

文章编号: 1001-0920(2010)09-1399-04

基于 DDT 方法的自适应总体最小二乘算法收敛性分析

李晓波, 樊养余, 彭 轲

(西北工业大学 电子信息学院, 西安 710129)

摘要: 针对确定连续时间(DCT)方法在研究最小主元分析(MCA)学习算法时,存在严格限制条件而较难实现的问题,基于确定离散时间(DDT)方法,研究了 AMEX MCA 学习算法的收敛条件.理论分析表明,只有在算法的学习因子和输入信号的自相关矩阵满足一定条件时,AMEX MCA 学习算法才能收敛到系统的总体最小二乘解.最后,仿真结果表明了收敛条件的正确性.

关键词: 总体最小二乘; 最小主元分析; 确定离散时间

中图分类号: TN911.7

文献标识码: A

Convergence analysis of adaptive total least square based on deterministic discrete time

LI Xiao-bo, FAN Yang-yu, PENG Ke

(College of Electronics Information, Northwestern Polytechnical University, Xi'an 710129, China. Correspondent: LI Xiao-bo, E-mail: su27baby@163.com)

Abstract: Deterministic continuous time(DCT) is a conventional method of studying the minor component analysis(MCA). Unfortunately, DCT is not used in practical systems because of its strict conditions. Therefore, the convergence condition of AMEX MCA learning algorithm is derived based on deterministic discrete time. Theoretical analysis shows that the total least square solution is not obtained until special conditions between learning factor and autocorrelation matrix of input signal are satisfied. Finally, simulation results show the correctness of the convergence condition.

Key words: Total least square; Minor component analysis; Deterministic discrete time

1 引言

总体最小二乘(TLS)算法是解决输入、输出观测数据均含有噪声的信号处理问题的一种有效方法^[1].早期求解 TLS 问题的困难主要在于 SVD(singular value decomposition)方法的复杂性.近年来,出现了大量求解总体最小二乘问题的自适应递推算法.总体最小二乘问题可转化为求解输入信号自相关矩阵最小特征值所对应的特征向量即最小主元分析(MCA)^[2].MCA 算法主要有文献[3]中的 OJA+MCA 学习算法、[4]中的 TLS MCA 算法、[5]中的 OJAm MCA 算法、[6]中的 MCA EXIN 算法、[7, 8]中的 MCA 算法、[9]中的 OJAn MCA 算法、[10]中的 SSMSR MCA 算法、[11]中的 AMEX MCA 学习算法等.这些 MCA 学习算法的区别在于自适应调整权值向量的递推公式的不同.其中, AMEX MCA 学习算法不受初始权值向量的影响且具有良好的全局收敛

性,是一种较好的 MCA 学习算法^[11].但文献[11]未给出 AMEX MCA 学习算法的收敛条件.然而,研究 MCA 学习算法的收敛性十分重要,只有最终收敛的 MCA 学习算法才具有实际意义.MCA 学习算法通常被描述为随机离散时间(SDT)系统,直接分析其收敛性存在困难.研究 SDT 系统收敛性的传统方法是确定连续时间(DCT)方法^[12].DCT 方法是根据随机逼近理论^[13],在满足某些严格限制条件的情况下,将随机离散时间(SDT)系统转换为与之对应的确定连续时间(DCT)系统.通过这样的转换,SDT 系统可被描述为对应的微分方程.但采用 DCT 方法的限制条件之一就是要求算法的学习因子必须趋于零^[13].而实际系统中,学习因子常被设置成大于零的常数,因而大大限制了 DCT 方法的应用.近年来,确定离散时间(DDT)方法被用于研究 SDT 系统的收敛性问题^[14-17].该方法通过取条件期望将随

收稿日期: 2009-08-02; 修回日期: 2009-12-16.

基金项目: 国家863计划项目(2007AA01Z324).

作者简介: 李晓波(1979—),男,陕西澄城人,工程师,博士生,从事自适应信号处理、无线信道均衡技术的研究;樊养余(1960—),男,西安人,教授,博士生导师,从事信号检测与恢复、信息安全等研究.

机离散时间(SDT)系统转换为对应的确定离散时间(DDT)系统. 相对于传统的DCT方法, DDT方法既能保持MCA学习算法的时间离散的特征, 又不要求算法的学习因子必须趋于零, 是一种更为实际合理的分析方法. 彭德中等^[17]利用DDT方法研究了AMEX MCA学习算法的收敛条件, 但经过仿真实验可知, 该收敛条件 $\eta\lambda_1 < 0.25$ 仅是算法收敛的充分条件而非必要条件.

