第 29 卷 第 2 期
Vol. 29 No. 2

控 制 与 决 策
Control and Decision

2014 年 2 月
Feb. 2014

文章编号: 1001-0920 (2014) 02-0358-05

DOI: 10.13195/j.kzyjc.2012.1546

丢失数据下的条件极大似然辨识

王建宏

(景德镇陶瓷学院 机械电子工程学院，江西 景德镇 333403)

摘　要: 针对仿射结构形式在丢失数据下的条件极大似然辨识问题, 首先引入交换矩阵将原随机矢量分解成观测和丢失部分; 然后确定出观测数据在丢失数据下的条件均值和条件方差, 以此建立条件似然函数; 进而从理论上给出了条件极大似然函数关于未知参数矢量、未知白噪声方差值和丢失数据的求导公式, 并从工程上给出一种可分离的优化算法; 最后通过仿真算例验证了该辨识方法的有效性.

关键词: 条件极大似然；丢失数据；交换矩阵；优化算法

中图分类号: TP273　　　　**文献标志码:** A

Conditional maximum likelihood identification under missing data

WANG Jian-hong

(School of Mechanical and Electronic Engineering，Jingdezhen Ceramic Institute，Jingdezhen 333403，China. E-mail：wangjianhong@nuaa.edu.cn)

Abstract: To the conditional maximum likelihood identification problem of an affine structure under missing data, a permutation matrix is used to divide a random vector into observed and missing parts. Then conditional mean and covariance under missing data are set up to obtain a conditional likelihood function. In the theory, expressions of the derivatives about the conditional maximum likelihood function on the unknown parameter vector, unknown white noise variance and missing data are derived. A separable optimum algorithm is given to be applied in engineering. Finally, simulation results show the effectiveness of the identification method.

Key words: conditional maximum likelihood；missing data；permutation matrix；optimum algorithm

0 引　言

数据丢失是工业过程中普遍存在的现象. 即使某辨识实验含有丢失数据, 也难以舍弃该辨识实验, 因每次辨识实验的实施都需要花费大量的物力和财力, 为此需要研究在丢失数据下的系统辨识问题.

文献 [1] 从时域上分析了线性系统的辨识及渐近性、收敛性; 文献 [2] 从频域上分析了线性和非线性系统的辨识; 文献 [3] 列举出各种系统辨识中广泛应用的优化方法, 如牛顿法、切平面法和捆绑法等; 文献 [4] 提出一种离散时间系统在丢失数据下的频域辨识思想; 文献 [5] 构造出连续时间系统在丢失数据下的目标准则函数; 文献 [6] 分析了ARX 系统的丢失数据辨识问题. 综上可见, 对于丢失数据的系统辨识研究目前还甚少, 迫切需要开展深入研究.

对于随机白噪声矢量的仿射结构式, 可以把观测数据与丢失数据结合作为随机矢量元素, 而未知参数矢量分别存在于其他矩阵和矢量中, 将未知参数矢量和丢失数据集同时作为未知辨识量, 辨识策略采用文献 [1] 中的极大似然估计法. 目前对于极大似然估计的研究较多, 如文献 [7-10] 系统地分析了极大似然估计的优化、渐近性及相应的滤波性. 本文采用了较少使用的条件极大似然辨识, 其难点在于, 需要利用交换矩阵分离原系统的观测数据和丢失数据, 采用统计信号处理中的条件均值、方差和矩阵论中的多种矩阵运算以构造出在丢失数据下观测数据的条件均值和条件方差式. 对此条件极大似然函数, 推导出该条件极大似然函数关于未知参数矢量、未知白噪声方差值和丢失数据的求导式, 这些导数式可应用于未知辨识量的精确优化求解. 在工程实践的可容许范围内, 结合文献 [3] 给出一种关于此三类待辨识量的可分离求解过程.

收稿日期: 2012-10-16；**修回日期:** 2013-05-12.

基金项目: 江西省科技厅青年科学基金项目(20122BAB211012).

作者简介: 王建宏(1980−), 男, 副教授, 博士, 从事系统辨识与优化等研究.

1 模型描述

考虑一 n 维的随机白噪声矢量 e, 其均值为 0, 方差矩阵为 λI_n ($\lambda > 0$) 且 I_n 为 n 维的单位矩阵. 设 n 维随机矢量 x 是白噪声 e 的仿射结构式, 有

$$\phi x + \gamma = e. \tag{1}$$

其中: 设矩阵 ϕ 是可逆矩阵, 随机矢量 x 以观测数据和丢失数据为元素, 矩阵 ϕ 和矢量 γ 由未知参数矢量 θ 构成. 由式(1)可得随机矢量 x 的无偏估计值为

$$\hat{x} = -\phi^{-1}\gamma + \phi^{-1}e = -\phi^{-1}(\gamma - e).$$

对其两端同时取期望可得随机矢量 x 的均值为

$$\mu = E\hat{x} = -\phi^{-1}\gamma.$$

进一步可得随机矢量 x 的方差矩阵为

$$\Sigma = \lambda(\phi^{\mathrm{T}}\phi)^{-1}.$$

由极大似然辨识法可知, 随机矢量 x 的概率密度函数为

$$p(x) = \frac{1}{\sqrt{(2\pi)^n \det\Sigma}} \exp\Big\{-\frac{1}{2}(x-\mu)^{\mathrm{T}}\Sigma^{-1}(x-\mu)\Big\}.$$

