

Article

Convergence and Stability Improvement of Quasi-Newton Methods by Full-Rank Update of the Jacobian Approximates

Peter Berzi 

Applied Informatics and Applied Mathematics Doctoral School, Óbuda University, Bécsi út 96/B,
1034 Budapest, Hungary; berzi.peter@uni-obuda.hu; Tel.: +36-20-389-0242

Abstract: A system of simultaneous multi-variable nonlinear equations can be solved by Newton's method with local q-quadratic convergence if the Jacobian is analytically available. If this is not the case, then quasi-Newton methods with local q-superlinear convergence give solutions by approximating the Jacobian in some way. Unfortunately, the quasi-Newton condition (Secant equation) does not completely specify the Jacobian approximate in multi-dimensional cases, so its full-rank update is not possible with classic variants of the method. The suggested new iteration strategy ("T-Secant") allows for a full-rank update of the Jacobian approximate in each iteration by determining two independent approximates for the solution. They are used to generate a set of new independent trial approximates; then, the Jacobian approximate can be fully updated. It is shown that the T-Secant approximate is in the vicinity of the classic quasi-Newton approximate, providing that the solution is evenly surrounded by the new trial approximates. The suggested procedure increases the superlinear convergence of the Secant method ($\varphi_S = 1.618 \dots$) to super-quadratic ($\varphi_T = \varphi_S + 1 = 2.618 \dots$) and the quadratic convergence of the Newton method ($\varphi_N = 2$) to cubic ($\varphi_T = \varphi_N + 1 = 3$) in one-dimensional cases. In multi-dimensional cases, the Broyden-type efficiency (mean convergence rate) of the suggested method is an order higher than the efficiency of other classic low-rank-update quasi-Newton methods, as shown by numerical examples on a Rosenbrock-type test function with up to 1000 variables. The geometrical representation (hyperbolic approximation) in single-variable cases helps explain the basic operations, and a vector-space description is also given in multi-variable cases.

Keywords: quasi-Newton methods; multi-variable nonlinear equations; full-rank Jacobian approximate update; Rosenbrock function; super-quadratic convergence; efficiency



Citation: Berzi, P. Convergence and Stability Improvement of Quasi-Newton Methods by Full-Rank Update of the Jacobian Approximates. *AppliedMath* **2024**, *4*, 143–181. <https://doi.org/10.3390/appliedmath4010008>

Academic Editors: Jai Prakash Jaiswal and Juan Ramón Torregrosa Sánchez

Received: 15 November 2023

Revised: 18 December 2023

Accepted: 12 January 2024

Published: 26 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

It is a common task in numerous disciplines (e.g., physics, chemistry, biology, economics, robotics, and engineering, social, and medical sciences) to construct a mathematical model with some parameters for an observed system which gives an observable response to an observable external effect. The unknown parameters of the mathematical model are determined so that the difference between the observed and the simulated system responses of the mathematical model for the same external effect is minimized (see e.g., [1–11]). This problem leads to finding the zero of a residual function (difference between observed and simulated responses). The rapidly accelerating computational tools and the increasing complexity of mathematical models with more and more efficient numerical algorithms provide a chance for better understanding and control of the surrounding nature.

As referenced above, root-finding methods are essential for solving a great class of numerical problems, such as data fitting problems with m sampled data $D = [D_j]$ ($j = 1, \dots, m$) and n adjustable parameters $x = [x_i]$ ($i = 1, \dots, n$) with $m \geq n$. This leads to the problem of least-squares solving of an over-determined system of nonlinear equations,

$$f(x) = 0, \quad (1)$$

($x \in \mathbb{R}^n$ and $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ($m \geq n$)), where the solution x^* minimizes the difference

$$\|f(x)\|_2 = \|\phi(x) - D\|_2 \quad (2)$$

between the data D and a computational model function $\phi(x)$. The system of simultaneous multi-variable nonlinear Equation (1) can be solved by Newton's method when the derivatives of $f(x)$ are available analytically and a new iterate,

$$x_{p+1} = x_p - J_p^{-1} f_p, \quad (3)$$

that follows x_p can be determined, where $f_p = f(x_p)$ is the function value and $J_p = J(x_p)$ is the Jacobian matrix of f at x_p in the p^{th} iteration step. Newton's method is one of the most widely used algorithm, with very attractive theoretical and practical properties and with some limitations. The computational costs of Newton's method is high, since the Jacobian J_p and the solution to the linear system (3) must be computed at each iteration. In many cases, explicit formulae for the function $f(x)$ are not available ($f(x)$ can be a residual function between a system model response and an observation of that system response) and the Jacobian J_p can only be approximated. The classic Newton's method can be modified in many different ways. The partial derivatives of the Jacobian may be replaced by suitable difference quotients (discretized Newton iteration, see [12,13]),

$$\left[\frac{\Delta f}{\Delta x} \right] = \left[\frac{f_j(x + \Delta x_k d^k) - f_j(x)}{\Delta x_k} \right] = \left[\frac{\Delta f_j(x)}{\Delta x_k} \right], \quad (4)$$

($k = 1, \dots, n$), ($j = 1, \dots, m$) with n additional function value evaluations, where d^k is the k^{th} Cartesian unit vector. However, it is difficult to choose the stepsize Δx . If any Δx_k is too large, then Expression (4) can be a bad approximation to the Jacobian, so the iteration converges much more slowly if it converges at all. On the other hand, if any Δx_k is too small, then $\Delta f_j(x) \simeq 0$, and cancellations can occur which reduce the accuracy of the difference quotients (4) (see [14]). The suggested procedure ("T-Secant") may resemble the discretized Newton iteration, but it uses a systematic procedure to determine suitable stepsizes for the Jacobian approximates. Another modification is the inexact Newton approach, where the nonlinear equation is solved by an iterative linear solver (see [15–17]).

It is well-known that the local convergence of Newton's method is q-quadratic if the initial trial approximate x_0 is close enough to the solution x^* , $J(x^*)$ is non-singular, and $J(x)$ satisfies the Lipschitz condition

$$\|J(x) - J(x^*)\| \leq L\|x - x^*\| \quad (5)$$

for all x close enough to x^* . However, in many cases, the function $f(x)$ is not an analytical function, the partial derivatives are not known, and Newton's method cannot be applied. Quasi-Newton methods are defined as the generalization of Equation (3) as

$$x_{p+1} = x_p - B_p^{-1} f_p \quad (6)$$

and

$$B_p \Delta x_p = -f_p \quad (7)$$

where

$$\Delta x_p = x_{p+1} - x_p \quad (8)$$

is the iteration step length and B_p is expected to be the approximate to the Jacobian matrix J_p without computing derivatives in most cases. The new iterate is then given as

$$x_{p+1} = x_p + \Delta x_p \quad (9)$$

and B_p is updated to B_{p+1} according to the specific quasi-Newton method. Martinez [18] has made a thorough survey on practical quasi-Newton methods. The iterative methods of the form (6) that satisfy the equation

$$B_{p+1} \Delta x_p = f_{p+1} - f_p \quad (10)$$

for all $k = 0, 1, 2, \dots$ are called “quasi-Newton” methods, and Equation (10) is called the fundamental equation of quasi-Newton methods (“quasi-Newton condition” or “secant equation”). However, the quasi-Newton condition does not uniquely specify the updated Jacobian approximate B_{p+1} , and further constraints are needed. Different methods offer their own specific solution. One new quasi-Newton approximate x_{p+1} will never allow for a full-rank update of B_{p+1} because it is an $n \times n$ matrix and only n components can be determined from the Secant equation, making it an under-determined system of equations for the elements $[B_{i,j,p+1}]$ ($i, j = 1, \dots, n$) if $n > 1$.

The suggested new strategy is based on Wolfe’s [19] formulation of a generalized Secant method. The function

$$x \rightarrow f(x), \text{ where } x \in \mathbb{R}^n \text{ and } f: \mathbb{R}^n \rightarrow \mathbb{R}^n, n > 1 \quad (11)$$

is locally replaced by linear interpolation through $n + 1$ interpolation base points $A_p, B_{p,k}$ ($k = 1, \dots, n$). The variables x and the function values f are separated into two equations and an auxiliary variable q^A is introduced. Then the Jacobian approximate matrix B_p is split into a variable difference ΔX_p and a function value difference ΔF_p matrix, and the zero x_{p+1}^A of the p^{th} interpolation plane is determined from the quasi-Newton condition (7) as

$$\begin{bmatrix} \Delta x_{p+1}^A \\ -f_p^A \end{bmatrix} = \begin{bmatrix} \Delta X_p \\ \Delta F_p \end{bmatrix} q_p^A \quad (12)$$

where

$$\Delta x_{p+1}^A = x_{p+1}^A - x_p^A. \quad (13)$$

The auxiliary variable q_p^A is determined from the second row of Equation (12), and the new quasi-Newton approximate x_{p+1}^A comes from the first row of this equation. Popper [20] made further generalization for functions

$$x \rightarrow f(x), \text{ where } x \in \mathbb{R}^n \text{ and } f: \mathbb{R}^n \rightarrow \mathbb{R}^m, m \geq n > 1 \quad (14)$$

and suggested the use of a pseudo-inverse solution for the over-determined system of linear equations (where n is the number of unknowns and m is the number of function values). The auxiliary variable q_p^A is determined from the second row of Equation (12) as

$$q_p^A = -\Delta F_p^+ f_p^A, \quad (15)$$

where $[.]^+$ stands for the pseudo-inverse, and the new quasi-Newton approximate x_{p+1}^A comes from the first row of this equation as

$$x_{p+1}^A = x_p^A - \Delta X_p \Delta F_p^+ f_p^A. \quad (16)$$

The new iteration continues with $n + 1$ new base points $A_{p+1}, B_{p+1,k}$ ($k = 1, \dots, n$). Details are given in Section 3.

Ortega and Rheinboldt [12] stated that a necessary condition of convergence is that the interpolation base points should be linearly independent and they have to be “in general position” through the whole iteration process. Experiences show that the low-rank update procedures often lead to a dead end because this condition is not satisfied. The purpose of the suggested new iteration strategy is to determine linearly independent base points providing that the Ortega and Rheinboldt condition is satisfied. The basic idea

of the procedure is that another new approximate \mathbf{x}_{p+1}^B is determined from the previous approximate \mathbf{x}_{p+1}^A and a new system of n linearly independent base points is generated. The basic equations of the Wolfe–Popper formulation (Equation (12)) were modified as

$$\begin{bmatrix} \Delta \mathbf{x}_{p+1} \\ -\mathbf{f}_p^A \end{bmatrix} = \begin{bmatrix} \mathbf{T}_p^X & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_p^F \end{bmatrix} \begin{bmatrix} \Delta \mathbf{X}_p \\ \Delta \mathbf{F}_p \end{bmatrix} \mathbf{q}_p^B \quad (17)$$

where

$$\Delta \mathbf{x}_{p+1} = \mathbf{x}_{p+1}^B - \mathbf{x}_{p+1}^A \quad (18)$$

$$\mathbf{T}_p^X = \text{diag}(t_{p,i}^X) = \text{diag}\left(\frac{x_{p+1,i}^B - x_{p+1,i}^A}{x_{p+1,i}^A - x_{p,i}^A}\right) \quad (19)$$

and

$$\mathbf{T}_p^F = \text{diag}(t_{p,j}^F) = \text{diag}\left(\frac{f_{p+1,j}^B - f_{p+1,j}^A}{f_{p+1,j}^A - f_{p,j}^A}\right). \quad (20)$$

The auxiliary variable \mathbf{q}_p^B is determined from the second row of Equation (17) as

$$\mathbf{q}_p^B = -\Delta \mathbf{F}_p^+ (\mathbf{T}_p^F)^{-1} \mathbf{f}_p^A = -\left[\sum_{j=1}^m \left(\Delta \mathbf{F}_{p,i,j} + \frac{f_{p,j}^A}{t_{p,j}^F} \right) \right], \quad (21)$$

and the new quasi-Newton approximate \mathbf{x}_{p+1}^B comes from the first row of Equation (17) as

$$x_{p+1,i}^B = x_{p+1,i}^A + \frac{(\Delta x_{p,i}^A)^2}{\Delta x_{p,i} \mathbf{q}_{p,i}^B} = x_{p+1,i}^A - \frac{(\Delta x_{p,i}^A)^2}{\sum_{j=1}^m \left(\Delta x_{p,i} \Delta \mathbf{F}_{p,i,j} + \frac{f_{p,j}^A}{t_{p,j}^F} \right)} \quad (22)$$

($i = 1, \dots, n$). The details of the proposed new strategy (“T-Secant method”) are given in Section 4. It is different from the traditional Secant method in that all interpolation base points A_p and $B_{p,k}$ ($k = 1, \dots, n$) are updated in each iteration (full-rank update), providing $n + 1$ new base points A_{p+1} and $B_{p+1,k}$ for the next iteration. The key idea of the method is very simple. The function value f_{p+1}^A (that can be determined from the new Secant approximate \mathbf{x}_{p+1}^A) measures the “distance” of the approximate \mathbf{x}_{p+1}^A from the root \mathbf{x}^* (if $f_{p+1}^A = 0$, then the distance is zero and $\mathbf{x}_{p+1}^A = \mathbf{x}^*$). The T-Secant method uses this information so that the basic equations of the Secant method are modified by a scaling transformation \mathbf{T} , and an additional new estimate \mathbf{x}_{p+1}^B is determined. Then, the new approximates \mathbf{x}_{p+1}^A and \mathbf{x}_{p+1}^B are used to construct the $n + 1$ new interpolation base points A_{p+1} and $B_{p+1,k}$.

The T-Secant procedure has been worked out for solving multi-variable problems. It can also be applied for solving single-variable ones, however. The geometrical representation of the latter provides a good view with which to explain the mechanism of the procedure as shown in Section 5. It is a surprising result that the T-Secant modification corresponds to a hyperbolic function

$$z_p(x) = \frac{a_p}{x - x_{p+1}^A} + f_p^A, \quad (23)$$

the zero of which gives the second approximate x_{p+1}^B in the single-variable case. A vector space interpretation is also given for the multivariable case in this section.

The general formulations of the proposed method are given in Section 6 and compared with the basic formula of classic quasi-Newton methods. It follows from Equation (16) that

$$S_p \Delta x_p^A = -f_p^A, \quad (24)$$

where

$$S_p = \Delta F_p \Delta X_p^{-1} = \begin{bmatrix} \frac{\Delta f_{1,1,p}}{\Delta x_1} & \cdots & \frac{\Delta f_{n,1,p}}{\Delta x_n} \\ \vdots & \vdots & \vdots \\ \frac{\Delta f_{1,m,p}}{\Delta x_1} & \cdots & \frac{\Delta f_{n,m,p}}{\Delta x_n} \end{bmatrix} = \begin{bmatrix} \frac{\Delta f_{k,j,p}}{\Delta x_{i,p}} \end{bmatrix} \quad (25)$$

is the Jacobian approximate of the traditional Secant method. It follows from the first and second rows of Equation (17) of the T-Secant method and from the Definition (25) of S_p that

$$S_{T,p} \Delta x_p^A = -f_p^A \quad (26)$$

is the modified Secant equation, where

$$S_{T,p} = T_p^F S_p (T_p^X)^{-1} = T_p^F \Delta F_p \Delta X_p^{-1} (T_p^X)^{-1} = \begin{bmatrix} \frac{t_{j,p}^F}{t_{i,p}^X} \frac{\Delta f_{k,j,p}}{\Delta x_{i,p}} \end{bmatrix}. \quad (27)$$

It is well known that the single-variable Secant method has asymptotic convergence for sufficiently good initial approximates x^A and x^B if $f'(x)$ does not vanish in $x \in [x^A, x^B]$ and $f''(x)$ is continuous at least in a neighborhood of the zero x^* . The super-linear convergence property has been proved in different ways, and it is known that the order of convergence is $\alpha = (1 + \sqrt{5})/2 = \varphi$ (where $\varphi = 1.618 \dots$ is the golden ratio). The convergence order of the proposed method is determined in Section 7, and it is shown that it has super-quadratic convergence with rate $\alpha^{TS} = \varphi + 1 = \varphi^2 = 2.618 \dots$ in the single variable case. It is also shown for the multi-variable case in this section that the second approximate x_{p+1}^B will always be in the vicinity of the classic Secant approximate x_{p+1}^A , providing that the solution x^* will evenly be surrounded by the $n + 1$ new trial approximates and matrix S_{p+1} will be well-conditioned.

A step-by-step algorithm is given in Section 8, and the results of numerical tests with a Rosenbrock-type test function demonstrates the stability of the proposed strategy in Section 9 for up to 1000 unknown variables. The Broyden-type efficiency (mean convergence rate) of the proposed method is studied in a multi-variable case in Section 10, and it is compared with other classic rank-one update and line-search methods on the basis of available test data. It is shown in Section 11 how the new procedure can be used to improve the convergence of other classic multi-variable root finding methods (Newton–Raphson and Broyden methods). Concluding remarks are summarized in Section 12. Among others, the method has been used for the identification of vibrating mechanical systems (foundation pile driving [21,22], percussive drilling [23]) and found to be very stable and efficient even in cases with a large number of unknowns.

The proposed method needs $n + 1$ function value evaluations in each iteration, and it is not using the derivative information of the function like the Newton–Raphson method is doing. On the other hand, it needs n more function evaluations than the traditional secant method needs in each iteration. However, this is an apparent disadvantage, as the convergence rate considerably increases ($\alpha^{TS} \cong 2.618 \dots$). Furthermore, the stability and the efficiency of the procedure has been greatly improved.

2. Notations

Vectors and matrices are denoted by bold-face letters. Subscripts refer to components of vectors and matrices; superscripts A and B refer to interpolation base points. Notations A and B are introduced so as to be able to clearly distinguish between the two new approx-

imates x^A and x^B . Vectors and matrices may also be given by their general elements. Δ refers to a difference of two elements. x and X denote unknown quantities. f and F denote function values and matrices. q, \mathbf{q}, t , and T denote multiplier scalars, vectors, and matrices. e, ε , and E denote approximate error. p is iteration counter, α is convergence rate, and ε^* is termination criterion. n is the number of unknowns, m is the number of function values, and i, j, k , and l are running indexes of matrix columns and rows. Superscripts S and TS refer to the traditional Secant method and to the proposed T-Secant method, receptively.

