Nonlinear matrix algebra and engineering applications.
Part 1: Theory and linear form matrix

C. L. Wu (*) and R. J. Adler (**)

ABSTRACT
A matrix vector formalism is developed for systematizing the manipulation of sets of non-linear algebraic equations. In this formalism all manipulations are performed by multiplication with specially constructed transformation matrices. For many important classes of nonlinearities, algorithms based on this formalism are presented for rearranging a set of equations so that their solution may be obtained by numerically searching along a single variable. Theory developed proves that all solutions are obtained.

1. INTRODUCTION
The problem of solving sets of simultaneous, nonlinear algebraic equations by manipulation is considered in this paper. Simultaneous nonlinear algebraic equations arise naturally in many fields of engineering and science, both as an original expression of a physical problem and as a finite difference approximation to differential equations.

A matrix vector formalism is developed for systematically manipulating nonlinear algebraic equations and eliminating variables. This formalism consists of expressing the problem in matrix vector notation, and performing operations with specially constructed transformation matrices. A vital organization is brought into manipulations in this manner, and the theory can be compactly stated. The gain and loss of solution sets is controlled.

Mathematicians concerned themselves with the manipulative solution of simultaneous nonlinear algebraic equations in the latter half of the nineteenth century, but usually worked with only two equations in two variables due to the tedious nature of the manipulations involved. Various references (1, 2, 4, 5) describe procedures such as Bezout's method (2), Sylvester's determinant (1), etc., which were for the most part developed by 1900. These methods were rather tedious and unrelated due to the lack of a general framework such as provided by this paper. Since about 1900 activity in this area has been at a minimum. More recently, the development of the high speed digital computer and numerical methods has made numerical iterative procedures the popular method of solving nonlinear equations. However, iterative numerical procedures leave something to be desired. They sometimes have difficulty in converging to a solution. Perhaps more serious is the difficulty of locating all solutions, when several are present, and unusual solutions such as a continuous arc or region. In short, numerical methods give little insight into, and understanding of, the character of a problem. Often specific numerical procedures must be designed for each new problem on a "cut and try" basis.

The introduction of matrix vector techniques makes it possible to treat four or five simultaneous equations by hand in a period of a few hours. Algorithms based on these same techniques can treat any number of equations, but the practical implementation of these algorithms for large sets of equations depends upon the future development of computers to perform the symbolic manipulations, that is, to perform the tedious algebra.

For broad classes of problems, the final results of applying these matrix vector techniques is a single equation in a single variable. This single equation is then solved numerically. In certain instances, matrix vector techniques permit the transformation of a set of equations into another form from which the solution sets are easily obtained without further manipulation. In general, the techniques presented are most suited to the treatment of equations which contain only multinomial terms, that is, terms of the form

\[ \sum_{i=1}^{n} K_i \alpha_i x_i^\beta \]

where

- \( K \) is a constant
- \( \alpha_i \) are integers
- \( x_i \) are variables

Nonlinearities in the form of transcendental functions may sometimes be handled, depending upon the nature of the problem. The presence of one variable as the argument of any number of transcendental functions can always be handled easily.

There are three important concepts upon which everything rests. First, that matrix vector notation is a generalizing concept which systematizes and simplifies the manipulation of equations. Second, that all

(*) Texaco Research Laboratory, Beacon, N. Y.
(**) Chemical Engineering Department, Case Institute of Technology, Cleveland, Ohio.

manipulations are performed by multiplications with specially constructed transformation matrices. Third, that the gain and loss of solutions depends only upon the nature of the transformation matrices used.

2. CLASSIFICATION AND REPRESENTATION
Quite often a set of \( n \) nonlinear equations in \( n \) unknowns can be represented in one or more of the following three forms.

**Linear Form**

\[
\sum_{j=1}^{n} f_{ij} x_j + C_i = 0 \quad i = 1, 2, \ldots, n \quad (2-1a)
\]

where

\( x_j \) (\( j = 1, 2, \ldots, n \)) are unknowns

\( C_i \) (\( i = 1, 2, \ldots, n \)) and \( f_{ij} \) are either constants or functions of one or more of the unknowns, restricted only in that \( C_i \) and \( f_{ij} \) are continuous functions.

**Polynomial Form**

\[
\sum_{j=1}^{m} g_{ij} x_k^{m+1-j} + d_i = 0 \quad i = 1, 2, \ldots, k, \ldots, n \quad (2-2a)
\]

where

\( x_k \) is one of the \( n \) unknowns

\( d_i \) (\( i = 1, 2, \ldots, n \)) and \( g_{ij} \) are defined in the same way as \( f_{ij} \) with the additional restriction that \( d_i \) and \( g_{ij} \) are not functions of \( x_k \).

**Group Form**

\[
\sum_{j=1}^{\ell} f_{ij} h_j + C_i = 0 \quad i = 1, 2, \ldots, n \quad (2-3a)
\]

where

\( h_j \) (\( j = 1, 2, \ldots, \ell \)) are any functional groups of one or more unknowns

\( C_i \) (\( i = 1, 2, \ldots, n \)) and \( f_{ij} \) are as defined in equation (2-1a).

Each of these forms is useful for attacking different types of equations. Linear form is useful for simultaneously eliminating linear variables; polynomial form, for eliminating polynomial variables, and group form, for eliminating groups of variables. The formal operations associated with each of these three forms are discussed separately in a series of several papers.

Matrix-vector notation is well suited to representing and treating these three forms. Linear form equations

\[
\begin{bmatrix}
  f_{11} x_1 + f_{12} x_2 + \ldots + f_{1n} x_n + c_1 \\
  f_{21} x_1 + f_{22} x_2 + \ldots + f_{2n} x_n + c_2 \\
  \vdots \\
  f_{n1} x_1 + f_{n2} x_2 + \ldots + f_{nn} x_n + c_n
\end{bmatrix} = 0
\]

have the matrix-vector representation

\[
[f_{ij}] [x_j] + [c_i] = 0
\]

where

\( i = 1, 2, \ldots, n \)

\( j = 1, 2, \ldots, n \)

Polynomial form equations

\[
\begin{bmatrix}
  g_{11} x_k^m + g_{12} x_k^{m-1} + \ldots + g_{1m} x_k + d_1 \\
  g_{21} x_k^m + g_{22} x_k^{m-1} + \ldots + g_{2m} x_k + d_2 \\
  \vdots \\
  g_{n1} x_k^m + g_{n2} x_k^{m-1} + \ldots + g_{nm} x_k + d_n
\end{bmatrix} = 0
\]

have the representation

\[
[g_{ij}] [x_k^{m+1-j}] + [d_i] = 0
\]

where

\( i = 1, 2, \ldots, n \)

\( j = 1, 2, \ldots, m \)

Finally, group form equations

\[
\begin{bmatrix}
  f_{11} h_1 + f_{12} h_2 + \ldots + f_{1\ell} h_\ell + c_1 \\
  f_{21} h_1 + f_{22} h_2 + \ldots + f_{2\ell} h_\ell + c_2 \\
  \vdots \\
  f_{n1} h_1 + f_{n2} h_2 + \ldots + f_{n\ell} h_\ell + c_n
\end{bmatrix} = 0
\]

have the representation

\[
[f_{ij}] [h_j] + [c_i] = 0
\]

where

\( i = 1, 2, \ldots, n \)

\( j = 1, 2, \ldots, \ell \)

Equations (2-1c), (2-2c) and (2-3c) are all of the form

\[
BX + C = 0
\]

Definition

The matrix \( B \) is called a coefficient matrix. The matrix formed by joining the column \( C \) to the columns of \( B \) by bordering on the right, is called an augmented matrix \( A \).

