Journal of Southeast University (English Edition) Vol. 22, No. 4, pp. 445 – 450 Dec. 2006 ISSN 1003—7985

Joint eigenvalue estimation by balanced simultaneous Schur decomposition

Fu Tuo Gao Xiqi

(National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China)

Abstract: The problem of joint eigenvalue estimation for the non-defective commuting set of matrices $\mathscr{A}$ is addressed. A procedure revealing the joint eigenstructure by simultaneous diagonalization of $\mathscr{A}$ with simultaneous Schur decomposition (SSD) and balance procedure alternately is proposed for performance considerations and also for overcoming the convergence difficulties of previous methods based only on simultaneous Schur form and unitary transformations. It is shown that the SSD procedure can be well incorporated with the balancing algorithm in a pingpong manner, i. e., each optimizes a cost function and at the same time serves as an acceleration procedure for the other. Under mild assumptions, the convergence of the two cost functions alternately optimized, i. e., the norm of $\mathscr{A}$ and the norm of the left-lower part of $\mathscr{A}$, is proved. Numerical experiments are conducted in a multi-dimensional harmonic retrieval application and suggest that the presented method converges considerably faster than the methods based on only unitary transformation for matrices which are not near to normality.

Key words: direction of arrival; multi-dimensional harmonic retrieval; joint eigenvalue; simultaneous Schur decomposition; balance algorithm

The problem of joint eigenvalue estimation for general non-defective matrices sharing the same set of eigenvectors is often encountered in many signal processing applications, e. g., the 2-D direction of arrival (DOA) estimation[1], the joint angle-delay estimation[2] and the multidimensional harmonic retrieval[3].

A number of algorithms have been proposed for solving this problem. For the specific case of two real matrices with real eigenvalues, the simple T algorithm is proposed in Ref. [4]. For the general case, one method is to compute the eigenvalues of each matrix individually and then associate them, see Ref. [5] and the references therein. The disadvantage of these algorithms is that association techniques are not reliable if certain conditions are not fulfilled or computationally prohibitive if the matrix dimension is not small due to combinatorial search. In Ref. [1], the algebraically coupled matrix pencil (ACMP) method is proposed. It computes the Schur form of the first matrix whose Schur vectors are used in the triangularization of other matrices. The simultaneous Schur decomposition algorithm cannot handle the case when close eigenvalues of the first matrix are perturbed by amounts which are not small with respect to their mutual distances. A similar strategy based on eigenvectors of one matrix[5] suffers the same problem.

Then it is sensible to handle the matrices simultaneously. Moreover, such a strategy will also benefit from the averaging effect if noise is present. In Ref. [3], a Jacobi-type algorithm tries to find the simultaneous Schur form of the matrices to be estimated by orthogonal similarity transformations. The simultaneous Schur decomposition algorithm (named the SSD algorithm) is extended to the complex case in Ref. [6]. Similar to the one-matrix case[7-8], which may converge only linearly or does not converge at all, this type of scheme suffers the same convergence difficulty as is shown in section 4. The algorithm in Ref. [9] tries to generalize the classical QR algorithm[10] to obtain the simultaneous Schur form by the simultaneous QR decomposition. However, unlike its one-matrix counterpart which is easily incorporated with acceleration strategies, this generalization loses essential properties for the one-matrix case and may result in a very slow convergence rate (if convergence occurs).

Another problem with the simultaneous Schur decomposition is that it may not correspond to a simultaneous eigenstructure; i. e., upper triangular matrices may not share the same set of eigenvectors. Hence, in the noisy case, even the cost function defined by the

Received 2006-04-28.

Foundation items: The National Natural Science Foundation of China (No. 60572072, 60496311), the National High Technology Research and Development Program of China (863 Program) (No. 2003AA123310), the International Cooperation Project on Beyond 3G Mobile of China (No. 2005DFA10360).

Biographies: Fu Tuo (1977—), male, doctor, associate research fellow, futuo@ seu. edu. cn; Gao Xiqi (1967—), male, doctor, professor, xqgao@ seu. edu. cn.

Euclidean norm of the below diagonal elements is minimized. This does not correspond a minimum least squares distance to the simultaneous eigenstructure. Thus the performance seems unpredictable and is hard to analyze.

