Signatures methods in finance

Christa Cuchiero

University of Vienna

Mini course

Paris, Bachelier Seminar, March 2024

Data driven risk inference

Why has modern machine learning...

... already quite successfully entered the world of dynamic stochastic modeling, mathematical finance and risk assessement?

Data driven risk inference

Why has modern machine learning...

... already quite successfully entered the world of dynamic stochastic modeling, mathematical finance and risk assessement?

- One answer: it enables realistic data driven (financial) risk inference to exploit upside risks and reduce downside risks.

Data driven risk inference

Why has modern machine learning...

... already quite successfully entered the world of dynamic stochastic modeling, mathematical finance and risk assessement?

- One answer: it enables realistic data driven (financial) risk inference to exploit upside risks and reduce downside risks.
- What is inference?

Data driven risk inference

Why has modern machine learning...

... already quite successfully entered the world of dynamic stochastic modeling, mathematical finance and risk assessement?

- One answer: it enables realistic data driven (financial) risk inference to exploit upside risks and reduce downside risks.
- What is inference? a conclusion reached on both, evidence and reasoning.

Data driven risk inference

Why has modern machine learning...

... already quite successfully entered the world of dynamic stochastic modeling, mathematical finance and risk assessement?

- One answer: it enables realistic data driven (financial) risk inference to exploit upside risks and reduce downside risks.
- What is inference? a conclusion reached on both, evidence and reasoning.

Evidence = Data

- Time series data
- Derivatives' price data
- Macro economic data
- ...

Data driven risk inference

Why has modern machine learning...

... already quite successfully entered the world of dynamic stochastic modeling, mathematical finance and risk assessement?

- One answer: it enables realistic data driven (financial) risk inference to exploit upside risks and reduce downside risks.
- What is inference? a conclusion reached on both, evidence and reasoning.

Evidence = Data

- Time series data
- Derivatives' price data
- Macro economic data
- ...

Reasoning

- Recognition of universal structures (statistics)
- First principles, e.g. no arbitrage
- Universal model classes and strategies

Data driven risk inference

Why has modern machine learning...

... already quite successfully entered the world of dynamic stochastic modeling, mathematical finance and risk assessement?

- One answer: it enables realistic data driven (financial) risk inference to exploit upside risks and reduce downside risks.
- What is inference? a conclusion reached on both, evidence and reasoning.

Evidence = Data

- Time series data
- Derivatives' price data
- Macro economic data
- ...

Reasoning

- Recognition of universal structures (statistics)
- First principles, e.g. no arbitrage
- Universal model classes and strategies

Combining machine learning with theory from mathematical finance allows to conciliate both sides - modeling as close as possible to high dimensional data while obeying well established principles.

Machine learning ingredients

1. Highly over parameterized and/or randomly initialized universal model classes serving as regression bases. Examples include
   - (random) signature to approximate paths functionals;
   - artificial neural networks to approximate functions (also on infinite spaces);
   - kernel methods, etc.
   - (physical) reservoirs of dynamical systems;

Machine learning ingredients

1. Highly over parameterized and/or randomly initialized universal model classes serving as regression bases. Examples include
   - (random) signature to approximate paths functionals;
   - artificial neural networks to approximate functions (also on infinite spaces);
   - kernel methods, etc.
   - (physical) reservoirs of dynamical systems;

Machine learning ingredients

1. Highly over parameterized and/or randomly initialized universal model classes serving as regression bases. Examples include
   - (random) signature to approximate paths functionals;
   - artificial neural networks to approximate functions (also on infinite spaces);
   - kernel methods, etc.
   - (physical) reservoirs of dynamical systems;
2. Optimization criterion coming with a
   - a loss function taylored to the problem, e.g., a calibration functional to match financial market data
   - certain metrics (e.g. generative adversarial distances).

Machine learning ingredients

1. Highly over parameterized and/or randomly initialized universal model classes serving as regression bases. Examples include
   - (random) signature to approximate paths functionals;
   - artificial neural networks to approximate functions (also on infinite spaces);
   - kernel methods, etc.
   - (physical) reservoirs of dynamical systems;
2. Optimization criterion coming with a
   - a loss function taylored to the problem, e.g., a calibration functional to match financial market data
   - certain metrics (e.g. generative adversarial distances).
3. Algorithm used for training, typically
   - (stochastic) gradient type algorithms;
   - linear regression methods (if the regression basis is linear);
   - tools from convex (quadratic) optimization (if the problem allows for such a formulation).

Focusing on signature

- We focus here on signature of some underlying stochastic process, used as linear regression basis for path functionals allowing to build
  1. universal strategies for optimal control problems comprising portfolio optimization, hedging, optimal execution, optimal stopping, etc.

Focusing on signature

- We focus here on signature of some underlying stochastic process, used as linear regression basis for path functionals allowing to build
  1. universal strategies for optimal control problems comprising portfolio optimization, hedging, optimal execution, optimal stopping, etc.
  2. universal model classes of dynamic processes, being able to approximate classical financial models.

Focusing on signature

- We focus here on signature of some underlying stochastic process, used as linear regression basis for path functionals allowing to build
  1. universal strategies for optimal control problems comprising portfolio optimization, hedging, optimal execution, optimal stopping, etc.
  2. universal model classes of dynamic processes, being able to approximate classical financial models.
- The optimization criteria and loss functions depend on the problem and include
  - maximizing expected utility;
  - minimizing a risk measure;
  - maximizing over stopping times e.g. for pricing American options;
  - minimizing certain distances to time series and option price data $\Rightarrow$ calibration functionals.

Focusing on signature

- We focus here on signature of some underlying stochastic process, used as linear regression basis for path functionals allowing to build
  1. universal strategies for optimal control problems comprising portfolio optimization, hedging, optimal execution, optimal stopping, etc.
  2. universal model classes of dynamic processes, being able to approximate classical financial models.
- The optimization criteria and loss functions depend on the problem and include
  - maximizing expected utility;
  - minimizing a risk measure;
  - maximizing over stopping times e.g. for pricing American options;
  - minimizing certain distances to time series and option price data $\Rightarrow$ calibration functionals.
- As the regression basis is linear, many problems reduce to linear regression or convex quadratic optimization problems.

Signature in data science - application areas

The importance of signature methods in machine learning and data science has steadily increased: they have been employed

- as feature maps for classification tasks related to streamed data (see, e.g., I. Chevyrev & A. Kormilitzin ('16))
- for Chinese character recognition (see B. Graham ('13))

Fig. 3. Handwritten examples of the 52 Chinese radicals. The actual label

- for machine learning models for psychiatric diagnosis (Y. Wue et al. ('22))
- for time series generation (see, e.g., N. Hao ('23))
- for image recognition: 2D signature (see, e.g., I. Horozov ('15), M. Ibrahim & T. Lyons ('21), D. Lee & H. Oberhauser ('23), J. Diehl ('24) et al.)
- in the context of signature SDEs to obtain universal model classes of dynamic processes, (to approximate classical financial models; see, e.g., I. Perez Arribas et al. ('20), C.C., G.Gazzani & S.Svaluto-Ferro ('22))
- to obtain universal strategies for optimal control problems (see, e.g., Kalsi et al.('19), Bayer et al. ('21))

Themes of this course

Part I Review and overview of the theory of signature

- Review of signature in a semimartingale setup
- Global universal approximation property of linear functions of the signature on weighted spaces

