Master Thesis

Reinforcement Learning in Stock Market

Author:
Pablo Carrera Flórez de Quiñones

Supervisors:
Valero Laparra Pérez-Muelas
Jorge Muñoz Marí

Contents

1 Introduction

2 Financial Markets

In this section we are providing an introduction to financial markets, first we are going to define them, to explain their properties and actors and to divide them by the type trade they involve. Then we are going deeper into one of the most known financial markets, the stock market and the characteristics of the assets involved in it. Finally, we are going to make a operational introduction to the activity known as trading.

2.1 Fundamentals

A financial market, as defined in [1] or [2], is a market in which people exchange, or trade, financial assets. Although the existence of these markets is not a necessary condition for the creation and exchange of a financial asset, actually these operations can be realized outside of them, in most economies financial assets are created and subsequently traded in some type of financial market.

The function of financial assets, is to transfer funds from those who have surplus funds to those who need funds to invest in tangible assets, in such a way that they also redistribute the unavoidable risks associated with tangible assets among those seeking and providing the funds. In this sense, financial markets have three properties:

- First, the interactions of buyers and sellers in a financial market determine the price of the traded asset. That is, the return on a financial asset is not given or fixed by prior in the market.
- Second, financial markets provide a mechanism for an investor to sell a financial asset. That is, they allow investors to get rid of the obligation of keeping an asset transferring it to another investor.
- Third, financial markets reduce the cost of transacting. The search costs, related to looking for a potential buyer or seller, and information costs, related to the assessment of the financial merits of an asset, are almost totally avoided.

so from these properties surge many ways to classify financial markets:

- Nature of claim: we can distinguish between debt markets, where funds are borrowed and lent, and equity markets, where ownership of securities are issued and subscribed.
- Maturity of claim: we can distinguish between money markets, that allow firms to borrow funds on a short term basis, and capital markets, that allow firms to gain long-term funding to support expansion.
- Seasoning of claim: we can distinguish between primary markets, which leads with newly issued claims, and secondary markets, which leads with financial claims previously issued.
- Seasoning of claim: we can distinguish between cash markets, that trade assets directly, and derivatives instruments markets, that trade financial instruments based on an underlying asset.

2.2 Stock Market

Stocks, also known as equity securities or equities, represent ownership shares in a corporation. Each share of stock entitles its owner to one vote on any matters of corporate governance that are put to a vote at the corporation's annual meeting and to a share in the financial benefits of ownership when those earning are distributed in the form of dividends. There are two types of stocks, the so called common stocks and the preferred stocks. The key difference between these two forms lies in the degree to which the participate in the distribution of earnings and the priority going to each in the distribution of earnings, typically preferred stockholders are entitled to a fixed dividend before common stockholders may receive dividends. We will center our discussion in common stocks since they are by far more the most common type of equity. The two most important characteristics of common stock as an investment are

- Residual claim: this means that stockholders are the last in line of all those who have a claim on the assets and income of the corporation. In a liquidation of the firm's assets the shareholders have a claim to what is left after all other claimants such as the tax authorities, employees, suppliers, bondholders, and other creditors have been paid. For a firm not in liquidation, shareholders have claim to the part of operating income left over after interest and taxes have been paid.
- Limited liability: this means that the most shareholders can lose in the event of failure of the corporation is their original investment. Unlike owners of unincorporated businesses, whose creditors can lay claim to the personal assets of the owner, corporate shareholders may at worst have worthless stock. They are not personally liable for the firm's obligations.

For a common stock, its total vale is the sum of the price of all the shares, that is, the price per share multiplied by the number of shares outstanding, and it is also known as the market capitalization of the stock. For a common stock investor, the return realized by holding one of these shares come from two sources:

- Dividend payments: dividends are distributions made by a corporations to its owners, typically in the form of cash or additional shares of the stock. The payment of dividends is not compulsory, usually young companies do not pay them, but as time goes on and they mature they start doing it.
- Changes in the price of the stock: while holding a share, if the price at a future date is higher than the purchase price there is a capital gain, and if the price at a future date is higher than the purchase price there is a capital loss. When a stockholder is calculating the return from holding a stock from the purchase date to a given point in time, it is actually calculating the gain or loss that he will get in case of selling the stock at this time.

The desire of stockholders to trade their shares has led to the establishment of stock exchanges, organizations which provide marketplaces for trading shares and other derivatives and financial products. Those exchanges can be physical locations, such as the New York Stock Exchange (NYSE), or digital platforms, such as the NASDAQ. These markets, according to the properties of the equities they trade are classified as secondary capital markets.

In the last decades

2.3 Trading

When an investor wants to buy or sell a share of common stock, the price and conditions under which the order is to be executed must be communicated to a broker.

