Pairs Trading: Random Weight Approach

Reza habibi
Iran Banking Institute, Central Bank of Iran

Abstract. A random weight approach is introduced for construction of pairs trading strategies. First, condition for market neutrality of portfolio is considered then, the probability of attaining the profit is maximized. The portfolio contains long position in random units of first asset and random units of another asset in short position. Strategies are given and their performances are evaluated throughout some examples.

Keywords: Beta distribution; Pairs trading; Random weight; Single factor models

1 Introduction. The pairs trading is one of market neutral strategies which belongs to statistical arbitrage strategies. Neutral portfolios with respect to market contains two securities with a long position in one security and a short position in the other, in a suitable ratio (Bondarenko (2003)). An important problem in pairs trading is the identification of pairs and an efficient trading algorithm. In this note, the random weight approach is discussed to obtain the pairs trading strategy. To this end, consider two capital assets, where at time t, their return satisfy in single factor models as follows

$$R_{1t} = \alpha_{1t} + \beta_{1t} R_{mt} + \varepsilon_{1t,}$$

$$R_{2t} = \alpha_{2t} + \beta_{2t} R_{mt} + \varepsilon_{2t},$$

here, R_{mt} is the return of market. In practice, parameters of above model are estimated using suitable estimation methods like the least square or the maximum likelihood methods, hence, in the current paper, it is assumed that parameters $\alpha_{it}, \beta_{it}, i = 1,2$ are known. Variables ε_{it} have known distribution F_{it} with mean μ_{it} and variance $\sigma_{it}^2, i = 1,2$. They are mutually independent. Let γ_t be the weight of first asset and assume that it is a beta distributed random variable with parameters θ_t and τ_t independent of all variables exist in the both models. Next, consider a portfolio of long position in γ_t units of R_{1t} and short position in $(1-\gamma_t)$ units of R_{2t}. Return of related portfolio is

$$R_{pt} = \gamma_t R_{1t} - (1-\gamma_t) R_{2t}$$

At time zero, γ_0 is chosen such that $R_{p0} = 0$ that is $\gamma_0 = \frac{R_{20}}{R_{10}+R_{20}}$. At time t, γ_t, R_{1t} and R_{2t} are chosen such that $P(R_{pt} > 0)$ is close to one.

Proposition 1. $E(R_{pt})$ is independent of $E(R_{mt})$ if and only if $\tau_t = \frac{\beta_{1t}}{\beta_{2t}} \theta_t$, in this case,

$$E(R_{pt}) = \frac{\alpha_{1t}\theta_t - \alpha_{2t}\tau_t}{\theta_t + \tau_t}.$$

Proposition 2. Let $\pi_t = P\left(R_{pt} > 0\right) = P(\gamma_t > Z_t) = 1 - E\left(F_{\gamma_t}(Z_t)\right)$ where $Z_t = \frac{R_{2t}}{R_{1t}+R_{2t}}$. The Monte Carlo estimate of π_t is

$$\hat{\pi}_t = 1 - \frac{1}{R} \sum_{i=1}^{R} F_{\gamma_t}(Z_{it}),$$

where $Z_{it}, i = 1,2, \dots, R$ is an iid sample from Z_t.

Remark 1. To make sure that $\hat{\pi}_t \to 1$, almost sure, it is enough to determine $\theta_t$and τ_t such that

$$\frac{1}{R}\sum_{i=1}^{R} F_{\gamma_t}(Z_{it}) \to 0, a.s.$$

A pairs selection strategy is to select R_{1t} and R_{2t} such that $R_{1t} \to 0$ and $R_{1t} \to \infty$, *a.s*.

Remark 2 (Portfolio construction). In previous section, it was seen that the portfolio constructed as $R_p = \gamma R_1 - (1-\gamma)R_2$ has a high probability for profitability. An important question is which values of (0,1) should be chosen for γ. The solution is to find θ and τ such that $P(\gamma > z) = 1-\alpha$ for some small α. Thus, z is the α −the quantile of distribution of γ. Thus, generate $z_1, \dots, z_R$ and calculate their average $\bar{z}$. Find θ and τ such that $\bar{z}$ is the α −the quantile of distribution of γ. Then, generate one sample from this distribution, a suitable value for γ is this number.

Example 1. In this example, assume that $\alpha_1 = 0.5, \alpha_2 = 0.75, \beta_1 = 0.5, \beta_2 = 1.7$. Variables ε_1 and ε_2 have normal distributions with zero means and variances 0.01 and 0.36. The market return has normal distribution with mean 1.5 and variance 0.16. $\theta \in \{0,0.1,0.02,\dots,5\}$ and $\tau = \frac{\beta_1}{\beta_2}\theta$. Here, the time series plot of π_k is drawn. The number of repetition is $R = 1000.$ This shows the high probability of existing arbitrage opportunity.

