

文章编号: 1001-0920(2012)02-0317-04

非正交联合对角化的罚函数粒子群算法

徐洪涛¹, 王跃钢¹, 杨波¹, 陈国栋²

(1. 第二炮兵工程学院 304 教研室, 西安 710025; 2. 第二炮兵工程设计所, 北京 100011)

摘要: 针对多个矩阵近似联合对角化盲分离问题, 提出一种新的非正交近似联合对角化算法. 首先采用罚函数法将联合对角化的非线性约束优化模型转化为无约束优化模型; 其次将粒子群优化算法引入无约束优化模型中实现目标函数的最优化, 从而完成矩阵组的联合对角化. 分析了惩罚因子的更新策略及算法的收敛性能, 并设计仿真实验进行对比分析以检验算法解决实际盲分离问题的能力.

关键词: 盲源分离; 非正交联合对角化; 惩罚函数; 惩罚因子; 粒子群优化

中图分类号: TP911

文献标识码: A

Non-orthogonal joint diagonalization based on penalty function particle swarm optimization

XU Hong-tao¹, WANG Yue-gang¹, YANG Bo¹, CHEN Guo-dong²

(1. The 304 Staff Room, The Second Artillery Engineering College, Xi'an 710025, China; 2. The Second Artillery Engineering Design Research Institute, Beijing 100011, China. Correspondent: XU Hong-tao, E-mail: xuhongtaopp@163.com)

Abstract: Aim to approximately joint diagonalization(JAD) method in solving blind source separation(BSS) problem, a novel non-orthogonal JAD(NJAD) algorithm is proposed. By using the penalty function, the constrained optimization problem of NJAD can be changed to an unconstrained optimization problem. Then the particle swarm algorithm is used to realize the optimization of the criterion. The updated strategy of penalty factor is given and the convergence of the algorithm is analyzed. Computer simulations show the capability of the proposed algorithm in solving the BSS problems compared to another one.

Key words: blind source separation; non-orthogonal joint diagonalization; penalty function; penalty factor; particle swarm optimization

1 引言

矩阵近似联合对角化是解决盲源分离的有效方法, 在谐波恢复^[1]、阵列信号处理^[2]以及多输入多输出系统盲辨识^[3]等领域均得到广泛应用. 考虑 K 个 $n \times n$ 矩阵构成的矩阵组 $M = \{M_1, M_2, \dots, M_K\}$, 联合对角化的目标是寻求一个非奇异对角化矩阵 W 以及 K 个相应的对角矩阵 $A = \{A_1, A_2, \dots, A_K\}$, 满足 $WM_kW^T = A_k$.

早期的对角化方法主要有 JADE 算法^[4], SOBI 算法^[5]等正交对角化方法. 此类方法需要对信号进行预白化, 但是该过程会影响盲分离精度^[6]. 因此, 涌现出许多无须对信号进行预白化的非正交联合对角化方

法, 典型算法有 ACDC 迭代法等^[7-10]. 为了避免算法陷入零解或奇异解, 必须对 W 施加一定的约束条件, 如 $(WM_0W^T)_{ii} = 1$, 其中 M_0 为正定矩阵, 可以是观测信号的协方差矩阵等. 因此, 从最优化理论上讲, 非正交联合对角化实质上是一类具有等式约束的非线性最优化问题, 可以用约束最优化方法解决.

本文研究一种基于罚函数思想的非正交联合对角化粒子群优化算法. 首先通过对原目标函数施加相应的约束惩罚项, 构建新的无约束优化目标函数, 将较为棘手的非线性约束最优化问题转化为相对简单的无约束优化问题; 然后运用粒子群优化算法完成新目标函数的优化, 进而实现整个矩阵组的对角化过程.

收稿日期: 2010-09-19; 修回日期: 2010-11-27.

基金项目: 国家安全重大基础研究项目(61355020301).

作者简介: 徐洪涛(1982-), 男, 博士生, 从事导弹自动检测技术与故障诊断的研究; 王跃钢(1958-), 男, 教授, 博士生导师, 从事导弹导航、制导与控制等研究.

