The Dimensionality of the Aliasing Problem in Models With Rational Spectral Densities[1]

Lars Peter Hansen

Carnegie-Mellon University

Thomas J. Sargent

University of Minnesota and
Federal Reserve Bank of Minneapolis

Research Department Staff Report 72
Federal Reserve Bank of Minneapolis

September 1981

ABSTRACT

This paper reconsiders the aliasing problem of identifying the parameters of a continuous time stochastic process from discrete time data. It analyzes the extent to which restricting attention to processes with rational spectral density matrices reduces the number of observationally equivalent models. It focuses on rational specifications of spectral density matrices since rational parameterizations are commonly employed in the analysis of time series data.

1. INTRODUCTION

For a variety of continuous time, linear stochastic models, the aliasing phenemenon prevents unique identification of the parameters of the continuous time stochastic process from equispaced discrete-time observations. This problem has been studied by Christopher Sims [12], P.C.B. Phillips [7,8,9], and John Geweke [3], among others. The literature indicates that prior restrictions on the continuous time model are needed to overcome the aliasing identification problem.

The purpose of this paper is to describe the dimensionality of the class of continuous time models that are observationally equivalent, given that the true continuous time spectral density matrix is rational. For some subclasses of these models, it has previously been thought that there is a countably infinite number of observationally equivalent models (see P.C.B. Phillips [9]). This paper proves that in general there is only a finite number of observationally equivalent models. In certain regions of the parameter space, there may even be no identification problem at all, though in general there is one.

Results such as these are useful because, roughly speaking, they indicate the size of the job that the additional prior restrictions, above and beyond the prior specification of a rational spectral density matrix, have to do in order to achieve identification. Put differently, the specification of a rational spectral density matrix in itself goes a greater distance toward resolving the aliasing problem than has heretofore been recognized.

This paper is a prologue to a paper [5] studying the identification of continuous time rational expectations models with rational spectral density matrices from discrete time data. In that paper, it is shown how the cross-equation restrictions characteristic of rational expectations can serve uniquely to identify the continuous time model. It is natural to inquire how big a job the cross-equation, rational expectations restrictions are performing in achieving identification. That query caused the present paper.

Section 2 briefly describes the aliasing identification problem in the case of a general, indeterministic, covariance stationary continuous time vector stochastic process. It is remarked that in this general case the class of observationally equivalent continuous time models is uncountably infinite. Section 3 describes the identification problem for the more restricted case usual in applications, of an assumed rational spectral density matrix. We briefly indicate a machinery for proving that the class of observationally equivalent continuous time models is in general finite. Section 4 then provides a more complete characterization of the situation in the special case studied by P.C.B. Phillips [9], in which the true continuous time model is a first order vector Markov process. We show that there generally exists a discrete sampling interval sufficiently fine that the continuous time model is identified.[2]

2. THE ALIASING PROBLEM UNDER COVARIANCE STATIONARITY

Consider an n dimensional continuous time stochastic process, x, that is covariance stationary and linearly regular. For simplicity we assume that the process has full rank and mean zero. Wold's Decomposition Theorem assures us that such a process has a moving average representation

$$x(t) = \int_0^\infty c(\tau)w(t-\tau)d\tau \tag{1}$$

where w is an n dimensional continuous time white noise process with intensity matrix I and where c is an (n x n) matrix function whose elements are square integrable.[3] We assume that the process w is fundamental for x, which means that linear combinations of current and past w's span the same space as linear combinations of current and past x's. Under these assumptions, the matrix function c is unique up to a post multiplication by an orthogonal matrix.[4]

Let $C(s) = \int_0^\infty c(t)e^{-st}dt$ be the Laplace transform of c. We adopt a convenient notation and write representation (1) as

$$x(t) = C(D)w(t) \tag{2}$$

where D is the time derivative operator.[5] The population covariogram of x is completely summarized by the matrix function c or equivalently by its Laplace transform C. Alternatively, the covariogram is characterized by its Fourier transfrom, the spectral density matrix. The spectral density matrix f is positive semidefinite at all frequencies ω and is related to C by

(3) $$f(\omega) = C(i\omega)C(-i\omega)' \qquad -\infty < \omega < \infty.$$

Here the prime denotes transposition (but not conjugation). Since the x process is real, the function f satisfies

(4) $$f(\omega) = \bar{f}(-\omega) = \bar{f}(\omega)',$$

where the bar denotes conjugation.

The aliasing phenomenon for models that reside in this class can be conveniently described by using spectral density matrices. Let F denote the spectral density matrix for the discrete time process X obtained by observing x at integer points in time. It is known that f and F are linked by the folding formula:

(5) $$F(\omega) = \sum_{j=-\infty}^{+\infty} f(\omega+2\pi j).$$

Since F completely summarizes the population covariance properties of X, formula (5) implies that the function f cannot be inferred from the discrete time data. This can be seen by noting that alternative Hermitian, positive semidefinite matrix functions f^* can be constructed that satisfy

(6) $$F(\omega) = \sum_{j=-\infty}^{+\infty} f^*(\omega+2\pi j)$$
$$f^*(\omega) = \bar{f}^*(-\omega)$$

and hence are observationally equivalent to f. Corresponding to each function f^* is a matrix of square integrable functions c^* with Laplace transforms C^* such that

$$f^*(\omega) = C^*(i\omega)C^*(-i\omega)',$$

and such that if w^* is an n dimensional continuous time white noise process with intensity matrix I, then w^* is fundamental for x^* where

$$x^*(t) = C^*(D)w^*(t).$$

Although the matrix function c^* cannot be obtained from c by post multiplying c by an orthogonal matrix, c^* is observationally equivalent to c with discrete time data. This is the conventional formulation of the aliasing problem in time series analysis.