代入随机矢量 x 的均值和方差, 可得

$$\begin{aligned}
&p(x) = \\
&(\sqrt{(2\pi)^n \det\lambda(\phi^{\mathrm{T}}\phi)^{-1}}\,)^{-1}\times \\
&\exp\Big\{-\frac{1}{2\lambda}(x+\phi^{-1}\gamma)^{\mathrm{T}}(\phi^{\mathrm{T}}\phi)(x+\phi^{-1}\gamma)\Big\} = \\
&(\sqrt{(2\pi)^n \det\lambda(\phi^{\mathrm{T}}\phi)^{-1}}\,)^{-1}\times \\
&\exp\Big\{-\frac{1}{2\lambda}(\phi x+\gamma)^{\mathrm{T}}(\phi x+\gamma)\Big\}.
\end{aligned} \tag{2}$$

式(2)即为系统辨识中常见的极大似然目标准则函数, 对该似然函数关于未知参数矢量 θ、未知白噪声方差 λ 求偏导即可求得待辨识量. 考虑丢失数据下的系统辨识, 其随机矢量 x 既含有观测数据, 又含有丢失数据, 而式(2)的似然函数却没有反映出此特征, 为此需对式(2)的似然函数作进一步研究.

2 丢失数据下的条件似然函数

为将式(2)的随机矢量 x 分解成观测数据 x_0 和丢失数据 x_m, 引入交换矩阵 T, 其具有如下性质:

$$T^{\mathrm{T}}T = TT^{\mathrm{T}} = I,$$

使得随机矢量 x 在交换矩阵 T 的作用下为

$$\begin{bmatrix} x_0 \\ x_m \end{bmatrix} = Tx = \begin{bmatrix} T_0 \\ T_m \end{bmatrix} x.$$

利用交换矩阵的性质有

$$\phi x + \gamma = \phi T^{\mathrm{T}}Tx + \gamma = \phi T^{\mathrm{T}} \begin{bmatrix} x_0 \\ x_m \end{bmatrix} + \gamma = e. \tag{3}$$

定义式(3)中 ϕT^{T} 的分块矩阵为

$$[\phi_0\ \ \phi_m] = \phi T^{\mathrm{T}} = \phi[T_0\ \ T_m]. \tag{4}$$

将式(4)代入(3)可得

$$\begin{aligned}
\phi x + \gamma = [\phi_0\ \ \phi_m]\begin{bmatrix} x_0 \\ x_m \end{bmatrix} + \gamma = \\
\phi_0 x_0 + \phi_m x_m + \gamma = e.
\end{aligned} \tag{5}$$

类似地, 有变量

$$\begin{aligned}
&\begin{bmatrix} \mu_0 \\ \mu_m \end{bmatrix} = T\mu = \begin{bmatrix} T_0 \\ T_m \end{bmatrix}\mu, \\
&\xi = T(x-\mu) = Tx - T\mu = \\
&\begin{bmatrix} x_0 \\ x_m \end{bmatrix} - \begin{bmatrix} \mu_0 \\ \mu_m \end{bmatrix} = \begin{bmatrix} x_0-\mu_0 \\ x_m-\mu_m \end{bmatrix} = \begin{bmatrix} \xi_0 \\ \xi_m \end{bmatrix}, \\
&\xi_0 = x_0 - \mu_0,\ \xi_m = x_m - \mu_m.
\end{aligned}$$

利用上述变量, 式(2)的二次指数函数式可整理为

$$\begin{aligned}
&(x-\mu)^{\mathrm{T}}\Sigma^{-1}(x-\mu) = \\
&(Tx-T\mu)^{\mathrm{T}}T\Sigma^{-1}T^{\mathrm{T}}(Tx-T\mu) = \\
&\xi^{\mathrm{T}}T\Sigma^{-1}T^{\mathrm{T}}\xi.
\end{aligned} \tag{6}$$

式(6)中间第3项的矩阵运算为

$$\begin{aligned}
&T\Sigma^{-1}T^{\mathrm{T}} = \frac{(T\phi^{\mathrm{T}})(\phi T^{\mathrm{T}})}{\lambda} = \\
&\frac{1}{\lambda}\begin{bmatrix} \phi_0^{\mathrm{T}} \\ \phi_m^{\mathrm{T}} \end{bmatrix}[\phi_0\ \ \phi_m] = \frac{1}{\lambda}\begin{bmatrix} \phi_0^{\mathrm{T}}\phi_0 & \phi_0^{\mathrm{T}}\phi_m \\ \phi_m^{\mathrm{T}}\phi_0 & \phi_m^{\mathrm{T}}\phi_m \end{bmatrix}.
\end{aligned} \tag{7}$$

