Learning to reflect: data-driven stochastic optimal control strategies for diffusions and Lévy processes

Berlin Probability Colloquium

Sören Christensen (CAU Kiel) Lukas Trottner (Universität Mannheim)
joint work with Claudia Strauch (Aarhus University)

30.06.2021

Outline

1. Two classical singular control problems
2. Data-driven approach to singular control
3. Construction of estimators

Two classical singular control problems

Framework (diffusions)

regular 1-dim. Itô diffusion

$$\mathrm{d}X(t) = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t,$$

with assumptions that guarantee an *invariant density*

$$\rho(x) \;\coloneqq\; \frac{1}{C\sigma^2(x)} \exp\left(2\int^x \frac{b(y)}{\sigma^2(y)} dy\right),$$

Framework (diffusions)

- *Singular control*: $Z = (U_t, D_t)_{t \geqslant 0}$, U, D non-decreasing, right-continuous and adapted,

$$dX_t^Z = b(X_t^Z)\,dt + \sigma(X_t^Z)\,dW_t + dU_t - dD_t,$$

Framework (diffusions)

- *Singular control*: $Z = (U_t, D_t)_{t \geqslant 0}$, U, D non-decreasing, right-continuous and adapted,

$$\mathrm{d}X_t^Z = b(X_t^Z)\,\mathrm{d}t + \sigma(X_t^Z)\,\mathrm{d}W_t + \mathrm{d}U_t - \mathrm{d}D_t,$$

- c continuous, nonnegative running cost function, $q_u, q_d > 0$.
  Minimize

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left(\int_0^T c(X_s^Z)\,\mathrm{d}s + q_u U_T + q_l D_T\right),$$

Solution for singular control problem (diffusions)

For each (c, d), the corresponding reflection strategy has value

$$C(c,d) = \frac{1}{M(c,d)} \left(\int_c^d c(x)\, \mathrm{d}M(x) + \frac{q_u}{S'(c)} + \frac{q_d}{S'(d)} \right),$$

M speed measure, S scale function

Solution for singular control problem (diffusions)

For each (c, d), the corresponding reflection strategy has value

$$C(c,d) = \frac{1}{M(c,d)} \left(\int_c^d c(x)\, \mathrm{d}M(x) + \frac{q_u}{S'(c)} + \frac{q_d}{S'(d)} \right),$$

M speed measure, S scale function

Theorem (Alvarez (2018))

Under some assumptions, the value for the singular problem is given by

$$V_{\text{sing}} = \min_{(c,d)} C(c,d).$$

and the reflections strategy for the minimizer (c^*, d^*) is optimal.

Problem formulation (Lévy processes)

- X a *Lévy process* on $\mathbb{R}$, $\mathbb{E}^0 X_1 \in (0, \infty)$

Problem formulation (Lévy processes)

- X a *Lévy process* on $\mathbb{R}$, $\mathbb{E}^0 X_1 \in (0, \infty)$
- for impulse controls $S = (\tau_n, \zeta_n)$

$$X_t^S = X_t - \sum_{n;\ \tau_n \leqslant t} (X_{\tau_n -}^S - \zeta_n)$$

and for a nice $\mathcal{C}^2$ reward function γ solve

$$V_{\text{sing}} \coloneqq \sup_S \liminf_{T \to \infty} \frac{1}{T} \mathbb{E}^x \sum_{n: \tau_n \leqslant T} \left(\gamma \left(X_{\tau_n, -}^S \right) - \gamma \left(\zeta_n \right) \right) \tag{1}$$

Solution for known dynamics (Lévy processes)

- For $T_y = \inf\{t : X_t > y\}$ define the auxiliary function

$$f(x) := \lim_{\varepsilon \searrow 0} \frac{\mathbb{E}^x \gamma(X_{T_{x+\varepsilon}}) - \gamma(x)}{\mathbb{E}^x T_{x+\varepsilon}}$$

  (Long term average reward when reflecting in x)

Solution for known dynamics (Lévy processes)

- For $T_y = \inf\{t : X_t > y\}$ define the auxiliary function

$$f(x) := \lim_{\varepsilon \searrow 0} \frac{\mathbb{E}^x \gamma(X_{T_{x+\varepsilon}}) - \gamma(x)}{\mathbb{E}^x T_{x+\varepsilon}}$$