The models in (1) are in general infinite parameter models, with the parameters in c being the objects whose identification is sought. At this level of generality, the aliasing problem is a local identification problem in the sense that there are observationally equivalent parameters c^* satisfying (5) and (6) that are arbitrarily close to the true parameter c. Here our measure of distance is the matrix L_2 norm

$$\int_0^\infty \text{trace}\{[c(\tau)-c^*(\tau)][c(\tau)-c^*(\tau)]'\}d\tau.$$

This suggests that there is an overwhelming number of c^*'s that are observationally equivalent to c. In fact, this number is uncountable.[6] Thus, at the general level of the model (1), the dimensionality of the class of observationally equivalent models given equispaced discrete time observations is uncountable.

3. THE ALIASING PROBLEM WITH A RATIONAL SPECTRAL DENSITY MATRIX

In applications it is necessary to adopt a finite parameterization of the matrix function c. A convenient parameterization is to assume that c has a rational Laplace transform. In particular, suppose that

$$(7) \qquad C(s) = \frac{(G_o + G_1 s + \ldots + G_{q-1} s^{q-1})}{(s-\lambda_1)(s-\lambda_2)\ldots(s-\lambda_q)} = \frac{G(s)}{g(s)}$$

where G_o, $G_1, \ldots,$ G_{q-1}, are real (n x n) matrices and λ_1, $\lambda_2, \ldots,$ λ_q are distinct complex numbers with negative real parts.[7] We assume that the zeroes of detG(s) have negative real parts and that for each j, $\lambda_j = \bar{\lambda}_k$ for some index k.[8] Finally, we assume that any two λ's with the same real part do not have imaginary parts that differ by an integer multiple of $2\pi i$. The λ's are called the poles of the complex matrix function C. A choice of C that satisfies the above restrictions is known to have an inverse Laplace transform matrix c whose elements are square integrable and are concentrated on the nonnegative real numbers.

We proceed to examine the spectral density matrix of a process with a C that satisfies specification (7). Following an approach that was used by A. W. Phillips [10], we form a partial fractions representation of the matrix function h,

$$h(s) = C(s)C(-s)' = \sum_{j=1}^{q} \left[\frac{Q_j}{s-\lambda_j} + \frac{Q_j'}{-s-\lambda_j}\right]$$

where

$$Q_j = \frac{G(\lambda_j)G(-\lambda_j)'}{g_j\, g_{-j}(-2\lambda_j)}$$

$$g_j = \lim_{s \to \lambda_j} \frac{g(s)}{(s-\lambda_j)}$$

$$g_{-j} = \lim_{s \to -\lambda_j} \frac{g(-s)}{(-s-\lambda_j)} .^9$$

Note that if λ_k is the complex conjugate of λ_j, then the elements of Q_k are complex conjugates of the elements of Q_j. The spectral density matrix of x is

$$f(\omega) = h(i\omega),$$

and the autocovariance function, which equals the Fourier transform of $f(\omega)$, is

$$(9)\qquad r(\tau) = \begin{cases} \sum_{j=1}^{q} Q_j e^{\lambda_j \tau} & \text{for } \tau \geq 0 \\ r(-\tau)' & \text{for } \tau < 0. \end{cases}$$

The autocovariance function R for the discrete time process X can be obtained by sampling r at the integers. We write this as

$$R(\tau) = \begin{cases} \sum_{j=1}^{q} Q_j (\rho_j)^{\tau} & \text{nonnegative integer } \tau \\ R(-\tau)' & \text{negative integer } \tau \end{cases}$$

where

$$(10)\qquad \rho_j = e^{\lambda_j}.$$

Suppose that we wish to construct a function r^* that is distinct from r but can be written in the form given in (9) and is equal to R at integer values of τ. Such a function r^* is a candidate for a continuous time autocovariance function that is observationally equivalent to r. To generate such a family of r^*'s we use equation (10) and the fact that $e^{2\pi i \tau} = 1$ for any integer τ. Since the function R can be inferred from discrete time data, it is evident that the complex matrices Q_j and the complex numbers ρ_j are identifiable from discrete time data. From equation (10) we see that the real parts of the poles λ_j are just the real logarithms of $|\rho_j|$. Hence the real parts of the poles are identifiable from discrete time data. On the other hand, the imaginary parts of the poles are not necessarily identifiable. If at least one of the poles is complex, then we can construct a countable infinity of real matrix functions r^* of the form given in (9) that are equal to R at integer values of τ.[10] Thus, suppose that the first two λ's form a complex conjugate pair. Let

$$\lambda_1^k = \lambda_1 + 2\pi i k$$

$$\lambda_2^k = \lambda_2 - 2\pi i k.$$

Now form the functions

$$r_k(\tau) = \begin{cases} Q_1 e^{\lambda_1^k \tau} + Q_2 e^{\lambda_2^k \tau} + \sum_{j=3}^{q} Q_j e^{\lambda_j \tau} & \text{for } \tau \geq 0 \\ r_k(-\tau)' & \text{for } \tau < 0. \end{cases}$$

The matrix functions r_k are equal to R at integer values of τ. Therefore we have generated a countable sequence of candidates for autocovariance functions of observationally equivalent models. However, in order for these functions to be legitimate autocovariance functions of a continuous time process, it is necessary and sufficient that the continuous Fourier transforms of these functions be positive semidefinite at all frequencies. That is the functions

$$f_k(\omega) = \frac{Q_1}{i\omega-\lambda_1^k} + \frac{Q_1'}{-i\omega-\lambda_1^k} + \frac{Q_2}{i\omega-\lambda_2^k} + \frac{Q_2'}{-i\omega-\lambda_2^k}$$

$$+ \sum_{j=3}^{q} \left[\frac{Q_j}{i\omega-\lambda_j} + \frac{Q_j'}{-i\omega-\lambda_j}\right]$$

must be positive semidefinite for all values of ω. While this condition is met for f it will <u>not</u> in general be satisfied for f_k. In fact it turns out that except in singular cases we can generate only a <u>finite</u> number of observationally equivalent models in this fashion. We state this result in the theorem provided below.

<u>Theorem 1</u>: Let $Q_1 = Q_{11}+Q_{12}i$ where Q_{11} and Q_{12} are real matrices. Suppose that Q_1 fails to satisfy one of the following conditions:

(i) $Q_{12}i$ is a Hermitian matrix,

(ii) $(Q_{11}+Q_{11}')$ is positive semidefinite.

