

基于随机梯度的变动量因子 自适应白化算法

欧世峰¹ 高颖¹ 赵晓晖²

摘要 针对自适应白化技术中算法的收敛速度问题,通过融入具有变动量因子特性的动量项,提出了一种快速的自适应白化算法.该算法利用动量项来加速系统的收敛速度,并基于随机梯度方法对动量因子进行自适应更新,有效提升了白化系统的整体性能.仿真实验表明本文算法在平稳和非平稳环境下具有良好的性能.

关键词 白化,动量项,自适应,盲源分离

引用格式 欧世峰,高颖,赵晓晖.基于随机梯度的变动量因子自适应白化算法.自动化学报,2012,38(8):1370-1374

DOI 10.3724/SP.J.1004.2012.01370

Stochastic Gradient Based Variable Momentum Factor Algorithm for Adaptive Whitening

OU Shi-Feng¹ GAO Ying¹ ZHAO Xiao-Hui²

Abstract In view of the convergence speed in adaptive whitening technique, by adding a momentum term with a variable momentum factor, a fast adaptive whitening algorithm is proposed. First, the presented algorithm incorporates a momentum term into the adaptive whitening algorithm to accelerate the convergence speed. Then, a variable momentum factor to improve the allover performance of the whitening system is obtained based on the stochastic gradient of the cost function. Experimental results demonstrate the good performance of the proposed fast algorithm both in stationary and non-stationary conditions.

Key words Whitening, momentum term, adaptive, blind source separation

Citation Ou Shi-Feng, Gao Ying, Zhao Xiao-Hui. Stochastic gradient based variable momentum factor algorithm for adaptive whitening. *Acta Automatica Sinica*, 2012, 38(8): 1370-1374

盲源分离是指在对源信号和传输通道都没有任何先验知识的情况下,根据信号的统计特性,仅利用观测信号来恢复源信号的过程.由于对信号源与传输通道的要求很少甚至不做要求,盲源分离在无线通信、生物医学工程、语音信号及数字图像处理等方面都有着重要的应用^[1].白化或盲去相关技术是盲源分离必要的预处理步骤,通过白化可以简化盲分离算法的分离过程,并能够辨识混合矩阵以及直接实现有色信号或非平稳信号的盲分离^[1-3],因此,对于白化算法的研究近年来逐渐得到学者们的广泛关注^[4-6].

现有的白化算法可分为批处理算法和自适应算法^[1],由于自适应白化算法复杂度较低,且可以实现混合信号的在线

处理,故得到了广泛的研究及应用^[7].但同大多自适应算法一样,自适应白化算法也存在收敛速度与稳态性能间的矛盾问题,其往往不能同时兼顾系统对于时变环境跟踪能力、算法收敛速度和稳态误差的要求.缓解此问题的常规途径是采用自适应变步长技术,即通过对步长参数进行在线调整来获取两者间的有效平衡^[8-11].但变步长算法往往会引入较多的调节参数,且部分算法易陷入局部极小点,不能实现对于突变环境的快速跟踪^[12].与常规方法不同,本文从另一个新角度出发,给出了缓解自适应白化算法系统收敛速度和稳态误差矛盾的可行方案.首先,算法通过融入动量项来加速系统的收敛过程,提高训练速度和避免算法陷入局部极小;然后,针对固定数值动量因子算法存在的问题,基于随机梯度方法对动量因子进行在线更新,以获取更快的收敛速度与较小的稳态误差.文中推导了动量因子的自适应更新规则,并利用仿真实验验证了提出算法在平稳和非平稳环境下的优良性能.

1 自适应白化

假设有 n 个相互独立的信源,通过未知的瞬时混合系统之后,得到 m 个观测信号,在不考虑外界噪声的情况下,观测信号与源信号的关系可表示为

$$\mathbf{x}(k) = A\mathbf{s}(k) \quad (1)$$

其中,向量 $\mathbf{x}(k) = [x_1(k), x_2(k), \dots, x_m(k)]^T$ 为观测信号, $\mathbf{s}(k) = [s_1(k), s_2(k), \dots, s_n(k)]^T$ 为源信号, $A \in \mathbf{R}^{m \times n}$ 表示混合矩阵,要求其列满秩,一般假设 $m = n$,即源信号与观测信号的个数相同.

