

고유포트폴리오

황정민

* Case study 1- Portfolio Management: Finding an Eigen Portfolio

- 포트폴리오 관리의 가장 중요한 목적은 리스크 대비 수익을 최대화하도록 자산을 배분하는 것이다.
- 차원축소를 활용해 포트폴리오를 구성하고, 백테스트를 시행함으로써 전략을 평가

1. 문제 정의

- 다우존스 산업평균 지수(DIJA)와 30개 기업의 주가 데이터를 사용
- 차원축소를 통해 만든 포트폴리오들의 수익률을 벤치마크와 비교하고, 백테스트를 진행

2. 데이터와 패키지 가져오기

2-1) 패키지 가져오기

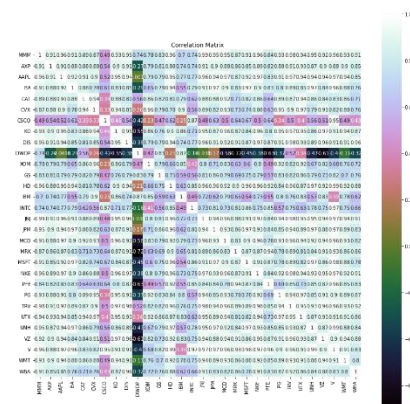
2-2) 데이터 가져오기

3. EDA

3-1) 데이터 살펴보기

3-2) 데이터 시각화

- 히트맵을 그려 봄으로써 각 기업들의 수정 증가들의 상관관계를 시각화



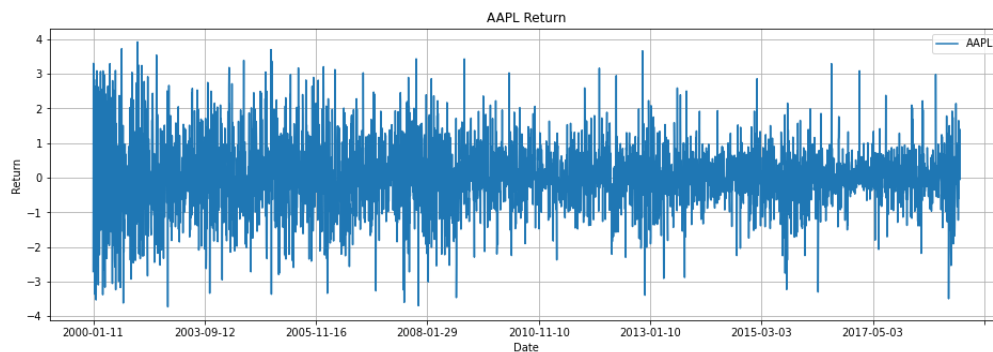
4. 데이터 전처리

4-1) 결측치 처리 및 수익률 계산

- 결측치의 비율이 30%가 넘는 기업들의 데이터는 사용하지 않기로 함
- 나머지 데이터들의 결측치는 이전의 값들로 채워 주기로 함.
- 수익률을 계산하고, 3σ 를 넘는 데이터는 이상치로 간주하여 제거하기로 함.
- 수익률 데이터를 사용해 이후의 작업을 진행

