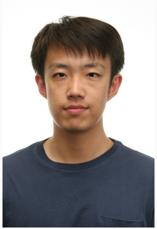
Factor Learning Portfolio Optimization

Presented by: Sinong Geng
Princeton University

October 18, 2022

Authors



Sinong Geng
Princeton University

Houssam Nassif
Amazon

Zhaobin Kuang
Google

A. Max Reppen
Boston University

Ronnie Sircar
Princeton University

Supported by Amazon AWS Credits.

Contents

Motivation

Motivation

→ Portfolio Optimization: Learn a policy for wealth allocation in order to

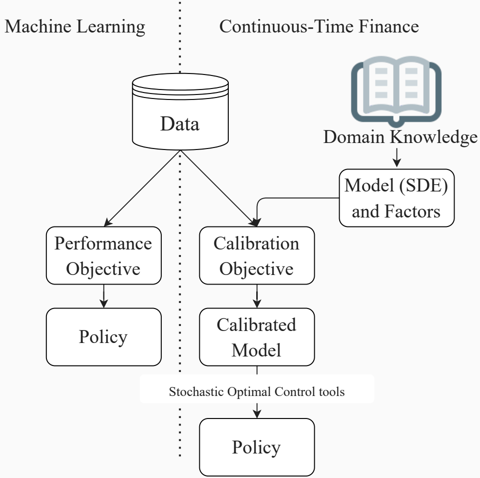
- *maximize return*,
- *minimize risk*.

→ Stochastic Factors like economic indexes and proprietary trading signals:

- Not controllable.
- Evolve over time stocahstically.
- Affect asset prices.

Motivation

→ Machine Learning v.s. Continuous-Time Finance



Machine Learning:

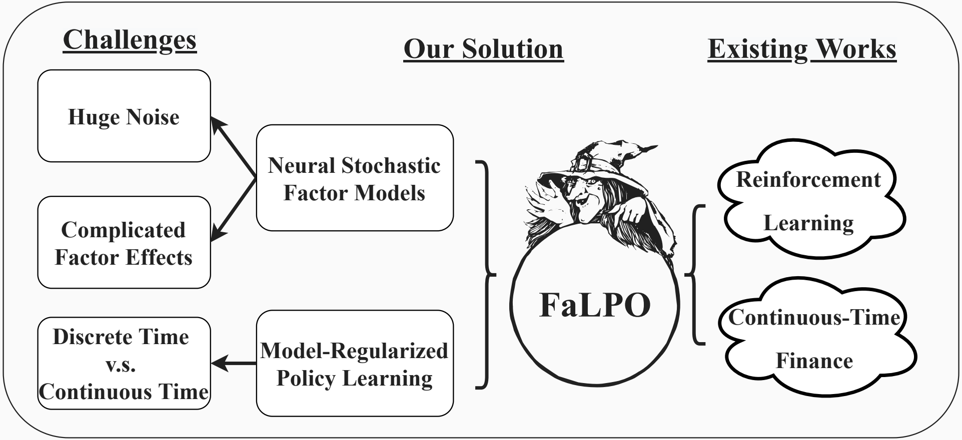
- ✔ Flexible Representation
- ✖ Poor sample complexity and tend to overfit

Continuous-Time Finance:

- ✔ Small sample complexity.
- ✖ Rely on domain knowledge and thus may end up with over-simplified models.

Motivation

→ Combine machine learning with continuous-time finance



Methodology

Neural Stochastic Factor Model
Model-Regularized Policy Learning

Methodology

→ Problem Formulation

- Assets: $S_t := [S_t^1, S_t^2, \cdots S_t^{d_S}]^\top$ and a risk-free money market account with, for simplicity, zero interest rate of return;
- Features: Y_t
- Factors: From Y_t, we can derive d_X factors denoted as X_t which
  - affect the dynamics of asset prices;
  - evolve over time stochastically;
  - are not affected by investment decisions.
- Policy: π_t as the fractions of wealth invested in the d_S assets at time point t.
- Wealth: Z_t^π.
- Performance Objective/Value Function:

$$\max_\pi V(\pi) \text{ with } V(\pi) := \mathbb{E}[U(Z_T^\pi)].$$

Neural Stochastic Factor Models

→ Stochastic Factor Models

$$\frac{dS_t^i}{S_t^i} = f_S^i(X_t;\theta_S^*)dt + \sum_{j=1}^{d_W} g_S^{ij}(X_t;\theta_S^*)dW_t^j, \quad i \in \{1, 2, \cdots, d_S\},$$
$$dX_t = f_X(X_t;\theta_S^*)dt + g_X(X_t;\theta_S^*)^\top dW_t.$$

