第 34 卷　第 1 期　　华侨大学学报（自然科学版）　　Vol. 34　No. 1
2013 年 1 月　　Journal of Huaqiao University (Natural Science)　　Jan. 2013

文章编号：1000-5013(2013)01-0049-07

一种基于非线性系统的直接加权优化辨识算法

姚杰，朱永红

（景德镇陶瓷学院 机械电子工程学院，江西 景德镇 333403）

摘要： 对于非线性系统的直接加权优化辨识算法，通过在原线性仿射函数形式中增加若干项关于输入观测数据序列的线性项，来增强逼近非线性，减少逼近的时间．对于增加若干线性项后展开式中的未知权重值的选取，分别从理论和实用上推导出这些未知权重值的选取过程．理论上的推导分析，可明确增加的未知权重值在整个逼近非线性系统的目的中起着辅助作用；实用上的推导分析，将某些复杂的最优化问题经过整理变换成常见的最优化问题，从而可利用最为基础的优化方法来求解．

关键词： 系统辨识；非线性系统；权重值；最优化问题；迭代

中图分类号： TP 273　　**文献标志码：** A

系统辨识只是利用数学的方法，从输入、输出观测数据序列中提炼出对象的数学模型，为控制器的设计作充分的模型准备基础．根据系统对象模型的类别，可将系统辨识的理论研究分为线性系统辨识和非线性系统辨识．线性系统辨识的研究较为成熟，已有现成的理论分析基础和系统辨识仿真软件可直接调用[1-10]；而非线性系统的辨识研究也正日益开展．文献[7]对多种特殊的非线性系统展开研究，如 Wiener 系统、Hammerstein 系统和二者之间的任意组合形式，提出了诸如最小概率法、互方差辅助变量法、盲极大似然法等多种辨识方法．文献[10]分析了正交基函数的构造方法，即当采用所构造的正交基函数形式来表示原非线性系统时，有限脉冲响应模型、Laruerre 模型和双参数 Kautz 模型都可以作为该正交基函数模型结构的特例．文献[3]中提出了一种新的非线性系统辨识方法——直接加权优化法．文献[4]将直接加权优化辨识方法的基本思想应用于对分段仿射系统中各个权重值的辨识．文献[5]分析了直接加权优化辨识法对某参数的摄动所带来的影响．本文在文献[3]的思想基础之上，进行直接加权优化辨识方法的研究．

1　问题描述

给定观测数据$\{\boldsymbol{\varphi}(t), y(t)\}_{t=1}^{N}$，其非线性系统可描述为

$$y(t) = f_0(\boldsymbol{\varphi}(t)) + e(t). \tag{1}$$

式(1)中：$f_0(\boldsymbol{\varphi}(t))$称为未知待辨识估计的非线性系统函数，$\boldsymbol{\varphi}(t)$为回归矢量．通常情况下，$\boldsymbol{\varphi}(t)$有两种形式分别对应于非线性有限脉冲响应形式和外部输入下的非线性自回归形式，即

$$\left.\begin{aligned}\boldsymbol{\varphi}(t) &= [y(t-1), \cdots, y(t-n_e)]^{\mathrm{T}}, \\ \boldsymbol{\varphi}(t) &= [u(t-1), \cdots, u(t-n_u); e(t-1), \cdots, y(t-n_e)]^{\mathrm{T}}.\end{aligned}\right\}$$

其中：$e(t)$为零均值的独立同分布的随机白噪声，对应的方差值记为σ_e^2．

设非线性系统$f_0(\boldsymbol{\varphi}(t))$的一个逼近线性仿射函数形式为

$$\hat{f}(\boldsymbol{\varphi}^*(t)) = a_0 + \sum_{t=1}^{N} a_t u(t) + \sum_{t=1}^{N} b_t y(t). \tag{2}$$

将式(2)与文献[3]中对应的式子相对比，可知式(2)增加了 N 个关于输入观测数据序列$\{u(t)\}_{t=1}^{N}$

收稿日期： 2012-06-15

通信作者： 姚杰(1982-)，男，讲师，主要从事模式识别、系统辨识的研究．E-mail：njyaojie@yahoo.com.cn.

基金项目： 国家自然科学基金资助项目(61164014)

的线性项,从而增加了 N 个待求解的未知权重值$\{a_t\}_{t=0}^{N}$. 由此可见,文献[3]中的线性仿射函数形式是本文中的特例,其中特殊之处在于本文中所有的$\{a_t\}_{t=1}^{N}$全取为零值.

因此,文中以下内容的目的就在于如何确定出由这 $2N+1$ 个未知权重值构成的参数矢量 $\boldsymbol{\theta}$. 即

$$\boldsymbol{\theta}=[a_0,a_1,\cdots,a_N;b_1,\cdots,b_N]^{\mathrm{T}}. \tag{3}$$

2 直接加权优化辨识

为了能够在任意指定点 $\boldsymbol{\varphi}^*(t)$处都得到良好的逼近程度,用 $\hat{f}(\boldsymbol{\varphi}^*(t))$来近似地逼近原非线性系统 $f_0(\boldsymbol{\varphi}(t))$,逼近的准确度依赖于权重值$\{a_t\}_{t=1}^{N}$和$\{b_t\}_{t=1}^{N}$的选取. 则建立衡量逼近性能的目标函数为

$$W(\boldsymbol{\varphi}^*,f_0,\boldsymbol{\theta})=E[\hat{f}(\boldsymbol{\varphi}^*(t))-f_0(\boldsymbol{\varphi}^*(t))]^2. \tag{4}$$

