53.2 Basic Properties of Positive Definite Kernels

Proposition 53.1 suggests a second approach to kernel functions which does not assume that a feature space and a feature map are provided. We will see in Section 53.3 that the two approaches are equivalent. The second approach is useful in practice because it is often difficult to define a feature space and a feature map in a simple manner.

Definition 53.2. Let X be a nonempty set. A function $\kappa\colon X \times X \to \mathbb{C}$ is a *positive definite kernel* if for every finite subset $S = \{x_1, \ldots, x_p\}$ of X, if K_S is the $p \times p$ matrix

$$K_S = (\kappa(x_j, x_i))_{1 \leq i,j \leq p}$$

called a *Gram matrix*, then we have

$$u^* K_S\, u = \sum_{i,j=1}^{p} \kappa(x_i, x_j) u_i \overline{u_j} \geq 0, \qquad \text{for all } u \in \mathbb{C}^p.$$

Observe that Definition 53.2 does not require that $u^* K_S\, u > 0$ if $u \neq 0$, so the terminology *positive definite* is a bit abusive, and it would be more appropriate to use the terminology *positive semidefinite*. However, it seems customary to use the term *positive definite kernel*, or even *positive kernel*.

Proposition 53.2. *Let $\kappa\colon X \times X \to \mathbb{C}$ be a positive definite kernel. Then $\kappa(x, x) \geq 0$ for all $x \in X$, and for any finite subset $S = \{x_1, \ldots, x_p\}$ of X, the $p \times p$ matrix K_S given by*

$$K_S = (\kappa(x_j, x_i))_{1 \leq i,j \leq p}$$

is Hermitian, that is, $K_S^* = K_S$.

Proof. The first property is obvious by choosing $S = \{x\}$. To prove that K_S is Hermitian, observe that we have

$$(u + v)^* K_S (u + v) = u^* K_S u + u^* K_S v + v^* K_S u + v^* K_S v,$$

and since $(u + v)^* K_S (u + v), u^* K_S u, v^* K_S v \geq 0$, we deduce that

$$2A = u^* K_S v + v^* K_S u \tag{1}$$

must be real. By replacing u by iu, we see that

$$2B = -iu^* K_S v + iv^* K_S u \tag{2}$$

must also be real. By multiplying Equation (2) by i and adding it to Equation (1) we get

$$u^* K_S v = A + iB. \tag{3}$$