2011

A Study on Interval Solutions of Nonlinear Systems

Mollah, Md. Shirazul Hoque

University of Rajshahi

A Study on Interval Solutions of Nonlinear Systems

By
Md. Shirazul Hoque Mollah

The thesis submitted to

Department of Mathematics

Faculty of Science

University of Rajshahi, Rajshahi-6205, Bangladesh in

partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Mathematics

June, 2011.

Statement of Originality

I declare that the works in my Doctor of Philosophy thesis entitled **"A Study on Interval Solutions of Nonlinear Systems"** is original and accurate to the best of my knowledge. I as declare that the materials contained in my research work have not been previously published or written by any person for any kind of degree.

20.06.2011

Md. Shirazul Hoque Mollah
Roll No. 122
Reg. No. 0006,
Session 2005-2006
Ph.D Research Fellow
Faculty of Science
Department of Mathematics
Rajshahi University, Rajshahi-6205, Bangladesh.

Dr. M. Zulfikar Ali
Professor
Department of Mathematics
University of Rajshahi
Rajshahi- 6205, Bangladesh

Phone : 750041-49/4108
Mobile : 01556-308978
Fax : 0721-750065
E-mail : zulfi1022002@yahoo.com

No. Date :.....

Certificate

I have the pleasure to certify that the Doctor of Philosophy thesis entitled **"A Study on Interval Solutions of Nonlinear Systems"** submitted by Md. Shirazul Hoque Mollah in fulfillment of the requirements for the Degree of Doctor of Philosophy in Mathematics, University of Rajshahi, Department of Mathematics, Rajshahi- 6205, Bangladesh has been completed under my supervision. I believe that the research work is an original one and it has not been submitted elsewhere for any kind of degree.

I wish him a bright future and every success in life.

Supervisor

20.6.11

Dr. M. Zulfikar Ali
Professor
Department of Mathematics
University of Rajshahi
Rajshahi- 6205, Bangladesh

Dedicated to the members of my family

Abstract

Suppose we seek a solution of the nonlinear system

$$f(x) = \{f_i(x_1, x_2, \ldots, x_n)\} = 0 \quad i = 1, 2, \ldots, n \tag{1}$$

where $f_1, f_2, f_3, \ldots, f_n$ are continuous functions on an open set D in $\mathbb{R}^n$.

There are good many methods for iterative interval solutions of system (1) for any such methods, R. E. Moore developed a technique for finding a safe starting point from which iterates converge, with a particular iterative method in mind, Krqwczyk's operator.

Minoru Urabe established an existence and uniqueness theorem (1965) which helps verify the existence and uniqueness of an exact solution and to know the error bound to an approximate solution of a system like (1). His theorem assumes that all the computations are to be carried out in real numbers exactly.

Our attempts will be made to combine M. Urabe's theorem and R. E Moore's technique theoretically as well as numerically considering interval version of Newton's method.

Acknowledgments

First of all, I would like to express my sincere admiration, appreciation and deepest gratitude to my honorable supervisor **Dr M. Zulfikar Ali**, Professor, Department of Mathematics, Rajshahi University, Rajshahi-6205, for his expert guidance and many suggestions, encouragement, and constant support during my research work. I am indebted to him for his willingness to devote both time and effort to me as well as for patience and generosity.

I am also very grateful to my honorable Chairman, Department of Mathematics, Rajshahi University, Rajshahi-6205, Bangladesh for providing me the Computer facilities of the department and making all the department feasibilities those been enabled me to work in comfortable atmosphere.

I am thankful to the University of Rajshahi and also very grateful to all the teachers of this department for their invaluable advice and co-operation in different respects during my stay in this department and also thankful to all of my friends, staffs and attendants of this department.

I am grateful to the Vice-Chancellor, Dhaka University of Engineering and Technology, Gazipur for his kind permission for admission to the University of Rajshahi, Department of Mathematics as a research fellow.

University of Rajshahi, Rajshahi-6205 Md.Shirazul Hoque Mollah

June , 2011

Contents

Chapter 1

Introduction

In mathematics, there are real numbers, a real arithmetic for combining them, and a real analysis for studying the properties of the numbers and the arithmetic. Interval mathematics is a generalization in which interval numbers replace real numbers, interval arithmetic replaces real arithmetic and interval analysis replaces real analysis. Using a pair $[a,b]$ of computer numbers to represent an interval of real numbers $a \leq x \leq b$, we will define an arithmetic for intervals and interval valued extensions of functions commonly used in computing. In this way, an interval $[a,b]$ has a dual nature. It is a new kind of number pair and it represents a set $[a,b] = \{x : a \leq x \leq b\}$. We combine set operations on intervals with interval function evaluations to get algorithms for computing enclosures of sets of solutions to computational problems.

Numerical analysis is the study of computing with real (and other kinds of) numbers. Theoretical numerical analysis might entail exact numbers and exact arithmetic, while practical numerical analysis consider the cases in which rounding errors occur. We will be concerned with both theoretical and practical interval

analysis for computing with interval numbers.

There are several types of mathematical computing errors. Anyone using arithmetic of finite precision on digital computers has come across the impact of rounding error and propagated error. These errors appear as a result of the uncertainty of initial data, finite representation of numbers in arithmetic units of computing machines, or uncertain values of parameters in mathematical models in physical and engineering sciences. When solving computational problems in many similar situations, we should compute intervals which contain the approximate result or the solution of some equation, providing thus the upper error bound related to the exact result or solution.

In solving some practical problems "exact" arithmetic cannot be carried out because of the limited precision of the used computer. The calculations in ordinary floating point arithmetic do not normally produce any information about the accuracy of the obtained result. In order to provide such information **R.E. Moore** has proposed in 1966 a new structure which has come to be called **Interval Analysis or, later, Interval Mathematics**. Applying interval arithmetic, we can compute the intervals containing the results of the arithmetic of "infinite" precision. Most of the aforementioned problems can be solved using the methods of interval mathematics- a new expanding and applicable branch of applied mathematics. At present, these methods have practical application in various scientific areas including physical and engineering science, economics study of the behavior of mathematical models as sensitivity analysis, effects of inaccurate measurements, parameter studied, the mathematical programming, numerical analysis, approximation and optimization theory, asteroid orbits, robotic, signal processing, computer graphics, behavioral ecology and other disciplines. Generally speaking, interval arithmetic methods can be effectively applied to any compu-

tation in the presence of roundoff error or propagation error in initial data or for controlling computational errors in various numerical algorithms implemented on digital computers.

In practice, the resulting intervals may be very large in size and, therefore, of little value. The aim of interval methods is to modify the existing numerical algorithms, or to develop new ones that produce intervals as small as possible. Interval techniques are a useful tool in applied mathematics, which is especially prominent, in controlling and analyzing computational error, constructing self-validated algorithms with guaranteed safe bounds (upper and lower bounds on sets of solution or computed result) and in providing natural stopping criteria for interval methods.

In particular, error bound procedures for solving certain problems in the complex realm require complex interval arithmetic. This type of arithmetic is a natural extension of real interval arithmetic to the complex plane and used, beginning with Boche(1966). A complex "interval" can be a rectangle, a circle; or a more complicated set. Intervals of magnitude and phase can also be used. When complex intervals are used, then the computed complex intervals (circles or rectangles) contain the exact results of some computation or the exact solution of an equation. The center of this interval is taken for the approximate result, and the radius of circle(disk) represents the upper error bound. In this manner, an automatic control of accuracy of the obtained result is provided, which is the main advantage of interval methods.

Interval analysis has been used in rigorous computer-assisted proofs, for example, Hales' proof of the Kepler conjecture.[4]

An interval Newton method has been developed for solving systems of nonlinear equations. While inheriting the local quadratic convergence properties of the ordinary Newton method, the interval Newton method can be used in an algorithm that is mathematically guaranteed to find all roots within a given starting interval.

Interval analysis permits us to compute interval enclosures for the exact values of integrals. Interval methods can be bound the solutions of linear systems with inexact data. There are rigorous interval branch-and bound methods for global optimization, constraint satisfaction, and parameter estimation problems. We also introduce INTLAB (INTerval LABoratory), a powerful and flexible MATLAB toolbox capable of performing interval calculations[22].

In Chapter 2, introduces the interval number system, defines the basic terms and concepts, set operations (intersection and union) and arithmetic operations (addition, substraction, multiplication,and division), rectangular complex interval arithmetics, algebraic properties of interval arithmetic and inclusion isotonicity of interval arithmetic, needed to work within this system.

In Chapter 3, review of some basic definitions and Fundamental results, deals with sequences of intervals and interval functions, material needed as preparation for the iterative methods to be treated in Chapter 4 (convergence test), in Chapter 5 (on matrices) and in Chapter 6 (on root finding). Finally, in Chapter 7, we apply M. Urabe's theorem to R. Krawczyk's algorithm for finding the interval (numerical) solution of a polynomial equation and the system of nonlinear equations. Finally we present an numerical approach for searching the safe starting regions to solutions of nonlinear system of equations, based on interval computation by applying S.T Jones and R.E. Moore method. Here, we are given an n-dimensional interval bisection procedure, whose stopping criterion is the satisfaction of com-

putationally verifiable tests for existence of a solution in a test region and convergence of a given iterative method from the test region. This is a new area and research is still in progress, looking for improved method.

Research, development, and application of interval methods is now taking place in many countries around the world, especially in Germany, but also in Austria, Belgium, Brazil, Bulgaria, Canada, China, Denmark, Finland France India, Japan, Norway, Spain Sweden, Russia, the U.K and USA.

Chapter 2

The Interval Number System

2.1 Basic Terms and Concepts

A fundamental notion in interval analysis is the notion of interval. Let $\mathbb{R}$ be the set of all real numbers. A subset of $\mathbb{R}$ of the form

$$[a,b] = \{x \in \mathbb{R} : a \leq x \leq b\} \tag{2.1}$$

where $a, b \in \mathbb{R}$ and $a \leq b$ is called a closed real interval. We can also regard an interval as a number by the ordered pair of two real numbers, namely, its endpoints a and b with $a \leq b$. Although various other types of intervals (open, half-open) appear throughout the mathematics, our work will center primarily on closed intervals. In this thesis paper the term interval will mean **closed interval**.

The set of all intervals over $\mathbb{R}$ is denoted by $\mathbb{IR}$ where

$$\mathbb{IR} = \{[a,b] : a,b \in \mathbb{R}, a \leq b\} \tag{2.2}$$

The elements of $\mathbb{IR}$ will also called interval numbers.

Degenerate Intervals

We say that X is degenerate if $a = b$. Such when an interval of the form $X = [a,a] = a$ consists of one point is called a point interval, and is sometimes denoted by a which is a real numbers. From this point of view the set of real numbers is contained in the set of interval numbers i.e, $\mathbb{R} \subset \mathbb{IR}$.

Endpoint Notation, Equality

We will adopt the convention of denoting intervals by capital letters X, Y, Z etc and their endpoints by small letters a,b,c etc. We call two intervals $X = [a,b]$ and $Y = [c,d]$ are equal if and only if their corresponding endpoints are equal, that is $X = Y$ iff $a = c$ and $b = d$.

Intersection, Union, and Interval Hull

The intersection of two intervals X and Y is empty, $X \cap Y = \phi$, if either $X < Y$ or $Y < X$. Otherwise, we may define the intersection of X and Y is again an interval

$$\begin{aligned} X \cap Y &= \{z : z \in X \text{and} \quad z \in Y\} \\ &= [\max(a,c), \min(b,d)] \end{aligned}$$

In this latter case, the union of X and Y is also an interval:

$$\begin{aligned} X \cup Y &= \{z : z \in X \text{or} \quad z \in Y\} \\ &= [\min(a,c), \max(b,d)] \end{aligned}$$

In general, the union of two intervals is not an interval. However, the interval hull of two intervals, defined by

$$X \underline{\cup} Y = [\min(a,c), \max(b,d)]$$

is always an interval and can be used in interval computations. We have

$$X \cup Y \subseteq X \underline{\cup} Y$$

for any two intervals X and Y.

Importance of Intersection

Intersection plays a key role in interval analysis. If we have two intervals containing a result of interest- regardless of how they were obtained- then the intersection, which may be narrower, also contains the result.

Width, Absolute Value, Midpoint

1. The width of an interval $X = [a,b]$ is the real number defined and denoted by

$$W(X) = b - a \tag{2.3}$$

2. The absolute value of an interval X denoted by $|X|$, is the maximum of the absolute values of its endpoints:

$$|X| = \max(|a|, |b|) \tag{2.4}$$

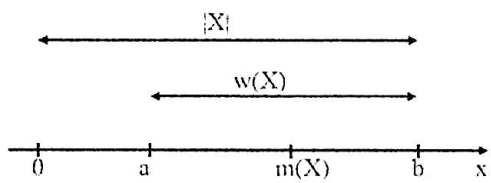

Figure 2.1: Width, absolute value and midpoint of an interval.

3. The midpoint(or center)of an interval X is the real number

$$m(X) = (a+b)/2 \tag{2.5}$$

2.2 Order Relations for Intervals

We know that the real numbers are ordered by the relation $<$. This relation is said to be transitive: if $a < b$ and $b < c$, then $a < c$ for any a,b and $c \in \mathbb{R}$. A corresponding relation can be defined for intervals, and we continue to use the same symbol for it: The elements of $\mathbb{IR}$ can be ordered in the following way: $X < Y$ iff $b < c$ and $X > Y$ iff $b > c$.
Another transitive order relation for intervals is set inclusion:

$$X \subseteq Y \quad \text{if and only if} \quad c \leq a \quad \text{and} \quad b \leq d \tag{2.6}$$

2.3 Definitions of the Arithmetic Operations

We are about to define the basic arithmetic operations between intervals. The key point in these definitions is that computing with intervals is computing with sets. Interval arithmetic operations are defined on $\mathbb{IR}$ such that the interval result encloses all possible real results. In this subsection arithmetic operations with intervals will be introduced. Let $\diamond$ denote one of the four arithmetic operators,$+,-,*,/$ on the real numbers. For $X,Y \in \mathbb{IR}$ we define an arithmetic operation on $\mathbb{IR}$ by

$$X \diamond Y = \{x \diamond y | a \leq x \leq b, c \leq y \leq d\} \tag{2.7}$$

with X/Y is undefined if $0 \in Y$

Let $X,Y \in \mathbb{IR}$. The sum of X and Y, denoted by $X+Y$, is defined by the set

$$X+Y = \{x+y : x \in X, y \in Y\} \tag{2.8}$$

It is seen that $X+Y$ is again an interval,i.e, $X+Y \in \mathbb{IR}$. Indeed, from (2.7) $a+c \leq x+y \leq b+d$. Thus we have the equivalent relation

$$X+Y = [a,b]+[c,d] = [a+c, b+d] \tag{2.9}$$

Although (2.8) and (2.9) are equivalent, formula(2.9) is by far more useful for practical applications since it permits to find the whole set $X+Y$ by computing its endpoints $a+c$ and $b+d$ using only the corresponding endpoint of X and Y.

We define the negative of an interval by the set $-X = \{-x : x \in X\}$. Similarly to the previous case we have $-X = -[a,b] = [-b,-a]$

The difference of two intervals X and Y is the set

$$X - Y = X + [-Y] = \{x - y : x \in X, y \in Y\} \tag{2.10}$$

or equivalently

$$X - Y = [a - d, b - c] \tag{2.11}$$

Obviously, $X - Y \in \mathbb{IR}$

The product $X.Y$ of two intervals X and Y is defined by the set

$$X.Y = \{xy : x \in X, y \in Y\} \tag{2.12}$$

It is not hard to see that

$$X.Y = [\min(ac,ad,bc,bd), \max(ac,ad,bc,bd)] \tag{2.13}$$

The end points of the product $Z = X.Y = [Z_1, Z_2]$ can be computed in a cheaper way if the signs of the end points of X and Y are taken into account. We have the following nine cases :

1. $Z_1 = ac, \quad Z_2 = bd \quad \text{if} \quad a \geq 0, \quad c \geq 0$

2. $Z_1 = ad, \quad Z_2 = bd \quad \text{if} \quad a < 0 < b, \quad c \geq 0$

3. $Z_1 = ad, \quad Z_2 = bc \quad \text{if} \quad b \leq 0, \quad c \geq 0$

4. $Z_1 = ad, \quad Z_2 = bc \quad \text{if} \quad a \geq 0, \quad c < 0 < d$

5. $Z_1 = ad, \quad Z_2 = ac \quad \text{if} \quad c \leq 0, \quad c < 0 < d$

6. $Z_1 = bc, \quad Z_2 = ad \quad \text{if} \quad a \geq 0, \quad c \leq 0$

7. $Z_1 = bc, \quad Z_2 = ac \quad \text{if} \quad a < 0 < b, \quad c \leq 0$

8. $Z_1 = bd, \quad Z_2 = ac \quad \text{if} \quad a \leq 0, \quad c \leq 0$

9. $Z_1 = \min\{ad, bc\}, \quad Z_2 = \max\{ac, bd\}$ if $a \leq 0 < b, c < 0 < d.$

It is seen that with the exception of the ninth case formula are twice more effective than (18). If X is an interval not containing the number 0, then we can define its reciprocal as follows:

$$\frac{1}{X} = \{\frac{1}{x} : x \in X\} \tag{2.14}$$

and hence

$$\frac{1}{X} = [\frac{1}{b}, \frac{1}{a}] \tag{2.15}$$

Finally, the quotient X/Y is defined as

$$X/Y = \{x/y : x \in X, y \in Y\} \tag{2.16}$$

provided that $0 \notin Y$.

A sophisticated extension of the interval arithmetic operations to unbounded intervals is needed. It is required for the interval Newton method. Alefeld (1968) was the first to use infinite intervals in Newton methods[18] . The following formulas are due to Hansen (1980):

Let $0 \in [c, d]$ and $c < d$, then

$$[a,b]/[c,d] = \begin{cases} [b/c,+\infty) & \text{if } b \le 0 \text{ and } d = 0, \\ (-\infty,b/d] \cup [b/c,+\infty), & \text{if } b \le 0, c < 0 \text{ and } d > 0, \\ (-\infty,b/d] & b \le 0 \text{ and } c = 0, \\ (-\infty,a/c] & a \ge 0 \text{ and } d = 0, \\ (-\infty,a/c] \cup [a/d,+\infty) & a > 0, c < 0 \text{ and } d > 0, \\ (-\infty,+\infty) & \text{if } a < 0 \text{ and } b > 0, \end{cases}$$

and furthermore $[a,b]/0 = (-\infty,+\infty)$.

These formulas are not applicable to every problem, but they are appropriate for solving linear equations in connection with the interval Newton method. There is also no need to implementing the above formulas on the machine explicitly since they are finally intersected with a bounded interval such that the result is always a bounded intervals,a pair of bounded intervals, or the empty set. We also have to shift unbounded intervals before intersecting them. This means that for $x, a \in \mathbb{R}$,

$$x + [a,\infty) = [x+a,\infty) \tag{2.17}$$

$$x + (-\infty,\infty) = (-\infty,\infty) \tag{2.18}$$

$$x + (-\infty,a] = (-\infty,x+a] \tag{2.19}$$

2.4 Interval Vectors

By an n-dimensional interval vector, we mean an ordered n-tuple of intervals

$$(X_1, X_2, \cdots, X_n)$$

Thus, if X is a two-dimensional interval vector then $X = (X_1, X_2)$ for some intervals $X_1 = [a_1, b_1]$ and $X_2 = [a_2, b_2]$.

A two-dimensional interval vector also represents a two-dimensional rectangle of

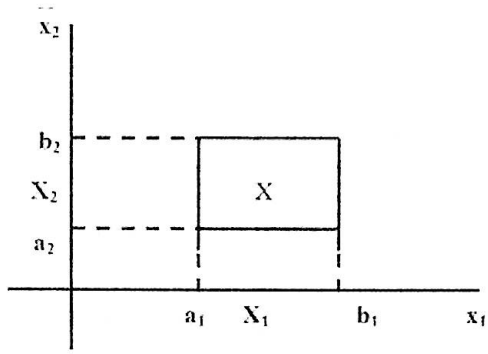

Figure 2.2: Two dimensional interval vector $X = (X_1, X_2)$.

points (x_1, x_2) such that $a_1 \leq x_1 \leq b_1$ and $a_2 \leq x_2 \leq b_2$. With suitable modifications, many of the notions for ordinary intervals can be extended to interval vectors.

1. If $x = (x_1, x_2, \cdots, x_n)$ is a real vector and $X = (X_1, X_2, \cdots, X_n)$ is an interval vector , we will write $x \in X$ if $x_i \in X_i$ for $i = 1, 2, 3, \cdots, n$.

2. The intersection of two interval vectors is empty if the intersection of any of their corresponding components is empty, otherwise, for $X = (X_1, X_2, \cdots, X_n)$ and $Y = (Y_1, Y_2, \cdots, Y_n)$ we have $X \cap Y = (X_1 \cap Y_1, \cdots, X_n \cap Y_n)$, which is again an interval vector. If two intervals X and Y have nonempty intersection their union,

$$X \cup Y = [\min(a, c), \max(b, d)] \tag{2.20}$$

is again an interval. The union of two intersecting interval vectors is not in general, an interval vector.

3. A useful relation for intervals is the set inclusion

$$X \subseteq Y \tag{2.21}$$

if and only if $a \leq c$ and $b \leq d$. If $X = (X_1, X_2, ..., X_n)$ and $Y = (Y_1, Y_2, \cdots, Y_n)$ are interval vectors, we have $X \subseteq Y$ if $X_i \subseteq Y_i$ for $i = 1, 2, 3, \cdots, n$.

4. The width of an interval vector $X = (X_1, X_2, \cdots, X_n)$ is the largest of the widths of any of its component intervals:

$$W(X) = \max(W(X_1), \cdots, W(X_n)) \tag{2.22}$$

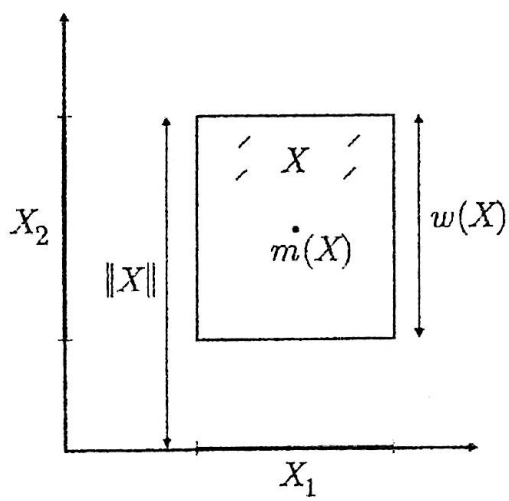

Figure 2.3: Width, norm, and midpoint of an interval vector $X = (X_1, X_2)$.

5. The midpoint of an interval vector $X = (X_1, X_2, \cdots, X_n)$ is

$$m(X) = (m(X_1), \cdots, m(X_n)) \tag{2.23}$$

6. The norm of an interval vector $X = (X_1, X_2, \cdots, X_n)$ is

$$\|X\| = \max(|X_1|, \cdots, |X_n|) \tag{2.24}$$

2.5 Rectangular Complex Arithmetic

We consider and analysis of various problems in the complex plane which either involve "inexact" data, or require some information on upper error bound of the obtained result or solution, dictate the need for a structure which is referred to as *complex interval arithmetic*.There are two reasonable choices for complex intervals: circular regions(disks) and rectangles in the complex plane. No new difficulties arise if we choose the Cartesian representation for complex numbers.

As is customary, we denote the complex numbers as an ordered pair of real numbers, (a_1, a_2).

Definition 2.5.1 *The sum of two complex numbers,* (a_1, a_2) *and* (b_1, b_2) *is the complex number* $(a_1 + b_1, a_2 + b_2)$.

Definition 2.5.2 *The product of two complex numbers,* (a_1, a_2) *and* (b_1, b_2), *is the complex number* $(a_1 b_1 - a_2 b_2, a_1 b_2 + a_2 b_1)$.

Under these definitions we note that the real number, a, corresponding to and may be identified with the complex number, $(a, 0)$. Traditionally the complex number, $(0, 1)$ is denoted by the letter i. Then since $(0, 1).(0, 1) = (-1, 0), \quad i^2 = -1$. In summary the complex number (a_1, a_2), could be represented as $(a_1, a_2) = (a_1, 0) + (0, a_2) = a_1 + a_2(0, 1) = a_1 + a_2 i$.
We wish to develop complex interval numbers in a similar manner.

Definition 2.5.3 *A complex interval number is an ordered pair of interval numbers* (X, Y).

Upper case script letters are used to denote complex interval numbers. As in the case with ordinary complex numbers, we shall have occasion to refer to the complex interval number, $([0, 1], [1, 1]$ as "i".

Let $X_1, X_2 \in \mathbb{IR}$. Then in set notation a complex interval number may be represented in the form

$$X = X_1 + X_2 i = [a,b] + [c,d]i = \{x+iy | a \leq x \leq b, \quad c \leq y \leq d\} \qquad (2.25)$$

Geometrically a complex interval number may be conceived of as a closed rectangular region in the complex plane. The set of this rectangles is denoted by $\mathbb{R}(\mathbb{C})$ where $\mathbb{C}$ is the set of complex numbers. A complex number $x+iy$ may be considered to be a complex point interval $X = [x,x] + i[y,y] \in \mathbb{R}(\mathbb{C})$.Futhermore, every real interval $X_1 \in \mathbb{R}(\mathbb{C})$ may be considered to be an element $X = X_1 + i[0,0] \in \mathbb{R}(\mathbb{C})$which evidently implies $\mathbb{IR} \subset \mathbb{R}(\mathbb{C})$. The midpoint(center) and the semidiagonal of a rectangle X will denoted by mid$X$ and sd$X$.[6]

Let $\diamond \in \{+,-,*,/\}$ be a binary operation on elements from $\mathbb{IR}$ and let $X = X_1 + iX_2, \quad Y = Y_1 + iY_2, \quad X,Y \in \mathbb{R}(\mathbb{C})$.

The basic operations of rectangular arithmetic are defined by

$$X \pm Y = X_1 \pm Y_1 + i(X_2 \pm Y_2)$$

$$X * Y = X_1 Y_1 - X_2 Y_2 + i(X_1 Y_2 + X_2 Y_1)$$

$$X/Y = \frac{X_1 Y_1 + X_2 Y_2}{Y_1^2 + Y_2^2} + i\frac{X_2 Y_1 - X_1 Y_2}{Y_1^2 + Y_2^2}, \; 0 \notin Y_1^2 + Y_2^2$$

2.6 Algebraic Properties of Interval Arithmetic

If X and Y are degenerate intervals, then (2.7) reduce to the ordinary arithmetic operations over real numbers. Thus, interval arithmetic can be regarded as a generalization of real arithmetic. Therefore, it is normal to expect that the properties of interval arithmetic will be similar to those of real arithmetic, which is really

the case. However, there are several striking dissimilarities that will be stressed below. It follows from the set-theoretic definitions (2.7) interval addition and multiplication is associative and commutative, that is , if $X, Y, Z \in \mathbb{IR}$.

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

$X + Y = Y + X$

$X(YZ) = (XY)Z, XY = YX$ zero and unity in $\mathbb{IR}$ are the degenerate intervals $[0,0]$ and $[1,1]$ which will be denoted by 0 and 1 respectively.

In other words: $X + 0 = 0 + X, \quad I.X = X.I$ for any $X \in \mathbb{IR}$.

With the identification of degenerate intervals and real numbers, interval arithmetic is an extension of real arithmetic and reduces to ordinary real arithmetic for intervals of zero width. Note that $X - X = 0$ and $X/X = 1$ only when X is of width zero.

It is important to underline that unlike real arithmetic $X - X \neq 0$

and $X/X \neq I$ when $W(X) > 0$. Indeed, $X - X = [a - b, b - a] = W(X)[-1, 1]$ and $X/X = [a/b, b/a]$ for $X > 0$ or $X/X = [b/a, a/b] \quad$ for $X < 0$.

Another interesting property of interval arithmetic is the fact that the distributive law

$$X(Y + Z) = XY + YZ \tag{2.26}$$

does not always hold.

For example, we have $[1,2](1,-1) = 0$; Whereas $[1,2].1 - [1,2].1 = [-1,1] \neq 0$.

