Monte Carlo Inference for Jump-Diffusion Processes

Flávio Gonçalves
Gareth Roberts

University of Warwick

20^{th} March, 2009

Content

Content

Jump-diffusions

A jump-diffusion process is defined as the solution $X := \{X_t \,:\, 0 \leq t \leq T\}$ of the SDE:

$$dX_t = b(X_{t-};\theta)dt + \sigma(X_{t-};\theta)dB_t + dJ_t, \quad X_0 = x_0.$$

J_t is a jump process where the jump times follow a Poisson process with rate $\lambda(t, X_{t-};\theta)$ and the jump sizes are given by a function $g(Z_t, X_{t-})$, where Z_t has a distribution $f(Z_t;\theta)$.

Jump-diffusions

A jump-diffusion process is defined as the solution $X := \{X_t \ : \ 0 \leq t \leq T\}$ of the SDE:

$$dX_t = b(X_{t-};\theta)dt + \sigma(X_{t-};\theta)dB_t + dJ_t, \quad X_0 = x_0.$$

J_t is a jump process where the jump times follow a Poisson process with rate $\lambda(t, X_{t-};\theta)$ and the jump sizes are given by a function $g(Z_t, X_{t-})$, where Z_t has a distribution $f(Z_t;\theta)$.

Problem

- $X_{t_1}, X_{t_2}, \ldots, X_{t_n}$ is observed.
- Make inference for θ given the observations.
- Only the likelihood function for the whole path is available $\rightarrow$ EM Algorithm.
- Aim: Use the EM Algorithm (MCEM, actually) to find the MLE of θ based on the observations.
- Beskos, Papaspiliopoulos, Roberts and Fearnhead (2006) proposed an MCEM Algorithm for diffusion processes.

Problem

- $X_{t_1}, X_{t_2}, \ldots, X_{t_n}$ is observed.
- Make inference for θ given the observations.
- Only the likelihood function for the whole path is available $\rightarrow$ EM Algorithm.
- Aim: Use the EM Algorithm (MCEM, actually) to find the MLE of θ based on the observations.
- Beskos, Papaspiliopoulos, Roberts and Fearnhead (2006) proposed an MCEM Algorithm for diffusion processes.

Problem

- $X_{t_1}, X_{t_2}, \ldots, X_{t_n}$ is observed.
- Make inference for θ given the observations.
- Only the likelihood function for the whole path is available $\rightarrow$ EM Algorithm.
- Aim: Use the EM Algorithm (MCEM, actually) to find the MLE of θ based on the observations.
- Beskos, Papaspiliopoulos, Roberts and Fearnhead (2006) proposed an MCEM Algorithm for diffusion processes.

Problem

- $X_{t_1}, X_{t_2}, \ldots, X_{t_n}$ is observed.
- Make inference for θ given the observations.
- Only the likelihood function for the whole path is available $\rightarrow$ EM Algorithm.
- Aim: Use the EM Algorithm (MCEM, actually) to find the MLE of θ based on the observations.
- Beskos, Papaspiliopoulos, Roberts and Fearnhead (2006) proposed an MCEM Algorithm for diffusion processes.

Content

General framework

Proposed by Dempster, Laird and Rubin (1977).

$X_{obs} \sim f(X_{obs}; \theta)$ (unknown).

X_{miss} is not observed.

$X = \{X_{obs}, X_{miss}\} \sim f(X; \theta)$ (known).

Aim: Find θ that maximises $l(X_{obs}; \theta)$.

General framework

Proposed by Dempster, Laird and Rubin (1977).

$X_{obs} \sim f(X_{obs};\theta)$ (unknown).

X_{miss} is not observed.

$X = \{X_{obs}, X_{miss}\} \sim f(X;\theta)$ (known).

Aim: Find θ that maximises $l(X_{obs};\theta)$.

General framework

Proposed by Dempster, Laird and Rubin (1977).

$X_{obs} \sim f(X_{obs};\theta)$ (unknown).

X_{miss} is not observed.

$X = \{X_{obs}, X_{miss}\} \sim f(X;\theta)$ (known).

Aim: Find θ that maximises $l(X_{obs};\theta)$.

Algorithm

- Set a initial value θ'.
- E-step

$$\mathbb{E}_{X_{miss}|X_{obs},\theta'}\left[l(X;\theta)\right] = Q(\theta,\theta')$$

- M-step
  Maximise $Q(\theta,\theta')$ w.r.t. θ and update θ'.

Algorithm

- Set a initial value θ'.
- E-step

$$\mathbb{E}_{X_{miss}|X_{obs},\theta'}\left[l(X;\theta)\right] = Q(\theta,\theta')$$

- M-step
  Maximise $Q(\theta,\theta')$ w.r.t. θ and update θ'.

Algorithm

- Set a initial value θ'.
- E-step

$$\mathbb{E}_{X_{miss}|X_{obs},\theta'}\left[l(X;\theta)\right] = Q(\theta,\theta')$$

- M-step
  Maximise $Q(\theta,\theta')$ w.r.t. θ and update θ'.

Monte Carlo EM

- Introduced in Wei and Tanner (1990).
- $\mathbb{E}_{X_{miss}|X_{obs},\theta'}[l(X;\theta)]$ is not analytically available.
- It is possible to simulate from $(X_{miss}|X_{obs},\theta')$
- Use the Monte Carlo estimator

$$\hat{Q}(\theta,\theta') = \frac{\sum_{i=1}^{M} l(X^{(i)};\theta)}{M}$$

- $\hat{Q}(\theta,\theta') \longrightarrow Q(\theta,\theta')$ as $M \to \infty$

Monte Carlo EM

- Introduced in Wei and Tanner (1990).
- $\mathbb{E}_{X_{miss}|X_{obs},\theta'}\left[l(X;\theta)\right]$ is not analytically available.
- It is possible to simulate from $(X_{miss}|X_{obs},\theta')$
- Use the Monte Carlo estimator

$$\hat{Q}(\theta,\theta') = \frac{\sum_{i=1}^{M} l(X^{(i)};\theta)}{M}$$

- $\hat{Q}(\theta,\theta') \longrightarrow Q(\theta,\theta')$ as $M \to \infty$

Monte Carlo EM

- Introduced in Wei and Tanner (1990).
- $\mathbb{E}_{X_{miss}|X_{obs},\theta'}\left[l(X;\theta)\right]$ is not analytically available.
- It is possible to simulate from $(X_{miss}|X_{obs},\theta')$
- Use the Monte Carlo estimator

$$\hat{Q}(\theta,\theta') = \frac{\sum_{i=1}^{M} l(X^{(i)};\theta)}{M}$$

- $\hat{Q}(\theta,\theta') \longrightarrow Q(\theta,\theta') \quad \text{as} \quad M \to \infty$

Monte Carlo EM

- Introduced in Wei and Tanner (1990).
- $\mathbb{E}_{X_{miss}|X_{obs},\theta'}\left[l(X;\theta)\right]$ is not analytically available.
- It is possible to simulate from $(X_{miss}|X_{obs},\theta')$
- Use the Monte Carlo estimator

$$\hat{Q}(\theta,\theta') = \frac{\sum_{i=1}^{M} l(X^{(i)};\theta)}{M}$$

- $\hat{Q}(\theta,\theta') \longrightarrow Q(\theta,\theta') \quad \text{as} \quad M \to \infty$

Content

Exact Algorithm (EA)

- Proposed by Beskos, Papaspiliopoulos and Roberts (2006).
- Exact simulation of a class of Itô's diffusions.
- Exact in the sense that there is no discretisation error.
- The EA performs retrospective rejection sampling by proposing paths from processes that we can simulate and accepting them according to appropriate probability density ratios.
- The novelty lies in the fact that the paths proposed are unveiled only at finite (but random) time instances and the decision whether to accept the path or not can be easily taken.

