Matrices

Dr. Andreas Hanke

Supplement to the lecture *Classical Mechanics*

Compiled from http://www.sosmath.com/matrix/matrix.html

A. Matrix Algebra

Matrices and Determinants were discovered and developed in the eighteenth and nineteenth centuries. Initially, their development dealt with transformation of geometric objects and solution of systems of linear equations. Historically, the early emphasis was on the determinant, not the matrix. In modern treatments of linear algebra, matrices are considered first. We will not speculate much on this issue.
Matrices provide a theoretically and practically useful way of approaching many types of problems including:

- Solution of Systems of Linear Equations,
- Equilibrium of Rigid Bodies (in physics),
- Graph Theory,
- Theory of Games,
- Leontief Economics Model,
- Forest Management,
- Computer Graphics, and Computed Tomography,
- Genetics,
- Cryptography,
- Electrical Networks,
- Fractals.

A.1. Multiplication of Matrices

We have

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{pmatrix}$$

In fact, we do not need to have two matrices of the same size to multiply them. The general rule says that in order to perform the multiplication AB, where A is a (mxn) matrix and B a (kxl) matrix, then we must have $n=k$. The result will be a (mxl) matrix. For example, we have

$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \\ \nu \end{pmatrix} = \begin{pmatrix} a\alpha+b\beta+c\nu \\ d\alpha+e\beta+f\nu \end{pmatrix}.$$

Remember that though we were able to perform the above multiplication, it is not possible to perform the multiplication

$$\begin{pmatrix} \alpha \\ \beta \\ \nu \end{pmatrix}\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}.$$

So we have to be very careful about multiplying matrices. Sentences like "multiply the two matrices A and B" do not make sense. You must know which of the two matrices will be to the right (of your multiplication) and which one will be to the left; in other words, we have to know whether we are asked to perform $A \times B$ or $B \times A$. Even if both multiplications do make sense (as in the case of square matrices with the same size), we still have to be very careful. Indeed, consider the two matrices

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

We have

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

and

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

So what is the conclusion behind this example? The matrix multiplication is not commutative, the order in which matrices are multiplied is important. In fact, this little setback is a major problem in playing around with matrices. This is something that you must always be careful with. Let us show you another setback. We have

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}; \text{ i.e.,}$$

the product of two non-zero matrices may be equal to the zero-matrix.

A.2. Algebraic Properties of Matrix Operations

In this page, we give some general results about the three operations: addition, multiplication, and multiplication with numbers, called **scalar multiplication**.

From now on, we will not write (mxn) but mxn.

Properties involving Addition. Let A, B, and C be mxn matrices. We have

1. $$A+B=B+A$$
2. $$(A+B)+C=A+(B+C)$$
3. $$A+\mathcal{O}=A$$

   where $\mathcal{O}$ is the mxn zero-matrix (all its entries are equal to 0);
4. $$A+B=\mathcal{O}$$

   if and only if $B=-A$.

Properties involving Multiplication.

1. Let A, B, and C be three matrices. If you can perform the products AB, $(AB)C$, BC, and $A(BC)$, then we have

   $$(AB)C=A\,(BC)$$

   Note, for example, that if A is 2x3, B is 3x3, and C is 3x1, then the above products are possible (in this case, $(AB)C$ is 2x1 matrix).
2. If α and β are numbers, and A is a matrix, then we have

   $$\alpha(\beta A)=(\alpha\beta)A$$

3. If α is a number, and A and B are two matrices such that the product $A\cdot B$ is possible, then we have

   $$\alpha(AB)=(\alpha A)B=A(\alpha B)$$

4. If A is an nxm matrix and $\mathcal{O}$ the mxk zero-matrix, then

   $$A\mathcal{O}=\mathcal{O}$$

   Note that $A\mathcal{O}$ is the nxk zero-matrix. So if n is different from m, the two zero-matrices are different.

Properties involving Addition and Multiplication.

1. Let A, B, and C be three matrices. If you can perform the appropriate products, then we have

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

and

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

2. If α and β are numbers, A and B are matrices, then we have

$$\alpha(A+B) = \alpha A + \alpha B$$

and

$$(\alpha+\beta)A = \alpha A + \beta A$$

Example. Consider the matrices

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, B = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \text{ and } C = \begin{pmatrix} 0 & 1 & 5 \end{pmatrix}.$$

Evaluate $(AB)C$ and $A(BC)$. Check that you get the same matrix.

Answer. We have

$$AB = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$$

so

$$(AB)C = \begin{pmatrix} -1 \\ -2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 5 \end{pmatrix} = \begin{pmatrix} 0 & -1 & -5 \\ 0 & -2 & -10 \end{pmatrix}.$$

On the other hand, we have

$$BC = \begin{pmatrix} 0 & 2 & 10 \\ 0 & -1 & -5 \end{pmatrix}$$

so

$$A(BC) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 2 & 10 \\ 0 & -1 & -5 \end{pmatrix} = \begin{pmatrix} 0 & -1 & -5 \\ 0 & -2 & -10 \end{pmatrix}.$$

Example. Consider the matrices

$$X = \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \text{ and } Y = \begin{pmatrix} \alpha & \beta & \nu & \gamma \end{pmatrix}.$$

It is easy to check that

$$X = a\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + b\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + c\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

and

$$Y = \alpha\begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix} + \beta\begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix} + \nu\begin{pmatrix} 0 & 0 & 1 & 0 \end{pmatrix} + \gamma\begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix}.$$

These two formulas are called **linear combinations.**

We have seen that matrix multiplication is different from normal multiplication (between numbers). Are there some similarities? For example, is there a matrix which plays a similar role as the number 1? The answer is yes. Indeed, consider the nxn matrix

$$I_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

In particular, we have

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The matrix I_n has similar behavior as the number 1. Indeed, for any nxn matrix A, we have

$$A\, I_n = I_n A = A$$

The matrix I_n is called the **Identity Matrix** of order n.

Example. Consider the matrices

$$A = \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix} \text{ and } B = \begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix}.$$

Then it is easy to check that

$$AB = I_2 \text{ and } BA = I_2.$$

The identity matrix behaves like the number 1 not only among the matrices of the form nxn. Indeed, for any nxm matrix A, we have

$$I_n A = A \text{ and } A I_m = A.$$

In particular, we have

$$I_4 \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}.$$

A.3. Invertible Matrices

Definition. An $n \times n$ matrix A is called **nonsingular** or **invertible** iff there exists an $n \times n$ matrix B such that

$$A\,B = B\,A = I_n$$

where I_n is the identity matrix. The matrix B is called the **inverse matrix** of A.

Example. Let

$$A = \begin{pmatrix} 2 & 3 \\ 2 & 2 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} -1 & 3/2 \\ 1 & -1 \end{pmatrix}.$$

One may easily check that

$$A\,B = B\,A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2 .$$

Hence A is invertible and B is its inverse.

Notation. A common notation for the inverse of a matrix A is A^{-1}. So

$$A\,A^{-1} = A^{-1}\,A = I_n .$$

Example. Find the inverse of

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}.$$

Write

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

Since

$$A\,A^{-1} = \begin{pmatrix} a+c & b+d \\ -a+2c & -b+2d \end{pmatrix} = I_2$$

we get

$$\begin{cases} a + c = 1 \\ -a + 2c = 0 \\ b + d = 0 \\ -b + 2d = 1 \end{cases}$$

Easy algebraic manipulations give

$$a = \frac{2}{3},\quad b = -\frac{1}{3},\quad c = \frac{1}{3},\quad d = \frac{1}{3}$$

or

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

The inverse matrix is unique when it exists. So if A is invertible, then A^{-1} is also invertible and

$$\left(A^{-1}\right)^{-1} = A.$$

The following basic property is very important:

If A and B are invertible matrices, then $A\,B$ is also invertible and

$$\left(A\,B\right)^{-1} = B^{-1}\,A^{-1}.$$

Remark. In the definition of an invertible matrix A, we used both $A\,B$ and $B\,A$ to be equal to the identity matrix. In fact, we need only one of the two. In other words, for a matrix A, if there exists a matrix B such that $A\,B = I_n$, then A is invertible and $B = A^{-1}$.

