



B A K K A L A U R E A T S A R B E I T

Einführung in Quasi-Monte Carlo Verfahren

für die Lehrveranstaltung
Seminar Bakkalaureat TM (Finanz- und Versicherungsmathematik),

ausgeführt am Institut für
Analysis und Computational Number Theory (Math A)
der Technischen Universität Graz

unter Anleitung von O.Univ.-Prof. Dr.phil. Robert Tichy

durch
Markus Zahrnhofer
Matr.-Nr. 0357944

im
Sommersemester 2007

Datum

Unterschrift

Inhaltsverzeichnis

1	Monte Carlo Verfahren	2
1.1	Einführung	2
1.2	Monte Carlo Verfahren	3
1.3	Quasi-Monte Carlo Verfahren	8
2	Quasi-Monte Carlo Verfahren	10
2.1	Diskrepanz	10
2.2	Fehlerschranken	16

Kapitel 1

Monte Carlo Verfahren

Für ein grundlegendes Verständnis der Quasi-Monte-Carlo Methode, sind Kenntnisse über das Monte-Carlo Verfahren hilfreich.

1.1 Einführung

Betrachten wir ein s -dimensionales numerisches Integrationsproblem. Für $s = 1$ gibt es bekannte Integrationsregeln wie z.B. Trapezregel oder Simpsonregel. Die Trapezregel ist ein Verfahren 2. Ordnung d.h. lineare Funktionen werden exakt integriert. Im Vergleich dazu ist die Simpsonregel ein Verfahren 4. Ordnung, womit eine exakte Integration kubischer Funktionen möglich ist. Wenden wir die Trapezregel auf das abgeschlossene Einheitsintervall $[0, 1]$ an und erhalten folgende Form

$$\int_0^1 f(u) du \approx \sum_{n=0}^m w_n f\left(\frac{n}{m}\right) \quad (1.1)$$

mit $m \in \mathbb{N}$ und den Integrationsgewichten w_n die wie folgt definiert sind: $w_0 = w_m = \frac{1}{2m}$ und $w_n = \frac{1}{m}$ für $1 \leq n \leq m - 1$. Wir erhalten einen Approximationsfehler von $O(m^{-2})$, vorausgesetzt f hat eine stetige 2. Ableitung auf $[0, 1]$.

Im mehrdimensionalen Fall, $s \geq 2$, wo über ein Intervall integriert wird, verwendet man für die numerische Integration das Kartesische Produkt von eindimensionalen Integrationsverfahren. Die Stützstellen werden als Kartesisches Produkt der eindimensionalen Stützstellen gewählt. Die Integrationsgewichte sind die entsprechenden Produkte der Eindimensionalen. Ein s -dimensionales Integral kann nun als Iteration von eindimensionalen Integralen betrachtet werden. Zur Illustration betrachten wir das Kartesische Produkt der Trapezregel (1.1):

$$\int_{\bar{I}^s} f(u) du \approx \sum_{n_1=0}^m \dots \sum_{n_s=0}^m w_{n_1} \dots w_{n_s} f\left(\frac{n_1}{m}, \dots, \frac{n_s}{m}\right), \quad (1.2)$$

mit $\bar{I} = [0, 1]^s$, dem s -dimensionale abgeschlossene Einheitswürfel und den Gewichten die analog wie in (1.1) zu wählen sind.

In (1.2) erhalten wir $N = (m + 1)^s$ Knoten. Aus der Fehlerabschätzung von (1.1) folgt für die s -dimensionale Approximation aus (1.2) die Fehlerschranke $O(m^{-2})$, vorausgesetzt $\frac{\partial^2 f}{\partial u_i^2}$

ist stetig auf \bar{I}^s für $1 \leq i \leq s$. Im Allgemeinen muss der Approximationsfehler aus (1.2) nicht kleiner als der des Eindimensionalen sein.

Stellt man die Fehlerschranke mit Hilfe der Anzahl der Knoten dar, erhält man $O(N^{\frac{-2}{s}})$. Man kann daraus erkennen, dass eine Erhöhung der Dimensionen die Brauchbarkeit der Schranke verringert. Jedoch um einen exakten Fehler garantieren zu können, z.B. $\leq 10^{-2}$, benötigt man beinahe 10^8 Knoten. Woraus folgt, dass die Anzahl Knoten exponentiell mit der Dimension s steigt. Diese Eigenschaft wird auch "Curse of Dimensionality" genannt.

Um das exponentielle Steigen der Anzahl der Knoten in Abhängigkeit der Dimension s zu überwinden wurde in den 40er Jahren die *Monte Carlo Methode* entwickelt. Diese Methode ist keine klassische numerische Integrationsmethode, wie z.B. Trapez- oder Simpsonregel, sondern eine auf einer Zufallsstichprobe basierende Approximationsmethode. Aufgrund dessen hat diese Methode einen sehr starken wahrscheinlichkeitstheoretischen und statistischen Einschlag. Ein positiver Aspekt dieser Methode ist, dass sie Integrationsfehler liefert, bei denen die Anzahl der Knoten unabhängig von der Dimension ist. Dies ist eine entscheidende Verbesserung gegenüber den klassischen numerischen Methoden.

Ein wichtiger Aspekt für die Monte Carlo Methode ist die richtige bzw. passende Wahl der Zufallsstichprobe, denn die Güte der Methode hängt entscheidend von der Qualität der Stichprobe ab; d.h. wie gut eine Stichprobe wahre Ereignisse wiedergibt.

Die Entwicklung der Monte Carlo Methode hat ein großes Interesse an der Generierung von Zufallszahlen und Zufallsvektoren hervorgerufen. Aufgrund der deterministischen Generierung spricht man auch von *Pseudozufallszahlen* und *Pseudozufallsvektoren* die heutzutage mit Hilfe von Computern abgehandelt wird. Erste Algorithmen zur Bestimmung wurden bereits in den 40ern entwickelt.

Wendet man die Monte Carlo Methode auf das numerische Integrationsbeispiel an, liefert es nur eine probabilistische Fehlerschranke, was sich als Nachteil erweisen kann. Jedoch, wie wir später sehen werden, ist es nicht so entscheidend, dass möglichst viele wahre Ereignisse durch die Zufallsstichprobe erzielt werden, sondern eher eine gleichmässige Verteilung dieser über den Integrationsbereich. Weiters werden wir sehen, dass durch die deterministische Wahl der Knoten eine deterministische Fehlerschranke erreicht werden kann. Die Idee, die Knoten so zu wählen, dass der Fehler möglichst gering wird, ist das Grundkonzept der *Quasi-Monte Carlo Methoden*. Dabei werden die Zufallszahlen durch exakt bestimmte Punkte ersetzt. Das Hauptziel der *Quasi-Monte Carlo Methoden* ist es die Punkte so zu wählen, dass ein geringerer Fehler als bei der Monte Carlo Methode erreicht wird.

1.2 Monte Carlo Verfahren

In diesem Abschnitt werden wir die Monte Carlo Methode im Zusammenhang mit der numerischen Integration analysieren. Dazu betrachten wir die approximative Berechnung folgenden Integrals $\int_B f(u)du$ mit dem Integrationbereich $B \subseteq \mathbb{R}^s$ und $0 < \lambda_s(B) < \infty$, wobei $\lambda_s(B)$ das s -dimensionale Lebesgue Maß ist. Sei B ein Wahrscheinlichkeitsraum mit dem Wahrscheinlichkeitsmaß $d\mu = \frac{du}{\lambda_s(B)}$. Weiters sei $f \in L^1(\mu)$ wir erhalten

$$\int_B f(u)du = \lambda_s(B) \int_B f d\mu = \lambda_s(B)E(f) \quad (1.3)$$

wobei $E(f)$ der Erwartungswert der Zufallsvariable f ist. Somit haben wir das Problem der

numerischen Integration auf ein Problem der approximativen Berechnung des Erwartungswertes reduziert. Der Erwartungswert lässt sich durch das arithmetische Mittel darstellen. Diese Idee kann auf allgemeine Wahrscheinlichkeitsräume ausgeweitet werden.

Sei f eine Zufallsvariable auf einem beliebigen Wahrscheinlichkeitsraum $(A, \mathcal{A}, \lambda)$. Für die *Monte Carlo Abschätzung* des Erwartungswertes, seien $a_1, \dots, a_N \in A$ unabhängige und λ -verteilte Zufallsstichproben und

$$E(f) \approx \frac{1}{N} \sum_{n=1}^N f(a_n). \quad (1.4)$$

Nach dem *Starken Gesetz der großen Zahlen* gilt:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(a_n) = E(f) \quad \lambda^\infty f.s$$

wobei λ^∞ das Produktmaß von überabzählbar vielen Kopien von λ ist. Um den probabilistischen Fehler des Erwartungswertes (1.4) bestimmen zu können brauchen wir die *Varianz*

$$\sigma^2(f) = \int_A (f - E(f))^2 d\lambda,$$

die endlich ist wenn $f \in L^2(\lambda)$ ist.

