# A SECANT METHOD BASED ON CUBIC INTERPOLATION FOR SOLVING ONE DIMENSIONAL OPTIMIZATION PROBLEMS

by

C. Kirjner Neto and E. Polak

Memorandum No. UCB/ERL M91/91

15 October 1991

# A SECANT METHOD BASED ON CUBIC INTERPOLATION FOR SOLVING ONE DIMENSIONAL OPTIMIZATION PROBLEMS

by

C. Kirjner Neto and E. Polak

Memorandum No. UCB/ERL M91/91

15 October 1991

**ELECTRONICS RESEARCH LABORATORY**

College of Engineering
University of California, Berkeley
94720

# A SECANT METHOD BASED ON CUBIC INTERPOLATION FOR SOLVING ONE DIMENSIONAL OPTIMIZATION PROBLEMS

by

C. Kirjner Neto and E. Polak

Memorandum No. UCB/ERL M91/91

15 October 1991

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

# A SECANT METHOD BASED ON CUBIC INTERPOLATION FOR SOLVING ONE DIMENSIONAL OPTIMIZATION PROBLEMS†

by

**C. Kirjner Neto and E. Polak**
Department of Electrical Engineering
and Computer Sciences
University of California
Berkeley, CA 94720, U.S.A.

**ABSTRACT**

We present two versions of a secant method for one-dimensional minimization, that use cubic interpolation through two successive points to obtain an estimate of the second derivative of the function. The first version uses both function and derivative values, while the second version uses only function values. Both versions are shown to be globally $R$-quadratically convergent on sufficiently smooth functions.

**Key words.** Line search, one dimensional minimization, secant method, cubic interpolation

**AMS(MOS) subject classification.** primary 49D37; secondary 65K10

---

† The research reported herein was sponsored by the Air Force Office of Scientific Research contract AFOSR-90-0068, the National Science Foundation grant ECS-8916168.

## 1. Introduction

Within the general realm of optimization algorithms, one-dimensional optimization algorithms play the important role of subprocedures for step length calculation, i.e., for the solution of problems of the form $\min_{x \in \mathbb{R}} g(y + xh)$, with $g : \mathbb{R}^n \to \mathbb{R}$. For example, the performance of various methods of conjugate directions, such as the Fletcher-Reeves and Polak- Ribière methods [4,12], and of some variable metric methods, such as the Davidon-Fletcher-Powell method [2,3,5], depends both on the efficiency and the accuracy of one-dimensional optimization methods.

The family of one-dimensional optimization algorithms includes the golden section search [8,11], versions of the secant method (see e.g. [6]), and algorithms based on either cubic, or quadratic interpolation of the function to be minimized (see, e.g. [8,14]). The most commonly used one-dimensional optimization algorithms for step length calculation are based on cubic and quadratic interpolation. The version of the quadratic interpolation method considered in [7] converges superlinearly with rate 1.3, but cannot be shown to be globally convergent. The version of the quadratic interpolation method presented in [8] converges globally, but has not been shown to be superlinear. The cubic interpolation method can be shown to be locally quadratically convergent on sufficiently smooth functions (see [15]), but it may not be globally convergent.

In this paper, we present two versions of a new global one-dimensional optimization algorithm that retains the robust structure of the globally convergent secant method presented in [6], while obtaining quadratic convergence by using cubic interpolation to construct approximations of the second derivative. The first version requires the computation of first derivatives, while the second version uses function values only. In Section 2, we present the first version of our secant algorithm and show that under a commonly used assumption, it converges with the secant $R$-rate of 1.618 on twice continuously differentiable functions, and $R$-quadratically on four times continuously differentiable functions. In Section 3, we present the second version of our algorithm and show that it retains the convergence properties of the first one. In Section 4, we present a few numerical examples which demonstrate that both versions of our algorithm perform as predicted, and yield results with adequately high accuracy for use in step length calculation. Although the discrete version of the algorithm is capable of less accuracy than the exact version, it is bound to be much more efficient in step length calculation problems, because of the high cost of derivative calculation in such problems.

## 2. The Cubic-Secant Algorithm

Consider the problem

$$\min f(x), \tag{2.1}$$

where $f : \mathbb{R} \rightarrow \mathbb{R}$ is at least twice locally Lipschitz differentiable. We will denote the first and second derivatives of $f(\cdot)$ by $f'(\cdot)$ and $f''(\cdot)$, respectively, while for $k > 2$, the $k\text{-}th$ derivative will be denoted by $f^{(k)}(\cdot)$.

We begin by recalling the "classical" stabilized secant method with Armijo step size rule [1] for solving (2.1) (see, e.g. [5]). This method is defined by the recursion

$$x_{i+1} = x_i - \lambda_i \tilde{f}''(x_i, x_{i-1})^{-1} f'(x_i) \tag{2.2}$$

where $\tilde{f}''(x_i, x_{i-1})$ is an approximation to the second derivative, $f''(x_i)$, of $f(\cdot)$ at $x_i$, defined by

$$\tilde{f}''(x_i, x_{i-1}) \triangleq \frac{f'(x_i) - f'(x_{i-1})}{x_i - x_{i-1}} \tag{2.3}$$

and the step size $\lambda_i$ is computed according to the rule

$$\lambda_i = \max_{k \in \mathbb{N}} \left\{ \beta^k \mid f(x_i - \beta^k \tilde{f}''(x_i, x_{i-1})^{-1} f'(x_i)) - f(x_i) \le -\alpha \beta^k \tilde{f}''(x_i, x_{i-1})^{-1} f'(x_i)^2 \right\} \tag{2.4}$$

with $\alpha \in (0, ½)$, and $\beta \in (0, 1)$.

Referring to [8], we see that on strictly convex problems, this method converges superlinearly, with $R$-rate at least 1.618. Moreover (as shown in [5]), it can be adapted for the solution of the much broader class of problems with local minimizers $\hat{x}$ such that $f''(\hat{x}) > 0$, by adding a crossover mechanism which switches over to the Armijo gradient method when $f''(x_i) > 0$ does not hold.

The algorithm that we will describe in this section differs from the "classical" stabilized secant method by the important feature that it does not approximate the second derivative $f''(x_i)$ by $\tilde{f}''(x_i, x_{i-1})$ defined in (2.3), but by the second derivative, $p''_{i,i-1}(x_i)$, of a cubic polynomial $p_{i,i-1}(x) \triangleq a_{3,i}x^3 + a_{2,i}x^2 + a_{1,i}x + a_{0,i}$, whose coefficients are computed by osculatory cubic-fit interpolation of the objective function through two points, as follows:

$$p_{i,i-1}(x_j) = f(x_j), \quad j = i-1, i \tag{2.5}$$

$$p'_{i,i-1}(x_j) = f'(x_j), \quad j = i-1, i. \tag{2.6}$$

The equations (2.5) and (2.6) can be rewritten in matrix form :

$$\begin{bmatrix} x_i^3 & x_i^2 & x_i & 1 \\ x_{i-1}^3 & x_{i-1}^2 & x_{i-1} & 1 \\ 3x_i^2 & 2x_i & 1 & 0 \\ 3x_{i-1}^2 & 2x_{i-1} & 1 & 0 \end{bmatrix} \begin{bmatrix} a_{3,i} \\ a_{2,i} \\ a_{1,i} \\ a_{0,i} \end{bmatrix} = \begin{bmatrix} f(x_i) \\ f(x_{i-1}) \\ f'(x_i) \\ f'(x_{i-1}) \end{bmatrix} . \tag{2.7}$$

After some manipulation of the system of equations (2.7), we obtain the following explicit expression for $p''_{i,i-1}(x) = 6a_{3,i}x + a_{2,i}$,

$$p''_{i,i-1}(x) = \frac{2c_i}{x_i - x_{i-1}} + \frac{2(x - x_{i-1}) + 2(2x - x_i - x_{i-1})}{(x_i - x_{i-1})^2} d_i , \tag{2.8a}$$

where

$$c_i \triangleq \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}} - f'(x_{i-1}) , \tag{2.8b}$$

$$d_i \triangleq f'(x_i) - 2\frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}} + f'(x_{i-1}) . \tag{2.8c}$$

We are now ready to state our new algorithm formally. Note that, like the secant method we have discussed earlier, it crosses over to the Armijo gradient method whenever $p''_{i,i-1}(x_i)$ is not sufficiently positive.