Themes of this course

Part I Review and overview of the theory of signature

- Review of signature in a semimartingale setup
- Global universal approximation property of linear functions of the signature on weighted spaces

Part II Signature methods in Stochastic Portfolio Theory (SPT)

- Introduction to SPT
- Signature-type portfolios
- Optimization tasks and approximation results
- Numerical results on simulated and real market data

Themes of this course

Part III An affine and polynomial perspective to signature based models

- An overview of affine and polynomial processes by means of Lévy's stochastic area formula
- Signature Stochastic Differential Equations (SDEs) from an affine and polynomial perspective

Themes of this course

Part III An affine and polynomial perspective to signature based models

- An overview of affine and polynomial processes by means of Lévy's stochastic area formula
- Signature Stochastic Differential Equations (SDEs) from an affine and polynomial perspective

Part IV Signature based asset price models for the joint calibration problem to SPX and VIX options

- Theoretical model ingredients and pricing of VIX and SPX options
- Calibration results on market data

Themes of this course

Part III **An affine and polynomial perspective to signature based models**

- An overview of affine and polynomial processes by means of Lévy's stochastic area formula
- Signature Stochastic Differential Equations (SDEs) from an affine and polynomial perspective

Part IV **Signature based asset price models for the joint calibration problem to SPX and VIX options**

- Theoretical model ingredients and pricing of VIX and SPX options
- Calibration results on market data

Part V **Signature of càdlàg rough paths, functional Itô formula and Taylor expansions**

- Marcus signature for càdlàg rough paths
- Functional Itô formula and Taylor expansions for non-anticipative maps of càdlàg rough paths

Part I

Review and overview of the theory of signature

- partly based on a course given jointly with Sara Svaluto-Ferro
- partly based on joint work with Philipp Schmocker and Josef Teichmann,
  C. Cuchiero, P. Schmocker and J. Teichmann, Global universal approximation of functional input maps on weighted spaces, 2023, https://arxiv.org/abs/2306.03303

Semimartingales as rough paths

- The most important class of stochastic processes in finance are semimartingales. We would therefore like to define their signature in line with the theory of (weakly) geometric rough paths.
- Let $\alpha \in (\frac{1}{2}, \frac{1}{3})$. Then for semimartingales with a.s. α-Hölder continuous trajectories, this can be realized via the Stratonovich lift which is a.s. a weakly geometric α-Hölder rough path.
- Denote by $\mathcal{C}_g^\alpha([0, T], \mathbb{R}^d)$ the set of weakly geometric α-Hölder rough paths.

Semimartingales as rough paths

- The most important class of stochastic processes in finance are semimartingales. We would therefore like to define their signature in line with the theory of (weakly) geometric rough paths.
- Let $\alpha \in (\frac{1}{2}, \frac{1}{3})$. Then for semimartingales with a.s. α-Hölder continuous trajectories, this can be realized via the Stratonovich lift which is a.s. a weakly geometric α-Hölder rough path.
- Denote by $\mathcal{C}_g^\alpha([0,T], \mathbb{R}^d)$ the set of weakly geometric α-Hölder rough paths.

Proposition

Let $\alpha \in (\frac{1}{2}, \frac{1}{3})$ and X be a continuous $\mathbb{R}^d$-valued semimartingale and $[X,X]^c$ its $(\mathbb{R}^d)^{\otimes 2}$-valued continuous quadratic variation. Then, $\mathbf{X}(\omega) = (X(\omega), \mathbb{X}^{(2)}(\omega)) \in \mathcal{C}_g^\alpha([0,T], \mathbb{R}^d)$ a.s., where, for $0 \leq s \leq t \leq T$,

$$\mathbb{X}^{(2)}_{s,t} := \int_s^t X_{s,r} \otimes dX_r + \frac{1}{2}[X,X]^c_{s,t} = \int_s^t X_{s,r} \otimes \circ dX_r$$

and the first integral is understood in Itô's sense and the second in Stratonovich

Signature Stratonovich SDE

Proposition

Let X be a continuous $\mathbb{R}^d$-valued semimartingale and $\mathbf{X}$ its Stratonovich lift. Then its unique Lyon's extension (used to define the signature for weakly geometric rough paths), denoted by $\mathbb{X}$, coincides a.s. with the following $G((\mathbb{R}^d))$-valued Stratonovich SDE

$$d\mathbb{X}_{s,t} = \mathbb{X}_{s,t} \otimes \circ dX_t, \quad \mathbb{X}_{s,s} = (1, 0, 0, \dots) \in G((\mathbb{R}^d)).$$

The explicit solution of this SDE are simply the interated integrals in Stratonovich sense, collected in the $G((\mathbb{R}^d))$ valued object

$$\mathbb{X}_{s,t} = 1 + \int_s^t \mathbb{X}_{s,r} \otimes \circ dX_r,$$

which in coordinate form, for a multi-index $I = (i_1, \dots, i_n)$, reads as

$$\mathbb{X}^{(n)}_{s,t;I} := \int_s^t \int_s^{u_n} \cdots \int_s^{u_2} dX^{i_1}_{u_1} \circ \cdots \circ dX^{i_n}_{u_n} \in \mathbb{R}.$$

Signature of continuous $\mathbb{R}^d$-valued semimartingales

- Hence the signature of an $\mathbb{R}^d$-valued continuous semimartingale X can be defined via

$$\mathbb{X}_{s,t} := \Big(1, \int_s^t \circ dX_s, \int_s^t \int_s^{u_2} \circ dX_{u_1} \otimes \circ dX_{u_2}, \ldots, \\ \cdots \int_s^t \int_s^{u_n} \cdots \int_s^{u_2} \circ dX_{u_1} \otimes \cdots \otimes \circ dX_{u_n}, \ldots \Big).$$

Signature of continuous $\mathbb{R}^d$-valued semimartingales

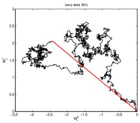
- Hence the signature of an $\mathbb{R}^d$-valued continuous semimartingale X can be defined via

$$\mathbb{X}_{s,t} := \Big(1, \int_s^t \circ dX_s, \int_s^t \int_s^{u_2} \circ dX_{u_1} \otimes \circ dX_{u_2}, \ldots, \\ \cdots \int_s^t \int_s^{u_n} \cdots \int_s^{u_2} \circ dX_{u_1} \otimes \cdots \otimes \circ dX_{u_n}, \ldots \Big).$$

- Visualizer of signature:
`https://zhy0.com/signature-visualizer/`

Geometric properties

- Consider the signature of order 2, i.e. $\mathbb{X}^{(2)}$. Then the Stratonovich product rule implies $Sym(\mathbb{X}^{(2)}_{s,t}) = \frac{1}{2}(X_t - X_s) \otimes (X_t - X_s)$, whence the symmetric part of $\mathbb{X}^{(2)}$ is fully determined by $\mathbb{X}^{(1)}_{s,t} = X_t - X_s$.
- To get rid of this redundancy one could only consider $Anti(\mathbb{X}^{(2)})$ given by $Anti(\mathbb{X}^{(2)}_{s,t})^{i,j} = \frac{1}{2}\left(\int_s^t (X^i_{s,u} - X^i_{s,s})dX^j_u - \int_s^t (X^j_{s,u} - X^j_{s,s})dX^i_u\right)$.
- This is the area (with orientation taken into account) between the curve $\{(X^i_u, X^j_u) : u \in [s,t]\}$ and the chord from (X^i_s, X^j_s) to (X^i_t, X^j_t).



