

Monte Carlo Methoden

Vorlesungsskript

Goethe-Universität Frankfurt

Sommersemester 2010

Thomas Gerstner

19. Juli 2010

Inhaltsverzeichnis

1	Einleitung	5
2	Allgemeines	7
2.1	Wahrscheinlichkeitsraum	7
2.2	Zufallszahlen	7
2.3	Verteilungen	8
2.4	Stichproben	9
2.5	Zentraler Grenzwertsatz, Gesetz der großen Zahlen	9
2.6	Stochastische Prozesse	10
3	Erzeugung von Zufallszahlen	11
3.1	Grundlagen	11
3.2	Kongruenzgeneratoren	12
4	Quasi-Zufallszahlen	19
4.1	Diskrepanz	19
4.2	Van der Corput-Folge	20
4.3	Halton-Folge	21
5	Allgemeine Verteilungen	23
5.1	Inversionsmethode	23
5.2	Generierung normalverteilter Zufallszahlen	24
5.3	Acceptance / Rejection Methode	25
5.4	Erzeugung von Zufallspfaden	26

6 Numerische Integration	29
6.1 Monte Carlo Integration	29
7 Varianzreduktion	33
Literaturverzeichnis	33

Kapitel 1

Einleitung

Monte Carlo Methoden sind eine Klasse von Algorithmen, die Zufallszahlen zur Berechnung des Resultats eines Problems berechnen. Die Zufallszahlen werden in Computersimulationen als Pseudo-Zufallszahlen generiert.

Beispiel 1.0.1 (Flächenberechnung)

Allg.:

Für ein Gebiet $\Omega \in \mathbb{R}^d$ und $x \in \mathbb{R}^d$ berechne

$$\int_{\Omega} 1 d\vec{x} = \int_{\Omega} \chi_{\Omega}(\vec{x}) d\vec{x} \approx \text{Vol}(H) \frac{M_N(\Omega)}{N}$$

für $H = [a_1, b_1] \times \dots \times [a_d, b_d] \supset \Omega$, $\chi_{\Omega}(\vec{x}) := \begin{cases} 1, & \text{falls } \vec{x} \in \Omega \\ 0, & \text{sonst} \end{cases}$ und

$M_N(\Omega) = \#\{x_i : x_i \in \Omega, i = 1, \dots, N\}$.

Bsp.:

- betrachte einen in ein Quadrat einbeschriebenen Kreis
- Fläche des Quadrats: 1
- Fläche des Kreises: $\pi/4$
- ziehe N Zufallszahlen in $[0, 1]^2$
- zähle die Punkte im Kreis: M_N
- \Rightarrow Näherung: $\pi_N = 4 \frac{M_N}{N}$
- für $N = 100$ und $M_N = 77$ ist $\pi_N = 3.08$

Beispiel 1.0.2 (Berechnung des Volumens eines Polyeders)

Halbebene: $a \circ x \geq b$ mit $a \in \mathbb{R}^d$, $b \in \mathbb{R}$ und Skalarprodukt „ \circ “.

Mehrere Halbenen: $a_i \circ x \geq b_i$ mit $a_i \in \mathbb{R}^d$, $b_i \in \mathbb{R}$ so dass $\Omega = \{x : Ax \geq b\}$ für $A = (a_{ij})_{ij}$, $b = (b_i)_i$, $j = 1, \dots, k$, $k \geq d$.

Frage: Wie bestimmt man $\int_{\Omega} 1 dx$?

Antwort: Man kann es mit Monte Carlo schätzen!

Wähle dazu \vec{x}_m zufällig in H (bounding box) und prüfe, ob \vec{x}_m innerhalb des Polyeders liegt. Berechne danach $\text{Vol}(H) \frac{M_N(\Omega)}{N}$.

Beispiel 1.0.3 (Numerische Integration)

Berechnung von bestimmten Integralen:

$\int_a^b f(x) dx \approx (b-a) \frac{1}{N} \sum_{i=1}^N f(x_i)$, x_i zufällig gleichverteilt in $[a, b]$, bzw.

$\int_{\Omega} f(\vec{x}) d\vec{x} \approx \text{Vol}(\Omega) \frac{1}{N} \sum_{i=1}^N f(\vec{x}_i)$, $\vec{x}_i \in \mathbb{R}$ zufällig gleichverteilt in Ω (z.B. H).

Beispiel 1.0.4 (Simulation)

Siehe Übungen Aufgabenblatt 1.

Beispiel 1.0.5 (Simulation von stochastischen Prozessen)

Modellierung von Wertpapierkursen im Black-Scholes-Modell

$$dS(t) = \mu S(t) dt + \sigma S(t) dW_t$$

mit Drift $\mu \in \mathbb{R}$, Volatilität $\sigma \in \mathbb{R}$ und Wiener Prozess $(W_t)_{t \geq 0}$. Die Lösung dieser stochastischen Differenzialgleichung lässt sich approximieren durch

$$S(t_i) = S(t_{i-1}) e^{(\mu - \frac{1}{2}\sigma^2)(t_i - t_{i-1})} e^{\sigma \sqrt{t_i - t_{i-1}} Z_i}, \quad Z_i \sim N(0, 1)$$

und wird bei der Optionspreisbewertung benutzt.

Beispiel 1.0.6 (Optimierung)

1. Problem des Handelsreisenden:

N Städte, $i = 1, \dots, N$ mit Abständen $d_{i,j}$ zwischen zwei Städten. Finde nun die Permutation (i_1, \dots, i_N) von $(1, \dots, N)$, so dass $(\sum_{j=1}^N d_{i_j, i_{j+1}}) + d_{i_N, i_1}$ minimal wird.

Eine Möglichkeit der Monte-Carlo-Optimierung dieses Problem zu lösen, wäre es M Touren zu würfeln und diejenige mit der kürzesten Länge zu wählen. Eine weitere Alternative wäre der Simulated Annealing Algorithmus.

Beispiel 1.0.7 (Integralgleichungen)

Kapitel 2

Allgemeines

2.1 Wahrscheinlichkeitsraum

Der Wahrscheinlichkeitsraum ist ein Tripel (Ω, Σ, P) . Die Ereignismenge Ω (meist überabzählbar, z.B. \mathbb{R}) entspricht dabei allen möglichen Ereignissen, wobei ein (Elementar-)Ereignis $\omega \in \Omega$ eine Teilmenge von Ω ist.

Die Ereignisalgebra (σ -Algebra) Σ ist eine Menge von Teilmengen von Ω und erfüllt die Bedingungen $\Omega \in \Sigma$, $A \in \Sigma \Rightarrow A^C \in \Sigma$ und $A_1, A_2, \dots \in \Sigma \Rightarrow \bigcup_{n \in \mathbb{N}} A_n \in \Sigma$, damit Ereignissen Wahrscheinlichkeiten zugeordnet werden können.

Für das Wahrscheinlichkeitsmaß $P : \Sigma \rightarrow [0, 1]$ gilt $P(\Omega) = 1$ und $P(\bigcup_{n \in \mathbb{N}} A_n) = \sum_{n \in \mathbb{N}} P(A_n)$ mit paarweisen disjunkten $A_n \in \Sigma$. Es ordnet jedem Ereignis $A \in \Sigma$ eine Wahrscheinlichkeit $P(A)$ zu.

