NGLM: 一类全局收敛的 Newton-GMRES 方法 *

安恒斌　白中治

(中国科学院数学与系统科学研究院, 计算数学与科学工程计算研究所)

(科学与工程计算国家重点实验室)

(北京 *100080*)

摘要　本文提出了一类具有全局收敛性质的 Newton-GMRES 方法 —NGLM 方法. 该方法是对经典 Newton-GMRES 方法的推广. NGLM 方法的全局策略是当在非精确 Newton 方向上后退不能成功时，转而在一个子空间上运用信赖域方法确定迭代步长. 理论分析与数值实验均表明, NGLM 方法改善了 Newton-GMRES 方法的强健性.

关键词　非线性方程组，非精确 Newton 法，广义极小残量法 (GMRES), 信赖域方法, Levenberg-Marquardt 方法，全局收敛性.

NGLM: A Globally Convergent Newton-GMRES Method

Heng-Bin An and Zhong-Zhi Bai

State Key Laboratory of Scientific/Engineering Computing

Institute of Computational Mathematics and Scientific/Engineering Computing

Academy of Mathematics and System Sciences

Chinese Academy of Sciences, Beijing 100080, P.R. China

Abstract: For large sparse system of nonlinear equations, we propose a globally convergent Newton-GMRES method, called as the Newton-GMRES with Levenberg-Marquardt strategy (NGLM) method. This method is a technical generalization of the standard Newton-GMRES method. In NGLM method, the adopted global strategy is using the trust region technique in a subspace to find a satisfactory iteration step when the backtracking along the inexact Newton direction fails. Both theoretical analysis and numerical results show that NGLM method is more robust than the Newton-GMRES method.

---

*国家重点基础研究项目“大规模科学计算研究（G1999032803)”专项经费资助课题

Keywords: System of nonlinear equations, inexact Newton method, generalized minimal residual (GMRES) method, trust region technique, Levenberg-Marquardt method, global convergence.

§1 引言

考虑大型稀疏非线性方程组

$$F(x) = 0, \tag{1}$$

其中 $F : \mathbb{R}^n \to \mathbb{R}^n$ 连续可微.

Newton-GMRES(NG) 方法是求解非线性方程组 (1) 的一种特殊而有效的非精确 Newton 算法 [2, 5, 7, 21]. 在 NG 方法的每步迭代中，均需利用 GMRES 方法 [22] 不精确求解 Newton 方程

$$F^{'}(x_k)s = -F(x_k)$$

而得到一个满足约束

$$\|F^{'}(x_k)\bar{s}_k + F(x_k)\| \le \bar{\eta}_k \|F(x_k)\| \tag{2}$$

的非精确 Newton 步 $\bar{s}_k$. 这里， $\bar{\eta}_k \in [0,1)$ 是给定的强制量. 我们称 (2) 为非精确 Newton 条件.

众所周知， NG 方法仅具有局部收敛性质 (见 [1]-[5] 和 [9, 10]). 如果在非精确 Newton 方向 $\bar{s}_k$ 上配以线搜索策略，则所得到的方法便是具有全局收敛性质的 Newton-GMRES with backtracking(NGB) 方法. 该方法可简单描述如下：

算法 1.1 (NGB 方法)

1. 给定 x_0, ϵ_0, $\eta_{max} \in [0,1)$, $\alpha \in (0,1)$, $0 < \theta_l < \theta_u < 1$. 置 $k := 0$.
2. 如果 $\|F(x_k)\| > \epsilon_0$
    2.1. 选取 $\bar{\eta}_k \in [0, \eta_{max}]$.
    2.2. 执行 **GNE** *(GMRES for the kth Newton equation)* 过程:
        (a) 选取 $s_k^0 \in \mathbb{R}^n$, 令 $r_k^0 = -F(x_k) - F^{'}(x_k)s_k^0$.
        (b) 执行 **GMRES** 迭代 m 次得到 *Krylov* 子空间

$$\mathcal{K}_m := span\left\{r_k^0,\, F^{'}(x_k)r_k^0,\, \left(F^{'}(x_k)\right)^2 r_k^0,\, \dots,\, \left(F^{'}(x_k)\right)^{m-1} r_k^0\right\}$$

及其标准正交基 $V_m := [v_1, v_2, \dots, v_m] \in \mathbb{R}^{n\times m}$, 使其满足

$$F^{'}(x_k)V_m = V_{m+1}\bar{H}_m \tag{3}$$

和 (2), 其中 $\bar{H}_m \in \mathbb{R}^{(m+1)\times m}$ 为上 Hessenberg 阵,

$$\bar{s}_k := \mathrm{argmin}_{s \in s_k^0 + \mathcal{K}_m} \|F^{'}(x_k)s + F(x_k)\|.$$

2.3. 执行 **BL** (*backtracking loop*):

(a) 置 $s_k := \bar{s}_k$, $\eta_k := \bar{\eta}_k$.

(b) 如果 $\|F(x_k + s_k)\| > [1 - \alpha(1 - \eta_k)] \|F(x_k)\|$

- 选取 $\theta \in [\theta_l, \theta_u]$.
- 置 $s_k := \theta s_k$, $\eta_k := 1 - \theta(1 - \eta_k)$.

2.4. $x_{k+1} = x_k + s_k$.

2.5. $k := k + 1$.

上述 NGB 方法的优点之一是可以不显式地计算和形成 Jacobi 矩阵. 事实上, 在算法的执行过程中, 我们只用到了 Jacobi 矩阵 $F^{'}(x)$ 与某一向量 y 的乘积, 而这可以通过有限差分

$$F^{'}(x)y \approx \frac{F(x + \epsilon y) - F(x)}{\epsilon} \tag{4}$$

来近似实现. 这里, ϵ 是差分步长 [3, 16]. 理论分析与数值实验均表明, NGB 方法具有很好的强健性, 用于求解大规模稀疏非线性方程组十分有效. 但是, 对于一些坏条件问题, NGB 方法却常常不能达到令人满意的计算效果, 甚至会出现不收敛的情形.

针对 NGB 方法的不足, 文 [1] 提出了求解非线性方程组 (1) 的 Newton-GMRES with quasi-conjugate-gradient backtracking(NGQCGB) 方法. 该方法的全局策略是: 首先, 在非精确 Newton 方向 $\bar{s}_k$ 上后退最多 N_b 步. 如果在这 N_b 步内能够得到当前迭代点 x_k 处的一个充分下降步 s_k, 则令 $x_{k+1} = x_k + s_k$, 并进行下一次的非线性迭代. 否则, 运用 QCGB(quasi-conjugate-gradient backtracking) 策略来产生一个充分下降步.

QCGB 策略作为 NGB 的一种补救措施, 其实质是通过在子空间

$$\Omega_k = \mathrm{span}\{\tilde{g}_k, \Delta_{k-1}\}$$

上极小化 $F(x)$ 在 x_k 的局部线性模型而得到一个方向 d_k[1], 即

$$d_k \in \mathrm{argmin}_{d \in \mathbb{R}^n} \|F(x_k) + F^{'}(x_k)d\|,$$

然后, 再通过线搜索技巧确定一个满足 "充分下降" 条件的步长因子 θ_k, 并令 $x_{k+1} = x_k + \theta_k d_k$. 注意到执行 QCGB 策略的前提是在当前的非精确 Newton 方向 $\bar{s}_k$ 上后退

[1] 这里 $\tilde{g}_k$ 是 $g_k := \nabla f(x_k)$ 在 Krylov 子空间上的投影, 其中 $f(x) = \frac{1}{2}\|F(x)\|_2^2$, $\Delta_{k-1} = x_k - x_{k-1}$.

N_b 步以后仍没有得到一个 "充分下降" 步，因此，我们可以断定在 NG 方法中得到的当前的非精确 Newton 方向 $\bar{s}_k$ 不是一个好方向．其原因可能是 $F(x)$ 在 x_k 处的性态不够好，以至于用 NG 方法计算出的 $\bar{s}_k$ 偏离 $f(x)=\frac{1}{2}\|F(x)\|_2^2$ 的最速下降方向 $-\bigtriangledown f(x_k)$ 太远，也可能是 $\|\bar{s}_k\|$ 太大，以至于 $F(x)$ 在 x_k 处的线性模型与 $F(x_k+\bar{s}_k)$ 的吻合太差所致.

基于上述分析，在本文中，我们提出了一种与 NGQCGB 方法相平行的具有全局收敛性质的 Newton-GMRES 方法，简称为 NGLM 方法. 该方法是将 NGQCGB 中的 QCGB 策略用 Levenberg-Marquardt(LM) 策略替代而得到的．在本质上， LM 策略是一种信赖域型的方法.

求解非线性方程组的信赖域方法是一种具有全局收敛性质的方法．在每步迭代中，这种方法首先确定当前迭代点 x_k 的一个邻域

$$N_{h_k}(x_k) := \{x \mid \|x-x_k\| \leq h_k\}$$

(其中 h_k 是信赖域半径), 然后通过求解信赖域模型

$$\min_{\|s\|\leq h_k} \|F(x_k)+F^{'}(x_k)s\| \tag{5}$$

去确定一个搜索方向 s_k. 如果 s_k 满足某种下降条件, 则采用 s_k, 并令 $x_{k+1}=x_k+s_k$. 否则，剔除该 s_k, 缩小信赖域半径 h_k, 并重新求解 (5), 直到得到满足下降条件的 s_k 为止．这里 $\|\cdot\|$ 可以取为 $\|\cdot\|_1$， $\|\cdot\|_2$ 或 $\|\cdot\|_\infty$.

