Introduction to matrices[1]

An $m \times n$ **matrix** ("over $\mathbb{R}$") is a rectangular array or table of elements of $\mathbb{R}$ arranged with m rows and n columns. It is usually written:

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

Matrices generalize vectors. A row vector of length n is a $1 \times n$ matrix. A column vector of length n is a $n \times 1$ matrix.

Basic definitions

The $(i,j)^{th}$ **entry** of A is a_{ij}. The i**th row of** A is

$$\begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix} \qquad (1 \le i \le m)$$

The j**th column of** A is

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \qquad (1 \le j \le n)$$

If A is an $m \times n$ matrix as above, the **transpose matrix**[2] ${}^TA$ is the $n \times m$ matrix, but with the rows of A written as columns and columns of A written as rows:

$${}^TA = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & & & \vdots \\ a_{1m} & a_{2m} & \dots & a_{nm} \end{pmatrix}.$$

Addition and scalar multiplication for matrices is just as for vectors.

The matrix A can be regarded as a column of row vectors,

(1) $$A = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_m \end{pmatrix}$$

where the ith row of A is

---

[1]Written by David Joyner, wdj@usna.edu.

[2]This is sometimes denoted as A^t or A^*. To avoid confusion with powers of a matrix, we use the notation ${}^TA$.

$$\mathbf{r}_i = \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix} \qquad (1 \le i \le m).$$

If

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

then the matrix times vector product is defined by

$$A\mathbf{x} = \begin{pmatrix} \mathbf{r}_1 \cdot^T \mathbf{x} \\ \mathbf{r}_2 \cdot^T \mathbf{x} \\ \vdots \\ \mathbf{r}_m \cdot^T \mathbf{x} \end{pmatrix}.$$

Example 1. *The* 2×2 *system*

$$\begin{cases} ax + by = r_1, \\ cx + dy = r_2, \end{cases}$$

may be visualized as two lines in a plane. Geometrically speaking, these lines could

- *intersect at a point (unique solution),*
- *be the same line written twice (infinitely many solutions)*[3]*,*
- *be parallel lines (no solutions)*[4]

Algebraically, this can be written as a matrix equation: $A\mathbf{x} = \mathbf{r}$*, which in the* 2×2 *case becomes*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}.$$

Linear systems of equations arise in many areas.

Example 2. *Consider the general form of the partial fraction decomposition,*

$$\frac{1}{(x-1)(x-2)(x+1)} = \frac{A}{x-1} + \frac{B}{x-2} + \frac{C}{x+1},$$

for some constants A, B, C*. Derive three equations in the three unknowns* A, B, C*. We cross multiply to clear denominators,*

$$\begin{aligned} 1 &= A(x-2)(x+1) + B(x-1)(x+1) + C(x-1)(x-2) \\ &= (A+B+C)x^2 + (-A-3C)x + (-2A-B+2C). \end{aligned}$$

Identifying like terms gives

$$A + B + C = 0, -A - 3C = 0, -2A - B + 2C = 1,$$

or

[3] For example, the system $x + y = 1$, $2x + 2y = 2$ has infinitely many solutions.

[4] For example, the system $x + y = 1$, $x + y = 2$ has no solutions since those equations imply $1 = 2$.

$$\begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & -3 \\ -2 & -1 & 2 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

The matrix A can also be regarded as a row of column vectors,

$$A = (\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n), \tag{2}$$

where the jth column of A is

$$\mathbf{c}_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \qquad (1 \le j \le n).$$

Note that

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

is a column of row vectors and

$${}^T A = ({}^T\mathbf{r}_1, {}^T\mathbf{r}_2, \ldots, {}^T\mathbf{r}_n),$$

is a row of column vectors.

A matrix A times a matrix B is defined similarly, when we think of B as a row of column vectors. If A is $m \times n$ and B is $n \times p$ and if B is a row of column vectors

$$B = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_p),$$

then

$$AB = (A\mathbf{b}_1, A\mathbf{b}_2, \ldots, A\mathbf{b}_p),$$

which is a $m \times p$ matrix.