**Cubic-Secant Algorithm 2.1**

*Parameters.* $\alpha \in (0, ½)$, $\beta \in (0, 1)$, $m > 0$, small.

*Data.* $x_0, x_{-1} \in [0, \infty)$.

*Step 0.* Set $i = 0$.

*Step 1.* Compute $p''_{i,i-1}(x_i)$ using (2.8a)-(2.8c).

If $p''_{i,i-1}(x_i) \geq m$, set

$$h_i = -p''_{i,i-1}(x_i)^{-1} f'(x_i) . \tag{2.9}$$

Else, set

$$h_i = -f'(x_i) . \tag{2.10}$$

*Step 2.* Compute the *Armijo step size*

$$\lambda_i = \max_{k \in \mathbb{N}} \left\{ \beta^k \mid f(x_i + \beta^k h_i) - f(x_i) \le \alpha \beta^k h_i f'(x_i) \right\} . \tag{2.11}$$

*Step 3.* Set $x_{i+1} = x_i + \lambda_i h_i$, replace $i$ by $i + 1$ and go to Step 1. □

**Assumption 2.2.** The function $f(\cdot)$ is continuously differentiable and bounded from below. □

**Theorem 2.3.** Suppose that Assumption 2.2 is satisfied and that the Cubic-Secant Algorithm has constructed an infinite sequence $\{ x_i \}_{i=0}^{\infty}$. Then every accumulation point $\hat{x}$ of $\{ x_i \}_{i=0}^{\infty}$ satisfies $f'(\hat{x}) = 0$.

*Proof.* It follows from (2.9) and (2.10) that for any $x_i \in \mathbb{R}$,

$$f'(x_i)h_i \le -\min\{ 1, \frac{1}{m} \} f'(x_i)^2 \le 0 . \tag{2.12}$$

$$| h_i | \le \max\{ 1, \frac{1}{m} \} | f'(x_i) | . \tag{2.13}$$

In view of (2.12) and (2.13), the desired result follows directly from the Polak-Sargent-Sebastian Theorem [12]. □

**Assumption 2.4.** The function $f(\cdot)$ is twice locally Lipschitz continuously differentiable and bounded from below. □

**Corollary 2.5.** Suppose that Assumption 2.4 is satisfied and that the sequence $\{ x_i \}_{i=0}^{\infty}$, constructed by the Cubic-Secant Algorithm 2.1, has an accumulation point $\hat{x}$ such that $f''(\hat{x}) \ge b$, for some $b > 0$. Then $x_i \to \hat{x}$ as $i \to \infty$.

*Proof.* Clearly, there exists a $\rho > 0$ such that for all $x \in B(\hat{x}, \rho)$, $f''(x) \ge b/2$, and hence $f(x) \ge f(\hat{x}) + \frac{1}{2} b (x - \hat{x})^2$ for all $x \in B(\hat{x}, \rho)$. Since $\hat{x}$ is an accumulation point of $\{ x_i \}_{i=0}^{\infty}$, there exists an infinite subset $K \subset \mathbb{N}$ such that $x_i \to^K \hat{x}$ as $i \to \infty$. Since the cost sequence $\{ f(x_i) \}_{i=0}^{\infty}$ is monotone decreasing, it converges to $f(\hat{x})$ and hence there exists an $i_0$ such that for all $i \ge i_0$, $f(x_i) \le f(\hat{x}) + \rho^2 b/8$, which implies that for all $i \ge i_0$, $i \in K$, $x_i \in B(\hat{x}, \rho/2)$. Next, there exists an $i_1 \ge i_0$ such that for all $i \ge i_1$, $i \in K$, $| h_i | \le \rho/2$. Since the step size $\lambda_i \le 1$, we conclude that for any $i \in K$, $i \ge i_1$, $x_{i+1} \in B(\hat{x}, \rho/2)$, i.e., that for all $i \ge i_1$, $x_i \in B(\hat{x}, \rho/2)$, which contains the unique stationary point $\hat{x}$. Since $B(\hat{x}, \rho/2)$ is compact, and $\hat{x}$ is the only stationary point in $B(\hat{x}, \rho)$, it now follows from Theorem 2.3 that $x_i \to \hat{x}$ as $i \to \infty$. □

We are now ready to establish the rate of convergence of the Cubic-Secant Algorithm 2.1. First we will show that it converges superlinearly.

**Lemma 2.6.** Suppose that Assumption 2.4 is satisfied. Then for any $\rho > 0$, and any $\hat{x} \in \mathbb{R}$, there exits a finite $L > 0$ such that

$$| p''_{i,i-1}(x_i) - f''(x_i) | \leq L \, | x_i - x_{i-1} | \,, \tag{2.14}$$

for all $x_i$ , $x_{i-1} \in B(\hat{x}, \rho) \triangleq [\hat{x} - \rho, \hat{x} + \rho]$ .

*Proof.* Since by Assumption 2.4, $f''(\cdot)$ is locally Lipshitz continuous, and, by inspection, so is $p''_{i,i-1}(\cdot)$, it follows that for any $\rho > 0$, and any $\hat{x} \in \mathbb{R}$, there exists a finite $L > 0$ such that for all $x_1, x_2 \in B(\hat{x}, \rho)$

$$| f''(x_1) - f''(x_2) | \leq \frac{2}{3} L \, | x_1 - x_2 | \,, \tag{2.15}$$

$$| p''_{i,i-1}(x_1) - p''_{i,i-1}(x_2) | \leq \frac{2}{3} L \, | x_1 - x_2 | \,. \tag{2.16}$$

Now consider now the system of equations (2.7). If we subtract the last equation from the third one and rearrange terms, we get from the Mean Value Theorem, that for some $s \in [0, 1]$

$$p''_{i,i-1}\left(\frac{x_i + x_{i-1}}{2}\right) = \frac{f'(x_i) - f'(x_{i-1})}{x_i - x_{i-1}} = f''(x_i + s(x_{i-1} - x_i)) \,. \tag{2.17}$$

Next, we conclude from the triangle inequality and (2.17) that

$$| p''_{i,i-1}(x_i) - f''(x_i) | \leq | p''_{i,i-1}(x_i) - p''_{i,i-1}\left(\frac{x_i + x_{i-1}}{2}\right) | + | p''_{i,i-1}\left(\frac{x_i + x_{i-1}}{2}\right) - f''(x_i) |$$

$$\leq | p''_{i,i-1}(x_i) - p''_{i,i-1}\left(\frac{x_i + x_{i-1}}{2}\right) | + | f''(x_i + s(x_{i-1} - x_i)) - f''(x_i) | \,. \tag{2.18}$$

The desired result now follows directly from (2.15), (2.16), and the fact that $s \in [0, 1]$. □

**Theorem 2.7.** Suppose that Assumption 2.4 is satisfied, and that the Cubic-Secant Algorithm 2.1 has constructed a sequence $\{ x_i \}_{i=0}^{\infty}$ that has an accumulation point $\hat{x}$ such that $f''(\hat{x}) \geq 3m$, with $m > 0$ as used in Algorithm 2.1. Then $x_i \to \hat{x}$, $R$-superlinearly, with root rate at least $\tau_1 \approx 1.618$.

*Proof.* Clearly, there exists a $\rho > 0$ such that for all $x \in B(\hat{x}, \rho)$ and any $s \in [0, 1]$, $f''(\hat{x} - s(\hat{x} - x)) \geq 2m$. Next, it follows from Corollary 2.5 and Lemma 2.6, that $x_i \to \hat{x}$ as $i \to \infty$, and that there exists an $i_0 \in \mathbb{N}$ such that for all $i \geq i_0$, $p''_{i,i-1}(x_i) \geq m$, so that the search direction $h_i$

is given by (2.9).

