الجمهورية العربية المتحدة

التخطيط القومي

معهد التخطيط القومي

Mem. No. 600

Lecture Notes on

Matrices.

by

Dr. Youssef Boutros

Operations Research Center

November 1965.

القاهرة

٣ شارع محمد مظهر - بالزمالك

CHAPTER I

DEFINITIONS AND FUNDAMENTAL THEOREMS.

I.1 Definitions.

An aggregate of numbers arranged in the form of a rectanglar table, is called a rectangular matrix. It will have m rows and n columns, and may be set in the form:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & & a_{mn} \end{bmatrix}$$

The first subscript, then, designating the row, the second designating the column, in which the element is located.

This may be abbreviated to the form:

$$A = [a_{ij}] \qquad (i = 1,2, \ldots, m; \quad j = 1,2,\ldots,n).$$

Two matrices are equal if their corresponding elements are equal.

Matrices composed of a single row are called simply rows (or row vectors). Matrices composed of a single column are called columns (or column vectors).

If the number of rows of a matrix is equal to the number, n, of columns, it is called square, and of the nth order..

Among square matrices, an important role is played by diagonal matrices, i.e. matrices of which only the elements along the principal (leading) diagonal are different from zero:

$$\begin{bmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & & \alpha_n \end{bmatrix}$$

If all the numbers α_i of such a matrix are equal to α, the matrix is said to be <u>scalar</u>, and if $\alpha = 1$, the matrix is said to be the <u>unit</u> matrix.

$$\begin{bmatrix} 0 & 1 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & & 1 \end{bmatrix} = I$$

Lastly, a matrix all of whose elements are equal to zero is called a <u>null</u> matrix, or zero matrix. We shall designate it by the symbol 0.

The determinant whose element are the elements of a square matrix (without disarrangement) is said to be the <u>determinant of that matrix</u>, and we write the determinant of A as $|A|$, or often as det (A).

A is said to be non-singular (or regular) if its determinant is not equal to zero, in the contrary case it is of course singular.

I.2. <u>The addition of matrices.</u>

A matrix C whose elements are the sums of the corresponding elements of A and B, matrices having like numbers of rows and columns, is called the <u>sum</u> of A and B:

$$C_{ik} = a_{ik} + b_{ik}, \quad (i=1,2,\ldots,m;\ k=1,2,\ldots,n).$$

I.3 Multiplication of a matrix by a number.

A matrix whose elements are obtained by multiplying all the elements of the matrix A by a number α is called the product of the number α and the matrix A:

$$\alpha A = [\alpha a_{ik}], \quad (i = 1,2,\ldots,m, \quad k=1,2,\ldots,n).$$

I.4 The multiplication of matrices.

Multiplication of the matrices A and B is defined only on the assumption that the number of columns of matrix A equals the number of rows of matrix B. On this assumption, the elements of the product , C = AB, are defined in the following manner: the element in the ith row and the jth column of the matrix C is equal to the sum of the products of the elements of the ith row of the matrix A by the correspondy elements of the jth column of matrix B. Thus:

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

$$(i=1,2, \ldots, m; \; j = 1,2, \ldots, p);$$

where

$$\underset{mxn}{A.} \quad \underset{nxp}{B} = \underset{mxp}{C.}$$

It is to be noted that the product of two rectangular matrices is again a rectangular matrix, the number of row of which is equal to the number of rows of the first matrix, and the number of columns of which is equal to the number of columns of the second matrix. So, for instance the product of a square matrix and a matrix composed of one column is a matrix of one column.

The commutative law for multiplication does not, generally speaking, hold. The matrices AB and BA make sense simultaneously only if both matrices A and B are square matrices and of the same order. Thus,

$$AB \neq BA$$

in general. In particular, the multiplication of a square matrix by the unit matrix, of the same order, is commutative, i.e.

$$AI = IA = A$$

Hence follows the special role of the unit matrix in the multiplication of matrices: amongst all square matrices of the same order, the unit matrix plays the same role as the number one does among numbers. If we interchange rows and columns in the matrix

$$A = [a_{ij}]$$

We obtain the <u>transposed matrix</u> or transpose $A^T = [b_{ij}]$

where

$$b_{ij} = a_{ji} \qquad (i=1,2, \ldots n \; ; \; j = 1,2, \ldots , m).$$

The following rule (the <u>reversal rule</u>) for a transposed product should be noted:

$$(AB)^T = B^T A^T .$$

In conclusion we shall remark that the determinant of a product of square matrices is equal to the product of the determinants of the multiplied matrices:

$$|AB| = |A| \cdot |B|$$

I.5. The quadratic form.

A homogeneous polynomial of the second degree in several variables $x_1, \dots, x_n$ is called a quadratic form. We shall consider only those with real coefficients. Any quadratic form may be written as

$$\Phi(x_1, x_2, \dots, x_n) = \sum_{i,k=1}^{n} a_{ik} x_i x_k,$$

where $a_{ik} = a_{ki}$.

It is obvious that the coefficients a_{ik} will form a square matrix of the order n, which fulfill the condition $a_{ik} = a_{ki}$. Such matrices are called symmetrical.

A quadratic form is said to be positive-definite if its values are positive for any real values of $x_1, \dots, x_n$, not all zero simultaneously.

It is evident that the diagonal coefficients of a positive-definite form are positive, for

$$a_{11} = \Phi(1,0, \dots, 0), \quad a_{22} = \Phi(0,1, \dots, 0), \dots$$

$$a_{nn} = \Phi(0,0, \dots, 1)$$

CHAPTER II

SYSTEMS OF LINEAR EQUATIONS.

This chapter is devoted to three problems that are closely related to each other: the problem of solving a non-homogeneous linear algebraic system, the problem of inverting a matrix, and the problem spoken of as elimination.

In theory all of these problems are soluble simply enough. However, if a matrix of high order is connected with the problem, the actual solution requires a great number of computational operations and storage space in computers.

Numerical methods solving the stated problems are divisible into two groups: exact and iterative methods. By exact methods we understand methods that give the solution of a problem by means of a finite number of elementary arithmetic operations. The number of computational operations necessary for the solution of the problem depends only upon the form of the computational scheme (Algorithm) and upon the order of the matrix defining the given problem. Inexactitude in the solution found occurs as the result of the inevitable rounding-off of the figures in the course of the computation. Along with this, one may run up against the phenomenon of the disappearance of significant figures in the course of the computation, as the result of the subtraction of two numbers differing little from each other. This loss of significant figures may occasion such an important reduction in the accuracy of the result that it is often necessary to alter the computational scheme because of it, or re-do the work with a greater number of significant figures in the intermediate calculations.

fundamental method of this group is the method based on the
of elimination . This algorithm of this method with
n the name of Gauss is associated, consists - when applied
he solution of a non-homogeneous linear system - of a chain of
essive eliminations by means of which the given system is
sformed into a system with a triangular matrix whose solution
ents no difficulty.

Many different algosithms have been developed and are
ently employed for all three of the problems mentioned.

Iterative methods afford a means by which a system of
ar equations may be solved approximately. The solution of
stem by iterative methods is obtained as the limit of succe-
e approximations computed by some uniform process. The
ergence of these approximations depends essentially on the
ents of the matrix defining the given problem. The rate
he convergence depends also on a happy choice of the initial
oximation on which the iterative process is founded.

The immense advnatage of the iterative schemes consists
he simplicity and uniformity of the operations to be effected,
therefore in the possibility of completely mechanizing the
ess of computation.

. Linear functions - Exchange.

simiplicity we consider a linear function y of two independent
ales x , x' :

$$y = a x + b x' + k,$$

e a, b are fixed given coefficients and the constant
a given number. In a spaces system of coordinates with
xes x, x', y, the linear function will be represented by
ne. If the constant k vanishes, the function

$$y = ax + b x'$$

is said to be a homogeneous linear function or shortly a linear form . The mentioned plane will then pass through the origin. Let now a second linear form

$$y' = cx + dx'$$

of the same variables x,x' be given. Assuming $a \neq o$ we can calculate x from the first form and substitute it in the second

$$x = \frac{1}{a}\, y - \frac{b}{a}\, x'$$

$$y' = \frac{c}{a}\, y + \left(d - \frac{bc}{a}\right) x'$$

this results in two new linear forms which are so characterized that the independent variable x has become a function, while y has become an indpendent variable. We say simply, x and y have been exchanged , and we call the whole process an exchange step . It forms the basis for almost all algorithms of linear algebra, which are mostly composed of several exchange steps.

