

利用固定矩阵计算亏损矩阵的幂级数之和

The Sum Of the Defective Matrix Power Series by the Regular Matrix

李大林

Li Dalin

(柳州职业技术学院基础部 柳州 545006)

(Dept. for Basic Courses, Liuzhou Vocational & Tech. Coll., Liuzhou, Guangxi, 545006, China)

摘要 通过方阵 A 的极小多项式 $h(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_s)^{n_s}$ 的指数来定义可变系数向量 $V(m)$, 并构成 A 的固定矩阵 D . 利用固定矩阵 D , 将计算亏损矩阵的幂级数公式 $A^m = PJ^mP^{-1}$ 改进为 $A^m = V(m)D^{-1}(E, A, \cdots, A^{w-1})^T$, 免去求若当链及 P^{-1} 的步骤.

关键词 亏损矩阵 幂级数 若当链 固定矩阵

中图分类号 O121.21

Abstract Define the transformable entries vector $V(m)$ on the basis of index of the minimal polynomial $h(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_s)^{n_s}$, and make up the regular matrix D . The classic formula $A^m = PJ^mP^{-1}$ is improved to $A^m = V(m)D^{-1}(E, A, \cdots, A^{w-1})^T$. Accordingly, the calculation of the Jordan chains and P^{-1} can be avoided.

Key words defective matrix, power series, Jordan chain, regular matrix

在工程技术中,常用到阶数很大的方阵的函数或矩阵幂级数,常规处理方法是将方阵化为对角形.如果该方阵为亏损矩阵,就要求它的若当链及过渡矩阵.实际操作难度很大^[1,2],本文试图解决这一问题.

设复数域上的 n 阶亏损矩阵 A 的极小多项式为 $h(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_s)^{n_s}$, 其中 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 为互异的特征根. $h(\lambda)$ 的次数记为 $w = \sum_{i=1}^s n_i$. A

的若当标准形记为 $J = \text{diag}(J_1, J_2, \cdots, J_s)$, 其中 $J_i = \text{diag}(J_{i1}(\lambda_i), J_{i2}(\lambda_i), \cdots, J_{in_i}(\lambda_i))$ 为 λ_i 对应的若当矩阵, 其中若当块

$$J_{ij}(\lambda_i) = \begin{bmatrix} \lambda_i & & & \\ & 1 & & \\ & & \cdots & \\ & & & \cdots & \\ & & & & 1 & \\ & & & & & \lambda_i \end{bmatrix}_{n_i(j) \times n_i(j)},$$

$$J_{ij}^m(\lambda_i) = \begin{bmatrix} \lambda_i^m & & & & & \\ C_m^1 \lambda_i^{m-1} & & & & \lambda_i^m & \\ C_m^2 \lambda_i^{m-2} & & & & C_m^1 \lambda_i^{m-1} & \\ C_m^3 \lambda_i^{m-3} & & & & C_m^2 \lambda_i^{m-2} & \cdots \\ \cdots & & & & \cdots & \cdots \\ C_m^{n_i(j)-1} \lambda_i^{m-n_i(j)+1} & & & & C_m^{n_i(j)-2} \lambda_i^{m-n_i(j)+2} & \cdots & \cdots & \cdots & \cdots & \cdots \\ C_m^2 \lambda_i^{m-2} & C_m^1 \lambda_i^{m-1} & \lambda_i^m \end{bmatrix}. \quad (1)$$

$J_{ij}(\lambda_i)$ 的阶记为 $n_i(j)$. 不妨设 $n_i(1) \leq \cdots \leq n_i(j) \leq \cdots \leq n_i(t(i))$. 显然 $n_i(t(i)) = n_i$. 由文献 [1], 可找到特征向量与广义特征向量构成过渡矩阵 P , 使 $A = PJP^{-1}$. 设

$$P = (P_1, P_2, \cdots, P_s), \text{ 其中 } P_i = (P_{i1}, P_{i2}, \cdots, P_{in_i}), P_{ij} = (P_{ij1}, P_{ij2}, \cdots, P_{ijn_i(j)}),$$

$$J_{ij}(\lambda_i) \text{ 对应的若当链链长为 } n_i(j), \text{ 满足}$$

$$AP_{ij1} = \lambda_i P_{ij1} + P_{ij2}, AP_{ij2} = \lambda_i P_{ij2} + P_{ij3}, \cdots,$$

$$AP_{ij, n_i(j)-1} = \lambda_i P_{ij, n_i(j)-1} + P_{ij, n_i(j)}, AP_{ij, n_i(j)} = \lambda_i P_{ij, n_i(j)}.$$

