Note di Matematica Vol. X, Suppl. n. 1, 167-172(1990)

THE REPRESENTATION THEOREM

R.S. BUCY

Dedicated to the memory of Professor Gottfried Köthe

1. INTRODUCTION

Representing the solutions of partial differential equations by integrals over function space has been done for various problems. In quantum mechanics, Feynman solved the Schroedinger equation in this way, see [5]. Mark Kac used integrals over Brownian paths to represent solutions of a generalized Fokker-Planck equation with particle birth and death, see [7]. Function space integrals with respect to Brownian paths have been considered by Wiener, [12] and Friedricks, [6]. Ito in [8], introduced path descriptions of Markov diffusion processes, stochastic differential equations, with these processes having probability distributions satisfying generalized Fokker-Planck equations. In the early 60's, Stratonovich derived a non-linear partial differential driven by stochastic term governing the evolution of the conditional density of the signal given the observations for the nonlinear filtering problem, see [10]. Controversy arose over the form of the stochastic driving term in this equation which hinged on the stochastic calculus used. In [1], Bucy proposed a solution to a version of the Stratonovich partial differential equation valid for the Ito calculus as a function space integral with respect to the signal process paths. This result was known as the representation theorem. Proofs were given for the representation theorem in [2], [3] and [9]. Duncan in his thesis, [4], resolved the statistical testing problem for processes using the representation theorem. This theorem was used to synthesize nonlinear filters with digital computers, see [2]. In this paper, we will derive the discrete time version of the representation theorem. It is interesting to do this, as the details are less technical than in the continuous time case. Further some interesting connections with Statistical Mechanics are apparent when this is done. Integrability conditions for systems of partial differential equations are used to characterize the solution of the nonlinear filtering problem. In the special case of the linear gaussian filtering problem, this characterization coincides with the Krein-Bellman equation.

2. THE FILTERING AND DECISION PROBLEM

Let v_n be an s-vector valued gaussian discrete time white noise process, that is $Ev_n = 0$ and $Ev_n v'_m = rI\delta_{n,m}$, the noise process. Let $x(n)$ be a discrete parameter d-vector valued stochastic process independent of v_n, the signal process. Let $h : R^n \to R^s$ be a measurable function satisfying $E||h(x(n))||^2 < \infty$ for all n, h is called the sensor. The observation process is $z(n) = h_n + v_n$, where $h_n = h(x(n))$. Let z^n be the vector with components z_j for $j < n$. There are two problems associated with this model;

1) The Filtering Problem, estimate $x(n)$ given $[z(i)]_{i=n-1,n-2,\dots,1}$ so as to minimize an appropriate loss function.

2) The Decision Problem, decide between the hypotheses $H_0 : z(n) = h_n + v_n$ and $H_1 : z(n) = v_n$, where we assume;

A_1: The densities under H_0 and H_1 are known twice continuously differentiable functions of z_i.

We denote by $\mathscr{L}$ the likelihood ratio,

$$(1) \qquad \mathscr{L}_1^n = \frac{p_{H_0}(z^n)}{p_{H_1}(z^n)}$$

where $z(k) = h_k + v_k$, and by

$$(2) \qquad \widehat{h}_{i|n} = Eh_i|z^n$$

the condition expectation of the sensor.

Lemma 1.

$$(3) \qquad \mathscr{L}_1^n = E^*(e^{\mathscr{H}_n})$$

where

$$\mathscr{H}_n = \beta \sum_{i=1}^{n} [(h_i, z(i)) - \frac{1}{2}(h_i, h_i)]$$

* denotes that z^n is fixed and the expectation is over $x(.)$ and $\beta = \dfrac{1}{r}$.

Proof. The relation between the joint and marginal densities yields,

$$(4) \qquad \mathscr{L}_1^n = \frac{p_{H_0}(z^n)}{p_{H_1}(z^n)} = \frac{1}{p_{H_1}(z^n)} \int p_{H_0}(z^n, x^n)\, dx^n$$

or

$$\frac{1}{p_{H_1}(z^n)} \int p(z^n|x^n) p(x^n) = \frac{1}{p_{H_1}(z^n)} \int p_{H_1}(z^n - h^n) p(x^n)\, dx^n$$

which in turn equals;

$$(5) \qquad \int exp(\mathscr{H}_n) p(x^n)\, dx^n$$

The above integral is finite as the integrated is bounded by an almost surely finite function. Let $\widehat{g}_{i|n}$ be the conditional expectation of $g(x_i)$ given $(z(i))_{i=1,\dots,n,}$ when it exists.

