	Lectures	Exercises/Home work problems	
Ordinary differential equations	Linear dynamics	E1	
	Non-linear dynamics	E2	
	Molecular dynamics		H1a/H1b
Stochastic methods	Monte Carlo integration	E3	
	Metropolis algorithm		H2a/H2b
	Brownian dynamics	E4	
Partial differential equations	Quantum structure	E5	H3a
	Quantum dynamics		H3b

The Monte Carlo method

Based on extensive use of random numbers

1. **Simulation of problems that are modelled as stochastic in nature.**

   Ex. Brownian dynamics (E4)

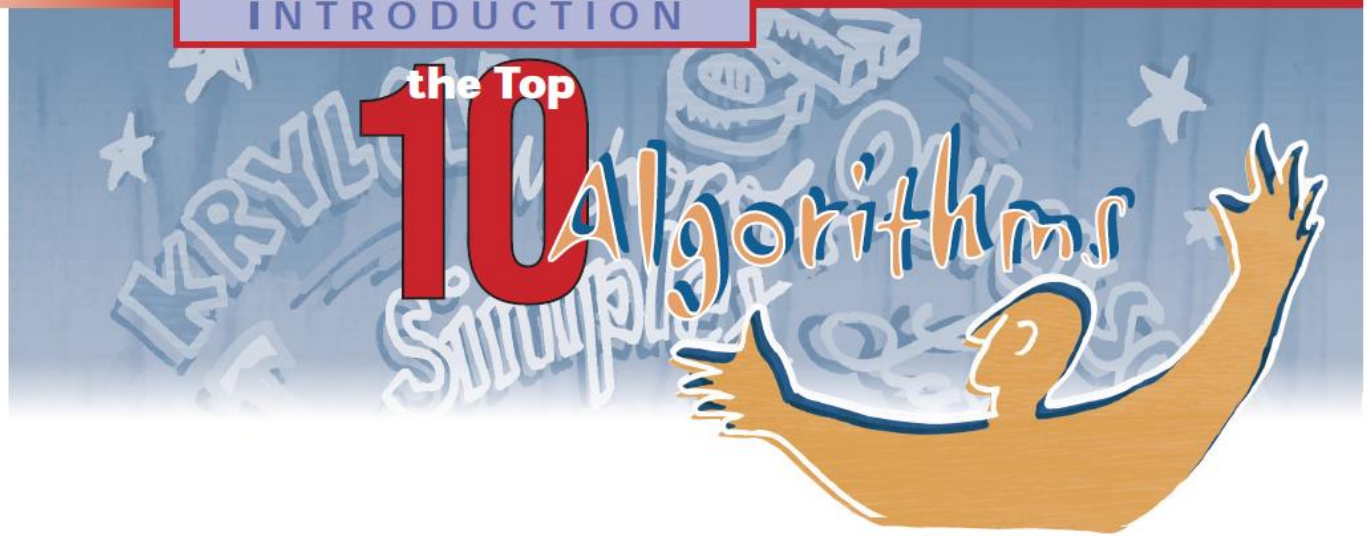
2. **Simulation of problems that are ”deterministic”, but reformulated such that a stochastic approach can be used.**

   Ex. Monte Carlo integration (E3)

   - equilibrium properties in statistical mechanics (H2a)
   - quantum stucture in high dimensions (H3a)

GUEST EDITORS' INTRODUCTION

the Top 10 Algorithms



In putting together this issue of *Computing in Science & Engineering*, we knew three things: it would be difficult to list just 10 algorithms; it would be fun to assemble the authors and read their papers; and, whatever we came up with in the end, it would be controversial. We tried to assemble the 10 algorithms with the greatest influence on the development and practice of science and engineering in the 20th century. Following is our list (here, the list is in chronological order; however, the articles appear in no particular order):

- Metropolis Algorithm for Monte Carlo
- Simplex Method for Linear Programming
- Krylov Subspace Iteration Methods
- The Decompositional Approach to Matrix Computations
- The Fortran Optimizing Compiler
- QR Algorithm for Computing Eigenvalues
- Quicksort Algorithm for Sorting
- Fast Fourier Transform
- Integer Relation Detection
- Fast Multipole Method