将式(2)代入到式(4)中,可得

$$W(\boldsymbol{\varphi}^*,f_0,\boldsymbol{\theta})=E[a_0+\sum_{t=1}^{N}a_tu(t)+\sum_{t=1}^{N}b_ty(t)-f_0(\boldsymbol{\varphi}^*(t))]^2. \tag{5}$$

将式(1)代入式(5)中,可得展开后的目标函数为

$$W(\boldsymbol{\varphi}^*,f_0,\boldsymbol{\theta})=[a_0+\sum_{t=1}^{N}a_tu(t)+\sum_{t=1}^{N}b_ty(t)-f_0(\boldsymbol{\varphi}^*(t))]^2+\sigma_e^2\sum_{t=1}^{N}b_t^2. \tag{6}$$

式(6)中的数学期望运算过程利用了噪声 $e(t)$为独立同分布的白噪声假设条件,并且白噪声 $e(t)$与输入、输出观测数据序列$\{u(t),y(t)\}_{t=1}^{N}$均为不相关. 引入符号表示 $\widetilde{\boldsymbol{\varphi}}(t)=\boldsymbol{\varphi}(t)-\boldsymbol{\varphi}^*(t)$后,在式(6)中进行增加和减去相同的两项,等式不变,可得

$$\begin{aligned}W(\boldsymbol{\varphi}^*,f_0,\boldsymbol{\theta})=&[a_0+\sum_{t=1}^{N}a_tu(t)+\sum_{t=1}^{N}b_t(f_0(\boldsymbol{\varphi}(t))-f_0(\boldsymbol{\varphi}^*(t))-\\&\nabla f_0(\boldsymbol{\varphi}^*(t))\widetilde{\boldsymbol{\varphi}}(t))+f_0(\boldsymbol{\varphi}^*(t))\times\\&(\sum_{t=1}^{N}(b_t-1)+\nabla f_0(\boldsymbol{\varphi}^*(t))\sum_{t=1}^{N}b_t\widetilde{\boldsymbol{\varphi}}(t)]^2+\sigma_e^2\sum_{t=1}^{N}b_t^2.\end{aligned} \tag{7}$$

式(7)中得到的平方项即为通常的平方偏差,后一项为由未建模引起的方差误差项. 从式(7)可看出:偏差项将会变得任意大,除非增加关于权重值$\{b_t\}_{t=1}^{N}$的两个约束条件:$\sum_{t=1}^{N}b_t=1$,$\sum_{t=1}^{N}b_t\widetilde{\boldsymbol{\varphi}}(t)=0$. 在上述两个施加的约束条件下,式(7)的目标函数可以简化为

$$\begin{aligned}W(\boldsymbol{\varphi}^*,f_0,\boldsymbol{\theta})=&[a_0+\sum_{t=1}^{N}a_tu(t)+\sum_{t=1}^{N}b_t(f_0(\boldsymbol{\varphi}(t))-\\&f_0(\boldsymbol{\varphi}^*(t))-\nabla f_0(\boldsymbol{\varphi}^*(t))\widetilde{\boldsymbol{\varphi}}(t))]^2+\sigma_e^2\sum_{t=1}^{N}b_t^2.\end{aligned} \tag{8}$$

利用泰勒级数公式,将非线性系统 $f_0(\boldsymbol{\varphi}(t))$在 $f_0(\boldsymbol{\varphi}^*)$处展开得

$$f_0(\boldsymbol{\varphi}(t))=f_0(\boldsymbol{\varphi}^*)+\frac{\mathrm{d}f_0(\boldsymbol{\varphi}(t))}{\mathrm{d}\boldsymbol{\varphi}(t)}(\boldsymbol{\varphi}(t)-\boldsymbol{\varphi}^*(t))+\frac{1}{2}\frac{\mathrm{d}f_0(\boldsymbol{\varphi}(t))}{\mathrm{d}\boldsymbol{\varphi}(t)}(\boldsymbol{\varphi}(t)-\boldsymbol{\varphi}^*(t))^2. \tag{9}$$

假设非线性函数 f_0 满足 Lipschitz 条件,即

$$\| f_0(\boldsymbol{\varphi}(t))-f_0(\boldsymbol{\varphi}^*(t))-\nabla f_0(\boldsymbol{\varphi}^*(t))\widetilde{\boldsymbol{\varphi}}(t))\| \leqslant \frac{L}{2}\widetilde{\boldsymbol{\varphi}}^2(t). \tag{10}$$

式(10)中:L 为一常数. 联合以上三式可得到式(10)均方误差期望的一个上界值为

$$W(\boldsymbol{\varphi}^*,f_0,w^N)\leqslant(|a_0|+\sum_{t=1}^{N}|a_t||u(t)|+\frac{L}{2}\sum_{t=1}^{N}|b_t|\|\widetilde{\boldsymbol{\varphi}}(t)\|^2)^2+\sigma_e^2\sum_{t=1}^{N}b_t^2. \tag{11}$$

