



Statistical Arbitrage: Pair Trade using PCA to construct ETF

經濟四 鄧力瑋

目錄

策略

交易標的

回測假設

策略簡介及Signal Generation

績效

樣本內














樣本外

全樣本



交易標的

市值前十的加密貨幣(去除穩定幣以及迷因幣(無實質應用))的永續合約

Rank	Name	Symbol
1	 Bitcoin	BTC
2	 Ethereum	ETH
3	 Tether USDt	USDT
4	 BNB	BNB
5	 Solana	SOL
6	 XRP	XRP
7	 USDC	USDC
8	 Cardano	ADA
9	 Avalanche	AVAX
10	 Dogecoin	DOGE
11	 TRON	TRX
12	 Polkadot	DOT
13	 Polygon	MATIC

日期:2024/1/7

回測假設

手續費:0.05% (幣安Regular User taker)

滑價:用Close買賣(較高)

時間頻率:日K

策略簡介

1.使用PCA建構代表市場的ETF

PCA建構ETF

PCA (主成分分析), 用於降低數據集的維度, 同時保留數據集中的主要信息(Principle Component, PC, 主成分)。

過程: 將資料標準化後的Covariance Matrix對角化後可以得到各eigen vector(PC)的值以及解釋度。本次策略選擇使用解釋度到80%的前4個PC

Covariance Matrix

	BTC	ETH	BNB	XRP	SOL	ADA	AVAX	DOT	MATIC	TRX
BTC	1.000000	0.815344	0.649649	0.598680	0.541740	0.645087	0.565482	0.709637	0.585679	0.632465
ETH	0.815344	1.000000	0.648022	0.590635	0.617179	0.677064	0.594812	0.754917	0.652531	0.629689
BNB	0.649649	0.648022	1.000000	0.532622	0.559945	0.567364	0.557698	0.611534	0.612036	0.562931
XRP	0.598680	0.590635	0.532622	1.000000	0.474603	0.539927	0.504921	0.558361	0.500174	0.586218
SOL	0.541740	0.617179	0.559945	0.474603	1.000000	0.503733	0.549613	0.523900	0.511775	0.434649
ADA	0.645087	0.677064	0.567364	0.539927	0.503733	1.000000	0.654896	0.675783	0.609997	0.574826
AVAX	0.565482	0.594812	0.557698	0.504921	0.549613	0.654896	1.000000	0.604552	0.560555	0.461982
DOT	0.709637	0.754917	0.611534	0.558361	0.523900	0.675783	0.604552	1.000000	0.590556	0.638113
MATIC	0.585679	0.652531	0.612036	0.500174	0.511775	0.609997	0.560555	0.590556	1.000000	0.527777
TRX	0.632465	0.629689	0.562931	0.586218	0.434649	0.574826	0.461982	0.638113	0.527777	1.000000

