University of Warsaw
Faculty of Economic Sciences

Krzysztof Wojdalski
Nr albumu: 310284

Master's thesis

Praca magisterska
na kierunku QUANTITATIVE FINANCE

The thesis written under the supervision of
dr Juliusz Jablecki Department of Quantitative Finance

2016-01-01

a aa

Data P

Statement of the Supervisor on Submission of the Thesis

I hereby certify that the thesis submitted has been prepared under my supervision and I declare that it satisfies the requirements of submission in the proceedings for the award of a degree.

Date Signature of the Supervisor:

Oświadczenie autora (autorów) pracy

s

Data Podpis autora (autorów) pracy

Statement of the Author(s) on Submission of the Thesis

Aware of legal liability I certify that the thesis submitted has been prepared by myself and does not include information gathered contrary to the law.

I also declare that the thesis submitted has not been the subject of proceedings resulting in the award of a university degree.

Furthermore I certify that the submitted version of the thesis is identical with its attached electronic version.

Date   Signature of the Author(s) of the thesis

Summary

The master's thesis is about the Reinforcement Learning application in the foreign exchange market. The author starts with describing the FX market, analyzing market organization, participants, and changes in the last years. He tries to explain current trends and the possible directions. The next part consists of theoretical pattern for the research - description of financial models, and the AI algorithms. Implementation of the RL-based approach in the third chapter, based on Q-learning, gives spurious results.

Key words*

FX, forex market, trading, reinforcement learning, Algo trading, artificial intelligence, quantitative finance

Area of study (codes according to Erasmus Subject Area Codes List)

11.2 Quantitative Finance

Theme classification

D. Software

The title of the thesis in Polish

Master's thesis

Contents

Chapter 1

Introduction

Financial markets has been interested in computer science methods since the early 1980s. Though there are ways to gain abnormal, positive returns by following traditional ways of investing, such as buy and hold, more modern methods gain on popularity. One of the most popular and emerging category among innovative approaches is the artificial intelligence-based trading. Machine learning has been employed because there is a belief algorithm can be at least as good as human in entrying and exiting positions. Such systems take different inputs but most of them are market-related.

The majority of systems described in the literature aim to maximize trading profits or some risk-adjusted measure such as the Sharpe ratio. Many attempts have been made to come up with a consistently profitable system and inspiration has come from different fields ranging from fundamental analysis, econometric modelling of financial markets, to machine-learning [5, 8]. Few attempts were successful and those that seemed most promising often could not be used to trade actual markets because of associated practical disadvantages. Among others these included large drawdowns in profits and excessive switching behaviour resulting in very high transaction costs. Professional traders have generally regarded those automated systems as far too risky in comparison to the returns they were themselves able to deliver. Even if a trading model was shown to produce an acceptable risk-return profile on historical data there was no guarantee that the system would keep on working in the future. It would cease working precisely at the moment it became unable to adapt to the changing market conditions. This paper aims to deal with the above problems to obtain a usable, fully automated and intelligent trading system. To accomplish this a risk management layer and a dynamic optimization layer are added to a known machine-learning algorithm. The middle layer manages risk in an intelligent manner so that it protects past gains and avoids losses by restraining or even shutting down trading activity in times of high uncertainty. The top layer dynamically optimizes the global trading performance in terms of a traderâ€™s risk preferences by automatically tuning the systemâ€™s hyper-parameters.

While the machine-learning system is designed to learn from its past trading experiences, the optimization overlay is an attempt to adapt the evolutionary behaviour of the system and its perception of risk to the evolution of the market itself. In the past an automated trading system based on 2 superimposed artificial intelligence algorithms was proposed [5]. This research departs from a similar principle by developing a fully layered system where risk management, automatic parameter tuning and dynamic utility optimization are combined. (For an earlier attempt in this direction see [13].) The machine learning algorithm combined with the dynamic optimization is termed adaptive reinforcement learning. Section 2 of this paper briefly discusses the RRL machine-learning algorithm underlying the trading system. Section 3 looks at the different layers of the trading system individually. In 3.1 the modifications to the standard algorithm are set out and in 3.2 and 3.3 the risk management and optimization layers are explained. The impressive performance of the trading system is demonstrated in Section 4 and the final section concludes. 2 2 THE MACHINE-L

1.1. Data

Datasets used for the purpose of this workpaper are from the following databases:

- Thomson Reuters Tick Database
- An aggregator tickdatabase

1.2. Structure

1.3. First chapter

The first part consists of the introduction to the problem. It outlines the whole concept of the AI-related fields in finance. It brought up historical background of finance and computer sciences, and its interdependency. Concretely, it includes the history of implementing first methods in early 80?s, the flash crash in October 1987, first recruitments of 'quants' on the Wall Street in the early 90?s.