In this paper, we propose a procedure revealing the joint eigenstructure by simultaneous diagonalization (the term "simultaneous diagonalization" is often used in the literature to denote the diagonalization of a set of matrices by congruence transformations[11]. Here, we use the term in the sense of simultaneous similarity diagonalization) of $\mathscr{A}$ with the SSD and balance procedure alternately for performance considerations and also for overcoming the convergence difficulties of previous methods based only on the simultaneous Schur form and unitary transformations. The procedure alternately optimizes two cost functions. One is the norm of $\mathscr{A}$, partly following the idea of Ref. [12] for the one-matrix case and is used to bring $\mathscr{A}$ closer to normality. The other is the norm of the left-lower part of $\mathscr{A}$, following the idea of Ref. [3] and is used for eigenvalue revealing. The key observation is that the two optimization sub-procedures can be well incorporated with each other in a pingpong manner, i. e., each optimizes a cost function and at the same time serves as an acceleration procedure for the other. The diagonalization of $\mathscr{A}$ is due to the fact that if a Schur form is balanced, it must be a diagonal form. We prove that the two cost functions are both convergent under mild assumptions. The balanced-SSD algorithm is provided as an alternative to the recently proposed shear-rotation algorithm[13].

1 Problem Formation

Consider a set $\mathscr{A} = \{\boldsymbol{A}_n \mid n = 1, 2, \ldots, N\}$ of N complex or real $M \times M$ matrices. When the matrices in $\mathscr{A}$ are diagonalizable commuting matrices, then $\mathscr{A}$ can be simultaneously diagonalized[14]. Hence, there is a matrix $\boldsymbol{P}$ such that

$$\boldsymbol{A}_n = \boldsymbol{P}\boldsymbol{\Lambda}_n\boldsymbol{P}^{-1} \qquad n = 1, 2, \ldots, N \tag{1}$$

where $\boldsymbol{\Lambda}_n$ is a diagonal matrix containing the eigenvalues of $\boldsymbol{A}_n$. We are interested in $\boldsymbol{\Lambda}_n$ and their associations; i. e., which eigenvalues correspond to the same simultaneous eigenvector.

In practice, $\mathscr{A}$ is corrupted by estimation errors due to noise and finite sample size effects. Then the off-diagonal elements can only be minimized but cannot generally be driven to zero by similarity transformations. The average eigenstructure corresponds only to an approximate simultaneous diagonalization. To gain insight into this average eigenstructure, we suppose that there are no identical joint eigenvalues (i. e., $\forall\, i_1, i_2 (1 \leqslant i_1 \neq i_2 \leqslant M)$, $\exists\, k (1 \leqslant k \leqslant N)$, such that $\lambda_{k(i_1)} \neq \lambda_{k(i_1)}$, where $\lambda_{k(i_1)}$ and $\lambda_{k(i_2)}$ are the i_1-th and i_2-th eigenvalue of $\boldsymbol{A}_k$ respectively) which is a case of significant practical importance. Taking the linear terms of the eigenvalue estimation error $\Delta\boldsymbol{\Lambda}_n$, see Ref. [2], we can express the estimation error of the i-th eigenvalue of $\boldsymbol{A}_n$ as

$$\Delta\lambda_{n(i)} = \boldsymbol{q}_i^{\mathrm{H}}\, \Delta\boldsymbol{A}_n \cdot \boldsymbol{t}_i \tag{2}$$

where $\boldsymbol{q}_i$ and $\boldsymbol{t}_i$ are the simultaneous left and right eigenvectors of error-free $\mathscr{A}$, respectively. Eq. (2) shows that the dominant term of the estimation error $\Delta\lambda_{n(i)}$ is only related to $\boldsymbol{q}_i$, $\boldsymbol{t}_i$ and $\Delta\boldsymbol{A}_n$ itself, but not related to the eigenvector error $\Delta\boldsymbol{q}_i$ and $\Delta\boldsymbol{t}_i$. This relation implies that the approximation to the exact minimizer of the off-diagonal elements is not critical to the estimation performance, assuming that the simultaneous left and right eigenvectors suffer only small perturbations.