3. Secant Method

The history of the Secant method in single-variable cases is several thousands of years old, its origin was found in ancient times. The idea of finding the scalar root x^* of a scalar nonlinear function

$$x \rightarrow f(x) \text{ (where } x \in \mathbb{R}^1 \text{ and } f: \mathbb{R}^1 \rightarrow \mathbb{R}^1 \text{)} \quad (28)$$

by successive local replacement of the function by linear interpolation (secant line) gives a simple and efficient numerical procedure. It has the advantage that it does not need the calculation of function derivatives, it only uses function values, and the order of asymptotic convergence is super-linear with a convergence of rate $\alpha^S \cong 1.618 \dots$

The function $f(x)$ is locally replaced by linear interpolation (secant line) through interpolation base points A and B , and the zero x^A of the Secant line is determined as an approximate to the zero x^* of the function. The next iteration continues with new base points, selected from available old ones. Wolfe [19] extended the scalar procedure to multidimensional

$$\mathbf{x} \rightarrow f(\mathbf{x}), \text{ where } \mathbf{x} \in \mathbb{R}^n \text{ and } f: \mathbb{R}^n \rightarrow \mathbb{R}^n, n > 1, \quad (29)$$

and Popper [20] made a further generalization

$$\mathbf{x} \rightarrow f(\mathbf{x}), \text{ where } \mathbf{x} \in \mathbb{R}^n \text{ and } f: \mathbb{R}^n \rightarrow \mathbb{R}^m, m \geq n > 1 \quad (30)$$

and suggested use of pseudo-inverse solution for the over-determined system of linear equations (where n is the number of unknowns and m is the number of function values).

The zero x^* of the nonlinear function $x \rightarrow f(x)$ has to be found, where $x \in \mathbb{R}^n$ and $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$. Let x^A be the initial trial for the zero x^* , and let the function $f(x)$ be linearly interpolated through $n + 1$ interpolation base points $A \begin{pmatrix} x^A & f^A \end{pmatrix}$ and $B_k \begin{pmatrix} x_k^B & f_k^B \end{pmatrix}$ ($k = 1, \dots, n$) and be approximated/replaced by the interpolation “plane” $y(x)$ near x^* . One of the key ideas of the suggested numerical procedure is that interpolation base points $B_k \begin{pmatrix} x_k^B & f_k^B \end{pmatrix}$ are constructed by individually incrementing the coordinates x_i^A of the initial trial x^A by an “initial trial increment” value Δx_i ($i = 1, \dots, n$) as

$$x_{k,i}^B = x_i^A + \Delta x_i, \quad (31)$$

or in vector form as

$$\mathbf{x}_k^B = \mathbf{x}^A + \Delta x_k \mathbf{d}^k, \quad (32)$$

where \mathbf{d}^k is the k^{th} Cartesian unit vector, as shown in Figure 1.

It follows from this special construction of the initial trials \mathbf{x}_k^B that $x_{k,i}^B - x_i^A = 0$ for $i \neq k$ and $x_{k,i}^B - x_i^A = \Delta x_i$ for $i = k$ providing that

$$\Delta \mathbf{x} = \begin{bmatrix} x_{1,i}^B - x_i^A \\ \vdots \\ x_{n,i}^B - x_i^A \end{bmatrix} = [\Delta x_i] \quad (33)$$

is the “initial trial increment vector”. Let

$$\Delta \mathbf{f}_k = \begin{bmatrix} f_{k,j}^B - f_j^A \end{bmatrix} \quad (34)$$

($j = 1, \dots, m$). Any point on the n dimensional interpolation plane $y(x)$ can be expressed as

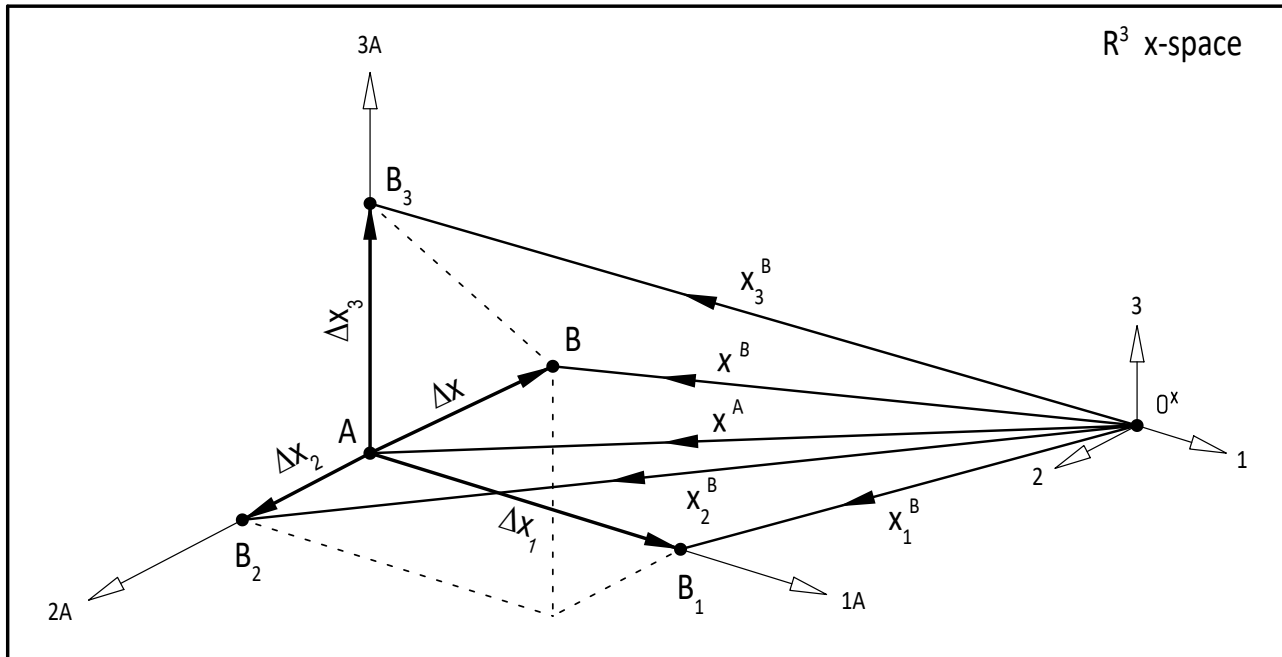


Figure 1. Formulation of a new set of base vectors ($n = 3$): x^A, x_1^B, x_2^B, x_3^B and interpolation base points A, B_1, B_2 and B_3 from new approximate x^A and from new trial increment (iteration stepsize) $\Delta x = x^B - x^A = [\Delta x_1 \quad \Delta x_2 \quad \Delta x_3]^T$.

$$\begin{bmatrix} x \\ y(x) \end{bmatrix} = \begin{bmatrix} x^A \\ f^A \end{bmatrix} + \begin{bmatrix} \Delta X \\ \Delta F \end{bmatrix} q^A, \quad (35)$$

where

$$\Delta X = [x_k^B - x^A] = \begin{bmatrix} x_{1,1}^B - x_1^A & \cdots & x_{n,1}^B - x_1^A \\ \vdots & \ddots & \vdots \\ x_{1,n}^B - x_n^A & \cdots & x_{n,n}^B - x_n^A \end{bmatrix} \quad (36)$$

$$\Delta F = [f_k^B - f^A] = \begin{bmatrix} f_{1,1}^B - f_1^A & \cdots & f_{n,1}^B - f_1^A \\ \vdots & \vdots & \vdots \\ f_{1,m}^B - f_m^A & \cdots & f_{n,m}^B - f_m^A \end{bmatrix} \quad (37)$$

($k = 1, \dots, n$), q^A is a vector with n scalar multipliers q_i^A ($i = 1, \dots, n$), and as a consequence of Equation (32),

$$\Delta X = [\Delta x_k] = \begin{bmatrix} \Delta x_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Delta x_n \end{bmatrix} = \text{diag}(\Delta x_k) \quad (38)$$

is a diagonal matrix that has great computational advantage. It also follows from Definition (34) that

$$\Delta F = [\Delta f_k] = \begin{bmatrix} \Delta f_{1,1} & \cdots & \Delta f_{n,1} \\ \vdots & \vdots & \vdots \\ \Delta f_{1,m} & \cdots & \Delta f_{n,m} \end{bmatrix}. \quad (39)$$

Let \mathbf{x}_{p+1}^A be the zero of the n -dimensional interpolation plane $\mathbf{y}_p(\mathbf{x})$ with interpolation base points $A_p \left(\mathbf{x}_p^A \quad f_p^A \right)$ and $B_{k,p} \left(\mathbf{x}_{k,p}^B \quad f_{k,p}^B \right)$ in the p^{th} iteration. Then, it follows from the zero condition that

$$\mathbf{y}_p(\mathbf{x}_{p+1}^A) = \mathbf{0} \quad (40)$$

and from the second row of Equation (35) that

$$\Delta \mathbf{F}_p \mathbf{q}_p^A = -f_p^A, \quad (41)$$

and the vector \mathbf{q}_p^A of multipliers $q_{p,i}^A$ can be expressed as

$$\mathbf{q}_p^A = -\Delta \mathbf{F}_p^+ f_p^A = - \left[\sum_{j=1}^m \left(\Delta F_{p,i,j}^+ f_{p,j}^A \right) \right], \quad (42)$$

where $[\cdot]^+$ stands for the pseudo-inverse. Let

$$\begin{bmatrix} \Delta \mathbf{x}_p^A \\ \Delta f_p^A \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{p+1}^A - \mathbf{x}_p^A \\ f_{p+1}^A - f_p^A \end{bmatrix} \quad (43)$$

be the iteration stepsize of the Secant method; then, it follows from the first row of Equation (35) and from Equation (42) that

$$\Delta \mathbf{x}_p^A = \Delta \mathbf{X}_p \mathbf{q}_p^A = -\Delta \mathbf{X}_p \Delta \mathbf{F}_p^+ f_p^A, \quad (44)$$

and from Definition (43), it follows that

$$\begin{bmatrix} \mathbf{x}_{p+1}^A \\ f_{p+1}^A \end{bmatrix} = \begin{bmatrix} \mathbf{x}_p^A + \Delta \mathbf{x}_p^A \\ f_p^A + \Delta f_p^A \end{bmatrix}, \quad (45)$$

and the new Secant approximate \mathbf{x}_{p+1}^A can be expressed from Equation (44) as

$$\mathbf{x}_{p+1}^A = \mathbf{x}_p^A + \Delta \mathbf{x}_p^A. \quad (46)$$

A new base point $A_{p+1}(\mathbf{x}_{p+1}^A, f_{p+1}^A)$ can then be determined for the next iteration. In a single-variable case ($m = n = k = 1$) with interpolation base points $A_p(x_p^A, f_p^A)$ and $B_p(x_p^B, f_p^B)$, Equation (42) will have the form

$$q_p^A = -\frac{f_p^A}{f_p^B - f_p^A} = -\frac{f_p^A}{\Delta f_p^A}, \quad (47)$$

and the new Secant approximate

$$x_{p+1}^A = x_p^A + \Delta x_p q_p^A = x_p^A - \frac{\Delta x_p}{\Delta f_p^A} f_p^A = \frac{x_p^A f_p^B - x_p^B f_p^A}{f_p^B - f_p^A} \quad (48)$$

can be determined according to Equation (46). The procedure then continues with new interpolation base points $A_{p+1}(x_{p+1}^A, f_{p+1}^A)$ and $B_{p+1}(x_{p+1}^B, f_{p+1}^B)$.

4. T-Secant Method

4.1. Single-Variable Case

The T-Secant method is different from the traditional Secant method in the way that all interpolation base points A_p and $B_{p,k}$ ($k = 1, \dots, n$) are updated in each iteration, providing $n + 1$ new base points A_{p+1} and $B_{p+1,k}$ for the next iteration. The key idea of the method is very simple. The function value f_{p+1}^A (that can be determined from the new Secant

approximate x_{p+1}^A) measures the “distance” of the approximate x_{p+1}^A from the root x^* (if $f_{p+1}^A = 0$, then the distance is zero and $x_{p+1}^A = x^*$). The T-Secant method uses this information to determine another approximate x_{p+1}^B . In a single-variable case ($m = n = k = 1$) with interpolation base points A_p and B_p , the basic equation

$$\Delta f_p q_p^A = -f_p^A \quad (49)$$

of the Secant method (Equation (41) in multi-variable case) is modified by a factor

$$t_p^f = \frac{f_{p+1}^A}{f_p^A} \quad (50)$$

that expresses the “improvement rate” of the new approximate x_{p+1}^A to the original approximate x_p^A , providing the T-Secant-modified basic equation

$$t_p^f \Delta f_p q_p^B = -f_p^A. \quad (51)$$

Then, the T-Secant multiplier

$$q_p^B = \frac{q_p^A}{t_p^f} = -\frac{(f_p^A)^2}{f_{p+1}^A \Delta f_p} \quad (52)$$

can be determined. The other basic equation

$$\Delta x_p^A = \Delta x_p q_p^A \quad (53)$$

of the Secant method (Equation (44) in multi-variable case) with iteration stepsize

$$\Delta x_p^A = x_{p+1}^A - x_p^A \quad (54)$$

is also modified in a similar way to that in the case of Equations (49) and (51) by a factor

$$t_p^x = \frac{x_{p+1}^B - x_{p+1}^A}{x_{p+1}^A - x_p^A} = \frac{\Delta x_{p+1}}{\Delta x_p^A} \quad (55)$$

that expresses the “improvement rate” of the new “T-Secant stepsize” Δx_{p+1} to the previous “secant stepsize” Δx_p^A , providing a new basic equation

$$\Delta x_p^A = t_p^x \Delta x_p q_p^B, \quad (56)$$

from which

$$\Delta x_p^A = \frac{\Delta x_{p+1}}{\Delta x_p^A} \Delta x_p q_p^B \quad (57)$$

and

$$x_{p+1}^A - x_p^A = -\frac{x_{p+1}^B - x_{p+1}^A}{x_{p+1}^A - x_p^A} (x_p^B - x_p^A) \frac{(f_p^A)^2}{f_{p+1}^A (f_p^B - f_p^A)}. \quad (58)$$

By re-ordering Equation (58), the T-Secant approximate

$$x_{p+1}^B = x_{p+1}^A + \frac{(\Delta x_p^A)^2}{\Delta x_p q_p^B} = x_{p+1}^A - \frac{(x_{p+1}^A - x_p^A)^2 (f_p^B - f_p^A) f_{p+1}^A}{(x_p^B - x_p^A) (f_p^A)^2} \quad (59)$$

can be determined, and it is used to update the original interpolation base point B_p to B_{p+1} . The new iteration will then continue with new base points A_{p+1} and B_{p+1} . Note that it follows from Equations (52), (53), and (59) that

$$\Delta x_{p+1} = x_{p+1}^B - x_{p+1}^A = \frac{(\Delta x_p^A)^2}{\Delta x_p q_p^B} = t_p^f \frac{(\Delta x_p^A)^2}{\Delta x_p q_p^A} = t_p^f \Delta x_p^A. \quad (60)$$

4.2. Multi-Variable Case

In the multi-variable case ($m \geq n > 1$) with $(n+1)$ interpolation base points $A_p \begin{pmatrix} x_p^A & f_p^A \end{pmatrix}$ and $B_{p,k} \begin{pmatrix} x_{p,k}^B & f_{p,k}^B \end{pmatrix}$ ($k = 1, \dots, n$), the basic equations of the Secant method (Equations (41) and (44)) are modified as

$$T_p^F \Delta F_p q_p^B = -f_p^A \quad (61)$$

and

$$\Delta x_p^A = T_p^X \Delta X_p q_p^B. \quad (62)$$

Then, a vector-based equation can be formulated, as in case of the traditional Secant method (see Equation (35)), in the following form:

$$\begin{bmatrix} x - \Delta x \\ z(x) \end{bmatrix} = \begin{bmatrix} x^A \\ f^A \end{bmatrix} + \begin{bmatrix} T^X & \mathbf{0} \\ \mathbf{0} & T^F \end{bmatrix} \begin{bmatrix} \Delta X \\ \Delta F \end{bmatrix} q^B, \quad (63)$$

where ΔX and ΔF are defined in (36) and (37), $z(x)$ is a function with zero at x_{p+1}^B , and the diagonal transformation matrix in the p^{th} iteration is

$$T_p = \begin{bmatrix} T_p^X & \mathbf{0} \\ \mathbf{0} & T_p^F \end{bmatrix}, \quad (64)$$

with T_p^X and T_p^F sub-diagonals, where

$$T_p^X = \text{diag}(t_{p,i}^X) = \text{diag}\left(\frac{x_{p+1,i}^B - x_{p+1,i}^A}{x_{p+1,i}^A - x_{p,i}^A}\right) = \text{diag}\left(\frac{\Delta x_{p+1,i}}{\Delta x_{p,i}^A}\right) \quad (65)$$

$$T_p^F = \text{diag}(t_{p,j}^F) = \text{diag}\left(\frac{f_{p+1,j}^B - f_{p+1,j}^A}{f_{p+1,j}^A - f_{p,j}^A}\right), \quad (66)$$

and T_p^F is approximated with the assumption $f(x) \simeq y_p(x_{p+1}^A) \simeq z_p(x_{p+1}^B)$ and according to the conditions $y_p(x_{p+1}^A) = 0$ and $z_p(x_{p+1}^B) = 0$ as

$$T_p^F \simeq \text{diag}\left(\frac{z_{p,j}(x_{p+1}^B) - f_{p+1,j}^A}{y_{p,j}(x_{p+1}^A) - f_{p,j}^A}\right) = \text{diag}\left(\frac{f_{p+1,j}^A}{f_{p,j}^A}\right) \quad (67)$$

($i = 1, \dots, n$), ($j = 1, \dots, m$), where $f_{p,j}^A \neq 0$. The vector of T-Secant multipliers

$$q_p^B = -\Delta F_p^+ (T_p^F)^{-1} f_p^A = -\left[\sum_{j=1}^m \left(\Delta F_{p,i,j} + \frac{f_{p,j}^A}{t_{p,j}^F} \right) \right] \quad (68)$$

can be determined from Equation (61), where $[\cdot]^+$ stands for the pseudo-inverse (ΔF_p^+ has already been calculated when q_p^A was determined from Equation (42)). The i^{th} element of the new approximate x_{p+1}^B can be expressed from the i^{th} row of Equation (62):

$$\Delta x_{p,i}^A = \frac{\Delta x_{p+1,i}}{\Delta x_{p,i}^A} \Delta x_{p,i} q_{p,i}^B = \frac{x_{p+1,i}^B - x_{p+1,i}^A}{\Delta x_{p,i}^A} \Delta x_{p,i} q_{p,i}^B, \quad (69)$$

and the T-Secant approximate x_{p+1}^B can be expressed as

$$x_{p+1,i}^B = x_{p+1,i}^A + \frac{(\Delta x_{p,i}^A)^2}{\Delta x_{p,i} q_{p,i}^B} = x_{p+1,i}^A - \frac{(\Delta x_{p,i}^A)^2}{\sum_{j=1}^m \left(\Delta x_{p,i} \Delta F_{p,i,j}^+ \frac{f_{p,j}^A}{t_{p,j}^F} \right)}, \quad (70)$$

where $\Delta x_{p,i} \neq 0$ and $q_{p,i}^B \neq 0$ ($i = 1, \dots, n$). Then, the next iteration continues with the new trial increment vector (iteration stepsize)

$$\Delta x_{p+1} = x_{p+1}^B - x_{p+1}^A \quad (71)$$

and with $n + 1$ new interpolation base points $A_{p+1} \left(\begin{matrix} x_{p+1}^A & f_{p+1}^A \end{matrix} \right) B_{k,p+1} \left(\begin{matrix} x_{k,p+1}^B & f_{k,p+1}^B \end{matrix} \right)$ ($k = 1, \dots, n$). Figure 1 shows the formulation of a set of new base vectors $x_{k,p+1}^B$ from x_{p+1}^A and x_{p+1}^B in the $n = 3$ case.

Let the ratio μ_i of constants $q_{p,i}^A$ and $q_{p,i}^B$ be introduced as

$$\mu_i = \frac{q_{p,i}^A}{q_{p,i}^B}. \quad (72)$$

Then, it follows from Equations (42), (44), (68), (70), and (71) that the i^{th} element of the new trial increment vector is

$$\Delta x_{p+1,i} = \frac{(\Delta x_{p,i}^A)^2}{\Delta x_{p,i} q_{p,i}^B} = \mu_i \frac{(\Delta x_{p,i}^A)^2}{\Delta x_{p,i} q_{p,i}^A} = \mu_i \Delta x_{p,i}^A. \quad (73)$$

The basic equations in single-variable and multi-variable cases are summarized in Table 1.

Table 1. Summary of the basic equations (single- and multi-variable cases).

	Single-Variable ($m = n = 1$)	Multi-Variable ($m \geq n > 1$)	Equations
1	x_p^A	x_p^A	
2	x_p^B	$x_{p,k}^B = x_p^A + \Delta x_{p,k} d^k$	(32)
3	$\Delta x_p = x_p^B - x_p^A$	$\Delta X_p = [x_{p,k}^B - x_p^A] = \text{diag}(\Delta x_{p,i})$	(36), (38)
4	$\Delta f_p = f_p^B - f_p^A$	$\Delta F_p = \begin{bmatrix} \Delta f_{p,1,1} & \cdots & \Delta f_{p,n,1} \\ \vdots & \ddots & \vdots \\ \Delta f_{p,1,m} & \cdots & \Delta f_{p,n,m} \end{bmatrix}$	(34), (37), (39)
5	$\Delta f_p q_p^A = -f_p^A$	$\Delta F_p q_p^A = -f_p^A$	(49), (41)
6	$q_p^A = -\frac{f_p^A}{\Delta f_p}$	$q_p^A = -\Delta F_p^+ f_p^A$	(47), (42)
7	$x_{p+1}^A = x_p^A + \Delta x_p q_p^A$	$x_{p+1}^A = x_p^A + \Delta X_p q_p^A$	(48), (46)
8	$\Delta x_p^A = x_{p+1}^A - x_p^A$	$\Delta x_p^A = x_{p+1}^A - x_p^A$	(54)

Table 1. Cont.

	Single-Variable ($m = n = 1$)	Multi-Variable ($m \geq n > 1$)	Equations
9	$t_p^f = \frac{f_{p+1}^A}{f_p^A}$	$T_p^F = \text{diag} \left(\frac{f_{p+1,j}^A}{f_{p,j}^A} \right)$	(50), (67)
10	$t_p^f \Delta f_p q_p^B = -f_p^A$	$T_p^F \Delta F_p q_p^B = -f_p^A$	(51), (61)
11	$q_p^B = \frac{q_p^A}{t_p^f} = -\frac{(f_p^A)^2}{f_{p+1}^A \Delta f_p}$	$q_p^B = -\Delta F_p^+ (T_p^F)^{-1} f_p^A$	(52), (68)
12	$t_p^x = \frac{x_{p+1}^B - x_{p+1}^A}{x_{p+1}^A - x_p^A} = \frac{\Delta x_{p+1}}{\Delta x_p^A}$	$T_p^X = \text{diag} \left(\frac{\Delta x_{p+1,i}}{\Delta x_{p,i}^A} \right)$	(55), (65)
13	$\Delta x_p^A = t_p^x \Delta x_p q_p^B$	$\Delta x_p^A = T_p^X \Delta X_p q_p^B$	(56), (62)
14	$x_{p+1}^B = x_{p+1}^A + \frac{(\Delta x_p^A)^2}{\Delta x_p q_p^B}$	$x_{p+1,i}^B = x_{p+1,i}^A + \frac{(\Delta x_{p,i}^A)^2}{\Delta x_{p,i} q_{p,i}^B}$	(59), (70)
15	$\Delta x_{p+1} = t_p^f \Delta x_p^A$	$\Delta x_{p+1,i} = \mu_i \Delta x_{p,i}^A$	(60), (73)

5. Geometry

5.1. Single-Variable Case

The T-Secant procedure has been worked out for solving multi-variable problems. It can also be applied for solving single-variable ones, however. The geometrical representation of the latter gives a good view with which to explain the mechanism of the procedure.

Find the scalar root x^* of a nonlinear function $x \rightarrow f(x)$, where $x \in \mathbb{R}^1$ and $f : \mathbb{R}^1 \rightarrow \mathbb{R}^1$. Let the function $f(x)$ be linearly interpolated through initial base points $A_p(x_p^A, f_p^A)$ and $B_p(x_p^B, f_p^B)$, providing a “secant” line $y_p(x)$ as shown on Figure 2, where $f_p^A = f(x_p^A)$ and $f_p^B = f(x_p^B)$ are the corresponding function values. An arbitrary point of the Secant $y_p(x)$ can be expressed as

$$\begin{bmatrix} x \\ y_p(x) \end{bmatrix} = \begin{bmatrix} x_p^A \\ f_p^A \end{bmatrix} + \begin{bmatrix} \Delta x_p \\ \Delta f_p \end{bmatrix} q_p^A, \quad (74)$$

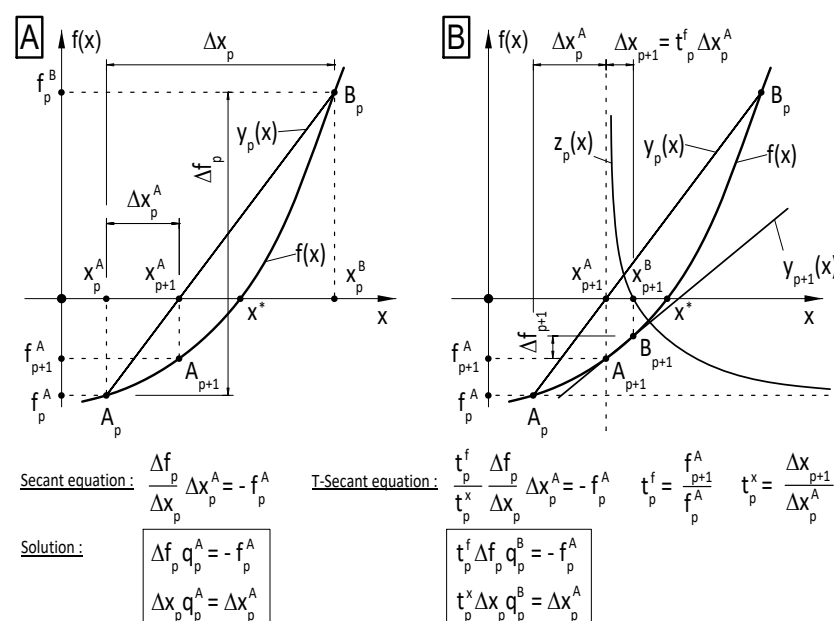


Figure 2. Geometrical representation of the Secant method in a single-variable case: (A) classic Secant method; (B) T-Secant modification.

where q_p^A is a scalar multiplier. Let a new approximate x_{p+1}^A be the root of the Secant $y_p(x)$ and let

$$\Delta x_p^A = x_{p+1}^A - x_p^A \quad (75)$$

be the iteration stepsize. It follows from condition

$$y_p(x_{p+1}^A) = 0 \quad (76)$$

and from the second row of Equation (74) that

$$\Delta f_p q_p^A = -f_p^A, \quad (77)$$

and the scalar multiplier can be determined as

$$q_p^A = -\frac{f_p^A}{\Delta f_p}. \quad (78)$$

From the first row of Equation (74), the iteration stepsize is given as

$$\Delta x_p^A = \Delta x_p q_p^A, \quad (79)$$

and the new approximate can be expressed as

$$x_{p+1}^A = x_p^A + \Delta x_p^A. \quad (80)$$

A new base point $A_{p+1}(x_{p+1}^A, f_{p+1}^A)$ (see Figure 2) can then be determined for the next iteration. Two out of the three available base points (A_p B_p A_{p+1}) are used for the next iteration by omitting either A_p or B_p in the case of the traditional secant method. The decision is not obvious and it may cause the iteration to unstable and/or not converge to the solution. Instead, an additional new approximate x_{p+1}^B is determined by the T-Secant procedure as a root of the function $z_p(x)$ near the first Secant approximate x_{p+1}^A , and iteration continues with new base points $A_{p+1}(x_{p+1}^A, f_{p+1}^A)$ and $B_{p+1}(x_{p+1}^B, f_{p+1}^B)$. An arbitrary point of the function $z_p(x)$ can be expressed as

$$\begin{bmatrix} x - \Delta x \\ z_p(x) \end{bmatrix} = \begin{bmatrix} x_p^A \\ f_p^A \end{bmatrix} + \begin{bmatrix} t^x & 0 \\ 0 & t^f \end{bmatrix} \begin{bmatrix} \Delta x_p \\ \Delta f_p \end{bmatrix} q_p^B, \quad (81)$$

where the transformation scalars for Δx_p and Δf_p at $x = x_p^B$ are

$$t_p^x = \frac{\Delta x_{p+1}}{\Delta x_p^A} = \frac{x_{p+1}^B - x_{p+1}^A}{x_{p+1}^A - x_p^A} \text{ and } t_p^f = \frac{f_{p+1}^A}{f_p^A}. \quad (82)$$

Then, it follows from condition

$$z_p(x_{p+1}^B) = 0 \quad (83)$$

and from the second row of Equation (81) that

$$t_p^f \Delta f_p q_p^B = -f_p^A \quad (84)$$

and

$$q_p^B = -\frac{f_p^A}{t_p^f \Delta f_p} = -\frac{(f_p^A)^2}{f_{p+1}^A (f_p^B - f_p^A)}. \quad (85)$$

The new approximate x_{p+1}^B can then be expressed from the first row of Equation (81) as

$$x_{p+1}^B = x_{p+1}^A + \frac{(\Delta x_p^A)^2}{\Delta x_p q_p^B}. \quad (86)$$

The new base point $B_{p+1}(x_{p+1}^B, f_{p+1}^B)$ (see Figure 2) can then be determined. Interpolation base points A_{p+1} and B_{p+1} are used for the next iteration. The scalar multiplier q_p^B can be expressed from Equation (86) as

$$q_p^B = \frac{(x_{p+1}^A - x_p^A)^2}{(x_p^B - x_p^A)(x_{p+1}^B - x_{p+1}^A)}. \quad (87)$$

By substituting it into the second row of Equation (81) and changing x_{p+1}^B to x , it turns into a hyperbolic function

$$z_p(x) = \frac{a_p}{x - x_{p+1}^A} + f_p^A \quad (88)$$

with vertical and horizontal asymptotes x_{p+1}^A and f_p^A , where

$$a_p = (x_{p+1}^A - x_p^A)^2 \frac{f_p^B - f_p^A}{x_p^B - x_p^A} \frac{f_{p+1}^A}{f_p^A}, \quad (89)$$

and the root x_{p+1}^B of the function $z_p(x)$ will be in the vicinity of x_{p+1}^A in “appropriate distance”, which is regulated by the function value f_{p+1}^A (see Figure 2). This virtue of the T-Secant procedure is that it ensures an automatic mechanism for having base vectors in general positions through the whole iteration process, providing a stable and efficient numerical performance.

5.2. Multi-Variable Case

Find the root x^* of a nonlinear function $x \rightarrow f(x)$, where $x \in \mathbb{R}^n$ and $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Let the function $f(x)$ be linearly interpolated through $n + 1$ base points $A_p(x_p^A, f_p^A)$ and $B_{k,p}(x_{k,p}^B, f_{k,p}^B)$ in the R^{n+m} space ($f(x)$ –space) in the p^{th} iteration as shown on Figure 3, where $k = 1, \dots, n$, given a set of approximates x_p^A and

$$x_{k,p}^B = x_p^A + \Delta x_{k,p} d^k \quad (90)$$

in the \mathbb{R}^n space (x –space) with $k = 1, \dots, n$, where d^k is the k^{th} Cartesian unit vector. Let the expression

$$\Delta F_p q_p^A = \begin{bmatrix} \Delta f_{1,1,p} & \cdots & \Delta f_{n,1,p} \\ \vdots & \vdots & \vdots \\ \Delta f_{1,m,p} & \cdots & \Delta f_{n,m,p} \end{bmatrix} q_p^A \quad (91)$$

represent the linear combination $q_p^A = [q_{p,k}^A]^T$ of n column vectors

$$[\Delta f_{k,j,p}] = [\Delta f_{k,p}] = [f_{k,p}^B - f_{k,p}^A]$$

in the \mathbb{R}^m space (f -space) with $k = 1, \dots, n$ column index, and with $j = 1, \dots, m$ row index, and the expression

$$\Delta \mathbf{X}_p \mathbf{q}_p^A = \begin{bmatrix} \Delta x_{1,1,p} & \cdots & \Delta x_{n,1,p} \\ \vdots & \vdots & \vdots \\ \Delta x_{1,n,p} & \cdots & \Delta x_{n,n,p} \end{bmatrix} \mathbf{q}_p^A \quad (92)$$

represents the same linear combination of n column vectors

$$[\Delta x_{k,j,p}] = [\Delta x_{p,k}] = [\mathbf{x}_{p,k}^B - \mathbf{x}_{p,k}^A], \quad (93)$$

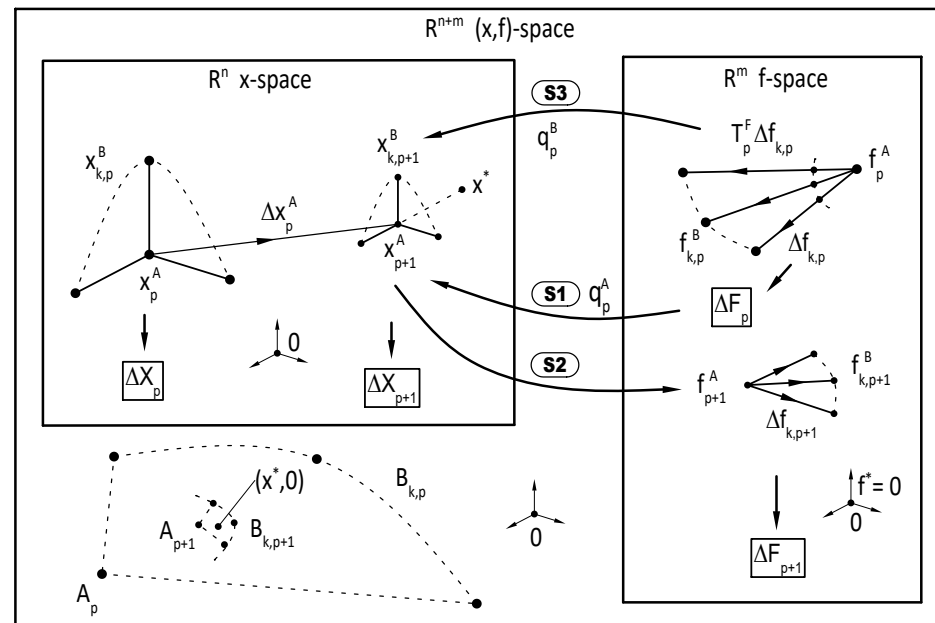


Figure 3. Vector space description of the T-Secant method in the multi-variable case ($k = 1, \dots, n$).

with $k = 1, \dots, n$ column index, and with $j = 1, \dots, n$ row index. The linear combination \mathbf{q}_p^A is determined from Equation (42) in step S1 (see Figure 3), providing a new approximate

$$\mathbf{x}_{p+1}^A = [\mathbf{x}_{p+1,k}^A] = \mathbf{x}_p^A + \Delta \mathbf{x}_p^A, \quad (94)$$

for the solution \mathbf{x}^* , and the corresponding \mathbf{f}_{p+1}^A vector is also determined in step S2 (see Figure 3). The column vectors $\Delta \mathbf{f}_{k,p}$ of $\Delta \mathbf{F}_p$ are then modified by a non-uniform scaling transformation,

$$\mathbf{T}_p^F = \text{diag} \left(\frac{f_{p+1,j}^A}{f_{p,j}^A} \right), \quad (95)$$

and a new linear combination $\mathbf{q}_p^B = [\mathbf{q}_{p,k}^B]^T$ is determined from Equation (68) in step S3 (see Figure 3), providing a new approximate \mathbf{x}_{p+1}^B for the solution \mathbf{x}^* with elements

$$x_{p+1,k}^B = x_{p+1,k}^A + \frac{(\Delta x_{p,k}^A)^2}{\Delta x_{p,k}^A \mathbf{q}_{p,k}^B}. \quad (96)$$

A new set of $n + 1$ approximates \mathbf{x}_{p+1}^A and

$$\mathbf{x}_{k,p+1}^B = \mathbf{x}_{p+1}^A + \Delta x_{k,p+1} \mathbf{d}^k \quad (97)$$

($k = 1, \dots, n$) can then be generated with iteration stepsize

$$\Delta x_{p+1} = [\Delta x_{k,p+1}] = x_{p+1}^B - x_{p+1}^A \quad (98)$$

for the next iteration.

5.3. Single-Variable Example

An example is given with function $x \rightarrow f(x)$, where $x \in \mathbb{R}^1$: $f: \mathbb{R}^1 \rightarrow \mathbb{R}^1$ and

$$f(x) = x^3 - 2x - 5 \quad (99)$$

with root $x^* \cong 2.0945514815423 \dots$. Figure 4 and Table 2 summarize the results of the first two iterations (left: x_1^A is the zero of $y_0(x)$, x_1^B is the zero of $z_0(x)$; right: x_2^A is the zero of $y_1(x)$, x_2^B is the zero of $z_1(x)$). Iterations were made with initial approximates $x_0^A = 3.0$ and $x_0^B = 1.0$, providing $f_0^A = 16$ ($p = 0$). The first Secant approximate $x_1^A = 1.545 \dots$ is found as the zero of the first Secant $y_0(x)$, and the first T-Secant appropriate $x_1^B = 1.945 \dots$ is found as the zero of the first hyperbola function $z_0(x)$ (Figure 4, left). Iteration then goes on ($p = 1$) with new interpolation base points $x_1^A = 1.545 \dots$ and $x_1^B = 1.945 \dots$ providing $f_1^A = -4.3997 \dots$, and new approximates $x_2^A = 2.158 \dots$ and $x_2^B = 2.0556 \dots$ are found as the zeros of the second Secant and the second hyperbola function $y_1(x)$ and $z_1(x)$, respectively (Figure 4, right). The next iteration ($p = 2$) will then continue with interpolation base point $x_2^A = 2.158 \dots$ and $x_2^B = 2.0556 \dots$ and with $f_2^A = 0.7367 \dots$. The iterated values of f_p^A , x_p^A and x_p^B are also indicated in the diagrams. Further diagrams for this example are shown in Section 7.3.

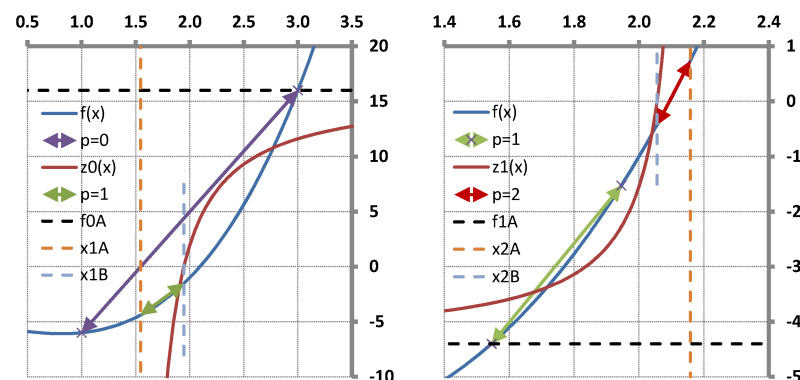


Figure 4. T-secant iterations with test function (99) with initial approximates $x_0^A = 3.0$ and $x_0^B = 1.0$ (Left: x_1^A is the root of $y_0(x)$, x_1^B is the root of $z_0(x)$; Right: x_2^A is the root of $y_1(x)$, x_2^B is the root of $z_1(x)$) (see also Table 2).

Table 2. T-Secant iteration results with $x_0^A = 3.0$ and $x_0^B = 1.0$ (see also Figure 4 for $p = 0, 1, 2$).

p	0	1	2	3	4
x_p^A	3.0	1.545	2.158	2.093	2.09455149745
Δx_p	−2.0	0.400	−0.103	0.0014	1.6×10^{-8}
f_p^A	16.0	−4.400	0.737	−0.015	1.8×10^{-7}
q_p^A	0.727	1.532	0.634	1.015	0.999
x_{p+1}^A	1.545	2.158	2.093	2.09455150	2.0945514815423
e_{p+1}^A	−0.549	0.064	0.0014	1.6×10^{-8}	2.7×10^{-14}
f_{p+1}^A	−4.400	0.737	−0.015	1.8×10^{-7}	
t_p^F	−0.275	−0.167	−0.021	1.2×10^{-5}	
q_p^B	−2.645	−9.149	−30.329	−88077	
x_{p+1}^B	1.945	2.056	2.09453	2.09455148153	

6. General Formulations

Re-ordering Equation (44) gives the general equation

$$\Delta F \Delta X^{-1} \Delta x^A = -f^A \quad (100)$$

of the Secant method. The initial trials are constructed according to Equation (32), providing that ΔX is a diagonal matrix with elements $(\Delta x_i) = (x_{i,i}^B - x_i^A)$ ($i = 1, \dots, n$). Let the Jacobean approximate of the secant-method be defined as

$$S = \Delta F \Delta X^{-1} \quad (101)$$

$$S = \begin{bmatrix} \frac{f_{1,1}^B - f_1^A}{x_{1,1}^B - x_1^A} & \dots & \frac{f_{n,1}^B - f_1^A}{x_{n,1}^B - x_1^A} \\ \vdots & \ddots & \vdots \\ \frac{f_{1,m}^B - f_m^A}{x_{1,1}^B - x_1^A} & \dots & \frac{f_{n,m}^B - f_m^A}{x_{n,1}^B - x_1^A} \end{bmatrix} = \begin{bmatrix} \frac{\Delta f_{1,1}}{\Delta x_1} & \dots & \frac{\Delta f_{n,1}}{\Delta x_n} \\ \vdots & \ddots & \vdots \\ \frac{\Delta f_{1,m}}{\Delta x_1} & \dots & \frac{\Delta f_{n,m}}{\Delta x_n} \end{bmatrix} = \begin{bmatrix} \Delta f_{k,j} \\ \Delta x_i \end{bmatrix} \quad (102)$$

($i = 1, \dots, n$), ($j = 1, \dots, m$), ($k = 1, \dots, n$) and

$$S^+ = \Delta X \Delta F^+. \quad (103)$$

Then, Equation (100) simplifies as

$$S \Delta x^A = -f^A \quad (104)$$

and

$$\Delta x^A = -S^+ f^A. \quad (105)$$

The i^{th} element of the new approximate x_{p+1}^A in the p^{th} iteration will then be

$$x_{p+1,i}^A = x_{p,i}^A + \Delta x_{p,i}^A = x_{p,i}^A - \sum_{j=1}^m (S_{p,i,j}^+ f_{p,j}^A) \quad (106)$$

($i = 1, \dots, n$). It follows from the first row of Equation (63) of the T-Secant method, from Equation (61) and from the Definition (103) of S^+ , that the p^{th} iteration stepsize is

$$\Delta x_p^A = -T_p^X S_p^+ (T_p^F)^{-1} f_p^A \quad (107)$$

and

$$T_p^F S_p (T_p^X)^{-1} \Delta x_p^A = -f_p^A. \quad (108)$$

Let the modified Jacobean approximate of the T-Secant method be defined as

$$S_{p,T} = T_p^F S_p (T_p^X)^{-1} \quad (109)$$

$$S_{p,T} = \begin{bmatrix} \frac{f_{p+1,1}^A}{f_{p,1}^A} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \frac{f_{p+1,m}^A}{f_{p,m}^A} \end{bmatrix} \begin{bmatrix} \frac{\Delta f_{p,1,1}}{\Delta x_{p,1}} & \dots & \frac{\Delta f_{p,n,1}}{\Delta x_{p,n}} \\ \vdots & \ddots & \vdots \\ \frac{\Delta f_{p,1,m}}{\Delta x_{p,1}} & \dots & \frac{\Delta f_{p,n,m}}{\Delta x_{p,n}} \end{bmatrix} \begin{bmatrix} \frac{\Delta x_{p,1}^A}{\Delta x_{p+1,1}} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \frac{\Delta x_{p,n}^A}{\Delta x_{p+1,n}} \end{bmatrix} \quad (110)$$

and in condensed form with general matrix elements (without the p index):

$$S_T = T^F S (T^X)^{-1} = \text{diag}(t_j^F) \left[\frac{\Delta f_{k,j}}{\Delta x_i} \right] \text{diag}\left(\frac{1}{t_i^X}\right) = \left[\frac{t_j^F}{t_i^X} \frac{\Delta f_{k,j}}{\Delta x_i} \right] \quad (111)$$

($i = 1, \dots, n$), ($j = 1, \dots, m$), ($k = 1, \dots, n$) and

$$S_T^+ = T^X S^+ (T^F)^{-1}. \quad (112)$$

Equations (107) and (108) then can be re-written as

$$\Delta x^A = -S_T^+ f^A \quad (113)$$

and

$$S_T \Delta x^A = -f^A \quad (114)$$

in a similar form as in case of the traditional Secant method (Equations (105) and (104)). The i^{th} element $x_{p+1,i}^B$ of the second new approximate x_{p+1}^B in the p^{th} iteration will then be

$$x_{p+1,i}^B = x_{p+1,i}^A + \Delta x_{p+1,i} = x_{p+1,i}^A - \frac{(\Delta x_{p,i}^A)^2}{\sum_{j=1}^m \left(S_{p,i,j}^+ \frac{f_{p,j}^A}{t_{p,j}^F} \right)}, \quad (115)$$

where $t_{p,j}^F \neq 0$, ($j = 1, \dots, m$) and ($i = 1, \dots, n$). Note that the T-Secant modification of the Jacobean approximate (102) is made with multipliers

$$t_{p,j}^F = \frac{f_{p+1,j}^A}{f_{p,j}^A} \quad \text{and} \quad t_{p,i}^X = \frac{\Delta x_{p+1,i}}{\Delta x_{p,i}^A} \quad (116)$$

to the difference quantities $\Delta f_{p,k,j}$ and $\Delta x_{p,i}$. The basic equations of the Secant method and the T-Secant method are summarized in Table 3; rows 1–4 are the elements (matrix T) of the basic equations, rows 5–6 are the explicit basic equations, row 7 depicts the Jacobean-type matrices, and rows 8–9 are the general formulations of the basic equations.

Table 3. Summary of the multi-variable Secant and T-Secant methods basic equations.

Secant Method	T-Secant Method	Equations
1 $\begin{bmatrix} \Delta X_p \\ \Delta F_p \end{bmatrix} = \begin{bmatrix} x_{p,k}^B - x_p^A \\ f_{p,k}^B - f_p^A \end{bmatrix} = \begin{bmatrix} \text{diag}(\Delta x_{p,i}) \\ \Delta f_{p,k,j} \end{bmatrix}$		(36), (37)
2	$T_p = \begin{bmatrix} T_p^X & \mathbf{0} \\ \mathbf{0} & T_p^F \end{bmatrix}$	(64)
3	$T_p^X = \text{diag}(t_{p,i}^X) = \text{diag}\left(\frac{\Delta x_{p+1,i}}{\Delta x_{p,i}^A}\right)$	(65)
4	$T_p^F = \text{diag}(t_{p,j}^F) \cong \text{diag}\left(\frac{f_{p+1,j}^A}{f_{p,j}^A}\right)$	(66), (67)
5 $\Delta F q^A = -f^A$	$T^F \Delta F q^B = -f^A$	(41), (61)
6 $\Delta F \Delta X^{-1} \Delta x^A = -f^A$	$T^F \Delta F \Delta X^{-1} (T^X)^{-1} \Delta x^A = -f^A$	(100), (108)
7 $S = \Delta F \Delta X^{-1} = \left[\frac{\Delta f_{k,j}}{\Delta x_i} \right]$	$S_T = T^F S (T^X)^{-1} = \left[\frac{t_j^F}{t_i^X} \frac{\Delta f_{k,j}}{\Delta x_i} \right]$	(102), (111)
8 $S \Delta x^A = -f^A$	$S_T \Delta x^A = -f^A$	(104), (114)
9 $\Delta x^A = -S^+ f^A$	$\Delta x^A = -S_T^+ f^A$	(105), (113)

7. Convergence

7.1. Single-Variable Case

As was shown in Section 4 (Equation (60)), the p^{th} iteration stepsize of the second new approximate x_{p+1}^B is

$$\Delta x_{p+1} = t_p^f \Delta x_p^A. \quad (117)$$

The Secant method is super-linear convergent, so the new approximate x_{p+1}^A is expected to be a much better approximate to the solution x^* than the previous one (x_p^A). Thus,

$$|f_{p+1}^A| \ll |f_p^A| \quad (118)$$

and

$$|t_p^f| = \left| \frac{f_{p+1}^A}{f_p^A} \right| \ll 1 \quad (119)$$

is expected to be a “small positive number”. It means that the T-Secant approximate x_{p+1}^B will always be in the vicinity of the classic Secant approximate x_{p+1}^A , and the approximate errors of the new approximates will be of a similar order, providing that the solution x^* will be evenly surrounded by the two new trial approximates x_{p+1}^A and x_{p+1}^B .

7.2. Convergence Rate

It is well known that the single-variable Secant method has asymptotic convergence for sufficiently good initial approximates x^A and x^B if $f'(x)$ does not vanish in $x \in [x^A, x^B]$ and $f''(x)$ is continuous at least in a neighborhood of the zero x^* . The super-linear convergence property has been proved in different ways and it is known that the order of convergence $\alpha^S = (1 + \sqrt{5})/2$ with asymptotic error constant

$$C = \left(\frac{1}{2} \left| \frac{f''(\xi)}{f'(\xi)} \right| \right)^{\frac{1}{\alpha}}. \quad (120)$$

The order of convergence of the T-Secant method is determined in this section. Let p be the iteration counter and the approximate error be defined in the p^{th} iteration as

$$e_p = x_p - x^*. \quad (121)$$

It follows from Equation (48) and from Definition (121) that the error e_{p+1}^A of the new Secant approximate x_{p+1}^A can be expressed as

$$e_{p+1}^A = \frac{e_p^A f_p^B - e_p^B f_p^A}{f_p^B - f_p^A} = \frac{x_p^B - x_p^A}{f_p^B - f_p^A} \frac{f_p^B / e_p^B - f_p^A / e_p^A}{x_p^B - x_p^A} e_p^A e_p^B. \quad (122)$$

It follows from the mean value theorem that the first factor of the right side of Equation (122) can be replaced with $1/f'(\eta_p)$, where $\eta_p \in (x_p^A, x_p^B)$, if $f(x)$ is continuously differentiable on (x_p^A, x_p^B) and $f'(\eta_p) \neq 0$. Let the function $f(x)$ be approximated around the root x^* by a second order Taylor series expansion as

$$f_p = f(e_p + x^*) = f(x^*) + e_p f'(x^*) + \frac{1}{2} (e_p)^2 f''(\xi_p), \quad (123)$$

where $\xi_p \in (x_p^A, x_p^B, x^*)$ in the remainder term. Since $f(x^*) = 0$, it follows from Equation (123) that

$$\frac{f_p}{e_p} = f'(x^*) + \frac{1}{2} f''(\xi_p) e_p. \quad (124)$$

Substituting this expression to Equation (122), and since $e_p^B - e_p^A = x_p^B - x_p^A$, we obtain

$$e_{p+1}^A = \frac{1}{2} \frac{f''(\xi_p)}{f'(\eta_p)} \frac{e_p^B - e_p^A}{x_p^B - x_p^A} e_p^A e_p^B = C_p e_p^A e_p^B \quad (125)$$

and

$$C_p = \frac{1}{2} \frac{f''(\xi_p)}{f'(\eta_p)}. \quad (126)$$

If the series $\{x_p^A\}$ converges to x^* , then ξ_p and $\eta_p \rightarrow x^*$ with increasing iteration counter p , and

$$C_p \rightarrow \frac{1}{2} \frac{f''(x^*)}{f'(x^*)} = \text{constant}. \quad (127)$$

It follows from Equation (59) with Definition (121) and from the mean value theorem (with $\eta_{p-1} \in (x_{p-1}^A, x_{p-1}^B)$), if $f(x)$ is continuously differentiable on (x_{p-1}^A, x_{p-1}^B) , that

$$x_p^B = x_p^A - \left(\frac{x_p^A - x_{p-1}^A}{f_{p-1}} \right)^2 f'(\eta_{p-1}) f_p^A, \quad (128)$$

and the error e_p^B of the T-Secant approximate x_p^B can be expressed as

$$e_p^B = e_p^A - \left(\frac{e_p^A - e_{p-1}^A}{f_{p-1}^A} \right)^2 f'(\eta_{p-1}) f_p^A. \quad (129)$$

