GENERALIZED GERSCHGORIN'S THEOREM FOR SOURCE NUMBER DETECTION

Olivier Caspary
CRAN, H. Poincaré University, Nancy 1,
11, rue de l'Université, 88100 Saint-Dié, France
Tel: +33 3 29516538; fax: +33 3 29516533
e-mail: olivier.caspary@iutsd.uhp-nancy.fr

ABSTRACT

A new family of source number estimators has appeared from the information provided by Gerschgorin radii and the centers of a unitary transformed covariance matrix. We suggest using a generalization of Gerschgorin's theorem developed for the eigenvalue problem $Ax = \lambda Bx$. This generalization can be applied to the perturbation of multiple eigenvalues and the usual theorem of Gerschgorin appears only as a particular case. For this, we need defining regions that bound a distance called the chordal metric. The techniques of diagonalization based on unitary transformation are necessary to exploit the estimated covariance matrix too. With sinusoidal signals embedded in a colored noise, the used criterion GDE_{dist} with this generalization shows a better detection rate compared to that obtained by the simple Gerschgorin theorem.

1 INTRODUCTION

In most of the applications containing harmonic signals, we must determine their number to separate them from the noisy background and then to estimate their parameters like frequency. In spectral analysis, it is a crucial problem for the high resolution methods such as MUSIC, ESPRIT where we must truncate the estimated covariance matrix in signal and noise subspaces. To cope with this problem, several source number detection criteria have been proposed in literature. Our choice concerns the criteria based on the Gerschgorin theorem applied to a covariance matrix. Those criteria are elaborated from the matrix perturbation theory and more particularly from the inclusion regions of eigenvalues of a matrix using the Gerschgorin theorem. By definition, the inclusion region of a matrix is a region of the complex plan that contains at least one of its eigenvalues. The eigenvalues $\lambda_1, ..., \lambda_n$ of a square matrix A, where $A \in \mathbb{C}^{nxn}$, depend on the elements of A and are the roots of the characteristic polynomial of A from the Cayley-Hamilton theorem. So, any perturbation that modifies one or several elements of A can change the eigenvalues in a small or high proportion. It is this notion of "proportion" that is taken into account by the inclusion regions and that we consider as useful information. We can underline that it is not necessary to search the smallest inclusion regions; in fact, we tend to the estimated eigenvalues provided by the numerical methods and the further information brought by the inclusion regions can become trivial. Moreover, the eigenvalues of the estimated covariance matrix are different from those of the theoretical covariance matrix, in particular about the multiplicity of eigenvalues associated to the noise subspace. Section 2 describes the Generalized Gerschgorin theorem used to define our inclusion regions. Now, two matrices are required. In section 3, we show a possible application of this generalization to detect the source number. In section 4, simulation results are given with criteria based on the GDE form. Finally, the last section deals with the conclusion.

2 THE GENERALIZED GERSCHGORIN'S THEOREM

The inclusion regions $\mathcal{G}_i$ of the eigenvalues λ_i of the matrix $A = [a_{ij}]$ of order N submitted to a perturbation matrix E, can be described by the following Gerschgorin theorem :

$$\mathcal{G}_i = \{\lambda \in \mathbb{C} : \mid \lambda - a_{ii} \mid \leq R_i = \|\mathbf{a}_i\|_1\} \qquad (1)$$

where $\|\mathbf{a}_i\|_1$ is the i^{th} row vector of $A - diag(A)$ and $diag(A)$ is the matrix that contains the main diagonal of A. These regions are called Gerschgorin disks with each disk D_i defined by a radius R_i and a center $O_i = a_{ii}$. The eigenvalues λ_i belong to the union of the N disks. However, the direct application of this theorem does not enable us to detect the source number from the disks because the radii are higher and the disks overlapped. Different solutions have been put forward in literature using a unitary transformation matrix applied to the covariance matrix to transform it into an almost diagonal matrix [1]. An extension of the Gerschgorin theorem is possible for the generalized eigenvalue problem $Ax = \lambda Bx$, always based on the perturbation theory for multiple eigenvalues. Let (A, B) be a regular pair, then we have the following theorem [2] :

$$\mathcal{D}_i = \left\{ \langle\alpha,\beta\rangle : | \beta a_{ii} - \alpha b_{ii} | \leq \sum_{j=1, j\neq i}^{N} | \beta a_{ij} - \alpha b_{ij} | \right\} \quad (2)$$

