Kalman Filtering and Pair Trading Notes (Pre-class)

Pair Trading - select two stocks which move similarly, sell high priced stock and buy low priced stock when theres a price divergence

*Cointegration - statistical property of time series
- Specifies co-movement of price (long-term relationship)

Correlation - specifies comovenment of return (short-term)

- At each step n

① Form a Prior Belief with last update

② Update your Belief w/ New measurment

Basics

1. Find 2 likely correlated stocks
2. Estimate Spreads
3. Check Stationarity
4. Create trading Signal
5. Apply kalman

Observation Explanation: Let Y_t, Y_{t-1} (data) $t, t-1$ (time)

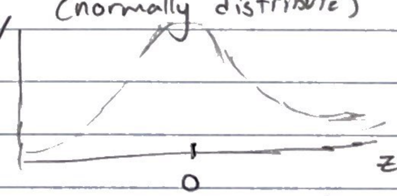
* We assume Y_t depends on an unobservable quantity θ_t (state of nature)

Main goal → to infer θ_t from Y_t

$$Y_t = \underset{\text{measurment matrix}}{F_t}\, \underset{\text{true state}}{\theta_t} + \underset{\text{noise}}{V_t}$$

Hidden State: $$\theta_t = \underset{\text{state transition matrix}}{G_t}\, \underset{\text{prev. state}}{\theta_{t-1}} + \underset{\text{noise}}{W_t}$$

V_t is observation error (normally distribute)



- Hand Derive K_t and include as pdf to assignment

CA - canadian ETF

$$K_t = \frac{S_{t,CA}}{S^2_{t,CA} + y^{-1}} \qquad y = \frac{P_{t,t-1}}{\sigma^2}$$

y < small we ignore observation b/c its too noisey
opp. relevant

Since the Kalman Filter is basically a Bayesian Estimator

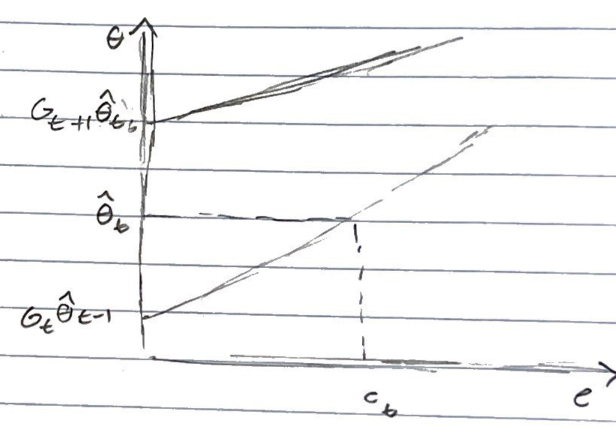
$$\underset{①}{P(\Theta_t | Y_t)} \propto \underset{②}{P(Y_t | \Theta_t, Y_{t-1})}\,\underset{③}{P(\Theta_t | Y_{t-1})}$$

① $P(\Theta_t | Y_t)$ → Posterior Distribution (our updated belief about Θ_t)

② $P(Y_t | \Theta_t, Y_{t-1})$ → Likelihood (How well does the measurement fit the estimate)

③ $P(\Theta_t | Y_{t-1})$ → Prior Distribution (Prediction from last step)

- Updates recursively so t_{-1} important to t
- Adjust e_b to adjust filters estimate



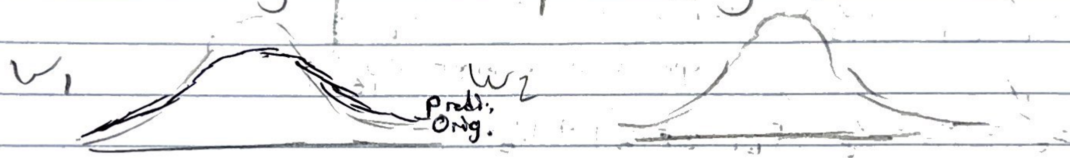
$y = \mu_y + \sigma_y w_1 \rightarrow N(0,1)$ w_1 and w_2 are correlated

$\theta = \mu_\theta + \sigma_\theta w_2 \rightarrow N(0,1)$ $w_2 = \rho w_1 + \sqrt{1-\rho^2}\, w_2$

We solve for $w_1 = \frac{y - \mu_y}{\sigma_y}$ and

plug into $\theta = \mu_\theta + \sigma_\theta \left(\rho \left(\frac{y-\mu_y}{\sigma_y} \right) + \sqrt{1-\rho^2}\, w_2 \right)$

