Time Series Econometrics for Macroeconomics (FEM21044-19)

W.M. Brus (484505) & C.A. Vriends (440435)

Question 1: Markov switching model [4 points]

For all the ML estimation procedures in this question, `fminunc` and `fmincon` are used as optimizers. The one with the lowest (highest) log-likelihood value is chosen if the optimizers do not reach a consensus.

1a) In the table one can see that for starting point one the σ_1 is larger than σ_2, indicating that μ_1 and σ_1 belong to the high volatility state. However, using the second starting point we see that σ_2 is larger than σ_1, indicating that μ_2 and σ_2 belong to the high volatility state here. Note further, that the differences between the μ values and the σ of the different starting points in fact result from using different starting points. Furthermore, this suggests that the parameters are switched if the first set of starting points are used.

ML parameters	p_{11}	p_{22}	μ_1	μ_2	σ_1	σ_2	LogL
starting point 1	0.88	0.95	-0.46	3.60	7.68	2.56	-520.95
starting point 2	0.96	0.82	4.13	-4.96	3.10	5.24	-518.67

1b) Figure 1 shows the plotted smoothed probabilities of being in the high volatility state. We see that they have very similar patterns. At the beginning they differ a bit due to the different starting points. However, later they converge. As the estimation of the states switched in the first set of starting points, we plot $Pr(S_t = 1|I_T)$ (red) for the first starting point and $Pr(S_t = 2|I_T)$ (green) for the second starting point.

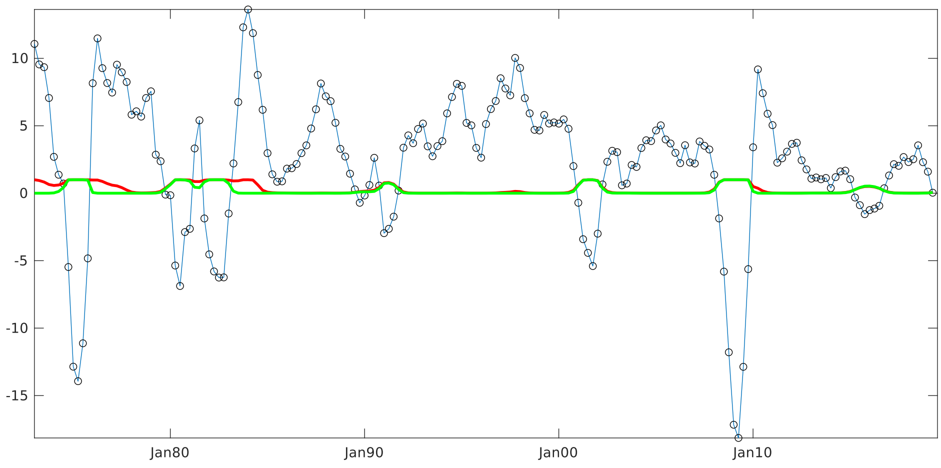

Figure 1: Smoothed probability of being in state 2

1c)

EM iterations	ρ_1	p_{11}	p_{22}	μ_1	μ_2	σ_1	σ_2	LogL
10	0.50	0.95	0.86	4.02	-1.78	2.70	6.98	-521.63
100	0	0.96	0.88	3.59	-0.46	2.56	7.69	-521.07
1,000	0	0.96	0.88	3.59	-0.46	2.56	7.69	-521.07

1d) We used starting point 2. One can observe that EM converges to the (switched) ML estimates of starting point 1. At first glance, this seems unexpected. However, the initialisation of the Hamilton filter is different, $Pr(S_0 = 1|I_0) = \frac{1}{2}$ for EM versus $Pr(S_0 = 1|I_0) = 1$ for ML. If one re-estimates the parameters (using EM with the ML filter initialisation), one will obtain (roughly) the same ML estimates of starting point 2 as in 1a. This hints at the likely cause of different initialisation.

