INFINITE DIMENSIONAL ESTIMATION THEORY APPLIED TO A WATER POLLUTION PROBLEM

Ruth F. Curtain
Control Theory Centre, University of Warwick, Coventry CV4 7AL, U.K.

1. INTRODUCTION

There is now a fairly complete theory for filtering, prediction and smoothing for linear infinite dimensional systems where one assumes a Gaussian white noise type disturbance in the system model (see [1], [2], [3], [5]). Recently in the finite dimensional stochastic control literature, there has been interest in problems involving jump processes (see [9]) as these have applications to a wide range of problems. For example in [8], Kwakernaak applies the filtering theory for linear systems excited by Poisson white noise with noisy observations corrupted by Gaussian white noise to solve a river pollution problem. In fact the model of the river is distributed and so he uses a finite dimensional approximation. Here we show how the same problem can be solved using an infinite dimensional model by appropriately modelling the Poisson type noise as an infinite dimensional stochastic process and applying the recent theory of estimation for linear systems disturbed by general noise processes in [7]. The necessary mathematical background of stochastic differential equations in Hilbert space is outlined in § 2 and the results on infinite dimensional estimation theory for linear systems corrupted by Poisson type noise processes are summarized in § 3. In the final section § 4, the model for river pollution is discussed in detail and comparisons are made with the model of Kwakernaak in [8].

2. STOCHASTIC DIFFERENTIAL EQUATIONS IN HILBERT SPACES

The theory we summarize here is a special case of the theory for stochastic equations with general white noise disturbance [4].

Let H, K be real separable Hilbert spaces, (Ω, P, μ) a complete probability space and $T = [0,T]$, a real finite time interval.

Then we shall use the following definitions

Definition 2.1

An H-valued random variable is a map $u: \Omega \to H$ which is measurable with respect to the μ measure.

If $u \in L_1(\Omega, \mu; H)$, we define its expectation

$$E\{u\} = \int_{\Omega} u \, d\mu$$

If $u \in L_2(\Omega, \mu; H)$, we define its covariance operator

$$Cov(u) = E\{(u - E\{u\}) \circ (u - E\{u\})\}$$

where $u_o u \in \mathcal{L}(H)$ is self adjoint nuclear operator given by

$$(u_o u)\, h = u \langle u, h \rangle \text{ for all } h \in H$$

<u>Definition 2.2</u>

An H-valued stochastic process is a map $u(\cdot,\cdot) : T \times \Omega \to H$ which is measurable on $T \times \Omega$ using the Lebesgue measure on T.

Consider the following estimation problem:
Estimate the random variable $x \in L_2(\Omega, \mu; H)$ from the random variable $y \in L_2(\Omega, \mu; K)$. Let $\tilde{L}_2(K, \sigma; H)$ be the closed subspace of $L_2(\Omega, \mu; H)$ which is isometric to $L_2(K, \sigma; H)$, where (K, σ) is the probability spaced induced by y. Then we shall be concerned with two types of estimates.

<u>Definition 2.3</u>

The best global estimate $\hat{x} = E\{x \mid y\}$ of x from y is the projection of x on $\tilde{L}_2(K, \sigma; H)$.

The best linear estimate $\bar{x}$ of x from y is $\bar{x} = \Lambda_o y$, where $\Lambda_o \in \mathcal{L}(K, H)$ minimizes $E\{\|x - \Lambda y\|^2\}$ over all $\Lambda \in \mathcal{L}(K, H)$. $\&$

If x and y are Gaussian, then the linear and global estimates are identical.

The following class of stochastic process is used for our application model and is a special case of the orthogonal increments process introduced in [4].

<u>Definition 2.4</u>

An H-valued compound Poisson process $\{q(t); t \in T\}$ is such that

$$q(t) = \sum_{i=0}^{\infty} q_i(t)\, e_i \tag{2.1}$$

where $\{e_i\}$ is a complete orthonormal basis for H and where $\{q_i(t); t \in T\}$ is a real compound Poisson process such that

$$\begin{cases} E\{q_i(t) - q_i(s)\} = \mu_i (t-s) \\ E\{(\bar{q}_i(t_1) - \bar{q}_i(s_1))(\bar{q}_j(t_2) - \bar{q}_j(s_2))\} = 0 \quad ; \ 0 \le s_1 < t_1 \le s_2 < t_2 \le T \\ E\{(\bar{q}_i(t) - \bar{q}_i(s))(\bar{q}_j(t) - \bar{q}_j(s))\} = \lambda_{ij} (t-s) \quad ; \ 0 \le s \le t \le T \end{cases} \tag{2.2}$$

