

MSBD 5007 HW4

RONG Shuo

April 13, 2025

Question1

Consider the function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ defined by

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

where $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$. Recall that the proximity operator of a function $g : \mathbb{R}^d \rightarrow \mathbb{R}$ is defined as

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

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

Question2

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

(a) The subdifferential of g is given by

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

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

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

Question 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

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

where $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$, 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 proximity operator $\text{prox}_{\beta_1 \|\cdot\|_1 + \frac{\beta_2}{2} \|\cdot\|_2^2}(\mathbf{y})$.

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

Question 4