最重要的一类信赖域方法是在 (5) 中取 $\|\cdot\|_2$ 而得到的 Levenberg-Marquardt 方法，即

$$\min_{\|s\|_2\leq h_k} \|F(x_k)+F^{'}(x_k)s\|_2.$$

该模型可以通过求解线性方程组

$$\left(F^{'}(x_k)^T F^{'}(x_k)+\mu I\right)s = -F^{'}(x_k)^T F(x_k)$$

来表征，其中 μ 是一个非负参数 [8].

如果在 NG 方法中所产生的非精确 Newton 方向 $\bar{s}_k$ 上后退 N_b 步以后, 仍然不能得到一个 "充分下降" 步, 则我们考虑在一个子空间上运用上述 Levenberg-Marquardt 方法来寻找一个新的搜索方向．这就是我们提出 Newton-GMRES with Levenberg-Marquardt(NGLM) 方法的基本思想.

在下面的讨论中，我们用 $\|\cdot\|$ 表示 Euclidean 范数，并约定

$$f_2(x)=\frac{1}{2}\|F(x)\|_2^2, \quad g_k=\bigtriangledown f_2(x_k).$$

容易验证， $g_k=F^{'}(x_k)^T F(x_k)$.

§2 NGLM 方法的建立

首先，我们考虑投影子空间的构造．在 NGB 方法的第 k 步迭代中，根据以下两种不同情况，我们可以分别构造出相应的子空间．

(C_1) $s_k^0 \in \mathcal{K}_m$. 在这种情形下，我们令

$$\Omega_k = \mathrm{span}\{\tilde{g}_k,\, \Delta_{k-1},\, v\}, \tag{6}$$

其中 $\Delta_{k-1} = x_k - x_{k-1}$, $\tilde{g}_k = V_m V_m^T g_k$, 而 v 是 Krylov 子空间 $\mathcal{K}_m$ 的基 V_m 中满足

$$|v^T g_k| = \max_{1\le j\le m} |v_j^T g_k|$$

的单位向量．这样选取的 v 是 V_m 中与 $\mathrm{span}\{g_k\}$ 的夹角最小的向量．由 (3) 知，

$$v_j^T g_k = v_j^T F'(x_k)^T F(x_k) = -\beta_k v_1^T F'(x_k) v_j = -\beta_k h_{1j}.$$

故当

$$|h_{1i}| = \max_{1\le j\le m} |h_{1j}|$$

时， $v = v_i$, 其中 $(h_{11}, h_{12}, \ldots, h_{1m})$ 为 Hessenberg 矩阵 $\bar{H}_m$ 的第一行．

(C_2) $s_k^0 \notin \mathcal{K}_m$. 在这种情形下，我们适当地扩大 Krylov 子空间 $\mathcal{K}_m$. 为此，令

$$\hat{\mathcal{K}} = \mathrm{span}\{\mathcal{K}_m,\, s_k^0\},$$

并设

$$Q = [q_1, q_2, \ldots, q_{m+1}] \in \mathbb{R}^{n\times(m+1)}$$

为 $\hat{\mathcal{K}}$ 的标准正交基．由于 Krylov 子空间 $\mathcal{K}_m$ 的标准正交基 V_m 是已知的，故我们取 $q_j = v_j,\ j = 1, 2, \ldots, m$. 令 $\tilde{g}_k$ 为 g_k 在子空间 $\hat{\mathcal{K}}$ 上的投影，即

$$\tilde{g}_k = QQ^T F'(x_k)^T F(x_k),$$

并选取 v 为 $\hat{\mathcal{K}}$ 的标准正交基 Q 中满足

$$|q^T g_k| = \max_{1\le j\le m+1} |q_j^T g_k|$$

的单位向量．与 (6) 相似，同样定义子空间 Ω_k. 若再记 $p = Q^T g_k$, 则有

$$\begin{aligned}
p &= Q^T F'(x_k)^T F(x_k) \\
&= \begin{pmatrix} V_m^T F'(x_k)^T F(x_k) \\ q_{m+1}^T F'(x_k)^T F(x_k) \end{pmatrix} \\
&= \begin{pmatrix} -\beta_k \bar{H}_m^T e_1 \\ q_{m+1}^T F'(x_k)^T F(x_k) \end{pmatrix}.
\end{aligned}$$

由此可得

$$\begin{aligned}\tilde{g}_k &= Q\begin{pmatrix} -\beta_k \bar{H}_m^T e_1 \\ q_{m+1}^T F^{'}(x_k)^T F(x_k) \end{pmatrix} \\ &= -\beta_k V_m \bar{H}_m^T e_1 + \left(F(x_k)^T F^{'}(x_k) q_{m+1}\right) q_{m+1}.\end{aligned}$$

因为 $F^{'}(x_k)q_{m+1}$ 可用有限差分公式 (4) 近似计算，因此 $\tilde{g}_k$ 的形成只需额外增加一次函数的赋值即可. 故我们的方法亦无需形成 Jacobi 矩阵 [3, 17]. 又由于

$$q_j^T g_k = q_j^T F^{'}(x_k)^T F(x_k) = p_j,$$

故当

$$|p_i| = \max_{1 \le j \le m+1} |p_j|$$

时， $v = q_i$, 其中 p_j 为向量 p 的第 j 个分量. 值得注意的是，向量 p 的前 m 个分量恰好构成了 Hessenberg 矩阵 $\bar{H}_m$ 的第一行.

现在，我们考虑如下信赖域模型问题：

$$\begin{array}{ll} \min_{s \in \Omega_k} & \|F(x_k) + F^{'}(x_k)s\|_2, \\ s.t. & \|s\|_2 \le h_k, \end{array} \tag{7}$$

其中 $h_k > 0$ 是信赖域半径， Ω_k 是如 (6) 所定义的子空间. 设 W_k 为 Ω_k 的标准正交基， $m(k) = \dim(\Omega_k)$ 为 Ω_k 的维数，则 (7) 可转化为

$$\begin{array}{ll} \min_{z \in \mathbb{R}^{m(k)}} & \|F(x_k) + F^{'}(x_k)W_k z\|_2, \\ s.t. & \|z\|_2 \le h_k. \end{array} \tag{8}$$

记

$$B_k = F^{'}(x_k)^T F^{'}(x_k),$$

则 (8) 等价于

$$\min_{\|z\|_2 \le h_k} (W_k^T g_k)^T z + \frac{1}{2} z^T (W_k^T B_k W_k) z. \tag{9}$$

由最优性条件易得如下结论：

命题 2.1 如果 $F^{'}(x_k)$ 可逆，则 z_k 为约束最优化问题 *(9)* 的解的充要条件是：存在 $\mu_k \ge 0$, 使得

$$\left(W_k^T B_k W_k + \mu_k I\right) z_k = -W_k^T g_k. \tag{10}$$

对于约束最优化问题 (9), 我们只要进行非精确求解即可. 事实上, 在实际计算中, 我们首先确定参数 μ_k, 再通过 (10) 计算出相应的 z_k, 然后令 $\tilde{s}_k = W_k z_k$. 参数 μ_k 的确定, 对于算法的成功起着关键性的作用. 我们所采用的确定方式如下 [11]:

$$\mu_k = \rho_k \|F(x_k)\|^{\tau},$$

其中 $\tau \in (0,1]$ 为给定的常数. 在每次迭代开始时, 对于给定的 ρ_k, 我们即可计算出相应的 $\tilde{s}_k$. 如果 $\tilde{s}_k$ 满足充分下降条件, 则得到下一迭代点 $x_{k+1} = x_k + \tilde{s}_k$. 否则, 适当增大 ρ_k 的值, 并重新计算相应的 $\tilde{s}_k$, 直到得到满足充分下降条件的 $\tilde{s}_k$ 为止. 这就是我们所说的 LM 策略.

在下面的讨论中, 我们定义函数 $F(x)$ 在点 x_k 处关于 s_k 的实际下降量 $\mathrm{Ared}_k(s_k)$ 及预计下降量 $\mathrm{Pred}_k(s_k)$ 分别为

$$\begin{aligned} \mathrm{Ared}_k(s_k) &= \|F(x_k)\| - \|F(x_k + s_k)\|, \\ \mathrm{Pred}_k(s_k) &= \|F(x_k)\| - \|F(x_k) + F^{'}(x_k) s_k\|. \end{aligned}$$

现在, 我们即可给出 NGLM 方法的具体描述.

算法 2.1 (NGLM 方法)

1. 给定初值 x_0, 精度 ϵ_0, 及参数 $\eta_{max} \in [0,1)$, $0 < \alpha < 1$, $0 < \tau \le 1$, $0 < \theta_l < \theta_u < 1$, $N_b \ge 0$, $\vartheta > 1$. 置 $k := 0$, $\Delta_{-1} := 0$.

2. 如果 $\|F(x_k)\| > \epsilon_0$

   2.1. 选取 $\bar{\eta}_k \in [0, \eta_{max}]$.

   2.2. 通过 **GNE** 计算非精确 *Newton* 步 $\bar{s}_k$, 使其满足

   $$\|F(x_k) + F^{'}(x_k)\, \bar{s}_k\| \le \bar{\eta}_k \|F(x_k)\|.$$

   2.3. 在 $\bar{s}_k$ 上最多后退 N_b 步, 得到 s_k , 并得到相应的 η_k.

