

LMS 和归一化 LMS 算法收敛门限与步长的确定¹

曾召华 刘贵忠 赵建平*

(西安交通大学电信学院 西安 710049)

*(中兴通讯股份有限公司西安研究所 西安 710065)

摘要 从 LMS 算法失调量的准确表达式出发, 根据输入信号特征值分布重新研究了 LMS, 归一化 LMS(Normalized LMS, NLMS) 算法收敛的必要条件, 推导出 LMS 和 NLMS 算法收敛的步长门限, 并分析了输入信号特征值分布、滤波器阶数对算法收敛步长门限的影响, 推导出满足性能失调下步长的自适应计算公式, 减小了应用 LMS, NLMS 算法时步长选取的盲目性. 与已有的算法相比, 具有计算简单、实用、自适应性能强, 同时可获得满意失调量的特点. 计算机模拟结果表明该方法的正确性.

关键词 LMS, 归一化 LMS, 稳态失调, 步长, 门限

中图分类号 TN911

1 引言

最小均方 (LMS) 算法和步长归一化的最小均方算法 (Normalized LMS, NLMS) 是自适应算法中应用广泛的一类算法^[1-7]. NLMS 使得算法在输入信号较大的情况下避免梯度噪声放大的干扰, 因而具有更好的收敛性能^[1-3]. 但是和 LMS 算法一样, 它的收敛速度和稳态误差也要受到步长的控制. 为克服常规 LMS 或 NLMS 自适应算法在收敛速率、跟踪速率与权失调噪声之间的要求上存在的较大矛盾, 人们发展了多种变步长 LMS 算法^[5-7], 其基本思想就是在跟踪阶段动态给定一个较大的步长, 使算法有较快的收敛速度; 然后随着收敛的加深逐渐减小步长来减小稳态误差. 然而这些算法都没有解决启动步长大到多大, 稳态收敛步长小到多小这些定量问题. 显然, 步长的确定是与根据实际要求给定的稳态失调量密切相关的, 并非启动步长越大越好、稳态收敛步长越小越好. 本文旨在研究 LMS, NLMS 算法的步长门限, 以及满足给定失调量的要求而确定的步长值, 为其所有改进算法提供理论依据.

长期以来, 人们一直认为 LMS 算法收敛的条件是其步长 μ 满足: $0 < \mu < 1/\lambda_{\max}(\lambda_{\max}$ 定义见 (3) 式)^[5-7]; NLMS 算法收敛的条件是其步长 η 满足: $0 < \eta < 2^{[1-3]}$. 文献 [4] 从全新的收敛概念角度出发, 给出了 LMS 算法收敛的新的步长门限, 并在保证权向量均值收敛和给定的稳态失调量情况下给出了一种步长的取值公式, 但计算较复杂, 同时随给定的稳态失调量增加时, 真实稳态失调量与给定的稳态失调量之间的误差急剧增加. 本文从 LMS 算法失调量的准确表达式出发, 重新研究了 LMS, NLMS 算法收敛的必要条件, 给出了它们收敛的步长门限, 统一了已有的结论. 分析了输入信号特征值分布对算法收敛步长门限的影响, 推导出在满足给定失调下步长计算公式, 具有含义清晰、计算简单、工程实用, 同时又满足给定失调量的要求.

2 LMS, NLMS 算法收敛的步长门限

2.1 步长门限的求解

标准 LMS 算法表达如下^[1,2]:

$$\left. \begin{aligned} e(j) &= d(j) - \mathbf{W}^T(j)\mathbf{X}(j) \\ \mathbf{W}(j+1) &= \mathbf{W}(j) + 2\mu e(j)\mathbf{X}(j) \end{aligned} \right\} \quad (1)$$

¹ 2002-05-23 收到, 2002-10-27 改回

国家教育部高等学校博士学科点专项科研基金 (No.2000069828)、陕西省教育厅自然科学专项基金 (No.01jk194)

