第 12 卷第 2 期
1995 年 4 月

控制理论与应用
CONTROL THEORY AND APPLICATIONS

Vol. 12, No. 2
Apr., 1995

线性大系统参数估计的目标补偿协调法

何建敏　达庆利
（东南大学经济管理学院·南京，210018）

摘要：为克服线性大系统参数估计问题的维数灾难，本文提出一种目标补偿协调算法．它的主要特点是不用 Lagrange 乘子作协调变量，而通过对子问题的目标引入补偿项来协调它们，从而简化了协调结构．理论分析和实例计算表明，它的收敛性较好，适宜在多处理机系统上实现，优于整体算法和文献[1]的平行分解算法．

关键词：线性大系统；分解-协调；最小二乘估计；目标补偿；递阶优化

1 引　言

考察下列线性大系统模型

$$y = X\theta + u. \tag{1}$$

式中，$y \in \mathrm{R}^m$ 为输出向量，$\theta \in \mathrm{R}^n$ 为待估计的模型参数向量；$u \in \mathrm{R}^m$ 为观测噪声向量，一般设为零均值 Gauss 分布，其中各元素同方差且互不相关，即：$E(u) = 0$、$E(uu^{\mathrm{T}}) = \sigma^2 I_m$，$X \in \mathrm{R}^{m\times n}$ 为满秩观测矩阵．模型(1)参数 θ 的最小二乘估计问题为

$$\min_{\hat{\theta}} \frac{1}{2}\hat{u}^{\mathrm{T}}\hat{u} = \min_{\hat{\theta}} \frac{1}{2}(y - X\hat{\theta})^{\mathrm{T}}(y - X\hat{\theta}), \tag{2}$$

其解由大家熟知的公式给出

$$\hat{\theta}^* = (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}y. \tag{3}$$

由于待估计参数 θ 的维数较高，直接用(3)式计算其计算量和存贮量均较大，常常不能满足实际要求．为避免这一不足，文献[1]将(3)式平行分解，得到了适宜在多处理机系统上求解的平行分解算法．实例计算表明，这一算法优于(3)式给出的整体算法．本文从另一角度出发，运用大系统分解-协调原理，将(2)分解为一系列最小二乘估计子问题．为了降低协调变量维数，简化协调结构，且使分解得到的各子问题仍为线性最小二乘形式，本文不用一般的目标协调法，而通过对各子问题目标分别嵌入二次补偿项进行协调．然后，合理确定二次补偿项中的待定参数以保证所得算法的解与(2)的解相等价．理论分析和实例计算结果表明，本文算法的收敛性和收敛速度均优于文献[1]的平行分解法．

2 目标补偿协调算法

在(2)中引入预估参数向量 $\hat{\theta}^0$ 和相应的约束 $\hat{\theta}^0 = \hat{\theta}$，将 $\hat{\theta}^0$ 及 $\hat{\theta}$ 分解成 N 个子和量：$\hat{\theta}^0 = [\hat{\theta}_1^{0\mathrm{T}}, \cdots, \hat{\theta}_N^{0\mathrm{T}}]^{\mathrm{T}}$；$\hat{\theta} = [\hat{\theta}_1^{\mathrm{T}}, \cdots, \hat{\theta}_N^{\mathrm{T}}]^{\mathrm{T}}$，$\hat{\theta}_i^0, \hat{\theta}_i, \in \mathrm{R}^{n_i}$；$\sum_{i=1}^{N} n_i = n$，输出向量 y 分解为：$y = [y_1^{\mathrm{T}}, \cdots, y_N^{\mathrm{T}}]$；$y_i \in \mathrm{R}^{m_i}$；$\sum_{i=1}^{N} m_i = m$，并相应地将矩阵 X 分块为

本文于 1993 年 12 月 28 日收到．1994 年 7 月 11 日收到修改稿．

$$X=\begin{bmatrix} X_{11} & \cdots & X_{1N} \\ \vdots & \cdots & \vdots \\ X_{N1} & \cdots X_{NN} & \end{bmatrix};\quad X_{ij}\in \mathrm{R}^{m_i\times n_j};\quad i,j=1,\cdots,N.$$