The previous example of an exchange step can be represente by the following scheme:

	x	x'
y =	a	b
y' =	c	d

⟶

	x	x'
x =	$\frac{1}{a}$	$-\frac{b}{a}$
y' =	$\frac{c}{a}$	$(d - \frac{bc}{a})$

The column x and the row y of the exchanged variables intersect at the element a which we call the pivot element or pivot. The transformation of the element d, which lies in the same diagonal opposite to a, is the most complicated; the correction $\frac{bc}{a}$ can be obtained by multiplying the two elements of the othe diagonal and dividing the result by the pivot. We shall call t

rule the rectangle rule.

We proceed now to linear forms of several variables, for example in the schematic form:

	x_1	x_2	x_3	x_4
$y_1 =$	a_{11}	a_{12}	a_{13}	a_{14}
$y_2 =$	a_{21}	a_{22}	a_{23}	a_{24}
$y_3 =$	a_{31}	a_{32}	a_{33}	a_{34}
$y_4 =$	a_{41}	a_{42}	a_{43}	a_{44}

Explicitly we write for example

$$y_2 = a_{21}\, x_1 + a_{22}\, x_2 + a_{23}\, x_3 + a_{24}\, x_4$$

r in general

$$y_i = \sum_{k=1}^{4} a_{ik}\, x_k \quad , \qquad (i = 1,2,3,4).$$

ow, let us exchange an indpendent variable with a dependent one, .g. $x_3 \longleftrightarrow y_2$. The pivot column x_3 intersects the pivot row t the pivot element a_{23}; the exchange is possible only if $a_{23} \neq 0$. x_2 lays the role of x in the previous elementary example, whereas 2 that of y. All other variables x_1, x_2, x_4 and y_1, y_3, y_4 ehave exactly in the same manner as x' & y' respectively. herefore, we get

	x_1	x_2	y_2	x_4
$y_1 =$	α_{11}	α_{12}	$\frac{a_{13}}{a_{23}}$	α_{14}
$x_3 =$	$-\frac{a_{21}}{a_{23}}$	$-\frac{a_{22}}{a_{23}}$	$\frac{1}{a_{23}}$	$-\frac{a_{24}}{a_{23}}$
$y_3 =$	α_{31}	α_{32}	$\frac{a_{33}}{a_{23}}$	α_{34}
$y_4 =$	α_{41}	α_{42}	$\frac{a_{43}}{a_{23}}$	α_{44}

where $\alpha_{11} = a_{11} - \frac{a_{21}\, a_{13}}{a_{23}}$ $\qquad \alpha_{12} = a_{12} - \frac{a_{22}\, a_{13}}{a_{23}}$

$$\alpha_{14} = a_{14} - \frac{a_{24}\, a_{13}}{a_{23}}$$

$$\alpha_{31} = a_{31} - \frac{a_{21}\, a_{33}}{a_{23}} \qquad \alpha_{32} = a_{32} - \frac{a_{22}\, a_{33}}{a_{23}}$$

$$\alpha_{34} = a_{34} - \frac{a_{24}\, a_{33}}{a_{23}}$$

$$\alpha_{41} = a_{41} - \frac{a_{21}\, a_{43}}{a_{23}} \qquad \alpha_{42} = a_{42} - \frac{a_{22}\, a_{43}}{a_{23}}$$

$$\alpha_{44} = a_{44} - \frac{a_{24}\, a_{43}}{a_{23}}$$

The rules for an exchange step can be summarized as follows:

1. The pivot is changed into its reciprocal value.
2. All other elements of the pivot column are divided by the
3. " " " " " " row " " " "

and then given opposite signs.

4. An element of the rest of the matrix is transformed by forming the rectangle of 4 elements, which contains the pivot in the apposite corner; then the rectangle rule is applied.

Example:

	x_1	x_2	x_3
$y_1=$	3	5	1
$y_2=$	2	4	5
$y_3=$	1	2	2
	-0.5		-1

	x_1	y_3	x_3
$y_1=$	0.5	2.5	-4
$y_2=$	0	2	1
$x_2=$	-0.5	0.5	-1

To simplify the calculations we wrote under the old matrix the new pivot row (excluding the element in the pivot column). We denote this row by the cellar row. Thus we can modify the previous rule 4 as follows:

4'. An element of the rest of the matrix is transformed by adding to it the product of the element below it in the cellar row and the element beside it in the pivot column.

II.2 The sum check.

Let us suppose that the sum of the elements in each row of the given matrix is equal to unity. This means: if we give the independent variables x the value 1, the dependent variables y will also have the value 1. This also holds after an exchange step so that the row sum must again be unity. The value 1 of the row sums can be artificially produced by introducing a new variable $\sum$ beside the x in the linear forms whose coefficient will be so chosen that the row sum becomes unity. In the previous example this appears as follows:

	x_1	x_2	x_3	Σ
$y_1 =$	3	5	1	-8
$y_2 =$	2	4	5	-10
$y_3 =$	1	2	2	-4
	-0.5		-1	2

	x_1	y_3	x_3	Σ
$y_1 =$	0.5	2.5	-4	2
$y_2 =$	0	2	1	-2
$x_2 =$	-0.5	0.5	-1	2

During an exchange step we treat the Σ column as if it were an additional x column. The sum check thus means that in the new scheme the row sums should be equal to unity.

II.3 The Gauss's Elimination Method (Single-division Scheme).

Let us have a system of linear equations as that of the following example:

$$3x_1 + 5x_2 + x_3 = -4$$
$$2x_1 + 4x_2 + 5x_3 = 9$$
$$x_1 + 2x_2 + 2x_3 = 3$$

This can be written in the form:

$$y_1 = 3x_1 + 5x_2 + x_3 + 4$$
$$y_2 = 2x_1 + 4x_2 + 5x_3 - 9$$
$$y_3 = x_1 + 2x_2 + 2x_3 - 3$$

or in the table form

	x_1	x_2	x_3	1
$y_1=$	3	5	1	4
$y_2=$	2	4	5	-9
$y_3=$	1	2	2	-3

It is required to find the values of x for which all y will be equal to zero. Using the exchange process we can exchange the x against the y and transform the last column, the column of the constants, in the same manner as before. After an exchange step we need not to write down in the new table either the y column or the x row of the exchanged variables, because, the y variable will be set equal to zero while the x row is exactly the cellar row under the previous table. Therefore, the sucessive exchange steps will yield the following:

	x_1	x_2	x_3	1
$y_1=$	3	5	1	4
$y_2=$	2	4	5	-9
$y_3=$	1	2	2	-3
$x_2=$	-0.5		-1	1.5

	x_1	x_3	1
	0.5	-4	11.5
	0	1	-3
$x_1=$		8	-23

x_3	1
1	-3

The last table means that $x_3-3=0$, or $x_3=3$. All other unknowns can be obtained by substitution in the cellar rows backwards, namely:

$$x_1 = 8 x_3 - 23 = 1$$

$$x_2 = -0.5 x_1 - x_3 + 1.5 = -2$$

So for solving the given system we first construct an auxiliary triangular system, and then solve it. The process of finding the coefficients of the traingular is called the forward course, and the process of obtaining its solution, the return course, or back solution.

The number of multiplications and divisions necessary for finding the solution of a system of n equations by the Gauss's method is equal to $\frac{n}{3}(n^2 + 3n - 1)$.

II.4 Dominating Diagonals-Concentrated Algorithm:

In the following we shall discuss a practically important case. Let us suppose that in the given matrix the elements of the leading diagonal strongly dominate the other elements of the matrix. This assumption allows us to choose the pivot at the left upper corner of the matrix. After an exchange step the matrix elements will suffer from certain corrections which are formed by the elements outside the leading diagonal and they will therefore be small according to the above assumption. Thus, the diagonal remains dominating and the new pivot can again be chosen at the left upper corner. This holds also for the next elimination steps.

