Stat arb paper

Note Title 1/7/2010

- PCA + sector ETF's

- using "trading time" (i.e. scaled by V) improves s.r.

- pairs trading: instruments P & Q;

$$\frac{dP_t}{P_t} = \alpha\, dt + \beta \frac{dQ_t}{Q_t} + \underline{dX_t} \qquad (1)$$

residual

Q: stationarity $\leftrightarrow$ mean reversion?

$\mathbb{E} X_t = \mathbb{E} X_0$

$\mathbb{E} X_t X_{t+\tau} = \mathbb{E} X_0 X_\tau$

ass.: stationary/mean-reverting

$\alpha\, dt \ll dX_t$

then: $\frac{dP}{P} - \beta \frac{dQ}{Q} \simeq dX$

MR in discr. time: $R_t = a(R_{t-1} - \mu) + u + \sigma w_t$

GMR (OUMR): $dX_t = \eta(\mu - X_t)\,dt + \sigma dW_t \qquad (2)$

$$\log(X_t/X_0) = \int \eta\left(\frac{\mu}{X_t} - 1\right) dt + \sigma W_t \quad ??$$

ICBST OUMR is stationary (cf wiki)
↳ by explicit stoch.-$\int$-ation

Thus, since X assumed MR, if dX is low, be contrarian & take $\{1\,P, -\beta\,Q\}$; if dX is high do the opposite.

Doing so for a sector basket wrt sector ETF nets a long-short portfolio.

General decomp.:

$$\frac{dP_t}{P_t} = \alpha\, dt + \sum \beta_j F_t^{(j)} + dX_t \qquad (3)$$

Q: how to pick a set of $\{F^{(j)}\}$ s.t. dX_t maximizes PNL?

A1: use PCA

A2: use industry-sector ETFs as proxies

PCA: generate M-day standardized returns:

for stock i k days ago the daily return is

(4) $$R_{ik} = \frac{S_{i,k-1} - S_{i,k}}{S_{i,k}} \qquad (S_{i,k} = S_i(t_0 - k\Delta t))$$

std returns: $Y_{ik} = \frac{R_{ik} - \bar{R}_i}{\bar{\sigma}_i}$; $\bar{R}_i$ - M-day avg

$\bar{\sigma}_i^2 = \frac{1}{M-1}\sum (R_{ik} - \bar{R}_i)^2$

sample avg

Define

(5) $$\rho_{ij} = \frac{1}{M-1}\sum_k Y_{ik} Y_{jk}$$

Estimating the parameters of mean reversion using ETF's as factors:

Going back to (1):

$$\frac{dP_t}{P_t} \simeq d\left(\log P_t/P_{t-1}\right) = \frac{dP_t}{P_t} = \frac{P_t - P_{t-1}}{P_t}$$

cont.

i.e. this is simply a stock return.

Now for every stock we will express it thru a combination of ETF returns + idiosync. component that follows OU process.

(6) $$R^S_n = \beta_0 + \beta R^I_n + \epsilon_n, \quad n = 1, 2, \dots 60$$

using 60-day regression for estimation

To correspond to model (6), let

$$\beta_0 = \alpha \Delta t \Rightarrow \alpha = 252 \beta_0$$

$$X_k = \sum_{j=1}^{k} \epsilon_j, \quad k = 1 \dots 60 \quad : \text{discrete OU}$$

(7) AR(1): $X_{n+1} = a + b X_n + \zeta_{n+1}, \quad n = 1 \dots 59$

(8) $dX_i = \kappa_i(\mu_i - X_i)\,dt + \sigma_i\,dW_t$ (2, rewritten)

$\mathbb{E}X_i = \mu_i, \quad \operatorname{Var} X_i = \sigma_i^2/2\kappa_i$

& param. estimates are

$$a = \mu\left(1 - e^{-\kappa\Delta t}\right)$$

(9) $$b = e^{-\kappa\Delta t}$$

$$\operatorname{Var}\zeta = \sigma^2\,\frac{1 - e^{-2\kappa\Delta t}}{2\kappa}$$

$\Rightarrow$

$$\kappa = 252\cdot(-\log b)$$

$$\mu = \frac{a}{1-b}$$

$$\sigma = \sqrt{\frac{\operatorname{Var}\zeta\cdot 2\kappa}{1-b^2}}$$

$$\sigma_{eq} = \sqrt{\frac{\operatorname{Var}\zeta}{1-b^2}}$$

(10) $$s := \frac{X(t) - \mu}{\sigma_{eq}} = \frac{-\mu}{\sigma_{eq}} = \frac{-a\sqrt{1-b^2}}{(1-b)\sqrt{\operatorname{Var}\zeta}}$$

