SYSTEM IDENTIFICATION AND
NONLINEAR FILTERING: LIE ALGEBRAS
AND
A SUMMARY OF APPROXIMATION METHODS FOR
NONLINEAR FILTERING PROBLEMS ARISING IN
SYSTEM IDENTIFICATION

P.S. KRISHNAPRASAD, S.I. MARCUS AND M. HAZEWINKEL
AND
P.S. KRISHNAPRASAD, M. HAZEWINKEL AND B. HANZON

REPRINT SERIES no. 299

This article appeared in "IEEE Conference on Decision &

ERASMUS UNIVERSITY ROTTERDAM - P.O. BOX 1738 - 3000 DR ROTTERDAM - THE NETHERLANDS
SYSTEM IDENTIFICATION AND NONLINEAR FILTERING: LIE ALGEBRAS

P.S. Krishnaprasad* Steven I. Marcus** Michiel Hazewinkel***

*Univ. of Md. College Park, Md. 20742  **Univ. of Texas Austin, Texas 78712  ***Erasmus Univ. Rotterdam, The Netherlands

Abstract

This paper is a continuation of our previous work ([1], [2], [3]) to understand the identification problem of linear system theory from the viewpoint of nonlinear filtering. The estimation algebra of the identification problem is a subalgebra of a current algebra. It therefore follows that the estimation algebra is embeddable as a Lie algebra of vector fields on a finite dimensional manifold. These features permit us to develop a Weierstrass-Weierstrass type procedure for the associated Cauchy problem and reveal a set of functionals of the observations that play the role of joint sufficient statistics for the identification problem.

1. Introduction

Consider the stochastic differential system:

\[ d\theta_t = 0 \]
\[ dx_t = A(\theta)x_t dt + b(\theta)dw_t \]
\[ dy_t = c(\theta,x_t)dt + dv_t \]

(1)

Here \( \{w_t\} \) and \( \{v_t\} \) are independent, scalar, standard, Wiener processes, and \( \{x_t\} \) is an \( \mathbb{R}^n \)-valued process. Assume that \( \theta \) takes values in a smooth manifold \( \Theta \subset \mathbb{R}^N \), and the map \( \theta : \Theta \rightarrow (A(\theta), b(\theta), c(\theta)) \) in a smooth map taking values in minimal triples. By the identification problem we shall mean the nonlinear filtering problem associated with eqn. (1); i.e. the problem of recursively computing conditional expectations of the form

\[ \tau_c(\theta) \overset{\text{def}}{=} \mathbb{E}[H(x_t,\theta) | Y_t] \]

(2)

where \( \tau_c(\theta) \) and \( Y_t \) are fixed volume elements on \( \mathbb{R}^n \) and \( \Theta \) respectively. Further if \( Q(t,\theta) \) denotes the unnormalized posterior density of \( \theta \) given \( t \), then it satisfies the Ito equation:

\[ \frac{dQ}{dt} = E[H(x_t,\theta) | Y_t] \frac{dQ}{dt} \]

(3)

Recent work in nonlinear filtering theory (see the proceedings [6]) shows that it is natural to look at eqn. (2) formally as a deterministic partial differential equation, and the Lie algebra of the identification problem, we shall mean the operator Lie algebra \( G \) generated by \( A_o \) and \( B_o \). For more general nonlinear filtering problems, estimation algebras analogous to \( G \) have been emphasized by Brockett and Clark [7], Brockett ([8] - [11]), Hitter ([12], [13]), Hazewinkel and Marcus [14] and others (see [6]) as being objects of central interest. In the papers ([1], [2]) the Lie algebra \( G \) is used to classify identification problems and to understand the role of certain sufficient statistics.

2. The Structure of the Estimation Algebra \( G \)

To understand the structure of the estimation algebra \( G \) it is well worth considering an example.

Example 1:

Let \( dx_t = 0, dw_t; \quad d\theta = 0 \)
\[ dy_t = x_t dt + dv_t \]

Then \( A_o = \frac{\partial^2}{\partial x^2}-\frac{\partial}{\partial x}-\frac{\partial}{\partial x}A(\theta)x^2/2 \)
\[ B_o = c(\theta, x) \]

(4)

\[ G = \{A_o, B_o\} \]

is spanned by the set of operators

\[ \frac{x^2}{2}, \quad \frac{\partial^2}{\partial x^2}, \quad \{x^{2n}\}_{n=1}^{\infty} \]

We then notice that, 

$$ G \subset \mathbb{R}[\mathbb{A}^2] \{ \frac{2}{3x}, \frac{2}{3x^2}, \frac{2}{3x} \} \cdot \mathbb{A} \cdot \mathbb{A} \cdot \mathbb{A} \cdot \mathbb{A} $$

is a subalgebra of the Lie algebra obtained by tensoring the polynomial ring $$ \mathbb{R}[\mathbb{A}^2] $$ with a 6 dimensional Lie algebra. 

The general situation is very much as in this example. Consider the vector space (over the reals) of operators spanned by the set,

$$ S_i := \{ \frac{2}{3x}, \frac{2}{3x^2}, \frac{2}{3x} \} \cdot \mathbb{A} \cdot \mathbb{A} \cdot \mathbb{A} \cdot \mathbb{A} \cdot \mathbb{A} $$

This space of operators has the structure of a Lie algebra henceforth denoted as $$ G_0 $$ (of dimension 3n^2 + 2n + 1) under operator commutation.

For each choice of $$ \theta $$, $$ a_0 $$ and $$ b_0 $$ take values in $$ G_0 $$, it follows that in general $$ A $$ and $$ B $$ are smooth maps from $$ \Theta $$ into $$ G_0 $$.

Theorem 1: The map $$ \phi : G \rightarrow \text{Lie algebra of vertical vector fields on a finite dimensional manifold fibered over } \Theta $$

defined by

$$ \phi(A) = a_0 $$, $$ \phi(B) = b_0 $$

is a faithful representation of the Lie algebra of the identification problem as a Lie algebra of (vertical) vector fields on a finite dimensional manifold fibered over $$ \Theta $$.

Example 2: To illustrate Theorem 1, consider the Lie algebra of example 1. The embedding equations (11) take the form

$$ d\theta = 0 $$

$$ dp = (\theta^2 - p^2)dt $$

$$ dz = -pdt + pdy $$

Then
The induced maps on Lie brackets are given by

\[ \phi_k(e^{2k\alpha}e^{\beta}) = e^{2k\alpha}e^{\beta} \]

where \( \phi_k \) is the map defined by \( \phi_k(x) = e^{2k\alpha}e^{\beta} \).

The induced maps on Lie brackets are given by

\[ \phi_k(e^{2k\alpha}e^{\beta}) = e^{2k\alpha}e^{\beta} \]

where \( \phi_k \) is the map defined by \( \phi_k(x) = e^{2k\alpha}e^{\beta} \).

The embedding equations have the following statistical interpretation. Assume that the initial condition for (12) is of the form,

\[ \rho_0(x,0) = 2(2\pi\text{det}E(\theta))^{-n/2} e^{-\alpha\cdot \mu(\theta)/(\ell - 1)} \]

where \( e^{\cdot \mu(\theta)} \) is a smooth map, \( \mu(0) > 0 \) and \( \mu(0) > 0 \) for \( 0 \leq \theta < 1 \).

Append to the system (11) an output equation,

\[ \mathcal{L} = \mathcal{L}(H;L) \]

Define,

\[ \mathcal{R} = C^\infty(M;\mathfrak{g}(m,R)) \]

\[ \mathcal{L} = C^\infty(M;L) \]

\[ \mathcal{D} = C^\infty(M;G) \]

Clearly \( \mathcal{R} \) is an algebra under pointwise multiplication and

\[ \mu(0) = \alpha \subset \mathcal{R} \]

Let \( \mathcal{U}_{\alpha}R \) be a \( C^\infty \) atlas for \( M \). Then for \( f_1, f_2 \in \mathcal{E} \), define

\[ \|f_1 - f_2\|_k = \left( \int d \text{vol}_{\alpha} \left| D^k(f_1 - f_2) \right|^2 \right)^{1/2} \]

where

\[ |f|^2 = \text{tr} \( f \cdot f \) \]

(Here \( k = d/2 + s \), \( s > 0 \)). Let \( \mathcal{R}_k \) be the completion of \( \mathcal{R} \) and \( \mathcal{D} \) the completion of \( \mathcal{D} \) in the norm \( \| \cdot \|_k \). By the Sobolev theorem, \( \mathcal{D} \) is a Banach algebra and the group operation

\[ \mathcal{D} \times \mathcal{D} \to \mathcal{D} \]

\[ (f_1, f_2) \mapsto f_1 f_2 \]

when \( (f_1 f_2)(m) = f_1(m) f_2(m) \) is continuous. Thus \( \mathcal{D} \) is a topological group.