As an example let us have the following system:

x_1	x_2	x_3	x_4	1
a_{11}	a_{12}	a_{13}	a_{14}	C_1
a_{21}	a_{22}	a_{23}	a_{24}	C_2
a_{31}	a_{32}	a_{33}	a_{34}	C_3
a_{41}	a_{42}	a_{43}	a_{44}	C_4
$x_1 =$	α_{12}	α_{13}	α_{14}	γ_1

The tables of the next Gauss-Eliminations will have the forms:

a_{22}	a_{23}	a_{24}	C_2
a_{32}	a_{33}	a_{34}	C_3
a_{42}	a_{43}	a_{44}	C_4

$x_2 =$ $\quad \alpha'_{23} \quad \alpha'_{24} \quad \gamma'_2$

a_{33}	a_{34}	C_3
a_{43}	a_{44}	C_4

$x_3 =$ $\quad \alpha''_{34} \quad \gamma''_3$

x_4	1
a_{44}	C_4

$x_4 =$ $\quad \gamma''_4$

In order to save the writing space we press all above table in one which contains only the pivot column and the corresponding cellar rows of the successive steps:

a_{11}	α_{12}	α_{13}	α_{14}	γ_1	$= x_1$
a_{21}	a'_{22}	α'_{23}	α'_{24}	γ'_2	$= x_2$
a_{31}	a'_{32}	a''_{33}	α''_{34}	γ''_3	$= x_3$
a_{41}	a'_{42}	a''_{43}	a'''_{44}	γ'''_4	$= x_4$

We shall show now that the elements of the last table can be directly calculate without any need of intermediate steps. We start with an element of the lower triangle (below the leading diagonal). The performance of the rules of the gauss elimination yields for example:

$$a'_{32} = a_{32} + a_{31}\alpha_{12}$$

or

$$a''_{43} = a'_{43} + a'_{42}\alpha_{23} = (a_{43} + a_{41}\alpha_{13}) + a'_{42}\,\alpha'_{23}.$$

The same holds for the elements of the diagonal

e.g. $$a'''_{44} = a''_{44} + a''_{43}\alpha''_{34} = (a'_{44} + a'_{42}\alpha'_{24}) + a''_{43}\,\alpha''_{34}$$

$$= a_{44} + a_{41}\alpha_{14} + a'_{42}\,\alpha'_{24} + a''_{43}\,\alpha''_{34}$$

Hence, the necessary calculations are summarized as follows:

1. An element of the new scheme which lies in the diagonal or below it, is obtained as the scalar product of the elements in its row on the left side and the elements in its column above it and add to the result the corresponding element of the original matrix.

   As an example for the elements above the diagonal take

$$\alpha'_{34} = -\frac{a''_{34}}{a''_{33}} = -\frac{1}{a''_{33}}(a'_{34}+a'_{32}\alpha'_{24}) = \frac{-1}{a''_{33}}(a_{34}+a_{31}\alpha_{14}+a'_{32}\alpha'_{24})$$

$$\gamma''_{3} = -\frac{c''_{3}}{a''_{33}} = -\frac{1}{a''_{33}}(c'_{3}+a'_{32}\gamma'_{2}) = \frac{-1}{a''_{33}}(c_{3}+a_{31}\gamma_{1}+a'_{32}\gamma'_{2})$$

Hence, we have

2. An element above the diagonal of the new scheme is obtained by performing the rule 1 at first, dividing the result by the diagonal element beside it and the changing the sign.

For the computation of the new table we start by the first column (which is exactly the first column of the given matrix) and the first row. The latter is obtained from the first row of the original matrix divided by $(-a_{11})$. Then we calculate by means of the rules 1,2 respectively, in the same way we read the lines of a book. The unknowns are obtained, as before, by backward substitution in the cellar raws which are ready at hand in the new table.

II.5 Decomposition of a Matrix into Factors.

A system of n linear equations in n unknowns can be written in matrix form

$$A X = F,$$

A being the given matrix, non singular, x and F being columns composed of the values of the unknowns and the constant terms, respectively, and which we shall speak of as vectors.

The Gauss method, performed with a fixed order of leading elements, consists in replacing the given system by an equivalent triangular system by combining the equations linearly; this reduces to combining linearly the rows of A and F. In the course of its application in using the single-divison scheme, we are obliged to add to the elements of some rows elements proportional to the elments of the preceding rows, i.e. to effect upon matrix A certin elementary transformations. The result of several transformations is equivalent to a premultiplication of the matrix A by some triangular matrix of the form.

$$\begin{bmatrix} \gamma_{11} & 0 & \cdots & 0 \\ \gamma_{21} & \gamma_{22} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ \gamma_{n1} & \gamma_{n2} & & \gamma_{nn} \end{bmatrix} = \Gamma$$

As a result of these transformations we arrive at a system with a triangular matrix

$$B = \begin{bmatrix} 1 & -\alpha_{12} & \cdots & -\alpha_{1n} \\ 0 & 1 & \cdots & -\alpha_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Thus, $\Gamma A = B$, i.e., $A = \Gamma^{-1} B$, and consequently the matrix A is factorable into the product of two triangular matrices. The concentrated scheme realizes this factorization. Indeed,

$$\Gamma^{-1} = \begin{bmatrix} a_{11} & 0 & 0 & \dots & 0 \\ a_{21} & a'_{22} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a'_{n2} & a''_{n3} & \dots & a^{n-1}_{nn} \end{bmatrix}$$

of the last table of II.4. Hence, $A = \Gamma^{-1} B$. Since the diagonal elements of B are equal to unity such a factorization is unique.

II.6 The Gauss-Cholesky (Square-Root) Method.

We shall show now that in case the matrix of the system is positive-definite, finding the solution may be rendered still easier, since in this case the matrix can be resolved into the product of two triangular matrices of which one is the transpose of the other.

Thus let $A = S^T S$,

where

$$S = \begin{bmatrix} s_{11} & s_{12} & \dots & s_{1n} \\ 0 & s_{22} & \dots & s_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & & s_{nn} \end{bmatrix}$$

Let us determine the elements of the matrix S. In view of the rule for the multiplication of matrices, we have

$$a_{ij} = s_{1i}\, s_{1j} + s_{2i}\, s_{2j} + \ldots + s_{ii}\, s_{ij}, \qquad i < j$$

$$a_{ii} = s_{1i}^2 + s_{2i}^2 + \ldots + s_{ii}^2, \qquad i = j$$

Hence we obtain the formulas for the determination of the s_{ij}:

$$s_{11} = \sqrt{a_{11}}, \qquad s_{1j} = \frac{a_{1j}}{s_{11}}$$

$$s_{ii} = \sqrt{a_{ii} - \sum_{\ell=1}^{i-1} s_{\ell i}^2}, \qquad i > 1$$

$$s_{ij} = \frac{a_{ij} - \sum_{\ell=1}^{i-1} s_{\ell i}\, s_{\ell j}}{s_{ii}} \qquad j > i$$

$$s_{ij} = 0, \qquad i > j.$$

Furthermore, the solution of the system reduces to the solution of two triangular systems. Indeed, the equation

$$AX = F,$$

is equivalent to the two equations

$$S^T K = F, \quad S X = K.$$

The elements of the vector K are determined by recurrence formulae analogous to the formulae for s_{ij}:

$$k_1 = \frac{f_1}{s_{11}}, \quad k_i = \frac{f_i - \sum_{\ell=1}^{i-1} s_{\ell i}\, k_\ell}{s_{ii}}; \qquad i > 1$$

The final solution is found by the formulae

$$x_n = \frac{k_n}{s_{nn}}, \quad x_i = \frac{k_i - \sum_{\ell=i+1}^{n} s_{i\ell}\, x_\ell}{s_{ii}} \qquad i < n.$$

With the square-root method, one has to record only the approximately $\frac{n^2}{2}$ elements of the matrix S and the 2n components of the vectors K and X. This method is now widely employed where the solution of symmetric systems is called for; it can be recommended as one of the most efficient methods.

II.7 The Inversion of a Matrix:

As has already been remarked in the introduction, the problem of solving a non homogeneous linear system and that of inverting a matrix are closely connected with each other.