4-2) 스케일링

- 차원축소를 잘 이뤄지기 위해서는 반드시 스케일링을 진행해줘야 함
- StandardScaler를 사용해 스케일링을 진행하고, 데이터를 시각화



5. 모델 사용하기

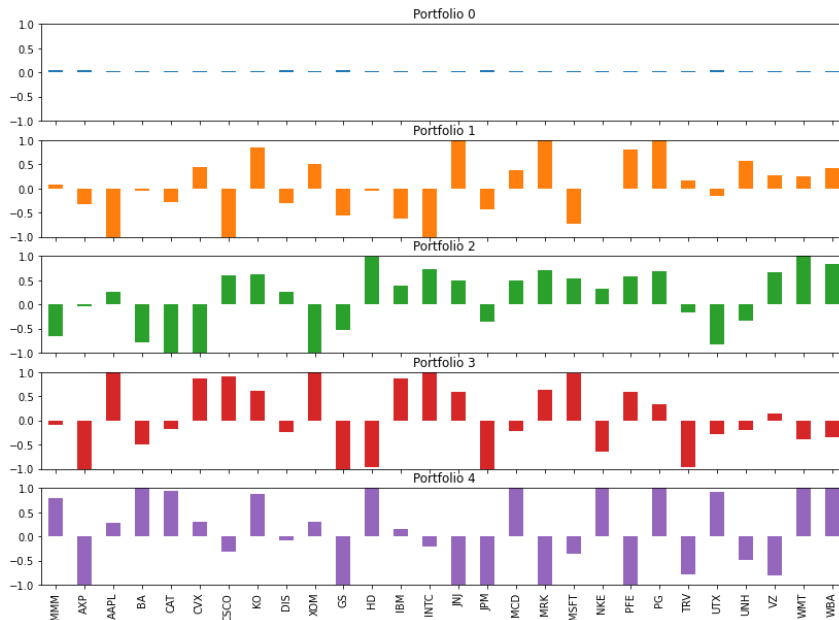
5-1) 훈련 세트와 테스트 세트 나누기

- 시간 순으로 앞에서 80%의 데이터를 훈련 세트로 나머지 20%의 데이터를 테스트 세트로 사용하기로 함

5-2) 모델 사용하고, 평가하기

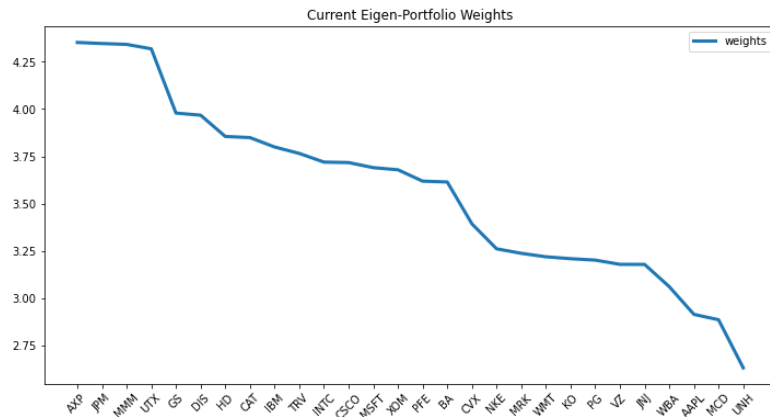
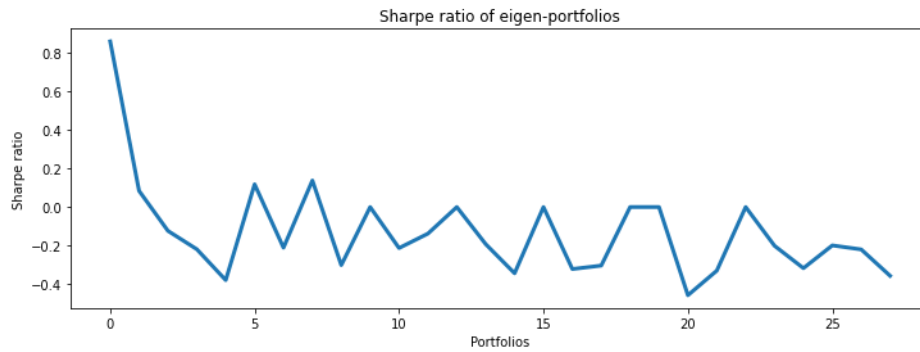
5.2.1) 설명된 분산(Explained Variance)

- PCA를 사용하고 설명된 분산을 살펴본 결과, 하나의 주성분이 40% 정도의 분산을 설명하는 것으로 나타남. 또한 약 10개의 주성분들 만으로 70% 이상의 분산을 설명하는 것을 알 수 있었음
- 가장 큰 Factor는 시장 Factor라고 볼 수 있다.



