

Research title:

Matrix theory and applications

By Ashraf Faraj

Eleventh grade

2015-2016

supervisor: Miss Clodia

Introduction

In some branches of mathematics, certain symbols are used to facilitate the process of solving different problems. Moreover, these symbols may have a numerical value like determinants or may not have any value like the coordinates of a point in the space  A matrix is an example of these symbols and it doesn't have any numerical value but it's just an ordered schedule of elements. Now I'm going to give the precise definition of the matrix. A matrix of order is an ordered schedule of  elements that are distributed over n rows and m columns.  refers to the element of the matrix A that lies on the Matrix theory started reaching every certain field of science including physics, chemistry, engineering, computer science and even the different branches of mathematics itself including geometry, linear algebra, probability , and number theory. So, because of the large importance of matrices in the different fields of science, I wrote this article to present the basis of this theory, and to write about it's applications in mathematics. I also wanted to answer the following questions:

* What are the uses of matrices in our daily life?
* Can we really use matrices to solve linear equations and differential equations in a faster and better way?
* Do matrices really have any practical application in science or were they just made for theoretical studies?
* Did matrices really find solutions for certain problems that science was unable to solve or they were just a way of facilitating different scientific processes?
* Is it really possible that some kinds of matrices are used to study and predict natural events? and how?

Note: The reader of this article should be aware of the simple notations of matrices and the basic operations concerning them like addition, subtraction, multiplication by a number, multiplication of matrices, the transpose, and a little bit about the inverse of the matrix. Also the reader must be familiar with some of the terminology concerning vector spaces.

Note 2: for the people who are interested in learning the smallest basics about matrices you can read any of the following books: "The Matrix Cookbook" by Kaare Brandt Petersen

and Michael Syskind Pedersen , "Matrix Operations for Scientists and Engineers" by the professor Allan Jeffrey , and the book "Matrix Analysis for Engineers" by Alan J.Laub.

Table of contents

**Introduction**  2

**Table of contents**  3

**Chapter one: The basics of the theory** 4

1.1 Generalities concerning matrices 4

1.1.1 Kinds of matrices 4

1.1.2 Definitions and identities 7

1.2 Determinants 9

1.2.1 Properties and identities 9

1.2.2 Other kinds of determinants and their applications 12

1.3 Eigenvalues and eigenvectors 13

1.3.1 Eigenvalues 13

1.3.2 Eigenvectors 15

**Chapter two: Applications of matrices**  16

2.1 The least squares method 16

2.2 Linear equations 16

2.2.1 Systems of linear equations 17

2.2.2 The Gaussian eliminations method 19

2.2.3 Cramer's law 19

2.2.4 Using the rank in solving the systems of linear equations 20

2.2.5 The homogenous system 22

2.3 Linear differential equations 22

2.3.1 Differentiation and integration of matrices 23

2.3.2 Systems of Homogenous constant coefficient differential equations 24

2.3.3 The homogenous system 26

2.3.4 The non- homogenous system 28

**Conclusion**  31

**References**  32

**The index of the pictures**  32

**Chapter one: The basics of the theory: 1.1 Generalities concerning matrices 1.1.1 Kinds of matrices:**

In this section we are going to talk about some of the most frequent kinds of matrices and their properties.

* **The square matrix:** is the matrix that has an equal number of rows and columns.
* **The rectangular matrix:** this matrix has a different number of rows and columns.
* **The zero matrix:** is the matrix where all elements are equal to zero. The zero matrix has the property that its multiplication with any matrix is the zero matrix also.
* **The symmetric matrix:** itis the matrix that satisfies the relation .
* **The skew-symmetric matrix:** is the matrix that satisfies the relation .The elements of the leading diagonal of this matrix are all equal to zero.
* **The hermitian matrix:** is a square matrix where it's elements are members of  (the complex numbers field) and it serves the relation where \* denotes the conjugation operation complex numbers.
* **The column and row vectors:** they are matrices of order respectively. **[[1]](#footnote-2)**

**Example 1.1.1.2:** the following are zero, symmetric, skew-symmetric , hermitian, row, and column matrices respectively:



* **The idempotent matrix:** is the matrix that serves the relation and that leads to the relation: .
* **The nilpotent matrix of order k:** is the matrix that satisfies the relation .
* **The involuntary matrix:** is the matrix that satisfies the relation . It has the property that .**[[2]](#footnote-3)**

**Example 1.1.1.3:** the following matrices are idempotent, nilpotent (of order 3), and involuntary respectively:



* **The diagonal matrix:** is the matrix where all it's elements are zero except for those on the leading diagonal.This matrix has the property that its determinant is equal to the product of the elements that lie on the leading diagonal.
* **The triangular matrix:** it has two kinds: upper triangular, and lower triangular. The upper triangular matrix is the matrix where all the elements that lie under the leading diagonal are equal to zero. While the lower triangular matrix is the matrix where all the elements that lie above the leading diagonal are equal to zero. Both the upper and lower triangular matrices have the property that their determinant is equal to the product of the elements that lie on the principal diagonal.

