Homework 12

Requirements:

1. Digital format (can be typeset or photos, ought to write clearly if written by hand), upload to `https://course.pku.edu.cn/`.
2. Submit by next class
3. A problem is not counted if nobody can work it out
4. Each homework 10 points; 1 point deducted for each week's delay

Problems:

1. Consider:

$$\min_{\mathbf{x}\in\mathbb{R}^N} \|\mathbf{y}-\mathbf{D}\mathbf{x}\|_2^2, \quad s.t. \quad \|\mathbf{x}\|_? \leq 1,$$

where $\|\mathbf{x}\|_?$ can be either the ℓ_1 norm or the ℓ_∞ norm. Randomly generate $\mathbf{D}\in\mathbb{R}^{200\times 300}$ and $\mathbf{y}\in\mathbb{R}^{200}$ and use Frank-Wolfe algorithm to solve it, for both ℓ_1 norm and ℓ_∞ norm. Further compare F-W algorithm with the projected gradient descent in convergence speed ($\log(f(\mathbf{x}_k)-f^*)$ vs. iteration number, where f^* is obtained by running the better algorithm many many iterations). Hand in your code and report.

2. For an unconstrained composite problem $\min_{\mathbf{x}} f(\mathbf{x})+g(\mathbf{x})$, we may introduce constraint $\mathbf{x}=\mathbf{y}$ to make it a separable problem with linear constraint:

$$\min_{\mathbf{x},\mathbf{y}} f(\mathbf{x})+g(\mathbf{y}), \quad s.t. \quad \mathbf{x}=\mathbf{y}.$$

Now let $f(\mathbf{x})=\|\mathbf{x}-\mathbf{a}\|_2$, $g(\mathbf{x})=\|\mathbf{x}-\mathbf{b}\|_1$, where $\mathbf{a}\neq\mathbf{b}\in\mathbb{R}^{100}$ are randomly generated vectors. Try whether the dual ascend can be used to solve it. If yes, compare it with ADMM in convergence speed ($\log(f(\mathbf{x}_k)-f^*)$ vs. iteration number, where f^* can be estimated by running ADMM many many iterations.). If no, explain why and only show the result of ADMM ($\log(f(\mathbf{x}_k)-f^*)$ vs. iteration number). When using ADMM, please use $\tau\beta$ in the update of the Lagrange multiplier, where $\beta>0$ is fixed at three values chosen by you and $\tau\in\{0.5,1,1.5\}$ (Namely you have to try 9 cases). Hand in your code and report.

3. Use LADMAP to solve a graph construction problem:

$$\min_{\mathbf{Z},\mathbf{E}} ||\mathbf{Z}||_* + \lambda||\mathbf{E}||_{2,1}, \quad \text{s.t.} \quad \mathbf{D}=\mathbf{D}\mathbf{Z}+\mathbf{E}, \mathbf{Z}^T\mathbf{1}=\mathbf{1}, \tag{1}$$

where $\mathbf{1}$ is an all-one vector. Randomly generate $\mathbf{D}\in\mathbb{R}^{200\times 300}$. Hand in your code and report.