Proceedings of the 37th IEEE
Conference on Decision & Control
Tampa, Florida USA • December 1998

Stationary Covariance Realization with a Specified Distribution of Amplitudes

Roger Brockett
Harvard University
brockett@hrl.harvard.edu

Abstract

The signals one encounters in examining image intensity data seldom appear to be even approximately Gaussian and as a consequence Gauss-Markov filtering theory, which vision researchers have found to be so useful in tracking and road following, has not been of much value in understanding the basic science involved in developing low level vision algorithms. In this paper we propose a new methodology for stochastic modeling which allows one to explore a class of models better fitted to the distribution of values taken on by the data while maintaining the ability to fit the autocorrelation function.

1 Introduction

Consider stochastic processes y generated by Itô differential equations of the form

$$dx = Axdt + Bdw$$

$$dz = \sum_{i=1}^{m} \phi_i(z)dN_i$$

$$y = c(z)dx + h(z)$$

Here w is vector-valued Browian motion and $N_1, ...N_m$ are Poisson counters of rates $\lambda_1, ...\lambda_m$, respectively. We assume that the functions ϕ_i are chosen, together with the allowable values of $z(0)$, in such a way as to have $z(t)$ evolve without leaving a finite set. The problem to be addressed is this.

Problem: Given a set of observed, stationary signals r_i, defined on an interval $(-\infty, \infty)$ suppose that the associated (empirical) autocorrelation is ψ and that the probability distribution for the values of $r_i(t)$ is $\rho(r)$, independent of the value of t. Find a model of the above class which matches exactly, or closely approximates, both the autocorrelation function and the probability density of the values of r. There are constraints implicit here. Of course the autocorrelation function must be a positive definite function and if the mean of the amplitude distribution is zero then

$$\psi(0) = \int_{-\infty}^{\infty} r^2 \rho(r) dr$$

Of course the standard solution of the covariance generation problem in terms of Gauss-Markov processes solves this problem in the very special case for which ρ is Gaussian. Such results are discussed in many textbooks. In an earlier paper [2] the author showed that one could achieve an arbitrarily good approximation to the power spectrum using finite state processes instead of Gaussian processes but in that work the probability distribution of the process being constructed, although certainly not Gaussian, was determined indirectly by the correlation function rather than being fitted to the distribution of the observed data. In both these situations one may say that the distribution of values of the process was sacrificed in favor of the shape of the autocorrelation function.

Although in many situations it is not useful to approximate a random variable by one whose density is Gaussian, a special case of an approximation theory result established by Wiener (see [1]) tells us that it is possible to approximate probability densities on the real line by convex combination of Gaussians. We will uses this fact below. The basic idea to be developed here may be introduced with the help of the following example. As in [2], we will make use of a continuous-time jump process taking on values on values in a set of orthonormal vectors. We take these to be the unit vectors $\{e_1, e_2, ...e_n\}$ in some n-dimensional Euclidean space.

Example: Suppose, that the given data $\{r_i(t)\}$ has a bimodal distribution with peaks at r_1 and r_2 and that the associated empirical probability distribution can be approximated by a sum of two Gaussians centered at these peaks. That is, consider

$$\rho_e(r) = \frac{1}{2\sqrt{2\pi\sigma_1}} e^{(r-r_1)^2/2\sigma_1} + \frac{1}{2\sqrt{2\pi\sigma_2}} e^{(r-r_2)^2/2\sigma_2} +$$

Consider a model whose state space is a product space, $S = \{e_1, e_2\} \times R^2$. We take the equations of evolution

[1]This work is supported in part by Army DAAH 04-93 -G-0330, U Brown Army DAAH 04-96-1-0445, Army DAAG 55 97 1 0114, MIT Army DAAL03-92-G-0115

to be

$$dx_1 = -a_1 dx_1 + b_1 dw_1$$

$$dx_2 = -a_2 dx_2 + b_2 dw_2$$

$$dz = \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix} z dN_{12} + \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} z dN_{21}$$

$$y(t) = \langle h, z(t) \rangle + x^T(t) C z(t)$$

and we ask that $z(0) \in \{e_1, e_2\}$.

