Cointegration Based Strategy

Abstract

Generalizes standard cointegration strategy for multiple common trends.

1 Model

The risky assets $Y = (Y_t^1, \dots, Y_t^n)$ satisfy the coupled system

$$\frac{dY_t^k}{Y_t^k} = \sum_{m=1}^{M} \Delta_{k,m}\alpha_{t,m}dt + \sum_{i=1}^{n} \sigma_{ki}dW_t^i$$

where

$$\alpha_{t,m} = a_{0,m} + \sum_{i=1}^{n} a_{i,m} \log Y_t^i.$$

Define $\Omega = \sigma\sigma^\top$, so that $\Omega_{ij} = \sum_{k=1}^n \sigma_{ik}\sigma_{kj}^\top = \sum_{k=1}^n \sigma_{ik}\sigma_{jk}$.

$$d\log Y_t^k = \left(\sum_{m=1}^{M} \Delta_{k,m}\alpha_{t,m} - \frac{1}{2}\Omega_{kk}\right)dt + \sum_{i=1}^{n} \sigma_{ki}dW_t^i = \left(\Delta_{k,\cdot}^\top\alpha_t - \frac{1}{2}\Omega_{kk}\right)dt + \sum_{i=1}^{n} \sigma_{ki}dW_t^i.$$

Also,

$$\begin{aligned} d\alpha_{t,j} &= d\left(a_{0,j} + \sum_{i=1}^{n} a_{i,j}\log Y_t^i\right) = \sum_{k=1}^{n} a_{k,j}\left(\sum_{m=1}^{M} \Delta_{k,m}\alpha_{t,m} - \frac{1}{2}\Omega_{kk}\right)dt + \sum_{k=1}^{n} a_{k,j}\sum_{i=1}^{n}\sigma_{ki}dW_t^i \\ &= \left(a_j^\top\Delta\alpha_t - \frac{1}{2}\mathrm{Tr}(A_j\Omega)\right)dt + a_j^\top\sigma dW_t \\ &= \kappa(\theta_t - \alpha_{t,j}) + a_j^\top\sigma dW_t \end{aligned} \tag{1}$$

where $\kappa = -\Delta_{\cdot,j}^\top a_j$ and $\theta_t = (\sum_{k=1}^n a_{k,j}\sum_{m=1 m\neq j}^{M}\Delta_{k,m}\alpha_{t,m} - \frac{1}{2}\mathrm{Tr}(A_j\Omega))/\kappa$. There is still mean reversion but the value to which $\alpha_{t,j}$ reverts depends on other trends.

Next, let $Z_t = \log Y_t$, then we have

$$dZ_t = (c + \alpha\beta^\top Z_t)dt + \sigma dW_t$$

where $c = \Delta a_0 - \mathrm{diag}(\Omega)/2$, $\alpha = \Delta$ and $\beta = [a_1, \dots, a_m]$ are $n \times M$ matrices. This has the form of an error correction VAR(1) model with m unit roots.

The wealth process and quadratic variations

$$\begin{aligned} dX_t^\pi &= \sum_{k=1}^{n} \pi_t^k \frac{dY_t^k}{Y_t^k} = \pi^\top\Delta\alpha_t dt + \pi^\top\sigma dW_t \\ d[Y^i, Y^j]_t &= Y_t^iY_t^j\sum_{k=1}^{n}\sigma_{ik}\sigma_{jk} = (y_t^\top\Omega y_t)_{ij}dt \\ d[X^\pi, X^\pi]_t &= (\pi_t^\top\Omega\pi_t)dt \\ d[X^\pi, Y^k]_t &= (\pi_t^\top\Omega)_k y_t^k dt. \end{aligned} \tag{2}$$

