

不相关高斯输入情况下 LMS 滤波器的收敛性分析*

李琳 路军 张尔扬

(国防科技大学电子科学与工程学院)

摘要: 在分析最小均方自适应滤波器(LMS AF)均方误差(MSE)的收敛性时, 文献常用统计自相关矩阵代替瞬时自相关矩阵以简化分析, 由此得出的收敛条件比较粗糙。本文指出: 不相关高斯输入情况下, 无需如上简化, 可依据高斯阶矩因式分解定理得到确切的 MSE 收敛条件, 相应的失调表式能更准确地预报失调。

关键词: LMS 算法; 均方误差; 高斯分布; 独立性假定

Convergence Analysis of Adaptive Filter with Uncorrelated Gaussian Input

Li Lin Lu Jun Zhang Eryang

(Institute of Electronic Science & Engineering, NUDT)

Abstract: When analyzing the convergence behavior of the mean squared error (MSE) of the least mean square (LMS) adaptive filter, many references often replace the instantaneous correlation matrix of the tap inputs with statistical correlation matrix in order to simplify analysis. It is clear that the condition derived from this method is coarse. In fact, for uncorrelated Gaussian input, the accurate necessary and sufficient conditions for the convergence of the MSE can be obtained by Gaussian moment factoring theorem, and the accurate misadjustment can be predicted by the corresponding misadjustment expression.

Key words: least mean square (LMS) algorithm; mean squared error (MSE); gaussian distribution; independence assumption

1 引言

自 1960 年 Widrow 提出 LMS 自适应算法以来, 大量文献对其收敛性能, 稳态性能等等作了分析。在分析最小均方自适应滤波器(LMS AF)均方误差(MSE)收敛性时, 往往用统计自相关矩阵代替瞬时自相关矩阵以简化分析^{[1][2]}, 由此得出的 MSE 收敛条件较为粗糙。对于不相关高斯输入情况, 不必作如上近似, 可通过矢量形式的高斯阶矩因式分解定理得到更确切的 MSE 收敛条件以及对自适应系统失调的预报更为准确的失调表式。

本文约定大写字母表示常量, 小写字母表示标量, 矢量用字母上带箭头表示, 上标 T 表示矩阵、矢量的转置。本文所有矢量均为列矢量。

2 预备知识

自适应系统根据 k 时刻的期望信号 $d(k)$ 与滤波器输出 $\bar{w}(k)^T \bar{x}(k)$ 之间的误差信号 $e(k) = d(k) - \bar{w}(k)^T \bar{x}(k)$ 来调整滤波器权值 $\bar{w}(k)$ 以逼近 $d(k)$ 。 $d(k)$ 用多线性回归模型建模为:

$$d(k) = \bar{w}_{opt}^T \bar{x}(k) + \varepsilon(k) \quad (1)$$

其中 \bar{w}_{opt} 为最优(最小均方(MMSE)意义下)权系数, $\varepsilon(k)$ 为零均值, 方差为 $E[\varepsilon^2(k)]$ 的序列。显然 $E[\varepsilon^2(k)]$ 等于最小均方误差 ξ_{min} 。

根据维纳滤波理论: $\varepsilon(k)$ 与输入 $\bar{x}(k)$ 正交, 即 $E[\varepsilon(k) \bar{x}(k)] = 0$ 。又 $\varepsilon(k)$ 零均值, 进而两者不相关。如果

收稿日期: 2003 年 5 月 27 日; 修回日期: 2003 年 8 月 11 日

*JS63 空间微波技术国防科技重点室基金试点项目(项目编号: 2000JS63.3.1.KG0111)

它们还是高斯过程, 则两者统计独立。

LMS 算法可描述为:

$$e(k) = d(k) - \bar{w}(k)^T \bar{x}(k) \quad (2)$$

$$\bar{w}(k+1) = \bar{w}(k) + \mu \bar{x}(k) e(k) \quad (3)$$

其中 μ 为收敛因子。

由于 $\bar{x}(k)$ 的自相关矩阵 $R = E\{\bar{x}(k) \bar{x}(k)^T\}$ 为对称矩阵, 那么存在正交矩阵 U 使得 $URU^T = \Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, 这里 $\lambda_1, \lambda_2, \dots, \lambda_n$ 为输入自相关矩阵 R 的特征值。

LMS 算法的收敛性分析包括权矢量均值 $E[\bar{w}(k)]$ 以及均方误差 $E[e^2(k)]$ 的收敛, 分析当中要采用著名的独立性假定^[1]。本文主要分析均方误差的收敛性。

将(1)式代入(2)式有:

$$e(k) = \varepsilon(k) - (\bar{w}(k) - \bar{w}_{opt})^T \bar{x}(k) \quad (4)$$

将(4)式代入(3)式得到:

$$\bar{w}(k+1) = \bar{w}(k) + \mu \bar{x}(k) \{\varepsilon(k) - (\bar{w}(k) - \bar{w}_{opt})^T \bar{x}(k)\} \quad (5)$$

称 $\bar{w}(k) - \bar{w}_{opt}$ 为权误差矢量。

在(5)式两端减去 \bar{w}_{opt} 后, 再左乘矩阵 U 有:

$$U(\bar{w}(k+1) - \bar{w}_{opt}) = U(\bar{w}(k) - \bar{w}_{opt}) + \mu U \bar{x}(k) \{\varepsilon(k) - \bar{x}(k)^T (\bar{w}(k) - \bar{w}_{opt})\} \quad (6)$$

