Frequency-Domain Identification of Continuous-Time ARMA Models from Sampled Data

Jonas Gillberg, Lennart Ljung
Division of Automatic Control
E-mail: gillberg@isy.liu.se, ljung@isy.liu.se

20th December 2010

Report no.: LiTH-ISY-R-2983
Accepted for publication in Automatica, Vol 45, pp 1371-1378, 2009.

Address:
Department of Electrical Engineering
Linköpings universitet
SE-581 83 Linköping, Sweden

WWW: http://www.control.isy.liu.se
Abstract
The subject of this paper is the direct identification of continuous-time autoregressive moving average (CARMA) models. The topic is viewed from the frequency domain perspective which then turns the reconstruction of the continuous-time power spectral density (CT-PSD) into a key issue. The first part of the paper therefore concerns the approximate estimation of the CT-PSD from uniformly sampled data under the assumption that the model has a certain relative degree. The approach has its point of origin in the frequency domain Whittle likelihood estimator. The discrete- or continuous-time spectral densities are estimated from equidistant samples of the output. For low sampling rates the discrete-time spectral density is modeled directly by its continuous-time spectral density using the Poisson summation formula. In the case of rapid sampling the continuous-time spectral density is estimated directly by modifying its discrete-time counterpart.

Keywords: System Identification
Frequency-Domain Identification of Continuous-Time ARMA Models from Sampled Data

Jonas Gillberg a Lennart Ljung b

a IPCOSAptitude Ltd
   Carthouse III, Copley Hill Farm Business Park, Cambridge Rd, Babraham
   Cambridge, CB22 3GN, United Kingdom
   Email: jonas.gillberg@ipcosaptitude.com

b Department of Electrical Engineering, Linköping University
   SE-581 83 Linköping, Sweden
   Email: ljung@isy.liu.se

Abstract

The subject of this paper is the direct identification of continuous-time autoregressive moving average (CARMA) models. The topic is viewed from the frequency domain perspective which then turns the reconstruction of the continuous-time power spectral density (CT-PSD) into a key issue. The first part of the paper therefore concerns the approximate estimation of the CT-PSD from uniformly sampled data under the assumption that the model has a certain relative degree. The approach has its point of origin in the frequency domain Whittle likelihood estimator. The discrete- or continuous-time spectral densities are estimated from equidistant samples of the output. For low sampling rates the discrete-time spectral density is modeled directly by its continuous-time spectral density using the Poisson summation formula. In the case of rapid sampling the continuous-time spectral density is estimated directly by modifying its discrete-time counterpart. Copyright ©2008 IFAC

Key words: System identification; Time-series analysis; Frequency domains; Continuous time systems; ARMA parameter estimation

1 Introduction

The topic of this paper is a novel method for the identification of continuous-time autoregressive moving average (ARMA) models. The method works on uniformly sampled data and operates primarily in the frequency domain. Its main advantage is that it is direct, i.e. does not make use of the sampled version of the continuous-time system. It is also proved that the method is equivalent to interpolating an estimate of the covariance function.

Parameter estimation of continuous-time systems is by no means a novel branch of system identification. See for example the survey article by Unbehauen and Rao, [25], the monograph by Sinha and Rao [20] or the PhD thesis by Mensler [11]. Research on identification of continuous-time models has primarily been concentrated to the time-domain, with approaches such as:

Poisson moment functionals, integrated sampling, orthogonal functions. A few authors on the other hand, have tackled the problem in the frequency domain. An early reference is by Shinbrot [19] followed by the Fourier modulating function approach introduced by Pearson et al. [14]. Frequency-domain analysis has also been the starting point for the work on input-output and noise models done by Pintelon et al. [15].

Recently there has been a renewed interest in continuous-time system identification. In particular continuous-time noise models[17],[10]. See for instance the articles by Larsson and Söderström on continuous-time AR [9] and ARMA [8] parameter estimation. The work on hybrid Box-Jenkins and ARMAX modeling by Pintelon et.al [16] also concerns this problem.

The central result in this paper is a method for the estimation of the continuous-time power spectral density

\[ \hat{\Phi}_c(i\omega) = F_{\ell+1,T_s}(i\omega)\hat{\Phi}_d(e^{i\omega T_s}) \]
by spectrally weighting the discrete-time periodogram. An optimal weighting would be

\[ F_{c,T_c}(i\omega) = \frac{\Phi_c(i\omega, \theta_0)}{\Phi_d(e^{i\omega T_c}, \theta_0)} \]

where the discrete-time power spectral density is computed using a version of the Poisson summation formula [13]

\[ \Phi_d(e^{i\omega T_c}, \theta_0) = \sum_{k=-\infty}^{\infty} \Phi_c(i\omega + \omega_s k, \theta_0) \tag{1} \]

where \( \omega_s = \frac{2\pi}{T_c} \) is the sampling frequency. Unfortunately, the optimal weighting depends on the true parameters of the time series model, which of course are unknown during identification. The solution presented in this paper to the above dilemma is to replace the true model by a chain of integrations such that

\[ \Phi_c(i\omega) \approx \frac{1}{|i\omega|^{2\ell}}, \]

where \( \ell = n - m \) is the relative degree or pole excess. This approximation can then be used in the summation formula (1) and will yield another weighting factor

\[ F_{c,T_c}(i\omega) = \frac{e^{i\omega T_c} - 1}{\omega T_c} \frac{\Pi_{2\ell-1}(e^{i\omega T_c})}{(2\ell - 1)!} \]

such that the spectrum can be estimated as

\[ \hat{\Phi}_{c,T}(i\omega) = F_{c,T_c}(i\omega) \hat{\Phi}_{d,N}(i\omega). \]

Here, the weighting only depends on \( \ell \) and the so called Euler-Frobenius polynomials \( \Pi_{2\ell-1}(z) \).

