第23卷第3期 上海师范大学学报(自然科学版) Vol. 23, No. 3
1994年9月 J. of Shanghai Teachers Univ. (Natural Sciences) Sep. , 1994

非线性等式约束分解信赖域算法(Ⅰ)

王政文 朱德通 蒋伟成
(数学系)

提 要 本文采用分解、合成的思想，求解非线性等式约束优化问题. 第一节，介绍了算法的发展；第二节，利用 Fletcher 罚函数，给出本文使用的两个算法：通常信赖域算法，非单调信赖域算法. 非单调信赖域算法是通常信赖域算法的推广，算法实践表明：非单调信赖域算法更具优越性，开始受到充分重视.

关键词 等式约束；信赖域；整体收敛；局部超线性收敛；非单调
中图法分类号 O241.7

考虑下列非线性等式约束的优化问题：

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c(x)=0 \end{aligned} \tag{1}$$

其中，$f: R^n \to R^1$，$c: R^n \to R^m$，$(m \le n)$ 都是两次连续可微的函数. 最近，许多学者提出了解决上述问题的方法. 其中，逐次二次规划方法(SQP)和止割方法(Secant method)最引人注目. 这些方法大体上分为两类：一类为完全 Hessioan 方法，即通过校正矩阵 B_k 来逼近问题(1)的 Lagrange 函数

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

关于 x 的 Hessian 阵

$$\nabla^2_{xx} L(x,\lambda)=\nabla^2 f(x)+\sum_{i=1}^{m}\lambda_i \nabla^2_{xx} c_i(x)$$

主要工作有 Fontecilla[14]，[15]，Tapia[25]，[26]. 另一类为投影 Hessioan 方法，即通过校正矩阵 B_k 逼近投影 Hessian 阵 $Z^T \nabla^2_{xx} L(x,\lambda) Z$，主要工作有 Byrd& Nocedal [7]，Coleman & Conn [10]，Nocedal& Overton [18]等.

Coleman& Conn [10]在第 k 次迭代时，须在 x_k 处构造零空间 $N[A_k^T]$ 的正交基 Z_k，并且考虑如下子问题：

$$\min \quad g_k^T d+\frac{1}{2} d^T Z_k B_k Z_k^T d,$$

本文于 1992年6月19日收到

$$\text{s.t.} \quad A_k^T d + C_k = 0$$

在合理的条件下，上述子问题有唯一解

$$d_k = h_k + v_k,$$

其中，

$$h_k = -Z_k B_k^{-1} Z_k^T g_k,$$
$$v_k = -A_k (A_k^T A_k)^{-1} C_k$$

Coleman & Conn 算法，实质上将迭代步长 d_k 分解为两个部分：h_k 和 v_k，h_k 属于 A_k^T 的零空间，即 $h_k \in N[A_k^T]$ 且 v_k 属于 A_k 的值空间 $R[A_k]$. 遗憾的是：最近的研究表明，由一般的正交分解得到的这种正交基 Z_k 通常是不连续的，是 Byrd & Schnabel [5].

Fontecilla [14]算法是 Coleman & Conn 算法的推广. 在算法中，为克服 Z_k 的不连续性，使用了连续的正交投影算子

$$P_k = I - A_k (A_k^T A_k)^{-1} A_k^T$$

上述这些方法，都只讨论算法的局部收敛性，并未涉及整体行为. 解决算法的整体收敛性. 主要有两种途径：线搜索技术和信赖域策略.

信赖域型的 SQP 方法，要在第 k 次迭代时，求解

$$\begin{aligned} \min \quad & g_k^T d + \frac{1}{2} d^T B_k d, \\ \text{s.t.} \quad & A_k^T d + c_k = 0 \\ & \| d \| \le \Delta_k, \end{aligned} \tag{2}$$

其中，Δ_k 为信赖域半径. 但是，问题(2)当 $C_k \neq 0$ 时，有可能使线性约束 $A_k^T d + c_k = 0$ 与信赖域约束 $\| d \| \le \Delta_k$ 的可行域互不相交，不为克服这种情形，主要有两种处理方法：一种方法是使用松驰因子，即考虑下列子问题：

$$\begin{aligned} \min \quad & g_k^T d + \frac{1}{2} d^T B_k d, \\ \text{s.t.} \quad & A_k^T d + \alpha C_k = 0, \\ & \| d \| \le \Delta_k, \end{aligned}$$

这里，$0 \le \alpha \le 1$. 这一方法首先由 Vardi(1985)引入. 但是，经过上述简单的松驰之后，Byrd 发现，当 x_k 接近 x_* 时，约束松驰后得到的超平面几乎平行于非线性约束的糟面，从而，信赖域约束迫使选择 d_k，使 $x_k + d_k$ 比 x_k 更不可行，于是，步长 d_k 将不被接受，破坏了收敛效率.

