

混合噪声中非线性 RLS 算法的性能分析

熊 鹰¹, 梁树雄², 尹俊勋¹

(1. 华南理工大学电子与通信工程系, 广东广州 510641; 2. 香港城市大学电子工程学系, 香港)

摘 要: 本文给出了在混合噪声中非线性递归最小均方误差算法的性能分析. 该算法即是非线性 RLS (NRLS). 对残差的饱和处理使用了广义限幅函数. 提出了改进的 NRLS 均方分析. 计算机数值模拟表明理论分析与模拟结果符合得很好. 根据该分析, 可以将 NRLS 的收敛和均方误差表示为非线性函数的斜率和限幅水平的函数. 基于归一化的均方误差 (mse), 引入了一个辅助变量, 导出了时变限幅函数, 加速了收敛并且得到了更小的均方误差.

关键词: 递归最小均方误差; RLS; 非线性; 均方误差分析; 混合噪声

中图分类号: TN713 **文献标识码:** A **文章编号:** 0372-2112 (2001) 07-0981-03

Performance Analysis of Nonlinear RLS in Mixture Noise

XIONG Ying¹, LEUNG Shu-hung², YIN Jun-xun¹

(1. Dept. of Electronic Engineering, South China University of Technology, Guangzhou, Guangdong 510641, China;

2. Dept of Electronic Engineering, City University of Hong Kong, China)

Abstract: This paper presents the performance analysis of recursive least square algorithm with error saturation in mixture noise. The algorithm is referred to nonlinear RLS (NRLS). Generalized clipping function is considered for the error saturation nonlinearity. An improved mean square behavior of NRLS is carried out. It is shown that the theoretical analysis and the simulation result are close to each other. From the analysis, we can relate the convergence and the mean square error in terms of the slope and the clipping level of the nonlinear function. Based on the normalized mean square error, an instrumental variable is derived for yielding a variable clipping function to provide fast convergence and small mean square error.

Key words: recursive least square; RLS; nonlinear; mean square analysis; mixture noise

1 引言

线性估计被广泛应用于系统辨识, 模拟, 预测, 均衡以及干扰消除^[1]. 一般地, 输出信号 $y(n)$ 可以表示为

$$y(n) = W^T X(n) + \eta(n) \quad (1)$$

其中 $\eta(n)$ 是观测噪声, W 是加权向量而 $X(n)$ 是输入信号向量.

在高斯噪声中, 递归最小均方误差 (RLS) 算法通常被用于对 W 的估计, 其收敛快并且均方误差较小^[1]. 实际上, 许多自然的噪声源都是非高斯的, 可以证明在非高斯噪声中 RLS 的性能将会下降. 最近, 人们留意到对残差的饱和处理能提高 RLS 在非高斯噪声中的性能^[2, 3]. 本文中, 使用广义限幅函数对残差的作非线性饱和处理. 下面可以看到, 利用该函数的斜率和限幅水平来控制收敛以及均方误差 (mse).

本文提出了一种改进的性能分析方法. 导出了 NRLS 的均方特性. 根据该特性, 可以适当选取广义限幅函数的斜率和限幅水平来加速收敛并同时减小均方误差. 模拟结果表明, 新的理论分析给出了比文[6]中方法和文[7]中通常方法更好的结果.

2 非线性 RLS 算法

假设式(1)中的观测噪声 $\eta(n)$ 为零均值独立同分布 (i. i. d.) 的随机变量, 其方差为 σ_η^2 . 使用混合模型来模拟非高斯噪声, 其密度函数为^[4, 5]

$$f_\eta(x) = (1-A)f_0(x) + Af_1(x) = \sum_{m=0}^1 p_m f_m(x),$$
$$p_0 = 1-A, p_1 = A \quad (2)$$

其中 A 是脉冲指标. $f_m(x)$ 是零均值的高斯密度函数, 其方差分别为 $\sigma_0^2 = \sigma_g^2$, $\sigma_1^2 = \sigma_g^2 + \sigma_i^2$. 其中 σ_g^2 是本底高斯噪声成分的方差, 而 σ_i^2 是脉冲噪声成分的方差. 本底高斯成分和脉冲成分的功率比定义为 $\Gamma = \sigma_g^2 / A\sigma_i^2$.