Using this we keep decreasing our variance



But since we know w_1 we can actually reduce variance for w_2

$\theta = \mu_\theta + \frac{[\sigma_\theta \rho \sigma_y]}{\sigma_y^2}$ ← covariance of x and y ← varianc

In-class Derivation

Kalman Filter

$Y_t = F_t \theta_t + V_t \to N(0, V_t)$

★ Using scalars G_t is not a matrix for simplicities sake

$\theta_t = G_t \theta_{t-1} + v_t \to N(0, W_t)$

$G_t \to Z_{t-1}$

$(\theta_t \mid Y_t) \sim N(G_t \hat{\theta}_{t-1}, \overbrace{W_2 + G_t^2 Z_{t-1}}^{R_t})$

$(\cancel{\theta_t} Y_t \mid Y_{t-1}) \sim N(\cancel{F_t \hat{\theta}_t}\ 0, F_t R_t + V_t)$

$cov(e_t, \theta_t \mid Y_t) = E[e_t(\theta_t - \hat{\theta}_t) \mid Y_{t-1}]$

Expectation of error of Y_{t-1}

$e_t = Y_t - \hat{Y}_t = F_t(\theta_t - \hat{\theta}_t) + V_t$

$= F_t E[(\theta_t - \hat{\theta}_t)^2 \mid Y_t]$

$= F_t R_t$

$\mu_\theta \mid Y_t = G_t \theta_{t-1} + \boxed{\frac{F_t R_t}{(V_t + F_t^2 R_t)}}(e) \to K$

$\sigma^2_{\theta_t \mid Y_t} = R_t - K$

$\mu_{\theta \mid R} = G_t \hat{\theta}_t + KF$

Observation Eq.

$$y_t = \beta_t x_t + \epsilon_t, \quad \epsilon_t \sim N(0, R)$$

y_t : this is the observed dependent variable at time t. (Return of SPY)
x_t : this is the independent variable at time t. (Return of DJIA)
β_t : time-varying hedge ratio, how many units of x_t we need to hedge one unit of y_t
ϵ_t : Measument noise or error, Normal variance, with zero mean and variance R

State Eq.

$$\beta_t = \beta_{t-1} + n_t, \quad n_t \sim N(0, Q)$$

The hedge ratio is random walk (follows no specific pattern). The current $\beta_t = \beta_{t-1}$ + random noise

Prediction State: $\hat{\beta}_{t|t-1} = \hat{\beta}_{t-1|t-1}$ since state Eq. is random walk our best prediction of the current state is prev. fitered Estimate $(\hat{\beta}_{t|t-1})$

Preddicted Error Covariance: $P_{t|t-1} = P_{t-1|t-1} + Q$
$P_{t-1|t-1}$ = error covarionce of prev. state Estimate
Q : We add to account for uncertanity introduced by random walk

Innovation (Measurement Residual):
Residual : $v_t = y_t - x_t \beta_{t|t-1}$ is the difference b/w actual y_t and predicted $x_t \beta_{t|t-1}$
Covariance: $S_t = x_t^2 P_{t|t-1} + R$ represents total uncertanity in our v_t
$x_t^2 P_{t|t-1}$: comes from propgating uncertanity of our state Prediction into Measurment
R : is variance of measument noise

Kalman Gain: $K_t = \frac{P_{t|t-1} x_t}{S_t} = \frac{P_{t|t-1} x_t}{x^2 P_{t|t-1} + R}$ tells us how much weight to give to the innovation when updating our estimate.
If $P_{t|t-1}$ (Prediction Uncertanity) is large relative to R (measurment noise)

1. State Equation

$\beta_t = \beta_{t-1} + w_t$, where $w_t \sim N(0, Q)$

2. Measurment Equation

$y_t = \beta_t x_t + \epsilon_t$, where $\epsilon_t \sim N(0, R)$

* Here y_t (OJIH price) depends on x_t (SPY price) based on time variance

Predicted state

$\beta_{t|t-1} = \beta_{t-1|t-1}$

Predicted Covariance

$P_{t|t-1} = P_{t-1|t-1} + Q$

Kalman Gain

$$K_t = \frac{P_{t|t-1} x_t}{x_t^2 P_{t|t-1} + R}$$

Updated Covariance

$\underline{P_{t|t}} = (1 - K_t x_t) P_{t|t-1}$

Substitute the $P_{t|t}$ to the Kalman Gain Equatio

$$\beta_t = \beta_{t-1} + \frac{(P_{t-1} + Q) x_t}{x_t^2 (P_{t-1} + Q) + R} (y_t - \beta_{t-1} x_t)$$

$y_t - \beta_{t-1} x_t$ is innovation term its the residual error between observed y_t and the predicted value using prior beta