Algorithmic Trading: Statistical Arbitrage
PIMS Summer School

Sebastian Jaimungal, U. Toronto

Álvaro Cartea,U. Oxford

many thanks to

José Penalva,(U. Carlos III)
Luhui Gan (U. Toronto)
Ryan Donnelly (Swiss Finance Institute, EPFL)
Damir Kinzebulatov (U. Laval)
Jason Ricci (Morgan Stanley)

July, 2016

Cointegration

Cointegration

- A process is said to be **stationary** if the **unconditional distribution is constant** (in time). For example,
  - Random walk

    $$y_t = y_{t-1} + \varepsilon_t \,, \text{ with } \varepsilon_t \sim \mathcal{N}(0, \sigma^2)$$

    is **non-stationary**
  - Auto-regressive of order 1: AR(1)

    $$y_t = a + b\, y_{t-1} + \varepsilon_t \,, \text{ with } b < 1, \text{ and } \varepsilon_t \sim \mathcal{N}(0, \sigma^2)$$

    is **stationary**

Cointegration

- A time-series is said to be **integrated of order d (i.e., $I(d)$)** if the **d-times difference is stationary**. For example,
  - Random walk

    $$y_t = y_{t-1} + \varepsilon_t\,,\ \text{with } \varepsilon_t \sim \mathcal{N}(0, sigma^2)$$

    is **I(1)**, since

    $$\mathcal{D}y_t := y_t - y_{t-1} = \varepsilon_t$$

    has a stationary distribution
  - AR(1) is **I(0)**
- Most Economic models are $I(0)$ or $I(1)$

Cointegration

- It is often the case that two (or more) time-series appear to be **non-stationary**, but a **linear combination is stationary**
- If $\boldsymbol{y}_t$ is a vector-valued process that **is** $I(d)$ and there exists a vector $\boldsymbol{b}$ such that $\boldsymbol{b}'\boldsymbol{y}_t$ **is** $I(d^*)$ with $d^* < d$, then $\boldsymbol{y}_t$ is said to be **cointegrated** and $\boldsymbol{b}$ is the **cointegrating vector**.
- Most of the time $d = 1$ and $d^* = 0$

Cointegration

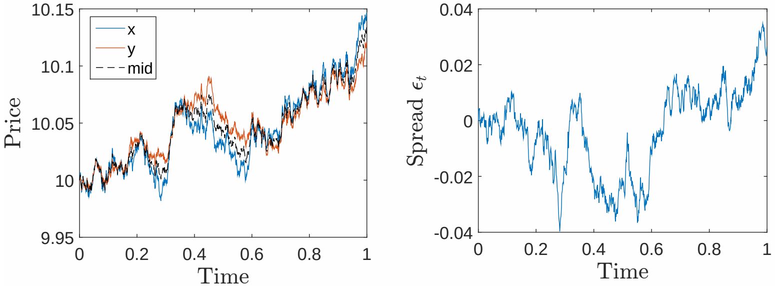
- For example, two price processes x and y are given by

$$x_t = (x_0 - \tfrac{1}{2}\varepsilon_0) + \sigma\, W_t + \tfrac{1}{2}\varepsilon_t,$$
$$y_t = (y_0 - \tfrac{1}{2}\varepsilon_0) + \sigma\, W_t - \tfrac{1}{2}\varepsilon_t,$$

where

$$d\varepsilon_t = \kappa\,(\theta - \varepsilon_t)\,dt + \eta\,dB_t\,,$$

and W and B are correlated Brownian motions.



Cointegration

- How to estimate from data? First need a model...
- A pair of prices $\mathbf{S} = (S_t^1, S_t^2)_{t \geq 0}$ satisfies the continuous analog of a **Vector Auto-Regressive** (VAR(1)) model

$$d\mathbf{S}_t = \boldsymbol{\kappa}(\boldsymbol{\theta} - \mathbf{S}_t)\, dt + \boldsymbol{\sigma}\, d\mathbf{W}_t$$

  where
  - $\boldsymbol{\kappa}$ is a positive semi-definite 2×2 matrix,
  - $\boldsymbol{\theta}$ is a 2×1 vector,
  - $\boldsymbol{\sigma}$ is a 2×2 matrix, equal to the **Cholesky decomposition** of $\boldsymbol{\Sigma}$, and
  - $\mathbf{W}$ is a 2×1 **independent Brownian motion**.

