Gibbs Sampling and Bayesian Networks

1 Bayesian Network Example

Consider a simple Bayesian network with nodes for Cloudy, Sprinkler, Rain, and WetGrass. The topological ordering of the nodes is as follows:

- Cloudy
- Sprinkler
- Rain
- WetGrass

The probability distributions are given as follows:

- $P(\text{Cloudy}) = 0.5$
- $P(\text{Sprinkler} \mid \text{Cloudy}) = \begin{pmatrix} 0.1 & 0.9 \\ 0.5 & 0.5 \end{pmatrix}$
- $P(\text{Rain} \mid \text{Cloudy}) = \begin{pmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{pmatrix}$
- $P(\text{WetGrass} \mid \text{Sprinkler}, \text{Rain}) = \begin{pmatrix} 0.99 & 0.9 \\ 0.9 & 0.0 \end{pmatrix}$

The query we are interested in is:

$$P(\text{Rain} \mid \text{Sprinkler} = \text{true}, \text{WetGrass} = \text{true})$$

1.1 Sampling Process

We sample from the network using the following procedure:

1. Sample from $P(\text{Cloudy}) = 0.5$, the value is true.
2. Sample from $P(\text{Sprinkler} \mid \text{Cloudy} = \text{true}) = \begin{pmatrix} 0.1 & 0.9 \end{pmatrix}$, the value is false.
3. Sample from $P(\text{Rain} \mid \text{Cloudy} = \text{true}) = \begin{pmatrix} 0.8 & 0.2 \end{pmatrix}$, the value is true.
4. Sample from $P(\text{WetGrass} \mid \text{Sprinkler} = \text{false}, \text{Rain} = \text{true}) = \begin{pmatrix} 0.9 & 0.1 \end{pmatrix}$, the value is true.

The final sampled event is $(\text{Cloudy} = \text{true}, \text{Sprinkler} = \text{false}, \text{Rain} = \text{true}, \text{WetGrass} = \text{true})$.

2 Markov Blanket and Gibbs Sampling

In Gibbs sampling, the non-evidence variables are sampled repeatedly in random order according to their conditional distribution, given the values of the other variables (the Markov blanket).

The Markov blanket for a node consists of its parents, its children, and the parents of its children.

For example, in our Bayesian network, the Markov blanket of Rain includes:

- Parents: Cloudy
- Children: WetGrass
- Parents of Children: Sprinkler

2.1 Gibbs Sampling Procedure

In Gibbs sampling, we:

1. Fix the evidence variables (e.g., Sprinkler = true, WetGrass = true).
2. Randomly select a non-evidence variable (e.g., Cloudy) and sample its value from the conditional distribution given the current values of its Markov blanket.
3. Repeat for the other non-evidence variables (e.g., Rain) and update their values based on their Markov blanket.

This process continues iteratively until the samples converge to the true posterior distribution.

3 Calculating Markov Blanket Distribution

Let us calculate the Markov blanket distribution for Cloudy given Sprinkler = true and Rain = false. The conditional probability is given by:

$$P(C \mid S = \text{true}, R = \text{false}) \propto P(C) \cdot P(S \mid C) \cdot P(R \mid C)$$

Substituting the values:

$$P(C \mid S = \text{true}, R = \text{false}) \propto 0.5 \cdot 0.1 \cdot 0.2$$

Thus, the normalized probabilities would be approximately $< 0.048, 0.952 >$.

4 Transition Kernel for Gibbs Sampling

The transition kernel $k(x \to x')$ defines the probability of moving from state x to state x'. Gibbs sampling ensures that only one variable changes at a time.

1. If the states x and x' differ in more than one variable, $k(x \to x') = 0$.

2. If the states differ in exactly one variable, say X_i, then the transition probability is given by:

$$k(x \to x') = P(X_i \mid \text{Markov blanket of } X_i)$$

4.1 Stationary Distribution

The stationary distribution $\pi(x)$ of the chain is the posterior probability of state x, given the evidence:

$$\pi(x) = P(x \mid e)$$

The ergodicity condition ensures that every state can be reached from any other state, and there are no cycles.