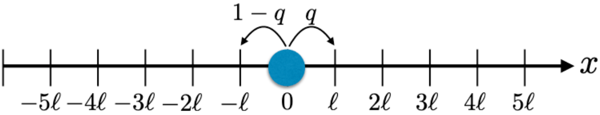
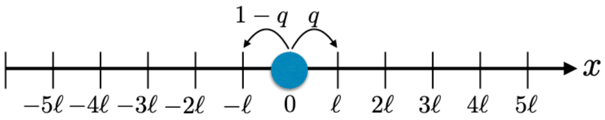
A 1 dimensional random walk



A particle starts at $x = 0$ and jumps left or right with probability q, left with probability $1 - q$.

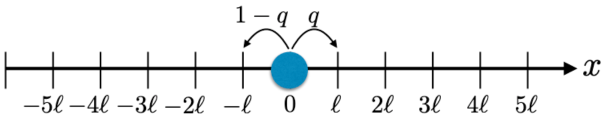
A 1 dimensional random walk



A particle starts at $x = 0$ and jumps left or right with probability q, left with probability $1 - q$.

What is the probability that we find the particle at position x after n jumps?

A 1 dimensional random walk



A particle starts at $x = 0$ and jumps left or right with probability q, left with probability $1 - q$.

What is the probability that we find the particle at position x after n jumps?

The particle is at position x, iff it has made k jumps to the right where $x = k - (n - k)$. The probability of having jumped k times to the right is a binomial:

$$P(k, n) = \binom{n}{k} q^k (1 - q)^{n-k}$$

Deriving diffusion

According to the central limit theorem, the Binomial distribution $P(k, n)$ approaches the Gaussian with mean $\mu = n(2q - 1)$ and variance $\sigma^2 = 4nq(1 - q)$[1]:

$$p(x, n) \approx \mathcal{N}\Big(n(2q - 1), 4nq(1 - q)\Big).$$

[1] For $q = \frac{1}{2}$, it's $\mathcal{N}(0, n)$.

Scaling the master equation

The master equation gives a recursive relation between the probability distributions at time n:

$$
\begin{aligned}
P(x, n+1) &= qP(x-1, n) + (1-q)P(x+1, n) \\
P(x, 0) &= \delta_0(x).
\end{aligned}
$$

[2]Since the variance of the position grows at rate $\sqrt{\Delta t}$

Scaling the master equation

The master equation gives a recursive relation between the probability distributions at time n:

$$P(x, n+1) = qP(x-1, n) + (1-q)P(x+1, n)$$
$$P(x, 0) = \delta_0(x).$$

What if we speed up the process and make the shrink the step size? The correct scaling[2] is to let $\Delta x = \sqrt{\Delta t}$.

Derivation: write $p + \delta t$

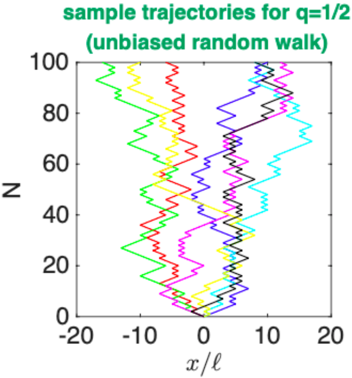
[2]Since the variance of the position grows at rate $\sqrt{\Delta} t$

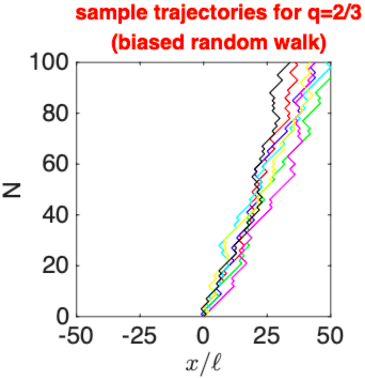
With drift

What if $q \neq \frac{1}{2}$? Then, there is some drift involved:

With drift

What if $q \neq \frac{1}{2}$? Then, there is some drift involved:





4

With drift

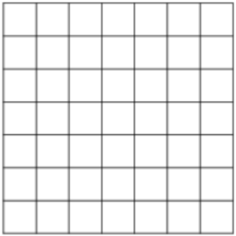
The Fokker-Planck equation is like the heat equation, but with drift:

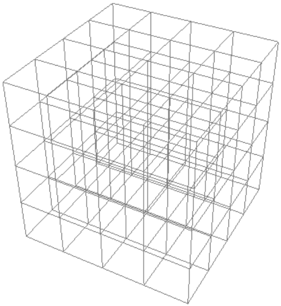
$$p_t = -vp_x + \Delta p$$

Derive this by Taylor-expanding the master equation,

Diffusion in n dimensions

All of this can be generalized to n dimensions, of course.

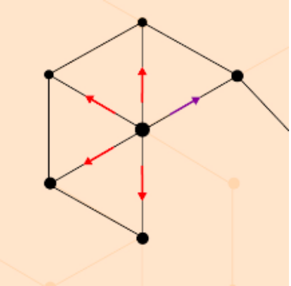




- Particle moves in each direction with probability $\frac{1}{2^n}$.
- Shizuo Kakutani: "A drunk man will find his way home, but a drunk bird may get lost forever."

