Algorithmic Trading

Chapter 2: Basics of Mean Reversion

Introduction

- “Sports Illustrated Jinx” - what goes up must come down
- Two kinds of Mean Reversion
  - Temporal
  - Cross-sectional
- Mean Reversion
  - Change in price in the next time step being proportional to the difference between the mean price and current price
- Stationarity
  - Prices diffuse slower than a geometric random walk

Let’s discuss:

- **What are some of the theories / explanation behind price reversals in financial markets?**

Augmented Dickey-Fuller (ADF)

- Test for Mean Reversion
- High level idea:
  - Basically relating the price **change** at timestep **t** with the price at previous timestep **t-1**
  - Start with the linear price change model (Eqn. (2.1)). Run a regression to obtain the coefficient **lambda** below (but we denote as **L** from now on) and standard error **SE**

$$\underbrace{\Delta y(t)}_{\textbf{Dep.}} = \lambda \underbrace{y(t-1)}_{\textbf{Indep.}} + \underbrace{\mu + \beta t + \alpha_1 \Delta y(t-1) + \cdots + \alpha_k \Delta y(t-k) + \epsilon_t}_{\textbf{Discard}}$$

  - Test **L/SE** for statistical significance at some confidence threshold, say 95%
- Sanity check: Test statistic **L/SE** should be < 0 since **L** < 0 as mean reverting

Hurst Exponent

- Test for Stationarity (alternately, indicator for MR or Trendiness)
- High level idea:
  - Characterize the diffusion speed of log prices (z) over some arbitrary lag (tau) by:

    $$\mathrm{Var}(\tau) = \langle |z(t+\tau) - z(t)|^2 \rangle \sim \tau^{2H}$$

    where H is the Hurst exponent, and tilde (~) means that the relationship becomes an equality with some proportional constant in the limit.

  - Compare this characterization against the geometric random walk 'benchmark':

    $$\langle |z(t+\tau) - z(t)|^2 \rangle \sim \tau$$

  - If H = 0.5, we obtain tau which means that log prices is geometric random walk
  - If H < 0.5, log prices are mean reverting; if H > 0.5, log prices are trending

Variance Ratio

- Not discussed in detail in the book
- Check for stationarity in the series
- Tests only random walk hypothesis (reject/do not reject) unlike Hurst (trend, random walk, mean reversion)

Interesting:

- **Not commonly used across pipelines (e.g. Quantopian), unlike Hurst or ADF - Possibly a supporting actor and not a main lead?**

Half Life of Mean Reversion I

- Recall the gradient **lambda (L)** from ADF slide. This can be viewed as a gauge of the time needed for mean reversion
- How do we go about it? Start with the linear model of price change:

$$\Delta y(t) = \lambda y(t-1) + \mu + \beta t + \alpha_1 \Delta y(t-1) + \cdots + \alpha_k \Delta y(t-k) + \epsilon_t$$

After some magical hand-waving, we arrive at the realm of stochastic calculus to get ...

OU for MR process: $dy(t) = (\lambda y(t-1) + \mu)dt + d\varepsilon$

… for which the expected price at t i.e. E(y(t)), can then be solved analytically

$$E(\,y(t)) = y_0 exp(\lambda t) - \mu/\lambda(1 - exp(\lambda t))$$

Half Life of Mean Reversion II

$$E(y(t)) = y_0 exp(\lambda t) - \mu/\lambda(1 - exp(\lambda t))$$

- Price decays exponentially to the highlighted term, with the half-life of decay being -log(2)/$\mathbf{L}$
- Recall that $\mathbf{L}$ < 0 as per ADF.
- So what's the point? It links the regression coeff. $\mathbf{L}$ to the half life of mean reversion i.e. useful for trading:
  - If $\mathbf{L}$ > 0, not mean reverting
  - If $\mathbf{L}$ approx. 0, the half life is very long (price series not 'choppy' enough to be profitable)
  - $\mathbf{L}$ provides a natural time scale for many parameters when we put together strategies
    - Setting the lookback equal to (a small multiple of the)half life

Cointegration I

- Process of linearly combining non-stationary (price) series such that resulting portfolio has a stationary (price) series (e.g. Equity L/S)
- **Hedge Ratio** is the key ingredient for building these portfolios
- How do we go about getting the Hedge Ratio?
  - For 2 variables: Use Cointegrated Augmented Dickey-Fuller (CADF)
  - For > 2 variables: Use Johansen Test

Let's discuss:

- **Apart from Equity L/S, what are some other interesting trades using cointegration?**

Cointegration II - Hedge Ratio

- Gives the ratio for which different variables should be combined in a portfolio
- High level idea (2 asset example using CADF):
    - Use 2 assets, say, A and B, with A as the independent variable
    - Run a regression - coefficient gives the Hedge Ratio i.e. how much to combine B with A. Let's call this HR1
    - Reverse the roles of A and B in the regression and obtain HR2 **(Important step: Only one of either HR1 or HR2 will give the stationary portfolio)**
    - Use the CADF to decide on HR1 or HR2 based on some confidence threshold
- For Johansen Test - we use the eigenvector with the largest eigenvalue

Let's discuss:

- **What's the difference between spurious regression and cointegration?**

Pros and Cons of Mean Reversion

- Pros:
  - Vast array of choices to construct portfolios
  - Span a great variety of time scales (different types of traders, strategies)
- Cons:
  - Risk management difficult - hard to implement stop losses since mean itself is a moving target

Key Takeaways

- Definitions for MR, Stationarity, Cointegration
- Various tests for MR (ADF), Stationarity (H, VR) and Cointegration (CADF, J)
- Ideas underlying MR and cointegration is remarkably simple, but the devil is in the details
  - Ordering in the CADF tests, etc.

Let's discuss:

- **How can we enhance MR-based strategies?**
- **What are some alternatives to the methods used in the chapter?**
- **What has each of you gotten out of the chapter?**

Thank you!

Useful Resources

https://www.quantopian.com/posts/enhancing-short-term-mean-reversion-strategies-1

https://www.quantopian.com/lectures/integration-cointegration-and-stationarity