Long and Short Term Risk Control for Online Portfolio Selection

Yizhe Bai, Jianfei Yin(✉), Shunda Ju, Zhao Chen, and Joshua Zhexue Huang

College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China
byz1994@yeah.net, {yjf,zx.huang}@szu.edu.cn, jushunda2017@email.szu.edu.cn, rirnel@163.com

Abstract. Online portfolio selection is to allocate the capital among a set of assets to maximize cumulative returns. Most of online portfolio selection algorithms focus on maximizing returns without effectively controlling risk of loss in return. Further, many risk control algorithms use the maximum drawdown, the Sharpe ratio, and others as risk indicators. However, these risk indicators are not sensitive to the short-term of loss in return. This paper proposes the Long and Short Term Risk (LSTR) control algorithm for online portfolio selection. LSTR achieves high return and low risk by combining the effects of two parameters. The first parameter learns the long-term risk of the market, and its posterior probability changes slowly according to the mean reversion theory. The second parameter senses the short-term risk of the market and makes a quick response to changes in short-term returns. Through the multiplication of the two parameters, the risk control ability of online portfolio selection is effectively improved. The experimental results of the six datasets demonstrate that the performance of LSTR is better than the online portfolio selection algorithms with risk control and those without risk control.

Keywords: Risk control · Long term learning · Short term control · Mean reversion theory

1 Introduction

The primary task of online portfolio select algorithms is to periodically adjust the capital ratio among a set of risky assets to maximize the final return. In the past few years, many online portfolio selection algorithms have been developed with this design goal in mind, such as UP [2], EG [3], ONS [1], CFR-OGD [4] and so on. However, these algorithms lack explicit risk control of loss in return, which leads to serious losses in the market with large reversals [6,12]. Therefore, people study the definition of risk from the perspective of probability measurement of returns,

G. Li et al. (Eds.): KSEM 2020, LNAI 12275, pp. 472–480, 2020.
https://doi.org/10.1007/978-3-030-55393-7_42

such as variance, semi-variance, and the probability of adverse outcomes [5]. Any portfolio selection strategy that implements any of the three risk definitions requires a long trading period to collect enough observations of returns to make the empirical estimation of probability of risk related events. Thus, the returns generated by the algorithms based on the three risk definitions show stability in the long term but volatility in short term.

The improvement of the existing online portfolio selection algorithms in the aspect of risk is helpful to obtain effective trading algorithms [9,12]. One approach is to track the maximum drawdown of each portfolio vector, i.e., each expert and give more capital to the expert with a smaller maximum drawdown at each trading period [9]. However, the maximum drawdown is a long-term variable that describes the market and is not sensitive to the performance changes of portfolios in the short term. Another popular solution is to use reinforcement learning. In order to control the risk, the variables describing the risk are added to reward functions [8,10]. The Sharpe ratio is the most commonly used reward variable for reinforcement learning with risk control. It is a measure of the excess return per unit of risk for a trading strategy. This approach shows some advantages in terms of annual performance statistics. However, the Sharpe ratio needs to calculate the mean and standard deviation of the portfolio returns. The mean and standard deviation are long-term indicators that describe the market. Their short-term changes are small.

Based on the above analysis, we design an algorithm to control the long-term and short-term risks of loss in return by changing the proportion of the risk-free asset directly, called Long and Short Term Risk Control (LSTR). We summarize our main contributions as follows, i) we define a capital-ratio updating equation for online portfolio selection, where the multiplication effect of the long-term risk control parameter λ and the short-term risk control parameter η is found to be effective in the equation; ii) A risk indicator random variable $\mathcal{C}$ is defined to parameterize the probability of λ and control η ; iii) Based on the definition of $\mathcal{C}$, the learning and control algorithms for λ and η are proposed; iv) Based on λ and η , the LSTR algorithm is designed to improve the long and short term risk control ability of portfolios.

2 Related Work

The most popular methods of defining risk are variance, semi-variance, and the probability of adverse outcomes [5]. Taking the return as a random variable, the variance of the return is defined as the risk by the variance method. The semi-variance method computes the variance of the return only when the return is below the mean of the return. The probability of adverse outcomes, given by Roy's Safety First Criterion, defines a probability of the loss with respect to a target return. That is, $P\{(\phi - \zeta) \geq z\}$, where ϕ denotes target return, and $\phi - \zeta$ denotes the loss of the portfolio.

