

# 基于LVW-LM-QN算法的Wiener系统辨识

于森, 汪万里, 魏永涛

(东北大学 信息科学与工程学院, 辽宁 沈阳 110819)

**摘要:** Wiener系统由线性动态子系统与静态非线性子系统串联组成, 广泛应用于石油、化工等过程中, 获得Wiener系统的模型具有重要意义. 本文针对Wiener系统提出一种基于线性变化权重-列文伯格马夸特-拟牛顿(linear variable weight-Levenberg Marquardt-quasi Newton, LVW-LM-QN)算法的非线性系统辨识方法. 将Wiener系统分成两个子系统分别处理, 对于线性动态部分, 采用规范变量分析(canonical variate analysis, CVA)算法的子空间识别方法进行参数估计; 对于非线性静态部分, 采用LVW-LM-QN算法进行辨识处理. 最后通过数值例子和双储罐系统液位控制的应用案例来评估该方法, 仿真结果验证了所提方法的有效性和精确性.

**关键词:** Wiener系统; 子空间方法; LVW-LM-QN算法; 神经网络; CVA算法

中图分类号: N 945.14 文献标志码: A 文章编号: 1005-3026(2025)10-0027-09

## Identification of Wiener Systems Based on LVW-LM-QN Algorithm

YU Miao, WANG Wan-li, WEI Yong-tao

(School of Information Science & Engineering, Northeastern University, Shenyang 110819, China. Corresponding author: YU Miao, E-mail: yumiao@neuq.edu.cn)

**Abstract:** Wiener systems, consisting of a linear dynamic subsystem and a static nonlinear subsystem in series, find extensive application in process industries such as petroleum and chemical engineering. Obtaining the model of Wiener systems holds significant importance. A nonlinear system identification method based on the linear variable weight-Levenberg Marquardt-quasi Newton (LVW-LM-QN) algorithm for Wiener systems was proposed. The Wiener system was divided into two subsystems for separate processing. For the linear dynamic part, the subspace identification method with the canonical variate analysis (CVA) algorithm was used for parameter estimation, whereas for the subsequent nonlinear static part, the LVW-LM-QN algorithm was employed for identification. Finally, the method was evaluated through numerical examples and an application case of liquid level control in a two-tank system, and the effectiveness and accuracy of the proposed method were verified by the simulation results.

**Key words:** Wiener system; subspace method; LVW-LM-QN algorithm; neural network; CVA algorithm

系统辨识是根据收集到的输入和输出数据集建立数学模型. 近几十年, 线性系统的研究领域取得了一些重大进展. 然而, 非线性系统广泛存在于工业过程中, 由于非线性系统的复杂性和多样性, 建立全局模型有困难, 因此非线性系统的辨识仍然具有挑战性.

面向块的模型, 例如Wiener<sup>[1]</sup>, Hammerstein, Hammerstein-Wiener<sup>[2]</sup>等, 用简单的块来构建模型, 以便找到足够灵活的结构来表示许多实际非线性系统<sup>[3-4]</sup>. 本文考虑建立Wiener系统, 它是面向块的非线性模型的典型块结构, 其特点是线性动态部分后面是静态非线性块.

收稿日期: 2024-03-15

基金项目: 国家自然科学基金资助项目(62003082); 河北省自然科学基金资助项目(F2021501018); 河北省教育厅科学技术研究资助项目(ZD2022148).

作者简介: 于森(1986—), 女, 黑龙江佳木斯人, 东北大学副教授.

工业过程的许多非线性特征都可以用 Wiener 系统来描述,例如连续搅拌釜反应器<sup>[5]</sup>、热交换系统<sup>[6]</sup>等. Wiener 系统的优点在于线性块包含了系统动力学的复杂性,而非线性的复杂性仅包含在静态块中,这一优点降低了辨识的难度,因为可以应用线性和非线性系统理论的分而治之策略来分别处理两个子系统<sup>[7]</sup>.有许多方法来辨识 Wiener 系统,区别主要在于动态线性块和静态非线性块的表示方式.用来表示动态线性块的结构包括阶跃响应模型<sup>[8-9]</sup>、外源输入自回归(ARX)模型<sup>[8]</sup>、状态空间模型<sup>[10]</sup>等.对非线性块提出的结构有分段线性函数<sup>[10]</sup>、多项式函数<sup>[11]</sup>、支持向量机<sup>[12]</sup>等.

多项式函数经常被用来表示 Wiener 系统的非线性部分,但对于具有高度非线性增益的过程,它存在一定限制<sup>[13]</sup>.还有许多性能更强大的非线性建模方法,例如在 Wiener 系统构造中引入 B 样条网络<sup>[14]</sup>和神经网络<sup>[15]</sup>.近年来,神经网络已成为分类和回归领域的常用方法<sup>[16]</sup>,对于黑盒模型识别具有几个有吸引力的特性:强大的非线性拟合能力、泛化能力良好、可以并行计算、具有鲁棒性. Mohsen 等<sup>[17]</sup>应用神经网络将非线性行为过程的子空间识别方法扩展到 Wiener 结构.

本文采用 CVA 子空间方法辨识的状态空间方程与考虑偏置参数的神经网络相结合构建了 Wiener 系统,该模型是 Mohsen 等<sup>[17]</sup>提出的 Wiener 系统的改进版本.并提出了 LVW-LM-QN 算法来识别 Wiener 系统的非线性静态部分,引入了线性变化权重,提高了算法效率和性能,数值例子以及应用案例的仿真结果验证了所提方法的有效性和优越性.

## 1 问题描述

Wiener 系统由动态线性块串联静态非线性块组成,如图 1 所示.其中,  $u(k)$ ,  $y(k)$ ,  $v(k)$  分别表示模型的输入、输出和噪声信号;  $x(k)$  表示线性状态空间子系统的状态向量;  $A$ ,  $B$ ,  $C$  和  $D$  是适当维度的系统矩阵;  $\varpi(k)$ ,  $s(k)$  是未知的中间变量;  $F(\cdot)$  表示系统的输出非线性函数.