Die Aussage A gilt „fast überall“ / „fast sicher“ bedeutet:

$\exists A \in \Sigma$ mit $P(A) = 0$ und A gilt für alle $\omega \in \Omega \setminus A$, d.h. die Aussage ist mit Wahrscheinlichkeit Null falsch.

2.2 Zufallszahlen

Reelle Zufallszahl/-größe/-variable:

messbare Abbildung $X : \Omega \rightarrow \mathbb{R}$ mit $\{X \leq x\} := \{\omega \in \Omega : X(\omega) \leq x\} = X^{-1}((-\infty, x])$ messbar für alle $x \in \mathbb{R}$, also $\{X \leq x\} \in \Sigma$.

(kumulative) Verteilungsfunktion $F(x)$ von X :

$F : \mathbb{R} \rightarrow [0, 1]$ mit $F(x) = P(X \leq x)$.

(Verteilungs-)Dichte von X :

$f(x) : \mathbb{R} \rightarrow [0, \infty)$ mit $P(a \leq x \leq b) = \int_a^b f(x) dx$ für $a < b$, bzw. $P(x \leq b) = F(b) =$

$\int_{-\infty}^b f(x)dx$, falls F absolutstetig und fast überall differenzierbar ist.

Erwartungswert:

$$\mu = E(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_{-\infty}^{\infty} x dF(x).$$

Varianz:

$$\sigma^2 = \text{Var}(X) = E((X - \mu)^2) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)dx = E(X^2) - \mu^2.$$

Standardabweichung:

$$\sigma = \sqrt{\text{Var}(X)}.$$

Eigenschaften:

$\alpha, \beta \in \mathbb{R}$, X, Y Zufallszahlen über dem gleichen Wahrscheinlichkeitsraum:

$$E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$$

$$\text{Var}(\alpha X + \beta Y) = \alpha^2 \text{Var}(X) + \beta^2 \text{Var}(Y) + 2\alpha\beta \text{Cov}(X, Y).$$

Kovarianz von X und Y :

$$\text{Cov}(X, Y) = E((X - E(X))(Y - E(Y))) = E(XY) - E(X)E(Y)$$

$$\Rightarrow \text{Var}(X \pm Y) = \text{Var}(X) + \text{Var}(Y) \pm 2\text{Cov}(X, Y)$$

Unabhängigkeit zweier Zufallszahlen X und Y :

$$P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$$

$$\Rightarrow E(XY) = E(X)E(Y)$$

$$\Rightarrow \text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) \tag{2.1}$$

Bedingte Wahrscheinlichkeit:

$$P(B | A) = \frac{P(A \cap B)}{P(A)}, \quad P(A) \neq 0.$$

2.3 Verteilungen

Definition 2.3.1 (Normalverteilung)

Dichte: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

$X \sim N(\mu, \sigma^2)$ heißt X ist normalverteilt mit Erwartungswert μ und Varianz σ^2 . Daraus folgt $Z = \frac{X-\mu}{\sigma} \sim N(0, 1)$ ist standard-normalverteilt. Die Werte der Verteilungsfunktion müssen numerisch berechnet werden.

Definition 2.3.2 (Multivariate Normalverteilung)

Dichte: $f(\vec{x}) = \frac{1}{(2\pi)^{d/2} \sqrt{\text{Det}(\Sigma)}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)}$ für $x \in \mathbb{R}^d$, Erwartungswertvektor $\mu \in \mathbb{R}^d$ und positiv definiten Kovarianzmatrix $\Sigma \in \mathbb{R}^{d \times d}$.

Definition 2.3.3 (Gleichverteilung)

eindimensional im Intervall $[a, b]$:

Dichte: $f(x) = \begin{cases} \frac{1}{b-a} & \text{für } a \leq x \leq b \\ 0, & \text{sonst} \end{cases}$
 $E(X) = \frac{1}{2}(a+b)$ und $\text{Var}(X) = \frac{1}{12}(b-a)^2$.

allgemein (multivariat):

Dichte: $f(x) = \begin{cases} \frac{1}{\text{Vol}(D)} & \text{für } x \in D \subset \mathbb{R}^d \\ 0, & \text{sonst} \end{cases}$

2.4 Stichproben

Gegeben M Zufallszahlen x_1, \dots, x_M (Realisierungen) ist das Stichprobenmittel (sample mean) definiert durch

$$\hat{\mu} = \frac{1}{M} \sum_{i=1}^M x_i$$

und die Stichprobenvarianz (sample variance) durch

$$\hat{s}^2 = \frac{1}{M-1} \sum_{i=1}^M (x_i - \mu)^2$$

und es gilt $E(\hat{\mu}) = \mu$ sowie $E(\hat{s}^2) = \sigma^2$ (Erwartungstreue Schätzung).

2.5 Zentraler Grenzwertsatz, Gesetz der großen Zahlen

Theorem 2.5.1 (Zentraler Grenzwertsatz)

Seien X_1, \dots, X_n u.i.v. (unabhängig, identisch verteilte) Zufallszahlen, $\mu = E(X_i)$, $S_n = \sum_{i=1}^n X_i$ und $\sigma^2 = E(X_i - \mu)^2$, dann gilt für jedes a

$$\lim_{n \rightarrow \infty} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq a\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^a e^{-\frac{z^2}{2}} dz = P(\xi \leq a), \quad \xi \sim N(0, 1).$$

Bemerkung: Spezialfall für die Binomialverteilung ist der Satz von Moivre-Laplace.

Schwaches Gesetz der großen Zahlen:

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{S_n}{n} - \mu \right| > \epsilon \right) = 0$$

Starkes Gesetz der großen Zahlen:

$$P \left(\lim_{n \rightarrow \infty} \frac{S_n}{n} = \mu \right) = 1$$

2.6 Stochastische Prozesse

Ein stochastischer Prozess ist eine Sammlung von Zufallszahlen $X(t) : \Omega \rightarrow Z$ für alle $t \in T$ (T : Indexmenge) und $X(t)$ Z -messbar, wobei (Ω, Σ, P) ein Wahrscheinlichkeitsraum ist und der Raum $(Z, \dot{\Sigma})$ mit der σ -Algebra $\dot{\Sigma}$ Z -messbar ist. Im zeitdiskreten Fall ist T abzählbar (z.B. $T \in \mathbb{N}_0$). Der Prozess ist wertediskret, wenn Z endlich oder abzählbar ist.

Definition 2.6.1 (Markov-Prozess) *Stochastischer Prozess mit*

$$P(X(t+s) = Y \mid X(u) = X(t) \forall u < t) = P(X(t+s) = Y \mid X(t) = X(t)), \forall s > 0.$$

Die bedingt Wahrscheinlichkeit für zukünftige Zustände hängt nur vom aktuellen Zustand ab.

Definition 2.6.2 (Wiener Prozess (Brownsche Bewegung)) *Zeit- und wertestetiger Markov-Prozess mit*

$$X(0) = 0$$

$X(t)$ ist fast sicher stetig

$$X(t) - X(s) \sim N(0, t - s) \text{ für } t > s$$

(im Computer sind stochastische Prozesse immer zeit- und wertediskret).