We do, however, always have the following algebraic property

$$X(Y + X) \subseteq XY + XZ \tag{2.27}$$

We call this property sub-distributivity. In certain special cases , distributivity holds. Some particularly useful cases are : The interval number X is of zero

width, $X(Y+Z) = XY+XZ$.

The cancellation law holds for interval addition : $X+Y = X+Z$ implies $Y = Z$. If $XY = XZ$ and $0 \notin X$, then the cancellation law for multiplication holds, and $Y = Z$.

2.7 Symmetric Intervals

By a symmetric interval, we mean an interval $X = [a,b]$ such that $a = -b$. Thus, X is symmetric if and only if $m(X) = 0$.

For a symmetric interval X, we have $|X| = W(X)/2$ and $X = |X|[-1,1]$.
If X and Y are symmetric intervals, then
$X+Y = X-Y = (|X|+|Y|)[-1,1]$
$XY = |X||Y|[-1,1]$.
If X Y and Z are symmetric intervals, then
$X(Y \pm Z) = XY+XZ = |X|(|Y|+|Z|)[-1,1]$

An arbitrary interval X can be written as the sum of a real number and a symmetric interval. Thus $X = m+W$ where $m = m(X)$ and $W = \frac{1}{2}W(X)[-1,1]$ Put another way, we can write

$$\begin{aligned} X &= m(X) + \left[-\frac{1}{2}w(X), \frac{1}{2}w(X)\right] \\ &= m(X) + \frac{1}{2}w(X)[-1,1] \\ &= \frac{a+b}{2} + \frac{b-a}{2}[-1,1] \end{aligned}$$

If Y is a symmetric interval and X is any interval, then $XY = |X|Y$. It follows that $X(Y+Z) = XY+XZ$ if Y and Z are symmetric, for any interval X.

2.8 Inclusion Isotonicity of Interval Arithmetic

Let $\diamond$ denote one of the four arithmetic operators,$+,-,*,/$ on the interval numbers. If I,J,K and L are intervals such that $I \subseteq K$ and $J \subseteq L$. Then

$$I \diamond J \subseteq K \diamond L \tag{2.28}$$

These relations follow directly from the above definitions. However, these relations are not purely algebraic; they serve to connect the algebraic set properties of interval arithmetic. Interval arithmetic is said to be inclusion isotonic.

Finally we note that the set of interval numbers forms an Abelian semi-group under the operation of addition and and also under the operation of multiplication.

The following four lemmas will be needed later for established the Fundamental theorem of Interval Analysis. If I,J,K and L be interval numbers such that $I \subset K$ and $J \subset L$, let us prove the following lemmas.

Lemma 2.8.1 $I+J \subset K+L$

Proof. Let $I=[a,b]$, $J=[c,d]$, $K=[e,f]$ and $L=[g,h]$. Then from the definition of interval addition, $I+J=[a+c,b+d]$, and $K+L=[e+g,f+h]$. Since $I \subset K$ and $J \subset L$,then the elements of I,J,K and L can be ordered in the following way: $e \leq a$, $g \leq c$, $b \leq f$ and $d \leq h$. Combining these relations $e+g \leq a+c$ and $b+d \leq f+h$. Hence $I+J \subset K+L$.

Lemma 2.8.2 $I-J \subset K-L$.

Proof. Let I,J,K and L be same as in Lemma 5.1. Then from the definition of interval subtraction, $I-J=[a-d,b-c]$, and $k-L=[e-h,f-g]$. Combining the relations $e \leq a$ and $d \leq h$, gives $e-h \leq a-d$. Similarly, from $b \leq f$ and $g \leq c$, we get $b-c \leq f-g$.
Hence $I-J \subset K-L$.

Lemma 2.8.3 $I.J \subset K.L$.

Proof. Let I, J, K and L be same as in the above Lemma. Then from the definition of interval multiplication with the set notation $I.J = \{w.x | a \leq w \leq b, c \leq x \leq d$ and $K.L = \{y.z | e \leq y \leq f, g \leq z \leq h\}$. Then $e \leq a \leq b \leq f$ and $g \leq c \leq d \leq h$, and we note that every element,$w.x$ of the set $I.J$ is also an element, $y.z$ of the set $K.L$. Therefore $I.J \subset K.L$.

Lemma 2.8.4 $I/J \subset K/L$ *if* $0 \notin L$.

Proof. Let I, J, K and L be same as in the above Lemma. Again, with the set notation, $I/J = \{w/x | a \leq w \leq b, c \leq x \leq d\}$, and $K/L = \{y/z | e \leq y \leq f, g \leq z \leq h\}$.Then, using $e \leq a \leq b \leq f$ and $g \leq c \leq d \leq h$, we find that every element,w/x, of the set I/J is also an element, y/z, of the set K/L. Therefore, if $0 \notin L$, $I/J \subset K/L$.

Chapter 3

Sequences of Intervals and Interval Functions

3.1 Review of some basic definitions and Fundamental results

Definition 3.1.1 ***Convergent Sequence.*** *In real analysis, we call a sequence* $\{x_k\}$ *convergent if there exists a real number* x^* *such that for every* $\varepsilon > 0$ *there is a natural number* $N = N(\varepsilon)$ *such that*

$$|x_k - x^*| < \varepsilon \tag{3.1}$$

whenever $k > N$. *In this case, we write*

$$x^* = \lim_{k \to \infty} x_k \tag{3.2}$$

and refer to x^* *as the limit of* $\{x_k\}$. *The quantity* $|x_k - x^*|$ *in (3.1) is just the distance between* x_k *and* x^* *as measured along the real line.*

The notion of sequence convergence appears throughout mathematics. One can consider convergent sequences of complex numbers, real-valued functions of real variable, etc. All that is needed is a suitable measure of 'distance' between the objects of interest.

Definition 3.1.2 ***Continuity.*** *We say that $f(x)$ is continuous at a point x_0 if for every $\varepsilon > 0$ there is a positive number $\delta = \delta(\varepsilon)$ such that*

$$|f(x) - f(x_0)| < \varepsilon \tag{3.3}$$

whenever $|x - x_0| < \delta$. In this case, $|f(x) - f(x_0)|$ is the distance between the y-axis points $f(x)$ and $f(x_0)$, corresponding to the distance $|x - x_0|$ along the x-axis.

It is said that convergence and continuity are the two central concepts of analysis. We see, in turn, how they both hinge on having a suitable way to express distance. The need to discuss convergence and continuity outside of ordinary real analysis has led to a powerful generalization of the distance idea.

3.2 Norms of Vector and Matrix

Let us consider the linear space(or vector space) X over the field F, whose elements(vectors) denoted by $x, y, z, \cdots$

Definition 3.2.1 ***Inner Product.*** *Let x, y, z be any three vectors in X.The inner product of two vectors x and y in X defined by (x, y) is a scalar satisfying the following axioms :*

- $(x, x) \geq 0$; $(x, x) = 0$ *if and only if* $x = 0$ *(positive definitions)*
- *(x,y)=(y,x) (symmetric property)*

- *(ax+by,z)=a(x,z)+b(y,z) (linearity), where* $a, b \in F$

Let $X = \mathbb{R}^n$, *and* $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, *then the inner product* (x,x) *is given by* $(x,x) = (|x_1|^2 + |x_2|^2 + \cdots + |x_n|^n)$.
By the norm $||.||$ *of a vector* x *of* X, *we mean a function* $||x|| : X \to \{t : 0 \leq \infty\}$ *such that*

- $||x+y|| \leq ||x|| + ||y||$
- $||ax|| = |a|||x||$
- $||x|| = 0$ *if an only if* $x = 0$

Definition 3.2.2 ***Uniform norm.*** *In mathematical analysis, the* ***uniform norm****assigns to real or complex-valued bounded functions* f *the nonnegative number*

$$\|f\|_\infty = sup\{|f(x)| : x \in domain\ of \quad f\}$$

This norm is also called the ***supremum norm*** *or the* ***Chebyshev norm.***

If f *is a continuous function on a closed interval then it is bounded and the supremum in the above definition is attained by the Weierstrass extreme value theorem, so we can replace the supremum by the maximum. In this case, the norm is also called the* ***maximum norm.*** *For the case of a vector* $x = (x_1, x_2, \ldots, x_n)$ *in finite dimensional coordinate space, it takes the form*

$$\|x\|_\infty = max\{|x_1|, \ldots, |x_n|\}$$

The reason for the subscript "∞*" is that*

$$\lim_{p \to \infty} ||f||_p = ||f||_\infty,$$

where

$$\|f\|_p = \left(\int_D |f|^p d\mu \right)^{1/p}$$

where D is the domain of f The binary function

$$d(f,g) = ||f-g||_\infty$$

is then a metric on the space of all bounded functions on a particular domain. A sequence $\{f_n : n = 1,2,3,\ldots\}$ *converges uniformly to a function* f *if and only if*

$$\lim_{n\to\infty} ||f_n - f||_\infty = 0.$$

Definition 3.2.3 ***Normed linear space.*** *A linear space X with a norm*$||.||$ *defined on it is called a normed linear space.*

Matrix norm

In mathematics, a **matrix norm** is a natural extension of the notion of a vector norm to matrices. **Properties of matrix norm**

Let us denote by K the field of real or complex numbers. We consider the space $K^{m\times n}$ of all matrices with m rows and n columns with entries in K. A matrix norm on $K^{m\times n}$ satisfies all the properties of vector norms. That is if $||A||$ is the norm of the matrix A, then

- $||A|| \geq 0$ and $||A|| = 0$ if and only if $A = 0$
- $||\alpha A|| = |\alpha|||A||$ for all $\alpha \in K$ and all matrices A in $K^{m\times n}$
- $||A+B|| \leq ||A|| + ||B||$
- $||AB|| \leq ||A||||B||$ where B is another element in $K^{m\times n}$
- $||A|| = ||A^*||$ where A^* is the conjugate transpose of A.

Induced norm

If vector norms on K^m and K^n are given, then one defines the corresponding induced norm or operator norm on the space of m-by-n matrices as the following

maxima:

$$\begin{aligned}||A|| &= \max\{||Ax|| : x \in K^n \text{with} ||X|| \leq 1\} \\ &= \max\{||Ax|| : x \in K^n \text{with} ||X|| = 1\} \\ &= \max\left\{\frac{||Ax||}{||x||} : x \in K^n \text{with} ||X|| \neq 0\right\}\end{aligned}$$

There are numerous ways by which matrix norms can be formed. The following are matrix norms on $K^{m\times n}$

$$(1) \quad ||A||_p = \max_{x\neq o} \frac{||Ax||_p}{||x||_p}$$

$$(2) \quad ||A||_1 = \max_{1\leq j\leq n} \sum_{i=1}^{m} |a_{ij}|$$

$$(3) \quad ||A||_\infty = \max_{1\leq i\leq m} \sum_{j=1}^{n} |a_{ij}|$$

"Entrywise" norms

$$(4) \quad ||A||_p = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^p\right)^{\frac{1}{p}}$$

Frobenius norm

$$(5) \quad ||A||_F^2 = \sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2 = trace(A^*A)$$

Trace norm

$$(6). \quad ||A||_{tr} = trace\left(\sqrt{A^*A}\right)$$

Max norm

$$(7). \quad ||A||_{max} = max\{|a_{ij}|\}$$

Definition 3.2.4 ***Linear operator.*** *Suppose we are given two linear spaces X and Y over the same scalar field F. An operator is a mapping P which map X into Y over the same field F such that for each $x \in X$ there is a uniquely defined $P(x) \in Y$.*

The mapping P is said to be linear if it satisfy two conditions:

1. $x_1 \to y_1$, $x_2 \to y_2$ *implies* $x_1 + x_2 \to y_1 + y_2$

2. $x \to y$ *implies* $\alpha.x \to \alpha.y$ *where* $x_1, x_2 \in X$, $y_1, y_2 \in Y$ *and* αF.

.

Definition 3.2.5 ***Non-linear operator.*** *An operator P from a linear space X into a linear space Y is said to be non-linear if it is not a linear operator from X into Y. A simple non-linear operator is one that gives, for all* $x \in X$. $P(x) = y_0$ *where* y_0 *is a fixed, nonzero elements.*

Definition 3.2.6 ***Inverse operator.*** *Let P be an operator defined on a vector subspace of X. An operator A defined on the ranges of P, R(P) is called the inverse of P if*

$$PAx = x \quad for\ all \quad x \in R(P)$$

and

$$APx = x \quad for\ all \quad x \in D(P), domain\ of \quad P.$$

Theorem 3.2.1 *If an operator has an inverse then it is unique.*

Theorem 3.2.2 *If A is a linear mapping from X into Y the* A^{-1} *exists if and only if* $Ax = 0$ *implies* $x = 0$.

Lemma 3.2.3 *Banach Lemma Suppose L is a bounded linear operator in X,* L^{-1} *exists if and only if there is a bounded linear operator M in X such that* M^{-1} *exists, and*

$$||M - L|| < \frac{1}{||M^{-1}||}.$$

If L^{-1} exists then

$$L^{-1} = \sum_{n=0}^{\infty} (1 - M^{-1}L)^n M^{-1}$$

and

$$||L^{-1}|| \leq \frac{||M^{-1}||}{1 - ||1 - M^{-1}L||} \leq \frac{||M^{-1}||}{1 - ||M^{-1}||||M - L||}$$

Definition 3.2.7 *The real number field $\mathbb{R}$ is itself a one-dimensional vector space over itself. Then any mapping of $(X,Y) \Rightarrow (F,F)$ where X is a normed space, is called a functional. If the mapping is linear, it is called linear functional.*

Definition 3.2.8 ***Bounded operators.*** *The concept of a bounded operator is closely connected with that of a continuous operator. Let U and V be two normed spaces and let $T : U \to V$ be a linear operator. We say that T is bounded if it is possible to find a number $K > 0$ such that*

$$||Tu|| \leq K||u|| \quad \text{for all} \quad u \in U$$

For all bounded linear operator T,
we have $||Tu|| \leq ||T||||u||$, where $||T|| = sup\{||T||/||u||, u \neq 0\}$.

We have the following theorem connecting the boundedness and continuity of operators:

Theorem 3.2.4 *A linear operator T from a normed space U to a normed space V is continuous if and only if it is bounded.*

An important class of normed linear space which is named after Stefen Banach (1892-1945), plays an important rule in the existence of the limit x^* of an infinite sequence $\{x_m\}$ of elements of normed linear space X. Consider the sequence $\{x_m\}$ of rational numbers defined by

$$x_0 = 1, x_m = \frac{1}{2}(x_{m-1} + \frac{2}{x_{m-1}}), \quad m = 1, 2, \ldots$$

There is no rational number x^* which can be the limit of this sequence.(with $||x|| = |x|$). However if the above sequence $\{x_m\}$ in $\mathbb{R}$, it has a limit x^* which is the solution of the nonlinear equation $x^2 = 2$.

Consequently , the space of real numbers has a property with respect to limits which the set of rational number does not. This property is defined precisely in the more abstract setting in a normed linear space by the following fundamental definition.

Definition 3.2.9 ***Cauchy or Fundamental sequence.*** *A sequence $\{x_n\}$ of elements of a normed linear space is called a Cauchy sequence if for every $\varepsilon > 0$ there exist a number λ such that*

$$||x_m - x_n|| < \varepsilon \quad \textit{for all} \quad m, n > \lambda.$$

Theorem 3.2.5 *If $\{x_n\}$ is a Cauchy sequence in a normed linear space, then the sequence of norms $\{||x_n||\}$ converges*

Theorem 3.2.6 *Let $(E, ||)$ be a normed space. The function d defined by*

$$d(x,y) = ||x - y||$$

is a distance on E satisfying the following conditions:

1. $d(x+z, y+z) = d(x,y)$ *(invariance by translation).*
2. $d(\lambda x, \lambda y) = |\lambda| d(x,y)$

According to theorem 1 every normed space is a metric space. The structure of a normed space combining the structures of a vector space and of a metric space, has a large role both in mathematical theory and in applications. Complete normed space, in particular, play a very important role.

Definition 3.2.10 ***Banach space.*** *Banach spaces are defined as complete normed linear spaces.This means that a Banach space is a linear space V over the real or complex number with a norm $\|.\|$ such that every Cauchy sequence in V has a limit in V.*

Theorem 3.2.7 *A Euclidean space $\mathbb{R}$ becomes a normed linear space when equipped with the norm*

$$||x|| = \sqrt{(x,x)}, \quad (x \in \mathbb{R})$$

3.3 Balls in normed linear space.

Let X be a normed linear space. The open and closed balls in X with the center $x_0 \in X$ and radius r can be defined respectively by the subsets

$$B(x_0,r) = \{x : x \in X \text{and} \quad ||x - x_0|| < r\}$$

and

$$B(x_0,r) = \{x : x \in X \text{and} ||x - x_0|| \leq r\}.$$

Example. In the case of $\mathbb{R}^n$, consider the balls

$$B_p(x,\varepsilon) = \{y \in \mathbb{R}^n \quad \text{such that} ||x - y|| \leq \varepsilon\}.$$

In the case where $n = 2$, $x = 0$, $\varepsilon = 1$, the balls $B_p(0,1)$ with center 0 and radius 1 are defined by

$$B_1(0,1) = \{y = (y_1,y_2) \quad \text{such that} \quad |y_1| + |y_2| \leq 1\}$$
$$B_2(0,1) = \{y = (y_1,y_2) \quad \text{such that} \quad |y_1|^2 + ||y_2|^2 \leq 1\}$$
$$B_\infty(0,1) = \{y = (y_1,y_2) \quad \text{such that} \quad (|y_1||y_2|) \leq 1\}$$

A subset of a normed linear space X is said to be bounded if it is contained in the same ball of finite radius.

Definition 3.3.1 ***Gateaux derivatives.*** *Suppose X and Y be two Banach spaces over the field F, and consider the operator $T : X \to Y$ with domain $D_T = X$. Suppose x is a fixed point of X. The operator $T : X \to Y$ is said to be Gateaux differentiable at x if there exists a continuous linear operator L such that*

$$\lim_{t \to 0} \frac{||T(x+th) - T(x) - L(h)||}{||t||} = 0$$

for every $h \in X$, where $t \to 0$ in F. The operator L is called the Gateaux derivative of T and x, and its value at h is denoted by

$$A(h) = dT(x,h)$$

The notation $dT(x,h)$ or $T'(x)h$ is also used.

Definition 3.3.2 ***Frechet derivatives.*** *Let x be a fixed point in a Banach space X. A continuous linear operator $A : X \to Y$ is called the Frechet derivative of the operator $T : X \to Y$ at x if*

$$T(x+h) - T(x) = Ah + \varphi(x,h).$$

provided

$$\lim_{||h|| \to 0} \frac{||\varphi(x,h)||}{||h||} = 0$$

or equivalently

$$\lim_{||h|| \to 0} \frac{||T(x,h) - T(x) - Ah||}{||h||} = 0$$

The Frechet derivative at x is usually denoted by $T'(x)$ or $dT(x)$.

We will now sate a theorem without proof that relates these two types of derivatives.

Theorem 3.3.1 *If a mapping has the Frechet derivative at a point x, then it has Gateaux derivative at that point and both derivatives are equal.*

Corollary 3.3.2 *If the Frechet derivative exists, it is unique.*

Example. If $f:\mathbb{R}^2 \to \mathbb{R}$ defined by $f(x,y) = \begin{cases} \frac{x^3y}{x^4+y^2} & \text{if } x \neq 0 \text{ and } y \neq 0 \\ 0 & \text{if } x = y = 0 \end{cases}$

It is easy to check that f is Gateaux differentiable at 0,and the Gateaux derivative at that point is 0. On the other hand ,since

$$\frac{|f(x,x^2)|}{||(x,x^2)||} = \frac{|x^3x^2|}{(x^4+x^4)} \frac{1}{\sqrt{(x^2+x^4)}} = \frac{1}{2\sqrt{(1+x^2)}} \to \frac{1}{2} asx \to 0$$

f is not Frechet differentiable at $(0,0)$. The above example suggests the following theorem.

Theorem 3.3.3 *The existence of Frechet derivative implies the Gateaux derivative but the converse is generally false.*

3.4 Metric space

A metric space is an ordered pair$(M,d)$where M is an non-empty set, and suppose that a real-valued function d is defined such that for any two elements $x,y \in M$ the following statements hold:

1. $d(x,y) \geq 0$ (non-negativity),
2. $d(x,y) = 0$ if and only if $x = y$
3. $d(x,y) = d(y,x)$ (symmetry),
4. $d(x,z) \leq d(x,y) + d(y,x), z \in M$ (Triangle inequality),

The first condition follows from the other three, since

$$2d(x,y) = d(x,y) + d(y,x) \geq d(x,x) = 0$$

These can be regarded as the essential characteristics of distance between the objects x and y, and they certainly hold in the real number system when $x, y \in \mathbb{R}$ and $d(x,y) = |x-y|$. The function d is called a metric on M, and (M,d) is known as a metric space.

Definition 3.4.1 *Cauchy sequences in metric space. Let (M,d)be a metric space, a sequence $x_1, x_2, x_3, \cdots$ is Cauchy, if for every positive real number $\varepsilon > 0$ there is a positive integer N such that for all natural numbers $m, n > N$, the distance $d(x_m, x_n) < \varepsilon$.*

Definition 3.4.2 Convergent sequence in a Metric space. *Let (M,d) be a metric space. A sequence $\{x_n\}$ of elements $x_n \in M$* ***converges to*** *x if for all $\varepsilon > 0$, there exists an $n_0 \in N$ such that for $n \geq n_0$, $\quad d(x_n, x) \leq \varepsilon$.*

Definition 3.4.3 Uniform Convergence. *Suppose X is a set and $f_n : X \to \mathbb{R}$ are real-valued functions for every natural number n. We say that the sequence $\{f_n\}$ is uniformly convergent with limit $f : X \to \mathbb{R}$ if for every $\varepsilon > 0$, there exists a natural number N such that for all $x \in X$ and all $n = N$, $|f_n(x) - f(x)| < \varepsilon$.*

Consider the sequence $a_n = sup|f_n(x) - f(x)|$ where the supremum is taken over all $x \in X$. Clearly $\{f_n\}$ converges to f uniformly if and only if a_n goes to 0.

The sequence $\{f_n\}$ is said to be locally uniformly convergent with limit f if for every x in some metric space X, there exists an $r > 0$ such that $\{f_n\}$ converges uniformly on $B(x,r) \cap S$

Definition 3.4.4 Uniform Continuous function in metric space. *Let (M,d) and (N,d) be two matric spaces and f be a function such that $f : (M,d) \mapsto (N,d)$. We called f is* ***uniformly continuous*** *if for all $\varepsilon > 0$, there exist $\eta = \eta(\varepsilon)$ depending on ε and independent of x such that $d_N(f(x), f(y)) \leq \varepsilon$ when $d_M(x,y) \leq \eta$.*

Proposition 3.4.1 *1. Every uniformly continuous function is continuous.*

2. *Every uniformly continuous function maps Cauchy sequences onto Cauchy sequences.*

Proposition 3.4.2 *If M is a metric space and if $A \subset M$ is nonempty, the function $x \to d(x,A)$ is uniformly continuous from M to $\mathbb{R}$.*

Proof. *The proposition is a consequence of the inequality:*

$$|d(x,A) - d(y,A)| \leq d(x,y)$$

This suggests the introduction of the following definition.

Definition 3.4.5 *We say that a function f from a metric space M to a metric space N is* **Lipschitz** *if there exists a constant $k > 0$ such that*

$$d_N(f(x), f(y)) \leq k d_M(x,y) \text{for all} \quad x,y \in M.$$

We say that a function f is a contraction if in addition $k < 1$. For example, the function $x \to d(x,A)$ is Lipschitz with $k = 1$

We remark from this definition that the following proposition holds.

Proposition 3.4.3 *Every Lipschitz function is uniformly continuous.*

Balls in general metric spaces *In mathematics, a ball is the space inside a sphere. It may be either a closed ball (including the boundary points) or an open ball (excluding them).*

These concepts are defined not only in three-dimensional Euclidean space but also for lower and higher dimensions, and for metric spaces in general. A ball in the Euclidean plane, for example, is the same thing as a disk, the area bounded by a circle.In mathematical contexts where ball is used, a sphere is usually assumed to be the boundary points only.

Now we consider a metric space (M,d).***The open(metric) ball with center*** x ***and radius*** ε *is the set*

$$B(x,\varepsilon) = \{y \in M \quad such\ that \quad d(x,y) < \varepsilon\},$$

and ***the closed ball with center*** x ***and radius*** ε *is the set*

$$\overline{B}(x,\varepsilon) = \{y \in M \quad such\ that \quad d(x,y) \leq \varepsilon\}.$$

Definition 3.4.6 Complete metric space. *In mathematical analysis, a metric space* M *is called* **complete** *if every Cauchy sequence of points in* M *has a limit that is also in* M *or alternatively if every Cauchy sequence in* M *converges in* M. *However, we must discuss continuity and convergence in the context of interval analysis., and for this we require a suitable metric.*

3.5 A metric for the Set of Intervals

Definition 3.5.1 *We have to organize metric topology for the set of real intervals and the set of real interval vectors as metric spaces*[14].
For $X,Y \in \mathbb{IR}$) *we define the distance between the two intervals* $X = [a,b]$ *and* $Y = [c,d]$ *as the nonnegative function*

$$d(X,Y) = max(|a-c|,|b-d|) \tag{3.4}$$

It is easy to prove that the defining properties of a metric are satisfied by $d(X,Y)$:

$$d(X,Y) = 0 \quad if\ and\ only\ if \quad X = Y$$
$$d(X,Y) \geq 0$$
$$d(X,Y) = d(Y,X)$$
$$d(X,Z) \leq d(X,Y) + d(Y,Z) \quad for\ any Z \in \mathbb{IR}$$

For degenerate interval $[a,a]$ *and* $[b,b]$ *we have*

$$d([a,a],[b,b]) = |a-b| \tag{3.5}$$

which is the usual topology on the real line.

For interval vectors $X = (X_1, X_2 ..... X_n)$ *and* $Y = (Y_1, Y_2 ..... Y_n)$, *we define the distance as*

$$d(X,Y) = \max_i d(X_i, Y_i) \tag{3.6}$$

Thus, also $\mathbb{IR}^n$ *is organized as a metric space.*

Definition 3.5.2 Convergence and Continuity in Interval Mathematics

A sequence of intervals $\{X_k\}$ *in* $\mathbb{IR}^n$ *is said to converge to* X *if* $d(X_k, X) \to 0$ *for* $k \to \infty$. *From the definition (3.6) it follows that this means component-wise,convergence, i.e,* $d(X_{(k)i}, X_i) \to 0$ *and from the definition (3.4) of the scalar distance this means that* $a_{(k)i} \to a_i$ *and* $b_{(k)i} \to b_i$ *i.e, the end points converge to the end points of the limit interval.*

An interval function $F : \mathbb{IR}^n \mapsto \mathbb{IR}^m$ *is continuous for* $X \subseteq A \in \mathbb{IR}^n$ *if*

$$\lim_{k\to\infty} F(X_k) = F(X) \quad for \quad \lim_{k\to\infty} X_k = X \tag{3.7}$$

Definition 3.5.3 ***Lipschitz Continuity.*** *Now we look at a presumption on interval functions which is stronger than continuity,viz.* Liplschitz continuity. *An ordinary function* $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ *defined on an interval* A *is said to be Lipschitz continuous if there exists a number* $K > 0$ *such that*

$$d(f(x), f(y)) \leq K d(x,y) \quad for x, y \in A$$

This definition is transferred literally to interval functions: An interval function F *from* $X \in \mathbb{IR}^n$ *into* $\mathbb{I}(\mathbb{R}^m)$ *is said to be Lipschitz continuous if there exists a number* $K > 0$ *such that*

$$d(F(X), F(Y)) \leq K d(X,Y) \quad for \quad X, Y \subseteq A.$$

Definition 3.5.4 ***Isometric Embedding.*** *We know that the interval number system represents an extension of the real number system. In fact, the correspondence* $[x,x] \leftrightarrow x$ *can be regarded as a function or mapping between the two systems. This mapping preserves distances between corresponding objects: we have*

$$d([x,x],[y,y]) = max\{|x-y|,|x-y|\} = |x-y|$$

for any x and y. For this reason, it is called an isometry, and we say that the real line is "isometrically embedded" in the metric space of intervals.

