# II Graphical models on general graphs

## 11. Markov random fields, Gibbs random fields

Notations

- $(V,E)$ – undirected graph or hypergraph
- $\partial M$ – outer boundary of $M \subset V$, i.e.
  $$\partial M = \{ i \in V \setminus M \mid \exists j \in M \text{ s.t. } \{i,j\} \in E \}$$
- $S = \{S_i \mid i \in V\}$ – a field (collection) of $K$-valued random variables $S_i$ indexed by graph nodes $i \in V$.
  $S_M$, $M \subset V$ denotes a subset of them, i.e. $S_M = \{S_i \mid i \in M\}$
- $p(s)$ – a joint p.d. defined on $K^V$

Definition 1 A joint p.d. defined on $K^V$ is a Markov random field w.r.t. the graph structure $(V,E)$ if

$$p(s_M, s_{\tilde{M}} \mid s_{\partial M}) = p(s_M \mid s_{\partial M})\, p(s_{\tilde{M}} \mid s_{\partial M})$$

holds for each $M \subset V$ and $\tilde{M} = V \setminus (M \cup \partial M)$.

It follows that an MRF has the property

$$p(s_M \mid s_{\partial M}, s_{\tilde{M}}) = p(s_M \mid s_{\partial M})$$

Definition 2 Let $(V, \mathcal{C})$ be a hypergraph. A joint p.d. defined on $K^V$ is a Gibbs random field w.r.t. the hypergraph structure $(V, \mathcal{C})$ if it factorises into a product of functions depending on $s_c$, $c \in \mathcal{C}$, i.e.

$$p(s) = \prod_{c \in \mathcal{C}} f_c(s_c).$$

if $p(s)$ is strictly positive, it can be written as

$$p(s) = \frac{1}{Z} \exp \sum_{c \in \mathcal{C}} u_c(s_c),$$

where $u_c : K^c \to \mathbb{R}$ are arbitrary functions (aka Gibbs potentials) and $Z$ is a normalising constant. ▣

<u>Theorem 1</u> (Hammersley, Clifford, 1971)

Let $(V,E)$ be a graph and let $\mathcal{C}$ denote the system of its cliques. Every strictly positive MRF w.r.t. $(V,E)$ is also a GRF w.r.t. $(V,\mathcal{C})$ and vice versa. ▣

<u>Remark 1</u> Def. 2 does not require that $\mathcal{C}$ has to be the system of cliques for some graph. Consider e.g. a complete graph $(V,E)$. Every p.d. on $K^V$ is an MRF w.r.t. $(V,E)$. However, the class of GRFs w.r.t. $C = V \cup E$ is a proper subclass of p.d.s on $K^V$!

<u>Example 1</u> (segmentation of images)

- $x : V \to \mathbb{R}^3$ – a colour image defined on $V \subset \mathbb{Z}^2$
- $s : V \to K$ – a segmentation with labels from $K$

A model for a joint p.d. $p(x,s) = p(x|s)\,p(s)$

(1) $p(s)$ is a GRF w.r.t. the lattice $(V,E)$

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

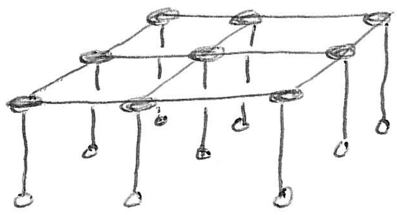
Simple variant: Potts model $u(k,k') = \alpha\,\delta_{kk'}$

(2) $p(x|s)$ is a conditionally independent appearance model

$$p(x|s) = \prod_{i \in V} p(x_i \mid s_i)$$

where $p(x_i|s_i)$ are e.g. (mixtures of) Gaussians

The model is a GRF w.r.t. the graph



Example 2 (segmentation of images)

$X, S$ as in the previous example. Now we model only $p(s|x)$: $(V, E)$ is the lattice as in the previous example, $\mathcal{C}$ is a system of subsets of $V$ (receptive fields) s.t. $\mathcal{C} = \{C_i \in V \mid i \in V \text{ and } i \in C_i\}$

$$p(s|x) = \frac{1}{Z(x)} \exp\Big[\sum_{ij \in E} u(s_i, s_j) + \sum_{i \in V} w(s_i, x_{C_i})\Big]$$

The functions $w(s_i, x_{C_i})$ can be e.g. implemented by a convolutional neural network.

Equivalent transformations for GRFs

Consider a GRF w.r.t. $C = V \cup E$ defined on $K^V$

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

Q: Are the functions $u_i, u_{ij}$ uniquely defined by $p(s)$?

(1) Clearly, adding a constant to any of them will not change $p$

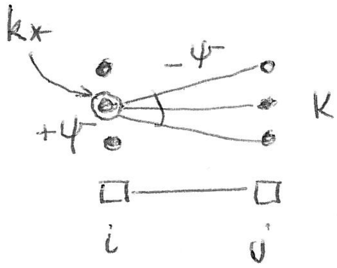
$$\tilde{u}_i(s_i) = u_i(s_i) + c$$

(2) Consider a node $i \in V$, an edge $\{i,j\} \in E$ and fix a state $k_* \in K$. Change

$$u_i(k_*) \longrightarrow u_i(k_*) + \varphi$$

$$u_{ij}(k_*, k) \longrightarrow u_{ij}(k_*, k) - \varphi \quad \forall k \in K$$

This will not change $p(s)$!




(3) „Elementary" transformations as in (2) can be applied for any triple $i \in V$, $\{i,j\} \in E$, $k \in K$:

$$\tilde{u}_i(s_i) = u_i(s_i) - \sum_{j \in N_i} \varphi_{ij}(s_i)$$

$$\tilde{u}_{ij}(s_i, s_j) = \varphi_{ij}(s_i) + u_{ij}(s_i, s_j) + \varphi_{ji}(s_j)$$

<u>Remark 2</u> The functions $u_{ij}$ are defined on undirected edges. We may think of them as

$$u_{ij}(s_i, s_j) = u_{ji}(s_j, s_i)$$

In contrast, the functions $\varphi_{ij}(s_i)$ are defined for oriented edges.

<u>Theorem 2</u> (w/o proof)

The equivalent transformations (aka reparametrisations) given above describe <u>all</u> possible equivalent transformations. $\boxtimes$