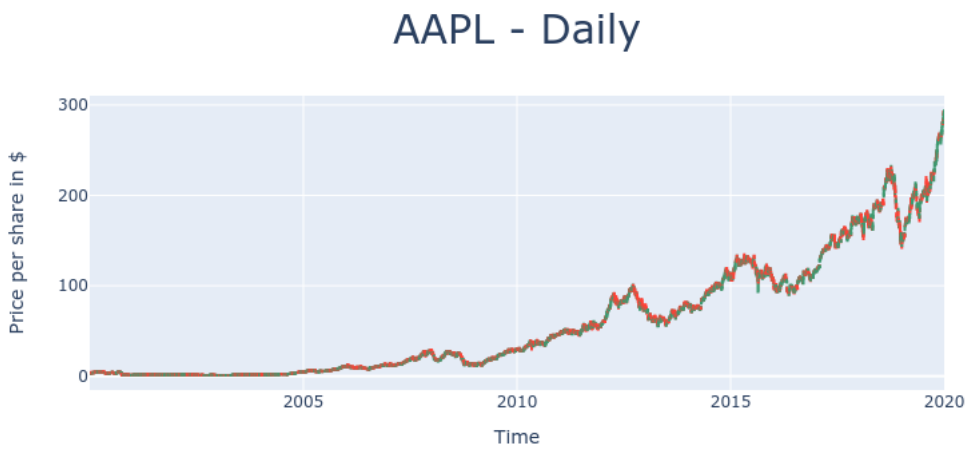

Figure 1: *Example of Candlestick diagram. Source: own.*

FALTA POR AÑADIR

- ORDERS AND ORDER BOOK
- SHORT SELLING
- TAXES/FEES
- INDICATORS/INDEXES
- TECHNICAL ANALYSIS
- ALGORITHMIC TRADING

3 Reinforcement Learning

In this section we are going to make a operational introduction to the field known as Reinforcement Learning. First we are going to introduce Markov Decision Processes, which are the classical formalization of sequential decision making, where actions influence not just immediate rewards, but also subsequent situations, or states, and through those future rewards. Then we are going to present the fundamentals of Temporal-Difference methods, the most extended methods in this field. Finally, we are going to introduce the use of approximate solutions to these methods, which allow us to use state-of-the-art algorithms, such as Neural Networks, to solve efficiently real-world problems.

[3]

3.1 Markov Decision Processes

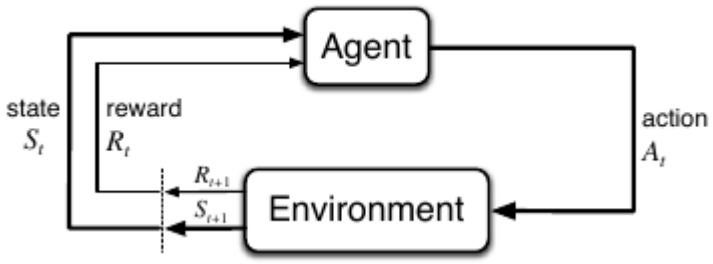

Figure 2: *Schema of the agent-environment interaction in a MDP. Source: [3]*

3.2 Temporal-Difference Methods

[4]

Sarsa (on-policy TD control) for estimating $Q \approx q_*$

Algorithm parameters: step size $\alpha \in (0, 1]$, small $\varepsilon > 0$
Initialize $Q(s, a)$, for all $s \in \mathcal{S}^+, a \in \mathcal{A}(s)$, arbitrarily except that $Q(terminal, \cdot) = 0$

Loop for each episode:
  Initialize S
  Choose A from S using policy derived from Q (e.g., ε-greedy)
  Loop for each step of episode:
    Take action A, observe R, S'
    Choose A' from S' using policy derived from Q (e.g., ε-greedy)
    $Q(S, A) \leftarrow Q(S, A) + \alpha\big[R + \gamma Q(S', A') - Q(S, A)\big]$
    $S \leftarrow S'$; $A \leftarrow A'$;
  until S is terminal

Figure 3: *Pseudocode for Sarsa algorithm. Source: [3].*

Q-learning (off-policy TD control) for estimating $\pi \approx \pi_*$

Algorithm parameters: step size $\alpha \in (0, 1]$, small $\varepsilon > 0$
Initialize $Q(s, a)$, for all $s \in \mathcal{S}^+, a \in \mathcal{A}(s)$, arbitrarily except that $Q(terminal, \cdot) = 0$

Loop for each episode:
  Initialize S
  Loop for each step of episode:
    Choose A from S using policy derived from Q (e.g., ε-greedy)
    Take action A, observe R, S'
    $Q(S, A) \leftarrow Q(S, A) + \alpha\big[R + \gamma \max_a Q(S', a) - Q(S, A)\big]$
    $S \leftarrow S'$
  until S is terminal

Figure 4: *Pseudocode for Q-learning algorithm. Source: [3].*

3.3 Approximate Solutions

[5]

4 Experimental Design

4.1 Environment

[6]

4.2 Agents

4.3 Algorithms

5 Results

6 Conclusions

References

[1] Frank J. Fabozzi, Franco P. Modigliani, and Frank J. Jones. *Foundations of Financial Markets and Institutions*. 4th. Pearson, 2013. ISBN: 9781292021775.

[2] Zvi Bodie, Alex Kane, and Alan Marcus. *Investments*. 11th. McGraw-Hill, 2018. ISBN: 9781259277177.

[3] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. 2nd. Adaptive Computation and Machine Learning. MIT Press, 2018. ISBN: 9780262039246.

[4] Gavin A. Rummery and Mahesan Niranjan. *On-line Q-learning using connectionist systems*. Technical Report CUED/F-INFENG/TR 166. Engineering Department, Cambridge University, 1994.

[5] Volodymyr Mnih et al. *Playing Atari with Deep Reinforcement Learning*. 2013. arXiv: 1312.5602 [cs.LG].

[6] Greg Brockman et al. *OpenAI Gym*. 2016. arXiv: 1606.01540 [cs.LG].

[7] Christopher J. C. H. Watkins and Peter Dayan. "Q-learning". In: *Machine Learning* 8 (1992), pp. 279–292. DOI: 10.1007/BF00992698.