Newton-like method for singular 2-regular system of nonlinear equations

Stanisław Grzegórski*, Edyta Łukasik

Department of Computer Science, Lublin University of Technology, Nadbystrzycka 36b, 20-618 Lublin, Poland

Abstract

In this article the problem of solving a system of singular nonlinear equations will be discussed. The theory of local and Q-superlinear convergence for the nonlinear operators is developed.

1. Introduction

Let $F : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a nonlinear operator. The problem of solving a system of nonlinear equations consist in finding a solution $x^* \in D$ of the equation

$$F(x) = 0.$$  \hspace{1cm} (1)

Definition 1

A linear operator $\Psi_2(h) : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $h \in \mathbb{R}^n$, is called 2-factor operator, if

$$\Psi_2(h) = F'(x^*) + P^\perp F\big( x^* \big) h,$$ \hspace{1cm} (2)

where

$P^\perp$ - denotes the orthogonal projection on $\left( \text{Im} \ F'(x) \right)^\perp$ in $\mathbb{R}^n$ [1].

Definition 2

Operator $F$ is called 2-regular in $x^*$ on the element $h \in \mathbb{R}^n$, $h \neq 0$, if the operator $\Psi_2(h)$ has the property:

$$\text{Im} \ \Psi_2(h) = \mathbb{R}^m.$$
Definition 3
Operator $F$ is called 2-regular in $x^*$, if $F$ is 2-regular on the set $K_2(x^*)\{0\}$, where

$$K_2(x^*) = \text{Ker}F^\prime(x^*) \cap \text{Ker}^2 P^\perp F^\prime(x^*), \quad (3)$$

$$\text{Ker}^2 P^\perp F^\prime(x^*) = \{ h \in R^n : P^\perp F^\prime(x^*)[h]^2 = 0 \}.$$  

We need the following assumption on $F$:
A1) completely degenerated in $x^*$:
$$\text{Im} F^\prime(x^*) = 0. \quad (4)$$

A2) operator $F$ is 2-regular in $x^*$:
$$\text{Im} F^\prime(x^*)h = R^m \text{ for } h \in K_2(x^*), h \neq 0. \quad (5)$$

A3)
$$\text{Ker} F^\prime(x^*) \neq \{0\}. \quad (6)$$

If $F$ satisfies A1 in $x^*$, then
$$K_2(x^*) = \text{Ker}^2 F^\prime(x^*) = \{ h \in R^n : F^\prime(x^*)[h]^2 = 0 \}.$$  

In [1] it was proved, that if $n=m$, then the sequence

$$x_{k+1} = x_k - \left\{ \hat{F}^\prime(x_k) + P_k^\perp F^\prime(x_k)h_k \right\}^{-1} \cdot \left\{ F(x_k) + P_k^\perp F^\prime(x_k)h_k \right\}, \quad (8)$$

where

$P_k^\perp$ - denotes orthogonal projection on $\left( \text{Im} \hat{F}^\prime(x_k) \right)^\perp$ in $R^n$,

$h_k \in \text{Ker} \hat{F}^\prime(x_k)$, $\|h_k\| = 1$

converges Q-quadratically to $x^*$.

The matrices $\hat{F}^\prime(x_k)$ obtained from $F^\prime(x_k)$ by replacing all elements, whose absolute values do not increase $\nu>0$, by zero, where $\nu = \nu_k = \|F(x_k)\|^{(1-\alpha)/2}$, $0<\alpha<1$.

In the case $n = m+1$ the operator

$$\left\{ \hat{F}^\prime(x_k) + P_k^\perp F^\prime(x_k)h_k \right\}^{-1}$$

in method (8) is replaced by the operator

$$\left[ \hat{F}^\prime(x_k) + P_k^\perp F^\prime(x_k)h_k \right]^+$$

and then the method converges Q-linearly to the set of solutions [2].

Under the assumptions A1-A3, the system of equation (1) is undetermined ($n>m$) and degenerated in $x^*$. 

