Tutorial 9, Apr 5, 2024

Graphical Probabilistic Models

- Dependencies between random variables can be represented in a Bayesian network as a DAG
  - DAGs are general enough to represent all possible factorizations of the joint probability
- Example: Let $X \sim \text{Ber}(\alpha)$ and $Y \sim \text{Ber}(\beta)$; let $Z = X \vee Y$ and $X \perp\!\!\!\perp Y$
  - $P(Z = 0) = (1 - \alpha)(1 - \beta)$
  - Note that, given Z, X and Y are no longer independent
  - $P(X = 1|Z = 1) = \dfrac{P(X = 1, Z = 1)}{P(Z = 1)} = \dfrac{\alpha}{1 - (1 - \alpha)(1 - \beta)}$
  - $P(Y = 0|Z = 1) = \dfrac{P(Y = 0, Z = 1)}{P(Z = 1)} = \dfrac{\alpha(1 - \beta)}{1 - (1 - \alpha)(1 - \beta)}$
  - $P(X = 1, Y = 0|Z = 1) = \dfrac{P(X = 1, Y = 0, Z = 1)}{P(Z = 1)} = \dfrac{\alpha(1 - \beta)}{1 - (1 - \alpha)(1 - \beta)} \neq P(X = 1|Z = 1)P(Y = 0|Z = 1)$
  - We have $P(X = 1|Y = 1, Z = 1) < P(X = 1|Z = 1)$ in this case, which is known as *explaining away* – being given more information makes the event less likely
- For Markov random fields, we have an undirected graph instead, which is allowed to have cycles
  - This is typically used in situations were there is no direct causal relationship, e.g. particles in a lattice
  - The conditional probability of a node is dependent only on its immediate neighbours
- The joint distribution factors as $p(\boldsymbol{x}) = \dfrac{1}{Z} \prod_C \psi_C(\boldsymbol{x}_C)$
  - Each individual ψ is a *potential function*
  - In the case of a Bayesian network, the potential functions are conditional probabilities; however in general for a Markov random field they need not be