J. KSIAM **Vol.5**, No.2, 101-115, 2001

NEW CONVERGENCE CONDITIONS OF SECANT METHODS VIA ALPHA THEORY

S. KIM

Abstract. Recent theoretical analysis of numerical methods for solving nonlinear systems of equations is represented by alpha theory of Newton method developed Smale et al. The theory was extended to Secant method by providing convergence conditions by Yakoubsohn which the Secant method is treated as an operator defined for analytical functions. We use Secant methods as an iterative scheme with approximations, which results in new convergence conditions. We compare the two conditions and show that the new conditions represent the features of Secant method in a more precise way.

1. Introduction

We consider solving nonlinear systems of algebraic equations

$$f(x) = 0, \tag{1}$$

where $f : E \to F$ with E and F two Euclidean spaces or more generally two Banach spaces. It is a classical problem with applications in many branches of engineering. Computationally popular numerical methods for (1) are Newton and Secant methods. The theoretical aspects of the methods have been dealt recently with Alpha theory [1]. Alpha theory for Newton method on solving (1) was developed by Smale *et al* and Yakoubsohn extended the theory to Secant type method in [5, 6].

The conventional convergence analysis of Newton method can be described as follows. Newton method defined as

$$N_f(x) = x - (f'(x))^{-1} f(x),$$

is considered to be an interation based on the map from R to itself, where $f'(x)$ is the derivative of f at x. The convergence rate of Newton method is q-quadratic convergence [2, 4], i.e., Newton method converges quadratically if an initial guess x_0 is in an open set $N(r, \epsilon)$ about a root r and there exists $r \in R^n,\ \gamma,\ \beta > 0$, such that $N(r,\gamma) \subset D$, with D domain, $f(r) = 0$, $f'(r)^{-1}$ exists with $\| f'(r)^{-1} \| < \beta$, and f' Lipschitz continuous in a region containing x_0, where f is assumed to be continuously differentiable.

AMS subject classification 65H10,90C53
Key Words: Alpha theory, Secant Method, Newton Method, Nonlinear Systems of Equations.
Research supported by Kosef R004-2001-00200

On the other hand, in Alpha theory,

$$\beta(f,x) = |f'(x)^{-1}f(x)|, \quad \alpha(f,x) = \beta(f,x) \sup_{k\geq 2} \left| \frac{f'(x)^{-1}f^{(k)}(x)}{k!} \right|^{1/(k-1)} < \alpha_0,$$

and

$$\gamma(f,x) = \sup_{k\geq 2} \left| \frac{f'(x)^{-1}f^{(k)}(x)}{k!} \right|^{1/(k-1)}$$

were introduced to verify the convergence of Newton method on a starting point where f is an analytic system of equations. They computed a ball centered at a root ξ of f that contains only approximate zeros, showed that N_f is a contraction map in a ball, and evaluated a neighborhood for a point x in the ball of contraction. The value of the universal constant *alpha* for Newton method is approximately 0.16. Alpha theory gives the size of the basin of attraction around the zeros in terms of the invariant $\gamma(f,x)$. It is a strong result in the sense that the convergence does not depend on initial guesses, but $\alpha(f,x)$ a number characterized by f and x.

Similar approach for theoretical analysis of Secant methods was made using Alpha theory in [6]. Basically, Secant method is an iterative scheme as Newton method with two given initial guesses for a root. The method reduces cost of computation $f'(x)$ by approximating, which can be very expensive operation. The q-superlinear convergence analysis of Secant method proved by [2] assumes a weak regularity for f. Yakoubsohn in [6] assumes that $f : E \to F$ is an analytic function and showed that Secant mapping defined as

$$S_f(y,x) = y - ([y,x]f)^{-1}f(y) = x - ([y,x]f)^{-1}f(x) \tag{2}$$

where

$$[y,x]f = \sum_{k\geq 1} \frac{D^k f(x)}{k!}(y-x)^{k-1}$$

is a contraction map and S_f maps $B(x_0, \frac{u_0}{\gamma(f,x_0)})$ into $B(\xi, \frac{u_0}{\gamma(f,\xi)})$ with a contraction constant less than or equal to $1/2$ under the given conditions.

The aim of this paper is to analyze the conditions for Secant method given by Yakoubsohn and give a new condition for contraction and robust alpha theorem. The new condition is derived from the fact that Secant method uses an approximation to the Jacobian of Newton method in practice, which was not considered in [6]. Wide use of Secant method also comes from its effectiveness as Newton method in obtaining a solution. In particular, the distance between x and y is very small, i.e., $\| x - y \| = \epsilon$,

$$[y,x]f = Df(x) + \boldsymbol{O}(\epsilon),$$

then the iterates produced from Secant map is nearly same as those of Newton method given the same initial point. In this case, a similar value of a universal constant to Newton method can be expected if a universal constant for Secant method is found and a contraction constant with which Secant map is a contractive map to be determined near that of Newton method.

The values of u that satisfies the conditions are in the range of $0 \leq u \leq 0.08$ approximately with varying α from 0 approximately to 0.008. Larger values of α does not satisfy the conditions. This range of values of u and α are very small compared to Newton method. This computation of α does not show the property of Secant method well in the sense that an approximation is used in the iterative framework and it can be very accurate as we choose ϵ very small. Therefore, it provides us room to improve the universal constant and the contraction constant of Secant method in [6] to represent its theoretical properties.

