Technische Universität Berlin

Optimal stopping with signatures

Sebastian Riedel | Technische Universität Berlin
joint work with C. Bayer (WIAS), P. Hager (TUB), and J. Schoenmakers (WIAS)

Outline

The setting

Signature stopping times and their optimality

Approximation

Outline

The setting

Signature stopping times and their optimality

Approximation

The optimal stopping problem

Goal: For a given càdlàg stochastic process $Y\colon [0,T] \to \mathbb{R}$, adapted to some right-continuous filtration $(\mathcal{F}_t)$, determine

$$\sup_{\tau \in \mathcal{S}[0,T]} \mathbb{E}[Y_\tau]$$

where the supremum ranges over all $(\mathcal{F}_t)$-stopping times τ with values in $[0,T]$.

The optimal stopping problem

Goal: For a given càdlàg stochastic process $Y\colon [0,T] \to \mathbb{R}$, adapted to some right-continuous filtration $(\mathcal{F}_t)$, determine

$$\sup_{\tau \in \mathcal{S}[0,T]} \mathbb{E}[Y_\tau]$$

where the supremum ranges over all $(\mathcal{F}_t)$-stopping times τ with values in $[0,T]$.

Motivation: Interesting in its own right, but also fundamental in pricing for American(-style) options on financial (or other) markets (Bensoussan '84, Karatzas '88).

Closed-form solution[1]

Define

$$Z_t^* = \operatorname*{ess\,sup}_{\tau \in \mathcal{S}[t,T]} \mathbb{E}(Y_\tau \mid \mathcal{F}_t).$$

[1] ...goes back to Fakeev '70, Bismut-Skalli '77.

Closed-form solution[1]

Define

$$Z_t^* = \operatorname*{ess\,sup}_{\tau \in \mathcal{S}[t,T]} \mathbb{E}(Y_\tau \mid \mathcal{F}_t).$$

One can prove that Z^* is the **Snell envelope** of Y, i.e. Z^* is the minimal right-continuous supermartingale that majorizes Y_t.

[1] ...goes back to Fakeev '70, Bismut-Skalli '77.

Closed-form solution[1]

Define

$$Z_t^* = \operatorname*{ess\,sup}_{\tau \in \mathcal{S}[t,T]} \mathbb{E}(Y_\tau \mid \mathcal{F}_t).$$

One can prove that Z^* is the **Snell envelope** of Y, i.e. Z^* is the minimal right-continuous supermartingale that majorizes Y_t. Moreover, the stopping times

$$\rho_t := \inf\{t \le s \le T \;:\; Z_s^* = Y_s\}, \quad t \in [0, T]$$

satisfy

$$\sup_{\tau \in \mathcal{S}[t,T]} \mathbb{E}(Y_\tau) = \mathbb{E}(Y_{\rho_t}).$$

In particular, ρ_0 solves the optimal stopping problem.

[1] ...goes back to Fakeev '70, Bismut-Skalli '77.

Closed-form solution[1]

Define

$$Z_t^* = \operatorname*{ess\,sup}_{\tau \in \mathcal{S}[t,T]} \mathbb{E}(Y_\tau \mid \mathcal{F}_t).$$

One can prove that Z^* is the **Snell envelope** of Y, i.e. Z^* is the minimal right-continuous supermartingale that majorizes Y_t. Moreover, the stopping times

$$\rho_t := \inf\{t \le s \le T \;:\; Z_s^* = Y_s\}, \quad t \in [0, T]$$

satisfy

$$\sup_{\tau \in \mathcal{S}[t,T]} \mathbb{E}(Y_\tau) = \mathbb{E}(Y_{\rho_t}).$$

In particular, ρ_0 solves the optimal stopping problem. However, this solution is of little use in practice. Calculating Z_0^* efficiently is a challenging problem, which has attracted many researchers' attention, and which is still actively studied (many (!!!) references missing...).

[1] ...goes back to Fakeev '70, Bismut-Skalli '77.

Remark. If $\mathcal{F}_t = \sigma(X_s \, : \, 0 \leq s \leq t)$, the solution of the optimal stopping problem depends on the law of the process (Y, X). The law of a stochastic process is uniquely determined by its **expected signature**. Therefore, the optimal stopping problem should have a reformulation in terms of the signature of a stochastic process.

Remark. If $\mathcal{F}_t = \sigma(X_s \,:\, 0 \leq s \leq t)$, the solution of the optimal stopping problem depends on the law of the process (Y, X). The law of a stochastic process is uniquely determined by its **expected signature**. Therefore, the optimal stopping problem should have a reformulation in terms of the signature of a stochastic process.

Notation.

