第32卷第4期 上海师范大学学报(自然科学版) Vol. 32, No. 4
2003年12月 Journal of Shanghai Normal University(Natural Sciences) Dec. 2003

Projected Hessian Algorithm with Backtracking Interior Point Technique for Linear Constrained Optimization

ZHU De-tong

(Mathematics and Sciences College, Shanghai Normal University, Shanghai 200234, China)

Abstract: We propose a new trust region projected Hessian algorithm with nonmonotonic backtracking interior point technique for linear constrained optimization. Based on performing QR decomposition of an affine scaling equality constraint matrix, the conducted subproblem in the algorithm is the general trust region subproblem defined by minimizing a quadratic function subject only to an ellipsoidal constraint. By using both trust region strategy and line search technique, each iterate switches to backtracking interior point step generated by the trust region subproblem. The global convergence and fast local convergence rate of the proposed algorithm are established under some reasonable conditions. A nonmonotonic criterion is used to speed up the convergence progress in some ill—conditioned cases.

Key words: trust region method; interior point backtracking; nonmonotone

CLC number: O221.2 **Document code**: A **Article ID**: 1000-5137(2003)04-0007-07

1 Introduction

We analyze the trust region interior point algorithm for solving the linear equality constrained optimization problem:

$$\min\ f(x) \quad \text{s.t.}\ Ax = b,\ x \geqslant 0,$$

where f: $\mathbf{R}^n \to \mathbf{R}$ is smooth and nonlinear, but not necessarily convex, $A \in \mathbf{R}^{m\times n}$ is a matrix and $b \in \mathbf{R}^m$ is a vector. Recently, COLEMAN and LI in [1] presented an interior double-trust region algorithm for minimization problem with simple bounds on the variables. Trust region method is a well—accepted technique in nonlinear optimization to assure global convergence. However, the search direction must be strictly feasible which should bring about the difficulty of computation, and the total computation for completing one iteration might be expensive. A combination idea of the trust region and line search method (see [4]) motivates to switch to the line search technique by employing the

Received date: 2002-12-05

Foundation item: The author gratefully acknowledges the partialsupports of the key project of Applied Mathematics of Shanghai Normal University

Biography: ZHU De-tong(1954-), male, professor, doctor, Mathematics and Sciences College, Shanghai Normal University.

backtracking steps at trial step which may be unaccepted in trust region strategy, since the trial step should provide a direction of sufficient descent. The nonmonotonic line search and trust region techniques for solving unconstrained optimization are respectively proposed by Grippo et. al in [3] and Deng et. al in [2]. The nonmonotonic idea also motivates to further study the trust region projected reduced Hessian algorithm with backtracking interior point technique for solving (1.1) in this paper, because monotonicity may cause a series of very small steps if the contours of objective function f are a family of curves with large curvature. In the paper, based on performing QR decomposition of an affine scaling equality constraint matrix, the conducted subproblem in the algorithm is the general trust region subproblem defined by minimizing a quadratic function subject only to an ellipsoidal constraint.

The paper is organized as follows. In Section 2, we describe the algorithm which combines the techniques of trust region, interior point, backtracking step and nonmonotonic search. In Section 3, the weak global convergence of the proposed algorithm is established. Some further convergence properties such as strong global convergence and local convergence rate are discussed in Section 4.

2 Algorithm

In this section, we propose a trust region projected Hessian method with nonmonotonic backtracking interior technique for linear constrained optimization. The backtracking step generated by the trust region subproblem involves choosing a scaling matrix D_k and a quadratic model $\bar{\varphi}_k(\bar{d}^z)$. We motivate our choice of scaling matrix by examining the optimality conditions for (1.1) and get the reduced Hessian by performing QR decomposition of an affine scaling equality constraint matrix.

