

HW2

2020年10月13日 13:59



HW2

MATH 3332 Data Analytic Tools

Homework 2

Due date: 19 October, 6pm, Monday

1. Let V be a vector space, and $\langle \cdot, \cdot \rangle$ be an inner product on V . Use the definition of inner product to prove the following.
 - (a) Prove that $\langle \mathbf{0}, \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{0} \rangle = 0$ for any $\mathbf{x} \in V$. Here $\mathbf{0}$ is the zero vector in V .
 - (b) Prove that the second condition

$$\langle \alpha \mathbf{x}_1 + \beta \mathbf{x}_2, \mathbf{y} \rangle = \alpha \langle \mathbf{x}_1, \mathbf{y} \rangle + \beta \langle \mathbf{x}_2, \mathbf{y} \rangle, \quad \forall \mathbf{x}_1, \mathbf{x}_2, \mathbf{y} \in V, \alpha, \beta \in \mathbb{R}$$

is equivalent to

$$\langle \mathbf{x}_1 + \mathbf{x}_2, \mathbf{y} \rangle = \langle \mathbf{x}_1, \mathbf{y} \rangle + \langle \mathbf{x}_2, \mathbf{y} \rangle \quad \text{and} \quad \langle \alpha \mathbf{x}, \mathbf{y} \rangle = \alpha \langle \mathbf{x}, \mathbf{y} \rangle, \quad \forall \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}, \mathbf{y} \in V, \alpha \in \mathbb{R}.$$

2. Let V be a vector space with a norm $\|\cdot\|$ that satisfies the parallelogram identity

$$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2, \quad \forall \mathbf{x}, \mathbf{y} \in V.$$

Note that we don't have an inner product on V so far. For any $\mathbf{x}, \mathbf{y} \in V$, define

$$f(\mathbf{x}, \mathbf{y}) := \frac{1}{2} (\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2)$$

Obviously, for all $\mathbf{x}, \mathbf{y} \in V$, we have $f(\mathbf{x}, \mathbf{x}) \geq 0$ and $f(\mathbf{x}, \mathbf{y}) = f(\mathbf{y}, \mathbf{x})$. Also, $f(\mathbf{x}, \mathbf{x}) = 0$ if and only if $\mathbf{x} = \mathbf{0}$.

- (a) Prove $f(\mathbf{x} + \mathbf{y}, \mathbf{z}) = f(\mathbf{x}, \mathbf{z}) + f(\mathbf{y}, \mathbf{z})$ for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$.
- (b) Prove $f(-\mathbf{x}, \mathbf{y}) = -f(\mathbf{x}, \mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in V$.
- (c) Prove $(f(\mathbf{x}, \mathbf{y}))^2 \leq f(\mathbf{x}, \mathbf{x})f(\mathbf{y}, \mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in V$.
- (d) (*Bonus question!*) Prove that $f(\mathbf{x}, \mathbf{y})$ is an inner product on V whose induced norm is $\|\cdot\|$. (*Hint: From Q1(b) and Q2(a), it suffices to prove $f(\alpha \mathbf{x}, \mathbf{y}) = \alpha f(\mathbf{x}, \mathbf{y})$ for all $\alpha \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y} \in V$. You first prove this identity for rational α , and then use limit and Q1(c) to show it for any real α .*)

This question proved that the parallelogram identity is also a sufficient condition for a norm to be induced by an inner product. Combined with the parallelogram law on inner product spaces, we see that the parallelogram identity is a necessary and sufficient condition for a norm to be an induced by an inner product.

3. Consider an inner product space V with the induced norm. Let $X = \{\mathbf{x}_1, \dots, \mathbf{x}_N\} \subset V$ be a set of vectors in V with $\|\mathbf{x}_i\| = 1$ for all i . Given a vector $\mathbf{y} \in V$ with $\|\mathbf{y}\| = 1$, show that the following two things are the same:

- finding the vector in X that has the smallest distance to \mathbf{y} (i.e., solving $\min_{\mathbf{x} \in X} \|\mathbf{x} - \mathbf{y}\|$)

- finding the vector in X that has the smallest angle to \mathbf{y} (i.e., solving $\min_{\mathbf{x} \in X} \arccos(\langle \mathbf{x}, \mathbf{y} \rangle)$)
4. Consider the polynomial kernel $K(\mathbf{x}, \mathbf{y}) = (\mathbf{x}^T \mathbf{y})^2$ for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$. Find an explicit feature map $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ satisfying $\langle \phi(\mathbf{x}), \phi(\mathbf{y}) \rangle = K(\mathbf{x}, \mathbf{y})$, where the inner product the standard inner product in \mathbb{R}^3 .
5. (*You don't need to do anything for this question.*) A good Matlab code and demonstration of kernel K-means can be found at
<http://www.dcs.gla.ac.uk/~srogers/firstcourseml/matlab/chapter6/kernelkmeans.html>
Read the code. Run the code in Matlab, if possible, to see how kernel K-means works for nonlinear data.

