Global Convergence of a Non-monotone Trust-Region Filter Algorithm for Nonlinear Programming

Nicholas I. M. Gould\(^1\) and Philippe L. Toint\(^2\)

\(^1\) Rutherford Appleton Laboratory, Computational Science and Engineering Department, Chilton, Oxfordshire, England. n.i.m.gould@rl.ac.uk
\(^2\) University of Namur, Department of Mathematics, 61, rue de Bruxelles, B-5000 Namur, Belgium. philippe.toint@fundp.ac.be

Summary. A non-monotone variant of the trust-region SQP-filter algorithm analyzed in Fletcher et al (SIAM J. Opt. 13(3), 2002, pp. 653–659) is defined, that directly uses the dominated area of the filter as an acceptability criterion for trial points. It is proved that, under reasonable assumptions and for all possible choices of the starting point, the algorithm generates at least a subsequence converging to a first-order critical point.

1 Introduction

Our objective is to define and analyze a new algorithm for solving constrained minimization problems where both the objective function and the constraints are smooth, that is

\[
\begin{align*}
\text{minimize } & f(x) \\
\text{subject to } & c_\ell(x) = 0 \\
& c_\ell(x) \geq 0,
\end{align*}
\]

where \( f \) is a twice continuously differentiable real valued function of the variables \( x \in \mathbb{R}^n \) and \( c_\ell(x) \) and \( c_\ell(x) \) are twice continuously differentiable functions from \( \mathbb{R}^n \) into \( \mathbb{R}^m \) and from \( \mathbb{R}^n \) into \( \mathbb{R}^p \), respectively. Let \( c(x)^T = (c_\ell(x)^T \quad c_\ell(x)^T) \). Note that no convexity assumption is made.

The algorithms that we discuss belongs is a trust-region filter method, and belong as such to a class of algorithms introduced by [FL02]. A global convergence theory for this class is proposed in [FLT98], in which the objective function is locally approximated by a linear function, leading, at each iteration, to the (exact) solution of a linear program. Similar results are shown in [FLT02], where the approximation of the objective function is quadratic, leading to a Sequential Quadratic Programming (SQP) method. However, this is accomplished at the (very high) cost of finding a global minimizer of the possibly nonconvex quadratic programming subproblem. This latter requirement
is relaxed in [FGLTW02], where the SQP step is decomposed in “normal” and “tangential” components.

The main purpose of the current paper, a companion of [FGLTW02], is to analyze an algorithm where the filter acceptance criterion for new iterates is relaxed to allow dominated iterates to be accepted in some cases. This is potentially important as it is known that SQP method can generate such iterates in their asymptotic fast convergence phase. The theory developed here therefore provides a possible convergence framework for a filter method with quadratic convergence properties without the need to introduce second-order corrections. Results along this line are already known for linesearch-filter methods [WB01], and for another variant of trust-region-filter methods where the definition of filter entries is modified [Ulb04]. Our objective is to introduce a framework suitable for trust-region-filter methods using the original definition of the filter entries. Moreover, the new theory is that it no longer needs the notion of a “margin” around the filter, a device which is common to all theoretical approaches of the filter method so far.

2 A Non-monotone Filter Algorithm

As indicated above, the algorithm that we are about to describe is of the SQP type. At a given iterate \( x_k \), Newton’s method is implicitly applied to solve (a local version of) the first-order necessary optimality conditions by solving the quadratic programming subproblem \( QP(x_k) \) given by

\[
\begin{align*}
\text{minimize} & \quad f_k + \langle g_k, s \rangle + \frac{1}{2} \langle s, H_k s \rangle \\
\text{subject to} & \quad c_G(x_k) + A_G(x_k) s = 0 \\
& \quad c_T(x_k) + A_T(x_k) s \geq 0,
\end{align*}
\]

where \( f_k = f(x_k) \), \( g_k = g(x_k) \) \( \text{def} \equiv \nabla_x f(x_k) \), where \( A_G(x_k) \) and \( A_T(x_k) \) are the Jacobians of the constraint functions \( c_G \) and \( c_T \) at \( x_k \) and where \( H_k \) is a symmetric matrix. We will not immediately be concerned about how \( H_k \) is obtained, but we will return to this point in Section 3. The solution of \( QP(x_k) \) then yields a step \( s_k \). If \( s_k = 0 \), then \( x_k \) is first-order critical for problem (1).

2.1 The composite SQP step

The step \( s_k \) is typically computed by solving, possibly approximately, a variant of (2). In the trust-region approach, one takes into account the fact that (2) only approximates our original problem locally: the step \( s_k \) is thus restricted in norm to ensure that \( x_k + s_k \) remains in a trust-region centred at \( x_k \), where we believe this approximation to be adequate. The subproblem \( QP(x_k) \) is thus replaced by its \( TRQP(x_k, \Delta_k) \) variant given by