显然每条若当链的最后一个向量 $P_{ij, n_i(j)}$ 为特征向量, 其余为广义特征向量. 计算 A 的幂的公式为^[3]

$$A^m = P \text{diag}(J_1^m, J_2^m, \cdots, J_s^m) P^{-1}, \text{ 其中 } J_i^m = \text{diag}(J_{i1}^m(\lambda_i), J_{i2}^m(\lambda_i), \cdots, J_{in_i}^m(\lambda_i)), \text{ 其中 } J_{ij}^m(\lambda_i) \text{ 如 (1) 式所示.}$$

用(1)式求 \$A\$ 的幂,必须先求出 \$n\$ 个线性无关的特征向量.广义特征向量,即找到足够的若当链构成 \$P\$,再求 \$P^{-1}\$.该方法计算量很大,即使知道 \$A\$ 的标准形,要找到足够的若当链也不容易\$^{[2]}\$.本文试图寻找一种既能利用若当块 \$J_{ij}^m(\lambda_i)\$ 的元素的规律性,又能避免求若当链的方法.

对于给定的过渡矩阵 \$P\$,令 \$P^{-1} = Q = (Q_1, Q_2, \dots, Q_s)^T\$, 其中 \$Q = (Q_1, Q_2, \dots, Q_{t(i)})^T, Q_j = (Q_{j1}, Q_{j2}, \dots, Q_{jn_j(j)})^T\$, 则

$$A^m = P \operatorname{diag}(J_1^m, J_2^m, \dots, J_s^m) P^{-1} = \sum_{i=1}^s P_i J_i^m Q = \sum_{i=1}^s (P_{i1}, P_{i2}, \dots, P_{it(i)}) \operatorname{diag}(J_{i1}^m(\lambda_i), J_{i2}^m(\lambda_i), \dots, J_{it(i)}^m(\lambda_i)) (Q_1, Q_2, \dots, Q_{t(i)})^T = \sum_{i=1}^s \sum_{j=1}^{t(i)} P_{ij} J_{ij}^m(\lambda_i) Q_j = \sum_{i=1}^s \sum_{j=1}^{t(i)} \sum_{k=1}^{n_i(j)} \sum_{h=0}^m C_m^h \lambda_i^{m-h} P_{ij, k+h} Q_{jk} = \sum_{i=1}^s \sum_{h=0}^m C_m^h \lambda_i^{m-h} \left(\sum_{j=1}^{t(i)} \sum_{k=1}^{n_i(j)} P_{ij, k+h} Q_{jk} \right), \quad (2)$$

令 \$A^{(h)} = \sum_{j=1}^{t(i)} \sum_{k=1}^{n_i(j)} P_{ij, k+h} Q_{jk}\$, 其中 \$h > m\$ 时, \$C_m^h = 0\$. 当 \$k+h > n_i(j)\$ 时, \$P_{ij, k+h}\$ 不存在, 故 \$P_{ij, k+h} Q_{jk}\$ 不存在. 又因 \$n(1) \leq n(2) \leq \dots \leq n(t(i)) = n_i\$, 故最后一个为 \$A^{(n_i-1)} = \sum_{j=1}^{t(i)} P_{ij, k+(n_i-1)} Q_{j1}\$, 其中 \$n_i(a) = n_i\$ 而 \$n_i(a-1) < n_i\$. 称 \$A_i^{(0)}, A_i^{(1)}, \dots, A_i^{(n_i-1)}\$ 为 \$A\$ 的属于 \$\lambda_i\$ 的广义特征矩阵. 由(2)式, 有

