Volumenabschätzung konvexer Körper

Seminar zur Stochastischen Analysis

Prof. Dr. Andreas Eberle
Referent: Harald G. Grohganz

1. *Inhaltsübersicht*
2. *Einführung und Problematik*
   - Motivation
   - Wiederholung
3. *Konvergenz des ball walk*
   - Stationäre Verteilungen des ball walk
   - Konvergenz ins Gleichgewicht
   - Beweis der Poincaré-Ungleichung
4. *Anwendung und Ausblick*
   - Implementierung
   - Volumenberechnung
   - Ausblick
5. *Literatur*

Problemstellung

Gegeben sei ein konvexer Körper $K \subseteq \mathbb{R}^n$, wobei n typischerweise groß ist. Eine direkte Berechnung des Volumens ist auf Grund der hohen Dimension praktisch nicht möglich, wir wählen daher ein Monte-Carlo-Verfahren.

Hierbei müssen wir das Problem im zwei Schritte unterteilen:

- Simuliere gleichverteilte Punkte auf K;
- Schätze das Volumen $Vol_n\, K$ von K ab.

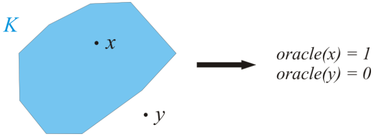
Hierbei ist vor allem die Simulation interessant, die Volumenabschätzung ergibt sich als Grenzwert der Summation genügend vieler Punkte und kann daher auf das erste Problem zurückgeführt werden.

Methodik: Orakelfunktion

Wir nehmen an, dass wir K durch eine Orakelfunktion gegeben haben, die für ein gegebenes $x \in \mathbb{R}^n$ angibt, ob $x \in K$ gilt. Ein Beispiel für eine Orakelfunktion ist die Indikatorfunktion:

$$oracle(x) = \begin{cases} 1 & x \in K \\ 0 & x \notin K \end{cases} .$$

In der Praxis geschieht dies oft durch Lineare Programmierung, aber es gibt auch Fälle, in denen ein exaktes Orakel nicht möglich ist.



Methodik: Fluch der Dimension

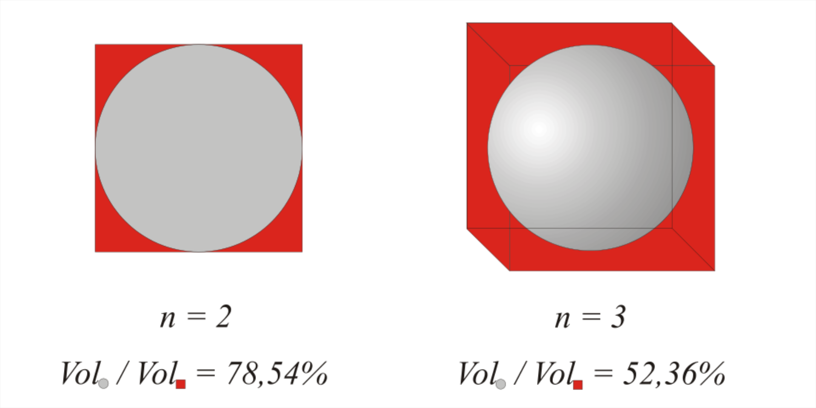
Ebenfalls ist ein direktes Monte-Carlo-Verfahren nicht möglich, da K im Vergleich zum möglichen Samplingbereich verschwindet gering sein kann.

Sei etwa $K = B_n(0,1)$ die Einheitskugel und $C = [-1,1]^n$ der kleinste Würfel, der K enthält. Hier kann $\operatorname{Vol}_n K$ exakt berechnet werden und wir erhalten

$$\frac{\operatorname{Vol}_n K}{\operatorname{Vol}_n C} = \frac{\frac{\pi^{n/2}}{n \cdot \Gamma(n/2)}}{2^n} \to 0$$

exponentiell schnell.

Abbildung: Fluch der Dimension



Methodik: Der ball walk

Daher nutzen wir das Prinzip des *ball walk*:

Gegeben sei ein Punkt $X_t \in K$, welcher der Position des random walk zur Zeit t entspricht. Wir wählen X_{t+1} i.i.d. aus $B(X_t, \delta) \cap K$, wobei $B(x, r)$ die n-dim. Kugel um x mit dem Radius δ beschreibt. Die geeignete Wahl von δ ergibt sich aus zusätzlichen Informationen, die wir über K haben.

Wir werden zeigen, dass diese Markovkette eine stationäre Verteilung hat, welche annähernd der Gleichverteilung auf K entspricht, und dass die Konvergenzrate polynomiell in n ist.

Algorithmen: Direkt vs. Metropolis

Warnung: Theorie und Praxis

Für unsere weiteren Betrachtungen müssen wir feststellen, dass der oben beschriebene Algorithmus, d.h. die Wahl eines Punktes $X_{t+1} \in B(X_t, \delta) \cap K$ i.i.d. für ein allgemeines K nicht implementierbar ist. Dennoch werden wir uns in der Theorie oft auf diesen *direkten Algorithmus* beziehen.

In der Praxis verwenden wir eine Art *Metropolis*-Algorithmus:

`Wähle` $y \in B(X_t, \delta)$ `i.i.d.`
`Wenn` $oracle(y) = 1$, `setze` $X_{t+1} := y$.
`Sonst setze` $X_{t+1} := X_t$.

Wir sehen später, dass die *Metropolis*-Variante nur um einen konstanten Faktor langsamer als die *direkte* Variante ist.

Notation bei raumkontinuierlichen Markovketten

Dadurch, dass bei kontinuierlichem Zustandsraum (mit diskreter Zeit) die Übergangswahrscheinlichkeit von einem Punkt $x \in K$ zu einem Punkt $y \in K$ fast sicher 0 beträgt, betrachten wir hier die Wahrscheinlichkeit, dass wir zur Zeit 1 in eine messbare Menge A gesprungen sind, m.a.W. $P[x, A] = \mathbb{P}[X_1 \in A \mid X_0 = x]$.

Die Wahrscheinlichkeit für Schritt t definieren wir induktiv mit $P^1 = P$ und

$$P^t(x, A) := \int_K P^{t-1}(x, dy)\, P(y, A) \qquad \text{für } t > 1.$$

Notation bei raumkontinuierlichen Markovketten

Im Falle des *ball walk* erhalten wir

$$P(x, A) = \frac{\mathrm{Vol}_n(B(x, \delta) \cap A)}{\mathrm{Vol}_n(B(x, \delta) \cap K)}$$

für jede messbare Menge $A \subseteq K$ und

$$P(x, dy) = \frac{dy}{\mathrm{Vol}_n(B(x, \delta) \cap K)}, \qquad (1)$$

gegeben $y \in B(x, \delta) \cap K$.