In Section 2, we define the approximate zero for Secant iteration as Newton and show convergence of Secant method using a constant. Section 3 includes Newton method's contraction theorem to observe the valid range of α, which provides a basis for Secant method's contraction theorem. In section 4, we describe Yakoubsohn's results and discuss values with figures. The conditions are examined and a new condition is derived. We show Yakoubshon's condition and the proposed condition with Figures and compare. Section 4 is devoted to concluding discussion.

2. Basic Analysis

The fact that Secant method involves less amount of work in computation of f' gives an edge over Newton method. The availability of f' and the degree of difficulty of computing f' determine work involved in each iteration. However, Secant method generally has a slower convergence rate than Newton method. The constant α indicates the rate of convergence of Secant method and the convergence of Secant method is

$$|x_n - r| \leq K|x_{n-1} - r|^{\alpha} \tag{3}$$

for some constant K [2].

We use the techniques of [1] to derive a convergence region on a given input for Secant method. But the main difference is the speed of convergence reflected on the constant K. To observe the difference clearly the definition of approximate zeros in Newton method for two initial guesses is given as follows.

Definition 2.1. Let $\lambda_n = \lambda_{n-1} + \lambda_{n-1}$, $\lambda_0 = \lambda_1 = 1$ and $\lambda = \frac{1+\sqrt{5}}{2}$. If, for given x_0 and x_1, $x_{i+1} = S_f(x_i, x_{i-1})$ is defined for $i \geq 2$ and there is a x_* such that $f(x_*) = 0$ with

$$|x_i - x_*| \leq \left(\frac{1}{2}\right)^{\lambda_i - 1} |x_0 - x_*|, \tag{4}$$

then, x_0 is an approximate zero of f and x_* is called as the associated zero.

Let

$$\gamma = \gamma(f, x) = \sup_{k \geq 2} \left| \frac{f'(x)^{-1} f^{(k)}(x)}{k!} \right|^{1/(k-1)},$$

$$u = \gamma(f,x)|\omega - x|,$$

and

$$E = \sum_{k=2}^{\infty} k \frac{f'(x)^{-1} f^{(k)}(x)(\omega - x)^{(k-1)}}{k!}.$$

The concept of approximate zeros and the convergence criteria by Smale is based on the series E to converge with the radius of convergence $u \leq 1 - \frac{\sqrt{2}}{2}$.

Lemma 2.2. [1] *If* $u < 1 - \frac{\sqrt{2}}{2}$, *then*
(a)

$$f'(x)^{-1} f'(\omega) = 1 + E$$

where $|E| \leq \frac{1}{(1-u)^2} - 1 < 1$.
(b)

$$|f'(\omega)^{-1} f'(x)| \leq \frac{(1-u)^2}{1 - 4u + 2u^2}.$$

Lemma 2.3. *For some* $\eta_n \in (x_{n-1}, x_n)$ *and* $u_n = \gamma(f,r)(\eta_n - r)$,

$$|f'(r)^{-1} f''(\eta_n)| \leq \gamma(f,r) \frac{1}{(1-u_n)^3}. \tag{5}$$

Proof.

$$\begin{aligned}
f''(\eta_n) &= f''(r) + f^{(3)}(r)(\eta_n - r) + \sum_{k=4}^{\infty} \frac{f^{(k)}(r)(\eta_n - r)^{(k-2)}}{(k-2)!}(\eta_n - r)^{(k-2)} \\
f'(r)^{-1} f''(\eta_n) &= \sum_{k=2}^{\infty} \frac{f'(r)^{-1} f^{(k)}(r)}{(k-2)!}(\eta_n - r)^{(k-2)} \\
|f'(r)^{-1} f''(\eta_n)| &\leq \sum_{k=2}^{\infty} k(k-1) u_n^{(k-2)} \\
&= \frac{1}{(1-u_n)^3}.
\end{aligned}$$

Lemma 2.4. *Let* $f(r) = 0$, *and* $u_n = \max(\gamma(f,r)|\eta_n - r|,\ \gamma(f,r)|\zeta_n - r|)$, *for some* $\eta_n, \zeta_n \in (x_n, x_{n-1})$, *and* $\Psi(u_n) = 1 - 4u_n + 2u_n^2$. *We also let*

$$c_n = \frac{\gamma}{2\Psi(u_n)(1-u_n)}.$$

Then,

$$\|S_f(x_n, x_{n-1}) - r\| < c_n \|x_n - r\| \|x_{n-1} - r\|$$

$$\|S_f^k(x_n, x_{n-1}) - r\| < c_n^{\lambda - 1} \|x_n - r\|^{\lambda}. \tag{6}$$