Kapitel 3

Erzeugung von Zufallszahlen

3.1 Grundlagen

Varianten zur Erzeugung von Zufallszahlen:

- Physikalische (Hardware-) Generatoren: radioaktiver Zerfall, thermisches Rauschen, elektrische Schwankungen
- Software Generatoren: Algorithmen, Pseudo-Zufallszahlen.

Vorgehen:

- Tabelle z.B. RAND-Corp. 1955 Buch mit 1 Mio. Zufallszahlen oder Marsaglia 1995 CD-Rom mit 4.8 Mrd. Zufallszahlen
- ad hoc Generierung.

Forderungen:

- Gleichverteilung (Bereich $[0, 1)$ als Gleitpunktzahl oder Bereich $[0, N)$ als ganze Zahl)
- Unvorhersagbarkeit (Konstruktionsmechanismus komplex)
- Reproduzierbarkeit (für Fehlersuche, Vergleich von Simulationen)
- Effizienz (geringer Speicherbedarf).

Erste Realisierungen:

- transzendente Zahlen, z.B. $\pi \rightarrow$ statistisch gleichverteilte Ziffernfolge, aber zu aufwändig zu bestimmen

- Middle-Square-Methode (1940, von Neumann, Metropolis): wähle eine 4-stellige Zahl
 → quadriere die Zahl (ergibt 8-stellige Zahl, falls nicht, von links mit Nullen auffüllen)
 → wähle die mittleren 4 Ziffern als neue Zahl

Beispiel 3.1.1 (Middle-Square-Methode)

$$X_1 = 7182 \rightarrow X_1^2 = 51581124$$

$$X_2 = 5811 \rightarrow X_2^2 = 33767721$$

$$X_3 = 7677 \text{ usw.}$$

Probleme:

- Periodizität, z.B. 8441, 2504, 2700, 2900, 4100, 8100, 6100, 2100, 4100, 8100, ...
- Iteration konvergiert, z.B. 7182, 5811, 7677, ..., 0012, 0001, 0000

Grundprinzip eines Zufallsgenerators:

M endlich Menge. Iteriere eine Funktion $f : M \rightarrow M$ (Generator) durch $x_{n+1} = f(x_n)$ (Zufallsfolge / Orbit) und bestimme die Folge $(x_n)_{n \in \mathbb{N}}$ mit Startwert (Seed) x_0 .

Allgemeiner: $x_{n+1} = R(x_n, x_{n-1}, \dots, x_{n-k})$, $n \in \mathbb{N}_0$ mit $R : M^{k+1} \rightarrow M$ und Startvektor (x_0, \dots, x_k) .

Eigenschaften:

- M endlich \Rightarrow nicht alle Folgenglieder können verschieden sein, d.h. $\exists k, l$ mit $x_k = x_l$ für $k > l$ mit $r = k - l \Rightarrow x_{n+r} = x_n \forall n \geq l$ also periodisch mit Periode r .
- Idealerweise ist $r = |M| \Rightarrow f$ ist surjektiv/bijektiv.

3.2 Kongruenzgeneratoren

Lineare Kongruenz-Generatoren (LCG) nach Lehmer (1948) basieren auf ganzen Zahlen und Restklassen modulo eines Moduls $m \in \mathbb{N}$, $m \geq 2$ sowie auf Division mit Rest, das heißt

$$\forall a \in \mathbb{Z}, m \in \mathbb{N} \exists \text{ eindeutige Zahlen } q, r \in \mathbb{Z} \text{ mit } a = mq + r \quad (0 \leq r \leq m - 1)$$

Schreibweise: $a \equiv r \pmod{m}$.

Restklassen $\mathbb{Z}_m := \{[0], [1], \dots, [m-1]\}$ mit $[i] := \{z \in \mathbb{Z}, m \text{ teilt } z - i\}$.

Addition: $i + j := k \in \{0, 1, \dots, m-1\}$ falls $i + j - k$ von m geteilt wird.

Multiplikation: $i \cdot j := k \in \{0, 1, \dots, m-1\}$ falls $i \cdot j - k$ von m geteilt wird.

Ist in modularer Arithmetik auf jedem Rechner implementiert, oft durch: $m = 2^e$, $e \in$

\mathbb{N} , $e = 32$ oder $e = 64$.

LCG: ist eine Funktion $f : M \rightarrow M$ mit

$$x_{n+1} = ax_n + b \bmod m \iff x_{n+1} = ax_n + b + mj, \quad j \in \mathbb{Z}$$

Rechenregeln:

$$\begin{aligned} (a + b) \bmod m &= (a \bmod m) + (b \bmod m) \\ (a \cdot b) \bmod m &= (a \bmod m) \cdot (b \bmod m) \\ (a \bmod m) \bmod m &= a \bmod m. \end{aligned} \tag{3.1}$$

Ziel:

Wähle $x_0, a, b, m \in \mathbb{N}_0$ so, dass x_i „gleichverteilt“ in $\{0, 1, \dots, m-1\}$.

Beispiel: $a = 1, b = 1, m = 10, x_0 = 7 \Rightarrow x_1 = 8, x_2 = 9, x_3 = 0, \dots$

Algorithm 3.2.1 (LCG)

Eingabe: Parameter a, b, m und Startwert $x_0 \in M = \{0, 1, \dots, m-1\}$.

Iterationsanzahl $k \ll m$,

für $i = 0, \dots, k$ berechne $x_{i+1} = ax_i + b \bmod m$.

Ausgabe: Pseudo-Zufallszahlen x_1, \dots, x_k .

Erzeugung von reellen Zufallszahlen in $[0, 1)$ durch $y_i = \frac{x_i}{m}$.

Wahl des Startwertes (seed) x_0 garantiert Reproduzierbarkeit.

Beispiel 3.2.2 (LCG)

$a = 1, b = 1, m = 10, x_0 = 5 \Rightarrow X = \{6, 7, 8, 9, 0, 1, \dots\} \Rightarrow$ maximale Periode

$a = 7, b = 7, m = 10, x_0 = 7 \Rightarrow X = \{6, 9, 0, 7, 6, 9, \dots\} \Rightarrow$ Periode = 4

Beispiel 3.2.3 (Praxisbeispiel)

a	b	m	Periode	Quelle
24298	99991	199017	?	TI 59
65539	0	2^31	2^29	RANDU
137	187	2^{16}	max	Knuth

Ziel: Wann ist die Periode maximal, das heißt, wann ist die Iterationsvorschrift $x_{i+1} = f(x_i)$ bijektiv?

Notwendige (nicht hinreichende Bedingung): a muss zu m teilerfremd sein.

Theorem 3.2.4 (Knuth) Mit $m, a, b \in \mathbb{Z}$, $m \geq 2$ betrachte die Abbildung $f : \{0, 1, \dots, m-1\} \ni x \rightarrow (ax + b \bmod m) \in \{0, 1, \dots, m-1\}$. Für beliebiges $x_0 \in \{0, 1, \dots, m-1\}$ sei die Folge $(x_n)_{n \in \mathbb{N}}$ definiert durch $x_{n+1} = f(x_n)$, $n \in \mathbb{N}_0$.

