Available online at www.sciencedirect.com
ScienceDirect
Computers and Mathematics with Applications 55 (2008) 2134–2141
An International Journal computers & mathematics with applications
www.elsevier.com/locate/camwa

A line search filter approach for the system of nonlinear equations☆

Pu-yan Nie[a,∗], Ming-yong Lai[b], Shu-jin Zhu[b], Pei-ai Zhang[c]

[a] Institute of Industrial Economics, Jinan University, Guangzhou, 510632, PR China
[b] College of Economics and Trade, Hunan University, Changsha, 410079, PR China
[c] Department of Mathematics, Jinan University, Guangzhou, 510632, PR China

Received 22 February 2006; received in revised form 2 August 2007; accepted 30 August 2007

Abstract

Some constrained optimization approaches have been recently proposed for the system of nonlinear equations (SNE). Filter approach with line search technique is employed to attack the system of nonlinear equations in this paper. The system of nonlinear equations is transformed into a constrained nonlinear programming problem at each step, which is then solved by line search strategy. Furthermore, at each step, some equations are treated as constraints while the others act as objective functions, and filter strategy is then utilized. In essence, constrained optimization methods combined with filter technique are utilized to cope with the system of nonlinear equations.

Keywords: Optimization; Nonlinear programming; Equality constraints; The system of nonlinear equations

1. Introduction

The system of nonlinear equations (SNE) plays crucial roles in economics, in engineering field and in the optimization community [1,2], and is formally stated as follows

$$c_i(x) = 0 \quad i = 1, 2, \ldots, m, \tag{1.1}$$

where $x \in R^n$ and $c_i : R^n \longrightarrow R$ for $i = 1, 2, \ldots, m$.

When (1.1) is tackled by iterative methods, x_k, for $k = 1, 2, \ldots$, is utilized to denote the successive iterates. There mainly exist three ways to tackle (1.1).

The apparent technique is based on successive linearization, in which iterating direction d_k is obtained on iteration k by handling the system of linear equations

$$c_i^{(k)} + (a_i^{(k)})^T d = 0 \quad i = 1, 2, \ldots, m, \tag{1.2}$$

☆ This work is supported by the National Natural Science Foundation of China (No. 10501019), the New Century Excellent Talents in University (NCET) of State Education Ministry, and the China Postdoctoral Science Foundation (No. 20060400875). It is also partly supported by the Second Phase Construction Item of Hunan University and Base for Economic Globalization and Development of Trade Philosophical Social Sciences.
∗ Corresponding author.
E-mail address: pynie2005@yahoo.com.cn (P.-y. Nie).

0898-1221/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.camwa.2007.08.037

where $c_i^k = c_i(x_k)$, $a_i^k = \{\nabla c_i(x_k)\}$ for $i = 1, 2, \ldots, m$. We can employ Newton's method to attack it. If $m = n$, (1.2) has local and second-order convergence near a regular solution. It is also possible that (1.2) is inconsistent.

The second approach is to pose (1.1) as a minimization problem

$$\text{minimize} \qquad h(x) = c(x)^T c(x).$$

This problem can also be handled by successive linearization. This idea helps to improve the global properties of Newton's method [3,4]. On the other hand, there still exist some potential difficulties. Powell [4] gives an example that the iterates based on the above minimization problem converge to a non-stationary point of $h(x)$, which is obviously unsatisfactory.

The third strategy is recently proposed in [5]. In [5], Nie divides the set $\{1, 2, \ldots, m\}$ into two subsets S_1 and S_2 at each step, where S_2 denotes the complement $\{1, 2, \ldots, m\}/S_1$. S_1 and S_2 can be defined according to the case and the constrained optimization algorithms are utilized. The problem thus becomes the following optimization problem with equality constraints

$$\begin{aligned} &\text{minimize} \quad \sum_{i \in S_1} c_i^2(x) \\ &\text{subject to} \quad c_j(x) = 0, \qquad j \in S_2. \end{aligned} \tag{1.3}$$

The choice of two sets S_1 and S_2 will be given below the algorithm in Section 2 in this paper. We aim to solve (1.3) with optimization approaches. Some other ways are also proposed to handle (1.3) [6–8].

