ELSEVIER

Signal Processing 81 (2001) 1089–1094

SIGNAL PROCESSING

www.elsevier.nl/locate/sigpro

Short communication

Some comments on "Complex AM signal model for non-stationary signals" by Sircar and Syali

Swagata Nandi, Debasis Kundu*

Department of Mathematics, Indian Institute of Technology Kanpur, P.O. Box I.I.T., Kanpur, Pin 208 016, India

Received 28 April 2000; received in revised form 28 July 2000

Abstract

Recently, Sircar and Syali (Signal Processing 53 (1996) 35–45) introduced the complex AM signal model for analyzing non-stationary speech data. It is assumed that the complex AM model has additive errors, which are independent and identically distributed. They provided an efficient estimation procedure of the unknown parameters using Prony's equation. We re-analyze both the data sets and it is observed that the error assumptions may not be correct in either of them. In this note, we propose a more general complex AM model with additive stationary errors. We use the standard least-squares estimators to estimate the unknown parameters of both the speech data sets. The error analysis indicates that proposed model assumption is correct. Visually, it seems that our estimation procedure provides a better fit than Sircar and Syali (Signal Processing 53 (1996) 35–45) for both the data sets. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: AM signal model; Stationary errors

1. Introduction

Recently, Sircar and Syali [4] introduced the following complex AM signal model for analyzing non-stationary speech data. The discrete time complex random process $y(t)$ consisting of M single-tone AM signals, which can be re-written as follows:

$$y(t) = \sum_{i=1}^{M} A_i(1 + \mu_i e^{j\nu_i t})e^{j\omega_i t} + X(t), \tag{1.1}$$

where $A_i(\neq 0)$ is the complex-valued carrier amplitude of the constituent signal, μ_i $(\neq 0)$ is the modulation index (may be complex valued), ω_i is the carrier angular frequency and ν_i is the independent modulating angular frequency. Note that in [4] the μ_i's are considered to be real only. It is assumed that $X(t)$'s are independent and identically distributed (i.i.d.) random variables with zero mean and finite variance.

Using the above assumption, Sircar and Syali [4] proposed an estimation procedure of the unknown parameters using Prony's difference type equations and analyzed two non-stationary speech data. Unfortunately, they did not perform any error analysis to validate the error assumption of the model. We re-analyze the same two speech data sets and it is observed that the error assumptions may not be correct. In fact it is observed that in both the cases the errors are correlated and for one data set the

* Corresponding author. Tel.: + 91-512-597-636; fax: + 91-512-590-007.

E-mail address: kundu@iitk.ac.in (D. Kundu).

0165-1684/01/$ - see front matter © 2001 Elsevier Science B.V. All rights reserved.
PII: S0165-1684(00)00267-X

errors are not even stationary. It may be mentioned that if the errors are correlated then the Prony's equation does not work, therefore the method proposed by Sircar and Syali [4] cannot be extended if the errors are not independent. Moreover, if the errors are not stationary then it may not be possible to develop any theoretical properties of the estimators. The main aim of this note is to introduce a more general AM model than [4], but at the same time it should be analytically tractable. To make a more general and analytically tractable model, we assume that the errors are from a stationary distribution. We assume the following assumption on the errors.

Assumption 1. $X(t)$ has the following error structure:

$$X(t) = \sum_{k=-\infty}^{\infty} a_k e(t-k), \tag{1.2}$$

where $e(t)$'s are i.i.d. complex-valued random variables with zero mean and finite variance for both the real and imaginary parts and they are uncorrelated. The constants a_k's satisfy

$$\sum_{k=-\infty}^{\infty} |a_k| < \infty . \tag{1.3}$$

Note that Assumption 1 is the standard assumption of a stationary linear process. The stationary AR, MA and ARMA models can be represented as (1.2), when the coefficients a_k's satisfy (1.3). Because of the present error structure (Assumption 1), the model would be able to handle correlated errors. Based on Assumption 1, we propose to use the least-squares estimators (LSEs), which can be obtained by minimizing the residual sums of squares, namely

$$Q(\boldsymbol{A},\boldsymbol{\mu},\boldsymbol{v},\boldsymbol{\omega}) = \sum_{t=1}^{N} \left| y(t) - \sum_{i=1}^{M} A_i(1+\mu_i e^{jv_i t}) e^{j\omega_i t} \right|^2, \tag{1.4}$$

