Pairs Trading Strategies

the Optimization in Decision-making Processes

Hungwei Chang
Jiaxin Li
Tianpei Zhu
Xiaohan Cheng
Yanlin Chen
Yuanhang Zhao

Duke University

April 30, 2021

Pairs Trading Strategy

1. Pairs Selection
2. Parameters Calculation
3. Risk Management
4. Performance
5. **Sub Strategy – Hedge Macro Risks**
6. **Data**
   - S&P Components
   - ETFs
   - China Concept Shares

Pair Selection

1. PCA&Clustering
Find the clusters in the Ticker set.
2. Filtered by Criteria
Correlation
Cointegration
Hurst
Fundamental Data

Parameters

1. OLS
2. Kalman Filter
Get the **DYNAMIC** Hedge Ratio.

Risk Management

1. Margin Constraits
2. Stop-loss Limit
3. VaR

Performance

1. Sharpe Ratio
2. Total Return
3. Max Drawdown

DBSCAN

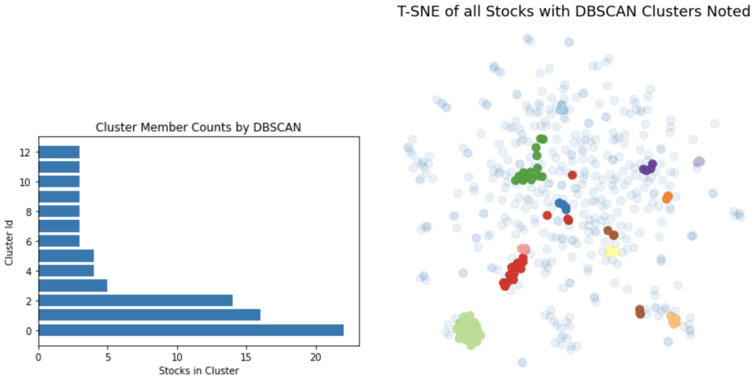
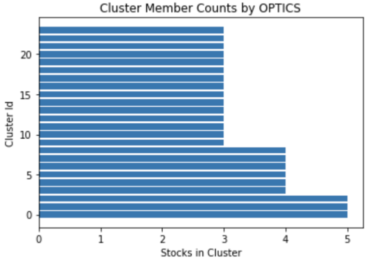


Figure: Clustering Result by DBSCN

OPTICS



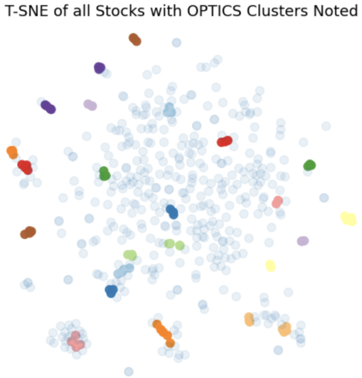


Figure: Clustering Result by OPTICS

Stage Optimization

Pairs Selection - Clustering Methodologies

OPTICS

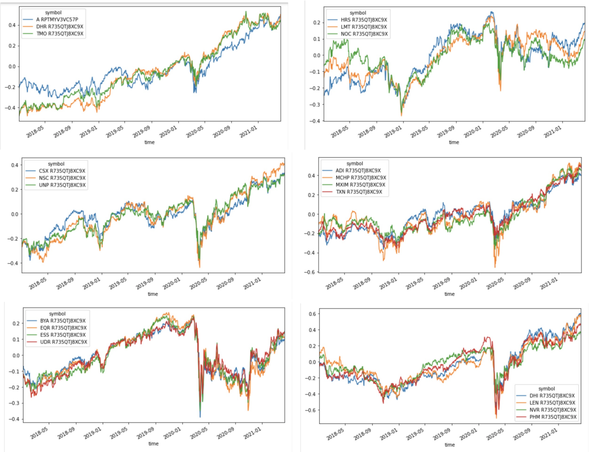


Figure: Clusters

- Correlation

Pairs	Correlaition
DHR, TMO	0.994439
DNB, MHP	0.992253
CMS, WPH	0.984816
AEE, CMS	0.976660
ESS, UDR	0.975056

Table: Ranked by Correlation

- Cointegration

Pairs	P-Value
MCHP, MXIM	0.000904
CHV, XON	0.002567
ALK, LUV	0.004910
DHR, TMO	0.007147
P, XON	0.007320

Table: Ranked by P-Value in Cointegration Test

Hurst Exponent

Hurst Exponent	Time Series
$H = 0.5$	random walk
$H < 0.5$	mean reversion
$H > 0.5$	persistent trend

Table: Hurst Exponent Criterion

Filter out pairs for which the half-life takes extreme values: less than one day or more than one year.

