doi:10.3772/j.issn.1002-0470.2018.09 – 10.010

比例最小均方/四阶算法及其在系统辨识中的应用①

王学成②* 张佳庚** 马文涛③***

(*陕西国际商贸学院信息与工程学院 西安 712046)

(**西安交通大学网络信息中心 西安 710049)

(***西安理工大学自动化与信息工程学院 西安 710048)

摘 要 针对非高斯噪声环境下稀疏系统参数辨识问题,提出一种基于比例更新机制的最小均方/四阶(LMS/F)自适应滤波算法(PLMS/F)。该方法以混合均方/四阶准则(MS/FE)为代价函数,其包含了误差的高阶项,具有解决非高斯噪声问题的优势。引入比例更新机制,从而可根据算法当前时刻权重变化特征来调整各权重参数的步长,因此具有良好的跟踪性能。使用梯度下降法设计了阈值参数自适应更新机制以进一步改进算法稳态性能。此外,分析了所提算法的平均和均方收敛性。应用具有稀疏特征的 FIR 系统参数模型对所提算法实现了在非高斯噪声环境中的参数辨识。仿真实验结果表明,该算法可以有效辨识模型参数,且具有较低的稳态误差和强的鲁棒性。

关键词 比例更新, 最小均方/四阶算法(LMS/F), 系统辨识, 非高斯噪声

0 引 言

自适应滤波作为一种系统辨识的有效工具已被广泛应用于工程实践中[1-5]。其中以均方误差准则(mean square error, MSE)为代价函数的最小均方(least mean square, LMS)算法和归一化最小均方算法(normalized LMS, NLMS)是其典型代表,它们具有复杂度低且易实现的特点。然而上述算法依赖于噪声能量比例[1],在实际应用中的特定条件下无法获得理想效果。以高阶项准则为代价函数的最小平均四阶(least mean fourth, LMF)和归一化 LMF(normalized LMF, NLMF)算法可有效平衡收敛速度和稳态性能[6]。尽管 LMF 算法具有优良特性,但其在高信噪比环境中存在稳定性和收敛速度慢的问题[7-9]。

近年来,学者们提出了最小均方/四阶算法(least mean square/fourth, LMS/F)[10,11],其以混合均方/四阶误差准则(mixed square/fourth error, MS/FE)为代价函数,其综合利用了 LMS 和 LMF 算法的优势。不仅保留了原算法简单和稳定的特性[11],而且优于标准 LMS 和 LMF 算法的性能。尽管 LMS/F 算法具有良好的性能,但在解决具有稀疏特征的系统参数辨识问题时无法达到更理想的结果。

为了有效解决稀疏系统辨识(sparse system identification, SSI)问题,如今已提出了许多稀疏自适应滤波算法(sparse aware adaptive filtering algorithms, SAFAs)[12-19]。对于现有的 SAFAs 的构造,主要包含基于稀疏惩罚约束[12-15]和比例更新机制(proportionate update scheme, PUS)[16-19]2 种策略。其中 PUS 策略主要是在原始自适应滤波算法中(如 LMS、NLMS 和 NLMF 等)引入增益矩阵根据当前时

① 国家自然基金(61472316),陕西省自然科学基础研究计划(2017JM6033),陕西省教育厅科研计划(17JK0550)和咸阳市科技成果推广计划(2015KT – 15)资助项目。

② 男,1978 年生,硕士;研究方向:数字信号处理,计算机应用等;E-mail: 50001980@qq.com

③ 通信作者,E-mail: mawt@xaut.edu.cn

(收稿日期:2018-04-10)

刻权重的变化来自适应更新步长,从而获得良好的系统特性跟踪能力。自从 Duttweiler[16] 首次将 PUS 引入 NLMS 算法并提出比例 NLMS(proportionate NLMS, PNLMS)算法且应用于稀疏回声干扰消除后,比例自适应滤波算法(proportionate adaptive filtering algorithms, PAFAs)的研究则获得了广泛关注,许多新的 PAFAs 被提出且成功用于解决 SSI 问题。文献[17]从收敛速度的角度提出了改进 PNLMS。文献[18]则将 PUS 机制引入到放射投影算法(affine projection algorithm, APA)构建了比例 APA。文献[19]针对稀疏分布式估计问题提出了一种比例扩散 LMS 算法。文献[20,21]进一步研究了比例更新类型的 NLMS 算法收敛性。上述 PAFAs 算法虽能有效解决 SSI 问题,但它们均是以 MSE 为准则,因此具有类似于标准 LMS 算法的缺陷,且在非高斯噪声环境(例如均匀分布或二值噪声)中其收敛性能将明显递降。研究表明误差高阶项为准则的 AFAs 对于非高斯噪声环境中的系统辨识问题将具有优于以 MSE 为代价函数的算法。Sayin 等人[22]提出了比例 NLMF(proportionate NLMF, PNLMF)算法来解决 PNLMS 算法的缺陷,而 PNLMF 也将在高信噪比环境下具有不稳定特性,且该文也主要在高斯噪声环境背景下研究了 PNLMF 算法的性能。

鉴于此,本文以 LMS/F 算法为基础设计比例更新 LMS/F 算法。首先,将比例更新机制 PUS 引入到 LMS/F 算法中,提出比例 LMS/F 算法(proportionate LMS/F, PLMS/F)并分析其特性。其次考虑到 LMS/F算法中阈值参数对算法性能的影响,应用梯度下降思想提出了变阈值 PLMS/F 算法。再次分析了算法的平均和均方收敛特性,从理论上说明算法的收敛特性。最后,本文在 Matlab 环境中,以非高斯噪声环境为研究前提,应用所提算法解决稀疏系统参数辨识问题以验证其性能。仿真实验结果表明,该方法相对于其他辨识算法具有优秀的辨识性能,即不仅具有理想的稳态精度而且具有较快的收敛速度。

