MSE-BASED REGULARIZATION APPROACH TO RANK DETERMINATION IN CLS AND TLS ESTIMATION

H. Kagiwada, Y. Aoki, J. Xin and A. Sano

Department of Electrical Engineering, Keio University
3-14-1 Hiyoshi, Kohoku-ku, Yokohama 223, Japan
Tel : +81 45 563 1141; fax : +81 45 563 2773
e-mail : sano@sano.elec.keio.ac.jp

ABSTRACT

The corrected least squares (CLS) approach using an overdetermined model is investigated to decide the number of sinusoids in additive white noise. Like the total least squares (TLS) approach, the CLS estimation is different from the ordinary least squares (LS) method in that the noise variance is subtracted from the diagonal elements of the correlation matrix of the noisy observed data. Therefore the inversion of the resultant matrix becomes ill-conditioned and then adequate truncation of the eigenvalue decomposition (EVD) should be done. This paper clarifies how to simultaneously estimate the noise variance and truncate the eigenvalues, since they are mutually dependent. By introducing a multiple number of regularization parameters and determining them to minimize the MSE of the model parameters, we can give an optimal scheme for the truncation of eigenvalues. Furthermore, an iterative algorithm using only observed data is also clarified.

1. INTRODUCTION

The eigenstructure approach is a powerful tool in high-resolution spectral analysis in low SNR cases. Application of the singular value decomposition (SVD) on the noisy observed data matrix is a low rank approximation for it in the sense of the LS estimate with the rank equal to the number of sinusoids [1]. The TLS approach is one of refinements of the LS method when there are observation noises in both the data matrix and observed signal vector [2][3]. However, when an overdetermined model is used, the number of sinusoids should be known *a priori* to obtain an estimate of the spectrum or frequencies of signals.

In this paper, we take the CLS and TLS approaches using an overdetermined model [4], which are different from the ordinary LS approach in that the noise variance is subtracted from the diagonal elements of the correlation matrix of the noisy observed data. In the TLS estimation, the noise variance is specified by the minimum eigenvalue of the matrix constructed by appending a noisy observed data vector to the noise-corrupted data matrix. The CLS estimate is given by using the inverse of the signal correlation matrix with the diagonal elements subtracted by the noise variance. However, these resultant matrices have a reduced rank and then adequate truncation for the EVD should be taken. Thus, in the overdetermined CLS approach, the estimation of the noise variance and the truncation of eigenvalues are mutually dependent [5][6], therefore it is significant to decide them simultaneously. We will clarify that the detection of the number of sinusoids is closely related to the optimal truncation for the EVD, which can be performed by introducing multiple regularization parameters. Further, an iterative data-adaptive algorithm using only received data is given to estimate the noise variance and the regularization parameters for the truncation simultaneously to determine the number of sinusoids.

2. CORRECTED AND TOTAL LS METHOD

2.1 Description of Signal Model

Let the observed signal consist of complex sinusoids in white Gaussian noise as

$$y(t) = x(t) + e(t), \qquad \text{for} \quad t = 1, 2, \cdots, n+m \tag{1}$$

$$x(t) = \sum_{j=1}^{M} c_j \exp\left(j2\pi f_j t \right) \tag{2}$$

where $\{x(t)\}$ is a signal part consisting of multiple sinusoids, $\{e(t)\}$ is a zero-mean white Gaussian noise with variance σ_e^2. f_j and c_j are the frequency and complex amplitude of the j-th sinusoid. The problem is to detect the number of sinusoids M from the observed $n+m$ data samples $\{y(t)\}$.

