

Kalman Filter-Based ETF Pairs Trading Strategy

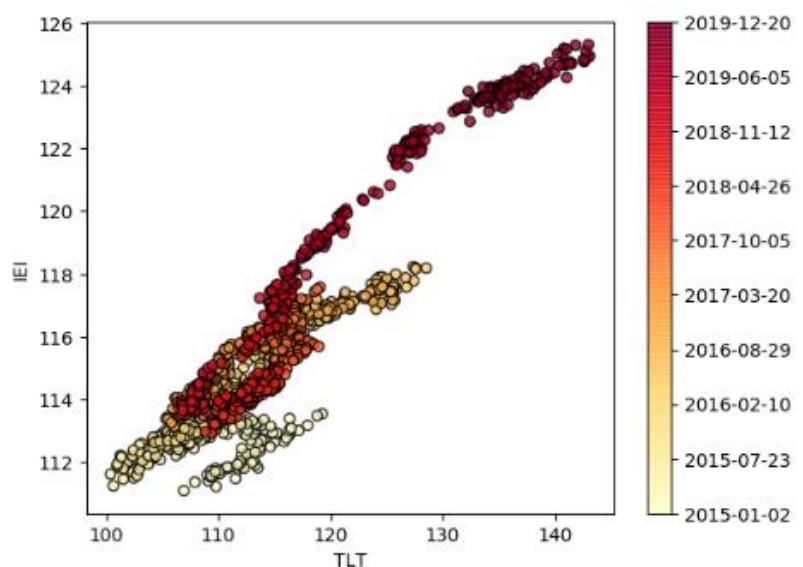
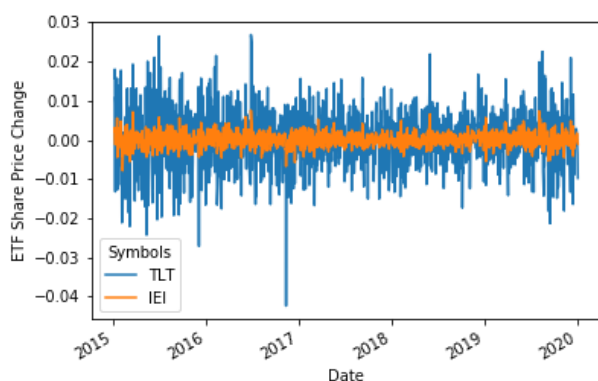
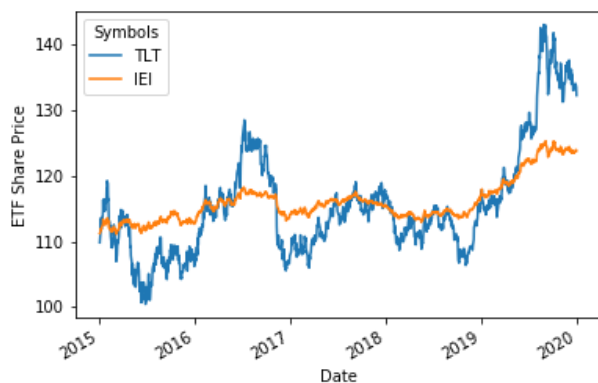
все данные и отчеты лежат на гитхабе https://github.com/romankop/CMF_Test_Case в папке “~”

В своем эссе я рассмотрю стратегию парного трейдинга, где в качестве базовых активов будут выступать ETF фонды, вкладывающиеся в казначейские облигации. Рассматриваемая стратегия была предложена в книге Advanced Algorithmic Trading за авторством Michael L. Halls-Moore и основана на результатах Ernest Chan в книге Algorithmic Trading: Winning Strategies and Their Rationale. Реализация этой стратегии приближена к оригинальному источнику, поэтому может не совпадать с принятыми методами реализации парного трейдинга. Стратегия дневная и торгуется на американской бирже, так как оба актива, а также бенчмарк торгуются на NASDAQ.

В качестве активов выбраны акции двух ETF, которые вкладываются в казначейские облигации США разной длительности.

