DETECTION OF CISOIDS IN NOISE

Joakim Sorelius
Systems and Control Group, Uppsala University,
SE-751 03 Uppsala, Sweden
e-mail: js@syscon.uu.se

ABSTRACT

In this paper, a rank test for detecting the number of cisoids in noise is presented. The method is based on Gaussian Lower triangular - Diagonal - Upper triangular (LDU) decomposition of a Hankel data matrix, and the method is especially useful for short data records.

1 INTRODUCTION

There is a considerable interest in methods for estimating the parameters of damped and undamped cisoids from noise-corrupted measurements. The cisoid-in-noise data model typically appears in applications such as radar, communications and nuclear magnetic resonance tomography. Most of these methods require the number of signals, or the *order* of the system, to be known. An important step in obtaining automated estimation procedures is thus the development of methods for order estimation.

Many classical order estimation schemes, such as the Final prediction error (FPE) and Akaike's information criterion (AIC) [1, 2], or improved versions of these methods, can be difficult to use because of their computational complexity and the multivariable search they require. In brief, these methods evaluate a criterion function, which depends on the parameters of the signals, for increasing orders of the system. The criterion function is then inspected, and the number of signals that gives its minimum value, or the number at which the decrease in the criterion function becomes insignificant, is chosen as an estimate of the order of the system.

Another important class of order estimators is formed by the so called *rank test methods*. These methods rely on the fact that a certain Hankel matrix has rank equal to the system order. Typically, the rank test methods perform a decomposition (e.g., the eigenvalue decomposition, EVD) of the Hankel matrix to obtain some quantity that (under the assumptions made) has a known distribution, and statistical methods can then be applied to estimate the rank of the matrix from the data. Most of the proposed methods work on a Hankel covariance matrix (see, e.g., [3, 4] and the references therein).

While the above methods often have an excellent performance in noisy environments, one drawback is that they require a relatively large number of data points to provide accurate estimates of the model order.

In some applications the signal-to-noise ratio (SNR) might be high, but the available amount of data might be small. One possibility is then to work on a Hankel matrix formed from the measured signal itself. One such order estimation scheme using the SVD, is presented in [5].

In this paper we propose an order estimation scheme for short data records that works on the LDU-decomposition (see Section 3 below) of the Hankel data matrix. The LDU-decomposition being a linear procedure, the distribution of the data is preserved, and in the case of Gaussian noise, a very simple significance test for the number of signals results.

The outline of the paper is as follows: Section 2 introduces the data model used throughout the paper and Section 3 introduces the LDU decomposition. In Section 4 the proposed algorithm is summarized. Some numerical examples illustrating the performance of the method are presented in Section 5. Section 6 finally states the conclusions.

2 NOTATIONS AND ASSUMPTIONS

Consider the following (possibly damped) sinusoidal signal:

$$\hat{x}(t) = x(t) + w(t) \tag{1}$$

$$= \sum_{k=1}^{r} a_k e^{(2\pi j\omega_k - \beta_k)t} + w(t) \tag{2}$$

In (2), the complex number a_k determines the initial amplitude and phase of the kth signal whereas ω_k and β_k are its angular frequency and damping factor respectively. We stack N snapshots of the measured signal $\hat{x}(t)$ in a $(p \times m)$-dimensional Hankel matrix as

$$\hat{X} = \begin{pmatrix} \hat{x}(1) & \hat{x}(2) & \cdots & \hat{x}(m) \\ \hat{x}(2) & \hat{x}(3) & \cdots & \\ \vdots & & \ddots & \\ \hat{x}(p) & & & \hat{x}(N) \end{pmatrix} \tag{3}$$

$$= X + W \tag{4}$$

where $m = N - p + 1$ (we chose $m \leq p$). The noise $w(t)$ is assumed to be complex Gaussian, so that $\mathrm{vec}\, W \sim \mathbb{C}\mathcal{N}(0, \mathcal{W})$. It is straightforward to show that

$$\mathrm{rank}\, X = r, \;\; r \leq \min(m, p) \tag{5}$$

so that the order of the system is given by the rank of X. A new method for determining the rank of X, given the noisy measurement $\hat{X}$, is outlined below.

3 THE LDU DECOMPOSITION

In this section, the Gaussian Lower triangular - Diagonal - Upper triangular (LDU) decomposition of $\hat{X}$ (defined in (3)) is presented (the derivations are patterned from [6]).

