

一种多趋势指标结合与择时引入峰值的投资组合优化系统

陈靖邦 潘俊哲 沈皓朗 谷培 扈明涛

暨南大学伯明翰大学联合学院 广州 511443

(cjb2018054910@stu2018.jnu.edu.cn)

摘要 趋势表达指标是投资组合优化领域上的一个重要话题。但是大部分基于趋势表达的投资组合优化系统仅仅考虑到了一种指标,而仅考虑到一种指标的系统在不同的数据集上的效果往往差别会比较大,因此文中使用了多趋势指标结合的系统。文中提出的投资组合优化系统使用了一系列径向基函数分别对应3种趋势表达指标(分别是简单移动平均线、指数移动均线、低延迟趋势线),并通过收盘价与短期均线价格之间的关系,对以上3种趋势进行择时,在股票出现上涨趋势的情况下加入最高价格指标(第4个指标)。在这个算法中,一系列的径向基函数会根据近期的投资情况选择最好的趋势表达指标(自适应选择),并根据以最大化下一期财富为目标的凸优化问题的解集进行投资。最后,对本系统和5种常见的投资组合优化系统在两个数据集中进行了横向对比,并取其中较为先进的两种系统在4个数据集上进行了更详细的比较,发现本系统均优于其他系统。

关键词: 投资组合优化系统;指数移动均线;低延迟趋势线;径向基函数;最高价格指标

中图法分类号 O224

Portfolio Optimization System Based on Multiple Trend Indices with Time Picking of Inducing Peak Prices

CHEN Jing-bang, PAN Jun-zhe, SHEN Hao-lang, GU Pei and HU Ming-tao

Jinan University-University of Birmingham Joint Institute, Jinan University, Guangzhou 511443, China

Abstract Trend representation index is an important topic in the field of portfolio optimization. However, most of the portfolio optimization systems based on trend representation only consider one index, and the effect of the system considering only one index is often quite different on different data sets, so we use multiple trend indices in our system. The portfolio optimization system proposed in this paper uses a series of radial basis functions corresponding to three trend representation indices (simple moving average line, exponential moving average line and low-lag trendline) respectively. This system uses the above three indices and adds the peak price index according to the relationship between the closed price and the short-term average price. In this system, the series of radial basis functions will select the best trend expression index (adaptive selection) according to the recent investment situation. Then, the system will make investment according to the solution set of the convex optimization problem which aims at maximizing the wealth of the next period. Finally, the system and five common portfolio optimization systems are compared on two data sets, two of which are chosen to be compared in more detailed on four data sets, and we conclude that our system is better than other systems.

Keywords Portfolio optimization system, Exponential moving average line, Low-lag trendline, Radial basis functions, Peak price index

1 引言

资产配置是经济、金融领域的一个基本而重要的课题,它指把财富配置到金融市场的各项资产当中,以实现某种长期的经济、金融目标。这些资产可以是股票、债券、银行存款、货

币基金、外汇等,因此投资组合优化的适用面很广。当前我国正处于经济发展的重要阶段,我国人民的收入逐步提高,对投资理财的需求也在日益增长。在这种社会背景下,开发高效的投资组合优化系统就显得尤为紧迫和必要。

近年来,随着大数据分析和计算机运算能力的迅猛发展,

基金项目:国家自然科学基金(61703182,62077028,61877029);中央高校基本科研经费(21617347,21617408,21619404,22wkzd10);广东科技计划项目(2017A040405029,2018KTSCX016,2019A050510024,2019A101002015);广州科技计划项目(201902010041);暨南大学‘国家大学生创新性实验计划’项目(202010559056)

This work was supported by the National Natural Science Foundation of China(61703182,62077028,61877029), in part by the Fundamental Research Funds for the Central Universities(21617347,21617408,21619404,22wkzd10), in part by the Science and Technology Planning Project of Guangdong(2017A040405029,2018KTSCX016,2019A050510024,2019A101002015), the part by the Science and Technology Planning Project of Guangzhou, China(201902010041) and part by the Project of ‘National University Student Innovative Experiment Program’ of Jinan University (202010559056).

通信作者:扈明涛(humingtiao2018051526@stu2018.jnu.edu.cn)

运用大数据进行分析决策的可行性大大提高。结合金融市场本身就具有大量交易数据的特点,通过机器学习算法对庞大的金融数据库进行趋势预测、波动率分析、相关性分析具有得天独厚的优势。相比之下,运用基本的金融市场原理和规律进行人工分析就显得非常局限。首先是受限于投资者本身的知识背景和主观因素的干扰;其次是计算机处理大数据的能力无论从处理速度还是误差上来说,都是人脑所无法比拟的。基于以上两点原因,在可以预见的未来,基于大数据分析和机器学习技术的投资组合优化系统具有无限的发展潜力。

由于目前大部分资产配置算法都存在只考虑了一种趋势表达的指标、运算时间长等缺陷,因此本文提出了一种多趋势表达与择时引入峰值的投资组合优化系统,运算时长短并且具有稀疏性特征。