3.6 Width of an interval

The final basic metric concept to consider is the width of an interval(vector).This was defined in Chapter 2 $w(X) = \begin{cases} b-a & \text{if } X \in \mathbb{IR} \\ max_i\{w(X_i)\} & \text{if } X \in \mathbb{IR}^n \end{cases}$
For $X \subseteq Y$ there is the following important relation between the width and distance

$$\frac{1}{2}(w(Y)-w(X)) \leq d(X,Y) \leq w(Y)-w(X) \tag{3.8}$$

Further,for $X,Y,Z,W \in \mathbb{IR}$, $a \in \mathbb{R}$ we have the following relations for the width

$$w(X+Y) = w(X)+w(Y)$$
$$w(X-Y) = w(X)-w(Y)$$
$$w(aX) = ||a||w(X)$$

$\max\{||Y||w(X),||X||w(Y)\} \leq w(X*Y) \leq ||Y||w(X)+||X||w(Y)$
$w(X/Y) \leq \frac{1}{cd}(||Y||w(X)+||X||w(Y))$ and for the distance

$$d(X+Z,Y+Z) = d(X,Y)$$
$$d(X+Y,Z+W) \leq d(X+Z)+d(Y,W)$$
$$d(X-Y,Z-W) \leq d(X,Z)+d(Y,W)$$
$$d(X*Y,Z*W) \leq ||X||d(Y,Z)+||Z||d(X,W)$$
$$d(\frac{1}{X},\frac{1}{Y}) \leq \frac{1}{||X||||Y||}d(X,Y)$$

Lemma 3.6.1 *Let $X,Y \in \mathbb{IR}$.Then*

$$d(X,Y) = max\left\{\max_{a\in X}\{\min_{b\in Y}|a-b|\}, \max_{b\in Y}\{\min_{a\in X}|a-b|\}\right\}$$

3.7 Set Images and United extension

Let f be a real-valued function of a single real variable x. Ultimately, we would like to know the precise range of values taken by $f(x)$ as x varies through a given interval X. In other words, we would like to be able to find the image of the set X under the mapping f

$$f(X) = \{f(x) : x \in X\} \tag{3.9}$$

More generally,, we are given a function

$$f = f(x_1,x_2,\ldots,x_n)$$

of several variables, we will wish to find the image set

$$f(X_1,X_2,\ldots,X_n) = \{f(x_1,x_2,\ldots,x_n) : x_1 \in X_1,\cdots,x_n \in X_n\} \tag{3.10}$$

where$X_1,X_2,\ldots,X_n$ are specified intervals.

Definition 3.7.1 *Let $g : M_1 \to M_2$ be a mapping between sets M_1 and M_2, and denote by $S(M_1)$ and $S(M_2)$ the families of subsets of M_1 and M_2, respectively. The united extension of g is the set- valued mapping $\overline{g} : S(M_1) \to S(M_2)$ such that*

$$\overline{g}(X) = \{g(x) : x \in X, X \in S(M_1)\} \tag{3.11}$$

The mapping $\overline{g}$ is sometimes of interest as a single-valued mapping on $S(M_1)$ with values in $S(M_2)$. For our purposes, however,it is merely necessary to note that

$$\overline{g}(X) = \bigcup_{x \in X} \{g(x)\} \tag{3.12}$$

i.e., that $\overline{g}(X)$ contain precisely the same elements as the set image $g(X)$.

3.8 Elementary Functions of Interval Arguments

Interval methods can also apply to functions which do not just use simple arithmetic, and we must also use other basic functions for redefining intervals, using already known monotonicity properties.

For monotonic function of one variable, the range of values is also easy, If $f : \mathbb{R} \to \mathbb{R}$ is monotonically rising or falling in the interval $X = [a,b]$ then for all values in the interval $a_1, b_1 \in [a,b]$ such that $a_1 \leq b_1$, one of the following inequalities applies:

$$f(a_1) \leq f(b_1), \quad \text{or} f(a_1) \geq f(b_1)$$

The range corresponding to the interval $[a_1, b_1] \subseteq [a,b]$ can be calculated by applying the function to the endpoints a_1 and b_1.

$$f([a_1, b_1]) = [\min\{f(a_1), f(b_1)\}, \max\{f(a_1), f(b_1)\}]$$

From this the following basic features for interval functions can easily be defined:

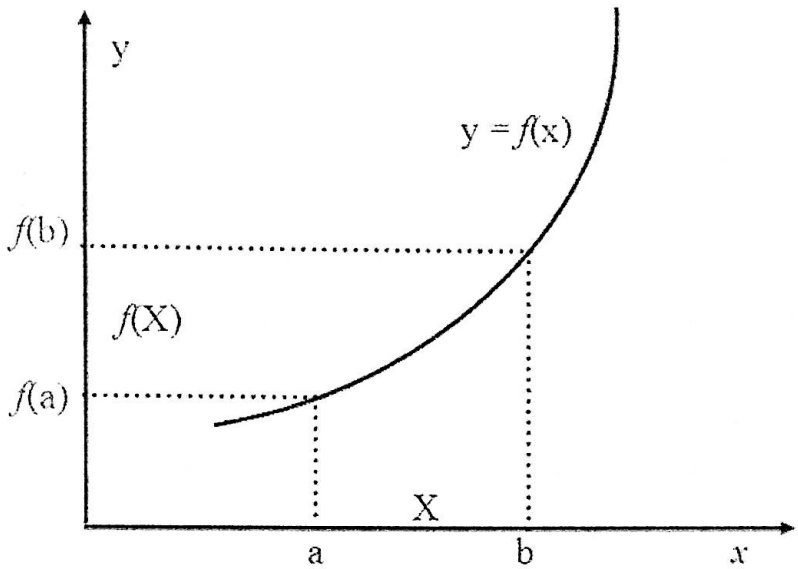

Figure 3.1: The image of an interval X under an increasing function $f(x)$.

- Exponential function : $e^{[a,b]} = [e^a, e^b]$
- Logarithm: log([a,b])=[loga,logb]) for positive intervals $[a,b]$
- Odd powers: $[a,b]^n = [a^n, b^n]$ for odd $n \in \mathbb{N}$

For even powers, the range of values being considered is important, and needs to be dealt with before doing any multiplication. For example x^n for $x \in [-1,1]$ should produce the interval $[0,1]$ when $n = 2,4,6,\cdots$. But if $[-1,1]$ is taken by applying interval multiplication of the form $[-1,1]\cdots[-1,1]$ then the result will appear to be $[-1,1]$, wider than necessary.

Instead consider the function x^n as a monotonically decreasing function for $x < 0$ and a monotonically increasing function for $x > 0$. So for even $n \in \mathbb{N}$:

- $[a,b]^n = [a^n, b^n]$,if $a \geq 0$

- $[a,b]^n = [b^n, a^n]$, if $b < 0$

- $[a,b]^n = [0, \max\{a^n, b^n\}]$, otherwise

3.9 Interval-Valued Extensions of Real Functions

We were able to define a few interval-valued functions by selecting a real-valued function f and computing the range of values $f(x)$ taken as x varied through some interval X. By definition, the result is equal to the set image $f(X)$. Here we will consider a different process: that of extending a given real-valued function f by applying its formula directly to interval arguments.

3.10 Formulas and Interval Extensions

We will begin with an example. We consider the real-valued function f given by

$$f(x) = 1 - x, \quad x \in \mathbb{R} \tag{3.13}$$

We have seen that a function is defined by two things:(1) a domain over which it acts, and (2) a rule that specifies how elements of that domain are mapped under the function. Both of these are specified in the above equation(8):the elements of Dom f are real numbers x, and the mapping rule is $x \mapsto 1 - x$. Taken in isolation, the entity

$$f(x) = 1 - x \tag{3.14}$$

is a formula-not a function.

Now suppose we take the formula(9) that describes function (8) and apply it to interval arguments. The resulting interval-valued function

$$F(X) = 1 - X, \quad X = [a,b], \tag{3.15}$$

is an extension of the function (8): we have enlarged the domain to include non degenerate interval X as well as the degenerate intervals $x = [x,x]$.

Definition 3.10.1 *We say that F is an interval extension of f, if for degenerate interval arguments, F agrees with f:*

$$F([x,x]) = f(x), \tag{3.16}$$

Let us compare $F(X)$ with the set image $f(X)$. We have, according to the laws of interval arithmetic,

$$F(X) = [1,1] - [a,b] = [1-b, 1-a].$$

On the other hand, x increases through the interval $[a,b]$, the values $f(x)$ given by (*) clearl decrease from $1-a$ to $1-b$; by definition then,

$$f(X) = [1-b, 1-a].$$

In this example we have, $F(X) = f(X)$: this particular extension of f, obtained by applying formula (9) directly to interval arguments, yields the desired set image (8). In other words, we have found the united extension of f:

$$f(X) = 1 - X$$

3.11 A Numerical Example

Consider the two real-valued function

$$f(x) = x(1-x), \quad x \in [0,1]. \tag{3.17}$$

and

$$g(x) = x - x^2, \quad x \in [0,1]. \tag{3.18}$$

These are mathematically equal, because in ordinary real arithmetic we have

$$x(1-x) = x - x^2$$

As x increases from 0 to 1, the values $f(x)$ and $g(x)$ both increases from 0 to $\frac{1}{4}$ then decrease back to 0. Therefore,

$$f([0,1]) = g([0,1]) = \left[0, \frac{1}{4}\right]$$

Let us form interval-valued extensions of f and g:

$$F(X) = X.(1-X), \quad X = [a,b] \tag{3.19}$$

and

$$G(X) = X - X^2, \quad X = [a,b] \tag{3.20}$$

Remembering that $X^2 \neq X.X$ in interval arithmetic, we work out the details of each extension separately for an interval $X = [a,b] \subseteq [0,1]$:

$$\begin{aligned} F(X) &= [a,b].([1,1] + [a,b]) \\ &= [a,b].([1,1] + [-b,-a]) \\ &= [a,b]([1-b, 1-a] \\ &= [\text{minS}, \text{maxS}], \end{aligned}$$

Where $S = \{a(1-b), a(1-a), b(1-b), b(1-a)\}$,while

$$\begin{aligned} G(X) &= [a,b] - [a,b]^2 \\ &= [a,b] - [a^2, b^2] \\ &= [a,b] + [-b^2, -a^2] \\ &= [a-b^2, b-a^2] \end{aligned}$$

Putting, say $X = [0,1]$, we see that $G(X) = F(X)$; the two formulas involved (*) and (*)-again, identical in ordinary arithmetic-give rise to different extensions,and neither of these maps the interval$[0,1]$ into the interval$[0,\frac{1}{4}]$. We have $F([0,1]) = [0,1]$ and $G([0,1]) = [-1,1]$

We stress that two expressions can be equivalent in real arithmetic but no equivalent in interval arithmetic. This is due to the lack of distributivity and additive and multiplicative inverses in interval arithmetic.

It turns out that the united extension of the original function f arises from use of a third equivalent formula:

$$h(x) = \frac{1}{4} - \left(x - \frac{1}{2}\right)^2.$$

We get

$$\begin{aligned} H(X) &= \frac{1}{4} - \left(X - \frac{1}{2}\right)^2 \\ &= \left[\frac{1}{4}, \frac{1}{4}\right] - \left([a,b] - \left[\frac{1}{2}, \frac{1}{2}\right]\right)^2 \\ &= \left[\frac{1}{4}, \frac{1}{4}\right] - \left[a - \frac{1}{2}, b - \frac{1}{2}\right]^2 \end{aligned}$$

It is easily verified that this is $f(X)$. We have $H([0,1]) = [0,\frac{1}{4}]$.

Subset Property of United Extension

We see that the united extension $\overline{g}$ has the following subset property:

$$X, Y \in S(M_1) \quad \text{with} \quad X \subseteq Y \Rightarrow \overline{g}(X) \subseteq \overline{g}(Y).$$

3.12 Interval Extensions of Multivariable Functions

So far, we have limited ourselves to functions of a single interval variable X. In principle, there is no reason we should avoid more general function

$$f = f(X_1, X_2, \ldots, X_n)$$

depending on n interval variables $X_1, \ldots, X_n$

Definition 3.12.1 *By an interval extension of f, we mean an interval-valued function F of n interval variables $X_1, \ldots, X_n$ such that for real arguments $x_1, \ldots, x_n$ we have*

$$F(x_1, \ldots, x_n) = f(x_1, x_2, \ldots, x_n) \tag{3.21}$$

That is, if the arguments of F are degenerate intervals, then the left-hand side of the above equation is a degenerate interval equal to the right-hand side[17].

Inclusion Isotonicity

Definition 3.12.2 ***Inclusion Isotonicity.*** *We say that $F = F(X_1, \ldots, X_n)$ is inclusion isotonic if*

$$Y_i \subseteq X_i \quad \text{for } i=1,\ldots n \Rightarrow F(Y_1, \ldots, Y_n) \subseteq F(X_1, \ldots, X_n)$$

3.13 The Fundamental Theorem

Theorem 3.13.1 *Fundamental Theorem of Interval Analysis). If F is an inclusion isotonic interval extension of f, then*

$$f(X_1, X_2, \ldots, X_n) \subseteq F(X_1, \ldots, X_n).$$

Definition 3.13.1 ***Lipschitz Interval Extensions.*** *An interval extension F is said to be Lipschitz in X_0 if there a constant L such that $w(F(X)) \leq Lw(X)$ for every $X \subseteq X_0$.*

Hence, the width of $F(X)$ approaches zero at least linearly with the width of X. Here X may be an interval or an interval vector $X = (X_1, X_2, \ldots, X_n)$.

3.14 Linear Interval Extensions

If F is an interval extension of f on an interval $A \subseteq \mathbb{IR}^n$ and there exists a $K > 0$, independent of X, so that

$$d(F(X)f(X)) \leq Kw(X) \text{for all} X \subseteq A$$

then we say that F is a linear interval extension. It is obvious that a linear interval extension satisfies the convergence criterion for $w(X) \to 0$.

3.15 Mean Value Forms

We have seen one general method for constructing interval extensions,viz. interval extensions for rational functions. Now, we shall see how the mean value theorem can be used to construct interval extensions for a much wider class of functions, the differentiable functions.

We consider $f : \mathbb{R}^n \mapsto \mathbb{R}$, which is differentiable in $A \subseteq \mathbb{R}^n$. For such a function the mean value theorem says

$$f(x) = f(\widehat{x}) + \sum_{i=1}^{n} \frac{\delta f}{\delta x_i}(\xi)(x_i - \widehat{x_i}),$$

with $\widehat{x}, x \in A$ and ξ is a point on the line segment between $\widehat{x}$ and x. Let F' be the interval extension of the gradient

$$f' = \left(\frac{\delta f}{\delta x_1}, \cdots \frac{\delta f}{\delta x_n}\right),$$

and let $X \subseteq A$ be an interval that contains both $\widehat{x}$ and x.Then

$$f(x) \in f(\widehat{x}) + F'(\widehat{x}).(X - \widehat{x})$$

where the multiplication in the last term is the inner product of two interval vectors. In this reformulation we got rid of ξ, and the expression only involves intervals and interval operations. In order for this to be an extension of f we must accompany the interval X with a specification of the choice of $\widehat{x}$.

If we choose $\widehat{x} = m(X)$, the midpoint of X , we say that the resulting interval function

$$F_M(X) = f(m(X)) + F'(X).(X - m(X)) \tag{3.22}$$

is a mean value form.

3.16 Quadratic Interval Extensions.

Theorem 3.16.1 *If F is an interval extension of f on an interval $A \subseteq \mathbb{IR}^n$ and there exists a $K > 0$, independent of X, so that*

$$d(F(X), \overline{f}(X)) \leq KW(X)^2 \quad \textit{for all} \quad X \subseteq A, \tag{3.23}$$

then we say that F is a quadratic interval extension.

3.17 Finite Convergence and Stopping Criteria

Definition 3.17.1 *An interval sequence $\{X_k\}$ is nested if $X_{k+1} \subseteq X_k$ for all k.*

we state the following important lemma without proof:

Lemma 3.17.1 *Every nested sequence $\{X_k\}$ converges and has the limit $\cap_1^\infty X_k$.*

Lemma 3.17.2 *Suppose* $\{X_k\}$ *is such that there is a real number* $x \in X_k$ *for all* k. *Define* $\{Y_k\}$ *by* $Y_1 = X_1$ *and* $Y_{k+1} = X_{k+1} \cap Y_k$ *for* $k = 1, 2, \ldots$. *Then* Y_k *is nested with limit* Y, *and*

$$x \in Y \subseteq Y_k \quad forall \quad k \tag{3.24}$$

Definition 3.17.2 ***Finite Convergence.*** *By the finite convergence of a sequence* $\{X_k\}$, *we mean there is a positive integer* K *such that* $X_k = X_K$ *for all* $k \geq K$. *Such a sequence converges in* K *steps.*

Example. It is not hard to see that $X_0 = [1,2]$, $X_{k+1} = 1 + X_k/3$ $(k = 0, 1, 2, \ldots)$ generates a nested sequence $\{X_k\}$. The rational interval function $F(X) = 1 + X/3$ is inclusion isotonic. Therefore

$$X_1 = F(X_0) = 1 + [1,2]/3 = [\frac{4}{3}, \frac{5}{3}] \subseteq X_0 = [1,2].$$

It follows that $X_{k+1} = F(X_k) \subseteq X_k$ for all k by finite induction. By the above lemma, the sequence has a limit X. If we compute $\{X_k\}$ using interval arithmetic, we will obtain a sequence $\{X_k^*\}$ with $X_k \subseteq \{X_k^*\}$ for all k. More precisely , let X_k^* be defined by $X_0^* = X_0 = [1,2]$ then we will obtain

$$X_1^* = [1.33, 1.67]$$

$$X_2^* = [1.44, 1.56]$$

$$X_3^* = [1.48, 1.52]$$

$$X_4^* = [1.49, 1.51]$$

$$X_5^* = [1.49, 1.51]$$

and $X_k^* = X^*$ for all $k \geq 4$. We have finite convergence in four steps.

Natural Stopping Criterion

For any fixed precision representation of machine numbers, there is a finite set of machine numbers represented by strings of bits $b_0, b_1, \ldots b_s$ with s fixed. Hence

there is only a finite set of intervals with machine number endpoints. *Any nested sequence of such intervals is necessarily finitely convergent.*

For any iterative interval method that produces a nested sequence with endpoints represented by fixed precision machine numbers, we have a natural stopping criterion. Since the sequence $\{X_k\}$ converges in a finite number of steps, we can compute the X_k until

$$X_{k+1} = X_k \tag{3.25}$$

If the X_k are generated by a procedure of the form

$$X_{k+1} = F(X_k) \tag{3.26}$$

such that X_{k+1} depends only on the previous X_k then it is clear that the equation (1.25) is sufficient to guarantee convergence.

In particular, if $F(X)$ is a rational expression in X and if X_0 is an interval such that $F(X_0) \subseteq X_0$, which can be tested on the computer, it follows that $\{X_k\}$ defined by

$$X_{k+1} = F(X_k) \quad (k = 0, 1, 2, 3 \ldots.) \tag{3.27}$$

is nested with

$$X_0 \supseteq X_1 \supseteq X_2 \supseteq \cdots.$$

and hence converges to some X^* with $X^* = F(X^*)$ and $X^* \subseteq X_k$
for all $k = 0, 1, 2, \ldots..$

With interval arithmetic, it may happen that $X_1 = F(X_0) \subseteq X_0$ but that $X_{k+1} \not\subseteq X_k$ for some k. If instead of (1.27), we may compute

$$X_{k+1} = F(X_k) \cap X_k \tag{3.28}$$

and stop when (1.25) is satisfied.

Chapter 4

Convergence and error bound of Newton's, Kantorovich and Horner's algorithm

4.1 Newton's Method in Real Euclidean Space.

Let $A \subseteq \mathbb{R}$ be an open set and let $f : A \to \mathbb{R}$ be a Frechet differential function and consider the equation $f(x) = 0$. If x_0 is a point of A near to a root of this equation, then a first approximation be the linear equation

$$f(x_0) + f'(x_0)(x - x_0) = 0$$

and this has the solution

$$x = x_0 - [f'(x_0)]^{-1} f(x_0) \tag{4.1}$$

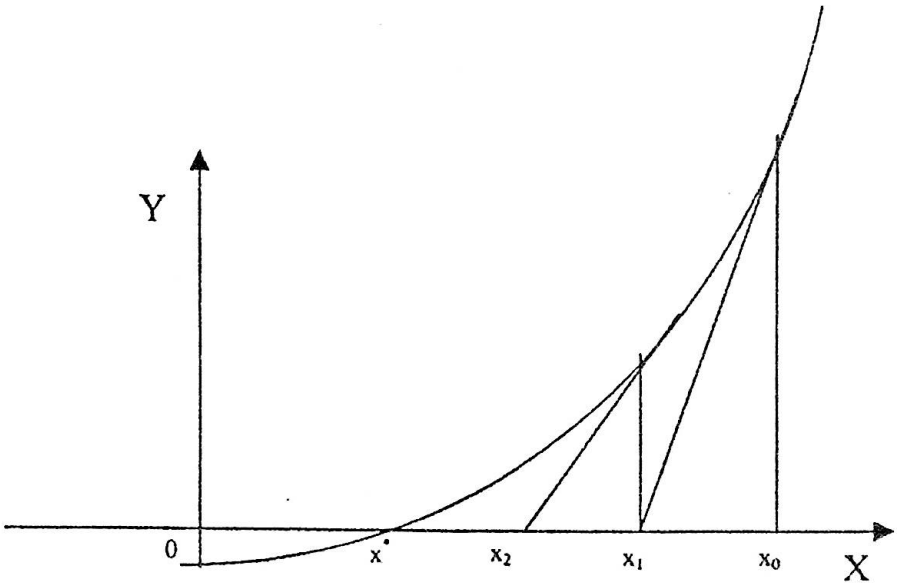

Figure 4.1: A two dimensional example

provided that the inverse $[f'(x_0)]^{-1}$ exists. Continuing in this manner, starting from the initial approximation x_0, we obtain points $x_1, x_2, x_3, \ldots$,given by

$$x_{k+1} = x_k - [f'(x_k)]^{-1} f(x_k), \quad k = 0, 1, 2, 3, \ldots. \tag{4.2}$$

x_{k+1} being defined so long as $x_1, x_2, x_3, \ldots, x_n \in A$ in effect that x_n are successive approximations for the equation

$$x = x - [f'(x)]^{-1} f(x).$$

It is intuitive that if we start from a point x_0 for which $f(x_0)$ is sufficiently small, and f' does not vary too much near x_0, then the recurrence relation will define a sequence $\{x_k\}$ that converges to a root x^* of the equation $f(x) = 0$.
This is particular transparent in the case of real-valued function f of real variable, because the recurrence formula becomes

$$x_{k+1} = x_k - \frac{f(x_k}{f'(x_k)},$$

so that x_{k+1} is the abscissa of the point where the tangent to the graph of f at meets the x-axis. This case was first considered by Newton and the sequence $\{x_k\}$ is given by the recurrence relation usually known as Newton sequence for the equation $f(x) = 0$.

An alternative possibility is to consider the recurrence relation

$$x_{k+1} = x_k - [f'(x_k)]^{-1} f(x_k), \quad k = 0, 1, 2, 3 \ldots \tag{4.3}$$

which defines successive approximations for the solution of the equation

$$x = x - [f'(x)]^{-1} f(x).$$

The successive approximations using the algorithm (4.3) is known as **modified Newton method**.

4.2 Newton's Method in Banach spaces

Let X and Y be (real or complex) Banach spaces, and F be an operator(linear or nonlinear) from X into Y which is twice differentiable in a suitable domain, Starting with an approximate solution x_0 of

$$F(x) = 0 \tag{4.4}$$

We consider the sequence defined by

$$x_{k+1} = x_k - [F'(x_k)]^1 F(x_k), \quad k = 0, 1, 2, 3 \ldots \tag{4.5}$$

Kantorovich first proposed to solve the functional equation(4.4) and was able to give theorems concerning the existence, convergence and uniqueness of solution of the equation (4.4). He also gave the error bound for the solution[9]. His two

fundamental theorems the first one guarantees the existence and convergence and the second one guarantees the uniqueness of the solution, is given in the open ball $M(x_0,\rho) = \{x \in X : \|x - x_0\| < \rho$

Theorem 4.2.1 (Kantorovich existence an convergence theorem) *Suppose the following conditions are satisfied :*

1. $F'(x_0)$ *maps* X *onto* Y *and has an inverse* $[F'(x_0)]^{-1}$ *for which* $\|[F'(x_0]^{-1}\| \leq \beta$

2. x_0 *is an approximate solution of* $F(x) = 0$ *such that*

$$\|[F'(x_0)]^{-1}F(x_0)\| \leq \eta$$

3. F *is twice differentiable in the open ball* $U_0(x_0,\rho_0)$ *and in this ball*

$$\|F''(x)\| \leq k$$

where k *is a constant and*

$$\rho_0 = (1 - \sqrt{1-2h}\frac{\eta}{h}$$

and for constants β,k,η *satisfying*

4. $h = \beta\eta k \leq \frac{1}{2}$.

Then $F(x) = 0$ *ha a solution* x^* *in the closed ball* $\overline{U_0}(x_0,\rho_0)$ *and the successive approximations defined by (1.5) converges to* x^*. *Further*

$$\|x_k - x^*\| \leq \frac{1}{2^{n-1}}(2h_0)^{2^{n-1}}\eta \tag{4.6}$$

Here we can make remark that $h_0 > 0$ hold always, since $h_0 = 0$if and only $\beta = 0$ or $\eta = 0$. We also see that the restriction on h viz.$0 < h \leq \frac{1}{2}$,gives $1 < (1 - \sqrt{1-2h})\frac{\eta}{h} \leq 2$. Thus the condition (3) of the above theorem holds if $\|F''(x)\| \leq k$ in $U_0(x_0,2\eta)$.

4.3 Numerical example of Kantorovich theorem.

Here we consider the third degree polynomial equation

$$F(x) = x^3 - 3x + 3 = 0.$$

This equation has only one root $x^* = \alpha = -2.103803402$.....Let $x_0 = -2.11$ be the approximate solution. Then we have $F(x_0) = -0.06$, $F'(x_0) = 10.35$ and $F''(x_0) = -12.66$ We calculate,

$$\|F'(x_0)\| = \left\| \frac{1}{10.35} \right\| = 0.096618357 \leq 0.099662 = \beta,$$

$$F'(x_0)]^{-1} F(x_0)\| = \left| \frac{-0.06}{10.35} \right| = 0.005771 \leq 0.00580 = \eta$$

and

$$\|F''(x_0)\| = |-12.66| \leq 12.66 = k$$

Now we choose $h = \beta\eta k = 0.09662 \times 0.00580 \times 12.67 = 0071002732 < \frac{1}{2}$. Thus the solution $x^* = -2.103803402...$ is the only real root of the above equation. By the Newton's algorithm,

$$x_1 = x_0 - \frac{F(x_0)}{F'(x_0)} = -2.104202899.$$

So, $\|x_1 - x^*\| = 0.000399497$, and from the inequality , $\frac{1}{2^{n-1}}(2h_0)^{2^{n-1}}\eta$ becomes .001160. Then $\|x_1 - x^*\| \leq 0.01160$ is an improvement of the approximation.

The local uniqueness of the solution x^* depends on the bound holding in a larger sphere(ball).