**Example 1.1.1.4:** the following matrices are diagonal, upper triangular, and lower triangular respectively:



* **Orthogonal matrix:**  is the matrix that satisfies the relation  .This matrix has a group of properties which are:

1. it's eigenvalues lie on the unit circle.(we will discuss the meaning of this term later).
2. Its eigenvectors are unitary, i.e have length one. (this will be discussed also).
3. The inverse of the orthogonal matrix is orthogonal too.
4. the following relations also hold for the orthogonal matrix A:



* **Ortho-Sym matrix:** is a matrix which is orthogonal and symmetric in the same time and it's denoted by  and it satisfies the two following rules:



The Ortho-sym matrix has an extremely important rule when calculating its power:



* **Ortho-skew matrix:** A matrix which simultaneously is orthogonal and skew-symmetric is called an Ortho-skew matrix. It is denoted by  and it follows the two rules:



like the Ortho-sym matrix, this matrix has an important rule for calculating its power:



where i is the imaginary unit.

* **Unitary matrix:** is the matrix that satisfies the relation .
* **Positive definite matrices:** a matrix is said to be positive definite if and we write .
* **Positive semi- definite matrices:** a matrix is said to be positive semi-definite if and we write .
* **Negative definite matrices:**  a matrix A is said to be negative definite if –A is positive definite and we write .
* **Negative semi-definite: :**  a matrix A is said to be negative semi-definite if –A is positive semi-definite and we write .

These previous four definitions allow us to compare between matrices. We say that A>B if A-B is positive-definite.

* **The Vander monde matrix:** it is the matrix defined as follows:



The transpose of this matrix is also a Vander monde matrix. It also has the following very important property:



* **The Toeplitz matrix:** A Toeplitz matrix T is a matrix where the elements of each diagonal is the same. In the square case, it has the following structure:



Here are some kinds of the Toeplitz matrix (the following matrices are Toeplitz symmetric, circular Toeplitz, lower triangular Toeplitz, and upper triangular Toeplitz respectively):



The circular matrix has an important property which is:

 [[3]](#footnote-4)

**1.2.1 Basic operations and identities:**

In this section, we are going to talk about the basic operations concerning matrices including the inverse and some important identities and laws.

* **Addition and subtraction:** Two matrices are said to be conformable for addition (subtraction) if they both have the same shape. The result of adding (subtracting) two matrices is a matrix of the same shape of the first two in which each element in it is the sum(difference) of the corresponding elements in the first two matrices.
* **Multiplication:** two matrices are said to be conformable for multiplication if the number of columns of the first matrix is equal to the number of rows of the second matrix. The product is a matrix defined by the relation:



Here are some laws concerning operations over matrices:



* **Inverses:** the inverse matrix A-1of the square matrix A is the matrix that satisfies the relation: where is the unit matrix. In this case the inverse matrix can be constructed using the adjoint matrix of A by the relation: 
* **The generalized inverses**: The generalized inverse of the matrix A is the matrix A- that satisfies the relation: .As we can see the matrix A- is not unique.
* **The Pseudo inverse**: the pseudo inverse or also called ( The Moore-Penrose inverse) is the matrix A+ that satisfies the following four conditions:



The matrix A+ is unique but it doesn’t always exist. We also should mention that the conditions of symmetry are changed to the conditions of being a Hermitian matrix in the case of complex valued matrices. Now we are going to discuss the method of construction of this inverse matrix.



The so called "broad" and "tall" versions can also be called the "right" and the "left" inverses respectively. The matrix that doesn't have an inverse is called a singular matrix and it's also the matrix with determinant zero.[[4]](#footnote-5)

* **The Kroencker product:** the Kroencker product between two matrices A and B is defined as follows:



As we can see the Kroencker product of two matrices is very different than the normal product. It doesn't have any condition. For example, let's assume that we have two matrices A and B of orders  respectively then their Kroencker product is a matrix of order  defined in the same way as above.

**Theorem 1.1.1 (The Woodbury identity):** the Woodbury identity states that:



**Theorem 1.1.2( The PR theorem):** If the matrices P and R are positive definite then the following relation holds:



**Theorem 1.1.3 (The Kailath Variant):** this theorem holds for all matrices A,B and C :



**Theorem 1.1.4 (The Sherman-Morrison identity):** this theorem holds for a matrix A and both the vectors b ,and c:[[5]](#footnote-6)



**1.2 Determinants: 1.2.1 Properties and theorems:**

In this section we are going to talk about the basic properties of determinants and we are going to mention some of their theorems. NOTE: this chapter assumes that the reader is aware of the cofactors and the minors of the determinant and their properties.

* **Basic properties:**

1. Multiplication of the elements of any row or column of det A by a constant k changes the value of the determinant to k.det A. Equivalently, multiplication of det A by k can be replaced by multiplying the elements of any one row (column) of det A by k.
2. If every element in a row or column of det A is zero, then det A = 0.
3. If two rows (columns) of det A are the identical, or proportional, then det A = 0.[[6]](#footnote-7)
4. The determinant of the matrix A is equal to zero if there exists a row that is a linear combination of the others.
5. If A is a square matrix of order n, then . Where is the transpose matrix of the matrix A.
6. The value of the determinant does not change if any row is added or subtracted from the other it also remains unchanged if we add the multiple of a row to another row or if we add a linear combination of any group of rows to a certain one.
7. If two rows are interchanged in a determinant, then the sign of the determinant changes..
8. If a row "I" in a determinant can be written as a sum of two rows "k" and "f" then the determinant as whole can be written as the sum of two determinants which are identical in all the rows except for row "I" where one of them has the row "f" instead of it and the other has the row "k" instead of the row "I".
9. If A is a square matrix of order n then [[7]](#footnote-8)
10. .
11. .