With the Taylor-series expansion (123) for f_{p-1}^A and f_p^A , where $\xi_{p-1} \in (x_{p-1}^A, x_{p-1}^B, x^*)$ and $\xi_p \in (x_p^A, x_p^B, x^*)$ in the remainder term, we obtain

$$e_p^B = e_p^A - e_p^A \left(\frac{e_p^A - e_{p-1}^A}{e_{p-1}^A} \right)^2 \gamma_p, \quad (130)$$

where

$$\gamma_p = \frac{\frac{f'(x^*)}{f'(\eta_{p-1})} + \frac{1}{2} \frac{f''(\xi_p)}{f'(\eta_{p-1})} e_p^A}{\left(\frac{f'(x^*)}{f'(\eta_{p-1})} + \frac{1}{2} \frac{f''(\xi_{p-1})}{f'(\eta_{p-1})} e_{p-1}^A \right)^2} \quad (131)$$

and $f'(\eta_{p-1}) \neq 0$. If the series $\{x_p^A\}$ converges to x^* , then, with increasing iteration counter p , $\xi_p, \xi_{p-1}, \eta_{p-1} \rightarrow x^*$, and $e_p^A, e_{p-1}^A \rightarrow 0$, it implies that

$$\frac{f'(x^*)}{f'(\eta_{p-1})} \rightarrow \frac{f'(x^*)}{f'(x^*)} = 1 \quad (132)$$

and $\gamma_p \rightarrow 1$. Substituting e_p^B (Equation (130)) into Equation (125) gives

$$e_{p+1}^A = C_p e_p^A \left(e_p^A - e_p^A \left(\frac{e_p^A - e_{p-1}^A}{e_{p-1}^A} \right)^2 \gamma_p \right), \quad (133)$$

and re-arranging

$$e_{p+1}^A = C_p e_p^A \left(\gamma_p (e_p^A)^2 \frac{2e_{p-1}^A - e_p^A}{(e_{p-1}^A)^2} + (1 - \gamma_p) e_p^A \right) \quad (134)$$

with $\{x_p^A\}$ converges to x^* , $\gamma_p \rightarrow 1$, and the above equation simplifies as

$$e_{p+1}^A = C_p e_p^A (e_p^A)^2 \frac{2e_{p-1}^A - e_p^A}{(e_{p-1}^A)^2}. \quad (135)$$

This means that e_{p+1}^A depends on e_p^A and e_{p-1}^A , and by assuming an asymptotic convergence, a power law relationship

$$|e_{p+1}^A| = C |e_p^A|^\alpha \quad (136)$$

can be established, where C is the asymptotic error constant and α is the convergence rate, also called the “convergence order” of the iterative method. It also follows from Equation (136), that

$$|e_p^A| = C |e_{p-1}^A|^\alpha \quad (137)$$

and

$$|e_{p-1}^A| = \left(\frac{|e_p^A|}{C} \right)^{\frac{1}{\alpha}}. \quad (138)$$

Let $E = |e_p^A|$ be introduced for simplifying purposes; then, it follows from Equations (133), (136)–(138) that

$$E^\alpha = \frac{C_p}{C} E^3 \frac{2 \left(\frac{E}{C} \right)^{\frac{1}{\alpha}} - E}{\left(\frac{E}{C} \right)^{\frac{2}{\alpha}}}, \quad (139)$$

where C_p and C are constants and, if the series $\{x_p^A\}$ converges to x^* , with increasing iteration, counter p , $E \rightarrow 0^+$. Taking the logarithms of both sides of Equation (139) and dividing by $\ln E$ gives

$$\alpha = \frac{\ln \frac{C_p}{C}}{\ln E} + 3 - \frac{2}{\alpha} \cdot \frac{\ln \left(\frac{E}{C} \right)}{\ln E} + \frac{\ln \left(2 \left(\frac{E}{C} \right)^{\frac{1}{\alpha}} - E \right)}{\ln E}. \quad (140)$$

If $\{x_p^A\}$ series converges to x^* , then, with increasing iteration, they counter p , $E \rightarrow 0^+$, $\ln E \rightarrow -\infty$ and

$$\lim_{E \rightarrow 0^+} \frac{\ln \frac{C_p}{C}}{\ln E} = 0 \quad (141)$$

$$\lim_{E \rightarrow 0^+} \ln \left(\frac{E}{C} \right) = \lim_{E \rightarrow 0^+} (\ln E - \ln C) = \ln E \quad (142)$$

$$\lim_{E \rightarrow 0^+} \frac{\ln \left(2 E^{\frac{1}{\alpha}} - E \right)}{\ln E} = \frac{1}{\alpha}, \quad (143)$$

and Equation (140) simplifies as

$$\alpha - 3 + \frac{1}{\alpha} = 0, \quad (144)$$

with root (convergence rate of the T-Secant method):

$$\alpha^{TS} = \frac{3 + \sqrt{5}}{2} \cong 2.618033988 \dots = \alpha^S + 1 = \varphi^2, \quad (145)$$

where $\alpha^S = \varphi \cong 1.618033988 \dots$ is the convergence rate of the traditional Secant method, and φ is the well-known golden ratio. It follows from Equation (140) that the actual values of α^* of α^{TS} depend on the approximate error $E = |e^A|$. Convergence rates $\alpha^*(E)$ were determined for different E values and are shown in Figure 5. The upper bound $\alpha^{TS} = \alpha^S + 1 = 2.618 \dots$ at $E \rightarrow 0^+$ is also indicated (horizontal dashed red line).

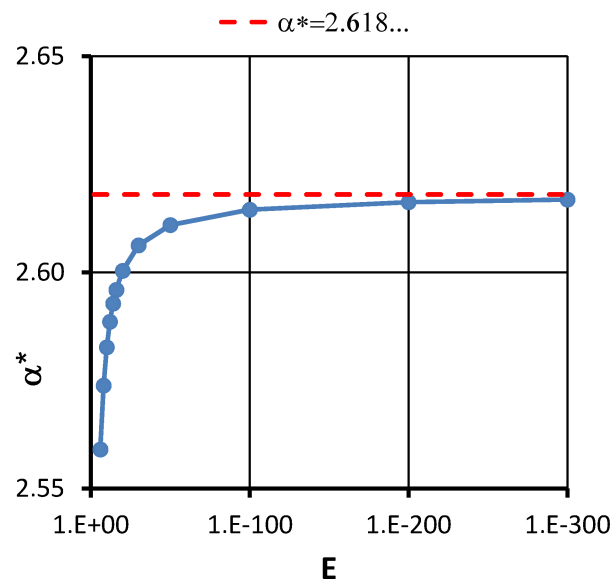


Figure 5. α^* convergence rate variation with decreasing $E \rightarrow 0^+$ (dashed red lines indicate $\alpha = \alpha^S + 1 \cong 2.618$ level, where $\alpha^S \cong 1.618$ is the convergence rate of the traditional Secant method).

7.3. Single-Variable Example

An example is given for demonstration purposes with a single-variable test function (99) with root $x^* \cong 2.09455 \dots$. Iterations were made with initial approximates $x_0^A = 3.5$ and $x_0^B = 2.5$, and the convergence rates α^S , α^N , and α^{TS} were determined for the traditional Secant method (Table 4, Figure 6), for the Newton–Raphson method (Table 5, Figure 7), and for the T-Secant method (Table 6, Figure 8), respectively. The cumulative number of function values (N_f) and derivative function values ($N_{f'}$) calculations are also indicated in the tables. Calculated convergence rates agree well with theoretical values $\alpha^S = 1.62 \dots$, $\alpha^N = 2.0$ and $\alpha^{TS} = 2.62 \dots$. Figure 9 summarizes the results of iterations with three different methods (Secant, Newton–Raphson, and T-Secant). Two groups of graphs show the absolute approximate error $|e_p^A|$ decrease and the calculated convergence rates α for the three compared methods. Results demonstrate that the convergence rate of the T-Secant method is higher than the convergence rate of the Newton–Raphson method.

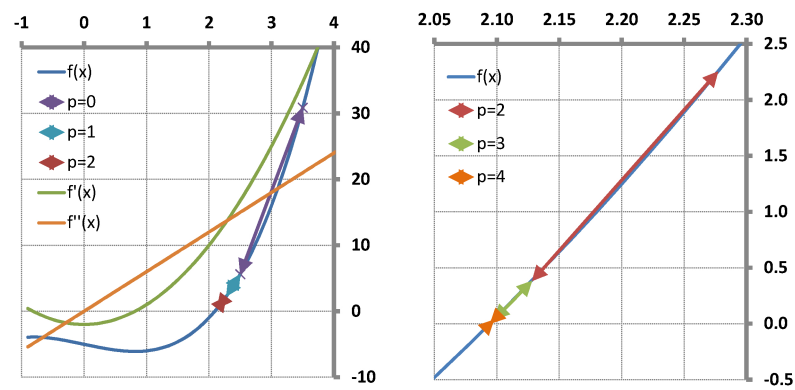


Figure 6. Secant iteration with test function (99) with initial approximates $x_0^A = 3.5$ and $x_0^B = 2.5$ (Left: $p = 0, 1, 2$; Right: $p = 2, 3, 4$ (see data in Table 4)).

Table 4. Secant method iteration and computed convergence rate, α^S (see Figure 6).

p	x_p^A	x_p^B	x_{p+1}^A	$ e_{p+1}^A $	α^S	N_f
0	3.5	2.5	2.2772	1.8×10^{-1}		2
1	2.5	2.2772	2.1282	3.4×10^{-2}		3
2	2.2772	2.1282	2.0977	3.2×10^{-3}	0.64	4
3	2.1282	2.0977	2.094611	5.9×10^{-5}	2.12	5
4	2.0977	2.094611	2.094552	1.1×10^{-7}	1.39	6
5	2.094611	2.09455216	2.09455148	3.6×10^{-12}	1.69	7
6	2.0945516	2.09455148	2.09455148154233	2.7×10^{-14}	1.59	8
7	2.09455148	2.09455148154233	2.09455148154233	2.7×10^{-14}	1.63	9

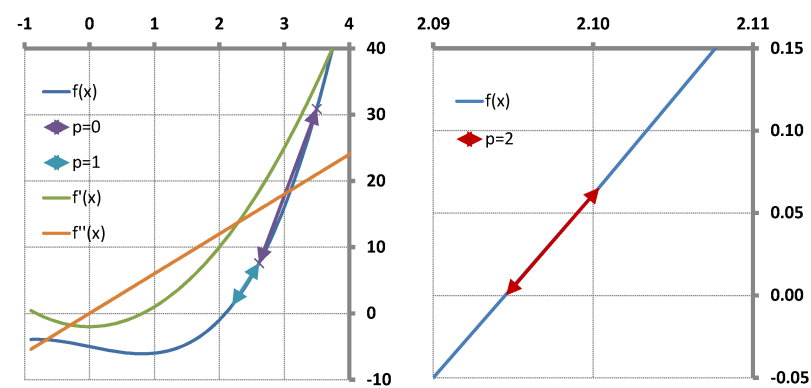


Figure 7. Newton iteration with test function (99) with initial approximate $x_0^A = 3.5$ (Left: $p = 0, 1$; Right: $p = 2$ (see data in Table 5)).

Table 5. Newton method iteration and computed convergence rate, α^N (see Figure 7).

p	x_p^A	x_{p+1}^A	$ e_{p+1}^A $	α^N	N_f	$N_{f'}$
0	3.5	2.61	5.2×10^{-1}		1	1
1	2.61	2.200	1.1×10^{-1}		2	2
2	2.200	2.10037	5.8×10^{-3}	1.58	3	3
3	2.10037	2.09457	1.9×10^{-5}	1.82	4	4
4	2.09457	2.09455148	2.0×10^{-10}	1.97	5	5
5	2.09455148	2.09455148154233	2.7×10^{-14}	2.00	6	6

Table 6. T-Secant method iteration and computed convergence rate, α^{TS} (see Figure 8).

p	x_p^A	x_p^B	x_{p+1}^A	$ e_{p+1}^A $	α^{TS}	N_f
0	3.5	2.5	2.28	1.8×10^{-1}		2
1	2.28	2.1879	2.1032	8.6×10^{-3}		4
2	2.1032	2.0957112	2.0945571	5.6×10^{-6}	1.50	6
3	2.0945571	2.09455151	2.09455148154242	1.2×10^{-13}	2.41	8
4	2.09455148154242	2.09455148154233	2.09455148154233	2.7×10^{-14}	2.40	10

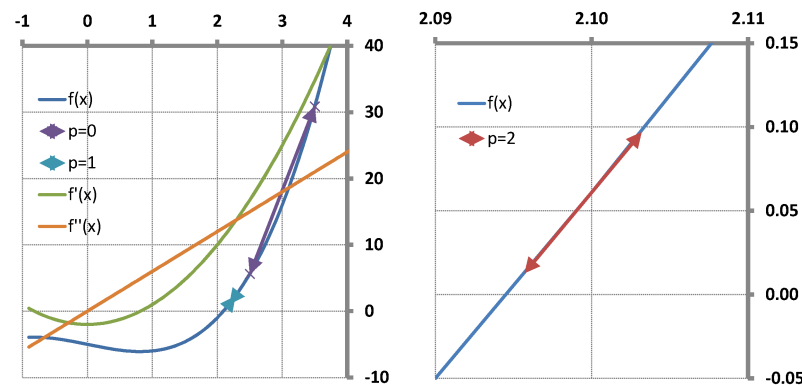


Figure 8. T-Secant iteration with test function (99) with initial approximates $x_0^A = 3.5$ and $x_0^B = 2.5$ (Left: $p = 0$ (with interpolation base points $A_0 B_0$) and $p = 1$ ($A_1 B_1$); Right: $p = 2$ ($A_2 B_2$) (see data in Table 6)).

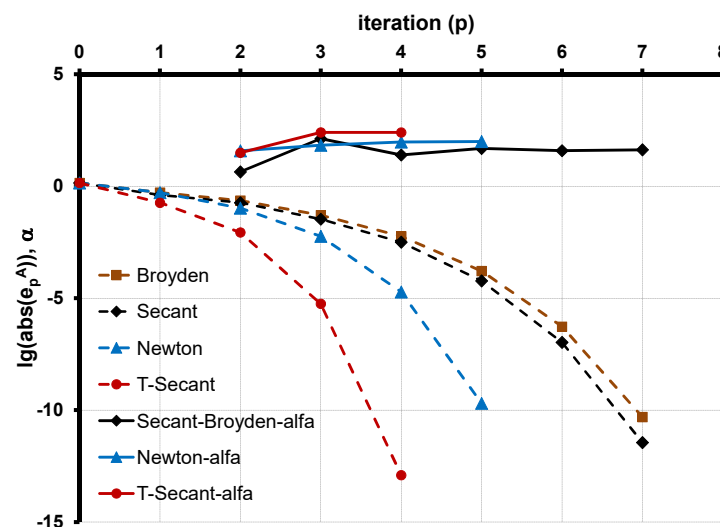


Figure 9. Absolute approximate error $|e_p^A|$ decrease (dashed lines) and computed convergence rates (α) (solid lines) of different methods (Broyden (brown line), Secant (black lines), Newton–Raphson (blue lines), and T-Secant (red lines) method).

7.4. Multi-Variable Convergence

Matrix S (see Equation (102)) corresponds to a divided difference approximation of the Jacobian. It is known (e.g., from Dennis-Schnabel [13]) that these values give a second-order approximation of the derivative in the midpoint. When considering Newton's iteration, it is assumed that the Jacobian has inverted in a neighborhood of x^* . If that condition holds, then there is a chance that the approximate Jacobian has also inverted in the same neighborhood.

It is known that the Secant method is locally q-super-linear convergent, so the new approximate \mathbf{x}_{p+1}^A is expected to be a much better approximate to the solution \mathbf{x}^* than the previous approximate \mathbf{x}_p^A . Thus,

$$\|\mathbf{f}_{p+1}^A\| \lll \| \mathbf{f}_p^A \| \quad (146)$$

and the diagonal elements

$$|t_{p,j}^F| = \left| \frac{f_{p+1,j}^A}{f_{p,j}^A} \right| \lll 1 \quad (147)$$

of the transformation matrix \mathbf{T}_p^F ($j = 1 \dots, m$) are expected to be “small numbers”. It follows from Equations (68), (70), (73), and (147) that

$$\mu_i = \frac{\sum_{j=1}^m \left(S_{p,i,j}^+ f_{p,j}^A \right)}{\sum_{j=1}^m \left(S_{p,i,j}^+ \frac{f_{p,j}^A}{t_{p,j}^F} \right)} \lll 1 \quad (148)$$

($i = 1 \dots, n$) and

$$\|\Delta \mathbf{x}_{p+1}\| \leq \|\boldsymbol{\mu}\| \|\Delta \mathbf{x}_p^A\| \lll \|\Delta \mathbf{x}_p^A\| \quad (149)$$

(see Figure 10). This means that the T-Secant approximate \mathbf{x}_{p+1}^B will always be in the vicinity of the classic Secant approximate \mathbf{x}_{p+1}^A and the approximate errors of the new approximates will be of similar order, providing that the solution \mathbf{x}^* will be evenly surrounded by the $n + 1$ new trial approximates \mathbf{x}_{p+1}^A and $\mathbf{x}_{k,p+1}^B$ ($k = 1 \dots, n$), and that matrix \mathbf{S}_{p+1} will be well-conditioned.

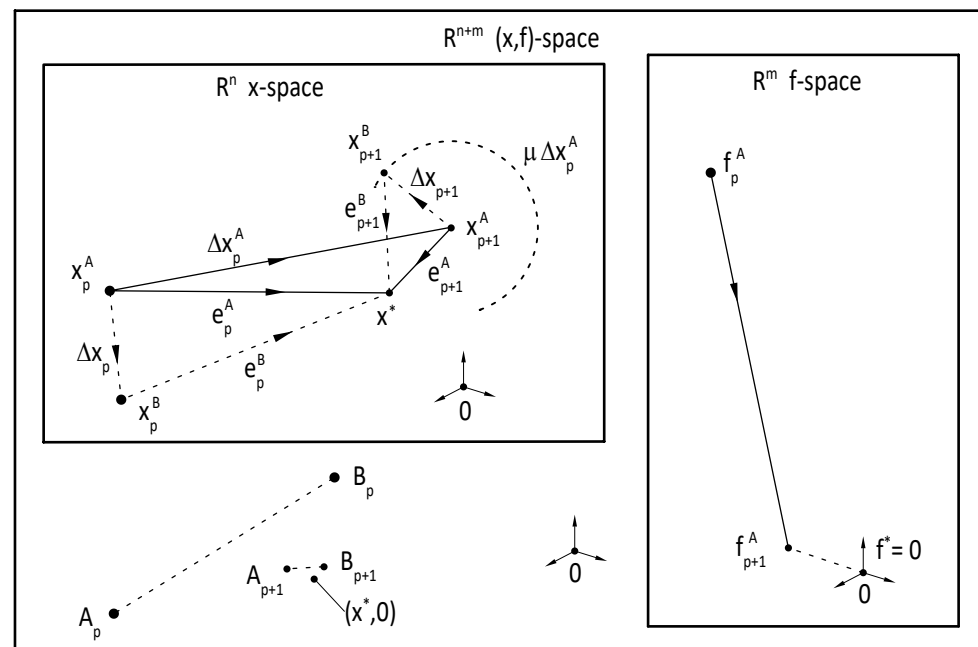


Figure 10. Geometrical representation of the T-Secant method convergence in multi-variable case (analogous to the convergence proof figure Dennis-Schnabel [13], p. 180).

