SIGNAL PROCESSING IV: Theories and Applications
J.L. Lacoume, A. Chehikian, N. Martin, and J. Malbos (Editors)
Elsevier Science Publishers B.V. (North-Holland)

RAYLEIGH ESTIMATES: PERFORMING LIKE SVD

M.A. Lagunas and M. Cabrera

Signal Processing in Communications
E.T.S.I. Telecomunicacion
Apdo. 30002
08080 BARCELONA S P A I N

Abstract

This paper describes the structure of the so-called Rayleigh estimates and the features they share with indirect SVD like procedures. The problem of finding procedures of high resolution in spectral estimation is faced under the framework of non-linear estimates of the autocorrelation matrix and the low rank approximation to the frequency estimation problem.

It is shown the existing relationship between the proposed estimates and the principal component analysis. The main advantages of the procedure is that the performance of the spectral estimates reported herein is almost equal to SVD techniques, yet preserving a good asymptotic convergence to the actual power spectral density. Also, the procedure could be viewed under variotional concepts revealing its potential under adaptive schemes and data adaptive windowing for spectral estimation. In summary, the work shows how classical constrained Wiener filtering with data adaptive windowing can enhance the performance of SVD methods with very low complexity.

1. INTRODUCTION

The basic principle in spectral estimation resides in the use of the autocorrelation function of the signal under analysis. The acf function has in its Fourier Transform its eigendecomposition. In other words, the steering vector $\underline{S}^H=(..,\exp(jnw),..)$ are the eigenvectors and the power spectral density Sx(w) are the corresponding eigenvalues.

$$r(n)=(1/2\pi)\int_{-\pi}^{\pi} Sx(w).e^{jnw}.dw \qquad (1)$$

Based on this principle, it could be stated that any candidate to be a spectral estimate, based in Q acf lags, have to converge to the actual spectral density and, as a consequence, to the eigenvalues of the acf matrix as the order Q tends to infinity. This provides an useful background to formulate non-linear functions of the data autocorrelation matrix, yet preserving the convergence mentioned before.

Thus, two approaches that can be encompassed in this direction should be mentioned; one is the well know principal component analysis, and, the second one the so-called power function estimates reported by Pisarenko /1/. Rayleigh estimates is a special case with particular features and sharing the advantages of both ideas.

Let us to review briefly the principal component analysis (PCA). The procedure consists in the selection of subspaces from the eigendecomposition, either of the data autocorrelation matrix R or over its inverse $\underline{\underline{R}}^{-1}$. Thus, depending on the procedure /2/, PCA versions of $\underline{\underline{R}}$ or $\underline{\underline{R}}^{-1}$ can be used in the structure of currently reported spectral estimates. As an example, a PCA version of R (the signal subspace) can be used in the Blackman-Tykey estimate,

$$\widetilde{S}x(w)= \underline{S}^H.\underline{\underline{R}}_{pca}.\underline{S} \qquad (2.a)$$

or a PCA version of $\underline{\underline{R}}^{-1}$, $\underline{\underline{R}}^{-1}_{pca}$ can be used in the linear prediction (2.b) or maximum likelihood estimates (2.c).

$$1//1^T.\underline{\underline{R}}^{-1}_{pca}.\underline{S}/^2 \qquad (2.b)$$

$$1/(\underline{S}^H.\underline{\underline{R}}^{-1}_{pca}.\underline{S}) \qquad (2.c)$$

This work is supported by CAYCIT grant number 21906/84 (Spain).

The main drawback of these procedures arise from the eigendecomposition itself /3/. Using the most familiar case which is the linear prediction approach and assuming (3) as the eigendecomposition of $\underline{\underline{R}}^{-1}$,

$$\underline{\underline{R}}^{-1}=\sum_{i=1}^{Q} \lambda_i \underline{U}_i \cdot \underline{U}_i^H \qquad (3)$$

with $\lambda_i > \lambda_{i+1}$ (i=1,Q-1), the corresponding PCA or low rank approximation will be (4).

$$\underline{\underline{R}}^{-1}_{pca}=\sum_{i=1}^{Q} \lambda_i \underline{U}_i \cdot \underline{U}_i^H \qquad (4)$$

The resulting estimate handles only the noise subspace of R and, because the orthogonality with the signal subspace, its location at the denominator of the spectral estimate formula produces the claimed high resolution of these procedures. Nevertheless, the resulting PCA estimate cannot longer be considered as a true spectral estimate due to the distortion involved in the rank reduction operation. This why better to refer this kind of procedures as frequency detectors than real power density estimates. As the reader can see when the order tends to infinity the estimate will tend to a cero eigenvalue instead of the actual value, breaking down the basic principle stated at the begining of this section. The point is the possibility of finding high resolution without loosing the desired natural convergence of the estimation process.