1. Let V be a vector space, and $\langle \cdot, \cdot \rangle$ be an inner product on V . Use the definition of inner product to prove the following.
- Prove that $\langle \mathbf{0}, \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{0} \rangle = 0$ for any $\mathbf{x} \in V$. Here $\mathbf{0}$ is the zero vector in V .
 - Prove that the second condition

$$\langle \alpha \mathbf{x}_1 + \beta \mathbf{x}_2, \mathbf{y} \rangle = \alpha \langle \mathbf{x}_1, \mathbf{y} \rangle + \beta \langle \mathbf{x}_2, \mathbf{y} \rangle, \quad \forall \mathbf{x}_1, \mathbf{x}_2, \mathbf{y} \in V, \alpha, \beta \in \mathbb{R}$$

is equivalent to

$$\langle \mathbf{x}_1 + \mathbf{x}_2, \mathbf{y} \rangle = \langle \mathbf{x}_1, \mathbf{y} \rangle + \langle \mathbf{x}_2, \mathbf{y} \rangle \quad \text{and} \quad \langle \alpha \mathbf{x}, \mathbf{y} \rangle = \alpha \langle \mathbf{x}, \mathbf{y} \rangle, \quad \forall \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}, \mathbf{y} \in V, \alpha \in \mathbb{R}.$$

|a). let $u \in V$, and $-u \in V$

$$\begin{aligned} \Rightarrow \langle u, \mathbf{x} \rangle &= \langle u - u, \mathbf{x} \rangle \\ &= \langle u, \mathbf{x} \rangle + \langle -u, \mathbf{x} \rangle \quad (\text{By linearity of inner product}) \\ &= \langle u, \mathbf{x} \rangle - \langle u, \mathbf{x} \rangle \\ &= 0 \end{aligned}$$

$$\therefore \langle 0, \mathbf{x} \rangle = 0$$

$$\Rightarrow \langle \mathbf{x}, 0 \rangle = \langle 0, \mathbf{x} \rangle = 0 \quad (\text{commutativity of inner product})$$

- (b) Prove that the second condition

|b)-

$$\langle \alpha \mathbf{x}_1 + \beta \mathbf{x}_2, \mathbf{y} \rangle = \alpha \langle \mathbf{x}_1, \mathbf{y} \rangle + \beta \langle \mathbf{x}_2, \mathbf{y} \rangle, \quad \forall \mathbf{x}_1, \mathbf{x}_2, \mathbf{y} \in V, \alpha, \beta \in \mathbb{R}$$

is equivalent to

$$\langle \mathbf{x}_1 + \mathbf{x}_2, \mathbf{y} \rangle = \langle \mathbf{x}_1, \mathbf{y} \rangle + \langle \mathbf{x}_2, \mathbf{y} \rangle \quad \text{and} \quad \langle \alpha \mathbf{x}, \mathbf{y} \rangle = \alpha \langle \mathbf{x}, \mathbf{y} \rangle, \quad \forall \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}, \mathbf{y} \in V, \alpha \in \mathbb{R}.$$

given $\langle \alpha \mathbf{x}_1 + \beta \mathbf{x}_2, \mathbf{y} \rangle = \alpha \langle \mathbf{x}_1, \mathbf{y} \rangle + \beta \langle \mathbf{x}_2, \mathbf{y} \rangle$ -①, $\forall \alpha, \beta \in \mathbb{R}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{y} \in V$

putting $\alpha=1, \beta=1$ in ①, we have

$$\langle \mathbf{x}_1 + \mathbf{x}_2, \mathbf{y} \rangle = \langle \mathbf{x}_1, \mathbf{y} \rangle + \langle \mathbf{x}_2, \mathbf{y} \rangle ..$$

putting $\beta=0$ in ①, we have

$$\langle \alpha \mathbf{x}_1, \mathbf{y} \rangle = \alpha \langle \mathbf{x}_1, \mathbf{y} \rangle$$

Conversely,
given $\left\{ \begin{array}{l} \langle \mathbf{x}_1 + \mathbf{x}_2, \mathbf{y} \rangle = \langle \mathbf{x}_1, \mathbf{y} \rangle + \langle \mathbf{x}_2, \mathbf{y} \rangle - ② \\ \langle \alpha \mathbf{x}_1, \mathbf{y} \rangle = \alpha \langle \mathbf{x}_1, \mathbf{y} \rangle - ③ \end{array} \right.$
 $\forall \mathbf{x}_1, \mathbf{x}_2, \mathbf{y}, \alpha \in \mathbb{R}$

Let $\alpha, \beta \in \mathbb{R}$, $x_1, x_2 \in V$

$$\therefore \alpha x_1, \beta x_2 \in V$$

So, replacing x_1 by αx_1 and x_2 by βx_2 in ②

$$\langle \alpha x_1 + \beta x_2, y \rangle = \langle \alpha x_1, y \rangle + \langle \beta x_2, y \rangle$$

Using ③, we have

$$\langle \alpha x_1 + \beta x_2, y \rangle = \alpha \langle x_1, y \rangle + \beta \langle x_2, y \rangle$$

Hence, both statements are equivalent.

- Q2. 2. Let V be a vector space with a norm $\|\cdot\|$ that satisfies the parallelogram identity

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2, \quad \forall x, y \in V.$$

Note that we don't have an inner product on V so far. For any $x, y \in V$, define

$$f(x, y) := \frac{1}{2} (\|x+y\|^2 - \|x\|^2 - \|y\|^2)$$

Obviously, for all $x, y \in V$, we have $f(x, x) \geq 0$ and $f(x, y) = f(y, x)$. Also, $f(x, x) = 0$ if and only if $x = \mathbf{0}$.

- (a) Prove $f(x+y, z) = f(x, z) + f(y, z)$ for all $x, y, z \in V$.
- (b) Prove $f(-x, y) = -f(x, y)$ for all $x, y \in V$.
- (c) Prove $(f(x, y))^2 \leq f(x, x)f(y, y)$ for all $x, y \in V$.
- (d) (Bonus question!) Prove that $f(x, y)$ is an inner product on V whose induced norm is $\|\cdot\|$. (Hint: From Q1(b) and Q2(a), it suffices to prove $f(\alpha x, y) = \alpha f(x, y)$ for all $\alpha \in \mathbb{R}$ and $x, y \in V$. You first prove this identity for rational α , and then use limit and Q1(c) to show it for any real α .)

This question proved that the parallelogram identity is also a sufficient condition for a norm to be induced by an inner product. Combined with the parallelogram law on inner product spaces, we see that the parallelogram identity is a necessary and sufficient condition for a norm to be an induced by an inner product.

$$\text{given } f(x, y) = \frac{1}{2} (\|x+y\|^2 - \|x\|^2 - \|y\|^2)$$

By parallelogram inequality, we have:

$$f(x, y) = \frac{1}{2} (2\|x\|^2 + 2\|y\|^2 - \|x+y\|^2 - \|x-y\|^2) \quad \text{--- ①}$$

$$f(x, y) = \frac{1}{2} (\|x\|^2 + \|y\|^2 - \|x-y\|^2) \quad \text{--- ①}$$

a). let $x, y, z \in V$,

$$f(x+y, z) = \frac{1}{2} (\|(x+y)+z\|^2 - \|x+y\|^2 - \|z\|^2)$$

$$f(x, z) = \frac{1}{2} (\|x+z\|^2 - \|x\|^2 - \|z\|^2)$$

$$\text{or...} = \frac{1}{2} (\|y+z\|^2 - \|y\|^2 - \|z\|^2)$$

$$f(y, z) = \frac{1}{2}(\|y+z\|^2 - \|y\|^2 - \|z\|^2)$$

$$f(y, z) = \frac{1}{2}(\|y\|^2 + \|z\|^2 - \|y-z\|^2) \quad (\text{By } ①)$$

for $f(x, z) + f(y, z)$,

$$f(x, z) + f(y, z) = \frac{1}{2}[\|x+z\|^2 - \|x\|^2 - \|z\|^2 + \|y\|^2 + \|z\|^2 - \|y-z\|^2]$$

$$= \frac{1}{2}[(\|x+z\|^2 + \|y\|^2) - (\|y-z\|^2 + \|x\|^2)]$$

$$f(x, z) + f(y, z) = \frac{1}{2} \left[\frac{1}{2}(\|(x+z)+y\|^2 + \|(x+z)-y\|^2) - \frac{1}{2}(\|(y-z)+x\|^2 + \|(y-z)-x\|^2) \right]$$

