# The Filtering Problem

Shino Takayama
School of Economics
University of Queensland

# The Filtering Problem

* Given the observations $Z_s$ satisfying: for $0 \le s < t$,

(observations) $dZ_t = c(t, X_t)dt + \gamma(t, X_t)dV_t;\ Z_0 = 0$ (6.1.6)

what is the best estimate $\hat{X}_t$ of the state $X_t$ of:

(system) $dX_t = b(t, X_t)dt + \sigma(t, X_t)dU_t$ (6.1.2)

based on these observations?

# More Detailed Description

* By saying that the estimate $\hat{X}_t$ is based on the observations $\{Z_s; s \le t\}$

 we mean that

$$\hat{X}_t(\cdot) \quad \text{is} \quad \mathcal{G}_t - \text{measurable},$$

$$\text{where} \quad \mathcal{G}_t \ \text{is the} \ \sigma\text{-algebra generated by} \ \{Z_s(\cdot); s \le t\}. \quad (6.1.7)$$

# “The Best Such Estimate”

- By saying that $\hat{X}_t$ is the best such estimate, we mean that

$$\int_{\Omega} | X_t - \hat{X}_t |^2 \, dP = E[|X_t - \hat{X}_t|^2] = \inf\{E[| X_t - Y |^2 ; Y \in \mathcal{K}\}. (6.1.8)$$

- $(\Omega, \mathcal{F}, P)$ is the probability space corresponding to the $(p+r)$-dimensional Brownian motion $(U_t, V_t)$ starting at $0$

$E$ denotes expectation w.r.t. $P$ and

$$\mathcal{K} := \mathcal{K}_t := \mathcal{K}(Z, t)$$

$$:= \{Y : \Omega \to \Re^n ; Y \in L^2(P) \text{ and } Y \text{ is } \mathcal{G}_t - \text{measurable}\} (6.1.9)$$

where $L^2(P) = L^2(\Omega, P).$

# Lemma 6.1.1

* Let $\mathcal{H} \subset \mathcal{F}$ be a $\sigma$-algebra and let $X \in L^2(P)$ be $\mathcal{F}$- measurable. Put

$$\mathcal{N} = \{Y \in L^2(P); Y \text{ is } \mathcal{H}-\text{measurable}\}$$

and let $\mathcal{P}_{\mathcal{N}}$ denote the orthogonal projection from the Hilbert space $L^2(P)$ into the subspace $\mathcal{N}$. Then,

$$\mathcal{P}_{\mathcal{N}}(X) = E[X \mid \mathcal{H}].$$

# Proof: Lemma 6.1.1

- Proof: The existence and uniqueness of $E[X \mid \mathcal{H}]$ comes from the Radon-Nikodym theorem:
  - Let $\mu$ be the measure on $\mathcal{H}$ defined by
$$\mu(H) = \int_H X dP; \quad H \in \mathcal{H}.$$
  - Then, $\mu$ is absolutely continuous w.r.t. $P \mid \mathcal{H}$, so there exists a $P \mid \mathcal{H}$-unique $\mathcal{H}$-measurable function $F$ on $\Omega$ such that
$$\mu(H) = \int_H F dP \quad \text{for all } H \in \mathcal{H}.$$
  - Then, $E[X \mid \mathcal{H}] := F$ is unique w.r.t. $P \mid \mathcal{H}$.

# Proof Continued

- Now, $\mathcal{P}_{\mathcal{N}}(X)$ is $\mathcal{H}$-measurable and
$$\int_{\Omega} Y(X - \mathcal{P}_{\mathcal{N}}(X))dP = 0 \quad \text{for all } Y \in \mathcal{N}.$$
- In particular,
$$\int_{A} (X - \mathcal{P}_{\mathcal{N}}(X))dP = 0 \quad \text{for all } A \in \mathcal{H}.$$
- Hence, by uniqueness,
$$\mathcal{P}_{\mathcal{N}}(X) = E[X \mid \mathcal{H}].$$

# Theorem 6.1.2

$$\hat{X}_t = \mathcal{P}_{\mathcal{K}_t}(X_t) = E[X_t \mid \mathcal{G}_t].$$

# The 1-Dimensional Linear Filtering Problem

* Consider the 1-dimensional case:

(System)

$$dX_t = F(t)X_t dt + C(t)dU_t; \quad F(t), C(t) \in \Re \ (6.2.3)$$

(Observations)

$$dZ_t = G(t)X_t dt + D(t)dV_t; \quad G(t), D(t) \in \Re \ (6.2.4)$$

* Assume that $F, G, C, D$ are bounded on bounded intervals.
* Assume that $Z_0 = 0$.
* Assume that $X_0$ is normally distributed and independent of $\{U_t\}, \{V_t\}$.
* Assume that $D(t)$ is bounded away from $0$ on bounded intervals.

