Prof. Dr. Johannes Blömer

21.01.2016
submission due: 28.01.2016, F1.110, 11:00

Clustering Algorithms

WS 2015/2016

Handout 10

Exercise 1:
Let $P \subset \mathbb{R}^d$ and let V be an arbitrary subspace of $\mathbb{R}^d$. Assume $\hat{\mathcal{C}} = \{\hat{C}_1, \ldots, \hat{C}_k\}$ is a k-clustering of $\pi_V(P)$ and denote by $C_i := \{p \in P : \pi_V(p) \in \hat{C}_i\}, \mathcal{C} := \{C_1, \ldots, C_k\}$ the corresponding k-clustering of P. Prove that

$$D(\pi(P), \hat{\mathcal{C}}) \leq \mathrm{cost}(P, \mathcal{C}).$$

Exercise 2:
Consider the point set $P = \{(1,0), (0,1), (2,1)\}$ in $\mathbb{R}^2$. Use the singular value decomposition to compute an optimal solution to the k-variance problem with input P and $k = 1$.

Exercise 3:
Prove Theorem 5.16 for the case $k = 2$.