THE SOURCE NUMBER ESTIMATION BASED ON GERSCHGORIN RADII

O. Caspary, P. Nus, T. Cecchin

Research Center for automatic control of Nancy, CRAN-CNRS
11, rue de l'Université, 88100 Saint-Dié, France

ABSTRACT

The GDE criterion is based on the estimation of the Gerschgorin disks' radii where the disks are separated in two distinct sets : one associated to the signal sources, the other related to the noise. We aim at modifying that criterion into a new one called SGDE by using the sum of the disks' radii. Besides, the SGDE criterion is modified with a simple deflation on the sum of the Gerschgorin radii to obtain a better estimation with sources of different power. We also suggest applying a deflation method to the covariance matrix before using the criteria based on the Gerschgorin radii. The transformed Gerschgorin radii can be connected to the Least-Squares through the transformed cross-correlation vector. So, two new criteria are put forward on the same principle as the SGDE criterion. These criteria can be applied in many situations : coloured or white noise, sources of different power.

1. INTRODUCTION

In spectral analysis, the estimation of the source number or frequency components is a crucial problem because the performances of most of the high resolution methods depend on the dimensions of signal and noise subspaces (MUSIC, ESPRIT, etc.). A lot of criteria have been studied in literature so far, among which direct criteria based on the eigenvalues profile of the covariance matrix [1], more elaborated criteria considering the statistical properties of noise (AIC, MDL, etc.) [2] and, recently, a criterion resting on the profile of the eigenvalues of noise [3]. However, if the number of samples is small, if the noise is not a Gaussian white noise, and if the sources are of very different power, then most of the criteria lose their efficiency or cannot be applied any longer. Still, there remains a new criterion, the GDE criterion (Gerschgorin Disk Estimator), the efficiency of which is not linked at all to the nature of the noise [4]. As its name indicates, this criterion makes use of the Gerschgorin theorem according to which the eigenvalues λ of a squared matrix A of a dimension (N,N) belong to the union of the N Gerschgorin disks, that is to say :

$$\lambda \in \bigcup_{k=1}^{N} D_k \tag{1}$$

where each Gerschgorin disk D_k is defined on the complex plane by a radius :

$$r_k = \sum_{\substack{l=1 \\ l \neq k}}^{N} \left| a_{kl} \right| \text{ and a center } O_k = a_{kk} \tag{2}$$

Yet, in reference [4], the authors have shown that the Gerschgorin theorem cannot be applied as such to estimate the number of sources, because the disks are tightly interlacing and the radii are so long that the signal and noise subspaces cannot be distinguished. Indeed, the membership of the eigenvalues associated to noise can be confused with the belonging field of the eigenvalues associated to signal. The following example shows the interlacing. We consider a signal x(t) of 20 samples composed of two sinusoids and of a white Gaussian noise n(t) of unitary variance, such as :

$$x(t) = A_1 \sin(2\pi f_1 t) + A_2 \sin(2\pi f_2 t) + n(t) \tag{3}$$

where $t = 1, 2, \ldots, 20$, $A_1 = A_2 = 4{,}47$ (10 dB) and the normalized frequencies are $f_1 = 0.20$, $f_2 = 0.25$. The covariance matrix C of size (8,8) is under the backward-forward form. The Gerschgorin theorem applied to C gives the figure 1.

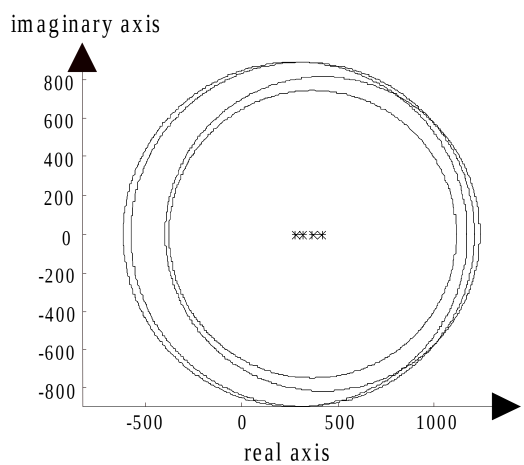

Fig. 1. Gerschgorin disks of C (∗ : disk centers).

