IERG5130 Homework 1

Due: Oct 19, 2018

Problem A: Conditional independence [10pt]

Consider the following joint distribution:

$$p(A, B, C) = p(A)p(B)p(C|A, B),$$

where all variables are binary variables that take values from $\{0, 1\}$. Particularly, we have $p(A = 1) = a$, $p(B = 1) = b$, and

$$p(C = 1|A, B) = \begin{cases} c_{00} & (A = 0, B = 0) \\ c_{01} & (A = 0, B = 1) \\ c_{10} & (A = 1, B = 0) \\ c_{11} & (A = 1, B = 1) \end{cases}.$$

Here, we assume that $a, b, c_{00}, c_{01}, c_{10}, c_{11} \in (0, 1)$, namely all probability values are positive.

Questions:

(1) [1pt] Please compute the marginal distribution $p(C)$.

(2) [2pt] Please compute $p(A, B)$ and show that $A \perp B$.

(3) [3pt] Let $a = b = 0.5$, please derive the setting of c (for example the equalities that $c_{00}, c_{01}, c_{10}, c_{11}$ have to satisfy), under which we have $A \perp B \mid C$.

(4) [2pt] Please derive the setting of a, b, c, under which we have $C \perp A$.

(5) [2pt] Please derive the setting of a, b, c, under which we have $C \perp (A, B)$.

Problem B: Bayesian Network [20pt]

Consider the following joint distribution:

$$p(A, B, C, D, E) = p(A)p(B|A)p(C|A)p(D|B, C)p(E|D).$$

Questions:

(1) [2pt] Please draw the graphical representation of this joint distribution.

(2) [6pt] Please determine which of the following statments are true:

(a) $A \perp D$
(b) $A \perp D \mid B$
(c) $A \perp D \mid (B, C)$
(d) $A \perp E \mid D$

(e) $B \perp C \mid A$

(f) $B \perp C \mid (A, D)$

(3) [2pt] Please derive the set of local independencies $\mathcal{I}_l(G)$.

(4) [2pt] Please derive the set of pairwise independencies $\mathcal{I}_p(G)$.

(5) [3pt] Please derive the set of global independencies $\mathcal{I}_g(G)$.

(6) [2pt] Please derive the moralized graph $\mathcal{M}[G]$.

(7) [3pt] Please derive the set of global independencies for the moralized graph, namely $\mathcal{I}_g(\mathcal{M}[G])$.

Problem C: Markov Random Field [15pt]

Consider the following joint distribution:

$$p_\alpha(A, B, C, D) = \frac{1}{Z}\phi_\alpha(A, B)\phi_\alpha(A, C)\phi_\alpha(B, D)\phi_\alpha(C, D),$$

where A, B, C, D are all binary variables, and the factor ϕ_α is defined to be

$$\phi_\alpha(X, Y) = \begin{cases} \alpha & (X = Y) \\ 1 & (X \neq Y) \end{cases}.$$

Questions:

(1) [2pt] Please draw the graphical representation of the joint distribution.

(2) [2pt] Please draw a factor graph for this model.

(3) [2pt] Please derive the set of global independencies $\mathcal{I}_g(G)$.

(4) [2pt] Please derive the setting of α, under which all variables are mutually independent.

(5) [2pt] Please compute the normalizing constant Z, express it as a function of α.

(6) [3pt] Let $q = \Pr(A = B = C = D)$. Please compute q as a function of α.

(7) [2pt] Please compute the value of α_0 and α_1, such that $q(\alpha_0) = 0.1$ and $q(\alpha_1) = 0.9$.

Problem D: Convex Analysis [20pt]

A lot of mathematical analysis in this course rely on basic concepts in convex analysis. This problem focus on this area.

Questions:

(1) [3pt] Please show that the ball $B(\mathbf{a}, r) \subset \mathbb{R}^d$ given as below is a convex set.

$$B(\mathbf{a}, r) = \{\mathbf{x} : \|\mathbf{x} - \mathbf{a}\| \leq r\}, \quad \text{with } r > 0.$$

(2) [3pt] Please show that a half-space given as $H = \{\mathbf{x} : \mathbf{a}^T\mathbf{x} >= b\}$ is a convex set.

(3) [3pt] Please show that $f(x) = ax^2 + bx + c$ is a convex function if and only if $a \geq 0$.

(4) [3pt] Please show that $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{a}^T\mathbf{x}$ is a convex function if and only if $\mathbf{Q}$ is positive semi-definite.

(5) [4pt] Let $\{f_\theta : \theta \in \Omega\}$ be a set of convex functions. Please show that the point-wise supreme g given as below is a convex function.

$$g(\mathbf{x}) := \sup_{\theta \in \Omega} f_\theta(\mathbf{x}).$$

(6) [4pt] Let f be a real-valued convex function defined on $\mathbb{R}^d$, and p be a distribution over $\mathbb{R}^d$. Please show that

$$f(E_p[X]) \leq E_p[f(X)].$$

Problem E: Exponential Family & Multivariate Gaussian [5pt]

A multivariate Gaussian distribution is a distribution over a vector space $\mathbb{R}^d$, with a probability density function given below:

$$p_{(\boldsymbol{\mu}, \boldsymbol{\Sigma})}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^d |\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right).$$

We usually use $X \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ to denote that a random vector X follows a multivariate Gaussian distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$.

Question:

(1) [3pt] Please show that the family of multivariate Gaussian distributions is an exponential family. Please give the canonical parameters and sufficient statistics.

(2) [2pt] Consider the exponential family derived for the question above. Is the representation overcomplete or minimal? Please justify your answer.

Problem F: Formulation of Bayesian Network [15pt]

Please formulate a dynamic Bayesian network to model a soccer game. The observation is a set of players in both teams. For the i-th player, we can observe its position at each time step t, denoted by $X_t^{(i)}$, the velocity $V_t^{(i)}$, and the type of action $a_t^{(i)}$ (for example, running, standing, kicking a ball, etc). Your formulation should (at least) take into account the following aspects:

1. Each player will actively make/adjust the choice of strategy.
2. The overall situation will affect the strategy choices. It is up to you to decide how to formalize the "overall situation" into random variables. But you have to justify your approach.
3. Of course, the actions of the players change not only their positions and velocities, but also the situation.

Write down the formulation in both a joint distribution formula and a graphical representation.

Problem G: Formulation of Markov Random Field [15pt]

Please formulate a Markov random field to model multiple objects (for example, chairs, desks, etc) in a room and their relationships. Here, the observation comprises a set of objects, where each object is associated with several observable attributes, including its position X_i, its orientation Y_i.

Your formulation should (at least) take into account the following aspects:

1. The configuration of the objects depends on the function (for example, living room, bedroom) as well as the size of the room.

2. The placement and the orientation of an object may depend on that of others.

3. You should consider not only the objects on the ground, but also those hang on the wall or the ceil.

Write down the formulation in both a joint distribution formula, and specify the factors.