We will treat with the sinusoidal signal $\{x(t)\}$ as an overdetermined model with the order m $(m \geq M)$ described by

$$x(t) = \varphi_x^T(t) a \tag{3}$$

where $\varphi_x(t) = (x(t-1), x(t-2), \cdots, x(t-m))^T$ and $a = (a_1, a_2, \cdots, a_m)^T$. Then the observed signal $\{y(t)\}$ is then given by

$$y(t) = \varphi_y^T(t) a + \overline{w}(t) \tag{4}$$

where $\varphi_y(t) = (y(t-1), y(t-2), \cdots, y(t-m))^T$, $\overline{w}(t) = e(t) - \varepsilon^T(t) a = \overline{e}^T(t)\overline{a}$, $\varepsilon(t) = (e(t-1), e(t-2), \cdots, e(t))^T$, $\overline{e}(t) = (e(t), \varepsilon^T(t))^T$ and $\overline{a} = (1, -a^T)^T$.

Thus the observed signals from $t = 1$ to $t = n$ are expressed in a compact form as

$$y = \Phi_y a + \overline{w} \tag{5}$$

where $y = (y_1, y_2, \cdots, y_n)^T$, $\Phi_y = [\varphi_y(1), \varphi_y(2), \cdots, \varphi_y(n)]^T$, $\overline{w} = (\overline{w}(1), \overline{w}(2), \cdots, \overline{w}(n))^T = \overline{E}\,\overline{a}$, $\overline{E} = [\overline{e}(1), \overline{e}(2), \cdots, \overline{e}(n)]^T$.

The data matrix Φ_y can also be written from (1) by

$$\Phi_y = \Phi_x + E \tag{6}$$

where $\Phi_x = [\varphi_x(1), \varphi_x(2), \cdots, \varphi_x(n)]^T$, $E = [e(1), e(2), \cdots, e(n)]^T$.

Thus it is noticed that the data matrix Φ_y and data vector y are contaminated by the additive noises as described by (5) and (6), where Φ_x and E are the corresponding true signal matrix and noise matrix in which the each component is given by (1).

2.2 Rank Deficiency Problem in CLS and TLS Estimation

In this paper, we take a CLS estimate which is given by

$$\hat{a}_{\mathrm{CLS}} = \left[\frac{1}{n}\Phi_y^H \Phi_y - \sigma_e^2 I_m\right]^{-1} g, \qquad g \equiv \frac{1}{n}\Phi_y^H y \tag{7}$$

It can be noticed from (7) that the CLS estimate requires the true noise variance σ_e^2 and that the $m \times m$ matrix in the bracket in (7) approaches a true signal correlation matrix with the rank M $(M < m)$, which becomes singular for a sufficiently large n. Therefore in detection of the number of sinusoids based on the CLS approach using the overdetermined model, we have to consider the problem how to simultaneously estimate both the noise variance σ_e^2 and the rank of the true signal correlation matrix. The TLS estimate is given by replacing σ_e^2 with the minimum eigenvalue $\lambda_{\min}[(\Phi_y, y)^H(\Phi_y, y)]$ and it is equivalent to the CLS estimate for $n \to \infty$.

3. OPTIMAL RANK DETERMINATION PROBLEM

3.1 MSE of Regularized CLS Estimate

By using the eigenvalue decomposition the CLS estimate (7) can be rewritten by

$$\hat{a}_{\mathrm{CLS}} = \sum_{i=1}^{m} \frac{1}{\lambda_i - \sigma_e^2} v_i v_i^H g \tag{8}$$

where $\{v_i\}$ and $\{\lambda_i\}$ are the eigenvectors and eigenvalues defined as follows:

$$\frac{1}{n}\Phi_y^H \Phi_y = V \Lambda V^H \tag{9}$$

where $V = [v_1, v_2, \cdots, v_m]$, $V^H V = V V^H = I$, $\Lambda = \mathrm{diag}[\lambda_1, \lambda_2, \cdots, \lambda_m]$, and $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_m$. Since the number of signals is M, it holds that $\lambda_{M+1} \cong \lambda_{M+2} \cong \cdots \cong \lambda_m \cong \sigma_e^2$. Therefore the denominators in (8) for $i \ge M+1$ become nearly zero and it causes the CLS estimate to be numerically unstable.

To overcome this problem, we introduce a regularization matrix R to obtaine the regularized CLS estimate as

$$\hat{a}_{RCLS}(R) = \left[\frac{1}{n}\Phi_y^H \Phi_y - \sigma_e^2 I_m + R\right]^{-1} g \tag{10}$$

The MSE of the estimate (10) can be evaluated by taking account of up to the fourth moment. The result is given as:

Theorem 1 [4]: The MSE of the estimate (10) is given for sufficiently large n as

$$\begin{aligned} MSE(R) &\equiv E\left(\left\| a - \hat{a}_{RCLS}(R)\right\|^2\right) \\ &= \frac{1}{n}(1 + \|a\|^2) tr[(\Sigma_x + \sigma_e^2 I_m)(\Sigma_x + R)^{-2}] \\ &+ \frac{1}{n}\sigma_e^4 a^H (\Sigma_x + R)^{-2} a + a^H R (\Sigma_x + R)^{-2} R a \end{aligned} \tag{11}$$

Now by specifying the regularization matrix by use of the eigenvectors as

$$R = \sum_{i=1}^{m} \gamma_i v_i v_i^H \tag{12}$$

we can rewrite the MSE in (11) as

$$\begin{aligned} \mathrm{MSE}(\{\gamma_i\}) &= E\left(\left\| a - \hat{a}(\{\gamma_i\})\right\|^2\right) \\ &= \frac{1}{n}\sigma_e^2(1 + \|a\|^2)\sum_{i=1}^{m} \frac{\lambda_i}{(\lambda_i - \sigma_e^2 + \gamma_i)^2} \\ &+ \frac{1}{n}\sigma_e^4 \sum_{i=1}^{m} \frac{\left|v_i^H a\right|^2}{(\lambda_i - \sigma_e^2 + \gamma_i)^2} + \sum_{i=1}^{m} \frac{\gamma_i^2 \left|v_i^H a\right|^2}{(\lambda_i - \sigma_e^2 + \gamma_i)^2} \end{aligned} \tag{13}$$

where $\hat{a}(\{\gamma_i\})$ is the regularizaed estimate expressed in terms of the regularization parameters $\{\gamma_i\}$, which is given by

$$\hat{a}(\{\gamma_i\}) = \sum_{i=1}^{m} \frac{1}{\lambda_i - \sigma_e^2 + \gamma_i} v_i v_i^H g \tag{14}$$

3.2 Minimum MSE-Based Truncation of Eigenvalues

It can be noticed from (8) and (14) that the whether the eigenvalue λ_i should be retained or discarded corresponds to the choice of $\gamma_i = 0$ or $\gamma_i = \infty$. Therefore the problem is formulated as follows: Determine $\{\gamma_i\}$ so as to minimize the MSE (13) under the constraint that $\{\gamma_i\}$ has only two values 0 and ∞. The solution is given in the next theorem.

Theorem 2: The optimal decision rule for retaining or discarding the eigenvalue λ_i so that the MSE in (13) can be minimized is given by

$$\gamma_i^{(T)} = \begin{cases} 0: & \text{if } \lambda_i \ge \gamma_i^* + \sigma_e^2 \\ \infty: & \text{if } \lambda_i < \gamma_i^* + \sigma_e^2 \end{cases} \tag{15}$$

where $\{\gamma_i^*\}$ minimize the MSE in (13) without the constraint and are given by

$$\gamma_i^* = \frac{\sigma_e^2 \lambda_i (1 + \|a\|^2)}{n(\lambda_i - \sigma_e^2)\left|v_i^H a\right|^2} + \frac{\sigma_e^4}{n(\lambda_i - \sigma_e^2)} \tag{16}$$

Remark 1: Theorem 2 gives the threshold value for deciding whether the eigenvalue λ_i should be retained or discarded. Therefore, if the regularization parameters in (15) satisfy that

$$\gamma_i^{(T)} = \begin{cases} 0: & \text{for } i = 1, \cdots, M \\ \infty: & \text{for } i = M+1, \cdots, m \end{cases} \tag{17}$$

then we can conclude that the number of sinusoid is M by finding the intersection of the plots of $\lambda_i - \sigma_e^2$ and γ_i^* as illustrated in Fig.1. However, the above truncation scheme requires the true values of a and σ_e^2, so they should be replaced with their estimates using the accessible finite data samples.

Remark 2: The true parameter vector a of the overdetermined model is included in (13). It is the parameter of the overdetermined model (3) obtained in a noise-free case, which can be represented by

$$a = \sum_{i=1}^{M} \frac{1}{\lambda_i^*} v_i^* v_i^{*H} g_x^* \tag{18}$$

where λ_i^* and v_i^* are the eigenvalues and eigenvectors of the signal correlation matrix $\Sigma_x = \lim_{n\to\infty}(\Phi_x^H \Phi_x / n)$ and $g_x^* = \lim_{n\to\infty}(\Phi_x^H x / n)$, where $x = (x(m+1), \cdots, x(M+n))^T$.

3.3 Estimation of Noise Variance

The problem considered in this section is how to estimate the noise variance σ_e^2. We employ an alternative regularized LS estimate $\hat{a}_{\mathrm{RLS}}$ which can minimize an weighted LS criterion

$$J' = \frac{1}{n}\| y - \Phi_y \hat{a}_{\mathrm{RLS}} \|^2 + \| \hat{a}_{\mathrm{RLS}} \|_R^2, \tag{19}$$

The regularized LS estimate is given by

$$\hat{a}_{\mathrm{RLS}} = \left[\frac{1}{n} \Phi_y^T \Phi_y + R \right]^{-1} g \tag{20}$$

which is a biased estimate but it plays an important role of estimating the noise variance. In order to deal with an ill-conditioned case when the noise variance is rather small, we employ an alternative regularization matrix R, which will be specified in the next section. Now the noise variance satisfies the following relation:

Theorem 3: For a sufficiently large data length n, the noise variance σ_e^2 satisfies that

$$\sigma_e^2 = \frac{\lim_{n\to\infty} J' - a^H R \lim_{n\to\infty} \hat{a}_{\mathrm{RLS}}}{1 + a^H \lim_{n\to\infty} \hat{a}_{\mathrm{RLS}}} \tag{21}$$

4 DATA-ADAPTIVE ITERATIVE ALGORITHM

4.1 Replacements with Their Estimates

(16) and (21) give the key relation for the minimum MSE-based rank determination in the overdetermined CLS estimation. However, since the true values of the unknown parameters a and σ_e^2 are included in (16) and (21), they should be replaced by their estimate using an alternative set of regularization parameters $\{\mu_j\}$ as

$$\hat{a}(\{\mu_i\}) = \sum_{i=1}^{m} \frac{1}{\lambda_i - \hat{\sigma}_e^2(\{\mu_j\}) + \mu_i} v_i v_i^H g \tag{22}$$

$$\hat{\sigma}_e^2(\{\mu_j\}) = \frac{J'(\{\mu_j\}) - \hat{a}^H(\{\mu_j\}) R(\{\mu_j\}) \hat{a}_{\mathrm{RLS}}(\{\mu_j\})}{1 + \hat{a}^H(\{\mu_j\}) \hat{a}_{\mathrm{RLS}}(\{\mu_j\})} \tag{23}$$

where

$$R(\{\mu_j\}) = \sum_{j=1}^{m} \mu_j v_j v_j^H \tag{24}$$

$$\hat{a}_{\mathrm{RLS}}(\{\mu_j\}) = \sum_{j=1}^{m} \frac{1}{\lambda_j + \mu_j} v_j v_j^H g. \tag{25}$$

