Brief paper

On the indirect approaches for CARMA model identification ☆

Kaushik Mahata*, Minyue Fu

Centre for Complex Dynamic Systems and Control, School of Electrical Engineering and Computer Science, University of Newcastle, Callaghan, NSW 2308, Australia

Abstract

In this paper, we study the problem of reconstructing a continuous-time (CT) model from an identified discrete-time (DT) model for a continuous-time stochastic process. We present a new necessary and sufficient condition for the existence of the solution. We also show that the solution is unique if it exists. Our results are useful in modeling multivariable processes as well. These results are then used to develop an algorithm where the intermediate discrete-time model estimation is not necessary. The performance of our algorithm is illustrated using numerical simulations.
© 2007 Elsevier Ltd. All rights reserved.

Keywords: Continuous-time; Stochastic process; CARMA model; Identification; Indirect approach; Solvability

1. Introduction

In practice, continuous-time (CT) systems are identified using sampled signals. Hence it is quite popular to identify an intermediate discrete-time (DT) model from the sampled data (Söderström, 1991) and transform the estimated DT model to an equivalent CT model (Söderström, 1991). This approach is often referred to as the *indirect* approach in literature. Another avenue for CT model identification is to use a direct approach, where no intermediate DT model is estimated. A direct approach is often preferred to the indirect approach because of several reasons: (i) at a fast sampling rate, the poles and the zeros of the associated DT system cluster close to the point $1 + i0$ in the complex plane, leading to a numerically ill-conditioned identification problem; (ii) there is no guarantee that a solution exists for DT to CT model conversion problem; and (iii) it is not known how the reconstructed CT model depends on the

realization of the DT model. These issues have been discussed in Larsson, Mossberg, and Söderström (2006) and Söderström (1991), and references therein.

Direct algorithms for CT autoregressive model identification have been an active research topic in recent years (Larsson & Söderström, 2002; Söderström & Mossberg, 2000). The popular algorithms use *delta operators* (Feuer & Goodwin, 1996). This approach is computationally efficient and avoids nonlinear transformation. However, the extension of this technique to continuous-time ARMA (CARMA) models is nontrivial. An alternative analytic interpolation framework is proposed in Mahata and Fu (2006), but it is not clear how to extend this algorithm for multivariable processes. In this paper, the indirect approach is examined in detail. We first focus on the DT to CT model transformation step. We present a necessary and sufficient condition for the solvability of the model transformation problem. The solution is shown to be unique, provided it exists. Unlike the previous results (El-Khoury & Crisalle, 1992; Söderström, 1990), our result is valid for multivariable processes of any order. Using the analysis we present several ways of indirect modeling approach with guaranteed solution. Next, we propose an alternative computationally efficient approach where the intermediate DT model identification step is not needed. At most it is required to solve a convex problem which can be solved globally and efficiently in

* Corresponding author.
E-mail addresses: Kaushik.Mahata@newcastle.edu.au (K. Mahata), Minyue.Fu@newcastle.edu.au (M. Fu).

polynomial time. In addition, for the single-variate case the proposed algorithm can achieve the Cramér–Rao bound, with a significantly low computational burden.

2. Preliminaries

Consider a CT stationary stochastic process $y(t) \in \mathbb{R}^m, t \geqslant 0$, given in terms of the linear stochastic differential equation

$$dz(t) = Az(t)\,dt + B\,de(t), \quad y(t) = Cz(t), \tag{1}$$

where $z(t) \in \mathbb{R}^n$, $m \leqslant n$. The process $e(t) \in \mathbb{R}^m$, is a Wiener process with unit incremental covariance matrix. The problem under consideration is to model $y(t)$ from a sampled version $y(kh)$, $k \in \{0, 1, \ldots, N-1\}$.

Since a proper rational model leads to unbounded variance of $y(t)$, we consider only strictly proper models (1). The sampled signal admits a DT state space representation (Söderström, 2002, p. 86)

$$z_{\mathrm{d}}(kh + h) = e^{Ah}z_{\mathrm{d}}(kh) + w(kh),$$

$$y(kh) = Cz_{\mathrm{d}}(kh), \tag{2}$$

where $w(kh)$ is a fictitious DT zero-mean white noise with

$$\mathscr{E}\{w(k_1 h)w'(k_2 h)\} = R_{\mathrm{d}}\delta_{k_1, k_2}.$$

Furthermore, R_{d} is given by

$$R_{\mathrm{d}} = \int_0^h e^{At} BB' e^{A't}\,dt. \tag{3}$$

Model (2) is the DT equivalent to the underlying CT model (1) in the sense that the second order statistics of the DT model is consistent with the CT process at the sampling instants. The CT state $z(t)$ and the DT state $z_{\mathrm{d}}(kh)$ have the same covariance matrix. Indeed, if P is the covariance matrix of $z(t)$ then it must satisfy the CT Lyapunov equation

$$AP + PA' + BB' = 0. \tag{4}$$

It can be shown (Söderström, 1991) that P also satisfies the DT Lyapunov equation

$$R_{\mathrm{d}} = P - FPF', \quad F = e^{Ah}. \tag{5}$$

The spectrum of the CT process can be written as

$$\begin{aligned}
\Phi_{\mathrm{c}}(s) &= C(sI - A)^{-1}BB'(-sI - A')^{-1}C' \\
&= -C(sI - A)^{-1}\{AP + PA'\}(-sI - A')^{-1}C' \\
&= K'(-sI - A')^{-1}C' + C(sI - A)^{-1}K, \tag{6}
\end{aligned}$$