Then there is a positive integer k^* such that when $|k| \geq k^*$, f_k is not a spectral density matrix for a continuous time process.

The proofs of all theorems are in the appendix.

The conditions (i) and (ii) of Theorem 1 will be met only for singular examples of the C(s) is given by (7). Theorem 1 illustrates the difficulty in generating a countable sequence of observationally equivalent continuous time models wihout violating the requirement that the implied continuous time spectral density matrix be positive semidefinite at all frequencies. When C has only one complex conjugate pair of poles, this theorem implies that in general there will only be a finite number of observationally equivalent models. On the other hand, the strategy described above for constructing observationally equivalent continuous time models does not exhaust all possible ways of constructing such models when there are more than one complex conjugate pair of poles.[11] Although we conjecture that the flavor of our results will remain intact by treating the case of multiple pairs of complex conjugate poles, we do not consider more general theorems for model (7). Instead we present a comprehensive analysis of the dimensionality of the special case of (7) that P.C.B. Phillips has studied. We turn to this analysis in the following section.

4. FIRST ORDER MARKOV MODELS

In this section we study identification of the parameters of continuous time first order Markov processes from discrete time data. We build upon and modify P.C.B. Phillips's [9] characterization of the aliasing phenomenon in this class of models.

Consider an x process that can be represented

$$(10) \qquad Dx(t) = A_o x(t) + \varepsilon(t)$$

where ε is a continuous time vector white noise with intensity matrix V_o. The square matrix A_o is real and has eigenvalues whose real parts are negative. From (10) we can derive an expression for a fundamental moving average representation as follows.

Assume that V_o has full rank and factor it according to $V_o = U_o'U_o$. Solve (10) for x(t) to obtain

(11) $$x(t) = \frac{\mathrm{adj}[DI-A_o]U_o'}{\det[DI-A_o]}w(t),$$

where $w(t) = U_o'^{-1}\varepsilon(t)$ and adj denotes adjoint. We can rewrite (11) as

$$x(t) = C(D)w(t)$$

where

$$C(D) = \frac{\mathrm{adj}[DI-A_o]}{\det[DI-A_o]}U_o' .$$

The white noise vector w(t) has intensity matrix I, the poles of C(s) are just the eigenvalues of A_o, and the Q_j matrices in the matrix partial functions decomposition of $h(s) = C(s)C(-s)'$ are rank one matrices formed from the eigenvectors of A_o. While we could proceed to discuss identification using the machinary of section 2, it is more convenient to adopt an alternative machinary appropriate for these first order Markov models, one that was used by Phillips [9].

The discrete time process obtained by sampling x at the integers has a first order autoregressive representation

(12) $$X(t) = B_0X(t-1)+\eta(t)$$

where

$$B_0 = \exp A_0$$

(13)

$$\eta(t) = {}_0\int^1 \exp(A_0\tau)\varepsilon(t-\tau)\,d\tau.$$

By the white noise nature of ε, it follows that η is a discrete time vector white noise disturbance when sampled at the integers. The contemporaneous covariance matrix of $\eta(t)$ is

(14) $$W_0 = {}_0\int^1 \exp(A_0\tau)V_0\exp(A_0'\tau)\,d\tau.$$

As noted by Phillips [9], the covariance properties of x sampled at the integers are completely characterized by (B_0,W_0). Given the pair (B_0,W_0), which is estimable from discrete time data, our goal is to identify the covariance properties of the continuous time process, which are completely characterized by (A_0,V_0). The version of the aliasing phenomenon considered by Phillips [9] is simply the fact that given (B_0,W_0) one cannot in general solve uniquely for (A_0,V_0) using equations (13) and (14).[12] We seek to characterize the dimensionality of the class of (A_0,V_0) pairs consistent with a given (B_0,W_0) pair.

To begin, we consider equation (13) and ask the question of whether the matrix equation

(15) $$\exp A^* = B_0 = \exp A_0.$$

implies that $A^* = A_0$. Without restrictions on the matrix A^*, the answer is in general no. If the matrix A_0 has complex eigenvalues, then there is a countable infinity of matrices A^* that satisfy (15). To see this, assume that the eigenvalues of A_0 are distinct and write the spectral decomposition of A_0,

(16) $$A_0 = T \Lambda T^{-1}$$

where Λ is a diagonal matrix of eigenvalues of A_0 and T is a matrix of eigenvectors of A_0. Without loss of generality, we are free to assume that the first n-2m diagonal elements are real and that the remainder occur in complex conjugate pairs as λ_{n-2m+1}, ..., λ_{n-m}, $\lambda_{n-m+1}=\bar{\lambda}_{n-m+1}$, ..., $\lambda_n=\bar{\lambda}_{n-m}$. We assume that the eigenvalues of A_0 do not differ by integer multiples of $2\pi i$. Following Phillips [9] and Coddington and Levinson [1], if a matrix A^* satisfies (15) then

(17) $$A^* = A_0 + 2\pi i T \begin{bmatrix} 0 & 0 & 0 \\ 0 & P & 0 \\ 0 & 0 & -P \end{bmatrix} T^{-1}$$

where P is an (m x m) diagonal matrix of integers. Any choice of integers for the diagonal elements of P will give rise to a solution of the matrix equation (15).

Phillips [9] asserted that the pair (A_o, V_o) is identifiable in (B_o, W_o) if and only if the matrix A_o is identifiable in B_o. This assertion would be true if, given a real valued matrix A^* of the form specified in (17), it were possible to find a positive semidefinite matrix V^* such that

$$(18) \qquad \int_0^1 \exp(A^*\tau)V^*\exp(A^{*\prime}\tau)d\tau = \int_0^1 \exp(A_0\tau)V_o\exp(A_0'\tau)d\tau.$$

Now Phillips's equation 4 shows how to compute a V^* satisfying (18) as a function of A^* and W_o. However, there is no guarantee that the resulting V^* is positive semidefinite, and so it need not be a legitimate intensity matrix of a white noise process. This fact indicates the presence of extra identifying information about A_o in the discrete innovation covariance matrix W_o, information summarized in equation (14).[13] It follows that Phillips's characterization of the identification problem must be modified to take account of the information about A_o that is contained in W_o. The question of whether V^* is positive semidefinite is equivalent to the question of whether the implied continuous time spectral density matrix is positive semidefinite at all frequencies.

