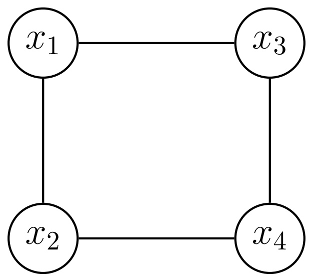
University of Illinois Spring 2018

CS 446: Machine Learning
Homework 7

Due on Tuesday, March 6, 2018, 11:59 a.m. Central Time

1. [**4 points**] Inference Methods for Discrete Markov Random Fields

   For this problem, consider the following Markov Random Field, where each node can be assigned one of the values in $\{1, 2, 3, 4, 5\}$:



   (a) To conduct MAP inference on this graph using exhaustive search, how many configurations must be checked?

Your answer: 5^4 configurations must be checked.

   (b) Can MAP inference be run on this graph using a dynamic programming algorithm? Why or why not?

Your answer: No we cannot. Because in this case the graph is a 'loopy graph' and dynamic programming approach can only be used in 'tree' structure. So we cannot use a dynamic programming algorithm.

   (c) To run MAP inference on this graph using loopy belief propagation, how many messages must be computed?

Your answer: If we perform loopy belief propagation on this graph, 40 ($8 \times 5 = 40$) messages have to be computed.

2. [**7 points**] ILP Inference formulation in Discrete Markov Random Fields

   (a) Suppose we have two variables $x_1 \in \{0, 1\}$ and $x_2 \in \{0, 1\}$ and their local evidence functions $\theta_1(x_1)$ and $\theta_2(x_2)$ as well as a pairwise function $\theta_{1,2}(x_1, x_2)$. Using this setup, inference solves $\arg\max_{x_1,x_2} \theta_1(x_1) + \theta_2(x_2) + \theta_{1,2}(x_1, x_2)$. Using

$$\theta_1(x_1) = \begin{cases} 1 & \text{if } x_1 = 0 \\ 2 & \text{otherwise} \end{cases} \qquad \theta_2(x_2) = \begin{cases} 1 & \text{if } x_2 = 0 \\ 2 & \text{otherwise} \end{cases}$$

$$\theta_{1,2}(x_1, x_2) = \begin{cases} -1 & \text{otherwise} \\ 2 & \text{if } x_1 = 0 \,\&\, x_2 = 1 \end{cases}$$

what is the integer linear programming formulation of the inference task? Make the cost function and constraints explicit for the given problem, i.e., do not use a general formulation.

Your answer: In this case we only care about whether $x_i = 0$ or $x_i \neq 0$, where i=1 or i=2 so I use notation $\overline{0}$ to indicate the condition that the value of x is not 0. Then we can get:

$$\max_{b_1,b_2,b_{12}} \begin{bmatrix} b_1(0) \\ b_1(1) \\ b_2(0) \\ b_2(1) \\ b_{1,1}(0,0) \\ b_{1,2}(1,0) \\ b_{1,2}(0,1) \\ b_{1,2}(1,1) \end{bmatrix}^\intercal \begin{bmatrix} \theta_1(0) \\ \theta_1(1) \\ \theta_2(0) \\ \theta_2(1) \\ \theta_{1,1}(0,0) \\ \theta_{1,2}(1,0) \\ \theta_{1,2}(0,1) \\ \theta_{1,2}(1,1) \end{bmatrix} = \begin{bmatrix} b_1(0) \\ b_1(1) \\ b_2(0) \\ b_2(1) \\ b_{1,1}(0,0) \\ b_{1,2}(1,0) \\ b_{1,2}(0,1) \\ b_{1,2}(1,1) \end{bmatrix}^\intercal \begin{bmatrix} 1 \\ 2 \\ 1 \\ 2 \\ -1 \\ -1 \\ 2 \\ -1 \end{bmatrix}$$

subject to:

$$b_r(x_r) \in \{0,1\} \quad \forall r, x_r;$$

$$b_r(x_r) \geq 0 \quad \forall r, x_r$$

$$\sum_r b_r(x_r) = 1 \quad \forall r$$

$$\sum_{x_p \setminus x_r} b_p(x_p) = b_r(x_r)$$

(b) What is the solution (value and argument) to the program in part (a).

Your answer: $x_1 = 0, x_2 = 1$, then the maximum of the function is $5 = 1 + 2 + 2$.

(c) Why do we typically not use the integer linear program for reasonably sized MRFs?

Your answer: Integer Linear Programming is very slow for larger problems.