Probabilistic graphical models

Siddharth Bhat

Spring 2020

Contents

Chapter 1

Background and aims

Consider a distribution of binary random variables $x_1, x_2, \ldots x_n, y$. Note that to define the value of $P(x_1)$, we need just one value: $P(x_1 = 0)$. We can derive $P(x_1 = 1) \equiv 1 - P(x_1 = 0)$.

However, the full joint distribution $P(x_1, x_2, \ldots, x_n, y)$ needs $2^{n+1} - 1$ values to fully define.

However, let us assume that $P(x_i|y)$ are all independent. Hence, we can rewrite the above distribution as $P(y) \prod_{i=1}^{n} P(x_i|y)$. Now, we need to know $P(x_i|y = 0), P(x_i|y = 1)$. Both of these are binary random variables which need one value to define. So in toto, we need $2n + 1$ values for the above (factored) joint distribution.

So, we will study how to represent, perform inference, and perfom bayesian updates (learning). Also, connections to boltzmann distributions and whatnot will be explored. Connections to graph theory as well. We are also going to study MCMC (Markov chain monte carlo) methods. I hope we study more than just metropolis hastings: I want to understand Hamiltonian and Lavengin Monte Carlo more deeply (NUTS sampling, slice sampling, their interactions with HMC, etc). Later, we will see some connections to Learning theory (PAC learning - defined by Valiant).

The textbook is "Kohler and Friedman".

1.0.1 A teaser problem

We start with an ordered deck. We propose a shuffling mechanism: take the top card and move it to somewhere in the deck. Eg. If we start form $(1, 2, 3)$, we can move this to $(2, 1, 3)$, or $(2, 3, 1)$. Now, when the card 3 comes to the top, note that we had placed all other numbers in the deck with uniform probability. So, when the card 3 comes to the top, all the other cards are uniformly distributed. We now need to place 3 uniformly in the deck.

Let T_1 be the random variable of the first round at which a single card is placed *underneath* n.

There are $n - 1$ slots where can place any top card, so the likelihood of hitting the bottom slot is $1/(n - 1)$.

$$
\begin{aligned}
P(T_1 = 1) &\equiv \frac{1}{n-1} \\
P(T_1 = 2) &\equiv \left(1 - \frac{1}{n-1}\right)\frac{1}{n-1} = \frac{n-2}{n-1} \\
P(T_1 = i) &\equiv (1 - P(T_1 = i-1))\frac{1}{n-1} = \left(1 - \frac{1}{n-1}\right)^{i-1}\frac{1}{n-1} = \frac{(n-2)^{i-1}}{(n-1)^i}
\end{aligned}
$$

This is a geometric distribution with parameter $\frac{1}{n-1}$. The expectation is going to be $\mathbb{E}[T_1] \equiv n-1$.

We now define T_2 to be the random variable which is the time from when the first card went underneath the nth card, to when the second card went underneath the nth card. We have two locations at the bottom. Eg. if we had $(1, 2, 3, 4)$ to start with, and after T_1, we are now at $(2, 3, 4, 1)$. We now have two positions $(2, 3, 4, \circ, 1, \circ)$ to be underneath the card 4.

$$
\begin{aligned}
P(T_2 = 1) &\equiv \frac{2}{n-1} \\
P(T_2 = i) &\equiv \left(1 - \frac{2}{n-1}\right)^{i-1}\frac{2}{n-1}
\end{aligned}
$$

This is a geometric distribution with parameter $\frac{2}{n-1}$. The expectation is going to be $\mathbb{E}[T_1] \equiv n-2$.

The total time for the nth card to reach the top is going to be $T \equiv T_1 + T_2 + \cdots + T_n$. So the expectation is going to be $\mathbb{E}[T] = \sum_i \mathbb{E}[T_i] = \sum_i \frac{1}{n-i}$