Matrix Algebra

Definition 1.1 (Matrix)

A ***matrix*** *is a rectangular array of numbers (real or complex) called the entries or elements of the matrix. A matrix with m rows and n columns is called a $m \times n$* ***matrix****. A $1 \times n$ matrix is called a* ***row vector*** *and a $m \times 1$ matrix is called a* ***column vector****. A 1×1 matrix is called a* ***scalar****.*

- For a $m \times n$ matrix, A we write

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

  where a_{ij} are the entries of A.
- **Convention:** We will use the convention that a vector written $\mathbf{x}$ will be a **column vector**. A **row vector** will be written as the transpose of a column vector; i.e. $\mathbf{x}^T$.

Matrix Algebra

- The **columns** of A are denoted by the column vectors

$$\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$$

  with

$$\mathbf{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}.$$

  We write

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$$

  in order to emphasize the **column structure** of A.

Matrix Algebra

- The **rows** of A are denoted by the row vectors

$$\mathbf{A}_1^T,\ \mathbf{A}_2^T,\ \ldots,\ \mathbf{A}_m^T$$

with

$$\mathbf{A}_i^T = \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix}.$$

We write

$$A = \begin{bmatrix} \mathbf{A}_1^T \\ \mathbf{A}_2^T \\ \vdots \\ \mathbf{A}_m^T \end{bmatrix}$$

in order to emphasize the **row structure** of A.

Matrix Algebra

- We see that the transpose of A is given by

$$A^T = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_n^T \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \cdots & \mathbf{A}_m \end{bmatrix} = \begin{bmatrix} a_{ji} \end{bmatrix}.$$

- In EMTH118/9 and MATH199, the focus was on the entries of a matrix. In this course we will focus more on the structure of a matrix; that is, on the columns and rows of a matrix. The basic matrix operations can be interpreted in terms of the column and row structure.

Matrix Addition

- Let A and B be **both** $m \times n$ matrices. Then

$$
\begin{aligned}
A + B &= \begin{bmatrix} a_{ij} + b_{ij} \end{bmatrix} && \text{entries} \\
&= \begin{bmatrix} \mathbf{a}_1 + \mathbf{b}_1 & \mathbf{a}_2 + \mathbf{b}_2 & \cdots & \mathbf{a}_n + \mathbf{b}_n \end{bmatrix} && \text{columns} \\
&= \begin{bmatrix} \mathbf{A}_1^T + \mathbf{B}_1^T \\ \mathbf{A}_2^T + \mathbf{B}_2^T \\ \vdots \\ \mathbf{A}_m^T + \mathbf{B}_m^T \end{bmatrix} && \text{rows}
\end{aligned}
$$

Scalar Multiplication

- Let A be a $m \times n$ matrix and c be a scalar.

$$
\begin{aligned}
cA &= \begin{bmatrix} ca_{ij} \end{bmatrix} && \text{entries} \\
&= \begin{bmatrix} c\mathbf{a}_1 & c\mathbf{a}_2 & \cdots & c\mathbf{a}_n \end{bmatrix} && \text{columns} \\
&= \begin{bmatrix} c\mathbf{A}_1^T \\ c\mathbf{A}_2^T \\ \vdots \\ c\mathbf{A}_m^T \end{bmatrix} && \text{rows}
\end{aligned}
$$

Matrix Multiplication

- Let A be a $m \times n$ matrix and B be a $n \times r$ matrix. The product $C = AB$ will be a $m \times r$ matrix with

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

In terms of the structure of A and B, the entry c_{ij} is given by

$$\begin{array}{cc} \mathbf{A}_i^T & \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \end{array} \quad \begin{array}{c} \mathbf{b}_j \\ \begin{bmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1r} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nj} & \cdots & b_{nr} \end{bmatrix} \end{array}$$

and so

$$c_{ij} = \mathbf{A}_i^T \mathbf{b}_j.$$

This is the "dot product" of the i^{th} row of A and the j^{th} column of B. For (column) vectors $\mathbf{x}$ and $\mathbf{y}$, the dot product is given by

$$\mathbf{x} \cdot \mathbf{y} \equiv \mathbf{x}^T \mathbf{y} = x_1 y_1 + \cdots + x_n y_n.$$

