

非对称非正交快速联合对角化算法

张伟涛¹ 楼顺天¹ 张延良^{1,2}

摘要 针对非对称联合对角化算法收敛速度慢以及有可能收敛到奇异解的问题, 首先提出一种基于最小二乘的非对称代价函数, 该代价函数在最小二乘标准的基础上增加了使对角化矩阵非奇异的约束项, 以保证算法不会收敛到奇异解. 然后利用一种循环最小化技术来优化提出的代价函数, 得到一种非对称非正交快速联合对角化算法. 算法的性能分析证明, 该算法不仅全局渐近收敛, 而且具有不变性. 左右对角化矩阵的关系也证明了非对称联合对角化的一般性. 实验仿真表明, 与原非对称联合对角化算法相比, 提出的算法收敛速度更快, 而且可以显著降低干扰信号比.

关键词 联合对角化, 盲信源分离, 最小二乘标准, 循环最小化, 非奇异

DOI 10.3724/SP.J.1004.2010.00829

Non-symmetrical Non-orthogonal Fast Joint Diagonalization Algorithm

ZHANG Wei-Tao¹ LOU Shun-Tian¹ ZHANG Yan-Liang^{1,2}

Abstract To overcome the drawbacks of slow convergence speed and possible singular solutions of existing non-symmetrical joint diagonalization algorithm, we first present a least-squares criteria based non-symmetrical cost function for joint diagonalization, in which a penalty term is added to the classical least-squares criteria to avoid singular solutions. Then a non-symmetrical non-orthogonal fast joint diagonalization algorithm is developed by using a cyclic minimizer technique. The performance analysis shows that the present algorithm globally asymptotically converges to the stable stationary point and has the invariance property. The relation between left and right diagonalization matrices is also investigated to show that the non-symmetrical joint diagonalization is a more general form for joint diagonalization problem. The simulation results show that the proposed algorithm converges faster than the original algorithm, and that the interference to signal ratio (ISR) is also significantly improved.

Key words Joint diagonalization, blind source separation, least-squares criteria, cyclic minimizer, nonsingular

多个矩阵的同时对角化是一种非常重要的数值计算问题, 它至少存在于三个不同的应用之中: 1) 公共主分量估计^[1-2]; 2) 盲波束形成^[3]; 3) 独立分量分析和盲信源分离^[4-8]. 公共主分量估计是由 Flury 提出的一种统计学问题, 它研究在多个随机过程具有共同主分量的情况下, 其公共主分量的估计问题. 通过对各随机过程的样本协方差矩阵进行联合对角化是估计公共特征向量的主要方法. 波束形成问题的关键在于阵列响应矩阵的估计, 在没有阵列流形先验信息的情况下, 只能对阵列响应矩阵进行盲辨识, 而多个特征矩阵的联合对角化是唯一稳健的盲估计方法. 在盲信源分离中, 观测信号来自于多个独立源信号的未知混合, 辨识分离矩阵是源信号恢复的关键. 观测信号一般用来构造一个目标矩阵的集合 (该矩阵一般由观测信号的二阶或高阶统

计量产生), 而分离矩阵则由目标矩阵联合对角化产生的对角化矩阵给出. 可以看出, 很多应用问题都采用联合对角化技术作为其主要的算法工具, 因此联合对角化算法的研究有着重要的应用价值.

现有的联合对角化算法大体可以分为两种: 正交联合对角化算法和非正交联合对角化算法. 其中, 正交联合对角化算法^[3-5] 约束对角化矩阵为正交矩阵, 而非正交联合对角化算法^[6-15] 并无此约束, 因此, 其对角化矩阵为普通矩阵. Cardoso 首先提出了 Givens 旋转方法^[3] 来解决正交联合对角化问题, 近来文献 [4] 也提出了一种基于 Householder 变换的高效正交联合对角化算法. 文献 [5] 中盲信源分离二阶盲辨识 (Blind identification using second order statistics, SOBI) 算法就是在正交联合对角化算法的基础上提出的. 然而, 对角化矩阵的正交性约束事实上破坏了最小二乘标准, 最终影响了联合对角化的性能. 因此文献 [6] 提出了最小二乘代价函数并得到了一种非正交联合对角化算法 — 列交替更新与对角化 (Alternating columns diagonal centers, ACDC), 然而该算法要交替更新两组待定参数, 因此其收敛速度较慢. 文献 [7] 中采用梯度下降来优化最小二乘代价函数. 为了得到较好的稳态性能, 其迭代步长一般很小, 因此算法很难快速收敛. 文献

收稿日期 2008-11-25 录用日期 2010-01-13
Manuscript received November 25, 2008; accepted January 13, 2010

国家自然科学基金 (60775013) 资助
Supported by National Natural Science Foundation of China (60775013)

1. 西安电子科技大学电子工程学院 西安 710071 2. 河南理工大学
计算机科学与技术学院 焦作 454001

1. School of Electronic Engineering, Xidian University, Xi'an
710071 2. Department of Computer Science and Technology,
Henan Polytechnic University, Jiaozuo 454001

[10–11] 使用了不同于最小二乘标准的信息论代价函数来解决联合对角化问题, 其算法收敛较快, 稳态性能较好, 然而该算法只能应用于对称正定的矩阵集合, 这使该算法的应用受到了极大的限制. 文献 [12] 使用子空间匹配得到了一种收敛速度极快的联合对角化算法, 然而算法极大的计算量 ($O(KN^6)$) 以及对初值的较高要求使其很难应用于高速实时信号处理. 文献 [14] 指出, 优化最小二乘代价函数的非正交联合对角化算法^[6, 8, 13] 都有可能收敛到奇异解, 这是非正交联合对角化的一个不足之处. 需要指出的是文献 [6–14] 中的算法都使用了对称的联合对角化结构, 即左右对角化矩阵完全相同. 为了使 ACDC 的优化变得简单, 文献 [15] 提出了非对称的最小二乘代价函数, 得到了性能更好的三迭代算法 (Triple iterative algorithm, TIA) 算法, 然而 TIA 算法进行一次迭代需要交替更新三组待定参数, 因此其收敛速度仍然较慢. 而且正如文献 [14] 中所述, 由于该算法没有对左右对角化矩阵进行任何约束, 因此仍有可能收敛到奇异解.