All \( A \) and \( B \) matrices may be classified on the basis of the number of variables contained in them. It is convenient to call a matrix a function of those variables which appear in the elements of that matrix. Throughout this paper, unless specifically mentioned otherwise, all elements contained in matrices are restricted to continuous functions.

3. EQUIVALENCE OF EQUATIONS
In order to solve a set of nonlinear equations it is often desirable to manipulate or transform them into a more convenient form. These manipulations may introduce or delete solution sets. It is, therefore, of great interest to consider a very useful class of trans-
formations and the conditions under which these transformations add or delete solution sets. This leads naturally to the notion of equivalent, subordinate, or dominant equations.

Before proceeding to an examination of these transformations, several definitions are given which make the Theorems given later in this section more precise and succinct.

Definition
An equation vector \( F = [f_i] \) is an \( nxl \) matrix representing the left hand side of any set of \( n \) non-linear equations containing \( n \) variables.

\[
\begin{align*}
f_1 &= 0 \\
f_2 &= 0 \\
\vdots \\
f_n &= 0
\end{align*}
\]

Definition
A set of equations \( F_2 = 0 \) is said to be equivalent \((\sim)\) to a set of equations \( F_1 = 0 \) if and only if the solution sets of \( F_2 = 0 \) are identical with the solution sets of \( F_1 = 0 \). In symbolic notation

\[
F_1 = 0 \sim F_2 = 0.
\]

Definition
A set of equations \( F_2 = 0 \) is said to be subordinate \((\subset)\) to a set of equations \( F_1 = 0 \), if and only if the solution sets of \( F_2 = 0 \) are a subset of the solution sets of \( F_1 = 0 \). In symbolic notation

\[
F_1 = 0 \subset F_2 = 0.
\]

Definition
A set of equations \( F_2 = 0 \) is said to be dominant \((\supset)\) to a set of equations \( F_1 = 0 \), if and only if the solution sets of \( F_1 = 0 \) are a subset of the solution sets of \( F_2 = 0 \). In symbolic notation

\[
F_1 = 0 \supset F_2 = 0.
\]

Definition
If a set of equations \( F_2 = 0 \) is dominant to a set of equations \( F_1 = 0 \), any solution sets of \( F_1 = 0 \) which are not solution sets of \( F_2 = 0 \) are called additional solution sets (with respect to \( F_1 = 0 \)). The notation

\[
F_1 = 0 \supset F_2 = 0 \quad [f=0]
\]

is defined to mean that \( F_2 = 0 \) is dominant to \( F_1 = 0 \) and if any additional solution sets exist, they must satisfy \( f = 0 \), where \( f \) is a continuous function.

Definition
If a set of equations \( F_2 = 0 \) is subordinate to a set of equations \( F_1 = 0 \), any solution sets of \( F_1 = 0 \) which are not solution sets of \( F_2 = 0 \) are called missing solution sets (with respect to \( F_1 = 0 \)). The notation

\[
F_1 = 0 \subset F_2 = 0 \quad [f=0]
\]

is defined to mean that \( F_2 = 0 \) is subordinate to \( F_1 = 0 \) and if any missing solution sets exist, they must satisfy \( f = 0 \), where \( f \) is a continuous function.

Definition
Any function which is zero for all sets of values of its variables is said to be identically zero.

Definition
A transformation matrix \( P = [p_{ij}] \) is any \( nxn \) matrix whose elements \( p_{ij} \) are either constants or continuous functions of one or more variables such that \( |P| \) is not identically zero.

**Theorem 3-1**

If (1) \( P \) is a transformation matrix

(2) \( F \) is an equation vector

then

\[
F = 0 \supset PF = 0 \quad [|P|\neq 0]
\]

**Proof**
The proof consists of two parts. First it is proved that \( PF = 0 \) is dominant to \( F = 0 \). Second it is proved that any additional solution sets, if they exist, must satisfy \( |P| = 0 \).

Clearly every solution set of \( F = 0 \) is also a solution set of \( PF = 0 \); thus, \( PF = 0 \) is dominant to \( F = 0 \). If there exists any additional solution set, i.e., a solution set which satisfies \( PF = 0 \) but which does not satisfy \( F = 0 \), let it be denoted by \( \gamma \). Substituting \( \gamma \) into \( PF = 0 \) and formally using Cramer's rule yields

\[
|P(\gamma)|F_j(\gamma) = 0 \quad j = 1, 2, \ldots, n
\]

As \( |P(\gamma)| \neq 0 \), one has a system of homogeneous equations with a solution different from null vector 0. Hence \( |P(\gamma)| = 0 \).

**Corollary**

If (1) \( P \) is a transformation matrix

(2) \( F \) is an equation vector

(3) \( |P| \neq 0 \) for any set of values of the variables contained in \( P \),

then

\[
F = 0 \sim PF = 0
\]

Any transformation matrix \( P \) has a formal inverse \( P^{-1} \) with the property

\[
P P^{-1} = P^{-1} P = I
\]

where \( I \) is the unit matrix. The construction of \( P^{-1} \) is the same whether \( P \) contains elements which are constants or functions. The inverse of any transformation matrix \( P \) can be always constructed without assigning a set of values to the variables contained in \( P \).

There may exist sets of values for the variables contained in \( P \) such that \( |P| = 0 \). Thus \( P^{-1} \) usually contains elements which are discontinuous functions.
The following Theorem deals with the special matrix $P^{-1}$ whose elements may not be continuous functions, but whose inverse $P$ has elements which are continuous functions.

**THEOREM 3-2**

If (1) $P$ is a transformation matrix
(2) $F$ is an equation vector
then
$$F = 0 \quad \Rightarrow \quad P^{-1}F = 0$$
$$|P| = 0$$

**Proof**

Let $G = [G_i]$ be the nx1 matrix and $G = P^{-1}F$.

According to Theorem 3-1
$$G = 0 \quad \Rightarrow \quad PG = 0$$
$$|P| = 0$$

Premultiplying $G = P^{-1}F$ by $P$ yields $PG = F$.

Substituting these two identities into the above equations yields
$$P^{-1}F = 0 \quad \Rightarrow \quad F = 0$$
$$|P| = 0$$

**THEOREM 3-3**

Any matrix $Q$ can always be expressed as the product $D^{-1}P$, where $D$ and $P$ are transformation matrices, and $D$ is diagonal.

**Proof**

Any matrix $Q$ can be represented by
$$Q = [q_{ij}] = \left[ \begin{array}{cccc} a_1 & b_1 & \cdots & b_n \\ a_2 & b_2 & \cdots & b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n & b_n & \cdots & b_n \end{array} \right]$$

where $q_{ij}$ may not be continuous, but where $q_{ij}$ and $q_{ik}$ are constants or continuous functions. The $q_{ij}$ are restricted to be not identically zero.