Die Folge ist genau dann periodisch mit maximaler Periodenlänge m , wenn die folgenden Bedingungen erfüllt sind:

- (a) b und m sind teilerfremd
 (b) $p \mid (a - 1)$ für alle Primteiler p von m ($a \mid b$: a ist Teiler von b)
 (c) $4 \mid (a - 1)$ falls $4 \mid m$

Lemma 3.2.5 Für $a \in \mathbb{Z}$ und $k \in \mathbb{N}_0$ setze $S_k(a) := 0$ falls $k = 0$ und $S_k(a) := 1 + a + \dots + a^{k-1}$ falls $k \geq 1$.

Dann gilt (in \mathbb{Z}):

$$S_{rk}(a) = S_r(a^k)S_k(a) \quad \forall r \in \mathbb{N}.$$

Beweis. (vollständige Induktion)

$r = 1$:

$$S_{rk}(a) = S_k(a) = S_1(a^k)S_k(a)$$

$r \rightarrow r + 1$:

$$\begin{aligned} S_{(r+1)k}(a) &= S_{rk+k}(a) = S_{rk}(a) + a^{rk} + \dots + a^{rk+k-1} \\ &= S_r(a^k)S_k(a) + a^{rk}S_k(a) \\ &= S_k(a)(S_r(a^k) + a^{rk}) \\ &= S_k(a)S_{r+1}(a^k) \end{aligned}$$

Lemma 3.2.6 Sei p eine Primzahl und a eine ganze Zahl mit $a \equiv 1 \pmod{p}$, bzw. $a \equiv 1 \pmod{4}$ falls $p = 2$, dann gilt für alle $n, k \in \mathbb{N}$:

- (a) $S_p(a^{p^n}) \equiv 0 \pmod{p}$, $S_p(a^{p^n}) \not\equiv 0 \pmod{p^2}$
 (b) $S_{p^n}(a) \equiv 0 \pmod{p^n}$, $S_{p^n}(a) \not\equiv 0 \pmod{p^{n+1}}$
 (c) $S_k(a) \equiv 0 \pmod{p^n} \iff p^n \mid k$

Beweis.

Zu (a):

Wähle $p > 2$. Mit der Voraussetzung ist $a = 1 + jp$ mit $j \in \mathbb{Z}$. Aus der Binomialformel folgt $a^t - 1 = tjp + cp^2$ mit einer Konstanten c , so dass

$$S_p(a^{p^n}) - p = \sum_{k=0}^{p-1} (a^{kp^n} - 1) = p^n jp \frac{p(p-1)}{2} + \hat{c}p^2$$

mit einer weiteren Konstanten \hat{c} .

Zu (b): Induktion nach n

$n = 1$: Fall $p = 2$ trivial, also wähle p ungerade. Nach Voraussetzung gilt $a = 1 + jp$ mit $j \in \mathbb{Z}$, daraus folgt $a^k = 1 + kjp \pmod{p^2}$, $k \in \mathbb{N}$, dann ist

$$S_p(a) - p = \sum_{k=0}^{p-1} (a^k - 1) = \sum_{k=0}^{p-1} kjp = jp \frac{p(p-1)}{2} = 0 \pmod{p^2}.$$

$n \rightarrow n + 1$: Ergibt sich aus Lemma 3.2.5 wegen

$$S_{p^{n+1}}(a) = S_p(a^{p^n})S_{p^n}(a)$$

Zu (c):

Sei p^m die höchste Potenz von p mit $p^m \mid k$, also $k = p^m l$, wobei $p \nmid l$. Mit Lemma 3.2.5 gilt: $S_k(a) = S_l(a^{p^m})S_{p^m}(a)$.

Es folgt: $S_l(a^{p^m}) = l \pmod p$ und $S_l(a^{p^m}) \neq 0 \pmod p$

$\Rightarrow p^n \mid S_k(a) \Leftrightarrow p^n \mid S_{p^m}(a)$ und $p^n \mid S_k(a) \Leftrightarrow n \leq m$.

Beweis von Satz 3.2.4.

Ist $m = \{p_1^{k_1}, \dots, p_r^{k_r}\}$ die Primfaktorzerlegung von m , so folgt mit dem chinesischen Restsatz, dass \mathbb{Z}_m isomorph zu $\mathbb{Z}_{p_1^{k_1}} \times \dots \times \mathbb{Z}_{p_r^{k_r}}$ ist. Der Beweis soll nun nur für den Fall, dass m eine Primzahlpotenz p^k ist, gezeigt werden. Es gilt:

$$x_{i+1} - x_i = f(x_i) - f(x_{i-1}) = a(x_i - x_{i-1}), \quad i = 1, 2, \dots$$

und daher

$$x_n - x_0 = \sum_{i=0}^{n-1} a^i (x_1 - x_0) = S_n(a)(x_1 - x_0) \pmod m.$$

„ \Leftarrow “

Voraussetzung: f erzeugt maximalen Zyklus.

Falls p nicht $a - 1$ teilt, sind m und $a - 1$ teilerfremd und die Gleichung $(a - 1)x = b$ ist in \mathbb{Z}_m lösbar. Damit besitzt f einen Fixpunkt und es kann kein Zyklus der Länge m existieren. \Rightarrow (b).

Es teile 4 die Zahl m , also $m = 2^k$ mit $k \geq 2$. Aus (b) ist bekannt, dass $2 \mid (a - 1)$, also ist a ungerade. Bleibt noch zu zeigen, dass $a = 3 \pmod 4$ nicht auftreten kann (Widerspruchsbeweis):

Angenommen $a = 3 \pmod 4$, dann wäre $S_2(a) = 1 + a = 0 \pmod 4$, so dass $S_{2i}(a) = S_i(a^2)S_2(a)$ und daraus folgt $S_{2i} = 0 \pmod 4$ für alle $i \in \mathbb{N}$. Weiter gilt $x_{2i} = x_0 \pmod 4$ und $x_{2i+1} = x_1 \pmod 4$. Damit ist kein maximaler Zyklus möglich für alle $i \in \mathbb{N}$. \Rightarrow (c)

In einem maximalen Zyklus kommt einmal das Element 0 vor. Setze o.B.d.A. $x_0 = 0$ und $x_1 = b$. Daraus folgt $x_n = S_n(a)b$ und $b \pmod m$ ist nicht invertierbar. Somit kann das Element 1 nicht im Zyklus auftreten. \Rightarrow (a)

„ \Rightarrow “

Es soll wieder $m = p^k$ sein. Ist $m=2$, dann ist $b = 1$ wegen (a) \Rightarrow Zyklus der Länge 2 existiert.

Im Fall $p = 2 \Rightarrow 4 \mid m \Rightarrow$ Voraussetzung zu Lemma 3.2.5.

Noch zu zeigen: für $x_0 = 0$ wird ein maximaler Zyklus erzeugt.

$x_n = S_n(a)b \pmod m$, da b invertierbar $\pmod m$. Weiter ist $x_n = x_0 = 0 \Leftrightarrow S_n(a) = 0 \pmod m$ und mit Lemma 3.2.5 gilt $m \mid n$ und somit die Behauptung.

Folgerungen:

- Falls $m = 2^k$ hat ein LCG volle Periode falls b ungerade und $a = 4n + 1$ für ein ganzzahliges n .
- Falls $b = 0$ und m Primzahl ist, hat der LCG volle Periode, falls $x_0 \neq 0$ ist und $a^{m-1} - 1$ ein Vielfaches von m sowie $a^j - 1$ kein Vielfaches von m für $j = 1, \dots, m-2$ ist. In diesem Fall ist a eine Primitivwurzel von m .

