Lectures 22, 23

This week we discuss prototype methods for classification — these are like $k$ nearest neighbors with the change that only selected 'prototype' points for regions of the feature space are kept, rather than the entire training set.

We will study unsupervised methods such as clustering — $k$-means clustering versus hierarchical clustering methods, which give very visualizable 'dendrograms' - hierarchical trees which subdivide families of feature vectors in a recursive way. Prior to this an application of clustering is also discussed - there are ways of using clustering not just to divide up the feature space into categories based on similarity, but to do (supervised) classification on this basis.

We discuss principal component analysis, which essentially finds the best $q$-dimensional subspace to project a dataset in $\mathbb{R}^p$ onto (by minimizing the distances between original and projected points). Most interesting is that this can also be adapted to become a projection onto a nonlinear surface.

Problems: 12.2, 13.3, 14.7