Efficient Simulation for Rare Events

Jose Blanchet

Columbia Department of IEOR
Lunteren Conference 2010

Agenda

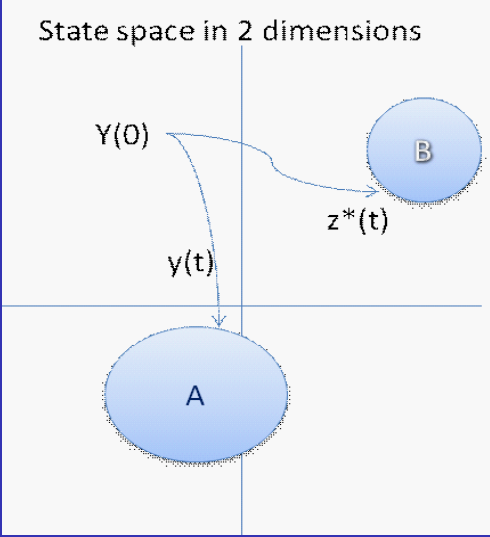
- **Review**
- Stylized Example: Light-tailed Random Walks
- Control Theory, Harmonic Functions and Doob's h-transform
- Lyapunov Inequalities
- Stylized Example: Heavy-tailed Random Walks

Generic Rare-event Simulation Problem

- A generic rare-event estimation problem:

$$P\left(\text{Hit } B \text{ prior to } A\right)$$

Generic Rare-event Simulation Problem



Description of State-independent Approach: Review

- Under suitable light-tailed assumptions:

$$
\begin{aligned}
& \Delta \log P\left(\text{Hit } B \text{ prior to } A\right) \\
\approx \; & -\inf\{J\left(z\right) : z\left(\cdot\right) \text{ is path that hits } B \text{ prior to } A\}
\end{aligned}
$$

Description of State-independent Approach: Review

- Under suitable light-tailed assumptions:

$$
\begin{aligned}
& \Delta \log P\,(\text{Hit } B \text{ prior to } A) \\
\approx\ & -\inf\{J\,(z) : z\,(\cdot) \text{ is path that hits } B \text{ prior to } A\}
\end{aligned}
$$

- $J\,(\cdot)$ is the action function and $z^*\,(\cdot)$ is the "optimal path"

- Under suitable light-tailed assumptions:

$$\begin{aligned} & \Delta \log P\left(\text{Hit } B \text{ prior to } A\right) \\ \approx & -\inf\{J\left(z\right) : z\left(\cdot\right) \text{ is path that hits } B \text{ prior to } A\} \end{aligned}$$

- $J\left(\cdot\right)$ is the action function and $z^{*}\left(\cdot\right)$ is the "optimal path"
- Tracking optimal path apply exponential tilting at time t to follow $\dot{z}^{*}\left(t\right)$

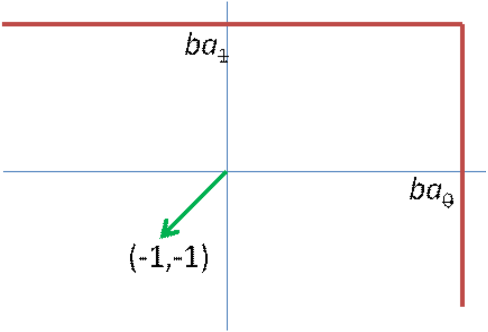
Counter-examples

- **Model:** $Y_t = (-1, -1)t + X_t$; X_t is Brownian motion & $B_b = \{(x, y) : x \geq a_0 b \text{ or } y \geq a_1 b\}$

Counter-examples

- **Model:** $Y_t = (-1, -1)t + X_t$; X_t is Brownian motion & $B_b = \{(x, y) : x \geq a_0 b \text{ or } y \geq a_1 b\}$
- **Estimate:** $u(b) = P(T_{B_b} < \infty)$ as $b \nearrow \infty$

First Passage Time Problem in two dimensions



Heavy-tails...

- **Consider estimating:** $P\left(T_b < \infty\right),\ T_b = \inf\{n \geq 0 : S_n > b\}$

Heavy-tails...

- **Consider estimating:** $P\left(T_b < \infty\right)$, $T_b = \inf\{n \geq 0 : S_n > b\}$
- X_i's i.i.d. reg. varying (heavy tails!) $EX_i < 0$

Heavy-tails...

- **Consider estimating:** $P(T_b < \infty)$, $T_b = \inf\{n \geq 0 : S_n > b\}$
- X_i's i.i.d. reg. varying (heavy tails!) $EX_i < 0$
- Turns out that

$$P(X_j \leq y | T_b < \infty) \longrightarrow P(X_j \leq y)$$

Heavy-tails...

- **Consider estimating:** $P(T_b < \infty)$, $T_b = \inf\{n \geq 0 : S_n > b\}$
- X_i's i.i.d. reg. varying (heavy tails!) $EX_i < 0$
- Turns out that

$$P(X_j \leq y | T_b < \infty) \longrightarrow P(X_j \leq y)$$

- **No clear way to mimic zero variance change-of-measure!**

Heavy-tails...

