# 특이값 접근방법에 의한 정현파의 수의 결정에 관한 연구

안 태천°, 류 창선 :원광대 전기공학과
이 동윤, 황 금찬 :연세대 전기공학과

# Determination of the number of sinusoidal frequencies by a new singular value approach

Tae-chon Ahn°, Chang-sun Ryu :Dept. of Electrical Eng. Wonkwang Univ.
Dong-yoon Lee, Keun-chan Whang : Dept. of Electrical Eng. Yonsei Univ.

Abstract

A new singular value approach is presented and analized in order to determine the number of multiple sinsoidal frequencies from the finite noisy data. Simulations are conducted for Akaike's information criterion(AIC), Rissanen's shortest data description(MDL) and a new singular value approach, in covariance matrix based methods. And then performances are compared.

## Ⅰ. Introduction

Estimation of multiple sinusoidal frequencies from finite noisy data is a very interisting and partical problem. It has been studied and used in many fields. With the rapid development of modern technology, the need for estimation of frequencies becomes ever increasing and therefore motivatives more and more researchers on this issue.

It is important to determine the number of sinusoids before estimating the frequencies. Generally it should be assumed that the number of sinusoids in the signal was known. In practice, however, this information is often unknown, it must be determined before estimating the frequencies. This is a difficult problem, particularly for short data.

A popular approach for selection of model order is, of course, the information theoretic criteria, introduced by Akaike[1][2] and Rassanen[3]. Wax and Kalaith[5] have extended them to the sinusoidal case. Chatterjee et al[6], have recently proposed a test rule for a generalized autoregressive (GAR) model, based on Bayesian approach, while Satorius and Alexander[7] used the determinant test on the covariance matrix. Another idea is to examine the singular values of certain matrices by distinguishing the signal singular values from noise singular values[4][8].

In the paper, a new singular value approach applied to the sinusoidal case is presented and analized in detail[9]. Using the approach, the number of multiple sinusoidal frequencies is to be determined from finite noisy data. Simulations are conducted for Akaike's information criterion(AIC), Rissanen's shortest data description(MDL) and a new singular value approach, in various examples. And then performances are compared.

## Ⅱ. A new singular value approach

First consider a covariance matrix determination of the number of sinusoids from N noisy data. Then Akaike's information criterion (AIC)[1][2] and Rissanen's shortest data description(MDL)[3][5] using the covariance matrix is discussed.

AIC and MDL were given by the follows.

$$AIC(n) = -2\log\left(\frac{\prod_{i=n+1}^{p} \sigma_i^{1/(p-n)}}{\frac{1}{p-n}\sum_{i=n+1}^{p} \sigma_i}\right)^{(p-n)N} + 2n(2p-n) \quad (1)$$

$$MDC(n) = -\log\left(\frac{\prod_{i=n+1}^{p} \sigma_i^{1/(p-n)}}{\frac{1}{p-n}\sum_{i=n+1}^{p} \sigma_i}\right)^{(p-n)N} + \tfrac{1}{2}n(2p-n)\log N \quad (2)$$

Where $\sigma$ is the ith singular values of matrix and n is the possible number of complex sinusoids, or twice the number of real sinusoids.

The correct number of n should be chosen as the one which makes AIC(n) or MDL(n) each its minimum. It is interisting to note that the term in the parenthesis of (1) and (2) is the ratio of giometric mean of noise singular values to the arithmatic mean.

Next a new singular value approach is presented analytically. It is known in the noiseless case that the data or covariance matrices[10] are of rank 2m, again m being the number of sinusoids. In such case, these matrices have 2m non-zero(greater than zero) signal singular values. In nosiy data case, all the singular values will be non-zero, but the signal singular values are usually much greater than the noise singular values. This information is much valuable when determining the number of sinusoids. Decisions can be made, for example, by inspecting the "gaps" between two succesive singular values. Here we will examine the singular values by observing how "close"a low rank matrix is to the originally full rank matrix. To do this, we will intro- duce the concept of low "effective rank" [4],[8].