Types of matrices

A matrix having as many rows as it has columns ($m = n$) is called a **square matrix**. The entries a_{ii} of an $m \times n$ matrix $A = (a_{ij})$ are called the **diagonal entries**, the entries a_{ij} with $i > j$ are called the **lower diagonal entries**, and the entries a_{ij} with $i < j$ are called the **upper diagonal entries**. An $m \times n$ matrix $A = (a_{ij})$ all of whose lower diagonal entries are zero is called an **upper triangular matrix**. This terminology is logical if the matrix is a square matrix but both the matrices below are called upper triangular

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & 6 & 7 \\ 0 & 0 & 8 & 9 \\ 0 & 0 & 0 & 10 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & 6 & 7 \\ 0 & 0 & 8 & 9 \\ 0 & 0 & 0 & 10 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

whether they look triangular or not! A similar definition holds for *lower* triangular matrices. The square $n \times n$ matrix with 1's on the diagonal and 0's elsewhere,

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

is called the $n \times n$ **identity matrix** and denoted I or I_n. This is both upper triangular and lower triangular. In general, any square matrix which is both upper triangular and lower triangular is called a **diagonal matrix**. A diagonal matrix of the form aI_n, for $a \in \mathbb{R}$, is called a **scalar matrix**. Note all the diagonal entries of a scalar matrix are the same.

The determinant as a measure of parallelopiped volume

The determinant is a function that attached to each square matrix, say A, with real entries, a real number, $\det(A)$. This number measures the volume of the parallelopiped spanned by the row vectors of A. For example, if $\mathbf{r}_1 = (a, b)$ and $\mathbf{r}_2 = (c, d)$ are the row vectors of a 2×2 matrix A and if P is the parallelogram formed by $\mathbf{r}_1, \mathbf{r}_2$ (see Figure 1), then

$$area(P) = |\det(A)|. \tag{3}$$

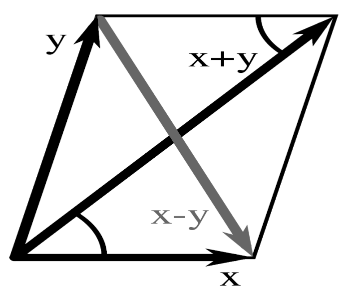

FIGURE 1. Parallelogram spanned by $\mathbf{x} = \mathbf{r}_1$ and $\mathbf{y} = \mathbf{r}_2$.
Source: `https://en.wikipedia.org/wiki/Parallelogram_law`.

Likewise, if $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_3$, are the row vectors of a 3×3 matrix A and if P is the parallelopiped formed by $\mathbf{u} = \mathbf{r}_1$, $\mathbf{v} = \mathbf{r}_2$, $\mathbf{w} = \mathbf{w} = \mathbf{r}_3$ (see Figure 2), then the 3-dimensional analog of (3), (4), holds:

$$vol(P) = |\det(A)|. \tag{4}$$

While interesting, this isn't a practical method of computing $\det(A)$.

The determinant cofactor expansion

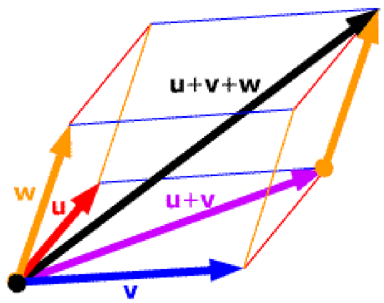

FIGURE 2. Parallelopiped spanned by **u** and **v**, **w**.
Source: `https://en.wikipedia.org/wiki/Parallelogram_law`.

For any $n \times n$ matrix $A = (a_{i,j})_{1 \le i,j \le n}$, let $A_{i,j}$ denote the $(n-1) \times (n-1)$ submatrix obtained by removing the ith row and jth column of A. The expression $\det(A_{i,j})$ is called the (i,j)-*minor* of A and $(-1)^{i+j}\det(A_{i,j})$ is called the (i,j)-*cofactor* of A.

