Market Inefficiency:

Pairs Trading with the Kalman Filter

Heather E. Dempsey, Sacred Heart University

December 8th, 2017

www.hedempsey.com

Abstract

Keywords: Pairs trading, Kalman Filter, Statistical arbitrage, Arbitrage Pricing Theory (APT)

JEL codes: C6 C15 C32 C88 G17

I. Introduction

- The motivation behind this study is to develop a profitable trading strategy as an individual investor and gain experience with the Kalman Filter (KF).
- This paper contributes to past research with application of the strategy on North American equities traded on the Nasdaq and NYSE, whereas much of the literature focused on the S&P 500. With emphasis on the individual investors application of the pairs strategy whereas the past literature has focused on the market makers perspective or that of individual investors.
- The research question –*is* pairs trading still profitable for the individual—with so many players?
- Interesting question because to profit with statistical arbitrage methods, seemingly contradicts efficient market hypothesis (EMH) and challenges our understanding of the widely accepted EMH.

  How does the paper address the question? With use of KF and state-space equations influenced heavily by the Hidden Markov Model to estimate the unobserved latent variable—the hedge-ratio. Theoretical model derived from statistical arbitrage equation relating the price of one stock to the other through a discrete value beta. What techniques and methods are used?

II. Literature review

There is an increasing interest in the use of the KF, sate-space models, and similar class of algorithms, because of the highly dynamic nature of the hedge ratio. The KF, a recursive algorithm, was first designed by Rudolf Kalman (1960's) to track a moving target. For example, the KF is used in NASA's trajectory estimations, the US Navy's ballistic missile submarines, radar technology, and more recently—global positioning systems (GPS), satellite tracking, robotics, econometrics, and many other uses. The KF is equipped to handle multi-dimensions in both the state and observation matrices. The KF uses a set of equations to iteratively measure successive observations with increasing accuracy using only the previous estimate and current estimate. One of the benefits is less memory required for computation, which ultimately increases speed, crucial to today's financial environment. The KF quickly "filters out the noise" (Martinelli & Rhoads, 2010) and narrows in on the true value of the hedge-ratio by reducing the errors in the estimate, and errors in the measurement Kalman gain (KG) a weight, which is responsible for putting greater or lesser importance when ultimately feeding into the calculation of the *new* current estimation (Van Biezen, 2015). In contrast, linear regression requires a set of data which has already been gathered. The KF, however needs only a few observations before estimating an optimized hedge ratio. KF is ideal for real-time applications. Financial market data, in general, is difficult to identify a trend and predict future movements with the KF. Historical data (120 observations minimum) is needed to estimate a trend. Market data is noisy, prices fluctuate arbitrarily from day-to-day and the variance often changes.

Engle and Granger (1987) in an influential work studying pairs trading used cointegration and vector error correction models to examine pairs of related equities. Often, the spread between equities were found to depart from equilibrium temporarily. A trading strategy which would short the high-priced stock and go long on the low-priced stock until market forces pushed the economy back into equilibrium and the pair would mean-revert. Johansen and Juselius (1990) extended their work to include baskets of cointegrated equities in Finnish and Dutch market data. The estimation was conducted using vector auto regression and used a maximum likelihood estimator with a Wald test for hypotheses on alpha and beta.

Vidyamurthy (2004) describes the best way to pick the pairs of cointegrated equities. In his book, a variety of methods are employed, including APT and cointegration models. He found proof the KF is optimal when the state-space and observation equations are linear and the noise follows a Gaussian distribution (Vidyamurthy, 2004). The algorithm is linearized in most cases but there does exist several extensions of the KF. For example, the Extended Kalman Filter and Unscented KF which deal with non-linearities (EKF), which handle non-linearities and the latter, for non-Gaussian distributions.
a thorough discussion of pairs trading. The underlying mathematical model used in KF, is rooted in L.E. Baum's (1972) Hidden Markov Model (HMM) and Baum-Welch algorithm which lays the foundation for state-transition and measurement equations when estimating unobservable or latent-state variables.

