

以PCA方式 執行統計套利

鄧力瑋

2024.05.23

策略資訊

商品資訊	<ul style="list-style-type: none">• 現貨交易版塊： RWA• 現貨交易標的： AVAXUSDT, ICPUSDT, MKRUSDT, SNXUSDT, RSRUSDT, CHRUSDT, DUSKUSDT（共7個標的）
交易設定	<ul style="list-style-type: none">• 資金分配: 每交易標的平均分配持有資金• 手續費： 0.036% (參考幣安VIP1設定)• 滑價： 0.05%(買進/賣出皆設5bps)
資料使用	<ul style="list-style-type: none">• 使用8h資料• 2022/01/01 – 2024/01/01• Rolling window方式進行回測(訓練集： 120筆 / 測試集： 30筆)

交易邏輯

Fama-French: $r = R_f + \beta_1 \times \text{市場風險溢酬} + \beta_2 \times \text{規模溢酬} + \beta_3 \times \text{股價淨值比溢酬}$

策略： $R_i = \alpha_i + \sum_{j=1}^m \beta_{ij} F_j + X_i$

1. 參考基準值：該現貨的預期報酬率

2. 殘差項：判定該現貨目前相對價格

1. 用PCA建風險因子F

2. 對殘差項(X_i)建模

- 1 使用訓練期資料，透過「相關性」篩選交易標的
- ➡ 防止做PCA時，將過大的weight放在與其他標的相關性低的標的上

	AVAXUSDT	ICPUSDT	MKRUSDT	SNXUSDT	RSRUSDT	CHRUSDT	DUSKUSDT
AVAXUSDT	1.000000	0.672344	0.584584	0.613869	0.624867	0.686663	0.623157
ICPUSDT	0.672344	1.000000	0.513324	0.538981	0.569503	0.647821	0.552286
MKRUSDT	0.584584	0.513324	1.000000	0.543653	0.491638	0.526965	0.456783
SNXUSDT	0.613869	0.538981	0.543653	1.000000	0.539038	0.581532	0.503377
RSRUSDT	0.624867	0.569503	0.491638	0.539038	1.000000	0.623140	0.607849
CHRUSDT	0.686663	0.647821	0.526965	0.581532	0.623140	1.000000	0.642309
DUSKUSDT	0.623157	0.552286	0.456783	0.503377	0.607849	0.642309	1.000000

所需參數

1. 相關性最低門檻
2. 高於界線個數
3. 交易標的最少個數

交易邏輯

2 將篩選後標的相關係數矩陣做PCA，達到特定解釋度所需的m個Eigenvector（PC個數）

→
$$\begin{pmatrix} V_1^{(1)} & V_1^{(2)} & \dots & V_1^{(m)} \\ V_2^{(1)} & V_2^{(2)} & \dots & V_2^{(m)} \\ \vdots & \vdots & \ddots & \vdots \\ V_7^{(1)} & \dots & \dots & V_7^{(m)} \end{pmatrix}, \quad V^{(m)} : \text{第 } m \text{ 個PC的Eigenvector}$$

所需參數
4. 解釋度

3 利用與實際報酬率與調整後的Eigenvector，做線性組合建立m個風險因子F

→
$$\underbrace{\begin{pmatrix} R_{11} & R_{12} & \dots & R_{17} \\ R_{21} & R_{22} & \dots & R_{27} \\ \vdots & \vdots & \ddots & \vdots \\ R_{T1} & \dots & \dots & R_{T7} \end{pmatrix}}_{\text{Actual Return}} \underbrace{\begin{pmatrix} \frac{V_1^{(1)}}{\sigma_1} & \frac{V_1^{(2)}}{\sigma_1} & \dots & \frac{V_1^{(m)}}{\sigma_1} \\ \frac{V_2^{(1)}}{\sigma_2} & \frac{V_2^{(2)}}{\sigma_2} & \dots & \frac{V_2^{(m)}}{\sigma_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{V_7^{(1)}}{\sigma_7} & \dots & \dots & \frac{V_7^{(m)}}{\sigma_7} \end{pmatrix}}_{\text{Adjusted Eigenvectors}} = \begin{pmatrix} F_{11} & F_{12} & \dots & F_{1m} \\ F_{21} & F_{22} & \dots & F_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ F_{T1} & \dots & \dots & F_{Tm} \end{pmatrix} \left\{ \begin{array}{l} R_{ij}: \text{第 } i \text{ 天第 } j \text{ 個標的實質報酬率} \\ \sigma_i : \text{第 } i \text{ 個標的在 } M \text{ 天內的標準差} \\ T : \text{訓練期期數} \end{array} \right.$$

