

[Back to Community](#)

Pair Trade with Cointegration and Mean-Reversion Tests



Justin Lent posted Jul 29, 2015

mean reversion

pairs trading

Show Q

tools and tips

Attached is a pair trading algo that allows the user to toggle on/off different tests for cointegration/mean-reversion of the pair's spread prior to taking any trades. If you choose to turn on one of the tests, the value from the test is recorded as a timeseries viewable from the backtest results page.

The pair being traded in this algo is the oil and gold ETFs (USO and GLD), but you can modify these as you wish.

The 3 different tests are:

- Augmented Dickey-Fuller test p-value
 - Effectively, this is a unit-root test for determining whether the spread is cointegrated
 - As well, a function is included showing how to use the critical-values from the ADF-test instead of p-value
- Half-life of mean-reversion, computed from an Ornstein-Uhlenbeck process
 - This is the theoretically computed time, based on a historical window of data, that it will take for the spread to mean-revert half of its distance after having diverged from the mean of the spread
- Hurst exponent
 - Effectively this returns a value between 0 and 1 that tells you whether a time-series is trending or mean-reverting. The closer the value is to 0.5 means the more "random" the time-series has behaved historically. Values below 0.5 imply the time-series is mean-reverting, and above 0.5 imply trending. The closer the value is to 0 implies greater levels of mean-reversion.
 - Trading literature is conflicted as to the usefulness of Hurst exponent, but I included it nonetheless, and have set the default switch to False in the algo.

The backtest results below incorporate two of these tests:

- ADF-test p-value, computed over a 63-day (e.g. 3-months) lookback window, with a required minimum p-value of 0.20
- Half-life days, computed over 126-day (6-month) lookback window, with a requirement of the half-life being between 1 and 42-days (2-months)

To modify the parameter values of the tests just look in the initialize function, for blocks of code that look like this. Here is how the ADF-test p-value parameters are defined:

```
context.stat_filter = {}
context.stat_filter['adf_p_value'] = {}

context.stat_filter['adf_p_value']['use'] = True
context.stat_filter['adf_p_value']['lookback'] = 63
context.stat_filter['adf_p_value']['function'] = adf_p_value
context.stat_filter['adf_p_value']['test_condition_min'] = 0.0
context.stat_filter['adf_p_value']['test_condition_max'] = 0.20
```

Here you see how there is a dictionary defined called 'stat_filter' which you can use to store the parameters of each test. First I create another dictionary inside of 'stat_filter' named 'adf_p_value' and then I load in all of the parameter values relevant to the ADF-test that I want to define when it is acceptable to enter a trade. These exact 5 parameters (e.g. keys of the dictionary) will be defined for all of the tests, as you'll see if you look at the algo code, and notice the adf_critical_value, half_life, hurst_exponent ones are defined following it. The 5 parameters are:

- 'use': Boolean, True if you want the algo to use this test
- 'lookback': Integer, value of how many lookback periods of the timeseries to be used in running the computation
- 'function': Function, this is the name of the function that will be called and return a value back to be compared to the _min and _max conditions below
- 'test_condition_min': Integer or Float, the minimum value returned by 'function' to determine if a trade can be triggered
- 'test_condition_max': Integer or Float, the maximum returned by 'function' to determine if a trade can be triggered

Support for Intraday Frequency

(Let me know if you run into issues with this, as I haven't done as much testing with it as I have with just daily freq)

You can configure this algo to be run on intraday minutely data as well. E.g. construct a pair spread using 15-min bar closing prices.

First, change the variable 'context.trade_freq' from 'daily' to 'intraday':

```
context.trade_freq = 'daily' # 'daily' or 'intraday'
```

Then, look for this code block below in the initialize() function, and specify the 'intraday_freq' value for the frequency of closing prices to use (E.g. 15 minute bars). Then, set 'run_trading_logic' to be how frequently you want the logic to be applied to market data. I chose 60 which means, run this logic every 60-minutes, but if you wish, change it to 1, and the logic will be run every single minute (beware though, as this will result in really long backtest times).

The variable 'check_exit_every_minute' can be set to True if you want the logic to be run *every minute if-and-only-if* you are currently in a trade. E.g. it checks to see whether you need to exit the trade every minute rather than waiting to the next N periods (e.g. 60 minutes, as specified in the 'run_trading_logic_freq' variable)

```
### START: INTRADAY FREQUENCY PARAMETERS
context.intraday_freq = 15          # only used if context.trade_freq='intraday'
context.run_trading_logic_freq = 60 # only used if context.trade_freq='intraday'
context.check_exit_every_minute = False # if True, and if in a trade, will check every minute whether to exit
### END: INTRADAY FREQUENCY PARAMETERS
```