1e) In both states the changes in Personal Income and Employment are positively correlated (positive covariances). Moreover, we see that in the high volatility state (second state), the covariances are higher, this is to be expected. Furthermore, these estimates imply that the correlation between the two series is lower in a high volatility state (37%) than in the lower volatility state (61%).

The estimates are estimated with the second starting point as initial values.

$$\mu_1 = \begin{pmatrix} 6.60 & 2.41 \end{pmatrix},\ \mu_2 = \begin{pmatrix} 5.17 & -0.76 \end{pmatrix},\ \Sigma_1 = \begin{pmatrix} 7.79 & 2.06 \\ 2.06 & 1.48 \end{pmatrix},\ \Sigma_2 = \begin{pmatrix} 31.44 & 4.77 \\ 4.77 & 5.18 \end{pmatrix}$$

$$\xi = \begin{pmatrix} 1.00 & 0.00 \end{pmatrix},\ p = \begin{pmatrix} 0.95 & 0.88 \end{pmatrix}$$

Question 2: State space model [6 points]
For all the ML estimation procedures in this question, `fminunc`, `fmincon` and `fminsearch` are used as optimizers. The one with the lowest (highest) log-likelihood value is chosen if the optimizers do not reach a consensus.

2a) The Kalman filter is initialized in the following manner and the results are provided below the initialization values:

$$\xi_{0|0} = 0,\ \mathbf{P}_{0|0} = 10^6$$

series	F	Q	R	$\mathrm{cor}(\widehat{y}_{i,t\|t-1}, y_{i,t})$	$\mathrm{cor}(\widehat{y}_{i,t\|t-4}, y_{i,t})$	LogL
$i = 1$	0.91	1.16	7.60	54%	40%	-488.16
$i = 2$	0.77	1.80	0.00	78%	25%	-324.23
$i = 3$	0.87	5.88	0.00	88%	7%	-433.99
$i = 4$	0.86	3.56	0.00	88%	12%t	-387.61

2b) The Kalman filter is initialized in the following manner and the results are provided below the initialization values:

$$\xi_{1|0} = \vec{\mathbf{0}},\ \mathbf{P}_{1|0} = 10^6 * \mathbf{I}_4$$

$$\mathbf{H} = \begin{pmatrix} 0.72 & 0.80 & 2.24 & 1.75 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \mathbf{F} = \begin{pmatrix} 0.88 & 0 & 0 & 0 & 0 \\ 0 & 0.38 & 0 & 0 & 0 \\ 0 & 0 & 0.48 & 0 & 0 \\ 0 & 0 & 0 & 0.69 & 0 \\ 0 & 0 & 0 & 0 & 0.54 \end{pmatrix}$$

$$\mathbf{Q} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 11.07 & 0 & 0 & 0 \\ 0 & 0 & 1.05 & 0 & 0 \\ 0 & 0 & 0 & 0.83 & 0 \\ 0 & 0 & 0 & 0 & 0.44 \end{pmatrix}, \text{ LogL} = -1471.57$$

series	$\text{cor}(\widehat{y}_{i,t\|t-1}, y_{i,t})$	$\text{cor}(\widehat{y}_{i,t\|t-4}, y_{i,t})$
$i = 1$	46%	13%
$i = 2$	79%	29%
$i = 3$	88%	9%
$i = 4$	87%	11%

2c) The P matrix represents the forecast errors made when predicting the states. The diagonal elements show the variance of the forecast error of each state. The off diagonal elements show how the forecast error of one predicted state is correlated to the forecast error of another predicted state. Most diagonal elements are nonzero. This means that for most states the forecast error is influenced by the prediction of the other states.

$$P = \begin{pmatrix} 0.11 & -0.08 & -0.09 & -0.25 \\ -0.08 & -0.06 & 0.06 & 0.18 \\ -0.09 & 0.06 & 0.07 & 0.20 \\ -0.25 & 0.18 & 0.20 & 0.55 \\ -0.19 & 0.14 & 0.16 & 0.43 \end{pmatrix}$$

