IE 539 Convex Optimization Assignment 2

Fall 2023

Out: 26th September 2023

Due: 9th October 2023 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:	25	25	25	25	100

1. (25 points) For a fixed z, derive closed form optimal solutions for
   (a) $\min_{x\in\mathbb{R}^d}\left\{\frac{1}{2}\|x-z\|_2^2+\lambda\|x\|_2^2\right\}$. (Here, $\lambda>0$.)
   (b) $\min_{x\in\mathbb{R}^d}\left\{\frac{1}{2}\|x-z\|_2^2+\lambda\|x\|_1\right\}$. (Here, $\lambda>0$.)

2. (25 points) Prove that (SOCP) is a conic program.

$$\begin{aligned}\text{minimize}\quad & f^\top x\\ \text{subject to}\quad & \|A_i x+b_i\|_2\le c_i^\top x+d_i\quad \text{for } i=1,\ldots,m,\\ & Ex=g.\end{aligned}\tag{SOCP}$$ eq:socp

3. (25 points) In this question we will work through the convergence proof of projected gradient descent and derive rates for Lipschitz continuous functions under various step sizes. The algorithm proceeds as follows:
   - Choose $x_1\in\mathbb{R}^d$.
   - For $t=1,2,3,\ldots,T+1$:
     - Obtain $\nabla f(x_t)$ and step size $\eta_t>0$.
     - Compute $x_{t+1}=x_t-\eta_t\nabla f(x_t)$.

   Assume that f is L-Lipschitz continuous in the ℓ_2-norm.
   (a) Set $\eta_t=\frac{1}{\sqrt{t}}$ for $t\ge 1$. Prove that

$$f\left(\left(\sum_{t=1}^T\eta_t\right)^{-1}\sum_{t=1}^T\eta_t x_t\right)-f(x^*)=O\left(\frac{\log T}{\sqrt{T}}\right).$$

   (b) Set $\eta_t=\frac{1}{t}$ for $t\ge 1$. Prove that

$$f\left(\left(\sum_{t=1}^T\eta_t\right)^{-1}\sum_{t=1}^T\eta_t x_t\right)-f(x^*)=O\left(\frac{1}{\log T}\right).$$

4. (25 points) In this question we will work through the convergence proof of projected gradient descent and derive rates for Lipschitz continuous functions. The algorithm proceeds as follows:
   - Choose $x_1\in C$.
   - For $t=1,2,3,\ldots,T+1$:
     - Obtain $\nabla f(x_t)$ and step size $\eta_t>0$.
     - Compute $x_{t+1}=\arg\min_{x\in C}\|x_t-\eta_t\nabla f(x_t)-x\|_2^2$.

   Assume that f is L-Lipschitz continuous in the ℓ_2-norm and that there exists some constant $R>0$ such that

$$\|x-y\|_2^2\le R^2$$

   for all $x,y\in C$.
   (a) Show that for each t, we have

$$\|x_{t+1}-x^*\|_2^2\le\|x_t-\eta_t\nabla f(x_t)-x^*\|_2^2$$

   where x^* is an optimal solution to $\min_{x\in C}f(x)$.
   (b) Set $\eta_t=\frac{\|x_1-x^*\|_2}{L\sqrt{T}}$. Then use part (a) to show that

$$f\left(\frac{1}{T}\sum_{t=1}^T x_t\right)-f(x^*)\le\frac{L\|x_1-x^*\|_2}{\sqrt{T}}.$$

   (c) Set $\eta_t=\frac{1}{\sqrt{t}}$. Prove that

$$f\left(\frac{1}{T}\sum_{t=1}^T x_t\right)-f(x^*)=O\left(\frac{1}{\sqrt{T}}\right)$$