end{array}$$

The entries $a_{11}, a_{22}, a_{33}, \ldots$ are the **diagonal entries**; they form the **main diagonal** of the matrix.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \quad \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

A **diagonal matrix** is a *square* matrix whose only nonzero entries are on the main diagonal.

$$\begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix}$$

The $n \times n$ **identity matrix** I_n is the diagonal matrix with all diagonal entries equal to 1. It is special because $I_n v = v$ for *all* v in $\mathbf{R}^n$.

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

More Notation for Matrices

Continued

The **zero matrix** (of size $m \times n$) is the $m \times n$ matrix 0 with all zero entries.

$$0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

The **transpose** of an $m \times n$ matrix A is the $n \times m$ matrix A^T whose rows are the columns of A. In other words, the ij entry of A^T is a_{ji}.

$$\underset{\text{flip}}{\overset{A}{\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}}} \rightsquigarrow \overset{A^T}{\begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{pmatrix}}$$

Addition and Scalar Multiplication

You add two matrices component by component, like with vectors.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} \end{pmatrix}$$

Note you can only add two matrices *of the same size*.

You multiply a matrix by a scalar by multiplying each component, like with vectors.

$$c \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} = \begin{pmatrix} ca_{11} & ca_{12} & ca_{13} \\ ca_{21} & ca_{22} & ca_{23} \end{pmatrix}.$$

These satisfy the expected rules, like with vectors:

$$\begin{aligned} A+B &= B+A & \qquad (A+B)+C &= A+(B+C) \\ c(A+B) &= cA+cB & \qquad (c+d)A &= cA+dA \\ (cd)A &= c(dA) & \qquad A+0 &= A \end{aligned}$$

Matrix Multiplication

Beware: matrix multiplication is more subtle than addition and scalar multiplication.

must be equal

Let A be an $m \times n$ matrix and let B be an $n \times p$ matrix with columns $v_1, v_2 \ldots, v_p$:

$$B = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_p \\ | & | & & | \end{pmatrix}.$$

The **product** AB is the $m \times p$ matrix with columns $Av_1, Av_2, \ldots, Av_p$:

The equality is a definition

$$AB \stackrel{\text{def}}{=} \begin{pmatrix} | & | & & | \\ Av_1 & Av_2 & \cdots & Av_p \\ | & | & & | \end{pmatrix}.$$

In order for $Av_1, Av_2, \ldots, Av_p$ to make sense, the number of columns of A has to be the same as the number of rows of B.

Example

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & -3 \\ 2 & -2 \\ 3 & -1 \end{pmatrix} = \left(\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} -3 \\ -2 \\ -1 \end{pmatrix} \right)$$

$$= \begin{pmatrix} 14 & -10 \\ 32 & -28 \end{pmatrix}$$

Composition of Transformations

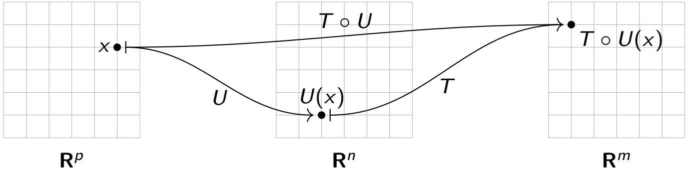
Why is this the correct definition of matrix multiplication?

Definition

Let $T\colon \mathbf{R}^n \to \mathbf{R}^m$ and $U\colon \mathbf{R}^p \to \mathbf{R}^n$ be transformations. The **composition** is the transformation

$$T \circ U\colon \mathbf{R}^p \to \mathbf{R}^m \qquad \text{defined by} \qquad T \circ U(x) = T(U(x)).$$

This makes sense because $U(x)$ (the output of U) is in $\mathbf{R}^n$, which is the domain of T (the inputs of T).



Fact: If T and U are linear then so is $T \circ U$.

Guess: If A is the matrix for T, and B is the matrix for U, what is the matrix for $T \circ U$?