That is why, in reference [4], a solution requiring a unitary transformation of the covariance matrix has been put forward. This solution does not modify the eigenvalues and reads as follows :

Let C be a covariance matrix :

$$C = \begin{pmatrix} C_1 & \mathbf{c} \\ \mathbf{c}^{\mathbf{H}} & c_{NN} \end{pmatrix} \tag{4}$$

where **c** is the last column of C except for the last element, then there is a unitary matrix :

$$U = \begin{pmatrix} U_1 & \mathbf{0} \\ \mathbf{0}^{\mathbf{H}} & 1 \end{pmatrix} \tag{5}$$

so that the transformed covariance matrix becomes :

$$S = U^H C\, U = \begin{pmatrix} S_1 & U_1^H \mathbf{c} \\ \mathbf{c}^{\mathbf{H}} U_1 & c_{NN} \end{pmatrix} \qquad (6)$$

The diagonal matrix S_1 contains eigenvalues λ' of the submatrix C_1. As U is a unitary matrix, the eigenvalues of S are the same as those of C, and it has been shown that the Gerschgorin disks have the following radii and centers :

$$r_k = \left| U_1^H \mathbf{c} \right| \quad , \quad O_k = \lambda' \qquad (7)$$

for k=1, ... , N-1. In theory, in the case of an infinite number of samples, the radii associated to the noise subspace are null, whereas those associated to the signal subspace are not. In practice, as the number of samples is finite, we look for a set of disks with long radii, for the signal subspace, which is to be distinct from a second set of disks with small radii, for the noise subspace. With the unitary transform, the figure 1 becomes :

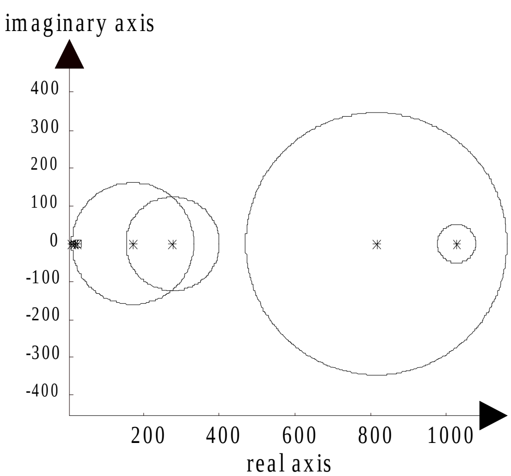

Fig. 2. Gerschgorin disks of C after unitary transform.

On the one hand, the disk centers associated to the signal subspace are distinct from the other centers gathered around zero. On the other hand, their radii are longer. We distinghish 4 sources corresponding to the 2 sinusoids.

To estimate the number of sources, we can search for a reasonable threshold to separate the two sets of disks or to apply the GDE criterion defined thus :

$$GDE(k) = r_k - \frac{F(L)}{N-1} \sum_{i=1}^{N-1} r_i \qquad (8)$$

where k = 1, ... , N-1. F(L) is a constant or an adjustable function, depending on the signal number of samples L. Generally, F(L)=1. This criterion is based on the hypothesis that the source number M is such that M<N-1. The source number is determined when the first negative value of GDE(k) is reached, so that we have M=k-1.

2. THE SGDE CRITERION

However, the GDE criterion is not satisfactory because, contrary to the eigenvalues that are arranged in a descending order, the radii are not. For a significative eigenvalue, we can have a smaller radius than those associated to other significative eigenvalues. Consequently, the GDE criterion stops at the first negative value, although the other values can still be positive. To avoid this situation as far as possible, we propose the SGDE (Sum of Gerschgorin Disk Estimators) criterion as follows :

$$SGDE(k) = \sum_{i=k}^{N-1} r_i - \frac{F(L)}{N-1} \sum_{l=1}^{N-1} \sum_{i=l}^{N-1} r_i \qquad (9)$$

This criterion is more robust than the GDE criterion and provides better results in all situations. What is more, we modify this criterion with a simple deflation technique on the sum of the Gerschgorin radii to obtain a better estimation with sources of different power. The deflation prevents the longest radii from hiding the others. Then, the previous criterion, associated to deflation and called SGDE_D, becomes :

$$SGDE\_D(k) = \left(\sum_{i=k}^{N-1} r_i - r_k \right) - \frac{F(L)}{N-1} \sum_{l=1}^{N-1} \left(\sum_{i=l}^{N-1} r_i - r_k \right) \qquad (10)$$

The source number is also determined when the first negative value is reached, so that M=k.