这里，可按系统结构或使 $\hat{\theta}_i^0(\hat{\theta}_i)$ 有适当的维数来选择 n_i，并参照 n_i 选择 m_i 使估计 $\hat{\theta}_i$ 有适当的样本点. 这样，问题(2) 等价为

$$\min_{\hat{\theta}} \frac{1}{2}\sum_{i=1}^{N}(y_i-\sum_{\substack{r=1\\r\neq i}}^{N}X_{ir}\hat{\theta}_r^0-X_{ii}\hat{\theta}_i)^{\mathrm{T}}(y_i-\sum_{\substack{r=1\\r\neq i}}^{N}X_{ir}\hat{\theta}_r^0-X_{ii}\hat{\theta}_r),$$

$$\text{s.t.}\quad \hat{\theta}_1^0=\hat{\theta}_i, i=1,\cdots,N. \tag{4}$$

现选 $\hat{\theta}_i^0(i=1,\cdots,N)$ 为协调变量，它们的值由协调级给定，可得第一级的 N 个子最小二乘估计问题

$$\text{spi}\quad \min_{\hat{\theta}_i}\frac{1}{2}\{(y_i-\sum_{r\neq i}X_{ir}\hat{\theta}_r^0-X_{ii}\hat{\theta}_i)^{\mathrm{T}}(y_i-\sum_{r\neq 1}X_{ir}\hat{\theta}_r^0-X_{ii}\hat{\theta})+(\hat{\theta}_i-b_i^0)^{\mathrm{T}}(\hat{\theta}_i-b_i^0)\}$$

$$=\min_{\hat{\theta}_i}{}_i J_i,\quad \text{s.t.}\quad \hat{\theta}_r^0 \text{ 由协调级给定}, r=1,\cdots,N,\ r\neq 1,\ i=1,\cdots,N.$$

这里，二次项 $\hat{\theta}_i-b_i^0)^{\mathrm{T}}(\hat{\theta}_i-b_i^0)$ 即为目标补偿协调项，b_1^0 为待定的参数向量. (spi) 的最优性条件为

$$\partial Ji/\partial\hat{\theta}_i=-X_{ii}^{\mathrm{T}}(y_i-\sum_{r\neq i}X_{ir}\hat{\theta}_i^0-X_{ii}\hat{\theta}_i)+\hat{\theta}_i-b_i^0=0, \tag{5}$$

而(2) 关于 $\hat{\theta}_i$ 的最优性条件为

$$-\sum_{r=1}^{N}X_{ri}^{\mathrm{T}}(y_r-\sum_{k=1}^{N}X_{rk}\hat{\theta}_k)=0, \tag{6}$$

为使协调算法收敛时，$(\hat{\theta}^0=\hat{\theta})$ 两者等价，取

$$b_i^0=\hat{\theta}_i^0+\sum_{r\neq i}X_{ri}^{\mathrm{T}}(y_r-\sum_{k=1}^{N}X_{rk}\hat{\theta}_k^0). \tag{7}$$

据(5)，得(spi) 的显式解为

$$\hat{\theta}_i=(I+X_{ii}^{\mathrm{T}}X_{ii})^{-1}[b_i^0+X_{ii}^{\mathrm{T}}(y_i-\sum_{r\neq i}X_{ir}\hat{\theta}_r^0)]. \tag{8}$$

于是第一级的迭代算法为

$$\hat{\theta}_i^{(l)}=(I+X_{ii}^{\mathrm{T}}X_{ii}{}^{-1}[b_i^{0(l)}+X_{ii}^{\mathrm{T}}(y^i-\sum_{r\neq i}X_{ir}\hat{\theta}_r^{0(l)})]. \tag{9}$$

协调级欲使等式约束 $\hat{\theta}^0=\hat{\theta}$ 成立，可采用一般形式迭代算法

$$\hat{\theta}^{0l+1}=\hat{\theta}^{0(l)}-sD(\hat{\theta}^{0(l)}-\hat{\theta}^{(l)};\quad D=\text{block-diag}\{D_{11},\cdots,D_{NN}\}, \tag{10}$$

或两步法

$$\hat{\theta}^{0l+1}=\hat{\theta}^{0(l)}+\alpha^{(l)}(\hat{\theta}^{(l)}-\hat{\theta}^{0(l)}+\beta^{(l)}(\hat{\theta}^{0(l)}-\hat{\theta}^{0(l-1)}). \tag{11}$$

其中，s 为松驰因子，D 为修正矩阵且 $\det(D)\neq 0$，$\alpha^{(l)}$，β^l 为两步法步长. 为使(11) 自开始(不需给出 $\hat{\theta}^{0(-1)}$)，取 $\beta^{(0)}=0$. D, $\alpha^{(l)l}$ 及 $\beta^{(l)}$ 的取法将在下节中讨论. 整个递阶算法步骤如下：

1) 给定初始 $\hat{\theta}^{(0)}$，允许误差 $\varepsilon>0$，令 $l=0$.

2) 对于 $i=1,\cdots,N$，据(9) 求得 $\hat{\theta}_i^{(l)}$.

3) 若 $\|\hat{\theta}^{0(l)}-\hat{\theta}^{(l)}\|^2\leqslant\varepsilon$,得最小二乘估计 $\hat{\theta}^*=\hat{\theta}^{(l)}$,算法终止,否则转 4).

4) 据(10)或(11)计算 $\hat{\theta}^{0(l+1)}$,令 $l=l+1$,转(2).

3　算法收敛法

1) (10)中,$D=I$,$S=1$,即

$$\hat{\theta}^{0(l+1)}=\hat{\theta}^{(l)}. \tag{12}$$

定理 1　设矩阵A非奇异,且矩阵$A^{-1}B$的谱半径满足:$\rho[A^{-1}B]<1$,则(12)式给出的协调算法收敛.这里

$$A=\text{block-diag}\{A_{11},\cdots,A_{NN}\}\in\mathrm{R}^{n\times n};\quad A_{ii}=I+X_{ii}^{\mathrm{T}}X_{ii},\quad i=1,\cdots,N, \tag{13}$$

$$B=\begin{bmatrix}B_{11}&\cdots&B_{1N}\\ \vdots&\cdots&\vdots\\ B_{N1}&\cdots&B_{NN}\end{bmatrix}\in\mathrm{R}^{n\times n};\quad B_{ij}=\begin{cases}-\sum\limits_{k=1}^{N}X_{ki}^{\mathrm{T}}X_{kj}, & i\neq j,\\ I-\sum\limits_{k\neq 1}X_{ki}^{\mathrm{T}}X_{kj}, & i=j,\end{cases}\quad i,j=1,\cdots,N. \tag{14}$$

证　对于协调级来说,(5)总是成立的,该式两端对 $\hat{\theta}_j^0$ 求导得

$$(I+X_{ii}^{\mathrm{T}}X_{ii})\frac{\partial\hat{\theta}_i}{\partial\hat{\theta}_j^0}+\sum_{k=1}^{N}X_{ki}^{\mathrm{T}}X_{kj}=0,\quad i\neq j,$$

$$(I+X_{ii}^{\mathrm{T}}X_{ii})\frac{\partial\hat{\theta}_i}{\partial\hat{\theta}_j^0}+\sum_{k\neq 1}X_{ki}^{\mathrm{T}}X_{kj}-I=0,\quad i=j,$$

即

$$A_{ii}\frac{\partial\hat{\theta}_i}{\partial\hat{\theta}_j^0}-B_{ij}=0.$$

由定理条件 $\det(A)=\prod\limits_{i=1}^{N}\det(A_{ii})\neq 0$,$A_{ii}^{-1}$ 存在,故有 $\partial\hat{\theta}_i/\partial\hat{\theta}_j^0=A_{ii}^{-1}B_{ij}$,即 $\hat{\theta}/(\hat{\theta}^0)=A^{-1}B$,亦即

$$\partial\hat{\theta}^{0(l+1)}/\partial\hat{\theta}^{0(l)}=A^{-1}B. \tag{15}$$

因此,$\rho[\partial\hat{\theta}^{0(l+1)}/\partial\hat{\theta}^{0(l)}]=\rho[A^{-1}B]<1$,于是,据 Ostrouski 定理[2],算法收敛.　　证毕.

