Lecture #7

3.2 Matrix Algebra

Early on, we saw matrices appear when we were solving linear systems of equations or taking linear combinations of a set of column vectors, but this was largely superficial in that the "matrix" was really just a notation thing, which meant the same thing as a linear combination. However, in the last section, we saw that matrices actually represented a large class of transformations, linear transformations, which show up everywhere. So, we are going to start defining an algebra (how to add, multiply, etc.) for matrices, all the while keeping our definitions consistent with how we would expect linear transformations to behave.

DEF (Addition and Scalar Multiplication of Matrices)

Let c be a scalar, and let

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{bmatrix}$$

be $n \times m$ matrices. Then $\boxed{\text{addition}}$ and $\boxed{\text{scalar multiplication}}$

are defined as follows:

(a) Addition: $A+B = \begin{bmatrix} (a_{11}+b_{11}) & (a_{12}+b_{12}) & \cdots & (a_{1m}+b_{1m}) \\ (a_{21}+b_{21}) & (a_{22}+b_{22}) & \cdots & (a_{2m}+b_{2m}) \\ \vdots & \vdots & & \\ (a_{n1}+b_{n1}) & (a_{n2}+b_{n2}) & \cdots & (a_{nm}+b_{nm}) \end{bmatrix}$

(Just add corresponding entries!)

(b) Scalar Multiplication:

$$cA = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1m} \\ ca_{21} & ca_{22} & \cdots & ca_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{n1} & ca_{n2} & \cdots & ca_{nm} \end{bmatrix}$$

(Just multiply each entry by the scalar!)

PROPERTIES

Let s and t be scalars, A, B, and C be matrices of dimension $n \times m$, and O_{nm} be the $n \times m$ matrix with all zero entries. Then

(a) $A+B = B+A$ (addition commutes!)

(b) $s(A+B) = sA+sB$

(c) $(s+t)A = sA+tA$

(d) $(A+B)+C = A+(B+C)$ (addition is associative!)

(e) $(st)A = s(tA)$

(f) $A + O_{nm} = A$

REMARK We call O_{nm} the "additive identity" since adding it to any matrix will give you the same matrix.

EXERCISE Use the definition of addition and scalar multiplication to prove a few of the properties.

EXERCISE Let $A = \begin{bmatrix} 4 & -1 \\ 2 & -3 \end{bmatrix}$, $B = \begin{bmatrix} 3 & -1 \\ 5 & 0 \end{bmatrix}$. Find $3A$ and $A-2B$.

Next, we will define multiplication between two matrices. First, recall how we multiply a matrix against a vector:

top row • column vector

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ 3x_1 + 4x_2 \end{bmatrix}$$

bottom row • column vector.

To multiply matrices, we will extend this multiplication in a natural way, for example:

first row in A • first column in B

first row in A • second column in B

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

2^{nd} row in A • 1^{st} column in B

2^{nd} row in A • 2^{nd} column in B.

In general, for $n \times m$ matrices, we need:

$$\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix}_{n \times k} \begin{bmatrix} & & \\ & B & \\ & & \end{bmatrix}_{k \times m}$$

(*) [A to have the same number of columns as B has rows!]

then we can multiply (otherwise the dot product makes no sense!).

DEF (Matrix multiplication)

Let A be an $n \times k$ matrix, B be $k \times m$ matrix, with entries

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{bmatrix}_{n \times k}, \quad B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & \ddots & \\ b_{k1} & b_{k2} & \cdots & b_{km} \end{bmatrix}_{k \times m}$$

then define $\boxed{\text{multiplication}}$ $AB = C$, where C is an $\underline{n \times m}$ matrix with entries

$$c_{ij} = (i^{th} \text{ row of } A) \bullet (j^{th} \text{ column of } B)$$

For example, $c_{21} = (2^{nd} \text{ row of } A) \bullet (1^{st} \text{ column of } B)$.

EXAMPLE 1

Compute AB where $A = \begin{bmatrix} 2 & -3 \\ 0 & -1 \\ 1 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 3 & -4 & 2 \\ 4 & -2 & 5 & -3 \end{bmatrix}$.