Proceeding as before, one can give a Sobolev completion of \( \mathcal{E} \) to obtain \( \mathcal{D} \) an infinite dimensional Lie algebra where once again by the Sobolev theorem the bracket operation

\[ \left\{ \mathcal{E} \right\} \to \mathcal{D} \]

\[ (f_1, f_2) \mapsto f_1 f_2 \]

with \( [f_1, f_2](m) = f_1(m) f_2(m) \) is continuous. Now for a small enough neighborhood \( \mathcal{U}(0) \) of \( 0 \mathcal{D} \), one can define

\[ \exp: \mathcal{D}(0) \to \mathcal{D} \]

\[ \mathcal{D} \mapsto \exp(\mathcal{D}) \]

by pointwise exponentiation. This permits us to provide a Lie group structure on \( \mathcal{D} \) with \( \mathcal{D} \) canonically identified as the Lie algebra of \( \mathcal{D} \).

The procedure outlined above appears to play a significant role in several contexts (the index theorem, Yang-Mills fields [24] [25] [26] [27].

For our purposes \( L \) will be identified with a faithful matrix representation of \( G_0 \). Thus we associate with the identification problem a Sobolev Lie group which is a subgroup of \( \mathcal{D} \) corresponding to \( G_0 \).
Remark:

One of the important differences between the problem of filtering and the problems of Yang-Mills theories is that in the latter case there are natural norms for Sobolev completion. This follows from the fact that in Yang-Mills theories, the algebra $L$ is compact (semi-simple) and one has the Killing form to work with. In filtering problems $L$ is never compact.

Remark:

We would like to acknowledge here that Prof. Sanjoy Mitter was kind enough to acquaint one of us (P.S.K) with the work of P.K. Mitter.

5. The Integration Problem & Sufficient Statistics

In (3) we look for a representation of the form,

$$p(t,x,y) = \exp(g_1(t,y)A_1)...\exp(g_n(t,y)A_n)p_0 \tag{18}$$

for the solution to the equation (8). In the case of example (1) this takes the form

$$p(t,x,y) = \exp(g_1(t,y)(\frac{2}{x^2} - \frac{y^2}{2}))(A_{2k})\exp(g_2(t,y)\frac{2\partial}{\partial x})\exp(g_3(t,y)x) \cdot \exp(g_4(t,y)\frac{x}{y})p_0 \tag{19}$$

Differentiating and substituting in (8) we get,

$$\frac{3g_1}{\partial t}(t,y) = 1$$

$$\frac{3g_2}{\partial t}(t,y) = \cosh(g_1(t,y))y$$

$$\frac{3g_3}{\partial t}(t,y) = -\frac{1}{\theta} \sinh(g_1(t,y))y$$

$$\frac{3g_4}{\partial t}(t,y) = \frac{3g_2}{\partial t}(t,y)g_2(t,y)$$

and $g_i(0,0) = 0$ for $i = 1,2,3,4,\theta \omega$. The above first-order partial differential equations may be easily solved by quadrature and one has the representation,

$$p(t,x,y) = \int \frac{1}{2\sinh(|y|t)} \exp\left(-\frac{1}{2\sinh^2(|y|t) + z}\right)$$

$$\cdot \exp\left(\frac{xz}{\sinh(|y|t)}\right) \cdot \exp(\frac{g_2(t,y)}{\theta \omega}) \cdot \exp(\frac{g_3(t,y)}{\theta \omega})p_0(\frac{g_3(t,y)^2}{\theta \omega} + |y|t, \theta \omega) \cdot \exp(\frac{g_3(t,y)}{\theta \omega}) \cdot \exp(\frac{g_3(t,y)^2}{\theta \omega} + |y|t, \theta \omega)$$

$$= \int \frac{1}{2\sinh(|y|t)} \exp\left(-\frac{1}{2\sinh^2(|y|t) + z}\right)$$

$$\cdot \exp\left(\frac{xz}{\sinh(|y|t)}\right) \cdot \exp(\frac{g_2(t,y)}{\theta \omega}) \cdot \exp(\frac{g_3(t,y)}{\theta \omega})p_0(\frac{g_3(t,y)^2}{\theta \omega} + |y|t, \theta \omega) \cdot \exp(\frac{g_3(t,y)}{\theta \omega}) \cdot \exp(\frac{g_3(t,y)^2}{\theta \omega} + |y|t, \theta \omega)$$

where $p_0(\theta, \omega)$ is a bounded set and of closure $\Omega$.

