

辨识VDE结构的白化残差技术

张 志 涌

(福州大学自动化所)

摘 要

用残差法(REM)从噪信比较大的观测数据中确定的结构指数往往偏高. 本文利用正交投影技术分析了REM的缺陷后, 提出了白化残差法(WREM). 白化残差可借助形成简便的白化投影算子生成. 运用白化残差法, 不仅可以准确地确定系统结构, 而且还可以指示所定结构下模型的性能.

多变量线性系统的结构辨识有三种基本方法: 经由Markov参数估计后从Hankel阵直接获最小实现法; 经过参数化模型的穷试法^[1]; 不经过参数化模型的残差法^[2,3,4,5,6]. 第三种方法由于计算量小而常被采用, 但是此法仅在噪信比较小的情况下有效. 本文将分析残差法的这个缺陷, 并进而引出新的白化残差法.

一、问题的描述

线性离散系统总可表述为向量差分方程(VDE)标准II型^[1]

$$\begin{aligned} y(t) &= A(q^{-1})y(t) + B(q^{-1})u(t) \\ &= \sum_{i=1}^{v_m} \{A^{(i)}y(t-i) + B^{(i)}u(t-i)\}, \end{aligned} \quad (1)$$

式中 $y(\cdot) \in R^m$, $u(\cdot) \in R^r$, q^{-1} 是后向移位算子, 且可观性指数 $v_1 \leq v_2 \leq \dots \leq v_m$ (这条件不影响表述的一般性). 对于 $v_{k-1} < i \leq v_k$ 的 $A^{(i)}$, $B^{(i)}$ 有如下规范形式

$$\begin{array}{c} \text{第 } k-1 \text{ 行} \\ \left(\begin{array}{cccc} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \\ * & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ * & * & \cdots & * \end{array} \right) \begin{array}{l} \text{第 } k \text{ 行,} \\ \\ \\ \\ \\ m \times x \end{array} \end{array}$$

在此, $x = m$ 或 r , 符号 $*$ 表示可能非平凡参数, k 属于正整数域 $[1, m]$, 并约定 $v_0 = 0$.

为逐个对子系统进行辨识, 可写出式(1)的第 i 个分量差分方程

$$y_i(t) = A_i(q_i^{-1})y(t) + B_i(q_i^{-1})u(t), \quad (2)$$

$$\text{式中 } A_i(q^{-1}) = \left[\left(\sum_{j=v_i-1}^0 a_{i+1,j} q^{j-v_i} \right) \cdots \left(\sum_{j=v_{i-1}-1}^0 a_{i,i-1,j} q^{j-v_i} \right) \right. \\ \left. \left(\sum_{j=v_i-1}^0 a_{i,i,j} q^{j-v_i} \right) \cdots \left(\sum_{j=v_i-1}^0 a_{i,m,j} q^{j-v_i} \right) \right], \\ B_i(q^{-1}) = \left[\left(\sum_{j=v_i-1}^0 b_{i+1,j} q^{j-v_i} \right) \cdots \left(\sum_{j=v_i-1}^0 b_{i,r,j} q^{j-v_i} \right) \right].$$

这里要特别指出: 在 $v_1 \leq v_2 \leq \cdots \leq v_m$ 条件下, $A_i(q^{-1})$ 向量的前 $(i-1)$ 分量的 Σ 求和下限分别为 $(v_1-1), \dots, (v_{i-1}-1)$, 而其余分量的 Σ 求和下限都是 (v_i-1) [注1]. 此时, $A_i(q^{-1}), B_i(q^{-1})$ 中的参数 a, b 构成了最大独立参数组. 但因结构确定前, $v_1, \dots, v_m$ 相对大小未知, 故为辨识用时, $A_i(q^{-1})$ 的前 $(i-1)$ 个分量的求和下限都应扩充至 (v_i-1) , 即

$$A_i(q^{-1}) = \left[\left(\sum_{j=v_i-1}^0 a_{i+1,j} q^{j-v_i} \right) \cdots \left(\sum_{j=v_i-1}^0 a_{i,m,j} q^{j-v_i} \right) \right].$$

当然这种修改使参数集 $\{a, b\}$ 可能包含非独立参数. [注2]

若输入、输出均受具有有理谱密度的观测噪声 $v_u(t), v_y(t)$ 干扰, 而有

$$\begin{cases} \tilde{u}(t) = u(t) + v_u(t), \\ \tilde{y}(t) = y(t) + v_y(t). \end{cases}$$