Matrix Multiplication

- Therefore

$$\begin{aligned} AB &= \begin{bmatrix} \mathbf{A}_i^T \mathbf{b}_j \end{bmatrix} && \text{entries} \\ &= \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_n \end{bmatrix} && \text{columns} \\ &= \begin{bmatrix} \mathbf{A}_1^T B \\ \mathbf{A}_2^T B \\ \vdots \\ \mathbf{A}_m^T B \end{bmatrix} && \text{rows} \end{aligned}$$

Matrix Algebra

- Let $\mathbf{e}_j$ be the *unit* vector

$$\mathbf{e}_j = \begin{array}{cl} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} & \begin{array}{l} \\ \\ \\ \leftarrow j^{\text{th}} \\ \\ \\ \\ \end{array} \end{array} .$$

Thus $\mathbf{e}_i^T$ will be the unit (row) vector $\mathbf{e}_i^T = \begin{array}{c} \begin{array}{ccccccc} & & & i^{\text{th}} & & & \end{array} \\ \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{bmatrix} \end{array}$.

Matrix Algebra

- Note

$$\begin{aligned} A\mathbf{e}_j &= \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \mathbf{e}_j &&= \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \\ &= 0\mathbf{a}_1 + \cdots + 1\mathbf{a}_j + \cdots + 0\mathbf{a}_n &&= \mathbf{a}_j . \end{aligned}$$

Thus $A\mathbf{e}_j$ is the j^{th} column of A.

- In particular

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$
$$= x_1\,\mathbf{a}_1 + x_2\,\mathbf{a}_2 + \cdots + x_n\,\mathbf{a}_n .$$

Therefore $A\mathbf{x}$ is a **linear combination of the columns of A**.

Outer Products

- If replace $\mathbf{x}$ by a matrix B represented by its rows, we have

$$A\,B = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} \mathbf{B}_1^T \\ \mathbf{B}_2^T \\ \vdots \\ \mathbf{B}_n^T \end{bmatrix} = \mathbf{a}_1\,\mathbf{B}_1^T + \mathbf{a}_2\,\mathbf{B}_2^T + \cdots + \mathbf{a}_n\,\mathbf{B}_n^T.$$

  The products $\mathbf{a}_i\,\mathbf{B}_i^T$ are matrices which are called **outer products**. This representation of the product is called the **outer product expansion** of $A\,B$.
- A and B need not be square matrices. If A is $m \times n$ and B is $n \times p$ then the outer products are $m \times p$ matrices.

Outer Products

Example

Evaluate $A\,B$ by the outer product expansion where

$$A = \begin{bmatrix} 1 & 3 & 0 \\ 2 & 1 & 4 \end{bmatrix} \qquad B = \begin{bmatrix} 4 & 6 & -1 \\ 1 & 0 & -1 \\ 2 & 5 & 1 \end{bmatrix}$$

Solution: We have

$$\mathbf{a}_1\,\mathbf{B}_1^T = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 4 & 6 & -1 \end{bmatrix} = \begin{bmatrix} 4 & 6 & -1 \\ 8 & 12 & -2 \end{bmatrix}$$

$$\mathbf{a}_2\,\mathbf{B}_2^T = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 3 & 0 & -3 \\ 1 & 0 & -1 \end{bmatrix}$$

$$\mathbf{a}_3\,\mathbf{B}_3^T = \begin{bmatrix} 0 \\ 4 \end{bmatrix} \begin{bmatrix} 2 & 5 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 8 & 20 & 4 \end{bmatrix}$$

Thus

$$A\,B = \mathbf{a}_1\,\mathbf{B}_1^T + \mathbf{a}_2\,\mathbf{B}_2^T + \mathbf{a}_3\,\mathbf{B}_3^T = \begin{bmatrix} 7 & 6 & -4 \\ 17 & 32 & 1 \end{bmatrix}.$$

Matrix Algebra

- Therefore

$$\mathbf{e}_i^T A = \begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \mathbf{A}_1^T \\ \mathbf{A}_2^T \\ \vdots \\ \mathbf{A}_m^T \end{bmatrix} = \mathbf{A}_i^T$$

  is the i^{th} row of A.
- Clearly

$$\begin{bmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \\ \vdots \\ \mathbf{e}_n^T \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n \end{bmatrix}$$

  is the identity matrix.