Now, making use of the second order Taylor expansion formula, we find that

$$f(x_i + h_i) - f(x_i) - \alpha h_i f'(x_i) = (1-\alpha) h_i f'(x_i) + \int_0^1 (1-s) f''(x_i + s h_i) h_i^2 \, ds$$

$$= -(1 - \frac{1}{2} - \alpha) h_i^2 p''_{i,i-1}(x_i) + \int_0^1 (1-s)[f''(x_i + s h_i) - p''_{i,i-1}(x_i)] h_i^2 \, ds \, , \tag{2.19}$$

where the second line was obtained by adding and subtracting $(1-s) h_i^2 p''_{i,i-1}(x_i)$ from the integrand, and using (2.9) to express $f'(x_i)$ in terms of $h_i$ and $p''_{i,i-1}(x_i)$. Adding and subtracting $(1-s) h_i^2 f''(x_i)$ from the integrand in the second line of (2.19), and using (2.14) and (2.15), we conclude that there exists an $i_1 \in \mathbb{N}$, $i_1 \geq i_0$, such that for all $i \in \mathbb{N}$, $i \geq i_1$,

$$f(x_i + h_i) - f(x_i) - \alpha h_i f'(x_i) \leq h_i^2 \left\{ -(\frac{1}{2} - \alpha) m + L \left[ \frac{h_i}{3} + \frac{|x_i - x_{i-1}|}{2} \right] \right\} \tag{2.20}$$

and, since both $x_i - x_{i-1} \to 0$ and $h_i \to 0$ as $i \to \infty$, there exists an $i_2 \in \mathbb{N}$, $i_2 \geq i_1$, such that $f(x_i + h_i) - f(x_i) \leq \alpha h_i f'(x_i)$ for all $i \geq i_2$, which implies that $\lambda_i = 1$, for all $i \geq i_2$.

Next, it follows from Assumption 2.4 and Lemma 2.6 that for some finite $L > 0$,

$$|f''(x_i) - f''(\hat{x})| \leq \frac{2}{3} L \, |x_i - \hat{x}| \tag{2.21}$$

and

$$|p''_{i,i-1}(x_i) - f''(x_i)| \leq L \, |x_i - x_{i-1}| \tag{2.22}$$

for all $x_i, x_{i-1} \in B(\hat{x}, \rho)$. The desired result now follows from Theorem 11.2.7 in [9]. □

fm: Command not found.

**Lemma 2.9** [15, Theorem A-1] Suppose Assumption 2.8 is satisfied. Then for any $x_{i-1}, x_i \in \mathbb{R}$, there exists a $\xi \in \mathbb{R}$ in the interval with end points $x_{i-1}$ and $x_i$, such that

$$f''(x_i) - p''_{i,i-1}(x_i) = \frac{1}{12} f^{(4)}(\xi)(x_i - x_{i-1})^2 \, . \tag{2.23}$$

□

**Theorem 2.10.** Suppose that Assumption 2.8 is satisfied, and that the Cubic-Secant Algorithm 2.1 has constructed an infinite sequence $\{x_i\}_{i=0}^{\infty}$ that has an accumulation point $\hat{x}$, such that $f''(\hat{x}) \geq 3m$ with $m > 0$, as used in the Cubic-Secant Algorithm. Then $x_i \to \hat{x}$, as $i \to \infty$, *R*-quadratically.

*Proof.* The fact that $x_i \to \hat{x}$ as $i \to \infty$ with $f'(\hat{x}) = 0$ follows from Theorem 2.3 and Corollary

2.5. As was shown in the proof of in Theorem 2.7, there must exist a $\rho > 0$ and an $i_1 \in \mathbb{N}$ such that for any $i \in \mathbb{N}$, $i \geq i_1$, we have that $x_i \in B(\hat{x}, \rho)$, $p''_{i,i-1}(x_i) \geq m$, $h_i$ is given by (2.9), and $\lambda_i = 1$. Hence it follows that by construction of the $x_i$, for all $i \in \mathbb{N}$, $i \geq i_1$,

$$p''_{i,i-1}(x_i)(x_{i+1} - x_i) = -f'(x_i) + f'(\hat{x}) = -\int_0^1 f''(x_i + s(x_i - \hat{x}))\,ds\,(x_i - \hat{x})\,. \tag{2.24}$$

Adding and subtracting $f''(x_i)$ from the integrand in (2.24) and reordering terms, we get that

$$|p''_{i,i-1}(x_i)|\,|x_{i+1} - \hat{x}| \leq \int_0^1 \Big[\, |f''(x_i - s(x_i - \hat{x})) - f''(x_i)| + |f''(x_i) - p''_{i,i-1}(x_i)| \Big]\, ds\, |x_i - \hat{x}|\,. \tag{2.25}$$

Let

$$J \triangleq \max_{\xi \in B(\hat{x}, \rho)} \frac{|f^{(4)}(\xi)|}{12}\,. \tag{2.26}$$

Then, using (2.15), (2.23), and the fact that for all $i \geq i_1$, $p''_{i,i-1}(x_i) \geq m$, we obtain that

$$|x_{i+1} - \hat{x}| \leq \frac{1}{m}\left[ \frac{L}{3}|x_i - \hat{x}|^2 + J\,|x_i - x_{i-1}|^2 |x_i - \hat{x}| \right]\,. \tag{2.27}$$

Let

$$K \triangleq \frac{1}{m} \max\{ \frac{L}{3}\,; 2J\,; \frac{m}{4} \}\,, \tag{2.28}$$

$$e_i \triangleq 4K\,|x_i - \hat{x}|\,, \quad i \in \mathbb{N}\,, \tag{2.29}$$

with L as in (2.15). Then, writing $(x_i - x_{i-1})$ as $(x_i - \hat{x}) + (\hat{x} - x_{i-1})$, and using (2.28), we obtain from (2.27) that

$$|x_{i+1} - \hat{x}| \leq K\Big\{ |x_i - \hat{x}|^3 + |x_i - \hat{x}|^2 + |x_{i-1} - \hat{x}|\,|x_i - \hat{x}|^2 + |x_{i-1} - \hat{x}|^2 |x_i - \hat{x}| \Big\}\,. \tag{2.30}$$

Making use of (2.29) to simplify notation, we see that (2.30) becomes

$$e_{i+1} \leq 4K^2 \left\{ \frac{e_i^3}{64K^3} + \frac{e_i^2}{16K^2} + \frac{e_{i-1}e_i^2}{64K^3} + \frac{e_{i-1}^2 e_i}{64K^3} \right\} \tag{2.31}$$

Since $4K \geq 1$, it follows from (2.31) that for all $i \geq i_1$,

$$e_{i+1} \leq \frac{1}{4}(e_i^3 + e_i^2 + e_i^2 e_{i-1} + e_i e_{i-1}^2)\,. \tag{2.32}$$

To complete our proof, we proceed by induction. Since $x_i \to \hat{x}$, as $i \to \infty$, there exists an $i_2 \in \mathbb{N}$, $i_2 \geq i_1$ such that $e_{i_2+1} < e_{i_2} < 1$. Let $\delta^2 \triangleq \frac{1}{2} \max \{ e_{i_2} + e_{i_2+1}, e_{i_2}^2 + e_{i_2} \}$. Then we see that $\delta \in (0, 1)$, and that $e_{i_2} \leq \delta$ and $e_{i_2+1} \leq \delta^2$.

Assume now that for $k = i_2, i_2+1, \ldots, i_2+n$, we have $e_{i_2+k} \leq \delta^{2^k}$. From (2.32), with $i = i_2 + n$, we get

$$e_{i_2+n+1} \leq \frac{1}{4}(\delta^{3\,2^n} + \delta^{2\,2^n} + \delta^{(2 2^n + 2^{n-1})} + \delta^{(2 2^{n-1} + 2^n)})$$

Hence,

$$e_{i_2+n+1} \leq \frac{1}{4}(\delta^{3\,2^n} + \delta^{2^{n+1}} + \delta^{(\frac{5}{2} 2^n)} + \delta^{2^{n+1}}) \leq \delta^{2^{n+1}} \tag{2.33}$$

which gives the desired result. □

## 3. The Discrete Cubic-Secant Algorithm

Although the Cubic-Secant Algorithm 2.1 converges quadratically, it has to evaluate derivatives at each step. When the objective function $f : \mathbb{R} \to \mathbb{R}$ is a composite function of the form $f(x) = g(z + xh)$, with $g : \mathbb{R}^n \to \mathbb{R}$, as in the case of step size evaluations in conjugate gradient methods, the computation of derivatives can impose a considerable computational burden. In such cases, it is desirable to approximate $f'(x)$ by a finite difference of the form

$$\bar{f}'_{\varepsilon}(x) \triangleq \frac{f(x+\varepsilon) - f(x)}{\varepsilon}. \tag{3.1}$$

When used in the Cubic-Secant Algorithm, such approximations also affect the estimates of the second derivative. Nevertheless, as we will now show, provided one uses a proper rule for selecting $\varepsilon$, one can preserve the quadratic $R$-rate of convergence of the Cubic-Secant Algorithm.