Indeed, if for the (square) matrix A its inverse A^{-1} is known, then on multiplying the equation.

$$A X = F$$

on the beft by A^{-1}, we obtain

$$X = A^{-1} F.$$

Conversely, the determination of the elements of the inverse matrix may be reduced to the solution of n systems of the form

$$\sum_{k=1}^{n} a_{ik} \alpha_{kj} = \delta_{ij} \qquad (j=1,\ldots,n;\ i = 1,\ldots,n),$$

where δ_{ij} is kronecker's symbol (=1 for i = j; = o for $i \neq j$).

The last relations follow from the definition of the inverse matrix ($AA^{-1} = I$) and the rule for matrix multiplication.

The determination of the n^2 elements of the inverse matrix we shall perform by the exchange method. Let A be the coefficient matrix of n linear forms of n variables X:

$$Y = A X.$$

This means that the dependent variables Y are expresed in terms of the independent variables X. Multiplying the last matrix equation on the left by A^{-1} we get:

$$X = A^{-1} Y.$$

In other words if we can express the X in term of the Y, the new matrix of the coefficients is actually the inverse function. The last step can easily be obtained using the exchange method. This is illustrated in the following example:

	x_1	x_2	x_3
y_1 =	3	5	1
y_2 =	2	4	5
y_3 =	1	2	2
	-0.5		-1

	x_1	y_3	x_3
y_1 =	0.5	2.5	-4
y_2 =	0	2	1
x_2 =	-0.5	0.5	-1
		-5	8

	y_1	y_3	x_3
x_1 =	2	-5	8
y_2 =	0	2	1
x_2 =	-1	3	-5
	0	-2	

	y_1	y_3	y_2
x_1 =	2	-21	8
x_3 =	0	-2	1
x_2 =	-1	13	-5

	y_1	y_2	y_3
x_1 =	2	8	-21
x_2 =	-1	-5	13
x_3 =	0	1	-2

In the last table we have rearranged the rows among themselves and also the column, so that the different variables of the same sort appear in order. The last table gives the inverse matrix.

It should be noted that by each exchange step the new pivot must never lie in a previous pivot row or pivot column. That is to say that the pivot candidates for any exchange step are those elements of the last table which remain after sweaping away all elements of previous pivot rows and columns.

II.8 Application of the Inversion to the Solution of Systems of Linear Equations.

As an example let us have the following three linear equations in the three unknowns x_1, x_2, x_3 :

$$3x_1 + 5x_2 + x_3 = -4$$

$$2x_1 + 4x_2 + 5x_3 = 9$$

$$x_1 + 2x_2 + 2x_3 = 3$$

The matrix of the coefficients is chosen to be that of the previous linear forms y_1, y_2, y_3. Our equations here demand that y_1, y_2, y_3 should take the numerical values -4, 9, 3. If we substitute these values in the inverse forms, given in the last table of the previous paragraph, we directly get:

$$x_1 = 1, \quad x_2 = -2 \quad , \quad x_3 = 3$$

for the solution of the above linear system of equations.

Often we are met with the problem of solving several systems of equations having the same matrix of the coefficients but differ only in the constants of the right-hand sides. In this case one inversion of the matrix is sufficient to obtain the inverse linear forms and then direct substitution of the successive right-hand sides will yield the solution of the different system respectively.

Exceptional cases: Let us consider the four-row matrix:

	x_1	x_2	x_3	x_4
y_1 =	3	5	1	2
y_2 =	2	-4	3	7
y_3 =	4	14	-1	-3
y_4 =	13	7	9	20
	-3	-5		-2

	x_1	x_2	y_1	x_4
x_3 =	-3	-5	1	-2
y_2 =	-7	-19	3	1
y_3 =	7	19	-1	-1
y_4 =	14	-38	9	2
	7	19	-3	

	x_1	x_2	y_1	y_2
x_3 =	-17	-43	7	-2
x_4 =	7	19	-3	1
y_3 =	0	0	2	-1
y_4 =	0	0	3	2

At this point a further exchange is not any more possible and the process of inversion stops. This is because all candidates for the next pivot - which must be in a y-row and a x-column- are all zeros. The exchange is only possible if the pivot element is different from zero. From this situation we deduce the following:

1. The remaining y-rows give the relations

$$y_3 = 2\,y_1 - y_2$$

$$y_4 = 3\,y_1 + 2\,y_2\,,$$

i.e. the given linear forms are dependent on each other. For example, the third form can be obtained by multiplying the first one by 2 and subtracting the second. Therefore, the exchange algorithm permitts to discover such dependence.

2. The x-rows will give:

$$x_3 = -17\,x_1 - 43\,x_2 + 7\,y_1 - 2\,y_2$$

$$x_4 = 7\,x_1 + 19\,x_2 - 3\,y_1 + y_2.$$

This can be considered as a partial inversion

Corollaries:

In an exchange algorithm it is forbidden to choose a pivot which is zero, and hence it is dangerous for the accuracy of numerical calculation to choose a pivot which is small. This is a mere application of the general principle. If an algorithm fails for certain exceptional value of the quantities palying part in it, then it will be inaccurate when the values of these quantities lie in the vicinity of these exceptional values.

Therefore, we have the following result: In order to avoid numerical instability during exchange processes, one should choose among all candidates for the pivot that one having the greatest absolute value.

II.9. The Gauss-Jordan's Method for the Solution of Systems of Linear Equations:

Consider the following m systems of linear equations which have the same matrix of coefficients:

$$A X + B_i = 0 \qquad i = 1, 2, \ldots, m.$$

They differ only in the constant vectors B_i. The present method enables us to solve all systems simulaneously. Moreover, it gives us the final tool which makes the subsequent solution of an additional system (with the same matrix) ready at hand, after a slight manipulation with the corresponding constant column. We shall explain the method in the following:

1. We write the given systems in a schematic form

	x_1	x_2	....	x_n	I	II	....	M	
$y_1=$	a_{11}	a_{12}	....	a_{1n}	b_{11}	b_{12}	....	b_{1m}	$= 0$
$y_2=$	a_{21}	a_{22}	....	a_{2n}	b_{21}	b_{22}	....	b_{2m}	$= 0$
- -	—	—	- - -	—	—	—	- - -	—	
$y_i=$	a_{i1}	a_{i2}	....	a_{in}	b_{i1}	b_{i2}	....	b_{im}	$= 0$
- -	—	—	- - -	—	—	—	- - -	—	
y_n	a_{n1}	a_{n2}	....	a_{nn}	b_{n1}	b_{n2}	....	b_{nm}	$= 0$

Where the column of the matrix (b_{ij}). represent the successive columns of the constants of the corresponding systems.

2. Applying the exchange method we bring the x_i on the left side and the y upwards. In the first step we exchange x_1; we seek in its column the largest element in absolute value and choose it as the first pivot a_{i1} say. The elements of the matrix B undergo the same transformations as those of the matrix A. Now, after the first exchange the scheme appears as follows:

	y_i	x_2	....	x_n	I	II	....	M
$y_1 =$	a_{11}	a_{12}	....	a_{1n}	b_{11}	b_{12}	....	b_{1m}
$y_2 =$	a_{21}	a_{22}	....	a_{2n}	b_{21}	b_{22}	....	b_{2m}
...	. . . . .				. . . . .			
$x_1 =$	$\boxed{a_{i1}}$	a_{i2}	....	a_{in}	b_{i1}	b_{i2}	....	b_{im}
...	. . . . .				. . . . .			
$y_n =$	a_{n1}	a_{n2}	....	a_{nn}	b_{n1}	b_{n2}	....	b_{nm}

Since the y_i will be set at last zero, therefore, it not necessary to calculate the new elements of the pivot column; we write instead the old pivot column and mark out the pivot element (this will be used later). This step (No.2) will be repeated wholely for exchanging x_2, x_3, ... , x_n each by its turn. It should be noted that after each exchange the numbers of candidates for the next pivot (in the next column) will be reduced by one, so that for exchanging x_n we shall find only one candidate for the last pivot. It is also to be noted that after each exchange all pivot columns will be coppied as they a

are in the new scheme, separated by vertical lines and the pivot element itself will be marked out. After the last exchange has been executed, the new b columns will represent the successive solutions of the corresponding systems. The last scheme will contain in place of the matrix A the pivot columns of the successive exchange steps. This may have the form (n=3):

	y_3	y_1	y_2	(o)	(1)	(2)	(3)
$x_2 =$	a_{11}	a'_{12}	a''_{13}	b_1	b'_1	b''_1	b'''_1
$x_3 =$	a_{21}	a'_{22}	a''_{23}	b_2	b'_2	b_2	b'''_2
$x_1 =$	a_{31}	a'_{32}	a''_{33}	b_3	b'_3	b''_3	b'''_3

