Maximum Cross Section Method in Optimal Filtering of Jump-Diffusion Random Processes

T. A. Averina; K. A. Rybakov

ICM & MG SB RAS, NSU; MAI

26 August 2019

Optimal estimation problem

Signal model is described by the Itô SDEs with a compound Poisson process:

$$dX(t) = f\big(t, X(t)\big)dt + \sigma\big(t, X(t)\big)dW(t) + \int_{\mathbb{R}^q} v\big(t, X(t^-), \xi\big)\nu(dt \times d\xi), \quad X(0) = X_0, \quad (1)$$

where

- $t \in \mathbb{T} = [t_0, T]$ is a time, $X \in \mathbb{R}^n$ is a state,
- $f(t,x)\colon \mathbb{T} \times \mathbb{R}^n \to \mathbb{R}^n$, $\sigma(t,x)\colon \mathbb{T} \times \mathbb{R}^n \to \mathbb{R}^{n\times s}$,
- $W(t)$ is s-dimensional Wiener process,
- $v(t,x,\xi)\colon \mathbb{T} \times \mathbb{R}^n \times \mathbb{R}^q \to \mathbb{R}^n$,
- ν is the Poisson random measure on $\mathbb{T} \times \mathbb{R}^q$ with the characteristic measure Π_ν, $\pi(t,x,\xi)\colon \mathbb{T} \times \mathbb{R}^n \times \mathbb{R}^q \to \mathbb{R}_+$,
- X_0 is an initial state with a probability density $\varphi_0(x)$.

Optimal estimation problem

Let $\lambda(t,x)\colon \mathbb{T}\times\mathbb{R}^n \to \mathbb{R}_+$ denote the intensity and let $\psi(t,\delta)\colon \mathbb{T}\times\mathbb{R}^n \to \mathbb{R}_+$ denote the probability density function for jumps (random increments of the state vector). Thus, for $v(t,x,\xi)=\xi$

$$\begin{aligned}\Pr\big(P(t+\Delta t)-P(t)=1 \,|\, X(t)=x\big)\\ =\lambda(t,x)\Delta t+o(\Delta t)\end{aligned} \qquad (2)$$

for small $\Delta t>0$, and

$$X(\tau_j)=X(\tau_j^-)+\Delta_j, \quad j=1,2,\dots, \qquad (3)$$

where Pr is a probability, $P(t)$ is the Poisson process, $\psi(\tau_j,\delta)$ is the probability density function for Δ_j, $\{\tau_j\}$ are points of the Poisson process $P(t)$, $\tau_0=0$.

Optimal estimation problem

Observation model is described by the Itô SDEs:

$$dY(t) = c\big(t, X(t)\big)dt + \zeta(t)dV(t), \quad Y(0) = Y_0 = 0, \qquad (4)$$

where

- $Y \in \mathbb{R}^m$ is an observation,
- $c(t,x)\colon \mathbb{T} \times \mathbb{R}^n \to \mathbb{R}^m$, $\zeta(t)\colon \mathbb{T} \to \mathbb{R}^{m\times d}$, $|\zeta(t)\zeta^{\mathrm{T}}(t)| \neq 0$,
- $V(t)$ is d-dimensional Wiener process ($W(t)$, $V(t)$ and X_0 are independent).

Optimal estimation problem

The optimal estimation problem is to find an estimate $\hat{X}(\theta)$ given the observations $Y_0^t = \{Y(\tau),\ \tau \in [0,t]\}$ such that $\hat{X}(\theta) = \psi(\theta, Y_0^t)$, where the function $\psi(\theta, \cdot\,)$ satisfies for all $\theta \in \mathbb{T}$ the following condition:

$$\mathrm{E}\big[\big(X(\theta) - \hat{X}(\theta)\big)^{\mathrm{T}}\big(X(\theta) - \hat{X}(\theta)\big)\big] \to \min_{\psi(\theta,\,\cdot\,)} .$$

This implies that $\hat{X}(\theta) = \psi(\theta, Y_0^t) = \mathrm{E}\big[X(\theta) | Y_0^t\big]$.

For $\theta = t$ we have the filtering problem, for $\theta < t$ and $\theta > t$ we have the smoothing problem and the prediction problem, respectively.

Main results in the theory of diffusion processes filtering

- Stratonovich–Kushner equation for the conditional probability density function (end of the 1950s).
- Kalman–Bucy filter (beginning of the 1960s).
- Duncan–Mortensen–Zakai equation for the unnormalized conditional probability density function (second half of the 1960s).
- Kallianpur–Striebel formula for the probabilistic representation of the conditional probability density function (end of the 1960s).

