Martingale problems and filtering

- Martingale problems for conditional distributions
- Filtering equations
- References
1. Martingale problems for conditional distributions

- Specifying Markov processes
- Examples of generators
- Characterization of Markov processes as solutions of martingale problems
- Conditional distributions for martingale problems
- Partially observed processes
- Filtered martingale problem
- Markov mapping theorem
- Burke’s theorem

Kurtz and Ocone (1988); Kurtz and Nappo (2011)
How to specify a Markov model

An $E$-valued process is *Markov* wrt $\{\mathcal{F}_t\}$ if

$$E[f(X(t + s))|\mathcal{F}_t] = E[f(X(t + s))|X(t)], \quad f \in B(E)$$

Ordinary differential equations: $\dot{X} = F(X)$

$$X(t + \Delta t) \approx X(t) + F(X(t))\Delta t$$

Stochastic differential equations:

$$X(t + \Delta t) \approx X(t) + F(X(t))\Delta t + G(X(t))\Delta W$$
Infinitesimal specification

Deterministic (ode) case:

\[ f(X(t + \Delta t)) \approx f(X(t)) + F(X(t)) \cdot \nabla f(X(t)) \Delta t \]

\[
\begin{align*}
    f(X(t + r)) - f(X(t)) &= \sum f(X(t_{i+1})) - f(X(t_i)) \\
    &\approx \sum F(X(t_i)) \cdot \nabla f(X(t_i))(t_{i+1} - t_i)
\end{align*}
\]

which suggests

\[
    f(X(t + r)) - f(X(t)) - \int_{t}^{t+r} F(X(s)) \cdot \nabla f(X(s)) ds = 0
\]
**Martingale properties**

“Infinitesimal changes of distribution”

\[
E[f(X(t + \Delta t)) | \mathcal{F}_t] \approx f(X(t)) + Af(X(t))\Delta t
\]

or

\[
E[f(X(t + \Delta t)) - f(X(t)) - Af(X(t))\Delta t | \mathcal{F}_t] \approx 0
\]

which suggests

\[
E[f(X(t + r)) - f(X(t)) - \int_t^{t+r} Af(X(s))ds | \mathcal{F}_t] = 0
\]

\[
f(X(t)) - f(X(0)) - \int_0^t Af(X(s))ds \quad \text{a martingale}
\]
Examples of generators: Jump processes

Poisson process \((E = \{0, 1, 2 \ldots\}, \mathcal{D}(A) = B(E))\)

\[
Af(k) = \lambda(f(k + 1) - f(k))
\]

Markov chain \((E \text{ discrete}, \mathcal{D}(A) = \{f \in B(E) : f \text{ has finite support}\})\)

\[
Af(k) = \sum_l q_{k,l}(f(l) - f(k))
\]

Pure jump process \((E \text{ arbitrary})\)

\[
Af(x) = \lambda(x) \int_E (f(y) - f(x)) \mu(x, dy)
\]
Examples of generators: Continuous processes

Standard Brownian motion \((E = \mathbb{R}^d)\)

\[
Af = \frac{1}{2} \Delta f, \quad \mathcal{D}(A) = C^2_c(\mathbb{R}^d)
\]

Diffusion process \((E = \mathbb{R}^d, \mathcal{D}(A) = C^2_c(\mathbb{R}^d))\)

\[
Af(x) = \frac{1}{2} \sum_{i,j} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} f(x) + \sum_i b_i(x) \frac{\partial}{\partial x_i} f(x)
\]

Reflecting diffusion \((E \subset \mathbb{R}^d)\)

\[
\mathcal{D}(A) = \{ f \in C^2_c(\overline{E}) : \eta(x) \cdot \nabla f(x) = 0, x \in \partial E \}
\]

\[
Af(x) = \frac{1}{2} \sum_{i,j} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} f(x) + \sum_i b_i(x) \frac{\partial}{\partial x_i} f(x)
\]
The martingale problem for $A$

$X$ is a solution for the martingale problem for $(A, \nu_0)$, $\nu_0 \in \mathcal{P}(E)$, if $P_X(0)^{-1} = \nu_0$ and

$$f(X(t)) - f(X(0)) - \int_0^t Af(X(s))ds$$

is an $\{\mathcal{F}_t^X\}$-martingale for all $f \in \mathcal{D}(A)$.

**Theorem 1.1** If any two solutions of the martingale problem for $A$ satisfying $P_X(0)^{-1} = P_X(0)^{-1}$ also satisfy $P_X(t)^{-1} = P_X(t)^{-1}$ for all $t \geq 0$, then the f.d.d. of a solution $X$ are uniquely determined by $P_X(0)^{-1}$

If $X$ is a solution of the MGP for $A$ and $Y_a(t) = X(a + t)$, then $Y_a$ is a solution of the MGP for $A$.

**Theorem 1.2** If the conclusion of the above theorem holds, then any solution of the martingale problem for $A$ is a Markov process.
A martingale lemma

Let $\{F_t\}$ and $\{G_t\}$ be filtrations with $G_t \subset F_t$.

**Lemma 1.3** Suppose that

$$M(t) = U(t) - U(0) - \int_0^t V(s) ds$$

is an $\{F_t\}$-martingale. Then

$$E[U(t)|G_t] - E[U(0)|G_0] - \int_0^t E[V(s)|G_s] ds$$

is a $\{G_t\}$-martingale.

**Proof.** The lemma follows by the definition and properties of conditional expectations. □
Martingale properties of conditional distributions

Corollary 1.4 If $X$ is a solution of the martingale problem for $A$ with respect to the filtration $\{\mathcal{F}_t\}$ and $\pi_t$ is the conditional distribution of $X(t)$ given $\mathcal{G}_t \subset \mathcal{F}_t$, then

$$\pi_t f - \pi_0 f - \int_0^t \pi_s A f \, ds$$

(1.1)

is a $\{\mathcal{G}_t\}$-martingale for each $f \in D(A)$. 
Technical conditions

**Condition 1.5**

1. \( A : \mathcal{D}(A) \subset C_b(E) \times C(E) \) with \( 1 \in \mathcal{D}(A) \) and \( A1 = 0 \).

