

基于P范数的核最小对数绝对差自适应滤波算法

火元莲 脱丽华 齐永锋 丁瑞博

Kernel least logarithm absolute difference algorithm based on P-norm

Huo Yuan-Lian Tuo Li-Hua Qi Yong-Feng Ding Rui-Bo

引用信息 Citation: *Acta Physica Sinica*, 71, 048401 (2022) DOI: 10.7498/aps.71.20211124

在线阅读 View online: <https://doi.org/10.7498/aps.71.20211124>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

您可能感兴趣的其他文章

Articles you may be interested in

非高斯冲激干扰下基于Softplus函数的核自适应滤波算法

Kernel adaptive filtering algorithm based on Softplus function under non-Gaussian impulse interference

物理学报. 2021, 70(2): 028401 <https://doi.org/10.7498/aps.70.20200954>

一种变尺度S型核分式低次幂自适应滤波算法

A variable-scale S-type kernel fractional low-power adaptive filtering algorithm

物理学报. 2021, 70(15): 158401 <https://doi.org/10.7498/aps.70.20210075>

混沌信号自适应协同滤波去噪

An adaptive denoising algorithm for chaotic signals based on collaborative filtering

物理学报. 2018, 67(6): 060501 <https://doi.org/10.7498/aps.67.20172470>

基于自适应Kalman滤波的平面阵列电容成像

Planar array capacitance imaging based on adaptive Kalman filter

物理学报. 2021, 70(11): 118102 <https://doi.org/10.7498/aps.70.20210442>

基于大脑情感学习模型和自适应遗传算法的混沌时间序列预测

Chaotic time series prediction based on brain emotional learning model and self-adaptive genetic algorithm

物理学报. 2018, 67(8): 080502 <https://doi.org/10.7498/aps.67.20172104>

放疗绝对剂量的数学算法模型

Mathematical algorithm model of absolute dose in radiotherapy

物理学报. 2021, 70(1): 018701 <https://doi.org/10.7498/aps.70.20200986>

基于 P 范数的核最小对数绝对差自适应滤波算法*

火元莲^{1)†} 脱丽华¹⁾ 齐永锋²⁾ 丁瑞博¹⁾

1) (西北师范大学, 物理与电子工程学院, 兰州 730000)

2) (西北师范大学, 计算机科学与工程学院, 兰州 730000)

(2021 年 6 月 14 日收到; 2021 年 11 月 24 日收到修改稿)

为了进一步提高在 α 稳定分布噪声背景下非线性自适应滤波算法的收敛速度, 本文提出了一种新的基于 p 范数的核最小对数绝对差自适应滤波算法 (kernel least logarithm absolute difference algorithm based on p -norm, P-KLLAD). 该算法结合核最小对数绝对差算法和 p 范数, 一方面利用最小对数绝对差准则保证了算法在 α 稳定分布噪声环境下良好的鲁棒性, 另一方面在误差的绝对值上添加 p 范数, 通过 p 范数和一个正常数 a 来控制算法的陡峭程度, 从而提高该算法的收敛速度. 在非线性系统辨识和 Mackey-Glass 混沌时间序列预测的仿真结果表明, 本文算法在保证鲁棒性能的同时提高了收敛速度, 并且在收敛速度和鲁棒性方面优于核最小均方误差算法、核分式低次幂算法、核最小对数绝对差算法和核最小平均 p 范数算法.

关键词: α 稳定分布噪声, 核自适应滤波算法, 最小对数绝对差准则, p 范数

PACS: 84.30.Vn, 05.40.Ca

DOI: 10.7498/aps.71.20211124

1 引言

核自适应滤波器^[1] (kernel adaptive filtering, KAF) 作为一种在重构核希尔伯特空间 (reproducing kernel Hilbert space, RKHS) 中开发的高效非线性近似方法, 利用核函数^[2] 把原始空间的输入数据映射到 RKHS 空间, 从而高效解决非线性问题. 在非线性系统辨识^[3]、回声消除^[4]、时间序列预测^[5–6]和信道均衡^[7–8]等方面有很广泛的应用. 在真实的应用环境中都存在脉冲噪声和非高斯噪声, 这些噪声通常由一些自然现象和电子设备产生, 例如水下声学信号、低频大气噪声、各种类型的人为噪声, 在电力通信、控制系统、雷达探测、语音处理等实际领域的应用中较为常见. 这些噪声特性在统计意

义下其概率密度分布呈现较为厚重的拖尾现象, 而 α 稳定分布可以对这类非高斯噪声进行很好的建模^[9–11]. 普通的核最小均方 (kernel least mean square, KLMS) 类算法通常在高斯噪声下表现得很好^[12], 但由于最小均方误差 (mean square error, MSE) 准则^[13] 只捕获误差信号的二阶统计量, 致使该类算法对异常值很敏感, 即缺乏鲁棒性. 于是一系列抗脉冲噪声干扰的核自适应滤波算法被提出. 文献^[14] 通过结合最小平均 p 范数 (mean p -power error, MPE) 准则^[15–16] 和核技巧, 提出了核最小平均 p 范数算法 (kernel least mean p -power, KLMP)、核递归最小平均 p 范数算法 (kernel recursive least mean p -power, KRLP), 在 α 稳定分布噪声环境下保证收敛速度的同时鲁棒性得到了提升. 文献^[17] 针对非高斯脉冲噪声环境提

* 国家自然科学基金 (批准号: 61561044) 资助的课题.