Phillips asserted that if A_o has complex eigenvalues, then without additional restrictions, there is a countable infinity of pairs $\{(A_k, V_k)\}_{k=1}^{\infty}$ that are observationally equivalent to (A_o, V_o) given discrete time data. Actually, however, the number of pairs (A_k, V_k) that are observationally equivalent to (A_o, V_o) is, except for singular cases, at most finite and in some cases is equal to one even if A_o has complex eigenvalues. We proceed to substantiate this claim by stating four theorems.

The first theorem specifies circumstances in which there is a countable infinity of pairs $\{(A_k,V_k)\}_{k=1}^{\infty}$ that are observationally equivalent to (A_o, V_o) given discrete time data.

<u>Theorem 2</u>: If there exists an $A^* \neq A_o$ such that

(i) $\exp A^* = B_o$

(ii) $\int_0^1 \exp(A^*\tau)V_o\exp(A^{*\prime}\tau)\,d\tau = W_o$,

then there is an infinite sequence of distinct matrices $\{A_k\}_{k=1}^{\infty}$ that satisfy (i) and (ii).

Theorem 2 states that if we can find an A^* of the form given in (17) that also satisfies (18) for $V^* = V_o$, then there exists a countable infinity of observationally equivalent pairs $\{(A_k, V_k)\}_{k=1}^{\infty}$ where $V_k = V_o$. That is, each of the A_k matrices is associated with the same intensity matrix V_o. The key feature is that the intensity matrix remains unaltered as we entertain admissible alterations in the continuous time coefficient matrix.

Theorem 2 delineates one class of circumstances in which there is a countably infinite number of continuous time models that are consistent with the discrete time observations. It happens that the class identified in Theorem 2 contains the <u>only</u> cases in which an infinite number of continuous time models are consistent with the discrete time observations. This is established in Theorem 3.

<u>Theorem 3</u>: If there does not exist an $A^* \neq A_o$ such that (i) and (ii) of Theorem 1 are satisfied, then there is only a finite number of distinct pairs (A_k, V_k) that satisfy

(i') $\exp A_k = B_o$

(ii') $\int_o^1 \exp(A_k \tau) V_k \exp(A_k' \tau) d\tau = W_o$

(iii') V_k is positive semidefinite.

The important feature of Theorem 3 is requirement (iii'). Phillips has shown if A_o has complex eigenvalues, then there is a countable infinity of $\{(A_k, V_k)\}_{k=1}^{\infty}$ that satisfy (i') and (ii'). Theorem 2 indicates that when V_k is required to be positive semidefinite, then in many circumstances there is only a finite number of pairs (A_k, V_k) that also satisfy (i') and (ii').

It remains to determine the size of the class of cases for which there is a countable infinity of observationally equivalent continuous time models. This question is answered by Theorem 4.

<u>Theorem 4</u>: If $R_o = T^{-1} V_o T^{-1'}$ does not have any zero elements, then there is at most a finite number of distinct pairs (A_k, V_k) that satisfy (i'), (ii'), and (iii') of Theorem 2.

Theorem 4 indicates that the class of cases in which there is a countable infinity (A_k, V_k) that are observationally equivalent to (A_o, V_o) is singular. Only when R_o has zero elements can this occur. Furthermore, there are many situations in which R_o has zero elements and there is still only a finite number of observationally equivalent models.

We have calculated an example to illustrate the messages of Theorems 2, 3, and 4. For this example, we chose

$$A_o = \begin{bmatrix} -1.0 & 1.0 \\ -1.0 & -2.0 \end{bmatrix}$$

$$V_o = \begin{bmatrix} 2.0 & v_{12} \\ v_{12} & 2.0 \end{bmatrix} .$$

We ask, how many observationally equivalent models could be generated for different values of v_{12}? The eigenvalue matrix for A_o is

$$\Lambda_o = \begin{bmatrix} -1.5+.86603i & 0 \\ 0 & -1.5-.86603i \end{bmatrix} .$$

We constructed candidates for observationally equivalent models by forming

$$A_p = T\Lambda_o T^{-1} + T\begin{bmatrix} 2\pi ip & 0 \\ 0 & -2\pi ip \end{bmatrix} T^{-1}$$

for integer values of p. We then computed the matrix V_p corresponding to V_p that satisfies equation (18). The eigenvalues of V_p were then calculated to determine if they were nonnegative, as is necessarily true when V_p is positive semidefinite. Table 1 indicates the range of values of the integer p over which the eigenvalues of V_p remained nonnegative for alternative selections for v_{12}. By construction $p = 0$ is always in that range.

The case in which $v_{12} = -1$ provides an example of Theorem 2. In this case all choices of integer p gave rise to the same $V_p = V_o$, and a countable infinity of observationally equivalent models emerged. Consistent with Theorem 4, R_o was verified to be a diagonal matrix in this case. The remaining selections of v_{12} exemplify Theorems 3 and 4 in that the R_o matrices had all nonzero elements and there was only a finite number of observationally equivalent models. As v_{12} moved further away from -1, the positive semidefiniteness constraint on V_p yielded fewer observationally equivalent models. In fact, for v_{12} greater than or equal to -.4 and v_{12} less than or equal to -1.5 we did not encounter an aliasing problem at all.

Table 1

Range of Observationally Equivalent Models*

v_{12}	Lower value of p	Upper value of p
-1.9 to -1.5	0	0
-1.4	-1	0
-1.3	-1	1
-1.2	-2	1
-1.1	-4	3
-1	$-\infty$	$+\infty$
-.9	-4	4
-.8	-2	2
-.7	-1	1
-.6	-1	1
-.5	-1	0
-.4 to 1.9	0	0

* The table gives the range of values for the integer p within which V_p remains a positive semidefinite matrix.