where the generalized eigenvalues of (A, B), let $\lambda(A,\ B) = \{\alpha_i/\beta_i : \beta_i \neq 0\}$, are included in the union of the N regions $\mathcal{D}_i$. If $B = I$, I being the identity matrix, we obtain the equation (1). Moreover, we have the following property again : if the regions $\mathcal{D}_i$ are pairwise disjoint, then each captures exactly one eigenvalue. However, as (α,β) appears on both sides, the regions $\mathcal{D}_i$ cannot be directly computed. But a judicious way has been found by Stewart to remove this dependence, based on the properties of norms and Cauchy's inequality (see [3], p. 296 for the demonstration). With $\mathbf{a}_i$ the row vectors of $A - diag(A)$ and $\mathbf{b}_i$ the row vectors of $B - diag(B)$, we have the bound :

$$\rho_i = \sqrt{\frac{\|\mathbf{a}_i\|_1^2 + \|\mathbf{b}_i\|_1^2}{|a_{ii}|^2 + |b_{ii}|^2}} \quad (3)$$

for the regions $\mathcal{G}_i$ defined by :

$$\mathcal{G}_i = \{ \langle\alpha,\beta\rangle : \mathcal{X}(\langle\alpha,\beta\rangle, \langle a_{ii}, b_{ii}\rangle) \leq \rho_i \} \quad (4)$$

and the regions $\mathcal{D}_i$ are included in $\mathcal{G}_i$. The chordal metric $\mathcal{X}$ is defined by :

$$\mathcal{X}(\langle\alpha,\beta\rangle, \langle a_{ii}, b_{ii}\rangle) = \frac{| \beta a_{ii} - \alpha b_{ii} |}{\sqrt{|a_{ii}|^2 + |b_{ii}|^2}\sqrt{|\alpha|^2 + |\beta|^2}} \quad (5)$$

This metric, never or rarely employed in signal processing, can be also used to measure the perturbation of eigenvalues. It corresponds to the distance between two matrix pairs. From the equation (3), we notice that the regions $\mathcal{G}_i$ only depend on the elements of the pair (A, B). So, we can avoid the computation of the generalized eigenvalues. The unitary transformations of the covariance matrix, previously applied, are always valid. So, in the next section, we show one possibility of exploiting the information provided by the regions $\mathcal{G}_i$.

3 DETECTION BASED ON THE COVARIANCE MATRICES

3.1 Choice of the matrices

The generalized eigenvalues are submitted to perturbation matrices (E, F) such as $A \rightarrow A_\varepsilon$ and $B \rightarrow B_\eta$ where $A_\varepsilon = A+E = D_A+\varepsilon\, O_A$ and $B_\eta = A+F = D_B+\eta\, O_B$. $D_A = diag(A)$ and O_A is composed of the off-diagonal elements of A. In the same way for D_B and O_B. We have $0 \leq (\varepsilon, \eta) \leq 1$ such as $(\varepsilon = 1, \eta = 1)$ gives (A_ε, B_η). Our goal is to reduce the values (ε, η) such as if (ε, η) tends to 0, we obtain the estimated generalized eigenvalue λ in the expression :

$$(D_A + \varepsilon\, O_A)x = \lambda(D_B + \eta\, O_B)x \quad (6)$$

or, under another form :

$$(D_B + \eta\, O_B)^{-1}(D_A + \varepsilon\, O_A)\, x = \lambda\, x \quad (7)$$

A problem of matrix conditioning may crop us in this equation. We must choose the matrix that has the best condition number to B_η between the pair (A, B) if this number is different. To connect this equation to our problem of source number detection, we must choose two matrices. Naturally, we would rather choose the covariance matrix that can be written under different forms. A possibility consists in taking the covariance matrix C of dimension (N,N) and a matrix called C_T obtained after a unitary transformation of C, such as $C_T = U^H C U$ with a unitary matrix U and a partition of C defined by :

$$U = \begin{pmatrix} U_1 & \mathbf{0} \\ \mathbf{0}^H & 1 \end{pmatrix},\ C = \begin{pmatrix} C_1 & \mathbf{c} \\ \mathbf{c}^H & c_{NN} \end{pmatrix} \quad (8)$$

