Nonlinear smoothing for random fields

Shin Ichi Aihara\textsuperscript{a,\!*}, Arunabha Bagchi\textsuperscript{b}

\textsuperscript{a} Department of Management and System Science, The Science University of Tokyo, Suma College, 5000-1 Toyohira, Chino, Nagano 391-02, Japan
\textsuperscript{b} Faculty of Applied Mathematics, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

Received January 1993; revised March 1994

Abstract

Stochastic nonlinear elliptic partial differential equations with white noise disturbances are studied in the countably additive measure set up. Introducing the Onsager–Machlup function to the system model, the smoothing problem for maximizing the modified likelihood functional is solved and the explicit form of nonlinear smoother is presented in the finitely additive observation noise set up.

\textit{AMS(MOS) Classifications:} 93E10, 93E12

\textit{Keywords:} Stochastic nonlinear elliptic equation; Smoothing; Finitely additive white noise; Maximum likelihood estimate

1. Introduction

Smoothing problem for stochastic elliptic partial differential equations has been studied by Bensoussan (1969), in which smoothing equations were first derived for the case where both the state and observation noises have nuclear covariance operators. For extending these results to the white noise disturbances, a certain functional was introduced and then smoothing equations were derived by maximizing the introduced cost. Recently, in a probabilistic framework, we formulated the smoothing problem with white noise disturbances and parameter identification problem in the sense of maximum likelihood estimate was studied in Bagchi and Aihara (1988). However, there is no apparent difference between the forms of smoothing equations derived in Bensoussan (1969) and Bagchi and Aihara (1988), because the problem considered is linear and Gaussian.

In this paper, we consider a class of nonlinear problems where the system is modeled by a nonlinear elliptic equation with white noise disturbance and the observation mechanism is linear with the finitely additive white noise as defined in

\textsuperscript{*}Corresponding author.

0304-4149/95/$09.50 \textcopyright 1995 Elsevier Science B.V. All rights reserved
\textit{SSDI} 0304-4149(93)00023-M
Kallianpur and Karandikar (1988). After formulating the stochastic nonlinear systems, we easily obtain the so-called Kallianpur–Striebel formula for the finitely additive measure set up. However, in the nonlinear problem it is almost impossible to derive the explicit form of the optimal smoother of the nonlinear boundary systems. The main idea of this paper is to use the Onsager–Machlup function corresponding to the system model considered here. Roughly speaking, from the Onsager–Machlup function we can evaluate the most probable deterministic path to the considered stochastic system state. The rigorous treatment of this function is found in Ikeda and Watanabe (1981) for the Ito equations and the related a posteriori estimation problem is studied by Zeitouni and Dembo (1987). For the stochastic nonlinear elliptic systems studied here, we use the measure transformation technique in Kuo (1971) to derive the related Onsager–Machlup function. This Onsager–Machlup function gives us the modified likelihood functional. In this paper, maximizing this modified likelihood functional, the optimal smoother will be derived. It should be noted that this optimization problem is converted into the deterministic optimal control problem as studied in Lions (1969), because the observation noise is formulated as a finitely additive white noise. Finally, from the necessary conditions of the converted optimal control problem, the explicit form of the optimal smoother can be derived.

2. Nonlinear boundary value processes with white noise inputs

2.1 Motivational example

In order to make our idea clear, first we consider a simple example in the one-dimensional case:

\[
\Sigma_e = \begin{cases} 
- a \frac{d^2 u(x)}{dx^2} + f(u(x)) = n(x) & \text{in } G = ]0, 1[, \\
u(0) = u(1) = 0, 
\end{cases}
\]

where \( n \) is a Gaussian white noise in \((H^{-1}(G), \mathcal{B}, \mu)\). Here \( \mu \) is the countably additive canonical Gauss measure on \( H^{-1}(G) \) and \( n \) is the identity map on \((H^{-1}(G), \mathcal{B}, \mu)\). For details see Bagchi and Aihara (1988). Assume that \( a \) is a positive constant and denote

\[
\langle A \phi_1, \phi_2 \rangle = \int_0^1 a \frac{\partial \phi_1}{\partial x} \frac{\partial \phi_2}{\partial x} \, dx, \quad \forall \phi_1, \phi_2 \in H^1_0(G). 
\tag{2.1}
\]

We assume that \( f \in C^1(\mathbb{R}) \) and

\[
(f(\phi_1) - f(\phi_2), \phi_1 - \phi_2) \geq 0, \quad \forall \phi_1, \phi_2 \in L^2(G) \tag{2.2}
\]

and where \((\cdot, \cdot)\) denotes the inner product in \( L^2(G) \).

In this example, the key property is that the injection from \( H^1_0(G) \) into \( L^2(G) \) is Hilbert–Schmidt, i.e., the white noise \( n \) is defined in the countably additive
measure set up in $H^{-1}(G)$. The precise form of $\Sigma_\epsilon$ is given by

$$\Sigma_\epsilon: \{ \langle Au, \phi \rangle + (f(u), \phi) = \langle n, \phi \rangle, \ \forall \phi \in H^1_0(G),$$

where $\langle \cdot , \cdot \rangle$ denotes the duality between $H^1_0(G)$ and $H^{-1}(G)$.