于是由式(2)可写出

$$\tilde{y}_i(t) = A_i(q^{-1}) \tilde{y}(t) + B_i(q^{-1}) \tilde{u}(t) + \varepsilon_i(t), \quad (3)$$

$$\varepsilon_i(t) = v_{y_i}(t) - A_i(q^{-1})v_y(t) + B_i(q^{-1})v_u(t). \quad (4)$$

根据观测数据可构成

$$\tilde{Y}_i(t) = \tilde{\Phi}(v_i)\theta + \varepsilon_i(t), \quad (5)$$

其中 $\tilde{\Phi}(v_i) = [\tilde{Y}_1(t-1) \cdots \tilde{Y}_m(t-1) \tilde{U}_1(t-1) \cdots \tilde{U}_r(t-1) \mid \cdots \mid \tilde{Y}_1(t-v_i)$

$$\cdots \tilde{Y}_m(t-v_i) \tilde{U}_1(t-v_i) \cdots \tilde{U}_r(t-v_i)]; \quad (6)$$

$$\tilde{Y}_i(t-j) = [\tilde{y}_i(t-j) \tilde{y}_i(t-j+1) \cdots \tilde{y}_i(t-j+l-1)]^T, (i=1, \dots, m, j=0, \dots, v_i);$$

$$\varepsilon_i(t) = [\varepsilon_i(t) \varepsilon_i(t+1) \cdots \varepsilon_i(t+l-1)]^T.$$

假设 $v_u(\cdot), v_y(\cdot)$ 和 $u(\cdot)$ 独立, 并且彼此独立, 则

$$R_{ye}(\tau) = E\{\tilde{y}_i(t-\tau)\varepsilon_i(t)\} = E\{v_{y_i}(t-\tau)[v_{y_i}(t) - A_i(q^{-1})v_y(t) + B_i(q^{-1})v_u(t)]\}$$

[注1] VDE标准 II 型的非平凡参数与 Luenberger 可观标准型的非平凡参数有一一对应关系.

[注2] 非独立参数的引入会使得参数递推估计算法中的 Hessian 矩阵变得病态或奇异. 从而在辨识结构时采用参数递推估计穷试法亦是不甚可靠的 [7].

$$\begin{aligned}
 &= E \left\{ v_{y_i}(t-\tau) \left[v_{y_i}(t-\tau) - \sum_{p=1}^m \sum_{k=0}^{v_i-1} a_{ipk} v_{yp}(t+k-v_i) \right. \right. \\
 &\quad \left. \left. + \sum_{p=1}^r \sum_{k=0}^{v_i-1} b_{ipk} v_{up}(t+k-v_i) \right] \right\} \\
 &= R_{v_{y_i}}(\tau) - \sum_{p=1}^m \sum_{k=0}^{v_i-1} a_{ipk} R_{v_{y_i p}}(\tau+k-v_i). \quad (7)
 \end{aligned}$$

同样可得

$$R_{u_i}(\tau) = \sum_{p=1}^r \sum_{k=0}^{v_i-1} b_{ipk} R_{v_{u_i p}}(\tau+k-v_i).$$

显然, $R_{y_i}(\tau) \neq 0$, $R_{u_i}(\tau) \neq 0$ ($\forall \tau = 1, 2, \dots$),

并因此可知

$$\begin{cases} [\tilde{Y}_j(t-\tau)]^T \mathbf{e}_i(t) \neq 0, & (j=1, \dots, m, \tau=1, 2, \dots), \\ [\tilde{u}_i(t-\tau)]^T \mathbf{e}_i(t) \neq 0, & (j=1, \dots, r, \tau=1, 2, \dots). \end{cases} \quad (8)$$

本文讨论的结构辨识是不通过参数估计, 而直接从观测数据确定式(3)的结构指数 $v_i (i=1, \dots, m)$.

