Computational explosion in the frequency estimation of sinusoidal data

Kaimeng Zhang^a, Chi Tim Ng^{1,a}, Myunghwan Na^a

^aDepartment of Statistics, Chonnam National University, Korea

Abstract

This paper highlights the computational explosion issues in the autoregressive moving average approach of frequency estimation of sinusoidal data with a large sample size. A new algorithm is proposed to circumvent the computational explosion difficulty in the conditional least-square estimation method. Notice that sinusoidal pattern can be generated by a non-invertible non-stationary autoregressive moving average (ARMA) model. The computational explosion is shown to be closely related to the non-invertibility of the equivalent ARMA model. Simulation studies illustrate the computational explosion phenomenon and show that the proposed algorithm can efficiently overcome computational explosion difficulty. Real data example of sunspot number is provided to illustrate the application of the proposed algorithm to the time series data exhibiting sinusoidal pattern.

Keywords: ARMA, sinusoidal data, computational explosion

1. Introduction

Time series exhibiting sinusoidal pattern is ubiquitous in digital signal processing, econometrics, agriculture, and biomedicine. For example, Stroebel *et al.* (2010) summarize statistical methods for detecting biological rhythms. Human breathing data collected during radiation therapy are studied in Choi *et al.* (2014). White *et al.* (2011) suggests that the location of the tumor that moves periodically due to the breathing motion during the therapy can be predicted based on autoregressive moving average (ARMA) model. Classical treatment of ARMA model based time series methods is Brockwell and Davis (1991). Unlike the applications in econometrics and agriculture where the period can be naturally considered to be one-year, the frequency and periodicity of certain biological data (e.g., breathing cycle) depend on the physical conditions of the subjects. In such situations, frequency has to be estimated. In addition, the seasonal autoregressive integrated moving average (SARIMA) model requires known periodicity fail to handle the unknown-frequency cases. Periodogram is another important frequency estimation method for the time series data with unknown frequency. However, interpretation of estimation results from periodogram can be difficult when the time series deviates from the sinusoidal pattern, for example, a linear combination of a sinusoidal function and a linear function. In such a situation, non-zero Fourier frequencies are no longer sparse.

The feasibility of analyzing sinusoidal data via ARMA model is studied in Platonod *et al.* (1992). It is illustrated that sine/cosine model perturbed with error terms can be rewritten as an ARMA model with unit complex characteristic roots. After fitting the ARMA model, the frequency can be inferred indirectly through these characteristic roots. In addition, the combined linear trend and sinusoidal

¹ Corresponding author: Department of Statistics, Chonnam National University, 77, Yongbong-ro, Buk-gu, Gwangju 61186, Korea. E-mail: easterlyng@gmail.com

pattern can be described via the multiplicity of complex characteristic roots. In the simulation study, Platonod *et al.* (1992) only consider examples with small sample size $n = 100$. As noted in p.292 of Ozaki (2012), most existing numerical nonlinear optimization procedure of ARMA model estimation face the risk of “computational explosion” due to the violation of invertibility condition of MA parameters during the iteration. Such a computational explosion difficulty becomes more serious in the sinusoidal data because the unit complex roots provide non-invertible MA coefficients. This hinders the applications of ARMA model to biomedicine data exhibiting periodic patterns.

Both stationarity condition and invertibility condition are violated in the sinusoidal data and make the situation more difficult. The maximum likelihood estimate cannot be obtained unless the initial values of the innovation terms are supplied by the user since the stationary distribution is not well-defined. The initial values are unknown, but is necessary as the starting point of the recursive procedure as described in Brockwell and Davis (1991). Without the invertibility condition, it is not guaranteed that the influence of such initial values on the maximum likelihood estimate is negligible.

In this paper, a new algorithm is proposed to estimate the ARMA model violating both stationarity and invertibility conditions. The important idea is that the univariate ARMA model always has multivariate vector state-space representation, see Chapter 12 of Brockwell and Davis (1991). By storing the normalized vector and the logarithm of the norm of the vector only, we are able to circumvent the “computational explosion” difficulties.

The paper is organized as follows. In Section 2, the relationship between sinusoidal data and non-invertible ARMA model is discussed. In Section 3, a procedure allowing estimation under non-invertibility is presented. The computational explosion difficulty is further dealt with in Section 4. Simulation studies are given in Section 5, followed by concluding remarks in Section 6.