**Theorem 2.1.** There exists a unique solution $u$ of $\Sigma_\epsilon$ such that

$$u \in L^2(\mu; H^1_0(G)), \quad (2.3)$$

where $L^2(\mu; X) = \{ \phi \mid \int_X |\phi|^2 \, d\mu \}.$

**Proof.** Following the method proposed by Lions (1969), we can easily prove this theorem. The finite dimensional approximation of $\Sigma_\epsilon$ becomes

$$(A^m u^m, \phi) + \langle \Pi^m f(u^m), \phi \rangle = \langle \Pi^m n, \phi \rangle, \quad \forall \phi \in H^1_0(G), \quad (2.4)$$

where, with $e_i$ denoting orthonormal basis in $L^2(G)$ with values in $H^1_0(G),

$$\Pi^m = \sum_{i=1}^m e_i(e_i, \cdot), \quad \Pi^m = \sum_{i=1}^m e_i \langle e_i, \cdot \rangle \quad (2.5)$$

and

$$A^m = \sum_{i=1}^m \langle A(\cdot), e_i \rangle e_i. \quad (2.6)$$

Choosing $\phi$ as $u^m$ in (2.4), we have

$$\langle Au^m, u^m \rangle + \langle f(u^m), u^m \rangle = \langle n, u^m \rangle. \quad (2.7)$$

It follows from (2.2) that

$$\langle f(u^m) - f(0), u^m \rangle \geq 0. \quad (2.8)$$

Hence,

$$a |u^m|_{H_0^1(G)}^2 \leq \langle f(0), u^m \rangle + \langle n, u^m \rangle$$

$$\leq \langle f(0), u^m \rangle + \varepsilon |u^m|_{H^1_0(G)}^2 + C(\varepsilon) |n|_{H^{-1}(G)}^2 \quad \text{for } \varepsilon > 0, \ \exists C(\varepsilon) > 0, \quad (2.9)$$

i.e.,

$$\frac{a}{2} |u^m|_{H_0^1(G)}^2 \leq \langle f(0), u^m \rangle + C(a/2) |n|_{H^{-1}(G)}^2.$$
This implies that
\begin{equation}
\text{weakly in } L^2(\mu, H^1_0(G))
\end{equation}
and
\begin{equation}
f(u^{m'}) \to \tilde{f} \text{ (some function) weakly in } L^2(\mu; L^2(G)).
\end{equation}

Taking a limit as \( m' \to \infty \) in (2.4), we obtain
\begin{equation}
\langle Au, \phi \rangle + (\tilde{f}, \phi) = \langle n, \phi \rangle, \quad \forall \phi \in H^1_0(G).
\end{equation}

The remaining problem is to show that \( \tilde{f} = f(u) \). To do this, we use the method of monotone (Lions, 1969). From (2.2), we have for \( \phi \in L^2(\mu; L^2(G)) \)
\begin{equation}
0 \leq (f(u^{m'}) - f(\phi), u^{m'} - \phi)
= (f(u^{m'}), u^{m'}) - (f(\phi), u^{m'}) + (f(\phi), \phi).
\end{equation}

Substituting (2.7) into (2.14), it follows that
\begin{equation}
0 \leq \langle Au^{m'}, u^{m'} \rangle + \langle n, u^{m'} \rangle - (f(\phi), u^{m'}) - (f(u^{m'}), \phi) + (f(\phi), \phi).
\end{equation}

Noting that
\begin{equation}
\liminf E\{\langle Au^{m'}, u^{m'} \rangle\} \leq -E\{\langle Au, u \rangle\},
\end{equation}
we get
\begin{equation}
0 \leq -E\{\langle Au, u \rangle\} + E\{\langle n, u \rangle\} - E\{(f(\phi), u)\} - E\{\langle \tilde{f}, \phi \rangle\} + E\{f(\phi), \phi\}.
\end{equation}

On the other hand, setting \( \phi = u \) in (2.13), we have
\begin{equation}
E\{\langle Au, u \rangle\} + E\{\langle \tilde{f}, u \rangle\} = E\{\langle n, u \rangle\}.
\end{equation}

Hence
\begin{equation}
0 \leq E\{\langle \tilde{f}, u \rangle\} - E\{\langle f(\phi), u \rangle\} - E\{\langle \tilde{f}, \phi \rangle\} + E\{f(\phi), \phi\}
= E\{\langle \tilde{f} - f(\phi), u - \phi \rangle\}, \quad \forall \phi \in L^2(\mu; L^2(G)).
\end{equation}

Choosing \( u + \beta \phi, \beta > 0 \), for \( \phi \) and with \( \phi \in L^2(\mu; L^2(G)) \), we have
\begin{equation}
0 \leq E\{\langle \tilde{f} - f(u + \beta \phi), \beta \phi \rangle\}.
\end{equation}

Hence, dividing (2.20) by \( \beta \) and taking limit as \( \beta \to 0 \), we obtain
\begin{equation}
0 \leq E\{\langle \tilde{f} - f(u), \phi \rangle\}, \quad \forall \phi \in L^2(\mu; L^2(G)),
\end{equation}

implying that \( \tilde{f} = f(u) \).

The uniqueness property can easily be proved by applying the monotone property of \( f \).  \( \square \)
2.2. Abstract formulation

We consider the abstract formulation of the example $\Sigma_e$. Let $V$ and $H$ be two Hilbert spaces and $V$ be dense in $H$. Identifying $H$ with its dual, then

$$V \subset H \subset V',$$

where $V'$ is the dual of $V$.

The linear operator $A \in \mathcal{L}(V; V')$ satisfies

$$\forall \| \phi \|^2_V \leq \langle A\phi, \phi \rangle \leq \beta \| \phi \|^2_V \text{ for every } \phi \in V,$$

where $\langle \cdot, \cdot \rangle$ denotes the duality between $V$ and $V'$.

Furthermore, we need the following assumption:

(A2) the injection from $V$ into $H$ is Hilbert–Schmidt.

