2. Random walks and diffusion

Diffusion

Diffusion is the intermingling of one substance into another substance based on the movement of the particles. For example, when you put a sugar cube into a cup of coffee, the first thing that happens is that the sugar dissolves. Once it dissolves, the sugar particles begin to spread through the coffee. This is known as diffusion. Eventually, even the top layer is sweet (even though there may be an undissolved layer of sugary coffee at the bottom).

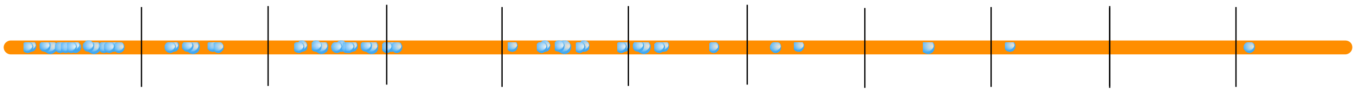
The diffusion process is ubiquitous throughout science: we may be interested in how heat diffuses in the handle of a frying pan, or how the hormones in birth control pills diffuse in the human body. To understand how diffusion works, we need a mathematical model.

Simplifying assumptions

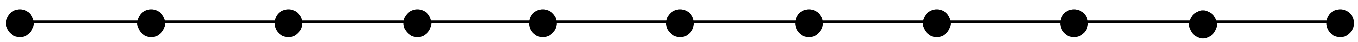
As a first step, we will consider a one dimensional model in which one substance (say sugar) diffuses along an interval of the real line (perhaps a very thin test tube of coffee).

We will model the sugar as discrete particles or molecules.

In addition, we will break up the interval into distinct subintervals, and look at how many particles move from one subinterval into the adjacent subintervals.



Once we make this restriction, we can model the real line as a series of nodes. At each moment in time, the number of particles at each node is determined by the number of particles on the corresponding subinterval.



