Pairs Trading Optimization using Q learning: A Cointegration Approach

ORIE 6590
Chaoping Deng
December 8, 2017

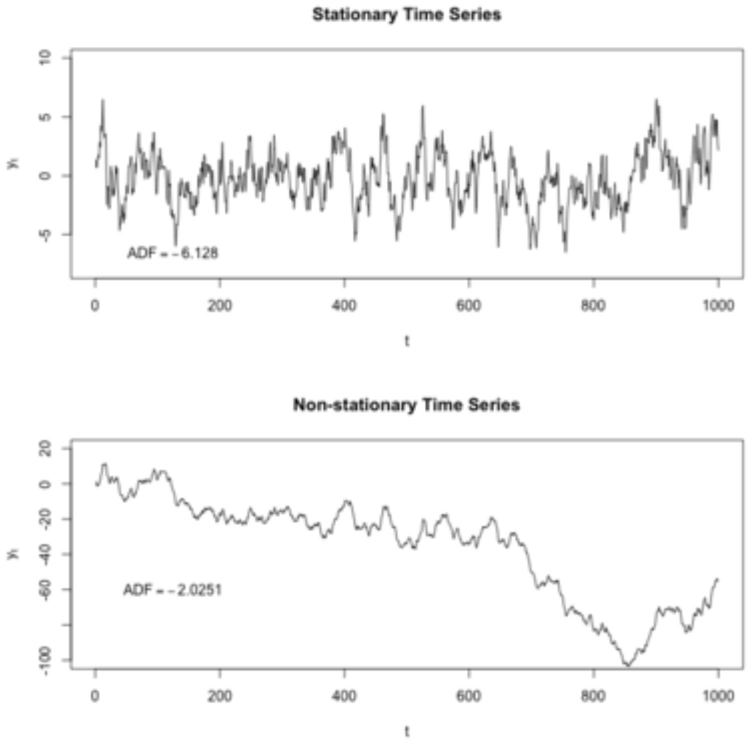
Stationarity

- A time series X(t) is weakly stationary if it satisfies:

$$\mathbb{E}[x(t)] = m_x(t) = m_x(t+\tau) \ \text{ for all } \tau \in \mathbb{R}$$

$$\mathbb{E}[(x(t_1) - m_x(t_1))(x(t_2) - m_x(t_2))] = C_x(t_1, t_2)$$

- Implication:
  - A stationary process never drift too far away from mean m
  - The spread of a stationary process is limited by its autocovariance



Stationary Process vs Non-Stationary Proces. *Wikipedia*.

AR(1) Process

- AR(1) process follows:

$$X_t = \alpha + \phi X_{t-1} + \epsilon_t$$
$$\epsilon_t = N(0, \sigma^2)$$

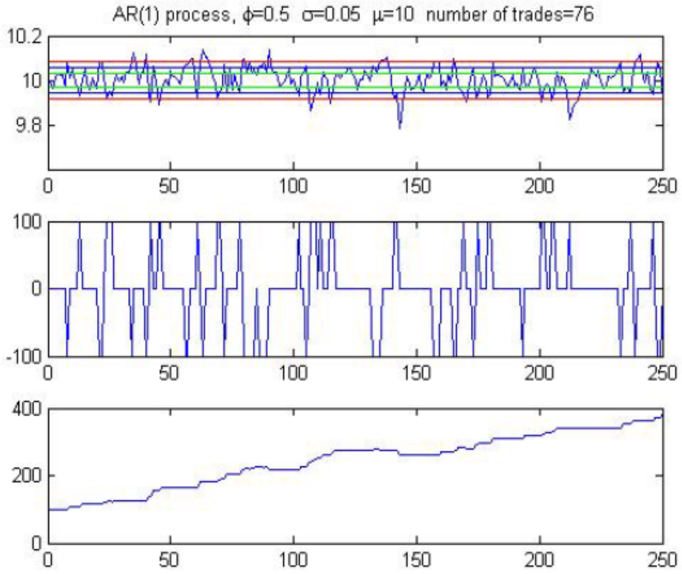
- Characteristics
  - Mean: $E(X_t) = \mu$
  - Variance: $Var(X_t) = \frac{\sigma^2}{(1-\phi^2)}$
  - Covariance: $Cov(X_t, X_{t+h}) = \phi^h \frac{\sigma^2}{(1-\phi^2)}$
  - Rate of Change: $E(X_t|X_0 = x) = \mu + \phi^\tau (x - \mu)$