Mohr et al. proposed two risk control algorithms: RAPS and CRAPS [9] based on UP [2]. The algorithms track the experts with the lowest maximum drawdown. RAPS allocates higher ratio of capital to the lower maximum drawdown experts so far. CRAPS takes the maximum drawdown and the winner's capital into consideration to obtain the final expert capital ratio. Shen Weiwei et al. [10] used Bandit Learning to solve online portfolio problems with risk. They modeled the portfolio as multiple arms. The reward function for each arm is expressed as the Sharpe ratio. They then used the upper confidence bound to select the optimal arm. Liang Zhipeng et al. [8] implemented two reinforcement learning algorithms, DDPG and PPO in portfolio management. They used the Sharp ratio objective function to carry on the experiment, but the method is not very useful.

3 Problem Definition

Consider a financial market with n trading days and m assets. The closing price of an asset during the t period is represented by a price vector $\boldsymbol{p_t} \in \mathbb{R}^m_{++}$, and it's component p^i_t denotes the closing price of the ith asset on the tth trading day. The change in market price can be expressed by an m-dimensional relative price vector $\boldsymbol{x_t} = (x^1_t, x^2_t, ..., x^m_t) \in \mathbb{R}^m_+, t = 1, 2, ..., n$, where the element x^i_t denotes the ratio of the closing price of the ith asset in the tth period and the closing price of the $t-1$th period, expressed as $x^i_t = \frac{p^i_t}{p^i_{t-1}}$. Historical market sequence starts from period 1 to n, that is, $\boldsymbol{x^n_1} = \{\boldsymbol{x_1}, \boldsymbol{x_2}, ..., \boldsymbol{x_n}\}$. At the beginning of the tth period, the capital allocation can be represented by a portfolio vector $\boldsymbol{b_t} = (b^1_t, b^2_t, ..., b^m_t) \in \mathbb{R}^m_+, t = 1, 2, ..., n$. The element b^i_t represents the proportion of the capital of the ith asset in the tth period. $\boldsymbol{b_t} \in \Delta_m$, where $\Delta_m = \{\boldsymbol{b} : \boldsymbol{b} \succeq 0, \boldsymbol{b}^\top \mathbf{1} = 1\}$. For the tth period, $s_t = \boldsymbol{b_t^\top} \boldsymbol{x_t}$ is the daily return. The final cumulative capital after a sequence of n periods is: $S(X^n_1) = S_0 \Pi^n_{t=1} \boldsymbol{b_t^\top} \boldsymbol{x_t} = S_0 \Pi^n_{t=1} s_t$. In this paper, we define $S_0 = 1$.

4 Long and Short Term Risk Control Algorithm

In this selection, we first proposed a capital-ratio updating equation for online portfolio selection. Then we define a risk indicator random variable to parameterize the long and short term risk control parameters. Finally, the LSTR algorithm is derived based on the long and short term control parameters.

4.1 Risk Control Method

In order to control the risk of loss in return in any market situation, we can reasonably add a risk-free asset as the first asset in a given portfolio. This paper assumes that the risk-free asset is cash, and its daily return is 1. Thus, the portfolio vector $\boldsymbol{b_t}$ becomes an $m+1$ dimensional vector, that is, $\boldsymbol{b_t} \in \mathbb{R}_+^{m+1}$. After many experiments, we found that $\boldsymbol{b_t}$ has excellent performance of risk control when it is defined by a capital-ratio updating equation:

$$\boldsymbol{b_t} = \lambda\eta\boldsymbol{e_1} + (1-\lambda\eta)(1-\boldsymbol{e_1}) \odot \boldsymbol{b_t}, \tag{1}$$

where $\odot$ denotes the element-wise product, the vector $\boldsymbol{e_1}$ is the unit vector with the first component is 1 and the others are 0. $\lambda \in (0,1)$ is the long-term risk control parameter. $\eta \in (0,1]$ is the short-term risk control parameter. λ and η are derived variables, which are parameterized by a risk indicator random variable $\mathcal{C}$ defined in the next section.

4.2 Long-Term Risk Control

The risk indicator random variable $\mathcal{C}$ is given as follows.

