CUB Optimization Methods in Machine Learning, Spring 2024

Test Exam

Exam total time is 60 minutes. During exam no materials can be used. For each task you may get 1 point. The exam grade is computed as a sum of points for all tasks divided by 10.

1. Give definition of function with Lipschitz Hessian. Give an example of function that belongs to this class and does not belong to this class.
2. Formulate necessary and sufficient conditions for unconstrained local minima of smooth function.
3. Formulate test of ratios for detecting convergence rate of some sequence. Give an example of some sequence with linear convergence rate for which the test of ratios is not applicable.
4. Formulate four main operations that preserve convexity of a given set.
5. Sort the following optimization methods in ascending order w.r.t. complexity of one optimization iteration (without considering oracle computation): 1) Gradient Descent, 2) Newton method, 3) Conjugate Gradient, 4) SR-1. Explain your answer.
6. Formulate the main regularity conditions (constraint qualifications) for KKT theorem. Give an example of non-regular optimization problem.
7. Solve the following constrained optimization problem:
$$\begin{aligned}&\boldsymbol{c}^T\boldsymbol{x} + \sum_{i=1}^n x_i \log(x_i) \to \min_{\boldsymbol{x}: x_i > 0\ \forall i},\\ &\sum_{i=1}^n x_i = 1.\end{aligned}$$
8. Let's consider the following constrained optimization problem:
$$\begin{aligned}&\boldsymbol{c}^T\boldsymbol{x} \to \min_{\boldsymbol{x}},\\ &\|\boldsymbol{x}\|_2^2 \le b.\end{aligned}$$
Here $\boldsymbol{x}, \boldsymbol{c} \in \mathbb{R}^n$, $b > 0$. Construct the dual optimization problem.
9. Let's consider the following optimization problem:
$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \max_{i=1,\dots,m} (\boldsymbol{a}_i^T\boldsymbol{x} - b_i).$$
Transform this problem to the equivalent Linear Programming problem.
10. For the following function find its subdifferential:
$$f(\boldsymbol{x}) = \sum_{1\le i<j\le n} |x_i - x_j|.$$
11. For one-dimensional function $f(x) = 1/x$, defined for $x > 0$, find its Fenchel conjugate $f^*(s)$.
12. Write down the general scheme of Proximal Gradient Method with constant step size α for minimizing composite function. Indicate also the stopping criterion. Is this method a descent optimization method?