第 7 卷 第 4 期
1990 年 12 月

控制理论与应用
CONTROL THEORY AND APPLICATIONS

Vol. 7, No. 4
Dec., 1990

条件数在大系统辨识数值分析中的应用

张红艺
(江西大学计算机科学系,南昌)

E. J. Davison
(Univ. of Toronto Canada)

摘要 本文利用条件数的链式性质,研究了大系统辨识中存在的数值问题.证明了关于参数辨识问题的条件数一般随着阶数增加而严格地单调增加.

关键词:大系统;条件数;数值问题

1. 引 言

考虑下面的单输入单输出模型:

$$y(k)+a_1y(k-1)+\cdots+a_ny(k-n)=b_0u(k)+b_1u(k-1)+\cdots+b_nu(k-n), \tag{1.1}$$

其中 k 表示离散时刻,u 是输入,y 是输出,a_i、b_i 是系统参数,n 是系统阶数.

假定系统(1.1)实际可量测的输入 u_m 和输出 y_m 分别带有噪音 ω_u 和 ω_y,即

$$u_m=u+w_u,\quad y_m=y+w_y.$$

因此可假定由系统(1.1)可获得一个输入/输出数据集:

$$Z^N\triangleq\{u_m(1),y_m(1),\cdots,u_m(N),y_m(N)\},$$

其中 $N\geqslant 2n+1$.

定义向量

$$U(1,N-i)\triangleq\begin{bmatrix}0\\ \vdots\\ 0\\ u(1)\\ \vdots\\ u(N-i)\end{bmatrix}\in R^N.$$

那么可以构造下列矩阵:

$$M_k\triangleq[U(1,N),U(1,N-1),\cdots,U(1,N-k),-Y(1,N-1),\cdots,-Y(1,N-k)],$$

$$\overline{M}_k\triangleq[U_m(1,N),\cdots,U_m(1,N-k),-Y_m(1,N-1),\cdots,-Y_m(1,N-k)],$$

$$\Delta M_k\triangleq[W_u(1,N),\cdots,W_u(1,N-k),-W_y(1,N-1),\cdots,-W_y(1,N-k)],$$

$$\eta\triangleq Y(1,N),\bar{\eta}\triangleq Y_m(1,N),\quad \Delta\eta\triangleq W_y(1,N),$$

其中 M_k、$\overline{M}_k$、$\Delta M_k\in R^{N\times(2k+1)}$; η、$\bar{\eta}$、$\Delta\eta\in R^N$.

由式(1.1)可得

本文于 1989 年 3 月 14 日收到. 1990 年 3 月 2 日收到修改稿.

$$M_n\theta = \eta,$$

其中 $\theta \triangleq [b_0\ b_1\cdots b_n, a_1\ a_2\cdots a_n]^T$. 那么参数估计最小二乘问题就是求

$$\|\overline{M}_n\theta_{est} - \bar{\eta}\|_2 = \min$$

的最小范数解 θ_{est}. 显然有[1]

$$\theta_{est} = \overline{M}_n{}^+ \bar{\eta},$$

其中 $\overline{M}_n{}^+ = (\overline{M}_n^T\overline{M}_n)^{-1}\overline{M}_n^T$.

如果 $\|\Delta M_n\|_2\|M_n^+\|_2<1$, 那么估计 θ_{est} 具有性质[1]:

$$\frac{\|\theta-\theta_{est}\|_2}{\|\theta\|_2} \leqslant \frac{co\#(M_n)}{1-\|\Delta M_n\|_2\|M_n^+\|_2}\left\{\frac{\|\Delta M_n\|_2}{\|M_n\|_2}+\frac{\|\Delta\eta\|_2}{\|\eta\|_2}\right\}, \tag{1.2}$$

其中 $co\#(M_n)=\|M_n\|_2\|M_n^+\|_2$ 称为矩阵 M_n 的条件数. 式(1.2)描述了参数估计精度与条件数的关系. 如果 $co\#(M_n)$ 是"大的", 那么意味着 M_n 是病态, 且估计 θ_{est} 是极不可靠的[1]. 因此在实际辨识问题中, 总是期望 $co\#(M_n)$ 尽可能地小.