where the vector $\mathbf{c}$ is the last column of C except the element c_{NN} and U_1 contains the eigenvectors of C_1 obtained after an eigendecomposition . The expression of C_T is :

$$C_T = \begin{pmatrix} S_1 & U_1^H\mathbf{c} \\ \mathbf{c}^H U_1 & c_{NN} \end{pmatrix} = \begin{pmatrix} \lambda_1^{'} & & & R_1 \\ & \lambda_2^{'} & & R_2 \\ & & \ddots & \vdots \\ R_1^H & R_2^H & \cdots & c_{NN} \end{pmatrix} \quad (9)$$

where $\lambda_i^{'}$ are the eigenvalues of C_1. A unitary transformation applied to the covariance matrix C does not change the eigenvalues, so the eigenvalues of C are equal to those of C_T. Another consequence is that the condition numbers are equal, that is to say we can equally choose the matrix C or C_T to be B_η. However, with the perturbation theory based on Gerschgorin's theorem, Wilkinson has shown that the off-diagonal elements of $O(\varepsilon)$ are reduced to order $O(\varepsilon^2)$ by diagonal similarity transformations [4]. If we consider $D_A = D_B$ and $\varepsilon = \eta$ but A_ε of order $O(\varepsilon^2)$ and now B_ε of order $O(\varepsilon)$, the equation (7) becomes :

$$(D_A + O(\varepsilon))^{-1}(D_A + O(\varepsilon^2))\, x = \lambda\, x \quad (10)$$

for ε sufficiently small. If there is no perturbation, the estimated generalized eigenvalues λ equals exactly one. With a high perturbation ($\varepsilon \rightarrow 1$), λ tend to 1 again. That is why we prefer to take C_T for A_ε and C for B_ε. The forms of C and C_T lead us to :

$$(D_C + O_C)^{-1}(D_{C_T} + O_{C_T})\, x = \lambda\, x \quad (11)$$

with $D_C = diag(C)$, $D_{C_T} = diag(C_T)$, O_C the matrix of off-diagonal elements of C of order $O(\varepsilon)$ and O_{C_T} the matrix of off-diagonal elements of C_T of order $O(\varepsilon^2)$. For the matrix C_T, the energy is concentrated in the main diagonal and is connected to the eigenvalues of C_1 while the diagonal elements of C tend to be equal. So, we can obtain two distinct signal and noise subspaces from the estimated generalized eigenvalues with any value of ε ($0 \leq \varepsilon \leq 1$). When ε is small, the smallest generalized eigenvalues seem to be less affected by perturbation than those in considering only C_T. We know it is possible to obtain two distinct sets of eigenvalues. Have we got the same possibility with the regions $\mathcal{G}_i$ by considering the matrices C and C_T ? The regions $\mathcal{G}_i$ are bounded by ρ_i from the equation (3). The bounds ρ_i are minimal when the numerator is small and the denominator high. It is true when off-diagonal elements are minimal, that is to say that the energy of the matrix C is concentrated on the main diagonal. The minimal perturbations in (3) tend to be around $1/\sqrt{2}$. For a commodity of visualization and to adopt a 2D representation like in [5], we make the transformation $\delta_i = 1/\rho_i$. Thus, with the couple (λ_i, δ_i), the values near the origin are supposed to be associated to the noise subspace again.

3.2 Detection criteria

To show the performances of the proposed method, we choose the criterion GDE_{dist} that is a heuristic criterion based on the normalized Euclidean distance called $dist(i)$ that takes into account the contributions of (λ_i, δ_i), let :

$$GDE_{dist}(k) = dist(k) - \frac{F(L)}{N}\sum_{i=1}^{P} dist(i) \qquad (12)$$

with $k = 1, \ldots, N$ and $F(L)$ is a constant value that can be connected to the sample number L. Usually, $F(L) = 1$ or $F(L)$ can vary in a range around 1 by short successive steps (for example $\Delta = 0.01$) and we retain the largest stages where the estimated source number M remains similar. The distance $dist(i)$ is described by :

$$dist(i) = \sqrt{(\lambda_i/\lambda_{max})^2 + (\delta_i/\delta_{max})^2} \qquad (13)$$