where $\bar{q}_i(t) = q_i(t) - \mu_i t$ and $\sum_{i=0}^{\infty} \mu_i < \infty$

For convenience we rewrite (2.2) as

$$(2.3)\quad \begin{cases} E\{q(t)\} = t \sum_{i=0}^{\infty} \mu_i e_i \\ Cov(\bar{q}(t) - \bar{q}(s)) = \Lambda (t-s) \qquad ; \quad 0 \le s \le t \le T \end{cases}$$

where $\bar{q}(t) = q(t) - t \sum_{i=0}^{\infty} \mu_i e_i$

and Λ is the covariance operator of $q(t)$ and is such that

$$(2.4)\qquad \Lambda e_i = \sum_{j=0}^{\infty} \lambda_{ij} e_j$$

We remark that $\lambda^2_{ij} \le \lambda_i \lambda_j \quad (\lambda_{ii} = \lambda_i)$ and

$$E\{\|\bar{q}(t) - \bar{q}(s)\|^2\} = trace \Lambda (t-s) = \sum_{i=0}^{\infty} \lambda_i (t-s)$$

We now define a stochastic integral for these processes.

Definition 2.5

Let $q(t)$ be as in definition 2.4 and suppose $\Phi \in \mathcal{B}_2(T; \mathcal{L}(H,K))$, the class of strongly-measurable $\mathcal{L}(H,H)$ -valued functions with $\int_T \|\Phi\|^2 ds < \infty$. Then we define the stochastic integral

$$\int_0^t \Phi(s)\, dq(s) = \sum_{i=0}^{\infty} \int_0^t \Phi(s)\, d\bar{q}_i(s) + \sum_{i=0}^{\infty} \mu_i \int_0^t \Phi(s) e_i\, ds$$

$\int_0^t \Phi(s)\, dq(s)$ is a well defined H -valued stochastic process with the properties

$$(2.5)\qquad E\left\{ \int_0^t \Phi(s)\, dq(s) \right\} = \sum_{i=0}^{\infty} \mu_i \int_0^t \Phi(s) e_i\, ds$$

$$(2.6)\qquad E\left\{ \left\| \int_0^t \Phi(s)\, dq(s) \right\|^2 \right\} \le trace \Lambda \int_0^t \|\Phi(s)\|^2 ds$$

We shall be concerned with the following stochastic evolution equation

$$(2.7)\qquad \begin{cases} du(t) = Au(t)dt + B\, dq(t) \\ u(0) = u_0 \end{cases}$$

where A is the infinitesimal generator of an analytic semigroup $\mathcal{J}_t$ on H, $B \in \mathcal{L}(H)$, $q(t)$ is as in definition 2.4, $u_0 \in L_2(\Omega, \mu; H)$ and $u(t)$ is an -valued stochastic process.

By a solution of (2.7) we shall mean the following:

Definition 2.6

(2.7) has a strong solution u, if $u \in C(T; L_2(\Omega,\mu; H))$, $u(t) \in \mathcal{D}(A(t))$ w.p.i. and $u(t)$ satisfies (2.8) almost everywhere on $T \times \Omega$. u is unique if whenever u_1 and u_2 are strong solutions,

$$\mu\{\omega : \sup_{t\in T} \| u_1(t) - u_2(t) \| = 0\} = 1$$

Theorem 2.1

(2.7) has the unique solution

$$u(t) = J_t u_0 + \int_0^t J_{t-s} B\, dq(s)$$

provided the following extra assumptions are satisfied.

$$\sum_{i=0}^{\infty} \lambda_i \int_0^t \| A J_t B e_i \|^2 ds < \infty \tag{2.8}$$

$$\sum_{i=0}^{\infty} \mu_i \int_0^t \| A J_t B e_i \| ds < \infty \tag{2.9}$$

3. ESTIMATION THEOREY FOR INFINITE DIMENSIONAL SYSTEMS WITH POISSON TYPE DISTURBANCE

We summarize the results of [7], specialized to time invariant systems corrupted by Poisson type noise disturbance.

With H, T and (Ω, P, μ) as in § 2, consider the following abstract system and observation models.

$$u(t) = J_t u_0 + \int_0^t J_{t-s} B\, dq(s) \tag{3.1}$$

$$z(t) = \int_0^t C u(s)\, ds + w(t) \tag{3.2}$$

where J_t is an analytic semigroup on H with generator A, $B \in \mathcal{L}(H)$, $q(t)$ is as in definition 2.4, $u_0 \in L_2(\Omega, \mu; H)$, $C \in \mathcal{L}(H, R^k)$ and W is a k-dimensional Wiener process with incremental covariance matrix the identitiy.

