

文章编号: 1001-0920(2011)05-0685-05

## 状态空间模型基于正交梯度搜索的预报误差辨识

袁路生<sup>1,2</sup>, 樊晓平<sup>1</sup>, 杨辉<sup>2</sup>, 瞿志华<sup>1</sup>

(1. 中南大学 信息科学与工程学院, 长沙 410083; 2. 华东交通大学 电气与电子工程学院, 南昌 330013)

**摘要:** 提出状态空间模型的正交梯度搜索的预报误差辨识方法. 通过极小化输出预报误差而获得系统的参数估计, 提出了在观测等价类相切面正交子空间进行参数更新的实现方法, 从而解决了全参数化引起的状态空间模型参数非唯一性问题; 给出了融合预报误差局部线性逼近性能的正则化因子自适应确定方法, 揭示了算法复杂度与系统的能控性、能观测性的内在联系. 仿真结果验证了所提出方法的有效性.

**关键词:** 系统辨识; 参数估计; 状态空间模型

中图分类号: TP273

文献标识码: A

## Output error identification of LTI state-space models by orthogonal gradient search

ZHONG Lu-sheng<sup>1,2</sup>, FAN Xiao-ping<sup>1</sup>, YANG Hui<sup>2</sup>, QU Zhi-hua<sup>1</sup>

(1. College of Information Science and Engineering, Central South University, Changsha 410083, China; 2. School of Electrical and Electronic Engineering, East China Jiaotong University, Nanchang 330013, China. Correspondent: ZHONG Lu-sheng, E-mail: lszhongzju@gmail.com)

**Abstract:** An output error identification method of projected gradient search is proposed for the parameter estimation of linear time-invariant(LTI) state-space models. The system parameters are estimated by optimizing an output-error cost function. The nonuniqueness of the fully parameterized state-space system is resolved by restricting the update of the parameters to the tangent space to the manifold of observationally equivalent state-space systems. In addition, the regularization parameter is adaptively determined by considering the local linear approximation of the output error. Moreover, the relation between the computational load and system properties such as the observability and controllability is also discussed in detail. Finally, simulation results show the effectiveness of the proposed method.

**Key words:** system identification; parameter estimation; state-space models

### 1 引言

长期以来, 状态空间模型的辨识问题受到了广泛关注. 这是因为: 1) 状态空间模型具有良好的数学映射能力, 其不仅能方便地描述系统和外界输入的作用关系, 还能揭示系统的内在动力学特性; 常用的多输入-多输出模型(如 ARX 模型, ARMAX 模型等)均可由状态空间模型等价表示<sup>[1]</sup>. 2) 由输入-输出观测数据建立系统的状态空间模型是对系统进行分析、设计和控制的重要环节<sup>[2-3]</sup>.

中外学者提出了许多辨识方法来估计状态空间模型的参数. 这些辨识方法依据辨识原理可以分为两类: 规范型辨识和全参数化辨识. 规范型辨识首先

将状态空间模型等价地变换为特定的规范形式, 如能控、能观规范型<sup>[4]</sup>, 多变量传递函数矩阵<sup>[5]</sup>规范型等; 然后利用最小二乘法<sup>[4]</sup>、递阶随机梯度法<sup>[5]</sup>等算法辨识系统参数. 规范型辨识仅适用于满足能控性、能观性<sup>[4]</sup>或最小相位<sup>[5]</sup>等特定条件的系统, 这在一定程度上限制了规范型方法的应用.

全参数化辨识是把状态空间模型的系数矩阵的所有参数看作多元解空间, 通过优化选定的目标函数(如输出预报误差<sup>[6-7]</sup>、极大似然函数<sup>[8]</sup>等)得到系统参数在多元解空间的估计. 因此, 全参数化辨识可归结为一个多元函数的无约束优化问题, 只要给定系统的状态维数  $n$ , 便可采用 Newton 法、共轭梯度法

收稿日期: 2010-04-28; 修回日期: 2010-08-30.

