Exercise: The Sum-Product Algorithm - A Numerical Calculation

Bookmark this page

Exercise: The Sum-Product Algorithm - A Numerical Calculation

3/3 points (graded)

Consider a chain graph $X_1 \leftrightarrow X_2 \leftrightarrow X_3$, with the following edge potentials:

ψ_{12}		$X_2 = 0$	$X_2 = 1$
X_1	0	5	1
	1	1	5

ψ_{23}		$X_3 = 0$	$X_3 = 1$
X_2	0	0	1
	1	1	0

The node potentials are just functions that always output 1.

- **(a)** Compute p_{X_1,X_3}. What is $p_{X_3|X_1}(0|0)$? For this part, do the calculation without using the sum-product algorithm, using what you learned from the first part of the course. Note that at times, you may have unnormalized quantities that you normalize (to sum to 1) at the end, which is fine.

  (Please be precise with at least 3 decimal places, unless of course the answer doesn't need that many decimal places. You could also put a fraction.)

$p_{X_3|X_1}(0|0) =$ 1/6 ✔ **Answer:** 1/6

- **(b)** Let's use the sum-product algorithm to show what the probability of $X_3 = 0$ is given that $X_1 = 0$.

  Start this problem by incorporating the observation to create a new graphical model that only has 2 nodes. Then compute the message $m_{2\to 3}$ (remember that this is a table). What is the table? In providing your answer, please normalize the message table so that its entries sum to 1.

  (Please be precise with at least 3 decimal places, unless of course the answer doesn't need that many decimal places. You could also put a fraction.)

  $m_{2\to 3}(0) =$ 1/6 ✔ **Answer:** 1/6

  $m_{2\to 3}(1) =$ 5/6 ✔ **Answer:** 5/6

  Use the last part of the sum-product algorithm that computes the marginal distribution using incoming messages. You should verify for yourself that the marginal distribution that you get for X_3 (in the graph that already accounts for conditioning on $X_1 = 0$) agrees with your answer to part (a) for the value of $p_{X_3|X_1}(0|0)$.

Solution:

- **(a)** Compute p_{X_1,X_3}. What is $p_{X_3|X_1}(0|0)$? For this part, do the calculation without using the sum-product algorithm, using what you learned from the first part of the course. Note that at times, you may have unnormalized quantities that you normalize (to sum to 1) at the end, which is fine.

  Solution: We have

$$\begin{aligned} p_{X_1,X_3}(x_1,x_3) &= \sum_{x_2} p_{X_1,X_2,X_3}(x_1,x_2,x_3) \propto \overbrace{\sum_{x_2} \psi_{12}(x_1,x_2)\psi_{23}(x_2,x_3)}^{\text{this is matrix multiplication}} \\ &= \psi_{12}(x_1,0)\psi_{23}(0,x_3) + \psi_{12}(x_1,1)\psi_{23}(1,x_3). \end{aligned}$$

Thus, p_{X_1,X_3} is proportional to the following table (you could get this from matrix multiplication or by brute force):

		X_3 = 0	X_3 = 1
X_1	0	$5 \cdot 0 + 1 \cdot 1 = 1$	$5 \cdot 1 + 1 \cdot 0 = 5$
	1	$1 \cdot 0 + 5 \cdot 1 = 5$	$1 \cdot 1 + 5 \cdot 0 = 1$

Upon normalization, we procure p_{X_1,X_3}:

p_{X_1,X_3}		X_3 = 0	X_3 = 1
X_1	0	$1/12$	$5/12$
	1	$5/12$	$1/12$

Thus,

$$p_{X_3|X_1}(0|0) = \frac{1/12}{1/12 + 5/12} = 1/6.$$

- **(b)** Let's use the sum-product algorithm to show what the probability of $X_3 = 0$ is given that $X_1 = 0$.

  Solution: We have

  $$p_{X_2,X_3|X_1}(x_2, x_3|0) \propto \underbrace{\psi_{12}(0, x_2)}_{\triangleq \tilde{\phi}_2(x_2)} \psi_{23}(x_2, x_3),$$

  which is actually just a two node graph over X_2 and X_3. The node potential for X_2 is $\widetilde{\phi}_2$, the edge potential is ψ_{23}, and the node potential for X_3 is identically 1. Computing a single message is sufficient:

  $$m_{2\to 3}(x_3) = \sum_{x_2} \underbrace{\psi_{12}(0, x_2)}_{\triangleq \tilde{\phi}_2(x_2)} \psi_{23}(x_2, x_3) = 5\psi_{23}(0, x_3) + 1\psi_{23}(1, x_3),$$

  which, in vector form, corresponds to 5 times the first row of table ψ_{23} (corresponding to $X_2 = 0$) plus 1 times the second row of table ψ_{23} (corresponding to $X_2 = 1$):

  $$m_{2\to 3} = 5 \cdot (0 \quad 1) + (1 \quad 0) = (1 \quad 5).$$

Thus, normalizing so the message table sums to 1:

- $m_{2\to3}(0) = \boxed{1/6}$
- $m_{2\to3}(1) = \boxed{5/6}$

Note that we can actually always normalize messages in this way because at the very end when we compute the marginal distributions, we will re-normalize anyways!