These equations imply that the jump process z jumps between the value e_1 and e_2. If $p_1(t)$ is the probability that $z(t) = e_1$ and if $p_2(t)$ the probability that $z(t) = e_2$ then the infinitesimal generator of the process is such that

$$\frac{d}{dt}\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} -\alpha & \beta \\ \alpha & -\beta \end{bmatrix}\begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$$

where α and β are the rates of the Poisson Counters N_{12} and N_{21}, respectively. Because of the particular values that z assumes, we have

$$\mathcal{E}z(t) = p(t) \ ; \ \mathcal{E}z(t)z^T(t) = \begin{bmatrix} \frac{\beta}{\alpha+\beta} & 0 \\ 0 & \frac{\alpha}{\alpha+\beta} \end{bmatrix}$$

The steady state probability distribution for p is just

$$p_\infty = \begin{bmatrix} p_1(\infty) \\ p_2(\infty) \end{bmatrix} = \begin{bmatrix} \frac{\beta}{\alpha+\beta} \\ \frac{\alpha}{\alpha+\beta} \end{bmatrix}$$

Matters being so, if a_1 and a_2 are negative so that the x processes tends to a unique steady state, then in steady state the expected values of y and y^2 are given by

$$\lim_{t\to\infty} \mathcal{E}y(t) = \langle p_\infty, h \rangle$$

$$\lim_{t\to\infty} \mathcal{E}y^2(t) = \mathrm{tr} h h^T p_\infty p_\infty^T + \mathrm{tr}\mathcal{E}C^T x(t) x^T(t) C D$$

where

$$D = \mathrm{Diag}\ (p_\infty p_\infty^T)$$

Moreover, we see easily that

$$\lim_{t\to\infty} \mathcal{E}x(t)x^T(t) = \begin{bmatrix} \frac{b_1}{2a_1} & 0 \\ 0 & \frac{b_2}{2a_2} \end{bmatrix}$$

In making these calculations we have used the fact that the Poisson and Wiener processes appearing are all independent.

The independence of the underlying processes allows us to compute the steady state value of the expectation of

$$\psi(\tau) = \lim_{t\to\infty} \mathcal{E}y(t)y(t+\tau)$$

In fact,

$$\psi(\tau) = \lim_{t\to\infty} \mathcal{E}Cz(t)z(t+\tau)(1 + \mathrm{tr}C^T \mathcal{E}x(t)x^T(t+\tau))$$

To be even more specific, if we let C be the identity, let $a_1 = -1$, $a_2 = -2$, $b_1 = 2$, and $b_2 = 4$ then the steady state values of the variances for the processes x_1 and x_2 are both one. If we let $h_1 = -h_2 = 1$ and let $\alpha = 1/2, \beta = 1/2$ we see that the expected value of z is zero and that the covariance of z is $e^{\tau|}$.

$$\lim_{t\to\infty} \mathcal{E}y(t)y(t+\tau) = e^{-|\tau|}(1 + e^{-|\tau|} + e^{-2|\tau|})$$

The density of the amplitude of $y(t)$ is the sum of two Gaussians, each of variance one, centered at -1 and +1, respectively. That is,

$$\rho(y) = \frac{1}{2\sqrt{2\pi}} e^{-(x-1)^2/2} + \frac{1}{2\sqrt{2\pi}} e^{-(x+1)^2/2}$$

In the standard linear theory one can not get nontrivial sums of Gaussians.

2 General Theory

The example just treated brings out the important aspects of a more general theory. However, with some additional generality we can introduce many Gaussians, via many decoupled linear systems, thus allowing the approximation of arbitrary amplitude densities. If the approximation involves ν Gaussians,

$$\rho(r) = \sum_{i=1}^{\nu} \frac{\mu_i}{\sqrt{2\pi\sigma_i}} e^{-(r-r_i)^2/2\sigma_i}$$

then we will take a finite state system with at least ν states. To get the amplitude density correct one can let $h_i = r_i$ and chose the coefficients of the linear systems such that the i^{th} system has variance σ_i. The coefficients μ_i are then fixed by fixing the steady state value of the probability vector associated with z. Of course this can be any set of nonnegative numbers that sum to one and in our setting these weights will be fixed by fixing the counting rates of the Poisson counters.