非线性 RLS 加权向量的更新采用

$$W(n+1) = W(n) + g[e(n)]k(n) \quad (3)$$

其中 $e(n)$ 为残差信号, $k(n)$ 为 Kalman 增益向量. $g(x)$ 是奇对称的残差饱和非线性函数. 将(3)中的 $k(n)$ 换成 $kX(n)$, 就成了文[4]中的非线性 LMS 算法.

本文使用的 $g(x)$ 是非线性的广义限幅函数, 其定义为

$$g(x) = \begin{cases} rsT_0, & x > T_0 \\ sx, & |x| \leq T_0 \\ -rsT_0, & x < -T_0 \end{cases} \quad (4)$$

使用这个非线性函数的首要原因是它很简单. 更重要的是 $g(x)$ 中的斜率参数 s 和限幅参数 r 可以方便地用于控制收敛以及均方误差.

3 非线性 RLS 算法分析

首先假设: 输入信号 $X(n)$ 是独立零均值高斯向量, 其相关矩阵为 $\sigma_x^2 \mathbf{I}$; 观测噪声 $\eta(n)$ 是独立同分布的且分布密度函数为式(2); $X(n)$ 和 $\eta(n)$ 是相互统计独立的. 为方便分析, 定义权重的误差向量为

$$\mathbf{V}(n) = \mathbf{W}(n) - \mathbf{W}_0 \quad (5)$$

其中 \mathbf{W}_0 是所需权重. 根据基本 RLS 算法^[1], 有递推式

$$\mathbf{V}(n+1) = \mathbf{V}(n) + \mathbf{P}(n) \mathbf{X}(n) g[e(n)] \quad (6)$$

其中 $\mathbf{P}(n)$ 是 N 维输入信号向量 $\mathbf{X}(n)$ 的相关矩阵 $\Phi(n)$ 的逆矩阵

$$\mathbf{P}(n) = \Phi^{-1}(n) = \left[\sum_{i=1}^n \omega^{n-i} \mathbf{X}(i) \mathbf{X}^T(i) + \omega^n \delta_0^{-1} \mathbf{I} \right]^{-1} \quad (7)$$

$\omega \in (0, 1]$ 是遗忘因子, $\delta_0 \mathbf{I}$ 是 $\mathbf{P}(n)$ 的初始值.

我们应用“平均原理(averaging principle)”^[7] 来简化式(6)的统计分析计算. 因为在 NRLS 过程中 $\mathbf{P}(n)$ 是缓变的而 $\mathbf{X}(n)$ 和 $e(n)$ 的变化快得多, 根据“平均原理”, 在求乘积的期望时我们可以近似认为 $\mathbf{P}(n)$ 与后者相独立.

3.1 均方特性分析

将式(6)右乘其转置并取数学期望, 得到 $\mathbf{V}(n)$ 相关矩阵 $\mathbf{K}(n)$ 的递推表示

$$\begin{aligned} \mathbf{K}(n+1) = & \mathbf{K}(n) + E\{\mathbf{V}(n)g[e(n)]\mathbf{X}^T(n)\mathbf{P}(n)\} \\ & + E\{\mathbf{P}(n)\mathbf{X}(n)g[e(n)]\mathbf{V}^T(n)\} \\ & + E\{\mathbf{P}(n)\mathbf{X}(n)g^2[e(n)]\mathbf{X}^T(n)\mathbf{P}(n)\} \end{aligned} \quad (8)$$

应用类似文[8]中的方法, 可以得到式(8)中第二和第三项为 $E\{\mathbf{V}(n)g[e(n)]\mathbf{X}^T(n)\mathbf{P}(n)\} = E\{\mathbf{P}(n)\mathbf{X}(n)g[e(n)]\mathbf{V}^T(n)\} = -\alpha_1 \gamma_1^{-1}(n) \mathbf{K}(n)$ (9)