式中 $\mathbf{W}(j)$, $\mathbf{X}(j)$, $d(j)$, $e(j)$ 分别为第 j 次迭代的滤波器权向量、输入信号向量、参考信号、误差信号, 其中向量都是 N 维列向量; μ 为算法步长; T 表转置。

令

$$\mu = \frac{\eta}{2\|\mathbf{X}(j)\|^2} = \frac{\eta}{2\text{tr}\mathbf{R}} \quad (2)$$

式中 $\text{tr}\mathbf{R}$ 为输入信号向量 $\mathbf{X}(j)$ 自相关函数 \mathbf{R} 的迹, 也表示输入信号向量 $\mathbf{X}(j)$ 的功率, η 为 NLMS 算法的步长。把 (2) 式中 μ 的表达式代入 (1) 式, 就可得到标准的 NLMS 算法。由此所有 NLMS 的研究都可通过 (2) 式与 LMS 联系起来。

LMS 算法的失调量定义为超量均方误差与最小均方之比, 其算法稳态失调量表示如下 [2]:

$$M = \sum_{i=1}^N \frac{\mu\lambda_i}{1-2\mu\lambda_i} \bigg/ \left(1 - \sum_{i=1}^N \frac{\mu\lambda_i}{1-2\mu\lambda_i} \right) \quad (3)$$

式中 λ_i 为输入信号向量 $\mathbf{X}(j)$ 的第 i 特征值 (也表示信号第 i 个分量的功率), 并对 $\forall i$, 都有 $\lambda_i > 0$ 。LMS 收敛还需满足

$$0 < \sum_{i=1}^N \frac{\mu\lambda_i}{1-2\mu\lambda_i} < 1 \quad (4)$$

由于步长恒大于零, 理论上 LMS 算法门限 μ^* 可由 $\sum_{i=1}^N \frac{\mu\lambda_i}{1-2\mu\lambda_i} = 1$ 解出。实际中 λ_i 是未知的, 必须探讨特定条件下 LMS 算法步长门限 μ^* 的取值, 然后根据 (2) 式, 得出 NLMS 算法步长门限 η^* 的取值。

定理 1 当信号各特征值不相等时, $\eta^* = 2/3$, $\mu^* = 1/(3\text{tr}\mathbf{R})$ 。

证明 设信号的 N 个特征值依次为 $\lambda_1, \lambda_2, \dots, \lambda_N$ 。不失一般性, 设 $\lambda_1 = \max\{\lambda_i, 1 \leq i \leq N\}$ 。根据常规 LMS 算法收敛的步长条件 [5-7]: $0 < \mu < 1/\lambda_{\max}$, 则有 $\mu\lambda_1 < 1$, 并对 $\forall i$, 都有 $\mu\lambda_i < 1$ 。由此 $\sum_{i=1}^N \frac{\mu\lambda_i}{1-2\mu\lambda_i} = \sum_{i=1}^N \mu\lambda_i \sum_{k=0}^{\infty} (2\mu)^k \lambda_i^k = \sum_{i=1}^N \lambda_i^{k+1} \sum_{k=0}^{\infty} 2^k \mu^{k+1} < (\frac{1}{N} \sum_{i=1}^N \lambda_i)^{k+1} \sum_{k=0}^{\infty} 2^k (N\mu)^{k+1} = \sum_{k=0}^{\infty} 2^k (N\mu\lambda_{av})^{k+1} = \frac{N\mu\lambda_{av}}{1-2N\mu\lambda_{av}} = \frac{\mu\text{tr}\mathbf{R}}{1-2\mu\text{tr}\mathbf{R}} < 1 \Rightarrow \mu < 1/(3\text{tr}\mathbf{R})$, 其中 $\lambda_{av} = \frac{1}{N} \sum_{i=1}^N \lambda_i = \frac{\text{tr}\mathbf{R}}{N}$, 为输入信号平均功率。

综合假设可得 $\mu^* = \min\{1/(3\text{tr}\mathbf{R}), 1/\lambda_1\} = 1/(3\text{tr}\mathbf{R})$ 。联系 (2) 式, 则可得 $\eta^* = 2/3$ 。

