Importance Sampling Methodology for Multidimensional Heavy-tailed Random Walks

Jose Blanchet (joint work with Jingchen Liu)

Columbia IEOR Department

RESIM

Agenda

- **Introduction**
- The One Dimensional Case
- Multidimensional Case
- Markov Random Walks
- Conclusions

Introduction

- **What is the general goal of this talk?**

- **What is the general goal of this talk?**
- *Discuss simulation methodology for estimating small expectations associated to rare events in multidimensional heavy-tailed settings.*

Introduction

- **What is the general goal of this talk?**
- *Discuss simulation methodology for estimating small expectations associated to rare events in multidimensional heavy-tailed settings.*
- **What is the objective of the talk?**

- **What is the general goal of this talk?**
- *Discuss simulation methodology for estimating small expectations associated to rare events in multidimensional heavy-tailed settings.*
- **What is the objective of the talk?**

1. *Illustrate our methodology for multidimensional regularly varying random walks.*

Introduction

- **What is the general goal of this talk?**
- *Discuss simulation methodology for estimating small expectations associated to rare events in multidimensional heavy-tailed settings.*
- **What is the objective of the talk?**

1. *Illustrate our methodology for multidimensional regularly varying random walks.*
2. *Discuss optimality properties of our proposed simulation procedure.*

Initial Notation

- Let X_1, X_2,... are iid regularly varying in R^d *(to be discussed –> TBD)*.
- $EX_i = \eta \in R^d$ and A is an appropriate subset of R^d.
- Given $b > 0$ we write $bA = \{ba : a \in A\}$.
- $S_n = X_1 + ... + X_n$, $(S_0 = s)$.
- $T_{bA} = \inf\{n \geq 0 : S_n \in bA\}$.
- Object of interest:

$$u_b(s, f) = E\left(f\left(S_0, S_1, ..., S_{T_{bA}}, T_{bA}\right) I\left(T_{bA} < \infty\right)\right),$$

for *any* $f(\cdot)$ such that $0 < \delta_f \leq f \leq \delta_f^{-1}$ and $u_b(s, 1) \longrightarrow 0$ as $b \nearrow \infty$ (TBD).

Efficiency and Performance Guarantee

Given any f for which there is $\delta_f \in (0, \infty)$, so that $\delta_f \leq f \leq \delta_f^{-1}$ estimate

$$u_b(s, f) = E\left(f\left(S_0, S_1, ..., S_{T_{bA}}, T_{bA}\right) I\left(T_{bA} < \infty\right)\right),$$

with good relative precision.

- The relative error of an estimator Z for $u_b(s, f)$ is

$$Rel.Error = \frac{(Var(Z))^{1/2}}{u_b(s)} + \left|\frac{EZ}{u_b(s)} - 1\right|$$

- Suppose $EZ = u_b(s, f)$ and that $Var(Z) = O\left(u_b(s, f)^2\right)$, then we say that Z is *strongly efficient*.

- If Z is unbiased, then sampling n iid replications of Z gives an estimator $\widehat{u}_b(n) = n^{-1}\sum_{j=1}^{n} Z_j$ such that

$$P\left(\left|\widehat{u}_b(n) - u_b(s,f)\right| \geq \varepsilon u_b(s,f)\right) \leq \frac{Var(Z)}{n\varepsilon^2 u_b(s,f)^2}.$$

- So, it takes $O\left(\varepsilon^{-2}\delta^{-1}Var(Z)/u_b(s,f)^2\right)$ replications to achieve ε-relative error with $(1-\delta)\cdot 100\%$ confidence.
- *Any* procedure for estimating $u_b(s,f)$ (for arbitrary f) must consider $(S_k : 0 \leq k \leq T_{bA})$ (this requires on average $O(E_s(T_{bA} | T_{bA} < \infty))$ operations).

- *In summary, the best performance that one can expect for ε-relative precision and $(1-\delta)\cdot 100\%$ confidence based on iid replications of a given estimator involves*

$$O\left(\varepsilon^{-2}\delta^{-1}E_s\left(\,T_b|\,T_b<\infty\right)\right)$$

operations.

An algorithm that achieves such performance is said to be optimal

Agenda

- Introduction
- **The One Dimensional Case**
- Multidimensional Case
- Markov Random Walks
- Conclusions

The One Dimensional Case

- There is a rich large deviations theory for heavy-tailed random walks based on *subexponential* rv's

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

as $b \longrightarrow \infty$.