Let $\bar{p}_{i,i-1}(\cdot)$ be a cubic polynomial whose coefficients are determined by the system of linear equations (2.7), with $f'(\cdot)$ replaced by $\bar{f}'_{\varepsilon_i}(\cdot)$. The explicit expression for $\bar{p}''_{i,i-1}(x)$ is given by (2.8a,b,c), with $f'(\cdot)$ replaced by $\bar{f}'_{\varepsilon_i}(\cdot)$, i.e.,

$$\bar{p}_{i,i-1}''(x) = \frac{2\bar{c}_i}{x_i - x_{i-1}} + \frac{2(x - x_{i-1}) + 2(2x - x_i - x_{i-1})}{(x_i - x_{i-1})^2} \bar{d}_i, \tag{3.2a}$$

where

$$\bar{c}_i = \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}} - \bar{f}'_{\varepsilon_i}(x_{i-1}), \tag{3.2b}$$

$$\bar{d}_i = \bar{f}'_{\varepsilon_i}(x_i) - 2\frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}} + \bar{f}'_{\varepsilon_i}(x_{i-1}) . \tag{3.2c}$$

Replacing $p_{i,i-1}''(x_i)$ by $\bar{p}_{i,i-1}''(x_i)$ in the Cubic-Secant Algorithm and adding a rule for adjusting the precision of the finite difference approximation parameter $\varepsilon$, we obtain the following "discrete" version of Algorithm 2.1.

**Discrete Cubic-Secant Algorithm 3.1.**

*Parameters.* $\alpha \in (0, 1/2)$, $\beta \in (0,1)$, $m \in (0,1)$, small, $\theta \in (0,1)$, and $\varepsilon_0$, small.

*Data.* $x_0, x_{-1} \in [0, \infty)$.

*Step 0.* Set $i = 0$.

*Step 1.* Set $\varepsilon = \min\{ \varepsilon_{i-1}, |x_i - x_{i-1}|^2, \theta^i \}$.

*Step 2.* Compute $\bar{f}'_{\varepsilon_i}(x_i)$.

*Step 3.* If $\varepsilon > |\bar{f}'_{\varepsilon_i}(x_i)|^{2.2}$, set $\varepsilon = ½\varepsilon_i$ and go to step 2. Else, set $\varepsilon_i = \varepsilon$.

*Step 4.* Compute $\bar{p}''_{i,i-1}(x_i)$.

If $\bar{p}''_{i,i-1}(x_i) \geq m$, set

$$h_i = -\bar{p}''_{i,i-1}(x_i)^{-1}\bar{f}'_{\varepsilon_i}(x_i) . \tag{3.3}$$

Else, set

$$h_i = -\bar{f}'_{\varepsilon_i}(x_i) . \tag{3.4}$$

*Step 5.* Compute the *Armijo step size*

$$\lambda_i = \max_{k \in \mathbb{N}} \left\{ \beta^k \mid f(x_i + \beta^k h_i) - f(x_i) \leq \alpha\beta^k h_i \bar{f}'_{\varepsilon_i}(x_i) \right\} . \tag{3.5}$$

*Step 6.* Set $x_{i+1} = x_i + \lambda_i h_i$, replace $i$ by $i+1$ and go to Step 1. □

**Lemma 3.2.** Suppose that Assumption 2.4 is satisfied and that there exists an $M > 1$ such that $|f''(x)| \leq M$, for all $x \in \mathbb{R}$. If $\{x_i\}_{i=0}^{\infty}$, $\{\varepsilon_i\}_{i=0}^{\infty}$, are infinite sequences constructed by the Discrete Cubic-Secant Algorithm 3.1, then there exists an $i_0$ such that for all $i \in \mathbb{N}$, $i \geq i_0$,

$$\varepsilon_i \leq \frac{2}{3M} |f'(x_i)| . \tag{3.6}$$

*Proof.* First, since $\theta \in (0,1)$, there exists $i_0 \in \mathbb{N}$ such that for all $i \in \mathbb{N}$, $i \geq i_0$,

$$\theta^i \le \left[ \frac{1}{2M} \right]^{\frac{2.2}{1.2}} . \tag{3.7}$$

Let $i \ge i_0$. First suppose that

$$|\bar{f}'_{\varepsilon_i}(x_i)| \le \left[ \frac{1}{2M} \right]^{\frac{1}{1.2}} . \tag{3.8}$$

Since by construction $\varepsilon_i \le |\bar{f}'_{\varepsilon_i}(x_i)|^{2.2}$, (3.8) implies that

$$\varepsilon_i \le |\bar{f}'_{\varepsilon_i}(x_i)|^{1.2} |\bar{f}'_{\varepsilon_i}(x_i)| \le \frac{1}{2M} |\bar{f}'_{\varepsilon_i}(x_i)| . \tag{3.9}$$

On the other hand, if (3.8) does not hold, then again we get that

$$\varepsilon_i \le \theta^i \le \left[ \frac{1}{2M} \right]^{\frac{2.2}{1.2}} = \frac{1}{2M} \left[ \frac{1}{2M} \right]^{\frac{1}{1.2}} \le \frac{1}{2M} |\bar{f}'_{\varepsilon_i}(x_i)| . \tag{3.10}$$

From (3.9) and (3.10) we conclude that for all $i \ge i_0$,

$$\varepsilon_i \le \frac{1}{2M} |\bar{f}'_{\varepsilon_i}(x_i)| . \tag{3.11}$$

Expanding of $f(\cdot)$ around $x_i$ to second order, we obtain that for some $s \in [0,1]$

$$f(x_i + \varepsilon_i) = f(x_i) + \varepsilon_i f'(x_i) + \frac{\varepsilon_i^2}{2} f''(x_i + s\,\varepsilon_i), \tag{3.12}$$

which implies that

$$\bar{f}'_{\varepsilon_i}(x_i) \triangleq \frac{f(x_i + \varepsilon_i) - f(x_i)}{\varepsilon_i} = f'(x_i) + \frac{\varepsilon_i}{2} f''(x_i + s\,\varepsilon_i) . \tag{3.13}$$

Therefore,

$$|f'(x_i)| \ge |\bar{f}'_{\varepsilon_i}(x_i)| - \frac{M}{2}\varepsilon_i . \tag{3.14}$$

The desired result then follows from (3.11) and (3.14). □

**Corollary 3.3.** If the assumptions of Lemma 3.2 are satisfied, then there exists an $i_0 \in \mathbb{N}$ such that for all $i \in \mathbb{N}$, $i \ge i_0$,

$$\frac{2}{3} |f'(x_i)| \le |\bar{f}'_{\varepsilon_i}(x_i)| \le \frac{4}{3} |f'(x_i)| . \tag{3.15}$$

□

**Theorem 3.4.** Suppose that Assumption 2.4 is satisfied and that there exists an $M > 1$ is such that

$|f''(x)| \le M$ for all $x \in \mathbb{R}$. If the Discrete Cubic-Secant Algorithm 3.1 has constructed the infinite sequence $\{x_i\}_{i=0}^{\infty}$, then every accumulation point $\hat{x}$ of $\{x_i\}_{i=0}^{\infty}$ satisfies $f'(\hat{x}) = 0$.

