

Random number generation

von Daniel Safonov

30. Oktober 2022

1 Einleitung

Zufallszahlen stellen Ausgänge von Zufallsexperimenten dar. Unter den Zufallszahlen unterscheidet man zwischen „echten“ Zufallszahlen, welche durch physikalische Zufallszahlengeneratoren wie den Münzwurf erzeugt werden, und „Pseudo-“ Zufallszahlen, welche durch feste, wiederholbare Verfahren erzeugt werden. Pseudozufallszahlengeneratoren (PRNG) erzeugen Zahlenfolgen, die scheinbar zufällig sind, jedoch aufgrund ihrer Erzeugung durch deterministische Algorithmen berechnet wurden. Bei jeder Berechnung eines Pseudozufallszahlengenerators mit gleichem Startwert, auch seed genannt, wird die selbe Zahlenfolge erzeugt. Dadurch kann bei der Beobachtung einer hohen Anzahl von Pseudozufallszahlen diese Folge reproduziert werden. Im Umkehrschluss kann ein „guter“ Pseudozufallszahlengenerator zum Beispiel an der Periodenlänge der Zahlenfolge bestimmt werden. Im folgenden Abschnitt betrachten wir wie man aus Zufallszahlen einer gegebenen Wahrscheinlichkeitsverteilung Zufallszahlen einer gewünschten Wahrscheinlichkeitsverteilung erzeugt. Hierzu betrachten wir beispielhaft den Tausch der Variablen, das rejection sampling und eine Methode um Stichproben der mehrdimensionalen Normalverteilung zu erzeugen. Im zweiten Abschnitt betrachten wir die Monte Carlo Methode, wessen Basis die Zufallszahlen darstellen.

Definitionen:

Zufallsvariable:

Sei $(\Omega, \Sigma, \mathcal{P})$ ein Wahrscheinlichkeitsraum und sei $(\bar{\Omega}, \bar{\Sigma})$ ein Messraum, so ist X eine $\bar{\Omega}$ -Zufallsvariable, falls X eine messbare Funktion von (Ω, Σ) nach $(\bar{\Omega}, \bar{\Sigma})$ ist.

unabhängige Zufallsvariablen:

Sei $(\Omega, \Sigma, \mathcal{P})$ ein Wahrscheinlichkeitsraum, seien $(\bar{\Omega}_1, \bar{\Sigma}_1)$, $(\bar{\Omega}_2, \bar{\Sigma}_2)$ zwei Messräume und gegeben seien zwei Zufallsvariablen $X_i : (\Omega, \Sigma) \rightarrow (\bar{\Omega}_i, \bar{\Sigma}_i)$, $i = 1, 2$.

Die Zufallsvariablen X_1, X_2 heißen unabhängig, wenn für jedes $A_1 \in \bar{\Sigma}_1$ und $A_2 \in \bar{\Sigma}_2$ gilt, dass

$$\mathcal{P}(X_1 \in A_1, X_2 \in A_2) = \mathcal{P}(X_1 \in A_1) \cdot \mathcal{P}(X_2 \in A_2) \quad .$$

Gaußsche Fehlerfunktion:

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

Insbesondere gilt : $\mathcal{P}(|X - \mu| \leq R) = \operatorname{erf}(\frac{R}{\sqrt{2}\sigma})$ für $X \sim N(\mu, \sigma^2)$

Beispiele von Wahrscheinlichkeitsdichtefunktionen:

$$\text{Uniformverteilung } U(a, b): \quad f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq x \leq b \\ 0 & \text{for } x < a \text{ or } x > b \end{cases}$$

$$\text{Normalverteilung } N(\mu, \sigma^2): \quad f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp(-\frac{1}{2}(\frac{x-\mu}{\sigma})^2)$$

$$\text{Cauchyverteilung } Cauchy(t, s): \quad f(x) = \frac{s}{\pi} \cdot (s^2 + (x - t)^2)^{-1}$$

Notation: Sei die Zufallsvariable X z.B. uniformverteilt, schreibe $X \sim U(a, b)$.

2 Methoden der Zufallszahlengenerierung

2.1 Tausch von Variablen

Beispiel:

Sei $B = \{x \in \mathbb{R}^2 : \|x\|_2 \leq 1\}$, wir wollen nun Stichproben $X \sim U(B)$ erzeugen. Hierfür schreiben wir $X = [x, y]^T = [R \cos \theta, R \sin \theta]^T$ in Polarkoordinaten. Die Wahrscheinlichkeitsverteilung auf dem Einheitskreis ist für alle Winkel θ uniform, dh. $\theta \sim U(0, 2\pi)$ und die Wahrscheinlichkeitsdichtefunktion der radialabhängigen Variable R ist proportional zum Umfang des jeweiligen Kreises, sodass $p_R(r) = 2r, r \leq 1$.

