MSBD 5004 Mathematical Methods for Data Analysis
Homework 3

Due date: October 31, Thursday

1. Determine whether each of the following scalar-valued functions of n-vectors is linear. If it is a linear function, give its inner product representation, i.e., an n-vector $\mathbf{a}$ for which $f(\boldsymbol{x}) = \mathbf{a}^T\boldsymbol{x}$ for all $\boldsymbol{x}$. If it is not linear, give specific $\boldsymbol{x}$, $\boldsymbol{y}$, α and β such that

$$f(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) \neq \alpha f(\boldsymbol{x}) + \beta f(\boldsymbol{y}).$$

   (a) The spread of values of the vector, defined as $f(\boldsymbol{x}) = \max_k x_k - \min_k x_k$.

   (b) The difference of the last element and the first, $f(\boldsymbol{x}) = x_n - x_1$.

2. Consider the regression model $y = \boldsymbol{x}^T\boldsymbol{a} + b$, where y is the predicted response, $\boldsymbol{x}$ is an 8-vector of features, $\boldsymbol{a}$ is an 8-vector of coefficients, and b is the offset term. Determine with reasoning whether each of the following statements is true or false.

   (a) If $a_3 > 0$ and $x_3 > 0$, then $y \geq 0$.

   (b) If $a_2 = 0$ then the prediction y does not depend on the second feature x_2.

   (c) If $a_6 = -0.8$, then increasing x_6 (keeping all other x is the same) will decrease y.

3. In linear regression models, we consider two data points $(\boldsymbol{x}_1, y_1)$ and $(\boldsymbol{x}_2, y_2)$ with $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^2$ and $y_1, y_2 \in \mathbb{R}$. For simplicity, we set the bias term $b = 0$. Let $\boldsymbol{X} \in \mathbb{R}^{2\times 2}$ have rows $\boldsymbol{x}_1^T$ and $\boldsymbol{x}_2^T$, and let $\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \in \mathbb{R}^2$. Assume the columns of $\boldsymbol{X}$, denoted by $\boldsymbol{x}^{(1)}$ and $\boldsymbol{x}^{(2)}$, are linearly dependent such that $\boldsymbol{x}^{(1)} = 2\boldsymbol{x}^{(2)}$.

   (a) Consider the least squares estimation:

$$\min_{\boldsymbol{\beta}\in\mathbb{R}^2} \|\boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{y}\|_2^2. \tag{1}$$

   What problem does the linear dependency among the columns of $\boldsymbol{X}$ cause when estimating $\boldsymbol{\beta}$ using least squares?

   (b) Now consider the ridge regression, which incorporates a regularization term:

$$\min_{\boldsymbol{\beta}\in\mathbb{R}^2} \|\boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{y}\|_2^2 + \lambda\|\boldsymbol{\beta}\|_2^2, \tag{2}$$

   where $\lambda > 0$ is a regularization parameter. Derive the solution $\hat{\boldsymbol{\beta}}$ of (2). What is the ratio between $\hat{\beta}_1$ and $\hat{\beta}_2$?

   (c) Discuss how varying the value of λ affects the solution and its ability to mitigate issues arising from linear dependency of columns of $\boldsymbol{X}$.

4. Let $\{(\boldsymbol{x}_i, y_i)\}_{i=1}^N$ be given with $\boldsymbol{x}_i \in \mathbb{R}^n$ and $y_i \in \mathbb{R}$. Consider the soft-SVM:

$$\min_{\boldsymbol{a}\in\mathbb{R}^n, b\in\mathbb{R}} \sum_{i=1}^{N} h\left(y_i(\langle \boldsymbol{a}, \boldsymbol{x}_i\rangle + b) - 1\right) + \lambda \|\boldsymbol{a}\|_2^2,$$

where $\lambda \in \mathbb{R}$ is a regularization parameter and $h(t) = \max\{0, -t\}$ is the hinge loss function. Prove that solving the above soft-SVM is equivalent to solving the following problem:

$$\begin{aligned} \min_{\boldsymbol{a}\in\mathbb{R}^n, b\in\mathbb{R}, \boldsymbol{\xi}\in\mathbb{R}^N} & \sum_{i=1}^{N} \xi_i + \lambda \|\boldsymbol{a}\|_2^2, \\ \text{s.t. } & y_i(\langle \boldsymbol{a}, \boldsymbol{x}_i\rangle + b) \geq 1 - \xi_i \text{ and } \xi_i \geq 0,\ i = 1, 2, \ldots, N \end{aligned}$$

5. Let V be a Hilbert space. Let S_1 and S_2 be two hyperplanes in V defined by

$$S_1 = \{\boldsymbol{x} \in V \mid \langle \boldsymbol{a}_1, \boldsymbol{x}\rangle = b_1\}, \quad S_2 = \{\boldsymbol{x} \in V \mid \langle \boldsymbol{a}_2, \boldsymbol{x}\rangle = b_2\}.$$

Assume $S_1 \cap S_2$ is non-empty. Let $\boldsymbol{y} \in V$ be given. We consider the projection of $\boldsymbol{y}$ onto $S_1 \cap S_2$, i.e., the solution of

$$\min_{\boldsymbol{x}\in S_1\cap S_2} \|\boldsymbol{x} - \boldsymbol{y}\|. \tag{3}$$

(a) Prove that $S_1 \cap S_2$ is a plane, i.e., if $\boldsymbol{x}, \boldsymbol{z} \in S_1 \cap S_2$, then $(1+t)\boldsymbol{z} - t\boldsymbol{x} \in S_1 \cap S_2$ for any $t \in \mathbb{R}$.

(b) Prove that $\boldsymbol{z}$ is a solution of (3) if and only if $\boldsymbol{z} \in S_1 \cap S_2$ and

$$\langle \boldsymbol{z} - \boldsymbol{y}, \boldsymbol{z} - \boldsymbol{x}\rangle = 0, \quad \forall \boldsymbol{x} \in S_1 \cap S_2. \tag{4}$$

(c) Find an explicit solution of (3).

(d) Prove the solution found in part (c) is unique.