  (Long term average reward when reflecting in x)
- *H ladder height process* of X: subordinator with characteristics (d_H, Π_H)

Solution for known dynamics (Lévy processes)

- For $T_y = \inf\{t : X_t > y\}$ define the auxiliary function

$$f(x) := \lim_{\varepsilon \searrow 0} \frac{\mathbb{E}^x \gamma(X_{T_{x+\varepsilon}}) - \gamma(x)}{\mathbb{E}^x T_{x+\varepsilon}}$$

  (Long term average reward when reflecting in x)
- *H ladder height process* of X: subordinator with characteristics (d_H, Π_H)
- f via the extended *generator*:

$$\begin{aligned} f(x) &= \mathcal{A}_H \gamma(x) \\ &= d_H \gamma'(x) + \int_0^\infty (\gamma(x+y) - \gamma(x))\, \Pi_H(\mathrm{d}y) \end{aligned}$$

Solution for known dynamics (Lévy processes)

- For $T_y = \inf\{t : X_t > y\}$ define the auxiliary function

$$f(x) := \lim_{\varepsilon \searrow 0} \frac{\mathbb{E}^x \gamma(X_{T_{x+\varepsilon}}) - \gamma(x)}{\mathbb{E}^x T_{x+\varepsilon}}$$

  (Long term average reward when reflecting in x)
- *H ladder height process* of X: subordinator with characteristics (d_H, Π_H)
- f via the extended *generator*:

$$\begin{aligned} f(x) &= \mathcal{A}_H \gamma(x) \\ &= d_H \gamma'(x) + \int_0^\infty (\gamma(x+y) - \gamma(x)) \, \Pi_H(\mathrm{d}y) \end{aligned}$$

Theorem (C., Sohr (2020))

Let f be unimodal with maximizer θ^* (+ technical assumptions). Then $V_{\text{sing}} = f(\theta^*)$ and reflecting in θ^* is optimal.

Data-driven approach to singular control

Questions

- Which are the relevant *characteristics* of X to *estimate* approximately optimal boundaries?

Questions

- Which are the relevant *characteristics* of X to *estimate* approximately optimal boundaries?
- How does controlling the process *influence* the estimation?

Main observation (diffusion)

Crucial characteristics: b (assume σ to be known).

Main observation (diffusion)

Crucial characteristics: b (assume σ to be known).

Singular problem: $V_{\text{sing}} = \min_{c,d \in K} C(c,d)$.

$$C(c,d) = \frac{1}{\int_c^d m(x)\,\mathrm{d}x} \left(\int_c^d c(x) m(x)\,\mathrm{d}x + \frac{q_u}{S'(c)} + \frac{q_d}{S'(d)} \right),$$

$$= \frac{1}{\int_c^d \rho(x)\,\mathrm{d}x} \left(\int_c^d c(x) \rho(x)\,\mathrm{d}x + \frac{q_u \sigma^2(c)}{2} \rho(c) + \frac{q_d \sigma^2(d)}{2} \rho(d) \right)$$

Main observation (diffusion)

Crucial characteristics: b (assume σ to be known).

Singular problem: $V_{\text{sing}} = \min_{c,d\in K} C(c,d)$.

$$C(c,d) = \frac{1}{\int_c^d m(x)\,\mathrm{d}x}\left(\int_c^d c(x)m(x)\,\mathrm{d}x + \frac{q_u}{S'(c)} + \frac{q_d}{S'(d)}\right),$$

$$= \frac{1}{\int_c^d \rho(x)\,\mathrm{d}x}\left(\int_c^d c(x)\rho(x)\,\mathrm{d}x + \frac{q_u\sigma^2(c)}{2}\rho(c) + \frac{q_d\sigma^2(d)}{2}\rho(d)\right)$$

Plug-in estimator: If $\widehat{\rho}_T$ is an estimator of ρ, then use

$$\widehat{C}_T(c,d) := \frac{1}{\int_c^d \widehat{\rho}_T(x)\,\mathrm{d}x}\left(\int_c^d c(x)\widehat{\rho}_T(x)\,\mathrm{d}x + \frac{q_u\sigma^2(c)}{2}\widehat{\rho}_T(c) + \frac{q_d\sigma^2(d)}{2}\widehat{\rho}_T(d)\right),$$

$$\widehat{(c,d)}_T \in \arg\min_{(c,d)} \widehat{C}_T(c,d)$$

Using estimators (diffusions)