Note: all of the previous properties are true for both the rows and columns of det A because of property 5.

* **Determinant theorems and methods of expansion:**

In what arrives, we refer to the group of rows of matrix A by Row(A) or shortly R. We also refer to the group of columns of A by Col(A) or shortly C. We also denote by I and J to the subgroups of the groups of the rows of A and columns of A respectively. And we denote by detA[I,J] the determinant of the submatrix A[I,J].

**Theorem 1.1.5 (Laplace expansion):**

For a square matrix A and i ∈ R, it holds that:



This formula that appears in identity 1 is called the Laplace expansion with respect to row

*i*. A more general form of this formula is an expansion with respect to a

subset I of the group R. The following identity is called the generalized Laplace

expansion with respect to row set I.

**Theorem 1.1.6 (Generalized Laplace expansion):**

For a square matrix A and , it holds that:



Here denotes the signature of the permutation:



where ,  with  and ,  with  **Example 1.2.1.1:** Let's consider the following matrix:



That's according to the generalized Laplace expansion in terms of the first two rows.

**Theorem 1.1.7 (Grassmann Plucker)**: Let A be a matrix with  (the number of its columns is bigger than the number of its rows).For with  and it holds that:

Where  is a short-hand notation for( in which the column j



is put at the position of column i in J; similarly for )[[8]](#footnote-9)

**Example 1.2.1.2:** Let's consider the following matrix A:



with by applying the previous theorem we have:



**Theorem 1.1.8 (Cauchy Binet theorem):** Suppose that A is a and B is a matrix and C=A.B, and  ,and .Then letting denote all subsets of {1,2,3,….k} of cardinality r then:



**Theorem 1.1.9 (Schur Complement):** If A is a square matrix and D is nonsingular(its determinant isn't equal to zero) then the following relation holds:



**Theorem 1.2.6 (The Hadamard inequality):** For any column vectors of length n with complex entries, it holds that:

[[9]](#footnote-10)

**1.2.2 Other kinds of determinants and their applications:** So far, we have only dealt with numerical determinants. But, there are other kinds of determinants called the functional determinants. These determinants have functions instead of numbers and for sure its value is a function also. In this section we are going to mention an example of a functional determinant and We are also going to talk about its properties and uses. The calculation of integrals over areas and volumes is often simplified by changing the variables involved to ones that are more natural for the geometry of the problem. When an integral is expressed in terms of the Cartesian coordinates x, y and z, a change of the coordinates to u1, u2 and u3 involves making a transformation of the form:



and when this is done a scale factor J enters the transformation to make up to the change of scales. The factor J is a functional determinant denoted by where:



and J is called the Jacobian of the transformation or, more simply, just the Jacobian. If the Jacobian vanishes at any point P, the transformation fails to establish a unique correspondence at that point between the point (xp, yP, zP) and the transformed point

(u1P, u2P, u3P). Here we illustrate an example of the usage of the Jacobian to transform from the Cartesian coordinates to the cylindrical coordinates.

**Example 1.2.2.1:** find the Jacobian for the cylindrical coordinates and for the spherical coordinates.

* solution for 1:

By calculations, we found that J = r. The Jacobian vanishes if r = 0, and the transformation fails because the angle  has no meaning when the magnitude is zero.

* solution for 2: By calculations, we found that . The Jacobian fails if r = 0 or . When r= 0 the previous angles have no meaning, and when  the angle  has no meaning.[[10]](#footnote-11)

**Picture 1**

**Picture 2**

**1.3 Eigenvalues and eigenvectors: 1.3.1 Eigenvalues:**

***Definition 1:*** let's consider the matrix A, which is a square matrix of order n. We call the following equation:



The characteristic equation of the matrix A. And we call the right hand side the characteristic polynomial of the matrix A. The characteristic equation can be also written in the following explicit form:



where I is the identity matrix of the same order of the matrix A. Let's denote to the characteristic polynomial of the matrix A by then we find by evaluating the determinant:



We can also notice that:

[[11]](#footnote-12)

***Definition 2:*** we call the real number ' f ' an eigenvalue of the matrix A, if that number is a root of the characteristic polynomial of the matrix. From this definition we can deduce that if the matrix A is of order n, then its characteristic polynomial is of order n too, so the matrix has n eigenvalues (some of them may be repeated).

***Definition 3:*** An eigenvalue  has algebraic multiplicity q if q is the largest number such that  is a root of the characteristic polynomial. The geometric multiplicity r is the dimension of the null space . An eigenvalue is said to be simple if r=q=1.

Depending on the rules of the polynomials, we find that (assuming that  are the eigenvalues of the matrix):



We also find that the sum of the eigenvalues is equal to the trace of the matrix ( the sum of the elements that lie on the leading diagonal).