2. \( \mathcal{D}(A) \) is closed under multiplication and separates points.

3. There exist \( \psi \in C(E), \psi \geq 1 \), and constants \( a_f \) such that \( f \in \mathcal{D}(A) \) implies \( |Af(x)| \leq a_f \psi(x) \).

4. Defining \( A_0 = \{(f, \psi^{-1}Af) : f \in \mathcal{D}(A)\} \), \( A_0 \) is separable in the sense that there exists a countable collection \( \{g_k\} \subset \mathcal{D}(A) \) such that every solution of the martingale problem for \( A_0^r = \{(g_k, A_0g_k) : k = 1, 2, \ldots\} \) is a solution for \( A_0 \).

5. \( A_0 \) is a pre-generator, that is, \( A_0 \) is dissipative and there are sequences of functions \( \mu_n : E \to \mathcal{P}(E) \) and \( \lambda_n : E \to [0, \infty) \) such that for each \( (f, g) \in A \) for each \( x \in E \)

\[
g(x) = \lim_{n \to \infty} \lambda_n(x) \int_E (f(y) - f(x))\mu_n(x, dy). \tag{1.2}
\]
Forward equation

A $\mathcal{P}(E)$-valued function $\{\nu_t, t \geq 0\}$ is a solution of the forward equation for $A$ if for each $t > 0$, $\int_0^t \nu_s \psi ds < \infty$ (see Condition 1.5) and for each $f \in \mathcal{D}(A)$,

$$\nu_t f = \nu_0 f + \int_0^t \nu_s A f ds. \quad (1.3)$$

Note that if $\pi$ satisfies (1.1), then $\nu_t = E[\pi_t]$ satisfies (1.3).

**Theorem 1.6** Under Condition 1.5, every solution of the forward equation corresponds to a solution of the martingale problem.
Martingale characterization of conditional distributions

**Theorem 1.7** Suppose that \( \{ \tilde{\pi}_t, t \geq 0 \} \) is a cadlag, \( \mathcal{P}(E) \)-valued process with no fixed points of discontinuity adapted to \( \{ \tilde{G}_t \} \) satisfying

\[
E \left[ \int_0^t \tilde{\pi}_s \psi ds \right] < \infty, \quad t > 0
\]

and that

\[
\tilde{\pi}_t f - \tilde{\pi}_0 f - \int_0^t \tilde{\pi}_s A f ds
\]

is a \( \{ \tilde{G}_t \} \)-martingale for each \( f \in \mathcal{D}(A) \). Then there exists a solution \( X \) of the martingale problem for \( A \), a \( \mathcal{P}(E) \)-valued process \( \{ \pi_t, t \geq 0 \} \) with the same distribution as \( \{ \tilde{\pi}_t, t \geq 0 \} \), and a filtration \( \{ G_t \} \) such that \( \pi_t \) is the conditional distribution of \( X(t) \) given \( G_t \).
Conditioning on a process

**Theorem 1.8** If \( \{ \tilde{G}_t \} \) in Theorem 1.7 is generated by a cadlag process \( \tilde{Y} \) with no fixed points of discontinuity and \( \tilde{\pi}(0) \), that is,

\[
\tilde{G}_t = \mathcal{F}_t^{\tilde{Y}} \vee \sigma(\tilde{\pi}(0)),
\]

then there exists a solution \( X \) of the martingale problem for \( A \), a \( \mathcal{P}(E) \)-valued process \( \{ \pi_t, t \geq 0 \} \), and a process \( Y \) such that \( \{ \pi_t, t \geq 0 \} \) and \( Y \) have the same joint distribution as \( \{ \tilde{\pi}_t, t \geq 0 \} \) and \( \tilde{Y} \) and \( \pi_t \) is the conditional distribution of \( X(t) \) given \( \mathcal{F}_t^Y \vee \sigma(\pi(0)) \).
Idea of proof

Enlarge the state space so that the current state of the process contains all information about the past of the observation $\tilde{Y}$.

Let $\{b_k\}, \{c_k\} \subset C_b(E_0)$ satisfy $0 \leq b_k, c_k \leq 1$, and suppose that the spans of $\{b_k\}$ and $\{c_k\}$ are bounded, pointwise dense in $B(E_0)$.

Let $a_1, a_2, \ldots$ be an ordering of the rationals with $a_i \geq 1$ and

$$
\tilde{V}_{ki}(t) = c_k(\tilde{Y}(0)) - a_i \int_0^t \tilde{V}_{ki}(s)ds + \int_0^t b_k(\tilde{Y}(s))ds \quad (1.4)
$$

$$
= c_k(\tilde{Y}(0))e^{-a_i t} + \int_0^t e^{-a_i(t-s)}b_k(\tilde{Y}(s))ds.
$$
Set $\tilde{V}(t) = (\tilde{V}_{ki}(t) : k, i \geq 1) \in [0, 1]^{\infty}$, 

$$\mathcal{D}(\hat{A}) = \{ f(x) \prod_{k,i=1}^{m} g_{ki}(v_{ki}) : f \in \mathcal{D}(A), g_{ki} \in C^1[0, 1], m = 1, 2, \ldots \}$$

and

$$\hat{A}(fg)(x, v, u) = g(v)Af(x) + f(x) \sum (-a_i v + b_k(u)) \partial_{ki} g(v),$$

For $fg \in \mathcal{D}(\hat{A})$,

$$\tilde{\pi}_t fg(\tilde{V}(t)) - \tilde{\pi}_0 fg(\tilde{V}(0))$$

$$- \int_0^t \left( g(\tilde{V}(s))\tilde{\pi}_s Af + \tilde{\pi}_s f \sum (-a_i \tilde{V}_{ki}(s) + b_k(\tilde{Y}(s))) \partial_{ki} g(\tilde{V}(s)) \right) ds$$

is a $\{\mathcal{F}_t^{\tilde{Y}}\}$-martingale and $\nu_t$ defined by

$$\nu_t(fgh) = E[\tilde{\pi}_t fg(\tilde{V}(t))h(\tilde{Y}(t))]$$

is a solution of the controlled forward equation for $\hat{A}$. 
Partially observed processes

Let $\gamma : E \to E_0$ be Borel measurable.