Remark. If $\mathcal{F}_t = \sigma(X_s \,:\, 0 \leq s \leq t)$, the solution of the optimal stopping problem depends on the law of the process (Y, X). The law of a stochastic process is uniquely determined by its **expected signature**. Therefore, the optimal stopping problem should have a reformulation in terms of the signature of a stochastic process.

Notation.

– $\mathbb{X}$ denotes a continuous, geometric, and random p-rough path defined on $[0, T]$.

Remark. If $\mathcal{F}_t = \sigma(X_s \, : \, 0 \leq s \leq t)$, the solution of the optimal stopping problem depends on the law of the process (Y, X). The law of a stochastic process is uniquely determined by its **expected signature**. Therefore, the optimal stopping problem should have a reformulation in terms of the signature of a stochastic process.

Notation.

– $\mathbb{X}$ denotes a continuous, geometric, and random p-rough path defined on $[0, T]$.
– $\hat{\mathbb{X}}$ denotes the rough path $\mathbb{X}$ where we add time $t \mapsto t$ as the first component to the underlying path X.

Remark. If $\mathcal{F}_t = \sigma(X_s \,:\, 0 \le s \le t)$, the solution of the optimal stopping problem depends on the law of the process (Y, X). The law of a stochastic process is uniquely determined by its **expected signature**. Therefore, the optimal stopping problem should have a reformulation in terms of the signature of a stochastic process.

Notation.

- $\mathbb{X}$ denotes a continuous, geometric, and random p-rough path defined on $[0, T]$.
- $\hat{\mathbb{X}}$ denotes the rough path $\mathbb{X}$ where we add time $t \mapsto t$ as the first component to the underlying path X.
- $\hat{\mathbb{X}}^{<\infty}$ denotes the **signature** of $\hat{\mathbb{X}}$, i.e. the collection of all (formal) iterated integrals with values in the **extended tensor algebra** $T((\mathbb{R}^{1+d}))$, the dual of the **tensor algebra** $T(\mathbb{R}^{1+d})$.

Remark. If $\mathcal{F}_t = \sigma(X_s \,:\, 0 \le s \le t)$, the solution of the optimal stopping problem depends on the law of the process (Y, X). The law of a stochastic process is uniquely determined by its **expected signature**. Therefore, the optimal stopping problem should have a reformulation in terms of the signature of a stochastic process.

Notation.

- $\mathbb{X}$ denotes a continuous, geometric, and random p-rough path defined on $[0, T]$.
- $\hat{\mathbb{X}}$ denotes the rough path $\mathbb{X}$ where we add time $t \mapsto t$ as the first component to the underlying path X.
- $\hat{\mathbb{X}}^{<\infty}$ denotes the **signature** of $\hat{\mathbb{X}}$, i.e. the collection of all (formal) iterated integrals with values in the **extended tensor algebra** $T((\mathbb{R}^{1+d}))$, the dual of the **tensor algebra** $T(\mathbb{R}^{1+d})$.

Our setting. Assume $\mathcal{F}_t = \sigma(\hat{\mathbb{X}}|_{[0,s]} \,:\, 0 \le s \le t)$ and that $Y\colon [0, T] \to \mathbb{R}$ is $(\mathcal{F}_t)$-adapted and continuous. We aim to calculate

$$\sup_{\tau \in \mathcal{S}[0,T]} \mathbb{E}[Y_\tau].$$

The setting

Signature stopping times and their optimality

Approximation

Let $\hat{\Omega}^p_t$ denote a separable rough path space where $\hat{\mathbb{X}}$ restricted to $[0, t] \subset [0, T]$ takes its values. The space of **stopped rough paths** is defined as

$$\Lambda_T := \bigcup_{t \in [0,T]} \hat{\Omega}^p_t .$$

It can be seen that Λ_T is Polish[2].

[2] Λ_T is a rough paths version of a space considered in [Dupire; 2009]. For rough paths, the space was first used in [Kalsi, Lyons, Perez Arribas; 2020].

Let $\hat{\Omega}^p_t$ denote a separable rough path space where $\hat{\mathbb{X}}$ restricted to $[0, t] \subset [0, T]$ takes its values. The space of **stopped rough paths** is defined as

$$\Lambda_T := \bigcup_{t \in [0,T]} \hat{\Omega}^p_t.$$

It can be seen that Λ_T is Polish[2].

Definition

(i) We set $\mathcal{T} := \mathcal{C}(\Lambda_T, \mathbb{R})$ and call it the space of **continuous stopping policies**.

[2] Λ_T is a rough paths version of a space considered in [Dupire; 2009]. For rough paths, the space was first used in [Kalsi, Lyons, Perez Arribas; 2020].

