On deformations, approximations and nonlinear filtering

Michiel HAZEWINKEL
Department of Mathematics, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

Received 3 March 1981

Given a (nonlinear) filtering problem there is associated to it a Lie algebra $L(\Sigma)$ of differential operators which is the Lie algebra generated by the two operators occurring in the Zakai equation for the unnormalized conditional density of the problem. The representability of $L(\Sigma)$ or quotients of $L(\Sigma)$ by means of vectorfields on a finite dimensional manifold is strongly related to the existence of exact finite dimensional recursive filters. However, in many cases, including the cubic sensor problem and the problem $dx = dw, dz = (x + ex^3) dt + dv, e \neq 0$, the algebra $L(\Sigma)$ is isomorphic to the Weyl algebra $W_1 = k(x, d/dx)$ which admits no nonzero homomorphisms into any Lie algebra of vectorfields on a finite dimensional manifold.

On the other hand the Lie algebra $L(\Sigma)$ mod $\epsilon^n$ is finite dimensional for all $n$ which opens up the possibility of the existence of a sequence of converging approximate filters.

Keywords: Nonlinear filtering, Lie algebra methods, Finite dimensional filters.

1. Introduction

Consider a stochastic system (in Ito form)
\[ \begin{align*}
\dot{x}_t &= f(x_t) \, dt + g(x_t) \, dw_t, \\
\dot{z}_t &= h(x_t) \, dt + dv_t,
\end{align*} \]
(1.1)
where $f, g, h$ are vector and matrix valued functions and $w, v$ are independent Wiener processes. For convenience we assume scalar observations. This paper is concerned with the problem of recursively filtering of the state $x_t$ given the past observations $z' = \{z_s; 0 \leq s \leq t\}$.

An unnormalized version of the conditional density of $x$, given $z'$ satisfies the so-called Zakai equation [14], which in Fisk–Stratonovic form looks as follows:
\[ d\rho(t,x) = \left( L\rho(t,x) - \frac{1}{2} h^2(x) \right) \, dt + h(x)\rho(t,x) \, dz_t, \]
(1.2)
where $L$ is the Fokker–Planck operator, and recently Brockett and Mitter have shown that the Lie algebra of differential operators generated by the two operators in this equation, i.e. $L - \frac{1}{2} h^2(x)$ and $h(x)$, plays an important role in (nonlinear) recursive filtering (cf. [1,2,3,11,12] and the contributions of Brockett and Mitter and others in [7]).

Now, by definition, a finite dimensional (exact) filter for some statistic $\gamma(x_t)$ of the conditional state is a system (on a finite dimensional manifold $M$) of the form
\[ d\eta_t = a(\eta_t) \, dt + b(\eta_t) \, dz_t, \]
\[ \gamma(x_t) = c(\eta_t), \]
(1.3)
i.e. it is a finite dimensional machine driven by the observations which calculates the estimate $\gamma(x_t) = E[\gamma(x_t) \mid z']$.

Now suppose that a filter (1.3) exists. We can assume that (1.3) is minimal. Then there are 2 ways of calculating $\gamma(x_t)$, viz. via (1.3) and also via (1.2) because given $\rho(x,t)$, $\gamma(x_t)$ can be obtained by first normalizing and then integrating $\gamma(x_t)$ against the normalized conditional density.

Via an as yet conjectural generalization of a result of Sussmann [13] this means that there must be a homomorphism of Lie algebra of the Lie algebra $L(\Sigma)$ generated by $L - \frac{1}{2} h^2(x)$ and $h(x)$ into the Lie algebra of vectorfields generated by the vectorfields $a(\eta), b(\eta)$ on the manifold $M$. In the case of the cubic sensor and the systems (2.1) and (2.2) to be considered in more detail below this result has been proved [6], and this implies (cf. below in Section 2) that for the cubic sensor
\[ d\eta_t = dw_t, \quad dz_t = x_t^3 \, dt + dv_t, \]
(1.4)
and for the perturbed linear systems
\[ d\eta_t = dw_t, \]
(1.5)
no finite dimensional exact filters exist for any nonzero statistic.

Let $L(\Sigma)$ be the Lie algebra of the Zakai equation of the system (1.5). It turns out (cf. Section 3 below) that the Lie algebra $L(\Sigma)$ mod $\epsilon^n$ is finite dimensional for all $n$. Now every finite
dimensional Lie algebra can be realized as a Lie algebra of vectorfields, which, conjecturally, would give us a convergent sequence of approximate filters. (In [8] and [9,10] there is positive evidence for the existence of filters corresponding to suitable homomorphisms of Lie algebras $L(\Sigma) \to V(M)$, where $V(M)$ is the Lie algebra of smooth vectorfields of a smooth manifold $M$.)