Theorem 4.3.1 *(Kantorovich, Uniqueness) Let the condition (1) to (4) of the above theorem hold with $\|F''(x)\| \leq k$ in the ball $U_0(x_0, \sigma)$ where*

$$\sigma = (1 + \sqrt{1 - 2h})\frac{\eta}{h}.$$

Then the x^* of the above theorem is the unique solution of $F(x) = 0$ in the same ball. Kantorovich in his theorem used the boundness of the second derivative of the operator. Fenyo I.(1954) first make the assumption of the condition of Lipschitzian of F' and gave the modified Kantorovich theorem as:

Theorem 4.3.2 *(Kantorovich; modified) Let $F : X \to Y$, X,Y Banach spaces, be Frechet differentiable function for $x \in U$, an open convex set in X [25] [15]. Suppose that $[F'(x_0)]^{-1} \in [Y \to X]$ at some $x_0 \in U$, and*

1. $\left\|[F'(x_0)]^{-1}\right\| \leq \beta,$
2. $\left\|[F'(x_0)]^{-1}F(x_0)\right\| \leq \eta,$
3. $\|F'(x) - F'(y)\| \leq K\|x-y\|$, $x,y \in U$ *for constants* β, k,η *satisfying* $h = \beta k\eta \leq \frac{1}{2}$ *and*
4. $U_0 \subset U$, *where*

$$U_o = \{x : \|x - x_0\| \leq (1 - \sqrt{1-2h})\frac{\eta}{h}\}$$

then the successive approximations of Newton's algorithm are defined for all n, $x_k \in U_0 \quad k = 0,1,2....$and converge to $x^* \in U_0$, which satisfies $F(x^*) = 0$.
Further,

$$\|x^* - x_k\| \leq \frac{\eta}{h}\frac{[1-\sqrt{1-2h}]^{2^k}}{2^k}, \quad k = 0,1,2,3... \tag{4.7}$$

An example of f is now shown which demonstrates that no stronger claim than the Kantorovich theorem can be made for existence and convergence. In this generality, then, no better theorem can be given.

Let $f(x) = \frac{1}{2}x^2 - x + h$ and x_0. We have $f'(x_0) = -1$ and $f''(x_0) = 1$. Here, we let $\beta = k = 1$, $\eta = h$(given),the conditions of the theorem are satisfied. The roots

of f are $1 \pm \sqrt{1-2h}$ if $h \leq \frac{1}{2}$. The smaller root

$$x^* = (1 - \sqrt{1-2h} = \frac{1}{h}(1 - \sqrt{1-2h}\eta$$

The other root is

$$\frac{1}{h}(1 + \sqrt{1-2h}\eta$$

and just excluded from the region. Presumably, Kantorovich obtained his theorem by comparison with the Newton series for $x^2 - a$.

Under the same conditions as for theorem (Kantorovich modified), it can be shown that if $\Gamma_0 = [F'(x_0)]^{-1}$ and $h < \frac{1}{2}$ the successive approximation given by the following theorem converge and is known as modified Newton's method.

Theorem 4.3.3 *(Modified Newton's method) Under the same conditions as in the previous theorem if $h < \frac{1}{2}$ the iterations $x_{k+1} = x_k - \Gamma_0 F(x_k)$ where $\Gamma_0 = [F'(x_k)]^{-1}$ are defined for all n and for any $x_0 \in U$, converge to a root $x^* \in U_0$. Further, $F(x) = 0$ has a unique root in U_0. Also*

$$\|x_k - x^*\| \leq 2(\frac{\eta}{h})[1 - \sqrt{1-2h}]^{k+1}, \quad k = 1,2,3...$$

Theorem 4.3.4 *(Newton-like or Quasi-Newton method) For a real function f of a real variable x, if we try to find an approximate root of function $f(x) = 0$ by the sequence of approximations*

$$x_{k+1} = x_k - [a_k]^{-1} f(x_k), k = 0,1,2,.....$$

where $\{a_k\}$ is a sequence of real numbers, then the algorithm above is the simplest form of the Newton-like or Quasi-Newton method.

Illustration of the applicability of Newton's method to nonlinear operator equation.

let us consider the Hammersteian equationn

$$x(t) + \int_0^1 K(s,t) f(s,x(s)) ds = 0 \tag{4.8}$$

on the space $C[0,1]$.We suppose that $f \in C^2([0,1] \times \mathbb{R})$ and $k \in C([0,1] \times [0,1])$. If we consider $F : C[0,1] \to C[0,1]$ by

$$F(x)(t) = x(t) + \int_0^1 K(s,t) f(s,x(s)) ds,$$

then our problem is to find a root of F. If $x_0 \in C[0,1]$ is an initial approximation, then for any $y \in C[0,1]$,

$$\begin{aligned}[F'(x_0)y](t) &= y(t) + \int_0^1 K(s,t) f_2'(s,x_0(s)) ds, \\ &= (1+k_0)y(t),\end{aligned}$$

where k_0 is the linear integral operator on $C[0,1]$ defined by

$$(k_0 y)(t) = \int_0^1 K(s,t) f_2'(s,x_0(s)) ds.$$

We suppose that $1+k_0$ is invertible and $\|(1+k_0)^{-1}\| \leq b$ .For a given $\delta > 0$,let

$$k_\delta = sup\{|k(s,t) f_2''(s,u)| : s,t \in [0,1], |u - x_0(s)| \leq \delta\}.$$

Then K_δ serves as a Lipschitz constant for F' on the set

$$S_\delta = \{x \in C[0,1]] : \|x - x_0\| \leq \delta\}$$

If $\|F(x_0)\| \leq p$, then

$$[F'(x_0)]^{-1} F(x_0)\| = (1+k_0)^{-1} F(x_0)\| \leq bp.$$

Therefore ,if $h = b^2 k_\delta p \leq \frac{1}{2}$ and $\delta \geq (bk_\delta)^{-1}$, then

$$\{x \in C[0,1] : \|x - x_0\| \leq t^*\} \subseteq S_\delta$$

and the following theorem guarantees that the functions x_n given by $x_{n+1} = x_n + y_n$, where the functions y_n are solutions of the linear integral equation

$$y_n(t) + \int_0^1 k(s,t) f_2'(s,x_n(s)) y - n(s) ds = -x_n(t) - \int_0^1 k(s,t) f(s,x_n(s)) ds,$$

converge to a solution of the first equation (1.8). Cherles W. Groetsch (1980) has discussed the additional applications of Newton's method.

Theorem 4.3.5 *Suppose that C is an open convex subset of a Banach space X and that Y is a Banach space. Let $F : C \to Y$ be differentiable on C and satisfy*

$$\|F'(x) - F'(y)\| \leq k\|x - y\|$$

for $x, y \in C$. Assuming that for some $x_0 \in C$, $G_0 = [F'(x_0)]^{-1}$ exists and that $\|G_0\| \leq b$ and $\|G_0 F(x_0)\| \leq \eta$, where $h = b\eta k \leq \frac{1}{2}$. We set

$$t^* = \frac{\eta(1 - \sqrt{1-2h})}{h}$$

and suppose that

$$S = \{x \in X : \|x - x_0\| \leq t^*\} \subseteq C$$

Then the Newton's sequence lies in S and converges to a root x^* of F.

Moreover,

$$\|x^* - x_k\| \leq \frac{\eta}{h} \frac{[1 - \sqrt{1-2h}]^{2^k}}{2^k}, \quad k = 0, 1, 2, 3..$$

Let X and Y be Banach spaces and D^0 is an open convex subset of X. Also let $F : D^0 \to Y$ be Frechet differentiable on D^0 with $\|F'(x) - F'(x')\| \leq \lambda\|x - x'\|$ for $x, x' \in D_0$.

Let $S(x, r)$ denotes the open ball $\{x' : \|x' - x\| < r\}$ and $\overline{S(x, r)}$ denote it closure. Let $x_0 \in D^0$ be such that $[F'(x_0)]^{-1} : Y \to X$ exists and

$$\|F'(x_0)]^{-1}\| \leq x, \|[F'(x_0)]^{-1} F_0(x_0)\| \leq \delta, h = 2x\lambda\delta \leq 1 and for S(x, t^*) \subset D^0, t^* = \frac{2}{h}(1 - \sqrt{1-h})\delta.$$

Then

1. The Newton sequence $\{x_k\}$ exists and $x_k \in S(x, t) \subset D^0$ for $n \geq 0$
2. $x^* = \lim x_k$ exists, $x^* \in \overline{S(x_0, t^*)} \subset D$ and $F(x^*) = 0$.
3. x^* is the only solution of $F(x) = 0$ in the set $S(x_0, t') \cap D^0$, $t' = \frac{2}{h}(1 + \sqrt{1-h})\delta$ if $h < 1$, and in $\overline{S(x_0, t')}$ if $h = 1$

With the above assumptions and Kantorovich hypotheses,Gragg W.B and Tapia R.A.(1974) were able to give the following best possible lower and upper bounds for error:

$$\|x^* - x_k\| \leq \frac{4\sqrt{1-h}}{h} \frac{\theta^{2^n}}{1-\theta^{2^n}} \|x_1 - x_0\|$$

and

$$\frac{2\|x_{k+1} - x_k\|}{1 + \sqrt{\frac{1+4\theta^{2^n}}{(1+\theta^{2^n})^2}}} \leq \|x^* - x_k\| \leq \theta^{2^{n-1}} \|x_k - x_{k-1}\|$$

where

$$\theta = \frac{1-\sqrt{1-h}}{1+\sqrt{1-h}} \leq 1.$$

In particular, the bounds for $h = 1$ are

$$\|x^* - x_0\| \leq 2^{n-1} \|x_1 - x_0\|$$

and

$$2(\sqrt{2}-1)\|x_{k+1} - x_k \leq \|x^* - x_k\| \leq \|x_k - x_{k-1}\|$$

holds for all $h \leq 1$

4.4 Horner's Method

Evaluation of a Polynomial.

Let the polynomial $P(x)$ of degree n have the form

$$P(x) = a_n x^n + a_{n-1}x^{n-1} + ... + a_k x^k + ... + a_2 x^2 + a_1 x + a_0 \qquad (4.9)$$

where $a_0, a_1, ...a_n$ are real numbers. **Horner's method** or **synthetic division** is a technique for evaluating polynomials. It can be thought of as nested multiplication. For example, a fifth-degree polynomial

$$P_5(x) = a_5 x^5 + a_4 x^4 + a_3 x^3 + a_2 x^2 + a_1 x + a_0$$

can be written in the "nested multiplication" form

$$P_5(x) = ((((a_5 x + a_4)x + a_3)x + a_2)x + a_1)x + a_0$$

Theorem 4.4.1 (Horner's method for polynomial evaluation) *Assume that $P(x)$ is the polynomial given in equation (4.9) and $x = z$ is a number for which $P(z)$ is to be evaluated.*

We set $b_n = a_n$ *and Compute*

$$b_k = a_k + z b_{k+1} \quad for \quad k = n-1, n-2, ..., 2, 1, 0 \qquad (4.10)$$

then $b_0 = P(z)$. Moreover, if we define

$$Q_0(x) = b_n x^{n-1} + b_{n-1}x^{n-2} + ... + b_3 x^3 + b_2 x + b_1 \qquad (4.11)$$

then

$$P(x) = (x - z)Q_0(x) + R_0 \qquad (4.12)$$

Where $Q_0(x)$ is the quotient polynomial of degree $n-1$ and $R_0 = b_0 = P(z)$is the remainder.

Proof. Substituting the right side of equation (4.11) for $Q_0(x)$ and b_0 for R_0 in equation (4.12) yields

$$P(x) = (x-z)(b_n x^{n-1} + b_{n-1}x^{n-2} + ... + b_3x^2 + b_2x + b_1) + b_0 \qquad (4.13)$$

The right side of (5) can be rewritten in decreasing powers of x

$$P(x) = b_n x^n + (b_{n-1} - zb_n)x^{n-1} + ... + (b_k - zb_{k+1})x^k + ... + (b_2 - zb_3)x^2 + (b_1 - zb_2)x + (b_0 - zb_1) \qquad (4.14)$$

The numbers b_k are determined by comparing the coefficients of x^k in equation (4.9) and (4.14) as shown in Table 1.1. The value of $P(z) = b_0$ is easily obtained by substituting $x = z$ in equation (4) and using the fact that $R_0 = b_0$

$$P(z) = (z-z)Q_0(z) + R_0 = b_0 \qquad (4.15)$$

The recursive formula for b_k given in (4.10) is easy to implement with a computer. A simple algorithm is

$a_n = b_n$

for $k = n-1 : -1 : 0$

$b_k = a_k + zb_{k+1}$

end.

Table 1.1 Coefficients b_k for Horner's Method

x^k	Comparing (1) and (3)	Solving for b_k
x^k	$a_n = b_n$	$b_n = a_n$
x^{n-1}	$a_{n-1} = b_{n-1} - zb_n$	$b_{n-1} = a_{n-1} + zb_n$
$\vdots$	$\vdots$	$\vdots$
x^k	$a_k = b_{n-1} - zb_{k+1}$	$b_k = a_k + zb_{k+1}$
$\vdots$	$\vdots$	$\vdots$
x^0	$a_0 = b_0 - zb_1$	$b_0 = a_0 + zb_1$

Table 1.2 Horner's table for Synthetic Division Process

Input	a_n	a_{n-1}	a_{n-2}	$\cdots$	a_k	$\cdots$	a_2	a_1	a_0
$x=z$		zb_n	zb_{n-1}	$\cdots$	zb_{k+1}	$\cdots$	zb_3	zb_2	zb_1
	b_n	b_{n-1}	b_{n-2}	$\cdots$	b_k	$\cdots$	b_2	b_1	$b_0 = P(z)$

Example 1. Use the "Horner's table" to find $P(3)$ for the polynomial

$$P(x) = x^5 - 6x^4 + 8x^3 + 8x^2 + 4x - 40 = 0.$$

Solution 1: In our example $n = 5 \quad a_5 = 1, a_4 = -6, a_3 = 8, a_2 = 8, a_1 = 4$ and $a_0 = -40$

Input	1	-6	8	8	4	-40
$x=3$		3	-9	-3	15	57
	1	-3	-1	5	19	$17 = P(3)$ Output

Lemma 4.4.2 (Horner's Method for Derivatives) *Assume that*

$$P(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_k x^k + \cdots + a_3 x^3 + a_2 x^2 + a_1 x + a_0$$

and x=z is a number for which $P(z)$ and $P'(z)$ are to be evaluated. We have already seen that $P(z) = b_0$ can be computed recursively as follows.

We set $b_n = a_n$ *and compute*

$b_k = a_k + zb_{k+1}$ *for* $k = n-1, n-2, ..., 2, 1, 0.$

The quotient polynomial

$$Q_0(x) = b_n x^{n-1} + b_{n-1}x^{n-2} + ... + b_3 x^3 + b_2 x + b_1$$

and remainder $R_0 = b_0 = P(z)$ form the relation

$$P(x) = (x-z)Q_0(x) + R_0$$

We can be computed $P'(z) = c_1$ *recursively as follows:*

We set $c_n = b_n$ *and compute*

$c_k = b_k + zc_{k+1}$ *for* $k = n-1, n-2, ..., 2, 1$

The quotient polynomial

$$Q_1(x) = c_n x^{n-2} + c_{n-1}x^{n-3} + \cdots + c_4 x^2 + c_3 x + c_2$$

and remainder $R_1 = c_1 = P'(z)$ *form the relation*

$$Q_0(x)(x) = (x-z)Q_1(x) + R_1$$

The Horner table was used for computing the coefficients is given bellow.

Input	a_n	a_{n-1}	a_{n-2}	$\cdots$	a_k	$\cdots$	a_2	a_1	a_0
$x = z$	$\downarrow$	$+zb_n$	$+zb_{n-1}$	$\cdots$	$+zb_{k+1}$	$\cdots$	$+zb_3$	$+zb_2$	$+zb_1$
	$\nearrow$	$\nearrow$	$\nearrow$	$\cdots$	$\nearrow$	$\cdots$	$\nearrow$	$\nearrow$	
	b_n	b_{n-1}	b_{n-2}	$\cdots$	b_k	$\cdots$	b_2	b_1	$b_0 = P(z)$ *Output*
	b_n	b_{n-1}	b_{n-2}	$\cdots$	b_k	$\cdots$	b_2	b_1	
	$\downarrow$	$+zc_n$	$+zc_{n-1}$	$\cdots$	$+zc_{k+1}$	$\cdots$	$+zc_3$	$+zc_2$	
	$\nearrow$	$\nearrow$	$\nearrow$	$\cdots$	$\nearrow$	$\nearrow$	$\nearrow$	$\nearrow$	
	c_n	c_{n-1}	c_{n-2}	$\cdots$	c_k	$\cdots$	c_2	$c_1 = P'(z)$ *Output*	

Using vector coefficients: As mentioned above, it is efficient to store the coefficients $\{a_{[k]}\}_{k=1}^{n+1}$ of a polynomial $P(x)$ of degree n in the vector $a = \{a_{[1]}, a_{[2]}, a_{[3]}, \cdots, a_{[n]}, a_{[n+1]}\}$. Notice that this is a shift of the index for $a_{[k]}$ and the polynomial $P(x)$ is written in the form

$$P(x) = \sum_{k=0}^{n} a_{[k+1]}x^k$$

to r Given the value $x = z$, the recursive formulas for computing the coefficients $\{b_{[k]}\}_{k=1}^{n+1}$ and $\{c_{[k]}\}_{k=2}^{n+1}$ of $Q_0(x)$ and $Q_1(x)$, we have the new form

$$b_{[n+1]} = a_{[n+1]}$$

$$b_{[k]} = a_{[k]} + zb_{[k+1]} \quad for \quad k = n, n-1, \cdots, 3, 2, 1.$$

$$c_{[n+1]} = b_{[n+1]}$$

$$c_{[k]} = b_{[k]} + zc_{[k+1]} \quad for \quad k = n, n-1, \cdots, 3, 2, 1.$$

Then $P(z) = b_{[1]}$ and $P'(z) = c_{[2]}$

4.5 Newton-Horner method

Assume that $P(x)$ is a polynomial of degree $n \geq 2$ and there exists a number $r \in [a,b]$,where $P(r) = 0$. If $P'(r) \neq 0$,then there exists a $\delta > 0$ such that the sequence $\{r_k\}_{k=0}^{\infty}$ defined by the Newton-Raphson iteration formula

$$r_{k+1} = r_k - f(r_k)/f'(r_k) \quad for \quad k = 0, 1, 2 \cdots$$

will converge to $r_0$for any initial approximation $r_0 \in [r-\delta, r+\delta]$. The recursive formulas in the Lemma can be adapted to compute $P(r_k) = b_{k,0} = b_{[1]}$ and $P'(r_k) = c_{k,1} = c_{[2]}$ and the resulting **Newton-Horner iteration formula** looks like

$$r_{k+1} = r_k - \frac{b_{[1]}}{c_{[2]}} \quad for \quad k = 0, 1, 2.....$$

Algorithm (Newton-Horner Iteration). To fine a root of $f(x) = 0$ given an initial approximation r_0 using iteration

$$r_{k+1} = r_k - \frac{b_{[1]}}{c_{[2]}} \quad for \quad k = 0, 1, 2..., max.$$

4.6 Horner's Method for Higher Derivatives

Assume that the coefficients $\{a_{[1,k]}\}_{k=1}^{n+1}$ of a polynomial $P[x]$ of degree n are stored in the first row of the matrix $[a_{i,j}]_{n+2\times n+1}$. Then the polynomial $P(x)$ can be written in the form

$$P(x) = \sum_{k=0}^{n} a_{[1,k+1]}x^k$$

Given the value $x = z$, the subroutine for computing all the derivatives $\{p^i[z]\}_{i=0}^{n}$ is

For [$i = 2, i \leq n+2, i++$,

$a_{[i,n+1]} = a_{[i-1,n+1]}$

For $[k = n, i-1 \leq k, k--$,

$a_{[i,k]} = a_{[i-1,k]} + za_{[i,k+1]};]$]and

$$p^{(i)}[z] == i!a_{[i+2,i+1]} \quad for \quad i = 0, 1, \cdots, n$$

Chapter 5

Interval Matrices

Definition 5.0.1 *By an interval matrix, we mean a matrix whose elements are interval components and the space of all $m \times n$ matrices is denoted by $\mathbb{IR}^{m\times n}$. For example, we might have*

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} [1,2] & [-1,1] \\ [0,4] & [6,8] \end{pmatrix} \tag{5.1}$$

If A is an interval matrix with elements A_{ij} and B is a matrix with real elements B_{ij} such that $B_{ij} \in A_{ij}$ for all i and j, then we write $B \in A$.

In this section we define two types of matrices. An interval matrix A is called an **M-matrix** if and only if $A_{ij} \leq 0$ for all $i \neq j$ and $Au > 0$ for some positive vector $u \in \mathbb{R}^n$. If the comparison matrix $\langle A \rangle$, where

$$\langle A \rangle_{ii} = min\{|\alpha| : \alpha \in A_{ii}\}$$

$$\langle A \rangle_{ik} = -max\{|\alpha| : \alpha \in A_{ik}\},$$

is an M-matrix then A is said to be an **H-matrix.**

5.1 Matrix Norm, Width and Midpoint

We use the matrix norm

$$\|A\| = \max_i \sum_{j=1}^{n} |A_{ij}| \tag{5.2}$$

for an interval matrix A. This is an interval extension of the maximum row sum norm for real matrices. If B is any real matrix contained in an interval matrix A, then $\|B\| \leq \|A\|$. We define the width $w(A)$ of an interval matrix A by

$$w(A) = \max_{ij} w(A_{ij}). \tag{5.3}$$

The midpoint of A is the real matrix $m(A)$ whose elements are the midpoints of the corresponding elements of A:

$$m(A_{ij}) = m(A_{ij}.$$

Lemma 5.1.1 *The product of two interval matrices using interval arithmetic is again an interval matrix consisting of interval elements each of which is exactly the range of values of the corresponding element of the product of a pair of real matrices whose elements are chosen independently from the corresponding elements of the interval matrices.*

Interval Matrices and Dependency

Even though, as indicated in the above lemma, the ij-th elements C_{ij} of the product $C = AB$ of an m by p interval matrix A and a p by n interval matrix B gives sharp bounds on the ranges

$$C_{ij} = \left\{ M_{ij} = \sum_{k=1}^{p} P_{ik}Q_{kj} : P_{ik} \in A_{ik} \quad \text{and} \quad Q_{kj} \in B_{kj} \quad \text{for} \quad 1 \leq k \leq p \right\}$$

for each i, $1 \leq i \leq m$, and each j, $1 \leq j \leq n$, the resulting interval matrix C may contain point matrices D that are not the result of the multiplication of point matrices $P \in A$ and $Q \in B$.

5.2 Systems of Linear Interval Equations

We know that the Linear systems of equations are a fundamental part of scientific calculations. In this section we consider bounding the solution set of interval linear systems.

The solution to linear systems of equations is prone to errors due to the finite precision of machine arithmetic and the propagation of error in the initial data. If the initial data is known to lie in specified ranges then the interval arithmetic enables computation of intervals containing the elements of the exact solution. Error bounds are provided in the calculation of the interval solution rather than estimating the error from an analysis of the error propagation after an approximate solution is obtained.

An interval linear system is of the form $AX = b$ where $A \in \mathbb{IR}^{n\times n}$ and $b \in \mathbb{IR}^{n\times n}$. The solution set

$$\sum(A,b) = \{\tilde{x} : \tilde{A}\tilde{x} = \tilde{b} \quad \text{for some} \quad \tilde{A} \in A \quad \tilde{b} \in b\}$$

is typically star-shaped. An example of such a solution set, we consider the system

$$A = \begin{pmatrix} [2,4] & [-2,1] \\ [-1,2] & [2,4] \end{pmatrix}, \quad b = \begin{pmatrix} [-2,2] \\ [-2,2] \end{pmatrix} \tag{5.4}$$

Exact solution set and hull of a linear system. The solution set is shown by the star-shaped area of Figure 1. The hull of the solution set is the interval vector with smallest radius containing $\sum(A,b)$ and is denoted $\overline{\sum(A,b)}$.

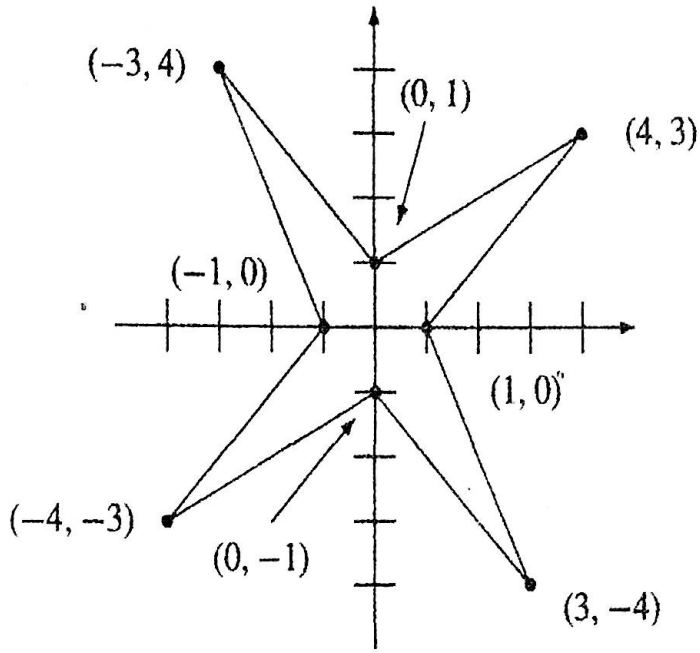

Figure 5.1: Exact solution set to Ax=b, A, and b in equation (5.4)

5.3 Interval Gaussian Elimination Method

An obvious approach is to use a generalization of Gaussian elimination adapted to deal with interval coefficients. A triangular system can be formed in the usual way but with interval arithmetic. By the **Fundamental theorem of Interval analysis** the solution of this triangular system will give an inclusion of the solution set.

The usual care has to be taken with division by zero. Column mignitude pivoting can be used to choose a pivot as the contender with the largest mignitude, where we recall that the mignitude of X is defined as

$$\text{mig}(X) = \begin{cases} \min(|a|,|b|), & \text{if } 0 \notin X\ ; \\ 0 & \text{otherwise.} \end{cases}$$

When interval Gaussian elimination is applied to a general $A \in \mathbb{IR}^{n\times n}$ and $b \in \mathbb{IR}^{n\times n}$. problems are soon encountered as n is increased. As interval calcula-

tions are carried out in the Gaussian elimination process the widths of the interval components are grow larger due to the nature of interval arithmetic. If a solution is obtained then it is likely that the width of the components are very large. Alternatively, at some stage in the Gaussian elimination process all contenders for pivot,or the bottom right element in the upper triangular system, contain zero which causes the algorithm to break down due to division by zero. For example, if the coefficient matrix is

$$A = \begin{pmatrix} [0.95,1.05] & [1.95,2.05] & [2.95,3.05] \\ [1.95,2.05] & [3.95,4.05] & [6.95,7.05] \\ [1.95,2.05] & [-0.05,0.05] & [0.95,1.05] \end{pmatrix}$$

then the upper triangular system is given by

$$U = \begin{pmatrix} [1.95,1.05] & [3.95,4.05] & [6.95,7.05] \\ [0,0] & [-4.31,-3.71] & [-6.46,-5.56] \\ [0,0] & [0,0] & [-1.23,0.23] \end{pmatrix}$$

This causes division by zero when using back-substitution. All the elements began a radius of 0.05, but the radius of U_{33} is 0.7301.

Despite interval Gaussian elimination not being effective in general, it is suitable for certain classes of matrices. In particular, realistic bounds for the solution set are obtained for M-matrices, H-matrices, diagonally dominant matrices, tridiagonal matrices and 2×2 matrices. In the case where A is an M-matrix the exact hull $\overline{\sum(A,b)}$. is obtained for many b;[13] shows that if $b \geq 0$, $b \leq 0$ or $0 \in b$ then the interval hull of the solution set is obtained.

An implementation written in INTLAB of interval Gaussian elimination with mignitude pivoting is given by the function intgauss.m below.

```
function x = intgauss (A, b)
%INTGAUSS  Interval Gaussian Elimination with mignitude pivoting.
%          x = INTGAUSS (A,b)
%          Solves Ax = b for linear systems of equations
%          INPUT  :   A   coefficient matrix
%                     b   right hand side vector
%          OUTPUT :   x   interval solution
 n = length (A);
for i = 1: n-1;
    [maxmig,index] =  max(mig(A(i:n,i)));
    if maxmig <=0
        error('All possible pivots contain zero.')
    end
    k = index + i-1;
    if  k  ~= i                  % Swap the rows if necessary.
    A([i,k],i:n) = A([k,i],i:n);
    b([i,k]) = b([k,i]);
   end

    for j = i+1:n;
      mult = A(j,i)/A(i,i);
     A(j,i+1:n) = A(j,i+1:n)-mult*A(i,i+1:n);
     b(j) = b(j) - mult*b(i);
     end
end
```

```
x(n) = b(n)/A(n,n);
for i = n-1:-1:1;
x(i) = (b(i)-A(i,i+1:n)*x(i+1:n,1))/A(i,i);
end
```

5.4 Krawczyk's Method

The linear interval system $AX = b$ can be preconditioned by multiplying by a matrix $Y \in \mathbb{R}^{n\times n}$. Here we choose Y to be the inverse of the mid point matrix of A, which often leads to the matrix YA being an H-matrix. If this is the case then Gaussian elimination can be used, but it is quicker to compute an enclosure of the solution by Krawczyk's method.