Let $\hat{\Omega}^p_t$ denote a separable rough path space where $\hat{\mathbb{X}}$ restricted to $[0, t] \subset [0, T]$ takes its values.
The space of **stopped rough paths** is defined as

$$\Lambda_T := \bigcup_{t \in [0,T]} \hat{\Omega}^p_t .$$

It can be seen that Λ_T is Polish[2].

Definition

(i) We set $\mathcal{T} := \mathcal{C}(\Lambda_T, \mathbb{R})$ and call it the space of **continuous stopping policies**.

(ii) Let Z be a strictly positive random variable independent of $\mathbb{X}$. For $\theta \in \mathcal{T}$, we define the randomized stopping time

$$\tau^r_\theta := \inf \left\{ t \geq 0 \; : \; \int_0^{t \wedge T} \theta(\hat{\mathbb{X}}|_{[0,s]})^2 \, \mathrm{d}s \geq Z \right\} \quad (\inf \emptyset := +\infty).$$

[2] Λ_T is a rough paths version of a space considered in [Dupire; 2009]. For rough paths, the space was first used in [Kalsi, Lyons, Perez Arribas; 2020].

Proposition

Assume $\mathbb{E}[\sup_{t\in[0,T]}|Y_t|] < \infty$. *Then*

$$\sup_{\theta\in\mathcal{T}} \mathbb{E}[Y_{\tau^r_\theta \wedge T}] = \sup_{\tau\in\mathcal{S}[0,T]} \mathbb{E}[Y_\tau].$$

Proposition

Assume $\mathbb{E}[\sup_{t\in[0,T]} |Y_t|] < \infty$. *Then*

$$\sup_{\theta\in\mathcal{T}} \mathbb{E}[Y_{\tau^r_\theta \wedge T}] = \sup_{\tau\in\mathcal{S}[0,T]} \mathbb{E}[Y_\tau].$$

Proof.

- Let $\tau \in \mathcal{S}[0,T]$. From the Doob-Dynkin lemma (and some further work), there exists a Borel measurable $\theta \colon \Lambda_T \to \{0,1\}$ such that

$$\theta(\hat{\mathbb{X}}|_{[0,t]}) = \mathbb{1}_{\{\tau\leq t\}}.$$

Proposition

Assume $\mathbb{E}[\sup_{t\in[0,T]} |Y_t|] < \infty$*. Then*

$$\sup_{\theta\in\mathcal{T}} \mathbb{E}[Y_{\tau_\theta^r \wedge T}] = \sup_{\tau\in\mathcal{S}[0,T]} \mathbb{E}[Y_\tau].$$

Proof.

- Let $\tau \in \mathcal{S}[0,T]$. From the Doob-Dynkin lemma (and some further work), there exists a Borel measurable $\theta\colon \Lambda_T \to \{0,1\}$ such that

$$\theta(\hat{\mathbb{X}}|_{[0,t]}) = \mathbb{1}_{\{\tau\le t\}}.$$

- From Lusin's Theorem, we can find $\tilde{\theta}_n \in \mathcal{T}$, $0 \le \tilde{\theta}_n \le 1$, such that $\tilde{\theta}_n(\hat{\mathbb{X}}|_{[0,t]}) \to \mathbb{1}_{\{\tau\le t\}}$ almost surely w.r.t. $\lambda|_{[0,T]} \otimes \mathbb{P}$.

Proof.

– Setting $\theta_n := (2\tilde{\theta}_n)^n$ gives

$$\lim_{n\to\infty} \theta_n(\hat{\mathbb{X}}|_{[0,t]}) \to \begin{cases} +\infty & \text{if } t \geq \tau, \\ 0 & \text{if } t < \tau. \end{cases}$$

Therefore, $\tau^r_{\theta_n} \to \tau$ a.s. as $n \to \infty$ and

$$\sup_{\theta \in \mathcal{T}} \mathbb{E}[Y_{\tau^r_\theta \wedge T}] \geq \sup_{\tau \in \mathcal{S}[0,T]} \mathbb{E}[Y_\tau],$$

using Lebesgue's dominated convergence theorem.

□

The shuffle product

– The basis elements $e_{i_1} \otimes \cdots \otimes e_{i_n}$ in the tensor algebra $T(\mathbb{R}^{1+d})$ can be identified with the words $i_1 \cdots i_n$ composed by the letters $1, \ldots, 1+d$.