第二种方法是在范数意义下松驰线性约束. 如 Celis，Dennis & Tapia (1984)使用下列子问题：

$$\begin{aligned} \min \quad & g_k^T d + \frac{1}{2} d^T B_k d, \\ \text{s.t.} \quad & \| A_k^T d + c_k \| \le Q_k \\ & \| d \| \le \Delta_k \end{aligned} \tag{3}$$

其中，$Q_k = \| A_k^T S_k^{cp} + c_k \|$，且 S_k^{cp} 是如下问题的解：

$$\begin{aligned} \min \quad & \| A_k^T s + c_k \| \\ \text{s.t.} \quad & \| s \| \le \Delta_k \end{aligned} \tag{4}$$

Powell& Yuan (1986)也求解(3)，但取 $Q_k = \| A_k^T S_k^* + c_k \|$，$s_k^*$ 为下列问题的解：

$$\min \quad \| A_k^T s + c_k \|^2,$$
$$\text{s.t.} \quad \| s \| \le \sigma \Delta_k, 0 < \sigma < 1$$

本文采用正割方法中的分解技术，将步长 S_k 分解成 d_k 和 h_k 的和，然后，在两个不同的子空间求解两个信赖域子问题，将分解技术和信赖域策略相结合，解决了求解问题(1)的算法的整体收敛性和局部超线性收敛速率.

本文采用的两个子问题为：

(D_k)
$$\min \quad g_k^T d + \frac{1}{2} d^T B_k d,$$
$$\text{s.t.} \quad A_k^T d = 0$$
$$\| d \| \le \Delta_k$$

(E_k)
$$\min \quad \| c_k + A_k^T A_k w \|^2$$
$$\text{s.t.} \quad \| w \| \le \Delta_k$$

设 (D_k) 和 (E_k) 最优解为 d_k 和 w_k，令

$$h_k = A_k w_k$$
$$s_k = d_k + h_k$$

和(3)，(4)相比，(D_k)，(E_k) 更易于求解.

无约束优化的非单调线搜索技术由 Grippo，et al 在[31]，[32]中提出，而这一思想，最早可以追溯到 ChamgerLain，Powell，Lemarechal& Pedersen 的 Watchdog 技术，它使传统的线搜索条件得以松驰，从而无需目标函数值每次迭代都单调下降. Deng，Xiao & Zhou [33]将非单调思想引入到信赖域中，对无约束问题，给出了非单调信赖域算法.数值结果表明，无论使用线搜索技术还是信赖域策略，非单调方法都只需较少的迭代次数和函数赋值，从而比通常的单调算法更具优越性.而上述文章，考虑的都是无约束问题，在本文的最后部分，我们将非单调技术应用到带有等式约束的优化问题中，得到了非单调算法.值得指出的是，非单调算法，以其自身的优点，必将越来越受到人们的重视.

算法

设

$$c(x) = (c_1(x), \ldots, c_m(x))^T \in R^m,$$
$$A(x) = [\nabla c_1(x), \ldots, \nabla c_m(x)] \in R^{n \times m},$$

并假定 $A(x)$ 是列满秩的，即 $\operatorname{rank}(A(x)) = m$. 定义投影矩阵为

$$p(x) = I - A(x)[A(x)^T A(x)]^{-1} A(x)^T \tag{5}$$

以及问题(1)的 Lagrange 函数为

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

其中，λ 采用投影乘子，即

$$\lambda(x) = -[A(x)^T A(x)]^{-1} A(x)^T g(x) \tag{6}$$

定义

$$W(x, \lambda) = \nabla_{xx}^2 L(x, \lambda)$$

是 $L(x, \lambda)$ 关于 x 的 Hessian 矩阵. 令

$$Q(x) = A(x)[A(x)^T A(x)]^{-1} A(x)^T \tag{7}$$

对 $A(x)$ 作 QR 分解

$$A(x)=[y(x),z(x)]\begin{bmatrix}R(x)\\0\end{bmatrix}=Y(x)R(x) \tag{8}$$

其中，$Y(x)$,$Z(x)$ 分别是值空间 $R[A(x)]$ 和零空间 $N[A(x)^T]$ 的标准正交基，$R(x)$ 是秩为 m、非奇异的上三角阵.