- TLT: iShares 20+ Year Treasury Bond ETF (инвестиции в облигации с оставшимся сроком длительности от 20 лет)
- IEI: iShares 3–7 Year Treasury Bond ETF (инвестиции в облигации с оставшимся сроком длительности от 3 до 7 лет)

Для парного трейдинга обычно выбираются активы, которые подвержены коинтеграции и в которых прослеживается зависимость процессов. Кроме того, они должны быть подвержены влиянию схожих экономических или любых других внешних факторов. Кроме того, очень удобно использовать активы, поведение которых и реакция на внешнюю среду достаточно устойчивая во времени. Акции ETF, вкладывающиеся в казначейские облигации, подходят под это определение. Среднесрочные и долгосрочные облигации обычно устойчивы в поведении и подвержены воздействиям одних и тех же факторов, обычно, макроэкономических, причем обычно на эти факторы они реагируют сонаправленно. Как мы видим на графике цен и процентного изменения цен этих активов на тренировочной выборке, акции ведут себя примерно одинаково, аналогично меняя свою динамику, из чего может следовать, что они примерно одинаково реагируют на среду. Также акции фонда долгосрочных облигаций более волатильны и сильнее падают в момент общего падения и быстрее растут. Судя по графику, можно также предположить, что они и сильнее реагируют на шоки.



Если мы посмотрим на совместное распределение цен, то заметим, что у цен активов прослеживается заметная линейная связь, причем тренд вырос в последние года выборки. Ряды явно связаны. Также я провел тест на коинтеграцию рядов Энгла-Грейнджера, в котором я отверг нулевую гипотезу об отсутствии коинтеграции на уровне 10% для цен и на любом уровне для дневных доходностей. Поэтому можно заключить, что коинтеграция у этих рядов присутствует.

Идея парного трейдинга заключается в том, чтобы использовать устойчивую статистическую связь между активами. Если две акции положительно скоррелированы, то имея определенную степень их соотношения можно зарабатывать на расхождении этих акций от степени их соотношения. Так, если одна из акций серьезно вырастает относительно второй и станет переоцененной, то открывая длинную позицию на недооцененную акцию и короткую на переоцененную, можно получить прибыль из-за схождения акций к установленной степени своего отношения. Оценивать степень отношения акций можно через корреляцию, но в книге используется линейная регрессия, поэтому мы ее и рассмотрим.

Линейная регрессия строится по показателю одной из акций на другую. Показателем может выступать, например, цена акции или доходность акции. Предпочтительно использовать второе, так как доходности скорее похожи на стационарный процесс и скорее будут следовать коинтеграции. Тем не менее, в книге стратегия использовали непосредственно цены на акции ETF. Можно представить связь между показателями акций как:

$$y_1 = \theta * y_2 + \gamma + e_t,$$

где y_1 — показатель первого актива (акции для простоты), y_2 — показатель второго актива (акции), а e_t — нормально распределенный случайный шум. В этом случае коэффициент θ как раз будет являться показателем соотношения между акциями, так как устанавливает количественную меру связи между показателями двух акций. Он же и определяет пропорцию, степень хеджирования, в которой надо открывать позицию второго актива относительно первого, противоположную, как отмечалось выше. Это верно, так как в такой пропорции ожидаемая доходность портфеля из двух акций будет равна нулю, а при значительном отклонении от ожидаемого соотношения одной из акций из-за обратного схождения к тренду можно извлечь из стратегии прибыль. Эта степень хеджирования тогда означает, что при сделке на k единиц первого актива, мы совершаем сделку на $[\theta * k]$ акций второго актива.

В моем случае, к слову, коинтеграция будет означать, что ряд из ошибок стационарен.

По сути, для определения возможности для дохода мы смотрим на реализованную ошибку прогноза. Если она значимо больше нуля, то мы считаем первую акцию переоцененной, так как она стоит сильно больше ожидаемого прогноза по второй акции (равно как отклоняется от ожидаемого соотношения между ними), поэтому можно ожидать конвергенцию цен, и надо продавать первую акцию и покупать вторую. Аналогично для случая, если ошибка значимо меньше нуля. Значимость определяется границами, вне которых значение ошибки может считаться полноценным торговым сигналом. Можно взять как фиксированные, так и границы, придуманные на основе любых эмпирик.

По сути, мы торгуем этой ошибкой. А, ошибка, умноженная на -1, по сути, показывает вес этого актива (ошибки) в портфолио, так как чем высокая положительная по величине ошибка приводит к продаже разности между первой акцией и прогноза первой акции по второй, ровно как продаже первой акции и покупке второй, ровно как продаже ошибки (отрицательный вес), и наоборот. Такая величина называется спредом, так как она показывает расхождение от ожидаемой зависимости. Можно также думать и с той стороны, что при высокой ошибке высоко ее отклонение от ожидаемого значения, поэтому ошибка переоценена, и надо ее шортить, что, в целом, понятно. Получается, что знак значения ошибки показывает направление торговли портфолио с длинной позицией по второй акции и с короткой по первой в пропорции, заданной степенью хеджирования. В целом, можно и спред воспринимать как актив, с ним все в точности до наоборот. Я дальше буду говорить про него, так как про него говорится в книге.