证毕

定理 2 当信号各特征值相等时, $\eta^* = 2/2(1+2/N)$, $\mu^* = 1/[(1+2/N)\text{tr}\mathbf{R}]$ 。

证明 由 $\lambda_1 = \lambda_2 = \dots = \lambda_N$, $\sum_{i=1}^N \frac{\mu\lambda_i}{1-2\mu\lambda_i} = 1$, 立即可得 $\mu\lambda/(1-2\mu\lambda) = 1/N$, 从而可以得出: $\mu^* = 1/[(N+2)\lambda] = 1/(\text{tr}\mathbf{R} + 2\lambda)$ 。联系 (2) 式, 则可得 $\eta^* = 2/(1+2/N)$ 。

证毕

综合定理 1、定理 2, 门限 μ^* , η^* 可表示为

$$\mu^* = \begin{cases} 1/(3\text{tr}\mathbf{R}), & \text{特征值不相等} \\ 1/(\text{tr}\mathbf{R} + 2\lambda), & \text{特征值相等} \end{cases} \quad (5)$$

$$\eta^* = \begin{cases} 2/3, & \text{特征值不相等} \\ 2/(1+2N), & \text{特征值相等} \end{cases} \quad (6)$$

2.2 μ^* , η^* 与信号特征值分布、滤波器阶数的关系分析

从 (5) 式可看出: (1) LMS 算法步长门限始终受输入信号向量总功率 $\text{tr}\mathbf{R}$ 的影响, 其门限随 $\text{tr}\mathbf{R}$ 的增大而减小。(2) 在信号平均功率 λ_{av} 一定时, 其门限表现为随滤波器阶数 N 的增大

而减小。

从 (6) 式可看出: (1) $\text{tr}\mathbf{R}$ 对 NLMS 算法步长门限没有影响。(2) 特征值不相等时, 滤波器阶数 N 对门限取值没有影响, 为定值 $2/3$ 。(3) 特征值相等时, 滤波器阶数 N 对门限取值有影响。当 $N=1$ 时, $\eta^* = 2/3$; $N \rightarrow \infty$ 时, $\eta^* \rightarrow 2$; 表现为增加滤波器阶数 N 可获得更大的步长门限。

综合 (5), (6) 式还可得出以下结论: (1) 特征值相等时的门限取值始终大于特征值不相等时的门限取值 ($N=1$ 时相等), 这说明输入信号 $\mathbf{X}(j)$ 的特征值分布越均匀, LMS 算法和 NLMS 算法步长可有更大的收敛门限, 在保证收敛条件下可以有更大的收敛速度。因此对于 LMS 或 NLMS 算法, 对其输入信号 $\mathbf{X}(j)$ 先进行去相关预处理, 可以提高收敛速度。(2) NLMS 算法比 LMS 算法要稳定一些, 其门限不受 $\text{tr}\mathbf{R}$ 的影响。

3 收敛步长的确定与数值分析

从方差和失调的定义出发, 文献 [3] 指出无偏估计不能保证稳态方差有界, 而有界方差估计也不能保证稳态失调量有界。与 LMS 算法一样, NLMS 算法在理论上也是统计平均收敛于维纳权向量; 在实际工程实践中常以稳态失调量的大小作为评价算法收敛的依据, 因此步长的选择就必然在满足收敛门限下, 同时依一定的稳态失调量的选取为前提。参照文献 [4] 中 LMS 算法收敛的定义, NLMS 算法收敛可定义如下:

定义 1 对于事先给定的失调量 M' 和正数 δ , 如果存在正数 η_1 和正整数 K , 使得 NLMS 算法中步长 η 和迭代次数 j 满足 $0 < \eta < \eta_1$ 和 $j > K$ 时, 存在: (1) $\|\mathbf{W}(j) - \mathbf{W}_0\| < \delta$; (2) $M(j) < M'$; 则称 NLMS 算法是收敛的。其中 \mathbf{W}_0 为维纳权向量, $M(j)$ 为动态失调量。LMS 算法收敛的定义可参考文献 [4]。