Göran Wahnström, Department of Physics, Chalmers

THE JOURNAL OF CHEMICAL PHYSICS VOLUME 21, NUMBER 6 JUNE, 1953

Equation of State Calculations by Fast Computing Machines

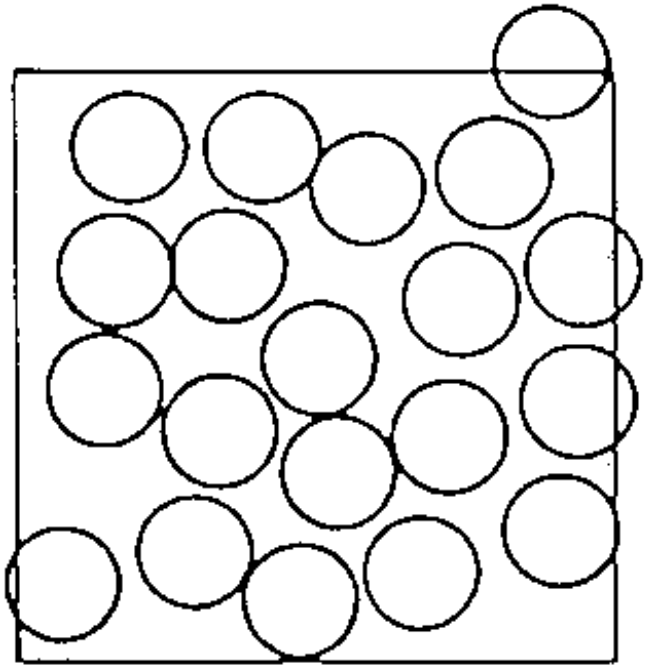
NICHOLAS METROPOLIS, ARIANNA W. ROSENBLUTH, MARSHALL N. ROSENBLUTH, AND AUGUSTA H. TELLER,
Los Alamos Scientific Laboratory, Los Alamos, New Mexico

AND

EDWARD TELLER,* *Department of Physics, University of Chicago, Chicago, Illinois*

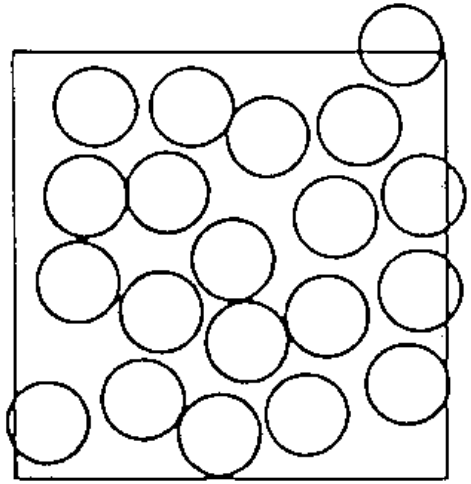
(Received March 6, 1953)

A general method, suitable for fast computing machines, for investigating such properties as equations of state for substances consisting of interacting individual molecules is described. The method consists of a modified Monte Carlo integration over configuration space. Results for the two-dimensional rigid-sphere system have been obtained on the Los Alamos MANIAC and are presented here. These results are compared to the free volume equation of state and to a four-term virial coefficient expansion.



Equation of state – system of many interacting particles

High-dimensional integrals



Probability distribution function

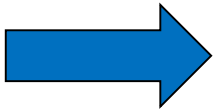
$$P(\mathbf{r}^N) = \frac{1}{Z} \exp\left[-\beta V(\mathbf{r}^N)\right]$$

Mean value, e.g. the potential energy

$$\langle V \rangle = \int d\mathbf{r}^N V(\mathbf{r}^N) P(\mathbf{r}^N)$$

Metropolis et al. introduced: Importance sampling

"instead of choosing configurations randomly, then weighting them with $\exp\left[-\beta V(\mathbf{r}^N)\right]$, we choose configurations with the probability proportional to $\exp\left[-\beta V(\mathbf{r}^N)\right]$ and weight them evenly"



Markov Chain Monte Carlo (MCMC)