所谓白化就是设法求得一满秩矩阵 W ,使得通过它可以从观测信号 $\mathbf{x}(k)$ 中获得白化信号 $\mathbf{y}(k)$:

$$\mathbf{y}(k) = W\mathbf{x}(k) \quad (2)$$

这里, $W \in \mathbf{R}^{n \times m}$ 表示白化矩阵,输出的白化向量 $\mathbf{y}(k)$ 满足以下特性:

$$E\{\mathbf{y}(k)\mathbf{y}^T(k)\} = WE\{\mathbf{x}(k)\mathbf{x}^T(k)\}W^T = I \quad (3)$$

其中, I 表示单位矩阵.令 $R_{\mathbf{x}\mathbf{x}}$ 为混合信号 $\mathbf{x}(k)$ 的自相关矩阵,并对其进行特征值分解,有:

$$R_{\mathbf{x}\mathbf{x}} = E\{\mathbf{x}(k)\mathbf{x}^T(k)\} = V_{\mathbf{x}}H_{\mathbf{x}}V_{\mathbf{x}}^T \quad (4)$$

其中, $V_{\mathbf{x}}$ 和 $H_{\mathbf{x}}$ 分别表示 $R_{\mathbf{x}\mathbf{x}}$ 的特征向量矩阵和特征值矩阵,则批处理白化算法对于 W 的求解结果为^[1]

$$W = H_{\mathbf{x}}^{-\frac{1}{2}}V_{\mathbf{x}}^T \quad (5)$$

批处理算法鲁棒性较好,但无法满足系统对于信号实时处理的要求,故对观测信号进行白化时,目前多采用文献 [1] 中具有自适应形式的白化算法,其结构如图 1 所示,该算法计算简单且白化及跟踪性能出色,其对白化矩阵 $W(k)$ 的估计可通过最小化式 (6) 的代价函数获得^[1].

收稿日期 2011-03-31 录用日期 2011-10-18
Manuscript received March 31, 2011; accepted October 18, 2011
国家自然科学基金 (61005021) 资助
Supported by National Natural Science Foundation of China (61005021)

本文责任编辑 封举富
Recommended by Associate Editor FENG Ju-Fu
1. 烟台大学光电信息科学技术学院 烟台 264005 2. 吉林大学通信工程学院 长春 130012

1. Institute of Science and Technology for Opto-electronic Information, Yantai University, Yantai 264005 2. College of Communication Engineering, Jilin University, Changchun 130012

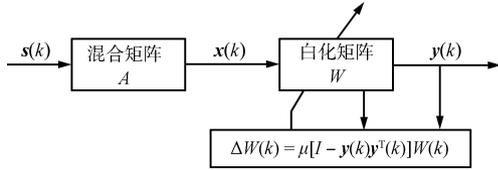


图 1 自适应白化算法结构图
Fig. 1 Basic scheme of adaptive whitening algorithm

$$J(k) = -\frac{1}{2} \left\{ \log \left[\det \left(W^T(k)W(k) \right) \right] - \sum_{i=1}^n E \left\{ y_i^2(k) \right\} \right\} \quad (6)$$

其中, $\det(Z)$ 表示对矩阵 Z 的行列式运算. 关于 $W(k)$ 对 $J(k)$ 的瞬时估计求导, 有:

$$\frac{\partial J(k)}{\partial W(k)} = - \left[I - \mathbf{y}(k)\mathbf{y}^T(k) \right] W(k) \quad (7)$$

基于式 (7) 可得自适应算法中白化矩阵 $W(k)$ 的更新公式为

$$W(k+1) = W(k) - \mu \frac{\partial J(k)}{\partial W(k)} = W(k) + \mu \left[I - \mathbf{y}(k)\mathbf{y}^T(k) \right] W(k) \quad (8)$$

其中, μ 为步长参数, 为保证算法收敛, 其取值应满足 $0 < \mu < \frac{2}{\sqrt{\lambda_{x-\max}(1 + \sqrt{\lambda_{y-\max}})}}$, 这里的 $\lambda_{x-\max}$ 和 $\lambda_{y-\max}$ 分别表示矩阵 $R_{\mathbf{x}\mathbf{x}}$ 与 $R_{\mathbf{y}\mathbf{y}} = E\{\mathbf{y}(k)\mathbf{y}^T(k)\}$ 最大的特征值.

