An Algorithm for Joint Diagonalization by a Congruence Transformation [1]

Lieven De Lathauwer, Bart De Moor and Joos Vandewalle[2]

May 2002

Published in:
Proc. 23rd Symp. on Information Theory in the Benelux,
May 29-31, 2002, Louvain-la-Neuve, Belgium, pp. 235-240.

[1] This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/SISTA/delathauwer/reports/ldl-02-78.ps.Z*

[2] K.U.Leuven, Dept. of Electrical Engineering (ESAT), Research group SCD-SISTA, Kasteelpark Arenberg 10, 3001 Leuven (Heverlee), Belgium, Tel. 32/16/32 11 04, Fax 32/16/32 19 70, WWW: *http://www.esat.kuleuven.ac.be/sista-cosic-docarch*. E-mail: *Lieven.DeLathauwer@esat.kuleuven.ac.be*, *Bart.DeMoor@esat.kuleuven.ac.be*, *Joos.Vandewalle@esat.kuleuven.ac.be*. Lieven De Lathauwer holds a permanent research position with the French Centre National de la Recherche Scientifique (C.N.R.S.); he also holds a honorary post-doctoral research mandate with the Fund for Scientific Research-Flanders (F.W.O.), affiliated with the K.U.Leuven. Bart De Moor and Joos Vandewalle are Full Professor with the K.U.Leuven. This work is supported by several institutions: (1) Research Council K.U.Leuven: Concerted Research Action GOA-MEFISTO-666 (Mathematical Engineering for Information and Communication Systems Technology), (2) Flemish Government: the F.W.O. projects G.0240.99 (Multilinear Generalisations of the Singular Value Decomposition and Applications in Signal Processing and System Identification) and G.0256.97 (Numerical Algorithms for Subspace System Identification, Extension to Special Cases), the F.W.O. Research Communities ICCoS (Identification and Control of Complex Systems) and ANMMM (Advanced Numerical Methods for Mathematical Modelling), (3) Belgian Federal Government: Interuniversity Poles of Attraction Programme IUAP IV-02 (1997-2001): Modeling, Identification, Simulation and Control of Complex Systems, and IUAP V-22 (2002-2006): Dynamical Systems and Control: Computation, Identification & Modelling. The scientific responsibility is assumed by its authors.

Abstract

Simultaneous matrix decompositions are nowadays popular for solving a variety of signal processing problems, especially problems related to signal separation. In this paper we derive a new algorithm for the approximate diagonalization of a set of matrices by means of a simultaneous non-unitary congruence transformation.

An Algorithm for Joint Diagonalization by a Congruence Transformation

Lieven De Lathauwer Bart De Moor Joos Vandewalle*

Abstract

Simultaneous matrix decompositions are nowadays popular for solving a variety of signal processing problems, especially problems related to signal separation. In this paper we derive a new algorithm for the approximate diagonalization of a set of matrices by means of a simultaneous non-unitary congruence transformation.

1 Introduction

Approximate diagonalization of a set of R given $(N \times N)$ matrices by means of a simultaneous congruence transformation forms the basis of several algorithms for source separation, such as the Analytical Constant Modulus Algorithm (ACMA) [15], algorithms for higher-order-only Independent Component Analysis (ICA) [3], ... In addition, many well-known source separation algorithms are based on an approximate diagonalization by means of a simultaneous unitary similarity transformation; the fact that the transformation is unitary is induced by a prewhitening. Examples are the ICA-algorithm described in [2], the second-order-only blind source separation method (SOBI) presented in [1] and the algorithm for the blind separation of non-stationary sources proposed in [12]. Actually, it seems strange to consider some second-order constraints on the mixing matrix as infinitely more reliable than other second- or higher-order constraints. Instead, one could give the different constraints an appropriate (finite) weight and deal with them simultaneously instead of sequentially. We called this principle "soft whitening" in [8] (complementarily to "hard whitening" or "without whitening"). Soft whitening versions of [2, 1, 12] also take the form of an approximate diagonalization by means of a simultaneous non-unitary congruence transformation.