Definition 1. *Let s_t denote the daily return of the online portfolio selection on the tth trading day, ϕ denote the daily target return set by the investor, and $z \geq 0$ denote the loss that the investor can withstand per trading day:*

$$\mathcal{C}(s_t) = \begin{cases} 0, & \phi - s_t > z \\ 1, & \phi - s_t \leq z \end{cases} \tag{2}$$

where $\{\mathcal{C}=0\}$ and $\{\mathcal{C}=1\}$ represent event $\{\phi - s_t > z\}$ and event $\{\phi - s_t \leq z\}$, respectively.

In Definition 1, the event $\{\phi - s_t \leq z\}$ and the event $\{\phi - s_t > z\}$ respectively indicate whether the loss investor is acceptable. We count the number of times $\mathcal{C}=0$ ($\mathcal{C}=1$) to predict the probability q of $\mathcal{C}=0$ ($\mathcal{C}=1$) on a single trading day. We use the Beta distribution to describe the probability value q. The Beta distribution is characterized by two shape parameters, α and β: $P(q;\alpha,\beta) = \frac{1}{B(\alpha,\beta)} q^{\alpha-1}(1-q)^{\beta-1}, 0 \leq q \leq 1, \alpha > 0, \beta > 0$, where $B(\alpha,\beta) = \int_0^1 q^{\alpha-1}(1-q)^{\beta-1}dq$ denotes the Beta function. The long-term risk control parameter λ obeys the Beta distribution, i.e., $\lambda \sim P(q;\alpha,\beta)$.

In the Bernoulli trials of N times, ν denotes the number of occurrences of the random event $\mathcal{C}=0$. We use the mean reversion theory [11] to describe the likelihood function $L(\nu|q)$, and the exponential of q is $N-\nu$ instead of ν: $L(\nu, N-\nu|q) = \binom{N}{\nu} q^{N-\nu}(1-q)^{\nu}$. Mean reversion theory points out that a stock's price will tend to move to the average price over time. The use of mean reversion strategies in the short term may be volatile, but it is very stable in the long-term run [11]. According to Bayes' theorem, we can get the following result: after a trading day, when $\mathcal{C}=0$, the posterior probability of q becomes:

$$P(q|\nu, N-\nu) = P(q;\alpha,\beta+1). \tag{3}$$

When $\mathcal{C} = 1$:

$$P(q|\nu, N-\nu) = P(q; 1+\alpha, \beta). \tag{4}$$

λ can be estimated by three values λ^S, λ^E and λ^M. λ^S takes values randomly from the Beta distribution. λ^E and λ^M denote the mean and mode of the Beta distribution, respectively.

Algorithm 1. Calculation method of η.

Input:
Adjust parameter τ;
Output:
Proportion of cash in the portfolio η;
Initialize $\eta = 1$, $\kappa = 0$;
for $t = 1, 2, ..., n$ **do**
  if $\mathcal{C} = 1$ **then**
    $\kappa{+}{+}$;
    Update η using Eq.(5);
  else
    $\kappa = 0$, $\eta = 1$;
  end if
end for

4.3 Short-Term Risk Control

In Eq. (1), η denotes the proportion of cash in the portfolio. It is defined by:

$$\eta = \frac{1}{1 + exp\{\kappa + \tau\}}, \tag{5}$$

where κ denotes the number of consecutive events $\mathcal{C} = 1$, and τ is a constant that determines how much η falls when $\kappa = 1$. When the $\kappa + \tau \in [-4, 4]$ in Eq. (5), η decreases exponentially in the region $(1, 0)$. So when $\kappa = 1$, to have $\kappa + \tau \in [-4, 4]$, the constant τ must be in $[-5, 3]$. By Eq. (5), we know that the parameter η can quickly respond to the short-term returns of a portfolio.

The calculation method of η is shown in Algorithm 1. The initial cash is $\eta = 1$, indicating that it has not entered the market. The count variable κ is set to be 0 initially. As the number of consecutive occurrences of $\mathcal{C} = 1$ increases, κ also increases. So that, η drops to near 0 to ensure that capital can be used to invest quickly in other risky assets according to Eq. (1). Once the event $\mathcal{C} = 0$ happens, the continuous count parameter κ will be set to 0 immediately, and the proportion of capital for cash will be restored to 1 in order to avoid the risk of short-term loss of earnings.

