

AD-A046 046

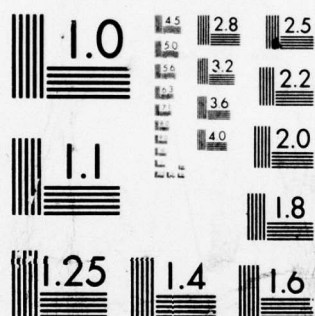
HARVARD UNIV CAMBRIDGE MASS DIV OF ENGINEERING AND --ETC F/6 12/1
STATIONARY COVARIANCE GENERATION WITH FINITE STATE MARKOV PROCE--ETC(U)
1974 R W BROCKETT N00014-75-C-0648

UNCLASSIFIED

| OF |
AD
A046046



END
DATE
FILMED
11-77
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD A 046046

DDC FILE COPY

14
TA26 - 12:30

STATIONARY COVARIANCE GENERATION WITH FINITE STATE MARKOV PROCESSES,

11 1974
12 4p.
10 ROGER W. BROCKETTDivision of Engineering and Applied Physics
Harvard University
and
Center for Mathematical System Theory
University of FloridaDDC
NOV 2 1977
F

Abstract

In this paper we study the stationary covariance generation problem, i.e. the problem of passing from a stationary covariance function to a dynamical system which generates a process having the given covariance, in the case where the dynamical system is a finite state, continuous time, Markov process. We find that strictly positive definite stationary covariances can be approximated to any degree of accuracy in this way. However the number of states required may approach infinity as the covariance approaches the boundary of the set of positive definite functions.

1. Introduction

The general area of stochastic realization theory consists of results on how to pass from a certain complete, or partial, description of a stochastic process to a differential equation (or difference equation) model for it with the stochastic process which drives the differential equation being simpler, or more "fundamental", than the original process. In the case where the given partial description is a stationary covariance ϕ and of a constant mean m it has been known for some time under what circumstances there exists a realization in the form of a finite dimensional linear Ito equation

$$\begin{aligned} dx(t) &= Ax(t)dt + bdw(t) + rdt; \quad y(t) = cx(t) \\ E\dot{y}(t) &= m \\ \phi(t, \tau) &= \phi(\tau - t) = E(y(t) - m)(y(\tau) - m) \end{aligned}$$

with constant coefficients. This result, which has its origin in the Bode-Shannon pre-whitening filter, states that ϕ is realizable if it is an even, $L_2(-\infty, \infty)$ function whose Fourier transform (power spectrum) is rational and nonnegative on the real axis. One then finds the system parameters A, b, c and r by a factorization of the power spectrum. Since the power spectrum is automatically nonnegative this result is quite satisfactory; the only price being paid for finite dimensionality of x is the rationality of the power spectrum.

In this paper we consider covariance generation but with a view toward generating a given

This work was supported in part by the U.S. Office of Naval Research under the Joint Services Electronics Program Contract (N00014-75-C-0648) and the Army Research Office under Grant DAAG29-76-C-0139.

covariance from a finite state continuous time Markov process. Specifically we ask the following.

Question: Given a stationary covariance $\phi(t, \tau)$ under what circumstances does there exist a finite state continuous time Markov process $x(\cdot)$ taking on values in a finite set X and a function $f: X \rightarrow \mathbb{R}$ (the real numbers) such that

- The transition probabilities are given by a time invariant infinitesimal generator whose null space is one dimensional.
- $\lim_{t \rightarrow \infty} E f[x(t)] = 0$ for all $x(0)$
- $\lim_{t \rightarrow \infty} E f[x(t)]f[x(t+\tau)] = \tilde{\phi}(\tau)$ for all $x(0)$

The reason for assuming that A has a one dimensional kernel is so that regardless of the initial state of the x -process, the steady state density is the same.

In this paper we derive necessary and sufficient conditions for ϕ to be realizable in this way. Roughly speaking, what we find is that any non-periodic stationary covariance can be approximated by one which is realizable in this way. Moreover those covariances with rational power spectra which can be realized exactly are characterized.

Since one knows [4] that any finite state continuous time Markov process is equivalent to one which can be, expressed as

$$dx(t) = \sum_{i=1}^{\delta} A_i x(t) dN_i(t); \quad x(t) \in \mathbb{R}^n$$

with $N_i(t)$ a standard Poisson counting process with rate λ_i , this work is a natural complement to the Gauss-Markov covariance generation problem.

