Random Walks and Markov Chains

An intuitive investigation into convergence/transience between 2/3D

A. McInerny[1]

[1]St Hilda's College
University of Oxford

Physics Talk

Quote!

'A drunk man will find his way home, but a drunk bird may get lost forever'

—Shizuo Kakatani

Rules of the game–the random walk

For a simple symetric random walk on a d-dimensional integer lattice $\mathbb{Z}^d$,

- The sequence $\{S_i\}_{i=0,\dots,N}$ defined by
  - it's locations at step n, $S_n = S_0 + \sum_{i=1}^{n} X_i$
  - where X_i is i.i.d from the d-dimensional standard unit vectors
  - and S_0 is an initial location

1. Markov Chains

A Markov chain is defined by 4 characteristics:

1. State space
2. Transition probabilities
3. Initial distribution
4. Memorylessness

State space

Set of states we can visit. This could really be anything...

(a) Cheesegrater 1

(b) Cheesegrater 2

Figure: Cheesegraters

State space for random walks

For our random walk, in 2D it'll look like $\mathbb{Z}^2$:

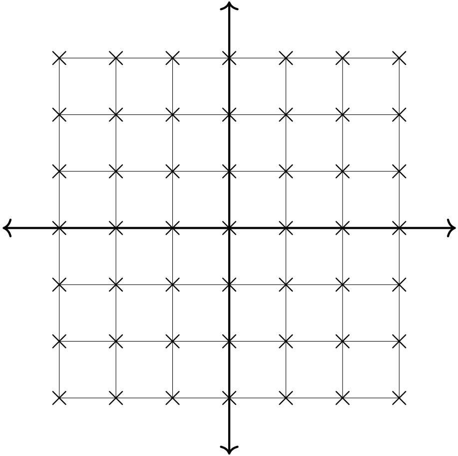


Figure: State space for 2D random walk

2. Transition probabilities

Probabilities of switching to any other state.

For a 1D Markov chain, it might look like this:

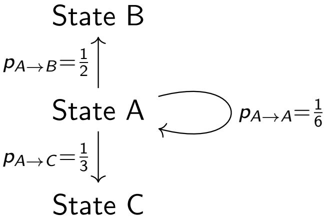


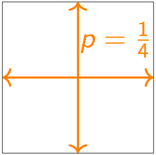
Figure: Possible transition probabilities

Transition probabilities for random walks

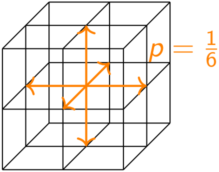
For our simple symmetric d-dimensional random walk, for neighbouring states we should have

$$\mathbb{P}(X_i = \hat{\mathbf{e}}_j) = \mathbb{P}(X_i = -\hat{\mathbf{e}}_j) = \frac{1}{2d}, \quad j = 1, ..., d$$

and 0 for any other states.



(a) 2D



(b) 3D

Figure: Symmetric random walks on $\mathbb{Z}^2$ and $\mathbb{Z}^3$

3. Initial distribution

Specified initial probabilities of starting in each state. Our walk starts at the origin:

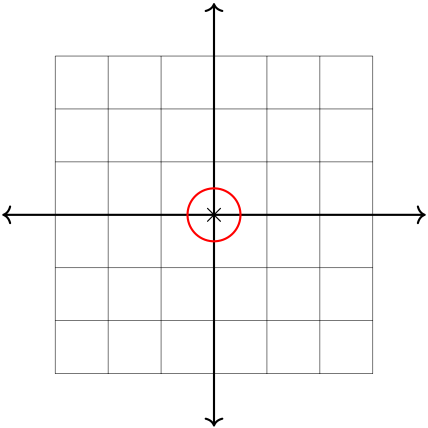
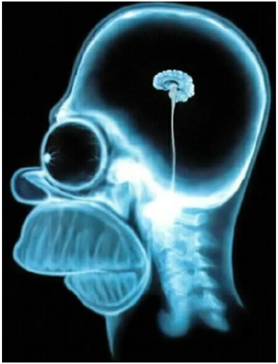


