Math 313 Lecture #9
§2.1: Matrix Operations

Entry-Wise Notation for Matrices. For an $m \times n$ matrix A we let a_{ij} denote the entry of A in the i^{th} row and the j^{th} column, and we call this the (i, j) entry of A.

Notational we write $A = [a_{ij}]$ (the text's notation) or $A = (a_{ij})$ (notation used in these lecture notes), so that

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{n1} \\ a_{21} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

For A written in terms of its columns $\begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \end{bmatrix}$ we have that

$$\vec{a}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \ \vec{a}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \dots, \ \vec{a}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

The **zero matrix** 0 is the matrix all of whose entries are zero.

The **diagonal entries** of A are the entries a_{11}, a_{22}, etc., and they form the **main diagonal** of the matrix.

A **diagonal matrix** is an $n \times n$ matrix $(m = n)$ where all of the nondiagonal entries are zero, e.g., the identity matrix.

The Beginnings of Matrix Algebra. Two matrices $A = (a_{ij})$ and $B = (b_{ij})$ are **equal** if they are of the same size (i.e., the same number of columns and rows), *and* $a_{ij} = b_{ij}$ for all i and j (i.e., entry-wise equality).

For $A = (a_{ij})$ an $m \times n$ matrix and α a scalar, the **scalar multiple** αA is the $m \times n$ matrix with entries (αa_{ij}). [This is multiplication done entry-wise.]

For two matrices $A = (a_{ij})$ and $B = (b_{ij})$ of the same size, their **sum** $A+B$ is the matrix whose entries are $(a_{ij} + b_{ij})$. [This addition is entry-wise.]

Theorem 1 (The Rules of Sums and Scalar Multiples). Let α and β be scalars, and A, B, and C matrices of the same size. Then

a. $A + B = B + A$ (matrix addition is commutative),
b. $(A + B) + C = A + (B + C)$ (matrix addition is associative),
c. $A + 0 = A$ (the zero matrix is an additive identity),
d. $\alpha(A + B) = \alpha A + \alpha B$, (scalar multiplication distributes over matrix addition)
e. $(\alpha + \beta)A = \alpha A + \beta A$,
f. $(\alpha\beta)A = \alpha(\beta A)$,
g. $1A = A$ (the scalar 1 is multiplicative identity for scalar multiplication).

Each of these rules is established by the showing that the i^{th} column of the left-hand side equals the i^{th} column of the right-hand side, or by showing the (i, j) entries of both sides are the same.

Matrix Multiplication.

Multiplying a vector $\vec{x}$ by a matrix B results in a vector $B\vec{x}$.

If we then multiply the vector $B\vec{x}$ by a matrix A, we get another vector $A(B\vec{x})$.

This is the *composition* of two matrix transformations and it is given by multiplying $\vec{x}$ by the matrix AB defined by

$$(AB)\vec{x} = A(B\vec{x}).$$

We will find a "formula" for the "product" AB.

For $A(B\vec{x})$ to make sense requires that number of columns of A equals the number of entries in $B\vec{x}$, i.e., the number of rows of B.

Let A be an $m \times n$ matrix and B an $n \times p$ matrix, and write $B = \begin{bmatrix} \vec{b}_1 & \vec{b}_2 & \cdots & \vec{b}_p \end{bmatrix}$.

Then with $\vec{x}$ in $\mathbb{R}^p$, we have $B\vec{x} = x_1\vec{b}_1 + x_2\vec{b}_2 + \cdots + x_p\vec{b}_p$, and so by the linearity of the matrix transformation $\vec{u} \mapsto A\vec{u}$ we have

$$\begin{aligned} A(B\vec{x}) &= A(x_1\vec{b}_1 + x_2\vec{b}_2 + \cdots + x_p\vec{b}_p) \\ &= x_1A\vec{b}_1 + x_2A\vec{b}_2 + \cdots + x_pA\vec{b}_p \\ &= \begin{bmatrix} A\vec{b}_1 & A\vec{b}_2 & \cdots & A\vec{b}_p \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}. \end{aligned}$$

This implies through the definition $(AB)\vec{x} = A(B\vec{x})$ that

$$AB = \begin{bmatrix} A\vec{b}_1 & A\vec{b}_2 & \cdots & A\vec{b}_p \end{bmatrix}.$$

This says that the i^{th} column of AB is the linear combination of the columns of A with the entries of the vector $\vec{b}_i$ as the weights.

For $A = (a_{ij})$ an $m \times n$ matrix and $B = (b_{ij})$ an $n \times p$ matrix, the (i, j) entry of the $m \times p$ matrix $C = AB$ is

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

This is the **row-column** pattern for multiplication of two matrices.

Example.

Let

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

Then

$$AB = \begin{bmatrix} 0 & -1 & -2 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 2 & 3 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 0(2)-1(2)-2(3) & 0(1)-1(3)-2(4) \\ 1(2)-1(2)+2(3) & 1(1)-1(3)+2(4) \end{bmatrix} = \begin{bmatrix} -8 & -11 \\ 6 & 6 \end{bmatrix}.$$

Observe that the first row of A as a matrix times B gives the first row of AB.