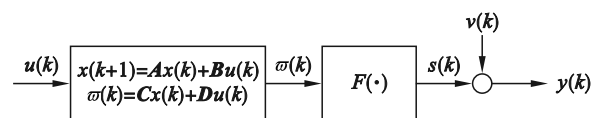


图 1 Wiener 系统框图

Fig. 1 Block diagram of Wiener system

系统的线性子系统以状态空间的形式描述为

$$\begin{cases} x(k+1) = Ax(k) + Bu(k), \\ \varpi(k) = Cx(k) + Du(k). \end{cases} \quad (1)$$

输出非线性子系统表示为

$$s(k) = F(\varpi(k)). \quad (2)$$

系统输出则为

$$y(k) = s(k) + v(k). \quad (3)$$

对图 1 的 Wiener 系统,有以下假设.

假设 1: 线性子系统是渐进稳定的,即其特征值严格位于单位圆内.

假设 2: 线性子系统是最小实现的(即对  $(A, C)$  是能观的,对  $(A, B)$  是能控的).

假设 3: 系统输入  $u(k)$  和系统输出  $y(k)$  是可用的持续激励信号<sup>[18]</sup>.

假设 4: 扰动  $v(k)$  是均值为零、方差为  $\sigma^2$  的高斯白噪声<sup>[19]</sup>.

用多层感知器神经网络来表示  $F(\cdot)$ , 可以得到非线性动态子系统为神经网络形式的 Wiener 模型,如图 2 所示,其中非线性部分的  $w_1(k) = [w_{11}(k) \cdots w_{1m}(k)]^T$ ,  $w_2(k) = [w_{21}(k) \cdots w_{2m}(k)]^T$  表示神经网络的权重,  $b_1(k)$ ,  $b_2(k) \in \mathbf{R}$  表示神经网络的偏置,  $\Phi(\cdot) \in \mathbf{R}^m$  表示隐藏层神经元的激活函数.

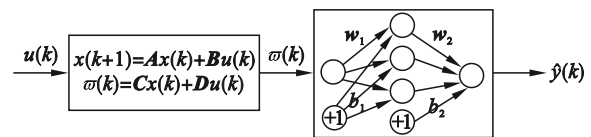


图 2 非线性子系统为神经网络的 Wiener 模型  
Fig. 2 Wiener model of nonlinear subsystem as a neural network

用该模型来辨识非线性动态系统,模型输出为  $\hat{y}(k) = w_2(k)^T \Phi(w_1(k)\varpi(k) + b_1(k)) + b_2(k)$ . (4)

辨识过程的成本函数定义为

$$V(k) = \frac{1}{2} f(k)^2, \quad (5)$$

$$f(k) = y(k) - \hat{y}(k). \quad (6)$$

$f(k)$  表示系统输出  $y$  和 Wiener 系统输出  $\hat{y}$  之间的误差.

本文的目标是使用所提出的 LVW-LM-QN 辨识方法最小化式(5),从而得到 Wiener 模型参数.

## 2 Wiener 系统辨识算法

### 2.1 Wiener 系统线性部分辨识

子空间方法在辨识中对模型结构先验知识

需求较少,在数值计算中具有一定的鲁棒性,以及广泛适用于多变量系统,因此子空间方法得到了辨识和控制领域的广泛关注.

考虑到假设1~3,子空间识别方法利用测量的输入输出信号序列,确定式(1)表示的状态空间子系统的系统矩阵  $A, B, C$  和  $D$  以及系统阶数.

下列输入输出矩阵方程在子空间识别的过程中发挥了重要作用,它表示未来输出  $Y_f$  块 Hankel 矩阵与未来输入块 Hankel 矩阵  $U_f$  和未来状态序列  $X_f$  线性相关:

$$Y_f = \Gamma_f X_f + H_f U_f, \quad (7)$$

其中扩展可观矩阵  $\Gamma_f$ :

$$\Gamma_f \stackrel{\text{def}}{=} \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{i-1} \end{bmatrix}. \quad (8)$$

下三角托普利茨矩阵  $H_f$  定义为

$$H_f \stackrel{\text{def}}{=} \begin{bmatrix} D & 0 & 0 & \cdots & 0 \\ CB & D & 0 & \cdots & 0 \\ CAB & CB & D & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ CA^{i-2}B & CA^{i-3}B & CA^{i-4}B & \cdots & 0 \end{bmatrix}. \quad (9)$$

输入输出块 Hankel 矩阵定义为

$$U_{0|j-1} \stackrel{\text{def}}{=} \begin{bmatrix} u_0 & u_1 & \cdots & u_{j-1} \\ u_1 & u_2 & \cdots & u_j \\ \vdots & \vdots & \ddots & \vdots \\ u_{i-1} & u_i & \cdots & u_{i+j-2} \end{bmatrix}, \quad (10)$$

$$Y_{0|j-1} \stackrel{\text{def}}{=} \begin{bmatrix} y_0 & y_1 & \cdots & y_{j-1} \\ y_1 & y_2 & \cdots & y_j \\ \vdots & \vdots & \ddots & \vdots \\ y_{i-1} & y_i & \cdots & y_{i+j-2} \end{bmatrix}. \quad (11)$$

简写为

$$\begin{aligned} U_p &\stackrel{\text{def}}{=} U_{0|j-1}, U_f \stackrel{\text{def}}{=} U_{i|2i-1}, \\ Y_p &\stackrel{\text{def}}{=} Y_{0|j-1}, Y_f \stackrel{\text{def}}{=} Y_{i|2i-1}, \end{aligned} \quad (12)$$

其中:下标  $p$  和  $f$  分别表示过去和未来. 包含输入  $U_p$  和输出  $Y_p$  的矩阵称为  $W_p$ :

$$W_p \stackrel{\text{def}}{=} \begin{bmatrix} Y_p \\ U_p \end{bmatrix}. \quad (13)$$

将状态序列  $X_f$  表示为

$$X_f \stackrel{\text{def}}{=} (x_i \ x_{i+1} \ x_{i+2} \ \cdots \ x_{i+j-1}). \quad (14)$$

子空间识别算法始终包含两个步骤:第1步对先前定义的Hankel矩阵生成的行空间进行加权投影,以得到扩展可观性矩阵  $\Gamma_i$  或未知系统状态序列  $X_i$  的估计  $\tilde{X}_i$ ;第2步从扩展可观性矩阵  $\Gamma_i$  或状态序列估计  $\tilde{X}_i$  中检索出系统矩阵  $A, B,$