1 比例最小均方/四阶算法

1.1 最小均方/四阶算法

图 1 所示为基于 AFAs 的系统辨识框架图。其中系统输出信号由 FIR 模型确定,即

$$d(i) = \boldsymbol{w}_o^{\mathrm{T}}\boldsymbol{u}(i) + v(i) \tag{1}$$

式(1)中 $\boldsymbol{w}_o = [w_{o,1}, w_{o,2}, \cdots, w_{o,N}]^{\mathrm{T}}$ 表示长度为 N 的待辨识权值向量。本文针对具有稀疏特征的参数辨识问题,即设 $\boldsymbol{w}_o$ 仅包含 K 个非零参数,且 $K << N$。$\boldsymbol{u}(i) = [u(i), u_2(i-1), \cdots, u(i-N+1)]^{\mathrm{T}}$ 为 N 维输入信号,而 $v(i)$ 表示方差为 σ_v^2 的输出测量噪声。定义输出误差为 $e(i) = d(i) - \boldsymbol{u}^{\mathrm{T}}(i)\boldsymbol{w}(i)$,其中 $\boldsymbol{w}(i) = [w_1(i), w_2(i), \cdots, w_N(i)]^{\mathrm{T}}$ 表示时刻 i 时 $\boldsymbol{w}_o$ 的估计。

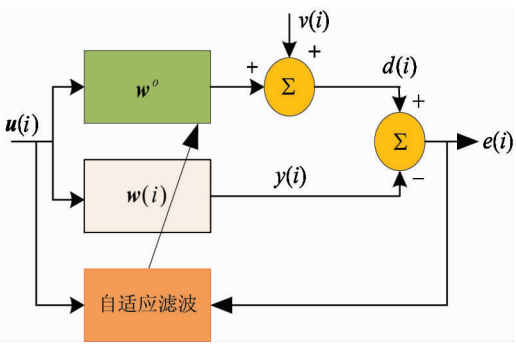

图 1 系统辨识框架图

基于以上模型,定义混合均方/四阶误差(MS/FE)准则为

$$J_{LMS/F}(i) = \frac{1}{2}e^2(i) - \frac{1}{2}\lambda \ln(e^2(i) + \lambda) \tag{2}$$

其中,λ 表示一正阈值参数,控制 LMS/F 算法的收敛速度和稳态精度。式(2)所示代价函数包含有误差平方的对数项,因此根据泰勒级数展开可知该代价函数隐含了误差的高阶项。

注记 1 文献[10]分析了 LMS/F 算法的稳定性,研究表明 LMS/F 算法具有与 LMS 算法相当的稳定性。此外,图 2 给出了误差的平方(MSE)、四阶及 MS/FE 准则的曲线图。从中可知,对于相对较大的误差值,混合均方/四阶代价函数表现出与均方代价函数相当的陡度,所以该算法则继承了 LMS 算法的

稳定性。

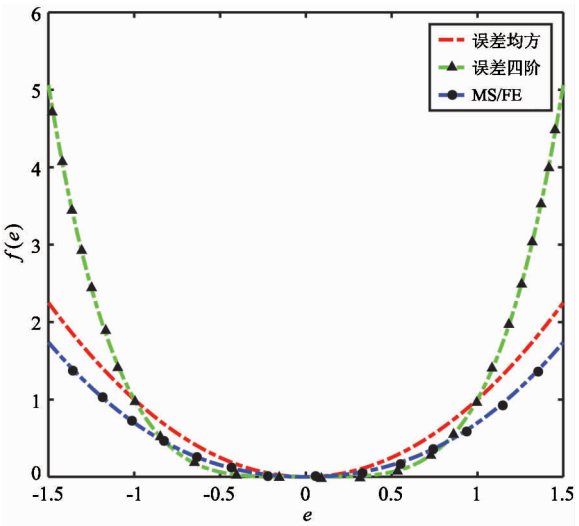

图 2 代价函数比较

以 MS/FE 为代价函数,结合梯度下降法,可得到 LMS/F 算法的迭代方程:

$$\boldsymbol{w}(i+1) = \boldsymbol{w}(i) - \mu \frac{\partial J_{LMS/F}(i)}{\partial \boldsymbol{w}(i)} = \boldsymbol{w}(i) + \mu \frac{e^3(i)\boldsymbol{u}(i)}{e^2(i) + \lambda} \tag{3}$$

其中 μ 为步长,在算法实现中需要合适地选择以平衡收敛速度和稳态精度之间的矛盾。

1.2 比例最小均方/四阶算法

比例自适应滤波算法构造的基本思想是,在传统自适应滤波算法中引入增益矩阵,其为对角矩阵,对角线上的每个元素均与当前时刻迭代权值有关,可根据权值的实时变化更新步长。因此具有良好的系统特性跟踪能力,非常适用于稀疏系统这种参数变化幅度较大的辨识问题。基于该思想,在式(3)中引入增益矩阵 $\boldsymbol{G}(i)$,即

$$\boldsymbol{w}(i+1) = \boldsymbol{w}(i) + \mu \frac{\boldsymbol{G}(i)e^3(i)\boldsymbol{u}(i)}{e^2(i) + \lambda} \tag{4}$$

其中,$\boldsymbol{G}(i) = diag(g_1(i), g_2(i), \cdots, g_N(i))$,对角线元素 $g_l(i)(l = 1, \cdots, N)$ 主要是修正权值向量中每个参数的步长,其满足:

$$g_l(i) = \gamma_l(i) / \sum_{l=1}^{N} \gamma_l(i) \quad 1 \leqslant l \leqslant N \tag{5}$$

其中,

$$\gamma_l(i) = \max[\delta \max(\xi, |w_1(i)|, |w_2(i)|, \cdots, |w_N(i)|)] \tag{6}$$

式(6)中,δ 和 ξ 表示正的参数,通常其取值为 $\delta = 5/N$, $\xi = 0.01$。式(4)则为 PLMS/F 算法的迭代公式。正是增益矩阵的引入,使得 PLMS/F 算法可以及时跟踪权值的变化特征,不断调整每个参数的步长因子,因此其在 SSI 应用中具有比 LMS/F 算法更优良的性能。