- *Focus on an important class* of a subexponential distributions is *the class of 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 $\beta > 0$.

The One Dimensional Case

- Let $EX_i = \eta < 0$ and set $A = [1, \infty)$, (we write $T_b = T_{bA}$) estimate ruin (Pakes, Veraberbeker, Cohen... see text of Asmussen '03)

$$u_b(s, 1) = u_b(s) = P_s(T_b < \infty).$$

The One Dimensional Case

- Let $EX_i = \eta < 0$ and set $A = [1, \infty)$, (we write $T_b = T_{bA}$) estimate ruin (Pakes, Veraberbeker, Cohen... see text of Asmussen '03)

$$u_b(s,1) = u_b(s) = P_s(T_b < \infty).$$

- **Strategy**: use importance sampling consider the Markov kernel $K(\cdot)$

$$\begin{aligned} K(s_0, s_1 + ds_1) &= r^{-1}(s_0, s_1)\, P(s_0 + X_1 \in s_1 + ds_1) \\ K(s_0, s_1)\, ds_1 &= r^{-1}(s_0, s_1)\, f_{X_1}(s_1 - s_0)\, ds_1 \qquad (1) \end{aligned}$$

for a positive function $r(\cdot)$. ((1) valid in presence of densities).

The One Dimensional Case

- Let $EX_i = \eta < 0$ and set $A = [1, \infty)$, (we write $T_b = T_{bA}$) estimate ruin (Pakes, Veraberbeker, Cohen... see text of Asmussen '03)

$$u_b(s,1) = u_b(s) = P_s(T_b < \infty).$$

- **Strategy**: use importance sampling consider the Markov kernel $K(\cdot)$

$$\begin{aligned} K(s_0, s_1 + ds_1) &= r^{-1}(s_0, s_1) P(s_0 + X_1 \in s_1 + ds_1) \\ K(s_0, s_1) ds_1 &= r^{-1}(s_0, s_1) f_{X_1}(s_1 - s_0) ds_1 \end{aligned} \quad (1)$$

for a positive function $r(\cdot)$. ((1) valid in presence of densities).

- **The importance sampling estimator is:**

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

the S_n's are simulated under $K(\cdot)$.

The One Dimensional Case

- **Classical result:** The conditional distribution of the random walk, given that $T_b < \infty$, gives an exact (zero variance) estimator for $u_b(s)$.

The One Dimensional Case

- **Classical result:** The conditional distribution of the random walk, given that $T_b < \infty$, gives an exact (zero variance) estimator for $u_b(s)$.
- **Moral of the story:**

The One Dimensional Case

- **Classical result:** The conditional distribution of the random walk, given that $T_b < \infty$, gives an exact (zero variance) estimator for $u_b(s)$.
- **Moral of the story:**
  - *Select an importance sampler that mimics the behavior of such conditional distribution.*

The One Dimensional Case

Theorem

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

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

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, the random walk has drift η and a large jump of size b occurs suddenly in $O(b)$ time...

The One Dimensional Case

Theorem

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

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

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, the random walk has drift η 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 $S_k \approx \eta k$ and the chance of reaching b in the next increment *given* that we eventually reach b ($T_b < \infty$)

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

The One Dimensional Case

- **Family of changes-of-measure:** Here s is the current position of the walk, f is the density (NOTE $p(s)$ and $a \in (0,1)$)

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

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

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

The One Dimensional Case

- **Lyapunov Inequalities for Variance Control:**

Lemma (B. & Glynn '07)

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,*

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

The One Dimensional Case

- **Testing the Inequality:**
- We 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 (*note that we need to select* θ, $\kappa > 0$ *to force Lyapunov ineq.*)

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

The One Dimensional Case

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

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

The One Dimensional Case

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

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

The One Dimensional Case

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

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

The One Dimensional Case

- **Corresponding Algorithm:**
- Select $a \in (0, 1)$, then choose θ and κ based on Lyapunov inequality
- AT EACH TIME STEP TEST
  -
    - IF $g(s) < 1$ apply Imp. Sampling according $p(s)$ - mixture
    - ELSE do NOT apply I.S. and continue until hitting.
  - OUTPUT PRODUCT OF LIKELIHOOD RATIOS Z