The shuffle product

- The basis elements $e_{i_1} \otimes \cdots \otimes e_{i_n}$ in the tensor algebra $T(\mathbb{R}^{1+d})$ can be identified with the words $i_1 \cdots i_n$ composed by the letters $1, \ldots, 1+d$.
- For two words u and v, we can consider the *shuffle product* $u \sqcup\!\sqcup\, v \in T(\mathbb{R}^{1+d})$. For example,

$$12 \sqcup\!\sqcup\, 3 = 123 + 132 + 312,$$

$$12 \sqcup\!\sqcup\, 24 = 2 \cdot 1224 + 1242 + 2124 + 2142 + 2412.$$

The shuffle product

- The basis elements $e_{i_1} \otimes \cdots \otimes e_{i_n}$ in the tensor algebra $T(\mathbb{R}^{1+d})$ can be identified with the words $i_1 \cdots i_n$ composed by the letters $1, \dots, 1+d$.
- For two words u and v, we can consider the *shuffle product* $u \sqcup\!\sqcup\, v \in T(\mathbb{R}^{1+d})$. For example,

$$12 \sqcup\!\sqcup\, 3 = 123 + 132 + 312,$$

$$12 \sqcup\!\sqcup\, 24 = 2 \cdot 1224 + 1242 + 2124 + 2142 + 2412.$$

- $\sqcup\!\sqcup$ can be bilinearly extended to define $l_1 \sqcup\!\sqcup\, l_2$ for every element $l_1, l_2 \in T(\mathbb{R}^d)$.

Definition

Let Z be a strictly positive random variable independent of $\mathbb{X}$. For $l \in T(\mathbb{R}^{1+d})$, we define the **randomized signature stopping time**

$$\tau_l^r := \inf\left\{t \geq 0 \,:\, \int_0^{t \wedge T} \langle l, \hat{\mathbb{X}}_{0,s}^{<\infty} \rangle^2 \, \mathrm{d}s \geq Z \right\}.$$

Lemma (Kalsi, Lyons, Perez Arribas '20)

Let μ be a probability measure on $(\hat{\Omega}^p_T, \mathcal{B}(\hat{\Omega}^p_T))$. Then, for every $\varepsilon > 0$, there is a compact set $\mathcal{K} \subset \hat{\Omega}^p_T$ such that:

1. $\mu(\mathcal{K}) > 1 - \varepsilon$.
2. *For every $\theta \in \mathcal{T}$ there is a sequence $l_n \in T(\mathbb{R}^{1+d})$ such that*

$$\sup_{\hat{\mathbb{X}} \in \mathcal{K};\ t \in [0,T]} \left| \langle l_n, \hat{\mathbb{X}}^{<\infty}_{0,t} \rangle - \theta(\hat{\mathbb{X}}|_{[0,t]}) \right| \to 0$$

   as $n \to \infty$.

Lemma (Kalsi, Lyons, Perez Arribas '20)

Let μ be a probability measure on $(\hat{\Omega}^p_T, \mathcal{B}(\hat{\Omega}^p_T))$. Then, for every $\varepsilon > 0$, there is a compact set $\mathcal{K} \subset \hat{\Omega}^p_T$ such that:

1. $\mu(\mathcal{K}) > 1 - \varepsilon$.
2. *For every $\theta \in \mathcal{T}$ there is a sequence $l_n \in T(\mathbb{R}^{1+d})$ such that*

$$\sup_{\hat{\mathbb{X}} \in \mathcal{K};\ t \in [0,T]} \left| \langle l_n, \hat{\mathbb{X}}^{<\infty}_{0,t} \rangle - \theta(\hat{\mathbb{X}}|_{[0,t]}) \right| \to 0$$

as $n \to \infty$.

Proof.

Stone-Weierstraß. To prove that $\mathcal{T}_{\mathrm{sig}}$ separates points, one needs the uniqueness result for the signature of a rough path proved in [Boedihardjo, Geng, Lyons, Yang; 2016]. □

Note: Convergence of the stopping policies does **not** imply convergence of the stopping times!

Note: Convergence of the stopping policies does **not** imply convergence of the stopping times!

Indeed, for $\vartheta, \vartheta_n \colon [0,3] \to [0,\infty)$ defined by

$$\vartheta(t) = \begin{cases} 1-t & \text{if } t \in [0,1] \\ 0 & \text{if } t \in [1,2] \\ t-2 & \text{if } t \in [2,3] \end{cases} \quad \text{and} \quad \vartheta_n(t) = \begin{cases} (1-\frac{1}{n})(1-t) & \text{if } t \in [0,1] \\ 0 & \text{if } t \in [1,2] \\ t-2 & \text{if } t \in [2,3]. \end{cases}$$

we have $\vartheta_n \to \vartheta$ as $n \to \infty$, but

$$\inf\left\{t \geq 0 \,:\, \int_0^{t\wedge 3} \vartheta(s)\,ds \geq \frac{1}{2}\right\} = 1 \qquad \text{and}$$
$$\inf\left\{t \geq 0 \,:\, \int_0^{t\wedge 3} \vartheta_n(s)\,ds \geq \frac{1}{2}\right\} > 2$$

for all $n \geq 1$.