Performance Source Code Migrated Source Code

```
# Backtest ID: 55b982b8bedc2f0c743a86d9
1 import numpy as np
2 import statsmodels.api as sm
3 import statsmodels.tsa.stattools as ts
4 import pandas as pd
5 from zipline.utils import tradingcalendar
6 import pytz
7
8
9 def initialize(context):
10     '''
11     START: USER CONFIGURABLE PARAMETERS FOR THE BACKTEST
12     '''
13     # Quantopian backtester specific variables
14     set_slippage(slippage.VolumeShareSlippage(volume_limit=0.025, price_impact=0.1))
15     set_commission(commission.PerShare(cost=0.0075, min_trade_cost=0.0))
16     #set_slippage(slippage.VolumeShareSlippage(volume_limit=0.99, price_impact=0.0))
17     #set_commission(commission.PerShare(cost=0.00, min_trade_cost=0.0))
18
19     # the ticker symbols for your pair
20     context.y = symbol('USO')
21     context.x = symbol('GLD')
22
23     # strategy specific variables
24     context.lookback = 20 # used for regression
25     context.z_window = 20 # used for zscore calculation, must be <= lookback
26     context.entry_z = 0.5 # trade entry triggered when spread is + or - entryZ
27     context.exit_z = 0.0 # trade exit triggered when spread is + or - entryZ
28     context.momentum_or_mean_reversion = 'mean-reversion' # 'momentum' or 'mean-reversion'
29     # This defines whether to use a hedge ratio computed N periods ago.
30     # Rationale for this is that since the algo is trading off of mean reversion
31     # that a hedge ratio that excludes N days of recency, e.g. when severe divergences
32     # could have occurred, and which this algo hopes to exploit, may be more aligned
33     # to the economic historical relationship befitting of the stock pair
34     context.use_hedge_ratio_lag = True
35     context.hedge_ratio_lag = 2
36
37
38     ##### START: TRADING FREQUENCY DEFINITIONS #####
39     context.trade_freq = 'daily' # 'daily' or 'intraday'
40
41     ### START: DAILY FREQUENCY PARAMETERS
42     # if context.trade_freq == 'intraday' then specify these inputs
43     # only run the trading logic at 3:30 Eastern time, which 30-minutes before market closes
44     context.daily_run_logic_hours = 15 # only used if context.trade_freq='daily'
45     context.daily_run_logic_minutes = 30 # only used if context.trade_freq='daily'
46     ### END: DAILY FREQUENCY PARAMETERS
47
48     ### START: INTRADAY FREQUENCY PARAMETERS
49     # if context.trade_freq == 'intraday' then specify these inputs
50     context.intraday_freq = 15 # only used if context.trade_freq='intraday'
51     context.run_trading_logic_freq = 60 # only used if context.trade_freq='intraday'
52     context.check_exit_every_minute = False # if True, and if in a trade, will check every minute whether to exit
53     ### END: INTRADAY FREQUENCY PARAMETERS
54
```

```

55 ##### END: TRADING FREQUENCY DEFINITIONS #####
56
57
58 ##### START: STATISTICAL FILTERS DEFINITIONS #####
59 # Specify 'statistical filters' (e.g. cointegration, half-life of mean reversion, etc)
60 # that the pair's spread must pass before before it will be considered for a trade
61 context.stat_filter = {}
62 context.stat_filter['adf_p_value'] = {}
63 context.stat_filter['adf_critical_value'] = {}
64 context.stat_filter['half_life_days'] = {}
65 context.stat_filter['hurst_exponent'] = {}
66
67 context.stat_filter['adf_p_value']['use'] = True
68 context.stat_filter['adf_p_value']['lookback'] = 63
69 context.stat_filter['adf_p_value']['function'] = adf_p_value # function defined below
70 context.stat_filter['adf_p_value']['test_condition_min'] = 0.0
71 context.stat_filter['adf_p_value']['test_condition_max'] = 0.20
72
73 context.stat_filter['adf_critical_value']['use'] = False
74 context.stat_filter['adf_critical_value']['lookback'] = 63
75 context.stat_filter['adf_critical_value']['function'] = adf_critical_value_test # function defined below
76 context.stat_filter['adf_critical_value']['test_condition_min'] = int(0) # see function description for return val
77 context.stat_filter['adf_critical_value']['test_condition_max'] = int(10) # see function description for return val
78
79 context.stat_filter['half_life_days']['use'] = True
80 context.stat_filter['half_life_days']['lookback'] = 126
81 context.stat_filter['half_life_days']['function'] = half_life # function defined below
82 context.stat_filter['half_life_days']['test_condition_min'] = 1.0
83 context.stat_filter['half_life_days']['test_condition_max'] = 42.0
84
85 context.stat_filter['hurst_exponent']['use'] = False
86 context.stat_filter['hurst_exponent']['lookback'] = 126
87 context.stat_filter['hurst_exponent']['function'] = hurst # function defined below
88 context.stat_filter['hurst_exponent']['test_condition_min'] = 0.0
89 context.stat_filter['hurst_exponent']['test_condition_max'] = 0.4
90 ##### END: STATISTICAL FILTERS DEFINITIONS #####
91
92 '''
93 END: USER CONFIGURABLE PARAMETERS
94 '''
95
96 # define a few more global variables based on the inputs above
97 context.intraday_history_lookback = context.lookback * context.intraday_freq + 10
98 context.spread = np.array([])
99 context.hedge_ratio_history = np.array([])
100 context.in_long = False
101 context.in_short = False
102
103 if not context.use_hedge_ratio_lag:
104     # a lag of 1 means to include the most recent price in the hedge_ratio calculation
105     # specifically, this is used for np.array[-1] indexing
106     context.hedge_ratio_lag = 1
107
108 # Will be called on every trade event for the securities you specify.
109 def handle_data(context, data):
110     if get_open_orders():
111         return
112     now = get_datetime()
113     exchange_time = now.astimezone(pytz.timezone('US/Eastern'))
114
115     # Only trade N-minutes before market close
116     if context.trade_freq == 'daily':
117         if not (exchange_time.hour == context.daily_run_logic_hours and exchange_time.minute == context.daily_run_logic_minutes):
118             return
119
120     if context.trade_freq == 'intraday':
121         # Only run trading logic every N minutes
122         if exchange_time.minute % context.run_trading_logic_freq > 0:
123             return
124
125     if context.trade_freq == 'daily':
126         prices = history(context.lookback, '1d', 'price').iloc[-context.lookback:]
127
128     if context.trade_freq == 'intraday':
129         prices = history(context.intraday_history_lookback, '1m', 'price')
130         indexes_at_freq = range(prices.shape[0]-1, 0, -context.intraday_freq)
131         indexes_at_freq.sort()
132         prices = prices.ix[ indexes_at_freq, :]
133
134     y = prices[context.y]
135     x = prices[context.x]
136
137     try:

```

```

138     hedge = hedge_ratio(y, x, add_const=True)
139 except ValueError as e:
140     log.debug(e)
141     return
142
143 context.hedge_ratio_history = np.append(context.hedge_ratio_history, hedge)
144 # Calculate the current day's spread and add it to the running tally
145 if context.hedge_ratio_history.size < context.hedge_ratio_lag:
146     return
147 # Grab the previous day's hedge ratio
148 hedge = context.hedge_ratio_history[-context.hedge_ratio_lag]
149 context.spread = np.append(context.spread, y[-1] - hedge * x[-1])
150 spread_length = context.spread.size
151 #record(spd=context.spread[-1])
152
153 # apply all the statistical filters to 'context.spread', if specified in initialize()
154 # only test for trading triggers if the spread passes all specified criteria
155 for stat_name in context.stat_filter:
156     if context.stat_filter[stat_name]['use']:
157         test_lookback = context.stat_filter[stat_name]['lookback']
158         if spread_length < test_lookback:
159             return
160         test_spread = context.spread[-test_lookback:]
161         test_func = context.stat_filter[stat_name]['function']
162         test_value = test_func(test_spread)
163         test_min = context.stat_filter[stat_name]['test_condition_min']
164         test_max = context.stat_filter[stat_name]['test_condition_max']
165
166         if stat_name == 'adf_p_value':
167             record(adf_p=test_value)
168         if stat_name == 'adf_critical_value':
169             record(adf_cv=test_value)
170         if stat_name == 'half_life_days':
171             record(half_life=test_value)
172         if stat_name == 'hurst_exponent':
173             record(hurst=test_value)
174         if (test_value < test_min) or (test_value > test_max):
175             return
176
177 if context.spread.size > context.z_window:
178     # Keep only the z-score lookback period
179     spreads = context.spread[-context.z_window:]
180
181     zscore = (spreads[-1] - spreads.mean()) / spreads.std()
182     record(Z=zscore)
183     #record(gr_lev=context.account.leverage, net_lev=context.account.net_leverage)
184
185     if context.in_short and zscore < context.exit_z:
186         order_target(context.y, 0)
187         order_target(context.x, 0)
188         context.in_short = False
189         context.in_long = False
190         #record(in_trade=0)
191         #record(stock_Y_pct=0, stock_X_pct=0)
192         return
193
194     if context.in_long and zscore > context.exit_z:
195         order_target(context.y, 0)
196         order_target(context.x, 0)
197         context.in_short = False
198         context.in_long = False
199         #record(in_trade=0)
200         #record(stock_Y_pct=0, stock_X_pct=0)
201         return
202
203
204
205 # only check for new trades every N minutes, but can exit existing trades _each_ minute
206 # which is why this check is here. Trade exits are handled above.
207 #if exchange_time.minute % context.run_trading_logic_freq > 0:
208 #    return
209
210 if zscore < -context.entry_z and (not context.in_long):
211     # Only trade if NOT already in a trade
212     y_target_shares = 1
213     x_target_shares = -hedge
214     context.in_long = True
215     context.in_short = False
216
217     (y_target_pct, x_target_pct) = compute_holdings_pct(y_target_shares,
218                                                         x_target_shares,
219                                                         y[-1], x[-1] )
220