A.4. Special Matrices: Triangular, Symmetric, Diagonal

We have seen that a matrix is a block of entries or two dimensional data. The size of the matrix is given by the number of rows and the number of columns. If the two numbers are the same, we called such matrix a square matrix.

To square matrices we associate what we call the **main diagonal** (in short the diagonal). Indeed, consider the matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Its diagonal is given by the numbers a and d. For the matrix

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & k \end{pmatrix}$$

its diagonal consists of a, e, and k. In general, if A is a square matrix of order n and if a_{ij} is the number in the i^{th}-row and j^{th}-colum, then the diagonal is given by the numbers a_{ii}, for i=1,..,n.

The diagonal of a square matrix helps define two type of matrices: **upper-triangular** and **lower-triangular**. Indeed, the diagonal subdivides the matrix into two blocks: one above the diagonal and the other one below it. If the lower-block consists of zeros, we call such a matrix **upper-triangular**. If the upper-block consists of zeros, we call such a matrix **lower-triangular**. For example, the matrices

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \text{ and } \begin{pmatrix} a & b & c \\ 0 & e & f \\ 0 & 0 & k \end{pmatrix}$$

are upper-triangular, while the matrices

$$\begin{pmatrix} a & 0 \\ c & d \end{pmatrix} \text{ and } \begin{pmatrix} a & 0 & 0 \\ d & e & 0 \\ g & h & k \end{pmatrix}$$

are lower-triangular. Now consider the two matrices

$$A = \begin{pmatrix} a & 0 & 0 \\ d & e & 0 \\ g & h & k \end{pmatrix} \text{ and } B = \begin{pmatrix} a & d & g \\ 0 & e & h \\ 0 & 0 & k \end{pmatrix}.$$

The matrices A and B are triangular. But there is something special about these two matrices. Indeed, as you can see if you reflect the matrix A about the diagonal, you get the matrix B. This operation is called the **transpose operation.** Indeed, let A be a nxm matrix defined by the numbers a_{ij}, then the transpose of A, denoted A^T is the mxn matrix defined by the numbers b_{ij} where $b_{ij} = a_{ji}$. For example, for the matrix

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & k \\ t & r & s \end{pmatrix}$$

we have

$$A^T = \begin{pmatrix} a & d & g & t \\ b & e & h & r \\ c & f & k & s \end{pmatrix}.$$

Properties of the Transpose operation. If X and Y are mxn matrices and Z is an nxk matrix, then

1. $(X+Y)^T = X^T + Y^T$
2. $(XZ)^T = Z^T X^T$
3. $(X^T)^T = X$

A **symmetric matrix** is a matrix equal to its transpose. So a symmetric matrix must be a square matrix. For example, the matrices

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix} \text{ and } \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$$

are symmetric matrices. In particular a symmetric matrix of order n, contains at most $\frac{n(n+1)}{2}$ different numbers.

A **diagonal matrix** is a symmetric matrix with all of its entries equal to zero except may be the ones on the diagonal. So a diagonal matrix has at most n different numbers. For example, the matrices

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \text{ and } \begin{pmatrix} a & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & b \end{pmatrix}$$

are diagonal matrices. Identity matrices are examples of diagonal matrices. Diagonal matrices play a crucial role in matrix theory. We will see this later on.

Example. Consider the diagonal matrix

$$A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}.$$

Define the power-matrices of A by

$$A^0 = I_2,\ A^1 = A,\ A^2 = AA,\ A^3 = AAA \text{ etc..}$$

Find the power matrices of A and then evaluate the matrices

$$I_2 + \frac{1}{1!}A + \frac{1}{2!}A^2 + \cdots + \frac{1}{n!}A^n$$

for n=1,2,....

Answer. We have

$$A^2 = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} a^2 & 0 \\ 0 & b^2 \end{pmatrix}$$

and

$$A^3 = A^2A = \begin{pmatrix} a^2 & 0 \\ 0 & b^2 \end{pmatrix}\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} a^3 & 0 \\ 0 & b^3 \end{pmatrix}.$$

By induction, one may easily show that

$$A^n = \begin{pmatrix} a^n & 0 \\ 0 & b^n \end{pmatrix}$$

for every natural number n. Then we have

$$I_2+\frac{1}{1!}A+\frac{1}{2!}A^2+\cdots+\frac{1}{n!}A^n = \begin{pmatrix} 1+\frac{a}{1!}+\frac{a^2}{2!}+\cdots+\frac{a^n}{n!} & 0 \\ 0 & 1+\frac{b}{1!}+\frac{b^2}{2!}+\cdots+\frac{b^n}{n!} \end{pmatrix}$$

for n=1,2,..

Scalar Product. Consider the 3x1 matrices

$$X = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \text{ and } Y = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}.$$

The **scalar product** of X and Y is defined by

$$X^TY = \begin{pmatrix} a & b & c \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = a\alpha + b\beta + c\gamma.$$

In particular, we have

$X^TX = (a^2 + b^2 + c^2)$. This is a 1 x 1 matrix.

A.5. Elementary Operations for Matrices

Elementary operations for matrices play a crucial role in finding the inverse or solving linear systems. They may also be used for other calculations. On this page, we will discuss these type of operations. Before we define an elementary operation, recall that to an nxm matrix A, we can associate n rows and m columns. For example, consider the matrix

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

Its rows are

$$\begin{pmatrix} 0 & 1 & -1 & 3 \end{pmatrix}, \begin{pmatrix} 0 & 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 & 2 & -3 \end{pmatrix}.$$

Its columns are

$$\begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ -3 \end{pmatrix}.$$

Let us consider the matrix transpose of A

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

Its rows are

$$\begin{pmatrix} 0 & 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 3 & 1 & -3 \end{pmatrix}.$$

As we can see, the transpose of the columns of A are the rows of A^T. So the transpose operation interchanges the rows and the columns of a matrix. Therefore many techniques which are developed for rows may be easily translated to columns via the transpose operation. Thus, we will only discuss elementary row operations, but the reader may easily adapt these to columns.

Elementary Row Operations.

1. Interchange two rows.
2. Multiply a row with a nonzero number.
3. Add a row to another one multiplied by a number.

Definition. Two matrices are **row equivalent** if and only if one may be obtained from the other one via elementary row operations.

Example. Show that the two matrices

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

are row equivalent.

Answer. We start with A. If we keep the second row and add the first to the second, we get

$$\begin{pmatrix} 3 & 0 & 1 \\ 2 & 1 & 1 \end{pmatrix}.$$

We keep the first row. Then we subtract the first row from the second one multiplied by 3. We get

$$\begin{pmatrix} 3 & 0 & 1 \\ 3 & 3 & 2 \end{pmatrix}.$$

We keep the first row and subtract the first row from the second one. We get

$$\begin{pmatrix} 3 & 0 & 1 \\ 0 & 3 & 1 \end{pmatrix}$$

which is the matrix B. Therefore A and B are row equivalent.

One powerful use of elementary operations consists in finding solutions to linear systems and the inverse of a matrix. This happens via **Echelon Form** and **Gauss-Jordan Elimination**. In order to appreciate these two techniques, we need to discuss when a matrix is row elementary equivalent to a triangular matrix. Let us illustrate this with an example.

