Solution 1: SVM – Kernel Trick

The polynomial kernel is defined as

$$k(x, \tilde{x}) = (x^T \tilde{x} + b)^d.$$

Furthermore, assume $x \in \mathbb{R}^2$ and $d = 2$.

(a) Derive the explicit feature map ϕ taking into account that the following equation holds:

$$k(x, \tilde{x}) = \langle \phi(x), \phi(\tilde{x}) \rangle$$

Solution:

$$\begin{aligned}
k(x, \tilde{x}) &= \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{pmatrix} + b \right)^2 \\
&= (x_1\tilde{x}_1 + x_2\tilde{x}_2 + b)^2 \\
&= (x_1\tilde{x}_1 + x_2\tilde{x}_2)^2 + 2(x_1\tilde{x}_1 + x_2\tilde{x}_2)b + b^2 \\
&= x_1^2\tilde{x}_1^2 + 2x_1\tilde{x}_1x_2\tilde{x}_2 + x_2^2\tilde{x}_2^2 + 2bx_1\tilde{x}_1 + 2bx_2\tilde{x}_2 + b^2 \\
&= \left\langle \begin{pmatrix} x_1^2 \\ \sqrt{2}x_1x_2 \\ x_2^2 \\ \sqrt{2b}x_1 \\ \sqrt{2b}x_2 \\ b \end{pmatrix}, \begin{pmatrix} \tilde{x}_1^2 \\ \sqrt{2}\tilde{x}_1\tilde{x}_2 \\ \tilde{x}_2^2 \\ \sqrt{2b}\tilde{x}_1 \\ \sqrt{2b}\tilde{x}_2 \\ b \end{pmatrix} \right\rangle \\
&= \langle \phi(x), \phi(\tilde{x}) \rangle
\end{aligned}$$

(b) Describe the main differences between the kernel method and the explicit feature map.

Solution:

Using the kernel method reduces the computational costs of computing the scalar product in the higher-dimensional features space after calculating the feature map.