Satz 1.1 Sei $f \in L^2(\lambda)$, dann gilt für jedes $N \geq 1$:

$$\int_A \dots \int_A \left(\frac{1}{N} \sum_{n=1}^N f(a_n) - E(f) \right)^2 d\lambda(a_1) \dots d\lambda(a_N) = \frac{\sigma^2(f)}{N}.$$

Beweis: Sei $g = f - E(f)$; dann ist $\int_A g d\lambda = 0$ und

$$\frac{1}{N} \sum_{n=1}^N f(a_n) - E(f) = \frac{1}{N} \sum_{n=1}^N g(a_n)$$

Somit folgt:

$$\begin{aligned} & \int_A \dots \int_A \left(\frac{1}{N} \sum_{n=1}^N f(a_n) - E(f) \right)^2 d\lambda(a_1) \dots d\lambda(a_N) \\ &= \int_A \dots \int_A \left(\frac{1}{N} \sum_{n=1}^N g(a_n) \right)^2 d\lambda(a_1) \dots d\lambda(a_N) \\ &= \frac{1}{N^2} \sum_{n=1}^N \int_A \dots \int_A g(a_n)^2 d\lambda(a_1) \dots d\lambda(a_N) \\ &+ \frac{2}{N^2} \sum_{1 \leq m < n \leq N} \int_A \dots \int_A g(a_m) g(a_n) d\lambda(a_1) \dots d\lambda(a_N) \\ &= \frac{1}{N} \int_A g^2 d\lambda = \frac{\sigma^2(f)}{N}. \end{aligned}$$

□

Satz 1.1 kann so interpretiert werden, dass der absolute Fehler von (1.4) im Mittel $\sigma(f)N^{-\frac{1}{2}}$ ist, wobei $\sigma(f) = (\sigma^2(f))^{\frac{1}{2}}$ die Standardabweichung von f ist. Betrachtet man den *zentralen Grenzwertsatz* erhält man weitere Informationen über den Fehler. Sei $0 < \sigma(f) < \infty$, dann gilt

$$\lim_{N \rightarrow \infty} P \left(\frac{c_1 \sigma(f)}{\sqrt{N}} \leq \frac{1}{N} \sum_{n=1}^N f(a_n) - E(f) \leq \frac{c_2 \sigma(f)}{\sqrt{N}} \right) = \frac{1}{\sqrt{2\pi}} \int_{c_1}^{c_2} e^{-\frac{t^2}{2}} dt$$

für alle Konstanten $c_1 < c_2$.

Betrachten wir nun die approximative Berechnung des Integrals (1.3) und verwenden wir (1.4) so erhalten wir die *Monte Carlo Abschätzung*

$$\int_B f(u) du \approx \frac{\lambda_s(B)}{N} \sum_{n=1}^N f(x_n), \quad (1.5)$$

wobei x_1, \dots, x_N unabhängige und μ -verteilte Zufallsvariablen aus B sind. Der absolute Fehler in (1.5) ist im Mittel $\lambda_s(B)\sigma(f)N^{-\frac{1}{2}}$. Daraus kann man bei der numerischen Integration, mit Hilfe der Monte Carlo Methode, eine probabilistische Fehlerschranke von $O(N^{-\frac{1}{2}})$ ableiten. Diese Größe hängt nur noch von der Anzahl der Knoten N ab und nicht mehr von der Dimension s , im Vergleich dazu die Fehlerschranke $O(N^{-2/s})$ der klassischen numerischen Integrationsmethoden. Für $s \geq 5$ ist die Monte Carlo Methode zu bevorzugen.

Im Falle, dass der Integrationsbereich B sehr kompliziert ist und das s -dimensionale Lebesgue-Maß schwierig zu berechnen ist, ist es nicht sinnvoll (1.5), sondern eine alternative Abschätzung anzuwenden. Dazu betrachtet man nach passender Variablentransformation, dass \bar{I}^s B beinhaltet.

$$\int_B f(u) du = \int_{\bar{I}^s} f(u) c_B(u) du$$

wobei c_B die charakteristische Funktion von B ist. Wird das Integral gleich wie (1.5) abgeschätzt erhält man folgende alternative *Monte Carlo Abschätzung*

$$\int_B f(u) du \approx \frac{1}{N} \sum_{\substack{n=1 \\ x_n \in B}}^N f(x_n), \quad (1.6)$$

mit x_1, \dots, x_N unabhängige Zufallsstichproben aus der Gleichverteilung auf \bar{I}^s . Haben dieselbe Fehlerschranke wie in (1.5).

Die vorangegangene Beschreibung gibt einen kurzen Überblick über den Ablauf der Monte Carlo Methode. Von der Auslegung eines numerischen Integrationsproblems als Erwartungswert von Zufallsvariablen, bis hin zur Analyse derer. Ein weiterer wichtiger Aspekt ist die Generierung der Zufallsvariablen. Es ist ratsam sich nicht mit einer einzigen Berechnung zufrieden zu geben, sondern die Abläufe wiederholt auszuführen um die Glaubwürdigkeit der Ergebnisse zu erhöhen. Solche Ablauffolgen sind die typische Arbeitsweise von *Stochastischen*

Simulationen, zu denen auch die Monte Carlo Methode zählt. Sie werden auch *Simulationenmethoden* genannt.

Trotz der vielfältigen Anwendungsmöglichkeiten der Monte Carlo Methode, ist sie jedoch kein Allheilmittel. Sie liefert nur eine probabilistische Fehlerschranke, was nie garantiert, dass die erwartete Exaktheit auch erreicht wird. In gewissen Bereichen, wo verlässliche Resultate unabdingbar sind, ist die Monte Carlo Methode unhaltbar.

Weiters muss darauf hingewiesen werden, dass die Fehlerschranke $O(N^{-\frac{1}{2}})$ unter sehr schwachen Voraussetzungen, nämlich der Quadratintegrierbarkeit der Funktion, hält. Es lässt sich kein Gewinn daraus erzielen, wenn die Funktion zusätzliche Eigenschaften aufweist. Im Vergleich dazu können in der klassischen numerischen Integration Funktionen unter zusätzlichen Bedingungen, eine höhere Konvergenzrate mit sich bringen. Im Generellen lassen sich folgende Nachteile der Monte Carlo Methode im Bereich der numerischen Integration zusammenfassen:

- Es existiert nur eine probabilistische Fehlerschranke
- Die Regularität der zu integrierenden Funktion wird nicht berücksichtigt
- Die Generierung der Zufallsstichproben ist schwierig.

Wie wir in Satz 1.1 gesehen haben, ist der MSE in der Monte Carlo Abschätzung $\frac{\sigma^2(f)}{N}$. Um die Effizienz der Methode zu verbessern, wurden einige Verfahren zur Verkleinerung der Varianz entwickelt. Eine davon heißt „*stratified sampling*“ (geschichtete Zufallsstichprobe) und läuft wie folgt ab. Sei f eine Zufallsvariable aus dem Wahrscheinlichkeitsraum $(A, \mathcal{A}, \lambda)$. Zerlege A in $A_1, \dots, A_k \in \mathcal{A}$ mit $\lambda(A_j) > 0$ für $1 \leq j \leq k$. Für jedes j $1 \leq j \leq k$, wähle N_j unabhängige μ_j -verteilte Zufallsstichproben $a_1^{(j)}, \dots, a_{N_j}^{(j)} \in A_j$, mit $\mu_j = \lambda(A_j)^{-1} \lambda$ dem Wahrscheinlichkeitsmaß auf A_j . Verwenden wir nun die Abschätzung:

$$E(f) = \sum_{j=1}^k \int_{A_j} f d\lambda = \sum_{j=1}^k \lambda(A_j) \int_{A_j} f d\mu_j \approx \sum_{j=1}^k \frac{\lambda(A_j)}{N_j} \sum_{n=1}^{N_j} f(a_n^{(j)}).$$

Der MSE dieser Abschätzung kann mit der selben Methode wie im Beweis von 1.1 berechnet werden. Woraus folgt:

$$\frac{\sigma^2(f)}{N} = \sum_{j=1}^k \frac{\lambda(A_j)}{N_j} \int_{A_j} \left(f - \frac{1}{\lambda(A_j)} \int_{A_j} f d\lambda \right)^2 d\lambda.$$

Um eine Varianzreduktion zu erreichen, müssen die N_j s passend gewählt werden. Sei $N = \sum_{j=1}^k N_j$ die Gesamtanzahl der Zufallspunkte, dann folgt:

Proposition 1.1 *Ist $N_j = \lambda(A_j)N$, $1 \leq j \leq k$, und ganzzahlig dann gilt:*

$$\sum_{j=1}^k \frac{\lambda(A_j)}{N_j} \int_{A_j} \left(f - \frac{1}{\lambda(A_j)} \int_{A_j} f d\lambda \right)^2 d\lambda \leq \frac{\sigma^2(f)}{N}.$$

Beweis: Durch die Form von N_j genügt folgendes zu zeigen:

$$\sum_{j=1}^k \int_{A_j} \left(f - \frac{1}{\lambda(A_j)} \int_{A_j} f d\lambda \right)^2 d\lambda \leq \int_A (f - E(f))^2 d\lambda$$

durch Ausquadrieren auf beiden Seiten erkennt man, dass es äquivalent zu

$$E(f)^2 \leq \sum_{j=1}^k \frac{1}{\lambda(A_j)} \left(\int_{A_j} f d\lambda \right)^2,$$

durch Anwendung der Cauchy-Schwarz-Ungleichung folgt:

$$\begin{aligned} E(f)^2 &= \left(\sum_{j=1}^k \int_{A_j} f d\lambda \right)^2 = \left(\sum_{j=1}^k \lambda(A_j)^{\frac{1}{2}} \frac{1}{\lambda(A_j)^{\frac{1}{2}}} \int_{A_j} f d\lambda \right)^2 \\ &\leq \left(\sum_{j=1}^k \lambda(A_j) \right) \sum_{j=1}^k \frac{1}{\lambda(A_j)} \left(\int_{A_j} f d\lambda \right)^2 = \sum_{j=1}^k \frac{1}{\lambda(A_j)} \left(\int_{A_j} f d\lambda \right)^2. \end{aligned}$$

□

Ein weiteres Verfahren zur Varianzverringern ist die “*antithetic variates*“ Methode. Wir betrachten folgendes Integral

$$E(f) = \int_0^1 f(u) du.$$

Weiters führen wir folgende Hilfsfunktion ein:

$$g(u) = \frac{1}{2}(f(u) + f(1-u)) \quad \text{für } 0 \leq u \leq 1 \quad (1.7)$$

sowie die Abschätzung:

$$E(f) = E(g) \approx \frac{1}{N} \sum_{n=1}^N g(x_n) = \frac{1}{2N} \sum_{n=1}^N (f(x_n) + f(1-x_n))$$

mit N unabhängigen und gleichverteilten Zufallspunkten $x_1, \dots, x_N \in [0, 1]$. Nach Satz 1.1 ist der MSE der Abschätzung $\frac{\sigma^2(g)}{N}$. Da aber $2N$ Funktionswerte von f in der Abschätzung vorkommen, sollte man $\frac{\sigma^2(g)}{N}$ mit $\frac{\sigma^2(f)}{2N}$ vergleichen. Die folgende Proposition zeigt wie eine Varianzreduktion erreicht werden kann.