基金项目: 国家自然科学基金项目(60870010, 60864004, 60904049); 国家863计划项目(2008AA04Z129).

作者简介: 袁路生(1979-), 男, 副教授, 从事系统辨识、信号处理的研究; 樊晓平(1961-), 男, 教授, 博士生导师, 从事智能控制、优化算法等研究.

等搜索方法通过多步迭代获得满足一定性能指标的系统参数估计<sup>[1,6-8]</sup>. 然而, 由于状态空间模型固有的相似变换特性, 导致全参数辨识在实际应用中遇到两个问题: 一是优化参数多, 计算量大<sup>[6-7]</sup>; 二是参数估计容易陷入局部极小值<sup>[1,3,8]</sup>. 目前, 围绕这两个问题已经提出了诸多改进的辨识方法. 文献[1,3]提出了正交梯度搜索的二步递阶优化辨识方法; [6]考虑算法的数值稳定性, 给出了辨识算法中Jacobian矩阵的动态确定策略; [7]提出了数据驱动的局部坐标(DDLC)辨识方法; [8]提出了基于期望极大原理的频域辨识方法. 以上方法在设计辨识算法时均采用固定的正则化因子, 难以反映估计参数与真实参数的逼近程度, 使得辨识算法在迭代后期失去寻优能力, 进而导致算法陷入局部极小值.

基于以上分析, 本文提出自适应正则化因子的正交梯度辨识方法, 给出了融合估计参数逼近真实参数信息的正则化因子自适应调整方法, 分析了辨识方法的算法复杂度, 获得了状态空间系统的能控性、能观测性与算法复杂度的关系. 最后通过仿真实验验证了所给出方法的有效性.

## 2 辨识问题描述

考虑如下离散随机状态空间系统:

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + Gv_t, \\ y_t &= Cx_t + Du_t + v_t. \end{aligned} \quad (1)$$

其中:  $x_t \in R^n$  是状态向量;  $u_t \in R^m$ ,  $y_t \in R^l$  分别是可观测的输入、输出向量; 参数矩阵  $A, B, C, D, G$  是适当维数的定常矩阵;  $v_t \in R^l$  是噪声向量, 服从高斯分布

$$v_t \sim N(0, \sigma^2 I_l), \quad (2)$$

$I_l$  表示  $l$  维的单位矩阵.

为后面描述方便, 定义参数向量

$$\theta = [\text{vec}(A)^T, \text{vec}(\hat{B})^T, \text{vec}(C)^T, \text{vec}(D)^T, \text{vec}(G)^T]^T. \quad (3)$$

其中: 符号  $\text{vec}\{\Delta\}$  表示对矩阵  $\Delta$  进行拉直运算,  $\Delta^T$  是矩阵  $\Delta$  的转置; 矩阵  $\hat{B} = (B : G)$ ,  $d = n^2 + n(m + 2l) + lm$  是待估计的系统参数总数目. 则辨识问题可表述为: 给定输入-输出观测序列  $(U_N, Y_N)$ , 其中

$$U_N = \{u_1, \dots, u_N\}, Y_N = \{y_1, \dots, y_N\}. \quad (4)$$

估计动态系统(1)的参数矩阵  $A, B, C, D, G$ , 使得辨识模型的预报误差达到最小, 即

$$J_N(\theta) = \sum_{t=1}^N \|e_t(\theta)\|_2^2 = E_N^T(\theta) E_N(\theta). \quad (5)$$

其中:  $e_t(\theta) = y_t - \hat{y}_{t|t-1}(\theta)$ ,  $E_N(\theta) = [e_1^T(\theta), \dots, e_N^T(\theta)]^T$ ,  $\hat{y}_{t|t-1}(\theta)$  表示基于  $(t-1)$  时刻对  $t$  时刻系