2. Perturbed linearly observed linear noise

Consider the following two one-dimensional systems:

\[
\begin{align*}
\frac{dx_i}{dt} &= dw_i, \\
\frac{dz_i}{dt} &= (x_i + ex_i^3) dt + dv_i,
\end{align*}
\]  

(2.1)

(2.2)

where $e$ is a (small) real parameter. The two operators $L_2(h_2(x))$ and $h(x)$ occurring in the Fisk–Stratonovich form of the Zakai equation are in these two cases equal to

\[
\begin{align*}
\frac{1}{2} \frac{d^2}{dx^2} - \frac{1}{2}(x + ex^2)^2, &\quad x + ex^2, \\
\frac{1}{2} \frac{d^2}{dx^2} - \frac{1}{2}(x + ex^3)^2, &\quad x + ex^3.
\end{align*}
\]  

(2.3)

(2.4)

For a given $e \neq 0$ let $L_2(e)$ be the Lie algebra generated by the two operators (2.3) and $L_3(e)$ the Lie algebra generated by the operators (2.4).

Let $W_1$ be the Weyl algebra of all differential operators (any order) with polynomial coefficients, i.e.

\[
W_1 = \mathbb{R} \langle x, \frac{d}{dx} \rangle.
\]  

(2.5)

A basis for $W_1$ consists of all operators

\[
e_{i,j} = x^i \frac{d^j}{dx^j}, \quad i, j \in \mathbb{N} \cup \{0\}.
\]  

(2.6)

We use $W'_1$ to denote the subalgebra of $W_1$ spanned by the operators (2.6) with $i - j$ even. Both $W_1$ and $W'_1$ are associative algebras but we shall only consider them as Lie algebras (with the product $[D_1, D_2] = D_1 D_2 - D_2 D_1$).

2.7. Theorem. For all $e \neq 0$, $L_2(e)$ is isomorphic to $W'_1$.

2.8. Theorem. For all $e \neq 0$, $L_3(e)$ is equal to $W_1$.

And of course $L_2(0)$ and $L_3(0)$ are both isomorphic to the four-dimensional oscillator algebra spanned by the operators

\[
\frac{1}{2} \frac{d^2}{dx^2} - \frac{1}{2} x^2, \quad x, \quad \frac{d}{dx}, \quad 1.
\]

Theorem 2.7 is not difficult to prove. The proof of Theorem 2.8 is long and computational. See the appendix.

However, we also have:

2.9. Theorem (see [4]). Let $M$ be a finite dimensional smooth manifold and $V(M)$ the Lie algebra of smooth vectorfields on $M$. Then there are no nonzero homomorphisms of Lie algebras $W_1 \to V(M)$.

The proof of this theorem can be adapted to yield the corresponding result for $W'_1$.

This implies [6] that no nonzero statistic of (2.3) and (2.4) for $e \neq 0$ admits a finite dimensional exact filter.

3. The Lie algebras $L_2(2), L_3(3) \mod e^n$

Let $W_1(e) = \mathbb{R} \langle x, e, \frac{d}{dx} \rangle$ be the Lie algebra of all differential operators with coefficients which are polynomial in $x$ and $e$. That is, a basis for $W_1(e)$ is the set of all operators

\[
x^ie^j \frac{d^k}{dx^k}, \quad i, j, k \in \mathbb{N} \cup \{0\}.
\]  

(3.1)

Let $A = \{A_1, \ldots, A_m\}$ be a set of elements of $W_1(e)$. Then $L(A)$ denotes the Lie algebra over $\mathbb{R}$ generated by $A_1, \ldots, A_m$, i.e. $L(A)$ is the $\mathbb{R}$-vectorspace spanned by $A_1, \ldots, A_m$ and all their iterated brackets. A somewhat larger algebra is $L'(A)$ which is the algebra generated by $A_1, \ldots, A_m$ over $\mathbb{R}[e]$. The algebra $L'(A)$ consists of all operators of the form $P(e)A, A \in L(A), P(e)$ a polynomial in $e$.

Now let $L \subset W_1(e)$ be a sub-Lie-algebra. Then $L \mod e^n$ is the Lie algebra obtained from $L$ by setting $e^i = 0$ for all $i \geq n$. More precisely, $e^n W_i(e)$
is an ideal in \( W_i(e) \) so that \( L \cap e^n W_i(e) \) is an ideal in \( L \) and we define
\[
L \text{ mod } e^n = L/L \cap e^n W_i(e).
\] (3.2)

### 3.3. Theorem
Let \( P(x) \) be a polynomial of degree \( n \).

\[
A = \left\{ \frac{1}{2} \frac{d^2}{dx^2} - \frac{1}{2} (x + \varepsilon P(x))^2, x + \varepsilon P(x) \right\}
\]

Then the Lie algebras \( L(A) \text{ mod } e^n \) and \( L'(A) \text{ mod } e^n \) are finite dimensional for all \( m \in \mathbb{N} \).