**Corollary 1.9** If in Corollary 1.8, $\tilde{Y}$ and $\tilde{\pi}$ satisfy

$$\int_E h \circ \gamma(x) \tilde{\pi}_t(dx) = h(\tilde{Y}(t)) \quad a.s. $$

for all $h \in B(E_0)$ and $t \geq 0$, then $Y(t) = \gamma(X(t))$.

The filtered martingale problem

**Definition 1.10** A $\mathcal{P}(E)$-valued process $\tilde{\pi}$ and an $E$-valued process $\tilde{Y}$ are a solution of the filtered martingale problem for $(A, \gamma)$ if

$$\tilde{\pi}_t f - \tilde{\pi}_0 f - \int_0^t \tilde{\pi}_s Af ds$$

is a $\{\mathcal{F}_t^{\tilde{Y}} \vee \sigma(\tilde{\pi}(0))\}$-martingale for each $f \in \mathcal{D}(A)$ and $\int_E h \circ \gamma(x) \tilde{\pi}_t(dx) = h(\tilde{Y}(t))$ a.s. for all $h \in B(E_0)$ and $t \geq 0$.

**Theorem 1.11** Let $\varphi_0 \in \mathcal{P}(\mathcal{P}(E))$ and define $\mu_0 = \int_{\mathcal{P}(E)} \mu \varphi_0(d\mu)$. If uniqueness holds for the martingale problem $(A, \mu_0)$, then uniqueness holds for the filtered martingale problem for $(A, \gamma, \varphi_0)$. If uniqueness holds for the filtered martingale problem for $(A, \gamma, \varphi_0)$, then $\{\pi_t, t \geq 0\}$ is a Markov process.

Note: $\tilde{\pi}_t = H(t, \tilde{Y}, \tilde{\pi}_0)$ implies $\pi_t = H(t, Y, \pi_0)$. 
Markov mappings

Theorem 1.12 $\gamma : E \to E_0$, Borel measurable.

$\alpha$ a transition function from $E_0$ into $E$ satisfying

$$\alpha(y, \gamma^{-1}(y)) = 1$$

Let $\mu_0 \in \mathcal{P}(E_0)$, $\nu_0 = \int \alpha(y, \cdot)\mu_0(\,dy)$, and define

$$C = \{(\int_E f(z)\alpha(\cdot, dz), \int_E Af(z)\alpha(\cdot, dz)) : f \in \mathcal{D}(A)\}.$$

If $\tilde{Y}$ is a solution of the MGP for $(C, \mu_0)$, then there exists a solution $Z$ of the MGP for $(A, \nu_0)$ such that $Y = \gamma \circ Z$ and $\tilde{Y}$ have the same distribution on $M_{E_0}[0, \infty)$.

$$E[f(Z(t))|\mathcal{F}_t^Y] = \int f(z)\alpha(Y(t), dz)$$

(at least for almost every $t$, all $t$ if $Y$ has no fixed points of discontinuity).
Uniqueness

**Corollary 1.13** If uniqueness holds for the MGP for \((A, \nu_0)\), then uniqueness holds for the \(M_{E_0}[0, \infty)\)-MGP for \((C, \mu_0)\). If \(\tilde{Y}\) has sample paths in \(D_{E_0}[0, \infty)\), then uniqueness holds for the \(D_{E_0}[0, \infty)\)-martingale problem for \((C, \mu_0)\).

Existence for \((C, \mu_0)\) and uniqueness for \((A, \nu_0)\) implies existence for \((A, \nu_0)\) and uniqueness for \((C, \mu_0)\), and hence that \(\tilde{Y}\) is Markov.
Intertwining condition

Let $\alpha(y, \Gamma)$ be a transition function from $E_0$ to $E$ satisfying

$$\alpha(y, \gamma^{-1}(y)) = 1,$$

and define $S(t) : B(E_0) \to B(E_0)$ by

$$S(t)g(y) = \alpha T(t)g \circ \gamma(y) \equiv \int_E T(t)g \circ \gamma(x)\alpha(y, dx).$$

Theorem 1.14 (Rogers and Pitman (1981), cf Rosenblatt (1966)) If for each $t \geq 0$,

$$\alpha T(t)f = S(t)\alpha f, \quad f \in B(E), \quad (S(t) \text{ is a semigroup})$$

and $X$ is a Markov process with initial distribution $\alpha(y, \cdot)$ and semigroup $\{T(t)\}$, then $Y$ is a Markov process with $Y(0) = y$ and

$$P\{X(t) \in \Gamma | \mathcal{F}^Y_t\} = \alpha(Y(t), \Gamma).$$
Generator for $Y$

Note that

$$\alpha T(t)f = S(t)\alpha f, \quad f \in B(E),$$

suggests that the generator for $Y$ is given by

$$C\alpha f = \alpha A f.$$
Burke’s output theorem \textcopyright{} Kliemann, Koch, and Marchetti (1990)

\[ X = (Q, D), \text{ an } M/M/1 \text{ queue and its departure process} \]

\[ Af(k, l) = \lambda (f(k + 1, l) - f(k, l)) + \mu \mathbf{1}_{\{k>0\}}(f(k - 1, l + 1) - f(k, l)) \]

\[ \gamma(k, l) = l \]

Assume \( \lambda < \mu \) and define

\[ \alpha(l, \{(k, l)\}) = (1 - \frac{\lambda}{\mu})(\frac{\lambda}{\mu})^{k-1}, \quad k = 0, 1, 2, \ldots \quad \alpha(l, \{(k, m)\}) = 0, \ m \neq l \]

Then

\[ \alpha Af(l) = \mu \sum_{k=1}^{\infty} (1 - \frac{\lambda}{\mu})\left(\frac{\lambda}{\mu}\right)^{k-1}(f(k - 1, l + 1) - f(k - 1, l)) \]

\[ = \lambda(\alpha f(l + 1) - \alpha f(l)) \]
Poisson output

Therefore, there exists a solution \((Q, D)\) of the martingale problem for \(A\) such that \(D\) is a Poisson process with parameter \(\lambda\) and

\[
P\{Q(t) = k|\mathcal{F}_t^D\} = (1 - \frac{\lambda}{\mu})(\frac{\lambda}{\mu})^{k-1},
\]

that is, \(Q(t)\) is independent of \(\mathcal{F}_t^D\) and is geometrically distributed.
Pitman’s theorem