Recall that for a linear, time-invariant, stable system characterized by the triple (A, B, C), the associated stochastic system

$$dx = Axdt + Bdw \ ; \ y = Cx$$

has steady state statistics

$$\mathcal{E}x = 0$$

$$\mathcal{E}xx^T = \Sigma_{xx} \ ; \ \Sigma_{xx}A^T + A\Sigma_{xx} = -BB^T$$

and also

$$\mathcal{E}x(t)x^T(t+\tau) = \Sigma_{xx} e^{A^T\tau} \ ; \ \tau \geq 0$$

On the other hand, for a time-invariant continuous time jump process defined on the unit vectors in E^n, the corresponding calculation involves the infinitesimal generator of the jump process. If the probability vector evolves according to $\dot{p} = Fp$, and if F is irreducible then it too has a unique steady state and in steady state

$$\mathcal{E}z(t)z^T(t) = \Sigma_{zz} Diag(p_1(\infty), ...p_n(\infty))$$

Using this we can write a linear differential in the variable τ and solve to get

$$\mathcal{E}z(t)z^T(t+\tau) = \Sigma_{zz} e^{F^T \tau} \; ; \; \tau \geq 0$$

Theorem: If y is generated by the stochastic equation

$$dx = Axdt + Bdw$$
$$dz = \sum_{i=1}^{m} \phi_i(z) dN_i$$
$$y = \langle h, z \rangle + x^T Cz$$

then under the assumption that infinitesimal generator for the probabilities for z is irreducible, and the assumption that the matrix A has eigenvalues with negative real parts, we can assert that y tends to a stationary process whose probability distribution is the sum of Gaussians and whose power spectrum is rational. Moreover, through choice of parameters in this setting (including the rates of the Poisson counters) one can approximate arbitrarily well any given probability distribution for the values of y.

Sketch of Proof: Again, using the independence of x and z we see that in steady state

$$\mathcal{E}x = 0$$
$$\mathcal{E}z = p_\infty$$
$$\mathcal{E}xx^T = \Sigma_{xx} \; ; \; \Sigma_{xx}A^T + A\Sigma_{xx} + BB^T = 0$$
$$\mathcal{E}zz^T = \Sigma_{zz} = \text{Diag} p_\infty P_\infty^T$$

and finally, for any set of indices i, j, k, l the independence lets us assert that

$$\mathcal{E}x_i x_j z_k z_l = (\mathcal{E}x_i x_j) \cdot (\mathcal{E}z_k z_l)$$

Using this we see that in steady state

$$\mathcal{E}C^T x(t) x^T(t) C z(t) z^T(t) = C^T \Sigma C \text{Diag} p_\infty p_\infty^T$$

If we introduce $J(\tau)$, defined as

$$J(\tau) = \lim_{t \to \infty} x(t) x^T(t+\tau)$$

then

$$\frac{dJ}{d\tau} = AJ(\tau) + J(\tau)F^T$$

so that

$$J(\tau) = e^{A\tau} J(0) e^{F^T \tau}$$

This allows us to see that the autocorrelation function has a time dependence that is expressible in terms of exponentials and thus that the power spectrum is rational. Appealing to the Wiener approximation theorem as suggested in the first paragraph of this section, we see that we can approximate an arbitrary amplitude density.

3 Further Remarks

We have not shown that our class of models is rich enough to solve the problem of jointly specifying the amplitude density and the covariance in all cases where this problem has a solution. What we see from the general development is that the time dependencies in the covariance function are of the form $e^{-\lambda_i t}$ with the λ_i either being an eigenvalues of the infinitesimal generator of the z process or else being the sum of the one of these eignevalues with an eigenvalue of A. Using the Bochner representation theorem, it was shown in [2] that it is possible to approximate an arbitrary positive definite autocorrelation function without using the Gauss-Markov part at all. However, if the autocorrelation has lightly damped exponentials then it is known to be necessary to use a very high dimensional system. It seems quite likely that in the present context autocorrelation functions that have only that have only lightly damped modes and, at the same time, multimodal amplitude densities will require high dimensional realizations. For example, it seems impossible to realize $e^{-t} \cos 20t$ as a autocorrelation function, given that the amplitude distribution is massed at -1 and 1. On the other hand, good approximations may exist.

4 References

1. N. I. Achieser, *Theory of Approximation*, Fredrick Ungar Pub. New York, 1956.
2. R. W. Brockett, "Stochastic Realization Theory and the Planck Law for Black Body Radiation," *Ricerche di Automatica*, Vol. 10 (1979) pp. 344-362.
3. R. W. Brockett, "Stationary Covariance Generation with Finite State Markov Processes," *Proceedings 1977 Joint Automatic Control Conference.*