THE UNIVERSITY OF CALGARY

SOLVING SYSTEMS OF NONLINEAR EQUATIONS USING INTERVAL ANALYSIS

BY

Ved P. Madan

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF COMPUTER SCIENCE

CALGARY, ALBERTA

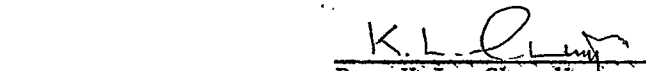
November, 1988

ISBN 0-315-50338-6

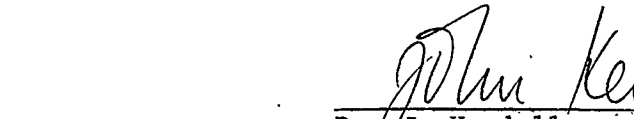
THE UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

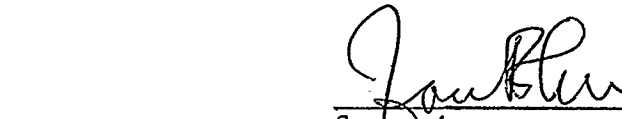
The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled, 'Solving Systems of Nonlinear Equations Using Interval Analysis', submitted by Ved P. Madan in partial fulfillment of the requirements for the master of computing science.



Dr. K.L. Chowdhury
Department of Mechanical Engineering



Dr J. Kendall
Department of Computer Science



Supervisor,
Dr. J. Rokne
Department of Computer Science

November 29, 1988

Abstract

This thesis examines the Hansen-Greenberg approach to solving a system of nonlinear equations using interval analysis. Starting with a brief introduction to interval mathematics and describing some common iterative procedures; we discuss various subalgorithms i.e. the Gauss-Seidel step, the real (noninterval) iteration and the elimination by the LU decomposition; that constitute the main algorithm. The method is illustrated with several numerical examples. Some procedural changes that further improve the efficiency of the algorithm have been included. Alternative procedures involving secondary iterations and the method of splitting are also explored. The thesis concludes with some suggestions for further work.

Acknowledgements

I would like to thank Professor J.G.Rokne for his supervision of this thesis, as well as for his invaluable constructive comments.

I am also grateful to Red Deer College for the sabbatical support, which enabled me to complete my studies at the University of Calgary.

Also, the kind devotion tended to me by my family members Munish, Suparna, and Sushma is greatly appreciated.

Contents

Abstract iii

Acknowledgements iv

1 Introduction 1
- 1.1 Purpose of Thesis . . . 1
- 1.2 Outline of Thesis . . . 3

2 Interval Analysis 6
- 2.1 Interval Arithmetic . . . 6
- 2.2 Fundamental Properties . . . 8
  - 2.2.1 Extended Interval Arithmetic . . . 10
  - 2.2.2 Combinations of Intervals . . . 12
- 2.3 Uses of Interval Analysis . . . 13

3 Solving Systems of Nonlinear Equations 15
- 3.1 The Basic Iterative Methods . . . 15
- 3.2 Solving Systems of Nonlinear Equations Using I.A. . . . 20

4 The Hansen / Sengupta / Greenberg Analysis 26
- 4.1 The Hansen-Sengupta Method . . . 26
  - 4.1.1 Hansen's Preconditioning Technique . . . 28
  - 4.1.2 Krawczyk's Successive Iteration . . . 29
  - 4.1.3 Gaussian Elimination Using Extended Interval Mathematics . 30
- 4.2 The Hansen And Greenberg's Method . . . 32
  - 4.2.1 Initial Hansen-Sengupta Step . . . 35
  - 4.2.2 The Real (Noninterval) Iteration . . . 36
  - 4.2.3 The Elimination Iteration . . . 37
  - 4.2.4 Order of Steps in Hansen-Greenberg Algorithm . . . 39
  - 4.2.5 Numerical Results . . . 43

5 Some Procedural Improvements 47
- 5.1 Improvement of Initial Interval Containing a Fixpoint . . . 47
- 5.2 Modified Gauss-Seidel Iteration . . . 54
- 5.3 Interval Iteration-Alternative to Real Iteration . . . 58

6 Other Suggested Procedures 61
6.1 Newton's Interval Method With Secondary Iterations 61
6.2 Solution By Splitting 68

7 Conclusions 72
7.1 Significant Results of the Thesis 72
7.2 Suggestions for Further Research 76

Bibliography 79

Chapter 1

Introduction

1.1 Purpose of Thesis

The purpose of this thesis is to discuss and analyse some of the basic theoretical results pertaining to the solution of systems of nonlinear equations using interval mathematics. Although the use of intervals in mathematics dates back at least 100 years [19], the real interest in the application of interval mathematics to solving a system of nonlinear equations was generated by the intriguing publication *Interval Analysis* by R.E. Moore [13] in 1966. Moore, who is now regarded as the "father figure" in this area, is credited with extending Newton's method to root finding for a system of nonlinear equations using interval mathematics. Since then, several mathematicians e.g. Krawczyk, Hansen, Sengupta, Greenberg, Nickel, Alefeld, Herzberger, Neumaier and many others (see references) have made significant contributions to the subject. The present work describes the research of Hansen and his colleagues Sengupta and Greenberg [5]. Hansen et. el. have made a noteworthy attempt in presenting a single algorithm of great effeciency for solving nonlinear equations. They utilize the best features available with regard to preconditioning, and they employ a real (noninterval) iteration for obtaining improved results with the Gauss-Seidel step and the elimination procedures. Extended interval mathematics [6] is employed to obtain bounded solution even if the Jacobian matrix contains zero in one or more diagonal elements. Thus, this interval mathematics approach can handle roots of

higher multiplicity, which the regular Newton method can not do effectively. It also tries to contain the growth in the width of intervals by employing a real iteration to locate a "good" point in a box.

Although, Hansen, Sengupta and Greenberg presented an almost complete set of procedures to approximately solve systems of nonlinear equations, the present thesis demonstrates with sufficient examples, how some procedural changes can produce further improvements to the overall algorithm. Under suitable assumptions, the thesis outlines procedures for improving the initial intervals containing a fixpoint, and a modified Gauss-Seidel iteration to enhance the relaxation procedure, if necessary. We also present an interval iteration for finding good starting intervals, to serve as an alternative to the real (noninterval) iteration introduced by Hansen and Greenberg [5].

Keeping in mind the objectives of the thesis, the concluding section describes the solution of systems of nonlinear equations following different methods. In particular, an interval form of Newton's method using a secondary Newton step is developed on the lines of the noninterval procedure. This method has some computational advantages besides offering "cubic convergence". The effectiveness of the method of splitting, when applicable, is also demonstrated.

1.2 Outline of Thesis

The thesis begins with an introductory Chapter on the main purpose of the thesis; which is to describe the solution of systems of nonlinear equations using interval Newton methods. Nonlinear systems are generally approximated by linear systems with the help of the mean value theorem. These linear systems are then solved directly or using iterative methods providing one step in an iteration for the nonlinear system. If the initial expansion is sufficiently close to a fixpoint then the convergence and accuracy of the approximated solution by the iterative procedure is ensured. In this thesis the focus is mainly on solving nonlinear equations using some methods based on interval analysis. Chapter 1 describes briefly the historical development of interval analysis and the interval Newton method as initiated by R.E.Moore [13]. Indepth studies of interval mathematics conducted by Alefeld [1], Hansen [4] and others generated further interest in this problem. Chapter 2 describes some basic results in interval mathematics; which enabled Hansen [5, 6] and his colleagues to deal with unbounded intervals. Properties associated with combinations of intervals and a brief description of the interval Newton method for root finding are also included here for a clear understanding of the ensuing principles.

The solution of systems of the interval linear equations occuring in the iteration steps is handled best with computers using iterative procedures. Chapter 3, therefore contains an account of the basic iterative methods. The methods of Jacobi,

Gauss-Seidel and successive overrelaxation (SOR) are employed in most solution procedures. The Chapter concludes with a brief description of the algebraic techniques, matrix operations and associated steps which will lead to approximate solutions of nonlinear systems, as illustrated in Chapter 4 with the work of Hansen, Sengupta, and Greenberg. The first part of this Chapter contains mainly the contributions of Hansen and Sengupta [6]. The preconditioning technique, first used by Hansen [4], is introduced here. Krawczyk's [8] successive iteration as well as the Gaussian elimination operation using extended interval mathematics (due to Hansen-Sengupta) are examined. The second part of the Chapter studies the Hansen-Greenberg analysis in some detail, starting with the initial Hansen-Sengupta step, and a real iteration. The procedure concludes with the elimination iteration. The order in which these steps must be carried out such that an efficient Hansen-Greenberg algorithm is obtained is discussed. The Chapter concludes with a few illustrations showing the performance of the described algorithm.

We describe some procedural improvements of our own in Chapter 5. The *conciseness* of the initial starting interval is of outmost importance for any iterative procedures associated with interval arithmetics. Accordingly, we first show how the initial interval containing a fixpoint can be improved greatly by finding a *better* approximating matrix for the identity matrix using the Hansen-Greenberg preconditioned system. We also introduce the modified Gauss-Seidel step for obtaining an improved solution, when necessary. The Chapter 5 concludes with the introduction of an interval iteration which may serve as an alternative for the real iteration in the Hansen-Greenberg's analysis.

Chapter 6 deals with the other procedures that may be used effectively for the so-

lution of nonlinear system of equations using interval analysis. Iterative noninterval methods have been used to solve systems of nonlinear equations. Interval techniques also exist in which higher derivatives are used. We present a modified Newton's interval method which uses one step of the Newton operator in the successive iterations. This iteration involves only the first order derivative in the computations. Besides the computational advantages, the method offers *cubic convergence.* A method for solving nonlinear systems of equations by splitting the Jacobian matrix, although limited in scope, is also illustrated in the concluding Section.

Finally Chapter 7 summarizes the significant results of the thesis and also contains some suggestions for further research.

Chapter 2

Interval Analysis

2.1 Interval Arithmetic

The algebraic operations and the terminology of interval arithmetic used in this thesis are presented briefly in this section. For a more extensive introduction see for example Alefeld and Herzberger [1].

The field of real numbers $\Re$ is denoted by lower case letters $x,y,z,...$A subset of $\Re$ of the form

$$X = [x_1, x_2] = \{x | x_1 \leq x \leq x_2, x_1, x_2 \in \Re\}$$

is called a closed real interval or an interval. The set of all closed intervals X, Y, Z in upper case, are members of $I\Re$. Real numbers $x \in \Re$ may be considered special members $[x, x]$ from $I\Re$.

Let $* \in \{+, -, \cdot, /\}$ be a binary operation on the set of real numbers $\Re$, then

$$X * Y = \{z = x * y | x \in X, y \in Y\},$$

defines a binary operation on $I\Re$. It is assumed that $0 \ni Y$ in the case of division.

The operations on intervals $X = [x_1, x_2]$ and $Y = [y_1, y_2]$ may be defined as follows:

$X + Y = [x_1 + x_2, y_1 + y_2]$,

$$X - Y = [x_1 - y_2, x_2 - y_1],$$

$$X \cdot Y = [\min\{x_1y_1, x_1y_2, x_2y_1, x_2y_2\}, \max\{x_1y_1, x_1y_2, x_2y_1, x_2y_2\}],$$

$$X/Y = [x_1, x_2] \cdot [1/y_2, 1/y_1].$$

If X, Y, Z are members of $I\Re$, then it can be shown easily that

$X + Y = Y + X$,

$Y + X = X + Y$, (commutativity).

$(X + Y) + Z = X + (Y + Z)$,

$(X \cdot Y) \cdot Z = X \cdot (Y \cdot Z)$, (associativity).

$X \cdot (Y + Z) \subseteq X \cdot Y + X \cdot Z$, (subdistributivity).

The distributive law is, however, not valid in general. $X = [0, 0]$ and $Y = [1, 1]$ are neutral elements with respect to addition and multiplication.

It should also be noted that $I\Re$ has no zero divisors and $X = [x_1, x_2]$ with $x_1 \neq x_2$ has no inverse elements with respect to $+$ and "$\cdot$".

In fact $X - Y = [0, 0]$, iff $X = Y = [x, x]$

and $X \cdot Y = [1, 1]$, iff $x = [x, x]$ and $Y = [1/x, 1/x]$.

2.2 Fundamental Properties

A fundamental property of interval computations is *inclusion monotonicity* as described by the following:

Theorem 1 *Let* X_k, $Y_k \in I\Re$, $k = 1, 2...$ *such that*

$X_k \subseteq Y_k, k = 1, 2...$

Then for the operations $* \in \{+, -, \cdot, /\}$, *it follows that*

$X_i * X_j \subseteq Y_i * Y_j$.

Hence if $x \in X$ and $y \in Y$ then $x * y \in X * Y$, a property that may be used to embed real algorithms in interval algorithms.

Definition

If $g(x)$ is a continuous unary operation on $\Re$, then

$g(x) = [\min g(x)\ , \max g(x)]\ , x \in X$

defines a corresponding unary operation on $I\Re$.

The unary operations $g(x)$ have the following properties:

$X \subseteq Y \Rightarrow g(X) \subseteq g(Y)$

$x \in X \Rightarrow g(x) \in g(X)$.

We denote the width of an interval $[x_1, x_2]$ by

$w([x_1, x_2]) = x_2 - x_1$.

The distance between a pair of intervals is defined as the nonnegative function

$$d(X, Y) = d([x_1, x_2], [y_1, y_2]) = max(|x_1 - y_1|, |x_2 - y_2|) .$$

Thus a sequence of intervals $A_1, A_2, ..., A_n$, where $A_n = [a_n, b_n]$, converges to the interval $A = [a, b]$ if $d(A_n, A)$ converges to zero i.e. $a_n \to a$ and $b_n \to b$.

