GRAPHICAL MARKOV MODELS (WS2014)
4. SEMINAR

Assignment 1. Derive an EM-algorithm for learning the following conditional independent probability model (sometimes called "naive Bayes probabilistic model")

$$p(x, y, k) = p(x \mid k)p(y \mid k)p(k),$$

where $x \in X$, $y \in Y$, $k \in K$ and all three sets are finite. The training data for learning are i.i.d. realisations of pairs (x, y) with empirical distribution $p^*(x, y)$ (the corresponding values of k are not observed).

Assignment 2. Prove equivalence of Definitions 1a and 1b given in the lecture for a Markov model on a tree (see sec. 11).

Assignment 3. Let $T = (V, E)$ be an undirected tree and s_i, $i \in V$ be a field of K-valued random variables. Suppose that $v_{ij}(k, k')$, $k, k' \in K$ is a system of pairwise probabilities associated with the edges $\{i, j\} \in E$ of the tree. Consider the set $\mathcal{P}(\mathbf{v})$ of all joint probability distributions $p(s)$, which have $\mathbf{v}$ as pairwise marginals, i.e.

$$\sum_{s \in K^{|V|}} p(s)\delta_{s_i k}\delta_{s_j k'} = v_{ij}(k, k') \quad \forall \{i, j\} \in E,\ \forall k, k' \in K.$$

We want to find the distribution wit highest entropy

$$H(p) = -\sum_{s \in K^{|V|}} p(s) \log p(s)$$

in $\mathcal{P}(\mathbf{v})$. Prove that the unique maximiser is the Markov model on the tree T defined by the edge marginals $\mathbf{v}$.
Hint: Formulate and solve the constrained optimisation task by using its Lagrange function.

Assignment 4. Let $\mathbf{a}^m$, $m = 1, 2, \ldots, M$ be vectors in $\mathbb{R}^n$. Consider the following pair of functions defined on $\mathbb{R}^n$

$$g(\mathbf{x}) = \log \sum_{m=1}^{M} \exp \langle \mathbf{a}^m, \mathbf{x} \rangle$$

$$h(\mathbf{y}) = \inf_p \Big\{ \sum_{m=1}^{M} p_m \log p_m \;\Big|\; \sum_{m=1}^{M} p_m = 1,\ \sum_{m=1}^{M} p_m \mathbf{a}^m = \mathbf{y} \Big\}.$$

Show that g is the Fenchel conjugate of h.
Hint: Substitute h into the definition of its Fenchel conjugate. Carry the infimum (possibly changing it to a supremum because of a sign) to the left and swap it with the

supremum in the definition of the Fenchel conjugate. Solve the resulting optimisation task and substitute its solution.
Remark: We assume here that the value of $x \log x$ is $+\infty$ if x is negative.