Remark: A *feature map* is often called a *feature embedding*, but this terminology is a bit misleading because it suggests that such a map is injective, which is not necessarily the case. Unfortunately this terminology is used by most people.

Example 53.1. Suppose we have two feature maps $\varphi_1 \colon X \to \mathbb{R}^{n_1}$ and $\varphi_2 \colon X \to \mathbb{R}^{n_2}$, and let $\kappa_1(x, y) = \langle \varphi_1(x), \varphi_1(y) \rangle$ and $\kappa_2(x, y) = \langle \varphi_2(x), \varphi_2(y) \rangle$ be the corresponding kernel functions (where $\langle -, - \rangle$ is the standard inner product on $\mathbb{R}^n$). Define the feature map $\varphi \colon X \to \mathbb{R}^{n_1+n_2}$ by

$$\varphi(x) = (\varphi_1(x), \varphi_2(x)),$$

an $(n_1 + n_2)$-tuple. We have

$$\begin{aligned}\langle \varphi(x), \varphi(y) \rangle &= \langle (\varphi_1(x), \varphi_2(x)), (\varphi_1(y), \varphi_2(y)) \rangle = \langle \varphi_1(x), \varphi_1(y) \rangle + \langle \varphi_2(x), \varphi_2(y) \rangle \\ &= \kappa_1(x, y) + \kappa_2(x, y),\end{aligned}$$

which shows that the map κ given by

$$\kappa(x, y) = \kappa_1(x, y) + \kappa_2(x, y)$$

is the kernel function corresponding to the feature map $\varphi \colon X \to \mathbb{R}^{n_1+n_2}$.

Example 53.2. Let X be a subset of $\mathbb{R}^2$, and let $\varphi_1 \colon X \to \mathbb{R}^3$ be the map given by

$$\varphi_1(x_1, x_2) = (x_1^2, x_2^2, \sqrt{2}x_1x_2).$$

Figure 53.1 illustrates $\varphi_1 \colon X \to \mathbb{R}^3$ when $X = \{((x_1, x_2) \mid -10 \leq x_1 \leq 10, -10 \leq x_2 \leq 10\}$.

Observe that linear relations in the feature space $H = \mathbb{R}^3$ correspond to quadratic relations in the input space (of data). We have

$$\begin{aligned}\langle \varphi_1(x), \varphi_1(y) \rangle &= \langle (x_1^2, x_2^2, \sqrt{2}x_1x_2), (y_1^2, y_2^2, \sqrt{2}y_1y_2) \rangle \\ &= x_1^2y_1^2 + x_2^2y_2^2 + 2x_1x_2y_1y_2 \\ &= (x_1y_1 + x_2y_2)^2 = \langle x, y \rangle^2,\end{aligned}$$

where $\langle x, y \rangle$ is the usual inner product on $\mathbb{R}^2$. Hence the function

$$\kappa(x, y) = \langle x, y \rangle^2$$

is a kernel function associated with the feature space $\mathbb{R}^3$.

If we now consider the map $\varphi_2 \colon X \to \mathbb{R}^4$ given by

$$\varphi_2(x_1, x_2) = (x_1^2, x_2^2, x_1x_2, x_1x_2),$$

we check immediately that

$$\langle \varphi_2(x), \varphi_2(y) \rangle = \kappa(x, y) = \langle x, y \rangle^2,$$

which shows that the same kernel can arise from different maps into different feature spaces.