If we need now to solve a new system having the same matrix of coefficients but with the new constants (b_1, b_2, b_3), we add these as a last column (no. (o)). The first exchange has been done with the pivot a_{31} and the elements of the pivot column are all what we need for the calculation of the new elements b'_1, b'_2, b'_3. Whence

$$b'_1 = b_1 - \frac{a_{11}\, b_3}{a_{31}}, \quad b'_2 = b_2 - \frac{a_{21}\, b_3}{a_{31}}, \quad b'_3 = -\frac{b_3}{a_{31}}.$$

The second exchange with the pivot a'_{12} will yield b''_1, b''_2, b''_3 while the third exchange with the pivot a''_{23} (last pivot) delivers the solution:

$$(x_1, x_2, x_3) \equiv (b'''_3, b'''_1, b'''_2).$$

Example: Solve the two systems of equations:

$$\left.\begin{aligned} 3x_1 + 2x_2 + 4x_3 - 1 &= 0 \\ 6x_1 + 3x_2 + 10x_3 - 1 &= 0 \\ 10x_1 + 4x_2 + 20x_3 - 1 &= 0 \end{aligned}\right\} \quad (I) \qquad \left.\begin{aligned} 3x_1 + 2x_2 + 4x_3 &= 0 \\ 6x_1 + 3x_2 + 10x_3 + 1 &= 0 \\ 10x_1 + 4x_2 + 20x_3 + 2 &= 0 \end{aligned}\right\} \quad (II)$$

In schematic form we have:

	x_1	x_2	x_3	I	II
y_1=	3	2	4	-1	0
y_2=	6	3	10	-1	1
y_3=	**10**	4	20	-1	2

	y_3	x_2	x_3	I	II
y_1=	3	**0.8**	-2	-0.7	-0.6
y_2=	6	0.6	-2	-0.4	-0.2
x_1=	**10**	-0.4	-2	0.1	-0.2

	y_3	y_1	x_3	I	II
x_2=	3	**0.8**	2.5	0.875	0.75
y_2=	6	0.6	**-0.5**	0.125	0.25
x_1=	**10**	-0.4	-3	-0.25	-0.5

	y_3	y_1	y_2	I	II
x_2=	3	**0.8**	2.5	1.6	2
x_3=	6	0.6	**-0.5**	0.25	0.5
x_1=	**10**	-0.4	-3	-1	-2

Therefore, the solution of the first system (I) is:

$(x_1, x_2, x_3) \equiv (-1, 1.5, 0.25)$

and that of the second system (II) is:

$(x_1, x_2, x_3) \equiv (-2, 2, 0.5)$.

Now, if we want to solve a new system subsequently we write down the new constant column in an adjoining column to the last

A matrix and then transform it according to the successive exchanges. Let the new constants be (-1,0,1) then:

	y_3	y_1	y_2	(o)	(1)	(2)	(3)
x_2 =	3	0.8	2.5	-1	-1.3	1.625	3.5
x_3 =	6	0.6	-0.5	0	-0.6	0.375	0.75
x_1 =	10	-0.4	-3	1	-0.1	-0.75	-3

In other words, the solution is:

$$(x_1, x_2, x_3) \equiv (-3, 3.5, 0.75)$$

It should be noted that the constants of the last system are chosen in a manner that each constant is the sum of two corresponding constants in the previous systems (I) & (II). The solution of the last system will then be the sum of the two solutions.

In many cases, we are met with systems of equations whose coefficients and constants are all integers (as in the last example). In such cases it is advisable to work out the problem with integer numbers. The possibility to obtain fractional numbers during processing such systems is only due to the division over the pivot element. If we retard this division to the end of the next exchange step, all intermediate results can be made integer. For, if we have

	x_1	x_2	x_3	
$y_1 =$	a_{11}	a_{12}	a_{13}	
$y_2 =$	a_{21}	a_{22}	a_{23}	divis : 1
$y_3 =$	$\boxed{a_{31}}$	a_{32}	a_{33}	

exchanging x_1 and y_3 without dividing over the pivot a_{31} will yield

	y_3	x_2	x_3	
$y_1 =$	a_{11}	$\boxed{a_{12}\,a_{31} - a_{11}\,a_{32}}$	$a_{13}\,a_{31} - a_{11}\,a_{33}$	
$y_2 =$	a_{21}	$a_{22}\,a_{31} - a_{21}\,a_{32}$	$a_{23}\,a_{31} - a_{21}\,a_{33}$	divis: a_{31}
$x_1 =$	$\boxed{a_{31}}$	$-a_{32}$	$-a_{33}$	

We want to prove now that exchanging x_2 and y_1 (say) will yield in the third column elements other than those in the pivot row which change the sign only) which are divisible by a_{31} (the previous pivot). For example, take the element of the second row and third column, its new value will be:

$$\frac{1}{a_{31}}\Big[\,(a_{23}\,a_{31} - a_{21}\,a_{33})(a_{12}\,a_{31} - a_{11}\,a_{32}) - (a_{13}\,a_{31} - a_{11}\,a_{33}) \cdot (a_{22}\,a_{31} - a_{21}\,a_{32})\Big]$$

$$= \frac{1}{a_{31}}\Big[\,a_{31}\,(\text{integer}) + a_{21}\,a_{33}\,a_{11}\,a_{32} - a_{31}\,(\text{integer}) - a_{11}\,a_{33}\,a_{21}\,a_{32}\Big]$$

$$= \text{integer.}$$

Also the last element of the third column will give

$$-\frac{1}{a_{31}}\left[-a_{33}\,(a_{12}\,a_{31}-a_{11}\,a_{32})+a_{32}\,(a_{13}\,a_{31}-a_{11}\,a_{33})\right]$$

$$=\frac{-1}{a_{31}}\left[a_{31}\,(\text{integer})+a_{33}\,a_{11}\,a_{32}+a_{31}(\text{integer})-a_{32}\,a_{11}\,a_{33}\right]$$

= integer

Hence, for integer problems we have the following rules for an exchange step:

1. Choice of the pivot as before.
2. Pivot column remains unchanged.
3. Pivot row changes the sign.
4. Other elements: apply rectangle rule and divide by the previous pivot (the so called divisor or divis besides. the corresponding scheme).

Applying this method to the last example we have:

	x_1	x_2	x_3		
y_1	3	2	4	-1	0
y_2	6	3	10	-1	1
y_3	**10**	4	20	-1	2

divis : 1

	y_3	x_2	x_3		
y_1	3	**8**	-20	-7	-6
y_2	6	6	-20	-4	-2
x_1	**10**	-4	-20	1	-2

divis : 10

	y_3	y_1	x_3		
x_2	3	[8]	20	7	6
y_2	6	6	[-4]	1	2
x_1	[10]	-4	-24	-2	-4

divis : 8

	y_3	y_1	y_2		
x_2	3	[8]	20	-6	-8
x_3	6	6	[-4]	-1	-2
x_1	[10]	-4	-24	4	8

divis : -4

To get the solutions the last columns should be divided by the last divis, i.e., by $\div 4$; these are the same as before.

II.10 <u>The Inversion of Matrix by the Method of Gauss-Jordan:</u>

As said before, the determination of the inverse matrix $A^{-1}=(\alpha_{ik})$ of the matrix $A=(a_{ik})$ so that $A\,A^{-1} = I$, leads to the solution of n systems of linear equations each having n equations of n unknowns. All these systems have the same matrix of coefficients so that we can make the best use of the Gauss-Jardan's method to obtain simultaueously the required solutions.