In this paper, we focus on the case $\mathscr{A}$ and its eigenvalues are both real. Then $\mathscr{A}$ can be simultaneously diagonalized with a real $\boldsymbol{P}$ and all the calculations are in the real domain. The extension to the complex case is not difficult[6, 15].

2 Brief Review of the SSD Algorithm

We here present a brief introduction to the SSD algorithm proposed in Ref. [3] and make some remarks on the convergence property.

Let $\mathscr{A} = \{\boldsymbol{A}_n \in \mathbf{R}^{M \times M} \mid n = 1, 2, \ldots, N\}$ be the set of matrices to be simultaneously upper-trianglized and $\boldsymbol{\Theta}$ be an orthogonal matrix composed of a product of elementary Jacobi rotations

$$\boldsymbol{\Theta} = \prod_{\text{sweeps}} \prod_{q=2}^{M} \prod_{p=1}^{q-1} \boldsymbol{\Theta}_{qp} \tag{3}$$

where the classical Jacobi rotations $\boldsymbol{\Theta}_{qp}$ are defined by modifying the identity matrix with $\boldsymbol{\Theta}_{qp}(p, p) = \boldsymbol{\Theta}_{qp}(q, q) = c$ and $\boldsymbol{\Theta}_{qp}(p, q) = -\boldsymbol{\Theta}_{qp}(q, p) = s$, where $c = \cos\theta_{qp}$ and $s = \sin\theta_{qp}$. The objective of the SSD algorithm is to choose the rotation angle θ_{qp} at each particular step such that the following cost function ($\boldsymbol{\Theta}^{-1} = \boldsymbol{\Theta}^{\mathrm{T}}$ since $\boldsymbol{\Theta}$ is orthogonal)

$$\psi(\boldsymbol{\Theta}) = \| \mathscr{L}(\boldsymbol{\Theta}^{-1}\mathscr{A}\boldsymbol{\Theta}) \|_{\mathrm{E}}^2 = \sum_{n=1}^{N} \| \mathscr{L}(\boldsymbol{\Theta}^{-1}\boldsymbol{A}_n\boldsymbol{\Theta}) \|_{\mathrm{E}}^2 \tag{4}$$

is decreased as much as possible, where $\|\cdot\|_{\mathrm{E}}$ denotes the Euclidean norm and $\mathscr{L}(\cdot)$ denotes an operator that extracts the strictly lower triangular part of its matrix-valued argument by setting the upper triangular part and the elements on the main diagonal to zero. The optimal θ_{qp} can be chosen by solving a fourth-order polynomial equation as derived in Ref. [3].

Ref. [7] gives a detailed example showing that convergence difficulty arises for this kind of strategy when the norm of the upper-right part is considerably larger than the norm of lower-left part. The large unbalanced norm results in a very small feasible rotation angle of θ_{qp} at each step and makes this piecemeal optimization strategy difficult to converge.

Suppose that $\boldsymbol{A}_n$ has a large deviation from normality, by the measure of deviation from normality of Ref. [16],

$$\Delta_{\mathrm{E}}(\boldsymbol{A}_n) = \{ \| \boldsymbol{A}_n \|_{\mathrm{E}}^2 - \| \boldsymbol{A}_n \|_{\mathrm{E}}^2 \}^{\frac{1}{2}} \qquad (5)$$

is a large value. Then as the algorithm converges to a certain stage, the lower-left and diagonal part of $\boldsymbol{A}_n$ are relatively near to zero and $\boldsymbol{A}_n$, respectively. Since the SSD algorithm does not change $\| \boldsymbol{A}_n \|_{\mathrm{E}}^2$, by Eq. (5) we must have a large norm on the upper-right part of $\boldsymbol{A}_n$. Thus the convergence difficulty arises. This observation makes us to resort to a norm reducing procedure which serves as the basic idea for the next section.