Cointegration

- **Diagonalize** $\boldsymbol{\kappa}$, so that $\boldsymbol{\kappa} = \boldsymbol{U}\,\tilde{\boldsymbol{\kappa}}\,\boldsymbol{U}^{-1}$ where $\boldsymbol{U}$ is the **matrix of eigenvectors** of $\boldsymbol{\kappa}$, and $\tilde{\boldsymbol{\kappa}}$ is a diagonal matrix.
- Then, $\tilde{\mathbf{S}}_t = \boldsymbol{U}^{-1}\mathbf{S}_t$ will satisfy decoupled SDEs

$$
\begin{aligned}
d\tilde{\mathbf{S}}_{t,1} &= \tilde{\boldsymbol{\kappa}}_{t,1}(\tilde{\boldsymbol{\theta}}_1 - \tilde{\mathbf{S}}_{t,1})\,dt + (\tilde{\sigma}\,d\mathbf{W}_t)_1 \\
d\tilde{\mathbf{S}}_{t,2} &= \tilde{\boldsymbol{\kappa}}_{t,2}(\tilde{\boldsymbol{\theta}}_2 - \tilde{\mathbf{S}}_{t,2})\,dt + (\tilde{\sigma}\,d\mathbf{W}_t)_2
\end{aligned}
$$

- these are the **cointegrating factors**

Cointegration

- With this model, we can estimate from data by regressing

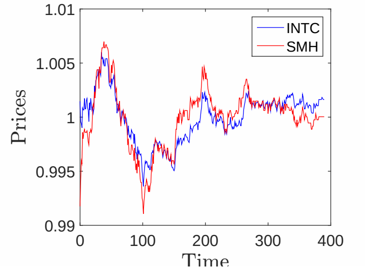
$$\mathbf{S}_{n+1} = \mathbf{A} + \mathbf{B}\,\mathbf{S}_n + \varepsilon_n$$

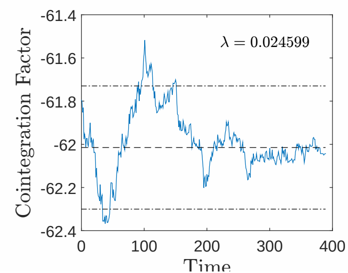
Then,

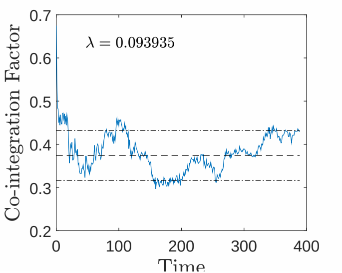
$$\widehat{\boldsymbol{\kappa}} = \frac{1}{\Delta t}(\mathbb{I} - \widehat{\mathbf{B}}),$$
$$\widehat{\boldsymbol{\theta}} = (\mathbb{I} - \widehat{\mathbf{B}})^{-1}\,\widehat{\mathbf{A}}$$

Cointegration

- Using INTC and SMH prices at 1-minute intervals







Pairs Trading

Naive Pairs Trading

- **Pairs trading** assumes that two assets are **cointegrated** and often are behave as a vector autoregressive (**VAR**) model

$$\Delta \boldsymbol{S}_t = \boldsymbol{A} + \boldsymbol{B}\,\boldsymbol{S}_{t-1} + \varepsilon_t\,,$$

  ε_t are iid bivariate normal with mean zero.

- It can be seen as a discrete version of the continuous time model

$$d\boldsymbol{S}_t = \boldsymbol{\kappa}\,(\boldsymbol{\theta} - \boldsymbol{S}_t)\,dt + \boldsymbol{\sigma}\,d\boldsymbol{W}_t,$$

- To estimate the model, regress the vector of price changes on the price at the interval start.

- The eigenvector with the largest eigenvalue represents the cointegration factor that you trade on:

$$\zeta_t = \alpha\,S_t^{(1)} + \beta\,S_t^{(2)} \quad \text{and} \quad d\zeta_t = \kappa_\zeta(\theta_\zeta - \zeta_t)\,dt + \sigma_\zeta\,dW_t^\zeta\,.$$

Naive Pairs Trading

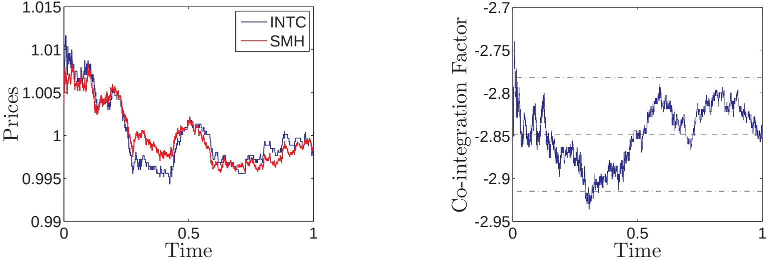


Figure: INTC and SMH on November 1, 2013: (left) midprice relative to mean midprice; (right) co-integration factor. The x-axis is time in terms of fractions of the trading day. The dashed line indicates the mean-reverting level; the dash-dotted lines indicate the 2 standard deviation bands.