记 $\bar{v}(k) = U(\bar{w}(k) - \bar{w}_{opt})$, 并带入(6)式得到:

$$\begin{aligned} \bar{v}(k+1) &= \bar{v}(k) + \mu U \bar{x}(k) \{\varepsilon(k) - \bar{x}(k)^T (\bar{w}(k) - \bar{w}_{opt})\} \\ &= \bar{v}(k) - \mu U \bar{x}(k) \bar{x}(k)^T U^T \bar{v}(k) + \mu \varepsilon(k) U \bar{x}(k) \end{aligned} \quad (7)$$

下面给出采用近似方法分析得到的 MSE 收敛条件; 不相关高斯输入情况下, 不作近似得到的收敛条件。

3 近似方法分析 MSE 的收敛性以及失调表式

3.1 MSE 收敛性分析

k 时刻的均方误差为:

$$\begin{aligned} \xi(k) &= E\{e^2(k)\} \\ &= E\{[\varepsilon(k) - (\bar{w}(k) - \bar{w}_{opt})^T \bar{x}(k)]^2\} \\ &= E\{e^2(k)\} - 2E\{\varepsilon(k)(\bar{w}(k) - \bar{w}_{opt})^T \bar{x}(k)\} \\ &\quad + E\{(\bar{w}(k) - \bar{w}_{opt})^T \bar{x}(k) \bar{x}(k)^T (\bar{w}(k) - \bar{w}_{opt})\} \end{aligned} \quad (8)$$

根据(5)式知 $\bar{w}(k)$ 只与 k 时刻以前的 $\bar{x}(i)$ ($i=k-1, k-2, \dots$) 有关, 再根据独立性假定: 不同时刻的输入矢量 $\bar{x}(k)$ 互相独立, 从而 $\bar{w}(k)$, $\bar{v}(k)$ 与 $\bar{x}(k)$, $\varepsilon(k)$ 独立。又零均值序列 $\varepsilon(k)$ 与 $\bar{x}(k)$ 不相关, 则(8)式第二项为零。注意到 $(\bar{w}(k) - \bar{w}_{opt})^T \bar{x}(k) \bar{x}(k)^T (\bar{w}(k) - \bar{w}_{opt})$ 为标量, 它的迹就等于本身。利用这些条件可将(8)式化简为:

$$\begin{aligned} \xi(k) &= \xi_{\min} + E\{(\bar{w}(k) - \bar{w}_{opt})^T \bar{x}(k) \bar{x}(k)^T (\bar{w}(k) - \bar{w}_{opt})\} \\ &= \xi_{\min} + E\{tr\{(\bar{w}(k) - \bar{w}_{opt})^T \bar{x}(k) \bar{x}(k)^T (\bar{w}(k) - \bar{w}_{opt})\}\} \\ &= \xi_{\min} + E\{tr\{(\bar{w}(k) - \bar{w}_{opt})(\bar{w}(k) - \bar{w}_{opt})^T \bar{x}(k) \bar{x}(k)^T\}\} \\ &= \xi_{\min} + tr\{E\{(\bar{w}(k) - \bar{w}_{opt})(\bar{w}(k) - \bar{w}_{opt})^T \bar{x}(k) \bar{x}(k)^T\}\} \\ &= \xi_{\min} + tr\{E\{(\bar{w}(k) - \bar{w}_{opt})(\bar{w}(k) - \bar{w}_{opt})^T\} E\{\bar{x}(k) \bar{x}(k)^T\}\} \\ &= \xi_{\min} + tr\{(U^T C(k) U)(U^T \Lambda U)\} \\ &= \xi_{\min} + tr\{U^T C(k) \Lambda U\} \\ &= \xi_{\min} + tr\{C(k) \Lambda U U^T\} \\ &= \xi_{\min} + tr\{C(k) \Lambda\} \\ &= \xi_{\min} + \bar{\lambda}^T \bar{\sigma}(k) \end{aligned} \quad (9)$$

其中 $\bar{\sigma}(k)$ 为权误差矢量的协方差矩阵 $C(k) = E\{\bar{v}(k) \bar{v}(k)^T\}$ 主对角线上元素组成的矢量, 特征值矢量 $\bar{\lambda} = \Lambda \bar{1} = (\lambda_1, \lambda_2, \dots, \lambda_n)^T$, $\bar{1}$ 为 n 维全 1 列矢量。推导当中用到矩阵迹的性质 $tr(AB) = tr(BA)$ 。

由(9)式知: 只要 $\bar{\sigma}(k)$ 收敛, 均方误差 $\xi(k)$ 就收敛且收敛到 $\xi_{\min} + \bar{\lambda}^T \bar{\sigma}(\infty)$, 从而可将均方误差的收敛问题归结为 $\bar{\sigma}(k)$ 的收敛问题。

为简化分析, 文献用统计自相关矩阵 R 代替瞬时自相关 $\bar{x}(k) \bar{x}(k)^T$ ^{[1][2]}, 则(7)式变为:

$$\bar{v}(k+1) = \bar{v}(k) - \mu \Lambda \bar{v}(k) + \mu \varepsilon(k) U \bar{x}(k) \quad (10)$$

进而:

$$\begin{aligned} \bar{v}(k+1) \bar{v}(k+1)^T &= [\bar{v}(k) - \mu \Lambda \bar{v}(k) + \mu \varepsilon(k) U \bar{x}(k)] [\bar{v}(k) - \mu \Lambda \bar{v}(k) + \mu \varepsilon(k) U \bar{x}(k)]^T \\ C(k+1) &= E\{\bar{v}(k+1) \bar{v}(k+1)^T\} \\ &= E\{\bar{v}(k) \bar{v}(k)^T\} - \mu E\{\bar{v}(k) \bar{v}(k)^T\} \Lambda - \mu \Lambda E\{\bar{v}(k) \bar{v}(k)^T\} \\ &\quad + \mu^2 \Lambda E\{\bar{v}(k) \bar{v}(k)^T\} \Lambda + \mu^2 E\{\varepsilon^2(k)\} \Lambda \end{aligned} \quad (11)$$

