GRAPHICAL MARKOV MODELS (WS2014)
4. SEMINAR

Assignment 1. 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. Move the infimum (changing it to a supremum because of the minus 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.

Assignment 2. Consider the task of finding the most probable sequence of (hidden) states for a (Hidden) Markov model on a chain.
a) Show that the Dynamic Programming approach applied for this task can be interpreted as an equivalent transformation (re-parametrisation) of the model.
b) Show that the transformed functions (potentials) encode an explicit description of *all* optimisers of the problem

Assignment 3. Consider a GRF for binary valued labellings $x \colon V \to \{0, 1\}$ of a graph (V, E) given by

$$p(x) = \frac{1}{Z} \exp \Big[\sum_{ij \in E} u_{ij}(x_i, x_j) + \sum_{i \in V} u_i(x_i) \Big].$$

Show that is is always possible to find an equivalent transformation (re-parametrisation)

$$u_{ij} \to \tilde{u}_{ij}, \quad u_i \to \tilde{u}_i$$

such that the new pairwise functions $\tilde{u}_{ij}$ have the form

$$\tilde{u}_{ij}(x_i, x_j) = \alpha_{ij} |x_i - x_j|$$

with some reals $\alpha_{ij} \in \mathbb{R}$.

Assignment 4. Let K be a completely ordered finite set. We assume w.l.o.g. that $K = \{1, 2, \ldots, m\}$. For a function $u \colon K \to \mathbb{R}$ define its discrete "derivative" by $Du(k) = u(k+1) - u(k)$.

a) Let u be a function $u\colon K^2 \to \mathbb{R}$ and denote by D_1 and D_2 the discrete derivatives w.r.t. its first and second argument. Prove the following equality

$$D_1 D_2 u(k_1, k_2) = u(k_1 + 1, k_2 + 1) + u(k_1, k_2) - u(k_1 + 1, k_2) - u(k_1, k_2 + 1).$$

Conclude that all mixed derivatives $D_1 D_2 u(k_1, k_2)$ of a submodular functions are negative.

b) Prove that the condition established in a) is not only necessary but also sufficient for a function to be submodular.

Hint: Start from the observation that the following equality holds for a function of one variable

$$u(k + l) - u(k) = \sum_{i=k}^{k+l-1} Du(i)$$

and generalise it for functions of two variables.

c) Prove that any function $u\colon K^2 \to \mathbb{R}$ can be represented as a sum of a submodular and a supermodular function.

Hint: Consider the mixed derivative $D_1 D_2 u(k_1, k_2)$, decompose it into its negative and positive part and "integrate" them back separately.