统输出的一步超前预报. 根据动态系统(1)的模型结构信息, 系统输出预报  $\hat{y}_{t|t-1}(\theta)$  由下式定义的均方一步最优估计器进行计算:

$$\begin{aligned} \hat{x}_{t+1|t} &= \bar{A}\hat{x}_{t|t-1} + \bar{B}u_t + Gy_t, \\ \hat{y}_{t|t-1} &= C\hat{x}_{t|t-1} + Du_t, \end{aligned} \quad (6)$$

其中

$$\bar{A} = A - GC, \bar{B} = B - GD.$$

如果动态系统(1)能控能观, 并且状态转移矩阵的特征值满足  $\lambda_1 = \max|\lambda(\bar{A})| < 1$ ,  $\lambda_2 = \max|\lambda(A)| < 1$ , 则由式(6)可得到系统状态向量以及输出向量的一致渐近估计<sup>[3]</sup>.

从上述分析可知, 通过极小化(5)定义的预报误差而得到系统的参数向量  $\theta$  估计是一个非线性、非凸的优化问题, 无法直接获得参数向量  $\theta$  的解析表达式, 因此, 需要设计辨识算法并通过多步迭代搜索得到系统的参数估计.

## 3 梯度搜索的预报误差辨识

本节首先讨论了状态空间模型的梯度优化辨识方法; 然后给出了在观测等价类垂空间更新参数的正交梯度辨识方法; 最后提出了融合估计参数逼近真实参数信息的辨识算法.

### 3.1 梯度优化辨识

为了获得系统参数(3)和预报误差(5)之间的解析关系, 本文采用局部线性化方法对预报误差(5)进行近似处理. 为此, 假定第  $k$  步迭代系统参数的估计值为  $\theta_k$ , 则误差向量(5)按一阶泰勒公式展开得

$$E_N(\theta) = E_N(\theta_k) + R(\theta_k)(\theta - \theta_k) + o(\|\theta - \theta_k\|). \quad (7)$$

其中:  $o(\cdot)$  表示高阶小量, 雅可比矩阵  $R(\theta_k) \in R^{lN \times d}$  是误差向量关于系统参数  $\theta_k$  的一阶导数, 根据导数的极限定义有

$$\begin{aligned} R(\theta_k)_{(:,i)} &= \frac{\partial E_N(\theta)}{\partial \theta} \Big|_{\theta=\theta_k} = \\ &= \lim_{\gamma \rightarrow 0} \frac{E_N(\theta_k + \gamma I_{(:,i)}) - E_N(\theta_k)}{\gamma}, \end{aligned} \quad (8)$$

其中  $I_{(:,i)}$  表示单位矩阵的第  $i$  列.

将式(7)代入(5), 并忽略高阶小量  $o(\cdot)$ , 可得

$$\begin{aligned} J_N(\theta) &\cong E_N^T(\theta_k) E_N(\theta_k) + g_k^T (\theta - \theta_k) + \\ &= \frac{1}{2} (\theta - \theta_k)^T T_k (\theta - \theta_k). \end{aligned} \quad (9)$$

其中

$$\begin{aligned} g_k &= 2R(\theta_k)^T E_N(\theta_k) \Big|_{\theta=\theta_k}, \\ T_k &= 2R(\theta_k)^T R(\theta_k) \Big|_{\theta=\theta_k}. \end{aligned} \quad (10)$$

至此, 式(5)的非线性、非凸优化问题便转化为式(9)的线性二次优化问题. 将式(9)对参数  $\theta$  求导, 并令导数为零, 得到  $\theta$  的解并赋给  $\theta_{k+1}$ , 有

$$\theta_{k+1} = \theta_k + p_k, \quad (11)$$

$$p_k = -(T_k)^\dagger g_k = -[2R(\theta_k)^T R(\theta_k)]^\dagger g_k. \quad (12)$$

其中:  $\Delta^\dagger$  是矩阵  $\Delta$  的伪逆, 向量  $p_k$  称为搜索方向.

