



论 文

基于多项式预测模型的最优自适应滤波算法

檀甲甲, 张建秋*

复旦大学电子工程系, 上海 200433

* 通信作者. E-mail: jqzhang@fudan.edu.cn

收稿日期: 2009-11-07; 接受日期: 2010-04-27

国家自然科学基金 (批准号: 60872059) 资助项目

摘要 文中提出了一个通过多项式预测模型来描述待辨识系统冲击响应系数的最优自适应滤波算法. 该算法首先利用具有时不变参数的多项式预测模型, 来描述系统冲击响应的时变/时不变系数. 当视描述的模型为其待辨识系统冲击响应时变/时不变系数进化的状态方程, 而待辨识系统的输入和输出联系视为对这些状态的观测方程时, 自适应滤波问题可以在 Kalman 滤波的框架下得以解决. 由于在 Gauss 白噪声环境中以及状态方程准确的情况下, Kalman 滤波是最大似然、最大后验和最小均方等统计意义下的最优滤波, 因此当待辨识的系统冲击响应系数可以由多项式模型建模时, 文中模型和相应算法也是这些统计意义下的最优自适应滤波. 在分析的结果验证上述结论的同时, 仿真的结果也表明: 文中提出的自适应滤波算法的性能优于已知的自适应滤波算法.

关键词 自适应滤波 多项式预测模型 冲击响应系数 状态方程 Kalman 滤波 最优

1 引言

自适应信号处理, 在噪声和回声的抵消、谱线的增强、信道的均衡、系统的辨识和延迟时间的估计等应用中起着非常重要的作用^[1-4]. 在过去几十年中, 针对这些应用, 人们已经提出了许多不同的自适应滤波算法. 这些方法包括最小均方 (LMS, least mean squares) 以及迭代最小二乘法 (RLS, recursive least squares)^[1-6] 等自适应滤波法. 在这些算法中, LMS 算法是一种经典的自适应滤波算法, 其最优性和稳定性在文献 [7, 8] 中得到了充分的讨论. 在此基础上发展的归一化最小均方 (NLMS, normalized least-mean-square) 的自适应滤波算法, 由于其具有较好的鲁棒性和易于使用, 而成为最常用的自适应滤波算法^[1]. 这个算法的关键在于其步长参数的选取, 这是因为它的选取能折中算法的收敛性和跟踪性能. 因此如何选取好而可靠的步长就成为该自适应滤波算法的研究热点^[9-13]. 近年来, 针对不同的步长选取策略, 人们提出了许多具有不同形式的 NLMS 自适应滤波方法. 这些方法多以变步长为其特点^[14-16]. 通过分析不难发现: 虽然这些算法有不依赖于待辨识的系统冲击响应系数的先验信息的优点, 但是这些算法的步长的获取都依赖于先验误差信号的模型和对系统观测噪声方差的估计. 然而, 由于没有考虑待辨识的系统冲击响应系数的先验信息, 所以通常对先验误差信号模型的估计是有一定误差的. 同时, 系统噪声的方差的估计通常也很难达到十分准确. 由于这两点, NLMS 类的算法的性能都会受到了一定程度的影响. 此外, 同样可以发现, 传统或经典的自适应滤波算法都是通过待辨识系统输

入输出之间的联系, 试图在诸如最小均方或最小二乘等统计意义下, 获取自适应滤波的算法^[2]. 虽然, 在最大似然和最大后验等统计意义下的自适应滤波算法在文献中也有相关的研究报道^[17,18], 但是由于其算法过于复杂而未成为应用中的主流.

本文通过利用多项式预测模型来对待辨识系统冲击响应的时变/时不变系数进行建模. 并将建立的模型视为其系数进化的状态方程, 而将待辨识系统的输入输出之间的联系视为其状态的观测方程. 那么, 自适应滤波问题就可以在 Kalman 滤波的框架下得以解决. 由于在 Gauss 白噪声环境中以及状态方程准确的情况下, Kalman 滤波是最大似然、最大后验和最小均方等统计意义下的最优滤波^[19-21], 因此本文的模型和相应的算法在 Gauss 白噪声环境中模型准确的情况下, 可以称之为上述全部统计意义下的最优自适应滤波算法. 在分析的结果验证上述结论的同时, 在仿真的结果中, 本文提出的自适应滤波算法的表现也优于对比的自适应滤波算法. 本文中采用的对比算法为 NPVSS-NLMS 算法. 该对比算法在文献中称其性能优于其他的传统算法^[16].

本文的第 2 节给出了滤波器冲击响应系数的多项式预测模型; 第 3 节以一阶多项式预测模型为例, 对该预测模型进行了状态空间描述并给出了相应的迭代最优自适应滤波算法以及算法的性能分析; 第 4 节给出了相关的仿真和分析结果. 最后给出了结论.

2 系统冲击响应系数的多项式预测模型

Weierstrass 逼近定理^[22]告诉我们: 如果 $g(x)$ 是定义在区间 $[a, b]$ 上的连续实函数, 对任意小的正数 $\varepsilon > 0$, 存在区间 $[a, b]$ 上的多项式 $p(x)$, 对任意的 $x \in [a, b]$, 满足 $|g(x) - p(x)| < \varepsilon$. 也就是说, 任一有界闭区间的连续函数都可以由该区间内的多项式以任意精度一致逼近. 这意味着对任意连续信号, 可以将该函数的定义域分割成若干小区间, 如果每个小区间可以充分小, 那么就可以由一个低阶多项式近似到任意精度^[23].

假设待辨识的系统冲击响应系数是时变的, 据 Weierstrass 逼近定理, 我们知道如果该系统冲击响应中任意一个系数, 其关于时间的函数的定义域可以分割成若干连续的小区间, 并在其每一小区间内是可以用多项式以任意精度一致逼近. 这样, 假设在时刻 n , 一个长度为 M 的待辨识系统冲击响应系数的向量是: $\mathbf{h}(n) = [h(n), h_1(n), \dots, h_{M-1}(n)]^T$. 该系统冲击响应的任意一个系数 $h_i(n) (i = 0, 1, \dots, M-1)$ 在某一时间区间内可以由如下 L 阶多项式描述:

$$h_i(n) = \sum_{l=0}^L p(l)n^l, \quad (1)$$

其中, $p(l) (l = 0, \dots, L)$ 为多项式第 l 次幂的系数, L 是该多项式的最高阶.