2 联合对角化盲分离

2.1 盲分离模型

假定源信号与观测信号的维数相同且为 n , 记源信号向量为 $s = \{s_1, s_2, \dots, s_K\}$, 观测向量为 $x = \{x_1, x_2, \dots, x_K\}$, 信号混合过程可以用下式表示:

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

其中 $A = \{a_{ij}\}$ 为 $n \times n$ 满秩混合矩阵. 盲分离的目标是寻找一非奇异分离矩阵 W , 使得下式成立:

$$\mathbf{y}(t) = W\mathbf{x}(t) = WAs(t) = DP\mathbf{s}(t) \quad (2)$$

其中: D 为 $n \times n$ 阶非奇异对角矩阵; P 为 $n \times n$ 阶置换矩阵.

根据对源信号假设^[5], 源信号(观测信号)二阶统计量矩阵一般满足

$$M_s(\tau) = \text{diag}\{\sigma_1(\tau) \sigma_2(\tau) \cdots \sigma_n(\tau)\}, \quad (3)$$

$$M_x(\tau) = A \text{diag}\{\sigma_1(\tau) \sigma_2(\tau) \cdots \sigma_n(\tau)\} A^T. \quad (4)$$

2.2 近似联合对角化

矩阵对角化即是寻找矩阵 W , 使下式成立:

$$\begin{aligned} WM_x(\tau)W^T &= WAM_s(\tau)(WA)^T = \\ DPM_s(\tau)(DP)^T &= \Lambda(\tau). \end{aligned} \quad (5)$$

从而, n 个源信号的盲分离问题可通过对 K 个矩阵 $\{M(\tau_k), k = 1, 2, \dots, K\}$ 的近似联合对角化来实现, 其基于最小二乘思想的目标函数为

$$\min L(W) = \sum_k \alpha_k \sum_{i \neq j} (WM(\tau_k)W^T)_{ij}^2. \quad (6)$$

其中: $(\bullet)_{ij}$ 代表矩阵 \bullet 的第 i 行、第 j 列的元素; α_k 是矩阵 $M(\tau_k)$ 权系数, 一般情况下均为 $1/K$.

如上文所述, 为了避免算法陷入奇异解或无意义解, 必须对 W 施加相应的约束条件, 即

$$(WM(\tau_0)W^T)_{ii} = 1, \quad (7)$$

其中 $M(\tau_0)$ 为正定矩阵, 如观测信号的协方差阵.

3 惩罚函数法及粒子群实现

3.1 约束优化模型

由 2.2 节可知, 近似联合对角化问题是在满足约束条件(7)的情况下, 关于目标函数(6)求最小值, 结合优化理论和式(6), (7), 可建立如下的约束优化模型:

$$\begin{cases} \min L(W) = \sum_k \alpha_k \sum_{i \neq j} (WM(\tau_k)W^T)_{ij}^2, \\ \text{s.t. } (WM(\tau_0)W^T)_{ii} = 1, i, j = 1, 2, \dots, n. \end{cases} \quad (8)$$

模型(8)为一类带有非线性等式约束的非线性函数最优化问题, 求解此类问题的一种有效方法是将非线性约束优化问题转化为无约束优化问题.

3.2 惩罚函数法

惩罚函数法是一种间接处理约束优化问题的有

效方法^[11], 其基本思想为: 给原目标函数加上一项由约束函数构成的惩罚项, 当某个解不可行(不满足约束条件)时, 需对其处以惩罚; 但当某解可行(满足约束条件)时, 则不作任何惩罚, 从而迫使迭代点不断逼近并最终位于可行域内, 以便找到原约束问题的最优解.

构造二次平方“外点罚函数”惩罚项如下:

$$P(W) = \|\text{diag}(WM(\tau_0)W^T - I)\|_1^2. \quad (9)$$

其中: $P(W)$ 称为“惩罚项”, $\|\bullet\|_1$ 为矩阵 1-范数. 结合式(6)和(9), 建立模型(8)的无约束优化目标函数为

$$\begin{aligned} F(W, \lambda) &= L(W) + \lambda P(W) = \\ &= \sum_k \alpha_k \sum_{i \neq j} (WM(\tau_k)W^T)_{ij}^2 + \\ &= \lambda \|\text{diag}(WM(\tau_0)W^T - I)\|_1^2, \end{aligned} \quad (10)$$

其中 $\lambda > 0$ 为惩罚因子.

3.3 粒子群算法

基于随机学习理论的粒子群优化算法在非线性、不可微以及非凸函数的优化问题中得到了广泛的应用^[12], 运用该理论可以有效解决式(10)的优化问题.