- These properties imply that the correct state space for $\mathbb{X}^2$ is $G^2(\mathbb{R}^d)$, the free-step-2-nilpotent Lie group.
- We therefore will often view the Stratonovich lift $\mathbf{X} = \mathbb{X}^2 = (1, X, \mathbb{X}^{(2)})$ as stochastic process with values in $G^2(\mathbb{R}^d)$.

Linear functions of signature and universal approximation

- For a multi-index $I = \{i_1, ..., i_m\} \in \{1, \dots, d\}^m$ we denote by $\epsilon_I := \epsilon_{i_1} \otimes ... \otimes \epsilon_{i_m}$ the basis elements of $(\mathbb{R}^d)^{\otimes m}$.
- We call

$$L(\mathbb{X}_{s,t}) = \sum_{0 \leq |I| \leq n} \alpha_I \langle \epsilon_I, \mathbb{X}_{s,t} \rangle \text{ for } n \in \mathbb{N}$$

with $\alpha_I \in \mathbb{R}$ linear functions of the signature.

Linear functions of signature and universal approximation

- For a multi-index $I = \{i_1, ..., i_m\} \in \{1, \dots, d\}^m$ we denote by $\epsilon_I := \epsilon_{i_1} \otimes ... \otimes \epsilon_{i_m}$ the basis elements of $(\mathbb{R}^d)^{\otimes m}$.
- We call

$$L(\mathbb{X}_{s,t}) = \sum_{0 \leq |I| \leq n} \alpha_I \langle \epsilon_I, \mathbb{X}_{s,t} \rangle \text{ for } n \in \mathbb{N}$$

  with $\alpha_I \in \mathbb{R}$ linear functions of the signature.

Key properties to obtain a Universal Approximation Theorem (UAT) for linear functions of the signature

- Point-separation: for $(\hat{X}_t)_{t \geq 0} := (t, X_t)_{t \geq 0}$, its signature $\hat{\mathbb{X}}_{s,t}$ determines $\hat{X}_{[s,t]}$ uniquely.
- Algebra: the product of linear functions of the signature is again a linear function of the signature, precisely $\langle \epsilon_I, \mathbb{X}_{s,t} \rangle \langle e_J, \mathbb{X}_{s,t} \rangle = \langle \epsilon_I \sqcup\!\sqcup \epsilon_J, \mathbb{X}_{s,t} \rangle$.

$\Rightarrow$ Use the Stone-Weierstrass Theorem to approximate continuous (with respect to the α-Hölder norm) path functionals $f(X_{[0,t]})$ via $L(\hat{\mathbb{X}}_{0,t})$ uniformly in time on and compact sets of paths.

Towards a global UAT on a weighted space

based on joint work with P. Schmocker and J. Teichmann ('23)

- For applications approximation on compacts is often unsatisfactory in particular in stochastic setups.
- In Chevyrev & Oberhauser ('22), the strict topology going back to Giles ('71) is used to go beyond compact sets of paths.
- As one needs to work with bounded continuous functions a so-called tensor normalization has to be introduced to make signature bounded.
- This, however, destroys many tracatability properties of signature, e.g. in view of expected signatures.

⇒ Goal: global approximation result for linear functions of the signature (without normalization) for functions defined on a weighted space, corresponding to appropriate generalizations of continuous functions on paths spaces.

⇒ Tool: weighted Stone-Weierstrass theorem

Towards a global UAT on a weighted space - setup

based on joint work with P. Schmocker and J. Teichmann ('23)

- For $\alpha \in (1/3, 1/2)$ we consider the following path space of Hölder continuous maps

$$\widehat{C}_o^\alpha([0,T]; G^2(\mathbb{R}^{d+1})) := \left\{ \hat{\mathbb{X}}^2_{[0,T]} \in C_o^\alpha([0,T]; G^2(\mathbb{R}^{d+1})) : \hat{X}_t = (t, X_t), t \in [0,T], X_0 = 0 \right\}.$$

Here, $G^2(\mathbb{R}^{d+1})$ denotes the free-step-2-nilpotent Lie group where $\hat{\mathbb{X}}^2$ takes values. Moreover, $\hat{\mathbb{X}}^2_{[0,T]}$ denotes a path in $\widehat{C}_o^\alpha([0,T]; G^2(\mathbb{R}^{d+1}))$ (here, not necessarily induced by a semimartingale).

- We equip the space with an α-Hölder norm adapted to the group structure (Carnot-Caratheodory norm), denoted by $\| \cdot \|_{CC,\alpha}$.
- As topology on $\widehat{C}_o^\alpha([0,T]; G^2(\mathbb{R}^{d+1}))$ we however consider the weaker $C^{\alpha'}$-topology for $0 \leq \alpha' < \alpha$ or the weak-$*$-topology.

Towards a global UAT on a weighted space - setup

based on joint work with P. Schmocker and J. Teichmann ('23)

- We choose a weight function $\psi = \exp(\beta \| \cdot \|^{\gamma}_{CC,\alpha})$ for $\beta > 0$ and $\gamma > 2$. Then for both topologies, $(\widehat{C}^{\alpha}_{o}([0,T]; G^2(\mathbb{R}^{d+1})), \psi)$ becomes a weighted space, i.e. every pre-image

$$K_R := \psi^{-1}(]0,R]) = \left\{ \hat{\mathbb{X}}^2_{[0,T]} \in \widehat{C}^{\alpha}_{o}([0,T]; G^2(\mathbb{R}^{d+1})) : \psi(\hat{\mathbb{X}}^2_{[0,T]}) \leq R \right\}$$

is compact for all $R > 0$.

Towards a global UAT on a weighted space - setup

based on joint work with P. Schmocker and J. Teichmann ('23)

- We choose a weight function $\psi = \exp(\beta \| \cdot \|^{\gamma}_{CC,\alpha})$ for $\beta > 0$ and $\gamma > 2$. Then for both topologies, $(\widehat{C}^{\alpha}_{o}([0,T]; G^2(\mathbb{R}^{d+1})), \psi)$ becomes a weighted space, i.e. every pre-image

$$K_R := \psi^{-1}(]0, R]) = \left\{ \hat{\mathbb{X}}^2_{[0,T]} \in \widehat{C}^{\alpha}_{o}([0,T]; G^2(\mathbb{R}^{d+1})) : \psi(\hat{\mathbb{X}}^2_{[0,T]}) \leq R \right\}$$

  is compact for all $R > 0$.

- Define functions on this weighted space via
$B_{\psi} = \left\{ f : \widehat{C}^{\alpha}_{o}([0,T]; G^2(\mathbb{R}^{d+1})) \to \mathbb{R} : \sup_{\hat{\mathbb{X}}^2_{[0,T]} \in \widehat{C}^{\alpha}_{o}} \frac{|f(\hat{\mathbb{X}}^2_{[0,T]})|}{\psi(\hat{\mathbb{X}}^2_{[0,T]})} < \infty \right\}$, i.e. functions which are controlled by the growth of the weight function ψ. We work on $\mathcal{B}_{\psi}$ defined as the $\| \cdot \|_{\mathcal{B}_{\psi}(X)}$-closure of C_b-functions in B_{ψ}.