Proof. From Secant iteration (2),

$$\begin{aligned} S_f(x_n, x_{n-1}) - r &= x_n - \left([f(x_n) - f(x_{n-1})]^{-1}(x_n - x_{n-1})\right) f(x_n) - r \\ &= [f(x_n) - f(x_{n-1})]^{-1}[f(x_n)(x_{n-1} - r) - f(x_{n-1})(x_n - r)] \\ &= (f(x_n) - f(x_{n-1}))^{-1}(x_n - x_{n-1})(x_n - x_{n-1})^{-1} \\ &\quad [f(x_n)(x_n - r)^{-1} - f(x_{n-1})(x_{n-1} - r)^{-1}](x_n - r)(x_{n-1} - r) \\ &= \frac{1}{2} f'(\zeta_n)^{-1} f''(\eta_n)(x_n - r)(x_{n-1} - r), \end{aligned} \tag{7}$$

for some $\zeta_n, \eta_n \in (x_n, x_{n-1})$. Now,

$$\begin{aligned} |S_f(x_n, x_{n-1}) - r| &= \frac{1}{2} \left| f'(\zeta_n)^{-1} f''(\eta_n)(x_n - r)(x_{n-1} - r) \right| \\ &= \frac{1}{2} \left| f'(\zeta_n)^{-1} f'(r) f'(r)^{-1} f''(\eta_n)(x_n - r)(x_{n-1} - r) \right| \\ &\leq \frac{1}{2} \left| f'(\zeta_n)^{-1} f'(r) \right| \left| f'(r)^{-1} f''(\eta_n) \right| \left| (x_n - r)(x_{n-1} - r) \right|. \end{aligned}$$

$$|S_f(x_n, x_{n-1}) - r| \leq \frac{1}{2} \left| f'(\zeta_n)^{-1} f'(r) \right| \left(\frac{1}{(1 - u_n)^3} \right) |(x_n - r)(x_{n-1} - r)|$$

By Lemma 2.2 and 2.3,

$$\begin{aligned} |S_f(x_n, x_{n-1}) - r| &\leq \frac{1}{2} \sum_{k=0}^{\infty} |E|^k \left(\frac{1}{(1 - u_n)^3} \right) |(x_n - r)(x_{n-1} - r)| \\ &\leq \frac{1}{2\Psi(u_n)(1 - u_n)} |(x_n - r)(x_{n-1} - r)| \end{aligned} \tag{8}$$

We let for $k = 1, \cdots,\ S_f^k(x_1, x_0) = S_f(x_k, x_{k-1})$.

$$S_f(x_1, x_0) \leq c_0 |x_1 - r||x_0 - r|.$$

where $u_0 = \gamma |x_0 - r|$. We assume that we are given initial guesses x_0, x_1 such that $|x_1 - r| < |x_0 - r|$. Let $\lambda_n = \lambda_{n-1} + \lambda_{n-2}$, $\lambda_0 = \lambda_1 = 1$ and $\lambda = \frac{1+\sqrt{5}}{2}$. Then,

$$|S_f^2(x_0, x_1) - r| \leq c_0^2 |x_0 - r|^3,$$

$$|S_f^k(x_0, x_1) - r| \leq c_0^{\lambda_{k+1} - 1} |x_0 - r|^{\lambda_{k+1}}.$$

From $|x_{k+1} - r| = |S_f^k(x_0, x_1) - r|$,

$$\begin{aligned} c_0 |x_{k+1} - r| &\leq |x_0 - r|^{\lambda_{k+1}} c_0^{\lambda_{k+1}}, \\ |x_{k+1} - r| &\leq \frac{1}{c_0} |x_0 - r|^{\lambda_{k+1}} c_0^{\lambda_{k+1}}, \\ &= \frac{d_0^{\lambda_{k+1}}}{c_0}, \end{aligned}$$

where $d_0 = c_0|x_0 - r|$. It follows that

$$\begin{aligned} |S_f^k(x_1, x_0) - r| &\leq c_k|x_k - r|^{\lambda}|x_k - r|^{(1-\lambda)}|x_{k-1} - r| \\ &\leq c_0|x_k - r|^{\lambda}\left(\frac{d_0^{\lambda_k}}{c_0}\right)^{1-\lambda}\left(\frac{d_0^{\lambda_{k-1}}}{c_0}\right) \\ &= |x_k - r|^{\lambda}(c_0)^{\lambda-1}d_0^{\lambda_{k+1}-\lambda\lambda_k}. \end{aligned}$$

Since $\lambda_{n+1} - \lambda\lambda_n \to 0$ as $n \to \infty$, we show that (6) follows from (8):

For $k = 0$, (6) is trivial. By induction, we assume for $k \geq 1$ that

$$|S_f^{k-1}(x_1, x_0) - r| \leq (c_0|x_0 - r|)^{\lambda_k - 1}|x_0 - r|.$$

Using (8), we have

$$\begin{aligned} |S_f^k(x_1, x_0) - r| &\leq c_0\left[(c_0|x_0 - r|)^{\lambda_k - 1}|x_0 - r|\right]\left[(c_0|x_0 - r|)^{\lambda_{k-1} - 1}|x_0 - r|\right] \\ &= (c_0|x_0 - r|)^{\lambda_{k+1} - 1}|x_0 - r|, \end{aligned}$$

which holds for x_n and x_{n-1}.

Theorem 2.5. *Let $u = \gamma(f, x)|x - r|$ and $\rho(u) = \frac{1}{2\Psi(u)(1-u)} - \frac{1}{2} = 0$. Suppose that $f(r) = 0$ and $f'(r)^{-1}$ exists. If*

$$|x - r| \leq \frac{1}{2\Psi(u)(1 - u)} \quad \textit{for}$$

and u is less than the small root y of $\rho(u)$, then x is an approximate zero of f with associate zero r.

