

유한 크기 효과를 고려한 비선형 의존성 지표를 활용한 계층적 리스크 패리티 모형 기반 포트폴리오 최적화

최인수¹, 김우창^{2*}

¹KAIST 산업및시스템공학과

²KAIST 산업및시스템공학과

jl.cheivly_@kaist.ac.kr, wkim_@kaist.ac.kr

Hierarchical Risk Parity Portfolio Optimization via Nonlinear Measures Considering Finite Size Effects

Insu Choi¹, Woo Chang Kim

¹Dept. of Industrial and Systems Engineering, KAIST

²Dept. of Industrial and Systems Engineering, KAIST

요 약

본 연구는 계층적 리스크 패리티 (Hierarchical Risk Parity, HRP) 포트폴리오 방법론과 정규화된 상호 정보 거리의 결합을 연구하였다. 이때, 한정된 이동창에서 발생할 수 있는 유한 크기 효과(finite size effects) 문제를 극복하기 위해 무작위로 섞인 NID 값에 대한 평균치를 제공함에 따라 NID를 활용한 새로운 포트폴리오 최적화 방법을 제안한다. 본 연구의 결과는 NID를 통합한 HRP 포트폴리오가 기존 방법론에 비해 통계적 장점과 함께 더욱 효율적이며 안정적인 성능을 보여준다.

1. 서론

포트폴리오 최적화의 중심에는 자산들 간의 상관 관계의 이해가 있다. 상관계수는 전통적으로 이러한 관계를 측정하는 데 널리 사용되는 지표다. 그러나, 상관계수는 주로 선형 상관관계에만 이용된다. 반면 상호 정보 거리의 경우 선형 관계를 포함하여 자산들 간의 서로 공유되는 정보의 양을 정량화하는 방법을 제공한다.

본 연구는 Hierarchical Risk Parity (HRP) 포트폴리오 방법과 정규화된 상호 정보 거리를 결합한 새로운 접근법을 제안한다. HRP 포트폴리오 방법은 자산들 간의 상관 관계를 기반으로 하여, 자산 분배를 결정하는 방법론이다. 이 연구의 주된 목적은 NID를 활용하여 HRP 포트폴리오의 효율성과 안정성을 더욱 높이는 것이다. 또한, NID의 계산을 위해 특정 상황에서 발생할 수 있는 유한 크기 효과 문제를 제거하기 위하여 무작위로 시계열을 섞은 뒤, 이러한 값을 제거하는 방법을 활용하여 해당 크기를 제거한 방법을 활용한다 (Dimpfl and Peter, 2013).

효율적이고 안정적인 포트폴리오 구축은 투자자들에게 중요한 이익을 제공하며, 기존의 상관계수를 기반으로 한 방법론에 더하여, 다른 통계적 의존성을 보는 것은 의미가 있으며, 이러한 점에서 선형 정보에 대한 상관관계를 포함하는 NID를 사용할 경우 더욱 정확하고 안정적인 결과를 얻는 데에 기여할 수 있을 것으로 판단된다.

2. 연구 방법론

HRP(Hierarchical Risk Parity)는 전통적인 포트폴리오 최적화 방법과는 다르게 자산들 간의 상관 관계를 중심으로 한 방법이다. HRP는 de Prado (2016)에 의해 제안되었다.

이러한 장점들은 HRP 포트폴리오 방법론이 현대 포트폴리오 이론의 전통적인 접근법에 비해 다양한 시장 조건에서 더 나은 성과를 내는 데 기여할 수 있다는 것을 의미한다.

Normalized Mutual Information (NID)는 상호 공유되는 정보의 양을 측정하며, 다음과 같이 정의된다.

$$NID = 1 - \frac{I(X, Y)}{\max(H(X), H(Y))} \quad (1)$$

$I(X, Y)$ 는 X 와 Y 사이의 상호 정보 거리이며, $H(X)$ 와 $H(Y)$ 는 각각 X 와 Y 의 엔트로피이다.