2. Preliminaries: ARMA versus sinusoidal pattern

Suppose that the model for the data is generated by

$$y_t = \sum_{k=1}^K A_k \sin 2\pi\omega_k t + \sum_{k=1}^K B_k \cos 2\pi\omega_k t + \epsilon_t,$$

where ϵ_t , $t = 1, 2, \dots, T$ are independent and identically-distributed $N(0, 1)$ random variables. Direct estimation based on the above model is extremely difficult because the local minimum of

$$\sum_{t=1}^T \left(y_t - \sum_{k=1}^K A_k \sin 2\pi\omega_k t - \sum_{k=1}^K B_k \cos 2\pi\omega_k t \right)^2$$

is not unique due to the periodicity of sine/cosine function.

Can the ARMA model be used to handle the sinusoidal pattern? The answer is positive, see Platonod *et al.* (1992). The crucial point is that sine/cosine function is the solution to a recursive relationship. To illustrate some basic ideas, consider a special case where the coefficients are constants of time,

$$y_t = Ar^t \sin 2\pi\omega t + Br^t \cos 2\pi\omega t + \epsilon_t.$$

Here, A, B, r, ω are unknown parameters. Note that $y_t - \epsilon_t = Ar^t \sin 2\pi\omega t + Br^t \cos 2\pi\omega t$ fulfills

$$(y_t - \epsilon_t) = \phi_1(y_{t-1} - \epsilon_{t-1}) + \phi_2(y_{t-2} - \epsilon_{t-2}) \quad (2.1)$$

when the characteristic equation $1 - \phi_1 z - \phi_2 z^2 = 0$ has complex roots $r^{-1} \cos 2\pi\omega \pm r^{-1} \sin 2\pi\omega \sqrt{-1}$. See Brockwell and Davis (1991) for the application of recursive relationships to time series analysis. Transposing terms, Equation (2.1) can be rewritten as

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} + \epsilon_t,$$

where $\theta_1 = \phi_1$ and $\theta_2 = \phi_2$. This is a special case of ARMA model with MA coefficients equaling to the AR coefficients. Then, r and ω can be obtained from the roots of the characteristic equation $1 - \phi_1 z - \phi_2 z^2 = 0$. Given r and ω , the coefficients A and B can further be obtained using the usual regression method and the amplitude is estimated as $(A^2 + B^2)^{1/2}$.

If the true model is

$$y_t = \sum_{k=1}^K A_k r_k^t \sin 2\pi\omega_k t + \sum_{k=1}^K B_k r_k^t \cos 2\pi\omega_k t + \epsilon_t. \quad (2.2)$$

One can consider the following model fitting procedure

Procedure 1. (*Finding frequency via ARMA model*)

1. Fit the ARMA(2K, 2K) model.
2. Find all roots $z_1, \dots, z_K$ (excluding the conjugates $\bar{z}_1, \dots, \bar{z}_K$) of the AR characteristic polynomial.
3. Estimate $\hat{\omega}_K = \text{Arg}(1/z_K)/2\pi$ and $\hat{r}_k = |1/z_k|$.
4. Estimate A_k and B_k , $k = 1, 2, \dots, K$ by regressing y_t against $\hat{r}_k^t \sin 2\pi\hat{\omega}_k t$ and $\hat{r}_k^t \cos 2\pi\hat{\omega}_k t$, $k = 1, 2, \dots, K$.

3. Obtaining $\hat{\epsilon}_t$, $t = 1, 2, \dots$ under non-invertible cases

Consider the model

$$y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q},$$

where ϵ_t , $t = 1, 2, \dots, n$ are independent $N(0, \sigma^2)$ random variables.

Consider the special cases with $p = q = 2K$ and

$$\begin{aligned} y_t - \phi_1 y_{t-1} - \dots - \phi_p y_{t-p} &= \prod_{k=1}^K (1 - 2r_k \cos 2\pi\omega_k L + r_k^2 L^2) y_t \\ &= \prod_{k=1}^K (1 - 2r_k \cos 2\pi\omega_k L + r_k^2 L^2) \epsilon_t \\ &= \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q}, \end{aligned} \quad (3.1)$$

where L is the back-shift operator. Note that y_t exhibits periodic pattern. In such a situation, the MA polynomial has complex roots and the model is non-invertible when $r_k \geq 1$.

