
Cointegration-Based Pairs Trading with Rolling Z-Score: A Walk-Forward Backtesting Framework

이름: 지윤승 | 학번: 20259431



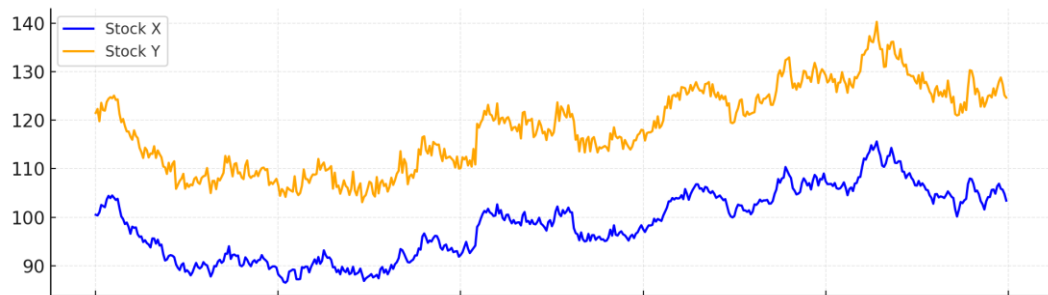
Motivation & Background

페어트레이딩의 핵심 아이디어

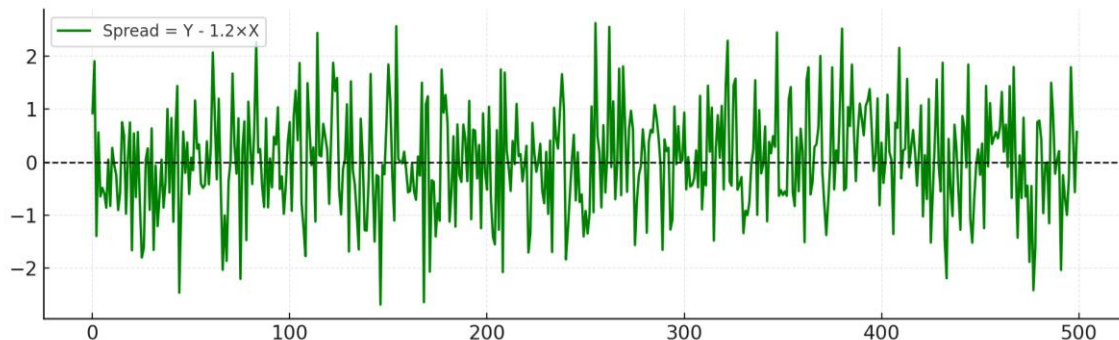
- 시장 중립 기반의 상대가치 전략이다.
- 가격 관계가 유사한 두 자산을 대상으로, 저평가된 자산 매수, 고평가된 자산 매도를 통해 가격 관계가 평균으로 회귀할 때 수익을 얻는 방식이다.

공적분 기반 접근의 필요성

- 단순 상관관계만으로는 가격이 장기적으로 괴리될 수 있다.
- 공적분은 두 자산 가격 사이의 안정적인 장기 균형 관계를 보장한다.
- 이를 통해 스프레드가 정상성을 가지므로, 진입/청산 시점이 통계적으로 더 견고해진다.



→ 함께 움직이면서도 일정한 장기 균형 관계를 보임



→ 두 시계열 간 스프레드가 0을 중심으로 정상성을 보임

Data & Test Setting

Data

항목	내용
데이터 소스	sp500_return.csv(일별 수익률), sp500_list.csv(S&P500 편입/편출 일자), stock_id.csv(PERMNO-종목명 매핑)
기간	1960년 ~ 1975년
테스트 대상 기간	1961-01-01 ~ 1975-12-31
유니버스 구성	S&P 500 종목
전처리	누락값 제거 후, 누적 수익률로 변환하여 가격 시계열 생성

- 포메이션 구간 시작을 위해 1960년 데이터부터 로드하여 1961년 1월부터 본격적으로 백테스트를 시작했다.
- 전체 기간 동안 Rolling Formation-Test 윈도우를 적용하여 전략을 검증했다.