Assuming an interval vector X_i is known such that $\overline{\sum(A,b)} \subseteq X_i$ then

$$\tilde{A^{-1}}\tilde{b} = Y\tilde{b} + (I - Y\tilde{A})\tilde{A^{-1}}\tilde{b} \in Yb + (I - YA)X_i$$

holds for all $\tilde{A} \in A$ and $\tilde{b} \in b$, so that

$$\overline{\sum(A,b)} \subseteq X_i \Rightarrow \overline{\sum(A,b)} \subseteq (Yb + (I - YA)X_i \cap X_i \tag{5.5}$$

This gives the Krawczyk's iteration

$$X_{k+1} = (Yb + (I - YA)X_i \cap X_i \tag{5.6}$$

To start the iteration we require an initial vector X_0 such that the solution $\tilde{x} \in X_0$ and $\overline{\sum(A,b)} \subseteq X_0$. A possible X_0 can be found with the aid of the following theorem.

Theorem 5.4.1 *If Y satisfies $\|I - Y\tilde{A}\| = \beta < 1$, $\tilde{x} = \tilde{A^{-1}}\tilde{b}$ and $\|.\|$ is any subordinate norm , then*

$$\|\tilde{x}\| = \frac{\|Y\tilde{b}\|}{1-\beta}.$$

Proof. From $\tilde{A}\tilde{x} = \tilde{b}$ we have $\tilde{x} = Y\tilde{b} + (I - Y\tilde{A})\tilde{x}$, and hence

$$\begin{aligned} \|\tilde{x}\| &\le \|Y\tilde{b}\| + \|I - Y\tilde{A}\|\|\tilde{x}\| \\ &\le \|Y\tilde{b}\| + \beta\|\tilde{x}\| \end{aligned}$$

which gives the result. Since $\|Y\tilde{b}\|_\infty \le \|Yb\|_\infty$ and $\beta < 1$ is very likely for Y being the inverse of the midpoint of A, we define the initial interval vector to be

$$X_0 = ([-\alpha, \alpha], \ldots, [-\alpha, \alpha])^T \quad \text{with} \quad \alpha = \frac{\|Yb\|_\infty}{1 - \beta}.$$

The iterations can be terminated if the radii of the components of X_i are no longer rapidly decreasing.

```
This routine solves an interval linear system of equations
by Krawczyk'smethod.
function x = Kraw (A, b)
%KRAW    Solves  Ax = b for interval linear systems.
%        x = Kraw (A,b)
%     Solves linear systems of equations using Krawczyk's method.
%     If A and b have all real components then x is a verified bou
%     on the solution. If A or b are of interval type then an oute
%     estimate of the interval hull of the system is given.
%     INPUT  :  A   coefficient matrix
%               b   right hand side vector
%     OUTPUT :  x   interval solution
n = length (A);
midA = mid(A);
Y = inv(mid(A));                    % Preconditions system.
b = Y*b;
YA = Y*intval(A);
A = eye(n)-YA;
```

```
btea = norm(A,inf);
if beta >= 1;
     error ('Algorithm not suitable for this A')
end;
alpha = norm (b, inf)/(1-beta);
 x(1:n,1) = infsup(-alpha, alpha);
s_old = inf;
s = sum(rad(x));
mult = (1+beta)/2;
while  s <  mult*s_old
       x = intersect (b+A*x,x);          %Krawczyk's iteration.
       s_old = s ;
       s = sum(rad(x));
 end
```

5.5 The Hansen-Bliek-Rohn-Ning-Kearfott-Neumaier Method

A bound for the interval hull of the solution set of linear interval equations is given by Hansen[5] for the case where the midpoint matrix of A is the identity.This result was also found by Bliek[2], but it was Rohn[21] who first gave a rigorous proof.Ning and Kearfott[23], generalised the result for the case when A is an H-matrix. This is of particular interest since the coefficient matrix can be preconditioned in an attempt to produce an H-matrix.

The method is based on the following theorem which uses the comparison matrix,$\langle A \rangle$.

Theorem 5.5.1 *Let $A \in \mathbb{IR}^{m\times n}$ be an H-matrix, $b \in \mathbb{IR}^n$ a right hand side,*

$$u = \langle A\rangle^{-1}|b|, \quad d_i = (\langle A\rangle^{-1})_{ii},$$

and

$$\alpha_i = \langle A_{ij}\rangle - 1/d_i, \quad \beta_i = u_i/d_i - |b_i|.$$

Then $\overline{\sum(A,b)}$ is contained in the vector x with components

$$x_i = \frac{b_i + [-\beta_i, \beta_i]}{A_{ii} + [-\alpha_i, \alpha_i]}$$

A simplified proof is given by Neumaier[13] .

In order to give a rigorous enclosure of the interval hull using floating point arithmetic, rigorous upper bounds are required for α_i and β_i. These are obtained if a rigorous bound B for $\langle A\rangle^{-1}$ is used. The following explanation of how this is achieved is based on that given in[13].

A property of the H-matrix A is that $\langle A\rangle^{-1}$ is nonnegative. This suggests that an upper bound B for $\langle A\rangle^{-1}$ can be expressed in terms of $\tilde{B}$ an approximation to $\langle A\rangle^{-1}$ and vectors $v \in \mathbb{R}^n$, $u \in \mathbb{R}^n$ satisfying $I - \langle A\rangle\tilde{B} \leq \langle A\rangle vw^T$ by

$$B = \tilde{B} + vw^T.$$

By the definition of an H-matrix, there exists a vector $v > 0$ such that $u = \langle A\rangle v > 0$ This vector v can be used to satisfy the above relation by taking the vector w with components

$$w_k = \max_i \frac{-R_{ik}}{u_i},$$

where $R = \langle A\rangle\tilde{B} - I$. It is now left to find the vector v. Assuming there is a positive vector $\tilde{u}$ such that $v = \tilde{B}\tilde{u} \approx \langle A\rangle^{-1}\tilde{u} > 0$ then A is an H-matrix, and if $u = \langle A\rangle v \approx \tilde{u}$ is positive then the approximation B is a good enough. Since $\langle A\rangle^{-1}$ is nonnegative, the vector $\tilde{u} = (1,\ldots,1)$ is sufficient to produce $\langle A\rangle^{-1}\tilde{u} > 0$.

The values u and R must be calculated with downward rounding, w and B calculated with upward rounding while $\tilde{B}$ and v can be calculated with nearest rounding.
Now we will be implemented the above method in INTLAB described below.

This routine solves an interval linear system of equations by a method based on work by Hansen Bliek, Rohn, Ning, Kearfott and Neumaier.

```
function x = hsolve (A, b)
%HSOLVE    Solves  Ax = b for interval linear systems.
%         x = HSOLVE (A, b)
%     Solves linear systems of equations using a method motivated
%     by Hansen, Bliek, Rohn, Ning, Kearfott and Neumaier
%     If A and b have all real components then x is a verified boun
%     on the solution. If A or b are of interval type then an outer
%     estimate of the interval hull of the system is given.
%     INPUT  :  A   coefficient matrix
%               b   right hand side vector
%     OUTPUT :  x   interval solution


  n = dim(A);
Y = inv(mid(A));
A = Y*A ;
b = Y*b;
dA = diag(A);                    % Diagonal entries of A.
A = compmat(A);                  % Comparison matrix.
B = inv(A);
v = abss(B*ones(n,1));
setround (-1)
```

```
u=A*v ;
if ~ all(min(u)>0)                % Check positivity of u.
      error ('A is not an H-matrix')
else
    dAc = diag(A);
    A = A*B-eye(n);               % A contains the residual matrix.
    setround (1)
    w = zeros(1, n);
    for  i = 1:n,
         w = max(w,(-A(i,:))/u(i));
    end;
    dlow = v.*w'- diag(B);
    B = B+v*w;                    % Rigorous upper bound for exact B.
    u = B*abss(b);
    d = diag(B);
    alpha = dAc+(-1)./d;
    beta =  u./dlow-abss(b);
    x =  (b+midrad(0,beta))./(dA+midrad(0,alpha));
 end
```

Chapter 6

Interval Newton Methods

6.1 Interval Version of Newton's Method in One Dimension

We consider an interval X that contains a root x^* of f, where f is continuously differentiable. Then the mean value theorem tells us that there exists a $\xi \in X$ such that

$$0 = f(x^*) = f(x) + (x^* - x)f'(\xi)$$

Assuming that $f'(\xi) \neq 0$, this leads to

$$x^* = x - \frac{f(x)}{f'(\xi)}$$

Now,let F' be an interval extension of f' and assume that $F'(X)$ does not contain 0. Then

$$x^* \in N(x,X) \equiv x - \frac{f(x)}{F'(X)}, \tag{6.1}$$

Where we have introduced the Newton operator $N(x,X)$. The root is also contained in X, and therefore it must lie in the intersection between X and $N(x,X)$,

$$x^* \in X \cap N(x,X).$$

Thus, we can formulate the interval version of Newton's method[20]. We will start with X_0 containing x^* and computing a nested sequence of intervals $X_{(1)}, X_{(2)}, \ldots\ldots$ by the formula

$$X_{k+1} = X_k \cap N(x_k, X_k) \quad with \quad x_k \in X_k, k = 0, 1 \ldots \tag{6.2}$$

The term "nested" is used because each new interval is contained in the previous,$X_0 \supseteq X_1 \supseteq \ldots\ldots$ This implies that the widths decrease, $W(X_0) \geq W(X_1) \geq \ldots\ldots$, and since all the X_k are contained in X_0, they are bounded. Therefore, there exists a limit X containing x^*.

6.1.1 Numerical Example

Consider the function $f(x) = x^2 - 2$ with $f'(x) = 2x$ and $X_0 = [1,2]$. We choose $x_k = m(X_k)$ and $F'(X) = 2X$, so that the Newton operator is

$$N(x_k, X_k) = m(x_k) - (m(X_k) - (m(X_k))^2 - 2)/(2X_k).$$

The first step with (2.2) is

$$N(x_0, X_0) = \frac{3}{2} - \frac{1}{4}/[2,4] = [\frac{22}{16}, \frac{23}{16}]$$

$$X_1 = [1,2] \cap [[\frac{22}{16}, \frac{23}{16}] = [1.3749999999998, 1.43750000000001]$$

We observe that $X_1 subset X_0$. Continuing, we compute from (2.2) the results obtained by the MATLAB Toolbox INTLAB, and displayed with format long.

$$X_2 = [1.41406249999998, 1.41441761363637]$$

$$X_3 = [1.41421355929452, 1.41421356594719]$$

$$X_4 = [1.41421356237309, 1.41421356237310]$$

$$\vdots$$

For this example, we have seen that the interval version of Newton's method gives a very rapidly contracting sequence of intervals containing $\sqrt{2}$

The condition that $F'(X)$ does not contain 0 is necessary for the application of the Newton operator, and it has an interesting implication.

Lemma 6.1.1 *If $f : \mathbb{R} \to \mathbb{R}$ is continuously differentiable on the interval X and $0 \notin F'(X)$, then X either contains a simple root x^* or no roots.*

Now we will state the following theorems without Proof.

Theorem 6.1.2 *If x^* is a root of f in X_0 and $F'(X_0)$ does not contain 0, then the sequence $X_1, X_2, X_3, \ldots$. defined by (2.2) converges to x^*.*

Theorem 6.1.3 *If X is an interval with $0 \notin F'(X)$ and there exists an $x \in X$ such that $X \cap N(x,X) = \phi$ then f has no root in X.*

Theorem 6.1.4 *If f is continuously differentiable, X is an interval with $0 \notin F'(X)$, then there exists an $x \in X$ such that $N(x,X) \subseteq X$, then f has a root in X.*

Theorem 6.1.5 *If there is a root in $X \subseteq X_0$, F' is a linear interval extension of f', and $0 \notin F'(X)$,then there exists a constant $K > 0$ such that*

$$w((N(x,X)) \leq Kw(X)^2 for all x \in X$$

6.2 Interval Version of Multidimensional Newton's Method

We will consider continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}^n$, and seek a solution $x^* \in \mathbb{R}^n$ such that $f_i(x^*), i = 0, 1, 2..., n$. This is the system of n nonlinear

equations in n unknowns, the components $x_1^*,...,x_n^*$ of x^*. Again, we start with mean value theorem, but this time in its multidimensional version. Let $x^* \in \mathbb{R}^n$ be a root for x. There exist vectors $\xi_{[1]},...,\xi_{[n]} \in X$ such that

$$0 = f(x^*) = f(x) + J(x,x^*)(x^* - x),$$

where $J \in \mathbb{R}^{n \times n}$ is the Jacobian with the elements

$$(J(x,x^*))_{i,j} = \frac{\delta f_i}{\delta x_j}(\xi_{[i]})$$

Assume that the Jacobian is nonsingular for all $x \in X$. Then the inverse of this matrix exists, and

$$x^* = x - (J(x,x^*))^{-1} f(x) \in x - V(X) f(x)$$

Where V is an interval matrix containing all possible J^{-1},

$$\{(J(x,x^*))^{-1}, x \in X\} \subseteq V(X)$$

Thus from X_0 containing the root x^* we can construct a nested sequence of intervals, that all contains x^*,

$$X_{k+1} = X_k \cap N(x_k, X_k), \quad k = 0,1,2,3.... \tag{6.3}$$

with

$$N(x_k, X_k) = x_k - V(X) f(x_k), \quad x_k \in X_k$$

If there is a zero of f in X_0 then the zero is also in $N(X_0)$ and therefore in X_k for $k = 0,1,2,3..$

6.2.1 Numerical Example-1

Let us consider the following system

$$\begin{aligned} f_1(x_1,x_2) &= x_1^2+x_2^2-1=0 \\ f_2(x_1,x_2) &= x_1-x_2=0 \end{aligned} \tag{6.4}$$

For an interval vector $X=(X_1,X_2)$, we can take $F(X)=\begin{pmatrix} X_1^2+X_2^2-1 \\ X_1-X_2 \end{pmatrix}$ and $F'(X)=\begin{pmatrix} 2X_1 & 2X_2 \\ 1 & -1 \end{pmatrix}$ as the respective interval extensions of the vector function $\mathbf{f(x)}=\begin{pmatrix} f_1(x_1,x_2) \\ f_2(x_1,x_2) \end{pmatrix}=\begin{pmatrix} x_1^2+x_2^2-1 \\ x_1-x_2 \end{pmatrix}$ and the Jacobian matrix

$$f'(x)=\begin{pmatrix} 2x_1 & 2x_2 \\ 1 & -1 \end{pmatrix}=J(x,x^*)=\begin{pmatrix} 2\xi_{11} & 2\xi_{12} \\ 1 & -1 \end{pmatrix}, \quad \xi_{[1]}=(\xi_{11},\xi_{12})$$

We considered the given function defined on the interval $X_0=[\frac{1}{2},1],[\frac{1}{2},1]$

Now $(J(x,x^*))^{-1}=\frac{1}{\xi_{11}+\xi_{12}}\begin{pmatrix} \frac{1}{2} & \xi_{12} \\ \frac{1}{2} & -\xi_{11} \end{pmatrix}$, and we can use

$$V(X)=\frac{1}{X_1+X_2}\begin{pmatrix} \frac{1}{2} & X_2 \\ \frac{1}{2} & -X_1 \end{pmatrix}.$$

If we choose the midpoint $x_k=m(X_k)$ in () we get the following results.

k	$X_k,1=X_k,2$	$w(X_k)$
1	[0.68749999999998, 0.71875000000001]	$3.12e-02$
2	[0.70703124999999, 0.70720880681819]	$1.78e-04$
3	[0.70710677964726, 0.70710678297359]	$3.33e-09$
4	[0.70710678118654, 0.70710678118655]	$1.11e-16$
5	[0.70710678118654, 0.70710678118655]	$1.11e-16$

We see that a fast (quadratic) convergence to the solution, $x_1^*=x_2^*=\sqrt{0.5}$. The width of X_4 is half of the machine accuracy, and we cannot get closer to x^* because

of rounding errors.

6.2.2 Numerical Example-2

Let us consider the following system

$$\begin{aligned} f_1(x_1,x_2) &= x_1^2+x_2^2-5=0 \\ f_2(x_1,x_2) &= x_1x_2-2=0 \end{aligned} \tag{6.5}$$

For an interval vector $X=(X_1,X_2)$, we can take $F(X)=\begin{pmatrix} X_1^2+X_2^2-5 \\ X_1X_2-2 \end{pmatrix}$ and $F'(X)=\begin{pmatrix} 2X_1 & 2X_2 \\ X_2 & X_1 \end{pmatrix}$ as the respective interval extensions of the vector function $\mathbf{f(x)}=\begin{pmatrix} f_1(x_1,x_2) \\ f_2(x_1,x_2) \end{pmatrix}=\begin{pmatrix} x_1^2+x_2^2-5 \\ x_1x_2-2 \end{pmatrix}$ and the Jacobian matrix

$$f'(x)=\begin{pmatrix} 2x_1 & 2x_2 \\ x_2 & x_1 \end{pmatrix}=J(x,x^*)=\begin{pmatrix} 2\xi_{11} & 2\xi_{12} \\ \xi_{22} & \xi_{21} \end{pmatrix}$$

We considered the given function defined on the interval $X_0=[1.6,2],[1,1.4]$ and it has the root $x^*=(2,1)$. It is nonsingular in X_0 and we can use

$$V(X)=\frac{1}{2(X_1X_1-X_2X_2}\begin{pmatrix} X_1 & -2X_2 \\ -X_2 & 2X_1 \end{pmatrix}.$$

If we choose the midpoint $x_k=m(X_k)$ in () we get the following sequence.

k	$X_k,1$	$X_k,2$	$w(X_k)$
1	[1.938666666,2]	[1,1.06133334]	$6.13e-02$
2	[1.998451819,2]	[1,1.001548181]	$1.55e-04$
3	[1.999999001,2]	[1,1000000998]	$9.98e-07$
4	[1.9999999999,2]	[1,1000000001]	$4.15e-13$
5	[1.9999999999,2]]	[1,1.0000000001]	$2.22e-16$
6	[2,2]	[1,1]	0

The only new feature is that now the components of the root can be represented in floating point without error.

If, instead we take $X_0 = [2, 1.4]$ we get

$$N(x_0, X_0) = \begin{pmatrix} [0.666666666, 3.610666667] \\ [-1.10666667, 2.093333334] \end{pmatrix} \supseteq X_0$$

Therefore , the sequence defined by (6.3) get stuck at X_0 and it shows that the choice of $x_0 \in X_0$ can have great importance.

6.3 Fixed Point Theorems.

If T is an operator that maps the Banach space X into itself, then any $x \in X$ such that

$$x = Tx \tag{6.6}$$

is called a fixed point of the operator T. For example, the operator $Tx = x^2$ in the space $\mathbb{R}$ of real numbers has the fixed points $x = 0$ and $x = 1$.The linear operator

$$Tx = x(0) + \int_0^1 x(t)dt$$

in $C[0, 1]$ has any function $x = x(s)$ of the form

$$x(s) = ce^s, \quad 0 \leq s \leq 1,$$

as a fixed point, where c is a real constant.

The method for finding a fixed point of an equation $f(x) = 0$ is an iterative method. This method is based on the principle of finding a sequence $\{x_k\}$ each element of which successively approximates a root x^* of the equation $f(x) = 0$ in some interval $[a, b]$. Therefore, there is a deep connection between the study of fixed point theorems and Newton's method of iteration.

The iteration process $x_{k+1} = f(x_k)$ leads to a solution of the equation $x = f(x)$, where f maps the real line into itself if the mapping $f(x)$ is contractive. The Newton's algorithm

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

for finding the real roots of the algebraic and transcending equations $f(x) = 0$ is also an iterative process. Let us consider the mapping

$$y = x - \frac{f(x_k)}{f'(x_k)}$$

of the real line into itself, where $\frac{f(x)}{f'(x)}$ is continuous on a closed interval $[a,b]$ and differentiable on the open interval (a,b). If x_1 and x_2 are any two points of $[a,b]$ which maps into y_1 and y_2, thus

$$\begin{aligned} y_1 - y_2 &= x_1 - x_2 - \left\{ \frac{f(x_1)}{f'(x_2} - \frac{f(x_2)}{f'(x_2} \right\} \\ &= (x_1 - x_2)\frac{f(\xi)f''(\xi)}{(f'(\xi))^2} \end{aligned}$$

where ξ lies between x_1 and x_2. Hence if

$$\left\| \frac{f(x)f''(x)}{(f'(x))^2} \right\| \leq k < 1$$

on $[a,b]$, the mapping is a contraction mapping of $[a,b]$ onto a closed interval of the real line. Hence we can conclude state in below theorem that if there is a point $x_0 \in (a,b)$ such that $\left| \frac{f(x_0)}{f'(x_0)} \right| < \delta|1-k|$ the mapping has a unique fixed point $\alpha \in (x_0 - \delta, x_0 + \delta) \cap [a,b]$ and that Newton's sequence $\{x_k\}$ converges to α.

Here we will present an important theorem relating to the fixed point iteration method.

Definition 6.3.1 ***Contraction mapping principle.*** *A map $T : (X,d_1) \to (Y,d_2)$ of metric spaces that satisfies $d_2(Tx,Tz) \leq L'd_1(x,z)$ for some fixed constant L' and $x,z \in X$, is called Lipschitzian ; the smallest such L' is called Lipschitz constant*

$L(T)$ of T. If $L(T) < 1$, the map T is called contractive with contraction constant $L(T)$; if $L(T) = 1$, the map T is said to non-expansive.

Banach fixed point theorem

The **Banach fixed point theorem** (also known as the **contraction mapping theorem or contraction mapping principle**) is an important tool in the theory of metric spaces ; it guarantees the existence and uniqueness of fixed points of certain self maps of metric spaces, and provides a constructive method to find those fixed points. This theorem is named after Stefan Banach (1892-1945), and was first stated by Banach in 1992.

Theorem 6.3.1 *Let (X,d) be a non-empty complete metric space. Let $T : X \to X$ be a contraction mapping on X, i.e,: there is a nonnegative real number $q < 1$ such that*

$$d(Tx,Ty) \leq q.d(x,y)$$

for all $x,y \in X$. Then the map T admits one and only one fixed point x^* in X(this means $Tx^* = x^*$)
Furthermore, this fixed point can be found as follows: Start with an arbitrary element x_0 in X and define an iterative sequence by

$$x_k = Tx_{k-1} \quad for \quad k = 1,2,3,\ldots.$$

This sequence converges, and its limits is x^*.The following inequality describes the speed of convergence:

$$d(x^*,x_k) \leq \frac{q^k}{1-q} d(x_1,x_0).$$

Equivalently,

$$d(x^* x_{k+1}) \leq \frac{q}{1-q} d(x_{k+},x_k).$$

and

$$d(x^*,x_{k+1}) \leq qd(x_k,x^*).$$

The smallest such value of q is called the Lipschitz Constant.

Definition 6.3.2 ***Cluster point.*** *Let E be a metric space with matric d. let $\{x_k\}$ be a sequence in E. We say that " x is a cluster point" of the sequence $\{x_k\}$ if one of the following three equivalent conditions is satisfied.*

1. *x is the limit point of a subsequence $\{x_{k_p}\}_p$ of the sequence $\{x_k\}_k$*
2. *for all $\varepsilon > 0$ and n, there exists $m \geq n$ such that $d(x, x_m) \leq \varepsilon$,*
3. *for all n,x belongs to the closure of the set say, $A_n = \{x_m\}$, $m \geq n$*

Definition 6.3.3 ***Compact set.*** *We say that a subset K of E is "compact" if every infinite sequence $\{x_k\}$ of elements x_k of K has at least one cluster point belonging to K.*

We now state, without proof, the Brouwer[19] and Schauder's fixed point theorems which is the origin of most of the theorems of nonlinear analysis.

Theorem 6.3.2 ***(Brouwer)*** *Let $T : E \to E$ be a continuous mapping from a nonempty, compact convex set $E \subset \mathbb{R}^n$ into itself, then there is a $x^* \in E$ such that $x^* = Tx^*$(i.e., x^* is a fixed point of the mapping T.)*

Theorem 6.3.3 ***(Schauder)*** *Let T be a continuous and compact mapping of a Banach space X into itself, such that the set*

$$\{x \in X : x = \lambda Tx \quad \text{for some} \quad 0 \leq \lambda \leq 1$$

is bounded. Then T has a fixed point.

6.4 R. Krawczk's Version of Newton's Method

R. Krawczyk's (1969) presented a version of Newton's method that avoids the inversion of interval matrices[8]. This method shares the nice properties of Newton's method, but may have slightly slower convergence. We start by seeking at an interval X containing a root x^* of f. Let $y \in X$ and let Y be an arbitrary matrix in $\mathbb{R}^{n\times n}$. Then we have

$$K(y,X) = y - Yf(y) + \{I - YF'(X)\}(X - y) \tag{6.7}$$

Here, $I = diag(1,\ldots,1)$ is the unit matrix and we introduced the so-called Krawczyk operator $K(y,X)$. Suppose a region (interval vector) X_0 that contains a solution of nonlinear systems[11] then the Krawczyk operator is given as for $k = 0,1,2,3,\ldots$

$$X_{k+1} = X_k \cap K(X_k) \tag{6.8}$$

where

$$K(X_k) = m(X_k) - Y_k F(m(X_k)) + R_k(X_k - m(X_k))$$

with $R_k = \{I - Y_k F'(X_k)\}$, for $k = 0,1,2,3,\ldots$. generates a nested sequence of intervals all of which contain the root x^*.

Theorem 6.4.1 *If there exists an x in the interval X and a matrix* $Y \in \mathbb{R}^{n\times n}$ *such that* $X \cap K(x,X) = \Phi$, *then there is no root in X.*

Lemma 6.4.2 *If* $P(x) = x - Yf(x)$ *maps X into itself, then* $f(x) = 0$ *has a solution in X.*

Proof: The continuity of P follows from that of f. Since P maps the convex, compact set X into itself, P has a fixed point in X by the Brouwer's fixed point theorem. From the non-singularity of Y, a fixed point of P is a solution of $f(x) = 0$ and the lemma is proved.

Lemma 6.4.3 *If A is an interval matrix and X is an interval vector, then*

$$w(A(X-m(X))) \leq ||A||w(X).$$

Proof: Let X and Z be intervals.

From $w(Z(X-m(X))) = |Z|w(X)$ and $w(X++Z) = W(X)+w(Z)$ it follows that

$$w(A(X-m(X))) = \max_i w\left(\sum_{j=1}^{n} A_{ij}(X-m(X))\right) = \max_i \left(\sum_{j=1}^{n} w(A_{ij}(X-m(X)))\right) = \max_i \sum_{j=1}^{n} |A_{ij}|w($$

and

$$w(X) = \max_j W(X_j) \geq w(X_j),$$

We have

$$w(A(X-m(X))) \leq \max_i \sum_{j=1}^{n} |A_{ij}|w(X) = ||A||(X).$$

Theorem 6.4.4 *(R.E.Moore's Theorem) Suppose that a region(interval vector) X_0, a point y_0 in X_0 ,and a real matrix Y_0 have been found such that (i.) $K(X_0) \subseteq X_0$. Then, there is a solution $x \in X_0$ to the system $f(x) = 0$. Consider the algorithm*

$$X_{k+1} = X_k \cap K(X_k), \quad k = 0,1,2,3,\ldots$$

$$K(X_k) = y_k - Y_k f(y_k) + \{I - Y_k F'(X_k)\}(X_k - y_k)$$

where y_k and Y_k, $k = 1,2,3,\ldots$, are chosen as follows:

$$y_{(k)} = m(X_{(k)})$$

$$\text{and } Y_{(k)} = \begin{cases} Y, & \text{an approximation to } [mF'(X_{(k)})]^{-1} \text{ for } k=0 \text{ and for } k>0 \\ & \text{if } \left\|I - YF'(X_{(k)})\right\| \leq \left\|R_{(k-1)}\right\|, \\ Y_{(k-1)}, & \text{otherwise.} \end{cases}$$

with $r_k = \|I - Y_k F'(X_k)\|, \quad k = 0,1,2,3,\ldots.$

If the condition (ii). $r_0 < 1$ is satisfied, then there is a unique solution x to $f(x) = 0$ in X_0 and the following hold:

1. $x \in X_k \subseteq X_{k-1}$ *for* $k = 1,2,3\ldots.$

2. $w(X_k) \leq r_0^k w(X_0).$

Thus, $\{X_k\}$ *is a nested sequence of interval containing and converging at least linearly to the unique solution* x *in* X_0.