Hinweis zur Implementierung: Problem: ax_i ist nicht mehr als Gleitpunktzahl darstellbar.

Abhilfe:

- Wähle a, m , so dass kein Überlauf auftreten kann: im 64-Bit-Rechner $n = 2^{40}$, $a = 2^{23}$.
- Verwendung von Multi-Precision-Arithmetic (langsam).
- Stelle $a = 2^\alpha a_1 + a_2$ mit $a_1, a_2 \leq 2^\alpha$ dar und nutze $(ax_i) \bmod m = (a_1(2^\alpha x_i \bmod m) + a_2 x_i \bmod m) \bmod m$.

Beispiel: $\alpha = 16, m = 2^{31} - 1 \Rightarrow$ Zwischenergebnisse maximal 2^{47} , obwohl $ax_i \leq 2^{62}$.

Verteilungseigenschaften:

Seien x_1, x_2, x_3 gleichverteilte Zufallszahlen in $[0, 1)$, so dass (x_1, x_2) gleichverteilt in $[0, 1)^2$ und (x_1, x_2, x_3) gleichverteilt in $[0, 1)^3$ ist.

Gütekontrolle eines Generators durch Betrachtung der Verteilung konsekutiver Paare/Tripel/etc. von Zufallszahlen in $[0, 1)^\alpha$.

Empfehlungen von Knuth:

- Wähle m möglichst groß.
- Wähle $\sqrt{m} < a < m - \sqrt{m}$.
- Ziffernmuster von a sollte unregelmäßig sein.
- b sollte ungerade und nicht durch 5 teilbar sein.

Definition 3.2.7 (Inverse Kongruenzgeneratoren (ICG))

Konstruktionsvorschrift: $x_{i+1} = (ax_i + b) \bmod m$, wobei \bar{x}_i das multiplikative „Inverse“ Element zu x_i ist, das heißt $\bar{x}_i x_i = 1 \bmod m$.

Falls m Primzahl ist, dann besitzen alle Elemente ungleich Null ein Inverses (Geometrische Korrelation tritt nicht auf). Nachteil dieser Methode ist der Aufwand zur Berechnung der Inversen der Größenordnung $O(\log_2(m))$.

Definition 3.2.8 (Fibonacci-Generatoren)

Konstruktionsvorschrift: $x_{i+1} = (x_i + x_{i-1}) \bmod m$ und Startwerte x_0, x_1 .

Mit Verzögerung (lags): $x_{i+1} = (x_{i-k} + x_{i-1}) \bmod m$ für $k \leq l$, $i = l + 1, l + 2, \dots$ und Startvektor der Länge l (oder durch LCG erzeugt).

Beispiele

1. Rand 17: $x_i = (x_{i-5} + x_{i-17}) \bmod 2^{32}$ und 17 Startwerte, besitzt eine Periodenlänge von $2,8 \cdot 10^{14}$.

2. Multiplikativer Generator: $x_i = (x_{i-k}x_{i-l}) \bmod m$ hat für $m = 2^d$ die Periode der Länge $(2^{d-3})(2^l - 1)$ für $d > 3$.

Definition 3.2.9 (Feedback Shift-Register Generatoren)

Zufallsbit-Generator: $b_i = (a_1b_{i-1} + a_2b_{i-2} + \dots + a_kb_{i-k}) \bmod 2$ und $a_i, b_i \in \{0, 1\}$.

Konstruktion ganzer Zahlen: $x_k = \sum_{i=0}^{m-1} b_{bk_{m+i}} 2^i$.

Konstruktion reeller Zahlen in $[0, 1)$: $x_k = \sum_{i=1}^m b_{bk_{m+i}} 2^{-i}$.

Beispiel 3.2.10 (Mersenne-Twister)

Twisted generalized feedback shift register generator.

Selbst in 623 Dimensionen gleichverteilt mit Periodenlänge $2^{19937} - 1 \approx 41 \cdot 10^{6001}$.

Gütetests für Zufallszahlen-Generatoren Hyperflächen auf denen konsekutive Zufallszahlen in \mathbb{R}^+ liegen. Statistische Tests:

- Chi-Quadrat-Test (zähle Anzahl der Zufallszahlen pro Teilintervall)
- Kolmogorov-Smirnov-Test
- Run-Tests
- Poker-Tests

Kapitel 4

Quasi-Zufallszahlen

Ziel: Erzeugung von Punktemengen, die möglichst gute Gleichverteilungseigenschaften besitzen. „Zufälligkeit“ spielt keine Rolle.

4.1 Diskrepanz

Als Maß für die Gleichverteilung führt man den Begriff der Diskrepanz ein.

Definition 4.1.1 (Diskrepanz)

Sei $X = \{x_1, \dots, x_N\}$ eine endliche Menge von Punkten $x_i \in [0, 1]^d \forall i$, Q ein Hyperquader $Q \subset [0, 1]^d : Q = [a_1, b_1] \times \dots \times [a_d, b_d]$ mit $0 \leq a_i \leq b_i \leq 1 \forall i$. Die Diskrepanz von X ist dann definiert als

$$D_N(X) := \sup_{Q \subset [0,1]^d} \left| \frac{\#\{i \mid x_i \in Q\}}{N} - \text{vol}(Q) \right|.$$

Die Stern-Diskrepanz von X ist definiert als

$$D_N^*(X) := \sup_{Q^* \subset [0,1]^d} \left| \frac{\#\{i \mid x_i \in Q^*\}}{N} - \text{vol}(Q^*) \right|,$$

wobei Q^* ein Quader ist, der im Nullpunkt verankert ist, das heißt $Q^* = [0, b_1] \times \dots \times [0, b_d]$ für $0 < b_i \leq 1$.

Je gleichmäßiger die Punkte x_i im Raum verteilt sind, desto kleiner ist die Diskrepanz $D_N(X)$. Eine Folge von Punktmengen X_N heißt gleichverteilt, wenn gilt $\lim_{N \rightarrow \infty} D_N(X_N) = 0$.

Eigenschaften:

- $D_N, D_N^* \geq 0$
- $D_N^* \leq D_N \leq 2^d D_N^*$

- $D_N^* \geq \frac{1}{2N}$ für $d = 1$

Definition 4.1.2 (Niederdiskrepanz)

Eine Menge $X = \{x_1, \dots, x_N\}$ heißt *Niederdiskrepanz Folge (low discrepancy series)*, wenn gilt

$$D_N(X) \leq c_d \frac{(\ln N)^d}{N}$$

mit einer von N unabhängigen Konstanten $c_d \in [0, \infty)$.

4.2 Van der Corput-Folge

Definition 4.2.1 (Van der Corput-Folge)

Das P -te Folgenglied x_P der Van der Corput-Folge wird dadurch generiert, dass die Zahl P zur Basis p (p Primzahl) geschrieben wird, das heißt $P = \sum_{k=0}^j d_k p^k$, wobei die $d_k \in \{0, \dots, p-1\}$ die j Ziffern der Zahlendarstellung sind. Dann ist das Folgenglied x_P definiert als die radikal Inverse (also die Spiegelung am „Dezimal“-punkt) der Zahl P , das heißt $x_i = \sum_{k=0}^j d_k p^{-k-1}$.