2. The Model

We find it convenient to associate with each state of an n state Markov process a particular point in \mathbb{R}^n . This lets us visualize the process as jumping between points in a vector space and allows us to use certain very familiar formulas from linear system theory. Let e_i be the i th standard basic element in \mathbb{R}^n (prime denotes transpose)

$$e_i = [0, 0, \dots, 0, 1, 0, \dots, 0]'$$

↑
ith coordinate

and let x be a process which takes on the values in the set $\{e_1, e_2, \dots, e_n\}$. If $p_i(t)$ is the probability

that $x(t)=e_i$ then our assumptions imply that there exists a constant matrix A such that

$$\begin{bmatrix} \dot{p}_1(t) \\ \dot{p}_2(t) \\ \vdots \\ \dot{p}_n(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} p_1(t) \\ p_2(t) \\ \vdots \\ p_n(t) \end{bmatrix}$$

Notice that because of the way we have embedded the states in \mathcal{R}^n we have

$$p(t) = Ex(t)$$

According to the well known theory of such processes, see, e.g. Karlin [1], the entries of A satisfy the condition

$$(i) \quad a_{ij} \begin{cases} \geq 0 & \text{if } i \neq j \\ \leq 0 & \text{if } i = j \end{cases}$$

$$(ii) \quad \sum_{i=1}^n a_{ij} = 0$$

Such matrices are called here infinitesimally stochastic. Of course Peron-Frobenius theory implies that A has a nontrivial null space. If we ask that the null space be one dimensional then we are assured that there is a unique steady state probability distribution. For these reasons we will assume

(iii) the kernel of A is one dimensional.

Such processes are called irreducible.

As an immediate consequence of the definitions we see that

$$Ex(t) = e^{At} x(0) \quad (E)$$

Because the i th and j th components of x are never simultaneously nonzero and because the components take on only the values zero and one

$$\Sigma(t) = Ex(t)x'(t) = \begin{bmatrix} p_1(t) & 0 & \dots & 0 \\ 0 & p_2(t) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & p_n(t) \end{bmatrix}$$

Moreover an elementary application of Bayes' rule yields

$$Ex(t)x'(\tau) = \Sigma(t)e^{A'(\tau-t)}; \quad \tau > t \quad (V)$$

The equations (E) and (V) together with one more remark will yield a restatement of our question in linear algebraic form. The additional remark is this. Any map of the state space of the Markov process into \mathcal{R}^n is of the form $c'x$ for some c in \mathcal{R}^n . Simply let the i th component of c be $f(e_i)$ where $f: X \rightarrow \mathcal{R}$. Bringing these remarks together we have the following theorem.

Theorem 1: A stationary covariance $\phi(t-\tau)=\phi(\tau,t)$ is realizable by a zero mean process which is a function of a finite state, continuous time irreducible Markov process if and only if

$$\tilde{\phi}(t) = c' \Sigma e^{At} c; \quad t > 0 \quad (S)$$

for some pair (A,c) where A is an n by n matrix satisfying (i), (ii), (iii), c is an n vector such that $c'x=0$ for all x in the kernel of A and

$$\Sigma = \begin{bmatrix} p_1 & 0 & \dots & 0 \\ 0 & p_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & p_n \end{bmatrix}$$

with (p_1, p_2, \dots, p_n) being the probability vector in the kernel of A .

3. Realization with A Circulant

The determination of what functions can be expressed in the form required by equation (S) is made especially hard by the requirement that A be infinitesimally stochastic. To get around the awkwardness of this constraint we focus attention on a special class of matrices. By a circulant matrix we understand a square matrix of the form

$$M = \begin{bmatrix} m_0 & m_1 & m_2 & \dots & m_{n-1} \\ m_{n-1} & m_0 & m_1 & \dots & m_{n-2} \\ \dots & \dots & \dots & \dots & \dots \\ m_1 & m_2 & m_3 & \dots & m_0 \end{bmatrix}$$

Associated with each such M there is a polynomial $\hat{m}(z) = m_0 + m_1 z + \dots + m_{n-1} z^{n-1}$. The eigenvalues of a circulant matrix are simply the values

$$\lambda_k = \hat{m}(e^{ik\theta}); \quad \theta = \frac{2\pi}{n}; \quad k=0,1,\dots,n-1$$

Thus M meets condition (i), (ii), (iii) if and only if

(i') the coefficients of $\hat{m}(z)$ are all nonnegative except for the constant term.

(ii') $\hat{m}(1) = 0$

(iii') $\hat{m}(z)$ does not vanish for z an n th root of unity unequal to one.

Under these assumptions one sees easily that the solution of

$$\dot{p} = Mp$$

for $p(0)$ a probability vector, tends to

$$p_0 = \frac{1}{n} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

Thus using such an M we see that the matrix Σ in equation (S) becomes $\frac{1}{n} I$ and we have

$$\tilde{\phi}(t) = \frac{1}{n} c' e^{Mt} c; \quad t > 0$$

The condition that the mean should vanish, $c'p_0=0$, is also easily interpretable. In fact if

$$c' = (c_0, c_1, \dots, c_{n-1})$$

we introduce

$$\hat{c}(z) = c_0 + c_1 z + \dots + c_{n-1} z^{n-1}$$

In this notation $c'p_0 = 0$ becomes

$$(b') \quad \hat{c}(1) = 0$$

We can also express $\tilde{\phi}$ succinctly in terms of $m(z)$ and $c(z)$ (see [3]).

$$\tilde{\phi}(t) = \sum_{\{z: z^n=1\}} c(z)c(z^{-1})e^{m(z)t}; \quad t > 0$$

The following lemma gives a somewhat more satisfactory form of this.

