

주식 포트폴리오를 위한 알파 팩터 구축

~시장 비효율성을 활용한 체계적 접근~

2025-01-18

팩터 투자란?

- 정의: 자산 가격에 영향을 미치는 시장 비효율성 (예: 가치, 크기, 모멘텀)
- 목적:
 - **위험 조정 수익 극대화**: 더 적은 위험으로 높은 수익을 추구하는 투자 전략
 - **원하지 않는 위험 최소화**: 의도하지 않은 손실 가능성을 줄이는 방법
 - **트레이딩 엣지 확보 (비무작위적 기회 활용)**: 시장의 비효율성이나 패턴을 이용해 꾸준히 이익을 얻는 기술

팩터의 역할

- 투자의 기본 구성 요소
 - 자산 가격에 지속적으로 영향을 미침
 - 다양한 자산 클래스에서 수익 동인으로 작용
- 시장 비효율성 활용
 - 이상 현상을 기반으로 이익 창출
 - 트레이더의 수익 원천 제공

팩터 분석의 핵심

- 팩터 식별
 - 데이터 분석으로 잠재적인 수익 동인 탐색
- 포트폴리오 민감도 파악
 - 각 팩터가 포트폴리오에 미치는 영향을 평가
- 액션 실행
 - 원하지 않는 위험: 헷지(hedge)로 제거
 - 원하는 팩터 노출: 전략적으로 증가

다루는 내용

- 주성분 분석을 사용하여 잠재적인 수익 동인을 식별하기
- 선형 회귀를 사용한 포트폴리오 베타 계산 및 베타 헷징
- Fama-French 팩터에 대한 포트폴리오 민감도 분석
- 변동성을 기반으로 시장 비효율성 평가
- ~~Zipline Reloaded Pipeline API를 사용하여 팩터 순위 모델 준비~~

주성분 분석(PCA)를 사용한 잠재적 수익 동인 식별

- PCA란?

- 차원 축소 기법: 데이터의 복잡성을 줄이면서 중요한 정보를 유지.
- 주성분: 데이터의 최대 분산을 설명하는 새로운 특징 집합.
- 목적: 큰 변수 집합을 더 작은 집합으로 변환해 분석 효율성 향상.