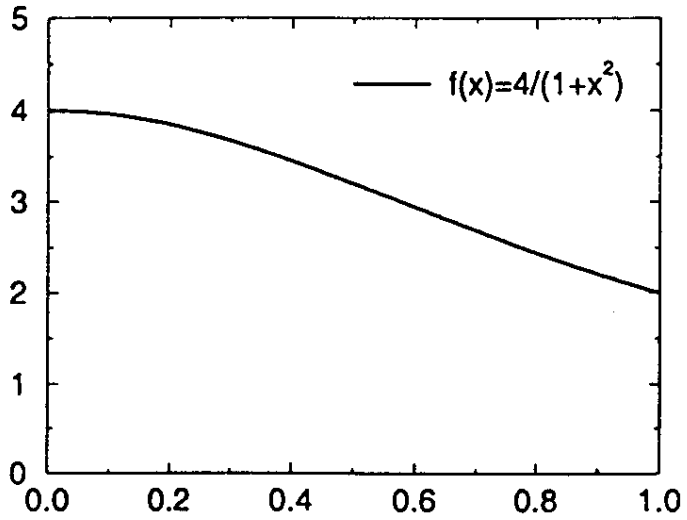
Monte Carlo

Content:

- **Monte Carlo integration**
- **Variance reduction / Importance sampling**
- **Markov chains**
- **The Metropolis algorithm**
- **Error estimate**

Monte Carlo integration

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



1. Choose N points x_i at random with uniform probability within the integration interval [0,1]

2. Determine the mean value

$$I_N = \langle f \rangle \equiv \frac{1}{N}\sum_{i=1}^{N} f(x_i) \equiv \frac{1}{N}\sum_{i=1}^{N} f_i$$

and the variance

$$\sigma_f^2 = \left\langle (f - \langle f \rangle)^2 \right\rangle = \left\langle f^2 \right\rangle - \langle f \rangle^2$$

3. Approximate the value of the integral as

$$I = I_N \pm \frac{\sigma_f}{\sqrt{N}}$$

Central limit theorem

The sum

$$I_N = \frac{1}{N}\sum_{i=1}^{N} f_i$$

is approximately Gaussian distributed

$$P(I_N) = \frac{1}{\sqrt{2\pi}\sigma_I}\exp\left[-\frac{(I_N-\mu)^2}{2\sigma_I}\right]$$

with mean value

$$\mathrm{E}[I_N] = \mu = \langle f \rangle$$

and variance

$$\mathrm{Var}[I_N] = \sigma_I^2 = \sigma_f^2/N$$

The requirements are:

1. The variables $(f_1, \ldots, f_i, \ldots, f_N)$ have to be statistically independent.

2. The mean value $\mu = \langle f \rangle$ and the variance $\sigma_f^2 = \left\langle (f - \langle f \rangle)^2 \right\rangle$ have to exist.

3. N has to be sufficiently large.

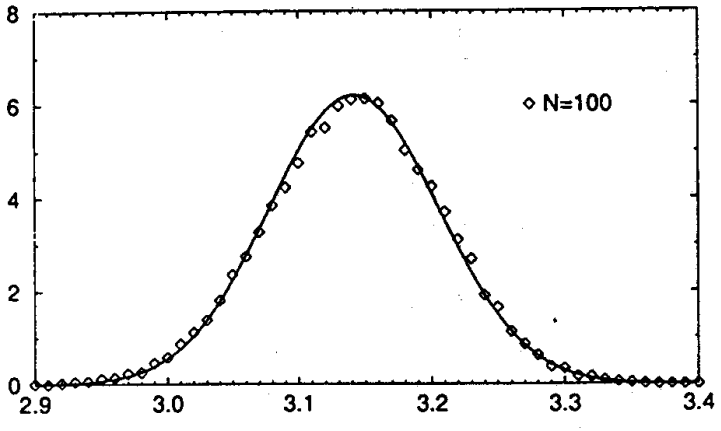
Notice: This is independent on the actual distribution for f.

