CS 391L: Machine Learning **Summer 2023**

Homework 5 - Theory

Lecture: Prof. Qiang Liu

1. **Gaussian Multivariate**

Assume we have a multivariate normal random variable $X = [X_1, X_2, X_3, X_4]^\top$, whose covariance matrix Σ and inverse covariance matrix Q are

$$\Sigma = \begin{bmatrix} 0.71 & -0.43 & 0.43 & 0 \\ -0.43 & 0.46 & -0.26 & 0 \\ 0.43 & -0.26 & 0.46 & 0 \\ 0 & 0 & 0 & 0.2 \end{bmatrix} \qquad Q = \begin{bmatrix} 5 & 3 & -3 & 0 \\ 3 & 5 & 0 & 0 \\ -3 & 0 & 5 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix}.$$

Note that Q is simply the inverse of Σ, i.e., $Q = \Sigma^{-1}$.

(a) [**5 points**] Are X_3 and X_4 correlated?

(b) [**5 points**] Are X_3 and X_4 conditionally correlated given the other variables? That is, does $\mathrm{cov}(X_3, X_4 \mid X_1, X_2)$ equal to zero?

(c) [**5 points**] Please find the Markov blanket of X_2. Recall that the Markov blanket of X_i is the set of variables (denoted by X_{M_i}), such that

$$X_i \perp X_{\neg\{i\}\cup M_i} \mid X_{M_i},$$

where $\neg\{i\} \cup M_i$ denotes all the variables outside of $\{i\} \cup M_i$.

(d) [**5 points**] Assume that $Y = [Y_1, Y_2]^\top$ is defined by

$$Y_1 = X_1 + X_4$$
$$Y_2 = X_2 - X_4.$$

Please calculate the covariance matrix of Y.

2. **Expectation Maximization (EM)**

Assume we have a dataset of two points $\{x^{(1)},\ x^{(2)}\}$:

$$x^{(1)} = -1, \qquad\qquad x^{(2)} = 1.$$

Assume $x^{(i)}$ is drawn **i.i.d.** from a simple mixture distribution of two Gaussian components:

$$f(x \mid \mu_1,\ \mu_2) = \frac{1}{2}\phi(x \mid \mu_1,\ 1) + \frac{1}{2}\phi(x \mid \mu_2,\ 1),$$

where $\phi(\cdot \mid \mu_i, 1)$ denotes the probability density function of Gaussian distribution $\mathcal{N}(\mu_i, 1)$ with mean μ_i and unit variance. We want to estimate the unknown parameters μ_1 and μ_2.

(a) **[5 points]** Assume we run EM starting from an initialization of $\mu_1 = -2$ and $\mu_2 = 2$. Please decide the value of μ_1 and μ_2 at the next iteration of EM algorithm. (You may find it handy to know that $1/(1 + \exp(-4)) \approx 0.98$).

(b) **[5 points]** Do you think EM (when initialized with $\mu_1 = -2$ and $\mu_2 = 2$) will eventually converge to $\mu_1 = -1$ and $\mu_2 = 1$ (i.e., coinciding with the two data points). Please justify your answer using either your theoretical understanding or the result of an empirical simulation.

(c) **[5 points]** Please decide the fixed point of EM when we initialize it from $\mu_1 = \mu_2 = 2$.

(d) **[5 points]** Please decide the fixed point of **K-means** when we initialize it from $\mu_1 = -2$ and $\mu_2 = 2$.