其中 $\alpha_1 = \sum_{m=0}^1 p_m \phi_1(g, \sigma_e)$, $\gamma_1(n) = \rho_1(n) + \omega^n \delta_0^{-1} \sigma_x^{-2}$, $\phi_1(g, \sigma) = E_G \{g^{(1)}(x)\}$, $\sigma_m^2 = \sigma_m^2 + \sigma_x^2 \text{Tr}[\mathbf{K}(n)]$,

$$\rho_1(n) = \sum_{i=1}^n \omega^{n-i} = (1 - \omega^n) / (1 - \omega)$$

式(8)中最后一项是整个分析中的难点, 这里我们考虑仅取其对角线向量. 利用恒等式 $E[\Phi(n)\mathbf{P}(n)\mathbf{A}\mathbf{P}(n)\Phi(n)] = \mathbf{A}$ 展开 $\Phi(n)$, 经过冗长运算, 我们得到

$$\begin{aligned} \text{diag}\{E\{\mathbf{P}(n)E\{\mathbf{X}(n)g^2[e(n)]\mathbf{X}^T(n)\}\mathbf{P}(n)\}\} \\ = \sigma_x^{-4} \mathbf{G}(n) \mathbf{d}_L \end{aligned} \quad (10)$$

其中 $\mathbf{G}(n) = \gamma_2^{-1}(n) [\mathbf{I} - \mathbf{1}\mathbf{1}^T / (\gamma_2(n) \rho_2^{-1}(n) + N)]$,

$$\mathbf{d}_L = \text{diag}\{E\{\mathbf{X}(n)g^2[e(n)]\mathbf{X}^T(n)\}\},$$

$$\gamma_2(n) = \rho_2(n) + \gamma_1^2(n),$$

$$\rho_2(n) = (1 - \omega^{2n}) / (1 - \omega^2), \mathbf{1} = (1, 1, \dots, 1)^T$$

令 $\mathbf{d}_k(n) = \text{diag}[\mathbf{K}(n)]$, 从文[8]中, 我们知道

$$\mathbf{d}_L = \alpha_2 \sigma^2 \mathbf{1} + \alpha_2 \sigma_x^4 \mathbf{1}\mathbf{1}^T \mathbf{d}_K(n) + \alpha_3 \sigma_x^4 \mathbf{d}_K(n) \quad (11)$$

其中 $\alpha_{20} = \sum_{m=0}^1 p_m \sigma_m^2 \phi_2(g, \sigma_e)$, $\alpha_2 = \sum_{m=0}^1 p_m \phi_2(g, \sigma_e)$, $\alpha_3 = \sum_{m=0}^1 p_m \phi_3(g, \sigma_e)$, $\phi_2(g, \sigma) = E_G \{g^2(x) / \sigma^2\}$, $\phi_3(g, \sigma) = E_G \{2gg^{(1)}(x)\}$, $\mathbf{1} = (1, 1, \dots, 1)^T$

将式(11)代入式(10), 连同式(9)代入式(8)并取对角线向量, 我们得到

$$\mathbf{d}_k(n+1) = \mathbf{H}(n) \mathbf{d}_k(n+1) + \alpha_2 \sigma_x^2 \mathbf{G}(n) \mathbf{1} \quad (12)$$

其中 $\mathbf{H}(n) = [1 - 2\alpha_1 \gamma_1^{-1}(n)] \mathbf{I} + \mathbf{G}(n) (\alpha_2 \mathbf{1}\mathbf{1}^T + \alpha_3 \mathbf{I})$

我们定义均方误差(mse)为 $\sigma_{\text{mse}}^2(n) = \sigma_x^2 \text{Tr}[\mathbf{K}(n)]$, 若仅考虑 $\mathbf{H}(n)$ 的最大特征值, 可以写出递推式

$$\sigma_{\text{mse}}^2(n+1) = \lambda_1 \sigma_{\text{mse}}^2(n) + N \alpha_2 \rho_2(n) + \gamma_2(n) \quad (13)$$

其中 $\lambda_1 = 1 - 2\alpha_1 \gamma_1^{-1}(n) + (N \alpha_2 + \alpha_3) [N \rho_2(n) + \gamma_2(n)]^{-1}$.

由式(13), 我们可以递推计算非线性 RLS 的均方误差性能. 更重要地, 由下一节我们可以看到, 我们可以通过在每一次递归中即时调整式(4)中非线性函数的参数, 斜率 s 和限幅水平 r , 使均方误差达极小, 从而提高性能. 模拟实验表明, 新的算法能得到更快的收敛速度和更小的均方误差.