Let A be a data or covariance matrix of dimension MxL, which contains the information about the sinusoids-in-noise process. Perform a SVD for the matrix A

$$A = U\Sigma V^T \quad (3)$$

where

$$U = (u_1, u_2, \cdots, u_M) \quad (4)$$

$$V = (v_1, v_2, \cdots, v_L) \quad (5)$$

$$\Sigma = diag\{\sigma_1, \sigma_2, \cdots, \sigma_q\},\ q = min(M,L) \quad (6)$$

Use the normalized square Frequenious norm sense as

$$Per(n) = \frac{\|A^{(n)}\|_F^2}{\|A\|_F^2} = \frac{\sum_{i=1}^{n}\sigma_i^2}{\sum_{i=1}^{q}\sigma_i^2} \quad (7)$$

where

$$A^{(n)} = \sum_{i=1}^{n}\sigma_i u_i v_i^T \quad (8)$$

Obviously Per(n) is a nondecreasing function of n. Since matrix A in our case is generally of full rank, Per(n) will monotonically increasing to one as n approaches min(M,L). Because the signal singular values of A are much larger than the noise singular values, Per(n) is close to one for some small number.

The low effective rank of A is then defined as the number which is much smaller than min (M,L) and which makes Per(n) very close to one. Any larger number than this low effective rank makes no significant contribution to Per(n).

The phase "very close to one"might be vague. Therefore the decision made may depend on user's opinion. However, expreience shows that it is useful. In the following, two examples are given to demonstrate how the low effective rank should be chosen and how the design parameters affect Per(n). Clearly we will take this low effective rank as the "signal rank", or twice the number of sinusoids. Hence this low effective rank must be taken as an even integer.

## Ⅲ. Simulation examples

In this section, simulation examples are given for determination of the number of sinusoids of a sinusoids-in-noise process, using covariance. In all the examples, the signal was assumed to consist of two sinusoids. The SVD has been performed for different matrices and different design parameters.

Example 1. The data simulated is given by

$$y(t) = \sqrt{2}\sin(0.7226t) + \sqrt{2}\sin(1.036t) + e(t) \quad (9)$$

where e(t) is a zero mean white Gaussian process, the data Length is N=64, and the SNR is 10 dB, i.e. $SNR_i = 10\log(\alpha_i^2/2\sigma^2)$, i=1,2. AIC(n), MDL(n) and Per(n) are calculated with a number of choice of n. Results are given in table 1 through 3.

| p \ n | 2 | 4 | 6 | 8 |
|---|---|---|---|---|
| 12 | 1972.8 | 383.4 | 300.0 | 285.5 |
| 16 | 3899.5 | 911.1 | 705.7 | 677.2 |
| 20 | 4945.4 | 810.0 | 705.4 | 670.5 |
| 24 | 6096.0 | 1346.3 | 1301.8 | 1277.6 |

Table 1 performances of AIC(n), Example 1

| p \ n | 2 | 4 | 6 | 8 |
|---|---|---|---|---|
| 12 | 1033.9 | 278.1 | 266.6 | 280.9 |
| 16 | 2014.5 | 576.5 | 521.2 | 545.9 |
| 20 | 2554.7 | 560.4 | 572.9 | 611.6 |
| 24 | 3147.3 | 863.1 | 922.9 | 984.2 |

Table 2 Perfomance of MDL(n), Example 1

| L \ n | 2 | 4 | 6 | 8 |
|---|---|---|---|---|
| 12 | 0.8805 | 0.9994 | 0.9998 | 0.9999 |
| 16 | 0.6820 | 0.9992 | 0.9998 | 0.9999 |
| 20 | 0.5688 | 0.9993 | 0.9997 | 0.9998 |
| 24 | 0.6789 | 0.9993 | 0.9996 | 0.9998 |

Table 3. Per(n) for the covariance matrix, Example 1

Example 2.

The same signal as the example 1 has been used but the SNR has been changed to 0 dB. Computations are repeated as done in example 1, and performance are listed in Table 4 through 6.

| P \ n | 2 | 4 | 6 | 8 |
|---|---|---|---|---|
| 12 | 498.4 | 240.7 | 253.8 | 280.2 |
| 16 | 1144.3 | 450.7 | 484.9 | 479.3 |
| 20 | 1938.1 | 758.7 | 811.5 | 821.5 |
| 24 | 2790.1 | 1445.4 | 1495.6 | 1533.7 |

Table 4. Performances of AIC(n), Example 2

| P \ n | 2 | 4 | 6 | 8 |
|---|---|---|---|---|
| 12 | 296.7 | 206.7 | 243.5 | 278.3 |
| 16 | 636.9 | 346.2 | 410.8 | 446.9 |
| 20 | 1051.1 | 534.8 | 626.0 | 687.1 |
| 24 | 1494.4 | 912.7 | 1019.8 | 1112.3 |