注记 2 由式(4)可知,当 $\lambda >> e^2(i)$ 时,PLMS/F 算法将表现出 PLMF 算法的性能,此时其步长可定义为 $\mu/(\lambda + e^2(i))$。而当 $\lambda << e^2(i)$ 时,PLMS/F 将退化为 PLMS 。因此,PLMS/F 算法结合了 PLMS 和 PLMF 算法的优势。

考虑到阈值参数 λ 对算法性能的影响,受广义归一化梯度下降算法启发[20],基于梯度法变阈值机制可进一步改进 PLMS/F 算法的性能。根据代价函数式(2)求解关于 λ 的微分,即:

$$\frac{\partial J_{LMS/F}(i)}{\partial \lambda} = \frac{1}{2}\left\{\ln(e^2(i) + \lambda) + \frac{\lambda}{e^2(i) + \lambda}\right\} \tag{7}$$

则应用梯度下降法,可得到关于 λ 的更新方程:

$$\lambda(i+1) = \lambda(i) + \beta\left\{\ln(e^2(i) + \lambda(i)) + \frac{\lambda(i)}{e^2(i) + \lambda(i)}\right\} \tag{8}$$

其中 β 为步长因子。为保证 $\lambda(i)$ 的递增趋势且为正的特性,对式(8)进行修正,即:

$$\lambda(i+1) = \lambda(i) + \beta\left\{|\ln(e^2(i) + \lambda(i))| + \frac{\lambda(i)}{e^2(i) + \lambda(i)}\right\} \tag{9}$$

根据仿真实验可以验证该方法的可行性。结合式(4)和式(9),可得到变阈值 PLMS/F 算法的更新方程:

$$\boldsymbol{w}(i+1) = \boldsymbol{w}(i) + \mu \frac{G(i)e^3(i)\boldsymbol{u}(i)}{e^2(i) + \lambda(i)} \tag{10}$$

本文所提算法总结如表 1 所示。

表 1　PLMS/F 与 PVLMS/F 算法

权值初始化: $\boldsymbol{w}(0) = \boldsymbol{0}$

参数设置: $\lambda, \mu, \delta, \xi, \beta$

对于迭代次数 i

$$e(i) = d(i) - \boldsymbol{u}^T(i)\boldsymbol{w}(i)$$

PLMS/F

$$\boldsymbol{w}(i+1) = \boldsymbol{w}(i) + \mu \frac{G(i)e^3(i)\boldsymbol{u}(i)}{e^2(i) + \lambda}$$

PVLMS/F

$$\lambda(i+1) = \lambda(i) + \beta\left\{ \left| \ln(e^2(i) + \lambda(i)) \right| + \frac{\lambda(i)}{e^2(i) + \lambda(i)} \right\}$$

$$\boldsymbol{w}(i+1) = \boldsymbol{w}(i) + \mu \frac{G(i)e^3(i)\boldsymbol{u}(i)}{e^2(i) + \lambda(i)}$$

结束

2　收敛性分析

本节将分析所提 PLMS/F 算法的平均和均方收敛特性。为了后续分析描述简单,将更新方程式(4)重写为

$$\boldsymbol{w}(i+1) = \boldsymbol{w}(i) + \mu \boldsymbol{G}(i) f(e(i)) e(i) \boldsymbol{u}(i) \tag{11}$$

其中 $f(e(i)) = \dfrac{e^2(i)}{e^2(i) + \lambda}$, 且做如下假设。

假设 1 输入信号 $\{\boldsymbol{u}(i)\}$ 服从独立同分布零均值高斯分布。

假设 2 噪声信号 $\{v(i)\}$ 均值为零,且独立于 $\{\boldsymbol{u}(i)\}$。

假设 3 误差非线性函数 $f(e(i))$ 与输入信号 $\{\boldsymbol{u}(i)\}$ 相互独立。

上述假设通常被应用于自适应滤波算法的收敛性分析中[23,24]。而权值向量 $\boldsymbol{w}(i)$ 位于最优解 $\boldsymbol{w}_o$ 的领域时是有效的。

2.1　平均收敛性

为了分析算法的收敛性,首先定义权值误差向量:

$$\widetilde{\boldsymbol{w}}(i) = \boldsymbol{w}(i) - \boldsymbol{w}_o \tag{12}$$

结合式(11)和(12),可得:

$$\begin{aligned}
\widetilde{\boldsymbol{w}}(i+1) &= \widetilde{\boldsymbol{w}}(i) + \mu f(e(i)) \boldsymbol{G}(i) e(i) \boldsymbol{u}(i) \\
&= \widetilde{\boldsymbol{w}}(i) + \mu f(e(i)) \boldsymbol{G}(i) (d(i) - \boldsymbol{w}^T(i)\boldsymbol{u}(i)) \boldsymbol{u}(i) \\
&= \widetilde{\boldsymbol{w}}(i) + \mu f(e(i)) \boldsymbol{G}(i) (\boldsymbol{w}_o^{\mathrm{T}} \boldsymbol{u}(i) + v(i) \\
&\quad - \boldsymbol{w}^T(i)\boldsymbol{u}(i)) \boldsymbol{u}(i) \\
&= \widetilde{\boldsymbol{w}}(i) - \mu f(e(i)) \boldsymbol{G}(i) \boldsymbol{u}(i) \boldsymbol{u}^T(i) \widetilde{\boldsymbol{w}}(i) \\
&\quad + \mu f(e(i)) \boldsymbol{G}(i) v(i) \boldsymbol{u}(i) \\
&= [\boldsymbol{I} - \mu f(e(i)) \boldsymbol{G}(i) \boldsymbol{u}(i) \boldsymbol{u}^T(i)] \widetilde{\boldsymbol{w}}(i) \\
&\quad + \mu f(e(i)) \boldsymbol{G}(i) v(i) \boldsymbol{u}(i)
\end{aligned} \tag{13}$$

