Example 53.5. For any positive real constant $R > 0$, the constant function $\kappa(x, y) = R$ is a kernel function corresponding to the feature map $\varphi\colon X \to \mathbb{R}$ given by $\varphi(x, y) = \sqrt{R}$.

By definition, the function $\kappa_1'\colon \mathbb{R}^n \to \mathbb{R}$ given by $\kappa_1'(x, y) = \langle x, y\rangle$ is a kernel function (the feature map is the identity map from $\mathbb{R}^n$ to itself). We just saw that for any positive real constant $R > 0$, the constant $\kappa_2'(x, y) = R$ is a kernel function. By Example 53.1, the function $\kappa_3'(x, y) = \kappa_1'(x, y) + \kappa_2'(x, y)$ is a kernel function, and for any integer $d \geq 1$, by Example 53.4, the function κ_d given by

$$\kappa_d(x, y) = (\kappa_3'(x, y))^d = (\langle x, y\rangle + R)^d,$$

is a kernel function on $\mathbb{R}^n$. By the binomial formula,

$$\kappa_d(x, y) = \sum_{m=0}^{d} R^{d-m}\langle x, y\rangle^m.$$

By Example 53.1, the feature map of this kernel function is the concatenation of the features of the $d+1$ kernel maps $R^{d-m}\langle x, y\rangle^m$. By Example 53.3, the components of the feature map of the kernel map $R^{d-m}\langle x, y\rangle^m$ are reweightings of the functions

$$\varphi_{(i_1,\ldots,i_n)}(x) = x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}, \quad i_1 + i_2 + \cdots + i_n = m,$$

with $(i_1, \ldots, i_n) \in \mathbb{N}^n$. Thus the components of the feature map of the kernel function κ_d are reweightings of the functions

$$\varphi_{(i_1,\ldots,i_n)}(x) = x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}, \quad i_1 + i_2 + \cdots + i_n \leq d,$$

with $(i_1, \ldots, i_n) \in \mathbb{N}^n$. It is easy to see that the dimension of this feature space is $\binom{m+d}{d}$.

There are a number of variations of the polynomial kernel κ_d; all-subsets embedding kernels, ANOVA kernels; see Shawe–Taylor and Christianini [159], Chapter III.

In the next example the set X is not a vector space.

Example 53.6. Let D be a finite set and let $X = 2^D$ be its power set. If $|D| = n$, let $H = \mathbb{R}^X \cong \mathbb{R}^{2^n}$. We are assuming that the subsets of D are enumerated in some fashion so that each coordinate of $\mathbb{R}^{2^n}$ corresponds to one of these subsets. For example, if $D = \{1, 2, 3, 4\}$, let

$$\begin{array}{llll}
U_1 = \emptyset & U_2 = \{1\} & U_3 = \{2\} & U_4 = \{3\} \\
U_5 = \{4\} & U_6 = \{1, 2\} & U_7 = \{1, 3\} & U_8 = \{1, 4\} \\
U_9 = \{2, 3\} & U_{10} = \{2, 4\} & U_{11} = \{3, 4\} & U_{12} = \{1, 2, 3\} \\
U_{13} = \{1, 2, 4\} & U_{14} = \{1, 3, 4\} & U_{15} = \{2, 3, 4\} & U_{16} = \{1, 2, 3, 4\}.
\end{array}$$

Let $\varphi\colon X \to H$ be the feature map defined as follows: for any subsets $A, U \in X$,

$$\varphi(A)_U = \begin{cases} 1 & \text{if } U \subseteq A \\ 0 & \text{otherwise.} \end{cases}$$