Stat521A Spring 2009: homework 2

1 Independencies in Gaussian graphical models

1. Consider the DAG $X1 \leftarrow X2 \rightarrow X3$. Assume that all the CPDs are linear-Gaussian. Which of the following matrices *could* be the covariance matrix?

$$A = \begin{pmatrix} 9 & 3 & 1 \\ 3 & 9 & 3 \\ 1 & 3 & 9 \end{pmatrix}, B = \begin{pmatrix} 8 & -3 & 1 \\ -3 & 9 & -3 \\ 1 & -3 & 8 \end{pmatrix}, C = \begin{pmatrix} 9 & 3 & 0 \\ 3 & 9 & 3 \\ 0 & 3 & 9 \end{pmatrix}, D = \begin{pmatrix} 9 & 03 & 0 \\ -3 & 10 & -3 \\ 0 & 03 & 9 \end{pmatrix} \tag{1}$$

2. Which of the above matrices could be inverse covariance matrix?

3. Consider the DAG $X1 \rightarrow X2 \leftarrow X3$. Assume that all the CPDs are linear-Gaussian. Which of the above matrices could be the covariance matrix?

4. Which of the above matrices could be the inverse covariance matrix?

5. Let three variables x_1, x_2, x_4 have covariance matrix $\mathbf{\Sigma}_{(1:3)}$ and precision matrix $\mathbf{\Omega}_{(1:3)} = \mathbf{\Sigma}^{-1}_{(1:3)}$ as follows

$$\mathbf{\Sigma}_{(1:3)} = \begin{pmatrix} 1 & 0.5 & 0 \\ 0.5 & 1 & 0.5 \\ 0 & 0.5 & 1 \end{pmatrix}, \mathbf{\Omega}_{(1:3)} = \begin{pmatrix} 1.5 & -1 & 0.5 \\ -1 & 2 & -1 \\ 0.5 & -1 & 1.5 \end{pmatrix} \tag{2}$$

   Now focus on x_1 and x_2. Which of the following statements about their covariance matrix $\mathbf{\Sigma}_{(1:2)}$ and precision matrix $\mathbf{\Omega}_{(1:2)}$ are true?

$$A : \mathbf{\Sigma}_{(1:2)} = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}, \; B : \mathbf{\Omega}_{(1:2)} = \begin{pmatrix} 1.5 & -1 \\ -1 & 2 \end{pmatrix} \tag{3}$$

2 Two filter approach to smoothing in HMMs

Assuming that $\Pi_t(i) = p(S_t = i) > 0$ for all i and t, derive a recursive algorithm for updating $r_t(i) = p(S_t = i|\mathbf{x}_{t+1:T})$. Hint: it should be very similar to the standard forwards algorithm, but using a time-reversed transition matrix. Then show how to compute the posterior marginals $\gamma_t(i) = p(S_t = i|\mathbf{x}_{1:T})$ from the backwards filtered messages $r_t(i)$, the forwards filtered messages $\alpha_t(i)$, and the stationary distribution $\Pi_t(i)$.

3 E step for HMMs with mixture of Gaussian observations

Let $\hat{\alpha}_t(j) = p(S_t = j, \mathbf{x}_{1:t})$ be the unnormalized filtering distributions in the FB algorithm. In Rabiner's HMM tutorial **?**, he gives the following equation (p267, just after his equation 54) for $\gamma_t(j,k) = p(S_t = k, M_t = k|\mathbf{x}_{1:T})$:

$$\gamma_t(j,k) = \left[\frac{\hat{\alpha}_t(j)\beta_t(j)}{\sum_{j'} \hat{\alpha}_t(j')\beta_t(j')} \right] \left[\frac{W_{jk}\mathcal{N}(\mathbf{x}_t|\boldsymbol{\mu}_{j,k}, \mathbf{\Sigma}_{j,k})}{\sum_m W_{jm}\mathcal{N}(\mathbf{x}_t|\boldsymbol{\mu}_{j,m}, \mathbf{\Sigma}_{j,m})} \right] \tag{4}$$

Prove that this is equivalent to the following (Equation 1.127 in the HMM handout from my book, dated 29Jan09)

$$\gamma_t(j,k) \quad = \quad \frac{a_t(j)\beta_t(j)W_{j,k}\mathcal{N}(\mathbf{x}_t|\boldsymbol{\mu}_{j,k}, \mathbf{\Sigma}_{j,k})}{p(\mathbf{x}_{1:T})} \tag{5}$$

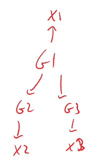

Figure 1: A simple DAG representing inherited diseases.

4 Message passing on a tree

Consider the DGM in Figure 1 which represents the following fictitious biological model. Each G_i represents the genotype of a person: $G_i = 1$ if they have a healthy gene and $G_i = 2$ if they have an unhealthy gene. G_2 and G_3 may inherit the unhealthy gene from their parent G_1. $X_i \in \mathbb{R}$ is a continuous measure of blood pressure, which is low if you are healthy and high if you are unhealthy. We define the CPDs as follows

$$p(G_1) = [0.5, 0.5] \tag{6}$$

$$p(G_2|G_1) = \begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix} \tag{7}$$

$$p(G_3|G_1) = \begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix} \tag{8}$$

$$p(X_i|G_i = 1) = \mathcal{N}(X_i|\mu = 50, \sigma^2 = 10) \tag{9}$$

$$p(X_i|G_i = 2) = \mathcal{N}(X_i|\mu = 60, \sigma^2 = 10) \tag{10}$$

The meaning of the matrix for $p(G_2|G_1)$ is that $p(G_2 = 1|G_1 = 1) = 0.9$, $p(G_2 = 1|G_1 = 2) = 0.1$, etc.

1. Suppose you observe $X_2 = 50$, $X_3 = 50$, and X_1 is unobserved. What is the message that G_2 sends up to G_1? What is the message that G_3 sends up to G_1?
2. What is the posterior belief on G_1, i.e., $p(G_1|X_2 = 50, X_3 = 50)$?
3. Now suppose you observe $X_2 = 50$, but do not observe X_3. (And again you do not observe X_1.) What is $p(G_1|X_2)$? Explain your answer intuitively.
4. Now suppose $X_2 = 60$, $X_3 = 60$. What is $p(G_1|X_2, X_3)$? Explain your answer intuitively.
5. Now suppose $X_2 = 50$, $X_3 = 60$. What is $p(G_1|X_1, X_2)$? Explain your answer intuitively.

5 Multi-dice dishonest casino (Matlab)

Download PMTK from `http://www.cs.ubc.ca/~murphyk/pmtk/`. (The latest version is 1.3, 29 Jan 09.) Make sure you understand the `casinoDemo` in the `examples/hmmDistExamples` directory.

Now consider the following generalization of the dishonest casino. When the casino is using a fair dice, each face shows up with probability 1/6. However, it has 6 different kinds of loaded dice: type t puts probability 5/10 on face t and probability 1/10 on the other faces. Suppose the casino uses the fair dice for on average 20 rolls, and uses an unfair dice for on average 10 rolls. Your task is to figure out which type of unfair dice the casino is using: it might be using just a single type, or it might using be a uniform mixture of different unfair dice. That is, the data will appear as coming from a mixture of uniform distribution and one or more skewed distributions.

Figure out how to create an HMM which can determine which kind of unfair dice (one or more) the casino is using. (You might think of other, simpler ways of estimating this; You are encouraged to try these.) Apply your method to the following datasets: `casinoDataA500.mat`, `casinoDataB500.mat`, `casinoDataC500.mat`. In each case, estimate which unfair dice is/are being used to generate the data.