From this assumption the white noise $n$ can be defined in the countably additive measure set up in $V'$ (see Bagchi and Aihara, 1988; Hida, 1980). For the nonlinear function $f$, we assume that

(A3) $(f(\phi_1) - f(\phi_2), \phi_1 - \phi_2) \geq 0$ for $\phi_1, \phi_2 \in H$.

Theorem 2.2. Under (A1)–(A3), the nonlinear stochastic elliptic equation

$$\Sigma_e: \langle Au, \phi \rangle + (f(u), \phi) = \langle n, \phi \rangle \text{ for every } \phi \in V$$

has a unique solution $u$ in $L^2(\mu; V)$.

The proof of this theorem is almost the same as that of Theorem 2.1 under the assumptions (A1)–(A3).

Remark 2.1. For $G = \mathbf{]0, 1[} \times \mathbf{]0, 1[} \subset \mathbb{R}^2$, it is well known that the injection from $V = H^1_0(G)$ into $H = L^2(G)$ is not Hilbert–Schmidt. This implies that Theorem 2.2 does not cover the second-order partial differential operator case, e.g., $\partial^2(\cdot)/\partial x_1^2 + \partial^2(\cdot)/\partial x_2^2$. For the linear case studied in Bagchi and Aihara (1988), assumption (A2) is not necessary. For the nonlinear case, this strong assumption cannot be relaxed. However, in the multidimensional case, we may be able to choose $V$ and $H$ appropriately. For example, for $G = \mathbf{]0, 1[} \times \mathbf{]0, 1[}$, set

$$V = H^2_0(G) \subset H = L^2(G) \subset V' = H^{-2}(G).$$

In this case, we can show that the injection from $V$ into $H$ is Hilbert–Schmidt and the linear operator $A$ is for example given by

$$\langle A\phi_1, \phi_2 \rangle = \int_G \left\{ \frac{\partial^2 \phi_1}{\partial x_1^2} \frac{\partial^2 \phi_2}{\partial x_1^2} + \frac{\partial^2 \phi_1}{\partial x_2^2} \frac{\partial^2 \phi_2}{\partial x_2^2} \right\} dx_1 dx_2, \forall \phi_1, \phi_2 \in V.$$
3. Nonlinear transformation of Gaussian measures

In the assumption (A3), we add

\[(A3)' \quad f \in C^2 \text{ and } \sup_{x \in \mathbb{R}^r} |df(x)/dx| < \text{const.},\]

and

\[(A4) \quad \text{all coefficients of } A \text{ are regular such that } A \in \mathcal{L}(D; H) \text{ where } D = \{ \phi \in V' | A\phi \in H \}.\]

Let \( \tilde{u} \) be a solution of

\[
\langle A\tilde{u}, \phi \rangle = \langle n, \phi \rangle \quad \text{for every } \phi \in V. \tag{3.1}
\]

From Bagchi and Aihara (1988), let \( v \) be the countably additive measure induced by \( \tilde{u} \) and from Theorem 2.2, we already have

\[
\tilde{u} \in L^2(v; V). \tag{3.2}
\]

**Theorem 3.1.** Let \( \mu \) and \( v \) be countably additive measures induced by \( u \) and \( \tilde{u} \), respectively. The Radon–Nikodym derivative of \( \mu \) with respect to \( v \) is given by

\[
\frac{d\mu}{dv}(u) = \gamma(u), \tag{3.3}
\]

where

\[
\gamma(u) = \exp\{ - \langle Au, f(u) \rangle - \frac{1}{2} |f(u)|^2 + \log[\det(I + A^{-1}Df(u))] \} \tag{3.4}
\]

and where \( Df(u) \) is a Fréchet derivative with respect to \( u \).

**Proof.** First we consider the finite dimensional case. Let \( e_i \) be an orthonormal basis in \( H \) with values in \( D(A) \) and \( H^m = \text{span}[e_1, e_2, \ldots, e_m] \). Define the orthonormal projector \( \Pi^m \in \mathcal{L}(H; H^m) \),

\[
\Pi^m = \sum_{i=1}^m (\cdot, e_i)e_i, \tag{3.5a}
\]

and its extension to \( \mathcal{L}(V'; H^m) \),

\[
\tilde{\Pi}^m = \sum_{i=1}^m \langle \cdot, e_i \rangle e_i. \tag{3.5b}
\]

From (3.1) we have

\[
A^m\tilde{u}^m = n^m, \tag{3.6}
\]

where \( A^m = \tilde{\Pi}^mA \) and \( n^m = \tilde{\Pi}^mn \).

For every \( \phi \in V' \), \( |\tilde{\Pi}^m\phi - \phi|_{V'} \to 0 \) as \( m \to \infty \) and then we have

\[
\tilde{u}^m \to \tilde{u} \text{ strongly in } V \text{ and } v\text{-a.s.} \tag{3.7}
\]
From the nonlinear system $\Sigma_z$ and (3.6), we obtain
\[ A^m u^m + \Pi^m f(u^m) = A^m \hat{u}^m. \quad (3.8) \]

From (A1), $A^m$ becomes invertible and then
\[ u^m + (A^m)^{-1} \Pi^m f(u^m) = \tilde{u}^m. \quad (3.9) \]

Define the following nonlinear operator $T^m$:
\[ u^m = T^m(\tilde{u}^m) = (I + K^m)(\tilde{u}^m), \quad (3.10) \]
where
\[ K^m(\tilde{u}^m) = -(A^m)^{-1} \Pi^m f(u^m) = -(A^m)^{-1} f(u^m). \]