where $K = PC'$. The observed DT process spectrum is given by

$$\begin{aligned}
\Phi_{\mathrm{d}}(z) &= C(zI - F)^{-1}R_{\mathrm{d}}(z^{-1}I - F')^{-1}C' \\
&= C(zI - F)^{-1}\{P - FPF'\}(z^{-1}I - F')^{-1}C' \\
&= L(z) + L'(z^{-1}), \tag{7}
\end{aligned}$$

$$L(z) = D + C(zI - F)^{-1}H, \tag{8}$$

where $H = FPC'$ and $D + D' = CPC'$ (Söderström, 2002, p. 96). The function $L(z)$ will be referred to as the half-spectrum of the observed DT process. In the indirect approach to CARMA modeling we identify the equivalent DT model first (Larsson, 2005), and then transform the identified DT model to an equivalent CT model. Given the estimates $\hat{F}$, $\hat{C}$ and $\hat{R}_{\mathrm{d}}$, the key steps for the DT to CT model transformation algorithm (Söderström, 1991) are the following:

Algorithm 1. (1) Estimate A as, see Golub and Van Loan (1989, p. 556),

$$\hat{A} = \frac{1}{h}\log(\hat{F}) := \frac{1}{h}\sum_{k=1}^{\infty}(-1)^{k-1}\frac{\hat{F}^k}{k}. \tag{9}$$

(2) Estimate P by solving the DT Lyapunov equation; see (5)

$$\hat{P} - \hat{F}\hat{P}\hat{F}' = \hat{R}_{\mathrm{d}}.$$

(3) The CT transfer function $G(s) := C(sI - A)^{-1}B$ is estimated by solving a spectral factorization problem

$$\begin{aligned}
\hat{G}(s)\hat{G}'(-s) &= \hat{\Phi}_{\mathrm{c}}(s) \\
&:= \hat{C}(sI - \hat{A})^{-1}\hat{P}\hat{C}' + \hat{C}\hat{P}(-sI - \hat{A}')^{-1}\hat{C}'. \tag{10}
\end{aligned}$$

Assumption 1 (*Larsson, 2005; Söderström, 1991*). For any eigenvalue $\lambda = \lambda_r + i\lambda_i$ of A it holds that $-\pi/\lambda_i < h < \pi/\lambda_i$. In addition, the spectrum of $\hat{F}$ lies in the interior of the open unit disc, and none of the eigenvalues of $\hat{F}$ lies on the interval $(-1, 0]$.

Assumption 1 ensures that $\hat{A}$ is a consistent estimate of A. The spectral factorization problem (10) has no solution when $\hat{\Phi}_{\mathrm{c}}(s)$ fails to be positive definite on the imaginary axis. It turns out that the zeros of the system (2) cannot be arbitrary for the factorization problem in (10) to admit a solution (El-Khoury & Crisalle, 1992; Larsson, 2005; Söderström, 1990; Wahlberg, 1988; Weller, Moran, Ninness, & Pollington, 2001). Also for a given $\hat{F}$, $\hat{C}$ and $\hat{R}_{\mathrm{d}}$ and a nonsingular T, we know $T\hat{F}T^{-1}$, $\hat{C}T^{-1}$ and $T\hat{R}_{\mathrm{d}}T'$ gives an alternative realization. But the converse is not true, i.e. we can find many other realizations of the DT spectrum estimate for which there exist no such transformation T. The following lemma can be used to characterize all possible realizations of the DT spectrum.

Lemma 1. *Given* $M_i \in \mathbb{R}^{(m+n)\times(m+n)}$, $i \in \{1, 2\}$, *Define*

$$\Phi_i(z) := \begin{bmatrix} (zI - F')^{-1}C' \\ I \end{bmatrix}' M_i \begin{bmatrix} (z^{-1}I - F')^{-1}C' \\ I \end{bmatrix}. \tag{11}$$

Then $\Phi_1(z) = \Phi_2(z)$, $\forall z$ *if and only if there exists* $Q = Q' \in \mathbb{R}^{n \times n}$ *such that*

$$M_1 = M_2 + \mathscr{H}(F, Q, C),$$

$$\mathscr{H}(F, Q, C) := \begin{bmatrix} Q - FQF' & -FQC' \\ -CQF' & -CQC' \end{bmatrix}. \tag{12}$$

Proof. See Hassibi, Sayed, and Kailath (1999). □

Note that Lemma 1 does not require $\Phi_i(z)$ in (11) to be a valid spectrum. Also it is not necessary to have $M_i > 0$ to ensure $\Phi_i(z) > 0$ on $|z| = 1$. The necessary and sufficient condition to ensure $\Phi_i(z) > 0$ on $|z| = 1$ is given by Kalman–Yakubovitch–Popov lemma (Rantzer, 1996).

