

一种基于均方偏差分析的通用最小均方算法

谢小平 史雄坤*

(湖南大学汽车车身先进设计制造国家重点实验室 长沙 410012)

摘 要: 无论是传统的定步长还是最近新提出的变步长最小均方(LMS)算法,在处理特定数学特征的信号时需要
对算法参数进行先验的估计才能达到较好的效果。但在实际信号处理过程中,算法参数的估计本就是一个很困难
的过程。该文分析了LMS算法的均方偏差及收敛特性,并提出一种以相对误差为变量的变步长LMS算法,能够实
现步长控制参数的自估计;可以自适应不同数学特征的信号,具体算例表明新算法有更快的收敛速度和较小的均
方误差。

关键词: 均方偏差分析; 自适应滤波; 相对误差; 通用性

中图分类号: TN911.7

文献标识码: A

文章编号: 1009-5896(2021)08-2249-09

DOI: 10.11999/JEIT200639

A General Least Mean Square Algorithm Based on Mean Square Deviation Analysis

XIE Xiaoping SHI Xiongkun

(State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body,
Hunan University, Changsha 410012, China)

Abstract: Whether it is the traditional fixed step size or the newly proposed Least Mean Square (LMS) algorithm, a priori estimation of the algorithm parameters is required to achieve better results when processing signals of specific mathematical features. However, in the actual signal processing process, the estimation of the algorithm parameters is a very difficult process. In this paper, the mean square deviation and convergence characteristics of LMS algorithm are analyzed, and a variable step size LMS algorithm with relative error as variable is proposed, which can realize self-estimation of step control parameters. It can adapt signals of different mathematical features. The example shows that the new algorithm has faster convergence speed and smaller mean square error.

Key words: Mean square deviation analysis; Adaptive filter; Relative error; Universality

1 引言

最小均方(Least Mean Square, LMS)算法属于随机梯度算法的一种,可以表示为一种横向线性滤波器,如图1所示。因其算法简单且易于实现而广泛应用于系统辨识、自适应均衡、自适应噪声消除、自适应谱线增强、自适应波束形成等领域^[1]。收敛速度和均方误差(Mean Square Error, MSE)是衡量LMS算法的两个重要指标。对于传统的定参数LMS算法,采用较大的步长可以加快算法的收敛速度,但收敛时误差值较大;采用较小的步长收敛速度较慢,收敛时稳态误差较小。

为了解决收敛速度和稳态误差之间的矛盾,人们提出了各种变步长LMS算法,这些算法大致可以

分为如下4个大类。第1类是通过输出响应和期望响应之差的大小来控制步长的大小,其步长的变化趋势与S型曲线Sigmoid^[2]函数一致;即误差较大和较小时步长变化的梯度较小,而中间过渡过程的步长变化梯度较大。这种S型的步长变化曲线使算法在误差较大时保持较快的收敛速度,误差较小时保持较小的稳态误差。该类算法典型的代表有:基于Sigmoid函数的最小均方(Least Mean Square based on Sigmoid function, GSVSLMS)算法^[3]、基于Sigmoid框架的非负最小均方(Non-Negative Least Mean Square based on Sigmoid framework, SNNLMS)算法^[4]等。这类算法的参数选取没有一个很好的数学准则,没有给出严格的数学证明,只能根据经验选择合适的大小。第2类是通过加入一些约束准则使误差均方代价函数最小^[5],然后求解出在这种约束下的步长更新公式。该类算法的代表

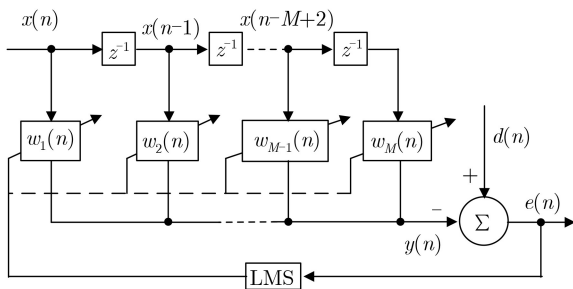


图1 横向自适应LMS滤波器

有：重新加权0吸引最小均方(Reweighted Zero-Attracting Least Mean Square, RZA-LMS)算法^[6]、分数最小均方(Fractional Least Mean Squares, FLMS)算法^[7]等。这一类算法的优点是其算法参数都是满足代价函数最小准则，并且在数学上是严格收敛的；缺点是在引入这些约束后使算法步长更新公式变得异常复杂，算法参数所受约束比较严格，其性能也将被削弱。第3类是一些步长调节组合类算法，这类算法组合了两种步长约束准则；根据设定的条件选择相应的约束准则进行步长的迭代更新。这类算法的代表有：组合正则化参数归一化最小均方(Combined Regularization Parameter for Normalized Least Mean Square, CRP-NLMS)算法^[8]、鲁棒可变加权系数扩散最小均方(Robust Variable Weighting Coefficients Diffusion Least Mean Square, RVWC-DLMS)算法^[9]、分数阶修正最小均方(Fractional Order Modified Least Mean Square, FOMLMS)算法^[10]等。该类算法的性能介于两种组合算法之间，性能比较均衡，但复杂度一般都比较较高。第4类是频域内的块LMS自适应算法，如改进的频域块最小均方(Modified Frequency-domain Block Least Mean Square, MFBLMS)算法^[11]、距离-多普勒-空间最小均方(Range Doppler Space Least Mean Square, RDS-LMS)算法^[12]等。这类算法通过快速傅里叶变换(Fast Fourier Transform, FFT)将时域信号转变到频域再做自适应；在频域的自适应可以减弱输入序列的线性相关性，起到一定加速收敛的作用；当自适应权值向量的维数较大时，频域自适应算法的性能才明显好于传统时域算法。此外还有递归最小二乘(Recursive Least Squares, RLS)算法和总体最小二乘(Total Least Squares, TLS)算法；其代表有：鲁棒递归最小二乘(Robust Recursive Least Squares, RLS)算法^[13]和对数总体最小二乘(Logarithmic Total Least Square, L-TLS)算法^[14]等。该类算法以最小二乘法为基础，具有较快的收敛速度和较小的误差；但该类算法涉及一些矩阵求逆的问题，计算复杂度较高。

前面提到的一些算法虽然有的算法都给出了参数的取值范围，但在对具体信号处理时还需要选定合适的值。当信号信噪比、信号功率等参数改变时就需要重新选定参数，因此并不具备通用性。

本文分析了基本LMS算法的收敛特性，确定了LMS算法收敛临界参数值与输入信号、期望信号之间的关系；并引入了采用相对误差作为步长调节因子的新算法。即使所处理信号特征发生改变，但相对误差值仍然可以反映算法收敛情况。通过分析算法收敛特性临界值结合相对误差快速调节步长参数，可以使新算法自适应不同特征信号处理。计算机仿真表明，新算法可以快速收敛并达到较小的均方误差，是一种高度自适应的通用LMS算法。