Optimality conditions for problem (1.1) are well established. Assuming feasibility, first-order necessary conditions for x_* to be a local minimizer are that there exist $0\leqslant \nu_*\in \mathbf{R}^n$ and $\lambda_*\in \mathbf{R}^m$ such that

$$g_* + A^T\lambda_* - \nu_* = 0,\ Ax_* = b,\ \nu_*^T x_* = 0. \tag{2.1}$$

Equivalently,

$$\begin{cases}(g_* + A^T\lambda)_i = 0, & \text{if } (x_*)_i > 0\\ (g_* + A^T\lambda)_i \geqslant 0, & \text{if } (x_*)_i = 0\end{cases} \tag{2.2}$$

where $(g_* + A^T\lambda_*)_i$ and $(x_*)_i$ are the ith components of $(g_* + A^T\lambda_*)$ and x_*, respectively. We now define a vector function $\gamma(x)$: $\mathbf{R}^n \to \mathbf{R}^n$ and the ith component of the vector function defined componentwise as follows:

$$\gamma_i(x) \stackrel{\text{def}}{=} \begin{cases}x_i, & \text{if } (g+A^T\lambda)_i \geqslant 0,\\ -1 & \text{if } (g+A^T\lambda)_i < 0.\end{cases} \tag{2.3}$$

Defining $D(x) \stackrel{\text{def}}{=} \mathrm{diag}\{|\gamma_1(x)|^{\frac{1}{2}},\cdots,|\gamma_n(x)|^{\frac{1}{2}}\}$, which arise naturally from examining the first—order necessary conditions for the problem (1.1)

$$D^{-2}(x)\{g(x) + A^T\lambda\} = 0,\ Ax = b, \tag{2.4}$$

where the Lagrange multiplier λ is the solution vector of the least squares problem

$$\min_{\lambda} \| A^T\lambda + g(x) \|_{D(x)^{-2}}.$$

By solving the normal equation of the above problem, we have

$$\lambda(x) = -(\bar{A}(x)\bar{A}(x)^T)^{-1}\bar{A}(x)\bar{g}(x). \tag{2.5}$$

where set $\bar{A}(x) \stackrel{\text{def}}{=} D(x)^{-1}A$ and $\bar{g}(x) \stackrel{\text{def}}{=} D(x)^{-1}g(x)$. We define the following sets:

$$\mathscr{F} \stackrel{\text{def}}{=} \{x\in \mathbf{R}^n \mid Ax = b,\ x\geqslant 0\},\ \text{and}\ \mathscr{F}^0 \stackrel{\text{def}}{=} \{x\in \mathbf{R}^n \mid Ax = b,\ x > 0\}$$

so that $\mathcal{F}$ is the set of feasible points and $\mathcal{F}^0$ is the set of 'strictly feasible' interior points. In the sequel, we assume that $\mathcal{F}$ is bounded and $\mathcal{F}^0$ is nonempty, system (2.4) is continuous but not everywhere differentiable. Nondifferentiability occurs when $\gamma_i = 0$. Discontinuity of γ_i may also occur when $(g + A^T\lambda)_i = 0$. Assume that $x_k \in \mathcal{F}^0$, a Newton step for (2.4) satisfies

$$(D_k^{-2}\nabla^2 f(x_k) + \operatorname{diag}\{g_k + A^T\lambda_k\}J_k^\gamma)d_k + D_k^{-2}A^T\Delta\lambda_k = -D_k^{-2}(g_k + A^T\lambda_k),\ Ad_k = 0 \quad (2.6)$$

where $\operatorname{diag}\{g_k + A^T\lambda_k\} = \operatorname{diag}\{(g_k + A^T\lambda_k)_1, \cdots, (g_k + A^T\lambda_k)_n\}$ and $J^\gamma(x) \in \mathbf{R}^{n\times n}$ is the Jacobian matrix of $|\gamma(x)|$ whenever $|\gamma(x)|$ is differentiable. Each diagonal component of the diagonal matrix J^γ equals zero or 1.

