中国农业大学学报　1997,2(6):17～22
Journal of China Agricultural University

不精确线搜索条件下锥模型 Broyden 凸族算法的收敛速率①

李正锋②　陈　静　邓乃扬
(中国农业大学工程基础科学部)

摘　要　在一定假设条件下的锥模型Broyden凸族算法的局部收敛性和全局收敛性已有人研究过。本文中进一步研究不精确线搜索条件下锥模型Broyden凸族算法的收敛速率。证明了如果初值 x_1 充分接近强局部极小点 x_*，那么族中任一算法所产生的点列都是R-局部收敛的，且其R-收敛阶至少是 $\tau\geqslant\sqrt[n]{2}$，而不需要假设 A_1 充分靠近海色阵 $\nabla^2 f(x_*)$。

关键词　无约束极小化；Broyden凸族；锥模型；局部收敛
中图分类号　O221.2

Convergence Rate of Conic Broyden Convex Family Under Imperfect Line Search

Li Zhengfeng　Chen Jing　Deng Naiyang
(College of Applied Engineering Sciences, CAU)

Abstract　After the study on local and global convergence of conic Broyden convex family under certain conditions, the convergence rate of it under the imperfect line search is extended. It is proved that all the members in the family are convergent locally with R-order $\tau\geqslant\sqrt[n]{2}$ at least without the hypothesis of A_1 close to $\nabla^2 f(x_*)$ if the starting point x_1 is close to x_* enough.

Key words　unconstrained minimization; Broyden convex family; conic model; local convergency

对于无约束极小化问题，Davidon[1]提出了一种基于目标函数的锥近似算法。此类算法可以看作是Broyden凸族算法的一种扩展。在假定初始值 x_1 和 A_1 分别充分接近极小值点 x_* 及其海色阵 $\nabla^2 f(x_*)$ 时，锥模型Broyden凸族算法的局部收敛性已被Sorensen[2]以及Ariyawansa和Lau[3]讨论过；其全局收敛性可见文献[4]。笔者进一步研究不精确线搜索条件下锥模型Broyden凸族算法的收敛速率，获得了新的结果：证明了如果初值 x_1 充分接近强局部极小点 x_*，那么族中任一算法所产生的点列都是R-局部收敛的，且其R-收敛阶至少是 $\tau\geqslant\sqrt[n]{2}$，而不需要假设 A_1 充分靠近海色阵 $\nabla^2 f(x_*)$。

收稿日期：1997-08-06
①国家自然科学基金资助项目
②李正锋，北京清华东路17号 中国农业大学(东校区)71信箱，100083

1 锥模型 Broyden 凸族算法

考虑求解下列无约束优化问题：

$$\min_{x\in\mathbf{R}^n} f(x) \tag{1}$$

其中 $f(x)\in C^2$。

锥模型 Broyden 凸族算法是一种迭代算法。在每一迭代步，构造一个锥函数来近似目标函数，使其与目标函数在当前迭代点和前一迭代点处有相同的函数值及梯度值，然后基于此锥函数的极小点确定下一步的搜索方向。与以往 Broyden 凸族算法不同，锥模型 Broyden 凸族算法生成的搜索方向不仅与目标函数的梯度值有关，而且还依赖于目标函数的函数值，因此可望能提高计算效率。

下面给出锥模型 Broyden 凸族算法，其详细导出过程见文献[4]。

算法 1

1)选取初始点 $x_1\in\mathbf{R}^n$ 和一对称正定矩阵 A_1，令 $k=1$。

2)计算 $g_k=\nabla f(x_k)$，如果 $g_k=0$，则停，否则转 3)。

3)令 $s_k=-A_k^{-1}g_k$。

4)沿 s_k 方向作一维线搜索，确定步长 $\lambda_k>0$ 使之满足

$$\lambda_k=\min\{\lambda>0,g(x_k+\lambda_k s_k)^{\mathrm{T}}s_k=\mu_k g_k^{\mathrm{T}}s_k\} \tag{2}$$

