Quantitative Methoden Wissensbasierter Systeme

Probabilistische Netzwerke

Norman Heino

Universität Leipzig

16. Januar 2008

Übersicht

Übersicht

Motivation I (Quantitative Logik)

- Symbolische Methoden repräsentieren Wissen nur absolut
- Faktum/Regel wird entweder geglaubt (in Wissensbasis enthalten) oder nicht (in Wissensbasis nicht enthalten)
- Widerspruch zur Realität $\longrightarrow$ Niemand der zur Vorlesung geht, kann mit absoluter Gewissheit sagen, dass diese stattfinden wird
- Alternativ: Quantitative Methoden, die Aussagen, Formeln mit assoziertem **Grad der Überzegung**, **Stärke der Einflussnahme**, **Zugehörigkeitsgrad zu einer Menge (fuzzy set)** versehen
- Häufig werden dazu probabilistische Methoden verwandt

Repetitorium: Wahrscheinlichkeitstheorie

- Zufallsvariable $X \rightarrow$ Realisierung x (Elementarereignis)
- Wahrscheinlichkeitsfunktion $P : \Omega \longrightarrow [0, 1]$
- Zugehörige Wahrscheinlichkeit $P(X = x) = 0.5$

Probabilistische Logik I

- **Vollkonjunktionen:** Konjunktionen, über alle möglichen Literale der Sprache (Elementarereignisse im probabilistischen Sinne).
- Beispiel: $\mathcal{L}$ Sprache mit der Signatur $\Sigma = \{R, W\}$. Sämtliche Vollkonjuktionen über $\mathcal{L}$ sind:

  $$\Omega := \begin{cases} R \wedge W & R \wedge \neg W \\ \neg R \wedge W & \neg R \wedge \neg W \end{cases}$$

  Zu n Aussagen gibt es 2^n Vollkonjunktionen.

Probabilistische Logik II

- Formel über $\mathcal{L}$ stellt Menge von Vollkonjunktionen $\Omega_A := \{\omega \in \Omega \mid A(\omega) = 1\}$ dar, die A implizieren.
- Setze $P(A) := P(\Omega_A)$, dann ist Wahrscheinlichkeit für Formel A gerade die Wahrscheinlichkeit ihrer Primimplikanten. $P(A)$ symbolisiert Grad der Gewissheit mit der A wahr ist.

Motivation II (Probabilistische Netze)

- Problem des probabilistischen Schließens:

$$P(A \wedge B) \leq P(A) \cdot P(B).$$

  Gleichheit gilt nur bei probabilistischer Unabhängigkeit!
- Also: zusätzliches Wissen über (Un-)Abhängigkeiten nötig
- Am Geeignetsten in Graphen repräsentiert
- Im Folgenden: $V = \{X_1, \ldots, X_n\}$ endl. Menge von Aussagenvariablen, G_V Graph, dessen Knoten Elemente von V sind und der Abhängigkeitsstruktur zwischen den Variablen widerspiegelt.

Übersicht

Einführung

Definitionen

Konstruktion

Beispiele

Graphen und Unabhängigkeit I

- Idee der **Separation** in Graphen: $A, B, C \subseteq G$, paarweise disjunkt. C separiert A und B oder

$$A \perp\!\!\!\perp_G B \mid C$$

gdw. jeder Weg zwischen einem Knoten in A und einem Knoten in B mindestens einen Knoten aus C enthält.

- Probabilistische **bedingte Unabhängigkeit**:

$$\begin{aligned} A \perp\!\!\!\perp_P B \mid C \quad &\iff \quad P(A, B \mid C) = P(A \mid C) \cdot P(B \mid C) \\ &\iff \quad P(A \mid C, B) = P(A \mid C) \end{aligned}$$

Graphen und Unabhängigkeit II

Wünschenswert wäre allgemeine Gültigkeit der Äquivalenz:

$$A \perp\!\!\!\perp_P B \mid C \iff A \perp\!\!\!\perp_G B \mid C?$$

Diese jedoch nicht gegeben, da graphische Separation ausdrucksstärker ist

Abhängigkeitsgraph, Unabhängigkeitsgraph

- G heißt **Abhängigkeitsgraph (dependency map, D-map)** zu P, wenn die bedingten Unabhängigkeiten in P durch G repräsentiert werden können:

$$A \perp\!\!\!\perp_P B \mid C \Longrightarrow A \perp\!\!\!\perp_G B \mid C.$$

- G heißt **Unabhängigkeitsgraph (independency map, I-map)** zu P, wenn graphisch separierte Variablenmengen auch probabilistisch unabhängig sind:

$$A \perp\!\!\!\perp_G B \mid C \Longrightarrow A \perp\!\!\!\perp_P B \mid C.$$

- G heißt **perfekter Graph (perfect map)** zu P, wenn G sowohl Abhängigkeits- als auch Unabhängigkeitsgraph zu P ist:

$$A \perp\!\!\!\perp_G B \mid C \Longleftrightarrow A \perp\!\!\!\perp_P B \mid C.$$

Markov-Netz

G heißt **minimaler Unabhängigkeitsgraph (minimal I-map)** oder **Markov-Netz** zu P, wenn G ein Unabhängigkeitsgraph zu P ist, der keine überflüssigen Kanten enthält.

Übersicht

Einführung

Definitionen

Konstruktion

Beispiele

Wahrscheinlichkeitsverteilung und Markov-Netz

1. Existiert zu einer beliebigen Wahrscheinlichkeitsverteilung P ein Markov-Graph?
2. Kann man zu einem bel. Graphen G eine Wahrscheinlichkeitsverteilung P so definieren, dass G ein Unabhängigkeitsgraph zu P ist?

