GEORGIA INSTITUTE OF TECHNOLOGY

Mathematical Foundations of Machine Learning, Quiz #1

September 28, 2022

Name: Solution.
Last, First

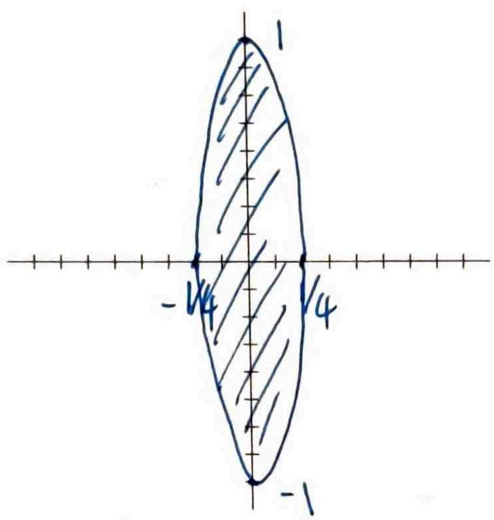
- Closed book, closed notes, one $8\frac{1}{2}'' \times 11''$ handwritten sheet is allowed.
- Seventy-five (75) minute time limit.
- No calculators. None of the problems require involved calculations.
- There are four problems, each are worth 25 points. Subproblems are given equal weight.
- All work should be performed on the quiz itself. If more space is needed, use the backs of the pages.
- This quiz will be conducted under the rules and guidelines of the Georgia Tech Honor Code and no cheating will be tolerated.
- **Write your final answers in the boxes provided.**
- **Turn in your "cheat sheet" by placing it in between the first and second pages.**

Problem 1:

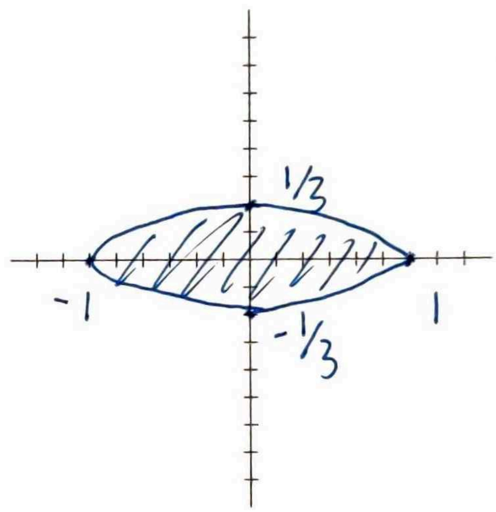
(a) Let $\|\cdot\|_A$ and $\|\cdot\|_B$ be the following valid norms on $\mathbb{R}^2$

$$\|\boldsymbol{x}\|_A = \sqrt{16x_1^2 + x_2^2}, \quad \|\boldsymbol{x}\|_B = \sqrt{x_1^2 + 9x_2^2}.$$

Sketch the unit balls on the axes below.



$\mathcal{B}_A = \{\boldsymbol{x} \;:\; \|\boldsymbol{x}\|_A \leq 1\}$



$\mathcal{B}_B = \{\boldsymbol{x} \;:\; \|\boldsymbol{x}\|_B \leq 1\}$

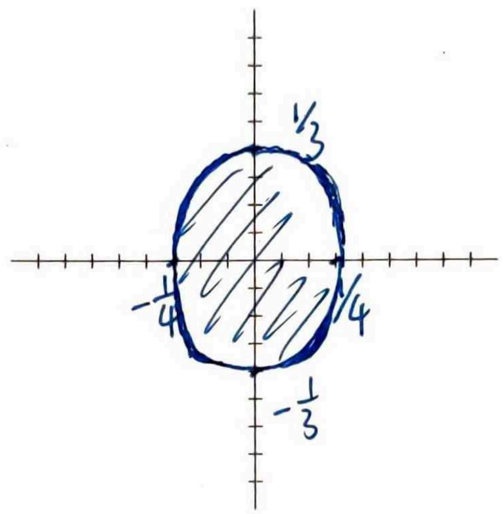
① Must indicate $(\pm\frac{1}{4}, 0), (0, \pm 1)$ on B_A
$(\pm 1, 0), (0, \pm\frac{1}{3})$ on B_B.

