1 Introduction

Neighbourhood component analysis

Neighbourhood component analysis (NCA) builds on the well-known classification method k nearest neighbours (kNN). Instead of using the plain Euclidean metric for computing the distances, NCA proposes to learn a more general metric from the training data. In this case, the metric used is a Mahalanobis-like distance. For the classic Mahalanobis metric, we have $\mathbf{S} = (\mathrm{cov}\{\mathcal{X}\})^{-1}$. We can generalize and use any $\mathbf{S}$, as long at it is positive definite. This allows to learn different metrics that are specific for our tasks: on the same data set multiple tasks can be performed, so $\mathbf{S}$ its dependent on the task.

A nice remark is that learning a Mahalanobis distance metric is equivalent to learning a projection matrix:

$$\begin{aligned} d &= \sqrt{(\mathbf{x}_i - \mathbf{x}_j)^{\mathrm{T}}\mathbf{S}(\mathbf{x}_i - \mathbf{x}_j)} \\ &= \sqrt{(\mathbf{x}_i - \mathbf{x}_j)^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{A}(\mathbf{x}_i - \mathbf{x}_j)} \\ &= \sqrt{(\mathbf{A}\mathbf{x}_i - \mathbf{A}\mathbf{x}_j)^{\mathrm{T}}(\mathbf{A}\mathbf{x}_i - \mathbf{A}\mathbf{x}_j)}. \end{aligned}$$

The matrix $\mathbf{S}$ or $\mathbf{A}$ are achieved by maximizing an objective function that is equivalent to the number of correctly classified points.

Summary:

- Learns a Mahalanobis distance metric S:

$$d = \sqrt{(\mathbf{x}_i - \mathbf{x}_j)^{\mathrm{T}}\mathbf{S}(\mathbf{x}_i - \mathbf{x}_j)}.$$

- Objective function: expected number of correctly classified points

$$f(\mathbf{S}) = \sum_i p_i,$$