Towards a global UAT on a weighted space - setup

based on joint work with P. Schmocker and J. Teichmann ('23)

- We choose a weight function $\psi = \exp(\beta \| \cdot \|^{\gamma}_{CC,\alpha})$ for $\beta > 0$ and $\gamma > 2$. Then for both topologies, $(\widehat{C}^{\alpha}_{o}([0,T]; G^2(\mathbb{R}^{d+1})), \psi)$ becomes a weighted space, i.e. every pre-image

$$K_R := \psi^{-1}(]0, R]) = \left\{ \hat{\mathbb{X}}^2_{[0,T]} \in \widehat{C}^{\alpha}_{o}([0,T]; G^2(\mathbb{R}^{d+1})) : \psi(\hat{\mathbb{X}}^2_{[0,T]}) \leq R \right\}$$

  is compact for all $R > 0$.

- Define functions on this weighted space via
  $B_{\psi} = \left\{ f : \widehat{C}^{\alpha}_{o}([0,T]; G^2(\mathbb{R}^{d+1})) \to \mathbb{R} : \sup_{\hat{\mathbb{X}}^2_{[0,T]} \in \widehat{C}^{\alpha}_{o}} \frac{|f(\hat{\mathbb{X}}^2_{[0,T]})|}{\psi(\hat{\mathbb{X}}^2_{[0,T]})} < \infty \right\}$, i.e. functions which are controlled by the growth of the weight function ψ. We work on $\mathcal{B}_{\psi}$ defined as the $\| \cdot \|_{\mathcal{B}_{\psi}(X)}$-closure of C_b-functions in B_{ψ}.

- We can then apply a weighted version of the Stone-Weierstrass theorem to obtain a global UAT for linear functions of the signature approximating $\mathcal{B}_{\psi}-$functions.

Weighted Stone-Weierstrass Theorem for $\mathcal{B}_\psi$

based on joint work with P. Schmocker and J. Teichmann ('23)

For the weighted version of the Stone-Weierstrass theorem we need additionally to point separation and the algebra property a growth condition.

Definition

A subalgebra $\mathcal{A} \subset \mathcal{B}_\psi$ is called point separating and of moderate growth if there exists a point separating vector subspace $\widetilde{\mathcal{A}} \subseteq \mathcal{A}$ s.t. $x \mapsto \exp(|\widetilde{a}(x)|) \in \mathcal{B}_\psi$, for all $\widetilde{a} \in \widetilde{\mathcal{A}}$.

Theorem (C.C., P. Schmocker & J. Teichmann ('23))

Let $\mathcal{A} \subset \mathcal{B}_\psi$ be a subalgebra, that is point separating and of moderate growth and vanishes nowhere. Then $\mathcal{A}$ is dense in $\mathcal{B}_\psi$.

Global UAT for linear functions of the signature on $\mathcal{B}_\psi$

based on joint work with P. Schmocker and J. Teichmann ('23)

Recall that $\psi = \exp(\beta \| \cdot \|^{\gamma}_{CC,\alpha})$ for $\beta > 0$ and $\gamma > 2$.

Theorem (C.C., P. Schmocker, J. Teichmann ('23))

The linear span of the set $\left\{ \hat{\mathbb{X}}^2_{[0,T]} \mapsto \langle \epsilon_I, \hat{\mathbb{X}}_{0,T} \rangle : I \in \{0, \dots, d\}^m, m \in \mathbb{N} \right\}$ *is dense in* $\mathcal{B}_\psi$*, i.e. for every* $f \in \mathcal{B}_\psi$ *and* $\varepsilon > 0$ *there exists a linear function* L *of the signature (at time* T*) such that*

$$\sup_{\hat{\mathbb{X}}^2_{[0,T]} \in \widehat{C}^\alpha_o} \frac{\left| f(\hat{\mathbb{X}}^2_{[0,T]}) - L(\hat{\mathbb{X}}_{0,T}) \right|}{\psi(\hat{\mathbb{X}}^2_{[0,T]})} < \varepsilon.$$

Global UAT for linear functions of the signature on $\mathcal{B}_\psi$

based on joint work with P. Schmocker and J. Teichmann ('23)

Recall that $\psi = \exp(\beta \| \cdot \|^{\gamma}_{CC,\alpha})$ for $\beta > 0$ and $\gamma > 2$.

Theorem (C.C., P. Schmocker, J. Teichmann ('23))

The linear span of the set $\left\{ \hat{\mathbb{X}}^2_{[0,T]} \mapsto \langle \epsilon_I, \hat{\mathbb{X}}_{0,T} \rangle : I \in \{0, \dots, d\}^m, m \in \mathbb{N} \right\}$ is dense in $\mathcal{B}_\psi$, i.e. for every $f \in \mathcal{B}_\psi$ and $\varepsilon > 0$ there exists a linear function L of the signature (at time T) such that

$$\sup_{\hat{\mathbb{X}}^2_{[0,T]} \in \widehat{C}^\alpha_o} \frac{\left| f(\hat{\mathbb{X}}^2_{[0,T]}) - L(\hat{\mathbb{X}}_{0,T}) \right|}{\psi(\hat{\mathbb{X}}^2_{[0,T]})} < \varepsilon.$$

- In stochastic setups this allows to obtain global approximations in probability under exponential moment conditions.

Part II

Signature methods in Stochastic Portfolio Theory

- based on joint work with Janka Möller
  C. Cuchiero and J. Möller, Signature methods in stochastic portfolio theory, 2023, https://arxiv.org/abs/2310.02322

Overview on Stochastic Portfolio Theory (SPT)

Major goals of Stochastic Portfolio Theory (SPT) are

- ... to specify only a few normative assumptions on the market (not necessarily absence of arbitrage);
- ... to analyze the relative performance of a portfolio with respect to the market portfolio, corresponding to major indices like S&P500;
- ... to develop and analyze models which allow for relative arbitrage with respect to the market portfolio;
- ... to understand various aspects of relative arbitrages, in particular properties of portfolios generating them, e.g., so-called functionally generated portfolios.

A (very incomplete) literature overview of SPT

- The first instance of the ideas of SPT is the article "Stochastic Portfolio Theory and Stock Market Equilibrium" by Robert Fernholz and Brian Shay.
- Robert Fernholz further developed it in several papers and the monograph "Stochastic Portfolio Theory" (2002).

A (very incomplete) literature overview of SPT

- The first instance of the ideas of SPT is the article "Stochastic Portfolio Theory and Stock Market Equilibrium" by Robert Fernholz and Brian Shay.
- Robert Fernholz further developed it in several papers and the monograph "Stochastic Portfolio Theory" (2002).
- Since then a lot of research has been conducted in this area, in particular by Adrian Banner, Daniel Fernholz, Robert Fernholz, Ioannis Karatzas, Constantinos Kardaras, Martin Larsson, Soumik Pal, Johannes Ruf, etc., which is partly summarized in the...
- ... overview articles and recent book
  - Stochastic Portfolio Theory: an Overview (2009) by Robert Fernholz Ioannis Karatzas;
  - Topics in Stochastic Portfolio Theory (2015) by Alexander Vervuurt;
  - Portfolio Theory and Arbitrage: A Course in Mathematical Finance (2021) by Ioannis Karatzas and Constantinos Kardaras.

