

Der Propp-Wilson-Algorithmus

Marie Cuno

6. November 2008

Der Propp-Wilson-Algorithmus ist ein Algorithmus zum Ziehen von exakten Stichproben auf einem gegebenen, endlichen Zustandsraum. Das Besondere an dem Algorithmus ist, dass man die stationäre Verteilung exakt erhält und nicht wie bei der klassischen Markov-Chain-Monte-Carlo-Methode nur approximativ, und das bei einem klaren Abbruchkriterium.

Die Grundidee des Algorithmus ist zum einen, dass in jedem möglichen Zustand eine Kopie der Markovkette gestartet wird. Zum anderen werden die Ketten nicht zum Zeitpunkt 0 gestartet und immer *weiter in Zukunft* iteriert, sondern man produziert Stücke der Ketten, die immer *tiefer in der Vergangenheit* gestartet werden und bis zum Zeitpunkt 0 laufen. Deshalb ist der Algorithmus auch unter dem Begriff *Coupling from the past* bekannt.

Beschreibung des Algorithmus

Konstruiere eine irreduzible und aperiodische Markovkette auf einem endlichen Zustandsraum $S = \{s_1, \dots, s_k\}$ mit stationärer Verteilung π , Übergangsmatrix P , Übergangsfunktion $\phi : S \times [0, 1] \rightarrow S$ und N_1, N_2, N_3, \dots eine aufsteigende Folge positiver ganzer Zahlen.¹ Sei weiterhin $(U_i)_{i \leq 0}$ eine Folge unabhängiger, identisch verteilter Zufallsvariablen aus $[0, 1]$.

Der Propp-Wilson-Algorithmus funktioniert wie folgt:

Beginne mit $m = 1$.

1. Für jedes $s \in \{s_1, \dots, s_k\}$ simuliere eine Markovkette mit Startzeit $-N_m$ und Startzustand s bis zur Endzeit 0. Die Übergänge ergeben sich durch ϕ und die Zufallsvariablen $U_{-N_m+1}, U_{-N_m+2}, \dots, U_{-1}, U_0$.

Bemerkung: Diese Zufallsvariablen sind für jede der k Ketten gleich.

2. Sind alle k Ketten zur Zeit 0 im gleichen Zustand $s' \in S$, stoppt der Algorithmus mit der Ausgabe s' ; ansonsten erhöhe m um eins und führe Schritt zwei erneut durch.

Für die Korrektheit des Algorithmus ist es notwendig, dass die im m -ten Durchlauf benötigten Zufallsvariablen $U_{-N_m+1}, \dots, U_{-1}, U_0$ den Zufallsvariablen $U_{-N_{m-1}+1}, \dots, U_0$ aus dem $(m-1)$ -ten Durchlauf entsprechen und lediglich die Zufallsvariable U_{-N_m+1} neu ist.

Die Korrektheit des Algorithmus

Satz: Sei P die Übergangsmatrix einer irreduziblen und aperiodischen Markovkette auf dem Zustandsraum $S = \{s_1, \dots, s_k\}$ mit stationärer Verteilung $\pi = (\pi(s_1), \dots, \pi(s_k))$ und sei ϕ eine Übergangsfunktion für P . Betrachte nun den Propp-Wilson-Algorithmus wie beschrieben. Unter der Annahme, dass der Algorithmus mit Wahrscheinlichkeit 1 terminiert und Y ausgibt, gilt für jedes $i \in \{1, \dots, k\}$

$$P(Y = s_i) = \pi(s_i).$$

¹In der Regel wählt man $(N_1, N_2, N_3, \dots) = (1, 2, 4, 8, \dots)$.

„Sandwiching“

In jedem Iterationsschritt des Algorithmus müssen $|S|$ Markovketten parallel laufen, was für großes $|S|$ nicht effizient realisierbar ist. Unter gewissen Struktureigenschaften der Markovkette kann die Anzahl der notwendigen Ketten stark reduziert werden. Eine Methode ist das sogenannte *Sandwiching*, es nutzt eine gewisse Monotonie der Markovkette aus.

Wir nehmen an, der Zustandsraum $|S|$ ist halbgeordnet und es existieren ein größtes Element $\hat{1}$ und ein kleinstes Element $\hat{0}$ in S , so dass gilt:

$$\hat{0} \leq i \leq \hat{1} \quad \text{für alle } i \in S.$$

Lässt sich Übergangsfunktion so wählen, dass die Ordnung in jedem Iterationsschritt der Kette beibehalten wird, so liegt Verschmelzung aller Ketten genau dann vor wenn die in $\hat{0}$ und in $\hat{1}$ gestarteten Ketten verschmolzen sind. Dementsprechend genügt es, lediglich diese beiden Ketten zu betrachten.

Anwendung: Das Ising-Modell

Sei $G = (V, E)$ ein Graph. Das Ising-Modell ist eine Methode, aus $\{-1, 1\}^V$ ein zufälliges Element zu ziehen.

Sei $\beta \geq 0$ ein fester Parameter (*inverse Temperatur*) und

$$H(\xi) = - \sum_{(x,y) \in E} \xi(x)\xi(y)$$

die *Energie* der Konfiguration $\xi \in \{-1, 1\}^V$. Im Ising-Modell wird eine zufällige Konfiguration nach folgender Verteilung $\pi_{G,\beta}$ gezogen: Für $\xi \in \{-1, 1\}^V$ sei

$$\pi_{G,\beta}(\xi) = \frac{1}{Z_{G,\beta}} \exp(-\beta H(\xi)).$$

Das Modell, wie die Bezeichnung der Parameter, ist physikalisch motiviert, die Konfigurationen versteht man als Spinkonfigurationen $+1/-1$ von Atomen in ferromagnetischen Stoffen.

Man beachte, dass die ± 1 für $\beta > 0$ auf den Knoten nicht gleichverteilt sind, es liegt eine Art Nachbarschaftsstruktur vor. Die Wahrscheinlichkeit, dass einem bestimmten Knoten $v \in V$ eine $+1$ zugeordnet wird, hängt von der Belegung der Nachbarknoten ab. Je mehr benachbarte Knoten eine $+1$ haben, desto wahrscheinlicher wird es, dass auch dem Knoten v eine $+1$ zugeordnet wird. Das System neigt also zur Verklumpung.