Test Setting

항목	설정 내용
포메이션 기간	12개월 (공적분 페어 선정에 사용)
테스트 기간	6개월 (선정된 페어로 트레이딩 수행)
롤링 방식	포메이션-테스트 구간을 순차적으로 이동하는 워킹-포워드(rolling) 구조
거래비용	1회 체결 10 bps (왕복 약 20 bps)
슬리피지	1회 체결 5 bps
룩어헤드 방지	Z-score 계산 시 t-1 데이터만 사용하여 신호 생성
포트폴리오 제약	총 레버리지 1배, 페어당 노출 상한 설정

- 각 포메이션 구간에서 공적분 검정을 통해 페어를 선정하고, 이어지는 6개월 테스트 구간에서 Z-score 기반으로 트레이딩을 수행했다.
- 이 과정을 전체 테스트 기간 동안 반복하여 전략 성과를 평가했다.

Strategy Overview

전략 단계 요약

단계	내용
① 데이터 준비	1960~1975년 S&P 500 수익률 데이터 사용
② 포메이션 (12개월)	공적분 검정(OLS + ADF + Half-life + FDR)으로 정상성 페어 선정
③ 테스트 (6개월)	롤링 Z-score 기반 진입·청산·손절 신호 생성
④ 포트폴리오 구성	선정된 페어를 동일가중 방식으로 합산, 총 레버리지 1배 제약
⑤ 비용 반영	거래비용 10bps + 슬리피지 5bps, 턴오버 기반 일일 차감
⑥ 룩어헤드 방지	t-1 데이터만으로 신호 생성
⑦ 워킹-포워드 백테스트	12개월 포메이션 + 6개월 테스트를 순차적으로 반복 적용 (1980-1990)
⑧ 성과·유의성 검정	Newey-West 및 Bootstrap 검정을 통해 전략 수익의 통계적 유의성 검증

- 본 전략은 공적분 관계를 이용해 평균회귀 성향이 강한 페어를 선별하고, 롤링 Z-score 기반 신호로 진입·청산을 수행한다.
- 거래비용·슬리피지·룩어헤드 방지 및 레버리지 제약 등을 모두 반영하여 실전 운용 환경에 근접하게 설계하였으며, 워킹-포워드 방식으로 전체 기간(1961년 ~ 1975년)에 걸쳐 전략을 평가한다.
- 포트폴리오는 동일가중 방식으로 구성되며, 마지막 단계에서 통계적 유의성(Newey-West, Bootstrap)을 검정한다.

Why This Strategy?

① 장기균형 관계에 기반한 안정적 페어 선택

공적분 검정을 통해 장기균형 관계가 성립하는 종목쌍을 선택하면 스프레드가 평균으로 회귀할 가능성이 높기 때문에, 이러한 경제적·통계적 기반을 갖춘 페어를 사용하면 보다 안정적이고 재현성 있는 차익 기회를 포착할 수 있을 것으로 기대하였다.

② Z-score를 통한 체계적 진입·청산 규칙

스프레드의 과도한 이탈을 정량적으로 식별할 수 있는 z-score 규칙을 적용하면 시장의 일시적 오버슈팅 상황에서 효율적으로 진입할 수 있고, 평균 회귀 구간에서 체계적으로 수익을 실현할 수 있어 예측 가능하고 일관된 수익 구조를 만들 수 있을 것으로 기대하였다.

③ 실거래 환경을 반영한 현실적 구현

턴오버 기반 거래비용, 레버리지 및 노출 제한, 워킹-포워드 검증 등 실거래 환경을 고려한 구조를 도입하면 비용과 리스크를 감안한 후에도 전략이 실제 시장 조건에서 유효하게 작동할 가능성이 높아질 것으로 기대하였다.