**Proof.** Observe that
\[
\begin{bmatrix}
x' \
x'' \\
\frac{dx'}{dx}, \frac{dx''}{dx}
\end{bmatrix}
\]
is a linear combination of terms \( x^i \frac{d^j}{dx^j} \) with \( i + j \leq r + s + l + m - 2 \). Now let \( L \) be the sub-Lie-algebra of \( W_i(e) \) consisting of all linear combinations of the expressions
\[
e^x \frac{d^k}{dx^k}, \quad j + k - (n - 1)i \leq 2.
\] (3.4)

It follows immediately from the remark made above that \( L \) is indeed a sub-Lie-algebra of \( W_i(e) \).

(We are giving \( e \) degree \( (n - 1) \) so to speak.)

Now for every \( m \in \mathbb{N} \) there are only finitely many pairs \( (j, k), j, k \in \mathbb{N} \cup \{0\} \) such that \( j + k \leq 2 + (n - 1)m \). It follows that \( L \) mod \( e^n \) is finite dimensional for all \( m \). (In fact the dimension of \( L \) mod \( e^n \) equals \( \frac{1}{2} \sum_{i=0}^{m-1} (3 + (n - 1)i)(4 + (n - 1)i) \).

Now \( \frac{1}{2} \frac{d^2}{dx^2} - \frac{1}{2} (x + \varepsilon P(x))^2 \) and \( x + \varepsilon P(x) \) are both in \( L \), and if \( A \in L \) then so is \( Q(e)A \) for every polynomial \( Q(e) \). It follows that \( L(A) \) and \( L'(A) \) are both sub-Lie-algebras of \( L \) and consequently \( L(A) \text{ mod } e^n \) and \( L'(A) \text{ mod } e^n \) are sub-Lie-algebras of \( L \) mod \( e^n \) and hence finite dimensional.

### 3.5. Example
Taking \( P(x) = x^2 \) and \( x^3 \) we find the Zakai–Lie algebras \( L_2(e), L_3(e) \) of the systems (2.3) and (2.4). Thus the Zakai–Lie algebras of these systems are finite dimensional mod \( e^n \) for all \( m \gg 0 \). And in fact we find for \( m = 2 \) that \( L_2(e) \) mod \( e^2 \) is the 10-dimensional Lie algebra with basis
\[
\frac{1}{2} \frac{d^2}{dx^2} - \frac{1}{2} x^2 - \varepsilon x^3, \quad x, \quad \frac{d}{dx}, \quad 1,
\]
and \( L_3(e) \) mod \( e^2 \) turns out to be the 14-dimensional algebra with basis
\[
\begin{align*}
\frac{d^2}{dx^2} & - \frac{1}{2} x^2 - \varepsilon x^3, \quad x, \quad \varepsilon \frac{d}{dx}, \\
1, \quad \varepsilon x, \quad \varepsilon \frac{d}{dx}, \quad \varepsilon \frac{d^2}{dx^2}, \quad \varepsilon x^2.
\end{align*}
\] (3.7)

### 3.8. Remark
Note that for the truth of Theorem 3.3 and various obvious generalizations it seems quite important that we are dealing with algebras generated by second order differential operators (and with systems which are deformations of linear systems).

### 4. Implications and conclusions
We have seen that the Zakai–Lie algebra of the perturbed linear system (2.2) is equal to \( W_i \) for all fixed \( \varepsilon \neq 0 \), and this implies that no nonzero statistic of the conditional state admits an exact finite dimensional filter. However, treating \( \varepsilon \) as an indeterminate the Zakai–Lie algebra of this system is finite dimensional mod \( e^n \) for all \( m \). Every finite dimensional Lie algebra \( L \) has a faithful representation \( L \to \mathfrak{gl}_m(\mathbb{R}) \) for some \( m \in \mathbb{N} \) and \( \mathfrak{gl}_m(\mathbb{R}) \) is realizable as a Lie algebra of vectorfields on \( \mathbb{R}^m \) by means of the embedding
\[
(a_{ij}) \to \sum_{i,j} a_{ij} \frac{\partial}{\partial x_i}.
\]

Thus the quotient Lie algebras \( L_2(3) \text{ mod } e^2 \) are realizable as Lie algebras of vectorfields which, conjecturally, means that there exist corresponding filters which then will be approximate filters for the conditional density. If the conditional density \( \rho(x,t,e) \) admits a power series expansion in \( \varepsilon \),
\[
\rho(x,t,e) = \rho_0(x,t) + \varepsilon \rho_1(x,t) + \varepsilon^2 \rho_2(x,t) + \cdots,
\]
then the filter corresponding to \( L_2(3) \text{ mod } e^2 \) calculates \( \rho_0(x,t), \rho_1(x,t), \ldots, \rho_{n-1}(x,t) \).

