

Problem Set 9, due May 11, 2019 (Frank-Wolfe)

Convergence of Frank-Wolfe

Exercise 1:

Assuming $h_0 \leq 2C$, and the sequence h_0, h_1, \dots satisfies

$$h_{t+1} \leq (1 - \gamma)h_t + \gamma^2 C \quad t = 0, 1, \dots$$

for $\gamma = \frac{2}{t+2}$, prove that

$$h_t \leq \frac{4C}{t+2} \quad t = 0, 1, \dots$$

Applications of Frank-Wolfe

Exercise 2:

Derive the LMO formulation for matrix completion, that is

$$\min_{Y \in X \subseteq \mathbb{R}^{n \times m}} \sum_{(i,j) \in \Omega} (Z_{ij} - Y_{ij})^2$$

when $\Omega \subseteq [n] \times [m]$ is the set of observed entries from a given matrix Z .

Where our optimization domain X is the unit ball of the trace norm (or nuclear norm), which is defined the convex hull of the rank-1 matrices

$$X := \text{conv}(\mathcal{A}) \quad \text{with} \quad \mathcal{A} := \left\{ \mathbf{u}\mathbf{v}^\top \mid \begin{array}{l} \mathbf{u} \in \mathbb{R}^n, \|\mathbf{u}\|_2 = 1 \\ \mathbf{v} \in \mathbb{R}^m, \|\mathbf{v}\|_2 = 1 \end{array} \right\}.$$

1. Derive the LMO for this set X for a gradient at iterate $Y \in \mathbb{R}^{n \times m}$.
2. Dervie the *projection* step onto X . How does the computational operations (or costs) needed to compute the LMO and the projection step compare?

Practical Implementation

Follow the Python notebook provided here:

github.com/epfml/OptML_course/tree/master/labs/ex09/