where λ_{max} and δ_{max} are respectively the maximal values of λ_i and δ_i. The criterion stops when a first negative value appears and the estimated source number becomes $M = k - 1$. We equally consider this criterion with only the values δ_i in order to avoid the calculation of the generalized eigenvalues. We apply the proposed method to the Marple signal [6]. This signal of 64 samples is composed of 4 complex sinusoids in a colored noise. Two sinusoids are very close to normalized frequencies 0.2 and 0.21 and they are more powerful by 20 dB than the two others placed at the normalized frequencies 0.1 and -0.15. With the Gerschgorin's radii R_i and centers O_i calculated from the matrix C_T and used in the criterion GDE_{dist}, we find two complex sinusoids. If we replace this information by (λ_i, δ_i) in the criterion (see figure 1), we obtain 4 sources by varying $F(L)$ in the range up to [0.6 1.4] around 1 (see figure 2). A larger range brings about 3 estimated sources.

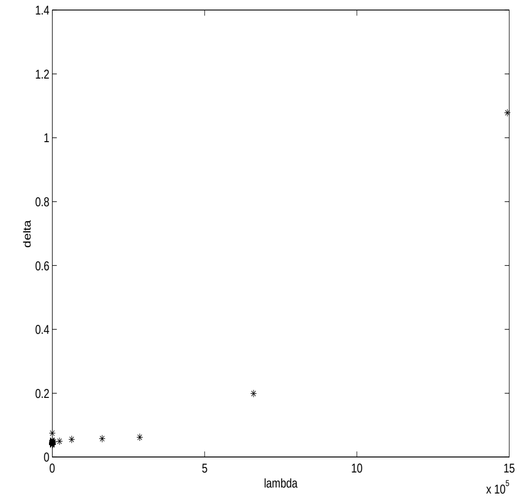

Figure 1: 2D representation of the Marple signal

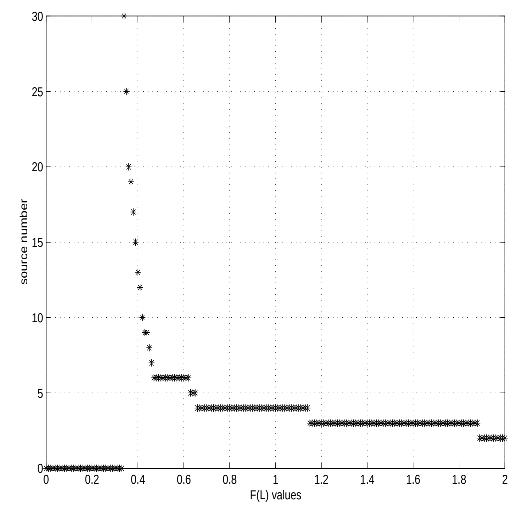

Figure 2: Variations of $F(L)$ ($\Delta = 0.01$), Marple signal

4 SIMULATION RESULTS

The signal under consideration is composed of 2 sinusoids embedded in a noise $n(t)$, let :

$$x(t) = A_1\ sin(2\pi f_1 t) + A_2\ sin(2\pi f_2 t) + n(t) \qquad (14)$$

where $t = 1, \ldots, 64$, the normalized frequencies $f_1 = 0.2, f_2 = 0.2 + 1/64$. The parameters A_1 and A_2 depend on simulations (same form of signal as [7]). The estimated matrix C is under the modified covariance form and of dimension $(32, 32)$. For each result, 200 simulations of Monte-Carlo are carried out. The GDE_{dist}

criterion is applied to C_T (see [5]) and to (C_T, C) from the equation (12) with $F(L) = 1$.

Case 1 (identical parameters A_i), the parameters $A_1 = A_2$ and vary from -5 db to +15 dB and $n(t)$ is a white Gaussian noise. The results obtained by the GDE_{dist} criterion, applied to the information (O_i, R_i) of C_T, slightly outperform those obtained by the information (λ_i, δ_i) of (C_T, C) of one dB but reach 100% detection rate with high SNR (Signal to Noise Ratio) (see figure 3). The simulations are carried out with the same parameters as previously with $n(t)$ a nonwhite Gaussian noise obtained through an AR(1) of coefficient 0.9. The results are slightly better with the $GDE_{dist}(\lambda_i, \delta_i)$ criterion but the best detection rate is not obtained (see figure 3).

