数理解析研究所講究録
第 453 巻 1982 年 174-196

The Subroutine Package NAES for the Solution of a System of Nonlinear Equations —— Deflation Algorithm

Takeo Ojika (Osaka Kyoiku Univ.)
Satoshi Watanabe (Yamagata Univ.)
Taketomo Mitsui (RIMS, Kyoto Univ.)

Abstract. The authors have been developing a subroutine package NAES (Nonlinear Algebraic Equations Solver)[14] for the numerical solution of the system of nonlinear equations. The purpose of the paper is to present an algorithm, in the package, termed here as the deflation algorithm, for determining multiple roots for a system of nonlinear equations, and to show the effectiveness of the algorithm by solving a numerical example.

I. Introduction

The purpose of this paper is to present a method, termed here as the deflation algorithm, for finding roots of a system of nonlinear equations

$$F(x) = 0, \qquad F: \ R^n \to R^n, \tag{1.1}$$

where the Jacobian matrix F_x of F is singular at the root x^*, i.e.,

$$F_x(x^*) = 0. \tag{1.2}$$

If the Jacobian matrix F_x at a root $x^* \in F^{-1}(0)$ is nonsingular, it is well known that the Newton (also Newton-Raphson) iteration:

$$^{k+1}x = {}^{k}x - [F_x({}^{k}x)]^{-1}F({}^{k}x), \qquad k = 0,1,2,... \tag{1.3}$$

converges to X^* from any initial guess ${}^0X$ in a sufficiently small ball centered at X^* [1-6].

However, in the case of a singular Jacobian matrix $F_x(X^*)$, the classical theory is not applicable and except for the one dimensional problem only few results are available [7-13].

In this paper, several properties of the multiple roots of a system of nonlinear equations are studied first. Then a practical algorithm is proposed to determine the multiple root. Finally the effectiveness of the proposed algorithm is illustrated by a numerical example.

II. Newton Method for Multiple Roots

2.1. The One-Dimensional Case

It is instructive to consider first the one-dimensional case of a real-valued function f of a real variable x, i.e.,

$$f(x) = 0 . \tag{2.1}$$

In general, a root x^* of the nonlinear equation (2.1) is said to have multiplicity m if

$$f(x) = (x - x^*)^m \bar{f}(x), \qquad 0 \neq |\bar{f}(x^*)| < \infty, \tag{2.2}$$

where $\bar{f}(x)$ is twice continuously differentiable at the root x^*.

Starting from an initial guess ${}^0x$ in a neighborhood $D \in R$ of x^*, the Newton method defines the sequence of approximations

$${}^{k+1}x = {}^kx + \Delta^k x, \qquad \Delta^k x = -[f_x({}^kx)]^{-1} f({}^kx), \qquad k = 0,1,2,\ldots \tag{2.3}$$

Let ${}^k\eta$ be the error of ${}^kx$ from x^*, i.e.,

$${}^{k}\eta = {}^{k}x - x^*. \tag{2.4}$$

Then, from (2.2) and (2.3), we have

$${}^{k+1}\eta = {}^{k}\eta - \frac{({}^{k}\eta)^m \bar{f}({}^{k}x)}{m({}^{k}\eta)^{m-1}\,\bar{f}({}^{k}x)\{1 + \frac{{}^{k}\eta}{m\bar{f}({}^{k}x)}\,\bar{f}_x({}^{k}x)\}}\ . \tag{2.5}$$

From (2.5) and Taylor's theorem, it follows that

$${}^{k+1}\eta = (\frac{m-1}{m}){}^{k}\eta + O({}^{k}\eta^2). \tag{2.6}$$

Consequently, if the sequence $\{{}^{k}x\}$ is convergent to x^*, the sequence $\{{}^{k}\eta\}$ will converge to 0 with the speed of a geometric progression with ratio $(m-1)/m$. If $m = 1$, it follows from (2.6) that

$${}^{k+1}\eta = O({}^{k}\eta^2), \tag{2.7}$$

and the sequence $\{{}^{k}x\}$ is said to converge quadratically to x^* which is called a simple root; if $m > 1$, the convergence is said to be geometric, with ratio $(m-1)/m$.

From the above discussion, we now have the following theorem [16, 17].

Theorem 2.1. If the sequence $\{{}^{k}x\}$ defined by (2.3) converges to x^*, and $\bar{f}(x)$ has a Taylor series expansion at x^* which converges in some neighborhood D of x^*, then the following asymptotic relations hold:

$$\text{(i)} \qquad \lim_{k\to\infty} \frac{f({}^{k+1}x)}{f({}^{k}x)} = (\frac{m-1}{m})^m , \tag{2.8a}$$

$$\text{(ii)} \qquad \lim_{k\to\infty} \frac{f_x({}^{k+1}x)}{f_x({}^{k}x)} = \begin{cases} 1, & \text{if } m = 1, \\ (\frac{m-1}{m})^{m-1}, & \text{if } m \geqq 2. \end{cases} \tag{2.8b}$$

Proof. From (2.2), (2.4), (2.6) and a Taylor series expansion, we have

$$\frac{f({}^{k+1}x)}{f({}^{k}x)} = \left(\frac{m-1}{m}\right)^m \cdot \frac{\bar{f}(x^*) + \left(\frac{m-1}{m}\right)\bar{f}_x(x^*)\,{}^{k}\eta + O({}^{k}\eta^2)}{\bar{f}(x^*) + \bar{f}_x(x^*)\,{}^{k}\eta + O({}^{k}\eta^2)}$$

$$= \left(\frac{m-1}{m}\right)^m\{1 + O({}^{k}\eta)\}. \tag{2.9}$$

Since the sequence $\{{}^{k}x\}$ is convergent to x^* as $k \to \infty$, the sequence $\{{}^{k}\eta\}$ is also convergent to 0. Hence, from (2.9), we have (2.8a).