Stationäre Verteilung

Definition

Eine Wahrscheinlichkeitsverteilung μ heißt *stationär* oder *invariant* für P, wenn sie folgende Bedingung erfüllt:

$$\mu(A) = \int_K P(x, A)\ \mu(dx)$$

für alle messbaren Mengen $A \in K$.
Die stationäre Verteilung ist *eindeutig*, wenn $P^t(x, A) \stackrel{t\to\infty}{\longrightarrow} \mu(A)$ für alle x und alle A gilt.

Erste Anforderungen an δ

Um sinnvolle Aussage über stationäre Verteilungen treffen und somit mit den vorgestellten Algorithmen arbeiten zu können, ist es notwendig, eine Bedingung an δ aufzustellen:

$$\delta \leq \sup\{\|x - y\| \ : \ x, y \in K\},$$

d.h. δ muss kleiner sein als der Durchmesser von K, da wir sonst im ersten Schritt bereits die Gleichverteilung als stationäre Verteilung erhalten – somit haben wir das Problem, einen Punkt i.i.d. auf ganz K zu erzeugen. Dies ist aber exakt das Ausgangsproblem!

Für die Anwendung des Metropolis-Algorithmus sollte ferner das Verhältnis

$$\frac{\mathrm{Vol}_n\, B(X_t, \delta) \cap K}{\mathrm{Vol}_n\, B(X_t, \delta)}$$

nicht zu klein sein.

Krümmungsbedingung

Ist K sehr lang und dünn, wird die Konvergenzrate des Metropolis-Algorithmus signifikant sinken: δ muss klein genug sein, dass der ball walk überhaupt Punkte in K trifft – aber andererseits groß genug, dass der Prozess ganz K erreichen kann.

Für das weitere Vorgehen stellen wir daher eine Krümmungsbedingung auf, die insbesondere in der Theorie wesentliche Vereinfachungen zulässt:

$$\forall\, x \in K \;\exists\, y \in K \;:\; x \in B(y, 1) \subseteq K. \tag{2}$$

Anschaulich bedeutet dies, dass wir ganz K mit 1-Bällen überlagern können.

Lokale Leitfähigkeit

Die gesuchte stationäre Verteilung ist bei der direkten Methode zwar nicht gleichverteilt, für kleines δ jedoch annähernd. Um die stationäre Verteilung zu beschreiben, definieren wir die Funktion $\ell : K \to \mathbb{R}$ wie folgt:

$$\ell(x) := \frac{\mathrm{Vol}_n\, B(x,\delta) \cap K}{\mathrm{Vol}_n\, B(x,\delta)} \tag{3}$$

ℓ heißt *lokale Leitfähigkeit* und beschreibt anschaulich die Wahrscheinlichkeit, dass ein zufälliger Punkt aus dem δ-Ball um x in K liegt.

Wir normalisieren $\ell(x)$, um eine Dichtefunktion zu erhalten, welche sich als die Dichte der stationären Verteilung des *ball walk* erweisen wird:

$$\mu(A) := \frac{\int_A \ell(x)dx}{\int_K \ell(x)dx}. \tag{4}$$

μ ist stationäre Verteilung

Lemma

Wenn X_0 μ-verteilt ist, dann gilt dies auch für X_1.

Zum Beweis bezeichne μ_1 die Verteilung von X_1 und wir setzen $L := \int_K \ell(x)dx$.

Beweis des ersten Lemmas

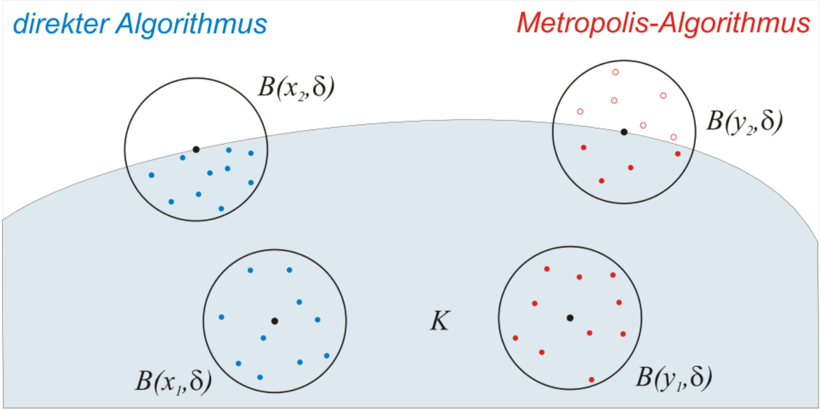
$$
\begin{aligned}
\mu_1(A) = \int_A \mu_1(dy) &= \int_A \int_K P(x,dy)\mu(dx) \\
&= \int_A dy \int_{B(y,\delta)\cap K} \frac{\mu(dx)}{\mathrm{Vol}_n B(x,\delta)\cap K} \quad \text{(mit 1)} \\
&= \frac{1}{L}\int_A dy \int_{B(y,\delta)\cap K} \frac{\ell(x)dx}{\mathrm{Vol}_n B(x,\delta)\cap K} \quad \text{(mit 4)} \\
&= \frac{1}{L}\int_A dy \int_{B(y,\delta)\cap K} \frac{dx}{\mathrm{Vol}_n B(x,\delta)} \quad \text{(mit 3)} \\
&= \frac{1}{L}\int_A dy \, \frac{\mathrm{Vol}_n B(y,\delta)\cap K}{\mathrm{Vol}_n B(y,\delta)} \\
&= \frac{1}{L}\int_A \ell(y)dy = \mu(A) \quad \text{(mit 3, 4)}
\end{aligned}
$$

qed.

Bemerkung: Gleichvert. stationär bei Metropolis

Analog zu oben zeigt man, dass die Gleichverteilung eine stationäre Verteilung für die *Metropolis*-Variante ist.

Anschaulich ist dies klar:



Laufzeitvergleich Metropolis vs. direkt

Lemma

Wir nehmen die Krümmungsbedingung (2) an und wählen $\delta \leq \frac{c_1}{\sqrt{n}}$, wobei c_1 eine dimensionsunabhängige Konstante ist. Dann gilt:

$$\frac{4}{10} \leq \ell(x) \;\leq\; 1$$

für alle $x \in K$.

