Lecture 7: Matrix Operations

1 Matrix Notation

An $m \times n$ *matrix* is an array of numbers with m rows and n columns.

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix}$$

or more compactly

$$A = [a_{ij}]_{m \times n} \quad \text{or} \quad A = [a_{ij}]$$

Sometimes we also use $(A)_{ij}$ to mean the same as a_{ij}, for example if

$$B = \begin{pmatrix} 1 & -4 \\ 3 & 5 \end{pmatrix}$$

then $(B)_{21} =$

A matrix with only one row ($m = 1$) is a row vector

$$\left(a_{1,1} \; a_{1,2} \; \ldots \; a_{1,n}\right)$$

and a matrix with only one column ($n = 1$) is a column vector

$$\begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{m,1} \end{bmatrix}$$

and a matrix with the same number of rows and columns ($m = n$) is known as a **square matrix**.

For example, B above is a 2×2 square matrix

2 Equality of Matrices

We say that two matrices A and B are **equal** if they have the same number of rows, the same number of columns, and their respective entries are the same

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 & 0 \\ 3 & 4 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 \\ 3 & 5 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

3 Sum of Matrices

If A and B are both $m \times n$ matrices, we can add them by adding their elements component-wise, e.g.,

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

You can not add two matrices unless they have both the same numbers of rows and the same number of columns!

4 Negation of Matrices

The negative of an $m \times n$ matrix is the matrix of the same size whose entries have all been negated

$$-\begin{pmatrix} 1 & -2 \\ 4 & -5 \end{pmatrix} =$$

5 Difference of Matrices

If A and B are both $m \times n$ matrices, we can subtract one from the other them by subtracting their elements component-wise, e.g.,

$$\begin{pmatrix} 1 & -3 \\ 7 & 8 \end{pmatrix} - \begin{pmatrix} 1 & -2 \\ 0 & 6 \end{pmatrix} =$$

You can not subtract one matrix from another unless they have both the same numbers of rows and the same number of columns!

6 Multiplication of a Matrix by a Scalar

Just multiply each component by the scalar. The resulting matrix is the same size as the original

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

7 Rows of a Matrix

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix}$$

We write $\mathbf{r}_1(A)$ to mean the first row of A

$$\mathbf{r}_1(A) = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \end{pmatrix}$$

and more generally $\mathbf{r}_i(A)$ means the ith row

$$\mathbf{r}_i(A) = \begin{pmatrix} \quad \end{pmatrix}$$

For example, if

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

then

$\mathbf{r}_2(A) =$

So we can think of a matrix as being a stack of rows

$$A = \begin{pmatrix} \text{———} \mathbf{r}_1(A) \text{———} \\ \text{———} \mathbf{r}_2(A) \text{———} \\ \vdots \\ \text{———} \mathbf{r}_m(A) \text{———} \end{pmatrix}$$

8 Columns of a Matrix

We write $\mathbf{c}_1(A)$ to mean the first column of A,

$$\mathbf{c}_1(A) = \begin{pmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{m,1} \end{pmatrix}$$

and more generally $\mathbf{c}_j(A)$ means the jth column

$$\mathbf{c}_j(A) = \begin{pmatrix} \\ \\ \vdots \\ \\ \end{pmatrix}$$

For example, if

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

then

$\mathbf{c}_1(A) =$

And we can see the matrix as being built from a bunch of columns

$$A = \left[\qquad \qquad \qquad \qquad \right]$$

9 Product of a Matrix and a Column Vector

Recall that we write the linear system

$$
\begin{aligned}
a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &= b_1 \\
a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n &= b_2 \\
&\vdots \\
a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n &= b_m,
\end{aligned}
$$

as the augmented matrix

$$
\left(\begin{array}{cccc|c}
a_{1,1} & a_{1,2} & \cdots & a_{1,n} & b_1 \\
a_{2,1} & a_{2,2} & \cdots & a_{2,n} & b_2 \\
\vdots & \vdots & & \vdots & \vdots \\
a_{m,1} & a_{m,2} & \cdots & a_{m,n} & b_m
\end{array}\right)
$$