Central limit theorem - example

Distribution of 20 000 Monte Carlo evaluations of

$$I = \int_0^1 \frac{4}{1+x^2}\, dx$$

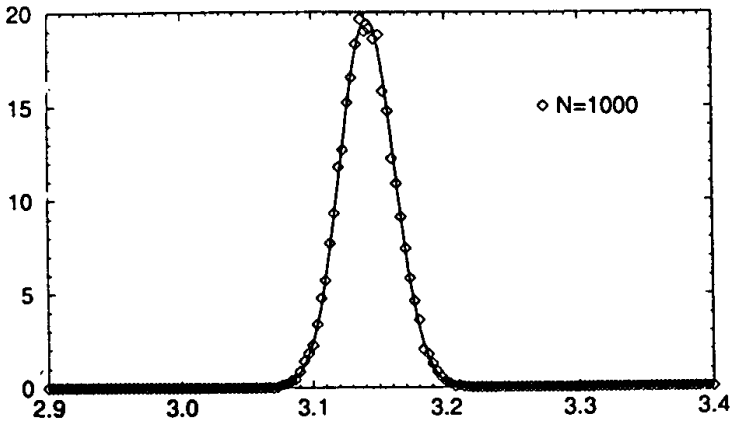
using $N = 100$ and $N = 1000$ points, respectively, for each integral evaluation.



Analytical result:

$$\mu = \pi$$

$$\sigma_I = \sqrt{(4 + 2\pi - \pi^2)}/\sqrt{N}$$



Monte Carlo integration - example

Numerical example

$$I = \int_0^1 \frac{4}{1+x^2}\, dx$$

The probability to find the correct value within

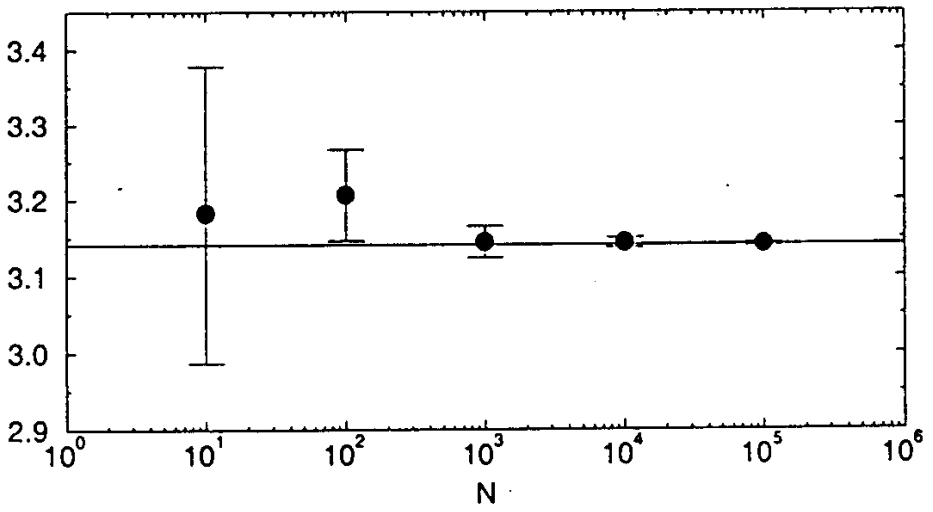
$$\pm\sigma_I \quad \text{is} \quad 68\%$$

and within

$$\pm 2\sigma_I \quad \text{it is} \quad 95\%$$

Notice: The error is independent on dimension, in comparison with more conventional numerical integration schemes.

	N	I_N	σ_f	$\sigma_I = \sigma_f/\sqrt{N}$
MC	10^1	3.18266	0.61954	0.19592
MC	10^2	3.20677	0.60315	0.06032
MC	10^3	3.14463	0.65030	0.02056
MC	10^4	3.14380	0.63989	0.00640
MC	10^5	3.14096	0.64438	0.00204
exact		3.14159	0.64310	