3 Balanced SSD

3.1 Balance procedure

The balancing algorithm described here is an extension of Ref. [15] to multiple matrices in order to serve for the balanced-SSD algorithm. The balancing algorithm iteratively looks for a diagonal matrix $\boldsymbol{D}$ such that the Euclidean norm of the m-th row of $\boldsymbol{D}^{-1}\mathscr{A}\boldsymbol{D}$ is nearly equal to that of the m-th column, where $1 \leqslant m \leqslant M$. The pseudo-code is shown as

$(\mathscr{B}, \boldsymbol{D}) = \mathrm{Balance}(\mathscr{A})$
① $\boldsymbol{D} \leftarrow \boldsymbol{I}$
② $\mathscr{B} \leftarrow \mathscr{A}$
③ repeat
④ for m = 1 to M
⑤ $f \leftarrow \left(\dfrac{\| \mathscr{B}'(m,:) \|_{\mathrm{E}}}{\| \mathscr{B}'(:,m) \|_{\mathrm{E}}} \right)^{\frac{1}{2}}$
⑥ $\mathscr{B}'(:,m) \leftarrow \mathscr{B}'(:,m) f$
⑦ $\mathscr{B}'(m,:) \leftarrow \mathscr{B}'(m,:)/f$
⑧ $\boldsymbol{D}(m,m) \leftarrow \boldsymbol{D}(m,m) f$
⑨ endfor
⑩ until entries of $\boldsymbol{D}$ do not change much in an iteration

where $\mathscr{B}$ and $\mathscr{A}$ each denote a set of N matrices and

$$\| \mathscr{B}'(m,:) \|_{\mathrm{E}}^2 = \sum_{n=1}^{N} \| \boldsymbol{B}'_n(m,:) \|_{\mathrm{E}}^2 \qquad (6)$$

with the prime indicates the omission of diagonal elements. The handling of the zero denominator is the same as Ref. [15].

3.2 Balanced SSD algorithm

Now we can construct the balanced-SSD algorithm. Considering the effect of the SSD algorithm after one sweep, we tend to have a larger norm of the upper-right part of $\mathscr{A}$ and a smaller norm of the lower-left part as shown in Fig. 1(a). In Ref. [15], the author proved that balancing in the Euclidean norm is equivalent to minimizing the Euclidean norm and it is easy to verify that the larger f (line ⑤ in the pseudo-code of balance), the larger reduction of $\| \mathscr{A} \|_{\mathrm{E}}$ will be achieved. Examine the coefficient f for $m = 1$, by Fig. 1(a) we find that f^2 is the quotient of the norm of larger elements divided by the norm of smaller elements. This makes the balance procedure work effectively and $\mathscr{B}$ is closer to normality since the norm of $\mathscr{B}$ is reduced. Then examining f for $m = 2, 3, \ldots, M$, see Fig. 1(b), it is easily shown that the balance procedure also tends to be accelerated with the structure obtained by the SSD procedure. On the other hand, after the balance procedure, since we have balanced elements on the off-diagonal as shown in Fig. 1(c), the convergence difficulty caused by small feasible rotation angles due to the large unbalanced norm of the upper-right and lower-left part is also solved. Then the SSD procedure works more like the normal case[17] which has a good convergence property. Thus we have seen that the SSD procedure can be well incorporated with the balancing algorithm in a pingpong manner. This observation leads to the following balanced-SSD algorithm. For the SSD sub-procedure, there is no constraint on the rotation angle to be "inner" or "outer". We conjecture that the implied convergence property for the nonrepetition hypothesis (see section 1 of Ref. [18]), which is of great importance in practical applications, applies well to our case since $\mathscr{A}$ is iteratively brought to normality.

$(\mathscr{B}, \boldsymbol{D}) = \mathrm{BalSSD}(\mathscr{A})$
① $\mathscr{B} \leftarrow \mathscr{A}$
② repeat
③ $(\mathscr{B}, \boldsymbol{D}) = \mathrm{Balance}(\mathscr{B})$
④ $(\mathscr{B}) = \mathrm{SSD}(\mathscr{B}, \text{sweepnum})$
⑤ until stop criterion is reached

The parameter sweepnum of the function SSD(·) denotes the number of sweeps to be performed and can be set to 1 since the balance procedure is computationally more efficient.

$$\begin{bmatrix} d & l & l & l \\ s & d & l & l \\ s & s & d & l \\ s & s & s & d \end{bmatrix} \quad \begin{bmatrix} d & b & b & b \\ b & d & l & l \\ b & s & d & l \\ b & s & s & d \end{bmatrix} \quad \begin{bmatrix} d & b & b & b \\ b & d & b & b \\ b & b & d & b \\ b & b & b & d \end{bmatrix}$$