As the name suggests, the LDU-decomposition partitions the matrix $\hat{X}$ into a product of three matrices: L, which is lower triangular, D, which is diagonal, and U, which is upper triangular. The decomposition is usually performed by successive Gaussian elimination. Some kind of pivoting operation is necessary to ensure numerical stability of the decomposition procedure. We shall see that the test to be developed below requires the pivoting to be complete. With complete pivoting we mean that at each step of the Gaussian elimination the current sub-matrix is searched for its largest element (in absolute magnitude), which is shifted to the top left corner by column and row interchanges (this is in contrast to partial pivoting which only shifts the largest element in the first column of the sub-matrix). Pivoting is discussed, e.g., in [7] and the numerical implementation of the LDU-decomposition is treated in detail in [8].

LDU-decomposition with complete pivoting of X yields

$$PXQ = LDU \tag{6}$$

where P and Q are permutation matrices corresponding to the pivoting and D is a diagonal matrix which will have a certain structure due to the row and column pivoting as will be explained below. As mentioned before, matrices L and U^T are lower triangular, and they are normalized to have ones along the diagonals. If the $p \times m$ matrix X has rank r, then the LDU-decomposition with complete pivoting can be partitioned as

$$LDU = \begin{pmatrix} L_{11} & 0 & 0 \\ L_{21} & L_{22} & 0 \\ L_{31} & L_{32} & I_{m-p} \end{pmatrix} \begin{pmatrix} D_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} U_{11} & U_{12} \\ 0 & I_{p-r} \\ 0 & 0 \end{pmatrix} \tag{7}$$

The corresponding LDU-decomposition for $\hat{X}$ is denoted

$$\hat{P}\hat{X}\hat{Q} = \hat{L}\hat{D}\hat{U}, \tag{8}$$

and is partitioned as

$$\hat{L}\hat{D}\hat{U} = \begin{pmatrix} \hat{L}_{11} & 0 & 0 \\ \hat{L}_{21} & \hat{L}_{22} & 0 \\ \hat{L}_{31} & \hat{L}_{32} & I_{m-p} \end{pmatrix} \begin{pmatrix} \hat{D}_1 & 0 & 0 \\ 0 & \hat{D}_2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{U}_{11} & \hat{U}_{12} \\ 0 & \hat{U}_{22} \\ 0 & 0 \end{pmatrix}. \tag{9}$$

In (7) and (9), the row as well as the column partition are by r, $p-r$ and $m-p$. An important property of the complete pivoting is that it ensures that the $r \times r$ diagonal matrix D_1 in (7) is nonsingular and that the diagonal of D contains exactly $m-r$ zeros, placed as indicated in (7). Also, it follows from (6) that L_{11}, $\hat{L}_{11}$, U_{11}^T and $\hat{U}_{11}^T$ are unit lower triangular, and choosing $L_{22} = I_{p-r}$ and $L_{32} = 0$ ensures that the decomposition (7) is unique [9]. Since $\hat{X}$ is assumed to have full column rank, $\hat{L}_{22}$ as well as $\hat{U}_{22}^T$ will be unit lower triangular (but different from their exact counterparts) so that also the decomposition (9) is unique.

As the SNR grows to infinity, the matrix $\hat{X}$ tends to X and it can be shown that the roofed quantities in the sample LDU-decomposition (9) converge in probability to the true ones, given by (7) (see, e.g., [6]). In other words,

$$\hat{P}\hat{X}\hat{Q} = \hat{L}\hat{D}\hat{U} \;\xrightarrow{p}\; PXQ = LDU. \tag{10}$$

Particularly, $\hat{D}_2 \xrightarrow{p} 0$ and we shall see that under the null hypothesis,

$$H_0 : \; \mathrm{rank}\, X = r \tag{11}$$

we can derive the statistical properties of the $p-r$ vector

$$\hat{d}_2 \triangleq \mathrm{diag}\,(\hat{D}_2) \tag{12}$$

which will enable us to develop the test for determining the rank of X. To do this, we need some additional notation.

Let Δ_i be a $(p-r) \times (p-r)$ matrix that has 1 as its (i,i)-element and zeros elsewhere. Also define

$$\Delta = (\Delta_1 \; \Delta_2 \cdots \Delta_{p-r})^T. \tag{13}$$

It is easily verified that the $(p-r)^2 \times (p-r)$ matrix Δ satisfies the orthogonality property

$$\Delta^T \Delta = I_{p-r} \tag{14}$$

and also that

$$\mathrm{vec}\, \hat{D}_2 = \Delta \hat{d}_2 \tag{15}$$

with $\hat{d}_2$ as defined in (12). From (10), we infer that, for sufficiently high SNR,

$$\hat{P}(\hat{X} - X)\hat{Q} \;\xrightarrow{p}\; (\hat{L}\hat{D}\hat{U} - LDU). \tag{16}$$

For notational convenience, introduce the matrices

$$\begin{array}{ll} \hat{L}_1 = \begin{pmatrix} \hat{L}_{11} \\ \hat{L}_{21} \end{pmatrix} & \hat{L}_2 = \begin{pmatrix} 0 \\ \hat{L}_{22} \end{pmatrix} \\ \hat{U}_1 = \begin{pmatrix} \hat{U}_{11} & \hat{U}_{12} \end{pmatrix} & \hat{U}_2 = \begin{pmatrix} 0 & \hat{U}_{22} \end{pmatrix} \end{array} \tag{17}$$

and similarly define L_1, L_2, U_1 and U_2. Then it is easy to see that under H_0

$$\begin{aligned} \hat{L}\hat{D}\hat{U} - LDU &= (\hat{L}_1\hat{D}_1\hat{U}_1 + \hat{L}_2\hat{D}_2\hat{U}_2 - L_1D_1U_1) \\ &= (\hat{L}_1 - L_1)\hat{D}_1\hat{U}_1 + L_1\left(\hat{D}_1 - D_1\right)\hat{U}_1 \\ &\quad + L_1D_1(\hat{U}_1 - U_1) + \hat{L}_2\hat{D}_2\hat{U}_2 \end{aligned} \tag{18}$$

and using (10) and (16) we obtain, as the SNR increases, that

$$\hat{P}(\hat{X}-X)\hat{Q} = (\hat{L}_1 - L_1)\hat{D}_1\hat{U}_1 + \hat{L}_1(\hat{D}_1 - D_1)\hat{U}_1 + \hat{L}_1\hat{D}_1(\hat{U}_1 - U_1) + \hat{L}_2\hat{D}_2\hat{U}_2. \quad (19)$$

Introduce the matrices $\hat{H}$ and $\hat{K}$ defined by

$$\hat{H} = \begin{pmatrix} -\hat{L}_{22}^{-1}\hat{L}_{21}\hat{L}_{11}^{-1} & \hat{L}_{22}^{-1} & 0 \end{pmatrix} \quad (20)$$

$$\hat{K} = \begin{pmatrix} -\hat{U}_{11}^{-1}\hat{U}_{12}\hat{U}_{22}^{-1} \\ \hat{U}_{22}^{-1} \end{pmatrix} \quad (21)$$

These matrices satisfy the readily verified properties

$$\begin{array}{ll} \hat{H}\hat{L}_1 = 0, & \hat{H}\hat{L}_2 = I_{p-r}, \\ \hat{U}_1\hat{K} = 0, & \hat{U}_2\hat{K} = I_{p-r}. \end{array} \quad (22)$$

so that pre- and postmultiplying equation (19) with $\hat{H}$ and $\hat{K}$ respectively yields

$$\hat{D}_2 = \hat{H}\hat{P}\sqrt{N}(\hat{X}-X)\hat{Q}\hat{K} \quad (23)$$

and by (12) and (14)–(15) we have that

$$\begin{aligned} \hat{d}_2 &= \Delta^T \mathrm{vec}\,(\hat{D}_2) \\ &= \Delta^T \mathrm{vec}\,(\hat{H}\hat{P}(\hat{X}-X)\hat{Q}\hat{K}) \\ &= \Delta^T (\hat{K}^T\hat{Q}^T \otimes \hat{H}\hat{P})\mathrm{vec}\,(\hat{X}-X) \\ &= \Delta^T (\hat{K}^T \otimes \hat{H})(\hat{Q}^T \otimes \hat{P})\mathrm{vec}\,(\hat{X}-X) \quad (24) \end{aligned}$$

Now, from the distribution of $\mathrm{vec}\,(\hat{X}-X) = \mathrm{vec}\,(W)$ we finally obtain

$$\hat{d}_2 \sim \mathbb{C}\mathcal{N}(0, \mathcal{D}) \quad (25)$$

with $\mathcal{D}$ being the asymptotic counterpart of

$$\hat{D} = \Delta^T(\hat{K}^T \otimes \hat{H})(\hat{Q}^T \otimes \hat{P})\widehat{\mathcal{W}}(\hat{Q} \otimes \hat{P}^T)(\hat{K} \otimes \hat{H}^T)\Delta \quad (26)$$

The result (25) has a clear potential for developing an algorithm for order estimation. To see this, we first introduce a transformation of $\hat{d}_2$ which turns it into a real-valued quantity. Let

$$v = (v_{1R} + iv_{1I} \cdots v_{nR} + iv_{nI})^T \quad (27)$$

be a complex-valued vector. Then define

$$f(v) \triangleq (v_{1R}\ v_{1I}\ v_{2R}\ v_{2I} \cdots v_{nR}\ v_{nI})^T \quad (28)$$