Theorem (C., Strauch, T. (2021+))

Assume that we have a data-driven estimator $\widehat{\rho}_T$ for ρ. Then

$$\mathbb{E}_b\left[V_{\text{sing}} - C(\widehat{(c,d)}_T)\right] \leqslant 2\max_{(c,d)}\left|C(c,d) - \widehat{C}_T(c,d)\right|$$
$$\lesssim \mathbb{E}_b\left[\|\widehat{\rho}_T - \rho_b\|_{L^\infty}\right]$$

Using estimators (diffusions)

Theorem (C., Strauch, T. (2021+))

Assume that we have a data-driven estimator $\widehat{\rho}_T$ for ρ. Then

$$\mathbb{E}_b\left[V_{\text{sing}} - C(\widehat{(c,d)}_T)\right] \leqslant 2 \max_{(c,d)} \left| C(c,d) - \widehat{C}_T(c,d) \right| \lesssim \mathbb{E}_b\left[\|\widehat{\rho}_T - \rho_b\|_{L^\infty}\right]$$

Need nonparametric bounds for $\mathbb{E}_b\left[\|\widehat{\rho}_T - \rho_b\|_{L^\infty}\right]$.

Exploration vs. exploitation (diffusions)

Central Assumption in Stochastic Control

The dynamics of the underlying process is known.

What to do if this is not the case?

Exploration vs. exploitation (diffusions)

Simple-minded idea:

- Estimate the optimal boundary based on the controlled process.
- Use the strategy based on the estimated boundary

Exploration vs. exploitation (diffusions)

Simple-minded idea:

- Estimate the optimal boundary based on the controlled process.
- Use the strategy based on the estimated boundary

Problem

Exploration vs. Exploitation!

Exploration vs. exploitation (diffusions)

Combine Exploration- and Exploitation-Cycles...

Exploration vs. exploitation (diffusions)

Combine Exploration- and Exploitation-Cycles...

- S_t: (random) time in exploration-cycles until t
- $X' := (X^K_{S_s^{-1}})_{s \geqslant 0}$ is a diffusion process
- In the exploitation cycles, use the estimated boundaries based on the information obtained in the exploration cycles.

Exploration vs. exploitation (diffusions)

Theorem (C., Strauch, T. (2021+))

Assume that we have a data-driven estimator $\widehat{\rho}_T$ for ρ with

$$\mathbb{E}_b^0\left[\|\widehat{\rho}_T - \rho_b\|_{L^\infty}\right] \in \mathrm{O}\left(\sqrt{\frac{\log T}{T}}\right)$$

and consider S such that $S_T \approx T^{2/3}$. Then, the regret (difference of optimal reward rate and the expected data driven reward rate) is of order $\mathrm{O}\left(\sqrt{\log T}\, T^{-1/3}\right)$.

Strategy for Lévy processes

- no exploration/exploitation problem by spatial homogeneity of Lévy processes $\rightsquigarrow$ recover path from controlled process following the same distribution as path from the uncontrolled process by "undoing" controls
- All we need is a good estimator $\widehat{f}_T$ of $f = \mathcal{A}_H \gamma$ wrt. sup-norm risk

$\rightsquigarrow$ continuously update estimated boundary via greedy strategy $\widehat{\theta}_T = \sup_x \widehat{f}_T(x)$

Strategy for Lévy processes

- no exploration/exploitation problem by spatial homogeneity of Lévy processes $\rightsquigarrow$ recover path from controlled process following the same distribution as path from the uncontrolled process by "undoing" controls
- All we need is a good estimator $\widehat{f}_T$ of $f = \mathcal{A}_H\gamma$ wrt. sup-norm risk

$\rightsquigarrow$ continuously update estimated boundary via greedy strategy $\widehat{\theta}_T = \sup_x \widehat{f}_T(x)$

Theorem (C., Strauch, T. (2021+))

Assume that $\theta^* \in D$ for some open bounded set D and we have a data-driven estimator $\widehat{f}_T$ for f satisfying

$$\mathbb{E}^0\left[\left\|\widehat{f}_T - f\right\|_{L^\infty(D)}\right] \in \mathrm{O}(\eta(T)),$$

for some $\eta \in \mathrm{o}(1)$. Then, the regret is of order $\mathrm{O}(\eta(T))$ as well.

Construction of estimators

Invariant density estimation for diffusions

- missing piece for data-driven estimator of optimal reflection boundaries: estimator $\widehat{\rho}_T$ of stationary density ρ with sup-norm rate $O(\sqrt{\log T/T})$
- assumption: continuous record $X^T = (X_t)_{t\in[0,T]}$ available
- classical candidate: kernel density estimator

$$\widehat{\rho}_{h,T}(x) = \frac{1}{hT}\int_0^T K\left(\frac{x - X_t}{h}\right)\mathrm{d}t, \quad x \in \mathbb{R}$$

Controlling the sup-norm risk

Two approaches:

1. make use of specific structure of diffusions by employing local time and continuous martingale techniques (Aeckerle-Willems and Strauch, 2021)
2. use mixing properties to control the long-time transitional behavior and heat-kernel bounds on the transition density for the short time behavior (Dexheimer, Strauch, T., 2021+)

Controlling the sup-norm risk

Two approaches:

1. make use of specific structure of diffusions by employing local time and continuous martingale techniques (Aeckerle-Willems and Strauch, 2021)
2. use mixing properties to control the long-time transitional behavior and heat-kernel bounds on the transition density for the short time behavior (Dexheimer, Strauch, T., 2021+)

Both approaches allow to handle deviation inequalities and moment bounds for suprema of empirical processes of the form

$$\sup_{g \in \mathcal{G}} \Big| \underbrace{\frac{1}{\sqrt{T}} \int_0^T g(X_s)\,\mathrm{d}s}_{=:\mathbb{G}_T(g)} \Big|, \quad \mathcal{G} \subset L^\infty(\mathbb{R}),$$

via Talagrand's generic chaining device

Controlling the sup-norm risk

For (general) stationary, exponentially β-mixing Markov processes with inv. distribution μ, i.e., $\beta(t) = \int \|P_t(x,\cdot) - \mu\|_{TV}\, \mu(\mathrm{d}x) \lesssim \exp(-\kappa t)$, we obtain for $m_T \leqslant T/4, \tau \in [m_T, 2m_T]$,

$$\left(\mathbb{E}^\mu\left[\sup_{g\in\mathcal{G}} |\mathbb{G}_T(g)|^p\right]\right)^{1/p} \leqslant C_1 \int_0^\infty \log \mathcal{N}\big(u, \mathcal{G}, \tfrac{2m_T}{\sqrt{T}} d_\infty\big)\, \mathrm{d}u + C_2 \int_0^\infty \sqrt{\log \mathcal{N}(u, \mathcal{G}, d_{\mathbb{G},\tau})}\, \mathrm{d}u$$
$$+ 4 \sup_{g\in\mathcal{G}} \left(\frac{2m_T}{\sqrt{T}} \|g\|_\infty c_1 p + \|g\|_{\mathbb{G},\tau} c_2 \sqrt{p} + \frac{1}{2}\|g\|_\infty c_\kappa \sqrt{T} \mathrm{e}^{-\frac{\kappa m_T}{p}}\right),$$

where $d_{\mathbb{G},\tau}(f,g) = \mathrm{Var}(\mathbb{G}_\tau(f-g))$.

Controlling the sup-norm risk

For (general) stationary, exponentially β-mixing Markov processes with inv. distribution μ, i.e., $\beta(t) = \int \|P_t(x,\cdot) - \mu\|_{\mathrm{TV}}\, \mu(\mathrm{d}x) \lesssim \exp(-\kappa t)$, we obtain for $m_T \leqslant T/4, \tau \in [m_T, 2m_T]$,

$$\left(\mathbb{E}^\mu\left[\sup_{g\in\mathcal{G}} |\mathbb{G}_T(g)|^p\right]\right)^{1/p} \leqslant C_1 \int_0^\infty \log \mathcal{N}\big(u, \mathcal{G}, \tfrac{2m_T}{\sqrt{T}} d_\infty\big)\, \mathrm{d}u + C_2 \int_0^\infty \sqrt{\log \mathcal{N}(u, \mathcal{G}, d_{\mathbb{G},\tau})}\, \mathrm{d}u$$
$$+ 4 \sup_{g\in\mathcal{G}} \left(\frac{2m_T}{\sqrt{T}} \|g\|_\infty c_1 p + \|g\|_{\mathbb{G},\tau} c_2 \sqrt{p} + \frac{1}{2}\|g\|_\infty c_\kappa \sqrt{T} \mathrm{e}^{-\frac{\kappa m_T}{p}}\right),$$