Graph random walk

What about a random walk on the nodes of a graph?



The movement of the random walker has to do with the view of the adjacency matrix as a linear operator!

Graph random walk

The normalized adjacency matrix is

$$W(i,j) = \frac{1}{\deg(i)} A(i,j)$$

E.g.

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \Rightarrow \quad W = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{pmatrix}$$

Graph random walk

The normalized adjacency matrix is

$$W(i,j) = \frac{1}{\deg(i)} A(i,j)$$

E.g.

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \Rightarrow \quad W = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{pmatrix}$$

Row i can be thought of as a discrete probability distribution of the next position given that we start at node i.

Graph random walk

How does the normalized adjacency matrix W behave as a linear operator?

$$W : \mathbb{R}^n \to \mathbb{R}^n$$

Graph random walk

How does the normalized adjacency matrix W behave as a linear operator?

$$W : \mathbb{R}^n \to \mathbb{R}^n$$

Let $v = e_k$ for some basis vector e_k of $\mathbb{R}^n$. We can think of this as an initial state: a particle at node k.

Then Wv represents the probability distribution after one walker step. W^t represents the probability distribution after t steps. W is a Markov transition matrix.

Graph random walk

How does the normalized adjacency matrix W behave as a linear operator?

$$W : \mathbb{R}^n \to \mathbb{R}^n$$

Let $v = e_k$ for some basis vector e_k of $\mathbb{R}^n$. We can think of this as an initial state: a particle at node k.

Then Wv represents the probability distribution after one walker step. W^t represents the probability distribution after t steps. W is a Markov transition matrix.

v can also be any normalized vector.

Graph random walk

We can apply the theory of stochastic processes and Markov chains to the graph random walk problem.

- Existence of stationary distributions
- Convergence to stationary distributions
- Eigenvalues and eigenvectors of W. Property: its largest eigenvalue is always $\lambda = 1$ and all the rest have $|\lambda_i| < 1$.

Application: PageRank

Let the graph be the (directed) graph of hyperlinks on the web. Each vertex is a webpage, with edges pointing towards the websites it links.

[3] The algorithm actually has some stochasticity, where we click a link with probability α and go to a completely new webpage with probability $1 - \alpha$.

Application: PageRank

Let the graph be the (directed) graph of hyperlinks on the web. Each vertex is a webpage, with edges pointing towards the websites it links.

The random walk tells us what would happen if we started at a webpage and follow a random link on the page, and repeat this process.[3]

[3] The algorithm actually has some stochasticity, where we click a link with probability α and go to a completely new webpage with probability $1 - \alpha$.

Application: PageRank

Let the graph be the (directed) graph of hyperlinks on the web. Each vertex is a webpage, with edges pointing towards the websites it links.

The random walk tells us what would happen if we started at a webpage and follow a random link on the page, and repeat this process.[3]

The stationary distribution gives us a ranking of the most influential/popular pages on the internet.

[3] The algorithm actually has some stochasticity, where we click a link with probability α and go to a completely new webpage with probability $1 - \alpha$.

Diffusion equation

$$u_t = \Delta u.$$

[4] I tried to derive the graph heat equation from the random walk/taylor expansion method, but it was unclear what scaling to use.

Diffusion equation

$$u_t = \Delta u.$$

$$\Delta u = \mathrm{div} \nabla u(x) = \partial_{x_1} u + \partial_{x_2} u + \cdots + \partial_{x_n} u$$

What does the graph version look like?[4]

[4] I tried to derive the graph heat equation from the random walk/taylor expansion method, but it was unclear what scaling to use.

Graph diffusion

The directional derivatives are replaced by derivatives in the direction of each neighboring node; the graph analogue of $\mathrm{div}\nabla u$ is then

$$
\begin{aligned}
\Delta u(i) &= \sum_{j \in N(i)} u_j - u_i \\
&= \sum_{j=1}^{n} A_{ij}(u_j - u_i) \\
&= \sum_{j=1}^{n} A_{ij} u_j - u_i \mathrm{deg}(i) \\
&= (A - D)u
\end{aligned}
$$

where $D = \mathrm{diag}(\mathrm{deg}(i))$.

$A - D =: L$ is called the graph Laplacian.

Graph diffusion

What do we do with this?

- $u_t = Lu$ is a linear system of ODEs, so is solvable.
- Since L is symmetric: real eigenvalues, and orthogonal eigenvectors
- L is singular with eigenvalue 0 and corresponding eigenvector $\mathbf{1}$.

$$\boxed{\vec{u}(t) = \sum_{i=1}^{n} a_i(0) e^{-c\lambda_i t} \vec{v}_i}$$

Spectral Solution

Graph diffusion: uses

The eigenvalues of the graph Laplacian tell us a lot about the structure of the graph

- $\lambda_1 = 0$ always. The number of eigenvectors corresponding to λ_1 is the number of connected components of the graph.
- All eigenvalues must be nonnegative[5]

[5] If there was a negative e-value, then the corresponding term in the spectral solution would go to ∞ which is not possible

- diffusion from random walk
- diffusion on graphs