Proof: From the definition of P in the above lemma, we have

$$\begin{aligned} P(x) &= x - Yf(x) \\ &= y - Yf(y) + x - y - Y(f(x) - f(y)) \end{aligned}$$

Now using the mean value theorem,

$$f(x) = f(y) + \sum_{j=1}^{n} \left.\frac{\partial f}{\partial y_j}\right|_{y=y} (x_j - y_j)$$

We have

$$f(x) - f(y) \in F'(X)(x-y) \quad \text{for all} \quad x, y \in X$$

Hence we have

$$x - y - Y(f(x) - f(y)) \in x - y - YF'(X)(x-y) \in (I - YF'(X)(X-Y)$$

Thus we have $P(x) \in y - Yf(y) + I - YF'(X)(X-y) \equiv K(x,X)$ That is $P(x) \in K(X)$ for all $x \in X$. If $K(X_0) \subseteq X_0$,P maps X_0 into itself; and by the above lemma $f(x) = 0$ has a solution in X_0. If $f(x) = 0$ for $x \in X$,then $x = P(x)$ is also in $K(X)$. Since $X_k = X_{k-1} \cap K(X_{k-1}$, the solution x in X_0 is also in X_k for all $k = 0, 1, 2, 3, \ldots$ which proves (i). It remains to show that the inequality 2. holds. From the definition of the algorithm, we have $w(X_{k+1}) \leq w(K(X_k))$. But we have

$$w(K(X_k)) = w\left(y_k - Y_k f(y_k) + \{I - Y_k F'(X_k)\}(X_k - y_k)\right)$$

where

$$y_k = m(X_k) \quad \text{and} \quad Y_k = [m(F'(X_k))]^{-1}$$

Since

$$w(y_k) = 0, \quad w(Y_k f(y_k)) = 0$$

we have

$$w(K(X_k)) = w\{I - Y_k F'(X_k)\}(X_k - y_k) \leq r_k w(X_k)$$

from lemma 2. Hence we have

$$w(K(X_k)) \leq r_k w(X_k)$$

Thus we obtain

$$\begin{aligned} w(X_k) &\leq w(K(X_{k-1})) \\ &\leq r_{k-1} w(X_{k-1}) \\ &\leq r_{k-1} w(K(K_{k-2})) \\ &\leq r_{k-1} r_{k-2} w(X_{k-2}) \\ &\leq r_{k-1} r_{k-2} \cdots r_0 w(X_0) \end{aligned}$$

Since $\{r_k\}$ is non-increasing by constructive, we have $w(X_k) \leq r_0^k w(X_0)$ which proves (ii)

6.4.1 Numerical Example

Let us consider the following system

$$\begin{aligned} f_1(x_1, x_2) &= x_1^2 + x_2^2 - 1 = 0 \\ f_2(x_1, x_2) &= x_1^2 - x_2 = 0 \end{aligned} \tag{6.9}$$

For an interval vector $X = (X_1, X_2)$, we can take $F(X) = \begin{pmatrix} X_1^2 + X_2^2 - 1 \\ X_1^2 - X_2 \end{pmatrix}$ and $F'(X) = \begin{pmatrix} 2X_1 & 2X_2 \\ 2X_1 & -1 \end{pmatrix}$ as the respective interval extensions of the vector function

$$\mathbf{f(x)} = \begin{pmatrix} f_1(x_1,x_2) \\ f_2(x_1,x_2) \end{pmatrix} = \begin{pmatrix} x_1^2+x_2^2-1 \\ x_1^2-x_2 \end{pmatrix}$$ and the Jacobian matrix

$$f'(x) = \begin{pmatrix} 2x_1 & 2x_2 \\ 2x_1 & -1 \end{pmatrix} = J(x) \text{ (say).}$$

Here, $x^* = (0.7861513377, 0.618033988)$ is an exact solution of the above system correct to nine decimal places.

Let $x_0 = (0.80, 0.62)$ be an approximate solution obtained by some non-interval method. We can use Krawczyk's method both to test whether there is a root close to x_0, say in $X_0 = ([0.7, 0.9], [0.5, 0.7])$ containing the exact root and to get better approximation if this is satisfied.

To apply Krawczyk algorithm with X_0 we need the following results :

$$f(y_0) = \begin{pmatrix} 0.0244 \\ 0.02 \end{pmatrix},$$

$$f'(y_0) = \begin{pmatrix} 1.6 & 1.24 \\ 1.6 & -1 \end{pmatrix},$$

and also

$$F'(X_0) = \begin{pmatrix} [1.53, 1.67] & [1.17, 1.31] \\ [1.53, 1.67 & [-1, -1] \end{pmatrix}$$

For Y, we may take the approximation of $[f'(y_0)]^{-1}$:

$Y_0 = [f'(y_0)]^{-1} = \begin{pmatrix} 0.279017857 & 0.345982142 \\ 0.446428571 & -0.446428571 \end{pmatrix}$. With this results using rounded interval arithmetic, we compute that for Krawczyk transformation

$$K(X_0) = y_0 - Y_0 f(y_0) + \{I - Y_0 F'(X_0)\}(X_0 - y_0)$$

$$K(X_0) = \begin{pmatrix} [0.7657, 0.8064] \\ [0.5872, 0.6481] \end{pmatrix} \subseteq X_0$$

and on the other hand

$$\|R_0\| = \|I - Y_0 F'(X_0)\| = \begin{pmatrix} [-0.1250, 0.1250] & [-0.0446, 0.670] \\ [-0.1784, 0.1784] & [-0.704, 0.1080] \end{pmatrix}$$

We get $\|R_0\| = 0.2864 < 1$, and Now for X_0, we have

$$X_1 = K(X_0) \cap X_0 = \begin{pmatrix} [0.784, 0.788] \\ [0.615, 0.621] \end{pmatrix} \subseteq X_0$$

In the table below we will find the results from this algorithm with $y_k = m(X_k)$ and the stopping criterion $w(X_k) \leq 10^{-6}$.

k	$X_k, 1$	$X_k, 2$	$w(X_k)$
1	[0.7657323, 0.8063645]	[0.5872375, 0.6480857]	$6.083e-02$
2	[0.7815712, 0.7907271]	[0.6111227, 0.6249431]	$1.38e-02$
3	[0.7851161, 0.7871866]	[0.6164709, 0.6195970]	$3.13e-03$
4	[0.7859172, 0.7863856]	[0.6176805, 0.6183875]	$7.07e-04$
5	[0.7860984, 0.7862044]	[0.6179540, 0.6181140]	$1.60e-04$
6	[0.7861394, 0.7861634]	[0.6180159, 0.6180521]	$3.62e-05$
7	[0.7861486, 0.7861541]	[0.6180298, 0.6180381]	$8.18-06$
8	[0.7861507, 0.7861520]	[0.6180330, 0.6180350]	$1.85-06$
9	[0.7861512, 0.7861516]	[0.6180337, 0.6180342]	$4.18e-07$

The true Krawczyk method ()with the choices() converges considerably faster ; quadratically.

k	$X_k, 1$	$X_k, 2$	$w(X_k)$
1	[0.7657323, 0.8063645]	[0.5872375, 0.6480857]	$6.083e-02$
2	[0.7850995, 0.7872034]	[0.6164672, 0.6196009]	$3.13e-03$
3	[0.7861485, 0.7861542]	[0.6180298, 0.6180382]	$8.35-06$
4	[0.7861513, 0.7861514]	[0.61803379, 0.6180340]	$5.93e-11$

We have seen seen that $w(X_{(k+1)}) \leq \|R_{(k)}\| w(X_{(k)})$ and $\{w(X_{(k)}\} \to 0$ are satisfied

In 1965 **Minoru Urabe** established an existence and uniqueness theorem that helps us verify the existence and uniqueness of an exact solution and to know

the error bound to approximate solution of the nonlinear system. His existence theorem assumes that all the computations are to be carried out in real number system exactly. We now state the following theorem of Minoru Urabe[24]

Theorem 6.4.5 *Let f(x) be continuously differentiable function on the domain* $D \subset \mathbb{R}^n$. *Let* $x_0 \in D$ *and suppose* $J(x_0)$ *be regular. Also suppose that the following three conditions are satisfied for a positive number* δ *and a non-negative* k $(0 \leq k < 1)$.

1. $\Omega_\delta = \{x \in \mathbb{R}^n : \|x - x_0\| \leq \delta\} \subset D$,
2. $\|J(x) - J(x_0\| \leq \dfrac{k}{M}(x \in \Omega_\delta)$,
3. $\dfrac{Mr}{1-k} \leq \delta$, *where* $\|f(x_0\| \leq r$, *and* $\left\|J^{-1}(x_0)\right\| \leq M$.

Suppose $x^* \in \Omega_\delta$ *be the unique solution of the equation* $f(x) = 0$ *and* $J(x^*)$ *be regular.*
Then the error estimation $\|x_0 - x^*\| \leq \dfrac{Mr}{1-k}$ *is satisfied.*

6.5 Safe Starting Regions For Iterative Methods

R.E Moore and S.T. Jones discussed a search procedure based on interval computation for finding safe starting regions in n-dimensions for iterative methods for solving systems of nonlinear equations. The procedure can search an arbitrary n-dimensional rectangle for a safe starting region for quadratically convergent iterative method. We have seen that the procedure is more powerful than continuation methods[12] .

There are many good methods for iterative solution of systems of nonlinear equations([7],[16]). For any such method, however, the problem remains of finding a safe starting point: an initial approximation from which the iterates will

converge to a solution. R.E. Moore developed a technique of finding such a safe starting region with a particular iterative method in mind, an interval version of Newton's method due to R.Kraczyk[8] which possesses computationally verifiable sufficient conditions for existence and convergence.It will become clear how the procedure could be used, mutatis mutandis, with other iterative methods.

By recasting the problem of solving a system in the form of a fixed point problem, we can make use of sufficient conditions for the existence of a solution based on fixed point theorems such as those of Brouwer[4] or Schauder. The interval binary search procedure to be described in this section can start with an arbitrary n-dimensional rectangle and an arbitrary algorithm for finding zeros of a system. As a stopping criterion for the search procedure we can use: satisfaction of the criteria for existence and convergence appropriate to the algorithm to which the search procedure is to be applied. When supplied with a suitable algorithm and a specific system of equations, the search procedure will be ready to search an arbitrary rectangle B. It will terminate in a finite number of steps with one of the following three results: a subregion of B which is a safe starting region for the algorithm and we will discover that there are no solutions in B or a list of small sub-regions of B which would require higher precision machine arithmetic to yield a result of one of the first two types.

In addition to finding safe starting regions for iterative methods, the search procedure could be used to find safe starting regions for algorithms based on simlicial subdivisions such as those of Scarf , Stenger , and Todd providing they can begin with a safe starting region in the form of an n-dimensional rectangle. The conditions to be tested on a region for such algorithms may be based on Miranda's theorem[3] or on the Brouwer fixed point theorem directly.

Now we seek a solution of the nonlinear system of equations in n unknowns

$$f_i(x_1, x_2, \ldots, x_n) = 0 \quad i = 1, 2, 3, \cdots, n \tag{6.10}$$

where $f_1, f_2, \ldots, f_n$ are real-valued functions and continuously differentiable in the open domain $D \subseteq \mathbb{R}^n \to \mathbb{R}^n$. Equation (6.10) also can be written in the vector form as

$$\mathbf{f}(\mathbf{x}) = \mathbf{0} \tag{6.11}$$

with $x = (x_1, x_2, \ldots, x_n)$ and $f = (f_1, f_2, f_3, \ldots, f_n)^T$.
Suppose we could compute the exact range of values,

$$f_i(X) = [\underline{f_i}(X), \overline{f_i}(X)], \quad \underline{f_i}(X) = \min_{x \in X} f_i(X), \overline{f_i}(X) = \max_{x \in X} f_i(X)$$

of each f_i when x lies in an n-dimensional rectangle

$$X = \{x : \underline{x_i} \leq x_i \leq \overline{x}_i, i = 1, 2, \cdots, n\} \subseteq D.$$

Then we could use the following simple cyclic n-dimensional bisection procedure. Suppose X_0 is any n-dimensional rectangle in D for which $f_i(X_0)$ includes the number zero for all $i = 1, 2, 3 \cdots$.Then X_0 may contain a solution to the above system . However if there is some i for which $f_i(X_0)$ does not contain zero, then there is no solution in X_0. Either $\underline{f_i}(X_0) > 0$ or $\overline{f}_i(X_0) < 0$. This is a nonexistence theorem. We can make it into an exclusion principle.

We can perform a cyclic sequence of bisections of X_0 as follows.
We will bisect X_0 in co-ordinate x_1. Exclude half of X_0 for which $o \notin f_i(X_0)$ for some i. We will construct list (1) from the half or halves of $X_{(0)}$ for which 0 is in the range of value of f_i for all $i = 1, 2, \cdots, n$. Now we will bisect, in coordinate direction x_2, each region in list (1); and construct list (2) from all resulting regions X for which $0 \in f_i(X), i = 1, 2, \cdots, n$. Continue in this way with list $(kn + j)$ resulting from bisection in coordinate direction x_i of the regions in list $(kn + j - 1)$.

Any solutions which the system (6.10) may have will lie in regions contained in list(m) for any m. If the list (m) becomes empty for some m, then there are no solutions of (4.10) in X_0.

In one dimension this procedure, applied to a function which crosses the axis

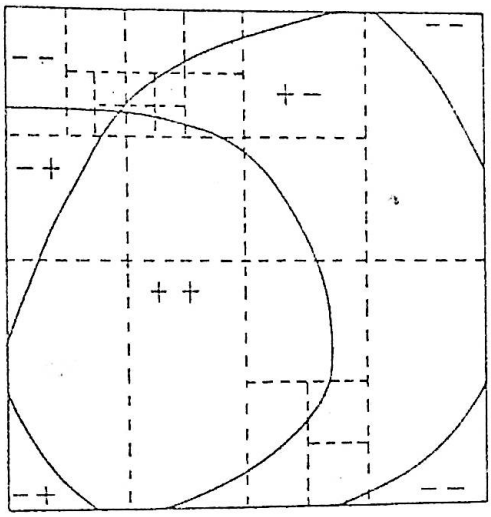

Figure 6.1: Figure: Bisection in two dimensions.

once in $[a,b]$, will produce the same sequence of subintervals of $[a,b]$,as the usual bisection method. One the other hand, unlike the usual bisection procedure, this bisection procedure can also find multiple zeros.

A two dimensional example may help to make the general procedure clear. In the above figure. twenty-three bisections carried out just as described produces a smaller rectangle $\frac{1}{16}$th the diameter of the initial rectangle containing the solution. The solid curves represent loci of zeros of f_1 or f_2 separately. The regions in between are marked ++,+-,-+ or – according to the signs of f_1 and f_2 in those regions. At the stage shown, only the small rectangle containing the intersection of $f_1 = 0$ with $f_2 = 0$ need be further bisected. Thus a rectangle need be further bisected only when it contains a zero of f_1 and a zero of f_2. Even when we cannot compute the exact ranges of values of the functions

$f_1, f_2, \cdots, f_n$ it is still possible to use bisection procedures in n dimensions.If interval extensions $F_1, F_2, \cdots, F_n$, ([20][11]) of the functions $f_1, f_2, \cdots, f_n$ are available, then we can compute intervals $F_1(X), \cdots, F_n(X)$ containing the ranges of values of $f_1, f_2, \cdots, f_n$ for x in an n-dimensional rectangle X.If any one of the intervals $F_i(X), i = 1, 2, \cdots, n$ does not contain the number 0,then there is no solution of (6.10)in X.Since $F_i(X), i = 1, 2, \cdots, n$ may be wider than the actual range of values of f_i for x in X,this may require more bisections than if we were able to compute the exact range of values of $f_1, \cdots, f_n$. Nevertheless,such an approach can be used effectively,as will be shown in this paper, to find safe starting regions for iterative methods.

As a stopping criterion for the bisection procedure we can make use of computationally verifiable sufficient conditions for convergence of a chosen iterative method to a solution of (6.10).Thus, the bisection procedure will terminate when a safe starting region is found for a more efficient iterative method.

An interval search procedure for R. Krawczyk's method.

We assume that $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable in the open set D. We assume that the interval extensions F_i, F'_{ij} of the functions $f_i, i = 1, \cdots, n$in (1) and of the coefficients f'_{ij} of the Jacobian matrix of the system. Denote by $I^n[D]$ the set of n-dimensional rectangles contained in D. An element of$I^n[D]$ can be represented by an interval vector $X = (X_1, \cdots, X_n)$ where X_i is a closed bounded interval of real numbers.

The interval valued functions F_i, F'_{ij} are assumed to have the following properties.

1. $F_i, F'_{ij} (i, j = 1, 2, \cdots, n)$ are interval valued functions defined on $I^n[D]$.

2. $f_i(x) = F_i(x), f'_{ij}(x) = F'_{ij}(x)$ for all x in D

3. (inclusion monotonicity[20]) For all X, B in $I^n[D]$, $X \subseteq B$ implies $F_i(X) \subseteq F_i(B)$ and $F'_{ij}(X) \subseteq F'_{ij}(X)$ for all $i, j = 1, \cdots, n$

We know the following notations. For an interval $[a,b] = \{x : a \leq x \leq b\}$, the magnitude of an interval is defined by $|[a,b]| = \max(|a|,|b|)$. The width of an interval is given by $w([a,b]) = b-a$. The midpoint of an interval $[a,b]$ is given by $\frac{a+b}{2}$. We define

$$\|X\| = \max_i(|X_i|), \quad w(X) = \max_i w(X_i).$$

For an interval matrix A , we define the norm for an interval matrix

$$\|A\| = \max_i \sum_{j=1}^{n} |A_{ij}|$$

and $m(A)$=the real matrix with components $m(A_{ij})$.

R.Krawczyk introduced another form of interval version of Newton's method which does not require the inversion of an interval matrix is given by[8]with the Krawczyk operator

$$K(X) = y - Yf(y) + \{I - YF'(X)\}(X - y) \tag{6.12}$$

where y is a point chosen from X and Y is an arbitrary non-singular real matrix.

Suppose a region (interval vector) X_0 that contains a solution of (6.10) then the Krawczyk algorithm is given as for $k = 0, 1, 2, 3, \ldots$

$$X_{k+1} = X_k \cap K(X_k) \tag{6.13}$$

where

$$K(X_k) = m(X_k) - Y_k F(m(X_k)) + R_k(X_k - m(X_k))$$

with $R_k = \{I - Y_k F'(X_k)\}$, for $k = 0, 1, 2, 3, \ldots$ (I is the identity matrix)

where $y_{(k)}$ and Y_k, $k = 1, 2, 3 \ldots$, are chosen as follows :

$$y_k = m(X_k)$$

$$\text{and } Y_k = \begin{cases} Y, & \text{an approximation to } [mF'(X_k)]^{-1} \text{ for } k=0 \text{ and for } k>0 \\ & \text{if } \|I - YF'(X_k)\| \leq \|R_{k-1}\|, \\ Y_{k-1}, & \text{otherwise.} \end{cases}$$

The algorithm(6.13) satisfies the following properties:

1. $X_{k+1} \subseteq X_k$ for all $k = 0, 1, 2, 3, \cdots$.

2. If X_0 contains a solution of (6.10), then so does $K(X_0)$.Thus,if $K(X_0 \cap X_0$ is empty ,then there is no solution in X_0. This is another computationally verifiable sufficient condition for nonexistence of a solution in an n-dimensional rectangle X_0 (in addition to $0 \notin F(X_0$.

3. The sequence $\{\|R_k\|\}$ is non-increasing and$w(X_{k+1}) \leq \|R_k\| w(X_k)$.

4. If $K(X_0) \subseteq X_0$, then there is a solution of (6.10) in X_0. This is a computationally verifiable sufficient condition for existence of a solution of (6.10) in an n-dimensional rectangle X_0.

5. If $K(X_0) \subseteq X_0$ and $\|R_0\| \leq 1$, then there is a unique solution of (6.10) in X_0. Furthermore, the solution is in X_k for all $k = 0, 1, 2, 3, \cdots$ and

$$\{w(X_k\} \to 0.$$

This is a computationally verifiable sufficient condition for convergence of the iterative algorithm (6.13)to a solution of (6.10). The convergence is at least linear ; in fact

$$w(X_k) \leq \left\{ \prod_{i=0}^{k-1} \|R_i\| \right\} w(X_0) \leq \|R_0\|^k w(X_0).$$

Under certain conditions, the convergence is quadratic. Assuming the conditions of property 5) above, we have the following:

Theorem 6.5.1 *If the Jacobian of the system (6.10) is nonsingular in a neighborhood of the solution (for instance, if* $\|[F'(X_0)]^{-1}\| \leq b$*); if the elements of F satisfy*

a Lipschitz condition: for some $L > 0$. $w(F(X)_{ij} \leq Lw(X)$ *for all* $X \subseteq X_0$ *and all* $i, j = 1, 2, \ldots, n$ *and if* Y_k *is sufficiently good approximation to* $[m(F(X_k))]^{-1}$ *so that* $Y_k = [m(F(X_k))]^{-1} + E_k$ *with* $\|E_k\| \leq Cw(X_k)$, *then the convergence of the Krawczyk's algorithm is quadratic; that* $w(X_{k+1}) \leq q[w(X_k)]^2$ *for some* q *independent of* k.

Proof. We have

$$\|R_k\| \leq \|E_k\|\|F(X_k)\| + \frac{1}{2}\|[[m(F(X_k))]^{-1}\|nLw(X_k) \leq \{C\|F(X_0)\| + \frac{1}{2}bnL\}w(X_k).$$

Therefore,

$$w(X_{k-1}) \leq w(K(X_k)) \leq \|R_k\|w(X_k) \leq \{C\|F(X_0)\| + \frac{1}{2}bnL\}[w(X_k)]^2$$

We can take $q = C\|F(X_0)\| + \frac{1}{2}bnL$ and the theorem is proved.
We will present a search procedure for finding safe starting regions for the Krawczyk's algorithm . Let Bbe an arbitrary n-dimensional rectangle contained in the domain D of the functions in the system (6.10). We do not require the non-singularity of the Jacobian matrix f'in B.

We begin by noting that for any region $X \subseteq B$ exactly one of the following conditions will hold:

1. X satisfies the existence and convergence criteria of Krawczyk's algorithm.

2. X satisfies one of the nonexistence criteria of Krawczyk's algorithm

3. X satisfies neither 1 nor 2.

Determination of which of the above conditions holds is called analysis of the region X.

The search procedure consists of a recursive application of this analysis, beginning with the region $X = B$. At each level, we analyze the region X. If X satisfies condition 1, we designate it a safe starting region for the solution of the system (6.10). If X satisfies condition 2 , then no solution of (6.10) is contained in X and this region is excluded from further consideration. If condition 3 holds, we bisect X, if possible, in some appropriately chosen coordinate direction and select one of the two resulting half-regions for analysis at the next level. If condition 3 holds and it is not possible to bisect X, there may still be a solution of (6.10) in X. In this case we add X to a list of regions too small for further analysis and continue the search as if this region had been excluded.Since B is of finite dimension and its components are of finite width, there can be only a finite number of such regions obtained by bisection.

Thus, the search procedure, described in detail below, will in a finite number of steps, do one of the following three things:

1. find a safe starting region X within B for convergence of thekrawczyk's algorithm to a solution of (6.10);

2. discover that there are no solutions of (6.10) in B;

3. terminate with a list of small regions within B which might contain solutions of (6.10).

Search procedure for algorithm (6.13):

List T is the list of subregions of B yet to be tested.

List P is the list of subregions of B which may contain a solution to (6.10) but which are too small for further analysis.

Unless otherwise indicated, Step $m+1$ follows step m.

Step 1. (Initialization). Set list T to empty; set list P to empty; set X to B.

Step 2. Compute $F(X)$.

Step 3. (Exclusion). If $0 \notin F(X)$ go to step 11.

Step 4. Compute $Y = [m(F'(x))]^{-1}$; if not possible, go to step 9.

Step 5. Compute $||R||$ and $K(X)$.

Step 6. (Exclusion). If $X \cap K(X)$ is empty, go to step 11.

Step 7. (Existence). If $K(X) \subseteq X$,then X(and also$K(X)$) contains a solution,continue; if not, go to step 9.

Step 8. (Test$||R||$, If $||R|| < 1$,then X is a safe starting region X_0 for the algorithm (6.13)-terminate search-; otherwise set B to $K(X)$ and go to step 1.

Step 9. (Bisection). Bisect X according to rules described next section; if no further bisection is possible, add X to list P and go to step 11.

Step 10. Set X to half-region selected according to bisection rules; add remaining half-region to head of list T; go to step 2.

Step 11. (Test list T). If list T is empty goto step 12; otherwise set X to region at head of list T, delete this region from list T and go to step 2.

Step 12 (Test list P). If list P is empty, terminate with no solution in B; otherwise print list P and terminate.

We now discuss bisection rules for step 9 of the search procedure. When bisection of a region X is to be carried out, there are two decisions to be made:

1. In which coordinate direction we should be bisected X ?

2. Which bisected half of X we should be searched first ?

Note that the procedure always saves the untested half of bisected region for possible examination later.

If X cannot be further bisected within the limited machine number precision being used, then we can indicate "possible solution in X" and go to step 11.

We only bisect a region X after it has been determined that: $0 \in F(X)$; and either

1. we can not find an approximation $Y = [m(F'(x))]^{-1}$;

2. $X \cap K(X)$ is not empty, but $K(X)$ is not contained in X. Thus we only bisect a region X when it still might contain a solution. The number of bisections required to find a safe starting region will depend on the rules used to choose coordinate directions and bisected halves.

In order to guarantee that $X_{(0)}$ is a safe starting region for algorithm (6.13), we need

$$||R_0|| = ||I - Y_0 F'(X_0)|| < 1 \text{and} \quad K(X_0) \subseteq X_0$$

We will assume that Y_0 has been found as an approximation to $[m(F'(x))]^{-1}$. We can write

$$F'(X_0) = m(F_0') + W$$

where

$$W_{ij} = \frac{1}{2}[-1,1]w(F'(X_0)_{ij}).$$

Then $||R_0||$ is approximately $||Y_0 W||$. This motivates our choice of a rule for selecting a coordinate direction, in which to bisect $X^{(0)}$. We attempt to decrease $||R||$ by decreasing

$$w[F'(X)] = \max_{ij} w(F'(X)_{ij})$$

To choose a coordinate direction for the bisection of X:

1. we find a pair i, j for which $w(F'(X)_{ij} = w(F'(X))$.

2. we bisect X in each of the coordinate directions in turn, which occur in $F'(X)_{ij}$ and choose the first one for which

$$w\{F'(X_1)_{ij} \cup F'(X_2)_{ij}\}$$

is minimum,where $X = X_1 \cup X_2$ and X_1, X_2, are the halves of X.

Having chosen, in this way, a bisection direction, we select one of the resulting halves X_h, h=1,2, for which

$$\sum_{i=1}^{n} |m(F_j(X_h))| \quad \text{is minimum.}$$

This choice is motivated by an attempt to select a bisected half of X which is most likely to contain a solution of (6.10).It is symmetry test-selecting a half of X on which the values of the functions $f_1, f_2, \ldots, f_n$ seem to be closest to zero" on the average".