本文首先研究了矩阵随着列数的增加, 其条件数严格地单调增加的链式性质. 给出了文[4]中部分结果的证明. 即关于辨识问题的条件数几乎总是随着系统待估计阶数的增加而严格地单调增加.

2. 主要结果

首先讨论矩阵条件数的链式性质. 令 $A\in C_r^{n\times m}$ 表示 $A\in C^{n\times m}$, 且 $\mathrm{rank}(A)=r$.

性质 2.1 令 $A\subset B$ 表示 A 的所有列包含在 B 的所有列中, 那么链 $A_1\subset A_2\subset\cdots\subset A_n$ 具有性质:

$$co\#(A_1)\leqslant co\#(A_2)\leqslant\cdots\leqslant co\#(A_n). \tag{2.1}$$

证明可直接由文[1]得到.

下面研究式(2.1)中不等号严格成立的充要条件. 为此由文[3]引入下面的结果.

引理 2.1 假定 $\overline{A}=\begin{bmatrix}A & b\\ c^T & d\end{bmatrix}$, 其中 $A\in R^{n\times n}, b、c\in R^n, d\in R$; 那么对 $\forall\ s\in C$,

$$\det(sI-\overline{A}) = (s-d)\det(sI-A) - c^T\mathrm{adj}(sI-A)b.$$

引理 2.2 给定 A_{k+1}, 令 A_k 表示从 A_{k+1} 去掉最后一列所得的矩阵, 即 $A_{k+1}=[A_k\quad b_{k+1}]\in R_{k+1}^{N\times(k+1)}$. 假定 $\{\alpha_i, i\in \underline{k+1}\}$、$\{\beta_i, i\in \underline{k}\}$ 分别是 A_{k+1} 和 A_k 按降序排列的奇异值. 则

$$\alpha_1>\beta_1>\alpha_2>\cdots>\alpha_k>\beta_k>\alpha_{k+1},$$

当且仅当对 $\forall\ \sigma\in\{\beta_i, i\in\underline{k}\}$,

$$b_{k+1}^TA_k\mathrm{adj}(\sigma^2I-A_k^TA_k)A_k^Tb_{k+1}\neq 0. \tag{2.2}$$

证 由[1]有

$$\alpha_1\geqslant\beta_1\geqslant\alpha_2\geqslant\cdots\geqslant\alpha_k\geqslant\beta_k\geqslant\alpha_{k+1}.$$

必要性: 利用引理2.1, 可推得

$$\det(\sigma^2I-A_{k+1}^TA_{k+1}) = (\sigma^2-b_{k+1}^Tb_{(k+1)})\det(\sigma^2I-A_k^TA_k) - b_{k+1}^TA_k\mathrm{adj}(\sigma^2I-A_k^TA_k)A_k^Tb_{k+1}.$$

于是对所有 β_i, 有

$$\det(\beta_i^2 I - A_{k+1}^T A_{k+1}) = - b_{k+1}^T A_k \operatorname{adj}(\beta_i^2 I - A_k^T A_k) A_k^T b_{k+1} \neq 0.$$

充分性：如果存在 $\alpha_i = \beta_i$，那么

$$\det(\alpha_i^2 I - A_{k+1}^T A_{k+1}) = - b_{k+1}^T A_k \operatorname{adj}(\beta_i^2 I - A_k^T A_k) A_k^T b_{k+1} = 0.$$

于是由矛盾得证.

由性质2.1和引理2.2可得到下面的结果.

性质 2.2　假定 $A_{k+1} = [A_k \quad b_{k+1}] \in R_{k+1}^{N\times(k+1)}$. 那么

$$co\#(A_k) < co\#(A_{k+1}),$$

当且仅当

$$b_{k+1}^T A_k \operatorname{adj}(\sigma_{\max}^2(A_k) I - A_k^T A_k) A_k^T b_{k+1} \neq 0, \tag{2.3a}$$

或者

$$b_{k+1}^T A_k \operatorname{adj}(\sigma_{\min}^2(A_k) I - A_k^T A_k) A_k^T b_{k+1} \neq 0, \tag{2.3b}$$

其中 $\sigma_{\max}$、$\sigma_{\min}$表示矩阵最大和最小奇异值.

为便于理解式(2.3)，下面用系统理论观点来描述性质2.2.