When (1.1) is infeasible, a local minimization of $h(x) > 0$ is found or a point is located at which the linearized system (1.2) is infeasible, which is regarded as local infeasibility. For convenience, we give a definition of the solution about (1.1) based on (1.3), which will help us to understand the algorithm.

Definition 1. If $x^\star$ is a local minimization of (1.3) but not a solution to (1.1), we call $x^\star$ a local infeasibility point of (1.1).

Just like the global minimization of $h(x)$, it is very difficult to describe global infeasibility. We consequently discuss local infeasibility point in this paper. Moreover, there exist rare papers devoted to the utilization of constrained optimization strategies to solve the system of nonlinear equations. In this paper, we try to employ constrained optimization methods to attack (1.1). Furthermore, line search technique is employed in this paper, which is different from that in [5]. There are other techniques for the system of nonlinear equations (NSE), such as exclusion region algorithms [9] and differential-free algorithms[10].

Filter approaches, in which constraints and objective function are efficiently balanced, are recently proposed by Fletcher and Leyffer [11]. Filter strategies are focused on because of promising numerical results [12–22]. In [17], the local properties of filter methods are researched. Filter approach is also extended to handle the system of nonlinear equations(SNE) in [6].

The paper is organized as follows: In Section 2, a new algorithm based on filter line search technique similar to that in [20,21] is put forward for the system of nonlinear equations(SNE). The algorithm is analyzed in Section 3. Some numerical results and remarks are given in Section 4.

2. A line search filter algorithm for the system of nonlinear equations

To deal with (1.3) based on (1.1), we consider the KKT conditions to (1.3), which is given as follows:

$$\begin{aligned} &g(x) + A_{S_2}(x)\lambda = 0 \\ &c_{S_2}(x) = 0, \end{aligned} \tag{2.4}$$

where $g(x) = \nabla \sum_{i \in S_1} c_i^2(x)$ and $A_{S_2} = \nabla c_{S_2}(x)^T$, with the Lagrangian multipliers λ. Its KKT conditions are linearized at the beginning of the kth iteration x_k and the following formula is obtained:

$$\begin{bmatrix} B_k & A_{S_2}^k \\ (A_{S_2}^k)^T & 0 \end{bmatrix} \begin{bmatrix} s_k \\ \lambda_k^+ \end{bmatrix} = - \begin{bmatrix} g_k \\ c_{S_2}^k \end{bmatrix}, \tag{2.5}$$

where B_k is the Hessian or approximate Hessian matrix to $L(x,\lambda) := \sum_{i\in S_1} c_i^2(x) + c_{S_2}(x)^T\lambda$, $c_{S_2}^k = c_{S_2}(x_k)$, $(g_k)_i = \frac{\partial \sum_{i\in S_1} c_i^2(x_k)}{\partial (x_k)_i}$ and $A_{S_2}^k = \nabla c_{S_2}(x_k)$. Actually, (2.5) may be inconsistent. In [7], Celis–Dennis–Tapia subproblem (CDT) approach is employed to avoid insistence. In this paper, the subproblem is based on (2.5). In each step, line search is also utilized to obtain the step size. Further, we need just the approximate solution to (2.5).

After a direction s_k has been obtained, a step size $\alpha_k \in (0, 1]$ is determined to obtain the new iterate

$$x_k(\alpha_k s_k) := x_k + \alpha_k s_k,$$

where s_k is obtained from (2.5) and α_k is the step size following from some type of line search strategy. When a new point is obtained, the sets S_1 and S_2 are all updated by some strategy.

To handle (1.1), we solve (1.3) by employing line search techniques. The search direction is obtained by attacking (2.5). When (2.5) has no solution, restoration algorithm is employed to find a new iterate so that (2.5) is consistent at the new point. The restoration algorithm aims to find a feasible point and the following subproblem is solved:

$$\min_{r_k} \theta_k(x_k + r_k) = \|c_{S_2}(x_k) + \nabla c_{S_2}^T r_k\|^2,$$

where 2-norm is always employed.

