Exact Importance Sampling for Affine Processes

Brennan Hall
Supervised by: Alex Shkolnik
bhall@ucsb.edu

June 4, 2019

Department of Statistics & Applied Probability
University of California, Santa Barbara

Setting

Affine processes

For a Borel set $\mathbb{D} \subseteq \mathbb{R}^d$, a homogeneous $\mathbb{D}$-valued Markov process Y on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ is called *affine* if, letting $\mathcal{U} = \{u \in \mathbb{C}^d \mid \sup_{x \in \mathbb{D}} \Re(u \cdot x) < \infty\}$, for every pair $(t, u) \in [0, \infty) \times \mathcal{U}$ there are $\phi(t, u) \in \mathbb{C}$ and $\varphi(t, u) \in \mathbb{C}^d$ s.t.

$$\mathbf{E}_y(e^{u \cdot Y_t}) = e^{\phi(t,u) + \varphi(t,u) \cdot y} \tag{1}$$

where $y \in \mathbb{D}$ is the starting point of Y.

e.g. Brownian motion, Lévy, OU, Bessel, CIR.

Affine structure

An affine jump process Y on state space $\mathbb{D} \subset \mathbb{R}^d$ can be expressed as the solution to an SDE of the form

$$dX_t = a(X_t)dt + b(X_t)dW_t + dJ_t \tag{2}$$

where W is a standard Brownian Motion, J is a marked point process with jump size measure K where a, b and K are affine functions of the form

$$\begin{aligned} a(x) &= a + x_1\alpha^1 + \cdots + x_d\alpha^d \\ b(x) &= b + x_1\beta^1 + \cdots + x_d\beta^d \\ K(x, d\xi) &= m(d\xi) + x_1\mu^1(d\xi) + \cdots + x_d\mu^d(d\xi) \end{aligned} \tag{3}$$

and $a(x)$ is p.s.d. in $\mathbb{R}^{d\times d}$, $b(x) \in \mathbb{R}^d$ and K is a (nondegenerate) Radon measure on $\mathbb{R}^d$ s.t.. $\int_{\mathbb{R}^d}(|\xi|^2 \wedge 1)K(x, d\xi) < \infty$.

Simple example

Consider the strong solution $X \in \mathbb{R}_+$ to the SDE

$$dX_t = \kappa(\mu - X_t)dt + \sigma\sqrt{X_t}\,dW_t + \delta dN_t \tag{4}$$

for $\kappa, \mu, \sigma, \delta, x \in \mathbb{R}_+$ and W, a standard Brownian motion. Arrivals of counting process N occur with an intensity X.

For $n \in \mathbb{N}$ and $T \in [0, \infty)$, compute the tail probability

$$\mathbf{P}(\mathscr{E}_n) = \mathbf{P}(N_T \geq n). \tag{5}$$

Monte Carlo: sample $\mathbf{1}_{\mathscr{E}_n}$ and average.

Monte Carlo

Let $\{\mathscr{E}_n\}_{n\in\mathbb{N}}$ denote a rare-event sequence:

n – rare-event parameter (problem specific),

$\mathbf{P}(\mathscr{E}_n) \to 0$ *as* $n \uparrow \infty$.

The Monte Carlo (MC) estimator $\mathbf{1}_{\mathscr{E}_n}$ is highly inefficient.

(6) **Relative Error:** $$\frac{\sqrt{\mathbf{Var}(\mathbf{1}_{\mathscr{E}_n})}}{\mathbf{P}(\mathscr{E}_n)} \sim 1/\sqrt{\mathbf{P}(\mathscr{E}_n)}$$

Classic Approach

Entails a change of measure from $\mathbf{P}$ to $\mathbf{Q}$.

Let $Z \geq 0$ on $(\Omega, \mathcal{F})$ satisfy $\mathbf{E}(Z) = 1$. Then,

$$\mathbf{Q}(\mathscr{A}) = \mathbf{E}(\mathbf{1}_{\mathscr{A}} Z) \qquad \mathscr{A} \in \mathcal{F}. \tag{7}$$

Given an equivalent of measure, i.e., $Z > 0$,

Sample the estimator $\mathbf{1}_{\mathscr{E}_n}/Z$ *under* $\mathbf{Q}$, *i.e.,*

$$\mathbf{E}_{\mathbf{Q}}(\mathbf{1}_{\mathscr{E}_n}/Z) = \mathbf{P}(\mathscr{E}_n). \tag{8}$$

Exponential change of measure

For $\theta \in \mathbb{R}$, take (Siegmund, 1976)

$$m(\theta) = \mathbf{E}(e^{\theta N_T}) \tag{9}$$

and consider $\mathbf{Q} = Z\mathbf{P}$ for $Z = e^{\theta N_T}/m(\theta)$.

