Home assignment 2

Alexandr Katrutsa

The strict deadline to submit your solution is October 25, 01:00 MSK. Template to compose a solution is available at
`https://www.overleaf.com/read/vknkchxdwsmk`
The submission form is `https://forms.gle/XfQRzVybTvnEpa9X8`

1. Matrix calculus and autodiff (3.5 pts + 3 pts)

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. The programming part of this problem is evaluated separately and the maximum number of points for it is 3.

1. (1 pts) Compute the gradients with respect to $\mathbf{U} \in \mathbb{R}^{n\times n}$ and $\mathbf{V} \in \mathbb{R}^{n\times n}$ of function $J(\mathbf{U}, \mathbf{V}) = \|\mathbf{U}\mathbf{V} - \mathbf{Y}\|_F^2 + \frac{\lambda}{2}(\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2)$.

2. (1 pts) Compute the gradient and hessian of the following function $f(\mathbf{w}) = \sum\limits_{i=1}^{m} \log(1 + e^{-y_i \mathbf{w}^\top \mathbf{x}_i})$, where $\mathbf{x}_i \in \mathbb{R}^n$, $y_i \in \mathbb{R}$. Is this function convex or not and why?

3. (1 pts) Compute the Jacobi matrix of the following function $f : \mathbb{R}^n \to \mathbb{R}^n$, $f(\mathbf{w})_j = \dfrac{e^{w_j}}{\sum\limits_{k=1}^{n} e^{w_k}}$.

4. (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`

2. The shortest path in the graph (2 pts)

Let $G = (V, E)$ be a given weighted oriented graph. Denote by $\mathbf{c}$ the vector of weights assigned to the edges in this graph. Verify convexity/concavity of the function $p_{ij}(\mathbf{c})$, which is equal to the length of the shortest path between given pair of vertices (i, j) and depends on the vector $\mathbf{c}$.

3. Subdifferential computations (2.5 pts)

1. (1.5 pts) Compute the subdifferential of the following function $f(\mathbf{x}) = \|\mathbf{A}\mathbf{x} + \mathbf{b}\|_1$. Is this function convex or not and why?

2. (1 pts) Compute the subdifferential of the following function $L(\mathbf{w}) = \frac{1}{2}\|\mathbf{w}\|_2^2 + \sum\limits_{i=1}^{m} \max(0, 1 - y_i(\langle \mathbf{w}, \mathbf{x}_i \rangle + b_i))$. Is this function convex or not and why?