If it is not possible, write "NOT POSSIBLE."

First, notice A is a 3×2 matrix and B is a 2×4 matrix, so we can multiply. AB will be a 3×4 matrix

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

$$= \begin{bmatrix} -8 & 12 & -23 & 13 \\ -4 & 2 & -5 & 3 \\ 18 & -5 & 16 & -10 \end{bmatrix}_{3\times4}$$

EXERCISE

Do the same for

a) $A = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 5 & 6 \\ 1 & 2 \end{bmatrix}$.

b) $A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \\ -5 & 6 \\ 2 & -7 \end{bmatrix}$. (careful!)

Next, we will identify some special matrices that play an important role in the algebra we are defining. We've actually already identified one:

<u>DEF</u> (Additive Identity)

Recall from that for any matrix $A_{n \times m}$, if we add to it the matrix O_{nm} (where every entry is 0), we have the following property

$$A + O_{nm} = O_{nm} + A = A$$

Since addition against O_{nm} returns the non-zero matrix, we say O_{nm} is the <u>additive identity</u>.

There is another matrix, a multiplicative identity, that we can try to find. We want a matrix I such that $AI = A$ ... and perhaps a different I such that $IA = A$. We can reverse engineer what I needs to be by using the definition of matrix multiplication.

First, let A be an $n \times m$ matrix. If $AI = A$, we need for I to be an $\underline{\underline{m \times m}}$ square matrix. If $IA = A$, we need I to be an $\underline{\underline{n \times n}}$ matrix.

So, the identity when you multiply on the right may be different than the identity when you multiply on the left. If A is a square matrix, then this identity might be the same... (it will be!).

Now, if $AI = A$, then

$$a_{ij} = (i^{th} \text{ row of } A) \cdot (j^{th} \text{ row of } I)$$

dot product between vectors!

$$= (a_{i1}, a_{i2}, \ldots, a_{im}) \cdot (j^{th} \text{ row of } I).$$

a_{ij} is one of these!

So, we need the j^{th} row of I to "pick out a_{ij}."

We can do this by making the row: $(0, 0, \ldots, 0, 1, 0, \ldots, 0)$

j^{th} spot!

so that the dot product becomes

$$= (a_{i1}, a_{i2}, \ldots, a_{ij}, \ldots, a_{im}) \cdot (0, 0, \ldots, 1, \ldots, 0)$$

$$= a_{ij}$$

as desired.

In other words, I in this case needs to be an <u>$m \times m$</u> matrix with one's along the diagonal, and 0's for all other entries. We'll call it I_m:

$$I_m = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}_{m \times m}$$

Similarly, if we look for I such that $IA = A$, we will see that it needs to be $n \times n$ and

$$I_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & & \\ \vdots & \vdots & & & \\ 0 & 0 & & & 1 \end{bmatrix}_{n \times n}.$$

<u>EXERCISE</u> Justify this last statement using the argument for I_m.

Next, notice that if $m = n$, $I_m = I_n$.

<u>DEF</u> We call I_m and I_n $\boxed{\text{\underline{identity} \underline{matrices}}}$.

<u>NOTE</u> We could call I a multiplicative identity, but it is so important, we just call it the identity.

Next, let's examine what properties our matrix algebra has:

PROPERTIES

Let s be a scalar, and let A, B, and C be matrices. Then each of the following holds in the cases where the indicated operations are defined:

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

(b) $A(B+C) = AB + BC$

(c) $(A+B)C = AC + BC$

(d) $s(AB) = (sA)B = A(sB)$

(e) $AI = A$

(f) $IA = A$.

EXERCISE

Let $A = \begin{bmatrix} 2 & -3 \\ -1 & 5 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 7 \\ 4 & -2 \end{bmatrix}$, $C = \begin{bmatrix} -3 & -4 \\ 0 & -1 \end{bmatrix}$.

Verify that $A(BC) = (AB)C$ and $A(B+C) = AB + AC$.

EXAMPLE 2

Compute AB and BA where $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$ and $B = \begin{bmatrix} -4 & 6 \\ 2 & -3 \end{bmatrix}$.