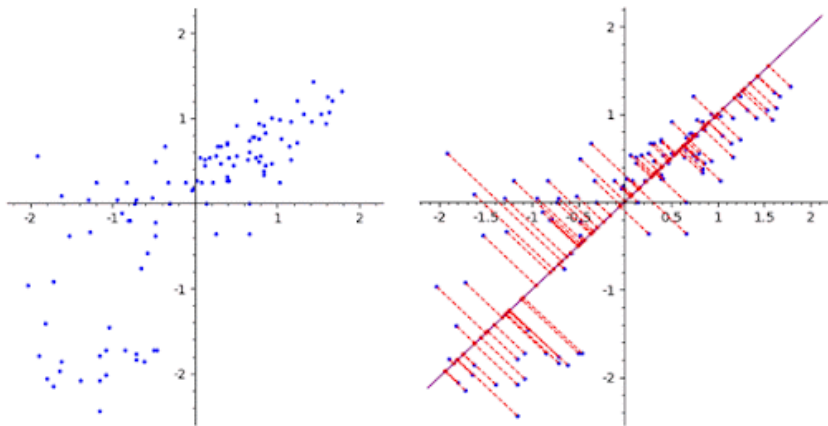


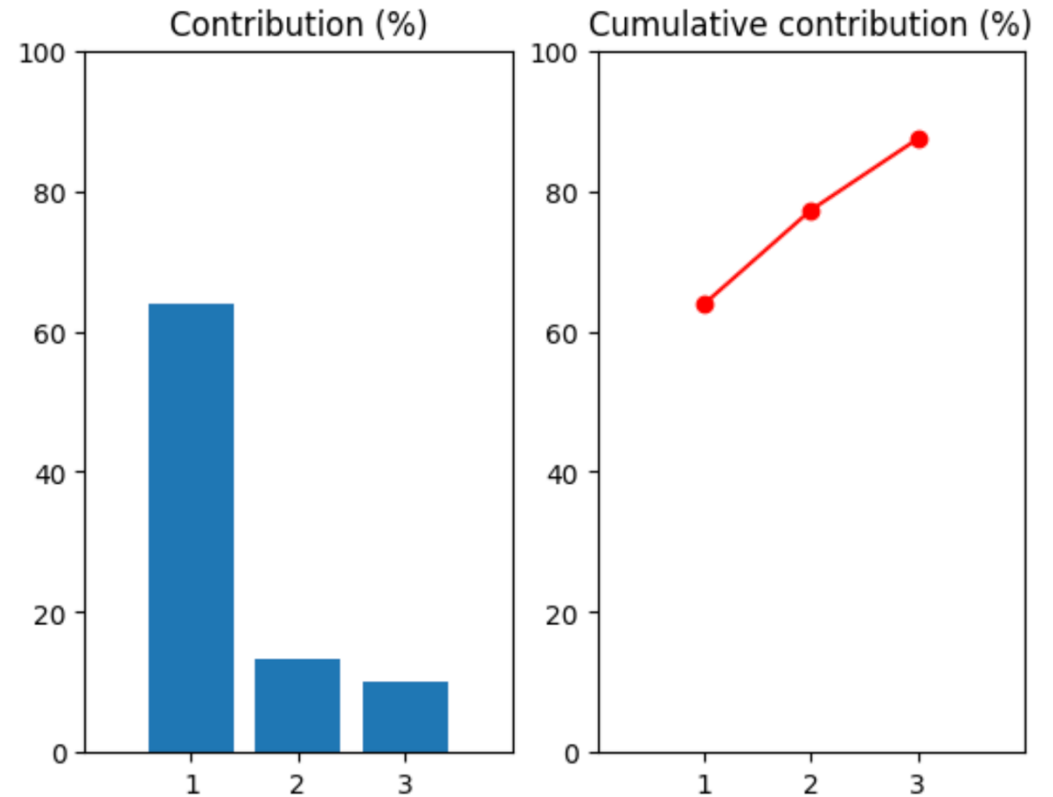
그림 : 105명의 GPA 데이터 분포와 PCA를 통해 추출된 주요 주성분 방향을 시각화한 그래프
(소스 : <http://matrix.skku.ac.kr/math4ai-intro/W12/>)

- PCA의 역할

- 수익률 분해:
 - 포트폴리오의 수익률을 최대 분산을 설명하는 주성분으로 분해
 - 각 주성분은 주요 위험 요인을 나타냄
- 주요 위험 요인 파악:
 - 처음 몇 개의 주성분이 전체 분산의 대부분을 설명
 - 이를 통해 포트폴리오의 핵심 위험을 정량화

