67-5422

BROWN, Kenneth Myron, 1936-
A QUADRATICALLY CONVERGENT METHOD FOR SOLVING SIMULTANEOUS NON-LINEAR EQUATIONS.

Purdue University, Ph.D., 1966
Mathematics

University Microfilms, Inc., Ann Arbor, Michigan

# A QUADRATICALLY CONVERGENT METHOD FOR SOLVING SIMULTANEOUS NON-LINEAR EQUATIONS

A Thesis

Submitted to the Faculty

of

Purdue University

by

Kenneth Myron Brown

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

August 1966

20487

Grad. School
Form No. 9
Revised

PURDUE UNIVERSITY

Graduate School

This is to certify that the thesis prepared

By Kenneth Myron Brown

Entitled A QUADRATICALLY CONVERGENT METHOD FOR SOLVING SIMULTANEOUS NON-LINEAR EQUATIONS

Complies with the University regulations and that it meets the accepted standards of the Graduate School with respect to originality and quality

For the degree of:

Doctor of Philosophy

Signed by the final examining committee:

S. D. Conte, chairman

J. S. Maybee

[illegible]

Saul Rosen

Approved by the head of school or department:

7/26 1966 S. D. Conte

To the librarian:

This thesis ~~is~~ is not to be regarded as confidential

S. D. Conte
Professor in charge of the thesis

# DEDICATION

To Sandy and David

# ACKNOWLEDGMENTS

I wish to express my appreciation to Professor Samuel D. Conte, my Major Professor, for the knowledge he has imparted to me, for his outstanding leadership in directing my research efforts and for his constant support and encouragement. I also wish to thank Dr. Conte in his capacity as Department Head for the opportunity to gain valuable teaching and research experience while in his employ.

I should like to thank the other members of my Advisory Committee, Professor Leonard D. Berkovitz, Professor Walter Gautschi, Professor John S. Maybee, and Professor Saul Rosen for the help they have given me during my graduate studies.

I should like to thank Professor Peter K. Henrici who guided my Master's Degree work and who proposed research along the present lines.

I wish to thank two truly fine people, Dr. and Mrs. Simon B. Matzkin, for their abundant and sincere support.

To my dear wife, Sandy, I owe my sincerest gratitude for her patience, her good-naturedness, and her continuous help and encouragement.

# TABLE OF CONTENTS

# LIST OF TABLES

# ABSTRACT

Brown, Kenneth Myron. Ph.D., Purdue University, August 1966. A QUADRATICALLY CONVERGENT METHOD FOR SOLVING SIMULTANEOUS NON-LINEAR EQUATIONS. Major Professor: Samuel D. Conte.

An iterative method for the numerical solution of a system of N simultaneous non-linear algebraic equations is presented. A proof of local convergence for the method is given in which it is shown that the convergence is quadratic in nature. Computable bounds on the region of convergence are obtained in connection with a non-local convergence theorem. Computer results for specific applications of the method are given and comparisons are made with other currently used methods. In particular it is shown that the method requires only $(\frac{N^2}{2} + \frac{3N}{2})$ function evaluations per iterative step as compared with $(N^2 + N)$ evaluations for Newton's Method. An ALGOL procedure for the method is also given.

# 1. INTRODUCTION

Let a system of N non-linear (or linear) algebraic equations in N unknowns be given as

$$
\begin{aligned}
&f_1(\underline{x}) = f_1\ (x_1,\ x_2,\ldots,\ x_N) = 0\\
&f_2(\underline{x}) = f_2\ (x_1,\ x_2,\ldots,\ x_N) = 0\\
(1.1)\qquad &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&f_N(\underline{x}) = f_N\ (x_1,\ x_2,\ldots,\ x_N) = 0
\end{aligned}
$$

or in vector notation as

$$(1.2)\qquad \underline{f}(\underline{x}) = 0\ .$$

In applying Newton's Method (see e.g. [1]) to find a zero of (1.1), it is necessary to evaluate for each step of the iteration N functions and $N^2$ partial derivatives. In this paper we propose a method which involves about one half as many evaluations but which, like Newton's Method, converges quadratically.

After a complete description of the method in part two, we proceed in part three to prove the method converges locally and that the convergence is quadratic in nature; i.e., we show that whenever we are "sufficiently close" to the solution the method converges quadratically, without being able to write down explicitly the bounds for such a region of "closeness." In part four we obtain explicit

Kantorovich-type bounds [2] for a region of convergence in connection with a non-local convergence theorem. These bounds are computable in terms of the norms of $\underline{f}$ (see 1.2) and its first and second partial derivatives. Part five presents computer results obtained by applying the method to some specific non-linear systems; a comparison is made with some of the better recent methods. An operations count reveals that the method requires only $(\frac{N^2}{2} + \frac{3N}{2})$ function evaluations at each iterative step. Finally the Appendix contains the ALGOL procedure that was used to implement the method on the computer.

# 2. DESCRIPTION OF THE METHOD

## 2.1 Notation

We shall introduce most of the notation as needed; however some comments concerning the symbols for partial differentiation are in order here. If we are given a function G(u, v) in which

$$u = u(x,y)$$
$$v = v(x,y),$$

then we adopt the following conventions:

$$(2.1.1) \quad G_x \equiv \frac{\partial G}{\partial x} = G_u \frac{\partial u}{\partial x} + G_v \frac{\partial v}{\partial x} = G_u u_x + G_v v_x$$

$$G_y \equiv \frac{\partial G}{\partial y} = G_u \frac{\partial u}{\partial y} + G_v \frac{\partial v}{\partial y} = G_u u_y + G_v v_y,$$

whereas

$$(2.1.2) \quad G_1 \equiv G_u$$
$$G_2 \equiv G_v;$$

thus $f_{i_{x_j}} \equiv \dfrac{\partial f_i}{\partial x_i}$ is meant in the sense of (2.1.1) and

$f_{ij} \equiv f_{i,j} \equiv (f_i)_j$ is meant in the sense of (2.1.2).

## 2.2 Newton's Method

Let $\underline{x}^o = (x_1^o, \ldots, x_N^o)$ be an initial approximation to a solution of (1.1). Newton's Method then consists of applying the iteration

$$\underline{x}^{n+1} = \underline{x}^n - J^{-1}(\underline{f}^n)\,\underline{f}^n \quad (n=0,\ 1,\ \ldots) \tag{2.2.1}$$

where $J(\underline{f})$ is the Jacobian matrix

$$J(\underline{f}) = \begin{bmatrix} f_{1x_1} & f_{1x_2} & \cdots & f_{1x_N} \\ f_{2x_1} & f_{2x_2} & \cdots & f_{2x_N} \\ \cdot & \cdot & \cdot & \cdot \\ f_{Nx_1} & f_{Nx_2} & \cdots & f_{Nx_N} \end{bmatrix} \tag{2.2.2}$$

and the superscript n means that all functions involved are to be evaluated at $\underline{x} = \underline{x}^n$. The following local convergence theorem is well known for Newton's Method (see e.g. [1]).

<u>THEOREM 2.2.1</u> <u>If</u>

(1) <u>in a closed region R containing the solution</u> $\underline{x} = \underline{r}$ <u>of</u> (1.2), $f_{i_{x_j}}$ <u>is continuous for all</u> $i = 1, \ldots, N;\ j = 1, \ldots, N$,

(2) $J(\underline{f})$ <u>is non-singular at</u> $\underline{x} = \underline{r}$, <u>and</u>

(3) $\underline{x}^o$ <u>is chosen sufficiently close to the root</u> $\underline{x} = \underline{r}$,

<u>then the iteration</u> (2.2.1) <u>is convergent to</u> $\underline{r}$.

## 2.3 A Modified Newton-Type Method

We again seek a solution of (1.1) and shall assume that conditions (1), (2) and (3) of theorem 2.2.1 hold. Let $\underline{x}^n$ denote an $n^{th}$ approximation to the root $\underline{x} = \underline{r}$ of (1.1). The method consists of applying the following steps:

Step 1. Expand $f_1(x)$ in a Taylor series about the point $\underline{x}^n$; retain only first order terms and thus obtain the linear approximation

$$f_1(\underline{x}) = f_1(\underline{x}^n) + f_{1_{x_1}}(\underline{x}^n)\,(x_1 - x_1^n)$$

$$(2.3.1) \qquad + f_{1_{x_2}}(\underline{x}^n)\,(x_2 - x_2^n) + \ldots + f_{1_{x_N}}(\underline{x}^n)(x_N - x_N^n).$$

Equate the right side of (2.3.1) to zero and solve for that variable, say $x_N$, whose corresponding partial derivative is largest in absolute value. By assumption (2) of theorem 2.2.1 there must exist at least one non-vanishing partial derivative, say $f_{1_{x_N}}(\underline{x}^n)$, so that such an explicit solution can always be carried out. Thus,

$$x_N = x_N^n - \frac{f_{1_{x_1}}(\underline{x}^n)}{f_{1_{x_N}}(\underline{x}^n)}\,(x_1 - x_1^n) -$$

$$(2.3.2) \qquad \frac{f_{1_{x_2}}(\underline{x}^n)}{f_{1_{x_N}}(\underline{x}^n)}\,(x_2 - x_2^n) - \ldots - \frac{f_{1_{x_{N-1}}}(\underline{x}^n)}{f_{1_{x_N}}(\underline{x}^n)}\,(x_{N-1} - x_{N-1}^n)$$

$$\frac{-f_1(\underline{x}^n)}{f_{1_{x_N}}(\underline{x}^n)}$$

The constants $f_{1_{x_j}}(\underline{x}^n)/f_{1_{x_N}}(\underline{x}^n)$, $j=1,\ldots,N-1$, and $f_1(\underline{x}^n)/f_{1_{x_N}}(\underline{x}^n)$ are stored for future use.

We rewrite (2.3.2) in the form

$$(2.3.3) \qquad b_N = x_N^n - \sum_{j=1}^{N-1} \frac{f_{1_{x_j}}(\underline{x}^n)}{f_{1_{x_N}}(\underline{x}^n)} (x_j - x_j^n) - \frac{f_1(\underline{x}^n)}{f_{1_{x_N}}(\underline{x}^n)} .$$

<u>Step 2</u>. Define

$$\bar{f}_2 = f_2(x_1, \ldots, x_{N-1}, b_N);$$

so that $\bar{f}_2$ is a function of $(x_1, \ldots, x_{N-1})$. Again, expand $\bar{f}_2$ in a Taylor series, this time about the point $(x_1^n, \ldots, x_{N-1}^n)$, linearize and solve for that variable, say $x_{N-1}$, whose corresponding partial derivative is largest in magnitude obtaining

$$(2.3.4) \qquad x_{N-1} = x_{N-1}^n - \sum_{j=1}^{N-2} \frac{\bar{f}_{2_{x_j}}}{\bar{f}_{2_{x_{N-1}}}} (x_j - x_j^n) - \frac{\bar{f}_2}{\bar{f}_{2_{x_{N-1}}}} .$$

We shall show in lemma 2.5.1 that this process is well-defined; i.e., that there actually exists at least one non-vanishing partial derivative at each stage of the process.

We think of (2.3.3) and (2.3.4) rewritten as

$$(2.3.5) \qquad b_N = x_N^n - \sum_{j=1}^{N-2} \frac{f_{1_{x_j}}(\underline{x}^n)}{f_{1_{x_N}}(\underline{x}^n)} (x_j - x_j^n) - \frac{f_{1_{x_{N-1}}}(\underline{x}^n)}{f_{1_{x_N}}(x^n)} \cdot (b_{N-1} - x_{N-1}^n) - \frac{f_1(\underline{x}^n)}{f_{1_{x_N}}(\underline{x}^n)}$$