2008 International Symposium on Nonlinear Theory and its Applications
NOLTA'08, Budapest, Hungary, September 7-10, 2008

A Representation by Power Series for the Sequence Generated by the Simplified Newton's Method

Takashi Ozeki†

†Department of Information Processing, Faculty of Engineering, Fukuyama University
1 Sanzo, Gakuen-cho Fukuyama, Hiroshima, 729-0292 Japan
Email: ozeki@fuip.fukuyama-u.ac.jp

Abstract—We discuss a property of the sequence of approximations obtained by the simplified Newton's method. In the case of nonlinear functions with one variable, it has been proved that approximations of the simplified Newton's method can be represented by a power series. Therefore, the convergence of the sequence can be accelerated by the ϵ algorithm and others. In this paper, we try to extend the result to several variables. It is shown that approximations of the simplified Newton's method for several variables can be represented in the same way of one variable if those eigenvalues don't idempotent each other. Moreover, in the case of more than two variables, we show exceptions that can't be represented by any power series with constant vectors.

1. Introduction

It is difficult to solve nonlinear equations analytically because they have complicated structure in general. Therefore, approximations of these equations are gotten by numerical calculations. Most of those methods refine approximations by iterations. For example, the Newton's method is one of famous and effective iterative methods. However, the method needs much times to obtain Jacobian matrices and requires differential calculations of nonlinear functions at each iteration. Hence, to omit calculating Jacobian matrices, many quasi Newton's methods that use an approximation matrix of the Jacobian Matrix have been proposed [1, 2]. Those methods have complicated procedures and require many memories in large problems. On the other hand, the simplified Newton's method [3], which doesn't calculate the Jacobian matrix at every time and uses a constant matrix in the Newton's method, doesn't need much calculate time and many memories at each iteration. This method has a weak point that the convergence is linear and takes much time to get solutions. However, if nonlinear functions have one variable, the sequence of approximations obtained by the method is represented by a power series [4]. Therefore, the sequence can be accelerated by the ϵ algorithm [5, 6] and the Limit Estimation [7] and others. In this paper, we try to extend the result to several variables. It is shown that approximations of the simplified Newton's method for nonlinear functions of several variables can be represented by a power series under some conditions. Moreover, in the case of more than two variables, it is shown that there exist exceptions that can't be represented by any power series with constant vectors. In the following sections, we discuss the case of only two variables to avoid complicated notations but all results hold in more than three variables.

2. Expression for Solutions By Power Series

The simplified Newton's method to solve nonlinear simultaneous equations

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$$

with two variables is given by the next iteration

$$\boldsymbol{x}_{n+1} = \boldsymbol{x}_n - \boldsymbol{G}\boldsymbol{f}(\boldsymbol{x}_n). \tag{1}$$

Here, $\boldsymbol{G}$ is a constant matrix with the size of 2×2 and an approximation of the inverse matrix of the Jacobian matrix $\boldsymbol{f}'(\boldsymbol{x})$ for the given function $\boldsymbol{f}(\boldsymbol{x})$. For example, it obtains from the matrix as follows:

$$\boldsymbol{G} = \boldsymbol{f}'(\boldsymbol{x}_0)^{-1}.$$

To simplify notations, a new function

$$\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{x} - \boldsymbol{G}\boldsymbol{f}(\boldsymbol{x})$$

is introduced. Then, the Simplified Newton's method is expressed by a simple iteration

$$\boldsymbol{x}_{n+1} = \boldsymbol{g}(\boldsymbol{x}_n). \tag{2}$$

Now, we discuss a property of the sequence $\{\boldsymbol{x}_n\}$. We assume that the function $\boldsymbol{g}(\boldsymbol{x})$ is differential any times around a fixed point $\boldsymbol{p}$. Then, the nonlinear function $\boldsymbol{g}(\boldsymbol{x})$ can be represented by a Taylor series like

$$\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{p} + \sum_{\substack{i+j=1 \\ 0 \le i,j}}^{\infty} (x-p)^i (y-q)^j \begin{pmatrix} a_{ij} \\ b_{ij} \end{pmatrix} \tag{3}$$

with constant vectors $\begin{pmatrix} a_{ij} \\ b_{ij} \end{pmatrix}$. Here, the vectors are given by

$$\begin{pmatrix} a_{ij} \\ b_{ij} \end{pmatrix} = \left(\begin{array}{c} \frac{1}{(i+j)!} \cdot \frac{\partial^{i+j} g_1}{\partial x^i \partial y^j}(\boldsymbol{p}) \\ \frac{1}{(i+j)!} \cdot \frac{\partial^{i+j} g_2}{\partial x^i \partial y^j}(\boldsymbol{p}) \end{array} \right).$$