其中,$\boldsymbol{I}$ 表示 $N \times N$ 单位矩阵。对式(14)两边取期望,且应用假设 1 ~ 3, 可得:

$$\begin{aligned}
&E[\widetilde{\boldsymbol{w}}(i+1)] \\
&= [I - \mu \boldsymbol{G}(i) E[f(e(i)) \boldsymbol{u}(i) \boldsymbol{u}^{\mathrm{T}}(i)]] E[\widetilde{\boldsymbol{w}}(i)] \\
&= [I - \mu \boldsymbol{G}(i) \boldsymbol{R}_f] E[\widetilde{\boldsymbol{w}}(i)]
\end{aligned} \tag{14}$$

其中, $\boldsymbol{R}_f = E[f(e(i)) \boldsymbol{u}(i) \boldsymbol{u}^{\mathrm{T}}(i)]$ 表示输入信号的加权自相关矩阵。由于 $\boldsymbol{R}_f$ 为对称且半正定,可将其分解为

$$\boldsymbol{R}_f = \boldsymbol{Q} \boldsymbol{\Lambda} \boldsymbol{Q}^{\mathrm{T}} \tag{15}$$

其中, $\boldsymbol{\Lambda} = diag(\lambda_1, \lambda_2, \cdots, \lambda_M)$ 为特征矩阵, M 表示特征值的个数;$\boldsymbol{Q}$ 为由相应特征向量构成的正交矩阵。式(14)两边同乘 $\boldsymbol{Q}^{\mathrm{T}}$, 同时用 $\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{\mathrm{T}}$ 替换 $\boldsymbol{R}_f$, 且根据 $\boldsymbol{Q}$ 的唯一性可得:

$$\widetilde{\boldsymbol{w}}'(i+1) = [\boldsymbol{I} - \mu \mathbf{G}(i) \boldsymbol{\Lambda}] \widetilde{\boldsymbol{w}}'(i) \tag{16}$$

其中 $\widetilde{\boldsymbol{w}}'(i) = \boldsymbol{Q}^{\mathrm{T}} E[\widetilde{\boldsymbol{w}}(i)]$, 而 $[\boldsymbol{I} - \mu \boldsymbol{G}(i) \boldsymbol{\Lambda}]$ 是对角阵。当且仅当对于所有的特征值满足条件 $[1 - \mu g_k(i) \lambda_k] < 1$ 时, $\widetilde{\boldsymbol{w}}'(i)$ 的每个元素则将独立地演进且收敛到 0。因此,该算法仅当它收敛于最大收敛模式(即对于 $\lambda_{\max}$)时收敛。因为 $\lambda_{\max} \leqslant tr[\boldsymbol{R}_f]$, 对于所提算法的收敛条件可根据式(16)得到,即当步长满足如式(17)所示的条件时,则权值误差向量 $\widetilde{\boldsymbol{w}}(i)$ 将随着 $i \to \infty$ 时而趋近于 0。

$$0 < \mu < \frac{2}{tr[\boldsymbol{G}(i) \boldsymbol{R}_f]} \tag{17}$$

根据假设 1 和 3,得到 $\boldsymbol{R}_f = E[f(e(i))] \sigma_x^2 \boldsymbol{I} (\sigma_x^2)$ 是输入方差。另外,使用 $tr[\boldsymbol{A}\boldsymbol{B}] \leqslant tr[\boldsymbol{A}] tr[\boldsymbol{B}]$ 和 $tr[\boldsymbol{G}(i)] = N$, 可知:

$$0 < \mu < \frac{2}{N \sigma_x^2 E[f(e(i))]} \tag{18}$$

上述结果当 $f(e(i)) = 1$ 时与参考文献[20]中一致。因此,当步长满足条件式(18)时,算法将满足平均收敛。

2.2 均方收敛性

此小节将进一步分析算法的均方收敛性,首先定义 $\tilde{\boldsymbol{w}}(i)$ 的自相关矩阵为

$$\boldsymbol{S}(i) = E[z(i)z^{\mathrm{T}}(i)] \tag{19}$$

其中 $z(i)$ 表示为:

$$z(i) = \tilde{\boldsymbol{w}}(i) - E[\tilde{\boldsymbol{w}}(i)] \tag{20}$$

结合式(20),(13)和(14),可得:

$$\begin{aligned} z(i+1) &= [\boldsymbol{I} - \mu f(e(i))\boldsymbol{G}(i)\boldsymbol{u}(i)\boldsymbol{u}^{\mathrm{T}}(i)]\tilde{\boldsymbol{w}}(n) \\ &\quad + \mu f(e(i))\boldsymbol{G}(i)v(i)\boldsymbol{u}(i) \\ &\quad - [I - \mu\boldsymbol{G}(i)\boldsymbol{R}_f]E[\tilde{\boldsymbol{w}}(i)] \\ &= [\boldsymbol{I} - \mu f(e(i))\boldsymbol{G}(i)\boldsymbol{u}(i)\boldsymbol{u}^{\mathrm{T}}(i)]z(i) \\ &\quad + \mu f(e(i))\boldsymbol{G}(i)v(i)\boldsymbol{u}(i) \\ &\quad + \mu[\boldsymbol{G}(i)\boldsymbol{R}_f \\ &\quad - f(e(i))\boldsymbol{G}(i)\boldsymbol{u}(i)\boldsymbol{u}^{\mathrm{T}}(i)]E[\tilde{\boldsymbol{w}}(i)] \\ &= \boldsymbol{\Phi}(i)z(i) + \mu\boldsymbol{\Psi}(i)E[\tilde{\boldsymbol{w}}(i)] \\ &\quad + \mu f(e(i))\boldsymbol{G}(i)v(i)\boldsymbol{u}(i) \end{aligned} \tag{21}$$