Let $\hat{\epsilon}_{-(q-1)} = \hat{\epsilon}_{-(q-2)} = \dots = \hat{\epsilon}_0 = 0$ be the initial values. If the invertibility condition holds, the innovation terms for $t = 1, 2, \dots$ can be obtained recursively as

$$\hat{\epsilon}_t(\boldsymbol{\phi}, \boldsymbol{\theta}) = y_t - \phi_1 y_{t-1} - \dots - \phi_p y_{t-p} + \theta_1 \hat{\epsilon}_{t-1}(\boldsymbol{\phi}, \boldsymbol{\theta}) + \dots + \theta_q \hat{\epsilon}_{t-q}(\boldsymbol{\phi}, \boldsymbol{\theta}). \quad (3.2)$$

However, if the invertibility condition is violated, the influence of the misspecified initial values $\hat{\epsilon}_{-(q-1)} = \hat{\epsilon}_{-(q-2)} = \dots = \hat{\epsilon}_0 = 0$ on $\hat{\epsilon}_t(\phi, \theta)$ is not guaranteed to diminish as $t \rightarrow \infty$. Our solution to the above-mentioned non-invertibility problem is based on the idea that the influences of the unknown true initial values $(\epsilon_{-(q-1)}, \epsilon_{-(q-2)}, \dots, \epsilon_0)$ can be eliminated by regressing $\hat{\epsilon}_t(\phi, \theta)$ against q covariates $\hat{\eta}_t(\phi, \theta) = (\hat{\eta}_{t1}(\phi, \theta), \hat{\eta}_{t2}(\phi, \theta), \dots, \hat{\eta}_{tq}(\phi, \theta))$ defined in Procedure 2. For $i = 1, 2, \dots, q$, the i^{th} component $\hat{\eta}_{ti}(\phi, \theta)$ measures the sensitivity of $\hat{\epsilon}_t(\phi, \theta)$ against unit change in the i^{th} misspecified initial value $\hat{\epsilon}_{t-q}$. Denote the vector of regression coefficients by α . The details are shown in the following procedure and the rigorous justifications are given after the presentation of the procedure.

Procedure 2. (New definition of conditional sum of squares)

1. Initialize $\hat{\epsilon}_{-(q-1)} = \hat{\epsilon}_{-(q-2)} = \dots = \hat{\epsilon}_0 = 0$ and $[\hat{\eta}_{-(q-1)}, \hat{\eta}_{-(q-2)}, \dots, \hat{\eta}_0] = -\mathbf{I}_q$.
2. Compute $\hat{\epsilon}$ and $\hat{\eta}$ via the recursive relationships,

$$\begin{aligned}\hat{\epsilon}_t(\phi, \theta) &= y_t - \phi_1 y_{t-1} - \dots - \phi_p y_{t-p} + \theta_1 \hat{\epsilon}_{t-1}(\phi, \theta) + \dots + \theta_q \hat{\epsilon}_{t-q}(\phi, \theta), \\ \hat{\eta}_t(\phi, \theta) &= \theta_1 \hat{\eta}_{t-1}(\phi, \theta) + \dots + \theta_q \hat{\eta}_{t-q}(\phi, \theta).\end{aligned}$$

3. Minimize the sum of squares

$$Q(\phi, \theta, \alpha) = \sum_{t=1}^T \left(\hat{\epsilon}_t(\phi, \theta) - \alpha^T \hat{\eta}_t(\phi, \theta) \right)^2.$$

Given (ϕ, θ) , the optimal α is

$$\alpha = \left(\sum_{t=1}^T \hat{\eta}_t(\phi, \theta) \hat{\eta}_t^T(\phi, \theta) \right)^{-1} \left(\sum_{t=1}^T \hat{\epsilon}_t(\phi, \theta) \hat{\eta}_t(\phi, \theta) \right).$$

Note that the usual conditional sum of squares method minimizes $\sum_{t=1}^T \hat{\epsilon}_t(\phi, \theta)$ instead.

To illustrate the rationale underlying the procedure, notice that the model can be rewritten as

$$\Theta_L \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_T \end{pmatrix} + \Theta_U \begin{pmatrix} \epsilon_{-(q-1)} \\ \epsilon_{-(q-2)} \\ \vdots \\ \epsilon_0 \end{pmatrix} = \begin{pmatrix} y_1 - \phi_1 y_0 - \dots - \phi_p y_{-(p-1)} \\ y_2 - \phi_1 y_1 - \dots - \phi_p y_{-(p-2)} \\ \vdots \\ y_T - \phi_1 y_{T-1} - \dots - \phi_p y_{T-p} \end{pmatrix},$$