To determine whether to accept or deny the trial point, some criterion is employed. The merit function is thus utilized.

$$m_k(x) = \sum_{i\in S_1} c_i(x)^2,$$

and

$$\theta_k(x_k) = \|c_{S_2}(x)\|_2^2.$$

The filter criterion is defined as follows

$$\begin{aligned} &m_k(x_k(\alpha_k s_k)) \le m_k(x_i) + \gamma_m \theta_k(x_k(\alpha_k s_k)) \\ &or \theta_k(x_k(\alpha_k s_k)) \le \gamma_\theta \theta_k(x_i), i \in \mathcal{F}_k, \end{aligned} \tag{2.6}$$

where $0 \le \gamma_m, \gamma_\theta < 1$ are two constants. If $x_k(\alpha_k s_k)$ satisfies (2.6), $x_k(\alpha_k s_k)$ is accepted and $x_{k+1} := x_k(\alpha_k s_k)$. Similar to the method in [20], to avoid obtaining a feasible point while not an optimal solution, the following switching condition is employed:

$$u_k(\alpha_{k,l}) < 0, \qquad -u_k(\alpha_{k,l})\alpha_{k,l} > \delta\theta(x_k), \tag{2.7}$$

where the constant $\delta > 0$, $u_k(\alpha) := \alpha g_k^T s_k$ and the following formulation is met

$$m_k(x_k)(\alpha_{k,l}) \le m_k(x_k) + \tau_3 u_k(\alpha_{k,l}) \tag{2.8}$$

with a fixed constant $\tau_3 \in (0, \frac{1}{2})$. For some constant $\eta_f > 0$, we also define the following actual reduction as follows:

$$Ared(\alpha_{k,l} s_k) = -(m_k(x_k + \alpha_{k,l} s_k) - m_k(x_k)) \ge \eta_f u_k(\alpha_{k,l}). \tag{2.9}$$

We note that a trial point which satisfied (2.7)–(2.9) is a f-type point. In this way, $x_{k+1} := x_k(\alpha_{k,l} s_k)$ but the filter set is not augmented. For convenience, we assume that the solution s_k to (2.5) is denoted $Z_k v_k$ where Z_k satisfies $Z_k^T A_{S_2}^k = 0$ and $Z_k^T Z_k = I$. The algorithm is formally given as follows, in which some line search technique is employed:

Algorithm 1. (*A Line Search Filter Algorithm for the System of Nonlinear Equations*)

Step 0. Choose x_0, S_1^0, S_2^0 and ϵ, compute g_0, $c_i(x_0)$, A_k for $i \in S_2^0$ and Z_0. Set $k := 0$ and $\mathcal{F}_0 = \{0\}$.

Step 1. If $\|c(x_k)\| \le \epsilon$ then **stop**. Give $\alpha_k^{\min} := \begin{cases} \min\{\gamma_\theta, \frac{\gamma_m \theta_k(x_k)}{-g_k^T s_k}\}, & g_k^T s_k < 0 \\ \gamma_\theta & \text{Otherwise.} \end{cases}$

Step 2. **Compute** (2.5) to obtain s_k. If there exists no solution to (2.5), goto **Step 6**.

Step 3. If $\|s_k\| \le \epsilon$ then **stop**. Otherwise, goto step 4.

Step 4. If (2.7)–(2.9) is satisfied for some $\alpha_{k,l} \geq \alpha_k^{\min}$, $x_{k+1} := x_k + \alpha_{k,l} s_k$ and goto Step 5. If $\alpha_{k,l} < \alpha_k^{\min}$, goto Step 6. If (2.6) is violated, goto Step 6. Otherwise, $x_{k+1} := x_k + \alpha_k s_k$ and update $\mathcal{F}_{k+1} := \mathcal{F}_k \bigcup \{(\theta, m) \in R^2 : \theta \geq (1-\gamma_\theta)\theta_k(x_k), m \geq m_k(x_k) - \gamma_m \theta_k(x_k)\}$.
Step 5. **Compute** $g_{k+1}, B_{k+1}, A_{k+1}, S_1^{k+1}, S_2^{k+1}, Z_{k+1}$. Let $k := k+1$, and goto Step 1.
Step 6. (Restoration Stage) Find $x_k^r := x_k + \alpha_k^r s_k^r$ such that x_k^r is accepted by the filter and the infeasibility measure θ is reduced. Goto step 1.