The estimation problem is to find the best linear unbiased estimate $\hat{u}(t|t_0)$ of the state $u(t)$ at time t, based on the observations $z(s)$; $0 \le s \le t$.

Under the above assumptions, the optimal filter $\hat{u}(t) = \hat{u}(t|t)$ is given by

(3.3) $$\hat{u}(t) = v(t) + \int_0^t \mathcal{Y}(t,s) P(s) C^* dz(s)$$

where $$v(t) = \sum_{i=0}^{\infty} \mu_i \int_0^t \mathcal{Y}(t,s) B e_i ds$$

and $\mathcal{Y}(t,s)$ is the unique solution of

(3.4) $$\mathcal{Y}(t,s)x = \mathcal{J}_{t-s}x - \int_s^t \mathcal{J}_{t-\rho} P(\rho) C^* C \mathcal{Y}(\rho,s) x \, d\rho$$

and $P(t)$ is the unique solution of the Riccati equation in the class of self adjoint weakly continuous operator-valued functions with $\langle x, P(t) y\rangle$ absolutely continuous for $x, y \in \mathcal{D}(A^*)$.

(3.5) $$\begin{cases} \frac{d}{dt} \langle P(t)x, y\rangle - \langle P(t)x, A^* y\rangle - \langle A^* x, P(t) y\rangle \\ \qquad + \langle P(t) C^* C P(t) x, y\rangle = \langle B \Lambda B^* x, y\rangle \quad \text{a.e.} \\ P(0) = P_0 \qquad ; \qquad x, y \in \mathcal{D}(A^*). \end{cases}$$

Moreover, $$P(t) = \mathrm{Cov}\,(u(t) - \hat{u}(t)).$$

The optimal smoother $\hat{u}(t \mid t_0)$, where $t_0 > t$ is given by

(3.6) $$\hat{u}(t \mid t_0) = \hat{u}(t) + P(t)\lambda(t)$$

where $$\lambda(t) = \int_t^{t_0} \mathcal{Y}^*(s,t) C^* (dz(s) - C\hat{u}(s) ds)$$

The optimal predictor $\hat{u}(t \mid t_0)$, where $t > t_0$ is given by

(3.7) $$\hat{u}(t \mid t_0) = \sum_{i=0}^{\infty} \mu_i \int_0^t \mathcal{J}_{t-s} B e_i ds + \mathcal{J}_{t-t_0} \hat{u}(t_0)$$

To obtain differential equations for the estimators we need to impose the following extra assumptions

(3.8) $$\sum_{i=0}^{\infty} \lambda_i \int_0^t \| \mathcal{A} \mathcal{J}_{t-s} \mathcal{B} e_i \| ds < \infty$$

(3.9) $$\sum_{i=0}^{\infty} \lambda_i \| \mathcal{A} \mathcal{J}_t P_0 e_i \| < \infty$$

Then writing $\hat{u}(t) = v(t) + y(t)$, $y(t)$ is the unique solution of the stochastic evolution equation

$$(3.10) \quad \begin{cases} dy(t) = (A - P(t)C^*C)\,y(t)\,dt + P(t)C^*C\,u(t)\,dt + P(t)C^*dw(t) \\ y(0) = 0 \end{cases}$$

and $V(t)$ is the unique solution of the deterministic differential equation

$$(3.11) \quad \begin{cases} \dot{V}(t) = (A - P(t)C^*C)\,V(t) + \sum_{i=0}^{\infty} \mu_i B e_i \\ V(0) = 0 \end{cases}$$

For the smoothing problem $t_0 > t$, writing

$\hat{u}(t|t_0) = V(t) + y(t, t_0)$, we have that $y(t, t_0)$ is the unique strong solution of the stochastic evolution equation

$$(3.12) \quad \begin{cases} dy(t, t_0) = (A - P(t)C^*C)\,y(t, t_0)\,dt + P(t)C^*C\,\hat{u}(t)\,dt + B\Lambda B^*\lambda(t)\,dt \\ y(t_0, t_0) = y(t_0) \end{cases}$$

For the prediction problem $t > t_0$, writing

$$\hat{u}(t|t_0) = \bar{u}(t) + y(t, t_0)$$

where $\bar{u}(t) = E\{u(t)\}$, we have that $y(t, t_0)$ is the unique strong solution of the stochastic evolution equation

$$(3.13) \quad \begin{cases} dy(t, t_0) = A\,y(t, t_0)\,dt \\ y(t, t_0) = y(t_0) \end{cases}$$

and $\bar{u}(t)$ is the unique solution of

$$(3.14) \quad \begin{cases} \dot{\bar{u}}(t) = A\,\bar{u}(t) + \sum_{i=0}^{\infty} \mu_i B e_i \\ \bar{u}(0) = 0 \end{cases}$$