We now describe briefly the interval arithmetic relating to *matrix operations* :

Firstly an interval vector is a vector whose components are intervals $X \in I\Re$. In same manner, an interval matrix is a matrix whose elements are intervals. The product of two interval matrices using interval arithmetic is an interval matrix consisting of interval elements. The set of interval vectors with n components and the set of interval $m \times n$ matrices are denoted respectively by $I\Re^n$ and $I\Re^{m \times n}$. The mid-point, radius, and absolute value of a matrix $\mathcal{M} \in I\Re^{m \times n}$ are understood componentwise and denoted by $\hat{\mathcal{M}} = \text{mid}\mathcal{M}$, $\rho(\mathcal{M}) = \text{rad}(\mathcal{M})$ and $|\mathcal{M}|$ respectively. An interval vector is also called a box in the sequel.

A linear interval equation $AX = b$ with coefficient matrix A and right hand side b is defined as the family of linear equations:

$\hat{A}\hat{X} = \hat{b}$

where $\hat{A} \in A, \hat{b} \in b$. The solution set of $AX = b$ is the set $\mathcal{S}$

$$\mathcal{S} = \{\hat{X} \in \Re^n | \hat{A}\hat{X} = \hat{b} \, for some \, \hat{A} \in A, \hat{b} \in b\}.$$

The solution set $\mathcal{S}$ is bounded if A is regular i.e. if all matrices $\hat{A} \in A$ have rank n. A sufficient condition for the regularity of a $n \times n$ interval matrix is that $\hat{A}^{-1}$ exists for each $\hat{A} \in A$. (See Ortega and Reinboldt [17]).

2.2.1 Extended Interval Arithmetic

The extended interval arithmetic introduced first by Alefeld [1] and independently by Hansen [4] makes the following assertions with regard to the binary operation of division.

Let $X = [x_1, x_2]$ and $Y = [y_1, y_2]$. Suppose $0 \in Y$.
We now define
$X/Y =$

$[x_2/y_1, \infty)$, if $x_2 \leq 0$ and $y_2 = 0$,

$(-\infty, x_2/y_2] \cup [x_2/y_1, +\infty)$, if $x_2 \leq 0$, and $y_2 > 0$,

$(-\infty, x_2/y_2]$, if $x_2 \leq 0$ and $y_1 = 0$,

$(-\infty, x_1/y_1]$, if $x_1 \geq 0$ and $y_2 = 0$,

$(-\infty, x_1/y_1] \cup [x_1/y_2, \infty)$, if $x_1 > 0$, $y_1 < 0$ and $y_2 > 0$,

$[x_1/y_2, +\infty)$, if $x_1 \geq 0$ and $y_1 = 0$,

$(-\infty, +\infty)$, if $x_1 \leq 0$ and $x_2 > 0$.

Some basic operations with the above intervals, useful to our calculations with the interval Newton method are:

$$x_i - [c_i, +\infty) = (-\infty, x_i - c_i],$$

$$x_i - (-\infty, d_i] = [x_i - d_i, \infty),$$

$$x_i - (-\infty, \infty) = (-\infty, \infty),$$

$$x_i - (-\infty, d_i] \cup [c_i, +\infty) = (-\infty, x_i - c_i] \cup [x_i - d_i, \infty).$$

Also

$$x_i + [-c_i, \infty) = [x_i - c_i, \infty),$$

$$x_i + (-\infty, \infty) = (-\infty, \infty),$$

$$x_i + (-\infty, d_i] = (-\infty, x_i + d_i),$$

$$x_i, c_i, d_i \in \Re.$$

2.2.2 Combinations of Intervals

Formulas relating to combinations of intervals may be deduced using interval arithmetic. Alefeld-Herzberger [1] lists a number of such formulas.

Some important results in this regard are:
$w([x, y]) = y - x$, $\text{mid}([x, y]) = (x + y)/2$,

$w(xX \pm yY) = |x|w(X) + |y|w(Y)$, ($w$ = width)

$\text{mid}(xX \pm yY) = x \cdot \text{mid(X)} \pm y \cdot \text{mid(Y)}$.

Also if $x, y \in \Re$ and $X = [-x, x]$, $Y = [y_1, y_2]$, then

$XY = X \max(|y_1|, |y_2|)$,

$X/Y = X/y_1$ if $y_1 > 0$,

$X/Y = X/y_2$ if $y_2 < 0$,

$$w(XY) = 2(x)\max(|y_1|, |y_2|).$$

The usual set theoretic operations of intersection and union are applicable for intervals, but these operations can lead to sets that are no longer intervals.

2.3 Uses of Interval Analysis

After the seminal work by Moore [13] a large amount of interest was shown in interval analysis. After all, it was ideally suited to deal with the roundoff errors in floating point computations automatically. This follows from the inclusion monotonicity which allows algorithms developed in real arithmetic to be transformed into machine interval arithmetic algorithms via the use of a machine interval arithmetic embedding the usual interval arithmetic. When the machine operations are performed on intervals then the machine interval arithmetic must be so designed that it rounds outwards. For the details of this process see for example Alefeld and Herzberger [1]. It is now natural to ask why is interval arithmetic not in universal use since it provides guaranteed error bounds. The reason for this is that the simple interval arithmetic versions of numerical algorithms provide only very pessimistic results in general. If the 7×7 Hilbert matrix is inverted using the interval Gaussian inversion procedure then the width of the results (using 26 bit mantissa) is of the order 10. Research in interval arithmetic therefore focusses on developing methods that are taylored to generate narrow intervals. Many of these methods are found in Alefeld Herzberger [1] and Ratschek Rokne [20] as well as in the published literature.

In this thesis we extensively discuss the extension of the usual Newton's method:

$$x^{(n+1)} = x^{(n)} - (J(x^{(n)}))^{-1} \cdot f(x^{(n)})$$

to interval spaces given by the iteration

$$Y^{(n+1)} = m(X^{(n)}) - (J(X^{(n)}))^{-1} \cdot f(m(X^{(n)}))$$

with $X^{(n+1)} = Y^{(n+1)} \cap X^{(n)}$, ($X^{(0)}$ is given) (see Moore [13] and also the development in the following chapters).

In this algorithm, only $J(X^{(n)})$ is evaluated for an interval argument as opposed to some root finding procedures where the evaluation of f over an interval of nonzero width is required. This means that the interval Newton iteration produces sharper results in general. It can be shown that the convergence of the method is quadratic to a simple root X^*

i.e. $||X^{(n+1)} - X^*|| \leq \gamma ||X^{(n)} - X^*||^2$, $n = 0, 1, 2...$

where γ =constant.

Convergence also seems to occur for multiple roots, but it can only be linear convergence. Newton's method also can not be used to bound multiple roots for if $0 \in J(X)$ then $(J(X))^{-1}$ would be undefined unless the extended interval arithmetic approach as described in the previous section is used.

Notice that the obvious extension

$$X^{(n+1)} = X^{(n)} - (J(X^{(n)}))^{-1} \cdot f(X^{(n)})$$

would lead to intervals that grow in size in each iteration, supporting the general requirement of having to taylor, not just translate the methods to interval spaces.

Chapter 3

Solving Systems of Nonlinear Equations

3.1 The Basic Iterative Methods

We now consider methods for solving nonlinear equations:

$$f(x) = 0 \tag{3.1}$$

where $f = (f_1, ..., f_n)^T$ and $f_i(x)$ is a nonlinear function of a real vector $x = (x_1, ..., x_n)^T$.

Using the Mean Value Theorem, we can approximate the system (3.1) at a point x with

$$Ax = b \tag{3.2}$$

where A is an $n \times n$ matrix and b is a $n \times 1$ column vector (see the development in the sequel). The case of principal interest for the applications is one in which A is nonsingular i.e. A^{-1} exists, then system (3.2) has a unique solution $x = A^{-1}b$.

A powerful method for attacking nonlinear and also linear equations is the method of successive substitutions or the method of iteration. This procedure originated in the writings of Heron of Alexanderia [2] in the second century B.C. in connection with the extraction of roots. In modern times Cauchy and Picard [25] have used this technique to establish the existence of solutions of differential equations. An abstract

formulation of these results as the contraction principle was achieved by Banach [3]. A full survey of the application of this principle to theoretical and computational problems can be found in the literature. An equation which can be approximated with sufficient accuracy can be solved numerically, once the convergence of the sequence and approximate error bounds have been established. As an example Newton's method for solving scalar algebraic and transcendental equations is well known for its simplicity and effectiveness. Ostrowski [18] generalized Newtons's method to systems of equations and Kantorovic [11] extended it to the operator equations in Banach spaces. The method now provides a powerful tool for theoretical as well as numerical investigations of nonlinear operator equations.

Using the Mean Value linearization repeatedly gives Newton's iteration formula:

$$x^{(k+1)} = x^{(k)} - (f'(x^{(k)}))^{-1} \cdot f(x^{(k)}), k = 0, 1, 2... \tag{3.3}$$

Procedure (3.3) is restricted by the assumption that $(f'(x^{(k)}))^{-1}$ exists for all k. The advantage to linearization of a nonlinear equation is that the highly developed theory and machinery for solving linear equations or inverting linear operators can be employed. The concept of iteration proceeds from the idea that, if $x^{(0)}$ is close to the actual solution x^*, then $x^{(1)}$ is possibly closer and the iteration procedure may be repeated to obtain successive approximations $x^{(2)}, x^{(3)}, ...$ to x^*. Equation (3.3) is sometimes written as:

$$f'(x^{(k)})x^{(k+1)} = f'(x^{(k)}) \cdot x^{(k)} - f(x^{(k)}). \tag{3.4}$$

Equation (3.4) has some computational advantages over equation (3.3) since the inverse operator is no longer required explicitly.

Variants of Newton's Method:

Many computational methods for the solution of nonlinear operator equations are related in some simple way to Newton's method. For example if an approximation to $(f'(x^{(k)})^{-1}$ is used then a variant of Newton's method would result. This could for example include the computational error in the generation of the exact Newton sequence. Perhaps the variant of Newton's method of greatest interest is the modified form of Newton's method:

$$x^{(k+1)} = x^{(k)} - [f'(x^{(0)})]^{-1} \cdot f(x^{(k)}), k = 0, 1, 2... \tag{3.5}$$

For this procedure the labor of calculating $f'(x^{(0)})$ and its inverse is done only once, and the fixed point is obtained using the modified Newton operator:

$$N(x) = x - [f'(x^{(0)})]^{-1} \cdot f(x). \tag{3.6}$$

The price paid for the computational simplicity of the modified form of Newton's method is a slower rate of convergence.

The Basic Iterative Methods:

We will now discuss briefly the four *basic* iterative methods, namely the Jacobi, Gauss-Seidel, Successive Overrelaxation (SOR), and the Symmetric Successive Overrelaxation (SSOR) methods. For illustration purposes, we apply these methods to the system of linear equations (3.2). One of the simplest iterative methods is that of Jacobi. The iterates $x_r^{(k+1)}$ are given by [17]

$$x_i^{(k+1)} = (1/a_{ii})[b_i - \sum_{j \neq i} a_{ij} x_j^{(k)}], \tag{3.7}$$

where $i = 1, 2, \ldots$ and $A = [a_{ij}], a_{ii} \neq 0$.

If we let $D = \text{diag}\,(a_{11}, \ldots, a_{nn})$ and $B = D - A$, then (3.7) may be written as :

$$x^{(k+1)} = D^{-1}Bx^{(k)} + D^{-1}b, \quad k = 0, 1 \ldots \tag{3.8}$$

If we assume that the computations in (3.7) are done sequentially for $i = 1, 2, \ldots n$, then at the time we are ready to compute $x_i^{(k+1)}$ the components $x_1^{(k+1)}, \ldots, x_{i-1}^{(k+1)}$ are available and we may compute $x_i^{(k+1)}$ using

$$x_i^{(k+1)} = (1/a_{ii})[b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k)}], \quad i = 1, 2, \ldots n. \tag{3.9}$$

The procedure (3.9) is called the Gauss-Seidel iteration. If we let $A = D + L + U$, where L is strictly lower triangular and U is strictly upper triangular part of A, then iteration (3.9) may be written as

$$x^{(k+1)} = -(D + L)^{-1}Ux^{(k)} + (D + L)^{-1}b, \; k = 0, 1, 2, \ldots. \tag{3.10}$$

Both iterations (3.8) and (3.10) are of the form

$x^{(k+1)} = Gx^{(k)} + d, k = 0, 1, 2, \ldots$

and

$x^* = Gx^* + d$, iff $Ax^* = b$.

If $A \in \Re^{n\times n}$ is an M-matrix and $b \in \Re^n$, then the Jacobi iterates (3.8) and the Gauss-Seidel iterates (3.10) converge to $A^{-1}b$ for any $x^{(0)}$ (Varga [23]).

A modification of the Gauss-Seidel iteration is known as Successive Overrelaxation (SOR). In this iteration, the Gauss-Seidel iterate $\hat{x}_i^{(k+1)}$ is computed as before by

$$\hat{x}_i^{(k+1)} = (1/a_{ii})[b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}] \tag{3.11}$$

but the new value $x_i^{(k+1)}$ is taken to be

$$x_i^{(k+1)} = x_i^{(k)} + \omega[\hat{x}_i{}^{(k+1)} - x_i^{(k)}] \tag{3.12}$$

for some parameter ω. If $\omega = 1$, then $x_i^{(k+1)}$ is just the Gauss-Seidel iterate (3.9). The procedure (3.11-3.12) may be written as

$$x^{(k+1)} = H_\omega x^{(k)} + \omega(D + \omega L)^{-1}b, \quad k = 0, 1 \ldots \tag{3.13}$$

where

$$H_\omega = (D + \omega L)^{-1}[(1-\omega)D - \omega U]. \tag{3.14}$$

It can be shown that the iterates (3.13) converge for all $x^{(0)}$ if $\rho(H_\omega) < 1$, and hence $0 < \omega < 2$, ($\omega = real$). In fact, the SOR iterates (3.12) converge to $A^{-1}b$ for any $x^{(0)}$ and $b \in \Re^n$, if $A \in \Re^{(n\times n)}$is symmetric positive definite and $0 \leq \omega \leq 2$. (Ostrowski- Reich Theorem [23]).