为了提高非对称联合对角化算法的收敛速度, 并避免算法收敛到奇异解, 本文首先分析了 TIA 中的代价函数, 并提出可以避免奇异解的非对称联合对角化代价函数, 然后使用循环最小化技术^[16] 来优化提出的代价函数, 得到一种快速非对称非正交联合对角化算法: 非对称非正交联合对角化 (Nonsymmetric nonorthogonal joint diagonalization, NNJD). 算法的性能分析表明了该算法具有全局渐近收敛性和不变性, 实验仿真验证了算法的有效性.

1 联合对角化问题及代价函数

对于给定的矩阵集合 $\mathcal{R} = \{R_1, \dots, R_K\}$, $\forall k \in \{1, \dots, K\}$, $R_k \in \mathbf{C}^{N \times N}$, 联合对角化的目标就是要寻找非奇异的对角化矩阵 W , 使矩阵 $W^H R_k W$, $\forall k \in \{1, \dots, K\}$ 变成对角矩阵. 事实上, 联合对角化是矩阵特征值问题的推广. 当集合中的矩阵个数 $K = 1, 2$ 时, 联合对角化分别对应于单个矩阵的特征值分解和矩阵束的广义特征分解^[17], 因此对角化可以准确得到; 当 $K > 2$ 时, 联合对角化只能通过优化某一代价函数近似得到. 其中经典的代价函数为 Yeredor^[6] 提出的最小二乘对称代价函数

$$J(A, \Lambda_1, \dots, \Lambda_K) = \sum_{k=1}^K \|R_k - A\Lambda_k A^H\|_F^2 \quad (1)$$

其中, A 为混合矩阵, Λ_k 为对角矩阵, $\|\cdot\|_F$ 为 Frobenius 范数. 可以看出由式 (1) 定义的代价函数是混合矩阵的四次函数, 因此代价函数关于 A 的直接最小化不易求解. 为了简化计算, 文献 [15] 在文

献 [6] 的基础上提出了如下非对称代价函数

$$\tilde{J}(A, \Lambda_1, \dots, \Lambda_K, B) = \sum_{k=1}^K \|R_k - A\Lambda_k B^H\|_F^2 \quad (2)$$

由式 (2) 定义的非对称代价函数是矩阵 A 和 B 的二次函数, 因此优化问题变得简单. 而且文献 [15] 中指出, 当算法收敛时, A 和 B 是本质相等的, 即 $A = EB$, 其中 E 是一个广义交换矩阵 (每一行和每一列只有一个非零元素).

由代价函数 (1) 和 (2) 的定义可知, 文献 [6] 和 [15] 求解的是联合对角化的逆问题. 以代价函数 (2) 为例, 当算法完全收敛时有 $R_k = A\Lambda_k B^H$ 成立, 那么可取双边不同的对角化矩阵 $W = A^{-H}$, $V = B^{-H} = E^H A^{-H} = E^H W$, 使 $W^H R_k V = \Lambda_k$, $k = \{1, \dots, K\}$ 成为对角矩阵. 然而优化代价函数 (2) 得到的联合对角化算法有三个缺点:

1) 非对称代价函数 (2) 不仅是混合矩阵 A 和 B 的函数, 而且也是对角矩阵 Λ_k 的函数, 因此在每一步迭代中不仅要更新矩阵 A 和 B , 也要相应地更新所有的对角矩阵 Λ_k , $k = \{1, \dots, K\}$. 然而对角矩阵 Λ_k 在盲信源分离中是没有意义的, 因此优化代价函数 (2) 无疑增加了额外的运算.

2) 优化 (2) 需要更新三组待定参数 (A, B 以及对角矩阵 Λ_k), 尤其当 K 较大时, 需要更新的对角矩阵较多, 那么对角矩阵的更新将严重影响 A 和 B 的更新方向, 从而使算法整体收敛较慢 (从数值实验可以看出).

3) 代价函数 (2) 没有关于 A 和 B 的任何非奇异性约束, 因此根据文献 [14] 可知, 非对称代价函数 (2) 也有可能收敛到奇异解. 而奇异解将导致源信号不能完全分离.

针对上述问题, 本文提出如下代价函数:

$$J(W, V) = J_1(W, V) - J_2(W, V) \quad (3)$$

其中

$$J_1(W, V) = \sum_{k=1}^K \sum_{i \neq j}^N |[W^H R_k V]_{ij}|^2 \quad (4)$$

$$J_2(W, V) = \log |\det(WV)| \quad (5)$$

式中 $[\cdot]_{ij}$ 代表矩阵的第 i 行第 j 列元素, $\det(\cdot)$ 表示矩阵行列式. 与式 (2) 相比, 非对称代价函数 (3) 试图从正向解决联合对角化问题. 式 (3) 中的第一项 $J_1(W, V)$ 计算矩阵 $W^H R_k V$ 非对角元素的平方和, 因此, $J_1(W, V)$ 的大小衡量了矩阵接近对角矩阵的程度. 式 (3) 中的第二项 $-J_2(W, V) = \log(1/|\det(WV)|)$ 事实上是使 W 和 V 非奇异的约束项, 当 W 和 V 接近奇异时就会导致代价函数

(3) 趋于无穷大. 那么对代价函数 (3) 的最小化不仅可以完成矩阵 $W^H R_k V$ 的对角化, 而且保证了左右对角化矩阵的非奇异性. 因此, 最小化式 (3) 的物理意义可以解释为

$$\min_{W, V} J_1(W, V), \text{ s.t. } \det(W) \neq 0, \det(V) \neq 0 \quad (6)$$

与式 (2) 相比, 优化代价函数 (3) 有两个优点:

1) 代价函数 (3) 的优化与对角矩阵无关, 只需要更新两个待定参数, 因此不需要额外的计算. 另一方面. 由于不存在对角矩阵的优化, W 和 V 的更新将不受未知的对角矩阵的影响, 因此所得算法有望快速收敛;

2) 由于代价函数 (3) 增加了使 W 和 V 非奇异的约束项, 因此算法不会收敛到奇异解.