Hint: the determinant of a matrix is different from zero if all the eigenvalues of the matrix are different from zeros. And as a consequence, the matrix A is nonsingular if the fixed term (a0) is non-zero .[[12]](#footnote-13)

**Example 1.3.1.1:** Find the characteristic polynomial and the eigenvalues of the matrix A:



* Solution: We find the characteristic polynomial of the matrix A and then we find its roots. The characteristic polynomial is:



**Theorem 1.3.1 (Calley-Hamilton):** any square matrix of order n is a root of its characteristic polynomial which means that if:  ­­­then .

**Theorem 1.3.2:** If  is the characteristic polynomial of the matrix A and if A is nonsingular then its inverse is given by the relation:



**Theorem 1.3.3:** If A is a block diagonal matrix, with its block matrices If is the characteristic polynomial of the block matrix then the characteristic polynomial of the matrix A is given by the relation:[[13]](#footnote-14)



**Theorem 1.3.4:** if the matrices A and B are similar to each other then they have the same eigenvalues with the same algebraic and geometric multiplicity. The algebraic and geometric multiplicities q and r respectively satisfy the following inequality:[[14]](#footnote-15)

**1.3.2 Eigenvectors:**

I'm not going to talk about all of the properties of the eigenvectors, I'm just going to mention all the information needed to understand the rest of the article.

***Definition 4:*** Given a matrix ,a vector is a right eigenvector for the eigenvalue if it satisfies the following equation:



***Definition 4:*** Given the same matrix A, a vector y that satisfies:

is called a left eigenvector.



Note that, even though the corresponding eigenvalues are the same, the right and left eigenvectors of a matrix are usually different.[[15]](#footnote-16)

**Example 1.3.2.1:** Let's consider the following matrix A:



find the eigenvectors of the matrix A.

* Solution: First of all we find the eigenvalues of the matrix A and they are:



The first eigenvector corresponding to the first eigenvalue has coordinates that satisfy the system of equations:



by using the method of solving the homogenous system of equations we find that the coordinates are found in the terms of an arbitrary variable let's say x1(1). by setting  we find that :



And we know that if we multiply the eigenvector with a nonzero real number then the eigenvectors stays an eigenvector so we can set that the number k1 is equal to 1 so:

**

With the same algebraic manipulation we find that the two left eigenvectors. And then:

 [[16]](#footnote-17)

**Chapter two: 2.1 linear equations: 2.1.1 Systems of linear equations:**

In this section, we are going to talk about the solution of the systems of linear equations using matrices. And for that, we present the following system of n equations and m variables (these variables are usually denoted by x1,x2,……,xm):



where both aij and bi are real constants (bi is also called the fixed term)

from now on, we are going to connect the previous system with two matrices which are:



The first matrix is called the coefficient matrix of the system. The second matrix is called the augmented matrix of the system. It's clear that we get the augmented matrix from adding the column to the coefficient matrix.

***Definition 1:*** We say that the numbers form a solution of the given system, if all the equations where satisfied after putting c1 in the place of each x1 and c2 in the place of every x2 and cm in the place of every xm.

***Definition 2:*** We say that the previous system is solvable if it has at least one solution. And we say that it is inconsistent if it has no solution at all. In what comes later, we will see that the system of linear equations is either solvable or inconsistent. And in the first case, the system either has a unique solution, or an endless number of solutions.

***Definition 3:*** We say that two systems of linear equations (with the same number of variables) are equivalent, if they both have the same solutions.

If we denote by B the column matrix formed from the fixed terms of the previous system, and we refer with X to the column matrix formed from the variables of the system i.e:



Then the system can be written in matrix form AX=B where A is the coefficient matrix.

**2.1.2 The Gaussian elimination method:**

Gauss's method is considered to be one of the best and easiest methods in solving the systems of linear equations. This method transforms the system of linear equations to an equivalent system that is easier to solve. And with solving the new system the old system is solved. The essence of the Gaussian elimination method is changing the augmented matrix of the system, depending on the elementary row operations, to an echelon form matrix. Then we replace the original system with this new system and then solve it. We call the resulting system an echelon linear system. The formal definition of the echelon form will come eventually in this section.

***Definition 4:*** we call each of the following operations elementary system operations:

* Multiplying any of the equations by a constant.
* Adding any of the equations after multiplying it with a number to another equation.
* interchanging between two equations of the system.

***Definition 5:*** A matrix **A** is said to be in echelon form and denoted by **A**E if:

1. All rows of A containing nonzero elements lie above the rows that contain only zeros.
2. The first nonzero entry in a row of A, called the leading entry in the row, lies in a column to the right of the leading entry in the row above. The echelon matrix has the form:



**Theorem 2.1:** if we get a system of linear equations from another system as a cause of applying a finite number of elementary row operations then the two systems are equivalent (have the same solutions).

Earlier we have connected between the system of linear equations and two matrices. The coefficient matrix and the augmented matrix. And every operation done on the system (from the previous elementary operations) corresponds to a certain operation done on the augmented matrix of the system. For example, multiplying a certain equation in the system by a constant corresponds with multiplying its coefficient row in the augmented matrix by the same constant. The second operation is clear. The third operation is equivalent to interchanging two rows in the augmented matrix. And now we are ready to illustrate the Gaussian elimination method.

Let’s consider the following system of linear equations and its augmented matrix respectively:



then we change the previous augmented matrix, depending on the elementary row operations, into an echelon matrix. After that we write the system corresponding to the echelon matrix that we produced. But before solving this new system we delete the rows that are all zeros. And now we have a new system of equations that is simple and easy to solve, and that's because the zero rows represent equations that are fulfilled for all the values of x1,x2,……,xm . Of course, deleting these rows from the system is equivalent to deleting the zero rows from the augmented matrix of the system.