二、残 差 法

若用 $\tilde{\Phi}(n)$ 记述结构指数试验值为 n 时按式(6)写出的由观测数据 l 维列向量构成的矩阵, 则 $S(n) = \text{span}\{\tilde{\Phi}(n)\}$ 是 l 维空间中的一个随 n 变化的子空间. $\tilde{Y}_i(t)$ 向 $S(n)$ 投影后得残差向量 $R_n(t) = [I - \tilde{\Phi}(n)\tilde{\Phi}^+(n)]\tilde{Y}_i(t)$. $[I - \tilde{\Phi}(n)\tilde{\Phi}^+(n)]$ 即是 $S(n)$ 正交补空间的投影算子, 并记为 M_n .

$$(1) \quad n < v_i$$

令 $S(v_i) = S(n) \oplus \bar{S}$, M_v 、 P 分别是 $S(v_i)$ 的补投影算子及 $\bar{S}$ 的投影算子. 那末, $M_n = M_v + P$. 又由式(5)可知 $[\tilde{Y}_i(t) - \mathbf{e}_i(t)]$ 在 $S(v_i)$ 空间内, 即 $M_v \tilde{Y}_i(t) = M_v \mathbf{e}_i(t)$ 因此

$$\begin{aligned}
 M_n \tilde{Y}_i(t) &= M_v \tilde{Y}_i(t) + P \tilde{Y}_i(t) \\
 &= M_v \mathbf{e}_i(t) + P \tilde{Y}_i(t). \quad (9)
 \end{aligned}$$

据投影算子的正交幂等性可得

$$\|M_n \tilde{Y}_i(t)\|^2 = \|M_v \mathbf{e}_i(t)\|^2 + \|P \tilde{Y}_i(t)\|^2,$$

$$\text{即 } \|R_n(t)\|^2 = \|R_v(t)\|^2 + \|P \tilde{Y}_i(t)\|^2. \quad (10)$$

式中 $\|R_v(t)\|$ 与噪信比关系密切. 噪信比愈大, $\|R_v(t)\|$ 愈大; 噪信比小, $\|R_v(t)\|$ 就接近于零, 且 $\|R_v(t)\| \ll \|P \tilde{Y}_i(t)\|$. 而 $\|P \tilde{Y}_i(t)\|$ 在持续激励下与 $(v_i - n)$ 差数关系密切.

$(v_i - n)$ 愈大, $\|P\tilde{Y}_i\|(t)$ 愈大. 记 $\dot{\sigma}_i(n) = \|R_n(t)\|/\sqrt{t}$, 并绘出 $\dot{\sigma}_i(n)$ 在 $n < v_i$ 时的曲线, 可发现它是一条斜率绝对值很大的直线. 噪信比愈小, 它就显得愈峭.

(2) $n \geq v_i$

$S(n)$ 的补投影算子作用于式 (5) 得

$$M_n \tilde{Y}_i(t) = M_n [\tilde{\Phi}(v_i)\theta] + M_n \varepsilon_i(t), \quad (10)$$

即

$$R_n(t) = M_n \varepsilon_i(t).$$

这是由于 $S(v_i)$ 是 $S(n)$ 的子空间, 导致 $M_n [\tilde{\Phi}(v_i)\theta] = 0$. 同样的原因使 $\|R_n(t)\| < \|R_v(t)\|$. 绘 $n \geq v_i$ 段的 $\dot{\sigma}(n)$ 曲线, 可以看到, 在噪信比较小时, 由于 $\|R_v(t)\|$, $\|R_n(t)\|$ 都很小, $\dot{\sigma}(n)$ 呈现为一条几乎平行且紧挨着 n 轴的直线. 它与 $n < v_i$ 段的直线 $\dot{\sigma}(n)$ 交接于折点. 此折点处 n 值便指示了被试的结构指数 v_i . 文献[2, 3, 4, 5]的残差法便是据此而提出的.

三、白化残差法

若被识系统渐近稳定, 又虑及 $v_u(t)$, $v_y(t)$ 的假设, 可知方程噪声 $\varepsilon_i(t)$ 的谱密度亦是有理的, 且它可由如下 AR 模型生成,

$$\varepsilon_i(t) = G_i(q^{-1})\varepsilon_i(t) + e_i(t).$$

在此, $G_i(q^{-1}) = \sum_{k=1}^{n_g} g_{ik} q^{-k}$. $e_i(t)$ 是白噪声. 它与 $u(\cdot)$, $y(\cdot)$ 独立, 与 t 时刻前的 $v_u(\cdot)$, $v_y(\cdot)$, $\varepsilon_i(\cdot)$ 独立. 相应地, 可由序列数据写出

$$\begin{aligned} \varepsilon_i(t) &= G_i(q^{-1})\varepsilon_i(t) + E_i(t) \\ &= \sum_{k=1}^{n_g} g_{ik} \varepsilon_i(t-k) + E_i(t). \end{aligned} \quad (11)$$