$Z$ standard Brownian motion

$M(t) = \sup_{s \leq t} Z(s), \quad V(t) = M(t) - Z(t)$

$X(t) = (Z(t), M(t) - Z(t)) = (Z(t), V(t))$

$Y(t) = 2M(t) - Z(t) = \gamma (X(t) = 2V(t) + Z(t))$

$$Af(z, v) = \frac{1}{2}f_{zz}(z, v) - f_{zv}(z, v) + \frac{1}{2}f_{vv}(z, v) \quad \text{b.c.} \quad f_v(z, 0) = 0$$

$$F(y) = \alpha f(y) = \frac{1}{y} \int_0^y f(y - 2v, v) dv$$

$$\alpha Af(y) = \frac{1}{2}F''(y) + \frac{1}{y}F'(y)$$
2. Filtering equations

- Bayes and Kallianpur-Striebel formulas
- “Particle representations” of conditional distributions
- Continuous time filtering in Gaussian white noise
- Derivation of filtering equations
- Zakai equation
- Kushner-Stratonovich equation
- Point process observations
- Spatial observations with additive white noise
- Cluster detection
- Signal in noise with point process observations
- Exit time observations
- Uniqueness

Kurtz and Xiong (1999, 2001)
Bayes and Kallianpur Striebel formulas

Kallianpur and Striebel (1968)

Let $X$ and $Y$ be random variables defined on $(\Omega, \mathcal{F}, P)$. Suppose we want to calculate the conditional expectation

$$E^P[f(X)|Y].$$

If $P << Q$ with $dP = L dQ$, Bayes formula says

$$E^P[f(X)|Y] = \frac{E^Q[f(X) L|Y]}{E^Q[L|Y]}.$$

If $X = h(U, Y)$, $L = L(U, Y)$ and $U$ and $Y$ are independent under $Q$, then

$$E^P[f(X)|Y] = \frac{\int f(h(u, Y)) L(u, Y) \mu_U(du)}{\int L(u, Y) \mu_U(du)}$$

where $\mu_U$ is the distribution of $U$.

**Method:** Find a reference measure under which what we don’t know is independent of what we do know.
Monte Carlo integration

Let $U_1, U_2, \ldots$ be iid with distribution $\mu_U$. Then

$$E^P[f(X)|Y] = \lim_{n \to \infty} \frac{\sum_{i=1}^{n} f(h(U_i, Y))L(U_i, Y)}{\sum_{i=1}^{n} L(U_i, Y)}$$

Note that $(U_1, Y), (U_2, Y), \ldots$ is a stationary (in fact exchangeable) sequence. Consequently,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} f(h(U_i, Y))L(U_i, Y) = E^Q[f(U_1, Y)L(U_i, Y)|\mathcal{I}]$$

$$= E^Q[f(U_1, Y)L(U_i, Y)|Y]$$

$$= \int f(h(u, Y))L(u, Y)\mu_U(dx) \quad a.s. \ Q$$
Continuous time filtering in Gaussian white noise

To understand the intuition behind the standard “observation in Gaussian white noise” filtering model, suppose $X$ is the signal of interest and noisy observations of the form

$$O_n\left(\frac{k}{n}\right) = h(X(\frac{k}{n})) \frac{1}{n} + \frac{1}{\sqrt{n}} \zeta_k$$

are made every $n^{-1}$ time units. For large $n$, the noise swamps the signal at any one time point.

Assume that the $\{\zeta_k\}$ are iid with mean zero and variance $\sigma^2$. Then $Y_n(t) = \sum_{k=1}^{[nt]} O_n\left(\frac{k}{n}\right)$ is approximately

$$Y(t) = \int_0^t h(X(s))ds + \sigma W(t).$$  \hspace{1cm} (2.1)

Note, however, that $E[f(X(t))|\mathcal{F}_t^{Y_n}]$ does not necessarily converge to $E[f(X(t))|\mathcal{F}_t^{Y}]$; however, we still take (2.1) as our observation model.
Reference measure

By the Girsanov formula,

$$E^P[g(X(t))|\mathcal{F}^Y_t] = \frac{E^Q[g(X(t))L(t)|\mathcal{F}^Y_t]}{E^Q[L(t)|\mathcal{F}^Y_t]}$$

where under $Q$, $X$ and $Y$ are independent, $Y$ is a Brownian motion with mean zero and variance $\sigma^2 t$, and

$$L(t) = \exp\left\{ \int_0^t \frac{h(X(s))}{\sigma^2} dY(s) - \frac{1}{2} \int_0^t \frac{h^2(X(s))}{\sigma^2} ds \right\}$$

that is,

$$L(t) = 1 + \int_0^t \frac{h(X(s))}{\sigma^2} L(s) dY(s).$$

Under $Q$, $\sigma^{-1}Y$ is a standard Brownian motion, so under $P$,

$$W(t) = \sigma^{-1}Y(t) - \int_0^t \sigma^{-1}h(X(s)) ds$$

is a standard Brownian motion.
Monte Carlo solution

Pardoux (1991)

Let $X_1, X_2, \ldots$ be iid copies of $X$ that are independent of $Y$ under $Q$, and let

$$L_i(t) = 1 + \int_0^t \frac{h(X_i(s))}{\sigma^2} L_i(s) dY(s).$$

Note that

$$\phi(g, t) \equiv E^Q[g(X(s))L(s) | \mathcal{F}^Y_s] = E^Q[g(X_i(s))L_i(s) | \mathcal{F}^Y_s]$$

Claim:

$$\frac{1}{n} \sum_{i=1}^n g(X_i(s))L_i(s) \rightarrow E^Q[g(X(s))L(s) | \mathcal{F}^Y_s]$$
Zakai equation

