Applied Mathematics and Computation 263 (2015) 47–58

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Line search filter inexact secant methods for nonlinear equality constrained optimization ☆

Zhujun Wang [a,*], Li Cai [b,c], Detong Zhu [b]

[a] College of Science, Hunan Institute of Engineering, Xiangtan 411104, China
[b] Department of Mathematics, Shanghai Normal University, Shanghai 200234, China
[c] Shanghai Research Institute of Microwave Equipment, Shanghai 200063, China

ARTICLE INFO

MSC:
49M37
90C30
65K05

Keywords:
Constrained optimization
Filter method
Inexact method
Secant method
Maratos effect

ABSTRACT

We present inexact secant methods in association with line search filter technique for solving nonlinear equality constrained optimization. For large-scale applications, it is expensive to get an exact search direction, and hence we use an inexact method that finds an approximate solution satisfying some appropriate conditions. The global convergence of the proposed algorithm is established by using line search filter technique. The second-order correction step is used to overcome the Maratos effect, while the line search filter inexact secant methods have superlinear local convergence rate. Finally, the results of numerical experiments indicate that the proposed methods are efficient for the given test problems.

1. Introduction

We consider the problem

$$\min_{x\in\mathbb{R}^n} f(x) \text{ subject to } c(x) = 0, \tag{1.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $c : \mathbb{R}^n \to \mathbb{R}^m$ are twice continuously differentiable.

Fletcher and Leyffer [6] proposed a class of filter methods, which do not require any penalty parameter and have promising numerical results. Consequently, filter technique has employed to many approaches, for instance, SLP methods [4], SQP methods [5,12], interior point approaches [13], bundle techniques [9] and so on. In [10,11], Wächter and Biegler presented a line search filter method for nonlinear equality constrained programming and discussed the global convergence properties and superlinear local convergence rate.

Secant methods (two-step algorithms) as defined in [7] are one of the most successful methods for solving problem (1.1) and have a main advantage which rests in the use of an orthogonal projection operator which is continuous. The basic idea about the secant algorithms can be summarized as follows: at each iteration, the total step d_k is considered as the sum of a horizontal step u_k and a vertical step v_k. The general form of the secant algorithms follows, in each iteration

$$\lambda_{k+1} = U(x_k, \lambda_k, W_k), \tag{1.2a}$$

☆ The authors gratefully acknowledge the partial supports of the National Science Foundation Grant (11371253) of China and the Science Foundation Grant (11C0336) of Provincial Education Department of Hunan.

* Corresponding author. Tel.: +86 21 64323361.
E-mail addresses: wangzj1998@aliyun.com (Z. Wang), caili0039@163.com (L. Cai), dtzhu@shnu.edu.cn (D. Zhu).

http://dx.doi.org/10.1016/j.amc.2015.04.016
0096-3003/

$$W_k\hat{u}_k = -\nabla_x\mathcal{L}(x_k, \lambda_{k+1}), \tag{1.2b}$$

$$u_k = P_k\hat{u}_k, \tag{1.2c}$$

$$v_k = -\nabla c_k^{\dagger} c_k, \tag{1.2d}$$

$$y_k = \nabla_x\mathcal{L}(x_k + u_k, \lambda_{k+1}) - \nabla_x\mathcal{L}(x_k, \lambda_{k+1}), \tag{1.2e}$$

$$W_{k+1} = B(u_k, y_k, W_k), \tag{1.2f}$$

$$x_{k+1} = x_k + u_k + v_k, \tag{1.2g}$$

where $\nabla c_k^{\dagger}$ is the pseudo-inverse of ∇c_k^T, $\mathcal{L}$ is the Lagrangian function defined for $x \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}^m$ by $\mathcal{L}(x,\lambda) = f(x) + \lambda^T c(x)$, and B is a secant update formula generating matrices W_k approximating the Hessian of the Lagrangian at each x_k. The projection onto the null space of $\nabla c(x)^T$ can be either the orthogonal projection $P(x)$ given by

$$P(x) = I - \nabla c(x)[\nabla c(x)^T\nabla c(x)]^{-1}\nabla c(x)^T \tag{1.3a}$$

or the oblique projection defined by

$$P(x) = I - W^{-1}\nabla c(x)[\nabla c(x)^T W^{-1}\nabla c(x)]^{-1}\nabla c(x)^T. \tag{1.3b}$$