在式(12)中, 如果  $R(\theta_k)$  的列向量线性相关, 则  $T_k$  是奇异矩阵而不可逆, 导致式(11)的参数更新方法失效. 因而, 通常采用带正则化因子的改进方法, 即

$$\begin{aligned} \theta_{k+1} &= \theta_k + p_k = \\ \theta_k &- [2R(\theta_k)^T R(\theta_k) + \lambda_k I]^\dagger g_k. \end{aligned} \quad (13)$$

其中:  $\lambda_k > 0$ ,  $I$  是单位阵.

### 3.2 正交梯度分析

需要指出, 极小化预报误差(5)获得的系统参数的解并不唯一, 这是由于状态空间模型(1)在非奇异相似变换下具有等价的输入-输出特性<sup>[1,3,6-7]</sup>, 即: 引入非奇异矩阵  $T$  对动态系统(1)的状态进行相似变换  $\tilde{x}_t = T^{-1}x_t$ , 则在观测数据集  $(U_N, Y_N)$  下参数  $\theta$  所对应的预报误差  $J_N(\theta)$  等于参数  $S(\theta, T)$  所对应的预报误差  $J_N(S(\theta, T))$  (称  $S(\theta, T)$  是  $\theta$  的观测等价类), 其中

$$\begin{aligned} S(\theta, T) &= [\text{vec}(\tilde{A}); \text{vec}(\tilde{B}); \text{vec}(\tilde{C}); \text{vec}(\tilde{D}); \text{vec}(\tilde{G})], \\ \tilde{A} &= TAT^{-1}, \tilde{B} = TB\hat{B}, \tilde{C} = CT^{-1}, \tilde{D} = D. \end{aligned} \quad (14)$$

为揭示  $S(\theta, T)$  和  $\theta$  之间的关系, 取变换矩阵  $T = I_n + \delta T$  ( $\delta T \in R^{n \times n}$  是摄动矩阵), 并利用矩阵的拉直计算<sup>[1,3,6-7]</sup>公式  $\text{vec}(ABC) = (C^T \otimes I_n)\text{vec}(B)$ , 及矩阵近似求逆公式  $(I_n + \delta T)^{-1} \simeq I_n - \delta T$ , 同时忽略关于  $\delta T$  的高阶项, 则式(14)可表示为

$$S(\theta, I_n + \delta T) = \theta + Q\text{vec}(\delta T), \quad (15)$$

其中观测等价类相切面  $Q$  的表达式为

$$Q = \begin{bmatrix} -I_n \otimes A^T + A^T \otimes I_n \\ \hat{B}^T \otimes I_n \\ -I_n \otimes C \\ 0_{lm \times n^2} \end{bmatrix}. \quad (16)$$

式(15)和(16)表明, 观测等价类  $S(\theta, T)$  可由观测等价类相切面  $Q$  的列向量所张成的子空间表示. 将这一结论应用于式(11)~(13)的参数更新, 可得: 对于任意的  $\delta T \in R^{n \times n}$ , 如果参数搜索方向  $p_k = \theta_{k+1} - \theta_k = Q_k \text{vec}(\delta T)$ , 则参数更新为  $\theta_{k+1}$  后对应的预报误差将没有任何变化, 即  $J_N(\theta) = J_N(S(\theta, T))$ , 这是梯度搜索辨识法容易陷入局部极小值的一个重要原因.

在设计算法时应该避免参数搜索方向  $p_k$  落入  $Q_k$  的列向量张成的子空间. 一个可行的策略是沿  $Q_k$  列向量的正交(垂直)子空间  $Q_k^\perp$  确定参数搜索方向. 从  $Q_k$  获得其列向量的正交子空间  $Q_k^\perp$  的数值方法有QR法、SVD法等. 以SVD分解法为例, 如果  $Q_k$  的

奇异值分解为

$$Q_k = [U_1 \ U_2] \begin{bmatrix} S_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}, \quad (17)$$

则有

$$Q_k^\perp = U_2. \quad (18)$$

其中:  $S_1 \in R^{r \times r}$  是矩阵  $Q_k$  的  $r$  个非零奇异值组成的对角矩阵;  $U_1 \in R^{d \times r}$ ,  $U_2 \in R^{d \times (d-r)}$ ,  $V_1 \in R^{n^2 \times r}$ ,  $V_2 \in R^{n^2 \times (n^2-r)}$  分别是单位正交矩阵.