如果我们希望用 $h_i(n)$ 的前 K 个时刻的值 $h_i(n-K+1), \dots, h_i(n)$ 来预测 $h_i(n)$ 的第 N 个将来值 $h_i(n+N)$, 即

$$h_i(n+N) = \sum_{k=0}^{K-1} f(k)h_i(n-k), \quad (2)$$

$f(k) (k = 0, \dots, K-1)$ 是用来预测的系数. 显然, (2) 式是以 $f(k) (k = 0, \dots, K-1)$ 为系数的一个 FIR 滤波器. 将 (1) 式代入 (2) 式, 有

$$\sum_{l=0}^L p(l)[n+N]^l = \sum_{k=0}^{K-1} f(k) \sum_{l=0}^L p(l)[n-k]^l. \quad (3)$$

由于 (1) 式中不同次幂的系数 $p(l)$ 是独立的, 所以可以将 (3) 式中各个 $p(l)$ 联立为 $L+1$ 个等式, 有

$$\begin{aligned} p(0)[n+N]^0 &= \sum_{k=0}^{K-1} f(k)p(0)[n-k]^0, \\ p(1)[n+N]^1 &= \sum_{k=0}^{K-1} f(k)p(1)[n-k]^1, \\ &\dots \\ p(L)[n+N]^L &= \sum_{k=0}^{K-1} f(k)p(L)[n-k]^L. \end{aligned} \quad (4)$$

消去等式两边的 $p(l)(l=0, \dots, L)$, 可得等式组:

$$\begin{aligned} \sum_{k=0}^{K-1} k^0 f(k) &= (-N)^0, \\ \sum_{k=0}^{K-1} k^1 f(k) &= (-N)^1, \\ &\dots \\ \sum_{k=0}^{K-1} k^L f(k) &= (-N)^L. \end{aligned} \quad (5)$$

(5) 式称为多项式预测滤波器系数 $f(k)$ 的约束条件. 但由于通常 $L < K$, 该方程是欠定的, 所以仅通过 (5) 式无法唯一确定 $f(k)$, 还需要其他的约束条件.

正如前面所分析的, (2) 式是以 $f(k) (k=0, \dots, K-1)$ 为系数的一个 FIR 滤波器. 那么 $h_i(n)$ 可以理解为通过该滤波器的信号. 在算法收敛之前, 待辨识的系统冲击响应系数 $h_i(n)$ 是不准确的, 那么我们可以把 $h_i(n)$ 理解为混有噪声的信号. 当然, 这里的噪声假定为 Gauss 白噪声. 那么我们就希望其噪声通过该滤波器的增益最小. 为了使噪声通过该预测滤波器的增益最小, 我们有 [24]

$$NG = \frac{1}{\pi} \int_0^\pi |F(e^{j\omega})|^2 d\omega. \quad (6)$$

(6) 式中, $F(\cdot)$ 为滤波器的传递函数. 对于 (2) 式描述的 FIR 滤波器, (6) 式可改写为 [24]

$$NG = \sum_{k=0}^{K-1} |f(k)|^2. \quad (7)$$

利用 Lagrange 数乘法, 当 $N=1, L=1$ 和 $N=1, L=2$ 时, 在 (5) 式的约束下使得 (7) 式最小的最优解是 [25]

$$f(k) = \frac{4K - 6k - 4}{K(K-1)} (N=1, L=1), \quad (8)$$

$$f(k) = \frac{9K^2 + (-27 - 36k)K + 30k^2 + 42k + 18}{K^3 - 3K^2 + 2K} (N=1, L=2), \quad (9)$$

其中, $k=0, 1, \dots, K-1$. $N=1$ 表明该多项式预测模型对系统冲击响应的每一个系数 $h_i(n)$ 的预测为一步超前预测. $L=1$ 和 $L=2$ 分别表示采用的多项式的阶数分别为一阶和二阶. 多项式的阶数可以根据

实际需要进行选取. 当 N 和 L 为其他值时, 请详见文献 [25] 的计算方法及计算结果. 这里, K 表示用来进行一步超前预测的系数 $h_i(n+1)$ 的个数. 通过 (8) 和 (9) 式可以看出, K 的取值要求为 $K \geq L+1$. 在这个要求下, K 的选取是可以根据具体需要自由选取的. 在本文中, 为了算法的实时性和降低算法的复杂度, K 取 $L+1$. 例如, 当我们选取一阶多项式预测模型进行一步超前预测时, $K=2$.

(1)–(9) 式的推导告诉我们, 当 $f(k)$ ($k=0, \dots, K-1$) 由 (8) 或 (9) 式给出时, 一个遵循 (1) 式这样的多项式规律变化的时变系统冲击响应系数, 其变化规律可由一个时不变 FIR 预测滤波器 (2) 式在数学上严格等效地来建模. 同时, (8) 和 (9) 式也表明, (2) 式的预测滤波器系数 $f(k)$ ($k=0, \dots, K-1$) 是时不变的且只与 N, K, L 有关. 当 N, K, L 确定时, 预测滤波器系数 $f(k)$ ($k=0, \dots, K-1$) 的长度和数值大小也唯一确定.

3 多项式预测模型的状态空间描述及其最优的自适应滤波算法

第 2 节的分析表明: 对一个任意连续时变的系统冲击响应系数, 其中任意一个系数都可以用多项式预测滤波器对其建模. 如果系统系数是不变或缓慢变化的, 或者其在连续的区间中是不变或缓慢变化的, 那么一阶多项式可以足够精确描述或充分逼近每一小区间内不变或缓慢变化的每一个冲击响应系数. 为了描述简单, 下面我们仅仅介绍利用一阶多项式预测模型来建模待辨识系统冲击响应系数的状态空间模型及其相应的自适应滤波算法. 更高阶多项式的预测模型的描述方法及其相应的最优自适应滤波算法可以依理类推.