At the end of the paper, it will be shown that this type of spectral weighting is equivalent to interpolating an estimate of the covariance function \( \{\tilde{r}(kT_s)\}_{k=0}^{N_t-1} \) by cardinal B-splines. It will also be proved that this setup can be interpreted as time-domain construction of the continuous-time measurements \( \{y(t)\}_{t=0}^{T_c} \) with a spline based kernel \( G_{T_c}(t) \) such that

\[ \tilde{y}_c(t) = \sum_{k=0}^{N_t-1} y(kT_s) G_{T_c}(t-kT_s). \]

After the continuous-time periodogram is calculated, the parameters can then be estimated using the continuous-time Whittle likelihood approach

\[ \hat{\theta} = \arg \min_{\theta} \sum_{k=1}^{N_t} \frac{\hat{\Phi}_{c,T}(i\omega k)}{\hat{\Phi}_c(i\omega k, \theta)} + \Phi_c(i\omega k, \theta). \]

where there is no need to use the discrete-time power spectral density. All the ideas in this paper should be seen as a direct continuation of the groundbreaking work on CARMA estimation initiated by Erik K. Larsson and Torsten Söderström [9, 8].

2 Outline

First, the model structure for continuous-time series models used in the paper is introduced in Section 3. Then, in Section 4, a method for the indirect frequency-domain estimation of a continuous-time ARMA model is presented, in order to familiarize the reader with the subject. Then, the procedure for the direct estimation of the continuous-time model is introduced in Section 5. As explained before, this approach requires the estimation of the continuous-spectrum, and a method for this is explained in Section 6. Section 8 then contains a brief introduction to cardinal B-splines and their time- and frequency domain properties. Finally, in Section 9 it is shown that for uniformly sampled data this method is equivalent to interpolation of the covariance function by polynomial splines.

3 Model and Representations

The content of this paper will focus on frequency-domain identification of continuous-time autoregressive moving average (CARMA) models. In this context, observed data \( \{y(t)\}_{t=0}^{T_c} \) or \( \{y(kT_s)\}_{k=0}^{N_t} \) is modeled as Gaussian noise passed through a continuous-time dynamical system. The system is parameterized as

\[ y(t) = H_c(s, \theta) e(t, \theta) \]

where

\[ H_c(s, \theta) = \frac{B(s, \theta)}{A(s, \theta)} \] \tag{2}
where $c(t)$ is used informally to denote continuous-time white noise such that

$$Ee(t, \theta) = 0$$

$$E\{e(t, \theta)e(s, \theta)\} = \sigma^2 \delta(t - s).$$

The numerator and the denominator in the transfer function are

$$A(s, \theta) = s^n + a_1 s^{n-1} + a_2 s^{n-2} + \cdots + a_n$$

$$B(s, \theta) = s^m + b_1 s^{m-1} + \cdots + b_m$$

and the vector of parameters is defined as $\theta = [a_1 \ a_2 \ \ldots \ a_n \ b_1 \ b_2 \ \ldots \ b_m \ \lambda]^T$ where $\lambda = \sigma^2$ is the variance of the driving noise. Here $m < n$, so the system is strictly proper and $Ey^2(t) < \infty$. A schematic view of the setup is illustrated in Figure 1. This type of configuration could also be interpreted such that the true output would have the spectrum

$$\Phi(i\omega, \theta) = \lambda \left| \frac{B(i\omega, \theta)}{A(i\omega, \theta)} \right|^2.$$

It should be noted that securing identifiability for continuous-time models which are inferred from discrete-time data is a delicate matter. The transformation from continuous- to discrete-time could possibly destroy information and make two different continuous-time models seem identical. In this paper the authors will assume identifiability of the models and the model structure used for identification is the same as that of the true system. However, for the interested reader we refer to discussion in [3] and [12], together with the method for determining identifiability which is found in [24].

4 Indirect estimation

A straightforward way to do frequency domain identification of a continuous time series model from uniformly sampled data would be to compute the discrete-time periodogram from sampled data such that

$$\hat{\Phi}_d(e^{i\omega T_s}) = |Y_d(e^{i\omega T_s})|^2$$

where

$$Y_d(e^{i\omega T_s}) = \frac{1}{N_f} \sum_{k=1}^{N_f} y(kT_s)e^{-i\omega kT_s}.$$ 

The parameters of (2) would then be estimated using the discrete-time Whittle likelihood procedure

$$\hat{\theta} = \arg \min_{\theta} \sum_{k=1}^{N_f} \frac{\hat{\Phi}_d(e^{i\omega T_s}, \theta)}{\Phi_d(e^{i\omega T_s}, \theta)} + \log \Phi_d(e^{i\omega T_s}, \theta) \quad (4)$$

where the discrete-time spectrum is expressed in terms of continuous-time parameters by the relationship

$$\Phi_d(e^{i\omega T_s}, \theta) = \sum_{k=-\infty}^{\infty} \Phi_c(i\omega + i\omega k, \theta) \quad (5)$$

which is the well known Poisson summation formula. A drawback connected to using the formula in (5) is of course the infinite sum. An approximations can however be achieved with a limited number of terms

$$\Phi_d(e^{i\omega T_s}) = \sum_{k=-N_f}^{N_f} \Phi_c(i\omega + i\omega k) \quad (6)$$

when the continuous-time system is strictly proper. The idea of truncating the Poisson summation formula can now be used to illustrate the detrimental effect that comes with the exclusion of higher order terms. For more efficient ways of approximating the summation we refer the readers to [2] and [23].

4.1 Numerical Example

In Figure 2 and Table 4.1, we have estimated the second-order continuous-time auto regressive (AR) model

$$y(t) = H_c(s)\hat{e}(t)$$

$$H_c(s) = \frac{1}{s^2 + a_1 s + a_2} \quad (7)$$

with the true parameters $\sigma = 1$, $a_1 = 2$ and $a_2 = 2$. The duration of the data set was $T = 1000s$ with the sampling interval $T_s = 1s$. Estimates were produced using $N_{MC} = 250$ Monte-Carlo simulations and the information in the table and figure is based on the mean of those values. The method that has been employed is that of (4) where the discrete time spectrum has been approximated using (6).