PDF created with FinePrint pdfFactory Pro trial version www.pdffactory.com
2. Extending of the system of equation

Now we construct the operator $\Phi : \mathbb{R}^n \to \mathbb{R}^{n-1}$ with the properties (4), (5) and such that $\Phi(x^*)=0$ [2].

Assume

A4) Let $F(x)=[f_1(x), f_2(x), \ldots, f_m(x)]^T$, $n>m$ is two continuously differentiable in some neighbourhood $U \subset \mathbb{R}^n$ of the point $x^*$.

Denote:

$$H=\text{lin}\{h\} \quad \text{for} \quad h \in \text{Ker} F' \left( x^* \right), \ h \neq 0.$$  

$P = P_{H^\perp}$ denotes the orthogonal projection $\mathbb{R}^n$ on $H^\perp$.

$$j'_i\Phi(x) = P \left( f'_i \left( x \right) \right)^T \quad \text{for} \ i=1,2,\ldots,m.$$  

For each system of indices $i_1, i_2, \ldots, i_{n-m-1} \subset \{1, 2, \ldots, m\}$ and vectors $h_1, h_2, \ldots, h_{n-m-1} \subset \mathbb{R}^n$ we define

$$\Phi(x) = \begin{bmatrix} F'(x) h \\ \varphi(x) \end{bmatrix},$$

where

$$\varphi(x) : \mathbb{R}^n \to \mathbb{R}^r, \quad r=n-m-1,$$

$$\varphi(x) = PF' \left( x \right) \bar{h}, \quad \bar{h} = \begin{bmatrix} h_1, h_2, \ldots, h_r \end{bmatrix}^T,$$

$$\varphi(x) = M \begin{bmatrix} j'_i\Phi(x) h_i \\ \cdots \\ j'_r\Phi(x) h_r \end{bmatrix}. \quad (11)$$

In [2] it was proved, that the sequence

$$x_{k+1} = x_k - \left[ \Phi \left( x_k \right) \right]^+ \cdot \Phi \left( x_k \right), \quad k=0,1,2,\ldots$$

quadratically converges to the solution of (1).

3. New method

We propose the Newton-like method, where the sequence $\{x_k\}$ is defined by:

$$x_{k+1} = x_k - \left( B_k \right)^+ \cdot \Phi \left( x_k \right). \quad (13)$$

The operator $\Phi'$ will by approximated by matrices $\{B_k\}$.

Let

$$s_k = x_{k+1} - x_k. \quad (14)$$

We propose matrices $B_k$ which satisfy the secant equation:

$$B_{k+1}s_k = \Phi \left( x_{k+1} \right) - \Phi \left( x_k \right) \quad \text{for} \ k=0,1,2,\ldots$$

(15)

For example, to obtain the sequence $\{B_k\}$ we can apply the Broyden method:
\[
B_{k+1} = B_k - \frac{r_k s_k^T}{s_k^T s_k} \quad \text{for } k=0,1,2,...
\]  

(16)

where

\[ r_k = \Phi(x_{k+1}) - \Phi(x_k) - B_k s_k. \]

(17)

We will prove for this method:

**Q-linear convergence** to \( x^* \) i.e. there exists \( q \in (0,1) \) such, that

\[
\|x_{k+1} - x^*\| \leq q \|x_k - x^*\| \quad \text{for } k = 0,1,2,...
\]

(18)

and next **Q-superlinear convergence** to \( x^* \), i.e.:

\[
\lim_{k \to \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = 0.
\]

(19)

We present the theorem which is an analogue of the Bounded Deterioration Theorem (Broyden, Dennis and More - [3]) for the Newton-like methods, when the operator \( F' (x^*) \) is nonsingular.

**Theorem 1** (The Bounded Deterioration Theorem)

Let \( F \) satisfies the assumptions A1-A4. If exist constants \( q_1 \geq 0 \) and \( q_2 \geq 0 \) such that matrices \( \{B_k\} \) satisfy the inequality:

\[
\|B_{k+1} - \Phi ' (x^*)\| \leq (1 + q_1 r_k) \|B_k - \Phi ' (x^*)\| + q_2 r_k,
\]

(20)

then there are constants \( \varepsilon > 0 \) and \( \delta > 0 \) such, that if

\[
\|x_0 - x^*\| \leq \varepsilon \quad \text{and} \quad \|B_0 - \Phi ' (x^*)\| \leq \delta,
\]

then the sequence

\[ x_{k+1} = x_k - B_k^* \Phi(x_k) \]

converges Q-linearly to \( x^* \).