1.4. Second chapter

This chapter starts with the critical discussion of models from finance. It includes both classic models, such as CAPM, a gold standard in equity research, and modern ones. The part is descriptive as it regards implicit pros and cons of financial models.

The latter part of the literature review is specifically about algorithmic trading and the methodology of other similar researches, e.g. Sakowski et al. (2013). The last subchapter is about machine learning algorithms that are used in trading. (8-10 pages)

1.5. Third chapter

The third chapter will start with goals of the research. I want to make it clear why this work is important. It was partially discussed in Problem part of this text. This master?s thesis is to find an application of the Reinforcement Learning for financial data. This part will contain hypotheses which are as follows: Algorithms based on artificial intelligence can be fruitful for investors by outperforming benchmarks in both risk and return; Better performance turns out to be true in high-frequency trading and on longer period intervals; Algorithms can learn how to spot overreacting on markets and choose the most under/overpriced security by exploiting time series analysis tools. (2 pages)

Methodology This subchapter contains the description of methodology. It includes all formulas and steps that directed to final results. The algorithm itself will incorporate two environments:

- R - to incorporate libraries for machine learning
- C++ - for code efficiency, it will help in improving performance in bottlenecks

The used algorithm is based on dynamic optimization approach. Besides a value function, there will be several indicators, e.g. RSI, which serve as a base for decision taking of the algorithm. The methodology will include transactional costs, so that the optimization is going to be implemented in a real-like environment

(10 pages)

The value function will be based by several statistics, such as the Sharpe and the Diferental Sharpe Ratio to capture both risk and return. As of now, I cannot enclose the exact form of formulas used in the research but I will provide them as soon as I write the proper code. The output of my algorithm in R will be probably a set of positions $-1, 0, 1$, cumulated returns, and risk measures (not only the Sharpe ratio but also MD, MDD, the Sortino ratio, and others).

How am I going to measure the efficiency of my code? I will implement several benchmarks ? the most logical choice is a buy-and-hold strategy on underlying asset (equities, equity-like securities). The second obvious choice is sort of random walk process. By this, I mean that a part of the algorithm will generate random values for a domain of $-1, 0, 1$ and these values will serve as a position. Of course, this benchmark will not include any transactional costs as this obvious that this extreme case would have an enormous cumulated transactional cost (position would change in $frac23$ of states). When I have the data I am going to discuss my results with other works. Outline the possible directions of future research papers on the issue: What can be implemented? What additionally can be done and measured? Fourth chapter This part consists of conclusions. Once again, I will write what have been done in this master?s thesis, and everything that conclusions should contain.

Chapter 2

FX Market

```
##
##
## processing file:  chapters/subchapters/1_2_FX_Market_Organization.Rnw
##
  |
  |                                                                 |   0%
  |
  |.............                                                    |  20%
##    inline R code fragments

##
##
## processing file:  chapters/subchapters/../../masters_thesis.Rnw
## Quitting from lines 2-24 (chapters/subchapters/../../masters_thesis.Rnw)
## Error in parse_block(g[-1], g[1], params.src):  duplicate label 'r'
```

Chapter 3

Financial models

The following chapter introduces articles that correspond with the subject of the current thesis and are considered as fundamentals of modern finance. Specifically, the beginning contains financial market models. The next subchapter includes basic investment effectiveness indicators that implicitly or explicitly result from the fundamental formulas from the first subchapter.

3.1. Selected financial market models and theory

Works considered as a fundament of quantitative finance and investments are Sharpe (Sharpe 1964), Lintner (Lintner 1965), and Mossin (Mossin 1966). All these authors, almost simultaneously, formulated Capital Asset Pricing Model (CAPM) that describes dependability between rate of return and its risk, risk of the market portfolio, and risk premium. Assumptions in the model are as follows:

- Decisions in the model regard only one period,
- Market participants has risk aversion, i.e. their utility function is related with plus sign to rate of return, and negatively to variance of portfolio rate of return,
- Risk-free rate exists,
- Asymmetry of information non-existent,
- Lack of speculative transactions,
- Lack of transactional costs, taxes included,
- Market participants can buy a fraction of the asset,
- Both sides are price takers,

- Short selling exists,

Described by the following model formula is as follows:

$$E(R_P) = R_F + \frac{\sigma_P}{\sigma_M} \times [E(R_M) - R_F] \tag{3.1}$$

where:

- $E(R_P)$ – the expected portfolio rate of return,
- $E(R_M)$ – the expected market rate of return,
- R_F – risk-free rate,
- σ_P – the standard deviation of the rate of return on the portfolio,
- σ_M – the standard deviation of the rate of return on the market portfolio.