   2.4. 如果

   $$\|F(x_k + s_k)\| \le [1 - \alpha\,(1 - \eta_k)]\, \|F(x_k)\|,$$

   则令 $\Delta_k = s_k$, 转步 *2.6*.

   2.5. 执行 *LM* 策略:

      2.5.1. 构造子空间 Ω_k, 并计算出相应的基 W_k. 取 $\rho_k = 10^{-4}$, $r_k = 0$.

      2.5.2. 如果 $r_k < \alpha$

      - 通过求解

      $$\left(W_k^T B_k W_k + \rho_k \|F(x_k)\|^{\tau} I\right) z = -W_k^T g_k$$

      得到 z_k, 并计算 $\tilde{s}_k = W_k z_k$.

- 令 $r_k = \mathrm{Ared}_k(\tilde{s}_k)/\mathrm{Pred}_k(\tilde{s}_k)$， $\rho_k := \vartheta\rho_k$.

2.5.3. $\Delta_k := \tilde{s}_k$,

2.6. $x_{k+1} = x_k + \Delta_k$.

2.7. $k := k+1$.

算法 2.1 的步 2.5 是 LM 策略. 在每次执行 LM 策略时，首先取 $\rho_k = 10^{-4}$, 然后在 LM 策略的循环体内不断地以 ϑ 倍来扩大 ρ_k. 循环终止的条件是 $\tilde{s}_k$ 对应的实际下降量 $\mathrm{Ared}_k(\tilde{s}_k)$ 与预计下降量 $\mathrm{Pred}_k(\tilde{s}_k)$ 的比值 r_k 大于常数 α.

§3 收敛性分析

为了证明 NGLM 方法的全局收敛性，我们对非线性函数 $F:\mathbb{R}^n \to \mathbb{R}^n$ 做如下基本 **假定**:

A_1. 对给定的点 $x_0 \in \mathbb{R}^n$, 函数 $f_2(x) = \frac{1}{2}\|F(x)\|^2$ 在水平集

$$\mathcal{L} = \{x \mid f_2(x) \le f_2(x_0)\}$$

上连续可微，并且存在常数 $L > 0$, 使得

$$\|\bigtriangledown f_2(x) - \bigtriangledown f_2(y)\| \le L\|x-y\|, \quad \forall x, y \in \mathcal{L};$$

A_2. 对于由 *NGLM* 方法产生的序列 $\{x_k\}$, 下述关系式总成立:

$$\|F(x_k) + F'(x_k)\bar{s}_k\| \le \eta_{max}\|F(x_k)\|, \quad k = 0, 1, 2, \ldots,$$

其中 $\bar{s}_k$ 为非精确 *Newton* 步.

我们记

$$\left\{\begin{array}{rcl} z_k(\mu) & = & -\left[W_k^T B_k W_k + \mu I\right]^{-1} W_k^T g_k, \\ s_k(\mu) & = & W_k z_k(\mu), \\ \eta(\mu) & = & \dfrac{\|F(x_k) + F'(x_k)s_k(\mu)\|}{\|F(x_k)\|}, \end{array}\right. \qquad 0 \le \mu \le \infty. \tag{11}$$

关于 $s_k(\mu)$ 及 $\eta(\mu)$, 我们有下述两个命题. 由于它们的证明方法类似，故我们只给出命题 3.2 的证明.

命题 3.1 $s_k(\mu)$ 与 $-g_k$ 的夹角关于 μ 单调非增.

命题 3.2 $\eta(\mu)$ 是 μ 的严格单调增函数.

证明 为简化记号，我们在证明中略去所有下标，并记

$$F = F(x_k), \quad J = F'(x_k), \quad \hat{g} = W^T g, \quad s = s_k(\mu).$$

由具体计算可得

$$
\begin{aligned}
\eta'(\mu) &= \frac{1}{\|F\|\,\|F+Js\|}\cdot(F+Js)^T J\frac{ds}{d\mu} \\
&= \frac{1}{\|F\|\,\|F+Js\|}\cdot(F+Js)^T JW\left[W^TBW+\mu I\right]^{-2}\hat{g} \\
&= \frac{1}{\|F\|\,\|F+Js\|}\cdot\Big\{\hat{g}^T\left[W^TBW+\mu I\right]^{-2}\hat{g} \\
&\quad +\hat{g}^T\left[W^TBW+\mu I\right]^{-1}W^TBW\left[W^TBW+\mu I\right]^{-2}\hat{g}\Big\} \\
&:= \frac{1}{\|F\|\,\|F+Js\|}\cdot G.
\end{aligned}
$$

这里，我们定义

$$
G=\hat{g}^T\left[I+\left(W^TBW+\mu I\right)^{-1}W^TBW\right]\left(W^TBW+\mu I\right)^{-2}\hat{g}.
$$

设

$$
W^TBW=Q^T\Lambda Q,
$$

其中

$$
\Lambda=\mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_{m(k)}),\quad \lambda_j>0,\quad 1\le j\le m(k),
$$

并令

$$
Q\hat{g}:=u=(u_1,u_2,\ldots,u_{m(k)})^T,
$$

则

$$
\begin{aligned}
G &= u^T(\Lambda+\mu I)^{-2}u+u^T(\Lambda+\mu I)^{-1}\Lambda(\Lambda+\mu I)^{-2}u \\
&= \sum_{j=1}^{m(k)}\left[\frac{u_j^2}{(\lambda_j+\mu)^2}+\frac{\lambda_j u_j^2}{(\lambda_j+\mu)^3}\right] \\
&> 0.
\end{aligned}
$$

从而， $\eta'(\mu)>0$, 亦即 $\eta(\mu)$ 是 μ 的严格单调增函数. ∎

令

$$
\tilde{\eta}_k:=\frac{\|F(x_k)+F'(x_k)s_k(0)\|}{\|F(x_k)\|}=\eta(0),
$$

则易知 $\tilde{\eta}_k\in[0,1)$. 由命题 3.2 知，当 μ 从 0 连续变化到 $+\infty$ 时， $\eta(\mu)$ 相应地从 $\tilde{\eta}_k$ 连续变化到 1. 如果我们再令

$$
E_k(\eta(\mu)):=s_k(\mu),\quad 0\le\mu<+\infty, \tag{12}
$$

则

$$\|F(x_k)+F^{'}(x_k)E_k(\eta)\|=\eta\|F(x_k)\|,\quad \eta\in[\tilde{\eta}_k,1). \tag{13}$$

假定在 NGLM 方法的第 k 步迭代中执行了 LM 策略，并设相应的步长为

$$\Delta_k=s_k(\mu_k)=-W_k\left[W_k^TB_kW_k+\mu_kI\right]^{-1}W_k^Tg_k.$$

对应于这里的 μ_k, 令 $\eta_k:=\eta(\mu_k)$ [2], 其中 $\eta(\mu)$ 如 (11) 所定义. 则由 (13) 可得

$$\|F(x_k)+F^{'}(x_k)\Delta_k\|=\eta_k\|F(x_k)\|.$$

由算法 2.1 知,

$$r_k=\frac{\mathrm{Ared}_k(\Delta_k)}{\mathrm{Pred}_k(\Delta_k)}\geq\alpha,\quad \alpha\in(0,1).$$

于是，我们有

$$\|F(x_k+\Delta_k)\|\leq[1-\alpha(1-\eta_k)]\,\|F(x_k)\|.$$

上述分析表明， NGLM 方法归属于下面的算法框架:

1. 给定 $x_0\in\mathbb{R}^n$, $\alpha\in(0,1)$;

2. 对 $k=0,1,2,\ldots$ 直到 $\{x_k\}$ 收敛,

2.1. 确定 $\eta_k\in[0,1)$ 及 $\Delta_k\in\mathbb{R}^n$, 使得

$$\|F(x_k)+F^{'}(x_k)\Delta_k\|\leq\eta_k\|F(x_k)\| \tag{14}$$

并且

$$\|F(x_k+\Delta_k)\|\leq[1-\alpha\,(1-\eta_k)]\|F(x_k)\|. \tag{15}$$

2.2. 令 $x_{k+1}:=x_k+\Delta_k$.

需要指出的是， NGLM 方法中的 Δ_k 或者由 BL 策略沿着曲线

$$\sigma_k(\eta)=\frac{1-\eta}{1-\bar{\eta}_k}\bar{s}_k,\quad \eta\in[\bar{\eta}_k,1] \tag{16}$$

所确定，或者由 LM 策略沿着曲线 $E_k(\eta)\,(\tilde{\eta}_k\leq\eta\leq1)$ 所确定.

以下几个引理对于我们分析 NGLM 方法的收敛性是必要的. 其中前两个引理是 [9] 中的结果.

[2]由于执行了 LM 策略，故在 $\bar{s}_k$ 上的后退过程中所确定的 η_k 在以后的计算中勿需用到. 因此，我们在这里用符号 η_k 不致引起混淆.

引理 3.1 *[9, 定理 3.4]* 设序列 $\{x_k\} \subset \mathbb{R}^n$ 满足 *(14)* 及 *(15)*. 如果 $\sum_{k=0}^{+\infty}(1-\eta_k) = +\infty$, 则 $F(x_k) \to 0$ $(k \to +\infty)$.