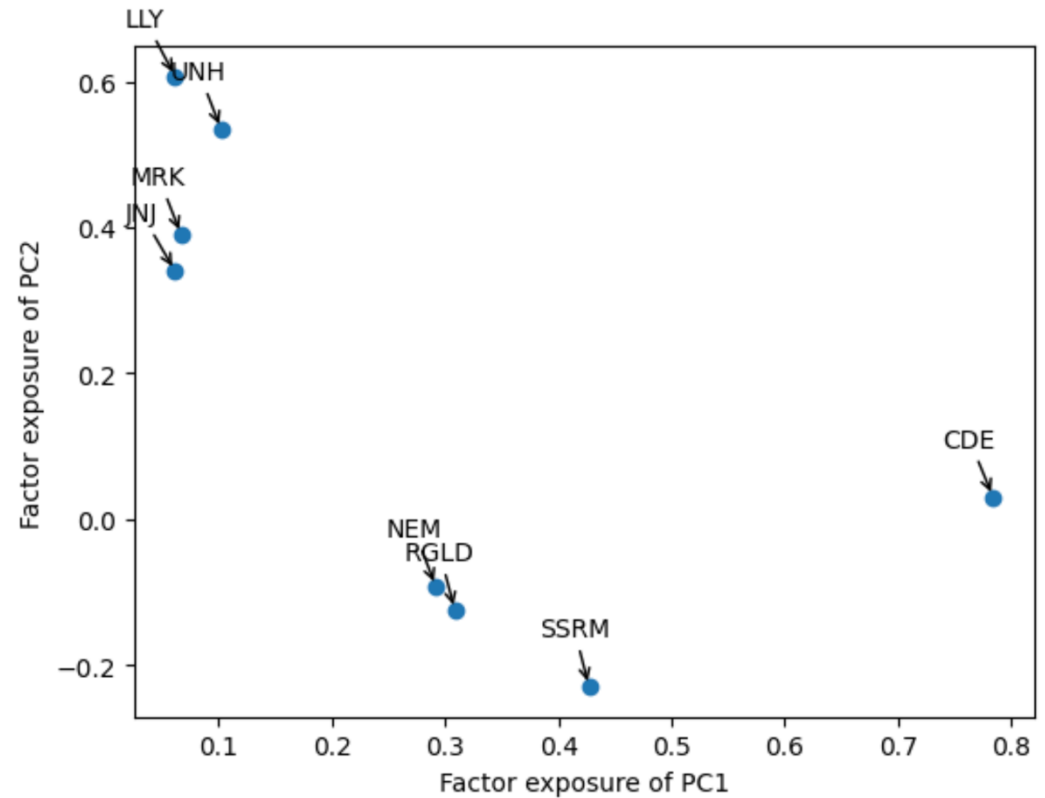
포트폴리오의 위험 요인 식별

- PCA를 통해 포트폴리오의 주요 위험 요인들을 찾아내고 있음
- 첫 번째 주성분이 전체 변동성의 약 64%를 설명
- 두 번째 주성분이 약 13%를 설명
- 세 번째 주성분이 약 10%를 설명



PCA를 활용한 포트폴리오 요인 분석

- 그룹
 - 광산/귀금속 관련 주식 (NEM, RGLD, SSRM, CDE)
 - 헬스케어 관련 주식 (LLY, UNH, JNJ, MRK)
- 주요 관찰
 - 금광 주식은 헬스케어 주식과 명확히 분리된 클러스터를 형성하고 있으며, 이는 두 섹터가 서로 다른 요인에 의해 주도된다는 것을 시사
 - 동일 섹터 내에서도 개별 주식들의 요인 노출도에는 차이가 있어, 다양한 리스크 및 수익 패턴을 반영



선형 회귀를 사용하여 포트폴리오 베타 계산 및 베타 헷징

- 목표 : 선택적 노출
 - 높은 수익 가능성이 있는 **특정 위험에 노출**
 - 불리하거나 불필요한 **다른 위험은 헷지**
 - **E.g.** : 시장 위험(전반적 수익률)을 헷지하면서, **주가수익비율(P/E)이 낮은 주식에 집중 투자**
- 팩터 모델 기본 형태
 - $Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$
 - **Y**: 포트폴리오/자산의 수익률
 - **α** : 기본 수익률 (팩터 영향 제외)
 - **β** : 각 팩터에 대한 민감도
 - **X**: 시장수익률, 기업규모 등 위험 요인
 - 포트폴리오 수익의 위험 요인에 대한 민감도는 베타(β)로 설명.
 - 헷지가 필요한 것은 베타, 위험 요인에 대한 노출을 집중적으로 관리해야

베타 헷징(Beta Hedging)