Pair Selection: Cointegration

① 로그가격 회귀 (OLS)

- 두 자산 X_t, Y_t 의 로그가격에 대해 회귀 수행

$$\log Y_t = \alpha + \beta \log X_t + \varepsilon_t$$

→스프레드 ε_t 추정

② 잔차 정상성 검정 (ADF)

- ε_t 에 대해 ADF 단위근 검정
- p-value ≤ 임계값(예: 5%)인 페어만 유지
→ 정상성(Stationarity) 확보

③ 다중검정 보정(FDR)

- 여러 페어를 동시에 검정할 때 1종 오류 제어
- Benjamini-Hochberg 방식으로 p-value 보정

$$p_i \rightarrow p_i^{adj}$$

④ Half-life 계산

- 스프레드의 평균 회귀 속도를 추정

$$\tau = -\frac{\ln 2}{\phi}$$

- ϕ : AR(1) 계수 → $\varepsilon_t = \phi \varepsilon_{t-1} + u_t$
→ 짧을수록 빠른 평균회귀

⑤ Top 3 페어 선정

- 점수 산식:

$$\text{Score} = \frac{-\ln(p_{adj})}{1 + \text{Half-life}}$$

→ 점수가 높은 상위 3개 페어 최종 선정

Trading Rules (Z-score)

① 스프레드 및 Z-score 계산

두 자산 X_t 와 Y_t 의 로그가격을 이용해 스프레드를 정의한 뒤, 롤링 윈도우를 사용하여 평균과 표준편차를 계산하고 이를 통해 Z-score를 산출한다.

$$\text{Spread}_t = \log Y_t - \alpha - \beta \log X_t, z_t = \frac{\text{Spread}_t - \mu_{t-1}}{\sigma_{t-1}}$$

여기서 μ_{t-1} 과 σ_{t-1} 는 t-1일까지의 데이터를 이용해 계산되며, 신호는 전일 정보를 기반으로 생성하고 실제 거래는 t일에 집행함으로써 록어헤드를 철저히 방지한다.

② 진입 규칙 (Entry)

Z-score의 절대값이 진입 임계값 k_{in} 이상이 되면 포지션을 개시한다.

$z_t > k_{in}$ 인 경우: 스프레드가 축소될 것으로 보고 Y를 매도, X를 매수

$z_t < -k_{in}$ 인 경우: 스프레드가 확대될 것으로 보고 Y를 매수, X를 매도

Entry when $|z_t| \geq k_{in}$

③ 청산 및 손절 규칙 (Exit / Stop)

Z-score가 청산 임계값 k_{out} 이하로 되돌아오면 포지션을 종료한다.

또한, 절대값이 손절선 k_{stop} 이상으로 확대될 경우에는 즉시 손절하여 손실을 제한한다.

Exit when $|z_t| \leq k_{out}$, Stop when $|z_t| \geq k_{stop}$

④ 보유 및 쿨다운 규칙 (Holding / Cooldown)

실거래 환경을 반영하기 위해 진입 후 최소 보유일을 설정하여, 일정 기간이 경과하기 전에는 청산 신호가 발생해도 포지션을 유지한다. 또한, 포지션 종료 후에는 쿨다운 기간을 두어 일정 기간 동안 동일 페어에 재진입하지 못하도록 하여 과도한 턴오버와 단기 노이즈 매매를 방지한다.

⑤ 포지션 스케일링 및 레버리지 관리

포지션은 시장중립을 유지하도록 β 헤지 비율을 반영해 두 자산의 비중을 조정하며, 전체 포트폴리오 레버리지와 페어당 최대 노출에 상한을 설정해 집중위험을 방지한다.