The noise variance estimate in (23) can also be rewritten by

$$\hat{\sigma}_e^2(\{\mu_j\}) = \frac{J'(\{\mu_j\}) - Q_1(\{\mu_j\})}{1 + Q_2(\{\mu_j\})} \tag{26}$$

where

$$J'(\{\mu_j\}) = \frac{1}{n}\| y \|^2 - \sum_{j=1}^{m} \frac{\left| v_j^H g \right|^2}{\lambda_j + \mu_j} \tag{27a}$$

$$Q_1(\{\mu_j\}) = \hat{a}^H(\{\mu_j\}) R(\{\mu_j\}) \hat{a}_{\mathrm{RLS}}(\{\mu_j\}) = \sum_{j=1}^{m} \frac{\mu_j \left| v_j^H g \right|^2}{(\lambda_j - \hat{\sigma}_e^2(\{\mu_i\}) + \mu_j)(\lambda_j + \mu_j)} \tag{27b}$$

$$Q_2(\{\mu_j\}) = \hat{a}^H(\{\mu_j\}) \hat{a}_{\mathrm{RLS}}(\{\mu_j\}) = \sum_{j=1}^{m} \frac{\left| v_j^H g \right|^2}{(\lambda_j - \hat{\sigma}_e^2(\{\mu_i\}) + \mu_j)(\lambda_j + \mu_j)} \tag{27c}$$

Thus, by substituting the estimates in (22) and (23) into (16), we calculate the regularization parameters in (16) in terms of these estimates as

$$\hat{\gamma}_i^*(\{\mu_j\}) = \frac{\hat{\sigma}_e^2(\{\mu_i\})\lambda_i(1 + \| \hat{a}(\{\mu_j\}) \|^2)}{n(\lambda_i - \hat{\sigma}_e^2(\{\mu_i\})) \left| v_i^H \hat{a}(\{\mu_j\}) \right|^2} + \frac{\hat{\sigma}_e^4(\{\mu_i\})}{n(\lambda_i - \hat{\sigma}_e^2(\{\mu_i\}))} \tag{28}$$

where

$$\| \hat{a}(\{\mu_j\}) \|^2 = \sum_{j=1}^{m} \frac{\left| v_j^H g \right|^2}{(\lambda_j - \hat{\sigma}_e^2(\{\mu_i\}) + \mu_j)^2} \tag{29a}$$

$$\left| v_i^H \hat{a}(\{\mu_j\}) \right|^2 = \frac{\left| v_i^H g \right|^2}{(\lambda_i - \hat{\sigma}_e^2(\{\mu_j\}) + \mu_i)^2} \tag{29b}$$

From (28) it is noticed that $\hat{\gamma}_i^*$ is expressed in term of $\{\mu_j\}$. Since $\hat{\gamma}_i^*$ should be equal to μ_i for $j = 1, \cdots, m$, then they should satisfy the nest nonlinear equation

$$\gamma_j^o(\{\mu_i\}) = \mu_j \quad \text{for } j = 1, \cdots, m \tag{30}$$

4.2 Data-Based Iterative Algorithm for Rank Decision

In this section, we give an iterative algorithm of calculating the solution of (30) by placing the constraint that $\mu_1 \le \mu_2 \le \mu_3 \le \cdots \le \mu_m$. It can be noticed from (28) and (30) that we can employ an iterative calculation by setting $\gamma_i^{(k)} = \mu_i$ and $\gamma_i^{(k+1)} = \hat{\gamma}_i^*$. As for the noise variance estimate, $\hat{\sigma}_e^2(\{\mu_i\})$ in the RHS of (28) and the denominator of (27)(29) may be replaced with $\hat{\sigma}_e^{2(k)}$, while $\hat{\sigma}_e^2(\{\mu_j\})$ in the LHS of (26) is replaced by $\hat{\sigma}_e^{2(k+1)}$.

(Step 1) Set $k = 0$ and the initial values of $\gamma_i^{(0)} = \varepsilon$ for $i = 1, \cdots, m$, where ε is chosen larger than machine precision.

(Step 2) Calculate the estimate of the noise variance by using (26) as

$$\hat{\sigma}_e^2(\{\gamma_i^{(k)}\}) = \frac{J'(\{\gamma_i^{(k)}\}) - Q_1(\{\gamma_i^{(k)}\})}{1 + Q_2(\{\gamma_i^{(k)}\})} \tag{31}$$

where $J'(\{\gamma_i^{(k)}\})$, $Q_1(\{\gamma_i^{(k)}\})$ and $Q_2(\{\gamma_i^{(k)}\})$ are defined by (27a), (27b) and (27c) respectively.