Now we reach to few results concerning the solution of the system of linear equations:

* If the augmented matrix had a row of the shape where a is nonzero then the system is inconsistent (unsolvable). So that means that we can know whether the matrix is solvable or not just by looking at the modified augmented matrix.
* If the matrix doesn’t have any row from the previous shape then the system is solvable. In this situation we find two cases:

1. the number of rows of the Augmented matrix is equal to the number of the variables. Then the system will have the following shape which we call the triangular shape:



where all the c's on the diagonal are nonzero. To find the solution of this case, we compute the value of the last variable from the last equation and we put its value in the previous equation and we find the value of the earlier variable and we put its value in the earlier equation and so on…. So, the previous system has only one solution.

1. The number of rows in the augmented matrix is less than the number of variables in the system. Then the echelon augmented matrix has the shape that we call The Trapezoidal shape:



In this case, we call the variables "main variables". And we call the rest of the variables "free variables". To find the solution of this system, we first move the free variables to the right side of the equations of the system. then we give these variables certain numerical values then the main variables take corresponding values too. So as a result, we say that this system has an endless number of solutions because with every choice of the free variables we find a different value for the main variables and because the choices are endless the solutions are endless too. [[17]](#footnote-18)

**2.1.3 Cramer's law:**

Cramer's rule is a method of solving properly determined systems of linear algebraic equations only (systems where the number of equations is equal to the number of the variables) , and it states that (assuming that the following matrix represents the system being solved):



where and  is the determinant of the coefficient matrix with the vector b instead of its first column, and  is the determinant of the coefficient matrix with the vector b instead of its nth column and so on.[[18]](#footnote-19)

**2.1.4 Using the rank in solving the system of linear equations:**

This section requires knowledge with the concept of the rank of the matrix (So because I wasn't able to discuss this wide concept you can go back to any of the references: Now I present the method of determining the solution of the system of linear equations using the rank:

Let the coefficient matrix A of the first system be an matrix and let b be an m element column matrix then:

1. A solution set exists if row rank(A) = row rank(A|b). The solution will be unique if row rank(A) =row rank(A|b) = n, but if row rank(A) = row rank(A|b) = r < n, then r of the unknowns  can be expressed in terms of the remaining n-r unknowns when specified as arbitrary parameters.
2. No solution set exists if row rank(A) < row rank(A|b).[[19]](#footnote-20)

**2.1.5 The homogenous system:**

So far, we have dealt with the non-homogenous systems of equations. Now we present the homogenous system. We say that the system of linear equations is homogenous if the column of the fixed terms is all zero. The general form of the system is:



Considering that A is the coefficient matrix and X is the column matrix formed from the variables this system can be written in matrix form:



We can easily see that the augmented matrix is formed by adding a row of zeros to the coefficient matrix so . As a result, we find that row rank(A)=row rank(). So according to the theorems in the previous section, we can say that the system is always solvable. In fact, that is true because we know that any homogenous system has the solution:



This solution is called the null solution or the lame solution. And we call any solution of the system in which at least one of the variables is nonzero, "the nonzero solution".

**Theorem 2.2:** let's consider the first system of linear homogenous equations. This system has infinitely many solutions if and only if row rank(A) is smaller than the number of variables. And it has one solution if and only if row rank(A) is equal to the number of variables.

As a direct consequence of this theorem, we get the following conclusion. If we have a homogenous system of n linear equations and n variables, then the condition for this system to have a solution different from the zero solution is that its determinant is equal to zero.

**Theorem 2.3:** If are all solutions of the homogenous system, then every linear combination of them i.e:  is a solution of the system too.

***Definition 6:*** Let's denote by V to the group of solutions of a certain homogenous system. We call this group a basic group, if all of these solutions are linearly independent and if every solution in the system can be written as a linear combination of them.

**Theorem 2.4:** Let's consider a homogenous system of linear equations with m variables. If row rank(A) is less than the number of variables then this system has a basic group of (m-r(A)) solutions.

***Definition 7:*** If we have a homogenous system of linear equations and if the group of solutions is a basic group then we call the solution:



(where all the c's are real numbers) the general solution of the system.

**Example 2.1.5.1:** Use Gaussian elimination, in the form of elementary row operations applied to the augmented matrix, to solve this system of linear equations:



* Solution: In this case we have that n=4 and m=3 so the system is undetermined. Now we apply certain row operations on the augmented matrix will produce the following matrix:



And as we can see the system can't be reduced more than that. Now we notice that the number of unknowns is bigger than the number of the equations, so we can set for example x4 as k then from equation 3 we have that: . By substituting x3 in the previous equation and so on we get that:

.[[20]](#footnote-21)

**2.2 Linear differential equations:**

In this section we are going to talk about the solution of the linear system of differential equations using matrices and the concept of the eigenvectors.

**2.2.1 Differentiation and integration of matrices:**

Solving the system of linear differential equations requires the knowledge with the differentiation and integration of matrices whose elements are functions of a certain variable say t. Let the  column vector have differentiable elements and let the matrix  have differentiable elements  . Then the derivatives of both x(t) and G(t) according to time are:



An important special case occurs when A is a constant matrix then its derivative is equal to the zero matrix.

