Machine Learning Assignment # 3
Universität Bern

Due date: 09/10/2019

Late submissions will incur a penalty. Submit your answers in ILIAS (as a pdf or as a picture of your written notes if the handwriting is very clear). Submission instructions will be provided via email. You are not allowed to work with others.

Probability theory review [Total 100 points]

Solve each of the following problems and show all the steps of your working.

1. Show that the covariance matrix is always symmetric and positive semidefinite. **[15 points]**

2. $X \in R^n$ and $Y \in R^m$ are independent random variables. Their expectations and covariances are $E[X] = 0$, $\text{Cov}[X] = I$, $E[Y] = \mu$ and $\text{Cov}[Y] = \sigma I$, where I is the identity matrix of the appropriate size and σ is a scalar. What is the expectation and covariance of the random variable $Z = AX + Y$, where $A \in R^{m \times n}$? **[20 points]**

3. Thomas and Viktor are friends. It is Friday night and Thomas does not have a phone. Viktor knows that there is a $2/3$ probability that Thomas goes to the party to downtown. There are 5 pubs in downtown and there is an equal probability of Thomas going to any of them if he goes to the party. Viktor already looked for Thomas in 4 of the bars.
What is the probability of Viktor finding Thomas in the last bar? **[15 points]**

4. Derive the mean for the Beta Distribution, which is defined as **[20 points]**

$$\text{Beta}(x|a, b) = \frac{1}{B(a, b)} x^{a-1}(1-x)^{b-1} \tag{1}$$

where $B(a, b)$, $\Gamma(a)$ are Beta and Gamma functions respectively:

$$B(a, b) \triangleq \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \tag{2}$$

and

$$\Gamma(x) \triangleq \int_0^\infty u^{x-1} e^{-u} du. \tag{3}$$

Hint: *Use integration by parts.*

5. Let $A \in \mathbb{R}^{n \times n}$ be a positive definite square matrix, $b \in \mathbb{R}^n$, and c be a scalar. Prove that **[20 points]**

$$\int_{x \in \mathbb{R}^n} e^{-\frac{1}{2}x^T A x - x^T b - c} dx = \frac{(2\pi)^{n/2}|A|^{-1/2}}{e^{c - \frac{1}{2}b^T A^{-1} b}}.$$

Hint: *Use the fact that the integral of the Gaussian probability density function of a random variable with mean μ and covariance Σ is 1.*

6. From the definition of conditional probability of multiple random variables, show that **[10 points]**

$$f(x_1, x_2, ..., x_n) = f(x_1) \prod_{i=2}^{n} f(x_i | x_1, ..., x_{i-1})$$

where $x_1, \ldots, x_n$ are random variables and f is a probability density function of its arguments.