针对文献[17]存在的缺点, 基于DDT方法, 本文推导出了保证AMEX MCA学习算法收敛的充要条件, 计算机仿真结果表明了所建立收敛条件的正确性.

2 DDT方法

考虑输入、输出观测数据均含有噪声的FIR自适应滤波系统, 系统权向量 $w(k) \in R^n$, $x(k) \in R^n$ 为输入信号, $y(k) = w^T(k)x(k)$ 为输出信号. 为了便于描述问题, 假定输入信号 $x(k)$ 为零均值随机平稳信号. MCA学习算法的目的就是使系统的权向量 $w(k)$ 自适应地收敛到输入信号自相关矩阵 R 的最小特征值所对应的特征向量, 从而得到系统的总体最小二乘解.

AMEX MCA学习算法权值向量的自适应递推公式为^[11]

$$w(k+1) = w(k) - \eta \left[x(k)x^T(k)w(k) - \frac{w(k)}{w(k)w^T(k)} \right], \quad (1)$$

其中 $\eta > 0$ 为学习因子. 式(1)两边同时求出条件期望 $E\{w(k+1)/w(0), x(i), i < k\}$. 在下一步迭代中用条件期望值代替原公式, 这样就得到了基于DDT方法的权值向量 $w(k)$ 的自适应递推公式, 即

$$w(k+1) = w(k) - \eta \left[Rw(k) - \frac{w(k)}{w(k)w^T(k)} \right], \quad (2)$$

其中 $R = E\{x(k)x^T(k)\}$ 为输入信号 $x(k)$ 的自相关矩阵. 输入信号自相关矩阵 R 是对称的非负定矩阵, 则其特征值 $\lambda_i \geq 0 (i = 1, 2, \dots, n)$, 令 λ_n 为最小特征值, λ_1 为最大特征值, 即 $\lambda_1 \geq \dots \geq \lambda_i \geq \dots \geq \lambda_n (1 \leq i \leq n)$. 相应的单位特征向量 $v_i (i = 1, 2, \dots, n)$ 构成了 R^n 的标准正交基. 实际系统中由于噪声的影响, 自相关矩阵 R 的最小特征值通常大于零, 不失一般性, 令 $\lambda_n > 0$. v_i 是 R^n 的标准正交基, 因此权值向量 $w(k)$ 可表示为

$$w(k) = \sum_{i=1}^n z_i(k)v_i, k \geq 0, \quad (3)$$

其中 $z_i(k)$ 为固定常数. 将式(3)代入(2)可得

$$z_i(k+1) = \left[1 - \eta\lambda_i + \frac{\eta}{w(k)w^T(k)} \right] z_i(k), \quad (4)$$

$$i = 1, 2, \dots, n$$

3 AMEX MCA学习算法收敛性分析

以下的讨论都是基于 $\eta\lambda_n < 0.5$ 的条件进行.

引理 1 若 $2\sqrt{\eta(1-\eta\lambda_n)} \leq x \leq \sqrt{\frac{1}{\lambda_n}}$, 则有

$$(1-\eta\lambda_n)x + \eta/x \leq \sqrt{1/\lambda_n}. \quad (5)$$

证明 令 $f(x) = (1-\eta\lambda_n)x + \eta/x$, 则有 $f'(x) = 1 - \eta\lambda_n - \eta/x^2$. 易知当 $x \geq \sqrt{\frac{\eta}{1-\eta\lambda_n}}$ 时, $f'(x) \geq 0$.