- General Autoregressive Model:

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t$$

Mean Reversion

- Prices have a tendency to "mean-revert“ Engle (1987), Engle and Granger (1993)
- How To Trade an asset following AR(1):
  - s = $\left|\frac{X_t-\mu}{\sigma_X}\right|$ s-score
  - Trade Entry:
    1. If $\frac{X_t-\mu}{\sigma_X}$ < -1(lower threshold), long asset;
    2. if $\frac{X_t-\mu}{\sigma_X}$ > 1(upper threshold), short asset
  - Trade Exit:
    1. If s < 0.5 closing
    2. If s > 1.5 stop loss
- Parameters to adjust: Estimation Window, Threshold, stop-loss level and trading window
- Prices are rarely stationary



Trade Execution on Simulated AR(1) process
Top to Bottom: Prices, Positions, Wealth

Pair Trading

- Objective: Many economic time series are non-stationary with an ascending or descending random trend, it is possible that a linear combination of these time series will always be stationary in a long run. Our objective is to identify two assets such that their linear combination produces a mean-reverting spread and thus they are identified as being co-integrated.
- For multivariate scenario: cointegration is defined as

$$\begin{aligned} e_t &= \beta_C' Y_t \qquad e_t \sim I(0) \\ &= \pm\beta_1 y_{1,t} \pm \beta_2 y_{2,t} \pm \cdots \pm \beta_n y_{n,t} \end{aligned}$$

- Hypothesis:
  - Given two asset X and Y, y = aX + bY can be a stationary process with long term mean relative mispricing persists. The “true long-term price” of combination of two assets X and Y being m, price difference m-y(t)
  - Pricing inefficiencies are identifiable with statistical models
- Cointegration exists where there is a correction of error in y from the equilibrium, and is therefore not a forecasting problem.

Cointegration

- Define two Time Series Xt and Yt

$$X_t = \gamma_o + \gamma_1 Z_t + \epsilon_t \sim I(1)$$
$$Y_t = \delta_o + \delta_1 Z_t + \eta_t \sim I(1)$$
$$Z_t \sim I(1)$$
$$\epsilon_t, \eta_t \sim I(0)$$

- $\delta_1 X_t - \gamma_1 Y_t (0)$ is stationary
- Cointegration assumes common I(1) processes shared (Ψ)

$$X_{1,t} = \alpha_1 + \gamma_1 Z_{1,t} + \gamma_2 Z_{2,t} + \cdots + \gamma_p Z_{p,t} + \epsilon_{1,t}$$
$$X_{2,t} = \alpha_2 + \phi_1 Z_{1,t} + \phi_2 Z_{2,t} + \cdots + \phi_p Z_{p,t} + \epsilon_{2,t}$$
$$\ldots$$
$$X_{m,t} = \alpha_m + \psi_1 Z_{1,t} + \psi_2 Z_{2,t} + \cdots + \psi_p Z_{p,t} + \epsilon_{m,t}$$

- Defines k = m - p as number of independent linear combinations
  - p = m: k =0, no cointegration
  - 0<p<m: cointegrated
  - p=0: all time series stationary

Cointegration: Continued

- Obtain residuals e from the long-run (equilibrium) relationship
  - $\hat{e} = y_t - bx_t - a$ (OLS regression)test it for the unit root.
  - If unit root does not exist, then process e is stationary.

- Plug e into ECM model to obtain β_1 and α

$$\Delta y_t = \beta_1 \Delta x_t - (1 - \alpha) e_{t-1} + \epsilon_t$$

Cointegration: Continued

- Start with VAR model:

$$y_t = a_{11} y_{t-1} + a_{12} z_{t-1} + \varepsilon_{yt}$$
$$z_t = a_{21} y_{t-1} + a_{12} z_{t-1} + \varepsilon_{zt}$$

- End With ECM

$$\Delta y_t = a_y \left(y_{t-1} - \beta z_{t-1} \right) + \varepsilon_{yt}$$
$$\Delta z_t = a_z \left(y_{t-1} - \beta z_{t-1} \right) + \varepsilon_{zt}$$