将式(7)代入(6)可得

$$\begin{aligned}
&\xi^{\mathrm{T}}T\Sigma^{-1}T^{\mathrm{T}}\xi = \\
&[\xi_0^{\mathrm{T}}\ \ \xi_m^{\mathrm{T}}]\frac{1}{\lambda}\begin{bmatrix} \phi_0^{\mathrm{T}}\phi_0 & \phi_0^{\mathrm{T}}\phi_m \\ \phi_m^{\mathrm{T}}\phi_0 & \phi_m^{\mathrm{T}}\phi_m \end{bmatrix}\begin{bmatrix} \xi_0 \\ \xi_m \end{bmatrix} = \\
&\frac{1}{\lambda}[\xi_0^{\mathrm{T}}\phi_0^{\mathrm{T}}\phi_0\xi_0 + \xi_m^{\mathrm{T}}\phi_m^{\mathrm{T}}\phi_0\xi_0 + \\
&\xi_0^{\mathrm{T}}\phi_0^{\mathrm{T}}\phi_m\xi_m + \xi_m^{\mathrm{T}}\phi_m^{\mathrm{T}}\phi_m\xi_m].
\end{aligned} \tag{8}$$

在条件极大似然辨识中需计算在丢失数据 x_m 条件下观测数据 x_0 的条件均值和条件方差. 利用统计信号处理中的条件高斯概率密度函数式, 有

$$\begin{aligned}
&\mu_1 = E(x_0/x_m) = \\
&E(x_0) + \frac{\mathrm{cov}(x_0,\ x_m)}{\mathrm{var}(x_m)}(x - E(x_m)), \\
&\Sigma_1 = \mathrm{var}(x_0/x_m) - \frac{1}{\mathrm{cov}^2(x_0,\ x_m)}\mathrm{var}(x_m).
\end{aligned} \tag{9}$$

将式(4)、(5)、(7)和(8)代入(9), 可得观测数据 x_0 的条件均值和条件方差矩阵分别为

$$\begin{aligned}
&\mu_1 = \mu_0 + (\phi_0^{\mathrm{T}}\phi_m)^{-1}(\phi_m^{\mathrm{T}}\phi_m)(x - E(x_m)), \\
&\Sigma_1 = (\phi_0^{\mathrm{T}}\phi_0)^{-1} - (\phi_0^{\mathrm{T}}\phi_m)^{-1}(\phi_m^{\mathrm{T}}\phi_m)(\phi_m^{\mathrm{T}}\phi_0)^{-1}.
\end{aligned} \tag{10}$$

根据式(10)关于观测数据 x_0 的条件均值和条件方差矩阵, 可得观测数据 x_0 的条件高斯概率密度函数为

$$\begin{aligned}
p(x_0/x_m) = \frac{1}{\sqrt{(2\pi)n_0 \det\Sigma_1}}\exp\Big\{-\frac{1}{2}(x_0 - \\
\mu_0)^{\mathrm{T}}\Sigma_1^{-1}(x_0-\mu_0)\Big\}.
\end{aligned} \tag{11}$$

n_0 为观测数据 x_0 的维数, 对应的条件似然函数为

$$
\begin{aligned}
&L(\theta,\lambda,x_m)=\\
&-\frac{n_0}{2}\log(2\pi)+\frac{1}{2}\log\varSigma_1+\\
&\frac{1}{2\lambda}(x_0-\mu_0)^{\mathrm{T}}\varSigma_1^{-1}(x_0-\mu_0)+\frac{n_0}{2}\log\lambda. \qquad (12)
\end{aligned}
$$

式(12)的条件似然函数将未知参数矢量θ, 未知白噪声方差λ及丢失数据x_m同时作为未知待辨识的优化变量. 因式(12)中的第1项不含有未知辨识量, 故可舍去, 而最后一项已是最简形式, 因此需着重分析中间两项的具体推导式. 对条件方差矩阵的行列式求解, 设

$$
A=\begin{bmatrix}(\phi_0^{\mathrm{T}}\phi_0)^{-1} & (\phi_0^{\mathrm{T}}\phi_m)^{-1}\\ (\phi_m^{\mathrm{T}}\phi_0)^{-1} & (\phi_m^{\mathrm{T}}\phi_m)^{-1}\end{bmatrix}.
$$