Um nun den Variablentausch durchzuführen nutzen wir den folgenden Satz:

Sei $Y : \Omega \rightarrow \mathbb{R}$ eine reelwertige Zufallsvariable und $g : (a, b) \rightarrow \mathbb{R}$ stetig differenzierbar. Ist p_Y die Wahrscheinlichkeitsdichtefunktion von Y , so ist die Wahrscheinlichkeitsdichtefunktion der Zufallsvariable $X : \Omega \rightarrow (a, b)$, $X := g^{-1}(Y)$ gegeben durch

$$p_X(x) = p_Y(g(x)) \cdot |g'(x)|, \quad a < x < b \quad .$$

Sei nun $S \sim U(0, 1)$, $g(r) = r^2$, so hat $R = g^{-1}(S)$ die Wahrscheinlichkeitsdichtefunktion $p_R(r) = 1 \cdot g'(r) = 1 \cdot 2r$ und kann erzeugt werden als $R = S^{1/2}$ für $S \sim U(0, 1)$.

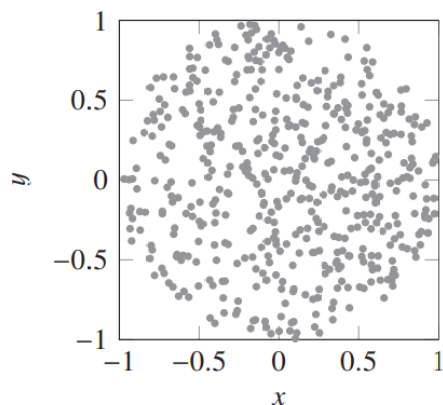


Abbildung 1: 500 unabhängige Stichproben von $U(B)$ aus Stichproben von $U(0,1)$

Beispiel (Cauchy Verteilung):

In diesem Beispiel wollen wir Stichproben X von $Cauchy(0, 1)$ erzeugen, wobei die Wahrscheinlichkeitsdichtefunktion der Cauchyverteilung gegeben ist durch $p_X(x) = (\pi(1 + x^2))^{-1}$. Setze nun $X = \tan(\theta)$, $\theta \in (0, \pi)$. Dann gilt nach obigen Satz des Variablentauschs mit $g(\theta) = \arctan(\theta)$, $g'(\theta) = (1 + \theta^2)^{-1}$, dass $X \sim Cauchy(0, 1)$ falls $\theta \sim U(0, \pi)$. Bedeutet zunächst muss eine Stichprobe $\theta \sim U(0, \pi)$ erzeugt werden und dann mittels $X = \tan(\theta)$ transformiert werden um eine Cauchyverteilte Zufallsvariable zu erzeugen.

2.2 Rejection sampling

Das rejection sampling, auch Verwerfungsmethode genannt, ist eine simple Methode um aus Zufallszahlen einer Hilfsverteilung eine andere Wahrscheinlichkeitsverteilungen zu erzeugen. Die verallgemeinerte Vorgehensweise ist

1. Erzeuge Stichproben Y durch die Hilfsverteilung.
2. Verwerfe alle Stichproben, die die Akzeptanzbedingung nicht erfüllen

Als einfaches Beispiel betrachte Y_j von unabhängig und identisch verteilten Stichproben von $U([-1, 1]^2)$. Als Akzeptanzbedingung definiere $Y = Y_n$ mit $n = \min\{j : \|Y_j\|_2 \leq 1\}$. Die damit akzeptierten Stichproben Y_n sind uniformverteilt, da die Stichprobengesamtheit Y_j uniformverteilt auf $[0, 1]^2$ sind, auch wenn die verworfenen Stichproben entfernt wurden.

Wie gut diese Methode ist, ist abhängig von der Anzahl der verworfenen Stichproben sowie der Generierungsgeschwindigkeit der Stichproben einer Verteilung. Sei p die Akzeptanz Wahrscheinlichkeit einer Stichprobe, so ist die Wahrscheinlichkeit, dass eine Stichprobe beim j -ten Versuch akzeptiert wird $(1-p)^{j-1} \cdot p$. Nach der Definition des Erwartungswertes gilt

$$\text{erw. Anzahl Versuche} = \sum_{j=1}^{\infty} j \cdot (1-p)^{j-1} \cdot p = -p \frac{d}{dp} \sum_{j=0}^{\infty} (1-p)^j = \frac{1}{p} \quad .$$

Dies bedeutet, dass es sinnvoll ist eine Verteilung zu wählen mit einer Akzeptanzwahrscheinlichkeit nahe 1.

Die Akzeptanzwahrscheinlichkeit im oberen Beispiel ist $p = \pi/4$ und damit ist die erwartete Anzahl an Versuchen gleich $4/\pi \approx 1,27$. Dies bedeutet 0,27 Versuche werden im Durchschnitt verworfen.

