

Übung 9

Niclas Kusenbach, 360227

21st January 2026

Aufgabe 1: Inferenz durch Sampling

a) Welches Problem stellt sich bei der Anwendung von Rejection Sampling?

Das Hauptproblem beim Rejection Sampling ist die **Ineffizienz bei seltenen Evidenzen**. Wenn die Wahrscheinlichkeit der beobachteten Evidenz $P(E = e)$ sehr klein ist, werden die meisten generierten Samples der Prior-Verteilung nicht mit der Evidenz übereinstimmen und müssen verworfen (rejected) werden. Dies führt dazu, dass sehr viele Samples generiert werden müssen, um eine statistisch signifikante Anzahl an gültigen Samples zu erhalten.

b) Wie addressiert Likelihood Weighting dieses Problem? Wie Gibbs Sampling?

- **Likelihood Weighting:** Bei diesem Verfahren werden Samples, die der Evidenz widersprechen, nicht verworfen. Stattdessen werden die Evidenzvariablen auf ihre beobachteten Werte fixiert. Um die dadurch entstehende Verzerrung auszugleichen, wird jedem Sample ein **Gewicht** (w) zugewiesen. Dieses Gewicht entspricht der Wahrscheinlichkeit ("Likelihood") der Evidenzvariablen gegeben den Werten ihrer Elternknoten im aktuellen Sample. Dadurch wird jedes generierte Sample genutzt.
- **Gibbs Sampling:** Gibbs Sampling ist ein Markov-Chain-Monte-Carlo (MCMC) Verfahren. Anstatt unabhängige Samples zu generieren und zu prüfen, wandert der Algorithmus durch den Zustandsraum. Die Evidenzvariablen bleiben fixiert. Die Nicht-Evidenzvariablen werden iterativ neu gesampelt, basierend auf der bedingten Wahrscheinlichkeit gegeben ihrer Markov Blanket (Eltern, Kinder, Eltern der Kinder). Da man sich so im Bereich der Posterior-Verteilung bewegt, müssen keine Samples aufgrund von falscher Evidenz verworfen werden (nach der Burn-In-Phase).

Aufgabe 2: Bayesische Netze und Variablenelimination

Gegeben ist die Struktur $A \rightarrow B, A \rightarrow C, B \rightarrow D, C \rightarrow D$.

2a) Faktorisierung der Verbundsverteilung

Ein Bayesisches Netz faktorisiert als das Produkt der bedingten Wahrscheinlichkeiten aller Knoten gegeben ihrer Eltern:

$$P(A, B, C, D) = P(A) \cdot P(B|A) \cdot P(C|A) \cdot P(D|B, C)$$

2b) Bedingte Unabhängigkeit

Die Struktur zeigt, dass B und C bedingt unabhängig sind, wenn A gegeben ist (divergierende Verbindung / "Common Cause"). Formal:

$$P(C | B, A) = P(C | A) \quad \text{bzw.} \quad B \perp C | A$$

Das Wissen über B liefert keine neuen Informationen über C , wenn die gemeinsame Ursache A bereits bekannt ist.

2c) $P(D)$ mittels Variablen-Elimination

Gesucht ist $P(D)$. Wir marginalisieren über A, B, C :

$$P(D) = \sum_A \sum_B \sum_C P(A, B, C, D)$$

Einsetzen der Faktorisierung:

$$P(D) = \sum_A \sum_B \sum_C P(A) \cdot P(B|A) \cdot P(C|A) \cdot P(D|B, C)$$

Umgruppieren der Terme (Pushing Sums), um die Summen so tief wie möglich zu verschachteln:

$$P(D) = \sum_A P(A) \sum_B P(B|A) \left[\sum_C P(C|A) \cdot P(D|B, C) \right]$$

Schritt 1: Elimination von C

Wir definieren einen neuen Faktor $\tau_1(A, B, D)$ für die innerste Summe:

$$\tau_1(A, B, D) = \sum_C P(C|A) \cdot P(D|B, C)$$

Die Gleichung wird zu: $P(D) = \sum_A P(A) \sum_B P(B|A) \cdot \tau_1(A, B, D)$

Schritt 2: Elimination von B

Wir summieren über B und erhalten $\tau_2(A, D)$:

$$\tau_2(A, D) = \sum_B P(B|A) \cdot \tau_1(A, B, D)$$

Die Gleichung wird zu: $P(D) = \sum_A P(A) \cdot \tau_2(A, D)$

Schritt 3: Elimination von A

Abschließend summieren wir über A :

$$P(D) = \sum_A P(A) \cdot \tau_2(A, D)$$