예를 들어, 페어별 총노출을 c_{pair} 로 제한하고, 전체 레버리지를 L_{max} 이하로 유지하도록 비중을 조정한다.

$$w_Y : w_X = 1 : |\beta|, \sum_i |w_i| \leq L_{max}$$

Realistic Implementation

① 거래비용 + 슬리피지 (turnover × bps)

$$\text{일일 턴오버} : \text{TOV}_t = \sum_i |w_{i,t} - w_{i,t-1}|$$

$$\text{비용(슬리피지 포함)} : \text{Cost}_t = \text{TOV}_t \times (c_{\text{bps}} + s_{\text{bps}}) \times 10^{-4}$$

$$\text{비용 차감 후 포트 수익} : r_t^{\text{port}} = \sum w_{i,t} r_{i,t} - \text{Cost}_t$$

- 매일 포지션 변경 규모(턴오버)에 거래비용(10bps)과 슬리피지(5bps)를
- 곱해 비용을 차감한다.
- 잦은 매매일수록 누적 비용이 커지며, 실제 수익률에 큰 영향을 미친다.
- 진입·청산이 한 번 일어날 때 왕복 약 30bps 수준의 비용이 발생한다.

② 레버리지 캡 + 페어당 캡

$$\text{페어당 캡} : \lambda_p = \min\left(1, \frac{c_{\text{pair}}}{|w_{y,t}^{\text{raw}}| + |w_{x,t}^{\text{raw}}|}\right), \quad (w_{y,t}, w_{x,t}) = \lambda_p (w_{y,t}^{\text{raw}}, w_{x,t}^{\text{raw}})$$

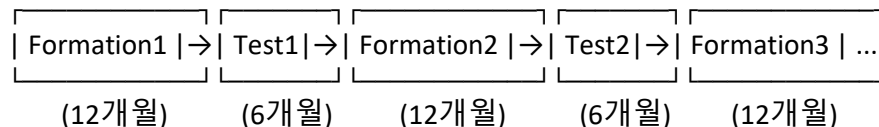
$$\text{포트폴리오 레버리지 캡} : G_t = \sum_i |w_{i,t}|, \quad \lambda_{\text{port}} = \min\left(1, \frac{L_{\text{max}}}{G_t}\right), \quad \tilde{w}_{i,t} = \lambda_{\text{port}} w_{i,t}$$

$$\text{시장중립 헤지} : w_Y : w_X = 1 : |\beta| \quad (\text{총노출} = c_{\text{pair}} \text{ 내})$$

- 페어당 캡은 한 페어에 과도하게 집중되는 것을 방지하기 위해, 각 페어의 총 노출을 일정 비율 이하로 제한한다.
- 포트폴리오 레버리지 캡은 전체 포트폴리오의 총 노출이 설정한 한도(예: 1.0)를 초과하지 않도록 비중을 비례적으로 축소해 관리한다.
- 이러한 제약을 통해 전략이 일부 페어에 쏠리거나 과도하게 레버리지를 사용하는 것을 방지하고, 위험을 통제하며 안정적인 운용이 가능하도록 한다.

③ 워킹-포워드 구조

- 윈도우: 12개월 포메이션 → 6개월 테스트를 연속 롤링
- 시점 분리(정보누출 방지)
 - 페어 선정(공적분·ADF·FDR·Half-life)은 오직 과거 12개월 데이터로만 수행
 - 테스트 구간의 신호는 전일 통계(μ_{t-1} , σ_{t-1})로 계산하고 t일 집행
- 흐름



④ 운영 디테일

- 최소 보유일 설정으로 진입 직후 단기 노이즈에 의한 과도한 매매를 방지한다.
- 쿨다운 기간을 두어 동일 페어에 잦은 재진입을 제한함으로써 턴오버와 비용을 억제한다.
- 신호 빈도와 포지션 조정의 현실적 제약을 반영해, 전략의 실행 가능성과 안정성을 높인다.