(a) (b) (c)

Fig. 1 Illustration for the balance procedure with $M = 4$. (a) $m = 1$; (b) $m = 2$; (c) After balance (d denotes diagonal element; l denotes larger element; s denotes smaller element; b denotes balanced element. Note that larger elements and smaller elements are in a probabilistic sense)

Remark Minimization of $\|\mathscr{A}\|_E$ is used to bring $\mathscr{A}$ closer to normality while minimization of $\|\mathscr{L}(\mathscr{A})\|_E$ is for eigenvalue revealing. The two optimization procedures each optimize a cost function and at the same time serve as an acceleration procedure for the other. Under mild assumptions, we prove that the two cost functions are both convergent as follows:

For $\|\mathscr{A}\|_E$, it is monotone decreasing for the balance procedure[15]. According to Ref. [19], $\|\mathscr{A}\|_E$ is lower bounded by the norm of the eigenvalues of $\mathscr{A}$. So the bounded monotone decreasing sequence $\|\mathscr{A}\|_E$ is convergent.

We make an empirical assumption that $\|\mathscr{A}\|_E$ converges within finite iterations since the balance procedure decreasing $\|\mathscr{A}\|_E$ is iteratively accelerated by the SSD procedure. Then the diagonal balancing matrix will be an identity matrix after the convergence of $\|\mathscr{A}\|_E$[15]. So ultimately $\|\mathscr{L}(\mathscr{A})\|_E$ will also be a monotone decreasing sequence and thus convergent since it is lower bounded by zero.

The convergence of $\|\mathscr{A}\|_E$ and $\|\mathscr{L}(\mathscr{A})\|_E$ implies that the two alternately optimization procedures will not constitute a loop and $\|\mathscr{A}\|_E$ and $\|\mathscr{L}(\mathscr{A})\|_E$ can be used as stop criterion for the balanced-SSD algorithm.

It should be noted that the convergence of the cost functions does not imply the global convergence of $\mathscr{A}$ to (approximate) diagonal form. Empirically, in our experience so far, it is observed that $\mathscr{A}$ does converge quickly to (approximate) diagonal form due to the effective pingpong manner of the SSD and the balance procedure, and we know that if a Schur form is balanced it must be a diagonal form. However, we have not been able to give a rigorous proof of this property.

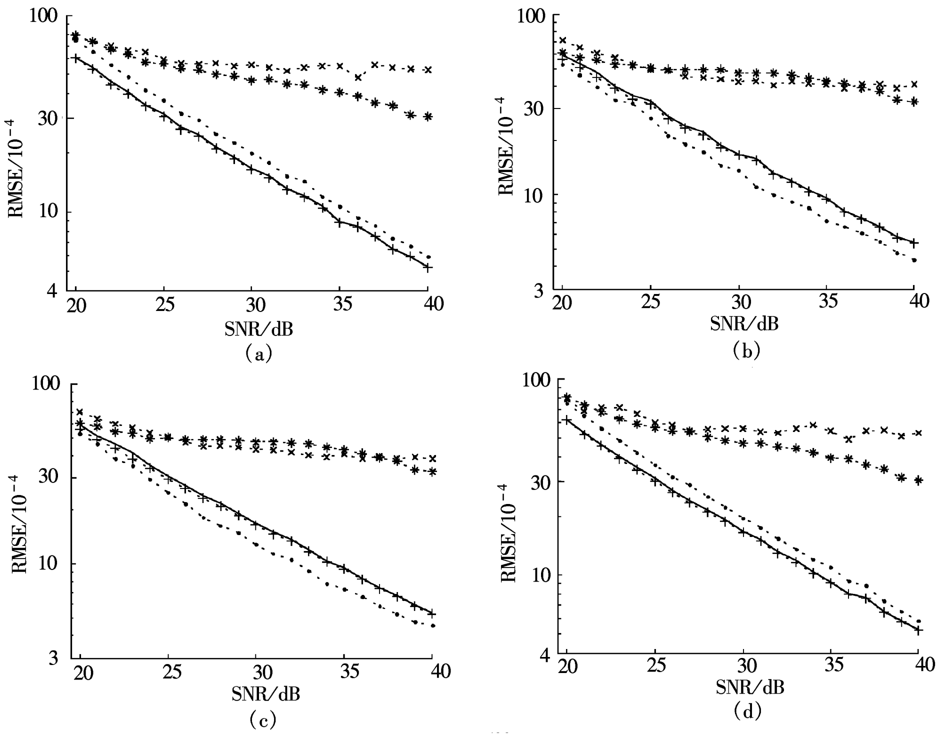
4 Simulation

We conduct our simulation in the scenario of a multidimensional harmonic retrieval application based on the multidimensional ESPRIT algorithm. The data model of Ref. [3] is applied where the signal s_i is assumed to be the white Gaussian process and is uncorrelated with each other. A 2-D uniform rectangular array (URA) with 6×6 elements is used. Four harmonic components are set as $\boldsymbol{\mu}_1 = \pi[0.30, 0.27]^T$, $\boldsymbol{\mu}_2 = \pi[0.32, 0.24]^T$, $\boldsymbol{\mu}_3 = \pi[0.34, 0.30]^T$, $\boldsymbol{\mu}_4 = \pi[0.36, 0.27]^T$. The number of snapshots is 512. SNR is defined as per source per element. 2 000 trials are conducted. Signal sub space estimation follows the real processing of Ref. [3]. To avoid frequency warping due to the Cayley transformation, the invariance equation is solved in the complex domain (except T algorithm[4]) by transforming the estimated real signal subspace back to complex domain with the inverse of the real-processing unitary matrix. In the joint eigenstructure estimation procedure, we will face two complex matrices which can be diagonalized with a real matrix $\boldsymbol{P}$. Thus, the problem is equivalent to simultaneously diagonalizing four real matrices by a real $\boldsymbol{P}$.

The plots in Figs. 2 and 3 show the RMSE (with $\boldsymbol{\mu}$ normalized by π) vs. the SNR for the harmonic estimates of various methods at different iteration counts. The resulting RMS errors of dimension 1 and 2 are depicted in Figs. 2 and 3, respectively. It is shown that the balanced-SSD algorithm (at each iteration of the balanced-SSD algorithm, the sub-iteration for SSD and balance are set as 1 and 5, respectively) converges to reliable estimates within 10 iterations where the SSD[3] and the simultaneous QR algorithm[9] still have large RMS errors even at 100 and 500 iterations, respectively. Unlike the one-dimensional case where the triangularization and diagonalization based method reach the same estimates if $\boldsymbol{P}$ is not constrained to be real, the ultimate performance of SSD with exhausted iterations (if convergence occurs) may be different from the balanced-SSD algorithm. The performance difference depends on each particular parameter and both algorithms have better and worse estimates than the other. The balanced-SSD algorithm is consistent with the T algorithm[4] when the latter works (the T algorithm may fail when $\boldsymbol{\mu}$ is relatively close to $\pm\pi$ and cannot be extended to cases of more than two dimensions), thus allowing a closed form performance analysis[2,20] while SSD eludes such an analysis.

