Wendelin Böhmer <j.w.bohmer@tudelft.nl>

Math and machine learning primer

Voluntary exercises

The following exercises do not have to be submitted as homework, but might be helpful to practice the required math and prepare for the exam. Some questions are from old exams and contain the used rubrik. You do not have to submit these questions and will not receive points for them. Solutions are available on Brightspace.

E1.1: Taylor expansion **(voluntary)**

For the function $\sqrt{1+x}$, write down the Taylor series around $x_0 = 0$ up to 3rd order.

E1.2: Critical points **(voluntary)**

Consider the two functions

$$f(x, y) := c + x^2 + y^2$$
$$g(x, y) := c + x^2 - y^2,$$

where $c \in \mathbb{R}$ is a constant.

(a) Show that $\boldsymbol{a} = (0, 0)$ is a critical point of both functions.

(b) Check for f and for g whether $\boldsymbol{a}$ is a minimum, maximum, or saddlepoint using the Hessian matrix.

Hint: A matrix is positive (negative) definite if and only if all its eigenvalues are positive (negative).

E1.3: Distributions and expected values **(voluntary)**

Let $x \in \mathbb{R}$ be a random variable with probability density $p : \mathbb{R} \to \mathbb{R}$ with:

$$p(x) = \begin{cases} c \cdot \sin(x), & x \in [0, \pi] \\ 0, & \text{elsewhere} \end{cases}$$

(a) Determine the parameter $c \in \mathbb{R}$ such that $p(x)$ is indeed a probability density.

(b) Determine the expected value $\mu := \mathbb{E}_p[x]$

(c) Determine the variance of x, $\mathbb{E}_p[(x - \mu)^2]$.

E1.4: Variance of the empirical mean (old exam question) (voluntary)

Prove that the variance of the empirical mean $f_n := \frac{1}{n}\sum_{i=1}^n x_i$, based on n samples $x_i \in \mathbb{R}$ drawn i.i.d. from the Gaussian distribution $\mathcal{N}(\mu, \sigma^2)$, is $\mathbb{V}[f_n] = \frac{\sigma^2}{n}$, *without* using the fact the variance of a sum of independent variables is the sum of the variables' variances.

E1.5: Unbiased variance estimate (voluntary)

Let $\{x_i\}_{i=1}^n$ be a data set that is drawn i.i.d. from the Gaussian distribution $x_i \sim \mathcal{N}(\mu, \sigma^2)$. Let further $\hat\mu := \frac{1}{n}\sum_{i=1}^n x_i$ denote the empirical mean and $\hat\sigma^2 := \frac{1}{n}\sum_{i=1}^n (x_i - \hat\mu)^2$ the equivalent empirical variance. Prove analytically that $\hat\mu$ is unbiased, i.e. $\mathbb{E}[\hat\mu] = \mu$, and that $\hat\sigma^2$ is biased, i.e. $\mathbb{E}[\hat\sigma^2] \neq \sigma^2$.

Bonus-question: Can you derive an unbiased estimator for the empirical variance?

Hint: If x_i and x_j are drawn i.i.d. from $\mathcal{N}(\mu, \sigma^2)$, then holds $\forall i$:

$$\mathbb{E}[x_i] = \mu\,, \qquad \mathbb{E}[(x_i - \mu)^2] = \sigma^2 \qquad \text{and} \qquad \mathbb{E}[(x_i - \mu)(x_j - \mu)] = 0 \quad \text{if} \quad i \neq j\,.$$

E1.6: Maximum dice (voluntary)

This question is designed to practice the use of Kronecker-delta functions and become more familiar with (discrete) probabilities. You are given 3 dice, a D6, a D8 and a D10, where Dx refers to a x-sided fair dice, where each of the x sides is numbered uniquely 1 to x and rolled with the exact same probability.

(a) Prove analytically that the probability that the D6 is among the highest (including equal) numbers when all 3 dice are rolled together is roughly $\rho \approx 19\%$.

(b) Prove analytically that the probability that the D8 rolls among the highest is $\rho' \approx 38\%$.

(c) Prove analytically that the probability that the D10 rolls among the highest is $\rho'' \approx 58\%$.

Hint: You can solve the question however you want, but you are encouraged to use Kronecker-deltas, e.g. $\delta(i > 5)$ is 1 if $i > 5$ and 0 otherwise. You will find that this can simplify complex sums enormously. If you do so, you can use the equalities $\sum_{i=1}^n i \overset{(1)}{=} \frac{n^2+n}{2}$ and $\sum_{i=1}^n i^2 \overset{(2)}{=} \frac{n(n+1)(2n+1)}{6}$.

Bonus-question: Why don't the above numbers sum up to 1?