For the convenience of description, we use the symbol $\tilde{\phi}$ such that
\[ \tilde{\phi} = [(\phi, e_1), (\phi, e_2), \ldots, (\phi, e_m)]', \quad (3.11) \]
and
\[ \tilde{A} = \text{matrix whose } ij\text{th component is } \langle Ae_i, e_j \rangle. \quad (3.12) \]

From (3.9), we have
\[ \tilde{u}^m + (\tilde{A}^m)^{-1} \tilde{f}(\tilde{u}^m) = \tilde{u}^m, \quad (3.13) \]
where
\[ \tilde{f}(\tilde{u}^m) = [(f(u^m), e_1), (f(u^m), e_2), \ldots, (f(u^m), e_m)]'. \]

In the sequel, we also use the symbol $\tilde{T}^m$ and $\tilde{K}^m$ for the corresponding operators as defined by (3.12). For the $m$-dimensional process $\tilde{u}^m$, we find that its distribution is Gaussian such that
\[ v(\tilde{u}^m \in B) = \left( \frac{1}{2\pi} \right)^{m/2} |\det [(\tilde{A}^m)^{-1} (\tilde{A}^m)^{-1}]|^{-1/2} \exp \left\{ \frac{1}{2} |\tilde{A}^m \tilde{u}^m|^2 \right\} \mathrm{d}\tilde{u}^m. \quad (3.14) \]

For any $g \in L^1(\mu; \mathbb{R}^1)$, we have
\[ \int_{\mathbb{R}^n} g(\tilde{u}^m) \mathrm{d}\mu_{\tilde{u}^m} = \int_{\mathbb{R}^n} g(\tilde{T}^m \tilde{u}^m) \mathrm{d}v_{\tilde{T}^m \tilde{u}^m} \]
\[ = \int_{\mathbb{R}^n} g(\tilde{T}^m \tilde{u}^m) (2\pi)^{-m/2} |\det [(\tilde{A}^m)^{-1} (\tilde{A}^m)^{-1}]|^{-1/2} \]
\[ \times \exp \left\{ -\frac{1}{2} |\tilde{A}^m \tilde{u}^m|^2 \right\} \mathrm{d}\tilde{u}^m \quad (\text{from the Jacobi theorem}) \]
\[ = \int_{\mathbb{R}^n} g(\tilde{u}^m) (2\pi)^{-m/2} |\det [(\tilde{A}^m)^{-1} (\tilde{A}^m)^{-1}]|^{-1/2} \]
\[ \times \exp \left\{ -\frac{1}{2} |\tilde{A}^m \tilde{T}^{-1} \tilde{u}^m|^2 \right\} \]
\[ \times \det \{ (I + (\tilde{A}^m)^{-1} \tilde{D}f(\tilde{u}^m)) \} \, d\tilde{u}^m \]

\[ = E_{\nu_{\tilde{u}^m}} \{ g(\tilde{u}^m) \exp \left\{ \frac{1}{2} |\tilde{A}^m \tilde{u}^m|^2 \right\} - \frac{1}{2} |\tilde{A}^m \tilde{T}^m - \tilde{u}^m|^2 \} \times \det \left\{ (I + (\tilde{A}^m)^{-1} \tilde{D}f(\tilde{u}^m)) \right\}. \] (3.15)

From (3.10), we have

\[ (T^m)^{-1} = I - K^m(T^m)^{-1}. \] (3.16)

Hence, (3.15) becomes

\[ = E_{\nu_{\tilde{u}^m}} \{ g(\tilde{u}^m) \exp \left\{ \frac{1}{2} |\tilde{A}^m \tilde{K}^m(\tilde{T}^m)^{-1} \tilde{u}^m, \tilde{A}^m \tilde{u}^m|^2 \right\} - \frac{1}{2} |\tilde{A}^m \tilde{K}^m(\tilde{T}^m)^{-1} \tilde{u}^m|^2 \} \times \det \left\{ (I + (\tilde{A}^m)^{-1} \tilde{D}f(\tilde{u}^m)) \right\}. \]

Noting that

\[ K^m(T^m)^{-1} \tilde{u}^m = K^m \tilde{u}^m = -(A^m)^{-1} f(\tilde{u}^m), \] (3.17)

we have (3.15)

\[ = E_{\nu_{\tilde{u}^m}} \{ g(\tilde{u}^m) \exp \left( -(\Pi^m f(\tilde{u}^m), A^m u^m) - \frac{1}{2} |\Pi^m f(\tilde{u}^m)|^2 \right) \times \det \left\{ (I + (A^m)^{-1} Df(\tilde{u}^m)) \right\}; \]

\[ = E_{\nu_{\tilde{u}^m}} \{ g(\tilde{u}^m) \exp \left( -(\Pi^m f(\tilde{u}^m), A^m u^m) - \frac{1}{2} |\Pi^m f(\tilde{u}^m)|^2 \right) \]

\[ + \log \{ \det \left\{ (I + (A^m)^{-1} Df(\tilde{u}^m)) \right\} \}. \]

Hence the Radon-Nikodym derivative \( \gamma \) is given by

\[ \gamma(\tilde{u}^m) = \exp \left\{ - \langle A\tilde{u}^m, \Pi^m f(\tilde{u}^m) \rangle - \frac{1}{2} |\Pi^m f(\tilde{u}^m)|^2 \right\} \]

\[ + \log \{ \det \left\{ (I + (A^m)^{-1} Df(\tilde{u}^m)) \right\} \}. \] (3.18)