$E(R_P)$ function is also known as Capital Market Line (CML). Any portfolio lies on that line is effective, i.e. its rate of return corresponds to embedded risk. The next formula includes all portfolios, single assets included. It is also known as Security Market Line (SML) and is given by the following equation:

$$E(R_i) = R_F + \beta_i \times [E(R_M) - R_F] \tag{3.2}$$

where:

- $E(R_i)$ – the expected i-th portfolio rate of return,
- $E(R_M)$ – the expected market rate of return,
- R_F – risk-free rate,
- β_i – Beta factor of the i-th portfolio.

3.2. The Modern Portfolio Theory

The following section discuss the Modern Portfolio Theory developed by Henry Markowitz (Markowitz 1952). The author introduced the model in which the goal (investment criteria) is not only to maximize the return but also to minimize the variance. He claimed that by combining assets in different composition it is possible to obtain the portfolios with the same return but different levels of risk. The risk reduction is possible by diversification, i.e. giving proper weights for each asset in the portfolio. Variance of portfolio value can be effectively reduced by analyzing mutual relations between returns on assets with use of methods in statistics (correlation and covariance matrices). It is important to say that any additional asset in portfolio reduces minimal variance for a given portfolio but it is the correlation what really impacts the magnitude. The Markowitz theory implies that for any assumed expected return there is the only one portfolio that minimizes risk. Alternatively, there is only one portfolio that maximizes return for the assumed risk level. The important term, which is brought in literature, is the effective portfolio, i.e. the one that meets conditions above. The combination of optimal portfolios on the bullet.

Bullet figure

The Markowitz concept is determined by the assumption that investors are risk-averse. This observation is described by the following formula:

$$E(U) < U(E(X)) \tag{3.3}$$

where:

- $E(U)$ – the expected value of utility from payoff;
- $U(E(X))$ – utility of the expected value of payoff.

The expected value of payoff is given by the following formula:

$$E(U) = \sum_{i=1}^{n} \pi_i U(c_i) \tag{3.4}$$

where:

- π_i – probability of the c_i payoff,
- $U(c_i)$ – utility from the c_i payoff.

One of the MPT biggest flaws is the fact that it is used for ex post analysis. Correlation between assets changes overtime so results must be recalculated. Real portfolio risk may be underestimated. Also, time window can influence the results.

3.3. The efficient market hypothesis

In 1965, Eugene Fama introduced the efficient market term (Fama 1965). Fama claimed that an efficient market is the one that instanteneously discounts the new information arrival in market price of a given asset. Because this definition applies to financial markets, it had determined the further belief that it is not possible to beat the market because assets are perfectly priced. Also, if this hypothesis would be true, market participants cannot be better or worse. Their portfolio return would be a function of new, unpredictable information. In that respect, the only role of an investor is to manage his assets so that the risk is acceptable.

3.4. Selected investment performance measures

Introduced articles does not include any indicator that would explicitly measure portfolio management effectiveness. Equations that result from the authors' work are important because some of further developed measures are CAPM-based. The most known are the Sharpe ratio, the Treynor ratio, and the Jensen's alpha. Popularity of these indicator comes from the fact that they are easy to understand for the average investor. (Marte 2012) In (Sharpe 1966), the author introduced the $\frac{R}{V}$ indicator, also known as the Sharpe Ratio (S), which is given by the following formula:

$$S_i = \frac{E(R_i - R_F)}{\sigma_i} \tag{3.5}$$

where:

- R_i – the i-th portfolio rate of return,
- R_F – risk-free rate σ_i – the standard deviation of the rate of return on the i-th portfolio.

Treynor (Treynor1965) proposed other approach in which denominator includes β_i instead of σ_i. The discussed formula is given by:

$$T_i = \frac{R_i - R_F}{\beta_i} \tag{3.6}$$

where:

- R_i – the i-th portfolio rate of return,
- R_F – Risk-free rate
- β_i – Beta factor of the i-th portfolio.

Both indicators, i.e. S and T are relative measures. Their value should be compared with a benchmark to determine if a given portfolio is well-managed. If they are higher (lower), it means that analyzed portfolios were better (worse) than a benchmark. The last measure, very popular among market participants, is the Jensen's alpha. It is given as follows:

$$\alpha_i = R_i - [R_F + \beta_i \times (R_M - R_F)] \tag{3.7}$$

where:

- R_i – the i-th portfolio rate of return,
- R_F – Risk-free rate
- β_i – Beta factor of the i-th portfolio.

The Jensen's alpha is an absolute measure and is calculated as the difference between actual and CAPM model-implied rate of return. The greater the value is, the better for the i-th observation.

