

# MA Advanced Macroeconomics:

## 5. Latent Variables: The Kalman Filter

Karl Whelan

School of Economics, UCD

Spring 2016

# Unobserved But Still There

- Sometimes in macroeconomics, we come across variables that play important roles in theoretical models but which we cannot observe.
- Examples include the concept of potential output. For example, in many Keynesian models, inflationary pressures are determined by how far actual output is from this time-varying potential output series.
- In reality, we do not observe potential output so is this concept even worth bothering with? Well, just because a variable isn't observable, that doesn't mean we can't make a guess as to how it is behaving.
- For example, if our data move in a way that would be consistent with a large increase in potential output (perhaps GDP rises a lot but there are no signs of inflationary pressures) then perhaps we should assume that it has indeed increased.
- In these notes, we will discuss methods for dealing with unobserved (or latent) variables in time series, building towards a method known as the **Kalman filter**. We will see the Kalman filter again when we discuss estimation of DSGE models.

# Extracting a Signal

- Going back to our earlier point, suppose we see a big increase in output in the latest quarterly data that is not accompanied by a burst of inflation?
- Does this mean we should assume there has been a big change in potential output?
- Probably not. Potential output probably doesn't move around a lot from quarter to quarter and it is likely that there is a lot of fairly random *noise* in the quarterly fluctuations in inflation.
- But there is also probably a useful *signal* in the data as well.
- So we are dealing with a type of signal extraction problem: What's the best way to extract a useful signal from information that also contains useless noise?
- We discuss these ideas in the next few slides, developing the tools that are used for the Kalman filter.

# Conditional Expectations

- Suppose we are interested in getting an estimate of the value of a variable  $X$ . We don't observe  $X$  but instead we observe a variable  $Z$  that we know to be correlated with  $X$ .
- Specifically, let's assume that  $X$  and  $Z$  are jointly normally distributed so that

$$\begin{pmatrix} X \\ Z \end{pmatrix} \sim N \left( \begin{pmatrix} \mu_X \\ \mu_Z \end{pmatrix}, \begin{pmatrix} \sigma_X^2 & \sigma_{XZ} \\ \sigma_{XZ} & \sigma_Z^2 \end{pmatrix} \right)$$

- In this case, the expected value of  $X$  conditional on observing  $Z$  is

$$E(X|Z) = \mu_X + \frac{\sigma_{XZ}}{\sigma_Z^2} (Z - \mu_Z)$$

- Alternatively, if  $\rho$  is the correlation between  $X$  and  $Z$  ( $\rho = \frac{\sigma_{XZ}}{\sigma_X \sigma_Z}$ ) then we can write

$$E(X|Z) = \mu_X + \rho \frac{\sigma_X}{\sigma_Z} (Z - \mu_Z)$$

- The amount of weight you put on the information in  $Z$  when formulating an expectation for  $X$  depends on how correlated  $Z$  is with  $X$  and on their relative standard deviation. If  $Z$  has a high standard deviation (so it's a poor signal) then you don't place much weight on it.

# Multivariate Conditional Expectations

- In the previous example, we considered only two variables. What about the case in which  $X$  is a  $1 \times n$  vector of variables and  $Z$  is a  $1 \times m$  vector? There is a straightforward generalisation of the formula just presented.
- Denote the covariance matrix of the variables in  $X$  as  $\Sigma_{XX}$ , the covariance matrix of the variables in  $Z$  as  $\Sigma_{ZZ}$  and the matrix of covariances between the entries in  $X$  and  $Z$  as  $\Sigma_{XZ}$ .

- If all the variables are jointly normally distributed then this can be written as

$$\begin{pmatrix} X \\ Z \end{pmatrix} \sim N \left( \begin{pmatrix} \mu_X \\ \mu_Z \end{pmatrix}, \begin{pmatrix} \Sigma_{XX} & \Sigma_{XZ} \\ \Sigma'_{XZ} & \Sigma_{ZZ} \end{pmatrix} \right)$$

- In this case, the expected value of  $X$  conditional on observing  $Z$  is

$$E(X|Z) = \mu_X + \Sigma_{XZ} \Sigma_{ZZ}^{-1} (Z - \mu_Z)$$

- This formula will play an important role in our explanation of the Kalman filter.