由 $\eta\lambda_n < 0.5$ 可知

$$\sqrt{\frac{\eta}{1-\eta\lambda_n}} < 2\sqrt{\eta(1-\eta\lambda_n)} < \sqrt{\frac{1}{\lambda_n}}.$$

因此, 有

$$f'(x) \geq 0, 2\sqrt{\eta(1-\eta\lambda_n)} \leq x \leq \sqrt{1/\lambda_n}.$$

即 $f(x)$ 在区间 $x \in [2\sqrt{\eta(1-\eta\lambda_n)}, \sqrt{1/\lambda_n}]$ 是单调递增的, 于是 $f(x) \leq f(\sqrt{1/\lambda_n}) = \sqrt{1/\lambda_n}$. \square

引理 2 若 $\eta \leq \frac{2}{\lambda_1 - \lambda_n}$, $w^T(0)v_n \neq 0$, 则当 $\|w(k)\|^2 \leq \sqrt{1/\lambda_n}$ 时, 有

$$\left| \frac{1 - \eta\lambda_i + \eta/\|w(k)\|^2}{1 - \eta\lambda_n + \eta/\|w(k)\|^2} \right| < 1, \quad (6)$$

$$i = 1, 2, \dots, n-1.$$

证明

$$\left| \frac{1 - \eta\lambda_i + \eta/\|w(k)\|^2}{1 - \eta\lambda_n + \eta/\|w(k)\|^2} \right| =$$

$$1 - \frac{\eta(\lambda_i - \lambda_n)}{1 - \eta\lambda_n + \eta/\|w(k)\|^2} <$$

$$1 - \frac{\eta(\lambda_1 - \lambda_n)}{1 - \eta\lambda_n + \eta/\|w(k)\|^2}.$$

由 $\|w(k)\|^2 \leq \sqrt{1/\lambda_n}$ 可知

$$1 - \eta\lambda_n + \eta/\|w(k)\|^2 \geq 1 \Rightarrow$$

$$\frac{\eta(\lambda_1 - \lambda_n)}{1 - \eta\lambda_n + \eta/\|w(k)\|^2} > 0 \Rightarrow$$

$$1 - \frac{\eta(\lambda_1 - \lambda_n)}{1 - \eta\lambda_n + \eta/\|w(k)\|^2} < 1.$$

由 $\eta \leq \frac{2}{\lambda_1 - \lambda_n} \Rightarrow 1 - \frac{\eta(\lambda_1 - \lambda_n)}{1 - \eta\lambda_n + \eta/\|w(k)\|^2} > -1$.

综上所述, $\left| \frac{1 - \eta\lambda_i + \eta/\|w(k)\|^2}{1 - \eta\lambda_n + \eta/\|w(k)\|^2} \right| < 1$ 得证. \square

定理 1 若 $\eta \leq \frac{2}{\lambda_1 - \lambda_n}$, $w^T(0)v_n \neq 0$, 则式(2)的最终收敛域为

$$S = \left\{ w \mid 2\sqrt{\eta(1-\eta\lambda_n)} \leq \|w\| \leq \sqrt{1/\lambda_n} \right\}.$$

即存在一个较大的正常数 k' , 当 $k \geq k'$ 时, 有

$$2\sqrt{\eta(1-\eta\lambda_n)} \leq \|w\| \leq \sqrt{1/\lambda_n}. \quad (7)$$

证明 1) 当 $2\sqrt{\eta(1-\eta\lambda_n)} \leq \|w\| \leq \sqrt{1/\lambda_n}$ 时, 有

$$\|w(k+1)\|^2 = \sum_{i=1}^n z_i^2(k+1) =$$

$$\begin{aligned} \sum_{i=1}^n z_i^2(k) \left[1 - \eta\lambda_i + \frac{\eta}{w^T(k)w(k)} \right]^2 &\leq \\ \sum_{i=1}^n z_i^2(k) \left[1 - \eta\lambda_n + \frac{\eta}{\|w(k)\|^2} \right]^2 &= \\ \|w(k)\|^2 \left[1 - \eta\lambda_n + \frac{\eta}{\|w(k)\|^2} \right]^2 &= \\ \left[(1 - \eta\lambda_n) \|w(k)\| + \frac{\eta}{\|w(k)\|} \right]^2. \end{aligned}$$

由引理 1 可得

$$\|w(k+1)\| \leq \|w(k)\| \leq \sqrt{1/\lambda_n}.$$

由 $w^T(0)v_n \neq 0$, 并根据式 (4) 可知 $z_n(0) \neq 0$, 即 $z_n(k) \neq 0$. 由此可得

$$\begin{aligned} \frac{z_i^2(k+1)}{z_n^2(k+1)} &= \left[\frac{1 - \eta\lambda_i + \eta/\|w(k)\|^2}{1 - \eta\lambda_n + \eta/\|w(k)\|^2} \right]^2 \frac{z_i^2(k)}{z_n^2(k)} = \\ \varepsilon z_i^2(k)/z_n^2(k), \quad i &= 1, 2, \dots, n-1, \end{aligned}$$

