Astérisque

R. W. Brockett

Remarks on finite dimensional nonlinear estimation

Astérisque, tome 75-76 (1980), p. 47-55
<http://www.numdam.org/item?id=AST_1980__75-76__47_0>

REMARKS ON FINITE DIMENSIONAL NONLINEAR ESTIMATION

by

R.W. BROCKETT

-:-:-:-

I.- INTRODUCTION.

In this paper we consider the problem of estimating the values taken on by a random process $y(t)$ of the form

$$dx = f(x)\,dt + g(x)\,dw \;;\quad dy(t) = h(x)\,dt + dv \tag{1}$$

where w and v are independent Wiener processes and the differential equation is to be interpreted as an Ito equation in $\mathbb{R}^n$. In particular, we investigate the existence of "recursive estimators", i.e. differential equations of the form

$$dz = a(z)\,dt + b(z)\,dy \tag{2}$$

$$\hat{y} = c(z) \tag{3}$$

where $\hat{y}$ may be for example, the conditional mean of y and z is finite dimensional. The main ideas involve the conditional density equation which, in unnormalized, Ito form is

$$d\hat{\rho}(t,x) = L_o\,\hat{\rho}(t,x)\,dt + L_1\,\hat{\rho}(t,x)\,dy \tag{4}$$

where $\hat{\rho}$ is, apart from a scale factor, the conditional density for x, given $y(s)$ for $0 \leqslant s \leqslant t$. Of course conditional expectations, etc. can be expressed in terms of $\hat{\rho}$ as

$$\hat{y}(t) = \int \psi(x)\,\hat{\rho}(t,x)\,dx\left[\int \hat{\rho}\,dx\right]^{-1} . \tag{5}$$

* Partially supported by the Army Research Office under Grant DAAG 29-76-C-0139, the National Aeronautics and Space Administration under Grant NSG-2265, and the U.S. Office of Naval Research under the Joint Services Electronics Program Contract N00014-75-C-0648.

Earlier work on the relationship between Lie-theoretic ideas and estimation theory appears in [1]. The main points of this paper involve the exploration of the following two ideas, both of which seem to be novel.

(a) If $\mathfrak{L}$ is the Lie algebra generated by the operators $L_o - \frac{1}{2} L_1^2$ and L_1 and $\mathfrak{F}$ is the Lie algebra of vector fields generated by $a - \frac{1}{2} b'b$, and b then under appropriate hypothesis $\mathfrak{F}$ will be a homomorphic image of $\mathfrak{L}$.

Conversely any homomorphism of $\mathfrak{L}$ onto a Lie algebra of complete vector fields on a finite dimensional manifold permits one to obtain some information about the conditional density by propagating the solution of a finite dimensional set of equations.

(b) Under appropriate hypothesis the input-output map defined by (4)-(5) is characterized by a Volterra series. This Volterra series may or may not have kernels which are separable in the sense of [2]. A necessary condition for the existance of a finite dimensional nonlinear estimator is that the kernels be separable.

Because of space limitations we can only sketch the basic ideas in this paper.

II. - GENERALITIES.

Consider the following notation. If $f : \mathbb{R}^n \to \mathbb{R}^n$ then we associate with f a vector field on $\mathbb{R}^n$ according to

$$L_f = \sum_{i=1}^{n} f_i \frac{\partial}{\partial x_i} .$$

Write for the adjoint

$$L_f^* = \sum_{i=1}^{n} \frac{\partial}{\partial x_i} f_i .$$

We can express equation (4) in this notation as

$$d\rho(t,x) = L_{\hat{f}}^* \, dt + \frac{1}{2} (L_g^*)^2 \hat{\rho} \, dy \tag{4}$$

where

$$\hat{f} = f - \frac{1}{2} \left(\frac{\partial g}{\partial x}\right) g \quad .$$