Corollary 1. *Suppose we are given* $\hat{R}_d \in \mathbb{R}^{n \times n}$ *and the associated spectrum estimate*

$$\hat{\Phi}_d(z) := \hat{C}(zI - \hat{F})^{-1}\hat{R}_d(z^{-1}I - \hat{F}')^{-1}\hat{C}'. \tag{13}$$

Then

$$\hat{\Phi}_d(z) = \bar{\Phi}_d(z) := \hat{C}(zI - \hat{F})^{-1}\bar{R}_d(z^{-1}I - \hat{F}')^{-1}\hat{C}' \quad \forall z,$$

for some $\bar{R}_d \neq \hat{R}_d$ *if and only if there exists* $Q = Q' \in \mathbb{R}^{n \times n}$ *satisfying* $Q\hat{C}' = 0$ *such that*

$$\bar{R}_d = \hat{R}_d + Q - \hat{F}Q\hat{F}'. \tag{14}$$

Proof. By Lemma 1, $\hat{\Phi}_d(z) = \bar{\Phi}_d(z)$, $\forall z$, if and only if there exists $Q = Q' \in \mathbb{R}^{n \times n}$ such that

$$\begin{bmatrix} \bar{R}_d & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \hat{R}_d & 0 \\ 0 & 0 \end{bmatrix} + \mathscr{H}(\hat{F}, Q, \hat{C}).$$

That is Q must be such that (14) holds, and

$$\hat{F}Q\hat{C}' = 0, \quad \hat{C}Q\hat{C}' = 0.$$

By Assumption 1, $\hat{F}$ is nonsingular ($\hat{F}$ has no eigenvalue at the origin). Hence $Q\hat{C}' = 0$. $\square$

From the above observations a question about the uniqueness arises naturally: if two different $\hat{R}_d$ leading to the same DT spectrum are used in Algorithm 1, will they lead to the same CT spectrum?

3. Existence and uniqueness

In this section we give existence and uniqueness results for the DT to CT model transformation algorithm, Algorithm 1. Our first result gives a necessary and sufficient condition for the solvability of $\hat{G}(s)$ in (10).

Theorem 1. *Let* $C_\perp \in \mathbb{R}^{n \times (n-m)}$ *be a full column-rank matrix such that* $\hat{C}C_\perp = 0$. *Then the spectral factorization problem* (10) *admits a solution if and only if*

$$\hat{A}C_\perp \Omega C_\perp' + C_\perp \Omega C_\perp' \hat{A}' > \hat{A}\hat{P} + \hat{P}\hat{A}', \tag{15}$$

for some $\Omega = \Omega' \in \mathbb{R}^{(n-m) \times (n-m)}$.

Proof. By Kalman–Yakubovitz–Popov lemma (Rantzer, 1996), $\hat{\Phi}_c(s)$ in (10) admits a stable minimum-phase spectral factor if and only if

$$\begin{bmatrix} \hat{A}S + S\hat{A}' & (S + \hat{P})\hat{C}' \\ \hat{C}(S + \hat{P}) & 0 \end{bmatrix} > 0,$$

for some $S = S' \in \mathbb{R}^{n \times n}$, which is equivalent to

$$\hat{A}S + S\hat{A}' > 0, \quad \hat{C}(S + \hat{P}) = 0. \tag{16}$$

Let the rank of $S + \hat{P}$ be r. The second condition in (16) implies that $r < n$. Since $S + \hat{P}$ is symmetric, it admits an eigenvalue decomposition

$$S + \hat{P} = U\Sigma U', \quad U \in \mathbb{R}^{n \times r}, \quad \Sigma \in \mathbb{R}^{r \times r}, \tag{17}$$

where Σ is a diagonal matrix having the r nonzero eigenvalues of $S + \hat{P}$ as its diagonal entries, while the corresponding eigenvectors constitute the columns of U. Note that Σ^{-1} exists and $U'U = I$. Hence the second condition in (16) gives

$$\hat{C}(S + \hat{P})U\Sigma^{-1} = \hat{C}U\Sigma U'U\Sigma^{-1} = \hat{C}U = 0.$$

Thus, there exists $\Sigma_1 \in \mathbb{R}^{(n-m) \times r}$ such that $U = C_\perp \Sigma_1$. Using this in (17) gives

$$S + \hat{P} = C_\perp \Omega C_\perp', \quad \Omega = \Sigma_1 \Sigma \Sigma_1'. \tag{18}$$

Since Σ is a diagonal matrix, we have $\Omega = \Omega'$. Combining the first condition in (16) and (18) we get (15). $\square$

Theorem 1 gives a feasibility test over Ω instead of over S (in KYP lemma) and hence reduces the search dimension from n to $n - m$.

Next we address the uniqueness issue.

Theorem 2. *Let* $R_1, R_2 \in \mathbb{R}^{n \times n}$, $R_1 \neq R_2$, *be such that*

$$\hat{\Phi}_d(z) = \hat{C}(zI - \hat{F})^{-1}R_i(z^{-1}I - \hat{F}')^{-1}\hat{C}',$$

for $i = 1, 2$. *Let* $P_1, P_2 \in \mathbb{R}^{n \times n}$ *satisfy*

$$R_i = P_i - \hat{F}P_i\hat{F}', \quad i = 1, 2. \tag{19}$$

Denote the reconstructed CT spectrum by

$$\Phi_{ci}(s) = \hat{C}(sI - \hat{A})^{-1}P_i\hat{C}' + \hat{C}P_i(-sI - \hat{A}')^{-1}\hat{C}'.$$

Then $\Phi_{c1}(s) = \Phi_{c2}(s)$, $\forall s$.