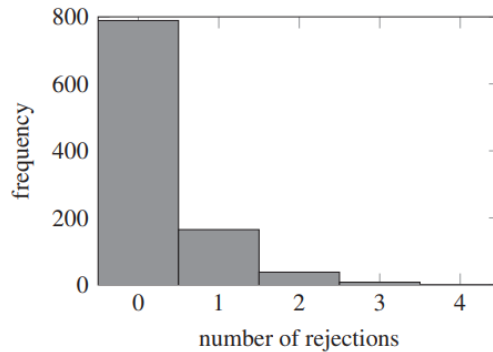


Abbildung 2: Histogramm der Anzahl von verworfenen Stichproben bei 1000 erzeugten Stichproben

2.3 Die mehrdimensionalen Normalverteilung $N(\mu, C)$

In diesem Abschnitt wird nun skizziert wie man mit der Cholesky Faktorisierung Stichproben erzeugen kann, welche mehrdimensional normalverteilt sind.

Cholesky Faktorisierung:

Sei $C \in \mathbb{R}^{n \times n}$ positiv definit, so kann C dargestellt werden als $C = L \cdot D \cdot L^T$ bzw. $C = R^T \cdot R$, wobei es sich bei L bzw. R um eine untere bzw. obere Dreiecksmatrix handelt und bei D um eine Diagonalmatrix.

Zunächst erzeuge man d viele unabhängige Stichproben $Z_1, \dots, Z_d \sim N(0, 1)$ und definiert im Anschluss $X := \mu + R^T \cdot Z$ mit $Z = [Z_1, \dots, Z_d]^T \sim N(0, I_d)$.

Es gilt :

$$E[X] = \mu + 0 = \mu$$

$$E[(X - \mu)(X - \mu)^T] = E[(R^T Z)(R^T Z)^T] = R^T E[ZZ^T]R = C$$

$$\Rightarrow X \sim N(\mu, C)$$

Das Verfahren ist einfach, jedoch aufgrund der Cholesky Faktorisierung in $\mathcal{O}(d^3)$. Glücklicherweise muss die Cholesky Faktorisierung nur einmalig berechnet werden für eine beliebige Anzahl Stichproben von $N(\mu, C)$.

3 Die Monte Carlo Methode

3.1 Einführung

Die Idee der Monte Carlo Methode basiert auf der wiederholten Zufallszahlengenerierung um damit eine gewünschte numerische Größe zu berechnen.

Von Berechnungen zu Neutronenstreuung in einer Atombombe, Studien von Proteinen, Klimawandel, Finanzen bis hin zur 3D Modellierung findet diese Methode vielerlei Anwendungen.

Die allgemeine Vorgehensweise kann wie folgt beschrieben werden:

1. Definiere eine Menge an möglichen Inputs.
2. Erzeuge solch zufällige Inputs durch eine Wahrscheinlichkeitsverteilung.
3. Deterministische Berechnung der Inputs.
4. Führe die Ergebnisse/Inputs zusammen.

Die Methode wird nun anhand des bekannten Beispiels der Zahl π Abschätzung veranschaulicht.

Beispiel:

Zunächst werden M viele Punkte unabhängig und uniformverteilt auf dem Einheitsquadrat $Q = [0,1] \times [0,1]$ erzeugt. Sei M_0 die Anzahl an Punkten im vom Quadrat Q eingeschlossenen Kreises.

Dann gilt:

$$\frac{\text{Fläche(Kreis)}}{\text{Fläche(Quadrat)}} = \frac{\pi \cdot \frac{1}{4}}{1} = \frac{\pi}{4}$$

Für $M \rightarrow \infty$ konvergiert das Verhältnis von M_0 zu M gegen das Verhältnis der Flächen $\frac{\pi}{4}$, was der Wahrscheinlichkeit eines Punktes im Kreis zu landen entspricht. Somit kann π abgeschätzt werden als

$$\pi = 4 \cdot \frac{M_0}{M} \quad \text{für } M \gg 1 \quad .$$

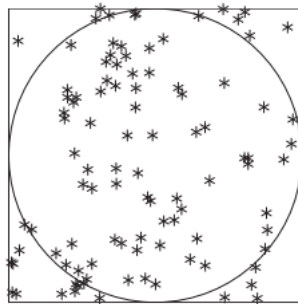


Abbildung 3: Repräsentatives Experiment mit 100 uniformverteilten Stichproben, $M=100$, $M_0= 76$ und π Abschätzung $4 \cdot 76/100 = 3,04$

Allgemein werden bei Monte Carlo Methoden durch verschiedene Inputs, Outputs X_j erzeugt und dann gemittelt, welche gegen den gewünschten Wert konvergiert(sofern gut gestellt).

$$\bar{X}_M = \frac{X_1 + X_2 + \dots + X_M}{M} \rightarrow \text{gewünschter Wert}$$

Im Beispiel oben wäre

$$X_j = \begin{cases} 0, & X_j \in M \setminus M_0 \\ 1, & X_j \in M_0 \end{cases}$$

Problem:

Die Konvergenzrate der Approximation ist $\mathcal{O}(M^{-\frac{1}{2}})$;

dh. pro zusätzliche genaue Kommastelle muss M um das 100-fache erhöht werden.