Beweis des zweiten Lemmas

Die obere Schranke ist trivial.

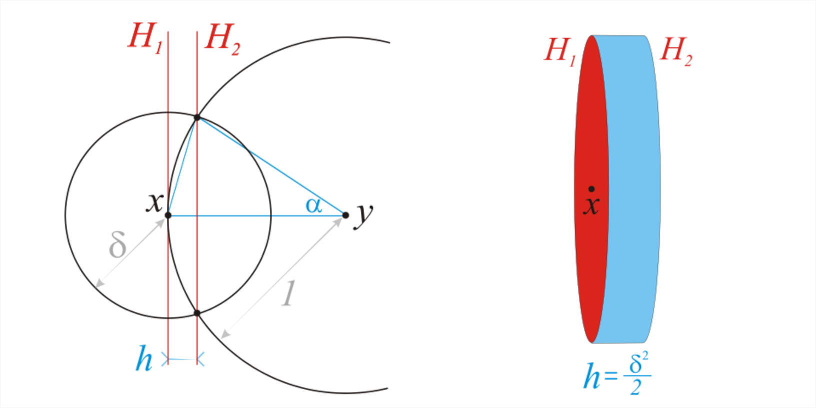
Unter Annahme der Krümmungsbedingung liegt jedes $x \in K$ in einem 1-Ball $B(y,1) \subseteq K$. Wir zeigen daher

$$\frac{\mathrm{Vol}_n\, B(x,\delta) \cap B(y,1)}{\mathrm{Vol}_n\, B(x,\delta)} \geq \frac{4}{10}.$$

Es genügt zu zeigen, dass dies für x auf dem Rand von $B(y,1)$ gilt.

Wir betrachten die Tangentialebene H_1 zu $B(y,1)$ durch x und eine parallele Hyperebene H_2 durch die beiden Schnittpunkte der Bälle.

Graphik zum Beweis des zweiten Lemmas



Fortsetzung des Beweises

Mit H_i^+ bezeichnen wir die Seite der Hyperebenen, die den Punkt y enthält. Es gilt

$$B(x,\delta) \cap H_2^+ \subseteq B(y,1)$$

für $\delta < 1$. Also genügt zu zeigen:

$$\mathrm{Vol}_n\, B(x,\delta) \cap H_2^+ \;\leq\; \frac{4}{10} \cdot \mathrm{Vol}_n\, B(x,\delta)$$

Wir zeigen daher, dass $B(x,\delta) \cap H_2^- \cap H_1^+$ höchstens $1/10$ des Volumens von $B(x,\delta)$ hat.

Dazu betrachten wir den Zylinder mit Grundfläche $B(x,\delta) \cap H_1$, dessen Höhe der Abstand von H_1 und H_2 ist. Diese Höhe beträgt genau $\delta^2/2$.

Fortsetzung des Beweises

Mit der Volumenformel für n-dim. Kugeln sowie die Stirlingsche Abschätzung der Γ-Funktion erhalten wir

$$\begin{aligned}
\frac{\mathrm{Vol}_{n-1} B(x,\delta) \cap H_1}{\mathrm{Vol}_n B(x,\delta)} &= \frac{\delta^{n-1}\pi^{\frac{n-1}{2}} \cdot n \cdot \Gamma\left(\frac{n}{2}\right)}{\delta^n \pi^{\frac{n}{2}} \cdot (n-1) \cdot \Gamma\left(\frac{n-1}{2}\right)} \\
&< \frac{n\sqrt{2\pi}\left(\frac{n+2}{2}\right)^{\frac{n+3}{2}} e^{-\frac{n-2}{2}}}{\sqrt{\pi}\delta(n-1) \cdot \sqrt{2\pi}\left(\frac{n+1}{2}\right)^{\frac{n+2}{2}} e^{-\frac{n-1}{2}}} \\
&= \frac{\sqrt{e}}{\sqrt{2\pi}\delta} \cdot \frac{n}{n-1} \cdot \sqrt{\frac{(n+2)^{n+3}}{(n+1)^{n+2}}} \\
&\overset{n>2}{\leq} \frac{\sqrt{e}}{\sqrt{2\pi}\delta} \cdot e \cdot 2^2 \cdot \sqrt{n} \;=:\; \frac{c\sqrt{n}}{\delta}
\end{aligned}$$

Abschluss des Beweises

Somit gilt:

$$\frac{\mathrm{Vol}_n(\text{Zylinder})}{\mathrm{Vol}_n B(x,\delta)} \leq \frac{1}{2} c\delta\sqrt{n}.$$

Setzen wir

$$c_1 = \frac{1}{5c} \quad \text{(unabhängig von } n \text{ !)}$$

erhalten wir die gewünschte Abschätzung.

qed.

"Partition der Varianz"

Wir zeigen nun, dass der *ball walk* schnell gegen die stationäre Verteilung konvergiert.

Lemma

Sei f eine messbare Funktion auf einer messbaren Menge M und $M_0, \ldots, M_{n-1}$ eine Partition von M. Dann gilt:

$$\int_M f^2 \, d\mu = \sum_{i=0}^{n-1} \int_{M_i} (f - \bar{f}_i)^2 \, d\mu + \sum_{i=0}^{n-1} \mu(M_i) \bar{f}_i^2 \ ,$$

wobei

$$\bar{f}_i = \frac{1}{\mu(M_i)} \int_{M_i} f \, d\mu.$$

Beweis des dritten Lemmas

$$\begin{aligned}
&\int_{M_i} (f - \bar{f}_i)^2 d\mu + \mu(M_i)\bar{f}_i^2 \\
&= \int_{M_i} f^2\, d\mu + \int_{M_i} \bar{f}_i^2\, d\mu - 2\int_{M_i} \bar{f}_i f\, d\mu + \mu(M_i)\bar{f}_i^2 \\
&= \int_{M_i} f^2\, d\mu + \mu(M_i)\bar{f}_i^2 - 2\mu(M_i)\bar{f}_i^2 + \mu(M_i)\bar{f}_i^2 \\
&= \int_{M_i} f^2\, d\mu.
\end{aligned}$$

qed.