$$D = \begin{bmatrix} 1 & 0 & \dots & 0 & 1 & 0 & \dots & 0 & \dots & 1 & 0 & \dots & 0 \\ \lambda_1 & 1 & \dots & 0 & \lambda_2 & 1 & \dots & 0 & \dots & \lambda_s & 1 & \dots & 0 \\ \lambda_1^2 & C_1^1 \lambda_1 & \dots & C_2^{n_1-1} \lambda_1^{3-n_1} & \lambda_2^2 & C_2^1 \lambda_2 & \dots & C_2^{n_2-1} \lambda_2^{3-n_2} & \dots & \lambda_s^2 & C_2^1 \lambda_s & \dots & C_2^{n_s-1} \lambda_s^{3-n_s} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \lambda_1^{w-1} & C_1^{w-1} \lambda_1^{w-2} & \dots & C_1^{n_1-1} \lambda_1^{w-n_1} & \lambda_2^{w-1} & C_2^{w-1} \lambda_2^{w-2} & \dots & C_2^{n_2-1} \lambda_2^{w-n_2} & \dots & \lambda_s^{w-1} & C_2^{w-1} \lambda_s^{w-2} & \dots & C_2^{n_s-1} \lambda_s^{w-n_s} \end{bmatrix} \quad (6)$$

式(6)可表示为 \$D = (V(0), V(1), V(2), \dots, V(w-1))^T\$, 称 \$D\$ 为 \$A\$ 的固定矩阵.

命题 1 \$A\$ 的固定矩阵 \$D\$ 可逆.
证明 把 \$D\$ 中的列向量第 1 个分量为 1 的作为第 1 组, 第 1 个分量为 0 而第 2 个分量为 1 的作为第 2 组, 依此类推, 因为 \$n_1 \leq n_2 \leq \dots \leq n_s\$, 所以最后一组(即第 \$n_s\$ 组)的每个向量前 \$n_s - 1\$ 个分量为 0, 而第 \$n_s\$ 个分量为 1. 显然不同两组的列向量是线性无关的. 下证每一组的列向量也线性无关. 对于第 \$i\$ 组, 设 \$n_j \geq i\$ 而 \$n_{j-1} < i\$, 则第 \$i\$ 个矩阵 \$D_i\$ 为 \$s-j+1\$ 阶方阵, 从第 \$i\$ 组每个列向量的第 \$i\$ 个元素(该元素为 1)取起,

定理 1 设 \$\lambda_1, \lambda_2, \dots, \lambda_s\$ 为 \$n\$ 阶亏损矩阵 \$A\$ 的互异的特征根, \$A_i^{(0)}, A_i^{(1)}, \dots, A_i^{(n_i-1)}\$ 为 \$A\$ 的属于 \$\lambda_i\$ 的广义特征矩阵, \$m \in N\$, 则

$$A^m = \sum_{i=1}^s \sum_{h=0}^m C_m^h \lambda_i^{m-h} A_i^{(h)}, \quad (3)$$

推论 1 \$\sum_{i=1}^s A_i^{(0)} = E\$. \quad (4)

证明 \$\sum_{i=1}^s A^{(0)} = \sum_{i=1}^s \sum_{j=1}^{t(i)} \sum_{k=1}^{n_i(j)} P_{ijk} Q_{jk} = \sum_{i=1}^s \sum_{j=1}^{t(i)} P_{ij} Q_j = \sum_{i=1}^s P_i Q = P Q = E\$. 证毕.

分别令 \$m = 0, 1, \dots, w-1\$, 代入(3)式(当 \$m = 0\$ 时 \$A^{(0)} = E\$, 见(4)式), 有方程组

$$\sum_{i=1}^s A_i^{(0)} = E, \quad \sum_{i=1}^s (\lambda_i A_i^{(0)} + A_i^{(1)}) = A, \dots, \sum_{i=1}^s \sum_{h=0}^{w-1} C_m^h \lambda_i^{w-1-h} A_i^{(h)} = A^{w-1}, \quad (5)$$