Let the elements of the diagonal matrix $D = [d_{ij}^{-1}]$ be given by
$$d_{ii} = \prod_{k=1, k \neq i}^{n} q_{ik} = q_{i1} q_{i2} \cdots q_{in}$$

The elements $d_{ii}$ are continuous since the $q_{ik}$ are continuous.

The $ij$th element of $DQ$, $(DQ)_{ij}$, is
$$\sum_{r=1}^{n} d_{ir} \overline{d_{rj}} q_{rj} = d_{ii} q_{ij}$$
$$= \prod_{k=1, k \neq j}^{n} \frac{q_{jk}}{q_{ij}} q_{ij} \prod_{k=1}^{n} q_{ik}$$

These elements are constants or continuous functions since the factors $\prod_{k=1}^{n} q_{ik}$ are constants or continuous functions. Since the elements $(DQ)_{ij}$ define a P matrix, it has been shown that
$$DQ = P$$

Since $|D| \neq 0$ identically, $D^{-1}$ exists. Premultiplying on the left by $D^{-1}$ yields
$$Q = D^{-1}P$$

In order to illustrate the application of Theorem 3-3, consider the following case when the $Q$ matrix is used for the premultiplication of $F = 0$. Then
$$QF = 0 \Rightarrow D^{-1}PF = 0$$
$$F = 0$$

according to Theorems 3-3, 3-2 and 3-1.

### 4. LINEAR FORM MATRICES

Linear form matrices are useful for eliminating variables which appear linearly in sets of equations. Whenever it is possible to select $m$ equations which contain $m$ variables in linear fashion only, it is possible to reduce $n$ equations in $n$ variables to $n-m$ equations in $n-m$ variables.

Section 4.1 describes the various operations which may be performed on linear form matrices. Section 4.2 develops the concepts of rank, linear dependence and nonlinear dependence. Sections 4.3, 4.4 and 4.5 describe three techniques of elimination. Systematic elimination is based upon a generalization of Cramer's rule. Singular elimination is based upon the concept of nonlinear dependence. Triangular elimination is the most mechanical of these elimination techniques.

#### 4.1 Linear form matrix operations

Before formal rules are given for the manipulation of linear form matrices, an important difference between sets of linear equations and sets of nonlinear equations should be noted. The linear form coefficient and augmented matrices $B$ and $A$ are unique for a set of fixed linear equations, but are not unique for a set of fixed nonlinear equations. The non-uniqueness of $B$, for example, is illustrated with the following set of nonlinear equations.

$$\begin{align*}
x_1 x_2^2 + x_2 + x_2^2 x_3 - 12 &= 0 \\
x_1 x_2 + x_2^2 x_3 - 2x_2^2 x_3 + 4 &= 0 \\
x_1 + x_2^2 - x_3 - 2 &= 0
\end{align*}$$

The linear form coefficient matrix $B$ may take a number of forms, three of which are

$$\begin{bmatrix}
x_2^2 & 1 & x_2 & x_2^2 & 1 & x_2 \\
x_2 & x_3 & -x_2^2 & 0 & (x_1 + x_3) & -x_2^2 \\
1 & x_2 & -1 & 1 & x_2 & -1
\end{bmatrix}$$

The last linear form coefficient matrix in (4.1-2) is generally preferred, since coefficient matrix $B$ is then a function of only one unknown.

#### 4.1.1. REARRANGEMENT OF ELEMENTS

All of the possible forms of the matrix $B$ are of course equivalent, but for ease of solution, $B$ should be chosen...
so as to contain the smallest possible number of variables. The various equivalent forms of $B$ may be obtained from each other by application of the following obvious matrix rearrangement rule.

**Linear form matrix rearrangement rule**

The following rule can be applied to any element $f_{ij}$ of a linear form augmented matrix $A$ or coefficient matrix $B$. Any element $f_{ij}$ ($i, j = 1, 2, \ldots, n$) can be replaced by zero providing $f_{ij}$ is multiplied by $x_j$, divided by $x_k$, and added to $f_{ik}$ ($k < n$).

The rule is useful whenever $f_{ij}$ can be factored into the form $\phi_j x_k$. In this paper it is assumed that the linear form matrix rearrangement rule has been applied to obtain matrices which contain the smallest possible number of unknowns.

### 4.1.2. COLUMN OPERATIONS

For the augmented matrix $A$, a useful extension of the above rearrangement rule can be stated.

**Column operation rule**

The entire $j$th column ($1 < j < n$) of the linear form augmented matrix $A$ may be replaced by zeros providing the entire $j$th column is multiplied, element by element, by $x_j$ and added, element by element, to the last column.

The column of zeros introduced by a column operation may be deleted if desired, reducing the number of columns in the matrix by one. The validity of the rule is obvious, since the operation performed is equivalent to a simple rearrangement of the original set of nonlinear equations.

Columns operations can be performed on any $nx(n+1)$ linear form augmented matrix $A$ by post multiplication with an $(n+1)x(n+1)$ matrix $R$. The matrix $R = [r_{ij}]$ defined by

$$
[r_{ij}] = \delta_{ij} - \sum_{k=s,t} \delta_{ik} \delta_{kj} + \sum_{k=s,t} \delta_{ik} \delta_{(n+1)jk} x_k
$$

(4.1.2-1)

operates on the columns $s, t, \ldots$ of $A$. For example, the $R$ matrix which operates only on the $s$ and $t$ columns of $A$ is

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

(4.1.2-2)

### 4.1.3. ROW OPERATIONS

Row operations may be performed on linear form augmented matrices. Since it is sometimes desirable to multiply or divide a row by a function, it is useful to distinguish between equivalent, subordinate and dominant row operations.

**Definition**

Any row operation performed on a linear form augmented matrix which produces an equivalent, subordinate or dominant augmented matrix is said to be an equivalent, subordinate or dominant row operation.

**Equivalent row operations**

The following are equivalent row operations:

1. The interchange of two rows.
2. The multiplication of the elements of a row by a constant other than zero.
3. The addition, to the elements of a row, of a constant times the elements of another row.
4. The addition, to the elements of a row, of $\phi$ times the corresponding elements of another row, where $\phi$ is a function of one or more variables, restricted only in that $\phi$ is a continuous function.

**Subordinate row operation**

The division of each element of a row by $\phi_k$, providing $\phi_k$ is a common factor of each element of that row.

Remark

By Theorem 3-2, any missing solution set must satisfy $\phi_k = 0$.

**Dominant row operation**

The following are dominant row operations:

1. The multiplication of the elements of the $k$th row by a function $\phi_k$, defined as above.
2. The addition, to $\phi_k$ times the elements of the $k$th row, of $\phi_j$ times the elements of the $j$th row, where $\phi_k$ and $\phi_j$ are defined as above.

Remark

By Theorem 3-1, if any additional solution sets are produced by either operation 1) or 2), these solution sets must satisfy $\phi_k = 0$.

### 4.2. Rank, linear dependence, and nonlinear dependence

A few useful concepts are singled out for special emphasis here. The concepts of rank, linear dependence, and nonlinear dependence help to clarify the algorithms of solution described in the next section.

#### 4.2.1. RANK

In linear algebra the concept of rank permits elegant statement of the conditions under which a set of linear equations has a solution. This concept of rank can readily be extended to sets of nonlinear equations represented in linear form. Since linear form matrices may contain one or more unknowns, it is
necessary to distinguish between unconditional rank and conditional rank.