Naive Pairs Trading

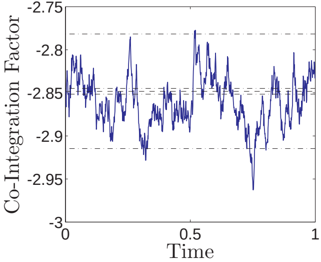
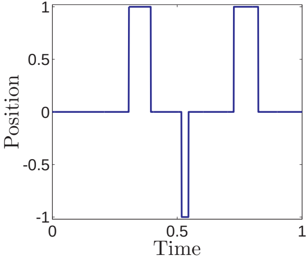
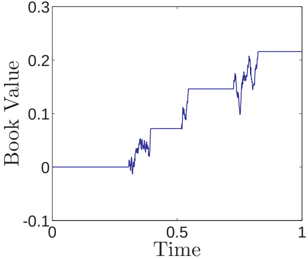


Figure: Traders often use **ad hoc bands** to decide when to enter and exit a long/short position in the cointegration factor... A sample path of the co-integration factor, the trading position, and the book value of the trade, using the two standard deviation banded strategy.





Naive Pairs Trading

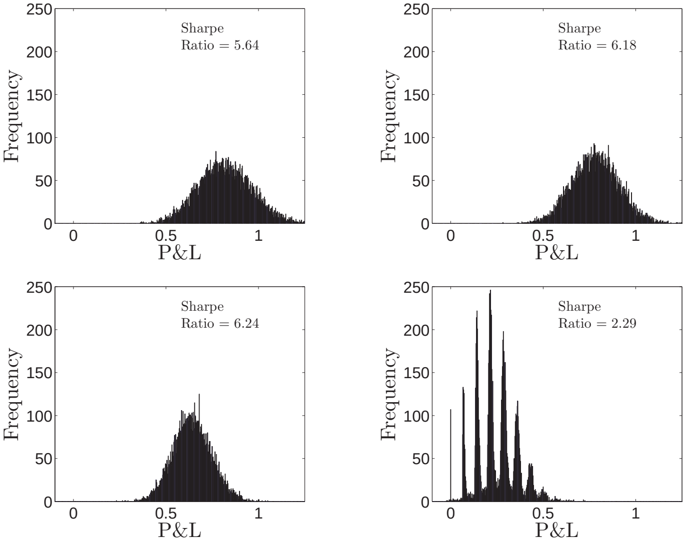


Figure: P&L histograms from 10,000 scenarios using the naive strategy with various trigger bands.

Pairs Trading: Optimal Band Selection

- It is possible to formulate an optimal band selection problem
- Consider the performance criteria for **exiting** a **long/short position**...

$$H_+^{(\tau)}(t,\varepsilon) = \mathbb{E}_{t,\varepsilon}\left[e^{-\rho(\tau-t)}\,(\varepsilon_\tau - c)\right]\,,$$
$$H_-^{(\tau)}(t,\varepsilon) = \mathbb{E}_{t,\varepsilon}\left[e^{-\rho(\tau-t)}\,(-\varepsilon_\tau - c)\right]\,,$$

- and consider the performance criteria for **entering** a **long/short position**...

$$\begin{aligned} G^{(\boldsymbol{\tau})}(t,\varepsilon) = \mathbb{E}_{t,\varepsilon}\Big[\quad & e^{-\rho(\tau_+-t)}\,(H_+(\tau_+,\varepsilon_{\tau_+}) - \varepsilon_{\tau_+} - c)\;\mathbb{1}_{\min(\tau_+,\tau_-)=\tau_+} \\ & + e^{-\rho(\tau_--t)}\,(H_-(\tau_-,\varepsilon_{\tau_-}) + \varepsilon_{\tau_-} - c)\;\mathbb{1}_{\min(\tau_+,\tau_-)=\tau_-}\Big]\,. \end{aligned}$$

Pairs Trading: Optimal Band Selection

- Variational inequality (**VI**) for **optimal exiting**
  - a **long position**

$$\max\left\{(\mathcal{L}-\rho)\,H_+(\varepsilon)\;;\;(\varepsilon-c)-H_+(\varepsilon)\right\}=0$$