*L. De Lathauwer is with the research group ETIS, UPRES-A 8051, 6 avenue du Ponceau, BP 44, F 95014 Cergy-Pontoise Cedex, France. Tel: +33-(0)1-30.73.66.10; fax: +33-(0)1-30.73.62.82; e-mail: delathau@ensea.fr; URL: http://www-rech.ensea.fr. B. De Moor and J. Vandewalle are with the group SCD-SISTA of the E.E. Dept. (ESAT) of the K.U.Leuven, Kard. Mercierlaan 94, B-3001 Leuven (Heverlee), Belgium. Tel: +32-(0)16-32.17.09; fax: +32-(0)16-32.19.70; e-mail: {demoor,vdwalle}@esat.kuleuven.ac.be; URL: http://www.esat.kuleuven.ac.be/sista-cosic-docarch.

For the computation of this simultaneous decomposition, an extended QZ-iteration has been proposed in [15]. Jacobi-algorithms have been derived in [4, 5, 9]. In this paper we show that the problem can be reduced to a simultaneous unitary congruence transformation.

In Section 2 we discuss simultaneous *congruence, in Section 3 simultaneous ^{T}congruence, and in Section 4 joint simultaneous * and ^{T}congruence (we adopt the terminology used in [10]). Section 5 is a remark of a numerical nature.

2 Simultaneous *congruence

We are dealing with the following problem. Given matrices $\mathbf{A_1}, \mathbf{A_2}, \ldots, \mathbf{A_K} \in \mathbb{C}^{N \times N}$, we want to find $\mathbf{W} \in \mathbb{C}^{N \times N}$ such that

$$\begin{aligned} \mathbf{A_1} &= \mathbf{W} \cdot \mathbf{\Lambda_1} \cdot \mathbf{W}^H \\ &\vdots \\ \mathbf{A_K} &= \mathbf{W} \cdot \mathbf{\Lambda_K} \cdot \mathbf{W}^H, \end{aligned} \tag{1}$$

with $\mathbf{\Lambda_1}, \mathbf{\Lambda_2}, \ldots, \mathbf{\Lambda_K} \in \mathbb{C}^{N \times N}$ diagonal. We assume that $\mathbf{W}$ is full rank. (If $\mathbf{W}$ has more rows than columns, then one may first perform a dimensionality reduction using the techniques explained in [4, 7]). For $K \geqslant 3$, the problem is in general overdetermined and will be solved in an approximate sense.

Let the SVD of $\mathbf{W}$ be given by

$$\mathbf{W} = \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^H, \tag{2}$$

in which $\mathbf{U}, \mathbf{V} \in \mathbb{C}^{N \times N}$ are unitary and $\mathbf{\Sigma} \in \mathbb{C}^{N \times N}$ is diagonal. Define, for $k = 1, 2, \ldots, K$, $\tilde{\mathbf{A}}_\mathbf{k}$ as

$$\tilde{\mathbf{A}}_\mathbf{k} = \mathbf{\Sigma}^{-1} \mathbf{U}^H \mathbf{A_k} \mathbf{U} \mathbf{\Sigma}^{-1} = \mathbf{V}^H \mathbf{\Lambda_k} \mathbf{V}. \tag{3}$$

The fact that all these matrices are commuting, may be expressed as

$$\mathbf{A_k} \cdot \mathbf{M} \cdot \mathbf{A_l} = \mathbf{A_l} \cdot \mathbf{M} \cdot \mathbf{A_k}, \quad \forall k, l, \tag{4}$$

with

$$\mathbf{M} = \mathbf{U} \cdot \mathbf{\Sigma}^{-2} \cdot \mathbf{U}^H. \tag{5}$$

Hence we may first obtain $\mathbf{M}$ from the overdetermined set of linear equations (4), then compute $\mathbf{U}$ and $\mathbf{\Sigma}$ from the EVD (5), and finally estimate $\mathbf{V}$ from the simultaneous EVD (3).

