MSc Econometrics and Management Science
Assignment for FEM21038-18 (Advanced Time Series Econometrics)

Instructions

This assignment covers two important topics from this course: state space and Markov switching models. The assignment should be done in pairs and you are not allowed to cooperate with other groups.

Your answers should consist of complete sentences, which can be understood independently of the questions.

Note 1: Motivate your answers! You will be rewarded points for showing you understand the technical aspects of the course, but also for intuitive, clearly written explanations, which demonstrate you master the subject at a deeper level.

Note 2: Feel free to use the Matlab demo codes provided during the course as your starting point. While these are provided for your convenience, note that they will not be sufficient to answer all questions. Of course you are free to use any other software.

Submission guidelines:

- Please answer the questions in the assignment as stated. In other words, there is no need to write additional material/explanation in the form of a 'report'. You may actually find it useful to use the equations/tables in this LaTeX file as your starting point.
- Reports should be be submitted via Canvas (under "Assignments" on the course pages). The submission deadline is **23 November 2018, 23:59**.
- Reports should be submitted as a PDF file with the name

  "`FEM21038-18-XXXXXX-YYYYYY.pdf`"

  where `XXXXXX` and `YYYYYY` are the student numbers of the two group members.
- Please make sure the file name is correct and your names and student numbers appear in the top-right corner of the first page of the report.
- One group submission is sufficient (as long as you name the file correctly, see above).
- Matlab codes (or any other code) should <u>not</u> be included in the main text.
- Instead, please submit a separate .zip file containing sub-folders with (Matlab) codes/files.

Good luck! For questions related to the assignment, please send an email to:
`lange@ese.eur.nl`

Data description

This assignment considers the following two time series:

- Real Gross Domestic Product (GDP), Billions of 2012 Dollars, Quarterly, Seasonally Adjusted Annual Rate
- Consumer Price Index (CPI) for All Urban Consumers: All Items, Index 1982-1984=100, Monthly, Seasonally Adjusted

These series, which are sampled at a quarterly and monthly frequency, respectively, can be found in the file `data.xls`. To obtain a balanced panel, we consider quarterly data for both series. Because both series contain *levels*, we compute annualised percentage changes by taking $y_{i,t} := 4 \times 100 \times (Y_{i,t}/Y_{i,t-1} - 1)$, where the time index t counts quarters. Here, $Y_{i,t}$ denotes the level, while $y_{i,t}$ denotes the (annualised) percentage change. The subscript $i = 1$ ($i = 2$) refers to GDP (CPI) data. Further, the index t runs from $t = 1$ corresponding to Q1 (April) 1947 to $T = 285$ corresponding to Q1 (April) 2018. The data transformation has already been performed for you in the sheet "data for assignment". The data *after this transformation* can also be found in `data.csv` or `data.mat`, where the last file can be loaded directly into Matlab. For your convenience, the Matlab demo file `assignment.m` plots these two series with correct dates displayed on the horizontal axis.

Question 1: Markov switching model [5 points]

1a) Consider the series $y_{1,t}$ corresponding to the GDP. Use maximum likelihood (ML) to estimate a Markov switching model with switching mean and variance, that is

$$y_{1,t} \sim \mathrm{N}(\mu_{S_t}, \sigma^2_{S_t})$$

where S_t is a random variable with $\Pr[S_t = 1|S_{t-1} = 1] = p_{11}$ and $\Pr[S_t = 2|S_{t-1} = 2] = p_{22}$. We use the convention that $S_t = 2$ denotes the high-volatility state. You are asked to investigate three possible initialisations. Initialise your model by imposing

- that the filter starts in state one, i.e. set $\Pr(S_0 = 1|\mathcal{I}_0) = 1$
- that the filter starts in state two, i.e. set $\Pr(S_0 = 2|\mathcal{I}_0) = 1$
- that the filter starts in the long-term mean, set $\Pr(S_0 = 1|\mathcal{I}_0) = (1 - p_{22})/(2 - p_{11} - p_{22})$ and $P(S_0 = 2|\mathcal{I}_0) = (1 - p_{11})/(2 - p_{11} - p_{22})$

You may consult (and use) the Matlab demo code you have been provided with during the course. Report your parameter estimates and resulting likelihood function in this table:

Initialisation	p_{11}	p_{22}	μ_1	μ_2	σ_1	σ_2	LogL
$\Pr(S_0 = 1\|\mathcal{I}_0) = 1$	·	·	·	·	·	·	·
$\Pr(S_0 = 2\|\mathcal{I}_0) = 1$	·	·	·	·	·	·	·
long-term mean	·	·	·	·	·	·	·

Display two decimal places for your parameter estimates, and one for the log likelihood.