② Must indicate that the interior of unit balls are parts of the unit ball.

(b) Let

$$\|\boldsymbol{x}\|_C = \max(\|\boldsymbol{x}\|_A, \|\boldsymbol{x}\|_B),$$

where $\|\cdot\|_A$ and $\|\cdot\|_B$ are the norms defined in the previous part. Sketch the unit ball on the axes below.



$$\mathcal{B}_C = \{\boldsymbol{x} : \|\boldsymbol{x}\|_C \leq 1\}$$

Drawing must indicate $B_C = B_A \cap B_B$.
(Don't have to be explicit).

(c) True or False: $\|\cdot\|_C$ is a valid norm on $\mathbb{R}^2$.

Circle one: True False

Justification:

Must show by definition.

(1) $\|\vec{x}\|_c = 0 \iff \max(\|X\|_A, \|X\|_B) = 0$

Since $\|X\|_A, \|X_B\| \geq 0$. $\vec{x} = \vec{0}$.

(2). $\|\alpha\vec{x}\|_c = \max(\|\alpha X\|_A . \|\alpha X\|_B)$

$= |\alpha| \max(\|\vec{x}\|_A, \|\vec{x}\|_B)$

$= |\alpha| \, \|\vec{x}\|_c$

(3). $\|\vec{x}+\vec{y}\|_c = \max(\|\vec{x}+\vec{y}\|_A, \|\vec{x}+\vec{y}\|_B)$

$\leq \max(\|\vec{x}\|_A + \|\vec{y}\|_A . \|\vec{x}\|_B + \|\vec{y}\|_B)$

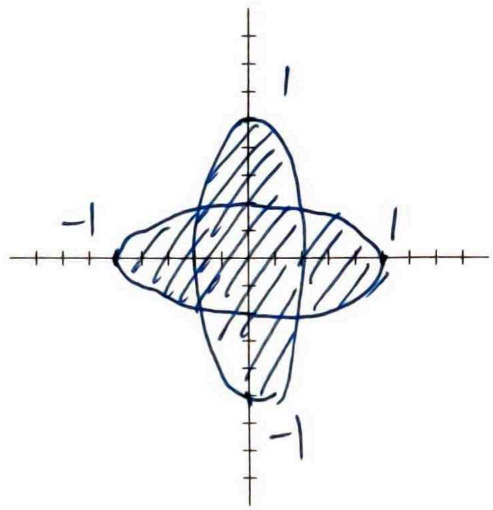
$\leq \max(\|\vec{x}\|_A, \|\vec{x}\|_B) + \max(\|\vec{y}\|_A, \|\vec{y}\|_B)$

$= \|\vec{x}\|_c + \|\vec{y}\|_c$.

(d) Let

$$\|\boldsymbol{x}\|_D = \min(\|\boldsymbol{x}\|_A, \|\boldsymbol{x}\|_B),$$

where $\|\cdot\|_A$ and $\|\cdot\|_B$ are the norms defined in the previous part.
Sketch the unit ball on the axes below.



$$\mathcal{B}_D = \{\boldsymbol{x} \ : \ \|\boldsymbol{x}\|_D \le 1\}$$

Should express the understanding that

$B_D = B_A \cup B_B$.

(e) True or False: $\|\cdot\|_D$ is a valid norm on $\mathbb{R}^2$.

Circle one: True False

Justification:

Must find counterexample.