Algorithm 2. LSTR algorithm.

Input:

Historical market sequence $\boldsymbol{x_1^n}$;
The maximum daily loss z that the investor can bear;
Adjust parameter τ;
The target return ϕ set by the investor;
An online portfolio selection A;
$\lambda \in \{\lambda^S, \lambda^E, \lambda^M\}$;

Output:

Final cumulative wealth S_n;

1: Initialize $S_0 = 1$, $b_1 = (\frac{1}{m+1}, ..., \frac{1}{m+1}) \in \mathbb{R}_+^{m+1}$, $\alpha = \beta = 2$, $\eta = 1$, $\kappa = 0$;
2: **for** $t = 1, 2, ..., n$ **do**
3: Using online portfolio strategy A to compute a portfolio b_t;
4: Calculate the daily return $s_t = \boldsymbol{b_t}^\top \boldsymbol{x_t}$;
5: Updates cumulative return $S_t = S_{t-1} \times s_t$;
6: Through Eq.(2), get $\mathcal{C} = 0$ or $\mathcal{C} = 1$;
7: **if** $\mathcal{C} = 1$ **then**
8: $\kappa + +$;
9: Update η using Eq.(5);
10: Update α and β by Eq.(4);
11: **else**
12: $\kappa = 0$, $\eta = 1$;
13: Update α and β by Eq.(3);
14: **end if**
15: Computes λ;
16: Adjust $\boldsymbol{b_t}$ using Eq.(1);
17: Portfolio strategy A updates the online portfolio selection rules;
18: **end for**

4.4 LSTR Algorithm

The LSTR algorithm is shown in Algorithm 2. First, the two shape parameters of the Beta distribution are initialized by $\alpha = \beta = 2$. When $\alpha > 1$ and $\beta > 1$, the Beta distribution has a unique mode, and λ^M works fine. LSTR obtains the portfolio vector $\boldsymbol{b_t}$ of the basic online portfolio selection algorithm through step (3). In step (7)–(14), LSTR automatically update the parameters through the risk indicator variable $\mathcal{C}$, including η, α, and β. By adjusting $\boldsymbol{b_t}$ in step (16), the adjusted $\boldsymbol{b_t}$ has a different proportion of cash than that of the portfolio selection A. Finally, the portfolio strategy is updated according to the rules of the online portfolio selection A.

5 Experimental Results

In Table 1, we show the six datasets used in this experiment. The datasets of this experiment have one more cash than the original datasets. The NYSE, DJIA, MSCI, and TSE datasets are all widely used[1] [7]. Two datasets, FSE and SP500,

[1] These datasets are downloaded from the https://github.com/OLPS/OLPS.

were collected at Yahoo Finance[2]. We choose six algorithms to compare with LSTR. UP [2], EG [3], ONS [1] are three online portfolio selection algorithms without risk control; RAPS and CRAPS [9] are two online portfolio selection algorithms with risk control.

Parameter Choices: For LSTR, $\phi = 1.003, z = 0.004$. In this experiment, let $\tau = 0$. This means that η directly drops from 1 to about $\frac{1}{1+e} \approx 0.268$. The parameters of other algorithms are set according to the suggestions of the authors. $LSTR^S$, $LSTR^E$, and $LSTR^M$ use three methods: λ^S, λ^E, and λ^M, respectively. We can see from Table 2, the results of $LSTR^S$ are slightly better than $LSTR^E$ and $LSTR^M$ in most cases. But $LSTR^E$ and $LSTR^M$ are more convenient and faster to calculate than $LSTR^S$.

Table 1. Summary of datasets

Dataset	Region	Time frame	Periods	Assets
NYSE	US	[1962.07.03–1984.12.31]	5651	37
DJIA	US	[2001.01.14–2003.01.14]	507	31
MSCI	Global	[2006.01.04–2010.03.31]	1043	25
FSE	GER	[2002.01.01–2003.06.16]	369	16
SP500	US	[2008.04.07–2011.01.14]	700	21
TSE	CA	[1994.04.01–1998.12.31]	1259	89

Cumulative Wealth (CW): In these six datasets, the cumulative wealth of $LSTR^M$ is superior to that of UBAH and the original algorithm. On NYSE, the cumulative wealth of $LSTR^M$ is about three times higher than that of UP and EG, and about seven times higher than that of ONS. The cumulative wealth of RAPS and CRAPS in six datasets is less than all LSTR.

