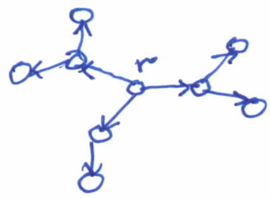
11. Hidden Markov Models on acyclic graphs

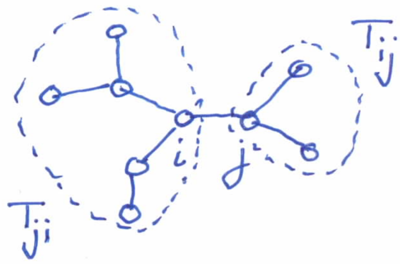
Notation

Let $T=(V,E)$ be an undirected, connected acyclic graph

- Fixing an arbitrary vertex $r \in V$, we denote by $\vec{E}_r$ the edge set of the corresponding rooted digraph



- For each oriented edge ij, $\{i,j\} \in E$ we define the subtree T_{ij} by
$$V(T_{ij}) = \{ m \in V \mid path(i,m) \ni j \}$$



A. The model

Definition 1a Let $T=(V,E)$ be an undirected tree and s_i, $i \in V$ be K-valued random variables. A p.d. for the random field $s \in K^{|V|}$ is a Markov model on T if

$$p(s) = p(s_r) \prod_{ij \in \vec{E}_r} p(s_j \mid s_i)$$

holds for any choice $r \in V$ (root). ■

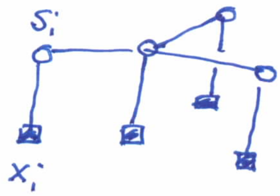
Definition 1b Equivalently, a p.d. $p(s)$ for $s \in K^{|V|}$ is a Markov model on T if

$$p(s) = \prod_{\{i,j\} \in E} p(s_i, s_j) \Big/ \prod_{i \in V} p^{n_i - 1}(s_i)$$

holds, where n_i is the degree of vertex $i \in V$. ■

An HMM on an undirected tree $T=(V,E)$ is a p.d. for pairs $s \in K^{|V|}$, $x \in F^{|V|}$, where F is a feature space and $p(x,s) = p(x|s)\,p(s)$, where

- $p(s)$ is a Markov model on T
- $p(x|s) = \prod_{i \in V} p(x_i | s_i)$
  (conditional independence)



B. Inference tasks

Consider the following task: Given an observation field $x \in F^{|V|}$, compute

$$p(x) = \sum_{s \in K^{|V|}}' p(x,s)$$

By substituting the model Def. 1a we get

$$p(x) = \sum_{s \in K^{|V|}}' p(s_r)\,p(x_r|s_r) \prod_{ij \in \vec{E}_r} p(s_j|s_i)\,p(x_j|s_j)$$

The problem has the form (fixed observation!)

$$\sum_{s \in K^{|V|}}' \prod_{i \in V} \varphi_i(s_i) \prod_{ij \in \vec{E}_r} f_{ij}(s_i, s_j)$$

where:

$$\varphi_i(s_i) = \begin{cases} p(s_r)\,p(x_r|s_r) & \text{if } i=r \\ 1 & \text{otherwise} \end{cases}$$

$$f_{ij}(s_i, s_j) = p(s_j|s_i)\,p(x_j|s_j)$$

The algorithm recomputes the φ-s starting from an arbitrary <u>leaf</u> $j \in V$. Let $ij \in \vec{E}_r$ denote the only incoming edge.

$$\varphi_i(s_i) := \varphi_i(s_i) \sum_{s_j \in K}' f_{ij}(s_i, s_j)\, \varphi_j(s_j)$$

Then the leaf j is removed. This is repeated until only the root r remains. Finally

$$p(x) = \sum_{k \in K}' \varphi_r(k)$$

Complexity: $|K|^2 |E|$

<u>Remark 1</u> The same approach is applied for solving the task

$$s_* \in \underset{s \in K^{|V|}}{\operatorname{argmax}} \log p(x, s)$$

simply by replacing operations

$$\times \mapsto +$$
$$+ \mapsto \max$$

1C. <u>Computing marginals</u>

Consider the following task: Given an observation field $x \in F^{|V|}$, compute marginal prob's

$$p(x, s_i), \quad \forall i \in V, \quad \forall s_i \in K$$

Recall the algorithm for computing marginals of an HMM on a chain (see sec. 5)

<u>Here:</u>

- It follows from Def 16 that

$$p(x, s_i) = p(s_i)\, p(x_i | s_i) \prod_{j \in N_i} p(x_{T_{ij}} | s_i)$$

- Let us denote $\varphi_{ij}(s_i) = p(x_{T_{ij}} \mid s_i)$. The φ-s fulfil the following system of equations

$$\varphi_{ij}(s_i) = \sum_{s_j}' p(s_j \mid s_i)\, p(x_j \mid s_j) \prod_{\substack{\ell \in N_j \\ \ell \neq i}} \varphi_{j\ell}(s_j)$$

- Two passes through all edges of T are needed to compute all of them $\Rightarrow$ complexity $2|\mathbf{K}|^2|E|$

D. Learning the tree structure

Given: i.i.d. sample of realisations $s \in K^{|V|}$

Task: estimate parameters and structure of the model

Denote by $\beta(s)$ the empirical p.d. assoc. with the sample.
MLE for known structure $T = (V, E)$

$$\sum_{s \in K^{|V|}}' \beta(s) \log p(s) \overset{\text{Def. 16}}{=} \cdots\cdots$$

$$= \sum_{\{ij\} \in E}' \sum_{s_i, s_j}' \beta(s_i, s_j) \log p(s_i, s_j) - \sum_{i \in V}'(n_i - 1) \sum_{s_i}' \beta(s_i) \log p(s_i)$$

Maximiser given by

$$p(s_i, s_j) = \beta(s_i, s_j) \quad \forall \{ij\} \in E$$

$$p(s_i) = \beta(s_i) \quad \forall i \in V$$

It remains to solve

$$\sum_{\{ij\} \in E}' \sum_{s_i, s_j}' \beta(s_i, s_j) \log \beta(s_i, s_j) - \sum_{i \in V}'(n_i - 1) \sum_{s_i}' \beta(s_i) \log \beta(s_i) \rightarrow \max_{E:\ \text{tree}}$$

Both terms depend on the edge structure E. Rearrange:

$$\sum_{\{ij\}\in E} \Big\{ \sum_{s_i s_j} \beta(s_i s_j) \log \beta(s_i s_j) - \sum_{s_i} \beta(s_i) \log \beta(s_i) - \sum_{s_j} \beta(s_j) \log \beta(s_j) \Big\} +$$

$$+ \sum_{i\in V} \sum_{s_i} \beta(s_i) \log \beta(s_i)$$

The last term does not depend on the edge structure. Denote the values of the curly brackets by h_{ij}:

$$\sum_{\{ij\}\in E} h_{ij} \to \max_{E:\,\text{tree}}$$

$\Rightarrow$ Maximum spanning tree problem. Can be solved in polyn. time

<u>Remark 1</u> Similar problem for chain $\to$ travelling salesman problem $\to$ NP hard.