DD2434/FDD3434 Machine Learning, Advanced Course
Assignment 1E, 2025

Harald Melin, Marzie Abdolhamdi, Jens Lagergren

Deadline, see Canvas

Read this before starting

You will present the assignment by a written report in PDF format, submitted before the deadline using Canvas. The assignment should be done in groups of two, and it will automatically be checked for similarities to other students' solutions as well as documents on the web in general. Although you are allowed to discuss the problem formulations with other groups, you are not allowed to discuss solutions, and any discussions concerning the problem formulations must be described in the solutions you hand in (including which group you discussed with).

From the report it should be clear what you have done and you need to support your claims with results. You are supposed to write down the answers to the specific questions detailed for each task. This report should clearly show how you have drawn your conclusions and explain your derivations. Your assumptions, if any, should be stated clearly. Show the results of your experiments using images and graphs together with your analysis and add your code as an appendix.

Being able to communicate results and conclusions is a key aspect of scientific as well as corporate activities. It is up to you as an author to make sure that the report clearly shows what you have done. Based on this, and only this, we will decide if you pass the task. No detective work should be required on our side. In particular, neat and tidy reports please!

The grading of the assignment 1E, 2E and 3E (20 points each) will be as follows,

E 40 points, with least 10 points from each Assignment.

- All points over 40 will be counted as bonus points for assignment 1AD and 2AD.

Good Luck!

1.1 Dependencies in a Directed Graphical Model

Consider the graphical models shown in Figures 1 and 2. We denote d-separation here by "$\perp$". You merely have to answer "yes" or "no" to each question.(0.5 point per correct answer)

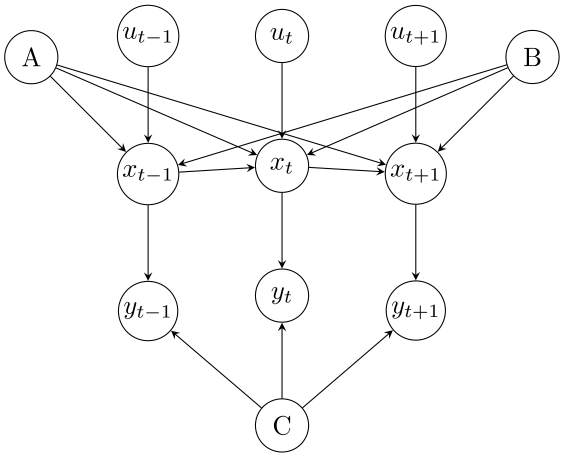

Figure 1: Extract of a structured state space model, the central model of e.g. Mamba.

Question 1.1.1: *In the graphical model of Figure 1, is* $y_{t-1} \perp y_t \mid x_t, C$*?*

Question 1.1.2: *In the graphical model of Figure 1, is* $x_{t-1} \perp x_{t+1} \mid x_t, C$*?*

Question 1.1.3: *In the graphical model of Figure 1, is* $A \perp B \mid x_t, C$*?*

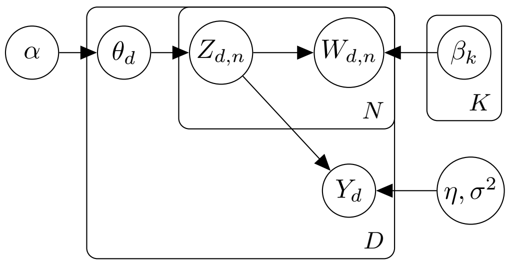

Figure 2: Bayesian network of Supervised Latent Dirichlet Allocation.

Question 1.1.4: *In the graphical model of Figure 2, is* $Z_{d,n} \perp Z_{d,n+1} \mid \theta_d, W_{d,n}, W_{d,n+1}$*?*

Question 1.1.5: *In the graphical model of Figure 2, is* $\alpha \perp \beta_k \mid Z_{d,n}, W_{d,n}$*?*

Question 1.1.6: *In the graphical model of Figure 2, is* $Z_{d,n} \perp \beta_k \mid W_{d,1:N}$*?*

1.2 Inference in a Bayesian Network

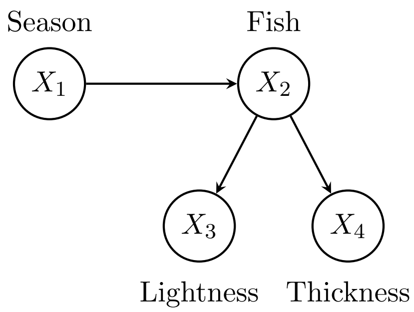

Figure 3: Fishing Bayes net.