Figure 2 illustrates the frequency-domain bias which could occur in the transfer function if only the central term ($N_f = 0$) in (6) have been used, which means that we have assumed that

$$\Phi_d(e^{i\omega T_s}, \theta) = \Phi_c(i\omega, \theta).$$

This will produce a biased estimate which is illustrated by the dashed-dotted line in the Bode diagram for the system. A set of systems with virtually no bias has also been estimated using $N_f = 5$ in (6). The mean value of these are also illustrated in the figure as a dotted line which is almost identical to the true system which is represented by the solid line.
Fourier transform of this signal system in (3). Then it would be possible to compute the identical to the solid curve of the true system. The system has been used. When $N_f = 0$ the estimated system will be biased. For the case $T$ is almost entirely remove this.

Assume for a moment that it would be possible to measure the true continuous-time output $\{\tilde{y}(t)\}_{t=0}^{T}$ of the system in (3). Then it would be possible to compute the Fourier transform of this signal

$$\hat{Y}(i\omega) = \frac{1}{T} \int_{0}^{T} \tilde{y}(t)e^{-i\omega t} dt,$$

and its periodogram

$$\hat{\Phi}(i\omega) = \left| \hat{Y}(i\omega) \right|^2 = \frac{1}{T} \int_{0}^{T} \tilde{y}(t)e^{-i\omega t} dt$$

would be distributed as

$$\hat{\Phi}(i\omega_k) = \left| Y_T(i\omega_k) \right|^2 \sim \text{AsExp} \Phi(i\omega_k, \theta_0).$$

The periodogram would also be asymptotically independent at the frequencies $\{\omega_k\}_{k=1}^{N_\omega}$ where

$$\omega_k = \frac{2\pi k}{T}$$

if $T$ is the time of observation of the output. Here, it is assumed that $T$ is so large so that the asymptotic expression can be considered valid.

When an estimate of the continuous-time power spectrum is available, a model can be identified by the following Maximum-Likelihood (ML) procedure described in the papers by Whittle [30] or Dzhaparidze [6]

$$L(\theta, \hat{\Phi}) = \frac{N_\omega}{k=1} \hat{\Phi}(i\omega_k) + \log \Phi(i\omega_k, \theta)$$

$$\hat{\theta} = \arg \min_{\theta} L(\theta, \hat{\Phi}).$$

The main problem here is that the discrete-time periodogram can be readily found from the uniformly sampled data while we need the continuous-time periodogram to use the method described above. A natural question is then to ask if it is possible to estimate the latter from the former? The answer to this question is yes, and the next section will explain how.

6 Estimating the Continuous-Time Spectrum

In this section, we will be looking for ways to approximate the continuous-time power spectral density. In particular expression such as

$$\hat{\Phi}_{c,T}(i\omega) = F_{2\pi T}^{-1}(i\omega) \hat{\Phi}_{d,N_\omega}(e^{i\omega T})$$

with a suitable function $F_{2\pi T}^{-1}(i\omega)$. In order to find such a function, we could argue as follows: if the true parameters $\theta_0$ of (3) were known, we would have

$$E\hat{\Phi}_{d,N_\omega}(e^{i\omega T}) = \Phi_d(e^{i\omega T}, \theta_0)$$

and

$$E\hat{\Phi}_{c,T}(i\omega) = \Phi_c(i\omega, \theta_0).$$

Table 1

<table>
<thead>
<tr>
<th>Method</th>
<th>$a_1 = 3$</th>
<th>$a_2 = 2$</th>
<th>$\sigma = 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_f = 0$</td>
<td>5.090122</td>
<td>3.217543</td>
<td>1.675111</td>
</tr>
<tr>
<td>$N_f = 5$</td>
<td>3.077730</td>
<td>2.044676</td>
<td>1.020823</td>
</tr>
</tbody>
</table>

In Table 1 the mean parameter values are also illustrated for $N_f = 0$ and $N_f = 5$ in (6). From the table we see that assuming that the discrete-time spectrum is equal to the continuous-time spectrum will also produce bias in the parameters. However, including $N_f = 5$ terms in (6) will almost entirely remove this.

Fig. 2. Bode diagram comparing the Whittle likelihood estimator with $N_f = 0$ (dashdot) and $N_f = 5$ (dashed) in (6) to the true system (solid). The system is $H(c) = \frac{\sigma}{\pi + a_1 + a_2}$ where $\sigma = 1$, $a_1 = 2$ and $a_2 = 2$. The sampling interval is $T_s = 0.8$ and frequencies up to the Nyquist frequencies have been used. When $N_f = 0$ the estimated system will be biased. For the case $N_f = 5$ the dashed curve is almost identical to the solid curve of the true system.
This means that the ideal spectral weighting in (8) would be characterized by

\[ F_{2\ell,T_s}(i\omega) = \frac{\Phi_c(i\omega, \theta_0)}{\Phi_d(e^{i\omega T_s}, \theta_0)}. \]

(9)

Since \( \theta_0 \) is unknown, we cannot construct \( F_{2\ell,T_s}(i\omega) \) in this fashion, but the point is that as \( T_s \to 0 \), \( F_{2\ell,T_s}(i\omega) \) in (9) will approach \( F_{2\ell,T_s}^c(i\omega) \), which is defined in (10). It turns out that \( F_{2\ell,T_s}^c(i\omega) \) is real, positive, and does not depend on the signal parameters \( \theta_0 \), but only on the relative degree (pole excess) \( \ell = n - m \) of the time series model. This is what we will show now.