PCA建構ETF

Covariance Matrix對角化

$$\begin{pmatrix} -16.195 & 1.992 & 0.205 & -0.559 & -0.256 & -4.443 & 17.037 & 0.923 & 0.351 & 1.123 \\ 22.159 & 0.611 & 0.651 & -0.289 & -1.191 & -4.145 & 19.182 & 0.658 & 0.096 & 1.163 \\ 2.502 & -0.526 & -1.897 & -0.344 & 3.397 & 5.286 & 23.302 & 4.728 & -0.174 & 1.042 \\ 0.801 & -0.644 & -0.271 & -0.451 & -0.960 & 3.531 & -51.390 & 6.854 & 0.630 & 0.968 \\ -3.763 & -0.232 & -0.873 & 0.524 & -0.811 & -3.593 & -3.207 & 12.053 & -0.992 & 0.938 \\ -0.320 & -0.039 & -6.307 & 0.175 & -0.450 & -1.088 & -12.598 & -9.695 & -0.124 & 1.070 \\ 1.442 & 1.078 & 3.683 & -0.020 & 1.623 & -0.336 & -34.569 & -7.160 & -0.810 & 0.998 \\ -5.287 & -3.245 & 2.284 & 0.010 & 0.163 & -4.574 & 9.761 & -4.261 & 0.247 & 1.110 \\ -3.593 & 0.008 & 1.581 & 0.088 & -2.458 & 9.664 & 22.800 & -3.363 & -0.331 & 1.017 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 0.165 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.279 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.301 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.378 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.383 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.465 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.502 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.541 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.660 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6.325 \end{pmatrix} = \begin{pmatrix} -0.020 & 0.027 & 0.003 & 0.001 & -0.005 & -0.000 & 0.002 & -0.006 & -0.004 & 0.001 \\ 0.112 & 0.034 & -0.030 & -0.036 & -0.013 & -0.002 & 0.061 & -0.182 & +0.000 & 0.056 \\ 0.003 & 0.010 & -0.028 & -0.004 & -0.013 & -0.094 & 0.055 & 0.034 & 0.024 & 0.015 \\ -0.275 & -0.142 & -0.169 & -0.222 & 0.258 & 0.086 & -0.010 & 0.005 & 0.043 & 0.492 \\ -0.010 & -0.049 & 0.139 & -0.039 & -0.033 & -0.018 & 0.066 & 0.007 & -0.100 & 0.041 \\ -0.021 & -0.020 & 0.026 & 0.017 & -0.017 & -0.005 & -0.002 & -0.022 & 0.047 & 0.005 \\ 0.003 & 0.003 & 0.004 & -0.009 & -0.001 & -0.002 & -0.006 & 0.002 & 0.004 & 0.000 \\ 0.002 & 0.002 & 0.012 & 0.018 & 0.031 & -0.025 & -0.018 & -0.011 & -0.009 & 0.003 \\ 0.104 & 0.028 & -0.051 & 0.186 & -0.293 & -0.037 & -0.239 & 0.073 & -0.098 & 0.295 \\ 0.103 & 0.106 & 0.095 & 0.089 & 0.086 & 0.098 & 0.091 & 0.102 & 0.093 & 0.092 \end{pmatrix}$$

PCA建構ETF

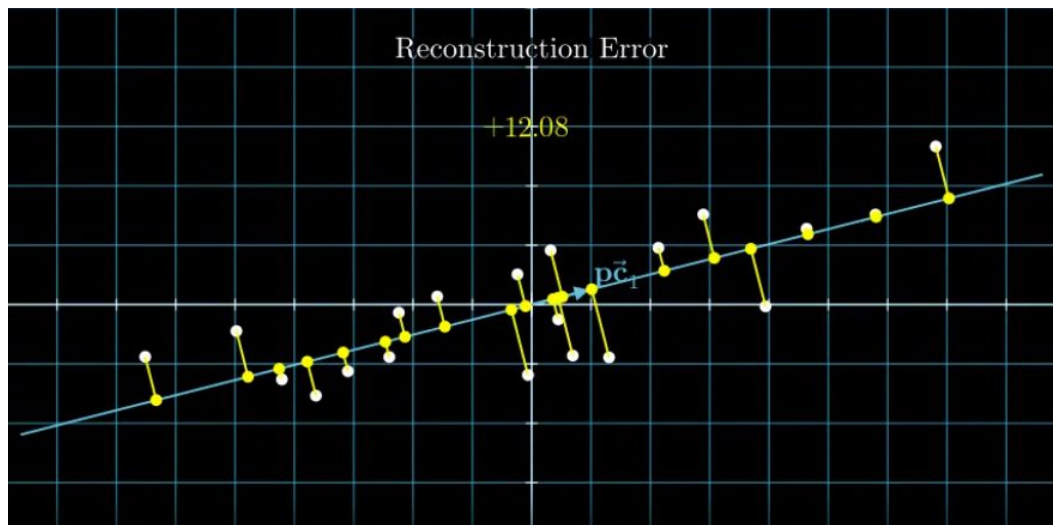
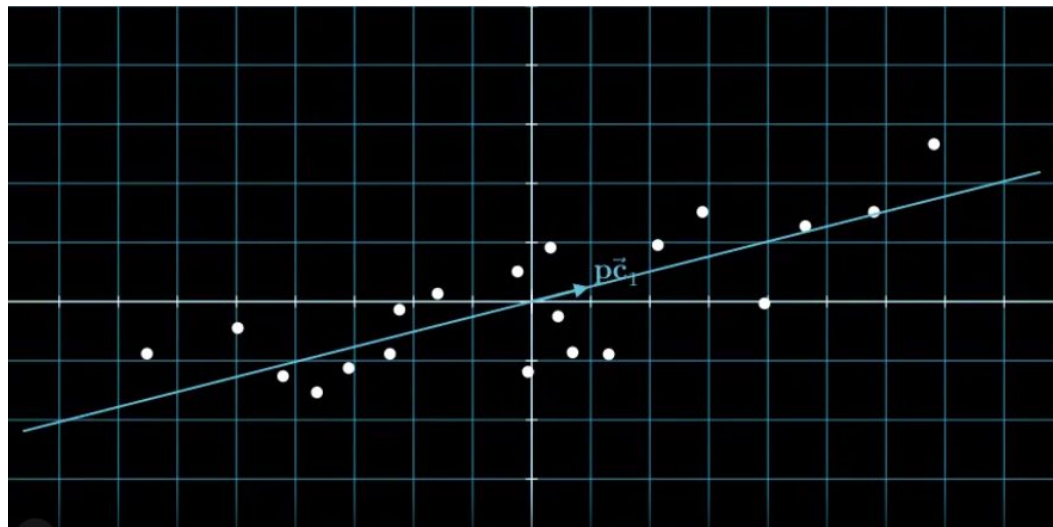
各PC的解釋度

```
[0.63579314 0.06477758 0.05225874 0.05004138 0.04662945 0.03818503  
0.03798176 0.03000957 0.02780516 0.01651819]
```