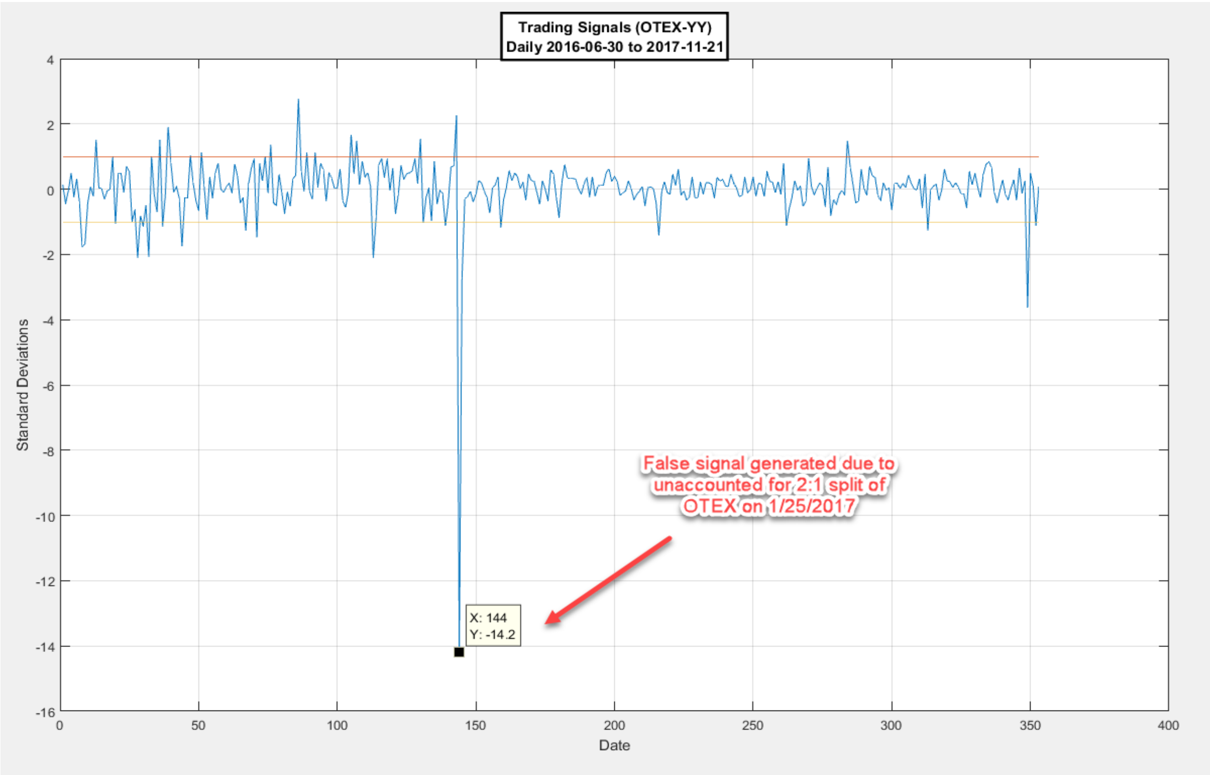
Gatev et al. (2006) used 40 years of daily US market data picked his pairs by minimizing the variance of two daily-price series once each was normalized. The strategy entails taking a long and short position when the prices between a pair deviate by more than plus or minus 2 standard deviations in anticipation of mean-reversion (De Moura, Pizzinga, & Zubelli, 2016). The current study follows these limits.

Elliot et al. (2005) focused solely on the state-space models including the KF, and mean-reverting Markov Chain Model to "monitor spreads, profiting from investment decisions based on this spread" Elliot et al found it useful for market—in this case dollar-neutral portfolios, hedge funds, and trading. finite state space Markov chains. Week 3 begins the transition to continuous time with finite time versions of the Gaussian processes. The idea is when the spread widens to predetermined degree, to then short the high-valued stock, and go long on the low-priced equity. Profits are shown to have been made from such strategy. However, speed is crucial to identify the deviation from the mean so as to enter into a position before equilibrium was restored.