为衡量算法的白化效果, 一般采用白化性能因子 (Performance factor, PF) 作为评价标准^[8], 其计算如下:

$$PF [G(k)] = \left\| I - G(k)G^T(k) \right\|_F^2 \quad (9)$$

其中, $G(k) = W(k)A$ 表示整个系统的传递矩阵. PF 值的大小反映了系统输出信号的白化效果, 其值较大时, 表示系统输出信号的相关性较大, 白化程度较差, 而 PF 越小, 则说明系统对于混合信号的白化效果越好.

2 融入动量项的自适应白化算法

同多数自适应处理算法一样, 前一节给出的自适应白化算法在信号处理过程中也无法兼顾收敛速度与稳态性能的要求. 为缓解该类问题, 人们在神经网络领域中通过结合动量项技术来改善误差反向传播 (Back propagation, BP) 算法的收敛性能^[13], 由于该技术具有计算量小, 能有效加速系统收敛速度和避免算法陷入局部极小点等优势, 其在稀疏信号处理、模糊控制等领域也得到了广泛应用^[14-15]. 因此, 本文考虑在 $W(k)$ 的更新过程中融入动量项, 以进一步加快算法的收敛速度, 具体操作如下式:

$$W(k+1) = W(k) - \mu \frac{\partial J(k)}{\partial W(k)} + \beta [W(k) - W(k-1)] = W(k) + \mu \left[I - \mathbf{y}(k)\mathbf{y}^T(k) \right] W(k) + \beta [W(k) - W(k-1)] \quad (10)$$

其中, $\beta [W(k) - W(k-1)]$ 表示动量项, β 为动量因子. 从式 (10) 中可以看出, 此算法利用了相邻时刻间白化矩阵的相关性, 在对 $W(k)$ 的更新过程中引入一个动量项, 其结果是

提高了算法的收敛速度和稳定性. 因为如果前一时刻的白化矩阵变化量较大, 也就是 $[W(k) - W(k-1)]$ 的数值较大时, 则 $W(k+1)$ 的修正量相应增加, 这样在白化矩阵的迭代过程中部分地加入这一修正量, 则能够起到提高跟踪速度、改善算法性能的目的.

通过式 (10), 可知算法在每次迭代中的更新量为

$$\Delta W(k) = W(k+1) - W(k) = \mu \sum_{i=1}^k \beta^{k-i} \left[I - \mathbf{y}(i)\mathbf{y}^T(i) \right] W(i) + \beta^k [W(1) - W(0)] \quad (11)$$

因此, 为保证算法的收敛性, β 的取值范围应为

$$\beta \in [0, 1) \quad (12)$$

图 2 给出的是一次白化实验中, 当动量因子 β 取不同数值时, 融入动量项后自适应白化算法的收敛曲线图. 其中 $\beta = 0$ 时, 表示算法没有加入动量项, 即回归至式 (8) 原始算法. 从图 2 结果不难看出, 融入动量项后, 算法的收敛速度能够获得有效的提升, 而且动量因子的数值越大, 系统的收敛及跟踪速度就越快. 但算法在获得较快收敛速度的同时, 将会不同程度地引入一定的稳态误差, 并且随着动量因子数值的增加, 算法的稳态性能就会越差.