Let the model be strictly proper, stable and \( \ell = n - m \) be its relative degree (or pole excess), i.e. the difference between the number of poles and zeros of the system. Then, at high frequencies, we can approximate the system as a chain of \( \ell \) integrators with the continuous-time spectrum

\[ \Phi_c(i\omega) = \frac{1}{|i\omega|^{2\ell}}. \]

The discrete-time spectrum for this model is then well known as [28]

\[ \Phi_d(e^{i\omega T_s}) = \frac{1}{|i\omega|^{2\ell}} + \sum_{k \neq 0} \frac{1}{|i\omega + i\omega_s k|^{2\ell}} = \frac{T_s^{2\ell}(-1)\ell e^{i\omega T_s} \Pi_{2\ell-1}(e^{i\omega T_s})}{(2\ell - 1)!} \left( e^{i\omega T_s} - 1 \right)^{2\ell}, \]

which means that we get the spectral weighting

\[ F_{2\ell,T_s}^c(i\omega) = \frac{\Phi_c(i\omega)}{\Phi_d(e^{i\omega T_s})} = \frac{1}{|i\omega|^{2\ell}} \frac{e^{i\omega T_s} - 1}{\Pi_{2\ell-1}(e^{i\omega T_s})} \left( e^{i\omega T_s} - 1 \right)^{-2\ell}, \]

(10)

where \( \Pi_{2\ell-1}(z) \) are the Euler-Frobenius polynomials, [29]

\[ \Pi_{\ell}(z) = b_1 z^{\ell-1} + b_2 z^{\ell-2} + \ldots + b_\ell \]

where

\[ b_k = \sum_{m=1}^{\ell} (-1)^{k-m} m^{\ell} \binom{n+1}{k-m}, k = 1, \ldots, \ell. \]

Finally, observe that since \( F_{2\ell,T_s}^c(i\omega) \geq 0 \) this actually means that

\[ F_{2\ell,T_s}^c(i\omega) = \frac{\left( \frac{e^{i\omega T_s} - 1}{\Pi_{2\ell-1}(e^{i\omega T_s})} \right)^{2\ell}}{\left( \frac{e^{i\omega T_s} - 1}{(2\ell - 1)!} \right)^{2\ell}}. \]

In Figure 3 we see that there is a very good correspondence between \( \Phi_c(i\omega, \theta_0)/\Phi_d(i\omega, \theta_0) \) and \( F_{2\ell,T_s}^c(i\omega) \) for the system in (7). This observation is verified by the following theoretical result.

**Theorem 1** Let the model in (2) be strictly proper, stable and \( \ell = n - m \) be its relative degree. Assume \( F_{2\ell,T_s}^c(i\omega) \) is defined as in (10). Then, for each \( \omega \)

\[ \lim_{T_s \to 0} F_{2\ell,T_s}^c(i\omega) \Phi_d(e^{i\omega T_s}, \theta_0) = \Phi_c(i\omega, \theta_0) \]

at the rate of \( T_s^{2\ell} \).

**Proof.** Let \( \Phi_c(i\omega, \theta_0) = \Phi_c(i\omega) \) and \( \Phi_d(e^{i\omega T_s}, \theta_0) = \Phi_d(e^{i\omega T_s}) \). Then for each \( \omega \) choose \( T_s \) such that \( -\omega N < \omega < \omega_N \) where \( \omega_N = \frac{\pi}{T_s} \) is the Nyquist frequency. As a consequence, \( F_{2\ell,T_s}(i\omega) > 0 \) since by Lemma 3.2 in [28]

\[ F_{2\ell,T_s}(i\omega) = \frac{1}{\left| i\omega \right|^{2\ell}} \frac{1}{\sum_{k \neq 0} \left| i\omega + i\omega_s k \right|^{2\ell}}. \]

and

\[ \frac{1}{\left| i\omega + i\omega_s k \right|^{2\ell}} \]

will only have singularities at \( \omega = \omega_s k \). This will in turn mean that

\[ \left| \Phi_c(i\omega) - F_{2\ell,T_s}^c(i\omega) \Phi_d(e^{i\omega T_s}) \right| = \left| \frac{F_{2\ell,T_s}^c(i\omega)}{F_{2\ell,T_s}(i\omega)} \Phi_c(i\omega) - \Phi_d(e^{i\omega T_s}) \right| \leq \left| \frac{1}{F_{2\ell,T_s}(i\omega)} \Phi_c(i\omega) - \Phi_d(e^{i\omega T_s}) \right| \]

since \( F_{2\ell,T_s}(i\omega) < 1 \). Using Lemma 3.2 in [28] again with the aid of the Poisson summation formula it possible to show that

\[ \left| \Phi_c(i\omega) i\omega \right|^{2\ell} \sum_{k=-\infty}^{\infty} \frac{\Phi_c(i\omega + i\omega_s k)}{\left| i\omega + i\omega_s k \right|^{2\ell}} = \sum_{k \neq 0} \Phi_c(i\omega + i\omega_s k) i\omega \]

Finally, observe that since \( F_{2\ell,T_s}(i\omega) \geq 0 \) this actually
Since the first terms in the sums cancel. Because the model in (2) is proper and stable with relative degree \( \ell \) it is possible to find a \( M > 0 \) such that \( \Phi_c(i\omega)|i\omega|^{2\ell} \leq M \) for all \( \omega \). Hence

\[
\hat{\Phi}_c(i\omega) = \frac{\sum_{k=0}^{\infty} \Phi_c(i\omega)|i\omega|^{2\ell}}{|i\omega + i\omega_k|^{2\ell}} \leq \frac{\sum_{k=0}^{\infty} \Phi_c(i\omega)|i\omega|^{2\ell}}{|i\omega + i\omega_k|^{2\ell}} \leq \sum_{k=0}^{\infty} \left( \frac{M}{|i\omega + i\omega_k|^{2\ell}} + \frac{M}{|i\omega_n + i\omega_k|^{2\ell}} \right) \leq \sum_{k=0}^{\infty} \left( \frac{M}{|i\omega_n - i\omega_k|^{2\ell}} + \frac{M}{|i\omega_n - i\omega_k|^{2\ell}} \right) \leq 2M \left( \frac{T_s}{2\pi} \right)^{2\ell} \left( 2 + \sum_{k=1}^{\infty} \frac{1}{k^{2\ell}} \right)
\]