Zwar ist die Konvergenz relativ langsam, jedoch ist die Methode simpel und besitzt zudem noch weitere Vorteile:

1. dimensionsunabhängige Konvergenzrate
2. Wiederholung deterministischer Algorithmen
3. Monte Carlo Methoden sind iterativ
Ein Experiment reicht für eine (schlechte) Approximation und kann durch weitere Experimente verbessert werden.

3.2 Analyse der Monte Carlo Methode

Seien X_j unabhängige und identisch verteilte reellwertige Zufallsvariablen des Wahrscheinlichkeitsraumes (Ω, Σ, P) mit je Erwartungswert μ und Varianz σ^2 .

Das Gesetz der großen Zahlen:

$$\sum_{i=1}^M \frac{X_i}{M} = \bar{X}_M \xrightarrow{M \rightarrow \infty} \mu \quad \text{bzw.} \quad \mathcal{P}(\lim_{M \rightarrow \infty} \bar{X}_M = \mu) = 1$$

Im obigen Beispiel ist dies

$$M_0 = \sum_{i=1}^M X_j \quad ; \quad \mathbb{E}[X_j] = \frac{\pi}{4} \quad ; \quad \frac{M_0}{M} \rightarrow \frac{\pi}{4}$$

Nun untersuchen wir die Konvergenzrate von $\bar{X}_M \rightarrow \mu$, hierzu betrachten wir zunächst die Varianz

$$\text{Var}(\bar{X}_M) = \text{Var}\left(\sum_{i=1}^M \frac{X_j}{M}\right) = \frac{1}{M^2} \sum_{i=1}^M \text{Var}(X_j) = \frac{M \cdot \sigma^2}{M^2} = \frac{\sigma^2}{M} \xrightarrow{M \rightarrow \infty} 0$$

Theorem:

Sei $X_n \rightarrow X$ im L^p Sinne, $p > 1$ und angenommen für $r > 0$, Konstante $K(p) > 0$ gilt

$$\|X_n - X\|_{L^p(\Omega, H)} = \mathbb{E}[\|X_n - X\|^p]^{\frac{1}{p}} \leq K(p) \cdot n^{-r}$$

So gilt : $X_n \rightarrow X$ in \mathcal{P} und $\mathcal{P}(\|X_n - X\| > n^{-r+\epsilon}) \leq K(p)^p \cdot n^{-p\epsilon}$; $\forall \epsilon > 0$

Mit $p = 2$, $r = \frac{1}{2}$ gilt für jedes $\epsilon > 0$:

$$\mathcal{P}(\|\bar{X}_M - \mu\| \geq M^{-\frac{1}{2}+\epsilon}) \leq \frac{\sigma^2}{M^{2\epsilon}}$$

dh. für $M \rightarrow \infty$ geht die Wahrscheinlichkeit, dass der Fehler größer ist als $M^{-\frac{1}{2}+\epsilon}$, gegen 0 für alle $\epsilon > 0$.

Durch den Zentralen Grenzwertsatz und der Berry-Esséen Ungleichung kann die Konvergenz nochmal näher untersucht werden.

Zentraler Grenzwertsatz:

Seien X_j Zufallsvariablen mit $\text{Var}(X_j) = \sigma^2$, so gilt, dass die Verteilungsfunktion von \bar{X}_M für $M \rightarrow \infty$ gegen die Verteilungsfunktion der Normalverteilung $N(\mu, \frac{\sigma^2}{M})$ mit Erwartungswert μ und Varianz $\frac{\sigma^2}{M}$ konvergiert.

Insbesondere konvergiert $X_M^* = M^{\frac{1}{2}}(\bar{X}_M - \mu)$ gegen $Z \sim N(0, \sigma^2)$.