Example. Consider the matrix

$$\begin{pmatrix} 0 & 0 & 1 & 3 \\ 2 & 4 & 0 & -8 \\ 1 & 2 & 1 & -1 \end{pmatrix}.$$

First we will transform the first column via elementary row operations into one with the top number equal to 1 and the bottom ones equal 0. Indeed, if we interchange the first row with the last one, we get

$$\begin{pmatrix} 1 & 2 & 1 & -1 \\ 2 & 4 & 0 & -8 \\ 0 & 0 & 1 & 3 \end{pmatrix}.$$

Next, we keep the first and last rows. And we subtract the first one multiplied by 2 from the second one. We get

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

We are almost there. Looking at this matrix, we see that we can still take care of the 1 (from the last row) under the -2. Indeed, if we keep the first two rows and add the second one to the last one multiplied by 2, we get

$$\begin{pmatrix} 1 & 2 & 1 & -1 \\ 0 & 0 & -2 & -6 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

We can't do more. Indeed, we stop the process whenever we have a matrix which satisfies the following conditions

1. any row consisting of zeros is below any row that contains at least one nonzero number;
2. the first (from left to right) nonzero entry of any row is to the left of the first nonzero entry of any lower row.

Now if we make sure that the first nonzero entry of every row is 1, we get a matrix in **row echelon form**. For example, the matrix above is not in echelon form. But if we divide the second row by -2, we get

$$\begin{pmatrix} 1 & 2 & 1 & -1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

This matrix is in echelon form.

A.6. Matrix Exponential

The matrix exponential plays an important role in solving system of linear differential equations. On this page, we will define such an object and show its most important properties. The natural way of defining the exponential of a matrix is to go back to the exponential function e^x and find a definition which is easy to extend to matrices. Indeed, we know that the Taylor polynomials

$$T_n(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!}$$

converges pointwise to e^x and uniformly whenever x is bounded. These algebraic polynomials may help us in defining the exponential of a matrix. Indeed, consider a square matrix A and define the sequence of matrices

$$A_n = I_n + \frac{1}{1!}A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots + \frac{1}{n!}A^n.$$

When n gets large, this sequence of matrices get closer and closer to a certain matrix. This is not easy to show; it relies on the conclusion on e^x above. We write this limit matrix as e^A. This notation is natural due to the properties of this matrix. Thus we have the formula

$$e^A = I_n + \frac{1}{1!}A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots + \frac{1}{n!}A^n + \cdots\cdots$$

One may also write this in series notation as

$$e^A = \sum_{n=0}^{\infty} \frac{1}{n!}A^n$$

At this point, the reader may feel a little lost about the definition above. To make this stuff clearer, let us discuss an easy case: diagonal matrices.

Example. Consider the diagonal matrix

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

It is easy to check that

$$A^n = \begin{pmatrix} 2^n & 0 \\ 0 & (-1)^n \end{pmatrix}$$

for $n = 1, 2, \cdots$. Hence we have

$$I_n+\frac{1}{1!}A+\frac{1}{2!}A^2+\frac{1}{3!}A^3+\cdots+\frac{1}{n!}A^n=\begin{pmatrix} 1+\frac{2}{1!}+\cdots+\frac{2^n}{n!} & 0 \\ 0 & 1+\frac{(-1)}{1!}+\cdots+\frac{(-1)^n}{n!} \end{pmatrix}.$$

Using the above properties of the exponential function, we deduce that

$$e^A=\begin{pmatrix} e^2 & 0 \\ 0 & e^{-1} \end{pmatrix}.$$

Indeed, for a diagonal matrix A, e^A can always be obtained by replacing the entries of A (on the diagonal) by their exponentials. Now let B be a matrix similar to A. As explained before, then there exists an invertible matrix P such that

$$B=P^{-1}AP.$$

Moreover, we have

$$B^n=P^{-1}A^nP$$

for $n=1,2,\cdots$, which implies

$$I_n+\frac{1}{1!}B+\frac{1}{2!}B^2+\cdots+\frac{1}{n!}B^n=P^{-1}\left(I_n+\frac{1}{1!}A+\frac{1}{2!}A^2+\cdots+\frac{1}{n!}A^n\right)P.$$

This clearly implies that

$$e^B=P^{-1}\begin{pmatrix} e^2 & 0 \\ 0 & e^{-1} \end{pmatrix}P.$$

In fact, we have a more general conclusion. Indeed, let A and B be two square matrices. Assume that $A\sim B$. Then we have $e^A\sim e^B$. Moreover, if $B=P^{-1}AP$, then

$$e^B=P^{-1}e^AP.$$

Example. Consider the matrix

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

This matrix is upper-triangular. Note that all the entries on the diagonal are 0. These types of matrices have a nice property. Let us discuss this for this example. First, note that

$$A^2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ and } A^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \mathcal{O}.$$

In this case, we have

$$e^A = I + A + \frac{1}{2!}A^2 = \begin{pmatrix} 1 & 1 & 3/2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

In general, let A be a square upper-triangular matrix of order n. Assume that all its entries on the diagonal are equal to 0. Then we have

$$A^n = \mathcal{O}.$$

Such matrix is called a **nilpotent** matrix. In this case, we have

$$e^A = I_n + \frac{1}{1!}A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots + \frac{1}{(n-1)!}A^{n-1}.$$

As we said before, the reasons for using the exponential notation for matrices reside in the following properties:

Theorem. The following properties hold:

1.

$$e^{\mathcal{O}} = I_n$$

;

2.

if A and B commute, meaning $AB = BA$, then we have
$$e^{A+B} = e^A e^B;$$

3.

for any matrix A, e^A is invertible and

$$\left(e^A\right)^{-1} = e^{-A}.$$

B. Determinants

B.1. Introduction to Determinants

For any square matrix of order 2, we have found a necessary and sufficient condition for invertibility. Indeed, consider the matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

The matrix A is invertible if and only if $ad - bc \neq 0$. We called this number the **determinant** of A. It is clear from this, that we would like to have a similar result for bigger matrices (meaning higher orders). So is there a similar notion of determinant for any square matrix, which determines whether a square matrix is invertible or not?

In order to generalize such notion to higher orders, we will need to study the determinant and see what kind of properties it satisfies. First let us use the following notation for the determinant

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

Properties of the Determinant

1. Any matrix A and its transpose have the same determinant, meaning

$$\det A = \det A^T$$

This is interesting since it implies that whenever we use rows, a similar behavior will result if we use columns. In particular we will see how row elementary operations are helpful in finding the determinant. Therefore, we have similar conclusions for elementary column operations.

2. The determinant of a triangular matrix is the product of the entries on the diagonal, that is

$$\begin{vmatrix} a & b \\ 0 & d \end{vmatrix} = \begin{vmatrix} a & 0 \\ b & d \end{vmatrix} = ad.$$

3.

If we interchange two rows, the determinant of the new matrix is the opposite of the old one, that is

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = -\begin{vmatrix} c & d \\ a & b \end{vmatrix}$$

4.

If we multiply one row with a constant, the determinant of the new matrix is the determinant of the old one multiplied by the constant, that is

$$\begin{vmatrix} \lambda a & \lambda b \\ c & d \end{vmatrix} = \lambda \begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ \lambda c & \lambda d \end{vmatrix}.$$

In particular, if all the entries in one row are zero, then the determinant is zero.

5.

If we add one row to another one multiplied by a constant, the determinant of the new matrix is the same as the old one, that is

$$\begin{vmatrix} a+\lambda c & b+\lambda d \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c+\lambda a & d+\lambda b \end{vmatrix}.$$

Note that whenever you want to replace a row by something (through elementary operations), do not multiply the row itself by a constant. Otherwise, you will easily make errors (due to Property 4).

6.

We have

$$\det(AB) = \det(A)\det(B).$$

In particular, if A is invertible (which happens if and only if $\det(A) \neq 0$), then

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

If A and B are similar, then $\det(A) = \det(B)$.

Let us look at an example, to see how these properties work.