Even considering the interest of frequency detectors, the practical absence of low complexity approaches /4/ or adaptive schemes /5/, promotes the interest of non-exact solutions but suitable to beused under filtering or adaptive schemes constrained or unconstrained. These aspects will be covered by Rayleigh estimates in a very elegant manner providing, at the same time, a nice background for open procedures in this field.

2. NON-LINEAR FUNCTIONS OF THE AUTOCORRELATION MATRIX

The second approach to eigenvalue analysis of the power spectral estimation, mentioned in the previous section, is the use of non-linear functions of the autocorrelation matrix. To better understand the presentation we will focuss in the presentation the Capon maximum likelihood power level estimate $P_x(w)=(\underline{S}^H.\underline{\underline{R}}^{-1}.\underline{S})^{-1}$. Being H(.) a function of R and h(.) its inverse, a general family of spectral estimates can be done by (5); being H(.) a continuous and non-decreasing function for positive arguments.

$$\tilde{S}(w)=h(\underline{S}^H.H(\underline{\underline{R}}).\underline{S}) \qquad (5)$$

To prove the convergence of this estimate, the eigendecomposition of $\underline{\underline{R}}$ and the properties of function H(.) should be used in order to obtain a new formulation of the estimate.

$$\underline{\underline{R}}=\sum_{j=1}^{Q} \lambda_j \underline{U}_j \cdot \underline{U}_j^H \qquad (6)$$

$$H(\underline{\underline{R}})=\sum_{j=1,Q} H(\lambda_j) \cdot \underline{U}_j \cdot \underline{U}_j^H \qquad (7)$$

Thus, denoting U(w) as the Fourier transform magnitude of the eigenvector U_j (i.e. $\underline{U}(w)=/\underline{S}^H.\underline{U}_j/^2$), (8) results.

$$\tilde{S}(w)=h\left(\sum_{j=1}^{Q} H(\lambda_j).U_j(w)\right) \qquad (8)$$

As the order Q goes to infinity, the eigenvectors converge to the Q equally spaced steering vectors $\underline{S}$; under this circunstance, only one of the eigenvectors will produce a Fourier transform different from cero, and, $\tilde{S}(w)$ will converge to the corresponding eigenvalue.

$$\tilde{S}(w_i)_{Q\to\infty}=h(H(\lambda_i).U_i(w_i))=h(H(\lambda_i))=\lambda_i \qquad (9)$$

Facing now the problem of selecting a function H(.), it is clear that a PCA analysis will require a non-linear function as it is shown in (10).

$$H_{op}(x)=\begin{cases} x & \text{for } x \geq x_o \\ 0 & \text{for } 0 \leq x < x_o \end{cases} \qquad (10)$$

A more suitable choice for H(.) is a polynomial of x or, furthermore, a given power of x. This last choice arises to the so-called power function estimates.

$$\tilde{S}^q(w)=(\underline{S}^H.\underline{\underline{R}}^{-q}.\underline{S})^{-(1/q)} \qquad (11)$$

The relationship of the power q, in (11), with the PCA or low rank approximation can be viewed in Fig 1. Note also that q=-1 and q=1 represent the Blackman-Tukey and Capon MLM estimates respectively.

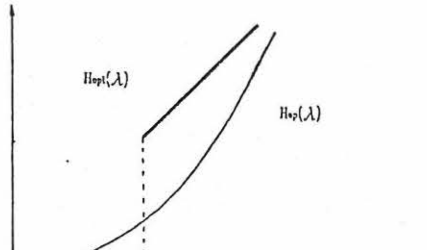

Fig 1. H(.) for the optimum low rank reduction problem $H_{op}(.)$ and the approximation of the power function estimates $H_{ap}(.)$.

Further details about (11) can be found in /1/, but the fundamental limitations concerning filtering framework and adaptive schemes supporting (11) remain unsolved. In other words, power function estimates could be viewed as an approximate solution to the PCA problem without eigendecomposition; but, there is not a good framework to support them and no adative schemes could be envisaged from their formulation.

