CS 189/289A Introduction to Machine Learning
Fall 2021 Jennifer Listgarten and Jitendra Malik **DIS12**

1 Probabilistic Graphical Models

Recall that we can represent joint probability distributions with directed acyclic graphs (DAGs). Let G be a DAG with vertices $X_1, ..., X_k$. If P is a (joint) distribution for $X_1, ..., X_k$ with (joint) probability mass function p, we say that G represents P if

$$p(x_1, \cdots, x_k) = \prod_{i=1}^{k} P(X_i = x_i | \text{pa}(X_i)), \tag{1}$$

where $\text{pa}(X_i)$ denotes the parent nodes of X_i. (Recall that in a DAG, node Z is a parent of node X iff there is a directed edge going out of Z into X.)

Consider the following DAG

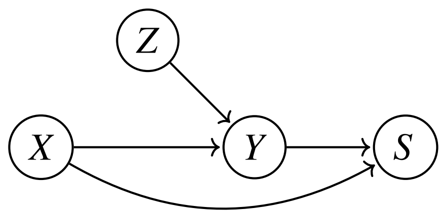

Figure 1: G, a DAG

(a) Write down the joint factorization of $P_{S,X,Y,Z}(s, x, y, z)$ implied by the DAG G shown in Figure 1.

(b) Is $S \perp Z \mid Y$?

(c) Is $S \perp X \mid Y$?

2 Hidden Markov Models: Math Review

A Hidden Markov Model is a Markov Process with unobserved (hidden) states.

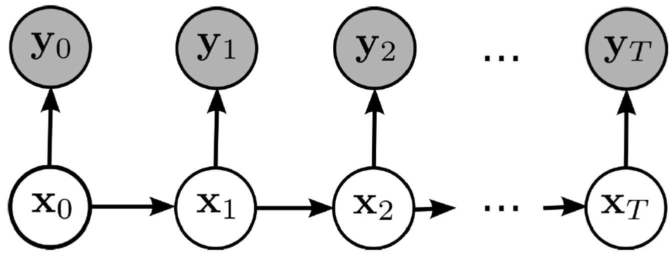

Figure 2: Example Hidden Markov Chain

Consider the following system in $\mathbb{R}^2$, where X_n is the true state at any given time n and Y_n is our observation. Given an initial state X_0, we move to future states by recursively multiplying our current state with transformation matrix A and adding i.i.d. Standard Normal Gaussian noise. When we take an observation Y_n of the true state X_n, we are also exposed to i.i.d. Standard Normal Gaussian Noise.

$$X_{n+1} = AX_n + N(0, I) \tag{2}$$
$$Y_n = X_n + N(0, I) \tag{3}$$

Where we have the $2x2$ transformation matrix A defined as follows:

$$A = \begin{bmatrix} .5 & -.25 \\ -.25 & .75 \end{bmatrix} \tag{4}$$

If we restrict the initial state X_0 to be a unit vector ($||X_0||_2 = 1$), determine the following

(a) What are the eigenvalues of A? Is A a positive semi-definite matrix? (Note that $\sqrt{5} = 2.236$)

(b) What is the $\|E[Y_\infty]\|_2$? Prove your assertion.

(c) Consider the Frobenius Norm of an arbitrary M x N matrix Q, defined as $||Q||_F = \sqrt{\sum_i \sum_j |Q_{i,j}|^2}$, which indicates the “magnitude” or “largeness” of a matrix. Is $\|Var[Y_\infty]\|_F$ finite or infinite? Prove your assertion.

You may find the following facts to be useful:

(i) Triangle Inequality: $||X + Y|| \leq ||X||+||Y||$

(ii) Cauchy Schwarz: $||XY|| \leq ||X||||Y||$

(iii) Geometric Sum: $\sum_{i=0}^{\infty} ar^i = \frac{a}{1-r} \qquad \forall r \text{ s.t. } 0 < r < 1; a, r \in \mathbb{R}$