Technische Universität Berlin

Lemma

Let F_Z denote the cumulative distribution function of Z. Then

$$\mathbb{E}(Y_{\tau_\theta \wedge T} \,|\, \hat{\mathbb{X}}) = \int_0^T Y_t \,\mathrm{d}\tilde{F}(t) + Y_T(1 - \tilde{F}(T)) = \int_0^T (1 - \tilde{F}(t)) \,\mathrm{d}Y_t + Y_0$$

where the second integral is a Young integral and

$$\tilde{F}(t) = F_Z\left(\int_0^t \theta(\hat{\mathbb{X}}|_{[0,s]})^2 \,\mathrm{d}s\right).$$

Lemma

Let F_Z denote the cumulative distribution function of Z. Then

$$\mathbb{E}(Y_{\tau_\theta \wedge T} \,|\, \hat{\mathbb{X}}) = \int_0^T Y_t \, \mathrm{d}\tilde{F}(t) + Y_T(1 - \tilde{F}(T)) = \int_0^T (1 - \tilde{F}(t)) \, \mathrm{d}Y_t + Y_0$$

where the second integral is a Young integral and

$$\tilde{F}(t) = F_Z\left(\int_0^t \theta(\hat{\mathbb{X}}|_{[0,s]})^2 \, \mathrm{d}s\right).$$

In particular, if Z has a density ϱ,

$$\mathbb{E}(Y_{\tau_\theta \wedge T}) = \mathbb{E}\left[\int_0^T Y_t \theta(\hat{\mathbb{X}}|_{[0,t]})^2 \varrho\left(\int_0^t \theta(\hat{\mathbb{X}}|_{[0,s]})^2 \, \mathrm{d}s\right) \mathrm{d}t + Y_T(1 - \tilde{F}(T))\right].$$

Proposition

Assume that Z has continuous density ϱ. Then

$$\sup_{\theta \in \mathcal{T}} \mathbb{E}(Y_{\tau^r_\theta \wedge T}) = \sup_{l \in T(\mathbb{R}^{1+d})} \mathbb{E}(Y_{\tau^r_l \wedge T}).$$

Proposition

Assume that Z has continuous density ϱ. Then

$$\sup_{\theta \in \mathcal{T}} \mathbb{E}\big(Y_{\tau^r_\theta \wedge T}\big) = \sup_{l \in T(\mathbb{R}^{1+d})} \mathbb{E}\big(Y_{\tau^r_l \wedge T}\big).$$

Definition

For $l \in T(\mathbb{R}^{1+d})$, we define the **signature stopping time**

$$\tau_l := \inf \left\{ t \geq 0 \,:\, \langle l, \hat{\mathbb{X}}^{<\infty}_{0,t} \rangle \geq 1 \right\}.$$

Proposition

Assume that Z has continuous density ϱ. Then

$$\sup_{\theta \in \mathcal{T}} \mathbb{E}\big(Y_{\tau_\theta^r \wedge T}\big) = \sup_{l \in T(\mathbb{R}^{1+d})} \mathbb{E}\big(Y_{\tau_l^r \wedge T}\big).$$

Definition

For $l \in T\big(\mathbb{R}^{1+d}\big)$, we define the **signature stopping time**

$$\tau_l := \inf \left\{ t \geq 0 \ : \ \langle l, \hat{\mathbb{X}}_{0,t}^{<\infty} \rangle \geq 1 \right\}.$$

Remark. Signature stopping times are hitting times of affine hyperplanes in $\bigoplus_{n=1}^{\infty} (\mathbb{R}^{1+d})^{\otimes n}$ of the process

$$t \mapsto (\hat{\mathbb{X}}_{0,t}^{(1)}, \hat{\mathbb{X}}_{0,t}^{(2)}, \ldots) \in \prod_{n=1}^{\infty} (\mathbb{R}^{1+d})^{\otimes n}.$$

Theorem (Bayer, Hager, R., Schoenmakers)

Assume that Z has a continuous density and that $\mathbb{E}[\sup_{t\in[0,T]} |Y_t|] < \infty$. Then

$$\sup_{\tau\in\mathcal{S}[0,T]} \mathbb{E}[Y_\tau] = \sup_{\theta\in\mathcal{T}} \mathbb{E}[Y_{\tau^r_\theta}] = \sup_{l\in T(\mathbb{R}^{1+d})} \mathbb{E}[Y_{\tau^r_l}] = \sup_{l\in T(\mathbb{R}^{1+d})} \mathbb{E}[Y_{\tau_l}].$$

