Homework 1: Due November 1 in class

Reading: Read chapter 1 of Lecture notes.

1. Let $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ and define $\phi\colon \mathbb{R}^n \to \mathbb{R}^m$ by $\phi(x) = Ax - b$.

   Prove that for every $i \in \{1, 2, \ldots, n\}$ one has that
$$\sum_{j=1}^{m} \frac{\partial \phi_j}{\partial x_i}(x)\phi_j(x) = a_i^T(Ax - b).$$

2. Let $A \in \mathbb{R}^{m\times n}$, where $m \le n$, and assume that $r(A) = m$.

   Prove that $A^\dagger = A^T(AA^T)^{-1}$.

3. Let $A \in \mathbb{R}^{m\times n}$.

   Prove that, if $A = U\Sigma V^T$ is a singular value decomposition for A, then $A^\dagger = V\Sigma^{-1}U^T$.

4. (a) Compute by hand a singular value decomposition and the pseudoinverse of $A = \begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \\ -2 & 1 \end{bmatrix}$.

   (b) Now try to do the same using Julia. Do you get what you expected? What happens if you compare the pseudoinverse obtained via the command `pinv` to the one obtained by taking $V\Sigma^{-1}U^T$? Produce a jupyter notebook documenting your work, including your comments on the behavior above.

5. Let $X \sim N(\mu, \sigma^2)$ for $\mu \in \mathbb{R}$ and $\sigma^2 > 0$. Prove that $\mathbb{E}X = \mu$ and $\mathrm{Var}(X) = \sigma^2$.

   Hint: it might be useful to recall that $\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\,dx = \sqrt{2\pi}$.

6. Let X and Y be two real random variables that are either:

   both discrete; both continuous, have respective densities f_X, f_Y and finite expected values, i.e., $\mathbb{E}(X), \mathbb{E}(Y) < \infty$.

   Prove that for all $a, b \in \mathbb{R}$ one has that $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$.

   Hint: use the transformation law (Lemma 1.25 in the notes) for $g(X, Y) = X + Y$ and that for every random variable $\mathbb{E}|X| < \infty$ if and only if $\mathbb{E}X < \infty$ (see Eq. (3.1.7) in Ash's book). The concept of marginal density might also be useful.

7. Let $\Omega := \{x_1, \ldots, x_n\}$ and $p_1, \ldots, p_n \ge 0$ with $p_1 + \cdots + p_n = 1$. Prove that the following algorithm generates a random variable $X \in \Omega$ with $P(X = x_i) = p_i$:

   define the numbers $w_k := \sum_{i=1}^{k} p_i$, $1 \le k \le n$, and $w_0 := 0$; draw $Y \sim \mathrm{Unif}([0, 1])$ (for instance, in `Julia` one can draw Y using the command `rand()`); let k be such that $w_{k-1} \le Y < w_k$; return x_k.

8. The element caesium-137 has a half-life of about 30,17 years. In other words, a single atom of caesium-137 has a 50 percent chance of surviving after 30,17 years, a 25 percent chance of surviving after 60,34 years, and so on.

   (a) Determine the probability that a single atom of caesium-137 decays (i.e., does not survive) after a single day. How would you model the random variable X that takes the value 1 when the atom decays and 0 otherwise?

   (b) Using `Julia`, simulate 1000 times the behaviour of a collection C of 10^6 caesium-137 atoms in a single day. How would you model the random variable $Y = |\{\text{atoms in } C \text{ decaying after a single day}\}|$?

   (c) The Poisson distribution with parameter λ is a discrete probability distribution that is used to "model rare events". When $Z \sim Pois(\lambda)$, one has that $P(Z = k) = \dfrac{\lambda^k e^{-\lambda}}{k!}$. Plot the Poisson distribution with $\lambda = 10^6 \cdot p$, where p is the probability computed in part (a).

   (d) Compare the empirical distribution in part (b) to the theoretical distribution in part (c).

   Some `Julia` packages that might be useful: `Distributions`, `StatsPlots`.