  - **short position**

$$\max\left\{(\mathcal{L}-\rho)H_-(\varepsilon)\;;\;(-\varepsilon-c)-H_-(\varepsilon)\right\}=0\,.$$

- VI for **optimal entry** is

$$\max\Big\{(\mathcal{L}-\rho)\,G(\varepsilon)\;;\\ (H_+(\varepsilon)-\varepsilon-c)-G(t,\varepsilon)\\ (H_-(\varepsilon)+\varepsilon-c)-G(t,\varepsilon)\Big\}=0\,.$$

Pairs Trading: Optimal Band Selection

- Two **fundamental solutions** to $(\mathcal{L} - \rho)F = 0$ are

$$F_+(\varepsilon) = \int_0^\infty u^{\frac{\rho}{\kappa}-1}\, e^{-\sqrt{\frac{2\kappa}{\sigma^2}}\,(\theta-\varepsilon)\,u-\frac{1}{2}\,u^2}\, du\,,$$

$$F_-(\varepsilon) = \int_0^\infty u^{\frac{\rho}{\kappa}-1}\, e^{+\sqrt{\frac{2\kappa}{\sigma^2}}\,(\theta-\varepsilon)\,u-\frac{1}{2}\,u^2}\, du\,.$$

Pairs Trading: Optimal Band Selection

- H_+ and H_1 admit the solution

$$H_+(\varepsilon) = A\,F_+(\varepsilon)\,\mathbb{1}_{\varepsilon<\varepsilon^*} + (\varepsilon - c)\,\mathbb{1}_{\varepsilon\geq\varepsilon^*}\,,$$

$$H_-(\varepsilon) = A\,F_-(\varepsilon)\,\mathbb{1}_{\varepsilon>\varepsilon^*_-} - (\varepsilon + c)\,\mathbb{1}_{\varepsilon\leq\varepsilon^*_-}\,,$$

- G admits the solution

$$\begin{aligned} G(\varepsilon) = &\ (A\,F_+(\varepsilon) + B\,F_-(\varepsilon))\,\mathbb{1}_{\varepsilon\in(\varepsilon_{*+},\varepsilon_{*-})} \\ &+ (H_+(\varepsilon) - \varepsilon - c)\ \mathbb{1}_{\varepsilon\leq\varepsilon_{*+}} + (H_-(\varepsilon) + \varepsilon - c)\ \mathbb{1}_{\varepsilon\geq\varepsilon_{*-}}\,. \end{aligned}$$

Pairs Trading: Optimal Band Selection

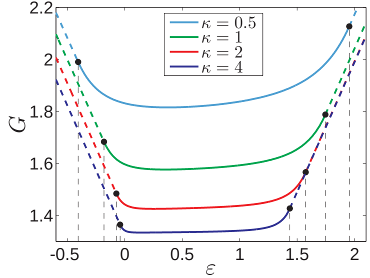
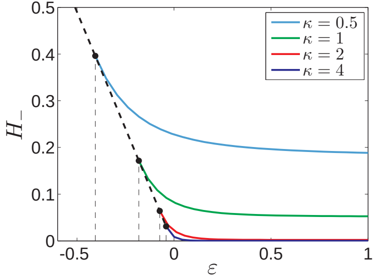
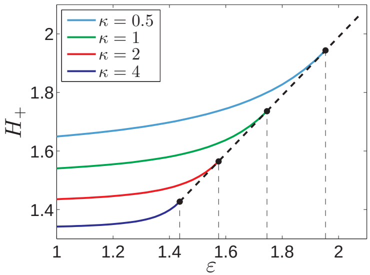


Figure: The optimal entry trigger level and corresponding value function for the double entry-exit problem.





Pairs Trading: Multiple Assets

- Both of the previous approaches hardwire the portfolio... what about dynamically changing the positions?
- A model with **short-term alpha** in log prices

$$dY_t^k = Y_t^k \left(\delta_k \, \alpha_t \, dt + \sum_{i=1}^{n} \sigma_{ki} \, dW_t^i \right) ,$$

where $\alpha_t = a_0 + \sum_{i=1}^n a_i \log Y_t^i$.