The probability $\mathbf{Q}$ is the (unique) asymptotically optimal importance sampling measure for estimator $\mathbf{1}_{\mathscr{E}_n}/Z$ of $\mathbf{P}(\mathscr{E}_n)$,

> *i.e., per fixed level of precision, the number of importance sampling trials of $\mathbf{1}_{\mathscr{E}_n}/Z$ is subexponential in n (efficient).*

$$\liminf_{n\uparrow\infty} \frac{\log \mathbf{Var}_{\mathbf{Q}}(\mathbf{1}_{\mathscr{E}_n}/Z)}{\log \mathbf{P}(\mathscr{E}_n)} = 2.$$

Transform formula

The process $Y = (N, X)$ is affine $\left(\mathbf{E}(e^{u \cdot Y_t}) = e^{\phi(t,u)+\varphi(t,u)\cdot y}\right)$.

We can compute $m(\theta)$. More generally,

$$\mathbf{E}(e^{\theta(N_T - N_t)} \mid \mathcal{F}_t) = e^{p(t)+q(t)X_t}. \tag{10}$$

where for $p(T) = q(T) = 0$ the (p, q) solve the ODEs

$$\begin{aligned} \dot{p} &= -\kappa\mu q \\ \dot{q} &= \kappa q - \frac{1}{2}\sigma^2 q^2 - (e^{\theta+\delta q} - 1) \end{aligned} \tag{11}$$

How do we sample (N, X) under $\mathbf{Q}$?

ECM dynamics (time-dependence)

The dynamics of $Y = (N, X)$ under $\mathbf{Q}$ are

- *N admits an intensity $e^{\theta+\delta q}X$.*
- *X starts at x and, for B a standard Brownian motion, solves*

$$dX_t = \kappa_q(t)\left(\mu_q(t) - X_t\right)dt + \sigma\sqrt{X_t}\,dB_t + \delta\,dN_t \tag{12}$$

where for the ODE solutions (p, q) we have

$$\kappa_q = \kappa - \sigma^2 q, \qquad \mu_q = \frac{\kappa\mu}{\kappa_q}. \tag{13}$$

The drift is now time dependent! Introduces sampling bias that can materially effect small probability estimates.

Sources of Bias

Two sources of bias in sampling X:

1. Standard method to sample the jump times through a time change of a standard Poisson process. Compensator of N

   $$A_T = \int_0^t X_s ds \tag{14}$$

   is computed by approximation.
2. Time-dependent drift coefficient makes transition law of X unknown under $\mathbf{Q}$.

New Approach

Approximate ECM

Approximate ECM allows for exact sampling of the process $Y = (N, X)$ under $\mathbf{Q}$.

- Define a time grid $0 = t_0 < t_1 < \cdots < t_{K_n} = T$
- Approximate X with step functions X^n along time grid
- Approximate time-dependent variables with step functions $p_n(t)$ and $q_n(t)$ along time grid
- Simulate exactly over each time grid.
- Bias can be corrected

Exact Sampling

Let $q_n(t) = q(t_k)$, $\kappa_k = \kappa_{q_n}(t)$, $\mu_k = \mu_{q_n}(t)$ and $X_t^n = X_{t_k}$ on each $[t_k, t_{k+1})$.

The dynamics of $Y = (N, X)$ under $\mathbf{Q}$ are given by

- N admits an intensity $e^{\theta + \delta q_n} X^n$.
- X starts at x and, for B a standard Brownian motion, solves

$$dX_t = \kappa_k(\mu_k - X_t)dt + \sigma\sqrt{X_t}\,dB_t + \delta dN_t \quad t_k \leq t < t_{k+1}$$

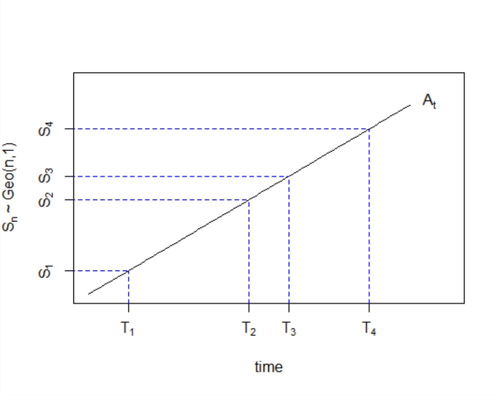
The (N, X) may be sampled exactly under $\mathbf{Q}$ on each $[t_{k-1}, t_k)$.