Proposition 1.2 *Ist f eine stetige und monotone Funktion auf $[0, 1]$ und g wie in (1.7) dann gilt,*

$$\sigma^2(g) \leq \frac{1}{2} \sigma^2(f).$$

Beweis:

$$\begin{aligned} \sigma^2(g) &= \int_0^1 (g(u) - E(g))^2 du = \int_0^1 \left(\frac{1}{2}f(u) + \frac{1}{2}f(1-u) - E(f) \right)^2 du \\ &= \frac{1}{2} \int_0^1 f^2(u) du + \frac{1}{2} \int_0^1 f(u)f(1-u) du - E(f)^2. \end{aligned}$$

Somit folgt, dass die Ungleichung äquivalent zu

$$\int_0^1 f(u)f(1-u) du \leq E(f)^2 \quad \text{ist.} \quad (1.8)$$

Durch die Substitution von f durch $-f$ kann erreicht werden, falls notwendig, dass f nicht fallend ist.

$$F(u) = \int_0^u f(1-u)dt - E(f)u$$

hat auf dem abgeschlossenen Einheitsintervall eine nicht steigende Ableitung $F' = f(1-u) - E(f)$. Da $F(0) = F(1) = 0$ folgt, dass $F(u) \geq 0$ für alle $u \in [0, 1]$. Daraus folgt

$$\int_0^1 F(u)df(u) \geq 0.$$

Durch partielles Integrieren erhalten wir

$$\int_0^1 f(u)dF(u) = \int_0^1 f(u)F'(u)du \geq 0.$$

Ersetzen wir in dieser Formel $F'(u)$, durch den obigen Ausdruck, so erhalten wir (1.8).

□

1.3 Quasi-Monte Carlo Verfahren

Wir erinnern uns, dass wir bei der Monte Carlo Methode mit N zufällig gewählten Punkten im Durchschnitt einen Fehler von $N^{-\frac{1}{2}}$ erhalten. Natürlich existieren Folgen von Punkten wo der absolute Fehler nicht größer als im Mittel ist. Es wäre jedoch hilfreich, wenn man gezielt solche Folgen von Punkten konstruieren könnte. Die *Quasi-Monte Carlo Methode* für die numerische Integration zielt darauf ab, deterministische Folgen von Punkten zu konstruieren, die eine wesentliche Verbesserung des in den Monte Carlo Methoden erreichten Fehlers liefert. Die Integrationsregel für die Quasi-Monte Carlo Methode wird aus der Monte Carlo Abschätzung übernommen.

Betrachten wir den s -dimensionalen Einheitswürfel als Integrationsbereich und dazu die *Quasi-Monte Carlo Approximation*

$$\int_{\bar{I}^s} f(u) \approx \frac{1}{N} \sum_{n=1}^N f(x_n). \quad (1.9)$$

Die Ähnlichkeit mit der Monte Carlo Abschätzung ist hier nicht zu übersehen, jedoch ist der Unterschied, dass $x_1, \dots, x_N \in \bar{I}^s$ eine Folge deterministisch gewählter Punkte ist. Diese sollte so gewählt werden, dass ein kleiner Fehler in (1.9) garantiert werden kann. Analog zur Abschätzung (1.6) folgt die

Quasi-Monte Carlo Approximation

$$\int_B f(u)du \approx \frac{1}{N} \sum_{\substack{n=1 \\ x_n \in B}}^N f(x_n), \quad (1.10)$$

mit B eine Teilmenge von \bar{I}^s und x_1, \dots, x_N deterministisch bestimmte Punkte $\in \bar{I}^s$. In Kapitel 2 werden wir den Fehler der Abschätzungen (1.9) und (1.10) genauer untersuchen.

Herr Scheibelhofer wird in seinem Vortrag *“Folgen mit kleiner Diskrepanz“* die Konstruktion von Folgen betrachten, die einen kleinen Integrationsfehler garantieren.

In (1.9) liefert die Monte Carlo Methode im Mittel einen Fehler von $O(N^{-\frac{1}{2}})$, im Gegensatz dazu erreicht man mit der quasi-Monte Carlo Methode, für passend gewählte Punktfolgen, eine Fehlerschranke von $O(N^{-1}(\log(N))^{s-1})$. Diese Verbesserung ist natürlich ein Vorteil der *quasi-Monte Carlo Methode*. Da diese Methode ein vollständiger deterministischer Prozess ist, garantiert er uns eine präzise Fehlerschranke. Weiters kann mit demselben Aufwand an Leistung ein besseres Ergebnis erzielt werden. Dadurch ist die Quasi-Monte Carlo Methode, die nicht nur für die numerische Integration existiert, der Monte-Carlo Methode überlegen.

Die Quasi-Monte Carlo Methode wurde in den 50ern erfunden und machte danach eine rapide Entwicklung durch. Jedoch erst durch den Einsatz von Computern wurde sie populärer und leichter berechenbar.

Bemerkung: Der Aufbau und die Aussagen dieses Kapitels sind im Allgemeinen analog dem Kapitel 1 aus H. Niederreiter [4].

Kapitel 2

Quasi-Monte Carlo Verfahren

Wie wir bereits festgestellt haben ist die Grundidee der Quasi-Monte Carlo Methoden die Zufallsstichproben bei den Monte Carlo Methoden durch Folgen von deterministischen Punkten zu ersetzen. Die Wahl der Punkte hängt vom zugrundeliegenden numerischen Problem ab und unterliegt dem Konzept der gleichverteilten Folgen und der Diskrepanz. Die Diskrepanz kann als quantitatives Maß der Abweichung von der Gleichverteilung gesehen werden.

2.1 Diskrepanz

Wir betrachten als Integrationsbereich $\bar{I}^s := [0, 1]^s$. Weiters betrachten wir die Quasi-Monte Carlo Approximation

$$\int_{\bar{I}^s} f(u) \approx \frac{1}{N} \sum_{n=1}^N f(x_n), \quad (2.1)$$

mit $x_1, \dots, x_N \in \bar{I}^s$. Im idealisiertem Model ersetzen wir x_1, \dots, x_N durch eine unendliche Folge x_1, x_2, \dots von Punkten aus \bar{I}^s . Eine Mindestanforderung an diese Folge muss sein, dass wir ein Konvergenzverhalten in (2.1) erhalten. Daher wollen wir eine Folge x_1, x_2, \dots dass,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(x_n) = \int_{\bar{I}^s} f(u) du \quad (2.2)$$

gilt.

Definition 2.1 Die Folge x_1, x_2, \dots aus \bar{I}^s , heißt gleichverteilt auf \bar{I}^s wenn

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(x_n) = \int_{\bar{I}^s} f(u) du$$

für alle stetigen Funktionen $f \in \bar{I}^s$.

Eine äquivalente Definition dazu ist: x_1, x_2, \dots ist gleichverteilt auf \bar{I}^s wenn

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N c_J(x_n) = \lambda_s(J) \quad (2.3)$$

für s -dimensionale Subintervalle J aus \bar{I}^s , mit c_J die charakteristische Funktion von J und λ_s das s -dimensionale Lebesgue Maß ist. Wenn x_1, x_2, \dots gleichverteilt sind, dann hält (2.2) nur für Riemann-integrierbare Funktionen siehe [3]. Die Betrachtung liefert uns also, dass die "guten" Knoten in (2.1) jene sind, wo die empirische Verteilung der Gleichverteilung nahe kommt. Anders gesagt die Punkte x_1, \dots, x_N sollten auf \bar{I}^s "gleichmäßig" verteilt sein. Wie bereits erwähnt, ist die *Diskrepanz* ein quantitatives Maß für die Abweichung von der Gleichverteilung.

Sei P eine Punktmenge x_1, \dots, x_N aus \bar{I}^s Für eine beliebige Teilmenge B aus \bar{I}^s , definieren wir

$$A(B; P) = \sum_{n=1}^N c_B(x_n),$$

mit c_B der charakteristischen Funktion von B . $A(B; P)$ ist eine Zählfunktion, die die Anzahl der $x_n \in B$ mit $1 \leq n \leq N$ zählt.

Die eindimensionale Diskrepanz:

Definition 2.2 Sei x_1, \dots, x_N eine endliche Folge, dann heißt

$$D_N = D_N(x_1, \dots, x_N) = \sup_{0 \leq u < v \leq 1} \left| \frac{A([u, v]; P)}{N} - (v - u) \right|$$

Diskrepanz der gegebenen Folge.