5)令 $x_{k+1}=x_k+\lambda_k s_k$。

6)计算 $g_{k+1}=\nabla f(x_{k+1})$，如果 $g_{k+1}=0$，则停，否则计算以下各量：

$$p_k=x_{k+1}-x_k$$

$$\rho_k=+\sqrt{[f(x_k)-f(x_{k+1})]^2-(p_k^{\mathrm{T}}g_{k+1})(p_k^{\mathrm{T}}g_k)}$$

$$\beta_k=\frac{f(x_k)-f(x_{k+1})+\rho_k}{-g_k^{\mathrm{T}}p_k}$$

$$q_k=\beta_k g_{k+1}-\beta_k^3 g_k$$

$$A_{k+1}=A_k-\frac{A_k p_k p_k^{\mathrm{T}}A_k}{p_k^{\mathrm{T}}A_k p_k}+\frac{q_k q_k^{\mathrm{T}}}{q_k^{\mathrm{T}}p_k}+\psi_k v_k v_k^{\mathrm{T}}$$

其中 $\psi_k\in[0,1]$，且

$$v_k=\sqrt{p_k^{\mathrm{T}}A_k p_k}\left(\frac{q_k}{q_k^{\mathrm{T}}p_k}-\frac{A_k p_k}{p_k^{\mathrm{T}}A_k p_k}\right)$$

7)令 $k=k+1$，转步 3)。

2 锥模型 Broyden 凸族算法的局部收敛性

现在研究锥模型 Broyden 凸族算法 1 的局部收敛性。其中要求式(2)中的参数 μ_k 满足

$$|\mu_k|\leqslant\bar{\mu}<1 \tag{3}$$

易知，大多数线搜索准则均满足式(3)。

与文献[5]一样，需作如下定义和假定。

定义 A 给定 $f:\mathbf{R}^n\to\mathbf{R}$，$f\in C^2(D)$，其中 $D\subseteq\mathbf{R}^n$ 是一个开凸集。如果存在 $x_*\in D$，使得

$\nabla f(x_*)=0$ 并且海色阵 $G_*=\nabla^2 f(x_*)$ 是正定的，则称 $x_*\in D$ 为 f 的强局部极小点。

假定 B　设 $f\in C^2(D)$，其中 $D\subseteq \mathbf{R}^n$ 是一开凸集，点 $x_*\in D$ 是 f 的一个强局部极小点，而且存在 x_* 的一个邻域 $N\subseteq D$ 和常数 $L>0$，使得对所有 $x,y\in N$ 有

$$\|\nabla^2 f(x)-\nabla^2 f(y)\|\leqslant L\|x-y\| \tag{4}$$

和

$$\|\nabla f(x)-\nabla f(y)\|\leqslant L\|x-y\| \tag{5}$$

成立，其中 $\|\cdot\|$ 表示向量或矩阵的欧几里德范数。易知，由假设 B，存在 x_* 的 1 个邻域 N 和 2 个正常数 α_1 和 α_2，使对于所有的 $x\in N$，下列不等式成立：

$$\alpha_1\|x-x_*\|^2\leqslant f(x)-f(x_*)\leqslant \alpha_2\|x-x_*\|^2 \tag{6}$$

和

$$\alpha_1\|g(x)\|^2\leqslant f(x)-f(x_*)\leqslant \alpha_2\|g(x)\|^2 \tag{7}$$

其中 $g(x)=\nabla f(x)$。

下面 2 个引理将确保存在 x_* 的一个邻域 N，使得当起始点 $x_1\in N$ 时，算法 1 是适定的。为书写简便，在无混淆情况下，略去下标 k，并用“+”代替 $k+1$。

引理 2.1　设 $f:\mathbf{R}^n\to\mathbf{R}$ 满足假定 B，则存在强局部极小点 x_* 的一个邻域 N，使得 1)对所有 $x\in N$，$\nabla^2 f(x)$ 是正定的，并且式(4)和(5)成立；2)对所有 $x,x_+\in N$，$x\neq x_+$，有

$$(f-f_+)^2-(g_+^{\mathrm T}p)(g^{\mathrm T}p)\geqslant \alpha_3\|p\|^4 \tag{8}$$

和

$$|1-\beta|/\|p\|\leqslant \alpha_4 \tag{9}$$

其中 $\beta=\beta(x_+,x)=(f-f_++\rho)/(-g^{\mathrm T}p)$，$p=x_+-x$，$f=f(x)$，$f_+=f(x_+)$，$g=\nabla f(x)$，$g_+=\nabla f(x_+)$，$\rho=\sqrt{(f-f_+)^2-(g_+^{\mathrm T}p)(g^{\mathrm T}p)}$，$\alpha_3>0$ 和 $\alpha_4>0$ 是常数。证明见文献[2]。