Lemma 1: The set of finite state continuous time realizable covariances includes those covariances expressible as

$$\tilde{\phi}(t) = \sum_{k=1}^{n-1} r_k e^{m(e^{2\pi i k/n})t} ; t \geq 0$$

with m satisfying (i'), (ii') and (iii') and the r_k real and nonnegative. In particular

$$\psi(t) = re^{-\sigma t} \cos \omega t ; t \geq 0$$

is so realizable if r and σ are real and positive and ω is real.

Proof: Of course $c(z)c(z^{-1})$ is, for z on the unit circle, real and nonnegative. Since $c(1)=0$ we must have $c(z)c(z^{-1})$ vanishing at $z=1$ but otherwise we may pick the coefficients so that $c(z)$ has arbitrary complex values at the n th roots of unity consistent with $c(z) = \overline{c(\bar{z})}$. Adding up the contribution from $\rho = e^{2\pi i k/n}$ we get

$$\tilde{\phi}_k(t) = 2|c(\rho)|^2 e^{\operatorname{Re} m(\rho)t} \cos \operatorname{Im} m(\rho)t$$

But since $c(\rho)$ is arbitrary we see that $|c(\rho)|^2$ can be any real nonnegative number. The general form given in the lemma then follows.

To show that the specific ψ given in the lemma is expressible in this way we make a particular choice of n and $m(z)$. Let

$$m(z) = a(1-\alpha z - (1-\alpha)z^2) ; a, \alpha, > 0$$

and let $\sigma > 0$ and ω be given. At $z = e^{i2\pi/n}$ $\cos(2\pi/n) + i \sin(2\pi/n)$ the ratio of the real to the imaginary parts of m is

$$\gamma = \frac{1-\alpha \cos(2\pi/n) - (1-\alpha) \cos(4\pi/n)}{-\alpha \sin(2\pi/n) - (1-\alpha) \sin(4\pi/n)}$$

Inspection of this equation shows that for any negative γ we can choose an integer n large enough so as to have a solution for α . (As γ approaches zero $n(\gamma)$ goes to infinity.) Thus we can, with this choice of $m(z)$ adjust the magnitude and argument of $m(e^{2\pi i/n})$ as needed to get the function $\psi(\cdot)$ of the lemma. Of course we pick c in such a way as to vanish on all n th roots of unity except the two which enter in this discussion.

4. The Approximation Lemma

Lemma 1 makes it clear how to realize non-negative linear combinations of the basic terms labeled there ψ ; take the direct sum of realizations of the type constructed in its proof. However these realizations

- (a) will not, in general, satisfy the irreducibility condition, and
- (b) give no suggestion as to how to realize covariances such as $\phi(t) = 10e^{-|t|} - 5|t|$ which are differences of positive definite functions but still positive definite.

In this section we establish the results necessary to get around these difficulties.

If $\phi(t)$ is a continuous, even, positive definite function then according to the well known representation theorem of Bochner it can be expressed as

$$\phi(t) = \int_{[0, \infty)} \cos \omega t d\mu$$

for some nonnegative measure μ . Of course if μ is absolutely continuous with respect to Lebesgue measure then we can write

$$\phi(t) = \int_0^\infty \cos \omega t \phi(\omega) d\omega ; \phi(\omega) \geq 0$$

displaying the power spectrum explicitly.

However, if we assume that ϕ is not only positive definite but in addition it is strictly positive definite in the sense that

$$\phi_\epsilon(t) = \phi(t)e^{2\epsilon|t|}$$

is for, some $\epsilon > 0$, also square integrable and positive definite then we can express ϕ as

$$\phi(t) = \int_0^\infty e^{-\epsilon|t|} \cos \omega t \phi(\omega) d\omega$$

with ϕ analytic. (This follows from Paley-Wiener theory; the Fourier transform of

$\phi(t)e^{\epsilon|t|}$ analytic in a strip of width 2ϵ centered on the ω -axis and ϕ is its Fourier transform.) Let $\{t_i\}_{i=1}^r$ be any finite set of real numbers. We can approximate simultaneously the integrals $\phi(t_i)$ by Riemann sums and thus obtain

$$\phi(t) = \sum_{i=1}^m e^{-\epsilon t} \cos \omega_i t \phi(\omega_i) s_i + \epsilon(t)$$

with $|\epsilon(t_i)|$ less than any preassigned positive number. Both ϕ and the approximation go to zero as $|t| \rightarrow \infty$; in view of the continuity of ϕ we see that it can be uniformly approximated by a linear combination of ψ -like terms with positive coefficients. The following theorem summarizes.