其中 $\varepsilon = \frac{1 - \eta\lambda_i + \eta/\|w(k)\|^2}{1 - \eta\lambda_n + \eta/\|w(k)\|^2}$. 由引理 2 可知 $|\varepsilon| < 1$, 故

$$\frac{z_i^2(k+1)}{z_n^2(k+1)} = |\varepsilon|^2 \frac{z_i^2(k)}{z_n^2(k)} = |\varepsilon|^{2k} \frac{z_i^2(0)}{z_n^2(0)}.$$

因此, $\lim_{k \rightarrow \infty} \frac{z_i^2(k)}{z_n^2(k)} = 0 (i = 1, 2, \dots, n-1)$.

因为 $|w(k+1)| \leq |w(k)| \leq \sqrt{1/\lambda_n}$, 即 $|w(k)|$ 有边界, 则 $z_n(k)$ 必定有边界, 所以有

$$\lim_{k \rightarrow \infty} z_i(k) = 0, i = 1, 2, \dots, n-1. \quad (8)$$

假定存在一个非常大的正常数 k' , 当 $k \geq k'$ 时, 有

$$w(k) \approx z_n(k)v_n, \quad (9)$$

将式 (9) 代入 (4) 可得

$$z_n(k+1) = z_n(k) \left[1 - \eta\lambda_n + \frac{\eta}{z_n^2(k)} \right], \quad (10)$$

即 $|w(k+1)| = |z_n(k+1)| \geq 2\sqrt{\eta(1-\eta\lambda_n)}$.

由以上证明可知 $w(k) \in S$.

2) 当 $\|w(k)\| > \sqrt{1/\lambda_n}$ 时, 有

$$\|w(k+1)\|^2 \leq \|w(k)\|^2 \left[1 - \eta\lambda_n + \frac{\eta}{\|w(k)\|^2} \right]^2.$$

令 $\beta = 1 - \eta\lambda_n + \frac{\eta}{\|w(k)\|^2}$, 因为 $\|w(k)\| > \sqrt{1/\lambda_n}$, 所以 $\beta < 1$, 即 $|w(k+1)| < \beta|w(k)|$, 则必存在一个正常数 k'' , 当 $k \geq k''$ 时, 有 $|w(k)| \leq \sqrt{1/\lambda_n}$, 即 $w(k) \in S$.

3) 当 $0 < \|w(k)\| < 2\sqrt{\eta(1-\eta\lambda_n)}$ 时, 类似于 1) 的证明过程可知, 存在一个大的正常数 k''' , 当 $k \geq k'''$ 时, 有

$$\|w(k+1)\| = |z_n(k+1)| \geq 2\sqrt{\eta(1-\eta\lambda_n)},$$

故 $w(k) \in S$.

综合 1)~3) 可知, 当 $k \geq \max(k', k'', k''')$ 时, 式

(2)的最终收敛域为 S . \square

定理 2 若 $\eta \leq \frac{2}{\lambda_1 - \lambda_n}$, $w^T(0)v_n \neq 0$, 则有

$$\lim_{k \rightarrow \infty} z_n(k) = \pm\sqrt{1/\lambda_n}, \quad i = 1, 2, \dots, n-1.$$

证明 由式 (10) 可知,

$$|z_n(k+1)| = |z_n(k)| \left[1 - \eta\lambda_n + \frac{\eta}{z_n^2(k)} \right].$$

根据引理 1, $|z_n(k+1)|$ 在区间 $[2\sqrt{\eta(1-\eta\lambda_n)}, \sqrt{1/\lambda_n}]$ 上单调递增, 故 $|z_n(k+1)| \leq \sqrt{1/\lambda_n} (k \geq k')$, 也就是说当 $k \rightarrow \infty$ 时, $|z_n(k)|$ 必定收敛到其上边界 $\sqrt{1/\lambda_n}$, 即

$$\lim_{k \rightarrow \infty} z_n(k) = \pm\sqrt{1/\lambda_n}, \quad i = 1, 2, \dots, n-1. \quad \square$$