2 算法推导

代价函数 (3) 的最小化可以使用经典的梯度下降法, 然而梯度下降法的收敛速度一般较慢, 而且要在算法收敛速度和稳态误差之间做权衡来选择合适的迭代步长, 因此其步长选择也是一个难题. 为了使算法快速收敛, 我们采用一种循环最小化技术^[16], 其实现过程如下:

当 $m = 0$ 时, 选择一个初始值 $V(0) = [\mathbf{v}_1(0), \dots, \mathbf{v}_N(0)] \in \mathbf{C}^{N \times N}$, 对于迭代次数 $m = 1, 2, \dots$, 按照下面步骤交替更新 W 和 V 直到算法收敛.

$$W(m) = \arg \min_W J(W, V(m-1)) \quad (7)$$

$$V(m) = \arg \min_V J(W(m), V) \quad (8)$$

上面的算法也可以用 $W(0) = [\mathbf{w}_1(0), \dots, \mathbf{w}_N(0)] \in \mathbf{C}^{N \times N}$ 作为初始值进行迭代, 相应地对换式 (7) 和 (8) 中的 W 和 V 即可.

可以看出, 上面的循环最小化方法将代价函数 (3) 的优化问题转化为式 (7) 和 (8) 两个比较简单的子优化问题. 式 (7) 的优化可以通过逐列更新矩阵 W 来最小化代价函数 $J(W, V(m-1))$, 而式 (8) 的优化则通过逐列更新矩阵 V 来最小化代价函数 $J(W(m), V)$. 由于式 (7) 和 (8) 的最小化在实现步骤上是完全相同的, 因此下面仅给出式 (7) 的最小化过程.

为了方便表示, 令 $U = [\mathbf{u}_1, \dots, \mathbf{u}_N] = V(m-1)$, $W = [\mathbf{w}_1, \dots, \mathbf{w}_N]$, 那么式 (7) 可以表示为

$$W(m) = \arg \min_W J(W, U) \quad (9)$$

$\forall n \in \{1, \dots, N\}$, 现更新矩阵 W 的第 n 列 \mathbf{w}_n 来最小化式 (9).

首先将式 (4) 化简得到:

$$J_1(W, U) = \sum_{k=1}^K \sum_{i \neq j}^N |\mathbf{w}_i^H R_k \mathbf{u}_j|^2 = \sum_{k=1}^K \sum_{i \neq j}^N \mathbf{w}_i^H R_k \mathbf{u}_j \mathbf{u}_j^H R_k^H \mathbf{w}_i = \sum_{i=1}^N \mathbf{w}_i^H \left\{ \sum_{k=1}^K R_k \left[\sum_{j \neq i}^N \mathbf{u}_j \mathbf{u}_j^H \right] R_k^H \right\} \mathbf{w}_i \quad (10)$$

再将式 (5) 中矩阵行列式按照 W 的第 n 列展开得到

$$J_2(W, U) = \log |\det(W)| + \log |\det(U)| = \log |\det(\boldsymbol{\omega}_n^H \mathbf{w}_n)| + \log |\det(U)| \quad (11)$$

其中, $\boldsymbol{\omega}_n = [\omega_{1n}, \dots, \omega_{Nn}]^H$ 是由 \mathbf{w}_n 中的元素对应的代数余子式组成的列向量, ω_{in} 是矩阵 W 中元素 W_{in} 对应的代数余子式.

令 $Q_i = \sum_{k=1}^K R_k \left[\sum_{j \neq i}^N \mathbf{u}_j \mathbf{u}_j^H \right] R_k^H$, 并将式 (10) 和式 (11) 代入式 (9), 得到化简后的代价函数为

$$J(W, U) = \sum_{i=1}^N \mathbf{w}_i^H Q_i \mathbf{w}_i - \log |\det(\boldsymbol{\omega}_n^H \mathbf{w}_n)| - \log |\det(U)| \quad (12)$$

计算代价函数 (12) 关于 \mathbf{w}_n 的共轭梯度及 Hessian 矩阵分别为

$$\nabla_{\mathbf{w}_n^*} J(W, U) = Q_n \mathbf{w}_n - \frac{\boldsymbol{\omega}_n}{2\boldsymbol{\omega}_n^H \mathbf{w}_n} \quad (13)$$

$$\nabla_{\mathbf{w}_n^*} \nabla_{\mathbf{w}_n}^T J(W, U) = Q_n \quad (14)$$

由于 Q_n 总是正定的, 因此令式 (13) 等于零, 得到 \mathbf{w}_n 的最优解为

$$\mathbf{w}_n^{\text{opt}} = \frac{Q_n^{-1} \boldsymbol{\omega}_n}{\sqrt{2\boldsymbol{\omega}_n^H Q_n^{-1} \boldsymbol{\omega}_n}} \quad (15)$$