文中提出的投资组合优化系统使用了一系列径向基函数算法分别对应3种趋势表达指标(分别是简单移动平均线、指数移动均线、低延迟趋势线),并通过收盘价与短期均线价格之间的关系,对以上3种趋势进行择时,在股票趋于上涨的情况下加入最高价格指标(第4个指标)。在这个算法中,一系列的径向基函数会根据近期的投资情况选择最好的趋势表达指标(自适应选择)。经实验证明,本系统具有投资累计收益高、算法运行时间短、交易成本低等优良性能,广泛适用于资产配置领域。

2 文献综述

2.1 资产配置与量化交易的发展现状

Markowitz 于 1952 年正式提出“资产配置”(Portfolio Selection)^[1]的概念,从此开启了数量金融领域的研究。马科维茨提出的有效前沿的理论是资产配置的开端,建立了风险与收益的数量概念。Cover 于 1991 年提出了一个全局资产组合模型^[2],建立了一个按投资期的时间顺序更新资产组合的普适性设定。这个系统为如何利用各项资产的历史数据来更新未来的资产组合提供了一套适用于机器学习算法的规范准则,是资产配置的学术研究鼻祖。但由于该理论过于理想,许多假设不符合现实,因此对其的使用渐渐变少。19 世纪 60 年代,Sharpe 提出了资本资产定价理论^[3]以及一种风险调整后的资产收益率指标(夏普比率)^[4],Fama 提出了证券市场线理论^[5],这些工作为量化资产配置研究打下坚实基础。

2.2 基于趋势表示的资产配置算法

设计资产配置算法,基本上要用到资产价格的趋势表示法。目前,利用机器学习进行资产配置的主要途径有以下3种^[6]:趋势跟踪、趋势反转、趋势模式匹配。

趋势跟踪是一种在金融市场得到广泛应用的比较成熟的技术分析思路和手段,因此设计相应的机器学习算法也较为直接。不失一般性,下文提到的趋势跟踪都指向上趋势,因为空方头寸的向下趋势相当于多方头寸的向上趋势。早期在资产配置领域的机器学习算法多属于趋势跟踪。Helmbold 等提出了一种基于指数梯度的算法^[7]。Gaivoronski 等提出一种通过随机非平稳优化来寻求全局资产组合的方法^[8]。Agarwal 等提出一种利用牛顿梯度方向更新资产组合的方法^[9]。很多机器学习模型都是通过优化问题的形式表现出来的,因此运用梯度方法求解就成为了一种直接又简单的手段。

与趋势跟踪准则相反的是趋势反转准则。根据 Fama 的有效市场假说^[10],一个成熟的金融市场将没有多少投机的空

间存在。Jegadeesh 提出了“均值反转”(Mean Reversion)准则^[11],即一项在短期内被高估或低估的资产终归回到某种平均价格水平。Borodin 等提出把投资从前期表现好的资产转移到与之具有反相关关系的当前表现差的资产上^[12]。Li 等提出了一种被动与主动相结合的均值反转策略^[13]。不久,Li 等认为除了要考虑资产收益率信息,还应考虑收益率的波动性,也就是利用资产收益率的二阶信息(即各项资产收益率的协方差矩阵)^[14],仅围绕前一交易日的资产价格进行均值反转的方法容易受剧烈价格波动的干扰而影响稳定性。受资产交易的技术分析的启发,Li 等提出了滑动平均反转(Moving Average Reversion)准则(OLMAR)^[15],认为资产价格有回归滑动平均线的趋势。Huang 等认为滑动平均依然容易受到噪声以及价格剧烈波动的干扰,并提出以 L1 中位数为参考的反转策略(RMR)^[16]。这两种趋势反转策略较有效地抓住价格趋势变化模式,取得了较为先进的资产配置效果。

除了趋势跟踪、趋势反转这两种基于人的投资思维而产生的思路外,还有一种趋势模式匹配的思路。这类算法把近期的资产价格趋势与历史上相类似的价格趋势相匹配,进而利用历史信息来推断未来的资产价格趋势。Györfi 等提出了一种非参数核方法^[17],以欧氏距离来度量近期价格与历史价格的相似度,从而找出一组相似集,再利用相似集上的收益率信息更新资产组合。随后,Györfi 等再次提出了一种基于 K 近邻算法寻找相似集的策略^[18]。Li 等提出了一种利用近期与历史两个时间窗口之间的资产相关度寻找相似集的策略^[19],从而整合了不同资产、不同时期的相关性信息。尽管趋势模式匹配方法运用机器学习算法的程度最高,但寻找相似集所耗费的计算量也相对较大。

2.3 带稀疏限制的机器学习算法

近年来,带稀疏限制的机器学习算法取得了快速的发展并获得了大量的关注。稀疏的意思是在一组待选择的样本中,只有少数几个有效样本对于学习任务起到关键作用。Wright 等提出了一个稀疏表示分类器(Sparse Representation Classification, SRC)^[20],用于人脸识别。Lai 等提出了一种判别紧致稀疏编码的分类器,用于稳健的人脸识别^[21]。Zhan 等提出了一种基于压缩感知(Compressed Sensing, CS)^[22]的快速多类字典学习方法^[23]用于机器学习领域中,同时被广泛应用在图像处理、人脸识别、医学成像、数据重构等方面,取得了先进的效果。在资产配置领域,Lai 等提出了一种基于交替方向乘子准则的短期稀疏资产组合优化系统^[24],填补了之前的研究空白。

