AMATH 563 - Homework 1 (Theory)

Nate Whybra

May 2025

Problem 1

As Γ is a PDS kernel, if $X = \{x_1, x_2\}$, then we we have for all vectors $y \in \mathbb{R}^2$, with $k_{ij} = \Gamma(x_i, x_j)$:

$$\begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \geq 0$$

By multiplying everything out, and noting that $k_{12} = k_{21}$, we see:

$$(k_{11})y_1^2 + (2k_{12})y_1y_2 + (k_{22})y_2^2 \geq 0$$

Now without loss of generality assume that $k_{22} \geq 0$ andlet $y_1 = \sqrt{k_{22}}$ and $y_2 = -k_{11}/\sqrt{k_{22}}$. If $k_{22} < 0$, then $\sqrt{k_{22}}$ will be a complex number (which is not allowed). However we can easily mitigate this by saying that in the case that $k_{22} < 0$, to just factor -1 from the whole expression, and then just proceed with the above definitions. With this in mind, the above becomes:

$$k_{11}k_{22} - 2k_{12}^2 + k_{12}^2 \geq 0$$

$$\implies k_{12}^2 \leq k_{11}k_{22}$$

Or equivalently, as desired:

$$\Gamma^2(x_1, x_2) \leq \Gamma(x_1, x_1)\Gamma(x_2, x_2)$$

Continuing, for functions of the form $f(u) = c^T K(X, u)$ and $g(v) = b^T K(Y, v)$ we'd like to show that the following defines an inner product over $\mathbb{R}$:

$$\langle f, g \rangle_0 = c^T K(X, Y)b$$

To do so first suppose $a \in \mathbb{R}$. Then:

$$\langle af, g \rangle_0 = (ac^T)K(X, Y)b = a\left(c^T K(X, Y)b\right) = a\langle f, g \rangle$$

So h satisfies absolute homogeneity. Now take f_1, f_2 such that $f_1(u) = c_1^T K(X, u)$ and $f_2(u) = c_2^T K(X, u)$, then $(f_1 + f_2)(u) = \left(c_1^T + c_2^T\right) K(X, u) = (c_1 + c_2)^T K(X, u)$ and:

$$\langle f_1 + f_2, g \rangle_0 = (c_1 + c_2)^T K(X, Y)b$$

$$= c_1^T K(X,Y)b + c_2^T K(X,Y)b$$

$$= \langle f_1, g \rangle_0 + \langle f_2, g \rangle_0$$

So we see the distributivity property holds. Now consider:

$$\begin{aligned}
\langle g, f \rangle_0 &= b^T K(Y,X)c \\
&= b^T K(X,Y)c \quad \text{(K is symmetric)} \\
&= cK(X,Y)b^T \quad \text{(everything is real valued)} \\
&= \langle f, g \rangle_0
\end{aligned}$$

So we also see that $\langle \cdot, \cdot \rangle_0$ satisfies the symmetric property (where we have ignored the complex conjugate because everything here is real valued). Finally now suppose we have:

$$\langle f, f \rangle_0 = c^T K(X,X)c = \sum_{i,j=1}^{n} c_i c_j k_{ij} = 0$$

As $K(X,X)_{ij} = k(x_i, x_j)$, the inequality from the first part tells us that each element must be greater than or equal to 0. So unless K is is identically 0, the only way for the above expression to equal 0 is if $c = 0$. Therefore we have definitness, and since K is PDS we clearly have positive definitness. Therefore $\langle \cdot, \cdot \rangle$ defines an inner product.

Problem 2

$\hat{K}$ is obviously symmetric from the properties of K. Suppose $X = \{x_1, \cdots, x_n\}$, then $K := K(X, X)$ is a PDS matrix, and for all vectors $y \in \mathbb{R}^n$ we have:

$$y^T K y \geq 0$$

For now suppose that $k_{ii} > 0$. Define the matrix $D \in \mathbb{R}^{n \times n}$ such that $D_{ii} = 1/\sqrt{k_{ii}} > 0$. Then we have $\hat{K} = DKD$, so that:

$$S := y^T \hat{K} y = y^T DKDy = (Dy)^T K(Dy) \geq 0$$

D is full rank, and therefore is a bijection from $\mathbb{R}^n$ b $\mathbb{R}^n$, meaning for every vector $z \in \mathbb{R}^n$ there exists a $y \in \mathbb{R}^n$ sucj that $Dy = z$, and we get that $z^T \hat{K} z \geq 0$, meaning $\hat{K}$ is PDS. Now suppose that for any i, such that $k_{ii} = 0$, then from **Problem 1**, we have:

$$k_{ij}^2 \leq k_{ii} k_{jj} = 0 \implies k_{ij} = 0$$

Put simply, if any of the diagonal entries are 0, then that corresponding row is 0, and since K is symmetric, the corresponding column will also be 0. Thus that row and column will contribute nothing to the sum S from above, and $S \geq 0$. Therefore, K is PDS.

Problem 3

From lecture, we know that the linear kernel $K(x, x') = x^T x' + c$ is PDS for any $c > 0$. We also know that if we define a function:

$$K(x, x') = f(k(x, x'))$$

And f can be represented as a non-negative power series with non-negative coefficients, and k is also a PDS kernel, then K is also PDS. The polynomial kernel is PDS since we can take $f(u) = u^\alpha$ and g to be the linear kernel. The exponential kernel is PDS since we can take $f(u) = \exp(u) = \sum_{i \geq 0}(u^i/u!)$ and g to be the linear kernel. For the RBF kernel, we can write:

$$\begin{aligned}
K(x, x') &= \exp(-\gamma^2 \left\|x - x'\right\|_2^2) \\
&= \exp(-\gamma^2 (x - x')^T (x - x')) \\
&= \exp(-\gamma^2 (x^T x - 2x^T x' + x'^{\,T} x')) \\
&= \exp(-\gamma^2 x^T x) \exp(2\gamma^2 x^T x') \exp(-\gamma^2 x'^{\,T} x') \\
&= C \exp(a x^T x')
\end{aligned}$$

Where $C = \exp(-\gamma^2 x^T x) \exp(-\gamma^2 x'^{\,T} x')$ is a positive constant in $\mathbb{R}$, and $a = 2\gamma^2 > 0$ in $\mathbb{R}$. Now the RBF kernel is PDS since we can take $f(u) = C \exp(au)$ and g as the linear kernel, so we are done.

Problem 4

Let $X = \{x_1, \cdots, x_m\} \subset \mathbb{R}^d$, then for any vector $y \in \mathbb{R}^n$ we have:

$$\begin{aligned}
y^T K(X,X) y &= \sum_{i=1}^{m} \sum_{k=1}^{m} c_i c_k K(x_i, x_k) \\
&= \sum_{i=1}^{m} \sum_{k=1}^{m} c_i c_k \sum_{j=1}^{n} \lambda_j \psi_j(x_i) \psi_j(x_k) \\
&= \sum_{j=1}^{n} \lambda_j \sum_{i=1}^{m} \sum_{k=1}^{m} c_i c_k \psi_j(x_i) \psi_j(x_k) \\
&= \sum_{j=1}^{n} \lambda_j \left(\sum_{i=1}^{m} c_i \psi_j(x_i) \sum_{k=1}^{m} c_k \psi_j(x_k) \right) \\
&= \sum_{j=1}^{n} \lambda_j \left(\sum_{i=1}^{m} c_i \psi_j(x_i) \right)^2 \\
&:= \lambda_j d_j
\end{aligned}$$

Since the ψ functions are continuous and the sums are finite we were free to swap the summation order above. Now since each $\lambda_j, d_j \geq 0$, we see the quaratic form above is ≥ 0. The symmetric property is trivially satisfied as well. Therefore K is PDS.