GATE Material for Mathematics

Lesson 1:

Linear Algebra:

Introduction to Matrices:

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.

Introduction and Basic Operations:

Matrices, though they may appear weird objects at first, are a very important tool in expressing and discussing problems which arise from real life cases.

Our first example deals with economics. Indeed, consider two families A and B (though we may easily take more than two). Every month, the two families have expenses such as: utilities, health, entertainment, food, etc... Let us restrict ourselves to: food, utilities, and health. How would one represent the data collected? Many ways are available but one of them has an advantage of combining the data so that it is easy to manipulate them. Indeed, we will write the data as follows:

If we have no problem confusing the names and what the expenses are, then we may write

This is what we call a Matrix. The size of the matrix, as a block, is defined by the number of Rows and the number of Columns. In this case, the above matrix has 2 rows and 3 columns. You may easily come up with a matrix which has m rows and n columns. In this case, we say that the matrix is a (mxn) matrix (pronounce m-by-n matrix). Keep in mind that the first entry (meaning m) is the number of rows while the second entry (n) is the number of columns. Our above matrix is a (2x3) matrix.

When the numbers of rows and columns are equal, we call the matrix a square matrix. A square matrix of order n, is a (nxn) matrix.

Back to our example, let us assume, for example, that the matrices for the months of January, February, and March are

To make sure that the reader knows what these numbers mean, you should be able to give the Health-expenses for family A and Food-expenses for family B during the month of February. The answers are 250 and 600. The next question may sound easy to answer, but requires a new concept in the matrix context. Indeed, what is the matrix-expense for the two families for the first quarter? The idea is to add the three matrices above. It is easy to determine the total expenses for each family and each item, then the answer is

So how do we add matrices? An approach is given by the above example. The answer is to add entries one by one. For example, we have

Clearly, if you want to double a matrix, it is enough to add the matrix to itself. So we have

we get

which implies

This suggests the following rule

and for any number , we will have

Let us summarize these two rules about matrices.

Addition of Matrices: In order to add two matrices, we add the entries one by one.

Note: Matrices involved in the addition operation must have the same size.

Multiplication of a Matrix by a Number: In order to multiply a matrix by a number, you multiply every entry by the given number.

Keep in mind that we always write numbers to the left and matrices to the right (in the case of multiplication).

What about subtracting two matrices? It is easy, since subtraction is a combination of the two above rules. Indeed, if M and N are two matrices, then we will write

M-N = M + (-1)N

So first, you multiply the matrix N by -1, and then add the result to the matrix M.

Example. Consider the three matrices J, F, and M from above. Evaluate

Answer. We have

and since

we get

To compute J-M, we note first that

Since J-M = J + (-1)M, we get

And finally, for J-F+2M, we have a choice. Here we would like to emphasize the fact that addition of matrices may involve more than one matrix. In this case, you may perform the calculations in any order. This is called associativity of the operations. So first we will take care of -F and 2M to get

Since J-F+2M = J + (-1)F + 2M, we get

So first we will evaluate J-F to get

to which we add 2M, to finally obtain

For the addition of matrices, one special matrix plays a role similar to the number zero. Indeed, if we consider the matrix with all its entries equal to 0, then it is easy to check that this matrix has behavior similar to the number zero. For example, we have

and

Matrices, though they may appear weird objects at first, are a very important tool in expressing and discussing problems which arise from real life cases.

Our first example deals with economics. Indeed, consider twofamilies A and B (though we may easily take more than two). Every month, the two families have expenses such as: utilities, health, entertainment, food, etc... Let us restrict ourselves to: food, utilities, and health. How would one represent the data collected? Many ways are available but one of them has anadvantage of combining the data so that it is easy to manipulate them. Indeed, we will write the data as follows:

$$\text{Month} = \begin{pmatrix} \text{Family} & \text{Food} & \text{Utilities} & \text{Health} \\ A & a & b & c \\ B & d & e & f \end{pmatrix}$$

If we have no problem confusing the names and what theexpenses are, then we may write

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

This is what we call a Matrix. The size of the matrix, as a block, is defined by the number of Rows and the number ofColumns. In this case, the above matrix has 2 rows and 3 columns. You may easily come up with a matrix which has m rows and n columns. In this case, we say that the matrix is a(mxn) matrix (pronounce m-by-n matrix). Keep in mind that the first entry (meaning m) is the number of rows while the second entry (n) is the number of columns. Our above matrix is a (2x3) matrix.

When the numbers of rows and columns are equal, we call the matrix a square matrix. A square matrix of order n, is a (nxn) matrix.

Back to our example, let us assume, for example, that the matrices for the months of January, February, and March are

$$J = \begin{pmatrix} 600 & 250 & 350 \\ 550 & 180 & 400 \end{pmatrix}, \quad F = \begin{pmatrix} 650 & 330 & 250 \\ 600 & 270 & 400 \end{pmatrix}, \quad \text{and } M = \begin{pmatrix} 580 & 270 & 350 \\ 625 & 350 & 410 \end{pmatrix}, \text{ respectively.}$$

To make sure that the reader knows what these numbersmean, you should be able to give the Health-expenses for family A and Food-expenses for family B during the month of February. The answers are 250 and 600. The next question may sound easy to answer, but requires a new concept in the matrix context. Indeed, what is the matrix-expense for the twofamilies for the first quarter? The idea is to add the three matrices above. It is easy to determine the total expenses for each family and each item, then the answer is

$$\text{First Quarter} = \begin{pmatrix} 1830 & 850 & 950 \\ 1775 & 800 & 1210 \end{pmatrix}$$

So how do we add matrices? An approach is given by the above example. The answer is to add entries one by one. For example, we have

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

Clearly, if you want to double a matrix, it is enough to add the matrix to itself. So we have

we get

$$\text{double of } \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} + \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$$

which   implies

$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} + \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} = \begin{pmatrix} a+a & b+b & c+c \\ d+d & e+e & f+f \end{pmatrix} = \begin{pmatrix} 2a & 2b & 2c \\ 2d & 2e & 2f \end{pmatrix}.$$

This   suggests   the   following   rule

$$2 \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} = \begin{pmatrix} 2a & 2b & 2c \\ 2d & 2e & 2f \end{pmatrix}$$

and   for   any number λ, we   will have

$$\lambda \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} = \begin{pmatrix} \lambda a & \lambda b & \lambda c \\ \lambda d & \lambda e & \lambda f \end{pmatrix}$$

Let us summarize these two rules about matrices.

Addition of Matrices: In order to add two matrices, we add the entries one by one. Note: Matrices involved in the addition operation must have the same size.

Multiplication of a Matrix by a Number: In order to multiply a matrix by a number, you multiply every entry by the given number.

Keep in mind that we always write numbers to the left and matrices to the right (in the case of multiplication).

What about subtracting two matrices? It is easy, since subtraction is a combination of the two above   rules.   Indeed,   if M and N are   two   matrices,   then we   will write

M-$N = M + (-1)N$

So first, you multiply the matrix N by -1, and then add the result to the matrix M.

Example. Consider the three matrices J, F, and M from above. Evaluate

$$J + 2F, \ J - M, \ \text{and} \ J - F + 2M.$$

Answer. We have

$$J + 2F = \begin{pmatrix} 600 & 250 & 350 \\ 550 & 180 & 400 \end{pmatrix} + 2 \begin{pmatrix} 650 & 330 & 250 \\ 600 & 270 & 400 \end{pmatrix}$$

and   since

$$2\begin{pmatrix} 650 & 330 & 250 \\ 600 & 270 & 400 \end{pmatrix} = \begin{pmatrix} 1300 & 660 & 500 \\ 1200 & 540 & 800 \end{pmatrix}$$

we   get

$$J + 2F = \begin{pmatrix} 1900 & 910 & 850 \\ 1750 & 720 & 1200 \end{pmatrix}.$$

To compute *J-M*, we note first that

$$(-1)M = \begin{pmatrix} -580 & -270 & -350 \\ -625 & -350 & -410 \end{pmatrix}.$$

Since *J-M = J +*   (-1)*M,*   we   get

$$J-M = \begin{pmatrix} 600 & 250 & 350 \\ 550 & 180 & 400 \end{pmatrix} + \begin{pmatrix} -580 & -270 & -350 \\ -625 & -350 & -410 \end{pmatrix} = \begin{pmatrix} 20 & -20 & 0 \\ -75 & -170 & -10 \end{pmatrix}$$

And finally, for *J-F+2M*, we have a choice. Here we would like to emphasize the fact that addition of matrices may involve more than one matrix. In this case, you may perform the calculations in any order. This is called associativity of the operations. So first we will take care of   -*F* and   2*M* to   get

$$-F = (-1)F = \begin{pmatrix} -650 & -330 & -250 \\ -600 & -270 & -400 \end{pmatrix} \text{ and } 2M = \begin{pmatrix} 1160 & 540 & 700 \\ 1250 & 700 & 820 \end{pmatrix}.$$

Since *J-F+2M = J +*   (-1)*F +*   2*M,*   we   get

$$J-F+2M = \begin{pmatrix} 600 & 250 & 350 \\ 550 & 180 & 400 \end{pmatrix} + \begin{pmatrix} -650 & -330 & -250 \\ -600 & -270 & -400 \end{pmatrix} + \begin{pmatrix} 1160 & 540 & 700 \\ 1250 & 700 & 820 \end{pmatrix}.$$

So first we will evaluate *J-F* to get

$$J - F = \begin{pmatrix} -50 & -80 & 100 \\ -50 & -90 & 0 \end{pmatrix}$$

to which we add 2*M*, to finally obtain

$$J-F+2M = \begin{pmatrix} -50 & -80 & 100 \\ -50 & -90 & 0 \end{pmatrix} + \begin{pmatrix} 1150 & 540 & 700 \\ 1250 & 700 & 820 \end{pmatrix} = \begin{pmatrix} 1110 & 460 & 800 \\ 1200 & 610 & 820 \end{pmatrix}.$$

For the addition of matrices, one special matrix plays a role similar to the number zero. Indeed, if we consider the matrix with all its entries equal to 0, then it is easy to check that this matrix has behavior similar to the number zero. For example, we have

$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$$

and

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

What about multiplying two matrices? Such operation exists but the calculations involved are complicated.

What about multiplying two matrices? Such operation exists but the calculations involved are complicated.

Multiplication of Matrices:

Before we give the formal definition of how to multiply two matrices, we will discuss an example from a real life situation. Consider a city with two kinds of population: the inner citypopulation and the suburb population. We assume that every year 40% of the inner city population moves to the suburbs, while 30% of the suburb population moves to the inner part of the city. Let *I* (resp. *S*) be the initial population of the inner city (resp. the suburban area). So after one year, thepopulation of the inner part is

$0.6\,I + 0.3\,S$

while   the population of   the suburbs is
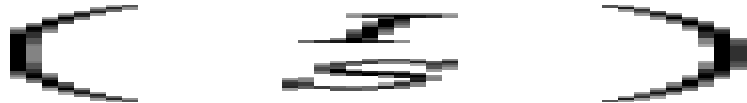
$0.4\,I + 0.7\,S$

After   two   years,   the population of   the   inner   city   is

$0.6\,(0.6\,I + 0.3\,S) + 0.3\,(0.4\,I + 0.7\,S)$

and   the   suburban population is   given   by

$0.4\,(0.6\,I + 0.3\,S) + 0.7(0.4\,I + 0.7\,S)$

Is there a nice way of representing the two populations after a certain number of years? Let us show how matrices may be helpful to answer this question. Let us represent the two populations in   one   table   (meaning   a   column   object   with   two   entries):

$$\begin{pmatrix} I \\ S \end{pmatrix}$$

So   after one   year the   table which   gives   the   two   populations   is

$$\begin{pmatrix} 0.6\,I + 0.3\,S \\ 0.4\,I + 0.7\,S \end{pmatrix}$$

If   we   consider   the   following   rule   (the   product of   two   matrices)

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} I \\ S \end{pmatrix} = \begin{pmatrix} aI + bS \\ cI + dS \end{pmatrix},$$

then the populations after one year are given by the formula

$$\begin{pmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{pmatrix} \begin{pmatrix} I \\ S \end{pmatrix}.$$

After two years the populations are

$$\begin{pmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{pmatrix} \left(\begin{pmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{pmatrix} \begin{pmatrix} I \\ S \end{pmatrix} \right).$$

Combining this formula with the above result, we get

$$\begin{pmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{pmatrix} \begin{pmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{pmatrix} = \begin{pmatrix} 0.6 \times 0.6 + 0.3 \times 0.4 & 0.6 \times 0.3 + 0.3 \times 0.7 \\ 0.4 \times 0.6 + 0.7 \times 0.4 & 0.4 \times 0.3 + 0.7 \times 0.7 \end{pmatrix}.$$

In other words, 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. Above, we did multiply a (2x2) matrix with a (2x1) matrix (which gave a (2x1) matrix). In fact, 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 abovemultiplication, 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.

Algebraic Properties of Matrix Operations:

In this page, we give some general results about the three operations: addition, multiplication, and multiplication withnumbers, 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.

where is the mxn zero-matrix (all its entries are equal to 0);

4.

                                                                        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 O the   mxk   zero-matrix,   then

$$A\,O = O$$

Note that $A0$ 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

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

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**. More on linear combinations will be discussed on a different page.

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

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

In particular, we have

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

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

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

Then it is easy to check that

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

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

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

In particular, we have

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

Invertible Matrices:

Invertible matrices are very important in many areas ofscience. For example, decrypting a coded message uses invertible matrices (see the coding page). The problem offinding the inverse of a matrix will be discussed in a different page.

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

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

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

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

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 $B = A^{-1}$.

Applications:

Application of Invertible Matrices: Coding

There are many ways to encrypt a message. And the use of coding has become particularly significant in recent years (due to the explosion of the internet for example). One way to encrypt or code a message uses matrices and their inverse. Indeed, consider a fixed invertible matrix A. Convert the message into a matrix B such that AB is possible to perform. Send the message generated by AB. At the other end, they will need to know A^{-1} in order to decrypt or decode the message sent. Indeed, we have

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

which is the original message. Keep in mind that whenever an undesired intruder finds A, we must be able to change it. Sowe should have a mechanical way of generating simple matrices A which are invertible and have simple inverse matrices. Note that, in general, the inverse of a matrix involves fractions which are not easy to send in an electronic form. The best is to have both A and its inverse with integers as their entries. In fact, we can use our previous knowledge to generate such class of matrices. Indeed, if A is a matrix such that its determinant is ± 1 and all its entries are integers, then A^{-1} will have entries which are integers. So how do we generate such class of matrices? One practical way is to start with an upper triangular matrix with ± 1 on the diagonal and integer-entries. Then we use the elementary row operations to change the matrix while keeping the

determinant unchanged. Do not multiply rows with non-integers while doing elementary row operations. Let us illustrate this on an example.

Example. Consider the matrix

$$\begin{pmatrix} -1 & 5 & -1 \\ 0 & 1 & 8 \\ 0 & 0 & 1 \end{pmatrix}.$$

First we keep the first row and add it to the second as well as to the third rows. We obtain

$$\begin{pmatrix} -1 & 5 & -1 \\ -1 & 6 & 7 \\ -1 & 5 & 0 \end{pmatrix}.$$

Next we keep the first row again, we add the second to the third, and finally add the last one to the first multiplied by -2. We obtain

$$\begin{pmatrix} -1 & 5 & -1 \\ -2 & 11 & 7 \\ 1 & -5 & 2 \end{pmatrix}.$$

This is our matrix A. Easy calculations will give $det(A) = -1$, which we knew since the above elementary operations did not change the determinant from the original triangular matrix which obviously has -1 as its determinant. We leave the details of the calculations to the reader. The inverse of A is

$$\begin{pmatrix} 57 & -5 & 46 \\ 11 & -1 & 9 \\ -1 & 0 & -1 \end{pmatrix}.$$

Back to our original problem. Consider the message

I LOVE MONICA

To every letter we will associate a number. The easiest way to do that is to associate 0 to a blank or space, 1 to A, 2 to B, etc... Another way is to associate 0 to a blank or space, 1 to A, -1 to B, 2 to C, -2 to D, etc... Let us use the second choice. So our message is given by the string

$$\begin{array}{ccccccccccccc} I & & L & O & & V & E & & M & O & N & I & C & A \\ 5 & 0 & -6 & 8 & -11 & 3 & 0 & 7 & 8 & -7 & 5 & 2 & 1 \end{array}$$

Now we rearrange these numbers into a matrix B. For example, we have

$$B = \begin{pmatrix} 5 & 8 & 0 & -7 & 1 \\ 0 & -11 & 7 & 5 & 0 \\ -6 & 3 & 8 & 2 & 0 \end{pmatrix}.$$

Then we perform the product AB, where A is the matrix found above. We get

$$AB = \begin{pmatrix} -1 & 5 & -1 \\ -2 & 11 & 7 \\ 1 & -5 & 2 \end{pmatrix} \begin{pmatrix} 5 & 8 & 0 & -7 & 1 \\ 0 & -11 & 7 & 5 & 0 \\ -6 & 3 & 8 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -66 & 27 & 30 & -1 \\ -52 & -116 & 133 & 83 & -2 \\ -7 & 69 & -19 & -28 & 1 \end{pmatrix}.$$

The encrypted message to be sent is

$$1, \ -52, \ -7, \ -66, \ \cdots, \ -1, \ -2, \ 1$$

Complex numbers as Matrices:

In this section, we use matrices to give a representation of complex numbers. Indeed, consider the set

$$C = \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} ; \text{ where } a \text{ and } b \text{ are real numbers.} \right\}.$$

We will write

$$M_{a,b} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

Clearly, the set C is not empty. For example, we have

$$M_{0,0} = O, \text{ and } M_{1,0} = I_2.$$

In particular, we have

$$M_{a,b} = M_{c,d} \text{ if and only if } a = c \text{ and } b = d$$

for any real numbers a, b, c, and d.

Algebraic Properties of $\mathcal{C}$

1.

Addition: For any real numbers a, b, c, and d, we have

$M_{a,b} + M_{c,d} = M_{a+c,b+d}$.

In other words, if we add two elements of the set $\mathcal{C}$, we still get a matrix in $\mathcal{C}$. In particular, we have

$-M_{a,b} = M_{-a,-b}$.

2.

Multiplication by a number: We have

$$\lambda M_{a,b} = \lambda \begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} \lambda a & \lambda b \\ -\lambda b & \lambda a \end{pmatrix} = M_{\lambda a, \lambda b}.$$

So a multiplication of an element of $\mathcal{C}$ and a number gives a matrix in $\mathcal{C}$.

2.

Multiplication: For any real numbers a, b, c, and d, we have

$$M_{a,b} M_{c,d} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} ac - bd & ad + bc \\ -bc - ad & ac - bd \end{pmatrix}.$$

In other words, we have

$$M_{a,b}\, M_{c,d} = M_{ac-bd,\ ad+bc}.$$

This is an extraordinary formula. It is quite conceivable given the difficult form of the matrix multiplication that, a priori, the product of two elements of $\mathcal{C}$ may not be in $\mathcal{C}$ again. But, in this case, it turns out to be true.

The above properties infer to $\mathcal{C}$ a very nice structure. The next natural question to ask, in this case, is whether a nonzero element of $\mathcal{C}$ is invertible. Indeed, for any real numbers a and b, we have

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \begin{pmatrix} a^2+b^2 & 0 \\ 0 & a^2+b^2 \end{pmatrix} = (a^2+b^2)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

$$a^2 + b^2 \neq 0$$

So, if $\qquad\qquad$, the matrix $M_{a,b}$ is invertible and

$$M_{a,b}^{-1} = \frac{1}{a^2+b^2}\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \frac{1}{a^2+b^2}M_{a,-b}$$

In other words, any nonzero element $M_{a,b}$ of $\mathcal{C}$ is invertible and its inverse is still in $\mathcal{C}$ since

$$M_{a,b}^{-1} = M_{\frac{a}{a^2+b^2},\ \frac{-b}{a^2+b^2}}$$

In order to define the division in $\mathcal{C}$, we will use the inverse. Indeed, recall that

$$a \div b = \frac{a}{b} = a \cdot \frac{1}{b} = a \cdot b^{-1}$$

So for the set $\mathcal{C}$, we have

$M_{a,\ b} \div M_{c,\ d} = M_{a,\ b} \times M_{c,\ d}^{-1} = M_{a,\ b} \times M$

which implies

$$M_{a,b} \div M_{c,d} = M$$,   -

The matrix $M_{a,-b}$ is called the conjugate of $M_{a,b}$. Note that the conjugate of the conjugate of $M_{a,b}$ is $M_{a,b}$ itself.

Fundamental Equation. For any $M_{a,b}$ in $\mathcal{C}$, we have

$$M_{a,b} = a\, M_{1,0} + b\, M_{0,1} = a\, I_2 + b\, M_{0,1}.$$

Note that

$$M_{0,1}\, M_{0,1} = M_{-1,0} = -I_2.$$

Remark. If we introduce an imaginary number i such that $i^2 = -1$, then the matrix $M_{a,b}$ may be rewritten by

$$a + bi$$

A lot can be said about $\mathcal{C}$, but we will advise you to visit the page on complex numbers.

Markov Chains:

In a previous page, we studied the movement between the city and suburbs. Indeed, if I are S are the initial populationof the inner city and the suburban area, and if we assume that every year 40% of the inner city population moves to thesuburbs, while 30% of the suburb population moves to the inner part of the city, then after one year the populations are given by

$$\begin{pmatrix} 0.6I + 0.3S \\ 0.4I + 0.7S \end{pmatrix} = \begin{pmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{pmatrix} \begin{pmatrix} I \\ S \end{pmatrix}.$$

The matrix

$$P = \begin{pmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{pmatrix}$$

is very special. Indeed, the entries of each column vectors are positive and their sum is 1. Such vectors are calledprobability vectors. A matrix for which all the column vectors are probability vectors is called transition orstochastic matrix. Andrei Markov, a russian mathematician, was the first one to study these matrices. At the beginning of this century he developed the fundamentals of the Markov Chain theory.
A Markov chain is a process that consists of a finite number ofstates and some known probabilities p_{ij}, where p_{ij} is theprobability of moving from state j to state i. In the example above, we have two states: living in the city and living in thesuburbs. The number p_{ij} represents the probability of moving from state i to state j in one year. We may have more than two states. For example, political affiliation: Democrat, Republican, and Independent. For example, p_{ij} represents theprobability of a son belonging to party i if his father belonged to party j.
Of particular interest is a probability vector p such that

$$Ap = p$$

, that is, an eigenvector of Aassociated to the eigenvalue 1. Such vector is called asteady state vector. In the example above, the steady state vectors are given by the system

$$\begin{pmatrix} 0.6 & 0.3 \\ 0.4 & 0.7 \end{pmatrix} \cdot X = X \quad \text{or} \quad \begin{pmatrix} -0.4 & 0.3 \\ 0.4 & -0.3 \end{pmatrix} \cdot X = \mathcal{O}.$$

This system reduces to the equation $-0.4\,x + 0.3\,y = 0$. It is easy to see that, if we

set $$x = 0.3\alpha$$, then

$$X = \begin{pmatrix} x \\ y \end{pmatrix} = \alpha \begin{pmatrix} 0.3 \\ 0.4 \end{pmatrix}.$$

$$\mathbf{P_1} = \begin{pmatrix} 0.3 \\ 0.4 \end{pmatrix}$$

So the vector is a steady state vector of the matrix above. So if the populations of the city and the suburbs are given by the vector $\mathbf{P_1}$, after one year the proportions remain the same (though the people may move between the city and the suburbs).

Let us discuss another example on population dynamics.

Example: Age Distribution of Trees in a Forest
Trees in a forest are assumed in this simple model to fall into four age groups: $b(k)$ denotes the number of baby trees in the forest (age group 0-15 years) at a given time period k; similarly $y(k), m(k)$ and $o(k)$ denote the number of young trees (16-30 years of age), middle-aged trees (age 31-45), and old trees (older than 45 years of age), respectively. The length of one time period is 15 years. How does the age distribution change from one time period to the next? The model makes the following three assumptions:

- A certain percentage of trees in each age group dies.

- Surviving trees enter into the next age group; old trees remain old.

- Lost trees are replaced by baby trees.

Note that the total tree population does not change over time.

We obtain the following difference equations:

$$b(k+1) \quad = d_b \cdot b(k) + d_y \cdot y(k) + d_m \cdot m(k) + d_o \cdot o(k) \tag{1}$$

$$y(k+1) \quad = (1-d_b)\, b(k) \tag{2}$$

$$m(k+1) \quad = (1-d_y)\, y(k) \tag{3}$$

$$o(k+1) \quad = (1-d_m)\, m(k) + (1-d_o)\, o(k) \tag{4}$$

Here $0 < d_b, d_y, d_m, d_o < 1$ denote the loss rates in each age group in percent.