The following result on the convergence of Jacobi, Gauss-Seidel and the SOR iterates is particularly useful:

Theorem 2 *If the matrix A is an irreducibly diagonally dominant M matrix and is also symmetric positive definite, then the Jacobi and Gauss-Seidel iterates converge to the unique solution of the system (3.2) and also, the SOR iterates converge for any* $\omega \in (0, 2)$.

The Symmetric Successive Overrealaxation method (SSOR), (Sheldon[21]), is the same as the SOR method except that one computes $x_i^{(k+1/2)}$ based on $x_i^{(k)}$ using a forward sweep and then computes $x_i^{(k+1)}$ based on $x_i^{(k+1/2)}$ using a backward sweep. The forward sweep is the same as the SOR–method; thus the values $x_i^{(k+1/2)}$ are the same as one would get for $x_i^{(k+1)}$ using the SOR–method. One then applies the SOR–method using the equations in the reverse order to complete the iterative step.

3.2 Solving Systems of Nonlinear Equations Using I.A.

We now turn to the main topic of the thesis, the solution of non-linear equations using interval analytic techniques.

Mathematical Developments

Let a vector $f = (f_1, ..., f_n)$ of n real, nonlinear functions of a real vector $x = (x_1, ..., x_n)^T$ be given. We consider Newton-like methods for finding solutions to $f(x) = 0$. Moore [13] first introduced an interval analytic method for finding and bounding a solution y which we now develop here. Let the equation to be solved be

$$f(x) = 0. \tag{3.15}$$

Using the Mean Value Theorem and expanding $f(x)$ about y, we obtain

$$f(x) = f(y) + J(\xi)(x - y) \tag{3.16}$$

where $x \leq \xi \leq y$, and where $J(\xi)$ is the Jacobian evaluated at a point ξ.
Thus if y is a solution of f, equation (3.16) yields

$$J(\xi)(x - y) = f(x). \tag{3.17}$$

If ξ was known then solving (3.17) would yield the solution y for a given x. Since ξ is not known approximations have to be used. Replacing ξ by x and iterating yield (3.3) or (3.4).

Moore [13] instead observed that if X is an interval vector containing x and y, then $\xi \in X$. Hence the set N of points z satisfying

$$J(X)(x - z) = f(x) \tag{3.18}$$

where $J(X) \supseteq \{J(x)|x \in X\}$ contains y for any x. For simplicity x is taken as the mid (X).

From (3.18), we obtain

$$N(X) = \text{mid}(X) - (J(X))^{-1} \cdot f(\text{mid}((X)). \tag{3.19}$$

where $N(X)$ is called the interval Newton function. $(J(X))^{-1}$ is a kind of inverse of $J(X)$ in that we require that for all $A \in J(X)$, $A \in R^{n\times n}$, $A^{-1} \in (J(X))^{-1}$. Choosing $X^{(0)}$, we define the sequence of intervals $X^{(1)}, X^{(2)}, \ldots$ by

$$X^{(n+1)} = N(X^{(n)}) \cap X^{(n)} \tag{3.20}$$

$X^{(0)}$ is given.

Equations (3.19)-(3.20) represent an interval version of Newton's method. A necessary condition for $N(X)$ to be defined by (3.19) is that X must contain at most one zero and such a zero be a simple root i.e. not simultaneously a zero of f and $J(x), x \in X^{(0)}$. An interesting result is that if a simple root $x \in X$ then $x \in N(X)$, and if $N(X) \bigcap X$ is empty then X does not contain a zero of f (see [13]). These simple properties form the basis of most interval Newton-like methods.

Hansen [4] pointed out that it is not necessary to find the exact inverse $(J(X))^{-1}$, as required by Moore's procedure. He multiplied equation (3.18) by an approximate inverse B of mid $J(X)$ and applied Gaussian elimination procedure to

$$M(X)(x - z) = b \tag{3.21}$$

where $M(X) = BJ(X), b = Bf(x)$ and $x = \text{mid } (X)$.

Krawczyk [8] introduced a variation of the interval Newton method which avoided Gaussian elimination of an interval matrix by not attempting a solution of (3.21). He in fact computes the box

$$K(X) = x - Bf(x) + [I - BJ(X)](X - x). \tag{3.22}$$

This box contains every solution of equation (3.18). The equation (3.22) solves equation (3.21) approximately for a bound $K_i(X)$ on the i^{th} component of the solution set Z.

Hansen and Sengupta [6] suggest the iteration

$$X^{(i+1)} = X^{(i)} \cap K(X^{(i)}), (i = 0, 1, 2...) \tag{3.23}$$

with (3.22) to ensure successive convergence.

Simultaneous iterations for components K_i computed from

$$K_i = x_i - b_i + \sum_{j=1}^{i-1} L_{ij}(K_j' - x_j) + \sum_{j=1}^{n} L_{ij}(X_j - x_j) \tag{3.24}$$

where $b = Bf(x), L = I - BJ(X)$ and $K_j' = K_j \cap X_j$ serve as a major improvement to Krawczyk's original procedure. Inspite of the modification, Hansen-Sengupta [6] noted that even though box determined by (3.24) bounds the solution set of equation (3.21), it is not the smallest such box. Thus Hansen-Sengupta, following a different approach, obtained a box, which is generally smaller than $K(X)$. Noticing that the matrix $M(X)$ is almost a diagonally dominant matrix resembling the identity, they write $M(X)$ as

$M = D + L + U$

where D, L and U are diagonal, lower triangular and upper triangular matrices respectively. The approximate solution X' of (3.21) may now be obtained as

$$Y = x - D^{-1}[b + L(X' - x) + U(X' - x)], \tag{3.25}$$

$$X' = Y \cap X. \tag{3.26}$$

When a D_{ii} contains zero, Hansen-Sengupta [6] make use of the extended interval arithmetic introduced by Alefeld [1] and Kahan [9] as well. The intersections in the system (3.26) produce a finite result.

As mentioned earlier, Hansen and Greenberg [5] have made a noteworthy attempt in producing a simple but most efficient algorithm that combines the best features of known algorithms. They have utilized the preconditioning, a relaxation procedure, a real (noninterval) iteration to speed up the elimination procedure. The extended interval mathematics is employed as before to obtain a finite solution. The order in which these operations must be performed is important for achieving efficient results. The main steps are summarized in the following mathematical procedures:

A. **Preconditioning**

$$M(X)(x - z) = b$$

where $M = BJ(X), b = Bf(x), B = \text{mid}(J(X))^{-1}$.

M is an almost diagonally dominant matrix close to unit matrix I .

B. **Hansen-Sengupta Step**

$$Y_i = x_i - (1/M_{ii})[b_i + \sum_{j=1, j\neq i}^{n} M_{ij}(X_j - x_j)].$$

C. **The Real Iteration** $x^{(k+1)} = x^{(k)} - Bf(x^{(k)}), k = 0, 1, 2...$

where $x^{(k)} = \text{mid } (X^{(k)})$.

Iteration terminates when

$$||f(x^{(r)})|| < 2 \cdot ||f(x^{(r+1)})||$$

or

$$||f(x^{(r)})|| < 10^{-3}.$$

D. **The Elimination Iteration**

$$LU(x - Y) = b, \ \ (M = LU)$$

$\mathrm{Z}_i = x_i - [G_i - \sum_{j=i+1}^{n} U_{ij}(x_j - X_j)]/U_{ii}$

$\mathrm{X}_i := X_i \bigcap Z_i$

where for U_{ij} and G_i, we have

$Lg = b$,

and $g \in G = U(x - Y)$ which is the box containing the required solution.

Chapter 4

The Hansen / Sengupta / Greenberg Analysis

4.1 The Hansen-Sengupta Method

In this Chapter, we examine the application of interval mathematics to bound a solution of a system of nonlinear equations arising from a given vector equation $f(x) = 0$, where

$$f(x) = (f_1, ..., f_n)^T.$$

Each component $f_i(x), (i = 1, 2..., n)$ is a nonlinear function of a real vector $x = (x_1, ..., x_n)^T$.

Moore[13], the 'father' of interval mathematics is one of the early researchers to have discussed this problem using interval analysis. As pointed out in Section 3.2 he is credited for implementing the interval mathematics into the multidimensional Newton's method, thus establishing a new framework for further research. Hansen's implementation is more sophisticated and hence it is described here.

We consider the equation:

$$f(x) = 0 \qquad or \qquad f_i(x) = 0, (i = 1, 2..n). \tag{4.1}$$

Using the mean value theorem about the point y, we obtain

$$f(x) = f(y) + \sum_{i=1}^{n} (x_i - y_i)(\partial/\partial x_i) f[y + \theta_i(x - y)] \tag{4.2}$$

where $\theta_i \in [0,1]$. If y is a solution vector for equation (4.1), then $f(y) = 0$ and equation (4.2) reduces to

$$f(x) = J(\xi)(x - y). \tag{4.3}$$

Here $J(\xi)$ is the Jacobian matrix with the elements

$$J_{ij}(\xi) = (\partial/\partial x_i) f_j[y + \theta_i(x - y)]$$

where $\xi = y + \theta_i(x - y), \theta_i \in [0,1]$.

Let $x, y \in X$. Suppose $J(X)$ is the natural interval extension of the matrix $J(\xi)$ and X is the interval containing y. We may write (4.3) as

$$J(X)(x - z) = f(x). \tag{4.4}$$

The solution of equation (4.4) is defined as the set

$$Y(x, X) = \{z | A(x - z) = f(x), A \in J(X)\}. \tag{4.5}$$

The set Y of points z satisfying (4.5) contains y.

Starting with an initial box $X^{(0)}$, new iterates for the solution are obtained by the procedure

$$X^{(n+1)} = Y^{(n+1)}(x^{(n)}, X^{(n)}) \cap X^{(n)}. \tag{4.6}$$

To achieve good results, the intersection is done for a component of $Y^{(n+1)}$ as soon as it is obtained.

4.1.1 Hansen's Preconditioning Technique

We now return to the solution of (4.4). Hansen [4] was perhaps first to point out that equation (4.4) is solved best by multiplying both sides of it with an approximate inverse of mid $(J(X))$. This results into a **preconditioning** of equation (4.4) most likely giving rise to a regular system. Let B denote this approximate inverse. Multiplying (4.4) with B yields

$$M(X)(x - z) = b \tag{4.7}$$

where $M(X) = BJ(X)$ and $b = Bf(x)$. The system (4.7) now contains an almost diagonally dominant matrix $M(X)$, providing the widths of the Jacobian entries are not too large. Such systems are amenable to Gauss-Seidel type iterations because of the diagonal dominance. Interval Newton methods differ in their approach for solving equation (4.7). They also differ in how the arguments of $J(X)$ are chosen. For simplicity the x in (4.7) is taken as the midpoint of X.

Moore [13] who introduced the first interval Newton method solved equation (4.7) by finding a *kind* of inverse of the interval matrix $J(X)$. This meant that an interval matrix $[M(X)]^{-1}$ of equation (4.7) is computed that contains the inverse of each noninterval matrix $D \in M(X)$. Thus each $J(X)$ had to be nonsingular. Some subsequent works, however, no longer make this stipulation. Application of extended interval arithmetic coupled with equation (4.6) give a bounded solution. This modification is described in the work of Hansen and Sengupta [6] and discussed here later.

4.1.2 Krawczyk's Successive Iteration

A variation of the interval Newton's method, which avoids the Gaussian elimination for a sharp solution of equation (4.7) is due to Krawczyk [8]. Historical development of Krawczyk's successive iteration presented by Hansen and Sengupta [6] was included in Section 3.3. Hansen Sengupta have used Krawczyk's successive iteration in conjunction with the Gaussian elimination described in next section. In this section, we review Krawczyk's technique in light of Hansen Sengupta's work [6]. Krawczyk introduced a box

$$K(X) = x - Bf(x) + [I - BJ(X)](X - x) \tag{4.8}$$

which contains every solution of (4.4). With this approach, one is solving the i^{th} equation of (4.7) for a bound $K_i(X)$ on the i^{th} component of the solution set $Y(X)$. Krawczyk [8] showed that if a solution y of (4.3) is contained in a box X, then it is also contained in box $K(X)$. But since $K(X)$ may not be contained in X, we use the iterative intersections:

$$X^{(n+1)} = X^{(n)} \cap K(X^{(n)}), n = 0, 1, 2... \tag{4.9}$$

where the initial box $X^{(0)}$ is known.

The convergence of the simultaneous iteration is improved, if it is used in a successive iteration mode. That is the component $K_i(X), (i = 1, 2, ...n)$ is computed as demonstrated by Wolfe [24] e.g.

$$K_i = x_i - b_i + \sum_{j=1}^{i-1} L_{ij}(K_j' - x_j) + \sum_{i=1}^{n} L_{ij}(X_j - x_j) \tag{4.10}$$

where

$$b = Bf(x), \;\; L = I - BJ(X), \;\; K_j' = K_j \cap X_j.$$

The intersection K_j' of K_j and X_j is found as soon as K_j is determined using the best currently available values. Each iteration of (4.10) determines a new Krawczyk box for bounding the solution of linearized system (4.7), however, it may not be the smallest such box as one would expect. To overcome this situation Hansen and Sengupta[6] have suggested a different approach for bounding the solution. The box containing the solution in their method is generally smaller than $K(X)$. Moreover each iteration in Hansen and Sengupta's method produces a greater reduction in the current box than does Krawczyk's method. Fewer steps are therefore required for the overall numerical convergence.

