Exercises For Wednesday Evening

1. Presented by: Jeremy Zuchuat

Suppose that Y_t follows the AR(1) model with heteroskedastic errors: $Y_t = \phi_1 Y_{t-1} + \varepsilon_t$, where $\varepsilon_t | Y_{0:t-1} \sim N(0, \sigma^2_{t-1})$ where $\sigma^2_{t-1} = \omega + \alpha Y^2_{t-1}$ for $t = 2, \ldots, T$. Suppose that $Y_0 = 0$.

(a) Write the explicit joint density/likelihood function for $Y_{1:T}$ (conditional on $Y_0 = 0$).

(b) The file **518_EX1_5.csv** contains a realization of T =100 observations for Y_t.

(i) Suppose $\phi = 0.8$, $\omega = 1$ and $0.0 \leq \alpha \leq 0.9$. Compute the MLE of α.

(ii) You are Bayesian. Before looking at the data you know that $\phi = 0.8$ and $\omega = 1$, but are unsure about the value of α. Your prior is $P(\alpha = 0.4) = 0.6$ and $P(\alpha = 0.7) = 0.4$. Use the data to find the posterior for α.

2. Presented by: Juliette Cattin

Consider the model used in the notes to derive the Kalman filter. Assume that $E(w_t v_t') = G$. (In the notes $G = 0$). Derive the Kalman filter.

3. Presented by: Giulia Sabbadini

Consider the model $y_t = s_t + \varepsilon_t$ where $\varepsilon_t \sim$ i.i.d. N(0,1) and s_t is a 0-1 binary random variable with $P(s_t = 1 | s_{t-1} = 0) = 0.3$ and $P(s_t = 1 | s_{t-1} = 1) = 0.8$.

(a) Suppose that the history of information on y tells you that $P(s_{t-1} = 1 \mid y_{1:t-1}) = 0.6$. You observe $y_t = 1.5$. Compute $P(s_t = 1 \mid y_{1:t})$.

(b) Generalize your calculations and derive a recursive algorithm for computing $P(s_t = 1 \mid y_{1:t})$ as a function of y_t and $P(s_{t-1} = 1 \mid y_{1:t-1})$. Explain how this result can be used the compute the likelihood function/joint density of $y_{1:T}$.

4. Presented by: Lorenz Driussi

Suppose that $Y_t = \tau_t + \varepsilon_t$, where $\tau_t = \tau_{t-1} + \eta_t$ and $\{\varepsilon_t\}$ and $\{\eta_t\}$ are mutually independent sequences of zero-mean normal random variables with standard deviations σ_ε and σ_η. The initial value $\tau_0 \sim N(0,\kappa^2)$ and is independent of (ε_t, η_t) for $t > 0$.

(a) Let $\tau_{t|t} = E(\tau_t \mid Y_{1:t})$. Show that $E(Y_{t+h} \mid Y_{1:t}) = \tau_{t|t}$ for all $h \geq 1$.

(b) The spreadsheet **518_EX1_6_FRED.xlsx** contains quarterly values on the GDP deflator (P) for the U.S. from 1990:Q1-2018:Q3. Let $Y_t = 400 \times \ln(P_t/P_{t-1})$ denote the inflation rate (in percentage points at an annual rate).

(i) Compute the sample variance of ΔY_t and the sample covariance between ΔY_t and ΔY_{t-1}. Use these to compute estimates of $(\sigma_\varepsilon, \sigma_\eta)$.

(ii) Use the estimates of $(\sigma_\varepsilon, \sigma_\eta)$ from (i) and a reasonable value for κ^2, the variance of τ_0, to compute $\tau_{t|t}$ for 1990:Q3 $\leq t \leq$ 2018:Q3.

(iii) Use the results in (ii) to forecast the average level of inflation in 2019. How precise is the forecast likely to be − that is, what is the mean and variance of the forecast error?

Some Additional Exercices

1. Suppose

$$y_t = \xi_t + w_t$$
$$\xi_t = 0.8\xi_{t-1} + v_t$$

where $w_t \sim iidN(0,2)$ and $v_t \sim iidN(0,3)$ and $\{w_t\}$ and $\{v_t\}$ are independent. Suppose that you know $\xi_{t-1} = 3.4$, and $y_t = 4.1$. Find $E(\xi_t \mid \xi_{t-1} = 3.4,\ y_t = 4.1)$ and var($\xi_t \mid \xi_{t-1} = 3.4,\ y_t = 4.1$).

2. Suppose that y_t follows the AR(1) model $y_t = \phi y_{t-1} + \varepsilon_t$, for $t = 1, 2, ..., 100$, with $y_0 = 0$ and ε_t ~ Niid(0, σ^2). Suppose data on y_{50} is missing.

Suppose that you know the values of ϕ and σ^2.