where $d_{\mathbb{G},\tau}(f,g) = \mathrm{Var}(\mathbb{G}_\tau(f-g))$. With the decomposition

$$\mathbb{E}^\mu\left[\left\|\widehat{\rho}_{h,T} - \rho\right\|_{L^\infty(D)}\right] = \|\mathbb{E}^\mu[\widehat{\rho}_{h,T}(\cdot)] - \rho\|_{L^\infty(D)} + \mathbb{E}^\mu\left[\left\|\widehat{\rho}_{h,T} - \mathbb{E}^\mu[\widehat{\rho}_{h,T}(\cdot)]\right\|_{L^\infty(D)}\right] =: \mathbf{B} + \mathbf{V},$$

we can use the general result to bound the statistical error $\mathbf{V}$ via

$$\mathbf{V} = \frac{1}{\sqrt{Th}} \mathbb{E}^\mu\left[\sup_{g\in\mathcal{G}} |\mathbb{G}_T(g)|\right], \quad \mathcal{G} = \left\{K\big(\tfrac{x-\cdot}{h}\big) - \mu\big(K\big(\tfrac{x-\cdot}{h}\big)\big) : x \in D \cap \mathbb{Q}\right\}.$$

The bias $\mathbf{B}$ is taken care of by assuming $\rho|_D \in \text{Hölder}_D(\beta)$ such that $\mathbf{B} \lesssim h^\beta$.

Central statistical result for the diffusion case

Proposition (C., Strauch, T. (2021+))

Suppose that $\rho|_D \in \text{Hölder}_D(\beta+1)$, $\beta > 0$, and K be of order $\lfloor \beta + 1 \rfloor$. Under standard assumptions on b, σ guaranteeing

1. exponential ergodicity of the marginals, i.e., $\|P_t(x,\cdot) - \mu\|_{\mathrm{TV}} \lesssim V(x)\exp(-\kappa t)$;
2. heat kernel bound, i.e., $\sup_{x,y\in\mathbb{R}} p_t(x,y) \lesssim t^{-1/2}$, $t \in (0,1)$,

it follows for the smoothness independent choice $h \equiv h(T) = \log^2 T/\sqrt{T}$ that

$$\mathbb{E}^0\left[\left\|\widehat{\rho}_{h,T} - \rho\right\|_{L^\infty(D)}\right] \in \mathrm{O}\left(\sqrt{\log T/T}\right).$$

Statistical challenge in the Lévy case

- missing piece: estimator of generator function $f = \mathcal{A}_H\gamma$, where H is ascending ladder height process
- $H_t = X_{\mathsf{L}_t^{-1}}$, where L is local time at the supremum, i.e., local time of reflected process $Y_t = \sup_{s \leqslant t} X_s - X_t$ in 0

Problem

L can generally not be observed, even with a continuous record $X^T = (X_t)_{t\in[0,T]}$ available

Statistical challenge in the Lévy case

- missing piece: estimator of generator function $f = \mathcal{A}_H \gamma$, where H is ascending ladder height process
- $H_t = X_{\mathrm{L}_t^{-1}}$, where L is local time at the supremum, i.e., local time of reflected process $Y_t = \sup_{s \leqslant t} X_s - X_t$ in 0

Problem

L can generally not be observed, even with a continuous record $X^T = (X_t)_{t \in [0,T]}$ available

$\rightsquigarrow$ plug-in type estimator not feasible for $\mathcal{A}_H \gamma$

- problem formulation is ergodic in nature $\rightsquigarrow$ how can this be reflected in the statistical analysis, even though Lévy processes are not ergodic in time?