Exact Algorithm (EA)

- Proposed by Beskos, Papaspiliopoulos and Roberts (2006).
- Exact simulation of a class of Itô's diffusions.
- Exact in the sense that there is no discretisation error.
- The EA performs retrospective rejection sampling by proposing paths from processes that we can simulate and accepting them according to appropriate probability density ratios.
- The novelty lies in the fact that the paths proposed are unveiled only at finite (but random) time instances and the decision whether to accept the path or not can be easily taken.

Exact Algorithm (EA)

- Proposed by Beskos, Papaspiliopoulos and Roberts (2006).
- Exact simulation of a class of Itô's diffusions.
- Exact in the sense that there is no discretisation error.
- The EA performs retrospective rejection sampling by proposing paths from processes that we can simulate and accepting them according to appropriate probability density ratios.
- The novelty lies in the fact that the paths proposed are unveiled only at finite (but random) time instances and the decision whether to accept the path or not can be easily taken.

Exact Algorithm (EA)

- Proposed by Beskos, Papaspiliopoulos and Roberts (2006).
- Exact simulation of a class of Itô's diffusions.
- Exact in the sense that there is no discretisation error.
- The EA performs retrospective rejection sampling by proposing paths from processes that we can simulate and accepting them according to appropriate probability density ratios.
- The novelty lies in the fact that the paths proposed are unveiled only at finite (but random) time instances and the decision whether to accept the path or not can be easily taken.

Idea

- Proposes paths from a Brownian Bridge.
- $\frac{d\mathbb{Q}}{d\mathbb{W}} \propto \exp\left\{-\int_0^T \phi(X_t)dt\right\} \leq 1$

Theorem

Let X be any continuous mapping from $[0,T]$ to $\mathbb{R}$, and $M(X)$ an upper bound for the mapping $t \mapsto \phi(X_t)$, $t \in [0,T]$. If Φ is a homogeneous Poisson process of unit intensity on $[0,T] \times [0,M(X)]$ and N is the number of points of Φ found below the graph $\{(t,\phi(X_t));\ t \in [0,T]\}$, then

$$P(N=0|X) = \exp\left\{-\int_0^T \phi(X_t)dt\right\}.$$

Idea

- Proposes paths from a Brownian Bridge.
- $\frac{d\mathbb{Q}}{d\mathbb{W}} \propto \exp\left\{-\int_0^T \phi(X_t)dt\right\} \leq 1$

Theorem

Let X be any continuous mapping from $[0, T]$ to $\mathbb{R}$, and $M(X)$ an upper bound for the mapping $t \mapsto \phi(X_t)$, $t \in [0, T]$. If Φ is a homogeneous Poisson process of unit intensity on $[0, T] \times [0, M(X)]$ and N is the number of points of Φ found below the graph $\{(t, \phi(X_t));\ t \in [0, T]\}$, then

$$P(N = 0|X) = \exp\left\{-\int_0^T \phi(X_t)dt\right\}.$$

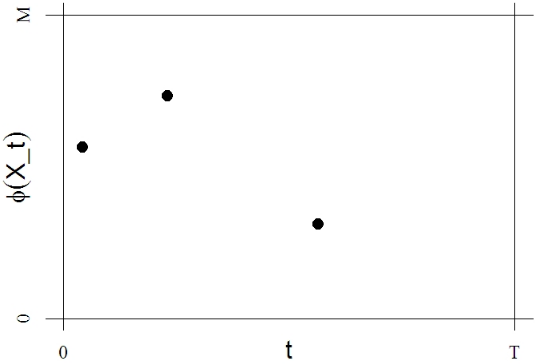
Idea

- Proposes paths from a Brownian Bridge.
- $\dfrac{d\mathbb{Q}}{d\mathbb{W}} \propto \exp\left\{-\int_0^T \phi(X_t)dt\right\} \leq 1$

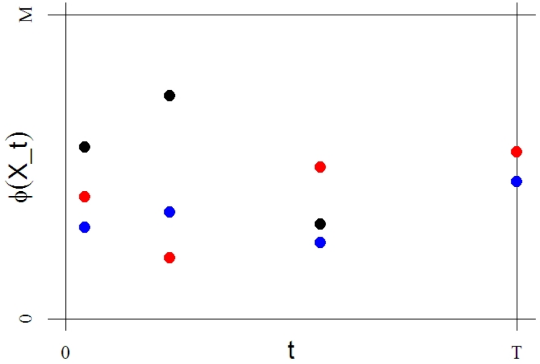
Theorem

Let X be any continuous mapping from $[0, T]$ to $\mathbb{R}$, and $M(X)$ an upper bound for the mapping $t \mapsto \phi(X_t)$, $t \in [0, T]$. If Φ is a homogeneous Poisson process of unit intensity on $[0, T] \times [0, M(X)]$ and N is the number of points of Φ found below the graph $\{(t, \phi(X_t));\ t \in [0, T]\}$, then

$$P(N = 0|X) = \exp\left\{-\int_0^T \phi(X_t)dt\right\}.$$

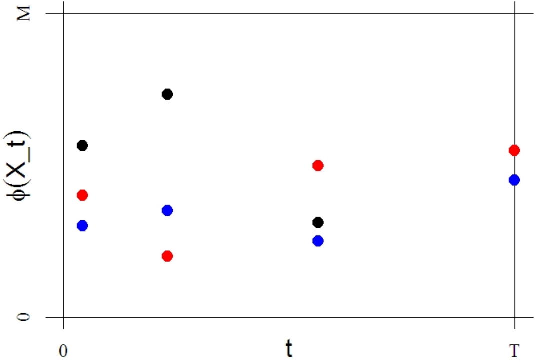


Returns a skeleton of the processes in $[0, T]$.
Any additional information on the process can be easily simulated.



Returns a skeleton of the processes in $[0, T]$.

Any additional information on the process can be easily simulated.

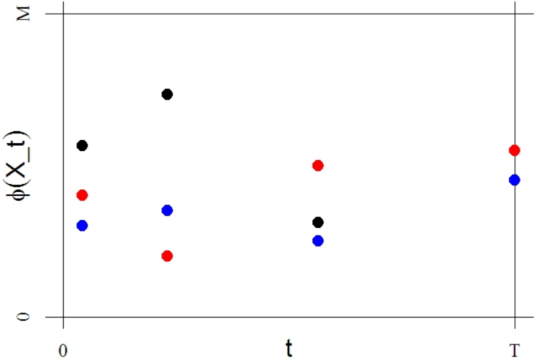


X is accepted and

X is rejected.

Returns a skeleton of the processes in $[0, T]$.

Any additional information on the process can be easily simulated.

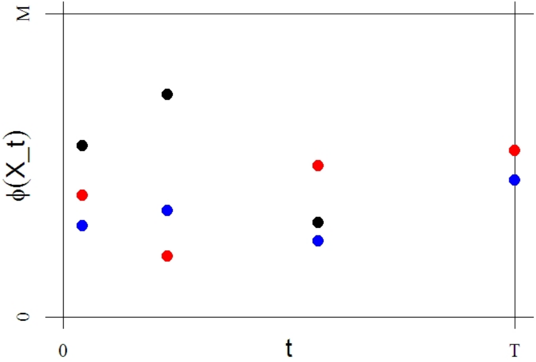


X is accepted and

X is rejected.

Returns a skeleton of the processes in $[0, T]$.

Any additional information on the process can be easily simulated.



X is accepted and

X is rejected.

Returns a skeleton of the processes in $[0, T]$.
Any additional information on the process can be easily simulated.

Content

Preliminaries

$$dX_t = b(X_{t-};\theta)dt + \sigma(X_{t-};\theta)dB_t + dJ_t, \quad X_0 = x_0.$$

$$J_t = \{PP(\lambda(t, X_{t-};\theta)); g(Z_t, X_{t-}), f(Z_t;\theta)\}$$

Transformation: $Y_t = \eta(X_t) = \int_z^{X_t} \frac{1}{\sigma(u;\theta)} du$

$$dY_t = \alpha(Y_{t-};\theta)dt + dB_t + dJ_t, \quad Y_0 = x.$$

$$J_t = \{PP(\lambda(t, \eta^{-1}(Y_t);\theta)); g_1(Z_t, Y_{t-}), f(Z_t;\theta)\}$$

Preliminaries

$$dX_t = b(X_{t-};\theta)dt + \sigma(X_{t-};\theta)dB_t + dJ_t, \quad X_0 = x_0.$$

$$J_t = \{PP(\lambda(t, X_{t-};\theta)); g(Z_t, X_{t-}), f(Z_t;\theta)\}$$

Transformation: $Y_t = \eta(X_t) = \int_z^{X_t} \frac{1}{\sigma(u;\theta)} du$

$$dY_t = \alpha(Y_{t-};\theta)dt + dB_t + dJ_t, \quad Y_0 = x.$$

$$J_t = \{PP(\lambda(t, \eta^{-1}(Y_t);\theta)); g_1(Z_t, Y_{t-}), f(Z_t;\theta)\}$$

Preliminaries

$$dX_t = b(X_{t-};\theta)dt + \sigma(X_{t-};\theta)dB_t + dJ_t, \quad X_0 = x_0.$$

$$J_t = \{PP(\lambda(t, X_{t-};\theta)); g(Z_t, X_{t-}), f(Z_t;\theta)\}$$

Transformation: $Y_t = \eta(X_t) = \int_z^{X_t} \frac{1}{\sigma(u;\theta)} du$

$$dY_t = \alpha(Y_{t-};\theta)dt + dB_t + dJ_t, \quad Y_0 = x.$$

$$J_t = \left\{PP(\lambda(t, \eta^{-1}(Y_t);\theta)); g_1(Z_t, Y_{t-}), f(Z_t;\theta)\right\}$$

General idea: Use Retrospective Rejection Sampling like EA.

1st candidate: $dY_t = dB_t + dJ_t, \quad Y_0 = x, \; Y_T = y$

$J_t = \{PP(\lambda); f\}$

2nd candidate: $Y = J + B^*$

$J_t = \{PP(\lambda); f\}$ and $B^* = BB(0, x; T, y - J_T)$

General idea: Use Retrospective Rejection Sampling like EA.

1st candidate: $dY_t = dB_t + dJ_t, \quad Y_0 = x, \; Y_T = y$

$J_t = \{PP(\lambda); f\}$

2nd candidate: $Y = J + B^*$

$J_t = \{PP(\lambda); f\}$ and $B^* = BB(0, x; T, y - J_T)$

General idea: Use Retrospective Rejection Sampling like EA.

1st candidate: $dY_t = dB_t + dJ_t, \quad Y_0 = x,\ Y_T = y$

$J_t = \{PP(\lambda); f\}$

2nd candidate: $Y = J + B^*$

$J_t = \{PP(\lambda); f\}$ and $B^* = BB(0, x; T, y - J_T)$

Main result

$$\frac{d\tilde{\mathbb{P}}}{d\tilde{\mathbb{D}}} \propto \exp\left\{-\int_0^T \left(\frac{\alpha^2+\alpha'}{2}\right)(Y_{t-};\theta)+\lambda(t,Y_{t-};\theta)dt - \sum_{j=1}^{NJ}\int_{Y_{t_j-}}^{Y_{t_j}} \alpha(u;\theta)du\right\}$$

$$\prod_{j=1}^{NJ}\frac{\lambda(t,Y_{t_j-};\theta)f_g(Y_{t_j}-Y_{t_j-};\theta)}{\lambda f(Y_{t_j}-Y_{t_j-})}$$

$$\exp\left\{-\frac{1}{2T}(y-J_T-x)^2\right\}$$

Main result

$$
\begin{aligned}
\frac{d\tilde{\mathbb{P}}}{d\tilde{\mathbb{D}}} \quad \propto \quad & \exp\left\{-\int_0^T \left(\frac{\alpha^2+\alpha'}{2}\right)(Y_{t-};\theta)+\lambda(t,Y_{t-};\theta)dt \right.\\
& \left. \qquad -\sum_{j=1}^{NJ}\int_{Y_{t_j-}}^{Y_{t_j}}\alpha(u;\theta)du\right\} \\
& \prod_{j=1}^{NJ}\frac{\lambda(t,Y_{t_j-};\theta)f_g(Y_{t_j}-Y_{t_j-};\theta)}{\lambda f(Y_{t_j}-Y_{t_j-})} \\
& \exp\left\{-\frac{1}{2T}(y-J_T-x)^2\right\}
\end{aligned}
$$

Main result

$$\frac{d\tilde{\mathbb{P}}}{d\tilde{\mathbb{D}}} \propto \exp\left\{-\int_0^T \left(\frac{\alpha^2+\alpha'}{2}\right)(Y_{t-};\theta) + \lambda(t, Y_{t-};\theta)dt \right.$$
$$\left. - \sum_{j=1}^{NJ} \int_{Y_{t_j-}}^{Y_{t_j}} \alpha(u;\theta)du \right\}$$
$$\prod_{j=1}^{NJ} \frac{\lambda(t, Y_{t_j-};\theta) f_g(Y_{t_j} - Y_{t_j-};\theta)}{\lambda f(Y_{t_j} - Y_{t_j-})}$$
$$\exp\left\{-\frac{1}{2T}(y - J_T - x)^2\right\}$$

Main result

$$\frac{d\tilde{\mathbb{P}}}{d\tilde{\mathbb{D}}} \propto \exp\left\{-\int_0^T \left(\frac{\alpha^2+\alpha'}{2}\right)(Y_{t-};\theta) + \lambda(t, Y_{t-};\theta)dt - \sum_{j=1}^{NJ}\int_{Y_{t_j-}}^{Y_{t_j}} \alpha(u;\theta)du\right\}$$

$$\prod_{j=1}^{NJ} \frac{\lambda(t, Y_{t_j-};\theta) f_g(Y_{t_j} - Y_{t_j-};\theta)}{\lambda f(Y_{t_j} - Y_{t_j-})}$$

$$\exp\left\{-\frac{1}{2T}(y - J_T - x)^2\right\}$$

Main result

$$\frac{d\tilde{\mathbb{P}}}{d\tilde{\mathbb{D}}} \propto \exp\left\{-\int_0^T \left(\frac{\alpha^2+\alpha'}{2}\right)(Y_{t-};\theta) + \lambda(t, Y_{t-};\theta)dt - \sum_{j=1}^{NJ} \int_{Y_{t_j-}}^{Y_{t_j}} \alpha(u;\theta)du\right\}$$

$$\prod_{j=1}^{NJ} \frac{\lambda(t, Y_{t_j-};\theta) f_g(Y_{t_j} - Y_{t_j-};\theta)}{\lambda f(Y_{t_j} - Y_{t_j-})}$$

$$\exp\left\{-\frac{1}{2T}(y - J_T - x)^2\right\}$$

Content

Log-likelihood

$$
\begin{aligned}
l(Y;\theta) \quad \propto \quad & A(y;\theta) - A(x;\theta) - \sum_{j=1}^{NJ} \int_{Y_{t_j-}}^{Y_{t_j}} \alpha(u;\theta) du \\
& - \int_0^T \left(\frac{\alpha^2 + \alpha'}{2} \right) (Y_{t-};\theta) + \lambda(t, Y_{t-};\theta) dt \\
& \sum_{j=1}^{NJ} \log(\lambda(t, Y_{t_j-};\theta)) + \log(f_g(Y_{t_j} - Y_{t_j-};\theta))
\end{aligned}
$$

$$A(x,\theta) = \int_0^x \alpha(u;\theta)du$$

$$\mathbb{E}_{X_{miss}|X_{obs},\theta'}\left[\int_0^T f(Y_t)dt\right] = \mathbb{E}_{X_{miss},U|X_{obs},\theta'}\left[Tf(Y_U)\right]$$

$$U \sim U(0,T)$$

The expectation of the log-likelihood depends on $\{X_U,\ NJ,\ t_j,\ ,X_{t_j-},\ X_{t_j}\}$.