目前投资组合优化领域的算法大多都运用了趋势表达策略,但是大部分策略都只考虑了一种指标,或是指标固定,没有动态加入指标这一做法,抑或不具有稀疏性。因此,本文提出具有稀疏性的多指标结合并择时引入其他指标的算法具有一定的创新性。

3 多趋势指标结合与择时引入峰值的投资组合

3.1 问题设置

本系统采用目前得到普遍运用且符合交易规则与标准的逐期序列性资产配置问题设置。设整个投资期共有 n 期,所研究的金融市场共有 d 项资产可供配置。本文把第 t 期的资产价格记为一个 d 维非负向量 $\mathbf{p}_t \in \mathbb{R}_+^d$ 。同时,定义价比向量 \mathbf{x}_t 如下:

$$\mathbf{x}_t = \frac{\mathbf{p}_t}{\mathbf{p}_{t-1}} \quad (1)$$

其中,两个向量的除法表示逐个元素相除。价比表示在前一期投入一单位的某项资产会在当期末有多少相对比例的产出(可能盈利也可能亏损)。相应地,一个资产组合向量表示在第 t 期初期以何种比例组合将财富分配到各个资产当中,定义资产组合向量 \mathbf{b}_t 为:

$$\mathbf{b}_t \in \Delta_d = \mathbf{b} \in \mathbb{R}_+^d : \sum_{i=1}^d b_i = 1 \quad (2)$$

这表示该金融市场不允许卖空,不能借入货币,且每期投资均用上所有财富。

记第 t 期的累积财富为 S_t ,则当期财富变化因子为 $\mathbf{b}_t^\top \mathbf{x}_t$,即:

$$S_t = S_{t-1} \cdot (\mathbf{b}_t^\top \mathbf{x}_t) \quad (3)$$

不失一般性,假设初始资产为 $S_0 = 1$,则最终累积财富 S_n 可表示为每期财富因子的累乘:

$$S_n = \prod_{t=1}^n (\mathbf{b}_t^\top \mathbf{x}_t) \quad (4)$$

因此,资产配置问题可归结如下:为了使最终累积财富最大,需要在每期期末找到一个合适的资产组合向量 \mathbf{b}_{t+1} ,使得下一期期末的累积财富尽可能地大。

3.2 4种趋势表达指标

以下是本系统用到的4种趋势表达指标,这4种趋势表达指标都是用价比的形式来表示。

3.2.1 简单移动平均线

简单移动平均线(Simple moving average line,SMA)表示最近 ω 个投资期中的平均价格,其价比的表达式为:

$$\hat{x}_{1,t+1}(\omega) = \frac{\sum_{k=0}^{\omega-1} \mathbf{p}_{t-k}}{\omega \mathbf{p}_t} \quad (5)$$

3.2.2 指数移动平均线

与简单移动平均线不同,指数移动平均线(Exponential moving average line,EMA)不设置时间窗口限制。此外,在计算平均值时,它还对不同时期的价格赋予不同的权重,认为在越靠近当前时期的价格应当赋予更大的权重。其价比计算公式如下:

$$\begin{aligned} \hat{x}_{2,t+1}(\vartheta) &= \frac{EMA_t(\vartheta)}{\mathbf{p}_t} \\ &= \frac{\vartheta \mathbf{p}_t + (1-\vartheta)EMA_{t-1}(\vartheta)}{\mathbf{p}_t} \\ &= \vartheta \mathbf{1} + (1-\vartheta) \frac{EMA_{t-1}(\vartheta)}{\mathbf{p}_{t-1}} \frac{\mathbf{p}_{t-1}}{\mathbf{p}_t} \\ &= \vartheta \mathbf{1} + (1-\vartheta) \frac{\hat{x}_{2,t}}{\mathbf{x}_t} \end{aligned} \quad (6)$$

其中, $\mathbf{1}$ 代表一个 d 维的单位向量, $0 < \vartheta < 1$ 是平滑参数, \mathbf{x}_t 则代表当前 t 时期的价比。

3.2.3 低延迟趋势线

低延迟趋势线(Low-lag Trendline,LLT)是对指数平均线EMA指标的改进。EMA指标实际上是信号处理中的一种一阶线性滤波器,EMA在指数的拐点部分通常存在较大的延迟。通过二阶滤波器对EMA信号的高频部分过滤掉,保留低频部分,从而达到降低延迟且保持良好的平滑性的目的。

Z变换是一种针对离散信号的常用的时频变换方法,将EMA指标 $EMA_t(\alpha) = \alpha * \mathbf{p}_t + (1-\alpha) * EMA_{t-1}(\alpha)$ 进行Z变换后,代入传输函数 $H(z)$, $H(z) = \frac{EMA(z)}{price(z)}$,考虑Z变换

