Projected Quasi-Newton Algorithm with Trust Region for Constrained Optimization

J. Z. Zhang\(^2\) and D. T. Zhu\(^3\)

Communicated by D. F. Shanno

Abstract. In Ref. 1, Nocedal and Overton proposed a two-sided projected Hessian updating technique for equality constrained optimization problems. Although local two-step Q-superlinear rate was proved, its global convergence is not assured. In this paper, we suggest a trust-region-type, two-sided, projected quasi-Newton method, which preserves the local two-step superlinear convergence of the original algorithm and also ensures global convergence. The subproblem that we propose is as simple as the one often used when solving unconstrained optimization problems by trust-region strategies and therefore is easy to implement.

Key Words. Two-sided projected Hessians, trust regions, differentiable penalty functions, global convergence, two-step Q-superlinear rate, constrained optimization.

1. Introduction

Consider the following optimization problem with nonlinear equality constraints:

\[
\begin{align*}
\min & \quad f(x), \\
\text{s.t.} & \quad c(x) = 0,
\end{align*}
\]

where \( f: \mathbb{R}^n \to \mathbb{R} \), \( c: \mathbb{R}^n \to \mathbb{R}^m \), \( m \leq n \). Nocedal and Overton proposed a two-sided projected Hessian updating method in Ref. 1. Its basic idea can be summarized as follows.

---

\(^1\) This research was supported in part by the National Natural Science Foundation of China.

\(^2\) Professor, Department of Mathematics, Shanghai Normal University, Shanghai, China. Currently, Department of Applied Mathematics, City Polytechnic of Hong Kong, Hong Kong.

\(^3\) Lecturer, Department of Mathematics, Shanghai Normal University, Shanghai, China.
Let
\[ g(x) = \nabla f(x) \in \mathbb{R}^n, \]
\[ c(x) = (c_1(x), \ldots, c_m(x)) \in \mathbb{R}^m, \]
\[ A(x) = \nabla c(x) = [\nabla c_1(x), \ldots, \nabla c_m(x)] \in \mathbb{R}^{n \times m}. \]

Assuming that \( A(x) \) has full column rank, one can make a QR-decomposition of \( A(x) \) as follows:
\[ A(x) = [Y(x), Z(x)] \begin{bmatrix} R(x) \\ 0 \end{bmatrix} = Y(x)R(x), \tag{2} \]
where \([Y, Z]\) is an orthogonal matrix, \( R(x) \) is a nonsingular upper triangular matrix of order \( m \), and \( Z(x) \in \mathbb{R}^{n \times t} \), where \( t = n - m \). The column vectors of \( Z(x) \) form an orthonormal basis for the null space \( N(A(x)^T) \); i.e.,
\[ A(x)^T Z(x) = 0. \tag{3} \]

The columns of \( Y(x) \in \mathbb{R}^{n \times m} \) form a normal orthogonal basis of the range space \( R(A(x)) \) of \( A(x) \). Clearly,
\[ Y(x)^T Y(x) = I_m, \quad Z(x)^T Z(x) = I_t, \tag{4a} \]
\[ Y(x)^T X(x)^T + Z(x) Z(x)^T = I_n. \tag{4b} \]

Let
\[ L(x, \lambda) = f(x) - A(x)^T c(x) \]
be the Lagrangian of problem (1), where \( \lambda \) is the solution vector of the least-square problem
\[ \min_{\lambda} \| A(x) \lambda - g(x) \|. \]

From (2), we have
\[ \lambda(x) = (A(x)^T A(x))^{-1} A(x)^T g(x) = R(x)^{-1} Y(x)^T g(x). \tag{5} \]

Let
\[ W(x, \lambda) = \nabla^2_{xx} L(x, \lambda) \tag{6} \]
be the Hessian matrix of the function \( L(x, \lambda) \) with respect to \( x \). A principal distinction between the Nocedal-Overton method and the usual quasi-Newton methods is that, in the former, the updating matrix \( B \in \mathbb{R}^{t \times t} \) is an approximation of the square matrix \( Z(x)^T W(x, \lambda) Z(x) \) of order \( t \), whereas in the latter the updating matrices approximate \( W(x, \lambda) \).