Chapter 3 – Linear Transformations

From chapter 2, we understand how matrices can be used to solve systems of linear equations. In this chapter we will view them very differently: as functions from $\mathbb{R}^n$ to $\mathbb{R}^m$.

§3.1 – Operations on Matrices

Let's put our understanding of systems aside for the time being and instead think of a matrix as just a block of numbers.

A matrix A is said to have size $m \times n$ ("m by n") if it has m rows and n columns.

Ex: $\begin{bmatrix} 1 & -1 \\ 2 & 4 \\ 3 & 3 \end{bmatrix}$ is 3×2 $\qquad$ $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ is 2×2

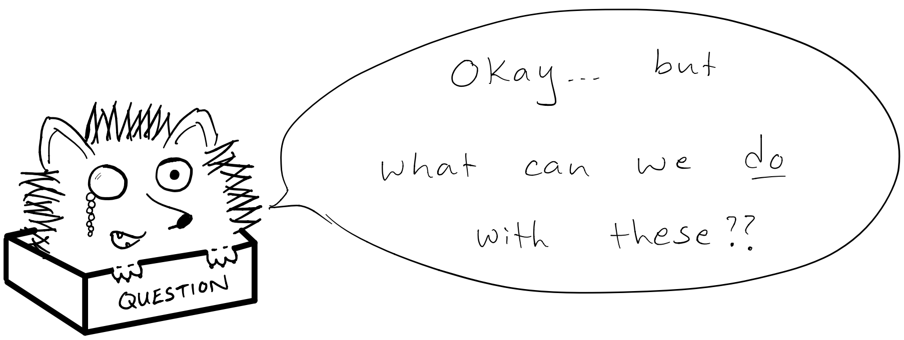
An $n \times n$ matrix is called a square matrix

As in the above example, we sometimes use a_{ij} to denote the entry of A in row i, column j.

When dealing with matrices abstractly, we may write

$$\boxed{A = (a_{ij})}$$

Two matrices $A = (a_{ij})$ and $B = (b_{ij})$ equal if they have the same size and the same entries (i.e., $a_{ij} = b_{ij}$ for all i, j)



Well... just like with vectors we can add two matrices of the same size:

If $A = (a_{ij})$, $B = (b_{ij})$, then $A + B = (a_{ij} + b_{ij})$

Ex:

$$\begin{bmatrix} 1 & -1 \\ 2 & 4 \\ 3 & 3 \end{bmatrix} + \begin{bmatrix} 0 & 2 \\ 1 & 1 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 3 & 5 \\ 6 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 2 \\ 1 & -2 \end{bmatrix} + \begin{bmatrix} 7 & 0 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 10 & 2 \\ 0 & -3 \end{bmatrix}$$

What is $\begin{bmatrix} 1 & -1 \\ 2 & 4 \\ 3 & 3 \end{bmatrix} + \begin{bmatrix} 3 & 2 \\ 1 & -2 \end{bmatrix}$? Not defined! The sizes are different!

Instead of a zero vector, we have a zero matrix.

The $m \times n$ zero matrix is $O_{m,n}$

Ex: $O_{2,3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

We can also multiply a matrix by a scalar:

If $A = (a_{ij})$ and $t \in \mathbb{R}$, then $tA = (t \cdot a_{ij})$

Ex: $2\begin{bmatrix} 1 & -1 \\ 2 & 4 \\ 3 & 3 \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ 4 & 8 \\ 6 & 6 \end{bmatrix}$, $(-1)\begin{bmatrix} 3 & 2 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} -3 & -2 \\ -1 & 2 \end{bmatrix}$

These operations satisfy properties similar to what we know for vectors

e.g. $t(A+B) = tA + tB$, $O_{m,n} + A = A$,

$A + B = B + A$, etc...

Matrix Multiplication

In order to define the notion of a product of matrices, let's first see how to define multiplication of a matrix A and a vector $\vec{x}$.

Suppose $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ and $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

Define the product $A\vec{x}$ to be the vector

$$\vec{y} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix}$$

So y_1 = dot product of 1^{st} row of A with $\vec{x}$.

y_2 = dot product of 2^{nd} row of A with $\vec{x}$.

Ex:

$$\begin{bmatrix} 3 & 2 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 5 \\ 4 \end{bmatrix} = \begin{bmatrix} 3(5) + 2(4) \\ 1(5) - 2(4) \end{bmatrix} = \begin{bmatrix} 23 \\ -3 \end{bmatrix}$$

Why would we define it like this??

Because this new vector resembles the left hand side of a system of equations! So instead of saying that $\vec{x}$ is a solution to $[A|\vec{b}]$, we can simply write $A\vec{x} = \vec{b}$.

Just like we know for systems (§2.3), the product $A\vec{x}$ is a linear combination of the columns of A:

$$A\vec{x} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix}$$

$$= x_1 \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}$$

Now suppose we have a second matrix $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$ and want to define a matrix product BA.

How should we do it?

