GAUSSIAN QUADRATURE NEWTON'S METHOD FOR FINDING ROOTS OF NONLINEAR EQUATIONS

A Thesis

By

G. H. J. Lanel

Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science

In

Industrial Mathematics

Department of Mathematics

Faculty of Graduate Studies

University of Sri Jayewardenepura

October 2002

DECLARATION

I hereby certify that this project report is my own work and it has never been submitted for any degree program.

.... Date: 05. 12. 2002

G. H. J. Lanel

We approve the Masters degree in Industrial Mathematics thesis of G. H. J. Lanel.

.... Date: 2002. 12. 05

Prof. Sunethra Weerakoon
Principal Supervisor
Department of Mathematics
University of Sri Jayewardenepura.

.... Date: 15/12/2002

Dr. G. K. Watugala
Supervisor
Department of Mechanical Engineering
University of Moratuwa.

.... Date: 05/12/2002

Mr. H. K. G. de Z. Amarasekera
Supervisor
B. Sc.; M. Phil. (Sri Jayewardenepura).
38 Gunananda Mawatha,
Panadura.

ABSTRACT

In this research we approximate the area under the curve appearing in the Newton's theorem by 2-point Gaussian quadrature formula. With the help of that we present an improvement to Newton's method for root finding. This iterative method converges to the root much faster and we have proved that it is third order convergent. The Established theory is supported by computed results by applying the new method to a wide range of functions and comparing it with the Newton's method and evaluating the computational order of convergence.

Algorithms constructed were implemented by using the computer language Turbo C^{++} and mathematical package Maple (version 6) was used for graphics.

Keywords: Newton's formula, Gaussian quadrature, Iterative methods, Order of convergence, Function evaluations.

CONTENTS

ACKNOWLEDGEMENT

I would like to take this opportunity to convey my heartfelt gratitude to my supervisors Prof. Sunethra Weerakoon, Dr. G. K. Watugala and Mr. H. K. G. de Z. Amarasekera for their tremendous contribution throughout the study.

I am grateful to Mr. N. W. A. I. Thilakasiri, Ms. P. S. S. Kumarasinghe, Mr. U. Angammana and Mr. W. A. S. C. Gunathilake for nurturing this project throughout its duration.

Special thanks go to the Head and the staff of the Department of Mathematics and Computer Sciences of the Open University of Sri Lanka for encouraging me during this period of research work.

I wish to express my sincere appreciation to my loving parents and family members for providing the kind of support that allowed me to focus on this study.

Finally I would like to thank all who have helped me to make this endeavour a success.

CHAPTER 1

INTRODUCTION

Problem of solving nonlinear equations may arise in many situations. Let us discuss such a case. The growth of large populations can be modeled over a short periods of time by assuming that the population grows continuously with time at a rate proportional to the number of individuals present at that time. If we let $M(t)$ denote the number of individuals at time t and a denote the constant birth rate of the population, the population satisfies the differential equation

$$\frac{dM(t)}{dt} = a\,M(t) \tag{1.1}$$

This model is valid only when the population is isolated with no immigration from outside the community. If immigration is permitted at a constant rate b, the differential equation governing the situation becomes

$$\frac{dM(t)}{dt} = a\,M(t) + b \tag{1.2}$$

whose solution is

$$M(t) = M_0 e^{at} + \frac{b}{a}(e^{at} - 1) \tag{1.3}$$

Suppose a certain population contains five million individuals initially, those 700,000 individuals immigrate into the community in the first year, and that 1,600,000 individuals are present at the end of one year. Determination of the birth rate of this population necessitates solving for a in the equation

$$10e^{a} + \frac{7}{a}(e^{a} - 1) = 16. \tag{1.4}$$

The numerical methods are used to find approximations to solutions of equations of this type (1.4), when the exact solutions cannot be obtained by algebraic methods.

The problem is called a root finding problem and it consists of finding values of the variable x that satisfy the equation $f(x)=0$, for a given function *f*. A solution to this problem is called a zero of *f* or a root of $f(x)=0$. This is one of the oldest numerical approximation problems and yet research is continuing actively in this area even at present. The most popular procedure is the Newton-Rapshon method, basically developed by Issac Newton over 300 years ago.