的线性性质以及时位移性质,进一步推导得出EMA的传输函数为: $H(z) = \frac{\alpha}{1 - (1-\alpha) * z^{-1}}$ 。进行一系列实验后,可得二阶滤波器效果较好。构造出的二阶低通滤波器即:

$$H(z) = \frac{\left(\alpha - \frac{\alpha^2}{4}\right) + \left(\frac{\alpha^2}{2}\right)z^{-1} - \left(\alpha - \frac{3\alpha^2}{4}\right)z^{-2}}{1 - 2(1-\alpha) * z^{-1} + (1-\alpha)^2 * z^{-2}} \quad (7)$$

再通过Z变换性质得到最终LLT的表达式为:

$$LLT_t(\alpha) = \left(\alpha - \frac{\alpha^2}{4}\right) * \mathbf{p}_t + \left(\frac{\alpha^2}{2}\right) * \mathbf{p}_{t-1} * LLT_{t-1}(\alpha) - (1-\alpha)^2 * LLT_{t-2}(\alpha) \quad (8)$$

其价比计算公式如下:

$$\hat{x}_{3,t+1}(\alpha) = \frac{LLT_t(\alpha)}{\mathbf{p}_t} \quad (9)$$

3.2.4 最高价格

在投资时间区段的最高价格(Peak Price,PP)是指在最近 ω 个投资期中的最高价格。

$$\hat{p}_{t+1}^{(d)} = \max_{0 \leq k \leq \omega-1} \hat{p}_{t-k}^{(d)}, i=1,2,\dots,d \quad (10)$$

相应的价比 \hat{x}_{t+1} 预测表示不同资产的增长潜力^[25]:

$$\hat{x}_{pp,t+1} = \frac{\hat{p}_{t+1}}{\mathbf{p}_t} \quad (11)$$

3.3 自适应选择复合趋势表达算法

3.3.1 3种趋势表达预测价比与PP进行复合

本算法在后续输入核函数更新资产配置用到的是前3种趋势表达方法分别与PP指标进行复合后的指标,结合方法如下:

如果第 t 天收盘价 $p_t \geq SMA_t$,也就是第 t 天的收盘价大于或等于第 t 天的均线价格,算法判断股票上涨的几率更大,此时原来3个指标的预测值 $\hat{x}_{n,t+1}$ 将同时加入PP指标;如果第 t 天收盘价 $p_t < SMA_t$,原来3个指标预测值将保持不变,把产生的新指标记作 $\hat{x}_{t,t+1}$ 。

在这里把第 t 天收盘价大于第 t 天的均线价格记为 $\gamma_t = 1$,否则 $\gamma_t = 2$,即:

$$\gamma_t = \begin{cases} 1, & p_t > SMA_t \\ 2, & p_t \leq SMA_t \end{cases} \quad (12)$$

$$\hat{x}_{l,t+1} = \begin{cases} \frac{1}{2} * (\hat{x}_{n,t+1} + \hat{x}_{pp,t+1}), & \gamma_t = 1 \\ \hat{x}_{n,t+1}, & \gamma_t = 2 \end{cases}, n \in 1,2,3 \quad (13)$$

3.3.2 计算径向基函数

典型的径向基函数形式如下:

$$y = \mathbf{A}(x), \boldsymbol{\phi} = [\phi_1, \phi_2, \dots, \phi_L]^\top \quad (14)$$

$$\phi_l(x) = \exp\left(-\frac{\|x - \mu_l\|^2}{2\sigma_l^2}\right) \quad (15)$$

其中, x 和 y 对应函数的输入与输出; \mathbf{A} 为函数的权重矩阵; d 则是函数输出的维度; $\boldsymbol{\phi}$ 表示高斯基函数的向量; μ_l 和 σ_l 对应中心和尺度参数。

首先将各趋势指标 $\{\hat{x}_{l,t+1}\}_{l=1}^L$ 投影到 d 维的单纯形上,其中 L 表示趋势的数量:

$$\tilde{x}_{l,t+1} = \operatorname{argmin}_{x \in \Delta_d} \|x - \hat{x}_{l,t+1}\|^2, l=1,\dots,L \quad (16)$$

接着,利用 $\{\tilde{x}_{l,t-k}\}_{k=0}^{\omega-1}$ 在最近的时间窗口来回测最近 l 趋势的投资表现:

$$R_{l,t-k} = \tilde{x}_{l,t+1}^\top x_{t-k}, k=0,\dots,\omega-1 \quad (17)$$

其中, $R_{l,t-k}$ 表示 l 趋势在最近 $(t-k)$ 周期的增长因子。

然后可以计算出3.3.1节中各趋势的增长因子,通过比

较其中各趋势的近期投资绩效,使用下方自适应选择出表现最好的趋势 $\tilde{x}_{*,t+1}$, * 作为径向基函数的自适应输入:

$$\tilde{x}_{*,t+1}, * \triangleq \underset{1 \leq l \leq 3}{\operatorname{argmax}} \min_{0 \leq k \leq \omega-1} R_{l,t-k} \quad (18)$$

