Chapter 5: Matrices

We have already studied matrices as blocks of coefficients of linear systems. We will now study matrices for their own properties.

Definition

An $m \times n$ **matrix** A is a block of numbers written with m rows and n columns. If $m = n$ we say that A is a **square** matrix. The set of all $m \times n$ matrices is denoted by M_{mn}.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

We use the notation a_{ij} to denote the general element in the matrix A.

Ex:

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 9 & -2 & \frac{1}{2} & 0 \\ 5 & 4 & 3 & 2 \end{pmatrix}$$

is a 3×4 matrix and $a_{23} = \frac{1}{2}$.

In M_{mn} there is a special matrix called the **zero matrix** or $\mathbf{0}$ which has all its entries equal to 0.

Ex:

$$\mathbf{0} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

is the zero matrix in M_{34}.

In M_{nn} (square matrices!) there is another special matrix I_n called the **identity** matrix, consisting of 1's down the main diagonal and zeros everywhere else.

For example

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

is the identity matrix in M_{22} and

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

is the identity matrix in M_{33}.

We sometimes just write I for the identity matrix if n is understood from context.

Definition

We say that two matrices A and B are equal if they have the same size (meaning the same number of rows and columns) and $a_{ij} = b_{ij}$ for each i, j.

We can only add or subtract matrices which have the same size, and we do this by simply adding or subtracting the entries in the same position:

If $A = (a_{ij})$ and $B = (b_{ij})$ are matrices of the same size, then

$$(A + B)_{ij} = a_{ij} + b_{ij}.$$

Ex: Find $A+B$ and $A-B$ if $A = \begin{pmatrix} 1 & 2 & 3 \\ 9 & -2 & 5 \\ -1 & 3 & -5 \end{pmatrix}$ and

$B = \begin{pmatrix} 5 & 3 & 9 \\ -2 & 4 & -6 \\ 8 & -4 & 2 \end{pmatrix}$.

Ex: If $A = \begin{pmatrix} 4 & -1 & 2 \\ 4 & -7 & 3 \\ 1 & 2 & -8 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 2 \\ 9 & -2 \\ -1 & 3 \end{pmatrix}$ then

Matrix addition is commutative and associative:

$$A + B = B + A, \quad (A + B) + C = A + (B + C),$$

(as long as A, B, and C are all the same size.)

As with vectors, (a vector is really just an $n \times 1$ matrix), we can multiply a matrix by a scalar in the obvious way. If A is a matrix and λ a scalar, then $B = \lambda A$ is a matrix with entries $b_{ij} = \lambda a_{ij}$.

For example, if $A = \begin{pmatrix} 1 & 2 & 3 \\ 9 & -2 & 5 \\ -1 & 3 & -5 \end{pmatrix}$ then

$3A =$

Note that for all matrices A, we have

$$A + \mathbf{0} = \mathbf{0} + A = A \text{ and } \; 0A = \mathbf{0}.$$

Let's try to motivate matrix multiplication. Suppose we have 2 houses H_1 and H_2 which order a certain number of bottles of milk, cream and a certain number of newspapers. The milk costs \$ 2 , the cream \$ 3 and the newspaper \$ 1. This information can be represented using matrices as follows:

$$\left(\begin{array}{c|ccc} & M & C & N \\ H_1 & 1 & 2 & 3 \\ H_2 & 2 & 2 & 1 \end{array} \right) \begin{pmatrix} \$ \\ 2 \\ 3 \\ 1 \end{pmatrix}$$

To work out the total cost to each house, we would do the following. The cost to the first house will be

$$1 \times 2 + 2 \times 3 + 3 \times 1 = 11$$

and similarly for the second house it is

$$2 \times 2 + 2 \times 3 + 1 \times 1 = 11.$$

This example suggests how we might define multiplication of an $m \times n$ matrix by an $n \times 1$ matrix (i.e. a vector with n entries).

Ex: Multiply $A = \begin{pmatrix} 1 & 2 & 3 \\ 9 & -2 & 5 \\ -1 & 3 & -5 \end{pmatrix}$ by $\mathbf{v} = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}$.