Let

$$x(k) = \begin{pmatrix} b(k) \\ y(k) \\ m(k) \\ o(k) \end{pmatrix}$$

be the ``age distribution vector''. Consider the matrix

$$A = \begin{pmatrix} d_b & d_y & d_m & d_o \\ 1 - d_b & 0 & 0 & 0 \\ 0 & 1 - d_y & 0 & 0 \\ 0 & 0 & 1 - d_m & 1 - d_o \end{pmatrix}.$$

Then we have

$$\mathbf{x}(k+1) = A \cdot \mathbf{x}(k).$$

Note that the matrix A is a stochastic matrix! If, $d_b=.1, d_y=.2, d_m=.3$ and $d_o=.4$, then

$$A = \begin{pmatrix} 0.1 & 0.2 & 0.3 & 0.4 \\ 0.9 & 0 & 0 & 0 \\ 0 & 0.8 & 0 & 0 \\ 0 & 0 & 0.7 & 0.6 \end{pmatrix}.$$

After easy calculations, we find the steady state vector for the age distribution in the forest:

$$\mathbf{p} = \begin{pmatrix} 1/3.88 \\ 0.9/3.88 \\ 0.72/3.88 \\ 1.26/3.88 \end{pmatrix} = \frac{1}{3.88} \begin{pmatrix} 1 \\ 0.9 \\ 0.72 \\ 1.26 \end{pmatrix}.$$

Assume a total tree population of 50,000 trees. Suppose the forest is newly planted, i.e.

$$\mathbf{x}(0) = \begin{pmatrix} 50,000 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

After 15 years, the age distribution in the forest is given by

$$\mathbf{x}(1) = \begin{pmatrix} 0.1 & 0.2 & 0.3 & 0.4 \\ 0.9 & 0 & 0 & 0 \\ 0 & 0.8 & 0 & 0 \\ 0 & 0 & 0.7 & 0.6 \end{pmatrix} \mathbf{x}(0) = 50,000 \begin{pmatrix} 0.1 \\ 0.9 \\ 0 \\ 0 \end{pmatrix}.$$

After 30 years, we have

$$\mathbf{x}(2) = \begin{pmatrix} 0.1 & 0.2 & 0.3 & 0.4 \\ 0.9 & 0 & 0 & 0 \\ 0 & 0.8 & 0 & 0 \\ 0 & 0 & 0.7 & 0.6 \end{pmatrix} \mathbf{x}(1) = 50,000 \begin{pmatrix} 0.19 \\ 0.09 \\ 0.72 \\ 0 \end{pmatrix}$$

and after 45 years

$$\mathbf{x}(3) = 50{,}000 \begin{pmatrix} 0.19 \\ 0.09 \\ 0.72 \\ 0 \end{pmatrix}$$

$$n = 1, 2, \cdots$$

After 15n years, where   , the age distribution in the forest is given by

$$\mathbf{x}(n) = \begin{pmatrix} 0.1 & 0.2 & 0.3 & 0.4 \\ 0.9 & 0 & 0 & 0 \\ 0 & 0.8 & 0 & 0 \\ 0 & 0 & 0.7 & 0.6 \end{pmatrix}^{n} \mathbf{x}(0) = 50{,}000 \begin{pmatrix} 0.1 & 0.2 & 0.3 & 0.4 \\ 0.9 & 0 & 0 & 0 \\ 0 & 0.8 & 0 & 0 \\ 0 & 0 & 0.7 & 0.6 \end{pmatrix}^{n} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

So the problem is to find the nth-power of the matrix A. We have seen that diagonalization technique may be helpful to solve this problem. Another problem deals with the long term behavior of the sequence x(n) when n gets large.

The calculations on the example above becomes tedious. Let us illustrate the problem on a small matrix.

Example. Consider the stochastic matrix

$$A = \begin{pmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{pmatrix}.$$

Note this is a symmetric matrix. The characteristic polynomial of A is

$$p(\lambda) = (0.8 - \lambda)^2 - 0.2^2 = (1 - \lambda)(0.6 - \lambda)$$

An eigenvector associated to 1 is

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

and an eigenvector associated to 0.6 is

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

If we set

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

then we have

$$P^{-1}AP = D = \begin{pmatrix} 1 & 0 \\ 0 & 0.6 \end{pmatrix}.$$

So, we have

$$A^n = PD^nP^{-1} = P \begin{pmatrix} 1 & 0 \\ 0 & (0.6)^n \end{pmatrix} P^{-1}.$$

When n gets large, the matrices A^n get closer to the matrix

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

So the sequence of vectors defined by

$$X(n+1) = AX(n), \quad \text{given } X(0)$$

will get closer to

$$X(\infty) = P \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} P^{-1}X(0)$$

$$X(0) = \begin{pmatrix} a \\ b \end{pmatrix}$$

when n gets large. If $\quad$, then we have

$$X(\infty) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \frac{a+b}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

$$X(\infty)$$

Note that the vector $\quad$ is proportional to the unique steady state vector of A

$$P = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

This is not surprising. In fact there is a general result similar to the one above for any stochastic matrix.

Systems of Linear Equations: Introduction

Many books on linear algebra will introduce matrices via systems of linear equations. We tried a different approach. We hope this way you will appreciate matrices as a powerful tool useful not only to solve linear systems of equations. Basically, the problem of finding some unknowns linked to each others via equations is called a system of equations.

For example,

$$\begin{cases} 2x - 3y &= 1 \\ x + 3y &= -2 \end{cases}$$

and

$$\begin{cases} x^2 + y^2 &= 1 \\ x + 3y &= -2 \end{cases}$$

are systems of two equations with two unknowns (x and y), while

$$\begin{cases} 2x - 3y^2 &= -1 \\ x + y + z &= 1 \end{cases}$$

is a system of two equations with three unknowns (x, y, and z).

These systems of equations occur naturally in many real life problems. For example, consider a nutritious drink which consists of whole egg, milk, and orange juice. The food energy and protein for each of the ingredients are given by the table:

	Food Energy (Calories)	Protein (Grams)
1 egg	80	6
1 cup milk	160	9
1 cup orange juice	110	2

A natural question to ask is how much of each ingredient do we need to produce a drink of 540 calories and 25 grams of protein. In order to answer that, let x be the number of eggs, y the amount of milk (in cups), and z the amount of orange of juice (in cups). Then we need to have

$$\begin{cases} 80x + 160y + 110z = 540 \\ 6x + 9y + 2z = 25 \end{cases}$$

The task of **Solving** a system consists of finding the unknowns, here: x, y and z. A solution is a set of numbers once substituted for the unknowns will satisfy the equations of the system. For example, (2,1,2) and (0.325, 2.25, 1.4) are solutions to the system above.

The fundamental problem associated to any system is to find all the solutions. One way is to study the structure of its set of solutions which, in some cases, may help finding the solutions. Indeed, for example, in order to find the solutions to a linear system, it is enough to find just a few of them. This is possible because of the rich structure of the set of solutions.