2 基本LMS算法收敛特性分析

2.1 系统模型

考虑如图2所示的LMS系统辨识模型。 $\mathbf{x}(n)$ 为参考信号向量， $\mathbf{x}(n) = [x(n) \ x(n-1) \ \cdots \ x(n-M+1)]^T$ 。 $y(n)$ 为系统的期望响应， $y(n) = \mathbf{x}^T(n)\mathbf{w}(n)$ 。 $\mathbf{w}(n)$ 为初始待辨识的权值向量， $\mathbf{w}(n) = [w_1(n) \ w_2(n) \ \cdots \ w_M(n)]^T$ 。 $d(n)$ 为系统输入的可观测信号， $d(n) = y(n) + z(n)$ ； $z(n)$ 为系统存在的加性噪声。 $\hat{y}(n)$ 为系统输出的估计值， $\hat{y}(n) = \mathbf{x}^T(n)\hat{\mathbf{w}}(n)$ ； $\hat{\mathbf{w}}(n)$ 为对待辨识系统权向量的估计值。 $e(n)$ 为待辨识信号与输出估计的观测误差， $e(n) = d(n) - \hat{y}(n)$ 。 $\mathbf{v}(n)$ 为权向量估计值与最优值的偏差， $\mathbf{v}(n) = \mathbf{w}(n) - \hat{\mathbf{w}}(n)$ 。

$\text{MSD}(n) = E\{\mathbf{v}^T(n)\mathbf{v}(n)\}$ 为偏差的均方值。由于 $\mathbf{w}(n)$ 为需要估计的量，无法直接求出 $\text{MSD}(n)$ 的值。为了便于计算，本文做出如下假设^[15-17]。

假设1 $x(n)$ 是独立同分布的，均值为0的稳态序列；存在均方 $E\{x^2(n)\} = \sigma_x^2$ ，4次均方 $E\{x^4(n)\}$ ，峰度系数 k

$$k = \frac{E\{x^4(n)\}}{E\{x^2(n)\}E\{x^2(n)\}} > 1 \quad (1)$$

假设2 可观测数据 $d(n)$ 可由一个线性回归模型来描述，该模型与维纳滤波器完全匹配，即

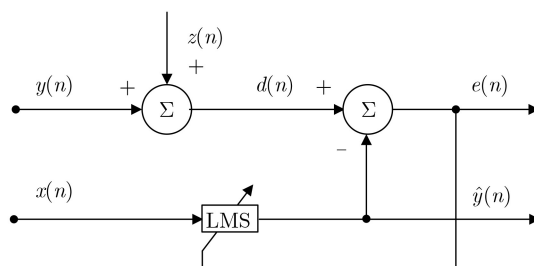


图2 LMS算法系统结构

$$d(n) = \mathbf{x}^T(n)\mathbf{w}(n) + z(n) \quad (2)$$

其中, $z(n)$ 是服从高斯分布的, 均值为0, 均方为 σ_z^2 的随机噪声信号。 $z(n)$ 与 $\mathbf{x}(n)$ 独立。

假设3 权值偏差 $\mathbf{v}(n)$ 独立于 $\mathbf{x}(n)$ 。

2.2 均方偏差分析

LMS算法权向量的迭代更新公式如式(3)所示

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \mu e(n)\mathbf{x}(n) \quad (3)$$

其中, μ 为步长参数, 则系统辨识误差可表示为

$$e(n) = \mathbf{x}^T(n)\mathbf{v}(n) + z(n) \quad (4)$$

那么权值误差迭代更新可以用式(5)表达

$$\mathbf{v}(n+1) = \mathbf{v}(n) - \mu \mathbf{x}(n)(\mathbf{x}^T(n)\mathbf{v}(n) + z(n)) \quad (5)$$

根据假设1~3, 式(5)两边同时乘于其转置并取期望得

$$\begin{aligned} E\{\mathbf{v}(n+1)\mathbf{v}^T(n+1)\} &= E\{\mathbf{v}(n)\mathbf{v}^T(n)\} + E\{\mu^2 z^2(n)\mathbf{x}(n)\mathbf{x}^T(n)\} \\ &\quad + E\{\mu^2 \mathbf{x}(n)\mathbf{x}^T(n)\mathbf{v}(n)\mathbf{v}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\} \\ &\quad - E\{2\mu \mathbf{x}(n)\mathbf{x}^T(n)\}E\{\mathbf{v}(n)\mathbf{v}^T(n)\} \end{aligned} \quad (6)$$

根据假设1有

$$E\{\mathbf{x}(n)\mathbf{x}^T(n)\} = \sigma_x^2 \mathbf{I} \quad (7)$$

根据假设3有

$$\begin{aligned} E\{\mu^2 \mathbf{x}(n)\mathbf{x}^T(n)\mathbf{v}(n)\mathbf{v}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\} &= \frac{M+k-3}{M} \text{tr}[\mathbf{v}(n)\mathbf{v}^T(n)]\sigma_x^4 \mathbf{I} \\ &\quad + 2\sigma_x^4 \mathbf{v}(n)\mathbf{v}^T(n) + \Delta \end{aligned} \quad (8)$$

$\text{tr}[\cdot]$ 表示对矩阵取迹, $[\cdot]^T$ 表示转置, $E\{\cdot\}$ 表示取数学期望; M 为滤波器的阶数; $\mathbf{I}$ 表示 $M \times M$ 的单位矩阵。式(8)的详细证明见附录。

将式(7)、式(8)代入式(6)可得

$$\begin{aligned} E\{\mathbf{v}(n+1)\mathbf{v}^T(n+1)\} &= E\{\mathbf{v}(n)\mathbf{v}^T(n)\} - 2\mu\sigma_x^2 E\{\mathbf{v}(n)\mathbf{v}^T(n)\} \\ &\quad + \mu^2 \left(\frac{M+k-3}{M} \text{tr}[\mathbf{v}(n)\mathbf{v}^T(n)]\sigma_x^4 \mathbf{I} \right. \\ &\quad \left. + 2\sigma_x^4 \mathbf{v}(n)\mathbf{v}^T(n) + \Delta \right) + \mu^2 \sigma_x^2 \sigma_z^2 \mathbf{I} \end{aligned} \quad (9)$$

因为

$$\text{MSD}(n) = E\{\mathbf{v}^T(n)\mathbf{v}(n)\} = E\{\text{tr}[\mathbf{v}(n)\mathbf{v}^T(n)]\} \quad (10)$$

所以, 根据式(10)、式(9)两边进行迹运算可得

$$\begin{aligned} \text{MSD}(n+1) &= (1 - 2\mu\sigma_x^2 + \mu^2\sigma_x^4(M+k-1)) \\ &\quad \cdot \text{MSD}(n) + M\mu^2\sigma_x^2\sigma_z^2 \end{aligned} \quad (11)$$