We now investigate the limiting behavior as one samples the continuous time process more frequently. Let h denote the length of time between observation and suppose that

$$X(t) = x(ht)$$

integer values of t.[14] In this circumstance

$$B_o = \exp(hA_o)$$

$$W_o = \int_o^h \exp(A_o\tau)V_o\exp(A_o'\tau)\,d\tau.$$

The question is what happens to the number of observationally equivalent models as h gets small. Theorem 5 provides an answer to this question.

<u>Theorem 5</u>: If $R_o = T^{-1}V_oT^{-1'}$ does not have any zero elements, then there is an h^* such that for $h \leq h^*$ the model (A_o,V_o) is identified from (B_o,W_o).

The content of Theorem 5 is that except for singular cases, it is possible to sample the continuous time process at fine enough time intervals so that the aliasing problem vanishes. This result is in sharp contrast to what happens to identification in cases in which the underlying continuous time process is <u>a priori</u> restricted only to be covariance stationary. In the latter circumstance, there is an uncountable infinity of observationally equivalent models for <u>any</u> choice of h.

Summarizing our results in this section, we have shown that even in cases in which A_o has complex eigenvalues, equations (15) and (18) will in most circumstances have only a finite number of solutions and in many cases have only one solution. The upshot of this situation is that for certain values of the continuous time parameters (A_o, V_o), the identification problem can be much less drastic than was suggested by Phillips's characterization.

5. CONCLUSION

The preceding results provide a notion of the role of the prior assumption of a rational spectral density matrix, or a vector first order Markov process, in resolving the aliasing identification problem. Previously, it was known that in the general covariance stationary case the dimensionality of the identification problem was uncountably infinite; and it was believed that for the rational spectral density case, in particular, in the first-order Markov case, the dimensionality was countably infinite. Realizing that in this latter case the dimensionality is finite better indicates the relative contributions of the restriction to a rational spectral density matrix, on the one hand, and any additional prior restrictions such as exclusion or cross-equation restrictions, on the other hand, in achieving identification. The role of additional prior restrictions of various kinds in achieving unique identification is described in Phillips [9] and Hansen and Sargent [5].

APPENDIX

Here we prove the five theorems of Sections 2 and 3.

Proof of Theorem 1: Suppose that μ is an n dimensional complex vector. If r_k is the autocovariance function for a continuous time process, then $\bar{\mu}' r_k \mu$ is the autocovariance function for a complex valued one dimensional process. In particular, it will be the case that

$$\mathrm{Re}[\bar{\mu}' r_k(\tau)\mu] \leq \bar{\mu}' r_k(0)\mu$$

for any real τ where Re(s) is equal to the real part of s. Now

$$\begin{aligned}
&\mathrm{Re}[\bar{\mu}' Q_1 \mu \exp(\lambda_1\tau + 2\pi i k\tau) + \bar{\mu}'\bar{Q}_1\mu\exp(\bar{\lambda}_1\tau - 2\pi i k\tau)] \\
&= \exp(\alpha_1\tau)[\bar{\mu}'(Q_{11}+Q_{11}')\mu\cos(\alpha_2\tau+2\pi k\tau) \\
&\quad + \bar{\mu}'(Q_{12}'-Q_{12})\mu\sin(\alpha_2\tau+2\pi k\tau)]
\end{aligned}$$

where $\lambda_1 = \alpha_1+\alpha_2 i$ with α_1 and α_2 real. If $\bar{\mu}'(Q_{12}'-Q_{12})\mu$ is different from zero or if $\bar{\mu}'(Q_{11}+Q_{11}')\mu$ is negative, then

$$\begin{aligned}
\text{(A1)}\qquad &\mathrm{Re}[\bar{\mu}' Q_1 \mu \exp(\lambda_1\tau + 2\pi i k\tau) + \bar{\mu}'\bar{Q}_1\mu\exp(\bar{\lambda}_1\tau - 2\pi i k\tau)] \\
&= \exp(\alpha_1\tau)\psi\cos[\theta+(\alpha_2+2\pi k)\tau]
\end{aligned}$$

where $\psi>0$ and $\theta\varepsilon(0,2\pi)$. In fact, if either condition (i) or (ii) of Theorem 1 are not satisfied then there exists a μ for which the corresponding θ in (A1) is indeed in the open interval $(0,2\pi)$. We shall construct a two-sided sequence $\{\tau_k\}_{k=-\infty}^{+\infty}$ where $\tau_k > 0$ such that for sufficiently large $|k|$

(A2) $$\exp(\alpha_1 \tau)\psi\cos[\theta + (\alpha_2 + 2\pi k)\tau_k] > \psi\cos(\theta) + \delta$$

for some $\delta > 0$. Let

$$\tau_k = \frac{-\theta}{\alpha_2 + 2\pi k}$$

For sufficiently large $|k|$,

$$\exp(\alpha_1 \tau_k)\psi\cos[\theta + (\alpha_2 + 2\pi k)\tau_k] = \exp(\alpha_1 \tau_k)\psi > \exp(\alpha_1 \tau_k)\psi\cos(\theta).$$

Furthermore, $\tau_k \to 0$ as $|k| \to \infty$. Hence, for sufficiently large $|k|$

$$\exp(\alpha_1 \tau_k)\psi > \psi\cos(\theta) + \delta$$

for some positive δ. This verifies (A2). However, (A2) implies that for sufficiently large $|k|$

$$\mathrm{Re}[\bar{\mu}' r_k(\tau_k)\mu] > \bar{\mu}' r_k(0)\mu$$

contradicting the presumtpion that r_k is an autocovariance function for a continuous time process.