Differentiating (2.2) with respect to x, it follows that

$$f_x({}^{k+1}x) = ({}^{k+1}\eta)^{m-1}\{m\bar{f}({}^{k+1}x) + {}^{k+1}\eta\bar{f}_x({}^{k+1}x)\}$$

$$= m({}^{k+1}\eta)^{m-1}\bar{f}(x^*)\{1 + O({}^{k}\eta)\}. \tag{2.10}$$

From (2.6) and (2.10), we have

$$\frac{f_x({}^{k+1}x)}{f_x({}^{k}x)} = \left(\frac{m-1}{m}\right)^{m-1}\{1 + O({}^{k}\eta)\} \tag{2.11}$$

which shows that since the sequence $\{{}^{k}\eta\}$ is convergent to 0, (2.8b) holds as $k \to \infty$. Thus the proof of the theorem are complete.

We now have the following corollary which is the same result previously obtained by Rall [12].

Corollary 2.1. Suppose that the conditions of the Theorem 2.1 hold. Then, from (2.3), we have

$$\lim_{k\to\infty} \frac{\Delta^{k+1}x}{\Delta^{k}x} = \frac{m-1}{m}\,. \tag{2.12}$$

Proof of the corollary is obvious from (2.3), (2.8a) and (2.8b).

From Theorem 2.1 and Corollary 2.1, it is easily seen that some properties of the approximate solution ${}^{k}x$ in the neighborhood of the simple root, multiple root or singular manifold are given in Table 2.1. The convergence tendencies to the simple root $(m = 1)$ and the multiple root $(m \geqq 2)$ are shown schematically in Fig. 2.1.

Table 2.1 Properties of roots

	simple	multiple	singular
$\lvert f({}^{k+1}x) \rvert$	$<< 1$	$<< 1$	$>> 1$
$Y = \lvert f({}^{k+1}x)/f({}^{k}x) \rvert$	$<< 1$	$(\frac{m-1}{m})^{m}$	$>> 1$
$\lvert d({}^{k+1}x) \rvert$	$\bar{f}(x^*)$	$<< 1$	$<< 1$
$z = \lvert d({}^{k+1}x)/d({}^{k}x) \rvert$	1	$(\frac{m-1}{m})^{m-1}$	$<< 1$
$x = \lvert \Delta^{k+1}x/\Delta^{k}x \rvert$	$<< 1$	$(\frac{m-1}{m})$	$>> 1$

On the other hand, from (2.2), we have

$$x - x^* = [\frac{f(x)}{\bar{f}(x)}]^{1/m}, \qquad m \geqq 1 \tag{2.13}$$

which shows that if one wished to calculate x^* to a single precision accuracy on a computer, one must compute the value of $f(x)$ using *m-precision arithmetic*. Otherwise, $f({}^{k}x) = f(x^* + {}^{k}\eta)$ can vanish before ${}^{k}\eta$ becomes negligible with respect to the accuracy desired in x^*, thus terminating the Newton process prematurely.

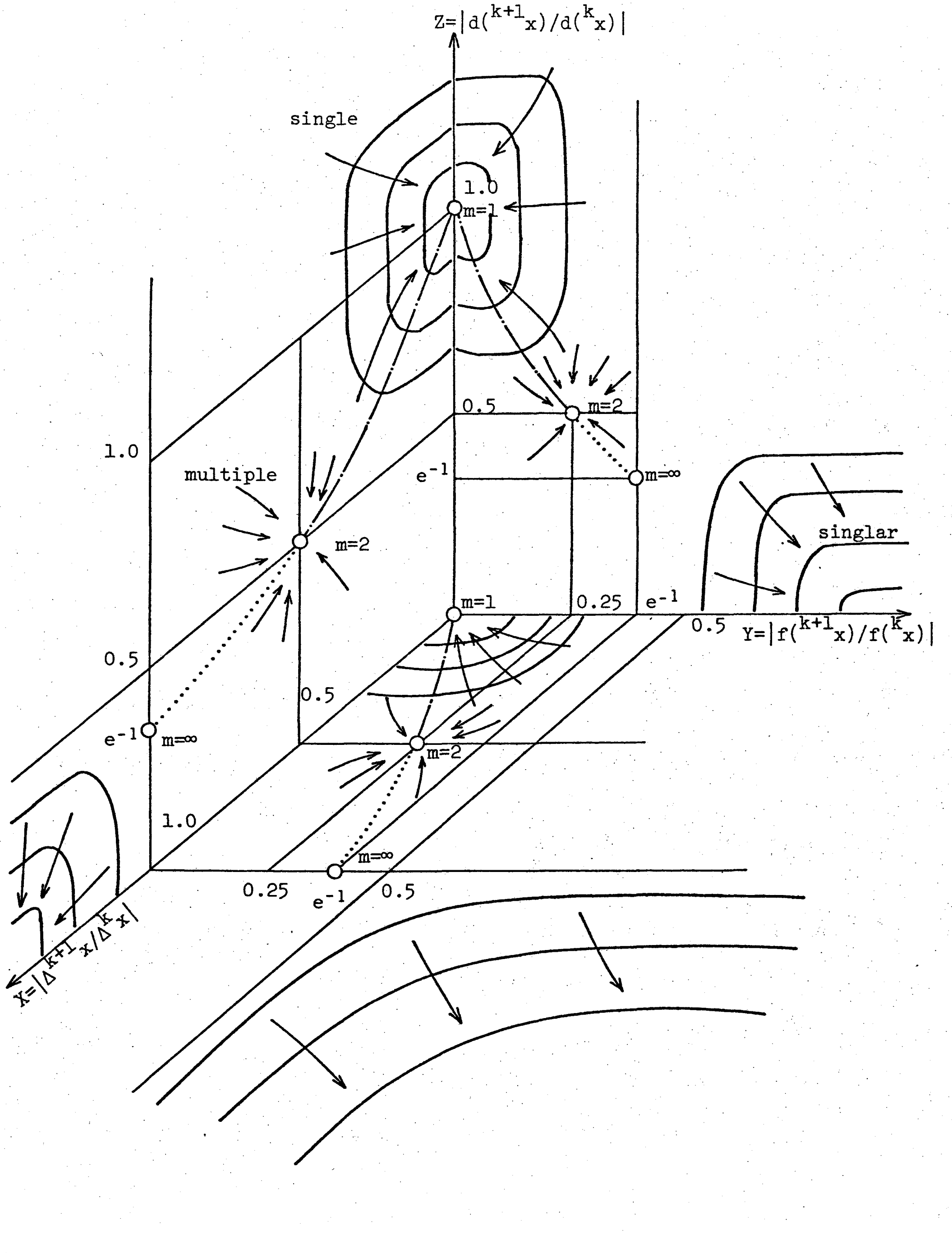

Fig. 2.1. Properties of roots

2.2. Systems of Nonlinear Equations

We now return to the problem of finding multiple roots of a vector function. For simplicity, let $F : R^2 \to R^2$, and consider the following nonlinear equations:

$$F(X) = \begin{pmatrix} f(X) \\ \\ g(X) \end{pmatrix} = \begin{pmatrix} (x + a_1 y - b_1)^{m_1} \bar{f}(X) \\ \\ (a_2 x + y - b_2)^{m_2} \bar{g}(X) \end{pmatrix} = 0, \quad m_1, m_2 \geqq 2, \tag{2.14}$$

$$X = (x,y), \quad 0 \neq |\bar{f}(X^*)|, \quad |\bar{g}(X^*)| < \infty,$$

where

$$\det \begin{bmatrix} 1 & a_1 \\ a_2 & 1 \end{bmatrix} = 1 - a_1 a_2 \neq 0, \tag{2.15}$$

and $\bar{f}$ and $\bar{g}$ are m_1 and m_2 times continuously differentiable in a neighborhood of a root X^* which satisfies the following relations:

$$x^* + a_1 y^* - b_1 = a_2 x^* + y^* - b_2 = 0. \tag{2.16}$$

Taking (2.15) and (2.16) into account, from (2.14), we have, say, for x:

$$x - x^* = \det \begin{bmatrix} (f(X)/\bar{f}(X))^{1/m_1} & a_1 \\ \\ (g(X)/\bar{g}(X))^{1/m_2} & 1 \end{bmatrix} \Big/ (1 - a_1 a_2). \tag{2.17}$$

This fact shows that if one wished to calculate x^* to a single precision accuracy on a computer, one must compute the values of $f(x)$ and $g(x)$ using m_1- and m_2- precision arithmetics, respectively.

In what follows, we denote, say, by $\partial_x^j f$ the j-th partial derivative of f with respect to x. Then the chain rule

$$\begin{cases} \partial_x^j f(X) = \sum_{i=0}^{j} \binom{j}{i} \frac{m_1!}{i!} (x + a_1 y - b_1)^{m_1-i} \cdot \partial_x^{j-i} \bar{f}(X), & j=0,1,\ldots,m_1 \\ \partial_y^j g(X) = \sum_{i=0}^{j} \binom{j}{i} \frac{m_2!}{i!} (a_2 x + y - b_2)^{m_2-i} \cdot \partial_y^{j-i} \bar{g}(X), & j=0,1,\ldots,m_2, \end{cases} \tag{2.18}$$

and (2.14) give the following relations:

$$\begin{cases} \partial_x^{m_1-1} f(X) = (x + a_1 y - b_1) U_{m_1-1}(X) = 0, \\ \partial_y^{m_2-1} g(X) = (a_2 x + y - b_2) V_{m_2-1}(X) = 0, \end{cases} \tag{2.19}$$

where

$$\begin{cases} U_j(X) = \partial_x^j f(X)/(x + a_1 y - b_1)^{m_1-j}, & j = 1,2,\ldots,m_1-1, \\ V_k(X) = \partial_y^k g(X)/(a_2 x + y - b_2)^{m_2-k}, & k = 1,2,\ldots,m_2-1. \end{cases} \tag{2.20}$$

From the above, we now have the following [16, 17].

Theorem 2.2. Let $X^* = (x^*, y^*)$ be the solution of (2.14) and satisfy (2.16). Then the following equations hold:

(i) $$\partial_x^{m_1-1} f(X^*) = \partial_y^{m_1-1} f(X^*) = \partial_x^{m_2-1} g(X^*) = \partial_y^{m_2-1} g(X^*) = 0, \tag{2.21}$$

(ii) $$U_{m_1-1}(X^*) = m_1! \bar{f}(X^*) \neq 0, \tag{2.22a}$$

$$V_{m_2-1}(X^*) = m_2! \bar{g}(X^*) \neq 0, \tag{2.22b}$$

(iii) $$\det [J_{ij}(X^*)] = 0, \quad i = 1,2,\ldots,m_1-2; \quad j = 1,2,\ldots,m_2-2, \tag{2.23a}$$

$$\det [J_{m_1-1,m_2-1}(X^*)] = (1 - a_1 a_2) m_1! m_2! \bar{f}(X^*) \bar{g}(X^*) \neq 0, \tag{2.23b}$$

where

$$J_{ij}(X) = \begin{bmatrix} \partial_x^{i+1} f(X) & \partial_y \partial_x^i f(X) \\ \partial_x \partial_y^j g(X) & \partial_y^{j+1} g(X) \end{bmatrix} . \tag{2.23c}$$

Proof. The equation (2.21) can easily be seen from (2.18). The equations (2.22a) and (2.22b) are obvious from the definitions of $U^{(i)}$ and $V^{(j)}$ given by (2.20).

On the other hand, from (2.19), we have

$$\partial_x^{m_1} f(X) = U_{m_1-1}(X) + (x + a_1 y - b_1)\cdot \partial_x U_{m_1-1}(X), \tag{2.24}$$

$$\partial_y \partial_x^{m_1-1} f(X) = a_1 U_{m_1-1}(X) + (x + a_1 y - b_1)\cdot \partial_y U_{m_1-1}(X).$$

Since $(x^* + a_1 y^* - b_1) = 0$, we have

$$U_{m_1-1}(X^*) = m_1!\bar{f}(X^*) \neq 0. \tag{2.25}$$

Similar results can be obtained for $\partial_y^{m_2-1} g(X)$. Substituting these results into (2.23c), we have (2.23a) and (2.23b). Thus the proofs of the theorem are complete.