Sensitivity of Risk: The Sharpe Ratio (SR) is an indicator that combines both return and risk. In Table 2, On the five datasets NYSE, MSCI, FSE, SP500, and TSE, the Sharpe ratios of LSTR are higher than those of the original algorithms. On the DJIA dataset, the Sharpe ratios of all algorithms are negative. The Maximum Drawdown (MDD) is the maximum observed loss of the portfolio from peak to valley before reaching a new peak. In these six datasets, the MDDs of LSTR are lower than that of UBAH and the original algorithms. Table 2 show that the MDDs and SRs of LSTR are better than those of RAPS and CRAPS on the six datasets. The results show that the risk sensitivity of LSTR is higher than that of RAPS and CRAPS.

[2] These datasets are collected from the https://finance.yahoo.com.

Table 2. Results of performance

Dataset	indices	UBAH	RAPS	CRAPS	UP	$LSTR_{UP}^{S}$	$LSTR_{UP}^{E}$	$LSTR_{UP}^{M}$	EG	$LSTR_{EG}^{S}$	$LSTR_{EG}^{E}$	$LSTR_{EG}^{M}$	ONS	$LSTR_{ONS}^{S}$	$LSTR_{ONS}^{E}$	$LSTR_{ONS}^{M}$
NYSE	CW	17.6211	24.5010	24.4553	24.5299	79.2908	79.9822	**80.1453**	24.4025	**78.0626**	77.8457	77.8463	17.2372	**122.0882**	122.0	121.9517
	SR	0.3080	1.0027	1.0062	1.0035	1.8397	1.8467	**1.8476**	1.0012	**1.8289**	1.8252	1.8254	1.0414	**1.3518**	1.3516	1.3516
	MDD	0.5350	0.1155	0.1164	0.1153	0.0654	0.0653	**0.0652**	0.1138	**0.0653**	0.0655	0.0655	0.1348	**0.1376**	0.1381	0.1381
DJIA	CW	0.7666	0.8382	0.8381	0.8374	0.8464	**0.8593**	0.8586	0.8127	0.8505	0.8528	**0.853**	1.5193	1.5131	1.53763	**1.5378**
	SR	-0.4755	-0.3091	-0.3031	-0.3091	-0.415	-0.359	-0.3615	-0.327	-0.4022	-0.3951	-0.3948	0.7608	0.9037	**0.9403**	**0.9403**
	MDD	0.3710	0.3673	0.365	0.3675	0.294	**0.2894**	0.2901	0.3673	**0.2822**	0.2894	0.2893	0.3362	0.2221	0.2221	**0.2220**
MSCI	CW	0.9010	0.9137	0.9032	0.9203	1.4072	**1.4254**	1.403	0.9246	**1.4233**	1.4154	1.4106	0.881	**1.4025**	1.3779	1.3834
	SR	-0.0878	-0.0521	-0.0601	-0.0502	0.4441	**0.4542**	0.4321	-0.0458	**0.4597**	0.4516	0.4472	-0.0233	**0.406**	0.3839	0.389
	MDD	0.6289	0.6242	0.6314	0.6279	0.3949	**0.3917**	0.3961	0.6274	**0.3893**	0.3934	0.3924	0.6612	**0.3989**	0.402	0.4009
FSE	CW	0.7316	0.7777	0.7706	0.7766	**1.0318**	1.0092	1.0069	0.8078	**1.0419**	1.0226	1.0219	1.8893	**1.9703**	1.9438	1.9434
	SR	-1.0232	-0.5613	-0.609	-0.5723	**0.0635**	-0.0335	-0.0407	-0.4799	0.093	0.0376	**0.0356**	0.7751	0.916	**0.9057**	0.9054
	MDD	0.4628	0.5007	0.4969	0.4994	**0.3631**	0.3698	0.3705	0.4994	**0.3559**	0.3703	0.3708	0.6706	0.3423	**0.339**	**0.339**
SP500	CW	1.0698	1.2357	1.2059	1.2164	**1.2711**	1.2604	1.2607	1.2227	**1.2685**	1.2638	1.2636	0.1943	**1.5815**	1.5697	1.5709
	SR	0.1503	0.3353	0.3065	0.3164	**0.4142**	0.3973	0.3979	0.3333	**0.4113**	0.4053	0.4051	0.2112	**0.5979**	0.5955	0.5958
	MDD	0.4993	0.5635	0.5607	0.5686	**0.4907**	0.4933	0.4932	0.5726	0.4966	0.4938	**0.4937**	0.9773	**0.8493**	0.852	0.8519
TSE	CW	1.5691	1.5155	1.5286	1.546	2.3443	2.3601	**2.3692**	1.5396	2.3097	2.3566	**2.3591**	1.4561	**1.9003**	1.8843	1.8714
	SR	0.5903	0.5570	0.5732	0.582	1.726	1.7293	**1.7387**	0.5663	1.6839	1.7393	**1.7427**	0.3249	**0.5828**	0.575	0.5697
	MDD	0.2985	0.3372	0.3355	0.3334	0.1552	0.1553	**0.1548**	0.3337	0.1554	0.154	**0.1539**	0.5167	**0.3819**	0.3838	0.3832