Example. Evaluate

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

Let us transform this matrix into a triangular one through elementary operations. We will keep the first row and add to the second one the first multiplied by $\frac{1}{2}$. We get

$$\begin{vmatrix} 2 & 1 \\ -1 & 3 \end{vmatrix} = \begin{vmatrix} 2 & 1 \\ 0 & \frac{7}{2} \end{vmatrix}.$$

Using the Property 2, we get

$$\begin{vmatrix} 2 & 1 \\ 0 & \frac{7}{2} \end{vmatrix} = 2 \cdot \frac{7}{2} = 7.$$

Therefore, we have

$$\begin{vmatrix} 2 & 1 \\ -1 & 3 \end{vmatrix} = 7$$

which one may check easily.

B.2. Determinants of Matrices of Higher Order

As we said before, the idea is to assume that previous properties satisfied by the determinant of matrices of order 2, are still valid in general. In other words, we assume:

1. Any matrix A and its transpose have the same determinant, meaning

$$\det A = \det A^T.$$

2. The determinant of a triangular matrix is the product of the entries on the diagonal.

3. If we interchange two rows, the determinant of the new matrix is the opposite of the old one.

4.

If we multiply one row with a constant, the determinant of the new matrix is the determinant of the old one multiplied by the constant.

5.

If we add one row to another one multiplied by a constant, the determinant of the new matrix is the same as the old one.

6.

We have

$$\det(AB) = \det(A)\det(B).$$

In particular, if A is invertible (which happens if and only if $\det(A) \neq 0$), then

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

So let us see how this works in case of a matrix of order 4.

Example. Evaluate

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 2 & 6 & 4 & 8 \\ 3 & 1 & 1 & 2 \end{vmatrix}.$$

We have

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 2 & 6 & 4 & 8 \\ 3 & 1 & 1 & 2 \end{vmatrix} = 2\begin{vmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 1 & 3 & 2 & 4 \\ 3 & 1 & 1 & 2 \end{vmatrix}.$$

If we subtract every row multiplied by the appropriate number from the first row, we get

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 1 & 3 & 2 & 4 \\ 3 & 1 & 1 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \\ 0 & 1 & -1 & 0 \\ 0 & -5 & -8 & -10 \end{vmatrix}.$$

We do not touch the first row and work with the other rows. We interchange the second with the third to get

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \\ 0 & 1 & -1 & 0 \\ 0 & -5 & -8 & -10 \end{vmatrix} = -\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & -1 & 0 \\ 0 & -4 & -8 & -12 \\ 0 & -5 & -8 & -10 \end{vmatrix}.$$

If we subtract every row multiplied by the appropriate number from the second row, we get

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & -1 & 0 \\ 0 & -4 & -8 & -12 \\ 0 & -5 & -8 & -10 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -12 & -12 \\ 0 & 0 & -13 & -10 \end{vmatrix}.$$

Using previous properties, we have

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -12 & -12 \\ 0 & 0 & -13 & -10 \end{vmatrix} = -12\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -13 & -10 \end{vmatrix}.$$

If we multiply the third row by 13 and add it to the fourth, we get

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -13 & -10 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 3 \end{vmatrix}$$

which is equal to 3. Putting all the numbers together, we get

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 2 & 6 & 4 & 8 \\ 3 & 1 & 1 & 2 \end{vmatrix} = 2 \cdot (-1) \cdot (-12) \cdot 3 = 72.$$

These calculations seem to be rather lengthy. We will see later on that a general formula for the determinant does exist.

Example. Evaluate

$$\begin{vmatrix} 1 & 2 & 0 \\ -1 & 1 & 1 \\ 1 & 2 & 3 \end{vmatrix}.$$

In this example, we will not give the details of the elementary operations. We have

$$\begin{vmatrix} 1 & 2 & 0 \\ -1 & 1 & 1 \\ 1 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{vmatrix} = 9.$$

Example. Evaluate

$$\begin{vmatrix} 1 & 1 & 2 \\ 0 & 1 & 0 \\ 2 & 1 & -1 \end{vmatrix}.$$

We have

$$\begin{vmatrix} 1 & 1 & 2 \\ 0 & 1 & 0 \\ 2 & 1 & -1 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 2 \\ 0 & 1 & 0 \\ 0 & -1 & -5 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \end{vmatrix} = -5.$$

General Formula for the Determinant Let A be a square matrix of order n. Write $A = (a_{ij})$, where a_{ij} is the entry on the row number i and the column number j, for $i = 1, \cdots, n$ and $j = 1, \cdots, n$. For any i and j, set A_{ij} (called **the cofactors**) to be the determinant of the square matrix of order (n-1) obtained from A by removing the row number i and the column number j multiplied by $(-1)^{i+j}$. We have

$$\det(A) = \sum_{j=1}^{j=n} a_{ij} A_{ij}$$

for any fixed i, and

$$\det(A) = \sum_{i=1}^{i=n} a_{ij} A_{ij}$$

for any fixed j. In other words, we have two type of formulas: along a row (number i) or along a column (number j). Any row or any column will do. The trick is to use a row or a column which has a lot of zeros.

In particular, we have along the rows

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & k \end{vmatrix} = a\begin{vmatrix} e & f \\ h & k \end{vmatrix} - b\begin{vmatrix} d & f \\ g & k \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

or

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & k \end{vmatrix} = -d\begin{vmatrix} b & c \\ h & k \end{vmatrix} + e\begin{vmatrix} a & c \\ g & k \end{vmatrix} - f\begin{vmatrix} a & b \\ g & h \end{vmatrix}$$

or

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & k \end{vmatrix} = g\begin{vmatrix} b & c \\ e & f \end{vmatrix} - h\begin{vmatrix} a & c \\ d & f \end{vmatrix} + k\begin{vmatrix} a & b \\ d & e \end{vmatrix}.$$

As an exercise write the formulas along the columns.

Example. Evaluate

$$\begin{vmatrix} 3 & 2 & 1 \\ 2 & 1 & -3 \\ 4 & 0 & 1 \end{vmatrix}.$$

We will use the general formula along the third row. We have

$$\begin{vmatrix} 3 & 2 & 1 \\ 2 & 1 & -3 \\ 4 & 0 & 1 \end{vmatrix} = 4\begin{vmatrix} 2 & 1 \\ 1 & -3 \end{vmatrix} - 0\begin{vmatrix} 3 & 1 \\ 2 & -3 \end{vmatrix} + 1\begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix} = 4(-6-1)+1(3-4) = -29.$$

Which technique to evaluate a determinant is easier ? The answer depends on the person who is evaluating the determinant. Some like the elementary row operations and some like the general formula. All that matters is to get the correct answer.

Note that all of the above properties are still valid in the general case. Also you should remember that the concept of a determinant only exists for square matrices.

B.3. Determinant and Inverse of Matrices

Finding the inverse of a matrix is very important in many areas of science. For example, decrypting a coded message uses the inverse of a matrix. Determinant may be used to answer this problem. Indeed, let A be a square matrix. We know that A is invertible if and only if $\det(A) \neq 0$. Also if A has order n, then the cofactor $A_{i,j}$ is defined as the determinant of the square matrix of order (n-1) obtained from A by removing the row number i and the column number j multiplied by $(-1)^{i+j}$. Recall

$$\det(A) = \sum_{j=1}^{j=n} a_{ij}A_{ij}$$

for any fixed i, and

$$\det(A) = \sum_{i=1}^{i=n} a_{ij}A_{ij}$$

for any fixed j. Define the **adjoint** of A, denoted $adj(A)$, to be the transpose of the matrix whose ij^{th} entry is A_{ij}.

Example. Let

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

We have

$$adj(A) = \begin{pmatrix} -2 & 5 & -1 \\ 5 & -7 & 8 \\ 6 & -4 & 3 \end{pmatrix}^T = \begin{pmatrix} -2 & 5 & 6 \\ 5 & -7 & -4 \\ -1 & 8 & 3 \end{pmatrix}.$$

Let us evaluate $A \cdot adj(A)$. We have

$$A \cdot adj(A) = \begin{pmatrix} 1 & 3 & 2 \\ -1 & 0 & 2 \\ 3 & 1 & -1 \end{pmatrix}\begin{pmatrix} -2 & 5 & 6 \\ 5 & -7 & -4 \\ -1 & 8 & 3 \end{pmatrix} = \begin{pmatrix} 11 & 0 & 0 \\ 0 & 11 & 0 \\ 0 & 0 & 11 \end{pmatrix}.$$

Note that $\det(A) = 11$. Therefore, we have

$$A \cdot adj(A) = \det(A) I_3.$$

Is this formula only true for this matrix, or does a similar formula exist for any square matrix? In fact, we do have a similar formula.