Estimation strategy

- Basic observation:

$$f(x) = \mathcal{A}_H\gamma(x) = d_H\gamma'(x) + \int_{0+}^{\infty} (\gamma(x+y) - \gamma(x))\,\Pi_H(\mathrm{d}y) = \int_0^{\infty} \mathbb{E}^0[X_1]\gamma'(x+y)\,\mu(\mathrm{d}y),$$

where

$$\mu(\mathrm{d}y) = \frac{1}{\mathbb{E}^0[H_1]}\big(d_H\delta_0(\mathrm{d}y) + \mathbf{1}_{(0,\infty)}(y)\Pi_H((y,\infty))\,\mathrm{d}y\big), \quad y \geqslant 0,$$

(note: $\mathbb{E}^0[H_1] = \mathbb{E}^0[X_1]$ by chosen scaling of local time such that $\mathbb{E}^0[\mathsf{L}_1^{-1}] = 1$)

Estimation strategy

- Basic observation:

$$f(x) = \mathcal{A}_H \gamma(x) = d_H \gamma'(x) + \int_{0+}^{\infty} (\gamma(x+y) - \gamma(x))\, \Pi_H(\mathrm{d}y) = \int_0^{\infty} \mathbb{E}^0[X_1] \gamma'(x+y)\, \mu(\mathrm{d}y),$$

  where

$$\mu(\mathrm{d}y) = \frac{1}{\mathbb{E}^0[H_1]} \Big(d_H \delta_0(\mathrm{d}y) + \mathbf{1}_{(0,\infty)}(y) \Pi_H((y,\infty))\, \mathrm{d}y \Big), \quad y \geqslant 0,$$

  (note: $\mathbb{E}^0[H_1] = \mathbb{E}^0[X_1]$ by chosen scaling of local time such that $\mathbb{E}^0[\mathrm{L}_1^{-1}] = 1$)
- μ is associated to overshoots of X: Let $\mathcal{O}_x \coloneqq X_{T_x} - x$ with $T_x \coloneqq \inf\{t \geqslant 0 : X_t > x\}$. Then, $\mathcal{O}_x \Longrightarrow \mu$ as $x \to \infty$.

Estimation strategy

- Basic observation:

$$f(x) = \mathcal{A}_H \gamma(x) = d_H \gamma'(x) + \int_{0+}^{\infty} (\gamma(x+y) - \gamma(x))\, \Pi_H(\mathrm{d}y) = \int_0^{\infty} \mathbb{E}^0[X_1]\gamma'(x+y)\, \mu(\mathrm{d}y),$$

  where

$$\mu(\mathrm{d}y) = \frac{1}{\mathbb{E}^0[H_1]}\Big(d_H \delta_0(\mathrm{d}y) + \mathbf{1}_{(0,\infty)}(y)\Pi_H((y,\infty))\,\mathrm{d}y\Big), \quad y \geqslant 0,$$

  (note: $\mathbb{E}^0[H_1] = \mathbb{E}^0[X_1]$ by chosen scaling of local time such that $\mathbb{E}^0[\mathrm{L}_1^{-1}] = 1$)
- μ is associated to overshoots of X: Let $\mathcal{O}_x \coloneqq X_{T_x} - x$ with $T_x \coloneqq \inf\{t \geqslant 0 : X_t > x\}$. Then, $\mathcal{O}_x \Longrightarrow \mu$ as $x \to \infty$.
- Even better: $\mathcal{O} = (\mathcal{O}_x)_{x \geqslant 0}$ is a Feller Markov process and μ is unique stationary distribution of $\mathcal{O}$

Estimation strategy

- Basic observation:

$$f(x) = \mathcal{A}_H\gamma(x) = d_H\gamma'(x) + \int_{0+}^{\infty} (\gamma(x+y) - \gamma(x))\, \Pi_H(\mathrm{d}y) = \int_0^{\infty} \mathbb{E}^0[X_1]\gamma'(x+y)\, \mu(\mathrm{d}y),$$

  where

$$\mu(\mathrm{d}y) = \frac{1}{\mathbb{E}^0[H_1]}\Big(d_H\delta_0(\mathrm{d}y) + \mathbf{1}_{(0,\infty)}(y)\Pi_H((y,\infty))\,\mathrm{d}y\Big), \quad y \geqslant 0,$$

  (note: $\mathbb{E}^0[H_1] = \mathbb{E}^0[X_1]$ by chosen scaling of local time such that $\mathbb{E}^0[\mathrm{L}_1^{-1}] = 1$)
- μ is associated to overshoots of X: Let $\mathcal{O}_x \coloneqq X_{T_x} - x$ with $T_x \coloneqq \inf\{t \geqslant 0 : X_t > x\}$. Then, $\mathcal{O}_x \Longrightarrow \mu$ as $x \to \infty$.
- Even better: $\mathcal{O} = (\mathcal{O}_x)_{x\geqslant 0}$ is a Feller Markov process and μ is unique stationary distribution of $\mathcal{O}$
- First step: consider spatial mean estimator

$$\widetilde{f}_Y(x) = \frac{1}{Y}\int_0^Y \mathbb{E}^0[X_1]\gamma'(x+\mathcal{O}_y)\,\mathrm{d}y,$$

  based on sample $(X_{T_y})_{y\in[0,Y]}$

Central convergence results

Theorem (Döring, T. (2021+))

