Diversification in Multiple Pairs Trading

Implementation based on work from

Ning, B. (2024). *Quantitative Methods of Statistical Arbitrage.*

Contents

Basic Pairs Trading Framework

For each pair,

1. Construct a spread
2. Estimate the hedge ratio (eg. OU approach)
3. Trade on spread

Part 1. Pairs Construction

Create a single time series that combines two assets so that it can be traded like a single instrument. The spread should be stationary and means reverting, meaning it drifts away from and returns to equilibrium predictably.

1.1. Spreads

A spread is defined as

$$X_t = S_t^1 - B \cdot S_t^2$$

where
- X_t is the portfolio value
- S_t^1, S_t^2 are two highly correlated assets
- B is the hedge ratio

The hedge ratio defines the spread.

A common way to estimate the hedge ratio is to select the value such that the spread resembles an Ornstein-Uhlenbeck process as much as possible.

1.2. OU process

$$\mathrm{d}X_t = \mu(\theta - X_t)\mathrm{d}t + \sigma \mathrm{d}W_t$$

- μ: speed of mean reversion. How quickly deviations appear. Large μ means trades close quickly.
- θ: long term mean of the spread
- σ: volatility. Even if spreads revert, there's noise. High σ means spread is wide & high fluctuation.
- W_t; standard Brownian motion under historical measure $\mathbb{P}$

The OU process is a simple continuous time reverting model, with parameters that are directly relevant to stat-arb related features. By interpreting a spread as an OU process, a lot of critical information is gained.

1.3. Parameter Estimation

Estimate OU parameters to choose the optimal hedge ratio estimator $\hat{B}$.

1. Consider a list of possible hedge ratio values

$$(b_1, ..., b_n)$$

2. For each b_i, find the OU parameters that define the OU model that fits the data most closely

With MLE,

$$(\mu^*(b_i), \theta^*(b_i), \sigma^*(b_i)) = \arg\max_{\mu,\theta,\sigma} l(\mu, \theta, \sigma \mid X_t(b_i))$$

$$l^*(b_i) = l(\mu^*(b_i), \theta^*(b_i), \sigma^*(b_i) \mid X_t(b_i))$$

where $l^*(b_i)$ is the best OU likelihood achieved under b_i.

4. Find the final estimate, $\hat{B}$, by selecting the b_i with the OU model with the highest likelihood

$$\hat{B} = \arg\max_b l^*(b)$$

1.4. MLE Implementation

1. Solve the OU process for discrete time steps.

An OU process has the continuous stochastic differential equation

$$\mathrm{d}X_t = \mu(\theta - X_t)\mathrm{d}t + \sigma \mathrm{d}W_t$$

Spreads are observed at discrete times $t_0, t_1, \ldots, t_n$ with a fixed step $\Delta t = t_i - t_{i-1}$.

Solving the SDE for discrete times,

$$X_i = X_{i-1} \cdot e^{-\mu \Delta t} + \theta\left(1 - e^{-\mu \Delta t}\right) + \varepsilon_i$$

where ε is a normally distributed random variable with a mean of 0 and variance

$$\delta_\varepsilon^2 = \delta^2 \frac{1 - e^{-2\mu \Delta t}}{2\mu}$$

So, given X_{i-1}, the next value X_i is normally distributed with

$$\text{Mean: } m_i(\mu, \theta) = X_{i-1} \cdot e^{-\mu \Delta t} + \theta\left(1 - e^{-\mu \Delta t}\right)$$

$$\text{Variance: } \delta_\varepsilon^2(\mu, \sigma) = \delta^2 \frac{1 - e^{-2\mu \Delta t}}{2\mu}$$

$$X_i \mid (X_{i-1}, \mu, \theta, \sigma) \sim \text{Normal}\left(m_i(\mu, \theta), \delta_\varepsilon^2(\mu, \sigma)\right)$$

2. Under a specific hedge ratio, use MLE to estimate for the OU parameters.

For a fixed hedge ratio b_i, form the spread series

$$x_i = X_{t_i}(b_i)$$

For the past L observations, the probability density of seeing the sequence $x_{t_0 - L + 1}, \ldots, x_{t_0}$ is

$$\prod_{i = t_0 - L + 1}^{t_0} f(x_i \mid x_{i-1}; \mu, \theta, \sigma)$$

Since each step is gaussian, the density for one step is

$$f(x_i \mid x_{i-1}) = \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \cdot \exp\left(-\frac{(x_i - m_i)^2}{2\sigma_\varepsilon^2}\right)$$

Log likelihood:

$$\begin{aligned} l(\mu, \theta, \sigma, b) &= \frac{1}{L} \sum_{i = t_0 - L + 1}^{t_0 - 1} \log f(x_i \mid x_{i-1}; \mu, \theta, \sigma) \\ &= -\frac{1}{2}\ln(2\pi) - \ln(\sigma_\varepsilon) - \frac{1}{2L\sigma_\varepsilon^2} \sum_{i = t - L + 1}^{t_0} \left[x_i - x_{i-1}e^{-\mu \Delta t} - \theta\left(1 - e^{-\mu \Delta t}\right)\right]^2 \end{aligned}$$

Focus on the residual sum:

$$S := \sum_i \left[x_i - x_{i-1}e^{-\mu\Delta t} - \theta\left(1 - e^{-\mu\Delta t}\right)\right]^2$$

Let $\phi = e^{-\mu\Delta t}$ and $c = 1 - \phi$.