Definition 2.3 Sei x_1, \dots, x_N eine endliche Folge, dann heißt

$$D_N^* = D_N^*(x_1, \dots, x_N) = \sup_{0 \leq u \leq 1} \left| \frac{A([0, u]; P)}{N} - u \right|$$

Sterndiskrepanz der gegebenen Folge.

Satz 2.1 Die Folge S ist gleichverteilt genau dann, wenn $\lim_{N \rightarrow \infty} D_N(S) = 0$.

Beweis: Die hinreichende Bedingung ist klar.

Wähle $m \geq 2$. Für $0 \leq k \leq m - 1$, sei $I_k = [k/m, (k+1)/m)$. Da S gleichverteilt ist, existiert ein $N_0 = N_0(m) \in \mathbb{N}$ sodass für $N \geq N_0$ und jedes $k = 0, 1, \dots, m - 1$

$$\frac{1}{m} \left(1 - \frac{1}{m}\right) \leq \frac{A(I_k; P)}{N} \leq \frac{1}{m} \left(1 + \frac{1}{m}\right) \quad (2.4)$$

Betrachten wir nun beliebige Teilintervalle $J = [\alpha, \beta)$ von I . Es existieren Intervalle J_1 und J_2 , die endliche Vereinigungen der I_k sind, sodass $J_1 \subseteq J \subseteq J_2$, $\lambda(J) - \lambda(J_1) < 2/m$ und $\lambda(J_2) - \lambda(J) < 2/m$. Aus (2.4) erhalten wir für alle $N \geq N_0$:

$$\lambda(J_1) \left(1 - \frac{1}{m}\right) \leq \frac{A(J_1; P)}{N} \leq \frac{A(J; P)}{N} \leq \frac{A(J_2; P)}{N} \leq \lambda(J_2) \left(1 + \frac{1}{m}\right).$$

wir erhalten:

$$\left(\lambda(J) - \frac{2}{m}\right) \left(1 - \frac{1}{m}\right) < \frac{A(J; P)}{N} < \left(\lambda(J) + \frac{2}{m}\right) \left(1 + \frac{1}{m}\right),$$

durch $\lambda(J) \leq 1$ folgt:

$$-\frac{3}{m} - \frac{2}{m^2} < \frac{A(J; P)}{N} - \lambda(J) < \frac{3}{m} + \frac{2}{m^2} \quad \text{für alle } N \geq N_0. \quad (2.5)$$

Da die Schranken in (2.5) unabhängig von J sind, erhalten wir $D_N(S) \leq (3/m) + (2/m^2)$ für alle $N \geq N_0$. $(3/m) + (2/m^2)$ kann beliebig klein gemacht werden, somit ist alles bewiesen. \square

Satz 2.2 Für jede Folge S vom Umfang N gilt:

$$\frac{1}{N} \leq D_N \leq 1.$$

Beweis: Die rechte Ungleichung folgt direkt aus der Definition.

Wählen $\epsilon > 0$, und sei $x \in I$ ein Element der Punktmenge. Weiters betrachten wir das Intervall $J = [x, x + \epsilon) \cap I$. Da $x \in J$, gilt: $A(J; P)/N - \lambda(J) \geq (1/N) - \lambda(J) \geq (1/N) - \epsilon$. Daraus folgt $D_N \geq (1/N) - \epsilon$, damit ist die gewünschte Ungleichung gezeigt. \square

Die mehrdimensionale Diskrepanz:

Sei \mathcal{B} eine nichtleere Familie von Lebesgue-messbaren Teilmengen von \bar{I}^s , dann ist die allgemeine Form der Diskrepanz der Punktmenge P gegeben durch

$$D_N(\mathcal{B}; P) = \sup_{B \in \mathcal{B}} \left| \frac{A(B; P)}{N} - \lambda_s(B) \right|. \quad (2.6)$$

Wobei $0 \leq D_N(\mathcal{B}; P) \leq 1$ liegt. Bei angemessener Spezialisierung der Mengenfamilie \mathcal{B} , erhalten wir weitere wichtige Diskrepanzen. Wir setzen $I^s = [0, 1)^s$.

Definition 2.4 Die Sterndiskrepanz $D_N^*(P) = D_N^*(x_1, \dots, x_N)$ der Punktmenge P ist definiert durch $D_N^*(P) = D_N(\mathcal{J}^*; P)$, wo \mathcal{J}^* eine Familie aller Teilintervalle von I^s der Form $\prod_{i=1}^s [0, u_i)$ ist.

Definition 2.5 Die Diskrepanz $D_N(P) = D_N(x_1, \dots, x_N)$ der Punktmenge P ist definiert durch $D_N(P) = D_N(\mathcal{J}; P)$, wo \mathcal{J} eine Familie aller Teilintervalle von I^s der Form $\prod_{i=1}^s [u_i, v_i)$ ist.

Bemerkung 2.1 Sind die Punkte von P aus I^s , dann gilt $D_N^*(P) = D_N(\mathcal{J}_c^*; P)$ und $D_N(P) = D_N(\mathcal{J}_c; P)$, mit \mathcal{J}_c^* der Familie der Teilintervalle von I^s mit der Form $\prod_{i=1}^s [0, u_i]$ und \mathcal{J}_c die Familie der abgeschlossenen Teilintervalle von I^s .

Proposition 2.1 Für jedes P deren Punkte in \bar{I}^s liegen, gilt

$$D_N^*(P) \leq D_N(P) \leq 2^s D_N^*(P).$$

Beweis: Die erste Ungleichung folgt direkt aus der Definition.

Für $s = 1$ betrachten wir $A([u, v]; P) = A([0, v]; P) - A([0, u]; P)$ mit $0 \leq u < v \leq 1$ und $\lambda_1([u, v]) = \lambda_1([0, v]) - \lambda_1([0, u])$ daraus folgt:

$$\left| \frac{A([u, v]; P)}{N} - (v - u) \right| \leq \left| \frac{A([0, v]; P)}{N} - v \right| + \left| \frac{A([0, u]; P)}{N} - u \right|$$

verwenden jetzt noch das Supremum und die Ungleichung folgt.

Für $s = 2$:

Betrachte

$$\begin{aligned} J &= \{(x_1, x_2) \in I^2 : u_1 \leq x_1 < v_1 \text{ und } u_2 \leq x_2 < v_2\} \\ &= [u_1, v_1) \times [u_2, v_2) \text{ mit } 0 \leq u_i < v_i \leq 1, \text{ für } i = 1, 2. \end{aligned}$$

Daraus folgt:

$$\begin{aligned} J &= \{([0, v_1) \times [0, v_2) \setminus ([0, u_1) \times [0, v_2))\} \setminus \{([0, v_1) \times [0, u_2) \setminus ([0, u_1) \times [0, u_2))\} \\ &= (J_1^* \setminus J_2^*) \setminus (J_3^* \setminus J_4^*) \end{aligned}$$

Weiters gilt:

$$\lambda(J) = \lambda(J_1^*) - \lambda(J_2^*) - \lambda(J_3^*) + \lambda(J_4^*) \text{ mit } A(J; P) = A(J_1^*; P) - A(J_2^*; P) - A(J_3^*; P) + A(J_4^*; P)$$

Durch das selbe Argument wie für $s = 1$ folgt die Ungleichung.

Für $s \geq 3$ analog.

□

Es folgt, dass folgende Aussagen äquivalent sind:

- S ist gleichverteilt auf \bar{I}^s ;
- $\lim_{N \rightarrow \infty} D_N(S) = 0$;
- $\lim_{N \rightarrow \infty} D_N^*(S) = 0$.

In diesem Sinn, kann die Diskrepanz und die Sterndiskrepanz als Quantifizierung der Definition von gleichverteilten Folgen in \bar{I}^s gegeben in (2.3) gesehen werden.

Für den Fall $s = 1$ können explizite Formeln für $D_N^*(P)$ und $D_N(P)$ angegeben werden.

Lemma 2.1 *Wenn die $x_1, \dots, x_N, y_1, \dots, y_N \in [0, 1]$, $|x_n - y_n| \leq \epsilon$ für $1 \leq n \leq N$ erfüllen dann gilt:*

$$\begin{aligned} |D_N^*(x_1, \dots, x_N) - D_N^*(y_1, \dots, y_N)| &\leq \epsilon, \\ |D_N(x_1, \dots, x_N) - D_N(y_1, \dots, y_N)| &\leq 2\epsilon \end{aligned}$$

Beweis: Sei P die Punktmenge x_1, \dots, x_N und Q die Punktmenge y_1, \dots, y_N . Weiters betrachten wir $J = [0, u] \subseteq [0, 1]$. Wenn $y_n \in J$, dann ist $x_n \in J_1 := [0, u + \epsilon] \cap [0, 1]$.

Daraus folgt:

$$\frac{A(J; Q)}{N} - \lambda_1(J) \leq \frac{A(J_1; P)}{N} - \lambda_1(J_1) + \epsilon \leq D_N^*(P) + \epsilon.$$

Wenn $x_n \in J_2 := [0, u - \epsilon]$, dann ist $y_n \in J$.

Daraus folgt:

$$\frac{A(J; Q)}{N} - \lambda_1(J) \geq \frac{A(J_2; P)}{N} - \lambda_1(J_2) - \epsilon \geq -D_N^*(P) - \epsilon.$$

Daher gilt $D_N^*(Q) \leq D_N^*(P) + \epsilon$. Durch Vertauschen von P und Q erhalten wir $D_N^*(P) \leq D_N^*(Q) + \epsilon$, woraus $|D_N^*(P) - D_N^*(Q)| \leq \epsilon$ folgt. Die 2. Ungleichung wird analog gezeigt. □

Für $s = 1$ kann man die Punktmenge x_1, \dots, x_N ordnen.