Suppose that $\Pi|_{(a,b)} \ll \mathrm{Leb}|_{(a,b)}$ for some $(a, b) \subset \mathbb{R}_+$. Then, $\mathcal{O}_x \xrightarrow{\mathrm{TV}} \mu$ as $x \to \infty$. If moreover $\int_1^\infty \mathrm{e}^{\lambda x}\, \Pi(\mathrm{d}x) < \infty$, then X is even exponentially ergodic.

Central convergence results

Theorem (Döring, T. (2021+))

Suppose that $\Pi|_{(a,b)} \ll \mathrm{Leb}|_{(a,b)}$ for some $(a,b) \subset \mathbb{R}_+$. Then, $\mathcal{O}_x \xrightarrow{\mathrm{TV}} \mu$ as $x \to \infty$. If moreover $\int_1^\infty e^{\lambda x}\,\Pi(\mathrm{d}x) < \infty$, then X is even exponentially ergodic.

Use this together with uniform moment bounds on Markovian integral functionals to obtain the following for the mean estimator $\widetilde{f}_Y = \frac{1}{Y}\int_0^Y \mathbb{E}^0[X_1]\gamma'(x + \mathcal{O}_y)\,\mathrm{d}y$:

Proposition (C., Strauch, T. (2021+))

Given the above assumptions, we have for any bounded open set $D \subset \mathbb{R}$,

$$\mathbb{E}^0\left[\left\|\widetilde{f}_Y - f\right\|_{L^\infty(D)}\right] \lesssim \frac{1}{\sqrt{Y}}.$$

From spatial to temporal estimator

- So far: spatial mean estimator $\widetilde{f}_Y$ based on overshoots up to level Y

$\rightsquigarrow$ how to translate this into temporal estimator $\widehat{f}_T$ based on Lévy sample $(X_t)_{t\in[0,T]}$ up to time T?

From spatial to temporal estimator

- So far: spatial mean estimator $\widetilde{f}_Y$ based on overshoots up to level Y

$\rightsquigarrow$ how to translate this into temporal estimator $\widehat{f}_T$ based on Lévy sample $(X_t)_{t\in[0,T]}$ up to time T?

- good candidate: $\widehat{f}_T(x) = \frac{1}{X_T}\int_0^{X_T} \gamma'(x + \mathcal{O}_y)\,\mathrm{d}y \mathbf{1}_{(0,\infty)}(X_T)$.

From spatial to temporal estimator

- So far: spatial mean estimator $\widetilde{f}_Y$ based on overshoots up to level Y

$\rightsquigarrow$ how to translate this into temporal estimator $\widehat{f}_T$ based on Lévy sample $(X_t)_{t\in[0,T]}$ up to time T?

- good candidate: $\widehat{f}_T(x) = \frac{1}{X_T}\int_0^{X_T} \gamma'(x + \mathcal{O}_y)\,\mathrm{d}y \mathbf{1}_{(0,\infty)}(X_T)$.
- Then, for any $\varepsilon > 0$,

$$\mathbb{E}^0\Big[\big\|\widehat{f}_T - f\big\|_{L^\infty(D)}\Big] \lesssim \mathbb{E}^0\Big[\big\|\widetilde{f}_{\mathbb{E}^0[X_1]T} - f\big\|_{L^\infty(D)}\Big] + \frac{\varepsilon}{\mathbb{E}^0[X_1]} + \mathbb{P}^0\Big(\Big|\frac{X_T}{T} - \mathbb{E}^0[X_1]\Big| > \varepsilon\Big)$$
$$\leqslant \frac{1}{\sqrt{\mathbb{E}^0[X_1]T}} + \frac{\varepsilon}{\mathbb{E}^0[X_1]} + \mathbb{P}^0\Big(\Big|\frac{X_T}{T} - \mathbb{E}^0[X_1]\Big| > \varepsilon\Big).$$