Then the node marginal at X_3 is

$$p_{X_3}(x_3) \propto \phi_3(x_3)m_{2\to3}(x_3) = m_{2\to3}(x_3) = \begin{pmatrix} 1/6 & 5/6 \end{pmatrix},$$

from which we conclude that $p_{X_3|X_1}(0|0) = 1/6$ (here we re-introduce X_1 into the notation, remembering that the 2 node graphical model we are working with actually was conditioned on $X_1 = 0$).

Note that if we did not normalize the message table, then that is fine too since in the last step, we would instead have

$$p_{X_3}(x_3) \propto \phi_3(x_3)m_{2\to3}(x_3) = m_{2\to3}(x_3) = \begin{pmatrix} 1 & 5 \end{pmatrix},$$

at which point we would normalize anyways to get the same answer for what the marginal p_{X_3} at X_3 is, again where it's actually the posterior distribution $p_{X_3|X_1}(\cdot \mid 0)$.

Submit

You have used 2 of 5 attempts

✔ Correct (3/3 points)

Hide Discussion

Add A Post

Need help for part a

discussion posted 3 days ago by **lokeshhh**

I can't solve part (a). I know that:

$$p_{X_1,X_3}(x_1,x_3) = \sum_{x_2} \frac{1}{Z}\phi_1(x_1)\phi_2(x_2)\phi_3(x_3) \cdot \psi_{12}(x_1,x_2)\psi_{23}(x_2,x_3)$$

But how do I find ϕ_1, ϕ_2 and ϕ_3

This post is visible to everyone.

Add A Response

3 responses

lokeshhh

3 days ago

I did some reading and I think ϕ_1, ϕ_2 and ϕ_3 should be 1. Right?

Yes.

Collapse discussion

Confusing: Exercise: The Sum-Product Algorithm - A Numerical Calculation

discussion posted a day ago by **Methanogen**

The numerical example with the values of 5 and 1 is confusing to me, as was the commentary from the prior week (which led to very many questions). How are these inflated values to be interpreted? For example, are we to infer that if X1 has the value 1, then the probability of X2 also having the value 1 is : 5/(5+1) ? Any gracious clarifications much appreciated!

This post is visible to everyone.

Add A Response

1 response

Mark_B2 **Community TA**

about 14 hours ago

That's correct. One can say the odds of $X_2 == 1$ are 5 to 1.

Add a comment

Collapse discussion

part b: Incorporating the observation?

discussion posted 3 days ago by **ripande**

The problem says: "Start this problem by incorporating the observation to create a new graphical model that only has 2 nodes." I am not sure...

This post is visible to everyone.

+ Expand discussion

Need help in understanding this

discussion posted 2 days ago by **ripande**

$P(X2, X3|X1 = 0) \propto \psi_{1,2}(X1 = 0, X2) * \psi_{2,3}(X2, X3)$

But is it not the case that it is $P(X2, X3, X1 = 0)$ that is proportional...

This post is visible to everyone.

+ Expand discussion

Numerical example for Sum-Product algorithms

discussion posted 4 days ago by **nar3k**

Hey guys!

For people who like to see examples there is a presentations where you can find it

http://www.cse.psu.edu/~rtc12/CSE586/lectures/cse586GMplusMP_6pp.pdf

...

This post is visible to everyone.

+ Expand discussion

What is the difference between probability & potential functions

discussion posted 4 days ago by **ripande**

1. I am still a little confused about the difference between probability and potential functions. I had thought that potential functions are relative weights which might not sum to one, but once you normalize them they should be same as probability. However, in a discussion from previous thread, the instructor had explained that it is possible for potentials to not be same as probabilities even after you normalize them. Hence, I am confused as to what potentials really are. Can someone explain please? Is there any good reference where I can gain more clarity about how potentials are different from probabilities.

2. Should I normalize the potential tables to convert them into probability tables before working on them for part a?

This post is visible to everyone.

Add A Response

2 responses

kiwitrader **Community TA**

4 days ago

Collapse discussion

What two nodes are we supposed to start with on the second part? What observation are we supposed to "incorporate"?

discussion posted 4 days ago by **EzraSchroeder**

The instructions for part (b) say in part "Start this problem by incorporating the observation to create a new graphical model that only has 2 nodes." What observation are we supposed to incorporate? What two nodes would that leave us with in the new graphical model? Thanks.

This post is visible to everyone.

Add A Response

1 response

studymcr

4 days ago

"(b) Let's use the sum-product algorithm to show what the probability of X3=0 is given that X1=0."

"given that X1=0" <-- is the observation

Collapse discussion

1