由矩阵计算公式有

$$
\begin{aligned}
\det A=&|(\phi_0^{\mathrm{T}}\phi_0)^{-1}||(\phi_0^{\mathrm{T}}\phi_0)^{-1}-\\
&(\phi_0^{\mathrm{T}}\phi_m)^{-1}(\phi_m^{\mathrm{T}}\phi_m)(\phi_m^{\mathrm{T}}\phi_0)^{-1}|=\\
&|(\phi_0^{\mathrm{T}}\phi_0)^{-1}|\det\varSigma_1.
\end{aligned}
$$

即有

$$
\begin{aligned}
&\det\varSigma_1=\\
&\frac{\det|(\phi_0^{\mathrm{T}}\phi_0)^{-1}(\phi_m^{\mathrm{T}}\phi_m)^{-1}-(\phi_m^{\mathrm{T}}\phi_0)^{-1}(\phi_0^{\mathrm{T}}\phi_m)^{-1}|}{\det(\phi_0^{\mathrm{T}}\phi_0)^{-1}}. \qquad (13)
\end{aligned}
$$

对式(13)取对数运算可得

$$
\begin{aligned}
&\log\det\varSigma_1=\\
&\log\det|(\phi_0^{\mathrm{T}}\phi_0)^{-1}(\phi_m^{\mathrm{T}}\phi_m)^{-1}-\\
&(\phi_m^{\mathrm{T}}\phi_0)^{-1}(\phi_0^{\mathrm{T}}\phi_m)^{-1}|+\log\det(\phi_0^{\mathrm{T}}\phi_0). \qquad (14)
\end{aligned}
$$

为计算条件方差矩阵$\varSigma_1$的逆矩阵$\varSigma_1^{-1}$, 利用矩阵求逆引理可得

$$
\begin{aligned}
\varSigma_1^{-1}=&(\phi_0^{\mathrm{T}}\phi_0)\{I+((\phi_0^{\mathrm{T}}\phi_0)^{-1}(\phi_m^{\mathrm{T}}\phi_0)\times\\
&(\phi_m^{\mathrm{T}}\phi_m)^{-1}(\phi_0^{\mathrm{T}}\phi_m)-I)^{-1}\}. \qquad (15)
\end{aligned}
$$

将式(13)～(15)分别代入条件似然函数(12), 可得极大最优化的目标准则函数为

$$
L(\theta,\lambda,x_m)=f_1+f_2+f_3+f_4. \qquad (16)
$$

其中各个量为

$$
\begin{aligned}
&f_1=\frac{n_0}{2}\log\lambda,\ f_3=\frac{1}{2}\log\det(\phi_0^{\mathrm{T}}\phi_0),\\
&f_2=\frac{1}{2}\log\det|(\phi_0^{\mathrm{T}}\phi_0)^{-1}(\phi_m^{\mathrm{T}}\phi_m)^{-1}-\\
&(\phi_m^{\mathrm{T}}\phi_0)^{-1}(\phi_0^{\mathrm{T}}\phi_m)^{-1}|,\\
&f_4=\frac{1}{2\lambda}(x_0-\mu_0)^{\mathrm{T}}\varSigma_1^{-1}(x_0-\mu_0).
\end{aligned}
$$

对式(16)中目标函数$L(\theta,\lambda,x_m)$的求解可采用文献[3]中的若干优化方法, 在任何一种优化策略中都需求解该目标函数关于未知待辨识量(θ,λ,x_m)的一阶偏导式. 略去繁琐的推导过程, 直接给出所有一阶偏导式分别为