The remaining problem is to check the convergence of \( \gamma(\tilde{u}^m) \) as \( m \to \infty \). In (3.18), \( u^m \) is \( \nu_{\tilde{u}^m} \)-Gaussian and then for \( u^m \) process, the property (3.7) holds, i.e.,

\[ u^m \to u \text{ strongly in } V \text{ and } \nu \text{-a.s.} \] (3.19)

Then, from (A3)', i.e., \( \sup_{x \in \mathbb{R}^1} |df(x)/dx| \leq \text{const.} \), we have for \( \phi \in V \)

\[ f(\phi) \in V. \] (3.20)

Hence,

\[ \langle A\tilde{u}^m, \Pi^m f(\tilde{u}^m) \rangle \to \langle A\tilde{u}, f(\tilde{u}) \rangle \text{ \( \nu \)-a.s., as } m \to \infty \] (3.21)

and

\[ \frac{1}{2} |f(\tilde{u}^m)|^2 \to \frac{1}{2} |f(\tilde{u})|^2 \text{ \( \nu \)-a.s., as } m \to \infty . \] (3.22)

We must check the convergence of the \( \det \{ I + (A^m)^{-1} Df(\tilde{u}^m) \} \) term, as \( m \to \infty \). It is easy to show that

\[ (A^m)^{-1} Df(\tilde{u}^m) \to A^{-1} Df(\tilde{u}) \text{ strongly in } L_1(H; H) \text{ and } \nu \text{-a.s., as } m \to \infty . \] (3.23)
Hence, from the fact that \( \text{tr}[A^{-1}Df(u)] < \infty \) a.s., we find that \( \det \{ I + (A^m)^{-1} Df(u^m) \} \) converges to a nonzero value. The original proof for a more abstract situation can be found in Kuo (1971). This completes the proof. \[ \square \]

**Theorem 3.2.** For any deterministic \( v \in \{ v \in D \mid \langle A\phi, Av \rangle < \infty \text{ for every } \phi \in V \} \),

\[
\mu \left\{ \| u - v \|_V \leq \varepsilon \right\} = E_v \left\{ \exp \left\{ - \langle A(u), f(u) \rangle - \frac{1}{2} |f(u)|^2 + \log \left[ \det \{I + A^{-1}Df(u)\} \right] \right\} \right\}.
\]

\[
\| u \|_V \leq \varepsilon.
\]

**Proof.** From Theorem 3.1, we have

\[
\mu \left\{ \| u - v \|_V \leq \varepsilon \right\} = E_v \left\{ \exp \left\{ - \langle A(u), f(u) \rangle - \frac{1}{2} |f(u)|^2 + \log \left[ \det \{I + A^{-1}Df(u)\} \right] \right\} \right\}.
\]

From \( \langle A\phi, Av \rangle < \infty \) for every \( \phi \in V \), for the transformation \( u - v = \tilde{u} \), we have

\[
\frac{dv_u}{dv_{\tilde{u}}} = \exp \left\{ - \langle A\tilde{u}, Av \rangle - \frac{1}{2} |Av|^2 \right\}.
\]

Hence, (3.25) leads to (3.24). \[ \square \]

**Theorem 3.3.** For \( v \in D \) and \( u \) (solution of (22s)), we have

\[
\lim_{\varepsilon \to 0} \frac{\mu \left\{ \| u - v \|_V \leq \varepsilon \right\}}{v \left\{ \| A^{-1}n \|_V \leq \varepsilon \right\}} = \exp \left\{ - L(v) \right\},
\]

where

\[
L(v) = \langle Av, f(v) \rangle + \frac{1}{2} |f(v)|^2 + \frac{1}{2} |Av|^2 - \log \left[ \det \{I + A^{-1}Df(v)\} \right].
\]

**Proof.** First, we determine the sequence \( v^k \in \{ v^k \in D \mid \langle A\phi, Av^k \rangle < \infty \text{ for every } \phi \in V \} \) with \( v^k \to v \) strongly in \( D(A) \). Then, for \( v^k \), it follows from Theorem 3.2 that

\[
\lim_{\varepsilon \to 0} \frac{\mu \left\{ \| u - v^k \|_V \leq \varepsilon \right\}}{v \left\{ \| \tilde{u} \|_V \leq \varepsilon \right\}} = \exp \left\{ - L(v^k) \right\}.
\]

Hence, we can take the limit in (3.29) as \( k \to \infty \). Noting that \( \tilde{u} = A^{-1}n \), (3.28) follows. \[ \square \]

**Remark 3.1.** From (3.27), we can evaluate the most probable deterministic path to \( u \) such that for two processes \( v_1 \) and \( v_2 \in D \)

\[
\lim_{\varepsilon \to 0} \frac{\mu \left\{ \| u - v_1 \|_V \leq \varepsilon \right\}}{\mu \left\{ \| u - v_2 \|_V \leq \varepsilon \right\}} = \lim_{\varepsilon \to 0} \frac{\mu \left\{ \| u - v_1 \|_V \leq \varepsilon \right\}}{\mu \left\{ \| A^{-1}n \|_V \leq \varepsilon \right\}} \frac{\mu \left\{ \| u - v_2 \|_V \leq \varepsilon \right\}}{\mu \left\{ \| A^{-1}n \|_V \leq \varepsilon \right\}} = \exp \left\{ - L(v_1) + L(v_2) \right\}.
\]
4. Nonlinear smoothing