2 Optimization

The agent's criterion is $H^\pi(t,x,y) = \mathbb{E}_{t,x,y}\left[-\exp(-\gamma X_T^\pi)\right]$ and the value function $H(t,x,y) = \sup_{\pi\in\mathcal{A}} H^\pi(t,x,y)$. The DPE is given by

$$\partial_t H + \alpha^\top \Delta^\top D_y H + \frac{1}{2} D_{yy}^\Omega H + \sup_\pi \left\{ \pi^\top \Delta\alpha \partial_x H + \frac{1}{2}\pi^\top \Omega \pi \partial_{xx} H + \pi^\top \Omega D_{xy} H \right\} = 0,$$

subject to $H(T,x,y) = -\exp(-\gamma x)$. Here,

$$\begin{aligned} D_y H &= \left(y^1 \partial_{y^1} H, \ldots, y^n \partial_{y^n} H\right)^\top \\ D_{yy}^\Omega H &= \sum_{i,j=1}^n y^i \Omega_{ij} y^j \partial_{y^i y^j} H \\ D_{xy} H &= \left(y^1 \partial_{xy^1} H, \ldots, y^n \partial_{xy^n} H\right)^\top \end{aligned} \tag{3}$$

It can be shown that

$$\pi_t^* = -\frac{\Omega^{-1}\Delta\alpha\partial_x H + D_{xy} H}{\partial_{xx} H} \text{ and, } M = -\frac{1}{2}\frac{L^\top H \Omega^{-1} L H}{\partial_{xx} H}$$

where $LH = \Delta\alpha\partial_x H + \Omega D_{xy} H$.

3 Solution

Take the ansatz $H(t,x,y) = -\exp\left(-\gamma(x + h(t,\alpha^1,\ldots,\alpha^m)\right)$.

$$\begin{aligned}
\partial_{y^k} H &= -\gamma H \sum_{m=1}^M \frac{a_{k,m}}{y^k} \partial_{\alpha^m} h \\
D_y H &= (-\gamma H) \sum_{m=1}^M a_m \partial_{\alpha^m} h \\
\partial_{y^j y^k} H &= \gamma^2 H \left(\sum_{m=1}^M \frac{a_{j,m}}{y^j}\partial_{\alpha^m} h\right)\left(\sum_{m=1}^M \frac{a_{k,m}}{y^k}\partial_{\alpha^m} h\right) - \gamma H \sum_{m,l=1}^M \frac{a_{j,m}a_{k,l}}{y^j y^k}\partial_{\alpha^m\alpha^l} h \\
&\quad + \delta_{jk}\gamma H \sum_{m=1}^M \frac{a_{k,m}}{(y^k)^2}\partial_{a^m} h \\
D_{yy}^\Omega H &= \sum_{i,j=1}^n \Omega_{ij}\left[\gamma^2 H \left(\sum_{m=1}^M a_{i,m}\partial_{\alpha^m} h\right)\left(\sum_{m=1}^M a_{j,m}\partial_{\alpha^m} h\right) - \gamma H \sum_{m,l=1}^M a_{i,m}a_{j,l}\partial_{\alpha^m\alpha^l} h + \delta_{ij}\gamma H \sum_{m=1}^M a_{j,m}\partial_{a^m} h\right] \\
\partial_{xy^k} H &= \gamma^2 H \sum_{m=1}^M \frac{a_{k,m}}{y^k}\partial_{\alpha^m} h \\
D_{xy} H &= \gamma^2 H \sum_{m=1}^m a_m \partial_{\alpha^m} h \\
LH &= -\gamma H \Delta\alpha + \gamma^2 H \sum_{m=1}^M \Omega a_m \partial_{\alpha^m} h = \gamma H\left(\gamma \sum_{m=1}^M \Omega a_m \partial_{\alpha^m} h - \Delta\alpha\right) \\
L^\top H \Omega^{-1} L H &= \gamma^2 H^2 \left(\gamma^2 \sum_{m,l=1}^M a_m^\top \Omega a_l \partial_{\alpha^m} h \partial_{\alpha^l} h + \alpha^\top\Delta^\top\Omega^{-1}\Delta\alpha - 2\gamma \sum_{m=1}^M \alpha^\top \Delta^\top a_m \partial_{\alpha^m} h\right) \\
M &= -\frac{H}{2}\left(\gamma^2 \sum_{m,l=1}^M a_m^\top \Omega a_l \partial_{\alpha^m} h \partial_{\alpha^l} h + \alpha^\top\Delta^\top\Omega^{-1}\Delta\alpha - 2\gamma \sum_{m=1}^M \alpha^\top \Delta^\top a_m \partial_{\alpha^m} h\right)
\end{aligned} \tag{4}$$

