Vorlesung Monte-Carlo-Verfahren

Technische Universität München, SS 2010

Dozentin: Caroline Lasser
Tutor: Thomas März

(Stand vom 22. Juli 2010)

Termine:
Vorlesung dienstags 10:15-11:45 Uhr, Raum MI 02.10.011
Vorlesung donnerstags 10:15-11:45 Uhr, Raum MI 02.10.011
Übung dienstags 12:15-13:45 Uhr, Raum MI 02.10.011

Inhaltsverzeichnis

1 Einführung

1.1 Motivation [L] (20.4.)

Gesetz der großen Zahl und zentraler Grenzwertsatz erklären die prinzipielle Dimensionsunabhängigkeit von Monte-Carlo-Verfahren

1.2 Anwendungen [L, ES] (22.4.)

Statistische Physik: zweidimensionales Ising-Modell, Integration bezüglich der Boltzmann-Verteilung auf hoch-dimensionalem Zustandsraum; Moleküle: Modellierung im $\mathbb{R}^{3k}$ für k Atome mit (singulärem) Lennard-Jones-Potential; Herausforderung der hohen Dimension: Volumen der Einheitskugel im $\mathbb{R}^n$ geht gegen Null für $n \to \infty$

1.3 Stochastik-Grundlagen [G] (27.4.)

Satz von Vitali, σ-Algebra, Wahrscheinlichkeitsmaß, bedingte Wahrscheinlichkeit, unabhängige Ereignisse, Zufallsvariable, Verteilung einer Zufallsvariablen, kumulative Verteilungsfunktion, Erwartungswert, Chebyshev-Ungleichung, Varianz, unabhängige Zufallsvariablen

1.4 Stochastik- und Statistikgrundlagen [G] (4.5)

Kovarianz, stochastische Konvergenz, schwaches Gesetz der großen Zahlen, fast sichere Konvergenz, starkes Gesetz der großen Zahlen, Konvergenz in Verteilung, zentraler Grenzwertsatz, statistische Modelle, Schätzer, Maximum Likelihood Prinzip, Erwartungstreue

2 Zufallsvariablen

2.1 Gleichverteilung & Inversion [ES, MN] (6.5.)

Periode einer Folge von Pseudozufallszahlen, lineare Kongruenzmethoden, Satz von Hull und Dobell (1962) zur Charakterisierung der maximalen Periode (ohne Beweis), mehrfach rekursive Matrixmethoden, Mersenne Twister als Generator mit Periode $2^{19937} - 1$, Definition der inversen Verteilungsfunktion, Inversionsmethode für Exponential-, Bernoulli- und Gleichverteilung auf $\{0, \dots, M-1\}$, Formulierung der Verwerfungsmethode

2.2 Verwerfungsmethode [D] (11.5.)

Formulierung der Verwerfungsmethode (v. Neumann 1951), Beweis, dass die Verwerfungsmethode fast sicher stoppt und eine Zufallsvariable mit gewünschter Dichte f erzeugt, die Anzahl der Iterationen ist geometrisch verteilt, Normierung der Dichte f, Formulierung der Squeeze-Methode (Marsaglia 1977), Beweis der Gleichung von Wald (1944), die erwartete Anzahl der f-Aufrufe in der Squeeze-Methode ist gleich der Fläche des einhüllenden Schlauchs.

2.3 Adaptive Verwerfung [GW, RC] (20.5., 10:15)

Definition logarithmisch konkav, Normalverteilung und Gamma-Verteilung als Beispiele für logarithmisch konkave Dichten, Konstruktion einer stückweise linearen oberen und unteren Einhüllenden für konkave und logarithmisch konkave Dichten, Diskussion der Kompositionsmethode, Formulierung der adaptiven Verwerfungsmethode.

3 Monte-Carlo-Integration

3.1 Numerische Integration [KU, Hi] (20.5., 12:15)

Beweis, dass die adaptive Verwerfungsmethode die gewünschte Verteilung liefert, Konditionszahlen für Integration in Abhängigkeit vom Integranden und den Integrationsgrenzen, Konditionszahlen für Quadraturformeln und Summation, Vorwärtsfehler der rekursiven und paarweisen Summation.

3.2 Einfache Monte-Carlo-Integration [KU, Ha] (27.5.)

Quadraturformeln über empirische Mittelwerte von n unabhängigen, identisch verteilten Zufallsvariablen, das starke Gesetz der großen Zahlen garantiert die fast sicher Konvergenz für $n \to \infty$, der mittlere quadratische Fehler ist der Quotient aus der Standardabweichung des Integranden und $\sqrt{n}$, das arithmetische Mittel und die Sampel-Varianz aus N unabhängigen Quadraturläufen sind erwartungstreue Schätzer für den Integralwert und den mittleren quadratischen Fehler, Varianzreduktion durch Approximation und Importance Sampling, Fehler-Abschätzungen für geschichtetes Sampling (Haber 1966)

