1 Kernel methods

Kernel methods help in using linear decision boundaries on non-linearly separable data by morphing the feature space. They can also help in disentangling data for clustering. Kernels can incorporate domain knowledge into your model.

1.1 Introduction to Kernels

Kernels can be seen in two ways: As similarity measures between data-points or transformations of the data-points into a higher dimensional space. For two points x and y a Kernel is given by:

$$K(x, y) = \sum_{i=1}^{n} h_i(x) h_i(y) = \langle h(x), h(y) \rangle, \tag{1}$$

where $h(x)$ is a transformation-function and $\langle \cdot \rangle$ is the inner product.

1.2 Kernel SVM