3. RAYLEIGH ESTIMATES

Starting from the bank filtering approach /2/, it should be pointed out that the data snapshot $\underline{X}_n$ in a spectral estimate procedure is handled in two ways: one is to design the dedicated filter A steered at a given frequency (i.e. $\underline{S}^H.\underline{A}=1$); and second one to produce the output residual which power provides the power level estimate $P_x(w)$. After, the bandwidth normalization allow to obtain the power density estimate $S_x(w)=P_x(w)/(\underline{A}^H.\underline{A})$.

In order to obtain the family of Rayleigh estimates it is necessary to introduce a transformation $\underline{\underline{T}}$, over the data snapshot $\underline{X}_n$, to modify the design of the dedicated filter $\underline{A}$. The new data will be:

$$\underline{\tilde{X}}_n=\underline{\underline{T}}.\underline{X}_n \qquad (12)$$

and the design equations for $\underline{A}$ will be as it is shown in (13).

$$\underline{A}^H.E(\underline{\tilde{X}}_n.\underline{\tilde{X}}_n^H).\underline{A} \;/_{min} \qquad \underline{A}^H.\underline{S}=1 \qquad (13)$$

It is worthwhile to note that the transformation matrix includes non-data adaptive window methods just reducing $\underline{T}$ to be a diagonal matrix, with elements equal to the window weigths along its main diagonal.

At this point, if an enphasis should be performed in the signal eigenvalues an adequate choice for the transformation matrix, among many others, is the data autocorrelation matrix $\underline{\underline{R}}$ or a power of it.

$$\underline{\underline{T}}=\underline{\underline{R}}^s \;;\quad \underline{\tilde{X}}_n=\underline{\underline{R}}^s.\underline{X}_n \qquad (14)$$

Again, it is important that a review of the window theory, under this formulation, opens new oportunities because 2-D windows for 1-d problems can be implemented using the above formulation.

Going back to the filter design, solving (13) for vector $\underline{A}$ and using (14), the desired solution is (15).

$$\underline{A}=\underline{\underline{Q}}^{-1}.\underline{S}.(\underline{S}^H.\underline{\underline{Q}}^{-1}.\underline{S}) \qquad (15.a)$$

$$\underline{\underline{Q}}=\underline{\underline{T}}.\underline{\underline{R}}.\underline{\underline{T}}=\underline{\underline{R}}^{2s+1} \qquad (15.b)$$

In order to obtain the power level estimate it is necessary to compute the output filter residual $\underline{A}^H.\underline{\underline{R}}.\underline{A}$. It is remarkable the difference between the output residual, which is obtained from the original data, and the objective, in the design equations (13), which is performed over the transformed data. Once the power level estimate is obtained, the white noise bandwidth normalization provides the final estimate shown in (16).

$$\tilde{S}(w)=\frac{\underline{A}^H.\underline{\underline{R}}.\underline{A}}{\underline{A}^H.\underline{A}}=\frac{\underline{S}^H.\underline{\underline{R}}^{-4s-1}.\underline{S}}{\underline{S}^H.\underline{\underline{R}}^{-4s-2}.\underline{S}} \qquad (16)$$

This will be refered hereafter as Rayleigh estimates. Next section will be devoted to the special properties associated with Rayleigh estimates.

4. RAYLEIGH ESTIMATES AND THE LOW RANK REDUCTION PROBLEM

From the formulation of Rayleigh estimates, results clear that the denominator can be viewed as the low rank approximation mentioned in section 2. In fact, this expression without the quadratic form of the numerator, can be used succesfully as a frequency detector. Of course, the exponetial enphasis of the eigenvalues done in the denominator alone produces distortion, this is why is better to consider it as a frequency estimate than a spectral density estimate. It is the numerator which compensates the denominator weigthing and removes, in this way, the distortion introduced in the eigenspace. To be more especific, (17.a) has the convergence desired for a spectral estimate, (17.b) is the propossed frequency detector, and (17.c) recovers the convergence from the frequency detector by forming the Rayleigh quotient.

$$(\underline{S}^H.\underline{\underline{R}}^{-4s-2}.\underline{S})^{-4s-2} \qquad (17.a)$$

$$(\underline{S}^H.\underline{\underline{R}}^{-4s-2}.\underline{S})^{-1} \qquad (17.b)$$

$$(\underline{S}^H.\underline{\underline{R}}^{-4s-1}.\underline{S}).(\underline{S}^H.\underline{\underline{R}}^{-4s-2})^{-1} \qquad (17.c)$$

The main difference between (17.a) and (17.c) resides in the framework which support the last one. A constrained design filter supporting (17.c) allows its adaptive implementation, it provides an easy extension to 2-D problems, angle of arrival estimation, open new fields and perspective to the topic of data adaptive windows and produces a SVD-like procedure with low computational burden. This last sentence summarizes the main properties associated with these estimates. Rayleigh estimates enhance the performance of frequency detectors and principal component analysis without eigendecomposition.

