

1 Dependencies in a Directed Graphical Model

1.1 In the graphical model of Figure 1, is $W_{d,n} \perp W_{d,n+1} \mid \theta_d, \beta_{1:K}$?

Yes.

1.2 In the graphical model of Figure 1, is $\theta_d \perp \theta_{d+1} \mid Z_{d,1:N}$?

No.

1.3 In the graphical model of Figure 1, is $\theta_d \perp \theta_{d+1} \mid \alpha, Z_{1:D,1:N}$?

Yes.

1.4 In the graphical model of Figure 2, is $W_{d,n} \perp W_{d,n+1} \mid \Lambda_d, \beta_{1:K}$?

No.

1.5 In the graphical model of Figure 2, is $\theta_d \perp \theta_{d+1} \mid Z_{d,1:N}, Z_{d+1,1:N}$?

No.

1.6 In the graphical model of Figure 2, is $\Lambda_d \perp \Lambda_{d+1} \mid \Phi, Z_{1:D,1:n}$?

No.

2 Likelihood of a Tree Graphical Model

2.1 Implement a dynamic programming algorithm that for a given T, Θ, β computes $p(\beta \mid T, \Theta)$

If we say that we have a 3 layered binary tree T made up of 7 nodes, A, B, C, D, E, F, G , where A is the parent of B and B is the parent of D, E , all of the probability distributions that can be represented by T are given by

$$P(A, B, C, D, E, F, G) = P(A)P(B \mid A)P(C \mid A)P(D \mid B)P(E \mid B)P(F \mid C)P(G \mid C)$$

In order to calculate the probability of an assignment to our leaf nodes, we want to obtain a joint probability density of the leaf nodes, and in order to do that we need to marginalize out all non-leaf nodes

$$P(D, E, F, G) = \sum_A \sum_B \sum_C P(A, B, C, D, E, F, G)$$

We can see that there is only one term containing each of the leaf nodes, and that all of the leaf nodes are in disjoint probability statements. Thus we know that our joint probability density containing only the leaf nodes will be a product of independent factors, one for each leaf node. Thus we can independently calculate our leaf probabilities as follows

$$P(D) = \sum_B P(D, B)$$

and

$$P(D \mid B) = \frac{P(D, B)}{P(B)}$$

We can conclude that

$$P(D) = \sum_B P(D \mid B)P(B)$$

Similarly we can look at the terms containing B to arrive at

$$P(B) = \sum_A P(B \mid A)P(A)$$

However, doing this recursive type of calculation for all of the leaf nodes will lead to us doing the same work multiple times, for instance in the given example we will calculate $P(B)$ when attempting to arrive at both $P(D)$ and $P(C)$.