3.1 一阶多项式预测模型的状态空间描述

由于这里采用一阶多项式一步超前预测模型, 即 $N=1, L=1$, 那么采用 (8) 式来计算该多项式预测模型的系数 $f(k)$. 根据第 2 节的分析, K 取 2. 那么此时得到的一阶多项式预测模型的系数是 $f(0)=2, f(1)=-1$. 采用第 2 节同样的假设, 在时刻 n , 一个长度为 M 的待辨识系统冲击响应系数的向量是: $\mathbf{h}(n) = [h_0(n), h_1(n), \dots, h_{M-1}(n)]^T$. 因此, 由 (2) 式知, 其待辨识系统中每一个冲击响应系数可以用下面的模型预测为

$$h_i(n+1) = 2h_i(n) - h_i(n-1), \quad i = 0, \dots, M-1. \quad (10)$$

由 (10) 式可以看到, 对于每一个系统冲击响应的系数, 其一阶多项式预测模型的形式是一致的. 所以, 由全部冲击响应系数组成的向量就满足如下的多项式预测模型:

$$\mathbf{h}(n+1) = 2\mathbf{h}(n) - \mathbf{h}(n-1). \quad (11)$$

这里, (11) 式可以改写成如下状态空间方程:

$$\mathbf{H}(n) = \mathbf{F}\mathbf{H}(n-1), \quad (12)$$

其中

$$\mathbf{H}(n) = \begin{bmatrix} \mathbf{h}(n+1) \\ \mathbf{h}(n) \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} 2\mathbf{I}_M & -\mathbf{I}_M \\ \mathbf{I}_M & \mathbf{0}_M \end{bmatrix},$$

$\mathbf{I}_M$ 表示 $M \times M$ 的单位阵, M 为系统冲击响应系数 $\mathbf{h}(n)$ 的向量长度, $\mathbf{0}_M$ 为 $M \times M$ 的零矩阵.

3.2 一阶多项式预测模型下的最优的自适应滤波算法

待辨识系统的输入输出模型可以描述为^[16]

$$y(n) = \mathbf{h}(n)^T \mathbf{x}(n) + v(n), \quad (13)$$

其中, n 为时刻序号, T 表示转置, $\mathbf{h}(n) = [h_0(n), h_1(n), \dots, h_{M-1}(n)]^T$ 表示 M 个待辨识系统冲击响应的系数向量, $\mathbf{x}(n) = [x(n), x(n-1), \dots, x(n-M+1)]^T$ 为系统最近的 M 个输入信号, $v(n)$ 是与待辨识系统输入信号 $x(n)$ 独立的系统观测噪声, 通常假定是 Gauss 白噪声, $y(n)$ 为待辨识系统的输出. $\sigma_v^2 = E[v^2(n)]$ 为系统观测噪声的方差.

对于这样的系统辨识问题, 一个直接的方法是采用最小二乘法 (LS, least square) 来解决. 使用最小二乘法的明显的好处就是简单快捷. 但是最小二乘法的精度是随数据量递增而提高的, 要想得到足够精度则需要大量的数据. 这个数据量可能数以万计的. 这样的话, 计算复杂度就太大了. 同时由于通常的最小二乘法对于数据是批处理的, 难以应用于实时处理. 而实际应用中通常需要进行实时处理. 所以这样的系统辨识问题通常都采用基于最小均方、归一化最小均方以及迭代最小二乘法^[1-16]等的自适应滤波方法. 本文提出的基于多项式预测模型的最优自适应滤波算法也是解决这样一个问题的自适应滤波算法. 本文接下来的部分将会具体介绍该自适应滤波算法, 并证明了在模型准确的情况下, 本算法是最优的.

对于待辨识的时不变系统, 其系统冲击响应系数虽然是未知的, 但是它是不随时间的改变而改变的. 即对于任意时刻 n 其时不变冲击响应系数向量都应该满足如下递推关系:

$$\mathbf{h}(n+1) = \mathbf{h}(n) = \mathbf{h}(n-1), \quad (14)$$

即对于任意时刻 n 每一个滤波器系数都满足

$$h_i(n+1) = h_i(n) = h_i(n-1), \quad i = 0, \dots, M-1. \quad (15)$$

如将 (15) 式代入 (10) 式, 可以发现 (10) 式成立. 因此 (15) 式为 (10) 式的一个特例, 即 (14) 式为 (11) 式的一个特例. 这说明待辨识时不变系统其冲击响应系数也可以用一阶多项式预测模型准确描述. 即 (12) 式的状态空间方程也可以描述待辨识的时不变系统冲击响应系数向量 $\mathbf{h}(n)$.

(13) 式可以视为对待辨识系统的冲击响应 $\mathbf{h}(n)$ 的观测, 对应 (12) 式将其改写为

$$\mathbf{Y}(n) = \mathbf{X}(n)\mathbf{H}(n) + \mathbf{w}(n), \quad (16)$$

其中 $\mathbf{Y}(n) = [y(n+1), y(n)]^T$, $\mathbf{w}(n) = [v(n+1), v(n)]^T$,

$$\mathbf{X}(n) = \begin{bmatrix} \mathbf{x}^T(n+1) & \mathbf{0} \\ \mathbf{0} & \mathbf{x}^T(n) \end{bmatrix}, \quad \mathbf{H}(n) = \begin{bmatrix} \mathbf{h}(n+1) \\ \mathbf{h}(n) \end{bmatrix}.$$

其噪声 $\mathbf{w}(n)$ 的协方差阵为

$$\mathbf{R} = E[\mathbf{w}(n)\mathbf{w}^T(n)] = \begin{bmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_v^2 \end{bmatrix}.$$

一旦将 (12) 和 (16) 式分别看作 Kalman 滤波器的状态方程和观测方程, 那么系统 $\mathbf{h}(n)$ 的自适应滤波问题可以转化为一个 Kalman 滤波问题来解决. 即: 利用一阶多项式预测模型描述的待辨识时

变/时不变系统冲击响应系数的自适应滤波算法, 可以递推表示为^[21]

$$\mathbf{P}(n) = \mathbf{F}\mathbf{\Sigma}(n-1)\mathbf{F}^T, \quad (17)$$

$$\mathbf{G}(n) = \mathbf{P}(n)\mathbf{X}^T(n)[\mathbf{X}(n)\mathbf{P}(n)\mathbf{X}^T(n) + \mathbf{R}]^{-1}, \quad (18)$$

$$\mathbf{\Sigma}(n) = [\mathbf{I} - \mathbf{G}(n)\mathbf{X}(n)]\mathbf{P}(n), \quad (19)$$

$$\hat{\mathbf{H}}(n) = \mathbf{F}\hat{\mathbf{H}}(n-1) + \mathbf{G}(n)(\mathbf{Y} - \mathbf{X}(n)\mathbf{F}\hat{\mathbf{H}}(n-1)). \quad (20)$$

其中, $\mathbf{P}(n)$ 和 $\mathbf{\Sigma}(n)$ 分别为 Kalman 滤波的预测误差的方差和 Kalman 滤波的估计误差的方差; $\mathbf{G}(n)$ 为 Kalman 增益; $\hat{\mathbf{H}}(n) = \begin{bmatrix} \hat{\mathbf{h}}^{(n+1)} \\ \hat{\mathbf{h}}(n) \end{bmatrix}$ 为 Kalman 滤波器对 $\mathbf{H}(n)$ 的估计.