For simplicity, assume $\sigma = 1$ and $X$ is a diffusion

$$X(t) = X(0) + \int_0^t \sigma(X(s))dB(s) + \int_0^t b(X(s))ds,$$

where under $Q$, $B$ and $Y$ are independent standard Brownian motions. Since

$$g(X(t)) = g(X(0)) + \int_0^t g'(X(s))\sigma(X(s))dB(s) + \int_0^t Ag(X(s))ds$$

$$g(X(t))L(t) = g(X(0)) + \int_0^t L(s)dg(X(s)) + \int_0^t g(X(s))dL(s)$$

$$= g(X(0)) + \int_0^t L(s)g'(X(s))\sigma(X(s))dB(s)$$

$$+ \int_0^t L(s)Ag(X(s))ds + \int_0^t g(X(s))h(X(s))L(s)dY(s)$$
Monte Carlo derivation of Zakai equation

\[ X_i(t) = X_i(0) + \int_0^t \sigma(X_i(s))dB_i(s) + \int_0^t b(X_i(s))ds, \]

where \((X_i(0), B_i)\) are iid copies of \((X(0), B)\).

\[ g(X_i(t))L_i(t) = g(X_i(0)) + \int_0^t L_i(s)g'(X_i(s))\sigma(X_i(s))dB_i(s) \]
\[ + \int_0^t L_i(s)Ag(X_i(s))ds \]
\[ + \int_0^t g(X_i(s))h(X_i(s))L_i(s)dY(s) \]

and hence

\[ \phi(g, t) = \phi(g, 0) + \int_0^t \phi(Ag, s)ds + \int_0^t \phi(gh, s)dY(s) \]
Kushner-Stratonovich equation

\[
\pi_t g = E^P[g(X(t))|\mathcal{F}_t^Y] = \frac{\phi(g,t)}{\phi(1,t)}
\]

\[
= \frac{\phi(g,0)}{\phi(1,0)} + \int_0^t \frac{1}{\phi(1,s)} d\phi(g, s) - \int_0^t \frac{\phi(g,s)}{\phi(1,s)^2} d\phi(1,s)
+ \int_0^t \frac{\phi(g,s)}{\phi(1,s)^3} d[\phi(1,\cdot)]_s - \int_0^t \frac{1}{\phi(1,s)^2} d[\phi(g,\cdot), \phi(1,\cdot)]_s
\]

\[
= \pi_0 g + \int_0^t \pi_s Ag ds + \int_0^t (\pi_s gh - \pi_s g \pi_s h) dY(s)
+ \int_0^t \sigma^2 \pi_s g \pi_s h^2 ds - \int_0^t \sigma^2 \pi_s gh \pi_s h ds
\]

\[
= \pi_0 g + \int_0^t \pi_s Ag ds + \int_0^t (\pi_s gh - \pi_s g \pi_s h)(dY(s) - \pi_s h ds)
\]

Note: \(\pi_t\) in Theorem 1.11 is \(\pi_t \times \delta_{Y(t)}\) in the current notation.
Spatial observations with additive white noise

Signal:

\[
X(t) = X(0) + \int_0^t \sigma(X(s))dB(s) + \int_0^t b(X(s))ds + \int_{S_0 \times [0,t]} \alpha(X(s), u)W(du \times ds)
\]

\[
= X(0) + \int_0^t \sigma(X(s))dB(s) + \int_{S_0 \times [0,t]} \alpha(X(s), u)Y(du \times ds)
\]

\[
+ \int_0^t (b(X(s))) - \int_{S_0} \alpha(X(s), u)h(X(s), u)\mu_0(du))ds
\]

Observation:

\[
Y(A, t) = \int_0^t \int_A h(X(s), u)\mu_0(du)ds + W(A, t)
\]

Under \(Q\), \(Y\) is Gaussian white noise on \(S_0 \times [0, \infty)\) with

\[
E[Y(A, t)Y(B, s)] = \mu_0(A \cap B)t \wedge s,
\]

and \(dP|_{\mathcal{F}_t} = L(t)dQ|_{\mathcal{F}_t}\) where

\[
L(t) = 1 + \int_{S_0 \times [0,t]} L(s)h(X(s), u)Y(du \times ds)
\]
Apply Itô’s formula

Under $P$, $X$ is a diffusion with generator

$$Af(x) = \frac{1}{2} \sum a_{ij}(x) \partial_i \partial_j f(x) + \sum b_i(x) \partial_i f(x)$$

where

$$a(x) = \sigma(x) \sigma(x)^T + \int_{S_0} \alpha(x,u) \alpha(x,u)^T \mu_0(du)$$

Then

$$f(X(t)) L(t)$$

$$= f(X(0)) + \int_0^t L(s) \nabla f(X(s))^T \sigma(X(s)) dB(s)$$

$$+ \int_{S_0 \times [0,t]} L(s)(\nabla f(X(s)) \cdot \alpha(X(s),u)) Y(du \times ds)$$
Particle representation

$B_i$ independent, standard Brownian motions, independent of $Y$ on $(\Omega, \mathcal{F}, P_0)$. Let

$$X_i(t) = X_i(0) + \int_0^t \sigma(X_i(s)) dB_i(s) + \int_{S_0 \times [0,t]} \alpha(X_i(s-), u) Y(du \times ds)$$

$$+ \int_0^t \int_{S_0} (b(X_i(s)) - \alpha(X_i(s), u) h(X_i(s), u)) \mu_0(du) ds$$

$$L_i(t) = 1 + \int_{S_0 \times [0,t]} L_i(s) h(X_i(s), u) Y(du \times ds)$$

Then

$$\phi(f, t) = E^{P_0}[f(X(t))L(t)|\mathcal{F}_t^Y] = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^n f(X_i(t))L_i(t)$$
Zakai equation