其方法是,首先在各趋势组合的时间窗口中寻找最小的增长因子,然后从中挑选出增长因子最大的趋势。通过这种方法,可以求得在最恶劣的金融环境下表现最好的趋势表示,提高了整个系统的稳定性。

本算法利用所有趋势的投资组合 $\{\tilde{x}_{l,t+1}\}$ 作为径向基函数的中心,其中表现最好的 $\tilde{x}_{*,t+1}$ 则作为函数的输入。最终得到如下径向基函数:

$$\Delta b_{t+1} = \hat{\mathbf{X}}_{t+1} \boldsymbol{\phi}(\tilde{x}_{*,t+1}) \quad (19)$$

$$\boldsymbol{\phi} = [\phi_1, \phi_2, \dots, \phi_L]^T \quad (20)$$

$$\phi_l(\tilde{x}_{*,t+1}) = \exp\left(-\frac{\|\tilde{x}_{*,t+1} - \tilde{x}_{l,t+1}\|^2}{2\sigma_l^2}\right) \quad (21)$$

其中, Δb_{t+1} 是第 $(t+1)$ 期的投资组合增量, $\hat{\mathbf{X}}_{t+1}$ 则作为其权重矩阵。

3.3.3 更新资产配置

本算法将结合径向基函数(Radial basis functions, RBF)来建立一个资产选择模型以获得更好的投资表现。根据前文描述,需要基于 $t+1$ 期以前的价格信息最大化增量 $b_{t+1}^T x_{t+1}$ 。尽管 x_{t+1} 在现阶段未知,但可以用 $\{\hat{x}_{l,t+1}\}$ 来替代它。此外,所提出的RBF也可以组合成一个广义的增长因子,具体如下:

$$b_{t+1} = \underset{b}{\operatorname{argmax}} \operatorname{tr}(\mathbf{B}\boldsymbol{\Phi}\hat{\mathbf{X}}_{t+1}^T), \mathbf{B} = \mathbf{b}\mathbf{1}_{(L)}^T \quad (22)$$

$$\boldsymbol{\Phi} = \operatorname{diag}(\boldsymbol{\phi}), \hat{\mathbf{X}}_{t+1} = [\hat{x}_{1,t+1}, \dots, \hat{x}_{L,t+1}] \quad (23)$$

$$\text{s. t. } \mathbf{b} \in \Delta_d, \|\mathbf{b} - \hat{\mathbf{b}}_t\| \leq \epsilon, \epsilon > 0 \quad (24)$$

其中, tr 是迹算子, \mathbf{B} 包含 \mathbf{b} 的 L 个副本, $\boldsymbol{\Phi}$ 是由 $\boldsymbol{\phi}$ 构成的对角矩阵,并且 $\hat{\mathbf{X}}_{t+1}$ 是 L 个不同的趋势表示。 $\mathbf{1}_{(L)}$ 是 L 维元素为1的向量。 \mathbf{b} 与 $\hat{\mathbf{b}}_t$ 相距 ϵ 。

与PPT^[25]中原始增长因子 $b_{t+1}^T \hat{x}_{t+1}$ 相比, $\operatorname{tr}(\mathbf{B}\boldsymbol{\Phi}\hat{\mathbf{X}}_{t+1}^T)$ 是以 $\boldsymbol{\Phi}$ 作为一个内核的广义增长因子,且可以调整不同趋势表示。表现最好的趋势代表 $\hat{\mathbf{X}}_{*,t+1}$ 影响最大,其他则影响较小,这取决于它们与 $\hat{\mathbf{X}}_{*,t+1}$ 的相似性。

根据梯度投影原理, \mathbf{b} 的单纯形约束可以放宽到 $\mathbf{1}^T \mathbf{b} = 1$,先求 $\hat{\mathbf{b}}_t$,然后投影到单纯形。另外,设 $\mathbf{c} = \mathbf{b} - \hat{\mathbf{b}}_t$,则上式可以进一步简化为:

$$c_{t+1} = \underset{c}{\operatorname{argmax}} (\mathbf{C}\boldsymbol{\Phi}\hat{\mathbf{X}}_{t+1}^T) \quad (25)$$

$$\text{s. t. } \mathbf{1}^T \mathbf{c} = 0, \|\mathbf{c}\| \leq \epsilon \quad (26)$$

$$\mathbf{C} = \mathbf{c}\mathbf{1}_{(L)}^T \cdot \hat{\mathbf{b}}_t + c_{t+1} \quad (27)$$

其原因是 $\hat{\mathbf{b}}_t$ 固定和 $\mathbf{1}^T \mathbf{c} = \mathbf{1}^T \mathbf{b} - \mathbf{1}^T \hat{\mathbf{b}}_t = 0$ 。现在,优化目标已变为更新增量 c_{t+1} 。

接下来将进行模型求解算法展示。求解方法可以总结为以下定理。

定理 1 如果 $(\mathbf{I} - \frac{1}{d}\mathbf{1}\mathbf{1}^T)\hat{\mathbf{X}}_{t+1}\boldsymbol{\Phi}\mathbf{1}_{(L)} = 0$,则 $c_{t+1} = 0$;否则上述公式有唯一解:

$$c_{t+1} = \frac{\epsilon \left(\mathbf{I} - \frac{1}{d}\mathbf{1}\mathbf{1}^T \right) \hat{\mathbf{X}}_{t+1} \boldsymbol{\Phi} \mathbf{1}_{(L)}}{\left\| \left(\mathbf{I} - \frac{1}{d}\mathbf{1}\mathbf{1}^T \right) \hat{\mathbf{X}}_{t+1} \boldsymbol{\Phi} \mathbf{1}_{(L)} \right\|} \quad (28)$$

定理的证明在相关文献[26]中查看。

通过计算出 c_{t+1} 的值,最终可以求得 $\hat{\mathbf{b}}_{t+1}$:

$$\hat{\mathbf{b}}_{t+1} = \hat{\mathbf{b}}_t + c_{t+1} \quad (29)$$

$$\hat{\mathbf{b}}_{t+1} = \underset{\mathbf{b} \in \Delta_d}{\operatorname{argmin}} \|\mathbf{b} - \hat{\mathbf{b}}_{t+1}\|^2 \quad (30)$$

算法流程图如图1所示。

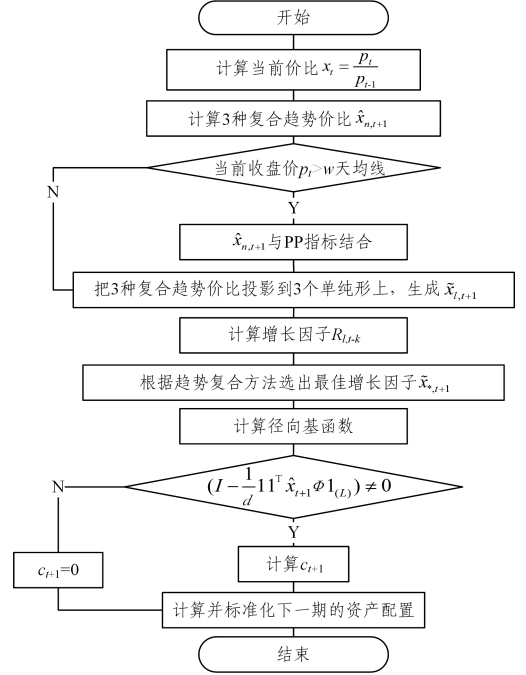


图1 MTTP算法流程图

Fig. 1 Flowchart of MTTP algorithm.

3.4 实验及结果分析

3.4.1 累计财富

如图2和图3所示,纵轴表示累积财富,横轴表示时间周期。本实验用MTTP和其他5种常见的资产配置算法进行了横向对比,MTTP是本文提出的新算法,其他5种分别是AICTR^[26], OLMAR^[15], PPT^[25], ANTICOR^[12]和UBAH(uniform buy-and-hold)。图2是算法在2010年8月18日到2015年5月15日的沪深300指数上运行的结果,图3是算法在2013年6月13日到2017年8月2日摩根士丹利资本国际公司(Morgan Stanley Capital International)所编制的证券指数(MSCI)上运行的结果,可以看出MTTP都是6种算法里表现最好的。

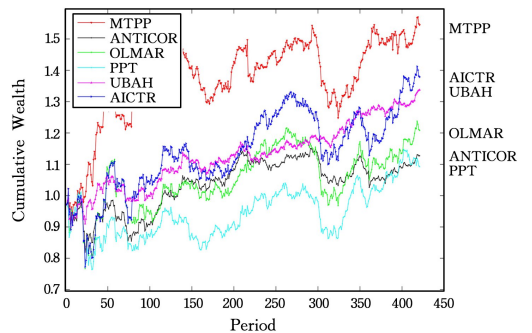


图2 6个投资组合系统在沪深300指数上的累计财富
Fig. 2 Cumulative wealth of six portfolio selection systems on HS300

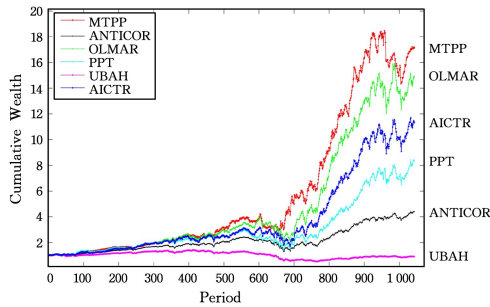


图3 6个投资组合系统在摩根士丹利证券指数上的累计财富

Fig.3 Cumulative wealth of six portfolio selection systems on MSCI

本实验选出了其他算法中较为有效的两个算法 AICTR 和 OLMAR,再与 MTTP 进行其他指标的比较。

本实验使用了4个不同的数据集,分别是 HS300(沪深300指数)、NYSE19(纽约证券交易所数据)、MSCI(摩根士丹利资本国际公司所编制的证券指数)和 SSEC(上证指数),它们是不同的股票(基金)指数的成分股(基金)的数据。累计财富数据如表1所列。

表1 投资组合系统在4个数据集上的累计财富

Table 1 Cumulative Wealth of portfolio selection systems on four data sets

Algorithm	HS300	NYSE19	MSCI	SSEC
OLMAR	1.2083	0.9892	14.9353	0.3761
AICTR	1.3780	1.2669	11.3176	0.3021
MTTP	1.5450	1.8832	17.1310	1.9064