Newton's method, which approximates the roots of a nonlinear equation in one variable using the value of the function and its derivative, in an iterative fashion, is probably the best known and most widely used algorithm and it converges to the root quadratically.

In this study we suggest an improvement to the iterations of Newton's method. Derivation of Newton's method involves an indefinite integral of the derivative of the function and the relevant area is approximated by a rectangle. In the proposed method we approximate this indefinite integral by applying a Gaussian Quadrature formula instead of a rectangle, thereby reducing the error in the approximation. It is shown that the order of convergence of the new method is three, and computed result supports this theory. Even though we have shown that the order of convergence is three, in several cases, computational order of convergence is even higher. For most of the functions we tested, the order of convergence for Newton's method was less than two and for the new method it was always close to three.

It is also shown that for quite a number of nonlinear functions the number of iterations required for the new method is less than that of the Newton's method.

CHAPTER 2
PRELIMINARIES

Mean Value Theorem 2.1

If $f \in C[a, b]$ and f is differentiable on (a, b), then a number $c, a < c < b,$ exists such that

$$f'(c) = \frac{f(b) - f(a)}{b - a} \tag{2.1}$$

Intermediate Value Theorem 2.2

If $f \in C[a, b]$ and k is any number between $f(a)$ and $f(b)$, then there exists c in (a, b) for which $f(c) = k$.

Taylor's Theorem 2.3

Suppose $f \in C^n[a, b]$ and f^{n+1} exists on $[a, b]$. Let $x_0 \in [a, b]$. For every $x \in [a, b]$, there exists $\zeta(x)$ between x_0 and x with

$$f(x) = P_n(x) + R_n(x),$$

where

$$P_n(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

$$= \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k, \tag{2.2}$$

and

$$R_n(x) = \frac{f^{(n+1)}(\zeta(x))}{(n+1)!}(x - x_0)^{n+1}.$$

Here $P_n(x)$ is called the n^{th} - degree Taylor polynomial for f about x_0 and $R_n(x)$ is called the remainder term associated with $P_n(x)$. The infinite series obtained by taking the limit of $P_n(x)$ as $n \to \infty$ is called the Taylor series for f about x_0.

Definition 2.4: Orthogonal set of functions

$\{\phi_0, \phi_1, \cdots, \phi_n\}$ is said to be an orthogonal set of functions for the interval $[a, b]$ with respect to the weight function w if

$$\int_a^b w(x)\phi_j(x)\phi_k(x)dx = \begin{cases} 0, & \text{whenever} \quad j \neq k, \\ \alpha_k, & \text{whenever} \quad j = k. \end{cases} \tag{2.3}$$

where $\alpha_k \in \mathbb{R}$

Definition 2.5: Lagrange interpolating polynomial

If $x_0, x_1, \cdots, x_n$ are $(n+1)$ distinct numbers and f is a function whose values are given at these numbers, then there exists a unique polynomial P of degree at most n with the property that

$$f(x_k) = P(x_k) \qquad \text{for each } k = 0, 1, 2, \cdots, n.$$

This polynomial is given by

$$P(x) = f(x_0)l_{n,0}(x) + \cdots + f(x_n)l_{n,n}(x) = \sum_{k=0}^{n} f(x_k)l_{n,k}(x), \tag{2.4}$$

where

$$l_{n,k}(x) = \frac{(x-x_0)(x-x_1)\cdots(x-x_{k-1})(x-x_{k+1})\cdots(x-x_n)}{(x_k-x_0)(x_k-x_1)\cdots(x_k-x_{k-1})(x_k-x_{k+1})\cdots(x_k-x_n)}$$

$$= \prod_{\substack{i=0 \\ i\neq k}}^{n} \frac{(x-x_i)}{(x_k-x_i)} \quad \text{for each } k = 0, 1, 2, \cdots, n. \tag{2.5}$$

Definition 2.6: Legendre polynomial $\{p_n\}$

The set of Legendre polynomials $\{p_0, p_1, \cdots, p_n\}$ defined in the following way is orthogonal on $[a,b]$ with respect to the weight function w.

$$p_0(x) \equiv 1, \quad p_1(x) \equiv x - B_1, \qquad \text{for each } a \leq x \leq b,$$

where $B_1 = \dfrac{\int_a^b x\, w(x)[p_0(x)]^2\, dx}{\int_a^b w(x)[p_0(x)]^2\, dx}$;

and when $k \geq 2$,

$$p_k(x) = (x - B_k)p_{k-1}(x) - C_k p_{k-2}(x) \qquad \text{for each } a \le x \le b, \tag{2.6}$$

where $B_k = \dfrac{\int_a^b x\, w(x)[p_{k-1}(x)]^2\, dx}{\int_a^b w(x)[p_{k-1}(x)]^2\, dx}$ and $C_k = \dfrac{\int_a^b x\, w(x)p_{k-1}(x)p_{k-2}(x)dx}{\int_a^b w(x)[p_{k-2}(x)]^2\, dx}$.