Proof. We need to show that $c_n \leq \frac{1}{2}$ from the definition 2.1 of the approximate zero. From Lemma 2.4, if $u < y$, then $\frac{\gamma}{2\Psi(u)(1-u)} < \frac{1}{2}$.

3. Convergence of Newton Maps

In this section, we first describe Newton contraction theorem and robust α theorem. The theorems include conditions on u and α to satisfy. We show the conditions in the theorems with Figures to describe the valid range of the values. The range is used in the next section to be compared to that of Secant methods.

Theorem 3.1. *(N-Gamma theorem)* [1] *Suppose that $f(\xi) = 0$ and that $Df(\xi)^{-1}$ exists. If*

$$\| x - \xi \| \leq \frac{3 - \sqrt{7}}{2\gamma(f, \xi)},$$

then x is an approximate zero of f with associated zero ξ, i.e., the sequence

$$x_0 = x,\ x_{k+1} = N_f(x_k),\ k \geq 0$$

is well defined and satisfies

$$\| x_k - \xi \| \leq \left(\frac{1}{2}\right)^{2^k-1} \| x - \xi \|, \; k \geq 0.$$

Theorem 3.2. *(N-Contraction theorem)* [1] *Let $x \in E$ and $u > 0$ such that the two conditions hold:*
1.

$$c = \frac{2\alpha(f,x) + u}{\Psi(u)^2} < 1,$$

2.

$$\alpha(f,x) + cu \leq u.$$

Then N_f is a contraction map of the ball $B\left(s, \frac{u}{\gamma(f,x)}\right)$ into itself with contraction constant c. Hence there is a unique root ξ of f in $B\left(x, \frac{u}{\gamma(f,x)}\right)$ and for all $y \in B\left(x, \frac{u}{\gamma(f,x)}\right)$ tend to ξ under iteration of N_f.

The left graph of Fig 1. shows the graph of $\frac{2\alpha(f,x)+u}{\Psi(u)^2} - 1$ in the variables α and u, To satisfy the condition 1 in Theorem 3.2, $c - 1$ should be negative. Therefore, the range of values that hold the condition 1 is approximately $0 \leq u \leq 1$. The right graph of Fig 1. shows $\alpha(f,x) + cu - u$ in Theorem 3.2. The condition 2 requires $\alpha(f,x) + cu - u \leq 0$, therefore, approximately $0 \leq \alpha \leq 0.015$ satisfies the condition 2. Figure 2 indicates the range of the values for u satisfying the conditions 1 and 2 simultaneously, which is approximately $0 \leq u \leq 0.1$.

4. Secant Iteration

We describe Yakoubsohn's results and show the conditions in his contraction theorem with Figures.

Secant iteration is defined in [6] as

$$S_f(x_n, x_{n-1}) = x_{n-1} - A(x_n, x_{n-1})^{-1} f(x_{n-1})$$

where

$$A(x_n, x_{n-1}) = \sum_{k\geq 1} \frac{Df^k(x_{n-1})}{k!}(x_n - x_{n-1})^{k-1},$$

and the secant operator A satisfies

$$f(x_n) - f(x_{n-1}) = A(x_n, x_{n-1})(x_n - x_{n-1}),$$

If A^{-1} is available, then

$$A^{-1}(x_n, x_{n-1})(f(x_n) - f(x_{n-1})) = x_n - x_{n-1},$$

It is basically the same map as (2). The following theorems are derived using the definition of the Secant map.

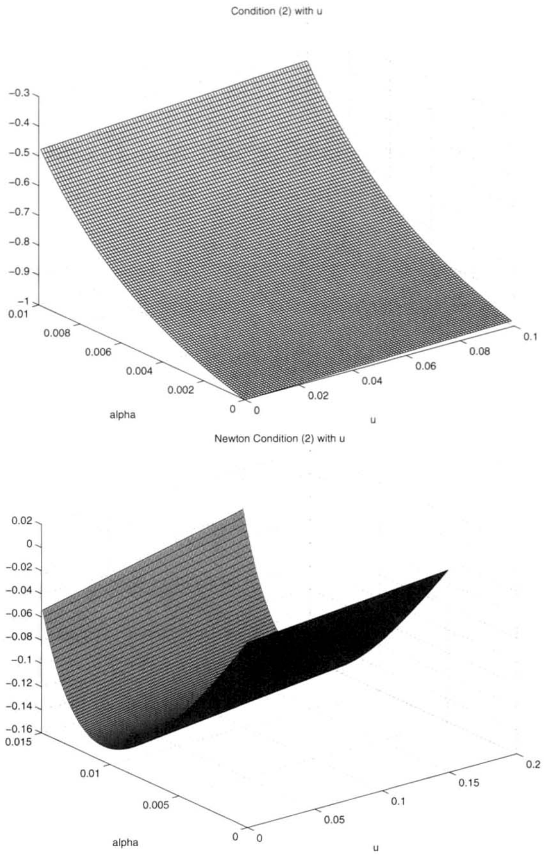

FIGURE 1. Newton Condition 1 and 2 with u

Theorem 4.1. *(S-Contraction) Let $x_0 \in E$ and $u_0 > 0$ such that*

$$\Psi(u_0)(1-u_0) - 4u_0 > 0.$$

Suppose
1.