Partitioned Matrices

- Matrix multiplication is a costly operation.
- Whilst you might wonder about the use of outer products to compute a matrix product (we will see a use for them later with the spectral theorem and SVD), if a matrix has structure then we may be able to reduce the cost of multiplication by *partitioning* the matrix (for an artificial example, see Poole Example 3.12).
- *Block diagonal matrices* occur frequently in applications. For example

$$A = \begin{bmatrix} 2 & 3 & 5 & 0 & 0 \\ 1 & 1 & 4 & 0 & 0 \\ 7 & 2 & 5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 2 & 2 \end{bmatrix}$$

  is a block diagonal matrix.

Partitioned Matrices

- We partition this matrix

$$A = \left[\begin{array}{ccc|cc} 2 & 3 & 5 & 0 & 0 \\ 1 & 1 & 4 & 0 & 0 \\ 7 & 2 & 5 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 2 & 2 \end{array}\right] = \begin{bmatrix} A_1 & O \\ O & A_2 \end{bmatrix}$$

  where O represents a 3×2 zero matrix in the upper right position and a 2×3 zero matrix in the lower left position.
- In partitioned form A "looks like" a 2×2 matrix (albeit with matrix entries).
- In order to compute $A\,B$ using this structure, we need to partition B into *compatible blocks*; that is blocks that can be multiplied by the appropriate blocks in A.

Partitioned Matrices

- Suppose B is a 5×8 matrix (and so the product will be 5×8 matrix). Since A_1 is 3×3 we must partition

$$B = \left[\begin{array}{cccccccc} * & * & * & * & * & * & * & * \\ * & * & * & * & * & * & * & * \\ * & * & * & * & * & * & * & * \\ \hline * & * & * & * & * & * & * & * \\ * & * & * & * & * & * & * & * \end{array}\right] = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

  which looks like a 2×1 matrix.
- The product becomes

$$A\,B = \begin{bmatrix} A_1 & O \\ O & A_2 \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} A_1\,B_1 + O\,B_2 \\ O\,B_1 + A_2\,B_2 \end{bmatrix} = \begin{bmatrix} A_1\,B_1 \\ A_2\,B_2 \end{bmatrix}.$$

- Rather than multiplying a 5×5 matrix by a 5×8 matrix, we have managed to find the product by multiplying a 3×3 matrix by a 3×8 matrix and a 2×2 matrix by a 2×8 matrix.

Partitioned Matrices

- If B has structure then we can partition it vertically to take advantage of this structure.
- Suppose

$$B = \left[\begin{array}{cccccccc} * & * & * & * & * & 1 & 0 & 0 \\ * & * & * & * & * & 0 & 1 & 0 \\ * & * & * & * & * & 0 & 0 & 1 \\ \hline 1 & 0 & * & * & * & * & * & * \\ 0 & 1 & * & * & * & * & * & * \end{array}\right].$$

- We would then partition B to take advantage of these identity matrices.

$$B = \begin{bmatrix} B_3 & I_3 \\ I_2 & B_4 \end{bmatrix}$$

Partitioned Matrices

- The product now becomes

$$A\,B = \begin{bmatrix} A_1 & O \\ O & A_2 \end{bmatrix} \begin{bmatrix} B_3 & I_3 \\ I_2 & B_4 \end{bmatrix} = \begin{bmatrix} A_1\,B_3 + O\,I_2 & A_1\,I_3 + O\,B_4 \\ O\,B_3 + A_2\,I_2 & O\,I_3 + A_2\,B_4 \end{bmatrix}$$

$$= \begin{bmatrix} A_1\,B_3 & A_1 \\ A_2 & A_2\,B_4 \end{bmatrix}.$$