(Step 3) Update the regularization parameters in the nest iteration by using (28) and (29) as

$$\gamma_i^{(k+1)} = \frac{\hat{\sigma}_e^2(\{\gamma_i^{(k)}\})\lambda_i(1 + \| \hat{a}(\{\gamma_i^{(k)}\}) \|^2)}{n(\lambda_i - \hat{\sigma}_e^2(\{\gamma_i^{(k)}\})) \left| v_i^H \hat{a}(\{\gamma_i^{(k)}\}) \right|^2} + \frac{\hat{\sigma}_e^4(\{\gamma_i^{(k)}\})}{n(\lambda_i - \hat{\sigma}_e^2(\{\gamma_i^{(k)}\}))} \tag{32}$$

(Step 4): Let the increment of $\gamma_i^{(k)}$ be denoted by $\Delta\gamma_i^{(k+1)} = \gamma_i^{(k+1)} - \gamma_i^{(k)}$. If $\Delta\gamma_i^{(k+1)} > \Delta\gamma_{i+1}^{(k+1)}$, set

$$\Delta\gamma_i^{(k+1)} = \Delta\gamma_{i+1}^{(k+1)} \quad \text{for } i = m-1, m-2, \cdots, 2, 1$$

If it is satisfied that $\left| \Delta\gamma_i^{(k+1)} \right| \le \delta$ for all converged $\gamma_i^{(k+1)}$, then stop the iteration. Otherwise, by using the correction term $\Delta\gamma_i^{(k+1)}$, update the regularization parameters as

$$\gamma_i^{(k+1)} = \gamma_i^{(k)} + \Delta\gamma_i^{(k+1)} \tag{33}$$

Return to Step 2 and repeat the above procedure until $\gamma_i^{(k+1)}$ converges to a constant or until it increases over a specified large value. Thus, it gives the optimal truncation number $K (= M)$ by detecting the intersection of the plots of $\{\lambda_i - \hat{\sigma}_e^2(\{\gamma_i^{(k)}\})\}$ and $\{\gamma_i^{(k+1)}\}$.

We can give sufficient conditions for the convergence of the above algorithm on a basis of the contraction mapping, in a similar way [4] which indicates the relation between the truncation and the detection of signal number theoretically. The details are omitted in this paper.

5. NUMERICAL EXAMPLE

Let the received signal be composed of three complex sinusoids and the independent complex white Gaussian noise. The frequencies are $f_1 = 0.30$, $f_2 = 0.33$ and $f_3 = 0.35$. The number of data was 64 where $m = 30$ and $n = 34$. As shown in Fig.1, the truncation can be determined by the intersection of the two plots between $i = 3$ and $i = 4$ which gives $M = 3$. In the iterative algorithm proposed in 4.2, $\gamma_i^{(k)}$ converges or diverges in a few iteration steps. Fig.2 plots the detection probability per 100 samples for various SNR from -10 dB to 2 dB and shows the comparison of the various approaches. The SNR implies the power ratio of the single sinusoid and the noise. The AIC and MDL criteria were given in [7]. The ABIC was proposed by one of the authors [8] where the low rank approximation for the ordinary LS method [1] was utilized. The proposed algorithm could attain the best performance of detection probability, as shown in Fig.2.