- **Consider estimating:** $P(T_b < \infty)$, $T_b = \inf\{n \geq 0 : S_n > b\}$
- X_i's i.i.d. reg. varying (heavy tails!) $EX_i < 0$
- Turns out that

$$P(X_j \leq y | T_b < \infty) \longrightarrow P(X_j \leq y)$$

- **No clear way to mimic zero variance change-of-measure!**
- **No clear way to apply the systematic approach!**

Solution?

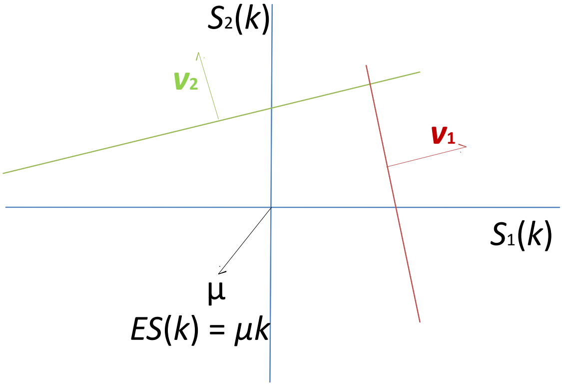
State – dependent importance sampling

Agenda

- Review
- **Stylized Example: Light-tailed Random Walks**
- Control Theory, Harmonic Functions and Doob's h-transform
- Lyapunov Inequalities
- Stylized Example: Heavy-tailed Random Walks

Back to Counter-example in Light-tailed Setting

- Two dimensional random walk
- $A_n = \{s : v_2^T s \geq 1\}$ and $B_n = \{s : v_1^T s \geq 1\}$



- **Efficiently estimate as** $n \nearrow \infty$

$$u_n(0) = P_0[S_k / n \textbf{ hits } A \textbf{ OR } B \textbf{ Eventually}]$$

Elements of Associated Large Deviations

- $S_{\lfloor nt \rfloor} = X_1 + ... + X_{\lfloor nt \rfloor}$, X_k's are i.i.d. with density $f(\cdot)$

Elements of Associated Large Deviations

- $S_{\lfloor nt \rfloor} = X_1 + ... + X_{\lfloor nt \rfloor}$, X_k's are i.i.d. with density $f(\cdot)$
- $W_n(t) = S_{\lfloor nt \rfloor}/n + x$

Elements of Associated Large Deviations

- $S_{\lfloor nt \rfloor} = X_1 + ... + X_{\lfloor nt \rfloor}$, X_k's are i.i.d. with density $f(\cdot)$
- $W_n(t) = S_{\lfloor nt \rfloor}/n + x$
- $Z_k^{(1)} = v_1^T X_k$ and $Z_k^{(2)} = v_2^T X_k$

Elements of Associated Large Deviations

- $S_{\lfloor nt \rfloor} = X_1 + ... + X_{\lfloor nt \rfloor}$, X_k's are i.i.d. with density $f(\cdot)$
- $W_n(t) = S_{\lfloor nt \rfloor}/n + x$
- $Z_k^{(1)} = v_1^T X_k$ and $Z_k^{(2)} = v_2^T X_k$
- Note $EZ_k^{(1)} = v_1^T \mu < 0$ and $EZ_k^{(2)} = v_2^T \mu < 0$

Elements of Associated Large Deviations

- $S_{\lfloor nt \rfloor} = X_1 + ... + X_{\lfloor nt \rfloor}$, X_k's are i.i.d. with density $f(\cdot)$
- $W_n(t) = S_{\lfloor nt \rfloor}/n + x$
- $Z_k^{(1)} = v_1^T X_k$ and $Z_k^{(2)} = v_2^T X_k$
- Note $EZ_k^{(1)} = v_1^T \mu < 0$ and $EZ_k^{(2)} = v_2^T \mu < 0$
- Assume there are $\theta_1^*, \theta_2^* > 0$ such that

$$
\begin{aligned}
E \exp(\theta_1^* Z_k^{(1)}) &= 1 \ \& \ E \exp(\theta_2^* Z_k^{(2)}) = 1 \\
E[\exp(\theta_1^* Z_k^{(1)}) Z_k^{(1)}] &< \infty \ \& \ E[\exp(\theta_2^* Z_k^{(2)}) Z_k^{(2)}] < \infty
\end{aligned}
$$

Large Deviations for the Stylized Example

- As $n \nearrow \infty$

$$
\begin{aligned}
u_n(x) &= P_x[W_n(t) \ \textbf{hits } A \ \textbf{OR } B] \\
&\sim c_1 \exp(-n\theta_1^*(1 - v_1^T x)) + c_2 \exp(-n\theta_2^*(1 - v_2^T x)) \\
&= \exp(-nh(x) + o(n)),
\end{aligned}
$$

where

$$
h(x) = \min[\theta_1^*(1 - v_1^T x), \theta_2^*(1 - v_2^T x)].
$$

Agenda

- Review
- Stylized Example: Light-tailed Random Walks
- **Control Theory, Harmonic Functions and Doob's h-transform**
- Lyapunov Inequalities
- Stylized Example: Heavy-tailed Random Walks

Isaacs Equation for I.S. (Dupuis - Wang)

- Let $\psi(\lambda) = \log E \exp\left(\lambda^T X_k\right)$

Isaacs Equation for I.S. (Dupuis - Wang)

- Let $\psi(\lambda) = \log E \exp\left(\lambda^T X_k\right)$
- $f_\lambda(x) = \exp\left(\lambda^T x - \psi(\lambda)\right) f(x)$ <– Controls

