Advanced Method in ML - Exercise 2

Avia Efrat 300928538
Mor Shoham 300953965

March 30, 2019

Question 2

(a) To simplify the notations below, we assume WLOG that i=1 and j=2.
We wish to find $p(x_1, x_2)$, and show it is same as $\mu_{ij}(x_1, x_2)$.
It is given that $p(x_1, x_2, ..., x_n) \propto \prod\limits_{a,b\in E} \phi_{ab}(x_a, x_b)$.
Therefore, by summing on other variables we get:

$$p(x_1, x_2) = \sum_{x_3,\ldots,x_n} p(x_1, x_2, ..., x_n) \propto \sum_{x_3,\ldots,x_n} \prod_{a,b\in E} \phi_{ab}(x_a, x_b)$$

We don't sum on $x_1$and $x_2$so we can write:

$$\sum_{x_3,\ldots,x_n} \prod_{a,b\in E} \phi_{ab}(x_a, x_b) = \phi_{12}(x_1, x_2) \sum_{x_3,\ldots,x_n} \prod_{a,b\in E\setminus\{(x_1,x_2)\}} \phi_{ab}(x_a, x_b)$$

The graph is a tree, so we can devide the variables $x_3, ..., x_n$ to 2 disjoint sets, V_1 and V_2, such that $V_1$are the variables in the subtree of x_1 excluding the part in the direction of x_2, and V_2 are the variables in the subtree of x_2 excluding the part in the direction of x_1.

Let's also define the sets of edges for each subtree: $E_1 = \{(x_a, x_b)|(x_a, x_b) \in E, x_a, x_b \in V_1\}$ and $E_2 = \{(x_a, x_b)|(x_a, x_b) \in E, x_a, x_b \in V_2\}$

Now we can split the sum:

$$\phi_{12}(x_1, x_2) \sum_{x_3,\ldots,x_n} \prod_{a,b\in E\setminus\{(x_1,x_2)\}} \phi_{ab}(x_a, x_b) = \phi_{12}(x_1, x_2) \sum_{V_1\setminus\{x_1\}} \prod_{a,b\in E_1} \phi_{ab}(x_a, x_b) \sum_{V_2\setminus\{x_2\}} \prod_{a,b\in E_2} \phi_{ab}(x_a, x_b)$$

And now, because it is a tree, we can also divide V_1 without $x_1$to disjoint sets, one for each neighbor of x_1 in V_1, marked as $V_{N_k}$for the neighbor k, with edges E_{N_k}.

Thus we can know:

$$\sum_{V_1\setminus\{x_1\}} \prod_{a,b\in E_1} \phi_{ab}(x_a, x_b) = \prod_{k\in N(1)\setminus\{2\}} \left[\phi_{k1}(x_k, x_1) \sum_{V_{N_k}} \prod_{a,b\in E_{N_k}} \phi_{ab}(x_a, x_b)\right] = \prod_{k\in N(1)\setminus\{2\}} m_{k,1(x_1)}$$

The second equality is also because it is a tree.
Due to symmetry, we get:

$$\sum_{V_2\setminus\{x_2\}} \prod_{a,b\in E_2} \phi_{ab}(x_a,x_b) = \prod_{k\in N(2)\setminus\{1\}} m_{k,2(x_2)}$$

Plugging back what we learned, we get:

$$p(x_1,x_2) \propto \sum_{x_3,\ldots,x_n} \prod_{a,b\in E} \phi_{ab}(x_a,x_b) = \phi_{12}(x_1,x_2) \prod_{k\in N(1)\setminus\{2\}} m_{k,1(x_1)} \prod_{k\in N(2)\setminus\{1\}} m_{k,2(x_2)}$$

And this is exactly $\mu_{ij}(x_1,x_2)$, like we wanted.

(b) We start by developing $\sum\limits_{x_j}\mu_{ij}$:

$$\sum_{x_j}\mu_{ij} \propto \sum_{x_j}\left[\phi_{ij}(x_i,x_j) \prod_{k\in N(i)\setminus\{j\}} m_{k,i(x_i)} \prod_{k\in N(j)\setminus\{i\}} m_{k,j(x_j)}\right] =$$

$$= \sum_{x_j}\left[\phi_{ij}(x_i,x_j) \prod_{k\in N(j)\setminus\{i\}} m_{k,j(x_j)}\right] \prod_{k\in N(i)\setminus\{j\}} m_{k,i(x_i)}$$

$\sum\limits_{x_j}\left[\phi_{ij}(x_i,x_j) \prod\limits_{k\in N(j)\setminus\{i\}} m_{k,j(x_j)}\right]$ is the message from j to i that is being updated every round, and if we are at fixed point of LBP, then they are equal.
Plugging it back we get:

$$\sum_{x_j}\mu_{ij} \propto m_{ji}(x_i) \prod_{k\in N(i)\setminus\{j\}} m_{k,i(x_i)} = \prod_{k\in N(i)} m_{k,i(x_i)} \propto \mu_i$$

