Connections between Lyapunov inequalities and Subsolutions of an Isaacs equation for importance sampling

(joint work with P. Glynn and K. Leder)

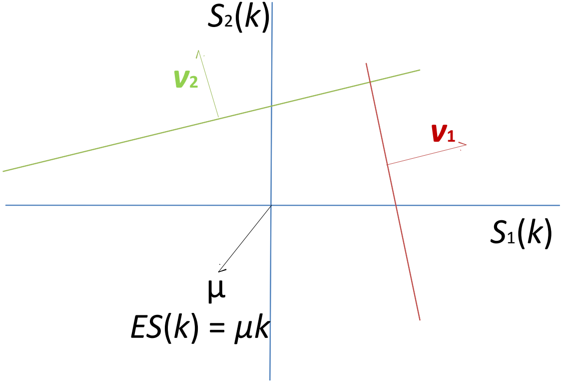
Jose Blanchet

Columbia Department of IEOR

Outline

Stylized Example

- State-space of a 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** $n \nearrow \infty$

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

Stylized Example

- $S_{\lfloor nt \rfloor} = Y_1 + ... + Y_{\lfloor nt \rfloor}$, Y_k's are i.i.d. with density $f(\cdot)$

Stylized Example

- $S_{\lfloor nt \rfloor} = Y_1 + ... + Y_{\lfloor nt \rfloor}$, Y_k's are i.i.d. with density $f(\cdot)$
- $W_n(t) = S_{\lfloor nt \rfloor}/n + x$

Stylized Example

- $S_{\lfloor nt \rfloor} = Y_1 + ... + Y_{\lfloor nt \rfloor}$, Y_k's are i.i.d. with density $f(\cdot)$
- $W_n(t) = S_{\lfloor nt \rfloor}/n + x$
- $Z_k^{(1)} = v_1^T Y_k$ and $Z_k^{(2)} = v_2^T Y_k$

Stylized Example

- $S_{\lfloor nt \rfloor} = Y_1 + ... + Y_{\lfloor nt \rfloor}$, Y_k's are i.i.d. with density $f(\cdot)$
- $W_n(t) = S_{\lfloor nt \rfloor}/n + x$
- $Z_k^{(1)} = v_1^T Y_k$ and $Z_k^{(2)} = v_2^T Y_k$
- Note $EZ_k^{(1)} = v_1^T \mu < 0$ and $EZ_k^{(2)} = v_2^T \mu < 0$

Stylized Example

- $S_{\lfloor nt \rfloor} = Y_1 + ... + Y_{\lfloor nt \rfloor}$, Y_k's are i.i.d. with density $f(\cdot)$
- $W_n(t) = S_{\lfloor nt \rfloor}/n + x$
- $Z_k^{(1)} = v_1^T Y_k$ and $Z_k^{(2)} = v_2^T Y_k$
- 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

- Then

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

as $n \nearrow \infty$, where

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

Outline

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

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

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

- Let $\psi(\lambda) = \log E \exp\left(\lambda^T Y_k\right)$
- Put $I(x) = \max_\lambda [\lambda^T x - \psi(\lambda)]$

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

- Let $\psi(\lambda) = \log E \exp\left(\lambda^T Y_k\right)$
- Put $I(x) = \max_\lambda [\lambda^T x - \psi(\lambda)]$
- $f_\lambda(y) = \exp\left(\lambda^T y - \psi(\lambda)\right) f(y)$ <– Controls

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

- Let $\psi(\lambda) = \log E \exp\left(\lambda^T Y_k\right)$
- Put $I(x) = \max_{\lambda}[\lambda^T x - \psi(\lambda)]$
- $f_{\lambda}(y) = \exp\left(\lambda^T y - \psi(\lambda)\right) f(y)$ <– Controls
- HJB eqn. to minimize 2nd moment...

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

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

- Let $\psi(\lambda) = \log E \exp\left(\lambda^T Y_k\right)$
- Put $I(x) = \max_\lambda [\lambda^T x - \psi(\lambda)]$
- $f_\lambda(y) = \exp\left(\lambda^T y - \psi(\lambda)\right) f(y)$ <– Controls
- HJB eqn. to minimize 2nd moment...

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

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

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

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

- Let $\psi(\lambda) = \log E \exp\left(\lambda^T Y_k\right)$
- Put $I(x) = \max_\lambda [\lambda^T x - \psi(\lambda)]$
- $f_\lambda(y) = \exp\left(\lambda^T y - \psi(\lambda)\right) f(y)$ <– Controls
- HJB eqn. to minimize 2nd moment...

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

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

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

- Subject to $g(w) = 0$ on $A \cup B$.