<u>Proof of Theorem 2</u>: Since (i) is satisfied for A^* it follows that

(A3) $$A^* = T\Lambda^* T^{-1}$$

where

$$\Lambda^* = \Lambda_o + 2\pi i P^*$$

and P^* is a diagonal integer matrix. At least two of the elements of P^* are not zero. Also

(A4) $$\exp A^* = T\exp\Lambda^* T^{-1} = T\exp\Lambda_o T^{-1}.$$

Following Phillips [17, page 354], from relations (ii) we can deduce an alternative relationship

(A5) $$\exp A^* V_o \exp A^{*\prime} - V_o = A^* W_o + W_o A^{*\prime} .$$

Substituting (A3) and (A4) into (A5) we have

(A6) $$T\exp\Lambda_o T^{-1} V_o T' \exp(\bar{\Lambda}_o') T - V_o = T\Lambda^* T^{-1} W_o + W_o T^{-1\prime} \bar{\Lambda}^* T'$$

Premultiplying and postmultiplying both sides of (A6) by T^{-1} and $T^{-1\prime}$, respectively, we obtain

(A7) $$\exp\Lambda_o R_o \exp\bar{\Lambda}_o - R_o = \Lambda^* S_o + S_o \bar{\Lambda}^*$$

where

$$R_o = T^{-1} V_o T^{-1\prime}$$
$$S_o = T^{-1} W_o \bar{T}^{-1\prime}.$$

Using analagous reasoning, it follows that

(A8) $$\exp\Lambda_o R_o \exp\bar{\Lambda}_o - R_o = \Lambda_o S_o + S_o \Lambda_o.$$

Equations (A7) and (A8) together imply that

$$(\lambda_j^o + \bar{\lambda}_\ell^o) s_{j,\ell}^o = (\lambda_j^* + \bar{\lambda}_\ell^*) s_{j,\ell}^o$$

where

$$[s_{j,\ell}^o] = S_o$$
$$\text{diag }(\lambda_1^o, \lambda_2^o, \ldots, \lambda_n^o) = \Lambda_o$$
$$\text{diag }(\lambda_1^*, \lambda_2^*, \ldots, \lambda_n^*) = \Lambda^*.$$

Let j_o be an index such that $\lambda_{oj}^o \neq {}_o\lambda_j^*$. By assumption, there must be at least one such j_o, since $\Lambda_o \neq \Lambda^*$. Let $J(j_o)$ be the set of indexes of eigenvalues which undergo identical perturbations, i.e., have the characteristic that $\lambda_j^* - \lambda_j^o = \lambda_{j_o}^* - \lambda_{j_o}^o$, or equivalently, $\lambda_{j_o}^o + \bar{\lambda}_j^o = \lambda_{j_o}^* + \bar{\lambda}_j^*$. Let $K(j_o)$ be the set of all indexes ℓ such that $\bar{\lambda}_\ell^o = \lambda_j^o$ where $j \in J(j_o)$. We define

(A9) $$\lambda_j^k = \begin{cases} \lambda_j^o + 2\pi i k & \text{if } j \in J(j_o) \\ \lambda_j^o - 2\pi i k & \text{if } j \in K(j_o) \\ \lambda_j^o & \text{otherwise.} \end{cases}$$

where k is an arbitrary integer. Now

(A10) $$(\lambda_j^o+\bar{\lambda}_\ell^o)s_{j,\ell}^o = (\lambda_j^k + \bar{\lambda}_\ell^k)s_{j,\ell}^o$$

for all j, ℓ = 1, 2,...., n. This follows from the fact that for any $j\varepsilon J(j_o)$ and $\ell\varepsilon J(j_o)$, the corresponding element $s_{j,\ell}^o = 0$. Let

$$A_k = T\Lambda_k T^{-1}$$

where

$$\Lambda_k = \text{diag}(\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k).$$

Using the definition (A9) and relation (A10), we conclude that (i) and (ii) are satisfied for $\{A_k\}_{k=1}^{\infty}$.

Proof of Theorem 3: Suppose to the contrary that there is an infinite sequence of distinct pairs $\{(A_k,V_k)\}_{k=1}^{\infty}$ that satisfy (i') and (ii'), and that there does not exist an $A^* = A_o$ satisfying (i) and (ii) of Theorem 2. We define

$$\Lambda_k = T^{-1}A_kT.$$

Since (i') is satisfied, it follows that Λ_k is diagonal and that

$$\Lambda_k-\Lambda_o = 2\pi i P_k$$

where P_k is a diagonal integer matrix. From Phillips [17] equation (4) it follows that for a given choice of A^* satisfying (i') there is one V^* that satisfies (ii'). Let

$$\Lambda_k = \mathrm{diag}(\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k).$$

Because the elements of the infinite sequence $\{A_k\}_{k=1}^{\infty}$ are distinct, there must exist a j_o such that $\{\lambda_{j_o}^k\}_{k=1}^{\infty}$ is unbounded. Let $J(j_o)$ be the set of indexes ℓ such that the sequence $\{\mu_\ell^k\}_{k=1}^{\infty}$ is bounded where

$$\mu_\ell^k = \lambda_{j_o}^k - \lambda_\ell^k.$$

Let $K(j_o)$ be the set of all indices ℓ such that $\{\theta_\ell^k\}_{k=1}^{\infty}$ is bounded where

$$\theta_\ell^k = \bar{\lambda}_{j_o}^k - \lambda_\ell^k.$$

Let

$$\lambda_j^* = \begin{cases} \lambda_j^o + 2\pi i & \text{if } j \in J(j_o) \\ \lambda_j^o - 2\pi i & \text{if } j \in K(j_o) \\ \lambda_j^o & \text{otherwise.} \end{cases}$$

Finally, let

$$\Lambda^* = \mathrm{diag}(\lambda_1^*, \lambda_2^*, \ldots, \lambda_n^*)$$
$$A^* = T\Lambda^* T.$$

We will show that A^* satisfies (i) and (ii) of Theorem 2.