The sequence $\{\boldsymbol{x}_n\}$ converges to the fixed point $\boldsymbol{p}$ if the given initial approximation $\boldsymbol{x}_0$ is sufficiently close to the fixed point $\boldsymbol{p}$ and the function $\boldsymbol{g}(\boldsymbol{x})$ is a contraction mapping. In other words, two eigenvalues α, β of the matrix

$$\left(\begin{array}{cc} a_{10} & a_{01} \\ b_{10} & b_{01} \end{array} \right),$$

which is made from the coefficients with one degree of the Taylor series (3), satisfy the following condition

$$0 < |\alpha|, |\beta| < 1.$$

2.1. Non-idempotent Eigenvalues

The sequence of the simplified Newton's method has a following property if those eigenvalues don't have a idempotent relation each other.

Theorem 1 Let non-idempotent conditions $\alpha^m \neq \beta^n (m, n \geq 1)$ satisfy. Then, the sequence $\{\boldsymbol{x}_n\}$ obtained from the iteration (2) can be represented by a power series of two eigenvalues α, β and constant vectors $\begin{pmatrix} A_{kl} \\ B_{kl} \end{pmatrix}$ as follows:

$$\boldsymbol{x}_n = \boldsymbol{p} + \sum_{\substack{k+l=1 \\ 0 \leq k,l}}^{\infty} \alpha^{kn} \beta^{ln} \begin{pmatrix} A_{kl} \\ B_{kl} \end{pmatrix} \tag{4}$$

for all $n \geq N$. Here, N is a sufficient large integer.

[Remark] In this theorem, we can't change the condition $n \geq N$ to $n \geq 0$ Because, sometimes, it happens $\sum_{k,l}^{\infty} A_{kl} = \infty$ or $\sum_{k,l}^{\infty} B_{kl} = \infty$. However, if the initial solution $\boldsymbol{x}_0$ is given sufficiently close to the fixed point $\boldsymbol{p}$, it holds

$$\boldsymbol{x}_0 = \boldsymbol{p} + \sum_{k+l=1}^{\infty} \begin{pmatrix} A_{kl} \\ B_{kl} \end{pmatrix}.$$

[Outline of Proof] In the following proof, we can put $\boldsymbol{p} = \boldsymbol{0}$ by the transformation of the coordinate system where the fixed point $\boldsymbol{p}$ displaces to the origin $\boldsymbol{0}$. Moreover, from the assumption of Theorem 1, the matrix

$$\left(\begin{array}{cc} a_{10} & a_{01} \\ b_{10} & b_{01} \end{array} \right)$$

has different two eigenvalues. Therefore, it can be transformed to a diagonal matrix by a regular matrix P like

$$\left(\begin{array}{cc} a_{10} & a_{01} \\ b_{10} & b_{01} \end{array} \right) = P^{-1} \left(\begin{array}{cc} \alpha & 0 \\ 0 & \beta \end{array} \right) P.$$

So, by a suitable transformation of the coordinate system, we can put

$$\left(\begin{array}{cc} a_{10} & a_{01} \\ b_{10} & b_{01} \end{array} \right) = \left(\begin{array}{cc} \alpha & 0 \\ 0 & \beta \end{array} \right) \tag{5}$$

from the beginning of this discussions without loss of generality.