$$A(x,\theta) = \int_0^x \alpha(u;\theta)du$$

$$\mathbb{E}_{X_{miss}|X_{obs},\theta'}\left[\int_0^T f(Y_t)dt\right] = \mathbb{E}_{X_{miss},U|X_{obs},\theta'}\left[Tf(Y_U)\right]$$

$$U \sim U(0,T)$$

The expectation of the log-likelihood depends on $\{X_U,\ NJ,\ t_j,\ ,X_{t_j-},\ X_{t_j}\}$.

$$A(x, \theta) = \int_0^x \alpha(u; \theta) du$$

$$\mathbb{E}_{X_{miss}|X_{obs},\theta'} \left[\int_0^T f(Y_t) dt \right] = \mathbb{E}_{X_{miss},U|X_{obs},\theta'} \left[T f(Y_U) \right]$$

$$U \sim U(0, T)$$

The expectation of the log-likelihood depends on $\{X_U,\ NJ,\ t_j,\ ,X_{t_j-},\ X_{t_j}\}$.

MCEM Algorithm

1. Choose θ';
2. Perform CJEA M times to obtain M samples of $\{X_U,\ NJ,\ t_j,\ ,X_{t_j-},\ X_{t_j}\}$;
3. Compute the MC estimator of the expectation of the log-likelihood;
4. Maximise the expectation w.r.t. θ;
5. Update θ' and go back to 2.

Other alternatives

- Propose from a modified process to cancel the term $-\sum_{j=1}^{NJ}\int_{Y_{t_j-}}^{Y_{t_j}}\alpha(u;\theta)du$ in the acceptance probability.
- Propose from $\hat{\mathbb{D}}$ and use Importance Sampling (IS) on the E-step.
- Propose from a similar jump-diffusion via CJEA and use IS (embedded Rejection Sampling).

Other alternatives

- Propose from a modified process to cancel the term $-\sum_{j=1}^{NJ}\int_{Y_{t_j-}}^{Y_{t_j}} \alpha(u;\theta)du$ in the acceptance probability.
- Propose from $\hat{\mathbb{D}}$ and use Importance Sampling (IS) on the E-step.
- Propose from a similar jump-diffusion via CJEA and use IS (embedded Rejection Sampling).

Other alternatives

- Propose from a modified process to cancel the term $-\sum_{j=1}^{NJ}\int_{Y_{t_j-}}^{Y_{t_j}}\alpha(u;\theta)du$ in the acceptance probability.

- Propose from $\hat{\mathbb{D}}$ and use Importance Sampling (IS) on the E-step.

- Propose from a similar jump-diffusion via CJEA and use IS (embedded Rejection Sampling).

Other alternatives

- Propose from a modified process to cancel the term $-\sum_{j=1}^{NJ} \int_{Y_{t_j-}}^{Y_{t_j}} \alpha(u;\theta)du$ in the acceptance probability.
- Propose from $\hat{\mathbb{D}}$ and use Importance Sampling (IS) on the E-step.
- Propose from a similar jump-diffusion via CJEA and use IS (embedded Rejection Sampling).

Thank you!

Contact: F.B.Goncalves@warwick.ac.uk

References

- Beskos, A; Papaspiliopoulos, O. and Roberts, G. O. (2006), Retrospective exact simulation of diffusion sample paths with applications. *Bernoulli*, **12**(6), 1077-1098.

- Beskos, A., Papaspiliopoulos, O., Roberts, G, and Fearnhead, P. (2006), Exact and computationally efficient likelihood-based inference for discretely observed diffusion processes (with discussion), *J. R. Stat. Soc. B* **68**(3), 333-382.