where

$$\Theta_L = \begin{pmatrix} 1 & & & & & & \\ -\theta_1 & 1 & & & & & \\ -\theta_2 & -\theta_1 & 1 & & & & \\ \vdots & \vdots & \vdots & \ddots & & & \\ -\theta_q & -\theta_{q-1} & \dots & -\theta_1 & 1 & & \\ & -\theta_q & -\theta_{q-1} & \dots & -\theta_1 & 1 & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots \\ & & & -\theta_q & -\theta_{q-1} & \dots & -\theta_1 & 1 \end{pmatrix}$$

and

$$\Theta_U = \begin{pmatrix} -\theta_q & -\theta_{q-1} & \cdots & -\theta_1 \\ & -\theta_q & \cdots & -\theta_2 \\ & & \ddots & \vdots \\ & & & -\theta_q \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Equivalently,

$$\Theta_L^{-1} \mathbf{Z} = \Theta_L^{-1} \begin{pmatrix} y_1 - \phi_1 y_0 - \cdots - \phi_p y_{-(p-1)} \\ y_2 - \phi_1 y_1 - \cdots - \phi_p y_{-(p-2)} \\ \vdots \\ y_T - \phi_1 y_{T-1} - \cdots - \phi_p y_{T-p} \end{pmatrix} = \Theta_L^{-1} \Theta_U \begin{pmatrix} \epsilon_{-(q-1)} \\ \epsilon_{-(q-2)} \\ \vdots \\ \epsilon_0 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_T \end{pmatrix}.$$

One can consider this as a regression model with a left-hand-side response, the coefficients are the initial values $(\epsilon_{-(q-1)}, \epsilon_{-(q-2)}, \dots, \epsilon_0)$, and the design matrix is $\Theta_L^{-1} \Theta_U$. It is possible to eliminate the influence of the initial guesses via the above regression model. The computation of the quantities $\Theta_L^{-1} \mathbf{Z}$ and $\Theta_L^{-1} \Theta_U$ can be simplified by noting that

$$[\Theta_U, \Theta_L] \begin{pmatrix} \mathbf{0} \\ \Theta_L^{-1} \mathbf{Z} \end{pmatrix} = \mathbf{Z} \quad \text{and} \quad [\Theta, \Theta_L] \begin{pmatrix} -\mathbf{I}_q \\ \Theta_L^{-1} \Theta_U \end{pmatrix} = \mathbf{0}. \quad (3.3)$$

Set $\hat{\epsilon}_{-(q-1)} = \hat{\epsilon}_{-(q-2)} = \cdots = \hat{\epsilon}_0 = 0$ and $[\hat{\eta}_{-(q-1)}, \hat{\eta}_{-(q-2)}, \dots, \hat{\eta}_0] = -\mathbf{I}_q$. Denote the elements of the response vector $\Theta_L^{-1} \mathbf{Z}$ by $\hat{\epsilon}_t(\phi, \theta)$ and the rows of the design matrix by $\hat{\eta}_t(\phi, \theta)$. Equations (3.3) is equivalent to the recursive formulas in Step 2 of the procedure since all rows in $[\Theta_U, \Theta_L]$ are the same excepting zeros entries. The conditional sum of squares defined in Step 3 can be minimized by using any suitable numerical optimization routines. In the simulation studies presented in Section 5, this is performed using the “unconstrained optimization by quadratic approximation (UObyQA)” algorithm in Powell (2002). Note that in invertible cases, $\hat{\eta}_t(\phi, \theta)$ has the exponentially decaying pattern. However, this is not true in non-invertible cases. Neglecting such terms gives the usual maximum conditional likelihood estimation method. They are not negligible under non-invertible cases.

4. Computational explosion issues

To avoid the difficulties related to the exploding $\hat{\epsilon}_t(\phi, \theta)$ and $\hat{\eta}_t(\phi, \theta)$, consider an algorithm that stores only the logarithms of the conditional sum of squares $Q(\phi, \theta, \alpha)$. Denote

$$E_t = (1, \hat{\epsilon}_t(\phi, \theta), \dots, \hat{\epsilon}_{t-q+1}(\phi, \theta), \hat{\eta}_t(\phi, \theta), \dots, \hat{\eta}_{t-q+1}(\phi, \theta))$$

and $\mathbf{e}_t = \mathbf{E}_t / \|\mathbf{E}_t\|_2$ be the normalized vector. Let