$$f(x, z) + f(y, z) = \frac{1}{4} [\|x+y+z\|^2 + \|x+z-y\|^2 - \|x+y-z\|^2 - \|y-z-x\|^2]$$

$$f(x, z) + f(y, z) = \frac{1}{4} [\|x+y+z\|^2 + \|y+z-y\|^2 - \|x+y-z\|^2 - \|x+z-y\|^2]$$

$$f(x, z) + f(y, z) = \frac{1}{4} [\|x+y+z\|^2 - \|x+y-z\|^2]$$

$$f(x, z) + f(y, z) = \frac{1}{4} [2\|x+y+z\|^2 - 2(\|x+y\|^2 + \|z\|^2)]$$

$$f(x, z) + f(y, z) = \frac{1}{2}[\|(x+y)+z\|^2 - \|(x+y)-z\|^2]$$

$$f(x, z) + f(y, z) = f(x+y, z),$$

b)- $f(-x, y) = \frac{1}{2}(\|-x+y\|^2 - \|-x\|^2 - \|y\|^2)$

$$= \frac{1}{2}(\|y-x\|^2 - \|x\|^2 - \|y\|^2)$$

$$= \frac{1}{2}(\|x\|^2 + \|y\|^2 - \|y+x\|^2) \quad (\text{By parallelogram equality})$$

$$= -\frac{1}{2}(\|x+y\|^2 - \|x\|^2 - \|y\|^2)$$

$$= -f(x, y),$$

c)- $f(x, x) = \frac{1}{2}(\|2x\|^2 - \|x\|^2 - \|x\|^2)$

$$= \frac{1}{2}(4\|x\|^2 - 2\|x\|^2)$$

$$= 2\|x\|^2$$

$$\begin{aligned} f(y, y) &= \frac{1}{2} (\|2y\|^2 - \|y\|^2 - \|y\|^2) \\ &= \frac{1}{2} (4\|y\|^2 - 2\|y\|^2) \\ &= \|y\|^2 \end{aligned}$$

For $f(x, y)^2$,

$$\begin{aligned} f(x, y)^2 &= \frac{1}{4} (\|x+y\|^2 - \|x\|^2 - \|y\|^2)^2 \\ f(x, y)^2 &\leq \frac{1}{4} ((\|x\| + \|y\|)^2 - \|x\|^2 - \|y\|^2)^2 \\ &\quad (\text{By triangle inequality}) \\ &= \frac{1}{4} (\|x\|^2 + 2\|x\|\|y\| + \|y\|^2 - \|x\|^2 - \|y\|^2)^2 \\ &= \frac{1}{4} (4\|x\|\|y\|) \\ &= \|x\|\|y\| \\ &= f(x, x) f(y, y) \end{aligned}$$

$$\therefore (f(x, y))^2 \leq f(x, x) f(y, y) \cdot \forall x, y \in V$$

- d). (Bonus question!) Prove that $f(\mathbf{x}, \mathbf{y})$ is an inner product on V whose induced norm is $\|\cdot\|$. (Hint: From Q1(b) and Q2(a), it suffices to prove $f(\alpha \mathbf{x}, \mathbf{y}) = \alpha f(\mathbf{x}, \mathbf{y})$ for all $\alpha \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y} \in V$. You first prove this identity for rational α , and then use limit and Q1(c) to show it for any real α .)

Showing $f(\alpha \mathbf{x}, \mathbf{y}) = \alpha f(\mathbf{x}, \mathbf{y})$. $\forall \alpha \in \mathbb{R}$,

from (a), we have
 $f(\mathbf{x} + \mathbf{y}, z) = f(\mathbf{x}, z) + f(\mathbf{y}, z) \quad \forall x, y, z \in V$

let $n \in \mathbb{N}$, from (a), we have

$$\begin{aligned} \therefore f(n \cdot \mathbf{x}, \mathbf{y}) &= f((n-1)\mathbf{x}, \mathbf{y}) + f(\mathbf{x}, \mathbf{y}) \\ f(n \cdot \mathbf{x}, \mathbf{y}) &= f((n-2)\mathbf{x}, \mathbf{y}) + 2f(\mathbf{x}, \mathbf{y}) \\ &\vdots \\ f(n \cdot \mathbf{x}, \mathbf{y}) &= nf(\mathbf{x}, \mathbf{y}), \quad \forall n \in \mathbb{N}, \forall x, y \in V \end{aligned}$$