$$c := \frac{(\alpha(f,x_0)+u_0)(1-u_0)^2}{(\Psi(u_0)(1-u_0)-2u_0)(\Psi(u_0)(1-u_0)-4u_0)} + \frac{u_0}{(\Psi(u_0)(1-u_0)-u_0)} < 1.$$

2.

$$\frac{(1-u_0)^2\alpha(f,x_0)+u_0^2(3-2u_0)}{\Psi(u_0)} \le u_0.$$

Then,
1. S_f maps $B\left(x_0, \frac{u_0}{\gamma(f,x_0)}\right)^2$ into $B\left(x_0, \frac{u_0}{\gamma(f,x_0)}\right)$.

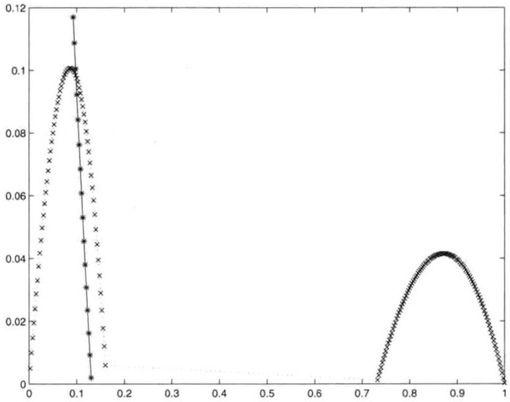

FIGURE 2. Newton Condition 1 and 2 with u

2. S_f is a contraction map with contraction constant c.
3. There is a unique root ξ of f such that

$$\| x_0 - \xi \| \leq \frac{u_0}{\gamma(f, x_0)}.$$

Theorem 4.2. *(S Robust α Theorem) Let u_0, α_0 be two positive numbers such that (a).*

$$c_0 := \frac{(\alpha_0 + u_0)(1-u_0)^2}{(\Psi(u_0)(1-u_0) - 2u_0)(\Psi(u_0)(1-u_0) - 4u_0)} + \frac{u_0}{(\Psi(u_0)(1-u_0) - u_0)} < \frac{1}{2}$$

(b).

$$\frac{(1-u_0)^2 \alpha(f, x_0) + u_0^2(3 - 2u_0)}{\Psi(u_0)} \leq u_0.$$

(c).

$$\left(u_0 + \frac{\alpha_0}{1-c_0}\right)\left(\frac{1}{\Psi(\frac{\alpha_0}{1-c_0})(1 - \frac{\alpha_0}{1-c_0})}\right) \leq \frac{3-\sqrt{7}}{2}.$$

(d).

$$\frac{1}{\Psi(\frac{\alpha_0}{1-c_0})(1 - \frac{\alpha_0}{1-c_0})} < \frac{1}{2c_0}.$$

If $\alpha(f, x_0) \leq \alpha_0$ then there is a root ξ such that

$$B\left(x_0, \frac{u_0}{\gamma(f(x_0))}\right) \subset B\left(x_0, \frac{3-\sqrt{7}}{2\gamma(f, \xi)}\right).$$

Moreover, S_f maps $B\left(x_0, \frac{u_0}{\gamma(f, x_0)}\right)$ into $B\left(\xi, \frac{u_0}{\gamma(f, \xi)}\right)$ with contraction constant less than or equal to 1/2.

We now describe a contraction theorem with a new contraction constance, different from Thoerem 4.2. The contraction constant shown here provides larger range of values than Theorem 4.2, which represents the features of Secant maps. Secant contraction theorem and S robust α algorithm in [6] and the assumptions used in the theorems are also shown in Figures. The ranges of values valid for the theorems are examined and compared.

We show that Secant method is a contraction mapping based on the following lemma.

Lemma 4.3. *([6] p.8.) Let $x, y, x_1, y_1 \in E$ and $u = \gamma(f,x) \parallel x - y \parallel, u_1 = \gamma(f,x) \parallel x - x_1 \parallel$, $v_1 = \gamma(f,x) \parallel y - y_1 \parallel, v = \gamma(f,x) \parallel x - y_1 \parallel$. Assume that u, u_1, v, and $v_1 < 1$. Then,*

$$1. \qquad \parallel Df(x)^{-1}([y_1,x_1]f - [y,x]f) \parallel \le \frac{u_1 + v_1 - u_1(v_1+u)}{(1-u_1)(1-v_1-u)(1-u)}$$

$$2. \qquad \parallel Df(x)^{-1}([y_1,x_1]f - Df(x)) \parallel \le \frac{u_1 + v - u_1 v}{(1-u_1)(1-v)}$$

Moreover if $2(1-u_1)(1-v) - 1 > 0$ *then*

$$3. \qquad \parallel [y_1,x_1]f^{-1} Df(x)) \parallel \le \frac{(1-u_1)(1-v)}{2(1-u_1)(1-v)-1}.$$