**Definition**
The unconditional rank of a linear form coefficient or augmented matrix is the order of the largest square array in that matrix (formed by deleting certain columns and rows) whose determinant does not vanish identically.

**Example**
The three equations in three unknowns
\[
\begin{align*}
x_2 x_1 + x_2^2 + (x_2 + x_3) x_3 - 3 &= 0 \\
x_1 x_1 + x_2^3 + (x_2 + x_3) x_3 + 3 &= 0 \\
x_1 - x_2 + (x_2 - 1) x_3 - 3 &= 0
\end{align*}
\] (4.2.1-1)
are associated with the linear form augmented matrix
\[
\begin{bmatrix}
x_2 & x_2 & x_2 + x_2 - 3 \\
x_2 & x_2 & x_2 & x_2 \\
1 & -1 & x_2 & x_2 - 3
\end{bmatrix}
\] (4.2.1-2)
The unconditional rank of the linear form coefficient matrix is two, since its determinant (which can be formed by deleting the last column) vanishes identically, and the determinant of any 2 x 2 square array does not vanish identically. However, the unconditional rank of the linear form augmented matrix (4.2.1-2) is three, since the determinant of 3 x 3 square arrays (which can be formed by deleting the third column) does not vanish identically. It should be pointed out that in linear algebra the rank of the coefficient matrix must be equal to the rank of the augmented matrix for a solution set to exist. It is not necessarily so in nonlinear algebra. A solution set, namely, \(x_1 = 2, x_2 = -1, x_3 = 3\), does exist for the system of equations (4.2.1-1), while the unconditional ranks of the coefficient and augmented matrices are not equal.

**Remark**
The unconditional rank of a linear form matrix is invariant under premultiplication or post multiplication by any conformable square matrix whose determinant does not vanish identically. The proof is based on the Binet-Cauchy Theorem concerning products of compound matrices and is well known (3). Of course, the use of the matrix rearrangement rule and column operations may affect the unconditional rank of a linear form matrix.

However, it should be noted in the above example that for certain numerical choices of \(x_2\) the rank of the above mentioned coefficient and augmented matrices may change. These possibilities lead to the notion of conditional rank.

**Definition**
Consider any linear form coefficient or augmented matrix containing \(I\) variables, say \(x_1, x_2, \ldots, x_I\) \((I \leq n)\). For each numerical set of these variables \((a_1, a_2, \ldots, a_I)\) where \(a_i \neq 0\) for \(i = 1, 2, \ldots, j\) and \(a_i = 0\) for \(i = j + 1, j + 2, \ldots, I\) \((0 \leq j \leq I)\), the conditional rank of the matrix is defined as the order of the largest square array whose determinant does not vanish, where the array is formed from the matrix after substituting the numerical values \((a_1, a_2, \ldots, a_I)\) and deleting the \((j + 1)\text{st}, (j + 2)\text{nd}, \ldots, I\text{th}\) columns.

**Example**
The conditional rank of (4.2.1-2) with respect to \(x_2 = 0\) is two after substituting \(x_2\) value and deleting the second column.

### 4.2.2. LINEAR DEPENDENCE
The concept of linear dependence can be extended to nonlinear sets of equations written in linear form. The concept applies equally well to linear form coefficient and augmented matrices. Let the row vectors of the coefficient matrix
\[
\begin{bmatrix}
f_{11} & f_{12} & \cdots & f_{1n} \\
f_{21} & f_{22} & \cdots & f_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
f_{m1} & f_{m2} & \cdots & f_{mn}
\end{bmatrix}
\] be denoted by \(X_i\) where \(i = 1, 2, \ldots, n\). Since the elements \(f_{ij}\) \((j = 1, 2, \ldots, n)\) may contain variables, two types of linear dependence are distinguished.

**Definition**
A set of \(m\) row vectors \((m \leq n)\) \(X_1, X_2, \ldots, X_m\) is said to be unconditionally linearly dependent if there exists a set of constants \(\alpha_1, \alpha_2, \ldots, \alpha_m\) (at least one of which is not zero) such that
\[
\alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_m X_m = 0 \quad (4.2.2-1)
\]
identically.

**Example**
The three row vectors
\[
\begin{align*}
X_1 &= (3x_1, 2x_2, 4x_1) \\
X_2 &= (x_1, x_2, 3x_1) \\
X_3 &= (2x_1, x_2, x_1)
\end{align*}
\]
when multiplied by \(\alpha_1 = 1, \alpha_2 = -1, \alpha_3 = -1\) and summed are equal to zero identically, and are therefore unconditionally linearly dependent.

**Definition**
A set of \(m\) row vectors \((m \leq n)\) \(X_1, X_2, \ldots, X_m\) is said to be conditionally linearly dependent with respect to a set of constants \(\alpha_1, \alpha_2, \ldots, \alpha_m\) \((k \leq l)\) if there exists a set of constants \(\alpha_1, \alpha_2, \ldots, \alpha_k\) \((k \leq l)\) such that
\[
\alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_k X_k = 0 \quad (4.2.2-1)
\] identically.

**Example**
The three row vectors
\[
\begin{align*}
X_1 &= (3x_1, 2x_2, 4x_1) \\
X_2 &= (x_1, x_2, 3x_1) \\
X_3 &= (2x_1, x_2, x_1)
\end{align*}
\]
when multiplied by \(\alpha_1 = 1, \alpha_2 = -1, \alpha_3 = -1\) and summed are equal to zero identically, and are therefore unconditionally linearly dependent.
of which is not zero) such that
\[ a_1X_1(a_1, \ldots, a_k, x_{k+1}, \ldots, x_1) + a_2X_2(a_1, \ldots, a_k, x_{k+1}, \ldots, x_1) + \ldots + a_mX_m(a_1, \ldots, a_k, x_{k+1}, \ldots, x_1) = 0 \]
when the constants \( a_1, a_2, \ldots, a_k \) have been substituted for the variables \( x_1, \ldots, x_k \) which appear in the row vectors.

**Example**
Let the first two row vectors of the linear form augmented matrix (4.2.1-2) be denoted by \( X_1(x_2) \) and \( X_2(x_2) \). If \( x_2 = -1 \), there exist \( \alpha_1 = 1, \alpha_2 = 1 \) such that
\[ X_1(-1) + X_2(-1) = 0. \]
Therefore \( X_1 \) and \( X_2 \) are conditionally linearly dependent with respect to \( x_2 = -1 \).

The following Theorem, based on the concept of linear dependence or conditional rank, is useful for testing the validity and character of a numerical solution set.

**Theorem 4-1**
Let \( A \) and \( B \) be \( nx(n+1) \) and \( nxn \) linear form augmented and coefficient matrices associated with a set of \( n \) nonlinear equations with \( n \) unknowns. Let these matrices contain \( I \) variables \( x_j(j = 1, 2, \ldots, I) \) where \( I \leq n \). Let \( A_C \) and \( B_C \) be the corresponding matrices after column operations on all \( j \) columns. A numerical set of values for \( x_j \), say \( \gamma_j(j = 1, 2, \ldots, I) \)

(1) is not a part of any solution set of the original nonlinear equations if and only if the conditional rank of \( A_C \) is not equal to the conditional rank of \( B_C \).

(2) is a part of a unique solution set of the original nonlinear equations if and only if the conditional rank of \( A_C \) is equal to the conditional rank of \( B_C = n - I \).