直接加权优化辨识算法中的最小化均方误差期望值 $W(\boldsymbol{\varphi}^*,f_0,w^N)$,可转化为最小化式(11)右边的上界值,即得到的最优化问题为

$$\left.\begin{aligned}&\min_{\theta^{2N+1}}\ (|a_0|+\sum_{t=1}^{N}|a_t||u(t)|+\frac{L}{2}\sum_{t=1}^{N}|b_t|\|\widetilde{\boldsymbol{\varphi}}(t)\|^2)^2+\sigma_e^2\sum_{t=1}^{N}b_t^2,\\&\text{s.t.}\quad \sum_{t=1}^{N}b_t=1,\qquad \sum_{t=1}^{N}b_t\widetilde{\boldsymbol{\varphi}}(t)=0.\end{aligned}\right\} \tag{12}$$

以上叙述内容即为文献[3]中所提出的直接加权优化辨识算法的核心思想，所不同的是文中的复杂度有所增加．在式(12)所示的最优化问题中多出了一项 $\sum_{t=1}^{N} |a_t||u(t)|$．

3　最优 Karush-Kuhn-Tucker 充要条件

为了从式(12)所示的带有约束条件的最优化问题中求解出未知权重值组成的未知参数矢量 $\boldsymbol{\theta}$，引入松弛变量 s_t, w_t，使得 $|b_t| \leqslant s_t, t=1,2,\cdots,N$；$|a_t| \leqslant w_t, t=0,1,\cdots,N$. 将引入的两松弛变量 s_t, w_t 应用于最优化问题式(12)中，可得最优化问题为

$$\left.\begin{aligned}
&\min_{\theta^{2N+1},\{s_t\}_{t=1}^{N},\{w_t\}_{t=0}^{N}} \left(w_0 + \sum_{t=1}^{N} w_t \,||\, u(t) \,| + \frac{L}{2}\sum_{t=1}^{N} s_t \,\|\, \widetilde{\boldsymbol{\varphi}}(t) \,\|^2\right)^2 + \sigma_e^2 \sum_{t=1}^{N} s_t^2, \\
&\text{s. t.}\quad s_t \geqslant b_t, \qquad s_t \geqslant -b_t, \qquad t = 1,\cdots,N, \\
&\qquad\; w_t \geqslant a_t, \qquad w_t \geqslant -a_t, \qquad t = 0,\cdots,N, \\
&\qquad\; \sum_{t=1}^{N} b_t = 1, \qquad \sum_{t=1}^{N} b_t \,\|\, \widetilde{\boldsymbol{\varphi}}(t) = 0.
\end{aligned}\right\} \tag{13}$$

要想得到最优化问题式(13)的最优解$(a_0,a_1,\cdots,a_N;b_1,\cdots,b_N;s_t|_1^N,w_t|_0^N)$，根据最优 Karush-Kuhn-Tucker 充要条件可知，对应的拉格朗日函数为

$$\begin{aligned}
&L(\theta^{2N+1}, s_t|_1^N, w_t|_0^N, \lambda_1, \lambda_2, \boldsymbol{\mu}_t^{\pm}|_0^N, \boldsymbol{\gamma}_t^{\pm}|_1^N) = \left(w_t + \sum_{t=1}^{N} w_t \,|\, u(t) \,| + \right. \\
&\left.\frac{L}{2}\sum_{t=1}^{N} s_t \,\|\, \widetilde{\boldsymbol{\varphi}}(t) \,\|^2\right)^2 + \sigma_e^2 \sum_{t=1}^{N} s_t^2 - \lambda_1\left(\sum_{t=1}^{N} b_t - 1\right) - \lambda_2\left(\sum_{t=1}^{N} b_t \widetilde{\boldsymbol{\varphi}}(t)\right) - \\
&\sum_{t=0}^{N} \boldsymbol{\mu}_t^{+}(w_t - a_t) - \sum_{t=0}^{N} \boldsymbol{\mu}_t^{-}(w_t + a_t) - \sum_{t=1}^{N} \boldsymbol{\gamma}_t^{+}(s_t - b_t) - \sum_{t=1}^{N} \boldsymbol{\gamma}_t^{-}(s_t + b_t).
\end{aligned} \tag{14}$$

式(14)中：λ_1 和 λ_2 为对应于等式约束的拉格朗日乘子；$\boldsymbol{\mu}_t^{\pm}|_0^N$ 和 $\boldsymbol{\gamma}_t^{\pm}|_1^N$ 为对应于 $2N+1$ 个不等式约束的拉格朗日乘子矢量，$\boldsymbol{\mu}_t^{\pm} = (\boldsymbol{\mu}_0^{\pm}, \boldsymbol{\mu}_1^{\pm}, \cdots, \boldsymbol{\mu}_N^{\pm})^{\mathrm{T}}$，$\boldsymbol{\gamma}_t^{\pm} = (\gamma_1^{\pm}, \cdots, \gamma_N^{\pm})^{\mathrm{T}}$.

