Factor Models: Kalman Filters

Learning Objectives

1. Understand dynamic factor models using Kalman filters.
2. Estimation of the parameters by maximum likelihood.
3. Applications to
   (a) Ex ante real interest rates
   (b) Stochastic volatility
   (c) Term structure of interest rates

Background Reading

1. Previous lecture notes on factor models in finance.

EViews Computer Files

1. kalman_exante.wf1
2. stochastic_volatility.wf1
3. yields_us.wf1
The discussion so far has concentrated on specifying and estimating factor models based on contemporaneous relationships amongst the observed variables.

In the case of the principal components estimator the aim is to decompose the covariance or correlation matrix of the $N$ observable variables in terms of a set of $K$ latent factors

$$s_1,t, s_2,t, \cdots, s_K,t$$

However, an important feature of many financial time series is that they exhibit dynamic patterns as the following example demonstrates.
### Example (Term Structure of Interest Rates)

The following table gives the autocorrelations for up to 10 lags on the 1-month, 1-year and 5-year U.S. Treasury yields.

<table>
<thead>
<tr>
<th>Autocorr.</th>
<th>Lag 1</th>
<th>Lag 2</th>
<th>Lag 3</th>
<th>Lag 4</th>
<th>Lag 5</th>
<th>Lag 6</th>
<th>Lag 7</th>
<th>Lag 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-month</td>
<td>0.977</td>
<td>0.948</td>
<td>0.921</td>
<td>0.887</td>
<td>0.852</td>
<td>0.819</td>
<td>0.778</td>
<td>0.731</td>
</tr>
<tr>
<td>1-year</td>
<td>0.980</td>
<td>0.950</td>
<td>0.917</td>
<td>0.883</td>
<td>0.849</td>
<td>0.815</td>
<td>0.779</td>
<td>0.739</td>
</tr>
<tr>
<td>5-year</td>
<td>0.936</td>
<td>0.855</td>
<td>0.786</td>
<td>0.727</td>
<td>0.670</td>
<td>0.630</td>
<td>0.600</td>
<td>0.576</td>
</tr>
</tbody>
</table>

Source: yields_us.wf1

The dynamics of the three series are very similar with the autocorrelations slowly decaying at an exponential rate. This suggests that a single factor could potentially capture the autocorrelation in all three yields.
As the previous example suggests that the dynamics of the interest rates can be explained by a common factor it is necessary to expand the factor structure as adopted in the principal components framework and replace the assumption that the factors are independent over time with a more dynamic specification.

In the case of $N$ variables and $K = 1$ factor, a potential specification is

$$y_{i,t} = \alpha_i + \beta_i s_t + u_{i,t}$$

$$s_t = \phi s_{t-1} + \nu_t,$$

where $u_{i,t} \sim N(0, \sigma_i^2)$ and $\nu_t \sim N(0, 1)$ are independent disturbance terms and

$$\{\alpha_1, \alpha_2, \ldots, \alpha_N; \beta_1, \beta_2, \ldots, \beta_N; \sigma_1, \sigma_2, \ldots, \sigma_N; \phi\},$$

are the unknown parameters.

Not only are the contemporaneous relationships captured by the factor $s_t$, but the dynamic relationships are as well.
An important special case is where there is no autocorrelation

$$\phi = 0$$

The factor $s_t$ is now an iid disturbance term given by

$$s_t = v_t$$

which is the specification underlying the principal components framework.

The expansion of the factor model to include a dynamic factor means that an alternative approach to the principal components estimator is needed.

The approach presented here is based on the Kalman filter.
Introduction

**Historical Background: Rudolf Kalman (1930 -)***

Rudy Kálmán was born in Hungary but educated in the U.S. where he spent most of his life. Is credited with inventing the filter commonly known as the Kalman filter, although others also contributed to the theory: often the filter is called the Kalman-Bucy filter.

The Kalman filter is applied in many areas, including econometrics, Bayesian learning and even the Apollo space program!
To understand the Kalman filter a simple model is specified consisting of a single observable variable ($y_t$) and a single latent factor ($s_t$)

$$
\begin{align*}
  y_t &= \beta s_t + u_t \\
  s_t &= \phi s_{t-1} + v_t
\end{align*}
$$

where $u_t \sim N(0, \sigma^2)$ and $v_t \sim N(0, 1)$ are independent disturbances, and $\{\beta, \phi, \sigma^2\}$ are unknown parameters.

This representation of the model is also known as a state-space system with the first equation representing the signal equation (the equation of the observable variable $y_t$) and the second representing the state equation (the equation of the unobservable variable $s_t$).
The Kalman Filter

The Univariate Model

- Define the conditional mean of \( y_t \) based on information at time \( t - 1 \)
  \[
  y_{t|t-1} = E_{t-1}[y_t]
  \]
  with variance
  \[
  V_{t|t-1} = E[(y_t - y_{t|t-1})^2]
  \]
- As \( s_t \) is unknown the aim of the Kalman filter is to estimate the factor \( s_t \) using the available information on the observable variable \( y_t \).
- The best estimator of the factor \( s_t \) based on information at time \( t - 1 \), is the conditional mean
  \[
  s_{t|t-1} = E_{t-1}[s_t]
  \]
  with variance
  \[
  P_{t|t-1} = E[(s_t - s_{t|t-1})^2]
  \]
But when information on $y_t$ becomes available then a better estimator of $s_t$ is given by the updated conditional mean

$$s_{t|t} = E_t [s_t]$$

with variance

$$P_{t|t} = E \left[ (s_t - s_{t|t})^2 \right]$$

This sequence of updating the estimate of $s_t$ as more information on $y_t$ becomes available is an important feature of the Kalman filter.

To understand the recursive nature of the algorithm it is assumed that the parameters

$$\beta, \sigma, \phi$$

are known, or at least represent some starting values. Issues of estimation are discussed below.
For the 1-factor model the Kalman filter equations are summarized as

**Prediction:**

\[ s_{t|t-1} = \phi s_{t-1|t-1} \]
\[ P_{t|t-1} = \phi^2 P_{t-1|t-1} + 1 \]

**Observation:**

\[ y_{t|t-1} = \beta s_{t|t-1} \]
\[ V_{t|t-1} = \beta^2 P_{t|t-1} + \sigma^2 \]

**Updating:**

\[ s_{t|t} = s_{t|t-1} + \frac{\beta P_{t|t-1}}{V_{t|t-1}} (y_t - y_{t|t-1}) \]
\[ P_{t|t} = P_{t|t-1} - \frac{\beta^2 P^2_{t|t-1}}{V_{t|t-1}} \]
At $t = 1$, starting values are needed for the two prediction equations $s_{1|0}, P_{1|0}$.

A typical choice of the mean of the factor is $s_{1|0} = 0$

although other values can be used. A typical choice of the variance of the factor is $P_{1|0} = 1/(1 - \phi^2)$

which is the variance of the unconditional distribution of an AR(1) process.

For given values of the parameters, the filter is computed for $t = 1, 2, \ldots, T$. 

Example (Numerical Example of the Filter)

Suppose that there are $T = 2$ observations on the variable $y_t$ given by $y_t = \{2, 5\}$. Assume that the parameters are $\beta = 0.5$, $\sigma = 0.1$, $\phi = 0.8$, and the initial estimate of the factor is chosen as $s_{1|0} = 0.1$. The first step $(t = 1)$ is

**Prediction:** $s_{1|0} = 0.1$

(initialization) $P_{1|0} = \frac{1}{1 - \phi^2} = \frac{1}{1 - 0.8^2} = 2.7778$

**Observation:** $y_{1|0} = \beta s_{1|0} = 0.5 \times 0.1 = 0.05$

$V_{1|0} = \beta^2 P_{1|0} + \sigma^2 = 0.5^2 \times 2.7778 + 0.1^2 = 0.7045$
Example (Numerical Example of the Filter continued)

**Updating:**

\[ s_{1|1} = s_{1|0} + \frac{\beta P_{1|0}}{V_{1|0}} (y_1 - y_{1|0}) \]
\[ s_{1|1} = 0.1 + \frac{0.5 \times 2.7778}{0.7045} \times (2 - 0.05) = 3.9444 \]

\[ P_{1|1} = P_{1|0} - \frac{\beta^2 P_{1|0}^2}{V_{1|0}} \]
\[ P_{1|1} = 2.7778 - \frac{0.5^2 \times 2.7778^2}{0.7045} = 0.0396 \]

Intuitively, the initial estimate of 0.1 for the factor at \( t = 1 \), results in an underestimate of the observed variable, 0.05 < 2. By updating the estimate of the factor to 3.9444 this yields a better estimate of \( y_1 \).
The Kalman Filter
The Univariate Model

Example (Numerical Example of the Filter continued)

The second step ($t = 2$) is

Prediction:  
\[
\begin{align*}
    s_{2|1} &= \phi s_{1|1} \\
    s_{2|1} &= 0.8 \times 3.9444 = 3.1555 \\
    P_{2|1} &= \phi^2 P_{1|1} + 1 \\
    P_{2|1} &= 0.8^2 \times 0.0396 + 1 = 1.0253
\end{align*}
\]

Observation:  
\[
\begin{align*}
    y_{2|1} &= \beta s_{2|1} \\
    y_{2|1} &= 0.5 \times 3.1555 = 1.5778 \\
    V_{2|1} &= \beta^2 P_{2|1} + \sigma^2 \\
    V_{2|1} &= 0.5^2 \times 1.0253 + 0.1^2 = 0.2663
\end{align*}
\]
Example (Numerical Example of the Filter continued)

The second step \( (t = 2) \) is

\[
\text{Updating: } s_{2|2} = s_{2|1} + \frac{\beta P_{2|1}}{V_{2|1}} (y_2 - y_{2|1})
\]

\[
s_{2|2} = 3.1555 + \frac{0.5 \times 1.0253}{0.2663} \times (5 - 1.5778) = 9.7435
\]

\[
P_{2|2} = P_{2|1} - \frac{\beta^2 P_{2|1}^2}{V_{2|1}}
\]

\[
P_{2|2} = 1.0253 - \frac{0.5^2 \times 1.0253^2}{0.2663} = 0.03840
\]
Consider a model where $N = 3$ variables and $K = 2$ factors

\[
y_{1,t} = \alpha_1 + \beta_{1,1}s_{1,t} + \beta_{1,2}s_{2,t} + u_{1,t} \\
y_{2,t} = \alpha_2 + \beta_{2,1}s_{1,t} + \beta_{2,2}s_{2,t} + u_{2,t} \\
y_{3,t} = \alpha_3 + \beta_{3,1}s_{1,t} + \beta_{3,2}s_{2,t} + u_{3,t}
\]

\[
s_{1,t} = \phi_{1,1}s_{1,t-1} + \nu_{1,t} \\
s_{2,t} = \phi_{2,2}s_{2,t-1} + \nu_{2,t}
\]

or in matrix notation

\[
\begin{bmatrix}
y_{1,t} \\
y_{2,t} \\
y_{3,t}
\end{bmatrix} = \begin{bmatrix}
\alpha_1 \\
\alpha_2 \\
\alpha_3
\end{bmatrix} + \begin{bmatrix}
\beta_{1,1} & \beta_{1,2} \\
\beta_{2,1} & \beta_{2,2} \\
\beta_{3,1} & \beta_{3,2}
\end{bmatrix} \begin{bmatrix}
s_{1,t} \\
s_{2,t}
\end{bmatrix} + \begin{bmatrix}
u_{1,t} \\
u_{2,t} \\
u_{3,t}
\end{bmatrix}
\]

\[
\begin{bmatrix}
s_{1,t} \\
s_{2,t}
\end{bmatrix} = \begin{bmatrix}
\phi_{1,1} & 0 \\
0 & \phi_{2,2}
\end{bmatrix} \begin{bmatrix}
s_{1,t-1} \\
s_{2,t-1}
\end{bmatrix} + \begin{bmatrix}
\nu_{1,t} \\
\nu_{2,t}
\end{bmatrix}
\]
For an extension the previous example, consider the case of $N$ variables $\{y_{1,t}, y_{2,t}, \ldots, y_{N,t}\}$ and $K$ factors $\{s_{1,t}, s_{2,t}, \ldots, s_{K,t}\}$. The multivariate version of the state-space system is

\[
y_t = A + Bs_t + u_t
\]
\[
s_t = \Phi s_{t-1} + v_t
\]

where the disturbances are distributed as

\[
u_t \sim N(0, R)
\]
\[
v_t \sim N(0, Q)
\]

where $E[u_t u'_t] = R$ and $E[v_t v'_t] = Q$ are respectively the covariances of $u_t$ and $v_t$.

The dimensions of the parameter matrices are as follows: $A$ is $(N \times 1)$, $B$ is $(N \times K)$, $\Phi$ is $(K \times K)$, $R$ is $(N \times N)$ and $Q$ is $(K \times K)$. 


The recursions of the multivariate Kalman filter are

**Prediction:**
\[ s_{t|t-1} = \Phi s_{t-1|t-1} \]
\[ P_{t|t-1} = \Phi P_{t-1|t-1} \Phi' + Q \]

**Observation:**
\[ y_{t|t-1} = Bs_{t|t-1} \]
\[ V_{t|t-1} = BP_{t|t-1} B' + R \]

**Updating:**
\[ s_t = s_{t|t-1} + P_{t|t-1} B' V_{t|t-1}^{-1} (y_t - y_{t|t-1}) \]
\[ P_t = P_{t|t-1} - P_{t|t-1} B' V_{t|t-1}^{-1} B P_{t|t-1} \]

The formulae for the multivariate version of the Kalman filter contain the univariate formulae with \( N = K = 1 \).
To start the recursion two cases are considered.

1. Stationary Latent Factors
   The initial values \( s_{1|0} \) and \( P_{1|0} \) for the multivariate \( K \) factor model are given by
   \[
   s_{1|0} = 0 \\
   \text{vec}(P_{1|0}) = (I_{K	imes K} - (\Phi \otimes \Phi))^{-1}\text{vec}(Q)
   \]

2. Nonstationary Latent Factors
   In the case the starting values for the variance would be undefined if the previous approach is adopted. To circumvent this problem, starting values are chosen as
   \[
   s_{1|0} = \psi \\
   P_{1|0} = \omega \text{vec}(Q)
   \]
   where \( \psi \) represents the best guess of starting value for the conditional mean and \( \omega \) is a positive constant whereby larger values of \( \omega \) correspond to the distribution of \( s_{1|0} \) being more diffuse.
The state-space model is under-identified unless some restrictions are imposed.

The difficulty is seen by noting that the volatility in the factor is controlled by $Q$, but the impact of the factor on $y_t$ is given by $B$.

There is an infinite number of combinations of $Q$ and $B$ that will be consistent with the volatility of $y_t$ ie in the case of $N = K = 1$, then

$$\text{var} (y_t) = \beta^2 \text{var} (s_t) + \text{var} (u_t)$$

Thus it is necessary to fix one of these quantities.

- A common approach is to set

$$Q = I$$

- Another approach is to place restrictions on $B$ and allow $Q$ to be estimated.
The discussion so far has concentrated on extracting the factor $s_t$, assuming given values for the population parameters

\[ \theta = \{A, B, \Phi, R, Q\} \]

In general, however, it is necessary to estimate these parameters.

If the factors are known, then the parameters are estimated by simply regressing $y_t$ on $s_t$ and regressing $s_t$ on $s_{t-1}$. But as $s_t$ is unobservable (latent), an alternative estimation strategy is needed.

The natural estimator of the parameters is the maximum likelihood estimator which constructs the log-likelihood function based on

\[ y_t \sim N(y_{t|t-1}, V_{t|t-1}) \]

As the likelihood is a nonlinear function of the parameters an iterative algorithm is required to obtain the maximum likelihood estimates.
For a sample of $t = 1, 2, \cdots, T$ observations on $y_t$, the log-likelihood function for the $t^{th}$ observation using the multivariate normal distribution is given by

$$\log L_t = -\frac{N}{2} \log(2\pi) - \frac{1}{2} \log |V_{t|t-1}| - \frac{1}{2} (y_t - y_{t|t-1})' V_{t|t-1}^{-1} (y_t - y_{t|t-1})$$

For the entire sample, the log-likelihood function is

$$\log L = \frac{1}{T} \sum_{t=1}^{T} \log L_t$$

This expression is a nonlinear function of the parameters

$$\theta = \{A, B, \Phi, R, Q\}$$

via $y_{t|t-1}$ and $V_{t|t-1}$ from the Kalman filter.
Maximum Likelihood Estimator

Using EViews

- Consider estimating a one-factor model of the spread between the one-year yield and the one-month yield

\[
YIELD\_Y1_t - YIELD\_M1_t = \alpha + \beta s_t + u_t, \quad u_t \sim N(0, \sigma^2)
\]

\[
s_t = \phi s_{t-1} + \nu_t, \quad \nu_t \sim N(0, 1)
\]

with starting values \( \{\alpha(0) = 0.1, \beta(0) = 0.1, \sigma^2(0) = 0.1, \phi(0) = 0.9\} \).

- The EViews commands are:

\[
\text{Object} / \text{New Object...} / \text{SSpace} / \text{OK}
\]

In the window type in the following commands

\[
\text{@signal yield\_y1-yield\_m1 = c(1) + c(2)*s + [var = c(3)]}
\]

\[
\text{@state s = c(4)*s(-1) + [var = 1]}
\]

\[
\text{@param c(1) 0.1 c(2) 0.1 c(3) 0.1 c(4) 0.9}
\]

Then click

\[
\text{Estimate} / \text{OK}
\]
Factor Extraction

- Once the algorithm has converged estimates of the latent factor $s_t$ at each point in time are available.
- In fact, three estimates can be calculated depending on the form of the conditioning information set used:

  - **One-step-ahead**: $s_{t|t-1} = E_{t-1}[s_t]$
  - **Filtered**: $s_{t|t} = E_t[s_t]$
  - **Smoothed**: $s_{t|T} = E_T[s_t]$

- The first two estimates, $s_{t|t-1}$ and $s_{t|t}$, are a by-product of the Kalman filter algorithm which are automatically available once the algorithm has converged.
- The third estimator $s_{t|T}$, is effectively obtained by running the Kalman filter algorithm in the reverse direction (from $T$ to $t - 1$) once the maximum likelihood estimates are obtained.
Factor Extraction

Using EViews

1. The one-step-ahead estimate of the factor $s_{t|t-1} = E_{t-1} [s_t]$

   View / State Views / Graph State Series...
   / One-step-ahead: Predicted States / OK

2. The filtered estimate of the factor $s_{t|t} = E_t [s_t]$

   View / State Views / Graph State Series...
   / Filtered: State Estimates / OK

3. The smoothed estimate of the factor $s_{t|T} = E_T [s_t]$

   View / State Views / Graph State Series...
   / Smoothed: State Estimates / OK
There exist two broad types of real interest rates:
(i) Ex post real interest rates (observed).
(ii) Ex ante real interest rates (unobserved).

The ex post real interest rate is observed (as given in the following Figure which gives the U.S. ex post 1-month real interest rate), but the ex ante real interest rate is not.

Source: kalman_exante.wf1
But it is the ex ante real interest rate that is important in finance and economics as it provides a measure of the real return on an asset between the present and the future.

How can the ex ante interest rate be measured?

There are two strategies:

(i) Proxy
Use the ex post real interest rate as a proxy for the ex ante interest rate.

(ii) Latent Factor
Treat the ex ante real interest rate as unknown using a latent factor model.
Estimating the Ex Ante Real Interest Rate

Formally the ex ante real interest rate is defined as

\[ r_t^e = i_t - \pi_t^e \]

where \( i_t \) is the nominal interest rate and \( \pi_t^e \) is the expected inflation rate defined as

\[ \pi_t^e = \log p_{t+1} - \log p_t \]

Whilst \( i_t \) is observed, \( \pi_t^e \) is not.

So it is the expected inflation rate that makes the ex ante real interest rate unobservable.
Estimating the Ex Ante Real Interest Rate

- Consider the ex post real interest rate

\[ r_t = i_t - \pi_t \]

which is observed where \( \pi_t = \log p_t - \log p_{t-1} \) is the actual inflation rate. Expanding this expression to allow for expected inflation, \( \pi^e_t \), gives

\[ r_t = i_t - \pi^e_t + \pi^e_t - \pi_t = i_t - \pi^e_t + u_t \]

- Defining \( s_t = i_t - \pi^e_t - \alpha \) as the ex ante real interest rate (adjusted by \( \alpha \)) and \( u_t = \pi^e_t - \pi_t \) as the inflation expectations error, this expression is written as a latent factor model as

\[ r_t = \alpha + s_t + u_t, \quad u_t \sim \mathcal{N}(0, \sigma_u^2) \]

- The key advantage of this formulation of the model is that it avoids the measurement error from using realized inflation and not expected inflation.
To estimate the ex ante real interest rate, monthly data starting in January 1971 and ending in December 2009 on the following U.S. series are used:

\[ \text{EURO}_1 \text{MTH} : \text{1-month Eurodollar rate, } \% \text{, p.a.} \]

\[ \text{CPI} : \text{Consumer price index} \]

The annualized percentage inflation rate is computed as

\[ \text{INF} = 1200 \times \text{DLOG} (\text{CPI}) \]

and the ex post real interest rate is computed as

\[ R = \text{EURO}_1 \text{MTH} − \text{INF} \]

This is the ex post real interest rate given in the previous Figure.
Estimating the Ex Ante Real Interest Rate

- Some summary statistics are given in the Figure below.

- Here the average real ex post interest rate is 2.174% p.a. over the sample period.
Estimating the Ex Ante Real Interest Rate

- The autocorrelation of the real ex post interest rate if given in the following figure.

![Autocorrelation Table]

Source: kalman_exante.wf1

- The correlogram shows strong evidence of first order autocorrelation. This result is important as identification of the parameters of the model require that there is significant autocorrelation.
The factor model of the ex ante real interest rate is specified as

\[ r_t = \alpha + s_t + u_t, \quad u_t \sim N \left( 0, \sigma_u^2 \right) \quad \text{[Signal equation]} \]

\[ s_t = \phi s_{t-1} + v_t, \quad v_t \sim N \left( 0, \sigma_v^2 \right) \quad \text{[State equation]} \]

where the unknown parameters are \( \theta = \{ \alpha, \phi, \sigma_u^2, \sigma_v^2 \} \).

The starting values for the parameters are chosen as follows:

- \( \alpha \) is based on the sample mean of \( r_t \), equal to 2.174.
- \( \phi \) is based on the first autocorrelation coefficient of \( r_t \), equal to 0.551.
- \( \sigma_u \) and \( \sigma_v \) are both set equal to half of the standard deviation of \( r_t \), equal to 4.333/2.
The EViews window to estimate the model is given below.

```
@signal r = c(1) + s + [var = c(3)^2]
@state s = c(2)*s(-1) + [var = c(4)^2]
@param c(1) 2.174 c(2) 0.551 c(3) 2.167 c(4) 2.167
Source: kalman_exante.wf1
```

where \( C (1) \) corresponds to \( \alpha \), \( C (2) \) corresponds to \( \phi \), \( C (3) \) corresponds to \( \sigma_u \), \( C (4) \) corresponds to \( \sigma_v \).

Note that it is the standard deviations \( \sigma_u \) and \( \sigma_v \) that are being estimated and not the variance. This choice of parameterization has the advantage that the variance is guaranteed to be positive. If either of the estimates of \( \sigma_u \) and \( \sigma_v \) happen to be negative, it is appropriate to just change the sign and report a positive estimate.
Estimating the Ex Ante Real Interest Rate

- The parameter estimates are contained in the following window.

Source: kalman_exante.wf1

- The estimated model is

\[
\begin{align*}
 r_t &= 2.174 + \hat{s}_t + \hat{u}_t \\
 \hat{s}_t &= 0.583 \hat{s}_{t-1} + \hat{v}_t
\end{align*}
\]

where \( \hat{\sigma}_u = 1.037 \), \( \hat{\sigma}_v = 3.411 \).
As it is the ex ante estimate of the real interest rate that is required, the one-step ahead factor $s_{t|t-1}$, is the appropriate quantity as it provides an estimate of the interest rate in the future at time $t$, based on information at time $t - 1$, without using current or future information.

The Eviews commands to extract the estimate of the one-step-ahead estimate of the factor $s_{t|t-1} = E_{t-1}[s_t]$, are

```
Proc / Make State Series...
```

Choose **One-step-ahead: Predicted states**, then for **Series names** choose

```
S_HAT
```
As the factor is defined as

\[ s_t = i_t - \pi_t^e - \alpha \]

the ex ante real interest rate is given by rearranging this expression as

\[ r_t^e = i_t - \pi_t^e = s_t + \alpha \]

Given that \( s_{t|t-1} \) is the appropriate conditional mean estimate of the factor, from the definition of the factor an estimate of the ex ante real interest rate is given by

\[ \hat{r}_t^e = \hat{s}_{t|t-1} + \hat{\alpha} \]

This quantity is computed using Genr as

\[ \text{RE\_HAT} = \text{S\_HAT} + 2.174 \]
Estimating the Ex Ante Real Interest Rate

- The estimate of the ex ante real interest rate ($\hat{r}_t^e$) and the ex post real interest rate ($r_t$) are compared in the following Figure.

![Graph showing comparison of $\hat{r}_t^e$ and $r_t$ over time]

Source: kalman_exante.wf1

- The estimate of the ex ante real interest rate $\hat{r}_t^e$ follows $r_t$ closely but exhibits less volatility.
Alternatively, as the ex ante real interest rate is a function of the expected inflation rate, then the latter can be estimated as

\[ \hat{\pi}_t^e = i_t - \hat{r}_t^e \]

Using the Genr command, \( \hat{\pi}_t^e \) is computed and plotted in the following Figure together with the actual inflation rate \( \pi_t \).

Source: kalman_exante.wf1
• Volatility is an important input into financial decision-making as it represents the risk of an asset.
• Consider the case where the asset is the UK/US exchange rate. The (demeaned) return on the UK/US exchange rate \( r_t \) is given in the following Figure from January 2nd 1979 to February 13th 2014.

Source: stochastic_volatility.wf1
A Stochastic Volatility Model of the Exchange Rate

- The aim is to extract a measure of the volatility of the exchange rate.
- One approach is to assume constant volatility. The following Figure yields an estimate of 0.005238.

![Graph showing distribution of exchange rate volatility](stochastic_volatility.wf1)

Source: stochastic_volatility.wf1

- Another approach is to assume time-varying volatility by specifying a GARCH model where the volatility is assumed to be a function of lagged (squared) shocks.
Another approach is the stochastic volatility model given by

\[ r_t = \sigma_t w_t \]  \hspace{1cm} \text{[Mean equation]} \]
\[ \log(\sigma_t^2) = \alpha + \phi \log(\sigma_{t-1}^2) + \nu_t \]  \hspace{1cm} \text{[Variance equation]} \]

where \( r_t \) is the (demeaned) exchange rate return, \( \sigma_t \) represents the exchange rate volatility, and \( w_t \) and \( \nu_t \) are disturbance terms with the properties \( w_t \sim N(0, 1) \) and \( \nu_t \sim N(0, \sigma_v^2) \).

An important feature of this model is the additional stochastic term given by \( \nu_t \), in the variance equation. For this reason the model is called the stochastic volatility model.

Estimating the stochastic volatility model is in general difficult arising from the presence of the additional disturbance term \( \nu_t \) as that now makes the volatility \( \sigma_t^2 \) stochastic as well.

One solution is to express the model as a latent factor model and use the Kalman filter to estimate the model by maximum likelihood methods.
The strategy consists of squaring both sides of the mean equation as

\[ r_t^2 = \sigma_t^2 w_t^2 \]

Now taking natural logarithms gives

\[ \log r_t^2 = \log (\sigma_t^2) + \log (w_t^2) \]

Redefine the variables as

\[
\begin{align*}
y_t &= \log r_t^2 \\
s_t &= \log (\sigma_t^2) \\
u_t &= \log (w_t^2) + 1.27
\end{align*}
\]

where the term 1.27 in the equation for \( u_t \) appears as it can be shown that \( E [\log (w_t^2)] = -1.27 \), so \( E [u_t] = 0 \).

Also, it can be shown that the variance of \( \log (w_t^2) \) and hence \( u_t \), is

\[ E [u_t^2] = \frac{\pi^2}{2} = 4.9348 \]
The stochastic volatility model is rewritten as a latent factor model as

\[
\begin{align*}
    y_t &= -1.27 + s_t + u_t \quad \text{[Mean equation]} \\
    s_t &= \alpha + \phi s_{t-1} + v_t \quad \text{[Variance equation]}
\end{align*}
\]

where \( y_t = \log r_t^2 \), the natural logarithm of the squared exchange rate.

The variable \( y_t \) is constructed using Genr in EViews.

To generate some starting values the following AR(1) model is estimated

\[
y_t = \beta_1 + \beta_2 y_{t-1} + w_t
\]

where \( w_t \sim N \left( 0, \sigma_w^2 \right) \).
The parameter estimates are given in the following window.

```
Dependent Variable: Y
Method: Least Squares
Date: 05/04/14  Time: 13:51
Sample (adjusted): 1/1/1979 2/13/2014
Included observations: 12825 after adjustments

                      Variable  Coefficient  Std. Error  t-Statistic  Prob.
                      C         -10.87204  0.134430    -80.87485  0.0000
                      Y(-1)       0.273568  0.008494    32.20855  0.0000

                      R-squared          0.074846  Mean dependent var -14.96627
                      Adjusted R-squared 0.074774  S.D. dependent var  5.149445
                      S.E. of regression  4.953184    Akaike info criterion  6.038094
                      Sum squared resid   314599.9    Schwarz criterion    6.039258
                      Log likelihood     -38717.28    Hannan-Quinn criter.  6.038483
                      F-statistic         1037.391  Durbin-Watson stat    1.783787
                      Prob(F-statistic)   0.000000

Source: stochastic_volatility.wf1
```
A Stochastic Volatility Model of the Exchange Rate

- The EViews window to estimate the model is given below.

```
View Proc Object Print Name Freeze Spec Estimate Stats Forecast
@signal y = -1.27 + s + [var=4.9348]
@state s = c(1) + c(2)*s(-1) + [var=c(3)*c2]
param c(1) -10.8720 c(2) 0.2736 c(3) 4.9532
```

Source: stochastic_volatility.wf1

where

- $C(1)$ corresponds to $\alpha$ with the starting value based on $\hat{\beta}_1 = -10.872$
- $C(2)$ corresponds to $\phi$ with the starting value $\hat{\beta}_2 = 0.2736$
- $C(3)$ corresponds to $\sigma_v$ with starting value based on $\hat{\sigma}_w = 4.9532$
The parameter estimates are contained in the following window.

The estimated model is

\[ y_t = -1.27 + \hat{s}_t + \hat{u}_t \]  
\[ \hat{s}_t = -9.8858 + 0.2779\hat{s}_{t-1} + \nu_t \]

where \( \hat{\sigma}_\nu = 4.4759 \).
As $s_t = \log(\sigma_t^2)$, an estimate of the volatility is

$$\hat{\sigma}_t = \exp\left(\frac{s_t}{2}\right)$$

If the strategy is to derive an historical estimate of the volatility the best estimates of the factor at each point in time is based on all of the sample information, namely $\hat{s}_{t|T}$, which is the smoothed estimate. Hence the volatility estimate is based on

$$\hat{\sigma}_t = \exp\left(\frac{\hat{s}_{t|T}}{2}\right)$$
A Stochastic Volatility Model of the Exchange Rate

- The volatility estimate is given in the following figure.

Source: stochastic_volatility.wf1

- The increase in volatility during times of financial crises is clear where the estimates of volatility reach 0.04.
A Stochastic Volatility Model of the Exchange Rate

- Descriptive statistics on the volatility series are given below.

An estimate of the mean of the volatility series is 0.0038 which is a little smaller than the constant volatility estimate of 0.005238.
Factor models are widely used in finance to model the term structure of interest rates. An important example is Cox, Ingersoll and Ross (1985) who derive a 1-factor model of the term structure of interest rates where the unobserved factor is the instantaneous interest rate.

Consider the following one-factor model of the term structure of interest rates

\[ r_{i,t} = \alpha_i + \beta_is_t + u_{i,t}, \quad i = 1, 2, \ldots, 9 \]
\[ s_t = \phi s_{t-1} + \nu_t \]
\[ u_{i,t} \sim N (0, \sigma_i^2), \quad \nu_t \sim N (0, 1) \]

There are 28 parameters. The starting parameters are chosen as

\[ \{ \alpha_i, \beta_i, \sigma_i^2 \} = 0.1 \]
\[ \phi = 0.9 \]
A Dynamic One-Factor Model of the Term Structure

- The EViews window to estimate the model is given below.

```
@signal yield_m1 = c(1) + c(10)*s + [var = c(19)]
@signal yield_m3 = c(2) + c(11)*s + [var = c(20)]
@signal yield_m6 = c(3) + c(12)*s + [var = c(21)]

@signal yield_y1 = c(4) + c(13)*s + [var = c(22)]
@signal yield_y2 = c(5) + c(14)*s + [var = c(23)]
@signal yield_y3 = c(6) + c(15)*s + [var = c(24)]

@signal yield_y5 = c(7) + c(16)*s + [var = c(25)]
@signal yield_y7 = c(8) + c(17)*s + [var = c(26)]
@signal yield_y10 = c(9) + c(18)*s + [var = c(27)]

@state s = c(28)*s(-1) + [var = 1]
@param c(1) 0.1 0.1 c(2) 0.1 c(3) 0.1 c(4) 0.1 c(5) 0.1 c(6) 0.1 c(7) 0.1 c(8) 0.1 c(9) 0.1 c(10) 0.1 c(11) 0.1 c(12) 0.1 c(13) 0.1 c(14) 0.1 c(15) 0.1 c(16) 0.1 c(17) 0.1 c(18) 0.1 c(19) 0.1 c(20) 0.1 c(21) 0.1 c(22) 0.1 c(23) 0.1 c(24) 0.1 c(25) 0.1 c(26) 0.1 c(27) 0.1 c(28) 0.9
```

Source: yields_us.wf1
The parameter estimates are contained in the following window.

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std. Error</th>
<th>z-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C(1)</td>
<td>-2.363971</td>
<td>13.06975</td>
<td>-0.180459</td>
</tr>
<tr>
<td>C(2)</td>
<td>-2.358248</td>
<td>13.27509</td>
<td>-0.177645</td>
</tr>
<tr>
<td>C(3)</td>
<td>-2.254918</td>
<td>13.42593</td>
<td>-0.167952</td>
</tr>
<tr>
<td>C(4)</td>
<td>-1.857111</td>
<td>12.60488</td>
<td>-0.147333</td>
</tr>
<tr>
<td>C(5)</td>
<td>-1.986897</td>
<td>10.96083</td>
<td>-0.088214</td>
</tr>
<tr>
<td>C(6)</td>
<td>-0.180924</td>
<td>9.449453</td>
<td>-0.019146</td>
</tr>
<tr>
<td>C(7)</td>
<td>1.181755</td>
<td>8.332502</td>
<td>0.141825</td>
</tr>
<tr>
<td>C(8)</td>
<td>2.150556</td>
<td>5.076758</td>
<td>0.423441</td>
</tr>
<tr>
<td>C(9)</td>
<td>2.933001</td>
<td>4.382718</td>
<td>0.669220</td>
</tr>
<tr>
<td>C(10)</td>
<td>0.168607</td>
<td>0.015277</td>
<td>11.03670</td>
</tr>
<tr>
<td>C(11)</td>
<td>0.171209</td>
<td>0.010041</td>
<td>17.05155</td>
</tr>
<tr>
<td>C(12)</td>
<td>0.172858</td>
<td>0.009675</td>
<td>17.86644</td>
</tr>
<tr>
<td>C(13)</td>
<td>0.162808</td>
<td>0.016945</td>
<td>9.007756</td>
</tr>
<tr>
<td>C(14)</td>
<td>0.139998</td>
<td>0.056018</td>
<td>2.499155</td>
</tr>
<tr>
<td>C(15)</td>
<td>0.120449</td>
<td>0.006705</td>
<td>1.492448</td>
</tr>
<tr>
<td>C(16)</td>
<td>0.087738</td>
<td>0.157020</td>
<td>0.558789</td>
</tr>
<tr>
<td>C(17)</td>
<td>0.064572</td>
<td>0.166674</td>
<td>0.389754</td>
</tr>
<tr>
<td>C(18)</td>
<td>0.046940</td>
<td>0.078354</td>
<td>0.599074</td>
</tr>
</tbody>
</table>

(continued on the next slide)
### A Dynamic One-Factor Model of the Term Structure

<table>
<thead>
<tr>
<th>C(19)</th>
<th>0.039699</th>
<th>0.014213</th>
<th>2.793184</th>
<th>0.0052</th>
</tr>
</thead>
<tbody>
<tr>
<td>C(20)</td>
<td>0.013433</td>
<td>0.006783</td>
<td>1.980373</td>
<td>0.0477</td>
</tr>
<tr>
<td>C(21)</td>
<td>0.002336</td>
<td>0.002098</td>
<td>1.113649</td>
<td>0.2654</td>
</tr>
<tr>
<td>C(22)</td>
<td>0.012359</td>
<td>0.007456</td>
<td>1.657509</td>
<td>0.0974</td>
</tr>
<tr>
<td>C(23)</td>
<td>0.089894</td>
<td>0.113655</td>
<td>0.790937</td>
<td>0.4290</td>
</tr>
<tr>
<td>C(24)</td>
<td>0.149941</td>
<td>0.267096</td>
<td>0.561376</td>
<td>0.5745</td>
</tr>
<tr>
<td>C(25)</td>
<td>0.210447</td>
<td>0.437776</td>
<td>0.480718</td>
<td>0.6307</td>
</tr>
<tr>
<td>C(26)</td>
<td>0.231194</td>
<td>0.418256</td>
<td>0.552756</td>
<td>0.5604</td>
</tr>
<tr>
<td>C(27)</td>
<td>0.204069</td>
<td>0.262756</td>
<td>0.776648</td>
<td>0.4374</td>
</tr>
<tr>
<td>C(28)</td>
<td>0.999234</td>
<td>0.007354</td>
<td>135.8760</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Final State</th>
<th>Root MSE</th>
<th>z-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>13.95279</td>
<td>1.025722</td>
<td>13.60200</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

- Log likelihood: -122.7663
- Parameters: 28
- Diffuse priors: 0

Source: yields_us.wf1

---

Jun YU  ()  ECON671 Factor Models: Kalman Filters  March 2, 2015  55 / 68
The log-likelihood value is

\[ \ln L(\hat{\theta}) = -121.2888 \]

The estimated loadings (\( \beta \)), given by parameters 10 to 18, show that the latent factor has its greatest impact on the shorter maturities (less than one year) which progressively diminishes in importance across the maturity spectrum.

The estimates of the idiosyncratic parameter (\( \sigma^2 \)), given by parameters 19 to 27, are smallest for the 6-month yield suggesting that this yield follows the factor more closely than the other yields.

As the intercept estimates (\( \alpha \)), given by parameters 1 to 9, increase over the maturity spectrum, this suggests an upward yield curve on average.

The parameter estimate of \( \phi \) is 0.999, suggesting that the latent factor is nonstationary.
The one-step ahead estimates of the latent factor $s_{t|t-1} = E_{t-1}[s_t]$, are given in the following Figure.

Source: yields_us.wf1
The confidence interval for the initial estimate of the factor is very wide representing a lack of information at this point in time. The confidence interval quickly narrows showing that the estimates for later points in time are more precise.

The factor is relatively flat in the first part of the period, then rises reaching a peak around by the end of 2006. As the loadings of the factor are relatively larger on the smaller maturities than the longer maturities, this increase in the factor is associated with a narrowing of the spreads.

From about mid-2007 the factor falls resulting in a widening of spreads, which eventually stabilize from 2009 onwards.
The prediction properties of the model are obtained by computing

\[ y_{i,t|t-1} = \hat{\alpha}_i + \hat{\beta}_i s_{t|t-1} \]

The EViews commands are

**View / Actual,Predicted,Residual Graph / OK**

This figure further highlights how the estimated factor follows the shorter maturities very closely, while the longer maturities tend to exhibit additional dynamics suggesting the need for a second factor.
A Dynamic One-Factor Model of the Term Structure

Source: yields_us.wf1
The state-space model represents a flexible framework which can easily accommodate a number of extensions.

Two important extensions are:

1. Dynamics
   Have focussed on a AR(1) representations of $s_t$ with the idiosyncratic disturbance $u_t$ being white noise. These restrictions can e relaxed.

2. Exogenous and Predetermined Variables
   Can allow for exogenous and predetermined variables in the signal and state equations.
An AR(2) Model of $s_t$

- Suppose that the latent factor is an AR(2) process
  \[ s_t = \phi_1 s_{t-1} + \phi_2 s_{t-2} + \nu_t \]

- This equation can be written as a vector AR(1) model
  \[
  \begin{bmatrix}
  s_t \\
  s_{t-1}
  \end{bmatrix} =
  \begin{bmatrix}
  \phi_1 & \phi_2 \\
  1 & 0
  \end{bmatrix}
  \begin{bmatrix}
  s_{t-1} \\
  s_{t-2}
  \end{bmatrix} +
  \begin{bmatrix}
  \nu_t \\
  0
  \end{bmatrix}
  \]

- The Kalman filter proceeds as before except now there are two factors, $s_t$ and $s_{t-1}$, with
  \[
  \Phi =
  \begin{bmatrix}
  \phi_1 & \phi_2 \\
  1 & 0
  \end{bmatrix}
  \]

- To accommodate the additional lag the signal equation becomes
  \[
  y_t =
  \begin{bmatrix}
  \beta & 0
  \end{bmatrix}
  \begin{bmatrix}
  s_t \\
  s_{t-1}
  \end{bmatrix} + u_t
  \]
Maximum Likelihood Estimator

Using EViews

Consider estimating a one-factor model of the spread between the one-year yield and the one-month yield

\[ \text{YIELD}_1 - \text{YIELD}_M = \alpha + \beta s_t + u_t, \quad u_t \sim N(0, \sigma^2) \]

\[ s_t = \phi_1 s_{t-1} + \phi_2 s_{t-2} + \nu_t \quad \nu_t \sim N(0, 1) \]

The EViews window to estimate the model is given below.

![EViews Window](source: yields_us.wf1)

Source: yields_us.wf1

Note that the second factor \( s_2 \), represents the lag of the first factor \( s_1 \), and thus does not have a disturbance term.
An AR(p) Model of $s_t$

- Consider an $AR(p)$ model of $s_t$

$$s_t = \phi_1 s_{t-1} + \phi_2 s_{t-2} + \cdots + \phi_p s_{t-p} + \nu_t$$

- The state equation is written as

$$\begin{bmatrix}
  s_t \\
  s_{t-1} \\
  s_{t-2} \\
  \vdots \\
  s_{t-p+1}
\end{bmatrix} = \begin{bmatrix}
  \phi_1 & \phi_2 & \cdots & \phi_{p-1} & \phi_p \\
  1 & 0 & \cdots & 0 & 0 \\
  0 & 1 & \cdots & 0 & 0 \\
  \vdots & \vdots & \cdots & \vdots & \vdots \\
  0 & 0 & \cdots & 1 & 0
\end{bmatrix} \begin{bmatrix}
  s_{t-1} \\
  s_{t-2} \\
  s_{t-3} \\
  \vdots \\
  s_{t-p}
\end{bmatrix} + \begin{bmatrix}
  \nu_t \\
  0 \\
  0 \\
  \vdots \\
  0
\end{bmatrix}$$

- In this case, the model is viewed as having $p$ factors

$$\{s_t, s_{t-1}, \cdots, s_{t-p+1}\}$$

although it is really just the first element of this set of factors that is of interest.
Extensions

Dynamics

**Idiosyncratic Dynamics**

- Consider the model

\[
\begin{align*}
y_{i,t} &= \beta_i s_t + \sigma_i u_{i,t}, \quad i = 1, 2, \ldots, 4 \\
s_t &= \phi_1 s_{t-1} + \phi_2 s_{t-2} + \nu_t \\
u_{i,t} &= \delta_i u_{i,t-1} + w_{i,t}
\end{align*}
\]

where \( u_{i,t} \sim N(0, I) \) and \( w_{i,t} \sim N(0, I) \).

- The state equation is now augmented to accommodate the dynamics in the idiosyncratic terms.

\[
\begin{bmatrix}
  s_t \\
  s_{t-1} \\
  u_{1,t} \\
  u_{2,t} \\
  u_{3,t} \\
  u_{4,t}
\end{bmatrix}
= \begin{bmatrix}
  \phi_1 & \phi_2 & 0 & 0 & 0 & 0 & 0 \\
  1 & 0 & 0 & 0 & 0 & 0 & 0 \\
  0 & 0 & \delta_1 & 0 & 0 & 0 & 0 \\
  0 & 0 & 0 & \delta_2 & 0 & 0 & 0 \\
  0 & 0 & 0 & 0 & \delta_3 & 0 & 0 \\
  0 & 0 & 0 & 0 & 0 & \delta_4 & 0
\end{bmatrix}
\begin{bmatrix}
  s_{t-1} \\
  s_{t-2} \\
  u_{1,t-1} \\
  u_{2,t-1} \\
  u_{3,t-1} \\
  u_{4,t-1}
\end{bmatrix}
+ \begin{bmatrix}
  \nu_t \\
  0 \\
  w_{1,t} \\
  w_{2,t} \\
  w_{3,t} \\
  w_{4,t}
\end{bmatrix}
\]
Extensions

Dynamics

- In this case, the model is viewed as having six factors

\[ \{ s_t, s_{t-1}, u_{1,t}, u_{2,t}, u_{3,t}, u_{4,t} \} \]

- In this scenario the idiosyncratic terms are redefined as factors.

- As there are now no disturbances terms, then the covariance matrix of the disturbances \( E[u_t'u_t'] = R \), reduces to

\[ R = 0 \]

- The full model in this alternative parameterization is

\[
\begin{align*}
y_t &= Bs_t \\
s_t &= \Phi s_{t-1} + \nu_t, \quad \nu_t \sim N(0, I)
\end{align*}
\]

where \( s_t \) represents the vector of six factors.
The state-space model is easily extended to include $M$ exogenous or lagged dependent variables, $x_t$. These variables can be included in one of two different ways.

The first approach is to include exogenous or predetermined variables in the signal equation

$$y_t = B f_t + \Gamma x_t + u_t,$$

where $\Gamma$ is $(N \times M)$ and $x_t$ is $(M \times 1)$.

This class of model is called a factor VAR model (F-VAR), where $x_t = y_{t-1}$ and $\Gamma$ is now a $(N \times N)$ diagonal matrix.

The second approach is to include the exogenous or predetermined variables in the state equation

$$s_t = \Phi s_{t-1} + \Gamma x_t + u_t,$$

where $\Gamma$ is now a $(K \times M)$ matrix of parameters.
End of Lecture