4.1.3 Gaussian Elimination Using Extended Interval Mathematics

Hansen and Sengupta start with the equation (4.7) as proposed by Hansen [4] i.e.

$$M(X)(x - z) = b \tag{4.11}$$

where $M(X) = BJ(X)$ and $b = Bf(x)$. The matrix $M(X)$ is believed to approximate closely the identity matrix and therefore normally exhibits diagonal dominance. The interval matrix $M(X)$ is now decomposed as the sum

$$M(X) = D(X) + L(X) + U(X)$$

where the matrices D, L, and U are the diagonal, lower and upper triangular matrices, respectively. We now consider the i^{th} equation in system (4.11) and solve it for the i^{th} variable, while replace others by their bounding intervals. Successive iteration is again employed as in Krawczyk's method.

The approximate solution X' is obtained from

$$Y = x - D^{-1}[b + L(X' - x) + U(X' - x)] \tag{4.12}$$

and $X' = Y \cap X$.

Each component $Y_i, (i = 1, 2...)$ is immediately intersected with X_i and the most recent value $X_i' = Y_i \cap X_i$ is then used for obtaining $Y_{i+1}, ..., Y_n$. The component form of equations (4.12) are:

$$Y_i = x_i - D_{ii}^{-1}[b_i + \sum_{j=1}^{i-1} M_{ij}(X_j' - x_j) + \sum_{j=i+1}^{n} M_{ij}(X_j - x_j)] \tag{4.13}$$

and $X_i' = Y_i \cap X_i, (i = 1, 2...n)$ where D_{ii} are the diagonal elements of matrix D. Hansen and Sengupta are also able to deal with the case in which the D_{ii} might contain zero for one or more values of i. In such a case, the extended interval arithmetic is employed to compute Y_i. Although the reciprocal of an interval containing a zero is the union of two semi-infinite intervals, the intersection of such a set with bounded intervals, during successive iterations, produces at most two bounded intervals. The method can therefore handle zero slopes which ordinary (noninterval) methods cannot do. That is, one can find roots of any multiplicity in higher dimensions as well, using the extended interval arithmetic described earlier in Section(2.1.1). The iter-

ation defined by (4.13) can be written in the form:

$$Y^{(k)} = x^{(k)} - (D^{(k)})^{-1}[b^{(k)} + L^{(k)}(X^{(k+1)} - x^{(k)}) + U^{(k)}(X^{(k)} - x^{(k)})] \qquad (4.14)$$

$X^{(k+1)} = Y^{(k)} \cap X^{(k)}, (k = 0, 1, 2...)$.

The following theorem due to Hansen and Sengupta [6] ensures the convergence of algorithm (4.14) under appropriate conditions.

Theorem 3 *Let* $\alpha_k = max\, w(X_i^{(k)})$, $\beta_k = max\, |M_{ii}^{(k)} - 1|$, *and* $\gamma_k = max \sum_{j=1, j\neq i}^{n} |M_{ij}^{(k)}|$ *each maximum is for* $i = 1, 2, ..n$ *and* $w(X_i^{(k)})$ *denotes the width of* $X_i^{(k)}$.

If f *has a single zero* x^* *in* $X^{(0)}$ *and if for some* $k = 0, 1, 2..$, *the conditions* $\beta_k \leq 2/3^{1/2} - 1$ *and* $\gamma \leq (1 - \beta)/2$ *hold, then* $X^{(k)} \rightarrow x^*$.

For sufficiently small $w(X^{(k)})$, the β_k and γ_k are monotonically decreasing with k and thus the hypothesis of the theorem are satisfied for any value of k. And $X^{(k+1)} \subset X^{(k)}$, while $X^{(k)} \rightarrow X^*$.

4.2 The Hansen And Greenberg's Method

In the following section, we will discuss and analyse the interval Newton method for bounding solutions of systems of nonlinear equations, as introduced by Hansen and Greenberg [5]. Basically, the method involves three subalgorithms. The first is a Gauss-Seidel type step as in Hansen and Sengupta's [6] work presented earlier. Secondly, a real (noninterval) Newton iteration is used to obtain an improved approximation for a zero, if one exists. Finally the solution of linearized equations is

attempted by performing elimination using the regular LU decomposition of the M matrix. Hansen and Greenberg have in fact combined the three known subalgorithms used in earlier works into a single algorithm of greater efficiency. In subsequent work of the thesis, we will show how some additional steps into the algorithm, can result in further improvement of the algorithm for a certain class of problems.

As shown earlier, we can expand the function $f(x)$ about y using the Mean Value theorem and if y is a zero of f, we obtain

$$J(\xi)(x-y) = f(x). \tag{4.15}$$

If X is an interval vector containing x and y then the point $\xi \in X$. Replacing $J(\xi)$ by the matrix $J(X)$, we find that y is contained in the set Z of points z satisfying

$$J(X)(x-z) = f(x). \tag{4.16}$$

As noted earlier, the interval Newton methods find an interval vector $Y(x,X)$ containing Z, which would satisfy (4.16) i.e.

$$Y(x,X) = \{z | A(x-z) = f(x),\ A \in X\}. \tag{4.17}$$

Starting with an initial bound $X^{(0)}$, a new iterate is obtained as

$$X^{(k+1)} = Y^{(k+1)}(x^{(k)}, X^{(k)}) \cap X^{(k)}. \tag{4.18}$$

For improved efficiency, the intersection (4.18) is done for each component of

$Y(x^{(k)}, X^{(k)})$ as soon as it is obtained.

Following the earlier works [4], Hansen and Greenberg multiply (4.16) with the approximate inverse of mid$J(X)$. The equation (4.16) then reduces to

$$M(X)(x - z) = b \tag{4.19}$$

where $M = BJ(X), b = Bf(x)$ and $B =$[mid$J(X)]^{-1}$.

It would be natural to attempt the solution of (4.19) in one step with the exact interval arithmetic using Gaussian elimination. However, this method cannot always be employed since it is possible that M could contain a singular matrix or that the interval growth results in a division by zero. The interval growth during the elimination process also produces poor results. If the interval growth did not occur, then the Gaussian elimination would be preferable to the earlier works of Krawczyk and Hansen-Sengupta, as the present technique offers Z and not just an inclusion for Z. Hansen and Greenberg, therefore apply Hansen and Sengupta's Gauss-Seidel step and use a real (noninterval) iteration to improve x, to evade growing intervals before elimination is attempted. Following Wolfe [24], Hansen and Greenberg also use inner iterations to reduce the number of times $J(X)$ and therefore $M(X)$ must be computed. The real iteration as described in the work below, in fact, serves as an inner iteration. Performed repetitively, this inner iteration improves the efficiency of the algorithm, since recalculations of $M(X)$ are reduced and the elimination yields a more accurate solution.

The method of Hansen and Greenberg can be summed up with the following procedures:

4.2.1 Initial Hansen-Sengupta Step

Performing the Hansen-Sengupta step on (4.19), we obtain

$$Y_i = x_i - (1/M_{ii})[b_i + \sum_{j=1}^{i-1} M_{ij}(X_j' - x_j) + \sum_{j=i+1}^{n} M_{ij}(X_j - x_j)]. \tag{4.20}$$

Since M was obtained by multiplying $J(X)$ with the approximate inverse B of mid$J(X)$, M should ideally approximate the identity matrix I. The diagonal elements M_{ii} may, however, contain a zero. If that is the case, then the extended interval arithmetic is used to solve the i^{th} equation of (4.19). Z_j have been replaced by X_j in (4.20) for all $j \neq i$. Thus the interval Y_i contains every solution z_i in X_i. Calculations are performed first for those i for which $0 \ni M_{ii}$ and each X_i is replaced by X_i' given by

$$X_i' = X_i \cap Y_i. \tag{4.21}$$

If the intersection (4.21) is empty, it means that there is *no* solution in the box X. When the intersection is not empty, the improved values of X_i are used to calculate the Y_i's for those i such that $0 \in M_{ii}$.

Two cases arise

1. If both $0 \in M_{ii}$ and $0 \in b_i + \sum_{j=1,j\neq i}^{n} M_{ij}(X_j - x_j)$, then $Y_i = (-\infty, \infty)$, which serves no useful purpose.

2. If $0 \in M_{ii}$ and $0 \ni b_i + \sum_{j=1,j\neq i}^{n} M_{ij}(X_j - x_j)$, then Y_i consists of two semiinfinite intervals which exclude a gap i.e. an open interval. The intersection $Y_i \cap X_i$ may be empty or consist of one or two intervals. The first two cases may be dealt with as we did when $0 \ni M_{ii}$. In the third case, where two new intervals arise we find the largest gap which would split the box , but save it until the other techniques have narrowed it down. This suggests that one should strive for a better value of X; an extended interval arithmetic calculations tend to provide a wider gap, when x is a *poor* approximation. Real iteration (noninterval) as illustrated below is indeed an attempt to obtain improved approximation. In subsequent analysis, the present thesis outlines steps that would further enhance the value of such an approximation.

4.2.2 The Real (Noninterval) Iteration

If $x^{(0)} = \text{mid}(X^{(0)})$ i.e. center of box $X^{(0)}$, then we compute $x^{(k+1)}$ using

$$x^{(k+1)} = x^{(k)} - Bf(x^{(k)}), (k = 0, 1, 2, ..., r). \tag{4.22}$$

If $x^{(k+1)}$ is not in $X^{(k+1)}$, we find the point on the boundary of $X^{(k+1)}$ which is on the line connecting $x^{(k)}$ to $x^{(k+1)}$ and use this as a replacement for $x^{(k+1)}$. This process may be stopped when

$$||f(x^{(r)})|| < 2.||f(x^{(r+1)})||$$

under some norm of f. This improved value of x is used for elimination or Hansen and Sengupta's iteration discussed earlier. The real iteration may also be stopped if $||f(x^{(r+1)})|| < 10^{-3}$. It should be noted at this point that a generalization of (4.22) discussed in Chapter (5) leads to a much improved value of x for a certain class of problems, offering a higher efficiency for the overall algorithm.

4.2.3 The Elimination Iteration

When an x has been found so that $||f(x)||$ is sufficiently small, the elimination method is applied to obtain an "exact" solution in a finite number of arithmetic operations assuming exact arithmetic. Under certain conditions, for example if $||f(x)||$ is very small, the interval arithmetic will also yield acceptable solution. These conditions would apply, especially if the matrix $M(X)$ is very near the identity matrix. The matrix $M(X)$ is strongly diagonally dominant and therefore pivoting is no longer required in the elimination procedure. However, if X is not small enough, we would return to the relaxation procedure described above, before resorting to this elimination iteration. This elimination procedure (a Gaussian elimination) adopted by Hansen and Greenberg is also described in detail by Ratschek and Rokne [20].
We obtain an LU decomposition of the interval matrix M so that

$$M \subseteq LU \tag{4.23}$$

where L and U are the interval lower triangular and upper triangular matrices. Equation (4.19) for the solution set Y containing points z can now be written as

$$LU(x - Y) = b. \tag{4.24}$$

Due to inclusion behaviour of matrix M in (4.24), it is clear that no zero of f is being lost through the use of (4.24).

It is possible that the decomposition of M into LU involves a division with an interval containing zero. Although the extended interval mathematics could be used for this case as well, further manipulation in this direction may weaken the approximation. It is best therefore to return to the relaxation procedures for a more favourable cycle of calculations.

. In order to solve equation (4.24), we first consider the simpler equation

$$Lg = b \tag{4.25}$$

where $g \in G \equiv U(x - Y)$ is the box containing the required solution Y.
The solution of interval equation (4.25) may also lead to a division by an interval containing zero, the process may be terminated and routed back to relaxation stage. Otherwise, we now proceed with the succeeding equation containing the solution i.e.

$$U(x - Y) = G. \tag{4.26}$$

Equation (4.26) is now solved for Y using the procedure

$$Z_i = x_i - [G_i - \sum_{j=i+1}^{n} U_{ij}(x_j - X_j)]/U_{ii}. \tag{4.27}$$

$$X_i := X_i \cap Z_i$$

where x_i, z_i, denote the components of x, and z, and Z is now the box containing the solution set Y.

4.2.4 Order of Steps in Hansen-Greenberg Algorithm

As noted in the preceding analysis, the solution to equation (4.16) involves a preconditioning, relaxation procedure, and local iteration, concluding with the elimination. Even though, we have resource to the extended interval mathematics for possible division by a zero, for convergence of results, we would prefer to terminate certain procedures with the hope that the increased relaxation and local iterations would perhaps produce more favourable results. The application of four main steps namely:

1. Preconditioning
2. Relaxation
3. Local Iteration and
4. Elimination

can proceed systematically in the following order:

1. Calculate $J(X)$ and the approximate inverse B of the mid $J(X)$.

2. Apply the preconditioning step to equation (4.16) and obtain
   $M(x - Y) = b$. Set $x = \text{mid } (X)$.

3. Apply the relaxation procedure i.e. the initial Hansen-Sengupta step to improve X and x by solving $M(x - Y) = b$. Perform first Hansen-Sengupta step for those i such that $0 \ni M_{ii}$. Then perform a Hansen-Sengupta step for those i such that $0 \in M_{ii}$ and save the largest gap for final splitting, as described in step (8).
   Thus

$$Y_i = x_i - (1/M_{ii})[\sum_{j=1}^{i-1} M_{ij}(X_j' - x_j) + \sum_{j=i+1}^{n} M_{ij}(X_j - x_j)]. \tag{4.28}$$
$$X_i' = X_i \cap Y_i.$$

   where $x = \text{mid}(X)$ and Y is the superset of the solution set of equation $M(x - Y) = b$.

4. Employ the local iteration i.e. the real (noninterval) procedure:
   $x^{(0)} = \text{Mid}(X^{(0)})$,
   $x^{(n+1)} = x^{(n)} - Bf(x^{(n)})$
   to improve x, until x is near the solution i.e. $f(x)$ is sufficiently small. If $f(x)$

is sufficiently small i.e. $||f(x)|| < 10^{-3}$ proceed to the elimination step (6).