利用 Karush-Kuhn-Tucker 最优性充要条件，可得在最优处存在成立的等式，即

$$\left.\begin{aligned}
&\frac{\partial L}{\partial a_t} = \boldsymbol{\mu}_t^{+} - \boldsymbol{\mu}_t^{-} = 0, \qquad t = 0,\cdots,N, \\
&\frac{\partial L}{\partial b_t} = -\lambda_1 - \lambda_2 \widetilde{\boldsymbol{\varphi}}(t) + \boldsymbol{\gamma}_t^{+} - \boldsymbol{\gamma}_t^{-} = 0, \qquad t = 1,\cdots,N, \\
&\frac{\partial L}{\partial s_t} = 2\left(w_0 + \sum_{t=1}^{N} w_t \,|\, u(t) \,| + \frac{L}{2}\sum_{t=1}^{N} s_t \,\|\, \widetilde{\boldsymbol{\varphi}}(t) \,\|^2\right) \|\, \widetilde{\boldsymbol{\varphi}}(t) \,\|^2 + 2\sigma_e^2 s_t - \boldsymbol{\gamma}_t^{+} - \boldsymbol{\gamma}_t^{-} = 0, \\
&\frac{\partial L}{\partial w_0} = 2\left(w_0 + \sum_{t=1}^{N} w_t \,|\, u(t) \,| + \frac{L}{2}\sum_{t=1}^{N} s_t \,\|\, \widetilde{\boldsymbol{\varphi}}(t) \,\|^2\right) - \boldsymbol{\mu}_t^{+} - \boldsymbol{\mu}_t^{-} = 0, \\
&\frac{\partial L}{\partial w_1} = 2\left(w_0 + \sum_{t=1}^{N} w_t \,|\, u(t) \,| + \frac{L}{2}\sum_{t=1}^{N} s_t \,\|\, \widetilde{\boldsymbol{\varphi}}(t) \,\|^2\right) |\, u(t) \,| - \boldsymbol{\mu}_t^{+} - \boldsymbol{\mu}_t^{-} = 0, \\
&\sum_{t=1}^{N} b_t = 1, \qquad \sum_{t=1}^{N} b_t \widetilde{\boldsymbol{\varphi}}(t) = 0, \\
&\boldsymbol{\mu}_t^{+}(w_t - a_t) = 0, \qquad \boldsymbol{\mu}_t^{-}(w_t + a_t) = 0, \\
&\boldsymbol{\gamma}_t^{+}(s_t - b_t) = 0, \qquad \boldsymbol{\gamma}_t^{-}(s_t + b_t) = 0, \\
&\boldsymbol{\mu}_0^{\pm} \geqslant 0, \qquad t = 0,\cdots,N; \\
&\boldsymbol{\gamma}_t^{\pm} \geqslant 0, \qquad t = 1,\cdots,N.
\end{aligned}\right\} \tag{15}$$

在最优解除包含有隐含的最优等式，$|b_t| = s_t, t=1,2,\cdots,N$；$|a_t| = w_t, t=0,1,\cdots,N$，故由第 1 个子式可知 $\boldsymbol{\mu}_t^{+} = \boldsymbol{\mu}_t^{-}$. 在第 9 个子式中，若 $a_t > 0$，则 $w_t + a_t = |a_t| + a_t = 2a_t > 0$，要使第 9 个子式成立必须有 $\boldsymbol{\mu}_t^{-}$，从而由第 1 个子式有，$\boldsymbol{\mu}_t^{+} = \boldsymbol{\mu}_t^{-} = 0$. 在第 2 个子式中，若 $a_t < 0$，则 $w_t - a_t = |a_t| - a_t = -2a_t > 0$. 同理要使得第 8 个子式成立，需要使得 $\boldsymbol{\mu}_t^{+} = 0$. 从而由第 1 个子式有 $\boldsymbol{\mu}_t^{+} = \boldsymbol{\mu}_t^{-} = 0$. 当 $a_t = 0$ 时即所有的

输入观测数据序列前的未知权重值都为0,在线性仿射函数形式中仅剩下输出观测数据序列项,从而简化为文献[3]中的特殊形式.

综合 $a_t>0$ 和 $a_t<0$ 两种情况,可得 $\boldsymbol{\mu}_t^+=\boldsymbol{\mu}_t^-=0,\dfrac{\partial L}{\partial a_t}=0,t=0,\cdots,N$. 代入到式(15),可得

$$\left.\begin{aligned}
&\lambda_1+\lambda_2\widetilde{\boldsymbol{\varphi}}(t)=\boldsymbol{\gamma}_t^+-\boldsymbol{\gamma}_t^-,\qquad t=1,\cdots,N,\\
&2(\mid a_0\mid+\sum_{t=1}^{N}\mid a_t\mid\mid u(t)\mid+\frac{L}{2}\sum_{t=1}^{N}\mid b_t\mid\ \|\widetilde{\boldsymbol{\varphi}}(t)\|^2)\|\widetilde{\boldsymbol{\varphi}}(t)\|^2+2\sigma_e^2\mid b_t\mid=\boldsymbol{\gamma}_t^++\boldsymbol{\gamma}_t^-,\\
&2(\mid a_0\mid+\sum_{t=1}^{N}\mid a_t\mid\mid u(t)\mid+\frac{L}{2}\sum_{t=1}^{N}\mid b_t\mid\ \|\widetilde{\boldsymbol{\varphi}}(t)\|^2)=0,\\
&\sum_{t=1}^{N}b_t=1,\qquad \sum_{t=1}^{N}b_t\widetilde{\boldsymbol{\varphi}}(t)=0,\\
&\boldsymbol{\gamma}_t^+(\mid b_t\mid-b_t)=0,\qquad \boldsymbol{\gamma}_t^-(\mid b_t\mid+b_t)=0.
\end{aligned}\right\}\tag{16}$$