从3个算法在4个不同数据集上的累计财富可以看出,本算法在各个数据集上的累计财富都处于领先地位。

3.4.2 夏普比率

夏普比率是一种风险调整后的收益率指标,其计算公式如下:

$$SR = \frac{\bar{r}_s - r_f}{\hat{\sigma}(r_s)} \quad (31)$$

其中, \bar{r}_s 是资产组合优化系统的平均单期收益率, r_f 是无风险利率, $\hat{\sigma}(r_s)$ 是资产组合优化系统的单期收益率的标准差。夏普比率表示投资组合每承担一单位总风险,会产生多少的超额收益。这个比例衡量了基金的操作风险与回报的关系,通常用于量化投资界比较策略的好坏,夏普比率越高,则系统的风险控制能力越强。从表2可以看出,本算法在各个数据集上的夏普比率都高于其他算法。

表2 投资组合系统在4个数据集上的夏普比率

Table 2 Sharpe ratio of portfolio selection systems on four data sets

Algorithm	HS300	NYSE19	MSCI	SSEC
OLMAR	0.0371	0.0181	0.1164	-0.0266
AICTR	0.0527	0.0260	0.1039	-0.0344
MTTP	0.0788	0.0382	0.1340	0.0440

3.5 SSEC(上证指数)不同时间的投资组合

该指数共30个资产,本实验选取了该策略在44,53,247,263,293,308这几个天数上投资组合的情况,如图4所示。从图中可以看出,本算法在每一天均对资产进行了投资,且每一天产生的资产组合是稀疏的,投资组合集中在少数几个资产上。

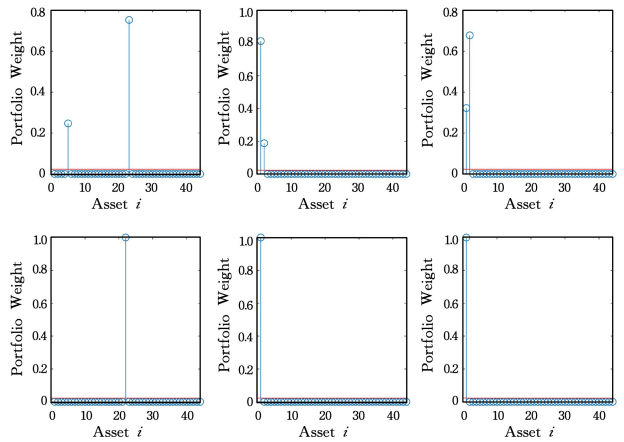


图4 不同时间投资组合的情况

Fig.4 Portfolios in different periods

3.6 累计财富与交易费用的关系

本实验还测量了MTTP与OLMAR和AICTR在承受一定交易费损耗的条件下所取得的最终累计财富,如图5所示。当交易费率从0增至0.05%时,本算法在所有数据集上的累计财富均大幅领先于OLMAR和AICTR,这表明本算法能承受一定的交易费损耗,适用于真实世界的金融环境。

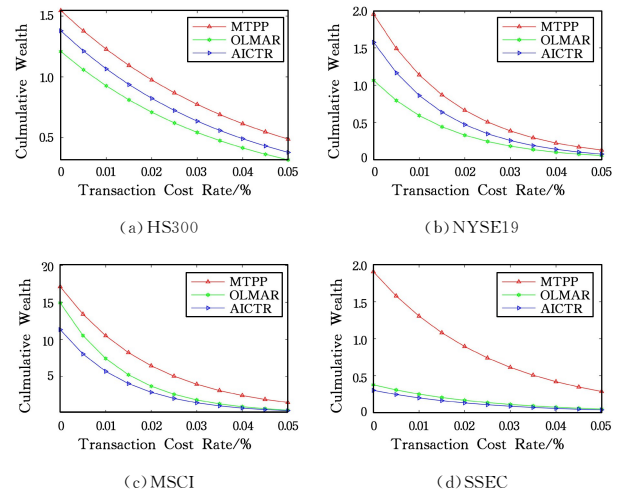


图5 投资组合系统在4个数据集上的累计财富-交易成本图
Fig.5 Cumulative wealth of portfolio selection systems with respect to transaction cost on four data sets

3.7 运行效率

本算法使用具有 Intel Core i7-8750H 的中央处理器和 16GB 内存的计算机完成。在 HS300, NYSE19, MSCI, SSEC 4 个数据集上的运行时间分别是 0.029 853 s, 0.085 580 s, 0.059 294 s, 0.057 600 s, 可以看出,本算法运行迅速,适用于真实的金融市场。

结束语 本文提出的资产配置系统使用了一系列径向基函数算法分别对应3种趋势表达指标(简单移动平均线、指数移动均线、低延迟趋势线)并通过收盘价与短期均线价格之间的关系,对以上3种趋势进行择时,在股票趋于上涨的情况下加入最高价格指标(第4个指标)。未来,本系统将在更新资产配置的部分引入深度学习框架,从而更好地结合3种趋势,并反映到投资组合上;另外争取提出更能反映股票趋势特征的新指标。