3.2 时变限幅函数

本节中, 我们引入一个新的辅助变量 $\sigma_l^2(n)$ 来调整限幅函数式(4)的斜率 s 和限幅水平 r 以期加速 NRLS 的学习过程. 定义

$$\sigma_l^2(n) = \sigma_{\text{mse}}^2(n) / \sigma_{\eta}^2 \quad (14)$$

为了能导出简单的时变限幅函数, 假设限幅水平 T_0 能随观测噪声自动在 $T'_0 \sigma_0^2$ 和 $T'_0 \sigma_{e_1}^2$ 间转换. 利用这个假设, 用 σ_{η}^2 归一化式(13), 得到

$$\sigma_l^2(n+1) = \lambda_1 \sigma_l^2(n) + N \alpha_2 [N \rho_2(n) + \gamma_2(n)]^{-1} \quad (15)$$

将 $\sigma_l^2(n+1)$ 对 s 和 r 求导并令其为零, 我们可以得到在时刻 $n+1$ 最小化 $\sigma_l^2(n)$ 即归一化 $\sigma_{\text{mse}}^2(n)$ 的 s 和 r 值, $s(n)$ 和 $r(n)$. 为节省篇幅起见, 略去冗长繁琐的运算过程而直接给出结果如下:

$$r(n) = \frac{f(t) \{2t^3 f(t) + [N(1 + \sigma_l^2(n)) + 2] (tf(t) - \text{erf}(t))\}}{2t (tf(t) - \text{erf}(t)) [N(1 + \sigma_l^2(n)) (1 - \text{erf}(t)) + f(t)]} \quad (16)$$

$$r(n) s(n) = \frac{[N \rho_2(n) + \gamma_2(n)] f(t)}{2 \gamma_1(n) t [N(1 + \sigma_l^2(n)) (1 - \text{erf}(t)) + tf(t)]} \quad (17)$$

其中 $t = T'_0 / \sqrt{2}$, $f(t) = \frac{2}{\sqrt{\pi}} \exp(-t^2)$,

$$\text{erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t \exp(-x^2) dx$$

由式(16)、(17)每次更新限幅函数(4)中的参数, 就能得到较小的 mse, 加快收敛速度. 观察得到, 要达到最佳的性能, $\sigma_l^2(n)$ 的初始值选取是与信噪比有关的.

4 模拟实验结果

用文[1]中的系统辨识例子作模拟实验以衡量使用广义

限幅函数的非线性 RLS 算法的性能. 实验中, 未知系统的脉冲响应设为

$$W_0 = (0.2, -0.4, 0.6, -0.8, 1.0, -0.8, 0.6, -0.4, 0.2)^T / 1.844$$

系统 W_0 的输入信号是零均值单位方差的高斯白噪声. 信噪比(SNR)定义为 σ_s^2/σ_n^2 . 实验中使用 $SNR = 9.63\text{dB}$. 在噪声模型中, 取脉冲指标 $A = 0.1$ 和功率比 $\Gamma = 0.1$. 自适应滤波器长度与 W_0 长度相同为 9, 且它的初始值为零.

第一个实验是用式(13)计算的均方误差比较线性 RLS 和 NRLS 性能(图 1). NRLS 使用固定的广义限幅函数, 其参数为 $s = 1$ 和 $r = 0, 1$, 限幅水平取为

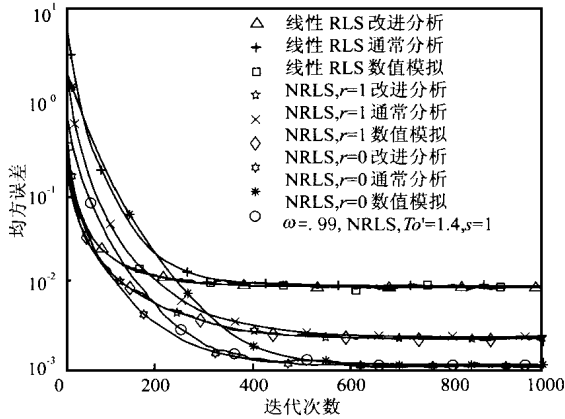


图 1 固定限幅函数的 NRLS 以及线性 RLS 的性能

$$T_0 = T'_0 \tilde{\sigma}_e^2(n)$$

其中残差的方差由下式递推计算

$$\tilde{\sigma}_e^2(n+1) = \beta \tilde{\sigma}_e^2(n) + (1-\beta)e^2(n)$$

使用平滑参数 $\beta = 0.99$, 固定 $T'_0 = 1.4$. 作为比较, 改进的分析方法和通常分析方法的结果连同模拟结果一同在图 1 中给出. 结果表明改进的分析方法较通常分析方法更接近模拟的结果. 同时使用限幅函数的 NRLS 较线性 RLS 改善了性能, 在稳态能达到更小的均方误差值.

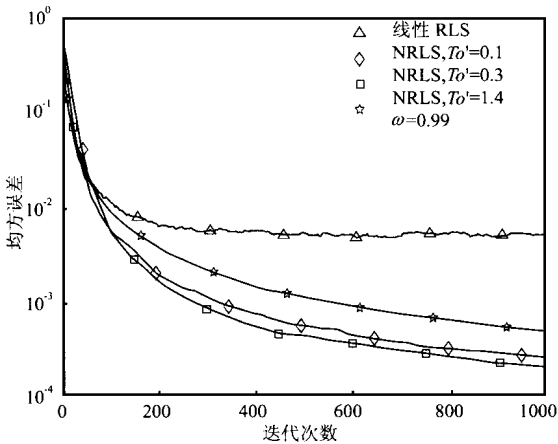


图 2 时变限幅函数的 NRLS 以及线性 RLS 的性能

第二个实验中, 我们给出了使用时变限幅函数的 NRLS 的性能(图 2). 广义限幅函数(4)的 r 和 s 由式(16)和式(17)实时算出, 限幅水平取值方法与上一实验相同, 固定 $T'_0 = 0.1, 0.3, 1.4$. 辅助变量 $\sigma_i^2(n)$ 的初值取 10. 模拟结果连同作为比较的线性 RLS 模拟结果一同在图 2 中给出. 结果表明, 时变限幅函数的 NRLS 性能较线性 RLS 和上一实验中的 NRLS 都有明显提高.

5 结论

本文提出了非线性 RLS 的改进分析方法. 模拟结果表明新的分析方法较通常分析方法更精确. 基于新的分析方法, 导出了一个时变广义限幅函数. 使用时变广义限幅函数的 NRLS 能更快收敛且达到更小的均方误差.

参考文献:

- [1] Haykin S. Adaptive Filter Theory [M]. 2nd Ed., Prentice Hall, New York, 1991.
- [2] Kim S R Efron. A adaptive robust impulse noise filtering [J]. IEEE Trans. Signal Processing, Aug. 1995, SP-43(8): 1855-1866.
- [3] Koike S Adaptive. Threshold nonlinear algorithm for adaptive filters with robustness against impulse noise [A]. Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing [C], Atlanta, USA, May 1996: 1644- 1647.
- [4] Masreliez C J. Approximate non gaussian filtering with linear state and observation relations [J]. IEEE Trans. Automatic Control, Feb. 1975, AC-20(2): 107- 120.
- [5] Aazhang B, Poor H V. And analysis of nonlinear direct sequence correlators [J]. IEEE Trans. Comm., July 1989, COM-37(7): 723- 731.
- [6] Leung S H, Weng J F. Adaptive filters with nonlinear RLS algorithm impulse noise [A]. Proc. IEEE Symp. Circuits, Systems [C], Orlando, Florida, USA, May 1999: III37- III40.
- [7] Bottomley G E, Alexander S T. A novel approach for stabilizing recursive least squares filters [J]. IEEE Trans. Signal Processing, Aug. 1991, SP-39(8): 1770- 1779.
- [8] Weng J F. Performance of DS/ CDMA two stage nonlinear detector in impulse noise [D]. PhD Dissertation, Dept. of Electron. Eng., City Univ. of Hong Kong, Hong Kong, Aug. 1997.

作者简介:

熊 鹰 男. 1963 年生于广东广州. 1991 年毕业于华南理工大学无线电与自动控制研究所, 获硕士学位. 现任华南理工大学电子与通信工程系讲师.

梁树雄 男. 博士. 香港城市大学电子工程学系副教授.