4. POLLUTION MODEL

The problem considered by Kwakernaak in [8] is an environmental problem of river pollution. For his model of the pollution by chemical wastes he assumes that the number of deposits in a section of the river of infinitesimal length dx (x being the distance coordinate along the river) behave according to a Poisson process with rate parameter $\lambda(x)dx$ where $\lambda(x)$ is a given function. He assumes that the number of deposits in non-overlapping sections are independent processes and that the amounts of chemical deposited each time at location x are independent stochastic variables with given distribution H_x and characteristic function ϕ_x. The time evolution of the concentration of the chemical at location x at time t, $y(t,x)$ is supposed to be given by

$$(4.1) \qquad \frac{\partial y(t,x)}{\partial t} = D\frac{\partial^2 y(t,x)}{\partial x^2} - V\frac{\partial y(t,x)}{\partial x} + \xi(t,x) \quad ; \quad 0 \le x \le \ell.$$

where D is the dispersion coefficient, V the water velocity and $\xi(t,x)$ is the rate of increase of the concentration at (t,x) due to the deposits of chemical wastes described above.

Along the river are a finite number of measuring stations which continuously measure the observed local concentrations of the chemical. Kwakernaak solves the problem of reconstructing the times, locations and amounts of deposits from the measured data by approximating (4.1) by a finite dimensional model and applying the finite dimensional filtering theory using a martingale approach.

Here we consider the problem of estimating the concentration of the chemical at any point along the river at any time in the past or future, based on continuous noisy measurements in time at a finite number of measuring stations along the river. Our approach is quite different in that we retain the infinite dimensional character of the problem throughout and show how the final equations can be solved by eigenfunction approximations.

4.1 Characterization of the pollution process

In order that (4.1) describes a stochastic differential equation it is necessary to define $\xi(t,x)$ as some suitable distributed stochastic process. Following the discussion in §2, we try a model of the following type

$$(4.2) \qquad \begin{aligned} dy(t) &= A_0\, y(t)\,dt + dq(t) \\ y(0) &= y_0 \end{aligned}$$

where $y(t) \in H = L_2(0,\ell)$ for each $t \in [0,T]$ models the amount of pollutant dumped at (t,x) and A_0 is the diffusion operator defined by

$$(4.3) \qquad \left\{ \begin{aligned} A_0 y &= D\frac{\partial^2 y}{\partial x^2} - V\frac{\partial y}{\partial x} \\ \mathcal{D}(A_0) &= \left\{ y \in H : \frac{\partial y}{\partial x}, \frac{\partial^2 y}{\partial x^2} \in H \text{ and } y = 0 \text{ at } x = 0, \ell \right\} \end{aligned} \right.$$

In order to apply the theory of stochastic differential equations on a Hilbert space from §2, we need to establish that $q(t)$ is a compound Poisson process according to definition 2.4.

In [8], Kwakernaak defines $q(t)$ formally in terms of a random functional, that is

$F_{q_t} : H \to L_2(\Omega \times T; R)$ is defined by

$$F_{q_t}(u) = \langle q(t), u \rangle_H \tag{4.4}$$

From his assumptions on the polluting process, he deduces that for each $u \in H$, $F_{q_t}(u)$ is a real valued compound Poisson process with rate parameter $\bar{\lambda} = \int_0^\ell \lambda(x)dx$ and jumps with characteristic function $1/\bar{\lambda} \int_0^\ell \lambda(x)\phi_x(\gamma\, u(x))\,dx$. $F_{q_t}(u)$ has characteristic function

$$E\{e^{i\gamma F_q(u)}\} = \exp\left\{ t \int_0^\ell \lambda(x)[\phi_x(\gamma u(x)) - 1]\,dx \right\} \tag{4.5}$$

We now show that provided $\lambda(x)$ and H_x are suitably chosen $F_{q_t}(\cdot)$ induces an H-valued stochastic process $q(t)$ defined by (4.4)

<u>Lemma 4.1</u>

Suppose $\{H_x ; 0 \le x \le \ell\}$ is a real stochastic process with characteristic function $\phi_x(\gamma)$ and $\lambda(x)$ is a real function on $(0,\ell)$.

