Also, if $\kappa\colon X \times X \to \mathbb{R}$ is a real symmetric positive definite kernel, then we see immediately that Theorem 53.8 holds with H_0 a real Euclidean space and H a real Hilbert space.

⊛ **Remark:** If $X = G$, where G is a locally compact group, then a function $p\colon G \to \mathbb{C}$ (not necessarily continuous) is *positive semidefinite* if for all $s_1, \ldots, s_n \in G$ and all $\xi_1, \ldots, \xi_n \in \mathbb{C}$, we have

$$\sum_{j,k=1}^{n} p(s_j^{-1}s_k)\xi_k\overline{\xi_j} \geq 0.$$

So if we define $\kappa\colon G \times G \to \mathbb{C}$ by

$$\kappa(s,t) = p(t^{-1}s),$$

then κ is a positive definite kernel on G. If p is continuous, then it is known that p arises from a unitary representation $U\colon G \to \mathbf{U}(H)$ of the group G in a Hilbert space H with inner product $\langle -, - \rangle$ (a homomorphism with a certain continuity property), in the sense that there is some vector $x_0 \in H$ such that

$$p(s) = \langle U(s)(x_0), x_0 \rangle, \qquad \text{for all } s \in G.$$

Since the $U(s)$ are unitary operators on H,

$$\begin{aligned} p(t^{-1}s) &= \langle U(t^{-1}s)(x_0), x_0 \rangle = \langle U(t^{-1})(U(s)(x_0)), x_0 \rangle \\ &= \langle U(t)^*(U(s)(x_0)), x_0 \rangle = \langle U(s)(x_0), U(t)(x_0) \rangle, \end{aligned}$$

which shows that

$$\kappa(s,t) = \langle U(s)(x_0), U(t)(x_0) \rangle,$$

so the map $\varphi\colon G \to H$ given by

$$\varphi(s) = U(s)(x_0)$$

is a feature map into the feature space H. This theorem is due to Gelfand and Raikov (1943).

The proof of Theorem 53.8 is essentially identical to part of Godement's proof of the above result about the correspondence between functions of positive type and unitary representations; see Helgason [90], Chapter IV, Theorem 1.5. Theorem 53.8 is a little more general since it does not assume that X is a group, but when G is a group, the feature map arises from a unitary representation.

53.4 Kernel PCA

As an application of kernel functions, we discuss a generalization of the method of principal component analysis (PCA). Suppose we have a set of data $S = \{x_1, \ldots, x_n\}$ in some input space $\mathcal{X}$, and pretend that we have an embedding $\varphi\colon \mathcal{X} \to F$ of $\mathcal{X}$ in a (real) feature space $(F, \langle -, - \rangle)$, but that we only have access to the kernel function $\kappa(x,y) = \langle \varphi(x), \varphi(y) \rangle$. We would like to do PCA analysis on the set $\varphi(S) = \{\varphi(x_1), \ldots, \varphi(x_n)\}$.

There are two obstacles: