Detecting Stealthy Behaviour using Markov Observation Models (MOM)

Michael Kouritzin

June 2022

Hidden Markov vs Markov Observation Model

- Hidden Markov state X with transitions $p_{x \to x'}$ for $x, x' \in E$. Discrete observation space O.
- HMM (old) - Baum & Petrie (1966); Baum & Eagon (1967) - distorted, corrupted partial observation Y - emission probabilities:
$$P\left(Y_n \in A \middle| X_n, X_{n-1}, ..., X_1\right) = P\left(Y_n \in A \middle| X_n\right) \quad \text{for } A \subset O.$$
- MOM (new) - Observations Markov, depending on hidden state:
$$P\Big(X_{n+1} = x, Y_{n+1} = y \Big| X_n = x_n, Y_n = y_n\Big) = p_{x_n \to x} p^Y_{y_n \to y}(x)$$
- Emission probability generalized to
$$P\left(Y_n \in A \middle| X_n, X_{n-1}, ..., X_1; Y_{n-1}, ..., Y_1\right) = p^Y_{Y_{n-1} \to Y_n}\left(X_n\right).$$
- Unseen initial distribution to be identified
$$P\left(X_0 \in dx_0, Y_0 \in dy_0\right) = \mu\left(dx_0, dy_0\right).$$

Expectation-Maximization (EM) Algorithm for MOM

- After Baum-Welch-like calculations

$$
\begin{aligned}
p^{k+1}_{x\to x'} &= \frac{\sum_{n=1}^{N-1} \alpha_n^k(x)\beta_n^k(x',x)}{\sum_{n=1}^{N-1}\sum_{x_{n+1}} \beta_n^k\left(x_{n+1},x\right)\alpha_n^k(x)} \\
p^{Y,k+1}_{i\to j}\left(x\right) &= \frac{\sum_{n=1}^{N-2} 1_{Y_n=i,Y_{n+1}=j}\alpha^k_{n+1}(x)\sum_{x_{n+2}}\beta^k_{n+1}\left(x_{n+2},x\right)}{\sum_{n=1}^{N-2} 1_{Y_n=i}\sum_{x_{n+2}}\beta^k_{n+1}\left(x_{n+2},x\right)\alpha^k_{n+1}(x)}.
\end{aligned}
$$

- $p^k_{x\to x'}$, $p^{Y,k}_{i\to j}(x)$ - converge ($k\to\infty$) to local minimum.

An Expectation-Maximization (EM) Algorithm

- α^k solved forward for $n = 2, 3, ..., N-1, N$

$$\alpha_n^k(x) \;=\; p^{Y,k}_{Y_{n-1}\to Y_n}(x) \sum_{x_{n-1}} \alpha^k_{n-1}(x_{n-1}) p_{x_{n-1}\to x} \;,$$

starting at $\alpha_1^k(x_1) = \int\int \mu(dx, dy) p^k_{x\to x_1}\, p^{Y,k}_{y\to Y_1}(x_1)$.

- β^k solved backward for $n = N-2, N-3, ..., 3, 2, 1, 0$

$$\beta_n^k(x_{n+1}, x) \;=\; \sum_{x'} \beta^k_{n+1}(x', x_{n+1}) p_{x\to x_{n+1}} p^{Y,k}_{Y_n\to Y_{n+1}}(x_{n+1}),$$

starting from $\beta^k_{N-1}(x_N, x_{N-1}) = p^k_{x_{N-1}\to x_N} p^{Y,k}_{Y_{N-1}\to Y_N}(x_N)$.

- α^∞ provides real time tracking filter

$$\pi_n(x) = P\left(X_n = x \middle| Y_1, ..., Y_n\right) = \frac{\alpha_n^\infty(x)}{\sum_\xi \alpha_n^\infty(\xi)},\ \forall x \in E$$

Bayes Factor for Model Selection and μ

- Number hidden states? Global max for $p_{i\to j}$? Initial distribution μ?
- Use Bayes factor to reference model with (some Q and) no hidden state dependence:

$$Q\left(X_{n+1}=x,\,Y_{n+1}=y\middle|X_n=x_n,\,Y_n=y_n\right)=p_{x_n\to x}\overline{p}^{Y}_{y_n\to y}.$$