Systems of Linear Equations: Gaussian Elimination

It is quite hard to solve non-linear systems of equations, while linear systems are quite easy to study. There are numerical techniques which help to approximate nonlinear systems with linear ones in the hope that the solutions of the linear systems are close enough to the solutions of the nonlinear systems. We will not discuss this here. Instead, we will focus our attention on linear systems.

For the sake of simplicity, we will restrict ourselves to three, at most four, unknowns. The reader interested in the case of more unknowns may easily extend the following ideas.

Definition. The equation

$$a\,x + b\,y + c\,z + d\,w = h$$

where a, b, c, d, and h are known numbers, while x, y, z, and w are unknown numbers, is called a **linear equation**. If $h=0$, the linear equation is said to be **homogeneous**. A **linear system** is a set of linear equations and a **homogeneous linear system** is a set of homogeneous linear equations.

For example,

$$\begin{cases} 2x - 3y &= 1 \\ x + 3y &= -2 \end{cases}$$

and

$$\begin{cases} x + y - z &= 1 \\ x + 3y + 3z &= -2 \end{cases}$$

are linear systems, while

$$\begin{cases} 2x - 3y^2 &= -1 \\ x + y + z &= 1 \end{cases}$$

is a nonlinear system (because of y^2). The system

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

is an homogeneous linear system.

Matrix Representation of a Linear System

Matrices are helpful in rewriting a linear system in a very simple form. The algebraic properties of matrices may then be used to solve systems. First, consider the linear system

$$\begin{cases} ax + by + cz + dw & = & e \\ fx + gy + hz + iw & = & j \\ kx + ly + mz + nw & = & p \\ qx + ry + sz + tw & = & u \end{cases}.$$

Set the matrices

$$A = \begin{pmatrix} a & b & c & d \\ f & g & h & i \\ k & l & m & n \\ q & r & s & t \end{pmatrix}, \quad C = \begin{pmatrix} e \\ j \\ p \\ u \end{pmatrix}, \quad \text{and } X = \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}.$$

Using matrix multiplications, we can rewrite the linear system above as the matrix equation

$$A \cdot X = C.$$

As you can see this is far nicer than the equations. But sometimes it is worth to solve the system directly without going through the matrix form. The matrix A is called the matrix coefficient of the linear system. The matrix C is called the **nonhomogeneous term**. When $C = O$, the linear system is homogeneous. The matrix X is the unknown matrix. Its entries are the unknowns of the linear system. The**augmented matrix** associated with the system is the matrix $[A|C]$, where

$$[A|C] == \begin{pmatrix} a & b & c & d & e \\ f & g & h & i & j \\ k & l & m & n & p \\ q & r & s & t & u \end{pmatrix}.$$

In general if the linear system has n equations with munknowns, then the matrix coefficient will be a nxm matrix and the augmented matrix an nx(m+1) matrix. Now we turn our attention to the solutions of a system.

Definition. Two linear systems with n unknowns are said to be **equivalent** if and only if they have the same set of solutions.

This definition is important since the idea behind solving a system is to find an equivalent system which is easy to solve. You may wonder how we will come up with such system? Easy, we do that through **elementary operations**. Indeed, it is clear that if we interchange two equations, the new system is still equivalent to the old one. If we multiply anequation with a nonzero number, we obtain a new system still equivalent to old one. And finally replacing one equation with the sum of two equations, we again obtain an equivalent system. These operations are called **elementary operations** on systems. Let us see how it works in a particular case.

Example. Consider the linear system

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

The idea is to keep the first equation and work on the last two. In doing that, we will try to kill one of the unknowns and solve for the other two. For example, if we keep the first and second equation, and subtract the first one from the last one, we get the equivalent system

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

Next we keep the first and the last equation, and we subtract the first from the second. We get the equivalent system

$$\begin{cases} x + & y + z & = & 0 \\ - & 3y + z & = & 4 \\ & y - 2z & = & 2 \end{cases}$$

Now we focus on the second and the third equation. We repeat the same procedure. Try to kill one of the twounknowns (*y* or *z*). Indeed, we keep the first and secondequation, and we add the second to the third after multiplying it by 3. We get

$$\begin{cases} x+ & y+ & z & = & 0 \\ - & 3y+ & z & = & 4 \\ & - & 5z & = & 10 \end{cases}$$

This obviously implies $z = $ -2. From the second equation, we get $y = $ -2, and finally from the first equation we get $x = 4$. Therefore the linear system has one solution

$$x = 4, \ y = -2, \ z = -2.$$

Going from the last equation to the first while solving for theunknowns is called **backsolving**.

Keep in mind that linear systems for which the matrix coefficient is upper-triangular are easy to solve. This is particularly true, if the matrix is in echelon form. So the trick is to perform elementary operations to transform the initial linear system into another one for which the coefficient   matrix   is   in   echelon   form. Using our knowledge about matrices, is there anyway we can rewrite what we did above in matrix form which will make our notation (or representation) easier? Indeed, consider the augmented matrix

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & -2 & 2 & 4 \\ 1 & 2 & -1 & 2 \end{array} \right).$$

Let us perform some elementary row operations on this matrix. Indeed, if we keep the first and second row, and subtract the first one from the last one we get

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & -2 & 2 & 4 \\ 0 & 1 & -2 & 2 \end{array} \right).$$

Next we keep the first and the last rows, and we subtract the first from the second. We get

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & -3 & 1 & 4 \\ 0 & 1 & -2 & 2 \end{array} \right).$$

Then we keep the first and second row, and we add the second to the third after multiplying it by 3 to get

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & -3 & 1 & 4 \\ 0 & 0 & -5 & 10 \end{pmatrix}.$$

This is a triangular matrix which is not in echelon form. The linear system for which this matrix is an augmented one is

$$\begin{cases} x + y + z = 0 \\ - 3y + z = 4 \\ - 5z = 10 \end{cases}$$

As you can see we obtained the same system as before. In fact, we followed the same elementary operations performedabove. In every step the new matrix was exactly the augmented matrix associated to the new system. This shows that instead of writing the systems over and over again, it is easy to play around with the elementary row operations and once we obtain a triangular matrix, write the associated linear system and then solve it. This is known as **Gaussian Elimination**. Let us summarize the procedure:

Gaussian Elimination. Consider a linear system.

1.

Construct the augmented matrix for the system;

2.

Use elementary row operations to transform the augmented matrix into a triangular one;

3.

Write down the new linear system for which the triangular matrix is the associated augmented matrix;

4.

Solve the new system. You may need to assign some parametric values to some unknowns, and then apply the method of back substitution to solve the new system.