引理 3.2 *[9, 定理 3.5]* 设序列 $\{x_k\} \subset \mathbb{R}^n$ 满足 *(14)* 及 *(15)*. 如果 $\{x_k\}$ 有一个聚点 x^*, 使得当 x_k 充分靠近 x^* 并且 k 充分大时，有

$$\|\Delta_k\| \leq \Gamma(1-\eta_k)\|F(x_k)\|,$$

则 $x_k \to x^*$ $(k \to +\infty)$. 这里， Γ 是一个与 k 无关的常数.

由 [1] 中的引理 4.4 及其证明过程，我们可得如下引理：

引理 3.3 在 *NGLM* 方法中，取 $\epsilon_0 = 0$, 并设已由该方法产生了一个无穷序列 $\{x_k\} \subset \mathbb{R}^n$. 如果 $\{x_k\}$ 有一个聚点 x^*, 使得 $F^{'}(x^*)$ 可逆，则存在与 k 无关的常数 $\lambda > 0$, 使得当 $x_k \in N_\delta(x^*)$ 时,

$$\begin{array}{rcl} \|F(x_k) + F^{'}(x_k)s_k(0)\| & \leq & \sqrt{1-\lambda^2}\|F(x_k)\|, \\ \dfrac{|\langle F(x_k), F^{'}(x_k)s_k(0)\rangle|}{\|F(x_k)\|\,\|F^{'}(x_k)s_k(0)\|} & \geq & \lambda. \end{array} \tag{17}$$

这里， $\delta > 0$ 为充分小的常数，使得当 $x \in N_\delta(x^*)$ 时， $F^{'}(x)$ 可逆且 $\|F^{'}(x)^{-1}\| \leq 2\|F^{'}(x^*)^{-1}\|$ [3].

引理 3.4 如果 $F^{'}(x_k)$ 可逆，则

$$\frac{|\langle F(x_k), F^{'}(x_k)s_k(\mu)\rangle|}{\|F(x_k)\|\,\|F^{'}(x_k)s_k(\mu)\|} \geq \frac{1}{\kappa(F^{'}(x_k))^2} \cdot \frac{|\langle F(x_k), F^{'}(x_k)s_k(0)\rangle|}{\|F(x_k)\|\,\|F^{'}(x_k)s_k(0)\|}.$$

这里, $\kappa(F^{'}(x_k))$ 表示矩阵 $F^{'}(x_k)$ 的 *Euclidean* 条件数 *(*亦即 $\kappa(F^{'}(x_k)) = \|F^{'}(x_k)\|\|F^{'}(x_k)^{-1}\|$*)*, $s_k(\mu)$ 由 *(11)* 定义. 下同.

证明 由命题 3.1 知,

$$\frac{|\langle F(x_k), F^{'}(x_k)s_k(\mu)\rangle|}{\|F^{'}(x_k)^T F(x_k)\|\,\|s_k(\mu)\|} \geq \frac{|\langle F(x_k), F^{'}(x_k)s_k(0)\rangle|}{\|F^{'}(x_k)^T F(x_k)\|\,\|s_k(0)\|}.$$

因此,

$$\begin{array}{rcl} \dfrac{|\langle F(x_k), F^{'}(x_k)s_k(\mu)\rangle|}{\|F(x_k)\|\,\|F^{'}(x_k)s_k(\mu)\|} & \geq & \dfrac{1}{\kappa(F^{'}(x_k))} \cdot \dfrac{|\langle F(x_k), F^{'}(x_k)s_k(\mu)\rangle|}{\|F^{'}(x_k)^T F(x_k)\|\,\|s_k(\mu)\|} \\ & \geq & \dfrac{1}{\kappa(F^{'}(x_k))} \cdot \dfrac{|\langle F(x_k), F^{'}(x_k)s_k(0)\rangle|}{\|F^{'}(x_k)^T F(x_k)\|\,\|s_k(0)\|} \\ & \geq & \dfrac{1}{\kappa(F^{'}(x_k))^2} \cdot \dfrac{|\langle F(x_k), F^{'}(x_k)s_k(0)\rangle|}{\|F(x_k)\|\,\|F^{'}(x_k)s_k(0)\|}. \end{array}$$

∎

[3] [20, 引理 2.3.3] 保证了这样的 δ 的存在性.

引理 3.5 在 *NGLM* 方法中，取 $\epsilon_0 = 0$，并设已由该方法产生了一个无穷序列 $\{x_k\} \subset \mathbb{R}^n$. 如果 $\{x_k\}$ 有一个聚点 x^*，使得 $F^{'}(x^*)$ 可逆，则存在与 k 无关的常数 $\Gamma > 0$，使得当 $x_k \in N_\delta(x^*)$ 时，

$$\|E_k(\eta)\| \le \Gamma(1-\eta)\|F(x_k)\|, \quad \eta \in [\tilde{\eta}_k, 1]. \tag{18}$$

这里，$E_k(\eta)$ 由 *(12)* 定义，而 $\delta > 0$ 为充分小的常数，使得当 $x \in N_\delta(x^*)$ 时，$F^{'}(x)$ 可逆并且

$$\|F^{'}(x)\| \le 2\|F^{'}(x^*)\|, \quad \|F^{'}(x)^{-1}\| \le 2\|F^{'}(x^*)^{-1}\|.$$

证明 由引理 3.3 知，存在与 k 无关的常数 $\lambda > 0$，使得当 $x_k \in N_\delta(x^*)$ 时，(17) 成立. 现令

$$æ = \max\{\|F^{'}(x^*)\|, \|F^{'}(x^*)^{-1}\|\}.$$

则由引理 3.4 知，当 $\tilde{\eta}_k \le \eta < 1$ 时，

$$\begin{aligned}
\frac{|\langle F(x_k), F^{'}(x_k)E_k(\eta)\rangle|}{\|F(x_k)\|\,\|F^{'}(x_k)E_k(\eta)\|} &= \frac{|\langle F(x_k), F^{'}(x_k)s_k(\mu)\rangle|}{\|F(x_k)\|\,\|F^{'}(x_k)s_k(\mu)\|} \\
&\ge \frac{1}{\kappa(F^{'}(x_k))^2} \cdot \frac{|\langle F(x_k), F^{'}(x_k)s_k(0)\rangle|}{\|F(x_k)\|\|F^{'}(x_k)s_k(0)\|} \\
&\ge \frac{1}{\kappa(F^{'}(x_k))^2} \cdot \frac{|\langle F(x_k), F^{'}(x_k)s_k(0)\rangle|}{\|F^{'}(x_k)^{-1}\|\|F^{'}(x_k)^T F(x_k)\|\|F^{'}(x_k)\|\|s_k(0)\|} \\
&\ge \frac{1}{\kappa(F^{'}(x_k))^3} \cdot \frac{|\langle F(x_k), F^{'}(x_k)s_k(0)\rangle|}{\|F^{'}(x_k)^T F(x_k)\|\,\|s_k(0)\|} \\
&\ge \frac{\lambda}{64æ^4} \\
&:= \gamma > 0.
\end{aligned}$$

故

$$\inf_{\tilde{\eta}_k \le \eta < 1} \frac{|\langle F(x_k), F^{'}(x_k)E_k(\mu)\rangle|}{\|F(x_k)\|\,\|F^{'}(x_k)E_k(\mu)\|} \ge \gamma.$$

再由 [9] 中的引理 7.2 可知，当 $\max\left\{\tilde{\eta}_k, (1-\lambda^2)^{\frac{1}{2}}\right\} \le \eta \le 1$ 时，

$$\begin{aligned}
\|E_k(\eta)\| &\le \frac{2\|F^{'}(x_k)^{-1}\|}{\gamma} \cdot (1-\eta)\|F(x_k)\| \\
&\le \frac{64æ^5}{\lambda} \cdot (1-\eta)\|F(x_k)\| \\
&:= \Gamma_1(1-\eta)\|F(x_k)\|.
\end{aligned}$$

特别，当 $1 \geq \eta \geq \tilde{\eta}_k \geq (1-\lambda^2)^{\frac{1}{2}}$ 或当 $\tilde{\eta}_k < (1-\lambda^2)^{\frac{1}{2}} \leq \eta \leq 1$ 时，上述不等式成立. 现只需考虑 $\tilde{\eta}_k \leq \eta < (1-\lambda^2)^{\frac{1}{2}}$ 的情形. 此时，由引理 3.3 可知,

$$\|F(x_k) + F^{'}(x_k)s_k(0)\| \quad \leq \quad \sqrt{1-\lambda^2}\|F(x_k)\|.$$

因此,

$$\begin{aligned}
\|E_k(\eta)\| &\leq \|s_k(0)\| \\
&\leq \|F^{'}(x_k)^{-1}\| \left(\|F(x_k) + F^{'}(x_k)s_k(0)\| + \|F(x_k)\|\right) \\
&\leq 2æ \cdot \frac{1+\sqrt{1-\lambda^2}}{1-\sqrt{1-\lambda^2}} \cdot (1-\eta)\|F(x_k)\| \\
&:= \Gamma_2(1-\eta)\|F(x_k)\|.
\end{aligned}$$

现令 $\Gamma = \max\{\Gamma_1, \Gamma_2\}$, 立得 (18). ∎

对于由 (16) 所定义的曲线 $\sigma_k(\eta)$, 根据 [9] 中定理 6.1 的证明，我们可知如下结论成立:

引理 3.6 *在 NGLM 方法中，取 $\epsilon_0 = 0$, 并设已由该方法产生了一个无穷序列 $\{x_k\} \subset \mathbb{R}^n$. 如果 $\{x_k\}$ 有一个聚点 x^*, 使得 $F^{'}(x^*)$ 可逆，则当 $x_k \in N_\delta(x^*)$ 时,*

$$\|\sigma_k(\eta)\| \leq \Gamma(1-\eta)\|F(x_k)\|,$$

其中

$$\Gamma = 2\|F^{'}(x^*)^{-1}\| \cdot \frac{1+\eta_{max}}{1-\eta_{max}},$$

而 $\delta > 0$ 为充分小的常数，使得当 $x \in N_\delta(x^*)$ 时， $F^{'}(x)$ 可逆并且 $\|F^{'}(x)^{-1}\| \leq 2\|F^{'}(x^*)^{-1}\|$.