Outline

Isaacs Equation & Harmonic Functions

- $u_n(x) = 1$ on $A \cup B$ and

$$u_n(x) = P_x(T_{A \cup B} < \infty) = E[u_n(x + Y_1/n)]$$

Isaacs Equation & Harmonic Functions

- $u_n(x) = 1$ on $A \cup B$ and

$$u_n(x) = P_x(T_{A\cup B} < \infty) = E[u_n(x + Y_1/n)]$$

- Zero-variance sampler is:

$$P^*(Y_{k+1} \in dy | S_k = nx) = f(y)\frac{u_n(x + y/n)}{u_n(x)}$$

Isaacs Equation & Harmonic Functions

- Zero-variance sampler

$$P^{*}\left(Y_{k+1}\in dy \right| S_{k}=nx) = f\left(y\right) \frac{u_{n}\left(x+y/n\right) }{u_{n}\left(x\right) }$$

Isaacs Equation & Harmonic Functions

- Zero-variance sampler

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

- Approximate sampler $u_n(x) \approx \exp(-nh(x))$

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

Isaacs Equation & Harmonic Functions

- Zero-variance sampler

$$P^{*}\left(Y_{k+1} \in dy \mid S_k = nx \right) = f\left(y\right) \frac{u_n\left(x + y/n\right)}{u_n\left(x\right)}$$

- Approximate sampler $u_n\left(x\right) \approx \exp\left(-nh\left(x\right)\right)$

$$\begin{aligned} & \widetilde{P}\left(Y_{k+1} \in dy \mid S_k = nx \right) \\ \approx \quad & f\left(y\right) \exp\left(-n[h\left(x + y/n\right) - h\left(x\right)]\right) \\ \approx \quad & f\left(y\right) \exp\left(-\partial h\left(x\right) \cdot y\right) \end{aligned}$$

- But

$$1 = \int \widetilde{P}\left(Y_{k+1} \in dy \mid S_k = nx \right) \Longrightarrow \psi\left(-\partial h\left(x\right)\right) = 0$$

Isaacs Equation & Harmonic Functions

- Zero-variance sampler

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

- Approximate sampler $u_n\left(x\right) \approx \exp\left(-nh\left(x\right)\right)$

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

- But

$$1 = \int \widetilde{P}\left(Y_{k+1} \in dy \middle| S_k = nx \right) \Longrightarrow \psi\left(-\partial h\left(x\right)\right) = 0$$

- Equivalent to Isaacs equation with $g\left(x\right) = 2h\left(x\right)$.

Outline

The Second Moment of a State-dependent Estimator

- Consider any sampler

$$P^{Q}\left(Y_{k+1} \in dy \middle| S_k = nx \right) = r^{-1}\left(x, x + y/n\right) f\left(y\right)$$

The Second Moment of a State-dependent Estimator

- Consider any sampler

$$P^{Q}\left(Y_{k+1} \in dy \middle| S_k = nx \right) = r^{-1}\left(x, x + y/n\right) f\left(y\right)$$

- Likelihood ratio

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

The Second Moment of a State-dependent Estimator

- Consider any sampler

$$P^Q\left(Y_{k+1} \in dy \middle| S_k = nx\right) = r^{-1}\left(x, x + y/n\right) f\left(y\right)$$

- 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

$$s(x) = E_x[r(x, x + Y/n)s(x + Y/n)]$$

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

The Lyapunov Inequality

Lemma

B. & Glynn '07: Lyapunov inequality

$$v(x) \geq E_x[r(x, x + Y/n)v(x + Y/n)]$$

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

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

The Lyapunov Inequalities and Subsolutions

- Lyapunov function $v(x) = \exp(-ng(x))$ & $\lambda = -\partial g(x)/2$

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

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

The Lyapunov Inequalities and Subsolutions

- Lyapunov function $v(x) = \exp(-ng(x))$ & $\lambda = -\partial g(x)/2$

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

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

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

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

The Lyapunov Inequalities and Subsolutions

- Lyapunov function $v(x) = \exp(-ng(x))$ & $\lambda = -\partial g(x)/2$

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

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

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

- Yields subsolution to the Isaacs equation (again smoothness)

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

Outline

The Mollification

- Back to random walk example

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

NOT smooth...

The Mollification

- Back to random walk example

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

NOT smooth...

- Mollification:

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

The Mollification

- Back to random walk example

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

NOT smooth...