Central limit theorem - example

Distribution of 20 000 Monte Carlo evaluations of

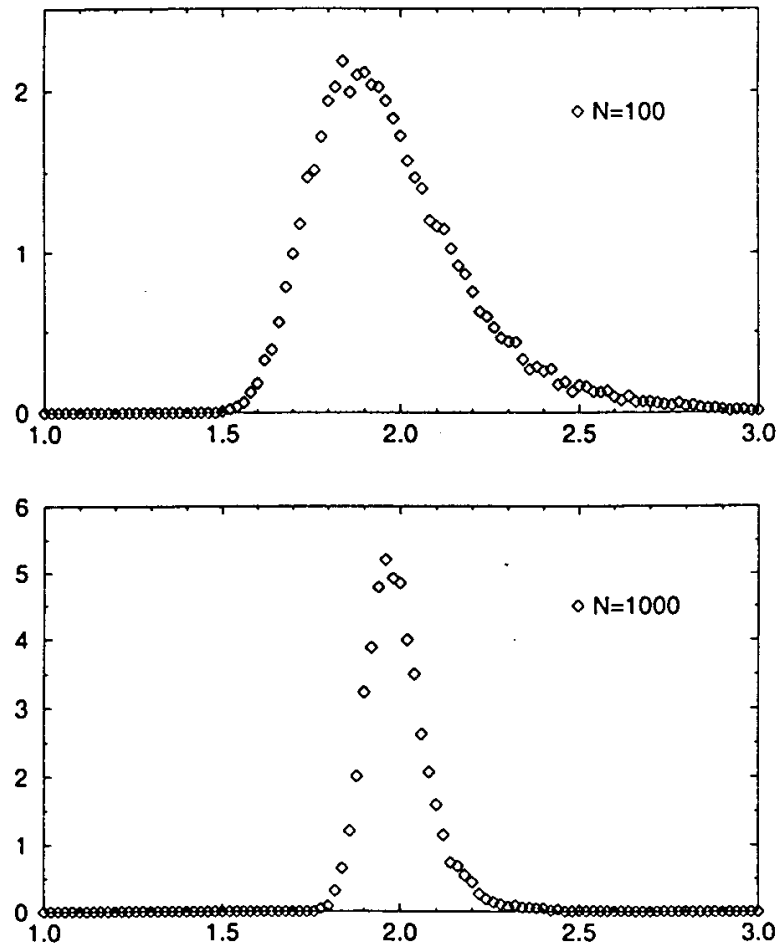
$$I = \int_0^1 \frac{1}{\sqrt{x}}\, dx$$

using $N = 100$ and $N = 1000$ points, respectively, for each integral evaluation.

Analytical result:

$\mu = 2$

σ_I undefined



Monte Carlo

Content:

- **Monte Carlo integration**
- **Variance reduction / Importance sampling**
- **Markov chains**
- **The Metropolis algorithm**
- **Error estimate**

Variance reduction / Importance sampling

$$I = I_N \pm \frac{\sigma_f}{\sqrt{N}}$$

The error can be reduced by either

- increasing the number of sampling points *N*

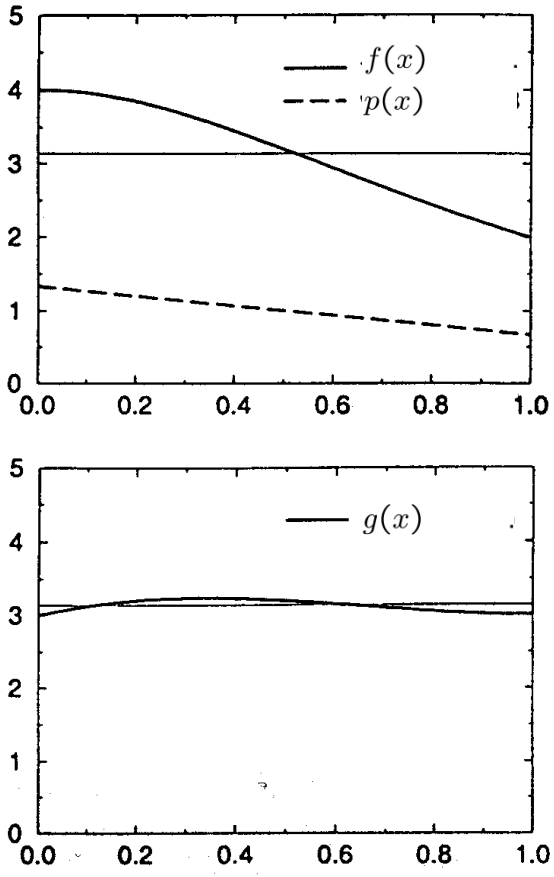
or by

- decreasing the variance by using importance sampling

Importance sampling

$$I = \int_a^b f(x)\ dx$$

$$= \int_a^b \frac{f(x)}{p(x)}\ p(x)dx \equiv \int_a^b g(x)\ p(x)dx$$



1. Choose N points x_i at random with probability $p(x)$ within the integration interval

2. Determine the mean value

$$I_N = \langle g \rangle_p \equiv \frac{1}{N} \sum_{i=1}^{N} g(x_i) \equiv \frac{1}{N} \sum_{i=1}^{N} f_i/p_i$$

and the variance

$$\sigma_g^2 = \left\langle (g - \langle g \rangle_p)^2 \right\rangle_p = \left\langle g^2 \right\rangle_p - \langle g \rangle_p^2$$