推论 1 若 $\eta \leq \frac{2}{\lambda_1 - \lambda_n}$, $w^T(0)v_n \neq 0$, 则有

$$\lim_{k \rightarrow \infty} w(k) = \pm\sqrt{\frac{1}{\lambda_n}}v_n, \quad i = 1, 2, \dots, n-1. \quad (11)$$

4 仿真结果与分析

4.1 最终收敛域分析

随机生成一个 4×4 的非负定对称矩阵, 即

$$R = \begin{bmatrix} 4.8171 & 4.6284 & 4.5330 & 3.4158 \\ 4.6284 & 5.2653 & 5.3094 & 3.4446 \\ 4.5330 & 5.3094 & 6.5037 & 4.0071 \\ 3.4158 & 3.4446 & 4.0071 & 3.7815 \end{bmatrix}.$$

矩阵 R 的最大特征值 $\lambda_1 = 18.0197$, 最小特征值 $\lambda_n = 0.0218$, 令学习因子 $\eta = 0.01$, 即满足推论 1 的收敛条件 $\eta \leq \frac{2}{\lambda_1 - \lambda_n}$. 图 1 给出了 10 条不同初始权值范数 $\|w(0)\| \in S$ 条件下 $\|w(k)\|$ 的学习曲线.

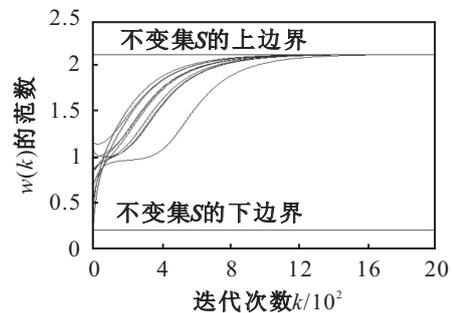


图 1 $\|w(0)\| \in S$ 时, $\|w(k)\|$ 范数的收敛情况

图 2 给出了 10 条不同初始权值范数 $\|w(0)\| \notin S$ 条件下 $\|w(k)\|$ 的学习曲线.

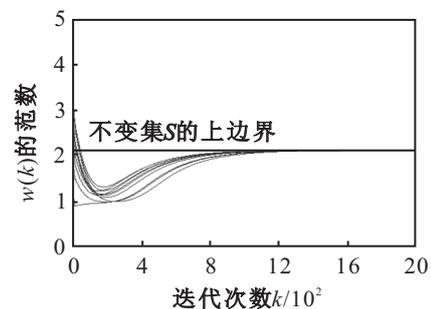


图 2 $\|w(0)\| \notin S$ 时, $\|w(k)\|$ 范数的收敛情况

由图 1, 图 2 可知, 无论初始权值范数 $\|w(0)\|$ 如何取值, $\|w(k)\|$ 都能稳定地收敛于最终收敛域 S . 这表明 AMEX MCA 学习算法不受初始权值向量的影响, 同时也验证了定理 1 的内容.

4.2 权值 $w(k)$ 的收敛性

由式 (3) 可知, $z_i(k) = w^T(k)v_i (i = 1, 2, 3, 4)$, 图 3 为任取一初始权值向量 $w(0) = [0.5, 1, 1.5, 0.8]$ 时, $z_i(k) (i = 1, 2, 3, 4)$ 的收敛情况.

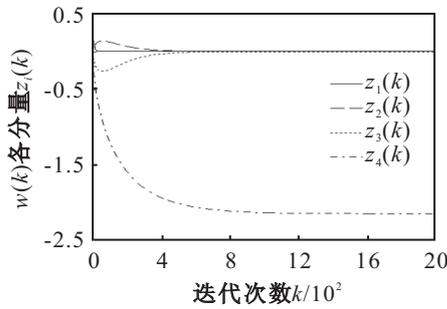


图 3 初始权值为 $w(0)$ 时, $z_i(k)$ 的收敛性

由图 3 可知, $z_i(k) (i = 1, 2, 3, 4)$ 最终收敛于零, 而对应于最小主元的 $z_4(k)$ 最终收敛于某一固定值 $\pm\sqrt{1/\lambda_n}$ (本文实验条件下 $\sqrt{1/\lambda_n} \approx 2.142$), 从而验证了定理 2.

图 4 为当学习因子 $\eta = 0.12$, 即 $\eta > \frac{2}{\lambda_1 - \lambda_n}$ 时, 权值 $w(k)$ 范数的收敛情况. 可以看出, 在学习因子 η 不满足 $\eta \leq \frac{2}{\lambda_1 - \lambda_n}$ 的收敛条件时, 随着迭代次数的增加, $\|w(k)\|$ 最终发散.