化简当中用到了 $\varepsilon(k)$ 是零均值的独立序列, 且 $\varepsilon(k)$, $\bar{x}(k)$, $\bar{v}(k)$ 互相独立的性质。

(11)式可整理为:

$$C(k+1) = C(k) - \mu C(k) \Lambda - \mu \Lambda C(k) + \mu^2 \Lambda C(k) \Lambda + \mu^2 \xi_{\min} \Lambda \quad (12)$$

由(12)式有:

$$\bar{\sigma}(k+1) = (I - 2\mu \Lambda + \mu^2 \Lambda^2) \bar{\sigma}(k) + \mu^2 \xi_{\min} \bar{\lambda} \quad (13)$$

显然, $\bar{\sigma}(k)$ 收敛当且仅当 $I - 2\mu \Lambda + \mu^2 \Lambda^2$ 的特征值在单位圆内, 即 $|1 - 2\mu \lambda_i + \mu^2 \lambda_i^2| < 1$, 也即 $0 < \mu < 2/\lambda_i$ ($i=1, 2, \dots, n$), 因此 MSE 收敛的充要条件为 $0 < \mu < 2/\lambda_{\max}$, 其中 λ_{\max} 为 R 的最大特征值。

3.2 失调表式

对(13)式两边取极限得到:

$$\begin{aligned} \bar{\sigma}(\infty) &= \mu^2 \xi_{\min} (I - (I - 2\mu \Lambda + \mu^2 \Lambda^2))^{-1} \bar{\lambda} \\ &= \mu^2 \xi_{\min} (2\mu \Lambda - \mu^2 \Lambda^2)^{-1} \bar{\lambda} \end{aligned} \quad (14)$$

根据(9)式有

$$\begin{aligned} \xi_{\infty} &= \xi_{\min} + \bar{\lambda}^T \bar{\sigma}(\infty) \\ &= \xi_{\min} + \mu^2 \xi_{\min} \bar{\lambda}^T (2\mu \Lambda - \mu^2 \Lambda^2)^{-1} \bar{\lambda} \end{aligned} \quad (15)$$

从而失调为

$$M_{LMS} = (\xi_{\infty} - \xi_{\min}) / \xi_{\min} \quad (16)$$

$$\begin{aligned} &= \mu^2 \bar{\lambda}^T (2\mu\Lambda - \mu^2 \Lambda^2)^{-1} \bar{\lambda} \\ &= \sum_{i=1}^n \frac{\mu \lambda_i}{2 - \mu \lambda_i} \end{aligned} \quad (17)$$

4 不相关高斯输入情况下 MSE 的收敛性分析以及失调表式

4.1 MSE 收敛性分析

不作如(10)式那样的简化, 直接由(7)式有:

$$\begin{aligned} E\{\bar{v}(k+1)\bar{v}(k+1)^T\} &= E\{\bar{v}(k)\bar{v}(k)^T\} - \mu E\{\bar{v}(k)\bar{v}(k)^T U \bar{x}(k)\bar{x}(k)^T U^T\} \\ &\quad - \mu E\{U \bar{x}(k)\bar{x}(k)^T U^T \bar{v}(k)\bar{v}(k)^T\} \\ &\quad + \mu^2 E\{U \bar{x}(k)\bar{x}(k)^T U^T \bar{v}(k)\bar{v}(k)^T U \bar{x}(k)\bar{x}(k)^T U^T\} \\ &\quad + \mu^2 E\{e^2(k)\} E\{U \bar{x}(k)\bar{x}(k)^T U^T\} \end{aligned}$$

因此

$$\begin{aligned} C(k+1) &= C(k) - \mu C(k)\Lambda - \mu \Lambda C(k) \\ &\quad + \mu^2 E\{U \bar{x}(k)\bar{x}(k)^T U^T \bar{v}(k)\bar{v}(k)^T U \bar{x}(k)\bar{x}(k)^T U^T\} + \mu^2 \xi_{\min} \Lambda \end{aligned} \quad (18)$$

对于零均值的高斯矢量 \bar{x} , 有如下的高斯阶因式分解定理:

$$E\{\bar{x}\bar{x}^T A \bar{x}\bar{x}^T\} = E\{\bar{x}\bar{x}^T\} E\{A \bar{x}\bar{x}^T\} + E\{\bar{x}\bar{x}^T A^T\} E\{\bar{x}\bar{x}^T\} + E\{\bar{x}\bar{x}^T\} \text{tr}\{E\{\bar{x}\bar{x}^T A^T\}\}$$