Note that if we define

$$
A = \begin{pmatrix}
a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\
a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\
\vdots & \vdots & & \vdots \\
a_{m,1} & a_{m,2} & \cdots & a_{m,n}
\end{pmatrix},
$$

and define

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad \text{and} \qquad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

Then the equations

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &= b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n &= b_2 \\ &\vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n &= b_m, \end{aligned}$$

can be written as

which we now write as

$$\begin{pmatrix} \text{———————} \mathbf{r}_1(A) \text{———————} \\ \text{———————} \mathbf{r}_2(A) \text{———————} \\ \vdots \\ \text{———————} \mathbf{r}_m(A) \text{———————} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

or, since the stack of the rows of A is just A

$$A\mathbf{x} = \mathbf{b}$$

In fact, we will now make a definition of the **product of a matrix and a column vector** that makes this notation mean what we want.

If A is the $m \times n$ matrix

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix},$$

and $\mathbf{x}$ is the $n \times 1$ column vector

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

we define $A\mathbf{x}$ to be the $m \times 1$ column vector

$$A\mathbf{x} = \begin{pmatrix} \mathbf{r}_1(A) \cdot \mathbf{x} \\ \mathbf{r}_2(A) \cdot \mathbf{x} \\ \vdots \\ \mathbf{r}_m(A) \cdot \mathbf{x} \end{pmatrix}$$

Note that the sizes of the matrix and the column vector must be compatible, otherwise we can not compute these dot products

Example: $\begin{pmatrix} 1 & 2 \\ 0 & -2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 11 \\ -3 \end{pmatrix} =$

We can also see $A\mathbf{x}$ as a linear combination of the columns of A (where the coefficient applied to the jth column is x_j)

$$A\mathbf{x} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$= \begin{pmatrix} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n \\ \vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n \end{pmatrix}$$

If A is the $m \times n$ matrix

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix},$$

and B is the $n \times p$ matrix

$$B = \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,p} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,p} \\ \vdots & \vdots & & \vdots \\ b_{n,1} & b_{n,2} & \cdots & b_{n,p} \end{pmatrix},$$

then AB is the $m \times p$ matrix

$$AB = \begin{pmatrix} \mathbf{r}_1(A)\cdot\mathbf{c}_1(B) & \mathbf{r}_1(A)\cdot\mathbf{c}_2(B) & \cdots & \mathbf{r}_1(A)\cdot\mathbf{c}_p(B) \\ \mathbf{r}_2(A)\cdot\mathbf{c}_1(B) & \mathbf{r}_2(A)\cdot\mathbf{c}_2(B) & \cdots & \mathbf{r}_2(A)\cdot\mathbf{c}_p(B) \\ \vdots & \vdots & & \vdots \\ \mathbf{r}_m(A)\cdot\mathbf{c}_1(B) & \mathbf{r}_m(A)\cdot\mathbf{c}_2(B) & \cdots & \mathbf{r}_m(A)\cdot\mathbf{c}_p(B) \end{pmatrix},$$

That is the ij entry of AB is the dot product of the ith row of A with the j column of B:

$$(AB)_{ij} = \mathbf{r}_i(A)\cdot\mathbf{c}_j(B)$$

Example: $\begin{pmatrix} 1 & 2 \\ 0 & -2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 11 & 0 \\ -3 & 2 \end{pmatrix} =$

Note that there is a size compatibility requirement for matrix multiplication:

If we consider B as a bunch of columns

$$B = \left[\begin{array}{c} \\ \\ \\ \end{array} \right]$$

then note that the kth column AB is just $A\mathbf{c}_k(B)$

We saw on slide 13 that if $\mathbf{x}$ is a column vector, then $A\mathbf{x}$ is a linear combination of columns of A