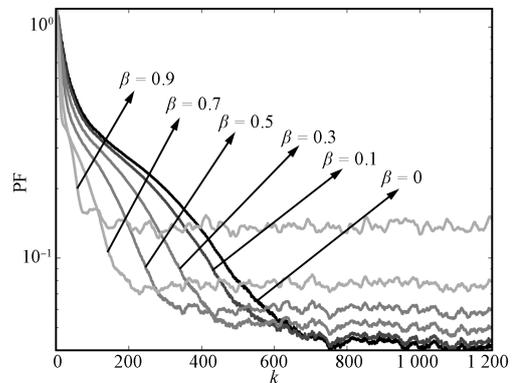


图 2 动量因子不同时算法的收敛曲线图
Fig. 2 Convergence curves with different momentum factors

3 变动物量因子算法

为了消除前一节定动量因子算法的不足, 本节考虑通过设定合理的自适应规则对动量因子进行在线更新, 使其在白化初期自适应地选择较大的数值来获取较快的收敛速度; 而在后期阶段, 则相应地选取数值较小的动量因子以获得较好的稳态效果, 从而有效避免固定数值动量因子的缺点, 确保算法在具有较快收敛速度的同时, 又保证输出信号的白化质量.

令动量因子数值可变, 则式 (10) 算法可写为

$$W(k+1) = W(k) + \mu \left[I - \mathbf{y}(k)\mathbf{y}^T(k) \right] W(k) + \beta(k) [W(k) - W(k-1)] \quad (13)$$

对于式 (13) 中动量因子提出采用下式的随机梯度方法进行自适应更新, 即沿代价函数最速下降的方向对其数值进行在

线调整:

$$\beta(k) = \beta(k-1) - \rho \nabla_{\beta} J(k) \Big|_{\beta=\beta(k-1)} \quad (14)$$

其中, ρ 为很小的常数, $J(k)$ 为式 (6) 中代价函数的瞬时估计. 定义两个矩阵的内积操作如下^[10, 16]:

$$\langle C, D \rangle = \text{tr}(C^T D) \quad (15)$$

其中, $\text{tr}(\cdot)$ 为矩阵的求迹运算, $\langle \cdot \rangle$ 表示内积过程. 根据式 (15), 式 (14) 中的随机梯度 $\nabla_{\beta} J(k) \Big|_{\beta=\beta(k-1)}$ 可以表示为

$$\begin{aligned} \nabla_{\beta} J(k) \Big|_{\beta=\beta(k-1)} &= \left\langle \frac{\partial J(k)}{\partial W(k)}, \frac{\partial W(k)}{\partial \beta(k-1)} \right\rangle = \\ &= \text{tr} \left[\left(\frac{\partial J(k)}{\partial W(k)} \right)^T \times \frac{\partial W(k)}{\partial \beta(k-1)} \right] \end{aligned} \quad (16)$$

由式 (7) 知:

$$\frac{\partial J(k)}{\partial W(k)} = - \left[I - \mathbf{y}(k) \mathbf{y}^T(k) \right] W(k) \quad (17)$$

根据式 (13), 有:

$$\begin{aligned} W(k) &= W(k-1) + \mu \left[I - \mathbf{y}(k-1) \mathbf{y}^T(k-1) \right] W(k-1) + \\ &+ \beta(k-1) [W(k-1) - W(k-2)] \end{aligned} \quad (18)$$

则

$$\frac{\partial W(k)}{\partial \beta(k-1)} = [W(k-1) - W(k-2)] \quad (19)$$

将式 (17) 和式 (19) 代入式 (16), 可得:

$$\begin{aligned} \nabla_{\beta} J(k) \Big|_{\beta=\beta(k-1)} &= \\ &= - \text{tr} \left\{ Q^T(k) \times [W(k-1) - W(k-2)] \right\} \end{aligned} \quad (20)$$

其中

$$Q(k) = \left[I - \mathbf{y}(k) \mathbf{y}^T(k) \right] W(k) \quad (21)$$

将式 (20) 代入式 (14), 最终可得动量因子的自适应更新规则:

$$\begin{aligned} \beta(k) &= \beta(k-1) + \\ &+ \rho \text{tr} \left\{ Q^T(k) \times [W(k-1) - W(k-2)] \right\} \end{aligned} \quad (22)$$

综合式 (17)、式 (19) 和式 (22) 可以看出, 当白化矩阵 $W(k)$ 的梯度负方向与 $[W(k-1) - W(k-2)]$ 相一致时, 说明系统处于收敛阶段, 算法通过式 (22) 可以自适应地提高动量因子的数值以加速收敛速度; 而当 $W(k)$ 的梯度负方向与 $[W(k-1) - W(k-2)]$ 不一致或出现更迭时, 表示白化系统正进入平稳阶段, 算法则相应地减小 $\beta(k)$ 以获取更好的稳态效果, 从而在整体上有效地缓解了算法收敛速度与稳态性能的矛盾问题. 式 (14) 中解的收敛性及 $\beta(k)$ 的收敛速度可见附录.