根据式(11)得LMS算法 μ 的边界条件

$$|1 - 2\mu\sigma_x^2 + \mu^2\sigma_x^4(M+k-1)| < 1 \quad (12)$$

对于2次函数

$$f(\mu) = \mu^2\sigma_x^4(M+k-1) - 2\mu\sigma_x^2 + 1 \quad (13)$$

当 $\mu = \frac{1}{\sigma_x^2(M+k-1)}$ 时, $f(\mu)$ 有最小值

$$f_{\min}(\mu) = \frac{M+k-2}{M+k-1} > 0 \quad (14)$$

显然有

$$0 < \mu^2\sigma_x^4(M+k-1) - 2\mu\sigma_x^2 + 1 < 1 \quad (15)$$

解之得

$$\mu < \frac{2}{\sigma_x^2(M+k-1)} \quad (16)$$

由式(16)可知, LMS步长 μ 主要与权值向量的维数 M 、输入信号峰度系数 k 、输入信号均方 σ_x^2 有关。在满足式(16)的条件下, 对 n 在式(11)两边取极限可得

$$\begin{aligned} \text{MSD}(\infty) &= (1 - 2\mu\sigma_x^2 + \mu^2\sigma_x^4(M+k-1)) \\ &\quad \cdot \text{MSD}(\infty) + M\mu^2\sigma_x^2\sigma_z^2 \end{aligned} \quad (17)$$

那么可得

$$\text{MSD}(\infty) = \frac{\mu M \sigma_z^2}{2 - \mu \sigma_x^2(M+k-1)} \quad (18)$$

理想状态下, μ 取值很小, 有 $\mu\sigma_x^2(M+k-1) \ll 2$ 。那么可得

$$\text{MSD}_{\min}(\infty) \approx \frac{\mu M \sigma_z^2}{2} \quad (19)$$

在实际自适应信号处理中总有 $\sigma_z^2 < \sigma_x^2$, $k > 1$, 则有

$$\mu M \sigma_z^2 < \mu \sigma_x^2(M+k-1) \quad (20)$$

令 $h = \mu \sigma_x^2(M+k-1)$ 则

$$\text{MSD}(\infty) = \frac{\mu M \sigma_z^2}{2 - \mu \sigma_x^2(M+k-1)} < \frac{h}{2-h} \quad (21)$$

考虑式(22)函数的图像, 如图3所示。

$$\text{MSD}(h) = \frac{h}{2-h} \quad (22)$$

由图3可知 h 与MSD存在反比关系。即小步长

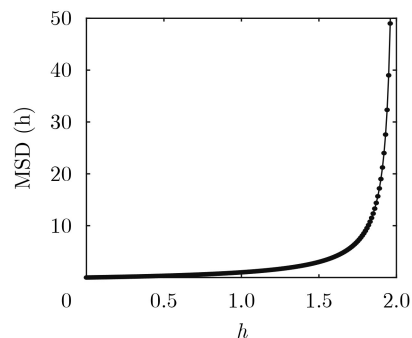


图3 h 与MSD变化关系

最终偏差较小,大步长偏差较大;这与实际信号处理过程相符合。

考虑式(16)以及信号在局部时段有峰值的变化,可以适当放宽式(16)的约束条件可得步长 μ 最大取值边界条件

$$\mu_{\max} = \frac{1}{\sigma_x^2(M+k-1)} \quad (23)$$

显然有

$$\mu_{\max} = \frac{1}{\sigma_x^2(M+k-1)} < \frac{2}{\sigma_x^2(M+k-1)} \quad (24)$$

显然新的步长边界值 $\mu_{\max}$ 是满足LMS算法严格收敛条件的。

3 新的变步长RELMS算法及分析

运用GSVSLMS这类算法的思想,结合LMS算法在实际信号处理过程的特性。即在前迭代时LMS观测误差较大,可以采用较大步长加速收敛;在算法接近收敛时,采用较小的步长使观测误差减小。

为了实现这一目标,本文关心的主要有3点。第1点:确定算法在迭代过程中的收敛程度,用 $q(n)$ 表示。第2点:判断算法接近收敛时 $q(n)$ 的大小 $q_{\min}$;第3点:拟合一个与理想步长调节过程相匹配的步长控制曲线。

对于第1点,本文可以采取短时功率形式的相对误差来度量算法的收敛程度,即

$$q(n) = \frac{\mathbf{e}^T(n)\mathbf{e}(n)}{\mathbf{d}^T(n)\mathbf{d}(n) + \mathbf{e}^T(n)\mathbf{e}(n)} < 1 \quad (25)$$

其中, $\mathbf{e}(n)$ 为观测误差 $e(n)$ 构成的 M 维向量(加粗的 $\mathbf{e}(n)$ 为向量), $\mathbf{e}(n) = [e(n)e(n-1) \cdots e(n-M+1)]^T$; $\mathbf{d}(n)$ 为输入观测信号 $d(n)$ 所构成的 M 维向量, $\mathbf{d}(n) = [d(n)d(n-1) \cdots d(n-M+1)]^T$ 。标度因子 $q(n)$ 采用相对形式,不受信号功率改变的影响。

对于第2点,本文需要确定算法收敛时 $q(n)$ 的最小值 $q_{\min}$ 。根据式(4)有

$$\begin{aligned} E\{e^2(n)\} &= E\{\mathbf{x}^T(n)\mathbf{v}(n)\mathbf{v}^T(n)\mathbf{x}(n)\} + E\{z^2(n)\} \\ &= \text{MSD}(n) + \sigma_z^2 \end{aligned} \quad (26)$$

考虑理想情况下的 $\text{MSD}(\infty)$,将式(19)、式(26)代入式(25)得

$$q_{\min} = \frac{\sigma_z^2 + \frac{\mu M \sigma_z^2}{2}}{\sigma_d^2 + \sigma_z^2 + \frac{\mu M \sigma_z^2}{2}} \quad (27)$$

式中 $\sigma_d^2 = E\{d^2(n)\}$,理想情况下 μ 值足够小,有

$$\frac{\mu M \sigma_z^2}{2} \ll \sigma_z^2 \quad (28)$$

式(27)可化简为

$$q_{\min} \approx \frac{\sigma_z^2}{\sigma_d^2 + \sigma_z^2} \quad (29)$$

对于第3点,本文在拟合曲线前需要确定必要的参数。在前面的分析中本文确定了最大的步长值 $\mu_{\max}$ 值,现在需要确定一个合适的最小步长 $\mu_{\min}$ 。

考虑基本LMS算法权值更新公式(3),当算法收敛时有

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \mu_{\min} e(n)_{\min} \mathbf{x}(n) \quad (30)$$

此时 $\hat{\mathbf{w}}(n)$ 接近理想权值 $\mathbf{w}(n)$,可得

$$\begin{aligned} \mathbf{v}_{\min}(n) &= \hat{\mathbf{w}}(n+1) - \hat{\mathbf{w}}(n) \\ &= \mu_{\min} e(n)_{\min} \mathbf{x}(n) \end{aligned} \quad (31)$$

