## 3. The most probable realisation of a GRF

$S = \{S_i \mid i \in V\}$ is a $K$-valued Gibbs random field w.r.t. the undirected graph $(V, E)$.

Task: Find the most probable realisation(s) $s_* \in \mathcal{S} = K^{|V|}$

$$s_* \in \operatorname*{argmax}_{s \in \mathcal{S}} \frac{1}{Z(u)} \exp \sum_{ij \in E} u_{ij}(s_i, s_j) = \operatorname*{argmax}_{s \in \mathcal{S}} \sum_{ij \in E} u_{ij}(s_i, s_j)$$

Remarks

- The task is NP complete (max-clique)
- The task is solvable in polynomial time if $(V, E)$ is a tree
- The task is solvable in polynomial time if $K$ is completely ordered and all functions $-u_{ij} : K^2 \to \mathbb{R}$ are submodular w.r.t. the ordering ∎

## LA: LP-relaxation

From an abstract view we need to solve the task

$$\operatorname*{argmax}_{s \in \mathcal{S}} \langle u, \varphi(s) \rangle$$

It can be relaxed to

$$\operatorname*{argmax}_{p \in \mathcal{P}} \sum_{s \in \mathcal{S}} p(s) \langle u, \varphi(s) \rangle = \operatorname*{argmax}_{p \in \mathcal{P}} \langle u, \mathbb{E}_p(\varphi) \rangle$$

For a GRF: $\mathbb{E}_p(\varphi) \mathrel{\hat{=}}$ pairwise marginal distributions.
Hence, the task reads

$$\begin{cases} \langle u, \mu \rangle \to \max\limits_{\mu} \\ \text{s.t. } \mu \in \operatorname{conv} \varphi(\mathcal{S}) \end{cases}$$

This is an LP-task, but $\operatorname{conv} \Phi(\mathcal{S})$ is „complicated".

↳ Relax the constraints to a simpler polytope $L \supset \operatorname{conv} \Phi(\mathcal{S})$

For a GRF:

$$\sum_{ij\in E} \sum_{s_i, s_j \in K} \mu_{ij}(s_i, s_j)\, u_{ij}(s_i, s_j) \to \max_{\mu}$$

s.t. $\displaystyle\sum_{s_i, s_j \in K} \mu_{ij}(s_i, s_j) = 1 \quad \forall \{ij\}\in E$

$$\sum_{s_j} \mu_{ij}(s_i, s_j) = \sum_{s_e} \mu_{ie}(s_i, s_e) \quad \forall i,j,e : \{ij\}, \{i,e\} \in E \quad \forall s_i \in K$$

$$\mu > 0$$

From abstract view, the relaxed task is

$$\langle u, \mu \rangle \to \max_{\mu}$$

s.t. $\mu \in \mathbb{R}^n_+ \cap \operatorname{aff} \Phi(\mathcal{S})$

Suppose $\operatorname{aff} \Phi(\mathcal{S}) = \{\mu \in \mathbb{R}^n \mid A\mu = b\}$, $A : \mathbb{R}^n \to \mathbb{R}^m$

The relaxed task reads

$$\langle u, \mu \rangle \to \max_{\mu}$$

s.t. $A\mu = b$

$\mu \geq 0$

Its dual task is

$$\langle b, \psi \rangle \to \min_{\psi}$$

s.t. $A^T \psi \geq u$

Remember: $\psi$ and $A^T\psi$ describe reparametrisations!

„Translate" this for the GRF → see seminar.

## 4. Computing marginal distributions for a GRF

$S = \{S_i \mid i \in V\}$ is a $K$-valued Gibbs random field w.r.t. the undirected graph $(V, E)$

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

Task: Compute marginal distributions $p(s_i)$, $i \in V$, $s_i \in K$ for vertices and $p(s_i, s_j)$, $\{ij\} \in E$, $s_i, s_j \in K$ for edges.

