SVC algorithms seek an "optimal" separating hyperplane H of equation $w^\top x - b = 0$. If some new data $x \in \mathbb{R}^n$ comes in, we can classify it by determining in which of the two half spaces determined by the hyperplane H they belong by computing the sign of the quantity $w^\top x - b$. The function $\mathrm{sgn}\colon \mathbb{R} \to \{-1, 1\}$ is given by

$$\mathrm{sgn}(x) = \begin{cases} +1 & \text{if } x \geq 0 \\ -1 & \text{if } x < 0. \end{cases}$$

Then we define the *(binary) classification function* associated with the hyperplane H of equation $w^\top x - b = 0$ as

$$f(x) = \mathrm{sgn}(w^\top x - b).$$

Remarkably, all the known optimization problems for finding this hyperplane share the property that the weight vector w and the constant b are given by expressions that *only involves inner products of the input data points u_i and v_j*, and so does the classification function

$$f(x) = \mathrm{sgn}(w^\top x - b).$$

This is a key fact that allows a far reaching generalization of the support vector machine using the method of *kernels*.

The method of kernels consists in assuming that the input space $\mathbb{R}^n$ is embedded in a larger (possibly infinite dimensional) Euclidean space F (with an inner product $\langle -, - \rangle$) usually called a *feature space*, using a function

$$\varphi\colon \mathbb{R}^n \to F$$

called a *feature map*. The function $\kappa\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ given by

$$\kappa(x, y) = \langle \varphi(x), \varphi(y) \rangle$$

is the kernel function associated with the embedding φ; see Chapter 53. The idea is that the feature map φ "unwinds" the input data, making it somehow more linear in the higher dimensional space F. Now even if we don't know what the feature space F is and what the embedding map φ is, we can pretend to solve our separation problem in F for the embedded data points $\varphi(u_i)$ and $\varphi(v_j)$. Thus we seek a hyperplane H of equation

$$\langle w, \zeta \rangle - b = 0, \quad \zeta \in F,$$

in the feature space F, to attempt to separate the points $\varphi(u_i)$ and the points $\varphi(v_j)$. As we said, it turns out that w and b are given by expression involving only the inner products $\kappa(u_i, u_j) = \langle \varphi(u_i), \varphi(u_j) \rangle, \kappa(u_i, v_j) = \langle \varphi(u_i), \varphi(v_j) \rangle$, and $\kappa(v_i, v_j) = \langle \varphi(v_i), \varphi(v_j) \rangle$, which form the symmetric $(p + q) \times (p + q)$ matrix $\mathbf{K}$ (a kernel matrix) given by

$$\mathbf{K}_{ij} = \begin{cases} \kappa(u_i, u_j) & 1 \leq i \leq p,\ 1 \leq j \leq q \\ -\kappa(u_i, v_{j-p}) & 1 \leq i \leq p,\ p + 1 \leq j \leq p + q \\ -\kappa(v_{i-p}, u_j) & p + 1 \leq i \leq p + q,\ 1 \leq j \leq p \\ \kappa(v_{i-p}, v_{j-q}) & p + 1 \leq i \leq p + q,\ p + 1 \leq j \leq p + q. \end{cases}$$