Since A^{-1} fulfills the above matrix equation, therefore,

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \cdots & \cdots & \cdots \\ \alpha_{n1} & & \alpha_{nn} \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 0 \\ & & \\ 0 & \cdots & 1 \end{pmatrix}$$

The unknowns are the α_{ik} $(i,k = 1,2,\ldots,n)$, i.e., n^2 unknowns. These do not appear altogether in one system but each n of them in one system, e.g., the system of equations for $\alpha_{11}, \alpha_{21}, \ldots \alpha_{n1}$ is:

$$a_{11}\alpha_{11} + a_{12}\alpha_{21} + \cdots + a_{1n}\alpha_{n1} - 1 = 0$$

$$a_{21}\alpha_{11} + a_{22}\alpha_{21} + \cdots + a_{2n}\alpha_{n1} = 0$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$a_{n1}\alpha_{11} + a_{n2}\alpha_{21} + \cdots + a_{nn}\alpha_{n1} = 0$$

In the Gauss-Jordan method the solution of this system appear in the column of the constants in the last scheme. Hence, we obtain the elements of the first column of A^{-1}. This can also be repeated for all other columns of A^{-1}; the successive constant columns will then be: (o, -1, o, ... , o), ... ,(0,0,...,-1) The matrix of the coefficients is all the same for all n systems of equations. Therefore, the method of Gauss-Jordan can directly be applied to obtain all solutions at the same time - Hence, if we fill the B matrix, in the Gauss-Jordan scheme, by (-I) and then execute this method till the end, we shall get in the place of (-I) the required inverse matrix (the rows should be rearranged, if necessar, according to the order of $x_1, \ldots, x_n$).

As an example, let us have the matrix

	x_1	x_2	x_3				
$y_1=$	[3]	5	1	-1	0	0	
$y_2=$	2	4	5	0	-1	0	divis : 1
$y_3=$	1	2	2	0	0	-1	

	y_1	x_2	x_3				
$x_1=$	[3]	-5	-1	1	0	0	
$y_2=$	2	[2]	13	2	-3	0	divis : 3
$y_3=$	1	1	5	1	0	-3	

	y_1	y_2	x_3				
$x_1=$	[3]	-5	21	4	-5	0	
$x_2=$	2	[2]	-13	-2	3	0	divis : 2
$y_3=$	1	1	[-1]	0	1	-2	

	y_1	y_2	y_3				
$x_1=$	[3]	-5	21	-2	-8	21	
$x_2=$	2	[2]	-13	1	5	-13	divis : -1
$x_3=$	1	1	[-1]	0	-1	2	

Dividing the last matrix by the last divis we get :

$$A^{-1} = \begin{vmatrix} 2 & 8 & -21 \\ -1 & -5 & 13 \\ 0 & 1 & -2 \end{vmatrix}$$

II.11. Correction of the solution by iteration.

Let us have the following system:

$$A X + B = 0$$

From the very beginning we can not guarantee that all intermediate results will be integers. Therefore, it is expectable that these results will contain fractions which need not to terminate, i.e. the number of decimal places representing this fraction need not to be finite. Since the computers represent numbers by a fixed finite number of digits, such fractions will be cut off at a certain decimal place and the last figure to the right is rounded off. This may result in an accumulation of the rounding error in an algorithm so that the results will not be accurate to all available decimal figures. To compensate this, the principle of correctors is introduced. It enables us to improve the results by adding to it certain correctors which are computed iteratively.

In the case of a system of linear equations let $X^{(1)}$ be the first result of solving this system by the method of Gauss-Jordan Since $X^{(1)}$ is not the exact solution, substitution of $X^{(1)}$ in the given system will not fulfill it, but certain residues $R^{(1)}$ will appear, i.e.

$$A X^{(1)} + B = R^{(1)}.$$

Suppose that the exact solution is

$$X = X^{(1)} + \delta X^{(1)},$$

where $\delta X^{(1)}$ represents the vector of correctors.

Hence, substituting this in the original system

$$A X + B = 0$$

we get

$$A (X^{(1)} + \delta X^{(1)}) + B = 0$$

$$A X^{(1)} + A \delta X^{(1)} + B = 0$$

or

$$A \delta X^{(1)} + R^{(1)} = 0.$$

In other words, after obtaining $X^{(1)}$ for the first approximation of the solution, we substitute it in the original system and calculate the residues $R^{(1)}$. With $R^{(1)}$ as new constants we compute the solution of the last system (which has the same matrix as the original system) for the correctors $\delta X^{(1)}$. Adding these correctors to $X^{(1)}$ we obtain a better solution $X^{(2)}$.

This procedure can be repeated several times to obtain new correctors $\delta X^{(1)}, \delta X^{(2)}, \delta X^{(4)}, \ldots$ and so on.

This method is very efficient and enables to improve the results to any degree of accuracy we want so long the whole iteration process is convergent.

As a numerical example we solve the following system:

	x_1	x_2	1
y_1=	137	-100	-1
y_2=	-100	73	-1

(1)

	y_1	x_2	
x_1=	137	0.72993	0.0072993
y_2=	-100	0.007	-1.7299

(2)

x_1=	137	0.72993	180.39
x_2=	-100	0.007	247.13

(3)

For the first solution we get:

$$X^{(1)} = \begin{pmatrix} 180.39 \\ 247.13 \end{pmatrix}$$

$$R^{(1)} = \begin{pmatrix} -0.57 \\ 0.49 \end{pmatrix}$$

The system for the first correctors will have form

	y_1	y_2	(o)	(1)	(2)
$\delta x_1^{(1)}$	137	0.72993	-0.57	0.0041606	-7.7054
$\delta x_2^{(1)}$	-100	0.007	0.49	0.07394	-10.562

The solution $\delta X^{(1)}$ will be

$$\delta X^{(1)} = \begin{pmatrix} \delta x_1^{(1)} \\ \delta x_2^{(2)} \end{pmatrix} = \begin{pmatrix} -7.7054 \\ -10.562 \end{pmatrix}$$

The better solution $X^{(2)}$ is

$$X^{(2)} = X^{(1)} + \delta X^{(1)} = \begin{pmatrix} 180.93 - 7.7054 \\ 247.13 - 10.562 \end{pmatrix} = \begin{pmatrix} 172.68 \\ 236.57 \end{pmatrix}$$

Repeating the whole process again we get.

$$X^{(3)} = X^{(2)} + \delta X^{(2)} = \begin{pmatrix} 173.01 \\ 237.02 \end{pmatrix}$$

The residues in this case are:

$$R^{(2)} = \begin{pmatrix} -0.011 \\ -0.022 \end{pmatrix}$$

and so on.

II.12. The Method of Iteration:

We shall pass now to a description of iterative methods of solving systems of linear equations. These method give the solution of the system in the form of the limit of a sequence of certain vectors, the construction of which is effected by a uniform process called the process of iteration.

Let the system of linear equations be given in the following forms:

$$X = A X + F,$$

where A is the matrix of the coefficients, F the vector of the constants.

Let us construct the indicated sequence of vectors as follows: in the capacity of an initial approximation let us take a certain vector $X^{(0)}$, chosen, generally speaking, quite arbitrarily. Next let us construct the vectors

$$X^{(1)} = A X^{(o)} + F$$

$$X^{(2)} = A X^{(1)} + F$$

. . . . . .

$$X^{(k)} = A X^{(k-1)} + F$$

. . . . . .

If the sequence $X^{(o)}, X^{(1)}, \ldots, X^{(k)}, \ldots$ has the limit X, this limit will be the solution of the system, for on passing to the limit in the equation $X^{(k)} = A X^{(k-1)} + F$ as $k \longrightarrow \infty$, we obtain $X = A X + F$, which proves our statement.

II.13. Gauss-Seidel's Methods

Let a systems of linear equations be given in the form

$$X = A X + F,$$

where A is the given matrix, F the given vector and X the sought vector with components $(x_1, \ldots, x_n)$.

The iterative method of seidel is reminiscent of the ordinary iterative process, with this difference, that is computing the kth approximation to the component x_i, one takes into consideration the kth approximations, already computed, to the components $x_1, \ldots, x_{i-1}$. Explicitly, the computation of the successive approximations is performed by the formulas

$$x_i^{(k)} = \sum_{j=1}^{i-1} a_{ij} x_j^{(k)} + \sum_{j=i}^{n} a_{ij} x_j^{(k-1)} + f_i$$

(instead of by $x_i^{(k)} = \sum_{j=1}^{n} a_{ij} x_j^{(k-1)}$ with the method of iteration).