Isaacs Equation for I.S. (Dupuis - Wang)

- Let $\psi(\lambda) = \log E \exp\left(\lambda^T X_k\right)$
- $f_\lambda(x) = \exp\left(\lambda^T x - \psi(\lambda)\right) f(x)$ <– Controls
- $I(z) = \max_\lambda[\lambda^T z - \psi(\lambda)]$ <– useful for large deviations

Isaacs Equation for I.S. (Dupuis - Wang)

- Let $\psi(\lambda) = \log E \exp\left(\lambda^T X_k\right)$
- $f_\lambda(x) = \exp\left(\lambda^T x - \psi(\lambda)\right) f(x)$ <– Controls
- $I(z) = \max_\lambda[\lambda^T z - \psi(\lambda)]$ <– useful for large deviations
- HJB eqn. to minimize 2nd moment...

$$C_n(w) = \min_\lambda E[e^{-\lambda^T X + \psi(\lambda)} C_n(w + X/n)]$$

Isaacs Equation for I.S. (Dupuis - Wang)

- Let $\psi(\lambda) = \log E \exp\left(\lambda^T X_k\right)$
- $f_\lambda(x) = \exp\left(\lambda^T x - \psi(\lambda)\right) f(x)$ <– Controls
- $I(z) = \max_\lambda[\lambda^T z - \psi(\lambda)]$ <– useful for large deviations
- HJB eqn. to minimize 2nd moment...

$$C_n(w) = \min_\lambda E[e^{-\lambda^T X + \psi(\lambda)} C_n(w + X/n)]$$

- $C_n(w) \approx \exp(-ng(w))$

$$\begin{aligned} 0 &\approx \min_\lambda \log E[e^{-\lambda^T X + \psi(\lambda) - n[g(w+X/n) - g(w)]}] \\ &\approx \min_\lambda \max_\beta [-\beta^T(\lambda + \partial g(w)) + \psi(\lambda) - I(\beta)] \end{aligned}$$

Isaacs Equation for I.S. (Dupuis - Wang)

- Let $\psi(\lambda) = \log E \exp\left(\lambda^T X_k\right)$
- $f_\lambda(x) = \exp\left(\lambda^T x - \psi(\lambda)\right) f(x)$ <– Controls
- $I(z) = \max_\lambda[\lambda^T z - \psi(\lambda)]$ <– useful for large deviations
- HJB eqn. to minimize 2nd moment...

$$C_n(w) = \min_\lambda E[e^{-\lambda^T X + \psi(\lambda)} C_n(w + X/n)]$$

- $C_n(w) \approx \exp(-ng(w))$

$$\begin{aligned} 0 &\approx \min_\lambda \log E[e^{-\lambda^T X + \psi(\lambda) - n[g(w+X/n) - g(w)]}] \\ &\approx \min_\lambda \max_\beta [-\beta^T(\lambda + \partial g(w)) + \psi(\lambda) - I(\beta)] \end{aligned}$$

- Game interpretation —> explains Isaacs equation name...

Explicit Solution to the HJB Equation

- HJB eqn. to minimize 2nd moment...

$$C_n(w) = \min_{\lambda} E[e^{-\lambda^T X + \psi(\lambda)} C_n(w + X/n)]$$

Explicit Solution to the HJB Equation

- HJB eqn. to minimize 2nd moment...

$$C_n(w) = \min_{\lambda} E[e^{-\lambda^T X+\psi(\lambda)} C_n(w + X/n)]$$

- $C_n(w) \approx \exp(-ng(w))$

$$\begin{aligned} 0 &\approx \min_{\lambda} \log E[e^{-\lambda^T X+\psi(\lambda)-n[g(w+X/n)-g(w)]}] \\ &\approx \min_{\lambda} \log E[e^{-\lambda^T X+\psi(\lambda)-\partial g(w)\cdot X}] \\ &= \min_{\lambda}[\psi(\lambda) + \psi(-\lambda - \partial g(w))] \end{aligned}$$

Explicit Solution to the HJB Equation

- HJB eqn. to minimize 2nd moment...

$$C_n(w) = \min_{\lambda} E[e^{-\lambda^T X + \psi(\lambda)} C_n(w + X/n)]$$

- $C_n(w) \approx \exp(-ng(w))$

$$\begin{aligned} 0 &\approx \min_{\lambda} \log E[e^{-\lambda^T X + \psi(\lambda) - n[g(w+X/n) - g(w)]}] \\ &\approx \min_{\lambda} \log E[e^{-\lambda^T X + \psi(\lambda) - \partial g(w) \cdot X}] \\ &= \min_{\lambda} [\psi(\lambda) + \psi(-\lambda - \partial g(w))] \end{aligned}$$

- **Solution:** $\lambda = -\partial g(w)/2$ and $\psi(-\partial g(w)/2) = 0$ subject to $g(w) = I(w \in A \cup B)$