引入给定的失调量 M' , (3) 式可改写为

$$\sum_{i=1}^N \frac{\mu\lambda_i}{1-2\mu\lambda_i} = \frac{M'}{1+M'} \quad (7)$$

对 LMS 算法, 步长的选取应使 μ 满足 $0 < \mu < \mu^*$, 其中 μ^* 由 (5) 式确定, 在此前提下, 由 (7) 式确定 μ 的大小。

对 NLMS 算法, 步长的选取应使 η 满足 $0 < \eta < \eta^*$, 其中 η^* 由 (6) 式确定, 在此前提下, 由 (2) 式, (7) 式联合确定 η 的大小。

3.1 特征值都相等时

由 (7) 式可知, 在特征值相等时, 有 $N\mu\lambda/(1-2\mu\lambda) = M'/(1+M')$, 立即可得:

$$\mu = 1/[(1+2/N+1/M')\text{tr}\mathbf{R}] \quad (8)$$

联系 (2) 式可得:

$$\eta = 2/(1+2/N+1/M') \quad (9)$$

3.2 特征值有一个极大, 其它都较小时

仿照上一小节, 则由 (7) 式可得:

$$\sum_{i=1}^N \frac{\mu\lambda_i}{1-2\mu\lambda_i} \approx \frac{\mu \sum_{i=1}^N \lambda_i}{1-2\mu \sum_{i=1}^N \lambda_i} = \frac{\mu \text{tr}\mathbf{R}}{1-2\mu \text{tr}\mathbf{R}} = \frac{M'}{1+M'} \quad (10)$$

由 (10) 式可得

$$\mu = 1/[(3 + 1/M')\text{tr}\mathbf{R}] \quad (11)$$

联系 (2) 式可得

$$\eta = 2/(3 + 1/M') \quad (12)$$

3.3 特征值分布不可知时

当特征值分布不可知, 则 (7) 式展开后 μ 的高次方程式的最小正根即是要求解的步长. 文献 [4] 采用高次方程的下限渐进解逼近真值 μ_0 , 随着逼近方程次数 p 的增大, 其精度增大; 但只有当 $p = 1$ 时, 其计算公式 (18) 式才可满足工程计算, $p > 1$ 时, 其计算公式 (19) 式非常复杂, 或者说不可能; 同时当失调量 M' 在 0.1 至 1.0 范围内变化时, 随着失调量的增大, 按其计算所产生的误差逐渐增大. 因此必须寻求一种新的计算方法.

考察 3.1 节的结论, 它是在一种非常理想情况下给出的 (特征值相等); 3.2 节的结论, 它是在一种极端恶劣情况下给出的 (特征值中有一个极大). 在实际情况中, 一般为上述两种情况的一种折中, 因此可对 3.1 节、3.2 节给出的结果作加权平均, 从而给出信号特征值一般分布时, 在满足给定失调量下收敛步长的表达式:

$$\mu = \left(\alpha \cdot \frac{1}{1 + 2/N + 1/M'} + \beta \cdot \frac{1}{3 + 1/M'} \right) / (2\text{tr}\mathbf{R}) \quad (13)$$

$$\eta = \alpha \cdot \frac{1}{1 + 2/N + 1/M'} + \beta \cdot \frac{1}{3 + 1/M'} \quad (14)$$

式中 α, β 为加权系数, 满足 $\alpha \cdot \beta \geq 0, \alpha + \beta = 1$.