类似于3.1节的分析, 预报误差在相切面的正交空间  $Q_k^\perp$  上关于估计参数的雅可比矩阵为

$$R(\theta_k^\perp)_{(:,i)} = \frac{\partial E_N(\theta)}{\partial \theta^\perp} \Big|_{\theta=\theta_k} = R(\theta_k)Q_{(:,i)}^\perp, \quad (19)$$

同时

$$g_k^\perp = 2R(\theta_k^\perp)^\perp E_N(\theta) \Big|_{\theta=\theta_k}. \quad (20)$$

综上, 可以得到  $Q_k^\perp$  空间上的系统参数更新方案, 即

$$\theta_{k+1} = \theta_k + p_k^\perp, \quad (21)$$

$$p_k^\perp = -[2R(\theta_k^\perp)^\perp R(\theta_k^\perp) + \lambda_k I]^\dagger g_k^\perp. \quad (22)$$

与式(13)相比, 式(21)和(22)具有两方面优点: 1) 式(22)能够克服因观测等价类而导致算法陷入局部极小值; 2) 式(22)能够降低算法复杂度, 提高算法效率.

### 3.3 辨识算法小结

基于梯度优化的辨识算法设计主要包括两步: 第1步是确定搜索方向  $p_k$ . 由式(13)和(22)可知, 一旦获得系统参数估计  $\theta_k$ , 则  $R(\theta_k, E_N(\theta_k))$  也随之确定, 因而搜索方向  $p_k$  仅与正则化因子  $\lambda_k$  有关. 目前还缺乏设定  $\lambda_k$  的有效方法, 已有工作通常将  $\lambda_k$  设为固定值<sup>[1-2,6,9]</sup>, 难以反映估计参数与真实参数的逼近程度, 进而影响参数收敛速度. 第2步是由一维搜索确定最佳步长. 常用的一维搜索方法有二分法、抛物线法、黄金分割点法(0.618法)等. 这些方法的共同点是通过多步试凑确定最佳步长  $\mu^*$ , 使得预报误差获得最大的下降量, 即

$$\mu^* = \arg \min(J_N(\theta_k + \mu p_k)), \mu \in (0, 1]. \quad (23)$$

记一维搜索中试凑的步数为  $n$ , 则最佳步长  $\mu^*$  是  $n$  的减函数<sup>[10]</sup>. 利用这一性质得到一种根据试凑步数  $n$  自适应确定正则化因子  $\lambda_k$  的方法, 具体为: 试凑步数  $n$  越大, 则对应的  $\mu^*$  越小, 进而使总的参数搜索方向  $\mu^* p_k$  的改变量较小, 表明此时的估计参数对真实参数的逼近性能较差, 下一步应减小参数搜索方向或增大正则化因子  $\lambda_k$ ; 反之, 试凑步数  $n$  越小, 则下一步应该减小  $\lambda_k$ . 具体实现见以下算法.