Basic definitions of Stochastic Portfolio Theory (SPT)

- Consider a finite time-horizon $T > 0$ and some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F})_{t \in [0,T]}, \mathbb{P})$.
- Market capitalizations of d companies given by a vector $S = (S^1, ..., S^d)$ of d positive continuous semimartingales.
- Portfolio: a vector $\pi = (\pi^1, ..., \pi^d)$ of predictable processes such that $\sum_{i=1}^d \pi_t^i \equiv 1$ for all $t \in [0, T]$. Each π_t^i represents the proportion of current wealth invested at time t in the i^{th} asset for $i \in \{1, \ldots, d\}$
- Market Portfolio: $\mu = (\mu^1, ..., \mu^d)$ with

$$\mu_t^i = \frac{S_t^i}{S_t^1 + ... + S_t^d}, \quad t \in [0, T].$$

- Denote the simplex of dimension d by

$$\Delta^d := \{(x^1, ..., x^d) \in \mathbb{R}^d | x^1 \geq 0, ..., x^n \geq 0 \text{ and } \sum_{i=1}^d x^i = 1\}.$$

Relative wealth process

- For a portfolio π the relative wealth process with respect to the market portfolio is given by

$$Y^{\pi} := \frac{V^{\pi}}{V^{\mu}}, \quad Y_0^{\pi} = 1,$$

where V^{π} (V^{μ} resp.) denotes the wealth process generated by the portfolio π (μ resp.).

- In this multiplicative setting, the dynamics of this relative wealth process are given by

$$\frac{dY_t^{\pi}}{Y_t^{\pi}} = \sum_{i=1}^{d} \pi_t^i \frac{d\mu_t^i}{\mu_t^i}, \quad Y_0^{\pi} = 1,$$

in perfect analogy with the usual wealth process dynamics where we have μ^i instead of S^i.

Relative arbitrage and functionally generated portfolios

Definition (Relative arbitrage opportunity)

A portfolio π is said to generate a relative arbitrage opportunity with respect to the market μ over the time horizon $[0, T]$ if

$$\mathbb{P}[Y_T^\pi \geq 1] = 1 \quad \text{and} \quad \mathbb{P}[Y_T^\pi > 1] > 0.$$

Relative arbitrage and functionally generated portfolios

Definition (Relative arbitrage opportunity)

A portfolio π is said to generate a relative arbitrage opportunity with respect to the market μ over the time horizon $[0, T]$ if

$$\mathbb{P}[Y_T^\pi \geq 1] = 1 \quad \text{and} \quad \mathbb{P}[Y_T^\pi > 1] > 0.$$

Under certain conditions on the market, e.g. diversity and ellipticity or sufficient volatility, so-called functionally generated portfolios have been shown to generate such relative arbitrage opportunities.

Definition (Functionally Generated Portfolios (Fernholz '02))

Consider a C^2-function $G : U \supset \Delta^d \to \mathbb{R}_+$ such that $x_i D_i \log G(x)$ is bounded on Δ^d. Then G defines the functionally generated portfolio via

$$\pi_t^i = \mu_t^i (D_i \log G(\mu_t) + 1 - \sum_{j=1}^{d} \mu_t^j D_j \log G(\mu_t)).$$

If G is concave, it holds that $\pi_t^i \geq 0$ for all $i \in \{1, ..., d\}$ and $t \in [0, T]$.

Fernholz's master equation

Proposition (Pathwise version of Fernholz's master equation)

Let π be a functionally generated by G and $(\mu_t)_{t\in[0,T]}$ a continuous path admitting a continuous S_+^d-valued quadratic variation $[\mu]$ along a refining sequence of partitions (in the sense of Föllmer).

Then the relative wealth process $(Y_t^\pi)_{t\geq 0}$ satisfies

$$\log(Y_t^\pi) = \log(G(\mu(t))) - \log(G(\mu(0))) + \mathfrak{g}_t, \quad t \in [0,T],$$

where $\mathfrak{g}_t = \int_0^t -\frac{1}{2G(\mu(t))} \sum_{i,j} D^{ij} G(\mu(t)) d[\mu^i, \mu^j]_t$.

Remark: Under certain market conditions it can be shown that after a sufficiently long time horizon t^*, the term $\mathfrak{g}_{t^*}$ dominates $\log(G(\mu(t))) - \log(G(\mu(0)))$ and thus creates relative arbitrage.

Signature portfolios

- Inspired by functionally generated portfolios and control problems in finance solved via signature methods (e.g. Kalsi et al. ('19) or Bayer et al. ('21)), we introduce path functional portfolios and signature portfolios.
- We denote here and throughout the signature of X by $\mathbb{X}_t := \mathbb{X}_{0,t}$.

Definition (Path-functional portfolios)

Consider a continuous semimartingale $(X_t)_{t\in[0,T]}$ and let $\hat{X}_t = (t, X_t)$. We define two types of path-functional portfolios, denoted by η and θ,

$$\eta_t^i = \mu_t^i(F^i(\hat{X}_{[0,t]}) + 1 - \sum_{j=1}^{d} \mu_t^j F^j(\hat{X}_{[0,t]})), \qquad (\eta\text{-portfolio})$$

$$\theta_t^i = F^i(\hat{X}_{[0,t]}) + \mu_t^i(1 - \sum_{j=1}^{d} F^j(\hat{X}_{[0,t]})). \qquad (\theta\text{-portfolio})$$

If $F^i(\hat{X}_{[0,t]}) = \sum_{0\leq|I|\leq n} \alpha_I^{(i)} \langle \epsilon_I, \hat{\mathbb{X}}_t \rangle$, then the path functional portfolio is called signature portfolio.

Optimizing performance functionals - logarithmic utility

- The goal is now to optimize certain performance functionals within the class of signature portfolios.
- We start with logarithmic utility for the relative wealth process, i.e. the goal is to optimize $\mathbb{E}[\log Y_t^\eta]$, by finding optimal parameters $\{\alpha_I^i\}_{0\leq I\leq n, i\in\{1,\ldots,d\}}$. A similar method also works for the θ-portfolio.
- Note that it is the same to optimize the (absolute) log portfolio wealth or the relative log portfolio wealth (w.r.t the market) as

$$\left(\max_{\{\alpha_I^i\}_{0\leq I\leq n, i\in\{1,\ldots,d\}}} \mathbb{E}[\log V_t^\eta]\right) \Leftrightarrow \left(\max_{\{\alpha_I^i\}_{0\leq I\leq n, i\in\{1,\ldots,d\}}} \mathbb{E}[\log V_t^\eta] - \mathbb{E}[\log V_t^\mu]\right)$$

$$\Leftrightarrow \left(\max_{\{\alpha_I^i\}_{0\leq I\leq n, i\in\{1,\ldots,d\}}} \mathbb{E}\left[\log Y_t^\eta\right]\right).$$

Optimizing logarithmic utility within signature portfolios

Theorem (C.C., Janka Möller ('23))

Consider a market of d stocks, let X and μ be a $\mathbb{R}^n$-valued and Δ^d-valued continuous semimartingales. Let $t_0 \geq 0$ be the time at which we start to invest. Consider an arbitrary but fixed labelling function $\mathcal{L}$. Then

$$\max_{\{\alpha_I^{(i)}\}_{i \in \{1,\dots,d\}, 0 \leq |I| \leq n}} \mathbb{E}\left[\log\left(Y_t^{\eta}\right)\right] \Leftrightarrow \min_{\mathbf{x}} \frac{1}{2} \mathbf{x}^T \mathbb{E}[\mathbf{Q}(t)]\mathbf{x} - \mathbb{E}[\mathbf{c}(t)]^T \mathbf{x}$$

where $\mathbf{x}$, $\mathbf{c}(t)$ *are vectors and* $Q(t)$ *is a matrix with coefficients*

$$\mathbf{x}_{\mathcal{L}(I,i)} = \alpha_I^{(i)}$$

$$(\mathbf{c}(t))_{\mathcal{L}(I,i)} = \int_{t_0}^{t} \langle \epsilon_I, \hat{\mathbb{X}}_s \rangle d\mu_s^i, \; (\mathbf{Q}(t))_{\mathcal{L}(I,i),\mathcal{L}(J,j)} = \int_{t_0}^{t} \langle \epsilon_I \sqcup\!\sqcup\, \epsilon_J, \hat{\mathbb{X}}_s \rangle d[\mu^i, \mu^j]_s.$$

The optimization task is a convex quadratic optimization problem.

