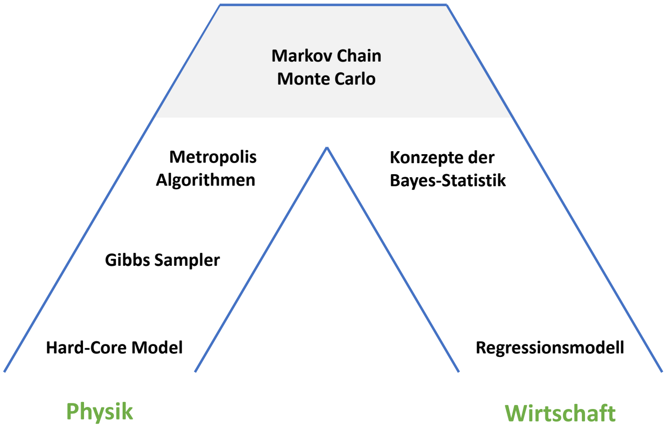
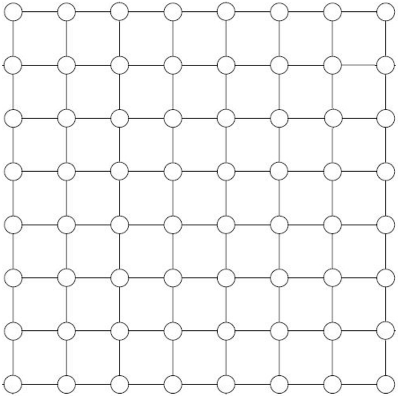
Markov-Chain-Monte-Carlo Simulation

Jonas Rothfuß
Marius Causemann
Fabian Neumann

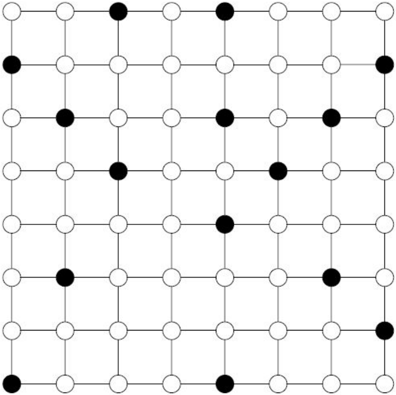
15. Juni 2015


Markov Chain
Monte Carlo
Metropolis
Algorithmen
Gibbs Sampler
Hard-Core Model
Konzepte der
Bayes-Statistik
Regressionsmodell
Physik
Wirtschaft

Darstellung des Hard-Core Models



Zulässige Konfiguration des Hard-Core Models



Grundlagen der Simulation

Satz (Gesetz der großen Zahlen)

Seien $X_1, X_2, \ldots$ unabhängig identisch verteilte Zufallsvariablen mit endlichem Mittelwert μ und endlicher Varianz σ^2. Sei $M_n = \frac{1}{n}(X_1 + \ldots + X_n)$. Dann gilt fast sicher:

$$\lim_{n\to\infty} M_n = \mu$$

Satz (Konvergenzsatz der Markovketten)

Sei $(X_0, X_1, ...)$ eine irreduzible und aperiodische Markov Kette mit Zustandsraum $S = \{s_1, ..., s_k\}$, Übergangsmatrix P und zufälliger Anfangsverteilung $\mu^{(0)}$. Dann gilt für die stationäre Verteilung π von P

$$\mu^{(n)} \quad \xrightarrow{n\to\infty} \quad \pi$$

Konstruktion der Markovkette zur Simulation

Man konstruiere zunächst eine irreduzible und aperiodische Markovkette $(X_0, X_1, ...)$ mit der stationären Verteilung μ_G und mit dem Zustandsraum

$$S = \{\xi \in \{0,1\}^V : \xi \text{ zulässig}\}$$

Wir konstruieren eine entsprechende Markovkette mithilfe der folgenden Übergangsregel zum Zeitpunkt $n+1$:

(1) Wähle zufällig einen Knoten $v \in V$ unter einer Gleichverteilung.

(2) Wirf eine faire Münze.

(3) Falls die Münze auf Kopf landet und zum Zeitpunkt n alle zu v adjazenten Knoten den Wert 0 annehmen, setze $X_{n+1}(v) = 1$. Andernfalls setze $X_{n+1}(v) = 0$.

(4) Für alle anderen Knoten $w \in V$ belasse den Wert unverändert, sodass $X_{n+1}(w) = X_n(w)$ gilt.

Beweis der Aperiodizität

Definition (Periode)

Sei $(X_0, X_1, \ldots)$ eine Markovkette mit Zustandsraum $S = \{s_1, \ldots, s_k\}$ und Übergangsmatrix P. Die Periode $d(s_i)$ eines Zustands $s_i \in S$ ist definiert durch

$$d(s_i) = \mathrm{ggT}\{n \geq 1 : (P^n)_{i,i} > 0\}$$

Definition (Aperiodizität eines Zustands)

Gilt $d(s_i) = 1$, so heißt der Zustand s_i aperiodisch.

Definition (Aperiodizität einer Markovkette)

Eine Markovkette heißt aperiodisch, wenn alle Zustände aperiodisch sind. Andernfalls nennen wir die Kette periodisch.

Beweis der Irreduzibilität

Definition (Kommunizierende Zustände)

Sei $(X_0, X_1, \ldots)$ eine Markovkette mit der Zustandsmenge $S = \{s_1, \ldots, s_k\}$ und Übergangsmatrix P. Der Zustand s_i kommuniziert mit Zustand s_j, falls ein $n \in \mathbb{N}$ existiert, sodass

$$\mathbb{P}(X_{m+n} = s_j | X_m = s_i) > 0$$

mit $m \in \mathbb{N}$ und $s_i, s_j \in S$.

Definition (Interkommunikation)

Gilt $s_i \to s_j$ und $s_j \to s_i$, so interkommunizieren die Zustände s_i und s_j. Man schreibt:

$$s_i \leftrightarrow s_j$$

Beweis der Irreduzibilität

Definition (Irreduzibilität)

Eine Markovkette $(X_0, X_1, \ldots)$ mit der Zustandsmenge $S = \{s_1, \ldots, s_k\}$ und Übergangsmatrix P heißt irreduzibel, wenn $\forall s_i, s_j \in S$ gilt:

$$s_i \leftrightarrow s_j$$

Sonst heißt die Markovkette reduzibel.

Beweis der Stationarität über Reversibilität

Satz (Reversibilität und Stationarität einer Markovkette)

Sei $(X_0, X_1, \ldots)$ eine Markovkette mit der Zustandsmenge $S = \{s_1, \ldots, s_k\}$ und Übergangsmatrix P. Ist π eine reversible Verteilung der Kette, so ist sie auch eine stationäre Verteilung der Kette.

Übergangsregel eines allgemeinen Gibbs Samplers

(1) Wähle zufällig einen Knoten $v \in V$ unter einer Gleichverteilung.

(2) Wähle $X_{n+1}(v)$ unter Berücksichtigung der vorgegebenen Verteilung gegeben alle anderen Knoten nehmen für X_n zulässige Werte an.

(3) Setze $X_{n+1}(w) = X_n(w)$ für alle Knoten $w \in V \setminus \{v\}$.

Bemerkung

Wenn man die Markov-Kette nach o.g. Übergangsregel konstruiert und zusätzlich die Bedingungen der Aperiodizität und Irreduzibilität erfüllt werden, hat sie eine eindeutige stationäre Verteilung π.

Bedingungen der Konstruktion von Metropolis-Ketten

Bedingungen an die Kantenkonfiguration des Graphen zur Konstruktion einer Metropolis-Kette:

1. Der Graph muss zusammenhängend sein, damit die resultierende Markov-Kette irreduzibel ist.
2. Ein Knoten sollte nicht zu viele Nachbarn haben, da die Markov-Kette sonst rechentechnisch zu aufwendig wird um in der Praxis Simulationszwecken zu dienen.

Metropolis Algorithmus im diskreten Fall

Sei $\pi = (\pi_1, \dots, \pi_k)$ eine diskrete Wahrscheinlichkeitsverteilung auf der endlichen Zustandsmenge $S = \{s_1 \dots, s_k\}$. Darüber hinaus sei $G = (S, E)$ ein zusammenhängender Graph mit der mit der Knotenmenge S. Eine Markov-Kette $Y = (X_n)_{n \in \mathbb{N}}$ mit dem Zustandsraum S und der Übergangsmatrix

$$P_{i,j} = \begin{cases} \frac{1}{d_i} \min\left\{\frac{\pi_j d_i}{\pi_i d_j}, 1\right\} & \text{falls } s_i \text{ und } s_j \text{ adjazent} \\ 0 & \text{falls } s_i \neq s_j \text{ nicht adjazent} \\ 1 - \sum_{l, s_l \sim s_i} \frac{1}{d_i} \min\left\{\frac{\pi_l d_i}{\pi_i d_l}, 1\right\} & \text{falls } i = j \end{cases}$$

mit d_i als Grad des Knotens s_i, ist reversibel und besitzt π als stationäre Verteilung.

