Kernel Trick in Logistic Regression

In logistic regression, the output is modeled as a logistic function:

$$y = \sigma(\mathbf{X}\mathbf{w})$$

where $\sigma(\cdot)$ is the sigmoid function. Using kernelization, we map the data to a higher-dimensional space using the feature map $\Phi(\mathbf{X})$:

$$y = \sigma(\Phi(\mathbf{X})\mathbf{w})$$

The kernelized version of the logistic regression then becomes:

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

Again, the kernel matrix K represents the dot products in the transformed space, and we compute the dual coefficients to obtain the model.