Preliminary problems

Deadline: 20/02/2021, 23:59

The purpose of these tasks is to give you an impression of the maths we are going to use in this module; to introduce some basic notation; and to give you joy of solving easy exercises.

1. Given $\alpha, w, x \in \mathbb{R}^k$ column vectors and $W \in \mathbb{R}^{n\times k}$ matrix, prove the following identities:
   a) $\nabla_\alpha(w^T\alpha) = w^T$
   b) $\nabla_x(Wx) = W$
   c) For $k = n$, $\nabla_x(x^TWx) = (W^T + W)x$

   (Note that for a function $f : \mathbb{R}^k \to \mathbb{R}$, ∇f is the row vector $\nabla f = [\partial_1 f, \ldots, \partial_k f]$, and for a function $g : \mathbb{R}^n \to \mathbb{R}^m$, ∇g is the matrix whose element in row i and column j is $\partial_j(g_i)$.)

2. Let $\xi_i \in \{0; 1\}$ $(i = 1, \ldots, N)$ independent, indentically distributed random variables with $\Pr(\xi_i = 1) = \mu$, and let $\nu = \frac{1}{N}\sum_{i=0}^{N} \xi_i$. What are the values of the following expressions?
   a) $\mathbb{E}\{\nu\}$
   b) $\operatorname{Var}\nu$
   c) $\mathbb{E}\{(\nu - \mu)^2\}$

3. Let $\pi(\theta)$ be a (one-dimensional) probability distribution for some parameter θ and $R : \mathbb{R} \to \mathbb{R}$ some function. Prove that

$$\frac{d}{d\theta}\Big(\mathbb{E}_{a\sim\pi(\theta)}R(a)\Big) = \mathbb{E}_{a\sim\pi(\theta)}\frac{d}{d\theta}\log\pi(a \mid \theta)R(a), \tag{1}$$

   where $\pi(a \mid \theta)$ is the likelihood of a in the distribution $\pi(\theta)$. ($\mathbb{E}_{a\sim\pi(\theta)}$ denotes the expected value of the random variable a that is distributed according to $\pi(\theta)$.)

4. In the Python file attached, for each function, write an equivalent implementation using only Numpy (vector or linear algebraic) operations. (Ie. eliminate all the loops.)

   There is a tester function at the end (called `main`), with which you can check your solution. (`np.vectorize` is cheating.) Only the first three tasks are mandatory, the last two are optional.