with respect to $\boldsymbol{A} = (A_1, \ldots, A_M)$, $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_M)$, $\boldsymbol{v} = (v_1, \ldots, v_M)$ and $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_M)$. The main reason to consider the LSEs is two folds. First of all, it is well known that if the errors are correlated then the usual Prony-estimators or their variants do not work well and secondly it should be possible to obtain the large sample properties of the LSEs, which are quite important in practice. Sometimes, the LSEs can be computationally very expensive, particularly if the initial guesses are poor. But it is observed that the Periodogram function (will be explained in the next section) can be used to obtain good initial guesses and from those initial guesses the LSEs can be obtained quite efficiently.

We re-analyze both the vowel sounds in Section 2 under Assumption 1 and obtain the estimated signals. In Section 3, we discuss some open problems and make some final conclusions.

2. Data analysis and discussions

In this section, we re-analyze the two sustained vowel sounds, namely "aaa" and "uuu". Both the data sets contain 512 signal values sampled at 10 kHz frequency. Sircar and Syali [4] fitted the AM model (1.1) under the i.i.d. error assumptions and with real μ_i's. We re-analyze both the data sets by fitting the model (1.1) with the error Assumption 1 and consider both real and complex μ_i's separately.

Data set 1 ("aaa"):

The original data set is presented in Fig. 1. We plot the Periodogram function

$$I(\lambda) = \frac{1}{N} \left| \sum_{t=1}^{N} y(t) e^{-j\lambda t} \right|^2 \tag{2.1}$$

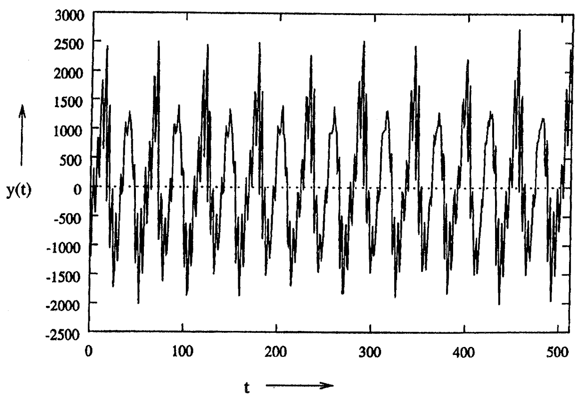

Fig. 1. Original "aaa" vowel sound.

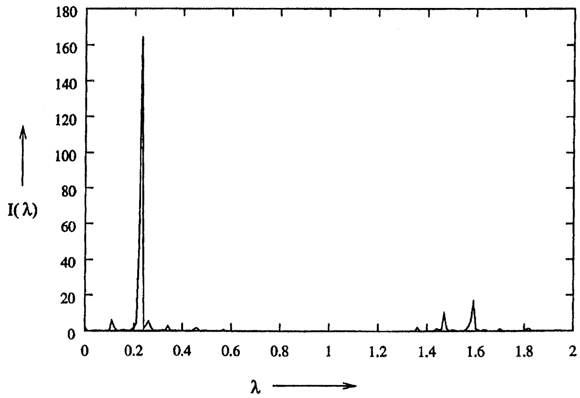

Fig. 2. Periodogram intensity function of "aaa" vowel sound.

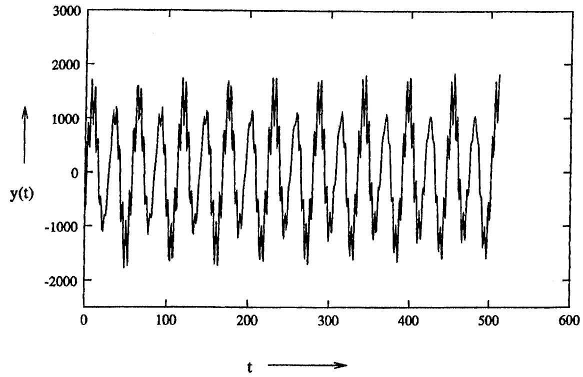

Fig. 3. Estimated "aaa" vowel sound with real modulation index.