Backtest Results

전략 / 지표	Terminal Equity	CAGR	Ann. Mean	Ann. Vol	Sharpe	Max Drawdown
Cointegration + Z-scoring	0.5113	-4.4036%	-4.3587%	5.3705%	-0.8116	-49.2911%

전략의 Terminal Equity가 0.51 수준으로 하락했고 CAGR과 Sharpe가 모두 음(-)의 값을 기록하여, 변동성 대비 수익 창출 능력이 부족하며 장기간 운용할수록 손실이 커지는 전략으로 나타났다. 또한 Max Drawdown이 약 49%까지 발생하여 손실 위험도 매우 높은 편으로 확인되었다.

① 초기 구간 (1961-1963)

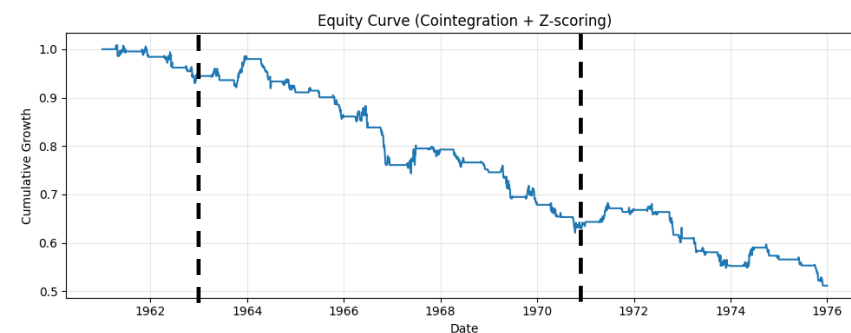
초기에는 누적수익 곡선이 1.0 수준을 대부분 유지하였고 비교적 안정적으로 움직였다. 평균회귀가 어느 정도 작동하여 약한 수익이 나타났지만, 거래비용이 수익을 대부분 상쇄하였다.

② 중반 구간 (1964-1971)

1964년 이후 누적수익이 서서히 하락하기 시작했고 0.8 수준까지 떨어졌다. 평균회귀 신호가 발생했지만 스프레드가 기대한 속도로 회귀하지 않아서 수익이 축적되지 않았다. 시장 환경이 변화하면서 페어 간 장기균형 관계가 약화된 것으로 보이며, 거래비용이 추가적으로 부담을 주었다. 이 구간에서 전략의 성과가 점차 악화되었다.

③ 후반 구간 (1972-1975)

후반부에 들어서 누적수익이 급격히 감소하였고 0.5 수준까지 떨어졌다. 공적분 기반 관계가 더 이상 유지되지 않았고 손실이 지속적으로 확대되었다. 손절 규칙을 적용했음에도 구조적 관계 붕괴를 막지 못했고 전략이 본질적으로 작동하기 어려운 환경이 이어졌다. 결과적으로 성과가 크게 악화되었다.



Statistical Significance

검정 항목	값
Newey–West t-stat	-3.542
표준오차 (SE)	0.000049
Stationary Bootstrap p-value (양측)	0.9853

① Newey–West t-stat = -3.542

이 검정은 평균 수익률이 정확히 0인지를 묻는 방식으로, 시계열 특유의 자기상관과 이분산을 보정하여 계산한다. 이번 결과에서 t-stat이 -3.542로 나타났기 때문에 평균 수익률이 0보다 유의하게 낮은 방향으로 존재한다는 결론을 내릴 수 있다.

② 표준오차 (SE) = 0.000049

이 값은 평균 수익률 추정치의 오차 범위가 어느 정도인지를 보여주는 지표이다. 이번 결과에서 표준오차가 매우 작은 값(0.000049)으로 나타났기 때문에, 평균 수익에 대한 추정이 상대적으로 정확한 추정에 기반하고 있다고 해석할 수 있다.