式(16)中:第4个和第5个子式所代表的等式关系已全部隐含在所构造的拉格朗日函数之中. 将第3个子式代入第2个子式中可得

$$2\sigma_e^2\mid b_t\mid=\boldsymbol{\gamma}_t^++\boldsymbol{\gamma}_t^-.\tag{17}$$

当 $b_t>0$ 时,由式(16)的第7个子式可得 $|b_t|+b_t=2b_t>0$. 从而要使第7个子式成立,需要有 $\boldsymbol{\gamma}_t^-=0$. 将 $\boldsymbol{\gamma}_t^-=0$ 代入到式(16)的第1个子式,可得

$$\boldsymbol{\gamma}_t^+=\lambda_1+\lambda_2\widetilde{\boldsymbol{\varphi}}(t),$$

再代入到式(17)即有

$$b_t=\frac{\boldsymbol{\gamma}_t^+}{2\sigma_e^2}=\frac{\lambda_1+\lambda_2\widetilde{\boldsymbol{\varphi}}(t)}{2\sigma_e^2}.\tag{18}$$

同理考虑当 $b_t<0$ 时,有

$$\boldsymbol{\gamma}_t^+=0,\qquad \lambda_1+\lambda_2\widetilde{\boldsymbol{\varphi}}(t)=-\boldsymbol{\gamma}_t^-,\qquad -2\sigma_e^2b_t=\boldsymbol{\gamma}_t^-=-(\lambda_1+\lambda_2\widetilde{\boldsymbol{\varphi}}(t)).$$

整理后同样可得,$b_t=\dfrac{\lambda_1+\lambda_2\widetilde{\boldsymbol{\varphi}}(t)}{2\sigma_e^2}$.

将式(18)所表示的权重值 $\{b_t\}_{t=1}^N$ 的表达式与文献[3]中的定理1进行比较可知,文中的 b_t 表达式更易于求解. 它仅为一个固定的形式,而文献[3]定理1中的表达式却是一个分段函数形式.

以上给出的是权重值 $\{b_t\}_{t=1}^N$ 的求解表达式,但在文中的线性仿射函数形式(2)中,还存在权重值 $\{a_t\}_{t=0}^N$,仍需进一步推导其求解过程.

将式(18)代入式(16)的第3个子式,可得

$$\mid a_0\mid+\sum_{t=1}^{N}\mid a_t\mid\mid u(t)\mid+\frac{L}{2}\sum_{t=1}^{N}\frac{\lambda_1+\lambda_2\widetilde{\boldsymbol{\varphi}}(t)}{2\sigma_e^2}\|\widetilde{\boldsymbol{\varphi}}(t)\|^2=0.$$

若在实数域中,必然有等式

$$\mid a_t\mid=0,\qquad \sum_{t=1}^{N}\mid a_t\mid\mid u(t)\mid=0,\qquad \sum_{t=1}^{N}(\lambda_1+\lambda_2\widetilde{\boldsymbol{\varphi}}(t))\|\widetilde{\boldsymbol{\varphi}}(t)\|^2=0$$

成立. 由此可解得

$$a_0=a_1=\cdots=a_N=0,\qquad \lambda_1\sum_{t=1}^{N}\|\widetilde{\boldsymbol{\varphi}}(t)\|^2+\lambda_2\sum_{t=1}^{N}\widetilde{\boldsymbol{\varphi}}(t)\|\widetilde{\boldsymbol{\varphi}}(t)\|^2=0.$$

但这并不是此处所期望得到的结果. 为此扩展到复数域中,考虑到

$$\mid a_t\mid+\sum_{t=1}^{N}\mid a_t\mid\mid u(t)\mid=\frac{L}{2}\sum_{t=1}^{N}\frac{\lambda_1+\lambda_2\widetilde{\boldsymbol{\varphi}}(t)}{2\sigma_e^2}\|\widetilde{\boldsymbol{\varphi}}(t)\|^2\tag{19}$$

的 $|u(t)|$ 表示的是输入激励信号的振幅值,当选择恒定振幅值的输入激励信号时,有 $|u(t)|=k$(k 为一正常数). 因此,式(19)可改写为

$$\mid a_t\mid+k\sum_{t=1}^{N}\mid a_t\mid=-\frac{L}{2}\sum_{t=1}^{N}\frac{\lambda_1+\lambda_2\widetilde{\boldsymbol{\varphi}}(t)}{2\sigma_e^2}\|\widetilde{\boldsymbol{\varphi}}(t)\|^2.\tag{20}$$

满足以上等式的未知权重值的选取方法有无穷多种,如

$$a_0=0,\qquad a_t=\pm\frac{1}{N}\frac{L}{2k}\sum_{t=1}^{N}\frac{\lambda_1+\lambda_2\widetilde{\boldsymbol{\varphi}}(t)}{2\sigma_e^2}\parallel\widetilde{\boldsymbol{\varphi}}(t)\parallel^2.$$

