Kalman Filters 101

Arnav Sheth
iCapital Network

January 17, 2023

Summary

What is a Kalman Filter?

- State space model or regime-switching model.
- Traditionally rooted in engineering applications.
  - For example, autonomous vehicles or trajectory analysis.
- Now used in economics/finance.
- Can be interpreted as a natural extension of traditional ARMA models[1].
  - Makes it easier to explain.

[1] See Harvey (1989), for example.

What is a Kalman Filter?

- Introduced in 1960 by Rudolf Kalman and Ruslan Bucy[2].

The Kalman filter is an estimator of the conditional moments of a Gaussian linear system. It is an optimal recursive algorithm.

- Used for:
  - the calibration of time series model predictor variables;
  - also for data smoothing applications.
  - Appropriate for dealing with multivariable time-varying systems and non-stationary stochastic processes.
    - Is an advantage over VAR since factors are non-stationary.
    - It's also a more realistic assumption.

[2] See Kalman (1960) and Kalman and Bucy (1961).

Why Use Kalman Filters?

- Kalman Filters have several desirable properties:
  1. They are fast; amenable to large datasets.
  2. Good at dealing with unknown structural breaks and regime changes[3].
  3. They preserve OLS property of being consistent, unbiased, and efficient.
  4. Can handle missing data in a fairly straightforward manner[4].
  5. Because they are constructed as recursions, can handle out-of-sample forecasting well.

[3] See Renzi-Ricci (2016).

[4] See Brockwell and Davis (1991).

A Simple Model

- In its simplest setup, we have[5]:

$$y_t = \mu_t + u_t \tag{1}$$
$$\mu_{t+1} = T_t \mu_t + \eta_t \tag{2}$$

  - Equation (3) is known as the *measurement* equation.
  - Equation (4) is known as the *transition* equation.
- For even more simplicity, let's assume T_t to be constant, or 1.
  - Then μ_t follows a random walk.

[5] These are variables but you can replace them with $K \times 1$ vectors (say).

A Simple Model

- In its simplest setup, we have[6]:

$$y_t = \mu_t + u_t \tag{3}$$
$$\mu_{t+1} = T_t \mu_t + \eta_t \tag{4}$$

  - u_t and η_t are called, respectively, *measurement* and *observation* noise.
  - $u_t \sim N(0, \sigma_u^2)$, $\eta \sim N(0, \sigma_\eta^2$, and $E[u_t \eta_t] = 0$.
- The latent state, μ_t is hidden (unobserved). E.g., the true beta of a company, or state of the economy.
  - The data gives us y_t.

[6] These are variables but you can replace them with $K \times 1$ vectors (say).

A Simple Model

- In its simplest setup, we have[6]:

$$y_t = \mu_t + u_t \tag{3}$$

$$\mu_{t+1} = T_t \mu_t + \eta_t \tag{4}$$

  - u_t and η_t are called, respectively, *measurement* and *observation* noise.
  - $u_t \sim N(0, \sigma_u^2)$, $\eta \sim N(0, \sigma_\eta^2$, and $E[u_t \eta_t] = 0$.
- The latent state, μ_t is hidden (unobserved). E.g., the true beta of a company, or state of the economy.
  - The data gives us y_t.
- To evaluate the size of σ_η^2, we use the ratio of variances: $\sigma_\eta^2 / \sigma_u^2$.

[6] These are variables but you can replace them with $K \times 1$ vectors (say).

A More Complex Model

$$y_t = \alpha + \beta_t x_t + u_t \tag{5}$$
$$\beta_{t+1} = \beta_t + \eta_t \tag{6}$$

- β_t is the hidden state variable here.
- For example:
  - y_t could be the excess return on a stock.
  - x_t could be the excess return on a market proxy.
- We could generalize further and allow Jensen's alpha to vary over time.

A More Complex Model

$$y_t = \alpha + \beta_t x_t + u_t \quad (5)$$
$$\beta_{t+1} = \beta_t + \eta_t \quad (6)$$

- β_t is the hidden state variable here.
- For example:
  - y_t could be the excess return on a stock.
  - x_t could be the excess return on a market proxy.
- We could generalize further and allow Jensen's alpha to vary over time.
  - Note that this is what we called μ_t in the previous example.

Parameter Estimation

- In traditional MLE, parameters are the slope and intercept.
    - Recall that these could be time-varying in state space models.
- In this setup, α, β, and the variances (σ_η^2 and σ_u^2), are all parameters.
- For a *filter*, you plug in variance values and estimate the slope and intercept.
- For a *smoother*, you go back and use MLE to find the values that minimize the variances.

General Steps

1. Specify an initial value for β, β_1.
2. Beginning with $t = 2$, apply the Kalman filter to β_t to provide an estimate for β_{t+1}.
   - Estimates are denoted with hats: $\widehat{\beta}_{t+1}$ and $\widehat{y}_{t+1}$.
   - At $t + 1$ compare with the actual value, and compute the prediction error, termed the *Kalman gain*.
   - Adjust $\hat{\beta}_{t+1}$ appropriately.
3. Use the adjusted estimate of β_{t+1} to get an estimate for β_{t+2}.
4. Rinse and repeat.

References

1. Brockwell, P. J., & Davis, R. A. (1991). *Time series: theory and methods*. Springer Science & Business Media.
2. Brooks, C. (2009). *Introductory Econometrics for Finance*. Cambridge University Press.
3. Harvey, A. C. (1990). *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press.
4. Jalles, J. T. (2009). Structural time series models and the Kalman filter: a concise review, SSRN.
5. Kalman, R. E. (1960). A new approach to linear filtering and prediction problems, *Transactions of the ASME - Journal of Basic Engineering*, 83 (Series D): 95-108.
6. Kalman, R. E., & Bucy, R. S. (1961). New results in linear filtering and prediction theory, *Transactions of the ASME - Journal of Basic Engineering*, 82 (Series D): 35-45.
7. Ljungqvist, L., & Sargent, T. J. (2018). *Recursive macroeconomic theory*. MIT press.
8. Renzi-Ricci, G. (2016). Estimating equity betas: What can a time-varying approach add. A comparison of ordinary least squares and the Kalman filter. *Insights in Economics*, 1-9.