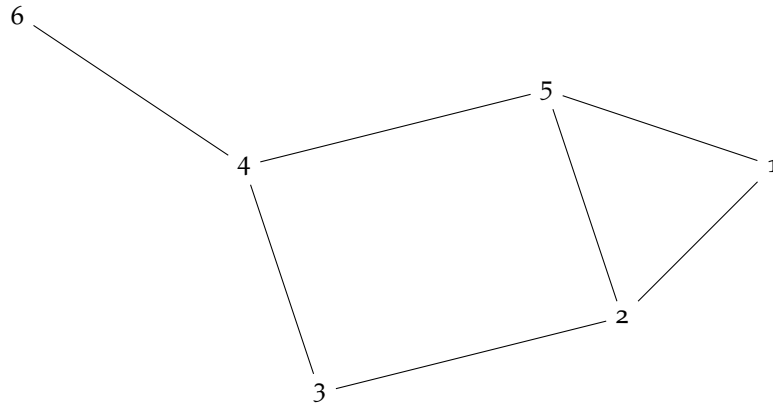


# Probabilistic Graphical Models

Atif Khan

October 15, 2012

## Bayesian Networks



## Tree Structure CPDs

Table CPDs are problematic when there are many many parents for a conditional variable. Tree structure CPDs can handle a large set of parents given that there is certain context provided.

We start from the root and traverse the branches. Each branch represents a CPD for a given set of conditions. For example if a student does not apply for a job ( $A = a^0$ ), then the probability of the student getting a job ( $J = j^1$ ) is

- $P(j^1|A, S, L) = 0.2$
- the probability is independent of  $S, L$

## Independence of Causal Influence

Tree CPS (§- Tree Structure CPDs) are good for adding context with many parents. But when the number of the parents is quite large and most of the parents contribute (see Figure 3), then the Tree CPD is not a good representation. For example, *Cough* can be caused by an array of diseases. We utilize a noisy OR CPD for this purpose.

## Noisy OR CPD

For each parent variable  $X$ , we introduce an intermediate variable  $Z$  (filter).  $Z$  represents the event of a parent  $X$  being *true*, causing  $Y$  to

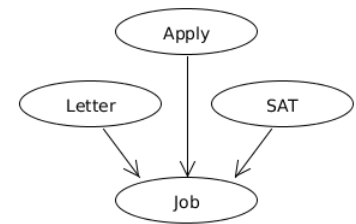


Figure 1: A simple CPD with four parents

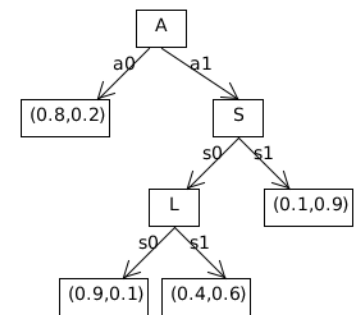


Figure 2: Tree based CPD for Figure 1

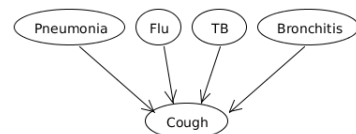
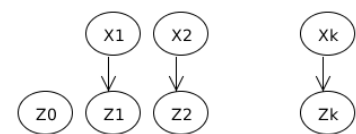


Figure 3: Multiple parents contributing towards a single variable. This does not lead it self into a Tree CPD.



be true by itself. Ultimately  $Y$  is true if any  $Z$  succeeded in making it true. Therefore,  $Y$  is a deterministic OR based of its parents  $Z$ .

$$P(Z_0 = 1) = \lambda_0 \quad \text{Leak}$$

$$P(Z_i = 1|X_i) = \begin{cases} 0 & X_i = 0 \\ \lambda_i & X_i = 1 \end{cases} \quad \text{Penetrance}$$

Now consider, what is the probability that  $Y = 0$  given all the parents  $X$ :

$$P(Y = 0|X_1, X_2, \dots, X_k) = (1 - \lambda_0) \prod_{i: X_i=1} (1 - \lambda_i)$$

where,  $\prod_{i: X_i=1} (1 - \lambda_i)$  represents the parents that are on.

