Maximum Likelihood Estimation

Kalman Filter

Tools for Macroeconomists: The essentials

Petr Sedláček

Kalman Filter

What was the Kalman filter originally developed for?

Kalman Filter

Time series model

Time series model

$$
\begin{aligned}
y_t =& H'\zeta_t + w_t, & \mathbb{E}(w_t, w_t') = R \quad & \forall t \\
\zeta_{t+1} =& F\zeta_t + v_{t+1}, & \mathbb{E}(v_t, v_t') = Q \quad & \forall t
\end{aligned}
$$

- y_t is observed, but ζ_t is not

Time series model

$$
\begin{aligned}
y_t &= H'\zeta_t + w_t, & \mathbb{E}(w_t, w_t') &= R \quad \forall t \\
\zeta_{t+1} &= F\zeta_t + v_{t+1}, & \mathbb{E}(v_t, v_t') &= Q \quad \forall t
\end{aligned}
$$

- y_t is observed, but ζ_t is not
- Kalman filter enables you to get an estimate of ζ_t

Some preliminaries

- for now assume coefficients are known
  - later on we will show how to estimate them
  - they could even be time-varying

Some preliminaries

- for now assume coefficients are known
  - later on we will show how to estimate them
  - they could even be time-varying
- initial conditions: ζ_1 has mean $\widehat{\zeta}_{1|0}$ and variance $P_{1|0}$

Some preliminaries

- for now assume coefficients are known
  - later on we will show how to estimate them
  - they could even be time-varying
- initial conditions: ζ_1 has mean $\widehat{\zeta}_{1|0}$ and variance $P_{1|0}$
- state and observation disturbances are
  - uncorrelated over time
  - uncorrelated with each other (at all leads and lags)
  - orthogonal to ζ_1

Kalman Filter

Main idea

Purpose of the Kalman filter

- calculate the expectation of the unobserved states
- given observations on y

$$\widehat{\zeta}_{t+1} = \widehat{\mathbb{E}}(\zeta_{t+1}|\mathcal{Y}_t),$$

$$\mathcal{Y}_t = (y_t', y_{t-1}', ..., y_1')$$

What is the objective?

How to pick the (right) forecast?

WHAT IS THE OBJECTIVE?

How to pick the (right) forecast?

- specify a *loss function* (objective function)
- evaluate the “usefulness” of the forecast according to it

What is the objective?

How to pick the (right) forecast?

- specify a *loss function* (objective function)
- evaluate the "usefulness" of the forecast according to it
- convenient results obtained with a quadratic loss function

$$\min E(\zeta_{t+1} - \widehat{\zeta}_{t+1|t})^2$$

What is the objective?

How to pick the (right) forecast?

- specify a *loss function* (objective function)
- evaluate the "usefulness" of the forecast according to it
- convenient results obtained with a quadratic loss function

$$\min E(\zeta_{t+1} - \widehat{\zeta}_{t+1|t})^2$$

- assuming a linear functional form
- $\rightarrow \widehat{\zeta}_{t+1|t}$ is a *linear projection* of ζ_{t+1} on its regressors

- closely related to OLS
- assume existence of following first and second moments:
  - $\bar{z} = E[z]$, $\bar{x} = E[x]$
  - $\Sigma_{z,z} = E[(z - \bar{z})(z - \bar{z})']$
  - $\Sigma_{x,x} = E[(x - \bar{x})(x - \bar{x})']$
  - $\Sigma_{z,x} = E[(z - \bar{z})(x - \bar{x})']$

Linear projection digression

- closely related to OLS
- assume existence of following first and second moments:
  - $\bar{z} = E[z]$, $\bar{x} = E[x]$
  - $\Sigma_{z,z} = E[(z - \bar{z})(z - \bar{z})']$
  - $\Sigma_{x,x} = E[(x - \bar{x})(x - \bar{x})']$
  - $\Sigma_{z,x} = E[(z - \bar{z})(x - \bar{x})']$
- linear projection (of z on x) is a function $\widehat{E}[z|x] = a + bx$