Theorem. For any square matrix A of order n, we have

$$A \cdot adj(A) = \det(A) I_n.$$

In particular, if $\det(A) \neq 0$, then

$$A^{-1} = \frac{1}{\det(A)} adj(A).$$

For a square matrix of order 2, we have

$$adj \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}^T = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

which gives

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

This is a formula which we used on a previous page.

B.4. Application of Determinant to Systems: Cramer's Rule

We have seen that determinant may be useful in finding the inverse of a nonsingular matrix. We can use these findings in solving linear systems for which the matrix coefficient is nonsingular (or invertible).

Consider the linear system (in matrix form)

$$AX = B$$

where A is the matrix coefficient, B the nonhomogeneous term, and X the unknown column-matrix. We have:

Theorem. The linear system $AX = B$ has a unique solution if and only if A is invertible. In this case, the solution is given by the so-called **Cramer's formulas:**

$$x_i = \frac{\det(A_i)}{\det A}, \quad \text{for } i = 1, \cdots, n$$

where x_i are the unknowns of the system or the entries of X, and the matrix A_i is obtained from A by replacing the i^{th} column by the column B. In other words, we have

$$x_i = \frac{b_1 A_{1i} + b_2 A_{2i} + \cdots + b_n A_{ni}}{\det(A)}$$

where the b_i are the entries of B.

In particular, if the linear system $AX = B$ is homogeneous, meaning $B = \mathcal{O}$, then if A is invertible, the only solution is the trivial one, that is $X = \mathcal{O}$. So if we are looking for a nonzero solution to the system, the matrix coefficient A must be singular or noninvertible. We also know that this will happen if and only if $\det(A) = 0$. This is an important result.

Example. Solve the linear system

$$\begin{pmatrix} 1 & 2 & 0 \\ -1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Answer. First note that

$$\begin{vmatrix} 1 & 2 & 0 \\ -1 & 1 & 1 \\ 1 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 2 & 3 \end{vmatrix} - 2 \begin{vmatrix} -1 & 1 \\ 1 & 3 \end{vmatrix} = 9$$

which implies that the matrix coefficient is invertible. So we may use the Cramer's formulas. We have

$$x=\frac{1}{9}\begin{vmatrix} 0 & 2 & 0 \\ 1 & 1 & 1 \\ 0 & 2 & 3 \end{vmatrix},\ y=\frac{1}{9}\begin{vmatrix} 1 & 0 & 0 \\ -1 & 1 & 1 \\ 1 & 0 & 3 \end{vmatrix},\ \text{and } z=\frac{1}{9}\begin{vmatrix} 1 & 2 & 0 \\ -1 & 1 & 1 \\ 1 & 2 & 0 \end{vmatrix}.$$

We leave the details to the reader to find

$$x=\frac{-6}{9}=-\frac{2}{3},\ y=\frac{3}{9}=\frac{1}{3},\ \text{and } z=0.$$

Note that it is easy to see that $z=0$. Indeed, the determinant which gives z has two identical rows (the first and the last). We do encourage you to check that the values found for x, y, and z are indeed the solution to the given system.

Remark. Remember that Cramer's formulas are only valid for linear systems with an invertible matrix coefficient.

C. Eigenvalues and Eigenvectors

C.1. Eigenvalues and Eigenvectors: An Introduction

The eigenvalue problem is a problem of considerable theoretical interest and wide-ranging application. For example, this problem is crucial in solving systems of differential equations, analyzing population growth models, and calculating powers of matrices (in order to define the exponential matrix). Other areas such as physics, sociology, biology, economics and statistics have focused considerable attention on "eigenvalues" and "eigenvectors"-their applications and their computations. Before we give the formal definition, let us introduce these concepts on an example.

Example. Consider the matrix

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

Consider the three column matrices

$$C_1 = \begin{pmatrix} 1 \\ 6 \\ -13 \end{pmatrix}, \ C_2 = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, \text{ and } C_3 = \begin{pmatrix} 2 \\ 3 \\ -2 \end{pmatrix}.$$

We have

$$AC_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \ AC_2 = \begin{pmatrix} 4 \\ -8 \\ -4 \end{pmatrix}, \text{ and } AC_3 = \begin{pmatrix} 6 \\ 9 \\ -6 \end{pmatrix}.$$

In other words, we have

$$AC_1 = 0C_1, \ AC_2 = -4C_2, \ \text{ and } \ AC_3 = 3C_3.$$

Next consider the matrix P for which the columns are C_1, C_2, and C_3, i.e.,

$$P = \begin{pmatrix} 1 & -1 & 2 \\ 6 & 2 & 3 \\ -13 & 1 & -2 \end{pmatrix}.$$

We have $det(P) = 84$. So this matrix is invertible. Easy calculations give

$$P^{-1} = \frac{1}{84}\begin{pmatrix} -7 & 0 & -7 \\ -27 & 24 & 9 \\ 32 & 12 & 8 \end{pmatrix}.$$

Next we evaluate the matrix $P^{-1}AP$. We leave the details to the reader to check that we have

$$\frac{1}{84}\begin{pmatrix} -7 & 0 & -7 \\ -27 & 24 & 9 \\ 32 & 12 & 8 \end{pmatrix}\begin{pmatrix} 1 & 2 & 1 \\ 6 & -1 & 0 \\ -1 & -2 & -1 \end{pmatrix}\begin{pmatrix} 1 & -1 & 2 \\ 6 & 2 & 3 \\ -13 & 1 & -2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

In other words, we have

$$P^{-1}AP = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Using the matrix multiplication, we obtain

$$A = P\begin{pmatrix} 0 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 3 \end{pmatrix}P^{-1}$$

which implies that A is similar to a diagonal matrix. In particular, we have

$$A^n = P\begin{pmatrix} 0 & 0 & 0 \\ 0 & (-4)^n & 0 \\ 0 & 0 & 3^n \end{pmatrix}P^{-1}$$

for $n = 1, 2, \cdots$. Note that it is almost impossible to find A^{75} directly from the original form of A.

This example is so rich of conclusions that many questions impose themselves in a natural way. For example, given a square matrix A, how do we find column matrices which have similar behaviors as the above ones? In other words, how do we find these column matrices which will help find the invertible matrix P such that $P^{-1}AP$ is a diagonal matrix?

From now on, we will call column matrices **vectors**. So the above column matrices C_1, C_2, and C_3 are now vectors. We have the following definition.

Definition. Let A be a square matrix. A non-zero vector C is called an **eigenvector** of A if and only if there exists a number (real or complex) λ such that

$$AC = \lambda C.$$

If such a number λ exists, it is called an **eigenvalue** of A. The vector C is called eigenvector associated to the eigenvalue λ.

Remark. The eigenvector C must be non-zero since we have

$$AO = O = \lambda O$$

for any number λ.

Example. Consider the matrix

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

We have seen that

$$AC_1 = 0C_1,\ AC_2 = -4C_2,\ \text{and}\ AC_3 = 3C_3$$

where

$$C_1 = \begin{pmatrix} 1 \\ 6 \\ -13 \end{pmatrix},\ C_2 = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix},\ \text{and}\ C_3 = \begin{pmatrix} 2 \\ 3 \\ -2 \end{pmatrix}.$$

So C_1 is an eigenvector of A associated to the eigenvalue 0. C_2 is an eigenvector of A associated to the eigenvalue -4 while C_3 is an eigenvector of A associated to the eigenvalue 3.