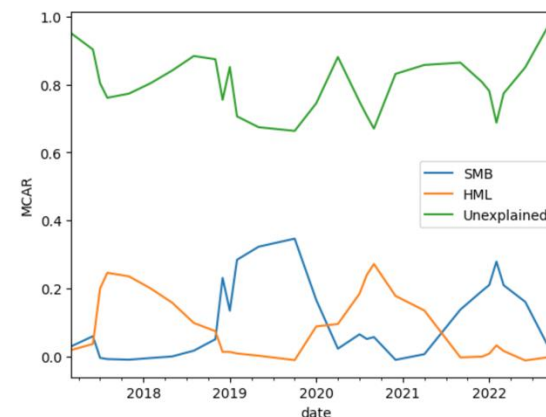
- 베타 헷징이란?
 - 베타 헷징은 특정 시장 위험에 대한 노출을 제거(헷지)하여, 포트폴리오의 성과가 시장 전반의 변동성과 무관하도록 만드는 방법
 - 목표는 원하지 않는 위험(시장 위험)을 제거하고, **알파(α)**를 기반으로 한 초과 수익을 유지하는 것
- 작동 방식
 - 1. 포트폴리오와 벤치마크 데이터 준비
 - 포트폴리오: 8개의 주식으로 구성된 포트폴리오의 일일 수익률을 합산.
 - 벤치마크: S&P 500을 추적하는 ETF(SPY)의 일일 수익률을 대리로 사용.
 - 포트폴리오와 SPY의 일일 수익률 계산.
 - 2. 선형 회귀를 통한 베타 계산
 - $Y = \alpha + \beta_{SPY}X$
 - α : 포트폴리오의 알파(초과 수익)
 - β_{SPY} : 포트폴리오의 시장 위험 민감도
 - 3. 헷지된 포트폴리오 생성
 - 시장 위험 제거: SPY에 대한 숏 포지션 설정.
 - 헷지된 포트폴리오 수익률 계산: $\text{hedged_portfolio_returns} = \text{portfolio_returns} - \beta_{SPY} \cdot \text{benchmark_returns}$
 - 헷지 효과: $\alpha + \beta_{SPY} - \beta_{SPY} = \alpha$
 - 시장 위험(β)을 제거하고, 포트폴리오 성과는 **알파(α)**로 제한.
 - 4. 헷지 검증
 - 헷지된 포트폴리오의 베타가 0에 가까운지 확인하기 위해 회귀를 다시 실행.
 - 성공적인 헷지: 시장 위험 민감도 제거(베타 ≈ 0).

Fama-French 요인에 대한 포트폴리오 민감도 분석

- Fama-French 요인이란?
 - 정의: 주식 수익률 변화를 설명하는 주요 요인으로, 경제학자 Eugene F. Fama와 Kenneth R. French가 개발.
 - Fama-French 3요인 모델:
 - 시장 요인 (Market Factor): 시장 전체 수익률과의 상관성.
 - 크기 요인 (Size Factor): 소형주(Small Cap) vs 대형주(Large Cap).
 - 가치 요인 (Value Factor): 고시장가치(P/B 비율 높은 주식) vs 저시장가치(P/B 비율 낮은 주식).
- 포트폴리오 민감도 계산:
 - 크기(size)와 가치(value) 요인에 대한 포트폴리오의 노출 정도를 측정.
 - 시장 움직임과 개별 요인에 따른 성과를 분리.

활성 위험에 대한 한계 기여도 (MCAR)

- MCAR이란?
 - 정의: 각 요인이 포트폴리오의 활성 위험(벤치마크 대비 초과 변동성)에 기여하는 정도를 정량화한 지표.
 - 목적:
 - 포트폴리오 내 요인별 위험 기여도 파악.
 - 다른 요인과의 상호작용을 고려해 리스크 관리를 최적화.
 - MCAR 계산 방법
 - b_i : 요인 i 의 민감도.
 - Σ_{ij} : 요인 i 와 j 간의 공분산.
 - σ_{active}^2 : 포트폴리오 활성 수익률의 분산(표준 편차 제곱).
- $$\text{MCAR}_i = \frac{(b_i \cdot \Sigma_{ij} \cdot b_j)}{\sigma_{\text{active}}^2}$$
- MCAR의 의미
 - 요인별 위험 기여도 분석:
 - 포트폴리오 내 각 요인이 얼마나 많은 위험을 추가하는지 계산.
 - 다른 요인과의 공분산 효과 반영.
 - 설명되지 않은 위험 기여도:
 - 분석 중인 요인 외의 기타 요인에 대한 노출로 발생하는 리스크.



