Kalman filter applications

May 26, 2011

The Kalman Filter

What we did last time:

- The scalar filter
  - Combining period t prior and signal is analogous to a simple minimum variance problem with two signals
- Derived the multivariate filter using
  - The projection theorem
  - Projecting onto orthogonal variables
  - The Gram-Schmidt procedure

The basic formulas

The state space system and the Kalman update equation

$$
\begin{aligned}
X_t &= AX_{t-1} + C\mathbf{u}_t : \mathbf{u}_t \sim N(0, I) \\
Z_t &= DX_t + \mathbf{v}_t : \mathbf{v}_t \sim N(0, \Sigma_{vv})
\end{aligned}
$$

$$X_{t|t} = AX_{t-1|t-1} + K_t \left(Z_t - DX_{t|t-1}\right)$$

where K_t is the Kalman gain and $X_{t|t} = E\left[X_t \mid Z^t, X_{0|0}\right]$

- Filter is also the *linear* minimum variance estimator of X_t even if shocks are non-gaussian.

The basic formulas

Most of the time, all you really need to know is how to put these formulas into a computer

$$X_{t|t} = AX_{t-1|t-1} + K_t \left(Z_t - DX_{t|t-1} \right)$$

$$\begin{aligned} K_t &= P_{t|t-1} D' \left(DP_{t|t-1} D' + \Sigma_{vv} \right)^{-1} \\ P_{t+1|t} &= A \left(P_{t|t-1} - P_{t|t-1} D' \left(DP_{t|t-1} D' + \Sigma_{vv} \right)^{-1} DP_{t|t-1} \right) A' \\ &\quad + CC' \end{aligned}$$

This will give you a recursive estimate of X_t

Kalman filter application: Estimating the output gap in real time

Consider the basic New-Keynesian model

$$\begin{aligned}
\pi_t &= \beta E_t \pi_{t+1} + \kappa (y_t - \overline{y}_t) + v_t^{\pi} : v_t^{\pi} \sim N(0, \sigma_{\pi}^2) \\
y_t &= E_t y_{t+1} - \gamma \left(i_t - E_t \pi_{t+1} \right) + v_t^{y} : v_t^{y} \sim N(0, \sigma_{y}^2) \\
i_t &= \phi \pi_t \\
\overline{y}_t &= \rho \overline{y}_{t-1} + u_t : u_t \sim N(0, \sigma_u^2)
\end{aligned}$$

Can we use the Kalman filter to form an estimate of potential output $\overline{y}_t$?

Estimating the output gap in real time

The solved model can be put in state space form

$$\begin{aligned}
\overline{y}_t &= \rho \overline{y}_{t-1} + u_t \\
\pi_t &= \frac{\kappa(1-\rho)}{\Psi}\overline{y}_t + v_t^{\pi} \\
y_t &= \frac{\kappa\gamma(\rho-\phi)}{\Psi}\overline{y}_t + v_t^{y}
\end{aligned}$$

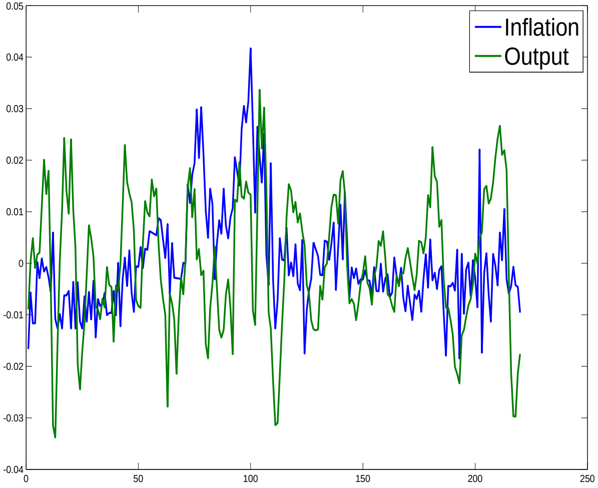
where

$$\Psi = \rho + \beta\rho - \beta\rho^2 - \kappa\gamma\phi + \kappa\gamma\rho - 1$$

so that $A = \rho$, $C = 1$ and $D = [\frac{\kappa(1-\rho)}{\Psi} \quad \frac{\kappa\gamma(\rho-\phi)}{\Psi}]'$

