MSBD 5004 Mathematical Methods for Data Analysis
Homework 2

Due date: October 14, Friday

1. Let $(V, \|\cdot\|)$ be a normed vector space.

   (a) Prove that, for all $\boldsymbol{x}, \boldsymbol{y} \in V$,
   $$|\|\boldsymbol{x}\| - \|\boldsymbol{y}\|| \leq \|\boldsymbol{x} - \boldsymbol{y}\|.$$

   (b) Let $\{\boldsymbol{x}_k\}_{k\in\mathbb{N}}$ be a convergent sequence in V with limit $\boldsymbol{x} \in V$. Prove that
   $$\lim_{k\to\infty} \|\boldsymbol{x}_k\| = \|\boldsymbol{x}\|.$$
   (*Hint: Use part (a).*)

   (c) Let $\{\boldsymbol{x}^{(k)}\}_{k\in\mathbb{N}}$ be a sequence in V and $\boldsymbol{x}, \boldsymbol{y} \in V$. Prove that, if
   $$\boldsymbol{x}^{(k)} \to \boldsymbol{x}, \quad \text{and} \quad \boldsymbol{x}^{(k)} \to \boldsymbol{y},$$
   then $\boldsymbol{x} = \boldsymbol{y}$. (In other words, the limit of the same sequence in a normed vector space is unique.)

2. 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 $\langle \boldsymbol{0}, \boldsymbol{x}\rangle = \langle \boldsymbol{x}, \boldsymbol{0}\rangle = 0$ for any $\boldsymbol{x} \in V$. Here $\boldsymbol{0}$ is the zero vector in V.

   (b) Prove that the second condition
   $$\langle \alpha\boldsymbol{x}_1 + \beta\boldsymbol{x}_2, \boldsymbol{y}\rangle = \alpha\langle \boldsymbol{x}_1, \boldsymbol{y}\rangle + \beta\langle \boldsymbol{x}_2, \boldsymbol{y}\rangle, \qquad \forall\ \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y} \in V\ , \alpha, \beta \in \mathbb{R}$$
   is equivalent to
   $$\langle \boldsymbol{x}_1 + \boldsymbol{x}_2, \boldsymbol{y}\rangle = \langle \boldsymbol{x}_1, \boldsymbol{y}\rangle + \langle \boldsymbol{x}_2, \boldsymbol{y}\rangle \quad \text{and} \quad \langle \alpha\boldsymbol{x}, \boldsymbol{y}\rangle = \alpha\langle \boldsymbol{x}, \boldsymbol{y}\rangle, \qquad \forall\ \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}, \boldsymbol{y} \in V,\ \alpha \in \mathbb{R}.$$

3. $\mathbb{R}^{m\times n}$ is a vector space over $\mathbb{R}$. Show that $\langle \boldsymbol{A}, \boldsymbol{B}\rangle = \text{trace}(\boldsymbol{A}^T\boldsymbol{B})$ for $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{m\times n}$ is indeed an inner product on $\mathbb{R}^{m\times n}$. Here $\text{trace}(\cdot)$ is the trace of a matrix, i.e., the sum of all diagonal entries.

4. Let V be a vector space with a norm $\|\cdot\|$ that satisfies the parallelogram identity
   $$\|\boldsymbol{x} + \boldsymbol{y}\|^2 + \|\boldsymbol{x} - \boldsymbol{y}\|^2 = 2\|\boldsymbol{x}\|^2 + 2\|\boldsymbol{y}\|^2, \qquad \forall\ \boldsymbol{x}, \boldsymbol{y} \in V.$$
   Note that we don't have an inner product on V so far. For any $\boldsymbol{x}, \boldsymbol{y} \in V$, define
   $$f(\boldsymbol{x}, \boldsymbol{y}) := \frac{1}{2}\left(\|\boldsymbol{x} + \boldsymbol{y}\|^2 - \|\boldsymbol{x}\|^2 - \|\boldsymbol{y}\|^2\right)$$

   (a) Prove $f(\boldsymbol{x}, \boldsymbol{x}) \geq 0$ for any $\boldsymbol{x} \in V$, and $f(\boldsymbol{x}, \boldsymbol{x}) = 0$ if and only if $\boldsymbol{x} = \boldsymbol{0}$.

   (b) Prove $f(\boldsymbol{x}, \boldsymbol{y}) = f(\boldsymbol{y}, \boldsymbol{x})$ for all $\boldsymbol{x}, \boldsymbol{y} \in V$.

(c) Prove $f(\boldsymbol{x}+\boldsymbol{y}, \boldsymbol{z}) = f(\boldsymbol{x}, \boldsymbol{z}) + f(\boldsymbol{y}, \boldsymbol{z})$ for all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V$.

(d) Prove $f(-\boldsymbol{x}, \boldsymbol{y}) = -f(\boldsymbol{x}, \boldsymbol{y})$ for all $\boldsymbol{x}, \boldsymbol{y} \in V$.

(e) Prove $(f(\boldsymbol{x}, \boldsymbol{y}))^2 \leq f(\boldsymbol{x}, \boldsymbol{x}) f(\boldsymbol{y}, \boldsymbol{y})$ for all $\boldsymbol{x}, \boldsymbol{y} \in V$.

(c)(d)(e) together with some other technique can show that $f(\alpha\boldsymbol{x} + \beta\boldsymbol{y}, \boldsymbol{z}) = \alpha f(\boldsymbol{x}, \boldsymbol{z}) + \beta f(\boldsymbol{y}, \boldsymbol{z})$. Therefore, we can finally prove f defines an inner product. This question showed 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.

5. Suppose that K_1 and K_2 are two kernel functions, and suppose that $a > 0$ is a constant. Prove that the following functions are also kernel functions:

   (a) $K(\boldsymbol{x}, \boldsymbol{y}) = K_1(\boldsymbol{x}, \boldsymbol{y}) + K_2(\boldsymbol{x}, \boldsymbol{y})$;

   (b) $K(\boldsymbol{x}, \boldsymbol{y}) = a \cdot K_1(\boldsymbol{x}, \boldsymbol{y})$.

6. Consider the polynomial kernel function $K : (\mathbb{R}^2, \mathbb{R}^2) \to \mathbb{R}$ defined by $K(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}^T\boldsymbol{y} + 1)^2$. The associated feature map ϕ and the feature space H are given explicitly as in the following

$$\phi : \boldsymbol{x} = (x_1, x_2) \to \phi(\boldsymbol{x}) = (\sqrt{2}x_1, \sqrt{2}x_2, x_1^2, x_2^2, \sqrt{2}x_1x_2, 1) \in \mathbb{R}^6 := H.$$

   This feature map takes the data from a two-dimensional to a six-dimensional space in a way that linear relations in the feature space correspond to quadratic relations in the input space. Prove that indeed ϕ and H satisfies

$$K(\boldsymbol{x}, \boldsymbol{y}) = \langle \phi(\boldsymbol{x}), \phi(\boldsymbol{y}) \rangle,$$

   and prove K is symmetric positive semi-definite (SPSD).

7. If $\boldsymbol{A}$ is a $d \times d$ symmetric positive semi-definite (SPSD) matrix, then the function $K : (\mathbb{R}^d, \mathbb{R}^d) \to \mathbb{R}$ given by $K(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ is a kernel function.