$C, D.$

使用几何工具如下:

运算符  $\Pi_B$  表示将某一矩阵的行空间投影到矩阵  $B \in \mathbf{R}^{q \times j}$  的行空间上:

$$\Pi_B \stackrel{\text{def}}{=} B^T \cdot (BB^T)^+ \cdot B. \quad (15)$$

$A/B$  表示将矩阵  $A \in \mathbf{R}^{p \times j}$  的行空间投影到矩阵  $B$  的行空间:

$$A/B \stackrel{\text{def}}{=} A \cdot \Pi_B = AB^T \cdot (BB^T)^+ \cdot B. \quad (16)$$

而  $\Pi_{B^\perp}$  是将某一矩阵的行空间投影到矩阵  $B$  的行空间的正交补上的几何运算符:

$$A/B^\perp \stackrel{\text{def}}{=} A \cdot \Pi_{B^\perp}. \quad (17)$$

$$\Pi_{B^\perp} = I_j - \Pi_B. \quad (18)$$

投影  $\Pi_B$  和  $\Pi_{B^\perp}$  的组合将矩阵  $A$  分解为行空间正交的2个矩阵:

$$A = A \cdot \Pi_B + A \cdot \Pi_{B^\perp} = A/B + A/B^\perp. \quad (19)$$

子空间识别算法的具体步骤:首先将  $Y_f$  的行空间投影到  $U_f$  行空间的正交补  $U_f^\perp$  上:

$$Y_f/U_f^\perp = \Gamma_f X_f/U_f^\perp + H_f U_f/U_f^\perp. \quad (20)$$

由式(15)~式(18),有

$$\begin{aligned} U_f/U_f^\perp &= U_f \cdot \Pi_{U_f^\perp} = U_f \cdot (I_j - \Pi_{U_f}) = \\ &U_f - U_f U_f^T \cdot (U_f U_f^T)^+ \cdot U_f = 0. \end{aligned} \quad (21)$$

于是  $Y_f/U_f^\perp = \Gamma_f X_f/U_f^\perp.$

使用矩阵  $W_1$  和  $W_2$  对该投影向左和向右加权:输入  $U_f$  和加权矩阵  $W_1$  和  $W_2$  不能任意选择,应满足以下2个条件:

$$\text{rank}(W_1 \cdot \Gamma_f) = \text{rank } \Gamma_f, \quad (22)$$

$$\text{rank}(X_f/U_f^\perp \cdot W_2) = \text{rank } X_f. \quad (23)$$

保证  $\Gamma_f X_f$  在投影到  $U_f^\perp$  上且由矩阵  $W_1, W_2$  加权后,其秩- $n$  特性不变.如果满足上述两个条件,可得

$$O_i \stackrel{\text{def}}{=} W_1 \cdot Y_f/U_f^\perp \cdot W_2 = W_1 \cdot \Gamma_f X_f/U_f^\perp \cdot W_2. \quad (24)$$

通过SVD分解:

$$O_i = (U_1 \ U_2) \begin{pmatrix} S_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix}. \quad (25)$$

得到以下重要属性:

$$\text{rank } O_i = n, \Gamma_f = W_1^{-1} U_1 S_1^{1/2}, X_f/U_f^\perp \cdot W_2 = S_1^{1/2} V_1^T. \quad (26)$$

显然,从矩阵  $W_1 \cdot Y_f/U_f^\perp \cdot W_2$  的奇异值分解中可以获得系统阶数  $n$ .此外,从对应于非零奇异值的左奇异向量可以找到扩展的可观测性矩阵  $\Gamma_i$  的相似变换,而右奇异向量包含了状态序列  $X_i$  的相关信息.适当选择权重矩阵  $W_2$ ,则

$$\tilde{X}_i \stackrel{\text{def}}{=} X_i/U_f^\perp \cdot W_2 \quad (27)$$

可被视为状态序列  $X_i$  的估计.

只要选择合适的权重矩阵  $W_1$  和  $W_2$ , LTI (linear-time-invariant) 系统的所有子空间算法都可以在上述框架中解释, 包括 N4SID, CVA, MOESP, IV-4SID 和 basic-4SID.

根据系统参数矩阵不同的估计方式, 可以将系统参数矩阵的计算方法分为两类: 一类是通过提取状态序列估计值  $\tilde{X}_i$  的列子空间检索系统矩阵, 为回归法, 如 N4SID, CVA; 另一类是通过提取扩展可观矩阵  $\Gamma_i$  的行子空间检索系统矩阵, 为实现法, 如 MOESP, IV-4SID, basic-4SID.

本文使用 CVA 算法来辨识线性动态系统, 算法原理如下.

已知 CVA 算法的权重矩阵取值:

$$W_1 = [(Y_f/U_f^+) \cdot (Y_f/U_f^+)^T]^{-1/2}, \quad (28)$$

$$W_2 = (W_p/U_f^+)^+ \cdot (W_p/U_f^+). \quad (29)$$

由式(20)~式(27)得到  $\tilde{X}_i$ , 通过类似的推理, 容易证明:

$$O_{i+1} \stackrel{\text{def}}{=} \bar{W}_1 \cdot Y_f / U_f^+ \cdot \bar{W}_2 = \bar{W}_1 \cdot \Gamma_{i-1} \cdot \tilde{X}_{i+1} \cdot \bar{W}_2. \quad (30)$$

其中:  $\bar{W} = W O_i$ ;  $U_f^+ = U_{i+1|2i-1}$ ;  $Y_f = Y_{i+1|2i-1}$ .

$$U_{i+1|2i-1} \stackrel{\text{def}}{=} \begin{pmatrix} u_{i+1} & u_{i+2} & \cdots & u_{i+j} \\ \vdots & \vdots & & \vdots \\ u_{2i-1} & u_{2i} & \cdots & u_{2i+j-2} \end{pmatrix}, \quad (31)$$

$$Y_{i+1|2i-1} \stackrel{\text{def}}{=} \begin{pmatrix} y_{i+1} & y_{i+2} & \cdots & y_{i+j} \\ \vdots & \vdots & & \vdots \\ y_{2i-1} & y_{2i} & \cdots & y_{2i+j-2} \end{pmatrix}. \quad (32)$$