Running the filter

1. Import time series of observable variables y_t and π_t
2. Choose "deep" parameters determining A, C, D and Σ_{vv}
3. Compute the filtered estimates using update equations for $X_{t|t}, K_t$ and $P_{t|t}$.
4. Plot $X_{t|t}$ for $t = 1, 2, ...T$



Inflation
Output
0.05
0.04
0.03
0.02
0.01
0
-0.01
-0.02
-0.03
-0.04
0
50
100
150
200
250

Parameters

Normally, the structural (or "deep") parameters need to be estimated or calibrated

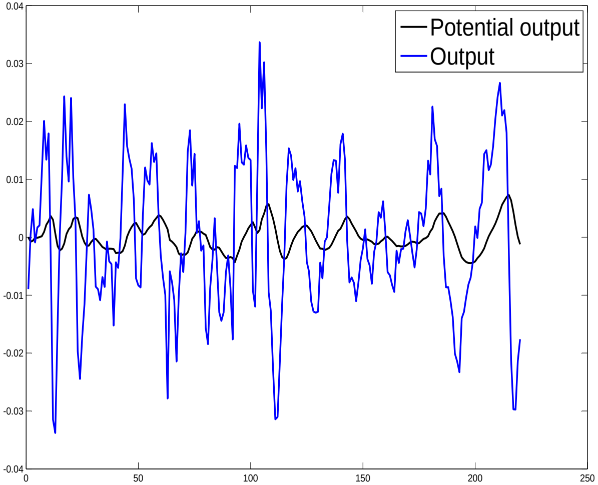
- Here we'll use

$$\begin{aligned} & \{\beta, \kappa, \gamma, \phi, \rho, \sigma_\pi^2, \sigma_y^2, \sigma_u^2\} \\ = \; & \{.99, 0.2, 2, 1.5, 0.98, 1, 1, 0.01\} \end{aligned}$$

  as a benchmark parameterization

Running the filter

Pugging the parameters into the state space model and running the filter gives a time series of $X_{t|t} = \overline{y}_{t|t}$.



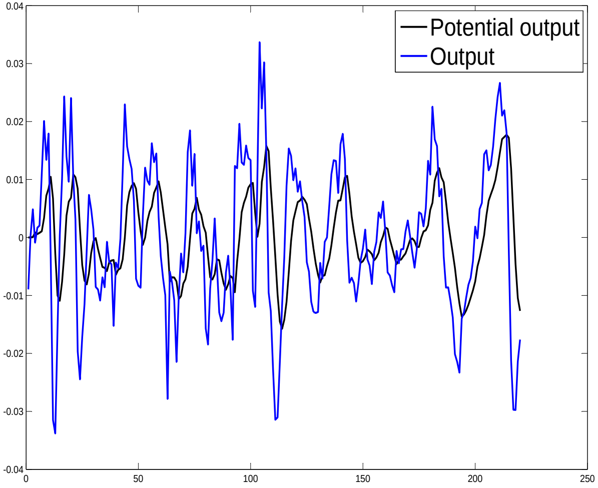
Potential output
Output
0.04
0.03
0.02
0.01
0
-0.01
-0.02
-0.03
-0.04
0
50
100
150
200
250

How does the parameters affect the state estimates?