Here are some laws concerning the differentiation of matrices:



By definition, if A(t) = [aij(t)] is an matrix, with i =1, 2, . . . , m and j = 1,2, . . . , n, then the indefinite integral of the element in the ith row and jth column of A(t) is  ,so the indefinite integral of the whole matrix is defined as :



Where of course, an arbitrary constant matrix should be added after the integration has been performed on each element of the matrix. Similarly, the definite integral of a matrix A which is between the limits , is by definition a matrix in which each element in it is the definite integration of the same corresponding element as follows:



**2.2.2 Systems of Homogenous constant coefficient differential equations:**

The most general first-order system of this kind involving the n unknown functions of the independent variable t has the following shape:



where aij and bij are real constants, and the hi(t) are arbitrary functions of the variable t. It will be assumed that the equations are all linearly independent, so no equation is can be written as a linear combination of the others. Let’s define the following important matrices:



then the previous system can be written in matrix form:



So by the hypothesis that the equations are linearly independent, the matrix B has an inverse which is the matrix B-1 , so by pre-multiplying by this matrix we see that:



Let's define the matrix A as A=B-1C and the column vector from order  as f(t)=B-1h(t) then the system will take the following shape:



the previous system is called non-homogenous if and it's called homogenous otherwise.

**2.2.3 The homogenous system:**

In this section we are going to consider the homogenous system:



and establish a connection between the general solution of the system and the eigenvectors of A. Our concern will be finding the general solution of the system ad then finding the solution for the initial value problem which is finding the solution for the system subject to a set of n initial conditions of the form with where the constants ki are the values the functions xi(t) are required to satisfy initially when t = t1.

And now we will attempt to find a solution for the system of the form:



where is a constant column vector. By substituting that in the first system we find:



and by dividing the equation by the non-vanishing term we find:



This shows that the permissible values of  are the eigenvalues of the matrix A, while the associated constant vectors  are the eigenvectors of the matrix A. When A has a full set of n linearly independent eigenvectors, the linearly independent solutions of the system are

. An matrix , with its columns the solution vectors of the system, is called the fundamental matrix for the system. The general solution of the system will an arbitrary linear combination of the n independent eigenvectors of the matrix A of the form:



A fundamental matrix is not unique, because the eigenvectors forming its columns can be arranged in different orders, and each eigenvector can be multiplied by a constant factor and still remain an eigenvector. This non-uniqueness of the fundamental matrix can cause the arbitrary constants in general solutions to appear differently, depending on how the fundamental matrix has been constructed. However, these different forms of the general solution of the system are unimportant, because the solution of a corresponding initial-value problem is unique, so when the arbitrary constants are chosen to make the xi(t) satisfy the n initial conditions, all forms of general solution in which arbitrary constants may appear differently will give rise to the same unique solution of the initial value problem.

**Example 2.2.3.1:** find the solution of the system of equations:



* Solution: first of all, the system can be written in matrix form as  where A is the matrix:



the eigenvalues and eigenvectors of the matrix A are:



Then we may take the fundamental matrix as:



and the general solution has the form:



Now we are going to illustrate the method of solving an nth order differential homogenous system using the previous information. The system is:



This equation can be replaced by n equivalent first order linear differential equations which by solving them we can establish the solution of the first equation. The matrix of the n linear differential equations is:



where



We mention here that an nth-order system can be reduced to a set of n first-order equations in more than one way, though the method of reduction used here is usually the simplest.

**2.2.4 The non-homogenous system:**

In this section we are going to deal with the system of non-homogenous linear differential equations and their solution.

A non-homogenous system has the form:



Recalling that , and setting that where P diagonalizes A then the system will take the following form:



by pre-multiplying with the constant matrix P-1 we find that:



setting that we find that:



and writing that , with ,where the functions are known in the terms of the non-homogenous vector f(t), the result becomes:



The solution of earlier system now simplifies to the solution of the n separate non-homogeneous equations of the form whereas before, the elements of the diagonal matrix D are the eigenvalues of the matrix A correspondent to the eigenvectors occurring in the diagonalization matrix P. Once the vector z(t) has been found, the vector x(t) cam be found by the relation: .

When no initial conditions are specified, each element of x(t) will be the sum of the general solution of the corresponding equation in the homogeneous system, to which is added a particular integral produced by the non-homogeneous term f(t). To solve an initial-value problem it is first necessary to find the general solution for x(t), and then to match the arbitrary constants involved to the initial conditions.

**Example 2.2.4.1:** Use diagonalization to find the solution of the non-homogeneous system:



* Solution: the eigenvalues and eigenvectors of the matrix A are:



so the diagonalizing matrix is:



Using these results we find that:



then separating the equations we find that:



For convenience in what follows, the method of solution of a general linear first-order differential equation by means of an integrating factor. Solving these linear equations we find that:



Therefore the solutions are:



After using the initial conditions, and after few calculations we find that:[[21]](#footnote-22)



**2.3 Matrices and The Least Squares Curve Fitting:**

A record of experimental or statistical is usually in the form of n discrete pairs of measurements  that show how a quantity y depends on the argument x, where both the xi and the yi are subject of experimental error. We will call these pairs "data points". when the set of data points is to be approximated by a smooth curve, this is most easily accomplished by approximating the discrete observations by a continuous curve . When representing experimental data points by a curve, it is usual to choose a curve in the form of a polynomial of low degree, and to fit it by using the method of least squares. If the plot of data points can reasonably be represented by a straight line, the equation is the equation we need. But when the data points appear to be in a parabolic shape the equation is what we need. Polynomials of still higher degree can also be fitted, though a cubic is usually the highest degree equation that is used. This is because when a higher-degree polynomial is fitted, the coefficients of the polynomial become very sensitive to the errors in the data points which can lead to a weaker approximation.

