

# KERNEL TRICK

Support vector classifiers can be written as a dot product:

$$b + \sum_{i=1} \alpha_i \underbrace{x^T x^{(i)}}_{\text{dot product}}$$

Annotations:   
 -  $b$  is labeled "bias" with a blue arrow.   
 -  $\alpha_i$  is labeled "parameters" with a red arrow.   
 -  $x^{(i)}$  is labeled "observation" with a green arrow.   
 - The entire term  $\alpha_i x^T x^{(i)}$  is indicated by a green arrow labeled "dot product".

The kernel trick is to replace the dot product with a Kernel:

$$b + \sum \alpha_i k(x, x^{(i)})$$

Annotation:   
 -  $k(x, x^{(i)})$  is labeled "Kernel" with a red arrow.

Allows for non-linear decision boundaries and computational efficiency.