We now discuss the sets S_1 and S_2. In this work, for some positive constant $n_0 > 0$, we define

$$c_{i_1}^2(x_k) \geq c_{i_2}^2(x_k) \geq \cdots \geq c_{i_n}^2(x_k)$$

$$S_1 = \{i_k | k \leq n_0\}, \tag{2.10}$$

$$S_2 = \{i_k | k \geq n_0 + 1\}. \tag{2.11}$$

The results in Section 3 are all based on (2.10) and (2.11).

There are some advantages of optimization techniques with constraints to attack the system of nonlinear equations. Firstly, it provides another way to attack the system of nonlinear equations. Secondly, when a point lies near a local infeasibility point, it may be very slow to find a local infeasibility point for other approaches. The local infeasibility point can be immediately found with Algorithm 1 because some second-order information (or second-order derivatives) of the equation is facilitated to the full. Finally, Algorithm 1 will help in balancing the differences in all equations.

3. Convergent properties of the algorithm

As we know, line search approach has globally convergent property with exact line search. We hope that there are also global convergence results for Algorithm 1. To give the global convergence, we make some assumptions as follows, which are called standard assumptions in general.

Assumption 1. (1) The set $\{x_k\} \in X$ is nonempty and bounded;
(2) $c_i(x), i = 1, 2, \ldots, m$ are all twice continuously differentiable on an open set containing X;
(3) The matrix sequence $\{B_k\}$ is bounded.

Assume that $\|B_k\| \leq M_1$ and $\| \bigtriangledown^2 c_i(x)\| \leq M_2$ for all k, where M_1 and M_2 are all positive constants independent of k. We then analyze the properties of Algorithm 1 based on the above assumption. When the algorithm terminates finitely, an ϵ solution of (1.1) or a local infeasibility point is obtained. It is apparent that the following result holds:

Theorem 1. *If Algorithm 1 terminates at Step 1, then an ϵ approximate solution to (1.1) is achieved. If the Algorithm terminates at Step 3, then a local infeasibility point is obtained.*

Proof. The first part is obvious by virtue of the terminating conditions in Step 1. When Algorithm 1 terminates at Step 3, it is not a solution to (1.1) because it does not satisfy the condition of Step 1, while it satisfies the KKT conditions to (1.3). It is hence a local minimization of (1.3). It is therefore a local infeasibility point. □

When the algorithm terminates infinitely, we investigate the iterate sequence. We then have the following results

Lemma 1. *If there are only finite or infinite number of iterates entering the filter, we then have*

$$\lim_{k\to\infty} \theta_k = 0. \tag{3.12}$$

Proof. Similar to that in [19] or Theorem 1 in [20], the result is immediately obtained. The detail proof is omitted. □

Further, we always assume that the solutions to (2.5) satisfy certain descent conditions.

Assumption 2. (1) $(A_{S_2}^k)^T$ has full column rank and Z_k is bounded for any k.
(2) $\|s_k\| \leq \gamma_s$ for all $k$with a positive constant γ_s.

We therefore have the following result, which explain that Algorithm 1 has the advantage of the least square strategy and the residual is reduced at each step.

Lemma 2. *Under Assumption 1, there exists the solution to (2.5) with exact (or inexact) line search which satisfies the following descent condition,*

$$|\theta_k(x_k+\alpha s_k)-(1-2\alpha)\theta_k(x_k)|\le \tau_1\alpha^2\|s_k\|^2$$
$$|m_k(x_k+\alpha s_k)-m_k(x_k)-u_k(\alpha)|\ge \tau_2\alpha^2\|s_k\|^2, \tag{3.13}$$

where $\alpha\in(0,1)$, τ_1 and τ_2 are all positive constants independent of k.