Beispiel 4.2.2 (Van der Corput Folge)

$p = 2$:

i	Binärdarstellung	Spiegelung	x_i
0	000	0.000	0
1	001	0.100	1/2
2	010	0.010	1/4
3	011	0.110	3/4
4	100	0.001	1/8
5	101	0.101	5/8
6	110	0.011	3/8
7	111	0.111	7/8

$p = 3$:

i	x_i
0	0
1	1/3
2	2/3
3	1/9
4	4/9
5	7/9
6	2/9

$p = 10, i = 4711 \Rightarrow x_i = 0.1174$

4.3 Halton-Folge

Mehrdimensionale Verallgemeinerung der Van der Corput-Folge.

Definition 4.3.1 (Halton-Folge)

Das P -te Folgenglied der Halton-Folge ist:

$$\vec{x}_i = ((x_{P_1})_i, \dots, (x_{P_d})_i),$$

wobei die x_{P_j} Van der Corput-Folgen zur Basis p_j sind und P_1, P_2, \dots, P_d die ersten d Primzahlen.

Beispiel 4.3.2 (Halton-Folge)

$d = 2$, also Basis 2 und 3 auf den Achsen:

$(0, 0), (1/2, 1/3), (1/4, 2/3), (3/4, 1/9), (1/8, 4/9), \dots$

Weitere mehrdimensionale Niederdiskrepanzfolgen

- Sobol-Folge
- Faure-Folge
- Niederreiter-Folgen (Netze)

Kapitel 5

Allgemeine Verteilungen

Ziel:

Generiere Zufallszahlen bezüglich einer beliebigen Verteilung basierend auf gleichverteilten Zufallszahlen.

5.1 Inversionsmethode

Wiederholung: $F : \mathbb{R} \rightarrow \mathbb{R}$ kumulierte Verteilungsfunktion $F(x) = P(X \leq x)$. F rechtsstetig, wenn $\lim_{x \searrow x_0} F(x) = F(x_0) \forall x_0 \in \mathbb{R}$.

Definition 5.1.1 Die verallgemeinerte Umkehrfunktion einer rechtsstetigen Funktion F ist definiert als $F^{-1}(y) = \inf \{u \in \mathbb{R} \mid F(u) \geq y\}$, $y \in \mathbb{R}$ sowie $F^{-1} = \infty$, falls $\{u \in \mathbb{R} \mid F(u) \geq 1\} \neq \emptyset$.

Theorem 5.1.2 Sei $F : \mathbb{R} \rightarrow \mathbb{R}$ eine Verteilungsfunktion einer Zufallsgröße X auf dem Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) . Sei U eine auf $[0, 1]$ gleichverteilte Zufallsgröße, also $U \sim U[0, 1]$. Dann ist $Y := F^{-1} \circ U$ eine reelle Zufallsgröße, die F als Verteilungsfunktion besitzt.

Beweis. Wir haben für $z \in [0, 1], x \in \mathbb{R}$ die Äquivalenz

a) $(F^{-1} \circ U)(z) \leq x$

b) $U(z) \leq F(x)$

a) \Rightarrow b): Sei $\epsilon > 0$. Dazu gibt es ein $u \in \mathbb{R}$ mit $u \leq x + \epsilon$ und $F(x + \epsilon) \geq F(u) \geq U(z)$. Aus der Rechtsstetigkeit von F folgt wegen $\epsilon > 0$ beliebig, dass $F(x) \geq U(z)$.

b) \Rightarrow a): Folgt, da $x \in \{u \in \mathbb{R} \mid F(u) \geq U(z)\}$. Damit gilt für $x \in \mathbb{R} : P(\{Y \leq x\}) = P(\{F^{-1} \circ U \leq x\}) = P(\{U \leq F(x)\}) = F(x)$.

Beispiel 5.1.3 Sei $F(x) = 1 - e^{-\lambda x}$, $x \geq 0$.

Erzeuge exponentialverteilte Zufallszahlen mit Hilfe von $F^{-1}(y) = -\frac{1}{\lambda} \ln(1 - y)$, $y \in (0, 1]$.

Da $1 - y$ ebenso wie y im Intervall $[0, 1]$ gleichverteilt ist, folgt dann, dass $x = F^{-1}(y) = -\frac{1}{\lambda} \ln y$ exponentialverteilt ist.

Falls F^{-1} nicht explizit bekannt ist, kann man mit Nullstellenverfahren die Inverse F^{-1} näherungsweise berechnen, das heißt finde x , so dass $F(x) = y$ bzw. $F(x) - y = 0$.

5.2 Generierung normalverteilter Zufallszahlen

Naiver Ansatz: Nutze den zentralen Grenzwertsatz.

Theorem 5.2.1 Sei $(X_i)_{i \in \mathbb{N}}$ eine Folge von identisch verteilten, unabhängigen Zufallsgrößen auf dem W -Raum (Ω, F, P) mit $E(X_i) = \mu$ und $\text{Var}(X_i) = \sigma^2$. Für den Mittelwert $\bar{X}_n = \frac{1}{n}(X_1 + \dots + X_n)$ gilt dann

$$\lim_{n \rightarrow \infty} P \left(\left\{ \frac{\bar{X}_n - \mu}{\sigma \sqrt{n}} \leq x \right\} \right) = N(x) \quad \forall x \in \mathbb{R}.$$

Anmerkung:

Existiert das dritte zentrale Moment $E((X_i - \mu)^3)$ und ist es endlich, dann ist die Konvergenz sogar gleichmäßig und die Konvergenzgeschwindigkeit ist wenigstens $O(\frac{1}{\sqrt{n}})$ (Satz von Berry-Esseen, siehe Übung).

Anwendung:

Erzeuge n $[0, 1]$ -gleichverteilte Zufallszahlen x_1, \dots, x_n , dann ist $\bar{x}_n = \frac{1}{n}(x_1 + \dots + x_n)$ näherungsweise normalverteilt. Im Falle von $n = 12$ mit Mittelwert $1/2$ und Varianz 1 . Dieses Verfahren ist jedoch ineffizient. Besser ist das **Box-Muller-Verfahren**:

Idee:

Verwende zwei gleichverteilte Zufallszahlen um zwei normalverteilte Zufallszahlen zu erzeugen.