Let (x_k, λ_k) be the kth iteration of the Newton process, which is defined by

$$\begin{pmatrix} x_{k+1} \\ \lambda_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ \lambda_k \end{pmatrix} + \begin{pmatrix} d_k \\ \Delta\lambda_k \end{pmatrix}$$

so that (2.6) can be rewritten as follows

$$(D_k^{-2}\nabla^2 f(x_k) + \operatorname{diag}\{g_k + A^T\lambda_k\}J_k^{\gamma})d_k = -D_k^{-2}(g_k + A^T\lambda_{k+1})$$
$$x_{k+1} = x_k + d_k,\ Ad_k = 0. \quad (2.7)$$

Multiplying the two sides of the first equation (2.7) by D_k, and defining

$$\bar{d}_k \stackrel{\text{def}}{=} D_k d_k,\quad \bar{g}_k \stackrel{\text{def}}{=} D_k^{-1}g_k,\quad \bar{A}_k \stackrel{\text{def}}{=} AD_k^{-1},$$
$$\bar{H}_k \stackrel{\text{def}}{=} D_k^{-1}\nabla^2 f(x_k)D_k^{-1},\quad \bar{C}_k \stackrel{\text{def}}{=} \operatorname{diag}\{g_k + A^T\lambda_k\}J_k^\gamma, \quad (2.8)$$

we can obtain that (2.8) is equivalent to

$$(\bar{H}_k + \bar{C}_k)\bar{d}_k = -(\bar{g}_k + \bar{A}_k^T\lambda_{k+1}),\quad \bar{A}_k\bar{d}_k = 0,\quad x_{k+1} = x_k + D_k^{-1}\bar{d}_k \quad (2.9)$$

Assuming $\bar{A}(x)$ has full row rank m, then QR decomposition can be performed, that is,

$$\bar{A}(x) = [\,R(x),\ 0\,]\begin{bmatrix} Y(x) \\ Z(x) \end{bmatrix} \quad (2.10)$$

where $\begin{bmatrix} Y(x) \\ Z(x) \end{bmatrix}$ is an orthogonal matrix, $R(x)$ is a nonsingular lower triangular matrix of order m. The row vectors of $Z(x)$ form an orthonormal basis for the null subspace $\mathcal{N}(\bar{A}(x))$, i.e., $\bar{A}(x)Z(x)^T = 0$. The rows of $Y(x)$ form an orthonormal basis of the range $\mathcal{R}(\bar{A}(x)^T)$. The central idea is to rewrite the first equation in the system (2.9) as

$$Z_k(\bar{H}_k + \bar{C}_k)\bar{d}_k = -Z_k\bar{g}_k.$$

Using the orthogonal matrix $\begin{bmatrix} Y(x) \\ Z(x) \end{bmatrix}$, (2.9) can be rewritten as

$$\begin{bmatrix} Z_k(\bar{H}_k + \bar{C}_k) \\ \bar{A}_k \end{bmatrix}[Y_k^T\ \ Z_k^T]\begin{bmatrix} Y_k \\ Z_k \end{bmatrix}\bar{d}_k = -\begin{bmatrix} Z_k\bar{g}_k \\ 0 \end{bmatrix}$$

that is,

$$\begin{bmatrix} Z_k(\bar{H}_k + \bar{C}_k)Y_k^T & Z_k(\bar{H}_k + \bar{C}_k)Z_k^T \\ R_k & 0 \end{bmatrix}\begin{bmatrix} \bar{d}_k^y \\ \bar{d}_k^z \end{bmatrix} = -\begin{bmatrix} Z_k\bar{g}_k \\ 0 \end{bmatrix}$$
$$\bar{d}_k = Y_k^T\bar{d}_k^y + Z_k^T\bar{d}_k^z,\quad x_{k+1} = x_k + D_k^{-1}\bar{d}_k. \quad (2.11)$$

This system now reduces to solving

$$[Z_k(\bar{H}_k + \bar{C}_k)Z_k^T]\bar{d}_k^z = -Z_k\bar{g}_k,\quad \bar{d}_k = Z_k^T\bar{d}_k^z,\quad x_{k+1} = x_k + D_k^{-1}\bar{d}_k, \quad (2.12)$$

since $\bar{d}_k^y = 0$ is the solution of $R_k\bar{d}_k^y = 0$ and replacing $Z_k(\bar{H}_k + \bar{C}_k)Z_k^T$ by an approximation matrix M_k, the trust region subproblem is as follows

$$(S_k) \qquad \min\ \bar{\varphi}_k(\bar{d}^z) \stackrel{\text{def}}{=} (Z_k\bar{g}_k)^T + \frac{1}{2}(\bar{d}^z)^T M_k\bar{d}^z$$
$$\text{s.t.}\ \ \|\bar{d}^z\| \leqslant \Delta_k$$

where $M_k \stackrel{\text{def}}{=} Z_k D_k^{-1}(\nabla^2 f_k + diag\{g_k + A^T\lambda_k\} J_k^v) D_k^{-1} Z_k^T$, and Δ_k is a trust region radius. The Lagrange multiplier λ_{k+1} can be obtained by solving the upper triangular equation

$$R_{k+1}{}^T\lambda_{k+1} = Y_{k+1}\bar{g}_{k+1}. \tag{2.13}$$

Based on solving the about general trust region subproblem (S_k), we give the following lemma (see [5]) which establishes the necessary and sufficient condition concerning ν_k and $\bar{d}_k^z$, when $\bar{d}_k^z$ solves the subproblem (S_k). The lemma also implies that x_k is a local minimizer of (1.1) if and only if $\bar{d}_k^z = 0$ is a solution of the subproblem (S_k).

Lemma 2.1 $\bar{d}_k^z$ is a solution of subproblem (S_k) if and only if there exists $0 \leqslant \nu_k \in \mathbf{R}^{n-m}$, such that

$$(M_k + \nu_k I_k)\bar{d}_k^z = -Z_k\bar{g}_k, \quad \nu_k(\Delta_k - \|\bar{d}_k^z\|) = 0 \tag{2.14}$$

holds and $M_k + \nu_k I_k$ is positive semidefinite.

Next we develop a trust region projected Hessian algorithm which combines nonmonotonic line search interior technique based on the trust region subproblem (S_k).

Algorithm

Initialization step

Choose parameters $\beta \in (0, \frac{1}{2})$, $\omega \in (0,1)$, $0 < \eta_1 < \eta_2 < 1$, $0 < \gamma_1 < \gamma_2 < 1 < \gamma_3$, $\varepsilon > 0$ and positive integer M. Let $m(0) = 0$. Choose a symmetric matrix B_0. Select an initial trust region radius $\Delta_0 > 0$ and a maximal trust region radius $\Delta_{max} \geqslant \Delta_0$, give a starting strictly feasible interior point $x_0 \in \mathscr{F}^0$. Set $k \leftarrow 0$, go to the main step.

Main step

1. Evaluate $f_k = f(x_k)$, $g_k = \nabla f(x_k)$, D_k and $\bar{A}_k$. Make a QR decomposition $\bar{A}_k$ to get Y_k, Z_k and R_k given in (2.10).
2. If $\|Z_k\bar{g}_k\| \leqslant \varepsilon$, stop with the approximate solution x_k.
3. Solve subproblem

$$(S_k) \qquad \min\ \bar{\varphi}_k(\bar{d}^z) \stackrel{\text{def}}{=} (Z_k\bar{g}_k)^T\bar{d}^z + \frac{1}{2}(\bar{d}^z)^T B_k \bar{d}^z$$
$$\text{s.t.}\ \|\bar{d}^z\| \leqslant \Delta_k$$

where B_k is either M_k or its approximation. Let $\bar{d}_k^z$ denote the solution of the subproblem (S_k).