3. Approximate the value of the integral as

$$I = I_N \pm \frac{\sigma_g}{\sqrt{N}}$$

Notice: $\sigma_g \ll \sigma_f$

Importance sampling - example

Numerical example

$$I = \int_0^1 \frac{4}{1+x^2}\, dx$$

$$p(x) = 1$$

	N	I_N	σ_f	$\sigma_I = \sigma_f/\sqrt{N}$
MC	10^1	3.18266	0.61954	0.19592
MC	10^2	3.20677	0.60315	0.06032
MC	10^3	3.14463	0.65030	0.02056
MC	10^4	3.14380	0.63989	0.00640
MC	10^5	3.14096	0.64438	0.00204
exact		3.14159	0.64310	

$$p(x) = (4-2x)/3$$

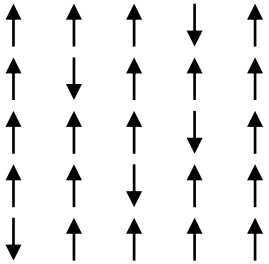
	N	I_N	σ_g	$\sigma_I = \sigma_g/\sqrt{N}$
MC	10^1	3.18155	0.06746	0.02133
MC	10^2	3.14107	0.08327	0.00833
MC	10^3	3.13931	0.08102	0.00256
MC	10^4	3.14119	0.08072	0.00081
MC	10^5	3.14200	0.08012	0.00025
exact		3.14159	0.08002	

Importance sampling

Multidimensional generalization of $I = \int_a^b g(x)\, p(x) dx$

Discrete system

e.g. the Ising model



Probability distribution function

$$P_i = \frac{1}{Z} \exp\left[-\beta E_i\right]$$

Partition function

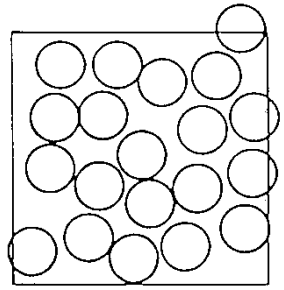
$$Z = \sum_i \exp\left[-\beta E_i\right]$$

Mean value, e.g. the energy

$$\langle E \rangle = \sum_i E_i P_i$$

Continuous system

e.g. an atomic system



Probability distribution function

$$P(\mathbf{r}^N) = \frac{1}{Z} \exp\left[-\beta V(\mathbf{r}^N)\right]$$

Partition function

$$Z = \int d\mathbf{r}^N \exp\left[-\beta V(\mathbf{r}^N)\right]$$

Mean value, e.g. the potential energy

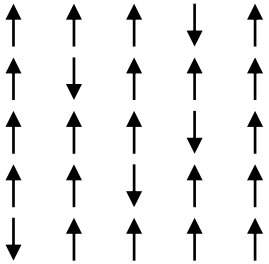
$$\langle V \rangle = \int d\mathbf{r}^N V(\mathbf{r}^N) P(\mathbf{r}^N)$$

Importance sampling

Multidimensional generalization of $I = \int_a^b g(x)\, p(x) dx$

Discrete system

e.g. the Ising model



Probability distribution function

$$P_i = \frac{1}{Z} \exp\left[-\beta E_i\right]$$

Partition function

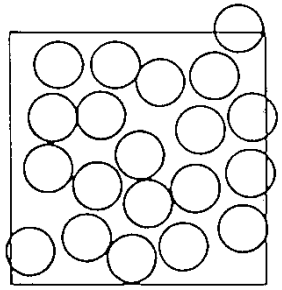
$$Z = \sum_i \exp\left[-\beta E_i\right]$$

Mean value, e.g. the energy

$$\langle E \rangle = \sum_i E_i P_i$$

Continuous system

e.g. an atomic system



Probability distribution function

$$P(\mathbf{r}^N) = \frac{1}{Z} \exp\left[-\beta V(\mathbf{r}^N)\right]$$

Partition function

$$Z = \int d\mathbf{r}^N \exp\left[-\beta V(\mathbf{r}^N)\right]$$

Mean value, e.g. the potential energy

$$\langle V \rangle = \int d\mathbf{r}^N V(\mathbf{r}^N) P(\mathbf{r}^N)$$

How to generate ?