- The likelihood ratio function that does the conversion is:

$$A_m=\prod_{0<n\leq m}\frac{p^{Y}_{Y_{n-1}\to Y_n}(X_n)}{\overline{p}^{Y}_{Y_{n-1}\to Y_n}}\ \forall m\in 0,\ldots,N,\ \frac{dP}{dQ}=A_N.$$

Theorem

$\{A_m\}$ is a $\{\mathcal{F}_m\vee\mathcal{F}^X_N\}_{m\in 0,\ldots,N}$-martingale with Q. Moreover, $\begin{pmatrix}X\\Y\end{pmatrix}$ is our MOM model with initial law $\mu(dx_0,dy_0)$ and

$$P\left(X_{n+1}=x,\,Y_{n+1}=y\middle|X_n=x_n,\,Y_n=y_n\right)=p_{x_n\to x}p^{Y}_{y_n\to y}(x)\ .$$

Bayes Factor for Model Selection and μ

- Fix $\overline{p}^Y_{i\to j}$ offline.
- Bayes factor $B^\mu_N = E[A_N|\mathcal{F}^Y_N]$ rates model and initial distribution.
- If $\rho_n(x_{n+1}) = E^Q\left[A_n p_{X_n\to x_{n+1}}\middle|\mathcal{F}^Y_n\right]$, then

$$B^\mu_N = \sum_x \frac{p^Y_{Y_{N-1}\to Y_N}(x)}{\overline{p}^Y_{Y_{N-1}\to Y_N}}\rho_{N-1}(x) \quad \text{and}$$

$$\rho_n(x_{n+1}) = \sum_{x_n} \frac{p_{x_n\to x_{n+1}} p^Y_{Y_{n-1}\to Y_n}(x_n)}{\overline{p}^Y_{Y_{n-1}\to Y_n}}\rho_{n-1}(x_n)$$

subject to $\rho_0(x_1) = \int_{E\times O} p_{x_0\to x_1}\mu(dx_0, dy_0)$.

- Let $Unif(A;M)$ be selection of M points from A without replacement.
- Just look for best $\mu = \delta_{x_0,y_0}$.

Calibration - BFF Outer layer

Let $D = \emptyset$ be points done; $C = Unif(E \times O; M)$ points considering.

Let $k = 1$ and $\{p^1_{x \to x'}\}, \{p^{Y,1}_{i \to j}(x)\}$ be first guesses.

While $C \neq \emptyset$ or transitions not converged do

- Do prior updates here for $\{p^{k+1}_{x \to x'}\}, \{p^{Y,k+1}_{i \to j}(x)\}, B^{\mu,k+1}_N$
- Move the μ with the m worst $B^{\mu,k+1}_N$ from C to D.
- Add $Unif(E \times O/(D \cup C); m)$ to C.
- For each added point randomly select $p^{k+1}_{x \to x'}, p^{Y,k+1}_{i \to j}(x), B^{\mu,k+1}_N$ from among the $M - m$ kept μ.

Issues

- Still have not used timing between packet information.
- Hidden state and observations should be continuous time.

Continuous MOM

- Key to unlocking information is next importance sampling result:

Theorem

Suppose Y is a Markov chain with rates $\overline{\gamma}_{i\to j}$ independent of X w.r.t. Q, N counts the transitions of Y and

$$A_t = \exp\left(\int_0^t \overline{\gamma}_{Y_s\to} - \gamma_{Y_s\to}(X_s)\,ds\right)\prod_{0<s\le t}\left[1+\left(\frac{\gamma_{Y_{s-}\to Y_s}(X_s)}{\overline{\gamma}_{Y_{s-}\to Y_s}}-1\right)\Delta N_s\right].$$

Then, A is a Q-martingale and Y has rates $\gamma_{i\to j}(X)$ under $\frac{dP}{dQ}=A_T$.

- Continuous-time MOM when X is Markov, say chain rates $\lambda_{x\to x'}$.
- Used notation $\gamma_{i\to}(x)=\sum_j \gamma_{i\to j}(x)$

Continuous MOM

- **Filter** $\pi_t(B) = P\left(X_t \in B \middle| \mathcal{F}_t^Y\right)$, for Borel $B \subset E$ tracks hidden state with model and back observations.
- **Unnormalized filter** $\sigma_t(B) = E^Q\left(A_t 1_{X_t \in B} \middle| \mathcal{F}_t^Y\right)$ provides the filter through Bayes rule $\pi_t(f) = \frac{\sigma_t(f)}{\sigma_t(1)}$, where $\pi_t(f) \doteq \int_E f d\pi_t$.
- Kalman & Bucy (1961), Zakai (1969), Fujisaki-Kallianpur-Kunita (1972) - filters with classical distorted, corrupted, partial observations.

Theorem

σ is the unique strong $D_{\mathcal{M}_f(E)}[0,\infty)$-valued solution to:

$$\sigma_t(f(\cdot)) = \sigma_0(f(\cdot)) + \int_0^t \sigma_s(Lf(\cdot))ds + \int_0^t \sigma_s(f(\cdot)(\overline{\gamma}_{Y_s\to} - \gamma_{Y_s\to}(\cdot))ds$$
$$+ \int_0^t \sigma_{s-}\left(\left[f(\cdot)\frac{\gamma_{Y_{s-}\to Y_s}(\cdot)}{\overline{\gamma}_{Y_{s-}\to Y_s}} - f(\cdot)\right]\right) dN_s \text{ a.s. } \forall f \in \overline{C}(E).$$

Continuous MOM Zakai with Hidden Chain

- X is Markov with $Lf(i) = \sum_{j \neq i} \lambda_{i \to j} [f(j) - f(i)]$ for $i \in \{1, 2, ..., m\}$.
- If $\sigma_t^i = \sigma_t(\delta_i)$ for $i = 1, 2, ..., m$, then get closed system of equations:

$$d \begin{bmatrix} \sigma_t^1 \\ \sigma_t^2 \\ \vdots \\ \sigma_t^m \end{bmatrix} = \begin{bmatrix} \frac{\gamma_{Y_{t-} \to Y_t}(1)}{\overline{\gamma}_{Y_{t-} \to Y_t}} - 1 & 0 & \cdots & 0 \\ 0 & \frac{\gamma_{Y_{t-} \to Y_t}(2)}{\overline{\gamma}_{Y_{t-} \to Y_t}} - 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & & \frac{\gamma_{Y_{t-} \to Y_t}(m)}{\overline{\gamma}_{Y_{t-} \to Y_t}} - 1 \end{bmatrix} \begin{bmatrix} \sigma_{t-}^1 \\ \sigma_{t-}^2 \\ \vdots \\ \sigma_{t-}^m \end{bmatrix} dN_t$$

$$+ \begin{bmatrix} \overline{\gamma}_{Y_t \to} - \gamma_{Y_t \to}(1) - \lambda_{1 \to} & \lambda_{2 \to 1} & \cdots & \lambda_{m \to 1} \\ \lambda_{1 \to 2} & \overline{\gamma}_{Y_t \to} - \gamma_{Y_t \to}(2) - \lambda_{2 \to} & \cdots & \lambda_{m \to 2} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{1 \to m} & \lambda_{2 \to m} & \cdots & \overline{\gamma}_{Y_t \to} - \gamma_{Y_t \to}(m) - \lambda_{m \to} \end{bmatrix} \begin{bmatrix} \sigma_t^1 \\ \sigma_t^2 \\ \vdots \\ \sigma_t^m \end{bmatrix} dt.$$

Continuous MOM Zakai with Hidden Chain

- Recognizing the adjoint suggests a mild solution with (state transition-observation weight) Trotter product semi-group using

$$S_t^n = \begin{bmatrix} P_t(1\to 1) & P_t(2\to 1) & \cdots & P_t(m\to 1) \\ P_t(1\to 2) & P_t(2\to 2) & \cdots & P_t(m\to 2) \\ \vdots & \vdots & \ddots & \vdots \\ P_t(1\to m) & P_t(2\to m) & \cdots & P_t(m\to m) \end{bmatrix}$$

$$* \begin{bmatrix} e^{t(\overline{\gamma}_{Y_{t_{n-1}}\to}-\gamma_{Y_{t_{n-1}}\to}(1))} & 0 & \cdots & 0 \\ 0 & e^{t(\overline{\gamma}_{Y_{t_{n-1}}\to}-\gamma_{Y_{t_{n-1}}\to}(2))} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{t(\overline{\gamma}_{Y_{t_{n-1}}\to}-\gamma_{Y_{t_{n-1}}\to}(m))} \end{bmatrix}$$