5. If $||f(x)||$ is not sufficiently small, perform a Hansen-Sengupta step for those components such that $0 \ni M_{ii}$. If X has improved significantly set $x = \text{mid}\ (X)$ and complete the Hansen-Sengupta step saving the largest gap.

6. Apply the elimination procedure, i.e. LU decomposition of M for improving X. The following possibilities exist.

   a) The elimination process is leading to a significant improvement in the width of X which means that the ratio of old box width to the new box width is less than 0.9. In this case, the elimination process is applied again with new value of X.

   b) The elimination process is not yielding any significant improvement in X, so we terminate elimination and proceed to splitting of X as indicated in step (8).

   c) The elimination process encounters a division by an interval containing zero. We terminate the elimination and perform simplified relaxation (step 7).

7. In the simplified relaxation procedure a Y_i is computed from (4.28) only if $0 \ni M_{ii}$. If X improves significantly we set $x := \text{mid}X$ and repeat the simplified relaxation procedure. (The Y_i with $0 \in M_{ii}$ are not calculated again, they

were already used to find the gaps). If X does not improve significantly even with the simplified relaxation, we will split X as in step (8) below.

8. The splitting of a box is extended only as a last resort when further improvement by other steps is unlikely. There are two possibilities:

   a) If the largest gap in step (3) occurs in the direction i, the box is split in the i^{th} direction.

   b) If the gap was obtained from step (7) due to lack of further improvement in the width X, we may split X at the mid-point of its largest component.

   All the split boxes are now put on a stack where they wait for further processing. A new iteration is done with the largest box and a check made for the termination criteria in step (9). If the stack has no more boxes, that can be chosen as a new box then f has no solution in the original box $X^{(0)}$ and the algorithm is terminated.

9. If the termination criteria is not fulfilled, then we must start with a new iteration at step 1.

4.2.5 Numerical Results

Example 1

We now apply Hansen-Greenberg's method to the solution of a nonlinear system

considered by Neumaier [16] namely

$$f_1(x_1, x_2) = x_1^2 + 9x_1 + x_2 - 36 = 0,$$

$$f_2(x_1, x_2) = x_1 + x_2^2 + 10x_2 - 3 = 0.$$

The algorithm for the interval Newton method, using Hansen-Greenberg analysis has been implemented in Fortran by Ratschek and Rokne [20]. The illustration uses the following notations:

n_1 = the number of interval Jacobian evaluations

n_2 = the number of interval matrix-vector products

n_3 = the number of real value iterations

n_4 = the number of interval LU decomposition procedures

n_5 = the number of interval elimination steps

n_6 = the number of executions of the relaxation procedure for those i with $0 \ni M_{ii}$

n_7 = the number of executions of relaxation procedure for those i with $0 \in M_{ii}$

s = significant improvement factor (when the ratio of the new box width to the old box width is smaller than s, then the box is said to have improved significantly by the relaxation procedure).

The algorithm was executed for the above system with three different boxes. The significant width improvement factor $s = 0.9$; and the procedure is terminated when each box on the list has width less than 10^{-6}. The following results are obtained:

$$Box1 = \begin{pmatrix} [-4, & 4] \\ [-4, & 4] \end{pmatrix}.$$

Solution(s) after 1 iteration(s) are

$x_1 = 0.3000000119E + 01$, Tolerance $= 0.1937150955E - 06$

$x_2 = 0.8566780352E - 07$, Tolerance $= 0.4408870744E - 06$

$f(x_1) = 0.2384185819E - 05$

$f(x_2) = 0.9983778142E - 06$

$n_1 = 1, n_2 = 16, n_3 = 2, n_4 = 0, n_5 = 0, n_6 = 16, n_7 = 0$

$$Box2 = \begin{pmatrix} [2.5, & 3.5] \\ [-2, & 2] \end{pmatrix}.$$

Solution(s) after 1 iteration(s) are

$x_1 = 0.300000030E + 01$, Tolerance $= 0.1490116119E - 07$

$x_2 = 0.3749835685E - 09$, Tolerance$= 0.3831625468E - 09$

$f(x_1) = 0.9536743306E - 06$

$f(x_2) = 0.4470348403E - 07$

$n_1 = 1, n_2 = 2, n_3 = 0, n_4 = 1, n_5 = 1, n_6 = 1, n_7 = 0$

$$Box3 = \begin{pmatrix} [2.9, & 3.1] \\ [-.1, & .1] \end{pmatrix}.$$

Solution(s) after 1 iteration(s) are

$x_1 = 0.3000000030E + 01$, Tolerance $=0.1490116119E - 07$

$x_2 = 0.2010322412E - 09$, Tolerance $= 0.4731546665E - 10$

$f(x_1) = 0.9536743306E - 06$

$f(x_2) = 0.4470348403E - 07$

$n_1 = 1, n_2 = 2, n_3 = 0, n_4 = 1, n_5 = 1, n_6 = 1, n_7 = 0$

All the three boxes enclose a single zero $x^* = (3, 0)^T$ as confirmed by Neumaier [16] following a different approach. The above example shows the relative frequencies of execution of the different components of the algorithm. In particular, the test results show a large number of n_2 and n_6 steps for Box1 compared with other boxes due to a larger width of Box1.

Example 2

Let

$f_1(x) = x_1^2 + x_2^2 - 1,$

$f_2(x) = x_1 - x_2.$

Suppose

$$X^{(0)} = \begin{pmatrix} [0.5, & 1.0] \\ [0.5, & 1.0] \end{pmatrix}.$$

Hansen-Greenberg's algorithm yields the following solution

$x_1 = 0.7071067914E + 00$, Tolerance $= 0.3352761269E - 07$

$x_2 = 0.7071067914E + 00$, Tolerance $= 0.3352761269E - 07$

with

$n_1 = 1$, $n_2 = 4$, $n_3 = 1$, $n_4 = 1$, $n_5 = 1$, $n_6 = 3$, $n_7 = 1$.

As before the significant factor is 0.9 and the tolerance limit is $1.0E - 06$.

Example 3

Let

$f_1(x) = x_1^2 - 2x_2 + 1,$

$f_2(x) = x_1 + 2x_2^2 - 3.$

Suppose

$$X^{(0)} = \begin{pmatrix} [0, & 2] \\ [0, & 2] \end{pmatrix}.$$

Solution(s) from Hansen-Greenberg's algorithm are:

$x_1 = 0.9999999627E + 00$, Tolerance= $0.4097819328E - 07$

$x_2 = 0.1000000030E + 01$, Tolerance= $0.1303851604E - 06$.

And

$x_1 = 0.1000000045E + 01$, Tolerance= $0.1490116119E - 07$

$x_2 = 0.1000000000E + 01$, Tolerance= $0.1490116119E - 07$.

With

$n_6 = 3$, $n_7 = 14$.

However, choosing

$$X^{(0)} = \begin{pmatrix} [0.9, & 1.1] \\ [0.9, & 1.1] \end{pmatrix}.$$

We obtain the following solution:

$x_1 = 0.1000000000E + 01$, Tolerance=$0.111758709E - 07$

$x_2 = 0.1000000000E + 01$, Tolerance=$0.111758709E - 07$.

With

$n_1 = 1$, $n_2 = 2$, $n_3 = 0$, $n_4 = 1$, $n_5 = 1$, $n_6 = 1$, $n_7 = 0$.

The smaller width for the initial interval results in fewer iterations overall.

Chapter 5

Some Procedural Improvements

5.1 Improvement of Initial Interval Containing a Fixpoint

Hansen and Greenberg [5] preconditioned the system

$$J(X)(x - Y) = f(x) \tag{5.1}$$

by multiplying it with an approximate inverse of mid$J(X)$, thereby obtaining the equation:

$$M(X)(x - Y) = b \tag{5.2}$$

where $M(X) = BJ(X)$, $b = Bf(x)$, and $B = [\text{mid}J(X)]^{-1}$.

As expected, the matrix $M(X)$ would then become almost diagonally dominant matrix containing the unit matrix if the widths of the elements of $J(X)$ were not too large. An approximate solution of equation (5.2) can be obtained by applying the noninterval quasi-Newton iteration:

$$x^{(n+1)} = x^{(n)} - Bf(x^{(n)}) \tag{5.3}$$

where $x^{(0)} = \text{mid}(X^{(0)})$.

The system (5.2) may be now amenable to the Gauss-Seidel iterations due to the diagonal dominance of $M(X)$. The detailed procedure described by Hansen and Greenberg examined in Chapter 4, nevertheless lead to a more acceptable solution.

In the following section, we describe an extension of (5.2), in which the matrix $M(X)$ is approximated by the unit matrix more closely. This additional condi-

tioning of the system (5.1) will result in a more exact analysis for the relaxation and the elimination procedures.

We use the following properties for the monotonic inclusion of inverse matrices A, setting $A = I$. See Alefeld and Herzberger [1].

Let $L^{(0)}$ and $U^{(0)}$ be two $n \times n$ matrices for which $L^{(0)} \geq 0$ and $L^{(0)} \leq I \leq U^{(0)}$. When the sequences $\{L^{(k)}\}$ and $\{U^{(k)}\}$, $k = 0, 1, 2...\infty$ are calculated according to

$$L^{(k+1)} = L^{(k)} + (I - L^{(k)}) \sum_{\nu=0}^{r-2} (I - L^{(k)})^{\nu} L^{(k)} \tag{5.4}$$

$$U^{(k+1)} = U^{(k)} + (I - U^{(k)}) \sum_{\nu=0}^{r-2} (I - U^{(k)})^{\nu} L^{(k)} \tag{5.5}$$

then the following statements are valid:

a) $0 \leq L^{(0)} \leq ..L^{(k)} \leq L^{(k+1)}.. \leq I.. \leq U^{(k+1)} \leq U^{(k)} \leq .. \leq U^{(0)}$.

b) Both sequences $\{L^{(k)}\}$ and $\{U^{(k)}\}$ converge to I iff the spectral radius $\rho(I - L^{(0)}) < 1$.

c) If the procedure (5.4-5.5) converges then the quantities

$d^{(k)} = ||U^{(k)} - L^{(k)}||$ satisfy the relation

$d^{(k+1)} \leq \gamma (d^{(k)})^r$, $\gamma \geq 0$, $r > 0$.

For briefness, we consider only values of $L^{(k+1)}$ and $U^{(k+1)}$ for $\nu = 0$. Equations (5.4-5.5) are then approximated with

$$L^{(k+1)} \approx L^{(k)} + (I - L^{(k)}) L^{(k)}, \tag{5.6}$$

$$U^{(k+1)} \approx U^{(k)} + (I - U^{(k)}) U^{(k)}. \tag{5.7}$$

or

$$L^{(k+1)} \approx (2I - L^{(k)}) L^{(k)}, \tag{5.8}$$

$$U^{(k+1)} \approx (2I - U^{(k)})U^{(k)}. \tag{5.9}$$

We write equation (5.2) as

$$([L^{(0)}, U^{(0)}])(x - Y) = b. \tag{5.10}$$

We may further express equation (5.10) into two separate equations as follows:

$$L^{(0)}(x - Y)_l = b_l, \tag{5.11}$$

and

$$U^{(0)}(x - Y)_u = b_u. \tag{5.12}$$

Multiplying both sides of equations (5.11) and (5.12) with $(2I - L^{(0)})$ and $(2I - U^{(0)})$ respectively, and using (5.8-5.9), we obtain

$$L^{(1)}(x - Y)_l = (2I - L^{(0)})b_l, \tag{5.13}$$

$$U^{(1)}(x - Y)_u = (2I - U^{(0)})b_u. \tag{5.14}$$

Equations (5.13-5.14) can be written back into a single equation

$$([L^{(1)}, U^{(1)}])(x - Y)_{l,u} = [(2I - L^{(0)})b_l, (2I - U^{(0)})b_u] \tag{5.15}$$

or

$$M^{(1)}(x - Y)_{l,u} = [(2I - L^{(0)})b_l, (2I - U^{(0)})b_u] \tag{5.16}$$

where $M^{(1)} = ([L^{(1)}, U^{(1)}])$ is a better approximating matrix for I than $M^{(0)} = ([L^{(0)}, U^{(0)}])$. Accordingly system (5.16) will yield a better approximation for Y than

the equations (5.2). The solution to system (5.16) may be obtained by solving the noninterval equations

$$Y_l^{(1)} = x_l^{(0)} - (2I - L^{(0)})b_l, \tag{5.17}$$

$$Y_u^{(1)} = x_u^{(0)} - (2I - U^{(0)})b_u. \tag{5.18}$$

We illustrate our modified procedure with the help of the following examples:

Example 1

Let $f(f_1, f_2) = 0$,

where

$$f_1(x_1, x_2) = x_1^2 - 1, \tag{5.19}$$

$$f_2(x_1, x_2) = x_2^3 - 1. \tag{5.20}$$

A fixpoint is clearly $x^* = (x_1, x_2) = (1, 1)$.

Suppose

$$X^{(0)} = \begin{pmatrix} [0.98, & 1.18] \\ [0.98, & 1.18] \end{pmatrix}.$$

Then

$$\hat{X}^{(0)} = \begin{pmatrix} 1.08 \\ 1.08 \end{pmatrix},$$

$$J(X^{(0)}) = \begin{pmatrix} 2X_1 & 0 \\ 0 & 3X_2^2 \end{pmatrix}.$$

$$J(X^{(0)}) = \begin{pmatrix} [1.96, 2.36] & 0 \\ 0 & [2.8812, 4.1772] \end{pmatrix},$$

$$\text{mid}J(X^{(0)}) = \begin{pmatrix} 2.16 & 0 \\ 0 & 3.529 \end{pmatrix},$$

$$[\text{mid}J(X^{(0)})]^{-1} = \begin{pmatrix} 0.4629 & 0 \\ 0 & 0.2833 \end{pmatrix}.$$

Therefore

$$M^{(0)} = \begin{pmatrix} [0.9072, 1.0924] & 0 \\ 0 & [0.8162, 1.1834] \end{pmatrix}.$$

Notice that

$$M^{(0)} \supset \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

$M^{(0)}$ is thus a **diagonally dominant** matrix containing the unit matrix as expected.