假设在一个 d 维目标搜索空间中, 由 m 个粒子组成一个种群, 其中第 i 个粒子在空间里的位置为 $X_i = (x_{i1} \cdots x_{id})^T, i = 1, 2, \dots, m$, 飞行速度为 $V_i = (v_{i1} \cdots v_{id})^T$, 历次飞行中个体最优位置为 $X_i^{\text{pbest}} = (x_{i1}^{\text{pbest}} \cdots x_{id}^{\text{pbest}})^T$, 群体所有粒子最优位置为 X_g^{pbest} , 在每一次迭代中, 第 i 个粒子的速度和位置更新规则如下:

$$\begin{aligned} V_i &= \omega V_i + c_1 r_1 (X_i^{\text{pbest}} - X_i) + \\ &= c_2 r_2 (X_g^{\text{pbest}} - X_i), \end{aligned} \quad (11)$$

$$X_i = X_i + V_i. \quad (12)$$

式中: r_1 和 r_2 为服从 $(0, 1)$ 均匀分布的随机数; c_1 和 c_2 为加速度系数; ω 为惯性权重, 本文采用一般的线性递减规则调节惯性权重.

罚函数的粒子群优化关键在于找到一个合适的惩罚因子 λ , 使得原目标函数与惩罚函数之间达到最佳平衡. 如果 λ 太小, 约束条件违反强度不够, 则找到的解可能违反了约束条件; 如果 λ 太大, 则会抑制算法对不可行解空间的搜索, 该问题在可行空间由一些不相交的可行区域组成的情况下尤为严重. 为了确保能够快速准确地找到问题的最优解, 算法采用自适应调整惩罚因子, 在第 k 次迭代过程中有

$$\lambda^{(k)} = \begin{cases} 1.5 * \lambda^{(k-1)}, & P(W) > \varepsilon; \\ 0.75 * \lambda^{(k-1)}, & P(W) \leq \varepsilon. \end{cases} \quad (13)$$

其中 ε 为约束条件精度.

3.4 罚函数的粒子群优化

根据3.2和3.3节的讨论,下面给出近似联合对角化的罚函数粒子群优化算法的实施步骤:

Step 1: 参数设置: 惩罚因子更新次数 N_0 , 粒子迭代次数 N , 粒子数 m , 最小惯性权重 ω_{\min} , 最大惯性权重 ω_{\max} , 速度限制 v_{\max} , 加速度系数 c_1 和 c_2 , 约束条件精度 ε , 初始惩罚因子 λ_0 , $i = 0$, $k = 1$.

Step 2: 随机生成 m 个 $n \times n$ 阶初始矩阵粒子 $W_j (j = 1, 2, \dots, m)$ 以及相对应的 $n \times n$ 阶初始飞行速度矩阵 $V_j, j = 1, 2, \dots, m$.

Step 3: 根据式(10)计算各个粒子的罚函数适应度值, 找到粒子个体最优矩阵 W_j^{pbest} 以及种群最优矩阵 W_g^{pbest} .

Step 4: $i = i + 1$, 根据更新规则(11), (12)完成粒子的速度和位置更新. 返回 Step 3 继续, 直至 $i = N$ 转至 Step 5.

Step 5: $k = k + 1$, 按照式(13)更新惩罚因子, 重复 Step 3~Step 4, 直至 $k = N_0$ 迭代结束, 输出最优解以及相应的 W_{opt} .

3.5 收敛性分析

引理 1 $F(W, \lambda)$ 是式(8)对应的罚函数, 且有 $\lambda^{(k+1)} > \lambda^{(k)}$. 如果 $F(W, \lambda^{(k+1)})$ 和 $F(W, \lambda^{(k)})$ 分别在 $W^{(k+1)}$ 与 $W^{(k)}$ 处达到无约束全局最小, 则有

$$F(W^{(k+1)}, \lambda^{(k+1)}) \geq F(W^{(k)}, \lambda^{(k)}), \quad (14)$$

$$P(W^{(k)}) \geq P(W^{(k+1)}), \quad (15)$$

$$L(W^{(k+1)}) \geq L(W^{(k)}). \quad (16)$$

引理 1 的证明参见文献[12](定理 7.2.1). 利用上述引理的结论, 可以证明本文算法是全局收敛的.