Proof. By Corollary 1 there exists $Q = Q' \in \mathbb{R}^{n \times n}$ such that

$$Q\hat{C}' = 0, \quad R_1 - R_2 = Q - \hat{F}Q\hat{F}'.$$

Combining with (19) we have

$$P_1 - P_2 - Q = \hat{F}(P_1 - P_2 - Q)\hat{F}'.$$

Since the spectrum of $\hat{F}$ lies in the interior of the open unit disc, we must have $P_1 - P_2 = Q$. Consequently,

$$\Phi_{c1}(s) - \Phi_{c2}(s)$$
$$= \hat{C}(sI - \hat{A})^{-1}Q\hat{C}' + \hat{C}Q(-sI - \hat{A}')^{-1}\hat{C}' = 0,$$

for all s, and the theorem follows. $\square$

4. DT to CT transformation

4.1. Transforming the innovations model

Standard indirect CARMA model identification algorithms estimate the innovations model, where one obtains the unique

minimum-phase ARMA transfer function

$$\hat{G}_{\mathrm{d}}(z) = \hat{C}(zI - \hat{F})^{-1}\hat{J} \tag{20}$$

and the DT spectrum is estimated as

$$\hat{\Phi}_{\mathrm{d}}(z) = \hat{G}_{\mathrm{d}}(z)\hat{G}'_{\mathrm{d}}(z^{-1}).$$

This requires computing $\hat{F}$, $\hat{C}$ and the associated unique full column rank $\hat{J} \in \mathbb{R}^{n \times m}$. In this case $\hat{R}_{\mathrm{d}} = \hat{J}\hat{J}'$. Algorithm 1 can be used to reconstruct the corresponding CT model provided (15) holds for some $\Omega = \Omega' \in \mathbb{R}^{(n-m)\times(n-m)}$. Otherwise, instead of step 2 in Algorithm 1, we solve the semidefinite program (Vandenberghe & Boyd, 1996)

$$\min_{\bar{P}, \Omega} \quad \|\hat{R}_{\mathrm{d}} - \bar{P} - \hat{F}\bar{P}\hat{F}'\|_{\mathrm{F}}^2$$

$$\text{s.t.} \quad \hat{A}C_{\perp}\Omega C'_{\perp} + C_{\perp}\Omega C'_{\perp}\hat{A}' > \hat{A}\bar{P} + \bar{P}\hat{A}',$$

$$\Omega = \Omega',$$

(here $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm) and set $\hat{P}$ as the argument minimizer with respect to $\bar{P}$.

4.2. Transformation of the half-spectrum

The innovations model (20) is computed using the prediction error method (PEM), which is computationally expensive and the underlying optimization problem is not tractable in general. However, Theorem 2 suggests that it is not necessary to work with the innovations model, we can work with a different $\hat{R}_{\mathrm{d}}$. An alternative way is to use the representation (7) and (8), and compute estimates $\hat{F}$, $\hat{H}$, $\hat{C}$ and $\hat{D}$ using a subspace algorithm (Mari, Stoica, & McKelvey, 2000; Van Overschee & De Moor, 1993).

Theorem 3. *Define*

$$\hat{L}(z) := \hat{D} + \hat{C}[zI - \hat{F}]^{-1}\hat{H}. \tag{21}$$

Let $\hat{A}$ be defined in (9), and $C_{\perp} \in \mathbb{R}^{n \times (n-m)}$ be a full column-rank matrix such that $\hat{C}C_{\perp} = 0$. Then the solution to the CT spectral density reconstruction problem for the DT function

$$\hat{\Phi}_{\mathrm{d}}(z) = \hat{L}(z) + \hat{L}'(z^{-1}) \tag{22}$$

exists if and only if

$$\hat{H} = \hat{F}Q\hat{C}', \quad \hat{D} + \hat{D}' = \hat{C}Q\hat{C}', \tag{23}$$

$$Q - \hat{F}Q\hat{F}' > 0. \tag{24}$$

$$\hat{A}C_{\perp}\Omega C'_{\perp} + C_{\perp}\Omega C'_{\perp}\hat{A}' > \hat{A}Q + Q\hat{A}', \tag{25}$$

for some $Q = Q' \in \mathbb{R}^{n \times n}$ and $\Omega = \Omega' \in \mathbb{R}^{(n-m)\times(n-m)}$. When the above holds, the reconstructed CT spectrum is given by

$$\hat{\Phi}_{\mathrm{c}}(s) = \hat{C}(sI - \hat{A})^{-1}Q\hat{C}' + \hat{C}Q(-sI - \hat{A}')^{-1}\hat{C}'. \tag{26}$$

Proof. The positive real lemma (Anderson, 1967) implies that (23) and (24) are the necessary and sufficient conditions for

$\hat{\Phi}_{\mathrm{d}}(z)$ to be strictly proper and $\hat{\Phi}_{\mathrm{d}}(z) > 0$, $\forall |z| = 1$. Using (22) and (23), we have

$$\hat{\Phi}_{\mathrm{d}}(z) = \begin{bmatrix} (zI - \hat{F}')^{-1}\hat{C}' \\ I \end{bmatrix}' M \begin{bmatrix} (z^{-1}I - \hat{F}')^{-1}\hat{C}' \\ I \end{bmatrix},$$

$$M = \begin{bmatrix} Q - \hat{F}Q\hat{F}' & 0 \\ 0 & 0 \end{bmatrix} + \mathscr{H}(\hat{F}, -Q, \hat{C}),$$

see (12). Therefore, by Lemma 1 we have

$$\hat{\Phi}_{\mathrm{d}}(z) = \hat{C}(zI - \hat{F})^{-1}\{Q - \hat{F}Q\hat{F}'\}(z^{-1}I - \hat{F}')^{-1}\hat{C}'.$$

In particular, $\hat{R}_{\mathrm{d}} = Q - \hat{F}Q\hat{F}'$ is an estimate of R_{d}. Consequently, by Theorem 2, the reconstructed CT spectral density function is given by (26). Now by Theorem 1, $\hat{\Phi}_{\mathrm{c}}(i\omega) > 0$, $\forall \omega$, if and only if (25) holds for some $\Omega = \Omega' \in \mathbb{R}^{(n-m)\times(n-m)}$. $\quad\square$

