Using Machine Learning Technique to Enhance the Portfolio Construction Based on PolyModel Theory

Siqiao Zhao, Dan Wang, Raphael Douady

October 15, 2023

Abstract

Factor models are widely employed in portfolio construction, optimization and empirical asset pricing, for the purpose of assessing the anticipated returns and risks pertaining to diverse assets or portfolios. In the field of factor models, two primary classifications of models stand out: the single-factor model and the multi-factor model. However, single-factor models usually underfit the complicated relations between assets and factors while multi-factor models suffer issues like collinearity, curse of dimensions, etc. To address these shortcomings, we propose a new method, the PolyModel theory.

When constructing portfolios, a key problem is that a lot of financial time series data are sparse, making it challenging to apply conventional machine learning techniques. Such scenarios are more common in Hedge funds, especially concerning the performance of hedge funds, where returns and AUMs are only reported on a monthly basis, and historical data is limited. Polymodel theory can solve those problems and demonstrate superiority in portfolio construction from following aspects: take the predicted future movements of different risk factors into the picture; utilizing long-term, more frequent time series of risk-factors in the estimation; getting a good estimation of the joint risks over all the risk factors to the portfolio.

To implement the PolyModel theory for constructing a hedge fund portfolio, we begin by identifying an asset pool, utilizing over 10,000 hedge funds for past 29 years' data. In addition to asset selection, PolyModel theory involves choosing a diverse set of risk factors, which includes various financial indices, currencies, and commodity prices. This comprehensive selection mirrors the complexities of the real-world environment. Leveraging the PolyModel theory, we create features like Long-term Alpha (LTA), Long-term Ratio (LTR), and SVaR (StressVaR), and other quantitative measures like the Sharpe ratio.

To enhance the performance for portfolio construction, we also employ machine learning techniques (for instance, XGBoost has the best performance) to capture the upward trend using all the features. This step largely improve our strategies, in particular, our strategies achieve better performance on Sharpe ratio and annualized return. The above process enables us to create multiple portfolio strategies aiming for high returns and low risks when compared to the benchmarks (HFRIFOF and HFRIFWI).

In the poster, we will

- Explain the principal of PolyModel Theory and provide mathematical justification for why PolyModel Theory can address the issue for traditional factor-model;
- Display the dataset used in portfolio construction and discuss about the challenges and limitations for financial data;
- Feature construction and provide the explanations of the significance and relevance of the features;
- Demonstrate how machine learning techniques are applied to the portfolio construction and show the final portfolio performance comparing with original methods and benchmarks ;
- Discuss some further directions and on-going work.

This is a joint work with Dan Wang and Prof. Raphael Douady.