8. Algorithm

Let p be the iteration counter, ε^* be the error bound for termination criterion, and

$$\mathbf{e}_p^A = \mathbf{x}_p^A - \mathbf{x}^* \quad (150)$$

be the approximate error vector of approximate x^A in the p^{th} iteration with elements $e_{p,i}^A$ ($i = 1, \dots, n$). Let the scalar approximate error

$$\varepsilon_p = \frac{\|e_p^A\|_2}{n} = \frac{\sqrt{\sum_{i=1}^n (e_{p,i}^A)^2}}{n} \quad (151)$$

be defined, where $\|\cdot\|_2$ is the Euclidean norm, and let the iteration be terminated when

$$\varepsilon_p < \varepsilon^* \quad (152)$$

holds. Let x_p^A be the initial trial and Δx_p be the trial increment (iteration stepsize) in the p^{th} iteration. Choose T_{min} and T_{max} as lower and upper bounds for $|t_{p,j}^F|$ ($j = 1 \dots, m$) and let f_{min} and q_{min} be lower bounds for $|f_{p,j}^A|$ ($j = 1, \dots, m$) and $|q_{p,i}^B|$ ($i = 1, \dots, n$), respectively.

- Initial step

Let $p=0$ and let the initial trial $x_p^A = (x_{p,1}^A \dots x_{p,n}^A)$ and the initial trial increment $\Delta x_p = (\Delta x_{p,1} \dots \Delta x_{p,n})$ be given. Calculate the corresponding function values f_p^A and assume that $f_{min} < |f_{p,j}^A|$ ($j = 1 \dots, m$).

- Step 1: Generate a set of n additional initial trials (interpolation base points)

$$x_{p,k}^B = x_p^A + \Delta x_{p,k} \cdot d^k \quad (153)$$

and evaluate function values $f_{p,k}^B$ ($k = 1, \dots, n$).

- Step 2 (Secant): Construct matrix

$$\Delta F_p = [\Delta f_{p,k}^i] = [f_{p,k}^B - f_p^A] \quad (154)$$

then calculate q_p^A from Equation (42). Let $q_{min} < |q_{p,i}^A|$, and determine x_{p+1}^A from Equation (46) and ε_p from Equation (151).

- Step 3: If $\varepsilon_p < \varepsilon^*$, then terminate iteration; otherwise, continue with Step 4.
- Step 4 (T-Secant): Calculate f_{p+1}^A and T_p^F from Equation (67). Let $T_{min} < |t_{p,j}^F| < T_{max}$ and determine q_p^B from Equation (68) (ΔF_p^+ has already been calculated when q_p^A was determined from Equation (42)). Let $q_{min} < |q_{p,i}^B|$. Calculate x_{p+1}^B from Equation (70).
- Step 5: Let the new initial trial be

$$\begin{bmatrix} x_{p+1}^A \\ f_{p+1}^A \end{bmatrix} \quad (155)$$

and the new initial trial increment (iteration stepsize) be

$$\Delta x_{p+1} = x_{p+1}^B - x_{p+1}^A, \quad (156)$$

and continue iteration with Step 1.

Iteration constants (δ_{min} f_{min} q_{min} T_{min} T_{max}) are necessary in order to avoid division by zero and to avoid computed values being near numerical precision. If p_{max} is the number of necessary iterations for satisfying the termination criterion $\varepsilon_p < \varepsilon^*$, and n is the number of unknowns to be determined, then the T-Secant method needs $n + 1$ function evaluations in each iteration, as well as

$$N_f = p_{max}(n + 1) \quad (157)$$

function evaluations to reach the desired termination criterion. p_{max} depends on many circumstances, such as the nature of the function $f(x)$, termination criteria (ε^* or others),

and the distance of the initial trial x^A from the solution x^* and from the iteration constants $(T_{min}, q_{min}^A, \dots)$.

9. Numerical Tests Results

9.1. Rosenbrock Test Function

A variant of the Rosenbrock function [24] has been used to test the numerical performance of the suggested method. The global minimum of the function

$$R(x) = \sum_{i=1}^{N-1} \left(100 \cdot (x_{i+1} - x_i^2)^2 + (1 - x_i)^2 \right) \quad (158)$$

has to be determined, where $x = (x_1 \dots x_N) \in \mathbb{R}^N$ and $N \geq 2$. $R(x)$ has exactly one minimum for $N = 3$ at $x^* = (1 \ 1 \ 1)$ and exactly two minima for $4 \leq N \leq 7$, i.e., a global minimum of all and a local minimum near $\hat{x} = (-1 \ 1 \dots 1)$. The sum of squares $R(x)$ will be minimum when all terms are zero, such that the minimization of the function $R(x)$ is equivalent to finding the zero of a function $x \rightarrow f(x)$, where $x \in \mathbb{R}^N$, $f: \mathbb{R}^N \rightarrow \mathbb{R}^{2(N-1)}$, and

$$f(x) = \begin{bmatrix} f_{2i-1}(x) \\ f_{2i}(x) \end{bmatrix} = \begin{bmatrix} 10 \cdot (x_{i+1} - x_i^2) \\ 1 - x_i \end{bmatrix} \quad (159)$$

($i = 1, \dots, N-1$). For $N > 7$, the function $R(x)$ has exactly one global minimum and has some local minima with some $x_i^* = -1$, with $x_i^* = 1$ for all other unknowns. The results were obtained by least squares solving of the simultaneous system of nonlinear equations $f(x) = 0$ via the T-Secant method.

9.2. $N = 2, N = 3$ and $N = 10$ Examples

In case the case of $N = n = m = 2$, the iterations terminated after $N_f = 6$ function evaluations ($p_{max} = 2$ iterations) in most cases. $f_2(x) = 1 - x_1$ is a linear function, and the first T-Secant iteration ($p = 0$) finds the exact value of x_1 in one step; then, $f_1(x) = 10(x_2 - x_1^2)$ also becomes linear. The exact value of x_2 was then determined in one additional step.

Let $N = n = 3$ and $m = 4$, $T_{min} = 0.01$ and $\varepsilon^* = 10^{-14}$. Let $p = 0$, and

$$\Delta x_{0,i} = 0.05 \cdot x_{0,i}^A \quad (160)$$

($i = 1 \dots 3$). The number of necessary function evaluations N_f varied between 20 and 36 within $p_{max} = 5-9$ iterations for different initial trials x_0^A . Iteration results are summarized in Table 7 and in Figure 11 with initial trial $x_0^A = (x_{0,i}^A) = (2.0 \ -1.5 \ -2.5)$. Termination criterion $\varepsilon_p < \varepsilon^*$ was satisfied after $p_{max} = 5$ iterations with $N_f = 20$ function evaluations.

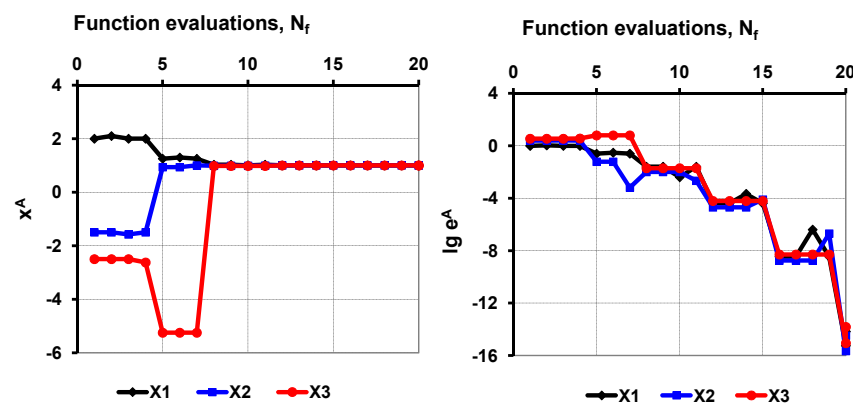


Figure 11. (Left) Variables $x_{p,i}^A$ and (Right) absolute approximate errors $\lg|e_p^A(x_{p,i}^A)|$ ($i = 1 \dots 3$) variation for initial trial $x_0^A = (2.0 \ -1.5 \ -2.5)$.

Table 7. Iteration results: $x_0^A = \begin{pmatrix} 2.0 & -1.5 & -2.5 \end{pmatrix}$, $T_{min} = 0.01$, $T_{max} = 1.5$.

p	0	1	2	3
x_p^A	$\begin{bmatrix} 2 \\ -1.5 \\ -2.5 \end{bmatrix}$	$\begin{bmatrix} 1.253 \\ 0.938 \\ -5.248 \end{bmatrix}$	$\begin{bmatrix} 1.026 \\ 0.990 \\ 0.980 \end{bmatrix}$	$\begin{bmatrix} 1.00004 \\ 0.99998 \\ 0.99994 \end{bmatrix}$
Δx_p	$\begin{bmatrix} 0.1 \\ -0.075 \\ -0.125 \end{bmatrix}$	$\begin{bmatrix} 0.046 \\ 0.061 \\ -0.026 \end{bmatrix}$	$\begin{bmatrix} -0.0217 \\ 0.0079 \\ -0.063 \end{bmatrix}$	$\begin{bmatrix} -3 \times 10^{-4} \\ 1 \times 10^{-4} \\ 2 \times 10^{-4} \end{bmatrix}$
f_p^A	$\begin{bmatrix} -55 \\ -1 \\ -47.5 \\ 2.5 \end{bmatrix}$	$\begin{bmatrix} -6.320 \\ -0.253 \\ -61.28 \\ 0.062 \end{bmatrix}$	$\begin{bmatrix} -0.621 \\ -0.026 \\ 0.005 \\ 0.010 \end{bmatrix}$	$\begin{bmatrix} -0.00102 \\ -0.00004 \\ -0.00021 \\ 0.00002 \end{bmatrix}$
q_p^A	$\begin{bmatrix} 7.47 \\ 32.5 \\ -22.0 \end{bmatrix}$	$\begin{bmatrix} 4.915 \\ -0.846 \\ 243.9 \end{bmatrix}$	$\begin{bmatrix} -1.184 \\ -1.269 \\ 0.307 \end{bmatrix}$	$\begin{bmatrix} -0.160 \\ -0.201 \\ -0.327 \end{bmatrix}$
x_{p+1}^A	$\begin{bmatrix} 1.253 \\ 0.938 \\ -5.248 \end{bmatrix}$	$\begin{bmatrix} 1.026 \\ 0.990 \\ 0.980 \end{bmatrix}$	$\begin{bmatrix} 1.00004 \\ 0.99998 \\ 0.99994 \end{bmatrix}$	$\begin{bmatrix} 0.9 \dots \\ 1.0 \dots \\ 1.0 \dots \end{bmatrix}$
e_{p+1}^A	$\begin{bmatrix} 0.253 \\ 0.062 \\ 6.248 \end{bmatrix}$	$\begin{bmatrix} 0.026 \\ 0.010 \\ 0.020 \end{bmatrix}$	$\begin{bmatrix} 4 \times 10^{-5} \\ 2 \times 10^{-5} \\ 6 \times 10^{-5} \end{bmatrix}$	$\begin{bmatrix} 3 \times 10^{-9} \\ 2 \times 10^{-9} \\ 5 \times 10^{-9} \end{bmatrix}$
$R(x_{p+1}^A)$	6.2×10^1	6.2×10^{-1}	1.0×10^{-3}	9.0×10^{-8}
ε_p	2.1×10^0	1.1×10^{-2}	2.6×10^{-5}	2.2×10^{-9}
f_{p+1}^A	$\begin{bmatrix} -6.32 \\ -0.253 \\ -61.3 \\ 0.062 \end{bmatrix}$	$\begin{bmatrix} -0.621 \\ -0.026 \\ 0.005 \\ 0.010 \end{bmatrix}$	$\begin{bmatrix} -0.00102 \\ -0.00004 \\ -0.00021 \\ 0.00002 \end{bmatrix}$	$\begin{bmatrix} 0.0 \dots \\ 0.0 \dots \\ 0.0 \dots \\ -0.0 \dots \end{bmatrix}$
t_p^F	$\begin{bmatrix} 0.115 \\ 0.253 \\ 1.290 \\ 0.025 \end{bmatrix}$	$\begin{bmatrix} 0.098 \\ 0.102 \\ -0.01 \\ 0.163 \end{bmatrix}$	$\begin{bmatrix} 0.01 \\ 0.01 \\ -0.044 \\ 0.01 \end{bmatrix}$	$\begin{bmatrix} -0.01 \\ -0.01 \\ -0.01 \\ -0.01 \end{bmatrix}$
q_p^B	$\begin{bmatrix} -120 \\ 1298 \\ -2365 \end{bmatrix}$	$\begin{bmatrix} 51.6 \\ -5.52 \\ -240000 \end{bmatrix}$	$\begin{bmatrix} -118 \\ -127 \\ 32 \end{bmatrix}$	$\begin{bmatrix} 16.0 \\ 20.1 \\ 32.7 \end{bmatrix}$
x_{p+1}^B	$\begin{bmatrix} 1.299 \\ 0.999 \\ -5.273 \end{bmatrix}$	$\begin{bmatrix} 1.004 \\ 0.998 \\ 0.917 \end{bmatrix}$	$\begin{bmatrix} 0.99978 \\ 1.00008 \\ 1.00013 \end{bmatrix}$	$\begin{bmatrix} 1.0 \dots \\ 0.9 \dots \\ 0.9 \dots \end{bmatrix}$
Δx_{p+1}	$\begin{bmatrix} 0.046 \\ 0.061 \\ -0.026 \end{bmatrix}$	$\begin{bmatrix} -0.0217 \\ 0.0079 \\ -0.063 \end{bmatrix}$	$\begin{bmatrix} -3 \times 10^{-4} \\ 1 \times 10^{-4} \\ 2 \times 10^{-4} \end{bmatrix}$	$\begin{bmatrix} 4 \times 10^{-7} \\ -2 \times 10^{-7} \\ -6 \times 10^{-7} \end{bmatrix}$

Let $N = n = 10$ and $m = 18$. Calculations were made with different, manually constructed initial trials $x_0^A = (x_{0,i}^A)$. Figure 12 (Left) shows the variation of $x_{p,i}^A$ for initial trial $x_0^A = (2.0 \ -1.5 \ -2.5 \ 1.5 \ -1.2 \ 3.0 \ -3.5 \ 2.5 \ -2.0 \ 3.5)$. Iterations terminated after $N_f = 154$ function evaluations ($p_{max} = 14$ iterations) for the $\varepsilon_p < \varepsilon^* = 10^{-14}$ condition. Table 8 shows a set of further initial trials for numerical tests. Test “3” failed, probably due to the large distance from the global optimal solution. Test “4” found a local zero $x^* = (-1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1)$. Figure 12 (Right) summarizes the results of numerical tests “1–6”. The graphs show the iteration paths in the $\lg|e_p^A| - R_p(x_p^A)$ plane. The graphs have an initial part, where the variation of $R_p(x_p^A)$ seems “chaotic”, while below $|e_p^A| \cong 0.01$ and $R_p(x_p^A) \cong 0.001$, the iterations run on similar paths.

Table 8. Initial trial vectors ($N = 10$, $n = 10$, $m = 18$), $(\mathbf{x}^* = (1 \dots 1))$.

	\mathbf{x}_0^A	p_{max}	N_f
1	(1.3 −1.5 −2.1 1.1 −1.3 1.8 −1.8 1.7 −2.0 2.1)	15	165
2	(3.1 −2.1 −4.3 1.2 −2.4 3.6 −1.6 2.7 −4.2 2.2)	21	231
3	(−4.1 1.1 −6.3 −3.2 −4.4 1.6 3.6 5.7 −2.2 3.2)	-	-
4	(−3.0 −3.1 2.3 −4.2 2.4 −1.6 −3.6 2.7 −2.2 4.2)	-	-
5	(2.1 3.1 −1.3 −2.2 −3.4 1.6 2.6 −1.7 2.2 −3.2)	16	176
6	(3.1 3.1 −4.3 −2.2 −3.4 2.6 1.6 −4.7 2.2 −2.2)	20	220

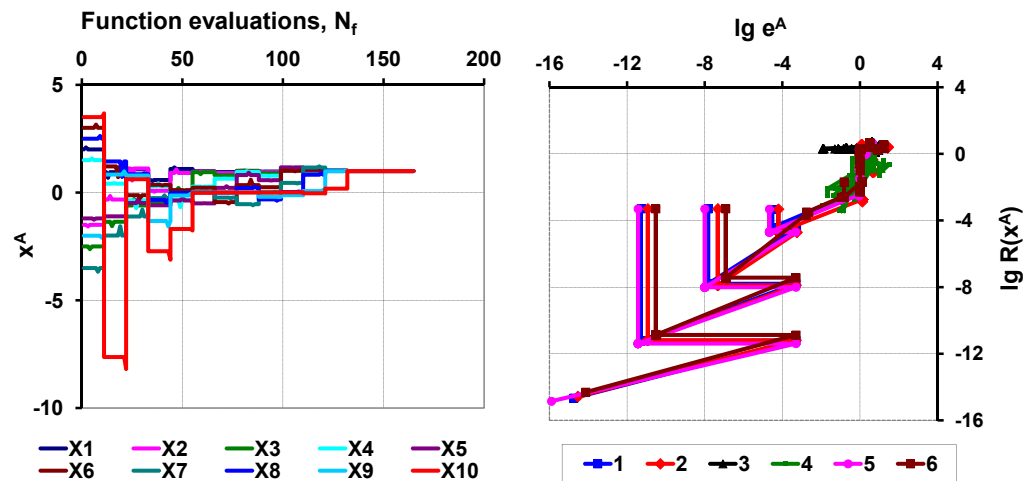


Figure 12. (Left) Variation of $x_{p,i}^A$ for $x_0^A = (2.0 \ -1.5 \ -2.5 \ 1.5 \ -1.2 \ 3.0 \ -3.5 \ 2.5 \ -2.0 \ 3.5)$ through iterations ($N = 10 = n = 10$, $m = 18$) with $p_{max} = 15$ and $N_f = 165$. (Right) The absolute approximate errors $\lg|e_p^A(x_{p,i}^A)|$ ($i = 1 \dots 10$) and the $R(x_p^A)$ function variation through iterations for different initial trials ($N = 10$, $n = 10$, $m = 18$) (see Table 8).