Harmonic Functions

- $u_n(w) = 1$ on $A \cup B$ & harmonic:

$$u_n(w) = P_w(T_{A\cup B} < \infty) = E[u_n(w + X/n)]$$

Harmonic Functions

- $u_n(w) = 1$ on $A \cup B$ & harmonic:

$$u_n(w) = P_w(T_{A\cup B} < \infty) = E[u_n(w + X/n)]$$

- Conditioning on $T_{A\cup B} < \infty$ (Doob's h-transform):

$$P^*(X_{k+1} \in dy | S_k = nw) = f(y)\frac{u_n(w + X/n)}{u_n(w)}$$

Harmonic Functions

- $u_n(w) = 1$ on $A \cup B$ & harmonic:

$$u_n(w) = P_w(T_{A\cup B} < \infty) = E[u_n(w + X/n)]$$

- Conditioning on $T_{A\cup B} < \infty$ (Doob's h-transform):

$$P^*(X_{k+1} \in dy | S_k = nw) = f(y)\frac{u_n(w + X/n)}{u_n(w)}$$

- Constant likelihood ratio $\Longrightarrow$ zero variance

$$u_n(S_0/n) = \frac{u_n(S_0/n)}{u_n(S_1/n)} \times \frac{u_n(S_1/n)}{u_n(S_2/n)} \times \ldots \times \frac{u_n(S_{T_{AUB}-1}/n)}{u_n(S_{T_{AUB}}/n)}$$

Isaacs Equation & Harmonic Functions

- Zero-variance sampler

$$P^*\left(X_{k+1} \in dy \middle| S_k = nw\right) = f\left(y\right) \frac{u_n\left(w + y/n\right)}{u_n\left(w\right)}$$

Isaacs Equation & Harmonic Functions

- Zero-variance sampler

$$P^{*}\left(X_{k+1}\in dy\middle|\,S_{k}=nw\right)=f\left(y\right)\frac{u_{n}\left(w+y/n\right)}{u_{n}\left(w\right)}$$

- Approximate sampler $u_{n}\left(w\right)\approx\exp\left(-nh\left(w\right)\right)$

$$\begin{aligned}&\widetilde{P}\left(Y_{k+1}\in dy\middle|\,S_{k}=nw\right)\\ \approx\;& f\left(y\right)\exp\left(-n[h\left(w+y/n\right)-h\left(w\right)]\right)\\ \approx\;& f\left(y\right)\exp\left(-\partial h\left(w\right)\cdot y\right)\end{aligned}$$

Isaacs Equation & Harmonic Functions

- Zero-variance sampler

$$P^{*}\left(X_{k+1}\in dy\middle| S_{k}=nw\right) =f\left(y\right) \frac{u_{n}\left(w+y/n\right) }{u_{n}\left(w\right) }$$

- Approximate sampler $u_{n}\left(w\right) \approx \exp \left(-nh\left(w\right) \right)$

$$\begin{aligned} & \widetilde{P}\left(Y_{k+1}\in dy\middle| S_{k}=nw\right) \\ \approx \quad & f\left(y\right) \exp \left(-n[h\left(w+y/n\right) -h\left(w\right)]\right) \\ \approx \quad & f\left(y\right) \exp \left(-\partial h\left(w\right) \cdot y\right) \end{aligned}$$

- But

$$1=\int \widetilde{P}\left(X_{k+1}\in dy\middle| S_{k}=nw\right) \Longrightarrow \psi \left(-\partial h\left(w\right) \right) =0$$

Isaacs Equation & Harmonic Functions

- Zero-variance sampler

$$P^{*}\left(X_{k+1}\in dy\middle| S_{k}=nw\right) =f\left(y\right) \frac{u_{n}\left(w+y/n\right) }{u_{n}\left(w\right) }$$

- Approximate sampler $u_{n}\left(w\right) \approx \exp \left(-nh\left(w\right) \right)$

$$\begin{aligned} & \widetilde{P}\left(Y_{k+1}\in dy\middle| S_{k}=nw\right) \\ \approx \; & f\left(y\right) \exp \left(-n[h\left(w+y/n\right) -h\left(w\right)]\right) \\ \approx \; & f\left(y\right) \exp \left(-\partial h\left(w\right) \cdot y\right) \end{aligned}$$

- But

$$1=\int \widetilde{P}\left(X_{k+1}\in dy\middle| S_{k}=nw\right) \Longrightarrow \psi \left(-\partial h\left(w\right) \right) =0$$

- **Equivalent to Isaacs equation with** $g\left(w\right) =2h\left(w\right)$ **–> best asymptotic rate**

Dynamic Programming, Isaacs Equation, Harmonic Functions: Summary

Agenda

- Review
- Stylized Example: Light-tailed Random Walks
- Control Theory, Harmonic Functions and Doob's h-transform
- **Lyapunov Inequalities**
- Stylized Example: Heavy-tailed Random Walks

The Second Moment of a State-dependent Estimator

- Consider sampler

$$P^{Q}\left(X_{k+1}\in dy \middle| S_{k}=nw\right) =r^{-1}\left(w,w+X/n\right) f\left(y\right) dy$$

The Second Moment of a State-dependent Estimator

- Consider sampler

$$P^{Q}\left(X_{k+1}\in dy\right| S_{k}=nw) =r^{-1}\left(w,w+X/n\right) f\left(y\right) dy$$

- Likelihood ratio

$$r\left(W_{n}(0),W_{n}\left(1/n\right) \right) ...r(W_{n}(T_{A\cup B}-1),W_{n}(T_{A\cup B}))$$

The Second Moment of a State-dependent Estimator

- Consider sampler

$$P^{Q}\left(X_{k+1}\in dy|\,S_k=nw\right)=r^{-1}\left(w,w+X/n\right)f\left(y\right)dy$$

- Likelihood ratio

$$r\left(W_n(0),W_n\left(1/n\right)\right)...r(W_n(T_{A\cup B}-1),W_n(T_{A\cup B}))$$

- Second moment of estimator

$$s(w)=E[r(w,w+X/n)s(w+X/n)]$$

subject to $s(w)=1$ for $w\in A\cup B$.

