1 HG48600/STAT34550 Lectures

1.1 Lecture: Graphical models: Introduction

Graphical models introduction

- A tool for visualizing the structure of a probabilistic model
- Provides insights, such as conditional independence properties, by inspection of the graph
- Complex computations can be expressed as operations on the graph
- Graphical models - also sometimes called Bayesian networks

Advantages of a graphical model:

- The **Chain Rule**
  - In its basic form:
  $$P(A, B) = P(B|A)P(A)$$
  - Which generalizes as:
  $$P(A_1, A_2, \dots, A_k) = P(A_1)P(A_2|A_1)\dots P(A_k|A_{k-1})$$
  - This result holds regardless of the ordering.
- A graph makes clear a simpler factorization than the chain rule:
$$P(\mathbf{X}) = \prod_{i=1}^{n} P(X_i|Pa_i)$$

Review abstractly: Conditional independence and three 3-node graphs

- Three example 3-node graphs (and how they behave when condtioning on X_2)
  - The linear chain graph
  - The "multiple offspring" graph
  - The v-structure graph
    - Definition 5: A v-structure (3) is an induced subgraph (a subset of nodes and all the edges between these nodes in G) of the form $X \rightarrow Y \leftarrow Z$ so that no edge exists between X and Z.
- d-seperation
  - Definition 6: Let G be a Bayesian network structure and $X_{x-1} - ... - X_n$ be a trail in G. Let E be a subset of nodes from X. There is an active trail between X_1 and X_n given evidence E if:
    - Whenever we have a v-structure $X_{i-1} \rightarrow X_i \leftarrow X_{i+1}$, then X_i or one of its descendants are in E.
    - No other node along the trail is in E. Intuitively, this means that the dependence can “flow” through every triplet $X_{i-1} - X_i - X_{i+1}$.
  - Definition 7:
    - Let X, Y, Z be three sets of nodes in G. We say that X and Z are d-separated given evidence Y, denoted $dsep_G(\mathbf{X}; \mathbf{Z}|\mathbf{Y})$, if there is no active trail between any node X in $\mathbf{X}$ and Z in $\mathbf{Z}$ given evidence Y (7).
    - $dsep_G$ is a property of the graph structure G that corresponds to the notion of conditional independence in the corresponding probability distribution P.

    - $Ind(G)$ is defined as the set of independence statements (of the form "X is independent of Z given Y") that are implied by G.
    - Using this formulation of d-separation, we can use an efficient graph algorithm whose running time scales linearly with the number of nodes in G to check whether any such conditional independence statement holds.

- Equivalence classes
  - Two graphs that imply the same set of conditional independences are in the same equivalence class. One can use PDAGs to represent equivalence classes but this beyond our scope

The elimination algorithm for helping query conditional probabilities on a graph

- Returning to the example... Suppose we want to query specific conditional probabilites (probability of grandparental genotype given child genotype?). How do we do so? I will leave it as an exercise because I want to use an alternative graph.
- Draw the example from Jordan Chapter 3.
- Basic recipe:
  - Write out sum over full joint distribution
  - Move sums as far in as possible (represents the elimination ordering)
  - Replace conditional probabilities in sum's with messages (e.g.$m(x_2)$) to be message passed from node X_2.
  - Continue building messages until all nodes eliminated.
- Example: Felsenstein Pruning algorithm (1981, JME)
  - Tree with data at tips. Compute probability of the data. Each message is pruning off of tips on the tree.

Additional notation:

- Plates for iid variables
- Examples:
  - regression
  - pedigree with multiple iid SNPs, add in a common mutation rate, add in observed genotpyes, add in common error rate for observations

Web resources

- Pe'er review: `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.159.7476&rep=rep1&type=pdf`
- Kollar video on elimination algorithm: `https://www.youtube.com/watch?v=jz02X3hByac`
- Relevant online course notes:
  - `http://www.inf.ed.ac.uk/teaching/courses/pmr/slides/elim-2x2.pdf`
  - `https://www.cs.cmu.edu/~aarti/Class/10701/readings/graphical_model_Jordan.pdf`
  - `http://www.cs.columbia.edu/~blei/fogm/2015F/notes/inference.pdf`
  - `http://www.cs.berkeley.edu/~jordan/courses/281A-fall04/`
- For drawing graphical models
  - `https://github.com/jluttine/tikz-bayesnet`