where the last steps comes from the definition of the sampling time \( \omega_n \) and Nyquist frequency \( \omega_n \). Since the system is strictly proper and \( \ell \geq 1 \) the sum \( \sum_{k=1}^{\infty} \frac{1}{k^{2\ell}} \) will be convergent [1]. This means that

\[
|\Phi_c(i\omega) - F_{2\ell,T_s}(i\omega)| \Phi_d(e^{i\omega T_s})| \leq CT_s^{2\ell}
\]

for some \( C > 0 \) and the result will follow. Q.E.D.

From the reasoning above, a reasonable estimate of the continuous-time spectrum would therefore be

\[
\hat{\Phi}_c(i\omega) = \frac{\sum_{k=0}^{\infty} \Phi_c(i\omega)|i\omega|^{2\ell}}{|i\omega_n + i\omega_k|^{2\ell}} \hat{\Phi}_d(e^{i\omega T_s}). \tag{12}
\]

It should be noted that theorem above only concern the relationship between the true continuous- and discrete-time power spectral densities as \( T_s \to 0 \). When the method in (12) is used and the duration of the dataset \( T \) is limited or short, spectral leakage will occur in the estimated spectrum \( \hat{\Phi}_d(e^{i\omega T_s}) \). This will in turn cause bias in the parameter estimates obtained from the Whittle method. However, this is a problem concerning small sample sizes and the Whittle estimator as such. The authors have thought about this issue, but feels that a rigorous treatment is outside the scope of this paper. For a more detailed empirical analysis of this difficult topic we refer the interested reader to [4].

\section{Numerical Illustration}

In Table 2 we have compared the performance of the approach in the section above, which we call Method 2, to using the discrete-time Whittle likelihood approach with the exact discrete time spectrum in (5), which we call Method 1. We have used different sampling intervals, and the mean parameter values have been estimated by \( N_{MC} = 250 \) Monte-Carlo simulations. The system is the one in (7) and the correspondence between the mean parameter estimates is good.

In Figure 4 we have compared the mean values of \( N_{MC} = 250 \) Monte-Carlo simulations of parameter estimates versus the sample time \( T_s \). The dotted line represents the method where we have assumed that \( \hat{\Phi}_c(T_s(i\omega)) = \hat{\Phi}_d(N_c(e^{i\omega T_s})) \). The solid line shows the result when \( \hat{\Phi}_c(T_s(i\omega)) = F_{2\ell,T_s}(e^{i\omega T_s}) \Phi_d(N_c(e^{i\omega T_s})) \) as in (10). Frequencies up to \( 2\pi \) rad/s have been used and the time of observation is \( T = 1000 \) s. For low sampling rates the difference between the two estimates is significant. This is due to the bias that occurs if no spectral weighting by \( F_{2\ell,T_s}(i\omega) \) is used.

\section{Introduction to Cardinal B-Splines}

In this section there will be a short introduction to some aspects of B-spline interpolation. The purpose is that of preparing the reader for the results in the remaining part of the paper. Most of the material is based on the two excellent papers by Unser [26,27] on B-splines and signal processing.
Table 2
Comparison of mean values for $N_{MC} = 250$ Monte-Carlo simulations of parameter estimates versus the sample time $T_s$. Method 1 is using the discrete-time Whittle likelihood approach with the exact discrete time spectrum. Method 2 employs the continuous-time Whittle likelihood approach with the spectral estimator (10). The system is $H_c(s) = \frac{\sigma}{s + \alpha_1 + \alpha_2}$ where $\alpha_1 = 1$, $\alpha_2 = 3$ and $\alpha_2 = 2$ and the performance of the methods are similar.

<table>
<thead>
<tr>
<th>$T_s$</th>
<th>Method</th>
<th>$a_1 = 3$</th>
<th>$a_2 = 2$</th>
<th>$\sigma = 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6</td>
<td>1</td>
<td>3.04 ± 0.22</td>
<td>2.02 ± 0.13</td>
<td>1.00 ± 0.047</td>
</tr>
<tr>
<td>0.6</td>
<td>2</td>
<td>2.99 ± 0.21</td>
<td>2.02 ± 0.13</td>
<td>1.01 ± 0.047</td>
</tr>
<tr>
<td>0.5</td>
<td>1</td>
<td>3.02 ± 0.16</td>
<td>2.02 ± 0.11</td>
<td>1.01 ± 0.034</td>
</tr>
<tr>
<td>0.5</td>
<td>2</td>
<td>3.05 ± 0.18</td>
<td>2.04 ± 0.12</td>
<td>1.02 ± 0.036</td>
</tr>
<tr>
<td>0.4</td>
<td>1</td>
<td>3.03 ± 0.18</td>
<td>2.02 ± 0.11</td>
<td>1.01 ± 0.038</td>
</tr>
<tr>
<td>0.4</td>
<td>2</td>
<td>3.04 ± 0.18</td>
<td>2.02 ± 0.11</td>
<td>1.01 ± 0.039</td>
</tr>
<tr>
<td>0.3</td>
<td>1</td>
<td>3.03 ± 0.17</td>
<td>2.01 ± 0.11</td>
<td>1.01 ± 0.035</td>
</tr>
<tr>
<td>0.3</td>
<td>2</td>
<td>3.03 ± 0.17</td>
<td>2.02 ± 0.12</td>
<td>1.01 ± 0.035</td>
</tr>
<tr>
<td>0.2</td>
<td>1</td>
<td>3.00 ± 0.19</td>
<td>2.01 ± 0.12</td>
<td>1.00 ± 0.038</td>
</tr>
<tr>
<td>0.2</td>
<td>2</td>
<td>3.00 ± 0.16</td>
<td>2.01 ± 0.12</td>
<td>1.00 ± 0.031</td>
</tr>
<tr>
<td>0.1</td>
<td>1</td>
<td>3.02 ± 0.18</td>
<td>2.02 ± 0.11</td>
<td>1.01 ± 0.035</td>
</tr>
<tr>
<td>0.1</td>
<td>2</td>
<td>3.02 ± 0.18</td>
<td>2.02 ± 0.12</td>
<td>1.01 ± 0.036</td>
</tr>
</tbody>
</table>