We know that for exponential families

$$p_u(s) = \frac{1}{Z(u)} \exp\langle \varphi(s), u \rangle \rightarrow \mathbb{E}_u(\varphi) = \nabla \log Z(u)$$

but this does not help.

Remark 1. The task is easy to solve if $(V, E)$ is a tree.

$\Rightarrow$ Belief propagation $\Rightarrow$ approx. algorithm for GRFs on arbitrary graphs: Loopy belief propagation (see seminar)

## L4. Mean field approximation for unary marginals

Idea: approximate $p_u(s)$ by a simpler distribution, e.g. assuming that $S_i$, $i \in V$ are independent.

$$D_{KL}(q \,\|\, p_u) = \sum_{s \in \mathcal{S}}' q(s) \ln \frac{q(s)}{p_u(s)} \rightarrow \min_q$$

where $q(s) = \prod_{i \in V} q_i(s_i)$

We get

$$\sum_{i\in V}\sum_{s_i\in K} q_i(s_i)\log q_i(s_i) - \sum_{ij\in E}\sum_{s_i,s_j\in K} u_{ij}(s_i,s_j)\,q_i(s_i)\,q_j(s_j) \to \min_q$$

s.t. $\sum_{s_i\in K} q_i(s_i) = 1 \quad \forall i\in V$

The task is convex for each single $q_i$ provided all other $q_j$, $j\neq i$ are fixed, but not convex.

Algorithm Start with some $q^{(0)}$. Keep iterating over all $i\in V$, each time solve for $q_i$ holding all other $q_j$, $j\neq i$ fixed. A single step reads

$$q_i(s_i) \leftarrow \frac{1}{Z_i}\exp\Big[\sum_{j\in N_i}\sum_{s_j\in K} u_{ij}(s_i,s_j)\,q_j(s_j)\Big]$$

Remark 2.

- The mean field approximation task is not convex $\Rightarrow$ algorithm terminates in local minima.
- The algorithm gives unary marginals only.

## B. Bethe approximation

From abstract view (exponential families) we know

- $\mu = \mathbb{E}_u(\varphi) = \nabla \underbrace{\log Z(u)}_{F(u)} \iff u \in \partial F^*(\mu)$ (FYI)

- $\log Z(u) = F(u) = \sup\big[\langle u,\mu\rangle - F^*(\mu)\big]$

- $F^*(\mu) = \inf_p \left\{ \sum_{s \in \mathcal{S}} p(s) \log p(s) \,\middle|\, \mathbb{E}_p(\Phi) = \mu,\ p \in \mathcal{P} \right\}$

(see assign. 3,5 seminar 1). This encapsulates $\mu \in \operatorname{conv} \Phi(\mathcal{S})$

Let us make two approximations

(1) relax $\operatorname{conv} \Phi(\mathcal{S})$ to $\mathbb{R}^n_+ \cap \operatorname{aff} \Phi(\mathcal{S})$ (see sec. 3)

(2) approximate the entropy $F^*(\mu)$ by

$$\tilde{F}(\mu) = \sum_{ij \in E} \sum_{s_i s_j} \mu_{ij}(s_i, s_j) \log \mu_{ij}(s_i, s_j) - \sum_{i \in V} (n_i - 1) \sum_{s_i} \mu_i(s_i) \log \mu_i(s_i)$$

(entropy of a tree)

Now solve

$$\left| \begin{array}{l} \langle u, \mu \rangle - \tilde{F}(\mu) \to \max\limits_{\mu} \\ \text{s.t.} \quad \mu \in \mathbb{R}^n_+ \cap \operatorname{aff} \Phi(\mathcal{S}) \end{array} \right.$$

<u>Problem</u>: If $(V, E)$ is a general graph (not a tree) then $\tilde{F}(\mu)$ is not convex on $\mathbb{R}^n_+ \cap \operatorname{aff} \Phi(\mathcal{S})$

↳ Apply a DC-algorithm (<u>D</u>ifference of <u>C</u>onvex functions) see later for details.