(3) is a part of an \( (n-I-r) \) fold infinity of solutions of the original nonlinear equations if and only if the conditional rank of \( A_C \) is equal to the conditional rank of \( B_C = r \), where \( r < n - I \).

All of the conditional ranks referred to are course with respect to the set of constants \( \gamma_j \).

**Proof**
Substitution of \( \gamma_j \) for \( x_j \) in \( A_C \) and \( B_C \) yields matrices with constant elements. These constant matrices \( A_C \) and \( B_C \) have the dimensions \( nx(n+1-I) \) and \( nx(n-I) \). Thus the nonlinear equations have been reduced conditionally to a set of \( n \) linear equations with \( n-I \) unknowns. The Theorem now follows directly from classical linear algebra.

### 4.2.3. Nonlinear Dependence

The concept of linear dependence presented in Section 4.2.2 can be extended to nonlinear dependence.

**Definition**
A set of \( m \) row vectors \( X_1, X_2, \ldots, X_m \) is said to be nonlinearly dependent if there exists a set of continuous functions or constants \( \phi_1, \phi_2, \ldots, \phi_m \) (at least one of which contains at least one variable) such that
\[ \phi_1X_1 + \phi_2X_2 + \ldots + \phi_mX_m = 0 \]
identically.

**Example**
The three row vectors of the coefficient matrix (4.2.1-2) when multiplied by \( \phi_1 = -1, \phi_2 = 1, \phi_3 = x_2^2 - x_2 \) and summed are equal to zero identically, and are therefore nonlinearly dependent.

**Theorem 4-2**
The \( p \) row vectors of any \( pxn \) matrix are nonlinearly dependent if and only if the unconditional rank of the matrix is less than \( p \). The \( p \) row vectors of a \( pxn \) matrix are nonlinearly independent if and only if the unconditional rank of the matrix is \( p \).

**Proof**
The proof of this Theorem is quite lengthy and is therefore not given in this paper (for proof see Ref (6)).

### 4.3. Systematic Elimination

Systematic elimination is a formal linear elimination technique which can often be used to eliminate certain linear variables and to reduce the number of equations which must be solved simultaneously. The technique is useful whenever it is possible to select \( m \) equations which contain \( m \) variables in linear fashion only, from the original \( n \) equations in \( n \) variables. After application of this technique, \( n-m \) equations in \( n-m \) variables remain to be solved simultaneously.

**Algorithm**
The algorithm for performing systematic elimination is as follows.

1. The augmented linear form matrix \( A \) which represents the set of \( n \) equations in \( n \) unknowns is written down.
2. The linear form matrix rearrangement rule (Section 4.1.1) is used to minimize the number of variables which appear in \( A \).
3. Select \( R \) rows which contain \( n-R \) or fewer unknowns so as to maximize \( R \).
4. Form a new \( Rx(n+1) \) matrix \( A' \) from these \( R \) rows.
5. Perform a column operation on \( A' \) for each variable contained in \( A' \), and delete the column of zeros created. If \( x_2 \) appears in \( A' \) operate on and then delete the second column, etc. If the resulting matrix contains more than \( R+1 \) columns, continue column operation and subsequent deletion on arbitrarily selected columns until exactly \( R+1 \) columns remain. Through step (5) strict equivalence has been maintained.
6. Using the generalized Cramer's rule (see Theorem 4-3, below) solve for the \( R \) variables which correspond to the first \( R \) columns in terms of the
remaining n-R variables. This step often introduces additional and missing solution sets as shown in the statement following Theorem 4-3.

(7) Substitute the results from step (6) into the unused n-R equations, thus eliminating R variables. The n-R equations obtained contain n-R variables.

(8) These n-R equations, which must be solved simultaneously using some other technique, together with the R equations from step (6), form a new set of equations which is dominant or subordinate to the original set of equations.

Cramer's rule may be generalized in nonlinear algebra. Theorem 4-3 shows that the use of Cramer's rule may introduce additional solution sets. The statement following the Theorem applies specifically to step (6) of the systematic elimination algorithm.

**Definition**

If P is a square transformation matrix, the matrix obtained from P by replacing each element by its cofactor and then interchanging rows and columns is called the adjoint of P, denoted by Adj P.

The following is considered to be the nonlinear generalization of Cramer's rule.

**THEOREM 4-3**

If (1) P is a transformation matrix
(2) F is an equation vector
then

\[
PF = 0 \quad \therefore \quad |P|F = 0
\]

Premultiplying by Adj P = \([P]_{ij}\) yields

\[
\begin{bmatrix}
P_{11} & P_{12} & \cdots & P_{1n} \\
P_{12} & P_{22} & \cdots & P_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
P_{1n} & P_{2n} & \cdots & P_{nn}
\end{bmatrix}
\begin{bmatrix}
P_{1n} & P_{21} & \cdots & P_{n1} \\
P_{12} & P_{22} & \cdots & P_{n2} \\
\vdots & \vdots & \ddots & \vdots \\
P_{1n} & P_{2n} & \cdots & P_{nn}
\end{bmatrix}
\begin{bmatrix}
F_1 \\
F_2 \\
\vdots \\
F_n
\end{bmatrix}
= 0
\]

Premultiplication of PF = 0 by Adj P, according to Theorem 3-1, may introduce additional solution sets which, if they exist, must satisfy \(|\text{Adj } P| = 0\).

In order to use the generalized Cramer's rule to solve for the vector X of the equation BX + C = 0, two operations are required, premultiplication by Adj B followed by premultiplication with

\[
\begin{bmatrix}
|B| \\
-1
\end{bmatrix}
\]

By Theorem 3-1 or 4-3, premultiplication by Adj B is a dominant operation and any additional roots if introduced must satisfy \(|\text{Adj } B| = 0\). By Theorem 3-2, premultiplication by (4.3-1) is a subordinate operation, and any missing solution sets must satisfy \(|B| = 0\).

**Remark**

If no variables are cancelled after premultiplying by (4.3-1), there is no possibility of missing any solution sets.

**Example**

The three equations in three unknowns

\[
\begin{align*}
xy^2 + y^2 + yz &= 0 \\
xy + y^2z + 1 &= 0 \\
x + y^2 - z - 2 &= 0
\end{align*}
\]

are associated with the linear form augmented matrix

\[
\begin{bmatrix}
y^2 & 1 & y & 0 \\
y & 0 & y^2 & 1 \\
1 & y & -1 & -2
\end{bmatrix}
\]

Using the generalized Cramer's rule on, e.g., the first and third equations yields

\[
\begin{bmatrix}
|P| \\
|P|
\end{bmatrix}
\begin{bmatrix}
F_1 \\
F_2
\end{bmatrix}
= 0
\]

Premultiplication of PF = 0 by Adj P, according to Theorem 3-1, may introduce additional solution sets which, if they exist, must satisfy \(|\text{Adj } P| = 0\).

In order to use the generalized Cramer's rule to solve for the vector X of the equation BX + C = 0, two operations are required, premultiplication by Adj B followed by premultiplication with

\[
\begin{bmatrix}
|B| \\
-1
\end{bmatrix}
\]

By Theorem 3-1 or 4-3, premultiplication by Adj B is a dominant operation and any additional roots if introduced must satisfy \(|\text{Adj } B| = 0\). By Theorem 3-2, premultiplication by (4.3-1) is a subordinate operation, and any missing solution sets must satisfy \(|B| = 0\).