Assume that we by interpolation want to approximate some $2\ell$ times differentiable function $r(t)$ by the B-splines such that

$$\hat{r}(t) = \sum_{k=-\infty}^{\infty} c(k) B_{2\ell,T_s}^c(t - k T_s)$$

(13)

where $B_{2\ell,T_s}^c$ is the B-spline interpolation kernel of order $2\ell$ which will be defined more precisely later on. Also, we define the sampled version of spline kernel function as

$$B_{2\ell,T_s}^d(k) = B_{2\ell,T_s}^c(k T_s).$$

Then, the interpolation conditions for the bi-infinite sequence $\{r(k T_s)\}_{k=-\infty}^{\infty}$ of equidistantly distributed data points will be

$$r(l T_s) = \hat{r}(l T_s), \quad l = -\infty, \ldots, \infty$$

(14)

$$= \sum_{k=-\infty}^{\infty} c(k) B_{2\ell,T_s}^d(l - k),$$

(15)

Taking the z-transform of this expression will then give

$$R(z) = \sum_{k=-\infty}^{\infty} r(k T_s) z^{-k} = C(z) B_{2\ell,T_s}^d(z),$$

where

$$C(z) = \sum_{k=-\infty}^{\infty} c(k) z^{-k},$$

$$B_{2\ell,T_s}^d(z) = \sum_{k=-\infty}^{\infty} B_{2\ell,T_s}^d(k) z^{-k}.$$

Extracting the z-transform of the coefficients is then just a matter of inversion, and we will get

$$C(z) = \left[ B_{2\ell,T_s}^d(z) \right]^{-1} R(z).$$

Hence, it remains is to compute the Z-transform $B_{2\ell,T_s}^d(z)$ in order to calculate $C(z)$.

8.1 Z-transform of a B-Spline

The B-spline basis function of order $2\ell$ is defined as [26]

$$B_{2\ell,T_s}^c(t) = (\ell + 1) T_s \frac{\Delta_{2\ell}^2((-t)(2\ell - 1))}{(\ell + 1)!}$$

$$= \sum_{l=0}^{2\ell} \frac{(-1)^l}{(2\ell - 1)!} \frac{(2\ell)}{l} \left( (2\ell - l) T_s - t \right)^{2\ell - 1}.$$
where $\Delta_{T_s}$ is the delta operator \[7\]
\[
\Delta_{T_s} f(t) = \frac{f(t + T_s) - f(t)}{T_s}
\]
and
\[
f_r^{2\ell-1} = \begin{cases} 
  f^{2\ell-1} & \text{if } f \geq 0 \\
  0 & \text{if } f < 0
\end{cases}
\]
In turn, this means that the transformed version of the spline basis function can be computed as \[26\]
\[
B_{2\ell, T_s}^d(z) = \sum_{k=-\infty}^{\infty} B_{2\ell, T_s}(k) z^{-k}
\]
\[
= z^{-(2\ell)} \sum_{l=0}^{2\ell} \frac{(-1)^l}{(2\ell - 1)!} \left( \frac{2\ell}{l} \right) z^{2\ell-1} \sum_{k=0}^{\infty} k^{2\ell-1} z^k
\]
\[
= z^{-(2\ell)} \frac{(1-z)^{2\ell}}{(2\ell - 1)!} \sum_{k=0}^{\infty} k^{2\ell-1} z^k. \tag{16}
\]
Because of the relationship \[18\]
\[
\sum_{k=0}^{\infty} k^{2\ell-1} z^k = \frac{\Pi_{2\ell-1}(z)}{(1-z)^{2\ell}}, \quad |z| \leq 1
\]
between the Euler-Frobenius polynomials $\Pi_l(z)$ and the summation in (16), the following explicit expression,
\[
B_{2\ell, T_s}(z) = z^{-2\ell+1} \Pi_{2\ell-1}(z) / (2\ell - 1)!
\]
is true for the sampled version of a B-spline.

8.2 Fourier Transform of a Spline

Altogether, the discussion above means that when a function is represented as a spline in (13), its continuous-time Fourier transform will be
\[
\hat{R}(i\omega) = \int_{-\infty}^{\infty} \hat{r}(t) e^{-i\omega t} dt
\]
\[
= \int_{-\infty}^{\infty} \left( \sum_{k=-\infty}^{\infty} c(k) B_{2\ell, T_s}^c(t - T_s k) \right) e^{-i\omega t} dt
\]
\[
= \sum_{k=-\infty}^{\infty} c(k) \int_{-\infty}^{\infty} B_{2\ell, T_s}^c(t - T_s k) e^{-i\omega t} dt
\]
\[
= \int_{-\infty}^{\infty} B_{2\ell, T_s}^c(t) e^{-i\omega t} dt \sum_{k=-\infty}^{\infty} c(k) e^{-i\omega T_s k}. \tag{17}
\]
If we use the relationship \[18\]
\[
B_{2\ell, T_s}^c(i\omega) = \int_{-\infty}^{\infty} B_{2\ell, T_s}^c(t) e^{-i\omega t} dt = \left( \frac{1 - e^{-i\omega T_s}}{i\omega} \right)^{2\ell},
\]
and insert this expression into (17) we get
\[
\hat{R}_c(i\omega) = F_{2\ell, T_s}^c(i\omega) R_d(e^{i\omega T_s})
\]
\[
= B_{2\ell, T_s}^c(i\omega) R_d(e^{i\omega T_s}),
\]
where
\[
F_{2\ell, T_s}^c(i\omega) = \frac{(e^{i\omega T_s} - 1)^{2\ell}}{(2\ell - 1)!}
\]