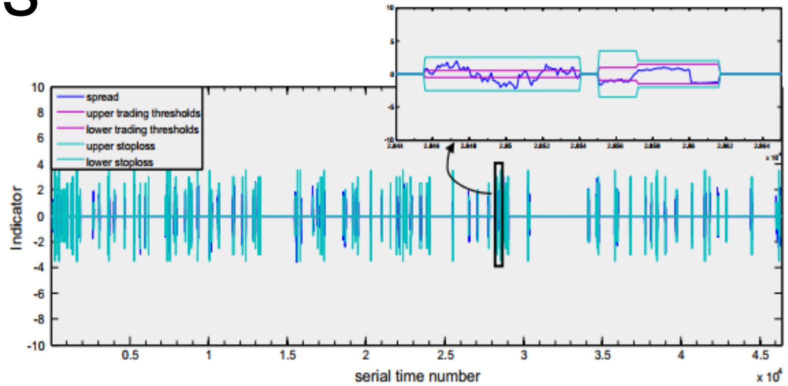
- Long term mean

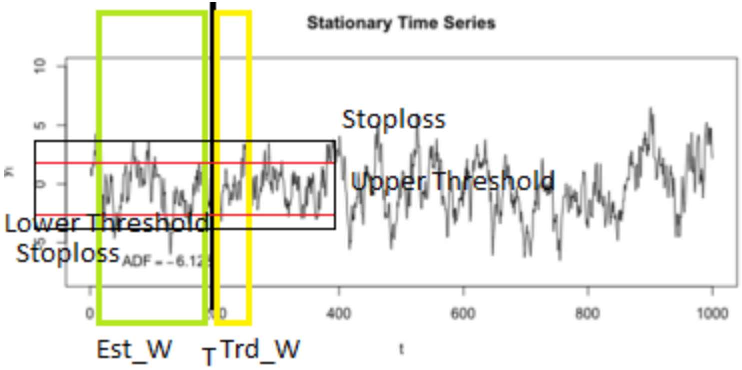
$$A_{t-1} - \beta B_{t-1} + \rho = \mu + \varepsilon_t$$

- Johansen Procedure (For multivariate scenario)
  - Test the rank of Cointegration Matrix (Johansen Test) and obtain β
  - Use cointegration vector from 2 to estimate other VECM parameters (speed of correction, intercept)

Q-learning: Decision Variables

- Estimation Window:
  - run cointegration test
  - re-estimate spread equations
- Trading Window:
  - Parameters from Estimation Window still apply
- Trading Threshold:
  - Threshold above and below spread
- Stop-loss:
  - Wider window than trading threshold to prevent loss





Q-learning: Objective Value

- Sortino Ratio
  - $S = \frac{R - T}{DR}$, $DR = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\mathrm{Min}(0, X_i - T))^2}$
  - Measure the return under downward volatilities.
  - Reward upward volatility but discourage downward volatility
  - Alternatives include: pure return, Sharpe ratio

Q-learning: Model

- Multi-Arm Bandit Model with each arm
  - A(t) $= (T_{est}, T_{trading}, Threshold, Stoploss)$
- SARSA

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_{t+1} + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)]$$

Initialize $Q(s, a)$, $\forall s \in \mathcal{S}$, $a \in \mathcal{A}(s)$, arbitrarily, and $Q(terminal - state, .) = 0$
Repeat (for each episode):

Initialize S
Choose A from S using policy derived from $Q(e.g., \epsilon - greedy)$
Repeat (for each step of episode):
Take action A, observe R, S'
Choose A' from S' using policy derived from $Q(e.g., \epsilon - greedy)$
$Q(S, A) \leftarrow Q(S, A) + \alpha[R + \gamma Q(S', A') - Q(S, A)]$
$S \leftarrow S'; A \leftarrow A'$;
until S is terminal

