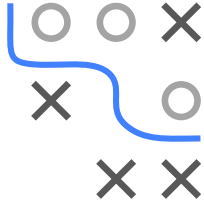
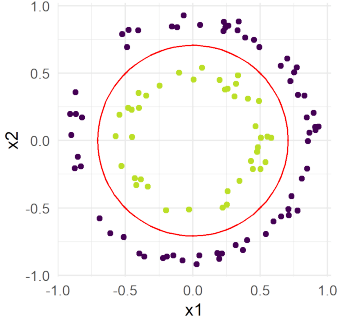
Introduction to Machine Learning



Nonlinear Support Vector Machines Feature Generation for Nonlinear Separation

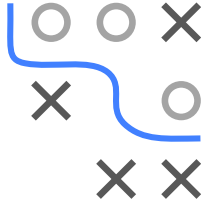


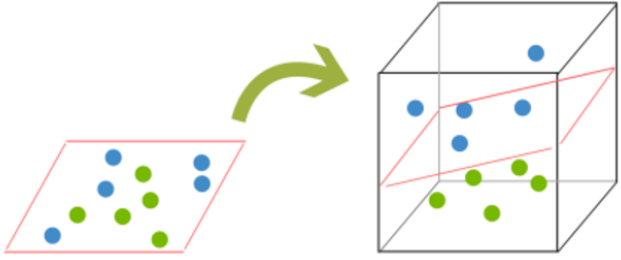
Learning goals

- Understand how nonlinearity can be introduced via feature maps in SVMs
- Know the limitation of feature maps

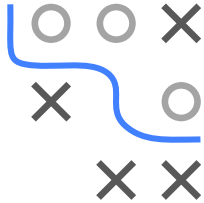
NONLINEARITY VIA FEATURE MAPS

- How to extend a linear classifier, e.g. the SVM, to nonlinear separation between classes?
- We could project the data from 2D into a richer 3D feature space!





FEATURE MAPS: COMPUTATIONAL LIMITATIONS



Let us have a look at a similar nonlinear feature map $\phi : \mathbb{R}^2 \to \mathbb{R}^5$, where we collect all monomial feature extractors up to degree 2 (pairwise interactions and quadratic effects):

$$\phi(x_1, x_2) = (x_1^2, x_2^2, x_1 x_2, x_1, x_2).$$

For p features vectors, there are k_1 different monomials where the degree is exactly d, and k_2 different monomials up to degree d.

$$k_1 = \binom{d+p-1}{d} \qquad k_2 = \binom{d+p}{d} - 1$$

Which is quite a lot, if p is large.