```

```

221         if context.momentum_or_mean_reversion == 'momentum':
222             y_target_pct = -1.0 * y_target_pct
223             x_target_pct = -1.0 * x_target_pct
224
225         order_target_percent(context.y, y_target_pct)
226         order_target_percent(context.x, x_target_pct)
227         #record(in_trade=1)
228         #record(stock_Y_pct=y_target_pct, stock_X_pct=x_target_pct)
229         return
230
231     if zscore > context.entry_z and (not context.in_short):
232         # Only trade if NOT already in a trade
233         y_target_shares = -1
234         x_target_shares = hedge
235         context.in_short = True
236         context.in_long = False
237
238         (y_target_pct, x_target_pct) = compute_holdings_pct(y_target_shares,
239                                                             x_target_shares,
240                                                             y[-1], x[-1] )
241
242         if context.momentum_or_mean_reversion == 'momentum':
243             y_target_pct = -1.0 * y_target_pct
244             x_target_pct = -1.0 * x_target_pct
245
246         order_target_percent(context.y, y_target_pct)
247         order_target_percent(context.x, x_target_pct)
248         #record(in_trade=1)
249         #record(stock_Y_pct=y_target_pct, stock_X_pct=x_target_pct)
250
251 def is_market_close(dt):
252     ref = tradingcalendar.canonicalize_datetime(dt)
253     return dt == tradingcalendar.open_and_closes.T[ref]['market_close']
254
255 def hedge_ratio(y, x, add_const=True):
256     if add_const:
257         x = sm.add_constant(x)
258         model = sm.OLS(y, x).fit()
259         return model.params[1]
260     model = sm.OLS(y, x).fit()
261     return model.params.values
262
263 def compute_holdings_pct(y_shares, x_shares, y_price, x_price):
264     y_dollars = y_shares * y_price
265     x_dollars = x_shares * x_price
266     notional_dollars = abs(y_dollars) + abs(x_dollars)
267     y_target_pct = y_dollars / notional_dollars
268     x_target_pct = x_dollars / notional_dollars
269     return (y_target_pct, x_target_pct)
270
271 def adf_p_value(input_ts):
272     # returns p-value from Augmented Dickey-Fullet test for cointegration, with lag=1
273     return ts.adfuller(input_ts, 1)[1]
274
275 def adf_critical_value_test(input_ts):
276     # returns 1: if the t-stat of the ADF-test is less than the 1% critical value
277     # returns 5: if the t-stat of the ADF-test is less than the 5% critical value (but greater than the 1% level)
278     # returns 10: if the t-stat of the ADF-test is less than the 10% critical value (but greater than the 1% and 5% levels)
279     # return 99: if the t-stat of the ADF-test is greater than the 10% critical value
280     adf_test = ts.adfuller(input_ts, 1)
281     t_stat = adf_test[0]
282     critical_values = adf_test[4]
283     if t_stat < critical_values['1%']:
284         return int(1)
285     if t_stat < critical_values['5%']:
286         return int(5)
287     if t_stat < critical_values['10%']:
288         return int(10)
289     return int(99)
290
291 def half_life(input_ts):
292     # returns the [theoretical, based on OU-process equations] number of periods to expect
293     # to have to wait for the spread to mean-revert half the distance to its mean
294     price = pd.Series(input_ts)
295     lagged_price = price.shift(1).fillna(method="bfill")
296     delta = price - lagged_price
297     beta = np.polyfit(lagged_price, delta, 1)[0]
298     half_life = (-1*np.log(2)/beta)
299     return half_life
300
301 def hurst(input_ts, lags_to_test=20):
302     # interpretation of return value
303     # hurst < 0.5 - input_ts is mean reverting

```

```

304 # hurst = 0.5 - input_ts is effectively random/geometric brownian motion
305 # hurst > 0.5 - input_ts is trending
306 tau = []
307 lagvec = []
308 # Step through the different lags
309 for lag in range(2, lags_to_test):
310     # produce price difference with lag
311     pp = np.subtract(input_ts[lag:], input_ts[:-lag])
312     # Write the different lags into a vector
313     lagvec.append(lag)
314     # Calculate the variance of the difference vector
315     tau.append(np.sqrt(np.std(pp)))
316 # linear fit to double-log graph (gives power)
317 m = np.polyfit(np.log10(lagvec), np.log10(tau), 1)
318 # calculate hurst
319 hurst = m[0]*2
320 return hurst

```

DISCLAIMER

The material on this website is provided for informational purposes only and does not constitute an offer to sell, a solicitation to buy, or a recommendation or endorsement for any security or strategy, nor does it constitute an offer to provide investment advisory services by Quantopian. In addition, the material offers no opinion with respect to the suitability of any security or specific investment. Quantopian makes no guarantees as to the accuracy or completeness of the views expressed in the website. The views are subject to change, and may have become unreliable for various reasons, including changes in market conditions or economic circumstances. All investments involve risk, including loss of principal. You should consult with an investment professional before making any investment decisions.