记 \$U = (A_1^{(0)}, A_1^{(1)}, \dots, A_1^{(n_1-1)}, A_2^{(0)}, A_2^{(1)}, \dots, A_2^{(n_2-1)}, \dots, A_s^{(0)}, A_s^{(1)}, \dots, A_s^{(n_s-1)})^T, V(m) = (C_m^0 \lambda_1^m, C_m^1 \lambda_1^{m-1}, \dots, C_m^{n_1-1} \lambda_1^{m-n_1+1}, C_m^0 \lambda_2^m, C_m^1 \lambda_2^{m-1}, \dots, C_m^{n_2-1} \lambda_2^{m-n_2+1}, \dots, C_m^0 \lambda_s^m, C_m^1 \lambda_s^{m-1}, \dots, C_m^{n_s-1} \lambda_s^{m-n_s+1})\$ (其中当 \$h > m\$ 时 \$C_m^h = 0\$), 则(3)式可表示为 \$A^m = V(m)U\$. 显然 \$V(m)\$ 是 \$A^m\$ 的 \$m\$ 次及 \$A\$ 的极小多项式 \$h(\lambda) = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{n_2} \dots (\lambda - \lambda_s)^{n_s}\$ 的指数 \$m_1, m_2, \dots, m_s\$ 的函数. 称 \$V(m)\$ 为 \$A\$ 的第 \$m\$ 个可变系数向量. 方程组(5)的系数矩阵如(6)式所示.

至第 \$i+s-j\$ 个元素止, 构成方阵 \$D_i\$. 其行列式

$$|D_i| = \begin{vmatrix} 1 & 1 & \dots & 1 \\ C_{i+1}^i \lambda_j & C_{i+1}^i \lambda_{j+1} & \dots & C_{i+1}^i \lambda_s \\ \dots & \dots & \dots & \dots \\ C_{i+s-j}^i \lambda_j^{s-j} & C_{i+s-j}^i \lambda_{j+1}^{s-j-1} & \dots & C_{i+s-j}^i \lambda_s^{s-j} \end{vmatrix} = C_{i+1}^{i-1} \dots C_{i+s-j}^i \begin{vmatrix} 1 & 1 & \dots & 1 \\ \lambda_j & \lambda_{j+1} & \dots & \lambda_s \\ \dots & \dots & \dots & \dots \\ \lambda_j^{s-j} & \lambda_{j+1}^{s-j-1} & \dots & \lambda_s^{s-j} \end{vmatrix}.$$

因为 \$\lambda_j, \lambda_{j+1}, \dots, \lambda_s\$ 互异, 故范德蒙行列式不为 0,

故 $|D_i| \neq 0$, 从而 D_i 可逆. 从而第 i 组列向量线性无关 ($i = 1, 2, \dots, s$). 故 D 的所有列向量线性无关, 从而 D 可逆. 证毕.

推论 2 方程组 (5) 有唯一解.

证明 由命题 1, 方程组 (5) 的系数矩阵 D 可逆, 故 (5) 有唯一解. 证毕.

定理 2 设 D 为 A 的固定矩阵, $m \in N$, 则 $A^m = V(m)D^{-1}(E, A, \dots, A^{w-1})^T$. (7)

证明 因为 $A^m = V(m)U$, 故方程组 (5) 可表示为 $DU = (E, A, \dots, A^{w-1})^T$, 从而 $U = D^{-1}(E, A, \dots, A^{w-1})^T$. 故 $A^m = V(m)U = V(m)D^{-1}(E, A, \dots, A^{w-1})^T$. 证毕.

用 (7) 求一系列幂 A^k ($k = w, w+1, \dots, m, \dots$) 时, 先求 $D^{-1}(E, A, \dots, A^{w-1})^T$, 再逐个求 $V(k)D^{-1}(E, A, \dots, A^{w-1})^T$ 即可. 在实际计算中, 也可直接利用初等行变换解方程组 (5), 即可用 E, A, \dots, A^{w-1} 表示各个 $A_i^{(k)}$. 再代入 (3) 式计算幂 A^k .

注意, 这里 D^{-1} 为 w 阶方阵, 而 $(E, A, \dots, A^{w-1})^T$ 看成一个由 w 个文字 E, A, \dots, A^{w-1} 组成的列向量, 不是分块矩阵. 事实上, 算出 $D^{-1}(E, A, \dots, A^{w-1})^T$, 就得到了 U . 本文推导 U 中的一系列广义特征矩阵时, 用到过渡矩阵 P , 也即用到特征向量与广义特征向量, 但由 $D^{-1}(E, A, \dots, A^{w-1})^T = U$ 知, 求 U 已不再需特征向量与广义特征向量了, 仅需列出固定矩阵 D , 再求 D^{-1} . 从而避免求若当链及 P^{-1} .

特别地, 当 A 为非亏损矩阵时, $n_1 = \dots = n_s = 1$, $V(k) = (\lambda_1^k, \lambda_2^k, \dots, \lambda_s^k)$, 因而固定矩阵 D 简化为范德蒙矩阵, 对这一特殊情形的结论^[5] 已用哈密顿-凯莱定理导出.

下面是本文给出的求矩阵幂的固定矩阵算法: 第一步: 求 A 的极小多项式; 第二步: 写出可变系数向量 $V(0), V(1), \dots, V(w-1)$, 构成固定矩阵 D ; 第三步: 求 D^{-1} ; 第四步: 用 (7) 式求 A^m .

命题 2 交换 $V(m)$ 中的分量的位置, 公式 (7) 仍然成立.

证明 交换 $V(m)$ 中第 i 个与第 j 个分量的位置相当于 $V(m)$ 右乘一个初等矩阵 $P(i, j)$, 即 $V(m)P(i, j)$. D 中第 i 列与第 j 列随之交换位置, 成为 $DP(i, j)$. 故

$$(V(m)P(i, j))(DP(i, j))^{-1} = V(m)P(i, j)P(i, j)^{-1}D^{-1} = V(m)D^{-1}. \text{证毕.}$$

定理 3 设 $f(t) = \sum_{k=0}^{\infty} a_k t^k$, 它的收敛半径为 r , λ_i 为 A 的特征根 ($i = 1, 2, \dots, s$), $|\lambda_i| < r$, 则矩阵幂级

数 $\sum_{k=0}^{\infty} a_k A^k$ 收敛, 且

$$\sum_{k=0}^{\infty} a_k A^k = (f(\lambda_1), \frac{f'(\lambda_1)}{\mathbf{1}}, \dots, \frac{f^{(n_1-1)}(\lambda_1)}{(n_1-1)!}, f(\lambda_2), \frac{f'(\lambda_2)}{\mathbf{1}}, \dots, \frac{f^{(n_2-1)}(\lambda_2)}{(n_2-1)!}, \dots, f(\lambda_s), \frac{f'(\lambda_s)}{\mathbf{1}}, \dots, \frac{f^{(n_s-1)}(\lambda_s)}{(n_s-1)!})D^{-1}(E, A, \dots, A^{w-1})^T.$$

证明 由 Lagrange-Sylvester 定理^[1] $\sum_{k=0}^{\infty} a_k A^k$ 收敛. 由定理 2,

$$\begin{aligned} \sum_{k=0}^{\infty} a_k A^k &= \sum_{k=0}^{\infty} a_k V(k)D^{-1}(E, A, \dots, A^{w-1})^T = \\ &= (\sum_{k=0}^{\infty} a_k V(k))D^{-1}(E, A, \dots, A^{w-1})^T = (\sum_{k=0}^{\infty} a_k C_k^0 \lambda_1^k, \\ &\sum_{k=0}^{\infty} a_k C_k^1 \lambda_1^{k-1}, \dots, \sum_{k=0}^{\infty} a_k C_k^{n_1-1} \lambda_1^{k-n_1+1}, \sum_{k=0}^{\infty} a_k C_k^0 \lambda_2^k, \dots, \\ &\sum_{k=0}^{\infty} a_k C_k^{n_s-1} \lambda_s^{k-n_s+1})D^{-1}(E, A, \dots, A^{w-1})^T = (f(\lambda_1), \\ &\frac{f'(\lambda_1)}{\mathbf{1}}, \dots, \frac{f^{(n_1-1)}(\lambda_1)}{(n_1-1)!}, f(\lambda_2), \frac{f'(\lambda_2)}{\mathbf{1}}, \dots, \frac{f^{(n_2-1)}(\lambda_2)}{(n_2-1)!}, \\ &\dots, f(\lambda_s), \frac{f'(\lambda_s)}{\mathbf{1}}, \dots, \frac{f^{(n_s-1)}(\lambda_s)}{(n_s-1)!})D^{-1}(E, A, \dots, A^{w-1})^T. \text{证毕.} \end{aligned}$$

定理 3 与著名的 Lagrange-Sylvester 定理等价, 但前者比后者实用, 因为无需求特征向量与广义特征向量. 定理 3 降低了^[1, 3, 4] 中方阵函数与矩阵幂级数的计算难度.

例 1 设 $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & -1 & -1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$, 求 $\sin A$.