The differential Sharpe ratio - this measure is a dynamic extension of Sharpe ratio. By using the indicator, it can be possible to capture a marginal impact of return at time t on the Sharpe Ratio. The procedure of computing it starts with the following two formulas:

$$A_n = \frac{1}{n} R_n + \frac{n-1}{n} A_{n-1} \tag{3.8}$$

$$B_n = \frac{1}{n} R_n^2 + \frac{n-1}{n} B_{n-1} \tag{3.9}$$

At $t = 0$ both values equal to 0. They serve as the base for calculating the actual measure - an exponentially moving Sharpe ratio on η time scale.

$$S_t = \frac{A_t}{K_\eta \sqrt{B_t - A_t^2}} \tag{3.10}$$

where:

- $A_t = \eta R_t + (1 - \eta) A_{t_1}$
- $B_t = \eta R_t^2 + (1 - \eta) B_{t_1}$
- $K_\eta = (\frac{1 - \frac{\eta}{2}}{1 - \eta})$

Using of the differential Sharpe ratio in algorithmic systems is highly desirable due to the following facts (Moody and Wu 1997):

- Recursive updating - it is not needed to recompute the mean and standard deviation of returns every time the measure value is evaluated. Formula for A_t (B_t) enables to very straightforward calculation of the exponential moving Sharpe ratio, just by updating for R_t (R_t^2)
- Efficient on-line optimization - the way the formula is provided directs to very fast computation of the whole statistic with just updating the most recent values
- Interpretability - the differential Sharpe ratio can be easily explained, i.e. it measures how the most recent return affect the Sharpe ratio (risk and reward).

The drawdown is the measure of the decline from a historical peak in an asset. The formula is given as follows:

$$D(T) = \max\{max_{0,t \in (0,T)} X(t) - X(\tau)\} \tag{3.11}$$

The Sterling ratio (SR)

The maximum drawdown (MDD) at time T is the maximum of the Drawdown over the asset history. The formula is given as follows:

$$MDD(T) = \max_{\tau \in (0,T)} [\max_{t \in (0,\tau)} X(t) - X(\tau)] \tag{3.12}$$

Chapter 4

Research

4.1. Research Objective

The primary research goal is to evaluate the Reinforcement Learning-based algorithm for multiasset trading. The main idea behind the algorithm deployment is that it can systematically outperform benchmarks in terms of both risk and return. The trading system will be able to spot non-trivial patterns, faster than human, and exploit them. Research will The goal of this project is to assess the possibility of using Reinforcement Learn- ing to create a trading agent which is capable of nding persistent similarities in nancial time series and which learns how to de

ne and exploit deviations from the expected, prevalent behaviour. We design and compare two approaches, a basic approach based on Monte Carlo Control and an extended approach based on Qlearn- ing and value function approximation. The

rst approach is aimed to provide more interpretability, the second approach is to provide better performance. We assess the outcome by the trading performance of the two agents on two weeks of out-of- sample fx market test data with one minute granulation. We set two benchmarks by which we measure trading performance. The percentage return and the Sharpe ratio of the trades the agent engages in should be higher than the return of a buy-and- hold strategy of either of the underlying cointegrated assets. A secondary goal is to draw conclusions about the interactions in parameter settings of the pair trading frame work and how they in uence pro

tability of the pair trade. The objectives are summarized in the following list: The objectives are as follows:

Design and deployment - this part

Bibliography

Fama, Eugene (1965). “The behavior of stock-market prices”. In: *Journal of business* 38.1, pp. 34–105. ISSN: 1098-6596. DOI: 10.1017/CBO9781107415324.004. arXiv: arXiv:1011.1669v3.

Lintner, John (1965). “Security prices, risk, and maximal gains from diversifivation”. In: *Jf* 20.4, pp. 587–615. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.1965.tb02930.x.

Markowitz, Harry (1952). “Portfolio Selection”. In: *1Journal of Finance* 7.1, pp. 77–91. ISSN: 00221082. DOI: 10.2307/2329297.

Marte, Jonelle (2012). *No Title*. URL: http://www.wsj.com/articles/SB10001424052702304692804577283

Moody, J. and Lizhong Wu Lizhong Wu (1997). “Optimization of trading systems and portfolios”. In: *Proceedings of the IEEE/IAFE 1997 Computational Intelligence for Financial Engineering (CIFEr)*, pp. 300–307. DOI: 10.1109/CIFER.1997.618952.

Mossin, Jan (1966). “Equilibrium in a capital asset market”. In: *Econometrica* 34.4, pp. 768–783. ISSN: 00129682. DOI: 10.2307/1910098. URL: http://www.jstor.org/stable/1910098?origin=crossref.

Sharpe, William F. (1964). “Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk”. In: *The Journal of Finance* 19.3, pp. 425–442. ISSN: 00221082. DOI: 10.2307/2329297. URL: http://www.jstor.org/stable/2977928.

— (1966). “Mutual Fund Performance”. In: *The Journal of Business* 39.1, pp. 119–138. ISSN: 0717-6163. DOI: 10.1007/s13398-014-0173-7.2. arXiv: arXiv:1011.1669v3.

List of Figures

List of Tables