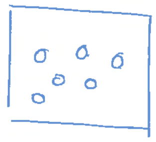
If you want a life problem try to prove anything about the Metropolis algorithm?

It is very hard and not much is known. This is one of the most used algorithms in scientific computing!

Original metropolis application: Hard disks in a box

Fix $n > 0$ disks and consider placing n disks of radius h in a box square so that they don't overlap.

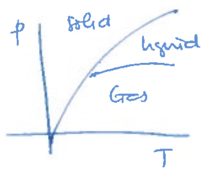


The set of all possible configs is a compact subset in $\mathbb{R}^{2n}$ and as such has a uniform distribution (normalised Lebesgue measure).

Problem: Pick a conf at random

open problem: how large/small must h be for both this to be true (e.g. connected, π_1).

Who cares?



phase space of many substances   why do they all look the same?

Kirk 1920s — there should be a phase transition if there are just repulsive forces. Simplest such is disks in a box. Are there phase transitions.

area covered > 0.705 $\Rightarrow$ disks all lined up

< 0.705 look random.

Metropolis invented an algorithm to sample from the uniform distribution for the disk problem.

In this case it goes: Start at x
pick disk at random and move its center in any dir ε. If ok go, if not stay.

To get estimates for one disk in one dimension is already extremely difficult.

Lecture 5   I will explain the first theorem I proved about shuffling cards:

Random transpositions — an introduction to group representations

In english: n cards face down on the table. Pick a card w/ left and then right and switch them. How many switches does it take to mix them.

S_n permutation grp.
$$Q(\pi) = \begin{cases} \frac{1}{n} & \text{if } \pi = id \\ \frac{2}{n^2} & \text{if } \pi = (ij) \\ 0 & \text{otherwise} \end{cases} \qquad (Q*Q)(\pi) = \sum_{\eta} Q(\eta)Q(\pi\eta^{-1})$$

The uniform distribution is $U(\pi) = \frac{1}{n!}$ $\|Q^{*k} - U\| = \max_A \|Q^{*k}(A) - U(A)\|$

Given $\varepsilon > 0$ how large should k be so $\|Q^{*k} - U\| < \varepsilon$

Theorem (Shah Shahani, D.) $k = \frac{1}{2} n \log n + cn$, $c > 0$

$$\|Q^{*k} - U\|_{TV} \leq 2e^{-c}$$

I was working at Bell Labs and trying to find an optimal strategy for some game. How do we tell a computer to pick a random permutation: the std algorithm for doing a random transposition. Pick a # between 1 and n and transpose. Then from 2 to n, etc.... This is still the best way we know.

TÉCNICO LISBOA

This program was thousands of lines and worked for 2 days. Something was not right - something was not monotone. The programmer had taken random to (when should be) random transposition!

$n = 52$, $\|\cdots\| < \frac{1}{1000}$ $c = \log 2000$

$n \log n \sim 150$

For a minute let G be a finite group, Q a probability distribution on the group.
We want to study the rate of convergence to the uniform distribution (this will happen if "Q is not stupid" as in switching the top 2 cards)
supp(Q) not contained in the coset of a subgroup.

<u>Def</u>: A representation of G is a map $\rho : G \to GL_d(V)$ (V vector space, $d = \dim V$) so that $\rho(st) = \rho(s)\rho(t)$

<u>Example</u>: S_n has a n-dimensional rep obtained by permuting coordinates.
Sign representation $\rho(\sigma) = \mathrm{sgn}(\sigma)$ 1-dimensional representation.
trivial representation sends everything to Id

C_n cyclic group $\rho_j(k) = e^{2\pi i j k/n}$

ρ is irreducible if $\nexists\, V_0 \subsetneq V$, $V_0 \neq 0$ so that $\rho(s)V_0 \subset V_0$ for all s.

<u>Ex</u>: The n dimensional rep is not irreducible. Take $V_0 = \{ v \mid \sum_j v_j = 0 \}$. It's orthogonal complement is also invariant.
this one is irreducible.

<u>Def</u>: Q probability, ρ representation. The Fourier transform of Q at ρ is

$$\hat{Q}(\rho) := \sum_{s \in G} Q(s)\rho(s) \quad \text{(it is a matrix)}$$

<u>FACTS</u>: (1) $\widehat{Q_1 * Q_2}(\rho) = \hat{Q}_1(\rho)\hat{Q}_2(\rho)$ so $\widehat{Q^{*k}}(\rho) = \hat{Q}(\rho)^k$

(2) If ρ is irreducible and $U(s) = \frac{1}{|G|}$ then $\hat{U}(\rho) = \begin{cases} 1 & \text{if } \rho \text{ is trivial} \\ 0 & \text{otherwise} \end{cases}$

(3) Fourier inversion theorem: $Q(s) = \frac{1}{|G|} \sum_{\rho \in \hat{G}} d_\rho \, \mathrm{Tr}\big(\hat{Q}(\rho)\, \rho(s^{-1})\big)$

set of irreducible reps

($\sum_{\rho \in \hat{G}} d_\rho^2 = |G|$ so there are only finitely many reps.)