Similarly, if A is an $(m|n)$ matrix with (μ,ν)th element $A_{\mu\nu} = a_{\mu\nu R} + ia_{\mu\nu I}$, introduce the matrix $f(A)$ of dimension $(2m|2n)$ as

$$f(A) = \begin{pmatrix} \tilde{a}_{11} & \cdots & \tilde{a}_{1n} \\ \vdots & & \vdots \\ \tilde{a}_{m1} & & \tilde{a}_{mn} \end{pmatrix} \quad (29)$$

where the block $\tilde{a}_{\mu\nu}$ is

$$\tilde{a}_{\mu\nu} = \begin{pmatrix} a_{\mu\nu R} & -a_{\mu\nu I} \\ a_{\mu\nu I} & a_{\mu\nu R} \end{pmatrix}$$

Then, defining the quantities

$$\bar{d}_2 = f(\hat{d}_2) \quad (30)$$

$$\bar{\mathcal{D}} = f(\hat{\mathcal{D}}) \quad (31)$$

and using (25), we obtain

$$\hat{\eta} \triangleq \bar{d}_2^T \bar{\mathcal{D}}^{-1} \bar{d}_2 \sim \chi^2(2(p-r)) \quad (32)$$

i.e., under the null hypothesis (11), the (real-valued) test quantity $\hat{\eta}$ is χ^2-distributed with $2(p-r)$ degrees of freedom.

4 THE ALGORITHM

Summarizing the results of the above derivations leads to the following rank test for determining the number of cisoids:

1. Perform the LDU decomposition of the data matrix $\hat{X}$. Set $j = 0$.
2. Partition the decomposition to obtain the $2(m-j)$-vector $\bar{d}_2$, and construct the covariance matrix $\bar{\mathcal{D}}$.
   - If $\hat{\eta} \le \chi^2_\alpha(2(m-j))$, then $\mathrm{rank}\,X = j$ and the number of signals is $r = j$.
   - Else, set $j = j+1$. If $j < m$, go to the beginning of Step 2, otherwise the test is terminated.
3. If the test terminates with $j = m$, then the rank of X could not be determined and m must be increased to find the system order. ■

The significance level α is defined as

$$\alpha = \mathrm{prob}\left(u > \chi^2_\alpha(m) | u \sim \chi^2(m)\right). \quad (33)$$

The parameter α is called the probability of false alarm; it is the probability of declaring $\mathrm{rank}\,X > r$ when in fact it holds true that $\mathrm{rank}\,X = r$. The threshold $\chi^2_\alpha(m)$ for different values of m and α can be read from a table of the χ^2 distribution, see, e.g., [10].

5 NUMERICAL EXAMPLES

In this section we present some numerical examples to illustrate the performance of the proposed detection scheme. Consider two (i.e., $r = 2$) undamped exponentials, the parameters of which are

$$\begin{array}{lll} \beta_1 = 0 & f_1 = 0.42 & a_1 = 1 \\ \beta_2 = 0 & f_2 = 0.52 & a_2 = e^{j\pi/4} \end{array}$$

(see (2)). The noise is assumed to be zero mean, circularly complex white Gaussian noise, i.e.,

$$\begin{array}{l} Ew(t) = 0,\ Ew(t)w(s) = 0 \text{ all } t,s \\ Ew(t)w^*(s) = \sigma^2\delta_{t,s} \end{array} \quad (34)$$

which means that $\mathcal{W} = \sigma^2 I$ (in the simulations we assume σ^2 to be known). Define the signal-to-noise ratio as SNR $= 10\log(1/\sigma^2)$ i.e., the ratio between the power of each signal and the noise power.

We will assess the performance of the method in terms of the probability of error, i.e., the percentage of times that the incorrect order is estimated for a number of runs, each on independent realizations of the data. In all the simulations we use 3000 Monte Carlo simulations.

In Figure 1 the (in-)dependency of the proposed algorithm on the number of data points is shown. The probability of false alarm is chosen to $\alpha = 0.01$ and $m = 6$. As expected, the method performs well even for short sample lengths for a variety of SNR values (note that for SNR=15dB, the correct order was always determined).

In Figure 2 we investigate the sensitivity of the method to the size of the data matrix $\hat{X}$. We used 25 data points and again α was set to 0.01. We note that a large m is not always desirable, but for high SNR, the choice of m has less impact on the performance of the algorithm. Note that $m > 4$ is required to separate the signals. This is due to the transformation $f(\cdot)$ introduced in (30) .