Notationen

Um die Konvergenzrate abzuschätzen, betrachten wir den Abfall der Varianz einer Testfunktionen f (mit $\mathbb{E}f = 0$ nach Konvention). Hierfür betrachten wir eine Funktion $h : K \to \mathbb{R}$ mit

$$
\begin{aligned}
h(x) &:= \frac{1}{2}\int_K P(x,dy)(f(x)-f(y))^2 \\
&= \frac{1}{2\,\mathrm{Vol}_n B(x,\delta)\cap K}\int_{B(x,\delta)\cap K}(f(x)-f(y))^2 dy \qquad (5)
\end{aligned}
$$

und führen folgende Bezeichnungen ein:

$$
\begin{aligned}
\mathrm{Var}_\mu f &:= \int_K f^2 d\mu \quad \text{(Varianz auf } K\text{)} \\
\mathcal{E}_P(f,f) &:= \int_K h\, d\mu \quad \text{(Dirichletform, "lokale Variation")}
\end{aligned}
$$

Die Poincaré-Ungleichung

Theorem (Poincaré-Ungleichung)

Sei $K \subset \mathbb{R}^n$ ein konvexer Körper mit Durchmesser D, der die Krümmungsbedingung (2) erfüllt. Wähle $\delta \leq c_1/\sqrt{n}$ wie im zweiten Lemma. Dann gilt für jede messbare Funktion $f : K \to \mathbb{R}$:

$$\mathcal{E}_P(f,f) \;\geq\; \lambda \cdot \mathrm{Var}_\mu f \;,$$

wobei gilt:

$$\lambda = \frac{c_2 \cdot \delta^2}{D^2 \cdot n}$$

für eine dimensionsunabhängige Konstante c_2.

Von der Ungleichung zur Laufzeit

Aus der Poincaré-Ungleichung lässt sich eine Beschränkung der Laufzeit herleiten:

Theorem (Laufzeitabschätzung)

Für ein $\varepsilon > 0$ sei $\tau(\varepsilon)$ die vom ball walk benötigte Zeit, die stationäre Verteilung μ bis auf Variationsnorm ε zu erreichen. Unter der Krümmungsbedingung (2) gilt:

$$\tau(\varepsilon) \ \leq \ O(\lambda^{-1}(\log \varepsilon^{-1} + i(\mu_0))),$$

wobei $i(\mu_0)$ die Abhängigkeit von der Startverteilung beschreibt und λ wie in der Poincaré-Ungleichung bestimmt ist.

In der Praxis ist $i(\mu_0)$ etwa in der Größenordnung $n \log(D/2\delta)$. (ohne Beweis)

Beweis: Grundidee

Wir beweisen nun die Poincaré-Ungleichung mit einem Widerspruchsbeweis. Annahme: Es existiert ein $f : K \to \mathbb{R}$ mit

$$\mathcal{E}_P(f,f) \; < \; \lambda \cdot \mathrm{Var}_\mu f \; . \tag{6}$$

Nun suchen wir kleiner werdende *schlechte Mengen* M, bei denen der Quotient

$$\frac{\int_M h \, d\mu}{\int_M (f - \bar{f})^2 d\mu} \tag{7}$$

klein ist, wobei $\bar{f} = \int_M f \, d\mu$. Setzen wir $M = K$, so erhalten wir

$$\frac{\int_K h \, d\mu}{\int_K (f - \bar{f})^2 d\mu} = \frac{\mathcal{E}_P(f,f)}{\mathrm{Var}_\mu f} < \lambda.$$

Wir versuchen, auch in Abhängigkeit von δ *kleine* M zu bekommen – denn dann muss f nahezu konstant auf M sein, obwohl die globale Varianz groß ist. Hier erreichen wir den Widerspruch.

Beh.: Schlechte Mengen sind platt

Als erstes zeigen wir, dass die schlechte Menge auf eine *platte Menge* K_1 eingeschränkt werden kann, d.h. K_1 ist in allen Dimensionen bis auf einer sehr klein. Natürlich achten bei unseren Einschränkungen darauf, dass das Verhältnis (7) durch λ beschränkt bleibt.

Behauptung

Angenommen, $f : K \to \mathbb{R}$ erfüllt die Ungleichung (6) und $\mathbb{E}_\mu = 0$. Dann gibt es für jedes $\varepsilon > 0$ eine konvexe Teilmenge $K_1 \subseteq K$, welche

$$\int_{K_1} h\,d\mu < \lambda \cdot \int_{K_1} f^2 d\mu \quad \text{und} \quad \int_{K_1} f\,d\mu = 0$$

erfüllt sowie K_1 liegt in der Box $[0, D] \times [0, \varepsilon]^{n-1}$ in einem geeigneten Koordinatensystem.

Beweisskizze zur ersten Behauptung

Wir nehmen an, dass für ein $j \geq 2$ der Körper K_j bereits eine schlechte Menge ist, die in $[0, D]^j \times [0, \varepsilon]^{n-j}$ liegt und für die $\int_{K_j} f \, d\mu = 0$ gilt, d.h. unsere schlechte Menge ist schon auf $n - j$ Dimensionen eingeschränkt.

Der Induktionsanfang $K_n = K$ ist selbstverständlich durch unsere Annahme (6) gegeben.

Für die Einschränkung entlang einer weiteren Dimension wird ein Divide-and-Conquer-Verfahren verwendet, welches wir hier nur grob ansehen werden.

Beweisskizze: Divide-and-Conquer-Verfahren

- Betrachte eine Projektion R auf die ersten beiden *pastösen* Dimensionen.
- Wähle $x \in \mathbb{R}$ so, dass jede Gerade durch x die Menge R in Gebiete mit jew. mind. einem Drittel der Gesamtfläche teilt. (So ein x existiert immer, ohne Beweis.)
- Wähle eine Hyperebene H durch x so, dass ihre Normale in R liegt und dass der Erwartungswert von f sowohl in $K_j \cap H^+$ als auch in $K_j \cap H^-$ gleich 0 ist. Beide Mengen sind dann ebenfalls schlecht.
- Setze dieses Verfahren auf eine diese Mengen fort, bis eine schlechte Menge mit Größe $\frac{1}{2}\varepsilon^2$ erreicht wird.
- Die Größe in einer Richtung ist dann maximal ε. (geometrischer Fakt, ohne Beweis.)
- Führe einen Koordinatenwechsel durch, sodass eine der beiden Projektionsdimensionen maximal Größe ε hat.

Per Induktion erreichen wir so eine *platte* Menge K_1.