定理 3.1 *在 NGLM 方法中，取 $\epsilon_0 = 0$, 并设已由该方法产生了一个无穷序列 $\{x_k\} \subset \mathbb{R}^n$. 如果 $\{x_k\}$ 有一个聚点 x^*, 使得 $F^{'}(x^*)$ 可逆，则 $x_k \to x^*$ $(k \to +\infty)$.*

证明 设 $\delta > 0$ 充分小，使得当 $x \in N_\delta(x^*)$ 时， $F^{'}(x)$ 可逆并且

$$\|F^{'}(x)\| \leq 2\|F^{'}(x^*)\|, \quad \|F^{'}(x)^{-1}\| \leq 2\|F^{'}(x^*)^{-1}\|.$$

再设 $x_k \in N_\delta(x^*)$, 并记 $\Delta_k = x_{k+1} - x_k$. 如果 Δ_k 是由 LM 策略所产生的，则由引理 3.5 知：存在与 k 无关的常数 $\Gamma^{'} > 0$, 使得

$$\|\Delta_k\| := \|E_k(\eta_k)\| \leq \Gamma^{'}(1-\eta_k)\|F(x_k)\|.$$

而如果 Δ_k 是由 BL 策略所产生的，则由引理 3.6 知：存在与 k 无关的常数 $\Gamma^{''} > 0$, 使得

$$\|\Delta_k\| \leq \Gamma^{''}(1-\eta_k)\|F(x_k)\|.$$

于是，若令 $\Gamma=\max\{\Gamma^{'},\Gamma^{''}\}$, 则对于一切 $x_k\in N_\delta(x^*)$, 总有

$$\|\Delta_k\|\le\Gamma(1-\eta_k)\|F(x_k)\|.$$

又因为 x^* 是 $\{x_k\}$ 的一个聚点，故由引理 3.2 可知 $x_k\to x^*$ $(k\to+\infty)$. ∎

类似于 [1] 中定理 4.2 的证明，我们可得如下定理:

定理 3.2 *在 NGLM 方法中，取 $\epsilon_0=0$, 并设已由该方法产生了一个无穷序列 $\{x_k\}\subset\mathbb{R}^n$. 如果 $\{x_k\}$ 有一个无穷子列 $\{x_{k_i}\}$ 是由 BL 策略产生的，满足 $x_{k_i}\to x^*$ $(i\to+\infty)$ 且 $F^{'}(x^*)$ 可逆，则 $F(x_k)\to 0$ $(k\to+\infty)$.*

引理 3.7 *假设在 NGLM 方法的第 k 步迭代中执行了 LM 策略. 如果 $F^{'}(x_k)$ 可逆，则*

$$\frac{\|W_k^Tg_k\|}{\|g_k\|}\ge\frac{1-\eta_{max}}{\kappa(F^{'}(x_k))(1+\eta_{max})}.$$

证明 据假定 A_2 及 [6] 中的引理 3.5, 我们得

$$\frac{\|V_m^Tg_k\|}{\|g_k\|}\ge\frac{1-\eta_{max}}{\kappa(F^{'}(x_k))(1+\eta_{max})},$$

其中 V_m 是第 k 步非线性迭代中的 Krylov 子空间 $\mathcal{K}_m(F^{'}(x_k),r_k^0)$ 的标准正交基. 注意到 $V_mV_m^Tg_k\in\Omega_k$, 故

$$\|W_kW_k^Tg_k\|\ge\|V_mV_m^Tg_k\|.$$

由此，引理得证. ∎

引理 3.8 *假设在 NGLM 方法的第 k 步迭代中执行了 LM 策略. 如果 $F(x_k)\ne 0$, 并且存在 $\Gamma>0$, 使得*

$$\|E_k(\eta)\|\le\Gamma(1-\eta)\|F(x_k)\|,\quad \eta\in[\tilde{\eta}_k,1],$$

其中 $E_k(\eta)$ 由 (12) 所定义，则 LM 策略中的循环将在有限步后终止，并且最后所得到的 ρ_k 满足

$$\rho_k\le\max\left\{\vartheta,\quad \frac{\vartheta\Gamma}{\delta}\cdot\|F^{'}(x_k)\|^2\|F(x_k)\|^{1-\tau}\right\}.$$

这里，$\delta>0$ 为充分小的常数，使得当 $x\in N_\delta(x_k)$ 时，成立

$$\|F(x)-F(x_k)-F^{'}(x_k)(x-x_k)\|\le\frac{1-\alpha}{\Gamma}\cdot\|x-x_k\|.$$

证明 为简化记号，我们在证明中做如下约定:

$$F_k=F(x_k),\quad J_k=F^{'}(x_k).$$

设 $\eta \in [\tilde{\eta}_k, 1]$, 满足 $\eta > 1 - \frac{\delta}{\Gamma\|F_k\|}$. 则

$$\|E_k(\eta)\| \leq \Gamma(1-\eta)\|F_k\| < \delta,$$

并且由 $E_k(\eta)$ 的性质 (13) 得

$$\begin{aligned}
r_k(E_k(\eta)) &= \frac{\mathrm{Ared}_k(E_k(\eta))}{\mathrm{Pred}_k(E_k(\eta))} \\
&= \frac{\|F_k\| - \|F(x_k + E_k(\eta))\|}{\|F_k\| - \|F_k + J_k E_k(\eta)\|} \\
&\geq \frac{\|F_k\| - \|F_k + J_k E_k(\eta)\| - \|F(x_k + E_k(\eta)) - F_k - J_k E_k(\eta)\|}{\|F_k\| - \|F_k + J_k E_k(\eta)\|} \\
&\geq \frac{(1-\eta)\|F_k\| - \frac{1-\alpha}{\Gamma} \cdot \|E_k(\eta)\|}{(1-\eta)\|F_k\|} \\
&\geq \frac{(1-\eta)\|F_k\| - (1-\alpha)(1-\eta)\|F_k\|}{(1-\eta)\|F_k\|} \\
&= \alpha.
\end{aligned}$$

因为 $\eta(\mu)$ 是 μ 的单调增函数，并且 $\lim_{\mu\to+\infty}\eta(\mu) = 1$, 再注意到 LM 策略中的 μ 将随着迭代步数的增大而增大，故循环定会在执行有限次后终止.

设 ρ_k 是循环停止时的值. 如果 $\rho_k = \vartheta$, 则结论显然成立. 如果 $\rho_k > \vartheta$, 则循环体至少已执行了两次. 令 ρ_k^- 是 ρ_k 的前一个值，并记

$$\mu_k^- = \rho_k^- \|F_k\|^\tau, \quad \eta_k^- = \eta(\mu_k^-).$$

则由前面的论证知，必有

$$\eta_k^- \leq 1 - \frac{\delta}{\Gamma\|F_k\|},$$

亦即有

$$\|F_k + J_k s_k(\mu_k^-)\| \leq \|F_k\| - \frac{\delta}{\Gamma}.$$

从而,

$$\begin{aligned}
\frac{\delta}{\Gamma} &\leq \|F_k\| - \|F_k + J_k s_k(\mu_k^-)\| \\
&\leq \|J_k s_k(\mu_k^-)\| \\
&\leq \|J_k\|\,\|s_k(\mu_k^-)\| \\
&= \|J_k\|\,\|W_k \left(W_k^T B_k W_k + \mu_k^- I\right)^{-1} W_k^T J_k^T F_k\| \\
&\leq \|J_k\|\,\|\left(W_k^T B_k W_k + \mu_k^- I\right)^{-1}\|\,\|W_k^T J_k^T F_k\| \\
&\leq \|J_k\|^2 \|F_k\| \cdot (\mu_k^-)^{-1}.
\end{aligned}$$

因此,

$$\mu_k^- \le \frac{\Gamma}{\delta}\|J_k\|^2\|F_k\|.$$

于是，我们可得

$$\rho_k^- \le \frac{\Gamma}{\delta}\|J_k\|^2\|F_k\|^{1-\tau}.$$

故

$$\rho_k \le \frac{\vartheta\Gamma}{\delta}\|J_k\|^2\|F_k\|^{1-\tau}.$$

至此，引理的结论得证. ∎

定理 3.3 在 $NGLM$ 方法中，取 $\epsilon_0 = 0$, 并设已由该方法产生了一个无穷序列 $\{x_k\} \subset \mathbb{R}^n$. 如果 $\{x_k\}$ 有一个无穷子列 $\{x_{k_i}\}$ 是由 LM 策略所产生的, 满足 $x_{k_i} \to x^*$ $(i \to +\infty)$ 且 $F'(x^*)$ 可逆，则 $F(x_k) \to 0$ $(k \to +\infty)$.