Definition: If $A = (a_{ij})$ is an $m \times n$ matrix and $\mathbf{x}$ is an $n \times 1$ matrix with entries x_i then we define the product of A and $\mathbf{x}$ to be the matrix $\mathbf{b}$, an $m \times 1$ matrix whose entries are given by

$$b_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = \sum_{k=1}^{n} a_{ik}x_k \quad \text{for} \quad 1 \leq i \leq m.$$

To multiply two general matrices together, we think of each column of the second matrix as a vector. For this to make sense we need the number of columns of the first matrix to equal to number of rows of the second.

We can multiply an $n \times m$ matrix by an $m \times p$ matrix and the answer will be an $n \times p$ matrix.

Ex. Multiply the matrices A and B where $A = \begin{pmatrix} 1 & 2 & 3 \\ 9 & -2 & 5 \\ -1 & 3 & -5 \end{pmatrix}$

and $B = \begin{pmatrix} 2 & -3 \\ 5 & 9 \\ -2 & 4 \end{pmatrix}$.

Ex: Let $A = \begin{pmatrix} 2 & 3 \\ -2 & 5 \\ 3 & -5 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 2 \\ 9 & -2 \end{pmatrix}$.

Definition: Suppose $A = (a_{ij})$ is an $m \times n$ matrix and $B = (b_{ij})$ is an $n \times p$ matrix, then we define the product $C = AB$ to be the $m \times p$ matrix whose entries, c_{ij} are given by

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

for $1 \leq i \leq m$ and $1 \leq j \leq p$.

Ex: Find AB and BA if $A = \begin{pmatrix} 2 & 3 \\ -2 & 5 \end{pmatrix}$ and $B = \begin{pmatrix} 9 & -2 \\ -1 & 3 \end{pmatrix}$.

As you see from the above example, in general AB does not equal BA. This is extremely important fact to keep in mind. In general, matrix multiplication is NOT commutative.

We write A^n for $\underbrace{A \times A \times \cdots \times A}_{n \text{ times}}$.

Ex: Prove by induction that for any positive integer n,

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix}.$$

Ex: Find AI and IA if $A = \begin{pmatrix} 2 & 3 \\ -2 & 5 \end{pmatrix}$ and $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Theorem: If A is any $n \times n$ matrix, and I the identity matrix of size n, then $AI = IA = A$. Also $I^n = I$ for any positive integer n.

The identity matrix I plays the same role for matrix multiplication that the number 1 does for regular multiplication.

Theorem

If A, B and C are matrices for which the necessary sums and products are defined and λ is a scalar, then:

- $A(\lambda B) = \lambda AB$
- $(AB)C = A(BC)$. *(Associativity).*
- $(A + B)C = AC + BC$ *and* $C(A + B) = CA + CB$. *(Distibutivity).*

Proof of associativity:

Suppose A is an $m \times n$ matrix, B an $n \times p$ matrix and C a $p \times q$ matrix.

We will show that the entries in $A(BC)$ equal those in the matrix $(AB)C$. Thus,

$$(A(BC))_{ij} = \sum_{r=1}^{n} a_{ir}(BC)_{rj} = \sum_{r=1}^{n} a_{ir} \left(\sum_{s=1}^{p} b_{rs} c_{sj} \right)$$

$$= \sum_{r=1}^{n} \sum_{s=1}^{p} a_{ir} b_{rs} c_{sj} = \sum_{s=1}^{p} \sum_{r=1}^{n} a_{ir} b_{rs} c_{sj}.$$

At this stage, we have swapped the order of summation. You should check this step carefully. Continuing,

$$(A(BC)_{ij} = \sum_{s=1}^{p} \left(\sum_{r=1}^{n} a_{ir} b_{rs} \right) c_{sj} = \sum_{s=1}^{p} (AB)_{is} c_{sj} = ((AB)C)_{ij}.$$

A matrix A is said to be *idempotent* if $A^2 = A$.