引理 2.3　假定 $A_{k+1} = [A_k \quad b_{k+1}] \in R_{k+1}^{N\times(k+1)}$，且 $b_{k+1}^T A_k \neq 0$. 那么 A_k 和 A_{k+1}至少有一个相同的奇异值当且仅当$(A_k^T A_k, A_k^T b_{k+1})$是不可控的.

证　充分性：假定$(A_k^T A_k, A_k^T b_{k+1})$是不可控的，那么至少存在一个 A_k 的奇异值 σ 及相应的非零向量 x 使得

$$x^T A_k^T b_{k+1} = 0,$$

$$x^T(\sigma^2 I - A_k^T A_k) = 0.$$

从而有

$$A_{k+1}^T A_{k+1} \begin{bmatrix} x \\ 0 \end{bmatrix} = \begin{bmatrix} A_k^T A_k x \\ b_{k+1}^T A_k x \end{bmatrix} = \sigma^2 \begin{bmatrix} x \\ 0 \end{bmatrix}.$$

因此 σ 也是 A_{k+1}的奇异值.

必要性：假定 σ 是 A_k 和 A_{k+1}的奇异值. 那么由引理2.2得

$$\det(\sigma^2 I - A_k^T A_k) = 0$$

及

$$b_{k+1}^T A_k \operatorname{adj}(\sigma^2 I - A_k^T A_k) A_k^T b_{k+1} = 0.$$

这说明系统$(b_{k+1}^T A_k, A_k^T A_k, A_k^T b_{k+1})$不是可控/可观的. 从而$(A_k^T A_k, A_k^T b_{k+1})$是不可控的.

由引理2.2及2.3可推得：

性质 2.3　假定 $A_{k+1} = [A_k \quad b_{k+1}] \in R_{k+1}^{N\times(k+1)}$，且 $b_{k+1}^T A_k \neq 0$. 那么

$$co\#(A_k) < co\#(A_{k+1}),$$

当且仅当 $\sigma_{\max}^2(A_k)$或 $\sigma_{\min}^2(A_k)$是$(A_k^T A_k, A_k^T b_{k+1})$的可控模.

引理 2.4[2]　假定 $A_k \in R^{N\times k}$，$N \geqslant k$，具有 $\operatorname{rank}(A_k) = k$；那么对几乎所有 $B_k \in R^{N\times k}$，有

$$\operatorname{rank}(A_k + B_k) = k.$$

现在把上面的结果应用于第1节中所定义的矩阵 M_k、$\overline{M}_k$ 及 ΔM_k 中，可以得到下面一些结果.

定理 2.1　考虑系统(1.1)，那么

1)如果系统没有噪声，即 $w_u = w_y = 0$.

$$\operatorname{rank}(M_k) = 2k + 1, \quad k \leqslant n,$$

$$\mathrm{rank}(M_k) < 2k+1, \quad k > n.$$

2)如果系统有噪声,即 $w_u \neq 0, w_y \neq 0$,则对所有正整数 k,$\mathrm{rank}(\overline{M}_k)=2k+1$对几乎所有 $w_u(i)$、$w_y(i)$, $i=1,2,\cdots,N$ 成立.

证 第一部分的证明可参见文[3].

对于第二部分,对几乎所有 $w_u(i)$、$w_y(i)$,$i=1,2,\cdots,N$,可假定 $\mathrm{rank}(\Delta M_k)=2k+1$. 再由引理2.4可得,$\mathrm{rank}(\overline{M}_k)=\mathrm{rank}(M_k+\Delta M_k)=2k+1$对几乎所有 $w_u(i)$、$w_y(i)$,$i=1,2,\cdots,N$ 成立.