*Proof.* Let $\{\varepsilon_i\}_{i=0}^{\infty}$ be the corresponding sequence constructed by Discrete Cubic-Secant Algorithm 3.1. First, it follows from (3.13) that for some $s \in [0,1]$,

$$f'(x_i)\bar{f}'(x_i) = |f'(x_i)|^2 + \frac{\varepsilon_i}{2} f''(x_i + s\varepsilon_i) f'(x_i), \tag{3.16}$$

which implies that

$$f'(x_i)\bar{f}'_{\varepsilon_i}(x_i) \ge |f'(x_i)|^2 - \frac{M}{2}\varepsilon_i |f'(x_i)| . \tag{3.17}$$

Therefore, by Lemma 3.2, there exists an $i_0 \in \mathbb{N}$ such that for all $i \in \mathbb{N}$, $i \ge i_0$,

$$f'(x_i)\bar{f}'_{\varepsilon_i}(x_i) \ge \frac{2}{3} |f'(x_i)|^2 . \tag{3.18}$$

In view of (3.3) and (3.4), we conclude that for all $i \ge i_0$

$$f'(x_i)h_i \le -\frac{2}{3}\min\{1, \frac{1}{m}\} |f'(x_i)|^2 . \tag{3.19}$$

Next, from the second inequality in (3.15), we conclude that for all $i \ge i_0$,

$$|h_i| \le \frac{4}{3}\max\{1, \frac{1}{m}\} |f'(x_i)| . \tag{3.20}$$

In view of (3.19) and (3.20), the desired result follows directly from the Polak-Sargent-Sebastian Theorem [12]. □

The proof of the following result is identical to that of Corollary 2.5 and is therefore omitted.

**Corollary 3.5.** Suppose that the assumptions of Theorem 3.4 are satisfied. If the Discrete Cubic-Secant Algorithm 3.1 has constructed the infinite sequence $\{x_i\}_{i=0}^{\infty}$ that has an accumulation point $\hat{x}$ such that $f''(\hat{x}) \ge b$ for some $b > 0$, then $x_i \to \hat{x}$ as $i \to \infty$. □

**Corollary 3.6.** Suppose that the assumptions of Theorem 3.4 are satisfied. If the Discrete Cubic-Secant Algorithm 3.1 has constructed the infinite sequences $\{x_i\}_{i=0}^{\infty}$ and $\{\varepsilon_{i_i}\}_{i=0}^{\infty}$, with $x_i \to \hat{x}$ as $i \to \infty$, then there exists an $i_0 \in \mathbb{N}$ such that for all $i \in \mathbb{N}$, $i \ge i_0$,

$$\varepsilon_i \le |x_i - \hat{x}|^2 . \tag{3.21}$$

*Proof.* Since $\theta \in (0,1)$ and $M > 1$, there exists an $i_0$ such that for all $i \in \mathbb{N}$, $i \ge i_0$,

$$\theta^i \le \left[ \frac{9}{16M^2} \right]^{11} . \tag{3.22}$$

Let $i \ge i_0$. First suppose that

$$|\bar{f}'_{\varepsilon_i}(x_i)| \le \left[ \frac{9}{16M^2} \right]^5 . \tag{3.23}$$

In this case, because by construction $\varepsilon_i \le |\bar{f}'_{\varepsilon_i}(x_i)|^{2.2}$ for all $i \in \mathbb{N}$,

$$\varepsilon_i \le |\bar{f}'_{\varepsilon_i}(x_i)|^{2.2} = |\bar{f}'(x_i)|^{0.2} |\bar{f}'(x_i)|^2 \le \frac{9}{16M^2} |\bar{f}'_{\varepsilon_i}(x_i)|^2 \le \frac{|f'(x_i)|^2}{M^2} , \tag{3.24}$$

where (3.23) and (3.15) were used to establish the third and fourth inequalities respectively.

On the other hand, if (3.23) does not hold, then

$$\varepsilon_i \le \theta^i \le \left[ \frac{9}{16M^2} \right]^{11} = \frac{9}{16M^2} \left[ \left[ \frac{9}{16M^2} \right]^5 \right]^2 \le \frac{9}{16M^2} |\bar{f}'_{\varepsilon_i}(x_i)|^2 \le \frac{|f'(x_i)|^2}{M^2} , \tag{3.25}$$

where (3.15) was used to establish the last inequality.

Since $f'(\hat{x}) = 0$, we have that for some $s \in [0,1]$

$$|f'(x_i)| = |(x_i - \hat{x}) f''(x_i + s(\hat{x} - x_i))| \le M |x_i - \hat{x}| \tag{3.26}$$

which together with (3.24) and (3.25) completes the proof of the Corollary. □

**Lemma 3.7.** Suppose that Assumption 2.4 is satisfied, that there exists an $M > 1$ such that $|f''(x)| \le M$ for all $x \in \mathbb{R}$, and that the Discrete Cubic-Secant Algorithm 3.1 has constructed the infinite sequences $\{ x_i \}_{i=0}^{\infty}$ and $\{ \varepsilon_{i_i} \}_{i=0}^{\infty}$. Then there exists a finite $K > 0$ such that for all $i$, $i-1 \in \mathbb{N}$,

$$|f''(x_i) - \bar{p}''_{i,i-1}(x_i)| \le K |x_i - x_{i-1}| . \tag{3.27}$$

*Proof.* Consider the expressions (2.8a,b,c) for $p''_{i,i-1}(\cdot)$, $c_i$, and $d_i$, as well as their counterparts (3.2a,b,c) for $\bar{p}_{i,i-1}''(\cdot)$, $\bar{c}_i$, and $\bar{d}_i$. Note that

$$p''_{i,i-1}(x_i) - \bar{p}''_{i,i-1}(x_i) = \frac{2}{x_i - x_{i-1}} \left[ (c_i - \bar{c}_i) + (d_i - \bar{d}_i) \right] , \tag{3.28}$$

and that for some $s \in [0,1]$

$$|(c_i - \bar{c}_i) + (d_i - \bar{d}_i)| = |f'(x_i) - \bar{f}'_{\varepsilon_i}(x_i)| = |\frac{\varepsilon_i}{2} f''(x_i + s\varepsilon_i)| \le M \frac{\varepsilon_i}{2} . \tag{3.29}$$

Since by construction $\varepsilon_i \le |x_i - x_{i-1}|^2$ for all $i \in \mathbb{N}$, we find that

$$|p''_{i,i-1}(x_i) - \bar{p}''_{i,i-1}(x_i)| = \frac{2}{|x_i - x_{i-1}|} |f'(x_i) - \bar{f}'_{\varepsilon_i}(x_i)| \le \frac{\varepsilon_i M}{|x_i - x_{i-1}|} \le M\,|x_i - x_{i-1}| \quad (3.30)$$

The desired result now follows from Lemma 2.6. □

**Theorem 3.8.** Suppose that Assumption 2.4 is satisfied and that there exists an $M > 1$ such that $f''(x) \le M$ for all $x \in \mathbb{R}$. If the Discrete Cubic-Secant Algorithm 3.1 has constructed a sequence $\{x_i\}_{i=0}^{\infty}$, with an accumulation point $\hat{x}$ such that $f''(\hat{x}) \ge 2m$, with $m \in (0,1)$ as in the Discrete Cubic-Secant Algorithm 3.1, then $x_i \to \hat{x}$, $R$-superlinearly, with root rate $\tau_1 \triangleq ½(1+\sqrt{5}) \approx 1.6180$.

*Proof.* It follows from Corollary 3.4 and Lemma 3.6, that $x_i \to \hat{x}$ as $i \to \infty$, and that there exists an $i_0 \in \mathbb{N}$ such that $\bar{p}''_{i,i-1}(x_i) \ge m$ for all $i \ge i_0$. Therefore, for all $i \ge i_0$, $h_i$ is given by (3.3).