When the system of equation is rectangular, the proof of the theorem is analogous to that for the nonsingular and quadratic system and we neglect it.

**Theorem 2** (Linear convergence)

Let \( F \) satisfies the assumptions A1-A4. Then the method

\[
x_{k+1} = x_k - \{B_k\}^* \cdot \Phi(x_k),
\]

\[
B_{k+1} = B_k - \frac{\{\Phi(x_{k+1}) - \Phi(x_k) - B_k s_k\} s_k^T}{s_k^T s_k}
\]

locally and Q-linearly converges to \( x^* \).

**Proof.**

To prove the Theorem we should prove the inequality (20) from Theorem 1.

Now we notice:
\[ \| B_{k+1} - \Phi'(x^*) \| \leq \| B_k - \Phi'(x^*) \| + \left\{ \Phi(x_{k+1}) - \Phi(x_k) - B_k s_k \right\} s_k^T \leq \| B_k - \Phi'(x^*) \| + \left\{ \Phi(x_{k+1}) - \Phi(x_k) - \Phi'(x^*) s_k + \Phi'(x^*) s_k - B_k s_k \right\} s_k^T \leq \| B_k - \Phi'(x^*) \| + \left\{ \Phi(x_{k+1}) - \Phi(x_k) - \Phi'(x^*) (x_{k+1} - x^*) s_k \right\} s_k^T + \left\{ \Phi(x_k) - \Phi'(x^*) (x_k - x^*) s_k^T \right\} s_k^T \leq \| \Phi'(x^*) - B_k \| (1 + q_i r_k) + c_1 \frac{\| x_{k+1} - x^* \|^2}{s_k^T s_k} \leq \| \Phi'(x^*) - B_k \| (1 + q_i r_k) + q_2 r_k, \]

where \( c_1 > 0, c_2 > 0, q_1 > 0, q_2 > 0, r_k = \max \{ \| x_{k+1} - x^* \|, \| x_k - x^* \| \}. \]

**Theorem 3** (Q-superlinear convergence)
Let \( F \) satisfies the assumptions A1-A4 and the sequence

\[ x_{k+1} = x_k - \left\{ B_k \right\}^{-1} \cdot \Phi (x_k), \]

\[ B_{k+1} = B_k - \left\{ \Phi (x_{k+1}) - \Phi (x_k) - B_k s_k \right\} s_k^T \]

linearly converges to \( x^* \). Then the sequence \( \{ x_k \} \) Q-superlinearly converges to \( x^* \).

**Proof.**
Matrices \( B_k \) satisfy secant equation (15), so

\[ B_{k+1} = P_{L_k} B_k \]  \hspace{1cm} (21)

where

\[ L_k = \{ X : X s_k = y_k, \text{ where } y_k = \Phi'(x_{k+1}) - \Phi'(x_k) \} \]  \hspace{1cm} (22)

Denote

\[ H_k = H (x_k, x_{k+1}) = \frac{1}{0} \Phi'(x_k + t (x_{k+1} - x_k)) dt. \]

We have \( H_k \in L_k \) [4].
From (21) and [3] it follows:

\[ \|B_{k+1} - B_k\|^2 + \|B_{k+1} - H_k\|^2 = \|B_k - H_k\|^2, \quad \text{for } i = 0, 1, 2, \ldots \]  

By lemma 2 [5] we get \( \sum_{k=1}^{\infty} \|B_{k+1} - B_k\|^2 < \infty \), thus we obtain

\[ \|B_{k+1} - B_k\| \to 0. \]

This denotes that the method (13)-(17) is Q-superlinearly convergent [6], which ends the proof. □

4. Summary

The proposed method is Q-superlinearly convergent and easier to apply than the method (12), without calculation of \( F''(x_k) \).

References