Betrachte dazu die Transformation

$$\vec{y} = G(\vec{x}), \quad (y_1, y_2) = \left(\sqrt{-2 \ln(x_1)} \cos(2\pi x_2), \sqrt{-2 \ln(x_1)} \sin(2\pi x_2) \right) \quad \text{mit } x_1, x_2 \in M = (0, 1)^2.$$

Auflösung nach x_1, x_2 ergibt wegen $y_1^2 + y_2^2 = -2 \ln(x_1)$ und $\frac{y_2}{y_1} = \tan(2\pi x_2)$

$$\begin{aligned} x_1 &= \exp\left(-\frac{|y|^2}{2}\right) \\ x_2 &= \frac{1}{2\pi} \arctan\left(\frac{y_2}{y_1}\right), \end{aligned}$$

so dass für die Umkehrabbildung $H := G^{-1}$ gilt

$$H(y) = \left(\exp\left(-\frac{|y|^2}{2}\right), \frac{1}{2\pi} \arctan\left(\frac{y_2}{y_1}\right) \right) \quad \text{für } (y_1, y_2) \in \mathbb{R}^2.$$

Anmerkung:

Betrachtet man x_1, x_2 als gleichverteilte Zufallsvariablen auf $[0, 1]^2$, so wird die Variable \bar{x} exponentialverteilt auf $[0, \infty)$ mit Erwartungswert 2, denn $P(\{R^2 \leq x\}) = 1 - e^{-\frac{x}{2}}$ für $x \geq 0$. Ist nun R gegeben, dann sind die Punkte (y_1, y_2) gleichverteilt auf dem Kreis mit Radius R . Es gilt somit für $Y = G \circ X$, dass die Dichte $g(y)$ gegeben ist durch

$$g(y) = -\frac{1}{2\pi} \exp\left(-\frac{y_1^2}{2}\right) \frac{1}{2\pi} \exp\left(-\frac{y_2^2}{2}\right).$$

Algorithm 5.2.2 (Box-Muller Methode)

Eingabe: Zwei $[0, 1]$ -gleichverteilte Zufallsgrößen U_1 und U_2 .

Setze:

$$\begin{aligned}\theta &= 2\pi U_2 \\ \rho &= \sqrt{-2 \ln(U_1)} \\ Z_1 &= \rho \cos(\theta) \\ Z_2 &= \rho \sin(\theta)\end{aligned}$$

Ausgabe: Z_1, Z_2 standardnormalverteilte Zufallszahlen.

Wie werden Zufallszahlen für allgemeine Verteilungen erzeugt?

Allgemein: Inversionsmethode

Speziell Normalverteilung: Box-Muller-Verfahren

Alternative: Polar-Marsaglia (vermeidet die Auswertung trigonometrischer Funktionen)

5.3 Acceptance / Rejection Methode

Idee: Approximation einer komplexen Verteilungsfunktion G durch eine einfache Verteilungsfunktion F , so dass für die Dichten g und f gilt: $g(x) \leq cf(x)$.

Akkzeptanzwahrscheinlichkeit: $a(x) = \frac{g(x)}{cf(x)} \in [0, 1]$ für $g(x), f(x) > 0$.

Algorithm 5.3.1 (Acceptance-Rejection Methode)

Eingabe: Wahrscheinlichkeitsdichten f, g mit Träger in (x_1, x_2) , Schranke $c \in [1, \infty)$ und N (Anzahl der zu erzeugenden Zufallszahlen).

Ausgabe: Zufallszahlen x_1, x_2, \dots, x_N , die nach der Dichte g verteilt sind.

Vorgehensweise:

0. $k = 1$

1. Erzeuge eine Zufallszahl $x \in (x_1, x_2)$ gemäß der Dichte f .

2. Erzeuge eine $U(0, 1)$ -gleichverteilte Zufallszahl U .

3. Ist $U \leq a(x) = \frac{g(x)}{cf(x)}$ akzeptiere x und setze $x_k = x$, $k = k + 1$, ansonsten verwerfe x .

4. Wiederhole Schritte 1-3 solange bis $k = N$.

Analyse:

$X \sim F$, $U \sim U(0, 1)$. Y Ergebnis von Algorithmus 5.3.1 bedingt auf $U \leq a(x) = \frac{g(x)}{cf(x)}$.

Betrachte Verteilung von Y :

Ereignisse $A := \{X \leq x\}$, $x \in \mathbb{R}$ und $B := \left\{U \leq \frac{g(x)}{cf(x)}\right\}$.

zu zeigen: $P(A | B) = \frac{P(A \cap B)}{P(B)} = \int_{-\infty}^x g(w)dw$.

$$\begin{aligned}
 P(B) &= P\left(\left\{U \leq \frac{g(x)}{cf(x)}\right\}\right) \\
 &= \int_{-\infty}^{\infty} P\left(\left\{U \leq \frac{g(x)}{cf(x)}\right\} \mid X = w\right) f(w)dw \\
 &= \int_{-\infty}^{\infty} \frac{g(x)}{cf(x)} f(w)dw = \frac{1}{c} \\
 P(A \cap B) &= P\left(\left\{U \leq \frac{g(x)}{cf(x)}, X \leq x\right\}\right) \\
 &= \int_{-\infty}^{\infty} P\left(\left\{U \leq \frac{g(x)}{cf(x)}, X \leq x\right\} \mid X = w\right) f(w)dw \\
 &= \frac{1}{c} \int_{-\infty}^x g(w)dw
 \end{aligned}$$

Damit hat Y die Dichte g .

5.4 Erzeugung von Zufallspfaden

Betrachte die stochastische Differentialgleichung (SDE)

$$dX_t = f(t, X_t)dt + g(t, X_t)dW_t.$$

Dies ist die symbolische Schreibweise für folgende Integralgleichung

$$X_t = X_0 + \int_0^t f(s, X_s)ds + \int_0^t g(s, X_s)dW_s.$$

Diskretisiere die SDE mit dem Euler-Maruyama-Verfahren.

Rückblick: Euler-Verfahren für gewöhnliche DGLen:

Betrachte: $u'(x) = g(x)$, $x \in [0, T]$

Anfangswertproblem: $u(0) = u_0$

Diskretisierung mit Schrittweite $\Delta x = \frac{T}{N}$ und N Schritten. Betrachte die DGL nur an den Gitterpunkten $x_i = i\Delta x$, $i = 0, \dots, N$.

Euler-Verfahren:

$$\begin{aligned} \frac{u(x_i + \Delta x) - u(x_i)}{\Delta x} &= g(u(x_i), x_i) \\ \Leftrightarrow u(x_i + \Delta x) &= \Delta x g(u(x_i), x_i) + u(x_i). \end{aligned} \tag{5.1}$$

Euler-Maruyama-Verfahren:

$$\begin{aligned} x_{t_{n+1}} &= x_{t_n} + \int_{t_n}^{t_{n+1}} f(t, X_t) dt + \int_{t_n}^{t_{n+1}} g(t, X_t) dW_t \\ &\approx x_{t_n} + \int_{t_n}^{t_{n+1}} f(t_n, X_{t_n}) dt + \int_{t_n}^{t_{n+1}} g(t_n, X_{t_n}) dW_t \\ &= x_{t_n} + f(t_n, X_{t_n}) \int_{t_n}^{t_{n+1}} 1 dt + g(t_n, X_{t_n}) \int_{t_n}^{t_{n+1}} 1 dW_t \\ \Rightarrow \tilde{x}_{t_{n+1}} &= \tilde{x}_{t_n} + f(t_n, \tilde{x}_{t_n}) \Delta t + g(t_n, \tilde{x}_{t_n}) \Delta W_n, \end{aligned} \tag{5.2}$$

für $\Delta t = T/N$ und $\Delta W_n = W_{t_{n+1}} - W_{t_n} \sim N(0, \Delta t)$.

Kapitel 6

Numerische Integration

6.1 Monte Carlo Integration

Problem: Berechne $\int_a^b f(x)dx$.