把式 (11) 代入式 (5) 得

$$\tilde{Y}_i(t) = \tilde{\Phi}(v_i)\theta + \sum_{k=1}^{n_g} g_{ik} \varepsilon_i(t-k) + E_i(t). \quad (12)$$

用 M_n , ($n \geq v_i$) 左乘此式, 再考虑到 $E_i(t)$ 总正交于 $S(n)$ 空间, 即 $M_n E_i(t) = E_i(t)$, 那末有

$$R_n(t) = \sum_{k=1}^{n_g} g_{ik} R_n(t-k) + E_i(t), \quad (n \geq v_i). \quad (13)$$

倘若记 $\Delta_n = \left\| \sum_{k=1}^{n_g} g_{ik} R_n(t-k) \right\|$, 再利用正交器等性可得

$$\|R_n(t)\|^2 = \Delta_n^2 + \|E_i(t)\|^2, \quad (n \geq v_i). \quad (14)$$

由式 (8) 或者由式 (12) 及 $S(n) = \text{span}\{\tilde{\Phi}(n)\}$ 可知, $\varepsilon_i(t-1)$, $\varepsilon_i(t-2)$, ..., $\varepsilon_i(t-1)$

3 期

n_e 既不在 $S(n)$ 中, 又不与 $S(n)$ 正交. $\varepsilon_i(t-k)$ 在 $S(n)$ 上的投影长度随 n 的增加、 $S(n)$ 的扩大而增加. 因此, Λ_n 随 n 增加而单调减地趋于零. $\|R_n(t)\|^2$ 随 n 增加单调减地趋于 $\|E_i(t)\|^2$. 这便是结构指数增大对外界干扰的自滤效应. 当 $\varepsilon_i(t)$ 的自相关函数具有较好截尾性时, 在 n 超出 v_i 的最初几个值上, 这种自滤效应的作用尤为明显. 也正由于这个原因, 在噪信比较大的情况下绘出的 $\hat{\sigma}_i(n)$ 折线往往出现两个折点, 从而导致确定结构指数困难. 本文的算例以及文[2, 3, 4]的仿真实例都表明了这个事实.

为克服残差法的缺陷, 引入 $S_R(n_e) = \text{span}\{R_n(t-k); k=1, \dots, n_e\}$ 的补投影算子 M_R . 又考虑到 $E_i(t)$ 与 $S_R(n_e)$ 正交, 所以当用 M_R 左乘式 (13) 时, 可得

$$M_R R_n(t) = E_i(t), \quad (n \geq v_i). \quad (15)$$

M_R 对残差 $R_n(t)$ 起了白化滤波作用, 并把 $M_R R_n(t) \triangleq W_n(t)$ 称为白化残差. 若令 $\hat{\sigma}_i(n) = \|W_n(t)\|/\sqrt{L}$, 那末由式 (15) 显然可见: 不管噪信比大或小, $\hat{\sigma}_i(n)$ 曲线都典型地呈现为具一折点的折线. 它的 $n \geq v_i$ 段相当好地平行于 n 轴. 折点的坐标不仅指示了被试的结构指数, 而且在不存在输入或输出观测干扰时能很好地给出输出观测噪声 $v_y(t) = D_y^{-1}(q^{-1})C_y(q^{-1})e(t)$ 或输入观测噪声 $v_u(t) = D_u^{-1}(q^{-1})C_u(q^{-1})e(t)$ 的 $e_i(t)$ 的估计值.

最后顺便指出, 在实际运算中, 由于 $R_n(t-k)$ 可由 $R_n(t)$ 简单移位而得, 因此, $S_R(n_e)$ 的补投影算子 M_R 可求取如下.

$$M_R = I - \Phi_R \Phi_R^+,$$

式中 $\Phi_R = [R_n(t-1) \dots R_n(t-n_e)]$. 至于 n_e 的选取, 与外界噪声性质有关. 当然较大 n_e 会取得较好的白化效果, 但同时计算量也增加. n_e 在许多场合下取 1 或 2 就足够了.