Since the matrix (2.23c) is nonsingular in the neighborhood of the root X^*, from (2.19), we now have the Newton iteration for multiple root:

$$\begin{pmatrix} {}^{k+1}x \\ {}^{k+1}y \end{pmatrix} = \begin{pmatrix} {}^{k}x \\ {}^{k}y \end{pmatrix} - [J_{m_1-1,m_2-1}({}^kX)]^{-1} \begin{bmatrix} \partial_x^{m_1-1} f({}^kX) \\ \partial_y^{m_2-1} g({}^kX) \end{bmatrix} . \tag{2.26}$$

It suggests that for the singular root of such the type as in (2.14) a convergent Newton iteration may be given by some partial differentiations for the original equations.

We note that, by virtue of (2.19), the sequences $\{^k x\}$ and $\{^k y\}$ generated by (2.26) will converge quadratically to x^* and y^*, respectively, and $m = m_1 \times m_2$ is called here *the multiplicity of the system* of equations given by (2.14).

III. Deflation Algorithm

In the previous section, we studied the multiple roots of a system of nonlinear equations. Let us now propose a practical algorithm, termed here as the deflation algorithm, for determining the multiple roots. For simplicity, let $F\colon R^n \to R^n$, and consider the system of general nonlinear equations given by

$$F^{[0]}(X) = (f_1^{[0]}(X),\ f_2^{[0]}(X),\ldots,\ f_n^{[0]}(X))' = 0,$$
$$X = (x_1,x_2,\ldots,x_n)', \tag{3.1}$$

where the Jacobian matrix $F^{[0]}(X)$ is singular at the root X^*, i.e.,

$$d^{[0]}(X^*) = \det\ [F_x^{[0]}(X^*)] = 0. \tag{3.2}$$

Here $[\cdot]^{[\ell]}$ denotes the value of $[\cdot]$ at the root X^* after the ℓ-th deflation process.

3.1. Deflation Process

The Newton iteration is now given by

$$F_x^{[\ell]}(^kX)(^{k+1}X - {}^kX) = -F^{[\ell]}(^kX), \quad k,\ell = 0,1,2,\ldots\ . \tag{3.3}$$

Assume now that the rank of Jacobian matrix at ^kX in the neighborhood of the root X^* is given by

$$r^* = \operatorname{rank} F_x^{[0]}(x^*), \qquad 0 \leqq r^* \leqq n-1, \tag{3.4a}$$

$$r = r^{[\ell]} = \operatorname{rank} F_x^{[\ell]}({}^k x), \qquad r^* \leqq r^{[\ell]} \leqq n, \tag{3.4b}$$

and that the system of linear equations (3.3) is solved by the familiar *Gaussian elimination method* [1, 6].

For simplicity, suppose ${}^k x_1$ has been eliminated from equations $2,\ldots,n$ of (3.3); hence x_1 remains only in the first equation, and x_1, x_2 from equations $3,\ldots,n$ and so on up to $x_1,\ldots, x_{r^*}$ from equations $r^*+1,\ldots, n$. Then we have the pivot matrix P_r defined by

$$P_r^{[0]}(x^*) = \begin{array}{cc} f_i^{[0]} & x_j^* \\ \left[\begin{array}{c} 1 \\ 2 \\ \vdots \\ r^* \end{array}\right. & \left.\begin{array}{c} 1 \\ 2 \\ \vdots \\ r^* \end{array}\right] \end{array}, \qquad P_r^{[\ell]}(x^*) = \begin{array}{cc} f_i^{[\ell]} & x_j^* \\ \left[\begin{array}{c} 1 \\ 2 \\ \vdots \\ r^{[\ell]} \end{array}\right. & \left.\begin{array}{c} 1 \\ 2 \\ \vdots \\ r^{[\ell]} \end{array}\right] \end{array}, \qquad \ell = 1,2,\ldots, \tag{3.5}$$

and the equation $f_i^{[\ell]}$ and variable x_j in (3.5) are called *the pivot equation and variable*, respectively.

Taking (3.5) into account, let us define an $(r^{[\ell]}+1) \times (r^{[\ell]}+1)$ Jacobian submatrix $D_s^{[\ell]}$, termed as *the deflation matrix*, by

$$D_s^{[\ell]}(x) = \begin{array}{c|cccc} & x_1 & \cdots & x_r & x_s \\ \hline \partial f_1^{[\ell]} & d_{11} & \cdots & d_{1r} & d_{1s} \\ \vdots & \vdots & & \vdots & \vdots \\ \partial f_r^{[\ell]} & d_{r1} & \cdots & d_{rr} & d_{rs} \\ \partial f_s^{[\ell]} & d_{s1} & \cdots & d_{sr} & d_{ss} \end{array}, \qquad s = r+1,\ldots, n, \tag{3.6}$$

where

$$\begin{aligned} d_{ij} &= \partial_{x_j} f_i^{[\ell]}, \qquad i,j = 1,\ldots,r,s, \\ r^* &= r^{[0]}, \qquad r = r^{[\ell]}, \qquad \ell = 1,2,\ldots\ . \end{aligned} \tag{3.7}$$

Note that d_{ij} is given in an analytic form.

