On the Superlinear Convergence of the Secant Method

Marco Vianello and Renato Zanovello

This note is devoted to filling a gap present in most of the numerical analysis textbooks, concerning the discussion on the superlinear convergence of the secant method.

Let us consider the secant method for the numerical solution of $f(x) = 0$ (cf., e.g., [**2**, §6.4])

$$x_{n+1} = x_n - f(x_n)\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}, \qquad f(x_n) \neq f(x_{n-1}), \quad n \geq 1. \quad (1)$$

It is well known that the method converges, for sufficiently good initial approximations x_0 and x_1, if $f'(\xi) \neq 0$ and $f(x)$ has a continuous second order derivative (at least in a neighborhood of the zero ξ). It is also known (cf., e.g., [**1**, §3.5]) that the fundamental three-term recurrence relation holds

$$e_{n+1} = -\frac{f[x_{n-1}, x_n, \xi]}{f[x_{n-1}, x_n]} e_n e_{n-1}, \qquad n \geq 1, \quad (2)$$

where $e_n = \xi - x_n$, and where $f[t_0, \dots, t_m]$ denotes the mth divided difference at the points $t_0, \dots, t_m$ (cf., e.g., [**1**, §2.3]). From (2) follows

$$|e_{n+1}| = c_n |e_n|\, |e_{n-1}|$$

$$c_n = \frac{1}{2}\left|\frac{f''(\xi_n)}{f'(\eta_n)}\right|, \qquad \xi_n \in \operatorname{conh}(x_{n-1}, x_n, \xi), \qquad \eta_n \in \operatorname{conh}(x_{n-1}, x_n), \quad (3)$$

$\operatorname{conh}(t_0, \dots, t_m)$ denoting the open convex-hull of the points $t_0, \dots, t_m$.

However, the determination of the order of convergence and of the asymptotic error constant is carried out unsatisfactorily in most of the textbooks. Indeed, either the discussion is heuristic in nature, or, having assumed that

$$\lim_{n \to \infty} \frac{|e_{n+1}|}{|e_n|^p} = C \quad (4)$$

for some positive constants p and C, their respective values are deduced.

Anyway for a rigorous treatment the usual reference is the classical book by Ostrowski [**3**], where, in addition, hypotheses on the third derivative $f'''(x)$ are introduced in order to study the asymptotic behavior.

In [**4**] the superlinear convergence property is proved without resorting to the third derivative, but such a proof is long and complex because of its generality, being addressed to a wide class of iterative methods.

The purpose of this note is to provide a rigorous and quite simple proof of the superlinear convergence of the secant method, under natural assumptions on $f(x)$.

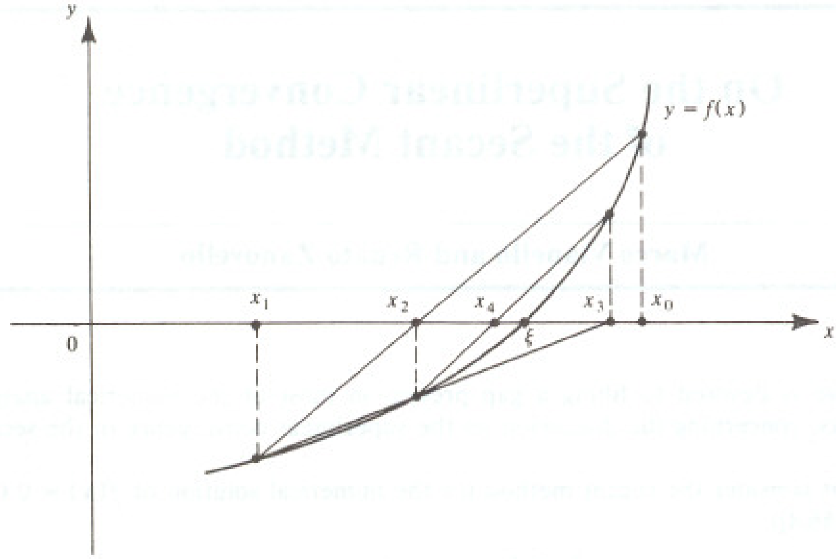

Figure 1. The Secant Method for finding roots. It is classically known that under mild restrictions, convergence is superlinear, and the errors satisfy a simple (Fibonacci-like) recurrence. There is a simple and elegant proof of this fact.

Following [1, §3.5, p. 103], setting $y_n = |e_n|/|e_{n-1}|^p$, $n \geq 1$ with $p > 0$, it is immediately seen from (3) that

$$y_{n+1} = c_n y_n^{-1/p}, \qquad n \geq 1 \tag{5}$$

if p is the positive solution of $t^2 - t - 1 = 0$, i.e. p is the "golden ratio" $(1+\sqrt{5})/2$. Now, assuming $f''(\xi) \neq 0$, we'll prove that

$$\lim_{n \to \infty} y_n = \left[\frac{1}{2}\left|\frac{f''(\xi)}{f'(\xi)}\right|\right]^{1/p} = C. \tag{6}$$

Taking logarithms in (5) and defining $z_n = \ln(y_n)$, $\alpha_n = \ln(c_n)$, we get the new first order linear difference equation

$$z_{n+1} = \alpha_n - \frac{1}{p} z_n, \qquad n \geq 1 \tag{7}$$

that can be immediately solved by recurrence, obtaining

$$z_n = \left(-\frac{1}{p}\right)^{n-1} z_1 + S_n, \qquad n \geq 2, \tag{8}$$

where

$$S_n = \sum_{j=0}^{n-2} \alpha_{n-j-1}\left(-\frac{1}{p}\right)^j. \tag{9}$$

Since $p > 1$, $\{z_n\}$ (and hence $\{y_n\}$) converges if and only if $\{S_n\}$ converges.

