Mean-Reversion StatArb

Zura Kakushadze

Quantigic® Solutions LLC, Stamford, CT, USA
Business School & School of Physics, Free University of Tbilisi, Georgia

zura@quantigic.com

(Guest lecture given at NYU Stern School of Business, Manhattan)
(Hedge Fund Strategies, FINC-GB.3121, Summer 2015)

July 18, 2015

Outline

Statistical Arbitrage

What is Mean-Reversion?

- Statistical Arbitrage (StatArb): *"refers to highly technical short-term mean-reversion strategies involving large numbers of securities (hundreds to thousands, depending on the amount of risk capital), very short holding periods (measured in days to seconds), and substantial computational, trading, and information technology (IT) infrastructure"* [Andrew Lo, 2010]
- "Mean-reversion": what is it?
- Simple idea: some quantities are historically correlated
- Correlations: temporarily undone by market conditions
- Expect (hope): correlations restored in the future
- StatArb: capture profit from temporary mispricings
- Trader/academic lingo: "mean-reversion" / "contrarian investment"

How Is It Done?

- Implementation: myriad ways of doing mean-reversion
- One approach (schematically):

  mean-reversion via demeaning $\rightarrow$

  regression $\rightarrow$

  weighted regression $\rightarrow$

  (constrained) optimization $\rightarrow$

  factor models

- Street lingo: D.E. Shaw / RBC style StatArb [ZK, 2014]

Pair Trading

- Simplest StatArb: pair trading
- 2 hist corr stocks (same sector): stock A (XOM) & stock B (RDS.A)
- Temp mispricing: A moves up (A is rich), B moves down (B is cheap)
- Pair trading strat: short A, buy B, net $$ neutral (hedge mkt risk)
- "Rich" & "cheap": how to quantify?
- A & B prices: diff, not constant (drift)
- "Rich" & "cheap": returns, not prices
- On average: A & B move in sync
- Say, A moves up more than B on *relative* basis: A is rich, B is cheap

Pair Trading (Cont'ed)

- Prices @t_1 (e.g., yest's close): $P_A(t_1)$ & $P_B(t_1)$
- Prices @t_2 (e.g., today's open): $P_A(t_2)$ & $P_B(t_2)$
- Ex-date today: $P_A(t_1)$ & $P_B(t_1)$ adj for splits/divs
- Returns (typically small, log def is OK):

$$R_A \equiv \ln\left(\frac{P_A(t_2)}{P_A(t_1)}\right)$$
$$R_B \equiv \ln\left(\frac{P_B(t_2)}{P_B(t_1)}\right)$$

- If $R_A > R_B$: A is rich, B is cheap, short A, buy B

Demeaned Returns

Two Stocks

- Mean return:

$$\overline{R} \equiv \frac{1}{2}\left(R_A + R_B\right)$$

- Demeaned returns:

$$\widetilde{R}_A \equiv R_A - \overline{R}$$
$$\widetilde{R}_B \equiv R_B - \overline{R}$$

- Demeaned ret $>$ ($<$) 0: stock is rich (cheap)
- Shares Q_i, $i = A, B$ (> 0 for buys, < 0 for shorts): fixed

$$P_A\ |Q_A| + P_B\ |Q_B| = I \quad \text{(total investment level)}$$
$$P_A\ Q_A + P_B\ Q_A = 0 \quad \text{(\$\$ neutrality)}$$

- P_i, $i = A, B$: prices @ establish

Demeaned Returns

Multiple Stocks

- N hist corr stocks (same sector): XOM, RDS.A, TOT, CVX, BP...
- N returns:

$$R_i = \ln\left(\frac{P_i(t_2)}{P_i(t_1)}\right), \quad i = 1, \dots, N$$

- Mean return:

$$\overline{R} \equiv \frac{1}{N}\sum_{i=1}^{N} R_i$$

- Demeaned returns:

$$\widetilde{R}_i \equiv R_i - \overline{R}$$

- Demeaned ret $>$ ($<$) 0: stock is rich/short (cheap/buy)

Demeaned Returns

Multiple Stocks (Cont'ed)

- 2 conditions, $N > 2$ unknowns (P_i prices @ establish):

$$\sum_{i=1}^{N} P_i \ |Q_i| = I \quad \text{(total investment level)}$$

$$\sum_{i=1}^{N} P_i \ Q_i = 0 \quad \text{(\$\$ neutrality)}$$

- \$\$ positions: $D_i \equiv P_i \ Q_i$
- Simple strat ($\gamma > 0$):

$$D_i = -\gamma \ \widetilde{R}_i$$

- γ fixed: $\gamma = I / \sum_{i=1}^{N} \left|\widetilde{R}_i\right|$
- *One* mean-reversion strat: myriad exist

Demeaned Returns

Multiple Clusters

- Multiple clusters: sectors, industries, sub-industries etc.
- E.g.: stocks in oil, tech & healthcare sectors
- K clusters: labeled by $A = 1, \dots, K$
- $N \times K$ binary (loadings) matrix: $\Lambda_{iA} = 1(0)$ if stock $i \in (\notin)$ cluster A

$$\Lambda_{iA} = \delta_{G(i),A}$$
$$G : \{1, \dots, N\} \mapsto \{1, \dots, K\}$$

- G: maps stocks to clusters
- Clusters don't overlap: do mean-reversion separately for each cluster
- Can do in one shot: regression