At the $(r^{[\ell]} + 1)$st formal elimination stage, the deflation matrix $D_s^{[\ell]}$ is transformed into the form:

$$E_s^{[\ell]}(X) = \begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1r} & e_{1s} \\ & e_{22} & \ddots & & \vdots \\ & & \ddots & e_{rr} & e_{rs} \\ & 0 & & & e_{ss} \end{bmatrix}^{[\ell]}, \quad \begin{array}{l} r = r^{[\ell]} \\ s = r^{[\ell]}+1,\ldots,n. \end{array} \tag{3.8}$$

Here the $(r^{[\ell]}+1) \times (r^{[\ell]}+1)$ upper triangular matrix $E_s^{[\ell]}$ is termed here as *the eliminated matrix* of $D_s^{[\ell]}$. We now have the following

<u>Theorem 3.1.</u> Suppose that $r^* = \text{rank } F_x^{[0]}(X^*)$, $0 \leqq r^* \leqq n-1$ and the pivot matrix $P_r^{[0]}(X^*)$ is given by (3.5). Let $e_{ii}^{[\ell]}(X)$ be the diagonal element of the eliminated matrix $E_s^{[\ell]}(X)$ obtained from the $(r^{[\ell]}+1)\times(r^{[\ell]}+1)$ deflation matrix $D_s^{[\ell]}(X)$. If the approximate solution ${}^kX$ of (3.3) is sufficiently close to the root X^*, then the following properties hold:

(i) $\quad |e_{ii}^{[\ell]}({}^{k+1}X)/e_{ii}^{[\ell]}({}^kX)| \approx 1, \quad i = 1,2,\ldots,r^{[\ell]},$ (3.9a)

(ii) $\quad |e_{s,s}^{[\ell]}({}^kX)| = |\det D_s^{[\ell]}({}^kX)| << 1, \quad s = r^{[\ell]}+1,\ldots,n,$ (3.9b)

(iii) $\quad \det D_s^{[\ell]}(X^*) = 0, \quad \text{if} \quad r^{[\ell]} \neq n,$ (3.9c)

(iv) $\quad 1/e < |e_{ss}^{[\ell]}({}^{k+1}X)/e_{ss}^{[\ell]}({}^kX)| \leqq 1/2, \quad k,\ell = 0,1,\ldots\ .$ (3.9d)

Proofs of the theorem are obvious from Theorem 2.1. The upper and lower bounds in condition (iv) can be easily derived from (2.8b) with $m = 2$ and $m = \infty$, respectively.

At the root X^*, the deflation matrix $D_s^{[\ell]}$ satisfies the relation (3.9c). Taking this fact into account, replace $f_{r+1}^{[\ell]}$ in (3.1) by $\det D_s^{[\ell]}=0$,

$s = r^{[\ell]}+1,\ldots, n$ and define a new set of equations in the next deflated stage by

$$F^{[\ell+1]}(x) = \begin{bmatrix} f_1^{[\ell+1]}(x) \\ \vdots \\ f_r^{[\ell+1]}(x) \\ f_{r+1}^{[\ell+1]}(x) \\ \vdots \\ f_n^{[\ell+1]}(x) \end{bmatrix} = \begin{bmatrix} f_1^{[\ell]}(x) \\ \vdots \\ f_r^{[\ell]}(x) \\ \det D_{r+1}^{[\ell]}(x) \\ \vdots \\ \det D_n^{[\ell]}(x) \end{bmatrix} = 0, \quad r = r^{[\ell]}. \tag{3.10}$$

Here $F^{[\ell+1]}(x)$ is termed as the $(\ell+1)$st deflated equations.

It is easily seen from the above discussion that, compared with $F^{[\ell]}$, the convergence of $F^{[\ell+1]}$ will be improved. In fact, for the system in Section 2, we have the following theorem.

Theorem 3.2. Assume that a system of equations is given by

$$f_i(x) = (a_{i1}x_1 + \ldots + a_{i,i-1}x_{i-1} + x_i + \ldots + a_{in}x_n)^{m_i}\bar{f}_i(x) = 0, \tag{3.11}$$

where

$$f_i(x^*) = 0, \quad \bar{f}_i(x^*) \neq 0, \quad i = 1,2,\ldots,n, \tag{3.12a}$$

$$m_i \geqq 1, \quad \tilde{m} = \max_i\,[m_i], \tag{3.12b}$$

$$\begin{vmatrix} 1 & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \\ a_{n1} & a_{n2} & \cdots & 1 \end{vmatrix} \neq 0. \tag{3.12c}$$

Then at the ℓ-th deflation process, the multiplicity $m^{[\ell]}$ of $F^{[\ell]}(x)$ is given by

$$m^{[\ell]} = \prod_{i=1}^{n} \{\max[1,\ m_i-\ell]\}, \qquad 0 \leqq \ell \leqq \tilde{m}-1, \tag{3.13a}$$

$$m^{[\tilde{m}]} = 1. \tag{3.13b}$$

From Theorem 2.2 and the definition of $F^{[\ell]}(X)$ given by (3.10), the theorem can be easily proved.

Let $m = m^{[0]}$ be the multiplicity of (3.11). Then this theorem shows that, for the system (3.11), $(\tilde{m}-1)$ deflation processes are necessary to obtain the m-ple root of (3.11) in the same accuracy as usual simple roots.

3.2. Computational Realization

As we have seen, the determinant of the deflated matrix (3.6) must be calculated in an analytic form. In the package NAES, a symbolic and algebraic manipulation language, REDUCE 2, is adopted for this purpose. However from a practical standpoint, it is often possible to simplify or skip computations of the determinants by using some properties of the deflation matrix[16, 17].

Consider the ℓ-th deflation process at k-th iteration given by

$$F^{[\ell]}({}^kX) = 0, \tag{3.14}$$

and suppose that the pivot matrix $P_r^{[\ell]}({}^kX)$ is given by

$$P_r^{[\ell]}({}^kX) = \begin{array}{c} \begin{array}{cc} f_i^{[\ell]} & {}^kx_j \end{array} \\ \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ \vdots & \vdots \\ r^{[\ell]} & r^{[\ell]} \end{bmatrix} \end{array}, \qquad k,\ell = 0,1,\dots\ . \tag{3.15}$$

It is noteworthy that while solving the linear equations (3.3) by the Gaussian elimination method, the pivot matrix can easily be obtained by checking the properties (3.9) in Theorem 3.1.