In view of Lemma 4.3, it follows

$$\begin{aligned}
\parallel ([y,x]f)^{-1} f(x) - ([z,x]f)^{-1} f(x) \parallel &\le \parallel ([z,x]f)^{-1} Df(x) \parallel\parallel Df(x)^{-1}([y,x] - [z,x]) f(x) \parallel \cdot \\
&\quad \parallel ([y,x]f)^{-1} Df(x) \parallel\parallel Df(x)^{-1} f(x) \parallel \\
&\le \frac{1}{(1-2u)^2} \beta(f,x)\gamma(f,x) \parallel z - y \parallel \\
&= \frac{1}{(1-2u)^2} \alpha(f,x) \parallel z - y \parallel .
\end{aligned}$$

From the definition of Secant map with $A(y,x) = [y,x]f$,

$$S(y,x) = y - A(y,x)^{-1} f(y), \tag{9}$$

where

$$A(y,x) = \sum_{k \ge 1} \frac{Df^k(x)}{k!} (y-x)^{k-1}.$$

Since $S(y,x)$ takes two variable x and y as input, $S(x,y)$ can be considered as a map from $2n$-dimensional space E to an n-dimensional space F. Hence, we consider $S(x,y)$ itself a function.

Notice that if we let $u = \gamma(f,x) \parallel y - x \parallel$ and $v = \gamma(f,y) \parallel x - y \parallel$, then,

$$\begin{aligned}
\parallel ([y,x]f)^{-1} f(x) \parallel &= \parallel ([y,x]f)^{-1} Df(x) \parallel\parallel Df(x)^{-1} f(x) \parallel \\
&\le \frac{\beta(f,x)}{(1-2u)^2}
\end{aligned} \tag{10}$$

and

$$\begin{aligned}\| ([x,y]f)^{-1}f(y) \| &= \| ([x,y]f)^{-1}Df(y) \| \| Df(y)^{-1}f(y) \| \\ &\leq \frac{\beta(f,y)}{(1-2v)^2}.\end{aligned}$$

We show $S(y,x)$ is a contractive map by differentiating $S(y,x)$ with respect to y and x.

Lemma 4.4. *Let $\beta_N(f,y) = \| Df(y)^{-1}f(y) \|$ be the Newton step at y, and $u = \gamma(f,x) \| x-y \|$. Then, the partial derivatives of Secant map with respect to x and y are*

$$\| \frac{\partial S(y,x)}{\partial y} \| \leq \frac{\alpha(f,x)}{(1-2u)^2}$$

$$\| \frac{\partial S(y,x)}{\partial x} \| \leq \frac{\alpha(f,y)}{(1-2v)^2}.$$

Proof. Using (2), the equation (9) can be written as

$$A^{-1}(y,x)(f(y)-f(x)) = y-x. \tag{11}$$

Let e_j denote a unit vector with 1 in jth position, for $j = 1, \ldots, n$. In view of (11),

$$\begin{aligned}\frac{-S(y+e_jh,x)+S(y,x)}{h} &= \frac{1}{h}[y+e_jh - A(x,y+e_jh)^{-1}f(y+e_jh) - y + A(x,y)^{-1}f(y)] \\ &= \frac{1}{h}[y+e_jh - A(x,y+e_jh)^{-1}f(y+e_jh) - y + A(x,y)^{-1}f(y) \\ &\quad +A(x,y+e_jh)^{-1}f(x) - A(x,y+e_jh)^{-1}f(x) \\ &\quad -A(x,y)^{-1}f(x) + A(x,y)^{-1}f(x)] \\ &= \frac{1}{h}\big(e_jh + (\cancel{x} - y - e_jh) + (y-x) + [A(x,y)^{-1} - A(x,y+e_jh)^{-1}]f(x)\big) \\ &= \frac{1}{h}\big([A(x,y)^{-1} - A(x,y+e_jh)^{-1}]f(x)\big)\end{aligned}$$

It can be shown from (10) that $\| S(x,y) - S(x,\hat{y}) \| \leq \frac{\beta_N(f,x)}{(1-2u)^2}\gamma(f,x) \| y - \hat{y} \|$, with $u = \gamma(f,x) \| x-y \|$, we have

$$\begin{aligned}\lim_{h\to 0}\frac{S(x,y+e_jh)-S(x,y)}{h} &= \lim_{h\to 0}\frac{(A^{-1}(x,y) - A(x,y+e_jh)^{-1})f(x)}{h} \\ &\leq \frac{\beta_N(f,x)}{(1-2u)^2}\gamma(f,x) \\ &\leq \frac{\alpha(f,x)}{(1-2u)^2}\end{aligned}$$

The partial derivative of $S(x,y)$ with respect to x is obtained as follows.

$$S(y, x+e_j h) - S(y,x) = y + A(x+e_j h, y)^{-1} f(y) - y - A(x,y)^{-1} f(y)$$

Using $\| S(x,y) - S(\hat{x},y) \| \leq \frac{\beta_N(f,y)}{(1-2v)^2}\gamma(f,y) \| x-\hat{x} \|, \quad v = \gamma(f,y) \| x-y \|$,

$$\begin{aligned}
\lim_{h\to 0} \frac{S(y,x+e_j h) - S(y,x)}{h} &= \lim_{h\to 0} \frac{(A^{-1}(x+e_j h, y) - A(x,y)^{-1})f(y)}{h} \\
&\leq \frac{\beta_N(f,y)\gamma(f,y)}{(1-2v)^2} \\
&\leq \frac{\alpha(f,y)}{(1-2v)^2}
\end{aligned}$$

We present a new contraction theorem as follows.