在输入 $\bar{x}(k)$ 服从高斯分布情况下, 可用 $U \bar{x}(k)$, $\bar{v}(k)$ $\bar{v}(k)^T$ 代替上式中的 \bar{x} 以及矩阵 A 得到^[5]:

$$\begin{aligned} &E\{U \bar{x}(k)\bar{x}(k)^T U^T \bar{v}(k)\bar{v}(k)^T U \bar{x}(k)\bar{x}(k)^T U^T\} \\ &= E\{(U \bar{x}(k))(U \bar{x}(k))^T\} E\{(\bar{v}(k)\bar{v}(k)^T) U \bar{x}(k)\bar{x}(k)^T U^T\} \\ &= E\{(U \bar{x}(k))(U \bar{x}(k))^T\} E\{(\bar{v}(k)\bar{v}(k)^T) (U \bar{x}(k))(U \bar{x}(k))^T\} \\ &\quad + E\{(U \bar{x}(k))(U \bar{x}(k))^T\} E\{(\bar{v}(k)\bar{v}(k)^T)^T\} E\{(U \bar{x}(k))(U \bar{x}(k))^T\} \\ &\quad + E\{(U \bar{x}(k))(U \bar{x}(k))^T\} \text{tr}\{E\{(U \bar{x}(k))(U \bar{x}(k))^T (\bar{v}(k)\bar{v}(k)^T)^T\}\} \} \\ &\text{又 } \bar{v}(k) \text{ 与 } \bar{x}(k) \text{ 统计独立, 所以} \end{aligned}$$

$$E\{U \bar{x}(k)\bar{x}(k)^T U^T \bar{v}(k)\bar{v}(k)^T U \bar{x}(k)\bar{x}(k)^T U^T\} = 2\Lambda C(k)\Lambda + \Lambda \text{tr}\{\Lambda C(k)\} \quad (19)$$

将(19)式代入(18)式得到:

$$C(k+1) = C(k) - \mu C(k)\Lambda + \Lambda C(k) + 2\mu^2 \Lambda C(k)\Lambda + \mu^2 \text{tr}\{\Lambda C(k)\}\Lambda + \mu^2 \xi_{\min} \Lambda$$

记 $C_{ii}(k)$ 为 $C(k)$ 对角线上的第 i 个元素, 由上式可得:

$$C_{ii}(k+1) = \gamma_{ii} C_{ii}(k) + \mu^2 \lambda_i \sum_{j=1}^n \lambda_j C_{jj}(k) + \mu^2 \xi_{\min} \lambda_i$$

其中 $\gamma_{ii} = 1 - 2\mu\lambda_i + 2\mu^2\lambda_i^2$ ($i=1, 2, \dots, n$), 进而有:

$$\begin{aligned} \sigma(k+1) &= (I - 2\mu\Lambda + 2\mu^2\Lambda^2)\sigma(k) + \mu^2(\bar{\lambda}^T \sigma(k))\bar{\lambda} + \mu^2 \xi_{\min} \bar{\lambda} \\ &= \{I - 2\mu\Lambda + 2\mu^2\Lambda^2 + \mu^2 \bar{\lambda} \bar{\lambda}^T\} \sigma(k) + \mu^2 \xi_{\min} \bar{\lambda} \\ &= F \sigma(k) + \mu^2 \xi_{\min} \bar{\lambda} \end{aligned} \quad (20)$$

其中 $F = I - 2\mu\Lambda + 2\mu^2\Lambda^2 + \mu^2 \bar{\lambda} \bar{\lambda}^T$.

$\bar{\sigma}(k)$ 收敛当且仅当 F 的全部特征值都在单位圆内。 F 的特征值满足 $|F - \zeta I| = 0$, 又

$$F - \zeta I = I - 2\mu\Lambda + 2\mu^2\Lambda^2 + \mu^2 \bar{\lambda} \bar{\lambda}^T - \zeta I$$

$$= (I - 2\mu\Lambda + 2\mu^2\Lambda^2 - \zeta I) \{I + \mu^2(I - 2\mu\Lambda + 2\mu^2\Lambda^2 - \zeta I)^{-1} \bar{\lambda} \bar{\lambda}^T\}$$

记 $\Lambda_{\gamma} = I - 2\mu\Lambda + 2\mu^2\Lambda^2 = \text{diag}(\gamma_{11}, \gamma_{22}, \dots, \gamma_{nn})$, 从而:

$$\begin{aligned} |F - \zeta I| &= |\Lambda_{\gamma} - \zeta I| |I + \mu^2(\Lambda_{\gamma} - \zeta I)^{-1} \bar{\lambda} \bar{\lambda}^T| \\ &= \left\{ \prod_{i=1}^n (\gamma_{ii} - \zeta) \right\} \left\{ 1 + \mu^2 \bar{\lambda}^T (\Lambda_{\gamma} - \zeta I)^{-1} \bar{\lambda} \right\} \\ &= \left\{ \prod_{i=1}^n (\gamma_{ii} - \zeta) \right\} \left\{ 1 + \mu^2 \sum_{i=1}^n \frac{\lambda_i^2}{\gamma_{ii} - \zeta} \right\} \\ &= g(\zeta) \end{aligned}$$

考虑方程

$$g(\zeta) = \left\{ \prod_{i=1}^n (\gamma_{ii} - \zeta) \right\} f(\zeta) \quad (21)$$

$$\text{其中 } f(\zeta) = \left\{ 1 + \mu^2 \sum_{i=1}^n \frac{\lambda_i^2}{\gamma_{ii} - \zeta} \right\}$$

显然 F 的特征值, 即 $g(\zeta)$ 的零点只能由 $f(\zeta)$ 的零点组成, 下面讨论 $f(\zeta)$ 的零点分布。

注意到 $f(\zeta)$ 的极点

$\gamma_{ii} = 1 - 2\mu\lambda_i + 2\mu^2\lambda_i^2 = (1 - \mu\lambda_i)^2 + \mu^2\lambda_i^2 > 0$, 位于正半轴。将极点按大小顺序排列为 $p_1 < p_2 < \dots < p_n$, 这里 $\{p_j\}_{j=1}^n = \{\gamma_{ii}\}_{i=1}^n$, 并用 η_j 替换 λ_i , 其中 $\{\eta_j\}_{j=1}^n = \{\lambda_i\}_{i=1}^n$ 。这样就有:

$$f(\zeta) = \left\{ 1 + \mu^2 \sum_{i=1}^n \frac{\lambda_i^2}{\gamma_{ii} - \zeta} \right\} = \left\{ 1 + \mu^2 \sum_{j=1}^n \frac{\eta_j^2}{p_j - \zeta} \right\}$$