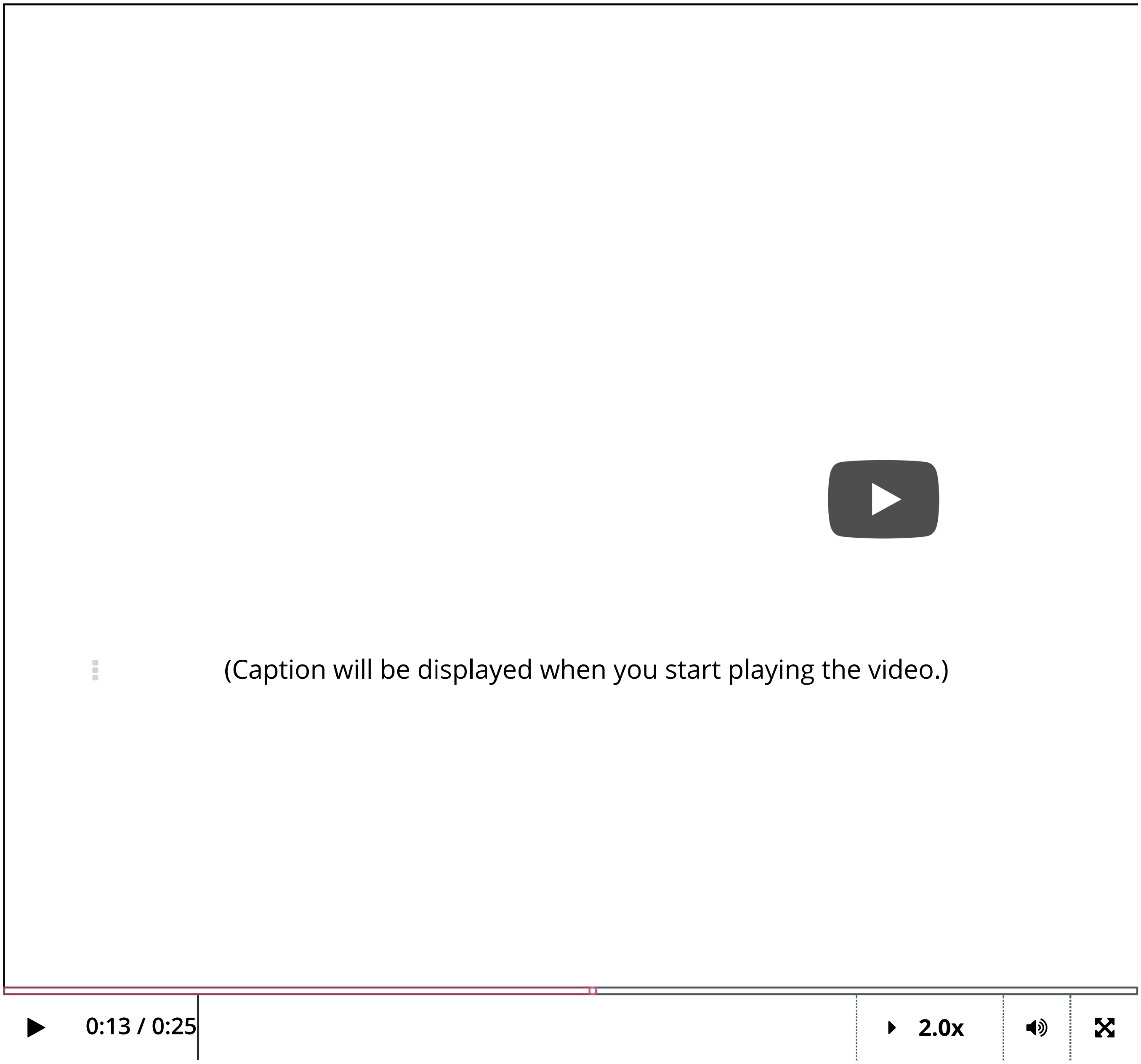
Random walks

The way we model the diffusion of sugar along this series of nodes is to consider the sugar particles as random walkers. A **random walk** is a path taken by the random walker, starting at one node, and moving to one or the other nodes at each step. The sense in which the path is random is that we consider more than one possible path, and assign a probability to each path.

For example, for each time step Δt, the random walker has an equal probability of moving one node to the right, or one node to the left. No walker stays stationary. Note that a random walker can only move into adjacent nodes, they cannot teleport! Also, if a random walker is all the way at the endpoint on the left, they can only move to the right, and vice versa.

The following animation shows a histogram of the distribution of particles at each node along a graph. The positions of these particles are shown over several consecutive time steps.

Histogram of random walk over several time steps



Video

[Download video file]

Transcripts

Download SubRip (.srt) file

Download Text (.txt) file

Let's consider a situation where our interval is broken up into 5 nodes.

(1)—(2)—(3)—(4)—(5)

At time 0, we dump 100 random walkers into node 3. Each walker has an equal probability of moving to node 2 or node 4. So after one time step, the expect value (average value) of the number of walkers is 50 walkers at node 2 and 50 at node 4. In the next step, some of the walkers at each node move right, and some move left. This gives an expected value of 25 walkers at nodes 1 and 5, and 50 walkers at node 3. How can we use a matrix to describe the evolution of the expected value of the number of walkers for any initial distribution of walkers among the nodes?

We describe the evolution of the expected value over one time step as a 5×5 matrix such that in row j, column i is the probability that a random walker from node i moves to node j. That is this value is $1/2$ if $j = i - 1$ or $j = i + 1$, and is 0 otherwise. This matrix is given by

$$\mathbf{M} = \begin{pmatrix} 0 & 1/2 & 0 & 0 & 0 \\ 1 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1 \\ 0 & 0 & 0 & 1/2 & 0 \end{pmatrix}.$$

The expected distribution after one time step is given by the matrix product

$$\begin{pmatrix} 0 & 1/2 & 0 & 0 & 0 \\ 1 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1 \\ 0 & 0 & 0 & 1/2 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 100 \\ 0 \\ 0 \end{pmatrix}.$$

