Math 304–504

Linear Algebra

Lecture 5:
Matrix algebra (continued).
Diagonal matrices.
Inverse matrix.

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 $c_{ij} = a_{ij} + b_{ij}$ 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 $d_{ij} = ra_{ij}$ 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.

Properties of linear operations

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

$$A + B = B + A$$

$$A + O = O + A = A$$

$$A + (-A) = (-A) + A = O$$

$$r(sA) = (rs)A$$

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

$$(r + s)A = rA + sA$$

$$1A = A$$

$$0A = O$$

Dot product

Definition. The **dot product** of n-dimensional vectors $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ is a scalar

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = \sum_{k=1}^{n} x_k y_k.$$

The dot product is also called the **scalar product**.

Matrix multiplication

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 $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$ for all indices i, j.

That is, matrices are multiplied **row by column**: c_{ij} is the dot product of the ith row of A and the jth column of B.

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

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

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

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

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

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

where

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

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

Another representation of this system:

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Theorem The above system is consistent if and only if the vector $\mathbf{b}$ is a *linear combination* of the column vectors of A.

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

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

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

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

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

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

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

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 $i = 1, 2, \ldots, n$:

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

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 $I_m A = AI_n = A$.

Transpose of a matrix

Definition. Given a matrix A, the **transpose** of A, denoted A^T, is the matrix whose rows are columns of A (and whose columns are rows of A). That is, if $A = (a_{ij})$ then $A^T = (b_{ij})$, where $b_{ij} = a_{ji}$.

Examples. $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$,

$$\begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix}^T = (7, 8, 9), \qquad \begin{pmatrix} 4 & 7 \\ 7 & 0 \end{pmatrix}^T = \begin{pmatrix} 4 & 7 \\ 7 & 0 \end{pmatrix}.$$

Definition. A square matrix A is said to be **symmetric** if $A^T = A$.

Properties of transposes:

- $(A^T)^T = A$
- $(A + B)^T = A^T + B^T$
- $(rA)^T = rA^T$
- $(AB)^T = B^T A^T$
- $(A_1 A_2 \ldots A_k)^T = A_k^T \ldots A_2^T A_1^T$

Proposition Given any matrix A, the products AA^T and A^TA are well defined symmetric matrices.

Proof: Suppose A is an $m \times n$ matrix. Then A^T is an $n \times m$ matrix. Hence AA^T and A^TA are well defined. AA^T is an $m \times m$ matrix while A^TA is an $n \times n$ matrix.

$$(AA^T)^T = (A^T)^T A^T = AA^T,$$
$$(A^TA)^T = A^T (A^T)^T = A^TA.$$

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}).

A non-invertible square matrix is called **singular**.

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

Basic properties of inverse matrices:

- If $B = A^{-1}$ then $A = B^{-1}$. In other words, if A is invertible, so is A^{-1}, and $A = (A^{-1})^{-1}$.

- If $AB = CA = I$ for some matrices $B, C \in \mathcal{M}_n(\mathbb{R})$ then $B = C = A^{-1}$.

Indeed, $B = IB = (CA)B = C(AB) = CI = C$.

- If matrices $A, B \in \mathcal{M}_n(\mathbb{R})$ are invertible, so is AB, and $(AB)^{-1} = B^{-1}A^{-1}$.

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I,$$
$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I.$$

- Similarly, $(A_1A_2\dots A_k)^{-1} = A_k^{-1}\dots A_2^{-1}A_1^{-1}$.