1. Random fields (undirected graphical models)

- A K-valued random field is a collection $\{S_i \mid i \in V\}$ of K-valued random variables
- $i \in V$ can be: pixels, object parts, etc.
- $k \in K$ can be: colours, segment labels, depth values etc.
- $s \in K^V \triangleq \mathcal{S}$ denotes realisations with $s_i \in K$, $i \in V$
- $p \in \mathcal{P}$ denotes a prob. distr. (density) $p: \mathcal{S} \to \mathbb{R}_+$

1A. Exponential families

- $S = \{S_i \mid i \in V\}$ is a K-valued random field
- $\Phi: \mathcal{S} = K^V \to \mathbb{R}^n$ is a random vector

Consider the task

$$\inf_{p \in \mathcal{P}} \Big\{ \sum_{s \in K^V} p(s) \log p(s) \;\Big|\; \mathbb{E}_p[\Phi] = \mu, \; \sum_{s \in K^V} p(s) = 1 \Big\}$$

Lagrange function

$$L(p) = \sum_{s \in K^V} p(s) \log p(s) - \langle u, \mathbb{E}_p[\Phi] - \mu \rangle - \lambda \Big[\sum_{s \in K^V} p(s) - 1 \Big]$$

$$\cdots \Rightarrow \quad p(s) = \exp\big[\langle u, \Phi(s) \rangle - \log Z(u) \big]$$

where $u \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$ are Lagrange multipliers and $Z(u)$ is the log-partition function

$$Z(u) = \sum_{s \in K^V} e^{\langle u, \Phi(s) \rangle}$$

Questions

(1) which $\mu \in \mathbb{R}^n$ qualify?

$$\mathbb{E}_p[\varphi] = \sum_{s\in K^V} p(s)\varphi(s) \qquad p(s)>0 \ \forall s\in K^V$$

$\Rightarrow$ μ is in the relative interior of conv $\varphi(\mathcal{S})$, where

$$\varphi(\mathcal{S}) = \{\varphi(s) \mid s\in\mathcal{S}\} \subset \mathbb{R}^n$$

(2) $u \mapsto \mu$? Yes, because

$$\nabla \log Z(u) = \frac{1}{Z(u)} \sum_{s\in K^V} e^{\langle u,\varphi(s)\rangle}\varphi(s) = \mathbb{E}_u[\varphi] = \mu$$

(3) $\mu \mapsto u$? (provided that $\mu \in ri(\text{conv}\,\varphi(\mathcal{S}))$)

No, not always. p_u is unique, but not u!

$$\langle u,\varphi(s)\rangle - \log Z(u) = \langle \tilde{u},\varphi(s)\rangle - \log Z(\tilde{u}) \quad \forall s\in K^V$$

i.e.

$$\langle u-\tilde{u},\varphi(s)\rangle = \text{const}_s \quad \forall s\in K^V$$

$$\{u\in\mathbb{R}^n \mid \langle u,\varphi(s)\rangle = \text{const}_s \ \forall s\in K^V\} = L^\perp$$

Two cases:

- $\text{aff}[\varphi(\mathcal{S})] = \mathbb{R}^n \Rightarrow L^\perp = \{0\} \Rightarrow \mu\mapsto u$ is a mapping
- $\text{aff}[\varphi(\mathcal{S})] = \{u\in\mathbb{R}^n \mid Au=b\} \neq \mathbb{R}^n$ with some $A:\mathbb{R}^n\to\mathbb{R}^m$, then $L=\text{Ker}A$, $L^\perp = \text{Im}A^T \Rightarrow$ we can re-parametrise u by

$$u \to u + A^T\psi, \quad \psi\in\mathbb{R}^m$$

and have

$$\langle u+A^T\psi,\varphi(s)\rangle = \langle u,\varphi(s)\rangle + \langle \psi, b\rangle$$

$$\log Z(u+A^T\psi) = \log Z(u) + \langle\psi,b\rangle$$

Remark 1 Recall the definition of an exponential family.

$s \in \mathcal{S}$ is a random variable with distr. (density)

$$p_u(s) = h(s) \exp\left[\langle \varphi(s), u\rangle - \log Z(u)\right]$$

with:

- $\varphi : \mathcal{S} \to \mathbb{R}^n$ is the sufficient statistic
- $u \in \mathbb{R}^n$ is the natural parameter
- $h(s)$ is a base measure
- $\log Z(u)$ is the log-partition function. □

B. Graphical models on undirected graphs

- $S = \{S_i \in K \mid i \in V\}$ is a K-valued random field
- (V, E) is an undirected graph
- fix marginal distributions $p(s_i, s_j) = \mu_{ij}(s_i, s_j)$ for all edges $\{i,j\} \in E$ and all $s_i, s_j \in K$

We search the p.d. with maximal entropy under the above constraints. We can apply A. because any marginal distr. $p(s_i, s_j)$ can be seen as expectation of the random variable $\varphi_{ij\,k,k'}(s) = \delta_{s_i k}\,\delta_{s_j k'}$.