参 考 文 献

- [1] MARKOWITZ M. Portfolio selection[J]. The Journal of Finance, 1952, 7(1): 77-91.
- [2] COVER T M. Universal portfolios[J]. Mathematical Finance, 1991, 1(1): 1-29.
- [3] SHARPE W F. Capital asset prices: A theory of market equilibrium under conditions of risk[J]. The Journal of Finance, 1964, 19(3): 425-442.
- [4] SHARPE W F. Mutual fund performance[J]. The Journal of Business, 1966, 39(1): 119-138.
- [5] FAMA E F. The behavior of stock-market prices[J]. The Journal of Business, 1965, 38(1): 34-105.
- [6] LI B, HOI S C H. Online portfolio selection: A survey[J]. ACM Computing Surveys (CSUR), 2014, 46(3): 35:1-35:36.
- [7] HELMBOLD D P, SCHAPIRE R E, SINGER Y, et al. On-line portfolio selection using multiplicative updates[J]. Mathematical Finance, 1998, 8(4): 325-347.
- [8] GAIVORONSKI A, STELLA F. Stochastic nonstationary optimization for finding universal portfolios[J]. Annals of Operations Research, 2000, 100: 165-188.
- [9] AGARWAL A, HAZAN E, KALE S, et al. Algorithms for portfolio management based on the newton method[C]// International Conference on Machine Learning. International Machine Learning Society, 2006: 9-16.
- [10] FAMA E F. Efficient capital markets: A review of theory and empirical work[J]. The Journal of Finance, 1970, 25(2): 383-417.
- [11] JEGADEESH N. Seasonality in stock price mean reversion: Evidence from the U. S. and the U. K.[J]. The Journal of Finance, 1991, 46(4): 1427-1444.
- [12] BORODIN A, EL-YANIV R, GOGAN V. Can we learn to beat the best stock[J]. Journal of Artificial Intelligence Research, 2004, 21(1): 579-594.
- [13] LI B, ZHAO P, HOI S, et al. PAMR: Passive aggressive mean reversion strategy for portfolio selection[J]. Machine Learning, 2012, 87(2): 221-258.
- [14] LI B, HOI S C H, ZHAO P, et al. Confidence weighted mean reversion strategy for online portfolio selection[J]. ACM Transactions on Knowledge Discovery from Data, 2013, 7(1): 4.
- [15] LI B, HOI S C H, SAHOO D, et al. Moving average reversion strategy for on-line portfolio selection[J]. Artificial Intelligence, 2015, 222: 104-123.
- [16] HUANG D, ZHOU J, LI B, et al. Robust median reversion strategy for on-line portfolio selection[J]. IEEE Transactions on Knowledge and Data Engineering, 2016, 28(9): 2480-2493.
- [17] GYÖRFI L, LUGOSI G, UDINA F. Nonparametric kernel-based sequential investment strategies[J]. Mathematical Finance, 2006, 16(2): 337-357.
- [18] GYÖRFI L, UDINA F, WALK H. Nonparametric nearest neighbor based empirical portfolio selection strategies[J]. Statistics and Decisions, 2008, 26(2): 145-157.
- [19] LI B, HOI S C H, GOPALKRISHNAN V. Corn: Correlation-driven nonparametric learning approach for portfolio selection[J]. ACM Transactions on Intelligent Systems and Technology, 2011, 2(3): article No. 21.
- [20] WRIGHT J, YANG A Y, GANESH A, et al. Robust face recognition via sparse representation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(2): 210-227.
- [21] LAI Z R, DAI D Q, REN C X, et al. Discriminative and compact coding for robust face recognition[J]. IEEE Transactions on Cybernetics, 2015, 45(9): 1900-1912.
- [22] DONOHO D L. Compressed sensing[J]. IEEE Transactions on Information Theory, 2006, 52(4): 1289-1306.
- [23] ZHAN Z, CAI J F, GUO D, et al. Fast multiclass dictionaries learning with geometrical directions in MRI reconstruction[J]. IEEE Transactions on Biomedical Engineering, 2016, 63(9): 1850-1861.
- [24] LAI Z R, YANG P Y, FANG L D, et al. Short-term sparse portfolio optimization based on alternating direction method of multipliers[J]. Journal of Machine Learning Research, 2018, 19.
- [25] LAI Z R, DAI D Q, REN C X, et al. A Peak Price Tracking-Based Learning System for Portfolio Selection[J]. IEEE Transactions on Neural Networks and Learning Systems, 2018, 29(7): 2823-2832.
- [26] LAI Z R, DAI D Q, REN C X, et al. Radial Basis Functions with Adaptive Input and Composite Trend Representation for Portfolio Selection[J]. IEEE Transactions on Neural Networks and Learning Systems, 2018, 29(12): 6214-6226.



CHEN Jing-bang, born in 2000, postgraduate. His main research interests include portfolio optimization and machine learning.



HU Ming-tao, born in 2000, postgraduate. His main research interests include portfolio optimization and machine learning.