很容易得出, 如果去掉由式(26)计算出的  $\Gamma_i$  的最后  $l$  (系统输出的数量) 行, 便得到  $\Gamma_{i-1}$ :

$$\Gamma_{i-1} = \underline{\Gamma}_i. \quad (33)$$

此时  $\tilde{X}_{i+1}$  为

$$\tilde{X}_{i+1} = (\bar{W}_1 \cdot \Gamma_{i-1}) \times O_{i+1} \bar{W}_2^+. \quad (34)$$

已经得到了  $\tilde{X}_i, \tilde{X}_{i+1}$ , 只使用系统的输入输出数据, 可以通过最小二乘法求解一组简单的超定方程来得到状态空间矩阵  $A, B, C$  和  $D$ :

$$\begin{pmatrix} \tilde{X}_{i+1} \\ Y_{ij} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} \tilde{X}_i \\ U_{ij} \end{pmatrix}. \quad (35)$$

在 Wiener 系统的辨识中, CVA 算法利用系统的输入输出数据辨识出模型的系统矩阵  $A, B, C$  和  $D$ , 即模型的线性动态块.

## 2.2 Wiener 系统非线性部分辨识

Gauss-Newton (GN) 算法是解决非线性最优问题的常见算法之一<sup>[20]</sup>, 它不直接计算 Hessian 矩阵  $H$ , 而是通过 Jacobi 矩阵  $J$  对  $H$  进行近似 ( $H = J^T J$ ):

$$h_{gn} = -(J(k)^T J(k))^{-1} g(k), \quad (36)$$

$$w_{new} = w + h_{gn}. \quad (37)$$

其中:  $w$  为神经网络权重;  $h_{gn}$  为 GN 算法的迭代步长;  $k$  是当前迭代次数;  $g = J^T f = \frac{\partial V}{\partial w}$  为成本函数  $V$  对  $w$  的梯度. Gauss-Newton 算法减少了计算量和所需的存储空间, 但如果  $J$  不是满秩矩阵, Gauss-Newton 算法不收敛.

因此引入单位矩阵  $H = J^T J + \mu I$  来保证近似矩阵的可逆性, 就得到了 Levenberg Marquardt (LM) 算法:

$$h_{lm} = -(J(k)^T J(k) + \mu I)^{-1} g(k), \quad (38)$$

$$w_{new} = w + h_{lm}. \quad (39)$$

其中: 正则化参数  $\mu \geq 0$ ;  $h_{lm}$  为 LM 算法的迭代步长.

假设  $\hat{w}$  为局部最小值点, 如果  $f(k)$  在这一点等于 0, LM 算法在  $\hat{w}$  处是二倍收敛<sup>[21]</sup>, 否则为线性收敛<sup>[21]</sup>.

为了得到更好的性能, Loke 等<sup>[22]</sup>将 LM 算法与 QN 算法相结合, 提出了一种混合算法 (LM-QN), 在  $f(k)$  于  $\hat{w}$  处  $\neq 0$  的条件下仍是超线性收敛<sup>[21]</sup>. QN 算法基于矩阵  $B$  对矩阵  $H$  进行近似:

$$h_{qn} = -B(k)^{-1} g(k), \quad (40)$$

$$w_{new} = w + h_{qn}. \quad (41)$$

其中:  $B$  为 Hessian 矩阵  $H$  的近似矩阵;  $h_{qn}$  为 QN 算法的迭代步长.

基本 LM-QN 算法的学习率是一个常数, 通常取 1. 然而对于一些复杂的问题, 由于存在许多局部解, LM-QN 算法可能在寻找全局最优解时性能不佳, 过早陷入局部最优解. 早熟收敛降低了优化效果, 为了避免这一缺陷, 本文引入线性变化权重 LVW (linear variable weight)<sup>[23]</sup> 策略, 应用于学习率  $\alpha$ :

$$\alpha(k) = \alpha_{max} - \frac{\alpha_{max} - \alpha_{min}}{k_{max}} k, \quad (42)$$

$$\text{LM: } w_{new} = w + \alpha(k) h_{lm}, \quad (43)$$

$$\text{QN: } w_{new} = w + \alpha(k) h_{qn}. \quad (44)$$

其中:  $k_{max}$  表示最大迭代次数;  $\alpha_{max}$  和  $\alpha_{min}$  分别表示学习率的最大值和最小值.

LVW-LM-QN 辨识算法的迭代过程如下:

LM-QN 算法的权重  $w$  在第 1 次迭代时, 算法默认为 LM, 根据式 (38) 计算出迭代步长  $h_{lm}$ , 再通过式 (43) 得到  $w_{new}$ , 此时根据增益率公式计算出  $\rho$  值:

$$\rho = (V(k) - V(k+1)) / (L(0) - L(h_{lm})). \quad (45)$$

该式表示函数值的实际下降与预测下降之间的比率,  $L$  为在当前迭代值  $w$  的邻域对  $V$  值的近似模型:

$$V(k+1) \approx L(h) \equiv V(k) + \mathbf{h}^T \mathbf{g} + \frac{1}{2} \mathbf{h}^T \mathbf{B} \mathbf{h}. \quad (46)$$

如果  $\rho$  值大于 0, 说明此次迭代使得  $V$  减小, 于是  $w$  更新为  $w_{\text{new}}$ . 否则, 不会进行  $w$  值的更新. 再根据式(47)进行判断:

$$\|g(k+1)\|_{\infty} < 0.02V(k+1). \quad (47)$$

如果式(47)连续满足 3 次, 表明获取的权重越来越接近  $w^+$  (全局最小值即  $w^+ = \operatorname{argmin}_w \{F(w)\}$ ), 其中  $g(w^+) = 0$  且  $V(w^+)$  显著非 0, 这可能导致缓慢的线性收敛<sup>[21]</sup>, 优化效果差, 于是会在下一次迭代时选择算法为 QN.

