Randolf Altmeyer

BZQ II: Stochastikpraktikum
Wintersemester 2016
Humboldt-Universität zu Berlin

Projektaufgaben Block 4

1 Markovketten zur Modellierung von Krebswachstum (5P)

In einer Studie wurde anhand von Autopsiedaten ermittelt wie sich Lungenkrebs im Körper ausbreitet (Beispiel 2.12 im Buch von Dobrow). Für insgesamt 50 verschiedene Positionen im Körper wurden die Übergangswahrscheinlichkeiten einer Markovkette berechnet, die stets in Zustand 23 (=Lunge) startet. Lade die Übergangsmatrix vom Web über die Datei *cancer.csv*. Berechne für drei verschiedene Positionen $x \in \{1, \ldots, 50\}$ im Körper deiner Wahl die mittlere Anzahl von Schritten die der Krebs braucht, um dorthin zu gelangen (mittlere Eintrittszeit $\mathbb{E}[\tau_x]$, siehe Vorlesung). Verwende die beiden folgenden Methoden, um $\mathbb{E}[\tau_x]$ zu approximieren:

1. Es gilt laut Vorlesung $\mathbb{E}[\tau_x] = \sum_{k=0}^{\infty} k\mathbb{P}(\tau_x = k)$. Weiterhin gilt (muss nicht gezeigt werden), dass $\mathbb{P}(\tau_x = k) = \sum_{i,j \neq x} \tilde{P}_{ij}^{k-1} P_{jx} \mu(i)$, wobei μ die Anfangsverteilung und P die Übergangsmatrix ist. Außerdem ist $(\tilde{P}_{ij})$ die Übergangsmatrix P, wenn man die x-te Spalte und die x-te Zeile von P durch Nullen ersetzt, und $\tilde{P}^{k-1} = \tilde{P}^{k-2}\tilde{P}$.
2. Simuliere 10.000 Mal die Markovkette und bestimme $\mathbb{E}[\tau_x]$ durch Mittelwertbildung. Wähle dazu einen maximalen Zeithorizont und verwende nur Pfade bei denen der Zielzustand auch erreicht wird.

2 Poisson-Prozess (5P)

Lade die Tabelle *traffic.csv* vom Web in R. Sie enthält Verkehrsdaten von vorbeifahrenden Autos an einem Autobahnabschnitt in Abständen von 15 Minuten. Bestimme eine Intensitätsfunktion $\tilde{\lambda}(t)$ für $0 \leq t < 24$ als mittlere Anzahl an beobachteten Autos pro Minute, indem du $\tilde{\lambda}(t)$ stückweise konstant definierst auf den 96 möglichen 15-Minuten-Intervallen. Wir wollen einen Stresstest für den Autobahnabschnitt betrachten anhand der beobachteten Daten. Simuliere dazu einen Poisson-Prozess auf $[0, 24)$ mit gesteigerter Intensität $\lambda(t) = 1.2\tilde{\lambda}(t)$. Wiederhole die Simulation 10.000 Mal (d.h. für 10.000 Tage) und bearbeite die Fragen unten. Diskutiere die Ergebnisse kurz.

1. Wieviele Autos kommen insgesamt in 24 Stunden im Schnitt an dem Autobahnabschnitt vorbei?
2. In welcher Stunde (0 - 24) fährt im Schnitt das 1000. Auto vorbei? Verwende 24 in der Berechnung, wenn es insgesamt weniger als 1000 Autos gab an einem Tag.
3. Wieviele Autos fahren durchschnittlich zwischen 1:30 und 1:45 vorbei?

4. Bestimme für jede der 10.000 Simulationen die einstündige Periode $[t, t+1)$, wobei $t \in \{0, 0.25, 0.5, 0.75, \ldots, 23.0\}$, in der die meisten Autos vorbeifahren. Erzeuge ein Histogramm für diese Perioden.

5. Erzeuge ein Histogramm für die Anzahl der vorbeifahrenden Autos in dieser Periode gemäß den 10.000 Simulationen.

3 Der betrunkene Vogel (5P)

Sei $B = (B_t)_{t\geq 0}$ eine d-dimensionale Brownsche Bewegung, $d \geq 1$, d.h. $B_t = (B_t^{(1)}, \ldots, B_t^{(d)})$, wobei die $B_t^{(i)}$ unabhängige eindimensionale Brownsche Bewegungen sind, $i = 1, \ldots, d$. Sei $T_{R,x} = \inf\{t > 0 : \|B_t + x\| = R\}$, für $R > 0$ und $x \in \mathbb{R}^d$, der erste Zeitpunkt zu dem $B_t + x$ den Rand des Balls mit Radius R im $\mathbb{R}^d$ trifft. Man kann zeigen, dass folgender Satz gilt:

Theorem (Polyá)**.** *Sei $x \in \mathbb{R}^d$ mit $\|x\| \geq R$. Dann ist* $\mathbb{P}(T_{R,x} < \infty) = \left(\frac{R}{\|x\|}\right)^{d-2}$.

