Lecture 6

Note Title 9/29/2005

Likelihood Weighting

- Sampe values for the hidden variables
- Reweight each sample according to the probabilities of the evidence

W.R.T Previous example

$Pr(C \mid D = A)$

1° Sample a value for A from $P(A)$
2° " " " " B from $P(B \mid A)$
3° " " " " C " $P(C \mid A)$
4° Weight sample by $Pr(D \mid B, C)$

$Pr(C \mid D=t, B=f)$?

1° Sample a value for A from $P(A)$
2° Set $W = Pr(B=f \mid A)$
3° Sample a value for C from $P(C \mid A)$
4° $W \leftarrow W * Pr(D=t \mid B=f, C)$

Samples :

$$\left.\begin{array}{ll} A=t, & C=f \quad 0.2 \\ A=f, & C=f \quad 0.3 \end{array}\right\} 0.5$$

$$\left.\begin{array}{ll} A=t, & C=t \quad 0.1 \end{array}\right\} 0.1$$

$$Pr(\underset{C=t}{C=f} \mid D=t, B=f) = 5/6$$
$$= 1/6$$

$$A \rightarrow B$$

$Pr(A=t) = 0.9$

$Pr(B \mid A)$

B \ A	T	F
T	1	0
F	0	1

$Pr(A \mid B=f)$?

Scenario: 9 samples with $A=t \Rightarrow B=t$
1 sample " $A=f \Rightarrow B=f$

To achieve good estimates, likelihood weghting often needs an exponential number of samples just like rejection sampling.

Importance Sampling

Let's say that we are trying to estimate $P(X)$. If we sample directly from $P(X)$ then all samples have a weight of 1. If instead we sample from another distribution, say $Q(X)$ then we need to weigth the samples: $W(S) = \frac{P(S)}{Q(S)}$

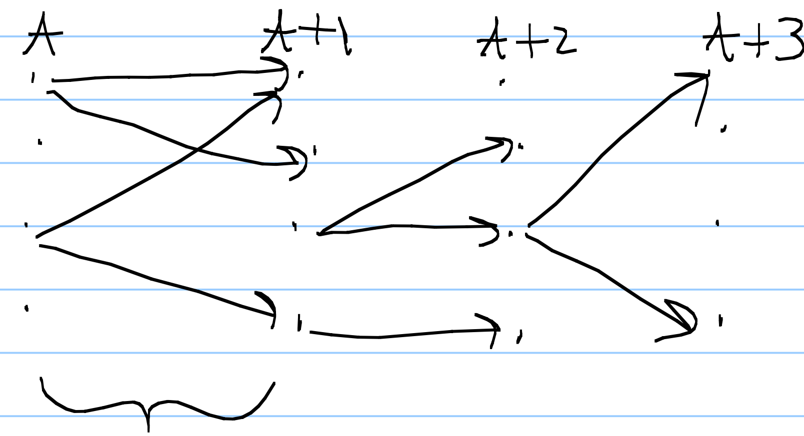
Self-importance sampling

Suppose you have already sampled N instances which give you an estimate $P'(x)$ for the true dist $P(x)$. Then sample from $P'(x)$ since it should be close to $P(x)$.

Repeat
- Sample from $P_i'(x)$
- Combine $P_i'(x)$ with new sample into a new dist $P_{i+1}'(x)$

Markov Chain Monte Carlo

Markov Chain:
- set of states
- Prob dist of going from state s to state s': $Pr(s'|s)$



$Pr(s'|s)$

If Markov chain is ergotic and aperiodic then there is a stationary distribution $\pi(S)$ that the process converges to.

Let $\pi(S)$ be a vector
and $P(S'|S)$ be a matrix

Then $\pi P = \pi$

Markov Chain Monte Carlo Sampling

pick a state S_0 to start
repeat n times
    sample S_{i+1} from $P(S_{i+1}|S_i)$
Frequency counts of sampled states are proportional to π.

MCMC for Bayes nets

Let $S_0 = \{A=t, B=f, C=t, D=t\}$

To sample S_1 sample from each conditional distribution using as evidence the values of the variables from the previous state.

Ex:
- 1° Sample a value from $P(A)$
- 2° " " " " $P(B \mid A=t)$
- 3° " " " " $P(C \mid A=t)$
- 4° " " " " $P(D \mid B=f, C=t)$

Gives S_1