IE 539 Convex Optimization Assignment 3

Fall 2023

Out: 23rd October 2023

Due: 8th November 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 be subject to a penalty. 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) Prove that for a positive definite matrix A,

$$f(x) = \frac{1}{2}x^\top Ax + b^\top x + c$$

is smooth and strongly convex in the ℓ_2-norm. Write down the smoothness constant and the strong convexity constant.

2. (25 points) In this question we prove the convergence of the projected subgradient method for functions that are strongly convex and Lipschitz continuous. Let $f : C \to \mathbb{R}$ be a function that is α-strongly convex with respect to the ℓ_2 norm and L-Lipschitz continuous in the ℓ_2 norm over a convex domain C. Recall that the projected subgradient method proceeds as follows.

   - Choose $x_1 \in C$.
   - For $t = 1, 2, 3, \ldots, T-1$:
     - Select any subgradient $g_t \in \partial f(x_t)$ and step size $\eta_t > 0$.
     - Compute $x_{t+1} = \mathrm{Proj}_C\{x_t - \eta_t g_t\}$.

   (a) Set $\eta_t = \frac{2}{\alpha(t+1)}$. Show that

   $$f\left(\sum_{t=1}^{T} \frac{2t}{T(T+1)} x_t\right) - f(x^*) \le \frac{2L^2}{\alpha(T+1)}$$

   where $x^* \in \arg\min_{x \in C} f(x)$.

   (b) Set $\eta_t = \frac{1}{\alpha t}$. Show that

   $$f\left(\frac{1}{T}\sum_{t=1}^{T} x_t\right) - f(x^*) \le \frac{L^2(1+\log T)}{2\alpha T}$$

   where $x^* \in \arg\min_{x \in C} f(x)$.

3. (25 points) In this question we will work through the convergence analysis of the online (projected) subgradient method for online convex optimization where the loss functions are strongly convex and Lipschitz continuous. Let $f_1, \ldots, f_T : C \to \mathbb{R}$ be a loss functions that are α-strongly convex with respect to the ℓ_2 norm and L-Lipschitz continuous in the ℓ_2 norm over a convex domain C. Recall that the online (projected) subgradient method proceeds as follows.

   - Choose $x_1 \in C$.
   - For $t = 1, 2, 3, \ldots, T-1$:
     - Observe f_t and Select any subgradient $g_t \in \partial f_t(x_t)$ and step size $\eta_t > 0$.
     - Compute $x_{t+1} = \mathrm{Proj}_C\{x_t - \eta_t g_t\}$.

   (a) Show that for each t, we have

   $$f_t(x_t) - f_t(x^*) \le \left(\frac{1}{2\eta_t} - \frac{\alpha}{2}\right)\|x_t - x^*\|_2^2 - \frac{1}{2\eta_t}\|x_{t+1} - x^*\|_2^2 + \frac{\eta_t}{2}\|g_t\|_2^2$$

   where x^* is an optimal solution to $\min_{x\in C}\sum_{t=1}^{T} f_t(x^*)$.

   (b) Set $\eta_t = \frac{1}{\alpha t}$. Then use part (a) to show that

   $$\sum_{t=1}^{T} f_t(x_t) - \min_{x\in C}\sum_{t=1}^{T} f_t(x) \le \frac{L^2}{2\alpha}(1 + \ln T).$$

4. (25 points) In this question we prove the convergence of stochastic gradient descent for functions that are strongly convex and Lipschitz continuous. Let $f : C \to \mathbb{R}$ be a function that is α-strongly convex with respect to the ℓ_2 norm and L-Lipschitz continuous in the ℓ_2 norm over a convex domain C. Recall that stochastic gradient descent proceeds as follows.

   - Choose $x_1 \in C$.
   - For $t = 1, 2, 3, \ldots, T-1$:
     - Obtain an unbiased estimator $\hat{g}_{x_t}$ of some $g \in \partial f(x_t)$.
     - Update $x_{t+1} = \mathrm{Proj}_C\left\{x_t - \eta_t \hat{g}_{x_t}\right\}$ for a step size $\eta_t > 0$.

Set $\eta_t = \frac{1}{\alpha t}$. Assuming $\|\hat{g}_{x_t}\|_2 \leq L$ for all t, show that

$$\mathbb{E}\left[f\left(\frac{1}{T}\sum_{t=1}^{T} x_t\right)\right] - f(x^*) \leq \frac{L^2(1+\log T)}{2\alpha T}$$

where $x^* \in \arg\min_{x \in C} f(x)$.