众所周知, 在 Kalman 滤波器开始工作的初始时刻以及状态发生突变的时刻, Kalman 滤波器需要经历一段过渡过程才能跟上状态的变化. 而对于 (12) 式描述的待辨识时变/时不变系统冲击响应系数的多项式预测模型, 虽然其状态模型 (12) 可以准确描述任意系数的一阶多项式变化的规律, 可是, 如果系统冲击响应系数的状态在某一时刻突变到另一个状态 (例如: 系统冲击响应系数发生了阶跃突变), 那么包含突变点的冲击响应系数的状态的变化规律就不能由一阶多项式完全描述. 由于这样的突变点, 基于 (12) 式描述的状态模型的 Kalman 滤波跟踪需要经历一段过渡过程才能跟上状态的变化. 针对这种状况, 下面我们将讨论如何探测待辨识时变/时不变系统冲击响应系数是否发生突变, 如何缩短 Kalman 滤波器跟踪时的过渡过程, 如何加速自适应算法收敛到稳态值的方法.

假如在从时刻 n 到时刻 $n+1$ 的时候, 某一待辨识的系统冲击响应系数发生了突变, 从而包含突变点的状态不能用 (12) 式精确描述, 即 $n+1$ 时刻的 $\mathbf{h}(n+1)$ 不能通过 (12) 式由前 2 个时刻的 $\mathbf{h}(n)$ 和 $\mathbf{h}(n-1)$ 准确预测. 那么我们可以将 (12) 式写为

$$\mathbf{H}(n) = \mathbf{F}\mathbf{H}(n-1) + \mathbf{u}(n), \quad (21)$$

其中 $\mathbf{u}(n)$ 为附加的状态噪声, $\mathbf{Q} = E[\mathbf{u}(n)\mathbf{u}^T(n)]$ 为状态噪声的协方差阵, 即认为此时状态方程存在模型不确定性, 而其不确定性的用一个附加的状态噪声描述. 文献 [26] 告诉我们只要该噪声取得足够大, 就能保证该状态方程 (21) 能够准确描述状态的变化. 此时, 就可以将 Kalman 递推方程 (17)–(20) 式改写为

$$\mathbf{P}(n) = \mathbf{F}\mathbf{\Sigma}(n-1)\mathbf{F}^T + \mathbf{Q}, \quad (22)$$

$$\mathbf{G}(n) = \mathbf{P}(n)\mathbf{X}^T(n)[\mathbf{X}(n)\mathbf{P}(n)\mathbf{X}^T(n) + \mathbf{R}]^{-1}, \quad (23)$$

$$\mathbf{\Sigma}(n) = [\mathbf{I} - \mathbf{G}(n)\mathbf{X}(n)]\mathbf{P}(n), \quad (24)$$

$$\hat{\mathbf{H}}(n) = \mathbf{F}\hat{\mathbf{H}}(n-1) + \mathbf{G}(n)(\mathbf{Y} - \mathbf{X}(n)\mathbf{F}\hat{\mathbf{H}}(n-1)). \quad (25)$$

那么如何来检测系统状态突变的发生呢? 由于当 Kalman 滤波的状态方程准确时, Kalman 滤波器的新息序列为零均值的 Gauss 白噪声^[27], 那么可以通过新息序列的均值检验来判断多项式预测模型是否准确, 即系统状态是否发生了突变. Kalman 滤波器的新息表达式为

$$\mathbf{S}(n) = \mathbf{G}(n)(\mathbf{Y} - \mathbf{X}(n)\mathbf{F}\hat{\mathbf{H}}(n-1)), \quad (26)$$

新息方差可以表示为

$$\mathbf{P}_s(n) = \mathbf{G}(n)\mathbf{P}(n)\mathbf{G}^T(n) + \mathbf{R}. \quad (27)$$

根据文献 [27], $\mathbf{S}^T(n)\mathbf{P}_s^{-1}(n)\mathbf{S}(n)$ 为具有 m 个自由度的 χ^2 变量, m 为新息的维数. 因此可以通过常用的 χ^2 检验法来验证新息均值是否为零. 即在显著性水平 α 之下, 如果 $\mathbf{S}^T(n)\mathbf{P}_s^{-1}(n)\mathbf{S}(n) > \chi_\alpha^2$, 则认为新息均值不为零, 即此时存在多项式预测的状态模型不能描述的系统状态突变情况. 此时, 由于 (12) 式的模型是存在模型误差的, 因此状态预测也是不准确的, 此时, 应设法加大观测值对状态估计的修正作用. 数学证明表明: 状态噪声 $\mathbf{Q}$ 值取得越大, 信息加权 $\mathbf{G}(n)$ 就越大 [27]. 因此, 当 (12) 式不能准确描述系统状态时, 采用 (21) 式以及足够大的状态噪声 $\mathbf{Q}$, 可以实现 Kalman 滤波器在系统状态发生突变的时候使算法的收敛. 因此我们可以在显著性水平 α , 当 $\mathbf{S}^T(n)\mathbf{P}_s^{-1}(n)\mathbf{S}(n) > \chi_\alpha^2$ 时, 即系统状态发生突变时, 取 $\mathbf{Q} = \beta\mathbf{R}$, β 为一个较大的正整数, 使得在 (12) 式多项式预测的状态模型不准确时, 用新息来加快 Kalman 滤波算法的收敛速度; 而在同样的显著性水平 α , 当 $\mathbf{S}^T(n)\mathbf{P}_s^{-1}(n)\mathbf{S}(n) < \chi_\alpha^2$ 时, 即 (12) 式多项式预测的状态模型能正确描述系统状态时, 让 $\mathbf{Q} = \mathbf{0}$.