2d) It is possible to make a one-step ahead forecast using the optimal parameters found at 2b. However, in this instance, we would not make use of the information given by the 187_{th} value of the series y_1, y_2, y_3. To incorporate this information, we first make a one-step ahead prediction for the 187_{th} value of y_4. We will substitute this value in the $\mathbf{y}$ matrix including the 187_{th} value of y_1, y_2 an y_3. Hence, we can then use this $\mathbf{y}_{4x187}$ matrix to estimate the filtered $\xi_{i,186:186}$ for $i = 1, 2, 3, 4$ and $\mathbf{F}, \mathbf{H}$ and $\mathbf{Q}$. These parameters include more information than the one-step ahead forecast with the $\mathbf{y}_{4x186}$ matrix and the re-estimated filtered $\xi_{i,186:185}$ is likely better than the filtered $\xi_{i,186:186}$ with the smaller dataset for all i. We can now use these estimated parameters to make a one-step ahead forecast of the 187_{th} observation of y_1, y_2, y_3 and y_4. This prediction is done in the following manner, where $\mathbf{F}$, $\mathbf{H}$ and $\xi_{186:186}$ are estimated using the $\mathbf{y}_{4x187}$ matrix:

$$y_{i,187:186} = \mathbf{H'F}\xi_{186:186}, \ i = 1,2,3,4 \tag{1}$$

Hence, we will use this forecast of y_4 as our optimal forecast. An indication of the quality of our forecast is how far the one-step ahead forecast of $y_{1,187}$, $y_{2,187}$ and $y_{3,187}$ lie from their actual value. The table below shows the differences between the observed values ($y_{1,187}$, $y_{3,187}$, and $y_{3,187}$) and their forecasted values ($\hat{y}_{1,187}$, $\hat{y}_{3,187}$, and $\hat{y}_{3,187}$). We see that for y_1 and y_3 the differences are quite far from zero, however for the y_2 we see the forecast is quite close to the observed value. Furthermore, we note that the difference between the forecast using the $\mathbf{y}_{4x187}$ (that includes the 187_{th} observation) and y_{4x186} matrix is rounded to zero and negligible. Hence, using the extra information does not alter the forecast of y_4 in a meaningful manner. The forecast for RMTIS observation for first of July 2019 is: 1.82

Series	y_1	y_2	y_3
$\widehat{y}_{i,187:186} - y_{i,187:187}$	-1.88	-0.05	-1.11

2e) The Kalman filter is initialized in the following manner and the results are provided below the initialization values:

$$\xi_{1|0} = \vec{\mathbf{0}}, \ \mathbf{P}_{1|0} = 10^6 * \mathbf{I}_4$$

$$\mathbf{R} = \begin{pmatrix} 7.58 & 0 & 0 & 0 \\ 0 & 0.15 & 0 & 0 \\ 0 & 0 & 0.01 & 0 \\ 0 & 0 & 0 & 0.01 \end{pmatrix}, \ \text{LogL} = -1450.61$$

	$\text{cor}(\widehat{y}_{i,t\mid t-1}, y_{i,t})$	$\text{cor}(\widehat{y}_{i,t\mid t-4}, y_{i,t})$
$i = 1$	49%	20%
$i = 2$	79%	30%
$i = 3$	88%	10%
$i = 4$	87%	13%

2f) We have estimated the set of parameters twice using the Expectation-Maximization (EM) algorithm. The first set of initial starting values is the following ($\hat{\Sigma}_y$ is the sample covariance matrix of the demeaned series):

$$\mathbf{F}_0 = \begin{pmatrix} 0.95 & 0.10 & 0.10 & 0.10 \\ 0.10 & 0.95 & 0.10 & 0.10 \\ 0.10 & 0.10 & 0.95 & 0.10 \\ 0.10 & 0.10 & 0.10 & 0.95 \end{pmatrix}, \ \mathbf{Q}_0 = 2*\hat{\Sigma}_y, \ \mathbf{R}_0 = 4*\hat{\Sigma}_y, \ \xi_{1|0} = \vec{\mathbf{0}}, \ \mathbf{P}_{1|0} = 10^6*\mathbf{I}_4$$