证明 首先，由定理 3.1 知

$$x_k \to x^* \ (k \to +\infty).$$

下面，我们进一步证明

$$\liminf \|g_k\| = 0. \tag{19}$$

假若此结论不成立，则存在 $\chi > 0$, 使得

$$\|g_k\| > \chi, \quad \forall\, k \ge 1.$$

令

$$æ = \max\{\|F'(x^*)\|, \|F'(x^*)^{-1}\|\},$$

并选取 $\delta_1 > 0$ 充分小，使得当 $x \in N_{\delta_1}(x^*)$ 时， $F'(x)$ 可逆并且成立

$$\|F'(x)\| \le 2\|F'(x^*)\|, \quad \|F'(x)^{-1}\| \le 2\|F'(x^*)^{-1}\|.$$

则由引理 3.5 知，存在与 k 无关的常数 $\Gamma > 0$, 使得当 $x_k \in N_{\delta_1}(x^*)$ 时，不等式 (18) 成立. 再选取 $\delta \in (0, \delta_1)$, 使得对于 $\forall\, x, y \in N_{2\delta}(x^*)$, 成立

$$\|F(y) - F(x) - F'(x)(y-x)\| \le \frac{1-\alpha}{\Gamma} \cdot \|y - x\|.$$

如果 $x_{k_i-1} \in N_\delta(x^*)$, 则一方面，由引理 3.7 知,

$$\begin{aligned}
\|W_{k_i-1}^T g_{k_i-1}\| &\ge \frac{1-\eta_{max}}{\kappa(F'(x_{k_i-1}))(1+\eta_{max})}\|g_{k_i-1}\| \\
&\ge \frac{1-\eta_{max}}{4æ^2(1+\eta_{max})}\|g_{k_i-1}\| \\
&\ge \frac{\chi(1-\eta_{max})}{4æ^2(1+\eta_{max})} \\
&:= C_1 > 0.
\end{aligned}$$

由于

$$\Delta_{k_i-1} = -W_{k_i-1}\left[W_{k_i-1}^T B_{k_i-1} W_{k_i-1} + \rho_{k_i-1}\|F_{k_i-1}\| I\right]^{-1} W_{k_i-1}^T g_{k_i-1} \to 0,$$

从而易得

$$\rho_{k_i-1} \to +\infty. \tag{20}$$

另一方面，又由引理 3.8 可知，

$$\begin{aligned}
\rho_{k_i-1} &\leq \max\left\{\vartheta, \quad \frac{\vartheta\Gamma}{\delta}\cdot\|F'(x_{k_i-1})\|^2\|F(x_{k_i-1})\|^{1-\tau}\right\} \\
&\leq \max\left\{\vartheta, \quad \frac{4\vartheta\Gamma æ^2}{\delta}\cdot\|F(x_0)\|^{1-\tau}\right\} \\
&:= C_2 > 0.
\end{aligned}$$

这与 (20) 相矛盾. 故 (19) 成立. 因为 $x_k \to x^*(k \to +\infty)$, 所以

$$\lim \|g_k\| = 0.$$

再注意到 $F'(x^*)$ 可逆，我们可立得 $F(x_k) \to 0\ (k \to +\infty)$. ∎

最后，基于定理 3.1, 定理 3.2 和定理 3.3, 我们即可得到 NGLM 方法的全局收敛性定理. 由于其证明类似于 [1] 中定理 4.4 的证明，故此略去.

定理 3.4 *在 NGLM 方法中，取 $\epsilon_0 = 0$, 并设已由该方法产生了一个无穷序列 $\{x_k\} \subset \mathbb{R}^n$. 如果 $\{x_k\}$ 有一个聚点 x^*, 使得 $F'(x^*)$ 可逆，则 $x_k \to x^*\ (k \to +\infty)$ 并且 $F(x^*) = 0$. 另外，当 k 充分大时，成立 $x_{k+1} = x_k + \bar{s}_k$.*

定理 3.4 表明， NGLM 方法的最终收敛速度完全取决于强制序列 $\{\bar{\eta}_k\}$.

§4 数值结果

我们通过比较 NGLM, NGQCGB, NGECB[3] 和 NGB 方法的计算结果来进一步说明 NGLM 方法的可行性，有效性和强健性.

§4.1 实验问题及有关参数

我们选取了文献中经常用到的 21 个非线性问题进行了数值实验. 这些问题都有各自的标准初值 x_s, 见表 1. 在表 1 中，我们还列出了问题的维数 n. 在这些问题中，既有坏条件的 (譬如，问题 P_1 和 P_2), 又有奇异的 (譬如，问题 P_{10}); 相对于 NGB 方法来说，它们中既有比较困难的 (譬如，问题 P_1, P_2, P_6, P_8-P_{10} 和 P_{16}-P_{21}), 也有比较容易的 (譬如，问题 P_3-P_5, P_7 和 P_{11}- P_{15}).

表 1: 测试问题

编号	问题名称	初始向量 x_s	维数 n
P_1	Augmented Powell Badly Scaled[13]	$(0,1,-4,0,1,-4,\ldots,)^T$	6000
P_2	Extended Powell Badly [18]	$(0,1,0,1,\ldots,)^T$	10000
P_3	Augmented Rosenbrock [13]	$(1.2,1,-1,20,1.2,1,-1,20,\ldots,)^T$	8000
P_4	Extended Rosenbrock [18]	$(-1.2,1,-1.2,1,\ldots,)^T$	8000
P_5	Generalized Rosenbrock($p=5$) [19]	$(1.2,1.2,\ldots,-1.2)^T$	5000
P_6	Modified Rosenbrock [13]	$(-1.8,-1,-1.8,-1,\ldots,)^T$	8000
P_7	Band Broyden($p=4$) [14]	$(-1,-1,\ldots,-1)^T$	3000
P_8	Broyden Tridiagonal Function [18]	$(-1,-1,\ldots,-1)^T$	3000
P_9	Broyden Tridiagonal Problem [18]	$(-1,-1,\ldots,-1)^T$	3000
P_{10}	Singular Broyden [18]	$(-1,-1,\ldots,-1)^T$	6000
P_{11}	Trigonometric-Exponential System1 [18]	$(0,0,\ldots,0)^T$	6000
P_{12}	Trigonometric-Exponential System2 [18]	$(0,0,\ldots,0)^T$	6000
P_{13}	Tri-diagonal System [18]	$(12,12,\ldots,12)^T$	6000
P_{14}	Five-diagonal System [18]	$(-2,-2,\ldots,-2)^T$	5000
P_{15}	Seven-diagonal System [18]	$(-3,-3,\ldots,-3)^T$	7000
P_{16}	Countercurrent Reactor [18]	$(0.1,0.2,0.3,0.4,0.5,0.4,0.3,0.2,\ldots,)^T$	8000
P_{17}	Extended Cragg and Levy Function [18]	$(1,2,2,2,1,2,2,2,\ldots,)^T$	4000
P_{18}	Structured Jacobian Problem [18]	$(-1,-1,\ldots,-1)^T$	5000
P_{19}	Discrete Integral Equation($c=0.999$) [18]	$(-1,-1,\ldots,-1)^T$	100
P_{20}	Tridimentional Valley [13]	$(-4,1,2,-4,1,2,\ldots,)^T$	6000
P_{21}	Trigonometric Function[23]	$(0,1,0,1\ldots,)^T$	300

我们取参数

$$\eta_{max} = 0.9, \quad \alpha = 10^{-4}, \quad \theta_l = 0.1, \quad \theta_u = 0.5, \quad \tau = 0.35$$

和

$$\vartheta = 2.$$

同时，我们取强制量 $\bar{\eta}_k$ 为

$$\bar{\eta}_k = \min\left\{\max\left\{0.9\left(\frac{\|F(x_k)\|}{\|F(x_{k-1})\|}\right)^2, 0.9\bar{\eta}_{k-1}^2\right\}, \eta_{max}\right\}.$$

需要说明的是，上述关于强制量的选取是 [10] 中所采用的一种策略. 对于 Jacobi 矩阵 $F'(x)$ 与某一向量 y 的乘积，我们采用差分公式 (4) 予以近似，其中取差分步长

$$\epsilon = \frac{\bar{\epsilon}\|x\|}{\|y\|}, \quad \bar{\epsilon} = 10^{-7},$$

见 [16].

在 GMRES 迭代中，我们没有采用重新启动策略，迭代初始向量选为 $s_k^0 = 0$, 最大迭代步数限制为 $m_{\max} = 40$, 而停机准则则取为

$$\max\left\{\frac{1}{\sqrt{n}}\|F(x_k)\|, \frac{\|F(x_k)\|}{\|F(x_0)\|}\right\} \leq 10^{-6}$$

或外迭代步数超过了最大允许非线性迭代步数 300. 若在进行了 $m = m_{max}$ 步 GMRES 迭代后，非精确 Newton 条件仍然不能得到满足，即有 $\|r_k^m\| > \bar{\eta}_k\|F(x_k)\|$, 则我们采用 [3] 中的策略，取

$$\begin{cases} \bar{s}_k & := & s_k^m, \\ \bar{\eta}_k & := & \frac{\|F'(x_k)\,\bar{s}_k + F(x_k)\|}{\|F(x_k)\|} \end{cases}$$

并结束 GMRES 过程. 在每步非线性迭代中，沿每条曲线的最大后退次数限定为 50. 在迭代过程中，若下列三种情形之一发生，则我们认为迭代失败：

F_1. 非线性迭代步数已经达到 $k_{\max} = 300$ 次，但停机准则还没有得到满足；

F_2. 在某一步非线性迭代中，后退次数 (不包括 NGLM, NGQCGB 和 NGECB 三种方法在非精确 Newton 方向 $\bar{s}_k$ 上的后退次数) 已经达到 50 次，但还没有得到充分下降步；

F_3. 在迭代过程中发生了 “停滞”(stagnation), 亦即 $|\,\|F(x_{k-1})\| - \|F(x_k)\|\,| \leq 10^{-6}\|F(x_k)\|$ [12, 15].

