Optimization for Data Science
September 21, 2018

1. (6 POINTS) Describe in depth the PageRank problem and the related mathematical model. Furthermore, propose a possible algorithm for its solution.
2. (7 POINTS) Describe in depth gradient method and accelerated gradient method, and explain the differences between the two.
3. (7 POINTS) Given the set of samples $V = \{v_1, \dots, v_n\}$, with $v_i \in \mathbb{R}^m, v_i \geq 0$, $i = 1, \dots, n$ and n very large, consider the following problem:

$$\min_{x \in conv\{v_1,\dots,v_n\}} f(x) = \|Ax - b\|_2^2 + \lambda \|x\|_1,$$

   with $A \in \mathbb{R}^{l \times m},\ b \in \mathbb{R}^l$ and $\lambda > 0$. Describe the properties of the problem and propose an efficient algorithm for its solution.
4. (8 POINTS) Consider the two sets:
   - $D_1 = \{x \in \mathbb{R}^n :\ a^\top x = b,\ x \geq 0\}$, with $a \in \mathbb{R}^n,\ b \in \mathbb{R},\ a, b > 0$;
   - $D_2 = \{x \in \mathbb{R}^n : \|Qx\|_2 \leq 1\}$, with $Q \in \mathbb{R}^{n \times n}$ symmetric positive definite matrix.

   Describe how to calculate the Frank-Wolfe direction for the problems

$$\min_{x \in D_1} f(x), \ \min_{x \in D_2} f(x).$$

   Furthermore give the computational costs for the two different tasks.
5. (8 POINTS) Consider the problem of solving the linear system:

$$Ax = b, \tag{1}$$

   with $A \in \mathbb{R}^{m \times n},\ b \in \mathbb{R}^m$. Let $m \geq n$ and assume A has full rank, so that there is a unique solution $x^* \in \mathbb{R}^n$. The problem can be rewritten as

$$\min_{x \in \mathbb{R}^n} f(x) = \frac{1}{2}\|Ax - b\|_2^2. \tag{2}$$

   Show[1] that the gradient method $x_{k+1} = x_k - \alpha \nabla f(x_k)$, with stepsize

$$\alpha = \frac{1}{\sigma_{max}(A)^2},$$

   applied to Problem (2) has the following rate:

$$\|x_{k+1} - x^*\|_2^2 \leq \left(1 - \frac{\sigma_{min}(A)^2}{\sigma_{max}(A)^2}\right) \|x_k - x^*\|_2^2.$$

---

[1] Keep in mind that

- $\sigma_{max}(A) = \|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2}$ and $\sigma_{min}(A) = \min_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2}$;
- For a given x, we have $\|Ax\|_2 \leq \|A\|_2 \|x\|_2$;
- $\sigma_{max}(I - \alpha A^\top A)^2 = 1 - \alpha \sigma_{min}(A)^2$.