基于以上分析, 得到如下的自适应正则化因子的正交梯度优化辨识算法:

**Step 1:** 设定近似导数步长  $\gamma$  和梯度范数阈值  $\xi$ , 设置正则化因子初始值  $\lambda_0 > 0$ , 正则化因子切换倍数  $\lambda_P > 1$  以及一维搜索试凑步数切换值  $n_{\max}$ , 同时选取  $\beta \in (0, 1)$ ;

**Step 2:** 设定系统参数初始值  $\theta_0$ , 并记迭代次数标记  $k = 0$ ;

**Step 3:** 依据式 (16) 计算观测等价类相切平面  $Q_k$ , 由式 (17), (18) 的奇异值分解得到正交垂平面, 根据式 (5), (8), (19) 和 (20) 分别计算  $R(\theta_k^\perp)$  和  $g_k^\perp$ , 如果  $\|g_k^\perp\| < \xi$ , 则停止计算; 否则, 转 Step 4;

**Step 4:** 设置一维搜索试凑步数  $n = 1$ , 设置初始步长  $\mu_1 = 1$ , 根据式 (22) 计算得到  $p_k^\perp$ , 并检验条件

$$J_N(\theta_k + \mu_n Q_k^\perp p_k^\perp) \leq J_N(\theta_k) + \beta (g_k^\perp)^T p_k^\perp \quad (24)$$

是否成立, 如果不成立, 则令  $n = n + 1$ , 并由一维搜索(黄金分割点法)确定新的步长  $\mu_n$ , 转 Step 4; 否则置  $\theta_{k+1} = \theta_k + \mu_n Q_k^\perp p_k^\perp$ , 如果  $n \geq n_{\max}$ , 置  $\lambda_{k+1} = \lambda_k \lambda_P$ , 如果  $n < n_{\max}$ , 置  $\lambda_{k+1} = \lambda_k$ ;

**Step 5:** 置  $k = k + 1$ , 转 Step 3.

说明: 在由 Step 1~Step 4 确定出  $\beta, Q_k^\perp, p_k^\perp, g_k^\perp$  以后, 通过多次一维搜索总能使得式 (24) 成立, 详细的分析见文献 [1,3].

#### 4 算法复杂度分析

本文将分析上述正交梯度辨识方法的算法复杂度. 3.3 小节的辨识算法主要包括: 为获得  $Q_k^\perp$  而对  $Q_k$  进行奇异值分解(式 (17))所涉及的计算量, 记为  $n_1$ ; 得到雅可比矩阵(式 (19))需要的计算量, 记为  $n_2$ ; 确定搜索方向需要计算  $R(\theta_k^\perp)^T R(\theta_k^\perp)$  (式 (22))而产生的计算量, 记为  $n_3$ .

依据算法复杂度定义得到  $n_1, n_2, n_3$  的表达式分别为

$$\begin{aligned} n_1 &= O(d^3), \quad n_2 = O(4Nn^2(d-r)), \\ n_3 &= O(N(d-r)^2l). \end{aligned} \quad (25)$$

其中:  $n_1$  仅与待辨识的系统参数量  $d$  (式 (3)) 有关, 而  $n_2, n_3$  则由样本数  $N$ , 非零奇异值个数  $r$  (式 (17), (18)) 以及  $d$  共同确定.

为满足辨识设计中的持续激励条件, 通常需要足够大的样本量, 即  $N \gg d^3$ , 因此辨识算法复杂度主要由  $n_2$  和  $n_3$  确定, 即

$$n_{\max} = \max(4Nn^2(d-r), N(d-r)^2l). \quad (26)$$

式 (26) 表明, 算法复杂度与非零奇异值个数  $r$  有关, 并且算法复杂度是  $r$  的减函数, 要减小算法复杂度, 则需增加  $r$  的取值. 以下定理给出了  $r$  取得最大值的充要条件.

**定理 1** 式 (17) 关于  $Q_k$  的奇异值分解中, 非零

奇异值数  $r$  的最大值为  $n^2$  的充要条件是系统完全能控且完全能观测, 即

$$\text{rank}(Q_{kc}) = n, \quad \text{rank}(Q_{ko}) = n, \quad (27)$$

其中  $Q_{kc}$  和  $Q_{ko}$  分别是第  $k$  步估计  $\theta_k$  所对应的能控性矩阵与能观测矩阵, 其定义为

$$\begin{aligned} Q_{kc} &= [\hat{B}_k, A_k \hat{B}_k, \dots, A_k^{n-1} \hat{B}_k], \\ Q_{ko} &= [(C_k)^T, (C_k A_k)^T, \dots, (C_k A_k^{n-1})^T]^T. \end{aligned}$$

利用矩阵奇异值分解和线性系统理论可以完成定理 1 的证明, 具体过程从略.

