Introduction to Optimization - Homework 1

Due October 3rd 2023, 15:00

Note: The problems have a logical order, but can be solved separately.

Let A be an $M \times N$ matrix, and let $b \in \mathbb{R}^M$. Consider the least-squares optimization problem, which consists in minimizing the function $f \colon \mathbb{R}^N \to \mathbb{R}$, defined by

$$\min_{x \in \mathbb{R}^N} f(x), \quad \text{where } f(x) = \frac{1}{2}\|Ax - b\|^2.$$

1. Prove the problem always has a solution. In other words, show that the minimum exists and is attained for all matrices A.

2. Show the solution to the problem is unique if A has full rank.

3. Characterize the set of solutions. If you get stuck, analyze the case $b \in \operatorname{ran}(A)$ first. You may provide your answer as a case distinction, where the answer differs based on whether A has full rank or not.

4. Show the Lipschitz constant of the gradient of f is $L = \max \sigma(A^T A)$, where $\sigma(\cdot)$ denotes the spectrum (set of eigenvalues) of a given matrix.
   Hint: You may use without proof that $\sup_{\mathbf{x} \in \mathbb{R}^N} \|A^T A\mathbf{x}\|/\|\mathbf{x}\| \leq \max \sigma(A^T A)$.

5. Given $\varepsilon > 0$, how many iterations of the gradient method, starting from $x^0 = 0$, are necessary to find a point x_ε such that $f(x_\varepsilon) - \min(f) \leq \varepsilon$?

6. Prove that for all $x \in \mathbb{R}^N$, it holds that

   $$\|\nabla f(x)\|^2 \geq 2\left[\min \sigma(A^T A)\right](f(x) - \min(f)).$$

   Hint: You may use without proof that $\inf_{\mathbf{x} \in \mathbb{R}^M} \|A^T\mathbf{x}\|/\|\mathbf{x}\| \geq \sqrt{\min \sigma(A^T A)}$, *where* $\sigma(\cdot)$ *denotes the spectrum (set of eigenvalues) of a given matrix.*

7. Does the result of 5 change your response to 4?

8. Show that the function $f_\mu(x) = f(x) + \frac{\mu}{2}\|x\|^2$ is strongly convex, and characterize its unique global minimum.

9. Given $\varepsilon > 0$, how many iterations of the gradient method, starting from $x^0 = 0$, are necessary to find a point x_ε such that $f_\mu(x_\varepsilon) - \min(f_\mu) \leq \varepsilon$?