(a) Find an expression for $E(y_{50} \mid y_{49})$

(b) Write down an expression that would allow you to calculate $E(y_{50} \mid \{y_{49}, y_{51}\})$

(c) How would you construct $E(y_{50} \mid \{\{y_1, y_{2,} ... y_{49}, y_{51}, y_{52}, ..., y_{100}\})$?

3. Suppose that y_{1t} and y_{2t} are scalar random variables with

$$y_{1t} = x_t + \varepsilon_{1t}$$

$$y_{2t} = x_t + \varepsilon_{2t}$$

where x_t, ε_{1t}, and ε_{2t} are mutually independent i.i.d. sequences of $N(0,1)$ random variables. A researcher has data on y_{1t} and y_{2t} and would like to use these data to estimate the value of x_t. He proposes the estimator $\hat{x}_t = \frac{1}{2}(y_{1t} + y_{2t})$.

(a) Compute the mean squared error of $\hat{x}_t$.

(b) A more general estimator is $\tilde{x}_t = \lambda_1 y_{1t} + \lambda_{2t} y_{2t}$, where λ_1 and λ_2 are two constants. What values of λ_1 and λ_2 yield the estimator with the smallest mean squared error?

4. Suppose that $y_t = x_t + \varepsilon_t$, where $x_t = 0.8x_{t-1} + e_t$, and were $\begin{bmatrix} \varepsilon_t \\ e_t \end{bmatrix} \sim iidN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} \right)$. Suppose you know that $x_0 = 2$ and $y_1 = 6$.

(a) Derive the minimum mean square error estimate of x_1.

(b) What is the mean squared error of the estimate in (a)?

Now suppose now that $\{\varepsilon_t\}$ and $\{e_t\}$ are mutually independent iid processes with (i) $\varepsilon_t = -2$ with probability 0.5 and $\varepsilon_t = 2$ with probability 0.5, and (ii) $e_t = -1$ with probability 0.5 and $e_t = 1$ with probability 0.5. Suppose you know that $x_0 = 2$ and $y_1 = 6$

(c) Derive the <u>linear</u> minimum mean square error estimate of x_1.

(e) What is the mean squared error of this estimate?

(f) Is the estimate in (e) the minimum mean squared estimate? Explain.

5. Suppose that $y_t = x_t + u_t$, where $x_t = \varepsilon_t + 0.8\varepsilon_{t-1}$, and

$$\begin{bmatrix} \varepsilon_t \\ u_t \end{bmatrix} \sim iidN\left(\begin{bmatrix} 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 9 & 3 \\ 3 & 4 \end{bmatrix}\right).$$ You are told that $y_{100} = 6$.

(a) Compute the best (minimum mean square error) estimate of x_{100}.

(b) Compute the best (minimum mean square error) estimate of x_{101}.

6. Suppose that $y_{it} = x_t + \varepsilon_{it}$, for $i = 1, \ldots, n$, $(x_t, \{\varepsilon_{it}\}_{i=1}^n)$ are *i.i.d.* through time, and normally distributed with $x_t \sim N(0, \sigma_X^2)$, $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$, and ε_{it} is independent of ε_{jt} (for $j \neq i$) and x_t.

(a) Show that $x_{t/t} = \lambda\, \bar{Y}_t$, where $\bar{Y}_t = \frac{1}{n}\sum_{i=1}^{n} y_{it}$, and derive an expression for λ.

(b) Show that $\lim_{n\to\infty}\lambda = 1$

(c) Show that $plim_{n\to\infty} x_{t/t} = x_t$.

(d) Show that $x_{t/t}$ converges in mean square to x_t as $n \to \infty$.

7. Y_t follows the stationary AR(2) model $Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$, $\varepsilon_t \sim Niid(0,\sigma^2)$. Write the explicit joint density/likelihood function for $Y_{1:T}$.

8. Y_t follows the MA(1) model $Y_t = \varepsilon_t - \theta\varepsilon_{t-1}$, where $\varepsilon_t \sim Niid(0,\sigma^2)$ for $t = 1, \ldots, T$, and $\varepsilon_t = 0$ for $t \leq 0$.

(a) Write the explicit joint density/likelihood function for $Y_{1:T}$. Discuss how you would compute the MLE of θ and σ^2.

(b) Does the result in (a) require that $|\theta| < 1$? Explain.

(c) Suppose $\varepsilon_t \sim Niid(0,\sigma^2)$ for $t \leq 0$. How would you modify your answer to (a) and (b)?

9. Y_t follows the stationary AR(1) model $Y_t = \phi_1 Y_{t-1} + \varepsilon_t$, $\varepsilon_t \sim Niid(0,\sigma^2)$. You have data for $Y_{1:100}$ and $Y_{102:T}$ (so that data are missing at time $t = 101$). Write the joint density/likelihood for ($Y_{1:100}$, $Y_{102:T}$).