Lemma 2.

$$\widehat{h}_{i|n}\mathscr{L}_1^n = E^*(h_i e^{\mathscr{H}_n}) \tag{6}$$

Proof.

$$\widehat{h}_{i|n} = \int h_i p(x^n|z^n)\,dx^n = \int h_i \frac{p(x^n, z^n)}{p_{H_0}(z^n)}\,dx^n \tag{7}$$

$$= \int h_i \frac{p(z^n|x^n)p(x^n)}{p_{H_0}(z^n)}\,dx^n = \int h_i \frac{p_{H_1}(z^n - h^n)}{p_{H_0}(z^n)} p(x^n)\,dx^n \tag{8}$$

$$= \frac{1}{\mathscr{L}_1^n} \int h_i \frac{p_{H_1}(z^n - h^n)}{p_{H_1}(z^n)} p(x^n)\,dx^n = \frac{E^*(h_i e^{\mathscr{H}_n})}{E^*(e^{\mathscr{H}_n})}. \tag{9}$$

*Remark.*Equations (7) - (9) hold for arbitrary i and n. When $E(g(x_i)) < \infty$, it is easily seen from the above proof that $\widehat{g}_{i|n}\mathscr{L}_i^n = E^*(g(x_i)e^{\mathscr{H}_n})$.

We introduce the innovations process, $I_k = z(k) - \widehat{h}_{k|k-1}$ and assume that;

A_2 : The σ field induced by the innovations I^k is identical to that induced by z^k for any k.

Lemma 3.

$$\frac{\partial \mathscr{H}_n}{\partial I_k} = \beta\left(h_k + \sum_{i=k+1}^{n} \frac{\partial \widehat{h}'_{i|i-1}}{\partial I_k} h_i\right) \tag{10}$$

for $k < n + 1$.

Proof. This follows by direct computation.

Now using the Lemma, since differentiation can be interchanged with expectation for z^n fixed, in the expression $E^*(e^{\mathscr{H}_n})$ as $\frac{\partial \mathscr{H}_n}{\partial I_k} e^{\mathscr{H}_n}$ is integrable with respet to the measure on x^n, it follows that;

$$\frac{\partial \mathscr{L}_1^n}{\partial I_k} = \beta\left(\widehat{h}_{k|n} + \sum_{i=k+1}^{n} \frac{\partial \widehat{h}'_{i|i-1}}{\partial I_k} \widehat{h}_{i|n}\right) \mathscr{L}_1^n \tag{11}$$

Now (11) is a system of partial differential equations for $\log(\mathscr{L}_1^n)$, and assuming $\log(\mathscr{L}_1^n)$ is C^2 then the well known integrability conditions yields.

Theorem 1. *Suppose $\widehat{h}_{i|n}$ is a C^2 function of $(I_m)_{m=1,\ldots,n}$ then,*

$$\widehat{h}_{a|n;b} + \sum_{k=a+1}^{n} \widehat{h}'_{k|k-1;a}\widehat{h}_{k|n;b} + U^{max(a,b)}(k)\widehat{h}'_{k|k-1;b;a}\widehat{h}_{k|n} \tag{12}$$

$$= \widehat{h}_{a|n;a} + \sum_{k=b+1}^{n} \widehat{h}'_{k|k-1;b}\widehat{h}_{k|n;a} + U^{max(a,b)}(k)\widehat{h}'_{k|k-1;a;b}\widehat{h}_{k|n} \tag{13}$$

where $g_{l|n;k} = \dfrac{\partial g_{l|n}}{\partial I_k}$ *and* $U^a(k) = 1$ *for* $k > a$ *and 0 otherwise.*

Notice that $\widehat{h}_{i|n} i = 1, \ldots, n$ is the least amount of information necessary to solve the filtering problem. In the special case where h is a linear function of the state, x_n, and x_n and v_n are gaussian, it is easy to see that;

$$\widehat{h}_k = E(h_k|I_1, \ldots, I_n) = E(h_k|I_1, \ldots, I_{k-1}) + E(h_k|I_k, \ldots, I_n) \tag{14}$$

where $k < n$.

Or

$$\frac{\partial \widehat{h}_{l|l-1}}{\partial I_k} = \frac{\partial \widehat{h}_{l|n}}{\partial I_k}$$

for $k < l$. This leads to the Krein-Bellman equation;

$$\frac{\partial h_{l|n}}{\partial I_k} - \frac{\partial h_{k|n}}{\partial I_l} = \mu(l,k) \sum_{j=min(l,k)+1}^{max(l,k)} \frac{\partial h'_{j|n}}{\partial I_k}\frac{\partial h_{j|n}}{\partial I_l}$$

where $\mu(a,b) = 1$ if $b > a$ and -1 otherwise.

Example. In this linear gaussian case in fact $\widehat{h}_{a|n;b} = C_{a,b}$, constants, so that;

$$C_{a,b} - C_{b,a} = \mu(a,b) \sum_{k=min(a,b)+1}^{max(a,b)} C'_{k,a}C_{k,b}$$

In this case we can evaluate $\mathscr{L}_1^n$ explicitly. Denote by $s_i, \widehat{h}_{i|n}$ then

$$s^n = Eh^n z^{n'}(Ez^n z^{n'})^{-1} z^n = K(K + rI)^{-1} z^n = \beta(\beta I + K^{-1})^{-1} z^n$$

where $K = E(h^n h^{n'})$. Now it easy to evaluate $\mathscr{L}_1^n$ since;

$$\mathscr{L}_1^n = E^* e^{\mathscr{H}_n} = \frac{1}{(2\pi)^{\frac{n}{2}} det(K)} \int e^{\beta(z^n, h^n)} e^{\frac{-\|h^n\|^2_{(K^{-1}+\beta I)}}{2}} dh^n \tag{15}$$

$$= \frac{1}{det(I+\beta K)} e^{\frac{\beta^2 \|z^n\|^2_{(K^{-1}+\beta I)^{-1}}}{2}} = \frac{e^{\frac{\beta(z^n, z^n)}{2}}}{det(I+\beta K)} \tag{16}$$

$$= \frac{e^{\frac{\beta}{2}\sum_{i=1}^{n}(z_i, \hat{h}_{i|n})}}{det(I+\beta K)} \tag{17}$$

The log Likelihood ratio $log(\mathscr{L}_1^n)$ is the test statistic to decide the decision problem, see [4].

3. CONCLUSIONS

We derive a new characterization of the discrete time nonlinear filtering under smoothness assumptions on the densities. In the linear gaussian case these yield well known relations. It would be interesting to find other cases where $\mathscr{L}_1^n$, the analog of the partition function in statistical mechanics can be evaluated explicitly. In general, when the free energy, $F = r log(\mathscr{L}_1^n)$ is known, the entropy and the internal energy can be found by differentiation, see [11].

REFERENCES

[1] R.S. Bucy, *Nonlinear Filtering*, IEEE Trans. Auto. Control. 198, 1965.

[2] R.S. Bucy, *Linear and Nonlinear Filtering*, Invited Tutorial Review Paper, Proc. of IEEE, 58, 6, 854-964, 1970.

[3] R.S. Bucy, P.D. Joseph, *Filtering for Stochastic Processes with applications to Guidance*, 2nd Edition, Chelsea, New York, 1987.

[4] T.E. Duncan, *Probability densities for Diffusion Processes*, AFOSR Tech. Report 7001-4, Stanford University, 1967.

[5] R.P. Feynman, A.R. Hibbs, *Quantum Mechanics and Path Integrals*, Mc Graw Hill, New York, 1965.

[6] K.O. Friedricks, *Function Space Integrals*, Lecture Notes, NYU, 1950.

[7] K. Ito, H.P. Mc Kean, *Diffusion Processes and their Sample Paths*, Springer Verlag, Berlin, 1965.

[8] K. Ito, *Lectures on Stochastic Processes*, Tata Institute of Fundamental Research, Bombay, 1961.

[9] R.E. Mortensen, *Optimal Control of Continuous Time Stochastic Systems*, Thesis, E.E. Dept. U.C., Berkeley, 1966.

[10] R.L. Stratonovich, *Conditional Markov Processes*, Theory Of Probability and Applications, 5, 2, 156-178, 1960.

[11] R.L. Stratonovich, *Theory of Information*, (in Russian), Soviet Radio, 1975.

[12] N. Wiener, *Nonlinear Problems in Random Theory*, MIT Press, Cambridge, 1958.

Received December 22, 1990
R.S. Bucy
Aerospace Engineering Department
University of Southern California
Los Angeles, CA 90089 - 1191
USA