Risk Related Portfolio Selection with Hedge Algorithm in Full Feedback Multi-armed Problem

Jin Tian & Zheng Zihao

2022 March

Abstract

In this paper, we used algorithm in multi-armed bandit problem to anlysis portfolio selection problem. We combined Hedge algorithm in full feedback problem with model in contextual bandit to banlance both return and risk. To our model, we assume data is independent identically distributed(i.i.d.). In the empirical part, we used data from Chinese market to check our model. It outperformed other simple models in many aspects.

1 Introduction

Multi-armed bandit problem is very useful in selection problems. Recently, developments in AI used this method in AlphaGo which attarcted much attention. Other applications like online-advertising and clinical trails were also widely discussed. This method usually assume data to be Markov Process and use elporation versus exploitation to construct the framework. In portfolio selection problems, there are many algorithm that could be used. Traditionally, algorithms like ϵ-greedy algorithm upper confidence bounds (UCB) algorithm raised were widely used.

Portfolio selction problem, with a long history, dated back to 1952 with Markowitz [1] problem which introduced the concept of risk-return balance. Also, in modern days, people used value at-risk, expected shortfall and others to quantify risks. Besides, Portfolio selection problem could be optimized with target like turnover, maxdrawndown etc.

To research portfolio selection problem under multi-armed bandit framework has recently drawn much attention. Traditional methods are used. Meanwhile, the target function could be risk-return balance function which was mentioned in Sani. et al.(2012) [2]. They studied the problem where the learner's objective is to minimize the mean variance defined as $\sigma^2 - \rho\mu$ and proposed algorithms MV-LCB and MV-DSEE. Others like Galichet et al.(2013) [3] used the conditional value-at-risk to be the objective and proposes the MARAB algorithm. In Huo & Fu(2017) [4], they used UCB algorithm and expected shortfall to construct new algorithm.

The new models that controls risks raised another problem about the measurement of risks. Clearly, there is no standard answer to it. Markowitz(1952) [1] used variance as a standard. Other methods like value at-risk and expected shortfall were also used. The optimization problems with target function like this were also well researched in academia. Like Ledoit (2004) [5], the author developed shrinkage methods on Markowitz problem to select sparse portfolio. Rockafellar & Uryasev(2000) [6] provided optimization of expected shortfall to select portfolio.

In this paper, we will combine Hedge algorithm with minimization of expected shortfall which could also be interprated as contextual bandit problem to construct a portfolio selection algorithm.

2 Models

Our models include two parts. The first part is the full feedback model, while the second part is a contextual bandit part with target function about expected shortfall.

In the first place, our algorithm chose Hedge algorithm in the full feedback problem. In detail, we define.

Protocol 1 Bandits with full feedback and adversarial costs in portfolio selection

Parameters: K arms, T rounds (both known)

In each round $t \in [T]$:

1. Adversary chooses costs $c_t(a) = \frac{1}{1+e^{-r_t}} \in (0,1)$ for each arm $a \in [K]$. Where r_t is the return in last period.
2. Algorithm picks arm $a_t \in [K]$
3. Algorithm incurs cost $c_t(a_t)$ for the chosen arm.
4. The costs of all arms, $c_t(a) : a \in [K]$, are revealed.

Here, we define r_t to be the return. And we used *sigmod* formula to generate the cost function. This is a reflection between true security return and $(0,1)$.

Now, we can define our full feedback part:

Algorithm 1 Hedge algorithm for portfolio selection

Parameters: $\epsilon \in (0, \frac{1}{2})$

Initialize the weights as $w_1(a) = 1$ for each arm a

For each round t

1: Let $p_t(a) = \frac{w_t(a)}{\sum_{a'=1}^{K} w_t(a')}$;

2: Sample an arm a_t from distribution ;

3: Observe cost $c_t(a)$ for each arm a

4: For each arm a , update its weight $w_{t+1}(a) = w_t(a) \cdot (1-\epsilon)^{c_t(a)}$

To this model,it was called the Hedge algorithm.It was proved that its regret which was defiened by:

$$R(T) = \mathrm{cost}(\mathrm{ALG}) - \min_{a \in [K]} \mathbb{E}[\mathrm{cost}(a)] \tag{1}$$

This regret number, is proved to be $O(\sqrt{T \log K})$ in Slivkins(2022) [7], which was quite well.