Satz 2.3 Sei $0 \leq x_1 \leq x_2 \leq \dots \leq x_N \leq 1$, dann gilt

$$D_N^*(x_1, \dots, x_N) = \max_{0 \leq n \leq N} \max \left(\left| \frac{n}{N} - x_n \right|, \left| \frac{n}{N} - x_{n+1} \right| \right) = \frac{1}{2N} + \max_{1 \leq n \leq N} \left| x_n - \frac{2n-1}{2N} \right|.$$

Beweis: Da D_N^* eine stetige Funktion von x_1, \dots, x_N (Lemma 2.1), können wir annehmen, dass $0 < x_1 < x_2 < \dots < x_N < 1$ gilt. Weiters setzen wir $x_0 = 0$ und $x_{N+1} = 1$. Sei P die Punktmenge die aus x_1, \dots, x_N besteht, dann gilt

$$\begin{aligned} D_N^*(P) &= \max_{0 \leq n \leq N} \sup_{x_n < u \leq x_{n+1}} \left| \frac{A([0, u]; P)}{N} - u \right| \\ &= \max_{0 \leq n \leq N} \sup_{x_n < u \leq x_{n+1}} \left| \frac{n}{N} - u \right| \\ &= \max_{0 \leq n \leq N} \max \left(\left| \frac{n}{N} - x_n \right|, \left| \frac{n}{N} - x_{n+1} \right| \right) \\ &= \max_{1 \leq n \leq N} \max \left(\left| \frac{n}{N} - x_n \right|, \left| \frac{n-1}{N} - x_n \right| \right) \\ &= \frac{1}{2N} + \max_{1 \leq n \leq N} \left| x_n - \frac{2n-1}{2N} \right|. \end{aligned}$$

□

Satz 2.4 Sei $0 \leq x_1 \leq x_2 \leq \dots \leq x_N \leq 1$, dann gilt:

$$D_N(x_1, \dots, x_N) = \frac{1}{N} + \max_{1 \leq n \leq N} \left(\frac{n}{N} - x_n \right) - \min_{1 \leq n \leq N} \left(\frac{n}{N} - x_n \right).$$

Beweis: Wie im Beweis von Satz 2.3, können wir $x_0 := 0 < x_1 < \dots < x_N < 1 := x_{N+1}$. Sei P die Punktmenge die aus x_1, \dots, x_N besteht, dann gilt

$$\begin{aligned} D_N(P) &= \max_{1 \leq i \leq j \leq N} \sup_{\substack{x_i < u \leq x_{i+1} \\ x_j < v \leq x_{j+1} \\ u < v}} \left| \frac{A([u, v]; P)}{N} - (v - u) \right| \\ &= \max_{1 \leq i \leq j \leq N} \sup_{\substack{x_i < u \leq x_{i+1} \\ x_j < v \leq x_{j+1} \\ u < v}} \left| \frac{j - i}{N} - (v - u) \right| \\ &= \max_{1 \leq i \leq j \leq N} \max \left(\left| \frac{j - i}{N} - (x_{j+1} - x_i) \right|, \left| \frac{j - i}{N} - (x_j - x_{i+1}) \right| \right). \end{aligned}$$

Setze $r_n = n/N - x_n$ für $0 \leq n \leq N + 1$ dann gilt:

$$\begin{aligned} D_N(P) &= \max_{1 \leq i \leq j \leq N} \max \left(\left| r_{j+1} - r_i - \frac{1}{N} \right|, \left| r_j - r_{i+1} + \frac{1}{N} \right| \right) \\ &= \max_{\substack{0 \leq i \leq N \\ 1 \leq j \leq N+1}} \left| \frac{1}{N} + r_i - r_j \right|. \end{aligned}$$

Ist der letzte Term eingeschränkt auf $1 \leq i, j \leq N$, dann ist der Wert eindeutig durch den Ausdruck aus dem Satz für $D_N(P)$ festgelegt. Durch

$$\max_{1 \leq n \leq N} r_n \geq r_N \geq 0, \quad \min_{1 \leq n \leq N} r_n \leq r_1 \leq \frac{1}{N},$$

kann man erkennen, dass im letzten Ausdruck für $D_N(P)$ die Terme für das Maximum entweder $i = 0$ oder $j = N + 1$ durch den Ausdruck aus dem Satz für $D_N(P)$ übertroffen werden. □

Hat man allgemeinere Integrationsbereiche als den \bar{I}^s , muss man neue Arten von Diskrepanzen betrachten. Die folgende ist für allgemeine konvexe Integrationsbereiche.

Definition 2.6 Die isotrope Diskrepanz $J_N(P) = J_N(x_1, \dots, x_N)$ einer Punktmenge P ist definiert durch $J_N(P) = D_N(\mathcal{C}; P)$, wobei \mathcal{C} die Familie aller konvexen Teilmengen von \bar{I}^s ist.

Es gilt folgende Ungleichung:

$$D_N(P) \leq J_N(P) \leq 4sDN(P)^{\frac{1}{s}}$$

Beweis siehe Niederreiter und Wills [7].

Aus dieser Ungleichung kann man erkennen, dass eine Folge S auf \bar{I}^s gleichverteilt ist, dann und nur dann wenn $\lim_{N \rightarrow \infty} J_N(S) = 0$ ist. Wobei $J_N(S)$ die isotrope Diskrepanz der ersten N Glieder von S ist.

Für Jordan-messbare Teilmengen von \bar{I}^s gibt es eine Unterteilung nach der Komplexität ihrer Ränder. Für $B \subseteq \bar{I}^s$ und $\epsilon > 0$, definieren wir

$$B_\epsilon = \{x \in \bar{I}^s : d(x, y) < \epsilon \text{ für endlich viele } y \in B\}$$

$$B_{-\epsilon} = \{x \in \bar{I}^s : d(x, y) \geq \epsilon \text{ für alle } y \in \bar{I}^s \setminus B\}$$

mit d als Euklidische Metrik im \mathbb{R}^s . Sei $b = b(\epsilon)$ eine nicht fallende Funktion definiert für alle $\epsilon > 0$ und erfüllt $\lim_{\epsilon \rightarrow 0+} b(\epsilon) = 0$. Weiters sei \mathcal{M}_b die Familie von allen Lebesgue-messbaren $B \subseteq \bar{I}^s$ für die folgendes gilt:

$$\lambda_s(B_\epsilon \setminus B) \leq b(\epsilon) \quad \text{und} \quad \lambda_s(B \setminus B_{-\epsilon}) \leq b(\epsilon) \quad \text{für alle } \epsilon > 0.$$

Jedes $B \in \mathcal{M}_b$ ist somit Jordan-messbar und umgekehrt jede Jordan-messbare Teilmenge von \bar{I}^s gehört zu \mathcal{M}_b für eine passende Funktion b .

Für \mathcal{M}_b betrachten wir nun die Diskrepanz $D_N(\mathcal{M}_b; P)$. Niederreiter und Wills [7] haben eine Abschätzung für diese Diskrepanz in den Termen von $D_N(P)$ gegeben. Wenn die Funktion b für alle $\epsilon > 0$, $b(\epsilon) \geq \epsilon$ erfüllt, dann kann wie folgt abgeschätzt werden:

$$D_N(\mathcal{M}_b; P) \leq 4b(2\sqrt{s}D_N(P))^{\frac{1}{s}}.$$

In vielen Fällen hat die Funktion b die Form $b(\epsilon) = C\epsilon$ für eine Konstante $C > 0$. Dann folgt die Abschätzung

$$D_N(\mathcal{M}_b; P) \leq (4C\sqrt{s} + 2C + 1)D_N(P)^{\frac{1}{s}}.$$

Da jede konvexe Teilmenge von \bar{I}^s zu \mathcal{M}_{b_0} mit $b_0 = 2s\epsilon$, erhalten wir

$$J_N(P) \leq D_N(\mathcal{M}; P)$$

wenn $b(\epsilon) \geq 2s\epsilon$ für alle $\epsilon > 0$ erfüllt. Wenn deine Folge S gleichverteilt auf \bar{I}^s ist, dann gilt $\lim_{N \rightarrow \infty} D_N(\mathcal{M}_b; S) = 0$, für jede beliebige Funktion b . $D_N(\mathcal{M}_b; S)$ ist die dazugehörige Diskrepanz der ersten N Glieder von S . Unter gewissen Voraussetzungen an die Funktion b gilt auch die Umkehrung: $\lim_{N \rightarrow \infty} D_N(\mathcal{M}_b; S) = 0$ impliziert, dass S gleichverteilt ist auf \bar{I}^s .

2.2 Fehlerschranken

Wir haben bereits die Fehlerschranke für die Quasi-Monte Carlo Approximation (2.1) sowie für allgemeinere Integrationsbereiche diskutiert.

Betrachten wir eingangs den 1-dimensionalen Fall. Ein wichtiges Resultat ist die *Koksma-Ungleichung* [2].

