SCHOOL OF COMPUTATION, INFORMATION AND TECHNOLOGY — INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Master's Thesis ... in Informatics

Portfolio Optimization with Gaussian Process Regression

Xiyue ZHANG

SCHOOL OF COMPUTATION, INFORMATION AND TECHNOLOGY — INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Master's Thesis ... in Informatics

Portfolio Optimization with Gaussian Process Regression

Titel der Abschlussarbeit

Author: Xiyue ZHANG
Supervisor: Prof. Dr. Hans-Jochim Bungartz
Advisor: Kislaya Ravi
Submission Date: December 3rd

I confirm that this master's thesis ... is my own work and I have documented all sources and material used.

Munich, December 3rd Xiyue ZHANG

Acknowledgments

Abstract

Contents

1 Introduction

1.1 Background and Motivation

Citation test [Lam94].

Acronyms must be added in `main.tex` and are referenced using macros. The first occurrence is automatically replaced with the long version of the acronym, while all subsequent usages use the abbreviation.

E.g. `\ac{TUM}, \ac{TUM}` ⇒ Technical University of Munich (TUM), TUM

For more details, see the documentation of the `acronym` package[1].

1.1.1 the US stock market

Sharpe Ratio. The idea combines not only return and volatility but also the fixed-income market. The fixed-income market is important because it shows what any risk-adverse investor would earn by simply being in the most conservative investment: the risk-free rate. The risk-free rate is the rate earned simply for investing in a risk-free bond. Since stocks are risky, they should earn returns in excess of that return. If a stock were to earn less than the risk-free return, then a rational investor would forgo the volatility of the stock market and stick with risk-free investments, which protect the principal while still earning some return.

The numerator in our example is the market premium. So, suppose for stock C, the market premium and standard deviation are 7% and 10%, respectively. Suppose for stock D, the market premium and standard deviation are 8% and 12%, respectively. Which is the preferred investment? In this case, stock C has a higher return per unit of volatility (risk) than stock D does. We would prefer the investment that has a higher amount of return per unit of risk. Now, this is the opposite of our coefficient of variation. When we express volatility units per return, we would like the lower number. When we express return units per volatility, we prefer the larger number. Here, the larger number means we expect to earn more return per unit of volatility.

In this case, we will name this ratio (keep a *sharp* lookout for it) and use it extensively to compare investments, but for now, the important thing is that we have seen how return and volatility can be combined to compare investments. This can just

[1] `https://ctan.org/pkg/acronym`

as easily work if these securities are from different asset classes. Return and risk help encapsulate the key statistical properties of a financial asset's performance. In the next lesson, we'll dive deeper into the details of these statistics.

1.1.2 Evolution of portfolio optimization techniques

1.1.3 Role of machine learning in financial forecasting

1.1.4 Challenges in time-series forecasting and traditional portfolio optimization methods

1.2 Research Objectives

1.2.1 Develop and evaluate a predictive portfolio optimization framework using GPR

1. Implement GPR for multi-asset return prediction 2. Compare various portfolio optimization strategies 3. Develop a probability-based dynamic strategy selection mechanism 4. Evaluate strategy performance considering transaction costs

1.2.2 Research Contributions

1. Novel integration of GPR with dynamic portfolio optimization 2. Probabilistic approach to strategy selection 3. Practical implementation considering transaction costs 4. Comparative analysis of different optimization strategies

Table 1.1: An example for a simple table.

A	B	C	D
1	2	1	2
2	3	2	3

!TeX root = ../main.tex

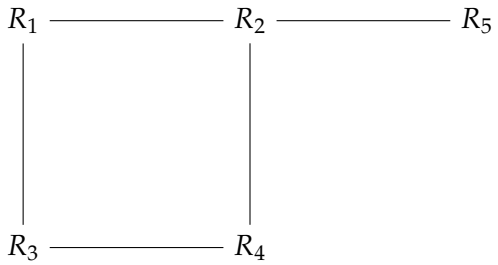

Figure 1.1: An example for a simple drawing.

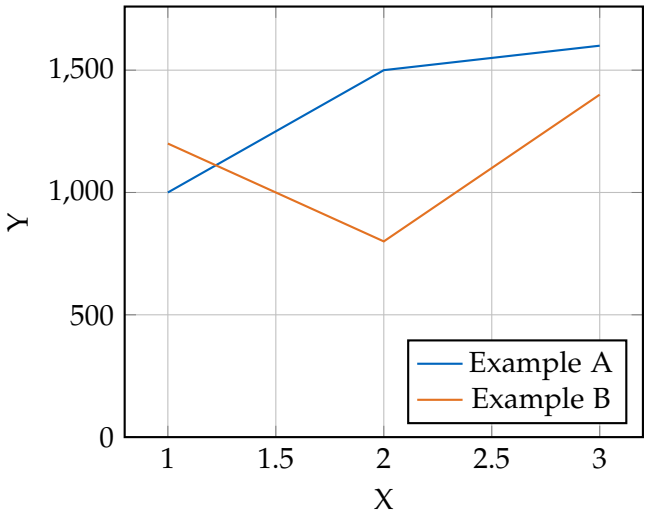

Figure 1.2: An example for a simple plot.

```
SELECT * FROM tbl WHERE tbl.str = "str"
```

Figure 1.3: An example for a source code listing.