If we think of $\mathbf{x} = \mathbf{c}_k(B)$, the kth column of AB is a linear combination of columns of A

$$\begin{aligned}\mathbf{c}_k(AB) &= A\mathbf{c}_k(B)\\ &= \mathbf{c}_1(A)b_{1k} + \mathbf{c}_2(A)b_{2k} + \cdots + \mathbf{c}_n(A)b_{nk}\end{aligned}$$

It is also true that each row of AB is a linear combination of rows of B:

$$\begin{aligned}\mathbf{r}_i(AB) &= \mathbf{r}_i(A)B\\ &= a_{i1}\mathbf{r}_1(B) + a_{i2}\mathbf{r}_2(B) + \cdots + a_{in}\mathbf{r}_n(B)\end{aligned}$$

11 Some algebraic rules for matrix Multiplication

If A is an $m \times n$ matrix and $\mathbf{x}$ and $\mathbf{y}$ are $n \times 1$ column vectors and c is a scalar, then

- $A(c\mathbf{x}) = cA\mathbf{x}$
- $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y}$

If A is the $m \times n$ matrix

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix},$$

then the **transpose of** A, is the $n \times m$ matrix we get from A by exchanging row indices with column indices: i.e., $(A^T)_{ji} = (A)_{ij}$

$$A^T = \begin{pmatrix} a_{1,1} & a_{2,1} & \cdots & a_{m,1} \\ a_{1,2} & a_{2,2} & \cdots & a_{m,2} \\ \vdots & \vdots & & \vdots \\ a_{1,n} & a_{2,n} & \cdots & a_{m,n} \end{pmatrix},$$

So the jth row of A^T is obtained from jth column of A

and the ith column of A^T is obtained from the ith row of A

Example: $\begin{pmatrix} 1 & -4 & 3 \\ 4 & 0 & 1 \end{pmatrix}^T =$

If we have a comma-delimited vector, say $\mathbf{x} = (1, 2, 3)$, and we want to use it in a matrix formula, we generally think of it as being a column vector by default. That is, if we write $A\mathbf{x}$, we are thinking of $\mathbf{x}$ as

$$\mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

If we want to take the inner product with $\mathbf{y} = (-1, 0, 3)$, we can accomplish this using the transpose.

$$\begin{aligned} \mathbf{x} \cdot \mathbf{y} &= \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \\ 3 \end{pmatrix} \\ &= (1)(-1) + (2)(0) + (3)(3) \\ &= 8 \\ &= \mathbf{x}^T \mathbf{y} \end{aligned}$$

Or, equally well,

$$\begin{aligned} \mathbf{y} \cdot \mathbf{x} &= \begin{pmatrix} -1 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \\ &= \mathbf{y}^T \mathbf{x} \end{aligned}$$

If, on the other hand, you try to calculate

$$\mathbf{x}\mathbf{y}^T = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \begin{pmatrix} -1 & 0 & 3 \end{pmatrix}$$

More generally if $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$, then $\mathbf{x}\mathbf{y}^T =$

This is the **outer product of x and y**

13 Trace of a Square Matrix

If A is the $n \times n$ square matrix

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix},$$

then the **trace of** A, written $\operatorname{tr}(A)$ is defined to be the sum of the elements on the diagonal

$$\operatorname{tr}(A) = a_{1,1} + a_{2,2} + \cdots + a_{n,n}$$

Example: $\operatorname{tr}\begin{pmatrix} 1 & 0 \\ -2 & -3 \end{pmatrix} =$

Look at the outer product $\mathbf{x}\mathbf{y}^T$ on the previous slide, and note that

$$\begin{aligned} \operatorname{tr}(\mathbf{x}\mathbf{y}^T) &= x_1 y_1 + \cdots + x_n y_n \\ &= \mathbf{x} \cdot \mathbf{y} \\ &= \mathbf{x}^T \mathbf{y}. \end{aligned}$$