Figure: Random walk starts at the origin

4. Memorylessness

The preceding path has no influence on the following steps.
Mathematically,

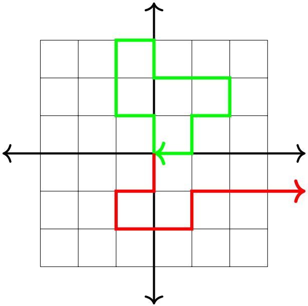
$$\mathbb{P}(S_{n+1} = s_{n+1} | S_1 = s_1, ..., S_n = s_n) = \mathbb{P}(S_{n+1} = s_{n+1})$$



The question

We want to know:

- Will a walk return to the origin ($\mathbb{P}(\text{return}) = 1 \implies$ **recurrent**)
- Or will it get lost ($\mathbb{P}(\text{return}) < 1 \implies$ **transient**)



How do we know $\mathbb{P}(\text{return})$?

Introducing variables to help: V

$V =$ number of returns to the origin (where $n \to \infty$)

For a recurrent walk:

$$\mathbb{P}(V = \infty) = 1$$
$$\implies \mathbb{E}[V] = \infty$$

Some strategic maths shows for a transient walk:

$$\mathbb{E}[V] < \infty$$

When we know about $\mathbb{E}[V]$, finished!

And now J!

Introduce...

$$J_n = \begin{cases} +1 & \text{if return to origin at step } n \\ 0 & \text{else} \end{cases}$$

So

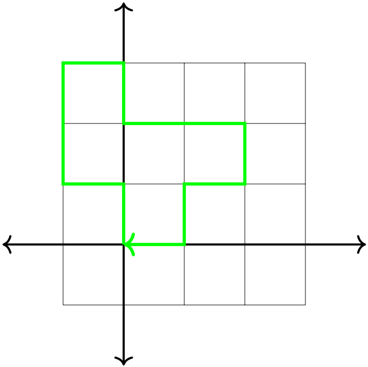
$$V = \sum_{n=0}^{\infty} J_n$$

And

$$\mathbb{E}[V] = \sum_{n=0}^{\infty} \mathbb{E}[J_n] = \sum_{n=0}^{\infty} \mathbb{P}(S_n = 0)$$

Now we can calculate $\mathbb{E}[V]$, and know a way to relate $\mathbb{E}[V]$ to transcience or recurrence. General result for Markov chains.

Computing $\mathbb{P}(S_n = 0)$



steps left $=$ steps right, steps up $=$ steps down

$\implies$ Total number of steps is even, so

$$\mathbb{P}(S_n = 0) = 0 \quad \text{for odd } n$$

Now computing $\mathbb{P}(S_{2n} = 0)$

Fix, say, $n = 16$;

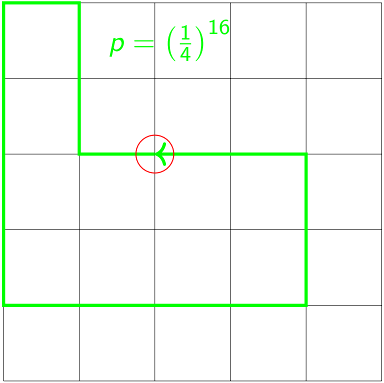


Figure: One possible recurrent path on $\mathbb{Z}^2$, $n = 16$

Over-counting

But there are many paths that return to the origin after 16 steps, possibly with multiple origin intersections.

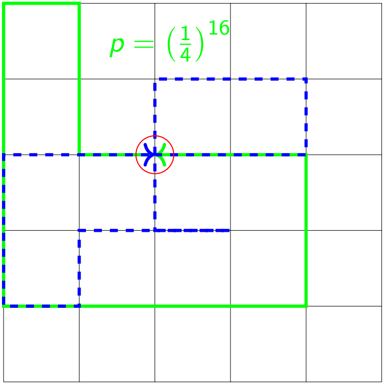


Figure: Two distinct paths with same n

(Somewhat) intuitively, there are $n! = 16!$ of these such paths.