Monte Carlo

Content:

- **Monte Carlo integration**
- **Variance reduction / Importance sampling**
- **Markov chains**
- **The Metropolis algorithm**
- **Error estimate**

The Metropolis algorithm

A general method of sampling arbitrary highly-dimensional probability distributions by taking a random walk through configuration space.

- It was introduced by Metropolis *et al.* 1953 to determine the equation of state for a hard sphere liquid.
- It uses the Markov chain technique to generate configurations with knowledge of only relative probabilities, no absolute probabilities have to be known.
- The algorithm guides the Markov chain to important regions by rejecting unlikely configurations.
- The generated configurations become correlated and care has to be taken when evaluating proper error bars.

Example – the weather in Göteborg

The weather in Göteborg is either sunny (1), cloudy (2) or rainy (3).

A sunny day is never followed by another sunny day. Rainy or cloudy weather is equally probable after a sunny day. A rainy or cloudy day is followed with 50% probability by another day with the same weather. If, on the other hand, the weather is changing from cloudy or rainy weather, the following day will be sunny only in half of the cases.

$$\mathbf{W} = \begin{bmatrix} 0 & 0.25 & 0.25 \\ 0.5 & 0.5 & 0.25 \\ 0.5 & 0.25 & 0.5 \end{bmatrix}$$

Example – the weather in Göteborg

$$\mathbf{P}(s) = \mathbf{W}^s \mathbf{P}(0)$$

s	p_1	p_2	p_3
0	1.00000	0.00000	0.00000
1	0.00000	0.50000	0.50000
2	0.25000	0.37500	0.37500
3	0.18750	0.40625	0.40625
4	0.20312	0.39844	0.39844
5	0.19922	0.40039	0.40039
6	0.20020	0.39990	0.39990
7	0.19995	0.40002	0.40002
8	0.20001	0.39999	0.39999
9	0.20000	0.40000	0.40000
10	0.20000	0.40000	0.40000
∞	0.20000	0.40000	0.40000

s	p_1	p_2	p_3
0	0.50000	0.00000	0.50000
1	0.12500	0.37500	0.50000
2	0.21875	0.37500	0.40625
3	0.19531	0.39844	0.40625
4	0.20117	0.39844	0.40039
5	0.19971	0.39990	0.40039
6	0.20007	0.39990	0.40002
7	0.19998	0.39999	0.40002
8	0.20000	0.39999	0.40000
9	0.20000	0.40000	0.40000
10	0.20000	0.40000	0.40000
∞	0.20000	0.40000	0.40000

Example – the weather in Göteborg

The weather in Göteborg is either sunny (1), cloudy (2) or rainy (3).

A sunny day is never followed by another sunny day. Rainy or cloudy weather is equally probable after a sunny day. A rainy or cloudy day is followed with 50% probability by another day with the same weather. If, on the other hand, the weather is changing from cloudy or rainy weather, the following day will be sunny only in half of the cases.

$$\mathbf{W} = \begin{bmatrix} 0 & 0.25 & 0.25 \\ 0.5 & 0.5 & 0.25 \\ 0.5 & 0.25 & 0.5 \end{bmatrix}$$

Solution of

$$\mathbf{P}^{\text{st}} = \mathbf{W}\mathbf{P}^{\text{st}}$$

implies that

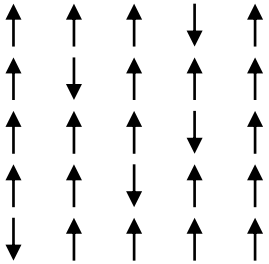
$$\mathbf{P}^{\text{st}} = \begin{bmatrix} 0.2 \\ 0.4 \\ 0.4 \end{bmatrix}$$