$$
\begin{aligned}
&\frac{\partial f_1}{\partial\lambda}=\frac{n_0}{2\lambda},\ \frac{\partial f_1}{\partial x_m}=0,\\
&\frac{\partial f_2}{\partial\lambda}=\frac{\partial f_3}{\partial\lambda}=\frac{\partial f_2}{\partial x_m}=\frac{\partial f_3}{\partial x_m}=0,\\
&\frac{\partial f_4}{\partial\lambda}=-\frac{1}{2\lambda^2}(x_0-\mu_0)^{\mathrm{T}}\varSigma_1^{-1}(x_0-\mu_0)=-\frac{1}{\lambda}f_4,\\
&\frac{\partial f_4}{\partial x_m}=\frac{1}{2\lambda}\Big(\frac{\partial\mu_1}{\partial x_m}\Big)^{\mathrm{T}}\varSigma_1^{-1}(x_0-\mu_0)-\\
&\qquad\frac{1}{2\lambda}(x_0-\mu_0)^{\mathrm{T}}\varSigma_1^{-1}\Big(\frac{\partial\mu_1}{\partial x_m}\Big),\\
&\frac{\partial\mu_1}{\partial x_m}=(\phi_0^{\mathrm{T}}\phi_m)^{-1}(\phi_m^{\mathrm{T}}\phi_m),\ \frac{\partial f_1}{\partial\theta_k}=0,\\
&\frac{\partial f_2}{\partial\theta_k}=\frac{1}{2}[(\phi_0^{\mathrm{T}}\phi_0)^{-1}(\phi_m^{\mathrm{T}}\phi_m)^{-1}-\\
&\qquad(\phi_m^{\mathrm{T}}\phi_o)^{-1}(\phi_0^{\mathrm{T}}\phi_m)^{-1}]\frac{\partial *}{\partial\theta_k},\\
&\frac{\partial f_3}{\partial\theta_k}=\frac{1}{2}(\phi_0^{\mathrm{T}}\phi_0)^{-1}\Big(\frac{\partial\phi_0^{\mathrm{T}}}{\partial\theta_k}\phi_0+\phi_0^{\mathrm{T}}\frac{\partial\phi_0}{\partial\theta_k}\Big),\\
&\frac{\partial *}{\partial\theta_k}=-(\phi_0^{\mathrm{T}}\phi_0)^{-\mathrm{T}}\Big(\frac{\partial\phi_0^{\mathrm{T}}}{\partial\theta_k}\phi_0+\phi_0^{\mathrm{T}}\frac{\partial\phi_0}{\partial\theta_k}\Big)\times\\
&\qquad(\phi_0^{\mathrm{T}}\phi_0)^{-1}(\phi_m^{\mathrm{T}}\phi_0)^{-1}-(\phi_0^{\mathrm{T}}\phi_0)^{-1}\times\\
&\qquad(\phi_m^{\mathrm{T}}\phi_m)^{-1}\Big(\frac{\partial\phi_m^{\mathrm{T}}}{\partial\theta_k}\phi_m+\phi_m^{\mathrm{T}}\frac{\partial\phi_m}{\partial\theta_k}\Big)(\phi_m^{\mathrm{T}}\phi_m),\\
&\frac{\partial f_4}{\partial\theta_k}=-\frac{\partial\mu_1}{\partial\theta_k}\varSigma_1^{-1}(x_0-\mu_1)-\\
&\qquad(x_0-\mu_1)^{\mathrm{T}}\varSigma_1^{-1}\frac{\partial\mu_1}{\partial\theta_k}-\\
&\qquad(x_0-\mu_1)^{\mathrm{T}}\varSigma_1^{-\mathrm{T}}\frac{\partial\varSigma_1}{\partial\theta_k}\varSigma_1^{-1}(x_0-\mu_1),\\
&\frac{\partial\mu_1}{\partial\theta_k}=\phi_0^{-\mathrm{T}}\frac{\partial\phi_0}{\partial\theta_k}\gamma-\phi_0^{-1}\frac{\partial\gamma}{\partial\theta_k}-\\
&\qquad(\phi_0^{\mathrm{T}}\phi_m)^{-\mathrm{T}}\Big(\frac{\partial\phi_0^{\mathrm{T}}}{\partial\theta_k}\phi_m+\phi_0^{\mathrm{T}}\frac{\partial\phi_m}{\partial\theta_k}\Big)\times\\
&\qquad(\phi_0^{\mathrm{T}}\phi_m)^{-1}(\phi_m^{\mathrm{T}}\phi_m)(x_m+\phi_m^{-1})\gamma+\\
&\qquad(\phi_0^{\mathrm{T}}\phi_m)^{-1}\Big(\frac{\partial\phi_m^{\mathrm{T}}}{\partial\theta_k}\phi_m+\phi_m^{\mathrm{T}}\frac{\partial\phi_m}{\partial\theta_k}\Big)\times\\
&\qquad(x_m+\phi_m^{-1})\gamma+(\phi_0^{\mathrm{T}}\phi_m)^{-1}(\phi_m^{\mathrm{T}}\phi_m)\times\\
&\qquad\Big(-\phi_m^{\mathrm{T}}\frac{\partial\phi_m}{\partial\theta_k}\phi_m^{\mathrm{T}}\gamma+\phi_m^{-1}\frac{\partial\gamma}{\partial\theta_k}\Big).
\end{aligned}
$$

由于极大似然辨识法得到的估计值是无偏的, 且对应的方差矩阵恰等于Fisher信息矩阵的逆, 对于高斯极大似然问题, 存在该Fisher信息矩阵的闭形式

$$
\begin{aligned}
&I(\theta,\lambda,x_m)=I(\bar\theta)=\\
&\frac{\partial\mu_1}{\partial\bar\theta_k}\varSigma_1^{-1}\frac{\partial\mu_1}{\partial\bar\theta_k}+\frac{1}{2}\mathrm{tr}\Big[\varSigma_1^{-1}\frac{\partial\varSigma_1}{\partial\bar\theta_k}\varSigma_1^{-1}\frac{\partial\varSigma_1}{\partial\bar\theta_k}\Big]. \qquad (17)
\end{aligned}
$$