Berry-Esséen Ungleichung:

Sei $E[|X_j|^3] < \infty$ so konvergiert $X_M^* \rightarrow N(0, \sigma^2)$ in $\mathcal{O}(M^{-\frac{1}{2}})$.
 Insbesondere für $Z \sim N(0, \sigma^2)$ gilt

$$\sup_{z \in \mathbb{R}} |\mathcal{P}(X_M^* \leq z) - \mathcal{P}(Z \leq z)| \leq \frac{E[|X_1 - \mu|^3]}{\sigma^3 \sqrt{M}}$$

Betrachte nun die Zufallsvariable X_M^* und sei $Z \sim N(0, \sigma^2)$, so gilt
 $\mathcal{P}(X_M^* \leq z) = \mathcal{P}(Z \leq z) + \mathcal{O}(M^{-\frac{1}{2}})$ und folglich

$$\begin{aligned} \mathcal{P}(|X_M^*| \leq R) &= \mathcal{P}(X_M^* \leq R) - \mathcal{P}(X_M^* < -R) \\ &= \mathcal{P}(Z \leq R) - \mathcal{P}(Z < -R) + \mathcal{O}(M^{-\frac{1}{2}}) \\ &= \mathcal{P}(|Z| \leq R) + \mathcal{O}(M^{-\frac{1}{2}}) \\ &= \text{erf}\left(\frac{R}{\sqrt{2}\sigma}\right) + \mathcal{O}(M^{-\frac{1}{2}}) \end{aligned}$$

wobei erf der Gaußschen Fehlerfunktion entspricht.

Für $R=2\sigma$: $\mathcal{P}(|\bar{X}_M - \mu| \leq \frac{2\sigma}{\sqrt{M}}) \approx 0,9545 + \mathcal{O}(M^{-\frac{1}{2}})$

Das Intervall $[\bar{X}_M - \frac{2\sigma}{\sqrt{M}}, \bar{X}_M + \frac{2\sigma}{\sqrt{M}}]$ entspricht dem 95% Erwartungsbereich, also das Intervall in dem μ zu 95% enthalten ist.

Beispiel (zur Abschätzung $\mathcal{O}(M^{-\frac{1}{2}})$):

Wie zuvor im π Beispiel, gäbe es M viele Zufallsvariablen X_j mit $\mathcal{P}(X_j = 1) = \frac{\pi}{4}$, $\mathcal{P}(X_j = 0) = 1 - \frac{\pi}{4}$ und $\mu = p = \frac{\pi}{4}$, $\sigma^2 = p(1 - p)$, so sind die Punkte $M_0 = M \cdot \bar{X}_M$ binomialverteilt, $B(M, p)$. Mit der Berry-Esséen Ungleichung für die Binomialverteilung erhält man:

$$\mathcal{P}(|\bar{X}_M - p| \leq \frac{2\sqrt{p(1-p)}}{\sqrt{M}}) > 0,95 - \frac{2(1-2p+2p^2)}{\sqrt{M} \cdot p(1-p)}$$

mit $p = \frac{\pi}{4}$: $\mathcal{P}(|\bar{X}_M - \frac{\pi}{4}| \leq \frac{0,8211}{\sqrt{M}}) > 0,95 - \frac{3,23}{\sqrt{M}}$

Beispiel (ODE mit zufälligen Anfangsdaten):

Seien u_1, u_2 zwei Populationen. Betrachte

$$\frac{d}{dt} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} u_1(1-u_2) \\ u_2(u_1-1) \end{pmatrix} ; \quad u(0) = u_0 \quad .$$

Durch solche Modell versucht man durch gegebene oder abgeschätzte Populationen, zukünftige Populationsgrößen vorherzusagen. Im folgenden betrachten wir den Erwartungswert sowie Fehler der Vorhersage durch Stichproben mit verschiedenen Anfangsbedingungen unter der Monte Carlo Methode.

Um die Ungewissheit der Anfangspopulation darzustellen, wird angenommen, dass $u_0 \sim U(D)$, also uniformverteilt auf

$$D = [u_{0,1} - \epsilon, u_{0,1} + \epsilon] \times [u_{0,2} - \epsilon, u_{0,2} + \epsilon] \quad .$$

Nun wollen wir den erwarteten Wert der ersten Komponente zur Zeit T , $E[u_1(T)]$, abschätzen. Da es keine analytische Lösung des Problems gibt, wenden wir das explizite Euler Verfahren an:

Sei $\Delta t > 0$ die Schrittweite, $u_0 \in \mathbb{R}^d$ dann kann die explizite Eulerapproximation u_n zur Lösung $u_n(t_n)$, $t_n = n \cdot \Delta t$, des Anfangswertsproblems $\frac{du}{dt} = f(u)$, $u(0) = u_0$ berechnet werden durch

$$u_{n+1} = u_n + \Delta t \cdot f(u_n) \quad .$$

Sei nun u_n die explizite Eulerapproximation der obigen ODE zur Zeit $t = n \cdot \Delta t$, welche zu jeder Schrittweite eine Zufallsvariable darstellt.

Definiere nun $\phi(u) := u_1$ und approximiere $E[\phi(u_N)]$ für $T = N \cdot \Delta t$ mittels der Monte Carlo Methode.