17 responses



Vladimir Yevtushenko Jul 30, 2015

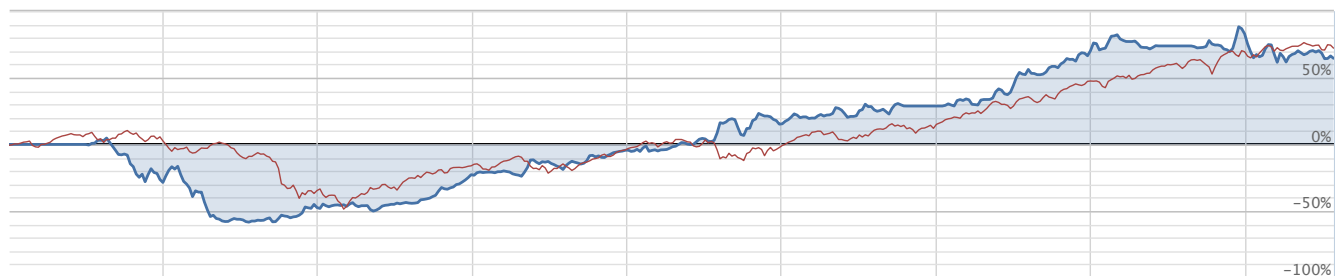
Same algo just start 9 month earlier.

Performance

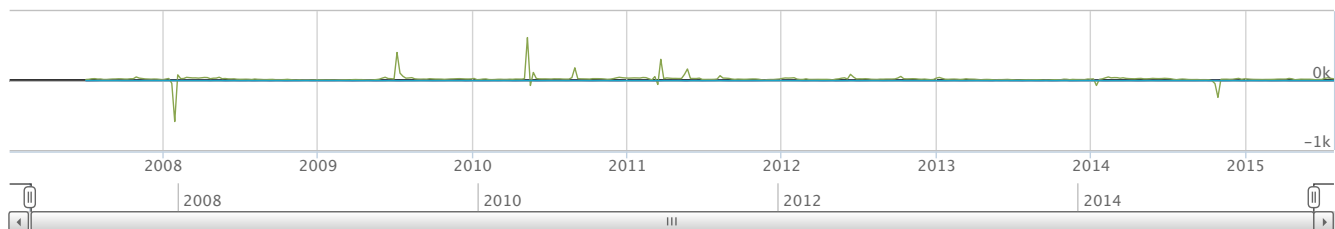
Source Code

Backtest from **2007-01-02** to **2015-07-28** with **\$100,000** initial capital

Algorithm Benchmark (SPY)



half_life Z adf_p



Jamie Lunn Sep 13, 2015

Hey Justin,

Thanks for the share. I've noticed that there is a coint function in statsmodels.tsa.stattools. Is there a significant difference between the coint function and the ADF test? Any sense in using both?

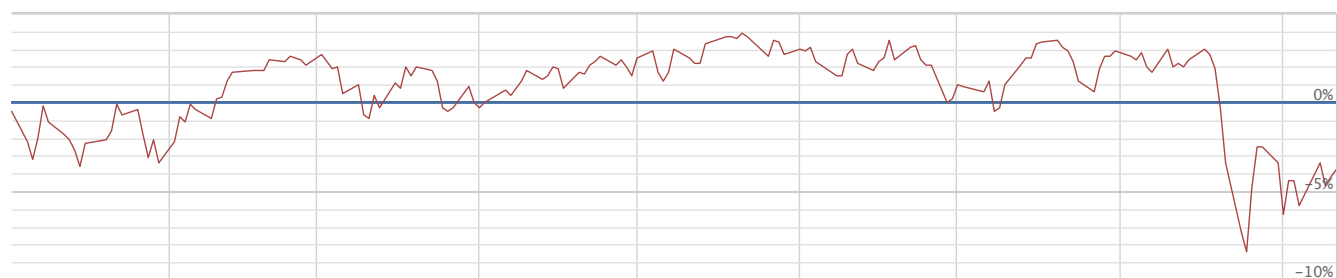
I've attached a backtest below that attempts to find the pvalue of both tests for each pair, every day. Disclaimer: what I often think is happening in python is actually not.