其中,

$$\begin{aligned} \boldsymbol{\Phi}(i) &= \boldsymbol{I} - \mu f(e(i))\boldsymbol{G}(i)\boldsymbol{u}(i)\boldsymbol{u}^{\mathrm{T}}(i) \\ \boldsymbol{\Psi}(i) &= \boldsymbol{G}(i)\boldsymbol{R}_f - f(e(i))\boldsymbol{G}(i)\boldsymbol{u}(i)\boldsymbol{u}^{T}(i) \end{aligned} \tag{22}$$

根据 $\boldsymbol{\Psi}(i)$ 的定义易知其均值为 0。基于假设 1 ~ 3,将式(21)带入式(19)可得:

$$\begin{aligned} \boldsymbol{S}(i+1) &= E[\boldsymbol{\Phi}(i)z(i)z^{\mathrm{T}}(i)\boldsymbol{\Phi}^{\mathrm{T}}(i)] \\ &\quad + \mu^2 E[\boldsymbol{\Psi}(i)\phi(i)\phi^{T}(i)\boldsymbol{\Psi}^{T}(i)] \\ &\quad + \mu^2 f^2(e(i))\boldsymbol{G}^2(i)\sigma_x^2\sigma_v^2 \end{aligned} \tag{23}$$

其中,σ_v^2 表示 $v(i)$ 的方差,且 $\phi(i) = E[\tilde{\boldsymbol{w}}(i)]$。利用高斯变量的四阶矩是方差平方 3 倍的结论,以及 $\boldsymbol{S}(i)$ 的对称特性[24],可得:

$$\begin{aligned} E[\boldsymbol{\Phi}(i)z(i)z^{\mathrm{T}}(i)\boldsymbol{\Phi}^{\mathrm{T}}(i)] &= \\ &(1 - 2\mu E[f(e(i))]\boldsymbol{G}(i)\sigma_x^2 \\ &+ 2\mu^2 E[f^2(e(i))]\boldsymbol{G}^2(i)\sigma_x^4)\boldsymbol{S}(i) \\ &+ \mu^2 E[f^2(e(i))]\boldsymbol{G}^2(i)\sigma_x^4 tr[\boldsymbol{S}(i)]\boldsymbol{I}_N \end{aligned} \tag{24}$$

$$\begin{aligned} E[\boldsymbol{\Psi}(i)\phi(i)\phi^{\mathrm{T}}(i)\boldsymbol{\Psi}^{\mathrm{T}}(i)] &= \\ &E[f^2(e(i))]\boldsymbol{G}^2(i)\sigma_x^4\{\phi(i)\phi^{\mathrm{T}}(i) \\ &+ tr[\phi(i)\phi^{T}(i)]\boldsymbol{I}_N\} \end{aligned} \tag{25}$$

根据式(23) ~ (25)可得:

$$\begin{aligned} tr[\boldsymbol{S}(i+1)] &= (1 - 2\mu E[f(e(i))]tr[\boldsymbol{G}(i)]\sigma_x^2 \\ &+ (M+2)\mu^2 E[f^2(e(i))]tr[\boldsymbol{G}^2(i)]\sigma_x^4)tr[\boldsymbol{S}(i)] \\ &+ (M+1)\mu^2 E[f^2(e(i))]\boldsymbol{G}^2(i)\sigma_x^4\phi^{\mathrm{T}}(i)\phi(i) \\ &+ M\mu^2 E[f^2(e(i))]\boldsymbol{G}^2(i)\sigma_x^2\sigma_v^2 \end{aligned} \tag{26}$$

由式(22)可知,$\phi(i)$ 和 $E[\boldsymbol{w}(i)]$ 收敛。因此,当且仅当算法 PLMS/F 满足下述条件时均方收敛。

$$\begin{aligned} &|(1 - 2\mu E[f(e(i))]tr[\boldsymbol{G}(i)]\sigma_x^2 \\ &+ (N+2)\mu^2 E[f^2(e(i))]tr[\boldsymbol{G}^2(i)]\sigma_x^4)| < 1 \end{aligned} \tag{27}$$

通过求解上述不等式,可获得如式(28)所示的均方稳定性条件。

$$0 < \mu < \frac{2E[f(e(i))]}{(N+2)E[f^2(e(i))]tr[\boldsymbol{G}^2(i)]\sigma_x^2} \tag{28}$$

3 仿真结果

本节将本文所提算法应用于解决 SSI 问题,以此来验证其性能,且与经典自适应滤波算法如 NLMS,PNLMS,NLMF 和 PNLMF 进行比较研究。在实验设置中,设稀疏系统参数向量长度 $N = 50$。定义系统的稀疏度(sparse rate, SR)为 $SR = K/N$。输入信号由均值为 0、方差为 4 的高斯随机序列产生。进行 500 次独立实验,通过计算均方偏差(mean square deviation, MSD)的均值作为性能评价指标,即:

$$MSD = E\{|\boldsymbol{w}(i) - \boldsymbol{w}_o|^2 / \|\boldsymbol{w}_o\|^2\} \tag{29}$$

算法中其他参数的设置均以获得其最优性能为前提,以保证性能相对的公平性。为了体现本研究所提算法在非高斯噪声环境中的优势,本实验主要选择服从均匀分布和二值分布的随机序列作为噪声序列。

算法的收敛性 在本仿真实验中设置稀疏度 $SR = 4/50$。各算法步长的设置是以保证它们均能获得相同初始收敛速度为前提条件,具体取值可见结果图。LMS/F,PLMS/F 以及 VPLMS/F 算法的正则化参数 λ 的取值为0.08,而其它算法取为0.0001。如图 3 和图 4 所示为叠加均匀分布噪声(均值为 0、方差为 6)和二值噪声的收敛曲线图。根据该结果可看出,一方面各算法在这 2 种非高斯噪声环境中均