The Lyapunov Inequality

Lemma

B. & Glynn '08: Lyapunov inequality

$$v(w) \geq E[r(w, w + X/n)v(w + X/n)]$$

subject to $v(w) \geq 1$ *for* $w \in A \cup B$. *Then,* $v(w) \geq s(w)$.

- **How to use the result?** 1) Identify a change-of-measure, 2) use heuristic / approx. to force $v(w) \approx u_n(w)^2$.

Lyapunov Inequalities and Subsolutions

- Subsolutions introduced by Dupuis & Wang '07

Lyapunov Inequalities and Subsolutions

- Subsolutions introduced by Dupuis & Wang '07
- Connections to Lyapunov inequalities B. & Glynn '08

Lyapunov Inequalities and Subsolutions

- Subsolutions introduced by Dupuis & Wang '07
- Connections to Lyapunov inequalities B. & Glynn '08
- Lyapunov function $v(w) = \exp(-ng(w))$ & $\lambda = -\partial g(w)/2$

$$1 \geq E[\exp(-\lambda^T X + \psi(\lambda)) \exp(-n[g(w + X/n) - g(w)])]$$

subject to $g(w) \leq 0$ for $w \in A \cup B$. Then, $v(w) \geq s(w)$.

Lyapunov Inequalities and Subsolutions

- Subsolutions introduced by Dupuis & Wang '07
- Connections to Lyapunov inequalities B. & Glynn '08
- Lyapunov function $v(w) = \exp(-ng(w))$ & $\lambda = -\partial g(w)/2$

$$1 \geq E[\exp(-\lambda^T X + \psi(\lambda)) \exp(-n[g(w + X/n) - g(w)])]$$

  subject to $g(w) \leq 0$ for $w \in A \cup B$. Then, $v(w) \geq s(w)$.
- Expanding as $n \nearrow \infty$ we get

$$1 + O(1/n) \geq \exp[2\psi(-\partial g(w)/2)]$$

Lyapunov Inequalities and Subsolutions

- Subsolutions introduced by Dupuis & Wang '07
- Connections to Lyapunov inequalities B. & Glynn '08
- Lyapunov function $v(w) = \exp(-ng(w))$ & $\lambda = -\partial g(w)/2$

$$1 \geq E[\exp(-\lambda^T X + \psi(\lambda)) \exp(-n[g(w + X/n) - g(w)])]$$

subject to $g(w) \leq 0$ for $w \in A \cup B$. Then, $v(w) \geq s(w)$.

- Expanding as $n \nearrow \infty$ we get

$$1 + O(1/n) \geq \exp[2\psi(-\partial g(w)/2)]$$

- Yields subsolution to the Isaacs equation (note smoothness)

$$\psi(-\partial g(w)/2) \leq 0 \text{ s.t. } g(w) \leq 0,\ w \in A \cup B$$

The Mollification

- Back to random walk example

$$
\begin{aligned}
h(w) &= \min[\theta_1^*(1 - v_1^T w), \theta_2^*(1 - v_2^T w)] \\
&= -\max[\theta_1^*(v_1^T w - 1), \theta_2^*(v_2^T w - 1)]
\end{aligned}
$$

NOT smooth...

The Mollification

- Back to random walk example

$$
\begin{aligned}
h(w) &= \min[\theta_1^*(1 - v_1^T w), \theta_2^*(1 - v_2^T w)] \\
&= -\max[\theta_1^*(v_1^T w - 1), \theta_2^*(v_2^T w - 1)]
\end{aligned}
$$

NOT smooth...

- Mollification:

$$
\begin{aligned}
& h_\varepsilon(w) \\
= \ & -\varepsilon \log[\exp(\theta_1^*(v_1^T w - 1)/\varepsilon) + \exp(\theta_2^*(v_2^T w - 1)/\varepsilon)]
\end{aligned}
$$

The Mollification

- Back to random walk example

$$
\begin{aligned}
h(w) &= \min[\theta_1^*(1 - v_1^T w), \theta_2^*(1 - v_2^T w)] \\
&= -\max[\theta_1^*(v_1^T w - 1), \theta_2^*(v_2^T w - 1)]
\end{aligned}
$$

NOT smooth...