Importance sampling

Multidimensional generalization of $I = \int_a^b g(x)\, p(x) dx$

Discrete system

e.g. the Ising model



Probability distribution function

$$P_i = \frac{1}{Z} \exp\left[-\beta E_i\right]$$

Partition function

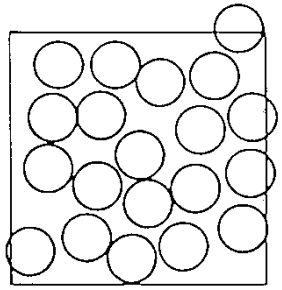
$$Z = \sum_i \exp\left[-\beta E_i\right]$$

Mean value, e.g. the energy

$$\langle E \rangle = \sum_i E_i P_i$$

Continuous system

e.g. an atomic system



Probability distribution function

$$P(\mathbf{r}^N) = \frac{1}{Z} \exp\left[-\beta V(\mathbf{r}^N)\right]$$

Partition function

$$Z = \int d\mathbf{r}^N \exp\left[-\beta V(\mathbf{r}^N)\right]$$

Mean value, e.g. the potential energy

$$\langle V \rangle = \int d\mathbf{r}^N V(\mathbf{r}^N) P(\mathbf{r}^N)$$

How to generate ?

$$\mathbf{P}^{\text{st}} = \mathbf{W}\mathbf{P}^{\text{st}}$$

The idea is to use $\mathbf{P}^{\text{st}}$ in a simulation study.

Importance sampling

Mean value, e.g. the potential energy

$$\langle V \rangle = \int d\mathbf{r}^N V(\mathbf{r}^N) P(\mathbf{r}^N)$$

However, we do not have control over $\mathbf{P}^{\text{st}}$.

It is a consequence of a given transition matrix $\mathbf{W}$.

How to construct the matrix W to obtain P^{st} ?

Basic idea

If you can find a matrix $\mathbf{W}$ with the following properties

1. $0 \leq w_{nm} \leq 1 \quad \forall\ n \text{ and } m$
2. $\sum_{n=1}^{M} w_{nm} = 1 \quad \forall\ m$
3. w_{nm} is ergodic
4. $w_{mn}p_n = w_{nm}p_m$

then the Markov process will, in the long run, produce states distributed according to the probability distribution $\mathbf{P}$.

Example – the weather in Göteborg

The weather in Göteborg is either sunny (1), cloudy (2) or rainy (3).

A sunny day is never followed by another sunny day. Rainy or cloudy weather is equally probable after a sunny day. A rainy or cloudy day is followed with 50% probability by another day with the same weather. If, on the other hand, the weather is changing from cloudy or rainy weather, the following day will be sunny only in half of the cases.

$$\mathbf{W} = \begin{bmatrix} 0 & 0.25 & 0.25 \\ 0.5 & 0.5 & 0.25 \\ 0.5 & 0.25 & 0.5 \end{bmatrix}$$

$$\mathbf{P}^{st} = \begin{bmatrix} 0.2 \\ 0.4 \\ 0.4 \end{bmatrix}$$

Example – the weather in Göteborg

The weather in Göteborg is either sunny (1), cloudy (2) or rainy (3).

A sunny day is never followed by another sunny day. Rainy or cloudy weather is equally probable after a sunny day. A rainy or cloudy day is followed with 50% probability by another day with the same weather. If, on the other hand, the weather is changing from cloudy or rainy weather, the following day will be sunny only in half of the cases.

$$\mathbf{W} = \begin{bmatrix} 0 & 0.25 & 0.25 \\ 0.5 & 0.5 & 0.25 \\ 0.5 & 0.25 & 0.5 \end{bmatrix}$$

$$\mathbf{P}^{st} = \begin{bmatrix} 0.2 \\ 0.4 \\ 0.4 \end{bmatrix}$$

$$\mathbf{W} = \begin{bmatrix} 0.4 & 0.2 & 0.1 \\ 0.4 & 0.3 & 0.5 \\ 0.2 & 0.5 & 0.4 \end{bmatrix}$$

Monte Carlo

Content:

- **Monte Carlo integration**
- **Variance reduction / Importance sampling**
- **Markov chains**
- **The Metropolis algorithm**
- **Error estimate**