1b) Which initialisation in question 1a do you believe to be more 'correct'? Plot the smoothed probability $\Pr(S_t = 2|\mathcal{I}_T)$ for this particular initialisation and, in no more than five sentences, comment on the result.

1c) Pick any initialisation as in question 1a and estimate the model again, but now enforce the restriction $p_{11} + p_{22} = 1$. For this restricted model, plot the smoothed probability $\Pr(S_t = 2|\mathcal{I}_T)$ over time and compare the resulting graph with the one in question 1b. In no more than five sentences, comment on the similarities/differences.

1d) Next, use the EM algorithm for Markov switching models to estimate the six parameters $p_{11}, p_{22}, \mu_1, \mu_2, \sigma_1, \sigma_2$ along with the starting points $\Pr(S_0 = 1) = \rho_1$ and $\Pr(S_0 = 2) = \rho_2$. Since $\rho_1 + \rho_2 = 1$, there are 7 parameters to estimate. As your initial parameter values, take $p_{11} = p_{22} = 0.8$, set μ_1 and μ_2 equal to the sample mean, and take σ_1 and σ_2 equal to 0.5 and 1.5 times the sample standard deviation, respectively. For your first EM run, set $\hat{\boldsymbol{\xi}}_{0|0} = (\rho_1, \rho_2)'$ with $\rho_1 = \rho_2 = 0.5$. Each EM step should update 7 parameters indicated in the table below. Display your results up to four decimal places.

- Run 40 EM steps and complete the following table:

EM estimates (40 iterations)	ρ_1	p_{11}	p_{22}	μ_1	μ_2	σ_1	σ_2	LogL
	.	.	.	.	.	.	.	.

- Plot the 'convergence' of these estimated parameters for the first 40 EM iterations, so that you can visually see that the estimates 'flatten out'. Paste this picture into your report.

- Run 1,000 EM steps and complete the following table:

EM estimates (10^3 iterations)	ρ_1	p_{11}	p_{22}	μ_1	μ_2	σ_1	σ_2	LogL
	.	.	.	.	.	.	.	.

- In no more than five sentences, discuss why the EM parameters after 1,000 iterations are similar and/or different to the parameters and log likelihood obtained in question 1a.

1e) *Bonus question.* The changes in GDP and CPI are expected to be correlated. Furthermore, this correlation might be higher in times of crisis. To investigate whether this is the case, consider again a two-state Markov switching model, except now the observation $\mathbf{y}_t$ is two dimensional. That is, the observation consists of two percentages (GDP growth and inflation), which are assumed to be jointly normally distributed with mean $\boldsymbol{\mu}_{S_t}$ and covariance matrix $\boldsymbol{\Sigma}_{S_t}$, where both the mean and the covariance depend on the state S_t. Estimate this model by EM and report your estimates. Interpret your results in no more than five sentences.

Question 2: State space model [5 points]

This question considers the same data as the previous question. However, for the purpose of this question, you should 'de-mean' the two data series $y_{1,t}$ and $y_{2,t}$. From this point onwards, we use the symbol $y_{i,t}$ with $i = 1, 2$ to denote the two de-meaned macroeconomic series.

2a) Estimate a simple 'AR(1) plus noise' model for both univariate series $y_{i,t}$ for $i = 1, 2$ and $t = 1 \ldots T$ by maximum likelihood (ML). That is, estimate the following model

$$y_{i,t} = \xi_{i,t} + w_t, \quad w_t \sim \mathrm{N}(0, R), \quad \xi_{i,t+1} = F\,\xi_{i,t} + v_t, \quad v_t \sim \mathrm{N}(0, Q),$$

for $i = 1, 2$. Use the Kalman filter with a diffuse initialisation; that is, take $\widehat{\xi}_{i,0|0} = 0$ and $P_{i,0|0} = 10^6$. Compute the log likelihood using the prediction-error decomposition and do not use a 'burning period' (i.e., sum over all observations when computing the log likelihood). Report your estimates in the table below (use two decimal places for your parameter estimates and one decimal place for the log likelihood) and paste a figure of $y_{2,t}$ along with your smoothed state estimates $\widehat{\xi}_{2,t|T}$ into your answer sheet. *Hint*: to answer this question, you may use the Kalman filter demo code you have been provided with.

series	F	Q	R	LogL
$i = 1$	·	·	·	·
$i = 2$	·	·	·	·

Note that, for computational reasons, `Matlab` minimizes negative log likelihoods; in your answer, please report (actual) log likelihoods.