其中某些偏导量为

$$
\begin{aligned}
&\frac{\partial\mu_1}{\partial\lambda}=\frac{\partial\varSigma_1}{\partial\lambda}=\frac{\partial\varSigma_1}{\partial x_m}=0,\\
&\frac{\partial\varSigma_1}{\partial\bar\theta_k}=\\
&(\phi_0^{\mathrm{T}}\phi_0)^{-\mathrm{T}}\Big(\frac{\partial\phi_0^{\mathrm{T}}}{\partial\theta_k}\phi_0+\phi_0^{\mathrm{T}}\frac{\partial\phi_0}{\partial\theta_k}\Big)(\phi_0^{\mathrm{T}}\phi_0)^{-1}+
\end{aligned}
$$

$$(\phi_0^{\mathrm T}\phi_m)^{-\mathrm T}\Big(\frac{\partial\phi_o^{\mathrm T}}{\partial\theta_k}\phi_m+\phi_0^{\mathrm T}\frac{\partial\phi_m}{\partial\theta_k}\Big)\times$$
$$(\phi_0^{\mathrm T}\phi_m)^{-1}(\phi_m^{\mathrm T}\phi_m)(\phi_0^{\mathrm T}\phi_m)^{-1}-$$
$$(\phi_0^{\mathrm T}\phi_m)^{-1}\Big(\frac{\partial\phi_m^{\mathrm T}}{\partial\theta_k}\phi_m+\phi_m^{\mathrm T}\frac{\partial\phi_m}{\partial\theta_k}\Big)(\phi_m^{\mathrm T}\phi_0)^{-1}+$$
$$(\phi_0^{\mathrm T}\phi_m)^{-1}(\phi_m^{\mathrm T}\phi_m)(\phi_m^{\mathrm T}\phi_0)^{-1}\times$$
$$\Big(\frac{\partial\phi_0^{\mathrm T}}{\partial\theta_k}\phi_m+\phi_0^{\mathrm T}\frac{\partial\phi_m}{\partial\theta_k}\Big)(\phi_m^{\mathrm T}\phi_0)^{-1}.$$

3 可分离的优化求解

似然函数的多个偏导式非常复杂, 因此不能应用于工程实践中. 工程实践算法讲究算法的快速性和高效率性, 容许某些未知估计量在一定小范围内小偏差地取值. 由未知估计量取值带来的小偏差或小扰动在控制器参数设计时需进行一定的校正补偿.

这里给出一种简单可行的优化求解算法——可分离优化法. 设其他两个量为定值, 利用最优化的必要条件得到一个未知估计值, 将此估计值代入原似然函数中, 得到仅含有两种未知待辨识量的似然函数. 重复上述推导过程, 直至求解出第 3 种未知待辨识量. 取式 (16) 的似然函数 $L(\theta,\lambda,x_m)$ 关于 λ 求偏导, 得

$$\frac{\partial L(\lambda)}{\partial\lambda}=\frac{n_0}{2\lambda}-\frac{1}{2\lambda^2}(x_0-\mu_0)^{\mathrm T}\varSigma_1^{-1}(x_0-\mu_0)=0.$$

进而可解出

$$\hat\lambda=\frac{(x_0-\mu_0)^{\mathrm T}\varSigma_1^{-1}(x_0-\mu_0)}{n_0}.\tag{18}$$

将式 (18) 代入原似然函数中, 整理可得

$$L_1(\theta,x_m)=$$
$$\frac{n_0}{2}\log\frac{(x_0-\mu_0)^{\mathrm T}\varSigma_1^{-1}(x_0-\mu_0)}{n_0}+\frac{1}{2}\log\det\varSigma_1.\tag{19}$$

对式 (19) 再关于丢失数据 x_m 求偏导, 可得

$$\frac{1}{2}\frac{n_0}{(x_0-\mu_0)^{\mathrm T}\varSigma_1^{-1}(x_0-\mu_0)}\Big[\Big(\frac{\partial\mu_1}{\partial x_m}\Big)^{\mathrm T}\times$$
$$\varSigma_1^{-1}(x_0-\mu_0)+(x_0-\mu_0)^{\mathrm T}\varSigma_1^{-1}\Big(\frac{\partial\mu_1}{\partial x_m}\Big)\Big]=0.$$

对该式进行简化, 可得

$$(x_0-\mu_0)^{\mathrm T}\varSigma_1^{-1}(\phi_0^{\mathrm T}\phi_m)^{-1}(\phi_m^{\mathrm T}\phi_m)=0.$$

展开上式, 若使等式左右两边相等, 则需要满足

$$x_0=\mu_1,$$

即有

$$x_0=\mu_0+(\phi_0^{\mathrm T}\phi_m)^{-1}(\phi_m^{\mathrm T}\phi_m)^{-1}(x_m+\phi_m^{-1}\gamma).$$

整理可得丢失数据 x_m 的估计值为

$$\hat{x_m}=(\phi_m^{\mathrm T}\phi_m)^{-1}(\phi_0^{\mathrm T}\phi_m)(x_0-\mu_0)-\phi_m^{-1}\gamma.\tag{20}$$