Satz 2.5 *Wenn f eine beschränkte Variation $V(f)$ auf $[0, 1]$ hat, dann gilt für alle $x_1, \dots, x_N \in [0, 1]$:*

$$\left| \frac{1}{N} \sum_{n=1}^N f(x_n) - \int_0^1 f(u) du \right| \leq V(f) D_N^*(x_1, \dots, x_N).$$

Beweis: Wir können annehmen, dass $x_1 \leq x_2 \leq \dots \leq x_N$ gilt. Weiters setzen wir $x_0 = 0$ und $x_{N+1} = 1$. Durch die Verwendung von partieller Summation und partieller Integration erhalten wir:

$$\begin{aligned} \frac{1}{N} \sum_{n=1}^N f(x_n) - \int_0^1 f(u) du &= - \sum_{n=0}^N \frac{n}{N} (f(x_{n+1}) - f(x_n)) + \int_0^1 u df(u) \\ &= \sum_{n=0}^N \int_{x_n}^{x_{n+1}} \left(u - \frac{n}{N}\right) df(u) \end{aligned}$$

für fixes n mit $0 \leq n \leq N$ haben wir

$$\left|u - \frac{n}{N}\right| \leq D_N^*(x_1, \dots, x_N) \quad \text{für } x_n \leq u \leq x_{n+1}$$

durch Satz 2.3 folgt die gewünschte Ungleichung sofort. □

Das Stetigkeitsmaß für eine stetige Funktion f auf $[0, 1]$ ist wie folgt definiert.

$$\omega(f; t) = \sup_{\substack{u, v \in [0, 1] \\ |u-v| \leq t}} |f(u) - f(v)| \quad \text{für } 0 \leq t \leq 1$$

Da f gleichmässig stetig auf \bar{I} ist, gilt $\lim_{t \rightarrow 0+0} \omega(f; t) = 0$.

Die folgenden Abschätzungen für stetige Integranden wurden von Niederreiter [5] aufgestellt.

Satz 2.6 *Ist f eine stetige Funktion auf $[0, 1]$, dann gilt für alle $x_1, \dots, x_N \in [0, 1]$*

$$\left| \frac{1}{N} \sum_{n=1}^N f(x_n) - \int_0^1 f(u) du \right| \leq \omega(f; D_N^*(x_1, \dots, x_N)).$$

Beweis: Wir können wieder $x_1 \leq x_2 \leq \dots \leq x_N$ annehmen. Durch den Mittelwertsatz für Integrale folgt:

$$\int_0^1 f(u) du = \sum_{n=1}^N \int_{(n-1)/N}^{n/N} f(u) du = \frac{1}{N} \sum_{n=1}^N f(t_n)$$

mit $(n-1)/N < t_n < n/N$. Daraus folgt:

$$\frac{1}{N} \sum_{n=1}^N f(x_n) - \int_0^1 f(u) du = \frac{1}{N} \sum_{n=1}^N (f(x_n) - f(t_n)).$$

Nun gilt $|x_n - t_n| \leq D_N^*(x_1, \dots, x_N)$ für $1 \leq n \leq N$ durch Satz 2.3, und das gewünschte Ergebnis folgt. □

Wir wollen nun die Koksma Ungleichung verallgemeinern und sie auch auf mehreren Dimensionen anwenden. Eingangs müssen wir erwähnen unter welchen Bedingungen eine Funktion in mehreren Dimensionen eine beschränkte Variation hat. Betrachten wir eine gegebene Funktion $f(x) = f(x^{(1)}, \dots, x^{(s)})$ aus \bar{I}^s mit $s \geq 2$. Unter einer *Partition* P von \bar{I}^s , verstehen wir eine Menge von s endlichen Folgen $\eta_0^{(j)}, \dots, \eta_{m_j}^{(j)}$ für $1 \leq j \leq s$ mit $0 = \eta_0^{(j)} \leq \eta_1^{(j)} \leq \dots \leq \eta_{m_j}^{(j)} = 1$ für $1 \leq j \leq s$. Dazu definieren wir für jedes $1 \leq j \leq s$, einen Operator Δ_j durch:

$$\begin{aligned} & \Delta_j f(x^{(1)}, \dots, x^{(j-1)}, \eta_i^{(j)}, x^{(j+1)}, \dots, x^{(s)}) \\ &= f(x^{(1)}, \dots, x^{(j-1)}, \eta_{i+1}^{(j)}, x^{(j+1)}, \dots, x^{(s)}) - f(x^{(1)}, \dots, x^{(j-1)}, \eta_i^{(j)}, x^{(j+1)}, \dots, x^{(s)}) \end{aligned} \quad (2.7)$$

für $0 \leq i < m_j$. Δ_{j_1, \dots, j_p} steht für $\Delta_{j_1}, \dots, \Delta_{j_p}$.

Definition 2.7 Für eine Funktion f aus \bar{I}^s setzen wir:

$$V^{(s)}(f) = \sup_P \sum_{i_1=0}^{m_1-1} \dots \sum_{i_s=0}^{m_s-1} \left| \Delta_{1, \dots, s} f(\eta_{i_1}^{(1)}, \dots, \eta_{i_s}^{(s)}) \right|,$$

wo sich das Supremum über alle Partitionen P aus \bar{I}^s erstreckt. Wenn $V^{(s)}(f)$ endlich ist, dann sagt man f hat eine beschränkte Variation im Sinne von Vitali.

Weitere Formulierungen:

$$V^{(s)}(f) = \sup_{\mathcal{P}} \sum_{J \in \mathcal{P}} |\Delta(f : J)|,$$

wobei sich das Supremum über alle Partitionen P von \bar{I}^s erstreckt.

$$V^{(s)}(f) = \int_0^1 \dots \int_0^1 \left| \frac{\partial^s}{\partial u_1 \dots \partial u_s} \right| du_1 \dots du_s \quad (2.8)$$

sie hält wenn die partielle Ableitung stetig auf \bar{I}^s ist.

Es folgt direkt aus der Definition des Δ -Operators, dass wenn die Funktion f aus \bar{I}^s von weniger als s Variablen abhängt, dann ist $V^{(s)}(f) = 0$. Deswegen betrachten wir folgende Definition einer Variation:

Definition 2.8 Sei f eine Funktion mit beschränkter Variation auf \bar{I}^s im Sinne von Vitali. Angenommen die Einschränkung von f auf jeden Bereich F von \bar{I}^s mit der Dimension $1, 2, \dots, s-1$ hat eine beschränkte Variation auf F im Sinne von Vitali, dann hat f eine beschränkte Variation auf \bar{I}^s im Sinne von Hardy und Krause.

Anders formuliert:

$$V(f) = \sum_{k=1}^s \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq s} V^{(k)}(f; i_1, \dots, i_k),$$

Wir führen einen weiteren Operator ein, der für $1 \leq j \leq s$ definiert ist:

$$\Delta_j^* f(x^{(1)}, \dots, x^{(s)}) = f(x^{(1)}, \dots, x^{(j-1)}, 1, x^{(j+1)}, \dots, x^{(s)}) - f(x^{(1)}, \dots, x^{(j-1)}, 0, x^{(j+1)}, \dots, x^{(s)})$$

Sei nun ein Ausdruck $F(r, \dots, r + p - 1; r + p, \dots, l)$ gegeben, der von der Partition der Variablen i_r, \dots, i_l in die Teile $\{i_r, \dots, i_{r+p-1}\}$ und $\{i_{r+p}, \dots, i_l\}$ abhängt, dann steht

$$\sum_{r, \dots, s; p} {}^*F(r, \dots, r + p - 1; r + p, \dots, l)$$

für die Summe aller Ausdrücke die von $F(r, \dots, r + p - 1; r + p, \dots, l)$, durch das sukzessive Ersetzen der gegebenen Partition $\{i_r, \dots, i_l\}$ durch alle anderen Partitionen von dieser Menge in eine Menge aus p und eine Menge aus $l - r - p + 1$ Variablen, abgeleitet wurde. Jede Partition wird genau einmal verwendet. Wenn entweder $p = 0$ oder $p = l - r + 1$ leer ist, dann interpretiert man die Summe, als zu einem Term reduzierte Summe.

Lemma 2.2 *Sei P eine Partition von \bar{I}^s , bestehend aus s Folgen $\eta_0^{(j)}, \dots, \eta_{i_{m_j}}^{(j)}$ für $1 \leq j \leq s$, und Q eine weitere Partition von \bar{I}^s bestehend wieder aus s Folgen $\xi_0^{(j)}, \dots, \xi_{i_{m_{j+1}}}^{(j)}$ für $1 \leq j \leq s$. Weiters seien $f(x)$ und $g(x)$ 2 gegebene Funktionen auf \bar{I}^s . Dann gilt:*

$$\begin{aligned} & \sum_{i_1=0}^{m_1-1} \dots \sum_{i_k=0}^{m_k-1} f(\xi_{i_1+1}^{(1)}, \dots, \xi_{i_s+1}^{(s)}) \Delta_{1, \dots, s} g(\eta_{i_1}^{(1)}, \dots, \eta_{i_s}^{(s)}) \\ &= \sum_{p=0}^s (-1)^p \sum_{1, \dots, s; p} {}^* \Delta_{p+1, \dots, s}^* \sum_{i_1=0}^{m_1} \dots \sum_{i_p=0}^{m_p} g(\eta_{i_1}^{(1)}, \dots, \eta_{i_p}^{(p)}, x^{(p+1)}, \dots, x^{(s)}) \\ & \quad \cdot \Delta_{1, \dots, p} f(\xi_{i_1}^{(1)}, \dots, \xi_{i_p}^{(p)}, x^{(p+1)}, \dots, x^{(s)}). \end{aligned} \quad (2.9)$$

Beweis: siehe [3] S. 148.