3.4 步长数值分析

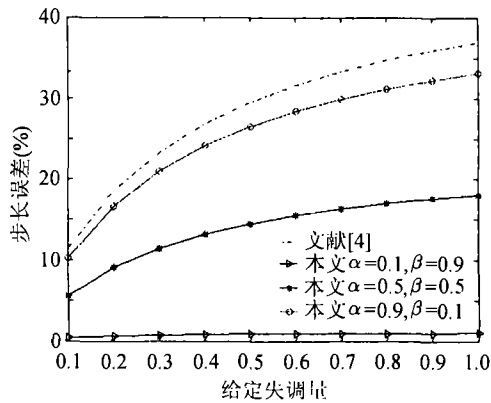
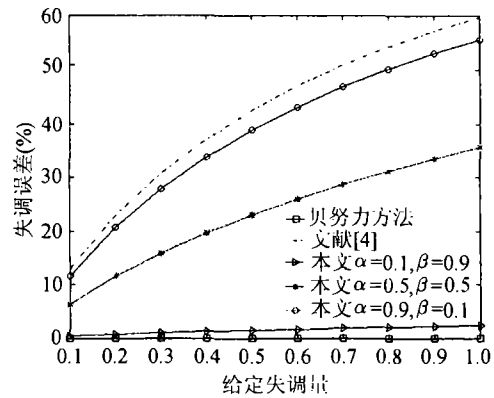
由于 $M' \geq 0$, 由 (8), (9), (11), (12) 式可知, 当 $M' \rightarrow 0$ 时, $\mu \rightarrow 0, \eta \rightarrow 0$; 当 $M' \rightarrow \infty$ 时, $\mu \rightarrow \mu^*, \eta \rightarrow \eta^*$. 显然确定的步长在满足失调量要求下, 又同时满足算法的收敛步长门限. 这说明期望失调量的大小极大地影响步长的取值, 也进一步说明, 工程实践中, 根据需要确定稳态失调量, 然后确定算法步长的意义.

(13), (14) 式具有含义清晰、计算简单的特点. 在工程实践中, 可根据对输入信号的了解程度, 来自适应调节 α, β . 当已对输入信号进行去相关预处理, 为了加速算法收敛, 可使 $\alpha \rightarrow 1, \beta \rightarrow 0$; 当输入信号的特征值分布较发散时, 为了保证算法收敛同时满足给定的失调量的要求, 可使 $\alpha \rightarrow 0, \beta \rightarrow 1$; 在大部分情况下, 可令 $\alpha = \beta = 0.5$.

4 计算机仿真实验

实验 给定失调量 M' 在 0.1 至 1 变化. 按文献 [1] 附录 A 的方法产生随机数扩大 20 倍作为 λ_i ($\lambda_1 = 9.029151, \lambda_2 = 14.401310, \lambda_3 = 16.1166510, \lambda_4 = 14.174570$), $N = 4$. 用贝努利法求 (7) 式的最小根为 μ_0 , 令 α 分别等于 0.1, 0.5, 0.9 按 (13) 式求得的步长为 μ_1, μ_2, μ_3 , 用文献 [4](18) 式求得步长为 μ_4 , 令 $\Delta\mu_i = (\mu_i - \mu_0)/\mu_0, i = 1 \sim 4$, 表示步长误差百分比. 将 μ_0 至 μ_4 代入 (3) 式分别求得实际失调量 M_0 至 M_4 , 令 $\Delta M_i = (M_i - M')/M', i = 0 \sim 4$, 表示失调误差百分比. 图 1 为步长误差随 M' 变化关系曲线, 图 2 为失调误差随 M' 变化关系曲线.

从图 1, 图 2 可以看出, 采用本文 (13) 式计算步长和相应的失调量, 它们的误差均小于文献 [4]. 特别是当 $\alpha = 0.1, \beta = 0.9$ 时, 其误差已小于 5%, 而计算复杂度远远小于贝努利方法.