算法为 QN 时, 根据式(40), 式(44)计算得到  $w_{\text{new}}$ , 此时将检验是否满足:

$$V(k+1) < V(k), \quad (48)$$

$$V(k+1) \leq (1 + \delta)V(k), \quad (49)$$

$$\|g(k+1)\|_{\infty} < \|g(k)\|_{\infty}. \quad (50)$$

其中  $\delta$  为浮点数的相对精度, 值为  $2.220 \times 10^{-16}$ . 式(48)的满足表明成本函数  $V$  减小; 式(49)的满足表明此时允许  $V$  值略微增加, 因为 QN 算法的主要作用在于使  $g(k)$  更接近于 0; 式(50)的满足说明  $g(k)$  的范数值在快速下降. 当式(48)和式(49)其中之一满足, 同时式(50)也满足时, 迭代算法具有超线性收敛的良好性能, 于是  $w$  更新为  $w_{\text{new}}$ , 否则不更新. 若

$$\|g(k+1)\|_{\infty} \geq \|g(k)\|_{\infty}, \quad (51)$$

说明  $V$  的梯度的范数下降得不够快, 此步迭代 QN 算法作用不明显, 收敛性能不佳, 于是将在下一次迭代时再切换算法为 LM.

该算法可以为解决许多非线性问题提供一种高效且具有前瞻性的方案, 也为寻找最优解提供了新的可能性.

1) 所有非线性优化方法都是迭代的: 从起点  $w_0$  (即权重初始值) 开始生成一系列向量  $w_1, w_2, \dots$ , 逐渐收敛到  $\hat{w}$ , 即给定函数的局部最小值. 当  $w_0$  远离  $\hat{w}$  时, 希望该方法向  $\hat{w}$  稳步迭代, 在迭代的全局阶段, 仅要求除了最初的迭代步骤之外  $e_k$  没有增加, 即  $\|e_{k+1}\| < \|e_k\|$ , 其中  $e_k = w_k - \hat{w}$ .

但在迭代的最后阶段,  $w_k$  越来越接近  $\hat{w}$ , 希望得到更快的收敛速度, 于是有

线性收敛:

$$\|e_{k+1}\| / \|e_k\| = a, \text{ 其中 } k \rightarrow \infty, 0 < a < 1. \quad (52)$$

超线性收敛:

$$\|e_{k+1}\| / \|e_k\| \rightarrow 0, \text{ 其中 } k \rightarrow \infty. \quad (53)$$

二倍收敛:

$$\|e_{k+1}\| = O(\|e_k\|^2), \text{ 其中 } k \rightarrow \infty. \quad (54)$$

2) 本文使用拟牛顿法更新矩阵  $\mathbf{B}$ , 其数值效果好, 并且具有全局收敛性和超线性收敛速度:

$$h(k) = w_{\text{new}} - w, \quad (55)$$

$$s(k) = \mathbf{J}(k+1)\mathbf{J}(k)\mathbf{h}(k) + (\mathbf{J}(k+1) - \mathbf{J}(k))^T f(k+1). \quad (56)$$

如果  $\mathbf{h}(k)^T \mathbf{s}(k) > 0$ , 则对矩阵  $\mathbf{B}$  作以下更新:

$$\mathbf{v}(k) = \mathbf{B}(k)\mathbf{h}(k). \quad (57)$$

$$\mathbf{B}(k+1) = \mathbf{B}(k) + \left( \frac{\mathbf{s}(k)}{\mathbf{h}(k)^T \mathbf{s}(k)} \right) \mathbf{s}(k)^T - \left( \frac{\mathbf{v}(k)}{\mathbf{h}(k)^T \mathbf{v}(k)} \right) \mathbf{v}(k)^T. \quad (58)$$

为简便表示, LVW-LM-QN 算法的伪代码如表 1 所示.

表 1 Wiener 系统 LVW-LM-QN 辨识算法  
Table 1 LVW-LM-QN identification algorithm for Wiener system

常量初始化: 迭代最大值 $k_{\text{max}}$ , 神经元权重 $w_1, w_2$ , 偏置参数 $b_1, b_2$ , 算法停止标准 $e_1, e_2$ , 学习率最值 $\alpha_{\text{max}}, \alpha_{\text{min}}$ , $\mu$ 值, 矩阵 $\mathbf{B}$
found = ( $\ g(k)\ _{\infty} \leq e_1$ )
While(not found) and ( $k \leq k_{\text{max}}$ )
$k = k + 1$ ;
通过式(42)计算学习率
if method = LM
通过式(38)计算 $h_{1m}$
初始化信赖域半径 $\Delta$
if $\ h_{1m}\  \leq e_2(\ w\  + e_2)$
found = true
else
通过式(43)更新 $w$
通过式(45)计算增益率 $\rho$
if $\rho > 0$
better = true
if 连续 3 次迭代都满足式(47), 则 method 切换到 QN
end if
end if
更新 $\mu$ 值
end if
else if method = QN
通过式(40)计算
if $\ h_{qn}\  \leq e_2(\ w\  + e_2)$
found = true

续表 1

```

else
    if  $\|h_{qn}\| > \Delta$ 
         $h_{qn} = (\Delta / \|h_{qn}\|) \cdot h_{qn}$  (步长超出信
        赖域的处理)
    end if
    通过式(44)更新  $w$ 
    better=式(48)式(49)and式(50)
    if 式(51)满足
        则 method 切换到 LM
    end if
    更新  $\Delta$  值
end if
if better=true
     $w = w_{new}$ 
end if
通过式(55)~式(58)更新  $B$ 
end While
    
```

所提辨识方法框图如图 3 所示. 辨识过程可概括如下: 子空间方法以状态空间形式来估计实际系统输入输出数据的线性动态近似, 该初始估计构成了 Wiener 系统的线性动态子系统(LDS). 前馈神经网络构成非线性静态子系统(NSS), 其权重使用 LDS 的输出数据和系统输入输出数据进行迭代训练. LDS 的参数是使用 CVA 算法确定的, 而 NSS 的参数是使用 LVW-LM-QN 算法

获得的, 从而获得 Wiener 整体模型参数. 该方法的主要优点之一是只需要系统的输入输出数据.