Now, by the second order Taylor expansion formula, for any $i$, there exists an $s_i \in [0,1]$ such that

$$f(x_i + h_i) - f(x_i) - \alpha h_i \bar{f}'_{\varepsilon_i}(x_i) = h_i f'(x_i) + ½ f''(x_i + s_i h_i) h_i^2 + \alpha h_i^2 \bar{p}''_{i,i-1}(x_i), \quad (3.31)$$

where (3.3) was used to express $\bar{f}'_{\varepsilon_i}(x_i)$ in terms of $h_i$ and $\bar{p}''_{i,i-1}(x_i)$. Adding and subtracting $½(\bar{p}''_{i,i-1}(x_i) + f''(x_i))h_i^2$ to the right-hand side of (3.31), we get, using (2.15) and (3.27), that for some finite $K, L > 0$, and some $i_1 \in \mathbb{N}$, $i_1 \ge i_0$, for all $i \ge i_1$,

$$f(x_i + h_i) - f(x_i) - \alpha h_{i} \bar{f}'_{\varepsilon_i}(x_i) \le h_i f'(x_i) + (\alpha + \frac{1}{2}) \bar{p}''_{i,i-1}(x_i) h_i^2 + \frac{L}{3} h_i^3 + \frac{K}{2} h_i^2 |x_i - x_{i-1}| \; . (3.32)$$

Next, making use of (3.13), we obtain from (3.32) that for some $s_i \in [0,1]$,

$$f(x_i + h_i) - f(x_i) - \alpha h_i \bar{f}'_{\varepsilon_i}(x_i) \le \left[ \bar{f}'_{\varepsilon_i}(x_i) - \frac{\varepsilon_i}{2} f''(x_i + s_i \varepsilon_i) \right] h_i$$

$$+ (\alpha + ½) \bar{p}''_{i,i-1}(x_i) h_i^2 + \frac{L}{3} h_i^3 + \frac{K}{2} h_i^2 |x_i - x_{i-1}| \; . \quad (3.33)$$

Since by construction $\varepsilon_i \le |\bar{f}'(x_i)|^{2.2}$, for all $i \in \mathbb{N}$, we obtain, using again (3.3) to express $\bar{f}'_{\varepsilon_i}(x_i)$ in terms of $\bar{p}''_{i,i-1}(x_i)$ and $h_i$, that for all $i \ge i_1$

$$f(x_i + h_i) - f(x_i) - \alpha h_i \bar{f}'_{\varepsilon_i}(x_i) \le h_i^2 \Bigg[ -(1 - \alpha - \frac{1}{2}) m + \frac{L}{2} h_i + \frac{K}{2} |x_i - x_{i-1}|$$

$$- \frac{\varepsilon_i}{2} f''(x_i + s_i \varepsilon_i) |h_i|^{0.2} |\bar{p}''_{i,i-1}(x_i)|^{2.2} h_i \Bigg] \; . \quad (3.34)$$

Since both $x_i - x_{i-1} \to 0$ and $h_i \to 0$ as $i \to \infty$, there exists an $i_2$, $i_2 \ge i_1$, such that

$f(x_i + h_i) - f(x_i) \le \alpha h_i \bar{f}'_{\varepsilon_i}(x_i)$, which implies that $\lambda_i = 1$, for all $i \ge i_2$. Therefore for all $i \ge i_2$, there exist an $s_i \in [0,1]$ and an $\eta_i \in [x_i, x_i + \varepsilon_i]$ such that

$$\bar{p}''_{i,i-1}(x_i)(x_{i+1} - x_i) = -\bar{f}'_{\varepsilon_i}(x_i) = -\frac{f(x_i + \varepsilon_i) - f(x_i)}{\varepsilon_i} = -f'(x_i) - f''(\eta_i)\frac{\varepsilon_i}{2}$$

$$= -\tfrac{1}{2} f''(x_i + s_i(x_i - \hat{x}))(x_i - \hat{x}) - \tfrac{1}{2} f''(\eta_i)\varepsilon_i \,. \tag{3.35}$$

Rearranging terms in (3.35), we get

$$\bar{p}''_{i,i-1}(x_i)(x_{i+1} - \hat{x}) = -\tfrac{1}{2}(f''(x_i + s_i(x_i - \hat{x})) - \bar{p}''_{i,i-1}(x_i))(x_i - \hat{x}) - \tfrac{1}{2} f''(\eta_i)\varepsilon_i \,. \tag{3.36}$$

Adding and subtracting $\tfrac{1}{2} f''(x_i)$ from the the right-hand side of (3.36), and making use of Corollary 3.6, (2.15) and (3.27), we conclude that there exists an $i_3 \in \mathbb{N}$, $i_3 \ge i_2$, such that for all $i \ge i_3$,

$$|x_{i+1} - \hat{x}| \le \frac{L}{3m}|x_i - \hat{x}|^2 + \frac{K}{2m}|x_i - x_{i-1}|\,|x_i - \hat{x}| + \frac{M}{2m}|x_i - \hat{x}|^2 \,. \tag{3.37}$$

Let

$$C \triangleq \frac{1}{m}\left\{ \frac{L}{3} + \frac{K}{2} + \frac{M}{2} \right\} . \tag{3.38}$$

Then, if we rewrite $x_i - x_{i-1}$ as $(x_i - \hat{x}) - (x_{i-1} - \hat{x})$, we obtain from (3.37) that

$$|x_{i+1} - \hat{x}| \le C \left\{ |x_i - \hat{x}|^2 + |x_i - \hat{x}|\,|x_{i-1} - \hat{x}| \right\} \tag{3.39}$$

For all $i \in \mathbb{N}$, let

$$e_i \triangleq 2C\,|x_i - \hat{x}| \,. \tag{3.40}$$

Then we get from (3.39) that

$$e_{i+1} \le \tfrac{1}{2}\{ e_i^2 + e_i e_{i-1} \} . \tag{3.41}$$

The rest of the proof follows by induction. Since $x_i \to \hat{x}$, there exists an $i_4 \in \mathbb{N}$, $i_4 \ge i_3$, such that $e_{i_4+1} < e_{i_4} < 1$. Let $\delta \in (0,1)$ be defined by $\delta^{\tau_1} = \frac{1}{2}\max\{ e_{i_4}^{\tau_1} + e_{i_4},\ e_{i_4+1} + e_{i_4} \}$. Then $e_{i_4} \le \delta$ and $e_{i_4+1} \le \delta^{\tau_1}$.

Suppose now that for $k = i_4, i_4 + 1, \ldots, i_4 + n$, we have $e_k \le \delta^{\tau_1^k}$. Then, from (3.41), we get that

$$e_{i_4+n+1} \le \tfrac{1}{2}(\delta^{2\tau_1^n} + \delta^{\tau_1^n}\delta^{\tau_1^{n-1}}) = \tfrac{1}{2}(\delta^{\tau_1^{n+1}}\delta^{(2-\tau_1)\tau_1^n} + \delta^{\tau_1^n(1+\tau_1^{-1})}) \,. \tag{3.42}$$

Since $1 + \tau_1^{-1} = \tau_1$, we conclude that

$$e_{i_4+n+1} \leq \delta^{\tau_1^{n+1}}, \tag{3.43}$$

which completes the inductive step and our proof.

□

From (3.39) and the fact that $x_i \to \hat{x}$ as $i \to \infty$ it is clear that - with $C$ as in (3.38) - there exists an $i_0 \in \mathbb{N}$ such that for all $i > i_0$,

$$|x_{i-1} - \hat{x}| \leq \frac{1}{4C} \tag{3.44}$$

and

$$|x_{i+1} - \hat{x}| \leq \frac{1}{2}|x_i - \hat{x}| \leq \frac{1}{2}|x_{i+1} - x_i| + \frac{1}{2}|x_{i+1} - \hat{x}| \tag{3.45}$$

Thus we have the following corollary

**Corollary 3.9.** If the assumptions of Theorem 3.8 are satisfied, then there exists an $i_0 \in \mathbb{N}$, such that for all $i \in \mathbb{N}$, $i \geq i_0$,

$$|x_i - \hat{x}| \leq |x_i - x_{i-1}| . \tag{3.46}$$

□

**Theorem 3.10** Suppose that Assumption 2.8 is satisfied and that there exists an $M > 1$ be such that $f''(x) \leq M$ for all $x \in \mathbb{R}$. If the Discrete Cubic-Secant Algorithm 3.1 has constructed a sequence $\{x_i\}_{i=0}^{\infty}$, with an accumulation point $\hat{x}$ such that $f''(\hat{x}) \geq 2m$, with $m \in (0,1)$ as in the Discrete Cubic-Secant Algorithm 3.1, Then $\{x_i\}_{i=0}^{\infty}$ converges to $\hat{x}$ quadratically.