约定 $p_0 = -\infty$, $p_{n+1} = \infty$, 从而 $f(\zeta)$ 在 $\bigcup_{j=0}^n (p_j, p_{j+1})$

上分段可微。由于

$$\frac{df(\zeta)}{d\zeta} = \mu^2 \sum_{j=1}^n \frac{\eta_j^2}{(p_j - \zeta)^2} > 0$$

所以 $f(\zeta)$ 在 $\bigcup_{j=0}^n (p_j, p_{j+1})$ 上分段单调递增。

又 $f(-\infty)=1$, $f(\infty)=1$ 以及

$$f(p_i^-) = \lim_{\zeta \rightarrow p_i^-} \left\{ 1 + \mu^2 \sum_{j=1}^n \frac{\eta_j^2}{p_j - \zeta} \right\} = +\infty$$

$$f(p_i^+) = \lim_{\zeta \rightarrow p_i^+} \left\{ 1 + \mu^2 \sum_{j=1}^n \frac{\eta_j^2}{p_j - \zeta} \right\} = -\infty$$

其中 $i=1, 2, \dots, n$ 。

易知 $f(\zeta) \sim \zeta$ 的关系曲线如图 1 所示, $f(\zeta)$ 的零点 ζ_l ($l=1, 2, \dots, n$) 满足:

$$0 < p_1 < \zeta_1 < p_2 < \zeta_2 < \dots < \zeta_{n-1} < p_n < \zeta_n < \infty$$

由于 $f(\zeta)$ 分段单调递增, 所以 $f(\zeta) > 0$ 只能说明 $\zeta_m, \zeta_{m-1}, \dots, \zeta_1 < t$, 这里 m 满足 $p_m < t < p_{m+1}$ 。为保证 $\bar{\sigma}(k)$ 收敛, 要求 F 的特征值, 即 $f(\zeta)$ 的零点 $\zeta_l < 1$, 这等价于 $t=1$, $m=n$ 情况, 因而:

$$\begin{cases} p_n < 1 \\ f(1) > 0 \end{cases}$$

即

$$\begin{cases} p_n < 1 \\ 1 + \mu^2 \sum_{i=1}^n \frac{\eta_i^2}{p_i - 1} > 0 \end{cases} \quad (22)$$

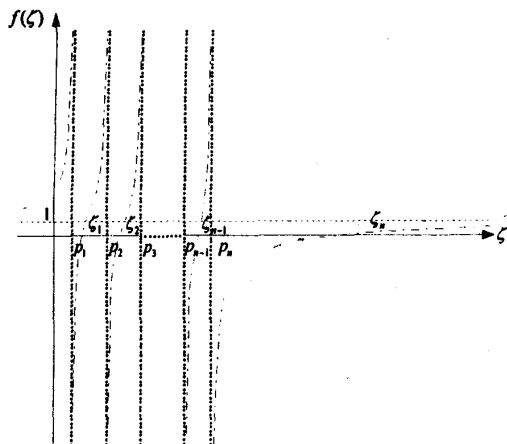


图 1 $\zeta, \sim \zeta$ 的关系曲线

根据 p_i 与 γ_{ii} 间的关系, (22)式可用 γ_{ii} 表示为:

$$\begin{cases} \gamma_{ii} < 1 & i = 1, 2, \dots, n \\ 1 + \mu^2 \sum_{i=1}^n \frac{\lambda_i^2}{\gamma_{ii} - 1} > 0 \end{cases}$$

将 γ_{ii} 的表达式代入并整理有:

$$\begin{cases} 0 < \mu < 1/\lambda_i & i = 1, 2, \dots, n \\ \sum_{i=1}^n \frac{\mu \lambda_i}{1 - \mu \lambda_i} < 2 \end{cases} \quad (23)$$

式(23)就是 F 的特征值在单位圆内的充要条件, 即 $\bar{\sigma}(k)$ 以及均方误差收敛的充要条件。因没有采用第三部分的近似, 所以得到的结论更确切, 容易验证 $0 < \mu < 2/\lambda_{\max}$ 不能使(23)式恒成立。

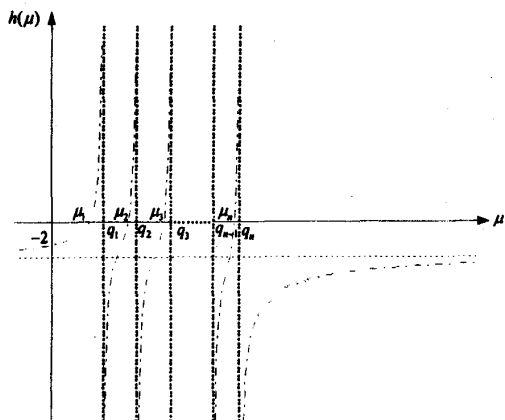


图 2 $h(\mu) \sim \mu$ 的关系曲线

式(23)所示的条件很难应用, 下面将其变通到对 μ 的限制上。

记 $h(\mu) = \sum_{i=1}^n \frac{\mu \lambda_i}{1 - \mu \lambda_i} - 2$, 其极点为 $1/\lambda_i$ ($i = 1, 2, \dots, n$)。将极点按大小顺序重新排列为 $q_1 < q_2 < \dots < q_n$, 这里 $\{q_j\}_{j=1}^n = \{1/\lambda_i\}_{i=1}^n$ 。采用与讨论 $f(\zeta)$ 的零点分布相似的方法, 易知 $h(\mu) \sim \mu$ 的关系曲线如图 2 所示。这里约定 $q_0 = -\infty$, $q_{n+1} = \infty$, $h(\mu)$ 的零点 μ_l 满足:

$$q_0 < 0 < \mu_1 < q_1 < \mu_2 < q_2 < \dots < \mu_{n-1} < q_{n-1} < \mu_n < q_n$$