结合式 (13)、式 (21) 和式 (22) 可得本文算法的结构如图 3 所示, 其中 $\Delta W(k-1) = W(k) - W(k-1)$, $\Delta W(k-2) = W(k-1) - W(k-2)$.

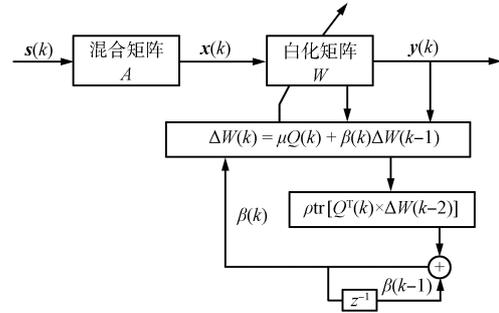


图 3 本文算法结构图

Fig. 3 Illustration of the proposed algorithm

4 仿真实验与结果分析

下面通过仿真实验来验证本文快速算法的整体性能. 为方便起见, 将式 (8) 算法和本文提出的式 (10) 算法及变动量因子算法分别简称为 AW (Adaptive whitening), MAW (Momentum term based adaptive whitening) 和 VMFAW (Variable momentum factor adaptive whitening).

仿真中所采用的 3 个源信号依次为: $s_1(k) = \sin(2\pi 200k)$, $s_2(k)$ = 区间 $[-0.5, 0.5]$ 内均匀分布的噪声, $s_3(k) = \text{sgn}(\sin(2\pi 100k))$. 混合矩阵 A 通过 Matlab 中函数 `randn(3)` 随机生成. 所有算法的步长大小均设置为 0.001, 两个动量因子数值不同的 MAW 算法, 其 β 值分别设为 0.8 和 0.2, VMFAW 算法中动量因子初值设为 $\beta(0) = 0.5$, 参数 $\rho = 0.4$. 在平稳环境即混合矩阵 A 的数值始终未发生变化的情况下, 同时运行 4 个算法, 100 次蒙特卡洛实验后所得算法的白化性能因子及 VMFAW 算法的动量因子变化曲线如图 4 和图 5 所示. 从图 4 中能够看出, 当动量因子选取较大数值时 ($\beta = 0.8$), MAW 算法的收敛速度较快, 其白化性能因子 PF 经 600 点左右迭代即可达到 $10^{-1.5}$. 但在稳态阶段, 其 PF 数值一直都维持在 $10^{-1.5}$ 附近, 说明该动量因子条件下算法对于混合信号的白化效果不够理想, 系统的稳态性能较差; 当动量因子设置为较小数值时 ($\beta = 0.2$), MAW 算法在平稳阶段拥有较好的稳态性能, 白化结果比较彻底, 但是收敛速度较慢, 其白化性能因子 PF 达到 $10^{-1.5}$ 时, 需经过近 2000 点左右的迭代, 从而影响系统的收敛速度和快速跟踪能力. 而本文提出的 VMFAW 算法则很好地避免了它们的缺点, 其在初始阶段具有更快的收敛速度, 而在稳态阶段拥有较好的稳态性能, 有效地消除了常规适应白化算法中收敛速度和信号白化质量间的矛盾关系. 从图 5 给出的 VMFAW 算法动量因子的自适应变化曲线中不难看出, 通过随机梯度设计的动量因子更新规则, 可以准确地快速地反映算法输出信号的白化状态, 自适应地调整动量因子数值的大小来获取更加优良的系统输出效果.