为避免常数太多，下面引入约当符号 $O(\cdot)$，例如 $a=O(\|\varepsilon\|)$ 表示存在一个仅依赖于式(4)～(9)中常数 $L,\alpha_1,\alpha_2,\alpha_3$ 和 α_4 的常数 γ，使得 $\|a\|\leqslant\gamma\|\varepsilon\|$。其中 a 是一个向量或矩阵，$\|\varepsilon\|$ 为一充分小的量。现在考虑算法 1 的一个代表步，即从 x 和 A 分别导出下一步 x_+ 和 A_+，其中 $x\in N$，A 是一正定矩阵。

对任何 $x\in N$，设 $\varepsilon=x-x_*$，$p=\lambda s=x_+-x$。由于只讨论局部收敛性，为不失一般性，可假定 ε 充分小。首先，给出 $f(x)-f(x_+)$ 的一个相当精确的估计。事实上，容易看出

$$\begin{aligned}f(x)-f(x_+)=&-\int_0^\lambda g(x+\tau s)^{\mathrm T}s\mathrm{d}\tau=\\&-\lambda g^{\mathrm T}s-\int_0^\lambda\left[\int_0^\tau s^{\mathrm T}\nabla^2 f(x+\eta s)s\mathrm{d}\eta\right]\mathrm{d}\tau=\\&-\lambda g^{\mathrm T}s-\frac{1}{2}\lambda^2 s^{\mathrm T}G_*s[1+O(\|\varepsilon\|)]\end{aligned} \tag{10}$$

式中用到了 $p=O(\|\varepsilon\|)$。另一方面，由式(2)知

$$(\mu-1)g^{\mathrm T}p=g(x+\lambda s)^{\mathrm T}p-g^{\mathrm T}p=\int_0^\lambda p^{\mathrm T}\nabla^2 f(x+\tau s)s\mathrm{d}\tau$$

由平均值定理和假定 B，有

$$\lambda=(1-\mu)\frac{-g^{\mathrm T}s}{s^{\mathrm T}\nabla^2 f(x+\tau s)s}$$

其中 τ 满足 $0<\tau<\lambda$。因此

$$\lambda=(1-\mu)\frac{-g^{\mathrm{T}}s}{s^{\mathrm{T}}G_{*}s}[1+O(\|\varepsilon\|)]$$

从而由式(10)得出，对所有 $|\mu|<1$，有

$$f(x)-f(x_{+})=\frac{1}{2}(1-\mu^{2})\frac{(g^{\mathrm{T}}s)^{2}}{s^{\mathrm{T}}G_{*}s}[1+O(\|\varepsilon\|)]$$

由于 $g^{\mathrm{T}}s=g^{\mathrm{T}}A^{-1}g\geqslant\|g\|^{2}/\|A\|$，有 $s^{\mathrm{T}}G_{*}s=g^{\mathrm{T}}A^{-1}G_{*}A^{-1}g\leqslant M\|A^{-1}\|^{2}\|g\|^{2}$，其中 $M=\|G_{*}\|$。注意到对于小的 $\|\varepsilon\|$，应用式(7)，可知存在常数 ξ，使得对于小的 $\|\varepsilon\|$，有

$$f(x)-f(x_{+})\geqslant\frac{\xi(1-\mu^{2})}{[\mathrm{Cond}(A)]^{2}}[f(x)-f(x_{*})]$$

其中 $\mathrm{Cond}(A)=\|A\|\cdot\|A^{-1}\|$，因此有：

引理 2.2 假设假定 B 成立，考虑算法 1，其中 $A_{1}\in\mathbf{R}^{n\times n}$ 正定及所有 k，$\theta_{k}\in[0,1]$，$|\mu_{k}|\in[0,\bar{\mu})$，$\bar{\mu}<1$，则存在常数 $\delta>0$ 和 α'，使得如果 $\|x_{1}-x_{*}\|<\delta$，则由算法 1 生成的序列 $\{x_{k}\}$ 或者有限，即对一些 m，$x_{m}=x_{*}$，或者满足对所有 $k\geqslant1$，有

$$f(x_{k+1})-f(x_{*})\leqslant\left(1-\frac{\alpha'}{[\mathrm{Cond}(A_{k})]^{2}}\right)[f(x_{k})-f(x_{*})]$$