Smaller variance of “measurement errors”, makes filter interpret more of variance as being due to structural innovations

- $\sigma_\pi^2 = 0.1$
- $\sigma_y^2 = 0.1$

$$\begin{aligned}
\overline{y}_t &= \rho \overline{y}_{t-1} + u_t \\
\pi_t &= \frac{\kappa(1-\rho)}{\Psi} \overline{y}_t + v_t^\pi \\
y_t &= \frac{\kappa\gamma(\rho-\phi)}{\Psi} \overline{y}_t + v_t^y
\end{aligned}$$



Potential output
Output
0.04
0.03
0.02
0.01
0
-0.01
-0.02
-0.03
-0.04
0
50
100
150
200
250

The Kalman gain and the variance of the measurement errors

Smaller m.e. errors imply more weight is put on signal relative to prior

$$X_{t|t} = AX_{t-1|t-1} + K_t \left(Z_t - DX_{t|t-1} \right)$$

$$K_t = P_{t|t-1} D' \left(DP_{t|t-1} D' + \Sigma_{vv} \right)^{-1}$$

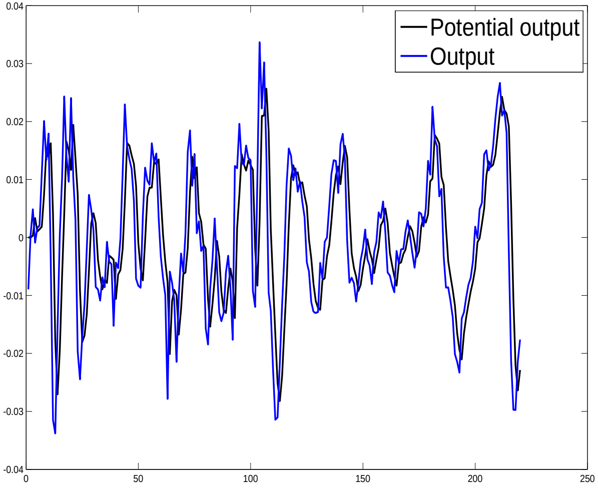
How does the parameters affect the state estimates?

Increasing variance of structural innovations has a similar effect to decreasing the variance of the measurement errors.

- $\sigma_u^2 = 1$

$$
\begin{aligned}
\bar{y}_t &= \rho \bar{y}_{t-1} + u_t \\
\pi_t &= \frac{\kappa (1-\rho)}{\Psi} \bar{y}_t + v_t^{\pi} \\
y_t &= \frac{\kappa \gamma (\rho - \phi)}{\Psi} \bar{y}_t + v_t^{y}
\end{aligned}
$$

One can think of the Kalman filter as splitting innovations in observables ("surprises") into the most likely combination of actual changes in the state and measurement errors.

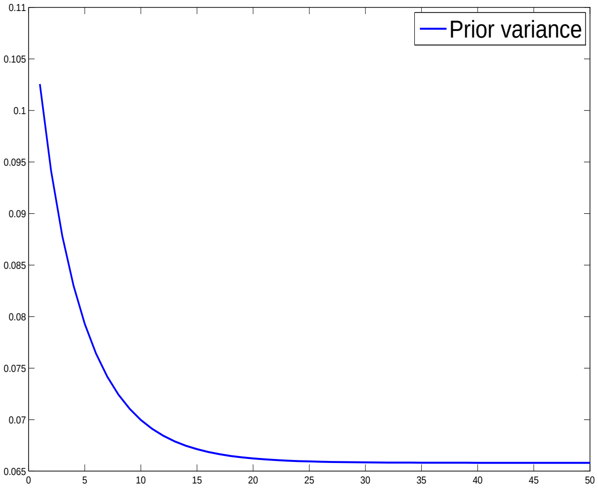


Potential output
Output
0.04
0.03
0.02
0.01
0
-0.01
-0.02
-0.03
-0.04
0
50
100
150
200
250

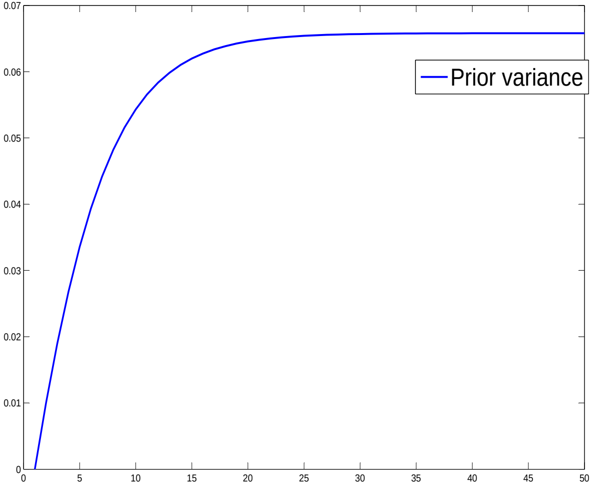
Convergence and the time invariant filter

If the matrices A, C, D and Σ_{vv} are constant, the Kalman filter recursions converge to steady state

Convergence of prior from above



Convergence of prior from below



Convergence of prior variance

Since K_t is function of constant matrices and $P_{t|t-1}$, K_t also converges

$$K_t = P_{t|t-1} D' \left(D P_{t|t-1} D' + \Sigma_{vv} \right)^{-1}$$

The Kalman Smoother

The standard filter gives an optimal real time estimate of the latent state

- Sometimes we are interested in the best estimate given the complete sample, i.e. $X_{t|T}$

$$X_{t|T} = E\left[X_t \mid Z^T, X_{0|0}\right]$$

The *Kalman smoother* can be used to find $X_{t|T}$

The Kalman Smoother: Implementation

Run filter forward, then backward.

$$X_{t|T} = X_{t|t} + J_{t-1}\left(X_{t+1|T} - X_{t+1|t}\right) \tag{1}$$

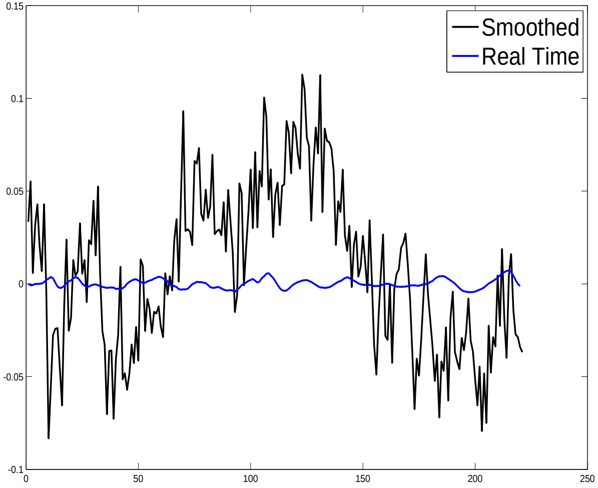
where

$$J_t = P_{t|t} A' P_{t+1|t}^{-1} \tag{2}$$

The covariances of the smoothed state estimation errors can be computed as

$$P_{t|T} = P_{t|t} + J_t\left(P_{t+1|T} - P_{t+1|t}\right) J_t'$$

(for more details, see Hamilton 1994).


Smoothed
Real Time
0.15
0.1
0.05
0
-0.05
-0.1
0
50
100
150
200
250

The Kalman Simulation Smoother

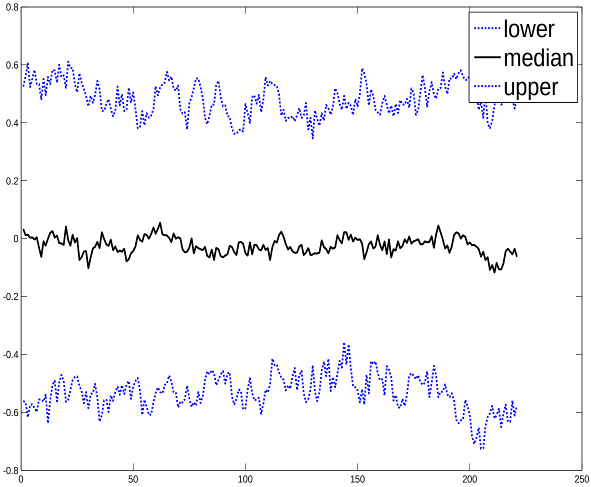
Sometimes we want to know something about the uncertainty of our smoothed estimate.

- One way to illustrate this is to use the *Kalman simulation smoother* to simulate the conditional distribution of X

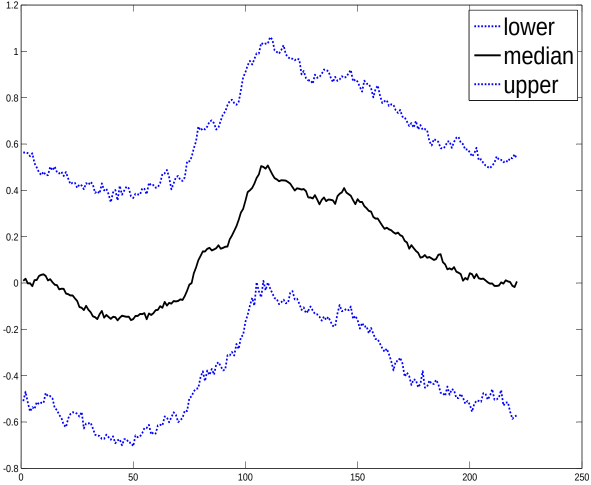
$$
\begin{aligned}
p\left(X_t \mid Z^T, X_{0|0}\right) &= N(X_{t|T}, P_{t|T}) \\
P_{t|T} &= P_{t|t} + J_t \left(P_{t+1|T} - P_{t+1|t}\right) J_t'
\end{aligned}
$$

See Durbin and Koopman (2002) for more details.

Posterior with large measurement error variance



Posterior with small measurement error variance



After the break:

The relationship between principal components and general state space models

The Kalman Smoother: Implementation

Previously, we used principal components to the factors of a system that was in state space form?

- What is the relation ship between the state X_t and the factors F_t
  - Can we find one from the other?
- When is state space models and filtering preferable to principal components?

Principal components factor models

Recall:

$$\underset{(r\times 1)}{F_t} = \Phi F_{t-1} + \mathbf{u}_t^F$$

$$\underset{(N\times 1)}{Y_t} = \underset{(N\times r)}{W}\underset{(r\times 1)}{F_t} + v_t$$

- W contains the eigenvectors of $EY_tY_t' = W\Lambda W'$ and $WW' = I$ so that $F_t = W'Y_t$
- Λ is a diagonal matrix containing the ordered eigenvalues

This looks like a special case of a state space system

Can we find a mapping from X_t to F_t?

Yes:

- When N is large or Σ_{vv} is small the following holds:

$$\begin{aligned} F_t &= W'Y_t \\ &= W'DX_t \end{aligned}$$

If D is of rank n (where $n =$ dimension of X), then $(W'D)^{-1}$ exists so the mapping works in both directions.

How can we find W for a state space model?

$$
\begin{aligned}
EZ_tZ_t' &= D\Sigma_{xx}D' + \Sigma_{vv} \\
&= W\Lambda W'
\end{aligned}
$$

where Σ_{xx} solves

$$\Sigma_{xx} = A\Sigma_{xx}A' + CC'$$

Doing the eigenvector/value decomposition of EZ_tZ_t' thus gives us W so that the factors can be computed as $F_t = W'DX_t$

How about the dynamics of the factors?

We have:

$$\begin{aligned} F_t &= \Phi F_{t-1} + \mathbf{u}_t^F \\ X_t &= A X_{t-1} + C \mathbf{u}_t \end{aligned}$$

To find Φ, use that $F_t = W'DX_t$ and $X_t = (W'D)^{-1} F_t$

$$\begin{aligned} F_t &= W'DAX_{t-1} + W'DC\mathbf{u}_t \\ &= W'DA \left(W'D\right)^{-1} F_{t-1} + W'DC\mathbf{u}_t \end{aligned}$$

to get

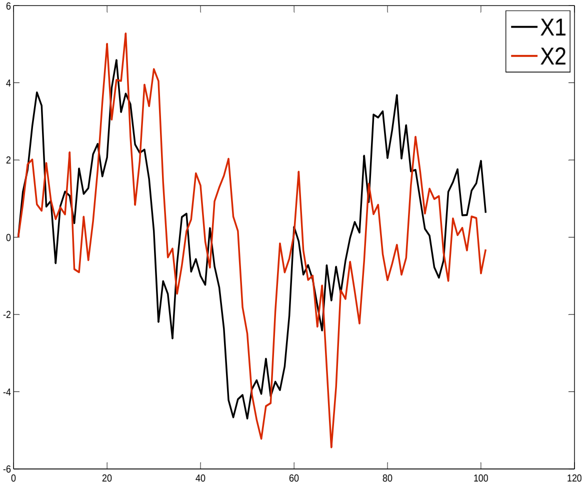
$$\Phi = W'DA \left(W'D\right)^{-1}, E\mathbf{u}_t^F \mathbf{u}_t'^F = W'DC \left(W'DC\right)'$$

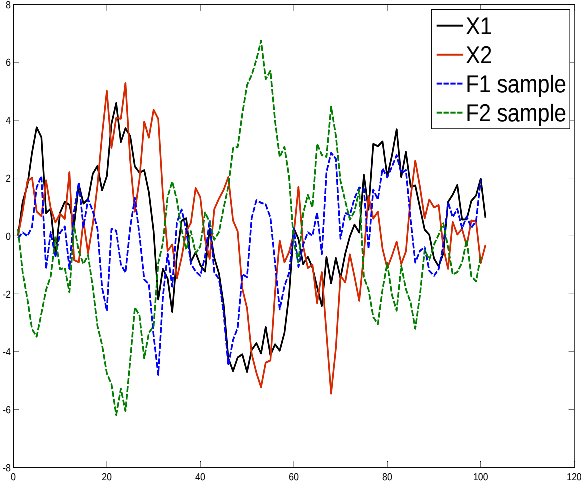
Example

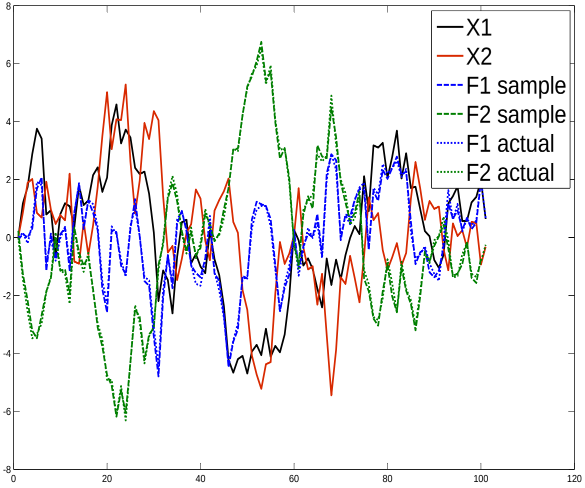
Simulate data using

$$
\begin{aligned}
A &= \begin{bmatrix} 0.9 & 0 \\ 0.2 & 0.7 \end{bmatrix} \\
C &= I, D = I, \Sigma_{vv} = 0.1 \times I
\end{aligned}
$$

with $T = 100$.


6
4
2
0
-2
-4
-6
0
20
40
60
80
100
120
X1
X2


X1
X2
F1 sample
F2 sample
8
6
4
2
0
-2
-4
-6
-8
0
20
40
60
80
100
120



X1
X2
F1 sample
F2 sample
F1 actual
F2 actual
8
6
4
2
0
-2
-4
-6
-8
0
20
40
60
80
100
120

T=100 sample and theoretical Ψ

Using $\Phi = W'DA\,(W'D)^{-1}$

$$\Phi = \begin{bmatrix} 0.69 & 0.008 \\ -0.19 & 0.91 \end{bmatrix}$$

$$\widehat{\Phi} = \begin{bmatrix} 0.71 & 0.025 \\ -0.19 & 0.93 \end{bmatrix}$$

Differences are due to small sample and non-zero measurement errors

Take homes from PCs and State Space models

State space systems that have the same implications for observables are not unique

- "Rotations" of state variables in X_t can give different interpretations
- Different rotations span the same space, so no difference in predictive content

Is this important?

- Depends on the question.
  - In factor models of the term structure, PC imply that the factors will have the level, slope and curvature interpretation.
  - In macro models, usually too few degrees of freedom to do any rotations since number of deep parameters is lower than free parameters in the state space system.

That's it for today.