The One Dimensional Case

- **Termination of the algorithm:** Intuitively, what can go wrong?
- Let M be the time until we get one large jump given $S_0 = 0$

$$
\begin{aligned}
P(M > tb) &\approx \prod_{j=0}^{kb} (1 - p(-j\eta)) \approx \exp\left(-\sum_{j=0}^{tb} \frac{\theta(\alpha - 1)}{b - \eta j} \right) \\
&\approx \exp\left(-\int_0^{tb} \frac{\theta(\alpha - 1)}{b - \eta s} ds \right) = \left(\frac{1}{1 - \eta t} \right)^{\theta(\alpha - 1)/(-\eta)}.
\end{aligned}
$$

So, if $\theta < -\eta/(\alpha - 1)$ then EXPECTED TERMINATION TIME COULD BE INFINITE!

The One Dimensional Case

- **Optimal Selection of Parameters:**

Lemma

For each $\varepsilon_1 > 0$ *there exists* $\delta_1, m > 0$ *such if* $\theta = -\eta - \delta_1$ *and* $a = 1 - \delta_1$ *then, if* $\alpha > 2$,

$$\left| E_0^K T_b - E_0 \left(T_b \middle| \, T_b < \infty \right) \right| \leq \varepsilon_1 E_0 \left(T_b \middle| \, T_b < \infty \right) \text{ for all } b \geq m$$

Corollary

The importance sampling algorithm based on the previous selection of parameters is **optimal.**

The One Dimensional Case

- **Implementation Issues:**

The One Dimensional Case

- **Implementation Issues:**
- Numerically test values of κ, θ and a before implementing

- **Implementation Issues:**
- Numerically test values of κ, θ and a before implementing
- Typically not too hard if one is willing to be somewhat conservative

The One Dimensional Case

- **Implementation Issues:**
- Numerically test values of κ, θ and a before implementing
- Typically not too hard if one is willing to be somewhat conservative
- BUT, as we saw, this can give infinite expected hitting time

- **Introducing a Controlled Bias:**

Note that

$$\begin{aligned} P_0\left(T_b < \infty\right) &= P_0\left(T_b < T_{-\beta b}\right) + P_0\left(T_{-\beta b} < T_b, T_b < \infty\right) \\ &\leq P_0\left(T_b < T_{-\beta b}\right) + P_{-\beta b}\left(T_b < \infty\right) \\ &\leq P_0\left(T_b < T_{-\beta b}\right) + g\left(-\beta b\right)^{1/2}. \end{aligned}$$

So, relative bias

$$0 \leq 1 - \frac{P_0\left(T_b < T_{-\beta b}\right)}{P_0\left(T_b < \infty\right)} \leq \frac{g\left(-\beta b\right)^{1/2}}{P_0\left(T_b < \infty\right)} \leq \frac{g\left(-\beta b\right)^{1/2}}{P_0\left(T_b < T_{-\beta b}\right)}$$

The One Dimensional Case

- **Complexity Count:**

$$\text{RELATIVE BIAS} = O\left(\beta^{-\alpha/2}\right).$$

- Want relative bias less than $\varepsilon/2$. Then, total complexity count (for ε-relative error with $(1-\delta)\cdot 100\%$ confidence)

$$O\left(\varepsilon^{-2}\delta^{-1}\varepsilon^{-2/\alpha}E_0\left(\left.T_b\right| T_b < \infty\right)\right).$$

The One Dimensional Case

- **Complexity Count:**

$$\text{RELATIVE BIAS} = O\left(\beta^{-\alpha/2}\right).$$

- Want relative bias less than $\varepsilon/2$. Then, total complexity count (for ε-relative error with $(1-\delta)\cdot 100\%$ confidence)

$$O\left(\varepsilon^{-2}\delta^{-1}\varepsilon^{-2/\alpha}E_0\left(T_b \mid T_b < \infty\right)\right).$$

- Or letting $\beta(b) \nearrow \infty$ one gets a complexity count of order

$$O\left(\varepsilon^{-2}\delta^{-1}\beta(b)\,E_0\left(T_b \mid T_b < \infty\right)\right).$$

Agenda

- Introduction
- The One Dimensional Case
- **Multidimensional Case**
- Markov Random Walks
- Conclusions

The Multidimensional Case

Assumptions on X

- **Multidimensional regular variation**: There exists a (Radon) measure $\mu(\cdot)$ such that

$$\frac{P(X \in Ab)}{P(|X| > b)} \Longrightarrow \mu(A)$$

(sense of vague convergence) for $A \in \mathcal{B}\left(R^d \backslash \{0\}\right)$. Resnick (2007), Hult et al (2005),...

