A Geometric Framework for Non-Unitary Joint Diagonalization of Complex Symmetric Matrices

Martin Kleinsteuber and Hao Shen

Department of Electrical Engineering and Information Technology
Technische Universität München, Germany
{kleinsteuber,hao.shen}@tum.de

Abstract. Non-unitary joint diagonalization of complex symmetric matrices is an important technique in signal processing. The so-called complex oblique projective (COP) manifold has been shown to be an appropriate manifold setting for analyzing the problem and developing geometric algorithms for minimizing the off-norm cost function. However, the recent identification of the COP manifold as a collection of rank-one orthogonal projector matrices is not a suitable framework for the reconstruction error function due to its large memory requirement compared to the actual dimension of the search space. In this work, we investigate the geometry of the COP manifold as a quotient manifold, which allows less memory requirement, and develop a conjugate gradient algorithm to minimize the reconstruction error function.

Keywords: Joint diagonalization of complex symmetric matrices, complex projective space, conjugate gradient algorithm.

1 Introduction

Joint diagonalization of a set of matrices plays an important role in various signal processing problems, such as blind source separation [1], beamforming [2], and direction of arrival estimation [3]. Early works on matrix joint diagonalization are restricted to unitary transformations, cf. [4]. However, it has been shown that unitary joint diagonalization (UJD) approaches may have a serious limit of degraded performance in the presence of additive noise, cf. [5]. To avoid such a limitation, non-unitary joint diagonalization (NUJD) has been proposed and actively studied, cf. [6,7,8]. In this work, we are interested in jointly diagonalizing a set of complex symmetric matrices, cf. [9,10].

In the literature, measurement of diagonality of matrices can be formulated in three different forms, namely, off-norm formulation [11], log-likelihood formulation [6], and reconstruction error formulation [12,13]. It is important to notice, that the log-likelihood based criterion only applies to the case with positive definite matrices, and the off-norm cost function is not *column-wise scale invariant* with respect to the demixing matrix in general, except if the matrix is an exact joint diagonalizer. In the work of [14], the complex oblique projective (COP) manifold is identified as a set of rank-one orthogonal projectors. Unfortunately,

F. Nielsen and F. Barbaresco (Eds.): GSI 2013, LNCS 8085, pp. 353–360, 2013.

such a representation requires a lot of memory compared to the actual dimension of the search space. In this work we study the representation of the COP manifold as a quotient manifold. Based on the derived geometry, we develop a conjugate gradient algorithm for solving the joint diagonalization problem.

This paper is organized as follows. In Section 2, we introduce the problem of joint diagonalization of complex symmetric matrices, and review second order statistics based approaches. Section 3 studies the geometry of the COP manifold. In Section 4, we develop a conjugate gradient algorithm for minimizing the reconstruction error cost function. Finally, Section 5 demonstrates the numerical performance of the developed CG algorithm.

2 Problem Description and Preliminaries

Let us start with some notations and definitions. In this work, $(\cdot)^\mathsf{T}$ denotes the matrix transpose, $(\cdot)^\mathsf{H}$ the Hermitian transpose, $\overline{(\cdot)}$ the complex conjugate, and $\Re(z)$ the real part of $z \in \mathbb{C}$. By $Gl(m)$ we denote the set of all invertible $(m \times m)$ complex matrices, by $\|\cdot\|_F$ the Frobenius norm of matrices, and by I_m the $(m \times m)$-identity matrix.