After 2 time steps the distribution is

$$\begin{pmatrix} 0 & 1/2 & 0 & 0 & 0 \\ 1 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1 \\ 0 & 0 & 0 & 1/2 & 0 \end{pmatrix}^2 \begin{pmatrix} 0 \\ 0 \\ 100 \\ 0 \\ 0 \end{pmatrix}.$$

In other words, after k time steps, the expected distribution of random walkers is

$$\mathbf{c}_k = \mathbf{M}\mathbf{c}_{k-1}, \qquad \mathbf{c}_{k-1} = \mathbf{M}^{k-1}\mathbf{c}_0,$$

where $\mathbf{c}_0$ is the initial distribution of walkers along the nodes.

Random walk setup (External resource) (1.0 / 1.0 points)

Random walkers set up

Suppose we dump 1000 random walkers into node 1 of a 10 node one dimensional graph at $t = 0$. At each time step Δt, a random walker has an equal probability of moving one step to the right, one step to the left, or staying stationary. **(At the end points it has a 2/3 probability of staying stationary and a 1/3 probability moving to the adjacent node.)** Because we allow the random walker to stay stationary, we call this a lazy random walker. Random walkers cannot move beyond adjacent nodes.

Create a matrix $\mathbf{A}$ whose ith column lists probabilities $a_{j,i}$ for a random walker in node i to get to node j over one time step. (Such a matrix is called a Markov matrix. A **Markov matrix** is a square matrix with nonnegative entries such that the sum of the entries in each column is 1. This is equivalent to the property that no mater what the starting distribution, the expected number of particles is conserved, even though these expectations are fractions of particles.)

The number of random walkers at node i at time $t = 0$ is c_i. Store the initial value of walkers in a column vector

$$\mathbf{c}0 = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_{10} \end{bmatrix}.$$

Your Script

Save Reset MATLAB Documentation (https://www.mathworks.com/help/)

```
%Enter the matrix A. (Hint it may be useful to look up the
%matlab documentation on diagonal and tridiagonal matrices.)
n = 10;
A = full(gallery('tridiag',n,1/3,1/3,1/3));
A(1,1) = 2/3;
A(n,n) = 2/3;
%A(1:2,1) = 0.5;
%A(n-1:n,n) = 0.5;
A

%Enter column vector of the initial state in a vector called c0.
c0 = zeros(n,1);
c0(1) = 1000;
%c0

%Store the expected distribution of walkers after one time step in a vector called x
x1 = A * c0;

%Store the expected distribution of walkers after two time steps in a vector called
x2 = A^2 * c0;

```

```
%Store the expected distribution of walkers after 9 time steps in a vector called x9
x9 = A^9 * c0;
```

▶ Run Script ❓ ()

Assessment: Incorrect

Submit ❓ ()

A is correct

Eigenvalues

3/3 points (graded)

We are interested in finding the steady state behavior of this system. To do so, we need to find $\mathbf{A}^n\mathbf{c}$ for large n. To study powers of a matrix, the easiest thing to do is to study its diagonalization.

Use the command **eig()** in MATLAB online to find the diagonal matrix of eigenvalues $\mathbf{D}$ and matrix of eigenvectors $\mathbf{S}$ so that $\mathbf{A} = \mathbf{S}\mathbf{D}\mathbf{S}^{-1}$. (You can copy and paste your matrix from the previous problem.)

What is the largest eigenvalue?

1 ✔

What do you notice about the other eigenvalues (for this particular matrix)? (Choose all that apply.)

- [ ] They are all positive.
- [ ] They are all negative
- [x] They are all real.
- [x] They all have absolute value less than 1.

✔

What is $\mathbf{D}^n$ as $n \to \infty$?