Ex: Suppose A is idempotent. Prove that $B = I - A$ is idempotent and that $AB = BA$.

Another operation on matrices is the *transpose*. To take the transpose of a matrix, we interchange the rows and columns.

For example if $B = \begin{pmatrix} 1 & 2 \\ 9 & -2 \\ -1 & 3 \end{pmatrix}$ then $B^T = \begin{pmatrix} 1 & 9 & -1 \\ 2 & -2 & 3 \end{pmatrix}$.

More formally, $(A^T)_{ij} = a_{ji}$.

Thus $(A^T)^T = A$.

An operation which when composed with itself gives the identity function is called an *involution*. What are other examples of involutions?

The transpose of a column vector is a row vector. If we take two (column) vectors, $\mathbf{a}, \mathbf{b}$, (of the same size) then the quantity $\mathbf{a}^T\mathbf{b}$ is a 1×1 matrix, whose sole entry is the dot product of $\mathbf{a}$ and $\mathbf{b}$.

Ex: Find $\mathbf{a}^T\mathbf{b}$ and $\mathbf{a}\mathbf{b}^T$ for $\mathbf{a} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$.

Theorem

If A and B are matrices such that $A+B$ and AB is defined, then

- $(A^T)^T = A$
- $(A+B)^T = A^T + B^T$
- $(\lambda A)^T = \lambda A^T$.
- $(AB)^T = B^T A^T$.

Proof of last part: Suppose A is an $m \times n$ matrix, then B is an $n \times p$ matrix.

Then

$$(AB)^T_{ij} = (AB)_{ji} = \sum_{k=1}^{n} a_{jk} b_{ki}.$$

Also

$$(B^T A^T)_{ij} = \sum_{k=1}^{n} (B^T)_{ik} (A^T)_{kj} = \sum_{k=1}^{n} b_{ki} a_{jk} = \sum_{k=1}^{n} a_{jk} b_{ki} = (AB)^T_{ij}.$$

Ex: Simplify $A^T(CBA)^TC$.

Definition

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

Note that symmetric matrix must necessarily be **square.**

Ex: Suppose A, B are $n \times n$ symmetric matrices. Prove that AB is symmetric if and only if A and B commute.

Definition

A matrix A is said to be **skew-symmetric** if $A^T = -A$.

The entries on the main diagonal of a skew-symmetric matrix are always zero. (Why?)

So far we have been considering matrices with real entries. For complex numbers, the basic matrix arithmetic is the same, but there is also complex conjugation, which means conjugating all the entries.

For example, if $A = \begin{pmatrix} 1+2i & i \\ 3 & 2-3i \end{pmatrix}$ then

$$\overline{A} = \begin{pmatrix} 1-2i & -i \\ 3 & 2+3i \end{pmatrix}.$$

To generalise the notions of symmetry and skew symmetry, the transpose is replaced with the *conjugate transpose*, and we write

$$A^* = \overline{A}^T.$$

So in the example above

$$A^* = \begin{pmatrix} 1-2i & 3 \\ -i & 2+3i \end{pmatrix}.$$

Definition

A complex matrix A is said to be

- **Hermitian** if $A^* = A$
- **skew-Hermitian** if $A^* = -A$.

Ex: The matrix $A = \begin{pmatrix} 1 & 1-i & 2 \\ 1+i & 3 & i \\ 2 & -i & 0 \end{pmatrix}$ is Hermitian.

Ex: Prove that $(AB)^* = B^* A^*$.

Ex: If A is square matrix, prove that AA^* is Hermitian.

Rotation Matrices: Suppose $\begin{pmatrix} x \\ y \end{pmatrix}$ is the position vector of a point P in 2-dim space, which has distance r from the origin and makes an angle of α with the positive x axis.

This vector can be written as $\begin{pmatrix} r\cos\alpha \\ r\sin\alpha \end{pmatrix}$.