The second term $\alpha^\top \Delta^\top D_y H$ and the first term in $D^\Omega_{yy} H$ cancel. We can also take out $-\gamma H$ from all terms.

$$
\begin{aligned}
0 &= \partial_t h + \sum_{i,j=1}^{n} \frac{\Omega_{ij}}{2} \left[\sum_{m,l=1}^{M} a_{i,m} a_{j,l} \partial_{\alpha^m \alpha^l} h - \delta_{ij} \sum_{m=1}^{M} a_{j,m} \partial_{\alpha^m} h \right] + \frac{1}{2\gamma} \alpha^\top \Delta^\top \Omega^{-1} \Delta \alpha \\
0 &= \partial_t h + \frac{1}{2} \sum_{m,l=1}^{M} a_m^\top \Omega a_l \partial_{\alpha^m \alpha^l} h - \sum_{i=1}^{n} \frac{\Omega_{ii}}{2} \sum_{m=1}^{M} a_{i,m} \partial_{\alpha^m} h + \frac{1}{2\gamma} \alpha^\top \Delta^\top \Omega^{-1} \Delta \alpha
\end{aligned} \tag{5}
$$

Next,

$$
\pi^* = \frac{\Omega^{-1} \Delta \alpha - \gamma \sum_{m=1}^{M} a_m \partial_{\alpha^m} h}{\gamma} = \frac{\Omega^{-1} \Delta \alpha}{\gamma} - \sum_{m=1}^{M} a_m \partial_{\alpha^m} h.
$$

4 Optimizing simple return

Let $H(t, x, y) = \sup_{\pi \in \mathcal{A}} \mathbb{E}_{t,x,y}[X_T^\pi]$. The DPE is given by

$$
\partial_t H + \alpha^\top \Delta^\top D_y H + \frac{1}{2} D^\Omega_{yy} H + \sup_\pi \left\{ \pi^\top \Delta \alpha \partial_x H + \frac{1}{2} \pi^\top \Omega \pi \partial_{xx} H + \pi^\top \Omega D_{xy} H \right\} = 0,
$$

subject to $H(T, x, y) = x$. Restrict π to strategies satisfying $\pi^\top \mathbf{1} \le x$ and $-x/N \le \pi_i \le x/N$.

Take the ansatz $H(t, x, y) = x + h(t, \alpha^1, \ldots, \alpha^M)$. The strategy invests in N assets at any given time (assuming $N \le n$) and its the indices corresponding to the top N values in $|\alpha^\top \Delta^\top|$. The value under the supremum is the average of the first N values times x. Let this operation be denoted by $g(\Delta\alpha)x$. Further assume that the amount invested is set at the start of the trading period and the agent always has enough to invest, thus the supremum is $g(\Delta\alpha)x_0$. Parameter N then controls the level of diversification.