显然，如果 $\mathrm{Cond}(A_{k})$ 一致有界，那么引理 2.2 表明 $f(x_{k})-f(x_{*})$ 至少 Q-线性收敛于 0。再由式(6)可知，则 $x-x_{*}$ 也至少 Q-线性收敛于 0。为了证明对于充分小的 $\|\varepsilon_{1}\|=\|x_{1}-x_{*}\|$，$\mathrm{Cond}(A_{k})$ 的有界性，这里应用 Fletcher[6] 的结果。

引理 2.3[6] 设 H 是一个正定矩阵，p 和 q 为向量，使得存在另一个正定矩阵 B 有 $H=B^{2}$，$Hq=p(p\neq0)$，则对所有 $\theta\in[0,1]$，由算法 1 生成的正定矩阵 A 和 A_{+} 具有下列性质：

1) $1\leqslant\|BA_{+}B\|\leqslant\max\{1,\|BAB\|\}$；

2) $1\leqslant\|(BA_{+}B)^{-1}\|\leqslant\max\{1,\|(BAB)^{-1}\|\}$。

现在应用这个引理来寻求关于 $K_{+}(A_{+})=\max\{1,\|G_{*}^{-1/2}A_{+}G_{*}^{-1/2}\|\}$，$K_{-}(A_{+})=\max\{1,\|G_{*}^{1/2}A_{+}^{-1}G_{*}^{1/2}\|\}$ 和 $C(A_{+})=K_{+}(A_{+})K_{-}(A_{+})$ 的界。由下面不等式给出 $\mathrm{Cond}(A_{+})$ 的界：

$$\mathrm{Cond}(A_{+})\leqslant\alpha_{5}C(A_{+})\tag{11}$$

式中 $\alpha_{5}=\max\{1,\|G_{*}\|\cdot\|G^{-1}\|\}=\max\{1,\mathrm{Cond}(G_{*})\}$。

引理 2.4 假设假定 B 成立。考虑算法 1，其中 A_{1} 正定且对所有 k，$\theta_{k}\in[0,1]$，$|\mu_{k}|\in[0,\bar{\mu})$，$\bar{\mu}<1$，则当 x_{1} 充分接近 x_{*} 时，算法 1 产生的任何一对正定矩阵 $A=A_{k}$，$A_{+}=A_{k+1}$ 满足估计式

$$\left.\begin{aligned}&K_{+}(A_{+})\leqslant K_{+}(A)[1+O(\|\varepsilon\|)]\\&K_{-}(A_{+})\leqslant K_{-}(A^{-1})[1+O(\|\varepsilon\|)]\\&C(A_{+})\leqslant C(A)[1+O(\|\varepsilon\|)]\end{aligned}\right\}\tag{12}$$

其中 $\varepsilon=\varepsilon_{k}=x_{k}-x_{*}$。

证明 现证明对于小的 $\|p\|\neq0$，存在正定矩阵 F，有

$$(I+F)G_{*}^{-1}(I+F)q=p,F=O(\|\varepsilon\|)\tag{13}$$

其中 $p=x_{+}-x$，并且 $q=\beta g_{+}-\beta^{3}g$。事实上，对于小的 $\|p\|\neq0$，存在正交阵 U，使得 $Up=(\|p\|/\sqrt{n})(1,1,\cdots,1)^{\mathrm{T}}$。由式(4)和(9)，并注意到 $p=O(\|\varepsilon\|)$，有

$$q=\beta g_{+}-\beta^{3}g=[1+O(\|p\|)]g_{+}-[1+O(\|p\|^{3})]g=[I+O(\|\varepsilon\|)]G_{*}p$$

$$UG_{*}^{-1}q=\frac{\|p\|}{\sqrt{n}}(1,1,\cdots,1)^{\mathrm{T}}$$

因此，对于小的$\|\varepsilon\|\neq0$，存在对角阵 D，有 $DUG_{*}^{-1}U^{\mathrm{T}}DUq=Up$，$\bar{F}=D-I=O(\|p\|)$，因此矩阵 $F=U^{\mathrm{T}}\bar{F}U$ 满足式(13)。