It may be interesting to know whether we found all the eigenvalues of A in the above example. In the next page, we will discuss this question as well as how to find the eigenvalues of a square matrix.

C.2. Computation of Eigenvalues

For a square matrix A of order n, the number λ is an eigenvalue if and only if there exists a non-zero vector C such that

$$AC = \lambda C.$$

Using the matrix multiplication properties, we obtain

$$(A - \lambda I_n)C = 0.$$

This is a linear system for which the matrix coefficient is $A - \lambda I_n$. We also know that this system has one solution if and only if the matrix coefficient is invertible, i.e. $det(A - \lambda I_n) \neq 0$. Since the zero-vector is a solution and C is not the zero vector, then we must have

$$det(A - \lambda I_n) = 0.$$

Example. Consider the matrix

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

The equation $det(A - \lambda I_n) = 0$ translates into

$$\begin{vmatrix} 1-\lambda & -2 \\ -2 & 0-\lambda \end{vmatrix} = (1-\lambda)(0-\lambda) - 4 = 0$$

which is equivalent to the quadratic equation

$$\lambda^2 - \lambda - 4 = 0.$$

Solving this equation leads to

$$\lambda = \frac{1+\sqrt{17}}{2}, \text{ and } \lambda = \frac{1-\sqrt{17}}{2}.$$

In other words, the matrix A has only two eigenvalues.

In general, for a square matrix A of order n, the equation

$$det(A - \lambda I_n) = 0$$

will give the eigenvalues of A. This equation is called the **characteristic equation** or **characteristic polynomial** of A. It is a polynomial function in λ of degree n. So we know that this equation will not have more than n roots or solutions. So a square matrix A of order n will not have more than n eigenvalues.

Example. Consider the diagonal matrix

$$D = \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{pmatrix}.$$

Its characteristic polynomial is

$$det(D - \lambda I_n) = \begin{vmatrix} a-\lambda & 0 & 0 & 0 \\ 0 & b-\lambda & 0 & 0 \\ 0 & 0 & c-\lambda & 0 \\ 0 & 0 & 0 & d-\lambda \end{vmatrix} = (a-\lambda)(b-\lambda)(c-\lambda)(d-\lambda) = 0.$$

So the eigenvalues of D are a, b, c, and d, i.e. the entries on the diagonal.

This result is valid for any diagonal matrix of any size. So depending on the values you have on the diagonal, you may have one eigenvalue, two eigenvalues, or more. Anything is possible.

Remark. It is quite amazing to see that any square matrix A has the same eigenvalues as its transpose A^T because

$$det(A - \lambda I_n) = det(A - \lambda I_n)^T = det(A^T - \lambda I_n).$$

For any square matrix of order 2, A, where

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

the characteristic polynomial is given by the equation

$$\begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = (a-\lambda)(d-\lambda) - bc = \lambda^2 - (a+d)\lambda + ad - bc = 0.$$

The number $(a+d)$ is called the **trace** of A (denoted $tr(A)$), and clearly the number $(ad-bc)$ is the determinant of A. So the characteristic polynomial of A can be rewritten as

$$\lambda^2 - tr(A)\lambda + det(A) = 0.$$

Let us evaluate the matrix

$$B = A^2 - tr(A)\, A + det(A)\, I_2.$$

We have

$$B = \begin{pmatrix} a^2+bc & ab+bd \\ ac+dc & bc+d^2 \end{pmatrix} - (a+d)\begin{pmatrix} a & b \\ c & d \end{pmatrix} + (ad-bc)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

We leave the details to the reader to check that

$$B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

In other word, we have

$$A^2 - tr(A)A + det(A)I_2 = O.$$

This equation is known as the **Cayley-Hamilton** theorem. It is true for any square matrix A of any order, i.e.

$$p(A) = O$$

where $p(\lambda) = det(A - \lambda I_n)$ is the characteristic polynomial of A.

We have some properties of the eigenvalues of a matrix.

Theorem. Let A be a square matrix of order n. If λ is an eigenvalue of A, then:

1. λ^m is an eigenvalue of A^m, for $m = 1, 2, \cdots$
2. If A is invertible, then $\frac{1}{\lambda}$ is an eigenvalue of A^{-1}.
3. A is not invertible if and only if $\lambda = 0$ is an eigenvalue of A.
4. If α is any number, then $\lambda + \alpha$ is an eigenvalue of $A + \alpha I_n$.
5. If A and B are similar, then they have the same characteristic polynomial (which implies they also have the same eigenvalues).

C.3. Computation of Eigenvectors

Let A be a square matrix of order n and λ one of its eigenvalues. Let X be an eigenvector of A associated to λ. We must have

$$AX = \lambda X \text{ or } \left(A - \lambda I_n\right)X = 0.$$

This is a linear system for which the matrix coefficient is $A - \lambda I_n$. Since the zero-vector is a solution, the system is consistent. In fact, we will in a different page that the structure of the solution set of this system is very rich. In this page, we will basically discuss how to find the solutions.

Remark. It is quite easy to notice that if X is a vector which satisfies $AX = \lambda X$, then the vector $Y = c\,X$ (for any arbitrary number c) satisfies the same equation, i.e. $AY = \lambda Y$. In other words, if we know that X is an eigenvector, then cX is also an eigenvector associated to the same eigenvalue.

Let us start with an example.

Example. Consider the matrix

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

First we look for the eigenvalues of A. These are given by the characteristic equation
$\det(A - \lambda I_3) = 0$
, i.e.

$$\begin{vmatrix} 1-\lambda & 2 & 1 \\ 6 & -1-\lambda & 0 \\ -1 & -2 & -1-\lambda \end{vmatrix} = 0.$$

If we develop this determinant using the third column, we obtain

$$\begin{vmatrix} 6 & -1-\lambda \\ -1 & -2 \end{vmatrix} + (-1-\lambda)\begin{vmatrix} 1-\lambda & 2 \\ 6 & -1-\lambda \end{vmatrix} = 0.$$

Using easy algebraic manipulations, we get

$$-\lambda(\lambda+4)(\lambda-3) = 0$$

which implies that the eigenvalues of A are 0, -4, and 3.
Next we look for the eigenvectors.

1. Case $\lambda = 0$: The associated eigenvectors are given by the linear system

$$AX = \mathcal{O}$$

which may be rewritten by

$$\begin{cases} x + 2y + z & = & 0 \\ 6x - y & = & 0 \\ -x - 2y - z & = & 0 \end{cases}$$

Many ways may be used to solve this system. The third equation is identical to the first. Since, from the second equations, we have $y = 6x$, the first equation reduces to $13x + z = 0$. So this system is equivalent to

$$\begin{cases} y & = & 6x \\ z & = & -13x \end{cases}$$

So the unknown vector X is given by

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ 6x \\ -13x \end{pmatrix} = x \begin{pmatrix} 1 \\ 6 \\ -13 \end{pmatrix}.$$

Therefore, any eigenvector X of A associated to the eigenvalue 0 is given by

$$X = c \begin{pmatrix} 1 \\ 6 \\ -13 \end{pmatrix},$$

where c is an arbitrary number.

2.

