Probabilistic Modelling for Adaptive Portfolio Optimization

Tomáš Procházka

Faculty of Nuclear Sciences and Physical Engineering, CTU
supervisor: Ing. Miroslav Kárný, DrSc., UTIA

June 21, 2024

Introduction

- Portfolio
- What is portfolio management?
- Different strategies

Motivation

- Get familiar with decision making and regression
- Apply accumulated knowledge on real-world data
- Possibly come up with a method that gives the investor an advantage

Problem Formulation

Simple formalism of discrete time Markov decision process [3]

Reward Function

$$R(\mathbf{s}_t, \mathbf{a}_t, \mathbf{a}_{t-1}) := 1 + \mathbf{s}_t \cdot \mathbf{a}_t - \sum_{i=1}^{N} \mathbf{c}_{t_i} |\mathbf{a}_{t_i} - \mathbf{a}_{t-1_i}|$$

Expected Return

$$E[R(\mathbf{s}_{t+1}, \mathbf{a}_{t+1}, \mathbf{a}_t) | \mathbf{a}_{t+1}, \mathbf{D}^{(t)}] = \int_{\mathcal{S}} p(\mathbf{s}_{t+1} | \mathbf{a}_{t+1}, \mathbf{D}^{(t)}) R(\mathbf{s}_{t+1}, \mathbf{a}_{t+1}, \mathbf{a}_t) d\mathbf{s}_{t+1}$$

$\mathbf{s}_t$ ... *returns*, $\mathbf{a}_t$ ... *action*, $\mathbf{D}^{(t)}$ ... *data*, $\mathbf{c}_t$ ... *transaction costs*

Methods for Solution

To model the returns of a single or a vector of multiple assets we will be using multivariate regression. In its general form it is stated as

$$\mathbf{s}_t = \mathbf{A}^T \mathbf{z}_t + \epsilon_t,$$

A *... matrix of regression coefficients,* ϵ_t *... normally distributed random noise*

Sufficient statistics matrix

$$\mathbf{V}_t^{-1} = \mathbf{L}_t \mathbf{L}_t^T$$

$$\mathbf{L}_t = \begin{pmatrix} \mathbf{L}_{f_t} & \mathbf{0} \\ \mathbf{L}_{zf_t} & \mathbf{L}_{z_t} \end{pmatrix}$$

The algorithm for effective matrix factorization is available in: [2]

Structure Estimation

- The underlying structure is unknown!
- Method for going through the hypotheses space ([1])
- MLE
- Returns form 7 previous trading days

Used Data

Past daily returns only

- Excluding weekends and holidays
- Including "Black Swan" events

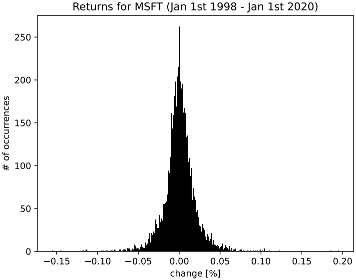
Other meaningful data

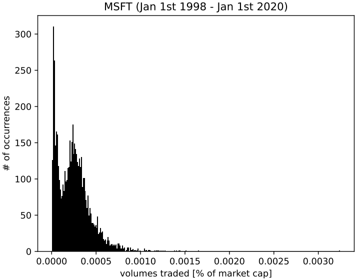
- 13F filings
- Volume traded
- Seasonality

Possible data transformation

Data source: YahooFinance

Used Data


Returns for MSFT (Jan 1st 1998 - Jan 1st 2020)
# of occurrences
0
50
100
150
200
250
−0.15
−0.10
−0.05
0.00
0.05
0.10
0.15
0.20
change [%]


MSFT (Jan 1st 1998 - Jan 1st 2020)
# of occurrences
0
50
100
150
200
250
300
0.0000
0.0005
0.0010
0.0015
0.0020
0.0025
0.0030
volumes traded [% of market cap]

Results

$$\hat{\epsilon}_{t+1} = \mathbf{s}_{t+1} - \hat{\mathbf{A}}_t^T \mathbf{z}_{t+1} = \mathbf{s}_{t+1} + \mathbf{L}_{f_t}^{-T} \mathbf{L}_{zf_t}^T \mathbf{z}_{t+1}$$

$$\Delta_{t+1} = \mathbf{s}_{t+1} - \mathbf{s}_t$$

$$\hat{\sigma} = \frac{1}{T} \sum_{t=t_0}^{T} \hat{\epsilon}_t^2 = / T = 250 / \approx 5.933 * 10^{-4}$$

$$\sigma = \frac{1}{T} \sum_{t=t_0}^{T} \Delta_t^2 = / T = 250 / \approx 1.26 * 10^{-3}$$

Simple 10 day moving average: $\sigma_{MA_{10}} \approx 6.6 * 10^{-4}$ *(10.6% improvement)*

Results

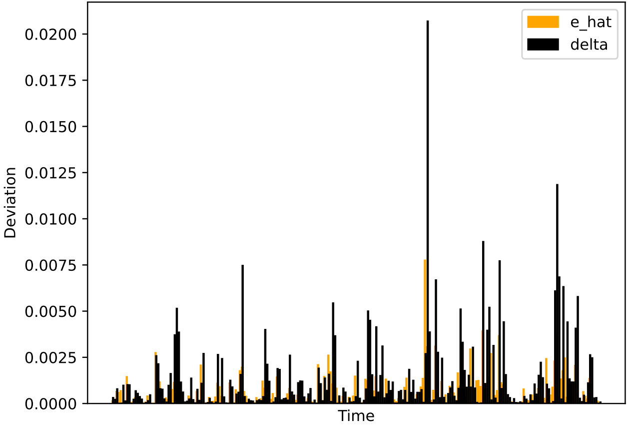


Figure: Errors Over Time

Conclusions and Improvements

- Further testing
- What could be improved?
- Optimization

THANK YOU FOR YOUR ATTENTION

References

M. Kárný and R. Kulhavý.
Structure determination of regression-type models for adaptive prediction and control.
Bayesian analysis of time series and dynamic models, pages 313–345, 1988.

V. Peterka.
A square root filter for real time multivariate regression.
Kybernetika, 11(1):53–67, 1975.

M. Puterman.
Markov decision processes.
Handbooks in operations research and management science, 2:331–434, 1990.