† 通信作者. E-mail: huoyuanlian@163.com

出了基于分数低阶统计误差准则的核分式低次幂算法 (kernel kernel fractional lower power, KFLP). 文献 [17] 在针对文献 [18] 收敛速度慢的问题, 在此基础上结合 Softplus 函数的非线性饱和特性提出了非高斯冲激干扰下基于 Softplus 函数的核自适应滤波算法. 文献 [19] 结合线性最小对数绝对差算法^[20] (least logarithmic absolute difference, LLAD) 和核技巧提出了核最小对数绝对差算法 (kernel least logarithmic absolute difference, KLLAD), 使最小对数绝对差应用到非线性系统并且保证了其良好的抗脉冲干扰能力.

KLLAD 算法可以很好地处理异常值, 但存在收敛速度慢的问题, 为了改进 KLLAD 算法的收敛速度, 本文结合核最小对数绝对差算法和 p 范数, 提出了一种新的鲁棒核自适应滤波算法, 基于 p 范数的核最小对数绝对差算法 (kernel least logarithm

absolute difference algorithm based on p -norm, P-KLLAD). 该算法在对数的基础上添加了 p 范数, 通过 p 范数和一个正常数 a 来控制代价函数的陡峭程度, 从而控制算法的收敛速度, 使该算法在保证鲁棒性的基础上提升其收敛速度. 当背景噪声为 α 稳定分布噪声时, 在非线性系统辨识和 Mackey-Glass 混沌时间序列预测的仿真表明, 该算法的在收敛速度和鲁棒性方面优于核最小均方算法、核分式低次幂算法、核最小绝对差算法和核最小平均 p 范数算法.

2 α 稳定分布模型

α 稳定分布和其他分布函数不同, 一般情况下没有闭式描述的概率密度函数表达式, 通常采用以下形式的特征函数来进行表示:

$$\phi(t) = \begin{cases} \exp \left\{ j\delta t - |\gamma t|^\alpha \left[1 - j\beta \operatorname{sgn}(t) \tan \left(\frac{\pi\alpha}{2} \right) \right] \right\}, & \alpha \neq 1, \\ \exp \left\{ j\delta t - |\gamma t| \left[1 - j\beta \operatorname{sgn}(t) \frac{2}{\pi} \tan(|t|) \right] \right\}, & \alpha = 1, \end{cases} \quad (1)$$

其中 j 表示复数单位, sgn 表示符号函数.

α 稳定分布主要有 α , δ , β 和 γ 四个参数共同决定.

1) $\alpha \in (0, 2]$: 特征指数, 它可以控制概率密度函数的拖尾的厚度. 当 α 取值越接近于 0 时, 信号或噪声出现脉冲的概率就越大. 当 α 为 2 时, 该分布退化为高斯分布.

2) $\delta \in R$: 位置参数, 它描述稳定分布的中心位置. 当 $\delta \in (1, 2]$ 时, δ 表示稳定分布的均值. 当 $\delta \in (0, 1]$ 时, δ 表示稳定分布的中值.

3) $\beta \in [-1, 1]$: 对称系数, 它决定着稳定分布相对于中心 δ 的倾斜程度. 当 $\beta = 0$ 时, 稳定分布是对称的. 当 $\beta \neq 0$ 时, 稳定分布是非对称的.

4) $\gamma \in R^+$: 分散系数, 它描述稳定分布相对于中心 δ 的扩散程度, γ 越大预示着脉冲噪声强度越大.

α 稳定分布具有三种特殊的情形: 当 $\alpha = 2$, $\beta = 0$ 时, α 稳定分布对应于高斯分布; 当 $\alpha = 1$, $\beta = 0$ 时, α 稳定分布对应于柯西分布; 当 $\alpha = 1/2$, $\beta = -1$ 时, α 稳定分布对应于皮尔逊分布.

3 核最小对数绝对差算法

核最小对数绝对差算法的代价函数:

$$J = |e(n)| - \frac{1}{a} \ln(1 + a|e(n)|), \quad (2)$$

式中 $e(n) = d(n) - \Omega(n-1)\varphi(n)$ 是迭代 n 时刻的预测误差, $d(n)$ 是观测到的期望信号, $\Omega(n-1)$ 是核自适应滤波器权重, $\varphi(n)$ 是核自适应滤波器输入. a 是常数, 它决定了代价函数 J 的陡峭程度. 利用负随机梯度可以得到算法的权重更新公式:

$$\Omega(n) = \Omega(n-1) + \eta \frac{ae(n)\varphi(n)}{1 + ae(n)}, \quad (3)$$

其中 η 表示步长, 则 $n+1$ 时刻的滤波器输出为

$$\begin{aligned} y(n+1) &= \Omega(n)^T \varphi(n+1) \\ &= \eta \sum_{j=1}^n \frac{ae(j)}{1 + ae(j)} \varphi(j)^T \varphi(n+1), \end{aligned} \quad (4)$$

根据 Mercer 定理, 任意核从输入空间 U 映射到特征空间 F 都可以表示为

$$\kappa(u, u') = \varphi^T(u) \varphi(u'), \quad (5)$$

其中 u 表示输入信号在 n 时刻的值, u' 表示输入信号在下一时刻的值, (5) 式被称为核技巧, 常用的核是高斯核.

$$\kappa(\mathbf{u}, \mathbf{u}') = \exp(-h\|\mathbf{u} - \mathbf{u}'\|^2), \quad (6)$$

其中 h 为高斯核的核宽.

4 基于 P 范数的核最小对数绝对差算法

4.1 算法推导

本文提出的基于 p 范数的核最小对数绝对差自适应滤波算法是将核最小对数绝对差算法和 p 范数相结合, 构造一个新的代价函数. 在误差绝对值上添加了 p 范数, 并且添加了一个正常数来控制算法的陡峭程度, 使算法的收敛速度相比于核最小对数绝对差算法得到了提升, 同时利用最小对数绝对差准则保证了算法具有良好的鲁棒性. 本文算法的代价函数为

$$J = a[|e(n)|^p - \ln(1 + |e(n)|^p)]. \quad (7)$$

p 范数和正常数 a 一起决定了代价函数 J 的陡峭程度, 来控制收敛速度和稳态误差. 对 (7) 式用最速下降法求梯度得:

$$\frac{\partial J(n)}{\partial \Omega(n-1)} = -\frac{ap|e(n)|^{2p-1}}{1+|e(n)|^p} \text{sgn}(e(n)) \varphi(n), \quad (8)$$

利用负随机梯度可以得到算法的权重更新公式:

$$\begin{aligned} \Omega(n) &= \Omega(n-1) - \eta \frac{\partial J(n)}{\partial \Omega(n-1)} \\ &= \Omega(n-1) + \eta \frac{ap|e(n)|^{2p-1}}{1+|e(n)|^p} \text{sgn}(e(n)) \varphi(n), \end{aligned} \quad (9)$$

其中 η 表示步长, 逐项递推得到如下形式:

$$\Omega(n) = \Omega(0) + \eta \sum_{j=1}^n \frac{ap|e(j)|^{2p-1}}{1+|e(j)|^p} \text{sgn}(e(j)) \varphi(j), \quad (10)$$

式中 $\Omega(0) = 0$, 则权重更新公式为

$$\Omega(n) = \eta \sum_{j=1}^n \frac{ap|e(j)|^{2p-1}}{1+|e(j)|^p} \text{sgn}(e(j)) \varphi(j), \quad (11)$$

本文 p 的取值为 1.1. 从权重更新公式可以看出, 当出现脉冲异常值的时候, $e(n)$ 会增大, 当 $e(n)$ 是小异常值时, 由于 p 的取值很接近于 1, 那么权重更新公式会停止更新. 但是如果 $e(n)$ 是大异常值时, 那么分母项就会远大于分子项, 该算法就不具

有很好的抗脉冲噪声效果. 综上, 滤波器 $n+1$ 时刻的输出为

$$\begin{aligned} y(n+1) &= \Omega(n)^T \varphi(n+1) = \\ &\eta \sum_{j=1}^n \frac{ap|e(j)|^{2p-1}}{1+|e(j)|^p} \text{sgn}(e(j)) \varphi^T(j) \varphi(n+1). \end{aligned} \quad (12)$$

因此得到本文提出的 P-KLLAD 算法如表 1 所示.

表 1 P-KLLAD 算法
Table 1. P-KLLAD algorithm.

P-KLLAD算法
初始化:
选择合适的参数 a , p 和核宽 h ,
$a_1(1) = \eta \frac{ap d(1) ^{2p-1}}{1+ d(1) ^p} \text{sgn}(d(1))$
更新:
每获得一对 $\{\mathbf{u}(i), d(i)\}$, 计算输出值:
$f_{i-1}(\mathbf{u}(i)) = \eta \sum_{j=1}^{i-1} a_j(i-1) \kappa(\mathbf{u}(i), \mathbf{u}(j))$
计算误差: $e(i) = d(i) - f_{i-1}(\mathbf{u}(i))$
更新系数: $a_i(i) = \eta \frac{ap e(i) ^{2p-1}}{1+ e(i) ^p} \text{sgn}(e(i))$

4.2 均方收敛分析

对本文提出的算法 P-KLLAD 进行均方收敛性能分析, 首先推导该算法的能量守恒关系, 再利用该能量守恒关系推导出最后的收敛公式. 针对非线性系统辨识, 可以得到未知系统输出为

$$d(n) = \Omega_0 \varphi(n) + v(n), \quad (13)$$

其中 $v(n)$ 是背景噪声, 则输出误差为

$$e(n) = d(n) - \Omega(n-1)^T \varphi(n), \quad (14)$$

将 (13) 式代入 (14) 式得到:

$$\begin{aligned} e(n) &= \Omega_0 \varphi(n) + v(n) - \Omega(n-1)^T \varphi(n) \\ &= \tilde{\Omega}(n-1)^T \varphi(n) + v(n) \\ &= e_a(n) + v(n), \end{aligned} \quad (15)$$

其中 $\tilde{\Omega}(n-1) = \Omega_0 - \Omega(n-1)$ 表示特征空间中权矢量的偏差, $e_a(n) = \tilde{\Omega}(n-1)^T \varphi(n)$ 表示第 n 次迭代时的先验误差. (9) 式两边同时减去 Ω_0 可得到:

$$\tilde{\Omega}(n) = \tilde{\Omega}(n-1) - \eta \frac{ap|e(n)|^{2p-1}}{1+|e(n)|^p} \text{sgn}(e(n)) \varphi(n). \quad (16)$$

令 $e_p(n) = \tilde{\boldsymbol{\Omega}}(n)^T \boldsymbol{\varphi}(n)$ 表示后验误差, 将 (16) 式代入可以表示为

$$\begin{aligned} e_p(n) &= \tilde{\boldsymbol{\Omega}}(n-1)^T \boldsymbol{\varphi}(n) - \eta \frac{ap|e(n)|^{2p-1}}{1+|e(n)|^p} \text{sgn}(e(n)) \boldsymbol{\varphi}^T(n) \boldsymbol{\varphi}(n) \\ &= e_a(n) - \eta \frac{ap|e(n)|^{2p-1}}{1+|e(n)|^p} \text{sgn}(e(n)) \kappa(\mathbf{u}, \mathbf{u}'), \end{aligned} \quad (17)$$