交易邏輯

$$\text{策略：} R_i = \alpha_i + \sum_{j=1}^m \beta_{ij} F_j + X_i$$

- 4 對風險因子F訓練迴歸模型，可得出訓練期的殘差值。

用訓練出的模型predict，得出預測期的殘差值

$$\rightarrow \begin{pmatrix} \epsilon_{11} & \epsilon_{12} & \dots & \epsilon_{17} \\ \epsilon_{21} & \epsilon_{22} & \dots & \epsilon_{27} \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon_{T1} & \dots & \dots & \epsilon_{T7} \end{pmatrix}, T: \text{訓練期的期數} \quad \begin{pmatrix} X_{11} & X_{12} & \dots & X_{17} \\ X_{21} & X_{22} & \dots & X_{27} \\ \vdots & \vdots & \ddots & \vdots \\ X_{t1} & \dots & \dots & X_{t7} \end{pmatrix}, t: \text{測試期的期數}$$

- 5 對每一幣別的殘差建模，並得出“s-score”： $s_i = \frac{X_{it} - m_i}{\sigma_{eq,i}}$

➡ 利用OU-process計算均值回歸速度，並用AR(1)模型估計OU-process參數

$$\sigma_{eq} = \frac{\sigma}{\sqrt{2\kappa}} \rightarrow \begin{cases} R_n^S = \beta_0 + \beta R_n^I + \epsilon_n \\ V_k = \sum_{j=1}^k \epsilon_j \\ V_{n+1} = a + bV_n + \delta_{n+1} \end{cases} \rightarrow \begin{cases} b = e^{-\kappa\Delta t} \\ \sigma_{eq} = \frac{\sigma}{\sqrt{2\kappa}} \end{cases}$$

得到 a, b, σ (模型的std)

進場時間點設定

所需參數	對策略影響
1. 相關性最低門檻	<ul style="list-style-type: none">設定太高，會降低交易頻率設定太低，會有風險因子失準的問題
2. 高於界線個數	<ul style="list-style-type: none">設定越高，可能間接影響交易頻率設定太低，相關性無法起到篩選作用
3. 交易標的最少個數	<ul style="list-style-type: none">不宜設定太低

	AVAXUSD
AVAXUSD	1.000000
ICPUSD	0.672344
MKRUSD	0.584584
SNXUSD	0.613869
RSRUSD	0.624867
CHRUSD	0.686663
DUSKUSD	0.623157



最終選定參數	
相關性最低門檻	0.6
高於界線個數	3
交易標最少個數	4

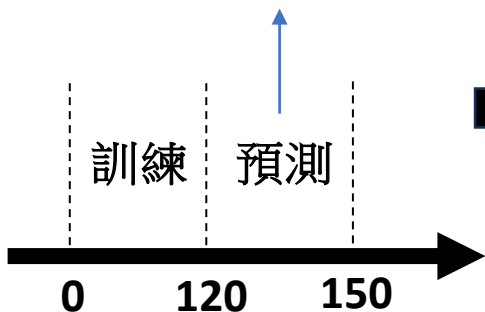
所需參數	對策略影響
4. 解釋度	<ul style="list-style-type: none">設定越低，預期報酬率可能會失準設定過高，預測出的殘差會變過小