下面的定理进一步给出了判定矩阵 $A^{-1}B$ 的谱半径 $\rho[A^{-1}B]<1$ 的充分条件:

定理 2　若矩阵 $X^{\mathrm{T}}X$ 满足下列任一条件:

1° $X^{\mathrm{T}}X$ 为严格对角优势矩阵;

2° $X^{\mathrm{T}}X$ 为不可约对角优势矩阵;

3° $X^{\mathrm{T}}X$ 为严格块对角优势矩阵,

则

$$\rho[A^{-1}B]<1.$$

证　仿文献[3],注意到 $A-B=X^{\mathrm{T}}X$ 即可证.

2) (10)中,$D=A$,即

$$\hat{\theta}^{0l+1}=\hat{\theta}^{0(l)}-sA(\hat{\theta}^{0(l)}-\hat{\theta}^{(l)}). \tag{16}$$

定理 3　设 $X^{\mathrm{T}}X$ 非奇异,$\det(A)\neq 0$,且 $o<s<\dfrac{2}{\lambda_{\max}}$,则(16)式给出的协调算法收敛.这里 $\lambda_{\max}$ 为 $X^{\mathrm{T}}X$ 的最大特征值.

证　由 $X^{\mathrm{T}}X$ 的非奇异性知,$X^{\mathrm{T}}X>0$(正定).

现记

$$G(\hat{\theta}^0)=\hat{\theta}^0-sA(\hat{\theta}^0-\hat{\theta}(\hat{\theta}^0)),$$

则有

$$\begin{aligned}\partial\hat{\theta}^{0(l+1)}/\partial\hat{\theta}^{0(l+1)} &= G'(\hat{\theta}^{0(l)}) \\ &= I - sA(I - A^{-1}B) \\ &= I - s(X^TX),\end{aligned}$$

因此，$\rho[\partial\hat{\theta}^{0(l+1)}/\partial\hat{\theta}^{0(l)}] = \max\limits_{\lambda}|1 - s\lambda|$，其中 λ 为 X^TX 的特征值. 由 $X^TX > 0$ 即有 $\lambda > 0$，再据定理条件得 $0 < s\lambda < \frac{2}{\lambda_{\max}} \cdot \lambda_{\max} = 2$，所以 $\max\limits_{\lambda}|1 - s\lambda| < 1$，即 $\rho[\partial\hat{\theta}^{0(l+1)}/\partial\hat{\theta}^{0(l)}] < 1$，满足 Ostrouski 定理条件[2]. 证毕.

对于(11) 式的两步法，记

$$e^{(l)} = \hat{\theta}^{(l)} - \hat{\theta}^{0(l)}, \tag{17}$$

由(5)，(7) 得：

$$A_{ii}e_i^{(l)} = \sum_{r=1}^{N}(X_{ri}^Ty_r - (X^TX)_{ir}\hat{\theta}_r^{0(l)}), \quad i = 1,\cdots,N. \tag{18}$$

这里 $(X^TX)_{ir} = \sum\limits_{k=1}^{N} X_{ki}^TX_{kr}$ 为 (X^TX) 的第 i,r 个分块.