图 6 和图 7 给出的是在非平稳环境下各算法的收敛曲线及 VMFAW 算法动量因子的变化曲线图, 所有算法所用参数同平稳环境下的设置相同. 这里的非平稳环境即指混合系统在算法运行过程中会发生变化的情况, 一般多利用 A 为缓慢变化或在某一时刻点发生突变来对此环境加以模拟. 本文仿真中选择外界环境变化最为恶劣的情况 (A 在某时刻点发生突变) 来验证算法的性能. 从两图结果能够看出, 在此环境下, 算法皆可以有效地跟踪外界环境的变化, 稳定实现混合信号的自适应白化, 但相对于 AW 算法和 MAW 算法, VMFAW 算法拥有最快的收敛速度, 且其动量因子能够及时地调整自

身数值的大小以适应外界环境的变化情况。

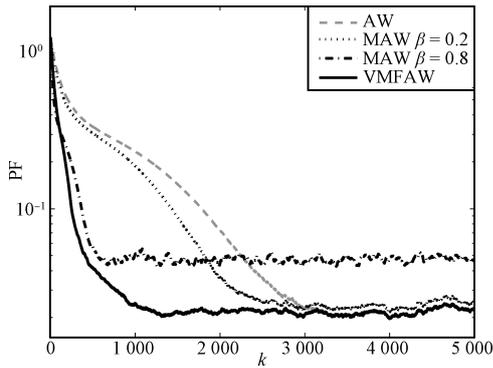


图 4 算法收敛曲线图 (平稳环境)
Fig. 4 Comparison of the convergence curves (stationary environment)

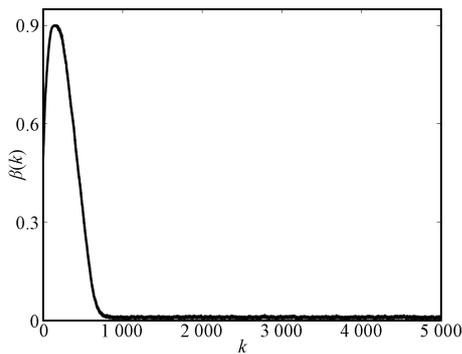


图 5 动量因子收敛曲线图 (平稳环境)
Fig. 5 Convergence curve of the momentum factor (stationary environment)

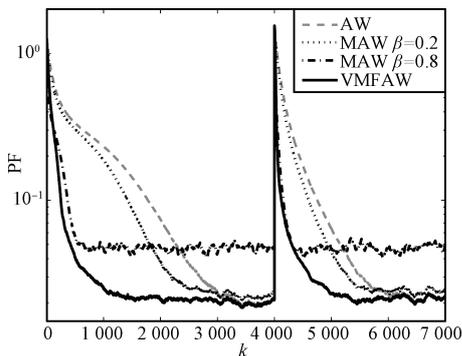


图 6 算法收敛曲线图 (非平稳环境)
Fig. 6 Comparison of the convergence curves (non-stationary environment)

除了算法的性能需要考虑外,在面向实际应用时,算法的计算复杂度也是一个重要的技术指标.本文以算法进行一次迭代所需的乘法次数来衡量 3 种算法的复杂度,其对比结果如表 1 所示.从表 1 中不难看出,VMFAW 算法所需乘法次数为 AW 算法的 2 倍,计算复杂度较高.另外,平稳环境下的仿真结果(100 次蒙特卡洛实验)也显示 3 种算法所消耗的 CPU 时间分别为 11.8045 秒、13.5190 秒和 24.4917 秒。

综合此两方面可以看出,VMFAW 算法所获取的优良性能是以增加计算量作为代价的。

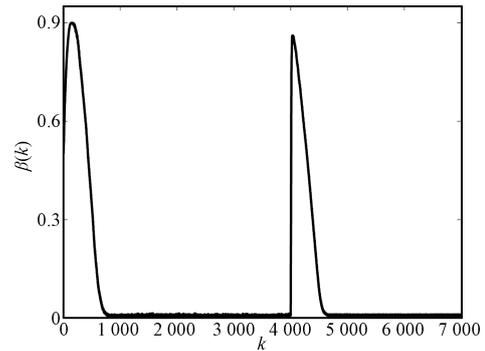


图 7 动量因子收敛曲线图 (非平稳环境)
Fig. 7 Convergence curve of the momentum factor (non-stationary environment)