Alternative Formulierung des stochastischen Übergangskerns

Sei $X_n = s_i$ mit $s_i \in S$

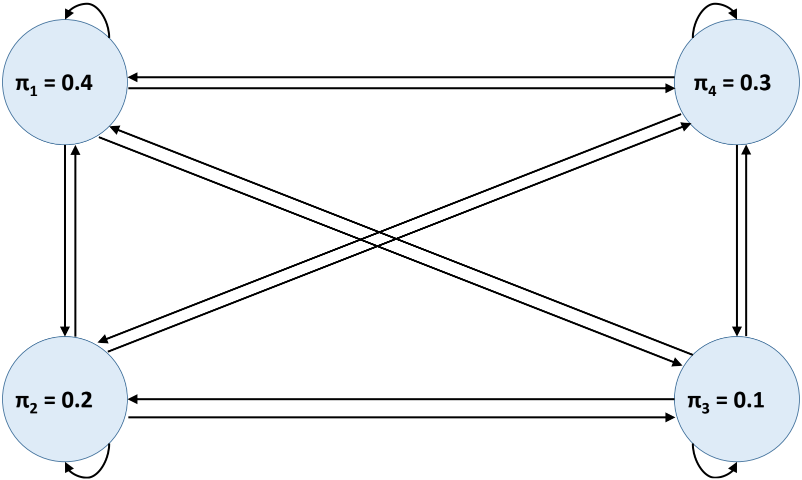
1. Wähle einen beliebigen Knoten $s_j \in S$ gemäß der Gleichverteilung

$$\mu(s_j) = \begin{cases} \frac{1}{d_i} & \text{falls } s_i \text{ und } s_j \text{ adjazent} \\ 0 & \text{sonst} \end{cases}$$

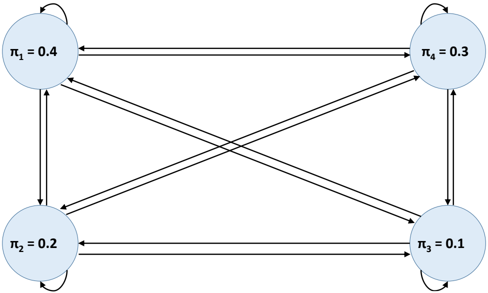
2. Setze

$$X_{n+1} = \begin{cases} s_j & \text{mit Wahrscheinlichkeit } \min\left\{\frac{\pi_j d_i}{\pi_i d_j}, 1\right\} \\ s_i & \text{mit Wahrscheinlichkeit } 1 - \min\left\{\frac{\pi_l d_i}{\pi_i d_l}, 1\right\} \end{cases}$$

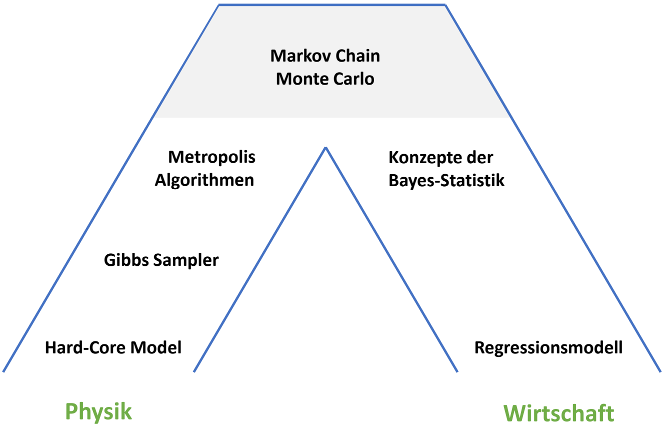
Vollständiger Graph bei beliebiger Verteilung



Vollständiger Graph bei beliebiger Verteilung



$$P^b_{i,j} = \begin{pmatrix} 1/2 & 1/4 & 1/6 & 1/12 \\ 1/3 & 1/3 & 2/9 & 1/9 \\ 1/3 & 1/3 & 1/6 & 1/6 \\ 1/3 & 1/3 & 1/3 & 0 \end{pmatrix}$$


Markov Chain
Monte Carlo
Metropolis
Algorithmen
Konzepte der
Bayes-Statistik
Gibbs Sampler
Hard-Core Model
Regressionsmodell
Physik
Wirtschaft

Wdhl: Metropolis-Hastings-Algorithmus im diskreten Fall

Sei $\pi = (\pi_1, \ldots, \pi_k)$ eine diskrete Wahrscheinlichkeitsverteilung auf der endlichen Zustandsmenge $S = \{s_1 \ldots, s_k\}$. Darüber hinaus sei $G = (S, E)$ ein zusammenhängender Graph mit der mit der Knotenmenge S. Eine Markov-Kette $Y = (X_n)_{n\in\mathbb{N}}$ mit dem Zustandsraum S und der Übergangsmatrix

$$P_{i,j} = \begin{cases} \frac{1}{d_i} \min\left\{\frac{\pi_j d_i}{\pi_i d_j}, 1\right\} & \text{falls } s_i \text{ und } s_j \text{ adjazent} \\ 0 & \text{falls } s_i \neq s_j \text{ nicht adjazent} \\ 1 - \sum_{l, s_l \sim s_i} \frac{1}{d_i} \min\left\{\frac{\pi_l d_i}{\pi_i d_l}, 1\right\} & \text{falls } i = j \end{cases}$$

mit d_i als Grad des Knotens s_i, ist reversibel und besitzt π als stationäre Verteilung.