解 易得 A 的极小多项式为 $h(\lambda) = (\lambda - 1)(\lambda - 2)^2$, $\lambda_1 = 1, \lambda_2 = 2$. $f(t) = \sin t = \sum_{k=1}^{+\infty} \frac{(-1)^{k-1}}{(2k-1)!} t^{2k-1}$ 的收敛半径为 $r = +\infty, |\lambda_1|, |\lambda_2| < +\infty$.

固定矩阵 $D = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ \mathbf{1}^2 & \mathbf{2}^2 & \mathbf{C}_2^2 \times \mathbf{2} \end{bmatrix}$,
 $D^{-1} = \begin{bmatrix} 4 & -4 & \mathbf{1} \\ -3 & 4 & -\mathbf{1} \\ 2 & -3 & \mathbf{1} \end{bmatrix}$, 由定理 3,
 $\sin A = (f(1), f(2), \frac{f'(2)}{\mathbf{1}})D^{-1}(E, A, A^2)^T =$
 $(\sin 1, \sin 2, \cos 2) \begin{bmatrix} 4 & -4 & \mathbf{1} \\ -3 & 4 & -\mathbf{1} \\ 2 & -3 & \mathbf{1} \end{bmatrix} \begin{bmatrix} E \\ A \\ A^2 \end{bmatrix} =$
 $= \sin 1(4E - 4A + A^2) + \sin 2(-3E + 4A - A^2) +$

参考文献

- 1 陈啟宗著. 线性系统理论与设计 [美]. 王纪文等译. 北京: 科学出版社, 1988. 29~ 45.
- 2 刘升德. 根子空间的几何理论与演化矩阵的实际计算. 阜新矿业学院学报, 1994, (4): 111~ 114.
- 3 张贤科, 许甫华. 高等代数学. 北京: 清华大学出版社, 1998. 172~ 173, 189~ 199.
- 4 陈大新. 矩阵理论. 上海: 上海交通大学出版社, 1991. 132~ 133, 142~ 143.
- 5 李德光. 求 n 阶 m 次方幂的一个公式. 湘潭矿业学院报, 1994, (4): 80~ 81.

(责任编辑: 黎贞崇)

$$\cos 2(2E - 3A + A^2) =$$

$$\sin \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \sin 2 \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} +$$

$$\cos 2 \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & -1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} =$$

$$\begin{bmatrix} \sin 1 & 0 & 0 & 0 \\ 0 & \sin 2 + \cos 2 & -\cos 2 & -\cos 2 \\ 0 & \cos 2 & \sin 2 - \cos 2 & -\cos 2 \\ 0 & 0 & 0 & \sin 2 \end{bmatrix}.$$

(上接第 25 页 Continue from page 257)

- 10 Perry J M. A class of conjugate algorithms with a two step variable metric memory, Discussion paper 269, Center for Mathematical Studies in Economics and Management Science Northwestern Northwestern University, 1977.
- 11 Powell M J D. A new algorithm for unconstrained optimization, In Nonlinear Programming, J B Rosen, O L Mangasarian et al eds. New York Academic Press, 1970.
- 12 Shanno D F. On the convergence of a new conjugate gradient algorithm. SIAM Journal on Numerical Analysis, 1978, (15): 1247~ 1257.
- 13 Wei Z, Qi L, Chen X. An SQP-type method and its application in stochastic programming. Journal of Optimization Theory and Applications, 2003, (116): 205~ 228.
- 14 Wei Z, Li G, Qi L. New Quasi-Newton methods for unconstrained optimization problems. Nanning Department of

- Mathematics and Information Science, Guangxi University, Guangxi, China, 2002.
- 15 Wei Z, Yu G, Yuan G et al. The superlinear convergence of a modified BFGS-type method for unconstrained optimization, to appear in Computational Optimization and Applications.
- 16 Dai Y, Ni Q. Testing different conjugate gradient methods for large-scale unconstrained optimization. Journal of Comput Math, 2003, 21(3): 311~ 320.
- 17 Griewank A. On automatic differentiation, In Mathematical Programming Recent Developments and Applications, M Iri, K Tanabe, eds. Kluwer Academic Publishers, 1989, 84~ 108.

(责任编辑: 蒋汉明 黎贞崇)