对于 $n = 1, \dots, N$, 逐列更新 W 完成一次更新, 从而实现式 (9) 的最小化. 注意到矩阵 Q_i 的计算中第二个求和号的直接计算复杂度为 $O(N^3)$, 而 W 的一次更新需要进行 N 次计算, 因此第二个求和号的直接计算会使 W 进行一次更新关于矩阵维度的复杂度高达 $O(N^4)$, 这也使整个算法的计算量增加. 事实上, Q_i 中的第二个求和号可以通过下式计算:

$$\sum_{j \neq i}^N \mathbf{u}_j \mathbf{u}_j^H = U U^H - \mathbf{u}_i \mathbf{u}_i^H \quad (16)$$

由循环最小化方法可知,在更新 W 的过程中,矩阵 U 是不变的,因此整个 W 的一次更新只需要计算一次矩阵乘积 UU^H ,而式 (16) 中等号右边第二项的计算复杂度仅为 $O(N^2)$,相对于矩阵乘积可以忽略,因此通过式 (16) 来计算 Q_i 可使 W 更新一次的计算复杂度回到 $O(KN^3)$. 现将整个最小化算法 (NNJD) 总结如下:

步骤 1. 随机初始化 $V(0) = [v_1(0), \dots, v_N(0)]$ (精确初始化可参考文献 [9]);

步骤 2. 根据如下子步骤完成式 (7) 的最小化,并对左对角化矩阵 W 进行一次更新;

- 1) 结合式 (16) 计算 Q_n ;
- 2) 计算与 w_n 相对应的代数余子式列向量 ω_n ;
- 3) 按照 (15) 计算最优解 w_n^{opt} 并用其替换 w_n ;
- 4) 重复 1)~3) 直到 W 的所有列向量完成更新.

步骤 3. 按照步骤 2 中类似的方法完成式 (8) 的最小化,并使右对角化矩阵 V 更新一次.

步骤 4. 重复步骤 2 和 3,直到算法收敛为止.

步骤 4 中算法的终止条件可如下确定:

$$\frac{|J^{m+1}(W, V) - J^m(W, V)|}{|J^m(W, V)|} \leq \varepsilon \quad (17)$$

其中,上标 m 表示迭代次数, $\varepsilon > 0$ 是一个很小的正数,实验中取 $\varepsilon = 10^{-3}$.

3 性能分析

本节将分析所提出的 NNJD 算法的一些基本性能,包括算法的全局渐近收敛性、算法相对于矩阵集合的不变性以及算法收敛后左右对角化矩阵的关系.

关于算法的收敛性,有如下定理:

定理 1. 若矩阵集合 \mathcal{R} 中的矩阵均具有对角化结构,即 $\forall k \in \{1, \dots, K\}, R_k = A\Lambda_k A^H$, 其中, A 和 Λ_k 是未知的混合矩阵和对角阵,那么对于定义的非对称代价函数 (3),提出的算法是全局渐近收敛的.

证明. 由于代价函数 (3) 的优化可以理解为式 (6) 的最小化,因此,我们从式 (6) 的角度来分析提出算法的收敛性能. 代价函数 $J_1(W, V)$ 关于其变量 W 和 V 是连续的,而且有 $J_1(W, V) = \sum_{k=1}^K \sum_{i \neq j}^N |W^H R_k V|_{ij}|^2 \geq 0$; 因此,代价函数是正定的. 当式 (3) 的优化理解为式 (6) 的最小化时,循环最小化式 (7) 和 (8) 可分别等效为

$$W(m) = \arg \min_W J_1(W, V(m-1)), \text{s.t. } \det(W) \neq 0 \quad (18)$$

$$V(m) = \arg \min_V J_1(W(m), V), \text{s.t. } \det(V) \neq 0 \quad (19)$$

其中, m 表示迭代次数. 由式 (18) 和 (19) 易知如下不等式成立:

$$\begin{aligned} J_1(W(m), V(m)) &= \min_V J_1(W(m), V) \leq \\ &J_1(W(m), V(m-1)) = \\ &\min_W J_1(W, V(m-1)) \leq \\ &J_1(W(m-1), V(m-1)) \end{aligned} \quad (20)$$

由式 (20) 可知: $J_1(W, V)$ 是迭代次数 m 的单调递减函数.

通过以上可知,代价函数 $J_1(W, V)$ 是 Lyapunov 函数. 由 Lyapunov 稳定性定理^[18] 知,左右对角化矩阵 W 和 V 全局渐近收敛到其稳定的平衡点. \square

本文提出的算法还具有联合对角化算法所期望的另外一个性能,即算法对于可对角化矩阵集合的不变性. 这一性质可描述为:若对角化矩阵 W 和 V 是联合对角化算法关于矩阵集合 \mathcal{R} 的解,那么对于任意一个非奇异矩阵 M ,有 $\tilde{W} = M^{-H}W, \tilde{V} = M^{-H}V$ 是该算法关于矩阵集合 $\tilde{\mathcal{R}} = \{\tilde{R}_1, \dots, \tilde{R}_K\}$ 的解,其中 $\tilde{R}_k = MR_k M^H, k = 1, \dots, K$. 本文提出的算法很明显满足上述不变性,由于对式 (4) 有

$$\begin{aligned} J_1(W, V) &= \sum_{k=1}^K \sum_{i \neq j}^N |W^H R_k V|_{ij}|^2 = \\ &\sum_{k=1}^K \sum_{i \neq j}^N |[(M^{-H}W)^H (MR_k M^H) (M^{-H}V)]_{ij}|^2 = \\ &\sum_{k=1}^K \sum_{i \neq j}^N |[\tilde{W}^H \tilde{R}_k \tilde{V}]_{ij}|^2 = J_1(\tilde{W}, \tilde{V}) \end{aligned} \quad (21)$$