First, note that A^* is real. This can be seen by verifying that if $j \varepsilon J(j_o)$ then λ_j^o is complex and the index ℓ of its complex conjugate $\bar{\lambda}_j^o = \lambda_\ell^o$ is contained in $K(j_o)$. Second, observe that $\exp A^* = B_o$. Third, note that if $j \varepsilon J(j_o)$ then $s_{j,\ell}^o = 0$ for all ℓ $J(j_o)$. This follows by letting

$$R_k = [r_{j,\ell}^k] = T^{-1} V_k T^{-1'}$$

and noting that since V_q satisfies (ii')

$$r_{j,j}^k = r_{j,j}^o$$
$$r_{\ell,\ell}^k = r_{\ell,\ell}^o$$

$$r_{j,\ell}^k = \frac{(\lambda_{k,j} + \bar{\lambda}_{k,\ell}) s_{j,\ell}^o}{\exp(\lambda_{o,\ell} + \bar{\lambda}_{o,\ell}) - 1}$$

for any choice of j, k = 1, 2, ..., n. Since (iii') is satisfied for V_k it follows that $r_{j,j}^k r_{\ell,\ell}^k - |r_{j,\ell}^k|^2 \geq 0$ for all k. Hence if $j \varepsilon J(j_o)$ and $\ell \not\varepsilon K(j_o)$, then $s_{j,\ell}^o = 0$. Fourth, observe that

$$r_{j,\ell}^o = r_{j,\ell}^*$$

for j, ℓ = 1, 2, ..., n where

$$r_{j,\ell}^* = \frac{(\lambda_j^* + \bar{\lambda}_\ell^*) s_{j,\ell}}{\exp(\lambda_j^o + \bar{\lambda}_\ell^o) - 1}$$

Finally, verify that $\int_0^1 \exp(A^* \tau) V_o \exp(A^{*\prime} \tau)\, d\tau = W_o$. This follows from the fact that

$$R^* = [r^*_{j,\ell}] = R_o = T^{-1} V_o T^{-1\prime}.$$

Thus we produced an $A^* \neq A_o$ that satisfies (i) and (ii) of Theorem 2. This contradiction proves Theorem 3.

<u>Proof of Theorem 4</u>: Suppose to the contrary that there exists an infinite sequence of distinct pairs $\{(A_k, V_k)\}_{k=1}^{\infty}$ that satisfy (i'), (ii') and (iii') of Theorem 3, and that $S_o = T^{-1} W_o T^{-1\prime}$ does not have any zero elements. As in the proof of Theorem 3, we know that

$$r^k_{j,j} = r^o_{j,j}$$
$$r^k_{\ell,\ell} = r^o_{\ell,\ell}$$

$$r^k_{j,\ell} = \frac{(\lambda_j + \bar{\lambda}^k_\ell)\, s^o_{j,\ell}}{\exp(\lambda^o_j + \lambda^o_\ell) - 1}$$

for all $j, \ell = 1, 2, \ldots, n$ and that $r^k_{j,j} r^k_{\ell,\ell} - |r^k_{j,\ell}|^2 \geq 0$ for all k. There exists an index j_o such that $\{\lambda^k_{o j}\}_{k=1}^{\infty}$ is unbounded and there exists an index ℓ_o such that $\lambda^k_{j_o} = \bar{\lambda}^k_{\ell_o}$. Therefore,

$$r^k_{j_o,\ell_o} = \frac{2\lambda_{j_o} s^o_{j_o,\ell_o}}{\exp(\lambda^o_{j_o}+\lambda^o_{\ell_o})-1} .$$

Since $\{r^k_{j_o,\ell_o}\}^{\infty}_{k=1}$ is unbounded, we have generated a contradiction.

Proof of Theorem 5: Define

(A10)
$$B_{o,h} = \exp(A_o h)$$
$$W_{o,h} = \int_o^h \exp(A_o \tau) V_o \exp(A_o' \tau) d\tau$$
$$S_{o,h} = T^{-1} W_{o,h} T^{-1'} = [s^{o,h}_{j,\ell}].$$

When the observation interval is h, the pair $(B_{o,h}, W_{o,h})$ can be inferred from discrete time data. Following the logic in the previous proofs,

$$s^{o,h}_{j,\ell,h} = \frac{\exp(h\lambda_{o,j}+h\bar{\lambda}_{o,\ell})-1}{(h\lambda_{o,j}+h\bar{\lambda}_{o,\ell})} r^o_{j,\ell}.$$

If there are not any complex conjugate pairs of eigenvalues, then the conclusion is immediate. Suppose there is one complex conjugate pair. Let λ_1 be complex and suppose λ_2 is its complex conjugate. By assumption $r^o_{1,2}$ is not zero. Hence $s^o_{1,2}$ is not zero. Let

$$\lambda_1^k = \lambda_1^o + k\frac{2\pi i}{h}$$

$$\lambda_2^k = \lambda_2^o - k\frac{2\pi i}{h}$$

$$A_{k,h} = T\,\Lambda_o \bar{T}^{-1} + \frac{2\pi i}{h} T \text{diag}\,\{k,\ -k,\ 0,\ldots 0\} T^{-1}.$$

The sequence $\{A_{k,h}\}_{k=-\infty}^{+\infty}$ contains all the solutions to

$$B_{o,h} = \exp(A^* h).$$

Let

$$R_{k,h} = T^{-1} V_{k,h} \bar{T}^{-1'}$$

where

$$W_{o,h} = \int_o^h \exp(A_{k,h}\tau) V_{k,h} \exp(A_{k,h}'\tau)\, d\tau.$$

Now

(A11) $$r_{1,2}^{k,h} = \frac{(h\lambda_{o,1} + h\bar{\lambda}_{o,2} + \frac{k4\pi i}{h})}{\exp(h\lambda_{o,1} + h\bar{\lambda}_{o,2}) - 1} s_{1,2}^{o,h}$$

$$= \frac{(\lambda_{o,1} + \bar{\lambda}_{o,2} + \frac{k4\pi i}{h})}{(\lambda_{o,1} + \bar{\lambda}_{o,2})} r_{1,2}^{o}$$

$$= r_{1,2}^{o} + \frac{4\pi k i}{h}\, r_{1,2}^{o}$$

For $V_{k,h}$ to be a legitimate intensity matrix, it must be that

$$r_{1,2}^o r_{2,2}^o - |r_{1,2}^{k,h}|^2 \geq 0.$$

However from |A11| it is clear that

$$\inf_{k \varepsilon \{1,-1,2,-2,\ldots\}} |r_{1,2}^{k,h}|$$

can be made arbitrarily large by making h go to zero. Thus there exists an h^* such that for $h < h^*$

$$r_{1,1}^{o} r_{2,2}^{o} - |r_{1,2}^{k,h}|^2 < 0$$

for all nonzero integer k. Hence there is no aliasing problem.