The determinant formula

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \det(A_{i,j})$$

is called the *cofactor expansion down the jth column*. The formula

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \det(A_{i,j})$$

is called the *cofactor expansion along the ith row.*

In the 2×2 case, this becomes the formula:

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

In the 3×3 case, this becomes the formula:

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = a \det \begin{pmatrix} e & f \\ h & i \end{pmatrix} - b \det \begin{pmatrix} d & f \\ g & i \end{pmatrix} + c \det \begin{pmatrix} d & e \\ g & h \end{pmatrix}$$

$$= aei - ahf - bdi + bfg + cdh - cge.$$

Theorem 3. *Let A, B be an $n \times n$ matrix. The following statements are true.*

- $\det(A^k) = \det(A)^k$.
- $\det({}^TA) = \det(A)$ *(where ${}^TA$ is the transpose of A).*
- $det(A)\det(B) = \det(AB)$.
- *If we write as as a column of row vectors,*

$$A = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_n \end{pmatrix}$$

then the elementary row operation $R_i + cR_j \to R_i$ *does not change the value of the determinant:*

$$\det \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_i \vdots \\ \mathbf{r}_n \end{pmatrix} = \det \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_i + c\mathbf{r}_j \vdots \\ \mathbf{r}_n \end{pmatrix}$$

- *If we write as as a column of row vectors,*

$$A = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_n \end{pmatrix}$$

then the elementary row operation $cR_i \to R_i$ *does changes the value of the determinant by a factor of* c*:*

$$c \det \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_i \vdots \\ \mathbf{r}_n \end{pmatrix} = \det \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ c\mathbf{r}_i \vdots \\ \mathbf{r}_n \end{pmatrix}$$

Theorem 4. *Let* A *be an* $n \times n$ *matrix with* $\det(A) \neq 0$*. The following statements are true.*

- $\det(A^{-1}) = \det(A)^{-1}$.
- *The system* $A\mathbf{x} = \mathbf{0}$ *has the unique solution* $\mathbf{x} = \mathbf{0}$.
- *The system* $A\mathbf{x} = \mathbf{b}$ *has a unique solution (namely,* $\mathbf{x} = A^{-1}\mathbf{b}$*).*
- *The rank of* A *is* n.
- $rref(A) = I_n$.
- $rref(A, I) = (I, A^{-1})$.
- $rref(A, \mathbf{b}) = (I, A^{-1}\mathbf{b})$.

Cramer's Rule: Write A as a vector of column vectors:

$$A = (\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n),$$

where

$$\mathbf{c}_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}, \qquad (1 \le j \le n).$$

Then the ith coordinate of the solution $\mathbf{x}$ to $A\mathbf{x} = \mathbf{b}$ is

$$x_i = \frac{\det(\mathbf{c}_1, \ldots, \mathbf{c}_{i-1}, \mathbf{b}, \mathbf{c}_{i+1}, \ldots, \mathbf{c}_n)}{\det(\mathbf{c}_1, \ldots, \mathbf{c}_{i-1}, \mathbf{c}_i, \mathbf{c}_{i+1}, \ldots, \mathbf{c}_n)}, \qquad (1 \le i \le n).$$

Exercise 1: Compute the determinant of

$$\begin{pmatrix} 1 & -3 \\ 2 & -6 \end{pmatrix}.$$

Exercise 2: Compute the determinant of

$$A = \begin{pmatrix} 1 & -3 & -2 \\ 2 & -6 & -4 \\ -1 & 1 & 0 \end{pmatrix}.$$

What is the volume of the parallelopiped spanned by the columns of A?

Exercise 3: Compute the determinant of

$$A = \begin{pmatrix} 5 & -2017 & -2018 \\ 0 & -4 & 2019 \\ 0 & 0 & -101 \end{pmatrix}.$$