The dimensionality of (4) is as follows. Because $\mathbf{M}$ is Hermitean, we have $N(N+1)/2$ complex unknowns. There are $R(R-1)/2$ different matrix equations, and each matrix equation provides $N(N+1)/2$ different complex linear equations — (4) can be rewritten as

$$\mathbf{A_k} \cdot \mathbf{M} \cdot \mathbf{A_l} = (\mathbf{A_k} \cdot \mathbf{M} \cdot \mathbf{A_l})^H. \tag{6}$$

To be a bit more accurate: since the diagonal of an Hermitean matrix is real, we have in fact $R(R-1)N^2/2$ real equations in N^2 real unknowns.

After calculating the best non-trivial solution to this set of linear equations (as the right singular vector corresponding to the smallest singular value of the coefficient matrix), and computing the EVD (5), the simultaneous EVD (3) remains. $\mathbf{V}$ may be obtained by means of a Jacobi-iteration, as explained in [2, 13]. The computation of an elementary rotation involves the computation of the dominant eigenvector of a real symmetric (3×3) matrix ($\{\mathbf{A_k}\}$ complex) or (2×2) matrix ($\{\mathbf{A_k}\}$ real).

In [15] the simultaneous congruence transformation (1) was translated into a simultaneous generalized Schur decomposition, for the computation of which an extended QZ-iteration was proposed. The latter involves an iteration over a sequence of SVDs of matrices of dimensions $(N\times K)$, $((N-1)\times K)$, $\ldots$, $(2\times K)$, respectively. An alternative is the Jacobi approach. For real matrices, computing the optimal rotation matrices amounts to the computation of the roots of a polynomial of degree 8 [4, 5]; for complex matrices, we have to compute the best rank-1 approximation of a partially symmetric complex $(2\times 2\times 2\times 2)$ tensor [9].

3 Simultaneous ^{T}congruence

For given $\mathbf{B_1}, \mathbf{B_2}, \ldots, \mathbf{B_P} \in \mathbb{C}^{N\times N}$, we may have, instead of Eqs. (1), a simultaneous decomposition of the type

$$\begin{aligned}\mathbf{B_1} &= \mathbf{W}\cdot\mathbf{\Omega_1}\cdot\mathbf{W}^T\\ &\vdots\\ \mathbf{B_P} &= \mathbf{W}\cdot\mathbf{\Omega_P}\cdot\mathbf{W}^T,\end{aligned}\tag{7}$$

with $\mathbf{\Omega_1}, \mathbf{\Omega_2}, \ldots, \mathbf{\Omega_P} \in \mathbb{C}^{N\times N}$ diagonal (e.g. [14]). Let the SVD of $\mathbf{W}$ again be given by Eq. (2). Then we define, for $p = 1, 2, \ldots, P$, $\tilde{\mathbf{B}}_\mathbf{p}$ as

$$\tilde{\mathbf{B}}_\mathbf{p} = \mathbf{\Sigma}^{-1}\mathbf{U}^H\mathbf{B_p}\mathbf{U}^*\mathbf{\Sigma}^{-1} = \mathbf{V}^H\mathbf{\Omega_p}\mathbf{V}^*.\tag{8}$$

The counterpart of (4) becomes:

$$\mathbf{B}_\mathbf{p}^*\cdot\mathbf{M}\cdot\mathbf{B}_\mathbf{q}^* = \mathbf{B}_\mathbf{q}^*\cdot\mathbf{M}\cdot\mathbf{B}_\mathbf{p}^*, \quad \forall p, q,\tag{9}$$

with $\mathbf{M}$ again given by (5).

Here we have, for the determination of $\mathbf{M}$, $R(R-1)N^2/2$ complex linear equations in $N(N+1)/2$ complex unknowns. (8) is a simultaneous Takagi factorization of the symmetric matrices $\{\tilde{\mathbf{B}}_\mathbf{p}\}$. In [6] we derived a Jacobi-algorithm for the determination of the optimal $\mathbf{V}$. It is similar to the Jacobi-algorithm for the simultaneous EVD (3): when $\{\mathbf{B_p}\}$ is complex, the calculation of an elementary rotation again involves the computation of the dominant eigenvector of a real symmetric (3×3) matrix; when $\{\mathbf{B_p}\}$ is real, both problems coincide. This result was independently rediscovered in [14].

4 Combined * and ^{T}congruence

In some problems, like the blind separation of non-circular sources [8] (the results may also be applied to, e.g., the problems addressed in [1, 12]), we have both Eqs. (1) and (7). $\mathbf{M}$ is then the solution to the resulting global set of linear equations. $\mathbf{V}$ is to be obtained from a joint simultaneous EVD / Takagi factorization. The real symmetric (3×3) matrix is here simply the sum of the matrices corresponding to the two subproblems.

5 Singular values of W are close

It is well-known that when eigenvalues of a matrix are close, the corresponding eigenvectors are ill-conditioned [11]. This does not necessarily pose a major problem for our method.

Call $\hat{\mathbf{M}}$ the estimate of the true matrix $\mathbf{M}$, with EVD

$$\hat{\mathbf{M}} = \hat{\mathbf{U}} \cdot \hat{\mathbf{\Sigma}}^{-2} \cdot \hat{\mathbf{U}}^H. \tag{10}$$

If

$$\hat{\mathbf{M}} = \mathbf{M} + O(\epsilon), \tag{11}$$

then we have

$$\hat{\mathbf{\Sigma}}^{-2} = \mathbf{\Sigma}^{-2} + O(\epsilon), \tag{12}$$

$$\hat{\mathbf{U}} = \mathbf{U} \cdot \mathbf{Q} + O(\epsilon), \tag{13}$$

in which $\mathbf{Q}$ is a block-diagonal matrix containing unitary blocks where eigenvalues are close. We have then

$$\hat{\mathbf{\Sigma}}^{-1}\hat{\mathbf{U}}^H\mathbf{A_k}\hat{\mathbf{U}}\hat{\mathbf{\Sigma}}^{-1} = \mathbf{Q}^H\mathbf{V}^H\mathbf{\Lambda_k}\mathbf{V}\mathbf{Q} + O(\epsilon) \tag{14}$$

such that the simultaneous diagonalization procedure may compensate for $\mathbf{Q}$.

6 Conclusion

In this paper we have derived a new algorithm for the approximate diagonalization of a set of R given $(N \times N)$ matrices by means of a simultaneous congruence transformation. It relies on the following 3 steps: (i) solving a big overdetermined set of linear equations, (ii) EVD of an Hermitian $(N \times N)$ matrix, (iii) joint diagonalization by a unitary congruence transformation of a set of R $(N \times N)$ matrices.

Acknowledgements

Lieven De Lathauwer holds a permanent research position with the French Centre National de la Recherche Scientifique (C.N.R.S.); he also holds a honorary post-doctoral research mandate with the Fund for Scientific Research-Flanders (F.W.O.), affiliated with the K.U.Leuven. Bart De Moor and Joos Vandewalle are Full Professor with the K.U.Leuven.

This work is supported by several institutions:

1. Research Council K.U.Leuven: Concerted Research Action GOA-MEFISTO-666 (Mathematical Engineering for Information and Communication Systems Technology),
2. Flemish Government: the F.W.O. projects G.0240.99 (Multilinear Generalisations of the Singular Value Decomposition and Applications in Signal Processing and System Identification) and G.0256.97 (Numerical Algorithms for Subspace System Identification, Extension to Special Cases), the F.W.O. Research Communities ICCoS (Identification and Control of Complex Systems) and ANMMM (Advanced Numerical Methods for Mathematical Modelling),
3. Belgian Federal Government: Interuniversity Poles of Attraction Programme IUAP IV-02 (1997-2001): Modeling, Identification, Simulation and Control of Complex Systems, and IUAP V-22 (2002-2006): Dynamical Systems and Control: Computation, Identification & Modelling.

The scientific responsibility is assumed by its authors.

References

[1] A. Belouchrani, et al., "A Blind Source Separation Technique using Second Order Statistics", *IEEE Trans. Signal Processing*, Vol. 45, No. 2, Feb. 1997, pp. 434–444.

[2] J.-F. Cardoso, A. Souloumiac, "Blind Beamforming for Non-Gaussian Signals", *IEE Proceedings-F*, Vol. 140, No. 6, 1994, pp. 362–370.

[3] L. De Lathauwer, B. De Moor, J. Vandewalle, "Independent Component Analysis Based on Higher-Order Statistics Only", *Proc. 8th IEEE SP Workshop on Statistical Signal and Array Processing (SSAP-96)*, Corfu, Greece, June 24–26, 1996, pp. 356–359.

[4] L. De Lathauwer, *Signal Processing Based on Multilinear Algebra*, Ph.D. thesis, K.U.Leuven, E.E. Dept., Belgium, 1997.

[5] L. De Lathauwer, J. Vandewalle, B. De Moor, "Jacobi-Algorithm for Simultaneous Generalized Schur Decomposition in Higher-Order-Only ICA", *Proc. IEEE Benelux Signal Processing Chapter Signal Processing Symposium (SPS'98)*, Leuven, Belgium, March 26–27, 1998, pp. 67–70.

[6] L. De Lathauwer, B. De Moor, J. Vandewalle, "ICA Techniques for More Sources than Sensors", *Proc. IEEE Signal Processing Workshop on Higher-Order Statistics (HOS'99)*, Caesarea, Israel, June 14-16, 1999, pp. 121–124.

[7] L. De Lathauwer, B. De Moor, J. Vandewalle, "On the Best Rank-1 and Rank-$(R_1, R_2, \ldots, R_N)$ Approximation of Higher-Order Tensors", *SIAM J. Matrix Anal. Appl.*, Vol. 21, No. 4, April 2000, pp. 1324–1342.

[8] L. De Lathauwer, B. De Moor, “On the Blind Separation of Non-circular Sources”, *Proc. XIth European Signal Processing Conf. (EUSIPCO-2002)*, Toulouse, France, Sept. 3–6, 2002.

[9] L. De Lathauwer, B. De Moor, J. Vandewalle, “Jacobi-iteration for the Complex Simultaneous Generalized Schur Decomposition”, Tech. Report ESAT/SCD-SISTA, K.U.Leuven, Belgium (in preparation).

[10] R.A. Horn, C.R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[11] T. Kato, *A Short Introduction to Perturbation Theory for Linear Operators*, Springer-Verlag, N.Y., 1982.

[12] D.-T. Pham, J.-F. Cardoso, “Blind Separation of Instantaneous Mixtures of Non-stationary Sources”, *IEEE Trans. Signal Processing*, Vol. 49, No. 9, pp. 1837–1848.

[13] A. Souloumiac, J.-F. Cardoso, “Jacobi Angles for Simultaneous Diagonalization”, *SIAM J. Matr. Anal. Appl.*, Vol. 17, No. 1, Jan. 1996, pp. 161–164.

[14] A. Souloumiac, E. Moreau, “Blind Source Separation via Joint Diagonalization of Symmetric Complex Matrices”, *Proc. 2002 IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP 2002)*, Orlando, FL, USA, May 13–17, 2002.

[15] A.-J. van der Veen, A. Paulraj, “An Analytical Constant Modulus Algorithm”, *IEEE Trans. Signal Processing*, Vol. 44, 1996, pp. 1136–1155.