更高阶的多项式预测模型的自适应滤波算法可以用相同的方法得到.

3.3 性能分析

这里 Kalman 滤波在极限情况下的形式取为

$$\boldsymbol{\Sigma} = \lim_{n \rightarrow \infty} \boldsymbol{\Sigma}(n), \quad \mathbf{P} = \lim_{n \rightarrow \infty} \mathbf{P}(n), \quad \mathbf{G} = \lim_{n \rightarrow \infty} \mathbf{G}(n), \quad (28)$$

将之代入 (17)–(19) 式后可得极限 $\mathbf{P}$ 是如下 Riccati 方程的唯一半正定解 [28]

$$\mathbf{P} = \mathbf{F}[\mathbf{P} - \mathbf{P}\mathbf{X}^T(n)(\mathbf{X}(n)\mathbf{P}\mathbf{X}^T(n) + \mathbf{R})^{-1}\mathbf{X}(n)\mathbf{P}]\mathbf{F}^T, \quad (29)$$

$$\Leftrightarrow \mathbf{F}\mathbf{P}\mathbf{F}^T - \mathbf{P} - \mathbf{F}\mathbf{P}\mathbf{X}^T(n)(\mathbf{X}(n)\mathbf{P}\mathbf{X}^T(n) + \mathbf{R})^{-1}\mathbf{X}(n)\mathbf{P}\mathbf{F}^T = \mathbf{0}. \quad (30)$$

通常, 该 Riccati 方程只能通过迭代的办法进行数值求解, 而没有解析解 [29]. 然而, 显然将 $\mathbf{P} = \mathbf{0}$ 代入可以使得 (29) 式成立. 既然 (29) 式只有唯一解, $\mathbf{P} = \mathbf{0}$ 又是 (29) 式的解, 那么可以得出结论 $\mathbf{P} = \mathbf{0}$ 为 (29) 式的唯一解. 因此, 递推 (17)–(20) 式中的 Kalman 滤波器的预测误差 $\mathbf{P}(n)$ 是逼近于 0 的. 根据 (19) 式, Kalman 滤波器的估计误差 $\boldsymbol{\Sigma}(n)$ 也是逼近于 0 的. 因此, 本算法的估计性能是最优的. 同时, 可以看到, $\mathbf{P} = \mathbf{0}$ 的结论与系统的观测噪声阵 $\mathbf{R}$ 没有关系. 因此这里还可以得到本算法对系统观测噪声的估计误差鲁棒的结论.

虽然 NPVSS-NLMS 算法在理论上也可以得到与本算法相同的结论 [16], 但是通过分析可以看出该算法中这样的结论的得出是建立在对先验误差的估计及系统的观测噪声的方差的准确已知的基础上的. 如果这两个量的估计不够准确, 那么 NPVSS-NLMS 算法就不能达到最优. 而 NPVSS-NLMS 算法中采用的估计先验误差的方法是一种对估计误差的进行滑动平均的方法. 这样的估计方法会导致一定的估计误差. 同时, 在实际中, 对系统观测噪声的估计也不能达到十分准确. 而上述分析表明: 本文算法则不需要像 NPVSS-NLMS 算法那样要估计先验误差和准确知道观测噪声的方差. 因此, 在实际应用的意义上 (即: 一般估计得到的先验误差和观测噪声的方差不准确的时候), NPVSS-NLMS 的算法性能将不如本文提出的自适应滤波算法. 当然, 本文算法也有其自身的局限性. 本文算法由于使用 Kalman 滤波器, 从而引入了很多矩阵相乘计算以及矩阵求逆的运算, 而 NPVSS-NLMS 算法的运算多为简单的标量运算. 另外, NPVSS-NLMS 算法不需要进行模型建模和额外的状态判断, 而本文算法需要对待辨识的系统冲击响应系数预先进行多项式建模, 而且需要对系统是否发生突变进行判断. 因此本文算法的计算量会比 NPVSS-NLMS 算法要大一些. 由于计算复杂度的分析并不是本文的重点, 因此这里就不进行详细的分析了.

另外, 众所周知在 Gauss 白噪声环境中以及状态模型准确的情况下, Kalman 滤波是最大似然、最大后验和最小均方等统计意义下的最优滤波^[19-21]. 这也从另一个角度证明了本算法的最优性.

本算法的性能优势是建立在多项式预测模型是否可以描述每一个滤波器系数变化的准确程度的基础上的. 因此, 我们需要通过多项式的阶数和采样数据的密度来进行折中, 显然, 如果使用的多项式模型不能准确描述每一个滤波器系数变化, 那么本算法的性能是不能达到最优的. 因此对于快变的非多项式规律变化的系统冲击响应系数的辨识的应用场景, 我们需要缩短数据的采样时间 (即加大数据的采集密度) 和使用高阶的多项式预测模型来提高算法性能; 而对于辨识缓变的或不变的系数则可以采用低阶多项式预测模型以及较低的数据采样密度.

由于更高阶的多项式预测模型下的算法采用的递推方程也满足 (17)–(20) 式的形式, 因此这里的性能分析也可以推广到高阶的多项式预测模型的情况.

4 算法仿真

这里, 假设需要辨识的系统冲击响应系数的长度为 $M=64$, 采样率为 8 kHz. $x(n)$ 采用一个白 Gauss 随机信号作为系统的输入. 系统输出 $y(n)$ 中加入了一个独立的 Gauss 噪声 $v(n)$. 输入信号与该噪声的信噪比为 30 dB. 仿真时间长度为 10 s. 采用的对比方法为文献 [16] 中的 NPVSS-NLMS、SM-NLMS 的方法以及一种定步长 NLMS 方法. 其中 NPVSS-NLMS 是最近的文献报道的一种新的自适应滤波算法. 该算法理论上也是一种最优算法, 其性能相比传统算法有一定优势. 但是根据第 3 节的分析, 本文提出算法的性能应优于该算法. 仿真的结果也验证了这一点. SM-NLMS 方法和定步长 NLMS 方法是用来作为对比的两种经典的自适应滤波算法^[13,14]. 采用的性能指标为自适应滤波领域中一个通用的性能指标, 称之为归一化偏差 (dB): $20 \log_{10}(\|\mathbf{h}(n) - \hat{\mathbf{h}}(n)\|_2 / \|\mathbf{h}(n)\|_2)$. 本文共仿真了 7 种不同的情况: 在系统 $\mathbf{h}(n)$ 为时不变的情况下, σ_v^2 准确知道的情况以及 σ_v^2 的估计值与真实值偏差 10% 和 50% 的情况. 另外还有 σ_v^2 准确知道的情况下, 两种系统 $\mathbf{h}(n)$ 呈线性变化的情况、系统 $\mathbf{h}(n)$ 呈正弦变化的情况以及系统 $\mathbf{h}(n)$ 发生阶跃式突变的情况. 所有仿真的算法的初始值相同: $\hat{\mathbf{X}}(0)$ 取 0 向量, $\Sigma(0)$ 取单位阵. 显著性水平 α 取为 0.01, β 取为 1000. 仿真的结果表明了在这几种情况下, 本算法的性能都优于对比算法.