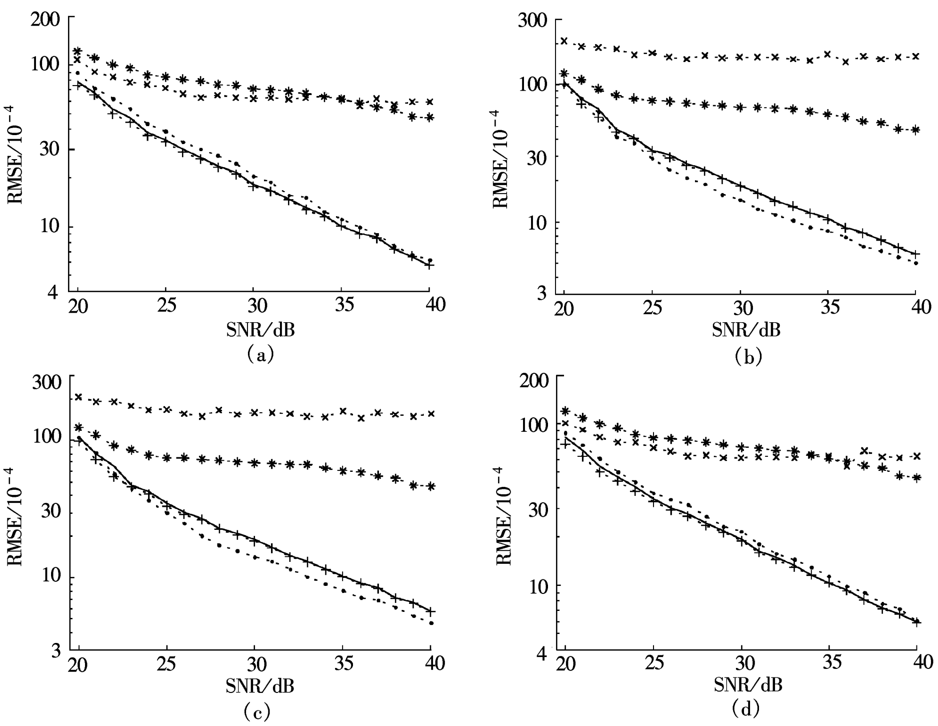
Numerous experiments are conducted while only one case is demonstrated here. It is shown that the convergence difficulty is severe for methods based on unitary transformation when the harmonic components are relatively close. This corresponds to a larger departure from normality of the matrices to be estimated. In a more extreme case, 20 harmonic components are equally spaced as $\boldsymbol{\mu}_1 = \pi[0, 0]^T$, $\boldsymbol{\mu}_2 = \pi[0.02, 0.02]^T, \ldots, \boldsymbol{\mu}_{20} = \pi[0.38, 0.38]^T$. For 20×20 URA, 256 snapshots and infinite SNR, the balanced-SSD algorithm will converge at around 40 iterations. While for SSD convergence is not observed even at one million iterations leaving the harmonics un-resolved.

The performance of the balanced-SSD algorithm is comparable to the shear-rotation algorithm (simulation results are omitted here in order to make Fig. 2 and Fig. 3 discriminable) of Ref. [13] and empirically the former appears to converge a bit faster.



——— T algorithm; ···×··· Simultaneous QR at 500 iterations; ···+·· Balanced-SSD at 10 iterations;
···✱·· SSD at 100 iterations; ···•··· Ultimate performance of SSD

Fig. 2 Performance comparison for dimension 1. (a) Source 1; (b) Source 2; (c) Source 3; (d) Source 4



——— T algorithm; ···×··· Simultaneous QR at 500 iterations; ···+·· Balanced-SSD at 10 iterations;
···✱·· SSD at 100 iterations; ···•··· Ultimate performance of SSD

Fig. 3 Performance comparison for dimension 2. (a) Source 1; (b) Source 2; (c) Source 3; (d) Source 4

5 Conclusion

We have proposed a simultaneous diagonalization algorithm to estimate the joint eigenvalues of non-defective matrices. The advantage of this algorithm is its good convergence property while computations cost per iteration increases only slightly since the computations of the balance procedure is proportional to $O(M^2)$ per matrix. And the performance is consistent with theoretical analysis. Numerical experiments suggest that the method presented here converges considerably faster than the methods based on only unitary transformation for matrices which are not near to normality.

References

[1] van der Veen A J, Ober P B, Deprettere E F. Azimuth and elevation computation in high resolution DOA estimation [J]. *IEEE Trans Signal Proc*, 1992, **40**(7): 1828 - 1832.

[2] Lemma A N, van der Veen A J, Deprettere E F. Analysis of joint angle-frequency estimation using ESPRIT [J]. *IEEE Trans Signal Proc*, 2003, **51**(5): 1264 - 1283.

[3] Haardt M, Nossek J A. Simultaneous Schur decomposition of several nonsymmetric matrices to achieve automatic pairing in multidimensional harmonic retrieval problems [J]. *IEEE Trans Signal Proc*, 1998, **46**(1): 161 - 169.