- **Example:** X follows a standard t with α degrees of freedom, then

$$P(X \in bA) = \int_A \frac{m}{(1 + y^T y/v)^{(\alpha+d)/2}} dy,$$

$$\mu(A) = \int_A \frac{dy}{(y^T y)^{(\alpha+d)/2}}$$

The Multidimensional Case

Assumptions on X

- **A bit of intuition about multivariate regular variation:**
- $\mu(\cdot)$ contains the information about the directions where large jumps occur

Assumptions on X

- **A bit of intuition about multivariate regular variation:**
- $\mu(\cdot)$ contains the information about the directions where large jumps occur
-
$$\mu(\{y : |y| > r; y/|y| \in S\}) = r^{-\alpha}\sigma(S)$$

The Multidimensional Case

Assumptions on X

- **A bit of intuition about multivariate regular variation:**
- $\mu(\cdot)$ contains the information about the directions where large jumps occur
- $$\mu(\{y : |y| > r; y/|y| \in S\}) = r^{-\alpha}\sigma(S)$$
- **Example:** For the t distribution $\sigma(\cdot)$ is the uniform distribution on the unit sphere.

Assumptions on $EX \in R^d$ and A

- **Assumption (CONE):** There are linearly independent vectors $v_1^*, v_2^*, ..., v_m^* \in R^d$ ($\|v_i^*\| = 1$) and $\delta^* > 0$ such that $v_i^{*T}\beta \leq -\delta^*$ for all $1 \leq i \leq m$ and for all $z \in A$ there is at least one i for which $v_i^{*T} z \geq \delta$.

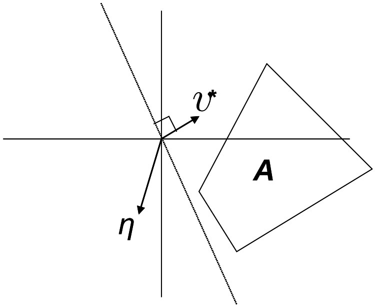
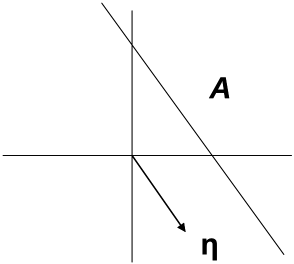


Diagram illustrating Assumption A or a two dimensional random walk

The Multidimensional Case

Assumptions on $EX \in R^d$ and A

- Previous assumption rules out the following situation:



- Here the drift goes parallel to A and the process will eventually hit A (this is the analogue of having $EX = 0$ in $d = 1$).

The Multidimensional Case

Assumptions on $\mu(\cdot)$ and A

- Avoid degenerate situations where A is "very thin" relative to $\mu(\cdot)$ (example A a line in the case of a t distribution)...

Assumptions on $\mu(\cdot)$ and A

- Avoid degenerate situations where A is "very thin" relative to $\mu(\cdot)$ (example A a line in the case of a t distribution)...
- **Assumption (POS):** Assume that $\alpha > 1$ (regularly varying index) and there exists $\delta_1 > 0$ such that

$$\inf_{0 \le h \le \delta_1} \mu(A - h\eta) > 0$$

Ruin Probabilities: (Hult, Lindskog, Mikosch and Samorodnitsky '05, Hult and Lindskog '06)

$$
\begin{aligned}
P_0\left(T_{Ab} < \infty\right) &\sim \int_0^\infty P\left(X + t\eta \in bA^*\right) dt \\
&\sim bP\left(|X| > b\right) \int_0^\infty \mu\left(A^* - t\eta\right) dt
\end{aligned}
$$

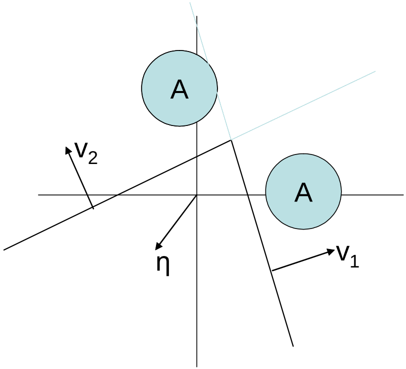
as $b \nearrow \infty$, where $A^* = \{A - t\eta : t \geq 0\}$

