Pairs Trading: A Kalman Filtering Approach

T. Dunlop, J. Llorens,
M. Keys, R. Knudsen

March 4, 2018

Introduction

- Natural demand for profitable strategies in financial markets
- Statistical arbitrage strategies capitalise on financial markets being out of equilibrium
- Pair Trading trades a pair of securities
  - that are "close" (e.g. stocks of Pepsi and Coca Cola)
  - that follow a certain joined process
  - and buys security A and sells security B (or vice versa) whenever the model indicates that the process is out of equilibrium

Mean-reverting hidden markov model of the spread

Hidden state process:

$$x_k = A + Bx_{k-1} + C\epsilon_k$$

where $\epsilon_k \sim \mathcal{N}(0,1)$, iid and $X(t)$ follows an Ornstein-Uhlenbeck process:

$$dX(t) = (a - bX(t))dt + \sigma dW(t)$$

where $\{W(t)|t \geq 0\}$ is a standard Brownian motion.

Observation process:

$$y_k = x_k + D\omega_k$$

where $\omega_k \sim \mathcal{N}(0,1)$

Estimating the hidden state

- How to optimally predict hidden states in a system characterised by measurement noise and a stochastic process driving the state?
- Kalman Filtering: Expresses state estimates as a weighted sum of measurements and apriori state estimates!

$$
\begin{aligned}
\hat{x}_{k+1|k} &= A + B\hat{x}_{k|k} \\
\Sigma_{k+1|k} &= B^2\Sigma_{k|k} + C^2 \\
\mathcal{K}_{k+1} &= \Sigma_{k+1|k}/(\Sigma_{k+1|k} + D^2) \\
\hat{x}_{k+1} &= \hat{x}_{k+1|k} + \mathcal{K}_{k+1}(y_{k+1} - \hat{x}_{k+1|k}) \\
\Sigma_{k+1} &= \Sigma_{k+1|k} - \mathcal{K}_{k+1}\Sigma_{k+1|k}
\end{aligned}
$$

Parameter Estimation

- As usual, we are interested in estimators of the form:
  - $\hat{\theta} = \arg\max \mathcal{L}(y|\theta)$
- Closed form solutions to this optimisation do not exist given the latent structure embedded in our model.
  - Why?
- What other options do we have?
  - There is special class of algorithms designed precisely for this problem: Expectation Maximisation (EM)
  - These are iterative algorithms that make use of the joint likelihood of observed and latent random variables, and are guaranteed to monotonically increase the observed likelihood at each iteration!

Expectation Maximisation

- Define $Q(\theta, \theta') \triangleq \mathbb{E}[log\mathcal{L}(x, y|\theta)|y_1, ..y_N, \theta']$.
- Derive parameter updates at each iteration as $\theta_{j+1} = \arg\max Q(\theta, \theta_j)$
- It converges to a (not necessarily global!) optimum.

EM Smoothing Algorithm

Input $y = (y_0, y_1, ..., y_n)$
Initialise $\theta_0 = (A_0, B_0, C_0, D_0)$, $\hat{x}_0 = y_0$ and $\Sigma_0 = D^2$
for $j = 0, 1, ...$ **do**
  Apply Kalman Filter with Σ_0, $\hat{x}_0$ and θ_j
  E-Step: Apply Kalman Smoothing with $\hat{x}_{N|N}$ and $\Sigma_{N|N}$ to compute $Q(\theta, \theta_j)$
  M-Step: Compute parameter updates as $\theta_{j+1} = \arg\max Q(\theta, \theta_j)$
  Update initial filter values for the following iteration as $\hat{x}_0 = \hat{x}_{0|N}$ and $\Sigma_0 = \Sigma_{0|N}$
  Continue until $\mathcal{L}(y|\theta_{j+1}) - \mathcal{L}(y|\theta_j) \leq \textit{tolerance}$ or $j =$ max iterations
Return θ