将(18) 写成整体形式

$$Ae^{(l)} = X^Ty - (X^TX)\hat{\theta}^{0(l)}, \tag{19}$$

若取

$$\alpha^0 = \gamma_0/q_0; \quad \alpha^l = \frac{\gamma_l}{ql}[\beta^{(l)} + 1]; \quad l \geqslant 1, \tag{20}$$

和

$$\beta^{(0)} = 0; \quad \beta^{(l)} = [1 - \gamma_l^2/(\gamma_{l-1}q\alpha^{(l-1)})]^{-1} - 1, \quad l \geqslant 1. \tag{21}$$

这里

$$\gamma_l = e^{T(l)}Ae^{(l)} \quad q_l = e^{T(l)}(X^TX)e^{(l)}, \quad l \geqslant 0. \tag{22}$$

则有

定理 4 设 X^TX 及 A 非奇异，则协调算法(11) 至多经 n 次迭代，就有 $e^{(l)} = \hat{\theta}^{(l)} - \hat{\theta}^{0(l)} = 0$.

证 不妨设 $e^k \neq 0$, $k = 0,\cdots,n-1$，则如果 $\{e^{(l)}, l = 0,\cdots,n\}$ 是两两 A-正交的，必有 $e^{(n)} = 0$. 现用数学归纳法证之.

$$\begin{aligned}\because \quad e^{T(0)}Ae^{(1)} &= e^{T(0)}(X^Ty - (X^TX)\hat{\theta}^{0(1)}) \\ &= e^{T(0)}(X^Ty - (X^TX)(\hat{\theta}^{0(0)} + \alpha^{(0)}e^{(0)})) \\ &= e^{T(0)}(Ae^{(0)} - \alpha^{(0)}(X^TX)e^{(0)}) \\ &= \gamma_0 - \alpha^{(0)}q_0 = \gamma_0 - \gamma_0 = 0,\end{aligned}$$

故进一步设 $i,j \leqslant k$ $(i \neq j)$ 时，有

$$e^{T(i)}Ae^{(j)} = 0. \tag{23}$$

则当 $i,j \leqslant k+1$ $(i \neq j)$ 时，只须利用归纳假设(23) 证 $e^{(k+1)}$ 与 $e^{(j)}, j = 0,\cdots,k$ 两两 A-正交. 将(11) 两端左乘 $-(X^TX)$，再加 X^Ty，由(19) 可得

$$Ae^{(l+1)} = Ae^{(l)}(X^TX)e^{(l)} - \beta^{(l)}(Ae^{(l-1)} - Ae^{(l)}). \tag{24}$$

(24) 中,令 $l=k$,分两步:

i) 乘以 $e^{T(k)}$ 得:

$$e^{T(k)}Ae^{(k+1)}=\gamma_k-\alpha^{(k)}\cdot q_k+\beta^{(k)}\gamma_k$$
$$=\gamma_k-\gamma_k(\beta^{(k)}+1)+\beta^{(k)}\gamma_k=0; \tag{25}$$

ii) 乘以 $e^{T(k-1)}$ 得

$$e^{T(k-1)}Ae^{(k+1)}=-\alpha^{(k)}e^{T(k-1)}(X^{\mathrm{T}}X)e^{(k)}-\beta^{(k)}\gamma_{k-1}. \tag{26}$$

(24) 中,令 $l=k-1$,乘以 $e^{T(k)}$ 得

$$\gamma_k=-\alpha^{(k-1)}e^{T(k)}(X^TX)e^{(k-1)}=-\alpha^{(k-1)}e^{T(k-1)}(X^{\mathrm{T}}X)e^{(k)},$$

即

$$e^{T(k-1)}(X^{\mathrm{T}}X)e^{(k)}=-\gamma_k/\alpha^{(k-1)}. \tag{27}$$

(27) 代入(26) 得

$$e^{T(k-1)}Ae^{(k+1)}=\alpha^{(k)}\frac{\gamma_k}{\alpha^{k-1}}-\beta^{(k)}\gamma_{k-1}$$
$$=\{\frac{\gamma_k^2}{\gamma_{k-1}q_k\alpha^{(k-1)}}[\beta^{(k)}+1]-\beta^{(k)}\}\gamma_{k-1}$$
$$=\{-[1-\frac{\gamma_k^2}{\gamma_{k-1}q_k\alpha^{(k-1)}}-1][\beta^{(k)}+1]-\beta^{(k)}\}\gamma_{k-1}$$
$$=\{-[1-\beta^{(k)}-1]-\beta^{(k)}\}\gamma_{k-1}=0. \tag{28}$$