定理 1 表明, 当系统满足能控性与能观测性条件时, 正交梯度搜索辨识方法的算法复杂度最低.

#### 5 仿真分析

本文应用所提出的辨识算法对状态空间模型进行数值仿真, 并将所得结果与文献 [7] 的 DDLC 方法所得到的结果进行对比分析.

考虑文献 [9] 的状态空间模型, 各参数矩阵的取值为

$$\begin{aligned} A &= \begin{bmatrix} 0.3274 & 0.4685 & -0.0608 & -0.3670 \\ 0.0379 & 0.7852 & -0.7436 & 0.1048 \\ 0.1169 & -0.0469 & 0.7596 & -0.2108 \\ 0.1272 & 0.0300 & 0.0660 & -0.3001 \end{bmatrix}, \\ B &= \begin{bmatrix} -0.0084 & -0.4142 \\ -0.0590 & 0.1041 \\ 0.0508 & 0.0516 \\ -0.0092 & -0.0949 \end{bmatrix}, \quad G = 0_{4 \times 2}, \quad D = 0_{2 \times 2}, \\ C &= \begin{bmatrix} -0.8937 & 0.1707 & 0.0906 & 0.0550 \\ -0.0862 & -0.4803 & -0.5923 & -0.4631 \end{bmatrix}. \end{aligned}$$

仿真的样本容量  $N = 500$ , 输入信号是独立同分布的正态高斯序列, 输出噪声是方差矩阵为  $\text{diag}(0.1, 0.1)$  的白噪声序列. 系统初始参数由 Matlab 函数 `drss()` 随机产生, 迭代步数设为 60.

利用本文提出的方法与文献 [7] DDLC 方法分别进行了目标函数(见式 (5))的优化分析和频率特性试验, 仿真结果如图 1 和图 2 所示.

图 1 是两种方法的目标函数与迭代步数的关系图. 从图 1 能够看出, 正交梯度辨识方法在第 9 步迭代就能使目标函数由初始值 73.36 下降到 0.015 的水平, 接近于白色噪声引起的固有偏差 0.01. 文献 [7] 方法在第 8 步迭代前对目标函数优化的效果与本文方法类似, 但第 8 步迭代后只能使目标函数缓慢下降, 迭代到第 60 步目标函数仍然为 1.362, 误差依然较大. 说明本文所提出方法的参数寻优速度快. 需要指出, 由于初始参数是随机产生的, 导致目标函数初始值比较大(为 73.36). 为减小初始目标函数值, 一般采用子空

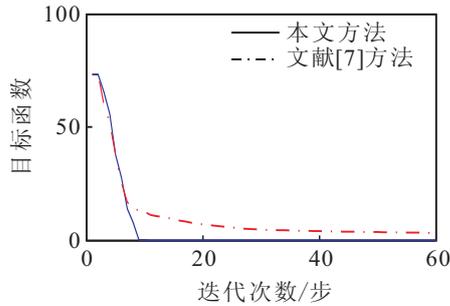
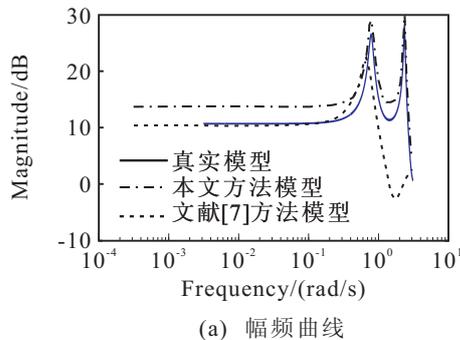