Continuous MOM Zakai with Hidden Chain

- Mild solution is $\begin{bmatrix} \sigma_t^1 \\ \sigma_t^2 \\ \vdots \\ \sigma_t^m \end{bmatrix} = \left[S^n_{\frac{t-t_{n-1}}{N}} \right]^N \begin{bmatrix} \sigma_{t_{n-1}}^1 \\ \sigma_{t_{n-1}}^2 \\ \vdots \\ \sigma_{t_{n-1}}^m \end{bmatrix}$ for $t \in [t_{n-1}, t_n)$,

  (some large N) with observation outcome updates

$$\begin{bmatrix} \sigma_{t_n}^1 \\ \sigma_{t_n}^2 \\ \vdots \\ \sigma_{t_n}^m \end{bmatrix} = \begin{bmatrix} \frac{\gamma_{Y_{t_{n-1}} \to Y_{t_n}}(1)}{\overline{\gamma}_{Y_{t_{n-1}} \to Y_{t_n}}} \sigma_{t_n-}^1 \\ \frac{\gamma_{Y_{t_{n-1}} \to Y_{t_n}}(2)}{\overline{\gamma}_{Y_{t_{n-1}} \to Y_{t_n}}} \sigma_{t_n-}^2 \\ \vdots \\ \frac{\gamma_{Y_{t_{n-1}} \to Y_{t_n}}(m)}{\overline{\gamma}_{Y_{t_{n-1}} \to Y_{t_n}}} \sigma_{t_n-}^m \end{bmatrix}.$$

Issue

- Multiple hidden models and unknown source in packet information?

Counting Measure MOM

- Independent behavior states $b = \{b^i\}_{i=1}^k$, b^i unique solution to:

$$m_t^h = h(b_t^i) - h(b_0^i) - \int_0^t L^{b,i} h(b_s^i) ds$$

  is a martingale for all $h \in D(L^{b,i})$.

- $(U, \mathcal{U}, \nu)$ is a measure space for marks U.
- $\{N^i\}$ are independent Poisson measure on $U \times [0, \infty)^2$ with intensity measure $\mu = \nu \times \ell^2$, ℓ being Lebesgue measure.
- i^{th} behavior count $\{B^i\}$ on A has rate $\int_A \lambda^i(u, b_{s-}^i, B_{s-}^i, s)\nu(du)$ so

$$B_t^i(A) = B_0^i(A) + \int_{A\times[0,\infty)\times[0,t]} 1_{[0,\lambda^i(u,b_{s-}^i,B_{s-}^i,s)]}(v) N^i(du \times dv \times ds)$$

- The observations are: $Y_t(A) = \sum_{i=1}^k B_t^i(A)$ for $A \in \mathcal{U}$.

Counting MOM - Reference Model Construction

- Under Q, rate $\overline{\lambda}^i$ does not depend on behavior and are independent:
  1. μ-Poisson measures $\{N^i\}_{i=1}^k$.
  2. (Behavior) processes $\{b^i\}_{i=1}^k$ satisfying the $L^{b,i}$-martingale problem.
  3. Categorical RVs $\{\xi_n(u,s), n \in \mathbb{N}, u \in U, s \geq 0\}$ with $Q(\xi_n(u,s) = e_j) = \frac{\overline{\lambda}^j(u,s)}{\sum_{l=1}^k \overline{\lambda}^l(u,s)}$ where $[e_1 e_2 \cdots e_k] = I_k$.
- Observations Poisson on $U \times [0,\infty)$ intensity $\sum_{i=1}^k \overline{\lambda}^i(u,s)\nu(du)ds$

$$Y_t(A) = Y_0(A) + \int_{A\times[0,\infty)\times[0,t]} \sum_{i=1}^k 1_{[0,\overline{\lambda}^i(u,s)]}(v) N^i(du \times dv \times ds).$$

- Behavior counts on Q satisfy vector equation:

$$B_t(A) = \int_{A\times[0,t]} \xi_{Y(U,s)}(u,s) Y(du \times ds) \quad \text{for all } A \subset U$$

so $Y_t = \sum_{i=1}^k B_t^i$.