Case $\lambda = -4$: The associated eigenvectors are given by the linear system

$$AX = -4X \text{ or } (A + 4I_3)X = O$$

which may be rewritten by

$$\begin{cases} 5x + 2y + z &= 0 \\ 6x + 3y &= 0 \\ -x - 2y + 3z &= 0 \end{cases}$$

In this case, we will use elementary operations to solve it. First we consider the augmented matrix $[A + 4I_3|O]$, i.e.

$$\left(\begin{array}{ccc|c} 5 & 2 & 1 & 0 \\ 6 & 3 & 0 & 0 \\ -1 & -2 & 3 & 0 \end{array}\right).$$

Then we use elementary row operations to reduce it to a upper-triangular form. First we interchange the first row with the first one to get

$$\left(\begin{array}{ccc|c} -1 & -2 & 3 & 0 \\ 5 & 2 & 1 & 0 \\ 6 & 3 & 0 & 0 \end{array}\right).$$

Next, we use the first row to eliminate the 5 and 6 on the first column. We obtain

$$\left(\begin{array}{ccc|c} -1 & -2 & 3 & 0 \\ 0 & -8 & 16 & 0 \\ 0 & -9 & 18 & 0 \end{array}\right).$$

If we cancel the 8 and 9 from the second and third row, we obtain

$$\left(\begin{array}{ccc|c} -1 & -2 & 3 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & -1 & 2 & 0 \end{array}\right).$$

Finally, we subtract the second row from the third to get

$$\left(\begin{array}{ccc|c} -1 & -2 & 3 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

Next, we set $z = c$. From the second row, we get $y = 2z = 2c$. The first row will imply $x = -2y+3z = -c$. Hence

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -c \\ 2c \\ c \end{pmatrix} = c\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}.$$

Therefore, any eigenvector X of A associated to the eigenvalue -4 is given by

$$X = c\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix},$$

where c is an arbitrary number.

2.

Case $\lambda = 3$: The details for this case will be left to the reader. Using similar ideas as the one described above, one may easily show that any eigenvector X of A associated to the eigenvalue 3 is given by

$$X = c\begin{pmatrix} 2 \\ 3 \\ -2 \end{pmatrix},$$

where c is an arbitrary number.

Remark. In general, the eigenvalues of a matrix are not all distinct from each other (see the page on the eigenvalues for more details). In the next two examples, we discuss this problem.

Example. Consider the matrix

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

The characteristic equation of A is given by

$$\begin{vmatrix} 3-\lambda & 2 & 4 \\ 2 & 0-\lambda & 2 \\ 4 & 2 & 3-\lambda \end{vmatrix} = -(\lambda+1)(\lambda+1)(\lambda-8) = -(\lambda+1)^2(\lambda-8) = 0.$$

Hence the eigenvalues of A are -1 and 8. For the eigenvalue 8, it is easy to show that any eigenvector X is given by

$$X = c\begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix},$$

where c is an arbitrary number. Let us focus on the eigenvalue -1. The associated eigenvectors are given by the linear system

$$AX = (-1)X \text{ or } (A + I_3)X = \mathcal{O}$$

which may be rewritten by

$$\begin{cases} 4x + 2y + 4z & = & 0 \\ 2x + y + 2z & = & 0 \\ 4x + 2y + 4z & = & 0 \end{cases}$$

Clearly, the third equation is identical to the first one which is also a multiple of the second equation. In other words, this system is equivalent to the system reduced to one equation

$$2x+y+2z= 0.$$

To solve it, we need to fix two of the unknowns and deduce the third one. For example, if we set $x = \alpha$ and $z = \beta$, we obtain $y = -2\alpha - 2\beta$. Therefore, any eigenvector X of A associated to the eigenvalue -1 is given by

$$X=\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}\alpha\\-2\alpha-2\beta\\\beta\end{pmatrix}=\alpha\begin{pmatrix}1\\-2\\0\end{pmatrix}+\beta\begin{pmatrix}0\\-2\\1\end{pmatrix}.$$

In other words, any eigenvector X of A associated to the eigenvalue -1 is a linear combination of the two eigenvectors

$$\begin{pmatrix}1\\-2\\0\end{pmatrix}\text{ and }\begin{pmatrix}0\\-2\\1\end{pmatrix}.$$

Example. Consider the matrix

$$\begin{pmatrix}1&-4\\4&-7\end{pmatrix}.$$

The characteristic equation is given by

$$\begin{vmatrix}1-\lambda&-4\\4&-7-\lambda\end{vmatrix}=(\lambda+3)^2=0.$$

Hence the matrix A has one eigenvalue, i.e. -3. Let us find the associated eigenvectors. These are given by the linear system

$$AX=(-3)X\ \text{ or }\ (A+3I_2)X=\mathcal{O}$$

which may be rewritten by

$$\begin{cases}4x-4y&=&0\\4x-4y&=&0\end{cases}$$

This system is equivalent to the one equation-system

$$x-y=0.$$

So if we set $x=c$, then any eigenvector X of A associated to the eigenvalue -3 is given by

$$X=\begin{pmatrix}x\\y\end{pmatrix}=c\begin{pmatrix}1\\1\end{pmatrix}.$$

Let us summarize what we did in the above examples.

Summary: *Let A be a square matrix. Assume λ is an eigenvalue of A. In order to find the associated eigenvectors, we do the following steps:*

1. Write down the associated linear system

$$AX = \lambda X \text{ or } (A - \lambda I_n)X = O.$$

2. Solve the system.
3. Rewrite the unknown vector X as a linear combination of known vectors.

The above examples assume that the eigenvalue λ is real number. So one may wonder whether any eigenvalue is always real. In general, this is not the case except for symmetric matrices. The proof of this is very complicated. For square matrices of order 2, the proof is quite easy. Let us give it here for the sake of being little complete.
Consider the symmetric square matrix

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

Its characteristic equation is given by

$$\det(A - \lambda I_2) = \begin{vmatrix} a-\lambda & b \\ b & c-\lambda \end{vmatrix} = \lambda^2 - (a+c)\lambda + ac - b^2 = 0.$$

This is a quadratic equation. The nature of its roots (which are the eigenvalues of A) depends on the sign of the discriminant

$$\Delta = (a+c)^2 - 4(ac - b^2).$$

Using algebraic manipulations, we get

$$\Delta = (a-c)^2 + 4b^2.$$

Therefore, Δ is a positive number which implies that the eigenvalues of A are real numbers.

Remark. Note that the matrix A will have one eigenvalue, i.e. one double root, if and only if $\Delta = 0$. But this is possible only if a=c and b=0. In other words, we have

$$A = a\, I_2.$$

C.4. Complex Eigenvalues

First let us convince ourselves that there exist matrices with complex eigenvalues.

Example. Consider the matrix

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

The characteristic equation is given by

$$\begin{vmatrix} 3-\lambda & -2 \\ 4 & -1-\lambda \end{vmatrix} = \lambda^2 - 2\lambda + 5 = 0.$$

This quadratic equation has complex roots given by

$$\lambda = \frac{2 \pm i\sqrt{16}}{2} = 1 \pm 2i$$

Therefore the matrix A has only complex eigenvalues.

The trick is to treat the complex eigenvalue as a real one. Meaning we deal with it as a number and do the normal calculations for the eigenvectors. Let us see how it works on the above example.

We will do the calculations for $\lambda = 1 + 2i$. The associated eigenvectors are given by the linear system

$$A\,X = (1+2i)\,X$$

which may be rewritten as

$$\begin{cases} (2-2i)x - 2y &= 0 \\ 4x - (2+2i)y &= 0 \end{cases}$$

In fact the two equations are identical since $(2+2i)(2-2i) = 8$. So the system reduces to one equation

$$(1-i)x - y = 0.$$

Set $x=c$, then $y = (1-i)c$. Therefore, we have

$$X = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c \\ c(1-i) \end{pmatrix} = c\begin{pmatrix} 1 \\ (1-i) \end{pmatrix}$$

where c is an arbitrary number.

Remark. It is clear that one should expect to have complex entries in the eigenvectors.

We have seen that $(1-2i)$ is also an eigenvalue of the above matrix. Since the entries of the matrix A are real, then one may easily show that if λ is a complex eigenvalue, then its conjugate $\bar{\lambda}$ is also an eigenvalue. Moreover, if X is an eigenvector of A associated to λ, then the vector $\bar{X}$, obtained from X by taking the complex-conjugate of the entries of X, is an eigenvector associated to $\bar{\lambda}$. So the eigenvectors of the above matrix A associated to the eigenvalue $(1-2i)$ are given by

$$X = c\begin{pmatrix} 1 \\ (1+i) \end{pmatrix}$$

where c is an arbitrary number.

