

Метод NCA для обучения метрики

Сомов Иван Сергеевич

МГУ имени М. В. Ломоносова, факультет ВМК, кафедра ММП

21 ноября 2017 г.

Определение параметров KNN:

- выбор значения k
- выбор метрики

NCA:

- позволяет обучить обобщённую метрику
- можно применить для снижения размерности пространства

- $d(\vec{x}, \vec{y}) = \sqrt{(\vec{x} - \vec{y})^T S^{-1} (\vec{x} - \vec{y})}$
- S – ковариационная матрица

- $d(\vec{x}, \vec{y}) = \sqrt{(\vec{x} - \vec{y})^T S^{-1} (\vec{x} - \vec{y})}$
- S – ковариационная матрица
- $S^{-1} = A^T A$

- $d(\vec{x}, \vec{y}) = \sqrt{(\vec{x} - \vec{y})^T S^{-1} (\vec{x} - \vec{y})}$
- S – ковариационная матрица
- $S^{-1} = A^T A$
- $d(\vec{x}, \vec{y}) = \sqrt{(A\vec{x} - A\vec{y})^T (A\vec{x} - A\vec{y})}$

Минусы LOO-error:

- неустойчива для малых изменений A
- недифференцируема

- p_{ij} – вероятность, что объект i окажется соседом j и примет его класс

- p_{ij} – вероятность, что объект i окажется соседом j и примет его класс
- $$p_{ij} = \frac{\exp(-\|Ax_i - Ax_j\|^2)}{\sum_{k \neq i} \exp(-\|Ax_i - Ax_k\|^2)}, p_{ii} = 0$$

- p_{ij} – вероятность, что объект i окажется соседом j и примет его класс
- $$p_{ij} = \frac{\exp(-\|Ax_i - Ax_j\|^2)}{\sum_{k \neq i} \exp(-\|Ax_i - Ax_k\|^2)}, p_{ii} = 0$$
- $p_i = \sum_{j \in C_i} p_{ij}$ – вероятность правильной классификации объекта i , где $C_i = \{j \mid c_i = c_j\}$

Функция потерь и оптимизационная задача

- $f(A) = \sum_i p_i$ – полученный функционал
- $f(A) \rightarrow \max_A$

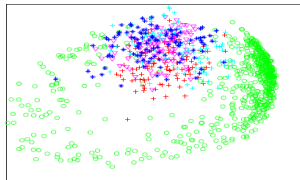
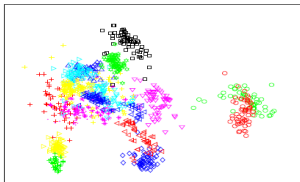
Функция потерь и оптимизационная задача

- $f(A) = \sum_i p_i$ – полученный функционал
- $f(A) \rightarrow \max_A$
- $\frac{\partial f}{\partial A} = -2A \sum_i \sum_{j \in C_i} p_{ij} \left(x_{ij} x_{ij}^T - \sum_k p_{ik} x_{ik} x_{ik}^T \right)$

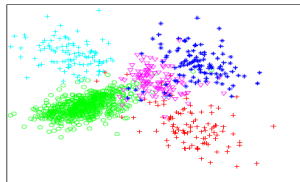
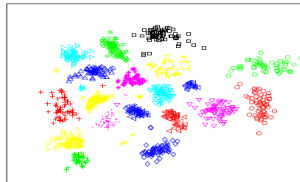
Понижение размерности пространства

- Пусть $A : \mathbb{R}^D \rightarrow \mathbb{R}^d$
- $A \in \mathbb{R}^{d \times D}$, $d \leq D$

Понижение размерности пространства

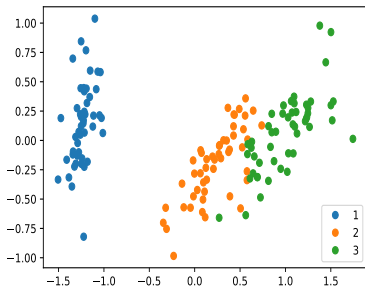


PCA

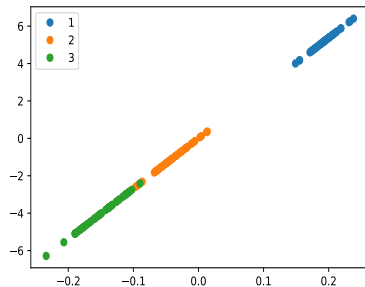


NCA

Снижение размерности для датасета iris



PCA



NCA

Плюсы:

- лучше, чем РСА
- не требует информации о структуре пространства
- легко реализовать

Минусы:

- долго обучать
- невыпуклый функционал
- мало эффективных реализаций метода