- Mollification:

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

- Implementation via mixtures:

$$
\begin{aligned}
-\partial h_\varepsilon(x) &= \theta_1^* v_1^T \frac{w_1^\varepsilon(x)}{w_1^\varepsilon(x) + w_2^\varepsilon(x)} + \theta_2^* v_2^T \frac{w_2^\varepsilon(x)}{w_1^\varepsilon(x) + w_2^\varepsilon(x)}, \\
w_1^\varepsilon(x) &= \exp(\theta_1^*(v_1^T x - 1)/\varepsilon), \\
w_2^\varepsilon(x) &= \exp(\theta_2^*(v_2^T x - 1)/\varepsilon).
\end{aligned}
$$

The Mollification

Theorem (Dupuis & Wang '07)

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

$$\text{2nd Moment of Est.} = \exp(-2nh(x) + o(n)).$$

A Lyapunov Inequality

- Select $\varepsilon_n = 1/n$

$$
\begin{aligned}
w_1(x) &= \exp(n\theta_1^*(v_1^T x - 1)) \\
w_2(x) &= \exp(n\theta_2^*(v_2^T x - 1))
\end{aligned}
$$

A Lyapunov Inequality

- Select $\varepsilon_n = 1/n$

$$
\begin{aligned}
w_1(x) &= \exp(n\theta_1^*(v_1^T x - 1)) \\
w_2(x) &= \exp(n\theta_2^*(v_2^T x - 1))
\end{aligned}
$$

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

$$
\frac{\widetilde{f}(y)}{f(y)} = \frac{w_1(x)}{w_1(x) + w_2(x)} \exp\left(\theta_1^* v_1^T y\right) + \frac{w_2(x)}{w_1(x) + w_2(x)} \exp\left(\theta_2^* v_2^T y\right)
$$

A Lyapunov Inequality

- Select $\varepsilon_n = 1/n$

$$\begin{aligned} w_1(x) &= \exp(n\theta_1^*(v_1^T x - 1)) \\ w_2(x) &= \exp(n\theta_2^*(v_2^T x - 1)) \end{aligned}$$

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

$$\frac{\widetilde{f}(y)}{f(y)} = \frac{w_1(x)}{w_1(x) + w_2(x)} \exp\left(\theta_1^* v_1^T y\right) + \frac{w_2(x)}{w_1(x) + w_2(x)} \exp\left(\theta_2^* v_2^T y\right)$$

- Lyapunov function

$$v(x) = (w_1(x) + w_2(x))^2 \geq 1$$

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

A Lyapunov Inequality

- $$\begin{aligned} v(x) &= [w_1(x) + w_2(x)]^2 \\ w_1(x + Y/n) &= w_1(x)\, e^{\theta^* v_1^T Y} \\ w_2(x + Y/n) &= w_2(x)\, e^{\theta^* v_2^T Y} \end{aligned}$$

A Lyapunov Inequality

- $$\begin{aligned} v(x) &= [w_1(x) + w_2(x)]^2 \\ w_1(x + Y/n) &= w_1(x)\, e^{\theta^* v_1^T Y} \\ w_2(x + Y/n) &= w_2(x)\, e^{\theta^* v_2^T Y} \end{aligned}$$

- $$\begin{aligned} & E\frac{v(x+Y/n)}{v(x)} \frac{1}{\frac{w_1(x)}{w_1(x)+w_2(x)} e^{\theta_1^* v_1^T Y} + \frac{w_2(x)}{w_1(x)+w_2(x)} e^{\theta_2^* v_2^T Y}} \\ = \quad & E\frac{w_1(x)\exp\left(\theta_1^* v_1^T Y\right) + w_2(x)\exp\left(\theta_2^* v_2^T Y\right)}{w_1(x)+w_2(x)} = 1. \end{aligned}$$

Moral...

- One can take $\varepsilon = 1/n$ as mollification parameter

Moral...

- One can take $\varepsilon = 1/n$ as mollification parameter
- $\varepsilon = 1/n$ is the optimal choice (bounded coef. of variation)

$$2nd \text{ Moment} \leq (v_1(0) + v_2(0))^2 \leq cu_n(0)^2$$

Outline

Markovian Tandem Networks

- Consider a stable tandem network with d stations.
- $\beta = \#$ of bottlenecks (i.e. stations with maximum load)

Theorem (B., Glynn and Leder '09)

There is selection of mollification parameters that guarantees the coefficient of variation of the Dupuis-Sezer-Wang '07 sampler to be $O\left(n^{2(d-\beta+1)}\right)$.

Remark: This guarantees *better* performance than solving the associated linear system of equations.

Outline

Conclusions

- Subsolutions can be understood from the standpoint of Lyapunov inequalities.
- Lyapunov inequalities help understand the nature of mollification parameters.
- Lyapunov inequalities guided by subsolutions can strengthen the performance analysis.
- Better complexity of importance sampling vs. linear system for tandem rigorously validated.