Sketch of the proof and remarks

- By the form of the η-portfolio the log relative wealth process is given by

$$
\begin{aligned}
\log(Y_t^\eta) &= \sum_{i=1}^{d} \int_{t_0}^{t} \frac{\eta_s^i}{\mu_s^i} d\mu_s^i - \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} \int_{t_0}^{t} \frac{\eta_s^i}{\mu_s^i} \frac{\eta_s^j}{\mu_s^j} d[\mu^i, \mu^j]_s \\
&= \sum_{i=1}^{d} \int_{t_0}^{t} F^i(\hat{X}_{[0,s]}) d\mu_s^i - \frac{1}{2} \sum_{i=1}^{d} \sum_{j=1}^{d} \int_{t_0}^{t} F^i(\hat{X}_{[0,s]}) F^j(\hat{X}_{[0,s]}) d[\mu^i, \mu^j]_s.
\end{aligned}
$$

  The linearity of F and the shuffle property of the signature yields the above convex quadratic optimization problem.

Sketch of the proof and remarks

- By the form of the η-portfolio the log relative wealth process is given by

$$\log(Y_t^\eta) = \sum_{i=1}^d \int_{t_0}^t \frac{\eta_s^i}{\mu_s^i} d\mu_s^i - \frac{1}{2}\sum_{i=1}^d \sum_{j=1}^d \int_{t_0}^t \frac{\eta_s^i}{\mu_s^i}\frac{\eta_s^j}{\mu_s^j} d[\mu^i,\mu^j]_s$$

$$= \sum_{i=1}^d \int_{t_0}^t F^i(\hat{X}_{[0,s]}) d\mu_s^i - \frac{1}{2}\sum_{i=1}^d \sum_{j=1}^d \int_{t_0}^t F^i(\hat{X}_{[0,s]}) F^j(\hat{X}_{[0,s]}) d[\mu^i,\mu^j]_s.$$

  The linearity of F and the shuffle property of the signature yields the above convex quadratic optimization problem.

- If $X = \mu$, then the components of $\mathbf{c}(t)$ and $\mathbf{Q}(t)$ are linear functions of the signature of $t \mapsto \widehat{\mu}_t = (t, \mu_t)$, whose expected value can then often easily be computed.

- Note that in practice the optimization is performed along the observed trajectory, i.e. without expected values. This allows to detect (path-)functionally generated relative arbitrages if they exist.

Remarks

- Suppose that μ has dt characteristics with drift b_t and diffusion matrix C_t. The general log-optimal portfolio is found by solving the quadratic optimization task

$$\inf_{\pi} \mathbb{E}[\int_{t_0}^{t} \frac{1}{2}(\frac{\pi_t}{\mu_t})^\top C_t(\frac{\pi_t}{\mu_t}) - b_t^\top \frac{\pi_t}{\mu_t})dt],$$

where the inf is taken over predictable processes with $\sum \pi_t^i = 1$.

Remarks

- Suppose that μ has dt characteristics with drift b_t and diffusion matrix C_t. The general log-optimal portfolio is found by solving the quadratic optimization task

$$\inf_{\pi} \mathbb{E}[\int_{t_0}^{t} \frac{1}{2}(\frac{\pi_t}{\mu_t})^\top C_t(\frac{\pi_t}{\mu_t}) - b_t^\top \frac{\pi_t}{\mu_t})dt],$$

  where the inf is taken over predictable processes with $\sum \pi_t^i = 1$.

- This optimization problem on the level of π is translated to a quadratic optimization problem over signature coefficients without constraints.

Remarks

- Suppose that μ has dt characteristics with drift b_t and diffusion matrix C_t. The general log-optimal portfolio is found by solving the quadratic optimization task

$$\inf_{\pi} \mathbb{E}[\int_{t_0}^{t} \frac{1}{2}(\frac{\pi_t}{\mu_t})^\top C_t(\frac{\pi_t}{\mu_t}) - b_t^\top \frac{\pi_t}{\mu_t})dt],$$

  where the inf is taken over predictable processes with $\sum \pi_t^i = 1$.

- This optimization problem on the level of π is translated to a quadratic optimization problem over signature coefficients without constraints.
- A similar convex quadratic optimization problem (with $\mathbf{Q}(t)$ of slightly different form) is obtained by replacing F^i by any linear function of some features, corresponding e.g. to
  - randomized signature (C.C., Gonon, Grigoryeva, Ortega, Teichmann);
  - random neural networks (Herrera, Krach, Teichmann).

General structure

Corollary (Quadratic Optimization Tasks)

Consider an optimization problem of the form

$$\inf_{\beta} \mathbb{E}[\int_{t_0}^{t} \beta_s^\top C_s \beta_s \nu_1(ds) - \int_{t_0}^{t} b_s^\top \beta_s \nu_2(ds)] \qquad (*)$$

over predictable processes β with values in $\mathbb{R}^d$, where b and C are stochastic processes with values in $\mathbb{R}^d$ and $\mathbb{S}^d$ resp., ν_i denotes signed measures on $[t_0, t]$.

If the controls β are parametrized via $\beta_t^i = \sum_{p \in \mathcal{P}} \alpha_p^i \varphi^p(t, X_{[0,t]})$, where $\{\varphi^p\}_{p \in \mathcal{P}}$ is a collection of feature maps and $\alpha_p^i \in \mathbb{R}$ are constant optimization parameters, then (*) is a quadratic optimization problem in $\{\alpha_p^i\}_{1 \le i \le d, p \in \mathcal{P}}$.

- A choice for φ^p is a version of randomized signature, $\varphi^p = \langle A^p, \widehat{\mathbb{X}}_t^N \rangle$, where A^p denotes the p-th row of a Johnson-Lindenstrauss projection matrix.
- Beside the log-optimal portfolio, a mean-variance type portfolio optimization can be cast into this framework.

Approximation by signature portfolios

Define the space of lifted stopped paths $\Lambda_T^2 = \bigcup_{t \in [0,T]} \{(\hat{\mathbb{X}}^2_{[0,t]})(\omega) \mid X \text{ cont. semi-mart.}, \hat{X}_s = (s, X_s),\ s \in [0,t]\}$ and equip it with an appropriate α-Hölder norm for $\alpha \in (1/3, 1/2)$.

Approximation by signature portfolios

Define the space of lifted stopped paths
$\Lambda_T^2 = \bigcup_{t\in[0,T]}\{(\hat{\mathbb{X}}^2_{[0,t]})(\omega) \,|\, X \text{ cont. semi-mart.}, \hat{X}_s = (s, X_s),\ s \in [0,t]\}$ and equip it with an appropriate α-Hölder norm for $\alpha \in (1/3, 1/2)$.