- Mollification:

$$
\begin{aligned}
& h_\varepsilon(w) \\
= \ & -\varepsilon \log[\exp(\theta_1^*(v_1^T w - 1)/\varepsilon) + \exp(\theta_2^*(v_2^T w - 1)/\varepsilon)]
\end{aligned}
$$

- Implementation via mixtures:

$$
\begin{aligned}
-\partial h_\varepsilon(w) &= \theta_1^* v_1^T \frac{\eta_1^\varepsilon(w)}{\eta_1^\varepsilon(w) + \eta_2^\varepsilon(w)} + \theta_2^* v_2^T \frac{\eta_2^\varepsilon(w)}{\eta_1^\varepsilon(w) + \eta_2^\varepsilon(w)}, \\
\eta_1^\varepsilon(w) &= \exp(\theta_1^*(v_1^T w - 1)/\varepsilon), \\
\eta_2^\varepsilon(w) &= \exp(\theta_2^*(v_2^T w - 1)/\varepsilon).
\end{aligned}
$$

The Form of a Typical Optimality Statement...

Theorem (Dupuis & Wang '07)

Let $g_{\varepsilon_n}(w) = 2h_{\varepsilon_n}(w)$ *and assume that* $n\varepsilon_n \longrightarrow \infty$ *apply corresponding sampler. Then,*

$$2nd\ Moment\ of\ Est. = \exp(-2nh(w) + o(n)).$$

Lyapunov Inequalities

- Select $\varepsilon_n = 1/n$

$$\begin{aligned} \eta_1(w) &= \exp(n\theta_1^*(v_1^T w - 1)) \\ \eta_2(w) &= \exp(n\theta_2^*(v_2^T w - 1)) \end{aligned}$$

Lyapunov Inequalities

- Select $\varepsilon_n = 1/n$

$$\begin{aligned} \eta_1(w) &= \exp(n\theta_1^*(v_1^T w - 1)) \\ \eta_2(w) &= \exp(n\theta_2^*(v_2^T w - 1)) \end{aligned}$$

- Mixture sampler from density $\widetilde{f}(x)$

$$\frac{\widetilde{f}(x)}{f(x)} = \frac{\eta_1(w)}{\eta_1(w) + \eta_2(w)} \exp\left(\theta_1^* v_1^T x\right) + \frac{\eta_2(w)}{\eta_1(w) + \eta_2(w)} \exp\left(\theta_2^* v_2^T x\right)$$

Lyapunov Inequalities

- Select $\varepsilon_n = 1/n$

$$\begin{aligned}\eta_1(w) &= \exp(n\theta_1^*(v_1^T w - 1)) \\ \eta_2(w) &= \exp(n\theta_2^*(v_2^T w - 1))\end{aligned}$$

- Mixture sampler from density $\widetilde{f}(x)$

$$\frac{\widetilde{f}(x)}{f(x)} = \frac{\eta_1(w)}{\eta_1(w) + \eta_2(w)} \exp\left(\theta_1^* v_1^T x\right) + \frac{\eta_2(w)}{\eta_1(w) + \eta_2(w)} \exp\left(\theta_2^* v_2^T x\right)$$

- Lyapunov function

$$v(w) = (\eta_1(w) + \eta_2(w))^2 \geq 1$$

for $v_1^T w \geq 1$ OR $v_2^T w \geq 1$... BOUNDARY CONDITION OK!

A Lyapunov Inequality

- $$\begin{aligned} v(w) &= [\eta_1(w) + \eta_2(w)]^2 \\ \eta_1(w + X/n) &= \eta_1(w)\, e^{\theta^* v_1^T X} \\ \eta_2(w + X/n) &= \eta_2(w)\, e^{\theta^* v_2^T X} \end{aligned}$$

A Lyapunov Inequality

- $$\begin{aligned} v(w) &= [\eta_1(w) + \eta_2(w)]^2 \\ \eta_1(w + X/n) &= \eta_1(w)\, e^{\theta^* v_1^T X} \\ \eta_2(w + X/n) &= \eta_2(w)\, e^{\theta^* v_2^T X} \end{aligned}$$

- $$\begin{aligned} & E \frac{v(w + X/n)}{v(w)} \frac{1}{\frac{\eta_1(w)}{\eta_1(w)+\eta_2(w)} e^{\theta_1^* v_1^T X} + \frac{\eta_2(w)}{\eta_1(w)+\eta_2(w)} e^{\theta_2^* v_2^T X}} \\ = \; & E \frac{\eta_1(w)\exp\left(\theta_1^* v_1^T X\right) + \eta_2(w)\exp\left(\theta_2^* v_2^T X\right)}{\eta_1(w) + \eta_2(w)} = 1. \end{aligned}$$

Theorem (B., Glynn and Leder (2009))

One can take $\varepsilon = 1/n$ as mollification parameter & in fact this is the optimal choice as it gives bounded coef. of variation

$$2nd\ Moment \leq (v_1(0) + v_2(0))^2 \leq c u_n(0)^2 .$$

Agenda

- Review
- Stylized Example: Light-tailed Random Walks
- Control Theory, Harmonic Functions and Doob's h-transform
- Lyapunov Inequalities
- **Stylized Example: Heavy-tailed Random Walks**

Setup

- Let $X_1, X_2, \ldots$ are heavy-tailed (TBD) and $EX_i = \mu < 0$
- $S_n = X_1 + \ldots + X_n$ given $S_0 = s$
- Object of interest:

$$u_b(s) = P_s(T_b < \infty) = \frac{\int_{b-s}^{\infty} P(X_i > u)\, du}{-\mu} (1 + o(1))$$

as $b - s \nearrow \infty$.

