SVM KERNELS (CONT.)

SYRACUSE UNIVERSITY
School of Information Studies

KERNEL FUNCTIONS

SVM algorithm maximizes the margin between the two separating hyperplanes by finding the maximum of the function:

$$W(\quad) = \mathop{}_{i=1}^{l} \quad {}_i \quad \frac{1}{2} \mathop{}_{i=1}^{l} \mathop{}_{j=1}^{l} \quad {}_i \quad {}_j y_i y_j \boxed{K(x_i, x_j)}$$

Subject to the constraints:

$$\mathop{}_{i=1}^{l} \quad {}_i y_i = 0, \quad {}_i \quad 0, i = 1,2,\ldots,l$$

SVM: KERNEL FUNCTIONS

Linear kernel: $K(X_i, X_j) = X_i \cdot X_j$ (cosine similarity)

Higher rank kernels: Instead of computing on the transformed data tuples, it is mathematically equivalent to instead applying a kernel function $K(X_i, X_j)$ to the original data, i.e., $K(X_i, X_j) = \Phi(X_i)\ \Phi(X_j)$

Typical kernel functions:

$$\text{Polynomial kernel of degree } h: \quad K(X_i, X_j) = (X_i \cdot X_j + 1)^h$$

$$\text{Gaussian radial basis function kernel}: \quad K(X_i, X_j) = e^{-\|X_i - X_j\|^2/2\sigma^2}$$

$$\text{Sigmoid kernel}: \quad K(X_i, X_j) = \tanh(\kappa X_i \cdot X_j - \delta)$$