Theorem 2: Any continuous, strictly positive definite function is the uniform limit of a sequence ϕ_n of the form

$$\phi_n(t) = \sum_{k=1}^{m(n)} \alpha_k e^{-\epsilon t} \cos \omega_k(t) ; \alpha_k > 0$$

We now address the second problem mentioned at the start of this section.

Lemma 2: Stationary covariances of the form appearing in theorem 2 can be realized by a pair (A, c) satisfying conditions (i), (ii), and (iii).

Proof: Let (A_k, c_k) be a realization of $\alpha_k e^{-\epsilon t/2} \cos \omega_k t$ of the form given in the proof of lemma 1; i.e. of the circulant form. (Note we have $\epsilon/2$ as the decay factor.) Then for

$$A = \begin{bmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & A_n \end{bmatrix} ; c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

it follows that

$$c^t e^{At} c = \sum_{k=1}^m \alpha_k e^{-\epsilon t/2} \cos \omega_k t$$

Now subtract from A the infinitesimally stochastic matrix

$$F = \frac{\varepsilon}{2} I - \frac{\varepsilon}{2n} \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 1 \end{bmatrix}$$

where n is the sum of the dimensions of the A_i . Clearly $A-F$ is infinitesimally stochastic and irreducible. Notice that

$$c'e^{(A-\varepsilon I/2)t}c = \sum_{k=1}^m \alpha_k e^{-\varepsilon t} \cos \omega_k t$$

(This time the decay factor agrees with theorem 2.) Finally, because of the null spaces of each of the A_i we see that

$$A \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 1 \end{bmatrix} = 0 ;$$

and

$$c' \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 1 \end{bmatrix} = 0$$

Together these imply $c'e^{(A-F)t}c = c'e^{(A-\varepsilon I/2)t}$ and so $(A-F, c)$ meets all requirements.

5. On the Number of States

In response to a question raised by Kolmogorov, about 30 years ago Dimitriev and Dynkin established that an n by n stochastic matrix cannot have any eigenvalues in the part of the unit circle bounded by a chord passing through the point $(1,0)$ in the complex plane and making angles of $(2\pi/p) + \pi/2$ with the real axis (see figure 1b) provided $p \geq n$ is an integer (see [2] for references). Since the exponential of an infinitesimal

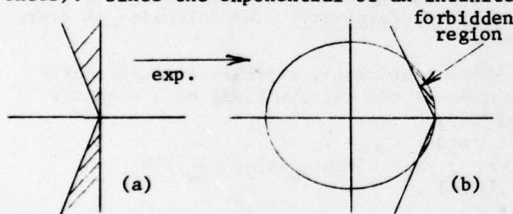


Figure 1: The infinitesimal version of the Dimitriev-Dynkin result.

stochastic matrix with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ is a stochastic matrix with eigenvalues $\exp \lambda_1, \exp \lambda_2, \dots, \exp \lambda_n$ this imposes certain constraints on the eigenvalues of an infinitesimally stochastic matrix. The space of infinitesimally stochastic matrices is a cone and so the excluded subset of the complex plane must be bound by straight lines passing through zero (see figure 1a). Some arithmetic establishes that $\sigma + i\omega$ cannot be an eigenvalue of an n by n infinitesimally stochastic matrix unless

$$\frac{\sigma}{\omega} \leq \frac{2n}{\pi} \ln \cos \frac{\pi}{2n} \quad (A)$$

Theorem 3: The power spectrum associated with the steady state covariance matrix of n -state continuous time Markov process is rational and has no poles in the region of the complex plane for which inequality (A) is violated.

References

- [1] S. Karlin, A First Course in Stochastic Processes, Addison-Wesley, Reading, Ma. 1966.
- [2] M. Marcus and H. Mink, Matrix Theory and Matrix Inequalities, Allyn and Bacon, Boston, Mass. 1964.
- [3] R.W. Brockett and Jacques Willems, "Discretized Partial Differential Equations: Examples of Control Systems Defined on Modules," Automatica, Vol. 10, pp. 507-515, Pergamon Press, 1974.
- [4] G. Blankenship and R.W. Brockett, to appear.

ACCESSION for	
NTIS	White Section <input checked="" type="checkbox"/>
DDC	Buff Section <input type="checkbox"/>
UNANNOUNCED	<input type="checkbox"/>
JUSTIFICATION	
BY	
DISTRIBUTION/AVAILABILITY CODES	
DI	SPECIAL
A SECRET	