(24) 中,再令 $l=k$,乘以 $e^{T(i)}(0\leqslant i\leqslant k-1)$ 有

$$e^{T(i)}Ae^{(k+1)}=-\alpha^{(k)}e^{T(i)}(X^{\mathrm{T}}X)e^{(k)}. \tag{29}$$

(24) 中,令 $l=i$,乘以 $e^{T(k)}$ 有(注意 $i+1)<k$)

$$0=-\alpha^{(i)}e^{T(k)}(X^{\mathrm{T}}X)e^{(i)}=-\alpha^{(i)}e^{T(i)}(X^{\mathrm{T}}X)e^{(k)},$$

即 $e^{\mathrm{T}(i)}(X^{\mathrm{T}}X)e^{(k)}=0$,代入(29) 得

$$e^{T(i)}Ae^{(k+1)}=0,\quad i=0,\cdots,k-2. \tag{30}$$

综合(25),(28),(30) 知,$e^{(k+1)}$ 与 $e^{(j)}(j=0,\cdots,k)$ 两两 A-正交. 因此,$\{e^{(l)},l=0,\cdots,n\}$ 是两两 A-正交的. 证毕.

定理 5 设(10) 或(11) 式给出的协调算法收敛于 $\hat{\theta}^*$,则有 $\hat{\theta}^*=(X^TX)^{-1}X^{\mathrm{T}}y$,为(2) 的解.

证 因为算法收敛时有 $\hat{\theta}^0=\hat{\theta}=\hat{\theta}^*$,将它代入(5).(7) 式得

$$-X_{ii}^T(y_i-\sum_{j=1}^N X_{ij}\hat{\theta}_j^*)+\hat{\theta}_i^*-\hat{\theta}_i^*-\sum_{j\neq i}X_{ji}^{\mathrm{T}}(y_j-\sum_{k=1}^N X_{jk}\hat{\theta}_k^*)=0,$$

即

$$-\sum_{j=1}^N X_{ji}^{\mathrm{T}}(y_j-\sum_{k=1}^N X_{jk}\hat{\theta}_k^*)=0,\quad i=1,\cdots,N.$$

写成整体形式

$$-X^{\mathrm{T}}(y-X\hat{\theta}^*)=0,$$

解得

$$\hat{\theta}^*=(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}y.$$

证毕.

4 计算量分析

据(3),可得整体算法的乘(除) 法次数为[4]

$$\frac{m(n+1)n}{2}+nm+\frac{n^3}{3}+n^2-\frac{n}{3}. \tag{31}$$

其中第 1,2 项为计算 $X^{\mathrm{T}}X, X^{\mathrm{T}}y$ 诸元素所需乘法次数，第 3 ～ 5 项为 Gauss 消去法方程 $(X^{\mathrm{T}}X)\hat{\theta}^*=X^{\mathrm{T}}y$ 所需乘(除)法次数.

本文所提算法的主要特点之一是便于在多处理机系统上实现. 若采用 N 个处理机的系统求解，由于 D 为分块对角形式，所以(12),(16) 与(18) 可改写成

$$\hat{\theta}_i^{0(l+1)}=\hat{\theta}_i^{0(l)}+A_{ii}^{-1}\left[\sum_{r=1}^{N}(X_{ri}^{\mathrm{T}}y_r-(X^{\mathrm{T}}X)_{ir}\hat{\theta}_r^{0(l)}\right], \tag{32}$$

$$\hat{\theta}_i^{0(l+1)}=\hat{\theta}_i^{0(l)}+s\left[\sum_{r=1}^{N}(X_{ri}^{\mathrm{T}}y_r-(X^{\mathrm{T}}X)_{ir}\hat{\theta}_r^{0(l)})\right]. \tag{33}$$