Let $H = L_2(0,\ell)$ and $\{e_k\}_{k=0}^\infty$ be the following orthonormal bases for H

$$e_k(x) = \sqrt{\tfrac{2}{\ell}}\, \sin\left(\frac{k\pi x}{\ell} + \epsilon_k\right) \quad ; \tan \epsilon_k = -\frac{2\pi k D}{\ell V} \quad ; k = 1, 2, . \tag{4.6}$$

Under the following assumptions on λ and H_x

$$(\lambda(x)E\{H_x\})^2,\ \lambda(x)E\{H_x^2\} \in L_2(0,\ell) \tag{4.7}$$
$$(\lambda(x)\, E\{H_x\})^2 = \sum_{k=1}^\infty \beta_k\, e_{2k}(x)$$
$$\lambda(x)E\{H^2 x\} = \sum_{k=1}^\infty \gamma_k\, e_{2k}(x)$$

$q(t)$ induced by (4.4) is a well-defined H-valued stochastic process $\in L_2(\Omega \times T; H)$; $q(t) = \sum_{k=1}^\infty q_k(t)\, e_k$, where $q_k(t)$ is a real compound Poisson process with characteristic function

$$\exp\left\{ t \int_0^\ell \lambda(x)[\phi_x(\gamma e_k(x)) - 1]\,dx \right\}$$

If we further assume that

(4.8) $$E\{(q_K(t)-E\{q_K(t)\})(q_i(t)-E\{q_i(t)\})\} = \lambda_{ij}\, t$$

then $q(t)$ is compound Poisson process according to definition 2.4 with

(4.9) $$\begin{cases} E\{q(t)\} = t\sum_{K=0}^{\infty} \mu_K e_K \\ Cov(q(t)) = t\,\Lambda \end{cases}$$

where Λ is a nuclear, self adjoint operator on H, given by

$$\Lambda e_K = \sum_{i=1}^{\infty} \lambda_{Ki} e_i \quad ; \quad trace\ \Lambda = \sum_{i=1}^{\infty} \lambda_{ii} < \infty .$$

and

(4.10 $$\begin{cases} \mu_K = \int_0^{\ell} \lambda(x) e_K(x) E\{H_x\}\, dx \\ \lambda_K = \lambda_{KK} = \int_0^{\ell} \lambda(x) e_K^2(x) E\{H_x^2\}\, dx \\ \lambda_{ij}^2 \le \lambda_i \lambda_j \end{cases}$$

Proof

From (4.4), $q(t) = \sum_{K=0}^{\infty} q_K(t) e_K$; where $q_K(t) = Fq_t(e_K)$.

So $q_K(t)$ is a real compound Poisson process and has the characteristic function $\exp\{t\int_0^{\ell}\lambda(x)[\phi_x(\gamma e_K(x))-1]\,dx\}$ and hence its moments are

(4.11) $$E\{q_K(t)\} = t\int_0^{\ell}\lambda(x)e_K(x)E\{H_x\}\,dx = \mu_K t$$, say

(4.12) $$E\{q_K^2(t)\} - E\{q_K(t)\}^2 = t\int_0^{\ell}\lambda(x)e_K^2(x)E\{H_x^2\}\,dx = \lambda_K t$$, say

and $q(t) \in L_2(\Omega \times T; H)$ provided $\sum_{K=0}^{\infty}\mu_K < \infty$ and $\sum_{K=1}^{\infty}\lambda_K < \infty$

From (4.6), it may be verified that

$$\int_0^{\ell} e_{2K}(x)\, e_i^2(x)\, dx = 0 \qquad \text{for} \qquad i \ne K$$

and $$\int_0^{\ell} e_{2K}(x)\, e_K^2(x)\, dx = O\left(\tfrac{1}{K}\right)$$

So $$\sum_{i=1}^{\infty}\lambda_i = \sum_{i=1}^{\infty}\gamma_i \langle e_{2i}, e_i^2\rangle = \sum_{i=1}^{\infty}\gamma_i\, O\left(\tfrac{1}{i}\right) < \infty .$$

and similarly $\sum_{i=1}^{\infty}\mu_i < \infty$. (Note that $\mu_0 = 0 = \lambda_0$).

So under assumption (4.7), $q(t) \in L_\nu(\Omega \times T; H)$

By construction $q_K(t)$ has independent increments and assumption (4.8) ensures $q(t)$ satsfies definition 2.4,(4.9) and (4.10) follow from (4.11), (4.12).

Remark

(4.7)(b) is necessary, since $\int_0^\ell e_k(x)\, e_k^2(x)\, dx \not\to 0$ as $k \to \infty$ for odd k.

