STRUCTURED MODEL LEARNING (SS2022)
1. SEMINAR

Assignment 1. Consider a collection $\{S_i \mid i = 1, 2, 3\}$ of three binary valued random variables, i.e., $s_i \in \{0, 1\}$ for $i = 1, 2, 3$. Suppose we fix all three pairwise marginal distributions $p_{ij}(s_i, s_j) = \mu(s_i, s_j)$, where

$$\mu(s_i, s_j) = \begin{cases} \frac{1}{2}\alpha & \text{if } s_i = s_j, \\ \frac{1}{2}(1-\alpha) & \text{otherwise} \end{cases}$$

and α is a fixed real number from the interval $[0, 1]$. We seek a simple joint distribution $p(s_1, s_2, s_3)$ that has the given marginals. Someone proposes the properly normalised product of μ-s

$$\bar{p}(s_1, s_2, s_3) = \frac{1}{Z(\alpha)} \mu(s_1, s_2)\mu(s_2, s_3)\mu(s_1, s_3).$$

Prove that the pairwise marginal distributions of $\bar{p}$ do not(!) coincide with the function μ.

Assignment 2. Consider a collection $\{S_i \mid i = 1, 2, 3\}$ of three binary valued random variables as in the previous assignment. Let us fix the following pairwise marginal distributions

$$p(s_1, s_2) = \mu(s_1, s_2)\ , \quad p(s_1, s_3) = \mu(s_1, s_3)\ , \quad p(s_2, s_3) = \tilde{\mu}(s_2, s_3)$$

where

$$\mu(s_i, s_j) = \begin{cases} 0.5 & \text{if } s_i = s_j, \\ 0 & \text{otherwise,} \end{cases} \qquad \tilde{\mu}(s_i, s_j) = \begin{cases} 0.5 & \text{if } s_i \neq s_j, \\ 0 & \text{otherwise,} \end{cases}$$

Do the μ-s represent a valid system of pairwise marginal distributions? I.e., is there a joint distribution $p(s_1, s_2, s_3)$ whose pairwise marginals coincide with the μ-s?

Assignment 3. Let $S = \{S_i \mid i \in V\}$ be a K-valued random field and let $\mathcal{P}$ denote the set of all possible joint probability distributions $p\colon K^{|V|} = \mathcal{S} \to \mathbb{R}_+$, s.t. $\sum_{s\in\mathcal{S}} p(s) = 1$.

a) Prove that the distribution $p \in \mathcal{P}$ with highest entropy is the uniform distribution. Prove that it factorises into the product of its unary marginal distributions.

b) Let us fix unary marginal distributions for each S_i, $i \in V$ by $p(s_i) = \mu_i(s_i)$. We assume that the functions $\mu_i\colon K \to \mathbb{R}_{++}$ fulfil $\sum_{k\in K} \mu_i(k) = 1$ for all $i \in V$.

Prove that the distribution

$$p(s) = \prod_{i\in V} \mu_i(s_i)$$

has the highest entropy among all joint distributions $p \in \mathcal{P}$ which have the given unary marginals. What happens if the functions μ_i are not necessarily strictly positive?

c) We equip V with the structure of an undirected graph (V, E). Let us fix pairwise marginal distributions for each pair of variables S_i, S_j where $\{i, j\} \in E$ by setting $p(s_i, s_j) = \mu_{ij}(s_i, s_j)$. All functions $\mu_{ij}\colon K^2 \to \mathbb{R}_{++}$ fulfil

$$\sum_{k,k'\in K} \mu_{ij}(k, k') = 1.$$

Furthermore, we assume that the system of μ-s represents a valid system of pairwise marginals, i.e. there exists at least one strictly positive joint distribution $\bar{p} \in \mathcal{P}$ whose pairwise marginal distributions coincide with μ-s.

Fill in details for the derivation in Section 1. of the lecture and prove that the distribution $p \in \mathcal{P}$ with highest entropy (among all those that have the given fixed pairwise marginals) has the form

$$p_u(s) = \frac{1}{Z(u)} \exp\Big[\sum_{ij \in E} u_{ij}(s_i, s_j)\Big],$$

where u-s are Lagrange multipliers which have to determined such that p has the required pairwise marginals.

Assignment 4.* (**Maximum entropy**) Define a convex function $h\colon \mathbb{R} \to (-\infty, +\infty]$ by

$$h(u) = \begin{cases} u \log u - u & \text{if } u > 0 \\ 0 & \text{if } u = 0 \\ +\infty & \text{if } u < 0 \end{cases}$$

and a convex function $f\colon \mathbb{R}^n \to (-\infty, +\infty]$ by

$$f(x) = \sum_{i=1}^{n} h(x_i).$$

(a) Prove f is strictly convex on $\mathbb{R}^n_+$ with compact level sets.
(b) Suppose the map $G\colon \mathbb{R}^n \to \mathbb{R}^m$ is linear with $G\hat{x} = b$ for some point $\hat{x}$ in the interior of $\mathbb{R}^n_+$. Prove for any vector c in $\mathbb{R}^n$ that the problem

$$\inf\big\{ f(x) + \langle c, x\rangle \;\big|\; Gx = b,\, x \in \mathbb{R}^n \big\}$$

has a unique solution $\bar{x}$ lying in $\mathbb{R}^n_{++}$.
(c) Prove that some vector λ in $\mathbb{R}^m$ satisfies $\nabla f(\bar{x}) = G^T\lambda - c$, and deduce $\bar{x}_i = \exp(G^*\lambda - c)_i$.