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

$$BA = \begin{bmatrix} -4 & 6 \\ 2 & -3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} 14 & 28 \\ -7 & -14 \end{bmatrix}$$

Notice $\boxed{\text{our matrix multiplication does not necessarily commute!}}$

And, even if A, B are non-zero, $AB = O_{nm}$ ...

THM Let A, B, C be nonzero matrices. Then

(1) It is possible that $AB \neq BA$.

(2) $AB = O$ does not mean $A = O$ or $B = O$.

(3) $AC = BC$ does not mean $A = B$ or $C = O$.

To see this last one, try the following exercise:

EXERCISE Let $A = \begin{bmatrix} -3 & 3 \\ 11 & -3 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 2 \\ 3 & 1 \end{bmatrix}$ and $C = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$.

Show that $AC = BC$.

Why is this an important theorem? Notice that for real numbers with multiplication, (1) $xy = yx$, (2) $xy = 0$ means $x = 0$ or $y = 0$; and (3) $xy = zy$ means either $y = 0$ or $x = z$.

In other words, matrix multiplication is very different, so we have to be careful!

This does leave us with a pretty big question though. Why would we define matrix multiplication like this if it is not so well behaved ... ?

Let's start with a thought experiment. Two matrices will correspond to two linear transformations, and as I mentioned before, we are interested in matrices because they are linear transformations. How do we multiply two linear transformations? That doesn't really make sense ... but if we have two linear transformations, I may be able to apply one after the other:

Let $T: \mathbb{R}^n \longrightarrow \mathbb{R}^k$ and $S: \mathbb{R}^k \longrightarrow \mathbb{R}^m$. Then, since T maps into $\mathbb{R}^k$, and S maps out of $\mathbb{R}^k$ into $\mathbb{R}^m$, I can apply T to a vector from $\mathbb{R}^n$ and then apply S to the result:

For $\vec{x}$ in $\mathbb{R}^n$, $T(\vec{x})$ is in $\mathbb{R}^k$, so $S(T(\vec{x}))$ makes sense. $S(T(\vec{x}))$ is then in $\mathbb{R}^m$.

Is $S(T(\vec{x}))$ a linear transformation? It is the **composition** of two linear transformations, S and T.

(DEF $S \circ T(\vec{x}) := S(T(\vec{x}))$ is the **composition**.)

just notation.

Let's see if $S \circ T$, the composition, is a linear transformation by trying to write it as a matrix...

First, let $T(\vec{x}) = A\vec{x}$ where A is a $k \times n$ matrix.

Then, let $S(\vec{x}) = B\vec{x}$ where B is an $m \times k$ matrix. Then

$$
\begin{aligned}
S \circ T(\vec{x}) &= S(T(\vec{x})) \\
&= S(A\vec{x}) \\
&= S(A\vec{x}) \\
&= S(x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n) \\
&= x_1 S(\vec{a}_1) + x_2 S(\vec{a}_2) + \cdots + x_n S(\vec{a}_n). \\
&= x_1 B\vec{a}_1 + x_2 B\vec{a}_2 + \cdots + x_n B\vec{a}_n \\
&= \begin{bmatrix} B\vec{a}_1 & B\vec{a}_2 & \cdots & B\vec{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\
&= (BA) \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = BA\vec{x}
\end{aligned}
$$

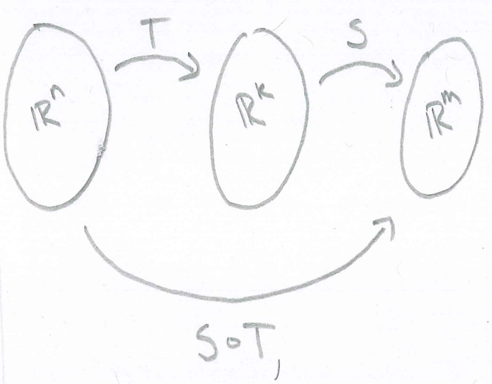
$\vec{x}$ is in $\mathbb{R}^n$

$\vec{a}_i$ is the i^{th} column vector.