4.2 State and observation models

In the filtering theory, both A and its adjoint appear in the equations and since A_0 is not self adjoint this makes an eigenfunction expansion approach difficult. So instead we transform the original model (4.2) to the following:

$$(4.13) \quad \begin{cases} du(t) = Au(t)\,dt + B\,dq(t) \\ u(0) = u_0 \end{cases}$$

where $y(t,x) = e^{ax} u(t,x)$; $y_0 = e^{ax} u_0$; $a = V/2D$.

A is a self adjoint operator on H given by

$$(4.14) \quad Ah = \frac{\partial^2 h}{\partial x^2} - a^2 h$$

$$\mathcal{D}(A) = \{ h \in H : \tfrac{\partial h}{\partial x}, \tfrac{\partial^2 h}{\partial x^2} \in H \text{ and } ah(x) + \tfrac{\partial h}{\partial x} = 0 \text{ at } x = 0, \ell \}$$

and $B \in \mathcal{L}(H)$ is given by

$$(4.15) \quad (Bh)(x) = e^{-ax} h(x) \quad \text{for} \quad h \in H$$

We suppose that the initial state u_0 has zero expectation and covariance operator P_0 given by

$$(4.16) \quad P_0 e_K = \alpha_K e_K \quad ; \quad \sum_{K=1}^{\infty} \alpha_K < \infty \quad ; \alpha_0 = 0.$$

A has the eigenfunctions $\{e_i\}$ given by (4.6) and generates the analytic semigroup $\mathcal{J}_t$ given by

$$(4.17) \quad (\mathcal{J}_t h)(x) = \sum_{n=1}^{\infty} \langle h, e_n \rangle e^{(a^2 - n^2/\ell^2)t} e_n(x) \quad \text{for} \quad h \in H$$

(4.13) has the well-defined mild solution

(4.18) $$u(t) = J_t u_0 + \int_0^t J_{t-s} B\, dq(s)$$

which we can take as our state model for the evolution of the concentration of pollutant in the river.

Now we show that (4.18) is a strong solution of (4.13) (see definition 2.6).

From theorem 2.1, we need to verify

(4.19) $$\sum_{i=1}^{\infty} \lambda_i \int_0^t \| A J_{t-s} B e_i \|^2 ds < \infty$$

(4.20) $$\sum_{i=1}^{\infty} \mu_i \int_0^t \| A J_{t-s} B e_i \| ds < \infty$$

Writing $B e_i = \sum_{j=0}^{\infty} \mu_{ij} e_j$, we find

$$A J_{t-s} B e_i = -\frac{\pi^2}{\ell^2} \sum_{j=1}^{\infty} \mu_{ij}\, j^2 e^{(a^2 - j^2/\ell^2)(t-s)} e_j$$

and
$$\int_0^t \| A J_{t-s} B e_i \|^2 ds \leq \text{const.} \sum_{j=0}^{\infty} \frac{j^4 \mu_{ij}^2}{(j^2 - a^2 \ell^2)}$$

By direct calculation, we find that for $i \neq j$

(4.21) $$\mu_{ij} = \int_0^\ell e^{-ax} e_i(x)\, e_j(x)\, dx$$

$$= \frac{\text{const.}\ \ ij\,(3a^2\ell^2 - i^2 - j^2)}{\sqrt{\pi^2 j^2 + a^2\ell^2}\ \sqrt{\pi^2 i^2 + a^2\ell^2}\ (a^2\ell^2 + \pi^2(i+j)^2)(a^2\ell^2 + \pi^2(i-j)^2)}$$

and since $\sum_{i=1}^{\infty} \lambda_i < \infty$, we have

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \frac{j^4 \lambda_i \mu_{ij}^2}{(j^2 - a^2\ell^2)} < \infty$$

and (4.19) is satisfied.

(4.20)is similarly verified.

For our observation process we take

(4.22) $$dz(t) = C u(t)\, dt + dw(t)$$

where $w(t)$ is a k dimensional Wiener process with covariance matrix the identity and $C \in \mathcal{L}(H, R^k)$ is given by

(4.23) $$(Cu)_j = \frac{1}{2\epsilon} \int_{x_j - \epsilon}^{x_j + \epsilon} u(x)\, dx \qquad \text{for small} \quad \epsilon > 0$$

This approximates point observations at the fixed locations $x_1, \ldots, x_k$.

4.3 Solution to the estimation problem

Our model in § 4.2 satisfies all the assumptions of the theory of § 3 and so there exists a unique optimal estimator, given by (3.3) - (3.7).