9.3. Large N (200 500 1000) Examples

A series of numerical tests has been performed with a large number of unknown variables. The values of the initial trials $\mathbf{x}_0^A = (x_{0,i}^A), (i = 1, \dots, N)$ were generated as

$$x_{0,i}^A = x_i^* + L_1 \cdot \frac{Random - \frac{1}{2}}{5} + L_2, \quad (161)$$

where “Random” is a random real number ($0 \leq Random < 1$), and $L_i (i = 1, 2)$ are parameters regulating the size and location of the interval in which the initial trial values are expected to vary. $\mathbf{x}^* = (x_i^*) = (1 \dots 1)$ ($i = 1, \dots, N$) is the known global optimal solution. Table 9 shows the results of T-Secant iterations with $N = 200$ and with initial trials $\mathbf{x}_0^A : 0.1 \leq x_{0,i}^A \leq 19.9$ ($L_1 = 99$, $L_2 = 9$). Figure 13 (Left) shows the variation of variables x_p^A through T-Secant iterations. The iteration counter p value is indicated below the graphs. Figure 13 (Right) shows the decrease in the approximate error $e_p^A = (e_{p,i}^A) (i = 1, \dots, 200)$, with the p iteration counter indication below the graphs. Table 10 shows the results of iterations with $N = 1000$ and initial trials $\mathbf{x}_0^A : 0.5 \leq x_{0,i}^A \leq 1.5$ ($L_1 = 5$, $L_2 = 0$). Figure 14 summarizes the results of numerical tests with a large number of unknowns $N = (200 \ 500 \ 1000)$. The norm ε_p of the approximate error e_p^A decrease is shown, and the number of function value evaluations N_f is indicated for $N = (200(blue) \ 500(red) \ 1000(green))$ and for initial trials $\mathbf{x}_0^A : 0.5 \leq x_{0,i}^A \leq 1.5$ (solid line) and $\mathbf{x}_0^A : 0.1 \leq x_{0,i}^A \leq 19.9$ (dashed line).

Table 9. Iteration results ($N = 200, L_1 = 99.9, L_2 = 9$) with initial trials $0.1 \leq x_{0,i}^A \leq 19.9$ (dashed blue line on Figure 14).

p	ε_p	$R(x_p^A)$	N_f
0	10.6925833405791	24123.43773726327	1
1	5.45917411911925	6895.1103569982861	201
2	2.13338434746463	1247.4064173528971	402
3	0.71430571273689	220.36900527956962	603
4	0.163511639031299	32.621494717337107	804
5	0.0145616620270659	2.4077509738413969	1005
6	0.000197003511771894	0.026366233831030046	1206
7	0.000000084768909602	0.000007982826913871	1407
8	0.00000000032791210	0.00000003114429023	1608
9	0.00000000000013862	0.00000000001333830	1809
10	0.00000000000000546	0.00000000000104185	2010

Table 10. Iteration results ($N = 1000, L_1 = 5, L_2 = 0$) with initial trials $0.5 \leq x_{0,i}^A \leq 1.5$ (solid green line on Figure 14).

p	ε_p	$R(x_p^A)$	N_f
0	0.287800987765134	212.38512786560364	1
1	0.121219403643695	57.87378211356512	1001
2	0.0396263348376487	13.743840511211417	2002
3	0.0298060844365720	9.6618077142097238	3003
4	0.0120370539008435	5.9465782106406841	4004
5	0.000705489922936629	0.42465246853444877	5005
6	0.000002762586723754	0.001324115254348589	6006
7	0.00000000990421380	0.000000388965253003	7007
8	0.00000000000433209	0.00000000155930410	8008
9	0.00000000000000860	0.00000000000363149	9009

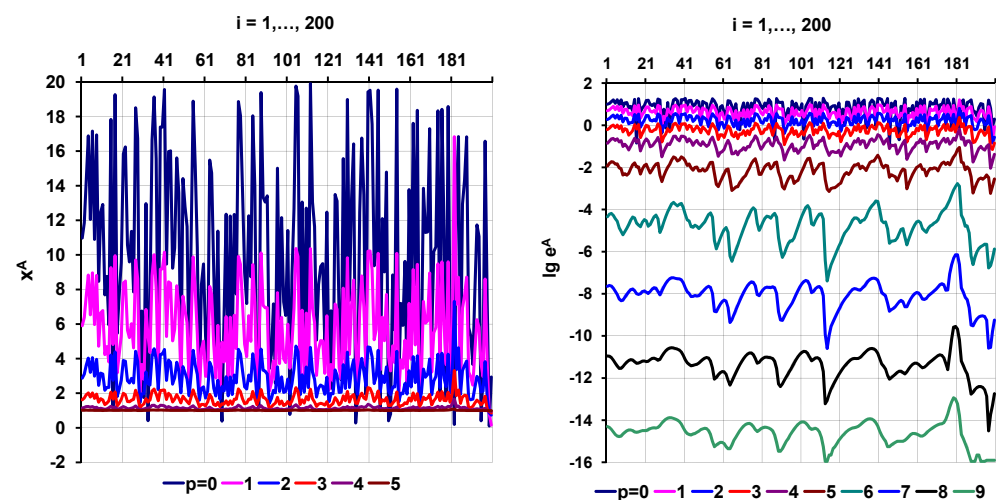


Figure 13. (Left) Variation of variables x_p^A through iterations. (Right) Decrease in approximate error $\lg e_p^A$ through iterations, $N = 200$ (with iteration counter p value indication below the graphs).

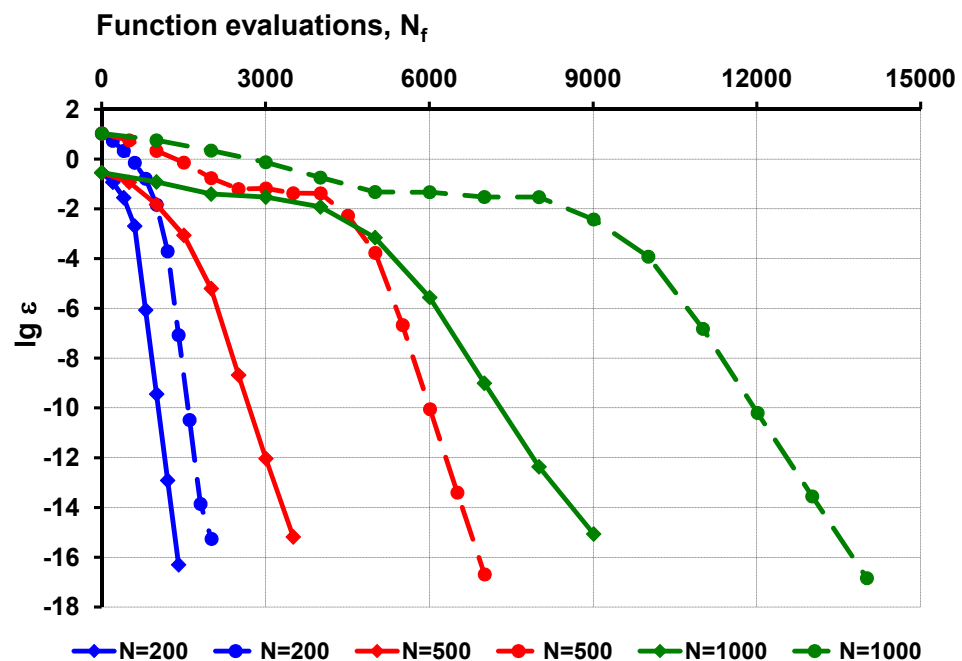


Figure 14. Number of function evaluations for $N = 200$ (blue), $N = 500$ (red), and $N = 1000$ (green) with initial trials $x_0^A : 0.5 \leq x_{0,i}^A \leq 1.5$ (solid line) and $x_0^A : 0.1 \leq x_{0,i}^A \leq 19.9$ (dashed line).

10. Efficiency

10.1. Single-Variable Case

The efficiency of an algorithm relates to the amount of computational resources used by the algorithm. For better efficiency, it is desirable to minimize resource usage. An algorithm is considered efficient if its resource consumption (computational cost) is below some acceptable level (it runs in a reasonable amount of time or space on an available computer). The efficiency of an algorithm for the solution of nonlinear equations is thoroughly discussed by Traub [25] as follows. Let p be the order of the iteration sequence such that for the approximate errors $e_i = x_i - x^*$, there exists a nonzero constant C (asymptotic error constant) for which

$$\frac{|e_{i+1}|}{|e_i|^p} \rightarrow C. \quad (162)$$

A natural measure of the information used by an algorithm is the “informational usage” d , which is defined as the number of new pieces of information (values of the function and its derivatives) required per iteration (called “horner” by Ostrowski [26]). Then, the efficiency of the algorithm within one iteration can be measured by the “informational efficiency”:

$$EFF = \frac{p}{d}. \quad (163)$$

An alternative definition of efficiency is

$$^*EFF = p^{\frac{1}{d}}, \quad (164)$$

called the “efficiency index” by Ostrowski [26]. Another measure of efficiency, called “computational efficiency”, takes into account the “cost” of calculating different derivatives. The concept of informational efficiency (EFF) and the efficiency index (*EFF) do not take into account the cost of evaluating f and its derivatives, nor do they take into account the total number of pieces of information needed to achieve a certain accuracy in the root of the function. If f is composed of elementary functions, then the derivatives are also composed

of elementary functions; thus, the cost of evaluating the derivatives is merely the cost of combining the elementary functions. Table 11 compares the efficiencies of classic (secant, Newton) and improved algorithms (T-Secant, T-Newton).

Table 11. Efficiencies of classic and improved algorithms.

Method	d	p	EFF [25]	$*EFF$ [26]
Secant	1	1.618...	1.618...	1.618...
Newton	2	2.0	1.0	1.414...
T-Secant	2	2.618...	1.309...	1.618...
T-Newton	3	3.0	1.0	1.442...

10.2. Multi-Variable Case

Very limited data are available to compare the performance of the T-Secant method with other methods, especially in cases with a large number of unknowns. Broyden [27] suggested the mean convergence rate

$$L = \frac{1}{N_f} \ln \frac{R(x_0^A)}{R(x_{p_{max}}^A)} \quad (165)$$

as a measure of efficiency of an algorithm for solving a particular problem, where N_f is the total number of function evaluations, x_0^A is the initial trial, and $x_{p_{max}}^A$ is the last trial for the solution x^* when the termination criteria is satisfied after p_{max} iterations. $R(x)$ is the Euclidean norm of $f(x)$. Efficiency results were given by Broyden [27] for the Rosenbrock function for $N = 2$ and for $x_0^A = (-1.2 \ 1.0)$. The calculated convergence rates for the two Broyden method variants [27], for the Powell's method [28], for the adaptive coordinate descent method [29] and for the Nelder–Mead simplex method [30] were compared with the calculated values for the T-Secant method in Table 12. Rows 1–5 are data from referenced papers, rows 6–8 are T-Secant results with the referenced initial trials, and rows 9–15 are calculated data for $N > 2$.

Results show that the mean convergence rate L (Equation (165)) for $N = 2$ is much higher for the T-Secant method ($\simeq 5.5$ – 6.9) than for the other listed methods ($\simeq 0.1$ – 0.6); however, it is obvious that the mean convergence rate values decrease rapidly with increasing N values (more unknowns need more function evaluations). A modified convergence rate

$$L_N = N * L = \frac{N}{N_f} \ln \frac{R(x_0^A)}{R(x_{p_{max}}^A)} \quad (166)$$

can be used as an “ N ” independent measure of efficiency (see Table 12). The values of L and L_N are at least 10 times larger for the T-Secant method than for the referenced classic methods for $N = 2$ (see Table 12). Note that the efficiency measures (L and L_N) are also dependant on the initial conditions (distance of the initial trial set from the optimal solution, termination criterion). Results from a large number of numerical tests indicate an average L_N value of around 7.4, with standard deviation 3.7 for the T-Secant method even for large N values. It has to be noted that if the value of $R(x_{p_{max}}^A)$ is zero, then the mean convergence rates (L and L_N) are not countable (zero in the denominator). A substitute value 10^{-25} was used when iterations ended with $R(x_{p_{max}}^A) = 0$ in the sample examples.

Table 12. Calculated values of the mean convergence rates (L and L_N) for the Rosenbrock function (¹: a substitute value 10^{-25} was used when $R(x_{p_{max}}^A) = 0$).

	N	Method	$R(x_0^A)$	$R(x_{p_{max}}^A)$	p_{max}	N_f	L	L_N
1	2	Broyden 1. [27]	4.9193	4.73×10^{-1}	-	59	0.391	0.78
2	2	Broyden 2. [27]	4.9193	2.55×10^{-1}	-	39	0.607	1.22
3	2	Powell [28]	4.9193	7.00×10^{-1}	-	151	0.150	0.30
4	2	ACD [29]	130.062	1.00×10^{-1}	-	325	0.086	0.17
5	2	Nelder-Mead [30]	2.0000	1.36×10^{-1}	-	185	0.127	0.25
6	2	T-secant [27,28]	4.9193	1.0×10^{-25} ¹	3	9	6.573 ¹	13.15 ¹
7	2	T-secant [29]	130.06	1.0×10^{-25} ¹	3	9	6.937 ¹	13.87 ¹
8	2	T-secant [30]	2.0000	6.66×10^{-15}	2	6	5.556	11.11
9	3	T-secant	72.722	1.41×10^{-14}	5	20	1.809	5.43
10	3		32.466	1.0×10^{-25} ¹	4	16	3.815 ¹	11.45 ¹
11	5		93.528	1.34×10^{-14}	8	48	0.760	3.80
12	5		7.193	5.90×10^{-14}	4	24	1.351	6.76
13	10		202.62	1.0×10^{-25} ¹	14	154	0.408 ¹	4.08 ¹
14	200		92.778	9.00×10^{-15}	10	2010	0.042	8.44
15	1000		212.39	3.63×10^{-13}	6	6006	0.006	5.66

11. Discussions

11.1. General

The suggested procedure needs the usual approximate x_{p+1}^A to be determined by any of a classic quasi-Newton iterative methods (Wolfe–Popper–Secant, Broyden, etc.). By using the “information” f_{p+1}^A , an additional and independent approximate x_{p+1}^B is determined, which provides the possibility for a full-rank update of the approximate derivatives (S_p for Secant or B_p for Broyden). Results and experience show that the suggested procedure considerably accelerates the convergence and the efficiency of the classic methods, and the full-rank update technique increases the stability of the iterative procedure. In multi-variable-case, it follows from Equation (107) that

$$\left(T_p^X\right)^{-1} \Delta x_p^A = -S_p^+ \left(T_p^F\right)^{-1} f_p^A, \quad (167)$$

and in explicit form after re-arrangement:

$$\left[\frac{\left(\Delta x_{p,i}^A\right)^2}{x_{p+1,i}^B - x_{p+1,i}^A}\right] = -\left[S_{p,i,j}^+\right] \left[\frac{f_{p,j}^A}{t_{p,j}^F}\right]. \quad (168)$$

Then, the i^{th} element of the new approximate x_{p+1}^B can be expressed from the i^{th} row of the above equation as

$$x_{p+1,i}^B = x_{p+1,i}^A - \frac{\left(\Delta x_{p,i}^A\right)^2}{\sum_{j=1}^m \left(S_{p,i,j}^+ \frac{f_{p,j}^A}{t_{p,j}^F}\right)}. \quad (169)$$

The mechanism of the procedure resembles to the mechanism of an engine’s turbocharger that is powered by the flow of exhaust gases (analogous to f_{p+1}^A or $t_{p,j}^F$).

11.2. Newton Method

Matrix S in the general formula (104) gives a direct connection between the Secant and Newton methods, as differences go to differentials,

$$S = \left[\frac{\Delta f_{kj}}{\Delta x_i} \right] = [S_{i,j}] \longrightarrow J = \left[\frac{\partial f_{kj}}{\partial x_i} \right] = [J_{i,j}] \quad (170)$$

where J is the Jacobian matrix of the function $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ($m \geq n$) with k and i column and with j row indexes, respectively. It follows from formula (111) of matrix S_T that the suggested full-rank update procedure can also be applied to the Newton method as

$$S_T = \left[\begin{array}{cc} t_j^F & \Delta f_{kj} \\ t_i^X & \Delta x_i \end{array} \right] \longrightarrow J_T = \left[\begin{array}{cc} t_j^F & \frac{\partial f_{kj}}{\partial x_i} \end{array} \right], \quad (171)$$

where J_T is the modified Jacobian matrix of the “T-Newton” method. In the single-variable case, with approximate x_p^A in the p^{th} iteration, with function value $f_p^A = f(x_p^A)$ and with derivative function value $f_p'^A = f'(x_p^A)$, the new Newton–Raphson approximate can be expressed as

$$x_{p+1}^A = x_p^A - \frac{\partial x_p}{\partial f_p} f_p^A = x_p^A - \frac{f_p^A}{f_p'^A}, \quad (172)$$

and the iteration stepsize is

$$\Delta x_p^A = x_{p+1}^A - x_p^A \quad (173)$$

with the hyperbolic function (Equation (88))

$$z_p(x) = \frac{a_p}{x - x_{p+1}^A} + f_p^A. \quad (174)$$

where

$$a_p = (x_{p+1}^A - x_p^A)^2 f_p'^A \frac{f_{p+1}^A}{f_p^A} \quad (175)$$

($\Delta f_p / \Delta x_p$ is replaced by $f_p'^A$), the new “T-Newton” approximate is

$$x_{p+1}^B = x_{p+1}^A - \frac{(\Delta x_p^A)^2 f_p'^A f_{p+1}^A}{(f_p^A)^2} \quad (176)$$

($\Delta f_p / \Delta x_p$ is again replaced by $f_p'^A$), similar to Equation (59) in case of the T-secant method. It can be seen from Tables 13 and 14 that the convergence rate is improved from $\alpha^N = 2$ to $\alpha^{TN} = 3$. In the multi-variable case, it follows from Equation (107) (S_p^+ is replaced by J_p^+) that

$$(T_p^X)^{-1} \Delta x_p^A = -J_p^+ (T_p^F)^{-1} f_p^A \quad (177)$$

and in explicit form after re-arrangement:

$$\left[\frac{(\Delta x_{p,i}^A)^2}{x_{p+1,i}^B - x_{p+1,i}^A} \right] = -[J_{p,i,j}^+] \left[\frac{f_{p,j}^A}{t_{p,j}^F} \right]. \quad (178)$$

Then, the i^{th} element of the new “T-Newton” approximate x_{p+1}^B can be expressed from the i^{th} row of the above equation as

$$x_{p+1,i}^B = x_{p+1,i}^A - \frac{(\Delta x_{p,i}^A)^2}{\sum_{j=1}^m \left(J_{p,i,j}^+ \frac{f_{p,j}^A}{t_{p,j}^F} \right)}, \quad (179)$$

similar to Equation (70) in case of the T-Secant method. Thus, the “hyperbolic” approximation accelerates the convergence of the Newton–Raphson method by only one additional function evaluation.

Table 13. Newton method iteration and computed convergence rate, α^N .

p	x_p^A	x_{p+1}^A	$ e_{p+1}^A $	α^N	N_f	$N_{f'}$
0	4.5	3.187	1.1×10^0		1	1
1	3.187	2.44965	3.6×10^{-1}		2	2
2	2.44965	2.14996	5.5×10^{-2}	1.42	3	3
3	2.14996	2.096188	1.6×10^{-3}	1.66	4	4
4	2.096188	2.094552	1.5×10^{-6}	1.89	5	5
5	2.094552	2.09455148	1.3×10^{-12}	1.99	6	6
6	2.09455148	2.09455148154233	3.6×10^{-15}	2.00	7	7

Table 14. T-Newton method iteration and computed convergence rate, α^{TN} (see Figure 15).

p	x_p^A	x_{p+1}^B	$ e_{p+1}^B $	α^{TN}	N_f	$N_{f'}$
0	4.5	2.830	7.4×10^{-1}		2	1
1	2.830	2.17760	8.3×10^{-2}		4	2
2	2.17760	2.09486	3.1×10^{-4}	1.84	6	3
3	2.09486	2.09455148	1.9×10^{-11}	2.56	8	4
4	2.09455148	2.09455148154233	3.6×10^{-15}	2.97	9	5

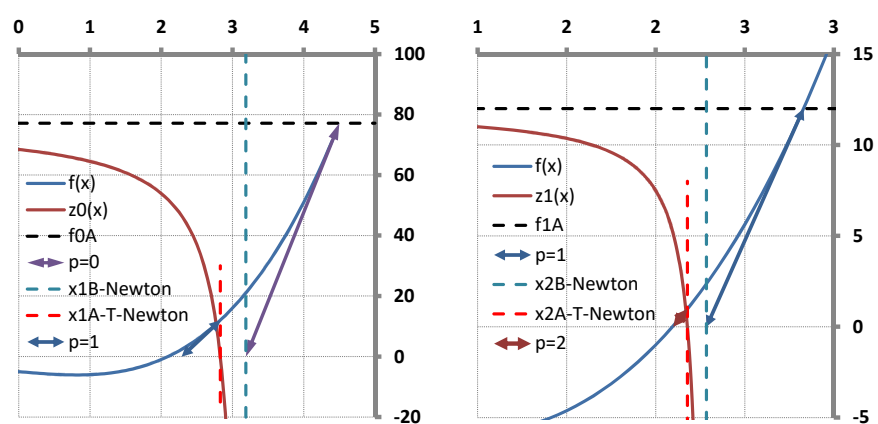


Figure 15. T-Newton iterations with test function (99) with initial approximate $x_0^A = 4.5$. **Left:** x_1^B is the root of the tangent line through f_0^A ; x_1^A is the root of $z_0(x)$. **Right:** x_2^B is the root of the tangent line through f_1^A ; x_2^A is the root of $z_1(x)$ (see data in Table 14).

11.3. Broyden's Method

Broyden's method is a special case of the Secant method. In the single-variable case, the derivative of the function is approximated as

$$f'_p \simeq B_p = B_{p-1} + \frac{\Delta f_p - B_{p-1} \Delta x_p}{|\Delta x_p|^2} \Delta x_p \quad (180)$$

in the p^{th} iteration step, and with

$$\frac{\Delta x_p}{|\Delta x_p|^2} = \frac{1}{\Delta x_p} \quad (181)$$

it is simplified as

$$B_p = B_{p-1} + \frac{\Delta f_p - B_{p-1} \Delta x_p}{\Delta x_p}. \quad (182)$$

The next Broyden-approximate is then determined as

$$x_{p+1}^A = x_p^A - \frac{f_p^A}{B_p}. \quad (183)$$

The convergence can similarly be improved by the new hyperbolic approximation procedure as in cases of the Secant and Newton methods. An additional new approximate

$$x_{p+1}^B = x_{p+1}^A - \frac{(\Delta x_p^A)^2 B_p f_{p+1}^A}{(f_p^A)^2} \quad (184)$$

can be determined, and the iteration continues with this value. Figure 16 demonstrates the effect of the hyperbolic approximation applied to the classic Broyden method. Not surprisingly, the convergence rate will be improved from $\alpha^B = \varphi \simeq 1.618$ to $\alpha^{TB} = \varphi^2 \simeq 2.618$, as in case of the Secant method. In the multi-variable case, the i^{th} element of the new the "T-Broyden" approximate x_{p+1}^B can be expressed as

$$x_{p+1,i}^B = x_{p+1,i}^A - \frac{(\Delta x_{p,i}^A)^2}{\sum_{j=1}^m \left(B_{p,i,j}^+ \frac{f_{p,j}^A}{t_{p,j}^F} \right)}, \quad (185)$$

similar to Equation (179) for the T-Newton method with $J_{p,i,j}^+$ replaced by $B_{p,i,j}^+$. The new approximate B_{p+1} to the Jacobian matrix can then be fully updated in a similar way as it was in the case of the T-Secant method.

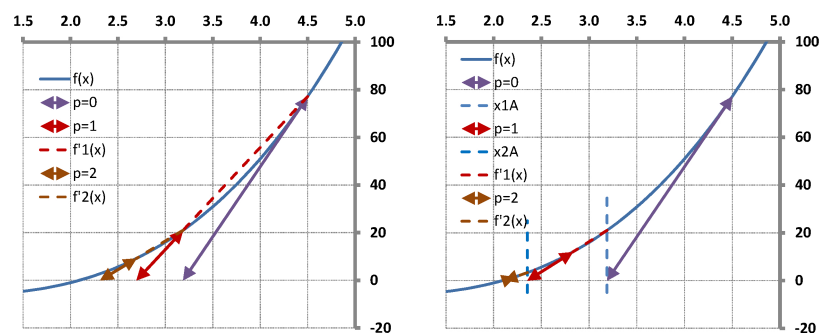


Figure 16. Broyden (Left) and T-Broyden (Right) iterations with test function (99) with initial approximates $x_0^A = 4.5$.

12. Conclusions

A completely new iteration strategy has been worked out for solving simultaneous nonlinear equations:

$$f(x) = 0, \quad (186)$$

$x \in \mathbb{R}^n$, and $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ($m \geq n$). It replaces the Jacobian matrix with finite-difference approximations. The stepsize Δx_{p+1} was determined as the difference between two new approximates

$$x_{p+1}^A = x_p^A + \Delta x_p^A \quad (187)$$

and x_{p+1}^B with elements

$$x_{p+1,i}^B = x_{p+1,i}^A + \Delta x_{p+1,i} = x_{p+1,i}^A - \frac{(\Delta x_{p,i}^A)^2}{\sum_{j=1}^m \left(S_{p,i,j}^+ \frac{f_{p,j}^A}{t_{p,j}^F} \right)} \quad (188)$$

($i = 1, \dots, n$) as

$$\Delta x_{p+1} = x_{p+1}^B - x_{p+1}^A. \quad (189)$$

The first one is a classic quasi-Newton approximate with stepsize Δx_p^A , while the second one was determined from a hyperbolic approximation governed by x_{p+1}^A and f_{p+1}^A , such that the classic Secant equation

$$S \Delta x^A = -f^A \quad (190)$$

was modified by a non-uniform scaling transformation

$$T = \begin{bmatrix} T^X & \mathbf{0} \\ \mathbf{0} & T^F \end{bmatrix} \quad (191)$$

with diagonal elements t_j^F ($j = 1, \dots, m$), t_i^X ($i = 1, \dots, n$) as

$$S_T \Delta x^A = -f^A, \quad (192)$$

where

$$S = \begin{bmatrix} \frac{\Delta f_{k,j}}{\Delta x_i} \end{bmatrix} \quad \text{and} \quad S_T = \begin{bmatrix} \frac{t_j^F}{t_i^X} \frac{\Delta f_{k,j}}{\Delta x_i} \end{bmatrix} \quad (193)$$

($k = 1, \dots, n$). It was shown that the new stepsize Δx_{p+1} is much smaller than the stepsize Δx_p^A of the classic quasi-Newton approximate, providing that x_{p+1}^B will always be in the vicinity of x_{p+1}^A . Having two new approximates, a set of $n + 1$ new independent trial approximates x_{p+1}^A and $x_{k,p+1}^B$ ($k = 1, \dots, n$) was constructed (see Equation (32)), providing that the new trial approximates are always in general positions, ensuring the stable behavior of the iteration. According to the geometrical representation in the single-variable case, the suggested procedure corresponds to finding the root of a hyperbolic function with vertical and horizontal asymptotes x_{p+1}^A and f_p^A . It was shown in Section 7 that the suggested method has super-quadratic convergence with a rate of $\alpha^{TS} = \varphi^2 = 2.618\dots$ (where $\varphi = 1.618\dots$ is the well-known golden ratio) in the single-variable case.

The suggested method needs two function evaluations in each iteration in single-variable cases and $n + 1$ evaluations in multi-variable cases. The efficiency of the proposed method was studied in Section 10 in the multi-variable case and compared with other classic low-rank-update and line-search methods on the basis of available data. The results show that the efficiency of the suggested full-rank-update procedure is considerably better than the efficiency of the other referenced methods. A Rosenbrock test function

(Equations (158) and (159)) with up to $n = 1000$ variables was used to demonstrate this efficiency in Section 9.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: A considerable part of the research work was conducted between the years 1988 and 1992 at the Technical University of Budapest (Hungary), at the TNO-BOUW Structural Division (The Netherlands), and at the Technical High-school of Lulea (Sweden). This work has been sponsored by the Technical University of Budapest (Hungary), by the Hungarian Academy of Sciences (Hungary), by TNO-BOUW (The Netherlands), by Sandvik Rock Tools (Sweden), by CP Test a/s (Denmark), and by Óbuda University (Hungary). The valuable discussions and personal support from Géza Petrasovits, György Popper, Peter Middendorp, Rikard Skov, Bengt Lundberg, Mario Martinez, and Csaba Hegedűs are greatly appreciated.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Růth, B.; Uekermann, B.; Mehl, M.; Birken, P.; Monge, A.; Bungartz, H.-J. Quasi-Newton waveform iteration for partitioned surface-coupled multiphysics applications. *Int. J. Numer. Methods Eng.* **2021**, *122*, 5236–5257. [\[CrossRef\]](#)
2. Barnafi, N.A.; Pavarino, L.F.; Scacchi, S. Parallel inexact Newton-Krylov and quasi-Newton solvers for nonlinear elasticity. *Comput. Methods Appl. Mech. Engrg.* **2022**, *400*, 115557. [\[CrossRef\]](#)
3. Ryu, J.; Jae, M. A quantification methodology of Seismic Probabilistic Safety Assessment for nuclear power plant. *Ann. Nucl. Energy* **2021**, *159*, 108296. [\[CrossRef\]](#)
4. Yahaya, M.M.; Kumam, P.; Awwal, A.M.; Aji, S. A structured quasi-Newton algorithm with nonmonotone search strategy for structured NLS problems and its application in robotic motion control. *J. Comput. Appl. Math.* **2021**, *395*, 113582. [\[CrossRef\]](#)
5. Schröter, M.; Sauer, O. Quasi-Newton Algorithms for Medical Image Registration. In Proceedings of the World Congress on Medical Physics and Biomedical Engineering, Munich, Germany, 7–12 September 2009; Dössel, O., Schlegel, W.C., Eds.; IFMBE Proceedings; Springer: Berlin/Heidelberg, Germany, 2019; Volume 25/4. [\[CrossRef\]](#)
6. Ludwig, A. The Gauss–Seidel–quasi-Newton method: A hybrid algorithm for solving dynamic economic models. *J. Econ. Dyn. Control.* **2007**, *31*, 1610–1632. [\[CrossRef\]](#)
7. Wülfingen, G.B. On some advantages of the application of Newton’s method for the solution of nonlinear economic models. In Proceedings of the IFAC Dynamic Modelling, Warsaw, Poland, 16–19 June 1980; pp. 339–347.
8. Schaefer, B.; Ghasemi, S.A.; Roy, S.; Goedecker, S. Stabilized quasi-Newton optimization of noisy potential energy surfaces. *J. Chem. Phys.* **2015**, *142*, 034112. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Kemeny, J.G.; Snell, J.L. *Mathematical Models in the Social Sciences. Introduction to Higher Mathematics*; Blaisdell Publishing Company, A Division of Ginn and Company: New York, NY, USA; Toronto, ON, Canada; London, UK, 1963; Volume VII, 145p.
10. Beregi, S.; Barton, D.A.W.; Rezgui, D.; Nield, S.A. Real-Time Hybrid Testing Using Iterative Control for Periodic Oscillations, arXiv. Available online: <https://arxiv.org/abs/2312.06362> (accessed on 15 December 2023).
11. Barnafi, N.A.; Pavarino, L.F.; Scacchi, S. Parallel inexact Newton-Krylov and Quasi-Newton Solvers for Nonlinear Elasticity, arXiv. Available online: <https://arxiv.org/abs/2203.05610> (accessed on 15 December 2023).
12. Ortega, J.M.; Rheinboldt, W.C. *Iterative Solution of Nonlinear Equations in Several Variables*; Academic Press: New York, NY, USA, 1970.
13. Dennis, J.E., Jr.; Schnabel, R.B. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1983.
14. Stoer, J.; Bulirsch, R. *Introduction to Numerical Analysis*, 3rd ed.; Springer: Berlin/Heidelberg, Germany, 2002.
15. Martinez, J.M.; Qi, L. Inexact Newton methods for solving nonsmooth equations. *J. Comput. Appl. Math.* **1995**, *60*, 127–145. [\[CrossRef\]](#)
16. Dembo, R.S.; Eisenstat, S.C.; Steihaug, T. Inexact Newton methods. *SIAM J. Numer. Anal.* **1971**, *19*, 400–408. [\[CrossRef\]](#)
17. Birgin, E.G.; Kojic, N.; Martinez, J.M. Globally convergent inexact quasi-Newton methods for solving nonlinear systems. *Numer. Algorithms.* **2003**, *32*, 249–260. [\[CrossRef\]](#)
18. Martínez, J.M. Practical quasi-Newton methods for solving nonlinear systems. *J. Comput. Applied Math.* **2000**, *124*, 97–121. [\[CrossRef\]](#)
19. Wolfe, P. The Secant Method for Simultaneous Nonlinear Equations. *Commun. ACM* **1959**, *2*, 12–13. [\[CrossRef\]](#)
20. Popper, G. Numerical method for least square solving of nonlinear equations. *Period. Polytech.* **1985**, *29*, 67–69.

21. Berzi, P. Model investigation for pile bearing capacity prediction. In Proceedings of the Euromech (280) Symposium on Identification of Nonlinear Mechanical Systems from Dynamic Tests, Ecully, France, 29–31 October 1991.
22. Berzi, P. Pile-Soil Interaction due to Static and Dynamic Load. In Proceedings of the XIII ICSMFE, New Delhi, India, 5–10 January 1994; pp. 609–612.
23. Berzi, P.; Beccu, R.; Lundberg, B. Identification of a percussive drill rod joint from its response to stress wave loading. *Int. J. Impact Eng.* **1994**, *18*, 281–290. [\[CrossRef\]](#)
24. Rosenbrock, H.H. An automatic Method for finding the Greatest or Least Value of a Function. *Comput. J.* **1960**, *3*, 175–184. [\[CrossRef\]](#)
25. Traub, J.F. *Iterative Methods for the Solution of Equations*, 1st ed.; Prentice-Hall, Inc.: Englewood Cliffs, NJ, USA, 1964.
26. Ostrowski, A.M. *Solution of Equations and Systems of Equations*; Academic Press: New York, NY, USA, 1966.
27. Broyden, C.G. A class of Methods for Solving Nonlinear Simultaneous Equations. *Math. Comput. Am. Math.* **1965**, *19*, 577–593. [\[CrossRef\]](#)
28. Powell, M.J.D. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *Comput. J.* **1964**, *7*, 155–162. [\[CrossRef\]](#)
29. Loshchilov, I.; Schoenauer, M.; Sebag, M. Adaptive Coordinate Descent. In Proceedings of the Genetic and Evolutionary Computation Conference (GECCO), Dublin, Ireland, 12–16 July 2011; ACM Press: New York, NY, USA, 2011; pp. 885–892.
30. Nelder, J.A.; Mead, R. A simplex method for function minimization. *Comput. J.* **1965**, *7*, 308–313. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.