THE RANDOM WALK INNER PRODUCTS: A FORWARD AND INVERSE PROBLEM

ERIC NITARDY

Abstract. In this paper, we develop a means to calculating expected number of steps in a random walk from boundary node to boundary node in a graph or resistance network with boundary nodes using elementary probability arguments. We express this expected random walk length in terms of a pair of inner products that are calculated using solutions to the Dirichlet problem on the graph or network. We, then, consider the expected number of steps in a random walk from boundary node to boundary node as a means for making an external measurement of a graph or network from which we can infer its properties and use the inner products as tools to analyze this inverse problem.

1. Introduction

[...]

In Megan McCormick's REU paper of 2005, *Metric Recoverability* [1], she considered the problem of recovering distances between interior vertices of a metric graph with a boundary if given the lengths of the shortest path between boundary nodes. One of the problems that made this difficult was this metric's tendency to provide information only on a few short paths through the graph leaving most of the graph unrecoverable. The motivation for this paper was the desire to develop an alternative metric which avoided this difficulty. This effort instead leads to a pair of inner products upon which we can build an inverse problem.

2. The Model

We begin with the following toy model: Imagine a set of boundary nodes and interior nodes connected in a network with resistors of varying value. Imagine further that there is no voltage across any of the nodes and we drop a conducting particle on one of the boundary nodes and let it diffuse into the network. The conducting particle will scatter off the atoms of the resistors and random walk through the network. The number off scattering atoms in a resistor should be proportional to the resistance so the higher the resistance the more time the particle will require random walk though a resistor. Eventually, the particle will random walk out of the network at some boundary node – possibly the one it started at. If we note the node it emerged at and the time it took, after repeating the experiment a large number of times, we could measure the the expected time the particle requires to get from one boundary node to another. Since we could build this resistance network from wire of constant resistance per unit, using different wire lengths for different resistor values, this expected time becomes a rough measure of the distance, in wire length, between to different boundary nodes. Unlike the

Date: August 9, 2008.

standard shortest path metric this new measure provides information about all portions of the network as it is an average over all possible paths from one boundary node to another.

This physical model has a mathematical counterpart involving graphs. Given a graph with vertices and undirected edges (V, E) where the vertices are partitioned into interior vertices I and boundary vertices ∂V. We may think of a resistor connecting two nodes in our physical network as a line of $n-1$ vertices and n edges connecting the two nodes where n is proportional to the resistance value of the resistor. In this way, we can convert, a resistance network into a graph where each edge represents a resistor with some small unit resistance.

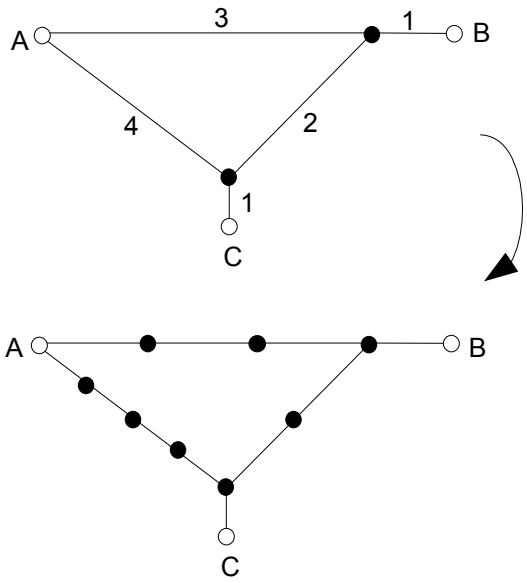

FIGURE 1. An example of converting a network with resistors of value $1, 2, 3, 4$ into one with unit resistors.

!!! A particle can then be fed into such a graph at one of the boundary vertices and made to step to the nearest interior vertex. From there it is allowed to random walk from vertex to vertex assuming it can step from the vertex it is on to each neighboring vertex with equal probability. In other words, at each step the particle can cross each available edge with equal probability. The random walk ends when the particle reaches a boundary vertex. Our proposed measurement is the expected number of steps from any two boundary nodes. Note that this is not a metric as the expected number of steps from a given boundary node back to itself is not zero.

At this point, we note that this is not the only way to model a resistance network with a random walk. The more typical approach is to build a random walk from node to node in the network without the intervening vertices to represent the resistance. Instead of giving an equal probability of crossing each available edge, we

would weight the probability according to the resistance: the probability of crossing from vertex i to vertex j, where j is among i's immediate neighbors $N(i)$, is

$$P_{ij} = \frac{1/r_{ij}}{\sum_{k\in N(i)} 1/r_{ik}}, \tag{1}$$

where $r_{ij}[= r_{ji}]$ is the resistance of the edge connecting vertex i and$j$.