$$\mathbf{B}_t = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & \mathbf{0}_q^T & \mathbf{0}_q^T & \cdots & \mathbf{0}_q^T \\ a_t & \theta_1 & \theta_2 & \cdots & \theta_q & \mathbf{0}_q^T & \mathbf{0}_q^T & \cdots & \mathbf{0}_q^T \\ 0 & 1 & 0 & \cdots & 0 & \mathbf{0}_q^T & \mathbf{0}_q^T & \cdots & \mathbf{0}_q^T \\ 0 & 0 & \ddots & \cdots & 0 & \mathbf{0}_q^T & \mathbf{0}_q^T & \cdots & \mathbf{0}_q^T \\ 0 & 0 & 0 & 1 & 0 & \mathbf{0}_q^T & \mathbf{0}_q^T & \cdots & \mathbf{0}_q^T \\ \mathbf{0}_q & \mathbf{0}_q & \mathbf{0}_q & \cdots & \mathbf{0}_q & \theta_1 \mathbf{I}_q & \theta_2 \mathbf{I}_q & \cdots & \theta_q \mathbf{I}_q \\ \mathbf{0}_q & \mathbf{0}_q & \mathbf{0}_q & \cdots & \mathbf{0}_q & \mathbf{I}_q & \mathbf{0}_q & \cdots & \mathbf{0}_q \\ \mathbf{0}_q & \mathbf{0}_q & \mathbf{0}_q & \cdots & \mathbf{0}_q & \mathbf{0}_q & \ddots & \cdots & \mathbf{0}_q \\ \mathbf{0}_q & \mathbf{0}_q & \mathbf{0}_q & \cdots & \mathbf{0}_q & \mathbf{0}_q & \cdots & \mathbf{I}_q & \mathbf{0}_q \end{pmatrix}$$

and

$$a_t = y_t - \phi_1 y_{t-1} - \cdots - \phi_p y_{t-p}.$$

Procedure 3. (Algorithm storing the logarithm of conditional sum of squares)

1. Set $g_0 = 0$ and $\mathbf{S}_0 = 0$.
2. Define recursively for $t = 1, 2, \dots, T$

$$\begin{aligned} c_t &= \|\mathbf{B}_t \mathbf{e}_{t-1}\|_2, \\ \mathbf{e}_t &= \frac{\mathbf{B}_t \mathbf{e}_{t-1}}{c_t}, \\ g_t &= g_{t-1} + \log c_t^2, \\ \mathbf{S}_t &= \mathbf{e}_t \mathbf{e}_t' + \frac{\mathbf{S}_{t-1}}{c_t^2}. \end{aligned}$$

3. Compute $\log Q(\boldsymbol{\phi}, \boldsymbol{\theta}, \boldsymbol{\alpha}) = g_T + \log \mathbf{b}^T \mathbf{S}_T \mathbf{b}$. Here,

$$\mathbf{b}^T = (0, 1, 0, \dots, 0, -\boldsymbol{\alpha}^T, \mathbf{0}_q^T, \dots, \mathbf{0}_q^T)$$

and $\boldsymbol{\alpha}$ can be obtained as $\mathbf{S}_{[3:(p+q+2), 3:(p+q+2)]}^{-1} \mathbf{S}_{[3:(p+q+2), 2]}$.

The equivalence between Procedures 2 and 3 can be verified easily by mathematical induction arguments.

5. Simulation study

The simulation study illustrates: (i) the computational explosion phenomenon, (ii) the necessity of introducing $\hat{\boldsymbol{\eta}}_t$ in Procedure 2, and (iii) the feasibility of frequency estimation. The proposed methods are implemented in CRAN R language. The optimization is carried out using the R package “powell” that implements the “UObYQA” algorithm in Powell (2002).

5.1. Computational explosion phenomenon

Consider the frequency representations of the ARMA model like Equation (2.2)

$$y_t = \sum_{k=1}^K A_k r_k^t \sin 2\pi\omega_k t + \sum_{k=1}^K B_k r_k^t \cos 2\pi\omega_k t + 0.5\epsilon_t,$$

with the following three settings,

1. Setting 1: $p = q = 2K = 4$, $r_1 = r_2 = 1$, $(A_1, A_2) = (2.5, -0.8)$, $(B_1, B_2) = (1.3, -0.5)$, and $2\pi\omega = (0.25, 0.5)$.
2. Setting 2: $p = q = 2K = 6$, $r_1 = r_2 = r_3 = 1$, $(A_1, A_2, A_3) = (2.5, -0.8, 1.2)$, $(B_1, B_2, B_3) = (1.3, -0.5, 0.3)$, and $2\pi\omega = (0.25, 0.5, 0.75)$.