Since

\[
\begin{align*}
f(X_i(t))L_i(t) &= f(X_i(0)) + \int_0^t L_i(s)\nabla f(X_i(s))^T \sigma(X_i(s))dB_i(s) \\
&\quad + \int_{S_0 \times [0,t]} L_i(s)(\nabla f(X_i(s)) \cdot \alpha(X_i(s), u)) \\
&\quad + \int_0^t L_i(s)Af(X_i(s))ds + f(X_i(s))h(X_i(s), u))Y(du \times ds),
\end{align*}
\]

\[
\phi(f, t) = \phi(f, 0) + \int_0^t \phi(Af, s)ds \\
&\quad + \int_{S_0 \times [0,t]} \phi(\nabla f \cdot \alpha(\cdot, u) + fh(\cdot, u), s)Y(du \times ds);
\]
Kushner-Stratonovich equation

It follows that

\[
\pi_t f = \frac{\phi(f, t)}{\phi(1, t)}
\]

\[
= \pi_0 f + \int_0^t \pi_s Af \, ds
\]

\[
+ \int_{S_0 \times [0, t]} \left( \pi_s (\nabla f \cdot \alpha(\cdot, u) + fh(\cdot, u)) - \pi_s f \pi_s h(\cdot, u) \right) Y(du \times ds)
\]

\[
+ \int_0^t \int_{S_0} \left( \pi_s f \pi_s h(\cdot, u) - \pi_s (\nabla f \cdot \alpha(\cdot, u) + fh(\cdot, u)) \right) \pi_s h(\cdot, u) \mu_0(du) \, ds
\]

\[
= \pi_0 f + \int_0^t \pi_s Af \, ds
\]

\[
+ \int_{S_0 \times [0, t]} \left( \pi_s (\nabla f \cdot \alpha(\cdot, u) + fh(\cdot, u)) - \pi_s f \pi_s h(\cdot, u) \right) \tilde{Y}(du \times ds)
\]

where

\[
\tilde{Y}(A, t) = Y(A, t) - \int_0^t \int_A \pi_s h(\cdot, u) \mu_0(du) \, ds
\]
Cluster detection: Possible applications

- Internet packets that form a malicious attack on a computer system.
- Financial transactions that form a collusive trading scheme.
- Earthquakes that form a single seismic event.
The model

The observations form a marked point process $O$ with marks in $E$.

$$O(A, t) = N(A, t) + C(A, t)$$

with

$$N(A, t) = \int_{A \times [0, \infty) \times [0, t]} 1_{[0, \gamma(u)]}(v) \xi_1(du \times dv \times ds)$$

$$C(A, t) = \int_{A \times [0, \infty) \times [0, t]} 1_{[0, \lambda(u, \eta_{s-})]}(v) \xi_2(du \times dv \times ds)$$

where $\xi_1$ and $\xi_2$ are independent Poisson random measures on $E \times [0, \infty) \times [0, \infty)$ with mean measure $\nu \times \ell \times \ell$, $\ell$ denoting Lebesgue measure.

$$\eta_t(A \times [0, r]) = \int_{A \times [0, t]} 1_A(u) 1_{[0, r]}(s) C(du \times ds)$$
Radon-Nikodym derivative

**Lemma 2.1** On \((\Omega, \mathcal{F}, Q)\), let \(N\) and \(C\) be independent Poisson random measures with mean measures \(\nu_0(du \times ds) = \gamma(u)\nu(du)ds\) and \(\nu_1(du \times ds) = \lambda(u)\nu(du)ds\) respectively that are compatible with \(\{\mathcal{F}_t\}\). Let \(L\) satisfy

\[
L(t) = 1 + \int_{E \times [0,t]} \left( \frac{\lambda(u, \eta_{s-})}{\lambda(u)} - 1 \right) L(s-)(C(du \times ds) - \lambda(u)\nu(du)ds). \tag{2.2}
\]

and assume that \(L\) is a \(\{\mathcal{F}_t\}\)-martingale.

Define \(dP|_{\mathcal{F}_t} = L(t)dQ|_{\mathcal{F}_t}\). Under \(P\), for all \(A\) such that \(\int_0^t \int_A \lambda(u, \eta_s)\nu(du)ds < \infty, \ t > 0\),

\[
C(A, t) - \int_{A \times [0,t]} \lambda(u, \eta_s)\nu(du)ds
\]

is a local martingale and \(N\) is independent of \(C\) and is a Poisson random measure with mean measure \(\nu_0\).
The general filtering equations

Theorem 2.2

\[ \phi(f, t) = \phi(f, 0) - \int_{E \times [0, t]} \phi(f(\cdot))(\lambda(u, \cdot) - \lambda(u)), s)\nu(du)ds \]

\[ + \int_{E \times [0, t]} \phi(f(\cdot + \delta(u, s))\frac{\lambda(u, \cdot)}{\lambda(u)} - f(\cdot), s-)\frac{\lambda(u)}{\lambda(u) + \gamma(u)}O(du \times ds) \]

and

\[ \pi_t f = \pi_0 f \]

\[ + \int_{E \times [0, t]} \frac{\pi_{s-}(f(\cdot + \delta(u, s))\lambda(u, \cdot)) - \pi_{s-} \lambda(u, \cdot)\pi_{s-}f}{\pi_{s-} \lambda(u, \cdot) + \gamma(u)}O(du \times ds) \]

\[ - \int_{E \times [0, t]} (\pi_s(f(\cdot)\lambda(u, \cdot)) - \pi_s f \pi_s \lambda(u, \cdot))\nu(du)ds \]
Simplify

**Problem:** The difficulty of computing the distribution: $2^{O(E,t)}$ possible states.

Need to compromise: compute $\pi_t f = E^P[f(\eta_s)|\mathcal{F}_s]$ for a “small” collection of $f$

Suppose one observes $u_i$ at time $\tau_i$ and $y_i = (u_i, \tau_i)$.

$\theta(y_i)(\cdot) = 1_{\{y_i \text{ is a point in the cluster}\}}$

$\theta_0(y_i)(\cdot) = 1_{\{y_i \text{ is the latest point in the cluster}\}}$

Need to be able to evaluate

$$\pi_t \lambda(u, \cdot)$$
A Markov scenario

Consider

$$\lambda(u, \eta_t) = \sum_{i=1}^{O(E,t)} \lambda(u, y_i) \theta_0(y_i) + \epsilon(u),$$

where $\theta_0(y_i) = 1_{\{y_i \text{ is the latest point in the cluster}\}}$.

Get a closed system for $\pi_t \theta_0(y_i)$

Let $\theta(y)(\cdot) = 1_{\{y \text{ is a point in the cluster}\}}$.

Get a closed system for

$$\pi_t \theta_0(y_i), \quad \pi_t \theta(y_i), \quad \pi_t \theta(y_i) \theta_0(y_j)$$
Exit time observations

Let $X^\tau$ be a diffusion process in a domain $D$ stopped at $\tau = \inf\{t : X^\tau(t) \notin D^o\}$. Let $Y_0$ satisfy

$$Y_0(t) = \int_0^t h(X^\tau(s))ds + W_0(t),$$

where $h(x) = 0$ for $x \in \partial D$. We are interested in the conditional distribution of $X^\tau(t)$ given $\mathcal{F}^{Y_0,\tau}_t = \sigma(Y_0(s), \tau \land s : s \leq t)$.

For $i = 1, 2, \ldots$, let $N_i$ be independent, unit Poisson processes and define

$$Y_i(t) = N_i\left(\int_0^t 1_{\partial D}(X^\tau(s))ds\right)$$

Let

$$\mathcal{F}^n_t = \sigma(Y_0(s), \ldots, Y_n(s) : s \leq t),$$

and $\tau_n = \inf\{t : \lor_{1 \leq i \leq n} Y_i(t) \geq 1\}$. Then $\tau = \lim_{n \to \infty} \tau_n$, and we have

$$\lim_{n \to \infty} E[f(X^\tau(t))|\mathcal{F}^n_t] = E[f(X^\tau(t))|\lor_n \mathcal{F}^n_t] = E[f(X^\tau(t))|\mathcal{F}^{Y_0,\tau}_t].$$

Krylov and Wang (2011)
Reference measure

Under the reference measure $Q_n$, $Y_0$ is a standard Browian motion and the $Y_i$, $i = 1, \ldots, n$ are independent Poisson processes with intensity $n^{-1}$.

Setting $\theta(t) = 1_{\partial D}(X(t\wedge \tau))$, $Y^n = (Y_0, Y_1, \ldots, Y_n)$, $S(Y^n(t)) = \sum_{i=1}^n Y_i(t)$,

\[
L_0(X^\tau, Y_0, t) = \exp\left\{ \int_0^t h(X^\tau(s))dY_0(s) - \frac{1}{2} \int_0^t h^2(X^\tau(s))ds \right\}
\]

\[
L_i(X^\tau, Y_i, t) = \exp\left\{ \int_0^t \log n\theta(s-)dY_i(s) - \int_0^t (\theta(s) - n^{-1})ds \right\}
\]

Then

\[
L_n(X^\tau, Y^n, t) = L_0(X^\tau, Y_0, t) \prod_{i=1}^n L_i(X^\tau, Y_i, t)
\]
SDE for Radon-Nikodym derivative

\[
L_n(X^\tau, Y^n, t) = L_0(X^\tau, Y_0, t) \exp\left\{ \sum_{i=1}^{n} \int_0^t \log n\theta(s-)dY_i(s) - \int_0^t (n\theta(s) - 1)ds \right\}
\]

\[
= L_0(X^\tau, Y_0, t) \mathbf{1}_{\{\tau_n > t\}} \exp\left\{- \int_0^t (n\theta(s) - 1)ds \right\}
+ L_0(X^\tau, Y_0, t) \mathbf{1}_{\{\tau < \tau_n \leq t\}} n^{S(Y^n(t))} \exp\{-n(t - \tau) + t\}ds \}.
\]

So

\[
L_n(X^\tau, Y^n, t) = 1 + \int_0^t L_n(X^\tau, Y^n, s)h(X^\tau(s))^T dY_0(s)
+ \sum_{i=1}^{n} \int_0^t L_n(X^\tau, Y^n, s-)(n\theta(s-) - 1)d(Y_i(s) - n^{-1}s).
\]
\textbf{Zakai equation}

Writing $L_n(t) = L_n(X^\tau, Y^n, t)$ and

\begin{align*}
M_\varphi(t) & = \varphi(X^\tau(t)) - \varphi(X(0)) - \int_0^t A^\tau \varphi(X^\tau(s)) ds \\
L_n(t)\varphi(X^\tau(t)) & = \varphi(X^\tau(0)) + \int_0^t \varphi(X^\tau(s))dL_n(s) + \int_0^t L_n(s)A^\tau \varphi(X^\tau(s)) ds \\
& \quad + \int_0^t L_n(s) dM_\varphi(s) \\
& = \varphi(X^\tau(0)) + \int_0^t \varphi(X^\tau(s))L_n(s) h(X^\tau(s))^T dY_0(s) \\
& \quad + \sum_{i=1}^n \int_0^t \varphi(X^\tau(s))L_n(s-)(n\theta(s-) - 1)d(Y_i(s) - n^{-1}s) \\
& \quad + \int_0^t L_n(s)A^\tau \varphi(X^\tau(s)) ds + \int_0^t L_n(s) dM_\varphi(s).
\end{align*}
Averaging

The Zakai equation becomes

\[
\langle V^n(t), \varphi \rangle = \langle V(0), \varphi \rangle + \int_0^t \langle V^n(s), A^T \varphi \rangle ds + \int_0^t \langle V^n(s), \varphi h^T \rangle dY_0(s) \\
+ \sum_{i=1}^n \int_0^t \langle V^n(s-), (n1_{\partial D} - 1) \varphi \rangle d(Y_i(s) - n^{-1}s).
\]
Kushner-Stratonovich equation

\[ \langle \pi^n(t), \varphi \rangle = \langle \pi(0), \varphi \rangle + \int_0^t \langle \pi^n(s), A^\tau \varphi \rangle ds \]

\[ + \int_0^t \left( \langle \pi^n(s), \varphi h^T \rangle - \langle \pi^n(s), \varphi \rangle \langle \pi^n(s), h^T \rangle \right) \left( dY_0(s) - \langle \pi^n(s), h \rangle ds \right) \]