故

$$Y(x)Y(x)^T+Z(x)Z(x)^T=I_n, \tag{9}$$

$$A(x)^TZ(x)=0 \tag{10}$$

为书写方便，用 c_k 表示 $c(x_k)$，下同.

求得 (D_k) 和 (E_k) 的最优解 d_k 和 w_k 后，为确定是否接受步长 s_k，即是否令 $x_{k+1}=x_k+s_k$，引入 Fletcher 罚函数

$$\varphi(x,p)=f(x)+\lambda^Tc(x)+\rho\|c(x)\|^2 \tag{11}$$

其中，λ 是(6)确定，ρ 是大于 0 的罚参数.

定义罚函数 $\varphi(x,p)$ 从 x_k 到 x_{k+1} 的实际改变量

$$\mathrm{Ared}_k(x_k)=\varphi(x_k,\rho_{k+1}^0)-\varphi(x_{k+1},\rho_{k+1}) \tag{12}$$

以及预计改变量

$$\begin{aligned}\mathrm{Pred}_k(x_k)=&-g_k^Td_k-\frac{1}{2}d_k^TB_kd_k+\frac{1}{2}[\lambda(x_k+d_k)-\lambda_k]^TA_k^Ts_k\\&-[\lambda(x_k+s_k)-\lambda_k]^T(c_k+\frac{1}{2}A_k^Ts_k)\\&+\rho_k[\|c_k\|^2-\|c_k+A_k^Ts_k\|^2]\end{aligned} \tag{13}$$

在给出算法之前，再引入如下记号：

$$L'_k=\begin{cases}\dfrac{\|\lambda(x_k+d_k)-\lambda_k\|}{\|d_k\|}, & 若\ \|d_k\|\neq 0\\ 0, & 否则.\end{cases} \tag{14}$$

$$L''_k=\begin{cases}\dfrac{\|\lambda(x_k+s_k)-\lambda_k\|}{\|s_k\|}, & 若\ \|s_k\|\neq 0\\ 0, & 否则.\end{cases} \tag{15}$$

β 是$(0,\frac{1}{2})$中任一常数，即 $\beta\in(0,\frac{1}{2})$.

$$\sigma=(\frac{1}{2}-\beta)/2 \tag{16}$$

$$\tau_k=\max\left\{\frac{L'_k}{\sigma},\frac{L''_k}{\sigma},\frac{1}{2\sigma}\right\} \tag{17}$$

$$L_k=L'_k+2L''_k, \tag{18}$$

$$l_k=L'_k+2L''_k(1+\|A_k\|), \tag{19}$$

$$\pi_k=(\beta\tau_k+l_k)\|(A_k^TA_k)^{-1}\| \tag{20}$$

$$\rho_k=\begin{cases}\rho_{k-1}, 若 & \rho_{k-1}\geqslant\pi_k,\\ \max\{(\rho_{k-1}+\rho_0,\pi_k\}, & 否则,\end{cases} \tag{21}$$

其中，ρ_0 是任一给定正常数.

通常信赖域算法(UTR)：

给定 $\mu \in (0,1)$，$\eta \in (\mu,1)$，$\nu > 1$，以及 $\gamma_0,\gamma_1,\gamma_2$ 和 γ_3 满足

$$0 < \gamma_0 \leq \gamma_1 < 1 < \gamma_2 \leq \gamma_3 \tag{22}$$

步 0　给定初始点 x_0, f_0 h 和 g_0，以及初始信赖域半径 Δ_0 和 β_0，置 $k=0, j=1, \varepsilon=\varepsilon_1>0$.

步 1　如果 $\|c_k\| + \|P_k g_k\| = 0$，停. 否则，令

$$\varepsilon \leftarrow \varepsilon_{j+1} = \min\left\{\frac{\|c_k\| + \|\rho_k g_k\|}{\gamma}, \varepsilon_j\right\}, \qquad j \leftarrow j+1$$

步 2　求解问题 (D_k) 和 (E_k)，得 d_k 和 w_k，置

$$h_k = A_k w_k, s_k = d_k + h_k$$

步 3　如果 $d_k = 0, c_k = 0$，停.