$$
\begin{aligned}
\partial_{y^k} H &= \sum_{m=1}^{M} \frac{a_{k,m}}{y^k} \partial_{\alpha^m} h \\
D_y H &= \sum_{m=1}^{M} a_m \partial_{\alpha^m} h \\
\partial_{y^j y^k} H &= \sum_{m,l=1}^{M} \frac{a_{j,m} a_{k,l}}{y^j y^k} \partial_{\alpha^m \alpha^l} h - \delta_{jk} \sum_{m=1}^{M} \frac{a_{k,m}}{(y^k)^2} \partial_{\alpha^m} h \\
D^\Omega_{yy} H &= \sum_{i,j=1}^{n} \Omega_{ij} \left[\sum_{m,l=1}^{M} a_{j,m} a_{k,l} \partial_{\alpha^m \alpha^l} h - \delta_{ij} \sum_{m=1}^{M} a_{i,m} \partial_{\alpha^m} h \right] \\
\partial_{x y^k} H &= 0 \\
0 &= \partial_t h + \sum_{m=1}^{M} \alpha^\top \Delta^\top a_m \partial_{\alpha^m} h + \frac{1}{2} \sum_{m,l=1}^{M} a_m^\top \Omega a_l \partial_{\alpha^m \alpha^l} h - \sum_{i=1}^{n} \frac{\Omega_{ii}}{2} \sum_{m=1}^{M} a_{i,m} \partial_{\alpha^m} h + g(\Delta\alpha) x_0
\end{aligned} \tag{6}
$$

5 Parameter estimation

Use the Var(1) error correction model with m common trends

$$
\Delta x_t = \mu + \alpha \beta^\top x_t + a_t
$$

where a_t are i.i.d. with mean 0 and covariance Σ. Then

$$
\Delta = \alpha, \quad [a_1, \ldots, a_m] = \beta, \quad \Omega = \Sigma
$$

and $a_0 = \min_a |\mu + \mathrm{diag}\Sigma/2 - \Delta a|$.

6 Backtests

Tests were performed by estimating parameters using a one month lookback and running the calculated policy for a three day period, at the end of which parameters were reestimated and the policy was recalculated. The strategy modeled a single common trend. Since it is difficult to predict the range of the future values of α and extreme values are possible, but rare, policy was calculated over a wide range of values $[5\alpha_{\min}, 5\alpha_{\max}]$, where the min and max values were calculated over the lookback period. Then the policy was normalized by the maximum $\sum |\pi_i|$ value over the range, so the maximum possible investment is \$1. This negates the effects of different γ values. To negate the effects of normalizing over such a wide range of α values, a leverage factor of 5 was used. The policy is then scaled up to invest 90% of the total value of the portfolio at the start of the period (leveraged 5x). This resulted in an average investment of about 75% and a maximum investment of about 450% of the available cash. The values can be further optimized to increase profits.

The strategy was tested over the basket of assets consisting ETH, BTC, LINK, BNB, SOL, which show cointegration. The tests were performed over a 1 year period from Jan '24 to Jan '25, with an initial investment of \$1000.

Basket	Return	Maximum Drawdown	Avg. invested	Max. invested
ETH, BTC, LINK	25.50%	-190.24	55%	244%
ETH, BTC, LINK, BNB	39.93%	-180.67	70%	432%
ETH, BTC, LINK, BNB, SOL	48.96%	-177.78	78%	443%

The second strategy is much easier to implement. It involves computing $\Delta\alpha$ values for each asset, then sorting them by decreasing magnitude and taking $1/N$ positions on the top N assets according to their sign. Here, parameters are reestimated at the end of each day. The total investment is set to be the total value of the portfolio at the start of the investment period. The results are much more volatile compared to the first strategy.

Basket	M	Diversification	Return	Maximum Drawdown
ETH, BTC, LINK, BNB	2	3	125.24%	-825.35
ETH, BTC, LINK, BNB	2	4	59.13%	-582.09
ETH, BTC, LINK, BNB, SOL	2	3	165.19%	-1572.42
ETH, BTC, LINK, BNB, SOL	2	4	175.75%	-1089.25
ETH, BTC, LINK, BNB, SOL	2	5	169.00%	-1028.45

7 Introducing fees

The strategies are profitable assuming 0.04% transaction costs. A simple way to reduce the volume traded is to reset the portfolio only after it has changed by a set threshold. Let ΔP denote the amount of volume required to be traded to reset the portfolio. The threshold can have form $\Delta P \geq$ some fixed dollar amount, some percentage of the existing value or some percentage of the current investment. The following table shows the returns and max. drawdowns using the second threshold. The drawdowns are significant.