而对于(11) 的两步算法，将(22) 改写成

$$\gamma_l=\sum_{i=1}^{N}e_i^{T(l)}A_{ii}e_i^{(l)},\quad q_l=\sum_{i,j}e_i^{T(l)}(X^{\mathrm{T}}X)_{ij}e_j^{(l)}=\sum_{i=1}^{N}e_i^{T(l)}p_i^{(l)}. \tag{34}$$

这里

$$p_i^{(l)}=\sum_{j=1}^{N}(X^{\mathrm{T}}X)_{ij}e_j^{(l)}. \tag{35}$$

故 γ_l 和 q_l 的主要计算工作也可分解到 N 个处理机上，再让某个处理机承担剩余的加法和计算 $\alpha^{(l)}$, $\beta^{(l)}$ 的不大计算量即可. 因此，据(32)，算法(12) 每个处理机所需乘(除) 法次数为

$$c_i=\frac{(n+1)n_i}{2}m+n_im+\left(\frac{n_i^3}{3}+n_i^2-\frac{n_i}{3}\right)+(n_in+n_i^2)L. \tag{36}$$

据(33)，算法(16) 每个处理机所需乘(除) 法次数为

$$c_i=\frac{(n+1)n_i}{2}m+n_im+(n_in+n_i)L. \tag{37}$$

据(18),(34),(35)，算法(11) 每个处理机所需乘(除) 法次数为

$$c_i=\frac{(n+1)n_i}{2}m+n_im+\left(\frac{n_i^3}{3}+n_i^2-\frac{ni}{3}\right)+(2n_in+4n_i+n_i^2)L. \tag{38}$$

其中某个处理机尚须增加计算 $\alpha^{(l)},\beta^{(l)}$ 的 $6L$ 次乘(除) 法. 这里 L 为收敛时的迭代次数.

现假定最小二乘估计问题中 $n=20, m=100$，将它分解为 $N=50, n_i=4, m_i=20, i=1,\cdots,5$，则对于协调算法(12)：$c_i=4636+96L$，对于协调算法(16)：$c_i=4600+84L$，对于协调算法(11)：$c_i+6L=4636+198L$，而据(31)，整体算法(3) 所需乘除法次数为 26060.

因此，对于协调算法(12),(16)，和(11)，只要 L 分别小于 224,256 和 109，就可比整体算法(3) 节省时间. 而由定理 4 知，(11) 的 $L\leqslant 20<109$，所以它无疑可节省相当多的计算时间.

5　仿真实例

设线性离散大系统的传递函数为

$$F(z) = (b_0 + b_1 z^{-1} + \cdots + b_p z^{-p})/(1 + a_1 z^{-1} + \cdots + a_p z^{-p}), \tag{39}$$

则相应的时域方程如下

$$y(k) + a_1 y(k-1) + \cdots + a_p y(k-p) = b_0 u(k) + \cdots + b_p u(k-p) + v(k). \tag{40}$$

式中,随机变量 $v(k)$ 描述系统的不确定性或观测噪声.显然这里 $n = 2p + 1$.

如果已得到(40)容量为 m 的样本,且 $m >> n$,则(40)易于改写成(1)的形式:

$$y = X\theta + e.$$

这里

$$X = \begin{bmatrix} -y(0) & -y(-1) & \cdots & -y(1-p) & u(1) & \cdots & u(1-p) \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ -y(m-1) & -y(m-2) & \cdots & -y(m-p) & u(m) & \cdots & u(m-p) \end{bmatrix},$$

$y = [y(1), y(2), \cdots, y(m)]^T$, $\theta = [a_1, \cdots, a_p, b_0, \cdots, b_p]^T$,

$e = [v(1), \cdots, v(m)]^T$.