eigen vectors(壓縮過資料的座標軸) 及轉換成ETF比重的資料

```
eig_vec_1 = np.array([1.123,1.166,1.041,0.962,0.947,1.078,1.008,1.116,1.021,1])  
eig_vec_2 = np.array([ 0.302, 0.071,-0.168, 0.721,-0.957,-0.124,-0.753, 0.187,-0.345,1])  
eig_vec_3 = np.array([ 2.048, 1.345, 7.601, 12.508, 22.996,-16.604,-11.979, -7.310, -8.273,1])  
eig_vec_4 = np.array([ 2.841, 2.865, 3.500,-7.743,-0.349,-2.102,-5.464, 1.562, 2.580,1])  
  
[0.10734085 0.11145097 0.09950296 0.09195183 0.09051807 0.10303957  
0.09634869 0.10667176 0.09759128 0.09558402]  
[ 0.06525497 0.0153414 -0.03630078 0.15579084 -0.20678479 -0.02679343  
-0.16270527 0.04040622 -0.07454624 0.21607606]  
[ 0.02234247 0.01467315 0.08292241 0.13645488 0.25087275 -0.18113981  
-0.1306838 -0.07974777 -0.09025353 0.01090941]  
[ 0.09468106 0.0954809 0.11664334 -0.25804839 -0.01163101 -0.07005266  
-0.18209691 0.05205626 0.0859828 0.03332667]
```


PCA建構ETF



PCA vs MarketCap ETF

聚焦在各daily return的關係，
不因市值大小錯過小幣漲幅

市值ETF只有一個因子，定
價CAPM return較不準確

策略簡介

- 1.使用PCA建構代表市場的ETF
- 2.以個別標的daily return為Y,對各個PC建構的ETF的daily return做多元線性回歸(建立各標的CAPM理論上的daily return)
- 3.將標的實際上的daily return減掉理論上的daily return並標準化
- 4.判斷當日漲跌幅相較於理論是溢價(漲太多or跌太少),或是折價(漲太少or跌太多),並在折價時做多,溢價時做空(論文作法)

