I. Graphical models on general graphs

II. Markov Random Fields & Gibbs Random Fields

Notations

- \((V, E)\) - undirected graph or hypergraph
- \(\partial M\) - outer boundary of \(M \subseteq V\), i.e.
  \[\partial M = \{i \in V \mid \exists j \in M \text{ s.t. } \{ij\} \in E\}\]
- \(S = \{S_i \mid i \in V\}\) - a field (collection) of \(K\)-valued random variables \(S_i\) indexed by graph nodes \(i \in V\). \(S_M, M \subseteq V\)
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

\[p(s_M, s_{\partial M} \mid s_{\partial \partial M}) = p(s_M \mid s_{\partial M}) \cdot p(s_{\partial M} \mid s_{\partial \partial M})\]

holds for each \(M \subseteq V\) and \(\partial = V \setminus (M \cup \partial M)\).

It follows that an MRF has the property

\[p(s_M \mid s_{\partial M}, s_{\partial \partial M}) = p(s_M \mid s_{\partial M})\]

Definition 2. Let \((V, C)\) be a hypergraph. A joint p.d. defined on \(K^V\) is a Gibbs Random Field (GRF) w.r.t. the hypergraph structure \((V, C)\) if it factorises into a product of functions depending on \(S_c\), \(c \in C\), i.e.

\[p(s) = \prod_{c \in C} f_c(S_c)\]

If \(p(s)\) is strictly positive, it can be written as

\[p(s) = \frac{1}{Z} \exp \sum_{c \in C} u_c(S_c),\]

where \(u_c : K^c \rightarrow \mathbb{R}\) 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 of 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)
- \(X: V \to R^3\) 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. \(p(x,s) = p(x|s)p(s)\) with

1. \(p(s)\) is a GRF w.r.t. the lattice \((V,E)\)
   
   \[
   p(s) = \frac{1}{Z} \exp \sum_{i,j \in E} u(s_i, s_j)
   \]

   Simple variant: Potts model \(u(k,k') = \delta[k = k']\)

2. \(p(x|s)\) is a conditionally independent appearance model
   
   \[
   p(x|s) = \prod_{i \in V} p(x_i|s_i)
   \]

   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, 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 \( C = \{ c_i \in V \mid i \in V \text{ and } i \in c_i \} \) be a system of subsets of \( V \) (receptive fields). We model

\[
p(s|x) = \frac{1}{Z(x)} \exp \left[ \sum_{i \in V} U_i(s_i; s_j) + \sum_{i \in E} W_i(s_i, x_c_i) \right].
\]

The functions \( W_i(s_i, x_c_i) \) can be e.g. implemented by a convolutional deep network.

Equivalent transformations for CRFs

Let us consider a CRF w.r.t. \( C = V \cup E \) defined on \( K^V \)

\[
p(s) = \frac{1}{Z(u)} \exp \left[ \sum_{i \in V} U_i(s_i) + \sum_{i \in E} U_{ij}(s_i, s_j) \right]
\]

Are the functions \( U_i, U_{ij} \) uniquely defined by \( p(s) \)?

\[
p(s; u) = \frac{1}{Z(u)} e^{U_i(s)} \equiv \frac{1}{Z(v)} e^{\psi(s)} = p(s; v) \iff U(s) = \psi(s) + \text{const.}
\]

(a) Clearly, adding a constant to any of the functions \( U_i, U_{ij} \) will not change \( p(s) \).

(b) Consider a node \( i \in V \) and an edge \( i; j \in E \). Choose any function \( \psi(s_i) \) and change the potentials

\[
U_i(s_i) \rightarrow U_i(s_i) + \psi(s_i)
\]
\[
U_{ij}(s_i, s_j) \rightarrow U_{ij}(s_i, s_j) - \psi(s_i).
\]

This will not change \( p(s) \).
"Elementary" transformations as in (2) can be applied for any pair \( i \in V, \ j \in E, \) giving

\[
\begin{align*}
U_i(s_i) & \Rightarrow U_i(s_i) - \sum_{j \in V} U_{ij}(s_i) \\
U_{ij}(s_i, s_j) & \Rightarrow U_{ij}(s_i) + U_{ij}(s_i, s_j) + U_{ji}(s_j)
\end{align*}
\]

\textbf{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 \( U_{ij}(s_i) \) are defined for oriented edges.

\textbf{Theorem 2 (w/o proof).}

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