$\mathbf{v}_{\min}(n)$ 为算法稳定时权值的扰动,其所带来的误差扰动为

$$\begin{aligned} \Delta e(n)_{\min} &= \mathbf{x}^T(n) \mathbf{v}_{\min}(n) \\ &= \mu_{\min} e(n)_{\min} \mathbf{x}^T(n) \mathbf{x}(n) \end{aligned} \quad (32)$$

若设定误差扰动的精度为 p ,则

$$p = \frac{\Delta e(n)_{\min}}{e(n)_{\min}} = \mu_{\min} \mathbf{x}^T(n) \mathbf{x}(n) \approx \mu_{\min} M \sigma_x^2 \quad (33)$$

p 可根据所需精度设置为0.0005~0.05的值。那么可得

$$\mu_{\min} = \frac{p}{M \sigma_x^2} \quad (34)$$

至此,LMS算法步长就有了 $\mu_{\min} \sim \mu_{\max}$ 的动态范围。

现在考虑如式(35)所示的函数

$$y = \frac{Ax}{1 + (A-1)x}, \quad x \in [0, 1] \quad (35)$$

如图4所示,当 $A=1$ 时,式(32)所示的函数退化为 $y=x$;当 $A=4$, $A=10$ 时, x 从0变化到1的过程中曲线向上凸起;曲线的梯度逐渐变小,变小的速度先快后慢。 y 在接近1的地方缓慢减小,接近0

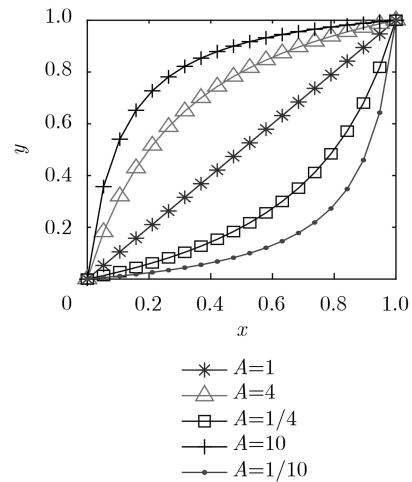


图4 y 随不同 A 的变化情况

的地方快速减小； A 越大这种趋势越明显，即越向上凸起。当 $A = 1/4$, $A = 1/10$ 时， x 从0变化到1的过程中曲线向下凹；曲线的梯度逐渐增大，增大的速度先慢后快； y 在接近1的地方快速减小，接近0的地方缓慢减小； A 越小这种趋势越明显，即越向下凹。

采用上述构造的函数，以 q 为自变量，建立 μ 关于 q 的函数。同时为了契合步长调节原则，应当在算法初始迭代时，即 q 的值 $q_{\max}$ 接近于1($q_{\max} < 1$ ，为了计算方便我们取 $q_{\max} = 1$)时采用较大的 A 值；当算法接近收敛时， q 趋于 $q_{\min}$ ($q_{\min} > 0$)时采用较小的 A 值。这样可以使算法前期维持较大的步长以加速收敛，在算法接近收敛时维持较小的步长以减少误差。

考虑到上述函数的值域为 $[0, 1]$ ，而需要拟合的自变量 q 的定义域为 $[q_{\min}, 1]$ ， μ 的值域为 $[\mu_{\min}, \mu_{\max}]$ 。拟合出如式(36)所示的函数

$$\mu(n) = \frac{A(n)q(n)\mu_{\max}}{1 + q(n)(A(n) - 1)} \quad (36)$$

该函数过点 $(q_{\min}, \mu_{\min})$ 和 $(1, \mu_{\max})$ ，将点 $(q_{\min}, \mu_{\min})$ 代入式(36)可以求得 $(q_{\min}, A_{\min})$

$$A_{\min} = \frac{\mu_{\min}(1 - q_{\min})}{q_{\min}(\mu_{\max} - \mu_{\min})} \quad (37)$$

将点 $(1, \mu_{\max})$ 代入式(36)时，求得 $A_{\max}$ 可以为任何值。为实现步长的动态调节，需要建立 A 关于 q 的函数。并且这个函数应有如下特点：当 q 趋近1时，有较大的 $A_{\max}$ 值；当 q 趋近于 $q_{\min}$ 时， $A_{\min}$ 值较小。而且 $A_{\min}$ 与 $A_{\max}$ 之间的曲线是向上凸出的，即当 q 由 $q_{\min}$ 增加至1的过程中，梯度是逐渐减小的；这有助于步长在前期的缓慢减小，加快收敛速度。本文发现在 $\sin(x)$ 的单调区间 $[0, 0.5\pi]$ 上的曲线特性刚好能满足这种要求，因此我们以此为基础拟合出函数为

$$A(n) = b \sin(0.5\pi q(n)) \quad (38)$$

且该曲线过点 $(1, A_{\max})$ 和点 $(q_{\min}, A_{\min})$ ，这里 $A_{\max}$ 与 $q_{\min}$ 的值有关。将点 $(q_{\min}, A_{\min})$ 代入式(38)可以求得

$$b = \frac{A_{\min}}{\sin(0.5\pi q_{\min})} \quad (39)$$

将式(39)代入式(38)得

$$A(n) = \frac{A_{\min} \sin(0.5\pi q(n))}{\sin(0.5\pi q_{\min})} \quad (40)$$

为了直观地显示 A 随不同 $q_{\min}$ 的变化情况，本文取 $\mu_{\min} = 0.001$, $\mu_{\max} = 1$ 作曲线 $A(n) - q(n)$ ，如图5所示。

如图5所示，当 $q_{\min}$ 一定时，随着 q 的增加 A 值逐渐增大， $q = 1$ 时 A 取得最大值 $A_{\max}$ ；且 q 在增大的过程中函数梯度逐渐减小，当 $q = 1$ 时梯度最小。 A 曲线整体走势主要由 $q_{\min}$ 决定， $q_{\min}$ 越小 $A_{\max}$ 值越大 A 曲线总体的梯度变化得越多，相应的 A 曲线越陡； $q_{\min}$ 越大 $A_{\max}$ 值越小， A 曲线总体的梯度变化得越少，且 A 曲线越平。这可以做如下解释，当 $q_{\min}$ 越小时，意味着算法能达到的稳态相对误差越小；那么也意味着达到收敛需要的迭代步数越多，时间越长。那么在前期的可以在较长一段时间以较大的步长进行迭代(需要较大的 A)，在接近收敛时再快速以较小的步长进行迭代。那么可以在迭代初始用较大的 $A_{\max}$ 得到较大的步长， A 再以加速的方式衰减至收敛时的 $A_{\min}$ 值，此时步长为 $\mu_{\min}$ 。这与本文的期望一致。而 $q_{\min}$ 越大，算法稳态相对误差越大；达到收敛所需要的迭代的步数越少，时间越短。那么步长需要由 $\mu_{\max}$ 快速变化到 $\mu_{\min}$ ，此时就不宜采用较大的 $A_{\max}$ ，而用较小的 $A_{\max}$ 。 A 曲线相对平的趋势使 $A_{\max}$ 与 $A_{\min}$ 相差不大(且都较小)，可以保证步长快速地由 $\mu_{\max}$ 变化到 $\mu_{\min}$ ；这也与本文的期望一致。

综上所述，步长控制函数及参数已全部给出，并整理为

$$\mu(n) = \frac{A(n)q(n)\mu_{\max}}{1 + q(n)(A(n) - 1)} \quad (41)$$

$$A(n) = \frac{A_{\min} \sin(0.5\pi q(n))}{\sin(0.5\pi q_{\min})} \quad (42)$$

$$A_{\min} = \frac{\mu_{\min}(1 - q_{\min})}{q_{\min}(\mu_{\max} - \mu_{\min})} \quad (43)$$

$$q(n) = \frac{\mathbf{e}^T(n)\mathbf{e}(n)}{\mathbf{d}^T(n)\mathbf{d}(n) + \mathbf{e}^T(n)\mathbf{e}(n)} \quad (44)$$

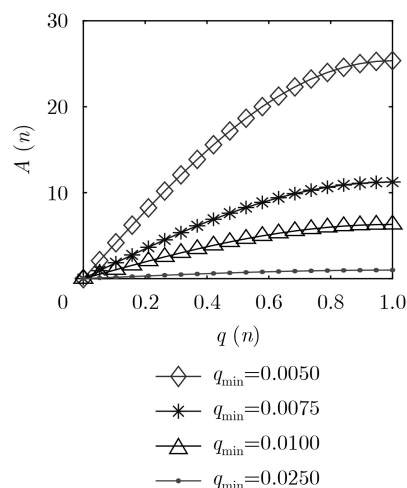


图5 $A(n)$ 随不同 $q_{\min}$ 的变化情况

$$q_{\min} = \frac{\sigma_z^2}{\sigma_d^2 + \sigma_z^2} \quad (45)$$

$$\mu_{\min} = \frac{p}{M\sigma_x^2}, \mu_{\max} = \frac{1}{\sigma_x^2(M+k-1)} \quad (46)$$

$\mu(n)$, $q(n)$, $A(n)$ 这三项是参与算法每一步迭代的;而后面的 $\mu_{\min}$, $\mu_{\max}$, $A_{\min}$, $q_{\min}$ 是根据已知的 M , σ_x^2 , σ_d^2 , σ_z^2 在算法初始时求得的定值, p 为设定的精度(一般设定为0.0005~0.0500, 值越小稳态误差越小)。

为验证拟合的步长控制曲线符合LMS算法期望的步长调整原则, 本文选取不同的 $q_{\min}$ 值以代表不同数学特征的信号。为了直观地显示步长的控制效果, 我们取 $\mu_{\min} = 0.001$, $\mu_{\max} = 1$; 作步长曲线 $\mu(n)$ 随 $q_{\min}$ 的变化曲线 $\mu(n) - q(n)$, 如图6所示。

如图6所示, 当 $q_{\min}$ 取不同值时这些曲线的变化趋势与本文前面的描述是一致的。当 $q_{\min}$ 较小, 算法迭代前期 $q(n)$ 值较大且接近于1时, 步长值 $\mu(n)$ 是缓慢减小的; 这将使算法在前期都能以相对较大的步长值进行迭代, 有助于加速算法的收敛速度; 当 $q(n)$ 接近于 $q_{\min}$ 时, 步长值迅速下降至最小步长 $\mu_{\min}$, 这有助于算法在接近收敛时降低稳态误差。当 $q_{\min}$ 较大时, 步长由 $\mu_{\max}$ 迅速下降至最小步长 $\mu_{\min}$ 。新算法 $q(n)$ 在 $q_{\min} \sim 1$ 的范围内变化时, 步长 $\mu(n)$ 有 $\mu_{\min} \sim \mu_{\max}$ 的动态范围, 其变化趋势与我们期望的步长调节方式一致。 $q_{\min}$ 值的不同代表了不同数学特征的信号, 结合 $q(n)$ 所表征的是相对误差; 因此, 新算法可以自适应不同数学特征信号的处理。

4 计算机仿真验证

本节通过计算机仿真来验证新算法的性能。由

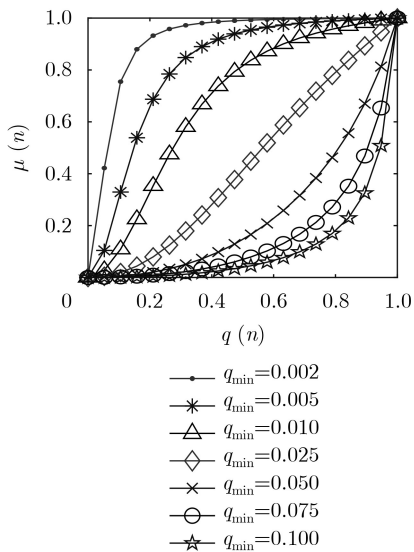


图6 $\mu(n)$ 随不同 $q_{\min}$ 的变化情况

于本文的LMS偏差分析基于假设1和假设3, 所以首先通过仿真验证其假设可靠性。结合假设2和假设3本文得出结果为

$$\begin{aligned} & E\{\mathbf{v}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{v}(n)\} \\ &= E\{\mathbf{v}^T(n)E\{\mathbf{x}(n)\mathbf{x}^T(n)\}\mathbf{v}(n)\} \\ &= \sigma_x^2 \text{MSD}(n) \end{aligned} \quad (47)$$

$$\begin{aligned} & E\{\mathbf{x}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{v}(n)\mathbf{v}^T(n)\mathbf{x}(n)\} \\ &= \sigma_x^4(M+k-1)\text{MSD}(n) \end{aligned} \quad (48)$$

本文用LMS算法辨识一个未知系统(辨识模型如图7所示)来验证假设2和假设3的合理性。取

$$f_1(n) = E\{\mathbf{v}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{v}(n)\} \quad (49)$$

$$g_1(n) = \sigma_x^2 \text{MSD}(n) \quad (50)$$

$$f_2(n) = E\{\mathbf{x}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{v}(n)\mathbf{v}^T(n)\mathbf{x}(n)\} \quad (51)$$

$$g_2(n) = \sigma_x^4(M+k-1)\text{MSD}(n) \quad (52)$$

建立如下的仿真环境:

自适应滤波器的维数 $M = 10$, 由MATLAB随机函数产生待辨识系统的抽头权值向量 $\mathbf{w}_0$; 滤波器输入 $x(n)$ 采用1阶AR模型, 如式(53)所示

$$x(n) = 0.09x(n-1) + z_1(n) \quad (53)$$

其中, $z_1(n)$ 是均值为0, 均方为0.01的高斯白噪声信号。参考输入信号如式(54)所示

$$d(n) = \mathbf{w}_0^T \mathbf{x}(n) + z(n) \quad (54)$$

其中, $z(n)$ 是均值为0, 其均方为0.001的高斯白噪声。取500次独立仿真结果做平均, 结果如图8、图9所示

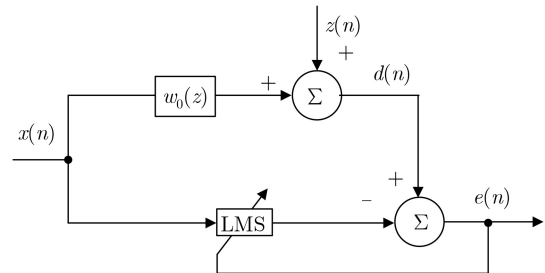


图7 系统辨识模型

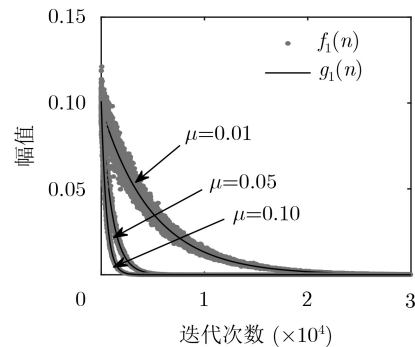


图8 式(47)的验证

通过图8、图9中 $f(n)$ 和 $g(n)$ 图像的对比结果来看, 根据假设2和假设3得出的曲线对实际值有较好的拟合效果, 因此可以认为假设2和假设3是合理的。当 μ 取较大值时, 假设值与实际值偏差变小, μ 越小与实际值偏差越大; 当LMS算法接近收敛时假设值与实际值偏差减小, 并最终趋于0。

采用式(53)的1阶AR模型作为输入进行系统辨识来验证新算法的性能。

仿真实验1: $z_1(n)$ 是均值为0, 均方为0.1的高斯白噪声; $z(n)$ 是均值为0, 均方为0.001的高斯白噪声; 本文算法 $p = 0.002$, LMS算法定步长 $\mu = 0.05$, GSVSLMS算法 $\alpha = 0.5$, $\beta = -200$ 。仿真结果见图10。

仿真实验2: $z_1(n)$ 是均值为0, 均方为0.1的高斯白噪声; $z(n)$ 是均值为0, 均方为0.0001的高斯白噪声; 本文算法 $p = 0.002$, LMS算法定步长 $\mu = 0.05$, GSVSLMS算法 $\alpha = 0.5$, $\beta = -200$ 。仿真结果见图11。

语音信号的仿真, 实验3: 实验过程不变, 将 $x(n)$ 换为一段语音信号, 如图12所示; $z(n)$ 均值为0, 均方为0.0001的高斯白噪声, 如图13所示; 取 $\mathbf{w}_0 = [1 - 0.8 \ 0.4 - 0.2 \ 0.1 \ 0.05 \ 0.1 - 0.02]^T$, 混合后的含噪声语音 $d(n)$ 如图14所示。本文算法 $p = 0.002$, LMS算法定步长 $\mu = 0.05$, GSVSLMS算法 $\alpha = 0.5$, $\beta = -50$ 。其仿真结果如图15、图16、图17所示。

通过仿真实验和仿真实验2, 两种不同数学特征信号条件下的系统辨识的结果可以得出: 新算法

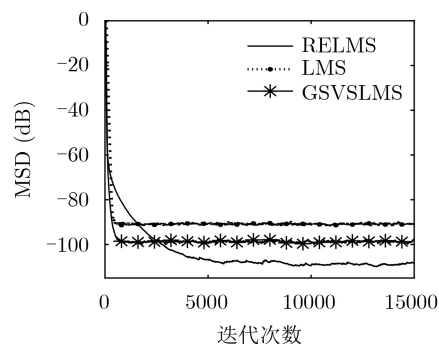


图 11 仿真实验2的结果

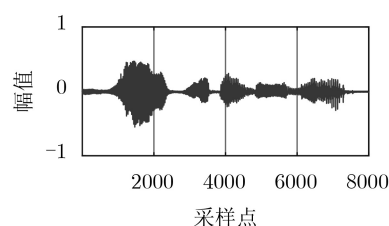
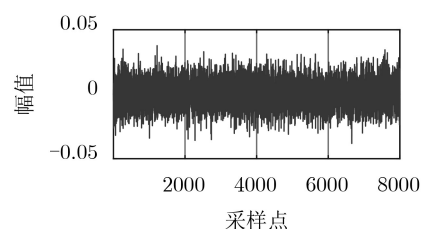
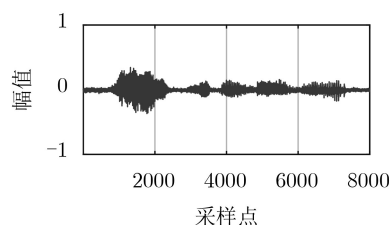
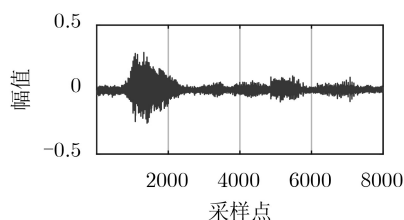
图 12 一段原始语音信号 $x(n)$ 图 13 加性噪声 $z(n)$ 图 14 含噪声语音信号 $d(n)$ 