Может такое случиться, что на новой выборке у акций будет новое соотношение или что нарушится коинтеграция. В таком случае удобно применить фильтр Калмана, который позволяет оценивать степень хеджирования динамически, подстраиваясь под новые значения при учете старых данных, не прибегая, например, к постоянной оценке степени хеджирования через некоторые промежутки времени. Модель фильтра Калмана выглядит следующим образом. Есть некоторый state (θ) размерности n , который через линейного оператора влияет на measurement (z) размерности m :

$$z_t = H * \theta_t + e_t, e_t \sim N(0, R),$$

где H – линейный оператор размерности (m, n) , а e_t – нормально распределенный шум с нулевым средним и ковариационной матрицей R размерности (m, m) . Замечу, что в нашем случае measurement — это цена первого актива размерности 1, а state – вектор размерности 2 из степени хеджирования и свободного члена регрессии. Также надо отметить, что в нашей модели линейный оператор H – это матрица размерности $(1, 2)$ из цены второго актива и единицы. Так как цена каждый день разная, то H будет зависеть от времени. Поэтому корректнее переписать это уравнение как

$$z_t = H_t * \theta_t + e_t, e_t \sim N(0, R).$$

Заметим, что в нашем случае при условии i.i.d. для e_t , ошибки будут стационарным рядом по определению. Далее State линейно зависит от своего предыдущего значения:

$$\theta_t = F * \theta_{t-1} + u_t, u_t \sim N(0, Q),$$

где F – некоторый линейный оператор размерности (n, n) , u_t – нормально распределенный шум с нулевым средним и ковариационной матрицей Q размерности (n, n) . В данном случае линейный оператор постоянен во времени.

Также мы считаем, что state $\theta_t \sim N(x_t, P_t)$, то есть распределен нормально с некоторым средним x_t и ковариационной матрицей P_t размерности (n, n) . Оценкой $\hat{\theta}_t$ для θ_t будет x_t . Эти параметры будут изменяться по мере получения новых наблюдений и подстраиваться под данные. Оценив эти параметры, мы сможем получить всю необходимую информацию из фильтра.

В каждый момент времени фильтр Калмана максимизирует параметры распределения θ_{t+1} при учете всех известных на данный момент времени параметров системы (D_t) и новых наблюдений measurement и вектора H_t в байесовском смысле:

$$x_{t+1}, P_{t+1} = \operatorname{argmax} P(\theta_{t+1} | z_t, D_t, H_t)$$

Новыми оптимальными значениями будут

$$x_{t+1} = x_t + K_t * y_t$$

$$P_{t+1} = (I - K_t * H_t) * P_t * (I - K_t * H_t)^T - K_t * R * K_t^T \text{ или } (I - K_t * H_t) * P_t,$$

но первое – более вычислительно устойчивое, где

$$K_t = P_t * H_t^T * S_t^{-1},$$

$$S_t = H_t * P_t * H_t^T + R,$$

$$y_t = z_t - H_t * x_t.$$

При этом y_t – будет ошибкой прогноза в момент времени t , а S_t – дисперсией ошибки прогноза, то есть,

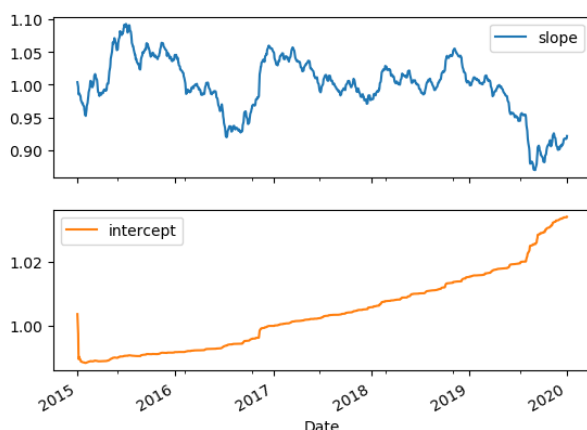
$$y_t = z_t - E(z_t | D_t),$$

$$S_t = \operatorname{Var}(z_t | D_t).$$

Именно эти значения и понадобятся для расчетов в стратегии.

Стоит отметить, что для инициализации фильтра требуется установить значения для неизменяемых во времени параметров, а также начальные значения для изменяющихся во времени параметров. Лучшие начальные значения обеспечат лучшую сходимость фильтра и получение более точных значений, что открывает возможности для подбора гиперпараметров.