4. Set $\bar{d}_k = Z_k^T\bar{d}_k^z$, $d_k = D_k^{-1}\bar{d}_k$ and $f(x_{l(k)}) = \max\limits_{0 \leqslant j \leqslant m(k)}\{f(x_{k-j})\}$. Choose $\alpha_k = 1, \omega, \omega^2, \cdots$, until the following inequality is satisfied

$$f(x_k + \alpha_k d_k) \leqslant f(x_{l(k)}) + \alpha_k\beta g_k^T d_k, \quad \text{with } x_k + \alpha_k d_k \geqslant 0. \tag{2.15}$$

5. Set

$$h_k = \begin{cases} \alpha_k d_k, & \text{if } x_k + \alpha_k d_k > 0, \\ \phi_k \alpha_k d_k & \text{otherwise,} \end{cases} \tag{2.16}$$

$$x_{k+1} = x_k + h_k. \tag{2.17}$$

Here assume that for some constant $\phi_l \in (0,1)$, $\phi_k \in [\phi_l, 1)$, $\phi_k - 1 = O(\|d_k\|)$.

6. Calculate

$$\text{Pred}(h_k) = -\bar{\varphi}_k(Z_k D_k h_k), \tag{2.18}$$

$$\widehat{\text{Ared}}(h_k) = f(x_{l(k)}) - f(x_k + h_k), \tag{2.19}$$

$$\hat{\rho}_k = \frac{\widehat{\text{Ared}}(h_k)}{\text{Pred}(h_k)}, \tag{2.20}$$

and take

$$\Delta_{k+1} = \begin{cases} [\gamma_1 \Delta_k, \gamma_2 \Delta_k], & \text{if } \hat{\rho}_k \leqslant \eta_1, \\ (\gamma_2 \Delta_k, \Delta_k], & \text{if } \eta_1 < \hat{\rho}_k < \eta_2, \\ (\Delta_k, \min\{\gamma_3 \Delta_k, \Delta_{\max}\}], & \text{if } \hat{\rho}_k \geqslant \eta_2 \end{cases}$$

7. Take $m(k+1) = \min\{m(k)+1, M\}$, and update B_k to obtain B_{k+1}. Then set $k \leftarrow k+1$ and go to step 1.

Remark The scalar α_k given in step 4 denotes the step size along d_k to the boundary $x_k + \alpha_k d_k \geqslant 0$, i. e., $\alpha_k \overset{\text{def}}{=} \min\{\frac{x_{k,i}}{d_{k,i}} \mid d_{k,i} < 0, i = 1, \cdots, m\}$ and $\frac{x_{k,i}}{d_{k,i}} \overset{\text{def}}{=} +\infty$ if $d_{k,i} = 0$, where $x_{k,i}$ and $d_{k,i}$ are the ith components of vectors x_k and d_k, respectively. A key property of this scalar α_k is that an arbitrary step $\alpha_k d_k$ to the point $x_k + \alpha_k d_k$ does not violate any nonnegative constraints. Further, it is easy to see that the usual monotone algorithm can be viewed as a special case of the proposed algorithm when $M = 0$.

3 Global convergence

Throughout this section we assume that $f: \mathbf{R}^n \to \mathbf{R}^1$ is twice continuously differentiable and bounded from below. Given $x_0 \in \mathcal{F}$, the algorithm generates a sequence $\{x_k\} \subseteq \mathbf{R}^n$. In our analysis, the level set of f is denoted by $\mathcal{L}(x_0) = \{ x \in \mathbf{R}^n \mid f(x) \leqslant f(x_0), Ax = b, x \geqslant 0\}$.

The following assumption is commonly used in convergence analysis of most methods for linear equality constrained optimization.

Assumption 1 Sequence $\{x_k\}$ generated by the algorithm is contained in a compact set $\mathcal{L}(x_0)$ on $\mathbf{R}^n$. Matrix A has full row-rank m.

