Prof. Dr. Johannes Blömer

6.11.2015
submission due: 19.11.2015, F1.110, 13:15

Clustering Algorithms

WS 2015/2016

Handout 3

Exercise 1:
The Itakura-Saito divergence is defined by

$$D_{IS}(p,q) = \sum_{i=1}^{d} \left(\frac{p_i}{q_i} - \ln \frac{p_i}{q_i} - 1 \right) \text{ , where } p, q \in \mathbb{R}^d \ .$$

Show that D_{IS} is a Bregman divergence associated to the (differentiable and strictly convex function) $\phi_{IS}(t) = \sum_{i=1}^{d} \ln \frac{1}{t_i}$.

Exercise 2:
Prove the following statements.

(a) Let D_ϕ be a Bregman divergence on domain $S \subseteq \mathbb{R}^d$. Show that for all $p, q \in S$ and $0 \leq \lambda \leq 1$ we have

$$D_\phi(\lambda p + (1-\lambda)q, r) \leq \lambda D_\phi(p,r) + (1-\lambda) D_\phi(q,r).$$

(b) Let $\phi, \psi : S \to \mathbb{R}$ be differentiable, strictly convex functions and let $\alpha, \beta > 0$ be arbitrary. Then we have

$$D_{\alpha\phi+\beta\psi} = D_{\alpha\phi} + D_{\beta\psi}$$

Hint: $\nabla(\alpha\phi + \beta\psi)(t) = \alpha\nabla\phi(t) + \beta\nabla\psi(t)$

($\star$) **Exercise 3:**
A k-clustering $\mathcal{C} = \{C_1, \dots, C_k\}$ is called linear separable if every two distinct clusters are separated by a hyperplane H, i.e., for each i, j, $i \neq j$, there exist some $a, b \in \mathbb{R}^d$ such that

$$C_i \subseteq H^+ = \{x \in \mathbb{R}^d \mid a^T x \leq b\}$$

and

$$C_j \subseteq H^- = \{x \in \mathbb{R}^d \mid a^T x > b\}.$$

Let $\mathcal{C} = \{C_1, \dots, C_k\}$ be a k-clustering with respect of a Bregman distance D_ϕ. Prove that $\mathcal{C}$ is linearly separable by defining the hyperplanes that separate the clusters.