The time series are then standardized so that the standard deviations always equal to one. The coefficients $\phi^* = \theta^*$ are obtained from the expansion

$$1 - \phi_1 L - \dots - \phi_p L^{t-p} = \prod_{k=1}^K (1 - 2r_k \cos 2\pi\omega_k L + r_k^2 L^2).$$

To illustrate the computational explosion phenomenon, the critical sample sizes for the stack overflow are investigated. The programs are implemented in CRAN R language. Procedure 2 requires the storage of the values of sums of squares Q . There are stack-overflow-related errors if Q is greater than 10^{308} since double precision is used in R to store numerical values. The experiment is conducted as follows. Perturb each component of the coefficients ϕ^* and θ^* by $N(0, 0.03^2)$ random variables. For each $T = 10, 20, 30, \dots, 2500$, the conditional sum of squares Q is evaluated at such perturbed coefficients using Procedures 2 and 3. The experiment is repeated for 100 times. Different perturbed coefficients are used in different trials. For each T , if there is no stack-overflow-related error in all 100 trials, the average $\log(Q)$ value is reported. For Procedure 2, the results of Settings 1 and 2 are presented in Figure 1, panels (a) and (b) respectively. Under Setting 1, computation of average $\log(Q)$ results in stack overflow problem after $T \approx 750$ and under Setting 2, the critical sample size of the explosion is ≈ 600 . Before the stack overflow occurs, average $\log(Q)$ is close to 400, or equivalently, the value of Q exceeds 10^{308} , the limit of double precision. This means that Procedure 2 alone can result in computational explosion problem. Therefore, Procedure 3 is necessary. For Procedure 3, the results of average $\log(Q)$ under Settings 1 and 2 are reported in Figure 1, panels (c) and (d). Procedure 3 results in no computational explosion problem.

5.2. Necessity of introducing $\hat{\eta}_t$ in Procedures 2 and 3

To see the necessity of introducing $\hat{\eta}_t$, Procedures 2 and 3 based on $Q(\phi, \theta, \alpha) = \sum_{t=1}^T (\hat{\epsilon}_t(\phi, \theta) - \alpha' \hat{\eta}_t(\phi, \theta))^2$ is compared to the usual conditional sum of squares $\sum_{t=1}^T \hat{\epsilon}_t(\phi, \theta)$. In the simulation, Procedure 3 is implemented in CRAN R language and conditional sum of squares based estimation is performed using the arima function provided by R. Consider the following three settings

1. Setting 1: $p = q = 2K = 4$, $r_1 = r_2 = 1$, $(A_1, A_2) = (2.5, -0.8)$, $(B_1, B_2) = (1.3, -0.5)$, and $2\pi\omega = (0.25, 0.5)$.
2. Setting 2: $p = q = 2K = 6$, $r_1 = r_2 = r_3 = 1$, $(A_1, A_2, A_3) = (2.5, -0.8, 1.2)$, $(B_1, B_2, B_3) = (1.3, -0.5, 0.3)$, and $2\pi\omega = (0.25, 0.5, 0.75)$.

Figure 1: The log sum of squares values against sample size T . (a) $p = 4$ case and (b) $p = 6$ case with sum of squares computed using Procedure 2. (c) $p = 4$ case and (d) $p = 6$ case with sum of squares computed using Procedure 3.

3. Setting 3: $p = q = 2K = 8$, $r_1 = r_2 = r_3 = r_4 = 1$, $(A_1, A_2, A_3, A_4) = (2.5, -0.8, 1.2, 0.7)$, $(B_1, B_2, B_3, B_4) = (1.3, -0.5, 0.3, -0.5)$, and $2\pi\omega = (0.25, 0.5, 0.75, 0.9)$.

For each setting, the experiment is repeated for $N = 100$ times with different sample sizes $T = 100, 300, 600, 1500$. The performance is evaluated via the 2% trimmed relative mean square error (RMSE) of the coefficients ϕ, θ ,

$$\text{RMSE}_\phi = \left(\frac{1}{0.8N} \sum_{n=1}^{0.8N} \left\| \frac{\phi^{(n)} - \phi^*}{\phi^*} \right\|_2 \right)^{\frac{1}{2}} \quad \text{and} \quad \text{RMSE}_\theta = \left(\frac{1}{0.8N} \sum_{n=1}^{0.8N} \left\| \frac{\theta^{(n)} - \theta^*}{\theta^*} \right\|_2 \right)^{\frac{1}{2}}.$$