since $X(t) = X_{60} = 0$

Finally, in the paper they choose to use

$\bar{\mu} := \frac{a}{1-b} - \left\langle \frac{a}{1-b} \right\rangle$, hence

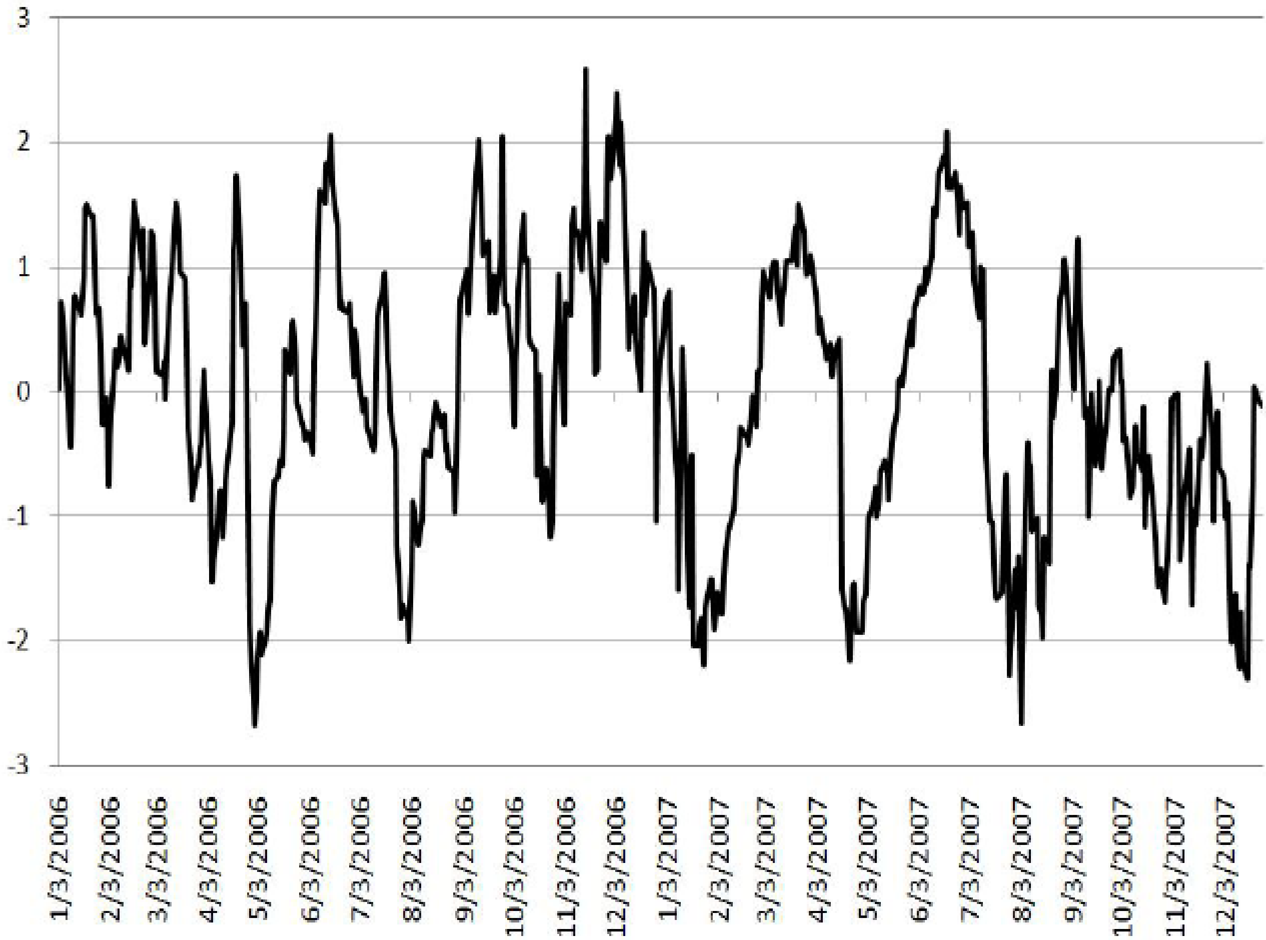
(11) $$s = \left(\quad \right)\left(-\frac{\sqrt{1-b^2}}{\sqrt{\operatorname{Var}\zeta}}\right)$$

So, the proc. is:

→ For every stock, use 15 ETF's to get the β's & residuals using simple regression

→ For the 60-term residuals series fit AR(1) using (7) & (9)

→ Compute the s-score from (11)



3
2
1
0
-1
-2
-3
1/3/2006
2/3/2006
3/3/2006
4/3/2006
5/3/2006
6/3/2006
7/3/2006
8/3/2006
9/3/2006
10/3/2006
11/3/2006
12/3/2006
1/3/2007
2/3/2007
3/3/2007
4/3/2007
5/3/2007
6/3/2007
7/3/2007
8/3/2007
9/3/2007
10/3/2007
11/3/2007
12/3/2007

Pairs trading: basic motivation:

$$ln(P_t/P_{t0}) = \alpha(t - t_0) + \beta ln(Q_t/Q_{t0}) \ + \ X_t$$

or, in its differential version,

$$\frac{dP_t}{P_t} \ = \ \alpha dt \ + \ \beta\frac{dQ_t}{Q_t} \ + dX_t,$$

Suppose $\{P_{t_0}, Q_{t_0}\} = \{41, 29\}$; $\beta = 1.25$

$\alpha \approx 0$; we invest 4220 in P, 5220 in Q

In the case of ETF factors, we work with the model

$$\frac{dS_i(t)}{S_i(t)} = \alpha_i \, dt + \beta_i \frac{dI(t)}{I(t)} + dX_i(t),$$

$$dX_i(t) = \kappa_i \left(m_i - X_i(t)\right) dt + \sigma_i dW_i(t), \ \ \kappa_i > 0.$$

suppose $\beta = 1.25$; we invest \$1 in P, -\$1.25 in Q

in Δt suppose P goes up 10%; this means Q goes up $\approx 12.5\%$

$\log \frac{P_t}{P_{t_0}} = \beta \log \frac{Q_t}{Q_{t_0}}$: $P_{t_0} = 1$, $P_t = 1.1$; $Q_{t_0} = 1$

$Q_t = 1.08$

$\Rightarrow$ the value of portfolio is $(1.1 - 1.35 = -0.25)$

Position allocation:

if working with N instruments, and the total equity is $\mathcal{E}$

$\Lambda = \frac{2}{N}$, so when going long, put $2\frac{\mathcal{E}}{N}$ into stock, $-\beta 2 \frac{\mathcal{E}}{N}$ into pair

To avoid overleveraging on small (e.g. test) portfolios, set $N = \max(100, \#\text{stks})$