这样以来(23)式就与下式等价

$$\begin{cases} 0 < \mu < q_1 \\ q_{l-1} < \mu < \mu_l & l = 1, 2, \dots, n \end{cases}$$

也即 $0 < \mu < \mu_1 = \min(\mu_l)$, 这就是(23)式变通到 μ 上的结果。下面来估计 $h(\mu)$ 最小的零点 μ_1 的大小。

$$\text{给方程 } \sum_{i=1}^n \frac{\lambda_i}{1/\mu - \lambda_i} - 2 = 0 \text{ 两边同乘以 } \prod_{j=1}^n (1/\mu - \lambda_j),$$

并整理得到:

$$\prod_{j=1}^n (1/\mu - \lambda_j) - (1/2) \sum_{i=1}^n \lambda_i \prod_{j=1, j \neq i}^n (1/\mu - \lambda_j) = 0$$

也即

$$\left(\frac{1}{\mu}\right)^n - a_1 \left(\frac{1}{\mu}\right)^{n-1} + a_2 \left(\frac{1}{\mu}\right)^{n-2} - \dots + (-1)^n a_n = 0 \quad (24)$$

其中

$$a_1 = (3/2) \sum_{i=1}^n \lambda_i, a_2 = \sum_{i=1}^n \sum_{j=1, j \neq i}^n \lambda_i \lambda_j, a_n = (1+n/2) \prod_{i=1}^n \lambda_i$$

方程(24)的解 (即 $h(\mu)$ 的零点) 为 μ_l 当且仅当

$$\left(\frac{1}{\mu}\right)^n - a_1 \left(\frac{1}{\mu}\right)^{n-1} + a_2 \left(\frac{1}{\mu}\right)^{n-2} - \dots + (-1)^n a_n = \prod_{i=1}^n \left(\frac{1}{\mu} - \frac{1}{\mu_i}\right)$$

从而有

$$a_1 = \sum_{i=1}^n (1/\mu_i) = (3/2) \sum_{i=1}^n \lambda_i \quad (25)$$

$$a_2 = (1/2) \sum_{i=1}^n \sum_{j=1, j \neq i}^n (1/\mu_i)(1/\mu_j) = \sum_{i=1}^n \sum_{j=1, j \neq i}^n \lambda_i \lambda_j \quad (26)$$

对于集合 $\{z_i > 0; i = 1, 2, \dots, n\}$ 有 (详见附录):

$$\max\{z_i\} \leq \frac{1}{n} \{s_1 + \sqrt{(n-1)(ns_2 - s_1^2)}\}$$

其中

$$s_1 = \sum_{i=1}^n z_i, s_2 = \sum_{i=1}^n (z_i)^2 = \left(\sum_{i=1}^n z_i\right)^2 - \sum_{i=1}^n \sum_{j=1, j \neq i}^n z_i z_j$$

将这个结论应用于集合 $\{1/\mu_i; i = 1, 2, \dots, n\}$ 有:

$$\frac{1}{\mu_1} \leq \frac{1}{n} \{a_1 + \sqrt{(n-1)[n(a_1^2 - 2a_2) - a_1^2]}\} \quad (27)$$

再将(25)、(26)两式代入可得:

$$\frac{1}{\mu_1} \leq \frac{1}{n} \left\{ (3/2) \sum_{i=1}^n \lambda_i + \sqrt{(n-1) \left\{ (3/2) \sum_{i=1}^n \lambda_i \right\}^2 - 2n \sum_{i=1}^n \sum_{j=1, j \neq i}^n \lambda_i \lambda_j} \right\}$$

上式可进一步整理为

$$\mu_1 \geq \frac{n}{\frac{3}{2} \sum_{i=1}^n \lambda_i + \sqrt{\frac{9(n-1)^2}{4} \left(\sum_{i=1}^n \lambda_i \right)^2 - 2n(n-1) \sum_{i=1}^n \sum_{j=1, j \neq i}^n \lambda_i \lambda_j}} = \bar{\mu}$$

因而 MSE 收敛的一个充分条件为 $0 < \mu < \bar{\mu}$ 。注意到

$$\bar{\mu} > \frac{n}{\frac{3}{2} \sum_{i=1}^n \lambda_i + \sqrt{\frac{9(n-1)^2}{4} \left(\sum_{i=1}^n \lambda_i \right)^2}} = \frac{2}{3tr(K_{xx})} \quad (28)$$

从而 MSE 收敛的一个充分条件为 $0 < \mu \leq 2/(3tr(K_{xx}))$ ，这个条件比 $0 < \mu < \mu_1$ 容易应用。

4.2 失调表式

对(20)式两边取极限可得：

$$\begin{aligned} \bar{\sigma}(\infty) &= \mu^2 \xi_{\min} (I - F)^{-1} \bar{\lambda} \\ &= \mu^2 \xi_{\min} (2\mu\Lambda - 2\mu^2\Lambda^2 - \mu^2 \bar{\lambda} \bar{\lambda}^T)^{-1} \bar{\lambda} \\ &= \mu^2 \xi_{\min} (\Lambda_i - \mu^2 \bar{\lambda} \bar{\lambda}^T)^{-1} \bar{\lambda} \end{aligned}$$