Given an initial value for $\hat{\rho}$ the pair of equations (4)-(5) then define what is called in control theory, an input-output map. If for a given initial value of z the pair (2)-(3) generate the same input-output map then we can deduce certain relationships concerning a, b, L_o and L_1. In order to facilitate these comparisons, which, for the most part involve Lie algebraic constructions, it is more natural to convert the two differential equations (2) and (4) to Fisk-Stratonovich form. In this way we can avoid the use of the somewhat unintuitive Ito calculus. [See, e.g.[3]). As is well known, in Fisk-Stratonovich form equations (2) and (4) are ($\dagger d$ distinguishes the Fisk-Stratonovich differentials from Ito differentials)

$$\text{(2')} \qquad \dagger dx = (a - \frac{1}{2}\frac{\partial b}{\partial x} b)dt + b\,\dagger dy$$

$$\text{(4')} \qquad \dagger d\rho = (L^{*}_{\tilde{f}} - \frac{1}{2}(L^{*}_{g})^{2} - \frac{1}{2}(h(x))^{2}\rho\, dt + h(x)\,\rho\,\dagger dy \quad .$$

Appending (3) to (2') and (5) to (4') we obtain, for each assignment of initial data, input-output systems.

Let us suppose now that the vectors a and b entering the differential equation (2') are real analytic maps of $\mathbb{R}^n$ into $\mathbb{R}^n$. Also, suppose that the pair (2')-(3) defines a minimal system in the sense that the Lie algebra of vector fields generated by $a - \frac{1}{2} b'b$, and b acts transitively at each point in $\mathbb{R}^n$ and that no two distinct initial states for x give rise to the same response for all smooth inputs. These assumptions have the effect of insuring that there is no redundancy in the pair (2')-(3).

The following observation is now appropriate. (cf.[4]-[5]).

Remark : Suppose that there exists for $L_o - \frac{1}{2} L_1^2$ and L_1 a common set of analytic vectors $\mathcal{D}$. If for some choice of initial condition z_o and $\rho(0) \in \mathcal{D}$ the pair (2)-(3) and the pair (4')-(5) generate the same input output map for all smooth inputs with (2)-(3) analytic and minimal then the map

$$\Phi : L_o - \frac{1}{2} L_1^2 \to a - \frac{1}{2} b'b$$

$$\Phi : L_1 \rightarrow b$$

extends to a Lie algebra homomorphism from the Lie algebra of operators on $\mathcal{B}$ generated by $L_o - \frac{1}{2} L_1^2$ and L_1 to the Lie algebra of vector fields generated by $a - \frac{1}{2} b'b$ and b.

Now of course relevance of this remark stems from the fact that under suitable assumptions it has been shown (see, e.g. [6], [7], [8]) the behaviour of a stochastic differential equation with white noise inputs can be deduced from its behaviour on smooth functions.

Reference [1] discusses a context in which this program is particularly easy to carry out in detail. It is the case where the stochastic process x takes on only a finite number of values and the operators L_o and L_1 act on finite dimensional spaces. In this case results available on the representation of Lie groups together with the above remark provide considerable guidance about the design of finite dimensional nonlinear filters.

III.- THE LINEAR PROBLEM.

We now illustrate these ideas in one nontrivial case. Even though the example is a linear problem our techniques give important new information about the role of "all pass" factors in simplifying the Kalman-Bucy filter.

The unnormalized conditional density equation associated with the estimation problem described by n simultaneous Ito equations

$$dx_i = \Sigma a_{ij} x_j \, dt + b_i \, dw \quad ; \quad dy = c_i x_i \, dt + dv$$

is given in Ito form by

$$(*) \qquad d\rho = \tilde{L}_o \rho \, dt + L_1 \rho \, dy$$

where

$$\tilde{L}_o = \frac{\partial}{\partial x_i} a_{ij} x_j + \frac{1}{2} \Sigma \frac{\partial^2}{\partial x_i \partial x_j} b_i b_j$$

and

$$L_1 = \Sigma c_i x_i \ .$$

Converting (*) into Fisk-Stratonovich form we get

(**) $$d\rho = L_o\rho\, dt + L_1\rho\, dy$$

where $L_o = \text{is } \tilde{L} - \frac{1}{2} L_1^2$.

The main point of [1] is the observation that the structure of the Lie algebra generated by L_o and L_1 is of overriding importance in understanding the estimation problem. (See also the work of Mitter [9]).

Lemma : The Lie algebra generated by L_o and L_1 has a basis $L_o, L_1, \ldots, L_{2n}, L_{2n+1}$ and commutation relations

$$\begin{aligned} L_i &= \mathrm{ad}_{L_o}^{i-1}(L_1) && i = 1, 2, \ldots, 2n \\ [L_i, L_j] &= \beta_{ij}\, L_{2n+1} && i, j = 1, 2, \ldots, 2n \\ [L_{2n+1}, L_i] &= 0 && i = 0, 1, \ldots, 2n+1 \end{aligned}$$

provided that $g(s) = c(Is-A)^{-1}b$ is of Mc Millan degree n and $g(s)\,g(-s)$ is of Mc Millan degree $2n$.

Remark : One easily verifies that if one regards the Lie algebra as a complex Lie algebra then the derived algebra in this case is isomorphic to the Heisenberg algebra $\mathbb{H}^n$.

Remark : The conditions on $g(s)$ corresponds to the requirement that the filter be minimal and that $g(s)$ have no factors of the form $(s-a)/(s+a)$. The contribution of such "all pass" factors to degeneracy in filtering problems is considerably clarified by this lemma.

Proof : From the form of L_o and the form of L_1 we see that $[L_o, L_1]$ is $\langle c, Ax\rangle + \Sigma (b_i \frac{\partial}{\partial x_i})(\Sigma b_i c_i)$, i.e. the sum of a linear function and a linear constant coefficient differential operator. Moreover if L is any such sum then one verifies easily that $[L_o, L]$ is as well. Moreover, $[L, L']$ is a constant if L and L' are each the sum of a linear function and a first order differential operator. Putting these remarks together we see that the Lie algebra $\{L_o, L_1\}_{LA}$ is at most a $2n+2$ dimensional algebra and that $L_o, L_1, \ldots, L_{2n+1}$

as defined above certainly contains $\{L_o, L_1\}_{LA}$ in their linear span. The only remaining question is that of the independence of the L_i. The key calculation is this. Write

$$L = \sum_{i=1}^{n} (\alpha_i \frac{\partial}{\partial x_i} + \beta_i x_i) \quad .$$

Then a calculation verifies that

$$[L_o, L] = \sum_{i=1}^{n} (\tilde{\alpha}_i \frac{\partial}{\partial x_i} + \tilde{\beta}_i x_i)$$

where

$$\begin{bmatrix} A & cc' \\ -bb' & -A' \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \tilde{\alpha} \\ \tilde{\beta} \end{bmatrix}$$

Thus the independence of $L_1, L_2, \ldots, L_{2n}$ is determined by the independence of the vectors

$$\begin{bmatrix} A & cc' \\ -bb' & -A' \end{bmatrix}^k \begin{bmatrix} c \\ 0 \end{bmatrix} \quad ; \quad k = 0, 1, \ldots, 2n-1 \; .$$

Now with a little work one can verify that the following generating function identity

$$\sum_{i=o}^{\infty} [0, c] \begin{bmatrix} A & cc' \\ -bb' & -A' \end{bmatrix}^i \begin{bmatrix} c \\ 0 \end{bmatrix} s^{-i-1} = \frac{g(s)\, g(-s)}{1+g(s)\, g(-s)} \quad .$$

We see from standard results in realization theory [10] that the Lemma holds.

Incidentally, this Lie algebra is isomorphic to the Lie algebra which appears in the recent Cauchy-Riemann theory and it has the Heisenberg algebra as its derived algebra (see [11]).

The equations of motion for the conditional density in this case can be written in vector/matrix notation as

$$dz = (A - P(\tau)\, cc')\, zdt + P(\tau)\, dy \, ; \; z(t) \in \mathbb{R}^n$$

$$\frac{d\tau}{dt} = 1 \quad .$$

Here $P(\tau)$ is an n by n matrix which satisfies the following differential equation

$$\frac{d}{d\tau} P(\tau) = AP(\tau)+P(\tau)A'+bb'-P(\tau)cc'P(\tau) .$$

We see that

$$\tilde{f} = \begin{bmatrix} (A-P(\tau)cc')x \\ 1 \end{bmatrix} \quad \text{and} \quad \tilde{g} = \begin{bmatrix} Pc \\ 0 \end{bmatrix}$$

are the local-coordinate descriptions of vector fields on $\mathbb{R}^{n+1}$. The vector fields in the Lie algebra which they generate are, with the exception of $\tilde{f}$ itself, of the form

$$h = \begin{bmatrix} Pn_1 + n_2 \\ 0 \end{bmatrix}$$

with n_1 and n_2 in $\mathbb{R}^n$. Clearly $\{\tilde{f},\tilde{g}\}_{LA}$ is then a Lie algebra of dimension $2n+1$ or less. Additional calculations, which we omit here, show that it is of dimension $2n+1$ under the hypothesis of the lemma.