In equation (21) the $g_i$'s should be viewed as canonical coordinates of the second kind on the corresponding Sobolev Lie group. Now expand $g_2$ and $g_3$ to obtain

$$g_2(t,y) = \Gamma(0)k^2 \int_0^t \frac{2k}{(2k+1)} d\omega k = 1,2,\ldots$$

$$g_3(t,y) = \Gamma(0)k^2 \int_0^t \frac{2k}{(2k+1)} d\omega k = 1,2,\ldots$$

It follows that all the "information" contained by the observations $y_0, y_0, \ldots$ about the joint unnormalized conditional density is contained in the sequence

$$T^t \left( \frac{\partial}{\partial t} y_0, d\omega; k = 0,1,2,\ldots \right)$$

Thus $T$ is nothing but a joint sufficient statistic for the identification problem.

Acknowledgements

Partial support for this work was provided by the Air Force of Scientific Research under grant AFOSR79-0025 and by the Department of Energy under Contract DEAC01-80-RAS0420-0001. One of us (P.S.K) would also like to acknowledge the hospitality of Erasmus University during a recent visit.

References


Approximation Methods for Nonlinear Filtering Problems Arising in System Identification

Summary

P.S. Krishnaprasad  Michiel Hazewinkel*  Ben Hanzon*

University of Maryland  *Erasmus University  Rotterdam, The Netherlands

College Park, MD 20742

Abstract: In this paper we investigate various approximate methods for computing the conditional density of a parameter. These techniques are related to the structure of certain Lie algebras of operators with the identification problem.

Consider the stochastic differential system:

\[ \begin{align*}
    d\theta &= 0 \\
    dx_t &= A(\theta)x_t dt + b(\theta) dw_t \\
    dy_t &= \langle c(\theta), x_t \rangle dt + dv_t.
\end{align*} \tag{1} \]

Here \( (w_t) \) and \( (v_t) \) are independent, scalar, standard Wiener processes and \( (x_t) \) is an \( \mathbb{R}^n \)-valued process. We let \( \theta \) take values in a smooth manifold \( \mathbb{M} \). Assume that the map \( \psi : (x, y, \theta) = (A(\theta), b(\theta), c(\theta)) \) is sufficiently smooth and takes values in the space of minimal triples.

Define two differential operators,

\[ \begin{align*}
    A_0 &= \frac{\partial}{\partial \theta} b(\theta) - \langle \partial / \partial \theta, A(\theta) x \rangle - \langle c(\theta), x \rangle^2 / 2 \\
    B_0 &= -\text{tr}(A(\theta)). \tag{3}
\end{align*} \]

The problem is to devise approximate finite dimensional, recursive techniques for calculating the conditional density of the parameter \( \theta \) given \( Y_t = \psi(x_t, y_t, \theta) \).

The general formulas are known:

\[ \begin{align*}
    P(t, x, \theta) &= e^{Q(t, \theta)} P(t, x, \theta) e^{-Q(t, \theta)} \\
    Q(t, \theta) &= \int_{[x, \theta]} \rho(t, x, \theta) dx d\theta \\
    \rho(t, x, \theta) &= e^{-Q(t, \theta)} P(t, x, \theta).
\end{align*} \tag{5} \]

where

\[ \begin{align*}
    \kappa_0 &= A_0 \\
    \kappa_1 &= \langle c(\theta), b(\theta) \rangle - \langle b(\theta), \partial / \partial \theta \rangle - \langle c(\theta), A(\theta) x \rangle \\
    \kappa_2 &= -\text{tr}(A(\theta)) \tag{8}
\end{align*} \]

Let \( Q(t, \theta) = e^{Q(t, \theta)} \). In this paper we consider approximations related to

(a) local series approximations

\[ S(t, \theta) = \sum_{k=0}^{l} \kappa_k(t) \theta^k \tag{1} \]

(b) Gaussian initial conditions:

\[ \rho(0, \theta) \text{ Gaussian for } \theta \]

Both these approximations are connected to the following algebraic objects:

(a) A sequence of Lie algebras \( G^{(k)} \)

where

\[ G^{(0)} = \{ A_0, B_0 \} \text{ L.A.} \]

\[ G^{(1)} = \left[ \begin{array}{ccc}
    A_0 & 0 & 0 \\
    0 & 3A_0 & 0 \\
    0 & 0 & 0
\end{array} \right] \text{ L.A.} \]

(b) Finite dimensional quotients of \( G^{(0)} \) in one-to-one correspondence with rings that are quotients of \( \mathbb{R}[\theta] \).

Our results use the fact that \( G^{(0)} \) is a subalgebra of a current algebra \( \{1, \theta\} \).

References