If (i) the i-th equation $f_i^{[\ell]}$ of (3.14) contains x_j explicitly and its partial derivative at the root x^* is zero, (ii) the equation $f_i^{[\ell]}$ does not contain x_s explicitly, then the (i,j) and (i,s) elements of the Jacobian matrix are given by

$$F_x^{[\ell]}(x)\Big|_{x=x^*} = i\ [\ \dots,\ \partial_{x_j} f_i^{[\ell]},\ \dots,\ \partial_{x_s} f_i^{[\ell]},\ \dots\ ]_{x=x^*}$$

$$= [\ \dots,\ 0,\ \dots,\ \bullet,\ \dots\ ]. \tag{3.16}$$

Here 0 and $\bullet$ in (3.16) are called numerical and algebraic zeros, respectively. As for the numerical zero, we have the following

<u>Theorem 3.3</u>. Suppose that the sequence $\{{}^k x\}$ defined by (3.3) converges to the root x^*. If the (i,j) element of the Jacobian matrix $F_x^{[\ell]}$ at the root is the numerical zero, then the following estimate at the ℓ-th deflation process holds:

$$\lim_{k\to\infty} |\partial_{x_j} f_i^{[\ell]}({}^{k+1}x)/\partial_{x_j} f_i^{[\ell]}({}^k x)| \leqq \tfrac{1}{2}. \tag{3.17}$$

Applying (2.8b) in Theorem 2.1, this theorem can be easily proved.

The estimation (3.17) is useful to identify the elements with numerical zeros in the Jacobian matrix $F_x^{[\ell]}$. In the following assume that the deflated equations and the pivot matrix at the ℓ-th deflation process in the k-th iteration are given by (3.10) and (3.15), respectively. From the computational standpoint, we first provide the following category.

<u>Category 1 (numerical zeros)</u>;

Suppose that the (i_u, j_v) element of the Jacobian matrix $F_x^{[\ell]}$ ($u=1,2,\ldots,u^{[\ell]}$, $v=1,2,\ldots,v^{[\ell]}$; $0 \leqq u^{[\ell]}, v^{[\ell]} \leqq n$) has a numerical zero. If the $(i_{\tilde{u}}, j_{\tilde{v}})$ element has the least total degree of variables which are not in the pivot variables in (3.15), it is called the minimum zero element. Let us explain the procedure of the category by showing an example with three numerical zero elements:

$$\partial_{x_{j_1}} f_{i_1}^{[\ell]} \Big|_{x=x^*} = x_1 \cdot x_{r^{[\ell]}+1} \Big|_{x=x^*} = 0, \tag{3.18a}$$

$$\partial_{x_{j_2}} f_{i_2}^{[\ell]} \Big|_{x=x^*} = x_{r^{[\ell]}+1} \cdot x^2_{r^{[\ell]}+2} \Big|_{x=x^*} = 0, \tag{3.18b}$$

$$\partial_{x_{j_3}} f_{i_3}^{[\ell]} \Big|_{x=x^*} = x^2_{r^{[\ell]}+1} \cdot x^2_{r^{[\ell]}+2} \Big|_{x=x^*} = 0. \tag{3.18c}$$

(i) It is easily seen that (3.18a) is the minimum numerical zero element. Since x_1 is already in the pivot matrix (3.15), put $x_{r^{[\ell]}+1}$ into the pivot variables and revise the matrix:

$$P_{r+1}^{[\ell]} = \begin{array}{c} \begin{array}{cc} f_i^{[\ell]} & {}^k x_j \end{array} \\ \begin{bmatrix} 1 & 1 \\ \vdots & \vdots \\ r^{[\ell]} & r^{[\ell]} \\ r^{[\ell]}+1 & r^{[\ell]}+1 \end{bmatrix} \end{array}, \tag{3.19a}$$

where $\partial_{x_{j_1}} f_{i_1}^{[\ell]}$ is replaced by $f_{r^{[\ell]}+1}^{[\ell]}$.

(ii) Deleting (3.18a), consider further (3.18b) and (3.18c). Since (3.18b) is now the minimum element, we have the following pivot matrix:

$$P_{r+2}^{[\ell]} = \begin{array}{c} \begin{array}{cc} f_i^{[\ell]} & {}^k x_j \end{array} \\ \begin{bmatrix} 1 & 1 \\ \vdots & \vdots \\ r^{[\ell]}+2 & r^{[\ell]}+2 \end{bmatrix} \end{array}. \tag{3.19b}$$

(iii) Since there is no new variable in (3.18c), the procedure is terminated. Thus, the equations $\det D^{[\ell]}_{r+1}$ and $\det D^{[\ell]}_{r+2}$ in (3.10) need not to be computed and can now be replaced by (3.18a) and (3.18b), respectively.

(iv) If $r^{[\ell]}+2 = n$, then replace ℓ by $\ell+1$ and terminate the ℓ-th deflation process. Otherwise proceed to Category 2.

<u>Category 2 (nontrivially proportional rows)</u>;

Suppose that, except for the elements with numerical or algebraic zeros, all the elements in the i- and j-th rows of the Jacobian matrix $F^{[\ell]}_x$ at the root X^* satisfy the following relation:

$$\partial_{x_s} f^{[\ell]}_i(X)/\partial_{x_s} f^{[\ell]}_j(X)\Big|_{X=X^*} = \alpha\ (\neq 0), \quad |\alpha| > \infty, \quad 1 \leqq {}^{\forall}s \leqq n, \tag{3.20}$$

where α is a constant. Then it is easily seen that the rank of $F^{[\ell]}_x(X^*)$ is degenerated by one. From (3.20), we form the $n(n-1)/2$ equations:

$$\partial_{x_u} f^{[\ell]}_i(X)\cdot\partial_{x_v} f^{[\ell]}_j(X) - \partial_{x_v} f^{[\ell]}_i(X)\cdot\partial_{x_u} f^{[\ell]}_j(X) = 0,$$

$$1 \leqq u \leqq n-1, \quad u+1 \leqq v \leqq n. \tag{3.21}$$

It is worth mentioning that (3.21) is generated by the REDUCE 2.