- The identity matrix.
- A matrix with 1's and -1's on the diagonal.
- A matrix with one 1 and one -1 on the diagonal, and the rest zeros.
- A matrix with one -1 on the diagonal, and the rest zeros.
- A matrix with one 1 on the diagonal, and the rest zeros. ✔
- A matrix of zeros.

Submit You have used 1 of 2 attempts

Steady state behavior

2/2 points (graded)

We are interested in the steady state when we start with 1000 random walkers in node j.

Find the expected distribution after infinite time.

(Enter as a column vector in matlab notation: e.g. [1; 0; 0 ; 0; 0].)

[100;100;100;100;100;10 ✔ **Answer:** [100; 100; 100; 100; 100; 100; 100; 100; 100; 100]

What if the distribution is 500 random walkers at node 1 and 500 at node 3. What is the steady state distribution?

(Enter as a column vector in matlab notation: e.g. [1; 0; 0 ; 0; 0].)

[100;100;100;100;100;10 ✔ **Answer:** [100; 100; 100; 100; 100; 100; 100; 100; 100; 100]

Solution:

We know that after n time steps, the expected distribution of random walkers along our 1-dimensional graph is given by $\mathbf{A}^n \mathbf{c}_0$, where $\mathbf{c}_0$ is the vector of initial values.

Because, $\mathbf{A}^n = \mathbf{S}\mathbf{D}^n\mathbf{S}^{-1}$, and as n tends to infinity $\mathbf{D}^n$ is a single one and all other zeros, we can calculate this explicitly using MATLAB.

First create a matrix $\mathbf{M}$ which is the matrix $\mathbf{D}^n$ as n tends to infinity.

```
M = zeros(10);
M(1,1) = 1;
```

Note that you must place the 1 in whichever column corresponds to the eigenvalue that is equal to 1. A quick MATLAB computation shows that $\mathbf{SMS}^{-1}\mathbf{c0}$ is the column vector whose entries are all 100.

Note that for any initial condition vector $\mathbf{c}_0$ whose sum of total walkers in each entry is 1000, the steady state vector is also the vector of all 100's.This is exactly what we expect from something that is modeling diffusion. At steady state, we expect the concentration of random walkers to be uniform at each node. Let's see why this is true.

Let $\mathbf{v}_1, \ldots, \mathbf{v}_{10}$ be the eigenvectors of $\mathbf{A}$. (You can check that you do in fact get 10 linearly independent, real eigenvectors for this matrix.) And let

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

be the column vector of all ones, which corresponds to the largest eigenvalue, which 1.

This matrix has distinct eigenvalues, hence is complete. Thus given any initial distribution $\mathbf{c}_0$, we can write this vector as a linear combination of the eigenvectors

$$\mathbf{c}_0 = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_{10} \mathbf{v}_{10}.$$

Therefore

$$\mathbf{A}^k\mathbf{c}_0 = \alpha_1\mathbf{v}_1 + \alpha_2\lambda_2^k\mathbf{v}_2 + \cdots + \alpha_{10}\lambda_{10}^k\mathbf{v}_{10} \longrightarrow \alpha_1\mathbf{v}_1, \quad \text{as } k \to \infty.$$

Note that $\mathbf{v}_1^T = \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix}$ is a row vector of all ones. And

$$\mathbf{v}_1^T\mathbf{c}_0 = 1000$$

is the total number of particles.

Because $\mathbf{A}$ is a Markov matrix, the condition that the entries in the columns sum to 1 is equivalent to the condition

$$\mathbf{v}_1^T\mathbf{A} = \mathbf{v}_1^T.$$

In particular, this tells us that

$$\mathbf{v}_1^T\mathbf{A}\mathbf{c}_0 = \mathbf{v}_1^T\mathbf{c}_0 = 1000,$$

and in particular after any number of time steps

$$\mathbf{v}_1^T\mathbf{A}^k\mathbf{c}_0 = \mathbf{v}_1^T\mathbf{c}_0 = 1000.$$

But we also have

$$\begin{aligned}\mathbf{v}_1^T(\mathbf{A}^k\mathbf{c}_0) &= \mathbf{v}_1^T\left(\alpha_1\mathbf{v}_1 + \alpha_2\lambda_2^k\mathbf{v}_2 + \cdots + \alpha_{10}\lambda_{10}^k\mathbf{v}_{10}\right) \\ &\longrightarrow \mathbf{v}_1^T\alpha_1\mathbf{v}_1 = \alpha_1\mathbf{v}_1^T\mathbf{v}_1 = \alpha_1 10\end{aligned}$$