Bei unbegrenzten Integrationsgebieten

- Abschneiden: $\int_{x_{\min}}^{x_{\max}} f(x)dx$

- Transformation

Beispiel: $\int_0^\infty \ln(1+x^2)e^{-x}dx$

Transformation: $t = 1 - e^{-x}$

$\Rightarrow I = \int_0^1 \ln(1 + \ln(1-t)^2)(1-t)dt$.

Achtung:

Transformation von einem unbegrenzten Gebiet in ein begrenztes Gebiet führt in der Regel zu Singularitäten an den Gebietsgrenzen.

Standardproblem:

$I f = \int_0^1 f(x)dx$ bzw.

$I f = \int_{[0,1]^d} f(\vec{x})d\vec{x} = \int_0^1 \dots \int_0^1 f(x_1, x_2, \dots, x_d)dx_1 dx_2 \dots dx_d$.

Genereller Quadraturansatz:

Quadraturformel: $Q_n f = \sum_{i=1}^n w_i f(\vec{x}_i)$ mit $\sum_{i=1}^n w_i = 1$.

Monte Carlo Ansatz: Gewichte fest, Stützstellen zufällig

$Q_n f = \sum_{i=1}^n \frac{1}{n} f(\vec{x}_i) = \frac{1}{n} \sum_{i=1}^n f(\vec{x}_i)$ mit $\vec{x}_i \sim U([0, 1]^d)$.

Quadraturfehler:

$$E_n f = |If - Q_n f|$$

Für Aussagen über die Güte der Approximation kann man das *Schwache Gesetz der großen Zahlen* zur Hilfe nehmen:

Für jedes $\epsilon > 0$ gilt: $\lim_{n \rightarrow \infty} P(\{|Q_n f - If| \geq \epsilon\}) = 0$, falls das Integral If existiert.

Oder den *Zentralen Grenzwertsatz*:

Existieren $\theta = E(f(X)) = If$ für $X \sim U([0, 1]^d)$ und $\sigma^2 = \text{Var}(f(X))$, dann ist $S_n = f(X_1) + \dots + f(X_n)$ approximativ normalverteilt $\sim N(n\theta, n\sigma^2)$, also

$$\lim_{n \rightarrow \infty} \left(\left\{ \frac{S_n/n - If}{\sigma\sqrt{n}} \leq x \right\} \right) = N(x), \quad x \in \mathbb{R}.$$

Definition 6.1.1 (Konvergenzgeschwindigkeit)

Sei $p \in [0, 1]$ und $n \in \mathbb{N}$. Ein Intervall der Form $[\theta - \epsilon, \theta + \epsilon]$ heißt *Konfidenzintervall* der Monte Carlo Schätzung $Q_n f$, falls gilt $P(\frac{1}{n} \sum_{i=1}^n f(x_i) \in I) = p$.

Theorem 6.1.2

Sei $p \in [0, 1]$, dann existiert ein $k > 0$ und eine Folge $(p_n)_{n \in \mathbb{N}}$ mit $\lim_{n \rightarrow \infty} p_n = p$, so dass die p_n -Konfidenzintervalle I_n von der Form $I_n = [\theta - \frac{k\sigma}{\sqrt{n}}, \theta + \frac{k\sigma}{\sqrt{n}}]$ sind.

Beweis. Wähle $x \in \mathbb{R}$, so dass mit der Standardnormalverteilung Φ gilt $\Phi(x) - \Phi(-x) = p$. Mit dem zentralen Grenzwertsatz existiert eine Folge δ_n^\pm mit $\lim_{n \rightarrow \infty} \delta_n^\pm = 0$, so dass

$$\left| P \left(\left\{ \frac{S_n - n\theta}{\sigma\sqrt{n}} \leq \pm x \right\} \right) - \Phi(\pm x) \right| = \delta_n^\pm.$$

Daraus folgt

$$P \left(\left\{ \left| \frac{S_n - n\theta}{\sigma\sqrt{n}} \right| \leq x \right\} \right) = \Phi(x) - \Phi(-x) + \delta_n^+ - \delta_n^- = p + \delta_n^+ - \delta_n^- =: p_n.$$

Es gilt somit $\lim_{n \rightarrow \infty} p_n = p$ und $\left| \frac{S_n - n\theta}{\sigma\sqrt{n}} \right| \leq x$ genau dann, wenn $\frac{1}{n} S_n \in [\theta - \frac{k\sigma}{\sqrt{n}}, \theta + \frac{k\sigma}{\sqrt{n}}]$ für $k > x$.

Kernaussage: Die Breite der Konfidenzintervalle schrumpft mit n wie $\frac{1}{\sqrt{n}}$.

Theorem 6.1.3 Sei $f : [0, 1]^d \rightarrow \mathbb{R}$ Lebesgue-integrierbar, λ_d das Lebesgue-Maß in \mathbb{R}^d und es gelte

$$\sigma_f^2 = \int_{[0,1]^d} \left(f(x) - \int_{[0,1]^d} f(u) du \right)^2 < \infty, \quad \text{endliche Varianz}$$

dann gilt:

(a) $\lim_{n \rightarrow \infty} Q_n f = \int_{[0,1]^d} f(x) dx$ λ_d -fast sicher

(b) $\lim_{n \rightarrow \infty} \lambda_d \left(\frac{\sigma_f}{\sqrt{n}} a < E_n f < \frac{\sigma_f}{\sqrt{n}} b \right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{z^2}{2}} dt$

(c) $|If - Q_n f| \leq \frac{\sigma_f}{\sqrt{n}}$ für $n \in \mathbb{N}$.

Beweis.

(a) folgt aus dem Gesetz der großen Zahlen.

(b) folgt aus dem zentralen Grenzwertsatz.

(c) Beweis für $d = 1$:

$$\begin{aligned}
 E(|If - Q_n f|)^2 &= E\left(If - \frac{1}{n} \sum_{i=1}^n f(x_i)\right)^2 \\
 &= E\left(\frac{1}{n} \sum_{i=1}^n (If - f(x_i))\right)^2 \\
 &\leq \frac{1}{n^2} \int_0^1 \dots \int_0^1 \left(\sum_{i=1}^n (If - f(x_i))\right)^2 dx_1 \dots dx_n \\
 &= \frac{1}{n^2} \int_{[0,1]^d} \left(\sum_{i=1}^n (If - f(x_i))^2 + 2 \sum_{i < j} (If - f(x_i))(If - f(x_j))\right) dx_1 \dots dx_n \\
 &= \frac{1}{n^2} \sum_{i=1}^n \int_0^1 (If - f(x_i))^2 dx_i + \frac{2}{n^2} \sum_{i < j} \int_0^1 (If - f(x_i)) dx_i \int_0^1 (If - f(x_j)) dx_j \\
 &= \frac{1}{n} \int_0^1 (If - f(x))^2 dx \\
 &= \frac{\sigma_f^2}{n}
 \end{aligned}$$

Anmerkungen:

Konvergenzaussagen sind probabilistischer Natur, das heißt es kann sein, dass der berechnete Integralwert stark vom wahren Wert abweicht.

Kapitel 7

Varianzreduktion

Siehe Skript von Prof. Baumeister, Kapitel 4:

http://www.math.uni-frankfurt.de/~numerik/lehre/Vorlesungen/Comp_Fin09/

Literaturverzeichnis