图 15 LMS算法误差

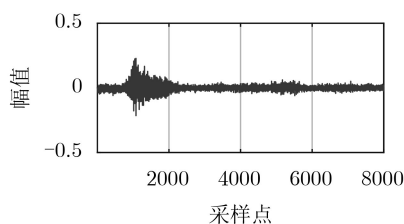


图 16 GSVSLMS算法误差

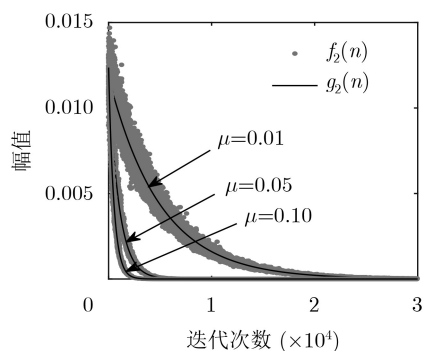


图 9 式(48)的验证

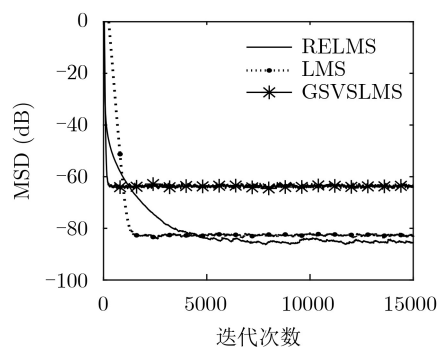


图 10 仿真实验1的结果

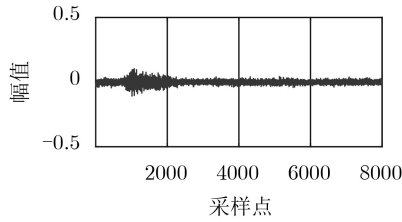


图 17 RELMS算法误差

在前期的收敛速度比定步长LMS算法和GSVSLMS类算法更快；并且当算法收敛时，新算法能达到定步长LMS算法几乎一样甚至更小的权值偏差。通过对语音信号的仿真实验3，可以看出：当信号发生功率上的突变时，本文算法的误差明显低于前面两种算法。本文采用相对误差形式的步长调节自变量，可以根据信号特征自动调整算法的步长控制曲线，特别适用于信号功率有突变的自适应信号处理。新算法作为一种通用的LMS算法，已经取得了良好的算法性能。

5 结论

本文通过合理的假设分析了基本LMS的收敛特性，得出了LMS算法可以接受的最大步长条件；并通过计算机仿真验证了这些假设的合理性。采用误差序列短时功率与期望序列信号短时功率的比值来表征算法的收敛程度，并据此提出了新的变步长RELMS算法；新算法步长调节方式契合LMS算法步长理想调整原则，相对误差形式的步长控制自变量可以自适应不同特征的信号，更适用于信号功率有突变的自适应信号处理。新算法避免了传统自适应算法参数选择的过程，是一种性能较好的通用性LMS算法。

6 附录

X 是均值为0，均方为 σ_x^2 ，峰度系数为 k 的独立同分布的随机变量。 $\mathbf{x}$ 是由 X 构成的 M 维列向量，即 $\mathbf{x} = [X_1 \ X_2 \ \cdots \ X_M]^T$ ， $\mathbf{x}\mathbf{x}^T$ 为其自相关矩阵。 $\mathbf{B}$ 为 $M \times M$ 的对称矩阵。则有

$$\begin{aligned} E\{\mathbf{x}\mathbf{x}^T \mathbf{B} \mathbf{x}\mathbf{x}^T\} &= \frac{M+k-3}{M} \text{tr}[\mathbf{B}] \sigma_x^4 \mathbf{I} \\ &\quad + 2\sigma_x^4 \mathbf{B} + \Delta \end{aligned} \quad (55)$$

证明 令 $\mathbf{D} = E\{\mathbf{x}\mathbf{x}^T \mathbf{B} \mathbf{x}\mathbf{x}^T\}$ ， $d_{i,j}$ 代表 $\mathbf{D}$ 矩阵的 i 行 j 列的元素；令 $\mathbf{A} = \mathbf{x}\mathbf{x}^T$ ， $a_{i,j}$ 为 $\mathbf{A}$ 矩阵 i 行 j 列的元素； $b_{i,j}$ 为 $\mathbf{B}$ 矩阵 i 行 j 列的元素， $\mathbf{Q} = \mathbf{A}\mathbf{B}$ ； $q_{i,j}$ 为 $\mathbf{Q}$ 矩阵 i 行 j 列的元素。则有

$$q_{i,j} = \sum_{m=1}^M a_{i,m} b_{m,j} \quad (56)$$

又 $\mathbf{D} = \mathbf{Q}\mathbf{A}$ ，根据式(56)则有

$$d_{i,j} = E\left\{\sum_{m=1}^M \sum_{n=1}^M a_{i,m} b_{m,n} a_{n,j}\right\} \quad (57)$$

而 $a_{i,j} = X_i X_j$ ，则有

$$d_{i,j} = E\left\{X_i X_j \sum_{m=1}^M \sum_{n=1}^M X_m X_n b_{m,n}\right\} \quad (58)$$

(1) 当 $i = j$ 时，有

$$d_{i,i} = E\left\{X_i^2 \sum_{m=1}^M \sum_{n=1}^M X_m X_n b_{m,n}\right\} \quad (59)$$

(a) 当 $m \neq i$ 或 $n \neq i$ ，且 $m = n$ 时，有

$$E\{X_i^2 X_m X_n b_{m,n}\} = 0 \quad (60)$$

(b) 当 $m \neq i$ 或 $n \neq i$ ，且 $m = n$ 时，有

$$\begin{aligned} &E\left\{X_i^2 \sum_{m=1, m \neq i}^M X_m^2 b_{m,m}\right\} \\ &= E\{X_i^2 X_i^2\} E\left\{\sum_{m=1, m \neq i}^M b_{m,m}\right\} \\ &= \sigma_x^4 (\text{tr}[\mathbf{B}] - b_{i,i}) \end{aligned} \quad (61)$$

(c) 当 $m = n = i$ 时，有

$$E\left\{X_i^2 \sum_{m=i}^M X_m^2 b_{i,i}\right\} = E\{X_i^4\} b_{i,i} \quad (62)$$

综上，当 $i = j$ 时，有

$$\begin{aligned} d_{i,i} &= \sigma_x^4 \text{tr}[\mathbf{B}] + \sigma_x^4 \left(\frac{E\{X_i^4\}}{E\{X_i^2\}E\{X_i^2\}} - 1\right) b_{i,i} \\ &= \sigma_x^4 \text{tr}[\mathbf{B}] + \sigma_x^4 (k-1) b_{i,i} \end{aligned} \quad (63)$$

(2) 当 $i \neq j$ ， $m \neq i$ ， j 或 $n \neq i$ ， j 时，有

$$E\{X_i X_j X_m X_n b_{m,n}\} = 0 \quad (64)$$

当 $i \neq j$ ， $m = i$ ， $n = j$ 或 $m = j$ ， $n = i$ 时，有

$$d_{i,j} = E\{X_i^2 X_j^2\} (b_{i,j} + b_{j,i}) = 2\sigma_x^4 b_{i,j} \quad (65)$$

以 $\Lambda(\mathbf{B})$ 表示只保留矩阵 $\mathbf{B}$ 主对角线元素，其他元素为0的对角矩阵，综合1, 2有

$$\mathbf{D} = 2\sigma_x^4 \mathbf{B} + \sigma_x^4 \text{tr}[\mathbf{B}] \mathbf{I} + (k-3)\sigma_x^4 \Lambda(\mathbf{B}) \quad (66)$$

对于 $\Lambda(\mathbf{B})$ ，本文做如式(67)的近似

$$\Lambda(\mathbf{B}) \approx \frac{\text{tr}[\mathbf{B}]}{M} \mathbf{I} + \Delta \quad (67)$$

将式(67)代入式(66)有

$$\mathbf{D} = \frac{M+k-3}{M} \text{tr}[\mathbf{B}] \sigma_x^4 \mathbf{I} + 2\sigma_x^4 \mathbf{B} + \Delta \quad (68)$$