Given the estimates $\hat{F}$, $\hat{C}$, $\hat{H}$ and $\hat{D}$, there may not exist any Q and Ω satisfying (23)–(25). In that case we need to solve a semidefinite program

$$\min_{\Omega, Q} \quad \left\| \begin{bmatrix} \hat{H} \\ \hat{D} + \hat{D}' \end{bmatrix} - \begin{bmatrix} \hat{F} \\ \hat{C} \end{bmatrix} Q\hat{C}' \right\|_{\mathrm{F}}^2$$

$$\text{s.t.} \quad Q - \hat{F}Q\hat{F}' > 0, \quad Q = Q', \quad \Omega = \Omega',$$

$$\hat{A}C_{\perp}\Omega C'_{\perp} + C_{\perp}\Omega C'_{\perp}\hat{A}' > \hat{A}Q + Q\hat{A}'.$$

Note that the proof of Theorem 1 does not depend on the choice of $C_{\perp}$. However, for a numerically sound implementation one may prefer a $C_{\perp}$ with mutually orthogonal columns.

5. Parameterizing the CT spectrum

In the previous sections we explored the possibilities of identifying the DT equivalent spectrum and focused on the transformation from the identified DT spectrum to the associated CT spectrum. However, if we use the parameterization (6) the intermediate DT model estimation step can be skipped. It is well known for $k \geqslant 0$

$$r_k := \mathscr{E}\{y(t + kh)y'(t)\} = C\mathrm{e}^{Ahk}PC' = CF^k K. \tag{27}$$

Hence we can use a standard step in subspace identification algorithms. We can compute the estimates $\hat{C}$, $\hat{F}$ and $\hat{K}$ from the singular value decomposition of the block Hankel matrix

$$\begin{bmatrix} \hat{r}_0 & \hat{r}_1 & \cdots & \hat{r}_q \\ \hat{r}_1 & \hat{r}_2 & \cdots & \hat{r}_{q+1} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{r}_q & \hat{r}_{q+1} & \cdots & \hat{r}_{2q} \end{bmatrix},$$

where $\hat{r}_k$ is an estimate of r_k computed from the data. The associated CT spectrum estimate is positive if and only if

$$\hat{K} + S\hat{C}' = 0, \quad \hat{A}S + S\hat{A}' > 0, \tag{28}$$

for some $S = S'$. This follows from KYP lemma and a calculation similar to (16). If condition (28) does not hold for any S,

then it is required to modify $\hat{K}$ and $\hat{C}$. The idea is to solve

$$\min_{\bar{K}, \bar{C}, S} \quad \|[\hat{K} - \bar{K} \quad \hat{C} - \bar{C}]\|_F^2$$

$$\text{s.t.} \quad \bar{K} + S\bar{C}' = 0, \quad \hat{A}S + S\hat{A}' > 0, \quad S = S'.$$

It is of interest to extend the canonical correlation type algorithms (Bauer, 2005) for the CT case for estimating K, F and C with optimal statistical accuracy.

5.1. The single variate case

For single variate case the optimal statistical accuracy can be achieved using weighted subspace fitting (Viberg & Ottersten, 1991) in the framework of approximate maximum likelihood estimation. We work in the observer canonical form (Kailath, 1980, p. 107):

$$|zI - F| = z^n + \gamma_1 z^{n-1} + \cdots + \gamma_n, \quad \gamma_0 = [\gamma_1 \quad \cdots \quad \gamma_n]',$$

$$C = [1 \ 0 \ \cdots \ 0], \quad F = \mathscr{F}(\gamma_0). \tag{29}$$

Define

$$\mathscr{F}(\gamma_0) = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ -\gamma_n & -\gamma_{n-1} & \cdots & -\gamma_1 \end{bmatrix},$$

$$O_q(\gamma_0) = \begin{bmatrix} C \\ C\mathscr{F}(\gamma_0) \\ \vdots \\ C\mathscr{F}^q(\gamma_0) \end{bmatrix}.$$

Note that we need not estimate C. Now (27) gives

$$r_q := [r_0 \ r_1 \ \cdots \ r_q] = O_q(\gamma_0)K.$$

Then the approximate maximum-likelihood estimates (Stoica, Friedlander, & Söderström, 1987) are given by

$$[\hat{\gamma}' \quad \hat{K}'] = \arg \min_{[\gamma' \ \beta']} \ell(\gamma, \beta), \tag{30}$$

$$\ell(\gamma, \beta) = \{\hat{r}_q - O_q(\gamma)\beta\}' \Sigma^{-1} \{\hat{r}_q - O_q(\gamma)\beta\}, \tag{31}$$

where $\hat{r}_q$ is the estimate of r_q, and $\Sigma = \mathscr{E}\{\hat{r}_q - r_q\}\{\hat{r}_q - r_q\}'$. It turns out for $N \to \infty$ that (Söderström & Stoica, 1989, p. 571).

$$[\Sigma]_{ij} = \frac{1}{N} \sum_{k=-N}^{N} \left[1 - \frac{|k|}{N} \right] (r_k r_{k+i-j} + r_{k-j-1} r_{k+i+1}). \tag{32}$$