### Appendix
Calculation of the Lie algebra of differential operators generated by \( \frac{d^2}{dx^2} - \frac{1}{2} P^2 \), \( P \) where \( P \) is the polynomial \( P = x + \varepsilon x^3, \varepsilon \neq 0 \).
For convenience write

\[ P_i = \frac{d^i P}{dx^i}, \]

\[ e_1 = \frac{1}{2} \frac{d^2}{dx^2} - \frac{1}{2} x^2 - e x^4 - \frac{1}{2} e^2 x^6, \]

\[ e_2 = x + e x^3, \]

\[ [e_1, e_2] = P_1 \frac{d}{dx} + \frac{1}{2} P_2 \]

\[ = (1 + 3e x^2) \frac{d}{dx} + 3ex = e_3, \]

\[ e_4 = [e_3, e_2] = P_1^2 = (1 + 3e x^2)^2, \]

\[ e_5 = [e_3, e_3] = P_1 (P_1^2), \]

\[ = 2P_1^3 P_2 = 2(1 + 3e x^2)^2 6ex, \]

\[ e_6 = [e_5, e_2] = P_2 (2 P_1^2 P_2), \]

\[ = 4P_2^2 P_2^2 + 2P_1^3 P_3 \]

\[ = 2(1 + 3e x^2)^2 (72e^2 x^2 + (1 + 3e x^2) 6e). \]

Combining \( e_4 \) and \( e_6 \) we can form any element of the form

\[(1 + 3e x^2)^2 (ax^2 + b) = \]

\[ = (1 + 6e x^2 + 9e^2 x^4) (ax^2 + b) \]

\[ = b + (6eb + a)x^2 \]

\[ + (9e^2 b + 6ea)x^4 + (9e^2 a)x^6. \]

Now take \( a = 1/18, b = 2/27e. \) Then this gives the element

\[ e_7 = \frac{2}{27e} + \left(\frac{12e}{27e} + \frac{1}{18}\right)x^2 \]

\[ + \left(\frac{18e^2}{27e^2} + \frac{6e}{18}\right)x^4 + \left(\frac{9e^2}{18}\right)x^6 \]

\[ = \frac{2}{27e} + \frac{1}{2} x^2 + e x^4 + \frac{1}{2} e^2 x^6. \]

Adding this to (1) we obtain the element

\[ e_8 = \frac{1}{2} \frac{d^2}{dx^2} + \frac{2}{27e} - \frac{1}{2} \frac{d^2}{dx^2} + c. \]

\[ [2e_8, e_2] = \left[ \frac{d^2}{dx^2} + 2c, P_1 \frac{d}{dx} + \frac{1}{2} P_2 \right] \]

\[ = 2P_2 \frac{d^2}{dx^2} + P_1 \frac{d}{dx} + P_3 \frac{d}{dx} + \frac{1}{2} P_4 \]

\[ = 12ex \frac{d^2}{dx^2} + 12e \frac{d}{dx}. \]

This gives us the elements

\[ e_9 = x \frac{d^2}{dx^2} + d \]

\[ e_{10} = \frac{1}{2} [2e_8, e_9] \]

\[ = \frac{1}{2} \left[ \frac{d^2}{dx^2} + 2c, x \frac{d^2}{dx^2} + d \right] \]

\[ = \frac{1}{2} \left( \frac{d^3}{dx^3} \right) = \frac{d^3}{dx^3}. \]

\[ [e_{10}, e_2] = \left[ \frac{d^3}{dx^3}, x + e x^3 \right] \]

\[ = 3 \frac{d^2}{dx^2} + 9ex \frac{d^2}{dx^2} + 18ex \frac{d}{dx} + 6e. \]

Combining this with \( e_8 \) we see that \( L \) contains an element

\[ e_{11} = x^2 \frac{d^2}{dx^2} + 2x \frac{d}{dx} + c_2, \quad c_2 \in \mathbb{R}. \]

\[ [e_8, e_{11}] = \left[ \frac{d^2}{dx^2} + c, x \frac{d^2}{dx^2} + 2x \frac{d}{dx} + c_2 \right] \]

\[ = 4x \frac{d^3}{dx^3} + 2 \frac{d^2}{dx^2} + 4 \frac{d^2}{dx^2}. \]