Consider a distributed observation mechanism of the form

\[ y = u + n_0 \quad \text{in } G, \tag{4.1} \]

where \( u \) is a solution of \( \Sigma_s \) and \( n_0 \) is a finitely additive white noise in \( H \), independent of \( n \). We denote the canonical sample space for the system state by \( V \) with \( \sigma \)-field \( \mathcal{A} \) and the measure \( \mu \). The observation noise \( n_0 \) is defined in \((H, \mathcal{C}, m)\) where \( m \) is the canonical Gauss measure. To make (4.1) meaningful, \( u \) and \( n_0 \) should be defined in a single probability space. To this end, define

\[ E = H \times V \quad \text{and} \quad \mathcal{F} = \bigcup_{P \in \mathcal{P}} \mathcal{G}_P \otimes \mathcal{A}, \]

where \( \mathcal{G}_P \otimes \mathcal{A} \) is the usual product \( \sigma \)-field and \( P \in \mathcal{P} = \text{the class of projection on } H \) with finite dimensional range. For \( P \in \mathcal{P} \), let \( \alpha_P \) be the usual product of \( m \) restricted to \( \mathcal{G}_P \) and \( \mu \). It is easy to see that the \( \alpha_P \)'s determine a finitely additive probability on \( \mathcal{F} \) such that \( \alpha = \alpha_P \) on \( \mathcal{G}_P \otimes \mathcal{A} \).

The basic result in signal estimation is the white noise version of the Kallianpur–Striebel formula in Kallianpur and Karandikar (1988).

**Proposition 4.1.** For any \( g \) integrable on \((V, \mu)\),

\[ E_\mu \{ g(u) | y \} = \frac{E_\mu \{ g(u) \exp \left\{ - \frac{1}{2} |u|^2 + (u, y) \right\} \}}{E_\mu \{ \exp \left( - \frac{1}{2} |u|^2 + (u, y) \right) \}} \tag{4.2} \]

and the likelihood functional \( \text{RN}(y) \) is given by

\[ \text{RN}(y) = E_\mu \{ \exp \left( - \frac{1}{2} |u|^2 + (u, y) \right) \}. \tag{4.3} \]

For a linear system, we can derive the explicit form of smoother from (4.2). For nonlinear systems, however, it is impossible to have an exact form of smoother. In this paper we try to derive a smoother by maximizing a modified likelihood functional with the aid of results derived in the previous section.

The likelihood functional \( \text{RN}(y) \) is modified by

\[ \text{RN}_\varepsilon(v, y) = E_\mu \{ \exp \left( - \frac{1}{2} |u|^2 + (u, y) \right) \mid |u - v|_V \leq \varepsilon \} \tag{4.4} \]

for \( v \in D(\mathcal{A}) \). Applying Theorem 3.3, we have

\[ \lim_{\varepsilon \to 0} \frac{\text{RN}_\varepsilon(v, y)}{\mu(\{|A^{-1}n|_V \leq \varepsilon \})} = \exp \left( - L(v) - \frac{1}{2} |v|^2 + (v, y) \right). \tag{4.5} \]

We call the right-hand side of (4.5) the modified likelihood functional, which is denoted by \( \text{MRN}(v, y) \), because of the fact that

\[ \lim_{\varepsilon \to 0} \frac{\text{RN}_\varepsilon(v_1, y)}{\text{RN}_\varepsilon(v_2, y)} = \lim_{\varepsilon \to 0} \frac{\text{RN}_\varepsilon(v_1, y)/\mu(\{|A^{-1}n|^2 \leq \varepsilon \})}{\text{RN}_\varepsilon(v_2, y)/\mu(\{|A^{-1}n|^2 \leq \varepsilon \})} = \frac{\text{MRN}(v_1, y)}{\text{MRN}(v_2, y)}. \tag{4.6} \]
Hence, the nonlinear smoothing problem is to seek an optimal \( v^0 \in D(A) \) which maximizes the cost \( \text{MRN}(v, y) \) with respect to \( v \in D(A) \). This optimization problem is equivalent to

\[
v^0 = \arg\max_{v \in D} \{ \text{MRN}(v, y) \}
\]

\[
= \arg\min_{v \in D} \{ -\log \text{MRN}(v, y) \}
\]

\[
= \arg\min_{v \in D} \left\{ \frac{1}{2} |Av + f(v)|^2 + \frac{1}{2} |v| - (v, y) - \log \det \{(I + A^{-1} Df(v))\} \right\}
\]

\[
= \arg\min_{v \in D} \left\{ \frac{1}{2} |Av + f(v)|^2 + \frac{1}{2} |v - y|^2 - \log \det \{(I + A^{-1} Df(v))\} \right\} \quad (4.7)
\]

Moreover, the optimization problem (4.7) is converted into the following optimal control problem:

\[
\Sigma_0:
\begin{align*}
\text{System model for } g \in H
\end{align*}
\]

\[
Av + f(v) = g,
\]

\[
\text{Cost functional}
\]

\[
J(g) = \frac{1}{2} |g|^2 + \frac{1}{2} |v - y|^2 - \log \det \{I + A^{-1} Df(v)\}.
\]

For formulating the above optimization problem, we need the following restriction:

\[
(A\phi, f(\phi)) \geq -C(1 + |A\phi| + |f(\phi)|) \quad \text{for every } \phi \in D(A), \quad \text{and some } C > 0.
\]

**Lemma 4.1.** Under all the assumptions stated above, \( Av + f(v) = g \) has a unique solution in \( D(A) \).

**Proof.** See Barbu (1983, p. 42). \( \square \)

For further development, we assume

(A6) \( f \in C^2 \) and \( \sup_{x \in \mathbb{R}^1} |d^2 f(x)/dx^2| \leq \text{const.} \)

**Lemma 4.2.** For \( \delta v \in H \),

\[
\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \left\{ \log \det \{I + A^{-1} Df(v + \varepsilon \delta v)\} - \log \det \{I + A^{-1} Df(v)\} \right\}
\]

\[
= \left( \delta v, \sum_{i=1}^{\infty} e_i \frac{d^2 f(v)}{d\delta v^2} (I - A^{-1})^* \left(I + A^{-1} \frac{d f(v)}{d\delta v} I\right)^{-1} e_i \right), \quad (4.8)
\]

where \( \{e_i\} \) is an orthonormal basis in \( H \).