Performance

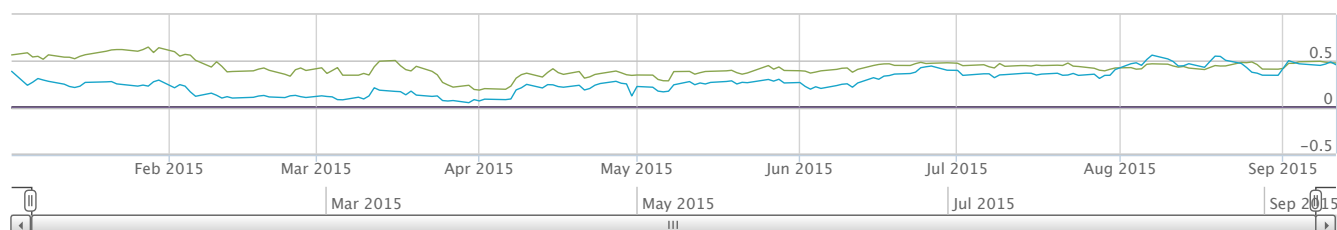
Source Code


Backtest from **2015-01-01** to **2015-09-11** with **\$100,000** initial capital

Algorithm **Benchmark (SPY)**



coint_p **lev** **adf_p**




 Justin Lent Sep 25, 2015

@Jamie,

I haven't tried the coint function in stattools yet, though I imagine it's very similar. I just took a quick glimpse at the code, and it's effectively running a regression of the lagged version of the input timeseries versus the unlagged version which is quite similar to ADF. The difference may lie in how the critical values are computed.

The Engle-Granger test is also sometimes used to test for co-integration, but I haven't looked at that implementation yet.

DISCLAIMER

 Adeel Imtiaz Sep 26, 2015

Thumbs up (y)

 Johy Reaper Sep 26, 2015

Great Algo. Its amazing. Very Helpful

 Adam Preston Sep 26, 2015

Hello Justin / All

Could you suggest how I can run this algo on multiple pairs, rather than just one pair?

Thanks!
Preston




James Christopher Sep 26, 2015

Try making a pairs trading class that keeps track of all the bookkeeping for each given pair. See David's generalized Kalman filters pair trading algo for a great example of class based pairs trading


DISCLAIMER

The material on this website is provided for informational purposes only and does not constitute an offer to sell, a solicitation to buy, or a recommendation or endorsement for any security or strategy, nor does it constitute an offer to provide investment advisory services by Quantopian. In addition, the material offers no opinion with respect to the suitability of any security or specific investment. Quantopian makes no guarantees as to the accuracy or completeness of the views expressed in the website. The views are subject to change, and may have become unreliable for various reasons, including changes in market conditions or economic circumstances. All investments involve risk, including loss of principal. You should consult with an investment professional before making any investment decisions.



Adeel Imtiaz Sep 27, 2015

Thanks for sharing info




Adam Preston Sep 27, 2015

Hi all,
I cloned Justin's algo, however when I run a backtest, the performance remains at 0% for the entirety of the backtest window.
I made no changes to the original source code.

Any ideas why this would be occurring?

Thanks!

