



에너지 선물 시장의 수익률 네트워크 구축과 선형 및 비선형 인과관계 네트워크 분석 결과의 예측 문제에서의 활용에 관한 연구

최인수¹, 김우창^{1,*}

¹한국과학기술원 산업및시스템공학과

서론

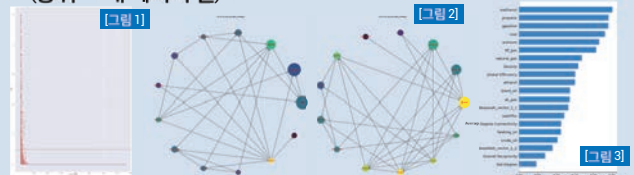
- 에너지는 사회 발전과 경제 성장을 위한 중요한 구성 요소
- 에너지 시장의 가격 변화는 에너지 시장에서 **다양한 자원의 흐름과 분배에 지대한 영향**을 미침
- 에너지 가격 예측의 중요성은 세 가지 범주로 나눌 수 있음:
 - (a) 정확한 에너지 가격 예측은 동적 비용 제어를 가능케 함
 - (b) 정확한 에너지 가격 예측은 시장 동향과 기회에 대한 정확한 이해를 용이하게 함
 - (c) 정확한 에너지 가격 예측은 정책 정의와 시장 규제에 대한 대한 과학적 기초를 제공
- 이에 본 연구에서는 에너지 선물에 대해서 수익률 흐름에 대한 **선형 및 비선형 인과관계 네트워크를 구축**
- 또한, 구축된 네트워크의 분석 결과, 즉 네트워크의 위상을 포함한 주요 네트워크 지표가 에너지 선물의 **가격 예측에서 활용될 수 있는지에 대해서 실험을 실시**

연구 방법론

- 그레인저 인과 관계 (Granger Causality)**
 - 기존의 가장 보편적인 인과관계 측도는 Granger(1969)의 Granger 인과관계(Granger Causality)로, 이 모델을 통하여 **선형적인 인과관계를 검증할 수 있음**
 - 두 시계열 X_t, Y_t 에 대하여 Y_t 의 정보가 포함되는 것이 X_t 의 선형 예측에 더욱 나은 효과를 보일 경우, 즉 A_{12} 와 A_{22} 가 통계적으로 유의미할 경우 이를 Y_t 가 X_t 를 **Granger 인과한다고 함**
 - Granger 인과관계를 검증하기 위해서는 X_t, Y_t 에 대하여 위와 같은 벡터자기회귀모형 $Var(p)$ 을 사용하며 해당 모형은 정규성을 가정해야 하나 본 연구의 데이터를 비롯한 대다수의 수익률 데이터를 기반으로 한 벡터자기회귀모형은 최적 시차를 설정하여도 **정규성을 만족하지 못한다는 한계가 있음**
 - $X_t = \alpha_1 + \sum_{j=1}^p A_{11} X_{t-j} + \sum_{j=1}^p A_{12} Y_{t-j} + \varepsilon_1(t)$
 - $Y_t = \alpha_2 + \sum_{j=1}^p A_{21} X_{t-j} + \sum_{j=1}^p A_{22} Y_{t-j} + \varepsilon_2(t)$
- 효율적 전이 엔트로피 (effective transfer entropy, ETE)**
 - 전이 엔트로피(transfer entropy, TE)는 두 시계열의 엔트로피 차이를 계산하여 두 시계열 사이의 **비선형적 정보 흐름을 측정하는 방법론**
 - 본 연구에서는 과거의 모든 시장정보를 반영되어 있음을 표현하기 적절한 시계열의 시차가 모두 1인 경우를 고려하며 전이 엔트로피는 아래 식과 같음
 - $H(X|Y) = - \sum_{x \in X, y \in Y} p(x, y) \log_2 \frac{p(x, y)}{p(y)}$
 - $TE_{Y \rightarrow X}^{(k,1)}(t) = H(X_{t+1}|X_t, \dots, X_{t-k+1}) - H(X_{t+1}|X_t, \dots, X_{t-k+1}, Y_t, \dots, Y_{t-1+1})$
 - TE는 비모수적(nonparametric) 방법이므로 순열 검증법을 활용하여 p-값을 계산하였으며, 해당 방법은 **Boba et al. (2015)에 의해 제시되었으며, 본 연구에서는 해당 방법론을 활용하였으며 총 10,000회의 시뮬레이션을 거침**
 - $ETE_{Y \rightarrow X}^{(1,1)}(t) = TE_{Y \rightarrow X}^{(1,1)}(t) - RTE_{Y \rightarrow X}^{(1,1)}(t) = TE_{Y \rightarrow X}^{(1,1)}(t) - \frac{1}{M} \sum_{i=1}^M TE_{Y \rightarrow X}^{(1,1)}(t)$

실험 결과

- 2016년 - 2021년까지의 모든 로그 수익률에 대해 7가지 정규성 검정을 실시, [그림 1]의 브렌트유에 대한 예시처럼 부분 데이터 집합의 수가 커질수록 7가지 정규성 검정의 결과에서 p-값이 감소 즉 **정규성이 오히려 감소하는 결과도 도출**되며 이러한 분포는 멱함수 분포(power law distribution)에 매우 적합이 잘 되는 결과를 도출 (모든 가능한 경우의 데이터 세트의 3.21%만이 정규성 만족)
- 이러한 결과는 **전이 엔트로피를 통한 비선형 인과관계 측정을 하는 통계적 근거로 해석 가능**
- 계산된 20일, 60일, 120일, 240일 (단기, 중단기, 중장기, 장기)의 그레인저 인과관계와 효율적 이전 엔트로피 값에 대해서 유의 수준 0.1 하에서 해당 측도 값을 **가중치(weight)**로 두는 인과관계 그래프 도식화 결과는 [그림 2]와 같음 (**2019년 6월 14일, 60일 이동창, 좌측이 그레인저 인과관계, 우측이 ETE**)
- 해당 결과에서 주요 네트워크 지표와 Deepwalk를 통한 네트워크 임베딩 결과를 활용하여 2017 - 2020년 데이터를 통해 2021년 데이터의 가격 등락을 XGBoost, LightGBM, CatBoost를 활용하여 예상한 결과, Accuracy를 계산한 결과 13개 데이터 세트의 지표가 정확도를 기준으로 57.36%에서 60.11%의 정확도로 개선되는 것을 나타냈으며 결과의 강건성을 위해 다양한 초모수로 실험을 진행하여도 역전되는 경우가 없었음
- [그림 3]은 Train 모형의 브렌트유에 대한 mean(SHAP) 값 순위 (상위 20개 데이터 열)



결론

- 본 연구에서는 에너지 선물에 대해서 수익률 흐름에 대한 **선형 및 비선형 인과관계 네트워크를 구축**
- 그 결과 주요 에너지 가격끼리 **선형 및 비선형 분석에서 모두 통계적으로 유의미한 인과관계가 발생**함을 확인
- 구축된 네트워크에 대한 위상을 포함한 **주요 네트워크 지표를 에너지 선물의 가격 예측에서 활용될 수 있는가에 대한 가설을 검증**하고자 기계 학습 모형을 통한 실험을 실시
- 그 결과, 예측력이 개선되었음을 성과 지표를 통해 확인하였으며 금융 시장이 복잡계 네트워크로 분류되기 시작한 최근의 흐름과 더불어 그 구조가 **금융 시장을 이해하고 예측하는 데에 통찰력을 제공할 수 있음**을 시사하며 향후 해당 연구를 발전시킬 예정

참고문헌

- Granger, C. W. (1969). Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica: Journal of the Econometric Society*, 424-438.
- Schreiber, T. (2000). Measuring Information Transfer. *Physical Review Letters*, 85(2), 461.
- Boba, P., Bollmann, D., Schoepe, D., Wester, N., Wiesel, J., & Hamacher, K. (2015). Efficient computation and statistical assessment of transfer entropy. *Frontiers in Physics*, 3, 10.
- Perozzi, B., Al-Rfou, R., & Skiena, S. (2014, August). Deepwalk: Online learning of social representations. In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 701-710).