DS801 Advanced Optimization for Data Science Assignment 2

Spring 2024

Out: 3rd May 2024

Due: 17th April 2024 at 11:59pm

Instructions

- Submit a PDF document with your solutions through the assignment portal on KLMS by the due date. Please ensure that your name and student ID are on the front page.
- **Late assignments will not be accepted** except in extenuating circumstances. Special consideration should be applied for in this case.
- It is **required** that you **typeset your solutions in LaTeX**. Handwritten solutions will not be accepted.
- Spend some time ensuring your arguments are **coherent** and your solutions **clearly** communicate your ideas.

Question:	1	2	3	4	Total
Points:	20	30	30	20	100

1. (20 points) Consider the function
$$f(x,y) = 2x^2 + y^4 - y^2.$$
Find all stationary points and classify each of them as a local minimum or a saddle point.

2. (30 points) Let $\phi : X \times Y \to \mathbb{R}$ be a L-Lipschitz continuous convex-concave function. Assume that $\|x_1 - x_2\|_2 \le R$ for any $x_1, x_2 \in X$ and $\|y_1 - y_2\|_2 \le R$ for any $y_1, y_2 \in Y$. Prove that gradient descent ascent with step size $\eta = R/L\sqrt{T}$ guarantees that for any $(x,y) \in X \times Y$,
$$\phi\left(\frac{1}{T}\sum_{t=1}^{t} x_t, y\right) - \phi\left(x, \frac{1}{T}\sum_{t=1}^{t} y_t\right) \le \frac{2LR}{\sqrt{T}}.$$

3. (30 points) Recall that optimistic gradient descent ascent proceeds with the following update rule.
$$\begin{aligned}
x_{t+\frac{1}{2}} &= x_t - \eta \nabla_x \phi\left(x_{t-\frac{1}{2}}, y_{t-\frac{1}{2}}\right),\\
y_{t+\frac{1}{2}} &= y_t + \eta \nabla_y \phi\left(x_{t-\frac{1}{2}}, y_{t-\frac{1}{2}}\right),\\
x_{t+1} &= x_t - \eta \nabla_x \phi\left(x_{t+\frac{1}{2}}, y_{t+\frac{1}{2}}\right),\\
y_{t+1} &= y_t + \eta \nabla_y \phi\left(x_{t+\frac{1}{2}}, y_{t+\frac{1}{2}}\right).
\end{aligned}$$
Prove that the update rule is equivalent to
$$\begin{aligned}
x_{t+1} &= x_t - 2\eta \nabla_x \phi(x_t, y_t) + \eta \nabla_x \phi(x_{t-1}, y_{t-1}),\\
y_{t+1} &= y_t + 2\eta \nabla_y \phi(x_t, y_t) - \eta \nabla_y \phi(x_{t-1}, y_{t-1}).
\end{aligned}$$

4. Let us consider a convex-concave function ϕ over $X \times Y$.
   (a) (10 points) Prove that for the minimiax optimization of ϕ over $X \times Y$, (x^*, y^*) is a saddle point if and only if
   $$\nabla_x \phi(x^*, y^*)^\top (x - x^*) - \nabla_y \phi(x^*, y^*)^\top (y - y^*) \ge 0 \quad \forall (x,y) \in X \times Y.$$
   (b) (10 points) Prove that the operator F given by
   $$F(x,y) = \begin{bmatrix} \nabla_x \phi(x,y) \\ -\nabla_y \phi(x,y) \end{bmatrix}$$
   is monotone.