最終選定參數	
解釋度	90%

進出場S-score點位設定

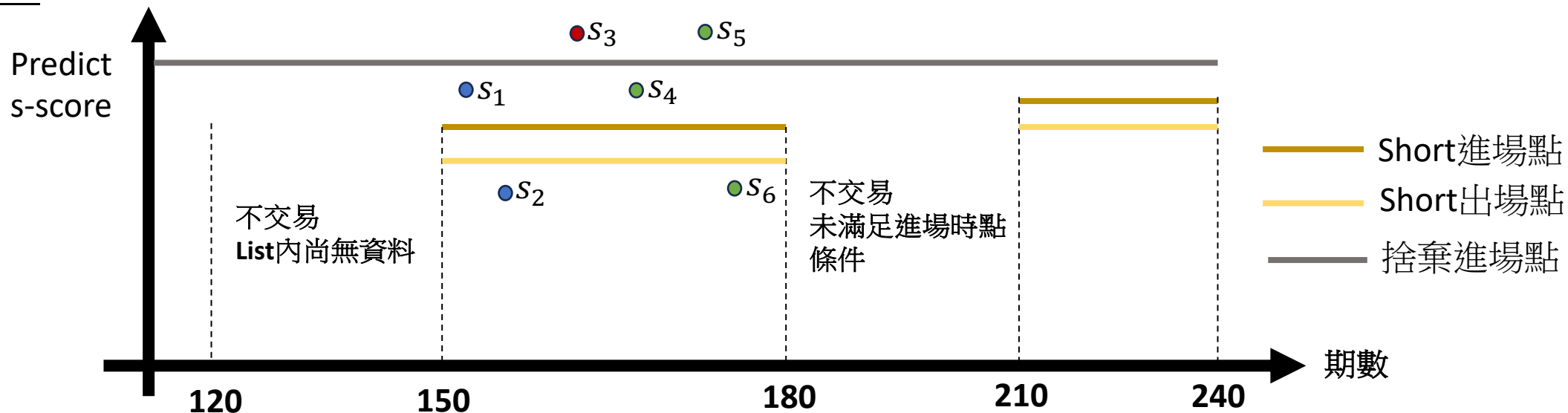
加進Predict_s-score List



1. 計算List的第5、10百分位數，作為下一訓練期long進場可能範圍
2. 計算List的第90、95百分位數，作為下一訓練期short進場可能範圍
3. Long出場點設為進場點的 $[+0.3 \sim +0.6]$ / 停損10%
4. Short出場點設為進場點的 $[-0.3 \sim -0.6]$ / 停損10%
5. 設置捨棄點 $+2$ ，放棄進場

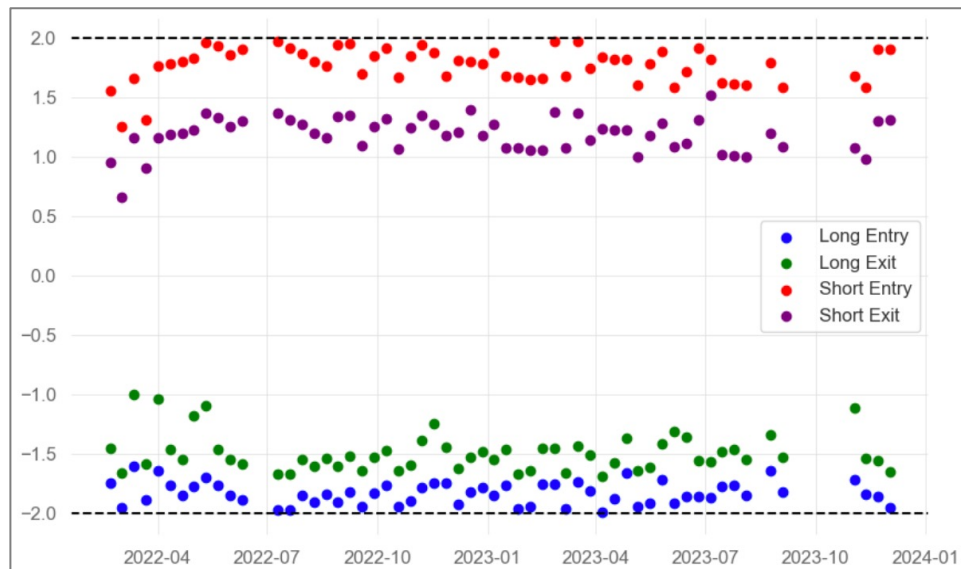
1. 找出訓練期中Sharpe最高的一組參數，用於預測期進行交易
2. 將新的這組Predict s-score加入List