In cases in which there are more than one complex conjugate pair of eigenvalues the above argument can be repreated for each conjugate pair.

FOOTNOTES

1. Thanks are due to Ian Bain and Judy Sargent who calculated the numerical examples. P.C.B. Phillips and John Taylor provided some very useful comments on an earlier draft. This research was supported in part by NSF Grant SES-8007016.

2. While this paper ends up modifying Phillips's characterization of identification, the analysis of section 3 obviously builds upon the machinary that he developed [9].

3. For an introduction to continuous time stochastic processes, see Kwakernaak and Sivan [6]. A vector white noise w with generalized covariance function $Ew(t)w(t-\tau) = V\delta(t-\tau)$ is said to have "intensity matrix" V_0, where $\delta(\cdot)$ is the Dirac delta generalized function.

4. This statement formally is only true once we treat all elements of the equivalence class of matrix functions c that are equal almost everywhere as the same matrix function. See Rozanov [11] for a discussion of fundamental representations.

5. This notation emerges because the Laplace transform of the time derivative operator D is the function $H(s) = s$.

6. This can be proved directly by noting that for any real $\omega^* > \pi$, it is possible to construct a bandlimited continuous time process x^*, with its spectral density matrix zero for $|\omega| > \omega^*$, but nonzero for $|\omega| < \omega^*$. This process can be chosen to be observationally equivalent to x from discrete time data. Since $\{\omega^* > \pi\}$ is an uncountable set, the class of observationally equivalent x^* processes in uncountably infinite.

7. A.W. Phillips [10] has studied the following subset of continuous time processes with rational spectral densities, those that can be represented as

$$D^p x(t) + a_1 D^{p-1} x(t) + \dots + a_p x(t) = b_1 D^{p-1} w(t) + b_2 D^{p-2} w(t) + \dots + b_p w(t).$$

We use the same set of tools as Phillips used to analyze the discrete time implications of continuous time stochastic processes. Phillips, however, did not address the aliasing problem.

8. The restrictions on the real parts of $\lambda_1, \lambda_2, \ldots, \lambda_q$ are sufficient to guarantee that the inverse Laplace transform of C is concentrated on the nonnegative real numbers. The restrictions on the zeroes of the detG(s) are sufficient to guarantee that the associated one-sided moving average representation is fundamental.

9. For the class of processes considered by A.W. Phillips [10], the matrices Q_j have rank one. For the class of processes considered here, these matrices are permitted to have ranks that exceed one.

10. This discussion uses the fact that the complex logarithm has multiple branches. In particular, if s is a complex variable then $\log(s) = \log|s| + \arg(s) + 2\pi ik$ for some integer k. The integer k indexes the multiple branches of the logarithmic function.

11. When there is more than one pair of complex conjugate poles, the possibility is raised of generating a countable sequence of observationally equivalent continuous time models by adding different integer multiples of $2\pi i$ to several of the complex conjugate pairs simultaneously. In the following section, this possibility is studied in detail for a special case of model (7).

12. A question related to this one occurs in determining whether the parameters of a continuous time Markov process can be inferred from Markov chain probabilities. This question has been treated, for example, by Singer and Spilerman [13]. They also consider the circumstances under which a Markov chain transition matrix can be embedded in a continuous time Markov process. There is an analogous question in our setting that asks: for which values (B_o, W_o) does there exist a pair (A_o, V_o) that satisfies (13) and (14)?

13. P.C.B. Phillips has correctly pointed out to us that if V_o is assumed to be singular, then when W_o is positive definite there is extra identifying information about A_o contained in W_o. (Phillips [7] has characterized cases in which W_o is nonsingular even though V_o is singular.) The argument which follows in the text extends Phillips's point by establishing that even if V_o is permitted to be nonsingular, W_o in general contains identifying information about A_o.

14. This discussion follows Phillips [9] and allows the interval of time between observations, h, to be different from unity. Phillips suggested to us that we investigate the identification problem as h goes to zero.

REFERENCES

[1] Coddington, E. A. and N. Levinson: Theory of Ordinary Differential Equations, New York: McGraw-Hill, 1955.

[2] Gantmacher, F. R.: The Theory of Matrices, Vol. I, New York: Chelsea, 1959.

[3] Geweke, J. B.: "Temporal Aggregation in the Multivariate Regression Model," Econometrica, 46 (1978), 643-662.

[4] Hansen, L. P. and T. J. Sargent: "Methods for Estimating Continuous Time Rational Expectations Models from Discrete Time Data," unpublished manuscript, 1980.

[5] Hansen, L. P. and T. J. Sargent: "Identification of Continuous Time Rational Expectations Models from discrete Time Data," unpublished manuscript, 1981.

[6] Kwakernaak, H. and R. Sivan: Linear Optimal Control Systems, New York: Wiley, 1972.

[7] Phillips, P.C.B.: "The Estimation of Some Continuous Time Models," Econometrica, 42 (1974), 803-824.

[8] Phillips, P.C.B.: "The Structural Estimation of a Stochastic Equation System," Econometrica, 40 (1972), 1021-1041.

[9] Phillips, P.C.B.: "The Problem of Identification in Finite parameter Continuous Time Models," Journal of Econometrics, 1 (1973), 351-362.

[10] Phillips, A.W.: "The Estimation of Parameters in Systems of Stochastic Differential Equations," Biometrika, 59 (1959), 67-76.

[11] Rozanov, Y. A.: Stationary Random Processes. San Francisco: Holden-Day, 1967.

[12] Sims, C. A.: "Discrete Approximations to Continuous Time Lag Distributions in Econometrics," Econometrica, 39 (1971), 545-564.

[13] Singer, B. and S. Spilerman: "The Representation of Social Processes by Markov Models," American Journal of Sociology, 82 (1976), 1-54.