

디리클레 분포를 이용한 트레이딩 알고리즘의 앙상블에 관한 연구

정재용, 이주홍*, 최범기**, 송재원***

인하대학교, *인하대학교, **㈜큐헷지, ***㈜밸류파인더스

eoehd4108@naver.com, *juhong@inha.ac.kr, **bgchoi666@gmail.com, ***jwsong@valuefinders.co.kr

A study on ensemble of trading algorithms using Dirichlet Distribution

Jae Yong Jeong, Ju-Hong* Lee, Bum Gi Choi**, Jae Wong Song***

Inha Univ., *Inha Univ., **QHedge Co. Ltd., ***ValueFinders Co. Ltd.

요약

금융시장에서 알고리즘 트레이딩은 중요한 역할을 한다. 시장의 변동성으로 인해서 알고리즘 트레이딩은 일관적인 결과가 나올 수 없고, 알고리즘마다 우수한 성능이 나오는 부분이 다르게 존재한다. 본 논문에서는 트레이딩 알고리즘의 결과에 대한 우수한 조합을 만들기 위한 앙상블 방법을 제안한다. 알고리즘 트레이딩에 적용하기 위한 다양한 앙상블 기법들이 존재하지만, 일부 앙상블 기법은 적절하지 않을 수 있다. 따라서 이론적으로 최적의 가중치 조합을 찾을 수 있는 BMA 기법을 변형한 앙상블 기법을 제안한다. 본 논문에서는 각 알고리즘의 가중치를 디리클레 분포로 가정하고, 분포로부터 샘플링을 수행하여 Transformer 모델로 알고리즘의 가중치를 학습한다. 실험 결과를 통해, 제안한 앙상블 기법이 알고리즘 트레이딩의 일반화 성능을 높였음을 보였다.

I. 서론

알고리즘 트레이딩이란 수학적 계산과 IT 시스템을 이용해 금융상품을 거래하는 것을 말한다. 2012년 미국에서 전체 거래의 85%를 차지할 정도로 알고리즘 트레이딩은 금융시장에서 중요한 역할을 한다[1]. 하지만, 시장의 변동성으로 인해서 알고리즘 트레이딩은 일관적인 결과가 나올 수 없고[2], 알고리즘마다 우수한 성능이 나오는 부분이 다르게 존재한다[3]. 따라서 알고리즘 트레이딩에 앙상블 기법을 사용하여 모델의 일반화 성능을 높이기 위한 다양한 시도들이 존재한다[2, 3, 4, 5].

[2]는 모델의 출력에 대한 동등한 가중치를 부여하여 조합한다. [3]는 가장 적절한 모델 1개를 선택하는 방식을 사용한다. [4]는 AND 연산자를 이용하여 불확실성에 대응하고자 한다. [5]는 앙상블 기법 중 대표적인 Bagging 기법을 적용한다. 이외에도 적용 가능한 앙상블 기법들을 [6, 7]에서 소개한다. 알고리즘 트레이딩에 적용하기 위한 다양한 앙상블 기법이 존재하지만, [8]는 기존에 알려진 앙상블 기법들이 주식 시장에 적용하기에 적절하지 않을 수 있음을 설명한다.

따라서 본 논문에서는 [9]에서 이론적으로 최적의 조합을 찾을 수 있다고 알려진 BMA(Bayesian Model Averaging)를 변형한 앙상블 기법을 제시한다. 기존의 BMA를 알고리즘 트레이딩에 적용하기 위해 2가지 접근을 사용한다. 첫째, 최근 정보에 집중하기 위해, 감소계수를 추가한다. 앙상블에서 적용될 모델들의 가중치가 시간에 따른 정보를 반영하기 위해서, 모델의 가중치를 디리클레분포로 가정하고, 훈련 구간에서 실제 모델의 성능을 기준으로 그 분포를 수정한다. 이때, 감소계수를 적용하여 가중치가 특정 모델에 지나치게 치우치는 것을 방지하고, 최근 정보에 가중치를 준다. 둘째, 훈련된 분포로부터 모델의 가중치를 샘플링하여 Transformer 모델을 훈련하여 다음에 나타나게 될 분포를 예측한다. 따라서, 특정 시점에 대한 과적합된 분포를 그대로 사용하지 않을 수 있다. 제안한 앙상블 방법을 암호 화폐를 대상으로 수행하였다. 그 결과로서 앙상블 방법은 암호화폐의 가격 흐름보다 변동성이 매우 적었으며 수익률도 훨씬 높게 나타나서 성능이 매우 우수함을 볼 수 있었다. 따라서 제안한 앙상블 방법의 일반화 성능이 우수함을 볼 수 있다.

이후 사용되는 표기법은 다음과 같다.