联合式 (16) 和式 (17) 消除 $e(n)$, 可得:

$$\tilde{\boldsymbol{\Omega}}(n) = \tilde{\boldsymbol{\Omega}}(n-1) + \frac{e_p(n) - e_a(n)}{\kappa(\mathbf{u}, \mathbf{u}')} \boldsymbol{\varphi}(n), \quad (18)$$

对 (18) 式两边取内积:

$$\|\tilde{\boldsymbol{\Omega}}(n)\|^2 = \|\tilde{\boldsymbol{\Omega}}(n-1)\|^2 + \frac{e_p^2(n)}{\kappa(\mathbf{u}, \mathbf{u}')^2} - \frac{e_a^2(n)}{\kappa(\mathbf{u}, \mathbf{u}')^2} + \gamma_q. \quad (19)$$

(19) 式表示 P-KLLAD 算法的能量守恒关系. 由于 $\|\tilde{\boldsymbol{\Omega}}(n)\|^2 = \tilde{\boldsymbol{\Omega}}(n)^T \tilde{\boldsymbol{\Omega}}(n)$, 则:

$$\gamma_q = \frac{1}{\kappa(\mathbf{u}, \mathbf{u}')^2} 2(e_p(n) - e_a(n)) \left\{ \tilde{\boldsymbol{\Omega}}(n-1)^T \boldsymbol{\varphi}(n) \kappa(\mathbf{u}, \mathbf{u}') - e_a(n) \right\}, \quad (20)$$

将 (17) 式和 (20) 式代入 (19) 式, 可得:

$$\begin{aligned} \|\tilde{\boldsymbol{\Omega}}(n)\|^2 &= \|\tilde{\boldsymbol{\Omega}}(n-1)\|^2 + \left(\eta \frac{ap|e(n)|^{2p-1}}{1+|e(n)|^p} \text{sgn}(e(n)) \right)^2 \\ &\quad - 2 \left(\eta \frac{ap|e(n)|^{2p-1}}{1+|e(n)|^p} \text{sgn}(e(n)) \right) \tilde{\boldsymbol{\Omega}}(n-1)^T \boldsymbol{\varphi}(n). \end{aligned} \quad (21)$$

对 (21) 式两边取期望:

$$\begin{aligned} E \left[\|\tilde{\boldsymbol{\Omega}}(n)\|^2 \right] &= E \left[\|\tilde{\boldsymbol{\Omega}}(n-1)\|^2 \right] + \eta^2 a^2 p^2 E \left[\left(\frac{|e(n)|^{2p-1}}{1+|e(n)|^p} \text{sgn}(e(n)) \right)^2 \right] \\ &\quad - 2\eta ap E \left[\frac{|e(n)|^{2p-1}}{1+|e(n)|^p} \text{sgn}(e(n)) \tilde{\boldsymbol{\Omega}}(n-1)^T \boldsymbol{\varphi}(n) \right], \end{aligned} \quad (22)$$

如果算法收敛必须满足:

$$\begin{aligned} E \left[\|\tilde{\boldsymbol{\Omega}}(n)\|^2 \right] &\leq E \left[\|\tilde{\boldsymbol{\Omega}}(n-1)\|^2 \right] \Leftrightarrow \eta^2 a^2 p^2 E \left[\left(\frac{|e(n)|^{2p-1}}{1+|e(n)|^p} \text{sgn}(e(n)) \right)^2 \right] \\ &\quad - 2\eta ap E \left[\frac{|e(n)|^{2p-1}}{1+|e(n)|^p} \text{sgn}(e(n)) \tilde{\boldsymbol{\Omega}}(n-1)^T \boldsymbol{\varphi}(n) \right] \leq 0 \Leftrightarrow \eta \leq \frac{2E \left[\tilde{\boldsymbol{\Omega}}(n-1)^T \boldsymbol{\varphi}(n) \right]}{ap E \left[\frac{|e(n)|^{2p-1}}{1+|e(n)|^p} \text{sgn}(e(n)) \right]}. \end{aligned} \quad (23)$$

所以保证该算法收敛的充分必要条件为

$$0 < \eta \leq \frac{2E \left[\tilde{\boldsymbol{\Omega}}(n-1)^T \boldsymbol{\varphi}(n) \right]}{ap E \left[\frac{|e(n)|^{2p-1}}{1+|e(n)|^p} \text{sgn}(e(n)) \right]}. \quad (24)$$

4.3 计算复杂度分析

对本文所提出的 P-KLLAD 算法的复杂度进

行分析, 将复杂度分为乘除、加减、指数、符号函数和分数幂几种运算类型, 并与 KLMS, KFLP, KLMP 和 KLLAD 算法作比较. 所有算法在迭代次数为 i 时分别计算其不同运算类型的次数, 结果如表 2 所示. 由表 2 可以看出, 所有算法的运算次数都随着迭代次数 i 的增加而增加, 这是因为核自适应滤波算法的网络尺寸随着输入数据的增加而增大. 在第 i 次迭代时, 本文算法 P-KLLAD 相比

于 KLLAD 算法增加了符号函数和分数幂的运算, 相比于 KLMP 算法增加了少量乘除和分数幂运算. 虽然 P-KLLAD 算法在一定程度上增加了计算复杂度, 但它在保持良好鲁棒性的同时又提高了收敛速度, 因此这并不妨碍该算法的实际应用.

表 2 算法的复杂度对比
Table 2. Complexity of the algorithm.

算法名称	乘法或 除法	加法或 减法	指数 运算	符号 函数	分数幂 运算
KLMS	$3i-1$	$2i-1$	$i-1$	—	—
KFLP	$3i-2$	$2i-1$	$i-1$	—	—
KLMP	$3i-2$	$2i-1$	$i-1$	1	1
KLLAD	$3i+1$	$2i-1$	$i-1$	—	—
P-KLLAD	$3i+1$	$2i-1$	$i-1$	1	2

4.4 参数选择

4.4.1 范数 p 对 P-KLLAD 算法性能影响

令 $a=3$, 范数 p 分别取 0.7, 0.9, 1.1, 1.3 和 1.5 时 P-KLLAD 算法的学习曲线如图 1 所示.