NID의 경우 정해진 분모의 값이 없기 때문에 여러 가지 버전이 존재하나 해당 버전은 이론적 상한인 1을 갖기 때문에 $[0, 1]$ 로 정의될 수 있으며 (이는 Pearson 상관 계수를 거리로 변환하여 사용하는 Lopez and Prado와 유사한 조건을 성립하게 만든다. NID를 활용한 HRP에서는 기존 클러스터링 결과와 새로운 클러스터링 결과의 유사성을 이 지표를 통해 평가할 수 있다. NID 값이 높을수록 두 클러스터링 결과는 유사하다고 판단하며 이를 바탕으로 계층적 클러스터링을 진행하게 된다.

이러한 두 시계열을 무작위로 섞는 방법을 사용하여 정규화된 상호 정보량(NID) 값을 보정하는 접근법은 데이터 분석에서 중요한 의미를 가진다. 특히, 작은 크기의 데이터 샘플에서 얻은 정보는 때때로 오류가 클 수 있어, 이를 보정함으로써 보다 정확한 상호작용의 측정치를 얻을 수 있다. 또한, 데이터를 무작위로 섞는 것은 과적합을 방지하는 데 큰 도움이 된다. 이는 상호 정보 거리와와 같은 통계적 의존성 지표가 특정 데이터 샘플에 너무 의존하지 않도록 하는 데 기여한다. 무작위로 섞은 데이터로부터 얻은 추정치는 원래의 시계열 데이터의 구조나 패턴에 영향을 받지 않아, 이를 원래의 NID 값에서 빼면 보다 객관적이고 신뢰성 있는 결과를 얻을 수 있다. 이 방법의 직관적인 특성은 복잡한 수학적 모델이나 알고리즘을 필요로 하지 않아 연구자나 응용 분야의 전문가들이 쉽게 이해하고 적용할 수 있게 한다. 또한, 다양한 시계열 데이터의 특성, 빈도, 또는 복잡도에 관계없이 적용 가능하다. 따라서, 두 시계열을 무작위로 섞어 randomized 된 값을 원래의 NID에서 빼는 방법은 유한 크기 효과를 효과적으로 제거할 수 있음을 알려져 있다(Boba et al., 2015). 시계열 데이터 간의 정보 상호작용을 보다 정확하게 측정하는 데 큰 도움이 된다.

$$ENID_{Y \rightarrow X}^{(1,1)}(t) = NID_{Y \rightarrow X}^{(1,1)}(t) - RNID_{Y \rightarrow X}^{(1,1)}(t) = NID_{Y \rightarrow X}^{(1,1)}(t) - \frac{1}{M} \sum_{i=1}^M NID_{Y \rightarrow X}^{(1,1)}(t), \quad (2)$$

3. 실험 결과

Table 1. 실험 대상

Full Name	Abbreviation	Tracking Index	Asset Class	Inception Date
iShares U.S. Real Estate ETF	IVR	Dow Jones U.S. Real Estate Capped Index	Equity	2000-06-12
SPDR S&P 500 ETF Trust	SPV	S&P 500 Index	Equity	1993-01-22
Materials Select Sector SPDR Fund	XLB	S&P Materials Select Sector Index	Equity	1998-12-16
Energy Select Sector SPDR Fund	XLE	S&P Energy Select Sector Index	Equity	1998-12-16
Financial Select Sector SPDR Fund	XLF	S&P Financial Select Sector Index	Equity	1998-12-16
Industrial Select Sector SPDR Fund	XLI	S&P Industrial Select Sector Index	Equity	1998-12-16
Technology Select Sector SPDR Fund	XLK	S&P Technology Select Sector Index	Equity	1998-12-16
Consumer Staples Select Sector SPDR Fund	XLP	S&P Consumer Staples Select Sector Index	Equity	1998-12-16
Utilities Select Sector SPDR Fund	XLU	S&P Utilities Select Sector Index	Equity	1998-12-16
Health Care Select Sector SPDR Fund	XLV	S&P Health Care Select Sector Index	Equity	1998-12-16
Consumer Discretionary Select Sector SPDR Fund	XLV	S&P Consumer Discretionary Select Sector Index	Equity	1998-12-16
Vanguard Communication Services ETF	VOX	MSCI U.S. Investable Market Communication Services 25% Index	Equity	2004-09-23

