1. a) Prove the chain rule for conditional probabilities:

$$P(A \cap B|C) = P(A|B \cap C) \cdot P(B|C)$$

Hint: $P(A|B) := P(A \cap B)/P(B)$.

The chain rule for conditional probabilities generalizes to probability densities as $f(x_1, x_2|x_3) = f(x_1|x_2, x_3) \cdot f(x_2|x_3)$, which implies for the likelihood function $L(\theta)$ of T sample observations $L(\theta) = f(y_1, \ldots, y_T|\theta)$ $= f(y_T|\mathcal{F}_{T-1}; \theta) \times f(y_{T-1}|\mathcal{F}_{T-2}; \theta) \times \cdots \times f(y_2|\mathcal{F}_1; \theta) \times f(y_1; \theta)$, where $\mathcal{F}_t = \{y_1, \ldots, y_t\}$. Let now S_1, \ldots, S_k be k state variables with possible values $S_i = \{0, 1, \ldots, n\}$. Then the law of total probability states that

$$f(x|x_1, x_2, \ldots) = \sum_{S_1=0}^n \cdots \sum_{S_k=0}^n f(x|S_1, \ldots, S_k; x_1, x_2, \ldots) P(S_1, \ldots, S_k|x_1, x_2, \ldots).$$

b) Given the information above, derive the log-likelihood function

$$\ell(\theta) = \sum_{t=1}^{T} \log \left[\sum_{S_t=0}^{1} \sum_{S_{t-1}=0}^{1} f(y_t | S_t, S_{t-1}, \mathcal{F}_{t-1}) P(S_t, S_{t-1} | \mathcal{F}_{t-1}) \right]$$

for the two state Markov switching model with the notation from the lecture notes.

c) Use the chain rule and the Markov property in order to show that the probability to find the two state Markov switching model for k periods in state 0 is

$$P(D_0 = k) := P(S_k = S_{k-1} = \dots = S_1 = 0 | S_0 = 1) = p^{k-1}(1-p),$$

where $p := P(S_t = 0 | S_{t-1} = 0).$

d) Show that

$$E(D_0) = \sum_{k=0}^{\infty} kP(D_0 = k) = \frac{1}{1-p}.$$

Hint: Taking the derivative of $\sum_{k=0}^{\infty} p^k = 1/(1-p)$ with respect to p gives a handy formula for $\sum_{k=0}^{\infty} kp^{k-1}$.

e) Use your result obtained in d) in order to show that

$$P(S_0 = 0|\mathcal{F}_0) = \frac{1-q}{2-p-q}$$
 and $P(S_0 = 1|\mathcal{F}_0) = \frac{1-p}{2-p-q}$,

where $p := P(S_t = 0 | S_{t-1} = 0)$ and $q := P(S_t = 1 | S_{t-1} = 1)$. Hint: The unconditional probability to find the model in a certain state is the expected duration of that state as a fraction of the sum of expected durations of all states.

2. Show that the multivariate dynamic regression model

$$\mathbf{y}_{t} = \mathbf{C} + \sum_{i=1}^{p} \mathbf{A}'_{i} \mathbf{y}_{t-i} + \sum_{i=0}^{p} \mathbf{B}'_{i} \mathbf{x}_{t-i} + \epsilon_{t}$$

with the notation from the lecture notes may be compiled in matrix form as $\mathbf{Y} = \mathbf{XB} + \mathbf{U}$ where $\mathbf{Y} = (\mathbf{y}_{p+1}, \dots, \mathbf{y}_T)', \mathbf{X} = (\mathbf{X}_{p+1}, \dots, \mathbf{X}_T)',$ and $\mathbf{U} = (\epsilon_{p+1}, \dots, \epsilon_T)'$ with $\mathbf{X}_t = (1, \mathbf{y}'_{t-1}, \dots, \mathbf{y}'_{t-p}, \mathbf{x}'_t, \dots, \mathbf{x}'_{t-p})$ and $\mathbf{B} = (\mathbf{C}, \mathbf{A}'_1, \dots, \mathbf{A}'_p, \mathbf{B}'_0, \dots, \mathbf{B}'_p)'.$

Hint: Write the regression equations for \mathbf{y}_t besides each other for all time points where they are defined, that is $t = p+1, \ldots, T$. Transposing yields the desired result.

3. With the notation from the lecture notes, show that the partitioned VAR

$$\mathbf{y}_t = \sum_{i=1}^p \mathbf{C}_{2i} \mathbf{x}_{t-i} + \sum_{i=1}^p \mathbf{D}_{2i} \mathbf{y}_{t-i} + \nu_{1t} \\ \mathbf{x}_t = \sum_{i=1}^p \mathbf{E}_{2i} \mathbf{x}_{t-i} + \sum_{i=1}^p \mathbf{F}_{2i} \mathbf{y}_{t-i} + \nu_{2t}$$

with $E(\nu_{it}\nu'_{jt}) = \Sigma_{ij}, i, j = 1, 2$ may be rewritten as

$$\begin{aligned} \mathbf{y}_t &= \sum_{i=0}^p \mathbf{C}_{3i} \mathbf{x}_{t-i} + \sum_{i=1}^p \mathbf{D}_{3i} \mathbf{y}_{t-i} + \omega_{1t} \\ \mathbf{x}_t &= \sum_{i=1}^p \mathbf{E}_{3i} \mathbf{x}_{t-i} + \sum_{i=0}^p \mathbf{F}_{3i} \mathbf{y}_{t-i} + \omega_{2t}, \end{aligned}$$

with block diagonal error covariance matrix $\Sigma_{\omega} = \begin{pmatrix} \Sigma_{\omega 1} & 0 \\ 0 & \Sigma_{\omega 2} \end{pmatrix}$.

Hint: Multiply the former system with the matrix given in the lectures.

4. Show that both the sum and the product of two lower triangular matrices are lower triangular as well, which implies that y does not react to a shock in x in a bivariate VAR system $\Phi(L)\begin{pmatrix} y_t\\ x_t \end{pmatrix} = \begin{pmatrix} \epsilon_{1,t}\\ \epsilon_{2,t} \end{pmatrix}$ if y is block-exogeneous with respect to x.