表 1 三种算法的计算复杂度

Table 1 Computational complexities of the three algorithms

算法	AW	MAW	VMFAW
乘法次数	$N^3 + 2N^2$	$N^3 + 3N^2$	$2N^3 + 4N^2$

5 结论

为进一步提高自适应白化算法的收敛速度,本文提出了一种快速的自适应白化算法,该算法通过融入动量项,利用随机梯度方法来自适应地调整动量因子,从而以新的角度给出了加速系统收敛速度的可行方案.但从文中的理论推导和仿真结果也能够看出,VMFAW 算法虽有效地提升了系统的收敛速度,但其在稳态过程的性能仍受到固定步长参数的限定,即当 3 种算法的步长参数设置相同时,无论动量因子 $\beta(k)$ 如何变化,VMFAW 算法最终的稳态误差仍不会优于 AW 算法.因此,后续的研究工作将考虑把本文算法与变步长技术相结合,综合两者的优势以获取更加优良的白化效果,同时将此变动物量项思想引入自适应盲源分离算法中,从而有可能成为自适应变步长理论的有益补充。

附录

在式 (10) 中引入动量项后,白化系统的代价函数为 (为便于数学表示,考虑连续时刻情况)

$$J'(t) = \frac{1}{2}m \frac{dW^T}{dt} \frac{dW}{dt} + J(t) \tag{A1}$$

此处 m 与动量因子 β 的关系为 $\beta = \frac{m}{m+\mu\Delta t}$, $J(t)$ 为式 (14) 中代价函数 $J(k)$ 的连续形式,式 (A1) 为单调下降的 Lyapunov 函数^[17].由于在平稳状态下 $J'(t)$ 与 $J(t)$ 中 W 的各个元素均不会发生变化,即在 W 收敛的末段,两个代价函数可以认为是相当的^[17],故式 (14) 中可直接利用 $J(t)$ 代替 $J'(t)$ 对 β 进行求导,此时有:

$$\frac{d\beta}{dt} = \frac{\beta(t) - \beta(t - \Delta t)}{\Delta t} = -\rho \nabla_{\beta} J'(t) = -\rho \nabla_{\beta} J(t) \tag{A2}$$

在极值点 β_0 附近对 $\nabla_{\beta} J(t)$ 利用泰勒级数展开,上式可近似为

$$\frac{\beta(t) - \beta(t - \Delta t)}{\Delta t} \approx -\rho h [\beta(t - \Delta t) - \beta_0] \tag{A3}$$

其中, $h = \frac{d^2 J(t)}{d\beta^2} \Big|_{\beta_0}$ 为一正实数^[17].不失一般性,设 $\beta_0 = 0$,稍作整

理后可得上式的离散形式为

$$\beta(k) = (1 - \rho h) \beta(k - 1) \quad (\text{A4})$$

假设 $\beta(k)$ 的初始值为 $\beta(0)$, 易得上式的解为

$$\beta(k) = (1 - \rho h)^k \beta(0) \quad (\text{A5})$$

为保证收敛性, 式 (A5) 中几何级数的公比须满足 $|1 - \rho h| < 1$, 即 ρ 的取值范围应为

$$0 < \rho < \frac{2}{h} \quad (\text{A6})$$

通过式 (A5) 可知, $\beta(k)$ 的收敛速度取决于 ρ 的大小; 在满足式 (A6) 的条件下, ρ 的数值越大, 则 $\beta(k)$ 的收敛速度越快。