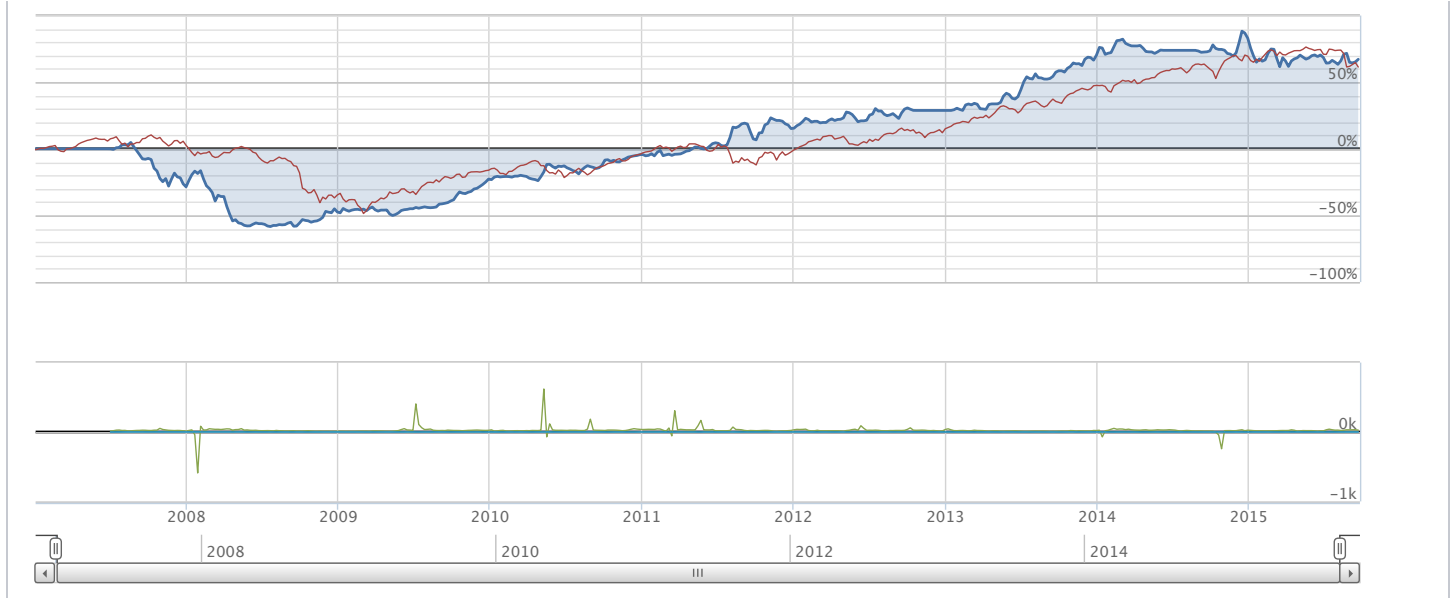



Vladimir Yevtushenko Sep 27, 2015

Adam,

You probably run algo in daily mode and it only work in minute mode.
Here is my latest backtest of original Justin's Lent algo started just 9 month earlier.

Performance	Source Code
Backtest from 2007-01-02 to 2015-09-25 with \$100,000 initial capital	
Algorithm	Benchmark (SPY)
half_life	Z
	adf_p



 Justin Lent Sep 27, 2015

All,

It's worth noting that when I post backtests, code, and research notebooks, the intent is to illustrate a methodology, and provide some code templates to spur the creative thought process of the community and save folks some time by providing cut-and-paste code fragments that can be integrated into their own code. By no means am I posting something that has been fully vetted, and immediately investable in it's exact form, by any stretch of the imagination. I often bias for simpler, rather than overly complex, examples as well, so as to benefit a broader spectrum of readers.

Vladimir,

I see you've recognized that the backtest I posted above seems to fail pretty badly over a different timeframe. We see this a lot with strategies we look at, many of which are overfit to just the 2 year period in the contests we run. We try to work with the algo owner and provide advice as to why it may have broken down over the different timeframes. Perhaps you can extend your analysis to provide me some advice as to improving this strategy? Maybe you have some recommendations as to how to incorporate a regime switching model which is very likely to help a strategy such as this given the time frame it seems to fail (the financial/commodity futures crises that occurred in late 2008). Perhaps a stochastic volatility regime switching model might help significantly. If you have experience in this area I'm sure the community would find it a solid addition to incorporate into strategies such as these to make them more robust. I know I would.

DISCLAIMER


 Adam Preston Sep 27, 2015

Justin,

Why did you choose the pair USO and GLD? I guess a broader question is can you suggest a process for scanning through a basket of stocks and determining if there are tradable pairs? I'm assuming tests for cointegration would be one method such as ADF as you used. It would be nice if there could be an algo to run through a basket of stocks and auto determine which would make "good" pairs.

Just a thought.

Adam

 Justin Lent Sep 27, 2015

Hey Adam,

I just chose USO/GLD in order to replicate this example that uses those same tickers, from this book: <http://www.amazon.com/Quantitative-Trading-Build-Algorithmic-Business/dp/0470284889/>

That book is a really good intro to stat arb pair trading (as well as his other books). All the code in the book is in Matlab, so my algo was an attempt to implement it in Python, in our backtester, and incorporate some of the other statistical techniques described throughout the book.

You are correct, that screening a bunch of potential pairs is a reasonable research idea, but you should be cognizant of simply datamining. You first want to determine a sensible economic basis for which the pairs of stocks should be tied (e.g. pairs of stocks in the same sector would be reasonable pairs of stocks to search across). Writing an algo in our backtester to accomplish this would be fairly straightforward: First you can use our Morningstar fundamentals database to grab all stocks in the Energy sector, perhaps even filtering down to stocks of companies of a certain band of marketcap (e.g. only mid-cap energy stocks), then in `before_trading_starts()`, you loop over each stock pair computing the ADF p-value (or other cointegration stat), keep all the stock pairs that meet your criteria, and then in `handle_data()` you just run the ones that meet the criteria through an algo similar to the one I shared to enter/exit the trades.

Myself or someone on our team here at Q can try to develop a template for this and share it.

As well you can look at this forum post that shows how to develop a single algo that trades a portfolio of multiple pairs:

It's the algo backtest in the first comment from David Edwards, here:


<https://www.quantopian.com/posts/quantopian-lecture-series-this-time-youre-more-wrong?c=1>

DISCLAIMER

 Adam Preston Apr 6, 2016

Justin,

I noticed in the blog section you have a notebook on using a Bayesian optimizer...would you know how i can pull it into Q? its currently on github..thanks!