Illustriere mit Hilfe dieses Satzes und anhand von Simulationen die folgende Aussage von Polyá:

> Ein betrunkener Mensch findet stets zurück nach Hause, ein betrunkener Vogel aber vielleicht nicht.

Hinweis: Man kann zeigen, dass $\mathbb{P}(T_{R,x} < \infty) = \lim_{r\to\infty} \mathbb{P}(T_{R,x} < r)$. Wähle r ausreichend groß und simuliere B_t für $0 \leq t \leq r$ mit einem hinreichend kleinen Gitter.

4 Geometrische Brownsche Bewegung und Option pricing (5P)

1. Sei X die geometrische Brownsche Bewegung aus der Vorlesung für $\mu \in \mathbb{R}$ und $\sigma > 0$.
   (a) Berechne Erwartungswert und Varianz von X_t für $t \geq 0$ anhand der exakten Lösung.
   (b) Wie verhält sich X_t für $t \to \infty$, abhängig davon ob $\mu - \frac{\sigma^2}{2} > 0, = 0$ oder < 0 gilt? Beweise die Aussage mit Hilfe des Gesetzes vom iterierten Logarithmus aus der Vorlesung und erkläre das Ergebnis mit Hilfe von Simulationen für von dir gewählte Werte von μ und σ.

2. Sei X_t der Preis einer Aktie für $0 \leq t \leq 1$. Wir nehmen an X_t folgt einer geometrischen Brownschen Bewegung. Wir möchten den arbitrage-freien Preis einer europäischen Option mit Auszahlungsfunktion $f(x) = (x - K)_+$ bestimmen. Dieser ist laut Vorlesung gegeben durch $C = e^{-r}\mathbb{E}^*[(X_1 - K)_+]$, wobei X unter dem Maß $\mathbb{P}^*$ eine geometrische Brownsche Bewegung mit den Parametern $\mu = r$ und σ ist. Sei im Folgenden $r = 0.04$, $\sigma = 0.2$ und $X_0 = K = 100$.
   (a) Berechne die Lösung mittels der Black-Scholes-Formel aus der Vorlesung.
   (b) Berechne C durch Simulation von $(X_t)_{0\leq t\leq 1}$ mittels der exakten Lösung der geometrischen Brownschen Bewegung (d.h. wir müssen nur die Brownsche Bewegung simulieren) mit Hilfe von Monte-Carlo-Simulationen. Verwende die Parameter $\mu = r$ und σ.
   (c) Berechne C mit Hilfe des Euler-Schemas aus der Vorlesung für $(X_t)_{0\leq t\leq 1}$. Verwende die Parameter $\mu = r$ und σ.

Freiwillig (3 Extrapunkte):

Sei $X = (X_t)_{t \geq 0}$ ein stochastischer Prozess, der die stochastische Differentialgleichung $dX_t = (\mu - X_t)dt + \sigma dB_t$ erfüllt für eine Zufallsvariable X_0, $\mu \in \mathbb{R}$, $\sigma > 0$. X heißt dann *Ornstein-Uhlenbeck-Prozess*. Man kann zeigen, dass die exakte Lösung durch $X_t = X_0 e^{-t} + \mu(1 - e^{-t}) + \int_0^t \sigma e^{s-t} dB_t$ gegeben ist, wobei X_0 der Anfangswert ist.

1. Bestimme Erwartungswert und Varianz von X_t über die exakte Lösung. Stelle ein Vermutung für eine Grenzverteilung von X_t an für $t \to \infty$.
2. Simuliere und plotte Pfade bezüglich X_t für verschiedene deterministische Startwerte X_0 und $\mu = 1$, $\sigma = 1$. Beschreibe das Ergebnis. Interpretiere die Rolle von μ und σ.

Hinweise zur Abgabe:

Alle Dateien (pdf der Auswertung, R-Code für jede einzelne Aufgabe in eigener Datei, aber OHNE Daten!!) „gepackt“ (z.B. mit zip) mit dem Namen *UE3_ Student1_ Student2* bis zum 31.01. per Email an *altmeyrx@math.hu-berlin.de* schicken.