K	알고리즘 트레이딩의 개수
w_i	i 번째 알고리즘 트레이딩의 가중치
$(0 \leq w_i \leq 1, \sum_{i=1}^K w_i = 1)$	
α_i	디리클레 분포의 i 번째 파라미터
D_t^i	i 번째 알고리즘 트레이딩의 t 시점 결정 ($M_t^i \in \{-1, 1\}$, -1 매도, 1 : 매수)
r_t^i	i 번째 알고리즘 트레이딩 t 시점 수익률

II. 본론

본 논문에서 제안하는 앙상블 방법을 다음 과정을 통해서 진행한다. 앙상블을 수행하기 위한 각 알고리즘 트레이딩의 가중치가 디리클레 분포를 따른다고 가정한다.

$$W \sim \text{Dir}(\alpha_1, \alpha_2, \dots, \alpha_K) \quad (1)$$

이때, 초기 $\alpha_i = 1$ ($i = 1, 2, \dots, K$)로 초기값을 설정한다.

알고리즘 트레이딩의 앙상블 $D_t^{i,j,k}$ 를 D_t^i, D_t^j, D_t^k 의 voting으로 얻은 결정이라고 하자. K 개의 알고리즘 트레이딩의 n 개의 앙상블은 $K C_n$ 개의 조합이 나오게 된다.

$K C_n$ 개의 앙상블 조합으로부터 상위 m 개의 알고리즘 트레이딩을 찾는다. 이때, 알고리즘 트레이딩의 성능은 샤프지수(수식 2.)를 기준으로 평가한다.

$$\text{Sharpe ratio} = \frac{\text{mean}(r)}{\text{stddev}(r)} \quad (2)$$

상위 m 개의 i 번째 알고리즘 트레이딩의 성능이 상위권이었던 횟수를 α'_i 라고 하자. 가중치(W)를 알고리즘 트레이딩이 상위 m 개의 모델에 포함되어 있을 확률의 의미로 해석하면, 다항분포의 모양으로 모델링할 수 있다.

$$P(\alpha' | W) = \frac{m!}{\alpha'_1! \alpha'_2! \dots \alpha'_K!} w_1^{\alpha'_1} w_2^{\alpha'_2} \dots w_K^{\alpha'_K} \quad (3)$$

수식 (1)은 (3)의 conjugate prior이기 때문에 베이저안 정리를 이용해

$$W \sim Dir(\alpha_1 + \alpha'_1, \alpha_2 + \alpha'_2, \dots, \alpha_K + \alpha'_K) \quad (4)$$

로 업데이트가 가능하다. 이때, 특정 모델에 가중치가 지나치게 집중되는 것을 막고 최근 정보에 대한 가중치를 주기 위하여 감소계수(c)를 적용한 업데이트를 한다.

$$W \sim Dir(\lfloor \alpha_1/c \rfloor + \alpha'_1, \lfloor \alpha_2/c \rfloor + \alpha'_2, \dots, \lfloor \alpha_K/c \rfloor + \alpha'_K) \quad (5)$$

따라서 특정 시점(t)의 가중치 분포 W_t 를 찾아낼 수 있다. 하지만, 알고리즘 트레이딩의 앙상블을 수행하기 위해서는 미래 시점(t')의 가중치 분포 $W_{t'}$ 를 알아야 하는데, 이는 알 수 없는 분포이다. 이를 보완하기 위해서 [10]에서 제안한 Transformer 모델을 사용하여 미래 시점의 분포를 예측한다. Transformer 모델을 훈련하기 위해서 과거 시점들의 데이터 분포로부터 H번의 랜덤 샘플링을 아래와 같이 수행하고,

$$x_{(p)}^{(t)} \sim Dir(W_t; \alpha) \quad (p = 1, 2, \dots, H)$$

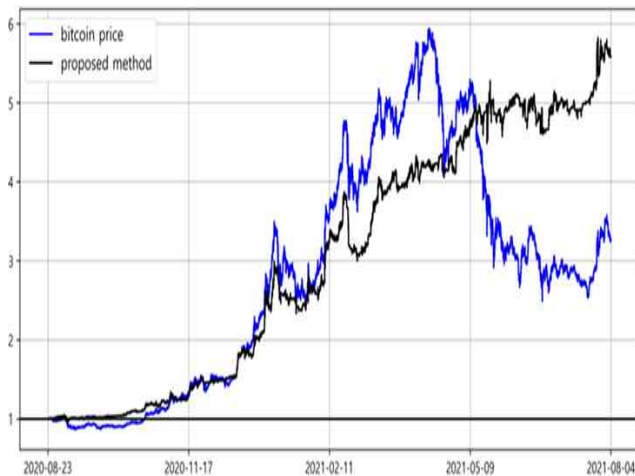
훈련을 위한 데이터 셋을 아래와 같이 정의한다.