# Step 1

- Let $\mathcal{L} = \mathcal{L}(Z,t)$ be the closure in $L^2(P)$ of functions which are linear combinations of the form

$$c_0 + c_1 Z_{s_1}(\omega) + \cdots + c_k Z_{s_k}(\omega), \quad \text{with } s_j \le t, c_j \in \Re.$$

- Let

$$\mathcal{P}_{\mathcal{L}} \text{ denote the projection from } L^2(P) \text{ onto } \mathcal{L}.$$

- Then, with $\mathcal{K}$ as in (6.1.9),

$$\hat{X}_t = \mathcal{P}_{\mathcal{K}}(X_t) = E[X_t \mid \mathcal{G}_t] = \mathcal{P}_{\mathcal{L}}(X_t).$$

- Then, the best $Z$-measurable estimate of $X_t$ coincides with the best $Z$-linear estimate of $X_t$ .

# Step 1: Lemma 6.2.2

* Let $X, Z_s; s \le t$ be random variables in $L^2(P)$ and assume that

$$(X, Z_{S_1}, Z_{S_2}, \cdots, Z_{S_n}) \in \Re^{n+1}$$

has a normal distribution for all $S_1, S_2, \cdots, S_n \le t, n \ge 1.$

Then,

$$\mathcal{P}_{\mathcal{L}}(X) = E[X \mid \mathcal{G}] = \mathcal{P}_{\mathcal{K}}(X).$$

# Proof: Lemma 6.2.2

* Let

$$\breve{X} = \mathcal{P}_{\mathcal{L}}(X), \tilde{X} = X - \breve{X}.$$

* Then, we claim that $\tilde{X}$ is independent of $\mathcal{G}$ .
* Recall that a random variable $(Y_1, \cdots, Y_k) \in \Re^k$ is normal if and only if $c_1 Y_1 + \cdots + c_k Y_k$ is normal for all choices of $c_1, \cdots, c_k \in \Re$.
* An $L^2$-limit of normal variables is again normal (by Appendix A).
* Therefore,

$$(\tilde{X}, Z_{s_1}, \cdots, Z_{s_n}) \text{ is normal for all } s_1 \cdots, s_n \le t.$$

# Proof Continued

* Since $E[\tilde{X}Z_{s_j}]=0$,

$$\tilde{X} \text{ and } Z_{s_j} \text{ are uncorrelated for } 1 \le j \le n.$$

* Thus,

$$\tilde{X} \text{ and } (Z_{s_j},\cdots,Z_{s_n}) \text{ are independent.}$$

* But then,

$$E[\chi_G(X-\breve{X})]=E[\chi_G\tilde{X}]=E[\chi_G]\cdot E[\tilde{X}]=0 \text{ for all } G\in\mathcal{G}.$$

* Since $\breve{X}$ is $\mathcal{G}$-measurable, we conclude:

$$\breve{X}=E[X\mid\mathcal{G}].$$

# Lemma 6.2.3

$$M_t = \begin{bmatrix} X_t \\ Z_t \end{bmatrix} \in \Re^2 \quad \text{is a Gaussian Process.}$$

* Proof: We may regard $M_t$ as the solution of a 2-dimensional linear stochastic differential equation of the form

$$dM_t = H(t)M_t dt + K(t)dB_t; M_0 = \begin{bmatrix} X_0 \\ 0 \end{bmatrix}; \text{(6.2.9)}$$

where $H(t) \in \Re^{2\times 2}; K(t) \in \Re^{2\times 2}$ and $B_t$ is a 2-dimensional Brownian motion.

# Proof: Lemma 6.2.3

* Put

$$M_t^{(n+1)} = M_0 + \int_0^t H(s) M_s^{(n)} ds + \int_0^t K(s) dB_s;\ n = 0,1,2,\cdots (6.2.10)$$

* Then, $M_t^{(n)}$ is Gaussian for all $n$ and $M_t^{(n)} \to M_t$ in $L^2(P)$.