Using $M^{(0)}(x^{(0)} - Y) = b$, we obtain

$$Y^{(1)} = \begin{pmatrix} [0.9029, & 1.10298] \\ [0.9064, & 1.10643] \end{pmatrix}.$$

Therefore

$$X^{(1)} = Y^{(1)} \cap X^{(0)} = \begin{pmatrix} [0.98, & 1.10298] \\ [0.98, & 1.10640] \end{pmatrix}.$$

Thus

$$\hat{X}^{(1)} = \begin{pmatrix} 1.04149 \\ 1.043215 \end{pmatrix}.$$

$\hat{X}^{(1)}$ is a better approximation for the fixpoint than the $\hat{X}^{(0)}$. In fact

$f_1(x_1^{(0)}, x_2^{(0)}) = 0.1664,$

$f_2(x_1^{(0)}, x_2^{(0)}) = 0.2597,$

while

$f_1(x_1^{(1)}, x_2^{(1)}) = 0.084700,$

$f_2(x_1^{(1)}, x_2^{(1)}) = 0.135328.$

We now use our procedure (5.8,5.9) to obtain a *better* approximating matrix $M^{(1)}$, for unit matrix I, than the matrix $M^{(0)}$ as used in above computations following Hansen-Greenberg [5].

Using the relations

$L^{(1)} \approx (2I - L^{(0)})L^{(0)}$

$U^{(1)} \approx (2I - U^{(0)})U^{(0)}$

we find that $M^{(1)} = ([L^{(1)}, U^{(1)}])$

$$= \begin{pmatrix} [0.99138, 0.99146] & 0 \\ 0 & [0.96621, 0.96636] \end{pmatrix}.$$

$M^{(1)} \subset M^{(0)}$. Also comparing $M^{(1)}$ with $M^{(0)}$, we see that $M^{(1)}$ is a better approximation for the unit matrix than $M^{(0)}$.

We now compute the approximate values of $Y_l^{(1)}$ and $Y_u^{(1)}$ using relations (5.17-5.18). In noninterval arithmetic

$$(2I - L^{(0)})b_l = \begin{pmatrix} 0.084167 \\ 0.087092 \end{pmatrix}$$

and

$$(2I - U^{(0)})b_u = \begin{pmatrix} 0.06990 \\ 0.06007 \end{pmatrix}.$$

Hence from (5.17) and (5.18), we obtain

$$X^{(1)} = Y^{(1)} = ([Y_l^{(1)}, Y_u^{(1)}]) = \begin{pmatrix} [0.895833, & 1.1101] \\ [0.892908, & 1.11992] \end{pmatrix}$$

and

$$\hat{X}^{(1)} = \begin{pmatrix} 1.002966 \\ 1.006414 \end{pmatrix}.$$

This gives

$f_1(\hat{X}_1, \hat{X}_2) = 0.005940,$

$f_2(\hat{X}_1, \hat{X}_2) = 0.019365.$

Clearly, the present value of $X^{(1)}$ is a better starting value than the $X^{(1)}$ computed in the Hansen-Greenberg algorithm. This approach will lead to more accurate results with fewer applications of the Hansen-Sengupta steps and Hansen-Greenberg's real iterations in the subsequent analysis.

Example 2

Let

$$f(x) = \begin{pmatrix} x_1^2 - 1 \\ x_2^2 - 0.01 \end{pmatrix}.$$

We take

$$X^{(0)} = \begin{pmatrix} [0.98, & 1.18] \\ [0, & 0.2] \end{pmatrix}$$

so that

$$\hat{X}^{(0)} = \begin{pmatrix} 1.08 \\ 0.1 \end{pmatrix}.$$

Also

$$J(X) = \begin{pmatrix} 2X_1 & 0 \\ 0 & 2X_2 \end{pmatrix}.$$

Hence $M^{(0)}$ =mid$[J(X^{(0)})]^{-1} \cdot J(X^{(0)}) =$

$$\begin{pmatrix} [0.9072, 1.0924] & 0 \\ 0 & [0, 2.0] \end{pmatrix} \supset \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Using (5.17-5.18), we obtain

$$X^{(1)} = (Y_l^{(1)}, Y_u^{(1)}) = \begin{pmatrix} [0.89586, & 1.1102] \\ [0, 0.2] & \end{pmatrix}.$$

Hence

$$\hat{X^{(1)}} = \begin{pmatrix} 0.994583 \\ 0.1 \end{pmatrix}.$$

The exact value for fixpoint is

$$X^* = \begin{pmatrix} 1.0 \\ 0.1 \end{pmatrix}.$$

Also

$f_1(x_1, x_2) = -.01080,$

$f_2(x_1, x_2) = 0.$

5.2 Modified Gauss-Seidel Iteration

Hansen and Greenberg [5] have performed an initial Hansen-Sengupta [6] step to the preconditioned equation:

$$M(X)(x - Y) = b. \tag{5.21}$$

The solution

$$Y_i = x_i - (M_{ii})^{-1}[b_i + \sum_{j=1}^{i-1} M_{ij}(X_j^{'} - x_j) + \sum_{j=i+1}^{n} M_{ij}(X_j - x_j)] \tag{5.22}$$

$$X_i^{'} = X_i \cap Y_i, (i = 1, 2, ..., n) \tag{5.23}$$

obtained in this manner is refined with the real iteration before the final elimination. The interval mathematics offers good results during elimination, if the approximating interval includes the solution precisely. Likewise the real iteration yields sufficiently small values of $||f(x)||$, if the Gauss-Seidel step results into a better approximating interval. In such a case, the Successive Overrelaxation (SOR) Gauss-Seidel procedure as outlined in Chapter 3 would be useful. This is described again as follows:

$$\hat{Y}_i^{(k+1)} = x_i^{(k)} - (1/M_{ii})[b_i + \sum_{j=1}^{i-1} M_{ij}(X_j^{(k+1)} - x_j^{(k)}) + \sum_{j=i+1}^{n} M_{ij}(X_j^{(k)} - x_j^{(k)})] \tag{5.24}$$

$$Y_j^{(k+1)} = [X_j^{(k)} + \omega(\hat{Y}_j^{(k+1)} - X_j^{(k)})] \cap X_j^{(k)} \tag{5.25}$$

and

$$X^{(k+1)} = X^{(k)} \cap Y^{(k+1)} \tag{5.26}$$

for some parameter ω. If $\omega = 1$, then $X^{(k+1)}$ is just the Gauss-Seidel iterate (5.22-5.23).

It can be established that the iterates (5.24-5.26) converge for all $X^{(0)}$ if $0 < \omega < 2$, (ω = real) [19]. This may constitute a sufficient convergence requirement for an important class of matrices.

Proof

Using (5.25) with (5.24) and rearranging, we have

$$X^{(k+1)} = H_\omega X^{(k)} + \omega(D + \omega L)^{-1}(Mx^{(k)} - b) \tag{5.27}$$

where

$$H_\omega = (D + \omega L)^{-1}[D(1-\omega) - \omega U]. \tag{5.28}$$

Since L is strictly lower triangular, $\det D^{-1} = \det\ (D + \omega L)^{-1}$.

Therefore,

$$detH_\omega = det(D + \omega L)^{-1}.det\{(1-\omega)D - \omega U\}$$

$$= detD^{-1}.det\{(1-\omega)D - \omega U\}$$

$$= det\{(1-\omega)I - \omega D^{-1}U\}$$

$$= det\{(1-\omega)I\}$$

$$= (1-\omega)^n,$$

since $D^{-1}U$ is strictly upper triangular. But det H_ω is the product of the eigenvalues of H_ω, hence

$$\rho(H_\omega) \geq |1-\omega|.$$

Kahan's Theorem [10].

The iterates (5.24) or (5.25) will converge for all values of $X^{(0)}$, if $\rho(H_\omega) < 1$ i.e. $0 < \omega < 2$, ω is real.

We deduce an extension of Ostrowski-Reich theorem [18].

Theorem 4 *Suppose $M \in I\Re^{m\times n}$, M^{-1} exists and $0 < \omega < 2$, then the SOR iterates (5.24-5.26) converge to $x - M^{-1}b$, where lim $\hat{x}^k = x$, as $k \to \infty$.*

Proof

$$X^{(k+1)} = H_\omega X^{(k)} + \omega(D + \omega L)^{-1}(Mx^{(k)} - b)$$

$$X^{(k+1)} = (D + \omega L)^{-1}[D(1-\omega) - \omega U]X^{(k)} + \omega(D + \omega L)^{-1}(Mx^{(k)} - b)$$

$$X^*[\omega(L + D + U)] = \omega Mx - \omega b$$

$$MX^* = M\hat{x} - b$$

$$M(\hat{x} - X^*) = b$$

$$\hat{x} - X^* = M^{-1}b.$$

Thus $X^* = \hat{x} - M^{-1}b$, here lim $X^{(k)} = X^*$ and lim $x^{(k)} = \hat{x}$, as $k \to \infty$.

The modified Gauss-Seidel iteration given by (5.24-5.26) can be very helpful, since it enables a faster convergence of the iterates. It may be shown that the eigenvalues μ_i and λ_i for the Jacobian and the SOR Gauss-Seidel iteration are related by (5.29)

$$(\lambda_i + \omega - 1)^2 = \lambda_i \omega^2 \mu_i^2. \tag{5.29}$$

Setting $\omega = 1$, we find that

$$\rho(GaussSeidel) = \rho(Jacobian)^2.$$

Therefore, if $\rho(Jacobian) < 1$, so that the Jacobian iteration is convergent then the asymptotic convergence of the Gauss-Seidel iteration is the square of the convergence of the Jacobian iteration (**quadratic convergence**). Using the modified Gauss-Seidel iteration, we may choose the parameter ω in such a way that ρ(modified Gauss-Seidel) is minimized and therefore convergence is maximized. The SOR Gauss-Seidel iterates should be preferred, where *more* precise interval inclusions are necessary for performing real iterations and subsequent eliminations in the Hansen-Greenberg's algorithm.

5.3 Interval Iteration-Alternative to Real Iteration

Hansen and Greenberg [5] have used mean value theorem to reduce a system of nonlinear equations $f(x) = 0$ to

$$J(X)(x - z) = f(x). \tag{5.30}$$

Some earlier development of above equation may be found also in Chapter 3. Multiplying both sides of equation (5.30) with the approximate inverse of mid$J(X)$, we obtain

$$M(X)(x - z) = b \tag{5.31}$$

where $M(X) = BJ(X)$, $b = Bf(x)$ and $B = \text{mid}[J(X)]^{-1}$. Equation (5.31) may be used to obtain the approximate superset

$$Y = x - M^{-1}b \tag{5.32}$$

or

$$Y = \text{mid}(X) - M^{-1}b.$$

Since M^{-1} closely approximates the identity matrix I, we may therefore consider the following iteration

$$X^{(k+1)} = \{m(X^{(k)}) - V^{(k)}b\} \cap X^{(k)}, \tag{5.33}$$

$$V^{(k+1)} = \{m(V^{(k)}) + V^{(k)}(I - M^{(k)}m(V^{(k)})\} \cap V^{(k)}, \tag{5.34}$$

$k \geq 0$, where $V^{(0)}$ is an interval matrix with the property that $M^{-1} \in V^{(0)}$, $M(X) = BJ(X)$. The equations (5.33-5.34) above yield the sequences of interval matrices and vectors

$X^{(0)} \supseteq X^{(1)} \supseteq X^{(2)} \supseteq ...$

$V^{(0)} \supseteq V^{(1)} \supseteq V^{(2)} \supseteq ...$

The procedure (5.33-5.34) consists of a simultaneous execution of $X^{(k)}$ and $V^{(k)}$. In fact a new interval matrix $V^{(k+1)}$ is being used for upgrading $X^{(k+1)}$.

Let $X^{(0)}$ be an interval vector and let $y \in X^{(0)}$ be a zero of the function $f(x)$. Let the derivative of $f(x)$ satisfy a Lipschitz condition in $X^{(0)}$. Furthermore let $V^{(0)}$ be an interval matrix containing the matrices $[M(X)]^{-1}$ for each $x \in X^{(0)}$. For the interval evaluation $J(X)$ and therefore $M(X)$ of the Frèchet derivative $J(X)$ the condition

$$||d(J(X))|| \leq c||d(X)||$$

or

$$||d(M(X))|| \leq \gamma||d(X)||$$

holds for all $x \subset X^{(0)}$ (c, γ are some constants).

The interval vectors $\{x^{(k)}\}, 0 \leq k \leq \infty$ and the interval matrices $\{V^{(k)}\}, 0 \leq k \leq \infty$ calculated according to (5.33-5.34), then satisfy the following:

1. Each interval vector $X^{(k)}$, $k \geq 0$, contains the zero y.

2. If all the matrices $V \in V^{(0)}$ are nonsingular, then it follows that

   Limit$X^{(k)} = y$ and limit $V^{(k)} = M^{-1}(y)$, $k \to \infty$.

3. Sequence $\{(X^{(k)}, V^{(k)})\}$, $0 \leq k \leq \infty$, converge at least quadratically toward $\{(y, M^{-1}(y)\}$.

The interval iteration as described above should provide a suitable alternative to the real iteration in the Hansen-Greenberg's algorithm.

Chapter 6

Other Suggested Procedures

6.1 Newton's Interval Method With Secondary Iterations

Newton's method for approximating roots of one nonlinear equation in one variable by the iteration

$$x^{(n+1)} = x^{(n)} - \frac{f(x^{(n)})}{f'(x^{(n)})} \tag{6.1}$$

is well known and may be found in several texts. The extension of this method to interval analysis was credited to the work of Moore [13] in previous Chapters. For $X \subset [x_1, x_2]$ such that $f(x_1).f(x_2) < 0$ and $m(X) =$ the midpoint of X, Moore defined the interval Newton function $N(X)$ by

$$N(X) = m(X) - \frac{f(m(X))}{F'(X)}, \tag{6.2}$$

where $F'(X)$ is the interval extension of $f'(x)$.