Wdhl: Metropolis-Hastings-Algorithmus im diskreten Fall

Sei $X_n = s_i$ mit $s_i \in S$

1. Wähle einen beliebigen Knoten $s_j \in S$ gemäß der Gleichverteilung

$$\mu(s_j) = \begin{cases} \frac{1}{d_i} & \text{falls } s_i \text{ und } s_j \text{ adjazent} \\ 0 & \text{sonst} \end{cases}$$

2. Setze

$$X_{n+1} = \begin{cases} s_j & \text{mit Wahrscheinlichkeit } \min\left\{\frac{\pi_j d_i}{\pi_i d_j}, 1\right\} \\ s_i & \text{mit Wahrscheinlichkeit } 1 - \min\left\{\frac{\pi_l d_i}{\pi_i d_l}, 1\right\} \end{cases}$$

Allgemeine Markov-Ketten

- Reversibilität:
  Understanding the Metropolis-Hastings Algorithm, Siddhartha Chib und Edward Greenberg, The American Statician, November 1995, Vol. 49, Nr. 4, Seite 328
- Irreduzibilität:
  General state space Markov chains and MCMC algorithms, Gareth O. Roberts und Jeffrey S. Rosenthal, Seite 30, 31
- Aperiodizität:
  General state space Markov chains and MCMC algorithms, Gareth O. Roberts und Jeffrey S. Rosenthal, Seite 32

Metropolis-Hastings Algorithmus (1)

Algorithmus

Sei $f(\theta) : \Theta \longrightarrow \mathbb{R}^+$ *die Dichtefunktion einer Wahrscheinlichkeitsverteilung. Gemäß dem Metropolis-Hastings-Algorithmus definiert man eine Markov-Kette* $(X_n)_{n \in \mathbb{N}}$ *auf dem Zustandsraum* Θ *und wählt eine Ausgangskonfiguration* $X_0 \in \Theta$ *mit* $f(X_0) > 0$.

Metropolis-Hastings Algorithmus (2)

Algorithmus

Der Übergangskern K von $(X_n)_{n\in\mathbb{N}}$ ist durch folgendes Schema charakterisiert:

1. *Erzeuge ausgehend von $X_n = \theta$ eine zufällige Realisation θ' einer Zufallsvariable gemäß der Vorschlagdichte $q(\theta'|\theta)$*
2. *Berechne die Annahmewahrscheinlichkeit $\alpha(\theta, \theta')$ von θ'*

$$\alpha(\theta,\theta') = min\left(1, \frac{f(\theta')q(\theta|\theta')}{f(\theta)q(\theta'|\theta)}\right) \tag{1}$$

3. *Setze*

$$X_{n+1} = \begin{cases} \theta' & \textit{mit Wahrscheinlichkeit } \alpha(\theta,\theta') \\ \theta & \textit{mit Wahrscheinlichkeit } 1-\alpha(\theta,\theta') \end{cases} \tag{2}$$

Metropolis-Hastings Algorithmus (3)

Lemma (Träger-Kriterium)

Sei $q_m(\theta'|\theta)$ die Dichte von $X_{n+m} = \theta'$ unter der Bedingung $X_n = \theta$ (m-stufige Vorschlagsdichte) und

$$\exists m \in \mathbb{N} : q_m(\theta'|\theta) > 0 \; \forall \; \theta', \theta \in supp(f) \tag{3}$$

so kann die, gemäß des Metropolis-Hastings-Algorithmus konstruierte, Markov-Kette $(X_n)_{n \in \mathbb{N}}$ mit $X_0 \in supp(f)$ alle $\theta \in \Theta$ mit $f(\theta) > 0$ erreichen.

Reversibilität und Stationarität der Metropolis-Kette

Satz

Sei $(X_n)_{n\in\mathbb{N}}$ eine Markov-Kette die gemäß des Metropolis-Hastings-Algorithmus konstruiert wurde. Für jede Vorschlagsdichte q, die das Träger-Kriterium erfüllt, gilt

(a) *Der Übergangskern $K(\theta', \theta)$ der Markov-Kette erfüllt die Detailed-Balance-Bedingung mit f und ist damit reversibel*

(b) *f ist die Dichte der stationären Verteilung der Markov Kette $(X_n)_{n\in\mathbb{N}}$*

Beweis des Satzes (1)

Um (a) zu beweisen, muss gezeigt werden, dass $K(\theta, \theta')$ die Detailed-Balance-Bedingung

$$f(\theta)K(\theta', \theta) = f(\theta')K(\theta, \theta') \; \forall \theta, \theta' \in \Theta \tag{4}$$

erfüllt ist und K somit reversibel ist.