Consider $\vec{x} = \begin{bmatrix}1\\0\end{bmatrix}$ $\vec{y} = \begin{bmatrix}0\\1\end{bmatrix}$

$\|\vec{x}\|_A = 1$ $\|\vec{x}\|_B = 3$. $\|\vec{x}\|_D = \min(1,3) = 1$

$\|\vec{y}\|_A = 3$ $\|\vec{y}\|_B = 1$, $\|\vec{y}\|_D = \min(3,1) = 1$

$\|\vec{x}+\vec{y}\|_A = \|\vec{x}+\vec{y}\|_B = \sqrt{10}$. $\rightarrow$ $\|\vec{x}+\vec{y}\|_D = \sqrt{10}$

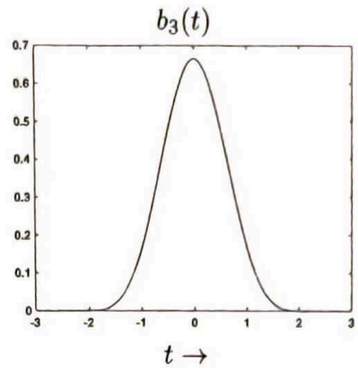
$\|\vec{x}+\vec{y}\|_D = \sqrt{10} > 2 = \|\vec{x}\|_D + \|\vec{y}\|_D$. $\Longrightarrow\!\Longleftarrow$.

Problem 2: Suppose that $f : \mathbb{R} \to \mathbb{R}$ is a third-order spline defined by the overlap of six B-splines:

$$f(t) = \sum_{k=0}^{5} \alpha_k b_3(t-k),$$

where $b_3(t)$ is the cubic B-spline function:

$$b_3(t) = \begin{cases} (t+2)^3/6 & -2 \le t \le -1 \\ -t^3/3 - t^2 + 2/3 & -1 \le t \le 0 \\ t^3/2 - t^2 + 2/3 & 0 \le t \le 1 \\ -(t-2)^3/6 & 1 \le t \le 2 \\ 0 & |t| \ge 2 \end{cases}.$$



Suppose

$$f(0) = -5, \quad f(1) = -1, \quad f(2) = 3, \quad f(3) = 0 \quad f(4) = -3, \quad f(5) = 7$$

Write the linear system of equations that we have to solve to find the unique α_k corresponding to these samples. (You do not have to solve the system.)

$$\begin{bmatrix} \frac{2}{3} & \frac{1}{6} & 0 & 0 & 0 & 0 \\ \frac{1}{6} & \frac{2}{3} & \frac{1}{6} & 0 & 0 & 0 \\ 0 & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} & 0 & 0 \\ 0 & 0 & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} & 0 \\ 0 & 0 & 0 & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \\ 0 & 0 & 0 & 0 & \frac{1}{6} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \end{bmatrix} = \begin{bmatrix} -5 \\ -1 \\ 3 \\ 0 \\ -3 \\ 7 \end{bmatrix}$$

$\uparrow$

index must be correct.

Problem 3: Given a 2×2 matrix $\boldsymbol{Q}$, define

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_Q = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{Q} \boldsymbol{y} \quad \text{for all } \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2.$$

Circle the matrices $\boldsymbol{Q}$ below that make $\langle \cdot, \cdot \rangle_Q$ a valid inner product on $\mathbb{R}^2$.

$$\boldsymbol{Q} = \begin{bmatrix} 0.999 & 0 \\ 0 & 0.001 \end{bmatrix}$$

$$\boldsymbol{Q} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

$$\boldsymbol{Q} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

$$\boldsymbol{Q} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

$$\boldsymbol{Q} = \begin{bmatrix} 4 & 2 \\ 3 & 4 \end{bmatrix}$$

Problem 4:

(a) Given 4 data points

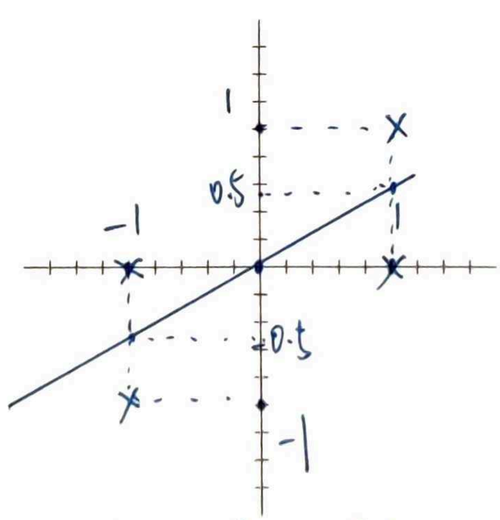
$$(x_i, y_i) \in \{(1,1), (1,0), (-1,0), (-1,-1)\},$$

find the vector $\widehat{\boldsymbol{\beta}} = \begin{bmatrix} \widehat{\beta}_0 \\ \widehat{\beta}_1 \end{bmatrix}$ such that

$$\widehat{\boldsymbol{\beta}} = \underset{\beta = [\beta_0, \beta_1]^\top \in \mathbb{R}^2}{\arg\min} \sum_{i=1}^{4} (y_i - \beta_0 - \beta_1 x_i)^2 .$$

You need to set up the least-squares problem in the matrix form by providing the design matrix $\boldsymbol{X}$ and the target vector $\boldsymbol{y}$ such that $\boldsymbol{y} \approx \boldsymbol{X}\widehat{\boldsymbol{\beta}}$. Sketch the linear regression line $\hat{f}(x) = \widehat{\beta}_0 + \widehat{\beta}_1 x$ on the axes below along with the 4 data points.

$$\boldsymbol{X} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \\ 1 & -1 \end{bmatrix} \qquad \boldsymbol{y} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix} \qquad \widehat{\boldsymbol{\beta}} = \begin{bmatrix} 0 \\ 0.5 \end{bmatrix}$$



Data points denoted by "x"

(b) Given 4 data points

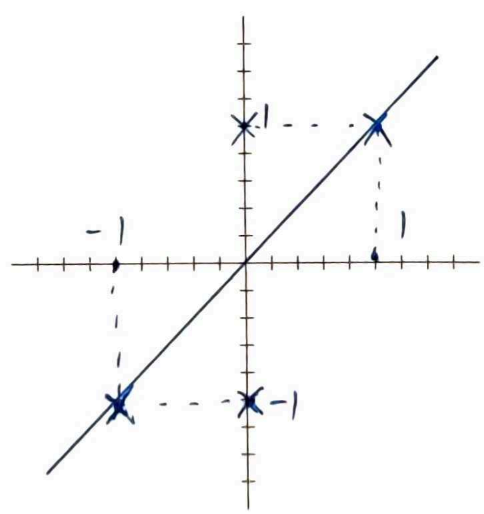
$$(x_i, y_i) \in \{(1,1), (0,1), (0,-1), (-1,-1)\},$$

find the vector $\widehat{\boldsymbol{\beta}} = \begin{bmatrix} \widehat{\beta}_0 \\ \widehat{\beta}_1 \end{bmatrix}$ such that

$$\widehat{\boldsymbol{\beta}} = \underset{\beta=[\beta_0,\beta_1]^\top \in \mathbb{R}^2}{\arg\min} \sum_{i=1}^{4} (y_i - \beta_0 - \beta_1 x_i)^2 .$$

You need to set up the least-squares problem in the matrix form by providing the design matrix $\boldsymbol{X}$ and the target vector $\boldsymbol{y}$ such that $\boldsymbol{y} \approx \boldsymbol{X}\widehat{\boldsymbol{\beta}}$. Sketch the linear regression line $\hat{f}(x) = \widehat{\beta}_0 + \widehat{\beta}_1 x$ on the axes below along with the 4 data points.

$$\boldsymbol{X} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & -1 \end{bmatrix} \quad \boldsymbol{y} = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} \quad \widehat{\boldsymbol{\beta}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$



Data points denoted by "x"