On C_n $\hat{Q}(j) = \sum_k Q(k) e^{2\pi i j k/n}$

$Q(k) = \frac{1}{n} \sum_{j=0}^{n-1} \hat{Q}(j) e^{-2\pi i j k/n}$ (this is the usual discrete Fourier transform)

<u>Example</u>: C_2 $Q(1) = Q(-1) = \frac{1}{2}$ $\hat{Q}(j) = \frac{1}{2}\left(e^{\frac{2\pi i j}{n}} + e^{\frac{-2\pi i j}{n}}\right) = \cos\left(\frac{2\pi j}{n}\right)$

Thus in order to show that $Q^{*k} \to U$ we want to prove $\hat{Q}(\rho)^k \to 0$ for $\rho \neq 1$ (want to do this quantitatively)

<u>Theorem</u> (upper bound) lemma $4\|Q^{*k} - U\|_{TV}^2 \leq \sum_{\rho \in \hat{G} \setminus \{\mathrm{triv}\}} d_\rho \|\hat{Q}(\rho)\|^k$ where $\|A\| = \mathrm{Tr}(AA^*)$

<u>Proof</u>: $4\left(\frac{1}{2}\sum_{s \in G} |Q^{*k}(s) - U(s)|\right)^2 \underset{\text{Cauchy Schwarz}}{\leq} |G| \sum_{s \in G} |Q^{*k}(s) - U(s)|^2 \underset{\text{Plancherel}}{=} \sum_{\rho \in \hat{G}} d_\rho \|\hat{Q}(\rho) - \hat{U}(\rho)\|$

(Fourier transform is an isometry)

1 if ρ trivial
hence we obtain the required formula $\square$

On C_n $\quad Q(1) = Q(-1) = \frac{1}{2}$

$$4\,\|Q^{*k} - U\|^2_{TV} \leq \sum_{j=1}^{n-1} \cos\left(\frac{2\pi j}{n}\right)^{2k} \leq 10\, e^{-k\pi^2 x / n^2}$$

use $\cos x = 1 - \frac{x^2}{2} + O(x^4)$

Back to S_n, Representations are indexed by partitions of n

$$\lambda = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$$

e.g $n=4$: $\quad 4 \quad 31 \quad 22 \quad 211 \quad 1111$

$p(n) = \#$ partitions of n $\quad = \#$ reps of S_n.

Q random transpositions. Observe $\rho(s^{-1})\,\hat{Q}(\rho)\,\rho(s) = \rho(s^{-1})\left(\sum Q(\pi)\rho(\pi)\right)\rho(s) = \sum Q(\pi)\,\rho(s^{-1}\pi s)$

$$= \sum Q(s^{-1}\pi s)\,\rho(\pi) = \hat{Q}(\rho)$$

($\hat{Q}$ is invariant under conjugation)

[because Q is supported on transpositions which form a conjugacy class]

<u>Schur's lemma</u>: (ρ_1, V_1) and (ρ_2, V_2) irreducible reps and $m: V_1 \to V_2$ a linear map commuting with the action of the group.

$$\rho_2(s)\,m = m\,\rho_1(s)$$

Then ① If $\rho_1 \neq \rho_2$ then $m = 0$

② If $\rho_1 = \rho_2$ then $m = c\,Id$

Schur's lemma for us says that $\hat{Q}(\rho) = c\,Id$ where $c = \frac{1}{n} + \frac{n-1}{n}\,\frac{\chi_\rho(12)}{d_\rho}$ where $\chi_\rho(12) = Tr(\rho(12))$

compute the trace.

We get $$4\,\|Q^{*k} - U\|^2_{TV} \leq \sum_{\substack{\lambda \text{ partition of } n \\ \lambda \neq 1\cdots1}} d_\lambda^2 \left(\frac{1}{n} + \frac{n-1}{n}\,\frac{\chi_\rho(12)}{d_\lambda}\right)^{2k}$$

square comes from trace of the identity

λ identified with rep it corresponds to

Roughly $\frac{\chi_\rho(12)}{d_\rho} \sim \frac{1}{2}$ when we look at tables of characters

Therefore RHS $\leq \left(\frac{1}{2}\right)^{2k} \sum_{\lambda \vdash n} d_\lambda^2 = \left(\frac{1}{2}\right)^{2k} n! \sim e^{-2k \log 2 + n\log n + n \cdots}$

partitions

But  for $\lambda = (n-1, 1)$ trace of a transposition is $n-3$. The term we get is

$$(n-1)^2\left(\frac{1}{n} + \frac{n-1}{n}\,\frac{n-3}{n-1}\right)^{2k} = (n-1)^2\left(1 - \frac{2}{n}\right)^{2k} \leq e^{-\frac{2k}{n} + 2\log n} \leq e^{-c}$$

not close to $\frac{1}{2}$

but it is small if $k = \frac{1}{2} n \log n + c$

so two different arguments suggest the answer is as required.

In 1900 Frobenius proved that $$\frac{\chi_\rho(12)}{d_\lambda} = \frac{1}{n(n-1)} \sum_{j=1}^{r} (\lambda_j - j)(\lambda_j - (j-1)) - j(j-1) \qquad \lambda = \lambda_1, \dots \lambda_j$$

Here very good approximations to this sum come from the hook-length formula for the dimension of the representations of the symmetric group

<u>How to learn more</u>: G. James + M. Liebeck - "Representation Theory"

Serre - Representations of finite groups

Bruce Sagan - Rep theory of the symmetric group

Diaconis - Use of group representations in Prob. and Statistics.

what it really takes to go further is <u>to pick a group you want to learn</u>.