Beweisskizze: Divide-and-Conquer-Verfahren

- Betrachte eine Projektion R auf die ersten beiden *pastösen* Dimensionen.
- Wähle $x \in \mathbb{R}$ so, dass jede Gerade durch x die Menge R in Gebiete mit jew. mind. einem Drittel der Gesamtfläche teilt. (So ein x existiert immer, ohne Beweis.)
- Wähle eine Hyperebene H durch x so, dass ihre Normale in R liegt und dass der Erwartungswert von f sowohl in $K_j \cap H^+$ als auch in $K_j \cap H^-$ gleich 0 ist. Beide Mengen sind dann ebenfalls schlecht.
- Setze dieses Verfahren auf eine diese Mengen fort, bis eine schlechte Menge mit Größe $\frac{1}{2}\varepsilon^2$ erreicht wird.
- Die Größe in einer Richtung ist dann maximal ε. (geometrischer Fakt, ohne Beweis.)
- Führe einen Koordinatenwechsel durch, sodass eine der beiden Projektionsdimensionen maximal Größe ε hat.

Per Induktion erreichen wir so eine *platte* Menge K_1.

Beweisskizze: Divide-and-Conquer-Verfahren

- Betrachte eine Projektion R auf die ersten beiden *pastösen* Dimensionen.
- Wähle $x \in \mathbb{R}$ so, dass jede Gerade durch x die Menge R in Gebiete mit jew. mind. einem Drittel der Gesamtfläche teilt. (So ein x existiert immer, ohne Beweis.)
- Wähle eine Hyperebene H durch x so, dass ihre Normale in R liegt und dass der Erwartungswert von f sowohl in $K_j \cap H^+$ als auch in $K_j \cap H^-$ gleich 0 ist. Beide Mengen sind dann ebenfalls schlecht.
- Setze dieses Verfahren auf eine diese Mengen fort, bis eine schlechte Menge mit Größe $\frac{1}{2}\varepsilon^2$ erreicht wird.
- Die Größe in einer Richtung ist dann maximal ε. (geometrischer Fakt, ohne Beweis.)
- Führe einen Koordinatenwechsel durch, sodass eine der beiden Projektionsdimensionen maximal Größe ε hat.

Per Induktion erreichen wir so eine *platte* Menge K_1.

Beweisskizze: Divide-and-Conquer-Verfahren

- Betrachte eine Projektion R auf die ersten beiden *pastösen* Dimensionen.
- Wähle $x \in \mathbb{R}$ so, dass jede Gerade durch x die Menge R in Gebiete mit jew. mind. einem Drittel der Gesamtfläche teilt. (So ein x existiert immer, ohne Beweis.)
- Wähle eine Hyperebene H durch x so, dass ihre Normale in R liegt und dass der Erwartungswert von f sowohl in $K_j \cap H^+$ als auch in $K_j \cap H^-$ gleich 0 ist. Beide Mengen sind dann ebenfalls schlecht.
- Setze dieses Verfahren auf eine diese Mengen fort, bis eine schlechte Menge mit Größe $\frac{1}{2}\varepsilon^2$ erreicht wird.
- Die Größe in einer Richtung ist dann maximal ε. (geometrischer Fakt, ohne Beweis.)
- Führe einen Koordinatenwechsel durch, sodass eine der beiden Projektionsdimensionen maximal Größe ε hat.

Per Induktion erreichen wir so eine *platte* Menge K_1.

Beweisskizze: Divide-and-Conquer-Verfahren

- Betrachte eine Projektion R auf die ersten beiden *pastösen* Dimensionen.
- Wähle $x \in \mathbb{R}$ so, dass jede Gerade durch x die Menge R in Gebiete mit jew. mind. einem Drittel der Gesamtfläche teilt. (So ein x existiert immer, ohne Beweis.)
- Wähle eine Hyperebene H durch x so, dass ihre Normale in R liegt und dass der Erwartungswert von f sowohl in $K_j \cap H^+$ als auch in $K_j \cap H^-$ gleich 0 ist. Beide Mengen sind dann ebenfalls schlecht.
- Setze dieses Verfahren auf eine diese Mengen fort, bis eine schlechte Menge mit Größe $\frac{1}{2}\varepsilon^2$ erreicht wird.
- Die Größe in einer Richtung ist dann maximal ε. (geometrischer Fakt, ohne Beweis.)
- Führe einen Koordinatenwechsel durch, sodass eine der beiden Projektionsdimensionen maximal Größe ε hat.

Per Induktion erreichen wir so eine *platte* Menge K_1.

Beweisskizze: Divide-and-Conquer-Verfahren

- Betrachte eine Projektion R auf die ersten beiden *pastösen* Dimensionen.
- Wähle $x \in \mathbb{R}$ so, dass jede Gerade durch x die Menge R in Gebiete mit jew. mind. einem Drittel der Gesamtfläche teilt. (So ein x existiert immer, ohne Beweis.)
- Wähle eine Hyperebene H durch x so, dass ihre Normale in R liegt und dass der Erwartungswert von f sowohl in $K_j \cap H^+$ als auch in $K_j \cap H^-$ gleich 0 ist. Beide Mengen sind dann ebenfalls schlecht.
- Setze dieses Verfahren auf eine diese Mengen fort, bis eine schlechte Menge mit Größe $\frac{1}{2}\varepsilon^2$ erreicht wird.
- Die Größe in einer Richtung ist dann maximal ε. (geometrischer Fakt, ohne Beweis.)
- Führe einen Koordinatenwechsel durch, sodass eine der beiden Projektionsdimensionen maximal Größe ε hat.

Per Induktion erreichen wir so eine *platte* Menge K_1.

Beh.: Schlechte Mengen sind noch platter!

Das oben besprochene Beweisvefahren kann leider die letzte Dimension nicht in eine ε-Box packen. Wir benötigen also einen neuen Ansatz für die letzte *pastöse* Dimension.