Konstruktion eines Graphen zur Verteilung P

1. Sei G vollständiger Graph auf V.
2. Konstruiere G', indem aus G alle Kanten (A, B) entfernt werden, für die $A \perp\!\!\!\perp_P B \mid (V \setminus \{A, B\})$ gilt.
3. Dann ist G' für alle strikt positiven Wahrscheinlichkeitsverteilungen ein eindeutig bestimmter Markov-Graph. □

Also: $G' = (V, E_0)$ mit $(A, B) \in E_0 \iff A \perp\!\!\!\perp_P B \mid (V \setminus \{A, B\})$ (paarweise Markov-Eigenschaft)

Markov-Decke, Markov-Rand

- Sei $A \in V$ Variable. Als **Markov-Decke (Markov blanket)** $bl(A)$ von A wird jede Variablenmenge $B \subseteq V$ bezeichnet, für die gilt
$$A \perp\!\!\!\perp_P (V \setminus (B \cup \{A\})) \mid B.$$
- Ein **Markov-Rand (Markov boundary)** $br(A)$ von A ist eine minimale Markov-Decke.
- Für strikt positive Verteilungen besitzt jedes Element $A \in V$ einen Markov-Rand, der gerade aus den Nachbarknoten $nb(A)$ besteht (lokale Markov-Eigenschaft):
$$A \perp\!\!\!\perp_P (V \setminus (nb(A) \cup \{A\})) \mid nb(A)$$

3 Markov-Eigenschaften

$$\text{global Markov} \Longrightarrow \text{lokal Markov} \Longrightarrow \text{paarweise Markov}^{1}$$

Damit lässt sich nachprüfen, ob ein bestimmter Graph ein Markov- bzw. Unabhängigkeitsgraph zu einer gegebenen Verteilung ist.

[1] oft gilt sogar Äquivalenz

Konstruktion einer Verteilung P zum Graphen G

Produktdarstellung

- G Graph mit Cliquen $C_1, \ldots, C_m$,
- jeder Clique wird nichtnegative Funktion auf den Vollkonjunktionen der entspr. Variablen zugeordnet, wobei

$$\sum_V \prod_{i=1}^m \psi_i(C_i) \neq 0$$

- Definiere Wahrscheinlichkeitsverteilung P durch

$$P(V) := K \prod_{i=1}^m \psi_i(C_i)$$

  mit K so, dass $\sum_V P(V) = 1$ ist.
- Dann ist G Unabhängigkeitsgraph zu P (P faktorisiert über den Cliquen von G) □

Potentialdarstellung

- Sei V eine endliche Menge von Aussagenvariablen,
- Sei P eine gemeinsame Verteilung über V,
- $\{W_i \mid 1 \leq i \leq p\}$, mit $\bigcup_{i=1}^{p} W_i = V$,
- $W_{\text{kon}} := \{w_i \mid w_i$ ist Vollkonjunktion von Variablen in $W_i, (1 \leq i \leq p)\}$,
- Def. Funktion $\psi : W_{\text{kon}} \longrightarrow \mathbb{R}^{\geq 0}$,
- $\{W_1, \ldots, W_p; \psi\}$ heißt **Potentialdarstellung** von P, falls gilt

$$P(V) = K \prod_{i=1}^{p} \psi(W_i).$$

$\square$

Faktorisierung triangulierter Graphen

Ist G ein triangulierter Unabhängigkeitsgraph zur Verteilung P, so faktorisiert P in der Form

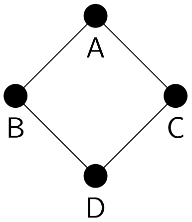
$$P(V) = \frac{\prod_{C_i} P(C_i)}{\prod_{S_j} P(S_j)},$$

wobei die C_i und S_j die Cliquen und Separatoren eines Cliquenbaumes von G sind. $\square$

Übersicht

Beispiel: Infektion (1)

- Vier Personen A, B, C, D, von denen jede mit genau zwei der drei anderen Kontakt hat (s. Graph).



- Vier Cliquen: $C_1 = \{A, B\}$, $C_2 = \{A, C\}$, $C_3 = \{B, D\}$, $C_4 = \{C, D\}$.
- Markov-Feld:
  $P(A, B, C, D) = K \cdot \psi_1(A, B)\psi_2(A, C)\psi_3(B, B)\psi_4(C, B)$.

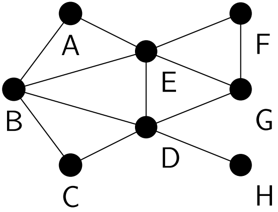
Beispiel: Infektion (2)

- Korrekte Deutung der Potentialfunktionen nicht klar! Denkbar bspw:

$$\psi_i(X,Y) := \begin{cases} \alpha_i, & \text{wenn } X \text{ und } Y \text{ entweder beide infiziert} \\ & \text{oder beide nicht infiziert sind} \\ \beta_i, & \text{wenn genau einer infiziert ist.} \end{cases}$$

- Wobei α_i und β_i bspw. die Kontakthäufigkeit des jeweiligen Paares widerspiegeln.

Beispiel: mit Triangulierung



- Cliquen: $\{B, C, D\}$, $\{B, D, E\}$, $\{A, B, E\}$, $\{D, E, G\}$, $\{E, F, G\}$, $\{D, H\}$.
- Separatoren: $\{B, D\}$, $\{B, E\}$, $\{D, E\}$, $\{E, G\}$, $\{D\}$.
- Markov-Feld: $P(A, B, C, D, E, F, G) =$

$$\frac{P(B, C, D)P(B, D, E)P(A, B, E)P(D, E, G)P(E, F, G)P(D, H)}{P(B, D)P(B, E)P(D, E)P(E, G)P(D)}$$