III. Methodology-modeling and testing procedure

1. Data

Figure 1. A ***false*** signal resulting from an unaccounted for 2:1 split on 1/25/2017



The Kalman Filter is a dynamic linear state-space model and has been shown to be an optimal estimator when one or more variables are unobservable.

The algorithm recursively approaches the true value of the linear variable (see Halls-Moore, 2015) beta; the hedge-ratio in this study, and does so in the presence of convoluting noise, inherent in financial market where the measurements coming in contain inaccuracies from which we wish to estimate the closest approximation of the unobserved 'latent-state' variable. The current estimate is defined as a function of itself with only the previous estimate needed plus some uncertainty. The current estimate($\hat{\beta}_t$), which represents the hedge ratio, is calculated through a set of equations in an iterative process. Accomplished by adding the delta between the new observation (Y_t) and previous estimate ($\hat{\beta}_{t-1}$), which is then weighted by the Kalman Gain (KG), then added to the $\hat{\beta}_{t-1}$to calculate the new $\hat{\beta}_t$ for each time-step.

The KG (Kalman Gain) calculation is simply the ratio of the error in the β_{t-1}, to the sum of error in the measurement *and* error in the previous estimate, which will decide whether the observation or the β_{t-1}estimate should carry more importance in deriving the next estimate.

IV. Main Results

Table 1.

In Sample Period 2017/06/30 - 2017/11/21 Pair VNET-NCR	Estimated using Kalman Filter	In Sample Period 2017/06/30 - 2017/11/21 Pair EEV-ERX	Estimated using Kalman Filter
APR	**31%**	**APR**	**47%**
Sharpe Ratio	**1.51**	**Sharpe Ratio**	**1.78**
Maximum Drawdown	**-9.8%**	**Maximum Drawdown**	**-10.8%**
Maximum Drawdown Days	**80**	**Maximum Drawdown Days**	**87**
Cumulative Return	**46%**	**Cumulative Return**	**74%**

Note: Kalman filter used for beta calculation on pair of screened securities, a) 21 Vianet 'VNET' and (b) NCR. Out-of-sample period estimated from 2017/11/22 – 2017/12/06 (24hrs using a quad-core processor, to analyze +20,000 pairs just like VNET-NCR. Strategy is still running, complete results made available by 2017/12/13. Results will be posted to data section at http://www.hedempsey.com. Table 1 inspired by (Drakos, 2016 & Dunis et al., 2010).

Figure 1. Cointegrated pair of securities. Morgan Stanley (MS) and Freeport-McMoRan, Inc. (FCX)