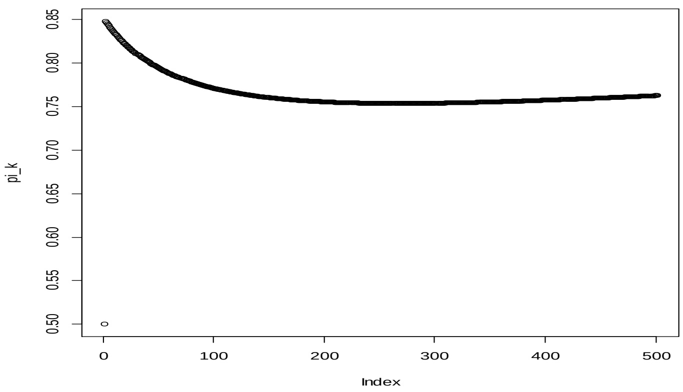

Figure 1: Probability of statistical arbitrage

Example 2. Here, for various α_i and β_i's suitable values for θ and τ are computed in Table 1. Again, it is assumed that ε_1 and ε_2 have normal distributions with zero means and variances 0.01 and 0.36. The market return has normal distribution with mean 1.5 and variance 0.16. here, it is assumed that $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2$. Parameters θ and τ are computed such that the probability of statistical arbitrage is maximized.

Table 1: Values for θ and τ

α_1	β_1	θ	τ
0.5	0.6	3.98	3.98
0.7	0.8	5	5
0.9	0.7	5	5
1.1	0.9	5	5
1.1	1.3	5	5
1.5	0.04	5	5
0.06	0.03	0	0

As follows, conditions are derived to ensure being the probability of statistical arbitrage to be high. The time t is fixed and therefore it is dropped from notation. Define $k_1 = \frac{\alpha_1+\alpha_2}{\alpha_2}$, $k_2 = \frac{\beta_1+\beta_2}{\beta_2}$, $k_3 = \frac{\sigma_1}{\sigma_2}$. Suppose that $\alpha_2 \to 0$ and $k_1 \to 1$. That is, $\frac{\alpha_1}{\alpha_2} \to 0$. The similar conditions happens for k_2 and β_1, β_2. Also, suppose that k_3 doesn't converges to zero and bounded away zero. It is easy to see that

$$z \approx \frac{\varepsilon_2^*}{\varepsilon_2^* + k_3\varepsilon_1^*},$$

almost sure, where ε_1^* and ε_2^* have standard normal distributions. When k_3 isn't close to zero, then z has Cauchy distribution and its expectation and variance is infinite. It is natural to expect that z has small and large values too much. So, it is reasonable to suggest the investor to trade when z is small. To see the Cauchy distribution of z, notice that for arbitrary number c, we have

$$z - c = \frac{\varepsilon_2^*}{\varepsilon_2^* + k_3\varepsilon_1^*} - c = \frac{(1-c)\varepsilon_2^* - k_3 c\varepsilon_1^*}{\varepsilon_2^* + k_3\varepsilon_1^*}.$$

Determine c such that two linear combination of two independent normal variables be orthogonal. That is $(1-c) - k_3^2 c = 0$. Then,

$$c = \frac{1}{1+k_3^2}.$$

Then,

$$U = \sqrt{\frac{1+k_3^2}{(1-c)^2 + k_3^2 c^2}}\, z = \frac{1+k_3^2}{k_3^2}\, z$$

is the fraction of two independent standard normal variables, therefore it is a Cauchy distributed random variable. Notice that

$$\pi_t = P(\gamma_t > z) = P\left(\gamma_t > \frac{k_3^2}{1+k_3^2} U\right).$$

Next, calculate U, when it is small the π_t is high. The following proposition summarizes the results.

Proposition 3. Let $k_1 = \frac{\alpha_1+\alpha_2}{\alpha_2}$, $k_2 = \frac{\beta_1+\beta_2}{\beta_2}$, $k_3 = \frac{\sigma_1}{\sigma_2}$. Suppose that $\alpha_2 \to 0$ and $k_1 \to 1$. That is, $\frac{\alpha_1}{\alpha_2} \to 0$. The similar conditions happens for k_2 and β_1, β_2. Also, suppose that k_3 doesn't converges to zero and bounded away zero. Generate a random sample from z and calculate U. Hence, trade when U is small.

References

[1] Bondarenko, O. (2003). Statistical Arbitrage and Securities Prices. *Review of Financial Studies* **16** : 875–919.