The multiplier updates $U(x_k, \lambda_k, W_k)$ in (1.2a) can be chosen from one of the following updates:

$$\lambda_{k+1}^P = -\left(\nabla c_k^T\nabla c_k\right)^{-1}\nabla c_k^T\nabla f_k, \tag{1.4a}$$

$$\lambda_{k+1}^S = -\left(\nabla c_k^T W_k^{-1}\nabla c_k\right)^{-1}\nabla c_k^T W_k^{-1}\nabla f_k, \tag{1.4b}$$

$$\lambda_{k+1}^N = \left(\nabla c_k^T W_k^{-1}\nabla c_k\right)^{-1}\left(c_k - \nabla c_k^T W_k^{-1}\nabla f_k\right), \tag{1.4c}$$

where $\lambda_{k+1}^P, \lambda_{k+1}^S, \lambda_{k+1}^N$ are projection update, null-space update and Newton update, respectively. The pseudo-inverse of ∇c_k^T, $\nabla c_k^{\dagger}$ will be either

$$\nabla c_k^{\dagger} = \nabla c_k\left(\nabla c_k^T\nabla c_k\right)^{-1} \tag{1.5a}$$

or

$$\nabla c_k^{\dagger} = W_k^{-1}\nabla c_k\left(\nabla c_k^T W_k^{-1}\nabla c_k\right)^{-1}. \tag{1.5b}$$

A drawback of many contemporary SQP algorithms is that they require explicit representations of exact derivative information and the solution of one or more linear systems during every iteration. For large-scale applications, it may be too expensive to get. In [2,3], Byrd et al. present inexact methods and give some conditions that guarantee the global convergence of inexact steps.

Motivated by the idea and methods above, we propose an inexact secant method in association with line search filter technique, which has both global convergence and superlinear local convergence rate. The method is globalized by line search and filter approach. It has been noted by Fletcher and Leyffer [6] that the filter approach, similar to a penalty function approach, can suffer from Maratos effect, hence it is necessary to overcome Maratos effect by solving second order correction (SOC) step in our algorithm. In addition, it is shown that the iterative point generated by our algorithm is superlinear convergent.

The paper is organized as follows. In Section 2 the algorithm is developed. Section 3 analyzes the global convergence properties. We characterize the local superlinear convergence in Section 4. The results of numerical experience with the method are discussed in Section 5.

2. Algorithm

Throughout the paper, we denote the Euclidean vector or matrix norm by $\|\cdot\|$, l_1 norm by $\|\cdot\|_1$, $A(x) := \nabla c(x)$, $A(x)^{\dagger} := \nabla c(x)^{\dagger}$ and $g(x) := \nabla f(x)$.

We will compute the horizontal step u_k inexactly, which means that (1.2b) will be substituted by

$$W_k\hat{u}_k = -\nabla_x\mathcal{L}(x_k, \lambda_{k+1}) + r_k, \tag{2.1}$$

where $\|r_k\| \le \eta_k\|c_k\|$ with $\eta_k \in [0, t]$ and $t \in (0, 1)$. We use the norm of c_k to control the residual because $\lim_{k\to\infty}\|c(x_k)\| = 0$, which will be shown in Theorem 3.1.

Because of adopting l_1 penalty function as a merit function, it is necessary to overcome Maratos effect by solving SOC step $\hat{d}_k$ in our algorithm. In this work, $\hat{d}_k$ can be either

$$\hat{d}_k = -A_k\left(A_k^T A_k\right)^{-1} c(x_k + d_k) \tag{2.2a}$$