证毕

特别地，当对矩阵 $\mathbf{D}$ 进行迹运算时，误差项 Δ 为0。

参考文献

- [1] 谢胜利, 何昭水, 高鹰. 信号处理的自适应理论[M]. 北京: 科学

- 出版社, 2006: 2–7.
- [2] HUANG Fuyi, ZHANG Jiashu, and ZHANG Sheng. A family of robust adaptive filtering algorithms based on sigmoid cost[J]. *Signal Processing*, 2018, 149: 179–192. doi: [10.1016/j.sigpro.2018.03.013](https://doi.org/10.1016/j.sigpro.2018.03.013).
- [3] 高鹰, 谢胜利. 一种变步长LMS自适应滤波算法及分析[J]. 电子学报, 2001, 29(8): 1094–1097. doi: [10.3321/j.issn:0372-2112.2001.08.023](https://doi.org/10.3321/j.issn:0372-2112.2001.08.023).
GAO Ying and XIE Shengli. A variable step size LMS adaptive filtering algorithm and its analysis[J]. *Acta Electronica Sinica*, 2001, 29(8): 1094–1097. doi: [10.3321/j.issn:0372-2112.2001.08.023](https://doi.org/10.3321/j.issn:0372-2112.2001.08.023).
- [4] 樊宽刚, 邱海云. 基于Sigmoid框架的非负最小均方算法[J]. 电子与信息学报, 2021, 43(2): 349–355. doi: [10.11999/JEIT200018](https://doi.org/10.11999/JEIT200018).
FAN Kuan'gang and QIU Haiyun. Robust nonnegative least mean square algorithm based on sigmoid framework[J]. *Journal of Electronics & Information Technology*, 2021, 43(2): 349–355. doi: [10.11999/JEIT200018](https://doi.org/10.11999/JEIT200018).
- [5] PEREZ F L, DE SOUZA F D C, and SEARA R. An improved mean-square weight deviation-proportionate gain algorithm based on error autocorrelation[J]. *Signal Processing*, 2014, 94: 503–513. doi: [10.1016/j.sigpro.2013.06.030](https://doi.org/10.1016/j.sigpro.2013.06.030).
- [6] JIN Danqi, CHEN Jie, RICHARD C, *et al.* Model-driven online parameter adjustment for zero-attracting LMS[J]. *Signal Processing*, 2018, 152: 373–383. doi: [10.1016/j.sigpro.2018.06.020](https://doi.org/10.1016/j.sigpro.2018.06.020).
- [7] BERSHAD N J, WEN Fuxi, and SO H C. Comments on “fractional LMS algorithm” [J]. *Signal Processing*, 2017, 133: 219–226. doi: [10.1016/j.sigpro.2016.11.009](https://doi.org/10.1016/j.sigpro.2016.11.009).
- [8] SHI Long, ZHAO Haiquan, WANG Wenyuan, *et al.* Combined regularization parameter for normalized LMS algorithm and its performance analysis[J]. *Signal Processing*, 2019, 162: 75–82. doi: [10.1016/j.sigpro.2019.04.014](https://doi.org/10.1016/j.sigpro.2019.04.014).
- [9] AHN D C, LEE J W, SHIN S J, *et al.* A new robust variable weighting coefficients diffusion LMS algorithm[J]. *Signal Processing*, 2017, 131: 300–306. doi: [10.1016/j.sigpro.2016.08.023](https://doi.org/10.1016/j.sigpro.2016.08.023).
- [10] CHENG Songsong, WEI Yiheng, CHEN Yuquan, *et al.* A universal modified LMS algorithm with iteration order hybrid switching[J]. *ISA Transactions*, 2017, 67: 67–75. doi: [10.1016/j.isatra.2016.11.019](https://doi.org/10.1016/j.isatra.2016.11.019).
- [11] YANG Feiran and YANG Jun. Mean-square performance of the modified frequency-domain block LMS algorithm[J]. *Signal Processing*, 2019, 163: 18–25. doi: [10.1016/j.sigpro.2019.04.030](https://doi.org/10.1016/j.sigpro.2019.04.030).
- [12] 杨鹏程, 吕晓德, 柴致海, 等. 基于频域分块RDS-LMS算法的机载外辐射源雷达杂波对消[J]. 电子与信息学报, 2017, 39(10): 2302–2310. doi: [10.11999/JEIT170190](https://doi.org/10.11999/JEIT170190).
YANG Pengcheng, LÜ Xiaode, CHAI Zhihai, *et al.* Clutter cancellation for airborne passive radar based on frequency-domain block RDS-LMS algorithm[J]. *Journal of Electronics & Information Technology*, 2017, 39(10): 2302–2310. doi: [10.11999/JEIT170190](https://doi.org/10.11999/JEIT170190).
- [13] SADIGH A N, TAHERINIA A H, and YAZDI H S. Analysis of robust recursive least squares: Convergence and tracking[J]. *Signal Processing*, 2020, 171: 107482. doi: [10.1016/j.sigpro.2020.107482](https://doi.org/10.1016/j.sigpro.2020.107482).
- [14] 赵海全, 李磊. 一种抗冲击噪声的对数总体最小二乘自适应滤波算法[J]. 电子与信息学报, 2021, 43(2): 284–288. doi: [10.11999/JEIT200344](https://doi.org/10.11999/JEIT200344).
ZHAO Haiquan and LI Lei. A logarithmic total least squares adaptive filtering algorithm for impulsive noise suppression[J]. *Journal of Electronics & Information Technology*, 2021, 43(2): 284–288. doi: [10.11999/JEIT200344](https://doi.org/10.11999/JEIT200344).
- [15] EWEDA E. A new approach for analyzing the limiting behavior of the normalized LMS algorithm under weak assumptions[J]. *Signal Processing*, 2009, 89(11): 2143–2151. doi: [10.1016/j.sigpro.2009.04.040](https://doi.org/10.1016/j.sigpro.2009.04.040).
- [16] ZHANG Sheng, ZHANG Jiashu, and SO H C. Mean square deviation analysis of LMS and NLMS algorithms with white reference inputs[J]. *Signal Processing*, 2017, 131: 20–26. doi: [10.1016/j.sigpro.2016.07.027](https://doi.org/10.1016/j.sigpro.2016.07.027).
- [17] BERMUDEZ J C M, BERSHAD N J, and EWEDA E. Stochastic analysis of the LMS algorithm for cyclostationary colored Gaussian inputs[J]. *Signal Processing*, 2019, 160: 127–136. doi: [10.1016/j.sigpro.2019.02.018](https://doi.org/10.1016/j.sigpro.2019.02.018).
- 谢小平: 男, 1978年生, 高级实验师, 研究方向为NVH.
史雄坤: 男, 1993年生, 硕士生, 研究方向为信号处理、主动降噪.

责任编辑: 余 蓉