Markov Chain of Greedy Algorithm for Gradient Sensing

Johannes Pausch

1 Aim

Use Markov Chain methods to analyse the performance of the greedy algorithm in gradient sensing situation.

Setup: Particles are emitted from a source (say the origin) and diffuse in space. A cell senses particles and tries to find the source. The greedy algorithm simply steps in the direction of the most recent arrival of a particle.

Link to Markov Chains: Let's regard the greedy algorithm as a random walk on a lattice where transition probabilities are biased towards the source. The bias should reflect the diffusive behaviour of the emitted particles.

Questions:

- Under what conditions / what biases exists an invariant distribution?
- What is the mean recurrence time for a greedy cell to return to the same position?
- What is the mean time to reach the source from any other position?
- If there is no invariant distribution, what is the chance to find the source?

2 Finding an invariant distribution

stationary does not imply detailed balance…

An invariant distribution $\pi P = \pi$ would fulfil the detailed balance equations $\pi_i p_{i,j} = \pi_j p_{j,i}$, which can be used to relate all elements of π to π_O, where O represents the origin.

Using the Manhattan metric, how many sites $x, y \in \mathbb{Z}^d$ would have equal distance $d_x = d_y$ to the origin and hence by symmetry equal invariant measures $\pi_y = \pi_x = a(d_x)\pi_O$?

The number of sites with distance d_x in the Manhattan metric can be found by considering compositions of natural numbers into d terms, where d is the dimension of the space. Here is why: a position vector x has Manhattan distance to the origin given by the sum of the absolute values of its entries

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix} \quad \Rightarrow \quad d_x = \sum_{\ell=1}^{d} |x_\ell| \tag{1}$$

If we're only interested in points that have $\forall \ell \in 1, \dots, d : x_\ell > 0$, then we need to count the compositions of d_x into d terms. Following the explanation on Wikipedia, we find this number by asking where to place pluses and where to place commas in a row of 1s of which there are d_x in total. The sums between commas are interpreted as the entries in the vector x. Hence $d - 1$ commas must be placed out of $d_x - 1$ possible positions.

$$x = (\underbrace{1 \overbrace{\square 1 \dots 1 \square}^{d-1 \text{ commas}} 1}_{d_x \text{ 1s}}) \tag{2}$$

Hence we find

$$\operatorname{card}\left\{x \in \mathbb{N}^d \,\middle|\, \sum_{\ell=1}^{d} x_\ell = d_x\right\} = \binom{d_x - 1}{d - 1} \tag{3}$$

However, we are interested in the number of such points in $\mathbb{Z}^d$. Before considering those that lie on an axis, we try to find out how many there are with $\forall \ell \in \{1, \dots, d\} : x_\ell \neq 0$. This can be found by noting that each coordinate of $x \in \mathbb{N}^d$ can be flipped to its negative pendant without changing d_x, resulting in a factor 2 for each dimension:

$$\operatorname{card}\left\{x \in \mathbb{Z}^d \,\middle|\, \sum_{\ell=1}^{d} |x_\ell| = d_x, \forall \ell : x_\ell \neq 0\right\} = 2^d \binom{d_x - 1}{d - 1} \tag{4}$$

Next we're interested in the number of sites that lie on an axis and have distance d_x. If the site lies on exactly one axis, then $d-1$ components of x equal zero. There are only $2d$ such sites. They are all of the form $x = \pm d_x e_i$ where e_i is a unit vector.

We can also take the opposite view and ask how many points are there that are distance d_x away from the origin and have exactly one component equal to zero. There are d choices for which exactly one vector component equals zero. For the remaining components, we simply need to count how many points there are whose distance is d_x and whose non-zero components are distributed over $d-1$ entries of x. The number of such sites equals

$$\operatorname{card}\left\{x \in \mathbb{Z}^d \,\middle|\, \sum_{\ell=1}^{d} |x_\ell| = d_x, \exists! j : x_j = 0, \forall \ell \neq j : x_\ell \neq 0\right\} = d 2^{d-1} \binom{d_x - 1}{d - 2} \tag{5}$$

If we extend this to exactly k components of x that have to equal zero, while x has distance d_x, we find

$$\operatorname{card}\left\{x \in \mathbb{Z}^d \,\middle|\, \sum_{\ell=1}^{d} |x_\ell| = d_x, \exists j_1 \neq \dots, \neq j_k : x_{j_i} = 0, \forall \ell \neq j_i : x_\ell \neq 0\right\} = \binom{d}{k} 2^{d-k} \binom{d_x - 1}{d - 1 - k} \tag{6}$$

For $k = d-1$ we recover the result from above for the number of sites that have distance d_x and lie on an axis. For $k = 0$ it coincides with the result above for number of sites with distance d_x and with no component equalling zero.

Thus, the total number of sites $x \in \mathbb{Z}^d$ that have distance d_x to the origin in the Manhattan metric equals

$$\begin{aligned}
\operatorname{card}\left\{x \in \mathbb{Z}^d \,\middle|\, \sum_{\ell=1}^{d} |x_\ell| = d_x\right\} &= \sum_{k=0}^{d-1} \binom{d}{k} 2^{d-k} \binom{d_x - 1}{d - 1 - k} &(7)\\
&= \begin{cases} 2 & d = 1 \\ 4d_x & d = 2 \\ 2 + 4d_x^2 & d = 3 \\ \frac{8}{3} d_x (2 + d_x^2) & d = 4 \\ \frac{2}{3}(3 + 10 d_x^2 + 2 d_x^4) & d = 5 \end{cases} &(8)
\end{aligned}$$

with the convention that those binomial coefficients equal zero which have negative entries or which have the top number strictly smaller than the bottom number.