Weighted in this fashion, the lower the resistance of an edge compared with the other edges available at a vertex the more likely it will be crossed. Since the resistances are encoded into the random walk according to their relative size, some information is lost in this approach, but it is a reasonable approach which may also lead to a good inverse problem. It is not the model used in this paper. Interestingly, however, most of the main results in this paper also apply when this model is used. We will state and prove results assuming the weighting of equation (1) wherever possible. Just keep in mind that the model we will have in mind imagines each edge have the same small unit resistance, and, consequently, from equation (1), the probability of crossing any edge connected to i is

$$P_{ij} = \frac{1}{\sum_{k\in N(i)} 1} = \frac{1}{d_i}, \tag{2}$$

where d_i is the degree of vertex i.

3. Some Definitions

For a undirected graph $G = (V, E)$ with boundary vertices $\{A, B, C \ldots\}$, let the space out outcomes be all possible paths on the graph from any vertex i_1 to a boundary vertex i_k (possibly the starting vertex) through interior vertices. So a path may begin and must end at a boundary node, but must pass through only interior nodes in between.

Definition 3.1. *The probability that a given path π from vertex i_1 to i_k following the series of neighboring vertices $i_1, i_2, i_3, \ldots, i_{k-1}, i_k$ is, from equation (1),*

$$prob(\pi) = P_{i_1 i_2} P_{i_2 i_3} P_{i_3 i_4} \cdots P_{i_{k-1} i_k}.$$

This reduces in our unit resistance model to

$$prob(\pi) = \frac{1}{d_{i_1}} \frac{1}{d_{i_2}} \frac{1}{d_{i_3}} \cdots \frac{1}{d_{i_{k-1}}}, \text{ again, where } d_i \text{ is the degree of vertex } i.$$

Definition 3.2. *Let p_{iA} be the probability that a random walk that starts at i arrives at the boundary first at node A. For an interior node i and boundary nodes, A and B, $A \neq B$, this can be calculated as*

$$p_{iA} = \sum_{\substack{all\ paths\ \pi \\ starting\ at\ i}} prob(\pi), \qquad p_{AA} = 1 \text{ and } p_{AB} = 0. \tag{3}$$

Let X_i be the random variable that represents the number of steps in a random walk path π leaving vertex i and arriving eventually at any boundary node. Let X_{iA} be the random variable that represents the number of steps in a random walk path π leaving vertex i and arriving at a given boundary node A, for any other path starting at i but not ending at boundary node A is zero.

Definition 3.3. *Let $l(\pi)$ be the length of a given path π. The expected value of X_i is*

$$E[X_i] = \sum_{\substack{\text{all paths } \pi \\ \text{starting at } i}} l(\pi) prob(\pi), \tag{4}$$

and the expected value of X_{iA} is

$$E[X_{iA}] = \sum_{\substack{\text{all paths } \pi \\ \text{starting at } i \\ \text{ending at } A}} l(\pi) prob(\pi). \tag{5}$$

Since all random walks beginning at i must eventually end at some boundary node $E[X_i]$ measures the expected number of steps in a random walk π leaving vertex i and arriving at the boundary. However, since $E[X_{iA}]$ measures number of steps for some paths, but ignores other paths, we must be careful so that we understand what it represents. Let's say we start a random walk at vertex i and wait until it goes out at some boundary node and note the number of steps it took to reach that node. If we repeat this a number of times (n times), each time writing down the number of steps it required in a table under the appropriate boundary node, and when we are done we total the number of step under each boundary node and divide by n, then as n gets large the number calculated under each boundary node $\{A, B, C \ldots\}$ approaches $E[X_{iA}], E[X_{iB}], E[X_{iC}], \ldots$.

Notice that $E[X_{iA}]$ is not the expected number of steps in a random walk given that it exits at A – that would be correctly calculated as a conditional expectation $E[X_{iA}]/p_{iA}$. Our initial intuition might suggest that this conditional expectation is the like measure that we are looking for, and, indeed, it may also form the basis of a good inverse problem. Instead, however, we will be using $E[X_{AB}]$ for any two boundary nodes A and B (possibly the same) as the measurements from which we will infer properties of the graph and base our inverse problem because, as the last paragraph indicates, it is a natural quantity to measure and because, as we shall see, it has such interesting properties.