Dollar Volume

Symbol	Number of Days on the top 500 List
KLAC	10
CSCO	10
EBAY	10
PLTR	10
PINS	10

Table: Company Classification

Fundamental Sector

Old Index	Symbol	Liquidity Count	Sector	Sector String
38	AEP	4	207	Utilities
39	CMS	2	207	Utilities
44	AEP	4	207	Utilities
45	ETR	1	207	Utilities
80	CMS	2	207	Utilities
81	ETR	1	207	Utilities
208	DHI	10	102	ConsumerCyclical
209	LEN	6	102	ConsumerCyclical

Table: Selected Pairs from Fundamental Persepective

Stage Optimization

One method in Spread Calculation - Kalman Filter

A Three-step Process of Prediction, Observation, and Correction

corrected state = predicted state + k (observation - prediction)

- (observation - prediction) is called the observation innovation. A fraction of the observation innovation is added as a correction to the predicted state. The value of this fraction k is known as the Kalman gain.
- k is decided such that the corrected state has the least amount of error variance associated with it.
- k is indeed optimal in the case where the mathematical models of state and observation are both linear and the errors are drawn from independent Gaussian distributions.

1. Evaluate $\hat{X}_{t|t-1}$ and $\hat{P}_{t|t-1}$ using the state equation.
$$\hat{X}_{t|t-1} = A\hat{X}_{t-1|t-1}$$
$$\hat{P}_{t|t-1} = A\hat{P}_{t-1|t-1}A^T$$
2. Find the observation Y_t and R by observing the system. Note we have the matrix H defined as follows:
$$Y_t = HX_t + v_t$$
3. Compute the Kalman gain K_t.
$$K_t = \hat{P}_t H^T (H\hat{P}_t H^T + R)^{-1}$$
4. Evaluate $\hat{X}_{t|t}$ given by
$$\hat{X}_{t|t} = \hat{X}_{t|t-1} + K_t(Y_t - H\hat{X}_{t|t-1})$$
5. Evaluate $\hat{P}_{t|t}$

Stage Optimization

One method in Spread Calculation - Kalman Filter

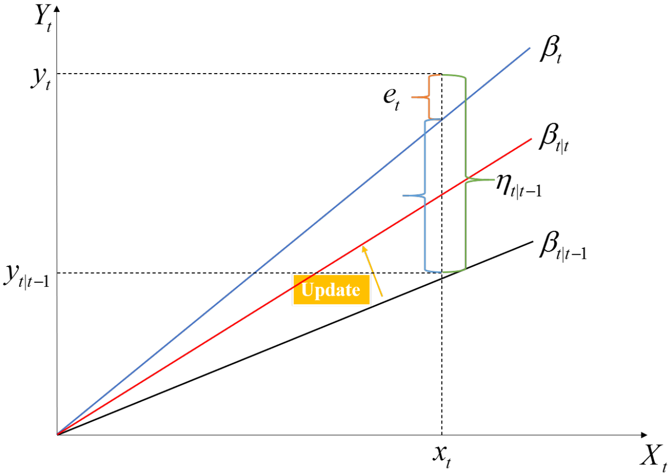


Figure: Illustration of Kalman Filter

Stage Optimization

Risk Management

1. **Margin Constraints:** 50%
2. **Stop-loss Limit:** $|Z| \geq 4$ or 15% of loss
3. **Value-at-Risk:** no more than 30,000

Stage Optimization
Performance Measurement

1. **Sharpe Ratio**

$$SR = \frac{R - R_f}{\sigma_i}$$

2. **Total Return:**

$$TR = \frac{V_t}{V_0} - 1$$

3. **Max Drawdown:**

$$MaxDD_t = max_{u \in [0,t]}(M_u - S_u)$$

Grid Research

Our optimization method is grid research. We backtest all the combinations of parameters to maximize the Sharpe Ratio with in-sample data from 2018/01/02 to 2021/04/15.