This tells us that $1000 = \alpha_1 10$ or $\alpha_1 = 100$ no matter what the initial distribution of random walkers is, as long as there are 1000 total random walkers at the beginning.

Submit

You have used 2 of 3 attempts

Answers are displayed within the problem

2. Random walks and diffusion

Hide Discussion

Topic: Unit 2: Linear Algebra, Part 2 / 2. Random walks and diffusion

Add a Post

All Posts

Statement: PROBLEM IS BEING UPDATED - Try the problem again later

question posted a day ago by **tanopag**

Is there something still to fix about problem 2. Random walks and diffusion ?

This post is visible to everyone.

Add a Response

2 responses

jfrench (Staff)
about 20 hours ago

My fault. Should be working now!

Does the grader work fine with the last question "Steady state behavior" ? How many decimals are expected ? I am pretty sure of my answers, I have entered them with 4 decimals (same rounding than MabLab) yet the grader rejects them.

posted about 17 hours ago by **yves-M** (Community TA)

i tried as an exact fraction in distribution terms (ie they sum to 1) and as well as relative to 1000 in rounded form. neither was accepted. And it doesn't allow me to multiply the distribution by 1000 for my 3rd attempt. the syntax for a 10 entry column vector to be entered in this format is a bit awkward.

posted about 14 hours ago by **amenon1** hmm, well i got it wrong i guess. i have to say the limit of the matrix as n goes to infinity from matlab is different than the solution. I will rework through the details of the grader answer but the distribution i ended up with is more intuitive than the supposed answer. i recall from boltzmann distributions of confined molecules endpoints are not uniform.

posted about 13 hours ago by **amenon1**

@yves-M I am also kind of sure that my answer is correct, but marked wrong by the grader. --- edited BB--- Can the [Staff] please verify?

posted about 13 hours ago by **sandipan_dey**

that is also what i got and what i continue to think is correct but that is not the accepted answer.

posted about 10 hours ago by **amenon1**

Dreaded red crosses here too. Seems fairly easy to solve analytically, but indeed the grader doesn't like my answers entered in the format `[a/b;c/d;...]`.

posted about 9 hours ago by **mrBB** (Community TA)

@sandipan_dey That was perhaps a bit too much info on what you feel the answer should look like. Please give others the opportunity to work on the problem without encountering too many spoilers.

posted about 9 hours ago by **mrBB** (Community TA)

@mrBB sorry if i revealed to much, but i still did not say the actual numbers :). The grader did not accept the decimal format either.

posted about 9 hours ago by **sandipan_dey**

No you didn't, but you made the solution space a few dimensions smaller. :-) No worries though.

I trust the course staff will look into this and let us know if there indeed is an issue. Perhaps we should have a bit of patience as it's still early in Boston.

posted about 8 hours ago by **mrBB** (Community TA)

I am also marked as incorrect. I would also like to know if there is no issue so that I can use up my tries and see the answer. Thanks.

posted about 2 hours ago by **byxrolland**

Add a comment

jfrench (Staff)
22 minutes ago

Ok. Sorry everyone. The issue is that the "problem broken" must have had some error, because it did not give everyone the most updated version of the problem. So in fact, I am sure your answers are correct, but none of my solutions following will make sense without the corrected problem statement. I am not sure what happened. But I can tell you that I did edit this problem, and somehow it wasn't saved properly.

Add a comment

Showing all responses

Add a response:

Preview

Submit

Learn About Verified Certificates