Because the measurements contain errors of observation, a curve cannot be expected to pass through each data point, so some compromise becomes necessary. The idea underlying the least-squares approximation involves choosing the coefficients in the equation to be fitted, like a0, a1 and a2 in a quadratic (parabolic) approximation, in such a way that the sum of the squares S of the differences between the points Yi on the curve Yi = a0+ ‏a1xi ‏+a2xi2 at the points xi, and the actual measurements yi at the points xi is minimized. So the expression S that is to be minimized is given by:



The quantity S is simply the sum of the squares of the vertical distances between Yi and the actual measurement yi at each of the n values xi is minimized. Here S is defined as the sum of the squares of these distances, because the quantities  take account of the magnitude of the differences between the Yi and the yi, without regard to the signs of the differences.

If the equation Yi = a0+ ‏a1xi ‏+a2xi2 is to be fitted, the sum S of the squares will be minimized when a0, a1 and a2 are chosen such that After differentiating with respect to a0, a1 and a2 we find three equations that are:



Instead of finding a0, a1 and a2 from these equations, we now show how a matrix argument can generalize these results. This approach has the advantage that the same form of matrix computation will enable a polynomial of any degree to be fitted to a set of data points.

Right now we are going to deal with a specific case which fitting a second degree polynomial to n pairs of data. This polynomial has the form:



To solve this problem let's consider this over-determined system of equations:



This system can be written in matrix form as:



Clearly this last matrix equation can't be solved as it's right now. It must be pre-multiplied by a 3 times n matrix M because then both MX and MY will be 3 times 3 matrices, and with a suitable choice of M the matrix MX will have an inverse leading to the solution of the system by:  but then the choice of the matrix M will control the resulting parameters. So, we are going to choose a matrix that coincides with we found earlier avoiding to introduce an arbitrary matrix. Let's try setting M=XT then the matrix product XT.X becomes:



So, in terms of matrices, the coefficients a0, a1 and a2 are the elements of the column vector a where:

****

If instead of a parabola a line is going to be fitted to the data points by the least squares method , then the matrix X is simplified to n times 2 matrix as this shape shows:



In statistics the fitting of a straight line to a data set by least squares is called "regression", and the straight line itself is called the "regression line", and the coefficient a1 that measures the slope of the regression line is called the "regression coefficient".

**Example 2.3.1.1:** Use the method of least squares to fit the quadratic to the set of data points: .

solution:



so:



and here is a plot of the previous function along with the experimental data points:[[22]](#footnote-23)



**Picture 3**

Conclusion

As we saw earlier, matrix theory has wide applications in different domains of our daily life. Now, let's answer the questions that I proposed at the beginning of the article. Matrices play a fundamental role in a lot of branches mathematics for example in statics ( the least squares curve fitting method) and in linear Algebra ( solving systems of linear equations) and Analysis ( solving a differential equation of any degree) which is considered to be an extremely important application for matrices because differential equations are used in all the sciences like physics and chemistry and of course mathematics. We asked a question whether matrices were found for only theoretical studies or do they have any practical applications. For sure, all the applications of matrices are used in theoretical studies, but all of these studies have a link to the real world which means that matrices have a great impact on the real world. Matrix theory wasn't only found for reduction of mathematical and scientific processes, for example fuzzy matrices were found to study the human behavior which is a new theory that started solving new problems and matrices were widely used in quantum physics through the density matrix theory. Now we come to the last question. Can matrices be used to study natural events and how? Actually matrices can be used to predict natural events through a concept called transition matrices. These matrices are built in a way such that the element aij indicates the probability of moving from state I to state J of a certain situation. Scientists use these kinds of matrices to model weather states in a certain region. And by multiplying any transition matrix (for example the weather transition matrix) by itself a lot of times the elements in these matrices converge to certain numbers that will denote in our case the climate in the studied region. So these matrices can be used by meteorologists to predict natural events for a long period of time ahead in a certain region. As a conclusion, matrix theory really went deeply in every branch of every science in the world not just simplifying processes but finding solutions for problems that were unsolvable for long times. I recommend every person to learn the techniques of matrix theory because maybe one day, this theory will be the basic theory used all over the world.