It is well known from solving the trust region algorithms that in order to assure the global convergence of the proposed algorithm, it is a sufficient condition to show that at the kth iteration the predicted reduction defined by $-\bar{\varphi}_k(\bar{d}_k^z)$ obtained by the step $\bar{d}_k^z$ from trust region subproblem, satisfies a sufficient descent condition (see [5]). Furthermore, we can also obtain that the direction of the trial step is a sufficiently descent direction (see [6]).

Lemma 3.1 Let the step $\bar{d}_k^z$ be the solution of the trust region subproblem. Then there exist a $\tau > 0$, and a $\tau_1 > 0$ such that the step $\bar{d}_k^z$ satisfies the following sufficient descent conditions:

$$-\bar{\varphi}_k(\bar{d}_k^z) \geqslant \tau \| Z_k \bar{g}_k \| \min\{\Delta_k, \frac{\| Z_k \bar{g}_k \|}{\| B_k \|}\}, \tag{3.1}$$

$$g_k^T d_k = \bar{g}_k^T \bar{d}_k = (Z_k \bar{g}_k)^T \bar{d}_k^z \leqslant -\tau_1 \| Z_k \bar{g}_k \| \min\{\Delta_k, \frac{\| Z_k \bar{g}_k \|}{\| B_k \|}\}, \tag{3.2}$$

for all $\bar{g}_k$, B_k and Δ_k. In fact $\tau = \frac{1}{2}$ and $\tau_1 = \frac{1}{4}$.

Assumption 2 B_k and $D^{-1}(x) \nabla^2 f(x) D^{-1}(x)$ are bounded, i. e., there exist a b, and a $\hat{b} > 0$ such that $\| B_k \| \leqslant b$, $\forall k$, and $\| D(x)^{-1} \nabla^2 f(x) D(x)^{-1} \| \leqslant \hat{b}$, $\forall x \in \mathcal{L}(x_0)$.

Similar to the proofs of Lemma 4.1 and Theorem 4.2 in [6], we can also obtain the following main result.

Lemma 3.2 Assume that Assumptions 1—2 hold. If there exists an $\varepsilon > 0$ such that

$$\| Z_k \bar{g}_k \| \geqslant \varepsilon \tag{3.3}$$

for all k, then there is an $\alpha > 0$ such that

$$\Delta_k \geqslant \alpha, \ \forall \ k. \tag{3.4}$$

Now we present only the following main result of the global convergence of the proposed algorithm whose proofs are omitted in the paper because of the limited space.

Theorem 3.4 Assume that Assumptions 1～2 hold. Let $\{x_k\} \subseteq \mathbf{R}^n$ be a sequence generated by the algorithm. Assume that the strict complementary of problem (1.1) at every limit point holds. Then

$$\liminf_{k\to\infty} \| Z_k \bar{g}_k \| = 0.$$

4 Local Convergent Rate

In order to get a stronger result and obtain the local convergent rate, we require more assumptions. However, because the paper is bounded, we present also only these main results of the proposed algorithm whose proofs are omitted.

Let the set of active constraints be denoted by

$$I(x) \stackrel{\text{def}}{=} \{i \mid x_i = 0, i = 1,\cdots,n\}, \tag{4.1}$$

which associates the optimization subproblem

$$(P)_I \qquad \min \ f(x); \quad \text{s.t.} \ Ax = b, \ x_I = 0. \tag{4.2}$$

Assumption 3 For all $I \subseteq \{1,\cdots,n\}$, the first order optimality system associated to $(P)_I$ has no nonisolated solutions and the strict complementary of problem (1.1) holds.

Assumption 4 The constraints of (1.1) are qualified in the sense that $(A^T\lambda)_i = 0, \ \forall i \notin I(\bar{x})$ implies that $\lambda = 0$.