Theorem (Bayer, Hager, R., Schoenmakers)

Assume that Z has a continuous density and that $\mathbb{E}[\sup_{t\in[0,T]}|Y_t|] < \infty$. Then

$$\sup_{\tau\in\mathcal{S}[0,T]} \mathbb{E}[Y_\tau] = \sup_{\theta\in\mathcal{T}} \mathbb{E}[Y_{\tau^r_\theta}] = \sup_{l\in T(\mathbb{R}^{1+d})} \mathbb{E}[Y_{\tau^r_l}] = \sup_{l\in T(\mathbb{R}^{1+d})} \mathbb{E}[Y_{\tau_l}].$$

Remark. If X is a Markov process in $\mathbb{R}^d$ and $Y_t = G(t, X_t)$ for a continuous function G, it is a classical result that

$$\sup_{\tau\in\mathcal{S}[0,T]} \mathbb{E}[Y_\tau] = \sup_{\tau\in\mathfrak{D}} \mathbb{E}[Y_{\tau\wedge T}]$$

where $\mathfrak{D}$ denotes the set of all hitting times of closed sets in $\mathbb{R}^{1+d}$ of the process $t \mapsto (t, X_t)$.

The setting

Signature stopping times and their optimality

Approximation

The following conclusion can be deduced from our former results:

Corollary

If $Z \sim \mathrm{Exp}(1)$ *and* $Y_0 = 0$,

$$\sup_{\tau \in \mathcal{S}[0,T]} \mathbb{E}[Y_\tau] = \sup_{l \in T(\mathbb{R}^{1+d})} \mathbb{E}\left[\int_0^T \exp\left(-\int_0^t \langle l, \hat{\mathbb{X}}^{<\infty}_{0,s}\rangle^2 \,\mathrm{d}s\right) \mathrm{d}Y_t\right].$$

Full linearisation

Assume that Y is the second component of X. Fix $l \in T(\mathbb{R}^{1+d})$. Then

$$\mathbb{E}\left[\int_0^T \exp\left(-\int_0^t \langle l, \hat{\mathbb{X}}_{0,s}^{<\infty}\rangle^2 \,\mathrm{d}s\right) \,\mathrm{d}Y_t\right] = \mathbb{E}\left[\int_0^T \exp\left(-\int_0^t \langle l \sqcup\!\sqcup\, l, \hat{\mathbb{X}}_{0,s}^{<\infty}\rangle \,\mathrm{d}s\right) \,\mathrm{d}Y_t\right]$$

Full linearisation

Assume that Y is the second component of X. Fix $l \in T(\mathbb{R}^{1+d})$. Then

$$\mathbb{E}\left[\int_0^T \exp\left(-\int_0^t \langle l, \hat{\mathbb{X}}^{<\infty}_{0,s}\rangle^2 \,\mathrm{d}s\right)\,\mathrm{d}Y_t\right] = \mathbb{E}\left[\int_0^T \exp\left(-\int_0^t \langle l \sqcup l, \hat{\mathbb{X}}^{<\infty}_{0,s}\rangle \,\mathrm{d}s\right)\,\mathrm{d}Y_t\right]$$
$$= \mathbb{E}\left[\int_0^T \exp\left(-\langle (l \sqcup l)\mathbf{1}, \hat{\mathbb{X}}^{<\infty}_{0,t}\rangle\right)\,\mathrm{d}Y_t\right]$$

Full linearisation

Assume that Y is the second component of X. Fix $l \in T(\mathbb{R}^{1+d})$. Then

$$\begin{aligned}\mathbb{E}\left[\int_0^T \exp\left(-\int_0^t \langle l, \hat{\mathbb{X}}^{<\infty}_{0,s}\rangle^2 \,\mathrm{d}s\right) \mathrm{d}Y_t\right] &= \mathbb{E}\left[\int_0^T \exp\left(-\int_0^t \langle l \sqcup\!\sqcup\, l, \hat{\mathbb{X}}^{<\infty}_{0,s}\rangle \,\mathrm{d}s\right) \mathrm{d}Y_t\right] \\ &= \mathbb{E}\left[\int_0^T \exp\left(-\langle (l \sqcup\!\sqcup\, l)\mathbf{1}, \hat{\mathbb{X}}^{<\infty}_{0,t}\rangle\right) \mathrm{d}Y_t\right] \\ &\stackrel{?}{=} \mathbb{E}\left[\int_0^T \langle \exp^{\sqcup\!\sqcup}(-(l \sqcup\!\sqcup\, l)\mathbf{1}), \hat{\mathbb{X}}^{<\infty}_{0,t}\rangle \,\mathrm{d}Y_t\right]\end{aligned}$$