其中 $\Lambda_i = 2\mu\Lambda - 2\mu^2\Lambda^2$

对矩阵 $(\Lambda_i - \mu^2 \bar{\lambda} \bar{\lambda}^T)^{-1}$ 应用矩阵求逆引理有^[1]：

$$(\Lambda_i - \mu^2 \bar{\lambda} \bar{\lambda}^T)^{-1} = \Lambda_i^{-1} + \Lambda_i^{-1} \bar{\lambda} (1/\mu^2 - \bar{\lambda}^T \Lambda_i^{-1} \bar{\lambda})^{-1} \bar{\lambda}^T \Lambda_i^{-1}$$

从而：

$$\begin{aligned} \bar{\sigma}(\infty) &= \mu^2 \xi_{\min} \{ \Lambda_i^{-1} + \Lambda_i^{-1} \bar{\lambda} (1/\mu^2 - \bar{\lambda}^T \Lambda_i^{-1} \bar{\lambda})^{-1} \bar{\lambda}^T \Lambda_i^{-1} \} \bar{\lambda} \\ &= \mu^2 \xi_{\min} \{ \Lambda_i^{-1} \bar{\lambda} + \Lambda_i^{-1} \bar{\lambda} (1/\mu^2 - \bar{\lambda}^T \Lambda_i^{-1} \bar{\lambda})^{-1} \bar{\lambda}^T \Lambda_i^{-1} \bar{\lambda} \} \\ &= \{ \mu^2 / (1 - \mu^2 \bar{\lambda}^T \Lambda_i^{-1} \bar{\lambda}) \} \xi_{\min} \Lambda_i^{-1} \bar{\lambda} \end{aligned}$$

因此失调可表示为：

$$\begin{aligned} M_{LMS} &= \{ \mu^2 / (1 - \mu^2 \bar{\lambda}^T \Lambda_i^{-1} \bar{\lambda}) \} \bar{\lambda}^T \Lambda_i^{-1} \bar{\lambda} \\ &= \frac{\sum_{i=1}^n \frac{\mu \lambda_i}{1 - \mu \lambda_i}}{2 - \sum_{i=1}^n \frac{\mu \lambda_i}{1 - \mu \lambda_i}} \quad (30) \end{aligned}$$

5 实验结果

以 6 个抽头的自适应滤波器为例。 k 时刻滤波器第 i 个抽头的输入为 $s(k) + n_i(k)$ ($i=1,2,\dots,6$)，这里 $s(k)$, $n_i(k)$ 均为零均值白色高斯序列，方差分别为 σ_s^2, σ_n^2 ，且 $s(k)$ 与 $n_i(k)$ 不相关。易知滤波器输入的自相关矩阵的特征值为 $\lambda_1 = \lambda_2 = \dots = \lambda_5 = \sigma_n^2$, $\lambda_6 = 6\sigma_s^2 + \sigma_n^2$ 。取 $\sigma_s^2 = \sigma_n^2 = 1$ ，则最大特征值 $\lambda_{\max} = 7$ 。

这个自适应系统的期望信号为 $s(k)$ ，则 $\bar{w}_{out} = (0.14286, 0.14286, \dots, 0.14286)^T$ ，最小均方误差 ξ_{\min} 为 0.14286， $h(\mu)$ 最小的零点为 $\mu_1 = 0.08634$ 。

取 $\mu = 0.09$ ，注意到 $\mu_1 < \mu < 2/\lambda_{\max}$ ，根据本文第四

部分的分析可知这种情况下均方误差发散。图 3 为 $\mu = 0.09$ 的学习曲线，从图明显可以看出 MSE 发散，因而 $0 < \mu < 2/\lambda_{\max}$ 的收敛条件对于不相关高斯输入而言欠准确。

取失调为 0.5，根据(17)式，(30)式求得两种分析方法对应的步长分别为 0.06934, 0.04357，这两个步长对应的学习曲线分别如图 4，图 5 所示，图中的水平线为失调等于 0.5 时均方误差的稳态值。显然，图 5 所示的学习曲线的失调约为 0.5，图 4 的失调远非 0.5，因而第二种分析方法得到的失调表式对失调的预报更准确。

6 结论

本文分析了不相关高斯输入情况下，最小均方自适应滤波器均方误差的收敛性。分析当中不采用将瞬时自相关矩阵近似为统计自相关矩阵的作法，而是借助高斯阶矩阵分解定理得到了更为确切的均方误差收敛的充分必要条件以及均方误差收敛前提下的失调表式，并给出均方误差收敛的一个充分条件 $0 < \mu \leq 2/(3tr(K_{xx}))$ 。实验结果进一步验证了分析得到的均方误差收敛条件以及失调表式的正确性。

参考文献

- [1] S.Haykin, Adaptive Filter Theory, Third Edition, Prentice-Hall, Inc, 1996.
- [2] Vijay K. Madisetti, Douglas B.Williams, The Digital Signal Processing Handbook, 1998 by CRC Press LLC.
- [3] A. Feuer, E.Weinstein, Convergence Analysis of LMS Filters with Uncorrelated Gaussian Data, IEEE Transactions on Acoustic, Speech, And Signal Processing, Vol.ASSP-33, 222-230, Feb. 1985
- [4] B.Widrow, E.Walach, Adaptive Inverse Control, 刘树棠, 韩崇昭译, 西安交通大学出版社, 2000.5.
- [5] Raymond H. Kwong, Edward W. Johnston, A Variable Step Size LMS Algorithm, IEEE Transactions on Signal Processing, Vol.40, 1633-1642, No. 7, July 1992.
- [6] 丁玉美, 阔永红, 高新波编著, 数字信号处理—时域离散随机信号处理, 西安电子科技大学出版社, 2002.12.