In a number of cases it turns out that the Seidel process converges faster than the ordinary process of iteration. Nevertheless, it does not always prove to have the advantage over the ordinary iterative process.

Sometimes the Seidel process converges more slowly than the process of iteration. It is even possible that the Seidel method may diverge while the method of iteration converges. The regions of convergence of these two processes are different, overlapping only partially.

II.14. The Inversion of Matrix by Partitioning

The handling of matrices of high orders requires, as a rule, a large number of operations. It is therefore often convenient to reduce a computation involving matrices of high orders to computations upon matrices of lower orders. Such a reduction can be effected by partitioning the given matrices: each matrix may be conceived as composed of several matrices of lower orders, and this subdivision may be carried through in many ways for example:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \left[\begin{array}{c|ccc} a_{11} & a_{12} & a_{13} & a_{14} \\ \hline a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right]$$

$$= \left[\begin{array}{cc|cc} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right]$$

2) We form the matrix

$$(D - C\,A^{-1}\,B) = \begin{pmatrix} 0.69785 & -0.15482 \\ -0.15482 & 0.53222 \end{pmatrix}$$

and find its inverse

$$N = (D - CA^{-1}\,B)^{-1} = \begin{pmatrix} 1.53183 & 0.44560 \\ 0.44560 & 2.00855 \end{pmatrix}.$$

3) We compute the matrices

$$M = -\,N\,C\,A^{-1} = \begin{pmatrix} -1.01148 & -0.26142 \\ -1.37834 & -0.44745 \end{pmatrix}$$

$$L = -\,A^{-1}\,B\,N = \begin{pmatrix} -1.01148 & -1.37834 \\ -0.26142 & -0.44745 \end{pmatrix}$$

and

$$K = A^{-1} - A^{-1}\,B\,M = \begin{pmatrix} 2.50757 & -0.12305 \\ -0.12305 & 1.33221 \end{pmatrix}$$

The sougth inverse matrix will thus be:

$$S^{-1} = \begin{pmatrix} 2.50757 & -0.12305 & -1.01148 & -1.37834 \\ -0.12305 & 1.33221 & -0.26142 & -0.44745 \\ -1.01148 & -0.26142 & 1.53183 & 0.44560 \\ -1.37834 & -0.44745 & -0.44560 & 2.00855 \end{pmatrix}$$

where $C_{ij} = A_{il} B_{lj} + A_{i2} B_{2j} + \ldots + A_{ik} B_{kj}$ $(i,j=1,\ldots,k)$

The last formulae show that operations with matrices partitioned in the manner indicated are to be performed just as if in place of each submatrix there were a number.

As said before, it is sometimes expedient to partition a matrix before inversion. Let us examine the formulae for inverting a matrix of the nth order partitioned into four cells by the scheme

$$S = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right),$$

where A and D are square matrices of orders P and q; $p + q = n$.

Let us seek the inverse matrix also in the form of a cellular matrix:

$$S^{-1} = \left(\begin{array}{c|c} K & L \\ \hline M & N \end{array}\right),$$

K and N being again square matrices of orders P & q.

In conformity with the rule for multiplying partioned matrices, the following matrix equations must hold:

$$\begin{aligned} A K + B M &= I \\ A L + B N &= O \\ C K + D M &= O \\ C L + D N &= I. \end{aligned}$$

Multiplying the third equation on the left by $B\,D^{-1}$ and subtracting it from the first, we obtain

$$(A - B\,D^{-1}\,C)\;K = I,$$

whence

$$K = (A - B\,D^{-1}\,C)^{-1}\ .$$

Furthermore, from the third equation we find that

$$M = - D^{-1}\,C\,K.$$

In like manner we find from the second and fourth equations that

$$N = (D - C\,A^{-1}\,B)^{-1}$$

and

$$L = - A^{-1}\ B\,N.$$

Of course these formulas have been derived on the assumption that all the indicated matrix inversions are realizable.

The inversion of a matrix of order n thus reduces to the inversion of four matrices, of which two are of the order p and two of order q, and to several matrix multiplications.

The formulae developed above can be altered so that for the computation of the matrices k,L,M and N only two matrices, of orders p and q, need be inverted. For, as can readily be verified,

$$N = (D - C\,A^{-1}\,B)^{-1}\ ,\qquad M = - N\,C\,A^{-1}\ ,$$

$$L = - A^{-1}\ B\,N\ ,\qquad K = A^{-1} - A^{-1}\,B\,M,$$

and, analogously,

$$K = (A - BD^{-1}C)^{-1} \quad , \quad L = KBD^{-1}$$

$$M = -D^{-1}CK \quad , \quad N = D^{-1} - D^{-1}CL.$$

The last formulae show that the method of particitioning is convenient by employed when any diagonal cell is easily inverted.

Let us take as an example the inversion of the matrix

$$\left(\begin{array}{cc|cc} 1.00 & 0.42 & 0.54 & 0.66 \\ 0.42 & 1.00 & 0.32 & 0.44 \\ \hline 0.54 & 0.32 & 1.00 & 0.22 \\ 0.66 & 0.44 & 0.22 & 1.00 \end{array}\right)$$

The computation will be performed as follows:

1) We compute the matrix A^{-1}:

$$A^{-1} = \begin{pmatrix} 1.21418 & -0.50996 \\ -0.50996 & 1.21418 \end{pmatrix}$$

and form the products

$$A^{-1}B = \begin{pmatrix} 0.49247 & 0.57698 \\ 0.11316 & 0.19767 \end{pmatrix}, \quad CA^{-1} = \begin{pmatrix} 0.49247 & 0.11316 \\ 0.57698 & 0.19767 \end{pmatrix},$$

$$C\,A^{-1}B = \begin{pmatrix} 0.30215 & 0.37482 \\ 0.37482 & 0.46778 \end{pmatrix}$$

2) We form the matrix

$$(D - C\,A^{-1}\,B) = \begin{pmatrix} 0.69785 & -0.15482 \\ -0.15482 & 0.53222 \end{pmatrix}$$

and find its inverse

$$N = (D - CA^{-1}\,B)^{-1} = \begin{pmatrix} 1.53183 & 0.44560 \\ 0.44560 & 2.00855 \end{pmatrix}.$$

3) We compute the matrices

$$M = -N\,C\,A^{-1} = \begin{pmatrix} -1.01148 & -0.26142 \\ -1.37834 & -0.44745 \end{pmatrix}$$

$$L = -A^{-1}\,B\,N = \begin{pmatrix} -1.01148 & -1.37834 \\ -0.26142 & -0.44745 \end{pmatrix}$$

and

$$K = A^{-1} - A^{-1}\,B\,M = \begin{pmatrix} 2.50757 & -0.12305 \\ -0.12305 & 1.33221 \end{pmatrix}$$

The sougth inverse matrix will thus be:

$$S^{-1} = \begin{pmatrix} 2.50757 & -0.12305 & -1.01148 & -1.37834 \\ -0.12305 & 1.33221 & -0.26142 & -0.44745 \\ -1.01148 & -0.26142 & 1.53183 & 0.44560 \\ -1.37834 & -0.44745 & -0.44560 & 2.00855 \end{pmatrix}$$

CHAPTER III

THE PROPER NUMBERS AND PROPER VECTORS OF A MATRIX.

III.1 The characteristic Polynomial Proper Numbers.

The equation

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} = 0$$

is called the characteristic , or secular, equation of the matrix $A = (a_{ij})$. The left member of this equation, which may be written in the abbreviated form $\phi(\lambda) = |A - \lambda I|$, bears the name characteristic polynomial (or characteristic function) of the matrix. The roots of the characteristic equation are called the proper numbers (or characteristic roots, latent roots, proper values, eigenvalues) of the matrix A.

III.2. The proper vectors of a Matrix.

By a proper vector characteristic vector, latent vector or eigenvector) of a matrix A is meant any non-zero vector X such that

$$A X = \lambda X,$$

where λ is any complex number.

III. 3 Determination of the proper Numbers and Proper vectors of a Matrix.

As has already been defined, the proper numbers of a matrix are the zeros of the characteristic polynomial $\phi(\lambda)$, i.e., the roots of the equations

$$\phi(\lambda) = |A - \lambda I| = (-1)^n \left[\lambda^n - p_1 \lambda^{n-1} - \dots - p_n\right] = 0.$$ The direct computation of the coefficients p_i is extractemely awkward and requires a huge number of operations.