(c) We start by developing the right side, by the definition of μ_i and μ_{ii}:

$$\prod_i \mu_i(x_i) \prod_{i,j\in E} \frac{\mu_{ij}(x_i,x_j)}{\mu_i(x_i)\mu_j(x_j)} \propto \prod_i \prod_{k\in N(i)} m_{k,i(x_i)} \prod_{i,j\in E} \frac{\phi_{ij}(x_i,x_j) \prod\limits_{k\in N(i)\setminus\{j\}} m_{k,i(x_i)} \prod\limits_{k\in N(j)\setminus\{i\}} m_{k,j(x_j)}}{\prod\limits_{k\in N(i)} m_{k,i(x_i)} \prod\limits_{k\in N(j)} m_{k,j(x_j)}} = \tag{1}$$

$$= \prod_i \left[\prod_{k\in N(i)} m_{k,i(x_i)}\right] \prod_{i,j\in E} \left[\frac{\phi_{ij}(x_i,x_j) \prod\limits_{k\in N(i)\setminus\{j\}} m_{k,i(x_i)} \prod\limits_{k\in N(j)\setminus\{i\}} m_{k,j(x_j)}}{m_{j,i}(x_i)m_{i,j}(x_j) \prod\limits_{k\in N(i)\setminus\{j\}} m_{k,i(x_i)} \prod\limits_{k\in N(j)\setminus\{i\}} m_{k,j(x_j)}}\right] = \tag{2}$$

$$= \prod_{i} \left[\prod_{k \in N(i)} m_{k,i(x_i)} \right] \prod_{i,j \in E} \left[\frac{\phi_{ij}(x_i, x_j)}{m_{j,i}(x_i) m_{i,j}(x_j)} \right] = \tag{3}$$

$$= \prod_{i,j \in E} [\phi_{ij}(x_i, x_j)] \prod_{i} \left[\prod_{k \in N(i)} m_{k,i(x_i)} \right] \prod_{i,j \in E} \left[\frac{1}{m_{j,i}(x_i) m_{i,j}(x_j)} \right] = \tag{4}$$

$$= \prod_{i,j \in E} [\phi_{ij}(x_i, x_j)] \left[\frac{\prod_{i} \left[\prod_{k \in N(i)} m_{k,i(x_i)} \right]}{\prod_{i,j \in E} m_{j,i}(x_i) m_{i,j}(x_j)} \right] \tag{5}$$

The move from (1) to (2) is extracting one element from each product in the denominator in order to get identical terms in the denominator and nominator, which we remove at (3).

The move from (3) to (4) and (4) to (5) is re organization of the products.

At (5), we have the term $\frac{\prod_{i} \left[\prod_{k \in N(i)} m_{k,i(x_i)} \right]}{\prod_{i,j \in E} m_{j,i}(x_i) m_{i,j}(x_j)}$. In the denominator we go over all edges and multiply the messages in both direction.

In the nominator we go over each edge twice, one from each vertex of it, and add the message to the overall product.

Both achieve the same result, and therefore:

$$\frac{\prod_{i} \left[\prod_{k \in N(i)} m_{k,i(x_i)} \right]}{\prod_{i,j \in E} m_{j,i}(x_i) m_{i,j}(x_j)} = 1$$

Plugging it back, we get:

$$\prod_{i} \mu_i(x_i) \prod_{i,j \in E} \frac{\mu_{ij}(x_i, x_j)}{\mu_i(x_i) \mu_j(x_j)} \propto \prod_{i,j \in E} [\phi_{ij}(x_i, x_j)] \propto p(x_1, x_2, ..., x_n)$$

Question 3

(a) Let G be a graph and π an elimination order.

Assume by contradiction that the induced graph G_π is not chordal.

Therfore, there must be a cycle of size greater than 3 that does not contain a chord. Let C be such a cycle.

Let x_i be the first vertex in C that is being eliminitaed in π, and let x_k and $x_m$be his adjacenct vertices on C.

When eliminating x_i, we add the edge (x_k, x_m) to G_π. But this is a contradiction to the fact we took C as a cycle of length greater than 3 without a chord. Therefore, G_π is chordal.

(b) Let $G = (V, E)$ a chordal graph.

Following is an algorithm to create an elimination order such that we never add a new edge.

1. $E' = E$
2. $V' = V$
3. $G' = (V', E')$
4. $\pi = []$
5. While V' is not empty:
   (a) Let **v** be a vertex with minimal degree in G'
      i. Append v to π
      ii. $V' = V' \setminus \{v\}$
      iii. $E' = E' \setminus \{(v, u) | (v, u) \in E'\}$
      iv. $G' = (V', E')$

We next prove that correctness of the algorithm.

Termination: The algorithm terminates, as it runs exactly $|V|$ iterations, because we remove one vertex at each iteration and stop when there are none.

Correctness:

We need to show that for every elimination step, we never add a new edge.

Let's look at the degree of the vertex that we remove at some iteration:

- If the degree is 0, we don't add an edge because it has no neighbors.
- If the degree is 1, we don't add an edge because it has one neighbor and the edge already exists.
- If the degree is 2, the vertex has to be part of a cycle, because if it is not a cycle, there is other vertex with degree 1 and it would be chosen.
  Assume by contradiction that there is no edge between the 2 neighbors. It means that the cycle is of size greater than 3, but we know that the graph is chordal,
  so if there is no edge between them, there must be a cycle of size greater than 3 without chord, which will be the minimal length cycle that contain the 3 vertices.
  This is contradiction to G being chrodal. Therefore the edge exists, and we don't add a new edge here.
- If the degree is 3 or more, we can look at the neighbors by couples, one couple at a time, and for each couple, we know by the same analysis that we did for degree=2, that there must be an edge between them.
  Therefire we don't edge in this case as well.

We covered all options, and thereore, in the elimination order we described, we never add a new edge. Thus $G = G_\pi$.