Approximation methods in the theory of diffusion processes filtering

- Linearization of the signal and observation equations (Kalman-type filters).
- Methods based on parametric or functional approximation of the conditional probability density.
- Conditionally optimal filtering (optimizing the filter structure).
- **Statistical modeling method, or Monte Carlo method** (Particle filters).

Duncan–Mortensen–Zakai equation

The equation for the unnormalized conditional probability density function $\varphi(t, x|Y_0^t)$:

$$\frac{\partial \varphi(t, x|Y_0^t)}{\partial t} = \mathcal{L}\varphi(t, x|Y_0^t) + \mu\left(t, x, \frac{dY(t)}{dt}\right)\varphi(t, x|Y_0^t)$$

with the initial condition $\varphi(t_0, x) = \varphi_0(x)$, where

$$\mathcal{L}\varphi(t, x|Y_0^t) = -\sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left[f_i(t, x)\varphi(t, x|Y_0^t)\right] +$$

$$+ \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j}\left[g_{ij}(t, x)\varphi(t, x|Y_0^t)\right] - \lambda(t, x)\varphi(t, x|Y_0^t) +$$

$$+ \int_{\mathbb{R}^n} \lambda(t, \xi)\psi(t, x - \xi)\varphi(t, \xi|Y_0^t)d\xi, \quad g(t, x) = \sigma(t, x)\sigma^{\mathrm{T}}(t, x),$$

$$\mu(t, x, z) = c^{\mathrm{T}}(t, x)\eta^{-1}(t)\, z - \frac{1}{2}c^{\mathrm{T}}(t, x)\eta^{-1}(t)c(t, x).$$

Statistical modeling method

To find the approximate solution of the filtering problem it is necessary to simulate M sample paths $X^i(t)$ of the random process $X(t)$ and the corresponding paths $\omega^i(t)$ of the weight function $\omega(t)$ by a numerical method for relations

$$dX(t) = f\big(t, X(t)\big)dt + \sigma\big(t, X(t)\big)dW(t) + \int_{\mathbb{R}^q} v\big(t, X(t^-), \xi\big)\nu(dt \times d\xi), \quad X(0) = X_0, \quad (1)$$

and

$$\omega(t) = \exp\left\{\int_0^t \mu\left(\tau, X(\tau), \frac{dY(\tau)}{d\tau}\right)d\tau\right\} = \exp\left\{\int_0^t c^{\mathrm{T}}\big(\tau, X(\tau)\big)\eta^{-1}(\tau)dY(\tau) - \frac{1}{2}\int_0^t c^{\mathrm{T}}\big(\tau, X(\tau)\big)\eta^{-1}(\tau)c\big(\tau, X(\tau)\big)d\tau\right\} \quad (5)$$

taking into account points of the Poisson process $P(t)$.

Statistical modeling method

For example, using the Euler–Maruyama method for SDE without a compound Poisson process we have

$$X_{k+1} = X_k + hf(t_k, X_k) + \sqrt{h}\sigma(t_k, X_k)\zeta_k,\ \ \zeta_k \sim \mathcal{N}(0, I_{s\times s}),$$
$$\omega_{k+1} = \omega_k \exp\Big\{ c^{\mathrm{T}}(t_k, X_k)\eta^{-1}(t_k)\big(Y(t_{k+1}) - Y(t_k)\big)$$
$$- \frac{1}{2} c^{\mathrm{T}}(t_k, X_k)\eta^{-1}(t_k)c(t_k, X_k)h \Big\}, \quad \omega_0 = 1,$$

where $\{t_k\}$ is a discretization of the time interval $\mathbb{T}$ with a step size $h > 0$:

$$t_{k+1} = t_k + h, \quad k = 1, 2, \ldots, N; \quad t_0 = 0, \quad t_N = T, \quad N = \frac{T}{h}.$$

Optimal estimation

The exact solution of the optimal filtering problem is

$$\hat{X}(\theta) = \mathrm{E}[X(\theta) \,|\, Y_0^t] = \frac{\mathrm{E}[\omega(t) X(\theta)]}{\mathrm{E}\,\omega(t)}.$$

The approximate solution of the optimal estimation problem is

$$\hat{X}(t_\kappa) \approx \hat{X}_\kappa = \left(\sum_{i=1}^{M} \omega_k^i\right)^{-1} \sum_{i=1}^{M} \omega_k^i X_\kappa^i,$$

where the index k corresponds to the current time $t = t_k$ and the index κ corresponds to the time $\theta = t_\kappa$ for which the state vector estimate is calculated. The higher order moments can be also found as well as estimations of the probability density function or distribution function of the state vector.