References

- Cichocki A, Amari S. *Adaptive Blind Signal and Image Processing: Learning Algorithms and Application*. New York: John Wiley and Sons, 2002. 129–175
- Choi S, Cichocki A, Park H M, Lee S Y. Blind source separation and independent component analysis: a review. *Neural Information Processing — Letters and Reviews*, 2005, **6**(1): 1–57
- Tang Ying, Li Jian-Ping. A new algorithm of ICA: using the parametrized orthogonal matrixes of any dimensions. *Acta Automatica Sinica*, 2008, **34**(1): 31–39
(汤影, 李建平. 利用参数表示任意维数正交矩阵的 ICA 新算法. 自动化学报, 2008, **34**(1): 31–39)
- Feng D Z, Zheng W X, Cichochi A. Matrix-group algorithm via improved whitening process for extracting statistically independent sources from array signals. *IEEE Transactions on Signal Processing*, 2007, **55**(3): 962–977
- Mollah M N, Eguchi S, Minami M. Robust prewhitening for ICA by minimizing β -divergence and its application to fastICA. *Neural Processing Letters*, 2007, **25**(2): 91–110
- Kocinski J, Libiszewski P, Sek A. Spatial efficiency of blind source separation based on decorrelation — subjective and objective assessment. *Speech Communication*, 2011, **53**(3): 390–402
- Coviello C M, Yoon P A, Sibul L H. Source separation and tracking for time-varying systems. *IEEE Transactions on Aerospace and Electronic Systems*, 2008, **44**(3): 1198–1214
- Douglas S C, Cichocki A. Adaptive step size techniques for decorrelation and blind source separation. In: *Proceedings of the 32th Asilomar Conference on Signals, Systems, and Computers*. Pacific Grove, USA: IEEE, 1998. 1191–1195
- Ou S F, Zhao X H, Gao Y. Variable step size technique for adaptive blind decorrelation. In: *Proceedings of the 8th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing*. Qingdao, China: IEEE, 2007. 823–826
- Yuan L X, Wang W W, Chambers J A. Variable step-size sign natural gradient algorithm for sequential blind source separation. *IEEE Signal Processing Letters*, 2005, **12**(8): 589–592
- Hsieh S T, Sun T Y, Lin C L, Liu C C. Effective learning rate adjustment of blind source separation based on an improved particle swarm optimizer. *IEEE Transactions on Evolutionary Computation*, 2008, **12**(2): 242–251
- Arenas-Garcia J, Gomez-Verdejo V, Figueiras-Vidal A R. New algorithms for improved adaptive convex combination of LMS transversal filters. *IEEE Transactions on Instrumentation and Measurement*, 2005, **54**(6): 2239–2249
- Mkadem F, Boumaiza S. Physically inspired neural network model for RF power amplifier behavioral modeling and digital predistortion. *IEEE Transactions on Microwave Theory and Techniques*, 2011, **59**(4): 913–923
- Zhang Z, Cui C, Liu K. Iterative hard-threshold algorithm with momentum. *Electronics Letters*, 2011, **47**(4): 257–259
- Xia C L, Guo C, Shi T N. A neural-network-identifier and fuzzy-controller-based algorithm for dynamic decoupling control of permanent-magnet spherical motor. *IEEE Transactions on Industrial Electronics*, 2010, **57**(8): 2868–2878
- Craven B D, Mond B. Linear programming with matrix variables. *Linear Algebra and Its Applications*, 1981, **38**: 73–80
- Qian N. On the momentum term in gradient descent learning algorithms. *Neural Networks*, 1999, **12**(1): 145–151

欧世峰 烟台大学光电信息科学技术学院副教授. 主要研究方向为盲信号处理与语音信号处理. E-mail: ousfeng@126.com

(OU Shi-Feng Associate professor at the Institute of Science and Technology for Opto-electronic Information, Yantai University. His research interest covers blind signal processing and speech signal processing.)

高颖 烟台大学光电信息科学技术学院讲师. 主要研究方向为盲信号处理. 本文通信作者. E-mail: claragaoying@126.com

(GAO Ying Lecturer at the Institute of Science and Technology for Opto-electronic Information, Yantai University. Her main research interest is blind signal processing. Corresponding author of this paper.)

赵晓晖 吉林大学通信工程学院教授. 主要研究方向为自适应信号处理与智能信息处理. E-mail: xhzhao@jlu.edu.cn

(ZHAO Xiao-Hui Professor at the College of Communication Engineering, Jilin University. His research interest covers adaptive signal processing and intelligent information processing.)