of the data set 1 in Fig. 2. The Periodogram function roughly indicates that there are four peaks and therefore M can be estimated as 2, the same was proposed in [4] also. Assuming $M = 2$ we can obtain the initial guesses of v_i's and ω_i's. First we consider the model (1.1) with real μ_i's as it was proposed in [4]. Knowing the initial guesses of the non-linear frequencies, we obtain the initial guesses of the A_i's and μ_i's by using the simple linear least squares regression technique, see [2]. Using those initial guesses, the LSEs of the different parameters are as follows:

$$\hat{A}_1 = -122.1314 - \mathrm{j}220.5148, \quad \hat{\mu}_1 = 4.7028,$$

$$\hat{\omega}_1 = 0.1119, \quad \hat{v}_1 = 0.1139,$$

$$\hat{A}_2 = 242.1911 + \mathrm{j}162.9930, \quad \hat{\mu}_2 = 0.7871,$$

$$\hat{\omega}_2 = 1.4720, \quad \hat{v}_2 = 0.1106.$$

The predicted values of $y(t)$, say $\hat{y}(t)$,

$$\hat{y}(t) = \hat{A}_1(1 + \hat{\mu}_1 \mathrm{e}^{\mathrm{j}\hat{v}_1 t})\mathrm{e}^{\mathrm{j}\hat{\omega}_1 t} + \hat{A}_2(1 + \hat{\mu}_2 \mathrm{e}^{\mathrm{j}\hat{v}_2 t})\mathrm{e}^{\mathrm{j}\hat{\omega}_2 t}$$

are plotted in Fig. 3. The predicted values match reasonably well with the observed values. We obtain the estimated errors as

$$\hat{X}(t) = y(t) - \hat{y}(t).$$

We want to test whether the errors are independent or not. We use the run test [1], and the test statistic value $z = -6.2932$ confirms that the residuals are not independent. The autocorrelation function and the partial autocorrelation function suggest that the residual should be an AR(3) process and it can be estimated as

$$\hat{X}(t) = 0.8353\hat{X}(t-1) - 0.7585\hat{X}(t-2) + 0.7174\hat{X}(t-3) + e(t). \tag{2.2}$$

Performing the run test on $\hat{e}(t)$ we obtain $z = -1.1679$, which verifies the independent assumption on $e(t)$ at the 95% level of significance. Since all the roots of the polynomial equation

$$z^3 - 0.8353z^2 + 0.7585z - 0.7174 = 0$$

are less than 1 in absolute value, therefore (2.2) can be written in the form of (1.2) when a_k''s satisfy (1.3). From (2.2), we can say that $X(t)$'s are from a stationary AR(3) process, and it satisfies Assumption 1.

Now we fit model (1.1) to the same data set assuming complex μ_i's. The results are as follows:

$$\hat{A}_1 = -125.2406 - \mathrm{j}181.0094,$$

$$\hat{\mu}_1 = 4.9096 - \mathrm{j}2.5431, \quad \hat{\omega}_1 = 0.1142,$$

$$\hat{v}_1 = 0.1133,$$

$$\hat{A}_2 = 206.4938 + \mathrm{j}170.3605,$$

$$\hat{\mu}_2 = 0.0635 - \mathrm{j}1.4827, \quad \hat{\omega}_2 = 1.4717,$$

$$\hat{v}_2 = 0.1150.$$

The predicted values of $y(t)$, say $\hat{y}(t)$, are plotted in Fig. 4. The predicted values match quite well with

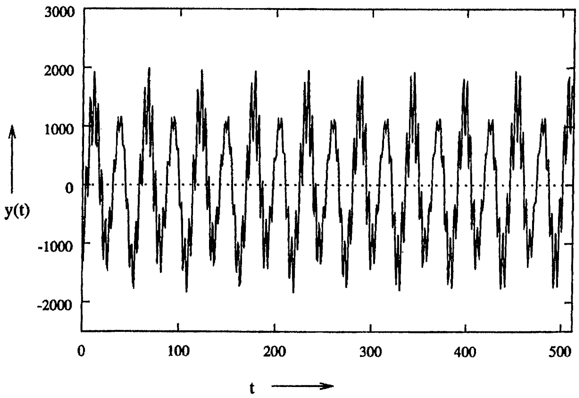

Fig. 4. Estimated "aaa" vowel sound with complex modulation index.