Now, let assume the expression

$$\boldsymbol{x}_n = \begin{pmatrix} x_n \\ y_n \end{pmatrix} = \sum_{k+l=1}^{\infty} \alpha^{kn} \beta^{ln} \begin{pmatrix} A_{kl} \\ B_{kl} \end{pmatrix}. \tag{6}$$

Then, arranging the next approximation $\boldsymbol{x}_{n+1}$ in formal power series for two eigenvalues α, β, from (2), (3) and (6), we obtain

$$\begin{aligned} \boldsymbol{x}_{n+1} &= \sum_{i+j=1}^{\infty} x_n^i y_n^j \begin{pmatrix} a_{ij} \\ b_{ij} \end{pmatrix} \\ &= \sum_{k+l=1}^{\infty} \alpha^{kn} \beta^{ln} \begin{pmatrix} C_{kl} \\ D_{kl} \end{pmatrix} \end{aligned} \tag{7}$$

where constant vectors $\begin{pmatrix} C_{kl} \\ D_{kl} \end{pmatrix}$ are gotten by using A_{kl}, B_{kl} and determined by

$$\begin{pmatrix} C_{kl} \\ D_{kl} \end{pmatrix} = \left(\begin{array}{c} \sum_{i+j=1}^{k+l} a_{ij} \prod_{p=1}^{i} A_{k_p l_p} \prod_{q=i+1}^{i+j} B_{k_q l_q} \\ \sum_{\substack{i+j=1 \\ 0 \leq i,j}} b_{ij} \prod_{p=1}^{i} A_{k_p l_p} \prod_{q=i+1}^{i+j} B_{k_q l_q} \end{array} \right).$$

Here, indexes k_p, l_p satisfy the following conditions

$$\begin{aligned} \sum_{p=1}^{i+j} k_p &= k \quad (0 \leq k_p \leq k) \\ \sum_{p=1}^{i+j} l_p &= l \quad (0 \leq l_p \leq l). \end{aligned}$$

To match the equation (7) with the equation

$$\boldsymbol{x}_{n+1} = \sum_{k+l=1}^{\infty} \alpha^{k(n+1)} \beta^{l(n+1)} \begin{pmatrix} A_{kl} \\ B_{kl} \end{pmatrix}$$

for any n, it is needed to satisfy

$$\alpha^{k(n+1)} \beta^{l(n+1)} \begin{pmatrix} A_{kl} \\ B_{kl} \end{pmatrix} = \alpha^{kn} \beta^{ln} \begin{pmatrix} C_{kl} \\ D_{kl} \end{pmatrix}$$

for all $k+l \geq 1$. From the comparison of coefficients of $\alpha^{kn}\beta^{ln}$ and (5), we obtain

$$\begin{pmatrix} \alpha-\alpha & 0 \\ 0 & \alpha-\beta \end{pmatrix}\begin{pmatrix} A_{10} \\ B_{10} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

for $k=1, l=0$. Moreover, since $\alpha \neq \beta$, we get

$$B_{10} = 0.$$

In the same way, we obtain

$$\begin{pmatrix} \alpha-\beta & 0 \\ 0 & \beta-\beta \end{pmatrix}\begin{pmatrix} A_{01} \\ B_{01} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and

$$A_{01} = 0$$

for $k=0, l=1$. In other cases, it holds the equation

$$\begin{pmatrix} \alpha^k\beta^l-\alpha & 0 \\ 0 & \alpha^k\beta^l-\beta \end{pmatrix}\begin{pmatrix} A_{kl} \\ B_{kl} \end{pmatrix} = \begin{pmatrix} C_{kl} \\ D_{kl} \end{pmatrix}.$$

Here, since $\alpha^i \neq \beta^j$, the left matrix has an inverse one and we obtain the next recurrent formula

$$\begin{pmatrix} A_{kl} \\ B_{kl} \end{pmatrix} = \begin{pmatrix} \alpha^k\beta^l-\alpha & 0 \\ 0 & \alpha^k\beta^l-\beta \end{pmatrix}^{-1}\begin{pmatrix} C_{kl} \\ D_{kl} \end{pmatrix}. \quad (8)$$

From this formula, the values of A_{kl} and B_{kl} are obtained by $A_{ij}, B_{ij} (i+j < k+l)$. As a result, every A_{kl} and B_{kl} can be calculated from only two coefficients A_{10} and B_{01}. Those are determined by the initial approximation $\boldsymbol{x}_0$.

[Remark] When a formula power series is expanded, the order of sum for double series is changed arbitrarily. If these double series converges absolutely, we can change the order of sum arbitrarily [8]. However, we omit the proof for want of space.

If given functions have only one variable, there is no idempotent eigenvalues obviously. Because there is only one eigenvalue in this case. Therefore, we immediately get the next corollary [4].