而且当 W 和 V 非奇异时,由于 M 非奇异,因此, \tilde{W} 和 \tilde{V} 也是非奇异的,即 $\det(\tilde{W}) \neq 0, \det(\tilde{V}) \neq 0$. 由式 (6) 和 (21) 知:用矩阵 W 和 V 将矩阵集合 \mathcal{R} 对角化等同于使用矩阵 \tilde{W} 和 \tilde{V} 将矩阵集合 $\tilde{\mathcal{R}}$ 对角化,而且后者也满足对角化矩阵的非奇异性约束,因此当 W 和 V 是提出算法关于矩阵集合 \mathcal{R} 的解时, \tilde{W} 和 \tilde{V} 也是算法关于矩阵集合 $\tilde{\mathcal{R}}$ 的解. 根据上面分析可知本文提出的算法具有不变性,可以验证, TIA 算法并不具有此特性. 事实上,不变性是联合对角化算法所期望具有的一个重要特性. 例如,盲信源分离的某些算法^[3-5] 需要首先对观测信号进行预白化,然后再利用正交联合对角化算法辨识正交的混合矩阵来进行源分离. 如果联合对角化算法具有不变性,那么可以省去预白化的过程,将算法直接应用于待对角化的统计量即可辨识出分离矩阵,而文献 [3-5] 中辨识出的正交分离矩阵也可以由具有

不变性的非正交联合对角化算法 (例如本文算法) 辨识出的分离矩阵和白化矩阵计算得到。

对于本文算法中的左右对角化矩阵 W 和 V , 有如下定理:

定理 2. 若矩阵集合 \mathcal{R} 中的矩阵均具有可对角化结构, 即 $\forall k \in \{1, \dots, K\}, R_k = A\Lambda_k A^H$, 其中, A 和 Λ_k 是未知的混合矩阵和对角阵, 且对于 Λ_k 中的对角元素 $\lambda_k(i), i = 1, \dots, N$, 有

$$\exists k \in \{1, \dots, K\}, \forall 1 \leq i \neq j \leq N, \lambda_k(i) \neq \lambda_k(j) \quad (22)$$

那么当 NNJD 算法全局收敛时, 有 $W = VD$ 成立, 其中, D 是任意的非奇异对角矩阵。

证明. 根据文献 [5] 中的定理 3, 式 (22) 是联合对角化问题中很容易满足的条件, 因此结合矩阵谱理论^[17] 可知, 当 NNJD 算法收敛时, 其左右对角化矩阵与混合矩阵的逆共轭转置是本质相等的, 即

$$W = A^{-H}P_1D_1, \quad V = A^{-H}P_2D_2 \quad (23)$$

其中, P_1, P_2 是交换矩阵, D_1, D_2 是非奇异的对角阵. 由于 $J_1(W, V) = 0$ 是 (6) 的全局最小点, 因此当 NNJD 算法收敛时有 $\forall k \in \{1, \dots, K\}, W^H R_k V = \hat{D}_k$, 其中, \hat{D}_k 是任意的对角矩阵. 注意到 R_k 的对角化结构并将 (23) 带入上式得到

$$P_1^H \Lambda_k P_2 = D_1^{-H} \hat{D}_k D_2^{-1} \quad (24)$$

由于式 (24) 等号右边仍是对角矩阵, 若 $P_1 \neq P_2$, 那么对于对角矩阵 Λ_k , 式 (24) 左边必为非对角矩阵, 因此必有 $P_1 = P_2$ 成立, 将其代入式 (23) 并令 $D = D_2^{-1}D_1$, 有 $W = VD$. \square

由定理 2 知, 从对称联合对角化的角度出发, 左右对角化矩阵 W 和 V 都可以用做矩阵集合 \mathcal{R} 的对角化矩阵, 因为当算法收敛时 $\forall k \in \{1, \dots, K\}, \hat{D}_k = W^H R_k V = D^H V^H R_k V$ 是一个对角矩阵, 所以 $V^H R_k V = D^{-H} \hat{D}_k$ 是对角矩阵. 同理 $W^H R_k W$ 也是对角矩阵. 事实上, 定理 2 表明, 非对称联合对角化是联合对角化问题的更一般的形式, 下面的实验仿真很好地验证了左右对角化矩阵 W 和 V 对于联合对角化具有相同的性能。

4 数值实验

本节使用两组数据对 NNJD 算法进行仿真研究. 第一个实验采用一组可近似对角化的矩阵集合 \mathcal{R} , 其每一个矩阵可由下面的混合模型产生

$$R_k = A\Lambda_k A^H + \sigma^2 \xi_k \xi_k^H, \quad k = 1, \dots, K \quad (25)$$

第二个实验将算法应用于语音信号的盲分离, 其待对角化的矩阵集合来自于混合信号的二阶统计量。

由于本文算法是一种非对称的联合对角化方法, 因此, 其性能将与文献 [15] 中的非对称联合对角化方法 (TIA) 进行比较。

联合对角化算法的性能由于干扰信号比 (Interference to signal ratio, ISR) 来描述, ISR 定义为

$$\text{ISR} = \frac{1}{2N} \sum_{i=1}^N \left[\sum_{j=1}^N \frac{|c_{ij}|^2}{\max_k |c_{ik}|^2} - 1 \right] + \frac{1}{2N} \sum_{j=1}^N \left[\sum_{i=1}^N \frac{|c_{ij}|^2}{\max_k |c_{kj}|^2} - 1 \right] \quad (26)$$

其中, $C = \{c_{ij}\} = W^H A$, 在盲信源分离中, C 代表分离-混合全局系统矩阵. ISR 描述了矩阵 C 接近广义置换矩阵的程度, 当 C 越接近广义置换矩阵, ISR 越小, 算法性能越好. 由于 TIA 求解的是联合对角化的逆问题, 因此 $C = \hat{A}^{-1}A$, 其中, \hat{A} 表示对 A 的估计. 事实上, 本文提出的算法与文献 [15] 中的算法也可以分别使用其右对角化矩阵 V 和 B 来计算 ISR。