Parameter	Min	Max	Step Size
enter	1	3	0.5
exit	0	0.5	0.1

Table: Grid Research Detail

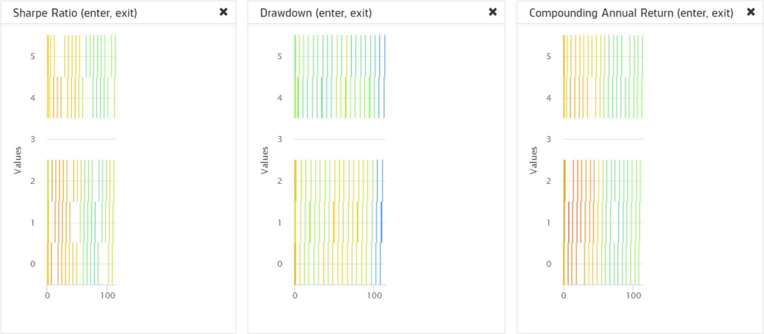


Figure: Optimization on Parameters

Stage Optimization

Optimization

Method	Sharpe Ratio	Enter	Exit	Return	MaxDD
OLS	1.427	2.0	0.1	16.479%	2.9%
Kalman	1.334	1.0	0.5	25.314%	4.0%

Table: In Sample: Optimal Back-testing Statistics on Selected SP500 Pairs

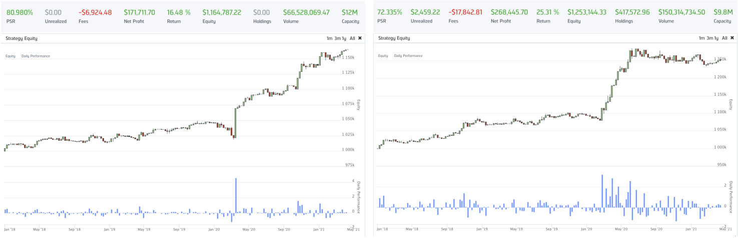


Figure: In Sample: Optimal Back-testing Result with OLS(left) Kalman(right)

Stage Optimization
Optimization

Method	Sharpe Ratio	Enter	Exit	Return	MaxDD
OLS	0.549	2.0	0.1	3.345%	3.1%
Kalman	1.485	1.0	0.5	10.420%	2.9%

Table: Out of Sample: Back-testing Statistics with Optimized Parameters

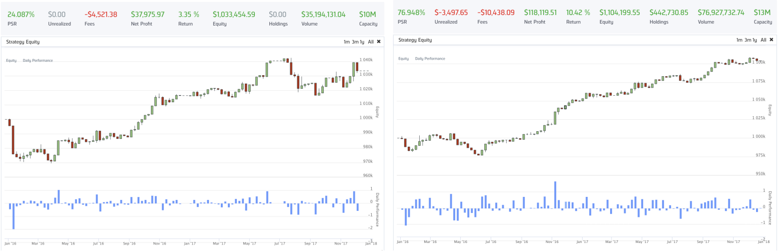


Figure: Out of Sample: Back-testing Result with OLS(left) Kalman(right)

Stage Optimization
Optimization

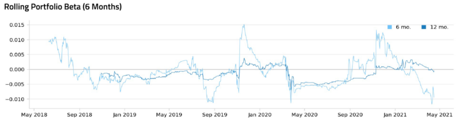


Figure: Rolling Portfolio Beta

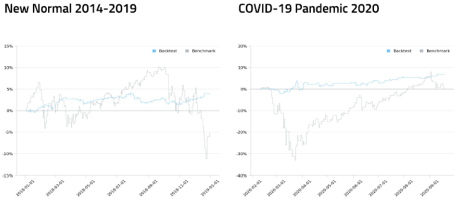


Figure: Covid-19 Pandemic 2020

Sub Strategy
Gold Trading

In order to **increase** returns and **hedge macro risks**, we have added a gold trading sub-strategy: **long** gold when the $\frac{gold}{S\&P\ 500\ Index}$ ratio is greater than the 5-day exponential moving average of the ratio and **long** S&P 500 index when the $\frac{gold}{S\&P\ 500\ Index}$ ratio is lower than the 5-day exponential moving average of the ratio.

Sub Strategy

Gold Trading

Method	Sharpe Ratio	Enter	Exit	Return	MaxDD
OLS	1.629	2.0	0.0	25.79%	5.6%
Kalman Filter	1.491	2.0	0.5	23.38%	4.1%

Table: In Sample: Optimal Back-testing Statistics on Selected SP500 Pairs with Gold Trading Strategy

Figure: In Sample: Optimal Back-testing Result with Gold Trading Strategy using OLS

Sub Strategy
Gold Trading

Method	Sharpe Ratio	Enter	Exit	Return	MaxDD
OLS	1.431	2.0	0.0	10.62%	2.9%
Kalman Filter	2.025	2.0	0.5	13.74%	2.1%

Table: Out of Sample: Back-testing Statistics with Optimized Parameters and Gold Trading Strategy

Figure: Out of Sample: Optimal Back-testing Result with Gold Trading Strategy using Kalman Filter

Live Trading

Go Live

Figure: Live Trading Result