Fall 1: $(\theta = \theta')$

Die Gleichung ist offensichtlich erfüllt, da $\theta = \theta'$

$$f(\theta)K(\theta', \theta) = f(\theta')K(\theta, \theta')$$

Beweis des Satzes (2)

Fall 2: *($\theta \neq \theta'$)*
Da $\theta \neq \theta'$ gilt, folgt $(1 - r(\theta'))\mathbb{1}_{\{\theta=\theta'\}}(\theta)) = 0$ bzw, $(1 - r(\theta))\mathbb{1}_{\{\theta'=\theta\}}(\theta')) = 0$. Für Detailed-Balance-Bedingung gilt deshalb

$$f(\theta)q(\theta'|\theta)\alpha(\theta,\theta') = f(\theta')q(\theta|\theta')\alpha(\theta',\theta)$$

$$\begin{aligned}
\iff \quad & f(\theta)q(\theta'|\theta) \cdot min\left(1, \frac{f(\theta')q(\theta|\theta')}{f(\theta)q(\theta'|\theta)}\right) \\
= \quad & f(\theta')q(\theta|\theta') \cdot min\left(1, \frac{f(\theta)q(\theta'|\theta)}{f(\theta')q(\theta|\theta')}\right)
\end{aligned}$$

Beweis des Satzes (3)

Fall 2.1: $\frac{f(\theta')q(\theta|\theta')}{f(\theta)q(\theta'|\theta)} \geq 1$

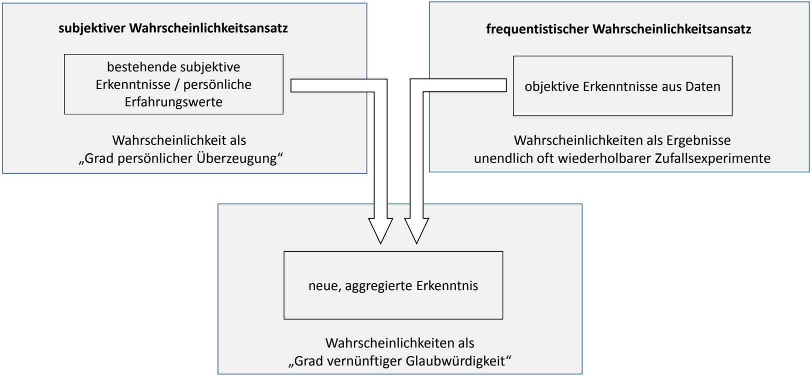
$$f(\theta)q(\theta'|\theta) = f(\theta')q(\theta|\theta')\frac{f(\theta)q(\theta'|\theta)}{f(\theta')q(\theta|\theta')}$$

Fall 2.2: $\frac{f(\theta')q(\theta|\theta')}{f(\theta)q(\theta'|\theta)} < 1$

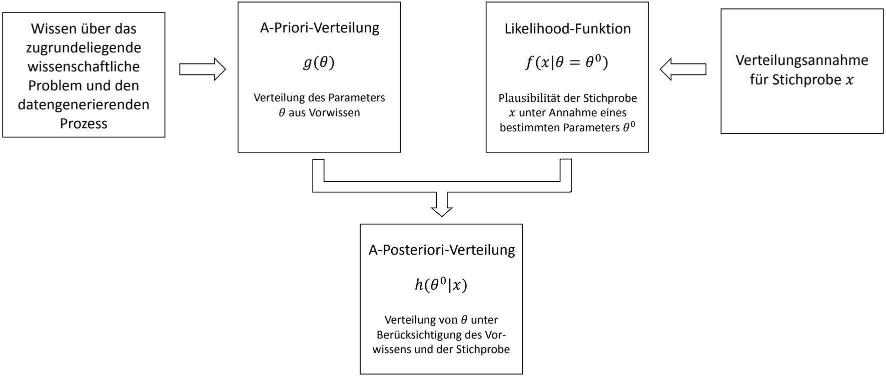
$$f(\theta)q(\theta'|\theta) * \frac{f(\theta')q(\theta|\theta')}{f(\theta)q(\theta'|\theta)} = f(\theta')q(\theta|\theta')$$

Hiermit wurde gezeigt, dass in beiden Fällen die Detailed-Balance-Bedingung erfüllt ist und (a) gültig ist. Die Behauptung (b) des Satzes folgt automatisch aus (a). □

Bayes Statistik



Bayes Inferenz



Beispiel zur Bayes-Inferenz

Beispiel (Ankunftsverhalten von Zügen im Streikfall)

Modellannahme:

- *Zwischenankunftszeiten der Züge in Stunden:* $X \sim exp(\theta)$
- *Die A-Priori-Verteilung des Parameters* θ *ist* $\gamma(p, b)$.
  D.h. die Dichte für $\theta \geq 0$ *lautet* $g(\theta) = \frac{b^p}{\Gamma(p)} \theta^{p-1} e^{-b\theta}$
  (Gammaverteilung), wobei $\Gamma(p) = \int_0^\infty t^{p-1} e^{-t} dt$ *(Gammafunktion)*

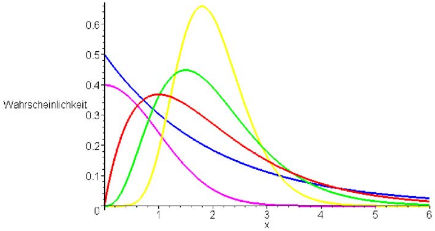


Abbildung: Mögliche Dichten der Gammaverteilung $\gamma(p, b)$

Berechnung der A-Posteriori-Dichte (1)

$$h(\theta|\vec{x}) = f(\vec{x}|\theta) \cdot g(\theta) \cdot \underbrace{\frac{1}{\int_0^\infty f(\vec{x}|\theta^\star) \cdot g(\theta^*) d\theta^\star}}_{c(\vec{x})}$$

$$= \theta^n \cdot \exp\left(-\theta \cdot \sum_{i=1}^{n} x_i\right) \cdot \frac{b^p}{\Gamma(p)} \cdot \theta^{p-1} \cdot e^{-b\theta} \cdot c(\vec{x})$$

$$= \underbrace{\theta^{(n+p)-1} \cdot \exp\left(-\theta\left(b + \sum_{i=1}^{n} x_i\right)\right)}_{D(\theta,\vec{x})} \cdot c^*(\vec{x})$$

Berechnung der A-Posteriori-Dichte (1)

$$h(\theta|\vec{x}) = f(\vec{x}|\theta) \cdot g(\theta) \cdot \underbrace{\frac{1}{\int_0^\infty f(\vec{x}|\theta^\star) \cdot g(\theta^*) d\theta^\star}}_{c(\vec{x})}$$

$$= \theta^n \cdot \exp\left(-\theta \cdot \sum_{i=1}^{n} x_i\right) \cdot \frac{b^p}{\Gamma(p)} \cdot \theta^{p-1} \cdot e^{-b\theta} \cdot c(\vec{x})$$

$$= \underbrace{\theta^{(n+p)-1} \cdot \exp\left(-\theta\left(b + \sum_{i=1}^{n} x_i\right)\right)}_{D(\theta,\vec{x})} \cdot c^*(\vec{x})$$

Berechnung der A-Posteriori-Dichte (1)

$$h(\theta|\vec{x}) = f(\vec{x}|\theta) \cdot g(\theta) \cdot \underbrace{\frac{1}{\int_0^\infty f(\vec{x}|\theta^\star) \cdot g(\theta^*) d\theta^\star}}_{c(\vec{x})}$$

$$= \theta^n \cdot \exp\left(-\theta \cdot \sum_{i=1}^{n} x_i\right) \cdot \frac{b^p}{\Gamma(p)} \cdot \theta^{p-1} \cdot e^{-b\theta} \cdot c(\vec{x})$$

$$= \underbrace{\theta^{(n+p)-1} \cdot \exp\left(-\theta\left(b + \sum_{i=1}^{n} x_i\right)\right)}_{D(\theta,\vec{x})} \cdot c^*(\vec{x})$$

Berechnung der A-Posteriori-Dichte (2)

Behauptung: $h(\theta|\vec{x})$ ist die Dichte von $\gamma(b^\star, p^\star)$ mit

$$p^\star = n + p$$

$$b^\star = b + \sum_{i=1}^{n} x_i$$

$$\int_\Theta D(\theta,\vec{x}) \cdot c^*(\vec{x}) d\theta = 1 = \int_\Theta \frac{b^{\star p^\star}}{\Gamma(p^\star)} \cdot D(\theta,\vec{x}) d\theta$$

$$\iff c^*(\vec{x}) = \frac{b^{\star p^\star}}{\Gamma(p^\star)}$$

$\Rightarrow h(\theta|\vec{x})$ entspricht der Dichte von $\gamma(b^\star, p^\star)$

$$\Rightarrow \hat{\theta} = \mathbb{E}(\theta) = \frac{p^\star}{b^\star}$$

Berechnung der A-Posteriori-Dichte (2)