4. Properties of p_{iA}

For $A \in \partial V$ a boundary node, and $i \in I$ an interior vertex on a general random walk graphs described by equation (1) – keep in mind our model is the special case where $r_{ij} = 1$ for all neighboring vertices i, j, and d_i be the degree of vertex i, let p_{iA} be the probability of a random walk beginning at vertex i and arriving at the boundary first at boundary node A as defined in equation (3). Let $u_A(i)$ be the solution to the Dirichlet problem on our graph with boundary values: one at boundary node A and zero at all other boundary nodes. More precisely, $u_A(i)$ is the unique function on the vertices of the graph, $u_A : V \to \mathbb{R}$, which has the three properties:

For each interior vertex, i, where $N(i)$ is the set of immediate neighbors of i,

$$u_A(i) = \left(\sum_{j \in N(i)} 1/r_{ij} \right)^{-1} \left(\sum_{j \in N(i)} (1/r_{ij}) u_A(j) \right)$$

$$\left[\begin{array}{c} \text{In the case of our model:} \\ u_A(i) = (1/d_i) \left(\sum_{j \in N(i)} u_A(j) \right) \end{array} \right] \tag{6}$$

and

$$u_A(A) = 1$$

while, for any other boundary node B,

$$u_A(B) = 0$$

Theorem 4.1. *Given the above definitions of p_{iA} and $u_A(i)$,*

$$p_{iA} = u_A(i)$$

Proof. For any interior vertex, i, if vertex $j \in N_i$, then, by equation (1),

$$P_{ij} = \frac{1/r_{ij}}{\sum_{k \in N(i)} 1/r_{ik}}.$$

P_{ij} is the probability of transition from i to one of its neighbors j. A random walk can only go from i to one of its neighbors, so since probability is conserved,

$$p_{iA} = \sum_{j \in N(i)} P_{ij} p_{jA} = \sum_{j \in N(i)} \left[\frac{(1/r_{ij}) p_{jA}}{\sum_{k \in N(i)} 1/r_{ik}} \right].$$

So, p_{iA} satisfies the first of the above three properties. The probability of getting to boundary node A first, starting at A, is one, and the probability of getting to A first, starting at another boundary node, B, is zero, i.e. $p_{AA} = 1$ and $p_{BA} = 0$. Thus, p_{iA} satisfies all the properties that $u_A(i)$ does, and since $u_A(i)$ is unique, they must be the same. □

Remark 4.2. If we compare the paths through the graph that are involved in calculating p_{iA} and $E[X_{iA}]$, when i is an interior vertex, the set of paths involved in the calculation are the same. When i is a boundary node, in calculating p_{iA}, the path never leaves the boundary, but, in calculating $E[X_{iA}]$, we assume they do. Thus, under the assumptions we make about paths for $E[X_{iA}]$, the probability of getting from interior vertex i to boundary node A is p_{iA}, but the probability of getting from boundary node B to boundary node A (even when $A = B$) is

$$\sum_{j \in N(B)} P_{Bj} p_{jA} = \sum_{j \in N(B)} \left[\frac{(1/r_{Bj}) p_{jA}}{\sum_{k \in N(B)} 1/r_{Bk}} \right].$$

And under the model that we will look at more closely, where the resistances are all one, this becomes

$$\left(\frac{1}{d_B} \right) \sum_{j \in N(B)} p_{jA} \text{ where } d_B \text{ is the degree of boundary node } B.$$

Remark 4.3. When we compare our preferred model of a random walk on a resistance network, where we divide up each resistor into strings of unit resistors with the alternative approach in which we weight the transition probabilities using equation (1), if we treat them only as resistance networks – ignoring the random walk aspects for the moment – they are identical. Consequently, their response to a unit voltage at boundary node A and zero voltage at all other nodes will, on the shared the nodes of the original network, be identical, that is, the solutions to the Dirichlet problem, $u_A(i)$, for node i in the original network, will be the same in both models. Hence, again, for original nodes i, p_{iA}, the probability that a random walk that starts at i arrives at the boundary first at node A, is the same in both models. We should expect, however, that $E[X_{iA}]$ for the two models to be quite different.

5. Properties of $E[X_i]$ and $E[X_{iA}]$

5.1. **The Leaving and Crossing Random Variables.** In order to calculate $E[X_i]$ and $E[X_{iA}]$, we need to define some additional random variables as tools.