This first set of starting values leads to the following estimated matrices. Note, that the precision for $\mathbf{F}$ is higher than the conventional precision of two decimal places. This is due to the fact that the elements in $\mathbf{F}$ would have been 0 otherwise.

$$\mathbf{F} = \begin{pmatrix} 5*10^{-4} & 3*10^{-4} & 9*10^{-4} & 7*10^{-4} \\ 4*10^{-4} & 4*10^{-4} & 1.1*10^{-3} & 9*10^{-4} \\ 9*10^{-4} & 1.1*10^{-3} & 3*10^{-3} & 2.5*10^{-3} \\ 5*10^{-4} & 9*10^{-4} & 2.3*10^{-3} & 2.1*10^{-3} \end{pmatrix}, \ \mathbf{Q} = \begin{pmatrix} 8.68 & 2.86 & 6.41 & 4.47 \\ 2.86 & 3.54 & 7.67 & 6.28 \\ 6.41 & 7.67 & 21.86 & 16.85 \\ 4.47 & 6.28 & 16.85 & 14.35 \end{pmatrix}$$

$$\mathbf{R} = \begin{pmatrix} 6.28 & 0.89 & 1.07 & 0.91 \\ 0.89 & 1.04 & 1.02 & 0.76 \\ 1.07 & 1.02 & 4.03 & 2.36 \\ 0.91 & 0.76 & 2.36 & 1.45 \end{pmatrix}, \ \mathrm{LogL} = -1676.96$$

series	$\mathrm{cor}(\widehat{y}_{i,t\|t-1}, y_{i,t})$	$\mathrm{cor}(\widehat{y}_{i,t\|t-4}, y_{i,t})$
$i=1$	35%	11%
$i=2$	76%	30%
$i=3$	89%	10%
$i=4$	87%	11%

The second set of initial startingvalues is the following:

$$\mathbf{F}_0 = \begin{pmatrix} 0.95 & 0.10 & 0.10 & 0.10 \\ 0.10 & 0.95 & 0.10 & 0.10 \\ 0.10 & 0.10 & 0.95 & 0.10 \\ 0.10 & 0.10 & 0.10 & 0.95 \end{pmatrix}, \ \mathbf{Q}_0 = 2*\hat{\Sigma}_y, \ \mathbf{R}_0 = 4*\hat{\Sigma}_y, \ \xi_{0|0} = \vec{\mathbf{0}}, \ \mathbf{P}_{0|0} = 10^6*\mathbf{I}_4$$

The second set of startingvalues leads to the following estimated matrices:

$$\mathbf{F} = \begin{pmatrix} 1.01 & -0.29 & -0.33 & 0.51 \\ 0.52 & -0.84 & -0.56 & 1.32 \\ 2.30 & -7.38 & -1.74 & 5.99 \\ 1.49 & -4.90 & -1.62 & 4.67 \end{pmatrix}, \ \mathbf{Q} = \begin{pmatrix} 1.56 & 0.47 & 0.53 & 0.43 \\ 0.47 & 0.59 & 0.51 & 0.92 \\ 0.53 & 0.51 & 1.53 & 1.26 \\ 0.43 & 0.92 & 1.26 & 1.73 \end{pmatrix}$$

$$\mathbf{R} = \begin{pmatrix} 7.33 & 0.052 & 0.62 & 0.68 \\ 0.05 & 0.72 & 0.19 & -0.13 \\ 0.62 & 0.19 & 0.22 & 0.14 \\ 0.68 & -0.13 & 0.14 & 0.22 \end{pmatrix}, \ \mathrm{LogL} = -1372.09$$

series	$\mathrm{cor}(\widehat{y}_{i,t\|t-1}, y_{i,t})$	$\mathrm{cor}(\widehat{y}_{i,t\|t-4}, y_{i,t})$
$i=1$	0.54	35%
$i=2$	83%	26%
$i=3$	92%	32%
$i=4$	90%	30%