Koksma-Hlawka Ungleichung

Satz 2.7 *Sei $f(x)$ eine Funktion mit beschränkter Variation auf \bar{I}^s im Sinne von Hardy und Krause. Weiters sei S eine endliche Folge mit den Punkten x_1, \dots, x_N in $I^s = [0, 1]^s$ und sei s_{j_1, \dots, j_p} die Projektion der Folge S auf den $(k-p)$ -dimensionalen Bereich von \bar{I}^s der wie folgt definiert ist: $x^{(j_1)} = \dots = x^{(j_p)} = 1$. Dann gilt:*

$$\left| \frac{1}{N} \sum_{n=1}^N f(x_n) - \int_{\bar{I}^s} f(x) dx \right| \leq \sum_{p=1}^s \sum_{1, \dots, s; p} {}^* D_N^*(S_{p+1, \dots, s}) V^{(p)}(f(\dots, 1, \dots, 1)),$$

wobei $V^{(p)}(f(\dots, 1, \dots, 1))$ die p -dimensionale Variation von $f(x^{(1)}, \dots, x^{(p)}, 1, \dots, 1)$ auf \bar{I}^p im Sinne von Vitali ist. Wo der Term der Summe $p = s$ ist, versteht man $D_N^*(S) V^{(s)}(f)$. Die Diskrepanz $D_N^*(S_{p+1, \dots, s})$ ist berechnet auf dem Bereich von \bar{I}^s wo sie enthalten ist.

Beweis: Für eine Teilmenge M von I^s , definieren wir die Zählfunktion $A(M; P)$ durch $A(M)$. Weiters definieren wir eine Funktion g auf \bar{I}^s durch

$$g(x) = g(x^{(1)}, \dots, x^{(s)}) = \frac{1}{N} A([0, x^{(1)}) \times \dots \times [0, x^{(s)}]) - x^{(1)} \dots x^{(s)}. \quad (2.10)$$

wir setzen

$$D_N^*(S) = \sup_{x \in I^s} |g(x)|,$$

und

$$D_N^*(S_{p+1,\dots,s}) = \sup_{(x^{(1)},\dots,x^{(p)}) \in \bar{I}^p} \left| g(x^{(1)}, \dots, x^{(p)}, 1, \dots, 1) \right|.$$

Für $1 \leq n \leq N$, setzen wir $x_n = (x_n^{(1)}, \dots, x_n^{(k)})$.

Mit einer zulässigen Partitionierung von \bar{I}^s durch P und Q meinen wir ein Paar (P, Q) .

P besteht aus s Folgen $\eta_0^{(j)}, \dots, \eta_{i_{m_j}}^{(j)}$ für $1 \leq j \leq s$ und Q besteht ebenfalls aus s Folgen $\xi_0^{(j)}, \dots, \xi_{i_{m_j+1}}^{(j)}$ für $1 \leq j \leq s$, die in folgender Relation zueinander stehen

$$0 = \xi_0^{(j)} = \eta_0^{(j)} \leq \xi_1^{(j)} < \eta_1^{(j)} \leq \xi_2^{(j)} < \eta_1^{(j)} \leq \dots < \eta_{i_{m_j}}^{(j)} = \xi_{i_{m_j+1}}^{(j)} = 1 \quad \text{für } 1 \leq j \leq s \quad (2.11)$$

Weiter soll für jedes $1 \leq j \leq s$ die Folge $\xi_0^{(j)}, \dots, \xi_{i_{m_j+1}}^{(j)}$ mindestens die $x_1^{(j)}, \dots, x_N^{(j)}$ enthalten.

Mit dieser zulässigen 2er Partition können wir nun das vorherige Lemma mit den Funktionen f und g (2.10) anwenden. Betrachten wir eingangs die linke Seite von (2.9) und wir erhalten:

$$\begin{aligned} & \sum_{i_1=0}^{m_1-1} \dots \sum_{i_k=0}^{m_k-1} f(\xi_{i_1+1}^{(1)}, \dots, \xi_{i_s+1}^{(s)}) \Delta_{1,\dots,s} g(\eta_{i_1}^{(1)}, \dots, \eta_{i_s}^{(s)}) \\ &= \frac{1}{N} \sum_{i_1=0}^{m_1-1} \dots \sum_{i_k=0}^{m_k-1} f(\xi_{i_1+1}^{(1)}, \dots, \xi_{i_s+1}^{(s)}) \Delta_{1,\dots,s} A([0, \eta_{i_1}^{(1)}] \times \dots \times [0, \eta_{i_s}^{(s)}]) \\ & \quad - \sum_{i_1=0}^{m_1-1} \dots \sum_{i_k=0}^{m_k-1} f(\xi_{i_1+1}^{(1)}, \dots, \xi_{i_s+1}^{(s)}) \Delta_{1,\dots,s} \eta_{i_1}^{(1)} \dots \eta_{i_s}^{(s)}. \end{aligned} \quad (2.12)$$

Nun:

$$\begin{aligned} & \Delta_{1,\dots,s} A([0, \eta_{i_1}^{(1)}] \times \dots \times [0, \eta_{i_s}^{(s)}]) \\ &= \Delta_{1,\dots,s-1} A([0, \eta_{i_1}^{(1)}] \times \dots \times [0, \eta_{i_{s+1}}^{(s)}]) - A([0, \eta_{i_1}^{(1)}] \times \dots \times [0, \eta_{i_s}^{(s)}]) \\ &= \Delta_{1,\dots,s-1} A([0, \eta_{i_1}^{(1)}] \times \dots \times [0, \eta_{i_{s-1}}^{(s-1)}] \times [\eta_{i_s}^{(s)}, \eta_{i_{s+1}}^{(s)}]) \\ &= \dots = A([\eta_{i_1}^{(1)}, \eta_{i_{1+1}}^{(1)}] \times \dots \times [\eta_{i_s}^{(s)}, \eta_{i_{s+1}}^{(s)}]). \end{aligned}$$

Somit ist der erste Term auf der rechten Seite von (2.12) reduziert zu

$$\frac{1}{N} \sum_{i_1=0}^{m_1-1} \dots \sum_{i_k=0}^{m_k-1} f(\xi_{i_1+1}^{(1)}, \dots, \xi_{i_s+1}^{(s)}) A([\eta_{i_1}^{(1)}, \eta_{i_{1+1}}^{(1)}] \times \dots \times [\eta_{i_s}^{(s)}, \eta_{i_{s+1}}^{(s)}]). \quad (2.13)$$

Es haben nur diese s -Tupel (i_1, \dots, i_s) Bedeutung die ein x_n für $1 \leq n \leq N$ im Intervall $[\eta_{i_1}^{(1)}, \eta_{i_{1+1}}^{(1)}] \times \dots \times [\eta_{i_s}^{(s)}, \eta_{i_{s+1}}^{(s)}]$ haben. Sollte dieser Fall eintreffen, dann folgt aus (2.11) und der zusätzlichen Bedingung an Q , dass $x_n = (\xi_{i_1+1}^{(1)}, \dots, \xi_{i_s+1}^{(s)})$ ist. Daher ist (2.13) nichts anderes als $(\frac{1}{N}) \sum_{n=1}^N f(x_n)$. Daraus folgt für die linke Seite von (2.9):

$$\frac{1}{N} \sum_{n=1}^N f(x_n) - \sum_{i_1=0}^{m_1-1} \dots \sum_{i_s=0}^{m_s-1} f(\xi_{i_1+1}^{(1)}, \dots, \xi_{i_s+1}^{(s)}) \Delta_{1,\dots,s} \eta_{i_1}^{(1)} \dots \eta_{i_s}^{(s)}. \quad (2.14)$$

Wir brauchen, dass $g(x) = 0$ ist, wenn mindestens eine Koordinate verschwindet und außerdem soll $g(1, \dots, 1) = 0$ sein. Wenn $p = 0$ ist in diesem Term, dann stimmt er mit der rechten Seite von (2.9) überein. Nämlich $\Delta_{1, \dots, s}^* g(x^{(1)}, \dots, x^{(s)}) f(x^{(1)}, \dots, x^{(s)})$ ist daher 0. Außerdem für $1 \leq p \leq s$ bleiben nur diese Terme über, wo die Variablen $x^{(p+1)}, \dots, x^{(s)}$ durch 1 ersetzt werden. Woraus folgt, dass die rechte Seite von (2.9) zu:

$$\sum_{p=1}^s (-1)^p \sum_{1, \dots, s; p}^* \sum_{i_1=0}^{m_1} \dots \sum_{i_p=0}^{m_p} g(\eta_{i_1}^{(1)}, \dots, \eta_{i_p}^{(p)}, 1, \dots, 1) \cdot \Delta_{1, \dots, p} f(\xi_{i_1}^{(1)}, \dots, \xi_{i_p}^{(p)}, 1, \dots, 1). \quad (2.15)$$

Wir haben nun einen Zusammenhang zwischen (2.14) und (2.15) gefunden. Nun wollen wir den Absolutwert von (2.15) abschätzen. Eine obere Schranke ist natürlich durch:

$$\sum_{p=1}^s \sum_{1, \dots, s}^* \sum_{i_1=0}^{m_1} \dots \sum_{i_p=0}^{m_p} \left| g(\eta_{i_1}^{(1)}, \dots, \eta_{i_p}^{(p)}, 1, \dots, 1) \right| \cdot \left| \Delta_{1, \dots, p} f(\xi_{i_1}^{(1)}, \dots, \xi_{i_p}^{(p)}, 1, \dots, 1) \right|.$$

Der Absolutwert bringt mit sich, dass g gleichmäßig beschränkt durch $D_N^*(S_{p+1, \dots, s})$ ist. Die restliche Summe über i_1, \dots, i_p ist dominiert durch $V^{(p)}(f(\dots, 1, \dots, 1))$.