# State-Space Models

- State-space models are a general class of linear time series models that mix together observable and unobservable variables. These models can be described using two equations.
- The first, known as the **state** or **transition equation**, describes how a set of unobservable state variables,  $S_t$ , evolve over time as follows:

$$S_t = FS_{t-1} + u_t$$

The term  $u_t$  can include either normally-distributed errors or perhaps zeros if the equation being described is an identity. We will write this as  $u_t \sim N(0, \Sigma^u)$  though  $\Sigma^u$  may not have a full matrix rank.

- The second equation in a state-space model, which is known as the **measurement equation**, relates a set of observable variables,  $Z_t$ , to the unobservable state variables

$$Z_t = HS_t + v_t$$

Again, the term  $w_t$  can include either normally-distributed errors or perhaps zeros if the equation being described is an identity. We will write this as  $v_t \sim N(0, \Sigma^v)$  though  $\Sigma^v$  may not have a full matrix rank.

# Estimation of State-Space Models: Intuition

- Before working through the details of the methods used to estimate state-space models via maximum likelihood, let's build-up some intuition first.
- The observed data are described by

$$Z_t = HS_t + v_t$$

where  $v_t$  contains some normally-distributed errors.

- We can't observe  $S_t$  but suppose we could replace it by an observable unbiased guess based on information available up to time  $t - 1$ . Call this guess  $S_{t|t-1}$  and suppose its errors are normally distributed with a known covariance matrix.

$$S_t - S_{t|t-1} \sim N\left(0, \Sigma_{t|t-1}^S\right)$$

- Then the observed variables could be written as

$$Z_t = HS_{t|t-1} + v_t + H(S_t - S_{t|t-1})$$

- Because  $S_{t|t-1}$  is observable and the unobservable elements ( $v_t$  and  $S_t - S_{t|t-1}$ ) are normally distributed, this model can be estimated via maximum-likelihood methods.

# Log Likelihood of Observables

- The observable data can be written as

$$Z_t = HS_{t|t-1} + v_t + H(S_t - S_{t|t-1})$$

- Then the variance of error term after conditioning on period  $t - 1$ 's estimate of the state variables is given by

$$v_t + H(S_t - S_{t|t-1}) \sim N(0, \Omega_t)$$

where

$$\Omega_t = \Sigma^v + H\Sigma_{t|t-1}^S H'$$

- Let  $\theta$  represent the parameters of the model, i.e.  $\theta = (F, H, \Sigma^v, \Sigma^u)$ . We will show later that  $\Sigma_{t|t-1}^S$  depends on these parameters.
- The log-likelihood function for  $Z_t$  given the observables at time  $t - 1$  is

$$\log f(Z_t | Z_{t-1}, \theta) = -\log 2\pi - \log |\Omega_t| - \frac{1}{2} (Z_t - HS_{t|t-1})' \Omega_t^{-1} (Z_t - HS_{t|t-1})$$



# Maximum Likelihood Estimation

- Given the initial estimates of the first-period unobservable state  $S_{1|0}$ , the combined likelihood for all the observed data is the product of all the period-by-period likelihoods

$$f(Z_1, Z_2, \dots, Z_T | S_{1|0}, \theta) = f(Z_1 | S_{1|0}, \theta) \prod_{i=2}^{i=T} f(Z_i | Z_{i-1}, \theta)$$

- So the combined log-likelihood function for the observed dataset is given by

$$\begin{aligned} \log f(Z_1, Z_2, \dots, Z_T | S_{1|0}, \theta) &= -T \log 2\pi - \sum_{i=1}^T \log |\Omega_i| \\ &\quad - \frac{1}{2} \sum_{i=1}^T (Z_i - HS_{i|i-1})' \Omega_i^{-1} (Z_i - HS_{i|i-1}) \end{aligned}$$

- The maximum-likelihood parameter estimates are the set of matrices  $\theta = (F, H, \Sigma^v, \Sigma^u)$  that provide the largest value for this function.

# Estimating the State Variables

- We have described how to estimate the model's parameters via MLE provided we have an unbiased guess based on information available up to time  $t - 1$ , which we called  $S_{t|t-1}$ , with normally distributed errors.
- Here we describe a method for generating these unbiased guesses known as the Kalman filter. It is an iterative method. Starting from one period's estimates of the state variables, it uses the observable data for the next period to update these estimates.
- Let's start with formulating an estimate of the state variable at time  $t$  given information at time  $t - 1$ . This is easy enough

$$S_t = FS_{t-1} + u_t \Rightarrow S_{t|t-1} = FS_{t-1|t-1}$$

- This means that in period  $t - 1$ , the expected value for the observables in period  $t$  are

$$Z_{t|t-1} = HS_{t|t-1} = HFS_{t-1|t-1}$$

- Then in period  $t$ , when we observe  $Z_t$  the question is how do we update our guesses for the state variable in light of the “news” in  $Z_t - HFS_{t-1|t-1}$ ?

# Conditional Expectation: The Kalman Filter

- The assumptions of the model imply that

$$\begin{pmatrix} S_t \\ Z_t \end{pmatrix} \sim N \left( \begin{pmatrix} FS_{t-1|t-1} \\ HFS_{t-1|t-1} \end{pmatrix}, \begin{pmatrix} \Sigma_{t|t-1}^S & (H\Sigma_{t|t-1}^S)' \\ H\Sigma_{t|t-1}^S & \Sigma^v + H\Sigma_{t|t-1}^S H' \end{pmatrix} \right)$$

- Now we can use our earlier result about conditional expectations to state that the minimum variance unbiased estimate of  $S_t$  given the observed  $Z_t$  is

$$E(S_t|Z_t) = S_{t|t} = FS_{t-1|t-1} + K_t (Z_t - HFS_{t-1|t-1})$$

where

$$K_t = \left( H\Sigma_{t|t-1}^S \right)' \left( \Sigma^v + H\Sigma_{t|t-1}^S H' \right)^{-1}$$

- The covariance matrices required to compute this  $K_t$  matrix (known as the **Kalman gain** matrix) are updated by the formulae:

$$\begin{aligned} \Sigma_{t|t-1}^S &= F\Sigma_{t-1|t-1}^S F' + \Sigma^u \\ \Sigma_{t|t}^S &= (I - K_t H) \Sigma_{t|t-1}^S \end{aligned}$$

# Initialising the Kalman Filter

- We still need an initial estimate  $S_{1|0}$  as well as its covariance matrix to start the filter process.
- In many macroeconomic models, the state variable can be assumed to have a zero mean without losing any generality, so that can work as a first guess for the state.
- We can estimate the unconditional variance by estimating what the variance of an estimate of the state variable would be from a large sample of data.

Recall that

$$\Sigma_{t|t-1}^S = F \Sigma_{t-1|t-1}^S F' + \Sigma^u$$

- The values of the covariance matrix generated by this equation will generally converge, so for our unconditional covariance matrix we can use a value of  $\Sigma$  that solves

$$\Sigma = F \Sigma F' + \Sigma^u$$

# The Kalman Smoother

- The Kalman filter is what is known as a **one-sided** filter: The estimates of states at time  $t$  are based solely on information available at time  $t$ . No data after period  $t$  is used to calculate estimates of the unobserved state variables.
- This is a reasonable model for how someone might behave if they were learning about the state variables in real time. But researchers have access to the full history of the data set, including all the observations after time  $t$ .
- For this reason, economists generally estimate time-varying models using a method known as the **Kalman smoother**. This is a **two-sided** filter that uses data both before and after time  $t$  to compute expected values of the state variables at time  $t$ .
- I don't want to overload the technique here so I won't go into how this is done: Basically, you do the Kalman filter first and then work backwards from the final estimates further exploiting joint distribution properties.

## Example: The Hodrick-Prescott Filter

- Previously, we discussed the HP filter as a way of choosing a trend for series that had a time-varying trend  $Y_t^*$  picked to minimize

$$\sum_{t=1}^N \left[ (Y_t - Y_t^*)^2 + \lambda (\Delta Y_t^* - \Delta Y_{t-1}^*)^2 \right]$$

- This seems fairly *ad hoc* but it can be viewed as an example of the Kalman filter. Consider the following state-space model

$$\begin{aligned} Y_t &= Y_t^* + C_t \\ \Delta Y_t^* &= \Delta Y_{t-1}^* + \epsilon_t^g \\ C_t &= \epsilon_t^c \end{aligned}$$

where  $\text{Var}(\epsilon_t^g) = \sigma_g^2$  and  $\text{Var}(\epsilon_t^c) = \sigma_c^2$

- It can be shown that for large samples the HP filter technique is the same as Kalman filter estimation of this model when we set  $\lambda = \frac{\sigma_c^2}{\sigma_g^2}$ . Hodrick and Prescott assumed  $C_t$  had a standard deviation of 5 percentage points while  $\epsilon_t^g$  had a standard deviation of one-eighth of a percentage point. Hence they chose  $\lambda = \frac{5^2}{(\frac{1}{8})^2} = (25)(64) = 1600$ .

## Example: Laubach and Williams (2001)

- This paper uses the Kalman filter to estimate a model featuring two unobservable time-varying series: Potential output and the natural rate of interest.
- The model has seven equations

$$\begin{aligned}\tilde{y}_t &= y_t - y_t^* \\ \tilde{y}_t &= A_y(L) \tilde{y}_{t-1} + A_r(L) (r_{t-1} - r_{t-1}^*) + \epsilon_{1t} \\ \pi_t &= B_\pi(L) \pi_{t-1} + B_y(L) \tilde{y}_{t-1} + B_x(L) x_t + \epsilon_{2t} \\ r_t^* &= c g_t + z_t \\ z_t &= D_z(L) z_{t-1} + \epsilon_{3t} \\ y_t^* &= y_{t-1}^* + g_{t-1} + \epsilon_{4t} \\ g_t &= g_{t-1} + \epsilon_{5t}\end{aligned}$$

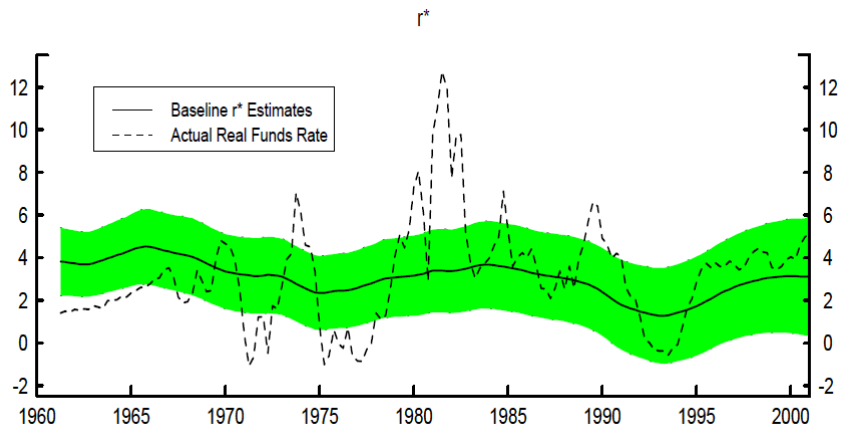
where  $y_t$  is log of real GDP,  $y_t^*$  is the log of potential output,  $\tilde{y}_t$  is the output gap,  $r_t$  is the real Federal funds rate,  $r_t^*$  is the “natural” real rate of interest,  $\pi_t$  is inflation,  $x_t$  is a set of additional variables that determine inflation and  $g_t$  is the growth rate of potential output.

## Laubach and Williams Results

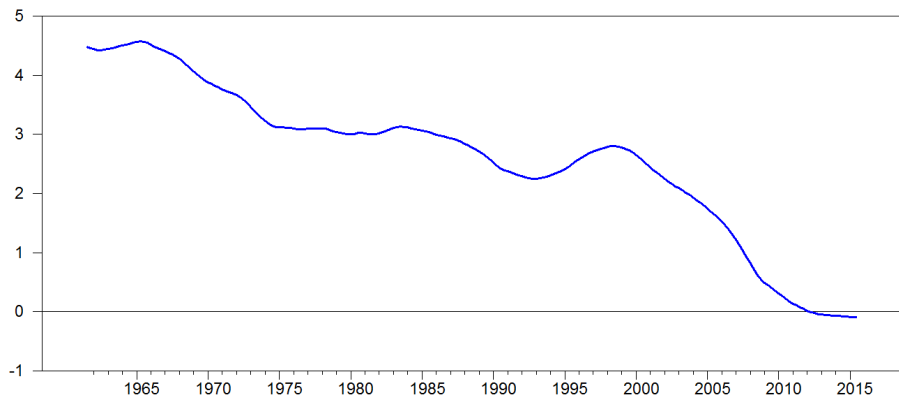
- In addition to the results in the original paper, updated estimates are available on the San Francisco Fed website.
- The next page shows the original paper's estimates of the natural real rate of interest (i.e. the real interest rate that stabilises the economy) with the green area showing standard errors,
- The following page shows the updated estimates of the natural rate using the two-sided Kalman smoother. While the original paper had shown the natural rate as a stationary series, the latest estimates show a steady downward trend so that the current “natural real rate” is negative.
- The model also sees a steady decline in the growth rate of potential output. It assigns much of the weak growth of recent years to structural factors (output gap in recent recession is quite small).
- Final pictures show differences between two-sided and one-sided estimates of natural rate of interest and potential output growth. You can see why they call it the Kalman smoother.
- RATS code for estimating the Laubach-Williams model is linked to on the website.



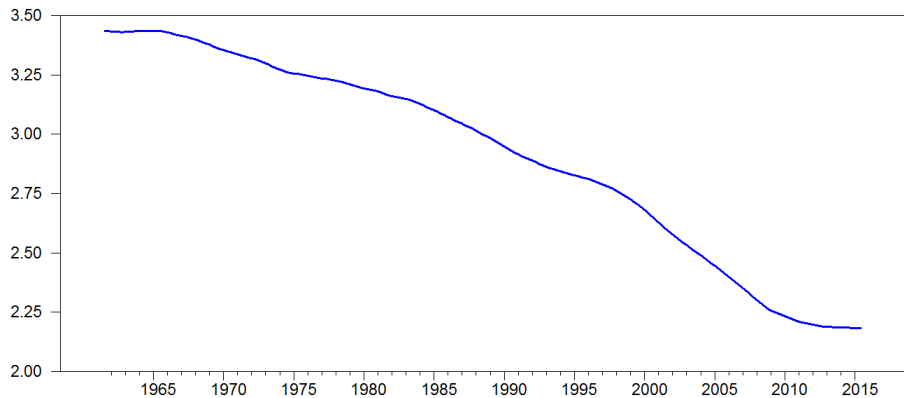
# Actual Versus Natural Real Interest Rates



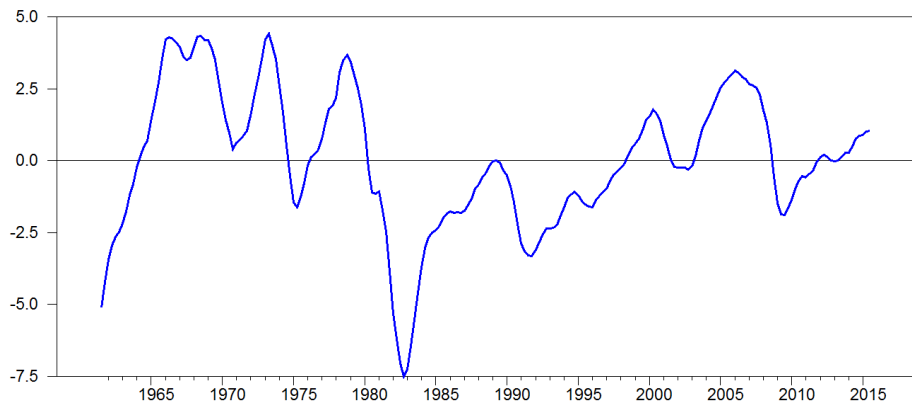
# Laubach-Williams Estimates of Natural Rate of Interest



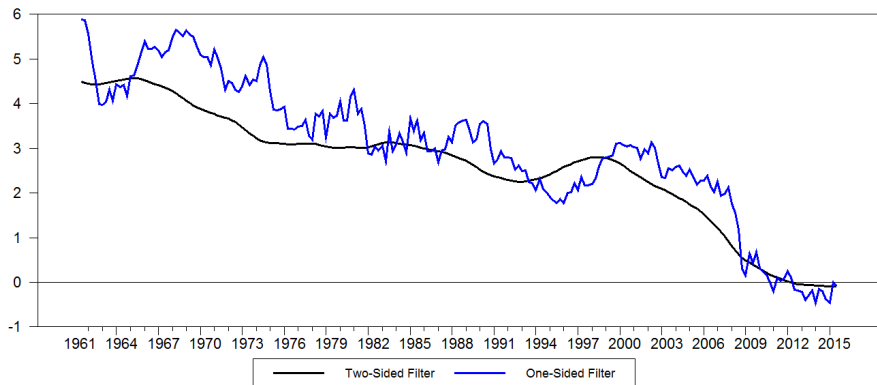
# Laubach-Williams Estimates of Potential Output Growth



# Laubach-Williams Estimates of Output Gap



# One-Sided and Two-Sided Estimates of Natural Rate of Interest



# One-Sided and Two-Sided Estimates of Potential Output Growth