- Asymptotics: Pakes, Veraberbeke, Cohen... see text of Asmussen '03

Basics on Heavy-tails

- Rich theory for heavy-tailed random walks based on *subexponentiality*

$$P\left(X_1 + X_2 > b\right) = 2P\left(X_1 > b\right)\left(1 + o\left(1\right)\right)$$

as $b \longrightarrow \infty$.

- *Focus on regularly varying distributions* (basically *power-law* type)

$$P\left(X_1 > t\right) = t^{-\alpha}L\left(t\right)$$

for $\alpha > 1$ and $L\left(t\beta\right)/L\left(t\right) \longrightarrow 1$ as $t \nearrow \infty$ for each $\beta > 0$.

State-dependent Importance Sampling

- **Importance sampling:** Markov kernel $K(\cdot)$

$$K(s_0, s_1) = r^{-1}(s_0, s_1) f_{X_1}(s_1 - s_0)$$

State-dependent Importance Sampling

- **Importance sampling:** Markov kernel $K(\cdot)$

$$K(s_0, s_1) = r^{-1}(s_0, s_1) f_{X_1}(s_1 - s_0)$$

- **Importance sampling estimator:** S_n's simulated under $K(\cdot)$

$$Z = \prod_{j=1}^{T_b - 1} r(S_j, S_{j+1}) I(T_b < \infty),$$

State-dependent Importance Sampling

- **Importance sampling:** Markov kernel $K(\cdot)$

$$K(s_0, s_1) = r^{-1}(s_0, s_1) f_{X_1}(s_1 - s_0)$$

- **Importance sampling estimator:** S_n's simulated under $K(\cdot)$

$$Z = \prod_{j=1}^{T_b - 1} r(S_j, S_{j+1}) I(T_b < \infty),$$

- **Recall to get good variance:** *Select an IS that mimics the conditional distribution.*

Description of the Conditional Distribution

Theorem

(Asmussen and Kluppelberg): Conditional on $T_b < \infty$, we have that

$$\left(\frac{S_{uT_b}}{T_b}, \frac{S_{T_b} - b}{b}, \frac{T_b}{b} \right) \Longrightarrow (\mu u, Z_1, Z_2)$$

on $D(0,1) \times R \times R$ as $b \nearrow \infty$, where Z_1 and Z_2 are Pareto with index $\alpha - 1$.

- **Interpretation**: Prior to ruin, random walk has drift μ and a large jump of size b occurs suddenly in $O(b)$ time...

Description of the Conditional Distribution

Theorem

(Asmussen and Kluppelberg): Conditional on $T_b < \infty$, we have that

$$\left(\frac{S_{uT_b}}{T_b}, \frac{S_{T_b} - b}{b}, \frac{T_b}{b} \right) \Longrightarrow (\mu u, Z_1, Z_2)$$

on $D(0,1) \times R \times R$ as $b \nearrow \infty$, where Z_1 and Z_2 are Pareto with index $\alpha - 1$.

- **Interpretation**: Prior to ruin, random walk has drift μ and a large jump of size b occurs suddenly in $O(b)$ time...
- So, given that a jump hasn't occurred by time k, then $S_k \approx \mu k$ and the chance of reaching b in the next increment *given* that we eventually reach b ($T_b < \infty$)

$$\frac{P(X > b - \mu k)}{\int_0^\infty P(X > b - \mu u)\, du} \approx \frac{-\mu P(X > b - \mu k)}{\int_b^\infty P(X > s)\, du} = O\left(\frac{1}{b}\right).$$

Good Family of Changes of Measure for Heavy Tails

- **Change-of-measure:** Here s is current position of the walk, f is the density

$$
\begin{aligned}
f_{X|s}(x|\,s) &= p(s)\,\frac{f_X(x)\,I(x>a(b-s))}{P(X>a(b-s))} \\
&\quad + (1-p(s))\,\frac{f_X(x)\,I(x\le a(b-s))}{P(X>a(b-s))}
\end{aligned}
$$

- In other words, $s_0 = s$ and $s_1 = s_0 + x$

$$
\begin{aligned}
r(s_0, s_1)^{-1} &= p(s_0)\,\frac{I(s_1-s_0>a(b-s_0))}{P(X>a(b-s_0))} \\
&\quad + (1-p(s_1))\,\frac{I(s_1-s_0\le a(b-s_0))}{P(X\le a(b-s_0))}
\end{aligned}
$$

- Introduced by Dupuis, Leder and Wang '06 for finite sums...

Lyapunov Inequality

- **Recall Lyapunov inequality**

Lemma (B. & Glynn '08)

Suppose that there is a positive function $g(\cdot)$ *such that*

$$E_s^K\left(\frac{g(S_1)\,r(s,S_1)^2}{g(s)}\right)=E_s\left(\frac{g(S_1)\,r(s,S_1)}{g(s)}\right)\leq 1$$

for all $s\leq b$ *and* $g(s)\geq 1$ *for* $s>b$. *Then,*

$$E_s^K Z^2=E_s^K\left(\prod_{j=1}^{T_b-1} r(S_j,S_{j+1})^2\, I(T_b<\infty)\right)\leq g(s).$$

Constructing Lyapunov Functions for Heavy-tails

- Wish to achieve *strong efficiency*, so we pick (for some $\kappa > 0$)

$$g(s) = \min\left(\kappa\left(\int_{b-s}^{\infty} P(X > u)\,du\right)^2, 1\right).$$