能收敛,但比例模式算法(如 PNLMS, PNLMF 和 PLMS/F)均表现出比其原始算法优秀的稳态性能。另一方面则是所提出的 PLMS/F 算法优于 PNLMS 和 PNLMF 算法,且 VPLMS/F 算法具有更加理想的性能。此外,通过比较图 3 和图 4,可知二值噪声环境下稳态误差低于均匀噪声,该结论类似于 LMF 算法的分析[7,8]。同时由图 4 可知,基于四阶矩的算法(NLMF 和 PNLMF)的收敛速度特别慢,而若将其步长增大以提高收敛速度,但这时将不能保证该 2 种算法收敛,这则体现了 NLMF 和 PNLMF 算法在特定条件下的不稳定性。然而所提出的算法在保证合理收敛速度的条件下均能收敛。

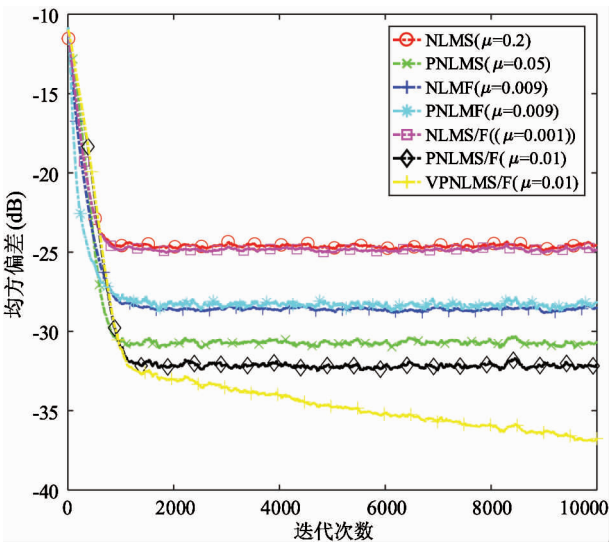

图 3　均匀分布噪声环境下收敛曲线

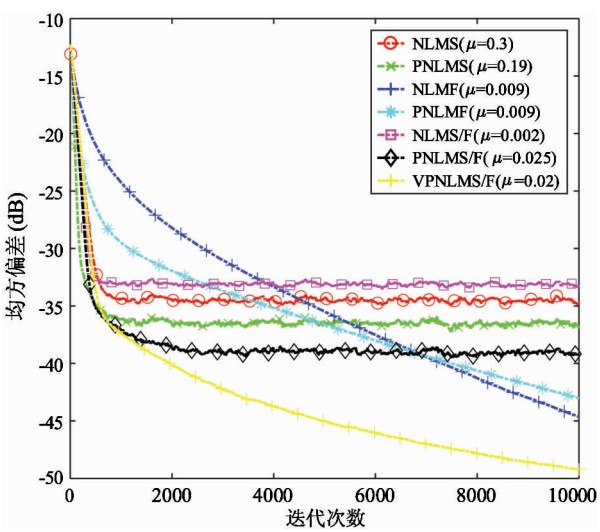

图 4　二值分布噪声环境下收敛曲线

算法的收敛速度 此仿真实验设置同算法的收敛性设置,而步长选择则主要是保证各算法具有相同的稳态精度。图 5 所示为二值噪声环境下各算法收敛曲线结果图。可知比例更新自适应滤波算法均具有比其原始算法较快的收敛速度,这主要是由于比例更新矩阵中每个元素可实时更新以调整各权重因子的步长,从而也可将其看作是一类变步长算法。该结果与其他比例更新算法的结论一致。

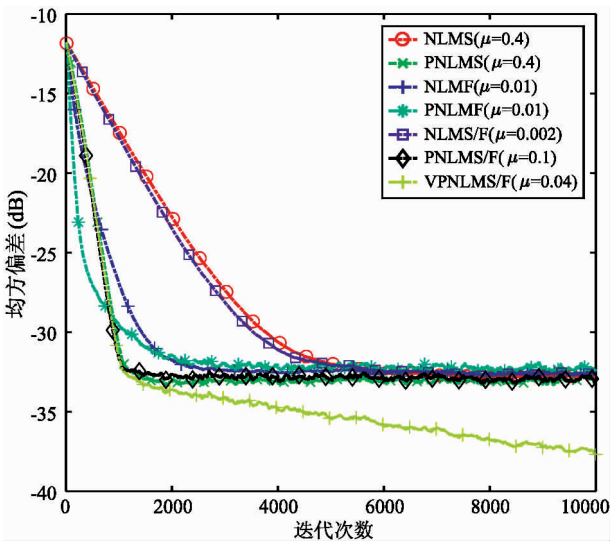

图 5　均匀分布噪声环境下收敛曲线

算法的性能 在此仿真实验中,系统的噪声设为均匀分布噪声,且设前5 000次迭代过程中的稀疏度 $SR = 3/50$,而在5 000 次迭代后 SR 改变为 8/50。图 6所示为各算法的收敛曲线图。从仿真结果中易

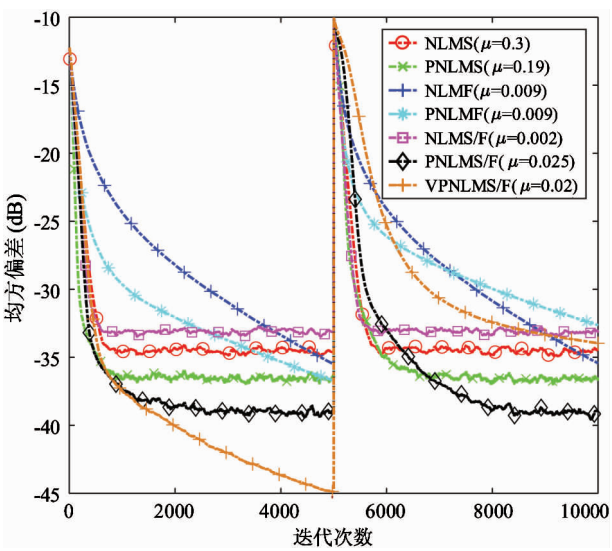

图 6　时变系统参数辨识收敛曲线

看出，PLMS/F 算法在 2 阶段始终具有较快的收敛速度和稳态精度。而 VPNLMS/F 算法在第 2 阶段的性能表现不够理想，这主要在于该算法要达到最终的稳态精度所需的迭代次数非常多，即虽具有好的稳态精度但却牺牲了收敛速度。

