MATH 311-504

Topics in Applied Mathematics

Lecture 8:
Matrix algebra (continued).

Matrices

Definition. An **m-by-n matrix** is a rectangular array of numbers that has m rows and n columns:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Notation: $A = (a_{ij})_{1\le i\le n,\, 1\le j\le m}$ or simply $A = (a_{ij})$ if the dimensions are known.

Matrix addition

Definition. Let $A = (a_{ij})$ and $B = (b_{ij})$ be $m \times n$ matrices. The **sum** $A + B$ is defined to be the $m \times n$ matrix $C = (c_{ij})$ such that $\boxed{c_{ij} = a_{ij} + b_{ij}}$ for all indices i, j.

That is, two matrices with the same dimensions can be added by adding their corresponding entries.

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \\ a_{31}+b_{31} & a_{32}+b_{32} \end{pmatrix}$$

Scalar multiplication

Definition. Given an $m \times n$ matrix $A = (a_{ij})$ and a number r, the **scalar multiple** rA is defined to be the $m \times n$ matrix $D = (d_{ij})$ such that $\boxed{d_{ij} = ra_{ij}}$ for all indices i, j.

That is, to multiply a matrix by a scalar r, one multiplies each entry of the matrix by r.

$$r \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} ra_{11} & ra_{12} & ra_{13} \\ ra_{21} & ra_{22} & ra_{23} \\ ra_{31} & ra_{32} & ra_{33} \end{pmatrix}$$

The $m \times n$ **zero matrix** (all entries are zeros) is denoted O_{mn} or simply O.

Negative of a matrix: $-A$ is defined as $(-1)A$.

Matrix **difference**: $A - B$ is defined as $A + (-B)$.

As far as the *linear operations* (addition and scalar multiplication) are concerned, the $m \times n$ matrices can be regarded as mn-dimensional vectors.

Matrix multiplication

The product of matrices A and B is defined if the number of columns in A matches the number of rows in B.

Definition. Let $A = (a_{ik})$ be an $m \times n$ matrix and $B = (b_{kj})$ be an $n \times p$ matrix. The **product** AB is defined to be the $m \times p$ matrix $C = (c_{ij})$ such that $\boxed{c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}}$ for all indices i, j.

That is, matrices are multiplied **row by column**:

$$\begin{pmatrix} * & * & * \\ \boxed{*} & \boxed{*} & \boxed{*} \end{pmatrix} \begin{pmatrix} * & * & \boxed{*} & * \\ * & * & \boxed{*} & * \\ * & * & \boxed{*} & * \end{pmatrix} = \begin{pmatrix} * & * & * & * \\ * & * & \boxed{*} & * \end{pmatrix}$$

$$A = \left(\begin{array}{cccc} a_{11} & a_{12} & \dots & a_{1n} \\ \hline a_{21} & a_{22} & \dots & a_{2n} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline a_{m1} & a_{m2} & \dots & a_{mn} \end{array}\right) = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_m \end{pmatrix}$$

$$B = \left(\begin{array}{c|c|c|c} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{np} \end{array}\right) = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_p)$$

$$\implies AB = \begin{pmatrix} \mathbf{v}_1 \cdot \mathbf{w}_1 & \mathbf{v}_1 \cdot \mathbf{w}_2 & \dots & \mathbf{v}_1 \cdot \mathbf{w}_p \\ \mathbf{v}_2 \cdot \mathbf{w}_1 & \mathbf{v}_2 \cdot \mathbf{w}_2 & \dots & \mathbf{v}_2 \cdot \mathbf{w}_p \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{v}_m \cdot \mathbf{w}_1 & \mathbf{v}_m \cdot \mathbf{w}_2 & \dots & \mathbf{v}_m \cdot \mathbf{w}_p \end{pmatrix}$$

Any system of linear equations can be represented as a matrix equation:

$$\left\{\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array}\right. \iff A\mathbf{x} = \mathbf{b},$$

where

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Properties of matrix multiplication:

$(AB)C = A(BC)$ (associative law)

$(A + B)C = AC + BC$ (distributive law #1)

$C(A + B) = CA + CB$ (distributive law #2)

$(rA)B = A(rB) = r(AB)$

Any of the above identities holds provided that matrix sums and products are well defined.