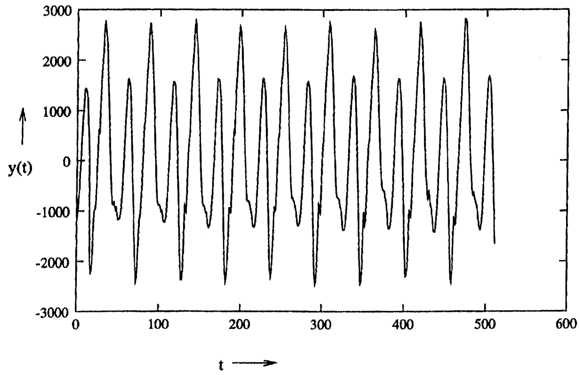

Fig. 5. Original "uuu" vowel sound.

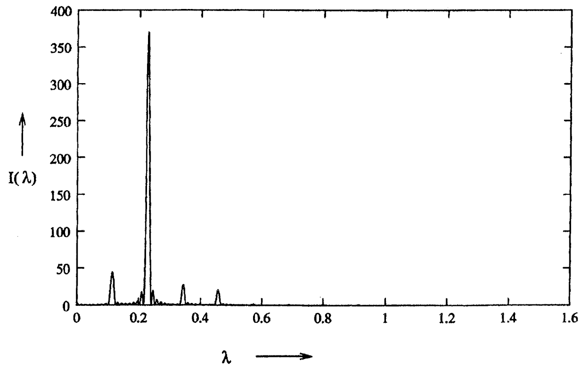

Fig. 6. Periodogram intensity function of "uuu" vowel sound.

the observed values. We compute the mean residual sums of squares for both the cases to compare their performances and it is observed that for real μ_i's the square root of the mean residual sums of squares is 466.872 and for the complex μ_i's it is 365.462. Therefore, clearly model (1.1) with complex μ_i's provides better fit than real μ_i's, which is not very surprising.

Similarly, as before, we perform the run test on the estimated errors and the test statistic value $z = -4.5829$ indicates that the residuals are not independent. It is observed that the residuals should be an AR(1) process and it can be estimated as

$$\hat{X}(t) = 0.3552\hat{X}(t-1) + e(t). \tag{2.3}$$

Performing the run test on $\hat{e}(t)$ we obtain $z = -1.6792$, which verifies the independent assumption on $e(t)$ and (2.3) suggests that $X(t)$'s are from a stationary AR(1) process, and it satisfies Assumption 1.

Data set 2 ("uuu"):

The data are plotted in Fig. 5. The Periodogram function is plotted in Fig. 6. Periodogram function indicates that $M = 2$. Similarly, as the vowel sound "aaa", first, we try to fit the model (1.1) with real μ_i's. We obtain the LSEs as before, but unfortunately in this case the estimated errors are not stationary. Therefore, model (1.1) cannot be used to fit the data set "uuu" with real modulation indexes.

Now we use model (1.1) with complex modulation indexes. Using the initial guesses as before we obtain the LSEs of the unknown parameters as follows:

$$\hat{A}_1 = -521.0508 + j378.3746,$$

$$\hat{\mu}_1 = -1.6067 + j2.0964, \quad \hat{\omega}_1 = 0.1146,$$

$$\hat{v}_1 = 0.1135,$$

$$\hat{A}_2 = -334.4524 + j290.0626,$$

$$\hat{\mu}_2 = 0.7591 - j0.2818, \quad \hat{\omega}_2 = 0.1146,$$

$$\hat{v}_2 = 0.3434.$$

The estimated $y(t)$'s are plotted in Fig. 7. Using the run test on the residuals, we obtain $z = -11.8920$, which clearly indicates that the errors are correlated. From the auto-correlation function and partial auto-correlation function we come to a conclusion that the errors are from a AR(3) process and it can be estimated as

$$\hat{X}(t) = 1.0904\hat{X}(t-1) - 0.5067\hat{X}(t-2) + 0.1065\hat{X}(t-3) + e(t). \quad (2.4)$$

The test statistic $z = -1.3297$ is obtained by performing the run test on $\hat{e}(t)$. The z value clearly indicates that $e(t)$'s are not dependent. Note that all the roots of the polynomial equation

$$z^3 - 1.0904z^2 + 0.5067z - 0.1065 = 0$$

are less than one in absolute value. Therefore, error assumptions are satisfied in this case.

