…explicitly mapping the data. This is achieved by using a kernel function $K(\mathbf{X}, \mathbf{X}')$, where:

$$K(\mathbf{X}, \mathbf{X}') = \langle \Phi(\mathbf{X}), \Phi(\mathbf{X}') \rangle$$

So, the weights are computed by the kernel matrix K as:

$$\mathbf{w} = \mathbf{K}\mathbf{a}$$

where $\mathbf{K} = \Phi(\mathbf{X})\Phi(\mathbf{X})^T$ and $\mathbf{a}$ is the vector of dual coefficients.

Kernel Trick in Polynomial Regression

In polynomial regression, the feature map is a polynomial function of the input features:

$$\Phi(\mathbf{X}) = \left[1, x_1, x_2, \ldots, x_n, x_1^2, x_1 x_2, x_2^2, \ldots\right]$$

This maps the input space to a higher-dimensional space, allowing for more complex, nonlinear relationships. The polynomial kernel is given by:

$$K(\mathbf{X}, \mathbf{X}') = (\langle \mathbf{X}, \mathbf{X}' \rangle + c)^d$$

where c is a constant and d is the degree of the polynomial.