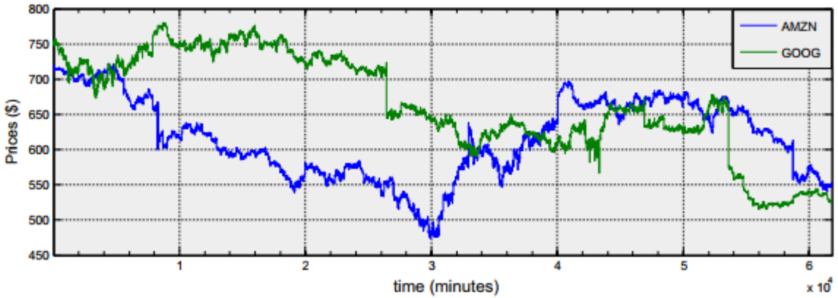
Input: *Stock A, Stock B*

1. *Initialization:*
   a. *action space*$_{n\times 4} \leftarrow [\delta_{n\times 1}$, *stop-loss*$_{n\times 1}$, *estimation window*$_{n\times 1}$, *trading window*$_{n\times 1}$]
   b. *number of iteration* $\leftarrow$ *max number of iteration*
2. **For** *episode* $\leftarrow$ 1 *to number of iteration*
3. *action* $\leftarrow$ *choose action based on the epsilon greedy policy*
4. *reward* $\leftarrow$ *perform action (action, A, B)*
5. *action update (action, reward)*
6. **End**
7. **Return** *Sortino ratio*

Output: *Sortino Ratio*

Simulation

- Data
  - 60000 mins of intraday min-level price data of Amazon and Google
- Decision Space
  - Estimation Window is from 60 to 600 minutes with increment of 5 mins
  - Trading Window is from 5 to 120 minutes with increment of 5 mins
  - Trading thresholds between 0 and 3 with increment of 0.5
  - Stop loss varies between 0 and 5 with increment of 0.5
- Parameter
  - 75% in-sample data and 25% out of sample data, Q-agent set $\alpha = 1, \epsilon = 1$ for training phase and $\alpha = 0.3, \epsilon = 0.3$ for testing



Performance

- Q-learning vs Static Optimization

	CPM		RLM	
Test and estimation window size (min)	380		Dynamic [60,600]	
Trading window size (min)	80		Dynamic [5,120]	
Δ (standard deviation)	1.5		Dynamic [0.5,3]	
Stop-loss (standard deviation)	3		Dynamic [1,5]	
	In-sample	Out-sample	In-sample	Out-sample
Return (%)	15.8	11.3	91.5	46.6
Downside volatility	0.51	0.18	0.61	0.12
Annualized return (%)	31.6	67.8	183	279.6
Sortino ratio	0.27	0.52	1.46	3.72
Number of trades	265	180	586	388
Average return per trade (%)	0.06	0.06	0.16	0.12
Average time per trade (min)	185.4	91	83.8	42.2

Appendix: Johansen Test

- STEP 1:
- Regression

$$\Delta Y_t = \sum_{k=1}^{p-1} \Gamma_k \Delta Y_{t-k} + R_{0t}$$

$$Y_{t-1} = \sum_{k=1}^{p-1} \Gamma_k \Delta Y_{t-k} + R_{kt}$$

$$\Delta Y_t = \Pi Y_{t-p} + \sum_{k=1}^{p-1} \Gamma_k \Delta Y_{t-k} + \epsilon_t$$

$$\Pi = I - \sum_{k=1}^{p} A_k, \quad \Gamma_k = \sum_{i=1}^{k} A_i - I.$$

- Concentrated Model

$$R_{0t} = \alpha\beta' R_{kt} + u_t$$

$$\Delta\widetilde{Y}_t = \alpha\beta' \widetilde{Y}_t + u_t$$

$$S_{ij} = \frac{1}{T} \sum_{t=1}^{T} R_{it} R'_{jt}, \quad i, j = [0, k]$$

Appendix: Johansen Test

- STEP 2:

$$\mathbf{S_{00}} = \frac{1}{T}\mathbf{R_{0t}R'_{0t}} \quad \mathbf{S_{kk}} = \frac{1}{T}\mathbf{R_{kt}R'_{kt}} \quad \mathbf{S_{k0}} = \frac{1}{T}\mathbf{R_{kt}R'_{0t}} \quad \mathbf{S_{0k}} = \frac{1}{T}\mathbf{R_{0t}R'_{kt}}$$

- Solve

$$\begin{aligned}
(S_{k0}S_{00}^{-1}S_{0k} - \lambda S_{kk})\beta &= 0 \\
|\lambda S_{kk} - S_{k0}S_{00}^{-1}S_{0k}| &= 0 \\
S_{kk}^{-1}S_{k0}S_{00}^{-1}S_{0k}\,\beta_i &= \lambda_i\beta_i
\end{aligned}$$

Appendix: Johansen Test

- STEP 3:

$$\lambda_{trace} = -T \sum_{i=r+1}^{n} \log(1 - \lambda_i)$$

$$\lambda_{max} = -T \log(1 - \lambda_{r+1})$$

Appendix: Johansen Test

- Trace Test
- H0: K = K0
- H0: K > K0
- Set K0 =0 representing no cointegration and test whether it can be rejected so that there is at least one cointegration relationship

- Maximum Eigenvalue Test
- H0: K = K0
- H0: K = K0+1
- Set K0 = 0,rejecting null hypothesis indicates there is 1 combination of stationary possible that leads to a stationary, less power than trace test. Increment K value.

