Maths Quiz for Machine Learning Fall 2023

The Swiss AI Lab IDSIA,
University of Applied Sciences and Arts of Southern Switzerland (SUPSI),
University of Lugano(USI)

Based on material for GI18, Gatsby Unit, University College London

Originally collected by Louis Kirsch

The purpose of this quiz is to help you and us assess any areas of background knowledge you may need to work on. **This quiz is not graded, and you do not need to turn it in**, but if you are not sure of any of your answers please speak with one of the TAs or instructors. If you cannot solve the majority of the questions, this course is probably not appropriate for you.

1. Compute the derivatives of the following functions w.r.t. x.
   - $f : \mathbb{R} \to \mathbb{R}, f(x) = 3x^2 + 2x$
   - $f : \mathbb{R} \to \mathbb{R}, f(x) = \log(\cosh(x))$
   - $f : \mathbb{R}^3 \to \mathbb{R}, f(\mathbf{x}) = \sin(x_1^2) + 3x_2^3 + e^{x_3}$, where $\mathbf{x} = (x_1, x_2, x_3)^T$
   - $f : \mathbb{R}^n \to \mathbb{R}, f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$, where A is a $n \times n$ matrix.

   What is the dimensionality of these derivatives?

2. Show that for small x
$$\log(1 + x) \approx x - \frac{x^2}{2}.$$
   [Hint: consider using a Taylor series expansion]

3. Consider two discrete random variables X and Y. Show that
$$P(X = x, Y = y) \leq P(X = x)$$

4. This is a Gaussian density with mean μ and variance σ^2:
$$p(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right].$$
   What is the point x^* with the highest probability density? Back up your claim by finding the stationary point(s) of the density function.

5. Let $\vec{x}$, $\vec{y}$ and $\vec{z}$ be $n \times 1$ column vectors and A be an $n \times n$ matrix. Assume that you know that $\vec{x}^T A\vec{y} = 1$, where the superscript ^{T} denotes the transpose operator. For each of the following statements answer **true**, **false**, or **dimensionally inconsistent** (if any part of the equation is undefined in terms of matrix algebra).
   $\vec{x}^T A\vec{y} = 1$ implies that:

   (a) $\vec{y}^T A\vec{x} = 1$

   (b) $A\vec{x}^T\vec{y} = 1$

   (c) $\vec{y}^T A^{-1}\vec{x} = 1$

   (d) $\vec{y}^T A^T\vec{x} = 1$

   (e) $\vec{x}^T A(\vec{y} - \vec{z}) = 1 - \vec{x}^T A\vec{z}$

   (f) $\mathsf{Trace}(A\vec{y}\vec{x}^T) = 1$

6. Let P be a plane in three-dimensional space which goes through the points (1,2,4), (2,6,3), (5,1,0). Write down its equation

   (a) in parametric form

   (b) in normal form.

7. The weather in London has probability p of being rainy on any given day. Assume the weather is independent across days. Let $X_i = 1$ if it is rainy on day i, and $X_i = 0$ if it is not rainy. Using the notation $E(X_i)$ to mean the "expected value of X_i" and $\mathrm{Var}(X_i)$ to mean the "variance of X_i", compute the values (in terms of p) of the following quantities:

   (a) $E(X_i)$

   (b) $E(X_i^2)$

   (c) $\mathrm{Var}(X_i)$

   (d) $E(\sum_{i=1}^{n} X_i)$

8. What is $\displaystyle\int_{\mu}^{\infty} e^{-(x-\mu)^2/2\sigma^2}\mathrm{d}x$?

9. Given a list of numbers $x_1, x_2, \ldots x_n$ as input, write down Python code that will output the largest number in the list (obviously, don't rely on the `max` function!).