3.3 Gleichverteilte Folgen und Diskrepanz [N] (1.6.)

Definition gleichverteilter Folgen über Intervalle, Zusammenhang mit gleichgewichteter Quadratur Riemann-integrierbarer Funktionen, Definition von Diskrepanz und Sterndiskrepanz, eine Folge ist genau dann gleichverteilt, wenn die Diskrepanz ihrer Anfangsstücke gegen Null konvergiert, Formel für eindimensionale Sterndiskrepanz, summierte Mittelpunktsregel basiert auf Folge mit minimaler Diskrepanz

3.4 Quasi-Monte-Carlo-Integration [N] (8.6.)

Definition van der Corput Folgen, van der Corput Folgen haben eine Diskrepanz von $O(n^{-1}\log(n))$, Definition der Variation für $f : [a, b] \to \mathbb{R}$, partielle Integration für das Lebesgue-Stieltjes Integral, Beweis der Koksma'schen Ungleichung, Definition von Halton-Folgen , Halton Folgen zu Basen $b_1, \dots, b_d$, die relativ prim sind, haben eine Diskrepanz von $O(n^{-1}\log(n)^d)$

3.5 Adaptives Importance Sampling I [OB, ES] (10.6.)

Definition der mehrdimensionalen Variation im Sinne von Hardy und Krause, Koksma-Hwlaka Ungleichung, Beweis, dass die Koksma-Hwlaka Ungleichung scharf ist, $w(x) = |f(x)|/\|f\|_1$ liefert die minimale Varianz für Importance Sampling, Ansatz der matching characteristics, adaptiver Algorithmus von Oh, Berger (1992), Formulierung der fast sicheren Konvergenz für den Vektor der Charakteristiken

3.6 Adaptives Importance Sampling II [ES, D] (15.6.)

Definition der bedingten Erwartung, elementare Eigenschaften der bedingten Erwartung, Definition eines Martingals, Summe einer Folge von fairen Münzwürfen als Beispiel eines Martingals, Martingalkonvergenzsatz, Beweis, dass der Vektor der Charakteristiken beim adaptiven Importance Sampling fast sicher konvergiert

4 Markov-Ketten

4.1 Grundlegendes [ES, G] (17.6.)

Beweisende für das adaptive Importance Sampling (Anwendung des Martingal-Konvergenzsatzes und des Kronecker-Lemmas), Definition einer Markov-Kette, effektiver Zustandsraum, Beschränkung auf zeitlich homogene Markov-Ketten, Definition des Übergangskerns

4.2 Stationarität & Ergodensatz [ES, R] (22.6.)

Darstellung der gemeinsamen Verteilung von $X_0, \dots, X_n$ mit Mehrfachintegralen über den Übergangskern, Definition einer stationären Verteilung, eine homogene Markov-Kette mit stationärer Anfangsverteilung hat für jedes X_n die selbe Verteilung, Definition von Stoppzeiten und ϕ-Irreduzibilität, Erinnerung an die Lebesgue-Zerlegung, Definition von aperiodisch, Formulierung des starken Ergodensatzes, Definition von Rekurrenz und Harris-Rekurrenz

4.3 Rekurrenz & Reversibilität [ES, MT] (24.6.)

rekurrente Markov-Ketten haben ein σ-endliches stationäres Maß, Harris-rekurrente ein eindeutig bestimmtes stationäres Wahrscheinlichkeitsmaß, Formulierung eines starken Gesetzes der großen Zahl für aperiodische, Harris-rekurrente Markov-Ketten und $f \in \mathcal{L}^1(\pi)$, Definition von ergodisch und V- gleichmäßig ergodisch, Formulierung eines zentralen Grenzwertsatzes für Harris-rekurrente, V-gleichmäßig ergodische Markov-Ketten und $f^2 \leq V$, Definition von reversibel, ist eine Markov-Kette bezüglich π reversibel, so ist π stationär, Formulierung eines zentralen Grenzwertsatzes für Harris-rekurrente, geometrisch ergodische, reversible Markov-Ketten und $f \in \mathcal{L}^2(\pi)$

5 Metropolis-Hastings Algorithmus

5.1 Definition & Eigenschaften [ES] (29.6.)

Formulierung des Metropolis-Hastings Algorithmus, Berechnung des Übergangskerns, Nachweis der Reversibilität, Diskussion symmetrischer Ketten

5.2 Konvergenz [RC] (1.7.)

Bedingungen, die Aperiodizität garantieren, $q(x, y) > 0$ für $x, y \in \mathrm{supp}(f)$ garantiert f-Irreduzibilität, Definition harmonischer Funktionen, Beweis, dass aperiodische, irreduzible Markov-Ketten mit stationärer Verteilung, welche nur die Konstanten als beschränkte harmonische Funktionen haben, Harris-rekurrent sind, Beweis, dass eine aperiodische und f-irreduzible Metropolis-Kette Harris-rekurrent ist, Diskussion von Independence-Ketten