 Justin Lent Apr 7, 2016

@Adam, At present it's not possible to use the Bayesian optimizer from the blog post in the Q environment. It was more of a proof of concept implementation idea. As you mentioned, the code I used for the blog post is on github and you can sign up for a trial with SigOpt to get a username/API key to work with it in your own python/zipline environment locally. Offering some of these alternative methods of optimization as a service is an interesting concept which we will have to think about as we develop our Q platform in the future. Thanks for the feedback!

DISCLAIMER

 Adam Preston Apr 7, 2016

Thanks Justin! Would be neat to be able to do that type of optimization and / or a particle swarm technique in Q. :)

 Shawn Emhe II Nov 27, 2016

Hello Justin,

I believe I found a gap in the trading logic. In the statistic filtering section (lines ~155-176) the algorithm immediately exits if a test fails. That prevents new trades from being opened but does nothing to handle existing trades. Open trades stay open until all of the statistical tests pass again and the algorithm reaches its standard exit logic.

By design we should also have a high likelihood of being in a trade when this happens so the impact could be quite high. The problem in detecting this is that if the relationship re-establishes quickly the performance won't suffer. But if we include a time period in which the relationship doesn't return quickly, as Vladimir did, the results are noticeable.

I added a few lines to close any positions that are open when the statistical tests break down. There are probably better ways of handling the exit logic, but this simple change shows the benefit of having it there. The algorithm doesn't do as well during the original test period but the performance improves over the extended period.

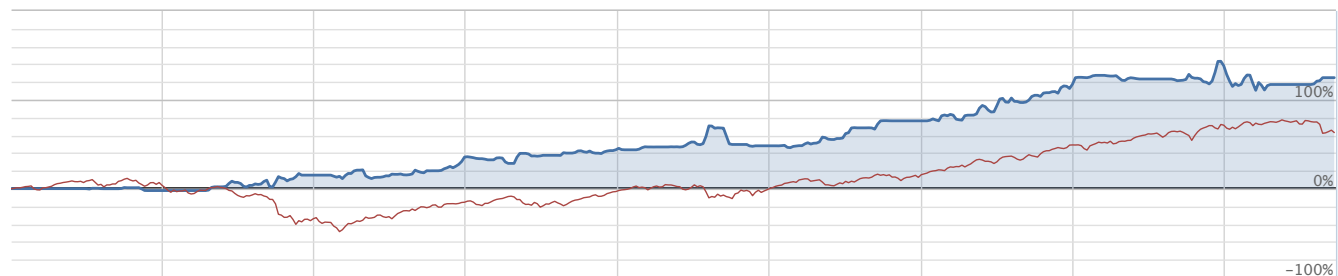
(I also made minor change on lines 20 and 21 to use `sid()` function to set x and y assets rather than `symbol()`. The rest of the algorithm is unchanged.)

Performance

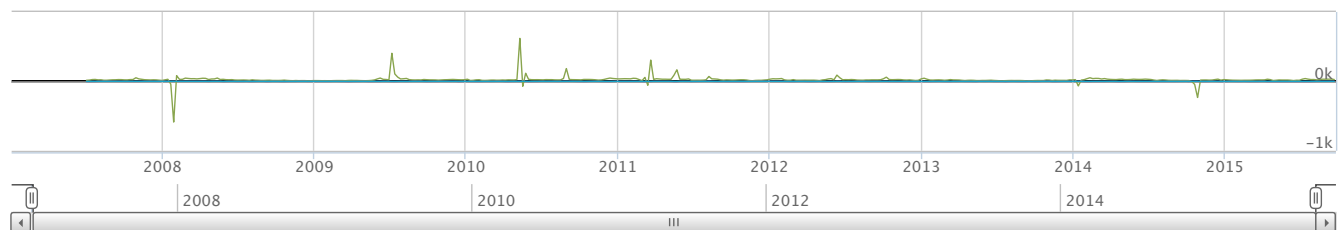
Source Code

Backtest from **2007-01-02** to **2015-09-25** with **\$100,000** initial capital

Algorithm **Benchmark (SPY)**



half_life **Z** **adf_p**



Please sign in or join Quantopian to post a reply.

About Quantopian

Careers

Community

Events

QuantCon

Help

Academia

Lectures

Workshops

Investor Relations

Status

Facebook

Twitter

LinkedIn

[Blog](#)

[Terms of Use](#)

[Privacy Policy](#)

The material on this website is provided for informational purposes only and does not constitute an offer to sell, a solicitation to buy, or a recommendation or endorsement for any security or strategy, nor does it constitute an offer to provide investment advisory or other services by Quantopian. In addition, the content of the website offers no opinion with respect to the suitability of any security or any specific investment.

Quantopian makes no guarantees as to accuracy or completeness of the views expressed in the website. The views are subject to change, and may have become unreliable for various reasons, including changes in market conditions or economic circumstances. All investments involve risk – including loss of principal. You should consult with an investment professional before making any investment decisions.