Matrix Algebra I.

Def$'\underline{n}$: An $(m \times n)$ matrix A is an array of real numbers of the form

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

where a_{ij} denotes the number in row i and column j.

Rks:

- a_{ij} is called the $(i,j)^{th}$ *component* of the matrix A
- We sometimes write $A = (a_{ij})$ or $A \in \mathbb{R}^{m \times n}$
- If $n = 1$, A is called a *column vector*, and we often write $A \in \mathbb{R}^m$ and write $A = (a_i)$
- If $m = 1$, A is called a *row vector*, and we often write $A \in \mathbb{R}^n$ and write $A = (a_j)$
- *vectors* are column vectors, unless otherwise specified
- If $m = n = 1$, A is a *scalar* (i.e., a number), $A \in \mathbb{R}$
- If $m = n$, A is called a *square matrix*.
- We could define matrices using other objects as components, such as complex numbers or functions, but we won't encounter them in this course. We will, however, consider in Section 5b and later in the course a straightforward generalization of our notion of matrix to one with components that are themselves matrices (i.e. a *partitioned matrix*).

Operations on Vectors and Matrices

1. Equality of matrices

Matrices $A = (a_{ij})$ and $B = (b_{ij})$ are said to be equal, denoted $A = B$, if
(i) they have the same number of rows and columns
(ii) $\forall i,j \;\; a_{ij} = b_{ij}$

Rk:

- $A = B \Leftrightarrow B = A$ (equality is a symmetric relationship)
- If A and B are not equal, we write $A \neq B$.

2. Addition of matrices

Suppose $A = (a_{ij})$ and $B = (b_{ij})$ have the same number of rows and columns, i.e. $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{m \times n}$. The sum $A + B \in \mathbb{R}^{m \times n}$ is the matrix $A + B = (a_{ij} + b_{ij})$.

Example:

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \quad B = \begin{bmatrix} 2 & 1 & 2 & 1 \\ 5 & 3 & 4 & 2 \end{bmatrix}$$

$$A + B = \begin{bmatrix} 3 & 3 & 5 & 5 \\ 10 & 9 & 11 & 10 \end{bmatrix}$$

Notation

- An $(m \times n)$ matrix whose components are all zero will be denoted 0_{mn} or, when clear from the context, just 0
- If $A = (a_{ij})$, then $-A = (-a_{ij})$

Properties of matrix sums

$$A + B = B + A \qquad \text{(commutative)}$$

$$(A + B) + C = A + (B + C) \qquad \text{(associative)}$$

$$A + 0 = A \qquad \text{(zero is the identity element)}$$

$$A + (-A) = 0 \qquad \text{(inverse)}$$

Rks:

- We usually write $A + (-B)$ as $A - B$
- Matrix addition is not defined unless the two matrices are *conformable for addition* (i.e., have the same number of rows and columns)
- We can add and subtract conformable matrices just like real numbers.

3. Multiplication by a scalar

Let $A \in \mathbb{R}^{m \times n}$ and $c \in \mathbb{R}$. We define their product by

$$cA = Ac = (c \cdot a_{ij}) \in \mathbb{R}^{m \times n}$$

Example

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \qquad c = 5$$

$$cA = \begin{bmatrix} 5 & 10 & 15 & 20 \\ 25 & 30 & 35 & 40 \end{bmatrix}$$

Properties of scalar multiplication

Suppose $c, d \in \mathbb{R}$ and $A, B \in \mathbb{R}^{m \times n}$

$$cA + dA = (c + d)A \qquad \text{(distributive property 1)}$$

$$cA + cB = c(A + B) \qquad \text{(distributive property 2)}$$

Rks:

- The $(n \times n)$ matrix with unity on the main diagonal and 0 else is called the *identity matrix*, denoted I_n, *i.e.*

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

- If $A = cI_n$ where I_n is the *identity* matrix, then we say A is a *scalar matrix*.
- The operations of matrix addition and scalar multiplication as we have defined them make the set of $(m \times n)$ matrices a mathematical object known as a *vector space* (over the reals). So, in particular (and reassuringly!) the vectors in $\mathbb{R}^m$ constitute a vector space.