Basket	0%	2%	4%
ETH, BTC, LINK	(3.92%, -243)	(5.15%, -251)	(6.03%, -243)
ETH, BTC, LINK, BNB	(10.34%, -180)	(12.46%, -177)	(10.05%, -177)
ETH, BTC, LINK, BNB, SOL	(13.85%, -157)	(16.52%, -145)	(15.06%, -149)

To deal with fees methodically, let B_t^i and S_t^i be adapted processes representing the total number of buys and sells in asset i given by $B_t^i = \int_0^t b_t^i dt$ and $S_t^i = \int_0^t s_t^i dt$. Let $\pi_t^i = Y_t^i(B_t^i - S_t^i)$. Rewrite the wealth equation

$$dX_t^\pi = (\pi^\top \Delta\alpha_t - fY_t^\top(b_t + s_t))dt + \pi^\top \sigma dW_t \tag{7}$$

The criterion is

$$H^{(b,s)}(t, x, y, b-s) = E_{t,x,y,b-s}[-\exp(-\gamma X_T^\pi)]$$

or

$$H^{(b,s)}(t, x, y, b-s) = E_{t,x,y,b-s}[X_T^\pi]$$

and the DPE is given by

$$\partial_t H + \alpha^\top \Delta^\top \partial_y H + \frac{1}{2} D_{yy}^\Omega H + \sup_{b_t, s_t} \left\{ (\pi^\top \Delta\alpha - f(b_t + s_t)^\top Y_t)\partial_x H + \frac{1}{2}\pi^\top \Omega \pi \partial_{xx} H + \pi^\top \Omega D_{xy} H + (b_t - s_t)^\top \partial_{b-s} H \right\} = 0.$$

Suppose trading in done in quanta κ_k and at each time, the agent has the option to buy a quantum, sell a quantum, or do nothing for each asset. Furthermore, the agent can hold at most $\bar{q}$ quanta in each direction for any asset.

In the linear case, take the ansatz $H(t, x, y, b-s) = x + h(t, \alpha, q_1, \ldots, q_n)$ where q_k are the number of shares held in asset k. Let $\mathcal{Q}$ define the space of potential actions at any given time. It is $\{-1, 0, 1\}^n$ unless $q_k = \bar{q}$ or $q_k = -\bar{q}$ and denote by $(l_1, \ldots, l_n)$ a vector in $\mathcal{Q}$. Let $\Delta^l h = h(t, \alpha, q_1 + l_1, \ldots, q_n + l_n) - h(t, \alpha, q_1, \ldots, q_n)$ for $l_i \in \{-1, 0, 1\}$. Then the DPE is

$$\begin{aligned} \partial_t h + \sum_{m=1}^M \alpha^\top \Delta^\top a_m \partial_{\alpha^m} h + \frac{1}{2} \sum_{m,l=1}^M a_m^\top \Omega a_l \partial_{\alpha^m \alpha^l} h - \sum_{i=1}^n \frac{\Omega_{ii}}{2} \sum_{m=1}^M a_{i,m} \partial_{\alpha^m} h + \pi^\top \Delta\alpha \\ + \sup_{l \in \mathcal{Q}} \left\{ \sum_{i=1}^n -f|l_i|\kappa_i y^k + \Delta^l h \right\} = 0 \end{aligned} \tag{8}$$

To eliminate y^k, set the price to some y_0^k at the start of the period. Further normalization and approximation can be done by taking $\kappa_k y_0^k = 1$. Then $\pi \approx q$.

The problem is very high-dimensional, but the strategy can be implemented online. A sample test on {ETH, BTC, LINK} for two weeks during Feb '24, where model was refitted at the end of each day and fees were assumed to be 0.04% of the volume traded, shows positive returns.

8 References

- Cartea et. al., Algorithmic and High Frequency Trading
- Tsay, Analysis of Financial Time Series
- Lütkepohl, New Introduction to Multiple Time Series Analysis