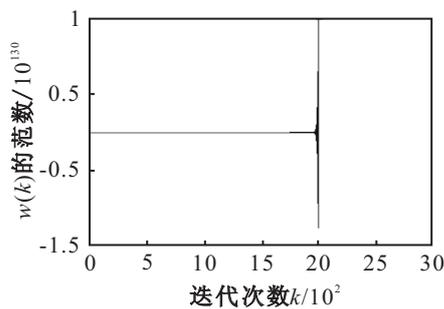


图 4 η 不满足收敛条件时, $\|w(k)\|$ 的发散情况

4.3 随机平稳输入信号状态下的收敛性分析

为了更好地说明 AMEX MCA 学习算法在使用随机输入信号时的收敛性, 可从下式生成的随机数据中提取最小主元:

$$x(k) = Ch(k). \tag{12}$$

其中: $C = \frac{\text{randn}(4, 4)}{4}$, $h(k) \in R^{4 \times 1}$ 为正态分布的随机变量. 利用式 (1) 进行权值向量的自适应递推, 学习因子 $\eta = 0.01$.

定义权值向量 $w(k)$ 的方向余弦为

$$\cos(\theta) = \frac{|w^T(k)v_n|}{\|w(k)\| \|v_n\|}. \tag{13}$$

显然, 如果 $\cos \theta$ 收敛到 1, 则 $w(k)$ 必然收敛到最小主元 v_n 所在的方向上.

在仿真实验中, 权值向量的方向余弦 $\cos \theta$ 最终收敛为 1, 如图 5 所示. 这意味着 AMEX MCA 学习算法给出了正确的收敛结果.

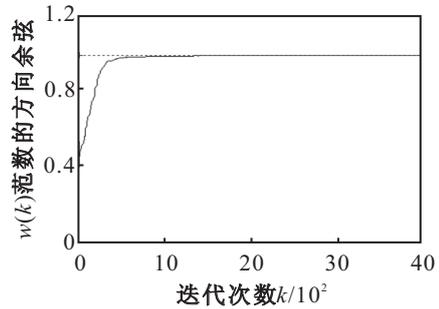


图 5 权值向量方向余弦的收敛性

5 结 论

针对 DCT 方法在研究 AMEX MCA 学习算法时存在限制条件的问题, 本文利用 DDT 方法对算法的收敛性展开研究, 推导出了算法的收敛条件. 该收敛条件是利用 AMEX MCA 学习算法自适应求解总体最小二乘问题的前提. 需要指出的是, 在 $\eta\lambda_n \geq 0.5$ (即不满足 $\eta\lambda_n < 0.5$) 时, AMEX MCA 学习算法的收敛条件还需要进一步研究.

参考文献(References)

- [1] Golub G H, Loan F V. An analysis of total least squares problem[J]. SIAM Numer Analysis, 1980, 17(6): 883-893.
- [2] 孔祥玉, 韩崇昭, 魏瑞轩, 等. 一种鲁棒的总体最小二乘自适应辨识算法[J]. 小型微型计算机系统, 2005, 26(6): 968-971.
(Kong X Y, Han C Z, Wei R X, et al. Robust total least mean square adaptive identification algorithm[J]. Mini-micro Systems, 2005, 26(6): 968-971.)
- [3] Oja E. Principal components, minor components and linear neural networks[J]. Neural Networks, 1992, 5(6): 927-935.
- [4] Feng D Z, Bao Z, Jiao L C. Total least mean squares algorithm[J]. IEEE Trans on Signal Processing, 1998, 46(8): 2122-2130.
- [5] Feng D Z, Zheng W X, Jia Y. Neural network learning algorithms for tracking minor subspace in high-dimensional data stream[J]. IEEE Trans on Neural Networks, 2005, 16(3): 513-521.
- [6] Cirrincione G, Cirrincione M, Herault J, et al. The MCA EXIN neuron for the minor component analysis[J]. IEEE Trans on Neural Networks, 2002, 13(1): 160-187.
- [7] Luo F L, Unbehauen R, Cichocki A. A minor component analysis algorithm[J]. Neural Networks, 1997, 10(2): 291-297.

(下转第 1407 页)