Corollary 1 The sequence of approximations generated by the simplified Newton's method applying to nonlinear functions with one variable is represented by a power series as follows

$$x_n = p + \sum_{i=1}^{\infty} A_i \lambda^{in}$$

for every $n \geq N$. Here, N is a sufficient large integer.

In order to verify Theorem 1, we practice an example.

Example 1 Let a nonlinear function $\boldsymbol{g}(\boldsymbol{x})$ be

$$\boldsymbol{g}(\boldsymbol{x}) = \begin{pmatrix} \alpha x \\ \beta y + xy \end{pmatrix}.$$

Here, it holds $0 < |\alpha|, |\beta| < 1$. Then, the function can be expanded to Taylor series at the origin $\mathbf{0}$ as follows:

$$\boldsymbol{g}(\boldsymbol{x}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + x\begin{pmatrix} \alpha \\ 0 \end{pmatrix} + y\begin{pmatrix} 0 \\ \beta \end{pmatrix} + xy\begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Therefore, we get

$$\begin{cases} a_{10} = \alpha, \quad b_{01} = \beta, \quad b_{11} = 1 \\ a_{ij} = b_{ij} = 0 \quad \text{(Others)}. \end{cases}$$

Since it holds $A_{01} = B_{10} = 0$ and from the recurrent formula (8), we obtain

$$A_{kl} = 0 \quad (A_{kl} \neq A_{10})$$

$$B_{kl} = \frac{A_{10}B_{k-1l}}{\alpha^k\beta^l - \beta} = \begin{cases} 0 & (l \neq 1) \\ \dfrac{A_{10}^k B_{01}}{\prod_{i=1}^{k}(\alpha^i\beta - \beta)} & (l = 1) \end{cases}.$$

Therefore, we get a power series as follows:

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} A_{10}\alpha^n \\ B_{01}\beta^n + B_{11}\alpha^n\beta^n + B_{21}\alpha^{2n}\beta^n + \cdots \end{pmatrix}. \quad (9)$$

From now on, we verify this power series by numerical calculations. Let put

$$\alpha = 0.5, \beta = 0.3$$

and

$$\boldsymbol{x}_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Then, from the power series (9), we get $A_{10} = 1$ and $B_{01} = \lim_{n\to\infty} \frac{y_n}{\beta^n}$. Table 1 shows the second variables y_n of the sequence $\{\boldsymbol{x}_n\}$ of (2) and the sequence $\{\tilde{\boldsymbol{x}}_n\}$ of (9). Here, to seek the value of $\tilde{\boldsymbol{x}}_n$ from (9), we need the value of B_{01}. Therefore, we use the approximation obtained by

$$B_{01} \approx \frac{y_{100}}{\beta^{100}} = 88.7212166719551192.$$

Also, in the Table 1, the value of $\tilde{y}_n$ is obtained by the sum of first hundred powers as follows:

$$\tilde{y}_n \approx \sum_{i=0}^{100} B_{i1}\alpha^{in}\beta^n.$$

The table shows that the values of y_n equal to the ones of $\tilde{y}_n$ for $n \geq 2$. Therefore, we can confirm that Theorem 1 holds. However, for $n = 0, 1$, two values of the sequences are different. The reason is that these formula power series don't converge absolutely and can't change the order of sum for double series. Nevertheless, we can use Wynn's acceleration method [6] because the sequence is represented by a power series. Table 2 shows the accelerated result for the sequence $\{\boldsymbol{x}_n\}$.

Table 1: Comparison of Solutions by the Simplified Newton's Method and by Infinite Power Series.

n	y_n	$\tilde{y}_n$
0	2.00000000000000	-1.37562681 $\times 10^{54}$
1	2.60000000000000	-5.29024380 $\times 10^{23}$
2	2.08000000000000	2.07999981789382
3	1.14399999999999	1.14399999999999
4	0.48620000000000	0.48619999999999
5	0.17624750000000	0.17624749999999
6	0.05838198437500	0.05838198437499
7	0.01842681381835	0.01842681381835
8	0.00567200362846	0.00567200362846
9	0.00172375735271	0.00172375735271
10	0.00052049391939	0.00052049391939
∞	0	0