在线性代数中，经常使用的选择方法是在增加一个$\sum\limits_{t=1}^{N}a_t=1$的约束条件下进行求解. 为了能够去掉式(20)中的绝对值符号，设 $N+1$ 个权重值中前 k_1+1 项为正，后 $N-k_1$ 为负，则式(19)，(20)可合并为

$$\begin{bmatrix}1 & 1 & \cdots & 1 & 1 & \cdots & 1\\ 1 & 1 & \cdots & 1 & 0 & \cdots & 0\\ 1 & 0 & \cdots & 0 & 1 & \cdots & 1\end{bmatrix}\begin{bmatrix}a_0\\ \cdots\\ a_{k_1}\\ a_{k_1+1}\\ \cdots\\ a_N\end{bmatrix}=\begin{bmatrix}1\\ \dfrac{1}{2}-\dfrac{L}{4}\sum\limits_{t=1}^{N}\dfrac{\lambda_1+\lambda_2\widetilde{\boldsymbol{\varphi}}(t)}{2\sigma_e^2}\parallel\widetilde{\boldsymbol{\varphi}}(t)\parallel^2\\ \dfrac{1}{2}+\dfrac{L}{4}\sum\limits_{t=1}^{N}\dfrac{\lambda_1+\lambda_2\widetilde{\boldsymbol{\varphi}}(t)}{2\sigma_e^2}\parallel\widetilde{\boldsymbol{\varphi}}(t)\parallel^2\end{bmatrix}. \tag{21}$$

在式(21)所示的奇异退化线性方程组中，通过选择 $N-2$ 个自由变量，总能得到一组未知权重值$\{a_t\}_{t=0}^{N}$的选取序列值，并在 $\sum\limits_{t=0}^{N}a_t=1$ 的约束条件下，总能保证线性仿射函数式(2)中至少有一个 a_t 的取值不为 0，以 a_t 的取值来对 b_t 起辅助逼近的作用.

4　未知权重值的迭代求解

假设 $a_0=w_0=0$，式(13)的优化决策变量包括三类变量 θ^{2N}，$\{s_t\}_{t=1}^{N}$，$\{w_t\}_{t=1}^{N}$. 设由此三类变量构成的 $4N$ 维的列矢量记为 $\boldsymbol{\eta}=(\theta^{2N};s_1,\cdots,s_N;w_1,\cdots,w_N)^{\mathrm{T}}$. 将最优化问题式(13)中的 $4N$ 个不等式约束条件整理成矩阵乘积的形式，有

$$\begin{bmatrix}-1 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & \cdots & -1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & \cdots & 1\\ 1 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & \cdots & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 1\\ 0 & \cdots & 0 & -1 & \cdots & 0 & 1 & \cdots & 0 & 0 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & \cdots & 0 & 0 & \cdots & -1 & 0 & \cdots & 1 & 0 & \cdots & 0\\ 0 & \cdots & 0 & 1 & \cdots & 0 & 1 & \cdots & 0 & 0 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & \cdots & 0 & 0 & \cdots & 1 & 0 & \cdots & 1 & 0 & \cdots & 0\end{bmatrix}\begin{bmatrix}a_0\\ \vdots\\ a_N\\ b_1\\ \vdots\\ b_N\\ s_1\\ \vdots\\ s_N\\ w_1\\ \vdots\\ w_N\end{bmatrix}\geqslant\boldsymbol{0}. \tag{22}$$

设上式左边 $4N\times 4N$ 维的矩阵记为 $\boldsymbol{A}$，右边 $\boldsymbol{0}$ 表示 $4N\times 1$ 维的零矢量，则不等式约束条件可简记为 $\boldsymbol{A\eta}\geqslant\boldsymbol{0}$. 同理，可将两等式约束条件简记为矩阵乘积的形式

$$\begin{bmatrix}1 & \cdots & 1 & -1 & \cdots & -1 & 0 & \cdots & 0 & 0 & \cdots & 0\\ 0 & \cdots & 0 & \widetilde{\boldsymbol{\varphi}}(1) & \cdots & \widetilde{\boldsymbol{\varphi}}(N) & 0 & \cdots & 0 & 0 & \cdots & 0\end{bmatrix}\times$$
$$[a_1,\cdots,a_N;b_1,\cdots,b_N;s_1,\cdots,s_N;w_1,\cdots,w_N]^{\mathrm{T}}=\boldsymbol{0}.$$

设上式左边 $2\times 4N$ 维矩阵记为 $\boldsymbol{B}$，右边 $\boldsymbol{0}$ 表示 2×1 的零矢量，则等式约束条件可记为 $\boldsymbol{B\eta}=\boldsymbol{0}$.

下面对目标函数进行整理. 很明显目标函数的第 2 项可改写为

$$\sigma_e^2\sum_{t=1}^{N}s_t^2=\sigma_e^2\boldsymbol{\eta}^{\mathrm{T}}\begin{bmatrix}0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & E & 0\\ 0 & 0 & 0 & 0\end{bmatrix}\boldsymbol{\eta}=\sigma_e^2\boldsymbol{\eta}^{\mathrm{T}}\boldsymbol{C}_2\boldsymbol{\eta}, \tag{23}$$

而目标函数中第 1 项括号内的式子可改写成

$$w_0+\sum_{t=1}^{N}w_t\mid u(t)\mid+\frac{L}{2}\sum_{t=1}^{N}s_t\parallel\widetilde{\boldsymbol{\varphi}}(t)\parallel^2=$$

$$[0,\cdots,0;\frac{L}{2}\parallel\widetilde{\boldsymbol{\varphi}}(1)\parallel^2,\cdots,\frac{L}{2}\parallel\widetilde{\boldsymbol{\varphi}}(N)\parallel^2;\mid u(t)\mid,\cdots,\mid u(N)\mid]=\boldsymbol{C}_1^{\mathrm{T}}\boldsymbol{\eta}. \tag{24}$$