因为沿非精确 Newton 方向 $\bar{s}_k$ 上的允许最大后退次数 N_b 会影响到相应方法的数值效果, 所以就不同的 N_b 值, 我们也进行了数值实验. 为明确起见, 用 NGLM(N_b), NGQCGB(N_b) 和 NGECB(N_b) 来分别表示最大允许后退次数为 N_b 所对应的 NGLM, NGQCGB 和 NGECB 方法. 在所有计算中，我们均取 $N_b \leq 5$[3].

表 2: 21 个问题， 363 次计算结果统计

方法	成功次数	成功率	失败次数	stagnation 次数	退化次数
NGB	246	68%	117	105	81
NGECB(3)	267	74%	96	49	180
NGQCGB(3)	271	75%	92	79	180
NGLM(3)	296	82%	67	54	180

§4.2 实验结果与分析比较

对于每个测试问题，我们给定 21 个初值:

$$x_0 = \pm j x_s,\ x_0 = \pm j e,\ j = 1, 2, \ldots, 5,\ \text{和}\ x_0 = 0,$$

其中 e 表示所有分量均为 1 的向量， 0 表示零向量. 对于某些问题，这 21 个初值可能会出现重复，也可能恰好是该问题的解. 在剔除这些或者是问题的解或者会有重复的初值之后，我们将所有剩余的初值 (总共 363 个) 做为测试问题的初值. 为简便起见，我们称这些初值为有效初值.

对于 NGLM, NGQCGB 和 NGECB 方法, 我们用不同的 N_b 值进行了数值实验. 结果表明，当 $N_b \leq 5$ 时，它们均表现出了良好的数值效果. 据此，我们固定 $N_b = 3$, 然后对于每个问题的所有有效初值，分别用 NGLM(3), NGQCGB(3), NGECB(3) 和 NGB 方法进行计算. 统计结果见表 2.

对于 NGB 方法来说，表 2 中的最后一列 " 退化次数" 是指在所有成功的次数中没有运用后退策略的次数，亦即 NGB 方法退化为 Newton-GMRES 方法的次数. 而对于其它三种方法而言，表 2 中最后一列 " 退化次数" 则是指在所有成功的次数中没有用到各自的备选策略的次数，亦即它们分别退化为 NGB 方法的次数. 由于后三种方法的参数 N_b 均为 3, 故它们的退化次数是相同的.

从表 2 可以看出, 后三种方法的总成功次数均要比 NGB 方法多. 这说明 NGLM(3), NGQCGB(3) 和 NGECB(3) 方法较之 NGB 方法具有更好的强健性.

显然， NGLM 方法取得了最好的计算效果. 在对于所有 21 个测试问题的总共 363 个有效初值的实验中， NGLM(3) 方法成功了 296 次，成功率达到了 82%, 而 NGQCGB(3) 和 NGECB(3) 方法的成功率则分别为 75% 和 74%. 在 NGB 方法的 246 次成功迭代中，只有 81 次没有用到后退策略，也即是说，它有 81 次退化为 NG(Newton-GMRES) 方法. 而对于 NGLM(3), NGQCGB(3) 和 NGECB(3) 三种方法来说，它们各自都有 180 次退化为 NGB 方法. 从表 2 还可以看出，这四种方法的失败情形主要是由 "停滞" 现象所造成的.

表 3: 7 个问题 (P_1, P_2, P_3, P_6, P_{13}, P_{19}, P_{21}), 134 次计算结果统计

方法	成功次数	成功率	失败次数	stagnation 次数	退化次数
NGB	68	50%	66	62	20
NGECB(3)	90	67%	44	14	37
NGQCGB(3)	101	75%	33	31	37
NGLM(3)	117	87%	17	14	37

根据上述计算结果，我们选出了 7 个 NGB 方法表现较差的问题，并对这 7 个问题的计算结果重新进行了统计，结果列于表 3.

从表 3 可以看出，在 134 次数值实验中， NGB 方法只成功了 68 次，成功率仅为 50%; 而 NGECB(3), NGQCGB(3) 和 NGLM(3) 三种方法的成功率则均高于 50%, 分别达到了 67%, 75% 和 87%. 与表 2 相比，可以看到 NGB 方法的成功率有明显下降 (从 67% 下将为 50%). 这是由于表 3 的统计结果中包括两个坏条件问题 (P_1 和 P_2) 的计算结果，而 NGB 方法对于这两个坏条件问题的数值效果特别差而导致的.

从表 3 还可以看出，在 NGLM(3) 方法的所有 117 次成功情形，其中有 37 次没有用到 LM 策略，换句话说，其中有 80 次用到了 LM 策略. 同样地，在 NGQCGB(3) 方法的所有 101 次成功情形，其中有 64 次用到了 QCGB 策略；而在 NGECB(3) 方法的所有 90 次成功情形，其中有 53 次用到了 ECB 策略. 这说明每个方法的备选策略对于该方法的成功都起了关键作用. 另外，就改善 NGB 方法的强健性的程度来说， ECB 策略的效果显然没有 QCGB 和 LM 策略明显.

基于函数的赋值次数，以及算法的非线性迭代步数，全局策略的后退次数和备选策略的使用次数，我们对 NGECB, NGQCGB 和 NGLM 三种方法进行了定量比较，结果列于表 4– 表 6. 这里，我们用 NI 表示非线性迭代总步数， FE 表示函数的总赋值次数， BT 表示总后退次数， SW 表示 NGLM, NGQCGB 和 NGECB 方法中使用各自备选策略的总次数，而用 $*$ 表示迭代失败.

由表 4 可以看出，在对于问题 P_1 的所有 21 次实验中， NGLM(3) 方法成功了 17 次， NGQCGB(3) 方法成功了 15 次，而 NGECB(3) 方法只成功了 11 次. 这说明 NGECB(3) 方法的强健性弱于 NGQCGB(3) 和 NGLM(3) 方法. 从迭代步数来看， NGECB(3) 方法最多， NGLM(3) 方法次之，而 NGQCGB(3) 方法则最少. 但是，就函数的赋值次数来说， NGLM(3) 方法最少， NGQCGB(3) 方法次之，而 NGECB(3) 方法则最多. 因此， NGLM(3) 的计算开销最小，而 NGECB(3) 方法的计算开销最大.

由表 5 可以看出, 在对于问题 P_2 的所有 20 次实验中，NGLM(3) 和 NGQCGB(3)

表 4: **Augmented Powell badly scaled 的计算结果**

x_0	NGLM(3)				NGQCGB(3)				NGECB(3)			
	NI	FE	BT	SW	NI	FE	BT	SW	NI	FE	BT	SW
x_s	*	*	*	*	*	*	*	*	*	*	*	*
$2x_s$	117	764	320	40	116	932	394	45	133	1042	480	53
$3x_s$	113	758	324	41	110	908	388	45	127	1026	479	54
$4x_s$	113	770	333	42	105	887	384	45	119	966	453	51
$5x_s$	104	718	313	42	100	868	377	46	297	3011	1586	245
$-x_s$	56	459	252	8	128	1084	475	54	*	*	*	*
$-2x_s$	70	447	183	4	*	*	*	*	143	1117	516	58
$-3x_s$	*	*	*	*	*	*	*	*	*	*	*	*
$-4x_s$	*	*	*	*	74	446	174	10	*	*	*	*
$-5x_s$	58	362	160	5	125	1017	428	50	*	*	*	*
e	121	809	345	45	114	919	389	45	126	979	449	50
$2e$	119	784	331	42	115	926	393	45	*	*	*	*
$3e$	114	767	329	42	111	914	391	45	128	1024	477	53
$4e$	111	734	310	39	107	874	371	43	128	1030	483	56
$5e$	105	723	316	42	101	870	377	46	298	3013	1586	245
$-e$	129	1021	510	50	141	1284	612	65	*	*	*	*
$-2e$	*	*	*	*	*	*	*	*	*	*	*	*
$-3e$	96	995	618	26	*	*	*	*	*	*	*	*
$-4e$	60	379	176	7	138	1155	509	57	56	298	106	5
$-5e$	56	283	94	1	101	879	443	29	160	1298	627	76
0	114	811	372	44	*	*	*	*	*	*	*	*