*Proof.* Exactly as in the proof of Theorem 3.8, we obtain that $x_i \to \hat{x}$ as $i \to \infty$, and that there exists an $i_0$ such that for all $i \in \mathbb{N}$, $i \geq i_0$, we have that $\bar{p}''_{i,i-1}(x_i) \geq m$, $h_i$ is given by (3.3), and $\lambda_i = 1$. It also follows that there exists an $i_1 \in \mathbb{N}$ and $s_i \in [0,1]$ and $\eta_i \in [x_i, x_i + \varepsilon_i]$ such that for all $i \in \mathbb{N}$, $i \geq i_1$, (3.36) holds.

Adding and subtracting $\frac{1}{2}[f''(x_i) + p''_{i,i-1}(x_i)](x_i - \hat{x})$ from the right-hand side of (3.36) and using (2.15), we get that there exists an $i_2$, $i_2 \geq i_1$, such that for all $i \in \mathbb{N}$, $i \geq i_2$,

$$|\bar{p}''_{i,i-1}(x_i)|\,|x_{i+1} - \hat{x}| \leq \frac{L}{3}|x_i - \hat{x}|^2$$
$$+ \frac{1}{2}\Big[\,|f''(x_i) - p''_{i,i-1}(x_i)| + |\bar{p}''_{i,i-1}(x_i) - p''_{i,i-1}(x_i)|\,\Big]\,|x_i - \hat{x}| + \frac{M\varepsilon_i}{2} \tag{3.47}$$

Hence, making use of (2.23) and (3.30), we get that there exists a positive constant $B$ such that

$$|\bar{p}''_{i,i-1}(x_i)|\,|x_{i+1}-\hat{x}\,| \le \frac{L}{3}|x_i-\hat{x}\,|^2 + \left[ B\,|x_i-x_{i-1}|^2 + \frac{M\varepsilon_i}{|x_i-x_{i-1}|} \right] |x_i-\hat{x}\,| + \frac{M\varepsilon_i}{2} \tag{3.48}$$

which, together with Corollary 3.5, implies that there exists an $i_3 \in \mathbb{N}$, $i_3 \ge i_2$ such that for all $i \in \mathbb{N}$, $i \ge i_3$,

$$|x_{i+1}-\hat{x}\,| \le \left[ \frac{L}{3m} + \frac{M}{2m} \right] |x_i-\hat{x}\,|^2 + \left[ \frac{B}{m}|x_i-x_{i-1}|^2 + \frac{M}{m}\frac{|x_i-\hat{x}\,|^2}{|x_i-x_{i-1}|} \right] |x_i-\hat{x}\,| \,. \tag{3.49}$$

Making use of Corollary 3.8 together with (3.49), and rewriting $x_i - x_{i-1}$ as $(x_i-\hat{x}\,)-(x_{i-1}-\hat{x}\,)$, we conclude that for some finite constant $C \ge 1/4$, there is an $i_4$, $i_4 \ge i_3$, such that for all $i \in \mathbb{N}$, $i \ge i_4$,

$$|x_{i+1}-\hat{x}\,| \le C\Big[\, |x_i-\hat{x}\,|^3 + |x_i-\hat{x}\,|^2 + |x_i-\hat{x}\,|^2|x_{i-1}-\hat{x}\,| + |x_{i-1}-\hat{x}\,|^2|x_i-\hat{x}\,| \Big] \,. \tag{3.50}$$

Let

$$e_i \triangleq 4C\,|x_i-\hat{x}\,| \tag{3.51}$$

It then follows from (3.50) and from the fact that $4C \ge 1$ that

$$e_{i+1} \le \frac{1}{4}(e_i^3 + e_i^2 e_{i-1} + e_i e_{i-1}^2 + e_i^2)\,. \tag{3.52}$$

The proof can now completed by induction exactly as for Theorem 2.5. □

## 4. Numerical Results

We will now present two numerical examples which illustrate the performance of the Cubic-Secant Algorithm (CSA) and of the Discrete Cubic-Secant Algorithm (DSCA). The cost functions for for these examples were generated as line searches, i.e., functions of the form

$$f(x) = g(y + xh)\,, \tag{4.1}$$

with $g : \mathbb{R}^n \to \mathbb{R}$, where $g(\cdot)$ is a standard test function used for testing unconstrained optimization algorithms. We compare the performance of our algorithms with that of stabilized versions of Newton's Algorithm (SNA) [10] and of the Secant Algorithm (SSA) [5], and with the Quadratic Interpolation Algorithm (QIA) [7].

In all of our experiments the parameters for the Armijo step size rule were $\alpha = 0.3$, $\beta = 0.9$; the remaining parameters were $m = 0.0001$ and $\theta = 0.01$. We initialized the CSA, DCSA, SNA and SSA with $x_0 = 0$, and $x_{-1} = 0.01$. The Quadratic Interpolation Algorithm was initialized with the bracket

$\{0, 0.05, 0.35\}$ for the first test function, and with the bracket $\{0, 0.01, 0.14\}$ for the second. These brackets were obtained using a priori knowledge of the minimizers that they contain. Note that, in general, a considerable number of function evaluations might be necessary to obtain initial brackets that contain minimizers, as required by the Quadratic Interpolation Algorithm. Our test problems were as follows:

**Problem ERF** 4-Dimensional Extended Rosenbrock's Function [9].

In this problem, $g : \mathbb{R}^4 \to \mathbb{R}$ is given by:

$$g(y) = 100[(y_2 - y_1^2)^2 + (y_4 - y_3^2)^2] + (1 - y_1)^2 + (1 - y_3)^2$$

In (4.1), we set $y = (-1.2, 1, -1, 1)^T$, and $h = -\nabla g(y) = (1, 0.40816, 0.01855, 0)^T$, the direction of steepest descent at $y$.

**Problem TF** Trigonometric Function [13].

In this problem, $g : \mathbb{R}^3 \to \mathbb{R}$, is defined as follows:

$$g(y) = \sum_{i=1}^{3} \{ 3 + i - \sum_{j=1}^{n} (a_{ij} \sin y_j + b_{ij} \cos y_j) \}^2$$

$$a_{ij} = \delta_{ij}$$

$$b_{ij} = i\delta_{ij} + 1$$

In (4.1), we set $y = (1/3, 1/3, 1/3)^T$ and $v = -\nabla g(y) = (-0.296450, 0.705533, 1)^T$.

Tables 1 and 2, below, summarize our results in terms of number of function evaluations (NF), gradient evaluations (NG), and second derivative evaluations (NH) required to satisfy $\| x_i - \hat{x} \| \leq \varepsilon$, for various values of $\varepsilon$, for each of the two test problems considered.

Table 1 - Results for ERF

| $\varepsilon$ | CSA | | DCSA | | SSA | | SNA | | | QIA | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | NF | NG | NF | NG | NF | NG | NF | NG | NH | NF | NG |
| 1e-02 | 6 | 3 | 9 | 0 | 6 | 3 | 5 | 2 | 1 | 8 | 0 |
| 1e-04 | 10 | 4 | 19 | 0 | 14 | 5 | 9 | 3 | 2 | 20 | 0 |
| 1e-06 | 10 | 4 | 19 | 0 | 14 | 5 | 9 | 3 | 2 | 36 | 0 |
| 1e-08 | 14 | 5 | 19 | 0 | 18 | 6 | 13 | 4 | 3 | 48 | 0 |
| 1e-12 | 14 | 5 | FAILS | | FAILS | | 13 | 4 | 3 | FAILS | |

| ε | CSA | | DCSA | | SSA | | SNA | | | QIA | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | NF | NG | NF | NG | NF | NG | NF | NG | NH | NF | NG |
| 1e-02 | 19 | 4 | 13 | 0 | 9 | 3 | 9 | 2 | 1 | 4 | 0 |
| 1e-04 | 19 | 4 | 23 | 0 | 21 | 5 | 18 | 3 | 2 | 28 | 0 |
| 1e-06 | 23 | 5 | 28 | 0 | 25 | 6 | 22 | 4 | 3 | 44 | 0 |
| 1e-08 | 23 | 5 | 33 | 0 | 29 | 7 | 26 | 5 | 4 | 60 | 0 |
| 1e-12 | 27 | 6 | FAILS | | 33 | 8 | 26 | 5 | 4 | FAILS | |

Table 2 - Results for TF

As one might expect, in view of the different convergence rates, the number of evaluations required by our algorithms is smaller than those required by the Stabilized Secant Algorithm (SSA) or the Quadratic Interpolation Algorithm (QIA). Indeed, our computational results suggest that our algorithms are competitive with Newton's Algorithm.