步长对算法收敛性的影响 根据第 3 节的分析可知，只有当步长的选择在一定范围内时，才能保证算法的均方收敛性。在此通过仿真实验来研究不同步长对所提 PLMS/F 算法收敛性的影响。图 7 所示为选择不同步长时算法的收敛曲线。从中可看出当步长选择大于0.3时，算法已无法收敛，而当步长选择太小时虽然提高了稳态精度，但是收敛速度将非常慢。因此在选择步长时，需要选择合适的步长以保证算法具有合理的收敛速度和稳态精度，如步长为 0.01 时的结果不仅具有低的稳态误差而且具有较快的收敛速度，这也验证了以上仿真实验中参数设置的合理性。如何选择最优的步长一直是自适应滤波算法研究的问题。

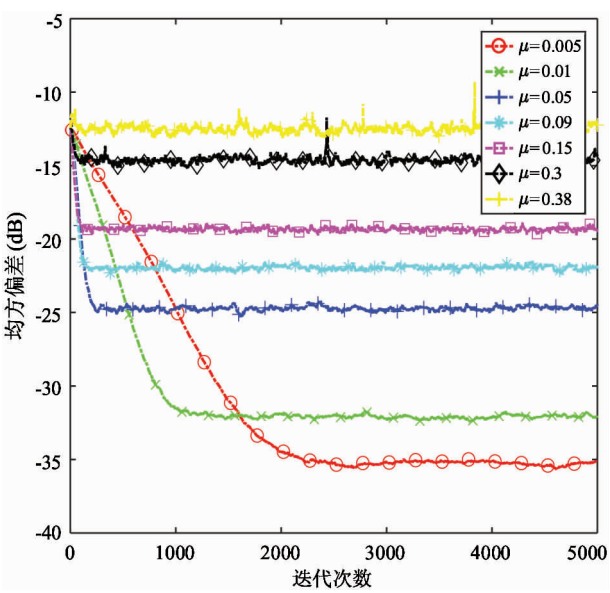

图 7　均匀分布噪声环境下收敛曲线

算法的鲁棒性 本实验选择输出噪声序列服从均匀分布，其方差分别为 3、5、6、7、8 和 10 来验证其性能。计算最后 200 次迭代的数据平均值，作为各算法稳态精度评价指标。图 8 所示为各算法的稳态均方偏差（MSD）比较结果。从中可知 PLMS/F 算法的性能在不同噪声能量下均具有最小的稳态误差。同时，图 9 给出了 PLMS/F 算法在上述噪声方差下的收敛曲线图，亦看出其具有稳定的收敛特性。

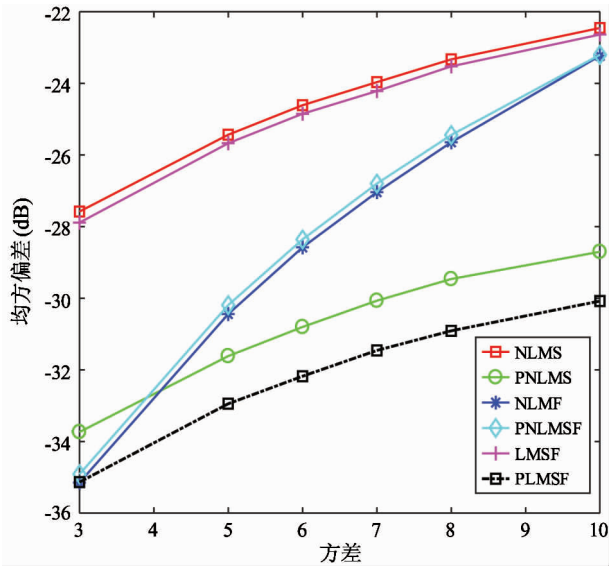

图 8　均匀分布噪声环境下稳态 MSD

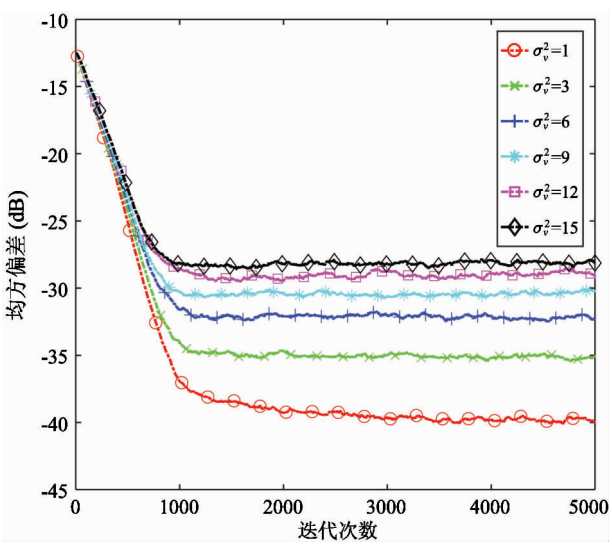

图 9　不同噪声方差条件下 PLMS/F 算法收敛曲线

4　结　论

针对非高斯噪声环境中稀疏系统参数辨识问题，本文以比例更新机制为手段，将其引入到 LMS/F 中，提出比例 LMS/F 算法 PLMS/F。同时考虑到 LMS/F 算法中阈值参数对算法性能的影响，设计了变阈值参数 PLMS/F 算法 VPLMS/F。所提算法将比例更新机制和 LMS/F 算法的优势相结合，一方面

可以有效跟踪稀疏系统参数的变化,另一方面可以有效利用误差高阶项的优势解决非高斯噪声的问题。同时对算法的收敛性进行了分析,将所提算法应用于非高斯噪声环境(均匀和二值噪声)下的 SSI 问题,仿真实验结果表明相对于其他算法,本文所提算法具有较低的稳态误差和快速的收敛速度。未来本课题将从如何提高 VPLMS/F 算法的收敛速度的角度来进行研究。