Difference between observed and hidden spread

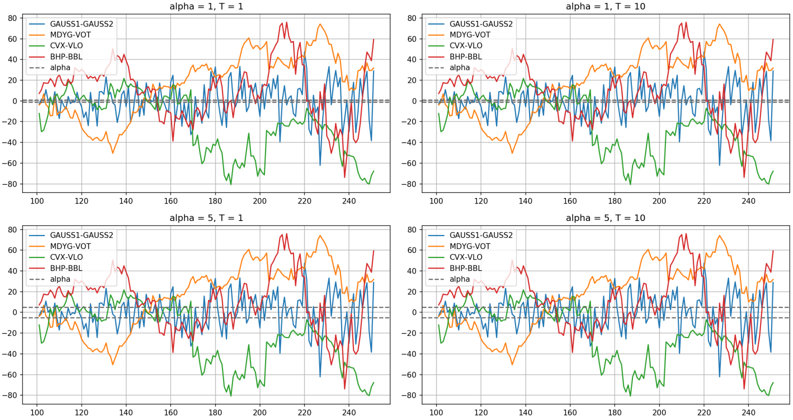

Figure: Difference between observed and hidden spread for four securities

Strategies chosen

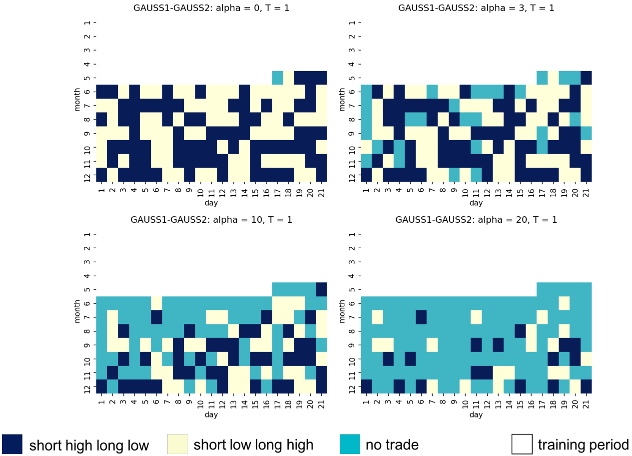

Figure: Strategies chosen for the four securities

Cumulative profits of strategies

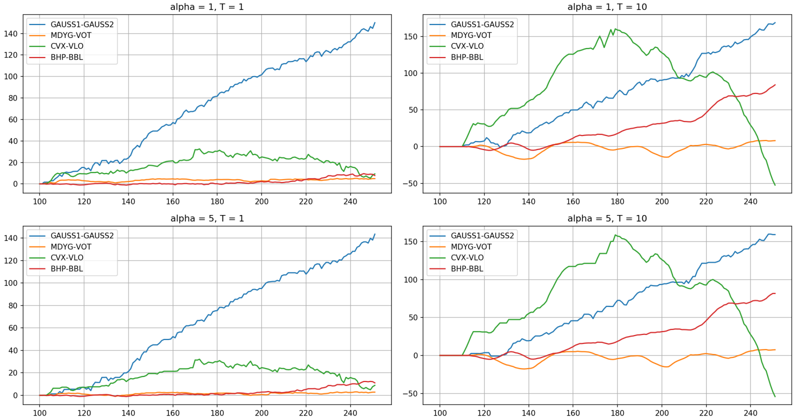

Figure: Cumulative profits for chosen securities using different strategies

Conclusion

- The strategy produces return before accounting for transaction costs.
- The model, however, assumes Gaussian noise, a clear limitation in financial markets.
- Extensions could incorporate more general distributions.
- Also, instead of using heuristics, one could find optimal trading rules, which is however hard.

Main references

- Elliott, R. J., Van Der Hoek*, J., and Malcolm, W. P. (2005). Pairs trading. *Quantitative Finance*, 5(3):271 276.
- Shumway, R. H. and Stoffer, D. S. (1982). An approach to time series smoothing and forecasting using the em algorithm. *Journal of time series analysis*, 3(4):253264.
- Vidyamurthy, G. (2004). *Pairs Trading: quantitative methods and analysis, volume 217*. John Wiley & Sons.