The determination of the components of a proper vector requires the solution of a system of n homogeneous equations in n unknowns; in order to compute all the vector of a matrix, one must solve, generally speaking n systems of the form

$$(A - \lambda_i I) \quad X_i = 0,$$

where $X_i = (X_{1i}, X_{2i}, \dots, X_{ni})$ is the ith proper vector of the matrix A.

As in the preceding chapter, we shall distinguish two groups of methods: exact and iterative. The exact methods, when applied to the first problem, will gives more or less convenient schemes for determining the coefficients p_i. The proper numbers will then be obtained as the solutions of an algebraic equation of the n^{th} degree.

The iterative methods make possible the direct determination of the proper numbers of the matrix, without resorting to the characteristic polynomial. The convergence of the iterative process is determined by the magnitude of the ratio of the moduli of different, neighboring proper numbers, and may be very slow. A proper vector is computed simulatneously with the proper number to which it belongs. As in the solution of linear

systems, the chief merit of the iterative processes consists in the simplicity and uniformity of the operations to be performed.

In all of the following sections we shall assume the elements of the matrix to be real.

III.4. An Exact Method.

Consider a matrix A, whose proper numbers and vectors are to be determined. Construct the matrix $(\lambda I - A)$, the inverse of which $(\lambda I - A)^{-1}$ is called the resolvent of A. Hence, we have $(\lambda I - A)(\lambda I - A)^{-1} = I$.

We shall show the steps of calculation in an example:

$$A = \begin{pmatrix} 1 & -1 & 1 \\ 4 & 6 & -1 \\ 4 & 4 & 1 \end{pmatrix}, \quad (\lambda I - A) = \begin{pmatrix} \lambda-1 & 1 & -1 \\ -4 & \lambda-6 & 1 \\ -4 & -4 & \lambda-1 \end{pmatrix}$$

We determine the resolvent of A by the method of Gauss-Jordan (in compact form):

$(\lambda-1)$	1	-1	-1	0	0	
-4	$(\lambda-6)$	1	0	-1	0	divis:1
-4	-4	$(\lambda-1)$	0	0	-1	
-1	1		1	0	0	
$\lambda^2-7\lambda+10$	$(\lambda-5)$		-4	$-\lambda+1$	0	divis:$(\lambda-1)$
$(-4+8)$	$(\lambda^2-2\lambda-3)$		-4	0	$-\lambda+1$	
. . . . .			. . . . .			divis $(\lambda^2-7\lambda+10)$

Resolvent $= (\lambda I - A)^{-1}$

$(\lambda^2-7\lambda+10$	$(-\lambda+5)$	$(\lambda-5)$
$(4\lambda-8)$	$(\lambda^2-2\lambda-3)$	$(-\lambda+5)$
$(4\lambda-8)$	$(4\lambda-8)$	$(\lambda^2+\lambda+10)$

divis: $\lambda^3 - 8\lambda^2 + 17\lambda - 10 = \Phi(\lambda)$

As already has been defined, if λ is a proper number of A, then, $|\lambda I - A| = 0$ and hence $|(\lambda I - A)^{-1}|$ will be infinite. In other words the proper numbers are the zeros of the common denominator of the resolvent, i.e. the zeros of the last divisor. Therefore, the last divisor is nothing else but the characteristic polynomial of A. The zeros are:

$$\lambda_1 = 5 \ , \ \lambda_2 = 2 \ , \ \lambda_3 = 1 \ \text{(proper numbers)}.$$

In order to determine the corresponding proper vectors we determine the partial fractions of the resolvent:

$$\text{Resolvent} = \sum_{i=1}^{3} \frac{F_i}{\lambda - \lambda_i} = \frac{F(\lambda)}{G(\lambda)}$$

Since the λ_i are simple, therefore,

$$F_i = \frac{F(\lambda_i)}{G'(\lambda_i)}$$

$$\text{Resolvent} = \frac{1}{\lambda - 1} \begin{pmatrix} 1 & 1 & -1 \\ -1 & -1 & 1 \\ -1 & -1 & 1 \end{pmatrix} + \frac{1}{\lambda - 2} \begin{pmatrix} 0 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} +$$

$$+ \frac{1}{\lambda - 5} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$$

The matrices F_i are called Frobenius covariants. They have the following properties:

a) $\sum F_i = I$

b) $F_i^2 = F_i$

c) $F_i F_j = 0$ for $i \neq j$

We shall prove now that the column vectors of the Frobenius convariants are the proper vectors of A. If $\phi(\lambda)$ denotes the characteristic polynomial of A, then, we have

$$(\lambda I - A)^{-1} = \sum_i \frac{F_i}{\lambda - \lambda_i}$$

Multiply both sides by $A \cdot \phi(\lambda)$

$$A \cdot \phi(\lambda)\ (\lambda I - A)^{-1} = \left(\sum_i \frac{A}{\lambda - \lambda_i} F_i\right) \phi(\lambda)$$

Denote by $B(\lambda)$ the matrix $\phi(\lambda)\ (\lambda I - A)^{-1}$

$$A \cdot B(\lambda) = \left(\sum_i \frac{A}{\lambda - \lambda_i} F_i\right) \phi(\lambda)$$

Putting $\lambda = \lambda_1$ we get:

$$B(\lambda_1) = \phi'(\lambda_1) \cdot F_1,$$

where $\phi'(\lambda_1) = \left.\frac{\phi(\lambda)}{\lambda - \lambda_1}\right|_{\lambda = \lambda_1}$

$$\therefore A\ B(\lambda_1) = A \phi'(\lambda_1) F_1$$

In general, we have

$$A(\lambda I - A)^{-1} = \lambda(\lambda I - A)^{-1} - I$$

$$\phi(\lambda)\ A(\lambda I - A)^{-1} = \lambda \phi(\lambda)(\lambda I - A)^{-1} - \phi(\lambda) I$$

$$\therefore \lambda \phi(\lambda)(\lambda I - A)^{-1} - \phi(\lambda) I \Big|_{\lambda = \lambda_1} = \lambda_1 B(\lambda_1)$$

$$A \phi'(\lambda_1) F_1 = \lambda_1 \phi'(\lambda_1) F_1$$

If $\phi'(\lambda_1) \neq 0$, then we get

$$A F_1 = \lambda_1 F_1$$

other words the columns of the matrix F_1 are proper vectors. the same manner we prove the columns of the other F_i are o proper vectors.

. 5. Determination of Proper Numbers and Proper Vectors by Iteration:

Let A be the given matrix and X_o be any vector. We construct et of vectors X_o, X_1, X_2, ... etc, so that

$$X_{k+1} = A X_k$$

se will have a definite limit (and that the proper vector corresponding to the largest proper number λ_1 in absolute value) the following two conditions are staisfied:

.) The proper number λ_1 of largest absolute value is distinct rom the others, i.e., $|\lambda_1| > |\lambda_2| >$ ..... etc.

) The starting vector X_o should not be perpendicular to V_1^T a proper vector of the transposed matrix which corresponds to the me proper number λ_1.

In this case we have for the proper vector

$$V_1 = \lim_{k \to \infty} \frac{X_k}{|X_k|}$$

and for the corresponding proper number:

$$\lambda_1 = \lim_{k \to \infty} \sqrt[k]{|X_k|}$$

$$\lambda_1 = \lim_{k \to \infty} \frac{|X_{k+1}|}{|X_k|}$$

Example:

$$\begin{pmatrix} 1 & -1 & 1 \\ 4 & 6 & -1 \\ 4 & 4 & 1 \end{pmatrix} \quad \underset{X_0}{\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}} \quad \underset{X_1}{\begin{pmatrix} 2 \\ 13 \\ 15 \end{pmatrix}} \quad \underset{X_2}{\begin{pmatrix} 4 \\ 71 \\ 75 \end{pmatrix}} \quad \underset{X_3}{\begin{pmatrix} 8 \\ 367 \\ 375 \end{pmatrix}} \quad \underset{X_4}{\begin{pmatrix} 16 \\ 1859 \\ 1875 \end{pmatrix}} \rightarrow \underset{V_1}{\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}}$$