Um nun den Mittelwert \bar{X}_M von M Stichproben von $\phi(u_N)$ zu erhalten, wird das explizite Euler Verfahren auf die ODE mit verschiedenen unabhängigen Anfangsbedingungen angewandt (siehe hierfür Algorithmus unten). Wir erwarten dann bei kleiner Schrittweite Δt und M groß, dass \bar{X}_M den Erwartungswert $E[\phi(u(T))]$ relativ gut approximiert.

Algorithmus:

Input: M, T, Δt , u_0 , ϵ

Output: \bar{x} (M.C. Abschätzung), sig95 (95% Erwartungsbereich)

```

1 function [bar_x, sig95]=pop_monte(M,T,Dt,baru0,epsilon)
2 u=[];
3 for j=1:M,
4     u0 = baru0+epsilon*(2*rand(1,2)-1); % sample initial data
5     u(j,:)=pop_solve(u0,T,Dt); % solve ODE
6 end;
7 [bar_x, sig95]=monte(u(:,1))% analyse first component
1 function [sample_av, conf95]=monte(samples)
2 M=length(samples);
3 conf95 = 2*sqrt(var(samples)/M); sample_av = mean(samples);

```

Der Fehler besteht bei diesem Vorgehen aus zwei Komponenten, dem Fehler der Monte Carlo Methode und der des expliziten Euler Verfahrens. Sei $\mu = E[\phi(u(T))]$ der exakte Erwartungswert. Dann ist der Fehler bei M Stichproben und N Zeitschritten gegeben durch

$$\epsilon_M = |\mu - \bar{X}_M| \leq \underbrace{|\mu - E[\phi(u_N)]|}_{\text{i) Eulerv. Fehler}} + \underbrace{|E[\phi(u_N)] - \bar{X}_M|}_{\text{ii) M.C.M. Fehler}} .$$

i) explizite Euler Verfahren Fehler:

Der Fehler ist bekannt und kann durch $\|u(T) - u_N\|_2 \leq K \cdot \Delta t$ abgeschätzt werden.

Aufgrund der Lipschitzstetigkeit von ϕ , folgt $|\phi(u(T)) - \phi(u_N)| \leq K \cdot L \cdot \Delta t$ und somit

$$|\mu - E[\phi(u_N)]| = |E[\phi(u(T))] - E[\phi(u_N)]| \leq K \cdot L \cdot \Delta t .$$

ii) Monte Carlo Verfahren Fehler:

Sei $Var(\phi(u_N)) = \sigma^2$, so erhalten wir mit der vorangegangenen Abschätzung

$$\begin{aligned} \mathcal{P}(|E[\phi(u_N)] - \bar{X}_M| \leq \frac{2\sigma}{\sqrt{M}}) &> 0,95 + \mathcal{O}(M^{-\frac{1}{2}}) \\ \Rightarrow \mathcal{P}(\epsilon_M \leq K \cdot L \cdot \Delta t + \frac{2\sigma}{\sqrt{M}}) &> 0,95 + \mathcal{O}(M^{-\frac{1}{2}}) . \end{aligned}$$

Um nun eine feste Genauigkeit δ zu erreichen, setzen wir $K \cdot L \cdot \Delta t = \delta/2$ und $\frac{2\sigma}{\sqrt{M}} = \delta/2$.

Damit ist $\Delta t \approx \frac{\delta}{2KL}$ und $M \approx \frac{16\sigma^2}{\delta^2}$.

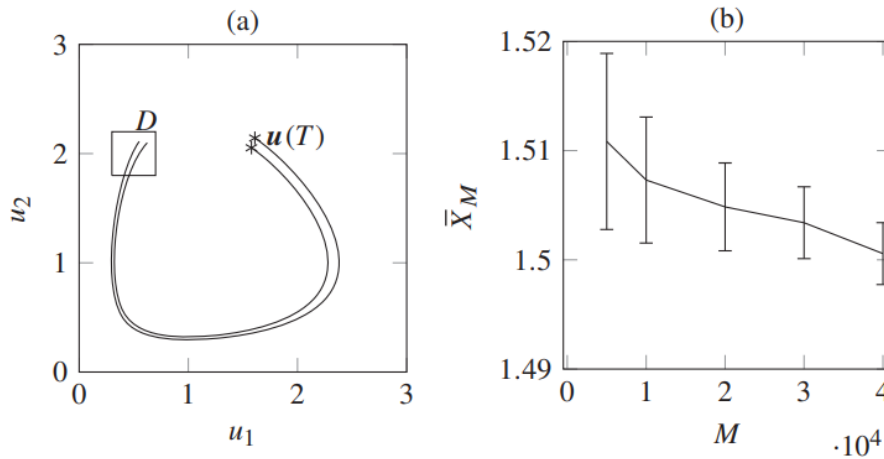


Abbildung 4: $u_0 = [\frac{1}{2}, 2]^T, \epsilon = 0, 2$; a) Bahnen für zwei zufällige Anfangsbedingungen, $T=6$; b) Verhältnis \bar{X}_M zu $E[u_1(T)]$ mit 95% Erwartungsbereich bei $\Delta t = T/10 \cdot M^{1/2}$