If A and B are $n\times n$ matrices, then both AB and BA are well defined $n\times n$ matrices.

However, in general, $AB \neq BA$.

Example. Let $A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Then $AB = \begin{pmatrix} 2 & 2 \\ 0 & 1 \end{pmatrix}$, $BA = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$.

If AB does equal BA, we say that the matrices A and B **commute**.

Problem. Let A and B be arbitrary $n \times n$ matrices. Is it true that $(A-B)(A+B) = A^2 - B^2$?

$$\begin{aligned}(A-B)(A+B) &= (A-B)A + (A-B)B \\ &= (AA - BA) + (AB - BB) \\ &= A^2 + AB - BA - B^2\end{aligned}$$

Hence $(A-B)(A+B) = A^2 - B^2$ if and only if A commutes with B.

Diagonal matrices

If $A = (a_{ij})$ is a square matrix, then the entries a_{ii} are called **diagonal entries**. A square matrix is called **diagonal** if all non-diagonal entries are zeros.

Example. $\begin{pmatrix} 7 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, denoted $\mathrm{diag}(7, 1, 2)$.

Let $A = \mathrm{diag}(s_1, s_2, \dots, s_n)$, $B = \mathrm{diag}(t_1, t_2, \dots, t_n)$.

Then $A + B = \mathrm{diag}(s_1 + t_1, s_2 + t_2, \dots, s_n + t_n)$,

$$rA = \mathrm{diag}(rs_1, rs_2, \dots, rs_n).$$

Example.

$$\begin{pmatrix} 7 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} -7 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 6 \end{pmatrix}$$

Theorem Let $A = \mathrm{diag}(s_1, s_2, \dots, s_n)$,
$B = \mathrm{diag}(t_1, t_2, \dots, t_n)$.

Then $A + B = \mathrm{diag}(s_1 + t_1, s_2 + t_2, \dots, s_n + t_n)$,

$$rA = \mathrm{diag}(rs_1, rs_2, \dots, rs_n).$$

$$AB = \mathrm{diag}(s_1 t_1, s_2 t_2, \dots, s_n t_n).$$

In particular, diagonal matrices always commute.

Example.

$$\begin{pmatrix} 7 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 7a_{11} & 7a_{12} & 7a_{13} \\ a_{21} & a_{22} & a_{23} \\ 2a_{31} & 2a_{32} & 2a_{33} \end{pmatrix}$$

Theorem Let $D = \mathrm{diag}(d_1, d_2, \ldots, d_m)$ and A be an $m \times n$ matrix. Then the matrix DA is obtained from A by multiplying the ith row by d_i for $i = 1, 2, \ldots, m$:

$$A = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_m \end{pmatrix} \implies DA = \begin{pmatrix} d_1\mathbf{v}_1 \\ d_2\mathbf{v}_2 \\ \vdots \\ d_m\mathbf{v}_m \end{pmatrix}$$

Example.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} 7 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 7a_{11} & a_{12} & 2a_{13} \\ 7a_{21} & a_{22} & 2a_{23} \\ 7a_{31} & a_{32} & 2a_{33} \end{pmatrix}$$

Theorem Let $D = \mathrm{diag}(d_1, d_2, \ldots, d_n)$ and A be an $m \times n$ matrix. Then the matrix AD is obtained from A by multiplying the ith column by d_i for $i = 1, 2, \ldots, n$:

$$A = (\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n)$$
$$\implies \quad AD = (d_1\mathbf{w}_1, d_2\mathbf{w}_2, \ldots, d_n\mathbf{w}_n)$$

Identity matrix

Definition. The **identity matrix** (or **unit matrix**) is a diagonal matrix with all diagonal entries equal to 1. The $n \times n$ identity matrix is denoted I_n or simply I.

$$I_1 = (1), \quad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

In general, $I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$.

