Mathematical Tools: Probabilistic Models

Lecture 2

Instructor: Su-In Lee

Probability theory review I

- Probability distribution

  Definition: A *probability distribution* P over (Ω,S) is a mapping from events in S (belongs to Ω) to real values that satisfies the following conditions:

[Koller and Friedman]

- Interpretation of probability

Probability theory review II

- Conditional probability
- Chain rule
- Bayes' rule
- Probabilistic independence

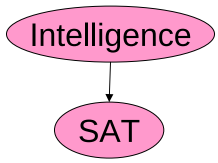
Probabilistic representation

- Joint distribution P over $\{x_1, \ldots, x_n\}$
  - x_i is binary
  - 2^n-1 entries
- If x's are independent
  - $P(x) = p(x_1) \ldots p(x_n)$

Conditional parameterization

- The *Student* example
  - Intelligence (I), SAT (S)
  - Val (I) = $\{i^1,i^0\}$, Val (S) = $\{s^1,s^0\}$

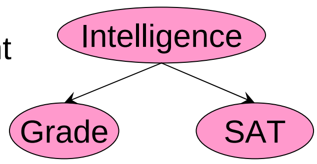
- P(I,S) = P(I) P(S|I)
  - P(I): Prior distribution
  - P(S|I): Conditional probabilistic distribution (CPD)



Naïve Bayes model - example

- Elaborating the *Student* example,
  - Intelligence (I), SAT (S), Grade (G)
  - Val (I) = $\{i^1,i^0\}$, Val (S) = $\{s^1,s^0\}$, Val (G) = $\{g^1,g^2,g^3\}$
  - 12 entries

- If S and G are independent given I,
  - P(I,S,G) = P(I)P(S|I)P(G|I)
  - 7 entries; more compact than joint



Naïve Bayes Model

- A *class* C where Val (C) = $\{c^1,\ldots,c^k\}$.
- *Finding* variables $x_1,\ldots,x_n$
- Naïve Bayes assumption
  - The *findings* are conditionally independent given the individual's *class*.
  - The model *factorizes* as:
- The *Student* example
  - class: intelligence, findings: grade, SAT

Naïve Bayes Model - example

- Medical diagnosis system
  - Class C: disease
  - Findings X: symptoms
- Computing the confidence:
- Drawbacks
  - Strong assumptions

Acknowledgement

- This set of slides is based on the following materials:
  – “Probabilistic Graphical Models: Principles and Techniques” by Daphne Koller and Nir Friedman