Homework 2: Probability and Information Theory

Dr. Benjamin Roth
Profilierungsmodul II

Due: Tuesday October 29, 2019 18:00

Please submit your solution to `[profilmodul1920 at cis dot uni-muenchen dot de]`
Please submit individually for this homework

Exercise 1: Probability [5 points]

Read Chapter 3 of the Deep Learning (DL) book (`www.deeplearningbook.org/`). **Note down and submit at least two questions that are still unclear after reading the chapter, and that you could not resolve easily** (e.g. by re-reading the paragraph or by consulting Wikipedia). (Only) If you don't have any questions, please answer the following:

- How can a n-dimensional multivariate normal distribution with mean vector μ and diagonal precision matrix β be written in terms of univariate normal distributions (and parameters mean μ_i and variance σ_i^2)?
- Given two discrete distributions $P(x) = \{\frac{1}{3}, \frac{1}{3}, \frac{1}{3} - \epsilon, \epsilon\}$ and $Q(x) = \{\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\}$, which of the following is bigger for small values of ϵ: $D_{KL}(P||Q)$ or $D_{KL}(Q||P)$?
- Which of the following is true, which is not, and why? (rationale, example or counter-example)
  - independence implies zero covariance
  - zero covariance implies independence
  - none implies the other
- What is the result of the following expectation: $E[X_1 + X_2 + \cdots + X_n]$, where all X_i are independent Bernoulli-distributed random variables and $P(X_i) = \phi$
- Does a probability density function $p(x)$ exist with $p(x) > 1$ for some state (value) x? If yes, give an example of such a distribution and corresponding state x. If no, why not (definition and rationale)?

- Does a probability density function $p(x)$ exist with $p(x) < 0$ for some state (value) x? If yes, give an example of such a distribution and corresponding state x. If no, why not (definition and rationale)?