- Interestingly, this model can be shown to be a **cointegration model of log-prices**
- We pose the trading problem as a portfolio optimization one and seek to maximize the performance criteria

$$H^{\pi}(t, x, \boldsymbol{y}) = \mathbb{E}_{t,x,\boldsymbol{y}} \left[- \exp\left(-\gamma \, X_T^{\pi} \right) \right] ,$$

Pairs Trading: Multiple Assets

- The value function, after using the feedback control, solves the non-linear PDE

$$\partial_t H + \alpha\,\boldsymbol{\delta}'\boldsymbol{\mathcal{D}}_y H + \tfrac{1}{2}\boldsymbol{\mathcal{D}}^{\boldsymbol{\Omega}}_{yy} H - \frac{\mathscr{L}' H\,\boldsymbol{\Omega}^{-1}\mathscr{L} H}{2\,\partial_{xx} H} = 0\,.$$

- Value function admits the ansatz

$$H(t,x,\boldsymbol{y}) = -\exp\left\{-\gamma\left(x + h\left(t,\ a_0 + \textstyle\sum_{i=1}^{n} a_i\,\log y^i\right)\right)\right\}$$

and

$$\partial_t h - \tfrac{1}{2}\mathrm{Tr}(\boldsymbol{A}\,\boldsymbol{\Omega})\,\partial_\alpha h + \tfrac{1}{2}(\boldsymbol{a}'\boldsymbol{\Omega}\boldsymbol{a})\,\partial_{\alpha\alpha} h + \frac{\boldsymbol{\delta}'\boldsymbol{\Omega}\boldsymbol{\delta}}{2\gamma}\,\alpha^2 = 0\,,$$

with the feedback control

$$\boldsymbol{\pi}^* = \tfrac{1}{\gamma}\,(\boldsymbol{\Omega}^{-1}\boldsymbol{\delta})\,\alpha - \boldsymbol{a}\,\partial_\alpha h\,.$$

Pairs Trading: Multiple Assets

- The function h can be solved exactly and leads to

$$h(t,\alpha) = \mathbb{E}^*_{t,\alpha}\left[\frac{\boldsymbol{\delta}'\boldsymbol{\Omega}\boldsymbol{\delta}}{2\gamma}\int_t^T \alpha_s^2\,ds\right],$$

the measure $\mathbb{P}^*$ is the one which renders Y_t $\mathbb{P}^*$-martingales

- Can also show that the relationship

$$\sup_{\boldsymbol{\pi}\in\mathcal{A}} \mathbb{E}_{t,x,\boldsymbol{y}}\left[-\exp\left(-\gamma\,X_T^{\boldsymbol{\pi}}\right)\right] = -\exp\left(-\gamma\,x - \tfrac{1}{2}\,\boldsymbol{\delta}'\,\boldsymbol{\Omega}\,\boldsymbol{\delta}\,\mathbb{E}^*_{t,x,\boldsymbol{y}}\left[\int_t^T \alpha_s^2\,ds\right]\right)$$

holds.

Pairs Trading: Multiple Assets

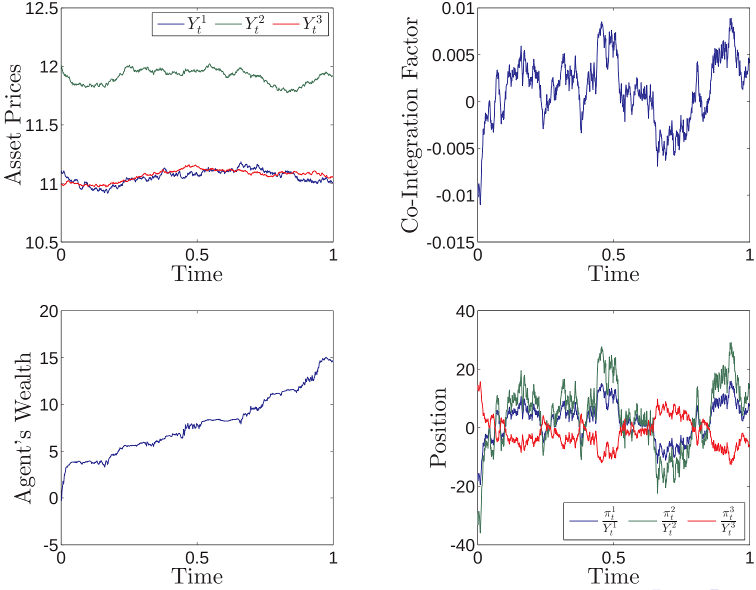


Figure: A single sample path of the co-integrated asset prices and the

Pairs Trading: Multiple Assets

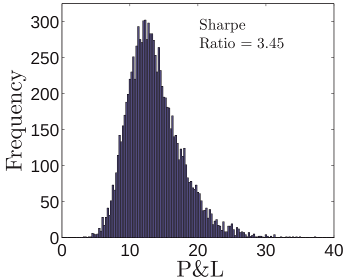


Figure: Histogram of the P&L of the optimal pairs trading strategy. Sharpe Ratio