Let us summarize what we did in the above example.

Summary: *Let A be a square matrix. Assume λ is a complex eigenvalue of A. In order to find the associated eigenvectors, we do the following steps:*

1. Write down the associated linear system
$$AX = \lambda X \text{ or } (A - \lambda I_n)X = O.$$
2. Solve the system. The entries of X will be complex numbers.
3. Rewrite the unknown vector X as a linear combination of known vectors with complex entries.
4. If A has real entries, then the conjugate $\bar{\lambda}$ is also an eigenvalue. The associated eigenvectors are given by the same equation found in 3, except that we should take the conjugate of the entries of the vectors involved in the linear combination.

In general, it is normal to expect that a square matrix with real entries may still have complex eigenvalues. One may wonder if there exists a class of matrices with only real eigenvalues. This is the case for symmetric matrices. The proof is very technical and will be discussed in another page. But for square matrices of order 2, the proof is quite easy. Let us give it here for the sake of being little complete.

Consider the symmetric square matrix

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

Its characteristic equation is given by

$$\det(A - \lambda I_2) = \begin{vmatrix} a-\lambda & b \\ b & c-\lambda \end{vmatrix} = \lambda^2 - (a+c)\lambda + ac - b^2 = 0.$$

This is a quadratic equation. The nature of its roots (which are the eigenvalues of A) depends on the sign of the discriminant

$$\Delta = (a+c)^2 - 4(ac - b^2).$$

Using algebraic manipulations, we get

$$\Delta = (a-c)^2 + 4b^2.$$

Therefore, Δ is a positive number which implies that the eigenvalues of A are real numbers.

Remark. Note that the matrix A will have one eigenvalue, i.e. one double root, if and only if $\Delta = 0$. But this is possible only if a=c and b=0. In other words, we have

$$A = a\, I_2.$$

C.5. Diagonalization

When we introduced eigenvalues and eigenvectors, we wondered when a square matrix is similarly equivalent to a diagonal matrix? In other words, given a square matrix A, does a diagonal matrix D exist such that $A \sim D$? (i.e. there exists an invertible matrix P such that $A = P^{-1}DP$)

In general, some matrices are not similar to diagonal matrices. For example, consider the matrix

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}.$$

Assume there exists a diagonal matrix D such that $A = P^{-1}DP$. Then we have

$$A - \lambda I_n = P^{-1}DP - \lambda I_n = P^{-1}DP - \lambda P^{-1}P = P^{-1}\left(D - \lambda I_n\right)P,$$

i.e $A - \lambda I_n$ is similar to $D - \lambda I_n$. So they have the same characteristic equation. Hence A and D have the same eigenvalues. Since the eigenvalues of D of the numbers on the diagonal, and the only eigenvalue of A is 2, then we must have

$$D = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = 2I_2.$$

In this case, we must have $A = P^{-1}DP = 2\,I_2$, which is not the case. Therefore, A is not similar to a diagonal matrix.

Definition. A matrix is **diagonalizable** if it is similar to a diagonal matrix.

Remark. In a previous page, we have seen that the matrix

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

has three different eigenvalues. We also showed that A is diagonalizable. In fact, there is a general result along these lines.

Theorem. Let A be a square matrix of order n. Assume that A has n distinct eigenvalues. Then A is diagonalizable. Moreover, if P is the matrix with the columns C_1, C_2, ..., and C_n the n eigenvectors of A, then the matrix $P^{-1}AP$ is a diagonal matrix. In other words, the matrix A is diagonalizable.

Problem: What happened to square matrices of order n with less than n eigenvalues?

We have a partial answer to this problem.

Theorem. Let A be a square matrix of order n. In order to find out whether A is diagonalizable, we do the following steps:

1. Write down the characteristic polynomial

$$p(\lambda) = \det\left(A - \lambda I_n\right).$$

2.

Factorize $p(\lambda)$. In this step, we should be able to get

$$p(\lambda) = (\lambda - \lambda_1)^{n_1} \cdot (\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_k)^{n_k}$$

where the λ_i, $i = 1, \cdots, k$, may be real or complex. For every i, the powers n_i is called the **(algebraic) multiplicity** of the eigenvalue λ_i.

3.

For every eigenvalue, find the associated eigenvectors. For example, for the eigenvalue λ_i, the eigenvectors are given by the linear system

$$A \cdot X = \lambda_i X \text{ or } \left(A - \lambda_i I_n\right) X = O.$$

Then solve it. We should find the unknown vector X as a linear combination of vectors, i.e.

$$X = \alpha_1 C_1 + \alpha_2 C_2 + \cdots + \alpha_{m_i} C_{m_i}$$

where α_j, $j = 1, \cdots, m_i$ are arbitrary numbers. The integer m_i is called the **geometric multiplicity** of λ_i.

4.

If for every eigenvalue the algebraic multiplicity is equal to the geometric multiplicity, then we have

$$m_1 + m_2 + \cdots + m_k = n$$

which implies that if we put the eigenvectors C_j, we obtained in **3.** for all the eigenvalues, we get exactly n vectors. Set P to be the square matrix of order n for which the column vectors are the eigenvectors C_j. Then P is invertible and

$$P^{-1} \cdot A \cdot P$$

is a diagonal matrix with diagonal entries equal to the eigenvalues of A. The position of the vectors C_j in P is identical to the position of the associated eigenvalue on the diagonal of D. This identity implies that A is similar to D. Therefore, A is diagonalizable.

Remark. If the algebraic multiplicity n_i of the eigenvalue λ_i is equal to 1, then obviously we have $m_i = 1$. In other words, $n_i = m_i$.

5.

If for some eigenvalue the algebraic multiplicity is not equal to the geometric multiplicity, then A is not diagonalizable.

Example. Consider the matrix

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

In order to find out whether A is diagonalizable, lt us follow the steps described above.

1.

The polynomial characteristic of A is

$$p(\lambda) = \begin{vmatrix} -1-\lambda & -1 & 1 \\ 0 & -2-\lambda & 1 \\ 0 & 0 & -1-\lambda \end{vmatrix} = (-1-\lambda)^2(-2-\lambda).$$

So -1 is an eigenvalue with multiplicity 2 and -2 with multiplicity 1.

2.

In order to find out whether A is diagonalizable, we only concentrate ur attention on the eigenvalue -1. Indeed, the eigenvectors associated to -1, are given by the system

$$\left(A + I_n\right)X = \begin{pmatrix} 0 & -1 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix} X = \mathcal{O}.$$

This system reduces to the equation $-y + z = 0$. Set $x = \alpha$ and $y = \beta$, then we have

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \\ \beta \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

So the geometric multiplicity of -1 is 2 the same as its algebraic multiplicity. Therefore, the matrix A is diagonalizable. In order to find the matrix P we need to find an eigenvector associated to -2. The associated system is

$$(A+2I_n)X = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} X = O$$

which reduces to the system

$$\begin{cases} x - y &= 0 \\ z &= 0 \end{cases}$$

Set $x = \alpha$, then we have

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \alpha \\ \alpha \\ 0 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Set

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

Then

$$P^{-1}AP = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

But if we set

$$P = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

then

$$P^{-1}AP = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

We have seen that if A and B are similar, then A^n can be expressed easily in terms of B^n. Indeed, if we have $A = P^{-1}BP$, then we have $A^n = P^{-1}B^nP$. In particular, if D is a diagonal matrix, D^n is easy to evaluate. This is one application of the diagonalization. In fact, the above procedure may be used to find the square root and cubic root of a matrix. Indeed, consider the matrix above

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

Set

$$P = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

then

$$P^{-1}AP = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = D.$$

Hence $A = P\,D\,P^{-1}$. Set

$$B = P\begin{pmatrix} -2^{1/3} & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}P^{-1},$$

Then we have

$$B^3 = A.$$

In other words, B is a cubic root of A.