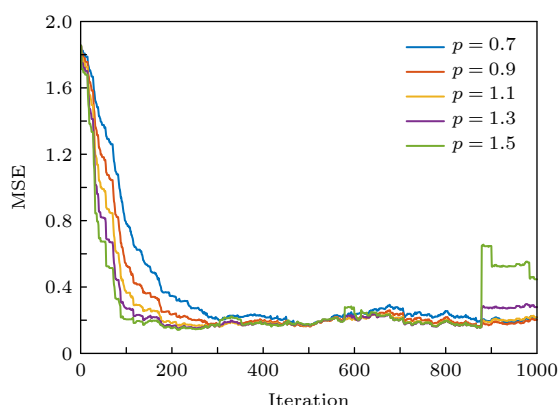


图 1 P-KLLAD 在不同 p 下的 MSE 曲线
Fig. 1. The MSE curves of P-KLLAD under different p .

从图 1 可以看出, 随着 p 的增大, 本文算法的稳态性能和收敛速度都会变得更优, 但是当 $p \geq 1.3$ 时该算法的抗脉冲噪声的能力会下降. 因此, 本文取 p 值为 1.1, 这在一定程度上平衡了收敛速度和鲁棒性.

4.4.2 正常数 a 对 P-KLLAD 算法性能影响

令 $p=1.1$, 正常数 a 分别取 1, 2, 3, 4 和 5 时 P-KLLAD 算法的学习曲线如图 2 所示.

从图 2 可以看出本文算法的收敛速度随着 a 的增大而增大, 但稳态误差趋于不稳定. 为了平衡收敛速度和稳态误差, 本文中的 a 取值为 3.

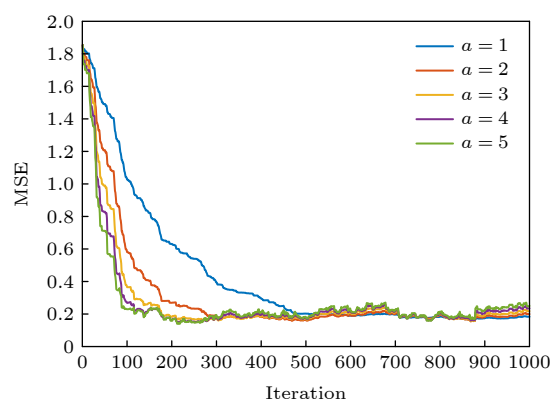


图 2 P-KLLAD 在不同 a 下的 MSE 曲线
Fig. 2. The MSE curves of P-KLLAD under different a .

4.5 α 稳定分布噪声不同参数对算法性能的影响

首先讨论 α 稳定分布噪声的特征指数 α 取不同值时对算法性能的影响. 文中令 $\gamma=0.1$, α 分别取 1.2, 1.5, 1.7 和 1.9, P-KLLAD 算法的学习曲线如图 3 所示. 从图 3 可以看出, 本文的算法在不同 α 下都具有良好的抗脉冲干扰能力, 随着 α 的增大, 脉冲出现的概率也逐渐较小.

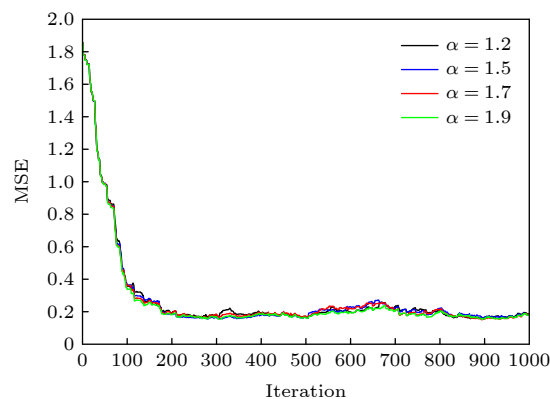
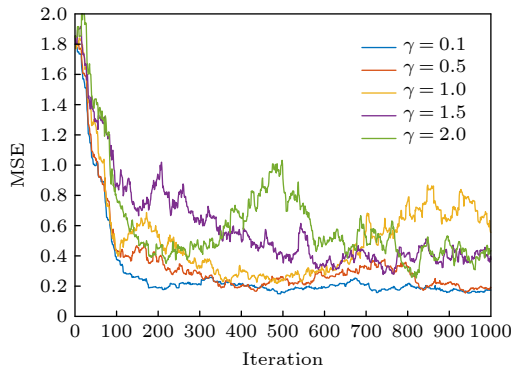


图 3 P-KLLAD 在不同 α 下的 MSE 曲线
Fig. 3. The MSE curves of P-KLLAD under different α .

图 4 给出了 α 稳定分布噪声的特征指数 $\alpha=1.3$, 分散系数 γ 取不同值时对算法性能的影响, 令 γ 分别取 0.1, 0.5, 1.0, 1.5 和 2.0. 从图 4 可以看出, 本文算法在不同 γ 下抗脉冲噪声干扰的效果也不同, 随着 γ 的增大噪声的脉冲会增大, 在小脉冲的时候算法具有良好的鲁棒性, 而在大脉冲的时候本文算法出现性能下降的问题.

综上, 在后续实验中选择 α 稳定分布噪声的参数 $\alpha=1.3$, $\gamma=0.1$ 来表示有脉冲干扰的噪声.