实验 1. 在本实验中, 矩阵集合 \mathcal{R} 中的元素由式 (25) 产生, 其中, 混合矩阵 A 和模拟噪声的干扰矩阵 ξ_k 中的元素, 以及矩阵 Λ_k 的对角线元素在每次独立实验中都是随机产生的, 且服从标准正态分布. σ 是一个用来控制干扰强度的标量. 在每次独立实验中, 两种算法的初始迭代矩阵都是随机产生的。

图 1 ~ 4 为两种算法独立运行 100 次的平均性能曲线. 选择矩阵维数 $N = 5$, 矩阵个数 $K = 10$, 图 1 给出了本文提出的 NNJD 算法与 TIA 算法在不同噪声强度下 ISR 随迭代次数增加的变化曲线. 从图 1 可以看出, 无论在何种噪声强度下, 本文提出的非对称算法的收敛速度都是最快的, 而且收敛后的干扰信号比也比 TIA 算法要好. 图 2 给出了两种算法的平均性能随噪声强度的变化曲线, 需要指出的是, 本文算法的迭代终止条件直接使用式 (17), 而 TIA 算法只需将式 (17) 中的代价函数用式 (2) 替换即可, 其中阈值 $\varepsilon = 10^{-3}$. 图 3 是在不同噪声强度下, ISR 的平均值随矩阵维数增加时的对比曲线. 可以看出, 当矩阵维数增加时, 两种算法的性能都将变差, 这主要是由于矩阵维数 N 的增加导致了式 (26) 的计算中求和项增加, 对算法而言相当于干扰项增加, 因此算法性能变差. 然而本文提出的算法较文献 [15] 中的算法仍然具有较好的抗干扰能力. 图 4 是分别使用左右对角化矩阵 (TIA 使用 A 和 B , NNJD 使用 W 和 V) 计算 ISR 所得到的性能曲线. 图 4 很好地验证了定理 2, 说明了两种算法得到的左右对角化矩阵对于联合对角化具有相同的性能。

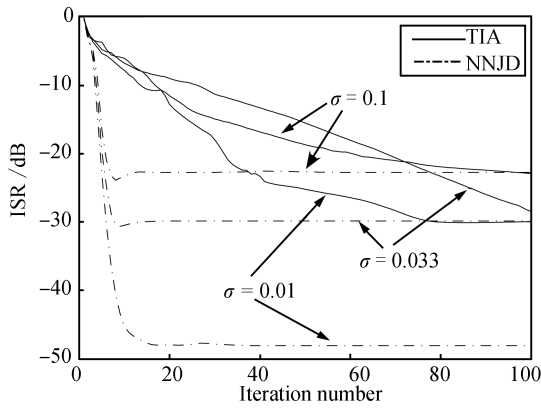


图 1 ISR 随迭代次数的变化曲线
Fig. 1 ISR vs. the number of iterations

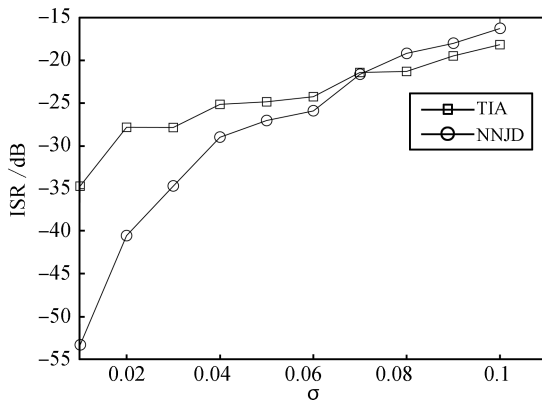


图 2 ISR 随噪声强度的变化曲线
Fig. 2 ISR vs. the noise level

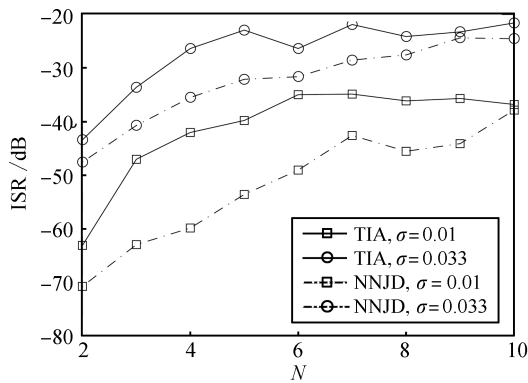


图 3 ISR 随矩阵维数的变化曲线
Fig. 3 ISR vs. the dimension of matrices

实验 2. 将两种非对称联合对角化算法分别应用于音频信号的盲分离, 源信号波形如图 5 所示, 其中, 第一个和第三个为语音信号, 第二个为乐器音, 3 个源信号的采样频率均为 8 kHz, 采样点数为 $T = 2 \times 10^4$. 矩阵集合 \mathcal{R} 仍由式 (25) 产生, 其中, 矩阵 A 和 ξ_k 的产生方式与实验 1 相同, 而对角矩

阵 Λ_k 取源信号的 k 时延相关矩阵 (与文献 [13] 中的仿真方法相同), 即

$$[\Lambda_k]_{ij} = \frac{1}{T} \sum_t s_i(t)s_j(t+k) \quad (27)$$

令 $k = 0, 5, \dots, 45$ 得到 $K = 10$ 个可对角化的矩阵. 需要指出的是, 与实验 1 不同的是, 实验 2 中的矩阵并非严格对角矩阵, 而是近似的对角矩阵.