As all sites are in one communicating class, if there is an invariant distribution π, then all sites $x \in \mathbb{Z}^d$ will have a non-zero contribution in that distribution, i.e. $\pi_x > 0$. Using detailed balance equations, all π_x can be expressed in terms of π_O. We assume a symmetric system that's only based on the Manhattan distance, i.e. the pre-factor only depends on the distance of x to the origin

$$\pi_x = f(d_x) \pi_O \tag{9}$$

Then, the invariant distribution has to be normalisable

$$1 = \sum_{x \in \mathbb{Z}^d} \pi_x = \pi_O \sum_{x \in \mathbb{Z}^d} f(d_x) = \pi_O \left(1 + \sum_{d_x=1}^{\infty} f(d_x) \sum_{k=0}^{d-1} \binom{d}{k} 2^{d-k} \binom{d_x - 1}{d-1-k}\right) \tag{10}$$

Hence we require that the sum in the large brackets is finite for the invariant distribution to exist. Therefore $f(d_x)$ has to have the following asymptotic behaviour with $\epsilon > 0$:

$$f(d_x) \sim \frac{1}{d_x^{d+\epsilon}} \tag{11}$$

This means that we have found now a condition for when the Markov Chain has an invariant distribution and hence a finite recurrence time. Is this condition met in a space where cue molecules are emitted by a single source? This is the topic of the next section.

3 Random Walk bias based on emittance of cue molecules which follow an SSRW

What is the steady state of a system in which particles are emitted at the origin in exponentially distributed time intervals and where particles can move diffusively? The action of a Doi-Peliti field theory of such a system is

$$\mathcal{A} = \int -\widetilde{\phi}(x,t)(\partial_t - D\Delta + r)\phi(x,t) + \delta(x)\gamma\widetilde{\phi}(x,t)\,\mathrm{d}t\mathrm{d}^d x, \tag{12}$$

where r is a regularising extinction rate. Hence, the steady state particle density can be calculated as follows:

$$\langle \phi(x,t) \rangle \hat{=} \longrightarrow\!\!\times \tag{13}$$

$$= \int \frac{\gamma e^{-ikx}}{Dk^2 + r} đ^d k \tag{14}$$

$$= \begin{cases} \frac{\gamma}{2\sqrt{Dr}} e^{-\sqrt{\frac{r}{D}}|x|} & d = 1 \\ \frac{\gamma}{2\pi D} K_0\left(\sqrt{\frac{r}{D}}|x|\right) & d = 2 \\ \frac{\gamma}{2\pi D|x|} e^{-\sqrt{\frac{r}{D}}|x|} & d = 3 \text{ and } = \frac{\gamma}{2\pi D|x|} \text{ for } r \to 0 \end{cases} \tag{15}$$

where $K_0(\bullet)$ is the modified Bessel function of the second kind. We observe that as $r \to 0$, the particle densities in $d = 1$ and $d = 2$ diverge, but for $d = 3$, a steady state exists.

In $d = 3$, we can conclude that the ratio of the average particle number at distance d_x over the average particle number at distance d_y equals d_y/d_x, which is independent of the diffusion constant D or spontaneous creation rate γ. Given a position x with distance d_x to the origin, there are on average $(d_x + h)/(d_x - h)$ times more particles one step of length h closer to the origin compared to one step further away from the origin. Therefore, we assume that the probability for the greedy algorithm to move

What do these results imply for the invariant distribution? In $d = 1$ and $d = 2$, the cue molecule density diverges. At finite extinction rates r, the particle distribution in $d = 1$ decays exponentially fast, which is faster than the required $\frac{1}{d_x^{(1+\epsilon)}}$. In $d = 2$, the asymptotic decay follows $\sim \frac{e^{-d_x}}{d_x}$ which is faster than $\frac{1}{d_x^{2+\epsilon}}$. Therefore invariant distributions exist for $r > 0$, which implies that the mean recurrence time is finite, which then implies that the random walker is expected to find the source within a finite time. In particular, this changes the behaviour of the random walker from being null recurrent when it's an SSRW to positive recurrent! If there is no extinction rate, i.e. $r = 0$, the entire space has a divergent cue particle density and we can't analyse any steady state behaviour, because it doesn't exist.

In $d = 3$, with a non-zero extinction probability, the invariant distribution exists, which implies that the random walk is recurrent and thus the expected first hitting time of the source is finite!

Furthermore, the limit $r \to \infty$ exists and results in a $\frac{1}{d_x}$ decay which is slower than the needed $\frac{1}{d_x^{3+\epsilon}}$. Hence, no invariant distribution exists, which means that the expected recurrence time is infinite, which then implies that the expected time to find the source would be infinite too.

However, there is a subtle but important distinction to be made: the expected recurrence time can be infinite while the recurrence probability can still equal 1. This is the case for the SSRW in 1 and 2 dimensions.