MSBD5007 Optimization and Matrix Computation
Homework 4

Due date: April 20, Sunday

1. Consider the function $f : \mathbb{R}^d \to \mathbb{R}$ defined by

$$f(\boldsymbol{x}) = \sum_{i=1}^{d} \max(0, 1 - x_i),$$

where $\boldsymbol{x} = [x_1, x_2, \ldots, x_d]^T$. Recall that the proximity operator of a function $g : \mathbb{R}^d \to \mathbb{R}$ is defined as

$$\operatorname{prox}_g(\boldsymbol{y}) = \arg\min_{\boldsymbol{x}\in\mathbb{R}^d} \left\{ g(\boldsymbol{x}) + \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{y}\|_2^2 \right\}, \quad \boldsymbol{y} \in \mathbb{R}^d.$$

Derive a closed-form expression for $\operatorname{prox}_f(\boldsymbol{y})$.

2. In this problem, we study two properties of the 2-norm function $g(\boldsymbol{x}) = \|\boldsymbol{x}\|_2$ defined on $\mathbb{R}^n$. Provide detailed derivations to show that:

   (a) The subdifferential of g is given by

$$\partial\|\boldsymbol{x}\|_2 = \begin{cases} \left\{ \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_2} \right\} & \text{if } \boldsymbol{x} \neq \boldsymbol{0}, \\ \{\boldsymbol{u} \in \mathbb{R}^n \mid \|\boldsymbol{u}\|_2 \leq 1\} & \text{if } \boldsymbol{x} = \boldsymbol{0}. \end{cases}$$

   (b) For any $\alpha > 0$, the proximity operator of $\alpha\|\cdot\|_2$ is

$$\operatorname{prox}_{\alpha\|\cdot\|_2}(\boldsymbol{y}) = \begin{cases} \left(1 - \frac{\alpha}{\|\boldsymbol{y}\|_2}\right)\boldsymbol{y} & \text{if } \|\boldsymbol{y}\|_2 \geq \alpha, \\ \boldsymbol{0} & \text{if } \|\boldsymbol{y}\|_2 \leq \alpha. \end{cases}$$

3. In this problem, we consider the elastic net regression model, which is widely used in statistics for regularized linear regression. The optimization problem is given by

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \lambda_1\|\boldsymbol{x}\|_1 + \frac{\lambda_2}{2}\|\boldsymbol{x}\|_2^2,$$

where $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{b} \in \mathbb{R}^m$, and $\lambda_1, \lambda_2 > 0$ are regularization parameters. Answer the following:

   (a) For any $\beta_1, \beta_2 > 0$, find a closed-form expression for the proximity operator $\operatorname{prox}_{\beta_1\|\cdot\|_1 + \frac{\beta_2}{2}\|\cdot\|_2^2}(\boldsymbol{y})$.

   (b) We apply the forward-backward splitting (i.e., proximal gradient) algorithm. In particular, we apply a forward step for $\frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2$ and a backward step for $\lambda_1\|\boldsymbol{x}\|_1 + \frac{\lambda_2}{2}\|\boldsymbol{x}\|_2^2$. Write down the iterative update rule for the resulting algorithm.

4. Let $g : \mathbb{R}^n \to \mathbb{R}$ be a convex function. Prove the following properties:

(a) For any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and for any $\boldsymbol{u} \in \partial g(\boldsymbol{x})$ and $\boldsymbol{v} \in \partial g(\boldsymbol{y})$, show that

$$\langle \boldsymbol{u} - \boldsymbol{v},\, \boldsymbol{x} - \boldsymbol{y} \rangle \geq 0.$$

Hint: Use the definition of the subdifferential.

(b) Prove that the proximity operator of g is nonexpansive; that is, for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$,

$$\| \operatorname{prox}_g(\boldsymbol{x}) - \operatorname{prox}_g(\boldsymbol{y}) \|_2 \leq \|\boldsymbol{x} - \boldsymbol{y}\|_2.$$

Hint: Apply the result from part (a).

(c) Show that a point $\boldsymbol{x}^*$ minimizes g if and only if

$$\boldsymbol{x}^* = \operatorname{prox}_g(\boldsymbol{x}^*).$$