Let's define it in terms of vector multiplication of the columns of A:

$$BA = B\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} B\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} & B\begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} \end{bmatrix}$$

Ex:

$$\begin{bmatrix} 3 & 2 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 4 & -1 \end{bmatrix} = \begin{bmatrix} 3(0) + 2(4) & 3(1) + 2(-1) \\ 1(0) - 2(4) & 1(1) - 2(-1) \end{bmatrix}$$

$$= \begin{bmatrix} 8 & 1 \\ -8 & 3 \end{bmatrix}$$

Note: This may look complicated, but it's not!

In row i column j of BA, the entry is the dot product of B's i^{th} row and A's j^{th} column!

This works because the # of columns of B is equal to the # of rows of A.

Let's use this example to define the product of matrices more generally:

Definition: If B is an $m \times n$ matrix and A is an $n \times p$ matrix, then BA is the $m \times p$ matrix whose (i,j)-entry is the dot product of the i^{th} row of B with the j^{th} column of A.

If the # of columns of B is not equal to the number of rows of A, the product BA is not defined.

Ex:

$$\begin{bmatrix} 3 & 5 & 0 & 0 \\ 2 & 1 & -1 & 4 \end{bmatrix} \begin{bmatrix} -1 & 3 \\ 2 & -1 \\ 0 & 1 \\ 9 & 1 \end{bmatrix} = \begin{bmatrix} 7 & 4 \\ 36 & 8 \end{bmatrix}$$

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

$$\begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 4 & 10 \\ 0 & 0 & 0 & 0 \\ 3 & 0 & 6 & 15 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 5 \\ 0 & 3 \\ 5 & 1 \end{bmatrix} = \text{Not defined!}$$

\# of columns of $\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} = 2$

\# of rows of $\begin{bmatrix} 1 & 5 \\ 0 & 3 \\ 5 & 1 \end{bmatrix} = 3$

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

We will think of a 1×1 matrix as a scalar!

So for us, $[-2] = -2$.

Properties: If A, B, C are matrices of the correct sizes, and $t \in \mathbb{R}$, then

(i) $A(B+C) = A + AC$

(ii) $(B+C)A = BA + CA$

(iii) $A(BC) = (AB)C$

(iv) $t(AB) = (tA)B = A(tB)$

Remarks:

(1) Matrix multiplication is not commutative:

$$\boxed{AB \neq BA\,!} \qquad \text{(in general)}$$

Ex: $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

(2) Matrix multiplication is not cancellative:

$$\boxed{\text{We can have } AB = AC \text{ but } B \neq C.}$$

Ex: $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, but $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

(3) The zero matrix times any matrix A is the zero matrix.

(4) Let the $n \times n$ identity matrix be

$$I_n = [\vec{e}_1 \ \vec{e}_2 \cdots \vec{e}_n]$$

where $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ is the standard basis for $\mathbb{R}^n$.

(so $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, etc...)

If A is any $m \times n$ matrix, then $\boxed{I_m A = A I_n = A}$

The equation $A = A I_n = [A\vec{e}_1 \; A\vec{e}_2 \; \cdots \; A\vec{e}_n]$ shows that the i^{th} column of A is Ae_i.

(This will be important in §3.2)

We sometimes just write I for the identity matrix when its size is clear from context.

(5) If A and B are two matrices that multiply all vectors $\vec{x}$ in the same way, then $A = B$.

Theorem: If A and B are $m \times n$ matrices and $A\vec{x} = B\vec{x}$ for all vectors $\vec{x} \in \mathbb{R}^n$, then $A = B$.

Proof: If $A\vec{x} = B\vec{x}$ for all vectors $\vec{x}$, then $A\vec{e}_i = B\vec{e}_i$ for each $\vec{e}_i$ in the standard basis.

By remark (4), this means that the i^{th} column of A and the i^{th} column of B are the same for all i.

Therefore, A and B have the same entries, so $A = B$ $\blacksquare$

Transpose of a Matrix

We can turn an $m \times n$ matrix $A = (a_{ij})$ into an $n \times m$ matrix by turning rows into columns and columns into rows. This matrix is called the transpose of A, denoted A^T.

Formally, the (i,j)-entry of A^T is a_{ji}, the (j,i)-entry of A.

Ex:
$$\begin{bmatrix} 1 & -1 \\ 2 & 4 \\ 3 & 3 \end{bmatrix}^T = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 4 & 3 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 2 \\ 1 & -2 \end{bmatrix}^T = \begin{bmatrix} 3 & 1 \\ 2 & -2 \end{bmatrix}, \qquad \begin{bmatrix} 5 \\ 4 \end{bmatrix}^T = \begin{bmatrix} 5 & 4 \end{bmatrix}$$

Properties: For all $m \times n$ matrices A and B,

(i) $(A^T)^T = A$;

(ii) $(A+B)^T = A^T + B^T$;

(iii) $(tA)^T = tA^T$ for all $t \in \mathbb{R}$.

(iv) $(AB)^T = B^T A^T$