表 5: **Extended Powell badly 的计算结果**

x_0	NGLM(3)				NGQCGB(3)				NGECB(3)			
	NI	FE	BT	SW	NI	FE	BT	SW	NI	FE	BT	SW
x_s	116	705	326	41	114	843	394	45	131	917	480	53
$2x_s$	114	700	326	41	112	838	394	45	129	912	480	53
$3x_s$	111	690	324	41	109	828	392	45	126	902	478	53
$4x_s$	108	680	322	41	106	818	390	45	123	892	476	53
$5x_s$	104	673	323	43	99	793	381	46	117	871	470	54
$-x_s$	69	447	239	12	72	594	319	22	*	*	*	*
$-2x_s$	133	953	516	51	66	457	221	15	*	*	*	*
$-3x_s$	132	944	508	51	133	1092	541	60	*	*	*	*
$-4x_s$	134	953	511	51	134	1095	541	60	*	*	*	*
$-5x_s$	131	886	458	47	135	1098	541	60	*	*	*	*
e	112	698	327	42	109	825	389	45	126	901	476	53
$2e$	115	702	326	41	113	840	394	45	130	914	480	53
$3e$	112	692	324	41	110	830	392	45	127	904	478	53
$4e$	109	682	322	41	107	820	390	45	124	894	476	53
$5e$	104	667	319	42	99	791	380	46	118	873	470	54
$-e$	61	327	145	7	132	1073	528	59	*	*	*	*
$-2e$	*	*	*	*	*	*	*	*	*	*	*	*
$-3e$	*	*	*	*	*	*	*	*	*	*	*	*
$-4e$	109	682	322	41	128	960	450	52	83	486	239	19
$-5e$	*	*	*	*	*	*	*	*	*	*	*	*

方法均成功了 17 次, 而 NGECB(3) 方法只成功了 11 次. 这进一步说明 NGLM(3) 和 NGQCGB(3) 方法具有比 NGQCGB(3) 方法更好的强健性. 同样地, 从迭代步数上来看, NGECB(3) 方法最多, NGLM(3) 方法次之, 而 NGQCGB(3) 方法则最少. 而从函数的赋值次数上来看, NGLM(3) 方法最少, NGQCGB(3) 方法次之, 而 NGECB(3) 方法则最多. 这进一步佐证了 NGLM(3) 方法的计算开销最小, 而 NGECB(3) 方法的计算开销最大.

需要指出的是, 对于问题 P_1 和 P_2, 我们运用 NGB 方法就表 5 中的所有初值做了计算. 结果发现, 只有问题 P_1 当取初值为 $-3e$ 时, NGB 方法获得了成功, 而对于其它所有情形, NGB 方法全都失败了.

表 6: Modified Rosenbrock 的计算结果

x_0	NGLM(1)				NGQCGB(1)				NGECB(1)			
	NI	FE	BT	SW	NI	FE	BT	SW	NI	FE	BT	SW
x_s	43	545	387	34	78	757	357	60	87	609	295	80
$2x_s$	66	762	519	53	174	2317	1352	155	212	1723	910	205
$3x_s$	95	1015	651	85	*	*	*	*	*	*	*	*
$4x_s$	192	1607	854	182	*	*	*	*	*	*	*	*
$5x_s$	*	*	*	*	*	*	*	*	*	*	*	*
$-x_s$	23	237	153	17	37	300	121	25	55	386	176	48
$-2x_s$	53	538	336	46	155	2031	1160	139	229	1908	1003	221
$-3x_s$	99	874	490	91	*	*	*	*	*	*	*	*
$-4x_s$	283	1967	845	276	*	*	*	*	*	*	*	*
$-5x_s$	*	*	*	*	*	*	*	*	*	*	*	*
e	3	9	0	0	3	9	0	0	3	9	0	0
$2e$	24	233	147	17	46	396	171	32	66	474	220	59
$3e$	42	435	277	35	108	1271	681	92	134	1056	531	126
$4e$	58	572	350	51	214	3042	1820	197	264	2226	1181	256
$5e$	82	749	433	74	*	*	*	*	*	*	*	*
$-e$	35	420	296	25	54	463	196	39	*	*	*	*
$-2e$	46	547	382	34	83	808	393	63	91	640	312	85
$-3e$	56	672	466	45	130	570	859	113	*	*	*	*
$-4e$	71	818	553	59	211	3036	1847	192	223	1824	967	215
$-5e$	87	953	622	76	*	*	*	*	*	*	*	*

我们还发现，当 $N_b \leq 5$ 时， NGLM(N_b), NGQCGB(N_b) 和 NGECB(N_b) 方法对于参数 N_b 的变化不太敏感. 为此，我们在表 6 中列出了 NGLM(1), NGQCGB(1) 和 NGECB(1) 方法对于问题 P_6 的计算结果，以便对照和比较. 由表 6 可见，在对于问题 P_6 的 20 次实验中， NGLM(1) 方法成功了 18 次， NGQCGB(1) 方法成功了 12 次，而 NGECB(1) 方法只成功了 10 次. 这进一步说明 NGLM(1) 方法的强健性最好，而 NGECB(1) 方法则最差. 另外，从迭代步数和函数赋值次数来看， NGLM(1) 方法均为最少，这表明 NGLM(1) 方法的计算开销也为最少.

§5 结论

对于大规模稀疏非线性方程组,我们将后退策略 (backtracking) 与 LM(Levenberg-Marquardt) 策略有机地组合,构造了一类新型的具有全局收敛性质的 Newton-GMRES 方法 —NGLM 方法. 这种方法既可以部分地克服线性模型不能很好逼近原问题的缺点,同时又能够保持 NGB 方法的所有优点. 数值结果表明,较之已有的 NGQCGB 和 NGECB 方法而言, NGLM 方法大大地提高了 NGB 方法的强健性,并且具有更高的计算效率. 因此, NGLM 方法是求解大型稀疏非线性方程组的一种可行,有效而强健的数值算法.

参考文献

[1] H.-B. An and Z.-Z. Bai, *A globally convergent Newton-GMRES method for large sparse systems of nonlinear equations*, Preprint, 2003.

[2] Z.-Z. Bai and P.-L. Tong, *On the affine invariant convergence theorems of inexact Newton method and Broyden's method*, J. Univ. Electr. Sci. Tech. China, 23:5(1994), 535-540. (In Chinese)

[3] S. Bellavia and B. Morini, *A globally convergent Newton-GMRES subspace method for systems of nonlinear equations*, SIAM J. Sci. Comput., 23 (2001), 940-960.

[4] S. Bellavia, M. Macconi and B. Morini, *A hybrid Newton-GMRES method for solving nonlinear equations*, Lecture Notes in Computer Science, L. Vulkov, J. Wasniewski and P. Yalamov eds., Vol. 1988, Springer-Verlag, 68-75, 2000.

[5] P.N. Brown and Y. Saad, *Hybrid Krylov methods for nonlinear systems of equations*, SIAM J. Sci. Statist. Comput., 11 (1990), 450-481.

[6] P.N. Brown and Y. Saad, *Convergence theory of nonlinear Newton-Krylov algorithms*, SIAM J. Optim., 4 (1994), 297-330.

[7] R.S. Dembo, S.C. Eisenstat and T. Steihaug, *Inexact Newton methods*, SIAM J. Numer. Anal., 19 (1982), 400-408.

[8] J.E. Dennis Jr. and R.B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, Englewood Cliffs, N.J., 1983.

[9] S.C. Eisenstat and H.F. Walker, *Globally convergent inexact Newton methods*, SIAM J. Optim., 4 (1994), 393-422.

[10] S.C. Eisenstat and H.F. Walker, *Choosing the forcing term in an inexact Newton method*, SIAM J. Sci. Comput., 17 (1996), 16-32.

[11] J.-Y. Fan and Y.-X. Yuan, *On the convergence of a new Levenberg-Marquardt method*, Research Report, AMSS, Chinese Academy of Science, 2001.

[12] D.R. Fokkema, G.L.G. Sleijpen and H.A. van der Vorst, *Accelerated inexact Newton schemes for large systems of nonlinear equations*, SIAM J. Sci. Comput., 19 (1998), 657-674.

[13] A. Friedlander, M.A. Gomes-Ruggiero, D.N. Kozakevich, J.M. Martínez and S.A. Santos, *Solving nonlinear systems of equations by means of quasi-Newton methods with a nonmonotone strategy*, Optim. Methods Software, 8(1997), 25-51.

[14] M.A. Gomes-Ruggiero, J.M. Martínez and A.C. Moretti, *Comparing algorithms for solving sparse nonlinear system of equations*, SIAM J. Sci. Comput., 13 (1992), 459-483.

[15] I.E. Kaporin and O. Axelsson, *On a class of nonlinear equation solvers based on the residual norm reduction over a sequence of affine subspaces*, SIAM J. Sci. Comput., 16 (1995), 228-249.

[16] C.T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*, SIAM, Philadelphia, 1995.

[17] D.A. Knoll and D.E. Keyes, *Jacobian-free Newton-Krylov methods: A survey of approaches and applications*, J. Comp. Phys., ??(2003), 1-41.

[18] L. Luksan, *Inexact trust region method for large sparse systems of nonlinear equations*, J. Optim. Theory Appl., 81 (1994), 569-591.

[19] N. Krejić and Z. Lužanin, *Newton-like method with modification of the right-hand-side vector*, Math. Comput., 71 (2001), 237-250.

[20] J.M. Ortega and W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[21] M. Pernice and H.F. Walker, *NITSOL: A Newton iterative sovler for nonlinear systems*, SIAM J. Sci. Comput., 19 (1998), 302-318.

[22] Y. Saad and M.H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Comput., 7 (1986), 856-869.

[23] L.C. William and R. Marcos, *Nonmonotone spectral methods for large-scale nonlinear systems*, Preprint, 2003.