将式(24)平方后可得

$$(w_0+\sum_{t=1}^{N}w_t\mid u(t)\mid+\frac{L}{2}\sum_{t=1}^{N}s_t\parallel\widetilde{\boldsymbol{\varphi}}(t)\parallel^2)^2=(\boldsymbol{C}_1^{\mathrm{T}}\boldsymbol{\eta})^2=\boldsymbol{\eta}^{\mathrm{T}}\boldsymbol{C}_1\boldsymbol{C}_1^{\mathrm{T}}\boldsymbol{\eta}. \tag{25}$$

由此可构成最优化问题为

$$\left.\begin{aligned}&\min_{\boldsymbol{\eta}^{4N}}\quad \boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{C}_1\boldsymbol{C}_1^{\mathrm{T}}+\sigma_e^2\boldsymbol{C}_2)\boldsymbol{\eta},\\&\text{s. t.}\quad \boldsymbol{A\eta}\geqslant\boldsymbol{0},\qquad \boldsymbol{B\eta}=\boldsymbol{0}.\end{aligned}\right\} \tag{26}$$

很明显,式(26)所表示的最优化问题中的目标函数为关于优化变量 $\boldsymbol{\eta}$ 的二次函数,不等式约束和等式约束均为关于优化变量 $\boldsymbol{\eta}$ 的线性函数. 即式(26)为一个二次规划问题,可采用内点法迭代地求解. 定义式(26)的拉格朗日函数为

$$\boldsymbol{L}(\boldsymbol{\eta},m,n)=\boldsymbol{\eta}^{\mathrm{T}}(\boldsymbol{C}_1\boldsymbol{C}_1^{\mathrm{T}}+\sigma_e^2\boldsymbol{C}_2)\boldsymbol{\eta}-m\boldsymbol{A\eta}-n\boldsymbol{B\eta}. \tag{27}$$

对式(27)关于 $\boldsymbol{\eta}$ 求偏导,可得

$$\left.\begin{aligned}&(\boldsymbol{C}_1\boldsymbol{C}_1^{\mathrm{T}}+\sigma_e^2\boldsymbol{C}_2)\boldsymbol{\eta}-\boldsymbol{A}^{\mathrm{T}}m-\boldsymbol{B}^{\mathrm{T}}n=0,\qquad m,n\geqslant 0,\\&m(\boldsymbol{A\eta})=\boldsymbol{0},\qquad \boldsymbol{A\eta}\geqslant\boldsymbol{0},\qquad \boldsymbol{B\eta}=\boldsymbol{0}.\end{aligned}\right\} \tag{28}$$

为了消去 $\boldsymbol{A\eta}\geqslant\boldsymbol{0}$ 不等式约束中的不等号,引入一个松弛变量 $z\geqslant 0$,对以上各式进行改写为

$$\left.\begin{aligned}&(\boldsymbol{C}_1\boldsymbol{C}_1^{\mathrm{T}}+\sigma_e^2\boldsymbol{C}_2)\boldsymbol{\eta}-\boldsymbol{A}^{\mathrm{T}}m-\boldsymbol{B}^{\mathrm{T}}n=0,\\&\boldsymbol{B\eta}=0,\qquad \boldsymbol{A\eta}-z=0,\qquad zm^{\mathrm{T}}=0.\end{aligned}\right\} \tag{29}$$

设由式(29)构成的矩阵记为

$$\boldsymbol{F}=[(\boldsymbol{C}_1\boldsymbol{C}_1^{\mathrm{T}}+\sigma_e^2\boldsymbol{C}_2)\boldsymbol{\eta}-\boldsymbol{A}^{\mathrm{T}}m-\boldsymbol{B}^{\mathrm{T}}n,\boldsymbol{A\eta}-z,\boldsymbol{B\eta},\boldsymbol{Z\Lambda}e-\varepsilon\boldsymbol{e}.]^{\mathrm{T}} \tag{30}$$

其中:$\boldsymbol{Z}=\mathrm{diag}(z_1,\cdots,z_{4N})$;$\boldsymbol{\Lambda}=\mathrm{diag}(m_1,\cdots,m_{4N})$;$\varepsilon\in[0,1]$;$\boldsymbol{e}=[1,1,\cdots,1]^{\mathrm{T}}$.

约束最小化可通过迭代地更新矢量 $\boldsymbol{\eta}$ 来获得,该最小化包含在于寻找拉格朗日函数的平稳点. 在最小化过程中,新的迭代值 $\boldsymbol{\eta}$ 是通过在此时估计值的基础之上增加一项 $\Delta\boldsymbol{\eta}$. 利用约束高斯-牛顿法时,$\Delta\boldsymbol{\eta}$ 应为等式