### 3 仿真实验

#### 3.1 数值示例

为了评估所提出的算法, 考虑以下带有色噪声干扰的 Wiener 系统:

$$x(k+1) = \begin{bmatrix} -1.5 & -3 \\ 3 & -1 \end{bmatrix} x(k) + \begin{bmatrix} 1.3 \\ 0 \end{bmatrix} u(k),$$

$$\varpi(k) = [1.15 \ 2.3] x(k),$$

$$\xi(k) = \varpi(k) + 0.3[\varpi(k)]^2 + 0.4[\varpi(k)]^3,$$

$$y(k) = \xi(k) + v(k).$$

仿真时, 输入  $u(k)$  采用单位方差的不相关零均值随机信号序列,  $v(k)$  采用方差  $\sigma^2=0.01$  的零均值高斯噪声序列. 通过使用 Matlab 软件仿真, 收集了 2 400 对输入输出样本的数据集 (见图 4).

使用前 600 对数据作为训练集, 剩余的 1 800 对数据作为验证集. 对于线性动态子系统(LDS)的辨识使用 CVA 算法<sup>[24]</sup>, CVA 算法先对系统训练集的输入输出块 Hankel 矩阵生成的行空间进行加权投影, 根据式(20)~式(28)得到未知系统状态序列  $X_i$  的估计  $\hat{X}_i$ , 再根据加权投影(式(30)), 由式(28)~式(35)得到  $\hat{X}_{i+1}$ , 此时求解最小二乘问题即可得到状态空间矩阵.

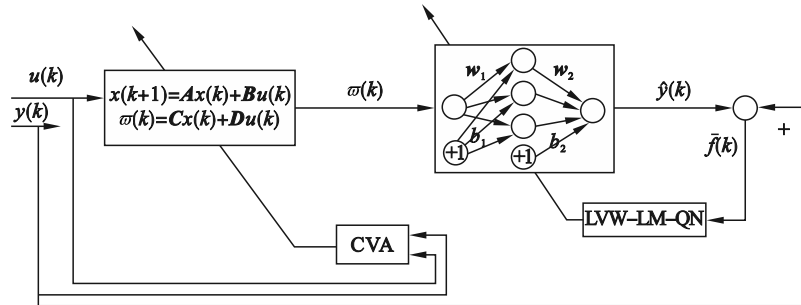


图 3 所提辨识方法框图

Fig. 3 Block diagram of proposed identification method

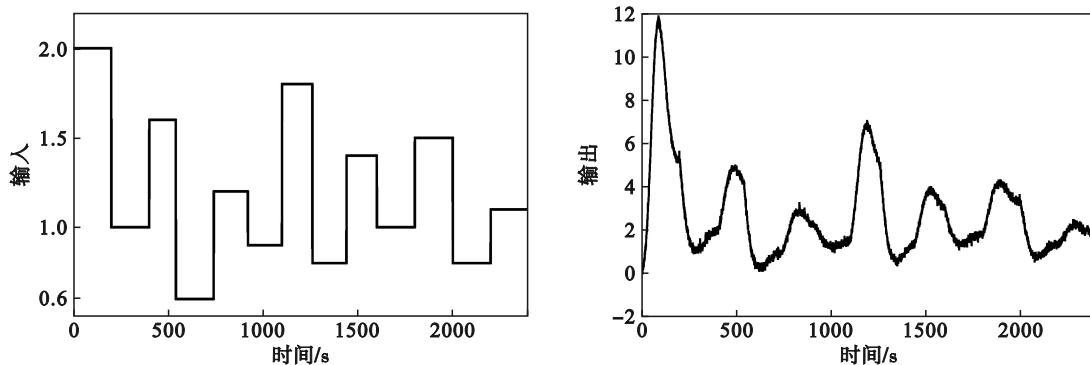


图 4 Wiener 系统的输入输出数据

Fig. 4 Input and output data of Wiener system

训练集的输入数据经过 CVA 辨识出的系统矩阵  $A, B, C, D$  构成的线性动态块, 得到中间变量  $\tilde{\omega}(k)$ , 即非线性静态块的输入. NSS 块中神经元的激活函数选择为  $\frac{1}{1+e^{-1.5x}}$ , 突触权重向量 ( $w_1, w_2$ ) 和偏置参数 ( $b_1, b_2$ ) 在训练开始时设定为  $[0, 1]$  之间的随机数. LVW-LM-QN 算法使用  $\tilde{\omega}(k)$  和系统输出数据, 通过反馈回来的误差  $f(k)$  进行迭代训练, 目的是最小化成本函数  $V$ , 最终得到构成 NSS 的多层感知器神经网络的权重  $w$ . 只需要训练集的输入输出数据, 便得到了 Wiener 模型在训练集上的线性块和非线性块的参数.

图 5 为本文所提出的 LVW-LM-QN 方法、Mohsen 等<sup>[17]</sup> 提出的 LM 方法以及 Gómez 等<sup>[25]</sup> 提出的 Subspace-based 方法在验证集的输出. 如图 5 所示, 本文方法优于 LM 以及 Subspace-based 方法. 为了评估所提出方法的准确性, 引入了均方根误差 (RMSE) 和平均绝对百分比误差 (MAPE):

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}, \quad (59)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{\hat{y}_i - y_i}{y_i} \right|. \quad (60)$$

其中:  $N$  为样本总数;  $y_i$  和  $\hat{y}_i$  分别是第  $i$  个样本的实际值和模型估计值. RMSE 和 MAPE 越小说明辨识效果越好. 辨识结果与实际值的均方根误差和平均绝对百分比误差如表 2 所示.

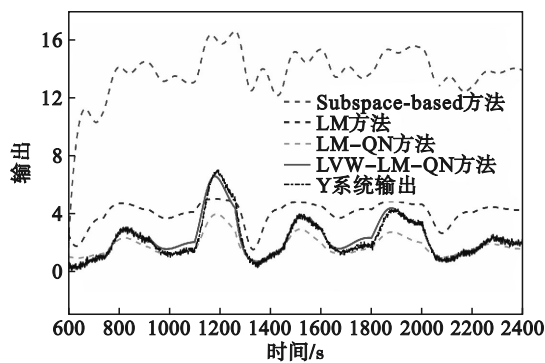


图 5 Wiener 系统和不同方法的输出信号

Fig. 5 Output signals of Wiener system and different methods

从表 2 可以看出, 相比其他两种方法, LVW-LM-QN 方法的 RMSE 和 MAPE 最小, 因此, 所提方法在数值例子中具有优越的辨识性能.