2b) Next, estimate a two-dimensional state space model on both series considered jointly. That is, consider the following state space model:

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} \xi_{1,t} \\ \xi_{2,t} \end{pmatrix} + \mathbf{w}_t, \qquad \begin{pmatrix} \xi_{1,t+1} \\ \xi_{2,t+1} \end{pmatrix} = \begin{pmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{pmatrix} \begin{pmatrix} \xi_{1,t} \\ \xi_{2,t} \end{pmatrix} + \mathbf{v}_t, \tag{1}$$

and where $\mathbf{w}_t \sim \mathrm{N}(\mathbf{0}, \mathbf{R})$ and $\mathbf{v}_t \sim \mathrm{N}(\mathbf{0}, \mathbf{Q})$ are independently normally distributed innovations with covariance matrices

$$\mathbf{R} = \begin{pmatrix} R_{11} & R_{12} \\ R_{12} & R_{22} \end{pmatrix}, \quad \text{and} \quad \mathbf{Q} = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{12} & Q_{22} \end{pmatrix}. \tag{2}$$

Adjust the demo code for the Kalman filter that you have been given and estimate 10 unknown parameters $(F_{11}, F_{12}, F_{21}, F_{22}, Q_{11}, Q_{12}, Q_{22}, R_{11}, R_{12}, R_{22})'$ by maximum likelihood (ML). In your answer sheet, display two decimal places for each parameter estimate and one decimal place for the resulting log likelihood:

$$\mathbf{F} = \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}, \quad \mathbf{Q} = \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} \qquad \mathbf{R} = \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}, \quad \mathrm{LogL} = \cdot$$

As above, use a diffuse initialisation in the Kalman filter and do not use a 'burning period' when computing the log likelihood (i.e. compute the log likelihood by summing over all observations).

Hint 1: You may use the Matlab optimisers `fmincon`, `fminunc` or `fminsearch`. For `fminunc` use the following settings:

```
options = optimset('fminunc');
options = optimset(options , 'MaxFunEvals' , 1e+6);
options = optimset(options , 'MaxIter' ,1e+6);
options = optimset(options , 'TolFun' , 1e-6);
options = optimset(options , 'TolX' , 1e-6);
```

and likewise for the other two optimisers. Depending on your computer, the optimisation should not take longer than two minutes (`fminunc` tends to be fastest, `fminsearch` slowest).

Hint 2: It may help to choose a 'good' starting point for the optimisation.

Hint 3: The Matlab function `mvnpdf` is know to give problems. Instead, use your own expression for the multivariate normal pdf. The multivariate normal distribution with parameters `mu` and `Sigma` evaluated at `y` is given by

```
1/sqrt(det(2*pi*Sigma))*exp(-1/2*(y-mu)'*inv(Sigma)*(y-mu))
```

Hint 4: The Matlab function `inv` can be slow or inaccurate. Instead, use `A\ B` rather than `inv(A)B` for two matrices `A` and `B` of appropriate size. In this case, the multivariate normal distribution above reads

```
1/sqrt(det(2*pi*Sigma))*exp(-1/2*(y-mu)'*((Sigma)\(y-mu)))
```

2c) In no more than five sentences, discuss and interpret your parameter estimates in question 2b.

2d) How is the obtained log likelihood in question 2b related to the log likelihoods obtained in equation 2a? Do you prefer the two univariate models or the multivariate model? Your answer should contain at most six sentences.