$$\begin{aligned} X_1 &: (x_{(1)}^{(t-5)}, x_{(1)}^{(t-4)}, x_{(1)}^{(t-3)}, x_{(1)}^{(t-2)}, x_{(1)}^{(t-1)}) \\ Y_1 &: x_{(1)}^{(t)} \\ \text{input data} &= (X_1, X_2, \dots, X_H) \\ \text{output data} &= (Y_1, Y_2, \dots, Y_H) \end{aligned}$$

이때, 목적함수는 크로스-엔트로피를 사용하여서 두 분포 사이의 차이를 줄여준다. 따라서, 미래 시점(t')의 데이터 분포는 이전 시점의 샘플링 $(x^{(t'-5)}, x^{(t'-4)}, x^{(t'-3)}, x^{(t'-2)}, x^{(t'-1)})$ 으로부터 추론 할 수 있다. 이때 추론된 값들을 $\widehat{x}_p^{(t')} \quad (p = 1, 2, \dots, H)$ 로 표현이 가능하다. 추론된 값들의 평균($\widehat{x}$)을 미래 시점(t')의 알고리즘 트레이딩의 가중치로 사용할 수 있다. 즉, 아래와 같이 알고리즘 트레이딩의 앙상블을 수행한다.

$$D_{t'}^{ensemble} = \widehat{x}_i D_{t'}^i$$

위에서 제안한 앙상블 방법을 암호 화폐(비트코인)으로 실험한 결과는 (그림 1.)과 같이 나타난다. 실험 데이터로서 2020년 8월 23일부터 2021년 8월 4일까지의 비트코인의 30분 단위 시계열을 사용하였다. x축은 시간을 의미하고, y축은 시간에 따른 수익률을 의미한다. 파란색으로 표시된 비트코인의 경우의 변동성이 제안된 앙상블 방법의 변동성보다 훨씬 큰 것을 알 수 있고, 제안된 앙상블 방법의 최종 수익률이 더 높음을 볼 수 있다.



(그림 1.) 제안한 앙상블 기법을 적용한 결과

III. 결론

본 논문에서는 시계열 데이터의 특성을 고려한 알고리즘 트레이딩의 앙상블 기법을 제안하였다. 각 알고리즘 트레이딩의 가중치를 디리클레 분포로 가정하고, 일정 시간 동안의 알고리즘 트레이딩 성능과 감소계수를 적용하여 분포를 업데이트한다. 또한 관측할 수 없는 미래 시점의 분포를 예측하기 위해서, 과거 시점에서의 디리클레 분포에서 데이터를 샘플링하여 Transformer 모델로 학습한다. 실제 트레이딩의 결과(그림1.)로서 제안된 앙상블의 성능이 모든 구간에서 꾸준히 유지되는 것을 볼 수 있다. 이를 통해 알고리즘 트레이딩의 일반화 성능이 우수함을 볼 수 있다.

ACKNOWLEDGMENT

이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (2019R1F1A1062094, 2017R1D1A1A02018319, NRF-2020R1F1A1069361)

참 고 문 헌

- [1] 안명호, 류미현 (2018) 머신러닝을 이용한 알고리즘 트레이딩 개발. 한빛미디어
- [2] Ullah, A. K. M., et al. "Combining machine learning classifiers for stock trading with effective feature extraction." arXiv preprint arXiv:2107.13148 (2021).
- [3] Yang, Hongyang, et al. "Deep reinforcement learning for automated stock trading: An ensemble strategy." Available at SSRN (2020).
- [4] Giacometti, Felipe, Renata Galante, and Adriano Pereira. "An algorithmic trading agent based on a neural network ensemble: a case of study in North American and Brazilian stock markets." 2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT). Vol. 2. IEEE, 2015.
- [5] Sun, Shaolong, Shouyang Wang, and Yunjie Wei. "A new ensemble deep learning approach for exchange rates forecasting and trading." Advanced Engineering Informatics 46 (2020): 101160.
- [6] Sagi, Omer, and Lior Rokach. "Ensemble learning: A survey." Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 8.4 (2018): e1249.
- [7] Huang, Faliang, Guoqing Xie, and Ruliang Xiao. "Research on ensemble learning." 2009 International Conference on Artificial Intelligence and Computational Intelligence. Vol. 3. IEEE, 2009.
- [8] Jacobsen, Ben, Fuwei Jiang, and Hongwei Zhang. "Ensemble machine learning and stock return predictability." The 4th International Workshop on Financial Markets and Nonlinear Dynamics (Paris). 2019.
- [9] Domingos, Pedro. "Bayesian averaging of classifiers and the overfitting problem." ICML. Vol. 747. 2000.
- [10] Vaswani, Ashish, et al. "Attention is all you need." Advances in neural information processing systems. 2017.