Если мы посмотрим на предсказанные на тренировочной выборке параметры state, то мы увидим, что они не постоянны и меняются со временем. Поэтому есть смысл в динамической линейной регрессии, а фильтр Калмана отличное решение для этого, так как он эффективный, использует всю историю и считает оптимальные значения для параметров регрессии, поэтому он может помочь с проблемами динамического хеджирования и коинтеграции.



Я для каждой пары цен в определенный момент буду смотреть, насколько сильно отличаются ценовые показатели, используемые в модели от ожидаемого предсказания. Это отличие будет выражаться в ошибке предсказания, y_t . Эта ошибка, умноженная на -1 , даст спред. Если ошибка значимо велика, то я буду открывать торговую позицию, реализуя торговую стратегию. Определять значимость я буду сравнением с величиной стандартного отклонения прогноза ($\sqrt{S_t}$), умноженного на некоторую константу масштаба ($scale$). Этот показатель (результат сравнения) будет величиной, формирующей сигнал. Я считаю, что это вполне уместно, так как мои ошибки по предположению распределены нормально, поэтому вполне уместно использовать идею доверительных интервалов. Если значимость по модулю превышает этот показатель, то ошибка значимо велика, и я реализую свою стратегию.

Моя стратегия предполагает следующее поведение.

- $y_t < -scale * (\sqrt{S_t})$ – открываю длинную позицию на спред (короткую на ошибку), так как ошибка значимо меньше нуля (меньше отмасштабированного отклонения). Открываю длинную позицию на N акций TLT и короткую позицию $[\theta_t * N]$ акций IEI.
- $y_t \geq -scale * (\sqrt{S_t})$ – закрываю длинную позицию на спред. Ошибка вернулась в незначимую область, стала незначимой, поэтому предполагаем, что отношение между активами вернулась в зону нормы.
- $y_t > scale * (\sqrt{S_t})$ – открываю короткую позицию на спред. Открываю короткую позицию на N акций TLT и длинную позицию $[\theta_t * N]$ акций IEI.
- $y_t \leq scale * (\sqrt{S_t})$ – закрываю короткую позицию на спред.

В моей стратегии для простоты N будет равно 2000.

В качестве бенчмарка я буду использовать ETF казначейских облигаций с оставшимся сроком длительности от 1 года (GOVT). Я выбрал его, так как это ETF, который торгуется наиболее широким подходящим портфелем из облигаций (без очень краткосрочных облигаций). Также он тоже торгуется на NASDAQ. Он используется для подсчета коэффициента Шарпа.

Моя стратегия позволяет подбирать к ней разные условные гиперпараметры. Например, какой масштаб отклонения использовать или использовать ли цены акций ETF или доходности этих акций. Кроме того, я посмотрю, нужно ли предобучить фильтр на данных, предшествующих старту бэктеста. Предобучение будет заключаться в получении оценок на параметры фильтра на данных по ценам ETF за некоторое количество лет до старта бэктеста, которое может варьироваться. Оценка ковариационных матриц и линейного оператора для изменения State будет получено из предположений и оценок линейной регрессии на всей выборке данных. Для оценки параметров распределения State на начало бэктеста я в предобучении прогоню имеющиеся данные через фильтр и настрою тем самым параметры распределения. Дополнительной опцией будет в качестве начальных параметров распределения перед прогоном на предобучении установить начальные параметры распределения, полученные из соответствующей линейной регрессии.

Бэктест будет проводиться с помощью библиотеки QSTrader. Я устанавливал ее специфическую версию, поэтому если Вы захотите повторить мои результаты, то напишите мне, я напишу, что надо установить. Я в ней изменил или расширил своими структурами некоторые классы, чтобы бэктест выдавал нужные мне результаты. Моя модель автоматически скачивает данные для нужных тикеров и в нужный промежуток с YahooFinance с помощью библиотеки pandas-datareader, поэтому я не скачивал заранее никаких данных, они скачиваются сами и скачаются у Вас при запуске моего кода.

Бэктест проводится по следующей схеме. На тренировочной выборке я тестирую стратегию с каждым допустимым набором гиперпараметров. Далее я оставляю те модификации, которые попали в топ-7 лучших по величине коэффициента Шарпа или отношения доходности к максимальной просадке. Стратегию на этих модификациях я прогоняю на тестовой выборке и по показателям выбираю лучшую модификацию. Далее стратегию с лучшей модификацией я тестирую на финальной выборке. Я предупреждаю, что я тренировочную выборку ограничил снизу началом 2015 года, чтобы у меня данные для предобучения не попадали в период Мирового финансового кризиса.