$$\begin{bmatrix}\boldsymbol{C}_1\boldsymbol{C}_1^{\mathrm{T}}+\sigma_e^2\boldsymbol{C}_2 & 0 & -\boldsymbol{A} & -\boldsymbol{B}\\ \boldsymbol{A} & -\boldsymbol{I} & 0 & 0\\ 0 & \boldsymbol{B} & 0 & 0\\ 0 & \boldsymbol{\Lambda} & \boldsymbol{Z} & 0\end{bmatrix}\begin{bmatrix}\Delta\boldsymbol{\eta}\\ \Delta z\\ \Delta m\\ \Delta n\end{bmatrix}=-\begin{bmatrix}(\boldsymbol{C}_1\boldsymbol{C}_1^{\mathrm{T}}+\sigma_e^2\boldsymbol{C}_2)\boldsymbol{\eta}-\boldsymbol{A}^{\mathrm{T}}m-\boldsymbol{B}^{\mathrm{T}}n\\ \boldsymbol{A\eta}-z\\ \boldsymbol{B\eta}\\ \boldsymbol{Z\Lambda}\boldsymbol{e}-\varepsilon\boldsymbol{e}\end{bmatrix}=-\boldsymbol{F} \tag{31}$$

的解. 则第 $k+1$ 次迭代更新值可取为

$$(\boldsymbol{\eta}^{k+1},z^{k+1},m^{k+1},n^{k+1})=(\boldsymbol{\eta}^{k},z^{k},m^{k},n^{k})+v(\Delta\boldsymbol{\eta},\Delta z,\Delta m,\Delta n). \tag{32}$$

其中:搜索方向的步长应满足不等式$(z^{k+1},m^{k+1})>0$ 成立,搜索方向由式(31)来求取.

为了避免奇异性现象的发生,可在 $\boldsymbol{C}_1\boldsymbol{C}_1^{\mathrm{T}}+\sigma_e^2\boldsymbol{C}_2$ 的基础上加上一项所谓的 Levenberg-Marquardt 参数 δ^2,使得式左(31)边矩阵的左上角(1,1)矩阵变化为 $\boldsymbol{C}_1\boldsymbol{C}_1^{\mathrm{T}}+\sigma_e^2\boldsymbol{C}_2+\delta^2\boldsymbol{I}$,从而保证该左上角矩阵存在可逆矩阵,并且其可逆矩阵为正定有界.

5 结论

对于非线性系统的直接加权优化辨识算法,通过在原线性仿射函数形式中,增加若干项关于输入观测数据序列的线性项来增强逼近非线性,减少逼近的时间. 对于增加若干线性项后展开式中的未知权重值的选取,分别从理论和实用上推导出这些未知权重值的选取过程.

理论上的推导分析,可明确增加的未知权重值在整个逼近非线性系统的目的中起着辅助作用;实用上的推导分析,则展示了如何将某些复杂的最优化问题经过整理变换成常见的最优化问题,从而可利用最为基础的优化方法来求解.

由于文中利用输入观测数据序列，而此输入观测数据需要人为事先选定以满足持续激励非线性系统，故对该非线性系统的选择持续激励的输入信号是目前需要深入研究的内容.

参考文献：

[1] LJUNG L. System identification: Theory for the user[M]. New Jersey:Prentice Hall,1999.

[2] BOYD S,VANDENBERGHE L. Convex optimization[M]. Cambridge:Cambridge University Press,2008.

[3] ROLL J,NAZINY A,LIUNG L. Non-linear system identification via direct weight optimization[J]. Automatica,2005,41(3):475-490.

[4] ROLL J,BEMPORAD A,LIUNG L. Identification of piecewise affine systems via direct mixed integer programming [J]. Automatica,2004,40(1):37-50.

[5] ROLL J. Piecewise linear solution paths with application to direct weight optimization[J]. Automatica,2008,44(11):2745-2753.

[6] PAOLETTI S,ROLL J,GARULLI A,et al. On the input-output representation of piecewise affine state space models[J]. IEEE Transactions of Automatic Control,2010,55(1):60-73.

[7] BAI E W. Recursive direct weight optimization in nonlinear system identification: A minimal probability approach [J]. IEEE Transactions of Automatic Control,2007,52(7):1218-1231.

[8] PINTELON R,SCHOUKENS J. System identification: A frequency domain approach[M]. New York:IEEE Press,2001.

[9] NOCEDAL J,WRIGHT S. Numerical optimization[M]. Berlin:Springer-Verlag,2002.

[10] ZEILINGE M N,JONES CN,MORARI M. Real-time suboptimal model predictive control using a combination of explicit MPC and online optimization[J]. IEEE Transactions of Automatic Control,2011,56(7):1524-1534.

Research of a Kind of Direct Weight Optimization Identification Algorithm Based on Nonlinear System

YAO Jie, ZHU Yong-hong

(School of Mechanical and Electronic Engineering, Jingdezhen Ceramic Institute, Jingdezhen 333403, China)

Abstract: As for the direct weighted optimization identification algorithm in the nonlinear system, enhancing approximate nonlinear and reducing approximate time can be realized by increasing several items about the linear items of input observed sequence in the original linear affine-functions. As for the unknown weights selection in expansion equation after increasing several linear items, this paper deduced the selection process of these unknown weight value from the theory and practical angle respectively . Theoretical analysis can increase definitely auxiliary role of unknown weights in the purpose of approximate nonlinear system;On the practical analysis is showing how to transform the complicated optimization problems to the common optimization problems and finally the based optimization solution can be used to solve these problems.

Keywords: system identification; nonlinear system; weights value; optimization problems; iterative choose

（**责任编辑**：陈志贤　　**英文审校**：吴逢铁）