Chapter 7

Computation of Interval Solutions to Nonlinear Systems

7.1 Computation of Interval solution to a polynomial equation

In this section we apply M. Urabe's theorem to R. Krawczyk's algorithm to find an interval solution enclosing a real root of a polynomial(nonlinear) equation. Let us consider the third degree polynomial equation

$$f(x) = x^3 - 3x + 3 = 0 \tag{7.1}$$

This equation has only one real root $x^* = \alpha = -2.103803402, \cdots$.

Let $x_0 = z = -2.11$ be the approximate real solution.

URABE'S THEOREM:

$r = ||F(z)|| = 0.06$,$||J^{-1}(z)|| = \frac{1}{F'(z)} = \frac{1}{10.35} = 0.096618 \cdots \leq 0.0967 = M$

1. in the closed domain $\Omega_\delta = \{x \in \mathbb{R}; ||x - x_0|| \leq \delta\}$

   $x+z = (x-z)+2z$leads to $|x+z| \leq 2|z|+\delta$ and

$$||J(x) - J(z)|| = 3|x+z||x-z| \leq (6|z|+3\delta)\delta$$

2. if $(6|z|+3\delta)\delta \leq \frac{k}{M}(x \in \Omega_\delta$ is satisfied.

   So $k \geq (6+|z|+3\delta)\delta M = (12.66+3\delta) \times 0.0967 = 1.224222\delta + 0.290\delta^2$

3. But $\frac{M\delta}{1-k} \leq \delta$ is satisfied if

   $\delta > \delta(1-k) \geq Mr = .005802$

   As the closed domain $\Omega_\delta = \{x \in \mathbb{R} : ||x - x_0|| \leq \delta\}$ is the δ-neighborhood of $x_0 = -2.11$. If we choose $\delta = 0.10$, then

$$\Omega_\delta = \{x \in \mathbb{R} : |x-(-2.11)| \leq 0.10\} = [-2.11, -2.01] \subset D[-3, -2]$$

   Therefore $k \geq 1.224222(0.10) + 0.2901(0.10)^2 = 0.1253232$

   So, if we choose $k = 0.13$, the all the conditions of Urabe's theorem are satisfied. In the closed domain Ω_δ, the $x^* = \alpha = -2.103803402\cdots$ is the unique solution of $f(x) = 0$.

   The error estimation gives

$$||z - x^*|| \leq \frac{Mr}{1-k} = 0.0066\cdots < .007$$

   That is $-2.11 - 0.007 < x^* < -2.11 + 0.007$.

we assume that $X_0 = [-2.117, -2.103]$ be the interval solution of (7.1). Here $y_0 = m(X_0) = -2.110$. Now we are using Horner's algorithm to Krawczyk's method in this interval .

Horner's Alogorithm

$$b^{-1}[k] = a_k, k = 0, 1, 2, ...n(n = 3) \qquad (7.2)$$

for j=0,1,2,...

$$b^{(j)}[0] = a_0, \quad b^{(j)}[k] = b^{(j-1)}[k] + b^{(j)}[k-1]y, (k = 1,2,...,n-j) \tag{7.3}$$

$$b^{(0)}[n] = f(y), \quad b^{(1)}[n-1] = f'(y) \tag{7.4}$$

Using the Horner's algorithm, we have the following tables.

Table 1: Computation of $f(y_0)$ and $f'(y_0$ by Horner's method

	$k=0$	$k=1$	$k=2$	$k=3$
$b^{(-1)}[k]$	1	0	-3	3
$b^{(0)}[k-1]y$		-2.110	4.452	-3.064
$b^{(0)}[k]$	1	-2.110	1.452	-0.064
$b^{(1)}[k-1]y$		-2.110	8.904	
$b^{(1)}[k]$	1	-4.220	10.356	

Table 2 : Computation of $F'(X_0)$ by Horner's method

	$k=0$	$k=1$	$k=2$	$k=3$
$b^{(-1)}[k]$	[1,1]	0	[-3,3]	[3,3]
$b^{(0)}[k-1]y$		[-2.117,-2.103]	[4.422,4.482]	[-3.138,-2.990]
$b^{(0)}[k]$	[1,1]	[-2.117,-2.103]	[1.422,1482]	[-0.138,0.010]
$b^{(1)}[k-1]y$		[-2.117,-2.103]	[8.845,8.964]	
$b^{(1)}[k]$	[1,1]	[-4.234,-4.206]	[10.267,10.446]	

Table 3: Computation of $f(y_1)$ and $f'(y_1$ by Horner's method

	$k=0$	$k=1$	$k=2$	$k=3$
$b^{(-1)}[k]$	1	0	-3	3
$b^{(0)}[k-1]y$		-2.1035	4.4247	-2.9968
$b^{(0)}[k]$	1	-2.1035	1.4247	-0.0032
$b^{(1)}[k-1]y$		-2.1035	8.849	
$b^{(1)}[k]$	1	-4.2070	10.2741	

Table 4 : Computation of $F'(X_1$ by Horner's method

	$k=0$	$k=1$	$k=2$	$k=3$
$b^{(-1)}[k]$	[1,1]	0	[-3,3]	[3,3]
$b^{(0)}[k-1]y$		[-2.117,-2.103]	[4.4226,4.4269]	[-3.0022,-2.9917]
$b^{(0)}[k]$	[1,1]	[-2.117,-2.103]	[1.4226,1.4269]	[-0.0022,0.0083]
$b^{(1)}[k-1]y$		[-2.117,-2.103]	[8.8452,8.8537]	
$b^{(1)}[k]$	[1,1]	[-4.204,-4.206]	[10.2678,10.2806]	

From table-1,we get

$$f(y_0) = -0.064, \quad f'(y_0) = 10.356$$

and also

$$Y_0 = [f'(y_0)]^{-1} = 0.09656$$

From table-2,we get,

$$F'(X_0) = [10.267, 10.446]$$

As

$$K(X_0) = y_0 - Y_0 f(y_0) + \{I - Y_0 F'(X_0)\}(X_0 - y_0)$$

$$= [-2.104, -2.103] \subseteq X_0 = [-2.117, -2.103],$$

and on the other hand

$$r_0 = ||I - Y_0 F'(X_0)|| = ||[-0.0087, 0.0087]|| = 0.0087 < 1,$$

We have

$$X_1 = K(X_0) \cap X_0 = K(X_0) = [-2.104, -2.103].$$

Again,

$$y_1 = m(X_1) = -2.1035$$

and

$$F'(X_1) = [10.2678, 10.2806].$$

From table-3 we get

$$f(y_1) = 0.0031, \quad f'(y_1) = 10.2741$$

and also

$$Y_1 = [f'(y_1)]^{-1} = 0.09733$$

From table-4 we have

$$F'(X_1) = [10.2678, 10.2806].$$

As

$$K(X_1) = y_1 - Y_1 f(y_1) + \{I - Y_1 F'(X_1)\}(X_1 - y_1)$$

$$= [-2.104, -2.103] \subseteq X_1 = [-2.104, -2.103]$$

then

$$X_2 = K(X_1) \cap X_1 \subseteq X_1$$

It can be noted here that for calculations rounded to three decimal places gives

$$X_1 \subseteq X_0$$

and

$$X_1 = X_2 = X_3 = \cdots$$

On the other hand

$$r_1 = ||I - Y_1 F'(X_1)|| = ||[-0.00062, 0.0007]|| = 0.0007 < r_0 < 1.$$

The initial interval converges to the interval $[-2.104, -2.103]$ rounded to five significant figures that contains the root. Higher precisions may increase the number of iterations.

7.2 Computation of Interval Solutions to Nonlinear Systems.

We have seen that R. Krawczyk introduced an interval version of Newton's method that does not require the inversion of an interval matrix. **Minoru Urabe** established a non interval theorem that deals with existence, uniqueness and convergence of solution to nonlinear system of equations. In this example we will verify[1] computationally the convergence of interval solutions of nonlinear system of equations by applying" **M. Urabe's** " theorem[24]to **R. Krawczyk's algorithm**[8].

We are given a system of n nonlinear equations in n unknowns

$$f_i(x_1, x_2, \ldots, x_n) = 0 \quad i = 1, 2, 3, \cdots, n \tag{7.5}$$

where $f_1, f_2, \ldots, f_n$ are real-valued functions and continuously differentiable in the open domain $D \subseteq \mathbb{R}^n \to \mathbb{R}^n$. Equation (7.5) also can be written in the vector form as

$$\mathbf{f(x)} = \mathbf{0} \tag{7.6}$$

with $x = (x_1, x_2, \ldots, x_n)$ and $f = (f_1, f_2, f_3, \ldots, f_n)^T$.

we assume that both of f and f' have continuous inclusion monotonic interval extensions F and F' defined on an interval contained in D. Therefore the interval extension of the system (7.5) is given as

$$F(X) = 0. \tag{7.7}$$

where $X = (X_1, X_2, \ldots, X_n)$ is an interval vector contained in D and $X_1, \ldots, X_n$ are closed bounded real intervals.

The following are the essential algorithms for nonlinear systems

1. **Newton's Algorithm.**

An interval version of Newton's method to solve (7.5) is introduced in[20], [11] which is given in the following form

$$X_{k+1} = X_k \cap N(X_k) \tag{7.8}$$

with the Newton interval operator N, defined for instance, by

$$N(X) = m(X) - Vf(m(X)) \tag{7.9}$$

where $m(X)$ is the midpoint of the interval vector X and V is an interval matrix containing $[f'(x)]^{-1}$ for all x in X. If there is a zero of f in X_0 then the zero is also in $N(X_0)$ and therefore in X_k, for $k = 0, 1, 2, 3, \ldots$

1. **R. Krawczyk Algorithm.**

R.Krawczyk introduced another form of interval version of Newton's method which does not require the inversion of an interval matrix is given by[8] with the Krawczyk operator

$$K(X) = y - Yf(y) + \{I - YF'(X)\}(X - y) \tag{7.10}$$

where y is a point chosen from X and Y is an arbitrary non-singular real matrix.

Suppose a region (interval vector) $X_{(0)}$ that contains a solution of (7.5) then the Krawczyk algorithm is given as for $k = 0, 1, 2, 3, \ldots$

$$X_{k+1} = X_k \cap K(X_k) \tag{7.11}$$

where

$$K(X_k) = m(X_k) - Y_k F(m(X_k)) + R_k(X_k - m(X_k))$$

with $R_k = \{I - Y_k F'(X_k)\}$, for $k = 0, 1, 2, 3, \ldots$ (I is the identity matrix)
where y_k and Y_k, $k = 1, 2, 3 \ldots$, are chosen as follows :

$$y_k = m(X_k)$$

$$\text{and } Y_k = \begin{cases} Y, & \text{an approximation to } [mF'(X_k)]^{-1} \text{ for } k = 0 \text{ and for } k > 0 \\ & \text{if } \|I - YF'(X_k)\| \leq \|R_{k-1}\|, \\ Y_{k-1}, & \text{otherwise.} \end{cases}$$

The algorithm(7.11) satisfies the following conditions.

1. $X_{k+1} \subseteq X_k$ for all $k = 0, 1, 2, 3, \cdots$.
2. If X_0 contains a solution of (7.5), then so does $K(X_0)$
3. The sequence $\{\|R_k\|\}$ is non-increasing and $w(X_{k+1}) \leq \|Rk\| w(X_k)$.
4. If $K(X_0) \subseteq X_0$, then there is a solution of (7.5) in X_0.
5. If $K(X_0) \subseteq X_0$ and $\|R_0\| \leq 1$, then there is a unique solution of (7.5) in X_0. Furthermore, the solution is in X_k for all $k = 0, 1, 2, 3, \cdots$ and $\{w(X_k\} \to 0$ (Test for convergence)

Theorem 7.2.1 :(M.Urabe) *Let f(x) be continuously differentiable function on the domain $D \subset \mathbb{R}^n$. Let $x_0 \in D$ and suppose $J(x_0)$ be regular. Also suppose that the following three conditions are satisfied for a positive number δ and a non-negative k $(0 \leq k < 1)$.*

1. $\Omega_\delta = \{x \in \mathbb{R}^n : \|x - x_0\| \leq \delta\} \subset D$,

2. $\|J(x) - J(x_0\| \leq \frac{k}{M} (x \in \Omega_\delta)$,

3. $\frac{Mr}{1-k} \leq \delta$, *where* $\|f(x_0\| \leq r$, *and* $\left\|J^{-1}(x_0)\right\| \leq M$

Suppose $x^* \in \Omega_\delta$ *be the unique solution of the equation* $f(x) = 0$ *and* $J(x^*)$ *be regular.*

Then the error estimation $\|x_0 - x^*\| \leq \frac{Mr}{1-k}$ *is satisfied*

Now we will apply this theorem to a two dimensional system of nonlinear equations. Let us consider the following system

$$\begin{aligned} f_1(x_1, x_2) &= x_1^2 + x_2^2 - 1 = 0 \\ f_2(x_1, x_2) &= x_1^2 - x_2 = 0 \end{aligned} \tag{7.12}$$

For an interval vector $X = (X_1, X_2)$, we can take $F(X) = \begin{pmatrix} X_1^2 + X_2^2 - 1 \\ X_1^2 - X_2 \end{pmatrix}$ and $F'(X) = \begin{pmatrix} 2X_1 & 2X_2 \\ 2X_1 & -1 \end{pmatrix}$ as the respective interval extensions of the vector function

$\mathbf{f(x)} = \begin{pmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \end{pmatrix} = \begin{pmatrix} x_1^2 + x_2^2 - 1 \\ x_1^2 - x_2 \end{pmatrix}$ and the Jacobian matrix

$f'(x) = \begin{pmatrix} 2x_1 & 2x_2 \\ 2x_1 & -1 \end{pmatrix} = J(x)$ (say).

Here, $x^* = (0.7861513377, 0.618033988)$ is an exact solution of the above system correct to nine decimal places.

Let $z = y_0 = (0.80, 0.62)$ be an approximate solution obtained by some non-interval method.

Then $J(z) = \begin{pmatrix} 1.6 & 1.24 \\ 1.6 & -1 \end{pmatrix}$. Now, we have $r = \|f(z)\| = 0.0244$ and $\|[J(z)]^{-1}\| = 0.892857142 \leq 0.90 = M(say)$.

In the closed domain $\Omega_\delta = \{x \in \mathbb{R}^2 : \|x - z\| \leq \delta\}$,

let $\|x-z\|_\infty \leq \delta \Rightarrow \left\| \begin{pmatrix} x_1 - z_1 \\ x_2 - z_2 \end{pmatrix} \right\|_\infty \leq \delta \Rightarrow \left\| \begin{pmatrix} x_1 - 0.80 \\ x_2 - 0.62 \end{pmatrix} \right\|_\infty \leq \delta.$

That is , $\max(|x_1 - 0.80|, |x_2 - 0.62|) \leq \delta$.

Again, $\|J(x) - J(z)\|_\infty$

$$= \left\| \begin{pmatrix} 2x_1 & 2x_2 \\ 2x_1 & -1 \end{pmatrix} - \begin{pmatrix} 1.6 & 1.24 \\ 1.6 & -1 \end{pmatrix} \right\|_\infty = \left\| \begin{pmatrix} 2(x_1 - 0.80) & 2(x_2 - 0.62) \\ 2(x_1 - 0.80) & 0 \end{pmatrix} \right\|_\infty$$

where $\|.\|_\infty$ = maximum norm.

Therefore, $\|J(x) - J(z)\|_\infty$

$= \max(2|x_1 - 0.80| + 2|x_2 - 0.62|, 2|x_1 - 0.80|)$

$= (2|x_1 - 0.80| + 2|x_2 - 0.62|) \leq 4\delta$.

But $\frac{Mr}{1-k} \leq \delta$ is satisfied if $\delta \geq \delta(1-k) \geq Mr = 0.90 \times 0.0244 = 0.02196$

and $k \geq 4 \times 0.90\delta = 3.6\delta$.

As the closed domain $\Omega_\delta = \{x \in \mathbb{R}^2 : \|x - z\| \leq \delta\}$ is the δ-neighborhood of $z = y_0 = (0.80, 0.62)$. We may choose $\delta = 0.10$,

then $\Omega_\delta = \{x \in \mathbb{R}^2 : \|x - z\| \leq 0.10\} = [0.70, 0.90] \times [0.52, 0.72] = D$

and $k \geq 4 \times 0.90\delta = 0.36$. So if we choose $k = 0.36$ then all the conditions of **M.Urabe's** theorem are satisfied.

In the closed domain Ω_δ, $x^* = (0.7861513377, 0.618033988)$ is the unique solution of $f(x) = 0$.

The error estimation gives $\left\|x^{(0)} - x^*\right\| \leq \frac{Mr}{1-k} = 0.0343125 \leq 0.035$ (say.)

i.e., $0.80 - 0.035 \leq x^* \leq 0.80 + .035$

$\Rightarrow x_1 = [0.765, 0.835]$ and $x_2 = [0.585, 0.655]$. *i.e.*, $(x_1, x_2) \in ([0.765, 0.835], [0.585, 0.655])$.

Thus we have the following rectangle $X_0 = ([0.765, 0.835], [0.585, 0.655])$ containing the exact root.

To apply Krawczyk algorithm with X_0 we need the following results :

$f(y_{(0)}) = \begin{pmatrix} 0.0244 \\ 0.02 \end{pmatrix}$, $f'(y_0) = \begin{pmatrix} 1.6 & 1.24 \\ 1.6 & -1 \end{pmatrix}$, and also

$F'(X_0) = \begin{pmatrix} [1.53, 1.67] & [1.17, 1.31] \\ [1.53, 1.67 & [-1, -1] \end{pmatrix}$.

For Y, we may take the approximation of $[f'(y_0)]^{-1}$:

$Y_0 = [f'(y_0)]^{-1} = \begin{pmatrix} 0.279017857 & 0.345982142 \\ 0.446428571 & -0.446428571 \end{pmatrix}$. With this results using rounded interval arithmetic, we compute that for Krawczyk transformation

$$K(X_0) = y_0 - Y_0 f(y_0) + \{I - Y_0 F'(X_0)\}(X_0 - y_0)$$

$K(X_0) = \begin{pmatrix} [0.784, 0.788] \\ [0.615, 0.621] \end{pmatrix} \subseteq X_0$ and on the other hand

$$\|R_0\| = \|I - Y_0 F'(X_0)\| = 0.09462 < 1.$$

Now for X_0,

we have $X_1 = K(X_0) \cap X_0 = \begin{pmatrix} [0.784, 0.788] \\ [0.615, 0.621] \end{pmatrix} \subseteq X_0$.

Again we find $y_1 = m(X_1) = \begin{pmatrix} 0.786 \\ 0.618 \end{pmatrix}$, $f(y_{(1)}) = \begin{pmatrix} -0.00024 \\ -.000204 \end{pmatrix}$,

$f'(y_1) = \begin{pmatrix} 1.572 & 1.236 \\ 1.572 & -1 \end{pmatrix}$, $F'(X_1) = \begin{pmatrix} [1.568, 1.576 & [1.23, 1.242] \\ [1.568, 1.576] & [-1, -1] \end{pmatrix}$

and $Y_1 = [f'(y_1)]^{-1} = \begin{pmatrix} 0.284 & 0.352 \\ 0.447 & -0.447 \end{pmatrix}$. Using rounded interval arithmetic for $K(X_1)$, we compute that for Krawczyk transformation

$K(X_1) = y_1 - Y_1 f(y_1) + \{I - Y_1 F'(X_1)\}(X_1 - y_1)$

$$K(X_1) = \begin{pmatrix} [0.786,0.786] \\ [0.618,0.618] \end{pmatrix} \subseteq X_1 \subseteq X_0$$

Thus $X_2 = K(X_1) \cap X_1 = \begin{pmatrix} [0.786,0.786] \\ [0.618,0.618] \end{pmatrix} \subseteq X_1$.

Therefore, we have the nested intervals $X_2 \subset X_1 \subset X_0$.

Also we have, $\|R_1\| = \|I - Y_1 F'(X_1)\| = 0.006766$.

In this example, we find that $\|R_0\| < 1$ and $\|R_1\| < 1$. As we have $w(X_1) = 0.006$. And $w(X_0) = 0.07$. It is seen that $w(X_{k+1}) \leq \|R_k\| w(X_k)$ and $\{w(X_k\} \to 0$ are satisfied.

The initial interval $X_{(0)}$ obtained by **M.Urabe's** theorem, thus guarantees rapid convergence for R. Krawczyk algorithm.

7.3 Search for Safe Starting Regions to Computate Interval Solutions

Here we are apply safe starting regions to a two dimensional system of nonlinear equations[10]. Let us consider the following system

$$\begin{aligned} f_1(x_1,x_2) &= x_1^2 + x_2^2 - 1 = 0 \\ f_2(x_1,x_2) &= x_1^2 - x_2 = 0 \end{aligned} \tag{7.13}$$

For an interval vector $X = (X_1, X_2)$, we can take $F(X) = \begin{pmatrix} X_1^2 + X_2^2 - 1 \\ X_1^2 - X_2 \end{pmatrix}$ and

$F'(X) = \begin{pmatrix} 2X_1 & 2X_2 \\ 2X_1 & -1 \end{pmatrix}$ as the respective interval extensions of the vector function

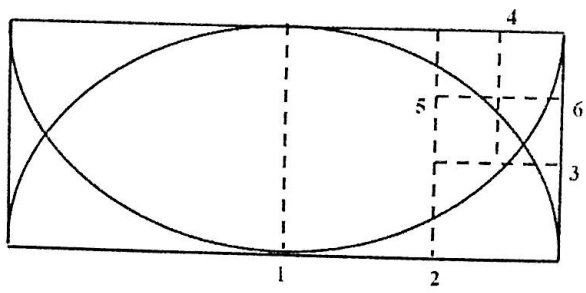

Figure 7.1: A two dimensional example

$\mathbf{f(x)} = \begin{pmatrix} f_1(x_1,x_2) \\ f_2(x_1,x_2) \end{pmatrix} = \begin{pmatrix} x_1^2+x_2^2-1 \\ x_1^2-x_2 \end{pmatrix}$ and the Jacobian matrix

$f'(x) = \begin{pmatrix} 2x_1 & 2x_2 \\ 2x_1 & -1 \end{pmatrix}$ respectively.

Note that the Jacobian matrix is singular at the point $x_1 = 0, x_2 \in [0,1]$ in B. This causes no difficulty for the search procedure.

The search proceeds as follows.(follow fig-1).

Initially, $X = B = ([-1,1],[0,1])$

Let $X = X^{(1)} \cup X^{(2)}$, where $X^{(1)} = ([-1,0],[0,1])$ and $X^{(2)} = ([0,1],[0,1])$

Symmetry test: We check

$$\min_h (\sum_{j=1}^{2} |m(F_j(X^{(h)}))|$$

$$\sum_{j=1}^{2} |m(F_j(X^{(1)}))| = |m(F_1(X^{(1)}))| + |m(F_2(X^{(1)}))| = |m([-1,0]^2+[0,1]^2-1)| + |m([-1,0]^2-[0,1]$$

$$= |m([-2,0)| + |m([-2,0])| = 2.$$

Again

$$\sum_{j=1}^{2} |m(F_j(X^{(2)}))| = |m(F_1(X^{(2)}))| + |m(F_2(X^{(2)}))| = |m([0,1]^2+[0,1]^2-1)| + |m([0,1]^2-[0,1])|$$

$$= |m([-1,1)| + |m([-1,1]| = 1.$$

As $1 < 2$, so we choose the region $X = X^{(2)} = ([0,1],[0,1])$. and test step 7 and step 8 from search procedure.

Here, $m(X) = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$, $F(m(X)) = \begin{pmatrix} (.5)^2 + (.5)^2 - 1 \\ (.5)^2 - .5 \end{pmatrix} = \begin{pmatrix} -0.5 \\ -0.25 \end{pmatrix}$

$$X - m(X) = \begin{pmatrix} [0,1] \\ [0,1] \end{pmatrix} - \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} = \begin{pmatrix} [-0.5,0.5] \\ [-0.5,0.5] \end{pmatrix}$$

$$\mathbf{Y} = \left[m \begin{pmatrix} 2[0,1] & 2[0,1] \\ 2[0,1] & [-1,-1] \end{pmatrix} \right]^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & -0.5 \end{pmatrix}$$

$$R = I - YF'(X) = I - Y \begin{pmatrix} [0,2] & [0,2] \\ [0,2] & [-1-1] \end{pmatrix} = I - \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & -0.5 \end{pmatrix} \begin{pmatrix} [0,2] & [0,2] \\ [0,2] & [-1-1] \end{pmatrix}$$

So $R = I - YF'(X) = \begin{pmatrix} [-1,1] & [-0.5,0.5] \\ [-1,1] & [-0.5,0.5] \end{pmatrix}$.

$$YF(m(X)) = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & -0.5 \end{pmatrix} \begin{pmatrix} -0.5 \\ -0.25 \end{pmatrix} = \begin{pmatrix} -0.375 \\ -0.125 \end{pmatrix}$$

Now, $R(X - m(X)) = \begin{pmatrix} [-1,1] & [-0.5,0.5] \\ [-1,1] & [-0.5,0.5] \end{pmatrix} \begin{pmatrix} [-0.5,0.5] \\ [-0.5,0.5] \end{pmatrix} = \begin{pmatrix} [-0.7501,0.7501] \\ [-0.7501,0.7501] \end{pmatrix}$

Therefore,

$$K(X) = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} + \begin{pmatrix} 0.375 \\ .125 \end{pmatrix} + \begin{pmatrix} [-0.7501,0.7501] \\ [-0.7501,0.7501] \end{pmatrix}$$

$$K(X) = \begin{pmatrix} 0.875 \\ .625 \end{pmatrix} + \begin{pmatrix} [-0.7501,0.7501] \\ [-0.7501,0.7501] \end{pmatrix} = \begin{pmatrix} [0.1249,1.6251] \\ [-0.1251,1.3751] \end{pmatrix} \not\subseteq X$$

So bisect; coordinate direction 1 chosen.

