Random Walks With Fixed Endpoints

Matthew Dutson

October 24, 2018

Motivation

Two of the major mathematical challenges in implementing the collagen fiber generator are:

1. Generating random lists of real numbers, potentially given a list of weights, which have a fixed weighted mean. The challenge is that these values (for instance, straightness) are restricted to a certain range. Normally, the approach would be to sample from a distribution with the desired mean and then add some Δ to each x_i, where $\Delta = \bar{x}_i - \mu$. However, this could potentially push values out of the allowed range.

2. Generating a discrete random path of fixed length between two points in 2D or 3D Euclidian space. Fibers are defined by their endpoints and the length of the path connecting them. This means that, for fibers with more than two segments, there are infinitely many ways to draw paths between the endpoints.

Both of these problems can be solved with a random walk between fixed points. Assuming the mean is not weighted, the first problem reduces to producing a random sequence $\{x_1, x_2, \dots x_n\}$ where

$$\sum_{i=1}^{n} x_i = n \cdot \mu$$

and the x_i values are from the range $x_{\min}$ and $x_{\max}$.

The second problem is similar. Assuming each step in the discrete path is of the same length l, the random choice at each step is an angle θ (or in the case of three dimensions, two angles θ and ϕ). The solution is a set of vectors $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ where $|\vec{v}_i| = l$ and

$$\sum_{i=1}^{n} \vec{v}_i = \vec{R}$$

$\vec{R}$ is the displacement between the starting and ending points.

Challenges

Here we focus on the second problem (2D and 3D). Let W be the set of all random walks of n steps. Let $S \subset W$ the subset of walks with displacement $\vec{R}$. Ideally, a solution

$\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ should sample evenly from the set S. In theory, this could be accomplished using the following method:

1. Let $\vec{p}_1$ and $\vec{p}_{n+1}$ be the starting and final points, respectively.

2. Assume the algorithm has progressed to some point $\vec{p}_i$.

3. The set of points directly reachable from $\vec{p}_i$ is a circle of radius l (or the higher-dimensional analog).

4. Let $P(n, \vec{R})$ be the differential probability of a random walk of n steps achieving a displacement of $\vec{R}$ (in reality, this probability should be symmetric in the direction of $\vec{R}$ and depend only on its magnitude $|\vec{R}|$).

5. Choose a point randomly from the circle/sphere of radius l, sampling from the probability distribution $P(n - i - 1, \vec{x} - \vec{p}_i)$, where $\vec{x}$ is on the circle of radius l. This is the probability that a random walk from $\vec{p}_n$ of $n - i - 1$ steps will arrive at the point $\vec{x}$ rather than some other point on the circle around $\vec{p}_i$.

6. Set $\vec{p}_{i+1}$ to the randomly chosen $\vec{x}$ and proceed to the next iteration.

The challenge in actually implementing the algorithm described above is that there is no general form for $P(n, \vec{R})$. For large n, the probability can be estimated using Stirling's approximation.

$$TODO : Inserthere$$

However, the problems with applying this approximation to our problem are obvious, especially when $|\vec{R}|$ is close to $n \cdot l$. The exponential in the equation above never reaches zero, which means that all steps are allowed, even if they lead to a configuration in which the end is no longer reachable in the number of remaining steps.

Unfortunately there is no known general expression for $P(n, \vec{R})$ (VERIFY THIS). The only way to estimate $P(n, \vec{R})$ is computationally, by generating a very large number of random walks of length $n = 1, 2, 3, \ldots$ and observing the distribution of their final displacements from the source. Such simulations were performed for up to $n = 8$

Workable Algorithms

Other Applications

Another application is the generation of random discrete functions $f(x)$ on an interval (x_1, x_2) such that the Riemann integral of f from x_1 to x_2 is a predetermined constant and the derivatives of f are subject to certain constraints. A simple example is generating a function with a minimum "smoothness", that is, a function whose slope is confined to some range $(-a, a)$.