Table 5. Performances of MDL(n), Example 2

| L \ n | 2 | 4 | 6 | 8 |
|---|---|---|---|---|
| 12 | 0.8434 | 0.9751 | 0.9881 | 0.9941 |
| 16 | 0.6680 | 0.9719 | 0.9824 | 0.9914 |
| 20 | 0.5522 | 0.9741 | 0.9819 | 0.9891 |
| 24 | 0.6444 | 0.9761 | 0.9833 | 0.9887 |

Table 6. Per(n) for covariance matrix, Example 2

The following comments are drawn from the above two examples:

( i ) The AIC and MDL tend to overestimate the number of sinusoids at relatively high SNR. However, MDL may provide correct results for large design paramdters. For low SNR, both AIC and MDL give consistent results. This may be argued from the following observa tions. At high SNR,the noise singular values diverse very much ( the ratio of the largest noise singular value to the smallest one is small ). The ratio of the geometric mean to the arithmatic mean is small ( note it is less than one ), and the first term in (1) and (2) will be very large. Therefore the second term, i.e. the penalty term due to parsimony principle, does not play important role in correction of overestimation. For low SNR, the noise singular values are closer to each other. The first term in (1) and (2) decreases,and the second term can give proper penalty for overestimating the number of signal ( note also the second term of (1) and (2) is constant with respect to SNR).

( ii ) The test Per(n) for covariance matrix shows that the low effective rank is 4, in both example, which means that there are two sinusoids in the measurement data. But for small design parameters,the low effective rank may underestimate the number of sinusoids. This can be explained by observations on the singular values. For small design parameters and two equal SNR's (SNR1=SNR2), the first two signal singular values are much greater than the other two. These small two signal singular values grow very fast (faster than the first two) with the increase of design parameters. Thus the first two signal singular values become less dominant. On the other hand, if the design parameters are too large, the signal singular values may become less dominant since there are too many noise singular values in this case. Finally the value of the SNR also affects the dominancy of the signal singular values.

## Ⅳ. Conclusions.

In the paper, a new singular value approach that determined the number of multiple sinusoidal frequencies from the finite noisy data was presented and analized.

Simulations were conducted for AIC, MDL and a new singular value approach in various methods.

The following conclusions are drawn:

( i ) The AIC and MDL tend to overestimate the number of sinusoids at relatively high SNR. However, MDL may provide correct results for large design parameters. For low SNR, both AIC and MDL give consistent results.

( ii ) The test of Per(n) for covariance matrix shows that the low effective rank is 4,in both examples, which means that there are two sinusoids in the measurement data. But for small design parameters, the low effective rank may underestimate the number of sinusoids.

## REFRERENCE


[1] H.Akaike "Information theory and an extention of the maximun likelihood principle", Proco and Int.Symp., Inform. Theory Suppl Problem of Control and Inform. Theory,PP.267-281,1973

[2] H.Akaike, "A new look at the satistical model identification," IEEE Trams. Automatic Control,Vol. AC-19,PP.713-723,1974

[3] J.Rissanen,"Modeling by shortest da-ta descsieption," Automatica, Vol.14,PP. 465-471,1978

[4] T.P.Bronez and J.A Cadzow, "Sigmal rank and model orderdetermination, "Proc.ICASSP,PP.14.8.1-14.8.4,1984

[5] M.Wax and T.Kailath, "Detection of signal by information theoretical criteria, "IEEE Trans. Acoustics, Speech and Siganl Processing,Vol.ASSP-33,PP.387-392, 1985

[6] C. Chatterjee, R.L. Kashyap and G. Boray, "Estimation of close sinusoids in coloured noise and model discrimination," IEEE Trans. Acoustics, Speech and Signal Processing,Vol.ASSP-35,PP.328-337,1987

[7] E.H.Satorius and S.T. Alexander,"High resolution analysis of simusoids in colour noise,"Rec.ICASSP,Tulsa,OK. PP.349-351,1978.

[8] J.A. Cadzow,"Spectral estimation:an overdetermined rational model equation approach, "Proc.IEEE,Vol.70,PP907-937 1982.

[9] T.C.Ahn,"Determination of the number of sinusoids by a new singular value approach, " To be appeared,1990

[10] T.C. Ahn, "Estimation of frequencies of multiple sinusoids from a data based covariance Matrix", Journal of KIEE, Vol.2, no.1,1989.