Theorem. Let A be an arbitrary $m \times n$ matrix. Then $I_m A = AI_n = A$.

Matrix polynomials

If B is not a square matrix then BB is not defined.

Definition. Given an n-by-n matrix A, let

$$A^2 = AA, \quad A^3 = AAA, \quad \dots, \quad A^k = \underbrace{AA\dots A}_{k \text{ times}}, \quad \dots$$

Also, let $A^1 = A$ and $A^0 = I_n$.

Associativity of matrix multiplication implies that all powers A^k are well defined and $A^j A^k = A^{j+k}$ for all $j, k \geq 0$. In particular, all powers of A commute.

Definition. For any polynomial

$$p(x) = c_0 x^m + c_1 x^{m-1} + \dots + c_{m-1} x + c_m,$$

let

$$p(A) = c_0 A^m + c_1 A^{m-1} + \dots + c_{m-1} A + c_m I_n.$$

Example. $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$.

$$A^2 = AA = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix},$$

$$A^3 = A^2 A = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 13 & 8 \\ 8 & 5 \end{pmatrix},$$

$$A^4 = A^2 A^2 = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} 34 & 21 \\ 21 & 13 \end{pmatrix}.$$

By the way, $1, 1, 2, 3, 5, 8, 13, 21, 34, \dots$ are famous *Fibonacci numbers* given by $f_1 = f_2 = 1$ and $f_n = f_{n-1} + f_{n-2}$ for $n \geq 3$.

Example. $p(x) = x^2 - 3x + 1, \;\; A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$

$$p(A) = A^2 - 3A + I = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}^2 - 3\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

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

Thus $A^2 - 3A + I = O$.

Properties of matrix polynomials

Suppose A is a square matrix, $p(x), p_1(x), p_2(x)$ are polynomials, and r is a scalar. Then

$$
\begin{aligned}
p(x) = p_1(x)+p_2(x) &\implies p(A) = p_1(A) + p_2(A)\\
p(x) = rp_1(x) &\implies p(A) = rp_1(A)\\
p(x) = p_1(x)p_2(x) &\implies p(A) = p_1(A)p_2(A)\\
p(x) = p_1(p_2(x)) &\implies p(A) = p_1(p_2(A))
\end{aligned}
$$

In particular, matrix polynomials $p_1(A)$ and $p_2(A)$ always commute.

If $A = \mathrm{diag}(s_1, s_2, \ldots, s_n)$ then

$$p(A) = \mathrm{diag}\big(p(s_1), p(s_2), \ldots, p(s_n)\big).$$

Examples.

- $(A - I)(A + I) = A^2 - I$
- $(A + I)^2 = A^2 + 2A + I$
- $(A - I)^2 = A^2 - 2A + I$
- $(A + I)^3 = A^3 + 3A^2 + 3A + I$
- $(A - I)^3 = A^3 - 3A^2 + 3A - I$
- $(A - I)(A^2 + A + I) = A^3 - I$
- $(A + I)(A^2 - A + I) = A^3 + I$

Inverse matrix

Let $\mathcal{M}_n(\mathbb{R})$ denote the set of all $n \times n$ matrices with real entries. We can **add**, **subtract**, and **multiply** elements of $\mathcal{M}_n(\mathbb{R})$. What about **division**?

Definition. Let $A \in \mathcal{M}_n(\mathbb{R})$. Suppose there exists an $n \times n$ matrix B such that

$$AB = BA = I_n.$$

Then the matrix A is called **invertible** and B is called the **inverse** of A (denoted A^{-1}).

$$\boxed{AA^{-1} = A^{-1}A = I}$$