Consider the Bayes net shown in Figure 3. Here, the nodes represent the following variables

$$X_1 \in \{\text{winter, spring, summer, autumn}\}, \qquad X_2 \in \{\text{salmon, sea bass}\} \tag{1.2.1}$$

$$X_3 \in \{\text{light, medium, dark}\}, \qquad X_4 \in \{\text{wide, thin}\} \tag{1.2.2}$$

The corresponding conditional probability tables are

$$p(x_1) = (\,.25 \;\; .25 \;\; .25 \;\; .25\,), \qquad p(x_2 \mid x_1) = \begin{pmatrix} .88 & .12 \\ .32 & .68 \\ .42 & .58 \\ .78 & .22 \end{pmatrix} \tag{1.2.3}$$

$$p(x_3 \mid x_2) = \begin{pmatrix} .33 & .34 & .33 \\ .78 & .12 & .10 \end{pmatrix}, \qquad p(x_4 \mid x_2) = \begin{pmatrix} .35 & .65 \\ .94 & .06 \end{pmatrix} \tag{1.2.4}$$

Note that in $p(x_4 \mid x_2)$, the rows represent x_2 and the columns x_4 (so each row sums to one and represents the child of the CPD). Thus $p(x_4 = \text{thin} \mid x_2 = \text{sea bass}) = 0.06$, $p(x_4 = \text{thin} \mid x_2 = \text{salmon}) = 0.65$, etc.

Answer the following queries. Show your derivations.

Question 1.2.7: *Suppose the fish was caught on December 20 — the end of autumn and the beginning of winter — and thus let*

$$p(x_1) = (\,.5,\ 0,\ 0,\ .5\,)$$

instead of the above prior. (This is called soft evidence, since we do not know the exact value of X_1, but we have a distribution over it.) Suppose the lightness has not been measured but it is known that the fish is thin. Classify the fish as salmon or sea bass. (3.5 points)

Question 1.2.8: *Suppose all we know is that the fish is thin and medium lightness. What season is it now, most likely? Use*

$$p(x_1) = (\,.25,\ .25,\ .25,\ .25\,).$$

(3.5 points)

1.3 CAVI

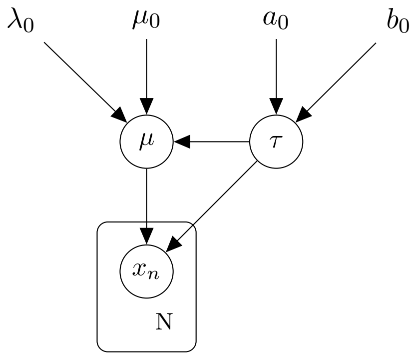

Figure 4: Bayesian network of the Normal-NormalGamma model.

Consider the model defined by Equation (10.21)-(10-23) in Bishop, for which DGM is presented in Figure 4. We are here concerned with the VI algorithm for this model covered during the lectures and in the book.

Use the template code provided on Canvas for your implementation. Solutions that have not used the template will not receive any points. You should use numpy and scipy for the implementation, do not use torch. Do not change the seeds provided in the notebook.

Question 1.3.9: *Implement a function that generates data points for the given model. Set $\mu = 1$, $\tau = 0.5$ and generate datasets with size N=10,100,1000. Plot the histogram for each of 3 datasets you generated. (1 points)*

Question 1.3.10: *Write out the expressions for the ML estimates of the variables μ and τ (without derivations). Implement a function that calculates and returns the ML estimates given data. (1 points)*

Question 1.3.11: *What is the exact posterior? (Show your derivations.) (2 points)*

Question 1.3.12: *Implement the VI algorithm for the variational distribution in Equation (10.24) in Bishop. Run the VI algorithm on the datasets. Plot the ELBO over iterations. Compare the inferred variational distribution with the exact posterior and the ML estimate. Visualize the results and discuss your findings. (6 points)*