-②

let $m \in \mathbb{N}$,

$$\text{Then } mf\left(\frac{1}{m}x, y\right) = f((m \cdot \frac{1}{m})x, y) \quad (\text{from ②})$$

$$mf\left(\frac{1}{m}x, y\right) = f(x, y)$$

$$\Rightarrow f\left(\frac{1}{m}x, y\right) = \frac{f(x, y)}{m} \quad -③$$

from ① & ③, we have:

$$f\left(\frac{n}{m}x, y\right) = \frac{n}{m}f(x, y) \quad \forall m, n \in \mathbb{N} \quad \forall x, y \in V.$$

from (b), we have

$$f(-\frac{1}{m}x, y) = -\frac{1}{m}f(x, y)$$

$$\therefore f(qx, y) = qf(x, y) \quad -④ \quad \begin{matrix} q \in \mathbb{Q}: \{\text{set of rationals}\}, \\ \forall x, y \in V \end{matrix}$$

Then, let $\alpha \in \mathbb{R}$, $\because \mathbb{Q}$ is dense in \mathbb{R}

$\therefore \exists \text{ seq}^{\text{q}} \text{ in } \mathbb{Q}, \text{ say } \{q_n\}_n, \text{ s.t. } \lim_{n \rightarrow \infty} q_n \rightarrow \alpha$.

Then, for $f(q_n x, y)$,

$$f(q_n x, y) = q_n f(x, y) \quad \begin{matrix} (\text{from ④}), \\ \forall n. \end{matrix}$$

$$\lim_{n \rightarrow \infty} f(q_n x, y) = \lim_{n \rightarrow \infty} q_n f(x, y)$$

$\because f: \mathbb{R}^n \rightarrow \mathbb{R}$ is a continuous function,

$$\therefore f\left(\lim_{n \rightarrow \infty} q_n x, y\right) = \alpha f(x, y)$$

$$f(\alpha x, y) = \alpha f(x, y) //$$

3. Consider an inner product space V with the induced norm. Let $X = \{x_1, \dots, x_N\} \subset V$ be a set of vectors in V with $\|x_i\| = 1$ for all i . Given a vector $y \in V$ with $\|y\| = 1$, show that the following two things are the same:

- finding the vector in X that has the smallest distance to y (i.e., solving $\min_{x \in X} \|x - y\|$)

- finding the vector in X that has the smallest angle to y (i.e., solving $\min_{x \in X} \arccos(x, y)$)

Let $X = \{x_1, \dots, x_N\} \subset V$ and
given vector $x \in X$ and $y \in V$ s.t. $\|y\| = 1$, $\|x\| = 1$

$$\|x-y\| = \sqrt{\langle x-y, x-y \rangle}$$

$$\|x-y\|^2 = \langle x-y, x-y \rangle \quad (\text{Solving } \min_{x \in X} \|x-y\| = \text{solving } \min_{x \in X} \|x-y\|^2)$$

$$\|x-y\|^2 = \langle x, x \rangle - 2\langle x, y \rangle + \langle y, y \rangle$$

$$\|x-y\|^2 = \|x\|^2 - 2\langle x, y \rangle + \|y\|^2$$

$$\|x-y\|^2 = 1 - 2\langle x, y \rangle$$

$$\therefore \text{solving } \min_{x \in X} \|x-y\|^2 = \text{solving } \max_{x \in X} \langle x, y \rangle$$

Let $A(x, y)$ be the angle of x to y , $x \in X, y \in V$.

??
..

4.

4. Consider the polynomial kernel $K(x, y) = (x^T y)^2$ for $x, y \in \mathbb{R}^2$. Find an explicit feature map $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ satisfying $\langle \phi(x), \phi(y) \rangle = K(x, y)$, where the inner product is the standard inner product in \mathbb{R}^3 .

Kernel function $K(x, y)$ satisfies $K(x, y) = \langle \phi(x), \phi(y) \rangle$.

$$K(x, y) = (x^T y)^2$$

$$K(x, y) = (x_1 y_1 + x_2 y_2)^2$$

$$K(x, y) = (x_1^2 y_1^2 + 2x_1 x_2 y_1 y_2 + x_2^2 y_2^2)$$

so, it is inner product with two vector

$$(x_1^2, \sqrt{2}x_1 x_2, x_2^2) \text{ and } (y_1^2, \sqrt{2}y_1 y_2, y_2^2)$$

$$\therefore \phi(x) = \begin{pmatrix} x_1^2 \\ \sqrt{2}x_1 x_2 \\ x_2^2 \end{pmatrix} \text{ and similar for } y$$