### 3.2 双储罐系统的液位控制

双储罐系统中的液位控制<sup>[26]</sup>是大多数过程

工业的组成部分, 例如化工厂、炼油厂、发电厂等. 液位控制的主要目的是根据施加到泵上的电压的变化, 控制罐中所需的液体液位. 在此示例中, 采样时间为 1 s, 使用 Matlab 软件收集了 3 000 对输入输出数据样本. 图 6 显示了该过程的输入和输出信号. 输入  $u(k)$  是施加到泵上的电压, 水泵产生流入上部水箱的流量, 水经过上部储罐底部的一个小孔进入下储罐, 下储罐的液位是双储罐系统的输出  $y(k)$ .

表 2 不同识别方法的均方根误差及平均绝对百分比误差

Table 2 RMSE and MAPE of different identification methods

方法	RMSE	MAPE
Subspace-based	10.574 6	7.377 9
LM	1.627 3	1.281 1
LVW-LM-QN	0.497 4	0.317 1

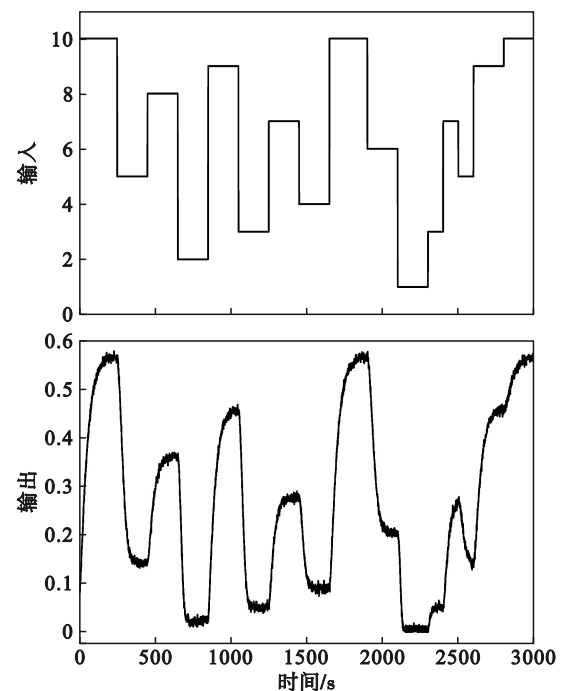


图 6 双储罐系统的输入输出数据

Fig. 6 Input and output data of two-tank system

使用前 500 对数据作为训练集, 剩余的 2 500 对数据作为验证集. 使用 CVA 算法辨识线性动态子系统 (LDS), 先对训练集的输入输出数据构成的块 Hankel 矩阵生成的行空间进行加权投影, 再根据式 (20)~式 (35) 得到未知系统状态序列  $X_i$  的估计  $\tilde{X}_i$  和  $\tilde{X}_{i+1}$ , 此时求解一组超定方程即可得到状态空间矩阵  $A, B, C$  和  $D$ .

训练集的输入数据经过 CVA 辨识的系统矩阵  $A, B, C, D$  构成的线性动态块, 得到中间变量  $\tilde{\omega}(k)$ ,

即非线性静态块的输入.NSS块中神经元的激活函数选择为  $\Phi(x) = \frac{1 - e^{-0.5x}}{1 + e^{-0.5x}}$ , 突触权重向量 ( $w_1, w_2$ ) 和偏置参数 ( $b_1, b_2$ ) 在训练开始时设定为  $[0, 1]$  之间的随机数.LVW-LM-QN 算法使用  $\tilde{\omega}(k)$  和系统输出数据以及误差  $f(k)$  进行迭代训练, 不断减小成本函数  $V$ , 最终实现  $V$  的最小化, 同时得到构成 NSS 的多层感知器神经网络的权重  $w$ . Wiener 模型在训练集数据上的线性块和非线性块的参数只需要训练集的输入输出数据就可辨识得到.

图 7 将所提出的 LVW-LM-QN 方法与 LM 方法以及 Subspace-based 方法在验证集上获得的输出进行比较. 所提出的方法明显更优. 为了进一步验证 LVW-LM-QN 的有效性, 辨识结果与实际值的均方根误差和平均绝对百分比误差如表 3 所示.

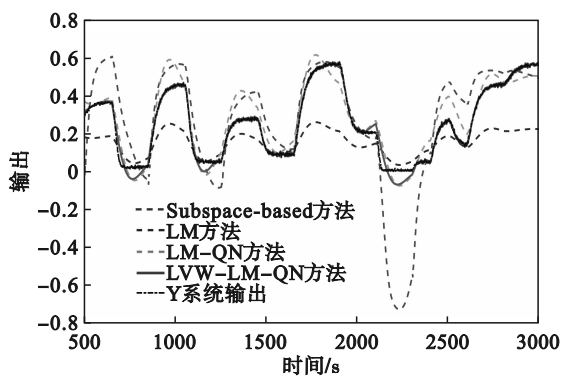


图 7 双储罐系统和不同方法的输出信号

Fig. 7 Output signals of two-tank system and different methods

表 3 不同识别方法的均方根误差及平均绝对百分比误差

Table 3 RMSE and MAPE of different identification methods

方法	RMSE	MAPE
Subspace-based	0.116 6	0.627 7
LM	0.123 2	0.854 3
LVW-LM-QN	0.042 5	0.242 7

从表 3 可以看出, 所提方法的误差指标明显低于其他两种方法, 体现了其在实际应用中的优越辨识性能, 这说明 LVW-LM-QN 方法可以实现 Wiener 模型的有效辨识. 另外, 所提算法辨识的 Wiener 模型结构简单. 线性动态子系统(LDS)的最佳近似值为 2 阶, 非线性静态子系统(NSS)具有 1 个输入、1 个输出和 3 个具有正切双曲激活函数的神经元.

## 4 结 语

本文研究 Wiener 系统的辨识, 提出了使用面向块的 Wiener 系统来辨识非线性动态系统的方法, 该模型在工程系统中具有广泛的应用前景和很强的实用性. 采用子空间方法来辨识模型的线性动态子系统, 多层感知器神经网络(MLP)则被用来表示静态非线性子系统. Wiener 模型的两个子系统的参数分别使用 CVA 和 LVW-LM-QN 方法辨识. 所提方法为 Wiener 系统辨识高度非线性系统提供了有效的辨识算法. 仿真实验证明了所提方法具有良好的辨识性能.