四、仿 真 算 例

算例取自文[1], 是 $v_1, v_2, v_3 = 1, 2, 3$ 的三输入三输出系统

$$\begin{aligned}
 A(q^{-1}) &= \begin{bmatrix} 0.2 & 0 & 0 \\ 0 & -1.1 & 0 \\ -0.5 & 0 & 0.5 \end{bmatrix} q^{-1} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & -0.28 & 0 \\ 0.1 & -0.5 & 0 \end{bmatrix} q^{-2} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0.3 \end{bmatrix} q^{-3}, \\
 B(q^{-1}) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} q^{-1} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1.4 & -0.4 \\ 0.5 & -1 & 0 \end{bmatrix} q^{-2} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0.6 & -0.25 \end{bmatrix} q^{-3}.
 \end{aligned}$$

输入 $u(t)$ 采用幅值为 1、长度为 $(2^8 - 1)$ 的 PRBS. 观测噪声 $v_y(t)$ 由标准差 $\sigma_i = 0.3$, ($i=1, 2, 3$) 的独立白噪声 $e(t)$ 形成. $e(t)$ 与 $u(t)$ 独立.

仿真中, 投影算子、残差、白化残差均采用 schmidt 正交法递推获得. 在不同观测噪声 $v_y(t)$ 下由残差法所得的 $\hat{\sigma}_i(n)$ 折线 (虚线所示) 与由白化残差法所得的 $\hat{\sigma}_i(n)$ 折线 (实线所示) 的比较由图 1、2、3、4 给出.

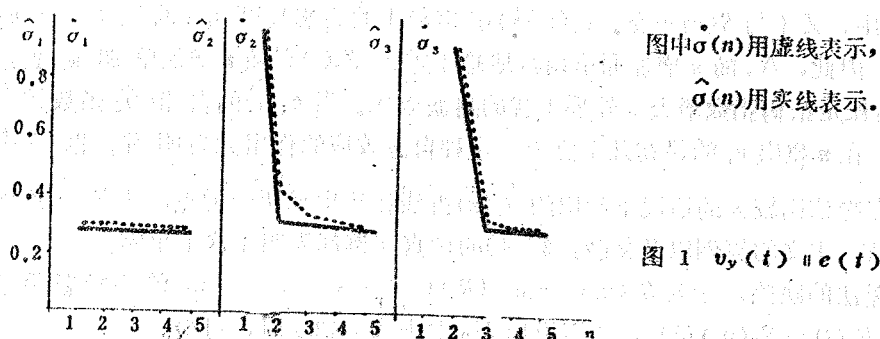
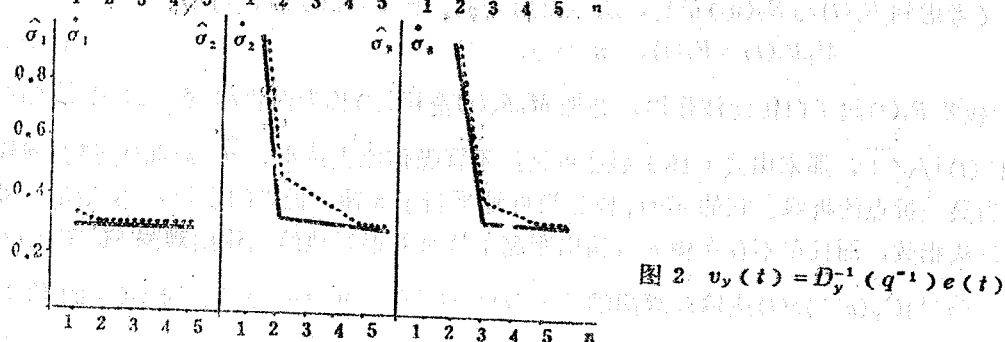
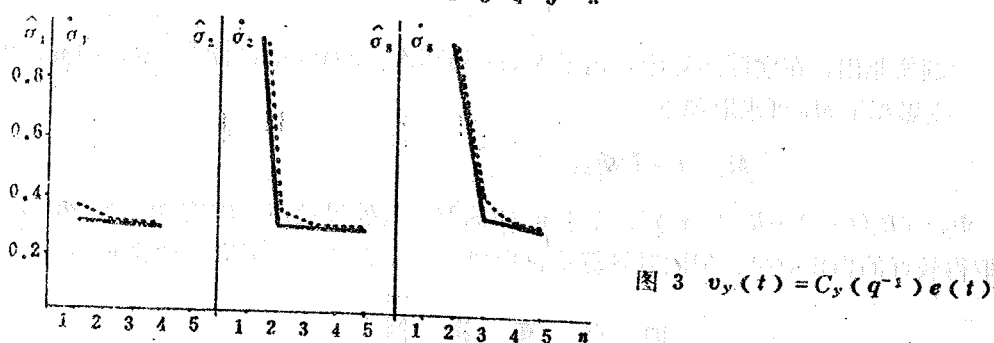
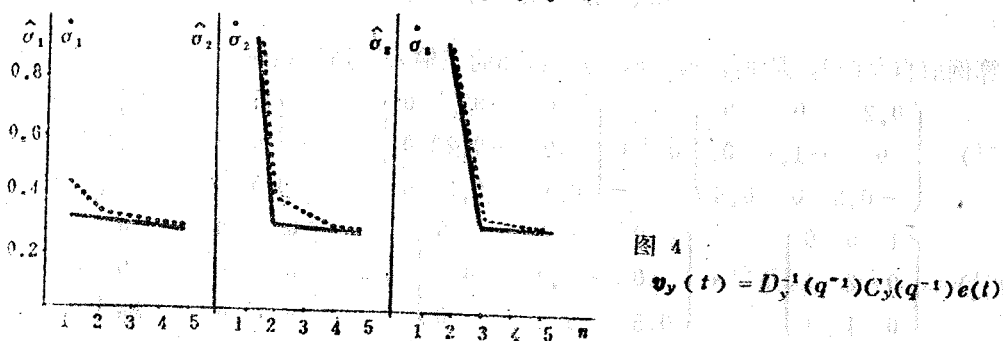
图 1 $v_y(t) = e(t)$ 图 2 $v_y(t) = D_y^{-1}(q^{-1})e(t)$ 图 3 $v_y(t) = C_y(q^{-1})e(t)$ 图 4
 $v_y(t) = D_y^{-1}(q^{-1})C_y(q^{-1})e(t)$