**Proof.** From the definition of \( \det \{ \cdot \} \), we have

\[
\det \{I + A^{-1} Df(v)\} = \prod_{i=1}^{\infty} \left( e_i + A^{-1} \frac{d f(v)}{d\delta v} e_i, e_i \right). \quad (4.9)
\]
Noting that the nonlinear function is monotone, we find that \( \frac{df(x)}{dx} \geq 0 \ \forall x \in \mathbb{R}^1 \). Then, from (A6) we have

\[
\lim_{\epsilon \to 0} \frac{1}{\epsilon} \log \left\{ \frac{\det \left[ I + A^{-1} Df(v + \epsilon \delta v) \right]}{\det \left[ I + A^{-1} Df(v) \right]} \right\} = \sum_{i=1}^{\infty} \left( A^{-1} \left( \frac{d^2 f(v)}{dv^2} \right) \delta v e_i, e_i \right) = \left( \delta v, \sum_{i=1}^{\infty} e_i \frac{d^2 f(v)}{dv^2} \right) (- A^{-1})^* \left( I + A^{-1} \frac{df(v)}{dv} I \right)^{-1} e_i \right) \quad \square \tag{4.10}
\]

**Proposition 4.2.** The map \( g \to v(g) \) is strongly continuous from \( H \) into \( V \).

**Proof.** Let \( g^n \) be a weakly convergence sequence in \( H \) converging to \( g \). Noting that the injection from \( H \) into \( V' \) is Hilbert-Schmidt, we have

\( g^n \to g \) strongly in \( V' \).

With the aid of the monotone property of \( f \), it is easy to show that

\[
|v(g^n) - v(g^n)|^2 \leq C |g^n - g^n|^2.
\]

Hence

\[
\lim_{n \to \infty} v(g^n) = v \ (\text{a solution of system state } \Sigma_0) \text{ strongly in } V. \quad \square
\]

**Theorem 4.1.** There exists at least one optimal control \( g^0 \) for minimizing the cost \( J(g) \).

**Proof.** Let \( d = \inf\{J(g); \ g \in H\} \). From the assumption (A3)', we find that

\[
|\log \left\{ \det \left[ \left( I + A^{-1} Df \right) \right] \right\}| \leq \text{const. for every } v \in H. \text{ Hence, we have } -\infty < d < \infty.
\]

Now let \( \{g^n\} \) be such that

\[
d < J(g^n) < d + 1/n.
\]

Noting that \( J(g) \) is a quadratic form with respect to \( g \), we find that \( g^n \) is weakly compact in \( H \). Hence from Proposition 4.2 we can extract a subsequence (which is still denoted by \( v^n \)) such that

\[
g^n \to g^* \text{ weakly in } H, \\
v^n \to v^* \text{ strongly in } V.
\]

Hence, the first and second terms of the cost \( J \) are quadratic and the last term \( \log \{ \cdot \} \) is continuous in \( V \) by Lemma 4.2, so that

\[
J(g^*) = \frac{1}{2} |g^*|^2 + \frac{1}{2} |v^* - y|^2 - \log \left\{ \det \left[ I + A^{-1} Df(v^*) \right] \right\} = d.
\]

In other words, \( g^* \) is an optimal control \( g^0 \). \quad \square
Theorem 4.2. The optimal control $g^0$ is given by

$$g^0 = -p,$$  \hspace{1cm} (4.11)

where

$$\left( p, A\phi + \frac{df(v^0)}{dv^0} \phi \right) = (v^0 - y, \phi) - \left( \sum_{i=1}^{\infty} e_i \frac{d^2f(v^0)}{dv^0^2} (-A^{-1})^* \left( I + A^{-1} \frac{df(v^0)}{dv^0} I \right)^{-1} e_i, \phi \right)$$

for every $\phi \in D$, and $v^0 = v(g^0)$. \hspace{1cm} (4.12)

Proof. We can directly apply the optimal control theory for the deterministic elliptic nonlinear systems studied by Lions (1969) and Barbu (1983).

Defining

$$\delta v = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} (v(g^0 + \varepsilon \delta g) - v(g^0)), \quad \delta g \in H,$$

we have

$$A\delta v + \frac{df(v^0)}{dv^0} \delta v = \delta g,$$  \hspace{1cm} (4.14)

and

$$\delta J(g^0) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} (J(g^0 + \varepsilon \delta g) - J(g^0))$$

$$= (g^0, \delta g) + (v^0 - y, \delta v) - \left( \sum_{i=1}^{\infty} e_i \frac{d^2f(v^0)}{dv^0^2} (-A^{-1})^* \left( I + A^{-1} \frac{df(v^0)}{dv^0} I \right)^{-1} e_i, \delta v \right).$$  \hspace{1cm} (4.15)

Then, in (4.12) setting $\phi = \delta v$, we have

$$\left( p, A\delta v + \frac{df(v^0)}{dv^0} \delta v \right) = (v^0 - y, \delta v)$$

$$- \left( \sum_{i=1}^{\infty} e_i \frac{d^2f(v^0)}{dv^0^2} (-A^{-1})^* \left( I + A^{-1} \frac{df(v^0)}{dv^0} I \right)^{-1} e_i, \delta v \right).$$  \hspace{1cm} (4.16)