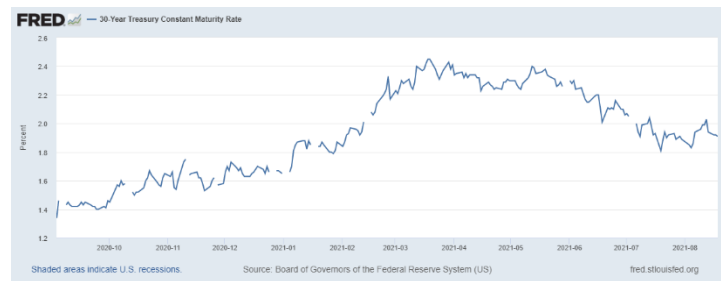
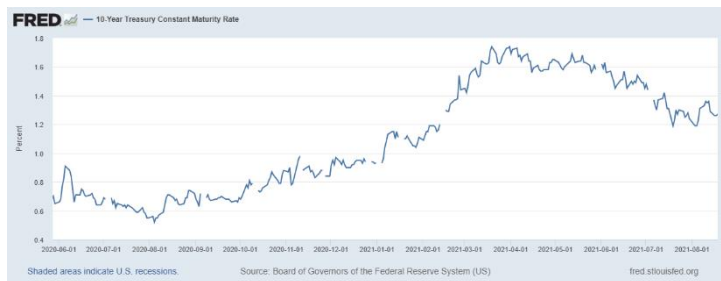
На тренировочной выборке я протестировал 88 модификаций стратегии и отобрал по критериям выше 8 стратегий. Все стратегии, которые попали на тестовый бэктест, использовали предобучение и использовали цены активов, а не доходности. 7 из 8 стратегий обучались на 4 или 5 годах до старта бэктеста и использовали масштаб 0.5 или 1. Лишь 1 стратегия выбивается из этого ряда, обучаясь на 3 годах и используя масштаб 2.

На тестовой выборке лучший результат показала стратегия с 4 годами обучения и масштабом 1. Её я и буду использовать в финальном прогоне.

Результаты финального бэктеста неутешительны. Стратегия еле вышла в ноль и в течение всего периода бэктеста не показывала значимый прирост стоимости портфолио. У стратегии случилась огромная просадка весной 2021-го года. Тем не менее, если мы посмотрим на бенчмарк, то мы увидим, что он подвергся такой же просадке. Значит, что проблема глобальная. **Коэффициент Шарпа равен 0.32, а отношение доходности к максимальной просадке равно -0.05.**

Проблема связана с тем, что весной 2021-го года вследствие ряда обстоятельств произошло резкое увеличение доходности (yield) как по среднесрочным, так и долгосрочным казначейским облигациям.





Из-за повышения доходностей упала цена облигаций, как и стоимость активов ETF и бенчмарка, что повлекло падение цен акций последних. Далее и стратегия, и бенчмарк корректируются, так как последствия такого шока были сглажены. Стоит заметить, что наша стратегия была более волатильной, чем бенчмарк. Это связано с тем, что наша стратегия зашортила спред, то есть, купила акции TLT и зашортила IEI. В нашей стратегии вес TLT был достаточно высоким, примерно половина портфеля по абсолютной стоимости, когда в бенчмарке облигации 20+ лет занимают значительно меньший вес:

as of Aug 19, 2021

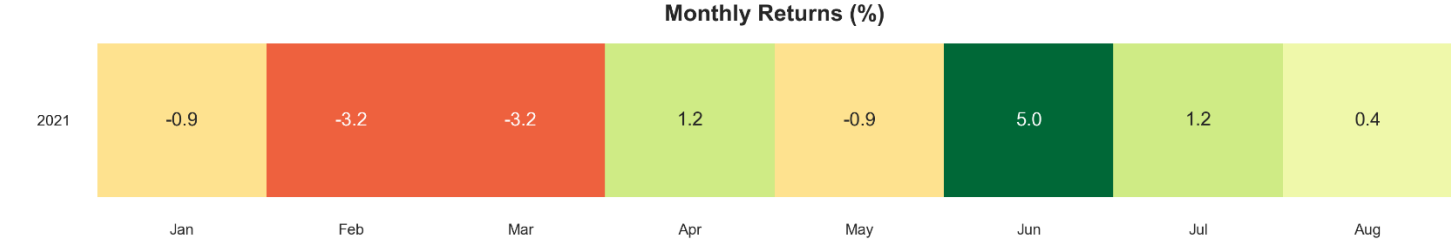
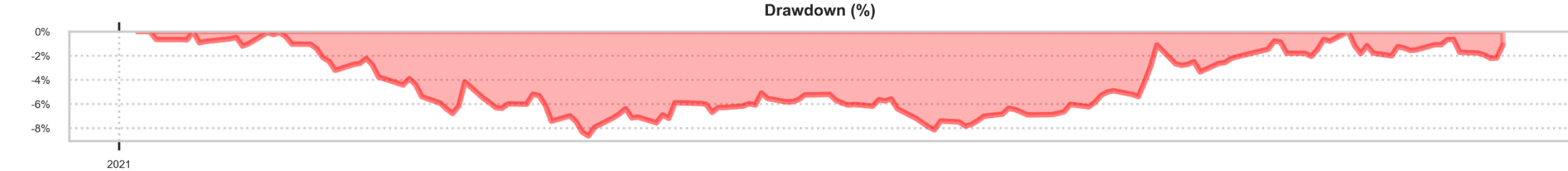
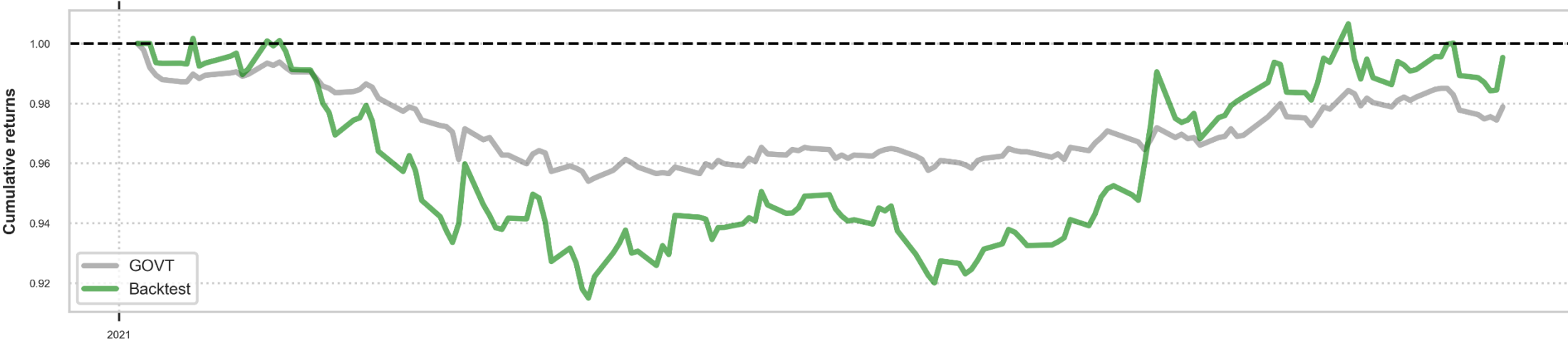
% of Market Value

| Type | ▲ Fund ► | Fund |
|----------------------|----------|------|
| 3 - 5 Years | 23.55 | |
| 20+ Years | 18.53 | |
| 5 - 7 Years | 16.94 | |
| 1 - 2 Years | 15.69 | |
| 2 - 3 Years | 9.66 | |
| 7 - 10 Years | 8.75 | |
| 0 - 1 Years | 3.56 | |
| 15 - 20 Years | 2.60 | |
| 10 - 15 Years | 0.07 | |
| Cash and Derivatives | 0.65 | |

Как мы видели, акции TLT более волатильны и сильнее реагируют на неожиданные события. Поэтому акции TLT сильнее прореагировали на такой шок, и падение их стоимости превысило хеджирование акциями IEI. По той же причине коррекция после шока была у нашей стратегии сильнее, чем у бенчмарка.

Замечу, что стратегия не совершила ни одной сделки после открытия короткой позиции на спред, так как из-за шока была высокая ошибка предсказаний, поэтому сигнал сохранялся. Возможно, проблема была в том, что фильтр работал с ценами, а цены менее склонны к коинтеграции, и в такой шок коинтеграция была нарушена, поэтому остатки регрессии там и не смогли сойтись обратно, и сделка не закрылась.