步 4　计算 $\mathrm{Pred}_k(x_k)$ 在 $\rho = \rho_k$ 处的值.如果

$$\mathrm{Pred}_k(x) \geqslant \beta\varepsilon\Delta_k \tag{23}$$

成立，转步 5. 否则，令 $\Delta_k \leftarrow \Delta_k \nu$，转步 2.

步 5　计算 $\mathrm{Ared}_k(x_k)$ 以及

$$\eta_k = \frac{\mathrm{Ared}_k(x_k)}{\mathrm{Pred}_k(x_k)} \tag{24}$$

如果

$$\eta_k \geqslant \eta, \tag{25}$$

令

$$x_{k+1} = x_k + s_k, \tag{26}$$

$$\Delta_{k+1} \in [\gamma_2\Delta_k, \gamma_3\Delta_k] \tag{27}$$

转步 6. 否则，令

$$x_{k+1} = x_k, \tag{28}$$

$$\Delta_{k+1} \in [\gamma_0\Delta_k, \gamma_1\Delta_k] \tag{29}$$

转步 2.

步 6　由 DFP 或 BFGS 校正公式校正矩阵 B_k，置 $k \leftarrow k+1$，转步 1.

如果 η_k 满足(25)，则称第 k 次迭代为成功迭代，否则称为不成功迭代，记 S 表示成功迭代的指标集.

设

$$\varphi(x_{e(k)}, \rho_{e(k)}) = \max(,0 \leq j \leq m(k))\{\varphi(x_{k-j}, \rho_{k-j}\} \tag{30}$$

$$M\mathrm{Ared}_k(x_k) = \varphi(x_{e(k)}, \rho_{e(k)}) - \varphi(x_k + s_k, \rho_k) \tag{31}$$

其中，$m(0)=0, 0 \leq m(k) = \min\{m(k-1)+1, M\}, k \geqslant 1$，且 M 是非负整数，$l(k)$ 满足

$$k - m(k) \leq l(k) \leq k \tag{32}$$

非单调信赖域算法(NTR)：

给定 $\mu \in (0,1)$，$\eta \in (\nu,1)$，$\gamma > 1$，$\gamma_0,\gamma_1,\gamma_2$ 和 γ_3,γ_4 满足

$$0 < \gamma_0 \leq \gamma_1 \leq \gamma_2 < 1 < \gamma_3 \leq \gamma_4 \tag{33}$$

步 0　给定初始点 x_0, f_0 和 g_0 以及初始信赖域半径 Δ_0 和 β_0，令 $k=0, j=1, \varepsilon=\varepsilon_1>0$.

步 1　若 $\|c_k\| + \|P_k g_k\| = 0$，停. 否则，令

$$\varepsilon \leftarrow \varepsilon_{j+1} = \min\left\{\frac{\|c_k\| + \|\rho_k g_k\|}{\gamma}, \varepsilon_j\right\}, j \leftarrow j+1. \tag{34}$$

步2 求解子问题 (D_k) 和 (E_k) 得 d_k 和 w_k，令

$$h_k = A_k w_k, s_k = d_k + h_k \tag{35}$$

步3 若 $d_k = 0, c_k = 0$，停.

步4 计算 $\mathrm{Pred}_k(x_k)$ 当 $\rho = \rho_k$ 的值，若

$$\mathrm{Pred}_k(x_k) \geqslant \beta\varepsilon\Delta_k \tag{36}$$

成立，转步5. 否则，令 $\Delta_k \leftarrow \Delta_{k/\gamma}$，转步2.

步5 计算 $\mathrm{Ared}_k(x_k)$ 和 $M\mathrm{Ared}_k(x_k)$，并令

$$\eta_k = \frac{\mathrm{Ared}_k(x_k)}{\mathrm{Pred}_k(x_k)}, \tag{37}$$

$$\mu_k = \frac{M\mathrm{Ared}_k(x_k)}{\mathrm{Pred}_k(x_k)} \tag{38}$$

步6 若

$$\mu_k \geqslant \mu, \tag{39}$$

令

$$x_{k+1} = x_k + s_k, \tag{40}$$

$$\Delta_{k+1} \in [\gamma_3\Delta_k, \gamma_4\Delta_k], \quad 若\ \eta_k \geqslant \eta_0 \tag{41}$$

$$\Delta_{k+1} \in [\gamma_1\Delta_k, \gamma_2\Delta_k], \quad 若\ \eta_k < \eta_0 \tag{42}$$

转步7. 否则，令

$$x_{k+1} = x_k, \tag{43}$$

$$\Delta_{k+1} \in [\gamma_0\Delta_k, \gamma_1\Delta_k] \tag{44}$$

转步2.