在引理 2.3 中，令 $B=(I+F)G_{*}^{-1/2}$即得所要的结论。

作为引理 2.2 和引理 2.4 的一个简单结果，有如下引理：

引理 2.5　假设假定 B 成立。考虑算法 1，其中 A_1 正定，且对所有 k，$\theta_k\in[0,1]$，$|\mu_k|\in[0,\bar{\mu})$，$\bar{\mu}<1$，则存在常数 $\delta>0$，$\alpha>0$，$\gamma>0$，使得如果$\|x_1-x_{*}\|<\delta$，则由算法 1 生成的序列$\{x_k\}$或者有限，即对一些 m，$x_m=x_{*}$或者下面估计式成立：

$$0<f(x_{k+1})-f(x_{*})\leqslant\left(1-\frac{\alpha}{C_k}\right)[f(x_k)-f(x_{*})]$$

$$1\leqslant C_{k+1}\leqslant C_k\left[1+\gamma\sqrt{f(x_k)-f(x_{*})}\right]$$

其中 $\alpha=\alpha'/\alpha_5$。

下面的引理表明，如果$\|\varepsilon_1\|C_1$ 充分小，则引理 2.5 中的常数 C_k 一致有界。

引理 2.6　除假定引理 2.5 的假设成立，还假定下面不等式成立：$C_1\sqrt{f(x_1)-f(x_{*})}\leqslant\alpha/2\gamma\mathrm{e}$，那么，对所有 $k\geqslant1$，$C_k\leqslant C_1\mathrm{e}$。其中 α 和 $C_k=C(A_k)$的定义与引理 2.5 和引理 2.3 中的一样。证明见文献[5]。

因此，由引理 2.5 和 2.6，可给出下面定理。

定理 2.1　假设假定 B 成立。考虑算法 1，其中对所有 k，参数 $\theta_k\in[0,1]$，$|\mu_k|\in[0,\bar{\mu})$，$\bar{\mu}<1$，则对任何常数 $\Gamma\geqslant1$，存在常数 $\delta>0$，使得对任何初始点 x_1，$\|x_1-x_{*}\|<\delta$，并且对任何正定矩阵 A_1，$\mathrm{Cond}(A_1)\leqslant\Gamma$，序列$\{x_k\}$，$\{A_k\}$或者有限，即对一些 m，$x_m=x_{*}$，或者无限。在后一种情况下，当 $k\to\infty$时，所有常数 $\mathrm{Cond}(A_k)$，$\|A_k\|$，$\|A_k^{-1}\|$一致有界，因此序列$\{x_k\}$至少 Q-线性收敛于 x_{*}。

3　收敛速率的估计

在 2 中，证明了由锥模型 Broyden 凸族算法产生的序列$\{x_k\}$至少局部 Q-线性收敛。确实有例子表明，对所有 k，使用固定的 $\bar{\mu}=\mu_k\neq0$，仅得到线性收敛速率(见文献[5])；所以，总的来说，如果 μ_k 只在远离零点处有界，则不可能得到更好的收敛速率。在下面引理中将看到，如果 μ_k 足够快地收敛于 0，即不精确线搜索渐近地变成精确线搜索，则有比较好的收敛速率。

引理 3.1　假设假定 B 成立。考虑算法 1，其中对所有 k，参数 $\theta_k\in[0,1]$，$0\leqslant\mu_k\leqslant\min\{\bar{\mu},C\|g_k\|\}$，$\bar{\mu}\in[0,1)$，$C>0$ 是常数。进一步，选择 x_1，A_1 使算法 1 关于参数 θ_k 和 μ_k 生成的 2 个无穷序列$\{x_k\}$，$\{A_k\}$，具有性质

1) $\lim\limits_{k\to\infty}x_k=x_{*}$；

2)对所有 k 和某个常数 Γ，有

$$\max\{\|A_k\|,\|A_k^{-1}\|,\mathrm{Cond}(A_k)\}<\Gamma \tag{14}$$

则存在不依赖于 k 的常数 $\bar{\gamma}_1,\bar{\gamma}_2,\cdots,\bar{\gamma}_n$ 使得对每一充分大的 k，存在一 j，$1\leqslant j\leqslant n$ 有$\|x_{k+j}-$

$x_*\| \leqslant \bar{\gamma}_j \|x_k - x_*\|^2$。特别地，对所有充分大的 k，有

$$\|x_{k+n} - x_*\| \leqslant \bar{\gamma} \|x_k - x_*\|^2, \bar{\gamma} = \max_{1 \leqslant j \leqslant n} \bar{\gamma}_j \tag{15}$$