본 연구에서는 상기한 미국 시장 섹터 ETF의 2017년 1월부터 2023년 6월까지의 데이터를 활용하였다. 벤치마크 비교를 위해서 세 가지 세분화된 설정을 통해서 포트폴리오를 생성하였으며, Pearson 상관 계수, Spearman 순위 상관 계수, Kendall 순위 상관 계수와 NID를 통해서 HRP 포트폴리오를 생성하였으며, 이때 변동성(표준편차) 최소화, Sharpe 지수 최대화, 동일 위험 포트폴리오의 세 가지 포트폴리오 최적화 방법론을 생성하였다. 마지막으로 계층적 포트폴리오 생성 방법론에서 활용되는 연결 방법(linkage method)에 대해서는 ‘단일’, ‘완전’, ‘평균’에 해당하는 거리 방식을 선택하여 총 $4 \times 3 \times 3 \times 4$ (144개)의 포트폴리오를 생성하였다.

그 결과 결과는 해당 기간 내에서 섹터 ETF에 대해 선형 및 비선형 지표가 고루 계층적 포트폴리오 최적화를 통해 개선된 포트폴리오 평가 지표를 보일 수 있음을 의미한다.

4. 논의

포트폴리오 최적화에서 자산 간의 상호 관계 파악은 중요한 과제다. 전통적인 상관계수는 이를 평가하는 데 일반적으로 사용되지만, 본 연구에서는 정규화된 상호 정보 거리가 이러한 문제를 해결하는 데 얼마나 기여할 수 있음을 탐구하였다. NID를 활용한 HRP 포트폴리오 방법은 기존 방법론에 비해 통계적인 장점과 함께 투자 결정을 보완할 수 있는 잠재력을 보여준다.

또한, HRP 포트폴리오의 구조적 특징은 NID와 매우 잘 어울린다. 계층적인 자산 분배 방법과 NID의 조합은 포트폴리오의 안정성을 높이며, 투자 리스크를 효과적으로 분산시킬 수 있다. 이런 결합은 포트폴리오 관리의 새로운 전략으로서의 가능성을 열어놓고 있다.

5. 결론

본 연구는 정규화된 상호 정보 거리를 활용하여 HRP 포트폴리오의 효율성과 안정성을 향상시키는 방법을 제시하였다. 실험 결과를 통해, 전통적인 상관 계수 대신 NID를 사용할 때의 장점과 함께 포트폴리오 성능이 경쟁력이 있음을 확인하였다. 이 연구 결과는 투자자와 포트폴리오 매니저들에게 새로운 투자 전략의 가능성을 제시하며, 포트폴리오 최적화 분야에 새로운 연구 방향성을 제안한다.

참고문헌

[1] Boba, P., Bollmann, D., Schoepe, D., Wester, N., Wiesel, J., & Hamacher, K. “Efficient computation and statistical assessment of transfer

- entropy” . *Frontiers in Physics*, 3, 10, 2015.
- [2] Dimpfl, T., & Peter, F. J. "Using transfer entropy to measure information flows between financial markets." *Studies in Nonlinear Dynamics and Econometrics*, 17(1), 85-102, 2013.
- [3] Lopez de Prado, M. "Building diversified portfolios that outperform out of sample." *Journal of Portfolio Management*, 42(4), 59-69, 2016.
- [4] Kraskov, A., Stögbauer, H., & Grassberger, P. "Estimating mutual information." *Physical review E*, 69(6), 066138, 2004.
- [5] Kvålseth, T. O. On normalized mutual information: measure derivations and properties. *Entropy*, 19(11), 631. 2017.
- [6] Hacine-Gharbi, A., & Ravier, P. "A binning formula of bi-histogram for joint entropy estimation using mean square error minimization." *Pattern Recognition Letters*, 101, 21-28, 2018.