【附录】

已知 $p_i > 0 (i=1,2,\dots,N)$, $\sum_{i=1}^N \rho_i = s_1$, $\sum_{i=1}^N \rho_i^2 = s_2$, 则

$$0 < \rho_i \leq \rho_M = \left\{ s_1 + \sqrt{(N-1)(Ns_2 - s_1^2)} \right\} / N$$

(1) 预备知识：

算术平均与均方根之间的关系为

$$\left| \sum_{i=1}^N a_i \right| / N \leq \sqrt{\sum_{i=1}^N a_i^2} / N$$

等号成立当且仅当各 $a_i (i=1, 2, \dots, N)$ 相等。

显然, 对于 $\rho_i > 0$ 有: $\left(\sum_{i=1}^N \rho_i \right)^2 \leq N \sum_{i=1}^N \rho_i^2$

(2) 根据算术平均与均方根之间的关系:

$$s_1^2 \leq N s_2 \quad (\text{A.1})$$

等号成立当且仅当各 $\rho_i (i=1, 2, \dots, N)$ 全相等。

以及

$$\left(\sum_{l=1, l \neq j}^N \rho_l \right)^2 \leq (N-1) \sum_{l=1, l \neq j}^N \rho_l^2 \quad (\text{A.2})$$

等号成立当且仅当各 $\rho_l (l=1, 2, \dots, j-1, j+1, \dots, N)$ 全相等。

注意到 $\sum_{l=1, l \neq j}^N \rho_l = s_1 - \rho_j$, $\sum_{l=1, l \neq j}^N \rho_l^2 = s_2 - \rho_j^2$, 从而(A.2)

式变为

$$(s_1 - \rho_j)^2 \leq (N-1)(s_2 - \rho_j^2) \quad (\text{A.3})$$

不等式(A.3)可整理成:

$$N\rho_j^2 - 2s_1\rho_j + s_1^2 - (N-1)s_2 \leq 0$$

其判别式为

$$\Delta = 4s_1^2 - 4N(s_1^2 - (N-1)s_2) = 4(N-1)(Ns_2 - s_1^2)$$

由(A.1)的 $s_1^2 \leq Ns_2$ 有: $\Delta \geq 0$, 从而不等式(A.3)的解为:

$$s_1 - \frac{\sqrt{(N-1)(Ns_2 - s_1^2)}}{N} \leq \rho_j \leq \frac{s_1 + \sqrt{(N-1)(Ns_2 - s_1^2)}}{N} \triangleq \rho_M$$

这里 $j=1, 2, \dots, N$ 。

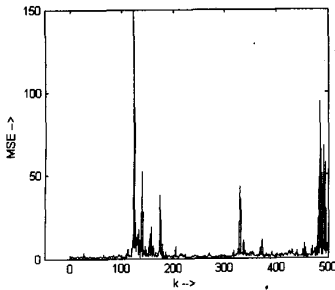


图 3 $\mu=0.09$ 时自适应滤波器的学习曲线

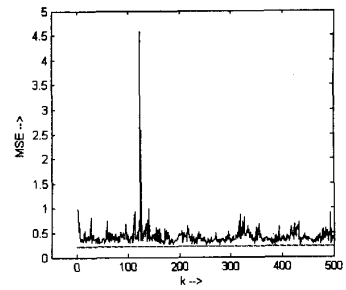


图 4 $\mu=0.06934$ 时自适应滤波器的学习曲线

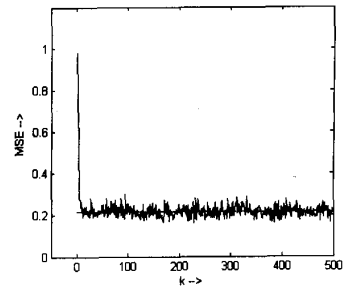


图 5 $\mu=0.04357$ 时自适应滤波器的学习曲线

作者简介

李琳, 国防科技大学电子科学与工程学院博士研究生。
研究方向扩频通信抗干扰技术和自适应信号处理。

微波滤波器设计培训——视频课程

易迪拓培训(www.edatop.com)由数名来自于研发第一线的资深工程师发起成立, 致力和专注于微波、射频、天线设计研发人才的培养, 是国内最大的微波射频和天线设计人才培养基地。客户遍布中兴通讯、研通高频、国人通信等多家国内知名公司, 以及台湾工业技术研究院、永业科技、全一电子等多家台湾地区企业。

我们推出的微波滤波器设计培训专题, 有资深工程师领衔主讲, 课程既有微波滤波器设计原理的详细解释, 也有各种仿真分析工具的实际设计应用讲解, 设计原理和设计仿真实践相结合, 向大家呈现各种结构的微波滤波器的完整设计流程。旨在帮助大家透彻地理解并实际的掌握各种微波滤波器的设计。



微波滤波器设计培训专题视频课程

高清视频, 专家授课, 中文讲解, 直观易学; 既有微波滤波器设计原理的详细解释, 也有像 ADS、CST、HFSS 各种仿真分析工具的实际设计应用讲解, 旨在帮助大家透彻地理解并实际的掌握各种微波滤波器的设计。

课程网址: <http://www.edatop.com/peixun/filter/>

更多专业培训课程:

- **HFSS 视频培训课程**

网址: <http://www.edatop.com/peixun/hfss/>

- **CST 视频培训课程**

网址: <http://www.edatop.com/peixun/cst/>

- **天线设计专业培训课程**

网址: <http://www.edatop.com/peixun/antenna/>