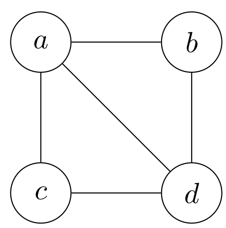
Problem Set 3

CS 6347

Due: 4/8/2024 by 11:59pm

Note: all answers should be accompanied by explanations for full credit. Late homeworks cannot be accepted. All submitted code **MUST** compile/run.

Problem 1: Gibbs Sampling for Weighted Coloring Problems (50 pts)



1. Use the Gibbs sampling algorithm to approximate the marginals of the probability distribution over weighted vertex colorings from problem 2 of problem set 2. Your solution should be written as a MATLAB function that takes as input an $n \times n$ matrix A corresponding to the adjacency matrix of a graph G, a vector of weights $w \in \mathbb{R}^k$, `burnin` which is the number of burn-in samples, and `its` which is an input that controls the total number of iterations of the Gibbs sampler after burn-in. The output should be an $n \times k$ matrix of marginals whose i^{th}, x_i^{th} entry is equal to the probability that $i \in V$ is colored with color $x_i \in \{1, \ldots, k\}$.

   ```
   function m = gibbs(A, w, burnin, its)
   ```

2. Run the Gibbs sampler to estimate the probability that $a \in V$ is colored with color 4 using weights $w_i = i$ for each $i \in \{1, 2, 3, 4\}$ in the graph above. Construct a table of the estimated marginal as a function of `burnin` versus `its` for your implementation where `burnin` and `its` are chosen from the set $\{2^6, 2^{10}, 2^{14}, 2^{18}\}$. Does your answer depend on the initial choice of assignment used in your Gibbs sampling algorithm?

Problem 2: Maximum Likelihood for Colorings (50 pts)

For this problem, we will use the same factorization as we have in past assignments. However, the weights will now be considered parameters of the model that need to be learned from samples.

1. Use the belief propagation algorithm that you wrote on Problem Set 2 to perform (approximate) maximum likelihood estimation for the coloring counting problem. Your solution should

be written as a MATLAB function that takes as input an $n \times n$ matrix A corresponding to the adjacency matrix of the graph G and `samples` which is an $n \times m$ k-ary matrix where $samples_{i,t}$ corresponds to observed color for vertex i in the t^{th} sample. The output should be a k-dimensional vector of weights w corresponding to the MLE parameters for the model.

```
function w = colormle(A, samples)
```

2. Validate your algorithm by showing that you obtain the correct maximum-likelihood parameters when the graph is a 3-node chain with 3 colors and weights $w_i = i$ using your Gibbs sampler that you wrote on Problem 1. Construct a table of the estimated partition function (from the MLE) after $10^1, 10^2, 10^3, 10^4$, and 10^5 samples.