In practice one estimates Σ from the data using (32). Now for any given γ,

$$\hat{\beta}(\gamma) := \arg \min_{\beta} \{\hat{r}_q - O_q(\gamma)\beta\}' \Sigma^{-1} \{\hat{r}_q - O_q(\gamma)\beta\}$$

$$= [O_q'(\gamma)\Sigma^{-1}O_q(\gamma)]^{-1} O_q'(\gamma)\Sigma^{-1}\hat{r}_q. \tag{33}$$

In addition from the theory of weighted least squares (Söderström & Stoica, 1989) it is known that

$$\ell_1(\gamma) := \ell(\gamma, \hat{\beta}(\gamma)) = \hat{r}_q' Z(\gamma)\hat{r}_q,$$

$$Z(\gamma) = \Sigma^{-1} - \Sigma^{-1}O_q(\gamma)[O_q'(\gamma)\Sigma^{-1}O_q(\gamma)]^{-1}O_q'(\gamma)\Sigma^{-1}. \tag{34}$$

Therefore, we have

$$\hat{\gamma} = \arg \min_{\gamma} \ell_1(\gamma), \quad \hat{K} = \hat{\beta}(\hat{\gamma}).$$

We can express (34) in a more convenient form. Let $M(\gamma) \in \mathbb{R}^{(q+1-n) \times (q+1)}$ be a full row rank matrix such that

$$M(\gamma)O_q(\gamma) = 0, \tag{35}$$

it follows that $Z(\gamma) = M'(\gamma)[M(\gamma)\Sigma M'(\gamma)]^{-1}M(\gamma)$:

$$\ell_1(\gamma) = \hat{r}_q' M'(\gamma)[M(\gamma)\Sigma M'(\gamma)]^{-1}M(\gamma)\hat{r}_q. \tag{36}$$

Using Caley–Hamilton theorem we can show for $\gamma = [\bar{\gamma}_1 \ \cdots \ \bar{\gamma}_n]'$,

$$M(\gamma) = \begin{bmatrix} \bar{\gamma}_n & \cdots & \bar{\gamma}_1 & 1 & 0 & \cdots & 0 \\ 0 & \bar{\gamma}_n & \cdots & \bar{\gamma}_1 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \bar{\gamma}_n & \cdots & \bar{\gamma}_1 & 1 \end{bmatrix}, \tag{37}$$

satisfies (35). Now by definition of $M(\gamma)$ it follows that $M(\gamma)\hat{r}_q = \hat{\bar{r}} + \hat{R}\gamma$, where $\hat{R}$ and $\hat{\bar{r}}$ are estimates of

$$R = \begin{bmatrix} r_0 & r_1 & \cdots & r_{n-1} \\ r_1 & r_2 & \cdots & r_n \\ \vdots & \vdots & \ddots & \vdots \\ r_{q-n} & r_{q-n+1} & \cdots & r_{q-1} \end{bmatrix}, \quad \bar{r} = \begin{bmatrix} r_n \\ r_{n+1} \\ \vdots \\ r_q \end{bmatrix},$$

respectively. Hence (36) gives

$$\hat{\gamma} = \arg \min_{\gamma} [\hat{\bar{r}} + \hat{R}\gamma]'[M(\gamma)\Sigma M'(\gamma)]^{-1}[\hat{\bar{r}} + \hat{R}\gamma]. \tag{38}$$

Solving (38) is a hard problem. However, note that

$$\hat{\gamma}_{iv} = \arg \min_{\gamma} [\hat{\bar{r}} + \hat{R}\gamma]'W[\hat{\bar{r}} + \hat{R}\gamma], \tag{39}$$

for any user defined positive definite W is the extended instrumental variable estimate of γ (Söderström & Stoica, 1989). Hence, we can use an iterative procedure where an arbitrary positive definite matrix W is used to first compute $\hat{\gamma}_{iv}$, which is then used to estimate the optimal weight

$$\hat{W}_* = [M(\hat{\gamma}_{iv})\Sigma M'(\hat{\gamma}_{iv})]^{-1}.$$

Subsequently, we compute a refined instrumental variable estimate

$$\check{\gamma} = \arg \min_{\gamma} [\hat{\bar{r}} + \hat{R}\gamma]' \hat{W}_* [\hat{\bar{r}} + \hat{R}\gamma].$$

It is shown in Stoica, Söderström, and Friedlander (1985) that $\check{\gamma}$ achieves the Cramér–Rao bound for large N. In addition,

the analysis in Viberg and Ottersten (1991) can be extended to show $\check{\gamma}$ and $\hat{\gamma}$ have the same asymptotic distribution. Similarly, $\check{K}$ and

$$\check{K} = \hat{\beta}(\check{\gamma})$$

have the same asymptotic distribution. Hence, $\check{\gamma}$ and $\check{K}$ achieve the Cramér–Rao bound. From $\check{\gamma}$ one can construct $\hat{F}$ in a straightforward manner, see (29).

6. Numerical simulation results

Algorithm in Section 5 is tested in a numerical simulation study. We consider a scalar process with a spectrum

$$\Phi_c(s) = \frac{c(s)c(-s)}{a(s)a(-s)},$$

$$a(s) = s^3 + 0.3s^2 + 9s + 0.9, \quad c(s) = s^2 + 0.5s + 6.$$

The correlation function of the chosen process has an oscillatory behavior and a large time constant. To obtain a reliable estimate of such a process, it is generally required to have a large observation time window. Here we work with a data length of 500 s, sampled at a frequency of 2.5 Hz. This choice of sampling frequency is driven by two factors. The sampling frequency is bounded below by the constraint posed in Assumption 1. On the other hand, it is well known that a large sampling frequency gives rise to numerical problems, see the discussion in the Section 7. We estimate the correlation function up to 5 s from the data ($q = 12$), and use it in our estimation routine. If the data length is very large compared to the observation time, the estimation accuracy should improve with increasing q. But when the data length is finite, due to the poor accuracy for the correlation estimates for the larger time lags, the estimation accuracy deteriorates when q is increased beyond a particular value.