Full linearisation

Assume that Y is the second component of X. Fix $l \in T(\mathbb{R}^{1+d})$. Then

$$\begin{aligned}
\mathbb{E}\left[\int_0^T \exp\left(-\int_0^t \langle l, \hat{\mathbb{X}}_{0,s}^{<\infty}\rangle^2 \, \mathrm{d}s\right) \, \mathrm{d}Y_t\right] &= \mathbb{E}\left[\int_0^T \exp\left(-\int_0^t \langle l \sqcup\!\sqcup\, l, \hat{\mathbb{X}}_{0,s}^{<\infty}\rangle \, \mathrm{d}s\right) \, \mathrm{d}Y_t\right] \\
&= \mathbb{E}\left[\int_0^T \exp\left(-\langle (l \sqcup\!\sqcup\, l)\mathbf{1}, \hat{\mathbb{X}}_{0,t}^{<\infty}\rangle\right) \, \mathrm{d}Y_t\right] \\
&\stackrel{?}{=} \mathbb{E}\left[\int_0^T \langle \exp^{\sqcup\!\sqcup}(-(l \sqcup\!\sqcup\, l)\mathbf{1}), \hat{\mathbb{X}}_{0,t}^{<\infty}\rangle \, \mathrm{d}Y_t\right] \\
&= \mathbb{E}\left[\langle \exp^{\sqcup\!\sqcup}(-(l \sqcup\!\sqcup\, l)\mathbf{1})\mathbf{2}, \hat{\mathbb{X}}_{0,T}^{<\infty}\rangle\right]
\end{aligned}$$

Full linearisation

Assume that Y is the second component of X. Fix $l \in T(\mathbb{R}^{1+d})$. Then

$$\begin{aligned}
\mathbb{E}\left[\int_0^T \exp\left(-\int_0^t \langle l, \hat{\mathbb{X}}^{<\infty}_{0,s}\rangle^2 \,\mathrm{d}s\right) \mathrm{d}Y_t\right] &= \mathbb{E}\left[\int_0^T \exp\left(-\int_0^t \langle l \sqcup\!\sqcup l, \hat{\mathbb{X}}^{<\infty}_{0,s}\rangle \,\mathrm{d}s\right) \mathrm{d}Y_t\right] \\
&= \mathbb{E}\left[\int_0^T \exp\left(-\langle (l \sqcup\!\sqcup l)\mathbf{1}, \hat{\mathbb{X}}^{<\infty}_{0,t}\rangle\right) \mathrm{d}Y_t\right] \\
&\overset{?}{=} \mathbb{E}\left[\int_0^T \langle \exp^{\sqcup\!\sqcup}(-(l \sqcup\!\sqcup l)\mathbf{1}), \hat{\mathbb{X}}^{<\infty}_{0,t}\rangle \,\mathrm{d}Y_t\right] \\
&= \mathbb{E}\left[\langle \exp^{\sqcup\!\sqcup}(-(l \sqcup\!\sqcup l)\mathbf{1})\mathbf{2}, \hat{\mathbb{X}}^{<\infty}_{0,T}\rangle\right] \\
&= \langle \exp^{\sqcup\!\sqcup}(-(l \sqcup\!\sqcup l)\mathbf{1})\mathbf{2}, \mathbb{E}[\hat{\mathbb{X}}^{<\infty}_{0,T}]\rangle.
\end{aligned}$$

Theorem (Bayer, Hager, R., Schoenmakers)

For $\kappa > 0$, *define*

$$S = S_\kappa = \inf\{t \geq 0 \;:\; \|\hat{\mathbb{X}}\|_{p-\mathrm{var};[0,t]} \geq \kappa\} \wedge T.$$

Then

$$\sup_{\tau \in \mathcal{S}[0,T]} \mathbb{E}[Y_\tau] = \lim_{\kappa \to \infty} \lim_{K \to \infty} \lim_{N \to \infty} \sup_{|l| + \deg(l) \leq K} \langle \exp^{\sqcup\!\sqcup}(-(l \sqcup\!\sqcup l)1)2, \mathbb{E}[\hat{\mathbb{X}}^{\leq N}_{0,S}] \rangle$$

where

$$\exp^{\sqcup\!\sqcup}(l) = \sum_{n=0}^{\infty} \frac{l^{\sqcup\!\sqcup n}}{n!}.$$

Theorem (Bayer, Hager, R., Schoenmakers)

For $\kappa > 0$*, define*

$$S = S_\kappa = \inf\{t \geq 0 \,:\, \|\hat{\mathbb{X}}\|_{p-\mathrm{var};[0,t]} \geq \kappa\} \wedge T.$$

