Home assignment № 2

Since in this assignment both formulas and code are required, then please submit your solution via the single Jupyter Notebook file, where analytical solutions are written in Markdown/ LaTeX format.

In the subproblems presented below derive the analytic expression of the requested objects. After that, implement Python functions to compute considered mathematical functions with JAX[1] and compare speed of gradient computations by autodiff approach and by implementation of analytic expressions that you derived. Jupyter Notebook with such comparison should be also included in your submission. Consider different dimensions and conclude what approach (autodiff vs. analytic expression) is asymptotically faster.

Problem 1

1. (1 + 2 pts) Compute the gradients with respect to $\mathbf{U} \in \mathbb{R}^{n\times k}$ and $\mathbf{V} \in \mathbb{R}^{k\times n}$, $k < n$ of function $J(\mathbf{U}, \mathbf{V}) = \|\mathbf{UV} - \mathbf{Y}\|_F^2 + \frac{\lambda}{2}(\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2)$.
2. (1 + 2 pts) Compute the gradient and hessian of the following function $f(\mathbf{w}) = \sum\limits_{i=1}^{m} \log(1 + e^{-y_i \mathbf{w}^\top \mathbf{x}_i}) + \frac{1}{2}\|\mathbf{w}\|_2^2$, where $\mathbf{x}_i \in \mathbb{R}^n$, $y_i \in \mathbb{R}$.
3. (0.5 pts) Derive analytical expression for the gradient of function $g(\mathbf{x}) = f(\mathbf{Ax} + \mathbf{b})$ if you know how to compute the gradient of f in any feasible point.

Problem 2

1. (1 pts) Compute the Jacobi matrix of the following function $f : \mathbb{R}^n \to \mathbb{R}^n$, $f(\mathbf{w})_j = \dfrac{e^{w_j}}{\sum\limits_{k=1}^{n} e^{w_k}}$.

   Also, consider how to compute this function in a stable manner if one of the elements in $\mathbf{w}$ is large.
2. (0.5 pts) Compute the gradient of the following functions with respect to matrix $\mathbf{X}$
   (a) $f(\mathbf{X}) = \sum\limits_{i=1}^{n} \lambda_i(\mathbf{X})$
   (b) $f(\mathbf{X}) = \prod\limits_{i=1}^{n} \lambda_i(\mathbf{X})$,

   where $\lambda_i(\mathbf{X})$ is the i-th eigenvalue of matrix $\mathbf{X}$.

[1] https://github.com/google/jax