

实验 1. 度量学习实验报告

MG1733098, 周华平, zhp@smail.nju.edu.cn

2017 年 11 月 5 日

综述

任务 1

度量函数学习目标

根据马氏距离的定义

$$\text{dist}_{mah}^2(x, y) = (x - y)^\top Q (x - y) = (Ax - Ay)^\top (Ax - Ay)$$

其中 Q 称为“度量矩阵”，而度量学习则是对 Q 的学习。为了保持距离非负且对称， Q 必须是（半）正定对称矩阵，即必有正交基 A 使得 Q 能写为 $Q = AA^\top$ 。

为了提高近邻分类器的性能，我们将 Q 直接嵌入到近邻分类器的评价指标中去，通过优化该性能目标相应地求得 Q 。在本实验中我们采用近邻成分分析 (Neighbourhood Component Analysis) 进行学习。

近邻分类器在进行判别时通常使用多数投票法，领域中的每个样本投 1 票，领域外的样本投 0 票。NCA 将其替换为概率投票法，对于任意样本 x_j ，它对 x_i 分类结果影响的概率为

$$p_{ij} = \frac{\exp(\|Ax_i - Ax_j\|^2)}{\sum_{k \neq i} \exp(\|Ax_i - Ax_k\|^2)}, \quad p_{ii} = 0$$

若以留一法正确率的最大化为目标，则可计算 x_i 的留一法正确率，即它被自身之外的所有样本正确分类的概率为

$$p_i = \sum_{j \in C_i} p_{ij}$$

其中 $C_i = \{j | c_i = c_j\}$ ，即与 x_i 属于相同类别的样本的下标集合。于是，整个样本集上的留一法正确率为

$$f(A) = \sum_i \sum_{j \in C_i} p_{ij} = \sum_i p_i$$

NCA 的优化目标是使得 $f(A)$ 最大化，即

$$\max_A \sum_i \sum_{j \in C_i} p_{ij}$$

通过求 f 对 A 的偏导，可以得到梯度公式（令 $x_{ij} = x_i - x_j$ ）

$$\frac{\partial f}{\partial A} = -2A \sum_i \sum_{j \in C_i} p_{ij} (x_{ij} x_{ij}^\top - \sum_k p_{ik} x_{ik} x_{ik}^\top)$$

根据该公式，使用梯度下降法即可求解 NCA 的目标函数，得到最大化近邻分类器留一法正确率的距离度量矩阵 Q 。

优化算法

任务 2