I Graphical models on general graphs

11. Markov Random Fields & Gibbs Random Fields

Notations

- (V, E) – undirected graph or hypergraph
- ∂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 (MRF) 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 (GRF) 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 $u_c \colon K^c \to \mathbb{R}$ are arbitrary functions (aka Gibbs potentials) and Z is a normalising constant. □

Theorem 1 (Hammersley, Clifford, 1971)

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

Remark 1 Def. 2 does not require that C has to be the system of cliques for some graph. Every p.d. on K^V is an MRF w.r.t. to the structure of a complete graph. However, the class of GRFs w.r.t. $C = V \cup E$ is a proper subclass of p.d.s on K^V even if (V,E) is the complete graph.

Example 1 (Segmentation of images, MRF)

- $x: V \to \mathbb{R}^3$ a color 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. $p(x,s) = p(x \mid s)\, p(s)$ with

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

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

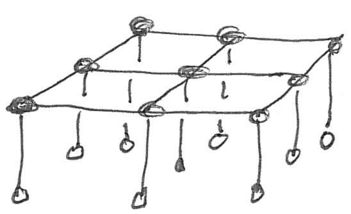
simple variant: Potts model $u(k,k') = d\,[\![k = k']\!]$

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

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

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

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



□

Example 2 (Segmentation of images, CRF)

Assume X, S as in the previous example. Now we model only $p(s|x)$. Let (V, E) denote the lattice as in the previous example and let $C = \{C_i \subset V \mid i \in V \text{ and } i \in C_i\}$ be a system of subsets of V (receptive fields). We model

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

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

Equivalent transformations for GRFs

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

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

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

$$p(s; u) = \frac{1}{Z(u)} e^{u(s)} \equiv \frac{1}{Z(v)} e^{v(s)} = p(s; v) \iff$$

$$u(s) \equiv v(s) + \text{const}.$$

(1) Clearly, adding a constant to any of the functions u_i, u_{ij} will not change $p(s)$

(2) Consider a node $i \in V$ and an edge $\{i,j\} \in E$. Choose any function $\varphi(s_i)$ and change the potentials

$$u_i(s_i) \to u_i(s_i) + \varphi(s_i)$$

$$u_{ij}(s_i, s_j) \to u_{ij}(s_i, s_j) - \varphi(s_i).$$

This will not change $p(s)$.

„Elementary" transformations as in (2) can be applied for any pair $i \in V$, $\{ij\} \in E$, giving

$$u_i(s_i) \to u_i(s_i) - \sum_{j \in N_i} \varphi_{ij}(s_i)$$

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

Remark 2 Recall that 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.

$\square$

Theorem 2 (w/o proof).

The equivalent transformations (aka reparametrisations) given above, describe all possible equivalent transformations. $\square$