We now provide the procedure for Category 2. (i) From (3.21), find the equation with the minimum toatl degree of variables and a new pivot variable which is not in (3.19b), and let $u = \tilde{u}$ and $v = \tilde{v}$. (ii) Denote (3.21) with $u = \tilde{u}$ and $v = \tilde{v}$ by $f^{[\ell]}_{r^{[\ell]}+3}$ and its new pivot variable by $x_{r^{[\ell]}+3}$, and put them into the pivot matrix:

$$P^{[\ell]}_{r+3} = \begin{array}{c} \begin{array}{cc} f^{[\ell]}_i & {}^k x_j \end{array} \\ \begin{bmatrix} 1 & 1 \\ \vdots & \vdots \\ r^{[\ell]}+3 & r^{[\ell]}+3 \end{bmatrix} \end{array}. \tag{3.19c}$$

Also replace $\det D^{[\ell]}_{r^{[\ell]}+1}(x)$ in (3.10) by $f^{[\ell]}_{r^{[\ell]}+3}$. (iii) If $r^{[\ell]}+3 = n$, then replace ℓ by $\ell+1$ and terminate the ℓ-th deflation process. (iv) Otherwise, delete (3.21) with $u = \tilde{u}$ and $v = \tilde{v}$, and repeat the procedure (i) $\sim$ (iv) until a new pivot variable can not be found in (3.21) for $\forall u$ and $\forall v$.

Category 3 (nontrivially proportional columns);

Similarly to Category 2, suppose that, except for the elements with numerical and algebraic zeros, all the elements in i- and j-th columns of the Jacobian matrix $F^{[\ell]}_x$ at the root x^* satisfy the following relation:

$$\partial_{x_i} f^{[\ell]}_s(x)/\partial_{x_j} f^{[\ell]}_s(x) \Big|_{x=x^*} = \beta\ (\neq 0), \quad |\beta| < \infty, \quad 1 \leqq \forall s \leqq n \tag{3.22}$$

where β is a constant. From (3.22), form the $n(n-1)/2$ equations:

$$\partial_{x_i} f^{[\ell]}_u(x)\cdot\partial_{x_j} f^{[\ell]}_v(x) - \partial_{x_i} f^{[\ell]}_v(x)\cdot\partial_{x_j} f^{[\ell]}_u(x) = 0,$$

$$1 \leqq u \leqq n-1, \quad u+1 \leqq v \leqq n. \tag{3.23}$$

Then the same procedures (i) $\sim$ (iv) in Category 2 also hold for Category 3.

Applying Categories 1 $\sim$ 3, computations of the deflation matrix (3.6) can *greatly be reduced*. However, if n pivot variables were not abtained by these categories, it is then necessary to compute some of the matrices by the following procedure.

Category 4;

Suppose that, from Categries 1 $\sim$ 3, the pivot matrix is given by

$$\begin{array}{cc} f^{[\ell]}_i & x_j \end{array}$$

$$P^{[\ell]}_{\tilde{r}} = \begin{bmatrix} 1 & 1 \\ \vdots & \vdots \\ \tilde{r} & \tilde{r} \end{bmatrix}, \quad r^{[\ell]} \leqq \tilde{r} < n. \tag{3.24}$$

Then the procedure is executed as follows:

(i) From the Jacobian matrix $F_x^{[\ell]}$, find the element with a new pivot variable, say, $x_{\tilde{r}+1}$ which is not in (3.24). (ii) Form the $(r^{[\ell]}+1) \times (r^{[\ell]}+1)$ deflation matrix $D_s^{[\ell]}$, $s = \tilde{r}+1,\dots,n-\tilde{r}$ given by (3.6) so that the element is included. (iii) If $\tilde{r}+1 = n$, then replace ℓ by $\ell+1$ and terminate the ℓ-th deflation process. (iv) Otherwise repeat the procedures (i) ∿ (iii).

IV. Numerical Example

Consider the same nonlinear equations in Šamanskii [18], i.e.,

$$f^{[0]}(X) = \begin{bmatrix} x_1 + x_2 + x_3 - 1 \\ 0.2x_1^3 + 0.5x_2^2 - x_3 + 0.5x_3^2 + 0.5 \\ x_1 + x_2 + 0.5x_3^2 - 0.5 \end{bmatrix} = 0. \tag{4.1}$$

This has a double root $X^* = (0, 0, 1)$. From (3.2), we have

$$\det F_x^{[0]}(X^*) = \det \begin{bmatrix} 1 & 1 & 1 \\ 0.6x_1^2 & x_2 & -1+x_3 \\ 1 & 1 & x_3 \end{bmatrix}_{X=X^*} \tag{4.2a}$$

$$= \det \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} = 0. \tag{4.2b}$$

It is easily seen that the rank of the Jacobian matrix is one.

Let us here define a convergence condition by

$$^kE^{[\ell]} = \{\frac{1}{n} f^{[\ell]}(^kX)' f^{[\ell]}(^kX)\}^{1/2} \leqq 10^{-14}. \tag{4.3}$$

(i) When the condition $^6E^{[0]} \leqq 10^{-4}$ at the sixth iteration was satisfied, the diagonal elements $e_{jj}({}^{k}x)$ in (3.8) were computed:

$$[|e_{jj}({}^{6}x)/e_{jj}({}^{5}x)|] = [1.000,\ 0.505,\ 0.500], \tag{4.4}$$

(ii) From the properties (3.9) in Theorem 3.1 and (4.4), the pivot matrix was given by

$$P_1^{[0]}({}^{6}x) = \begin{array}{cc} f_i^{[0]} & {}^{6}x_j \\ \end{array} \left[\begin{array}{cc} 1 & 1 \end{array}\right]. \tag{4.5}$$

(iii) Since there are numerical zeros in (4.2). Category 1 can be applied. In fact, from (4.2a) at $x = {}^{6}x$, the following pivot matrix and deflated equations were obtained by the REDUCE 2:

$$P_3^{[1]}({}^{6}x) = \begin{array}{cc} f_i^{[1]} & {}^{6}x_j \\ \end{array} \left|\begin{array}{cc} 1 & 1 \\ 2 & 2 \\ 3 & 3 \end{array}\right|, \tag{4.6}$$

$$f^{[1]}(x) = \left|\begin{array}{l} x_1 + x_2 + x_3 - 1 \\ x_2 \\ x_3 - 1 \end{array}\right| = 0. \tag{4.7}$$

(iv) It is easily seen that the rank of the Jacobian matrix at the solution corresponding to (4.7) is three. Thus the first deflated equations (4.7) has a simple root.