Proposition (C.C., Janka Möller ('23))

Consider for $t \in [0,T]$ path-functional portfolios of η- and θ-type of the form

$$\pi_t^i = \mu_t^i\Big(f^i(\hat{\mathbb{X}}^2_{[0,t]}) + 1 - \sum_j \mu_t^j f^j(\hat{\mathbb{X}}^2_{[0,t]})\Big) \quad \text{and} \quad \pi_t^i = f^i(\hat{\mathbb{X}}^2_{[0,t]}) + \mu_t^i\Big(1 - \sum_j f^j(\hat{\mathbb{X}}^2_{[0,t]})\Big),$$

where f^i are continuous non-anticipating path functionals on Λ_T^2 for every i.

- *Then portfolios of η- and θ-type can be approximated arbitrarily well by signature portfolios η^{Sig} (θ^{Sig} resp) uniformly in time and on compacts of Λ_T^2.*
- *Moreover, if $\mathbb{E}[\exp(\beta\|\hat{\mathbb{X}}_{[0,T]}\|^\gamma_{CC,\alpha})] < \infty$ for $\beta > 0$ and $\gamma > 2$, then for any $\varepsilon, \delta > 0$, there exists a signature portfolio η^{Sig} (θ^{Sig} resp) such that*

$$\mathbb{P}\Big[\sup_{t\in[0,T]} \|\pi_t - \eta_t^{Sig}\| > \varepsilon\Big] < \delta.$$

Approximation of the log-optimal portfolio

Proposition (C. C., Janka Möller ('23))

Consider a market model, where for all $i \in \{1, \dots, d\}$

$$dS_t^i = S_t^i \left(a^i \left(\hat{\mathbb{X}}^2_{[0,t]} \right) dt + \sum_{j=1}^{m} B^{ij} \left(\hat{\mathbb{X}}^2_{[0,t]} \right) dW_t^j \right),$$

with $m \geq d$ such that $(BB^T)^{-1}$ exists (and some integrability cond. are satisfied). Assume that for all $i \in \{1, \dots, d\}$, $j \in \{1, \dots, m\}$ a^i, B^{ij} are continuous non-anticipating path-functionals on Λ^2_T.

- *Then the log-optimal portfolio can be approximated arbitrarily well by signature portfolios θ^{Sig} uniformly in time and on compact sets of Λ^2_T.*
- *Moreover, if $\mathbb{E}[\exp(\beta \|\hat{\mathbb{X}}_{[0,T]}\|^{\gamma}_{CC,\alpha})] < \infty$ for $\beta > 0$ and $\gamma > 2$, then for any $\varepsilon, \delta > 0$, there exists a signature portfolio θ^{Sig} such that*

$$\mathbb{P}[\sup_{t \in [0,T]} \|\pi_t - \theta_t^{Sig}\| > \varepsilon] < \delta.$$

Learning the log-optimal portfolio

1. Correlated Black-Scholes Market:

$$dS_t^i = S_t^i(a^i dt + \sum_{j=1}^{m} B^{ij} dW_t^j), \quad 1 \leq i \leq d.$$

Learning the log-optimal portfolio

1. Correlated Black-Scholes Market:

$$dS_t^i = S_t^i(a^i dt + \sum_{j=1}^m B^{ij} dW_t^j), \quad 1 \leq i \leq d.$$

2. Volatility Stabilized Market:

$$\frac{dS_t^i}{S_t^i} = \frac{1+\gamma}{2} \frac{1}{\mu_t^i} dt + \sqrt{\frac{1}{\mu_t^i}} dW_t^i \quad 1 \leq i \leq d.$$

Learning the log-optimal portfolio

1. Correlated Black-Scholes Market:

$$dS_t^i = S_t^i(a^i dt + \sum_{j=1}^m B^{ij} dW_t^j), \quad 1 \leq i \leq d.$$

2. Volatility Stabilized Market:

$$\frac{dS_t^i}{S_t^i} = \frac{1+\gamma}{2}\frac{1}{\mu_t^i}dt + \sqrt{\frac{1}{\mu_t^i}}dW_t^i \quad 1 \leq i \leq d.$$

3. Signature Market:

$$dS_t^i = S_t^i(\mathbf{a}_t^i dt + \sum_{j=1}^m B^{ij}\, dW_t^j) \quad 1 \leq i \leq d$$

where $(\mathbf{a}_t^i) = \sum_{0 \leq |I| \leq N} \lambda_I^{(i)} \langle \epsilon_I, \hat{\mu} \rangle_t$ and $B \in \mathbb{R}^{d \times m}$.

Optimization procedure

For each market:

- We use a Monte-Carlo type optimization. Note

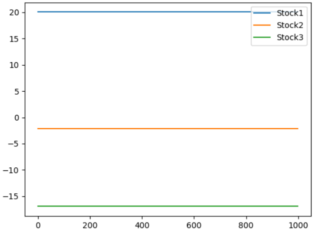
$$\left(\max_{\eta} \ \frac{1}{M}\sum_{m=1}^{M} \log Y_T^{\eta}(\omega_m)\right) \Leftrightarrow \left(\min_{\mathbf{x}} \ -\mathbf{x}^T\tilde{\mathbf{c}}(T) + \frac{1}{2}\mathbf{x}^T\tilde{\mathbf{Q}}(T)\mathbf{x}\right),$$

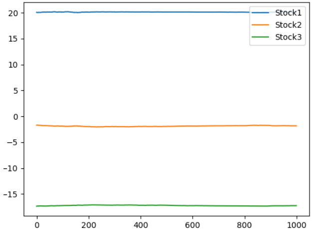
  for $\omega_1, \dots, \omega_M \in \Omega$ and where $\tilde{\mathbf{Q}}(T) = \frac{1}{M}\sum_{m=1}^{M} \mathbf{Q}(T, \omega_m)$ and $\tilde{\mathbf{c}}(T) = \frac{1}{M}\sum_{m=1}^{M} \mathbf{c}(T, \omega_m)$.

- We take here $d = 3$.
- Simulate $M \approx 100000$ in-sample trajectories to create $\tilde{\mathbf{Q}}(T)$, $\tilde{\mathbf{c}}(T)$.
- Evaluate performance on 100000 test samples and compare it to the respective theoretical log-optimal portfolio.
- Log-optimal weights are never shown to signature portfolios during training!

Results: Black-Scholes Market

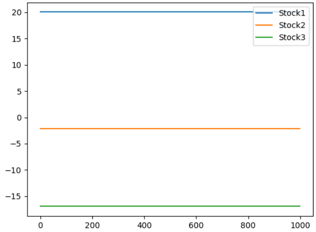
- We learned a signature portfolio of type η of degree three.
- Mean log-relative wealth equals 9.0115 in the theoretical log-optimal portfolio (left), while in the learned signature portfolio (right) it is 9.0122.

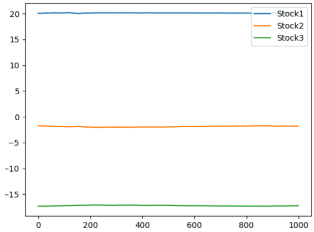




Results: Black-Scholes Market

- We learned a signature portfolio of type η of degree three.
- Mean log-relative wealth equals 9.0115 in the theoretical log-optimal portfolio (left), while in the learned signature portfolio (right) it is 9.0122.