8.3 Fundamental Polynomial B-Spline Function

An interesting consequence of the line of reasoning in the subsection above is that if we interpret $F_{2\ell, T_s}^c(i\omega)$ as the continuous-time Fourier transform of some function $F_{2\ell, T_s}(t)$ such that
\[
F_{2\ell, T_s}(i\omega) = \frac{B_{2\ell, T_s}^c(i\omega)}{B_{2\ell, T_s}^c(e^{i\omega T_s})} = \int_{-\infty}^{\infty} F_{2\ell, T_s}(t) e^{-i\omega t} dt.
\]
Then, $F_{2\ell, T_s}(t)$ is the so called fundamental spline function of order $2\ell$ (see for instance Lecture 4 in \[18\]) which corresponds to the solution of the interpolation problem
\[
\delta(t) = \sum_{k=-\infty}^{\infty} c(k) B_{2\ell, T_s}^c(T_s l - T_s k) \quad l = -\infty, \ldots, \infty
\]
where
\[
\delta(t) = \begin{cases} 
  1 & \text{if } l = 0 \\
  0 & \text{if } l \neq 0
\end{cases}
\]
is the Kronecker delta function. This means that we can also write our interpolation function in (13) as
\[
\hat{r}_c(t) = \sum_{k=-\infty}^{\infty} r(T_s k) F_{2\ell, T_s}^c(t - T_s k)
\]
which is called the Lagrange form \[5\] of the spline representation. The functions $F_{2\ell, T_s}^c(t - kT_s)$ can be thought of as an orthogonal basis for the linear mapping from the interpolation points to the spline of order $2\ell$ with equidistantly distributed knots.
8.4 Interpolation of the Auto-Correlation Function

It is a well known fact that the auto-correlation function
\[
\hat{r}_d(kT_s) = \frac{1}{N_t} \sum_{l=k+1}^{N_t} y(lT_s) y((k + l)T_s), \quad 0 \leq k \leq N_t
\]
is related to the discrete-periodogram \( \hat{\Phi}_{d,N_t}(e^{i\omega T_s}) \) such that [22]
\[
\hat{\Phi}_{d,N_t}(e^{i\omega T_s}) = \left| Y_{d,N_t}(e^{i\omega T_s}) \right|^2 = \sum_{k=-(N_t-1)}^{N_t-1} \hat{r}_d(kT_s)e^{-i\omega T_s}
\]
This implies that estimating the continuous-time periodogram from the discrete-time periodogram using the spectral weighting \( F^c_{2\ell,T_s}(i\omega) \) will yield
\[
\hat{\Phi}_{c,T}(i\omega) = F^c_{2\ell,T_s}(i\omega) \hat{\Phi}_{d,N_t}(e^{i\omega T_s}) = F^c_{2\ell,T_s}(i\omega) \sum_{k=-(N_t-1)}^{N_t-1} \hat{r}_d(kT_s)e^{-i\omega kT_s} = \int_{-T}^{T} \hat{r}_c(t)e^{-i\omega t} dt.
\]
Hence, the estimate of the continuous-time autocorrelation function can be represented as
\[
\hat{r}_c(t) = \sum_{k=-(N_t-1)}^{N_t-1} \hat{r}_d(kT_s) F^c_{2\ell,T_s}(t - kT_s)
\]
where \( F^c_{2\ell,T_s}(t) \) can also be interpreted as the fundamental spline basis function of order \( 2\ell \). In the case of continuous time series models and uniformly sampled data, one is therefore interpolating the covariance function instead of the output as in the case of input-output models.

9 Interpretations of Spectral Factorization

Imagine that it is possible to compute the spectral factor \( G^c_{\ell,T_s}(i\omega) \) of \( F^c_{2\ell,T_s}(i\omega) \) such that
\[
F^c_{2\ell,T_s}(i\omega) = |G^c_{\ell,T_s}(i\omega)|^2.
\]
Then, you could write the estimate of the continuous spectrum as
\[
\hat{\Phi}_{c,T}(i\omega) = F^c_{2\ell,T_s}(i\omega) \hat{\Phi}_{d,N_t}(e^{i\omega T_s}) = F^c_{2\ell,T_s}(i\omega) \left| Y_{d,N_t}(e^{i\omega T_s}) \right|^2 = \left| G^c_{\ell,T_s}(i\omega) Y_{d,N_t}(e^{i\omega T_s}) \right|^2
\]
This in turn would mean that we could define a function
\[
\tilde{y}_c(t) = \sum_{k=0}^{N_t-1} y(kT_s) G^c_{\ell,T_s}(t - kT_s)
\]
such that
\[
\tilde{Y}_c(i\omega) = G^c_{\ell,T_s}(i\omega) Y_{d,N_t}(e^{i\omega T_s}) = \sum_{k=0}^{N_t-1} y(kT_s)G^c_{\ell,T_s}(i\omega) e^{i\omega T_s}k = \sum_{k=0}^{N_t-1} y(kT_s) \int_{-\infty}^{\infty} G^c_{\ell,T_s}(t - kT_s)e^{-i\omega t} dt = \int_{-\infty}^{\infty} \tilde{y}_c(t)e^{-i\omega t} dt.
\]
Now, as the reader might have noticed, the entity \( F^c_{2\ell,T_s}(i\omega) \) is just the Fourier transform of the fundamental spline function \( F^c_{2\ell,T_s}(i\omega) \) introduced in Section 8.3 and its factorization would be
\[
F^c_{2\ell,T_s}(i\omega) = \left| \frac{e^{i\omega T_s-1}}{\omega^T_s} \right|^{2\ell} \frac{B_{2\ell-1}(e^{i\omega T_s})}{(2\ell-1)!} = C(e^{i\omega T_s}) \left( \frac{e^{i\omega T_s-1}}{\omega^T_s} \right)^\ell
\]
where
\[
|C(z)|^2 = \left| \frac{(2\ell - 1)!}{B_{2\ell-1}(z)} \right|.
\]
This is possible, since if \( z \) is a root of \( B_{2\ell-1}(z) \) then so is \( 1/z \) [21]. Hence we have
\[
G^c_{\ell,T_s}(i\omega) = C(e^{i\omega T_s}) \left( \frac{e^{i\omega T_s-1}}{\omega^T_s} \right)^\ell,
\]
and since the Fourier transform of a traditional B-spline of order \( \ell \) is
\[
B^c_{\ell,T_s}(i\omega) = \left( \frac{e^{i\omega T_s-1}}{\omega^T_s} \right)^\ell.
\]
the function $G_{c,T_s}^\ell$ can be expressed as