Table 1: The RMSE of coefficients (ϕ , θ) of ARMA and ARMAeta models

p	T	ARMA		ARMAeta	
		RMSE $_{\phi}$	RMSE $_{\theta}$	RMSE $_{\phi}$	RMSE $_{\theta}$
4	100	1.2098	0.9529	0.0714	0.0572
	300	1.2292	0.9467	0.0413	0.0350
	600	1.2227	0.9464	0.0418	0.0286
	1500	1.1796	0.9663	0.0390	0.0495
6	100	1.0972	0.9697	0.0297	0.0038
	300	1.1268	0.9651	0.0183	0.0027
	600	1.1357	0.9638	0.0342	0.0029
	1500	1.1204	0.9624	0.0609	0.0024
8	100	1.0189	1.0308	0.0035	0.0019
	300	1.0055	1.0227	0.0055	0.0014
	600	1.0614	0.9825	0.0090	0.0015
	1500	1.1555	0.9217	0.0095	0.0014

ARMA is obtained by arima function in R and ARMAeta is obtained by Procedure 3. RMSE = relative mean square error.

Here, $\phi^{(n)}$ and $\theta^{(n)}$ refers to the estimates obtained in the trial with n^{th} smallest mean relative errors and ϕ^* and θ^* refer to the true parameters. The coefficients $\phi^* = \theta^*$ are obtained from the expansion

$$1 - \phi_1 L - \dots - \phi_p L^{p-p} = \prod_{k=1}^K (1 - 2r_k \cos 2\pi\omega_k L + r_k^2 L^2).$$

Table 1 shows the results of the RMSE. It can be seen that ARMA models fitted by R function arima tend to have greater RMSE than those fitted by Procedure 3, especially for large p cases with $p = 6, 8$.

5.3. Feasibility of frequency estimation

The same settings as in Section 5.2 are considered. To study the performance of the frequency estimation under ARMAeta model, $2\pi\omega$ is estimated via Procedure 1. Step (1) of Procedure 1 is carried out by minimizing the sum of squares in Procedure 3 using UObyQA optimization algorithm. The computational explosion difficulty is avoided since the sum of squares is obtained from Procedure 3 and the performance is evaluated via the 2% trimmed RMSE of $2\pi\omega$,

$$\text{RMSE}_{\omega k} = \left(\frac{1}{0.8N} \sum_{n=1}^{0.8N} \|2\pi\omega_k^{(n)} - 2\pi\omega_k^*\|_2 \right)^{\frac{1}{2}},$$

where the ω_k^* is the true value and $\omega_k^{(n)}$ is the estimated value with n^{th} smallest mean relative errors.

Table 2 shows the results of RMSE of $\omega_1, \dots, \omega_k$. It can be see that the RMSE increases as k , T , and p increase.

6. Real data analysis

Consider the hysteresis of the solar cycles, that can be measured by the number of sunspots, see Dmitriev *et al.* (2002) and Suyal *et al.* (2012). We use the monthly Wolf's sunspot number data from 1749 to March 2018, with 3230 observations in total, see Figure 2. The data can be obtained from the website of Royal Observatory of Belgium; similar data is also used in the literature of time series analysis, for example, Li *et al.* (2012) and Tong (1990).

Table 2: The RMSE of coefficients $2\pi\omega$

p	T	ARMAeta			
		$\text{RMSE}_{\omega 1}$	$\text{RMSE}_{\omega 2}$	$\text{RMSE}_{\omega 3}$	$\text{RMSE}_{\omega 4}$
4	100	0.0079	0.0149		
	300	0.0096	0.0199		
	600	0.0639	0.0416		
	1500	0.1019	0.0613		
6	100	0.0723	0.0688	0.0498	
	300	0.1605	0.1197	0.1184	
	600	0.1687	0.1152	0.0850	
	1500	0.1836	0.1409	0.1101	
8	100	0.1724	0.0836	0.8350	0.0656
	300	0.2106	0.1068	0.1007	0.0737
	600	0.2078	0.1068	0.0989	0.0695
	1500	0.2104	0.1119	0.0969	0.0617

ARMAeta is fitted by Procedure 3 and $2\pi\omega$ is obtained by Procedure 1. RMSE = relative mean square error.

Figure 2: The monthly sunspot number data.

Table 3: The BIC values of ARMAeta models for the sunspot data

Models	ARMAeta(2, 2)	ARMAeta(4, 4)	ARMAeta(6, 6)	ARMAeta(8, 8)
BIC	55112.43	55498.78	69445.98	176840.6

BIC = Bayesian information criterion.