假设 $W^{(k)}$ 是序列 $\lambda^{(k)}$ 对应罚函数 $F(W, \lambda^{(k)})$ 的整体无约束极小点, 则有 $W^{(k)} \rightarrow \bar{W}$, 同时设 W^* 为式(10)的一个全局最优解, 故有 $P(W^*) = 0$, 从而由上述引理所述的 $F(W, \lambda^{(k)})$ 的单调递增可知

$$L(W^*) = L(W^*) + \lambda^{(k)} P(W^*) \geq F(W^{(k)}, \lambda^{(k)}). \quad (17)$$

这表明 $\{F(W^{(k)}, \lambda^{(k)})\}$ 为一单调增且有上界的序列, 则极限存在, 设为 \hat{F} . 另一方面, 因

$$L(W^{(k)}) \leq F(W^{(k)}, \lambda^{(k)}) \leq L(W^*), \quad (18)$$

同样由 $\{L(W^{(k)})\}$ 的单调增特性可知其必收敛到某一极限 \hat{L} . 于是有

$$\begin{aligned} \lim_{k \rightarrow \infty} \lambda^{(k)} P(W^{(k)}) &= \\ \lim_{k \rightarrow \infty} (F(W^{(k)}, \lambda^{(k)}) - L(W^{(k)})) &= \hat{F} - \hat{L}. \end{aligned} \quad (19)$$

因 $\lim_{k \rightarrow \infty} P(W^{(k)}) = 0$, 从而由 $W^{(k)} \rightarrow \bar{W}$ 与 $P(W)$ 的连续性有 $P(\bar{W}) = 0$, 即 \bar{W} 为式(10)的可行解. 因此, 由 W^* 的最优性可知 $L(W^*) \leq L(\bar{W})$, 但由上述内容

可知

$$L(\bar{W}) = \lim_{k \rightarrow \infty} (L(W^{(k)})) \leq L(W^*). \quad (20)$$

所以 $L(\bar{W}) = L(W^*)$, 结合 \bar{W} 的可行性表明了 \bar{W} 必为式(10)的一个全局最优解, 即算法是收敛的.

4 仿真实验

4.1 本文算法仿真

随机产生 $K = 11$ 个 $n \times n (n = 3)$ 阶对角矩阵 $D_{(k)} = \{D_1 \cdots D_K\}$, 每个矩阵的对角线元素为服从标准正态分布的随机数. 混合矩阵 A 为服从 $(0, 1)$ 间均匀分布的随机矩阵, 令 $M_k = AD_kA^T$, 生成相应的 K 个相关矩阵 $M_{(k)}$.

令 $\alpha_k = 1/K, k = 1, 2, \dots, K$, 采用如下的性能指标衡量对角化误差:

$$E = \frac{L(W)}{n^2 - n}. \quad (21)$$

参数设置: $N_0 = 100, N = 20, m = 80, \omega_{\min} = 0.4, \omega_{\max} = 0.9, v_{\max} = 0.3, c_1 = 2.1, c_2 = 2, \lambda_0 = 1, \varepsilon = 10^{-6}$. 对 3.4 节所提出的算法重复运行 5 次, 得到如图 1 所示的收敛曲线.

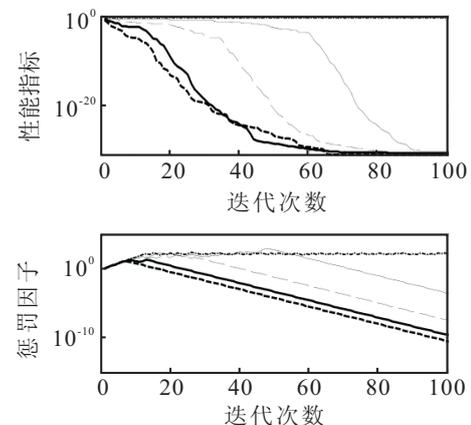


图 1 性能指标及惩罚因子收敛曲线

1) 5 次实验中性能指标有 4 次收敛在 10^{-30} 附近, 说明本文算法完全能够达到较高的精度, 剩余误差是由计算机计算精度引起的.

2) 算法的成功在于惩罚因子表现出一个较理想的收敛过程: 迭代开始阶段逐步增大, 允许具有一定的振荡现象, 此过程表明算法在较大的范围内探索所有可行解; 随后惩罚因子逐渐收敛至一个很小的正数, 该过程表明算法在可行解范围内寻找最优解, 而且一个小的惩罚因子更有利于最优解的搜索.