→ Representation Function

$$X_t = \phi(Y_t;\theta_\phi^*)$$

Model-Regularized Policy Learning

→ Policy Functional Form

- Using tools in stochastic optimal control, we can derive the functional form of an optimal continuous-time policy: $\tilde{\pi}_t^* = \Pi(t, S_t, Z_t, X_t; \theta_{\tilde{\pi}}^*)$.
- Use the funtional form in policy parameterization.

$$\pi(t, S_t, Z_t, Y_t; \theta_\phi, \theta_\pi) := \Pi(t, S_t, Z_t, \phi(Y_t; \theta_\phi); \theta_\pi).$$

→ Model Calibration

$$\max_{(\theta_\phi, \theta_\pi, \theta_S) \in \mathcal{A}} H(\theta_\phi, \theta_\pi, \theta_S),$$

$$H(\theta_\phi, \theta_\pi, \theta_S) := (1 - \lambda) V(\theta_\phi, \theta_\pi) + \lambda L(\theta_\phi, \theta_S).$$

Model-Regularized Policy Learning

→ Algorithm

Algorithm FaLPO

1: **Input:** number of iterations N.
2: Initialize θ_ϕ and θ_π .
3: **for** $n \in [N]$ **do**
4: Parameterize the policy function with Π.
5: Estimate the policy gradient for H.
6: Update θ_ϕ, θ_π, and θ_S.
7: **end for**
8: **Return** $\pi(\cdot; \boldsymbol{\theta}_\phi, \boldsymbol{\theta}_\pi)$

Example: Kim-Omberg Model

→ Neural Stochastic Factor Model

$$\frac{dS_t^i}{S_t^i} = X_t^i\,dt + \sum_{j=1}^{d_W} \sigma^{ij}\,dW_t^j,$$

$$dX_t = \mu(\omega - X_t)\,dt + v\,dW_t, \text{ and } X_t = \phi(Y_t;\theta_\phi^*).$$

→ Model-Regularized Policy Learning

- **Policy Functional Form:** For power utility $\Pi(t, S_t, Z_t, \phi(Y_t;\theta_\phi);\theta_\pi) = k_1(t;\theta_\pi)\phi(Y_t;\theta_\phi) + k_2(t;\theta_\pi)$; for exponential utility $\Pi(t, S_t, Z_t, \phi(Y_t;\theta_\phi);\theta_\pi) = k_1(t;\theta_\pi)\phi(Y_t;\theta_\phi)/Z_t + k_2(t;\theta_\pi)/Z_t$.
- **Model Calibration:**
  $L(\theta_\phi, \theta_S) := -\mathbb{E}\left[\sum_{i=1}^{d_S} \left[\log(S_{t+\Delta t}^i) - \log(S_t^i) - \phi^i(Y_t;\theta_\phi)\Delta t - \theta_S^i\right]^2\right]$

Results

Theory

→ Setup

Algorithm Projected FaLPO

1: **Input:** Number of iterations N and a ball $\mathcal{B}$.
2: **Output:** θ_ϕ, θ_π, and θ_S
3: **for** $n \in [N]$ **do**
4: Parameterize the policy function by Π.
5: Estimate the gradients of H.
6: Update θ_S and θ_R with learning rate η by gradients.
7: Project the achieved update to $\mathcal{B}$.
8: **end for**
9: **Return** θ_ϕ, θ_π, and θ_S.

→ In $\mathcal{B}$, we pose assumptions.