**Remark**

If no variables are cancelled after premultiplying by (4.3-1), there is no possibility of missing any solution sets.
which may be factored into

\[ y(y+1)(y+1)(y^2-y-1) = 0. \quad (4.3-6) \]

According to the statement and remark following Theorem 4-3, equations (4.3-5) and (4.3-6) are dominant to (4.3-2) and the following numerical values of \( y \) obtained from the numerator of (4.3-6) contain all of the solutions of the original problem (4.3-2).

\[ y = 0 \quad y = -1 \quad y = +1 \quad y = 1.618 \]
\[ y = -0.618 \quad (4.3-7) \]

If any additional solution sets are present in (4.3-7) they must satisfy \(|\text{Adj } B| = 0 \) which in this example is

\[ \begin{pmatrix} -1 & -y \\ -1 & y^2 \end{pmatrix} = 0 \quad (4.3-8) \]

Thus the roots \( y = 0 \) and \( y = -1 \) may correspond to additional solution sets. These two values are now checked by means of Theorem 4-1. Substituting \( y = 0 \) into (4.3-4) yields

\[ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & -2 \end{bmatrix} \quad (4.3-9) \]

Since the conditional rank of the coefficient matrix of (4.3-9) is one, and the conditional rank of the augmented matrix is two, \( y = 0 \) is not a solution set of the original problem (4.3-2). Substituting \( y = -1 \) into (4.3-4) yields

\[ \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & -1 \end{bmatrix} \quad (4.3-10) \]

Since the conditional rank of the coefficient matrix of (4.3-10) equals the conditional rank of the augmented matrix of (4.3-10) which equals one, \( y = -1 \) is part of a one-fold infinity of solutions of the original problem (4.3-2). By inspection this solution set is

\[ x - z - 1 = 0 \quad y = -1 \quad (4.3-11) \]

which represents a straight line.

The remaining roots of (4.3-7), \( y, y = 1, y = 1.618 \) and \( y = -0.618 \), when substituted into (4.3-5) yield the remaining solution sets

\[ x = 0 \quad y = 1 \quad z = -1 \]
\[ x = -0.618 \quad y = 1.618 \quad z = 0 \quad (4.3-12) \]
\[ x = 1.618 \quad y = -0.618 \quad z = 0 \]

Thus the original problem (4.3-2) has four solution sets as expressed by (4.3-11) and (4.3-12).

4.4. Singular elimination

Singular elimination is a flexible elimination technique which in favorable cases is more powerful than systematic elimination. Occasionally singular elimation can eliminate certain nonlinear variables. The technique is based on the notion of nonlinear dependence already described in Section 4.2.3.

The principle of singular elimination is as follows. Consider the nx(n+1) linear form augmented matrix \( A \) associated with \( n \) nonlinear equations in \( n \) unknowns. After any number \( n+1-m \) of column operations \((0 < m < n+1)\), the equivalent nxm matrix \( A_c \) takes the form

\[
\begin{bmatrix}
  f_{11} & f_{12} & \ldots & f_{1m-1} & f_{1m} \\
  \vdots & \vdots & & \vdots & \vdots \\
  f_{kn} & f_{k2} & \ldots & f_{km-1} & f_{km} \\
  \vdots & \vdots & & \vdots & \vdots \\
  f_{nm} & f_{n2} & \ldots & f_{nm-1} & f_{nm}
\end{bmatrix}
\begin{bmatrix}
  x_1 \\
  \vdots \\
  x_k \\
  \vdots \\
  x_n
\end{bmatrix}
\begin{bmatrix}
  f_{1m} \\
  \vdots \\
  f_{km} \\
  \vdots \\
  f_{nm}
\end{bmatrix}
\quad (4.4-1)
\]

where \( X_i \) \((i = 1, 2, \ldots, n)\) are the \( n \) row vectors \( f_{i1}, f_{i2}, \ldots, f_{im-1} \). If there exists a set of functions or constants \( \phi_i \) \((i = 1, 2, \ldots, m)\) as defined in Section 4.2.3, i. e.,

\[ \phi_1 X_1 + \phi_2 X_2 + \ldots + \phi_m X_m = 0, \]

then

\[
\begin{bmatrix}
  f_{1m} \\
  \vdots \\
  f_{km} \\
  \vdots \\
  f_{nm}
\end{bmatrix}
\begin{bmatrix}
  [\phi_k = 0] \\
  \vdots \\
  X_{k+1} \\
  \vdots \\
  X_n
\end{bmatrix}
= 0
\quad (4.4-2)
\]

so that all solution sets of \( A_c \) must satisfy

\[ \phi_1 f_{1m} + \phi_2 f_{2m} + \ldots + \phi_m f_{nm} = 0 \quad (4.4-3) \]

The validity of this statement is assured by Theorem 3-1, since the matrix on the right can be obtained from the matrix on the left by premultiplication with the transformation \( P \) matrix,

\[
\begin{bmatrix}
  1 \\
  \vdots \\
  \phi_1 \\
  \vdots \\
  \phi_k \\
  \vdots \\
  \phi_{k-1} \\
  \phi_k + 1 \\
  \vdots \\
  \phi_m 0 \ldots 0
\end{bmatrix}
\begin{bmatrix}
  1 \\
  \vdots \\
  \phi_k - 1 \\
  \phi_k + 1 \\
  \vdots \\
  \phi_m
\end{bmatrix}
\quad (4.4-4)
\]

whose determinant is \( \phi_k \).

Clearly is \( \phi_k \) a non-zero constant the matrix on
the right is equivalent to the matrix on the left. Thus, if one or more of the \( \phi \)'s are non-zero constants, it is possible to choose the \( k \)th row in such a way that no additional roots are introduced. If none of the \( \phi \)'s are non-zero constants, then it is usually desirable to assign the subscript \( k \) to the simplest \( \phi \).

The existence of a set of \( \phi \)'s for \( m \) row vectors \( X_i \) is assured since any \( m \) row vectors with \( m-1 \) components are nonlinearly dependent, according to Theorem 4.2.

\[
\begin{bmatrix}
\phi_1 \\
\phi_2 \\
\vdots \\
\phi_m
\end{bmatrix}
\begin{bmatrix}
X_1 \\
X_2 \\
\vdots \\
X_m
\end{bmatrix} = 0, \text{ i.e., } \phi X = 0
\]

Taking the transpose of this equation yields

\[
X^T \phi^T = 0, \text{ i.e., }
\begin{bmatrix}
f_1' \\
f_2' \\
\vdots \\
f_m'
\end{bmatrix}
\begin{bmatrix}
\phi_1 \\
\phi_2 \\
\vdots \\
\phi_m
\end{bmatrix} = 0
\]

Since (4.4-5) consists of \( m-1 \) equations (rows) and \( m \) unknowns \( (\phi_1, \phi_2, \ldots, \phi_m) \), one of the \( \phi \)'s, say \( \phi_k \), is not determined and may be assigned the arbitrary value \( \phi_k = -1 \). Placing the \( k \)th column of (4.4-5) on the right and denoting the new \( \phi \)'s with primes yields

\[
\begin{bmatrix}
f_1' \\
f_2' \\
\vdots \\
f_{m-1}'
\end{bmatrix}
\begin{bmatrix}
\phi_1' \\
\phi_2' \\
\vdots \\
\phi_{m-1}'
\end{bmatrix}
\begin{bmatrix}
k-11 \\
k+11 \\
\vdots \\
m1
\end{bmatrix} = 0
\]

The \( k \)th row is selected to insure that the coefficient matrix of (4.4-6) is non-singular, i.e., has an unconditional rank of \( m-1 \), if that is possible. If the coefficient matrix is non-singular, its inverse exists. Premultiplying by this inverse yields a solution \( \phi_1', \ldots, \phi_{k-1}', \phi_{k+1}', \ldots, \phi_{m}' \). If any of these terms are not continuous, the offending factors may be multiplied through, rendering the set of \( \phi \)'s, \( \phi_1', \phi_2', \ldots, \phi_{m}' \), continuous functions.