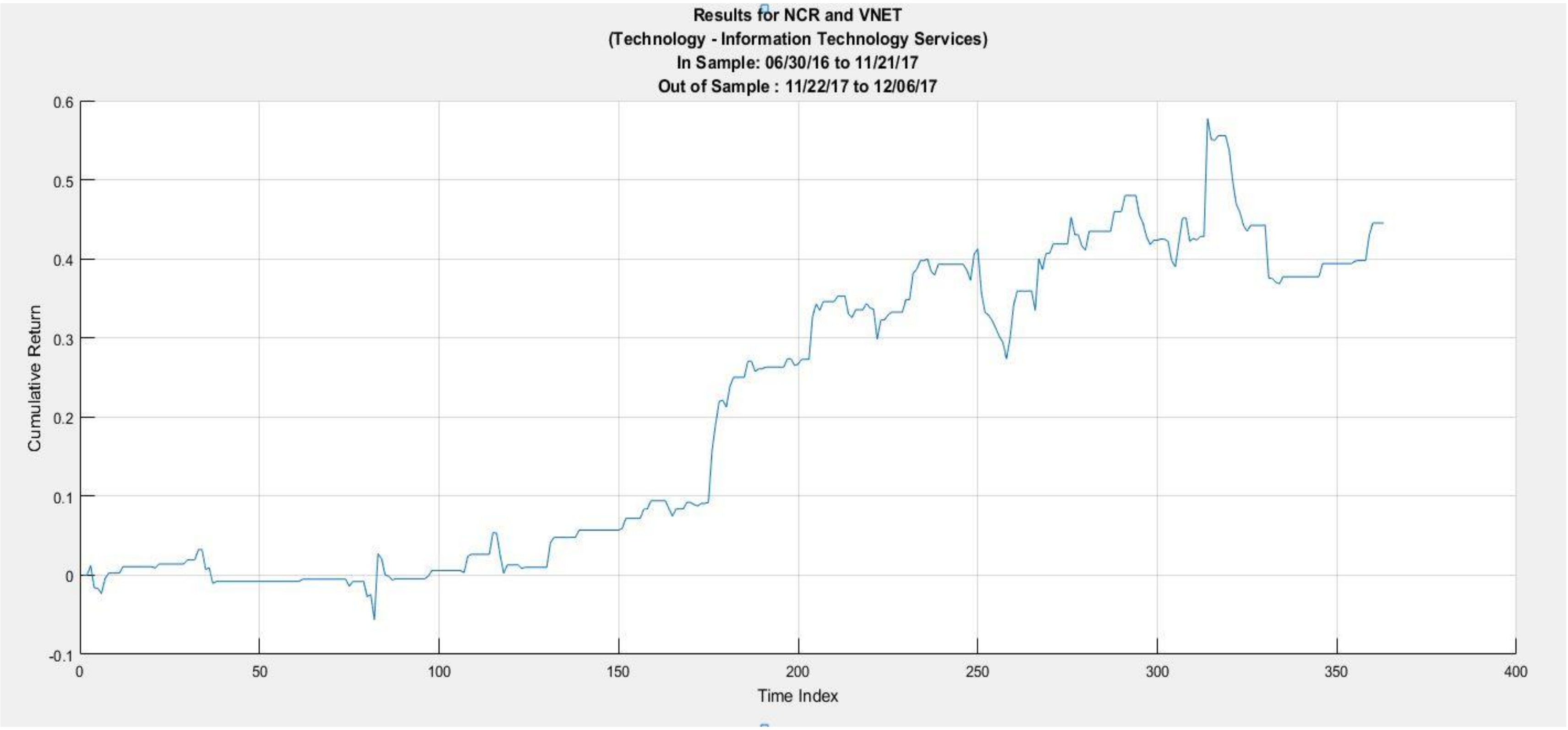


Figure 2. Cumulative Return pair NCR-VNET

V. **Conclusion**

Future recommendations, unscented Kalman Filter, neural networks and genetic algorithms to trade. Dynamic asset allocation with genetic algorithms or neural networks.

Appendix

Notation follows Chan (2013) *Algorithmic Trading: Winning Strategies and their Rationale* with code snippets in MATLAB.

$y(t) = x(t)\ \beta(t) + \epsilon(t)$ ("Measurement equation")

$\beta(t) = \beta(t-1) + \omega(t-1)$ ("State transition")

Denoting the expected $E[\ \beta_t]$ —given the previous observation with $\hat{\beta}(t|t-1)$, the expected value E [β] given the new data input $\hat{\beta}(t|t)$, and the E [y(t)] given the measurement at time denoted by $\hat{y}(t|t-1)$. Given the quantities $\hat{\beta}(t-1|t-1)$ and $R(t-1|t-1)$ at time $t-1$, one-step predictions are possible.

$\hat{\beta}(t|t-1) = \hat{\beta}(t-1|t-1)$ ("State prediction")

$R(t|t-1) = R(t-1|t-1) + V_w$ ("State covariance prediction")

$\hat{y}(t) = x(t)\hat{\beta}(t|t-1)$ ("Measurement prediction")

$Q(t) = x(t)R(t|t-1)x(t) + V\epsilon$ ("Measurement variance prediction")

In Chan's text, he describes the variance-covariance matrices as *"where $R(t|t-1)$ is $cov(\beta(t) - \hat{\beta}(t|t-1))$, measuring the covariance of the error of the hidden variable estimates. (It is a covariance instead of a variance because β has two independent components.) Similarly, $R(t|t)$ is $cov(\beta(t) - \hat{\beta}(t|t)$. Remembering that the hidden variable consists of both the mean of the spread as well as the hedge ratio, R is a 2×2 matrix. $e(t) = y(t) - x(t)\hat{\beta}(t|t-1)$ is the forecast error for $y(t)$ given observation at $t-1$, and $Q(t)$ is $var(e(t))$, measuring the variance of the forecast error".*