The estimation results obtained from 100 Monte-Carlo simulations are shown in Fig. 1(a), where the true spectrum is compared with estimated mean value ± standard deviation. As can be seen in Fig. 1(a), the estimated spectrum is slightly biased but is accurate. We have used the biased estimate of the correlation function (which is the popularly used maximum-likelihood estimate)

$$\hat{r}_\tau = \frac{1}{N} \sum_{t=1}^{N} y(t + \tau h)y(t).$$

However, if we use the unbiased estimate

$$\bar{r}_\tau = \frac{1}{N - \tau} \sum_{t=1}^{N} y(t + \tau h)y(t)$$

then the estimates are unbiased, but the mean square estimation error is slightly larger.

We give the parametric estimation results in Table 1. Here we consider the estimates of the coefficients of the polynomials $a(s)$ and $c(s)$, with a_k being the coefficient of s^{3-k} in $a(s)$, and c_k denoting the coefficient of s^{3-k} in $c(s)$.
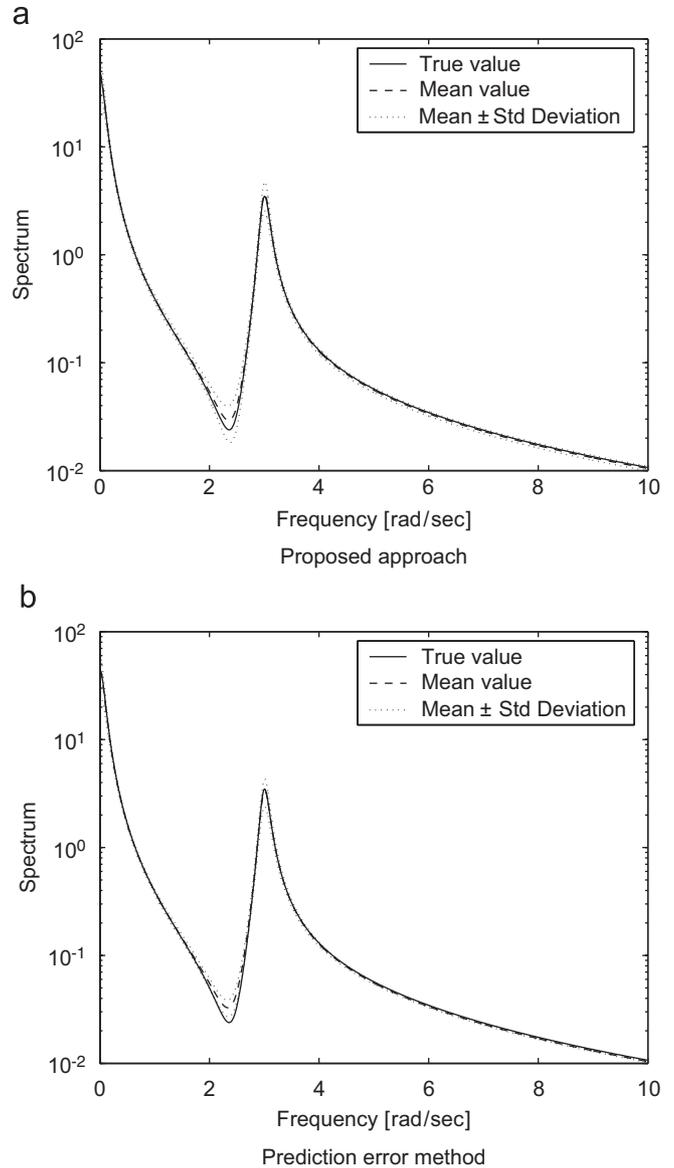
Fig. 1. Comparison of the mean of the estimated spectrum (dashed line) and the true spectrum (solid line). The mean ± standard deviation of the estimated spectrum is shown in dotted lines: (a) proposed approach; (b) prediction error method.

Table 1
Parameter estimation performance

Parameter	True vale	Mean	Std. deviation
a_1	0.3	0.3078	0.0404
a_2	9.0	9.0112	0.0995
a_3	0.9	0.9294	0.2153
c_1	1.0	0.9851	0.0254
c_2	0.5	0.5304	0.1150
c_3	6.0	5.9679	0.2033

The estimation performance of the modified PEM based approach in Section 4.1 is shown in Fig. 1(b). A comparison of Figs. 1(a) and (b) indicates that the performance of the

proposed approach is comparable with the PEM estimate. In about 25% of our simulations the PEM estimate of the DT spectrum did not have a CT counterpart. For these cases it is necessary to use the modification proposed in Section 4.1.

7. Conclusions

In this paper we have addressed the problem of reconstructing a CT model from a given DT model. A necessary and sufficient condition for the existence of the solution is given. It is also shown that the solution is unique if it exists. Based on our findings, we have presented several ways for reliable indirect identification. In addition, we have proposed a new approach where the intermediate DT model estimation step is not necessary. This method is statistically accurate and computationally efficient. It is, however, necessary to estimate F and we estimate A using (9) causing numerical problems with fast sampling of covariances. It is therefore of interest to explore the possibilities of estimating A directly.