多元線性迴歸

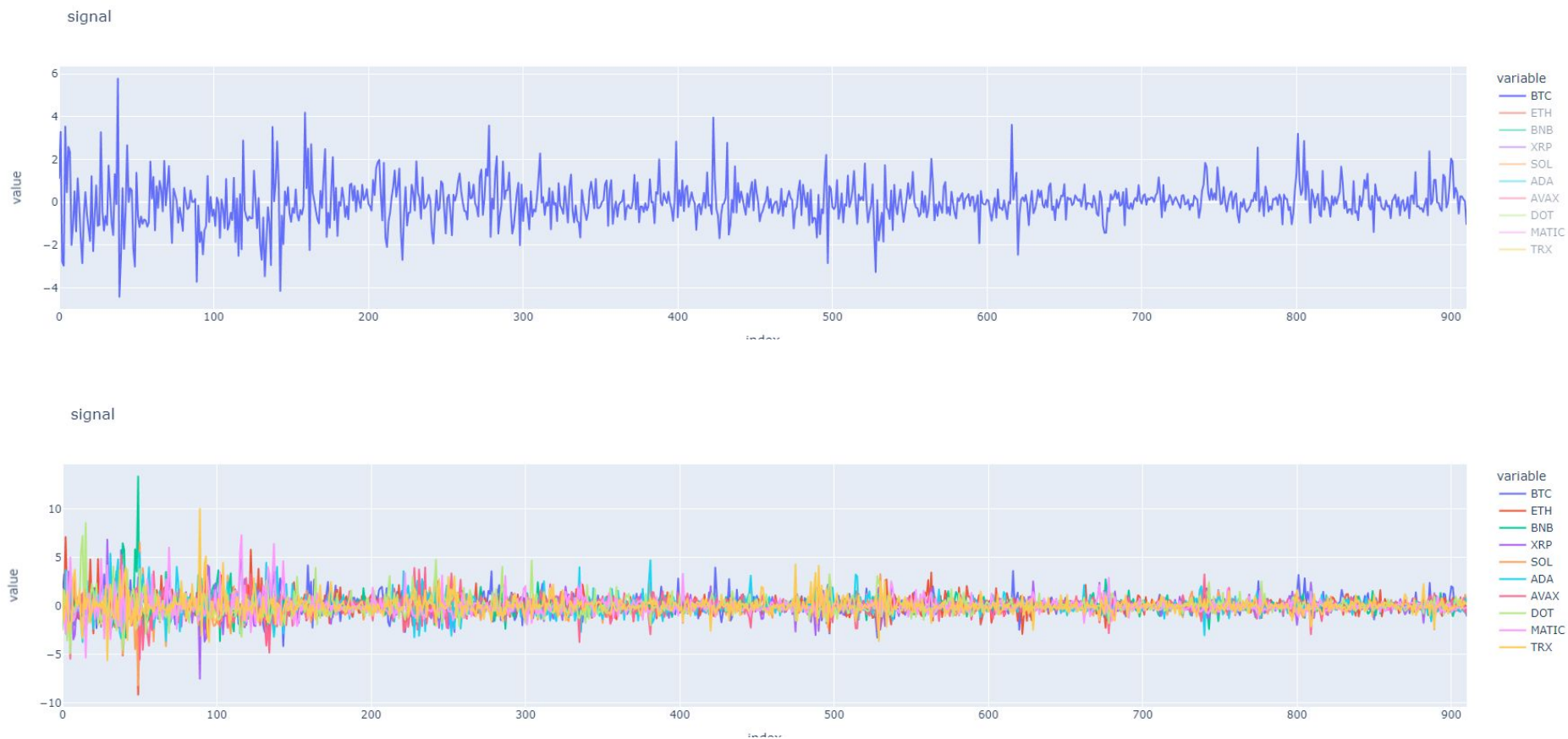
$$R_{CAPM} = a + B_1 * R_1 + B_2 * R_2 + B_3 * R_3 + B_4 * R_4$$

	B1	B2	B3	B4	a
BTC	0.711148	0.323819	0.071786	0.238998	-0.001787
ETH	0.96572	0.207722	0.086307	0.437559	-0.000818
BNB	1.023985	-0.046656	0.478242	0.71438	0.000711
XRP	1.056697	1.05256	0.812568	-1.535978	0.000831
SOL	0.871255	-1.487917	1.767535	-0.061935	-0.000739
ADA	0.925896	-0.038714	-0.942744	-0.296237	-0.000515
AVAX	0.866825	-1.314961	-1.150506	-1.431241	-0.000006
DOT	1.183667	0.470635	-0.427542	0.314308	-0.001792
MATIC	1.388334	-0.472153	-0.623384	1.326981	0.003491
TRX	1.018531	1.131705	0.141661	0.070193	0.001102

$$\text{residual} = R_{\text{Real}} - R_{\text{CAPM}}$$

$$\text{signal} = [\text{residual} - \text{mean}(\text{residual})] / \text{std}(\text{residual})$$