Woraus folgt:

$$\begin{aligned} & \left| \frac{1}{N} \sum_{n=1}^N f(x_n) - \sum_{i_1=0}^{m_1-1} \dots \sum_{i_s=0}^{m_s-1} f(\xi_{i_1+1}^{(1)}, \dots, \xi_{i_s+1}^{(s)}) \Delta_{1, \dots, s} \eta_{i_1}^{(1)} \dots \eta_{i_s}^{(s)} \right| \\ & \leq \sum_{p=1}^s \sum_{1, \dots, s; p}^* D_N^*(S_{p+1, \dots, s}) V^{(p)}(f(\dots, 1, \dots, 1)). \end{aligned} \quad (2.16)$$

Wir bemerken, dass $\Delta_{1, \dots, s} \eta_{i_1}^{(1)} \dots \eta_{i_s}^{(s)} = (\eta_{i_1+1}^{(1)} - \eta_{i_1}^{(1)}) \dots (\eta_{i_s+1}^{(s)} - \eta_{i_s}^{(s)})$, und daraus folgt, dass (2.11), die Summe über i_1, \dots, i_s auf der linken Seite von (2.16) nichts anderes ist als eine Riemann Summe für $\int_{\bar{I}^s} f(x) dx$, impliziert. Der andere Term in (2.16) ist unabhängig von der gewählten 2er Partition (P, Q) . Der Beweis wird dadurch komplettiert, dass man (P, Q) durch eine Folge von zulässigen 2er Partitionen mit

$$\max_{1 \leq j \leq s} \max_{0 \leq i < m_j} (\eta_{i+1}^{(j)} - \eta_i^{(j)}) \rightarrow 0$$

durchlaufen lässt.

□

Satz 2.8 Für alle $x_1, \dots, x_N \in I^s$ und für jedes beliebige $\epsilon > 0$, existiert eine Funktion $f \in C^\infty(\bar{I}^s)$ mit $V(f) = 1$ und

$$\left| \frac{1}{N} \sum_{n=1}^N f(x_n) - \int_{\bar{I}^s} f(u) du \right| < D_N^*(x_1, \dots, x_N) - \epsilon.$$

Beweis: Durch die Definition von $D_N^*(P) = D_N^*(x_1, \dots, x_N)$, existiert ein Intervall $J = \prod_{i=1}^s [0, v_i)$ mit $0 < v_i \leq 1$ für $1 \leq i \leq s$ und

$$\left| \frac{A(J; P)}{N} - \lambda_s(J) \right| > D_N^*(P) - \frac{\epsilon}{2}.$$

Weiters existiert ein Intervall $K = \prod_{i=1}^s [0, t_i]$ mit $0 \leq t_i \leq v_i$ und $v_i - t_i < \epsilon/(2s)$ für $1 \leq i \leq s$ so, dass $J \setminus K$ kein x_n enthält. Für gegebene $0 \leq t < v \leq 1$, können wir eine nichtsteigende Funktion $f_{t,v} \in C^\infty([0, 1])$ mit $f_{t,v}(u) = 1$ für $0 \leq u \leq t$ und $f_{t,v}(u) = 0$ für $v \leq u \leq 1$.

Dann ist

$$f(u) = f(u_1, \dots, u_s) = \prod_{i=1}^s f_{t_i, v_i}(u_i)$$

eine Funktion in $C^\infty([0, 1]^s)$ mit $0 \leq f(u) \leq 1$ für alle $u \in \bar{I}^s$, $f(u) = 1$ für $u \in K$, und $f(u) = 0$ für $u \notin J$. Aus der Gleichung (2.8) bekommen wir $V^{(s)}(f) = 1$. Da $f(u) = 0$ wenn eine Koordinate von $u = 1$ ist. Wir haben $V^{(k)}(f; i_1, \dots, i_k) = 0$ wenn $1 \leq k < s$ ist. Woraus wir $V(f) = 1$ erhalten. Es ist ersichtlich, dass

$$\frac{1}{N} \sum_{n=1}^N f(x_n) = \frac{A(J; P)}{N}$$

und dass

$$\lambda_s(K) \leq \int_{\bar{I}^s} f(u) du \leq \lambda_s(J).$$

Da $\lambda_s(J) - \lambda_s(K) \leq \sum_{i=1}^s (v_i - t_i) < \epsilon/2$, folgt, dass

$$\left| \frac{1}{N} \sum_{n=1}^N f(x_n) - \int_{\bar{I}^s} f(u) du \right| > \left| \frac{A(J; P)}{N} - \lambda_s(J) \right| - \frac{\epsilon}{2} > D_N^*(P) - \epsilon.$$

□

Es gibt auch eine mehrdimensionale Version des Satzes 2.6. Für eine stetige Funktion f auf \bar{I}^s definieren wir folgendes Stetigkeitsmaß

$$\omega(f; t) = \sup_{\substack{u, v \in \bar{I}^s \\ \|u-v\| \leq t}} |f(u) - f(v)| \quad \text{für } t \geq 0 \quad (2.17)$$

wo $\|u\| = \max_{1 \leq i \leq s} |u_i|$ für $u = (u_1, \dots, u_s) \in \mathbb{R}^s$. Die folgende Abschätzung geht auf Proinov [8] zurück.

Satz 2.9 *Ist f stetig auf \bar{I}^s , dann gilt für alle $x_1, \dots, x_N \in \bar{I}^s$:*

$$\left| \frac{1}{N} \sum_{n=1}^N f(x_n) - \int_{\bar{I}^s} f(u) du \right| \leq 4\omega(f; D_N^*(x_1, \dots, x_N)^{1/s}).$$

ohne Beweis

All diese Fehlerschranken laufen daraus hinaus, dass eine kleine Sterndiskrepanz einen kleinen Integrationsfehler der Quasi-Monte Carlo Methode garantiert. Wie man solche Punktfolgen konstruiert wird Herr Scheibelhofer in seinem Vortrag näher erläutern.

Betrachten wir noch etwas allgemeinere Integrationsbereiche. Durch Translation oder Kontraktion können wir annehmen, dass der Integrationsbereich B in \bar{I}^s liegt. Die folgende Fehlerschranke (in Termen von der isotropischen Diskrepanz) wurde von Zaremba [9] bewiesen.

Satz 2.10 Wenn $B \subseteq \bar{I}^s$ konvex ist und f eine beschränkte Variation $V(f)$ auf \bar{I}^s im Sinne von Hardy und Krause hat, dann gilt für jede Punktmenge P die aus $x_1, \dots, x_N \in \bar{I}^s$ besteht,

$$\left| \frac{1}{N} \sum_{\substack{C \\ x_n \in Bn=1}}^N f(x_n) - \int_B f(u) du \right| \leq (V(f) + |f(1, \dots, 1)|) J_N(P).$$

ohne Beweis

Die allgemeinste Situation ist, wenn wir einen B betrachten das aus einer Familie Jordanmessbarer Funktionen stammt. Der folgende Satz ist eine Verbesserung von de Clerck [1] und eine Idee von Niederreiter [6].

Satz 2.11 Wenn $B \in \mathcal{M}_b$ ist und f eine beschränkte Variation $V(f)$ auf \bar{I}^s im Sinne von Hardy und Krause hat, dann gilt für jede Punktmenge P die aus $x_1, \dots, x_N \in I^s$ besteht,

$$\left| \frac{1}{N} \sum_{\substack{C \\ x_n \in Bn=1}}^N f(x_n) - \int_B f(u) du \right| \leq (V(f) + |f(1, \dots, 1)|) D_N(\mathcal{M}_b; P).$$

ohne Beweis

Der Ausdruck $|f(1, \dots, 1)|$ in den letzten beiden Sätzen ist notwendig um auch für konstante Funktionen f zu halten.

Bemerkung: Der Aufbau und die Aussagen dieses Kapitels sind analog dem Kapitel 2 aus H. Niederreiter [4] und den Aussagen von L. Kuipers und H. Niederreiter [3] über die eindimensionale Diskrepanz und die Koksma-Hlawka Ungleichung.

Literaturverzeichnis

- [1] L. DE Clerck. A proof of Niederreiter's conjecture concerning error bounds for quasi-Monte Carlo integration. *Adv. in Appl. Math.* 2(1981), **pp. 1-6**.
- [2] J.F. Koksma. Een algemeene stelling uit de theorie der gelijkmatige verdeling modulo 1. *Mathematica B (Zutphen)* **26** (1942/1943),**pp. 7-11**.
- [3] L. Kuipers und H. Niederreiter. Uniform Distribution of Sequences. *John Wiley and Sons NY*, 1974.
- [4] H. Niederreiter. Random Number Generation and Quasi-Monte-Carlo Methods *Society for Industrial and Applied Mathematics Philadelphia, Pennsylvania*, 1992.
- [5] H. Niederreiter. *Methods for estimationg discrepancy*, in Applications of Number Theory to Numerical Analysis, S.K. Zraemba, ed., *Academic Press, New York*, 1972, **pp. 203-236**.
- [6] H. Niederreiter. *Application of diophantine approximations to numerical integration*, in Diophantine Approximation and Its Applications, *C.F. Osgood, ed., Academic Press, New York*, 1973 **pp. 129-199**.
- [7] H. Niederreiter und J.M. Wills. Diskrepanz und Distanz von Maßen bezüglich konvexer und Jordanscher Mengen. *Matz. Z.* 144 (1975),**pp. 125-134**; Berichtigung, *ibid.*, 148 (1976), **pp. 99**.
- [8] P.D. Proinov. Discrepancy and integration of continuous functions. *J. Approx. Theory*, 52 (1988),**pp. 121-131**.
- [9] Zaremba. la discrepance isotrope et l'integration numérique. *Ann. Mat. Pura Appl.*, 87 (1970),**pp. 125-136**.