* **English Electronic References:**

1. "**Matrix Operations for Engineers and Scientists** An Essential Guide in Linear Algebra" by **Prof. Dr ALAN JEFFRY**, University of Newcastle in the United Kingdom in the year 2010. Published by the library of Congress , first edition, pages: 17, 33 , 34, 101, 102 ,103, 63 , 64 and the pages from 159 to 174,
2. **"An Introduction to Random Matrices"** by **GREG W.ANDERSON, ALICE GUIONNET and OFER ZEITOUNI**, published by Cambridge university press in the U.S.A in 2010, the first edition, pages: 414 and 415
3. **"The Matrix Cookbook"** by **KAARE PETERSON, MICHAEL SYSKIND PEDERSON** the first edition published in November 15 2012, pages: 17 ,21 , 22 ,18.
4. **"Determinants: Manipulation and Evaluation"** by **YUFEI ZHAO**, published in 22 September 2009, pages: 1 and 2.
5. **"Matrices and Matroids for Systems Analysis"** by **KAUZO MURATO** published by the library of Congress in 2000, first edition, pages: 33 and 34.
6. **" Matrix Analysis and Algorithms"** by Andrew **Stuart and Jochen Voss**, published in 4th August 2009, pages: 7,8 and 9.
7. **"Introduction to Vectors and Tensors"** by **Ray.M.Bowen and C.C.Wang** , first volume pages: 4 and 6.
8. **"Handbook of Mathematics"** by  **I.N. Bronshtein , K.A. Semendyayev , G.Musiol , H.Muehlig**  published by the library of the congress, the fifth edition, page 260.
9. **"An Introduction to Matrices, Groups, and Sets for Science Students"** , by **G.Stephenson** , Dover publications, Inc. Newyork, page: 64.

* **Arabic Electronic References:**

1. **"المصفوفات النظرية و التطبيق**" by **professor Majdi al Taoeel** , published by the university of Cairo in 1996, pages: 11,12,13.

* **Paper References:**

1. The Arabic book **"المصفوفات "** by **professor Hasan Al Nakar** ,published by the university of Aleppo in Syria in 1980-1981,the second volume, pages:128 , 129 ,130 , 131 and pages from 140 to 157, and pages from 172 to 177.

* **Index of the pictures:**

|  |  |  |
| --- | --- | --- |
| Number of the picture | The page | Description |
| 1 | 12 | Cylindrical coordinates |
| 2 | 12 | Spherical coordinates |
| 3 | 30 | Example 2.3.1.1 |

1. "Introduction to Vectors and Tensors Volume 1" by Ray.M.Bowen and C.C.Wang , pages:4,6. [↑](#footnote-ref-2)
2. The book "المصفوفات النظرية و التطبيق" by the professor Majdi al Taoeel , pages:11,12,13. [↑](#footnote-ref-3)
3. ”Determinant: Evaluation and manipulation" by Yufei Zhao, pages: 1,2. [↑](#footnote-ref-4)
4. "The Matrix Cookbook" by Kaare Brandt Petersen and Michael Syskind Pedersen pages: 17,21,22. [↑](#footnote-ref-5)
5. "The Matrix Cookbook" Kaare Brandt Petersen and Michael Syskind Pedersen page 18. [↑](#footnote-ref-6)
6. "Matrix Operations for Engineers and Scientists", By professor Alan Jeffrey, page:17 [↑](#footnote-ref-7)
7. "Handbook of Mathematics", By I.N. Bronshtein · K.A. Semendyayev · G.Musiol · H.Muehlig, page 260. [↑](#footnote-ref-8)
8. "Matrices and Matroids for Systems Analysis" By Kauzo Murato, pages:33,34. [↑](#footnote-ref-9)
9. " An introduction to Random Matrices" By Greg W.Anderson, Alice Guionnet and Ofer Zeitouni, pages:414, 415. [↑](#footnote-ref-10)
10. "Matrix Operations for Scientists and Engineers" By prof Alan Jeffrey , pages:33, 34. [↑](#footnote-ref-11)
11. The book " " المصفوفات written by: professor Hasan Al Nakar, pages:128 and 129. [↑](#footnote-ref-12)
12. The book " Matrix Analysis and Algorithms" , by Andrew Stuart and Jochen Voss, page 8. [↑](#footnote-ref-13)
13. The book " " المصفوفات written by: professor Hasan Al Nakar, pages:130 and 131. [↑](#footnote-ref-14)
14. The book " Matrix Analysis and Algorithms" , by Andrew Stuart and Jochen Voss, page 9. [↑](#footnote-ref-15)
15. The book " Matrix Analysis and Algorithms" , By Andrew Stuart and Jochen Voss, pages 7 and 8. [↑](#footnote-ref-16)
16. The book " Matrix Operations for Engineers and Scientists", By professor Alan Jeffrey, pages:101 , 102 , 103. [↑](#footnote-ref-17)
17. "المصفوفات ", by professor Hassan Al Nakar, pages: from 140 to 153. [↑](#footnote-ref-18)
18. "An Introduction to Matrices, Groups, and Sets for Science Students", by G.Stephenson , page 64. [↑](#footnote-ref-19)
19. "Matrix Operations for Scientists and Engineers", by professor Allan Jeffrey , page: 84. [↑](#footnote-ref-20)
20. المصفوفات " ", by professor Hassan Al Nakar, pages: from 172 to 177. [↑](#footnote-ref-21)
21. this section was taken from "Matrix Operations for Scientists and Engineers", by Allan Jeffrey , pages: from 159 to 174 [↑](#footnote-ref-22)
22. "Matrix Operations for Engineers and Scientists" By professor Alan Jeffrey, pages:63, 64. [↑](#footnote-ref-23)