图1、2、3、4中的 $C_y(q^{-1})$, $D_y(q^{-1})$ 为

$$C_y(q^{-1}) = \begin{pmatrix} 1 - 0.5q^{-1} \\ 1 - 0.4q^{-1} + 0.2q^{-2} \\ 1 - 0.5q^{-1} + 0.24q^{-2} - 0.02q^{-3} \end{pmatrix}$$

$$D_y(q^{-1}) = \begin{bmatrix} 1 + 0.35q^{-1} & & \\ & 1 - 0.2q^{-1} + 0.37q^{-2} & \\ & & 1 - q^{-1} + 0.57q^{-2} - 0.082q^{-3} \end{bmatrix}.$$

参 考 文 献

- [1] Bokor, J. and Keviczky, L., Structure Properties and Structure Estimation of Vector Difference Equations, Int. J. Control, 36, (1982), 461—475.
- [2] Guidorzi, R. P., Losito, M. P. and Muratori, T., The Range Error Test in the Structural Identification of Linear Multivariable Systems, IEEE Trans. Automat. Control, AC-27, (1982), 1044—1053.
- [3] Suen, L. C. and Liu, R., Determination of the Structure of Multivariable Stochastic Linear Systems, IEEE Trans. Automat. Control, AC-23, (1978), 458—464.
- [4] El-Sherief, H. and Sinha, N. K., Determination of a Canonical Model for the Identification of Linear Multivariable Systems, IEEE Trans. Systems, man and Cyb., SMC-12, (1982), 668—673.
- [5] Sinha, N. K. and Kusza, B., Modeling and Identification of Dynamic Systems, New York, (1983).
- [6] Ljung, L. and Soderstrom, T., Theory and Practice of Recursive Identification, London, (1983).

Whitening Residual Error Technique to Determine the Structure of VDE

Zhang Zhiyong

(Institute of Automation, Fuzhou University)

Abstract

When the residual-error methods REM are applied to determine the structure indices from the contaminated observations for higher values of the noise-to-signal ratio, the estimated values of the structure indices are often larger than the actual values. After the reason of the drawback of REM is theoretically analysed with the orthogonal projection technique, the whitening residual-error method WREM is proposed. The so-called whitening residual errors WRE can be produced from a whitening projection operator, which is formed easily. Not only will the structure of the system to be identified be determined accurately, but also the performance of the model with this structure will be predicted if the WREM is used.