Then

$$\sup_{\tau \in \mathcal{S}[0,T]} \mathbb{E}[Y_\tau] = \lim_{\kappa\to\infty} \lim_{K\to\infty} \lim_{N\to\infty} \sup_{|l|+\deg(l)\leq K} \langle \exp^{\sqcup\!\sqcup}(-(l \sqcup\!\sqcup l)1)2, \mathbb{E}[\hat{\mathbb{X}}^{\leq N}_{0,S}]\rangle$$

where

$$\exp^{\sqcup\!\sqcup}(l) = \sum_{n=0}^{\infty} \frac{l^{\sqcup\!\sqcup n}}{n!}.$$

Remark. Nice approach, but does not prove itself in practice, unfortunately.

Partial linearisation - Direct Monte-Carlo approach

Idea: Discretize

$$\mathbb{E}\left[\int_0^T \exp\left(-\int_0^t \langle l, \hat{\mathbb{X}}_{0,s}^{<\infty}\rangle^2 \, \mathrm{d}s\right) \, \mathrm{d}Y_t\right]$$

in time and use a direct Monte-Carlo approach together with gradient descent to approximate l.

3

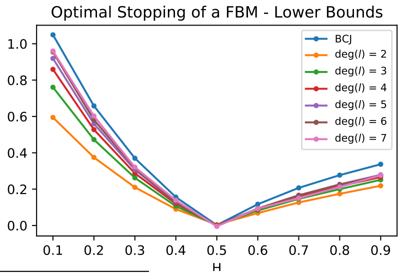
Partial linearisation - Direct Monte-Carlo approach

Idea: Discretize

$$\mathbb{E}\left[\int_0^T \exp\left(-\int_0^t \langle l, \hat{\mathbb{X}}^{<\infty}_{0,s}\rangle^2 \,\mathrm{d}s\right) \mathrm{d}Y_t\right]$$

in time and use a direct Monte-Carlo approach together with gradient descent to approximate l.

Example: Stopping a fractional Brownian motion[3]:



[3] Benchmark: [Becker, Cheridito, Jentzen; Deep optimal stopping; JMLR 2019].

Nonlinear approximation with neural networks

Corollary

$$\sup_{\tau \in \mathcal{S}[0,T]} \mathbb{E}[Y_\tau] = \sup_{\theta} \mathbb{E}\left[\int_0^T \exp\left(-\int_0^t \theta(\log^{\otimes} \hat{\mathbb{X}}^{<\infty}_{0,s})^2 \,\mathrm{d}s\right) \mathrm{d}Y_t\right].$$

where the supremum is taken over all continuous functions defined on the log-signature.

Nonlinear approximation with neural networks

Corollary

$$\sup_{\tau \in \mathcal{S}[0,T]} \mathbb{E}[Y_\tau] = \sup_{\theta} \mathbb{E}\left[\int_0^T \exp\left(-\int_0^t \theta(\log^{\otimes} \hat{\mathbb{X}}^{<\infty}_{0,s})^2 \, \mathrm{d}s\right) \mathrm{d}Y_t\right].$$

where the supremum is taken over all continuous functions defined on the log-signature.

To approximate θ, we use a neural network with ReLU activation function, 2 hidden layers, and $\mu_N + 20$ neurons on each layer where μ_N is the dimensionality of the truncated log-signature of level N.

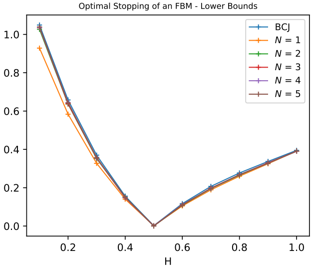
Nonlinear approximation with neural networks

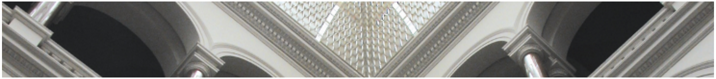
Corollary

$$\sup_{\tau \in \mathcal{S}[0,T]} \mathbb{E}[Y_\tau] = \sup_\theta \mathbb{E}\left[\int_0^T \exp\left(-\int_0^t \theta(\log^{\otimes} \hat{\mathbb{X}}^{<\infty}_{0,s})^2 \, \mathrm{d}s\right) \, \mathrm{d}Y_t\right].$$

where the supremum is taken over all continuous functions defined on the log-signature.

To approximate θ, we use a neural network with ReLU activation function, 2 hidden layers, and $\mu_N + 20$ neurons on each layer where μ_N is the dimensionality of the truncated log-signature of level N.





Thank you.