Example. Solve the following system via Gaussian elimination

$$\begin{cases} 2x- & 3y-z+ & 2w & +3v & = & 4 \\ 4x- & 4y-z+ & 4w & +11v & = & 4 \\ 2x- & 5y-2z+ & 2w & -v & = & 9 \\ & 2y+z & & +4v & = & -5 \end{cases}$$

The augmented matrix is

$$\left(\begin{array}{ccccc|c} 2 & -3 & -1 & 2 & 3 & 4 \\ 4 & -4 & -1 & 4 & 11 & 4 \\ 2 & -5 & -2 & 2 & -1 & 9 \\ 0 & 2 & 1 & 0 & 4 & -5 \end{array} \right).$$

We use elementary row operations to transform this matrix into a triangular one. We keep the first row and use it to produce all zeros elsewhere in the first column. We have

$$\left(\begin{array}{ccccc|c} 2 & -3 & -1 & 2 & 3 & 4 \\ 0 & 2 & 1 & 0 & 5 & -4 \\ 0 & -2 & -1 & 0 & -4 & 5 \\ 0 & 2 & 1 & 0 & 4 & -5 \end{array} \right).$$

Next we keep the first and second row and try to have zeros in the second column. We get

$$\left(\begin{array}{ccccc|c} 2 & -3 & -1 & 2 & 3 & 4 \\ 0 & 2 & 1 & 0 & 5 & -4 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & -1 & -1 \end{array} \right).$$

Next we keep the first three rows. We add the last one to the third to get

$$\left(\begin{array}{ccccc|c} 2 & -3 & -1 & 2 & 3 & 4 \\ 0 & 2 & 1 & 0 & 5 & -4 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right).$$

This is a triangular matrix. Its associated system is

$$\begin{cases} 2x- & 3y & - & z & + & 2w & + & 3v & = & 4 \\ & 2y & + & z & & & + & 5v & = & -4 \\ & & & & & & & v & = & 1 \end{cases}$$

Clearly we have $v = 1$. Set $z=s$ and $w=t$, then we have

$$y = -2 - \frac{1}{2}z - \frac{5}{2}v = -\frac{9}{2} - \frac{1}{2}s.$$

The first equation implies

$$x = 2 + \frac{3}{2}y + \frac{1}{2}z - w - \frac{3}{2}v.$$

Using algebraic manipulations, we get

$$x = -\frac{25}{4} - \frac{1}{4}s - t.$$

Putting all the stuff together, we have

$$
\begin{pmatrix} x \\ y \\ z \\ w \\ v \end{pmatrix}
=
\begin{pmatrix} -\dfrac{25}{4} - \dfrac{1}{4}s - t \\ -\dfrac{9}{2} - \dfrac{1}{2}s \\ s \\ t \\ 1 \end{pmatrix}.
$$

Example. Use Gaussian elimination to solve the linear system

$$
\begin{cases}
x - y & = & 4 \\
2x - 2y & = & -4
\end{cases}
$$

The associated augmented matrix is

$$
\left(\begin{array}{cc|c} 1 & -1 & 4 \\ 2 & -2 & -4 \end{array} \right).
$$

We keep the first row and subtract the first row multiplied by 2 from the second row. We get

$$\begin{pmatrix} 1 & -1 & \bigm| & 4 \\ 0 & 0 & \bigm| & -12 \end{pmatrix}.$$

This is a triangular matrix. The associated system is

$$\begin{cases} x & - & y & = & 4 \\ & & 0 & = & -12 \end{cases}$$

Clearly the second equation implies that this system has no solution. Therefore this linear system has no solution.

Definition. A linear system is called **inconsistent or overdetermined** if it does not have a solution. In other words, the set of solutions is empty. Otherwise the linear system is called **consistent**.

Following the example above, we see that if we perform elementary row operations on the augmented matrix of the system and get a matrix with one of its rows equal to $(0, 0, \cdots, 0, c)$, where $c \neq 0$, then the system is inconsistent.

Systems of linear equations in Two variables:

A system of equations is a collection of two or more equations with the same set of unknowns. In solving a system of equations, we try to find values for each of the unknowns that will satisfy everyequation in the system.
The equations in the system can be linear or non-linear. This tutorial reviews systems of linear equations.
A problem can be expressed in narrative form or the problem can be expressed in algebraic form. Let's start with an example stated in narrative form. We'll convert it to an equivalent equation in algebraic form, and then we will solve it.

Example **1:**
A total of $12,000 is invested in two funds paying 9% and 11% simple interest. If the yearly interest is $1,180, how much of the $12,000 is invested at each rate?

Before you work this problem, you must know the definition of simple interest. Simple interest can be calculated by multiplying the amountinvested at the interest rate.

Solution:
We have two unknowns: the amount of money invested at 9% and the amount of money invested at 11%. Our objective is to find these two numbers.
Sentence (1) "A total of $12,000 is invested in two funds paying 9% and 11% simple interest." can be restated as (The amount of money invested at 9%) $+$ (The amount of money invested at 11%) $=$$12,000.
Sentence (2) "If the yearly interest is $1,180, how much of the $12,000 is invested at each rate?" can be restated as (The amount of money invested at 9%) $\times$ 9% + (The amount of money invested at 11% $\times$ 11%) $=$ total interest of $1,180.

It is going to get tiresome writing the two phrases (The amount of money invested at 9%) and (The amount of money invested at 11%) over and over again. So let's write them in shortcut form. Call the phrase (The amount of money invested at 9%) by the symbol x and call the phrase (The amount of money invested at 11% times 11%) by the symbol y.

Let's rewrite sentences (1) and (2) in shortcut form.

$$(1) \quad : \quad x + y = \$12,000.$$

$$(2) \quad : \quad 0.09x + 0.11y = \$1180$$

We have converted a narrative statement of the problem to an equivalent algebraic statement of the problem. Let's solve this system of equations.

A   system   of   linear equations can   be   solved   four   different   ways:

Substitution,

Elimination,

Matrices,

Graphing.

The Method of Substitution:

The method of substitution involves five steps:

Step 1: Solve for y in equation (1).

$$x + y \ = \ \$12,000.$$

$$y \ = \ \$12,000 - x$$

Step 2: Substitute this value for y in equation (2). This will change equation (2) to an equation with just one variable, x.

$$0.09x + 0.11y \ = \ \$1180$$

$$0.09x + 0.11\left(\$12,000 - x\right) \ = \ \$1180$$

Step 3: Solve for x in the translated equation (2).

$$0.09x + 0.11\left(\$12,000 - x\right) \ = \ \$1180$$

$$0.09x + \$1,320 - 0.11x \ = \ \$1,180$$

$$-0.02x \ = \ -\$140$$

$$x \ = \ \$7,000$$

Step 4: Substitute this value of x in the y equation you obtained in Step 1.

$$(1) \quad : \quad x + y = \$12,000.$$

$$\$7,000 + y \quad = \quad \$12,000$$

$$y \quad = \quad \$5,000$$

Step 5: Check your answers by substituting the values of x and y in each of the original equations. If, after the substitution, the left side of the equation equals the right side of the equation, you know that your answers are correct.

$$(1) \quad : \quad x + y = \$12,000.$$

$$\$7,000 + \$5,000 \quad = \quad \$12,000$$

$$(2) \quad : \quad 0.09x + 0.11y = \$1180$$

$$0.09\,(\$7,000) + 0.11\,(\$5,000) \quad = \quad \$1180$$

The Method of Elimination:

The process of elimination involves five steps:

In a two-variable problem rewrite the equations so that when the equations are added, one of the variables is eliminated, and then solve for the remaining variable.

Step 1: Change equation (1) by multiplying equation (1) by -0.09 to obtain a new and equivalent equation (1).

(1) : $x + y = \$12,000.$

$(New\ 1)$: $-0.09x - 0.09y = -\$,1080$

Step 2: Add new equation (1) to equation (2) to obtain equation (3).

$(New\ 1)$: $-0.09x - 0.09y = -\$,1080$

(2) : $0.09x + 0.11y = \$1180$

(3) : $0.02y = 100$

$y = \$5,000$

Step 3: Substitute $y = \$5,000$ in equation (1) and solve for x.

$x + y = \$12,000$

$x + \$5,000 = \$12,000$

$x = \$7,000$

Step 4: Check your answers in equation (2). Does

$$0.09x + 0.11y \quad = \quad \$1,180$$

$$0.09\,(\$7,000) + 0.11\$5,000 = \$1,180$$

The Method of Matrices:

This method is essentially a shortcut for the method of elimination.

Rewrite equations (1) and (2) without the variables and operators. The left column contains the coefficients of the x's, the middle column contains the coefficients of the y's, and the right column   contains   the   constants.

$$
\begin{array}{c}
(1) \\
(2)
\end{array}
\left[
\begin{array}{cc|c}
1 & 1 & 12000 \\
0.09 & 0.11 & 1180
\end{array}
\right]
$$

The objective is to reorganize the original matrix into one that looks like

$$
\begin{array}{c} (1) \\ (2) \end{array}
\left[\begin{array}{cc|c}
1 & 0 & a \\
0 & 1 & b
\end{array} \right]
$$

where a and b are the solutions to the system.

Step 1. Manipulate the matrix so that the number in cell 11 (row 1-col 1) is 1. In this case, we don't have to do anything. The number 1 is already in the cell.

Step 2: Manipulate the matrix so that the number in cell 21 is 0. To do this we rewrite the matrix by keeping row 1 and creating a new row 2 by adding -0.09 x row 1 to row 2.

$$-0.09\,[Row\ 1] + [row\ 2] \; = \; [New\ Row\ 2]$$

$$
\begin{array}{c} (1) \\ (2) \end{array}
\left[\begin{array}{cc|c}
1 & 1 & 12000 \\
0 & 0.02 & 100
\end{array} \right]
$$

Step 3: Manipulate the matrix so that the cell 22 is 1. Do this by multiplying row 2 by 50.

$$
\begin{array}{c} (1) \\ (2) \end{array}
\left[\begin{array}{cc|c}
1 & 1 & 12000 \\
0 & 1 & 5000
\end{array} \right]
$$

Step 4: Manipulate the matrix so that cell 12 is 0. Do this by adding

$$-1[Row\ 2] + [Row\ 1] \quad = \quad [New\ Row\ 1]$$

$$
(1) \begin{bmatrix} 1 & 0 & | & 7000 \\ 0 & 1 & | & 5000 \end{bmatrix}
$$
(2)

You can read the answers off the matrix as x = $7,000 and y = $5,000.

Determinants:

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

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

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

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

Therefore, we have

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

which one may check easily.

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.

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 columnnumber 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.

Eigenvalues and Eigenvectors:

The eigenvalue problem is a problem of considerable theoretical interest and wide-ranging application. For example, this problem is crucial in solving systems of differentialequations, analyzing population growth models, andcalculating powers of matrices (in order to define the exponential matrix). Other areas such as physics, sociology,biology, economics and statistics have focused considerableattention on "eigenvalues" and "eigenvectors"-theirapplications 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

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

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.

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

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

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. $det\left(A - \lambda I_n\right) \neq 0$. Since the zero-vector is a solution and C is not the zero vector, then we must have

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

Example. Consider the matrix

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

The equation $det\left(A - \lambda I_n\right) = 0$ 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

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

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

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

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 $(ad\text{-}bc)$ 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

$$A^2 - tr(A)A + det(A)I_2 = \mathcal{O}.$$

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

$$p(A) = \mathcal{O}$$

where $\quad p(\lambda) = det\left(A - \lambda I_n\right) \quad$ 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 $\dfrac{1}{\lambda}$ 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).

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

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

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

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

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 $[A + 4I_3 | \mathcal{O}]$, i.e.

$$\begin{pmatrix} 5 & 2 & 1 & 0 \\ 6 & 3 & 0 & 0 \\ -1 & -2 & 3 & 0 \end{pmatrix}.$$

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

$$\begin{pmatrix} -1 & -2 & 3 & 0 \\ 5 & 2 & 1 & 0 \\ 6 & 3 & 0 & 0 \end{pmatrix}.$$

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

$$\begin{pmatrix} -1 & -2 & 3 & | & 0 \\ 0 & -8 & 16 & | & 0 \\ 0 & -9 & 18 & | & 0 \end{pmatrix}.$$

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

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

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

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

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 formore 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

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

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

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

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

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

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 realnumber. 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 I₂.

The Case of Complex Eigenvalues:

First let us convince ourselves that there exist matrices withcomplex 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 normalcalculations for the eigenvectors. Let us see how it works onthe above example.

$$\lambda = 1 + 2i$$

We will do the calculations for . 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 complexentries in the eigenvectors.

We have seen that (1-2*i*) 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 (1-2*i*) 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 acomplex eigenvalue of A. In order to find the associated eigenvectors, we do the following steps:*

1. Write down the associated linear system

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

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

$A = a\, I_2.$