요인 별 한계 위험 기여도 추이

변동성을 기반으로 시장 비효율성 평가

- 변동성 요인:
 - 변동성이 큰 주식의 가격 책정은 시장 비효율성을 반영.
 - 낮은 변동성 주식이 높은 수익을 낼 가능성이 더 크다.
 - 전통 금융 이론에 대한 모순: 위험이 높을수록 보상이 커야 한다는 가설과 상충.
- 활용 방안
 - 요인 기반 투자 전략:
 - 변동성 요인을 활용해 저변동성 주식 중심의 포트폴리오 구성.
 - 시장 비효율성 탐색:
 - 변동성과 수익률 간 관계를 분석해 예측 가능한 이상 현상 활용.
- 분석 과정
 - 미래 수익률 계산:
 - 특정 시점 이후의 증권 수익률.
 - 미래 수익률을 통해 요인의 예측력을 평가.
 - 스피어만 순위 상관계수 활용:
 - 요인과 미래 수익률 간의 단조 관계를 평가.
 - 높은 상관계수: 요인이 미래 수익률을 잘 예측.
 - 낮은 상관계수: 요인의 예측력이 부족.
 - 파킨슨 변동성 사용:
 - 고가와 저가를 기반으로 변동성을 더 정확하게 측정.
 - 고가와 저가는 종가보다 더 많은 가격 정보를 포함.

Parkinson 변동성 추정치

- 정의

- Parkinson 변동성 추정치는 주식의 일중 고가와 저가를 기반으로 보다 정밀하게 변동성을 측정하기 위한 방법입니다.
- 기존의 종가 기반 변동성 측정 방식보다 가격 변동의 실제 범위를 더 잘 반영하는 것이 특징입니다.

- Parkinson 변동성의 특징

- 고가와 저가 사용:
 - 종가만을 사용하지 않고, 일중 고가와 저가의 비율을 활용하여 가격 변동의 전체 범위를 계산.
 - 종가만 사용할 때 놓칠 수 있는 가격의 극단적 움직임을 반영.

- 변동성 추정 공식:

- 일중 고가와 저가의 로그 비율을 제공하여 변동성을 계산.

$$rs = \frac{1}{4 \cdot \ln(2)} \cdot \left(\ln \left(\frac{\text{High}}{\text{Low}} \right) \right)^2$$

- trading_days(연간 거래일 수)를 반영하여 연간화된 변동성으로 변환.

- 표준화:

- 계산된 변동성 값을 평균 0, 표준편차 1로 표준화하여 비교 가능하게 만들.

$$\text{표준화된 변동성} = \frac{\text{변동성 값} - \text{평균}}{\text{표준편차}}$$



```
def parkinson(data, window=14, trading_days=252):  
    # Parkinson 변동성 추정치를 계산하는 함수  
    # data: 주가 데이터프레임 (high, low 가격 포함)  
    # window: 이동 윈도우 크기 (기본값 14일)  
    # trading_days: 연간 거래일수 (기본값 252일)  
  
    # 일중 고가/저가 비율의 로그값을 제공하여 Parkinson 변동성 계산  
    rs = (1.0 / (4.0 * np.log(2.0))) * ((data.high / data.low).apply(np.log)) **  
2.0  
  
    def f(v):  
        # 연간화된 변동성으로 변환  
        return (trading_days * v.mean()) ** 0.5  
  
    # 이동 윈도우 적용하여 변동성 계산  
    result = rs.rolling(window=window, center=False).apply(func=f)  
  
    # 표준화된 변동성 값 반환 (평균=0, 표준편차=1)  
    return result.sub(result.mean()).div(result.std())
```

Q&A