For the probability that  $Y = 1$ , we have

$$P(Y = 1|X_1, X_2, \dots, X_k) = 1 - P(Y = 0|X_1, X_2, \dots, X_k)$$

**GENERALIZATION OF THE NOISY OR CPD:** Figure 5 represents the generalization of the noisy OR CPD. The variable  $Z$  is a deterministic variable that can represent different functions such as an *AND* operation, *MAX* operation etc.

### Sigmoid CPD

Given  $Z = w_0 + \sum_{i=1}^k w_i X_i$ , where  $Z_i = w_i X_i$ , a sigmoid CPD is

$$P(y^1|X_1, X_2, \dots, X_k) = \text{sigmoid}(Z)$$

Where  $\text{sigmoid}(z) = \frac{e^z}{1+e^z}$ ,  $z$  is a continuous variable. The result of the *sigmoid* function is to reduce the value of  $z$  to  $[0, 1]$ .

### Continuous Variables

Imagine that the temperature is a *continuous variable* and the sensor provides an approximation of the temperature. That is *sensor S* is a normal distribution defined using *linear Gaussian* as:

$$S \sim \mathcal{N}(T; \sigma_s^2)$$

No imagine that the temperature soon  $Temperature'$  depends on current temperature, outside temperature and the conditionally on the door being opened or closed (as shown in Figure 7). We have the following *conditional linear Gaussian* distributions:

$$T' \sim \mathcal{N}(\alpha_0 T + (1 - \alpha_0)O; \sigma_{0T}^2) \quad \text{when } D^0$$

$$T' \sim \mathcal{N}(\alpha_1 T + (1 - \alpha_1)O; \sigma_{1T}^2) \quad \text{when } D^1$$

*Penetrance* defines how good is  $X_i$  in turning  $Z_i$ , where as *Leak* defines  $Y$  turning on by itself.

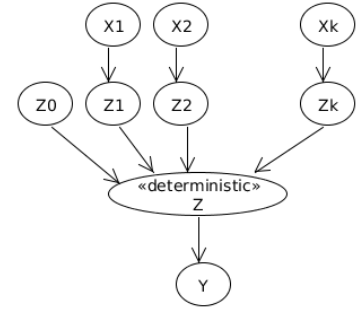


Figure 5: Generalization of the noisy OR CPD

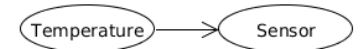


Figure 6: Example of continuous variables

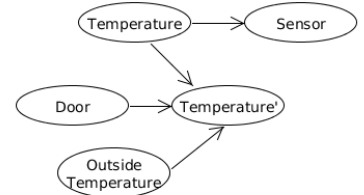


Figure 7: Example of continuous variables with condition

### Linear Gaussian

For the given graph in Figure 8, we have a variable  $Y$  with parents  $X$ , then we have a *linear Gaussian* defined as follows:

$$Y \sim \mathcal{N}(w_0 + \sum w_i X_i; \sigma^2)$$

where, the mean of the Gaussian distribution ( $w_0 + \sum w_i X_i$ ) is a linear function (of the parents  $X_i$ , and the variance  $\sigma^2$  does not depend on the parents.

### Conditional Linear Gaussian

We can now define a *conditional linear Gaussian* (see Figure 9) with a discrete parent variable  $A$  as follows:

$$Y \sim \mathcal{N}(w_{a0} + \sum w_i X_i; \sigma_a^2)$$

Note that the variance  $\sigma_a^2$  depends on the discrete parent  $A$  but not  $X$ .

### References

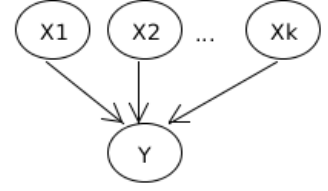


Figure 8: Model for linear Gaussian.

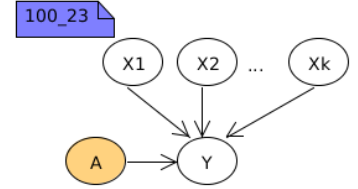


Figure 9: Model for conditional linear Gaussian. Variable  $A$  is a discrete parent. There can be more than one discrete parents.