Signal Generation



策略簡介

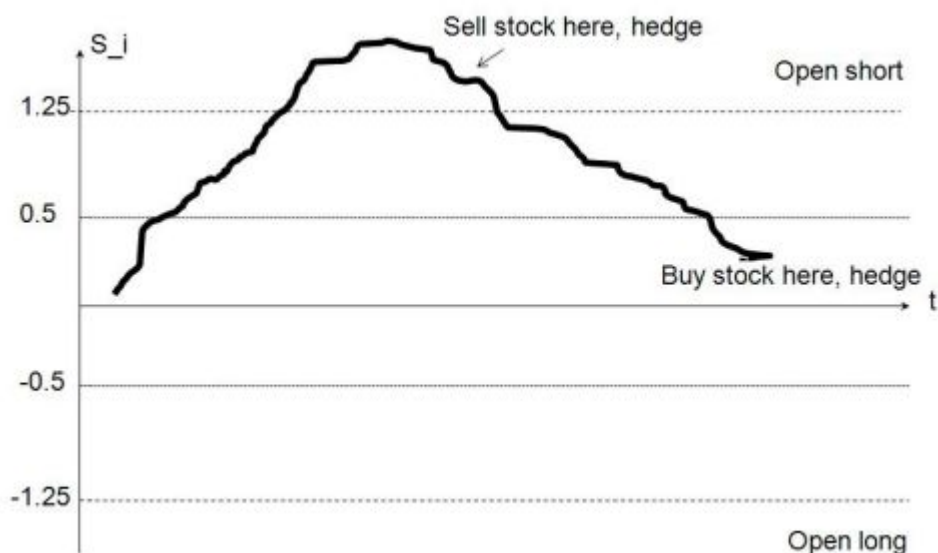
十個標的:把錢切成十份

開倉做多(做多標的、做空ETF): $\text{signal} < -1.25$

平倉做多: $\text{signal} > -0.5$

開倉做空(做空標的、做多ETF): $\text{signal} > 1.25$

平倉做空: $\text{signal} < 0.75$



策略簡介

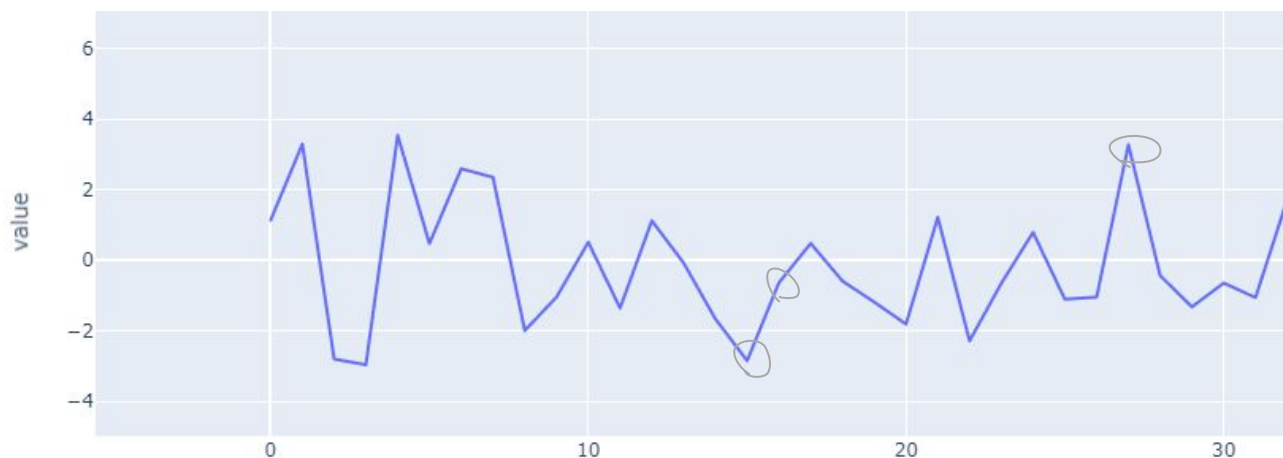
實測:輸爛

做多標的:小贏

避險的做空ETF 以及做空標的:大輸

改良思路: 只做多 平倉訊號改為原本做空信號