10. Hamilton (1994) derives the Kalman "smoother" as a recursive algorithm for computing $\xi_{t/T}$ and $P_{t/T}$ from ($\xi_{t+1/T}$, $P_{t+1/T}$, $\xi_{t+1/t}$, $P_{t+1/t}$, and $P_{t/t}$). The recursion (as reported in Hamilton) is

(1) $$J_t = P_{t/t} F' P_{t+1/t}^{-1}$$

(2) $$\xi_{t/T} = \xi_{t/t} + J_t(\xi_{t+1/T} - \xi_{t+1/t})$$

(3) $$P_{t/T} = P_{t/t} + J_t(P_{t+1/T} - P_{t+1/t})J_t'$$

Prove the validity of this algorithm.

11. Consider the model

$$y_t = \beta s_t + \varepsilon_t$$

where $\varepsilon_t \sim iidN(0,1)$, and s_t is a binary (0-1) variable that follows a Markov process with $p(s_t = 1|s_{t-1} = 0) = p_0$, $p(s_t = 0| s_{t-1} = 0) = 1-p_0$, $p(s_t = 1| s_{t-1} = 1) = p_1$, and $p(s_t = 0| s_{t-1} = 1) = 1-p_1$. Let $s_{t|t} = \mathrm{E}(s_t \mid y_{1:t})$.

(a) Derive a recursive algorithm that computes $s_{t|t}$ as a function of $s_{t-1|t-1}$ and y_t. (The filter will be a function of the model parameters β, p_0, and p_1. (Hint: recall that for a binary variable $E(s) = p(s = 1)$.)

(b) Derive $p(s_0 = 1)$ as a function of p_0 and p_1.

(c) In the spreadsheet EX2_1.xlsx you will find the realization $y_{1:100}$ constructed from a model with $\beta = 1$, $p_0 = 0.2$ and $p_1 = 0.7$.

(i) Plot the log-likelihood function $L(\beta)$ for $-1 \leq \beta \leq 3$ by computing the likelihood over a equally spaced grid of 100 values of β in this interval.

(ii) Use calculations like those in (i) to compute the MLE of β.

(iii) Suppose I have a prior that β is a draw from a truncated normal distribution with mean 1 and variance 1, and with $-1 \le \beta \le 3$.

(iii.a) Approximate this prior by a discrete prior on the 100 grid points from β in (i). Plot the prior probabilities.

(iii.b) Use the approximate prior from (iii.a) to compute the posterior for β. Plot the posterior.

12. Consider the state-space model:

$$y_t = H\xi_t + w_t,\ \ \xi_t = F\xi_{t-1} + v_t \text{ and where } \begin{bmatrix} w_t \\ v_t \end{bmatrix} \sim i.i.dN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} R & 0 \\ 0 & Q \end{bmatrix}\right)$$

and for simplicity, suppose that all variables are scalars, (H,F,R,Q) are known, and you know that $\xi_0 = 0$. Let $\xi_{1:T}$ denote the $T\times 1$ vector $(\xi_1, \xi_2, \dots \xi_T)'$, and similarly for $Y_{1:T}$. Suppose I observe $Y_{1:T}$. Because everything is Gaussian I know $\xi_{1:T} \mid Y_{1:T} \sim \text{N}(\mu,\Sigma)$ for suitable values of μ and Σ. I want to obtain a random draw from this $\text{N}(\mu,\Sigma)$ distribution. On was to do this is to use brute force to compute μ and Σ, and then compute the draw as $\xi_{1:T} = \mu + \Sigma^{1/2} z_{1:T}$, where $\Sigma^{1/2}$ satisfies $\Sigma=\Sigma^{1/2}\Sigma^{1/2\prime}$, and $z_{1:T}$ is a vector of iidN(0,1) random variables.

I want you to design a recursive algorithm to obtain a draw from the $\xi_{1:T} \mid Y_{1:T} = y_{1:T}$ distribution. Here's what I have in mind:

Step 1: Run a Kalman filter and find $\xi_{T|T}$ and $P_{T|T}$. Draw ξ_T from the $\text{N}(\xi_{T|T}, P_{T|T})$ distribution.

Step 2: Find the distribution of $\xi_{T-1} \mid (\xi_T, y_{1:T})$. It will be of the form $\text{N}(m_{T-1}, \sigma^2_{T-1})$, where you need to find m_{T-1} and σ^2_{T-1}. Draw ξ_{T-1} from the $\text{N}(m_{T-1}, \sigma^2_{T-1})$ distribution.

Step 3: Use the analogues of Step 2 to draw $\xi_{T-2}, \dots , \xi_1$.

As you derive the algorithm, You will need to recursively find distributions of, say, $\xi_t \mid (\xi_{t+1}, \dots , \xi_T , Y_{1:T})$. You should be able to show that these simplify, and that it suffices to find the distributions of $\xi_t \mid (\xi_{t+1}, Y_{1:t})$ Moreover, it should be possible to show that the mean and variances of these distributions can be computed using things already computed by the Kalman filter ($\xi_{t|t}$, $P_{t|t}$, and so forth) plus a few more things that will need to be calculated.