图 1 为系统 $\mathbf{h}(n)$ 时不变, σ_v^2 测量准确情况下 4 种算法的比较. (1) 为 SM-NLMS 算法, (2) 为 NPVSS-NLMS 算法, (3) 为定步长 NLMS 算法, 步长为 $0.055[\delta + \mathbf{x}^T(n)\mathbf{x}(n)]^{-1}$, (4) 为本文算法. 其中 δ 为避免除 0 而引入一个小的常数. 可以看到在系统 $\mathbf{h}(n)$ 时不变, σ_v^2 测量准确的情况下, 本文算法对比算法有将近 10 dB 的性能优势.

图 2, 3 分别为系统 $\mathbf{h}(n)$ 时不变, σ_v^2 测量偏差 10% 以及 50% 的情况下, 4 种算法的比较. (1) 为 SM-NLMS 算法, (2) 为 NPVSS-NLMS 算法, (3) 为定步长 NLMS 算法, 步长为 $0.055[\delta + \mathbf{x}^T(n)\mathbf{x}(n)]^{-1}$, (4) 为本文算法. 可以看到本文算法受 σ_v^2 测量偏差的影响较小, 而对比较算法, 尤其是 NPVSS-NLMS 算法受 σ_v^2 测量偏差的影响较大.

图 4, 5 皆为系统 $\mathbf{h}(n)$ 线性变化, σ_v^2 测量准确的情况下的 4 种算法的比较. (1) 为 SM-NLMS 算法, (2) 为 NPVSS-NLMS 算法, (3) 为定步长 NLMS 算法, 步长为 $0.055[\delta + \mathbf{x}^T(n)\mathbf{x}(n)]^{-1}$, (4) 为本文算法. 系统 $\mathbf{h}(n)$ 的变化情况分别设定为: $\mathbf{h}(t) = \mathbf{h}_0 + 0.4t$ 及 $\mathbf{h}(t) = \mathbf{h}_0 + 0.08t$, 其中 $\mathbf{h}_0$ 同上述各仿真,

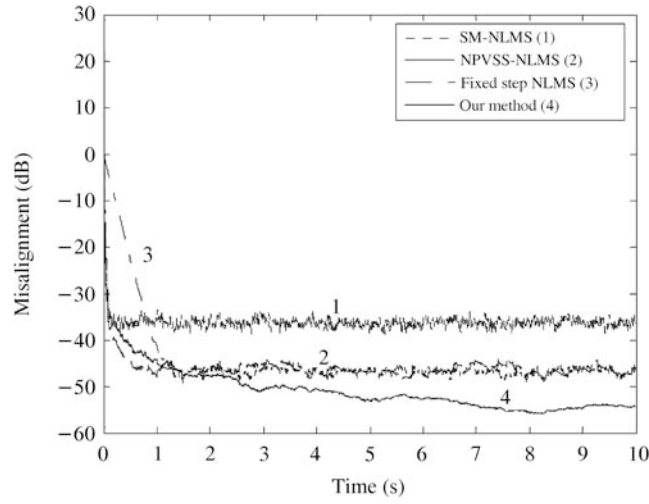
图 1 系统 $h(n)$ 时不变, σ_v^2 测量准确

Figure 1 Misalignment of the algorithms when $h(n)$ is time-invariant and the measurement of σ_v^2 is accurate.

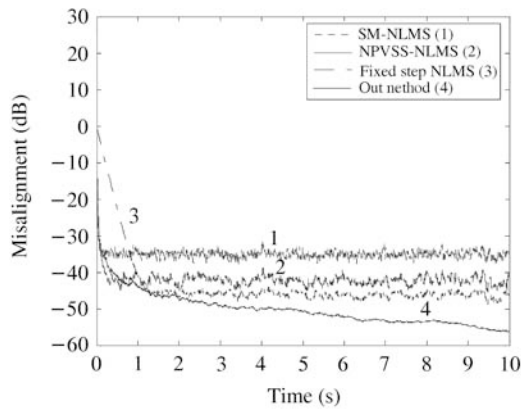
图 2 系统 $h(n)$ 时不变, σ_v^2 测量偏差 10%

Figure 2 Misalignment of the algorithms when $h(n)$ is time-invariant and the measurement of σ_v^2 has a deviation of 10%

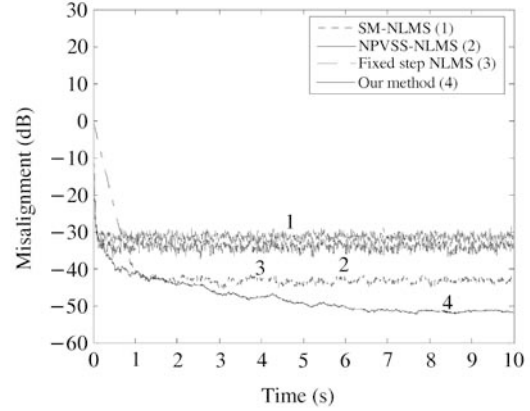
图 3 系统 $h(n)$ 时不变, σ_v^2 测量偏差 50%

Figure 3 Misalignment of the algorithms when $h(n)$ is time-invariant and the measurement of σ_v^2 has a deviation of 50%

t 为时间, 单位为秒 (s). 由于仿真中测量噪声的方差并没有随着系统的改变而改变, 因此系统的信噪比会随着系统的改变而提高. 这就解释了为什么图 4 的最终的性能指标要略好于图 5 的性能指标. 从图 4, 5 可以看出, 本算法相比对比算法的性能优势相当明显.