Maximum cross section method

Using the maximum cross section method:

- Solution of the transfer equation.
- Analysis of jump-diffusion systems.
- Analysis of switching diffusion systems (systems with random structure).
- Estimation in jump-diffusion systems.
- Estimation in switching diffusion systems (systems with random structure).

Maximum cross section method

If there exists λ^* such that $\lambda(t) \leqslant \lambda^*$, then the random time between neighboring points τ_j and τ_{j+1} should be simulated as follows

$$\tau = \theta_{\mathcal{N}}, \quad \mathcal{N} = \min\left\{\vartheta\colon \;\; \alpha_\vartheta \leqslant \frac{\lambda(\tau_j + \theta_\vartheta)}{\lambda^*}\right\}, \quad \theta_\vartheta = \sum_{i=1}^{\vartheta} \xi^i,$$

where $\xi^1, \xi^2, \ldots, \xi^\vartheta, \ldots$ is a sequence of independent random variables having the exponential distribution with the rate parameter λ^*: $\xi^i = -\ln \beta_i / \lambda^*$; $\alpha_1, \alpha_2, \ldots, \alpha_\vartheta, \ldots,$ $\beta_1, \beta_2, \ldots, \beta_\vartheta, \ldots$ is a sequence of independent random variables having the uniform distribution on the interval $(0, 1)$, and $\lambda(t) = \lambda(t, X(t))$.

Modified maximum cross section method

The modified maximum cross section method is more efficient due to fewer random number generator calls, and for this modified method the number $\mathcal{N}$ is defined by

$$\mathcal{N} = \min\left\{\vartheta\colon\;\; 1-\alpha > \prod_{i=1}^{\vartheta}\left(1-\frac{\lambda(\tau_j+\theta_i)}{\lambda^*}\right)\right\},$$

where α is a random variable having the uniform distribution on the interval $(0,1)$.

Numerical example

Signal and observation model includes two equations

$$
\begin{aligned}
dX(t) &= \mu X(t)dt + \sigma dW(t) \\
&\qquad + \int_{\mathbb{R}} v\big(t, X(t^-), \xi\big)\nu(dt \times d\xi), \quad X(0) = X_0 = 1, \\
dY(t) &= X^3(t)dt + \zeta dV(t), \quad Y(0) = Y_0 = 0,
\end{aligned}
$$

where $X, Y \in \mathbb{R}$, $t \in \mathbb{T} = [0, 1]$, $W(t)$ and $V(t)$ are one-dimensional standard Wiener processes. The compound Poisson process $P^c(t)$ is defined by the intensity $\lambda(t, x) = \varepsilon\, t\,(1 + \cos x)$, i.e., $\lambda^* = 2\,\varepsilon$, and the standard normal distribution for jumps; μ, σ, ε are parameters.

Numerical example

Filtering algorithms are used with following parameters: time discretization step $h = 0.001$ for the Euler–Maruyama method, sample size $M = 10^4$, $\mu = 0.15$, $\sigma = 0.1$, $\varepsilon = 1; 5; 10; 50; 100$.

Algorithms for the Poisson point modeling:

1. Algorithm based on the simple Poisson point modeling (points are modeled approximately so that they coincide with some points of the discretization $\{t_k\}$).
2. Algorithm based on the maximum cross section method.
3. Algorithm based on the modified maximum cross section method.

Numerical example

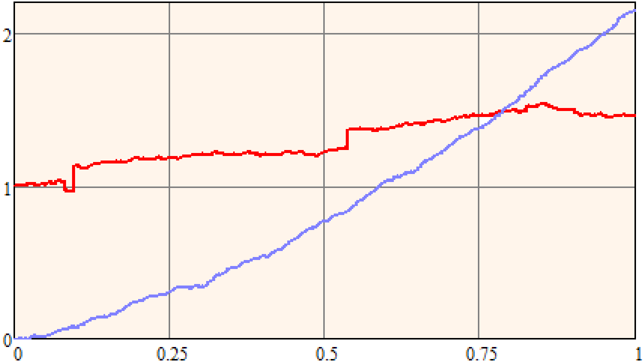
Random number generator calls:

ε	Algorithm		
	1	2	3
1	10000000	29864	28223
5	10000000	109848	88773
10	10000000	209526	158329
50	10000000	1011466	623500
100	10000000	2005794	1163052

Table illustrates random number generator calls to simulate points of the Poisson process only (random number generator calls to simulate the Wiener process increments and jumps are not taken into account).

Numerical example

Signal and observations ($\varepsilon = 5$):



Numerical example

Signal, mean, filtering result, smoothing result for $\theta \in [0, 0.25]$, prediction result for $\theta \in [0.75, 1]$: