Kernel Functions and Kernel K-Means Clustering

❑ Typical kernel functions:

    ❑ Polynomial kernel of degree h: $K(\boldsymbol{X}_i, \boldsymbol{X}_j) = (\boldsymbol{X}_i \cdot \boldsymbol{X}_j + 1)^h$

    ❑ Gaussian radial basis function (RBF) kernel: $K(\boldsymbol{X}_i, \boldsymbol{X}_j) = e^{-\|X_i - X_j\|^2 / 2\sigma^2}$

    ❑ Sigmoid kernel: $K(\boldsymbol{X}_i, \boldsymbol{X}_j) = \tanh(\kappa \boldsymbol{X}_i \cdot \boldsymbol{X}_j - \delta)$

❑ The formula for kernel matrix K for any two points x_i, $x_j \in C_k$ is $K_{x_i x_j} = \phi(x_i) \bullet \phi(x_j)$

❑ The SSE criterion of *kernel K-means*: $\quad SSE(C) = \sum_{k=1}^{K} \sum_{x_i \in C_k} \| \phi(x_i) - c_k \|^2$

    ❑ The formula for the cluster centroid:

$$c_k = \frac{\sum_{x_i \in C_k} \phi(x_i)}{|C_k|}$$

❑ Clustering can be performed without the actual individual projections $\phi(x_i)$ and $\phi(x_j)$ for the data points x_i, $x_j \in C_k$