Isotropic independent random walks

Wenbin Lu

Abstract

This note shows that the probability distribution of the position of a random walker after $N(\gg 1)$ steps is an d-dimensional Gaussian, provided that the steps are uncorrelated and isotropic in d dimensions.

General Consideration

At each step, a walker moves by $\boldsymbol{r}$ which is randomly drawn from a probability distribution $p(\boldsymbol{r})$. This probability distribution is assumed to be isotropic and hence $p(\boldsymbol{r})$ only depends on the magnitude of $r = |\boldsymbol{r}|$, and it has been normalized such that $\int p(\boldsymbol{r})\,\mathrm{d}^d\boldsymbol{r} = 1$, where $\mathrm{d}^d\boldsymbol{r}$ is the d-dimensional volume element. The probability distribution $p(\boldsymbol{r})$ can be an arbitrary scalar function, and the results in this note apply as long as the mean-squared stepsize is well defined

$$\langle r^2 \rangle = \int r^2\, p(\boldsymbol{r})\,\mathrm{d}^d\boldsymbol{r}. \tag{1}$$

Suppose the probability distribution of the walker's position $\boldsymbol{R}$ after N steps is $P_N(R)$. Then, the following recurrence relation must hold

$$P_{N+1}(\boldsymbol{R}) = \int p(\boldsymbol{r}) P_N(\boldsymbol{R} - \boldsymbol{r})\,\mathrm{d}^d\boldsymbol{r}. \tag{2}$$

In the limit $N \gg 1$, we expect that $P_N(\boldsymbol{R})$ varies on lengthscales much longer than the typical step size $\boldsymbol{r}$, so it is reasonable to Taylor expand the $P_N(\boldsymbol{R} - \boldsymbol{r})$ term as follows

$$P_N(\boldsymbol{R} - \boldsymbol{r}) \approx P_N(\boldsymbol{R}) - \sum_i r_i \partial_{R_i} P_N(\boldsymbol{R}) + \frac{1}{2}\sum_{ij} r_i r_j \partial_{R_i}\partial_{R_j} P_N(\boldsymbol{R}) + \ldots, \tag{3}$$

where i and j goes over all dimensions. The linear term does not contribute to the integral on the RHS of eq. (2) because $p(\boldsymbol{r})$ is forward-backward symmetric along any axis $i = 1, 2, \ldots, d$. As for the quadratic term involving $\sum_{ij} r_i r_j$, we only need to consider the terms with $i = j$ because each step is isotropic (meaning that there is no correlation between i and j directions if $i \neq j$). Thus, we obtain $\int p(\boldsymbol{r}) \left(\sum_{ij} r_i r_j\right) \mathrm{d}^d\boldsymbol{r} = \int p(\boldsymbol{r}) \left(\sum_i r_i^2\right) \mathrm{d}^d\boldsymbol{r} = \langle r^2 \rangle$. We also know that $P_N(\boldsymbol{R})$ is isotropic along all axes $i = 1, 2, \ldots, d$, so we write $\partial^2_{R_i} P_N(\boldsymbol{R}) = d^{-1} \sum_i \partial^2_{R_i} P_N(\boldsymbol{R}) = d^{-1}\nabla^2 P_N(\boldsymbol{R})$. Therefore, the probability distribution at step $N+1$ is given by

$$P_{N+1}(\boldsymbol{R}) = P_N(\boldsymbol{R}) + \frac{\langle r^2 \rangle}{2d}\nabla^2 P_N(\boldsymbol{R}). \tag{4}$$

In the limit $N \gg 1$, we expect that $P_N(\boldsymbol{R})$ varies on timescales much longer than the typical time step Δt, so we can approximate $P_N(\boldsymbol{R})$ as a continuous, time-dependent probability distribution $\rho(\boldsymbol{R}, t)$, which satisfies the following diffusion equation

$$\partial_t \rho(\boldsymbol{R}, t) = D\nabla^2 \rho\,(\boldsymbol{R}, t)\,, \quad D = \frac{\langle r^2 \rangle}{2d\Delta t}. \tag{5}$$

Here D is called a time-independent diffusion coefficient.

For a given initial condition $\rho(\boldsymbol{R}, t=0) = \delta(\boldsymbol{R})$, the above diffusion equation can be solved by separation of variables. It can be shown that $\rho(\boldsymbol{R}, t)$ has the following d-dimensional Gaussian form

$$\rho(\boldsymbol{R}, t) \propto \exp\left[-\frac{R^2}{2\sigma^2(t)}\right], \quad \sigma^2 = 2Dt = \frac{\langle r^2 \rangle}{d\Delta t} t = \frac{\langle r^2 \rangle}{d} N, \tag{6}$$

and it is normalized such that $\int \rho \, \mathrm{d}^d \boldsymbol{R} = 1$. Here $\Delta t = t/N$ is the time step. We see that the characteristic radius of the probability density d-dimensional "cloud" grows as the square root of time t or the number of steps N, i.e., $\sigma \propto t^{1/2} \propto N^{1/2}$. Sometimes, we also want to know the PDF for the distance from the origin R and it is given by

$$\rho(R, t) = A_d R^{d-1} \rho(\boldsymbol{R}, t), \tag{7}$$

where $A_d = 2\pi^{d/2}/\Gamma(d/2)$ (where $\Gamma(x)$ is the Gamma function) is the surface area of the unit sphere in d dimensions ($A_1 = 2, A_2 = 2\pi, A_3 = 4\pi, A_4 = 2\pi^2, \ldots$).

Another result is that, if the current distance to the origin is much greater than the step size $R \gg \sqrt{\langle r^2 \rangle}$, then each step corresponds to a *typical* intensity variation of the order $\sim R\sqrt{r^2}$, which is much greater than the naïve expectation of $\sqrt{r^2}$. Let us consider going from $\boldsymbol{R}$ (known) to $\boldsymbol{R}' = \boldsymbol{R} + \boldsymbol{r}$, where $\boldsymbol{r}$ is a random step. The average intensity variation is

$$\langle R'^2 - R^2 \rangle = 2\boldsymbol{R} \cdot \langle \boldsymbol{r} \rangle + \langle r \rangle^2 = \langle r^2 \rangle, \tag{8}$$

which is not surprising. However, the variance of the intensity variation is

$$\sqrt{\langle (R'^2 - R^2)^2 \rangle} = \sqrt{4\langle (\boldsymbol{R} \cdot \boldsymbol{r})^2 \rangle + \langle r^4 \rangle} = \sqrt{2R^2 \langle r^2 \rangle + \langle r^4 \rangle} \approx \sqrt{2} R \sqrt{\langle r^2 \rangle}, \tag{9}$$

where we have dropped the linear term (as the average is zero) and made use of $\langle (\boldsymbol{R} \cdot \boldsymbol{r})^2 \rangle = R^2 \langle r^2 \rangle / 2$, and the final approximation is for $R \gg \sqrt{\langle r^2 \rangle}$. The above variance describes the typical intensity variation in a step.

Example — 2D

For $d = 2$, we obtain the *Rayleigh distribution*

$$P_N(\boldsymbol{R}) = \frac{1}{\pi N \langle r^2 \rangle} \exp\left(-\frac{R^2}{N\langle r^2 \rangle}\right), \quad P_N(R) = \frac{2R}{N\langle r^2 \rangle} \exp\left(-\frac{R^2}{N\langle r^2 \rangle}\right). \tag{10}$$

Let us further consider various moments of the PDF of the distance R,

$$\langle R^n \rangle = \int_0^\infty P_N(R) R^n \, \mathrm{d}R = \Gamma(n/2 + 1) \left(N\langle r^2 \rangle\right)^{n/2}, \tag{11}$$

The average distance from the origin is $\langle R \rangle = 2^{-1}\sqrt{\pi N \langle r^2 \rangle}$. The variance of the distance to the origin is

$$\sqrt{\langle (R - \langle R \rangle)^2 \rangle} = \sqrt{\langle R^2 \rangle - \langle R \rangle^2} = \sqrt{1 - \frac{\pi}{4}} \left(N\langle r^2 \rangle\right)^{1/2}, \tag{12}$$

which is of the same order as the average distance to the origin (meaning that the uncertainty in R is quite large). A nice property of the Rayleigh distribution is that the variance in the intensity (which is proportional to the square of the amplitude R) is equal to the average intensity, and this is because

$$\langle (R^2 - \langle R^2 \rangle)^2 \rangle = \langle R^4 \rangle - \langle R^2 \rangle^2 = \langle R^2 \rangle^2 = \left(N \langle r^2 \rangle \right)^2 . \tag{13}$$