Constructing Lyapunov Functions for Heavy-tails

- Wish to achieve *strong efficiency*, so we pick (for some $\kappa > 0$)

$$g(s) = \min\left(\kappa\left(\int_{b-s}^{\infty} P(X > u)\, du\right)^2, 1\right).$$

- Pick for some $\theta > 0$

$$p(s) = \theta \frac{P(X > b - s)}{\int_{b-s}^{\infty} P(X > s)\, du}$$

Constructing Lyapunov Functions for Heavy-tails

- Wish to achieve *strong efficiency*, so we pick (for some $\kappa > 0$)

$$g(s) = \min\left(\kappa\left(\int_{b-s}^{\infty} P(X > u)\, du\right)^2, 1\right).$$

- Pick for some $\theta > 0$

$$p(s) = \theta \frac{P(X > b - s)}{\int_{b-s}^{\infty} P(X > s)\, du}$$

- **Must select κ and θ to verify Lyapunov inequality**

Testing the Lyapunov Inequality

- **Testing the Inequality on $g(s) < 1$ (note that $g \leq 1$):**

$$
\begin{aligned}
& E_s\left(\frac{g(S_1)\, r(s, S_1)}{g(s)}\right) \\
= \; & \frac{E(g(s+X); X > a(b-s))\, P(X > a(b-s))}{p(s)\, g(s)} \\
& + \frac{E(g(s+X); X \leq a(b-s))\, P(X \leq a(b-s))}{(1-p(s))\, g(s)}
\end{aligned}
$$

- **Testing the Inequality on $g(s) < 1$ (note that $g \leq 1$):**

$$
\begin{aligned}
& E_s\left(\frac{g(S_1)\, r(s, S_1)}{g(s)}\right) \\
= \quad & \frac{E(g(s+X); X > a(b-s))\, P(X > a(b-s))}{p(s)\, g(s)} \\
& + \frac{E(g(s+X); X \leq a(b-s))\, P(X \leq a(b-s))}{(1-p(s))\, g(s)} \\
\leq \quad & \frac{P(X > a(b-s))^2}{p(s)\, g(s)} + \frac{E(g(s+X); X \leq a(b-s))}{(1-p(s))\, g(s)}
\end{aligned}
$$

Testing the Lyapunov Inequality

- **Testing the Inequality on $g(s) < 1$ (note that $g \leq 1$):**

$$
\begin{aligned}
& E_s\left(\frac{g(S_1)\, r(s, S_1)}{g(s)}\right) \\
\leq \quad & \frac{P(X > a(b-s))^2}{p(s)\, g(s)} + \frac{E(g(s+X); X \leq a(b-s))}{(1-p(s))\, g(s)} \\
\approx \quad & \frac{a^{-\alpha} P(X > a(b-s))}{\theta\kappa \int_{b-s}^{\infty} P(X > u)\, du} + 1 + 2(\mu+\theta) \frac{P(X > (b-s))}{\left(\int_{b-s}^{\infty} P(X > u)\, du\right)}
\end{aligned}
$$

- **NOTE:** crucial that $\mu < 0$! Pick θ small and κ large

$$
\frac{a^{-\alpha}}{\theta\kappa} + 2\theta + 2\mu \leq 0
$$

Summary and Conclusions

- State-dependent Importance Sampling: Choose RIGHT parametric family of distributions (or controls)

Summary and Conclusions

- State-dependent Importance Sampling: Choose RIGHT parametric family of distributions (or controls)
- Find subsolutions / Lyapunov inequalities to enforce optimality (involves course approximation)

Summary and Conclusions

- State-dependent Importance Sampling: Choose RIGHT parametric family of distributions (or controls)
- Find subsolutions / Lyapunov inequalities to enforce optimality (involves course approximation)
- Lyapunov inequalities apply to light and heavy tails

Summary and Conclusions

- State-dependent Importance Sampling: Choose RIGHT parametric family of distributions (or controls)
- Find subsolutions / Lyapunov inequalities to enforce optimality (involves course approximation)
- Lyapunov inequalities apply to light and heavy tails
- Construct Lyapunov function from course large deviations analysis

Summary and Conclusions

- State-dependent Importance Sampling: Choose RIGHT parametric family of distributions (or controls)
- Find subsolutions / Lyapunov inequalities to enforce optimality (involves course approximation)
- Lyapunov inequalities apply to light and heavy tails
- Construct Lyapunov function from course large deviations analysis
- Lyapunov inequalities guide selection of mollification parameters and guarantee good performance

References

- Blanchet, J., Glynn, P. and Liu, J. C. (2007) Fluid heuristics, Lyapunov bounds and efficient importance sampling for a heavy-tailed G/G/1 queue. *QUESTA*. 57, 99-113.
- Blanchet, J. and Glynn, P. (2008) Efficient rare-event simulation for the maximum of a heavy-tailed random walk. *Ann. of Appl. Probab.* 18, 1351-1378.
- Dupuis, P. and Wang, H. (2004) Importance sampling, large deviations, and differential games. *Stoch. and Stoch. Reports*. 76, 481-508.
- Dupuis, P. and Wang, H. (2007) Subsolutions of an Isaacs equation and efficient schemes of importance sampling. *Mathematics of Operations Research*. 32, 723-757.