Tutoring Session 4

Multivariate Normal

Problem 1: Show that the sum of two independent Gaussian random variables ($\boldsymbol{X_1}$ and $\boldsymbol{X_2}$) is Gaussian. Some of the properties of Gaussians mentioned in the lecture can help.

Problem 2: Let $p(x) = \mathcal{N}(\mu_1, \sigma_1^2)$ and $q(x) = \mathcal{N}(\mu_2, \sigma_2^2)$. Show that the Kullback-Leibler divergence of q from p is $KL(p,q) = \log\frac{\sigma_2}{\sigma_1} + \frac{\sigma_1^2 + (\mu_1-\mu_2)^2}{2\sigma_2^2} - \frac{1}{2}$.

Problem 3: We can sample from any multivariate Gaussian by using:

$$\boldsymbol{X} = \boldsymbol{\mu} + \boldsymbol{L}\boldsymbol{Z} \Rightarrow \boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}),$$

with $\boldsymbol{L}\boldsymbol{L}^T = \boldsymbol{\Sigma}$ and $\boldsymbol{Z} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$. Show that this works by using the change of variable theorem.

Problem 4: The unbiased estimates for the covariance of a d-dimensional Gaussian based on n samples is given by

$$\hat{\boldsymbol{\Sigma}} = \boldsymbol{C}_n = \frac{1}{n-1}\sum_{i=1}^{n}(\boldsymbol{x} - \boldsymbol{\mu}_n)(\boldsymbol{x} - \boldsymbol{\mu}_n)^T$$

It is clear that it takes $O(nd^2)$ time to compute $\boldsymbol{C}_n$. If the data points arrive one at a time, it is more efficient to incrementally update these estimates than to recompute from scratch.

Show that the covariance can be sequentially udpated as follows

$$\boldsymbol{C}_{n+1} = \frac{n-1}{n}\boldsymbol{C_n} + \frac{1}{n+1}(\boldsymbol{x}_{n+1} - \boldsymbol{\mu}_n)(\boldsymbol{x}_{n+1} - \boldsymbol{\mu}_n)^T$$

Problem 5: We consider a partitioning of the components of x into three groups $\boldsymbol{x}_a$, $\boldsymbol{x}_b$, and $\boldsymbol{x}_c$, with a corresponding partitioning of the mean vector $\boldsymbol{\mu}$ and of the covariance matrix $\boldsymbol{\Sigma}$ in the form

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_a \\ \boldsymbol{\mu}_b \\ \boldsymbol{\mu}_c \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{aa} & \boldsymbol{\Sigma}_{ab} & \boldsymbol{\Sigma}_{ac} \\ \boldsymbol{\Sigma}_{ba} & \boldsymbol{\Sigma}_{bb} & \boldsymbol{\Sigma}_{bc} \\ \boldsymbol{\Sigma}_{ca} & \boldsymbol{\Sigma}_{cb} & \boldsymbol{\Sigma}_{cc} \end{pmatrix}$$

Find an expression for the conditional distribution $p(\boldsymbol{x}_a|\boldsymbol{x}_b)$ in which $\boldsymbol{x}_c$ has been marginalized out.

Problem 6: A very useful result from linear algebra is the Woodbury matrix inversion formula given by

$$(\boldsymbol{A} + \boldsymbol{B}\boldsymbol{C}\boldsymbol{D})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{C}^{-1} + \boldsymbol{D}\boldsymbol{A}^{-1}\boldsymbol{B})^{-1}\boldsymbol{D}\boldsymbol{A}^{-1}$$

By multiplying both sides by $(\boldsymbol{A} + \boldsymbol{B}\boldsymbol{C}\boldsymbol{D})$ prove the correctness of this result.