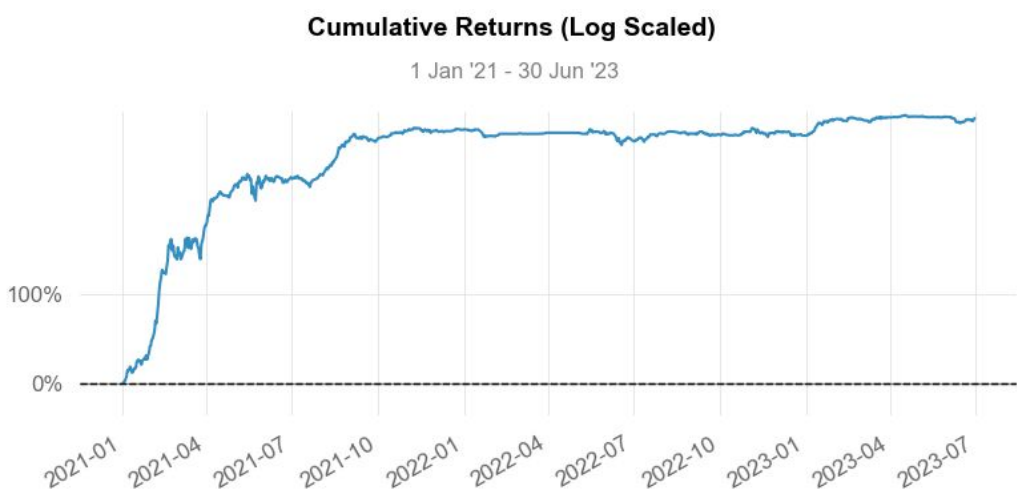
ex:原訊號為進場-2,出場-1.5改為進場-2出場+2



```
entry_list = [-2, -1.75, -1.25, -2, -1.75, -1.25, -1, -1.5, -1.25, -2]
exit_list = [2, 2, 2, -0.25, 2, 1, 0.5, -0.25, 1.25, 1]
```

拉長建倉時間, 使mispricing幅度增加

績效:樣本內



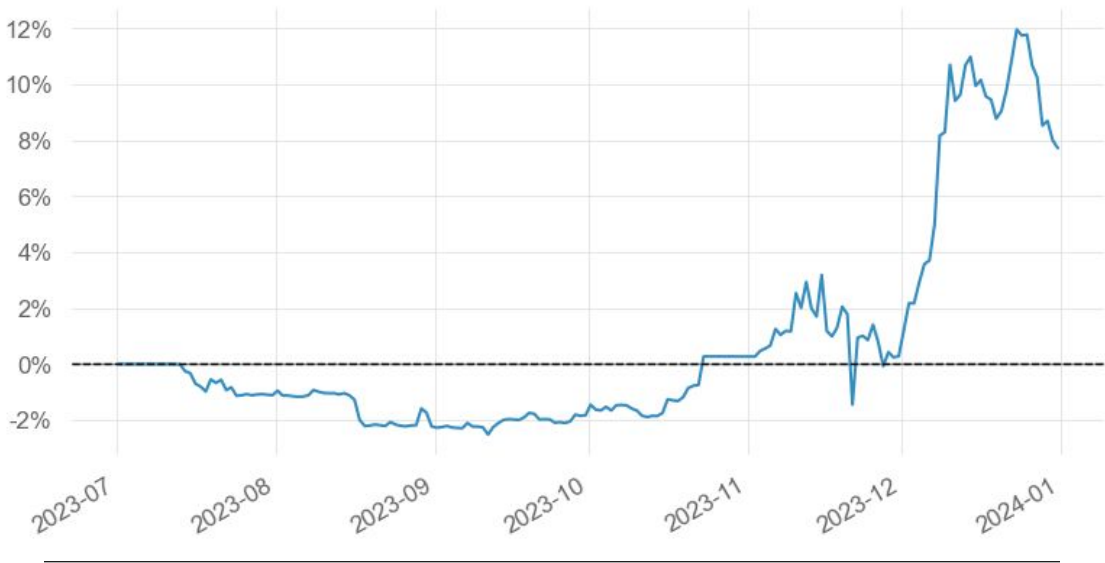
總報酬(單利):222.2%
年化sharpe:2.54
年化sortino:4.47
Max Drawdown:-24.01%