Further, it follows from (4.14) that

$$\left( A\delta v + \frac{df(v^0)}{dv^0} \delta v, p \right) = (\delta g, p).$$  \hspace{1cm} (4.17)

Hence, from (4.16) and (4.17), we have $\delta J(g^0) = (g^0 + p, \delta g)$. This implies (4.11) using (4.12). \hspace{1cm} $\Box$
Proposition 4.3. The exact form of the optimal smoother is given by

\[ A v^0 + f(v^0) = - \left( A^* + \frac{d f(v^0)}{d v^0} I \right)^{-1} (v^0 - y) \]

\[ - \left( \sum_{i=1}^{\infty} e_i \frac{d^2 f(v^0)}{d v^0} ( - A^{-1})^* \left( I + A^{-1} \frac{d f(v^0)}{d v^0} I \right)^{-1} e_i \right), \]

where \( e_i \) is an orthonormal basis in \( H \).

This proof is a direct consequence of Theorem 4.2.

5. An example and differential operator representation of the smoother

Physical examples of the nonlinear elliptic systems are often found in the free boundary problem (Lions, 1969; Barbu, 1983). Here we consider a simple obstacle, i.e., the system \( u \) has a constraint \( "u \geq 0" \). It is well known that such an obstacle problem can be formulated as a nonlinear system, although this nonlinearity does not have the regularity property stated in (A6). Usually for studying this problem, the regularization and penalization techniques are used to convert the original problem into a regular one. Considering the system \( \Sigma \) again with the constraint \( "u \geq 0" \) and using the regularization method in Lions (1969), the nonlinear function \( f \) is given by

\[
\begin{cases}
0 & \text{for } u \geq 0, \\
\frac{1}{6\delta} u^3 & \text{for } -\delta \leq u \leq 0, \\
\frac{1}{2} u^2 - \frac{\delta}{2} u - \frac{1}{6} \delta^2 & \text{for } u \leq -\delta,
\end{cases}
\]

where \( \delta > 0 \). (Taking a limit as \( \delta \to 0 \), we can formulate the original problem as the stochastic variational inequality. Here we do not discuss such a problem.)

We must check all the assumptions stated previously. It is sufficient to consider (A3), (A5) and (A6). From (5.1), we get

\[
\frac{d f}{du} = \begin{cases}
0 & \text{for } u \geq 0, \\
\frac{1}{2\delta} u^2 & \text{for } -\delta \leq u \leq 0, \\
- u - \frac{\delta}{2} & \text{for } u \leq -\delta,
\end{cases}
\]

and hence \( df/du \geq 0 \). Furthermore

\[
\frac{d^2 f}{du^2} = \begin{cases}
0 & \text{for } u \geq 0, \\
\frac{1}{\delta} u & \text{for } -\delta \leq u \leq 0, \\
- 1 & \text{for } u \leq -\delta
\end{cases}
\]
and we have $\sup_{\text{sup}} |d^2f/du^2| \leq 1$. The remaining problem is to check the assumption (A5). From (2.1), integrating by parts, we have

\[
(A\phi, f(\phi)) = a \int_0^1 \frac{\partial \phi}{\partial x} \frac{\partial f(\phi)}{\partial x} \, dx
\]

\[
= a \int_0^1 \left( \frac{\partial \phi}{\partial x} \right)^2 \frac{df(\phi)}{d\phi} \, dx
\]

\[
\geq 0 \quad \text{(from 5.2)).} \quad (5.4)
\]

Then, the system $\Sigma_e$ with the nonlinear function $f$ given by (5.1) satisfies all the assumptions stated in this paper.

For realizing the smoother by using a digital computer, the adjoint form given in Theorem 4.2 is more convenient than the explicit form (4.18). Set

\[
q(x) = \left( \sum_{i=1}^{\infty} e_i \frac{d^2f(v^0)}{dv^2} (I - A^{-1})^* \left( I + A^{-1} \frac{df(v^0)}{dv^0} I \right)^{-1} e_i \right). \quad (5.5)
\]

Then, from Theorem 4.2, the optimal smoother can be represented as the following differential forms:

\[
-a \frac{d^2v^0(x)}{dx^2} + f(v^0) = -p(x) \quad \text{for } x \in ]0, 1[.
\]

\[
v^0(0) = v^0(1) = 0 \quad (5.6)
\]

and

\[
-a \frac{dp(x)}{dx^2} + \frac{df(v^0)}{dv^0} p(x) = v^0(x) - y(x) - q(x) \quad \text{for } x \in ]0, 1[,
\]

\[
p(0) = p(1) = 0. \quad (5.7)
\]

At present we cannot obtain the exact differential form of $q(x)$.

6. Conclusions

Stochastic nonlinear elliptic systems with white noise disturbances are formulated and the related Onsager–Machlup function is derived. The nonlinear smoothing problem is solved by maximizing the modified likelihood functional related to the Onsager–Machlup function. It is possible to generalize the situation considered here to the stochastic variational inequality case. In this case, the final optimization problem becomes very complicated and the smoothing equation cannot be explicitly obtained.

References

A. Bensoussan, L'indentification and le Filtrage (IRIA, Cahier, No. 1, 1969).
T. Hida, Brownian Motion (Springer, Berlin, 1980).
J.L. Lions, Quelques Méthodes de résolution des problèmes aux limites non lineaires (Dunod, Paris, 1969).