Now rotate P anti-clockwise about the origin through an angle θ to the new point $P' = \begin{pmatrix} x' \\ y' \end{pmatrix}$. Then

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} r\cos(\alpha+\theta) \\ r\sin(\alpha+\theta) \end{pmatrix} = \begin{pmatrix} r\cos\alpha\cos\theta - r\sin\alpha\sin\theta \\ r\sin\alpha\cos\theta + r\cos\alpha\sin\theta \end{pmatrix}$$

$$= \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = A\mathbf{x}.$$

Note that if we apply the rotation twice, we have

$$A^2 = \begin{pmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix}, \text{ and so on.}$$

Given a square matrix A, can we find a matrix B such that $AB = I$, the identity matrix? In general the answer is NO! If such a B exists, we call B the *inverse* of A and write $B = A^{-1}$. A is then **invertible**.

For a 2×2 matrix we have

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

as long as $ad - bc \neq 0$.

The quantity $ad - bc$ is called the *determinant* of the matrix.

Ex: Find the inverse of $\begin{pmatrix} 3 & 1 \\ -2 & 5 \end{pmatrix}$.

For matrices of higher order, there is no simple formula for the inverse. We will develop an algorithm for finding the inverse if it exists.

Suppose we have a 3×3 invertible matrix A whose inverse X we seek to find. We know that

$$AX = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Suppose we write X as $\left(\ \mathbf{x_1} \ \middle| \ \mathbf{x_2} \ \middle| \ \mathbf{x_3} \ \right)$ where $\mathbf{x_1}$ etc are the columns of X. Thus we have

$$A \left(\ \mathbf{x_1} \ \middle| \ \mathbf{x_2} \ \middle| \ \mathbf{x_3} \ \right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

We can think of this equation as three sets of equations, namely

$$A\mathbf{x_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad A\mathbf{x_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad A\mathbf{x_3} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

To solve each of these equations, we would take A and augment it with each of the three right hand sides in turn and row reduce. If we completely row reduce to reduced echelon form, then the solutions would appear on the right hand side of each reduction.

To save time we could do all three reductions in one go, by augmenting the matrix A with the identity matrix and completely row reducing. This is the algorithm for finding the inverse of A.

Ex: Find the inverse of $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{pmatrix}$.

Ex: Does the matrix $A = \begin{pmatrix} 1 & 6 & 4 \\ 2 & 4 & -1 \\ -1 & 2 & 5 \end{pmatrix}$ have an inverse?

Properties of Inverses:

Suppose that A, B are invertible matrices of the same size. Then

(i) $(A^{-1})^{-1} = A$.

(ii) $(AB)^{-1} = B^{-1}A^{-1}$

Similarly, $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$ assuming that all the matrices have the same size and invertible. (Why?)

To see why (ii) is true, note that

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

Hence the inverse of AB is $B^{-1}A^{-1}$.

Ex: Simplify $HG(FHG)^{-1}FG$.

We can use inverses to solve equations. For example, solve

$$\begin{pmatrix} 2 & 4 \\ -1 & 7 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 2 \\ -3 \end{pmatrix}.$$

If we are solving $A\mathbf{x} = \mathbf{b}$ and A is invertible, we can write $\mathbf{x} = A^{-1}\mathbf{b}$.

Theorem

If A is a square matrix then $A\mathbf{x} = \mathbf{b}$ has a unique solution if and only if A is invertible.

Determinants: We have seen that the inverse of the general 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ exists if and only if $ad - bc \neq 0$.

This quantity is called the **determinant** of the matrix.

We use the notation

$$\det(A) = ad - bc \text{ or } \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

For higher order matrices, there are various ways to define the determinant. We will take a computational approach and develop a method of calculating the determinant.

Definition

Suppose A is a square matrix. We define $|A_{ij}|$, called the ijth **minor** of A, to be the determinant of the matrix obtained by deleting the ith row and jth column of A.

For example, if $A = \begin{pmatrix} 1 & 4 & 5 \\ -1 & 2 & 7 \\ 2 & -3 & 4 \end{pmatrix}$ then

$$|A_{23}| = \begin{vmatrix} 1 & 4 \\ 2 & -3 \end{vmatrix} = -11.$$

We now define the determinant recursively:

Definition

The **determinant** of an $n \times n$ matrix is

$$|A| = a_{11}|A_{11}| - a_{12}|A_{12}| + a_{13}|A_{13}| - \cdots + (-1)^{n-1}a_{1n}|A_{1n}|.$$

In other words, we multiply each entry in the top row by the determinant of the matrix obtained by deleting the row and column that the entry belongs to and then we take an alternating sum of all these quantities.

Ex: Find the determinant of $A = \begin{pmatrix} 1 & 4 & 5 \\ -1 & 2 & 7 \\ 2 & -3 & 4 \end{pmatrix}$.

For 4×4 matrices (and higher orders), we have to apply this definition iteratively.

Properties of the Determinant:
Suppose A and B are $n \times n$ matrices then

1. $|AB| = |A||B|$.
2. $|A| = |A^T|$.
3. $|I| = 1$.
4. $|A^{-1}| = \frac{1}{|A|}$ assuming that $|A| \neq 0$.

Proof of (iv): $AA^{-1} = I$ so $|AA^{-1}| = |I|$. Now using (i) and (iii) we see that this is equal to $|A||A^{-1}| = 1$ and the result follows.

By (ii), we can also compute the determinant by expanding down the first column instead of the across the first row. This can be very useful calculationally, for example:

The deteminant

$$\begin{vmatrix} 1 & 4 & 5 \\ 0 & 2 & 7 \\ 0 & -3 & 4 \end{vmatrix} = 1 \begin{vmatrix} 2 & 7 \\ -3 & 4 \end{vmatrix} = 29,$$

since if we expand down the first column, the other minors are multiplied by zero.

Ex: Suppose $|A| = 3, \quad |B| = -2$, and A, B are square matrices of the same size. Simplify $|A^2 B^T (B^{-1})^2 A^{-1} B|$.

Further Properties of the Determinant:

1. $|A| = 0$ if and only if A is NOT invertible.
2. If A has a row or column of zeros then $|A| = 0$.
3. If A has two rows or columns which are multiples of each other then $|A| = 0$.
4. If a row or column of A is multiplied by a scalar λ then the determinant of A is multiplied by λ. Note that if A is an $n \times n$ matrix, then $|\lambda A| = \lambda^n |A|$.
5. If we swap two rows or two columns of A then the determinant of the resulting matrix is -1 times $|A|$.
6. If U is in row echelon form, then $|U|$ is simply the product of the diagonal entries.
7. If U is the row-echelon form of A, which was achieved using only row swaps and row subtractions then $|A| = \epsilon |U|$, where $\epsilon = 1$ if there was an even number of row swaps and -1 if there was an odd number of row swaps.

From these results, we can find an efficient method to calculate the determinant of any square matrix.

We take the matrix A and reduce it to row echelon form U, without making any row swaps or scalar multiplications. Then $|A| = |U|$ and the latter is the product of the diagonal entries.

If we want to make row swaps or scalar multiplications, then we need to keep track of these and apply the above rules.

Ex: Find $\begin{vmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{vmatrix}$.

Ex: Find $\begin{vmatrix} 1 & -1 & 0 & 3 \\ 2 & -2 & 6 & -1 \\ 4 & -2 & 1 & 7 \\ 3 & 5 & -7 & 2 \end{vmatrix}$.

Ex: Find $\Delta = \begin{vmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^3 & b^3 & c^3 & d^3 \end{vmatrix}$.

This generalises and is known as the *Van der Monde* determinant, $\Delta(a, b, c, d)$.

Theorem

Suppose A is a square matrix. Then $A\mathbf{x} = \mathbf{b}$ has a unique solution if and only if A is invertible, which is true if and only if $\det(A) \neq 0$.

Ex: Without solving, determine whether or not $A\mathbf{x} = \mathbf{b}$ has a unique solution if $A = \begin{pmatrix} 1 & -1 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 4 \end{pmatrix}$.