Journal of Computational Mathematics, Vol.21, No.5, 2003, 613–624.

COMPOSITE-STEP LIKE FILTER METHODS FOR EQUALITY CONSTRAINT PROBLEMS *[1)]

Pu-yan Nie
(*LSEC, ICMSEC, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing 100080, China*)

Abstract

In a composite-step approach, a step s_k is computed as the sum of two components v_k and h_k. The normal component v_k, which is called the vertical step, aims to improve the linearized feasibility, while the tangential component h_k, which is also called horizontal step, concentrates on reducing a model of the merit functions. As a filter method, it reduces both the infeasibility and the objective function. This is the same property of these two methods. In this paper, one concerns the composite-step like filter approach. That is, a step is tangential component h_k if the infeasibility is reduced. Or else, s_k is a composite step composed of normal component v_k and tangential component h_k.

Key words: Composite-step like approaches, Filter methods, Equality constraints, Sequential quadratic programming(SQP) algorithms, Normal component, Tangential component, Convergence.

1. Introduction

In this paper, we consider the problem of minimizing a (linear or nonlinear) function f of n real variables x, which satisfy a set of (linear or nonlinear) constraints $c_i(x) = 0, i = 1, \cdots, m$, namely

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & c(x) = 0, \end{array} \tag{1.1}$$

where $x \in R^n$. The functions $c : R^n \to R^m, m < n, f : R^n \to R$ are assumed to be continuously differentiable. Then, we introduce composite-step like approaches and filter methods respectively.

1.1 Composite-Step Like Methods

An approach, whose every step s_k is consisted of two components v_k and h_k, is termed composite-step method. Where the normal component v_k is to degrade the degree of constraint violation, while the tangential component h_k aims to reduce a model of the merit functions. There are two kinds of composite-step like approaches. One is Vardi-like methods. The other is Byrd-Omojokun-like approaches.

A: Vardi-like methods

From (1.1) one recognizes that the set

$$F_k = \{d | c(x_k) + A(x_k)d = 0 \ \text{ and } \|d\| \leq \Delta_k\}, \tag{1.2}$$

may be empty, where $A(x_k) = \bigtriangledown c(x_k)$. Vardi[19] and Byrd, Schnable and Schultz[5] instead relax the linearized constraints so that $\alpha_k c(x_k) + A(x_k)d = 0$ for some $0 < \alpha_k \leq 1$ for which

$$F_k(\alpha_k) = \{d | \alpha_k c(x_k) + A(x_k)d = 0 \text{ and } \|d\| \leq \Delta_k\}, \tag{1.3}$$

* Received April 4, 2001; final revised March 3, 2002.
[1)] Supported partially by Chinese NNSF grants 19731010 and the knowledge innovation program of CAS.

is not empty. Clearly, $F_k(0)$ and any $\alpha_k \le \alpha_{max}$ is not empty, where α_{max} is the solution of the following problem.

$$\max_{\alpha\in(0,1]}\ \min_{\|d\|\le\Delta_k} \|\alpha c(x_k) + A(x_k)d\| = 0.$$

Certainly, it may be expensive to find α_{max}. In practice, to obtain the normal component an approximation v_k^c to $v^c(x_k)$ for which $c(x_k) + A(x_k)v = 0$ may be computed instead, and α_k subsequently found so that $v_k = \alpha_k v_k^c$ lies in the trust region.

The normal step found, the tangential component is chosen to reduce a model of the merit function. Specially, if we consider a merit function of the form

$$\phi(x,\sigma) = f(x) + \sigma\|c(x)\|. \tag{1.4}$$

Let $m_k(x_k + s) = q(x_k + s) + \sigma m_k^N(x_k + s)$ where

$$q(x_k + s) = f(x_k) + s^T g(x_k) + \frac{1}{2}s^T H_k s \text{ and } m_k^N(x_k + s) = \|c(x_k) + A(x_k)s\|. \tag{1.5}$$

In a tangential component, the following conditions are satisfied

$$m_k^N(x_k + v_k + h_k) = m_k^N(x_k + v_k) \text{ and } q(x_k + v_k + h_k) < q(x_k + v_k).$$

Thus, the tangential component is obtained by approximately solving the problem.

$$\begin{array}{ll} \text{minimize} & h^T(g(x_k) + H_k v_k) + \frac{1}{2}h^T H_k h \\ \text{subject to} & A(x_k)h = 0 \\ & \|h\| \le \Delta_k - \|v_k\|. \end{array} \tag{1.6}$$

Of course, the above requirements may readily be satisfied using suitable conjugate-gradient methods.

B: Byrd-Omojokun-like approaches

Byrd-Omojokun-like approach is proposed by Omojokun [14] and Byrd, Gilbert and Nocedal[3]. And it forms ETR, NITRO and BECTR algorithms, which is given by Lalee, Nocedal and Plantenga[13], Byrd, Hribar and Nocedal[4], and Plantenga[15], respectively. In a Byrd-Omojokun method, it is not required that the linearized constraints should be compatible. That is, the main difference lies in the computation of the normal step. Instead of shifting the linearized constraints, to obtain v_k one solves the following subproblem approximately.

$$\begin{array}{ll} \text{minimize} & \|c(x_k) + A(x_k)v_k\| \\ \text{subject to} & \|v_k\| \le \xi^N \Delta_k, \end{array} \tag{1.7}$$

for some $0 < \xi^N < 1$.

Apparently, (1.7) may have many solutions. Obtaining an exact solution to (1.7) may be costly. A cheaper choice is to calculate an approximate solution giving a reduction in $\|c(x_k) + A(x_k)v\|$ no worse than a fraction of Cauchy point for this problem, which is,

$$v_k^c = -\alpha_k^c A(x_k)^T c(x_k), \tag{1.8}$$

where

$$\alpha_k^c = \arg_{0\le\alpha\le\xi^N\Delta_k/\|A(x_k)^T c(x_k)\|} \min \|c(x_k) - \alpha A(x_k)A(x_k)^T c(x_k)\|.$$

At every step, (1.8) is satisfied if suitable conjugate-gradient method is used. Meanwhile, (1.5) is also met.

For Vardi-like approaches and Byrd-Omojokun-like methods, the superlinear convergence is obtained under suitable assumptions.

1.2. Filter Technique

Filter approach is proposed by Fletcher and Leyffer[8]. It has been applied extensively so far. In [9], filter method is used to SLP(sequential linear programming) and its global convergence