Definition 2.7

A function f is Lipschitz continuous with constant γ in a set X, written $f \in \mathbf{Lip}_\gamma(X)$, if for every $x, y \in X$,

$$| f(x) - f(y) | \le \gamma\, | x - y |. \tag{2.7}$$

Lemma 2.7.1

For an open interval A, let $f : A \to \mathbb{R}$ and let $f' \in \mathbf{Lip}_\gamma(A)$. Then for any $x, y \in A$,

$$| f(y) - f(x) - f'(x)(y - x) | \le \frac{\gamma(y - x)^2}{2}. \tag{2.8}$$

Lemma 2.7.2

For an open interval A, let $f : A \to \mathbb{R}$ and let $f' \in \mathbf{Lip}_\gamma(A)$. Then for any $x, y \in A$,

$$| f(y) - f(x) - f'(x)(y - x) | \le \frac{\gamma\, | y - x |^2}{2}. \tag{2.9}$$

Definition 2.8

An iterative method is said to be of *order p* or has the *rate of convergence p*, if p is the largest positive real number for which

$| e_{n+1} | \le c\, | e_n |^p$, where $e_n = x_n - \alpha$ is the error in the n^{th} iterate. (2.10)

The constant c is called the asymptotic error constant. It depends on various order derivatives of $f(x)$ evaluated at α and is independent of n.

The relation

$$e_{n+1} = c e_n^p + 0(e_n^{p+1}) \tag{2.11}$$

is called the *error equation*.

By substituting $e_i = x_i - \alpha$ for all i in any iteration method and simplifying we obtain the error equation for that method. The value of p thus obtained is called the order of this method.

Definition 2.9

Let α be a root of the function *f(x)* and suppose that x_{n-1}, x_n and x_{n+1} be the consecutive iterations closer to the root, α.

Then the *Computational Order of Convergence* ρ can be approximated by:

$$\rho \approx \frac{\ln\left|\frac{(x_{n+1} - \alpha)}{(x_n - \alpha)}\right|}{\ln\left|\frac{(x_n - \alpha)}{(x_{n-1} - \alpha)}\right|} \tag{2.12}$$

Definition: 2.10: Stopping criteria

We have to accept an approximate solution rather than the exact root, depending on the precision (ε) of the computer. So we adopt following stopping criteria for computer programs.

i. $|f(x_{n+1})| < \sqrt{\varepsilon}$

ii. $|x_{n+1} - x_n| < \sqrt{\varepsilon}$

Definition 2.11: A function of two variables

Let $A \subset \mathbb{R}$. A function f of two variables is a rule that assigns to each ordered pair (x, y) in A, a unique real number denoted by $f(x, y)$. The set D is the domain of f and its range is the set of values that f takes on, that is $\{f(x, y) \mid (x, y) \in A\}$. We often write $z = f(x, y)$ to make explicit the value taken on by f at the general point (x, y). The variables x and y are independent variables and z is the dependent variable.

Definition 2.12: Gradient of a function

A continuous function $f:\mathbb{R}^2 \to \mathbb{R}$ is said to be continuously differentiable at $\mathbf{x}=(x,y)^T \in \mathbb{R}^2$, if $f_x(x,y)$ and $f_y(x,y)$ exist and continuous. The gradient of f at (x,y) is then defined as

$$\nabla f(x,y) = [f_x(x,y), f_y(x,y)]^T \tag{2.13}$$

Lemma 2.12.1

Let $f:\mathbb{R}^2 \to \mathbb{R}$ be continuously differentiable in an open convex set $A \subset \mathbb{R}^2$. Then, for $\mathbf{x}=(x,y)^T \in A$ and nonzero perturbation $p=(p_1,p_2)^T \in \mathbb{R}^2$, the directional derivative of f at $\mathbf{x}$ in the direction $\mathbf{p}$, denoted by