现设定 $p = 8$,参数向量真值 $\theta = [1.2, -0.8, 1.7, 2.1, -1.0, 1.0, 1.0, 0.8, 0.0, 2.3, 0.5, 0.6, 1.4, -1.5, -1.6, 0.7, 1.5]^T$,样本容量 $m = 96$,且分解为 $n_i = 2$, $i = 1, \cdots, 7$, $n_8 = 3$, $m_i = 12$, $i = 1, \cdots, 8$,取初值 $\hat{\theta}^{0(0)} = 0$.则当 a) $s = 1.4$, $\omega = 1.4$, $y(k) = 0$, $k = 0, -1, \cdots, -p+1$. b) $s = 0.9$, $\omega = 1.2$, $y(k) = 3.73$, $k = 0, -2, \cdots, -6$, $y(k) = -1.23$, $k = -1, -3, \cdots, -7$ 时,各算法的仿真结果见表1.

据(31),整体算法(3)的乘除法次数为18241.由表1可见,本文所提算法在多处理机系统上实现时,可比整体算法大大节省计算时间且其收敛性和收敛速度均优于[1]的平行分解法.

表1 仿真结果表*

方法	文献[1]算法	算法(12)	算法(16)	算法(11)
收敛迭代次数	41	34	30	15
a) 子问题最大乘除法次数	5357	4937	4500	4742
$\Vert\theta - \theta^*\Vert$	2.52×10^{-4}	3.83×10^{-4}	1.98×10^{-4}	8.16×10^{-5}
收敛迭代次数	不收敛	不收敛	57	14
b) 子问题最大乘除法次数			5878	4619
$\Vert\theta - \theta^*\Vert$			3.17×10^{-4}	1.15×10^{-4}

6 结论

理论分析说明,本文利用目标补偿的基本思想导出的分解-协调算法具有相当好的收敛性,且收敛于原问题的解.和文献[1]的有关定理比较,文中定理1～4给出的收敛条件既较便于判别又易于满足.仿真实例表明,这一算法的收敛性和收敛速度也明显优于[1]的算法,且在多处理机系统上实现时,可比整体算法大大节省计算时间.

* 文献[1]计算量分析公式有误,这里按已更正公式统计

参 考 文 献

[1] Sultan, M. A., Hassan, M. F. Calvet, J. L. and Bilal, A. Y.. A First Order Parallel Decomposition Method for Parameter Estimation in Large Scale Systems. Large Scale Systems: Theory and Application, 1985, 9(1):51—62

[2] Ortega, J. M. and Rheinboldt, W. C.. Iterative Solution of Nonlinear Equations in Several Variables. New York: Academic Press, 1970, 300--301

[3] He Jianmin and Da Qingli. The Multilevel Parameter Estimation Approach for a Kind of Nonlinear Models with Applications. Preprints of 8th IFAC/IFORS Symposium on Identification and System Parameter Estimation, 1988, 1:627—632

[4] 李庆杨等. 数值分析. 武汉:华中工学院出版社, 1981, 355—356

An Objective Compensated Coordination Approach for Parameter Estimation in Large-Scale Linear Systems

He Jianmin　and　Da Qingli

(College of Economics and Management, Southeast University • Nanjing, 210018, PRC)

Abstract: In order to overcome the so called "dimensional disasters" caused by the parameter estimation problem of a large-scale linear system, an objective compensated coordination approach is proposed in the paper. Its main feature is that the coordination is realised by introducing compensation terms into the subproblems instead of using Lagrangian multiplier as the coordination variable, so the coordination structure is simplified. Theoretical analysis and applied calculation show that its convergence character and calculating time are better and less than those of the parallel decomposition approach proposed in [1] and the whole one

Key words: large-scale linear systems; decomposition-coordination; least-squares; objective compensated; hierarchical optimization

本文作者简介

何建敏　1956年生. 副教授. 1985年毕业于南京工学院自动控制理论及应用专业,并获硕士学位. 主要著作有《大系统理论与方法》(第二编著),《计算机应用系统软件开发基础》等. 主要研究领域为大系统算法、决策与决策支持.

达庆利　1945年生,教授. 1966年毕业于南京工学院自动控制系,1982年获同一专业硕士学位. 主要著作为《大系统理论与方法》(第一编著)等. 主要研究方向为大系统建模、规划与决策. 应用领域为环境经济、工厂企业等系统.