将式 (20) 再次代入似然函数 (19), 可简化为

$$L_2(\theta)=\frac{1}{2}\log\det\varSigma_1,$$
$$\varSigma_1=(\phi_0^{\mathrm T}\phi_0)^{-1}-(\phi_0^{\mathrm T}\phi_m)^{-1}(\phi_m^{\mathrm T}\phi_m)(\phi_m^{\mathrm T}\phi_0)^{-1}.\tag{21}$$

对式 (21) 的最小化运算为无约束优化问题

$$\theta=\arg\min_\theta L_3(\theta)=\arg\min_\theta\det\varSigma_1.\tag{22}$$

对于无约束优化问题 (22), 采用最速下降法迭代地求解未知参数矢量 θ. 最速下降法的计算过程如下.

Step 1: 给出初始估计值 $\theta(0)\in R^n$, $0\leqslant\varepsilon\leqslant 1$, k 表示迭代次数;

Step 2: 计算

$$-\frac{\partial L_3(\theta)}{\partial\theta(k)}=\varSigma_1^{-1}\frac{\partial\varSigma_1}{\partial\theta(k)},$$

如果 $\|\partial L_3(\theta)/\partial\theta(k)\|\leqslant\varepsilon$, 则算法停止;

Step 3: 步长因子 a_k 的选择应保证下式成立:

$$L_3\Big(\theta(k)-a_k\frac{\partial L_3(\theta)}{\partial\theta(k)}\Big)=\min_{a>0}L_3\Big(\theta(k)-a\frac{\partial L_3(\theta)}{\partial\theta(k)}\Big);$$

Step 4: 最速下降法的迭代式为

$$\theta(k+1)=\theta(k)-a_k\frac{\partial L_3(\theta)}{\partial\theta};\tag{23}$$

Step 5: $k=k+1$, 转 Step 2.

4 仿真算例

采用两个仿真算例来说明条件极大似然辨识法在丢失数据下参数辨识的可行性.

例 1 考虑如下 ARX 系统:

$$y(k)=\frac{0.8}{1+0.7q_{-1}+0.8q_{-2}}u(k)+$$
$$\frac{0.7}{1+0.7q_{-1}+0.8q_{-2}}e(k),\tag{24}$$

其中 q^{-1} 为时移算子, ARX 系统的输入激励信号选择为一随机变量序列, 且其以等概率的形式在 ±1 中取值. 观测噪声为均值为 0、方差为 0.5 的白噪声, 输入输出观测数据序列个数取为 1 000. 该 ARX 系统中 4 个未知参数的真实值为

$$a_1=0.7,\ a_2=0.8,\ b_1=0.8,\ b_2=0.7.$$

数据采集从时刻 $t=1$ 开始到时刻 $t=1\,000$ 结束, 输出观测数据中丢失比为 40/100, 且数据以随机形式丢失. 迭代算法的初始值都选取为零, 将 4 个待辨识的未知参数分为两个有序实数对 (a_1,b_1) 和 (a_2,b_2), 则 4 个未知参数恰对应于二维平面上的两个点. 参数值的变化将引起点的上下或左右移动. 图 1 为丢失数据下 4 个未知参数的极大似然辨识仿真图. 可以看出, 4 种迭代值无限接近于平面上的两点 (0.7, 0.8) 和 (0.8,

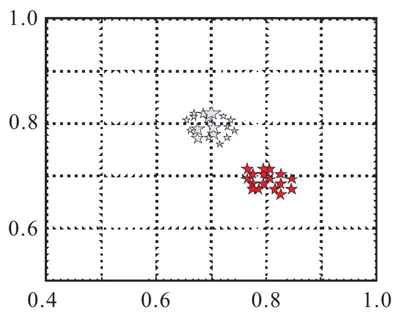

图 1　参数估计示意图

0.7). 此两点恰对应于4个参数的真实值 $a_1 = 0.7, a_2 = 0.8, b_1 = 0.8, b_2 = 0.7$.

分别在条件极大似然法和经典预测误差法下估计此4个未知参数, 所得估计值的均值和标准方差如表1所示. 由表1可见, 条件极大似然法下的估计值是无偏的, 且标准偏差程度较小.

表 1 参数估计值的采样均值和采样标准偏差

参数估计值	无偏	有偏	参数估计值	无偏	有偏
a_1 均值	0.700 3	0.723 6	b_1 均值	0.800 1	0.813 6
a_1 标准偏差	0.016 2	0.143 5	b_1 标准偏差	0.023 5	0.024 5
a_2 均值	0.800 7	0.790 9	b_2 均值	0.701 0	0.698 7
a_2 标准偏差	0.015 4	0.018 7	b_2 标准偏差	0.013 4	0.014 5