\[ - \int_0^t \langle \pi^n(s-), (n1_{\partial D} - 1) \varphi \rangle ds + \int_0^t \langle \pi^n(s-), \varphi \rangle \langle \pi^n(s-), (n1_{\partial D} - 1) \rangle ds \]

\[ + \sum_{i=1}^n \int_0^t \left( \frac{\langle \pi^n(s-), 1_{\partial D} \varphi \rangle}{\langle \pi^n(s-), 1_{\partial D} \rangle} - \langle \pi^n(s-), \varphi \rangle \right) dY_i(s) \]

\[ = \langle \pi(0), \varphi \rangle + \int_0^t \langle \pi^n(s), A^\tau \varphi \rangle ds \]

\[ + \int_0^t \left( \langle \pi^n(s), \varphi h^T \rangle - \langle \pi^n(s), \varphi \rangle \langle \pi^n(s), h^T \rangle \right) \left( dY_0(s) - \langle \pi^n(s), h \rangle ds \right) \]

\[ + \sum_{i=1}^n \int_0^t \left( \frac{\langle \pi^n(s-), 1_{\partial D} \varphi \rangle}{\langle \pi^n(s-), 1_{\partial D} \rangle} - \langle \pi^n(s-), \varphi \rangle \right) \left( dY_i(s) - \langle \pi^n(s-), 1_{\partial D} \rangle ds \right) \]
Stopped equation

Stopping at $\tau_n$, we have

$$\langle \pi^n(t \wedge \tau_n), \varphi \rangle = \langle \pi(0), \varphi \rangle + \int_0^{t \wedge \tau_n} \langle \pi^n(s), A^\tau \varphi \rangle ds$$

$$+ \int_0^{t \wedge \tau_n} \left( \langle \pi^n(s), \varphi h^T \rangle - \langle \pi^n(s), \varphi \rangle \langle \pi^n(s), h^T \rangle \right) (dY_0(s) - \langle \pi^n(s), h \rangle ds)$$

$$+ 1_{\{\tau_n \leq t\}} \left( \frac{\langle \pi^n(\tau_n^{-}), 1_{\partial D} \varphi \rangle}{\langle \pi^n(\tau_n^{-}), 1_{\partial D} \rangle} - \langle \pi^n(\tau_n^{-}), \varphi \rangle \right)$$

$$- n \int_0^{t \wedge \tau_n} \left( \frac{\langle \pi^n(s), 1_{\partial D} \varphi \rangle}{\langle \pi^n(s), 1_{\partial D} \rangle} - \langle \pi^n(s), \varphi \rangle \right) \langle \pi^n(s), 1_{\partial D} \rangle ds$$

Note that, assuming $P\{\tau_n < \infty\} = 1$, we must have

$$E[n \int_0^{\tau_n} \langle \pi^n(s), 1_{\partial D} \rangle ds] = 1$$

and for $E[\langle \pi^n(s), 1_{\partial D} \rangle 1_{\{s < \tau_n\}}] = E[\langle \pi^n(s), 1_{\partial D} \rangle 1_{\{\tau \leq s < \tau_n\}}] = O(n^{-1})$
Convergence

Let \( D^o = \{ x : \rho(x) > 0 \} \), \( |\nabla \rho(x)| > 0 \) on \( \partial D \). Let \( k_n \to \infty \) in such a way that, at least along a subsequence, for \( t < \tau \),

\[
\int_0^t k_n^2 \langle \pi^n(s), e^{-k_n \rho} \rangle ds \to \int_0^t \langle \beta(s), \varphi \rangle ds.
\]

Define

\[
\varphi_n(x) = \varphi(x) e^{-k_n \rho(x)}.
\]

Then for \( t < \tau \),

\[
\int_0^t n \langle \pi^n(s), 1_{\partial D} \varphi \rangle ds \to \int_0^t \langle \beta(s), \frac{1}{2} \sum_{i,j} a_{ij} \partial_i \rho \partial_j \rho \varphi \rangle ds,
\]

and we should have

\[
\frac{\langle \pi^n(\tau_n^-), 1_{\partial D} \varphi \rangle}{\langle \pi^n(\tau_n^-), 1_{\partial D} \rangle} \to \frac{\langle \beta(s), \sum_{i,j} a_{ij} \partial_i \rho \partial_j \rho \varphi \rangle}{\langle \beta(s), \sum_{i,j} a_{ij} \partial_i \rho \partial_j \rho \rangle}.
\]
Uniqueness for Kushner-Stratonovich equation

Each of the Kushner-Stratonovich equations is of the form

\[ \pi_t g = \pi_0 g + \int_0^t \pi_s Agds + \int_0^t G(\pi_{s-})d(Y(s) - \pi_s hds) \]

Each of the stochastic integrator would be a martingale if the \( \pi \) was the conditional distribution. Suppose that is not true.

In each case, under some restrictions on \( h \) and \( \pi \), we can do a change of measure to make \( Y(t) - \int_0^t \pi_s hds \) a martingale. Under the “new” measure, \((Y, \pi)\) is a solution of the corresponding filtered martingale problem.

By uniqueness of the filtered martingale problem, \( \pi_t = H(t, Y) \) for some appropriately measurable transformation. But this transformation doesn’t depend on the change of measure, so \( \pi \) was already the conditional distribution process.
References


Abstract

Martingale problems and filtering

Martingale problems for conditional distributions

Let $X$ be a Markov process characterized as the solution of a martingale problem with generator $A$, and let $Y(t)$ be given by a function of $X(t)$. The conditional distribution of $X(t)$ given observations of $Y$ up to time $t$ is characterized as the solution of a filtered martingale problem. Uniqueness for the original martingale problem implies uniqueness for the filtered martingale problem which in turn implies the Markov property for the conditional distribution considered as a probability-measure-valued process.

Derivation and uniqueness for filtering equations

The conditional distribution of a partially observed Markov process can be characterized as a solution of a filtered martingale problem.
In a variety of settings, this characterization in turn implies that the conditional distribution is given as the unique solution of a filtering equation. Previous results will be reviewed, and new uniqueness results based on local martingale problems and a local forward equation will be presented.