S is a linear transformation

!!

PICTURE



$S \circ T : \mathbb{R}^n \to \mathbb{R}^m$

EXERCISE Show that the j^{th} column of BA can be written as $B\vec{a}_j$.

So, we see two things: First, $S \circ T$ (the composition) is a linear transformation, and the associated matrix is BA, i.e., we multiply matrices!

NOTE Had we approached this geometrically, we could have used one of the defining geometric properties of linear transformations. Namely, linear transformations send lines to lines. It stands to reason that we should expect $S \circ T$ to be a linear transformation since T sends lines to lines, then S will send those lines to lines again, i.e. $S \circ T$ sends lines to lines.

NOTE Notice that we could have shortened our argument on the last page by recognizing the column vector as an "$n \times 1$" matrix:

$$\text{Then, } \quad S \circ T(\vec{x}) = S(T(\vec{x})) = S(A\vec{x}) = B(A\vec{x}) = (BA)\vec{x}$$

property (a) on page 9.

Next, we'll introduce a new sort of operation, something you can't really do with real numbers.

DEF (transpose)

The $\boxed{\text{transpose}}$ of a matrix A is denoted A^T and results from interchanging rows and columns.

In other words, if A is an $n \times m$ matrix, A^T is an $m \times n$ matrix. In addition, the entries of A^T, which we denote $a_{ij}^{(T)}$, are

$$a_{ij}^{(T)} = a_{ji}$$

entries from A

EXAMPLE 3

Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$. Compute A^T.

$\Rightarrow A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$. Notice, A is 2×3 and A^T is 3×2.

PROPERTIES (of the transpose)

Let A and B be $n \times m$ matrices, C an $m \times k$ matrix, and s a scalar. Then

(a) $(A+B)^T = A^T + B^T$

(b) $(sA)^T = sA^T$

(c) $(AC)^T = C^T A^T$.

EXERCISE Convince yourself (a) and (b) are true.

For (c), notice that the transpose does something a little weird when you are multiplying two matrices. Recall that if we let $B = AC$, then

$$b_{ij} = (i^{th} \text{ row of } A) \cdot (j^{th} \text{ column of } C)$$

So if we swap rows and columns, then the entries of B transpose will be

$$b_{ij}^{(T)} = b_{ji} = (j^{th}\text{-row of } A) \cdot (i^{th}\text{-column of } C).$$

Now, notice that $C^T A^T := \tilde{B}$ has entries

$$\begin{aligned}
\tilde{b}_{ij} &= (i^{th}\text{-row of } C^T) \cdot (j^{th}\text{-column of } A^T) \\
&\quad\ \ \Downarrow \text{transpose!} \qquad\qquad\qquad \Downarrow \text{transpose!} \\
&= (i^{th}\text{-column of } C) \cdot (j^{th}\text{-row of } A) \\
&= b_{ji} \\
&= b_{ij}^{(T)},
\end{aligned}$$

i.e., we see that $\tilde{B}$ is B^T, so $\overset{(B^T)}{(AC)^T} = C^T A^T$.

<u>EXERCISE</u> Review this argument. Do you see how the operation of exchanging rows and columns is working with how we defined multiplication?

<u>NOTE</u> The transpose will play a role later, and it will make sense why the order of the matrices swap.

Powers of a Matrix

We can also define what we mean by powers of a matrix by using our definition of multiplication:

Let

$$\boxed{A^k := \underbrace{AA\cdots A}_{\text{multiply } A \text{ } k \text{ times.}}.}$$

NOTE: What dimension matrix can A be? It must be square!

EXAMPLE 4

Let $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -3 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}$.