BICS Hierarchy Schematically

Industry Classification Hierarchy

Stocks	Sub-Industries	Industries	Sectors	"Market"
AWD1	AWD	AW	A	UNIV
AWD2				
AWE3	AWE			
AWE4				
AXF5	AXF	AX		
AXF6				
AXG7	AXG			
AXG8				
BYQ9	BYQ	BY	B	
BYQ10				
BYR11	BYR			
BYR12				
BZS13	BZS	BZ		
BZS14				
BZT15	BZT			
BZT16				

Regression

Binary Loadings

- Regress (X-sectionally) returns R_i over loadings Λ_{iA}:

$$R_i = \sum_{A=1}^{K} \Lambda_{iA}\ f_A + \varepsilon_i, \quad \sum_{i=1}^{N} \Lambda_{iA}\ \varepsilon_i = 0$$

- f_A / ε_i: regression coeff / residuals (no intercept, $\sum_{A=1}^{K} \Lambda_{iA} = 1$)

$$\overline{R}_A \equiv \frac{1}{N_A} \sum_{j \in J_A} R_j$$

$$\varepsilon_i = R_i - \overline{R}_{G(i)} = \widetilde{R}_i$$

- $\overline{R}_A$: mean return for cluster A (J_A spans N_A stocks in cluster A)
- $\widetilde{R}_i$: demeaned (w.r.t. cluster A) return for stock i in cluster A
- Demeaned returns: regression residuals (cluster neutral)

Regression

Non-Binary Loadings

- Non-binary loadings Ω_{iA}:

$$R_i = \sum_{A=1}^{K} \Omega_{iA}\ f_A + \varepsilon_i, \quad \sum_{i=1}^{N} \Omega_{iA}\ \varepsilon_i = 0$$

- "Regressed" returns: $\widetilde{R}_i = \varepsilon_i$ neutral w.r.t. Ω_{iA}

Weighted Regression

- Non-unit weights w_i:

$$R_i = \sum_{A=1}^{K} \Omega_{iA}\ f_A + \varepsilon_i, \quad \sum_{i=1}^{N} \Omega_{iA}\ w_i\ \varepsilon_i = 0$$

- $\widetilde{R}_i = w_i\ \varepsilon_i$: neutral w.r.t. Ω_{iA} ($$ neutral w/ intercept)

What Should Regression Weights Be?

- Our simple mean-reversion strategy:

$$D_i = -\gamma\ \widetilde{R}_i$$

- "Regressed" returns: $\widetilde{R}_i = w_i\ \varepsilon_i$
- Residuals: on average as volatile as returns R_i
- Strat with $w_i \equiv 1$: loads up on volatile stocks
- Simple idea: suppress weights by volatility $\sigma_i \equiv \sqrt{\mathrm{Var}(R_i)}$
- But how? $w_i = 1/\sigma_i$, $w_i = 1/\sigma_i^2$, ...?
- Naive: $w_i = 1/\sigma_i$; correct: $w_i = 1/\sigma_i^2$

Optimization

Share Ratio Maximization

- P&L, volatility & Sharpe ratio:

$$P = \sum_{i=1}^{N} R_i \; D_i$$

$$V = \sqrt{\sum_{i,j=1}^{N} C_{ij} \; D_i \; D_j}$$

$$S = \frac{P}{V}$$

- C_{ij}: sample covariance matrix of returns
- Optimization $S \rightarrow \max$: solution $D_i = \gamma \; \sum_{j=1}^{N} C_{ij}^{-1} \; R_j$
- Diagonal $C_{ij} = \sigma_i^2 \; \delta_{ij}$
- $D_i = \gamma \; w_i \; R_i$: weights $w_i = 1/\sigma_i^2$

Optimization

Factor Models (FM)

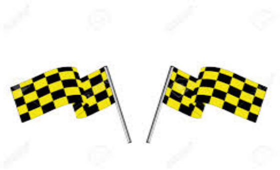
- Sample cov.mat C_{ij}: singular if $M \equiv \#(\text{observations}) < N + 1$
- Off-diag C_{ij}: not out-of-sample stable unless $M \gg N$
- $N \sim 1000 - 2500$ (liquid portfolios): 5 years $\sim$ 1260 daily observations
- Short-holding/ephemeral strats: long lookbacks not desirable/avail
- Factor risk & specific (idiosyncratic) risk: $R_i = \chi_i + \sum_A \Omega_{iA}\, f_A$
- Risk factors: f_A, $A = 1, \dots, K \ll N$
- Specific risk (SR) cov.mat: $\text{Cov}(\chi_i, \chi_j) = \xi_i^2\, \delta_{ij}$
- Factor risk cov.mat: $\text{Cov}(f_A, f_B) \equiv \Phi_{AB}$
- Uncorrelated: $\text{Cov}(\chi_i, f_A) = 0$
- Model cov.mat $\Gamma_{ij} \equiv \text{Cov}(R_i, R_j) = \xi_i^2\, \delta_{ij} + \sum_{A,B=1}^{K} \Omega_{iA}\, \Phi_{AB}\, \Omega_{jB}$
- Φ_{AB} positive-definite: Γ_{ij} positive-definite, invertible
- Opt w/ FM: $\rightarrow$ weighted regression w/ $w_i = 1/\xi_i^2$ in small SR limit

References

Lo, A.W. (2010) *Hedge Funds: An Analytic Perspective.* Princeton University Press, p. 260.

ZK (2014) Mean-Reversion and Optimization. *Journal of Asset Management* **16**(1): 14-40; http://ssrn.com/abstract=2478345.

The End



Thank you!