3. LINK WITH THE LEAST-SQUARES

For the standard Least-Squares solution, it is well-known that the minimization of the error $\boldsymbol{\varepsilon}$ in the following equation :

$$\boldsymbol{\varepsilon} = \mathbf{d} - X\, \mathbf{w} \qquad (11)$$

where X is a data matrix and $\mathbf{w}$ a weight vector, corresponds to the maximisation of the Euclidean norm E_y of the approximation $\mathbf{y}$ with $\mathbf{y} = X\,\mathbf{w}$, that is to say :

$$E_y = \mathbf{w}^{\mathbf{H}} X^H X\, \mathbf{w} = \mathbf{w}^{\mathbf{H}} C_1 \mathbf{w} \qquad (12)$$

where the covariance matrix C_1 of dimension (N-1,N-1) is defined by $C_1 = X^H X$. The LS solution for $\mathbf{w}$ is also defined by :

$$\mathbf{w} = C_1^{-1}\, \mathbf{c} \qquad (13)$$

with $\mathbf{c} = X^H \mathbf{d}$. If we replace this equation in (12), we obtain :

$$E_y = \mathbf{c}^{\mathbf{H}} C_1^{-1} \mathbf{c} = \mathbf{c}^{\mathbf{H}} U_1 S_1^{-1} U_1^H \mathbf{c} \qquad (14)$$

and the transformed cross-correlation vector introduced in the equation 5 of [5] corresponds to the equation of the radius r_k ($r_k = U_1^H \mathbf{c}$). So, the Euclidean norm E_y is defined by :

$$E_y = \sum_{k=1}^{N-1} \frac{r_k^2}{\lambda'_\kappa} = \frac{1}{\pi} \sum_{k=1}^{N-1} \frac{S_k}{\lambda'_\kappa} = \sum_{k=1}^{N-1} \theta_k \qquad (15)$$

where S_k is the surface of the Gerschgorin disks : $S_k = \pi r_k^2$ and where θ_k is a measure of the energy projected along the basis vectors forming U_1. So, the Euclidean norm uses the surface of the Gerschgorin disks, weighted by the inverse of the eigenvalues of C_1. We have a link between the transformed Gerschgorin radii and the measure of the energy θ_k. So, like for the GDE criterion, we can define the GDEE criterion (Gerschgorin Disk with Energy Estimators), the SGDEE criterion and its version with a simple deflation (SGDEE_D : Sum of Gerschgorin Disks with Energy Estimators and Deflation) :

$$\mathrm{GDEE}(k) = \theta_k - \frac{F(L)}{N-1} \sum_{i=1}^{N-1} \theta_i \qquad (16)$$

$$\mathrm{SGDEE}(k) = \left(\sum_{i=k}^{N-1} \theta_i \right) - \frac{F(L)}{N-1} \sum_{l=1}^{N-1} \left(\sum_{i=l}^{N-1} \theta_i \right) \qquad (17)$$

$$\mathrm{SGDEE\_D}(k) = \left(\sum_{i=k}^{N-1} \theta_i - \theta_k \right) - \frac{F(L)}{N-1} \sum_{l=1}^{N-1} \left(\sum_{i=l}^{N-1} \theta_i - \theta_k \right) \qquad (18)$$

For the GDEE criterion, the source number estimation M depends on the first negative value k , so that M = k-1. It is the same condition for the SGDEE and SGDEE_D criteria, except that M = k.

4. A DEFLATION TECHNIQUE

This technique makes it possible to separate slightly more easily the ″signal″ eigenvalues from the ″noise″ eigenvalues with a deflation technique of C, but the cost of calculation is higher. Let us add that all criteria described previously can be associated to this technique. The method consists in taking off the eigenvector of C corresponding to the highest eigenvalue to form a projection matrix P on the supposed noise subspace. Then, C is projected on this subspace. This method is repeated up to the last eigenvector and can be summed up by the following description :

1. eigendecomposition of C(N,N).

2. deflation technique :

for i = 1 to N

2.1 projection matrix $P = U_p U_p^H$ (with U_p containing the N+1-i last eigenvectors of U).

2.2 projection : $C_p = P\,C$.

2.3 formation of the matrix C_1 and of vector **c**.

2.4 eigendecomposition of C_1 in eigenvalues λ’ and eigenvectors (matrix U_1).