Compute A^3 and B^4.

$$A^3 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -3 \end{bmatrix}\begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -3 \end{bmatrix}\begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -3 \end{bmatrix}\begin{bmatrix} 4 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

$$= \begin{bmatrix} 8 & 0 & 0 \\ 0 & 64 & 0 \\ 0 & 0 & -27 \end{bmatrix} \left(\overset{\text{notice!}}{=} \begin{bmatrix} 2^3 & 0 & 0 \\ 0 & 4^3 & 0 \\ 0 & 0 & (-3)^3 \end{bmatrix} \right)$$

$$B^4 = \begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}$$

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

$$= \begin{bmatrix} 1 & 45 \\ 0 & 16 \end{bmatrix} \left(\overset{\text{notice!}}{=} \begin{bmatrix} 1^4 & 45 \\ 0 & 2^4 \end{bmatrix} \right)$$

In the last example, both A and B had a special form. In general, when you multiply arbitrary matrices, what we noticed is not necessarily going to happen.

Matrix A is a diagonal matrix:

<u>DEF</u> (diagonal matrix)

We say a matrix A is a $\boxed{\text{diagonal}}$ matrix if it is of the form

$$A = \begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ 0 & 0 & \ddots & & \\ \vdots & \vdots & & & \\ 0 & 0 & & & a_{nn} \end{bmatrix}_{n \times n}.$$

(i.e. the only elements allowed to be non-zero are along the diagonal. Note that a_{jj} could be zero, though.)

Matrix B from the last example is an upper triangular matrix.

<u>DEF</u> (upper triangular, lower triangular, triangular)

We say a matrix A is an $\boxed{\text{upper triangular}}$ matrix if it is of the form

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & & & a_{nn} \end{bmatrix}_{n \times n}$$

Similarly, a matrix A is said to be $\boxed{\underline{\text{lower triangular}}}$ if it is of the form

$$A = \begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ a_{31} & a_{32} & a_{33} & & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ & & & & a_{nn} \\ a_{n1} & a_{n2} & a_{n3} & \cdots & \end{bmatrix}_{n \times n}$$

If a matrix is either upper or lower triangular, then we say the matrix is $\boxed{\underline{\text{triangular}}}$.

Obviously not all matrices are diagonal or triangular, but if they are, they have nice properties when it comes to taking powers of the matrix.

<u>THM</u> (1) If a matrix A is diagonal, then A^k is diagonal for any k that is a natural number. Additionally,

$$A^k = \begin{bmatrix} a_{11}^k & 0 & \cdots & 0 \\ 0 & a_{22}^k & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & a_{nn}^k \end{bmatrix}_{n \times n}.$$

(2) If a matrix A is upper triangular, then A^k is upper triangular for any natural number k. Similarly, If A is a lower triangular matrix, then A^k is a lower triangular matrix for any natural number k.

Now, we have constructed enough machinery to make an observation about the elementary row operations we have been doing on matrices (and more often, augmented matrices).

Consider the following matrix:

$$A = \begin{bmatrix} 1 & 2 & 5 \\ -2 & 7 & 3 \end{bmatrix}.$$

Let's do a row operation: $2R_1 + R_2 \to R_2$. Then

$$\begin{bmatrix} 1 & 2 & 5 \\ -2 & 7 & 3 \end{bmatrix} \overset{2R_1 + R_2 \to R_2}{\sim} \begin{bmatrix} 1 & 2 & 5 \\ 0 & 11 & 13 \end{bmatrix} =: A_1$$

Call the equivalent matrix A_1.
Now ... lets do exactly the same row operation to I_2 (the 2x2 identity matrix).

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \overset{2R_1 + R_2 \to R_2}{\sim} \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} =: E.$$

Call the equivalent matrix E.

(*) (*)

And now notice what happens if we multiply A by E on the left...

this leaves the top row untouched!

$$EA = \begin{bmatrix} \boxed{1 \quad 0} \\ \boxed{2 \quad 1} \end{bmatrix} \begin{bmatrix} 1 & 2 & 5 \\ -2 & 7 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 5 \\ 0 & 11 & 13 \end{bmatrix} = A_1 \quad \ldots \quad (!)$$

notice how this is taking 2 times the top row and adding it to the bottom! $(2R_1 + R_2)$

So, for this row operation, we can find a matrix E that when we multiply A on the left by E, it does the row operation! This means that this row operation was actually a linear transformation. (In fact, you can think of it as a linear transformation on each of the column vectors of A ...)