3.3 Verbesserungen der Monte Carlo Methode

Im Anschluss betrachten wir nun Möglichkeiten zur Verbesserung der Konvergenzrate bei Monte Carlo Methoden. Als erstes betrachten wir die Quasi Monte Carlo Methoden. Im Vergleich zur Monte Carlo Methode, die Folgen von (Pseudo-)Zufallszahlen nutzt, benutzt die Quasi Monte Carlo Methode eine Folge von Quasi Zufallszahlen, auch low discrepancy sequence genannt. Die Diskrepanz einer Folge ist gering, falls die Anzahl an Punkte der Folge, die in einer Menge B liegen, proportional zum Maß von B ist.

Beispiel (van der Corput Folge zur Basis 3):

$$x_n = (0, \frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{4}{9}, \frac{7}{9}, \frac{2}{9}, \frac{5}{9}, \frac{8}{9}, \frac{1}{27}, \dots)$$

Diese ist dargestellt als $\frac{k}{3^j}$, wobei die Reihenfolge der k gegeben ist durch das rückwärtslesen der Basisrepräsentation.

Zb. : $4_3 = 11, 7_3 = 21, 2_3 = 02$ und somit die Reihenfolge $\frac{4}{9}, \frac{7}{9}, \frac{2}{9}$ weil $11 < 12 < 20$.

Die Folge ist uniform verteilt auf dem Einheitsintervall $[0,1]$.

Die Quasi Monte Carlo Methoden erhält man dann durch den Tausch der unabhängig und identisch verteilten Zufallszahlen von $U(0,1)$ der Monte Carlo Berechnung für $E[\phi(X)]$ mit $\phi \in C_c^\infty(0,1)$, $X \sim U(0,1)$ durch die van der Corput Folge der Länge M.

Dadurch kann der Fehler von $\mathcal{O}(M^{-1/2})$ zu $\mathcal{O}(M^{-2})$ verringert werden. Jedoch muss dabei beachtet werden, dass die Konvergenzrate von der Dimension d der Folge abhängt.

Eine weitere Möglichkeit sind die Methoden der Varianz Reduktion, diese haben, wie der Name sagt, das Ziel die Varianz der Zufallsvariablen zu verringern. Jede Output Zufallsvariable wird eine Varianz zugeordnet, welche die Genauigkeit des Experiments beeinflusst. Mittels der Varianz Reduktion kann also die Genauigkeit, der durch ein Experiment gegebenen Abschätzung erhöht werden.

Im Beispiel der π Abschätzung durch die Monte Carlo Methode kann die Varianz verringert werden, wenn man anstatt des Quadrates eine Fläche nimmt, die sich einem Kreis annähert.

Beispiel für eine Varianz Reduktionsmethode : antithetic sampling (antithetische Variate)

Zusätzlich zu den Stichproben X_1, \dots, X_M und \bar{X}_M der Monte Carlo Methode berechnet man weitere Stichproben X_1^a, \dots, X_M^a und \bar{X}_M^a . Es gilt

$$\bar{X}_M, \bar{X}_M^a \rightarrow E[X] \quad \text{sowie} \quad \frac{\bar{X}_M + \bar{X}_M^a}{2} \rightarrow E[X] \quad .$$

Ist $Cov(X_j, X_j^a) < 0$, dh. negativ korreliert, so heißen die Stichproben X_j^a antithetic Stichproben und die Approximation $(\bar{X}_M + \bar{X}_M^a)/2$ ist besser als \bar{X}_{2M} , wie man im folgenden Satz sehen kann.

Satz:

Seien X_j, X_j^a identisch verteilte reellwertige Zufallsvariablen, sodass $Cov(X_j, X_k) = Cov(X_j^a, X_k^a) = 0, j \neq k$. Dann gilt

$$Var\left(\frac{\bar{X}_M + \bar{X}_M^a}{2}\right) = Var(\bar{X}_{2M}) + \frac{1}{2} \cdot Cov(\bar{X}_M, \bar{X}_M^a) \leq Var(\bar{X}_M) \quad .$$

Beweis:

Sind X,Y unkorreliert, so gilt $Var((X+Y)/2) = (Var(X) + Var(Y))/4$. Ebenfalls gilt $Var(\bar{X}_{2M}) = Var(\bar{X}_M)/2 = Var(\bar{X}_M^a)/2$, da es sich bei \bar{X}_M, \bar{X}_M^a um Mittelwerte unkorrelierter Stichproben der selben Verteilung handelt.