Equation (6.2) provides an **interval version** of Newton's method, by choosing $X^{(0)}$ and defining the sequence of intervals $X^{(1)}, X^{(2)}, \ldots$ with

$$X^{(n+1)} = N(X^{(n)}) \cap X^{(n)}. \tag{6.3}$$

This procedure has been demonstrated by Moore with the following example

Let $f(x) = x^2 - 2 = 0$, starting with $X^{(0)} = [1, 2]$ we obtain

$$N(X^{(0)}) = m(X^{(0)}) - \frac{[m(X^{(0)})]^2 - 2}{2X^{(0)}}$$

And

$$X^{(1)} = N(X^{(0)}) \cap X^{(0)} = [1.375, 1.769]$$

In same manner, we obtain

$X^{(2)} = [1.41406..., 1.41441..]$

$X^{(3)} = [1.414213559.., 1.41213566..]$.

We find in this case that, the interval version of Newton's method gives a rapidly contracting sequence of intervals containing the root $2^{1/2}$.

Moore [13] has in fact shown that when $F'(X)$ is defined and does not vanish on $[x_1, x_2]$, the interval version of Newton's method as illustrated above is asymptotically **error-squaring**, that is to say, there is an interval $X^{(0)} \subset [x_1, x_2]$, containing the zero y and a positive real number k such that

$$\omega(X^{(n+1)}) \leq k(\omega(X^{(n)}))^2, \tag{6.4}$$

where ω denotes the width of the interval.

The property (6.4) makes the Newton's interval method more appealing due to the quadratic convergence of intervals containing the root. The Newton's interval method, however, has some disadvantage especially with the higher dimensions; since the computations of $F'(x^{(k)})$, (i.e. $J(F(X))$ in higher dimensions) involving n^2 elements can be somewhat more laborous and time consuming. To overcome this difficulty, the non-interval Newton methods, (see [22]), have employed iterative techniques in which fewer computations of the derivative are required, while better approximations for the root are being obtained. In this Section, we have introduced a modification to the Newton's interval method so that the Newton's function

$N(X^{(n)})$ computed utilizing *fewer* computations of the derivative. Effectiveness of the modified version is also discussed in terms of the convergence. Results of Moore's illustration are compared with those of the modified version described below:

We replace the equations(6.2)-(6.3) by

$$X^{(0,0)} := X^{(0)} \tag{6.5}$$

$$N^{(k,m)}(X^{(k)}) := m\{N^{(k,m-1)}(X^{(k)})\} - \frac{F(m(N^{(k,m-1)}(X^{(k)}))}{F'(X^{(k)})}, m = 1, 2.., p; k = 0, 1, 2... \tag{6.6}$$

$$X^{(k,m)} := N^{(k,m-1)}(X^{(k)}) \cap X^{(k,m-1)} \tag{6.7}$$

$$N^{(k,0)} := N^{(k-1,p)} \tag{6.8}$$

$$X^{(k,0)} \equiv X^{(k)} := X^{(k-1,p)}, k \geq 0 \tag{6.9}$$

$$N^{(0,0)}(X^{(0)}) := m(X^{(0)}) - \frac{F(m(X^{(0)}))}{F'(X^{(0)})}. \tag{6.10}$$

Here $X^{(0)}$ is the interval vector containing the solution of $f(x) = 0$.

As before, we let $f(x) = x^2 - 2$ and $X^{(0)} = [1, 2]$.

Using relations (6.5-6.10), we obtain

$$N^{(0,0)}(X^{(0)}) = m(X^{(0)}) - \frac{F(m(X^{(0)}))}{F'(X^{(0)})} = [1.375, 1.769] = X^{(0,1)}$$

$$N^{(0,1)}(X^{(0)}) = m(N^{(0,0)}(X^{(0)})) - \frac{F(m(N^{(0,0)}(X^{(0)})))}{F'(X^{(0)})} = [1.4118652, 1.4174804].$$

We find similarly $N^{(0,2)}(X^{(0)}) = [1.4140233, 1.4143481]$.

Thus if $p = 2$

$X^{(0,2)} \equiv X^{(1)} = N^{(0,2)}(X^{(0)}) \cap X^{(0)} = [1.4140233, 1.4143481]$.

The value for $X^{(1)}$ obtained by the modified Newton's interval version (6.5-6.10) compares with the value for $X^{(3)} = [1.414213559, 1.414213566]$ obtained by Moore [13] using the regular iteration process (6.2-6.3). We observe that $N^{(k,m)}(X^{(k)})$ requires $p-1$ evaluations of F, but only one evaluation of F'. If $p = 1$, then the interval generalization of Newton's method (6.5-6.10) reduces to Newton's method as introduced by Moore [13].

We now discuss the convergence of our iterated results. We write the fundamental equation (6.6) with supressed k as

$$N^{(j)}(X) = m\{N^{(j-1)}(X)\} - \frac{F(m(N^{(j-1)}(X)))}{F'(X)}. \tag{6.11}$$

Without loss of generality, we can assume that $F'(X) > 0$ for all $X \subset [x_1, x_2]$. There is a positive real number λ such that for $y_1 \leq y \leq y_2$, we have

$$F'([y_1, y_2]) = F'(y) + \lambda[-(y_2 - y_1), (y_2 - y_1)]. \tag{6.12}$$

For any $[y_1, y_2]$ containing the zero, y, of f such that $[y_1, y_2] \subset [x_1, x_2]$, we have from equations (6.5-6.10) and the inclusion monotonicity of interval arithmetic:

$$N^{(j)}([y_1, y_2]) \subset \frac{\underline{N}^{(j-1)} + \overline{N}^{(j-1)}}{2} - \frac{F((\underline{N}^{(j-1)} + \overline{N}^{(j-1)})/2)}{F'(y) + \lambda[-(y_2 - y_1), (y_2 - y_1)]}. \tag{6.13}$$

For small $y_2 - y_1$, $F'(y) - \lambda(y_2 - y_1)$ is positive. Since $F'(y)$ is supposed to be positive, we have

$$\begin{aligned} &\omega(N^{(j)}[y_1, y_2]) \\ &\leq \frac{2\lambda(y_2 - y_1)|F((\underline{N}^{(j-1)} + \overline{N}^{(j-1)})/2)|}{(F'(y))^2 - \lambda^2(y_2 - y_1)^2} \end{aligned} \tag{6.14}$$

$$\leq \frac{\lambda(y_2 - y_1)(\overline{N}^{(j-1)} - \underline{N}^{(j-1)})|F'([\underline{N}^{(j-1)}, \overline{N}^{(j-1)}])|}{(F'(y))^2 \qquad - (\lambda)^2(y_2 - y_1)^2}$$

$$\approx \frac{\lambda(y_2 - y_1)(\overline{N}^{(j-1)} - \underline{N}^{(j-1)})|F'([\underline{N}^{(j-1)}, \overline{N}^{(j-1)}])|}{(F'(y))^2}.$$

For $j = 2 = p$, we have

$$\omega(N^{(2)}[y_1, y_2]) \approx \frac{\lambda(y_2 - y_1)(\overline{N}^{(1)} - \underline{N}^{(1)})|F'([\underline{N}^{(1)}, \overline{N}^{(1)}])|}{(F'(y))^2}. \tag{6.15}$$

But

$$(\overline{N}^{(1)} - \underline{N}^{(1)}) = \omega(N^{(1)})$$

$$\approx \frac{\lambda(y_2 - y_1)(\overline{N}^{(0)} - \underline{N}^{(0)})|F'([\underline{N}^{(0)}, \overline{N}^{(0)}])|}{(F'(y)^2}$$

$$\leq k(y_2 - y_1)^2 = k(\omega([y_1, y_2]))^2.$$

Hence from (6.5), we have

$$\omega(N^{(2)}[y_1, y_2]) \leq K_1(\omega([y_1, y_2]))^3 \tag{6.16}$$

where

$$K_1 = \frac{\lambda k|F'([\underline{N}^{(1)}, \overline{N}^{(1)}])|}{(F'(y))^2}.$$

Choosing $X^{(0)} = [y_1, y_2]$ to satisfy (6.16), it follows that

$$\omega(N^{(2)}(X^{(0)}) \cap X^{(0)}) \leq K_1(\omega(X^{(0)}))^3.$$

Defining the sequence of intervals $X^{(k)}$ by (6.7), we find

$$X^{(k+1)} \subset X^{(k)}$$

and

$$\omega(X^{(k+1)}) \leq \omega(N^{(2)}(X^{(k)})) \leq K\omega((X^{(k)}))^3.$$

If $\max(p) = m = 2$, equations (6.5-6.10) may be written as

$$N^{(k,2)}(X^{(k)}) = m(X^{(k)}) - m(U^{(k)}) - \frac{F[m(X^{(k)}) - m(U^{(k)})]}{F'(X^{(k)})} \tag{6.17}$$

$$X^{(k+1)} = N^{(k,2)}(X^{(k)}) \cap X^{(k)} \tag{6.18}$$

where

$$U^{(k)} = \frac{F(m(X^{(k)}))}{F'(X^{(k)})}.$$

The iteration (6.17-6.18) represents a modification of the Newton's interval method in which a simplified Newton interval step is combined with a Newton interval step. Although, the simplified Newton iteration exhibits superlinear convergence only under restrictive conditions, the combined iteration (6.17-6.18) has at least **cubic convergence.** More generally in the m step method in which m simplified Newton interval steps are taken between each Newton interval step, the corresponding iteration will have convergence of order $m + 2$.

We sketch the proof:

$$||X^{(k,2)} - X^*|| \leq K||X^{(k)} - X^*||^3.$$

Then

$$||X^{(k,3)} - X^*||$$
$$\leq ||F'(X^{(k)})^{-1}||||F(X^{(k,2)}) - F(X^*) - F'(X^*)(X^{(k,2)} - X^*)||$$
$$+||(F'(X^*) - F'(X^{(k)}))(X^{(k,2)} - X^*)||$$

$$\leq C_1||X^{(k,2)} - X^*||^2 + C_2||X^{(k,2)} - X^*||.||X^{(k)} - X^*||$$

$$\leq \mu||X^{(k)} - X^*||^4.$$

where C_1, C_2 and μ are suitable constants.

Proceeding in this manner, one shows easily that

$$||X^{(k,m)} - X^*|| \leq \mu_m||X^{(k)} - X^*||^{m+1}. \tag{6.19}$$

This shows that the higher order iterative processes may be generated in interval mathematics by the composition of two lower-order processes in the same manner as the noninterval iterative processes. (See Traub [22]).

Equations (6.17-6.18) are useful in that the higher-order convergence may be achieved with only one evaluation of F' and *no* higher derivatives are required. Estimate (6.19) is unimportant from the standpoint of computing actual error bound, since even if constant μ_m were known X^* is not. However equation (6.19) indicates how the convergence proceeds to X^*, the larger the μ_m, the worse is the convergence.

We may summarize the advantages of iterating the Newton's interval operator in (6.5-6.10) or (6.16-6.17) as follows:

1. $[F'(X)]^{-1}$ is computed only once for any order $N^{(k,m)}$.

2. The recursive definition of $N^{(k,m)}$ permits its calculation in a simple loop on a computer.

3. The form of $N^{(k,m)}$ suggests generalizations to system of equations with $[F'(X)]^{-1}$ replaced by $[J(X)]^{-1}$, where $J(X)$ is the Jacobian of the system.

6.2 Solution By Splitting

For a special class of problems the equation:

$$J(X)(x - Y) = f(x) \tag{6.20}$$

may be solved iteratively by the method of "splitting". Gunter Mayer [12] has recently applied the method of splitting, as applicable in noninterval analysis, to the solution of interval equations of the type $AX = b$ for cases in which the interval matrix A admits a **strong splitting** in the sense of Varga [23], here b is an interval vector and $x \in \Re^n$.

Consider the splitting:

$\mathcal{A} = \mathcal{M} - \mathcal{N}$

where $\mathcal{A}$ and $\mathcal{M}$ are interval M-matrices $[M, M] \in I\Re^{n\times n}$ and $0 \leq \mathcal{N} = [N, N] \in I\Re^{n\times n}$. If $\mathcal{A}$ has a feasible splitting $\mathcal{M} - \mathcal{N}$ with $\mathcal{M}^{-1} \geq 0$ and $\mathcal{N} \geq 0$ then the equation $AX = b$ yields a unique fixpoint x^*. These assumptions form the basis of iterative method:

$$x^{(m+1)} = IGA[\mathcal{M}, \mathcal{N}x^{(m)} + b] \tag{6.21}$$

$m = 0, 1, 2, ...$ and $x^{(0)}$ is given.

(6.21) encloses the solution set of $Ax = b$, i.e.

$$\{x\} := \{x | Ax = b, A \in \mathcal{A}, b \in b\}$$

The IGA(B, c) denotes the interval vector obtained by the Gaussian algorithm [15] applied to $B \in I\Re^{n\times n}$. Gunter has also shown that the spectral radius $\rho(M^{-1}N) \leq 1$ is an upper bound of the convergence factor α of (6.21).