After observing the measurement at time t, the famous Kalman filter state estimate update and covariance update equations are

$\hat{\beta}(t|t) = \hat{\beta}(t|t-1) + K(t) * e(t)$ ("State update")

$R(t|t) = R(t|t-1) - K(t) * x(t) * R(t|T-1)$ ("State covariance update")

where $K(t)$ is called the Kalman gain and is given by

$K(t) = R(t|t-1) * x(t)/Q(t)$

References

Benhamou, E. (2016). *Trend Without Hiccups - A Kalman Filter Approach.* Thomson Reuters.

Chan, E. P. (2013). *Algorithmic Trading.* Hoboken: John Wiley & Sons.

De Moura, C. E., Pizzinga, A., & Zubelli, J. (2016). A Pairs Trading Strategy Based on Linear State Space Models and the Kalman Filter. *Quantitave Finance*, *16*(10), 1559-1573. doi:10.1080/14697688.2016.1164886

Doucet, A., & Johansen, A. M. (2008). A Tutorial on Particle Filtering and Smoothing: Fifteen years later. Tokyo, Japan: Institute of Statistical Mathematics.

Drakos, S. (2016). Statistical Arbitrage in S&P500. *Journal of Mathematical Finance*, 6, 166-177.

Dunis, C. L., Giorgioni, G., Laws, J., & Rudy, J. (2010). *Statistical Arbitrage and High-Frequency Data with an Application to Eurostoxx 50 Equities.* Liverpool: Center for International Banking, Economics and Finance.

Elliot, R. J., Van Der Hoek, J., & Malcolm, W. P. (2005). Pairs Trading. *Quantitative Finance*, 271-276. doi:10.1080/146976680500149370

Engle, R. F., & Granger, C. W. (1987, March). Cointegration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, *55*(2), 251-276. Retrieved from http://www.jstor.org/stable/1913236

Gatev, E., Goetzman, W. N., & Rouwenhorst, K. G. (2006). Pairs Trading: Performance of Relative-value Arbitrage Rule. *The Review of Financial Studies*, *19*(2), 797-827.

Halls-Moore, M. (2015, December 28). *State Space Models and the Kalman Filter*. Retrieved from QuantStart: https://www.quantstart.com/articles/State-Space-Models-and-the-Kalman-Filter

Johansen, S., & Juselius, K. (1990). Maximum Likelihood Estimation and Inference on Cointegration - With Applications to the Demand For Money. *Oxford Bulletin of Economics and Statistics*, *52*(2), 169-210.

Kim, P. (2011). *Kalman Filter for Beginners with MATLAB Examples.* Republic of Korea: A-JIN Publishing Company.

Martinelli, R., & Rhoads, N. (2010, January). Predicting Market Data Using the Kalman Filter. *Technical Analysis of Stock and Commodities*.

Miao, G. J. (2014). High Frequency and Dynamic Pairs Trading Based on Statistical Arbitrage Using a Two-Stage Correlation and Cointegration Approach. *International Journal of Economics and Finance, Vol. 6*(No. 3), 96-110. Retrieved from http://dx.doi.org/10.5539/ijef.v6n3p96

PennState Elberly College of Science. (2017, December 8). *State Space Models.* Retrieved from Applied Time Series Analysis | STAT 510: https://onlinecourses.science.psu.edu/stat510/node/27

Tsai, Y.-C., & Lyuu, Y.-D. (2017). A New Robust Kalman Filter for Filtering the Microstructure Noise. *Communications in Statistics - Theory and Methods*, *46*(10), 4691-4976. doi:10.1080/03610926.2015.1096390

Van Biezen, M. (2015, October 27). Special Topics - The Kalman Filter (9 of 55) The Multi-Dimension Model 3: The State Matrix.

Vidyamurthy, G. (2004). *Pairs Trading: Quantitative Methods and Analysis.* Hoboken: John Wiley & Sons, Inc.