Counting MOM - Change of Measure

- Suppose $E^Q\left[\prod_{i=1}^k A_t^i\right] = 1$ for all $t > 0$, where

$$\begin{aligned} A_t^i &= \exp\Bigg(\int_{U\times[0,t]}\Big[(\overline{\lambda}^i(u,s) - \lambda^i(u,b_{s-}^i,B_{s-}^i,s))\nu(du)ds \\ &\qquad + \ln\left(\frac{\lambda^i(u,b_{s-}^i,B_{s-}^i,s)}{\overline{\lambda}^i(u,s)}\right)B^i(du,ds)\Big]\Bigg). \end{aligned}$$

Theorem

Suppose $\{B^i\}$ are defined in terms of $Y, \{\xi\}$ as before and probability measure P satisfies $dP\big|_{\mathcal{F}_t} = A_t dQ\big|_{\mathcal{F}_t}$ for all $t \geq 0$. Then under P, $\{(b^i, B^i)\}_{i=1}^k$ are independent and b^i solves the $(L_s^{b,i}, \mu_0^i)$-local martingale problem B^i has $\lambda^i(u, b_{s-}^i, B_{s-}^i, s)$ rates for $i = 1, ..., k$.

Have equations for

$\pi_t(f) \doteq E^P[f(\{(b_t^i, B_t^i)\}_{i=1}^k)\big|\mathcal{F}_t^Y], \sigma_t(f) \doteq E^Q[A_t f(\{(b_t^i, B_t^i)\}_{i=1}^k)\big|\mathcal{F}_t^Y]$

Counting MOM - Branching Particle Filter

- Particles $\{b^{i,l}, B^{i,l}\}_{l=1}^N$ are independent copies of $\{b^i, B^i\}_{i=1}^k$ on Q.
- Weight particles with (likelihood ratio) weights

$$A_t^{i,l} = \exp\left(\int_{U\times[0,t]} \left[(\overline{\lambda}^i(u,s) - \lambda^i(u, b_{s-}^{i,l}, B_{s-}^{i,l}, s))\nu(du)ds \right.\right.$$
$$\left.\left. + \ln\left(\frac{\lambda^i(u, b_{s-}^{i,l}, B_{s-}^{i,l}, s)}{\overline{\lambda}^i(u,s)}\right) B^{i,l}(du,ds)\right]\right).$$

- SLLN implies $\mathbb{S}_t^N(f) \doteq \frac{1}{N}\sum_{l=1}^N A_t^{i,l} f(b_t^{i,l}, B_t^{i,l}) \to \sigma_t(f)$ a.s.
- $\{b^{i,l}, B^{i,l}\}_{l=1}^N$ spread and become unrepresentative of $\{b^i, B^i\}_{i=1}^k$ so branch high weights into several, kill low weights - unbiased.

Counting MOM - Branching Particle Filter

1. Weighted Particle Filter credited to Handschin (1970), Handschin and Mayne (1969).
2. Resampled particle filter Gordon, Salmond and Smith (1993).
3. A version of K (2017) is efficient and satisfies.

Theorem

For Q-almost all Y, the Residual Branching particle filter satisfies:

slln: $\mathbb{S}_t^N \Rightarrow \sigma_t$ *(i.e. weak convergence) a.s.* $[Q^Y]$;

Mlln: $\left|\mathbb{S}_t^N(f) - \sigma_t(f)\right| \overset{N}{\ll} N^{-\beta}$ *a.s.* $[Q^Y]\ \forall f \in \overline{C}(E)_+$, $0 \leq \beta < \frac{1}{2}$.

Thank you

michaelk@ualberta.ca

Auto-Regressive HMM

- Patch to HMM called AR-HMM with Observations:

$$Y_n = \beta_0^{(X_n)} + \beta_1^{(X_n)} Y_{n-1} + \cdots + \beta_p^{(X_n)} Y_{n-p} + \varepsilon_n,$$

- $\{\varepsilon_n\}_{n=1}^{\infty}$ i.i.d. RVs; and each β_1 is a functions of X_n.
- Rewrite as

$$\underbrace{\begin{bmatrix} Y_n \\ Y_{n-1} \\ \vdots \\ Y_{n-p+1} \end{bmatrix}}_{\mathcal{Y}_n} = \begin{bmatrix} \beta_1^{(X_n)} & \beta_2^{(X_n)} & \cdots & \beta_p^{(X_n)} \\ 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots 1 & 0 \end{bmatrix} \underbrace{\begin{bmatrix} Y_{n-1} \\ Y_{n-2} \\ \vdots \\ Y_{n-p} \end{bmatrix}}_{\mathcal{Y}_{n-1}} + \begin{bmatrix} \beta_0^{(X_n)} + \varepsilon_n \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

- Given X_n, an explicit formula in $\mathcal{Y}_{n-1}$ and independent noise $\Rightarrow \{\mathcal{Y}_n\}$ is conditionally Markov.