Proof. These inequalities directly follow from second-order Taylor expansion of $c_i(x)$ for $i=1,2,\ldots,n$ and (2.4). According to Taylor expansion of $\sum_{i\in S_1}c_i^2((x_k)+\alpha s_k)$ with $i\in S_1$, we then have

$$\left|\sum_{i\in S_1}c_i^2(x_k+\alpha s_k)-\sum_{i\in S_1}c_i^2(x_k)-\alpha g_k^T s_k\right| = \left|\frac{1}{2}\alpha^2(s_k)^T\,\nabla^2\sum_{i\in S_1}c_i^2(x_k+\varrho\alpha s_k)s_k\right|$$
$$\le \tau_2\alpha^2\|s_k\|^2,$$

where the last inequality follows the Assumption 1 and $\varrho\in[0,1]$. The second inequality in (3.13) therefore holds.

On the other hand, by virtue of the Taylor expansion of $c_i^2(x_k+\alpha s_k)$ with $i\in S_2$, we obtain

$$|c_i^2(x_k+\alpha s_k)-c_i^2(x_k)-2\alpha c_i(x_k)\,\nabla\,c_i(x_k)^T s_k|$$
$$=|\alpha^2 s_k^T[c_i(x_k)\,\nabla^2\,c_i(x_k+\zeta\alpha s_k)+\nabla c_i(x_k+\zeta\alpha s_k)\,\nabla\,c_i(x_k+\zeta\alpha s_k)^T]s_k|$$
$$\le\frac{1}{n}\tau_1\alpha^2\|s_k\|^2,$$

where the last inequality follows the Assumption 1 and $\zeta\in[0,1]$. Furthermore, from (2.4) we immediately obtain $-2\alpha c_i^2(x_k)-2\alpha c_i(x_k)\,\nabla\,c_i(x_k)^T s_k=0$. The first inequality in (3.13) consequently holds. The result holds and the proof is completed. □

We now show that a KKT point is obtained by the above algorithm. The following lemma is given to help prove the main result.

Lemma 3. *Under Assumptions 1 and 2, if $u_k(\alpha)\le-\alpha\epsilon_0$ for a positive constant ϵ_0 independent of k and for all $\alpha\in(0,1]$ and $\alpha\ge\alpha_{k,l}^{\min}$ with $(\theta_k(x_k),m_k(x_k))\notin\mathcal{F}_k$, then there exist $\gamma_1,\gamma_2>0$ so that $x_k+\alpha s_k\notin\mathcal{F}_k$ for all k and $\alpha\le\min\{\gamma_1,\gamma_2\theta_k(x_k)\}$.*

Proof. If $u_k(\alpha)\le-\alpha\epsilon_0$ for a positive constant ϵ_0 independent of k and for all $\alpha\in(0,1]$, the first part of (2.7) accordingly holds. The second part of (2.7) follows the definition of $\alpha_{k,l}^{\min}$. Let $(\theta_k(x_k),m_k(x_k))\notin\mathcal{F}_k$.

We further point a fact according to the definition of filter. If $(\bar\theta,\bar m)\notin\mathcal{F}_k$ and $\theta\le\bar\theta, m\le\bar m$, we obtain $(\theta,m)\notin\mathcal{F}_k$.

Define $\gamma_1:=\min\{1,(\epsilon_0+\frac{\theta_k(x_k)}{\alpha})/(\gamma_s^2\tau_2)\}$ and $\gamma_2:=1/(2\gamma_s^2\tau_1)$, where $\frac{\theta_k(x_k)}{\alpha}>0$ is upper bounded because of $\alpha\ge\alpha_{k,l}^{\min}$ and the definition of $\alpha_{k,l}^{\min}$ or directly follows (2.7). For $\alpha\le\gamma_1$, we correspondingly have $\alpha^2\le\frac{\alpha\epsilon_0+\theta_k(x_k)}{\gamma_s^2\tau_2}\le\frac{-u_k(\alpha)+\theta_k(x_k)}{\tau_2\|s_k\|^2}$. Namely, we obtain

$$u_k(\alpha)+\tau_2\alpha^2\|s_k\|^2-\theta_k(x_k)\le 0. \tag{3.14}$$

Combining (3.13) and (3.14), we have

$$m_k(x_k+\alpha s_k)\le m_k(x_k)-\theta_k(x_k). \tag{3.15}$$

Similarly, for $\alpha\le\gamma_2\theta_k(x_k)\le\frac{\theta_k(x_k)}{2\tau_1\|s_k\|^2}$, we have $-2\alpha\theta_k(x_k)+\tau_1\alpha^2\|s_k\|^2\le 0$. Combining (3.13), we also have

$$\theta_k(x_k+\alpha s_k)\le\theta_k(x_k). \tag{3.16}$$

On the other hand, from the update of S_1 and S_2, we have

$$\theta_{k+1}(x_k + \alpha s_k) \le \theta_k(x_k + \alpha s_k) \le \theta_k(x_k).$$

By combining (3.15) and (3.16) and the updating policy of S_1, S_2, we therefore have

$$m_{k+1}(x_k + \alpha s_k) \le m_k(x_k).$$

When $(\theta_k(x_k), m_k(x_k)) \notin \mathcal{F}_k$, then there exist $\gamma_1, \gamma_2 > 0$ so that $(\theta_{k+1}(x_k + \alpha s_k), m_{k+1}(x_k + \alpha s_k)) \notin \mathcal{F}_k$ for all k and $\alpha \le \min\{\gamma_1, \gamma_2\theta(x_k)\}$. The result consequently holds and the proof is complete. □

Remarks. It is reasonable that there exist $(\theta(x_k), m(x_k)) \notin \mathcal{F}_k$ if $u_k(\alpha) \le -\alpha\epsilon_0$ for a positive constant ϵ_0 independent of k and for all $\alpha \in (0, 1]$. We can show it by contradiction and the detailed proof is omitted. The convergence result is now given as follows.

Theorem 2. *Under Assumptions 1 and 2, for the sequence generated by Algorithm 1, we have*

$$\lim_{k\to\infty} \inf[\|c^k_{S_2^k}\| + \|Z_k^T g_k\|_2] = 0. \tag{3.17}$$

Namely, it has an accumulation which is the ϵ solution to (1.1) or a local infeasibility point. If the gradients of $c_i(x_k)$ are linear independent for all k and $i = 1, 2, \ldots, m$, then, the solution to SNE is obtained.

Proof. We now show that

$$\lim_{k\to\infty} \inf[\|c^k_{S_2^k}\| + \|s_k\|_2] = 0. \tag{3.18}$$

If (3.18) holds, (3.17) is also true, see [23]. We show (3.18) by contradiction. If the result were false there should have been a constant $\epsilon_1 > 0$ such that

$$\|c^k_{S_2^k}\| + \|s_k\|_2 > \epsilon_1 \tag{3.19}$$

for all k. Moreover, there exist the following results for sufficiently large k.

It is apparent that $\|c^k_{S_2^k}\| \le \frac{1}{2}\epsilon_1$ hold accordingly for large enough k to the Algorithm 1. It is reasonable that we have $\|s_k\|_2 \ge \frac{1}{2}\epsilon_1$ from (3.19).

Because $\|s_k\|_2 \ge \frac{1}{2}\epsilon_1$ for all k, let $\tau := \min\{\frac{\gamma_1}{2}, \gamma_2\theta_{k_0}(x_{k_0})\}$ for $(\theta_{k_0}(x_{k_0}), m_{k_0}(x_{k_0})) \notin \mathcal{F}_{k_0}$. We further assume that $\theta_{k_0}(x_{k_0}) < \frac{1}{32}\tau_2\tau^2\epsilon_1^2$. (This assumption is rational because of Lemma 1.) If $\frac{1}{2}\tau < \alpha_{k_0+j} < \tau$, for $j > 0$ and $k > k_0$, we then have

$$\theta_k(x_k) \le \theta_{k_0}(x_{k_0})$$
$$m_k(x_k) - m_k(x_k + \alpha_k s_k) \ge \tau_2\alpha_k^2\|s_k\|^2 \ge \frac{1}{4}\tau_2\alpha_k^2\epsilon_1^2 \ge \frac{1}{8}\tau_2\tau^2\epsilon_1^2,$$

and

$$\begin{aligned} m_k(x_k) + \theta_k(x_k) - m_{k+1}(x_k + \alpha_k s_k) - \theta_{k+1}(x_k + \alpha_k s_k) &\ge \tau_2\alpha_k^2\|s_k\|^2 - 2\theta_{k_0}(x_{k_0}) \\ \ge \frac{1}{4}\tau_2\alpha_k^2\epsilon_1^2 2\theta_{k_0}(x_{k_0}) \ge \frac{1}{8}\tau_2\tau^2\epsilon_1^2 - \frac{1}{16}\tau_2\tau^2\epsilon_1^2 = \frac{1}{16}\tau_2\tau^2\epsilon_1^2. \end{aligned}$$

These inequalities are immediately obtained with the similar techniques in Lemma 3. We therefore obtain

$$\infty > \sum_{k=k_0}^{\infty} (m_k(x_k) + \theta_k(x_k) - m_k(x_{k+1}) - \theta_{k+1}(x_{k+1})) \ge \sum_{k=k_0}^{\infty} \frac{1}{16}\tau_2\tau^2\epsilon_1^2, \tag{3.20}$$

where τ_2 is a constant in (3.13). Thus, (3.19) does not hold and the result holds.

If the gradients of $c_i(x_k)$ are linear independent for all k and $i = 1, 2, \ldots, m$, the solution to SNE is obtained by virtue of the result in [14]. The proof is complete. □

The global convergence to Algorithm 1 has been obtained and (3.13) holds. Actually, (3.13) is also satisfied for many inexact line searches. Therefore, the results in this paper can be easily extended.

Table 1
Numerical results for Example 1

Initial point	Iterate number	Number of functions	Number of gradients
(3, 1)	6	8	7
(6, 2)	8	11	10
(24, 8)	13	16	15

Table 2
Numerical results for Example 3

N	Iterate number	Number of functions	Number of gradients
10	12	14	15
20	17	21	22
40	26	32	33

4. Concluding remarks and numerical results

Employing (2.10) and (2.11) as the update rule about S_1 and S_2, some interesting results are obtained in Section 3. As for B_k, we can utilize one-side reduced Hessian update or two-side reduced Hessian update, which is also an important issue. Algorithm 1 is extremely flexible because we just assume that Assumptions 1 and 2 are satisfied.

We now give examples to illustrate Algorithm 1. The first example comes from [4], which converges to a non-stationary point if least square approach is employed. The second example comes from [24], which converges to a non-stationary point. We compute them in MATLAB 6.5 with exact line search. The tolerance is 10^{-5}.

Example 1. Consider the problem of finding a solution of nonlinear system with two variables

$$F(x, y) = \begin{pmatrix} x \\ 10x/(x+0.1) + 2y^2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The unique solution is $(x^\star, y^\star) = (0, 0)$. It has been proved in [4] that, starting from $(x_0, y_0) = (3, 1)$, the iterates converge to the point $z = (1.8016, 0.0000)$, which is not a stationary point. Utilizing the algorithm in this paper, we obtain a sequence of points converging to $(x^\star, y^\star)$. The toleration ϵ in this paper is always 1.0e–6. The iterate number is 43. The detail numerical result is given as follows (Table 1).

Example 2. Consider the following problem of finding a solution of nonlinear system

$$x_i^2 + x_{i+1} = 0, 1 \le i \le n-1$$
$$x_n^2 + x_1 = 0.$$

The root is $x^\star = 0$.

The detail numerical result is given as follows (Table 2) with the initial point $x^0_{N-2} = 0, x^0_i = 0, i \ne N-2$.

Example 3 (*[24]*). Consider the problem of finding a solution of nonlinear system with two variables

$$F(x, y) = \begin{pmatrix} x + 3y^2 \\ (x - 1.0)y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The result is obtained with the number of functions 15 and the number of gradients 12, while Newton iterates fail to find the solution.

Constrained optimization approaches attacking the system of nonlinear equations are exceedingly interesting and are further developed in this paper. Moreover, the local property of the algorithm is another further topic.