Bisection gives $X^{(1)} = ([0,0.5],[0,1])$, $X^{(2)} = ([.5,1],[0,1])$

Symmetry test: We check

$$\min_h(\sum_{j=1}^{2}|m(F_j(X^{(h)}))|$$

$$\sum_{j=1}^{2}|m(F_j(X^{(1)}))| = |m(F_1(X^{(1)}))| + |m(F_2(X^{(1)}))| = |m([0,0.5]^2 + [0,1]^2 - 1)| + |m([0,0.5]^2 - [0,1])|$$

$$= |m([-1,0.25])| + |m([-1,0.25])| = .75$$

Again

$$\sum_{j=1}^{2}|m(F_j(X^{(2)}))| = |m(F_1(X^{(2)}))| + |m(F_2(X^{(2)}))| = |m([.5,1]^2 + [0,1]^2 - 1)| + |m([0.5,1]^2 - [0,1])|$$

$$= |m([-0.75,1])| + |m([-0.75,1])| = .25$$

As $.25 < .75$, we choose the region $X = X^{(2)} = ([0.5,1],[0,1])$.

$$\text{Here,} \quad m(X) = \begin{pmatrix} 0.625 \\ 0.75 \end{pmatrix}, \quad F(m(X)) = \begin{pmatrix} -0.047 \\ -0.359 \end{pmatrix}$$

$$X - m(X) = \begin{pmatrix} [-0.125, 0.125] \\ [-0.25, 0.25] \end{pmatrix}$$

$$\mathbf{Y} = \left[m \begin{pmatrix} 2[0.5,1] & 2[0,1] \\ 2[0.5,1] & [-1,-1] \end{pmatrix}\right]^{-1} = \begin{pmatrix} 1.5 & 1 \\ 1.5 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} 0.3333 & 0.3333 \\ 0.5 & -0.5 \end{pmatrix}$$

$$R = I - YF'(X) = I - \begin{pmatrix} 0.3333 & 0.3333 \\ 0.5 & -0.5 \end{pmatrix} \begin{pmatrix} [1,2] & [0,2] \\ [1,2] & [-1,1] \end{pmatrix} R = I - YF'(X) =$$

$$\begin{pmatrix} [-0.3332, 0.3334] & [-0.3333.333] \\ [-0.5, 0.5] & [-0.5, 0.5] \end{pmatrix}.$$

$$\text{Now,} \quad R(X - m(X)) = \begin{pmatrix} [-0.2501, 0.2501] \\ [-0.3501, 0.3501] \end{pmatrix}$$

$$YF(m(X)) = \begin{pmatrix} 0.0417 \\ -0.0625 \end{pmatrix}$$

Therefore,

$$K(X) = \begin{pmatrix} 0.75 \\ 0.5 \end{pmatrix} - \begin{pmatrix} 0.0417 \\ -0.0625 \end{pmatrix} + \begin{pmatrix} [-0.2501, 0.2501] \\ [-0.3751, 0.3751] \end{pmatrix}$$

$$K(X) = \begin{pmatrix} [0.5416, 1.0417] \\ [.2499, 1.0000] \end{pmatrix} \not\subseteq X$$

We have seen that $K(X)$ does not contained in X. So bisect; coordinate direction 2 chosen; bisection gives

$$X^{(1)} = ([.5, 1], [0, 0.5]), \qquad X^{(2)} = ([0.5, 1], [0.5, 1])$$

Symmetry Test.

$$\sum_{j=1}^{2} |m(F_j(X^{(1)}))| = |m(F_1(X^{(1)}))| + |m(F_2(X^{(1)}))| = |m([0.5, 1]^2 + [0, 0.5]^2 - 1)| + |m([0.5, 1]^2 - [0, 0.$$

$$= |m([-.75, 0.251])| + |m([0.25, 1])| = .625$$

Again

$$\sum_{j=1}^{2} |m(F_j(X^{(2)}))| = |m(F_1(X^{(2)}))| + |m(F_2(X^{(2)}))| = |m([.5, 1]^2 + [0.5, 1]^2 - 1)| + |m([0.5, 1]^2 - [0.5, 1$$

$$= |m([-0.5, 1])| + |m([-0.75, 0.5])| = .375$$

As $0.375 < 0.625$, we choose the half $X = X^{(2)} = ([0.5, 1], [0.5, 1])$

Here, $m(X) = \begin{pmatrix} 0.75 \\ 0.75 \end{pmatrix}$, $F(m(X)) = \begin{pmatrix} (.75)^2 + (0.75)^2 - 1 \\ (.75)^2 - 0.75 \end{pmatrix} = \begin{pmatrix} 0.125 \\ -0.1875 \end{pmatrix}$

$$X - m(X) = \begin{pmatrix} [0.5, 1] \\ [0., 1] \end{pmatrix} - \begin{pmatrix} 0.75 \\ 0.75 \end{pmatrix} = \begin{pmatrix} [-0.25, 0.25] \\ [-0.5, 0.25] \end{pmatrix}$$

$$\mathbf{Y} = m\left[\begin{pmatrix} 2[0.5,1] & 2[0.5,1] \\ 2[0.5,1] & [-1,-1] \end{pmatrix}\right]^{-1} = \begin{pmatrix} 1.5 & 1.5 \\ 1.5 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} 0.267 & 0.4 \\ 0.4 & -0.4 \end{pmatrix}$$

$$R = I - YF'(X) = I - Y\begin{pmatrix} [1,2] & [1,2] \\ [1,2] & [-1-1] \end{pmatrix} = I - \begin{pmatrix} 0.267 & 0.4 \\ 0.4 & -0.4 \end{pmatrix}\begin{pmatrix} [1,2] & [1,2] \\ [1,2] & [-1,1] \end{pmatrix}$$

So $R = I - YF'(X) = \begin{pmatrix} [-0.334,0.4] & [-0.134.133] \\ [-0.4,0.4] & [-0.2,0.2] \end{pmatrix}$.

Now, $R(X - m(X)) = \begin{pmatrix} [-0.334,0.4] & [-0.134,.133] \\ [-0.4,0.4] & [-0.2,0.2] \end{pmatrix}\begin{pmatrix} [-0.25,0.25] \\ [-0.25,0.25] \end{pmatrix}$

$$= \begin{pmatrix} [-0.1335,0.1335] \\ [-0.6,0.6] \end{pmatrix}$$

$$YF(m(X)) = \begin{pmatrix} 0.267 & 0.4 \\ 0.4 & -0.4 \end{pmatrix}\begin{pmatrix} 0.125 \\ -0.1875 \end{pmatrix} = \begin{pmatrix} 0.072 \\ 0.125 \end{pmatrix}$$

Therefore,

$$K(X) = \begin{pmatrix} 0.75 \\ 0.75 \end{pmatrix} - \begin{pmatrix} 0.072 \\ 0.125 \end{pmatrix} + \begin{pmatrix} [-0.1335,0.1335] \\ [-0.6,0.6] \end{pmatrix}$$

$$K(X) = \begin{pmatrix} 0.675 \\ 0.625 \end{pmatrix} + \begin{pmatrix} [-0.1335,0.1335] \\ [-0.6,0.6] \end{pmatrix} = \begin{pmatrix} [0.6749,0.9084] \\ [0.4749,.7751] \end{pmatrix} \not\subseteq X$$

$$F(X) = F\begin{pmatrix} [0.5,1] \\ [0.5,1] \end{pmatrix} = \begin{pmatrix} [0.5,1]^2 + [0.5,1]^2 - 1 \\ [0.5,1]^2 - [0.5,1] \end{pmatrix} = \begin{pmatrix} [-0.5,1] \\ [-0.75,0.75] \end{pmatrix}$$

We have $0 \in F(X)$. As $K(X) \not\subseteq X$ and $0 \in F(X)$, so bisect; coordinate direction 1 chosen.

$$X^{(1)} = ([.75,1],[0.5,1]), \quad X^{(2)} = ([0.5,0.75],[0.5,1])$$

Symmetry Test.

$$\sum_{j=1}^{2} |m(F_j(X^{(1)}))| = |m(F_1(X^{(1)}))| + |m(F_2(X^{(1)}))|$$

$$= |m([0.75,1]^2 + [0.5,1]^2 - 1)| + |m([0.75,1]^2 - [0.5,1])| = .4375$$

Again

$$\sum_{j=1}^{2} |m(F_j(X^{(2)}))| = |m(F_1(X^{(2)}))| + |m(F_2(X^{(2)}))|$$

$$= |m([.5,.75]^2 + [0.5,1]^2 - 1)| + |m([0.5,.75]^2 - [0.5,1])| = .375$$

As $0.375 < 0.4375$, we choose the half $X = X^{(2)} = ([0.5,0.75],[0.5,1])$

$$\text{Here,} \quad m(X) = \begin{pmatrix} 0.625 \\ .75 \end{pmatrix}, F(m(X)) = \begin{pmatrix} -0.047 \\ -0.359 \end{pmatrix}$$

$$X - m(X) = \begin{pmatrix} [-0.125, 0.125] \\ [-0.25, 0.25] \end{pmatrix}$$

$$\mathbf{Y} = \left[m \begin{pmatrix} 2[0.5,0.75] & 2[0.5,1] \\ 2[0.5,0.75] & [-1,-1] \end{pmatrix} \right]^{-1} = \begin{pmatrix} 1.25 & 1.5 \\ 1.25 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} 0.32 & 0.48 \\ 0.4 & -0.4 \end{pmatrix}$$

$$R = I - YF'(X) = I - \begin{pmatrix} 0.32 & 0.48 \\ 0.4 & -0.4 \end{pmatrix} \begin{pmatrix} [1,1.5] & [1,2] \\ [1,1.5] & [-1-1] \end{pmatrix}$$

$$\text{So} \quad R = I - YF'(X) = \begin{pmatrix} [-0.2,0.2] & [-0.16,0.16] \\ [-0.2,0.2] & [-0.2,0.2] \end{pmatrix}.$$

$$\text{Now,} \quad R(X - m(X)) = \begin{pmatrix} [-0.065, 0.065] \\ [-0.075, 0.075] \end{pmatrix}$$

$$YF(m(X)) = \begin{pmatrix} -0.1874 \\ 0.1248 \end{pmatrix}$$

Therefore,

$$K(X) = \begin{pmatrix} 0.625 \\ 0.75 \end{pmatrix} - \begin{pmatrix} -0.1874 \\ 0.1248 \end{pmatrix} + \begin{pmatrix} [-0.065, 0.065] \\ [-0.075, 0.075] \end{pmatrix} = \begin{pmatrix} [0.7474, 0.8776] \\ [0.5499, 0.7001] \end{pmatrix} \not\subseteq X$$

$$F(X) = F\begin{pmatrix} [0.5, .75] \\ [0.5, 1] \end{pmatrix} = \begin{pmatrix} [0.5, .75]^2 + [0.5, 1]^2 - 1 \\ [0.5, .75]^2 - [0.5, 1] \end{pmatrix} = \begin{pmatrix} [-0.750, 0.0625] \\ [-0.75, 0.75] \end{pmatrix}$$

We have $0 \in F(X)$. As $K(X) \not\subseteq X$ and $0 \in F(X)$, so bisect; coordinate direction 2 chosen.Bisected halves are

$$X^{(1)} = ([0.5, 0.75], [0.75, 1]), \quad X^{(2)} = ([0.5, 0.75], [0.5, 0.75])$$

Symmetry Test.

$$\sum_{j=1}^{2} |m(F_j(X^{(1)}))| = |m(F_1(X^{(1)}))| + |m(F_2(X^{(1)}))|$$

$$= |m([0.5, .075]^2 + [0.75, 1]^2 - 1)| + |m([0.5, .75]^2 - [0.75, 1])| = .65625$$

Again

$$\sum_{j=1}^{2} |m(F_j(X^{(2)}))| = |m(F_1(X^{(2)}))| + |m(F_2(X^{(2)}))|$$

$$= |m([.5, .75]^2 + [0.5, .75]^2 - 1)| + |m([0.5, .75]^2 - [0.5, .75])| = .40625$$

As $0.40625 < 0.65625$, we choose $X = X^{(2)} = ([0.5, 0.75], [0.5, 0.75])$

$$\text{Here,} \quad m(X) = \begin{pmatrix} 0.625 \\ 0.625 \end{pmatrix} F(m(X)) = \begin{pmatrix} -0.21875 \\ -0.234375 \end{pmatrix}$$

$$X - m(X) = \begin{pmatrix} [-0.125, 0.125] \\ [-0.125, 0.125] \end{pmatrix}$$

$$\mathbf{Y} = \left[m \begin{pmatrix} 2[0.5, 0.75] & 2[0.5, 0.75] \\ 2[0.5, 0.75] & [-1, -1] \end{pmatrix}\right]^{-1} = \begin{pmatrix} 0.356 & 0.444 \\ 0.444 & -0.444 \end{pmatrix}$$

$$R = I - YF'(X) = I - \begin{pmatrix} 0.356 & 0.444 \\ 0.444 & -0.444 \end{pmatrix} \begin{pmatrix} [1,1.5] & [1,1.5] \\ [1,1.5] & [-1-1] \end{pmatrix}$$

So $R = I - YF'(X) = \begin{pmatrix} [-0.19985, 0.2001] & [-0.0885, 0.889] \\ [-0.2222, 0.2222] & [-0.111, 0.1112] \end{pmatrix}$.

Now, $R(X - m(X)) = \begin{pmatrix} [-0.19985, 0.2001] & [-0.0885, 0.889] \\ [-0.2222, 0.2222] & [-0.111, 0.1112] \end{pmatrix} \begin{pmatrix} [-0.125, 0.125] \\ [-0.25, 0.25] \end{pmatrix}$

$$= \begin{pmatrix} [-0.036, 0.036] \\ [-0.0416, 0.0416] \end{pmatrix}$$

$$YF(m(X)) = \begin{pmatrix} 0.0262 \\ -0.2012 \end{pmatrix}$$

Thus,

$$K(X) = \begin{pmatrix} 0.625 \\ 0.625 \end{pmatrix} + \begin{pmatrix} 0.0262 \\ -0.2012 \end{pmatrix} + \begin{pmatrix} [-0.036, 0.036] \\ [-0.0416, 0.0416] \end{pmatrix}$$

$$K(X) = \begin{pmatrix} 0.5988 \\ .8262 \end{pmatrix} + \begin{pmatrix} [-0.036, 0.036] \\ [-0.0416, 0.0416] \end{pmatrix} = \begin{pmatrix} [0.7708, .8431] \\ [0.5763, .6598] \end{pmatrix}$$

$K(X) \cap X = \phi$ for $X = ([0.5, 0.75], [0.5, 0.75])$ and $X = ([0.5, 0.75], [0.75, 1])$.

So exclude $X = ([0.5, 0.75], [0.5, 0.75])$ and $X = ([0.5, 0.75], [0.75, 1])$.

Also $K(X) \cap ([0.5, 0.75], [0.5, 1]) = \phi$, so exclude $X = ([0.5, 0.75], [0.5, 1])$,

We choose $X = ([0.75, 1], [0.5, 1])$

$$K(X) = m(X) - YF(m(X)) + R(X - m(X)).$$

Here, $m(X) = \begin{pmatrix} 0.875 \\ 0.75 \end{pmatrix} F(m(X)) = \begin{pmatrix} 0.328 \\ 0.016 \end{pmatrix}$

$$X - m(X) = \begin{pmatrix} [0.75, 1] \\ [0.5, 1] \end{pmatrix} - \begin{pmatrix} 0.875 \\ 0.75 \end{pmatrix} = \begin{pmatrix} [-0.125, 0.125] \\ [-0.25, 0.25] \end{pmatrix}$$

$$\mathbf{Y} = \left[m \begin{pmatrix} 2[0.75,1] & 2[0.5,1] \\ 2[0.75,1] & [-1,-1] \end{pmatrix}\right]^{-1} = \begin{pmatrix} 0.229 & 0.343 \\ 0.4 & -0.4 \end{pmatrix}$$

So $R = I - YF'(X) = \begin{pmatrix} [-0.143,0.143] & [-0.114,0.114] \\ [-0.2,0.2] & [-0.2,0.2] \end{pmatrix}$.

$$\text{Now,} \quad R(X - m(X)) = \begin{pmatrix} [-0.047,0.047] \\ [-0.075,0.075] \end{pmatrix}$$

$$K(X) = \begin{pmatrix} 0.875 \\ 0.75 \end{pmatrix} - \begin{pmatrix} 0.081 \\ 0.125 \end{pmatrix} + \begin{pmatrix} [-0.047,0.047] \\ [-0.075,0.075] \end{pmatrix} = \begin{pmatrix} [0.7482,0.8411] \\ [0.5499,0.7001] \end{pmatrix} \not\subseteq X$$

So bisect; coordinate direction 2 chosen. $X^{(1)} = ([.75,1],[.75,1]), X^{(2)} = ([0.75,1],[0.5,0.75])$

$F(X^{(1)}) = \begin{pmatrix} [0.125,1] \\ [-0.4375,0.251] \end{pmatrix}$, $0 \notin F(X^{(1)})$,

Therefore, the exclusion principle select the lower half $X = ([0.75,1],[0.5,0.75])$

$$\text{Here,} \quad m(X) = \begin{pmatrix} 0.875 \\ 0.625 \end{pmatrix} F(m(X)) = \begin{pmatrix} 0.156 \\ 0.141 \end{pmatrix}$$

$$X - m(X) = \begin{pmatrix} [-0.125,0.125] \\ [-0.25,0.125] \end{pmatrix}$$

$$\mathbf{Y} = \left[m \begin{pmatrix} 2[0.75,1] & 2[0.5,.75] \\ 2[0.75,1] & [-1,-1] \end{pmatrix}\right]^{-1} = \begin{pmatrix} 0.254 & 0.317 \\ 0.444 & -0.444 \end{pmatrix}$$

$YF(m(X)) = \begin{pmatrix} 0.084 \\ 0.007 \end{pmatrix}$. So $R = I - YF'(X) = \begin{pmatrix} [-0.142,0.144] & [-0.064,0.063] \\ [-0.222,0.222] & [-0.111,0.112] \end{pmatrix}$.

$$\text{Now,} \quad R(X - m(X)) = \begin{pmatrix} [-0.026,0.026] \\ [-0.042,0.042] \end{pmatrix}$$

$$Thus, K(X) = \begin{pmatrix} [0.7648,0.8165] \\ [0.5763,0.6598] \end{pmatrix} \subseteq X$$

Let

$$R = (R_{ij} = \begin{pmatrix} [-0.142, 0.144] & [-0.064, 0.063] \\ [-0.222, 0.222] & [-0.111, 0.112] \end{pmatrix}$$

Then, $||R|| = \max_i \sum_j |R_{ij}| = \max_i(|R_{ij}| + |R_{ij}|) = max(|R_{i1}| + |R_{i2}|$

$= max(|R_{11}| + |R_{12}|, (|R_{21}| + |R_{22}|)$

$= max(|[-0.142, 0.144]| + [-0.064, 0.063]|, |[-0.222, 0.222]| + [-0.111, 0.112]|)$

$= max(.144 + 0.064, 0.222 + 0.112)$

$= max(.208, .334) = .334 < .34 < 1.$

Therefore solution exists in X and a safe starting region has been found.

$$X^{(0)} = X = ([0.75, 1.0], [0.5, 0.75])$$

$$K(X^{(0)}) = \begin{pmatrix} [0.7648, 0.8165] \\ [.5763, 0.6598] \end{pmatrix} \subseteq X^{(0)}$$

and $||R^{(0)}|| =< .335 < .36 < 1.$

$$X^{(1)} = K(X^{(0)}) \cap X^{(0)} = \begin{pmatrix} [0.7647, 0.8166] \\ [.5763, 0.6599] \end{pmatrix}$$

$$K(X^{(1)}) = m(X^{(1)}) - Y^{(1)} F(m(X^{(1)}) + R(X^{(1)} - m(X^{(0)}))$$

Here, $m(X^{(1)}) = \begin{pmatrix} 0.7915 \\ 0.6185 \end{pmatrix}$, $F(m(X^{(1)})) = \begin{pmatrix} 0.0090144 \\ 0.0079722 \end{pmatrix}$

$$X^{(1)} - m(X^{(1)} = \begin{pmatrix} [-0.0345, 0.0345] \\ [-0.0415, 0.0415125] \end{pmatrix}$$

$$\mathbf{Y^{(1)}} = \left[m \begin{pmatrix} 2[0.757, .826] & 2[0.577, .660] \\ 2[0.757, .826] & [-1, -1] \end{pmatrix} \right]^{-1} = \begin{pmatrix} 1 & 1.23 \\ 1.583 & -1.583 \end{pmatrix}$$

$$Y^{(1)} F(m(X^{(1)})) = \begin{pmatrix} 0.0053303 \\ 0.0004659 \end{pmatrix}.$$

$$R^{(1)} = I - Y^{(1)}F'(X^{(1)} = \begin{pmatrix} [-0.0435878, .0435885] & [-0.0234385, 0.0234386] \\ [-0.06166898, .0616898] & [-.0371031, .0371035] \end{pmatrix}$$

$$R^{(1)}(X^{(1)} - m(X^{(1)})) = \begin{pmatrix} [-0.0024765, 0.0024765] \\ [-0.0036679, 0.0036679] \end{pmatrix}$$

$$K(X^{(1)}) = \begin{pmatrix} 0.7915 \\ 0.6185 \end{pmatrix} - \begin{pmatrix} 0.0053303 \\ 0.0004659 \end{pmatrix} + \begin{pmatrix} [-0.0024765, 0.0024765] \\ [-0.0036679, 0.0036679] \end{pmatrix}$$

$$= \begin{pmatrix} [0.7843, 0.7881] \\ [.6152, 0.6209] \end{pmatrix} \subseteq X^{(1)}$$

$||R^{(1)}|| < .09 < 1$

$$X^{(2)} = K(X^{(1)}) \cap X^{(1)} = \begin{pmatrix} [0.7842, 0.7882] \\ [.6151, 0.6210] \end{pmatrix}$$

$$K(X^{(2)}) = m(X^{(2)}) - Y^{(2)}F(m(X^{(2)}) + R(X^{(2)} - m(X^{(2)}))$$

Here, $m(X^{(2)} = \begin{pmatrix} 0.786 \\ 0.618 \end{pmatrix}$, $F(m(X^{(2)})) = \begin{pmatrix} -.00028 \\ -.000204 \end{pmatrix}$

$$X^{(2)} - m(X^{(2)} = \begin{pmatrix} [-0.003, 0.003] \\ [-0.004, 0.004] \end{pmatrix}$$

$$\mathbf{Y}^{(2)} = \left[m \begin{pmatrix} 2[0.783, .789] & 2[0.614, .622] \\ 2[0.783, .789] & [-1, -1] \end{pmatrix}\right]^{-1} = \begin{pmatrix} 1 & 1.236 \\ 1.572 & -1.572 \end{pmatrix}$$

$$Y^{(2)}F(m(X^{(2)})) = \begin{pmatrix} -0.000153 \\ -0.000034 \end{pmatrix}.$$

$$R^{(2)} = I - Y^{(2)}F'(X^{(2)} = \begin{pmatrix} [-0.0038163, .003872] & [-0.002276, 0.002276] \\ [-0.0053668, .0053668] & [-.0035773, .035783] \end{pmatrix}$$

$$R^{(2)}(X^{(2)} - m(X^{(2)})) = \begin{pmatrix} [-0.0000205, 0.00205] \\ [-0.0000304, 0.0000304] \end{pmatrix}$$

$$K(X^{(2)}) = \begin{pmatrix} 0.786 \\ 0.618 \end{pmatrix} - \begin{pmatrix} -0.000153 \\ -0.000034 \end{pmatrix} + \begin{pmatrix} [-0.0000205, 0.00205] \\ [-0.0000304, 0.0000304] \end{pmatrix}$$

$$= \begin{pmatrix} 0.7861513 \\ 0.618034 \end{pmatrix} + \begin{pmatrix} [-0.0000205, 0.00205] \\ [-0.0000304, 0.0000304] \end{pmatrix}$$

$$K(X^{(2)}) = \begin{pmatrix} [0.7861308, 0.7861718] \\ [0.6180036, 0.6180644] \end{pmatrix} \subseteq \begin{pmatrix} [0.7861, 0.7862] \\ [0.6180, 0.6181] \end{pmatrix} \subseteq X^{(2)}$$

and $||R^{(2)}|| < .08 < 1$

$X^{(3)} = K(X^{(2)}) \cap X^{(2)} = ([.7861, .7862], [.6180, .6181])$

Therefore we have the nested intervals $X^{(3)} \subseteq X^{(2)} \subseteq X^{(1)} \subseteq X^{(0)}$.

In this example ,we have $||R^{(0)}|| < 1$, $R^{(1)} < 1$, and also $||R^{(2)}|| < 1$. As we have $w(X^{(0)}) = 0.25$, $W(X^{(1)}) = 0.008$ and $W(X^{(2)}) = 0.006$. It is seen that $w(X^{(k+1)}) \leq ||R^{(k)}||w(X^{(k)})$ and $w(X^{(k)}) \to 0$ are satisfied.

An INTLAB implementation of the algorithm is given by the function nlinkraw.m.

```
This routine finds bounds on the solution of a nonlinear system of
equations using the Krawczyk operator.


   function  Y = nlinkraw (f, X)
%NLINKRAW   Bounds roots of  nonlinear systems of equations.
%
%          X = NLINKRAW (f,X)
%
%           Uses the Krawczyk operator to produce bounds on a root
%           in a given interval of a nonlinear equation f.
%
```

```
%          INPUT  :  f   A MATLAB function
%                    X   Initial interval
%          OUTPUT :  Y   interval solution


   f = inline(' [x(1)*x(1) + x(2)*x(2) - 1 ; x(1)*x(1) - x(2)]')
 X = infsup(.75,1);infsup(.5,.75)];
 n = length(X);
 ready = 0;  k = 0;
 N = intval(zeros (n,1));

   while ~ready
     k = k+1;
     F = feval(f, gradientinit(X));
     C = inv(mid(F.dx));
     x = mid(X);
     fx = feval(f,intval(x));
     K = x-C*fx+(eye(n)-C*(F.dx))*(X-x);  % Krawczyk operator.
     Xnew = intersect(K, X);

        if isempty(Xnew)
           error('No root in box')
     elseif  X == Xnew;
         ready = 1;
     else
         X = Xnew;
   end
end
```

```
disp('Number of iterations')
disp(k)
format long
intvalinit('displayinfsup')
 Y = X;
```

From the above INTLAB implementation, we have seen that the initial interval X_0 obtained by using the safe starting regions iterative process and for three iterations converges to the intervals $([0.786, 0.786], [0.618, 0.618])$ rounded to five significant figures that guarantees rapid convergence for R. Krawczyk algorithm. For six iterations also converges to the intervals $([0.78615137775774, 0..78615137775774], [0.6180339887498, 0.6180339887498])$. Here we conclude that the above degenerate intervals indicate that the approximate values of x_1 and x_2 are correct to fourteen decimal places. Higher precisions may increase the number of iterations.

References

[1] M. Zulfikar Ali and Md. Shirazul Hoque Mollah, *Computation of interval solution to nonlinear systems*, International Journal of Engineering and Indisciplinary Mathematics,India (Accepted 15 November2009).

[2] C.Bliek, *Computer methods design automation*, PhD Thesis, Manssachusetts Institute of Technology (1992).

[3] C.Miranda, *Un osservazione suun teorema di brower, boll.un.mat.ital.*, John Willey, New York **3** (1941), 5–7.

[4] J. Hales, *The kepler conjecture.http://www.math.pitt.edu/ thales/countdown.*, http://www.math.pitt.edu/ thales/countdown.

[5] E. R. Hansen, *On solving systems of equations using interval arithmetic*, Mathematics of Computation **22** (1968), no. 102, 374–384.

[6] Gareth I. Hargreaves, *Interval analysis in matlab*, Manchester Institute of Mathematical Sciences **1824** (2002).

[7] J.M.Ortega and w.c.Rheinboldt, *Iterative solution of nonlinear equations in several variables*, Academic Press, New York (1970).

[8] R. Krawczyk, *Newton-algorithm zur bestimmung von nullstellen mit fehlerschranken*, Computing **4** (1969), 187–201.

[9] Rall L.B, *Siam j. numericul anal.*, John Willey, New York **11** (1974), 34–36.

[10] M. Zulfikar Ali Md. Shirazul Hoque Mollah, *Search for safe starting regions to compute interval solution to a nonlinear systems.*, International Journal of Mathematics and Computations,India **13** (2011), no. D-11, 69–77.

[11] R.E. Moore, *A test for existence of solutions of nonlinear systems*, SIAM, T Numer,Anal **14** (1977), 611 – 615.

[12] R.E. Moore and S.T Jones, *Safe starting regions for iterative methods*, SAIM J. Numerical Analysis **14** (1977), 1051–1064.

[13] A. Neumaier, *Interval methods for system of equations*, Cambridge University, Press (1990).

[14] Kaj Madsen Ole Caprani and Hans Bruun Nielsen, *Introduction to interval analysis*, IMM,DTU (2002).

[15] Tapia R.A, *The kantorovich theorem for newton's method*, Amer.Math Monthly **78** (1971), 389–392.

[16] L.B. Rall, *Computational solution of nonlinear operator equations*, John Willey, New York (1969).

[17] R. Baker Kearfott Ramon E.Moore and Michael J. Cloud, *Introduction to interval analysis*, Socity for Industrial and Applied Mathematics, Philadelphia (2009).

[18] Helmut Ratschek and Jon Rokne, *New computer methods for global optimization*, University of Calgary, Calgary , Alberta (2007).

[19] T.Y.Li R.B. Kellogg and J. York, *A constructive proof of the brouwer fixed point theorem and computational results*, Tech. Note BN-810, Institute for Fluid Dynamics and applied Math., Univ.of Maryland (1975).

[20] R.E.Moore, *Interval analysis*, Prentice-Hall, Englewood Cliffs,NJ,USA (1966).

[21] J. Rohn, *The recent result by e.r. hansen can be made more efficient.*, Interval Computiations **4** (1993), 13 – 21.

[22] S.M.Rump, *Interval laboratory.*, In Tibor Csendes editor, Developments in Reliable Computing, Netherlands (1999), 77–105.

[23] S.Ning and R.B. Kearfott, *A comparsion of some methods for solving linear interval equations*, SIAM Journal on Numerical Analysis **34** (1997), no. 4, 1289 – 1305.

[24] Minoru Urabe, *Arch. rational mech.*, John Willey, New York **20** (1965), 120–152.

[25] Md. Majedur Rahaman ZulfikarAli, *On convergence of newton's method*, Ann. Bull of the Res. Inst. of Math and Inf. Sci, Japan **8** (2002), 99–103.