$$G_{c,T_s}^\ell(t) = \sum_{k=0}^{\infty} c(k) B_{c,T_s}^\ell(t-kT_s) \quad (18)$$

with the convolution property

$$F_{c,T_s}^\ell(t) = \int_{-\infty}^{\infty} G_{c,T_s}^\ell(t-\tau) G_{c,T_s}^\ell(\tau) d\tau.$$ 

The following example will illustrate the above line of reasoning for the case when $\ell = 2$.

**9.1 Example**

Let $\ell = 2$, and we will have

$$F_{c,T_s}^2(i\omega) = F_{c,T_s}^4(i\omega) = \frac{(\frac{\omega T_s}{\omega T_s})^4}{B_3(e^{i\omega T_s})},$$

where

$$B_3(z) = z^2 + 4z + 1.$$ 

Now, we can write

$$B_3(z) = z^2 + 4z + 1 = (z + 2 - \sqrt{3})(z + 2 + \sqrt{3}) = \frac{z}{2 - \sqrt{3}}(1 + z^{-1}(2 - \sqrt{3}))(1 + z(2 - \sqrt{3}))$$

which means that one can choose

$$C(z) = \sqrt{6}\sqrt{2 - \sqrt{3}} \frac{1}{1 + z^{-1}(2 - \sqrt{3})}.$$ 

Hence one will have

$$G_{c,T_s}^2(i\omega) = \sqrt{6}\sqrt{2 - \sqrt{3}} \frac{(\frac{\omega T_s}{\omega T_s})^2}{e^{i\omega T_s} + 2 - \sqrt{3}}$$

and

$$G_{c,T_s}^2(2)(i\omega) = \sqrt{6}\sqrt{2 - \sqrt{3}} \sum_{k=0}^{\infty} (-1)^k (2 - \sqrt{3})^k B_{c,T_s}^2(t-kT_s)$$

This function is illustrated in Figure 5.

**10 Conclusions**

In this paper, the problem of frequency domain estimation of continuous time series models was presented. The central result was a method for the estimation of the continuous-time spectrum from uniformly sampled data. It was proved that the method would approach the true spectrum when the sampling time approaches zero. It could also be interpreted as a way of reconstructing the continuous-time covariance function by means of polynomial splines. The method would manifest itself as way of spectrally weighting the discrete-time power spectral density in order to suppress the folding effects due to sampling. Finally, the spectral factor of the spectral weighting was found to correspond to a construction of the measured signals by means of polynomial spline functions.
References


Jonas Gillberg was born 1975 in Vänersborg, Sweden. He received his Bachelor of Science degree in Business Administration and Master of Science degree in Electrical Engineering in 2000, both at Linköping University. In 2002 he joined the Automatic Control group at Linköping University. He earned his Ph.D degree in Automatic Control in 2006. He currently works as an advanced process control consultant with IPCOSAptitude Ltd, serving the oil, gas and petrochemicals industry.

Lennart Ljung received his Ph.D in Automatic Control from the Lund Institute of Technology in 1974. Since 1976 he is Professor of the chair of Automatic Control in Linköping, Sweden, and is currently Director of the Competence Center “Information Systems for Industrial Control and Supervision” (ISIS). He has held visiting positions at Stanford and MIT and has written several books on System Identification and Estimation. He is an IEEE Fellow, an IFAC Fellow and an IFAC Advisor as well as a Member of the Royal Swedish Academy of Sciences (KVA), a Member of the Royal Swedish Academy of Engineering Sciences (IVA), an Honorary Member of the Hungarian Academy of Engineering and a Foreign Associate of the US National Academy of Engineering (NAE). He has
received honorary doctorates from the Baltic State Technical University in St. Petersburg, from the Uppsala University, Sweden, from the Technical University of Troyes, France, and from the Catholic University of Leuven, Belgium. In 2002 he received the Quazza Medal from IFAC, in 2003 the Hendryk W. Bode Lecture Prize from the IEEE Control Systems Society and he is the recipient of the IEEE Control Systems Award for 2007.
The subject of this paper is the direct identification of continuous-time autoregressive moving average (ARMA) models. The topic is viewed from the frequency domain perspective which then turns the reconstruction of the continuous-time power spectral density (CT-PSD) into a key issue. The first part of the paper therefore concerns the approximate estimation of the CT-PSD from uniformly sampled data under the assumption that the model has a certain relative degree. The approach has its point of origin in the frequency domain Whittle likelihood estimator. The discrete- or continuous-time spectral densities are estimated from equidistant samples of the output. For low sampling rates the discrete-time spectral density is modeled directly by its continuous-time spectral density using the Poisson summation formula. In the case of rapid sampling the continuous-time spectral density is estimated directly by modifying its discrete-time counterpart.