Acknowledgement

Sincere thanks to the anonymous reviewers for their helpful suggestions.

References

[1] J.E. Dennis Jr., R.B. Schnable, Numerical Methods for Unconstrained Optimization and Nonlinear Equations, Prentice-Hall, Englewood Cliffs, NJ, 1983.
[2] R. Fletcher, Practical methods of optimization, in: Constrained Optimization, vol. 2, John Wiley and Sons, New York, Toronto, 1981.
[3] J.Y. Fan, A modified Levenberg-Marquardt algorithm for singular system of nonlinear equations, Journal of Computational Mathematics 21 (2003) 625–636.
[4] M.J.D. Powell, A hybrid method for nonlinear equations, in: P. Rabinowitz (Ed.), Numerical Methods for nonlinear Algebraic Equations, Gordon and Breach, London, 1970.
[5] P.Y. Nie, A null space method for solving system of nonlinear equations, Applied Mathematics and Computation 149 (2004) 215–226.
[6] P.Y. Nie, Filter methods for system of nonlinear equations, Acta Nathematicae Applicatae Sinica 27 (2004) 407–415 (in Chinese).
[7] P.Y. Nie, CDT like approaches for the system of nonlinear equations, Applied Mathematics and Computation 172 (2006) 892–902.
[8] P.Y. Nie, An SQP approach with line search for the system of nonlinear equations, Mathematical and Computer Modelling 43 (2006) 368–373.
[9] H. Schichl, A. Neumaier, Exclusion regions for system of equations, SIAM Journal of Numerical Analysis 42 (2004) 383–408.
[10] P.Y. Nie, A derivative-free method for the system of nonlinear equations, Nonlinear Analysis: Real World Applications 7 (2006) 378–384.
[11] R. Fletcher, S. Leyffer, Nonlinear programming without a penalty function, Mathematical Programming 91 (2002) 239–269.
[12] R. Fletcher, N.I.M. Gould, S. Leyffer, et al., Global convergence of a trust-region SQP-filter algorithm for general nonlinear programming, SIAM Journal on Optimization 13 (2003) 635–659.
[13] N.I.M. Gould, S. Leyffer, Ph.L. Toint, A multidimension filter algorithm for nonlinear equations and least squares, SIAM Journal on Optimization 15 (2004) 17–38.
[14] P.Y. Nie, Sequential penalty quadratic programming filter methods for nonlinear programming, Nonlinear Analysis: Real World Applications 8 (2007) 118–129.
[15] R. Fletcher, S. Leyffer, P.L. Toint, On the global convergence of a filter SQP algorithm, SIAM Journal on Optimization 13 (2002) 44–59.
[16] M. Ulbrich, S. Ulbrich, L.N. Vicente, A globally convergent primal-dual interior-point filter method for nonconvex nonlinear programming, Mathematical Programming 100 (2004) 379–410.
[17] S. Ulbrich, On the superlinear local convergence of a filter-SQP methods, Mathematical Programming 100 (2004) 217–245.
[18] P.Y. Nie, J.Y. Fan, A derivate-free filter method for nonlinear complementarity problems, Applied Mathematics and Computation 161 (2005) 787–797.
[19] P.Y. Nie, Composite-step like filter methods for equality constraint problems, Journal of Computational Mathematics 21 (2003) 613–624.
[20] A. Wächter, L.T. Biegler, Line search filter methods for nonlinear programming: Motivation and global convergence, SIAM Journal on Optimization 16 (2005) 1–31.
[21] A. Wächter, L.T. Biegler, Line search filter methods for nonlinear programming: Local convergence, SIAM Journal on Optimization 16 (2005) 32–48.
[22] P.Y. Nie, A new penalty method for nonlinear programming, Computers and Mathematics with Applications 52 (2006) 883–896.
[23] J. Nocedal, S. Wright, Numerical Optimization, Springer, New York, NY, USA, 1999.
[24] R.H. Byrd, M. Marazzi M, J. Nocedal, On the convergence of Newton iterations to non-stationary points, Mathematical Programming 99 (2004) 127–148.