References

Anderson, B. D. O. (1967). A system theory criterion for positive real matrices. *SIAM Journal on Control*, 5, 171–182.

Bauer, D. (2005). Comparing the CCA subspace method to pseudo maximum likelihood methods in the case of no exogeneous inputs. *Journal of Time Series Analysis*, 26(5), 631–668.

El-Khoury, M., & Crisalle, O. D. (1992). Relative zero location for second-order sampled systems. *IEEE Transactions on Automatic Control*, 37(10),

Feuer, A., & Goodwin, G. C. (1996). *Sampling in digital signal processing and control*. Boston: Birkhäuser.

Golub, G. H., & Van Loan, C. F. (1989). *Matrix computations*. (2nd ed.), Baltimore, MD, USA: John Hopkins University Press.

Hassibi, B., Sayed, A. H., & Kailath, T. (1999). *Indefinite-quadratic estimation and control: A unified approach to H_2 and H_∞ theories*. PA: SIAM.

Kailath, T. (1980). *Linear systems*. Englewood Cliffs, NJ: Prentice-Hall.

Larsson, E. (2005). Limiting sampling results for continuous-time ARMA systems. *International Journal of Control*, 78(7), 461–473.

Larsson, E., Mossberg, M., & Söderström, T. (2006). An overview of important practical aspects of continuous-time ARMA system identification. *Circuits, Systems and Signal Processing*, 25(1), 17–46.

Larsson, E. K., & Söderström, T. (2002). Identification of continuous-time AR processes from unevenly sampled data. *Automatica*, 38(4), 709–718.

Mahata, K., & Fu, M. (2006). Modelling continuous-time processes via input-to-state filters. *Automatica*, 42(7), 1073–1084.

Mari, J., Stoica, P., & McKelvey, T. (2000). Vector ARMA estimation: A reliable subspace approach. *IEEE Transactions on Signal Processing*, 48(7), 2092–2104.

Rantzer, A. (1996). On the Kalman–Yakubovitz–Popov lemma. *Systems and Control Letters*, 28, 7–10.

Söderström, T. (1990). On zero locations for sampled stochastic systems. *IEEE Transactions on Automatic Control*, AC-35, 1249–1253.

Söderström, T. (1991). Computing stochastic continuous-time models from ARMA models. *International Journal of Control*, 53, 1311–1326.

Söderström, T. (2002). *Discrete-time stochastic systems*. (2nd ed.), London: Springer.

Söderström, T., & Mossberg, M. (2000). Performance evaluation of methods for identifying continuous-time autoregressive processes. *Automatica*, 36, 53–59.

Söderström, T., & Stoica, P. (1989). *System identification*. Hemel Hempstead, UK: Prentice-Hall International.

Stoica, P., Friedlander, B., & Söderström, T. (1987). Approximate maximum-likelihood approach for ARMA spectral estimation. *International Journal of Control*, 45(4), 1281–1310.

Stoica, P., Söderström, T., & Friedlander, B. (1985). Optimal instrumental variable estimates of the AR parameters of an ARMA process. *IEEE Transactions on Automatic Control*, AC-30, 1066–1074.

Van Overschee, P., & De Moor, B. (1993). Subspace algorithms for the stochastic identification problem. *Automatica*, 29(3), 649–660.

Vandenberghe, L., & Boyd, S. (1996). Semidefinite programming. *SIAM Review*, 38(1), 49–95.

Viberg, M., & Ottersten, B. (1991). Sensor array processing based on subspace fitting. *IEEE Transactions on Signal Processing*, 39(5), 1110–1121.

Wahlberg, B. (1988). Limit results for sampled systems. *International Journal of Control*, 48, 1267–1283.

Weller, S. R., Moran, W., Ninness, B., & Pollington, A. D. (2001). Sampling zeros and the Euler–Frobenius polynomials. *IEEE Transactions on Automatic Control*, 46(2), 340–343.

Kaushik Mahata received his M.E degree in Signal Processing from Indian Institute of Science, Bangalore, India in 2000, and a Ph.D. degree in Signal Processing from Uppsala University, Sweden in 2003. Currently he is a lecturer in the University of Newcastle, Australia. His research interest includes estimation identification, spectrum analysis and machine learning.

Minyue Fu received his Bachelor's Degree in Electrical Engineering from the University of Science and Technology of China, Hefei, China, in 1982, and M.S. and Ph.D. degrees in Electrical Engineering from the University of Wisconsin-Madison in 1983 and 1987, respectively. From 1983 to 1987, he held a teaching assistantship and a research assistantship at the Uiniversity of Wisconsin-Madison. He worked as a Computer Engineering Consultant at Nicolet Instruments, Inc., Madison, Wisconsin, during 1987. From 1987 to 1989, he served as an Assistant Professor in the Department of Electrical and Computer Engineering, Wayne State University, Detroit, Michigan. For the summer of 1989, he was employed by the Universite Catholoque de Louvain, Belgium, as a Maitre de Conferences Invited. He joined the Department of Electrical and Computer Engineering, the University of Newcastle, Australia, in 1989. Currently, he is a Chair Professor in Electrical Engineering and Head of School of Electrical Engineering and Computer Science. In addition, he was a Visiting Associate Professor at University of Iowa in 1995–1996, and a Senior Fellow/Visiting Professor at Nanyang Technological University, Singapore, 2002. His main research interests include control systems, signal processing and communications. He has been an Associate Editor for the IEEE Transactions on Automatic Control and Journal of Optimization and Engineering. He is a Fellow of IEEE.