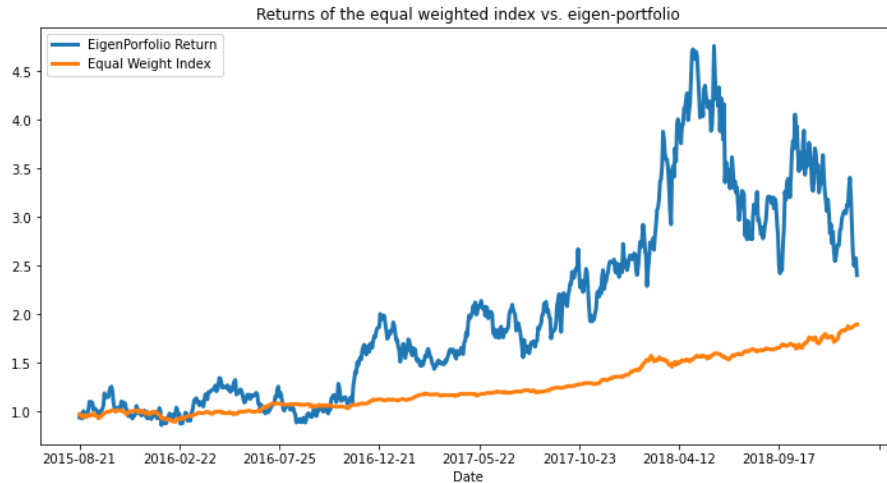
5.2.3 Finding the best eigen portfolio

- 샤프비율을 사용하여 위험 대비 수익성이 가장 좋은 포트폴리오를 선정
- 샤프비율을 기준으로 가장 수익률이 좋은 포트폴리오 4개는 고유포트폴리오 0,7,5,1이었다.

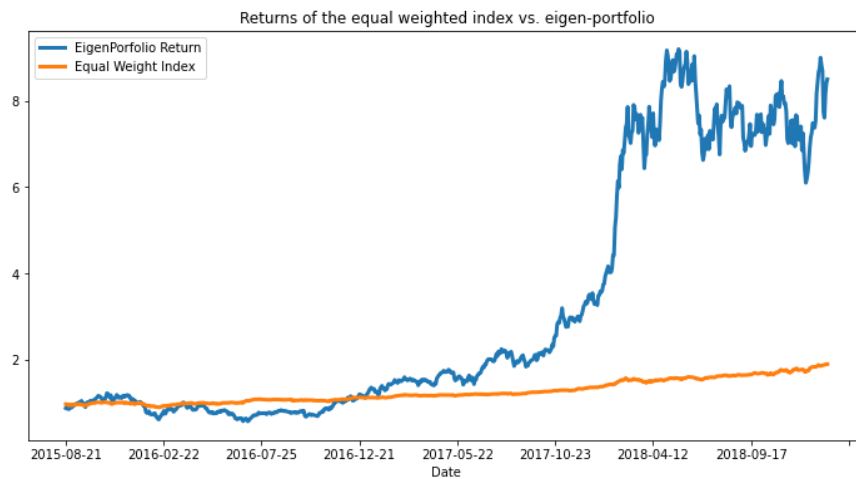


5.2.4 백테스팅

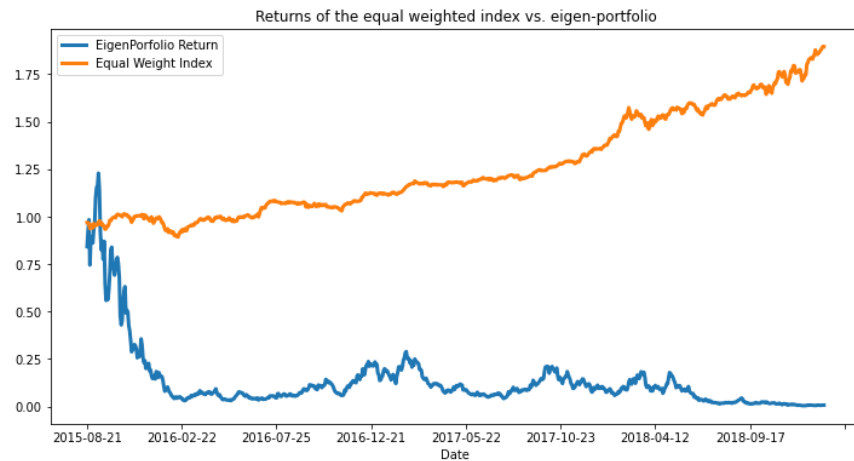
- 각각의 eigen-portfolio마다 기업들의 비중이 다르기 때문에 이에 따라 수익률의 차이가 발생
- 백테스팅 결과가 좋았던 포트폴리오와 나빴던 포트폴리오를 살펴본 결과, 고유포트폴리오 5와 1은 백테스트에서도 좋은 성능을 보여줌



5번 고유 포트폴리오



1번 고유 포트폴리오



14번 고유 포트폴리오

6. 결론

- PCA를 사용해 차원축소를 진행하고, eigen portfolio를 구성
- 가장 explained_variance_ratio가 높은 eigen portfolio는 체계적 위험을 나타냈고, 나머지 eigen portfolio는 산업과 관련되어 있음

* 코스피100에 포함되어 있는 종목들로 고유포트폴리오 구성하기: ipynb파일 참조