定理 2.2 考虑系统(1.1),

1)如果系统没有噪声,那么对几乎所有参数 $a_1,a_2,\cdots,a_n;b_0,b_1,\cdots,b_n$,和对几乎所有输入数据 $u(i)$,$i=1,2,\cdots,N$,$N\geqslant 2n+1$,有

$$co\#(M_1) < co\#(M_2) < \cdots < co\#(M_n).$$

2)如果系统有噪声,且 $\mathrm{rank}(\Delta M_k)=2k+1$,$k=1,2,\cdots$,那么对几乎所有扰动项 ΔM_k,$k=1,2,\cdots$,有

$$co\#(\overline{M}_k) < co\#(\overline{M}_{k+1}).$$

证 第一部分:令 t_k 是 M_{k+1}的最后一列. 于是由性质2.3可得,如果$\{M_k^TM_k, M_k^Tt_k\}$的可控性矩阵 Γ_k 是满秩的,则 $co\#(M_k)<co\#(M_{k+1})$,$k=1,2,\cdots,n-1$. 注意到若 Γ_k 不是满秩的,则意味着元素 $a_1,\cdots,a_n;b_0,\cdots,b_n;u(1),\cdots,u(N)$必须在参数空间 a_i、b_i、$u(i)$的某个超平面上. 可以证明这个超平面是正则的. 这是因为存在一类系统[3]:

$$y_k - y_{k-n} = u_{k-n}, \quad n \geqslant 1.$$

当 $u_1=1, u_i=0, i>1; y_i=0, i=1,2,\cdots,n$ 时有

$$co\#(M_k) < co\#(M_{k+1}), \quad 1 \leqslant k \leqslant n-1.$$

于是对几乎所有 a_i、b_i、$u(i)$有 $\mathrm{rank}(\Gamma_k)=2k+1$,$k=1,\cdots,n-1$. 故推得 $co\#(M_k)<co\#(M_{k+1})$,$k=1,\cdots,n-1$.

第二部分同样可以从定理2.1的第二部分性质2.3及引理2.4得到证明.

根据定理2.2,系统(1.1)可分成两类:

Ⅰ 当 $n\to\infty$时,系统的条件数 $co\#(M_n)$为无界;

Ⅱ 当 $n\to\infty$时,系统的条件数 $co\#(M_n)$有界.

如果阶数 n“很大”,则类型Ⅰ系统的参数估计是不可靠的;类型Ⅱ系统的估计精度取决于系统条件数的实际值.

3. 结 论

本文对于单输入单输出模型,证明了辨识问题的矩阵条件数一般随着阶数增加而严格地单调增加. 这表明,大系统辨识常常会遇到固有的数值困难.

参 考 文 献

[1] Lawson, C. L., and Hanson, R. J., Solving Least Squares Problems, Prentice—Hall, Englewood Cliffs, (1974).

[2] Davison, E. J., and Wang, S. H., Properties of Time Invariant Multivariable Systems Subject to Arbitrary Output and State Feedback, IEEE Trans. on Automatic Control, AC—18, (1973), 24—32.

[3] 张红艺,分散化和递阶控制系统的结构特性及大系统辨识的数值问题,北京航空航天大学博士论文,(1989).

[4] Zhang, H. Y., and Davison, E. J., Numerical Problems in Identification of Large Scale Systems, Proc. 1989 American Control Conference.

Applications of Condition Number to the Numerical Analysis in Identification of Large Scale Systems

Zhang Hongyi (Jiangxi University, Nanchang)

E. J. Davison (University of Toronto)

Abstract: In this paper, the numerical problems of identifying large scale systems are considered by use of the chain properties of the condition number. It is shown that generically the condition number associated with the parameter estimation problem strictly monotonically increases as the system order n increases.

Key words: large scale systems; condition number; numerical problems

(continued from page 79)

IFAC/IFORS Symposium (6th) Large Scale Systems: Theory and Applications	August 22—26	Beijing China, P. R.	*	Prof. Bao Liu, Inst. of Systems Engineering, College of Engg. Tianjin University, Tianjin China, P. R.
IFAC Symposium (7th) Automation in Mining, Mineral and Metal Processing	August 26—28	Beijing China, P. R.	Aug. 20 1991	Prof. Sheng Wei—Zhi IFAC MMM Secretariat POB 919 Beijing, China, P. R.
IFAC Workshop (2nd) System Structure and Control	Sept. 3—5	Prague CSFR	*	Dr. S. Kubik, Inst. of Inf. Theory and Automation, Pod vodarens kou vezi 4, CS—1182 08 Prague
IFAC Symposium (3rd) Low Cost Automation	Sept.	Vienna Austria	*	Prof. P. Kopacek c/o OePWZ Rockhgasse 6, A—1014 Vienna Austria

* not yet known —deadline past