Behauptung: $h(\theta|\vec{x})$ ist die Dichte von $\gamma(b^\star, p^\star)$ mit

$$p^\star = n + p$$
$$b^\star = b + \sum_{i=1}^{n} x_i$$

$$\int_\Theta D(\theta, \vec{x}) \cdot c^*(\vec{x}) d\theta = 1 = \int_\Theta \frac{b^{\star p^\star}}{\Gamma(p^\star)} \cdot D(\theta, \vec{x}) d\theta$$

$$\iff c^*(\vec{x}) = \frac{b^{\star p^\star}}{\Gamma(p^\star)}$$

$\Rightarrow h(\theta|\vec{x})$ entspricht der Dichte von $\gamma(b^\star, p^\star)$

$$\Rightarrow \hat{\theta} = \mathbb{E}(\theta) = \frac{p^\star}{b^\star}$$

Berechnung der A-Posteriori-Dichte (2)

Behauptung: $h(\theta|\vec{x})$ ist die Dichte von $\gamma(b^\star, p^\star)$ mit

$$p^\star = n + p$$
$$b^\star = b + \sum_{i=1}^{n} x_i$$

$$\int_\Theta D(\theta, \vec{x}) \cdot c^*(\vec{x}) d\theta = 1 = \int_\Theta \frac{b^{\star p^\star}}{\Gamma(p^\star)} \cdot D(\theta, \vec{x}) d\theta$$
$$\iff c^*(\vec{x}) = \frac{b^{\star p^\star}}{\Gamma(p^\star)}$$
$$\Rightarrow h(\theta|\vec{x}) \text{ entspricht der Dichte von } \gamma(b^\star, p^\star)$$
$$\Rightarrow \hat{\theta} = \mathbb{E}(\theta) = \frac{p^\star}{b^\star}$$

Berechnung der A-Posteriori-Dichte (2)

Behauptung: $h(\theta|\vec{x})$ ist die Dichte von $\gamma(b^\star, p^\star)$ mit

$$p^\star = n + p$$
$$b^\star = b + \sum_{i=1}^{n} x_i$$

$$\int_\Theta D(\theta, \vec{x}) \cdot c^*(\vec{x}) d\theta = 1 = \int_\Theta \frac{b^{\star p^\star}}{\Gamma(p^\star)} \cdot D(\theta, \vec{x}) d\theta$$
$$\iff c^*(\vec{x}) = \frac{b^{\star p^\star}}{\Gamma(p^\star)}$$
$$\Rightarrow h(\theta|\vec{x}) \text{ entspricht der Dichte von } \gamma(b^\star, p^\star)$$
$$\Rightarrow \hat{\theta} = \mathbb{E}(\theta) = \frac{p^\star}{b^\star}$$

Zur Erinnerung: Übergangskern des Metropolis-Hastings Algorithmus

Algorithmus

Der Übergangskern K von $(X_n)_{n\in\mathbb{N}}$ ist durch folgendes Schema charakterisiert:

1. *Erzeuge ausgehend von $X_n = \theta$ eine zufällige Realisation θ' einer Zufallsvariable gemäß der Vorschlagsdichte $q(\theta'|\theta)$*
2. *Berechne die Annahmewahrscheinlichkeit $\alpha(\theta, \theta')$ von θ'*

$$\alpha(\theta, \theta') = min\left(1, \frac{f(\theta')q(\theta|\theta')}{f(\theta)q(\theta'|\theta)}\right) \tag{5}$$

3. *Setze*

$$X_{n+1} = \begin{cases} \theta' & \text{mit Wahrscheinlichkeit } \alpha(\theta, \theta') \\ \theta & \text{mit Wahrscheinlichkeit } 1 - \alpha(\theta, \theta') \end{cases} \tag{6}$$

Ein MCMC-Regressionsmodell

Abbildung: Kursverlauf im Sinne der Random-Walk-Theorie

Ein MCMC-Regressionsmodell

Zielsetzung

- *Anwenden des Modells auf einen Aktienkursverlauf*
- *Schätzen der Modellparameter a, b, c und σ*

Lösungsskizze

Sei $\theta = (a, b, c, \sigma)^\intercal$ der Parametervektor und x ein Datensatz mit Kursdaten

1. *Leite Likelihood-Funktion $L(a, b, c, \sigma | x)$ aus dem Modell ab*
2. *Definiere eine A-Priori-Dichte $g(\theta)$*
3. *Konstruiere eine Markov-Kette $(X_n)_{n \in \mathbf{N}}$, die als stationäre Verteilung die A-Posteriori-Verteilung $h(\theta | x)$ besitzt*
4. *Erzeuge mithilfe von $(X_n)_{n \in \mathbf{N}}$ n Stichproben aus $h(\theta | x)$ und berechne $\hat{\theta}$ durch das arithmetische Mittel der Stichproben*

Ein MCMC Regressionsmodell

Likelihood-Funktion

$$L : \Theta \times D \longrightarrow [0,1], (\theta|x) \mapsto likelihood$$

mit $\Theta = \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+$ *und* $D = \mathbb{R}^{2\times d}$*, wobei* d *die Größe des Datensatzes angibt. Ein Datensatz* $x \in D$ *ist als* $2 \times d$ *Matrix* (x_{ij}) *zu verstehen. Dabei ist* x_{1j} *der Kurs, der zum Zeitpunkt* x_{2j} *gemessen wurde.*

$$L(a,b,c,\sigma|x) \overset{iid.}{=} \prod_{j=1}^{d} f_\sigma(x_{1j} - (a * x_{2j} + b + c * sin(x_{2j})) = likelihood$$

mit f_σ *als Dichtefunktion der Normalverteilung* $N(0,\sigma)$

Ein MCMC Regressionsmodell

A-Priori-Dichte

Annahme: Die vier Parameter a, b, c und σ sind unabhängig

$\Longrightarrow g(\theta) = g(a, b, c, \sigma) = g_1(a) * g_2(b) * g_3(c) * g_4(\sigma)$

- $g_1(a)$: *Gleichverteilung*
- $g_2(b)$: *Normalverteilung*
- $g_3(c)$: *Gleichverteilung*
- $g_4(\sigma)$: *Gleichverteilung*

Ein MCMC Regressionsmodell

A-Priori-Dichte

Markov-Kette $(X_n)_{n\in\mathbb{N}}$ *auf dem Zustandsraum* Θ *mit plausibler Ausgangskurfiguration* $X_0 \in \Theta$.
Der Übergangskern K *ist folgendermaßen charakterisiert:*

1. *Bestimme eine mögliche neue Parameterkonfiguration* θ' *in der nahen Umgebung der aktuellen Parameterkonfiguration* $X_n = \theta$, *anhand der Dichte einer vierdimensionalen Normalverteilung (symmetrische Vorschlagsdichte* $q(\theta'|\theta)$*)*
2. *Setzte*

$$X_{n+1} = \begin{cases} \theta' & \text{mit Wahrscheinlichkeit } \alpha(\theta,\theta') = \min\left(1, \frac{p(\theta',x)}{p(\theta_n,x)}\right) \\ \theta & \text{mit Wahrscheinlichkeit } 1-\alpha(\theta,\theta') \end{cases} \quad (7)$$

wobei $p(\theta, x) = L(\theta|x) * g(\theta)$

Ein MCMC Regressionsmodell

A-Priori-Dichte

Markov-Kette $(X_n)_{n\in\mathbb{N}}$ auf dem Zustandsraum Θ mit plausibler Ausgangskurfiguration $X_0 \in \Theta$.
Der Übergangskern K ist folgendermaßen charakterisiert:

1. *Bestimme eine mögliche neue Parameterkonfiguration θ' in der nahen Umgebung der aktuellen Parameterkonfiguration $X_n = \theta$, anhand der Dichte einer vierdimensionalen Normalverteilung (symmetrische Vorschlagsdichte $q(\theta'|\theta)$)*
2. *Setzte*

$$X_{n+1} = \begin{cases} \theta' & \textit{mit Wahrscheinlichkeit } \alpha(\theta, \theta') = \min\left(1, \frac{p(\theta', x)}{p(\theta_n, x)}\right) \\ \theta & \textit{mit Wahrscheinlichkeit } 1 - \alpha(\theta, \theta') \end{cases} \tag{7}$$

wobei $p(\theta, x) = L(\theta|x) * g(\theta)$

Ein MCMC Regressionsmodell

A-Priori-Dichte

Markov-Kette $(X_n)_{n\in\mathbb{N}}$ *auf dem Zustandsraum* Θ *mit plausibler Ausgangskurfiguration* $X_0 \in \Theta$.
Der Übergangskern K *ist folgendermaßen charakterisiert:*

1. *Bestimme eine mögliche neue Parameterkonfiguration* θ' *in der nahen Umgebung der aktuellen Parameterkonfiguration* $X_n = \theta$, *anhand der Dichte einer vierdimensionalen Normalverteilung (symmetrische Vorschlagsdichte* $q(\theta'|\theta)$*)*
2. *Setzte*

$$X_{n+1} = \begin{cases} \theta' & \text{mit Wahrscheinlichkeit } \alpha(\theta, \theta') = \min\left(1, \frac{p(\theta', x)}{p(\theta_n, x)}\right) \\ \theta & \text{mit Wahrscheinlichkeit } 1 - \alpha(\theta, \theta') \end{cases} \tag{7}$$

wobei $p(\theta, x) = L(\theta|x) * g(\theta)$

Ein MCMC Regressionsmodell

Simulation und Auswertung

- *Simuliere* $(X_n)_{n\in\mathbb{N}}$ *m Iterationen*
- $\hat{\theta} = \frac{1}{m}\sum_{n=1}^{m} X_n$