[4] Zoltowski M D, Haardt M, Mathews C P. Closed-form 2-D angle estimation with rectangular arrays in element space or beam space via unitary ESPRIT [J]. *IEEE Trans Signal Proc*, 1996, **44**(2): 316 - 328.

[5] Hua Y B, Abed-Meraim K. Techniques of eigenvalues estimation and association [J]. *Digital Signal Proc*, 1997, **7**(4): 253 - 259.

[6] Abed-Meraim K, Hua Y B. A least-squares approach to joint Schur decomposition [A]. In: *Proc IEEE ICASSP'98* [C]. Seattle, USA, 1998, **4**: 2541 - 2544.

[7] Froberg C E. On triangularization of complex matrices by two-dimensional unitary transformations [J]. *BIT Numer Math*, 1965, **5**: 230 - 234.

[8] Huang C P. A Jacobi-type method for triangularizing an arbitrary matrix [J]. *SIAM J Numer Anal*, 1975, **12**: 566 - 570.

[9] Strobach P. Bi-iteration multiple invariance subspace tracking and adaptive ESPRIT [J]. *IEEE Trans Signal Proc*, 2000, **48**(2): 442 - 456.

[10] Golub G H, Loan C F V. *Matrix computations* [M]. Johns Hopkins University Press, 1996.

[11] Yeredor A. Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation [J]. *IEEE Trans Signal Proc*, 2002, **50**(7): 1545 - 1553.

[12] Rutishauser H. Une méthod pour le calcul des valeurs propres des matrices non symétriques [J]. *Comptes Rendus des Séances de l'Academie des Sciences*, 1964, **259**: 2758.

[13] Fu T, Gao X Q. Simultaneous diagonalization with similarity transformation for non-defective matrices [A]. In: *Proc IEEE ICASSP'*2006 [C]. Toulouse, France, 2006, **4**: 1137 - 1140.

[14] Horn R A, Johnson C R. *Matrix analysis* [M]. Cambridge University Press, 1985. 52.

[15] Osborne E E. On pre-conditioning of matrices [J]. *Journal of the Association for Computing Machinery*, 1960, **7**(4): 338 - 345.

[16] Henrici P. Bounds for iterates, inverses, spectral variation, and fields of values of non-normal matrices [J]. *Numer Math*, 1962, **4**: 24 - 40.

[17] Bunse-Gerstner A, Byers R, Mehrmann V. Numerical methods for simultaneous diagonalization [J]. *SIAM J Matrix Anal Appl*, 1993, **14**(4): 927 - 949.

[18] Mascarenhas W F. On the convergence of the Jacobi method for arbitrary orderings [J]. *SIAM J Matrix Anal Appl*, 1995, **16**(4): 1197 - 1209.

[19] Mirsky L. On the minimization of matrix norms [J]. *Amer Math Monthly*, 1958, **65**: 106 - 107.

[20] Mathews C P, Haardt M, Zoltowski M D. Performance analysis of closed-form, ESPRIT based 2-D angle estimator for rectangular arrays [J]. *IEEE Signal Proc Letters*, 1996, **3**(4): 124 - 126.

基于平衡同时 Schur 分解的联合特征值估计

付　佗　　高西奇

（东南大学移动通信国家重点实验室,南京 210096）

摘要: 讨论可交换单纯矩阵族 $\mathscr{A}$ 的联合特征值估计问题. 为了克服基于同时 Schur 分解和酉变换算法的收敛和性能分析缺陷,提出了一种基于同时相似对角化的联合特征结构估计算法. 该算法通过对 $\mathscr{A}$ 交替进行同时 Schur 分解和范数平衡来实现矩阵族的对角化. 该算法的有效性在于:每个子过程在优化自身代价函数的同时,还对另一子过程的收敛起到加速作用. 在适当的假设条件下,可以证明该算法交替优化的 2 个代价函数(矩阵族范数和矩阵族下三角元素范数)的收敛性. 基于多维谐波提取的数值仿真显示该算法在矩阵族偏离正规阵时收敛速度显著快于基于同时 Schur 分解和酉变换算法,并且联合特征值的估计性能可以进行简洁的闭式分析.

关键词: 波达方向估计; 多维谐波提取; 联合特征值; 同时 Schur 分解; 平衡算法

中图分类号: TN911.7