5.3 Independence-Ketten & Irrfahrten [RC, MenT] (6.7)

äquivalente Charakterisierungen von gleichmäßiger Ergodizität, Beweis, dass die Independence-Ketten gleichmäßig ergodisch sind, falls es ein $C \geq 1$ mit $f(x) \leq Cg(x)$ für alle x im Träger von f gibt, Beispiel einer normalverteilten Vorschlagsdichte mit „falschem“ Erwartungswert, Beweis des ε-δ-Kriteriums für f-Irreduzibilität und Aperiodizität, Anwendung auf Irrfahrten-Metropolis-Ketten

6 Gibbs-Sampler

6.1 Definition & Konvergenz [ES, RP] (8.7.)

Definition von bedingten Wahrscheinlichkeitsdichten, Formulierung des Gibbs-Samplers, der Gibbs-Sampler erzeugt eine homogene Markov-Kette mit stationärer Dichte f, ist die Dichte des Übergangskerns positiv, so ist der Gibbs-Sampler f-irreduzibel, aperiodisch und Harris-rekurrent

6.2 Gibbs Sampler & hybride Ketten [RP, RC, T] (13.7.)

für positive und gleichgradig stetige Übergangsdichten ist der Gibbs-Sampler gleichmäßig ergodisch, Formulierung eines zweistufigen, reversiblen Gibbs-Samplers, Beweis der Reversibilität, Formulierung eines mehrstufigen reversiblen Gibbs-Samplers, Definition gemischter und zyklischer Ketten, Beweis, dass die gemischten und zyklischen Kette die stationäre Verteilung der Einzelketten erben, Beweis, dass sich die gleichmäßige Ergodizität einer Einzelkette auf die gemischte Kette überträgt

6.3 Gibbs-Sampler & Metropolis-Algorithmus [ES] (15.7.)

der Gibbs-Sampler ist eine zyklische Markov-Kette bestehend aus Metropolis-Ketten mit Akzeptanzwahrscheinlichkeit Eins, inverses Importance Sampling zur Berechnung von Normalisierungskonstanten

6.4 Verborgene und Kontrollvariablen [ES] (22.7.)

Methode der verborgenen Variablen an einem Beispiel, das mit multivariater Normalverteilung und Exponentialverteilung direkt gesampelt werden kann, Diskussion additiver Kontrollvariablen und Kriterium für Effektivität, Zusammenfassung des Inhalts der Vorlesung.

Literatur

[D] L. Devroye, Non-Uniform Random Variate Generation, Springer, 1986.

[Du] R. Durrett, Probability: theory and examples, 3rd edition, Brooks/Cole, 2005.

[ES] M. Evans, T. Schwartz, Approximating Integrals via Monte Carlo and Deterministic Methods, Oxford University Press, 2000.

[G] H.-O. Georgii, Stochastics, Introduction to Probability and Statistics, de Gruyter, 2008.

[GW] W. Gilks, P. Wild, Adaptive Rejection Sampling for Gibbs Sampling, Appl. Statist. Vol. 41, pp. 337–348, 1992.

[Ha] S. Haber, A modified Monte-Carlo quadrature, Math. Comp. Vol. 20, pp. 361–368, 1966.

[Hi] N. Higham, The accuracy of floating point summation, SIAM J. Sci. Comput. Vol. 14, pp. 783–799, 1993.

[KU] A. Krommer, C. Ueberhuber, Computational Integration, SIAM, 1998.

[L] J. Liu, Monte Carlo Strategies in Scientific Computing, Springer, 2001.

[MN] M. Matsumoto, T. Nishimura, Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator, ACM Trans. on Modeling and Computer Simulation Vol. 8, pp. 3–30, 1998.

[MenT] K. Mengersen, R. Tweedie, Rates of convergence of the Hasings and Metropolis Algorithms, Ann. Stat. Vol. 24, pp. 101–121, 1996.

[MT] S. Meyn, R. Tweedie, Markov Chains and Stochastic Stability, 2nd edition, Cambridge University Press, 2009.

[N] H. Niederreiter, Random Number Generation and Quasi-Monte Carlo Methods, SIAM, 1992.

[OB] M. Oh, J. Berger, Adaptive importance sampling in Monte Carlo integration, J. Statist. Comput. Simul. Vol. 41, pp. 143–168, 1992.

[RC] C. Robert, G. Casella, Monte Carlo Statistical Methods, 2nd edition, Springer, 2004.

[RP] G. Roberts, N. Polson, On the geometric convergence of the Gibbs Sampler, J. Roy. Statist. Soc. B 56, pp. 377–384, 1994.

[R] J. Rosenthal, A review of asymptotic convergence for general state space Markov chains, Far East J. Theor. Stat. Vol. 5, pp. 37–50, 2001.

[T] L. Tierney, Markov Chains for Exploring Posterior Distributions, Ann. Statist. 22, pp. 1701–1728, 1994.