$$\begin{aligned} Var\left(\frac{\bar{X}_M + \bar{X}_M^a}{2}\right) &= \frac{1}{4}(Var(\bar{X}_M) + Var(\bar{X}_M^a) + 2 \cdot Cov(\bar{X}_M, \bar{X}_M^a)) \\ &= Var(\bar{X}_{2M}) + \frac{1}{2} \cdot Cov(\bar{X}_M, \bar{X}_M^a) \quad . \end{aligned}$$

Behauptung: $Var(X+Y) \leq 2 \cdot (Var(X) + Var(Y))$

Beweis:

Mit der Cauchy Schwarz Ungleichung ist $Cov(X, Y)^2 \leq Var(X) \cdot Var(Y)$. Somit ist dann

$$\begin{aligned} 2 \cdot Cov(X, Y) &\leq 2 \cdot (Cov(X, Y)^2)^{1/2} \leq 2 \cdot (Var(X) \cdot Var(Y))^{1/2} \\ &\leq 2 \cdot \left(\frac{Var(X) + Var(Y)}{2}\right) \leq Var(X) + Var(Y) \end{aligned}$$

$$\Rightarrow Var(X) + Var(Y) + 2 \cdot Cov(X, Y) \leq 2 \cdot (Var(X) + Var(Y)) \quad \square$$

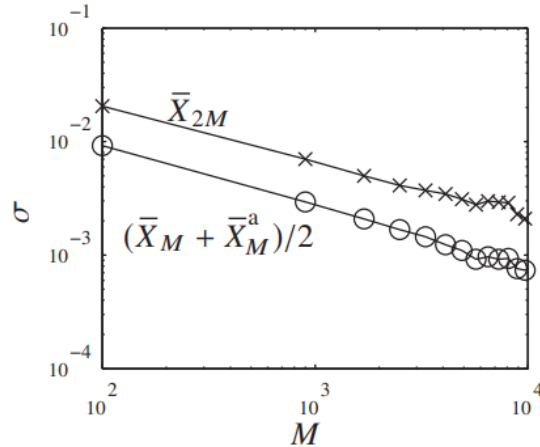
Somit folgt schlussendlich

$$Var\left(\frac{\bar{X}_M + \bar{X}_M^a}{2}\right) \leq 2 \cdot \left(Var\left(\frac{\bar{X}_M}{2}\right) + Var\left(\frac{\bar{X}_M^a}{2}\right)\right) = Var(\bar{X}_M) \quad \square$$

Ist also $Cov(\bar{X}_M, \bar{X}_M^a) < 0$, dann ist $Var((\bar{X}_M + \bar{X}_M^a)/2) < Var(\bar{X}_{2M})$.

Beispiel (ODE):

Wir führen das ODE Beispiel nun fort. Definiere $A(u) := 2u_0 - u$, ist $u \sim U(D)$ dann ist auch $A(u) \sim U(D)$. Dies bedeutet, dass die Approximation u_n^a des expliziten Euler Verfahrens mit Anfangsdaten $A(u_0)$ ist identisch verteilt wie die Approximation u_n mit Anfangsdaten $u_0 \sim U(D)$. Um nun das antithetic sampling anzuwenden um $E[\phi(u(T))]$ zu approximieren, berechnet man zuerst \bar{X}_M von M unabhängigen Stichproben $\phi(u_{N,j})$ von $\phi(u_N)$, sowie \bar{X}_M^a von M antithetic Stichproben $\phi(u_{N,j}^a)$ und berechnet das Mittel $(\bar{X}_M + \bar{X}_M^a)/2$.



In der Abbildung ist die Beziehung von Standardabweichung σ zur Anzahl an Stichproben M dargestellt. Man kann sehen, dass die Standardabweichung von $(\bar{X}_M + \bar{X}_M^a)/2$ geringer ist als die von \bar{X}_{2M} und somit auch der Erwartungsbereich der Approximation von $E[\phi(u(T))]$. Ebenfalls sichtbar ist die Konvergenzrate, welche für $M \rightarrow \infty$ in beiden Fällen in $\mathcal{O}(M^{-1/2})$ ist.

Literatur

- [1] Gabriel J. Lord, Catherine E. Powell, Tony Shardlow; AN INTRODUCTION TO COMPUTATIONAL STOCHASTIC PDES; Cambridge University Press (2014).
- [2] Ronald W. Shonkwiler, Franklin Mendivil; Explorations in Monte Carlo Methods; Springer Verlag (2009).
- [3] Enrico Zio; The Monte Carlo Simulation Method for System Reliability and Risk Analysis; Springer-Verlag London (2013).