Math 5490 – Homework 2: Due Feb 8 by 11:59pm

Instructions:

- Complete the problems below, and submit your solutions and Python code by uploading them to the Google form: https://forms.gle/zSteKSYk1KztzUT1A
- Submit all your Python code in a single .py file using the function templates given in each problem. I will import your functions from this file and test your code.
- If you use LaTeX to write up your solutions, upload them as a pdf file. Students who use LaTeX to write up their solutions will receive bonus points on the homework assignment (equivalent to 1/3 of a letter grade bump).
- If you choose to handwrite your solutions and scan them, please either use a real scanner, or use a smartphone app that allows scanning with you smartphone camera. It is not acceptable to submit photos of your solutions, as these can be hard to read.

Problems:

For some problems, we need the following definition.

Definition 1. The *Hadamard product* between two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, denoted as $\mathbf{x} \circ \mathbf{y}$, is the vector $\mathbf{z} \in \mathbb{R}^n$ with entries $z_i = x_i y_i$. Similarly, the *Hadamard product* of two matrices $A, B \in \mathbb{R}^{n\times m}$, denoted again as $A \circ B$, is the matric $C \in \mathbb{R}^{n\times m}$ with entries $C_{ij} = A_{ij}B_{ij}$.

The Hadamard product of two matrices or two vectors is simply the elementwise product. The reader should be careful to note that the Hadamard product is not the same as the matrix product. The Hadamard product applies only to vectors or matrices of the same size, and produces a vector or matrix of the same size. The Hadamard product is also clearly commutative, that is, $A \circ B = B \circ A$ and $\mathbf{x} \circ \mathbf{y} = \mathbf{y} \circ \mathbf{x}$, and satisfies the usual properties of multiplication (e.g., associativity).

1. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Show that
$$(\mathbf{x}\mathbf{x}^T) \circ (\mathbf{y}\mathbf{y}^T) = (\mathbf{x} \circ \mathbf{y})(\mathbf{x} \circ \mathbf{y})^T. \tag{1}$$
[Hint: Compare the components of the matrices on each side.]

2. Let A, B be symmetric and positive semidefinite matrices. Prove that $A \circ B$ is also symmetric and positive semidefinite. This is known as the *Schur Product Theorem*. [Hint: Since A and B are symmetric, they can be diagonalized and written in the form
$$A = \sum_{i=1}^{n} \lambda_i \mathbf{u}_i \mathbf{u}_i^T \quad \text{and} \quad B = \sum_{i=1}^{n} \mu_i \mathbf{v}_i \mathbf{v}_i^T,$$
where $\mathbf{u}_1, \ldots, \mathbf{u}_n$ and $\lambda_1, \ldots, \lambda_n$ are the orthonormal eigenvectors and eigenvalues of A, while $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and $\mu_1, \ldots, \mu_n$ are those of B. Continue from here, writing out $A \circ B$ using the expressions for A and B above, and then use part 1.]

3. Let A be positive definite and symmetric and let $\mathbf{x}$ be the solution of the linear system

$$A\mathbf{x} = \mathbf{b}.$$

Let $0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ be the eigenvalues of A and $\mathbf{v}_1, \ldots, \mathbf{v}_n$ the corresponding orthonormal eigenvectors.

(i) Show that $\mathbf{x}$ can be written as

$$\mathbf{x} = \sum_{i=1}^{n} \lambda_i^{-1} (\mathbf{b} \cdot \mathbf{v}_i) \mathbf{v}_i.$$

[Hint: Show by direct computation that $A\mathbf{x} = \mathbf{b}$.]

(ii) Define the spectrally truncated approximate solution

$$\mathbf{x}_N = \sum_{i=1}^{N} \lambda_i^{-1} (\mathbf{b} \cdot \mathbf{v}_i) \mathbf{v}_i. \tag{2}$$

Show that

$$\|\mathbf{x}_N - \mathbf{x}\| \leq \lambda_{N+1}^{-1} \|\mathbf{b}\|.$$

[Comment: In practice, it is often tractable to compute a few of the eigenvectors of A, say the first N, where $N \ll n$, even if computing all eigenvectors is intractable. In this case, the spectrally truncated approximate solution (2) can be an efficient way to solve $A\mathbf{x} = \mathbf{b}$, provided N is chosen so that λ_{N+1} is large.]

4. Assume we are in dimension $n = 1$. Show that the radial basis function kernel $\mathcal{K}(x, y) = e^{-\gamma(x-y)^2}$ can be expressed as

$$\mathcal{K}(x, y) = \sum_{m=0}^{\infty} \varphi_m(x) \cdot \varphi_m(y)$$

where the functions (i.e., feature maps) $\varphi_m : \mathbb{R} \to \mathbb{R}$ are given by

$$\varphi_m(x) = e^{-\gamma x^2} x^m \sqrt{\frac{(2\gamma)^m}{m!}}.$$

[Hint: Write

$$\mathcal{K}(x, y) = e^{-\gamma x^2} e^{-\gamma y^2} e^{2\gamma xy},$$

and use a Taylor expansion on the last exponential term.]