Behauptung

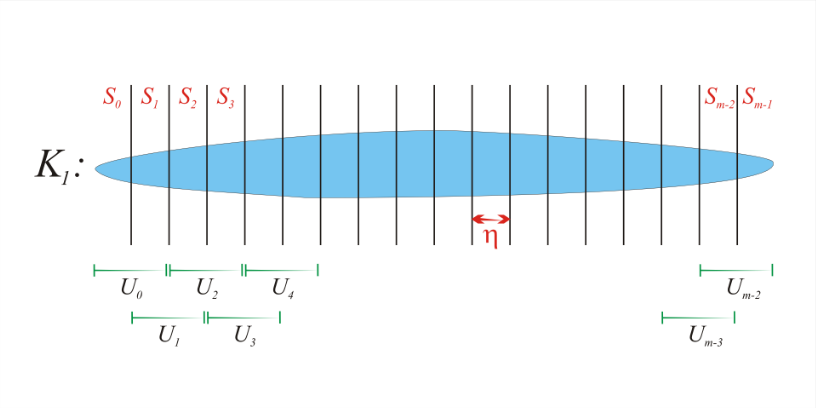
Seien K_1, f und δ wie oben und sei $\eta = \frac{c_3 \delta}{\sqrt{n}}$ mit einer Konstanten $c_3 > 0$. Unter der Krümmungsbedinung (2) existiert eine konvexe Teilmenge $K_0 \subseteq K_1$ mit

$$\int_{K_0} h \, d\mu < \frac{1}{10} \int_{K_0} (f - \bar{f})^2 d\mu \tag{8}$$

wobei $\bar{f} = \frac{1}{\mu(K_0)} \int_{K_0} f \, d\mu$ analog zu oben definiert ist.
Ebenfalls gilt: K_0 liegt in der Box $[-\eta, \eta] \cdot [0, \varepsilon]^{n-1}$.

Beweisidee für die zweite Behauptung

Wir zerlegen K_1 in m kleine Abschnitte und zeigen, dass mindestens einer dieser Abschnitte (oder die Vereinigung zweier benachbarter) die Ungleichung (8) erfüllt.



Beweisidee für die zweite Behauptung

- $\int_{K_1} f^2 d\mu = \sum_{i=0}^{m-1} \int_{S_i} (f - \bar{f}_i)^2 d\mu + \sum_{i=0}^{m-1} \mu(S_i) \bar{f}_i^2$
- $A :=$ erster Summand, $B :=$ zweiter Summand.
- Wenn $A > B$ gilt, folgt die Abschätzung schnell und es gilt sogar $\lambda = O(n^2)$.
- Für die Umkehrungen kann man zeigen, dass für die Vereinigungen U_k ebenfalls gilt: Eine globale große Varianz bedingt eine große lokale Varianz auf irgendeinem der U_k.

Der Beweis ist in Kapitel 6 der *Lecture Notes* von Mark Jerrum zu finden (Seiten 87-90).

Abschluss des Beweises

Wir haben nun eine schlechte Menge K_0 gefunden, die in einem 2η langem Prisma eingeschlossen ist, dessen restliche Seiten einem $(n-1)$ -dim. Würfel mit Seitenlänge ε entsprechen. Nun führen wir den Widerspruch herbei:

Sei C die Mittelachse des Prismas und seien z_1 und z_2 die Schnittpunkte von C mit den Seiten des Prismas. Wir setzen $\delta' := \delta - \varepsilon\sqrt{n}$ und wählen ε so klein, dass

$$\mathrm{Vol}_n\, B(0,\delta') \geq 0,9 \cdot \mathrm{Vol}_n\, B(0,\delta).$$

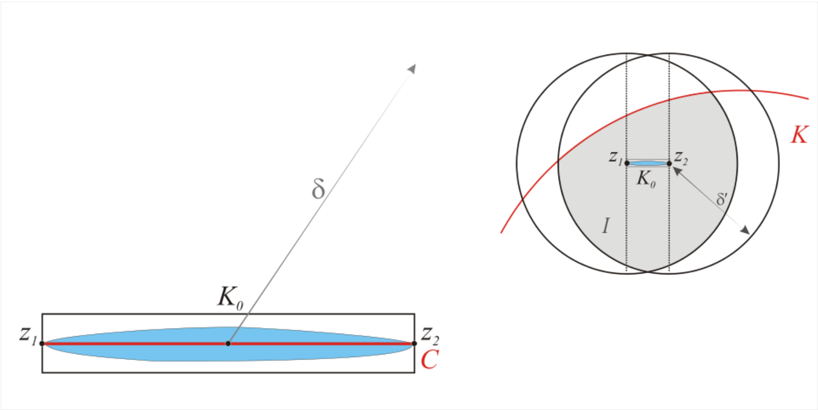
Sei $I := B(z_1,\delta') \cap B(z_2,\delta') \cap K$. Durch Änderung der Wahl von c_3 und damit Beeinflussung der Größe von η erreichen wir

$$\mathrm{Vol}_n\, B(z_1,\delta') \cap B(z_2,\delta') \geq 0,8 \cdot \mathrm{Vol}_n\, B(0,\delta),$$

wie in der Rechnung zum zweiten Lemma. Daraus folgt dann:

$$\mathrm{Vol}_n\, I \geq 0,2 \cdot \mathrm{Vol}_n\, B(0,\delta). \tag{9}$$

Graphik zum Beweisabschluss



Abschluss des Beweises II

Eine Schlüsseleigenschaft von I ist also, dass es nicht zu klein ist. Die andere ist die Tatsache, dass jeder Punkt in I von jedem Punkt aus K_0 in einem Schritt des ball walk erreicht werden kann. Nach Konstruktion gilt also

$$\sup\{\|x-y\| \,:\, x \in C \text{ und } y \in I\} \leq \delta',$$

woraus nach der Dreiecksungleichung folgt:

$$\sup\{\|x-y\| \,:\, x \in K_0 \text{ und } y \in I\} \leq \delta' + \varepsilon\sqrt{n} = \delta.$$

Da $I \subseteq K$, können wir hierfür auch schreiben:

$$I \subseteq B(x,\delta) \cap K \qquad \text{für alle } x \in K_0. \tag{10}$$

Abschluss des Beweises: Der Widerspruch

Nun erhalten wir

$$\begin{aligned}
\int_{K_0} h \, d\mu &\geq \frac{1}{2} \int_{K_0} \frac{\mu(dx)}{\mathrm{Vol}_n B(x,\delta) \cap K} \int_I (f(x) - f(y))^2 \, dy \quad \text{(5),(10)} \\
&\geq \frac{1}{2 \,\mathrm{Vol}_n B(0,\delta)} \int_{K_0} \mu(dx) \int_I (f(x) - f(y))^2 \, dy \\
&\geq \frac{1}{2 \,\mathrm{Vol}_n B(0,\delta)} \int_I dy \int_{K_0} (f(x) - f(y))^2 \, \mu(dx) \quad \text{(Fubini)} \\
&\geq \frac{1}{2 \,\mathrm{Vol}_n B(0,\delta)} \int_I dy \int_{K_0} (f - \bar{f})^2 \, d\mu \quad \text{(Varianzmin.)} \\
&\geq \frac{1}{10} \int_{K_0} (f - \bar{f})^2 d\mu \quad \text{(9)}
\end{aligned}$$

Dies ist aber ein Widerspruch zur zweiten Behauptung (8).