About the choice of the parameter s, it can be said that this parameter controls the convergence to the actual power density distribution when the order of the data autocorrelation matrix tends to infinity. But, more close to the practical use of the estimate, it is important to point out that s controls the peaky character of the resulting estimate /6/. In fact, it is worthwhile to remind that s, in some extent, is the degree of rank reduction carried out over the original data acf. Also, tha case of s equal one could be, for some applications, very high. Using the final formula (16), non-integer values can be set for parameter s in such a way that q=4s+1 must be an integer one.

When using the procedure with constrained adaptive algorithms, /7/,/8/, special attention should be payed to the existing trade-off between resolution and convergence rate. The associated high resolution to Rayleigh estimates promotes a lower convergence rate when compared with classical algorithms, like maximum likelihood or linear prediction. This slowness phenomena in convergence rate could be justifyed from the increase in the eigenvalue spread of the involved matrix Q refered to the eigenvalues of $\underline{\underline{R}}$.

Eigenvalue spread

$$\text{original}=\frac{\lambda_{max}}{\lambda_{min}} \qquad \text{Rayleigh}=\frac{\lambda_{max}}{\lambda_{min}}^{2s+1}$$

Finally, and as an example, in Fig 2 it can be viewed some examples for sucessive values of q (q=4s+1) in a 2-D problems of sinusoids in noise. From Fig 2 it is easy to conclude the similar performance of Rayleigh estimates and principal component analysis.

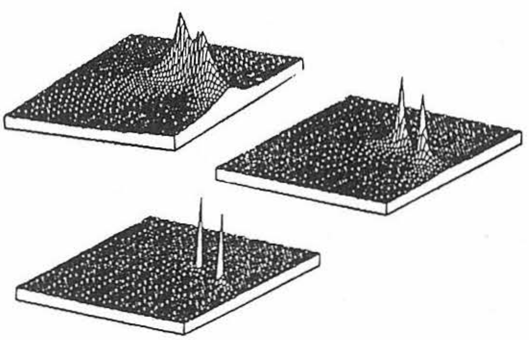

Fig 2.- Performance of Rayleigh estimates for the problem of sinusoids in noise for sucessive values of parameter q.

6. REFERENCES

/1/ V.P. Pisarenko. " On the Estimatetion of Spectra by Means of non-linear Functions of the Covariance Matrix". Geophys. J.R. Astro. Soc. (1972) 28, pp. 511-531.

/2/ M.A.Lagunas et al. "Maximum Likelihood Filters in Spectral Estimation Problems". Signal Processing 10, North-Holland, 1986, pp. 19-34

/3/ G.H. Golub and C.F. Van Loan, "Matrix Computation". Baltimore, MD: Univ. Press. 1983.

/4/ S.M. Kay and A.K. Shaw. "Frequency Estimation by Principal Component AR Spectral Estimation Method Without Eigendecomposition". IEEE Trans. on Acoustics Speech and Signal Processing, ASSP-36, no. 1, Jan. 1988, pp. 95-101.

/5/ J.F. Yang and M. Kaveh. "Adaptive Eigensubspace Algorithms for Direction or Frequency Estimation and Tracking. IEEE Trans. on Acoustics Speech and Signal Processing, ASSP-36, no. 2, Feb. 1988, pp. 241-251.

/6/ M.A. Lagunas and M. Cabrera. "Rayleigh Spectral Estimation". Submitted for publication.

/7/ O.L.Frost III. "An Algorithm for Linearly Constrained Adaptive Array Processing". Proc. IEEE, Vol. 55, no. 8, Aug. 1972, pp. 926-935.

/8/ L.J.Griffiths and K.M.Buckley. "Quiescent Pattern Control in Linearly Constrained Adaptive Arrays". IEEE Trans. on Acoustics Speech and Signal processing, ASSP-35, no. 7, Jul. 1987, pp. 917-926.