6 CONCLUSIONS

In this paper, an order estimation scheme for (possibly damped) cisoids in noise was presented. The advantages of the method is its simplicity and that it performs well for short data records and for a variety of medium-to-high SNR values, as verified by the numerical examples.

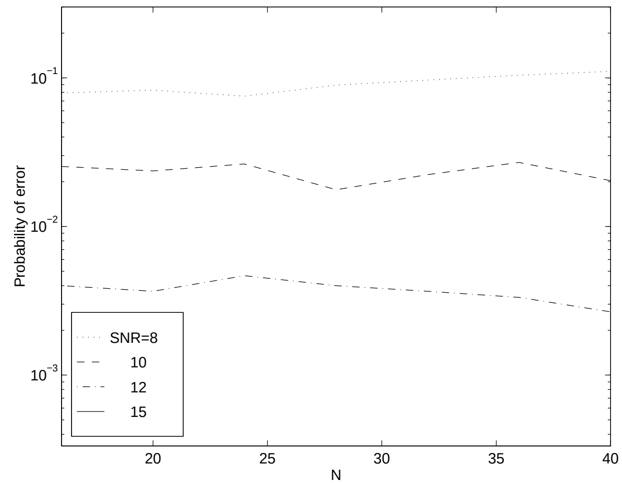

Figure 1: Probability of error as a function of the number of data points N ($\alpha = 0.01,\ m = 6$).

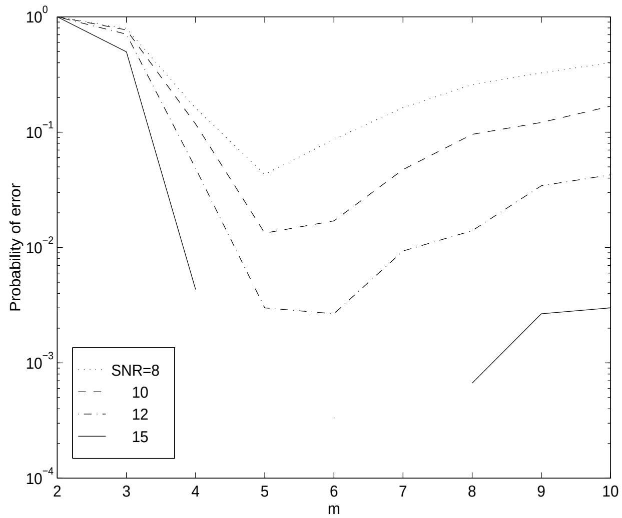

Figure 2: Probability of error as a function of m, the number of columns of the data matrix ($N = 25$, $\alpha = 0.01$).

References

[1] L. D. Davisson, “The prediction error of stationary Gaussian time series of unknown covariance,” *IEEE Transactions on Information Theory*, vol. 11, pp. 527–532, 1965.

[2] H. Akaike, “A new look at the statistical model identification,” *IEEE Transactions on Automatic Control*, vol. 19, pp. 716–723, 1974.

[3] J. Sorelius, T. Söderström, P. Stoica, and M. Cedervall, "Comparative Study of Rank Test Methods for ARMA Order Estimation". Chapter in *Statistical Methods in Control and Signal Processing*, T. Katayama and S. Sugimoto, eds., New York, USA: Mercel Dekker Inc., 1997.

[4] P. Stoica and M. Cedervall, “Detection tests for array processing in unknown correlated noise fields,” *IEEE Transactions on Signal Processing*, vol. 45, pp. 2351–2362, Sept. 1997.

[5] A. A. Shah and D. W. Tufts, “Determination of the dimension of a signal subspace from short data records,” *IEEE Transactions on Signal Processing*, vol. 42, pp. 2531–2534, Sept. 1994.

[6] L. Gill and A. Lewbel, “Testing the rank and definiteness of estimated matrices with applications to factor, state-space and ARMA models,” *Journal of the American Statistical Association*, vol. 87, no. 419, pp. 766–776, 1992.

[7] J. H. Wilkinson, *The Algebraic Eigenvalue Problem.* Oxford: Oxford University Press, 1965.

[8] G. H. Golub and C. F. Van Loan, *Matrix Computations.* Maryland, USA: The Johns Hopkins University Press, 1989.

[9] A. S. Householder, *The Theory of Matrices in Numerical Analysis.* New York, USA: Blaisdell, 1964.

[10] H. Cramér, *Mathematical Methods of Statistics.* Princeton: Princeton University Press, 1966.