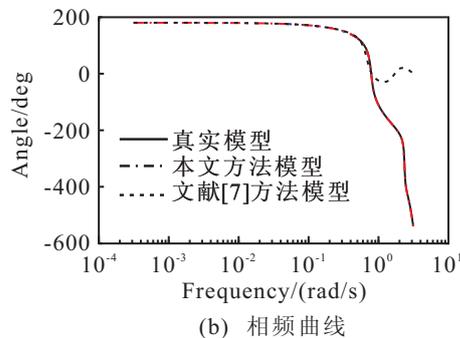
图1 目标函数与迭代步数关系

间辨识得到的模型作为仿真初始参数  $\theta_0$ .

图2是两种方法得到模型的输入1-输出1的频率响应曲线,其余的输入1-输出2,输入2-输出1等频率曲线的结果与图2类似,为节省篇幅,在此省略.仿真的频率范围是  $(0, \pi)$  弧度.由图2可知,本文方法的模型的幅频、相频曲线在整个仿真频段都能很好地拟合真实模型曲线,而文献[7]的模型频率曲线只能在低频段(小于0.8弧度)拟合真实模型曲线,而在高频段(大于0.8弧度)的幅频、相频曲线严重偏离真实模型的频率曲线.结果表明,本文方法获得的模型具有较好的频率响应性能.



(a) 幅频曲线



(b) 相频曲线

图2 模型频率特性

图1和图2的结论说明目标函数的优化性能与频率响应性能是一致的,或者说,目标函数越小的模型对应的频率响应曲线能以更高的精度拟合真实模型的频率响应曲线.

## 6 结论

本文针对状态空间模型的参数估计问题,提出了

自适应正则化因子的正交梯度搜索辨识方法.从状态空间模型的相似变换性质出发,得到了系统观测等价类的表达式,进而得到了在观测等价类相切面的正交子空间的系统参数更新策略.给出了根据一维搜索次数自适应确定正则化因子的实施方案,分析了正交梯度搜索辨识方法的算法复杂度,揭示了状态空间模型的能控性、能观测性与算法复杂度的内在联系.最后进行了目标函数优化和模型频率响应的仿真实验,结果验证了本文方法的有效性.

## 参考文献(References)

- [1] Zhong L S, Song Z H. Identification of composite local linear weighted state-space systems by orthogonal gradient search[J]. Control and Decision, 2008, 23(8): 879-882.
- [2] Verhaegen M, Dewilde P. Subspace model identification, Part 1: The output-error state-space model identification class of algorithms[J]. Int J of Control, 1992, 56(5): 1187-1210.
- [3] Zhong L S, Song Z H. Hierarchical optimization identification of LTI state-space systems by projected Gradient Search[J]. Acta Automatica Sinica, 2008, 34(6): 711-715.
- [4] 方崇智, 萧德云. 过程辨识[M]. 北京: 清华大学出版社, 1987.
- [5] Ding F, Chen T. Hierarchical gradient based identification of multivariable discrete-time systems[J]. Automatica, 2005, 41(2): 315-325.
- [6] Wills A, Ninness B. On gradient-based search for multivariable system estimates[J]. IEEE Trans on Automatic Control, 2008, 53(1): 298-306.
- [7] Mckelvey T, Helmersson A, Ribarits T. Data driven local coordinates for multivariable linear systems and their application to system identification[J]. Automatica, 2004, 40(9): 1629-1635.
- [8] Wills A, Ninness B, Gibson S. Maximum likelihood estimation of state space models from frequency domain data[J]. IEEE Trans on Automatic Control, 2009, 54(1): 19-33.
- [9] Bergboer H N, Verdult V, Verhaegen M. An efficient implementation of maximum likelihood identification of LTI state-space models by local gradient search[C]. Proc of the 41st IEEE Conf on Decision and Control. Las Vegas, Nevada, 2002: 616-621.
- [10] 邓乃扬. 无约束最优化计算方法[M]. 北京: 科学出版社, 1982.