Tutorial 7 - Markov Networks

COMP9418 – Advanced Topics in Statistical Machine Learning

Lecturer: Gustavo Batista

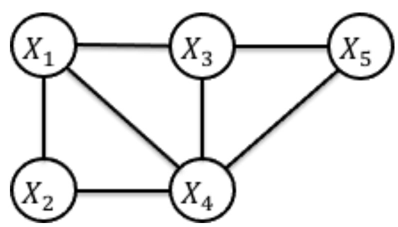
Lecture: Markov Networks

Topic: Questions from lecture topics

Last revision: Sunday 25^{th} October, 2020 at 21:51

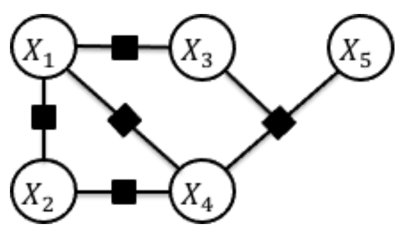
Question 1

Consider the following Markov network in the form of an undirected graph.



a. Determine the cliques in the graph.
b. Express the joint probability as a product of clique potentials.
c. Assuming all the variables are binary, how many parameters are there in this model?

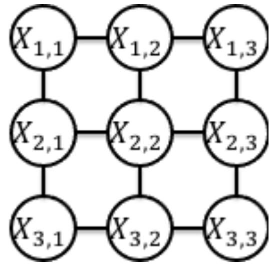
Now, consider the same Markov network expressed as a factor graph.



Answer the same questions (a,b and c) for this model.

Question 2

Consider the following pairwise Markov network in the form of a grid.



a. Express the pairwise Markov independencies of this graph for missing edges between the nodes $X_{1,1}$, $X_{1,2}$ and $X_{2,2}$.

b. List the nodes that compose the Markov blanket for each of nodes $X_{1,1}$, $X_{1,2}$ and $X_{2,2}$

c. Is $X_{1,1}$ separated from $X_{2,2}$ given $X_{1,2}$? If not, which minimal set of nodes is necessary to observe to guarantee the separation between these nodes?

Question 3

In an image smoothing application, the image is represented by a regular Markov Random Field with a neighbourhood defined as the grid of the previous question.

In this application, we have two kinds of potentials: $\phi_1(X_{i,j}, X_{i',j'})$ represents the correlation between neighboring pixels and $\phi_2(X_{i,j}, Y_{i,j})$ represents the correlation between the filtered image pixel ($X_{i,j}$) and the respective noisy image pixel ($Y_{i,j}$). A simple form of assigning values to these potentials, is using the absolute difference between pixel values, i.e, $\phi(a, b) = |a - b|$.

Also, we define the energy function for an image $\mathbf{X}$ as

$$E(\mathbf{X}) = \sum_{i,j}(\phi_1(X_{i,j}, X_{i+1,j}) + \phi_1(X_{i,j}, X_{i,j+1})) + \lambda \sum_{i,j} \phi_2(X_{i,j}, Y_{i,j})$$

with $\lambda = 4$, giving more weight to the observations. (For simplicity, this equation does not deal with the edge cases on the bottom and right edge of the image. Consider how these should be dealt with).

Suppose we want to smooth a small image represented by a 4×4 grid, where each site can take one of two values, 0 and 1. Given the initial configuration F

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

and the observation G,

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

a. Obtain the MAP configuration using the Iterated Conditional Modes (ICM) variant of the stochastic search algorithm. This involves considering each location one by one and checking whether changing that pixel will decrease the total energy.

b. Would you expect the results obtained by the other variants of the stochastic search algorithm to be different? Why?

Question 4

What is the time complexity of the stochastic search algorithm for its different variants?