Table 2: Acceleration by Wynn's method.

n	$\boldsymbol{x}_n$	ϵ_2	ϵ_4
0	(1.00, 2.00)		
1	(0.50, 2.60)	(0.51, 2.20)	
2	(0.25, 2.08)	(1.21, 2.88)	(0.36, 0.96)
3	(0.12, 1.14)	(0.11, -1.04)	(0.02, 0.16)
4	(0.06, 0.48)	(0.00, -0.09)	(0.00, 0.01)
5	(0.03, 0.17)	(0.00, -0.01)	
6	(0.01, 0.05)		
∞	(0, 0)	(0, 0)	(0, 0)

2.2. Idempotent Eigenvalues

When eigenvalues are idempotent each other, Theorem 1 doesn't holds. Now, we show such an example.

Example 2 Let a nonlinear function $\boldsymbol{g}(\boldsymbol{x})$ be

$$\boldsymbol{g}(\boldsymbol{x}) = \begin{pmatrix} \alpha x \\ \beta y + x^2 \end{pmatrix}$$

where $\beta = \alpha^2$ and $0 < |\alpha| < 1$ hold. From

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} \alpha x_n \\ \alpha^2 y_n + x_n^2 \end{pmatrix},$$

we obtain two power series as follows:

$$\begin{aligned} x_n &= \alpha^n x_0 \\ y_n &= \alpha^2 y_{n-1} + \alpha^{2(n-1)} x_0^2 \\ &= \alpha^2(\alpha^2 y_{n-2} + \alpha^{2(n-2)} x_0^2) + \alpha^{2(n-1)} x_0^2 \\ &= y_0 \alpha^{2n} + n x_0^2 \alpha^{2(n-1)} \\ &= y_0 \beta^n + n x_0^2 \alpha^{2(n-1)}. \end{aligned}$$

Here, the coefficient $n x_0^2$ of the power $\alpha^{2(n-1)}$ in the second term of y_n isn't constant and changes with n. Therefore, in this case, we can't express the sequence of approximations by power series (4). However, even in those bad cases, we can expect that the sequence of approximations are represented by the extended formula of power series like

$$\boldsymbol{x}_n = \boldsymbol{p} + \sum_{\substack{k+l=1 \\ 0 \le k,l}}^{\infty} \alpha^{kn} \beta^{ln} \begin{pmatrix} p_{kl}(n) \\ q_{kl}(n) \end{pmatrix} \quad (10)$$

where constant coefficients A_{kl} and B_{kl} are replaced by reasonable polynomials $p_{kl}(n)$ and $q_{kl}(n)$.

3. Conclusions

In this paper, we discussed a property of the sequence of approximations obtained by the simplified Newton's method. In the same way as nonlinear functions with one variable, it was shown that approximations of the simplified Newton's method for several variables can be represented by a power series if those eigenvalues don't idempotent each other. Also, in the case of functions having more than two variables, there existed an exception that can't be represented by any power series with constant vectors. However, in those cases, it is expected that nonlinear functions can be expressed by the extended power series with polynomials. The proof is left to the future works.

References

[1] S. Suzuki, H. Hiroshi, Y. Iida, T. Takashi and M. Tanaka, "Current Numerical Method", Ohm, 1994, pp.23–26.

[2] M. Iri, "Numerical Calculation", Asakura, pp.29–34, 1981.

[3] S. Oishi, "Numerical Calculation", Shoukabo, p.103, 1999.

[4] T. Ozeki and T. Iijima, "How does Parallel Code Method converge to a root of a nonlinear equation?", *Proceedings of NOLTA97*, 1997.

[5] P. Wynn, "On a Device for Computing the $e_m(S_n)$ Transformation", *Math. Tables Aids Comp.*, No. 10, pp. 91–96, 1956.

[6] P. Wynn, "Acceleration Techniques for Iterated Vector and Matrix Problems", *Math. of Comp.*, No. 16, pp. 301–322, 1962.

[7] T. Iijima, "A Method for the Limit Estimation of Convergent Sequence of Numbers," *IEICE Trans. Fund.*, J79-A, No.4, pp.947–953, 1996.

[8] M. Sugiura, "Introduction to Analysis I", University of Tokyo, pp.382–388, 1980.