MSBD 5004 Mathematical Methods for Data Analysis
Homework 2

Due date: 13 March, 9pm, Wednesday

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

2. Consider an inner product space V with the induced norm. Let $X = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\} \subset V$ be a set of vectors in V with $\|\boldsymbol{x}_i\| = 1$ for all i. Given a vector $\boldsymbol{y} \in V$ with $\|\boldsymbol{y}\| = 1$, show that the following two things are the same:
   - finding the vector in X that has the smallest distance to $\boldsymbol{y}$ (i.e., solving $\min_{\boldsymbol{x}\in X} \|\boldsymbol{x} - \boldsymbol{y}\|$)
   - finding the vector in X that has the smallest angel to $\boldsymbol{y}$ (i.e., solving $\min_{\boldsymbol{x}\in X} \arccos\langle \boldsymbol{x}, \boldsymbol{y}\rangle$)

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

4. Determine whether each of the following functions of vectors in $\mathbb{R}^n$ is linear. If it is a linear function, give its inner product representation, i.e., an vector $\boldsymbol{a} \in \mathbb{R}^n$ for which $f(x) = \langle \boldsymbol{a}, \boldsymbol{x}\rangle$ for all $\boldsymbol{x} \in \mathbb{R}^n$. If it is not linear, give specific $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$ for which superposition fails, i.e., $f(\alpha\boldsymbol{x} + \beta\boldsymbol{x}) \neq \alpha f(\boldsymbol{x}) + \beta f(\boldsymbol{y})$.
   (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, $f(\boldsymbol{x}) = x_n - x_1$.
   (c) The median of a vector, where we will assume $n = 2k + 1$ is odd. The median of the vector $\boldsymbol{x}$ is defined as the $(k + 1)$-st largest number among the entries of $\boldsymbol{x}$. For example, the median of $(7.1, 3.2, 1.5)$ is 1.5.
   (d) Vector extrapolation, defined as $x_n + (x_n - x_{n-1})$, for $n \geq 2$. (This is a simple prediction of what x_{n+1} would be, based on a straight line drawn through x_n and x_{n-1}.)