Composition of Linear Transformations

Let $T\colon \mathbf{R}^n \to \mathbf{R}^m$ and $U\colon \mathbf{R}^p \to \mathbf{R}^n$ be *linear* transformations. Let A and B be their matrices:

$$A = \begin{pmatrix} | & | & & | \\ T(e_1) & T(e_2) & \cdots & T(e_n) \\ | & | & & | \end{pmatrix} \quad B = \begin{pmatrix} | & | & & | \\ U(e_1) & U(e_2) & \cdots & U(e_p) \\ | & | & & | \end{pmatrix}$$

Question

What is the matrix for $T \circ U$?

We find the matrix for $T \circ U$ by plugging in the unit coordinate vectors:

$$T \circ U(e_1) = T(U(e_1)) = T(Be_1) = A(Be_1) = (AB)e_1$$

because Be_1 is the first column of B, which is $U(e_1)$. For any other i, the same works:

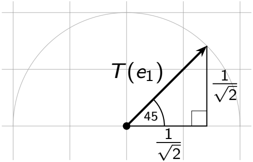
$$T \circ U(e_i) = T(U(e_i)) = T(Be_i) = A(Be_i) = (AB)e_i.$$

This says that the ith column of the matrix for $T \circ U$ is the ith column of AB.

> The matrix of the composition is the product of the matrices!

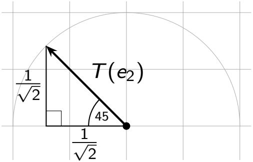
Composition of Linear Transformations
Example

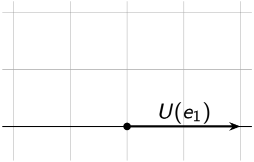
Let $T\colon \mathbf{R}^2 \to \mathbf{R}^2$ be rotation by 45°, and let $U\colon \mathbf{R}^2 \to \mathbf{R}^2$ be projection onto the x-axis. Let's compute their standard matrices A and B:



$$T(e_1) = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$$

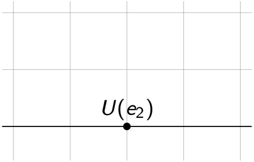
$$T(e_2) = \frac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\end{pmatrix}$$





$$U(e_1) = \begin{pmatrix}1\\0\end{pmatrix}$$

$$U(e_2) = \begin{pmatrix}0\\0\end{pmatrix}$$



$$\implies \quad A = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\1 & 1\end{pmatrix} \qquad B = \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}$$

Composition of Linear Transformations

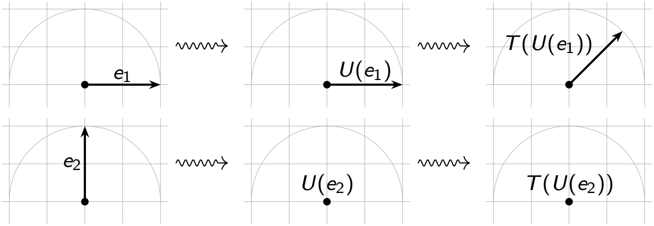
Example, continued

So the matrix C for $T \circ U$ is

$$C = AB = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

$$= \left(\tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \end{pmatrix} \right) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}.$$

Check:



$$T \circ U(e_1) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$T \circ U(e_2) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Longrightarrow \quad C = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \qquad ✔$$

Let

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & -3 \\ 2 & -2 \\ 3 & -1 \end{pmatrix}.$$

Let $T(x) = Ax$ and $U(y) = By$, so

$$T\colon \mathbf{R}^3 \longrightarrow \mathbf{R}^2 \qquad U\colon \mathbf{R}^2 \longrightarrow \mathbf{R}^3 \qquad T \circ U\colon \mathbf{R}^2 \longrightarrow \mathbf{R}^2.$$

Let's find the matrix for $T \circ U$:

$$T \circ U(e_1) = T(U(e_1)) = T(Be_1) = T\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = A\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 14 \\ 32 \end{pmatrix}$$

$$T \circ U(e_2) = T(U(e_2)) = T(Be_2) = T\begin{pmatrix} -3 \\ -2 \\ -1 \end{pmatrix} = A\begin{pmatrix} -3 \\ -2 \\ -1 \end{pmatrix} = \begin{pmatrix} -10 \\ -28 \end{pmatrix}$$

Before we computed $AB = \begin{pmatrix} 14 & -10 \\ 32 & -28 \end{pmatrix}$, so AB is the matrix of $T \circ U$.