2 Previous Literature

2.1 Portfolio Theory

2.1.1 Modern Portfolio Theory (Markowitz)

2.1.2 Post-modern portfolio theory developments

2.2 Risk measures and portfolio optimization

2.3 Machine Learning in Financial Markets

2.3.1 Overview of machine learning applications in finance

2.3.2 Gaussian Process Regression theory and applications

2.4 Time series prediction methods in financial markets

2.5 Dynamic Portfolio Management

2.5.1 Review of dynamic optimization strategies

2.5.2 Transaction costs and portfolio rebalancing

2.5.3 Existing approaches to strategy switching

!TeX root = ../main.tex

3 Methodology

3.1 Data Collection and Preprocessing

3.1.1 Asset selection and justification

Entropy, OrdianlEntroPy is a Python 3 package providing several time efficient, ordinal pattern based entropy algorithms for computing the complexity of one-dimensional time-series.

3.1.2 Data sources and time period

3.1.3 Feature engineering and preprocessing steps

3.1.4 Data normalization and scaling

3.1.5 Treatment of missing data and outliers

3.1.6 Data splitting and cross-validation

3.1.7 sliding window approach

3.1.8 Performance metrics

3.1.9 Baseline models

3.1.10 Hyperparameter tuning

3.1.11 Model evaluation and comparison

3.1.12 Portfolio optimization strategies

3.1.13 Transaction costs and rebalancing

3.1.14 Strategy selection mechanism

3.1.15 Implementation details

3.2 Gaussian Process Regression Model

3.2.1 Introduction to Gaussian Processes

3.2.2 Multi-input Gaussian Process Regression model

3.2.3 Kernel functions selection and hyperparameter optimization

3.2.4 Implementation of rolling window predictions

3.2.5 Model updating mechanism

3.3 Portfolio Optimization Strategies

3.3.1 Traditional Strategies

Maximum return strategy formulation Minimum volatility approach Maximum Sharpe ratio optimization Constraint specifications and justifications as Baseline models

3.3.2 Dynamic Strategy

Probability distribution modeling Strategy switching criteria Position holding logic Transaction cost considerations

4 Results and Analysis

4.1 Model Performance Analysis

4.1.1 Comparison with benchmark models

MSE, Present the prediction accuracy of the GPR models for the target assets.

4.1.2 Analysis of prediction intervals

4.1.3 Model robustness and generalization

4.2 Portfolio Optimization Outcomes

4.2.1 Strategy Performance Comparison

Return analysis Risk metrics Transaction costs impact Strategy switching frequency analysis

Comparative Analysis: Compare the performance of all strategies, highlighting the strengths and weaknesses of each.

4.2.2 Dynamic Strategy Evaluation

Probability threshold sensitivity Strategy switching effectiveness Portfolio turnover analysis Risk-adjusted performance metrics

4.3 Backtesting Results

Provide detailed results from the backtesting, including cumulative returns, volatility, Sharpe ratios, and transaction costs for each strategy.

4.3.1 Transaction Costs impact

4.4 Robustness Tests

4.4.1 Different market conditions

4.4.2 Hyperparameter sensitivity

4.4.3 Out-of-sample performance

4.4.4 Statistical significance tests

5 Discussion

5.1 Interpretation of Results

5.1.1 Key findings and insights

Discuss what the results mean in the context of your research objectives.

5.1.2 Dynamic Strategy Insights

Delve into why the dynamic strategy outperformed others, considering market conditions and model performance.

5.1.3 Model robustness and generalization

5.2 Implications for Practitioners

Explain how your findings can be applied in real-world portfolio management.

5.2.1 Limitations of the approach

Acknowledge any limitations in your study, such as data constraints, model assumptions, or external factors.

5.2.2 Future research directions

Suggest potential areas for further research based on your findings.

5.3 Comparative Analysis

Comparison with existing methods

5.3.1 Advantages and disadvantages

Discuss the pros and cons of your approach compared to traditional portfolio optimization strategies.

5.3.2 Implementation challenges

Discuss any practical difficulties in applying your approach to real-world scenarios.

5.3.3 Market impact considerations

6 Conclusion

6.1 Summary of Findings

6.1.1 Summary of Findings

Recap the main results and how they address your research questions.

6.1.2 Key findings and insights

Discuss what the results mean in the context of your research objectives.

6.2 Recommendations

Offer suggestions for practitioners based on your findings.

6.3 Future Research Directions

6.3.1 Model improvements

Discuss potential enhancements to your model or methodology.

6.3.2 Additional strategy considerations

Suggest new strategies or modifications to existing ones.

6.3.3 Alternative applications

Propose other areas where your approach could be useful.

6.3.4 Scalability considerations

Abbreviations

TUM Technical University of Munich

List of Figures

List of Tables

Bibliography

[Lam94] L. Lamport. *LaTeX : A Documentation Preparation System User's Guide and Reference Manual*. Addison-Wesley Professional, 1994.