Let $L_A(j)$ represent the number of times that random walk path π that starts at (and leaves) boundary node A leaves vertex j. Let $\lambda_A(j)$ be the expected value of $L_A(j)$, namely,

Definition 5.1. *Let $l(\pi)$ be the length of random walk path π,*

$$\lambda_A(j) = \sum_{\substack{\text{all paths } \pi \\ \text{starting at } A}} l(\pi) prob(\pi),$$

Note that, for any random walk path, $L_A(A) = 1$ and $L_A(B) = 0$, for boundary node B, $A \neq B$. So, $\lambda_A(A) = 1$ and $\lambda_A(B) = 0$

Similarly, for two boundary nodes A and B not necessarily different, let $L_{AB}(j)$ be the number of times that path starting at and leaving boundary node A leaves vertex j provided the path ends at B. $L_{AB}(j) = 0$ if the path does not end at B. $\lambda_{AB}(j)$ will be the expected value of $L_{AB}(j)$.

Definition 5.2. *Let $l(\pi)$ be the length of random walk path π.*

$$\lambda_{AB}(j) = \sum_{\substack{\text{all paths } \pi \\ \text{starting at } A \\ \text{ending at } B}} l(\pi) prob(\pi),$$

$\lambda_A(j)$ is the expected number of times a random walk beginning at A leaves vertex j, and $\lambda_{AB}(j)$ is the expected number of times a random walk beginning at A leaves vertex j, counting only those random walks that end at boundary node B. Notice that

$$\lambda_A(j) = \sum_{\text{all boundary nodes } B} \lambda_{AB}(j)$$

Moreover, since each step in a path involves leaving a vertex once,

$$E[X_A] = \sum_{\text{all vertices } j} \lambda_A(j)$$

and

$$E[X_{AB}] = \sum_{\text{all vertices } j} \lambda_{AB}(j).$$

We can do much the same thing with edge crossings. Let $C_{AB}(i,j)$ represent the number of times path π crosses edge $\{i,j\} \in E$ from i to j on a path that begins at A and ends at B. $C_{AB}(i,j)$ is zero on any other path. Let $\kappa_{AB}(i,j)$ be the the expected value of $C_{AB}(i,j)$. We observe that $\kappa_{AB}(i,j) + \kappa_{AB}(j,i)$ is the expected number times edge $\{i,j\}$ is crossed in a random walk from A to B. Thus, similarly to what happened above, since each step in a random walk involves crossing an edge once,

$$E[X_{AB}] = \sum_{\text{all edges } \{i,j\}} [\kappa_{AB}(i,j) + \kappa_{AB}(j,i)].$$

5.2. **A calculation of $\lambda_A(i)$.** If we consider a random walk starting boundary node A finishing at the next boundary node it visits, we know that the visits to an interior vertex i must be the result of leaving the neighboring vertices. So the expected number of leavings of interior vertex i is the sum of contributions from neighboring vertices. A random walk which is about to leave vertex $j \in N(i)$ has a probability P_{ji} visiting vertex i next. So, if $\lambda_A(j)$ is the expected number of leavings from vertex j, then j will contribute $P_{ji}\lambda_A(j)$ to the expected number of leavings from i. Consequently, for interior vertex i,

$$\lambda_A(i) = \sum_{j \in N(i)} P_{ji}\lambda_A(j) = \sum_{j \in N(i)} \frac{1/r_{ji}}{\sum_{k \in N(j)} 1/r_{jk}} \lambda_A(j). \tag{7}$$

An interior vertex in a random walk beginning at boundary node A, will only be visited from A or another interior node. So,

If i is an interior vertex that is not a neighbor of A,

$$\lambda_A(i) = \sum_{\substack{j \in N(i) \\ j \text{ interior vertex}}} P_{ji}\lambda_A(j). \tag{8}$$

If i is a neighbor of A, then $\lambda_A(i)$ gets an additional contributions from A:

$$\frac{1/r_{Ai}}{\sum_{k \in N(A)} 1/r_{Ak}} \lambda_A(A) = \frac{1/r_{Ai}}{\sum_{k \in N(A)} 1/r_{Ak}} = P_{Ai} \qquad (\text{as } \lambda_A(A) = 1)$$

or

$$\lambda_A(i) - P_{Ai} = \sum_{\substack{j \in N(i) \\ j \text{ interior vertex}}} P_{ji}\lambda_A(j). \tag{9}$$