In order to obtain computable solutions we first obtain recursive equations for our estimates and the covariance of the error process $P(t)$. From § 3, $P(t)$ is the unique solution of the differential Riccati equation

$$(4.24)\quad \begin{cases} \frac{d}{dt}\langle P(t)f,h\rangle - \langle P(t)f, Ah\rangle - \langle Af, P(t)h\rangle + \langle P(t)C^*Cf,h\rangle \\ \qquad\qquad = \langle B\Lambda B^* f, h\rangle \quad a.e. \\ P(0) = P_0 \quad \text{for} \quad f,h \in \mathcal{D}(A) \quad \text{(note that} \qquad \text{).} \end{cases}$$

We shall try for a solution $P(t)$ of the form

$$(4.25)\quad P(t)\,h = \sum_{i,j=1}^{\infty}\sum^{\infty} p_{ij}(t)\,e_i(x)\langle e_j, h\rangle \qquad ; \quad p_{ij}(t) = p_{ji}(t).$$

Substituting (4.24) into (4.23) and equating coefficients of $h = \sum_{i=1}^{\infty} h_i e_i$, $f = \sum_{i=1}^{\infty} f_i e_i$ we obtain

$$(4.26)\quad \begin{cases} \frac{d}{dt} p_{ij}(t) + \frac{\pi^2}{\ell^2}(i^2+j^2)\,p_{ij}(t) + \sum_{n,m=1}^{\infty}\sum^{\infty} p_{ni}(t)p_{mj}(t)\,A_{mn} = \rho_{ij} \\ p_{ij}(0) = \delta_{ij}\,\alpha_i \end{cases}$$

where $\rho_{ij} = \sum_{k=1}^{\infty}\sum_{n=1}^{\infty} \lambda_{kn}\,\mu_{in}\,\mu_{jk}$ (see 4.21)

and $A_{mn} = \sum_{r=1}^{K} a_{mr}\,a_{nr}$

$$a_{sr} = \frac{\sqrt{2\ell}}{s\pi\epsilon}\,\sin\frac{s\pi\epsilon}{\ell}\,\sin\!\left(\frac{s\pi x_r}{\ell} + \epsilon_s\right) = (Ce_s)_r$$

We now verify that the additional assumptions (3.8) and (3.9) are satisfied and hence $\hat{u}(t/t_0)$ may be expressed in differential form.

From § 3, we must verify the following

$$(4.28)\quad \sum_{i=1}^{\infty} \lambda_i \int_0^t \| A J_s e_i \|\, ds < \infty$$

$$(4.29)\quad \sum_{i=1}^{\infty} \lambda_i \| A J_t P_0 e_i \| < \infty \qquad \text{for} \quad t > 0$$

Now $$A J_t P_0 e_i = -\alpha_i \frac{i^2\pi^2}{\ell^2}\, e^{(a^2 - i^2/\ell^2)t}\, e_i$$

and so $\| A J_t P_0 e_i \| \leq$ constant for $t > 0$

and $\sum_{i=1}^{\infty} \lambda_i < \infty$, ensures that (4.29) holds.

$$\int_0^t \| A J_s e_i \| \, ds = \left| \frac{i^2 \pi^2}{(\ell^2 a^2 - i^2)} \left(1 - e^{(a^2 - i^2/\ell^2)t}\right) \right|$$

$$\leq \quad \text{const.}$$

and so again (4.28) is satisfied since $\sum_{i=1}^{\infty} \lambda_i < \infty$.

So $\hat{u}(t|t_0)$ may be expressed in differential form. For the optimal filter let us try the following expansion:

$$(4.30) \qquad \hat{u}(t) = \sum_{i=1}^{\infty} v_i(t) e_i + \sum_{i=1}^{\infty} \beta_i(t) e_i$$

where $v(t) = \sum_{i=1}^{\infty} v_i(t) e_i$ is deterministic. Substituting in (3.10) and (3.11) we obtain the following equations for $\beta_i(t)$ and $v_i(t)$

$$(4.31) \quad \begin{cases} d\beta_i(t) = -\dfrac{i^2\pi^2}{\ell^2} \beta_i(t)\, dt + \displaystyle\sum_{n=1}^{\infty} \sum_{r=1}^{\infty} A_{nr} \beta_r(t) p_{in}(t)\, dt \\ \qquad\qquad + \displaystyle\sum_{n=1}^{\infty} p_{ni}(t) \sum_{r=1}^{k} a_{nr}\, dz_r(t) \\ \beta_i(0) = 0 \end{cases}$$

$$(4.32) \quad \begin{cases} \dot{v}_i(t) = -\dfrac{i^2\pi^2}{\ell^2} v_i(t) + \displaystyle\sum_{n=1}^{\infty} \sum_{r=1}^{\infty} A_{nr} v_r(t) p_{in}(t) + \sum_{r=1}^{\infty} \mu_r \mu_{ri} \\ v_i(0) = 0 \end{cases}$$

So the filtering problem may be solved recursively from the infinite systems, (4.26), (4.31), (4.32) for $p_{ij}(t)$, $\hat{u}_i(t)$, $v_i(t)$ respectively, using the usual truncation methods.