$$\begin{aligned} S &= \sum_i \left[x_i - \phi x_{i-1} - \theta c\right]^2 \\ &= \sum_i \left[x_i^2 - 2\phi x_i x_{i-1} - 2\theta c x_i + \phi^2 x_{i-1}^2 + 2\phi\theta c x_{i-1} + \theta^2 c^2\right] \end{aligned}$$

This yields several sufficient statistics.

$$\begin{aligned} X_x &:= \sum x_{i-1} & X_y &:= \sum x_i & & \\ X_{xx} &:= \sum x_{i-1}^2 & X_{yy} &:= \sum x_i^2 & X_{xy} &:= \sum x_i x_{i-1} \end{aligned}$$

The series can be evaluated with these sufficient statistics.

$$l(\mu, \theta, \sigma, b) = -\ln(\sigma_\varepsilon) - \frac{1}{2L\sigma_\varepsilon^2} S\left(\phi, \theta; X_x, X_y, X_{xx}, X_{yy}, X_{xy}\right)$$

Optimal parameters are given by:

$$\begin{aligned} \theta^* &= \frac{X_y X_{xx} - X_x X_{xy}}{n\left(X_{xx} - X_{xy}\right) - \left(X_x^2 - X_x X_y\right)} \\ \mu^* &= -\frac{1}{\Delta t} \ln \frac{X_{xy} - \theta^* X_x - \theta^* X_y + n(\theta^*)^2}{X_{xx} - 2\theta^* X_x + n(\theta^*)^2} \\ (\sigma^*)^2 &= \frac{2\mu^*}{n(c^*)} \cdot \left(X_{yy} - 2\phi^2 X_{xy} + \phi^2 X_{xx} - 2\theta^* c^* \left(X_y - \phi X_x\right) + n(\theta^*)^2 c^{*^2}\right) \end{aligned}$$

3. Find the maximized average log-likelihood.

Maximize the average log-likelihood over (μ, θ, σ) for the specific b_i.

$$l^*(b_i) = l(\mu^*, \theta^*, \sigma^*, b^*)$$

4. Compare the maximized average log-likelihood over all b's.

$$B = \arg\max_b l^*(b)$$

The maximizer B is the hedge ratio that best fits the spread to an OU process.

In his paper, Ning suggests doing a grid search over values of b, computing the maximized log-likelihood function for each.

$$b = \{-2, -1.99, \ldots, 1.99, 2\}$$

Part 2. Trading

Given hedge ratio $\hat{B}$, how to trade the spread $X_t = S_t^1 - \hat{B} \cdot S_t^2$.

When the spread deviates far from its recent mean, bet that it will revert.

Define components & indicators:
- C_t := Capital available at time
- $POS(S_t^i)$:= current position of asset i in the spread
- $MA(X_t)$:= M-day moving average of spread
- $SD(X_t)$:= Standard deviation of spread over past M days
- K := Threshold parameter. How many SDs away from MA are required to enter a trade
- r := Stop-loss parameter. Defines acceptable adverse move from entry price before forcing exit

2.1. Trading Procedure

The spread is treated like a single instrument.

1. Estimate $\hat{B}$ at time 0

Use past L days of S^1 and S^2 to compute

$$\hat{B} = \arg\max_b l^*(b)$$

This $\hat{B}$ is fixed for the upcoming trading period.

2. Compute the spread $X_t = S_t^1 - \hat{B} \cdot S_t^2$ at each t, and track $MA(X_t)$ and $SD(X_t)$

3. Entry & Exit

Entry:

If spread is significantly low, then expect reversion upwards and enter a long position on the spread, which is long on asset 1 and short on asset 2:

$$POS(X_t) = 0 \text{ and } X_t < MA(X_t) - K \cdot SD(X_t)$$
$$\Longrightarrow POS(S_t^1) = \frac{C_t}{S_t^1} \text{ and } POS(S_t^2) = -\hat{\beta} \cdot POS(S_t^1)$$

If spread is significantly high, then expect reversion downwards, and enter a short position on the spread, which is short on asset 1 and long on asset 2:

$$POS(X_t) = 0 \text{ and } X_t > MA(X_t) + K \cdot SD(X_t)$$
$$\Longrightarrow POS(S_t^1) = -\frac{C_t}{S_t^1} \text{ and } POS(S_t^2) = -\hat{\beta} \cdot POS(S_t^1)$$

Exit:

If a long/short spread reverts back to equilibrium, reset positions and update capital:

$$POS(X_t) > 0 \text{ and } X_t > MA(X_t) \text{ (for long position)}$$
$$POS(X_t) < 0 \text{ and } X_t < MA(X_t) \text{ (for short position)}$$
$$\Longrightarrow C_t = C_t + X_t \cdot POS(X_t) \text{ and } POS(X_t, S_t^1, S_t^2) = 0$$

Stop-Loss:

Protects against large deviations in spread before mean reversion occurs.

In a long/short position, when the spread deviates too far below/above entry, cut the loss:

$$\begin{aligned} &POS(X_t) > 0 \quad \text{and} \quad X_t < X_0(1 - r) \text{ (for long position)} \\ &POS(X_t) < 0 \quad \text{and} \quad X_t < X_0(1 + r) \text{ (for short position)} \\ &\Longrightarrow C_t = C_t + X_t \cdot POS(X_t) \quad \text{and } POS(X_t, S_t^1, S_t^2) = 0 \end{aligned}$$

Stay Put:

If no entry or exit condition is met, hold the position:

$$POS(X_t) = POS(X_t) \quad \text{and} \quad C_t = C_t$$