ad hoc

June 7, 2019

This question concerns the **M-step**, when tying parameters together. We assume here one tree. Here's what they suggest in Crouse.
For:

- the ith element in the tree ($i = 1, \ldots, 2^J$),
- which belongs to a group, $[i]$, of tree-elements, which are to be tied together for estimating parameters,
- at the l-th iteration of the EM algorithm,
- with wavelet coefficient vector, $\mathbf{w}$ and parameter vector, $\boldsymbol{\theta}^l$,

$$P(S_i = m \mid \boldsymbol{\theta}) := \pi_i^{(m)} = P(S_i = m \mid \mathbf{w}, \boldsymbol{\theta}^l)$$

$$P(S_i = m \mid S_{p(i)} = n, \boldsymbol{\theta}) := \epsilon_{i,p(i)}^{mn} = \frac{P(S_i = m, S_{p(i)} = n \mid \mathbf{w}, \boldsymbol{\theta}^l)}{\pi_{p(i)}^{(n)}}$$

With tying, they suggest: (with $\mid [i] \mid$ being the number of elements in $[i]$)

$$\pi_i^{(m)} = \frac{1}{\mid [i] \mid} \sum_{j \in [i]} P(S_j = m \mid \mathbf{w}, \boldsymbol{\theta}^l)$$

$$\epsilon_{i,p(i)}^{mn} = \frac{1}{\pi_{p(i)}^{(n)}} \frac{1}{\mid [i] \mid} \sum_{j \in [i]} P(S_j = m, S_{p(j)} = n \mid \mathbf{w}, \boldsymbol{\theta}^l) \tag{1}$$

I was more thinking that expression (1) should be:

$$\epsilon_{i,p(i)}^{mn} = \frac{1}{\mid [i] \mid} \sum_{j \in [i]} \frac{P(S_j = m, S_{p(j)} = n \mid \mathbf{w}, \boldsymbol{\theta}^l)}{\pi_{p(j)}^{(n)}} \tag{2}$$

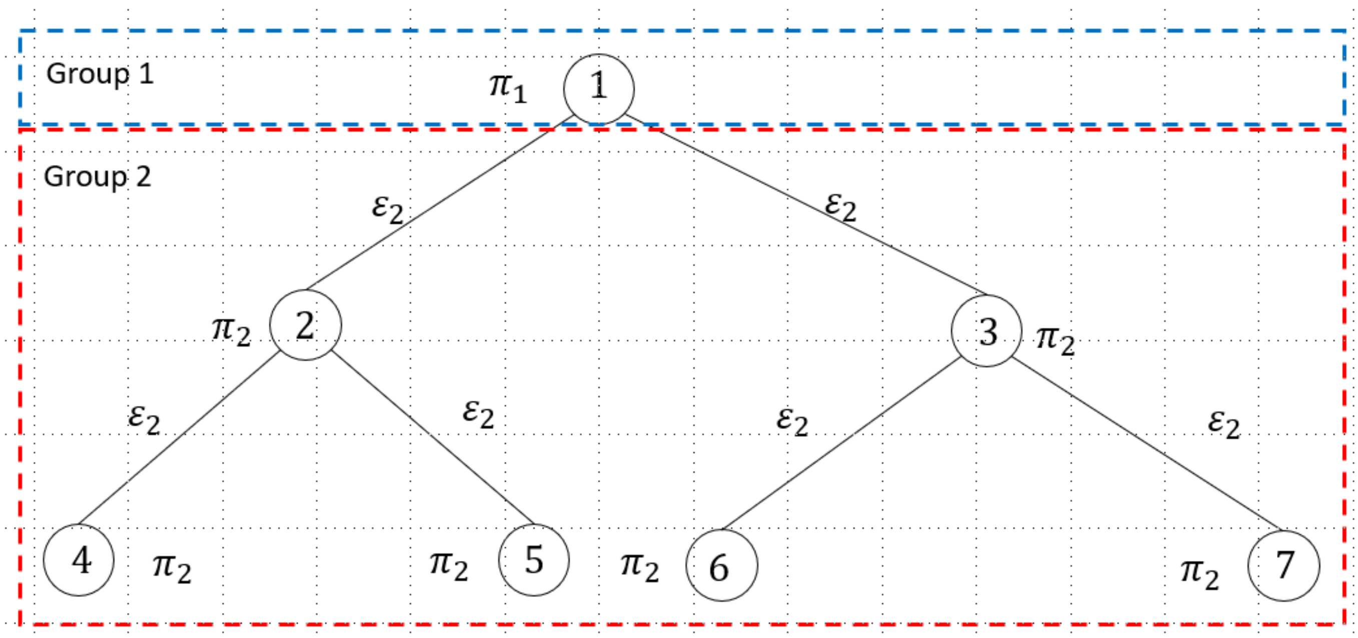
as different elements of the tree, j, within a group, may have different parents. Ie my approach is to average all the transition probabilities (based on the previously tied π), rather than assuming some sort of identical parent state probability for a whole group, i.

What I realised is that this formulation by Crouse is (I THINK) only specific to the tying example they've provided (tied by tree level). It makes sense here because all tree levels are tied, and therefore are parameterised to have the same π. Then, my formulation in (2) would be the same as (1), as $\pi_{p(j)}$ would be the same for all $j \in [i]$ (in english: all elements in a level have parents from the same level. These parents are the same as they're tied together). So in this case, I'd think mine is the more general expression...but i'd need to derive this to show it formally (or find the reference where tying is defined).

In my mind, there is one main situation where Crouse's formulation has an issue, and where I think my formulation would work more properly:

1. A group which has several tree levels worth of elements
   (a) The problem here is that there are different $\pi_{p(i)}$ values, depending on which level of the tree you are in.

Consider the example below:



Some have parents with π_1 and others with π_2, so in my opinion the parameterisation would need to be:

$$\epsilon^{mn}_{2,p(2)} = \frac{1}{6}\Bigg(\frac{P(S_2 = m, S_1 = n \mid \theta) + P(S_3 = m, S_1 = n \mid \theta)}{\pi_1^{(n)}} + \dots$$
$$\frac{P(S_4 = m, S_2 = n \mid \theta) + P(S_5 = m, S_2 = n \mid \theta) + P(S_6 = m, S_3 = n \mid \theta) + P(S_7 = m, S_3 = n \mid \theta)}{\pi_2^{(n)}}\Bigg)$$

where $\epsilon^{mn}_{2,p(2)}$ would need to be interpreted as 'transition probability to an element in group 2 from that element's parent (regardless of whether the parent was in group 1, or group 2)'.

In contrast to what you mentioned today, in the case where our 'top group' (the group including the root of the tree) ties together elements from several levels of the tree, I don't think there'd be an interpretation issue:

- A tied π_i just means we are attributing the same prior state probability for each element of the group – it doesn't imply anything about existence of transitions for the top group does it?
- Can still parameterise $\epsilon_{i,p(i)}$ based on the transitions which occur within the group, based on the above tied π_i, with its interpretation being identical to what we had for the example above.