The Row-Column Rule for Matrix Multiplication

Recall: A row vector of length n times a column vector of length n is a scalar:

$$\begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + \cdots + a_n b_n.$$

Another way of multiplying a matrix by a vector is:

$$Ax = \begin{pmatrix} — r_1 — \\ \vdots \\ — r_m — \end{pmatrix} x = \begin{pmatrix} r_1 x \\ \vdots \\ r_m x \end{pmatrix}.$$

On the other hand, you multiply two matrices by

$$AB = A \begin{pmatrix} | & & | \\ c_1 & \cdots & c_p \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | \\ Ac_1 & \cdots & Ac_p \\ | & & | \end{pmatrix}.$$

It follows that

$$AB = \begin{pmatrix} — r_1 — \\ \vdots \\ — r_m — \end{pmatrix} \begin{pmatrix} | & & | \\ c_1 & \cdots & c_p \\ | & & | \end{pmatrix} = \begin{pmatrix} r_1 c_1 & r_1 c_2 & \cdots & r_1 c_p \\ r_2 c_1 & r_2 c_2 & \cdots & r_2 c_p \\ \vdots & \vdots & & \vdots \\ r_m c_1 & r_m c_2 & \cdots & r_m c_p \end{pmatrix}$$

The ij entry of $C = AB$ is the ith row of A times the jth column of B:

$$c_{ij} = (AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}.$$

This is how everybody on the planet actually computes AB. Diagram ($AB = C$):

$$\begin{array}{cc} \begin{pmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ik} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mk} & \cdots & a_{mn} \end{pmatrix} & i\text{th row} \end{array} \cdot \begin{array}{c} \begin{pmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1p} \\ \vdots & & \vdots & & \vdots \\ b_{k1} & \cdots & b_{kj} & \cdots & b_{kp} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & b_{nj} & \cdots & b_{np} \end{pmatrix} \\ j\text{th column} \end{array} = \begin{array}{c} \begin{pmatrix} c_{11} & \cdots & c_{1j} & \cdots & c_{1p} \\ \vdots & & \vdots & & \vdots \\ c_{i1} & \cdots & c_{ij} & \cdots & c_{ip} \\ \vdots & & \vdots & & \vdots \\ c_{m1} & \cdots & c_{mj} & \cdots & c_{mp} \end{pmatrix} \\ ij \text{ entry} \end{array}$$

Example

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & -3 \\ 2 & -2 \\ 3 & -1 \end{pmatrix} = \begin{pmatrix} 1\cdot 1 + 2\cdot 2 + 3\cdot 3 & \square \\ \square & \square \end{pmatrix} = \begin{pmatrix} 14 & \square \\ \square & \square \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & -3 \\ 2 & -2 \\ 3 & -1 \end{pmatrix} = \begin{pmatrix} \square & \square \\ 4\cdot 1 + 5\cdot 2 + 6\cdot 3 & \square \end{pmatrix} = \begin{pmatrix} \square & \square \\ 32 & \square \end{pmatrix}$$

Properties of Matrix Multiplication

Mostly matrix multiplication works like you'd expect. Suppose A has size $m \times n$, and that the other matrices below have the right size to make multiplication work.

$$
\begin{aligned}
A(BC) &= (AB)C & \qquad A(B+C) &= (AB+AC) \\
(B+C)A &= BA+CA & c(AB) &= (cA)B \\
c(AB) &= A(cB) & I_n A &= A \\
AI_m &= A & &
\end{aligned}
$$

Most of these are easy to verify.

Associativity is $A(BC) = (AB)C$. It is a pain to verify using the row-column rule! Much easier: use associativity of linear transformations:

$$S \circ (T \circ U) = (S \circ T) \circ U.$$

This is a good example of an instance where having a conceptual viewpoint saves you a lot of work.

Recommended: Try to verify all of them on your own.

Warnings!

- AB is usually not equal to BA.

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 2 & 0 \end{pmatrix} \qquad \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix}$$

  In fact, AB may be defined when BA is not.

- $AB = AC$ does not imply $B = C$, even if $A \neq 0$.

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 5 & 6 \end{pmatrix}$$

- $AB = 0$ does not imply $A = 0$ or $B = 0$.

$$\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Other Reading

Read about powers of a matrix and multiplication of transposes in §2.1.