Theory

→ Main Results

Theorem

With the aforementioned projection-based FaLPO algorithm and assumptions, there exist positive constants C_1, C_2, C_3, and C_4 such that

$$\mathbb{E}[V^*_{\Delta t} - V(\bar{\theta})] \leq \frac{e_{\Delta t}}{1-\lambda} + \frac{H(\theta^*_{\Delta t}) - H(\theta^\dagger)}{1-\lambda} + \frac{C_1 \log(N)}{N(1-\lambda)} + \frac{C_1 \log(N)}{BN(1-\lambda)}\left[(1-\lambda)^2 C_2 + \lambda^2 C_3 + 2\lambda(1-\lambda)C_4\right], \tag{1}$$

where $\lambda \in [0,1]$. Also, $e_{\Delta t}$ is an error term not related to N or B but dependent on Δt with $\lim_{\Delta t \to 0} e_{\Delta t} = 0$.

Experiments

→ Synthetic:

Annual Volatility	0.1	0.2	0.3
FaLPO	$\mathbf{-0.465} \pm 0.446$	$\mathbf{-1.35} \pm 0.155$	$\mathbf{-2.737} \pm 0.219$
DDPG	-1.650 ± 0.456	-3.30 ± 1.294	-5.495 ± 1.269
SLAC	-0.750 ± 0.210	-5.50 ± 0.011	-6.160 ± 0.012
RichID	-3.350 ± 0.111	-5.65 ± 0.102	-6.325 ± 0.048
CT-MB-RL	-2.850 ± 0.014	-5.35 ± 0.020	-6.160 ± 0.026
MMMC	-4.723 ± 7.619	-5.602 ± 4.299	-6.124 ± 3.217

Table: Average terminal utility after tuning with standard deviation for synthetic data

Experiments

→ Real-world portfolio optimization:

Methods	Energy	Material	Industrials	Mix
FaLPO	$\mathbf{-2.4} \pm 1.9$	$\mathbf{-3.2} \pm 1.0$	$\mathbf{-6.3} \pm 2.3$	$\mathbf{-3.5} \pm 1.5$
DDPG	-6.6 ± 1.2	-7.3 ± 1.5	-7.3 ± 2.1	$-2.5 \times 10^4 \pm 3.3 \times 10^8$
SLAC	-6.8 ± 0.2	-7.0 ± 1.5	-342.4 ± 886.8	$-3.0 \times 10^8 \pm 4.3 \times 10^{12}$
RichID	-6.5 ± 0.1	-6.9 ± 1.4	-6.9 ± 0.4	-8.1 ± 3.9
CT-MB-RL	-4.2 ± 6.2	-5.4 ± 4.3	-11655 ± 32947.5	-5.7 ± 3.1
MMMC	-8.5 ± 7.6	-6.5 ± 1.7	-11.0 ± 5.4	-7.5 ± 4.4

Table: Average terminal utility for real-world data. Mix denotes a mix of stocks in the previous three sectors.

Thank you!

Appendix

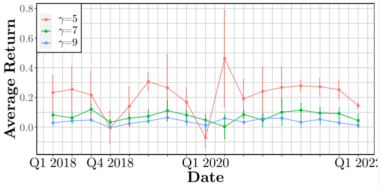
→ Competing Methods:

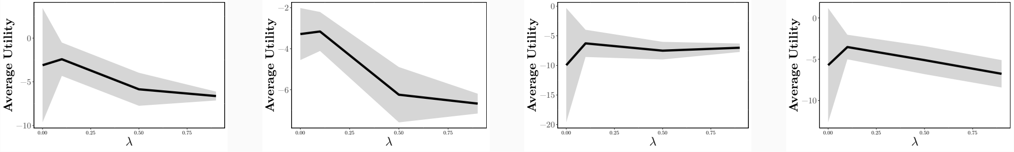
Methods	Explicit Factor Representation	Continuous-Time Model	Discrete-Time Model
MMMC	✖	✔	✖
DDPG	✖	✖	✖
SLAC	✔	✖	✔
RichID	✔	✖	✔
CT-MB-RL	✖	✔	✖
FaLPO	✔	✔	✖

Table: Competing methods and their characteristics.

Appendix

→ More results:





(a) Oil (b) Material (c) Industrials (d) Mix

Figure: Sensitivity analysis for λ