Applied Mathematics and Computation 240 (2014) 72–81

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

A dwindling filter trust region algorithm for nonlinear optimization

Chao Gu [a,*], Detong Zhu [b]

[a] School of Math. and Info., Shanghai LiXin University of Commerce, Shanghai 201620, PR China
[b] Department of Mathematics, Shanghai Normal University, Shanghai 200234, PR China

ARTICLE INFO

Keywords:
Nonlinear optimization
Trust region
Dwindling filter
Feasibility restoration phase
Convergence

ABSTRACT

In this paper, we develop dwindling filter idea and trust region strategy for solving nonlinear optimization. Compared with traditional filter trust region algorithm, the new filter algorithm has more flexibility for the acceptance of the trial step and the feasibility restoration phase is not needed. Under mild conditions, the global convergence of the new algorithm is analyzed. Preliminary numerical results are reported.

1. Introduction

Consider the following optimization problem with nonlinear equality constraint

$$\min\ f(x), \tag{1.1a}$$
$$\text{subject to } c(x) = 0, \tag{1.1b}$$

where $f(x) : \mathbb{R}^n \to \mathbb{R}^1$ and $c(x) : \mathbb{R}^n \to \mathbb{R}^m$ with $m < n$ are continuously differentiable.

Recently, Fletcher and Leyffer [6] presented filter methods for nonlinear programming (NLP), offering an alternative to merit functions, as a tool to guarantee global convergence of algorithms for nonlinear programming. The underlying concept is that trial points are accepted if they improve the objective function or improve the constraint violation. The practical results reported for the filter trust region sequential quadratic programming (SQP) method in [6] were encouraging, and subsequently global convergence results for related algorithms were established by Fletcher et al. [7,8]. Ulbrich [23] proved fast local convergence by making modifications to the filter SQP method. Nie and Ma [14,15] discussed the trust region filter method for general nonlinear programming. Su, Pu and Liu studied a nonmonotone filter trust region method [20] and a modified filter trust region method [21]. Filter strategies are focused on because of promising numerical results [1–3,9–11,13,19,22,24,25]. However, the filter methods following Fletcher and Leyffer's rule for step acceptance need a feasibility restoration phase, which may spend a large amount of computation. It is interesting for us to study the restoration-free filter algorithms.

In this paper, we develop dwindling filter idea and trust region strategy for solving nonlinear optimization. Compared with traditional filter trust region algorithm, the new algorithm has more flexibility for the acceptance of the trial step and the feasibility restoration phase is not needed. Under mild conditions, the global convergence of the new algorithm is analyzed. The paper is outlined as follows. In Section 2, we state the dwindling filter trust region algorithm; the global convergence of the algorithm is proved in Section 3; finally, we report some numerical experiments in Section 4.

* Corresponding author.
E-mail addresses: chaogumath@gmail.com (C. Gu), dtzhu@shnu.edu.cn (D. Zhu).

http://dx.doi.org/10.1016/j.amc.2014.04.101
0096-3003/

Notation. $\|\cdot\|$ is the ordinary Euclidean norm in the paper. Let $g(x)$ denote the gradient of the function $f(x)$, $A(x)^T$ denote the Jacobian of the constraint $c(x)$.

2. A dwindling filter trust region algorithm

Following Dennis, El Alem and Maciel [4], Omojokun [17], we obtain the trial step $s_k = s_k^t + s_k^n$ at the current iterate x_k by computing a quasi-normal step s_k^n and a tangential step s_k^t. The purpose of the quasi-normal step s_k^n is to improve feasibility. It is obtained as approximate solution of the trust-region subproblem

$$\min \|c_k + A_k^T s^n\|^2, \quad s.t.\ \|s^n\| \leqslant \Delta_k, \tag{2.1}$$

where the $\Delta_k > 0$ denotes the trust region radius. The requirements on the step s_k^n are that there exist constants $K_1, K_2 > 0$, such that s_k^n admits the upper bound

$$\|s_k^n\| \leqslant \min\{K_1\|c_k\|, \Delta_k\} \tag{2.2}$$

and satisfies the decrease condition

$$\|c_k\|^2 - \|c_k + A_k^T s_k^n\|^2 \geqslant K_2\|c_k\| \min\{\|c_k\|, \Delta_k\}. \tag{2.3}$$

Similar to [8], we replace our condition that $\|s_k^n\| \leqslant \Delta_k$ with the stronger requirement that

$$\|s_k^n\| \leqslant \kappa_\Delta \Delta_k \min\left[1, \Delta_k^\nu\right] \tag{2.4}$$

for some $\kappa_\Delta \in (0,1]$, $\nu \in (0,1)$. We define a quadratic model

$$q_k(s) = f_k + g_k^T s + \frac{1}{2} s^T H_k s \tag{2.5}$$

of $f(x)$ about the current point x_k, where H_k is a symmetric approximation of the Hessian of Lagrangian function

$$\ell(x, \lambda) = f(x) + \lambda^T c(x).$$

Based on the model, the tangential step s_k^t is computed as approximate solution of the trust region subproblem

$$\min q_k\left(s_k^n + s^t\right), \quad s.t.\ A_k^T s^t = 0, \quad \|s^t\| \leqslant \Delta_k \tag{2.6}$$

satisfying the decrease condition

$$q_k\left(s_k^n\right) - q_k\left(s_k^n + s_k^t\right) \geqslant K_3\left\|Z_k^T \nabla q_k\left(s_k^n\right)\right\| \min\left\{\left\|Z_k^T \nabla q_k\left(s_k^n\right)\right\|, \Delta_k\right\} \tag{2.7}$$

with a constant $K_3 > 0$, where $Z(x)$ denotes a matrix whose columns form a basis of the null space of $A(x)^T$. Assume that $A(x)$ has full column rank, then a QR decomposition can be performed, that is,

$$A(x) = \begin{pmatrix} Y(x) & Z(x) \end{pmatrix} \begin{pmatrix} R(x) \\ 0 \end{pmatrix},$$

where $\begin{pmatrix} Y(x) & Z(x) \end{pmatrix}$ is an orthogonal matrix, $R(x)$ is nonsingular upper triangular matrix of order m, and $Z(x) \in \mathbb{R}^{n\times(n-m)}$. The column vectors of $Z(x)$ form an orthonormal basis of the null space $N(A(x)^T)$ and the column vectors of $Y(x) \in \mathbb{R}^{n\times m}$ form an orthonormal basis of the range space $R(A(x))$.

To evaluate the descent properties of the step for the objective function, we use the predicted reduction of $f(x)$

$$pred_k = q_k(0) - q_k(s_k)$$

and the actual reduction of $f(x)$

$$ared_k = f_k - f(x_k + s_k).$$

The first order necessary optimality conditions at a local solution x^* of (1.1) can be written as

$$c(x^*) = 0, \quad Z(x^*)^T g(x^*) = 0.$$

Definition 1. $\psi(\Delta) : [0, +\infty) \to \mathbb{R}$ is a dwindling function if it is a monotonically increasing and continuous function such that

$$\psi(\triangle) = 0 \iff \triangle = 0, \tag{2.8}$$

$$\triangle \geqslant 1 \iff \psi(\Delta) = 1, \tag{2.9}$$

$$\lim_{\Delta\to 0} \frac{\psi(\Delta)}{\Delta} = 0. \tag{2.10}$$