2.5 Calculus of radii from (7).

2.6 Calculus of variables for the criteria (for example for the GDE criterion : $r_p(i) = r_1$)

end_for

3. Calculus of criteria (for example for the GDE : $\mathbf{r} = \mathbf{r_p}$).

The deflation technique applied to C improves the results presented in figure 2 : 4 disks are very distinct from the others (see fig. 3).

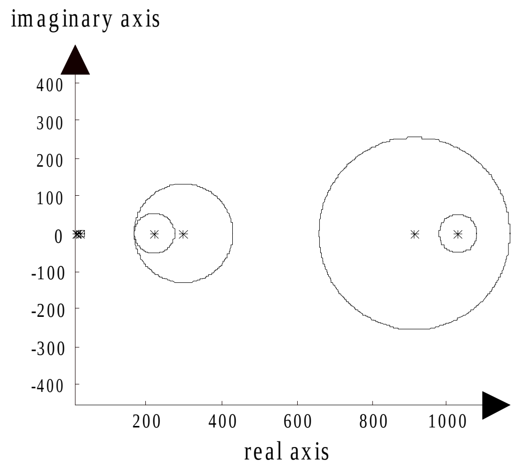

Fig. 3. Gerschgorin disks after the deflation of C.

The deflation enables to reduce the radii associated to the noise subspace, without any important consequences for the radii associated to the signal subspace. It provides the best results when it is associated to the SGDE_D and SGDEE_D criteria. There is no figure in this paper by lack of room.

5. SIMULATION RESULTS

The signal under consideration x(t) has 32 samples and contains 2 sinusoids of different power. Their normalized frequencies are $f_1 = 0.25$ and $f_2 = 0.27$. The following equation describes the signal x(t) :

$$x(t) = A_1 \sin(2\pi f_1 t) + A_2 \sin(2\pi f_2 t) + n(t) \qquad (19)$$

where t = 1, 2, ... , 32 , $A_1 = 4{,}47$ (10 dB) and $A_2 = A_1$ or varies from 10 dB to 25 dB according to the simulations. The estimated covariance matrix C, of dimension N = 16, is in the modified covariance form. 200 simulations of Monte-Carlo are carried out. n(t) is a Gaussian white noise with a unitary variance. In simulations, the noise n(t) is filtered by an AR(1) (coefficient 0.9) to obtain a Gaussian nonwhite noise.

The criteria GDE and GDEE are not presented here because their performances are always inferior to their improved version SGDE and SGDEE. In the case of Gaussian white noise and sources of equipower, the SGDE and SGDEE criteria without deflation have a detection rate superior by 3 dB compared to the same criteria with deflation (see fig. 4). But when the sources have a different power, the criteria with deflation show a difference of power superior to 10 dB and the SGDEE_D criterion gives the best results (see fig. 5). In the case of Gaussian nonwhite noise and sources of identical power, the criteria SGDE and SGDEE provide better rates of detection than SGDE_D and SGDEE_D, with an advantage for the SGDE criterion (see fig. 6). Nevertheless, if the sources have a different power, the

SGDE_D and SGDEE_D criteria are better (see fig. 7). We have 3 dB of improvement with SGDE_D and 7 dB with SGDEE_D compared to SGDEE. The criteria are not directly sensitive to the noise. So, we note little difference in performances between the figures 4 and 6, and the figures 5 and 7. The maximum detection rate remains around the well-known threshold of 5 dB.

6. CONCLUSIONS

The proposed SGDE criterion improves the performances of the GDE criterion. Associated to a simple deflation, this criterion is far less sensitive to the difference of power between the sources. Moreover, we have shown that the transformed Gerschgorin radii can be connected to the Least-Squares through the Euclidean norm. Then, we have deduced the family of the SGDEE criteria which is based on the surface of the Gerschgorin disks instead of the radii. This family presents the best results when the power of sources is not equal, and the results obtained in the case where the powers are equal or almost equal can be improved thanks to the deflation technique described in 4. The proposed criteria can be applied in most of the simulated or real cases.