4.2 与 SOBI 算法比较

实验将本文算法与典型的 SOBI 算法^[5]进行对比分析, 源信号为某飞行器实际振动数据 $\mathbf{s}(t) = \{s_1(t), s_2(t), s_3(t)\}$, 采样频率为 5 120 Hz, 数据总长度 L 为 4 096. 混合过程同 4.1 节, 混合信号的相关函数矩阵的无偏估计为

$$M_k = \frac{1}{L - \tau_k} (\mathbf{x}(t)\mathbf{x}^T(t + \tau_k)). \quad (22)$$

其中: $k = 1, 2, \dots, K (K = 11), \tau_k = 2(k - 1)$.

实验采用文献 [10] 中关于盲分离性能指数 PI 的定义来评价算法的分离性能, 取对数 $20 * \lg PI$. 图 2 给出了在不同的信噪比环境下两种算法的最终分离性能曲线, 显然本文算法较 SOBI 算法具有更好的分离性能.

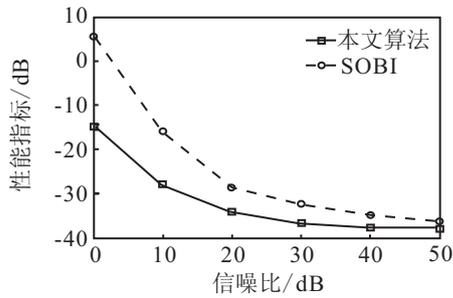


图 2 两种算法的信噪比-性能曲线

5 结 论

本文研究了一种新的非正交联合对角化方法, 并成功解决了飞行器振动信号的盲分离问题. 首先详细分析了非正交近似联合对角化目标函数及相应的约束条件, 提出了解决对角化问题的非线性约束优化模型. 通过惩罚函数法将约束优化模型转化为无约束优化模型, 方便了优化问题的求解. 最后采用粒子群优化算法实现无约束优化模型的求解, 成功解决了带有约束的非正交联合对角化问题. 相比正交对角化的 SOBI 算法, 本文算法能够成功解决实际信号的盲分离问题, 且具有更加优异的性能.

参考文献(References)

- [1] 聂卫科, 冯大政, 张斌. 谐波恢复的联合对角化算法[J]. 电子与信息学报, 2009, 31(2): 331-334.
(Nie W K, Feng D Z, Zhang B. Joint diagonalization algorithm for harmonic retrieval[J]. J of Electronics Information Technology, 2009, 31(2): 331-334.)
- [2] Zeng Wenjun, Li Xilin, Zhang Xianda. Direction-of-arrival estimation based on the joint diagonalization structure of

- multiple fourth-order cumulant matrices[J]. IEEE Signal Processing Letters. 2009, 16(3): 164-167.
- [3] McNeill S I, Zimmerman D C. A framework for blind modal identification using joint approximate diagonalization[J]. Mechanical Systems and Signal Processing, 2008, 22(7): 1526-1548.
- [4] Cardoso J F, Souloumiac A. Blind beamforming for non-Gaussian signals[J]. IEE Proceedings-F. 1993, 140(6): 362-370.
- [5] Belouchrani A, Abed-Meraim K, Cardoso J F, et al. A blind source separation technique using second-order statistics[J]. IEEE Trans on Signal Processing. 1997, 45(2): 434-444.
- [6] Li X L, Zhang X D. Nonorthogonal joint diagonalization free of degenerate solution[J]. IEEE Trans on Signal Processing, 2007, 55(5): 1803-1814.
- [7] Yeredor A. Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation[J]. IEEE Trans on Signal Processing, 2002, 50(7): 1545-1553.
- [8] Voland R, Obermayer K. Quadratic optimization for simultaneous matrix diagonalization[J]. IEEE Trans on Signal Processing, 2006, 54(9): 3270-3278.
- [9] Ziehe A, Laskov P, Nolte G, et al. A fast algorithm for joint diagonalization with non-orthogonal transformations and its application to blind source separation[J]. J of Machine Learning Research, 2004, 5(12): 777-800.
- [10] Souloumiac A. Nonorthogonal joint diagonalization by combining givens and hyperbolic rotations[J]. IEEE Trans on Signal Processing. 2009, 57(6): 2222-2231.
- [11] 徐成贤, 陈志平, 李乃成. 近代优化方法[M]. 北京: 科学出版社, 2002: 285-320.
(Xu C X, Chen Z P, Li N C. Modern optimization method[M]. Beijing: Scientific Publishing House, 2002: 285-320.)
- [12] Sedlaczek K, Eberhard P. Using augmented lagrangian particle swarm optimization for constrained problems in engineering[J]. Structural and Multidisciplinary Optimization, 2006, 32(4): 277-286.