If we index the interior vertices, putting those that neighbor A, first, P_{ij} forms a matrix and equations (7) and (8) become

$$[\lambda_A(i_1), \cdots, \lambda_A(i_k), \lambda_A(i_{k+1}), \cdots, \lambda_A(i_m)]P =$$
$$\left[\lambda_A(i_1) - P_{Ai_1}, \cdots, \lambda_A(i_k) - P_{Ai_k}, \lambda_A(i_{k+1}), \cdots, \lambda_A(i_m)\right]$$

or

$$[\lambda_A(i_1), \cdots, \lambda_A(i_k), \lambda_A(i_{k+1}), \cdots, \lambda_A(i_m)](I - P) = \tag{10}$$
$$\left[P_{Ai_1}, \cdots, P_{Ai_k}, 0, \cdots, 0\right]$$

With these considerations in mind we can show:

Lemma 5.3. *For vertex i and boundary node A in a resistance network, let $u_A(i)$ be the solution to the Dirichlet problem on the resistance network setting $u_A(A) = 1$ and $u_A(B) = 0$ at any other boundary node B, and let $\sigma_i = \sum_{k\in N(i)} 1/r_{ik}$. If $\lambda_A(i)$ is the expected number of leavings from vertex i on a random walk, governed by equation (1), from boundary node A to any node on the boundary, then*

$$\lambda_A(i) = \left(\frac{1}{\sigma_A}\right)\sigma_i u_A(i).$$

Note that for the model with unit resistances this formula becomes

$$\lambda_A(i) = \left(\frac{1}{d_A}\right) d_i u_A(i), \text{ where } d_j \text{ is the degree of vertex } j.$$

Proof. The j^{th} row of matrix P represents the probabilities that a random walk at vertex j will make its next step to any of the other interior vertices, so the sum of any row is less than or equal to one. If the vertex j is next to a boundary node, the sum of the j^{th} row is less than one, since the probability of leaving the interior from vertex j is positive. This implies that the matrix $I - P$ is diagonal dominant, and, thus, must have an inverse, and so, equation (10) has a unique solution. Let's show the solution is $\lambda_A(i) = (1/\sigma_A)\sigma_i u_A(i)$ for each interior vertex i.

Observe that since $u_A(i)$ is harmonic at any interior vertex i,

$$u_A(i) = \sigma_i^{-1}\left(\sum_{j\in N(i)} (1/r_{ij})u_A(j)\right)$$

or

$$\left(\frac{1}{\sigma_A}\right)\sigma_i u_A(i) = \left(\frac{1}{\sigma_A}\right)\left(\sum_{j\in N(i)} (1/r_{ij})u_A(j)\right)$$

$$= \sum_{j\in N(i)} \frac{(1/r_{ij})\left(\frac{1}{\sigma_A}\right)\sigma_j u_A(j)}{\sigma_j}$$

This is a restatement of equation (7) for $\lambda_A(i) = (1/\sigma_A)\sigma_i u_A(i)$. If i is not a neighbor of A, then, for any boundary node B, $B \in N(i)$, $u_A(B) = 0$, so B's contribution to the above sum is zero. Thus, equation (8) holds for $\lambda_A(i) = (1/\sigma_A)\sigma_i u_A(i)$.

If, on the other hand, i is a neighbor of A, then since $u_A(A) = 1$ the above sum becomes

$$\left(\frac{1}{\sigma_A}\right)\sigma_i u_A(i) = \left(\sum_{\substack{j\in N(i)\\ j \text{ interior vertex}}} \frac{(1/r_{ij})\left(\frac{1}{\sigma_A}\right)\sigma_j u_A(j)}{\sigma_j}\right) + \frac{(1/r_{iA})\left(\frac{1}{\sigma_A}\right)\sigma_A}{\sigma_A}$$

Since $P_{Ai} = \frac{(1/r_{iA})}{\sigma_A}$, equation (9) holds as well, and, thus, $\lambda_A(i)$ must equal $(1/\sigma_A)\sigma_i u_A(i)$ on the interior. We also know that $\lambda_A(A) = 1 = (1/\sigma_A)\sigma_A u_A(A)$, and, for any boundary node B, $A \neq B$, $\lambda_A(B) = 0 = (1/\sigma_A)\sigma_B u_A(B)$. Hence, $\lambda_A(i) = (1/\sigma_A)\sigma_i u_A(i)$ holds for all vertices i. □

5.3. **The vertex inner product.** We will use Lemma 5.3 and some probability arguments about paths from boundary node A to boundary node B to calculate $\lambda_{AB}(i)$. If we, in turn, calculate $E[X_{AB}] = \sum_{\text{all vertices } j} \lambda_{AB}(j)$., the result will involve an interesting inner product.