At this point it is clear that the problem can be reduced to studying the asymptotic behavior of a sequence like

$$\sigma_n = \sum_{j=0}^{n} a_{n-j} b_j \tag{10}$$

where

$$\lim_{n\to\infty} a_n = a \quad \text{and} \quad \sum_{j=0}^{\infty} |b_j| < \infty. \tag{11}$$

The sequence defined by (10) is usually termed the *convolution* of the two sequences $\{a_n\}$ and $\{b_n\}$. It naturally appears, for instance, as general term in the Cauchy product of the series $\Sigma^\infty a_n$, $\Sigma^\infty b_n$.

Now we'll prove that

$$\lim_{n\to\infty} \sigma_n = ab, \qquad b = \sum_{j=0}^{\infty} b_j, \tag{12}$$

exploiting essentially the approach used in proving the well-known Cesaro's theorem for sequences. A second, more abstract proof of (12), which we omit for brevity, could be given by the well-known dominated convergence theorem.

Proof: Without any loss of generality we can assume $a = 0$ in (11). In fact

$$\sigma_n - ab = \sum_{j=0}^{n} (a_{n-j} - a) b_j - a \sum_{j=n+1}^{\infty} b_j$$

and by the summability of $\{b_n\}$ the second term in the right-hand side above is infinitesimal as $n \to \infty$. Let us split the sum (10) in the following way

$$\sigma_n = \sum_{j=0}^{m} a_{n-j} b_j + \sum_{j=m+1}^{n} a_{n-j} b_j, \tag{13}$$

where $m \geq 0$, $n \geq m + 1$. Fix $\varepsilon > 0$. In view of (11) we can determine two positive indexes

$$\nu_1(\varepsilon) \text{ such that } |a_k| < \frac{\varepsilon}{2\beta} \text{ for } k \geq \nu_1$$

$$\nu_2(\varepsilon) \text{ such that } \sum_{j=\nu_2}^{\infty} |b_j| < \frac{\varepsilon}{2M},$$

where M is an upper bound for $|a_n|$ and $\beta = \Sigma_{j=0}^{\infty} |b_j|$. It follows from (13) with $m + 1 = \nu_2$ that

$$|\sigma_n| < \varepsilon \quad \text{for } n \geq \nu(\varepsilon) = \nu_1 + \nu_2.$$ ■

Going back to the sequence $\{S_n\}$ defined in (9) we finally obtain from (3), (5), (12)

$$\lim_{n\to\infty} S_n = \left(\lim_{n\to\infty} \alpha_n\right) \sum_{j=0}^{\infty} \left(-\frac{1}{p}\right)^j = \ln(c)\frac{p}{p+1} = \ln(c^{1/p}),$$

where $c = \frac{1}{2}|f''(\xi)/f'(\xi)|$ and hence by (8)

$$\lim_{n\to\infty} y_n = \lim_{n\to\infty} \exp(z_n) = c^{1/p} \neq 0,$$

i.e., the secant method has order of convergence $p = (1 + \sqrt{5})/2$ and asymptotic error constant

$$C = \left[\frac{1}{2}\left|\frac{f''(\xi)}{f'(\xi)}\right|\right]^{1/p}.$$

A last remark has to be made. The discussion above is based on the assumption $f''(\xi) \neq 0$. If, on the contrary, $f''(\xi) = 0$, excluding the trivial case that the method yields the root in a finite number of steps, we have $f''(\xi_n) \neq 0$ for all n. Then $\lim_{n\to\infty} \ln(c_n) = -\infty$ and hence $C = \lim_{n\to\infty} y_n = 0$. Thus the convergence order of the secant method may be greater than p. To conclude we can say, following e.g. [4], that the convergence of the secant method is superlinear.

REFERENCES

1. S. D. Conte and C. de Boor, *Elementary Numerical Analysis*, International Student Edition, McGraw-Hill Kogakusha, Tokyo, 1980.
2. G. Dahlquist, Å. Björck and N. Anderson, *Numerical Methods*, Prentice-Hall, Englewood Cliffs, N.J., 1974.
3. A. M. Ostrowski, *Solution of Equations and Systems of Equations*, Academic Press, New York and London, 1960.
4. J. F. Traub, *Iterative Methods for the Solution of Equations*, Prentice-Hall, Englewood Cliffs, N.J., 1964.

Dipartimento di Matematica Pura e Applicata
Università di Padova
via Belzoni 7
35131 Padova, ITALY