From the above two data analysis, it is clear that model (1.1) can be used to analyze various sustained vowel sounds. It is important to assume that the errors are from a stationary process. We also observe that if we assume that the modulation indexes are real valued, then the model may not work for both the cases. Since, for the real modulation indexes, the total number of parameters becomes less, we recommend first to use model (1.1) with real modulation indexes, if it is does not work then we should try the model with complex modulation indexes. On the other hand, if one prefers a generic model and better fit, then model (1.1) with complex μ's can be used. Proper testing procedure should be developed for this purpose.

3. Conclusions

In this note we introduce the AM signal model when the additive errors are from a stationary linear process. We propose to use the LSEs for analyzing any non-stationary signals and it is observed that the performances are quite satisfactory. We have not discussed two important issues. First, although we have proposed to use the LSEs we have not provided any theoretical properties of the LSEs. Note that if the errors are i.i.d, the results of [3] can be used to obtain the large sample properties of the LSEs. Based on that, we believe that the LSEs under Assumption 1 will be consistent and asymptotically normally distributed. These properties can be used to obtain the error bounds of the different estimators and also to test whether real or complex μ_i's should be used. The work is in progress in that direction and will be reported later.

Another important practical issue is the estimation of M. Although we have used the Periodogram function to estimate M, it is still quite subjective in nature. Sircar and Syali [4] also did not mention anything on that issue. Since there are two more small peaks in the Periodogram function (Fig. 2) of the data set 1, for curiosity, we have fitted model (1.1) to that data set with $M = 3$ and complex μ's. The estimated "aaa" vowel signal with $M = 3$ is plotted in Fig. 8. Note that Fig. 8 provides a much better fit to Fig. 1 than Fig. 4. This may not be very surprising, because when $M = 3$ the model has more parameters than when $M = 2$. It is expected

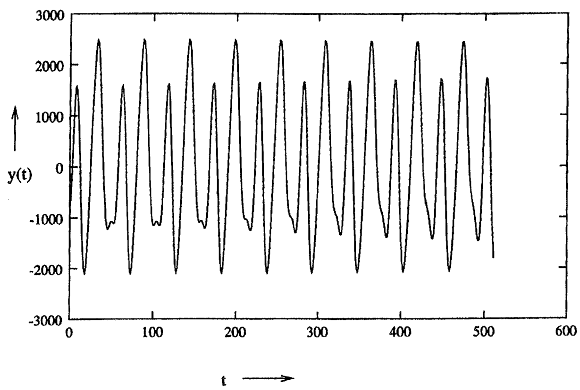

Fig. 7. Estimated "uuu" vowel sound with complex modulation index.

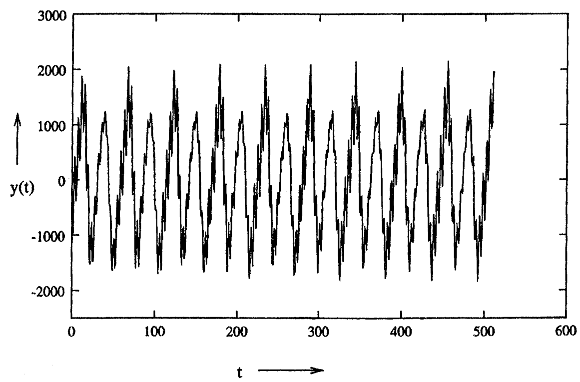

Fig. 8. Estimated "aaa" vowel sound considering the model with three components and complex modulation index.

that when $M = 3$ the standard errors of the estimators will be more. Therefore, the question remains, what M we should use for a given problem and how it can be estimated? More work is needed in that direction.

Acknowledgements

The authors would like to thank Professor G.C. Ray of I.I.T. Kanpur for providing the data sets and two anonymous referees for their constructive suggestions.

References

[1] N.R. Draper, H. Smith, Applied Regression Analysis, Wiley, New York, 1981, pp. 157–160 (Chapter 3).
[2] D. Kundu, Estimating the parameters of the undamped exponential signals, Technometrics 35 (1993) 215–218.
[3] D. Kundu, A. Mitra, On asymptotic properties and confidence intervals for exponential signals, Signal Processing 72 (1999) 129–139.
[4] P. Sircar, M.S. Syali, Complex AM signal for non-stationary signals, Signal Processing 53 (1996) 35–45.