图 6 为系统 $h(n)$ 呈正弦变化, σ_v^2 测量准确的情况下, 本方法与其他方法的比较. (1) 为 SM-NLMS 算法, (2) 为 NPVSS-NLMS 算法, (3) 为定步长 NLMS 算法, 步长为 $0.055[\delta + \mathbf{x}^T(n)\mathbf{x}(n)]^{-1}$, (4) 为本文算法. 系统 $h(n)$ 的变化情况为: $h(t) = h_0 + 0.5\sin(7\pi/16 + \pi t/160)$, 其中 h_0 为与上述仿真相同的初始系统响应, t 为时间, 单位为秒 (s). 可以看到本算法较对比算法的性能优势.

图 7 为系统 $h(n)$ 在第 5 s 时发生阶跃突变, σ_v^2 测量准确的情况下本算法和其他方法的比较. (1) 为 SM-NLMS 算法, (2) 为 NPVSS-NLMS 算法, (3) 为定步长 NLMS 算法, 步长为 $0.055[\delta + \mathbf{x}^T(n)\mathbf{x}(n)]^{-1}$,

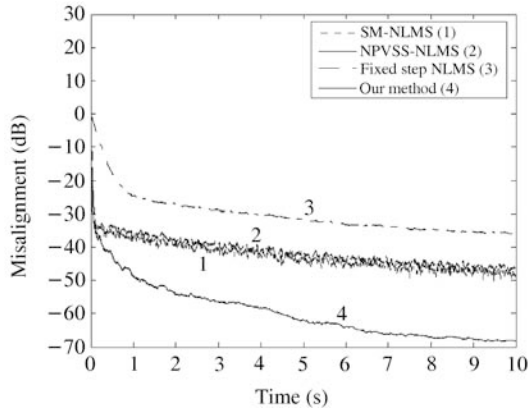


图 4 系统 $h(n)$ 线性变化: $h(t)=h_0+0.4t$, σ_v^2 测量准确

Figure 4 Misalignment of the algorithms when $h(n)$ is time-variant with $h(t) = h_0 + 0.4t$ and the measurement of σ_v^2 is accurate

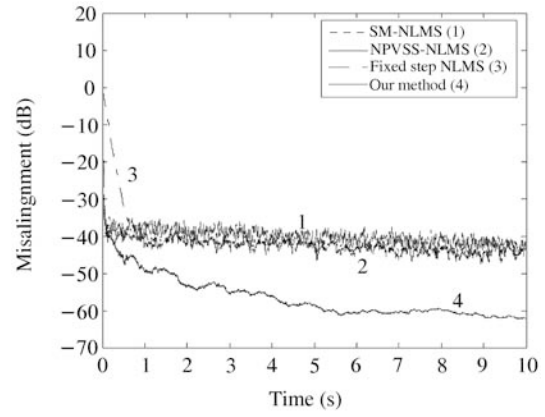


图 5 系统 $h(n)$ 线性变化: $h(t)=h_0+0.08t$, σ_v^2 测量准确

Figure 5 Misalignment of the algorithms when $h(n)$ is time-variant with $h(t) = h_0 + 0.08t$ and the measurement of σ_v^2 is accurate

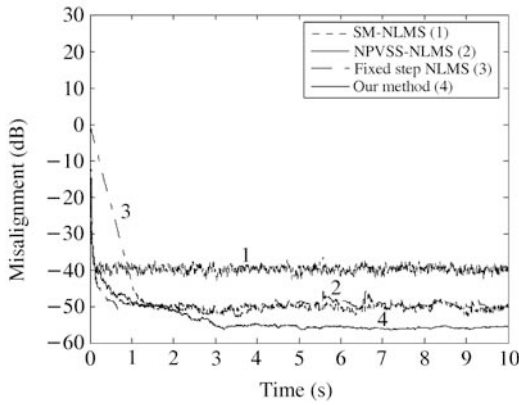


图 6 系统 $h(n)$ 呈正弦慢变: $h(t)=h_0+0.5\sin(7\pi/16+\pi t/160)$, σ_v^2 测量准确

Figure 6 Misalignment of the algorithms when $h(n)$ is time-variant with $h(t) = h_0 + 0.5\sin(7\pi/16 + \pi t/160)$ and the measurement of σ_v^2 is accurate

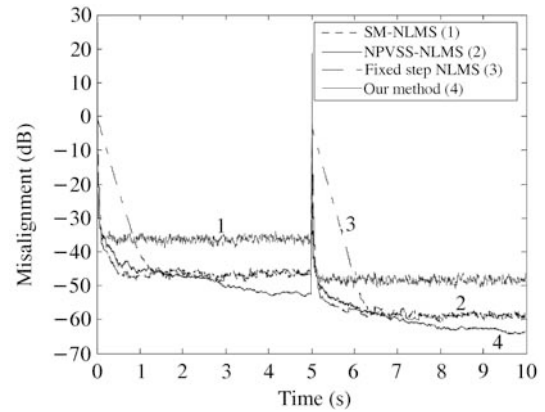


图 7 系统 $h(n)$ 在第 5 s 时发生突变, σ_v^2 测量准确

Figure 7 Misalignment of the algorithms when $h(n)$ is time-invariant but a step change happens at 5 s and the measurement of σ_v^2 is accurate

(4) 为本文算法. 突变方式为 $h = 4h_0$, 其中 h_0 为与上述仿真相同的初始系统响应, h 为发生阶跃突变后的系统响应. 可以看到在突变前后, 本算法的性能都优于对比算法. 同仿真 4、5, 由于突变后系统的观测噪声并没有改变, 因此系统的信噪比会变大, 性能指标会好于突变前. 图 1 至图 7 的仿真结果均表明: 在本文的几种仿真情况下, 本文算法的性能都优于已知的算法. 这是因为 Kalman 滤波理论特性保证了本文算法的性能.

5 结论

本文提出了一种描述待辨识的自适应滤波系统冲击响应系数的多项式预测模型以及相应的最优

自适应滤波算法. 该多项式预测模型可以准确或足够精确的逼近描述系统冲击响应的时变/时不变的系数的状态. 因此以此多项式预测模型为状态方程, 以系统输入输出关系作为观测方程, 用 Kalman 滤波的方法来解这个自适应滤波问题的方法是在最小均方误差意义下最优的. 同时, 由于该方法中的噪声是 Gauss 的, 因此该方法也是最大似然和最大后验意义下最优的. 文章给出了算法的模型、递推方程和性能分析, 并通过了大量不同的仿真结果验证了算法性能的优越性.

参考文献

- 1 Widrow B, Stearns S D. Adaptive Signal Processing. Englewood Cliffs, NJ: Prentice-Hall, 1985
- 2 Haykin S. Adaptive Filter Theory. 4th ed. Upper Saddle River, NJ: Prentice-Hall, 2002
- 3 Benesty J, Huang Y. Adaptive Signal Processing—Applications to Real-World Problems. Berlin: Springer-Verlag, 2003
- 4 Huang Y, Benesty J, Chen J. Acoustic MIMO Signal Processing. Boston, MA: Springer, 2006
- 5 Sayed A H. Fundamentals of Adaptive Filtering. New York: Wiley, 2003
- 6 Glentis G O, Berberidis K, Theodoridis S. Efficient least squares adaptive algorithms for FIR transversal filtering. IEEE Signal Process Mag, 1999, 16: 13–41
- 7 Guo L, Ljung L, Wang G J. Necessary and sufficient conditions for stability of LMS. IEEE Trans Automat Control, 1997, 42: 761–770
- 8 Hassibi B, Sayed A H, Kailath T. H^2 optimality of the LMS algorithm. IEEE Trans Signal Process, 1996, 44: 267–280
- 9 Evans J B, Xue P, Liu B. Analysis and implementation of variable step size adaptive algorithms. IEEE Trans Signal Process, 1993, 41: 2517–2535
- 10 Aboulnasr T, Mayyas K. A robust variable step-size LMS-type algorithm: analysis and simulations. IEEE Trans Signal Process, 1997, 45: 631–639
- 11 Pazaitis D I, Constantinides A G. A novel kurtosis driven variable step-size adaptive algorithm. IEEE Trans Signal Process, 1999, 47: 864–872
- 12 Mader A, Puder H, Schmidt G U. Step-size control for acoustic echo cancellation filters—an overview. Signal Process, 2000, 80: 1697–1719
- 13 Morgan D R, Kratzer S G. On a class of computationally efficient, rapidly converging, generalized NLMS algorithm. IEEE Signal Process Lett, 1996, 3: 245–247
- 14 Gollamudi S, Nagaraj S, Kapoor S, et al. Set-membership filtering and a set-membership normalized LMS algorithm with an adaptive step size. IEEE Signal Process Lett, 1998, 5: 111–114
- 15 Shin H C, Sayed A H, Song W J. Variable step-size NLMS and affine projection algorithms. IEEE Signal Process Lett, 2004, 11: 132–135
- 16 Benesty J, Rey H, Vega L R, et al. A nonparametric VSS NLMS algorithm. IEEE Signal Process Lett, 2006, 13: 581–584
- 17 Gendron P J. An empirical Bayes estimator for in-scale adaptive filtering. IEEE Trans Signal Process, 2005, 53: 1670–1683
- 18 Huang D Y, Rahardja S, Huang H B. Maximum a posteriori based adaptive algorithms. In: Asilomar Conference on Signals, Systems and Computers. Asilomar, 2007. 1628–1632
- 19 Steven M K. Fundamentals of Statistical Signal Processing, Volume: Estimation Theory; Fundamentals of Statistical Signal Processing, Volume: Detection Theory (in Chinese). Beijing: Publishing House of Electronics Industry, 2003
- 20 van der Merwe R. Sigma-point Kalman filters for probabilistic inference in dynamic state-space models. Dissertation for the Doctoral Degree. Cape Town, South Africa: Oregon Health & Science University, 2004
- 21 Welch G, Bishop G. An Introduction to the Kalman Filter. <http://www.cs.unc.edu/~welch/kalman/kalmanIntro.htm/>. 2004.
- 22 Weisstein E W. Weierstrass Approximation Theorem. From MathWorld—A Wolfram Web Resource. <http://mathworld.wolfram.com/WeierstrassApproximationTheorem.html>

- 23 Xu L J, Zhang J Q, Yan Y. A wavelet-based multisensor data fusion algorithm. *IEEE Trans Instrum Measur*, 2004, 53: 1539–1545
- 24 Ifeachor E C, Jervis B W. *Digital Signal Processing A Practical Approach*. 2nd ed. Beijing: Publishing House of Electronics Industry, 2003
- 25 Heinonen P, Neuvo Y. FIR-median hybrid filters with predictive FIR substructures. *IEEE Trans Acoust Speech Signal Process*, 1988, 36: 892–899
- 26 Fitzgerald R J. Divergence of the Kalman filter. *IEEE Trans Autom Control*, 1971, 16: 736–747
- 27 Zhang J K. *Linear Model Parameter Estimation and Its Improvement*. Changsha: Publishing House of National University of Defense Technology, 1999
- 28 Chui C K, Cheng G. *Kalman Filtering with Real-Time Application*. Berlin: Springer-Verlag, 1987
- 29 William F, Arnold III, Laub A J. Generalized eigenproblem algorithms and software for algebraic Riccati equations. *P IEEE*, 1984, 72: 1746–1754

An optimal adaptive filtering algorithm with a polynomial prediction model

TAN JiaJia & ZHANG JianQiu*

Department of Electronic Engineering, Fudan University, Shanghai 200433, China

*E-mail: jqzhang@fudan.edu.cn

Abstract A new approach to the optimal adaptive filtering is proposed in this paper. In this approach, a polynomial prediction model is used to describe the time-variant/invariant impulse response coefficients of an identified system. When the polynomial prediction model is viewed as the state equations of the identified impulse response coefficients and the relationships between the inputs and outputs of the system are regarded as the measurements of the states, our adaptive filtering can be achieved in the framework of the Kalman filter. It is understood that Kalman filter is optimal in the sense of the MAP (maximum a posteriori), ML (most likelihood) and MMSE (minimum mean square error) under the linear and Gaussian white noise conditions. As a result, our algorithm is also optimal in the statistical senses as Kalman filter does, provided that the impulse response coefficients can be modeled by a polynomial. Not only do the analytical results of the algorithm but also the simulation results show that our algorithm outperforms the traditional known algorithms.

Keywords adaptive filters, polynomial prediction model, impulse response coefficients, process equation, Kalman filter, optimal