$\rightsquigarrow$ need nonasymptotic controls on deviation from the mean

Convergence speed of LLN for Lévy processes

- Direct result: use Markov inequality + Burkholder–Davis–Gundy inequality in case $\mathbb{E}^0[|X_1|^p] < \infty$ for some $p \geqslant 2$ to obtain

$$\mathbb{P}^0\left(\left|\frac{X_T}{T} - \mathbb{E}^0[X_1]\right| > T^{-1/(2(1+p^{-1}))}\right) \leqslant T^{p/(1+2(1+p^{-1}))} T^{-p}\mathbb{E}^0\left[|X_T - \mathbb{E}^0[X_T]|^p\right]$$
$$\leqslant T^{-1/(2(1+p^{-1}))}.$$

Convergence speed of LLN for Lévy processes

- Direct result: use Markov inequality + Burkholder–Davis–Gundy inequality in case $\mathbb{E}^0[|X_1|^p] < \infty$ for some $p \geqslant 2$ to obtain

$$\mathbb{P}^0\left(\left|\frac{X_T}{T} - \mathbb{E}^0[X_1]\right| > T^{-1/(2(1+p^{-1}))}\right) \leqslant T^{p/(1+2(1+p^{-1}))} T^{-p} \mathbb{E}^0\left[|X_T - \mathbb{E}^0[X_T]|^p\right]$$
$$\leqslant T^{-1/(2(1+p^{-1}))}.$$

- Alternatively, if jumps are bounded, use Chernov-type argument + convenient form of char. exponent $\Psi(\theta) = \log \mathbb{E}^0[\exp(i\theta X_1)]$ to obtain

Theorem (C., Strauch, T. (2021+))

Suppose that X has bounded jumps. Then, there exists $\beta > 0$ and $T(p) > 0$ for $p > 0$ such that for any $T \geqslant T(p)$,

$$\mathbb{P}^0\left(|X_T - \mathbb{E}^0[X_T]| > \sqrt{\beta p T \log T}\right) \leqslant 2T^{-p/2}.$$

Putting the pieces together

Theorem (C., Strauch, T. (2021+))

Under regularity and exponential moment assumptions on $\Pi|_{(0,\infty)}$ such that $\mathcal{O}$ is exponentially ergodic, we have

$$\mathbb{E}^0\Big[\big\|\widehat{f}_T - f\big\|_{L^\infty(D)}\Big] \in \mathrm{O}\big(T^{-1/(2(1+p^{-1}))}\big),$$

provided that $\int_{|x|>1}|x|^p\,\Pi(\mathrm{d}x) < \infty$ for some $p \geqslant 2$ and

$$\mathbb{E}^0\Big[\big\|\widehat{f}_T - f\big\|_{L^\infty(D)}\Big] \in \mathrm{O}\Big(\sqrt{\log T/T}\Big),$$

if Π has bounded support. In particular, if $\theta^* \in D$, the regret of the greedy strategy $\widehat{\theta}_T = \sup_x \widehat{f}_T(x)$ for the Lévy control problem is of order $\mathrm{O}(T^{-1/(2(1+p^{-1}))})$ when X_1 has p-th moment and of order $\mathrm{O}(\sqrt{\log T/T})$ when jumps are bounded.

References

Aeckerle-Willems, C. and C. Strauch (2021). “Concentration of scalar ergodic diffusions and some statistical implications”. In: *Ann. Inst. Henri Poincaré Probab. Stat.*, to appear. arXiv: 1807.11331.

Alvarez, L. H. (2018). *A Class of Solvable Stationary Singular Stochastic Control Problems*. arXiv: 1803.03464.

Christensen, S. and T. Sohr (2020). “A solution technique for Lévy driven long term average impulse control problems”. In: *Stochastic Process. Appl.*

Christensen, S. and C. Strauch (2020). *Nonparametric learning for impulse control problems*. arXiv: 1909.09528.

Christensen, S., C. Strauch, and L. Trottner (2021). *Learning to reflect: A unifying approach for data-driven stochastic control strategies*. arXiv: 2104.11496.

Dexheimer, N., C. Strauch, and L. Trottner (2020). *Mixing it up: A general framework for Markovian statistics*. arXiv: 2011.00308.

Döring, L. and L. Trottner (2021). *Stability of Overshoots of Markov Additive Processes*. arXiv: 2102.03238.

Thank you!