参考文献

[1] Widrow B, Stearns D. Adaptive Signal Processing[M]. New Jersey:Prentice Hall, 1985. 50-82

[2] Falconer D D. Adaptive Filter Theory and Applications [M]. Springer Berlin Heidelberg, 1980. 35-67

[3] 李国栋, 焦杜娟, 张英杰,等. 基于自适应滤波的带宽测量方法研究[J]. 高技术通讯, 2011, 21(5):478-481

[4] 粟梅, 王莉娜, 张泰山,等. 基于自适应滤波器的电网谐波电流预测法[J]. 控制与决策, 2004, 19(12): 1429-1432

[5] Ma W, Qu H, Zhao J. Estimator with forgetting factor of correntropy and recursive algorithm for traffic network prediction[C]. In: Proceedings of the 25th Chinese Control and Decision Conference (CCDC), Guiyang, China, 2013. 490-494

[6] Walach E, Widrow B. The least mean fourth (LMF) adaptive algorithm and its family[J]. *IEEE Transactions on Information Theory*, 1984, 30(2):275-283

[7] Hübscher P I, Bermudez J C M. An improved statistical analysis of the least mean fourth (LMF) adaptive algorithm[J]. *IEEE Transactions on Signal Processing*, 2003, 51(3):664-671

[8] Eweda E. Global stabilization of the least mean fourth algorithm[J]. *IEEE Transactions on Signal Processing*, 2012, 60(3):1473-1477

[9] Eweda E. Dependence of the stability of the least mean fourth algorithm on target weights non-stationarity[J]. *IEEE Transactions on Signal Processing*, 2014, 62(7): 1634-1643

[10] Sayin M O, Vanli N D, Kozat S S. A novel family of adaptive filtering algorithms based on the logarithmic cost [J]. *IEEE Transactions on Signal Processing*, 2014, 62 (17):4411-4424

[11] Gui G, Xu L, Matsushita S. Improved adaptive sparse channel estimation using mixed square/fourth error criterion[J]. *Journal of the Franklin Institute*, 2015, 352 (10):4579-4594

[12] Gu Y, Jin J, Mei S. l_0-norm constraint LMS algorithm for sparse system identification[J]. *IEEE Signal Processing Letters*, 2009, 16(9):774-777

[13] Ma W, Qu H, Gui G, et al. Maximum correntropy criterion based sparse adaptive filtering algorithms for robust channel estimation under non-Gaussian environments[J]. *Journal of the Franklin Institute*, 2015, 352(7): 2708-2727

[14] Li Y, Wang Y, Jiang T. Norm-adaption penalized least mean square/fourth algorithm for sparse channel estimation[J]. *Signal processing*, 2016, 128:243-251

[15] Ma W, Chen B, Qu H, et al. Sparse least mean p-power algorithms for channel estimation in the presence of impulsive noise[J]. *Signal, Image and Video Processing*, 2016, 10(3): 503-510

[16] Duttweiler D L. Proportionate normalized least-mean-squares adaptation in echo cancellers[J]. *IEEE Transactions on Speech and Audio Processing*, 2000, 8(5):508-518

[17] Deng H, M. Doroslova $\check{c}$ki. Proportionate adaptive algorithms for network echo cancellation[J]. *IEEE Transactions on Signal Processing*, 2006, 54(5):1794-1803

[18] Zheng Z, Liu Z, Dong Y. Steady-state and tracking analyses of improved proportionate affine projection algorithm [J]. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2018, 65(11):1793-1793

[19] Das R L, Chakraborty M. On convergence of proportionate-type normalized least mean square algorithms[J]. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2015, 62(5):491-495

[20] Jelfs B, Mandic D P. A unifying framework for the analysis of proportionate NLMS algorithms[J]. *International Journal on Adaptive Control and Signal Processing*, 2015, 29(9):1073-1085

[21] Yim S H, Lee H S, Song W J. A proportionate diffusion LMS algorithm for sparse distributed estimation[J]. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2015, 62(10):992-996

[22] Sayin M O, Yilmaz Y, Demir A, et al. The Krylov-proportionate normalized least mean fourth approach: Formulation and performance analysis[J]. *Signal Processing*, 2015, 109: 1-13

[23] Das R L, Chakraborty M. Improving the performance of the PNLMS algorithm using norm regularization [J]. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*., 2016,24(7):1280-1290

[24] Shi K, Shi P. Convergence analysis of sparse LMS algorithms with l_1-norm penalty based on white input signal [J]. *Signal Processing*, 2010, 90(12):3289-3293

Proportionate least mean square/fourth algorithm and its application for system identification

Wang Xuecheng*, Zhang jiageng**, Ma Wentao***

(*School of Information and Engineering, Shannxi Institute of International Trade & Commerce, Xi'an 712046)

(**Center for Network and Information, Xi'an Jiaotong University, Xi'an 710049)

(***Shool of Automation and Information Engineering, Xi'an University of Technology, Xi'an 710048)

Abstract

For sparse system parameter identification problem under non-Gaussian noise environment, the least mean square/fourth (LMS/F) algorithm based on proportionate update scheme, namely PLMS/F, is proposed. The PLMS/F algorithm with the mixed square/fourth error criterion (including higher order term of error) has advantage of solving non-Gaussian problem. In addition, it has good tracking ability via the introduced proportionate update scheme which adjusts the step size of each weight parameter according to the weight change feature at instant time. A threshold adaptively update approach is developed to further improve the performance of the PLMS/F algorithm. Furthermore, we perform the mean and mean square convergence analysis of the proposed algorithm. Taking the FIR system parameter model with sparse feature as the object, the identification under the non-Gaussian environment is realized by using the proposed algorithm. Simulation results show that the proposed PLMS/F and BCPLMS/F algorithms can identify the model parameter efficiently, and have lower steady-state misalignment and stronger robustness in comparison to other algorithms.

Key words: proportionate update, least mean square/fourth (LMS/F), system identification, non-Gaussian noise