Optimization for Data Science
February 10, 2020

1. (8 POINTS) Describe in depth the randomized block coordinate gradient descent method and the Gauss-Southwell BCGD method highlighting their PROs and CONs. Furthermore, analyze the differences between the two methods.
2. (7 POINTS) Describe in depth the Frank-Wolfe Algorithm.
3. (7 POINTS). Describe the Page Rank problem and its mathematical model. Furthermore, propose a method to solve it.
4. (6 POINTS) Consider the problem of projecting a point $v \in \mathbb{R}^n$ over the ℓ_1-ball:

$$\begin{array}{ll} \min\limits_{w \in \mathbb{R}^n} & \|w - v\|_2^2 \\ \text{s.t.} & \|w\|_1 \leq r \end{array} \tag{1}$$

   with $r > 0$. Let w be an optimal solution of Problem (1). Prove that for all $i = 1, \ldots, n$, we have $w_i v_i \geq 0$.
   Taking into account the theoretical result described above, and assuming that an efficient procedure for projecting over the simplex

$$\Delta = \{w \in \mathbb{R}^n : \ e^\top w = r, \ w \geq 0\}$$

   is available, describe a method for efficiently projecting over the ℓ_1-ball.
5. (8 POINTS) Considering $r_k = -\nabla f(x_k)$, prove that the Away-step Frank-Wolfe direction d_k^{AS} and the Pairwise Frank-Wolfe direction d_k^{PW} satisfy a relation of the following form

$$r_k{}^\top d_k^{AS} \geq c \cdot r_k{}^\top d_k^{PW},$$

   with $c > 0$ suitably chosen (please specify c). Furthermore, consider the problem

$$\begin{array}{ll} \min\limits_{x \in \mathbb{R}^n} & f(x) \\ \text{s.t.} & x \in \Delta, \end{array} \tag{2}$$

   with Δ the unit simplex. Prove that at iteration k of the Away-step Frank-Wolfe method, when $(x_k)^i = 0, i \in 1, \ldots, n$ and $r_k{}^\top (e_i - x_k) < 0$, then $(x_{k+1})^i = 0$.