If the coefficient matrix of (4.4-6) is singular, i.e., if its unconditional rank is less than \( m-1 \), say \( r \), then \( m-1-r \) of the row vectors of (4.4-5) are nonlinearly dependent on \( r \) row vectors, according to the

\[
\phi_1 X_1 + \phi_2 X_2 + \ldots + \phi_r X_r = X_{r+1}
\]

Theorem 4.2. In this case it is necessary to select the \( r \) row vectors of (4.4-5) which are nonlinearly independent, to add an \( r+1 \)th row, and to search for a set of functions \( \phi_1', \phi_2', \ldots, \phi_r' \) such that

\[
\phi_1' X_1 + \phi_2' X_2 + \ldots + \phi_r' X_r = X_{r+1}
\]

The solution for \( \phi_1', \phi_2', \ldots, \phi_r' \) is unique, since if it is assumed that another set exists, say \( \phi_1', \phi_2', \ldots, \phi_r' \), it must satisfy

\[
\phi_1' X_1 + \phi_2' X_2 + \ldots + \phi_r' X_r = X_{r+1}
\]

Subtracting (4.4-8) from (4.4-7) yields

\[
(\phi_1 - \phi_1') X_1 + (\phi_2 - \phi_2') X_2 + \ldots + (\phi_r - \phi_r') X_r = 0
\]

But all of the coefficients \( (\phi_i - \phi'_i) \), \( i = 1, 2, \ldots, r \), must be zero because if any one should be non-zero, that would indicate that the set of row vectors \( X_1, X_2, \ldots, X_r \) is nonlinearly dependent, in contradiction with the original premise.

Since the \( \phi \)'s are unique they may be conveniently found from equation (4.4-7),

\[
\phi_1 X_1 + \phi_2 X_2 + \ldots + \phi_r X_r = X_{r+1}
\]

Taking the transpose of (4.4-9) yields

\[
X^T \phi^T = 0.
\]

Since the matrix \( X^T \) is of dimension \( (m-1) \times (r+1) \), \( m-1-r \) rows of \( X^T \) are deleted such that the determinant of the first \( r \) columns and the \( r \) remaining rows does not vanish identically. Then \( \phi_1, \phi_2, \ldots, \phi_r \) may be solved for by the same procedure used for \( \phi_1', \ldots, \phi_{k-1}', \phi_{k+1}', \ldots, \phi_m' \).

The process of singular elimination is usually performed on a linear form \( n \times (n+1) \) augmented matrix after \( n-m+1 \) column operations which leaves a \( n \times m \) column operated augmented matrix. This latter matrix can be treated by singular elimination to yield the result shown in (4.4-2). The treatment can be repeated on the matrix on the right hand side of (4.4-2) by choosing any \( m \) row vectors \( X_i \) which are not zero. In this fashion, the treatment can be repeated \( n-m+1 \) times, yielding \( n-m+1 \) equations of the form of (4.4-3). After these \( n-m+1 \) treatments the resulting matrix contains \( n-m+1 \) zero row vectors \( X_i \). If the number of column operations performed is equal to the number \( I \) of variables which appear in the linear form augmented matrix, and the column operations have been performed on the columns corresponding to these \( I \) variables, then \( n-m+1 = I \). Thus singular elimination can reduce the problem of solving \( n \) equations in \( n \) unknowns simultaneously to the problem of solving \( I \) equations in \( I \) unknowns simultaneously.

The reason for not allowing a zero vector to be chosen as one of the \( m \) row vectors \( X_i \) in the above
procedure is that the unconditional rank of the matrix formed from these \( m \) row vectors is often \( m-1 \). In this case, if the \( k \)th row vector is zero the only set of \( \phi \)'s which exist, namely \( \phi_1 = \ldots = \phi_{k-1} = \phi_{k+1} = \ldots = \phi_m = 0 \) and \( \phi_k = \) any non-zero function or constant, is not useful for further reduction by singular elimination.

Example

The same problem solved by systematic elimination in Section 4.3 is now treated by singular elimination.

\[
\begin{align*}
xy^2 + y + yz &= 0 \\
xy + y^2z + 1 &= 0 \\
x + y^2 - z - 2 &= 0
\end{align*}
\] (4.3-2)

The linear form augmented matrix corresponding to these equations is

\[
\begin{bmatrix}
y^2 & 1 & y & 0 \\
y & 0 & y^2 & 1 \\
1 & y & -1 & -2
\end{bmatrix}
\] (4.3-3)

Operating on the second column yields

\[
\begin{bmatrix}
y^2 & y & y \\
y & y^2 & 1 \\
1 & -1 & y^2 - 2
\end{bmatrix}
\] (4.3-4)

A set of \( \phi \)'s for carrying out the singular elimination is defined by

\[
[\phi_1 \phi_2 \phi_3] = 0
\]

Before solving for the \( \phi \)'s, it is convenient to rearrange the above matrix equation to

\[
[\phi_3 \phi_2 \phi_1] = 0
\]

Taking the transpose and assuming \( \phi_2 = -1 \) yields the following matrix associated with the unknowns \( \phi_3 \) and \( \phi_1 \)

\[
\begin{bmatrix}
1 & y^2 & y \\
-1 & y & y^2
\end{bmatrix}
\]

Reducing this matrix by a process similar to Gauss-Jordan reduction yields

\[
\begin{bmatrix}
1 & 0 & y - y^2 \\
0 & 1 & 1
\end{bmatrix}
\]

so that

\[
\begin{align*}
\phi_1 &= 1 \\
\phi_2 &= -1 \\
\phi_3 &= y - y^2
\end{align*}
\]

Forming the transformation matrix of the form of (4.4-4) by letting \( k = 2 \) yields

\[
\begin{bmatrix}
1 & 0 & 0 \\
1 & -1 & y - y^2 \\
0 & 0 & 1
\end{bmatrix}
\]

(4.4-10)

Premultiplying (4.3-4) by (4.4-10) yields the equivalent matrix

\[
\begin{bmatrix}
y^2 & y & y \\
y & y^2 & 1 \\
1 & -1 & y^2 - 2
\end{bmatrix}
\]

so that all of the solution sets of the original problem must satisfy

\[(y + 1)(y - 1)(y^2 - y - 1) = 0.\]

This result is identical with the results obtained when using systematic elimination, as shown in Section 4.3.