績效:樣本外

Cumulative Returns

1 Jul '23 - 31 Dec '23



總報酬(單利):7.74%

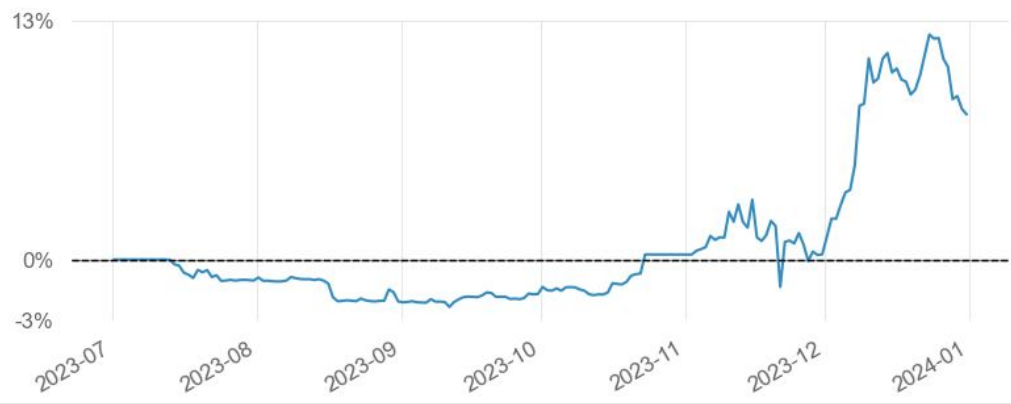
年化sharpe:1.39

年化sortino:2.18

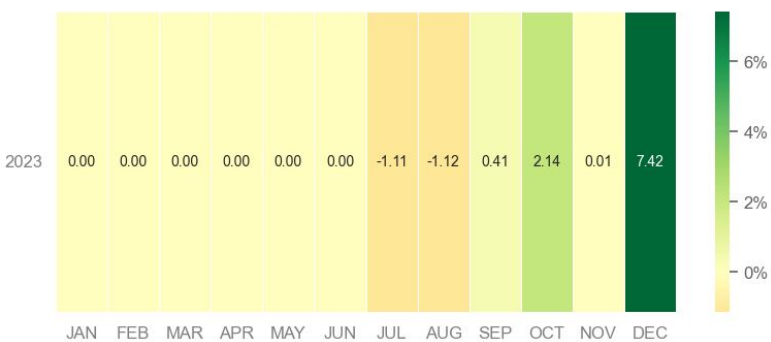
Max Drawdown:-4.49%

Cumulative Returns (Log Scaled)

1 Jul '23 - 31 Dec '23



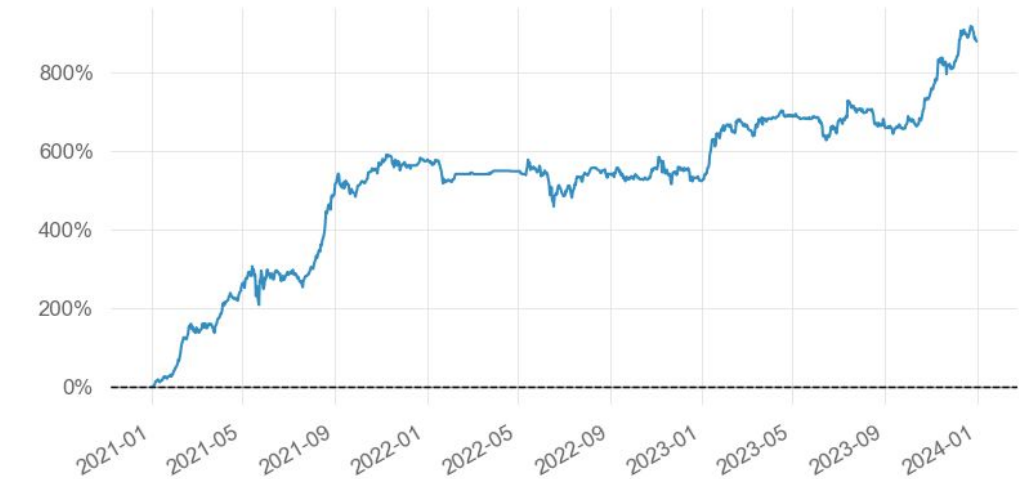
Strategy - Monthly Returns (%)



績效:全樣本

Cumulative Returns

1 Jan '21 - 31 Dec '23



Cumulative Returns (Log Scaled)

1 Jan '21 - 31 Dec '23



總報酬(單利):235.67%

年化sharpe:2.48

年化sortino:4.37

Max Drawdown:-24.06%

Strategy - Monthly Returns (%)





感謝聆聽，敬請指教