Assuming that $(\bar{\lambda}, \bar{v})$ is associated with a unique pair $\bar{x}$ which satisfies Assumption 3. Define the set of strictly active constraints as

$$J(\bar{x}) \stackrel{\text{def}}{=} \{i \mid \bar{v}_i > 0, i = 1,\cdots,n\} \tag{4.3}$$

and the extended critical cone as

$$\mathcal{T}(\bar{x}) \stackrel{\text{def}}{=} \{d \in \mathbf{R}^n \mid Ad = 0, d_i = 0, i \in J(\bar{x})\}. \tag{4.4}$$

Assumption 5 The solution x_* of problem (1.1) satisfies the strong second order condition, that is, there exists an $\alpha > 0$ such that

$$p^T H_k p \geqslant \alpha \| p \|^2, \ \forall \ p \in \mathcal{T}(x_k) \tag{4.5}$$

where $H_k = \nabla^2 f(x_k)$. This is a sufficient condition for the strong regularity.

Assumption 6

$$\lim_{k\to\infty} \frac{\| (B_k - Z_k D_k^{-1} H_k D_k^{-1} Z_k^T) \bar{d}_k^z \|}{\| \bar{d}_k^z \|} = 0. \tag{4.6}$$

Theorem 4.1 Assume that Assumptions 2～6 hold. Let $\{x_k\}$ be a sequence generated by the algorithm. Then $d_k \to 0$. Furthermore, if x_k is close enough to x_*, and x_* is a strict local minimum of the problem (1.1), then $x_k \to x_*$.

Theorem 4.2 Assume that Assumptions 2～6 hold. Let $\{x_k\}$ be a sequence generated by the algorithm. Then

$$\lim_{k\to\infty} \| Z_k \bar{g}_k \| = 0. \tag{4.7}$$

Theorem 4.3 Assume that Assumptions 2～6 hold. Then for sufficiently large k, the step $\alpha_k \equiv$

1 and the trust region constraint is inactive, that is, there exists a $\hat{\Delta} > 0$ such that $\Delta_k \geqslant \Delta$, $\forall\ k \geqslant K'$, where K' is a large enough index.

Theorem 4.3 means that the local convergence rate for the proposed algorithm depends on the reduced Hessian of objective function at x_* and the local convergence rate of the step d_k. If d_k becomes the projected quasi-Newton step, then the sequence $\{x_k\}$ generated by the algorithm converges x_* is superlinear.

参考文献:

[1] COLEMAN T F, LI Y. An interior trust region approach for minimization subject to bounds[J]. SIAM J Optimization, 1996, 6(3):418-445.

[2] DENG N Y, XIAO Y, ZHOU F J. A nonmonotonic trust region algorithm[J]. Journal of Optimization Theory and Applications, 1993, 76:259-285.

[3] GRIPPO L, LAMPARIELLO F, LUCIDI S. A nonmonotonic line search technique for Newton's methods[J]. SIAM Journal on Numerical Analysis, 1986, 23:707-716.

[4] NOCEDAL J, YUAN Y. Combining trust region and line search techniques[J]. Y. Yuan Advances in Nonlinear Programming (Kluver, Dordvechat) 1998: 153-175.

[5] SORENSEN D C. Newton's method with a model trust region modification[J]. SIAM J Numer Anal, 1982, 19: 409-426.

[6] ZHU D. Curvilinear paths and trust region methods with nonmonotonic back tracking technique for unconstrained optimization[J]. J of Computational Mathematics, 2001, 19:241-258.

投影 Hessian 的内点回代算法解线性约束优化问题

朱德通

(上海师范大学 数理信息学院, 上海 200234)

摘要: 提供非单调内点回代技术的信赖域投影 Hessian 算法解线性约束优化问题. 基于矩阵 QR 分解的技巧,将仿射零空间的信赖域子问题变换成通常的信赖域子问题,然后结合线搜索技术,在每次迭代信赖域子问题都将产生新的回代内点. 在合理的条件下,证明了算法不仅具有整体收敛性而且保持局部超线性收敛速率,引入非单调技术将克服病态问题,加速收敛性进程.

关键词: 信赖域方法;回代法;非单调技术;内点法