Theorem 4.5. *(Contraction) Let* $\beta_N(f,y) = \| Df(y)^{-1} f(y) \|$ *be the Newton step at* y*, and* $u = \gamma(f,x) \| x-y \|$*. If*

$$c_1 := \frac{\alpha(f,x)}{(1-2u)^2} \leq 1, \quad c_2 := \frac{\alpha(f,y)}{(1-2v)^2} \leq 1,$$

we let $c = \max(c_1, c_2)$ *and*

$$\frac{(1-u_0)^2\alpha_0 + u_0^2(3-u_0)}{\Psi(u_0)} \leq u_0,$$

then Secant iteration S_f *maps* $B\left(x_0, \frac{u_0}{\gamma(f,x_0)}\right)^2$ *into* $B\left(x_0, \frac{u_0}{\gamma(f,x_0)}\right)$*.* S_f *is a contraction mapping with constant* c*.*

Proof. From Lemma 4.4 and Lemma 2.2 in [6], the proof is straight to show that S_f is a contraction mapping with c.

The robust α theorem using the contraction constants in theorem 4.5 is derived as follows.

Theorem 4.6. *Let* u_0 *and* α_0 *be two real positive numbers such that*
(a)

$$c_0 := \max\left(\frac{\beta_N(f,y)(2v_0 - v_0^2) + 1 - 2v_0^2}{(1-v_0)^2}, \frac{1-u_0}{(1-2u_0)^2}\beta_N(f,x_0)\gamma(f,x_0)\right) < \frac{1}{2}$$

(b)

$$\frac{(1-u_0)^2\alpha_0 + u_0^2(3-2u_0)}{\Psi(u_0)} \leq u_0,$$

(c)

$$\left(u_0 + \frac{\alpha_0}{1-e_0}\right)\left(\frac{1}{\Psi(\frac{\alpha_0}{1-e_0})(1-\frac{\alpha_0}{1-e_0})}\right) \leq \frac{3-\sqrt{7}}{2},$$

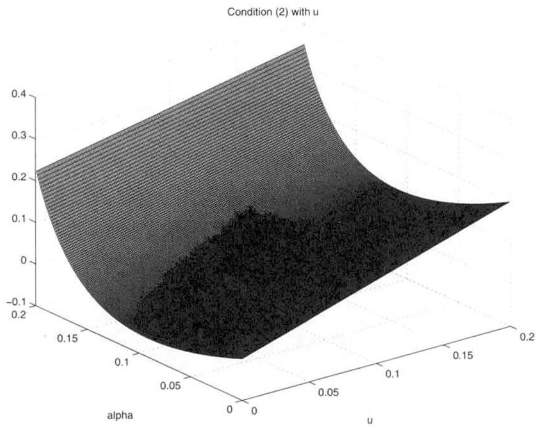

FIGURE 3. Common condition (b) with u

(d)

$$\left(\frac{1}{\Psi(\frac{\alpha_0}{1-e_0})(1-\frac{\alpha_0}{1-e_0})}\right) \leq \frac{1}{2e_0}.$$

If $\alpha(f,x_0) \leq \alpha_0$, then there is a root Ψ such that

$$B\left(x_0, \frac{u_0}{\gamma(f,x_0)}\right) \subset B\left(x_0, \frac{3-\sqrt{7}}{2\gamma(f,\Psi)}\right).$$

And, S_f maps $B\left(x_0, \frac{u_0}{\gamma(f,x_0)}\right)$ into $B\left(x_0, \frac{3-\sqrt{7}}{2\gamma(f,\Psi)}\right)$ with a constant less than or equal to $1/2$.

Proof. The result follows from Lemma 4.4, similarly to Theorem 1.2 in [6] with different c.

We now compare the constant c in Theorem 4.6 with Theorem 4.2. The convergence Theorem 4.1 and Theorem 4.2 require assumptions on the size of u. Each condition is described in Fig. 4 and Fig. 5.

The figures show the function values of

$$\boldsymbol{f} = \frac{(\alpha(f,x_0)+u_0)(1-u_0)^2}{(\Psi(u_0)(1-u_0)-2u_0)(\Psi(u_0)(1-u_0)-4u_0)} + \frac{u_0}{(\Psi(u_0)(1-u_0)-u_0)} - 1$$

$$\boldsymbol{g} = \frac{(1-u_0)^2\alpha(f,x_0)+u_0^2(3-2u_0)}{\Psi(u_0)} - u_0$$

The values of $\boldsymbol{f}$ and $\boldsymbol{g}$ less than 0 correspond to valid α and u.

Figure 3 shows the common condition (b) in 4.2 and 4.6. The valid range of values for α is $0 \leq 0.15$ for both cases approximately.