6 Conclusions

In this paper, we propose the LSTR algorithm to control the long and short term risks for the online portfolio selection. We define a capital-ratio updating equation for online portfolio selection, where the multiplication effect of the long and short term risk control parameters is found to be effective in the equation. We define a risk indicator random variable $\mathcal{C}$ to parameterize the probability of the long-term risk parameter and the control of the short-term risk parameter. Based on the definition of $\mathcal{C}$, we developed the learning and control algorithms for the risk control parameters. The experimental results on the six datasets show that the combination effect of long and short term risk control parameters is better than the existing risk control algorithms.

Our future work is to optimize for transaction costs. Besides, we will try more different methods to learn long and short term control parameters.

Acknowledgement. We thank the editors and reviewers for their expert comments and constructive suggestions, and the support of Natural Science Foundation of China No. 61972261. Professor Jianfei Yin is the corresponding author for this work.

References

1. Agarwal, A., Hazan, E., Kale, S., Schapire, R.E.: Algorithms for portfolio management based on the Newton method. In: Proceedings of the 23rd International Conference on Machine learning, pp. 9–16. ACM (2006)
2. Cover, T.M.: Universal portfolios. Math. Financ. **1**(1), 1–29 (1991)
3. Helmbold, D.P., Schapire, R.E., Singer, Y., Warmuth, M.K.: On-line portfolio selection using multiplicative updates. Math. Financ. **8**(4), 325–347 (1998)
4. Huang, D., Yu, S., Li, B., Hoi, S.C., Zhou, S.: Combination forecasting reversion strategy for online portfolio selection. ACM Trans. Intell. Syst. Technol. (TIST) **9**(5), 58 (2018)
5. Huang, X.: Portfolio selection with a new definition of risk. Eur. J. Oper. Res. **186**(1), 351–357 (2008)

6. Li, B., Hoi, S.C.: Online portfolio selection: a survey. ACM Comput. Surv. (CSUR) **46**(3), 35 (2014)
7. Li, B., Sahoo, D., Hoi, S.C.: OLPS: a toolbox for on-line portfolio selection. J. Mach. Learn. Res. **17**(1), 1242–1246 (2016)
8. Liang, Z., Chen, H., Zhu, J., Jiang, K., Li, Y.: Adversarial deep reinforcement learning in portfolio management. arXiv preprint arXiv:1808.09940 (2018)
9. Mohr, E., Dochow, R.: Risk management strategies for finding universal portfolios. Ann. Oper. Res. **256**(1), 129–147 (2017). https://doi.org/10.1007/s10479-016-2176-6
10. Shen, W., Wang, J., Jiang, Y.G., Zha, H.: Portfolio choices with orthogonal bandit learning. In: Twenty-Fourth International Joint Conference on Artificial Intelligence (2015)
11. Summers, L.H.: Mean reversion in stock prices: evidence and implications. J. Financ. Econ. **22**(1), 27–59 (1987)
12. Uziel, G., El-Yaniv, R.: Online learning of portfolio ensembles with sector exposure regularization. arXiv preprint arXiv:1604.03266 (2016)