2e) Use the expectation maximisation (EM) algorithm to estimate the parameters in the model given by (1) and (2). You may use the diffuse initialisation for $\boldsymbol{\xi}_0$ in each EM step, as in the EM demo code you received. As your starting point for the E-step of the algorithm, take

$$\mathbf{F} = \begin{pmatrix} 0.8 & 0 \\ 0 & 0.8 \end{pmatrix}, \quad \mathbf{Q} = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \qquad \mathbf{R} = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix},$$

Then report your parameter estimates after 10, 100 and 1,000 iterations using two decimal places as follows:

EM estimates after 10 iterations:

$$\mathbf{F} = \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}, \quad \mathbf{Q} = \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} \qquad \mathbf{R} = \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix},$$

EM estimates after 100 iterations:

$$\mathbf{F} = \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}, \quad \mathbf{Q} = \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} \qquad \mathbf{R} = \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix},$$

EM estimates after $1,000$ iterations:

$$\mathbf{F} = \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}, \quad \mathbf{Q} = \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} \qquad \mathbf{R} = \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix},$$

Hint 1: Due to bounded machine precision, the filtered/smoothed covariance matrix `P(:,:,t)` may become (slightly) unsymmetrical if the Kalman filter/smoother is computed many times. To avoid this from happening, you may calculate `P(:,:,t)` as usual and then set

```
P(:,:,t)=(P(:,:,t)+P(:,:,t)')/2
```

thereby entering a hard constraint that enforces symmetry.

Hint 2: Similarly, the covariance matrices in the M step of the EM-algorithm may become (slightly) unsymmetrical if the EM step is computed many times. To avoid this from happening, you may compute `Q` and `R` as usual, and then set `Q=(Q+Q')/2` and `R=(R+R')/2` to enforce symmetry before proceeding with another EM step.

2f) *Bonus question.* In question 2e, suppose $\boldsymbol{\xi}_0 \sim \mathrm{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$. Can you amend the EM algorithm to estimate $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$? Report your results using two decimal places. Does your estimate of $\boldsymbol{\Sigma}_0$ converge? If so, to what? Provide a detailed explanation of your answer.

Bonus question [1 point]

3) Consider the EM procedure proposed at the top of page 513 of 'Economic Forecasting' by Elliott & Timmermann (2016), which is paraphrased here for your convenience, and using our notation. They suggest to (i) pick some $\mathbf{H}$, $\mathbf{F}$, $\mathbf{Q}$ and $\mathbf{R}$, (ii) run the Kalman filter to obtain filtered state estimates $\hat{\boldsymbol{\xi}}_{t|t}$ and (iii) estimate the four system matrices by running linear regressions on the equations

$$\mathbf{y}_t = \mathbf{H}'\hat{\boldsymbol{\xi}}_{t|t} + \mathbf{w}_t, \quad \mathbf{w}_t \sim \mathrm{N}(\mathbf{0}, \mathbf{R}),$$

$$\hat{\boldsymbol{\xi}}_{t|t} = \mathbf{F}\hat{\boldsymbol{\xi}}_{t-1|t-1} + \mathbf{v}_t, \quad \mathbf{v}_t \sim \mathrm{N}(\mathbf{0}, \mathbf{Q}),$$

where $\mathbf{y}_t$ and $\hat{\boldsymbol{\xi}}_{t|t}$ are treated as known, while $\mathbf{H}$, $\mathbf{F}$, $\mathbf{Q}$ and $\mathbf{R}$ are estimated by ordinary least squares (OLS). The idea is to repeat steps (ii) and (iii) until convergence. Disprove their procedure by showing that the 'true' parameters are not a fixed point

of the proposed algorithm.

Hint: Suppose we are *given* the true matrices $\mathbf{H}$, $\mathbf{F}$, $\mathbf{Q}$ and $\mathbf{R}$ and we perform step (ii) of the proposed procedure. If the proposed procedure is correct, then $\mathbf{y}_t - \mathbf{H}'\hat{\boldsymbol{\xi}}_{t|t}$ should be distributed as $\mathrm{N}(\mathbf{0}, \mathbf{R})$. Similarly, $\hat{\boldsymbol{\xi}}_{t|t} - \mathbf{F}\hat{\boldsymbol{\xi}}_{t-1|t-1}$ should be distributed as $\mathrm{N}(\mathbf{0}, \mathbf{Q})$. Disprove both these claims by deriving the 'true' distribution of $\mathbf{y}_t - \mathbf{H}'\hat{\boldsymbol{\xi}}_{t|t}$ and $\hat{\boldsymbol{\xi}}_{t|t} - \mathbf{F}\hat{\boldsymbol{\xi}}_{t-1|t-1}$ under the true parameters.