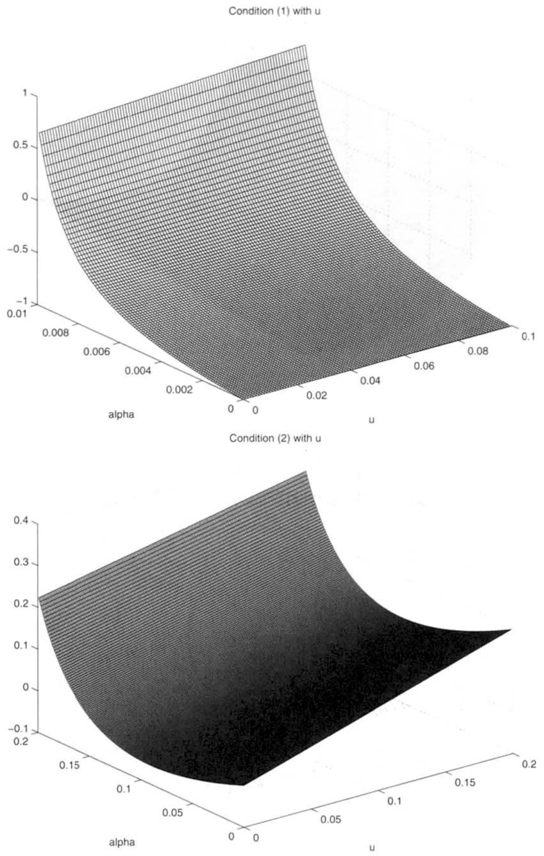

FIGURE 4. Yakoubsohn's condition (a) and New condition (a) with u

The range of α from the condition (a) of Theorem 4.2 is shown in the left side and the condition (a) of Theorem 4.6 in the right side of Figure 4 and 5. The function values less than zero are valid for the condition (a). The possible range of values of α for the condition (a) is $0 \leq \alpha \leq 0.177$. Yakoubsohn's condition (a) reduces the valid range of α to $0 \leq \alpha \leq 0.008$. Therefore, the range of α is very small compared to that of Newton method. On the other hand, as shown in Figure 5, the condition (a) in Theorem 4.6 maintains the valid range of the values for α. And, the range is closer to that of Newton map. Considering the fact that Secant iteration resembles Newton method, it is reasonable to expect α behaves like that of Newton.

5. CONCLUDING DISCUSSIONS

We have presented a new condition for contraction theorem for Secant method. The conditions are derived from the observation that Secant method approximates the

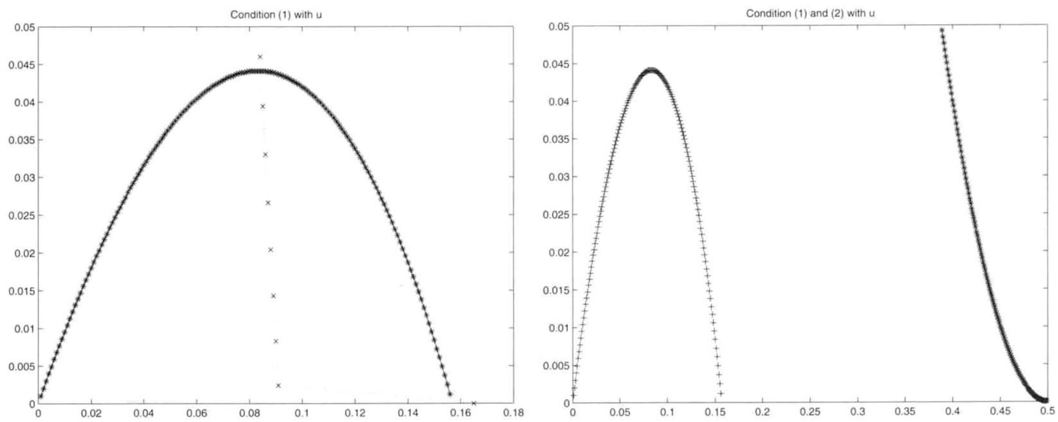

FIGURE 5. Yakoubsohn's condition (a) and (b), and New condition (a) and (b) with u

Jacobian in Newton's map, therefore, in a very close approximation, the behavior of the algorithm can be expected to be similar to that of Newton's method. We have shown that new conditions can be derived using the properties of Secant map, and also in the Figures that the proposed conditions have better representation of the features of Secant maps.

REFERENCES

[1] L. Blum, F. Cucker, M. Shub, S. Smale, *Complexity and real computation,* Springer-Verlag, New York, 1997.
[2] J. E. Dennis and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall Inc., Englewood Cliffs, New Jersey, (1983).
[3] J. J. Moré, B. S. Garbow, and K. E. Hillstrom, Testing Unconstrained Optimization Software, *ACM Transactions on Mathematical Software*, **7**, No. 1, 17-41, (1981).
[4] J. Ortega and V. Rheinbolt, *Numerical Solutions of Nonlinear Problems*, SIAM, Philadelphia, (1968).
[5] J. C. Yakoubsohn, Finding zeros of Analytic Functions, α Theory for Secant Type Methods, *J. of Complexity* **15**, 239-281, (1999).
[6] J. C. Yakoubsohn, Contraction, Robustness and Numerical Path-Following Using Secant Maps, *J. of Complexity* **16**, 286-310, (2000).

Department of Mathematics
Ewha Womans University
Dahyun-Dong, Seoul, Korea, 120-750
skim@mm.ewha.ac.kr,