Let Repeat $k = k + 1$ until transition probabilties converge:

For each $(x_0, y_0) \in C$ do

1. $\alpha_1^k(x_1)$, $\alpha_n^k(x) = p^{Y,k}_{Y_{n-1}\to Y_n}(x) \sum_{x'} \alpha_n^k(x') p_{x'\to x}$ for $n = 2, ..., N$.
2. β_{N-1}^k, $\beta_n^k(\xi, x) = \sum_{x'} \beta_{n+1}^k(x', \xi) p^k_{x\to\xi} p^{Y,k}_{Y_n\to Y_{n+1}}(\xi)$ for $n = N-2, ..., 0$
3. $p^{k+1}_{x\to x'}$, $p^{Y,k+1}_{i\to j}(x) = \dfrac{\sum\limits_{n\le N-2} 1_{Y_n=i, Y_{n+1}=j} \alpha^k_{n+1}(x) \sum\limits_{x_{n+2}} \beta^k_{n+1}(x_{n+2}, x)}{\sum\limits_{n\le N-2} 1_{Y_n=i} \sum\limits_{x_{n+2}} \beta^k_{n+1}(x_{n+2}, x)\, \alpha^k_{n+1}(x)}$,
4. $\rho_1^{k+1}, \rho_n^{k+1}(x_{n+1}) = \sum_{x_n} \dfrac{p^{k+1}_{x_n\to x_{n+1}} p^{Y,k+1}_{Y_{n-1}\to Y_n}(x_n)}{\overline{p}^Y_{Y_{n-1}\to Y_n}} \rho_{n-1}(x_n)$ for $n = 2, ..., N-1$.
5. $B_N^{\mu,k+1} = \sum_x \dfrac{p^{Y,k+1}_{Y_{N-1}\to Y_N}(x)}{\overline{p}^Y_{Y_{N-1}\to Y_N}} \rho^{k+1}_{N-1}(x)$

Adjoint $L^*p\,(j) = [L^*]p(j)$, $[L^*] = \begin{bmatrix} -\lambda_{1\to} & \lambda_{2\to 1} & \cdots & \lambda_{m\to 1} \\ \lambda_{1\to 2} & -\lambda_{2\to} & \cdots & \lambda_{m\to 2} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{1\to m} & \lambda_{2\to m} & & -\lambda_{m\to} \end{bmatrix}$

Continuous MOM Zakai with Hidden Chain

- Start at $\begin{bmatrix} \sigma_{t_0}^1 \\ \sigma_{t_0}^2 \\ \vdots \\ \sigma_{t_0}^m \end{bmatrix} = \begin{bmatrix} P(X_0 = 1) \\ P(X_0 = 2) \\ \vdots \\ P(X_0 = m) \end{bmatrix}$.
- Regain unnormalized filter measure $\sigma_t(\cdot) = \sum_{i=1}^{m} \sigma_t^i \delta_i(\cdot)$.

Counting MOM - Filter Equations

Theorem

σ is the unique strong $D_{\mathcal{M}_f(\mathcal{M}_f^c(U))}[0,\infty)$-valued solution to

$$
\begin{aligned}
\sigma_t(f) &= \sigma_0(f) + \sum_{i=1}^{k} \int_0^t \sigma_s(L^{b,i} f)ds \\
&+ \sum_{i=1}^{k} \int_{U\times[0,t]} \sigma_s(f(\cdot,\cdot)(\overline{\lambda}^i(u,s) - \lambda^i(u,\cdot,\cdot,s)))\nu(du)ds \\
&+ \int_{U\times[0,t]} \sigma_{s-}\left(\frac{\sum_{i=1}^{k} f(\cdot,\cdot + e_i^*\delta_u)\lambda^i(u,\cdot,\cdot,s)}{\sum_{j=1}^{k} \overline{\lambda}^j(u,s)} - f(\cdot,\cdot) \right) Y(du\times ds) \text{ a.s.}
\end{aligned}
$$