Let $\{C_i\}_{i=1}^n$ be a set of $m \times m$ complex symmetric matrices, constructed by

$$C_i = A\Omega_i A^\mathsf{T}, \qquad i = 1, \ldots n, \tag{1}$$

where $A \in Gl(m)$ is the mixing matrix and $\Omega_i = \operatorname{diag}\left(\omega_{i1}, \ldots, \omega_{im}\right) \in \mathbb{C}^{m \times m}$ with $\Omega_i \neq 0$ for all $i = 1, \ldots, n$. Both A and the set of $\{\Omega_i\}_{i=1}^n$ are unknown. The task is to find a matrix $X \in Gl(m)$ such that the matrices

$$X^\mathsf{T} C_i X, \quad i = 1, \ldots, n, \tag{2}$$

are simultaneously diagonalized. Clearly, the mixing matrix can only be identified up to permutation and scaling. In this work, we only take care of the scaling ambiguity, and define the set of all diagonal $(m \times m)$-matrices by

$$\mathcal{D}(m) := \{D \,|\, D \in Gl(m) \text{ is diagonal}\}. \tag{3}$$

Since $\mathcal{D}(m)$ admits a matrix group structure, we can define the following equivalence class on $\mathbb{C}^{m \times m}$, cf. [15].

Definition 1 (Equivalence Relation). *Let $X, Y \in Gl(m)$, then X is said to be* equivalent *to Y if there exists $D \in \mathcal{D}(m)$ such that $X = YD$.*

Accordingly, for a given $X \in Gl(m)$, we define the equivalent class of X as

$$\lfloor X \rfloor := \{XD \in Gl(m) \,|\, D \in \mathcal{D}\} \tag{4}$$

and the quotient space as

$$Op(m) := \{\lfloor X \rfloor \,|\, X \in Gl(m)\}, \tag{5}$$

which we call the *complex oblique projective (COP)* manifold. Note, that the COP manifold is a generalization of the so-called oblique manifold for solving real valued matrix joint diagonalization problems, cf. [16]. By adapting the reconstruction error based cost function proposed in [13] to the current setting of complex symmetric matrices, the cost function studied in this work is given by

$$f\colon Gl(m) \to \mathbb{R}, \qquad X \mapsto \sum_{i=1}^{n} \frac{1}{4} \left\| C_i - X^{-\mathsf{T}} \operatorname{ddiag}(X^{\mathsf{T}} C_i X) X^{-1} \right\|_F^2, \tag{6}$$

where $\operatorname{ddiag}(M)$ is the diagonal matrix whose diagonal entries are just those of M. It can be shown that the function is column-wise complex scaling invariant, i.e. $f(X) = f(XD)$ with a diagonal matrix $D \in \mathcal{D}$ and therefore induces a function on $Op(m)$ as

$$\widehat{f}\colon Op(m) \to \mathbb{R}, \qquad \lfloor X \rfloor \mapsto \sum_{i=1}^{n} \frac{1}{4} \left\| C_i - X^{-\mathsf{T}} \operatorname{ddiag}(X^{\mathsf{T}} C_i X) X^{-1} \right\|_F^2. \tag{7}$$

Since the COP manifold $Op(m)$ has no representation in terms of $(m \times m)$ complex matrices, we need to further explore its geometry to develop geometric gradient based numerical algorithms, which take into account the dimension of the underlying feasible set.

3 The Geometry of the COP Manifold

In this section, we derive the necessary geometric concepts of the COP manifold in order to implement a geometric conjugate gradient method. We refer to [17] for a detailed overview of optimization methods on matrix manifolds. Firstly, notice that $Op(m)$ is an open and dense Riemannian submanifold of the m-times product of $\mathbb{CP}^{m-1}$ with the Euclidean product metric, i.e.

$$\overline{Op(m)} = \underbrace{\mathbb{CP}^{m-1} \times \ldots \times \mathbb{CP}^{m-1}}_{m-\text{times}} =: \left(\mathbb{CP}^{m-1}\right)^m, \tag{8}$$

where $\overline{Op(m)}$ denotes the closure of $Op(m)$. Thus, the tangent spaces, the geodesics, and the parallel transport for $Op(m)$ locally coincide with those of $(\mathbb{CP}^{m-1})^m$. In what follows, we study the geometry of $\mathbb{CP}^{m-1}$ by considering it as the quotient space $\mathbb{CP}^{m-1} = \mathbf{S}^m / \mathbf{S}^1$, where

$$\mathbf{S}^m := \{x \in \mathbb{C}^m \mid x^{\mathsf{H}} x = 1\} \tag{9}$$

denotes the complex unit sphere and the equivalence classes $\lfloor x \rfloor \in \mathbb{CP}^{m-1}$ are defined through the relation

$$x \sim y \quad \Longleftrightarrow \quad \exists z \in \mathbf{S}^1 \text{ such that } x = yz. \tag{10}$$

The tangent space at $x \in \mathbf{S}^m$ is given by

$$T_x \mathbf{S}^m := \{h \in \mathbb{C}^m \mid \Re(h^{\mathsf{H}} x) = 0\}. \tag{11}$$

Endowing the tangent space with the metric

$$\langle g, h \rangle := \Re(g^{\mathsf{H}} h), \quad \text{for } g, h \in T_x \mathbf{S}^m, \tag{12}$$

turns $\mathbf{S}^m$ into a Riemannian manifold. Although the complex projective space $\mathbb{CP}^{m-1}$ cannot be represented in terms of vectors in $\mathbb{C}^m$, its tangent spaces do have a vector representation. This allows to derive geometric gradient descent and conjugate gradient descent methods for optimizing over $\mathbb{CP}^{m-1}$ where each iterate is a representative of the respective equivalence class and thus can be implemented in terms of vectors in $\mathbf{S}^m$. We identify the tangent space at $\lfloor x \rfloor \in \mathbb{CP}^{m-1}$ with the horizontal lift of the tangent space at x. In our case, this is just the intersection of all tangent spaces in the respective equivalence class, i.e.

$$\begin{aligned} T_{\lfloor x \rfloor} \mathbb{CP}^{m-1} &= \bigcap_{z \in \mathbf{S}^1} T_{xz} \mathbf{S}^m \\ &= \bigcap_{z \in \mathbf{S}^1} \{ h \in \mathbb{C}^m \mid \Re(h^{\mathsf{H}} xz) = 0 \} \\ &= \{ h \in \mathbb{C}^m \mid h^{\mathsf{H}} x = 0 \}. \end{aligned} \tag{13}$$

The orthogonal projection onto the tangent space is given in the following lemma.

Lemma 1. *The orthogonal projection of a vector $h \in \mathbb{C}^m$ onto the tangent space $T_{\lfloor x \rfloor} \mathbb{CP}^{m-1}$ with respect to the inner product $\langle x, y \rangle = \Re(x^{\mathsf{H}} y)$ is given by*

$$\pi_{\lfloor x \rfloor}(h) := (I_m - x x^{\mathsf{H}}) h. \tag{14}$$

Proof. It is easy to see that for $h \in \mathbb{C}^m$, we have

$$x^{\mathsf{H}} \pi_{\lfloor x \rfloor}(h) = x^{\mathsf{H}} (I_m - x x^{\mathsf{H}}) h = 0, \tag{15}$$

i.e. $\pi_{\lfloor x \rfloor}(h) \in T_{\lfloor x \rfloor} \mathbb{CP}^{m-1}$. To see that $\pi_{\lfloor x \rfloor}$ is orthogonal, we observe that

$$\begin{aligned} \langle h - \pi_{\lfloor x \rfloor}(h), \pi_{\lfloor x \rfloor}(h) \rangle &= \langle h, \pi_{\lfloor x \rfloor}(h) \rangle - \langle \pi_{\lfloor x \rfloor}(h), \pi_{\lfloor x \rfloor}(h) \rangle \\ &= \Re\big(h^{\mathsf{H}} (I_m - x x^{\mathsf{H}}) h\big) - \big((I_m - x x^{\mathsf{H}}) h\big)^{\mathsf{H}} \big(I_m - x x^{\mathsf{H}}\big) h \\ &= 0. \end{aligned} \tag{16}$$

Theorem 1. *The geodesics in $\mathbb{CP}^{m-1}$ through $\lfloor x \rfloor$ are given by $\lfloor \gamma(t) \rfloor$ with*

$$\gamma(t) := \mathbf{e}^{t(h x^{\mathsf{H}} - x h^{\mathsf{H}})} x. \tag{17}$$

Proof. Clearly, $\lfloor \gamma(t) \rfloor \subset \mathbb{CP}^{m-1}$ for all $t \in \mathbb{R}$. Taking the first derivative of γ yields

$$\dot{\gamma}(t) = \mathbf{e}^{t(h x^{\mathsf{H}} - x h^{\mathsf{H}})} (h x^{\mathsf{H}} - x h^{\mathsf{H}}) x \tag{18}$$

with $\dot{\gamma}(0) = h$. It remains to show that $\ddot{\gamma}(t)$ is orthogonal to the tangent space at $\lfloor \gamma(t) \rfloor$. The second derivative of γ is given by

$$\ddot{\gamma}(t) = \mathbf{e}^{t(h x^{\mathsf{H}} - x h^{\mathsf{H}})} (h x^{\mathsf{H}} - x h^{\mathsf{H}})^2 x. \tag{19}$$

Orthogonality to the tangent space holds because the orthogonal projection of $\ddot{\gamma}(t)$ vanishes, i.e.

$$\begin{aligned}\pi_{\lfloor\gamma(t)\rfloor}(\ddot{\gamma}(t)) &= \left(I_m - \gamma(t)(\gamma(t))^{\mathsf{H}}\right)\ddot{\gamma}(t)\\ &= \left(I_m - \mathbf{e}^{t(hx^{\mathsf{H}}-xh^{\mathsf{H}})}xx^{\mathsf{H}}\mathbf{e}^{-t(hx^{\mathsf{H}}-xh^{\mathsf{H}})}\right)\mathbf{e}^{t(hx^{\mathsf{H}}-xh^{\mathsf{H}})}(hx^{\mathsf{H}}-xh^{\mathsf{H}})^2x \\ &= 0.\end{aligned} \quad (20)$$

Theorem 2. *The parallel transport from $T_{\lfloor\gamma(0)\rfloor}\mathbb{CP}^{m-1}$ to $T_{\lfloor\gamma(t)\rfloor}\mathbb{CP}^{m-1}$ along the geodesic $\lfloor\gamma(t)\rfloor$ is given by*

$$\tau(t) := \mathbf{e}^{t(hx^{\mathsf{H}}-xh^{\mathsf{H}})}h. \quad (21)$$

Proof. We have to show that $\pi_{\lfloor\gamma(t)\rfloor}(\dot{\tau}(t)) = 0$. This holds true since, using $x^{\mathsf{H}}h = 0$, we have

$$\dot{\tau}(t) = -\|h\|^2\gamma(t), \quad (22)$$

and thus its projection to the tangent space vanishes.

Finally, by exploring the structure of the product manifold, the projection onto the tangent space of $Op(m)$, the geodesics and the parallel transport are verified straightforwardly and given without proof.

Lemma 2. *The orthogonal projection of a matrix $H \in \mathbb{C}^{m\times m}$ to the tangent space $T_{\lfloor X\rfloor}Op(m)$ with respect to the inner product $\langle X,Y\rangle = \operatorname{tr}(\Re(X^{\mathsf{H}}Y))$ is given by*

$$\Pi_{\lfloor X\rfloor}(H) = H - X\operatorname{ddiag}\left(\Re\left(X^{\mathsf{H}}H\right)\right). \quad (23)$$

Theorem 3. *The geodesics in the COP manifold through $\lfloor X\rfloor \in Op(m)$ are given by $\lfloor\Gamma(t)\rfloor$ with*

$$\Gamma(t) := \left[\mathbf{e}^{t(h_1x_1^{\mathsf{H}}-x_1h_1^{\mathsf{H}})}x_1,\ldots,\mathbf{e}^{t(h_mx_m^{\mathsf{H}}-x_mh_m^{\mathsf{H}})}x_m\right]. \quad (24)$$

Its associated parallel transport is given by

$$T(t) := \left[\mathbf{e}^{t(h_1x_1^{\mathsf{H}}-x_1h_1^{\mathsf{H}})}h_1,\ldots,\mathbf{e}^{t(h_mx_m^{\mathsf{H}}-x_mh_m^{\mathsf{H}})}h_m\right]. \quad (25)$$

4 A CG Algorithm for Simultaneous Non-unitary Diagonalization of Complex Symmetric Matrices

In this section, we derive a conjugate gradient algorithm for minimizing the cost function $\widehat{f}$ as defined in (7). For the sake of simplicity, we denote $E_i(X) := X^{-\mathsf{T}}\operatorname{ddiag}(X^{\mathsf{T}}C_iX)X^{-1}$. Firstly, we compute the derivative of $\widehat{f}$ as

$$\begin{aligned}\mathrm{D}\widehat{f}(\lfloor X\rfloor)H = \sum_{i=1}^{n}\tfrac{1}{2}\operatorname{tr}H^{\mathsf{H}}\overline{E_i(X)}(C_i - E_i(X))X^{-\mathsf{H}} + X^{-1}\overline{(C_i - E_i(X))}E_i(X)H \\ - \operatorname{ddiag}\left(H^{\mathsf{H}}\overline{C_iX}\right)\overline{X}^{-1}(C_i - E_i(X))X^{-\mathsf{H}} \\ - \operatorname{ddiag}\left(X^{\mathsf{T}}C_iH\right)X^{-1}\overline{(C_i - E_i(X))}X^{-\mathsf{T}}.\end{aligned} \quad (26)$$

It is easily seen that an exact joint diagonalizer $\lfloor X^* \rfloor$ is a critical point of the function $\widehat{f}$, i.e. $\mathrm{D}\,\widehat{f}(\lfloor X^* \rfloor)H = 0$ for all $H \in T_{\lfloor X^* \rfloor}Op(m)$. Then, by Lemma 2, the Riemannian gradient of $\widehat{f}$ is computed as

$$\operatorname{grad} \widehat{f}_{\lfloor X \rfloor} = \sum_{i=1}^{n} \Pi_{\lfloor X \rfloor}\Big(\overline{E_i(X)}(C_i - E_i(X))X^{-\mathsf{H}} - Z_i(X)\Big), \tag{27}$$

where $Z_i(X) := [z_{i1}(X), \ldots, z_{im}(X)]$ with

$$z_{ij}(X) := \overline{C_i x_j}(\overline{X}^{-1}(C_i - E_i(X))X^{-\mathsf{H}})_{jj}, \tag{28}$$

where $(\cdot)_{jj}$ denotes the j-th diagonal entry of a matrix.

We now construct a conjugate gradient algorithm for minimizing the function $\widehat{f}$ as defined in (7), and discuss a few details of it. By following the derivation above, we summarize a CG algorithm in Algorithm 1. Step 4 requires to find the local or global minimum of a restricted cost function, which is often unfeasible in practice. In this work, we employ a one-dimensional Newton step instead, i.e.

$$\lambda^* = -\frac{\frac{\mathrm{d}}{\mathrm{d}\,t}\widehat{f}\circ\lfloor \Gamma(t)\rfloor\Big|_{t=0}}{\frac{\mathrm{d}^2}{\mathrm{d}\,t^2}\widehat{f}\circ\lfloor \Gamma(t)\rfloor\Big|_{t=0}}, \tag{29}$$

where the numerator and the denominator can be obtained by a tedious but straightforward computation. Finally, for updating the direction parameter γ in Step 5, we confine ourselves to a formula, which was proposed in [18] and adapted to the manifold setting in [19], as

$$\gamma = \frac{\left\langle \mathcal{G}^{(j+1)}, \mathcal{G}^{(j+1)} - \mathcal{T}\mathcal{G}^{(j)} \right\rangle}{\left\langle \mathcal{H}^{(j)}, \mathcal{G}^{(j)} \right\rangle}. \tag{30}$$