Sampling jump times

Over each grid interval $[t_{k-1}, t_k)$ N has $\mathbf{Q}$-compensator

$$A_T = \int_0^T X_s^n ds = A_{t_{k-1}} + \int_{t_{k-1}}^T X_s^n ds \tag{15}$$

$T_j = \inf\left\{t > T_{j-1} \,:\, A_t \geq S_j\right\}.$



Bias correction

Bias correction term included in importance sampling measure $\mathbf{Q} = Z_n \mathbf{P}$.

Z_n may be expressed as $Z_n = Q_n Z$ for $Z = e^{\theta N_T}/m(\theta)$ and

$$Q_n = D_n^{ecm} \times D_n^{rate}$$

$$D_n^{ecm} = \exp\left\{ \sum_{k=1}^{K_n} \dot{q}_n(t_{k-1}) \int_{t_{k-1}}^{t_k} X_s\, ds - X_{t_{k-1}}(q_n(t_k) - q_n(t_{k-1})) \right\}$$

$$D_n^{rate} = \exp\left\{ \int_0^T e^{\theta + q_n(s)}(X_s - X_s^n)\, ds \right\} \prod_{j=1}^{N_T} \left(\frac{X_{T_j-}}{X_{T_j-}^n} \right)$$

$Q_n \implies 1$, ie. approximate ECM converges to ECM.

What we have so far

Quick review

- ECM is optimal importance sampling measure - for $Z = e^{\theta N_T}/m(\theta)$

$$\mathbf{Q} = Z\mathbf{P} \tag{16}$$

- Two sources of bias in sampling
  - Approximation required to sample jump times via compensator time change
  - ECM introduces sampling bias for affine jump processes due to time dependent drift
- Approximate ECM allows for exact sampling and preserves (asymptotic) optimality.

Main result

Theorem. The importance measure $\mathbf{Q} = Z_n \mathbf{P}$ is asymptotically optimal for estimating $\mathbf{P}(\mathscr{E}_n)$ and $\mathbf{1}_{\mathscr{E}_n}/Z_n$ may be sampled exactly.

Conditional Monte Carlo

Conditional Monte Carlo method uses

$$\mathbf{E}_{\mathbf{Q}}\left(\mathbf{1}_{\mathscr{E}_n} L_T\right) = \mathbf{E}_{\mathbf{Q}}\left(\mathbf{1}_{\mathscr{E}_n} \mathbf{E}_{\mathbf{Q}}(L_T \,|\, \mathcal{H})\right) \tag{17}$$

$$= \mathbf{E}_{\mathbf{Q}}\left(\mathbf{1}_{\mathscr{E}_n} L_T^{\mathcal{H}}\right) \tag{18}$$

with $\mathcal{H} = \sigma\left(\{N_{\tau_l}, X_{\tau_l}\}\right)$ and $L_T^{\mathcal{H}} = e^{-\theta N_T} m(\theta) G_T \mathbf{E}_{\mathbf{Q}}\left(H_T \,|\, \mathcal{H}\right)$

Define $\{\tau_l\}$ as the sequence of sorted times $\{t_k\} \cup \{T_j\}$.
Sample estimator $L_T^{\mathcal{H}} \mathbf{1}_{\mathcal{E}_n}$:

$$G_T = \exp\left\{ \sum_{k=1}^{K_n} X_{t_{k-1}} \left(e^{\theta + q_n(t_{k-1})} - \left(q_n(t_k) - q_n(t_{k-1}) \right) \right) \right\} \prod_{j=1}^{N_T} \left(\frac{X_{T_j-}}{X^n_{T_j-}} \right)$$

$$H_T = \prod_{l=1}^{K_n+N_T} \mathcal{L}_{\tau_{l-1}, X_{\tau_{l-1}}} \left(\tau_l, X_{\tau_l-} \,|\, \dot{q}_n(\tau_{l-1}) + e^{\theta + q_n(\tau_{l-1})} \right)$$

where $\mathcal{L}_{(s,x)}(t, y \,|\, z) = \mathbf{E}\left(\exp\left\{ z \int_s^t \tilde{X}_z dz \right\} \,|\, \tilde{X}_s, \tilde{X}_t = y \right)$ is the bridge transform.

Summary

Affine models have numerous applications.

The classical exponential change of measure approach leads to biased estimators (leads to a time-dependent drift coefficient).

We develop an approximate ECM scheme which inherits asymptotic optimality properties but allows for exact sampling.

Numerical results currently in progress.