图 1 步长误差随给定失调量 M' 变化关系曲线图 2 失调误差随 M' 变化关系曲线

5 结 论

通过本文的研究, 可得出如下结论:

(1) LMS, NLMS 算法收敛的步长门限由 (5), (6) 式确定。门限大小与输入信号向量的特征值分布直接相关, 分布越均匀, 门限值越大。因此对输入信号进行去相关预处理, 有利于提高步长门限。

(2) 可根据对输入信号的了解程度, 给定 (13), (14) 式中 α, β 参数来确定步长。输入信号特征值分布越均匀, α 可取越小的值, 有利于减小失调误差; 在对信号一无所知时, 可令 $\alpha = \beta = 0.5$ 。

(3) (13), (14) 式实际上提供了一种自适应控制步长的思想。在算法启动阶段, 为了获取较快的启动速度, 可令 $\alpha = 1, \beta = 0$; 在算法收敛阶段, 为了使稳态失调量不超过给定的 M' , 可令 $\alpha = 0, \beta = 1$ 。

(4) 对于其它的变步长算法, 在启动阶段, 为保证算法收敛, 同时有大的启动速度、避免大的振荡, 其最大值要受 (5), (6) 式给定的步长门限制, 具体取值可参考 (13), (14) 式在 $\alpha = 1, \beta = 0$ 时给定的步长值; 在算法收敛阶段, 为保证满足给定的失调量的要求, 同时又具有一定的跟踪能力, 以防止信道参数的突然改变, 其最小值可取 (13), (14) 式在 $\alpha = 0, \beta = 1$ 时的步长值。

参 考 文 献

- [1] B. Widrow, Adaptive Signal Processing[M], NJ: Prentice-Hall Inc, 1985, Chapter 3.
- [2] C. F. N. Cowan, Adaptive Filters[M], NJ: Prentice-Hall Inc, 1985, Chapter 4.
- [3] Jianfeng Liu, A novel adaptation scheme in the NLMS algorithm for echo cancellation[J], IEEE Signal Processing Letters, 2001, 8(1), 20-22.
- [4] 裴炳南, LMS 算法的收敛与步长选取 [J], 通信学报, 1994, 15(5), 106-111.
- [5] 曾召华, 韦力, 刘少亭, 一种改进的瞬变步长 LMS 自适应滤波算法 [J], 电子科学学刊, 1998, 20(4), 566-569.
- [6] R. H. Kwong, E. W. Johnston, A variable step size LMS algorithm[J], IEEE Trans. on Signal Processing, 1992, 40(7), 1633-1642.
- [7] K. Mayyas, T. Aboulnasr, A robust variable step size LMS-type algorithm: analysis and simulations[J]. IEEE Trans. on Signal Processing, 1997, 45(3), 631-639.

DETERMINING OF CONVERGENT THRESHOLD AND
STEP-SIZE FOR LMS AND
NORMALIZED LMS ALGORITHM

Zeng Zhaohua Liu Guizhong Zhao Jianping*

(School of Electron. & Info. Eng., Xi'an Jiaotong University, Xi'an 710049, China)

**(Zhongxing Telecom. Equipment Corp., Xi'an Institute, Xi'an 710065, China)*

Abstract Based on eigenvalue distributing of input signal, the convergent necessary conditions for LMS and Normalized LMS(NLMS) algorithm are again researched through the accurate analysis of the misadjustment. To avoid blindness for applying LMS and NLMS algorithm, the convergent threshold and the simple adaptive calculating formula for the step-size of them are proposed. The influence of eigenvalue distributing of input signal and order of filter on the threshold of step-size is also analyzed. Compared with an existing algorithm, the characters of lower computational complexity, practicality and stronger adaptive are shown and the satisfied misadjustment is achieved by adopting the presented method for calculating the step-size.

Key words Least Mean Square(LMS), Normalized LMS, Stead-state misadjustment, Step-size, Threshold

曾召华: 男, 1972年生, 博士后, 目前研究方向: 自适应信号处理技术及其在 CDMA 多用户检测算法中的应用.
刘贵忠: 男, 1962年生, 教授, 博士生导师, 目前研究方向: 非平稳信号处理及其在图像、视频与通信领域中的应用.
赵建平: 男, 1970年生, 博士, 中兴通讯西安研究所总工, 目前研究方向: 3G 关键技术及其工程实现.