CSC412/2506 Winter 2020: Probabilistic Learning and Reasoning

Tianyu Du

February 8, 2020

Contents

1   Introduction

2   Probabilistic Models

3   Directed Graphical Models

4   Exact Inference

Notation 4.1. Let X denote the set of all random variables in the model, and

1. X_E = The observed evidence;

2. X_F = The unobserved variable we want to infer;

3. $X_R = X - \{X_F, X_E\}$ = Remaining variables, extraneous to query.

The model defines the joint distribution of all random variables:

$$p(X_E, X_F, X_R) \tag{4.1}$$

Definition 4.1. The joint distribution over evidence and subject of inference is

$$p(X_F, X_E) = \sum_{X_R} p(X_F, X_E, X_R) \tag{4.2}$$

Definition 4.2. The conditional probability distribution for inference given evidence is
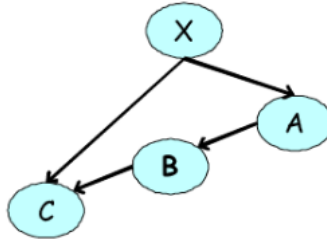
$$p(X_F | X_E) = \frac{p(X_F, X_E)}{p(X_E)} = \frac{p(X_F, X_E)}{\sum_{X_F} p(X_F, X_E)} \tag{4.3}$$

Definition 4.3. The distribution of evidence can be computed as

$$p(X_E) = \sum_{X_F, X_R} p(X_F, X_E, X_R) \tag{4.4}$$

4.1   Variable Elimination

4.2   Intermediate Factors

$$p(A, B, C) = \sum_X p(X)p(A|X)p(B|A)p(C|B, X) \tag{4.5}$$

$$= p(B|A) \underbrace{\sum_X p(X)p(A|X)p(C|B, X)}_{\text{unnormalized}} \tag{4.6}$$

Definition 4.4. A **factor** ϕ describes the local relation between random variables, meanwhile, $\int d\phi$ is <u>not</u> necessarily one.

Remark 4.1. Let $X_\ell \subseteq X$ be a group of local random variables, then $p(X_\ell)$ is automatically a factor $\phi(X_\ell)$.

$$p(A, B, C) = \sum_X \underbrace{p(X)p(A|X)p(B|A)p(C|B, X)}_{\text{from graphical representation}} \tag{4.7}$$

$$= \sum_X \underbrace{\phi(X)\phi(A, X)\phi(A, B)\phi(X, B, C)}_{\text{factor representation}} \tag{4.8}$$

$$= \phi(A, B) \sum_X \phi(X)\phi(A, X)\phi(X, B, C) \tag{4.9}$$

$$= \phi(A, B) \underbrace{\tau(A, B, C)}_{\text{another factor}} \tag{4.10}$$

4.3　Sum-Product Inference

Theorem 4.1. Consider a graphical model with random variables $X = Y \cup Z$. For an random variable Y in a <u>directed</u> or <u>undirected</u> model, $P(Y)$ can be computed using the **sum-product**

$$\tau(Y) = \sum_z \prod_{\phi \in \Phi} \phi(Scope[\phi] \cap Z, Scope[\phi] \cap Y) \tag{4.11}$$

where Φ is a set of factors.

Remark 4.2. For <u>directed models</u>,

$$\Phi = \{\phi_{x_i}\}_{i=1}^N = \{p(x_i| \text{ parents } (x_i))\}_{i=1}^N \tag{4.12}$$

4.4　Complexity of Variable Elimination Ordering

Theorem 4.2. The complexity of the variable elimination algorithm is

$$\mathcal{O}(mk^{N_{max}}) \tag{4.13}$$

where

　(i)　m is the number of initial factors $|\Phi|$;

　(ii)　k is the number of states each random variable takes, assumed to be equal;

　(iii)　N_i is the number of random variables within each summation;

　(iv)　$N_{max} = \max_i N_i$.

5   Message passing, Hidden Markov Models, and Sampling

6   Sampling and Monte Carlo Methods