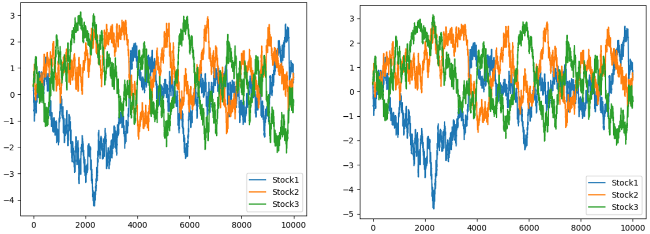


- The log-optimal portfolio in the B&S model, is a signature portfolio of type θ, but as we approximate it with an η-portfolio, the approximation task is actually

$$F^{(BS),i}(\mu_{[0,t]}) \approx \frac{c_i}{\mu_t^i}.$$

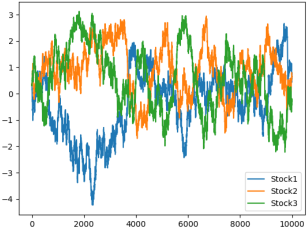
Results: Volatility Stabilized Market

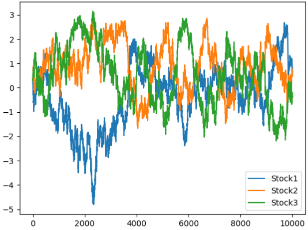
- We learned a signature portfolio of type η of degree three.
- Mean log-relative wealth equals 8.7619 in the theoretical log-optimal portfolio (left), while in the learned signature portfolio (right) it is 8.7417.



Results: Volatility Stabilized Market

- We learned a signature portfolio of type η of degree three.
- Mean log-relative wealth equals 8.7619 in the theoretical log-optimal portfolio (left), while in the learned signature portfolio (right) it is 8.7417.



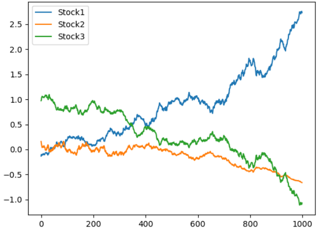


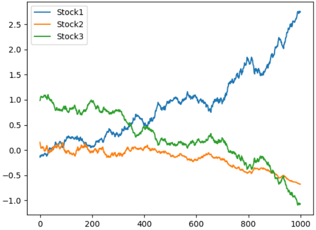
- The approximation task is here

$$F^{(Vol),i}(\mu_{[0,t]}) \approx \frac{\alpha+1}{2\mu_t^i} + \frac{d}{2}(\alpha-1).$$

Results: Signature Market

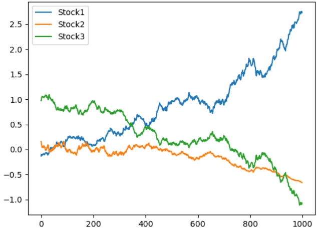
- We learned a signature portfolio of type θ of degree two.
- Mean log-relative wealth equals 0.2357 in the theoretical log-optimal portfolio (left), while in the learned signature portfolio (right) it is 0.2355.

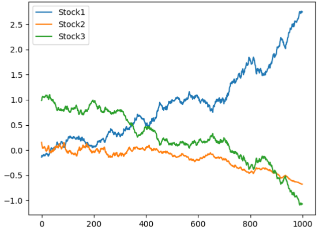




Results: Signature Market

- We learned a signature portfolio of type θ of degree two.
- Mean log-relative wealth equals 0.2357 in the theoretical log-optimal portfolio (left), while in the learned signature portfolio (right) it is 0.2355.





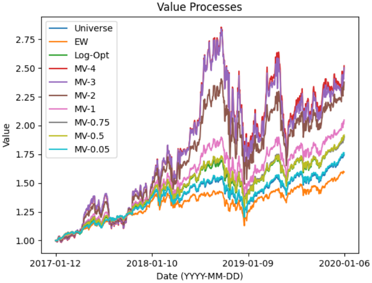
- Here, the log-optimal portfolio is a signature portfolio of type θ.

NASDAQ market

- We here consider the 100 dimensional NASDAQ market.
- Note that when working with real market data, we only have one realization available. Hence, we optimize just along the past observed trajectory (in other words we replace expectations by time averages).
- We choose X to be the ranked market weights.
- We apply a Johnson-Lindenstrauss projection of dimension 50 to the signature computed up to order 3 and then replace F^i in the η-portfolio by a linear map of this randomized signature.
- We perform both the log-utility and the mean-variance optimization with different risk aversion parameters.
- We take as an in-sample period 2000 trading days and as an out-of-sample period the following 750 trading days. The training is performed on historical data without estimating any drift or volatiltiy.

Results NASDAQ Market

We present out-of-sample results here without transaction costs.



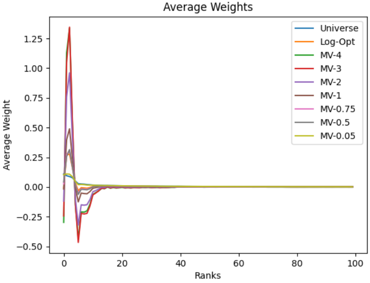


Figure: Left: Out-of-sample wealth processes entire NASDAQ, equally weighted portfolio, randomized signature portfolios optimizing log-utility and mean-variance.
Right: Average weights

Results S&P500 market

- We apply a similar procedure to the S&P 500, this time by choosing X to be the name-based market weights and by adding transaction costs.
- To keep the convex quadratic optimization structure we add the penalization term $\frac{\beta}{T}\sum_{t=0}^{T-1}\sum_i \left(\frac{\pi^i_{t+1}}{\mu^i_{t+1}} - \frac{\pi^i_t}{\mu^i_t}\right)^2$ accounting for transaction costs.

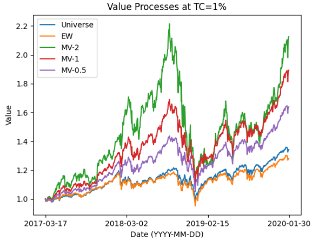


Figure: Out of sample wealth process with 1% prop. trans. costs, S&P500, equally weighted and randomized signature portfolio optimizing mean-variance.

- This picture suggests that a (strong) relative arbitrage opportunity even under transaction costs has been detected at least in this testing period.

Conclusion

- Signature portfolios can approximate a large class of path-functional portfolios including
  - classical functionally generated portfolios
  - log-optimal portfolios in a large class of non-Markovian markets.

  In some markets the log-optimal portfolios are exactly signature portfolios.
- Despite their versatility, optimizing the log-utiltiy or mean variance within the class of (randomized) signature portfolio leads to a convex quadratic optimization problem.
- Inclusion of transaction costs is possible, while preserving tractability of the optimization problem.
- The application to real market data points towards out-performance during the out-of-sample testing period we considered, also under transaction costs.

Bibliography and related literature

- C. Cuchiero and J. Möller: Signature methods in stochastic portfolio theory; https://arxiv.org/abs/2310.02322

Bibliography and related literature

- C. Cuchiero and J. Möller: Signature methods in stochastic portfolio theory; https://arxiv.org/abs/2310.02322

Related literature:

- O. Futter, B. Horvath, and M. Wiese: Signature Trading: A Path-Dependent Extension of the Mean-Variance Framework with Exogenous Signals; https://arxiv.org/abs/2308.15135
  - This paper treats mean-variance optimization with an additive approach where the trading strategies correspond to numbers of shares (inclusion of bank account is necessary to guarantee selfinancing).
- S. Campbell and L. Wong; Functional portfolio optimization in stochastic portfolio theory; https://arxiv.org/abs/2103.10925
  - This paper treats functional portfolio optimization over a family of ranked based portfolios.