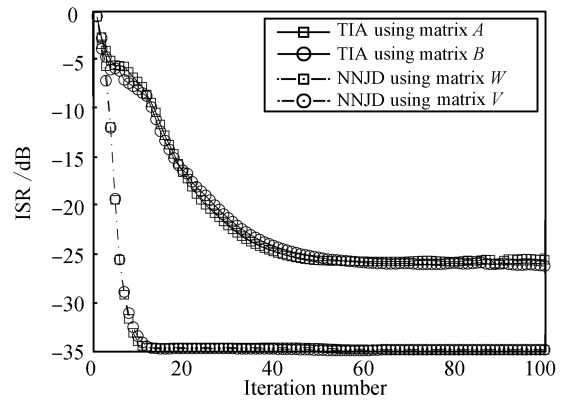


图 4 两种算法分别使用左右对角化矩阵计算 ISR
Fig. 4 The ISR of two algorithms using left and right diagonalizing matrices respectively

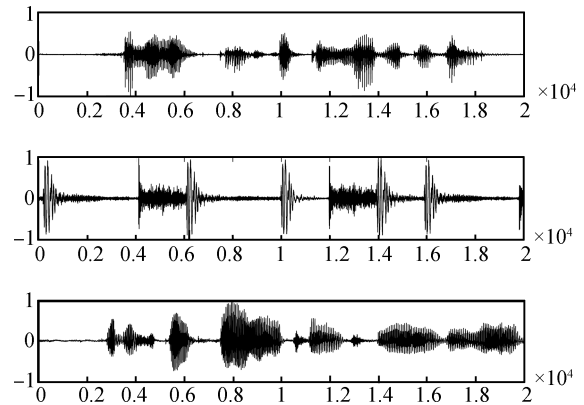


图 5 源信号波形
Fig. 5 Waveforms of source signals

取 $\sigma = 0.01$, 假设混合信号的个数与源的个数相同, 图 6 ~ 8 分别给出了三个源信号某一次的混合以及两种非对称算法的分离结果 (为了方便比较, 对信号幅度进行了归一化). 从图 7 可以看出, TIA 分离结果中的第一个信号与第三个信号都比较接近图 5 中的第一个源信号, 事实上并没有分离出图 5 中的第三个源信号, 因此可以判断, TIA 给出的解接近于奇异, 是病态解. 而从图 8 看出, 本文提出的 NNJD 算法则完全分离出了三个源信号的波形. 这主要是由于 NNJD 算法有对角化矩阵的非奇异性约

束, 而 TIA 则没有任何约束, 因此 TIA 很有可能收敛到病态解.

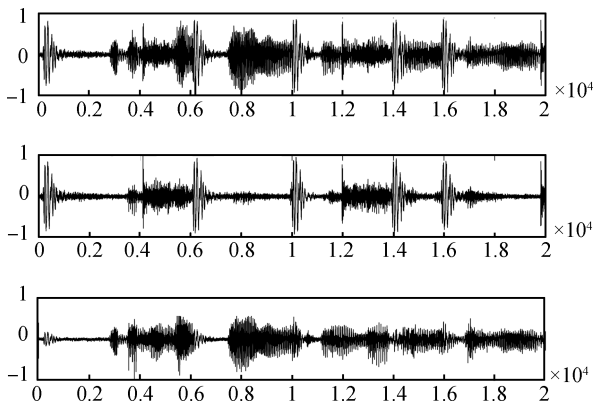


图6 观测到的信号波形

Fig. 6 Waveforms of observed signals

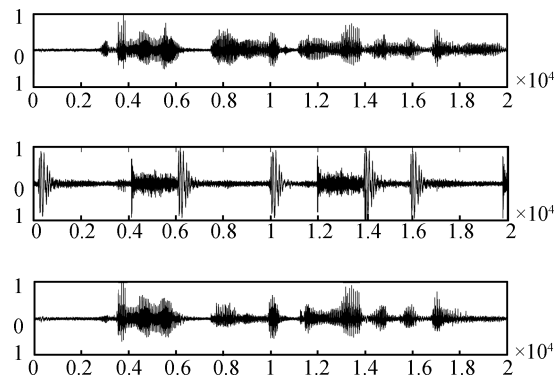


图7 TIA 分离结果

Fig. 7 Separation results resort to TIA

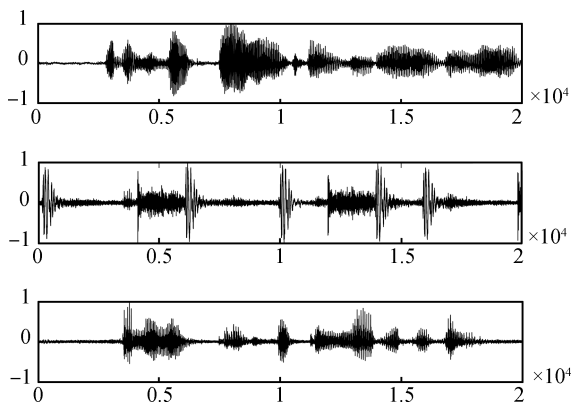


图8 NNJD 分离结果

Fig. 8 Separation results resort to NNJD

表1 比较了两种算法在音频信号盲分离中的平均性能, 其中, 源信号的估计误差 e 采用文献 [15] 中的方法计算

$$e = \frac{1}{NK} \left\| \hat{S} - (\hat{S}S^H)(SS^H)^{-1}S \right\|_F^2 \quad (28)$$

S 表示三个源信号的采样点组成的矩阵, \hat{S} 则是对 S 的估计. 由表1 可以看出, 本文算法的各项指标的平均性能较好, 可以使源信号的估计误差减小 12 dB, 干扰信号比减小了近 8 dB, 而且本文算法收敛所需的迭代次数也明显少于 TIA 算法.