In considering the numbers in our tables, one has to bear in mind that when $f(\cdot)$ is of the form (4.1), as in our examples, and the first and second derivatives of $f(\cdot)$ at $x_i$ are calculated by evaluating $\langle \nabla g(x_i), h \rangle$ and $\langle h, g_{yy}(x_i) h \rangle$ respectively, the computational work required for their evaluation is potentially equivalent to $n$ and $n^2$ function evaluations respectively. Moreover, the numbers for the Quadratic Interpolation Algorithm (QIA) do not include any evaluations that were used to determine brackets.

We also observe that the DCSA fails when the required precision is increased to $10^{-12}$. This is due to the inherent limitations of schemes that rely on finite-differences approximations.

Figures 1 and 2 illustrate the behavior of the five algorithms considered in the computational experiments. Once again, we observe that the behavior of the Cubic Secant Algorithm (CSA) and Newton's Algorithm (SNA) are quite similar.

## 5. Conclusion

We have presented two global, R-quadratically converging one-dimensional optimization algorithms for use as subprocedures for step length computation in various unconstrained optimization algorithms. They are more robust than the commonly used cubic interpolation method and faster than the commonly used quadratic interpolation method.

## 6. References

[1] L. ARMIJO, *Minimization of functions having continuous partial derivatives*, Pacific J. Math., 16 (1966), pp. 1-3.

[2] W. C. DAVIDON, *Variable Metric Methods for Minimization*, AEC Research and Development Rept ANL 5990 (Rev.), 1959.

[3] R. FLETCHER and M. J. D. POWELL, *A Rapidly Convergent Descent Method for Minimization*, Computer Journal, Vol. 6, (1963), pp. 163-168.

[4] R. FLETCHER and C. M. REEVES, *Function minimization by conjugate gradients*, Computer Journal, Vol. 7 (1964), pp 149-154.

[5] R. KLESSIG and E. POLAK, *Efficient Implementations of the Polak-Ribiere Conjugate Gradient Algorithm*, SIAM J. Control, Vol. 10, No. 3, pp. 524-549, 1972.

[6] J. KOWALIK and M. R. OSOBORNE, *Methods for Unconstrained Optimization Problems*, New York, American Elsevier Pub. Co., 1968.

[7] D. G. LUENBERGER, *Introduction to Linear and Nonlinear Systems*, Addison-Wesley, New York, 1983.

[8] J. M. ORTEGA and W. C. RHEINBOLDT, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[9] M. R. OSBORNE, *Some aspects of nonlinear least squares calculations*, in *Numerical Methods for Nonlinear Optimization*, F. A. Lootsma (Ed.), Academic Press, New York, 1972, pp. 171-189.

[10] E. POLAK, *Computational Methods in Optimization: a Unified Approach*, Academic Press, 1971.

[11] E. POLAK and G. RIBIERE, *Note sur la convergence de methodes de directions conjugees*, *Rev. Fr. Inform. Rech. Operation.*, 16-R1 (1969), pp.35-43.

[12] E. POLAK, R. W. H. SARGENT and D. J. SEBASTIAN, *On the Convergence of Sequential Minimization Algorithms*, JOTA, Vol. 14, No. 4, pp. 439-442, 1974.

[13] E. SPEDICATO, *Computational experience with quasi-Newton algorithms for minimization problems of moderately large size*, Rep. CISE-N-175, Segrate(Milano), 1975.

[14] A. TAMIR, *Line search techniques based on interpolating polynomials using function values only*, Management Sci., 5 (1976), pp.576-586.

[15] J. F. TRAUB, *Iterative Methods for the Solution of Equations*, Prentice Hall, Englewood Cliffs, 1964.

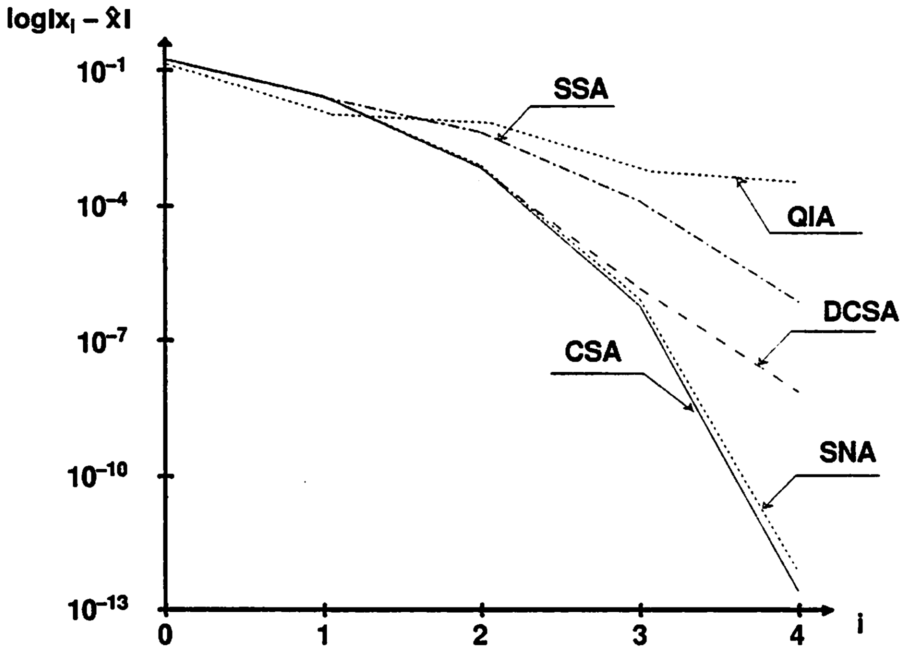


Fig.1. Performance of the Algorithms on Test Problem 1 - ERF

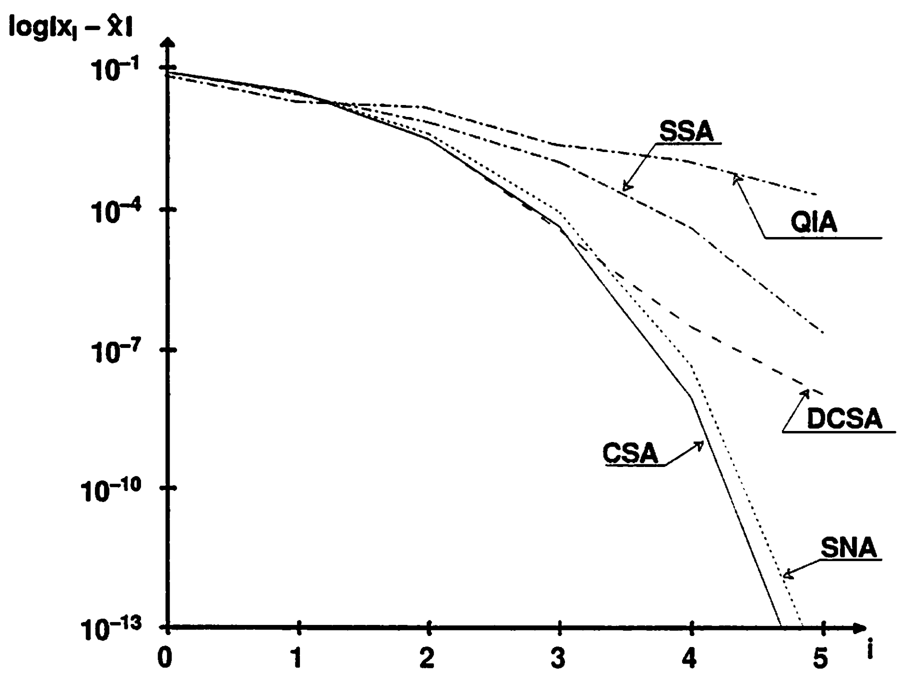


Fig.2. Performance of the Algorithms on Test Problem 2 - TF