Linear projection digression

Linear projection digression

linear projection picks a and b to minimize MSE:

$$b = \frac{\Sigma_{z,x}}{\Sigma_{x,x}}$$

$$a = \bar{z} - b\bar{x}$$

Linear projection digression

linear projection picks a and b to minimize MSE:

$$b = \frac{\Sigma_{z,x}}{\Sigma_{x,x}}$$

$$a = \bar{z} - b\bar{x}$$

$$\widehat{z} = \widehat{E}(z|x) = \bar{z} + \frac{\Sigma_{z,x}}{\Sigma_{x,x}}(x - \bar{x})$$

Linear projection vs. linear regression

Linear projection vs. linear regression

- linear regression (OLS) seeks effect of x on z
- keeping all else (including error term) constant
- linear projection is concerned “only” with forecasting
- $\rightarrow$ doesn’t matter if $x \rightarrow z$ or $z \rightarrow x$

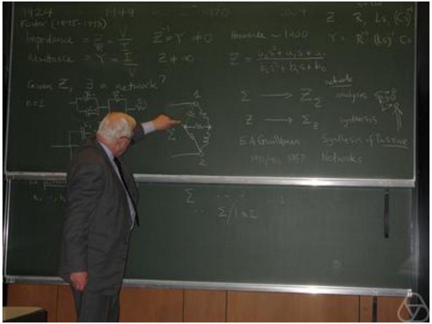
Kalman Filter

Derivation

Back to the Kalman filter

Back to the Kalman filter

Back to the Kalman filter



Back to the Kalman filter

- purpose of Kalman filter is to estimate unobserved states
- will do so using linear projections (given data and structure)
- main trick is to formulate it recursively

Main idea of the Kalman filter recursions:

Main idea of the Kalman filter recursions:

- given starting values for $\zeta_{1|0}$ and observation of y_1
- $\rightarrow$ forecast $\widehat{\zeta}_{2|1}$

Formulating a recursive scheme

Main idea of the Kalman filter recursions:

- given starting values for $\zeta_{1|0}$ and observation of y_1
- $\rightarrow$ forecast $\widehat{\zeta}_{2|1}$
- given $\widehat{\zeta}_{2|1}$ and observation of y_2
- $\rightarrow$ forecast $\widehat{\zeta}_{3|2}$ ...

For convenience, split the above into the following steps:

Formulating a recursive scheme

For convenience, split the above into the following steps:

- given starting values for $\zeta_{1|0}$ and observation of y_1
- $\rightarrow$*update* state $\widehat{\zeta}_{1|1}$

Formulating a recursive scheme

For convenience, split the above into the following steps:

- given starting values for $\zeta_{1|0}$ and observation of y_1
- $\rightarrow$*update* state $\widehat{\zeta}_{1|1}$
- given $\widehat{\zeta}_{1|1}$ and model structure
- $\rightarrow$ forecast $\widehat{\zeta}_{2|1}$ ...

1. Update step

1. Update step

use linear projection to produce update of ζ_t

1. Update step

use linear projection to produce update of ζ_t

- conditional on expectations from $t-1$
- conditional on observation of y_t

$$\widehat{z} = \widehat{E}(z|x) = \bar{z} + \frac{\Sigma_{z,x}}{\Sigma_{x,x}}(x - \bar{x})$$

1. Update step

use linear projection to produce update of ζ_t

- conditional on expectations from $t-1$
- conditional on observation of y_t

$$\widehat{z} = \widehat{E}(z|x) = \bar{z} + \frac{\Sigma_{z,x}}{\Sigma_{x,x}}(x - \bar{x})$$

$$\begin{aligned}\widehat{\zeta}_{t|t} =& \widehat{\mathbb{E}}[\zeta_t|\mathcal{Y}_t] = \widehat{\zeta}_{t|t-1} + \\ & \mathbb{E}\left[(\zeta_t - \widehat{\zeta}_{t|t-1})(y_t - \widehat{y}_{t|t-1})'\right] x \mathbb{E}\left[(y_t - \widehat{y}_{t|t-1})(y_t - \widehat{y}_{t|t-1})'\right]^{-1} x \\ & (y_t - \widehat{y}_{t|t-1})\end{aligned} \tag{1}$$