Short策略示意圖：

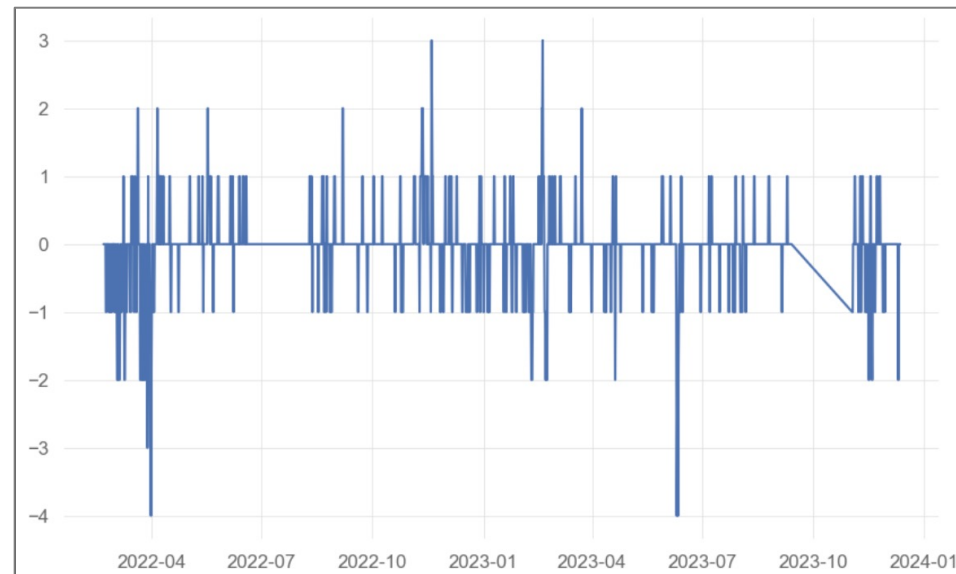


回測結果

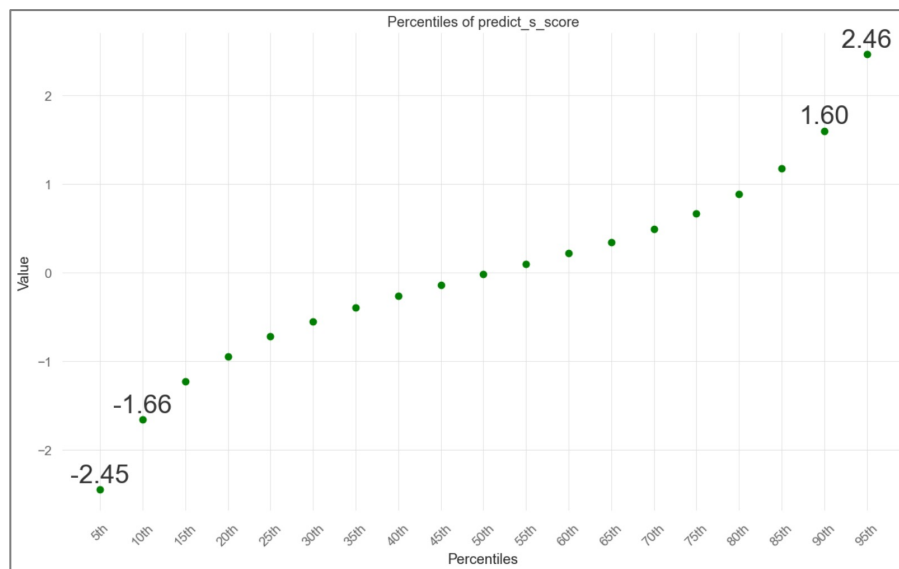
1. 進出場點



2. 單邊曝險



3. Predict S-score 百分位數



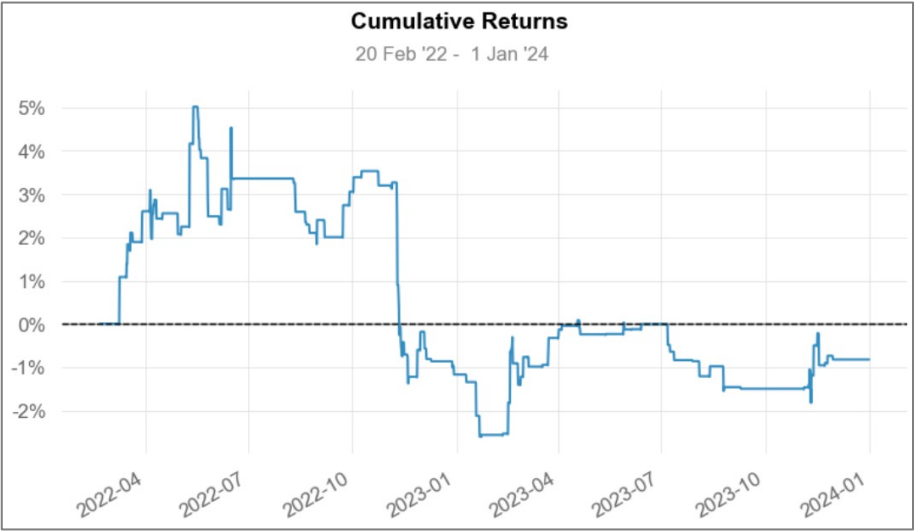
回測績效

1. 多空合併績效



指標	數值
總報酬率	15.31%
CAGR	5.43%
MDD	-6.31%
Sharpe	0.59
Sortino	1.32

2. 做多績效



指標	數值
總報酬率	-0.8%
CAGR	-0.3%
MDD	-7.2%
Sharpe	-0.05
Sortino	-0.08

3. 做空績效



指標	數值
總報酬率	16.3%
CAGR	5.75%
MDD	-2.45%
Sharpe	0.71
Sortino	1.9

感謝聆聽！