Remark: By assumption POS

$$
\int_0^\infty \mu\left(A^* - t\eta\right) dt \in (0, \infty).
$$

The Multidimensional Case

Selection of the change-of-measure:



The Multidimensional Case

Selection of the change-of-measure:

- $$C_i^a(s,b) = \{x : {v_i^*}^T x > a(\delta^* b - {v_i^*}^T s)\}$$

The Multidimensional Case

Selection of the change-of-measure:

- $$C_i^a(s,b) = \{x : v_i^{*T} x > a\left(\delta^* b - v_i^{*T} s\right)\}$$

- $$q_{X|s}(x) = p(s)\,\frac{f_X(x)\, I\left(\cup_{i=1}^m C_i^a(s,b)\right)}{P\left(\cup_{i=1}^m C_i^a(s,b)\right)} + (1-p(s))\,\frac{f_X(x)\, I\left(\cap_{i=1}^m \overline{C}_i^\alpha(s,b)\right)}{P\left(\cap_{i=1}^m \overline{C}_i^\alpha(s,b)\right)}.$$

The Multidimensional Case

Selection of Lyapunov Inequality:

- Heuristic (fluid) analysis suggests

$$
\begin{aligned}
P_s\left(T_{bA}<\infty\right) &\approx \int_0^\infty P\left(X+s+\eta t\in bA\right)dt \\
&\le \sum_{i=1}^m \int_0^\infty P\left({v_i^*}^T\left(X+s+\eta t\right)\ge b\delta^*\right)dt \\
&= \sum_{i=1}^m \frac{1}{-{v_i^*}^T\eta} G_i\left(b\delta^*-{v_i^*}^T s\right),
\end{aligned}
$$

where

$$
G_i\left(b\delta^*-{v_i^*}^T s\right)=\int_{b\delta^*-{v_i^*}^T s}^\infty P\left({v_i^*}^T X>u\right)du
$$

The Multidimensional Case

Lyapunov function to test:

- Heuristic (fluid) analysis suggests

$$g_b(s) = \min\left(\kappa h_b(s)^2, 1\right) = O\left(b^2 P(|X| > b)^2\right),$$

where

$$h_b(s) = \sum_{i=1}^{m} \frac{1}{-v_i^{*T}\eta} G_i\left(b\delta^* - v_i^{*T} s\right).$$

Theorem

Pick $p(s) = \theta P\left(\cup_{i=1}^{m} C_i^a(s,b)\right)/h_b(s)$ *and* $a \in (0,1)$*, then* κ*,* $\theta > 0$ *can be chosen so that* $g_b(\cdot)$ *satisfies the Lyapunov inequality. So, applying the proposed IS when* $g_b(s) < 1$ *gives a strongly efficient estimator.*

The Multidimensional Case

Implementation issues:

- Sampling transitions AND computing $p(s)$ involves a rare-event simulation problem

The Multidimensional Case

Implementation issues:

- Sampling transitions AND computing $p(s)$ involves a rare-event simulation problem
  - Right, but now the problem is finite dimensional and static

The Multidimensional Case

Implementation issues:

- Sampling transitions AND computing $p(s)$ involves a rare-event simulation problem
  - Right, but now the problem is finite dimensional and static
  - Sometimes one can use the spectral measure to do the sampling

Implementation issues:

- Sampling transitions AND computing $p(s)$ involves a rare-event simulation problem
  - Right, but now the problem is finite dimensional and static
  - Sometimes one can use the spectral measure to do the sampling
- Termination is even more complicated than in 1 dimension

The Multidimensional Case

Implementation issues:

- Sampling transitions AND computing $p(s)$ involves a rare-event simulation problem
  - Right, but now the problem is finite dimensional and static
  - Sometimes one can use the spectral measure to do the sampling
- Termination is even more complicated than in 1 dimension
  - As in $d = 1$ one can introduce a controlled bias

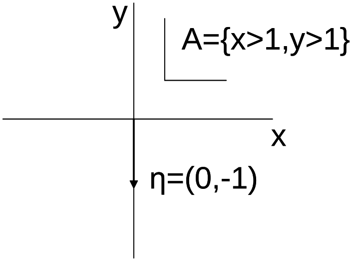
The Multidimensional Case

Implementation issues:

- Sampling transitions AND computing $p(s)$ involves a rare-event simulation problem
  - Right, but now the problem is finite dimensional and static
  - Sometimes one can use the spectral measure to do the sampling
- Termination is even more complicated than in 1 dimension
  - As in $d = 1$ one can introduce a controlled bias
- Determination of the v_i^*'s when A is union of polyhedra through linear programs

The Multidimensional Case

Example: Random walk with t distributed increments and $\alpha = 3$ degrees of freedom. Relative bias no more than 5% with at least 97% confidence.



b	Est	SD	Sample size
10	$1.83e-03$	$1.31e-02$	20000
20	$4.28e-04$	$2.54e-03$	20000
50	$8.80e-05$	$4.74e-04$	20000

The Multidimensional Case

Theorem

Assume CONE and POS. In addition suppose that the sampling and the evaluation of $K(\cdot)$ can be done in $O(1)$ operations (as $b \nearrow \infty$). Then, state-dependent importance sampling via Lyapunov control requires

$$O\left(\varepsilon^{-2}\delta^{-1}\varepsilon^{-2/\alpha}b\right)$$

operations to achieve ε-relative error with $(1-\delta)\cdot 100\%$ confidence.

Agenda

- Introduction
- The One Dimensional Case
- Multidimensional Case
- **Markov Random Walks**
- Conclusions

Summary of the methodology:

- *Select an appropriate change of measure.*

Summary of the methodology:

- *Select an appropriate change of measure.*
- *Use Lyapunov inequalities for variance / bias control and termination.*

Summary of the methodology:

- *Select an appropriate change of measure.*
- *Use Lyapunov inequalities for variance / bias control and termination.*
- *NOTE the difference between light vs heavy-tailed... discrete nature (jumps for heavy tails) forces control at very short time scales!*

Markov Random Walks

Summary of the methodology:

- *Select an appropriate change of measure.*
- *Use Lyapunov inequalities for variance / bias control and termination.*
- *NOTE the difference between light vs heavy-tailed... discrete nature (jumps for heavy tails) forces control at very short time scales!*
- **Goal: Apply Lyapunov inequalities at more than one time step**

Markov Random Walks

Basic idea (again let d = 1):

- To simplify, let $(Y_n : n \geq 0)$ be an irreducible, aperiodic, finite state-space recurrent Markov chain

$$P\left(X_n > t \mid Y_n = k\right) = L_k\left(t\right) t^{-\alpha}$$

Markov Random Walks

Basic idea (again let d = 1):

- To simplify, let $(Y_n : n \geq 0)$ be an irreducible, aperiodic, finite state-space recurrent Markov chain

$$P(X_n > t | Y_n = k) = L_k(t)\, t^{-\alpha}$$

- $E_\pi X_n < 0$ (but there maybe states $y = i$ for which $E_i X_n > 0$)

Markov Random Walks

Basic idea (again let d = 1):

- To simplify, let $(Y_n : n \geq 0)$ be an irreducible, aperiodic, finite state-space recurrent Markov chain

$$P(X_n > t | Y_n = k) = L_k(t) t^{-\alpha}$$

- $E_\pi X_n < 0$ (but there maybe states $y = i$ for which $E_i X_n > 0$)
- Application of Lyapunov inequality doesn't work (MIGHT NOT SEE $E_\pi X_n < 0$)

Markov Random Walks

Basic idea (again let d = 1):

- To simplify, let $(Y_n : n \geq 0)$ be an irreducible, aperiodic, finite state-space recurrent Markov chain

$$P\left(X_n > t \middle| Y_n = k \right) = L_k\left(t \right) t^{-\alpha}$$

- $E_\pi X_n < 0$ (but there maybe states $y = i$ for which $E_i X_n > 0$)
- Application of Lyapunov inequality doesn't work (MIGHT NOT SEE $E_\pi X_n < 0$)
- Test Lyapunov inequality and apply algorithm at time scales of order τ (for appropriately define stopping time τ).

Agenda

- Introduction
- The One Dimensional Case
- Multidimensional Case
- Markov Random Walks
- **Conclusions**

Conclusions

- Discussed rare-event simulation for multidimensional random walks
  - Discrete nature (sudden jumps) forces control at short time scales
  - Variance control via Lyapunov functions obtained from fluid analysis
  - Biased estimators to control termination, but *bias is quantifiable*
  - Provably efficient algorithms & interplay between linear programming, affine functions and regular variation.
- Markov random walks (and other settings, small noise processes?) force control at slightly larger time scales —> Lyapunov functions applied to τ transition kernel.