Probabilistic Graphical Models
(Bayesian networks, belief networks)

Leonid E. Zhukov

School of Applied Mathematics and Information Science
National Research University Higher School of Economics

5.12.2012
Basic Probability

- Joint probability $P(A, B)$
- Sum rule: $P(A) = \sum_B P(A, B)$
- Product rule: $P(A, B) = P(B|A)P(A)$
- Bayes Theorem

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} = \frac{P(A|B)P(B)}{\sum_B P(A|B)P(B)}$$
Basian Networks

- Basian Networks are directed acyclic graphs, DAG
- Nodes - random variables
- Edges - dependencies between random variables (direct influence)
Basian Networks

- Each node associated with conditional distribution $P(X|\text{parents})$
- Parents - nodes, sending arrows to $X$
- Root nodes associated with priors $P(X)$
Joint distribution

- \( P(x_1, x_2) = P(x_2|x_1)P(x_1) \)
- \( P(x_1, x_2, x_3) = P(x_3|x_1, x_2)P(x_2|x_1)P(x_1) \)
- \( P(x_1, \ldots, x_k) = P(x_k|x_1, \ldots, x_{k-1}) \ldots P(x_2|x_1)P(x_1) \)
- Fully connected, link between every pair of nodes
- Absence of links conveys information
- Factorization properties of joint distribution

\[
P(x_1 \ldots x_k) = \prod_{k=1}^{K} P(x_k|\text{parents}(x_k))
\]
\[ P(x_1 \ldots x_7) = \]
\[ P(x_1)P(x_2)P(x_3)P(x_4|x_1, x_2, x_3)P(x_5|x_1, x_3)P(x_6|x_4)P(x_7|x_4, x_5) \]
Factorization

\[ p(a, b, c) = p(a|c)p(b|c)p(c) \]

\[ p(a, b, c) = p(a)p(b)p(c|a, b) \]

\[ p(a, b, c) = p(a)p(c|a)p(b|c) \]
Example

\[ P(fo) = 0.15 \]
\[ P(bp) = 0.01 \]

- **family-out (fo)**
  - **light-on (lo)**
    - \[ P(lo | fo) = 0.6 \]
    - \[ P(lo | \sim fo) = 0.05 \]
  - **dog-out (do)**
    - \[ P(do | fo \sim bp) = 0.99 \]
    - \[ P(do | fo bp) = 0.90 \]
    - \[ P(do | \sim fo bp) = 0.97 \]
    - \[ P(do | \sim fo bp) = 0.3 \]

- **bowel-problem (bp)**
  - **hear-bark (hb)**
    - \[ P(hb | do) = 0.7 \]
    - \[ P(hb | \sim do) = 0.01 \]
Conditional independence

- Binary variables, fuel system on a car
- B - battery, B=1 charged, B = 0 dead
- F - fuel tank, F = 1 full F = 0 empty
- G - electric gauge, G=1 shows full tank, G =0 shows empty tank
- Prior distribution \( P(B = 1) = 0.9, P(F = 1) = 0.9 \)
- Gauge is unreliable! Conditional on gauge read "full"

\[
\begin{align*}
P(G = 1|B = 1, F = 1) &= 0.8 \\
P(G = 1|B = 1, F = 0) &= 0.2 \\
P(G = 1|B = 0, F = 1) &= 0.2 \\
P(G = 1|B = 0, F = 0) &= 0.1
\end{align*}
\]
Conditional independence

- Observe fuel gauge, reads "empty $G = 0$. Is tank really empty? $P(F = 0|G = 0)$- ?
- Bayes

\[
P(F = 0|G = 0) = \frac{P(G = 0|F = 0)P(F = 0)}{P(G = 0)}
\]

\[
P(G = 0|F = 0) = P(G = 0|B = 0, F = 0)P(B = 0) + P(G = 0|B = 1, F = 0)P(B = 1) = 0.9 \times 0.1 + 0.8 \times 0.9 = 0.81
\]

\[
P(G = 0) = \sum_{B \in \{0, 1\}} \sum_{F \in \{0, 1\}} P(G = 0|B, F)P(B)P(F) = 0.315
\]

\[
P(F = 0|G = 0) = 0.257 > P(F = 0) = 0.1
\]
Conditional independence

We also check (observe) battery, it is dead, $B = 0$

Bayes

$$P(F = 0|G = 0, B = 0) = \frac{P(G = 0|B = 0, F = 0)P(F = 0)P(B = 0)}{\sum_{F \in \{0,1\}} P(G = 0|B = 0, F)P(F)P(B = 0)}$$

$P(F = 0|G = 0, B = 0) = 0.111$

State of fuel tank and battery became dependent through the gauge observation
Learning networks

- Network - joint probability distribution of random variables
- Structure can be learned, often set up "by hand" using expert knowledge
- Probabilities can be estimated from data using MAP/MLE
Probabilistic inference

- Evaluate probability of some set of variables given the values (observations) of another set.
- Exact inference of arbitrary Bayesian network NP-hard
- Exact solutions for polytrees (no undirected cycles, exactly one undirected path between any two nodes)
- Approximate, numerical solutions, structure dependent
References

- Bayesian Networks without Tears. Eugene Charniak, AI magazine, vol 12, No4, 1991 pp 50-63
- Pattern Recognition and Machine Learning, Chapter 6, Christopher Bishop, Springer 2006