$$D_{\mathbf{p}} f(\mathbf{x}) = D_{\mathbf{p}} f(x,y) = \lim_{h\to 0} \frac{f(x+hp_{1,}, y+hp_2) - f(x,y)}{h}$$

exists and is equal to $\nabla f(\mathbf{x})^T$ **.p**. for any $\mathbf{x}, \mathbf{x}+\mathbf{p} \in \mathbf{A}$

$$f(\mathbf{x}+\mathbf{p}) = f(\mathbf{x}) + \int_0^1 \nabla f(\mathbf{x}+t\mathbf{p})^T.\mathbf{p}\, dt \tag{2.14}$$

and there exists $\mathbf{z} \in (\mathbf{x}, \mathbf{x}+\mathbf{p})$ such that

$$f(\mathbf{x}+\mathbf{p}) = f(\mathbf{x}) + \nabla f(\mathbf{z})^T\ .\mathbf{p} \tag{2.15}$$

Proof:

We simply parameterize *f* along the line through **x** and $\mathbf{x}+\mathbf{p}$ as a function of one variable

$$g:\mathbb{R} \to \mathbb{R},\ g:(t) = f(\mathbf{x}+t\mathbf{p}) = f(x+tp_1, y+tp_2)$$

and apply calculus of one variable to the function *g*.

Differentiating with respect to *t*

$$\begin{aligned} g'(t) &= f_x(x+tp_1, y+tp_2)p_1 + f_y(x+tp_1, y+tp_2)p_2 \\ &= \nabla f(x+tp_1, y+tp_2)^T.\mathbf{p} \\ &= \nabla f(\mathbf{x}+t\mathbf{p})^T.\mathbf{p} \end{aligned} \tag{2.16}$$

Then by the fundamental theorem of calculus or Newton's theorem,

$$g(1) = g(0) + \int_0^1 g'(t)dt$$

which, by the definition of g and (2.16), is equivalent to

$$f(\mathbf{x}+\mathbf{p}) = f(\mathbf{x}) + \int_0^1 \nabla f(\mathbf{x}+t\mathbf{p})^T.\mathbf{p}\, dt$$

and proves (2.14). Finally, by the mean value theorem for functions of one variable,

$$g(1) = g(0) + g'(\zeta), \qquad \zeta \in (0,1)$$

which by the definition of g and (2.16), is equivalent to

$$f(\mathbf{x}+\mathbf{p}) = f(\mathbf{x}) + \nabla f(\mathbf{x}+\zeta\mathbf{p})^T.\mathbf{p} \qquad \zeta \in (0,1)$$

and proves (2.15).

Definition 2.13

A continuously differentiable function $f:\mathbb{R}^2 \to \mathbb{R}$ is said to be twice continuously differentiable at $\mathbf{x} \in \mathbb{R}^2$, if $(\partial^2 f / \partial x_i \partial x_j)(\mathbf{x})$ exists and is continuous, $1 \le i, j \le 2$. The Hessian of f at $\mathbf{x}$ is then defined as the 2×2 matrix whose $(i,j)^{\text{th}}$ element is

$$\nabla^2 f(\mathbf{x})_{i,j} = \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j}, \quad 1 \le i, j \le 2 \tag{2.17}$$

Clairaut's Theorem 2.14

Suppose f is defined on $\boldsymbol{A} \subset \mathbb{R}^2$ that contains the point (a,b). If the functions f_{xy} and f_{yx} are both continuous on $\boldsymbol{A}$, then

$$f_{xy}(a,b) = f_{yx}(a,b). \tag{2.18}$$

Definition 2.15

A continuous function $\boldsymbol{F}:\mathbb{R}^2 \to \mathbb{R}^2$ is continuously differentiable at $\mathbf{x} \in \mathbb{R}^2$ if each component function $f(x,y)$ and $g(x,y)$ is continuously differentiable at $\mathbf{x}$. The derivative of $\boldsymbol{F}$ at $\mathbf{x}$ is called the Jacobian of $\boldsymbol{F}$ at $\mathbf{x}$, and its transpose is sometimes called the gradient of $\boldsymbol{F}$ at $\mathbf{x}$.