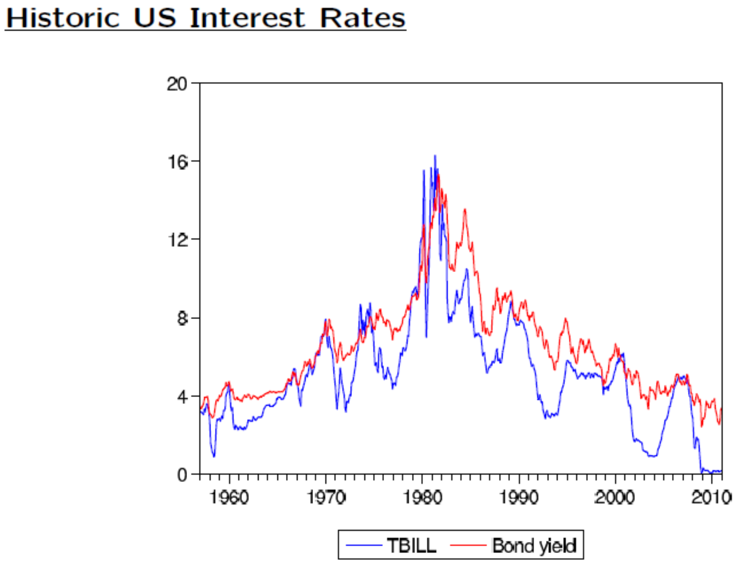
Appendix: Johansen Test

- Step 4

$$\Delta Y_t = \hat{\alpha}\widetilde{\beta}' Y_{t-1} + \sum_{k=1}^{p-1} \Gamma_k \Delta Y_{t-k} + \epsilon_t$$

$$\widehat{Y}_{T+h} = \widehat{\Pi}^h Y_T + \sum_{i=0}^{h-1} \widehat{\Pi}^i C$$

Appendix: Johansen Test



Appendix: Johansen Test

- Johansen Test

```
Sample (adjusted): 1957M04 2011M01
Included observations: 646 after adjustments
Trend assumption: No deterministic trend (restricted constant)
Series: Y R
Lags interval (in first differences): 1 to 2
```

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None	0.033251	24.33360	20.26184	0.0130
At most 1	0.003845	2.488511	9.164546	0.6798

No. of CE(s)	Eigenvalue	Statistic	Critical Value	Prob.**
None	0.033251	21.84509	15.89210	0.0051
At most 1	0.003845	2.488511	9.164546	0.6798

Appendix: Johansen Test

- Fit VCEM Model

$$\Delta y_t = 0.38\Delta y_{t-1} - 0.08\Delta r_{t-1} - 0.016e_{t-1} + \epsilon_t$$

$$\Delta r_t = 0.53\Delta y_{t-1} - 0.14\Delta r_{t-1} - 0.047e_{t-1} + \epsilon_t$$

```
=============================================
 Error Correction:       D(Y)        D(R)
=============================================
     CointEq1         -0.016142    0.047416
                      (0.00861)   (0.01450)
                     [-1.87572]  [ 3.26980]
     D(Y(-1))          0.380541    0.526983
                      (0.04554)   (0.07673)
                     [ 8.35696]  [ 6.86787]
     D(R(-1))         -0.076016    0.137130
                      (0.02650)   (0.04465)
                     [-2.86868]  [ 3.07103]
=============================================
```

Appendix: Johansen Test

- Cointegration

$$e_{t-1} = (y_{t-1} - 0.94\, r_{t-1} - 0.0159)$$

```
Sample (adjusted): 1957M03 2011M01
Included observations: 647 after adjustments
Standard errors in ( ) & t-statistics in [ ]
=============================================
 Cointegrating Eq:      CointEq1
=============================================
        Y(-1)           1.000000
        R(-1)          -0.940823
                        (0.07272)
                       [-12.9369]
          C            -0.015868
                        (0.00423)
                       [-3.75153]
=============================================
```