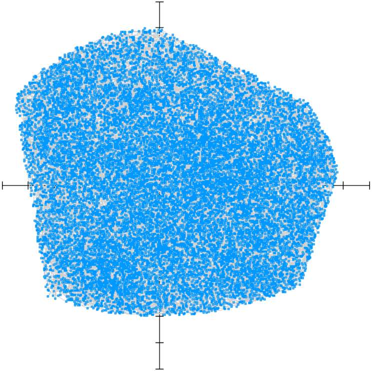
qed.

Das Verfahren in der praktischen Anwendung

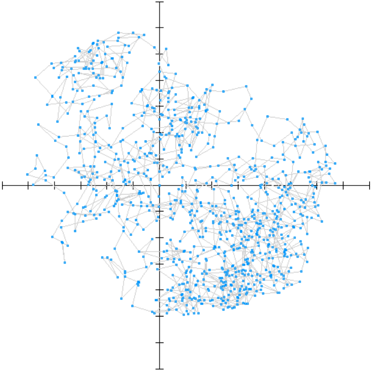
Zwar ist das Verfahren explizit für hochdimensionale Anwendungen ausgelegt, dennoch werden hier der Anschauung wegen nur zwei- und dreidimensionale konvexe Körper betrachtet.

Aus den am Anfang des Vortrags genannten Gründen kommt ausschließlich die Metropolis-Variante des Algorithmus für eine Implementierung in Frage. Ebenfalls vernachlässigen wir die Krümmungsbedingung, achten aber bei unseren Testkörpern darauf, dass keine spitzen Winkel auftreten.

Funktionsweise I

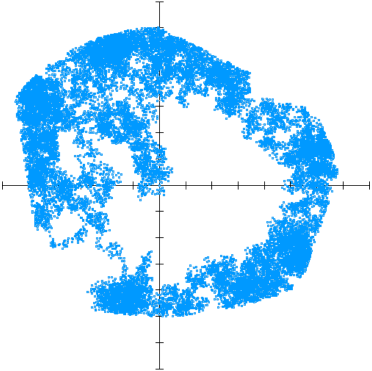


$\delta = 1;\ n = 20.000$
Gute Wahl...



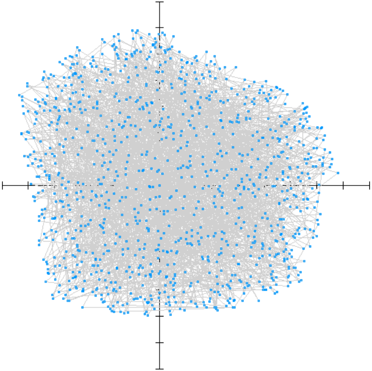
$\delta = 1;\ n = 1.000$
Zu wenig Schritte.

Funktionsweise II



$\delta = 0,005;\ n = 20.000$

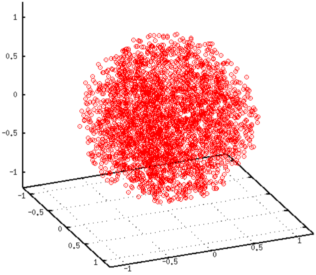
Zu kleine Schrittweite.



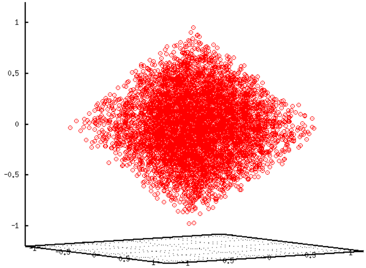
$\delta = 5.000;\ n = 20.000$

Zu große Schrittweite.

Konvexe dreidimensionale Körper



Einheitskugel



Doppelpyramide

Von den Punkten zum Flächeninhalt

Um aus den erzeugten Punkten einen Flächeninhalt zu generieren, nutzen wir den Ansatz des *Anteilsprodukts*:

Sei der konvexe Körper K gegeben. Wir betrachten eine Serie konzentrischer Kugeln $B_0 \subset B_1 \subset \cdots \subset B_k$, sodass $B_0 \subseteq K$ und $K \subseteq B_k$. Weiterhin fordern wir, dass das Volumen nicht zu stark ansteigt, etwa $\mathrm{Vol}_n B_{i+1} \leq 2 \cdot \mathrm{Vol}_n B_i$. Wir können die Anteile

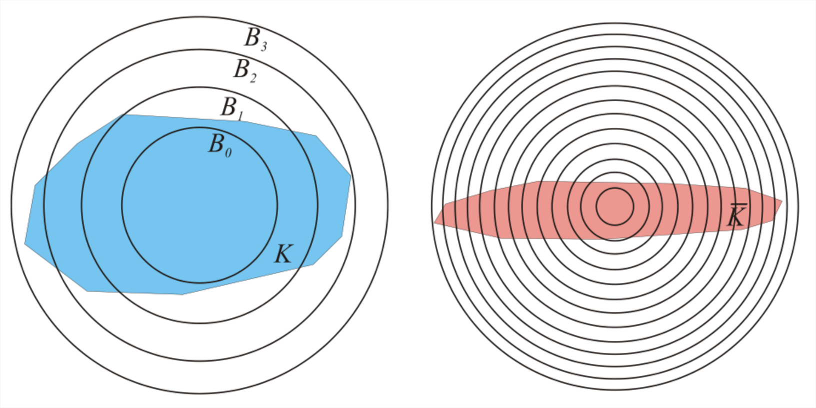
$$\rho_i = \frac{\mathrm{Vol}_n B_i \cap K}{\mathrm{Vol}_n B_{i+1} \cap K}$$

abschätzen durch wiederholtes Erzeugen von Punkten aus $B_{i+1} \cap K$ und Bestimmen des Anteils dieser Punkte, die auch in $B_i \cap K$ liegen.

