Lab 2 Clarifications

Rahul G. Krishnan

September 19, 2016

There were a few good questions by students whose clarifications might be beneficial to you all.

Bayes Ball

You should all be able to look at the graph in Figure 1 and say with confidence that X and Y are *not* conditionally independent given Z. The question we will concern ourself with is whether $X \perp\!\!\!\perp Y|K$.

A Bayes ball on the path $X - Z - Y$ will surely be blocked at Z since Z is not observed. However, consider the path, $X - Z - K - Z - Y$. Here the ball from X can travel to K at which point it bounces back and travels along $K - Z - Y$.

Formally, if we condition on a set of random variables C and consider a path that traverses a v-structure where Z is the common child, then X is d-separated from Y along that path only if Z and all of its descendants are *not* in the set C.

Intuition: In the generative story of our Bayesian network, the random variable Z encompasses some of the information about its children (since the child random variables are generated conditioned on the parents' value). Therefore the observation of any child of Z gives us some information about what value Z could have taken which in turn is responsible for allowing inference to flow from X to Y.

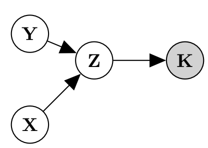

Figure 1: Bayesian Network

See the course readings for a concrete example of this. (Search for "guard" in the document)

Equivalence in Bayesian networks

There were a few questions about the equivalence of two Bayesian networks. The formal answer is that two Bayesian networks are mathematically equivalent if the set of independence statements they encode are identical.

In Figure 2, the two Bayesian networks that you see are equivalent since they encode the same set of independence statements (namely $X \perp\!\!\!\perp Y|Z$).

(a) Common parent (b) Chain

Figure 2: Equivalent Bayesian Network

You'll notice they do not however, have the same generative process. Figure 2a implies the joint distribution factorizes as $P(X, Y, Z) = P(Z)P(X|Z)P(Y|Z)$ while Figure 2b implies $P(X, Y, Z) = P(X)P(Z|X)P(Y|Z)$. How do we reconcile this? On the one hand the two networks are mathematically equivalent yet they represent the joint distribution over random variables in two distinct ways.

The factorization of the joint probability distribution is a function of the structure of your graphical model (evident from above) that posits a generative process for your data but equivalence is agnostic to this data generation process. When you ask whether two Bayesian networks are equivalent, you are asking a question of "representation", i.e whether or not your data can be represented equally well by these two Bayesian networks. Said differently, if you were given data comprising three random variables (called X, Y and Z of course) and an oracle told you that when you know the value of Z, X and Y behave independent of one another. Then either of the Bayesian networks above can *represent* your distribution. ie. there exists a setting of the conditional probability tables of both Bayesian networks that can explain your data.

However, there might still be reasons to prefer one structure over the other. For example, one Bayesian network's parameters might be easier to learn, or you might be interested in finding a good fit for the causal nature of your problem.

Finally, as an aside, there is a theorem as follows that might prove handy in checking whether two Bayesian networks are equivalent:

Theorem: If G, G' have the same skeleton and v-structures, then $I(G) = I(G')$.