For the prediction problem $t > t_0$ writing

$$\hat{u}(t|t_0) = \tilde{u}(t) + y(t, t_0)$$

$$= \sum_{i=1}^{\infty} \gamma_i(t) e_i + \sum_{i=1}^{\infty} y_i(t, t_0) e_i$$

and substituting in (3.13) and (3.14) we obtain the following equations for $\gamma_i(t)$, $y_i(t, t_0)$.

$$(4.33) \quad \begin{cases} dy_i(t, t_0) = -\dfrac{\pi^2}{\ell^2} i^2 y_i(t, t_0)\, dt \\ y_i(t_0, t_0) = \beta_i(t_0) \end{cases}$$

(4.34)
$$\begin{cases} \dot{\gamma}_i(t) = -\frac{\pi^2 i^2}{\ell^2}\gamma_i(t) + \sum_{k=1}^{\infty}\mu_k\mu_{ik} \\ \gamma_i(0) = 0 \end{cases}$$

Similarly for the smoothing problem $t_0 > t$, writing

$$\hat{u}(t/t_0) = v(t) + \sum_{i=1}^{\infty} y_i(t, t_0)$$

you may obtain equations for $y_i(t, t_0)$.

We remark that $\beta_i(t)$, $y_i(t, t_0)$ and $x_i(t, t_0)$ are unique strong solutions of stochastic evolution equations, whereas $v_i(t)$ and $\gamma_i(t)$ are deterministic functions.

Returning to the original problem of estimating the amount of concentration $\hat{y}(t/t_0; x)$ of pollutant at (t, x) based on measurements $y(x_1, s)$, .. $y(x_k, s); 0 \leq s \leq t_0$, we have

$$\hat{y}(t/t_0; x) = e^{\alpha x}\,\hat{u}(t/t_0)(x)$$

Finally we note that the estimators $\hat{u}(t/t_0)$ obtained are the best linear estimates and not the best global estimates, because the noise process $q(t)$ is Poisson-like. In the Gaussian case the best linear estimates are also the best global estimates (see § 2). The advantage with working with linear estimates is that you obtain Kalman-Bucy type recursive equations.

In his finite dimensional approximation Kwakernaak found the best global estimate and so obtained an infinite set of filtering equations which he then proceeded to develop approximation algorithms for. In his conclusions he notes that little if any improvement was obtained over the Kalman filter and this is not surprising since the Kalman filter approach obtains the best linear estimate to a high degree of accuracy; whereas the global estimate was obtained using approximating algorithms, which will incur errors.

In our infinite dimensional approach we again obtain Kalman Bucy type equations, but an infinite system this time. However they can be conveniently solved by an eigenfunction approach.

REFERENCES

1. BENSOUSSAN, A. "Filtrage Optimal des Systèmes Linéaires", Dunod, 1971.

2. CURTAIN, R.F. "Infinite Dimensional Filtering", SIAM J. Control, 1975.

3. CURTAIN, R.F. "A Survey of Infinite Dimensional Filtering", SIAM Review, 17, 1975.

4. CURTAIN, R.F. "Stochastic Evolution Equations with General White Noise Disturbance", Control Theory Centre Report No. 41, University of Warwick, 1975.

5. CURTAIN, R.F. "Infinite Dimensional Estimation Theory for Linear Systems", Control Theory Centre Report No. 38, University of Warwick, 1975.

6. CURTAIN, R.F. & PRITCHARD, A.J. "The Infinite Dimensional Riccati Equation for Systems Defined by Evolution Operators", 1975 (to appear in SIAM J. Control).

7. CURTAIN, R.F. "Estimation Theory for Abstract Evolution Equations Excited by General White Noise Processes", Control Theory Centre Report No. 40, University of Warwick, 1975.

8. KWAKERNAAK, H. "Filtering for Systems Excited by Poisson White Noise", Int. Symposium on Control Theory, Numerical Methods & Computer Systems Modelling, 1974. (Lecutre Notes in Economics & Math. Systems, 107, Springer Verlag.)

9. VARAIYA, P. "Filtering and Control of Jump Processes", Ibid.