4.5. Triangular elimination

Triangular elimination is an alternate reduction technique for treating linear form matrices. As with singular elimination this technique can also reduce \( n \) equations in \( n \) variables to \( I \) equations in \( I \) variables, where \( I \) is the number of variables contained in the linear form matrix. Triangular elimination is a repetitive procedure mechanically similar to the Gauss reduction in linear algebra and may therefore be the best of the linear form reduction techniques for computer implementation.

Algorithm

The algorithm for performing triangular elimination is as follows:

1. The augmented linear form matrix which represents the set of \( n \) equations in \( n \) unknowns is written down.
2. The linear form matrix rearrangement rule (Section 4.1.1) is used to minimize the number of variables which appear in the matrix.
3. Perform any number of column operations (Section 4.1.2). It is usually advantageous to perform \( I \) column operations on the columns corresponding to the \( I \) variables which appear in the matrix.
4. Rearrange the rows such that the 1,1 element of the matrix becomes the simplest possible non-zero function, preferably a non-zero constant. Denote this column operated augmented matrix by \( A_{c}^{[1]} \).

\[
A_{c}^{[1]} = \begin{bmatrix}
f_{11}^{[1]} & f_{12}^{[1]} & \ldots & f_{1m}^{[1]} \\
f_{21}^{[1]} & f_{22}^{[1]} & \ldots & f_{2m}^{[1]} \\
\vdots & \vdots & \ddots & \vdots \\
f_{n1}^{[1]} & f_{n2}^{[1]} & \ldots & f_{nm}^{[1]}
\end{bmatrix}
\] (4.5-1)
Prernultiplying (4.5-1) by the following transformation matrix

\[
\begin{bmatrix}
1 & -f^{[2]}_{32} & f^{[2]}_{22} \\
-f^{[2]}_{21} & f^{[1]}_{11} & 0 \\
-f^{[2]}_{31} & f^{[1]}_{31} & 0 \\
\vdots & \vdots & \vdots \\
-f^{[2]}_{n2} & f^{[1]}_{n2} & 0 \\
\end{bmatrix}
\]

produces the result

\[
\begin{bmatrix}
1 & f^{[1]}_{12} & f^{[1]}_{13} & \cdots & f^{[1]}_{1m} \\
-f^{[2]}_{21} & f^{[1]}_{12} & f^{[1]}_{13} & \cdots & f^{[1]}_{2m} \\
-f^{[2]}_{31} & f^{[1]}_{31} & f^{[1]}_{32} & \cdots & f^{[1]}_{3m} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
-f^{[2]}_{n2} & f^{[1]}_{n2} & f^{[1]}_{n3} & \cdots & f^{[1]}_{nm} \\
\end{bmatrix}
\]

where

\[
(4.5-1) \quad \Leftrightarrow \quad (4.5-3)
\]

\[
[f^{[1]}_{11} = 0]
\]

Rearrange the 2nd through nth rows of the matrix on the right hand side of (4.5-3) such that the 2, 2 element becomes the simplest possible non-zero function, preferably a non-zero constant. Denote this matrix by \(A^{[2]}_c\).

\[
\begin{bmatrix}
f^{[1]}_{11} & f^{[1]}_{12} & f^{[1]}_{13} & \cdots & f^{[1]}_{1m} \\
0 & f^{[2]}_{22} & f^{[2]}_{23} & \cdots & f^{[2]}_{2m} \\
f^{[2]}_{32} & f^{[2]}_{33} & f^{[2]}_{34} & \cdots & f^{[2]}_{3m} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & f^{[2]}_{n2} & f^{[2]}_{n3} & \cdots & f^{[2]}_{nm} \\
\end{bmatrix}
\]

Continuing in this manner it is possible to transform the original matrix (4.5-1) into a dominant matrix which has only zeros in its first m-1 columns below the principle diagonal. Thus

\[
\begin{bmatrix}
f^{[1]}_{11} & f^{[1]}_{12} & f^{[1]}_{13} & \cdots & f^{[1]}_{1m} \\
0 & f^{[2]}_{22} & f^{[2]}_{23} & \cdots & f^{[2]}_{2m} \\
0 & 0 & f^{[2]}_{33} & \cdots & f^{[2]}_{3m} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & f^{[2]}_{nm} \\
\end{bmatrix}
\]

where (4.5-4) \(\Leftrightarrow\) (4.5-6).

\[
f^{[2]}_{22} = 0
\]
If in step (3) 1 column operations are performed, then the solution sets of the matrix on the right hand side of (4.5-7) are found by solving the $n - m + 1 = t$ equations containing $t$ variables

\[
\begin{align*}
\ell_{m-1} & = 0 \\
\ell_{m+1} & = 0 \\
\vdots \\
\ell_{m} & = 0 \\
\end{align*}
\] (4.5-8)

and substituting back into the matrix on the right hand side of (4.5-7), providing none of the diagonal elements in the first $m-1$ columns of the matrix on the right hand side of (4.5-7) are zero.

(9) Any solution set of (4.5-8) which makes one or more of the diagonal elements of the matrix on the right hand side of (4.5-7) zero is tested by substituting it back into the original column operated augmented matrix $A_{c}^{[1]}$. Such solution sets of (4.5-8) are accepted or rejected on the basis of rank according to Theorem 4-1.

**Example**

The same problem solved by systematic elimination in Section 4.3 and singular elimination in Section 4.4

\[
\begin{bmatrix}
y^2 & y & y \\
y & y^2 & 1 \\
1 & -1 & y^2-2 \\
\end{bmatrix}
\] (4.3-4)

is now treated by triangular elimination.

First, the order of the rows is changed to place the third row first

\[
\begin{bmatrix}
1 & -1 & y^2-2 \\
y & y^2 & 1 \\
y^2 & y & y \\
\end{bmatrix}
\] (4.5-9)

Second, premultiplying by a transformation matrix

\[
\begin{array}{ccc}
1 & 0 & 0 \\
-y & 1 & 0 \\
y^2 & 0 & 1 \\
\end{array}
\begin{array}{ccc}
1 & -1 & y^2-2 \\
y & y & 1 \\
y^2 & y & y \\
\end{array}
= \\
\begin{array}{ccc}
1 & -1 & y^2-1 \\
y^2+y & 1-(y^2-2)y \\
y^2+y & y-y^2(y^2-2) \\
\end{array}
\] (4.5-10)

Third, premultiply by a constant transformation matrix

\[
\begin{array}{ccc}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & -1 & 1 \\
\end{array}
\begin{array}{ccc}
1 & -1 & y^2-1 \\
y^2+y & 1-(y^2-2)y \\
y^2+y & y-y^2(y^2-2) \\
\end{array}
= \\
\begin{array}{ccc}
1 & -1 & y^2-1 \\
y^2+y & 1-(y^2-2)y \\
0 & 0 & -(y+1)(y-1)(y^2-2y-1) \\
\end{array}
\] (4.5-11)

Since it has been possible to maintain equivalence throughout these transformations, the solution sets of the original problem must satisfy

\[
y + 1)(y-1)(y^2-y-1) = 0
\] (4.5-12)

i.e.,

\[
y = 1, y = -1, y = 1.618, y = -0.618
\]

Substituting into the matrix on the right hand side of (4.5-11) yields the four solution sets expressed by (4.3-11) and (4.3-12).

**9. BIBLIOGRAPHY**