The proposed rank determination is based on the minimization of the asymptotic MSE (13). Actually the truncation is given so that the estimate $\hat{a}(\{\gamma_i^{(k)}\})$ can be a best approximation to the true parameter vector a in (18). Fig.3 clarifies that the optima truncation can attain the good estimation,

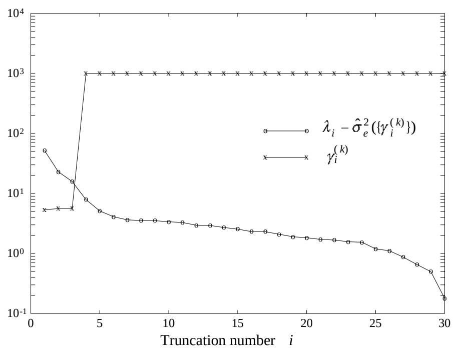

Fig.1 Plots of $\lambda_i - \hat{\sigma}_e^2(\{\gamma_i^{(k)}\})$ and $\{\gamma_i^{(k)}\}$. The truncation is determined by the intersection.

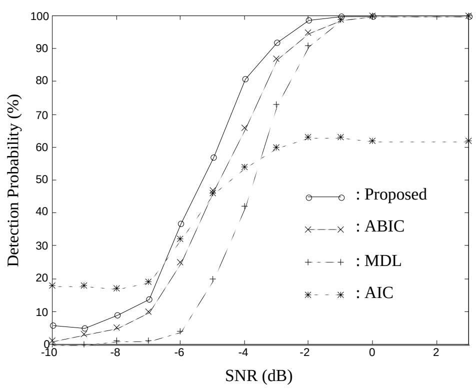

Fig.2 Comparison of detection probability for the roposed algorithm, AIC, MDL [7] and ABIC[8].

while many unstable poles appear if the adequate truncation is not employed.

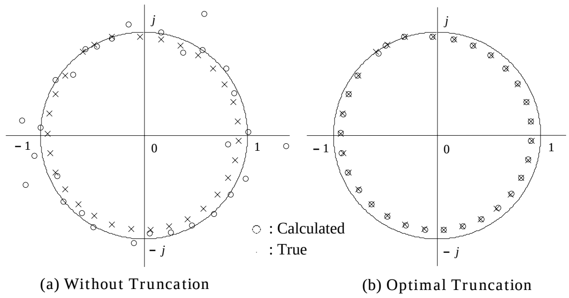

Fig.3 Plots of poles of the estimated model (indicated by the circles) with order 30, and the theoretical model (indicated by the crosses) calculated from (18).

6. CONCLUSION

We have presented the minimum MSE-based rank determination algorithm for the eigenvalue decomposition which is effective to the decision of the number of sinusoids imbedded in white noise. Further, the data-adaptive iterative algorithm using only accessible data has been clarified to estimate the noise variance and the regularization parameters simultaneously. Finally the simulation results indicated that the proposed scheme is effectively applied to a case with a small number of observed data.

REFERENCES

[1] D.W. Tuft and R. Kumaresan and, "Estimation of frequencies of multiple sinusoids: Making linear prediction perform like maximum likelihood", Proc. IEEE, vol.70, pp.975-989, 1982.

[2] M.A. Rahman and K.B. Yu, "Total least squares approach for frequency estimation using linear prediction", IEEE Trans. Acoust. Speech, Signal Processing, vol.ASSP-35, pp.1440-1154, 1987.

[3] S.V. Huffel and J. Vandewalle, The Total Least Squares Problems, SIAM, 1991.

[4] J. Xin, H. Ohmori and A. Sano, "Minimum MSE based regularization for system identification in the presence of input and output noise", Proc. the 34th IEEE CDC, 1995.

[5] C.E. Davila and H.L. Chiang, “An algorithm for pole-zero system model order estimation”, IEEE Trans. Signal Processing, vol.SP-43, pp.1013-1017, 1995.

[6] E. Karlsson and E. Sjostrom, “IV subspace system identification of noisy input-output system”, Proc. 10th IFAC Symp. System Identification, vol.2, pp.385-390, 1994.

[7] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria", IEEE Trans. Acoust. Speech, Signal Processing, vol.ASSP-33, pp.387-392, 1985.

[8] A. Sano and H. Tsuji, Proc. "Optimal determination of number of sinusoids in SVD-based spectral estimation", Proc. ICASSP'90, pp.2563-2566, 1990.