If we let $\mathrm{row}_i(D)$ denote the i^{th} row of a matrix D, then we have

$$\mathrm{row}_i(AB) = \mathrm{row}_i(A)B.$$

Theorem 2 (Properties of Matrix Multiplication). Let A be an $m \times n$ matrix, and let B and C be matrices for which the indicated sums and products are defined.

a. $A(BC) = (AB)C$ (matrix multiplication is associative),
b. $A(B+C) = AB + AC$ (left distribution),
c. $(B+C)A = BA + CA$ (right distribution),
d. $\alpha(AB) = (\alpha A)B = A(\alpha B)$ for any scalar α,
e. $I_m A = A = AI_n$ (identity for matrix multiplication).

Proof. Properties (b), (c), (d), and (e) are left for you to verify.

To verify property (a), we write $C = \begin{bmatrix} \vec{c}_1 & \vec{c}_2 & \cdots & \vec{c}_p \end{bmatrix}$.

Then

$$A(BC) = A\begin{bmatrix} B\vec{c}_1 & B\vec{c}_2 & \cdots & B\vec{c}_p \end{bmatrix} = \begin{bmatrix} A(B\vec{c}_1) & A(B\vec{c}_2) & \cdots A(B\vec{c}_p) \end{bmatrix}.$$

The definition of matrix multiplication is that $(AB)\vec{x} = A(B\vec{x})$, and so

$$\begin{bmatrix} A(B\vec{c}_1) & A(B\vec{c}_2) & \cdots A(B\vec{c}_p) \end{bmatrix} = \begin{bmatrix} (AB)\vec{c}_1 & (AB)\vec{c}_2 & \cdots & (AB)\vec{c}_p \end{bmatrix} = (AB)C.$$

This establishes the associativity of matrix multiplication. □

Notice that we did NOT put $AB = BA$ on this list of properties. Why?

If A is a 3×4 matrix and B is a 4×7 matrix then AB is a 3×7 matrix, but does BA make any sense? No, because the number of columns of B does not equal the number of rows of A.

For there to be a product BA requires that A be $m \times n$ and B be $n \times m$.

But then AB is an $m \times m$ matrix while BA is an $n \times n$ matrix.

So to compare AB with BA requires that $m = n$, that is, that A and B must be **square** matrices of the same size.

Okay then, if A and B are $n \times n$ matrices, is it always true that $AB = BA$?

Well, consider $A = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. Then

$$AB = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix} \quad \text{and} \quad BA = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}.$$

It is not always true that $AB = BA$:

$\boxed{\text{Matrix Multiplication is NOT Commutative.}}$

Does the cancellation law hold for matrix multiplication? That is, if $AB = AC$ then $B = C$?

You have a HW problem (#10) that shows the cancellation law does not always hold.

If $AB = 0$, then does $A = 0$ or $B = 0$?

You have a HW problem (#12) that shows it is possible for $AB = 0$ with $A \neq 0$ and $B \neq 0$.

The Transpose of a Matrix. The **transpose** of an $m \times n$ matrix A is the matrix A^T whose columns are the rows of A, i.e., the (i, j) entry of A^T is the (j, i) entry of A.

Example. If $A = \begin{bmatrix} 2 & 1 \\ 2 & 3 \\ 3 & 4 \end{bmatrix}$, then $A^T = \begin{bmatrix} 2 & 2 & 3 \\ 1 & 3 & 4 \end{bmatrix}$.

Theorem 3 (Properties of the Transpose). Let A and B be matrices whose sizes are appropriate for the indicated sums and products.

a. $(A^T)^T = A$.
b. $(A + B)^T = A^T + B^T$.
c. $(\alpha A)^T = \alpha A^T$ for any scalar α.
d. $(AB)^T = B^T A^T$.

The proof of the first three properties is "straight-forward."

The last property of the transpose may look wrong, but think about the sizes of the matrices involved.

The proof of property (d) is in the appendix of this lecture note.

Appendix. Proof of property (d) of the transpose.

Let $A = (a_{ij})$ be an $m \times n$ matrix and $B = (b_{ij})$ be an $n \times p$ matrix.

We need to show that the (i, j) entry of $(AB)^T$ is the same as that of $B^T A^T$.

The (i, j) entry of $(AB)^T$ is the (j, i) entry of AB which is the product of the j^{th} row of A with the i^{th} column of B.

The (i, j) entry of $B^T A^T$ is the product of the i^{th} row of B^T with the j^{th} column of A.

Thus the (i, j) entry of $B^T A^T$ is product of the i^{th} column of B with the j^{th} row of A.

But this is the same as the (i, j) entry of $(AB)^T$. So $B^T A^T = (AB)^T$. $\square$