4. Transpose

If $A = (a_{ij}) \in \mathbb{R}^{m \times n}$, then its transpose is $A' = (a_{ji}) \in \mathbb{R}^{n \times m}$

Rks:

- In words, the transpose takes the rows of A and makes them the columns of A'
- If $A = A'$, then A is called *symmetric*. Note that a symmetric matrix must be square. Why?
- We sometimes denote the transponse by $A^\top$

Properties of the transpose operator

$$(A')' = A$$
$$(A + B)' = A' + B'$$
$$(cA)' = cA'$$

5. Products of Matrices

5a Inner products (for vectors)

Let $x, y \in \mathbb{R}^m$. Explicitly, write

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} \qquad y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$

We define their *inner product* as

$$x'y = \sum_{i=1}^{m} x_i y_i$$

Rks:

- Notice that $x'y \in \mathbb{R}$, i.e. the inner product of two vectors is a scalar.
- The notation suggests that we first transpose x into a row vector and them multiply it with y. This is suggestive and consistent with the notion of matrix multiplication below.
- Other notations that you might see for the inner product are $x \circ y$ or $< x,\ y >$
- Other names for the inner product are the *dot product* or *scalar product*. DO NOT confuse the scalar product (i.e. the product of two vectors that is a scalar) with a *scalar multiple* (i.e. the product of a vector by a scalar which yields a vector)!

Example:

$$x = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \qquad y = \begin{bmatrix} 2 \\ 1 \\ -1 \\ 3 \end{bmatrix}$$

so

$$x'y = 1 \cdot 2 + 2 \cdot 1 + 3 \cdot (-1) + 4 \cdot 3$$
$$= 13$$

Properties of the inner product

Let $x, y, z \in \mathbb{R}^m$ and $c \in \mathbb{R}$. Then

$$x'y = y'x \qquad \text{(symmetry)}$$
$$x'(cy) = cx'y \qquad \text{(scalar product commutes with multiplication by scalar)}$$
$$x'(y+z) = x'y + x'z \qquad \text{(distributive law)}$$

5b Matrix product

Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$. It is convenient to write A in terms of its rows and B in terms of its columns

$$A = \begin{bmatrix} a_1' \\ a_2' \\ \vdots \\ a_m' \end{bmatrix} \text{ and } B = \begin{bmatrix} b_1 & b_2 & \cdots & b_p \end{bmatrix}$$

where $a_i' \in \mathbb{R}^{1 \times n}$ denotes the i^{th} row of A and b_j denotes the j^{th} column of B. Then we write the product $AB \in \mathbb{R}^{m \times p}$ as

$$AB \equiv \begin{bmatrix} a_1'b_1 & a_1'b_2 & \cdots & a_1'b_p \\ a_2'b_1 & a_2'b_2 & \cdots & a_2'b_p \\ \vdots & \vdots & \ddots & \vdots \\ a_m'b_1 & a_m'b_2 & \cdots & a_m'b_p \end{bmatrix}$$

that is,

$$(AB)_{ij} = a_i'b_j$$

Rks:

- A and B are said to be *conformable for multiplication* (i.e. the product AB is defined) iff the number of columns of A equals the number of rows of B.
- As we have defined the product, we say A is postmultiplied by B or, equivalently, that B is premultiplied by A.
- We write $AA \equiv A^2$ (Note tha A must be square for this to make sense. Why?). A^m is defined analogously for $m \in \mathbb{N}$.
- Recall

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

- If $A = A^2$, we say that A is *idempotent*
- Notice that $I^2 = I$ and $0^2 = 0$
- If A is square and $\exists B$ such that $AB = I$ then B is called the (multiplicative) *inverse* of A and we write $B = A^{-1}$
- If $A' = A^{-1}$, we say that A is an *orthogonal matrix*.
- Notice that $I' = I^{-1}$.

Properties of matrix multiplication

Let A, B, C denote matrices. Assume $A \in \mathbb{R}^{m \times n}$ and that the other matrices have dimensions that are implicitly determined so that they are conformable for the operations used.

1. $AB \neq BA$ (in general)

2. $A(B+C) = AB + AC$ (distributive law 1)

3. $(A+B)C = AC + BC$ (distributive law 2)

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

5. $(AB)' = B'A'$

6. $0_{rm}A = 0_{rn}; \ \ A0_{ns} = 0_{ms}$

7. $I_m A = AI_n = A$

8. $(A^{-1})' = (A')^{-1}$

9. $(AB)^{-1} = B^{-1}A^{-1}$ *provided both inverses exist*

Computing the matrix inverse: (2×2) case

Suppose

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

- The *determinant* of the matrix A is denoted $\det(A)$ and defined by
$$\det(A) = a_{11}a_{22} - a_{21}a_{12}$$
- If $\det(A) \neq 0$, then $\exists\, A^{-1}$ such that $AA^{-1} = A^{-1}A = I_2$ and it is given by
$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

Example

$$A = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$$

$$\det(A) = 10 - 12 = -2$$

$$A^{-1} = \frac{1}{-2} \begin{bmatrix} 5 & -3 \\ -4 & 2 \end{bmatrix}$$

Rks:

- If $\det(A) \neq 0$, then A is called *invertible* or *nonsingular*
- If $\det(A) = 0$, then A is called non*invertible* or *singular*