Online Learning and Decision Making Project Proposal
10/12/2019

I. Teammates

Our team members are Jessica Buzzelli (4th year), Soujanya Duggirala (4th year), Kiran Gite (4th year), Diego Granizo (4th year), and Annie Ni (5th year).

II. Main task

We want to determine a portfolio of stocks that will have the highest return and the weights of the stocks in that portfolio. We will first generate three optimal portfolios of stocks using the three full-information setting algorithms discussed in class (randomized weighted majority, multiplicative weight update, HEDGE). Experts will be individual stocks, each having a weight that will be updated based on whether the stock goes up or down. These weights will correspond to the proportion of funds invested in each stock.

Once the algorithms have run for a determined time T, we will then test the performance of each of the portfolios determined by the three full-information algorithms in a multi-armed bandit setting with Thomson sampling. We plan to do this live. The algorithm will select one portfolio to play at the beginning of a day, and then will receive feedback about how that portfolio performed during that day, allowing us to update the posterior distribution for that portfolio. By repeating this for many days, we hope to determine the portfolio with the highest average rate of return.

III. Data sources

Our data sources will be Alpha Vantage API for historical stock data and Yahoo Finance for live stock data.

IV. Deliverables and timeline (end of Oct, end of Nov, final project)

At the end of October, we plan to pull all of our historical data and format it in a way that our algorithms can use. We will also run our full-information algorithms in order to determine our three portfolio arms. We will then set up our code for live updates in the market prior to the midterm report due November 12th. Our implementation of live updates through Thomson Sampling in the stock market will end at the end of November. Then, we will prepare for our final reports on December 2nd and 10th, and our final presentation on December 3rd.

V. Concepts from class

Our experts in this project will first be stocks and then portfolios. This project will be using three update schemes: randomized weighted majority multiplicative weight update, and HEDGE to get the weights of the stocks for each portfolio.
Then, we are treating this as a multi-armed bandit problem using Thompson Sampling with a currently-undecided prior (Gaussian, Dirichlet) in order to converge upon the true return of the best portfolio.

VI. Current literature

Online learning has long been explored within the context of the stock market.Li & Hoi, (2014) provide a comprehensive survey of the research on the matter. The survey discusses various algorithms and approaches many which closely follow the on-line learning framework we've seen in class. Of the following papers we will discuss the most pertinent to our approach.

Classifications	Algorithms	Representative References
Benchmarks	Buy And Hold	
	Best Stock	
	Constant Rebalanced Portfolios	Kelly [1956]; Cover [1991]
Follow the Winner	Universal Portfolios	Cover [1991]; Cover and Ordentlich [1996]
	Exponential Gradient	Helmbold et al. [1996, 1998]
	Follow the Leader	Gaivoronski and Stella [2000]
	Follow the Regularized Leader	Agarwal et al. [2006]
	Aggregating-Type Algorithms	Vovk and Watkins [1998]
Follow the Loser	Anti-Correlation	Borodin et al. [2003, 2004]
	Passive Aggressive Mean Reversion	Li et al. [2012]
	Confidence Weighted Mean Reversion	Li et al. [2011b, 2013]
	Online Moving Average Reversion	Li and Hoi [2012]
	Robust Median Reversion	Huang et al. [2013]
Pattern-Matching–Based Approaches	Nonparametric Histogram Log-Optimal Strategy	Györfi et al. [2006]
	Nonparametric Kernel-Based Log-Optimal Strategy	
	Nonparametric Nearest Neighbor Log-Optimal Strategy	Györfi et al. [2008]
	Correlation-Driven Nonparametric Learning Strategy	Li et al. [2011a]
	Nonparametric Kernel-Based Semi-Log-Optimal Strategy	Györfi et al. [2007]
	Nonparametric Kernel-Based Markowitz-Type Strategy	Ottucsák and Vajda [2007]
	Nonparametric Kernel-Based GV-Type Strategy	Györfi and Vajda [2008]
Meta-Learning Algorithms	Aggregating Algorithm	Vovk [1990], [1998]
	Fast Universalization Algorithm	Akcoglu et al. [2002, 2004]
	Online Gradient Updates	Das and Banerjee [2011]
	Online Newton Updates	
	Follow the Leading History	Hazan and Seshadhri [2009]

Helmbold, Schapire, Singer, & Warmuth, (1998) provides an algorithm to determine the best constant-rebalanced portfolio. Similar to the first stage of our project, the paper proposes an on-line algorithm featuring a stock portfolio composed of N stocks, each with their respective weight used to determine the allocation of funds in the portfolio. The paper discusses multiplicative weight update but chooses an approach most similar to online regression. Choosing to select weights by maximizing the following:

$$F(w^{t+1}) = \eta(w^{t+1} \cdot x^t) - d(w^{t+1}, w^t)$$

The user picks a learning rate η, determining how much to change the stock weights, and a distance function $d(\bullet)$ used to mediate the change of weights. While different distance functions led to different results, Hembold favors relative entropy and through some manipulation, the paper concludes with the same multiplicative weight update seen in class. As a result the paper also proves the a similar upper bound where their algorithm will approach the performance of the best constant-rebalanced portfolio.

VII. Bibliography

Helmbold, D. P., Schapire, R. E., Singer, Y., & Warmuth, M. K. (1998). On-Line Portfolio Selection Using Multiplicative Updates. *Mathematical Finance*, *8*(4), 325–347. https://doi.org/10.1111/1467-9965.00058

Li, B., & Hoi, S. C. H. (2014). Online Portfolio Selection: A Survey. *ACM Comput. Surv.*, *46*(3), 35:1–35:36. https://doi.org/10.1145/2512962