③ Stationary Bootstrap p-value (two-sided) = 0.9853

이 검정은 시계열 특성을 그대로 유지하면서 데이터를 재표본추출하여, 관측된 평균 수익률이 실제로 수익성이 있다고 말할 수 있는지를 확인한다. 이번 결과의 p-value가 0.9853으로 매우 크게 나타났기 때문에, 관측된 평균 수익률이 0보다 크다고 주장할 수 있는 근거가 전혀 없고, 실질적인 수익성을 기대하기 어렵다고 해석할 수 있다.

Discussion & Implications

전략 부진의 원인 분석

① 공적분 관계 유지 실패

포메이션 구간에서 발견한 장기균형 관계가 테스트 구간에서 유지되지 않았고, 스프레드가 평균으로 회귀하지 않아 전략의 수익 근거가 무너졌다.

② 과도한 거래비용

진입과 청산 신호가 빈번하게 발생하여 수수료와 슬리피지 비용이 누적되면서 수익보다 비용이 더 크게 증가했다.

③ 신호품질 저하

60일 롤링 통계 기반 신호가 단기 변동성 변화에 계속 반응하여 의미 없는 시점에서도 포지션을 설정하고 해제하는 상황이 반복되었다.

④ 불완전한 헤지

B 값이 시장 환경 변화에 따라 달라지는데도 고정적으로 적용하여 스프레드가 시장 방향에 노출되었고, 헤지 효과가 충분하게 나타나지 않았다.

⑤ 느린 재학습 주기

6개월 동안 동일한 페어를 고정하여 사용하여 구조적 변화가 발생해도 대응하지 못했고, 붕괴된 관계를 계속 거래하면서 손실을 확대했다.

⑥ 환경 변화 미반영

평균회귀 가능성이 낮은 구간에서도 동일 기준으로 진입하여 손실을 반복했고, 시장 환경 변화에 따른 신호 유효성 검증 과정이 부족했다.

⑦ 음의 기대수익

Newey-West 검정에서 음(-)의 방향으로 유의한 결과가 나타나 전략이 구조적으로 손실을 유발하는 방향으로 작용했다.

⑧ 양의 알파 부재

Stationary Bootstrap p-value가 매우 커서 수익이 0보다 크다고 말할 근거가 없었고, 수익성이 존재할 가능성이 극단적으로 낮게 나타났다.

전략 개선 방향

① 관계 재평가 강화

공적분 관계가 시간이 지나며 약해질 수 있기 때문에, 포메이션과 테스트 주기를 짧게 조정하고 공적분 추정과 검정을 더 자주 수행하여 관계 붕괴 시점을 빠르게 파악할 필요가 있다.

② 거래비용 절감

진입과 청산 기준을 더 엄격하게 설정하여 불필요한 거래를 줄이고, 슬리피지 제어 기법과 시장충격 최소화 방식 등을 적용하여 비용 부담을 줄일 수 있다.

③ 신호 안정화

Z-score 계산 시 더 긴 롤링 기간 또는 시장 변동성을 반영하는 동적 표준화 방식을 사용하여 단기 노이즈에 과민하게 반응하지 않도록 신호 품질을 높여야 한다.

④ 동적 헤지 적용

시장 환경 변화에 따라 β 를 주기적으로 재추정하고 필요 시 고정 β 대신 시점별 동적 헤지를 적용하여 잔여 시장 노출을 줄인다.

⑤ 적응형 페어 선정

공적분 안정성 지표를 지속적으로 모니터링하여 기준을 충족하지 못하는 페어를 즉시 교체하도록 자동화하여 붕괴된 관계를 장기간 보유하지 않도록 대응한다.

⑥ 레짐 기반 필터링

평균회귀 강도가 약한 시장 국면(트렌드 장세)에서는 신호를 차단하고, 평균회귀 가능성이 높은 환경에서만 전략을 작동시키는 레짐 스위칭 구조를 적용한다.

⑦ 위험 중심 포지션 관리

레버리지 및 노출 한도를 시장 스트레스 수준에 따라 동적으로 조정하여 급격한 손실 확대를 미리 차단하도록 설계한다.