The ARMAeta is applied for the sunspot data and the optimal model is selected by minimizing the Bayesian information criterion (BIC). For ARMAeta(p, p) model, including α , there are a total of $3 \times p$ parameters. The BICs can be defined as

$$\text{BIC} = n * \ln(L) + 3p \ln(n),$$

where $\ln(L)$ is the fitted sum of squares of the ARMAeta model.

Table 3 shows the BIC values when $p = 2, 4, 6, 8$ and the coefficients (ϕ, θ, ω) of each model are shown in Table 4. From Table 3, the ARMAeta(2, 2) model is chosen because it has the smallest BIC. Table 4 indicates the estimated frequency is $2\pi\omega = 0.0511$ or $\omega \approx 123$. This suggests that there are $500/123 \approx 4$ cycles per 500 units of time and agrees with the pattern shown in Figure 2.

Table 4: The coefficients (ϕ , θ , ω) of ARMAeta models for the sunspot data

		p							
		1	2	3	4	5	6	7	8
ARMAeta(2, 2)	ϕ_p	-1.9967	0.9993						
	θ_p	-1.9986	1.0019						
	ω_k	0.0511							
ARMAeta(4, 4)	ϕ_p	-3.8922	5.7861	-3.8929	1.0004				
	θ_p	-3.8927	5.7862	-3.8929	1.0001				
	ω_k	0.0565	0.3249						
ARMAeta(6, 6)	ϕ_p	-5.6313	13.5533	-17.8440	13.6633	-5.6313	0.9990		
	θ_p	-5.6313	13.5533	-17.8439	13.5533	-5.6312	0.9990		
	ω_k	2.051×10^{-16}	0.3334	0.5162					
ARMAeta(8, 8)	ϕ_p	-7.0733	22.6740	-43.0201	52.8390	-43.0201	22.6739	-7.0733	1
	θ_p	-7.0733	22.6740	-43.0201	52.8390	-43.0201	22.6739	-7.0733	1
	ω_k	0.0498	0.3247	0.5174	0.7655				

7. Conclusions

The estimation problem of ARMA model is discussed under settings where the invertibility is not fulfilled. New algorithms are proposed to deal with non-invertibility difficulty and computational explosion difficulty. The methods developed are useful for the sinusoidal time series data where the frequency is unknown. This facilitates the application of the ARMA method to biomedicine and meteorology such as breathing motion data obtained during radiation therapy.

Acknowledgments

Chi Tim, Ng's work is supported by National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. NRF-2017R1C1B2011652).

References

- Brockwell PJ and Davis RA (1991). *Time Series: Theory and Method*, Springer, New York.
- Choi S, Chang Y, Kim N, Park S, Song S, and Kang H (2014). Performance enhancement of respiratory tumor motion prediction using adaptive support vector regression: Comparison with adaptive neural network method, *International Journal of Imaging Systems and Technology*, **24**, 8–15.
- Dmitriev AV, Suvorova AV, and Veselovsky IS (2002). Expected hysteresis of the 23-rd solar cycle in the heliosphere, *Advances in Space Research*, **29**, 475–479.
- Li GD, Guan B, Li WK, and Yu PLH (2012). Hysteretic autoregressive time series models, *Biometrika*, **99**, 1–8.
- Ozaki T (2012). *Time Series Modeling of Neuroscience Data*, CRC Press.
- Platonod AA, Gajo ZK, and Szabatin J (1992). General method for sinusoidal frequencies estimation using ARMA algorithms with nonlinear prediction error transformation, *IEEE*, **5**, 441–444.
- Powell MJD (2002). UOBYQA: unconstrained optimization by quadratic approximation, *Mathematical Programming, Series B*, **92**, 555–582.
- Stroebe AM, Bergner M, Reulbach U, Biermann T, Groemer TW, Klein I, and Kornhuber J (2010). Statistical methods for detecting and comparing periodic data and their application to the nycthemeral rhythm of bodily harm: a population based study, *Journal of Circadian Rhythms*, **8**, 1–10.

- Suyal V, Prasad A, and Singh HP (2012). Hysteresis in a solar activity cycle, *Solar Physics*, **276**, 407–414.
- Tong G (1990). *A Dynamical System Approach*, Oxford University Press, Oxford.
- White BM, Low DA, Zhao T, *et al.* (2011). Investigation of a breathing surrogate prediction algorithm for prospective pulmonary gating, *Medical Physics*, **38**, 1587–1595.

Received April 20, 2018; Revised June 1, 2018; Accepted June 8, 2018