例2 因任何线性系统都可改写为仿射结构式, 故仿射结构式常用来表示飞机颤振随机模型[11]. 通过仿射结构式的辨识可得到相应的飞机颤振模态参数. 考虑输入输出观测数据同时带有噪声时, 随机矢量 x 中的各个元素均需要改写成如下仿射结构式:

$$\begin{bmatrix} \theta_1 & \theta_2 & \theta_3 & \eta_1 & \eta_2 \\ 0 & 1 & \theta_1 & 0 & \eta_1 \end{bmatrix} \begin{bmatrix} y(3) \\ y(2) \\ y(1) \\ u(2) \\ u(1) \end{bmatrix} + \begin{bmatrix} -\tilde{y}(3) \\ -\theta_1 \tilde{y}(2) \\ -\theta_2 \tilde{y}(1) \\ \eta_1 \tilde{u}(2) \\ \eta_2 \tilde{u}(1) \end{bmatrix} = \begin{bmatrix} e(3) \\ e(2) \end{bmatrix}. \tag{25}$$

式(25)即为(1)中 $\phi x + \gamma = e$ 的结构形式, 其中 $\tilde{y}(t)$、$\tilde{u}(t)$ 分别为输出观测噪声和输入观测噪声. 5个参数的实际值为

$$\theta = [\theta_1 \ \ \theta_2 \ \ \theta_3] = [0.35 \ \ 0.24 \ \ 0.13],$$

$$\eta = [\eta_1 \ \ \eta_2] = [0.28 \ \ 0.12].$$

在仿真实验中, 输入、输出观测噪声和过程噪声均采用零均值、方差为1的白噪声序列, 输入-输出观测数据对 $\{u(t), y(t)\}_{t=1,2\cdots,1\,000}$ 在闭环实验条件下获取, 数据个数为1 000. 输出观测数据的丢失比仍取40/100, 利用极大似然法在丢失观测数据下辨识这5个未知参数值. 在迭代算法的初始化时, 5个初始值仍取为零. 运用迭代算法辨识的参数随着迭代次数的变化如图2所示. 可以看出, 当迭代次数增至70次时, 各个参数值将无限地趋近于其各自对应的真实值.

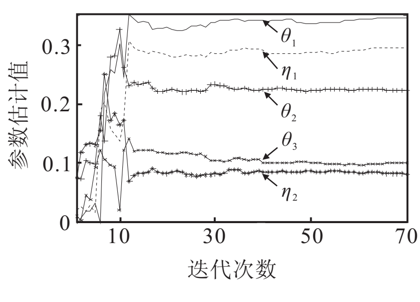

图 2 5个未知参数估计值的收敛曲线

5 结 论

对于仿射结构形式在丢失数据下的系统辨识问题, 本文分别从理论和工程上给出了优化求解算法. 但辨识过程并未分析条件极大似然辨识的准确性和渐近性, 该问题是下一步的研究内容.

参考文献(References)

[1] Ljung L. System identification: Theory for the user[M]. Upper Saddle River: Prentice Hall Press, 1999: 317-329.

[2] Pintelon R, Schoukens J. System identification: A frequency domain approach[M]. New York: IEEE Press, 2001: 208-222.

[3] Boyd S, Vandenberghe L. Convex optimization[M]. UK: Cambridge University Press, 2008: 560-578.

[4] Pintelon R. Frequency domain system identification with missing data[J]. IEEE Trans of Automatic Control, 2000, 45(2): 364-368.

[5] Pintelon R. Identification of continuous time systems with missing data[J]. IEEE Trans on Instrumentation and Measurement, 1999, 48(3): 736-740.

[6] Isaksson A. Identification of ARX models subject to missing data[J]. IEEE Trans of Automatic Control, 1993, 38(5): 813-819.

[7] Juan C Augero. Accuracy of linear multiple input multiple output models obtained by maximun likelihood estimation[J]. Automatica, 2012, 48(4): 632-637.

[8] Jakob K. A design algorithm using external perturbation to improve iterative feedback tuning convergence[J]. Automatica, 2011, 47(2): 2665-2670.

[9] Hakan Hjalmarsson, Brett Ninness. Least squares estimation of a class of frequency functions: A finite sample variance expression[J]. Automatica, 2006, 42(2): 589-600.

[10] Juan C Augero. A virtual closed loop method for closed loop identification[J]. Automatica, 2011, 47(8): 1626-1637.

[11] 王建宏. 基于先进辨识的控制策略研究及其应用[D]. 南京: 南京航空航天大学自动化学院, 2011.
(Wang J H. Research on control strategies based on advanced identification and their application[D]. Nanjing: College of Automation, Nanjing University of Aeronautics and Astronautics, 2011.)

[12] 王建宏, 王道波. 基于全局非线性可分离最小二乘法的飞机颤振模态参数辨识[J]. 振动与冲击, 2011, 30(2): 210-213.
(Wang J H, Wang D B. Subspace predictive control applied to active noise and vibration control[J]. J of Vibration and Shock, 2011, 30(2): 210-213.)

(责任编辑: 李君玲)