Without deflation : SGDE ( ) and SGDEE ( )

With deflation : SGDE_D ( ) and SGDEE_D ( )

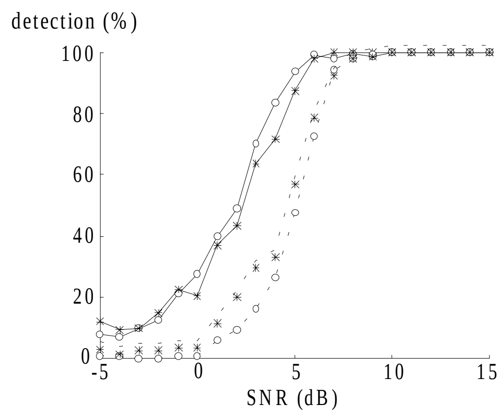

Fig. 4. Probability of detection with Gaussian white noise and $A_1 = A_2$ from -5 dB to +15 dB.

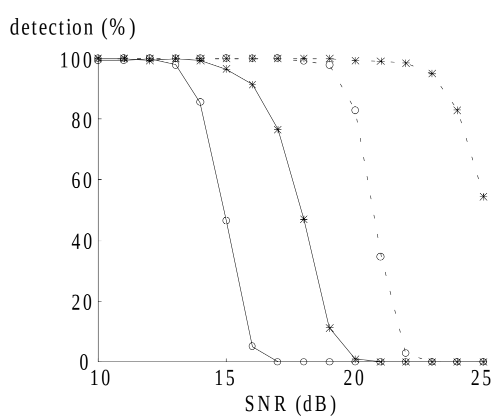

Fig. 5. Probability of detection with Gaussian white noise (A_1 = 10 dB, A_2 varies from 10 dB to 25 dB on the x-axis).

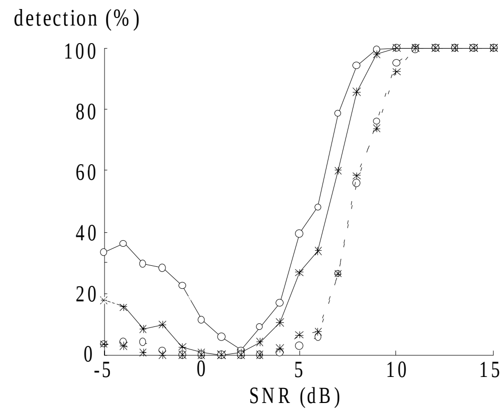

Fig. 6. Probability of detection with Gaussian nonwhite noise and $A_1 = A_2$ from -5 dB to +15 dB.

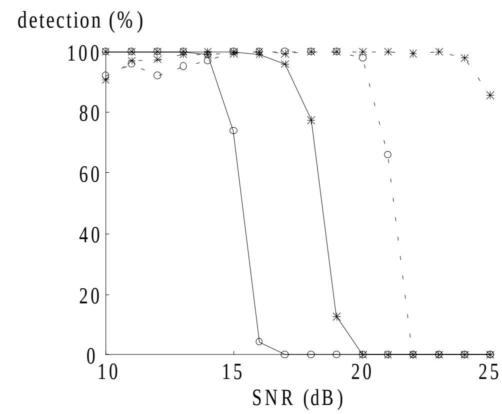

Fig. 7. Probability of detection with Gaussian nonwhite noise (A_1 = 10 dB, A_2 varies from 10 dB to 25 dB on the x-axis).

REFERENCES

[1] Konstantinides K., Yao K. "*Statistical analysis of effective singular values in matrix rank determination*". IEEE Trans. on ASSP, vol. 36, n°5, pp. 757-763, May 1988.

[2] Wax M., Kailath T. "*Detection of signals by information theoretic criteria*". IEEE Trans. on ASSP, vol. 33, n°2, pp. 387-392, April 1985.

[3] Grouffaud J., Larzabal P., Clergeot H. "*Some properties of ordered eigenvalues of a Wishart matrix : application in detection test and model order selection*". Proc. of ICASSP, Atlanta, 7-9 May, 1996.

[4] Wu H-T., Yang J-F., Chen F-K. "*Source number estimators using transformed Gerschgorin radii*". IEEE Trans. on SP, vol. 43, n°6, p. 1325, June 1995.

[5] Goldstein J. S., Reed I. S. "*Reduced-rank adaptive filtering*". IEEE Trans. on SP, vol. 45, n°2, pp. 492-496, February 1997.

[6] Caspary O., Nus P. "*Estimation du nombre de composantes d'un signal à l'aide du critère GDE amélioré*". Proc. of 16th GRETSI, Grenoble, p. 535, September 1997. (in French)