Using identical procedure, we may show under similar assumptions i.e. $[J(X)]^{-1} \geq 0$, $J(X) = \mathcal{M} - \mathcal{N}$ with $\mathcal{M}^{-1} \geq 0$ and $\mathcal{N} \geq 0$, equation:

$$J(X)(x - Y) = f(x) \tag{6.22}$$

offers a Newton-like iterative process

$$x^{(0,0)} := x^{(0)} \tag{6.23}$$

$$y^{(k+1,m)} := [\tilde{x}^{(k)}, \tilde{x}^{(k)}] - IGA\{(\mathcal{M}_k, \mathcal{N}_k([\tilde{x}^{(k)}, \tilde{x}^{(k)}] - x^{(k,m-1)}) + [f(\tilde{x}^{(k)}), f(\tilde{x}^{(k)})]\} \tag{6.24}$$

$$x^{(k,m)} := y^{(k+1,m)} \cap x^{(k,m-1)}, 1 \leq m \leq r_k \tag{6.25}$$

$$x^{(k+1)} := x^{(k,r_k)} \tag{6.26}$$

$$x^{(k+1,0)} := x^{(k+1)}, k \geq 0 \tag{6.27}$$

where $x^{(0)}$ is the interval vector containing the solution of (6.22). $\mathcal{M}_k - \mathcal{N}_k$ is a splitting of $\mathcal{A}_k$ according to [23] and $\{r_k\}$ is an appropriate sequence of positive integers normally increasing monotonically. The spectral radius $\rho(M^{-1}N) < 1$ ensures convergence to the fixed point x^*.

We illustrate the splitting method with the following example:

Example

Let $f = f(f_1, f_2) = 0$

where

$$f_1 \equiv 0.7x_1 - 0.5x_2 - 0.2,$$

$$f_2 \equiv -0.5x_1 + 0.4x_2 + 0.1.$$

Here

$$J(X) = \mathcal{M} - \mathcal{N} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0.3 & 0.5 \\ 0.5 & 0.6 \end{pmatrix},$$

$J^{-1}(X) \geq 0$, $\mathcal{M}^{-1} \geq 0$ and $\mathcal{N} \geq 0$.

We start with

$$x^{(0,0)} = \begin{pmatrix} [0.8, & 1.1] \\ [0.8, & 1.1] \end{pmatrix}.$$

The equation (6.23) yields

$$y^{(1,1)} = \begin{pmatrix} [0.84, & 1.08] \\ [0.78, & 1.11] \end{pmatrix}.$$

If $r_k = 1$

$$x^{(0,1)} = y^{(1,1)} \cap x^{(0,0)} = \begin{pmatrix} [0.84, & 1.08] \\ [0.8, & 1.1] \end{pmatrix}.$$

Similarly we obtain

$$y^{(2,1)} = \begin{pmatrix} [0.852, & 1.074] \\ [0.800, & 1.100] \end{pmatrix}$$

and

$$x^{(2)} = y^{(2,1)} \cap x^{(1,1)} = \begin{pmatrix} [0.852, & 1.074] \\ [0.800, & 1.100] \end{pmatrix}.$$

Hence $x_1 = 0.963$, and $x_2 = 0.950$.

The solution gives

$f_1(x_1, x_2) = -.0009$,

$f_2(x_1, x_2) = 0.0015$.

Thus, if the matrix $J(X)$ in equation (6.20) satisfies the requirements of strong splitting, we may solve the system with the method of splitting as indicated.

Chapter 7

Conclusions

7.1 Significant Results of the Thesis

This thesis describes the interval analysis approach to solving a system of nonlinear equations by the techniques put forth by Hansen, Sengupta and Greenberg [4, 5, 6]. The multidimensional interval extension of Newton's method by Moore in the mid 60's was restricted in practice. Further studies by Krawczyk, Nickel and Neumaier [8, 14, 15, 16], following different approaches, also obtained solutions to restricted class of problems, thus lacking full generality. Computations involving the solution of systems of nonlinear equations, using interval mathematics require careful analysis for devising an efficient algorithm. The present thesis describes the development of procedures leading to the main algorithm as proposed by Hansen and Greenberg [5]. The Hansen-Greenberg algorithm has resulted from the earlier investigations of Hansen and Smith [7], and Hansen and Sengupta [6]. Hansen and Greenberg have in fact deduced a single algorithm of greater efficiency, combining the three known subalgorithms. The time consuming and expensive task related to the elimination procedure is not started until the bounding interval containing a solution is reduced significantly as per requirement set for the problem. Division by an interval containing a zero is avoided as much as possible, and the extended interval arithmetic is used only when necessary. The subalgorithms constituting the main Hansen-Greenberg algorithm are:

1. The Mean Value linearization of the nonlinear system.

2. Initial preconditioning of the linearized system.

3. Hansen-Sengupta Step.

4. The Real (local/noninterval) iteration.

5. Gaussian elimination by LU decomposition.

The preconditioning of linearized system, Hansen-Sengupta step and the real iteration enable the algorithm to obtain a more precise solution with a single application of the elimination procedure. The order in which these subalgorithms must be carried out and repeated, when necessary, is important for achieving efficient results. The present thesis examines the justification for these procedures; which collectively seem to account for a more efficient algorithm.

The Hansen-Greenberg's method nevertheless offers a more efficient means of solving the system of nonlinear equations. But as noted earlier, the algorithm can involve **extensive** use of subalgorithms for narrowing down the widths of intervals containing the solution. Such extensive use of the subalgorithms for some problems may not produce as efficient results as expected. In this respect, the present thesis outlines further procedural changes that **can** contribute to the efficiency of the main algorithm. These procedural extensions are as follows:

1. The matrix $M^{(0)}(X)$ associated with the preconditioned system

$$M^{(0)}(x - Y) = b$$

can be approximated closer to the unit matrix I as shown in Section 5.1. We have, in fact, shown that the new matrix $M^{(1)}(X) = ([L^{(1)}, U^{(1)}])$ defined by

$$L^{(1)} \approx (2I - L^{(0)})L^{(0)},$$

$$U^{(1)} \approx (2I - U^{(0)})U^{(0)},$$

gives better approximating interval $[Y_l^{(1)}, Y_u^{(1)}]$ for the unit matrix, where

$$Y_l^{(1)} = x^{(0)} - (2I - L^{(0)})b_l,$$

$$Y_u^{(1)} = x^{(0)} - (2I - U^{(0)})b_u.$$

This new starting interval, bounds the solution more closely as demonstrated with examples.

2. The Gauss-Seidel iteration may be relpaced by the SOR (Successive Overrelaxation) Gauss-Seidel iteration described in Section 5.2. The real iteration yields sufficiently small values of $||f(x)||$, if the Gauss-Seidel step results into a better approximating interval. Thus the SOR Gauss-Seidel iterates should be preferred, where more precise inclusions are necessary for performing real iterations and subsequent eliminations in the Hansen-Greenberg algorithm. Section 5.2 also derives an extension of the Ostrowski-Reich theorem as it applies in interval analysis.

3. The Hansen-Greenberg method employs a noninterval quasi-Newton iteration $x^{(n+1)} = x^{(n)} - Bf(x^{(n)})$ in the main algorithm, to help the relaxation process. We have shown in Section 5.3, that, this noninterval iteration may be replaced by a pair of interval equations:

$$X^{(k+1)} = \{m(X^{(k)}) - V^{(k)}b\} \cap X^{(k)}$$

$$V^{(k+1)} = \{m(V^{(k)}) + V^{(k)}(I - M^{(k)}m(V^{(k)})\} \cap V^{(k)}$$

$k \geq 0$, and $V^{(0)}$ is an interval matrix with the property that $M^{-1}(X) \in V^{(0)}$. This modification will provide us with a procedure confined to the domain of interval analysis, besides offering better inclusions.

4. In Chapter 6.1, we examined the interval Newton's method with secondary steps. In particular, Newton's formula containing one Newton step has been shown to have at least a **cubic convergence.** The modified Newton's formula uses fewer computations of the first derivative. This fact is common knowledge for the noninterval iterations. But we have now established a proof, that, it also applies for interval procedures. A numerical example from Moore [13] has been chosen to demonstrate the effectiveness of this procedure.

5. Finally, it has been pointed out in Chapter 6.2 that some systems of the type $AX = b$ can be solved by the method of splitting, as shown recently by Mayer [12]. We have applied this procedure to solve the system $M(X)(x - Y) = b$, encountered in our work, in dealing with the nonlinear systems. These meth-

ods in conjunction with the Hansen-Greenberg approach may provide suitable alternatives for better solutions.

7.2 Suggestions for Further Research

Future research efforts in the following areas seem to hold promising results.

1. Use of a better Jacobian matrix.

   The Jacobian interval matrix $J(X)$ is used in the linearized system:

   $J(X)(x-y) = f(x)$

   where X is the interval containing both x and y. The above equation is further conditioned by multiplying it with the inverse of mid$J(X)$. Certain arguments in the Jacobian J could be real rather than interval. This argument could as well be the point x, about which the function $f(x)$ is expanded. If a real inner iteration is used then the improved value of $J(X)$ and, therefore $BJ(X)$, would also change. However, obtaining a better Jacobian requires more sophisticated analysis and extra programming. Hansen and Greenberg also point out that it is some what more efficient to use a better Jacobian than to perform only the real iterations. Further work in this direction can improve the algorithm.

2. Coupling the Splitting Method with the Hansen-Greenberg's Algorithm.

   It is quite possible that the method of splitting as illustrated in Section 6.2 and discussed in more detail by G.Mayer [12], can be coupled with the Gaussian elimination. Consider the splitting of $J(X) = M - N$ in $J(X)(x - Y) = f(x)$ and the iteration:

$$Y^{(k+1)} = \hat{x} + M^{-1}[NY^{(k)} - (N\hat{x} + f(\hat{x}))],$$

$$X^{(k+1)} = Y^{(k+1)} \cap X^{(k)}.$$

The convergence of above iteration depends upon the system being solved and the associated splitting. Some times it may be more practical to use the splitting technique after the linearized system has been conditioned to

$$M(X)(x - Y) = b.$$

Splitting the diagonally dominant matrix $M(X)$, (which is also approximating closely to the unit matrix), into $M_0 - I$, we obtain

$$Y^{(k+1)} = M_0^{-1}\{IY^{(k)} + (M\hat{x} - b)\}.$$

This iteration should provide a convergent sequence for bounding solution. Thus the method of splitting can serve an alternative to the LU decomposition in the Hansen-Greenberg's method, when appropriate.

3. Using the Method of Continuation:

As is well known, many iterative methods will converge to a solution x^* of $f(x) = 0$, only if the initial approximations are close to $x^{(n)}$. The continuation method:

$$G(t, x) = tf(x) + (1 - t)f(x^{(0)}),$$

$$x^{(n+1)} = G(t, x^{(n)}).$$

may be used as an attempt to obtain sufficiently close starting points. Although many researchers have discussed the numerical continuation approach, the feasibility of the numerical continuation process for general systems still

remains an open topic of investigation. Further research in this area can prove useful in the present investigations.

Bibliography

[1] G.Alefeld and J.Herzberger, Introduction to Interval Computations, Academic Press, New York, 1983.

[2] E.T.Bell, The Development of Mathematics, Second Edition, McGraw Hill, New York, 1945.

[3] S.Banach, Sur les ope'rations dans les ensembles abstraits et leur applications aux equations integreles. Fund. Math., 3(1922), 133-181.

[4] E.R.Hansen, On Solving Systems of Equations Using Interval Arithmetic, Math. Comp. 22, 374-384 (1968).

[5] E.R.Hansen and Greenberg, An Interval Newton Method, Applied-Mathematics And Computation, 12, 89-98, 1983.

[6] E.R.Hansen and Sengupta, Bounding Solutions of Systems of Equations Using Interval Analysis, BIT, 21, 203-211, 1981.

[7] E.R.Hansen, and, R.R.Smith, Interval arithmetic in matrix computations, part ii, SIAM Journal. Numer. Anal. 4(1967) 1-9.

[8] R.Krawczyk, Newton-Algrithmen zur Bestimmung von Nullstellen mit Fehlerschranken, Computing 4, 187-201, 1969.

[9] W.M.Kahan, A more complete interval arithmetic, Report, University of Toronto, 1968.

[10] W.M.Kahan, Gauss-Seidel methods for solving systems of linear equations. Doctoral thesis, University of Toronto, 1958.

[11] L.V.Kantorovic, Functional Analysis and Applied Math., Uspehi U Mat. Nauk., 3(1948) 89-185(Russian), Tr by C.D.Benster; Natl Bur. Std. Rept. No. 1509, Washington, 1952.

[12] G.Mayer, Comparison Theorems For Iterative Methods Based On Strong Splitting, SIAM J.Numer. Anal., Vol.24, No.1, 1987.

[13] R.E.Moore, Interval Analysis, Prentice-Hall, Englewood Cliffs, N.J. 1966.

[14] K.Nickel, On the Newton Method in Interval Analysis, MRC, Report. No.1136, Mathematics Research Center, University of Wisconsin, Madison, 1971.

[15] A.Neumaier, Linear Interval Equations, Lecture Notes in Computer Science, Edited by K.Nickel, Springer-Verlag, p109-120, Vol.212, 1985.

[16] A.Neumaier, Interval iteration for zeros of systems of equations, BIT, 25, 256-273, 1985.

[17] J.M.Ortega and W.C.Reinboldt, Iterative Solution of Nonlinear Equations In Several Variables, Academic Press, New York, 1970.

[18] A.M.Ostrowski, Solution of Equations and System of Equations, Second Ed., Academic Press, New York, 1986.

[19] H.Ratschek, Interval mathematics, Freiburger Intervallberichte, 87/4, Institut fur Angewandte Mathematik, Universitat Freiburg i.Br.

[20] H.Ratschek and J.G.Rokne, New Methods for Global Optimization, Horwood Chichester, 1988.

[21] J.Sheldon, On the numerical solution of elliptic equations, Math Tables Aids Comput, 9, 101-112, 1955.

[22] J.Traub, Iterative Methods for the Solutions of Equations, Prentice-Hall, 1966.

[23] R.S.Varga, Matrix Iterative Analysis, Prentice-Hall, 1962.

[24] M.A.Wolfe, A Modification of Krawczyk's Algorithm, SIAM J. Numer. Anal. 17, 376-379, 1980.

[25] A.Wouk, Direct Iteration, Existence and Uniqueness, In P.M. Anselone Ed., Nonlinear Integral Equations, University of Wisconsin, Madison, 3-33, 1964.