Результаты стратегии с лучшей модификацией на
финальном бэкteste



Curve vs. Benchmark

Trade

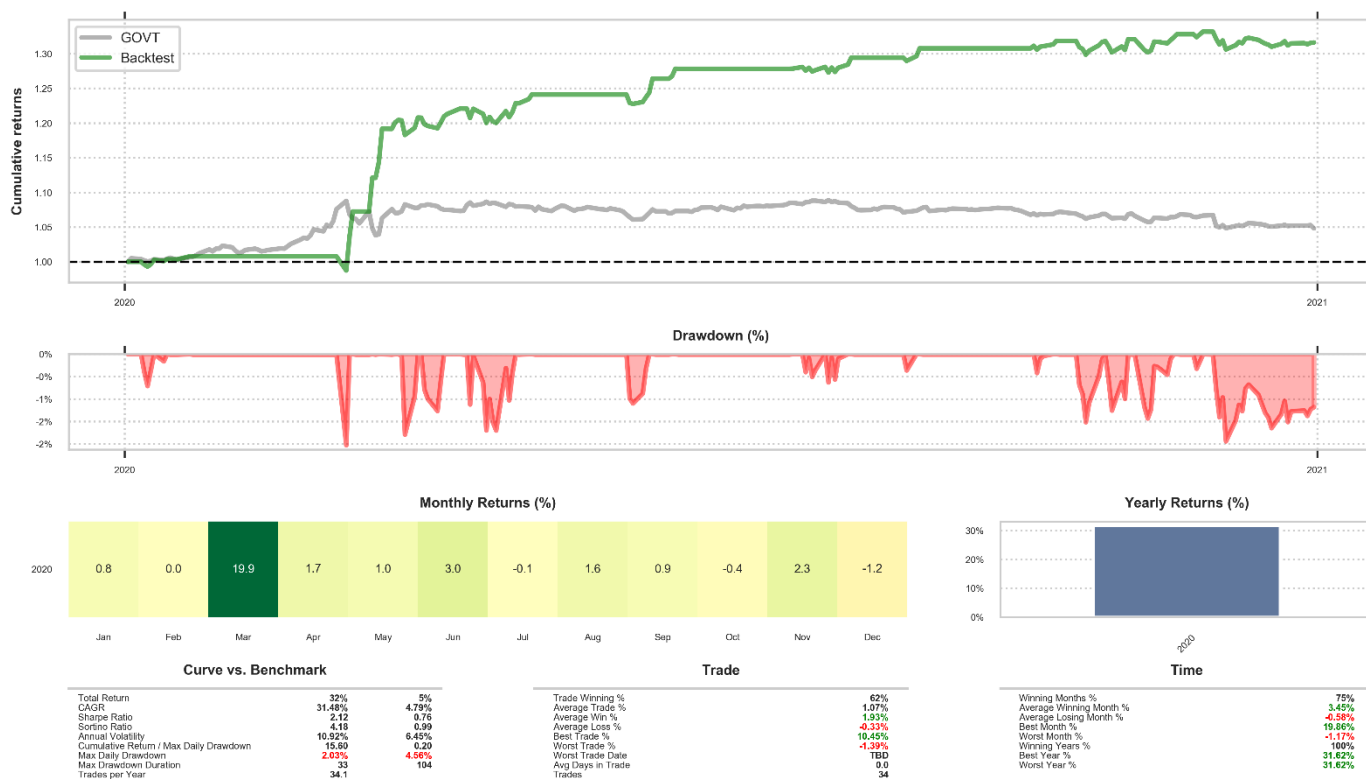
Time

| | | |
|--|--------|--------|
| Total Return | -0% | -2% |
| CAGR | -0.76% | -3.42% |
| Sharpe Ratio | 0.32 | -0.84 |
| Sortino Ratio | -0.05 | -1.38 |
| Annual Volatility | 9.65% | 4.06% |
| Cumulative Return / Max Daily Drawdown | -0.05 | 0.05 |
| Max Daily Drawdown | 8.65% | 4.39% |
| Max Drawdown Duration | 127 | 153 |
| Trades per Year | 0.0 | |