To illustrate these ideas in the simple case we consider the basic example

$$dx = dw \quad ; \quad dy = xdt + dv .$$

The filtering equations for the conditional mean z, given a gaussian initial distribution for x with variance k_o are

$$\begin{aligned} dz &= -kzdt + kdy \\ \dot{k} &= -k^2+1 . \end{aligned}$$

Introduce the vector fields

$$\begin{aligned} F &= -kz \frac{\partial}{\partial z}+(1-k^2)\frac{\partial}{\partial k} \\ G &= k\frac{\partial}{\partial z} . \end{aligned}$$

A calculation shows that

$$[F,G] = \frac{\partial}{\partial z} \overset{\text{def}}{=} H$$

and that

$$[F,H] = k\frac{\partial}{\partial z} = G .$$

On the other hand, for the conditional density equation we have in this case

$$d\hat{\rho} = L_o\hat{\rho}\,dt + L_1\hat{\rho}\,dy$$

where

$$L_o = \frac{1}{2}\left(\frac{\partial^2}{\partial z^2} - z^2\right), \quad L_1 = z \quad .$$

Now the Lie algebra of operators $\{L_o, L_1\}_{LA}$ satisfies

$$[L_o, L_1] = \frac{\partial}{\partial z} \stackrel{\text{def}}{=} L_2$$

$$[L_o, L_2] = z = L_1$$

$$[L_1, L_2] = 1 \stackrel{\text{def}}{=} L_4 \quad .$$

There is a Lie algebra homomorphism from $\mathcal{L}$ onto $\mathcal{G}$ which sends L_o into F, L_1 into G and has the operator L_4 in its kernel.

The Volterra series referred to in (b) of the introduction may, in this case, be computed to be

$$\hat{y}(t) = e^{\alpha(t)}\hat{y}(0) + \int_o^t e^{\int_o^t \beta(\rho)\,d\rho}\, dy(\rho)$$

which is separable and has no nonlinear terms. The expression for α and β involves solving for k above.

-:-:-:-

REFERENCES

[1] R.W. BROCKETT and J.M.C. CLARK.- On the Geometry of the Conditional Density Equation. Proc. of the Oxford Conference on Stochastic Control, Oxford, 1978, (to appear).

[2] R.W. BROCKETT.- Volterra Series and Geometric Control Theory. Automatica, Vol.12, n°2, March 1976, p.167-176.

[3] J.M.C. CLARK.- An Introduction to Stochastic Differential Equations on Manifolds. In Geometric Methods in System Theory, D.Q. Mayne and R.W. Brockett, es., Reidel, Dordrecht, 1973.

[4] H.J. SUSSMANN.- Existence and Uniqueness of Minimal Realizations of Nonlinear Systems. Math. System Theory, Vol.10, n°3, 1976/1977.

[5] A.J. KRENER and R. HERMANN.- Nonlinear Controllability and Observability. IEEE Trans. on Automatic Control, Vol.22, n°5, 1977, p.728-740.

[6] H.J. SUSSMANN.- On the Gap Between Deterministic and Stochastic Ordinary Differential Equations. Ann. of Prob., Vol.6, 1978, p.19-41.

[7] M.I. FREEDMAN and J.C. WILLEMS.- Smooth Representation of Systems with Differential Inputs. IEEE Trans. on Automatic Control, Vol.23, n°1, 1978, p.16-21.

[8] J.M.C. CLARK.- The Design of Robust Approximation to the Stochastic Differential Equations of Nonlinear Filtering in Communication Systems and Random Process Theory. J.K. Skwitzynski (ed) Sijthoff and Nordhoff, Alphen aan den Ryn, 1978.

[9] S.K. MITTER.- Modeling for Stochastic Systems and Quantum Fields. 1978 IEEE Conference on Decision and Control, IEEE New-York.

[10] R.W. BROCKETT.- Finite Dimensional Linear Systems. J. Wiley, New-York, 1970.

[11] A. KORANY and S. VAGI.- Ann. Scuola Norm. Pisa 25, (1971), p.575-648.

-:-:-:-

R.W. BROCKETT
Division of Engineering and
Applied Physics
Harvard University, Pierce Hall
CAMBRIDGE, MASSACHUSSETS 02138