Accounting for over-counting

So actually,

$$\mathbb{P}(S_{2n} = 0) = \left(\frac{1}{4}\right)^{2n} (\text{number of paths})$$

Degenerate paths:

- Swap 'pieces of path' and have no effect on route take.
- If i steps 'left':
  - also i steps 'right'
  - and $n - i$ steps 'up' and also 'down'.

So total number of return paths with $2n$ total steps and i steps 'left' is

$$\frac{(2n)!}{i!i!(n-i)!(n-i)!}$$

thus (since i ranges from 0 to n)

$$\mathbb{P}(S_{2n} = 0) = \left(\frac{1}{4}\right)^{2n} \sum_{k=0}^{n} \frac{(2n)!}{i!i!(n-i)!(n-i)!} \sim \frac{1}{n}$$

The answer

So...

$$\mathbb{E}[V] = \sum_{n=0}^{\infty} \mathbb{P}(S_{2n} = 0) \sim \sum_{n=1}^{\infty} \frac{1}{n} \to \infty$$

and a 2D simple random walk is guaranteed to pass the origin an infinite number of times!!!

Speedy generalisation to $\mathbb{Z}^3$

Now,

$$\mathbb{P}(S_{2n} = 0) = \left(\frac{1}{6}\right)^{2n} (\text{number of paths})$$

where (by the same logic) the total number of return paths with $2n$ total steps, where i go 'left', j 'up' and $n - i - j$ 'out', etc,

$$\frac{(2n)!}{i!i!j!j!(n-i-j)!(n-i-j)!}$$

and

$$\mathbb{P}(S_{2n} = 0) = \left(\frac{1}{6}\right)^{2n} \sum_{i,j=0}^{n} \frac{(2n)!}{i!i!j!j!(n-i-j)!(n-i-j)!} \sim \frac{1}{n^{3/2}}$$

The big reveal!

And in 3D:

$$\mathbb{E}[V] = \sum_{n=1}^{\infty} \left(\frac{1}{6}\right)^{2n} \sum_{i,j=0}^{n} \frac{(2n)!}{i!i!j!j!(n-i-j)!(n-i-j)!} \sim \sum_{n=1}^{\infty} \frac{1}{n^{3/2}} < \infty$$

So a 3D random walk isn't guaranteed to have an infinite number of returns to the origin - there's a chance it will get lost!

Higher dimensions

George Pólya, 1921

- $\mathbb{P}(\text{return in } 1D) = 1$
- $\mathbb{P}(\text{return in } 2D) = 1$
- $\mathbb{P}(\text{return in } 3D) = 0.3405373...$
- $\mathbb{P}(\text{return in } 4D) = 0.193206...$
- $\mathbb{P}(\text{return in } 5D) = 0.135178...$
- $\mathbb{P}(\text{return in } 6D) = 0.104715...$
- $\mathbb{P}(\text{return in } 7D) = 0.0858449...$
- $\mathbb{P}(\text{return in } 8D) = 0.0729126...$

Some intuition

Why does the return probability decrease with dimension?

Why is the recurrence/transience distinction between the 2nd and 3rd dimension?

- For any dimension, expected number of steps to reach distance r is r^2
  - 'Volume' of circle? $c_2 r^2$
  - Volume of sphere? $c_3 r^3$
  - 'Volume of d-sphere? $c_d r^d$
- ratio $\frac{\text{optimistic number of points visited when first } r \text{ away}}{\text{total number of points inside } r}$ tends to zero for $d > 2$

Applications

- Modelling stock market pricing in financial theory (boo)
- Statistics of population drift in population genetics
- Model cascades of neurons firing in brain research
- Estimate size of webs in computer science
- Model Brownian motion in statistical physics!

Thanks!

Thanks for your attention.

No mean questions, please!

References/further reading

- Random Walks, Wikipedia, hyperlink
- Markov Chains, Wikipedia, hyperlink
- Random Walks, University of Chicago, hyperlink
- Polyas Random Walk Constants, Wolfram Mathword, hyperlink
- Random Walks, Univesrity of Washington, hyperlink
- Why do random walks get lost in 3D?, Ari Seff, hyperlink
- How many paths intuition?, Atharva, human