5 Numerical Experiments

In our experiment, we investigate the performance of our method compared with the AC/DC algorithm in [20], and the CG algorithm on minimizing the off-norm algorithm [14]. The task of our experiment is to jointly diagonalize a set of complex symmetric matrices $\{C_i\}_{i=1}^n$ constructed by

$$C_i = A\Omega_i A^{\mathsf{T}} + \varepsilon E, \tag{31}$$

where $A \in Gl(m)$ is randomly picked, the diagonal entries of Ω_i are drawn from a uniform distribution on the interval $(0, 10)$, the matrix $E \in \mathbb{C}^{m\times m}$ is a complex symmetric matrix, whose real and imaginary parts are generated from a uniform distribution on the unit interval $(-0.5, 0.5)$, representing additive stationary noise, and $\varepsilon \in \mathbb{R}$ is the noise level.

We set $m = 5$, $n = 5$, $\varepsilon = 0.01$, and run 100 tests. The quartile based boxplot of Amari errors for each method are drawn in Figure 1. Our proposed CG approach outperforms consistently the other two methods in terms of both average performance and convergence stability.

Algorithm 1. A CG Joint Diagonalization Algorithm

Input : A set of matrices $\{C_i\} \subset \mathbb{C}^{m\times m}$ for $i = 1, \ldots, n$;

Output: A matrix representation of the joint diagonalizer $X \in Gl(m)$;

Step 1: Generate an initial guess $\lfloor X^{(0)} \rfloor = [\lfloor x_1 \rfloor \ldots, \lfloor x_m \rfloor] \in Op(m)$ and set $j = 1$;

Step 2: Compute $\mathcal{G}^{(1)} = \mathcal{H}^{(1)} \leftarrow -\operatorname{grad} \widehat{f}_{\lfloor X^{(0)} \rfloor}$ using Eq. (27) ;

Step 3: Set $j = j + 1$;

Step 4: Update $X^{j+1} \leftarrow \Gamma(\lambda_j)$, where λ_j is computed as in (29) ;

Step 5: Update $\mathcal{H}^{(j+1)} \leftarrow -\mathcal{G}^{(j+1)} + \gamma_j T(\lambda_j)$, where $\mathcal{G}^{(j+1)} = \operatorname{grad} \widehat{f}_{\lfloor X^{(j)} \rfloor}$, and γ_j is chosen according to Eq. (30) ;

Step 6: If $j \bmod 2m(m-1) - 1 = 0$, set $\mathcal{H}^{(j+1)} \leftarrow -\mathcal{G}^{(j+1)}$;

Step 7: If $\left\|\mathcal{G}^{(j+1)}\right\|$ is small enough, stop. Otherwise, go to Step 3;

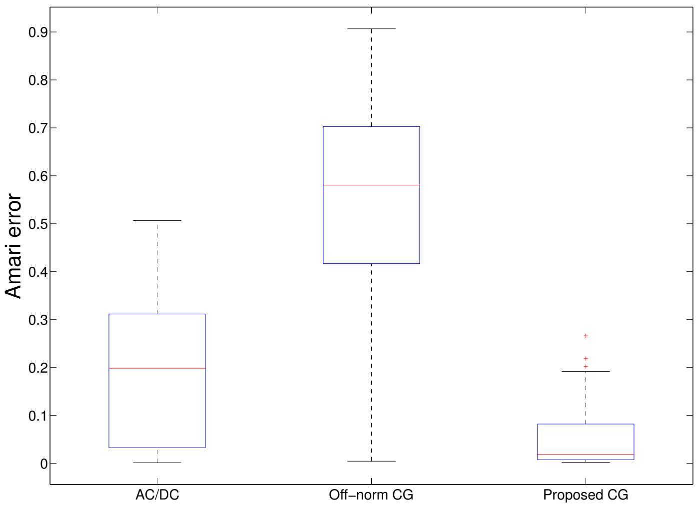

Fig. 1. Separation performance of the proposed algorithm

Acknowledgments. This work has been supported by the German DFG funded Cluster of Excellence CoTeSys - Cognition for Technical Systems.