The second part is to get a weight ω for minimising the expected shortfall, to define expected shortfall, we have to define $\alpha - VaR$ first:

$$\mathrm{VaR}_\alpha(X) \stackrel{\text{def}}{=} \inf\{x \in \mathbb{R} : \mathbb{P}(x + X < 0) \leq 1 - \alpha\} \tag{2}$$

The $\beta - ES$ is then:

$$\mathrm{ES}_\beta(X) \stackrel{\text{def}}{=} \frac{1}{1-\beta} \int_\beta^1 \mathrm{VaR}_\alpha(X) \mathrm{d}\alpha \tag{3}$$

In Rockafellar & Uryasev(2000) [6], the authors defined a method to slove the portfolio selection problem like this:

$$\min_{\boldsymbol{\omega} \in W} \quad ES_\beta \left(\boldsymbol{\omega}^\top \boldsymbol{R}_t\right) \tag{4}$$

It is the equivalent to solve this problem:

$$\min_{(\boldsymbol{\omega}, \boldsymbol{\alpha}') \in W \times \mathbb{R}} \tilde{F}_\beta(\omega, \alpha') = \alpha' + \frac{1}{q(1-\beta)} \sum_{k=1}^{q} \left[-\mathbf{x}^T \mathbf{y}_k - \alpha'\right]^+ . \tag{5}$$

Here, we can see this method used past history and put it as a kind of context, it could be interprated as a kind of contextual bandit to some extent.

Then, our algorithms could be combined.

$$\omega_t \stackrel{\text{def}}{=} \lambda\omega_t^H + (1-\lambda)\omega_t^{ES} \tag{6}$$

Where λ represents the relative weight to return and variance. In our paper, we set this to be 0.4. ω_t^H is a vector where the arm selected by Algorithm 1 a_t is 1, others are $0.\omega_t^{ES}$ is from (5).

Formally, our algorithm is:

Algorithm 2 Combined Algorithm

Parameters: β, λ

For each round t

1: Select arm a_t from Algorithm 1 and compute ω_t^H;
2: Generate ω_t^{ES} from (5) at the confidence level β ;
3: Generate ω_t from (6) with the given parameter λ ;
4: Observe returns $\mathbb{R}_t$ and update information for (5) and (6);

3 Data and Empirical Results

In this part we calculated several multi-armed bandit methods to compare with our algorithm. They included UCB, $\epsilon-$greedy and mean-average method. Mean-average means to calculate average return for past 8 periods and use it as prediction where we long the highest and short the lowest.

The second important thing is our data, we used weekly data from Chinese Market. We excluded some finance industry firms and randomly select seven firms for testing, we used 8 years data do finish the analyses. Later, we calculted annual return, standard deviation and Sharpe Ratio to compare results.

We have also drawn the cumulative return plot.For further details, please see appendix.

	return	std	sharpe
Hedge	-0.05057	0.511281	-0.09891
EpsilonGreedy	-0.05893	0.412309	-0.14292
UCB1	-0.16177	0.427973	-0.37798
MeanAverage	-0.29159	0.550303	-0.52987
Combined	0.000714	0.283736	0.002518

Table 1: performance Comparison

From this results we can see that our did outperformed other methods. Which means our results were quite useful.

4 Conclusions

From this paper, we developed our model to combine Hedge algorithm with ES. This combination was included more control variables in the model. We assumed data to be IID. Generally speaking, it was quite useful. Also, models could be improved like use other cross-validation methods to optimize paramters etc.

5 Reference and Appendix

Please see https://github.com/JinTian0717/Jin-Tian