Case 2 (different parameters A_i),we take $A_1 = 10$ dB and A_2 varies from -5 dB to +15 dB; $n(t)$ is a nonwhite Gaussian noise generated in the same conditions as previously. This case is difficult from the point of view of the frequency resolution, of the dynamic resolution and of the nature of the noise. The best results are obtained by the $GDE_{dist}(\lambda_i, \delta_i)$ criterion that can find 4 sources from 4 dB with a high detection rate compared to the $GDE_{dist}(O_i, R_i)$ (see figure 4). Those performances are almost similar to those of the figure 3 with a Gaussian white noise and show the possibility to discern the sources of different amplitude with a low SNR like in the case of Marple's signal.

5 CONCLUSION

The study of the generalized Gerschgorin's theorem and of the matrix perturbation theory can be a way of research to detect the source number. In this paper, we have presented one possibility based on the matrix pair (C_T, C) to improve the detection rate of sources in a critical situation but its cost of calculation is important. Other solutions can be found with a clever selection of estimated covariance matrices and criteria to exploit the information provided by these matrices, that is why other studies are under investigation.

References

[1] H. T. Wu, J. F. Yang and F. K. Chen, "Source number estimators using transformed Gerschgorin radii," IEEE Trans. on SP, vol. 43, n^{o}6, pp. 1325-1333, June 1995.

[2] G. W. Stewart, "Gerschgorin theory for the generalized eigenvalue problem $Ax = \lambda Bx$," Math. of Computation, vol. 29, n130, pp. 600-606, April 1975.

[3] G. W. Stewart and Ji-guang Sun, Matrix perturbation theory. Academic Press 1990.

[4] J. H. Wilkinson, The Algebraic Eigenvalue Problem. Oxford University Press, 1965.

[5] O. Caspary and T. Cecchin, "Detection of the source number by the Gerschgorin disks," Proc. of EUSIPCO, Tampere, Finland, 4-8 September, 2000.

[6] S. L. Marple, Digital spectral analysis with applications. Prentice Hall, 1987.

[7] P. Michel, J-Y. Tourneret and P. M. Djuric, "Online model selection of nonstationary time series using Gerschgorin disks," Proc. of ICASSP, Salt Lake City, 7-11 May, 2001.

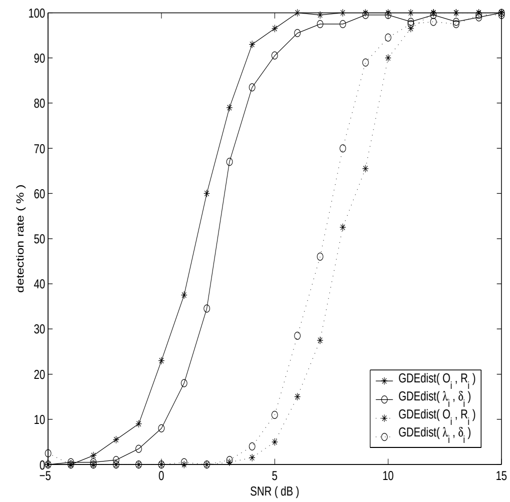

Figure 3: GDE_{dist} criterion, identical amplitudes (white noise : $-$, nonwhite noise : $\cdots$)

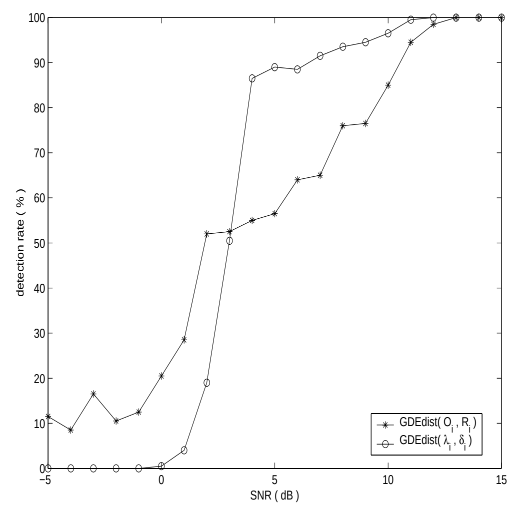

Figure 4: GDE_{dist} criterion, different amplitudes (nonwhite noise)