Using the model structure

covariance term

Using the model structure

covariance term

$$
\begin{aligned}
&\mathbb{E}\left[(\zeta_t - \widehat{\zeta}_{t|t-1})(y_t - \widehat{y}_{t|t-1})'\right] \\
&\quad = \mathbb{E}\left[(\zeta_t - \widehat{\zeta}_{t|t-1})(H'(\zeta_t - \widehat{\zeta}_{t|t-1}) + w_t)'\right] \\
&\quad = \mathbb{E}\left[(\zeta_t - \widehat{\zeta}_{t|t-1})(\zeta_t - \widehat{\zeta}_{t|t-1})'H\right] \\
&\quad = P_{t|t-1}H
\end{aligned}
\tag{2}
$$

$P_{t|t-1}$ is the related MSE

Using the model structure

variance term

variance term

$$
\begin{aligned}
&\mathbb{E}\left[(y_t - \widehat{y}_{t|t-1})(y_t - \widehat{y}_{t|t-1})'\right] \qquad (3)\\
&\quad = \mathbb{E}\left[H'(\zeta_t - \widehat{\zeta}_{t|t-1})(\zeta_t - \widehat{\zeta}_{t|t-1})H\right] + \mathbb{E}\left[w_t w_t'\right]\\
&\quad = H' P_{t|t-1} H + R
\end{aligned}
$$

Using the model structure

error term

error term

$$\widetilde{y}_{t|t-1} = y_t - \widehat{y}_{t|t-1} = y_t - H'\widehat{\zeta}_{t|t-1} \tag{4}$$

Our update of ζ_t given information up until period t:

$$\widehat{\zeta}_{t|t} = \widehat{\zeta}_{t|t-1} + P_{t|t-1} H (H' P_{t|t-1} H + R)^{-1} (y_t - H' \widehat{\zeta}_{t|t-1})$$

2. Forecast step

2. Forecast step

use model structure to forecast the state

2. Forecast step

use model structure to forecast the state

$$
\begin{aligned}
\widehat{\zeta}_{t+1|t} =& \widehat{\mathbb{E}}\left[\zeta_{t+1}|\mathcal{Y}_t\right] \\
=& F\widehat{\mathbb{E}}\left[\zeta_t|\mathcal{Y}_t\right] + \widehat{\mathbb{E}}\left[v_{t+1}|\mathcal{Y}_t\right] \\
=& F\widehat{\mathbb{E}}\left[\zeta_t|\mathcal{Y}_t\right]
\end{aligned} \tag{5}
$$

Collapsing the two steps

Combining (1) and (5) and using (2) to (4) we can write

$$\widehat{\zeta}_{t+1|t} = F\widehat{\zeta}_{t|t-1} + \underbrace{FP_{t|t-1}H(H'P_{t|t-1}H + R)^{-1}}_{\text{Kalman gain}}(y_t - H'\widehat{\zeta}_{t|t-1})$$

We're not done yet

We're not done yet

Still need to define recursions for P's:

We're not done yet

Still need to define recursions for P's:

Update step (use (1) to substitute out $\widehat{\zeta}_{t|t}$)

$$
\begin{aligned}
P_{t|t} =& \mathbb{E}\left[(\zeta_t - \widehat{\zeta}_{t|t})(\zeta_t - \widehat{\zeta}_{t|t})'\right] \\
=& \mathbb{E}\left[(\zeta_t - \widehat{\zeta}_{t|t-1})(\zeta_t - \widehat{\zeta}_{t|t-1})'\right] \\
& - \mathbb{E}\left[(\zeta_t - \widehat{\zeta}_{t|t-1})(y_t - \widehat{y}_{t|t-1})'\right] \times \\
& \mathbb{E}\left[(y_t - \widehat{y}_{t|t-1})(y_t - \widehat{y}_{t|t-1})'\right]^{-1} \times \\
& \mathbb{E}\left[(y_t - \widehat{y}_{t|t-1})(\zeta_t - \widehat{\zeta}_{t|t-1})'\right] \\
=& P_{t|t-1} - P_{t|t-1} H (H' P_{t|t-1} H + R)^{-1} H' P_{t|t-1}
\end{aligned} \tag{6}
$$

We're not done yet

Forecast step (use (5) to substitute out $\widehat{\zeta}_{t+1|t}$)

Forecast step (use (5) to substitute out $\widehat{\zeta}_{t+1|t}$)

$$\begin{aligned}
P_{t+1|t} =& \mathbb{E}\left[(\zeta_{t+1} - \widehat{\zeta}_{t+1|t})(\zeta_{t+1} - \widehat{\zeta}_{t+1|t})'\right] \\
=& \mathbb{E}\left[(F\zeta_t + v_{t+1} - F\widehat{\zeta}_{t|t})(F\zeta_t + v_{t+1} - F\widehat{\zeta}_{t|t})'\right] \\
=& F\mathbb{E}\left[(\zeta_t - \widehat{\zeta}_{t|t})(\zeta_t - \widehat{\zeta}_{t|t})'\right] F' + \mathbb{E}\left[v_{t+1}v'_{t+1}\right] \\
=& FP_{t|t}F' + Q
\end{aligned} \tag{7}$$

Collapsing the two steps

Combining (6) and (7) we can write

$$P_{t+1|t} = F\left[P_{t|t-1} - P_{t|t-1}H(H'P_{t|t-1}H + R)^{-1}H'P_{t|t-1}\right]F' + Q$$

Summary of recursive formulation

update:

$$\widehat{\zeta}_{t|t} = \widehat{\zeta}_{t|t-1} + P_{t|t-1}H(H'P_{t|t-1}H + R)^{-1}(y_t - H'\widehat{\zeta}_{t|t-1})$$

$$P_{t|t} = P_{t|t-1} - P_{t|t-1}H(H'P_{t|t-1}H + R)^{-1}H'P_{t|t-1}$$

forecast:

$$\widehat{\zeta}_{t+1|t} = F\widehat{\zeta}_{t|t}$$

$$P_{t+1|t} = FP_{t|t}F' + Q$$

Summary of recursive formulation

combined specification:

$$\widehat{\zeta}_{t+1|t} = F\widehat{\zeta}_{t|t-1} + K_t(y_t - H'\widehat{\zeta}_{t|t-1})$$

$$K_t = FP_{t|t-1}H(H'P_{t|t-1}H + R)^{-1}$$

$$P_{t+1|t} = F\left[P_{t|t-1} - P_{t|t-1}H(H'P_{t|t-1}H + R)^{-1}H'P_{t|t-1}\right]F' + Q$$

Kalman Filter

Uses

What was the Kalman filter originally developed for?

Journal of the
Royal Statistical Society

SERIES A
Statistics
in Society

J. R. Statist. Soc.

A dynamic bivariate Poisson model for analysing and forecasting match results in the English Premier League

Siem Jan Koopman and Rutger Lit

VU University Amsterdam, The Netherlands

[Received September 2012. Revised July 2013]

Summary. We develop a statistical model for the analysis and forecasting of football match results which assumes a bivariate Poisson distribution with intensity coefficients that change stochastically over time. The dynamic model is a novelty in the statistical time series analysis of match results in team sports. Our treatment is based on state space and importance sampling methods which are computationally efficient. The out-of-sample performance of our methodology is verified in a betting strategy that is applied to the match outcomes from the 2010–2011 and 2011–2012 seasons of the English football Premier League. We show that our statistical modelling framework can produce a significant positive return over the bookmaker's odds.

Journal of the
Royal Statistical Society
SERIES A Statistics in Society

J. R. Statist. Soc. A (2014)

A dynamic bivariate Poisson model for analysing and forecasting match results in the English Premier League

Siem Jan Koopman and Rutger Lit

VU University Amsterdam, The Netherlands

[Received September 2012. Revised July 2013]

Summary. We develop a statistical model for the analysis and forecasting of football match results which assumes a bivariate Poisson distribution with intensity coefficients that change stochastically over time. The dynamic model is a novelty in the statistical time series analysis of match results in team sports. Our treatment is based on state space and importance sampling methods which are computationally efficient. The out-of-sample performance of our methodology is verified in a betting strategy that is applied to the match outcomes from the 2010–2011 and 2011–2012 seasons of the English football Premier League. We show that our statistical modelling framework can produce a significant positive return over the bookmaker's odds.

Kalman Filter

Taking stock

Taking stock

Kalman filter

- method for estimating unobserved driving forces
  - given measurements of observed variables
  - and in the presence of uncertainty

Taking stock

Kalman filter

- method for estimating unobserved driving forces
  - given measurements of observed variables
  - and in the presence of uncertainty
- applied recursively to linear (state-space) systems
- has vast applications beyond Economics


That's all Folks!