步7 由 DFP 或 BFGS 校正公式校正 B_k，令 $k \leftarrow k+1, m(k) = \min\{m(k-1)+1, M\}$，转步1.

同样，称使(39)成立的迭代为成功迭代.

和通常信赖域算法相比，当 $M > 0$ 时的非单调信赖域算法 NTR，只要求 s_k 使 $\varphi(x_k + s_k, \rho_k)$ 相对于 $\varphi(x_{l(k)}, \rho_{l(k)})$ 有一定的下降量，因而，序列 $\{\varphi(x_{k+1}, \rho_k)\}$ 非单调下降. 同时，通常信赖域算法仅是非单调信赖域算法当 $M = 0$ 时的特殊情形.

参 考 文 献

[1] Boggs, Tolle & Wang, *SLAM J. Cont. Opti.*, 1982, 20

[2] Boggs, Tolle, *SLAM J. Num. Anal.*, 1984, 21(6)

[3] Boggs, Tolle, *SLAM J. Num. Anal.*, 1989, 26(3)

[4] Byrd, Math. Prog., 1985, 32

[5] Byrd, Schnabel, Math. Prog., 1986

[6] Byrd, Schnabel & Shultz, Math. Prog., 1990

[7] Byrd, Nocedal, Math. Prog., 1990

[8] Byrd, *SLAM J. Num. Anal.*, 1990, 27(1)

[9] Celis,Dennis & Tapia,In Numerical Optimization 1984,Philadelphia
[10] Coleman & Conn, *SLAM J. Num. Anal.*, 1984,21(4)
[11] Coleman & Sorensen,Math. Prog., 1984,**29**
[12] El-ALEM, *SLAM J. Num. Anal.*, 1991,28(1)
[13] Fontecilla,Steihaug & Tapia, *SLAM J. Num. Anal.*, 1987,**24**(3)
[14] Fontecilla, *SLAM J. Num. Anal.*, 1988,25(2)
[15] Fontecilla, *SLAM J. Num. Anal.*, 1990,27(1)
[16] Goodman,Math. Prog., 1985,**33**
[17] Lucidi,JOTA., 1990,67(2)
[18] Nocedal & Overton, *SLAM J. Num. Anal.*, 1985,22(5)
[19] Omojokun,University of Colorado,PHD thesis,1989
[20] Powell,Math. Prog., 1984,28
[21] Powell & Yuan,Report DAMTP 1986/NA2,Cambridge,England
[22] Powell,Math. Prog., 1986,**29**
[23] Powell & Yuan,Math. Prog., 1991,**49**
[24] Powell,Nonlinear Programming 3,Academic Press,New York
[25] Tapia, *J. Opti. Theor. Appl.*, 1977,**22**
[26] Tapia, *Math of Comp.*, 1988,**51**(183)
[27] Vardi, *SALM J. Num. Anal.*, 1985,22(3)
[28] Yuan. Math. Prog., 1985,**32**
[29] Zhang & Zhu, *JOTA*, 1990,**67**(2)
[30] Deng,Xiao & Zhou,to appear
[31] Grippo,Lampariello & Lucidi, *SLAM J. Num. Anal.*, 1986,**23**(4)
[32] Grippo,Lampariello & Lucidi, *JOTA*, 1989,**60**(3)

Trust Region Algorithm with Decomposition Technique for Nonlinear Equality Constrained Problem (I)

Wang Zhengwen Zhu Detong Jiang Weizheng

(Department of Mathematics)

Abstract

The nonlinear equality constrained optimization problem with a decomposition technique is discussed. The development of the algorithms is described in the first part. With the help of Fletcher's penalty function, we propose two algorithms: the usual trust region algorithm and the nonmonotone trust region algorithm in Part 2. Nonmonotone trust region algorithm is the extension of the usual trust region algorithm, and the behaviour of the algorithm indicates that the nonmonotone trust region algorithm has more advantages and therefore, is being paid special attention.

Keywords equality constrained optimization; trust region; global convergence; locally superlinear convergence; nonmontone