It turns out that the procedure we just followed works for any elementary row operation, so all elementary row operations can be represented by a matrix, which represents a linear transformation. i.e. all row operations are linear transformations.

DEF (Elementary matrix)

If we perform a single row operation on an Identity matrix I_n, the resulting matrix is called an **elementary matrix**.

On the last page, the matrix E is an elementary matrix.

EXAMPLE 5

Find a single matrix that you can use to multiply the matrix $A = \begin{bmatrix} 2 & 1 & -1 \\ 1 & -1 & 5 \\ -2 & 2 & -10 \end{bmatrix}$ on the left such that the result is in echelon form. (NOTE: Many possible answers since echelon form is not unique!)

First, notice

$$\begin{bmatrix} 2 & 1 & -1 \\ 1 & -1 & 5 \\ -2 & 2 & -10 \end{bmatrix} \overset{R_1 \leftrightarrow R_2}{\sim} \begin{bmatrix} 1 & -1 & 5 \\ 2 & 1 & -1 \\ -2 & 2 & -10 \end{bmatrix}$$

$$\overset{-2R_1 + R_2 \to R_2}{\sim} \begin{bmatrix} 1 & -1 & 5 \\ 0 & 3 & -11 \\ -2 & 2 & -10 \end{bmatrix}$$

$$\overset{2R_1 + R_3 \to R_3}{\sim} \begin{bmatrix} 1 & -1 & 5 \\ 0 & 3 & -11 \\ 0 & 0 & 0 \end{bmatrix} \quad \leftarrow \text{echelon form!}$$

Now, lets find the elementary matrix corresponding to each elementary row operation:

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \overset{R_1 \leftrightarrow R_2}{\sim} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} =: E_1$$

first row operation

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \overset{-2R_1 + R_2 \to R_2}{\sim} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} =: E_2$$

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \overset{2R_1 + R_3 \to R_3}{\sim} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} =: E_3$$

Now, notice: (check!)

$$E_1 A = \begin{bmatrix} 1 & -1 & 5 \\ 2 & 1 & -1 \\ -2 & 2 & -10 \end{bmatrix}$$

$$E_2 (E_1 A) = \begin{bmatrix} 1 & -1 & 5 \\ 0 & 3 & -11 \\ -2 & 2 & -10 \end{bmatrix}$$

$$E_3 (E_2 E_1 A) = \begin{bmatrix} 1 & -1 & 5 \\ 0 & 3 & -11 \\ 0 & 0 & 0 \end{bmatrix}$$

So, the matrix $E_3 E_2 E_1 A$ is in echelon form. This means $E_3 E_2 E_1$ is a matrix that answers the question:

$$E_3 E_2 E_1 = \overset{(E_3)}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}} \overset{(E_2)}{\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \overset{(E_1)}{\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}}$$

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

$$\boxed{= \begin{bmatrix} 0 & 1 & 0 \\ 1 & -2 & 0 \\ 0 & 2 & 1 \end{bmatrix}.} \Leftarrow$$

Notice! (check)

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -2 & 0 \\ 0 & 2 & 1 \end{bmatrix} \overset{(A)}{\begin{bmatrix} 2 & 1 & -1 \\ 1 & -1 & 5 \\ -2 & 2 & -10 \end{bmatrix}} = \begin{bmatrix} 1 & -1 & 5 \\ 0 & 3 & -11 \\ 0 & 0 & 0 \end{bmatrix}.$$

EXERCISE

If the previous example asked for the matrix needed to put A in reduced-row echelon form, what would it be? Would it be unique?

EXERCISE

Recall from M126 that $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$. (This is the Taylor series for the function e^x.) We can use this to define the exponential of a square matrix, $A_{n\times n}$:

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!},$$

where we use the convention that $A^0 = I_n$.

Compute

(a) $e^{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}}$

(b) $e^{\begin{bmatrix} 2 & -1 \\ 4 & -2 \end{bmatrix}}$

(c) $e^{\begin{bmatrix} 7 & 0 \\ 0 & 9 \end{bmatrix}}$ (Hint: You can sum in the matrix without worrying about convergence! At least in this class you can...)