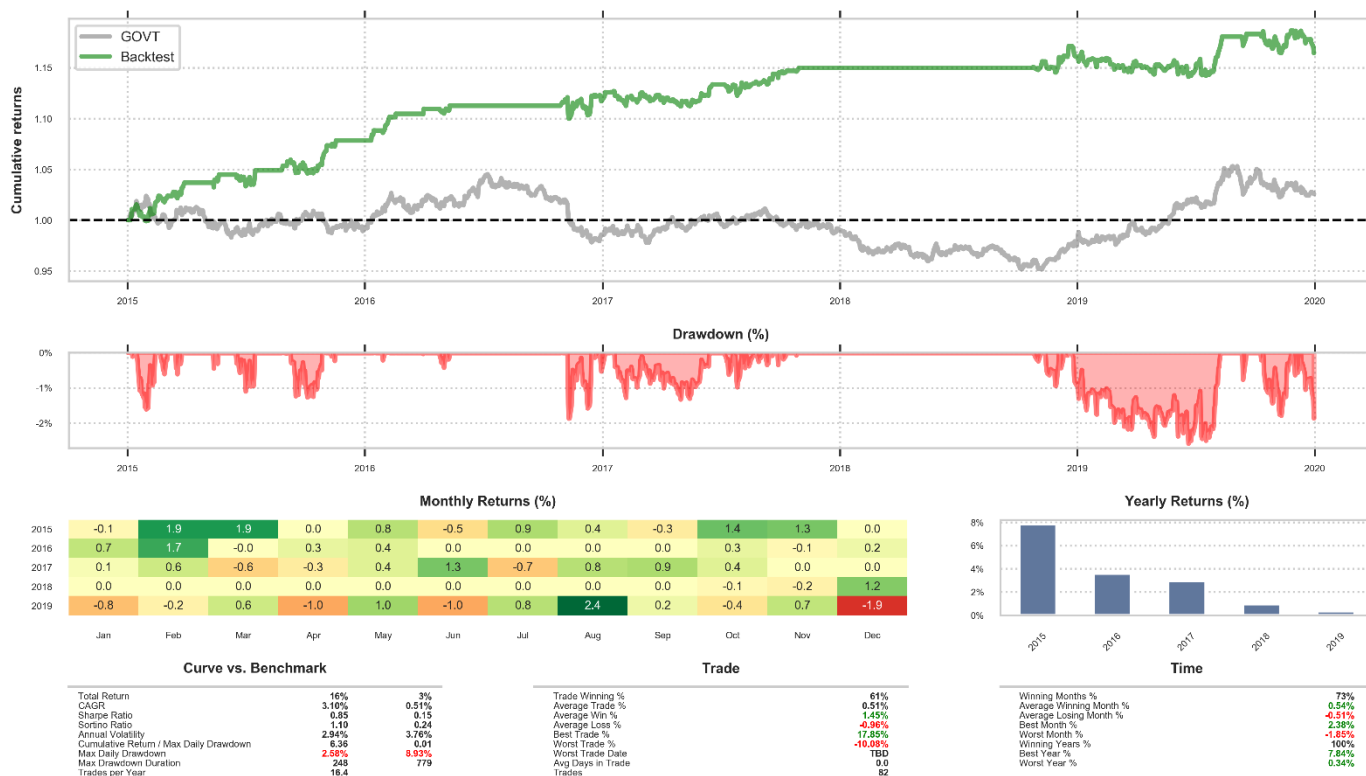
| | |
|-------------------|-----|
| Trade Winning % | N/A |
| Average Trade % | N/A |
| Average Win % | N/A |
| Average Loss % | N/A |
| Best Trade % | N/A |
| Worst Trade % | N/A |
| Worst Trade Date | TBD |
| Avg Days in Trade | 0.0 |
| Trades | 0 |

| | |
|-------------------------|--------|
| Winning Months % | 50% |
| Average Winning Month % | 1.97% |
| Average Losing Month % | -2.03% |
| Best Month % | 5.01% |
| Worst Month % | -3.18% |
| Winning Years % | 0% |
| Best Year % | -0.47% |
| Worst Year % | -0.47% |

Если мы посмотрим на результаты стратегии на тренировочной и тестовой выборках, то они выглядят намного лучше, чем на финальной. На тестовой выборке стратегия показывает устойчивый и вполне стабильный рост. Коэффициент Шарпа на тестовой выборке составляет 2.12, а отношение доходности к максимальной просадке равно 15.6. Хотя стоит признать, что основной доход случился из-за резкого падения процентных ставок в это время. Тем не менее, даже после такого, пускай и положительного шока, стратегия стабильно росла и добавила еще 10% к стоимости портфеля.



На тренировочной выборке стратегия тоже показывала, в целом, стабильную положительную динамику. Коэффициент Шарпа на этой выборке равен 0.95, а отношение доходности и максимальной просадки равно 6.36.



Вследствие этого я могу заключить, что нельзя с большой уверенностью заявлять, что по результатам финального бэктеста стоит признать стратегию невалидной, так как финальный период сопровождался непредвиденным отрицательным для стратегии шоком, из-за которого модель не смогла реализовать себя.

Тем не менее, с учетом того, что во время позитивного для стратегии шока в начале 2020-го стратегия много заработала, а в периоды без шоков скорее показывала стабильный рост, можно заключить, что эта стратегия скорее валидна. Позитивные и негативные шоки компенсируют друг друга, а в остальном модель скорее будет зарабатывать свое.

[Сцена после титров]

Чтобы посмотреть, теряется ли в финальном периоде коинтеграция между активами, я посмотрел на результаты той же модификации, только с доходностями вместо цен акций как данных для модели. В такой модификации стратегия совершила 46 сделок, что означает, что она сумела совладать с шоками и сохранить коинтеграцию, хоть и не удачно, судя по результатам, но это другой вопрос.