表1 算法性能比较

Table 1 Performance comparison between two algorithms

算法	e (dB)	ISR (dB)	迭代次数
TIA	-26.79	-14.81	63.28
NNJD	-38.99	-22.49	23.42

5 结论

多个目标矩阵的同时对角化是许多信号处理领域面临的共同课题, 因此, 研究更加稳健、高效的联合对角化算法有着重要的应用价值. 本文从正向解决联合对角化问题, 提出了联合对角化的非对称代价函数, 代价函数的第二项实际上是对角化矩阵的非奇异约束项, 因此, 利用循环最小化技术优化提出的代价函数得到了非奇异的对角化矩阵. 而且, 本文对提出的算法进行了性能分析, 并证明了提出算法的全局渐近收敛性以及算法对于待对角化的矩阵集合的不变性, 左右对角化矩阵的关系也说明了非对称联合对角化的一般性. 同时, 本文对提出的算法进行了实验仿真, 与另外一种非对称联合对角化算法进行了比较, 比较结果说明了本文算法能够更快速地收敛. 将算法应用于盲信源分离不仅可以完全分离源信号, 而且改善了分离效果.

与参考算法相比, 本文提出的算法从正向解决联合对角化问题, 由于对角化矩阵的更新不受未知对角矩阵的影响, 因此通过实验仿真可以看出, 本文算法在收敛速度上有很大的优势. 然而对联合对角化算法进行收敛速度的理论分析将使我们更好地理解算法的收敛过程, 这也将是我们下一步的研究内容.

References

- 1 Flury B. *Common Principal Components and Related Multivariate Models*. New York: Wiley, 1988
- 2 Flury B, Gautschi W. An algorithm for simultaneous orthogonal transformation of several positive definite symmetric matrices to nearly diagonal form. *SIAM Journal on Scientific and Statistical Computing*, 1986, 7(1): 169-184
- 3 Cardoso J F, Souloumiac A. Blind beamforming for non-Gaussian signals. *IEE Proceedings, Part F: Radar and Signal Processing*, 1993, 140(6): 362-370
- 4 Zhang W T, Liu N, Lou S T. Joint approximate diagonalization using bilateral rank-reducing Householder transform

- with application in blind source separation. *Chinese Journal of Electronics*, 2009, **18**(3): 471–476
- 5 Belouchrani A, Abed-Meraim K, Cardoso J F, Moulines E. A blind source separation technique using second-order statistics. *IEEE Transactions on Signal Processing*, 1997, **45**(2): 434–444
- 6 Yeredor A. Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation. *IEEE Transactions on Signal Processing*, 2002, **50**(7): 1545–1553
- 7 Joho M, Mathis H. Joint diagonalization of correlation matrices by using gradient methods with application to blind signal separation. In: Proceedings of IEEE Sensor Array and Multichannel Signal Processing Workshop. Rosslyn, USA: IEEE, 2002. 273–277
- 8 Degerine S, Kane E. A comparative study of approximate joint diagonalization algorithms for blind source separation in presence of additive noise. *IEEE Transactions on Signal Processing*, 2007, **55**(6): 3022–3031
- 9 Yeredor A. On using exact joint diagonalization for noniterative approximate joint diagonalization. *IEEE Signal Processing Letters*, 2005, **12**(9): 645–648
- 10 Pham D T. Joint approximate diagonalization of positive definite Hermitian matrices. *SIAM Journal on Matrix Analysis and Applications*, 2000, **22**(4): 1136–1152
- 11 Todros K, Tabrikian J. Fast approximate joint diagonalization of positive definite Hermitian matrices. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing. Hawaii, USA: IEEE, 2007. 1373–1376
- 12 van der Veen A J. Joint diagonalization via subspace fitting techniques. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing. Salt Lake City, USA: IEEE, 2001. 2773–2776
- 13 Vollgraf R, Obermayer K. Quadratic optimization for simultaneous matrix diagonalization. *IEEE Transactions on Signal Processing*, 2006, **54**(9): 3270–3278
- 14 Li X L, Zhang X D. Nonorthogonal joint diagonalization free of degenerate solution. *IEEE Transactions on Signal Processing*, 2007, **55**(5): 1803–1814
- 15 Zhang Hua, Feng Da-Zheng, Nie Wei-Ke, Xu Xian-Feng. Non-orthogonal joint diagonalization for blind source separation. *Journal of Xidian University (Natural Science)*, 2008, **35**(1): 27–31
(张华, 冯大政, 聂卫科, 徐先峰. 非正交联合对角化盲源分离算法. 西安电子科技大学学报 (自然科学版), 2008, **35**(1): 27–31)
- 16 Stoica P, Selen Y. Cyclic minimizers, majorization techniques, and the expectation-maximization algorithm: a refresher. *IEEE Signal Processing Magazine*, 2004, **21**(1): 112–114
- 17 Horn R A, Johnson C R. *Matrix Analysis*. Cambridge: Cambridge University Press, 1985
- 18 Vidyasagar M. *Nonlinear Systems Analysis (Second Edition)*. New Jersey: Prentice Hall, 1993



张伟涛 西安电子科技大学博士研究生。主要研究方向为盲信号分离、盲均衡。本文通信作者。

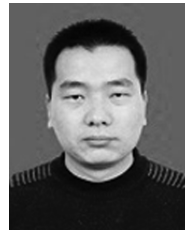
E-mail: zhwt-work@foxmail.com

(ZHANG Wei-Tao Ph.D. candidate at the School of Electronic Engineering, Xidian University. His research interest covers blind source separation and blind equalization. Corresponding author of this paper.)



楼顺天 西安电子科技大学教授。主要研究方向为模式识别、神经网络、模糊系统。E-mail: shtlou@mail.xidian.edu.cn

(LOU Shun-Tian Professor at the School of Electronic Engineering, Xidian University. His research interest covers pattern recognition, neural networks, and fuzzy systems.)



张延良 西安电子科技大学博士研究生, 河南理工大学讲师。主要研究方向为盲信源分离及其应用。

E-mail: ylz119@qq.com

(ZHANG Yan-Liang Ph.D. candidate at the School of Electronic Engineering, Xidian University, and he is also a lecturer in the Department of

Computer Science and Technology, Henan Polytechnic University. His research interest covers blind source separation and its application.)