图 4 P-KLLAD 在不同 γ 下的 MSE 曲线Fig. 4. The MSE curves of P-KLLAD under different γ .

5 算法仿真

为了进一步说明本文所提算法在实际应用中的有效性, 将本文所提出的算法与 KLMS, KLMP, KLLAD, KFLP 算法在非线性系统辨识和 Mackey-Glass 混沌时间序列中就收敛性和抗脉冲干扰性能进行对比. 所有实验高斯核核宽 $h = 1.0$, 在两个仿真系统中, 每种算法均采用 1000 个样本作为训练数据, 100 个样本用于测试, 得到的 MSE 性能曲线被用来评价算法性能.

5.1 非线性系统辨识

本实验用的非线性系统是由一个线性模型和一个非线性模型组合而成, 其中线性模型为: $H(z) = 1 + 0.2z^{-1}$, 非线性模型为: $f(i) = x(i) - 0.9x(i)^2$, 则系统的期望输出: $d(i) = x(i) - 0.9x(i)^2 + v(i)$. 其中 $v(i)$ 是环境噪声, 用 α 稳定分布噪声来表示. 为了使不同算法在各自最理想的情况下进行比较, 给出了不同算法的最优参数, 如表 3 所示.

表 3 各种算法的参数设定

Table 3. Parameter setting of various algorithms.

	KLMS	KLMP	KFLP	KLLAD	P-KLLAD
η	0.1	0.1	0.1	0.1	0.1
p		1.2	0.9		1.1
a				5	3

1) 5 种算法在高斯分布噪声下的性能对比. 实验中令 α 稳定分布噪声参数为 $\alpha = 2$, $\delta = 0$, $\beta = 0$ 和 $\gamma = 0.1$, 即此时该噪声是高斯分布噪声. 5 种算法的 MSE 曲线如图 5 所示.

从图 5 可以看出, 在高斯噪声环境下, 本文的算法和其他算法一样都有良好的稳态性能, 但是本

文算法收敛速度比其他算法快, 说明本文算法具有更优的收敛性能. 其原因在于本文算法中的 p 范数和正常数 a 的主要作用就是用来提升该代价函数的陡峭程度, 从而提高算法的收敛性能.

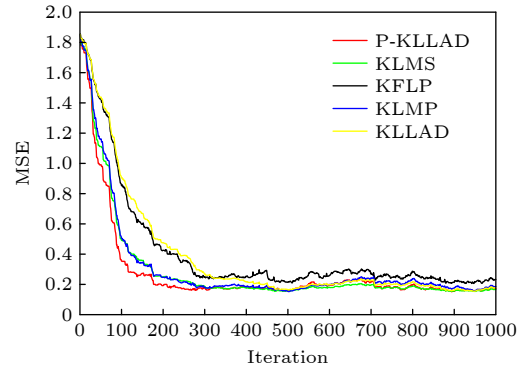


图 5 非线性系统辨识高斯噪声下的 MSE 曲线

Fig. 5. The MSE curve of nonlinear system identification under Gaussian noise.

2) 在脉冲噪声干扰下的性能比较. 实验中令 α 稳定分布噪声参数为 $\alpha = 1.3$, $\delta = 0$, $\beta = 0$ 和 $\gamma = 0.1$, 此时该噪声存在脉冲干扰. 实验结果如图 6 所示. 表 4 给出了各比较算法达到稳态时候的稳态误差的均值和标准偏差.

表 4 均值 \pm 标准偏差Table 4. Mean \pm standard deviation.

算法	MSE
KLMS	0.1952 ± 0.1909
KLMP	0.1921 ± 0.01339
KFLP	0.2311 ± 0.02468
KLLAD	0.1794 ± 0.0209
P-KLLAD	0.1792 ± 0.0130

从图 6 和表 4 可以看出在 α 稳定分布噪声下本文算法和 KLMP, KFLP, KLLAD 都能达到比较好的稳态误差, 但是本文算法的收敛速度都比其他几种算法都快, 并且对脉冲噪声有良好的鲁棒性. 这是由于最小对数绝对差准则可以抑制脉冲噪声的无限增长.

5.2 Mackey-Glass 混沌时间序列

Mackey-Glass 混沌时间方程由以下延迟微分方程生成:

$$\frac{dx(t)}{dt} = bx(t) + \frac{ax(t-c)}{1+x(t-c)^{10}}, \quad (25)$$

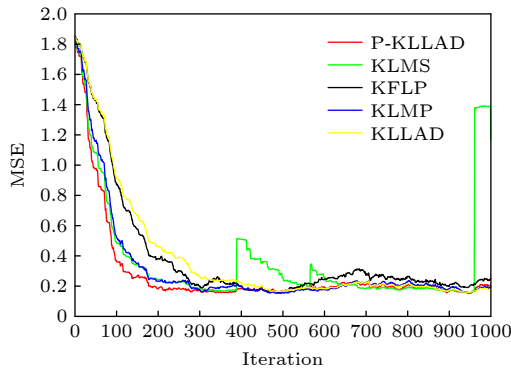
图 6 非线性系统辨识 α 稳定分布噪声下的 MSE 曲线

Fig. 6. The MSE curve of nonlinear system identification under α stably distributed noise.

其中, 参数设置为: $a = 0.2$, $b = 0.1$, $c = 10$. 时间序列通过一个 6 s 的采样周期进行离散. 根据 Takens 嵌入定理, 选择用最近的 7 个输入来预测当前的输入, 这个当前的输入就是期望信号. 这里将所有算法的参数设置为最优值, 其中 KLMS 和 KFLP 的步长分布为 0.2 和 0.07, KLMP, KLLAD 和 P-KLLAD 的步长均为 0.12, 其他参数设置同上述实验一致.

1) 设置 α 稳定分布噪声参数为 $\alpha = 2$, $\delta = 0$, $\beta = 0$ 和 $\gamma = 0.1$, 即此时该噪声是高斯分布噪声. 5 种算法性能曲线比较结果如图 7 所示. 表 5 是图 7 中各算法达到稳态时稳态误差的均值和标准偏差. 图 8 给出了基于本文算法的时间序列预测中的预测值与期望值的比较结果.

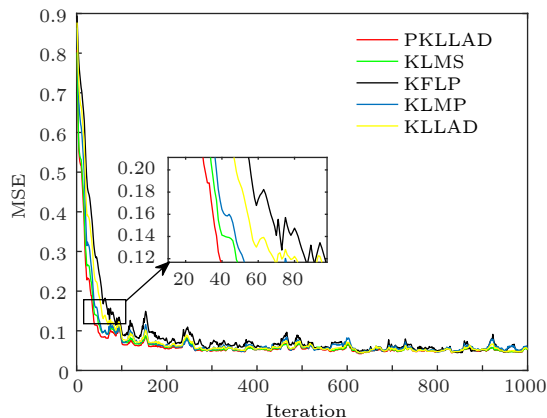


图 7 MG 混沌时间序列预测高斯噪声下的 MSE 曲线

Fig. 7. The MSE curve of MG chaotic time series prediction under Gaussian noise.

从图 7 可以看出, 在高斯噪声环境下, 本文的算法和 KLMS 算法都具有较快的收敛速度, 收敛性能优于 KLMP, KFLP 和 KLLAD 算法. 而从表 5

可以看出, 稳态时本文算法和 KLMS 算法的均值和标准偏差是最小的, 说明本文算法和 KLMS 都具有更优的收敛速度和稳态性能. 从图 8 可以看出本文算法对于 MG 混沌时间序列的预测值和期望值总体上基本保持一致.

2) 噪声存在脉冲异常值 (α 稳定分布噪声参数 $\alpha = 1.3$, $\delta = 0$, $\beta = 0$ 和 $\gamma = 0.1$) 时 5 种算法性能的比较. 结果如图 9 所示. 从图 9 可以看出, 存在脉冲噪声时, 本文算法比 KLMS 算法具有更好的

表 5 均值 \pm 标准偏差Table 5. Mean \pm standard deviation.

算法	MSE
KLMS	0.0467 ± 0.0058
KLMP	0.0485 ± 0.0086
KFLP	0.0507 ± 0.0118
KLLAD	0.0492 ± 0.0354
P-KLLAD	0.0423 ± 0.0050

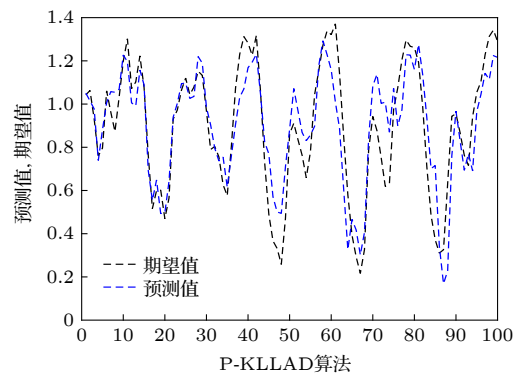


图 8 P-KLLAD 算法的预测值和期望值

Fig. 8. Predicted values of the P-KLLAD algorithm and target values.

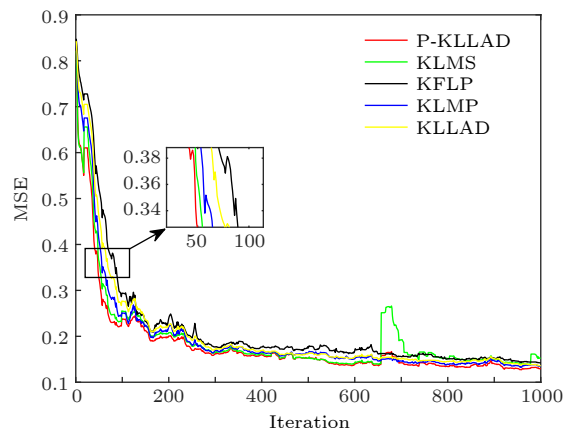
图 9 MG 混沌时间序列预测 α 稳定分布噪声下的 MSE 曲线

Fig. 9. The MSE curve of MG chaotic time series prediction under α stably distributed noise.

稳态性能和收敛速度, 并且对脉冲噪声有良好的鲁棒性.

综上, 本文的 P-KLLAD 算法虽牺牲了计算复杂度, 但在非线性系统辨识和 Mackey-Glass 混沌时间序列预测中不管是高斯噪声环境还是存在有小的脉冲噪声的情况下均具有良好的脉冲噪声抑制能力和较快的收敛速度.

6 结 论

针对 α 稳定分布噪声模型下的非线性自适应滤波问题, 本文中提出了一种基于 p 范数的核最小对数绝对差自适应滤波算法 P-KLLAD. 一方面利用最小对数绝对差准则来抑制 α 稳定分布噪声中的冲激干扰, 另一方面通过误差的 p 范数和一个正常数 a 来提高算法的收敛速度. 仿真结果表明, P-KLLAD 算法在 α 稳定分布噪声环境下的非线性系统辨识和 Mackey-Glass 混沌时间序列预测中在保证鲁棒性能的同时提高了收敛速度. 但是算法在 α 稳定分布噪声为大异常值的时候出现性能下降问题, 有待进一步研究完善. 针对于计算复杂度问题, 后续将进一步研究有效的优化方案来降低算法复杂度.

参考文献

- [1] Liu W F, Príncipe J C, Haykin S 2010 *Kernel Adaptive Filtering: A Comprehensive Introduction* (Hoboken, NJ, USA: John Wiley & Sons) pp16–32
- [2] Tian Z D, Gao X W, Shi T 2014 *Acta Phys. Sin.* **63** 70 (in Chinese) [田中大, 高宪文, 石彤 2014 物理学报 **63** 70]
- [3] Zhang W, Zhu J 2020 *Electronics*. **9** 940
- [4] Pauline S H, Samiappan D, Kumar R, Anand, A, Kar, A 2020 *Applied Acoustics*. **159** 107074
- [5] Shen L H, Chen J H, Ceng Z G, Jin J 2018 *Acta Phys. Sin.* **67** 030501 (in Chinese) [沈力华, 陈吉红, 曾志刚, 金健 2018 物理学报 **67** 030501]
- [6] Wang S Y, Shi C F, Qian G B, Wang W L 2018 *Acta Phys. Sin.* **67** 018401 (in Chinese) [王世元, 史春芬, 钱国兵, 王万里 2018 物理学报 **67** 018401]
- [7] Wu Q, Li Y, Xue W 2019 *Symmetry* **11** 1067
- [8] Wen, F 2013 *Electron. Lett.* **49** 1355
- [9] Jiao S B, Ren C, Huang W C, Liang Y M 2013 *Acta Phys. Sin.* **62** 210501 (in Chinese) [焦尚彬, 任超, 黄伟超, 梁炎明 2013 物理学报 **62** 210501]
- [10] Pelekanakis K, Chitre M 2014 *IEEE Wirel. Commun.* **13** 3183
- [11] Aalo V A, Ackie A, Mukasa C 2019 *Signal Process.* **154** 363
- [12] Liu W, Pokharel P P, Principe J C 2008 *IEEE Trans. Signal Process.* **56** 543
- [13] Seo B 2011 *Signal Process.* **91** 2623
- [14] Ma W, Duan J, Man W, Zhao H, Chen B 2017 *Eng. Appl. Artif. Intel.* **58** 101
- [15] Wu Z, Shi J, Xie Z, Ma W 2015 *Signal Process.* **117** 11
- [16] Zhao J Z, Jin M M 2017 *Appl. Res. Comp.* **34** 3308 (in Chinese) [赵知劲, 金明明 2017 计算机应用研究 **34** 3308]
- [17] Dong Q, Lin Y 2019 *Comp. Sci.* **046** 80 (in Chinese) [董庆, 林云 2019 计算机科学 **046** 80]
- [18] Huo Y L, Wang D F, Long X Q, Lian P J, Qi Y F 2021 *Acta Phys. Sin.* **70** 415 (in Chinese) [火元莲, 王丹凤, 龙小强, 连培君, 齐永锋 2021 物理学报 **70** 415]
- [19] Lin Y, Lei Y, Zeng J J 2016 *Appl. Electron. Tech.* **42** 78 (in Chinese) [林云, 雷洋, 曾俊俊 2016 电子技术应用 **42** 78]
- [20] Sayin M O, Vanli N D, Kozat S S 2014 *IEEE Trans. Signal Process.* **62** 4411

Kernel least logarithm absolute difference algorithm based on P-norm^{*}

Huo Yuan-Lian^{1)†} Tuo Li-Hua¹⁾ Qi Yong-Feng²⁾ Ding Rui-Bo¹⁾

1) (*College of Physics and Electronic Engineering, Northwest Normal University, Lanzhou 730000, China*)

2) (*College of Computer Science and Engineering, Northwest Normal University, Lanzhou 730000, China*)

(Received 14 June 2021; revised manuscript received 24 November 2021)

Abstract

The kernel adaptive filtering is an efficient and nonlinear approximation method which is developed in reproducing kernel Hilbert space (RKHS). Kernel function is used to map input data from original space to RKHS space, thus solving nonlinear problems is efficient. Impulse noise and non-Gaussian noise exist in the real application environment, and the probability density distribution of these noise characteristics shows a relatively heavy trailing phenomenon in the statistical sense. α stable distribution can be used to model this kind of non-Gaussian noise well. The kernel least mean square (KLMS) algorithms usually perform well in Gaussian noise, but the mean square error criterion only captures the second-order statistics of the error signal, this type of algorithm is very sensitive to outliers, in other words, it lacks robustness in α stable distribution noise. The kernel least logarithm absolute difference (KLLAD) algorithm can deal with outliers well, but it has the problem of slow convergence. In order to further improve the convergence speed of nonlinear adaptive filtering algorithm in α stable distributed noise background, a new kernel least logarithm absolute difference algorithm based on p-norm (P-KLLAD) is presented in this paper. The algorithm combining least logarithm absolute difference algorithm and p norm, on the one hand, the least logarithm difference criteria is ensure the algorithm to have good robustness in α stable distribution noise environment, and on the other hand, add p norm on the absolute value of error. The steepness of the cost function is controlled by p norm and a positive constant α to improve the convergence speed of the algorithm. The computer simulation results of Mackey-Glass chaotic time series prediction and nonlinear system identification show that this algorithm improves the convergence speed with good robustness, and the convergence speed and robustness better than the kernel least mean square algorithm, the kernel fractional lower power algorithm, the kernel least logarithm absolute difference algorithm and the kernel least mean p-norm algorithm.

Keywords: α stable distributed noise, kernel adaptive filtering algorithm, minimum logarithm absolute difference criterion, p norm

PACS: 84.30.Vn, 05.40.Ca

DOI: 10.7498/aps.71.20211124

^{*} Project supported by the National Natural Science Foundation of China (Grant No. 61561044).

[†] Corresponding author. E-mail: huoyuanlian@163.com