(v) The original equations (4.1) with $x = {}^{k}x$ (k=0,..,6) and deflated equations (4.7) with $x = {}^{k}x$ (k=7,8,9) were solved by the ε-secant method [15] with $\varepsilon = 10^{-8}$ which is a numerical realization of the Newton method.

The convergence tendency of ${}^{k}E^{[\ell]}$ with the deflation is shown in Table 4.1. That of the original equations without the deflation is also given in the table. As would be expected, the deflation algorithm resulted in faster convergence as well as the higher accuracy for the solution given in Table 4.2.

Table 4.1. Convergence tendencies

iteration	No. of deflation	with deflation	without deflation
0	0	0.103×10^{0}	0.103×10^{0}
1	0	0.569×10^{-1}	0.569×10^{-1}
2	0	0.157×10^{-1}	0.157×10^{-1}
3	0	0.394×10^{-2}	0.394×10^{-2}
4	0	0.986×10^{-3}	0.986×10^{-3}
5	0	0.247×10^{-3}	0.247×10^{-3}
6	0	0.617×10^{-4}	0.617×10^{-4}
7	1	0.739×10^{-2}	0.154×10^{-4}
8	1	0.758×10^{-13}	0.386×10^{-5}
9	1	0.0	0.964×10^{-6}
⋮			⋮
22			0.149×10^{-13}
23			0.368×10^{-14}

Table 4.2. Numerical solutions

	with deflation	without deflation
x_1	0.0	$0.32895108875564 \times 10^{-7}$
x_2	$-0.16787888226717 \times 10^{-18}$	$0.59028952604004 \times 10^{-7}$
x_3	1.0	0.99999990807594

V. Concluding Remarks

In this paper, several properties of the multiple roots for a system of nonlinear equations have been discussed first. Then the deflation algorithm for determining the multiple roots has been proposed. According to the algorithm, both convergency and accuracy can greatly be improved.

Finally a numerical example was solved. As would be expected, the ε-secant (the Newton) iteration with the deflation algorithm converged quadratically to the roots with sufficient accuracies.

We note that the deflation procedure can efficiently be executed by using a language for algebraic and symbolic manipulation, e.g., REDUCE 2.

All the numerical calculations were done on the DEC-System 2020 in the Computer Programming Laboratory of the Research Institute for Mathematical Sciences, Kyoto University, Kyoto.

References

[1] Blum, E. K. (1972) *Numerical Analysis and Computation Theory and Practice*, Addison-Wesley Publ. Comp., Reading, Mass..

[2] Dahlquist, G. and Björck, Å. (1974) *Numerical Methods*, Prentice-Hall, Englewood Cliffs, N.J.

[3] Ortega, J. M. and Rheinboldt, W. C. (1970) *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York.

[4] Ostrowski, A. M. (1973) *Solution of Equations in Euclidean and Banach Spaces*, Academic Press, New York.

[5] Rall, L. B. (1969) *Computational Solution of Nonlinear Operator Equations*, John Wiley & Sons, Inc., New York.

[6] Stoer, J. and Bulirsch, R. (1980) *Introduction to Numerical Analysis*, Springer-Verlag, New York.

[7] Branin, F. H., Jr. (1972) Widely Convergent Method for Finding Multiple Solutions of Simultaneous Nonlinear Equations, *IBM J. of Research and Development*, 16, 504-521.

[8] Decker, D. W. and Kelley, C. T. (1980) Newton's Method at Singular

Point. I, *SIAM J. Numer. Anal.*, 17, 66-70.

[9] Decker, D. W. and Kelley, C. T. (1980) Newton's Method at Singular Point. II, *SIAM J. Numer. Anal.*, 17, 465-471.

[10] Griewank, A. O. (1980) Starlike Domains of Convergence for Newton's Method at Singularities, *Numer. Math.*, 35, 95-111.

[11] Griewank, A. And Osborne, M. R. (1981) Newton's Method for Singular Problems When the Dimension of the Null Space Is > 1, *SIAM J. Numer. Anal.*, 18, 145-149.

[12] Rall, L. B. (1966) Convergence of the Newton Process to Multiple Solutions, *Numer. Math.*, 9, 23-37.

[13] Reddien, G. W. (1978) On Newton's Method for Singular Problems, *SIAM J. Numer. Anal.*, 15, 993-997.

[14] Ojika, T., Watanabe, S. and Mitsui, T., in preparation.

[15] Watanabe, S., Ojika, T. and Mitsui T., On Quadratic Convergence Properties of the ε-Secant Method for the Solution of System of Nonlinear Equations and Its Application to a Chemical Reaction Problem, to appear in *J. Math. Anal. Appl.*.

[16] Ojika, T. (1981) Deflation Algorithm for the Multiple Roots of Simultaneous Nonlinear Equations, *Memo. Osaka Kyoiku Univ.*, Ser. III, 30.

[17] Watanabe, S. (1982) On the Deflation Algorithm for Multiple Roots of Systems of Nonlinear Algebraic Equations and the Order of Convergence, *Bull. Yamagata Univ.*, 10, 245-263.

[18] Samanskii, V. E. (1967) The Application of Newton's Method in the Singular Case, *USSR J. Comp. Math. Math. Phys.*, 7, 774-783.