This gives (in combination with \( e_8 \)) an element of the form

\[ e_{12} = x \frac{d^3}{dx^3} + c_3. \]

\[ \left[ x \frac{d^3}{dx^3} + c_3, x + e x^3 \right] = \]

\[ = 3x \frac{d^2}{dx^2} + 9ex \frac{d^2}{dx^2} + 18ex \frac{d}{dx} + 6e. \]

Subtracting \( 3e_9 \) this gives us

\[ 9e x^3 \frac{d^2}{dx^2} + 18ex \frac{d^2}{dx^2} + 6e - 3 \frac{d}{dx}, \]

and subtracting \( 6e_3 \) from this gives

\[ 9e x^3 \frac{d^2}{dx^2} + 6e - 3 \frac{d}{dx} - 6 \frac{d}{dx} - 18ex \]
Volume 1, Number 1, July 1981

SYSTEMS & CONTROL LETTERS

giving us an element
\[ e_{13} = 2ex \frac{d^2}{dx^2} + 2dx - 2ex. \]
\[ [2e_9, e_{13}] = \left[ \frac{d^2}{dx^2} + c, e_{13} \right] \]
\[ = 6ex^2 \frac{d^2}{dx^2} + 6ex \frac{d^2}{dx^2} - 4e \frac{d}{dx}, \]
which gives us an element
\[ e_{20} = \left[ \frac{d^2}{dx^2} + c, e_{13} \right] \]
\[ = 2ex^2 \frac{d^2}{dx^2} + 3ex \frac{d^2}{dx^2} + 2dx \frac{d}{dx} + c_2 \]
\[ = 4x^2 \frac{d^2}{dx^2} + 2x \frac{d^2}{dx^2} + 2x \frac{d^2}{dx^2} + 2x \frac{d^2}{dx^2} \]
\[ + 2 \frac{d}{dx} - 2x^2 \frac{d^2}{dx^2} - 2x \frac{d^2}{dx^2} \]
\[ = 2x^2 \frac{d^2}{dx^2} + 6x \frac{d^2}{dx^2} + 2 \frac{d}{dx}, \]
which gives us the element
\[ e_{20} = \left[ \frac{d^2}{dx^2} + c, e_{13} \right] \]
\[ = 2x^2 \frac{d^2}{dx^2} + 3x \frac{d^2}{dx^2} + \frac{d}{dx}. \]

Now \( e_9, e_{14} \) and \( e_{20} \) are all linear combinations of
\( x^2 \frac{d^2}{dx^2}, \frac{d}{dx} \) and the coefficient matrix has determinant
\[ \det \begin{pmatrix} 0 & 1 & -1 \\ 3 & 3 & -2 \\ 1 & 3 & 1 \end{pmatrix} = \det \begin{pmatrix} 0 & 1 & 1 \\ 0 & -6 & -5 \\ 3 & 3 & 1 \end{pmatrix} = \det \begin{pmatrix} 1 & -6 & -5 \\ 1 & 3 & 1 \end{pmatrix} = 1. \]

And it follows that the elements
\[ e_{16} = \frac{dx}{dx}, \quad e_{17} = x \frac{d^2}{dx^2}, \quad e_{18} = x^2 \frac{d^2}{dx^2} \]
are all three in \( L \). The rest is easy.
\[ [e_{16}, e_2] = \left[ \frac{dx}{dx}, x + ex^2 \right] = 1 + 3ex^2, \]
\[ [e_{16}, 1 + 3ex^2] = 6ex, \]
\[ [e_{16}, 6ex] = 6e, \]
and this gives that (combined with \( e_4, e_5, e_6 \))
\[ e_{19} = 1, \quad e_{20} = x, \quad e_{21} = x^2, \quad e_{22} = x^3, \]
\[ e_{23} = x^4, \quad e_{25} = x^5, \quad e_{26} = x^6 \]
are all in \( L \). Now
\[ [e_3, x^n] = \left[ (1 + 3ex^2) \frac{d}{dx} + 3ex, x^n \right] \]
\[ = nx^{n-1} + 3nex^{n+1} \]
which gives us that
\[ x^n \in L \quad \text{for all } n = 0, 1, 2, \ldots. \]
Combined with \( x^2 / dx^2, dx / dx \in L \) this suffices to show that
\[ x^k \frac{d^k}{dx^k} \in L \quad \forall k, l \in \mathbb{N} \cup \{0\}. \]

References