### 参考文献:

- [1] 陈山, 宋樱, 房胜男, 等. 基于头脑风暴优化算法的 Wiener 模型参数辨识[J]. 控制与决策, 2017, 32(12): 2291-2295.  
(Chen Shan, Song Ying, Fang Sheng-nan, et al. Parameter identification of Wiener systems using brain storm optimization algorithm[J]. *Control and Decision*, 2017, 32(12): 2291-2295.)
- [2] 白晶, 毛志忠, 浦铁成. 多变量 Hammerstein-Wiener 模型的参数辨识[J]. 东北大学学报(自然科学版), 2018, 39(1): 6-10.  
(Bai Jing, Mao Zhi-zhong, Pu Tie-cheng. Parameter identification of multivariate Hammerstein-Wiener model [J]. *Journal of Northeastern University (Natural Science)*, 2018, 39(1): 6-10.)
- [3] Mao Y W, Ding F. Multi-innovation stochastic gradient identification for Hammerstein controlled autoregressive autoregressive systems based on the filtering technique [J]. *Nonlinear Dynamics*, 2015, 79(3): 1745-1755.
- [4] Ding F, Liu X M, Liu M M. The recursive least squares identification algorithm for a class of Wiener nonlinear systems [J]. *Journal of the Franklin Institute*, 2016, 353(7): 1518-1526.
- [5] Benítez M, Bermúdez A, Rodríguez-Calo J F. Adjoint method for parameter identification problems in models of stirred tank chemical reactors [J]. *Chemical Engineering Research and Design*, 2017, 123: 214-229.
- [6] Ghani F, Waser R, O'Donovan T S, et al. Non-linear system identification of a latent heat thermal energy storage system[J]. *Applied Thermal Engineering*, 2018, 134: 585-593.
- [7] Hsu Y L, Wang J S. A Wiener-type recurrent neural network and its control strategy for nonlinear dynamic applications [J]. *Journal of Process Control*, 2009, 19(6): 942-953.
- [8] Norquay S J, Palazoglu A, Romagnoli J A. Model predictive control based on Wiener models [J]. *Chemical Engineering Science*, 1998, 53(1): 75-84.
- [9] Norquay S J, Palazoglu A, Romagnoli J A. Application of Wiener model predictive control to a pH neutralization experiment [J]. *IEEE Transactions on Control Systems Technology*, 1999, 7(4): 437-445.
- [10] Shafiee G, Arefi M M, Jahed-Motlagh M R, et al. Nonlinear predictive control of a polymerization reactor based on piecewise linear Wiener model [J]. *Chemical*

- Engineering Journal*, 2008, 143: 282–292.
- [11] Andrzej J, Józef K. Two-stage instrumental variables identification of polynomial Wiener systems with invertible nonlinearities [J]. *International Journal of Applied Mathematics and Computer Science*, 2019, 29(3): 571–580.
- [12] Ma J H, Zhao L H, Han Z Z, et al. Identification of Wiener model using least squares support vector machine optimized by adaptive particle swarm optimization [J]. *Journal of Control Automation and Electrical Systems*, 2015, 26: 609–615.
- [13] Qian S Y, Ding Z Y, Li F. Estimation of Wiener model based on neural fuzzy network [C] // 2023 IEEE 12th Data Driven Control and Learning Systems Conference. Xiangtan, 2023: 1377–1380.
- [14] Hong X, Chen S. Modeling of complex-valued Wiener systems using B-spline neural network [J]. *IEEE Transactions on Neural Networks*, 2011, 22(5): 818–825.
- [15] 吴德会. 非线性动态系统的Wiener神经网络辨识法[J]. *控制理论与应用*, 2009, 26(11): 1192–1196.  
(Wu De-hui. Identification method for nonlinear dynamic system using Wiener neural network [J]. *Control Theory & Applications*, 2009, 26(11): 1192–1196.)
- [16] Fukushima K. Neocognitron: a self organizing neural network model for a mechanism of pattern recognition unaffected by shift in position [J]. *Biological Cybernetics* 1980, 36(4): 193–202.
- [17] Mohsen S, Mohammad F. Online identification of nonlinear dynamic systems by Wiener system using subspace method and neural networks [J]. *Transactions of the Institute of Measurement and Control*, 2016, 40(2): 666–674.
- [18] Narendra K S, Annaswamy A M. Persistent excitation in adaptive systems [J]. *International Journal of Control*, 1987, 45(1): 127–160.
- [19] Bai E W. An optimal two-stage identification algorithm for Hammerstein-Wiener nonlinear systems [J]. *Automatica*, 1998, 34(3): 333–338.
- [20] Frandsen R, Poul E, Kristian J, et al. Unconstrained optimization [M]. Lyngby: Technical University of Denmark, 1999.
- [21] Chiuso A. On the asymptotic properties of closed-loop CCA-type subspace algorithms: equivalence results and role of the future horizon [J]. *IEEE Transactions on Automatic Control*, 2010, 55(3): 634–649.
- [22] Loke M H, Dahlin T. A combined Gauss-Newton and quasi-Newton inversion method for the interpretation of apparent resistivity pseudo sections [C] // The 3rd Meeting of the European Association for Environmental and Engineering Geophysics. Aarhus, 1997: 139–142.
- [23] Bansal J C, Singh P K, Saraswat M, et al. Inertia weight strategies in particle swarm optimization [C] // 2011 Third World Congress on Nature and Biologically Inspired Computing. Salamanca, 2011: 633–640.
- [24] Larimore W. Canonical variate analysis in identification, filtering, and adaptive control [C] // Proceedings of the 29th IEEE Conference on Decision and Control. Honolulu, 1990: 596–604.
- [25] Gómez J C, Baeyens E. Subspace-based identification algorithms for Hammerstein and Wiener models [J]. *European Journal of Control*, 2005, 11(2): 127–136.
- [26] Chinthaned N, Sanposh P. Robust geometric control of a two-tank system [C] // 2016 13th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology. Chiang Mai, 2016: 1–4.