References

1. Comon, P., Jutten, C. (eds.): Handbook of Blind Source Separation: Independent Component Analysis and Applications. Academic Press Inc. (2010)
2. Huang, X., Wu, H.C., Principe, J.C.: Robust blind beamforming algorithm using joint multiple matrix diagonalization. IEEE Sensors Journal 7(1), 130–136 (2007)

3. Zeng, W.J., Li, X.L., Zhang, X.D.: Direction-of-arrival estimation based on the joint diagonalization structure of multiple fourth-order cumulant matrices. IEEE Signal Processing Letters 16(3), 164–167 (2009)
4. Cardoso, J.F., Souloumiac, A.: Blind beamforming for non Gaussian signals. The IEE Proceedings of F 140(6), 363–370 (1993)
5. Cardoso, J.F.: On the performance of orthogonal source separation algorithms. In: Proceedings of the 9th European Signal Processing Conference, 776–779 (1994)
6. Pham, D.T.: Joint approximate diagonalization of positive definite Hermitian matrices. SIAM Journal on Matrix Analysis and Applications 22(4), 1136–1152 (2001)
7. Ziehe, A., Laskov, P., Nolte, G., Müller, K.R.: A fast algorithm for joint diagonalization with non-orthogonal transformations and its application to blind source separation. Journal of Machine Learning Research 5, 777–800 (2004)
8. Souloumiac, A.: Nonorthogonal joint diagonalization by combining Givens and hyperbolic rotations. IEEE Transactions on Signal Processing 57(6), 2222–2231 (2009)
9. Li, X.L., Adalı, T.: Blind separation of noncircular correlated sources using Gaussian entropy rate. IEEE Transactions on Signal Processing 59(6), 2969–2975 (2011)
10. Kleinsteuber, M., Shen, H.: Uniqueness analysis of non-unitary matrix joint diagonalization. IEEE Transactions on Signal Processing (2013)
11. Joho, M., Rahbar, K.: Joint diagonalization of correlation matrices by using Newton methods with application to blind signal separation. In: Proceedings of the 2nd IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM 2002), Rosslyn, VA USA, pp. 403–407 (2002)
12. van der Veen, A.J.: Joint diagonalization via subspace fitting techniques. In: Proceedings of the 26th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), pp. 2773–2776 (2001)
13. Afsari, B.: Simple LU and QR based non-orthogonal matrix joint diagonalization. In: Rosca, J.P., Erdogmus, D., Príncipe, J.C., Haykin, S. (eds.) ICA 2006. LNCS, vol. 3889, pp. 1–7. Springer, Heidelberg (2006)
14. Shen, H., Kleinsteuber, M.: Complex blind source separation via simultaneous strong uncorrelating transform. In: Vigneron, V., Zarzoso, V., Moreau, E., Gribonval, R., Vincent, E. (eds.) LVA/ICA 2010. LNCS, vol. 6365, pp. 287–294. Springer, Heidelberg (2010)
15. Belouchrani, A., Meraim, K.A., Cardoso, J.F., Moulines, E.: A blind source separation technique based on second order statistics. IEEE Transactions on Signal Processing 45(2), 434–444 (1997)
16. Absil, P.A., Gallivan, K.A.: Joint diagonalization on the oblique manifold for independent component analysis. In: Proceedings of the 31st IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Toulouse, France, V945–V948 (2006)
17. Absil, P.A., Mahony, R., Sepulchre, R.: Optimization Algorithms on Matrix Manifolds. Princeton University Press, Princeton (2008)
18. Liu, Y., Storey, C.: Efficient generalized conjugate gradient algorithms, part 1: Theory. Journal of Optimization Theory and Applications 69(1), 129–137 (1991)
19. Kleinsteuber, M., Hüper, K.: An intrinsic CG algorithm for computing dominant subspaces. In: Proceedings of the 32nd IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP, pp. IV1405–IV1408 (2007)
20. Yeredor, A.: Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation. IEEE Transactions on Signal Processing 50(7), 1545–1553 (2002)