Hence, if $\mu_{ij}(s_i, s_j)$ define a valid set of marginal distributions, then the task has the solution

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

where

$$\varphi(s) = \left(\varphi_{ij\,kk'}(s)\right)_{ij \in E,\, kk' \in K} \in \mathbb{R}^{|E||K|^2}$$

However, the <u>potentials</u> $u_{ij}: K^2 \to \mathbb{R}$ are defined up to re-parametrisations only.

The distribution $p_u(s)$ is a <u>Gibbs random field</u> and factorises over the edges of the graph (V, E).

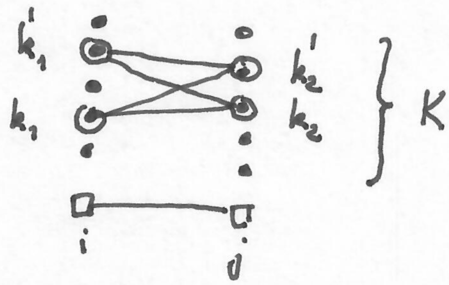
<u>Remark 2</u> All this can be generalised to (undirected) hypergraphs.

$\square$

Let us analyse the possible re-parametrisations: $u \to u + v$

a) Fix an edge $\{i,j\}$ and consider four labellings s^k, $k = 1,2,3,4$ which coincide on $V \setminus \{i,j\}$ and differ on $\{i,j\}$ as follows

$$(s_i^1, s_j^1) = (k_1, k_2)$$
$$(s_i^2, s_j^2) = (k_1, k_2')$$
$$(s_i^3, s_j^3) = (k_1', k_2)$$
$$(s_i^4, s_j^4) = (k_1', k_2')$$



We have

$$\langle v, \varphi(s^1) + \varphi(s^4) - \varphi(s^2) - \varphi(s^3) \rangle =$$
$$= v_{ij}(k_1, k_2) + v_{ij}(k_1', k_2') - v_{ij}(k_1, k_2') - v_{ij}(k_1', k_2) \overset{!}{=} 0$$

This holds for any edge $\{i,j\}$ and any k_1, k_2, k_1', k_2'.
It follows that

$$v_{ij}(s_i, s_j) = \psi_{ij}(s_i) + \psi_{ji}(s_j)$$

with some functions $\psi_{ij}(s_i)$ for <u>directed</u> edges

b) Fix a node $i \in V$ and consider two ~~···~~ labellings s^1, s^2 which differ on i only: $s_i^1 = k$, $s_i^2 = k'$.
We have

$$\langle v, \Phi(s^1) - \Phi(s^2) \rangle =$$

$$= \sum_{j \in \mathcal{N}_i} \varphi_{ij}(s_i = k) - \sum_{j \in \mathcal{N}_i} \varphi_{ij}(s_i = k') \overset{!}{=} 0$$

This holds for any node $i \in V$ and any $k, k' \in K$.
Thus it follows that

$$\sum_{j \in \mathcal{N}_i} \varphi_{ij}(s_i) = \text{const}_s$$

Concluding, we have all possible re-parametrisations given by

$$u_{ij}(s_i, s_j) \longrightarrow \varphi_{ij}(s_i) + u_{ij}(s_i, s_j) + \varphi_{ji}(s_j)$$

$$\text{s.t.} \quad \sum_{j \in \mathcal{N}_i} \varphi_{ij}(s_i) = c_i$$

Problems: All tasks

a) given potentials $u_{ij} : K^2 \to \mathbb{R}$ $\forall \{ij\} \in E$, compute $Z(u)$ and/or marginal prob's $p_u(s_i)$, $p_u(s_i, s_j)$

b) check whether $\mu_{ij} : K^2 \to \mathbb{R}_+$ $\forall \{ij\} \in E$ represent a consistent system of pairwise marginal prob's

c) given a consistent system of pairwise marginals $\mu_{ij} : K^2 \to \mathbb{R}_+$, $\forall \{ij\} \in E$, compute the potentials u_{ij}

d) given the potentials $u_{ij}\colon K^2 \to \mathbb{R}$ $\forall \{i,j\} \in E$, find the most probable realisations (labellings)

$$\underset{s \in K^V}{\operatorname{argmax}}\, p_u(s) = \underset{s \in K^V}{\operatorname{argmax}} \sum_{ij \in E} u_{ij}(s_i, s_j)$$

are <u>NP-hard</u>. Polynomial time complexity algorithms exist if (V, E) is acyclic or has low tree-width.