Sei Z_i ein so erzeugter Schätzer für ρ_i. Dann erhalten wir eine Abschätzung des Volumens von K durch

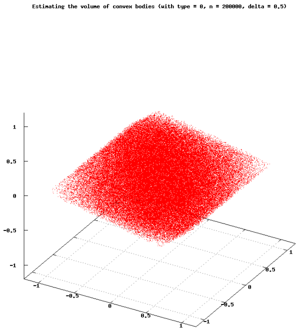
$$\mathrm{Vol}_n K \approx \mathrm{Vol}_n B_0 \cdot \prod_{i=0}^{k-1} \frac{1}{Z_i}.$$

Abbildung: Konzentrische Kugeln

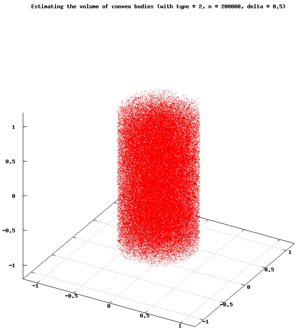


K annähernd rund $\rightarrow$ wenig Kugeln $\rightarrow$ gute Schätzung möglich.
$\overline{K}$ platt $\rightarrow$ viele Kugeln $\rightarrow$ große Ungenauigkeit!

Volumenberechnung zweier Testkörper

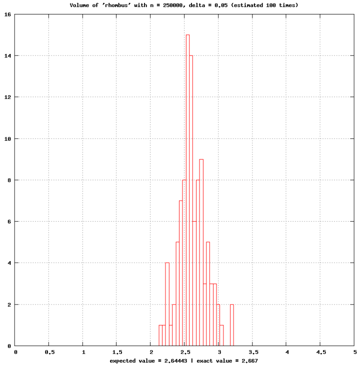


$\mathrm{Vol}_{exakt} = 8/3 \approx 2,667$
(Doppelpyramide)

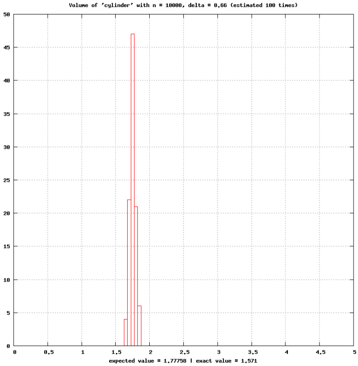


$\mathrm{Vol}_{exakt} = \pi/2 \approx 1,571$
(Zylinder)

Volumenberechnung: Simulation



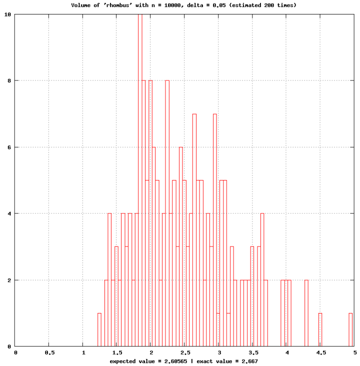
$\mathbb{E}[\text{Vol}] = 2,64443$
(Doppelpyramide)



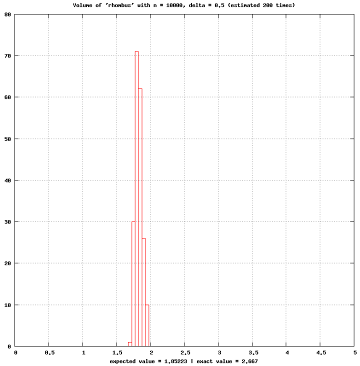
$\mathbb{E}[\text{Vol}] = 1,77758$
(Zylinder)

Wahl des richtigen δ ist essentiell!

Wir betrachten nun zwei Abschätzungen für die Doppelpyramide:



$\delta = 0.05;\ n = 10.000;\ m = 200$
$\mathbb{E}[\mathrm{Vol}] = 2,60565$



$\delta = 0.5;\ n = 10.000;\ m = 200$
$\mathbb{E}[\mathrm{Vol}] = 1,85223$

Probleme in der Praxis

- Implementierung der Orakelfunktion.
- Wie groß ist der Durchmesser D des konvexen Körpers?
- Wo lässt man den *ball walk* starten?
- Wahl des richtigen δ!
- Wegfall der Krümmungsbedingung (2)!
- Viele konzentrische Kugeln $\rightarrow$ große Varianz $\rightarrow$ große Fehler wahrscheinlich!
- Skalierung von K.

Fazit

Es müssen weitere Informationen über den konvexen Körper K gegeben sein und das Verfahren je nach Beschaffenheit angepasst werden!

Verwendung des Verfahrens

Die Berechnung hochdimensionaler Integrale über konvexe Mengen ist ein häufig auftretendes Problem in vielen mathematischen und außermathematischen Themenbereichen. Insbesondere erscheinen Fragestellungen dieser Art

- in der Finanzmathematik. [Risikooptimierung]
- in der Bayesschen Statistik.
- in der statistischen Physik. [Gibbs-Sampling]

Und nächste Woche sehen wir...

Unter gewissen Voraussetzungen an die Eingabedaten können direkte Fehlerschranken sowohl für den reinen Monte Carlo- als auch den Metropolis-Algorithmus mit zugrundeliegendem *ball walk* angegeben werden.

Diese Fehlerschranken hängen von der Dimension d, der Schrittweite δ des *ball walk* und einer Konstanten $\alpha < 1$ aus den Voraussetzungen an die Eingabedaten ab.

Markus Burkow wird hierzu konkrete Fehlerschranken (insbesondere also obere Grenzen) für beide Algorithmen angeben sowie als zentrale Aussage für die Metropolisvariante die Abschätzung

$$\lim_{n\to\infty} e\left(S_n^\delta, (f,\rho)\right)^2 \cdot n \quad \leq \quad \frac{8 \cdot 1600^2}{81\pi} (d+1) \cdot \frac{e^{2\alpha\delta}}{\delta^2}$$

herleiten und somit das polynomielle Wachstum der Komplexität in d.

Schlußwort

Die Folien sowie eine textgleiche Druckversion (mit allen Abbildungen) dieses Vortrags sind auch im Internet abrufbar unter:

www.grohganz.de/stochastic.html

Vielen Dank für Ihre Aufmerksamkeit.

Literaturverzeichnis

- JERRUM, MARK mit ALEX BELOW und CHRISTOPH AMBÜHL:
  Lecture Notes of "Counting, sampling and integrating: algorithms and complexity", Kapitel 6.
  homepages.inf.ed.ac.uk/mrj/ETHbook/chap6.ps
- FELLER, WILLIAM:
  An Introduction to Probability Theory and its Applications Vol I, 2nd Edition.
  Wiley, 1964, New York.