而且，序列 $\{x_k\}$ 至少 R-收敛于 x_*，收敛阶为 $\tau \geqslant \sqrt[n]{2}$。

证明 证明技巧与文献[5]中关于标准 Broyden 凸类所用的方法类似。其主要思想如下：对于固定的充分大的 k，定义 $\tilde{x}_k = x_k, \tilde{A}_k = A_k$ 并且把对于 $j \geqslant k$，带参数 $\tilde{\theta}_j = \theta_j, \tilde{\mu}_j = 0$ 及初始值 $\tilde{x}_k, \tilde{A}_k$ 的 Broyden 方法应用于二次函数 $\tilde{f}(x) = (x - x_*)^T G_* (x - x_*)/2$，这样，当 k 充分大时，得到迭代点 $\tilde{x}_{k+j}, \tilde{A}_{k-j}$ 分别近似 x_{k+j}, A_{k+j}。众所周知，所有带有 $\tilde{\theta}_k \geqslant 0$ 的精确 Broyden 方法用于严格凸二次函数 $\tilde{f}: \mathbf{R}^n \to \mathbf{R}$，至多在 n 步达到精确极小，即存在一个 $i \leqslant n$，使得 $\tilde{x}_{k+i} = x_*$；因此，如果证明差 $\Delta x_{k+i} = x_{k+i} - \tilde{x}_{k+i}$ 是小量的话，则 $x_{k+i} - x_*$ 是小量。这里细节略去。

定理 2.1 保证，如果 $\|x_1 - x_*\|$ 足够小，则引理 3.1 中的假设 1)和 2)可以对所有的 A_1 成立；因此算法 1 中的点列至少局部收敛且阶为 $\tau \geqslant \sqrt[n]{2}$。更精确地，有如下定理。

定理 3.1 假设假定 B 成立。考虑算法 1，其中对所有 k，参数 $\theta_k \in [0,1]$，$0 \leqslant \mu_k \leqslant \min\{\bar{\mu}, C\|g_k\|\}$，$\bar{\mu} \in [0,1)$，$C > 0$ 为常数，则对任何常数 $\Gamma \geqslant 1$，存在一 $\delta > 0$，使得对任何初始点 x_1，$\|x_1 - x_*\| < \delta$ 和任何正定矩阵 A_1 且 $\mathrm{Cond}(A_1) \leqslant \Gamma$，则存在不依赖于 k 的常数 $\bar{\gamma}_1, \bar{\gamma}_2, \cdots, \bar{\gamma}_n$ 使得对每一充分大的 k，存在一个 j，$1 \leqslant j \leqslant n$ 且 $\|x_{k+j} - x_*\| \leqslant \bar{\gamma}_j \|x_k - x_*\|^2$。特别地，对所有充分大的 k，有

$$\|x_{k+n} - x_*\| \leqslant \bar{\gamma} \|x_k - x_*\|^2, \bar{\gamma} = \max_{1 \leqslant j \leqslant n} \bar{\gamma}_j$$

而且序列 $\{x_k\}$ 至少 R-收敛于 x_*，收敛阶为 $\tau \geqslant \sqrt[n]{2}$。

参 考 文 献

1 Davidon W C. Conic approximations and collinear scalings for optimizers. SIAM J Numer Anal, 1980, 17(2): 268～281

2 Sorensen D C. The Q-superlinear convergence of a collinear scaling algorithm for unconstrained optimization. SIAM J Numer Anal, 1980, 17(1): 84～114

3 Ariyawansa K A, Lau D T M. Local and Q-superlinear convergence of a class of collinear scaling algorithms that extends quasi-Newton methods with Broyden's bounded-Φ class of updates. Optimization, 1992, 23(2): 323～339

4 Deng N Y, Li Z F. Some global convergence properties of a conic-variable metric algorithm for minimization with inexact line searchs. Optimization Methods and Softwares, 1995, 5(1): 105～122

5 Stoer J. On the convergence rate of imperfect minimization algorithm in Broyden's β-class. Math Prog, 1975, 9(2): 313～335

6 Fletcher R. A new approach to variable metric algorithms. Computer J, 1970, 13(2): 317～322