If we, now, consider a random walk starting boundary node A finishing at boundary node B (possibly equal to A), we might intuit that the expected number of leavings from vertex i on such random walks, $\lambda_{AB}(i)$, is equal to the expected number of leaving on a random walk from A to anywhere on the boundary, $\lambda_A(i)$, times the probability of getting from i to B, p_{iB} $[= u_B(i)]$, that is,

$$\lambda_{AB}(i) = \lambda_A(i)u_B(i) = \left(\frac{1}{\sigma_A}\right)\sigma_i u_A(i)u_B(i).$$

This turns out to be true for vertices i, $i \neq A$, but requires some care to demonstrate.

Lemma 5.4. *Given boundary nodes A and B possibly equal, let $\lambda_A(i)$ and $\lambda_{AB}(i)$ are as defined in section 5.1, namely, $\lambda_A(i)$ is the expected number of leavings from vertex i starting at boundary node A and ending the next time the boundary is reached, and $\lambda_{AB}(i)$ is the expected number of leavings from vertex i in such random walks counting only leavings for walks that end at boundary node B. Moreover, if $u_B(i)$ is the solution for all vertices i to the Dirichet problem on the resistance network satisfying the boundary conditions:*

$$u_B(B) = 1$$

while, for any other boundary node $C, C \neq B$

$$u_B(C) = 0$$

then, for all vertices i, $i \neq A$,

$$\lambda_{AB}(i) = \lambda_A(i)u_B(i)$$

Proof. For interior vertex i and boundary nodes A and B possibly equal, Let $p_{\underrightarrow{Ai}}$ be the probability of random walk going directly from A to i, that is, not visiting i only at the end of the walk. Let $p_{\underrightarrow{i}}$ be the probability of random walk going directly from i to the boundary, that is, not returning to i along the way. Let $p_{\underrightarrow{iB}}$ be the probability of random walk going directly from i to B, again, not returning to i along the way. Finally, let r_i be the probability of random walk returning to i.

Notice that if a random walk is at i it can go directly to the boundary or it can return to i, so $p_{\underrightarrow{i}} + r_i = 1$ or $p_{\underrightarrow{i}} = 1 - r_i$. Notice also that we can use these defined quantities to calculate $\lambda_A(i)$ directly as one times the probability of a one visit random walk plus two times the probability of a two visit random walk and so on. This translates into one times the probability of going from A directly to i and directly from i to the boundary plus two times the probability of going from A directly to i, returning to i once, and then going directly from i to the boundary etc, namely,

$$\begin{aligned}
\lambda_A(i) = \quad & 1\times \quad p_{\underrightarrow{Ai}}\, p_{\underrightarrow{i}} && \text{(one visit to } i \text{)} \\
+ \quad & 2\times \quad p_{\underrightarrow{Ai}}\, r_i\, p_{\underrightarrow{i}} && \text{(two visits to } i \text{)} \\
+ \quad & 3\times \quad p_{\underrightarrow{Ai}}\, r_i^2\, p_{\underrightarrow{i}} && \text{(three visits to } i \text{)} \\
& \vdots && \vdots
\end{aligned} \tag{11}$$

We can make similar calculation for $\lambda_{AB}(i)$.

$$
\begin{array}{rcll}
\lambda_{AB}(i) = & 1\times & p_{\underrightarrow{Ai}}\, p_{\underrightarrow{iB}} & \text{(one visit to } i \text{)} \\
+ & 2\times & p_{\underrightarrow{Ai}}\, r_i\, p_{\underrightarrow{iB}} & \text{(two visits to } i \text{)} \\
+ & 3\times & p_{\underrightarrow{Ai}}\, r_i^2\, p_{\underrightarrow{iB}} & \text{(three visits to } i \text{)} \\
& \vdots & & \vdots
\end{array} \tag{12}
$$

Comparing equations (11) and (12), we realize that $\lambda_{AB}(i) = \lambda_A(i)(p_{\underrightarrow{iB}}/p_{\underrightarrow{i}})$. Finally, using the same approach, we can calculate p_{iB} the probability a random walk going from i to B:

$$
\begin{array}{rcll}
p_{iB} = & & p_{\underrightarrow{iB}} & \text{(go directly to } B \text{)} \\
& + & r_i\, p_{\underrightarrow{iB}} & \text{(return to } i \text{ once, then go directly to } B \text{)} \\
& + & r_i^2\, p_{\underrightarrow{iB}} & \text{(return to } i \text{ twice, then go directly to } B \text{)} \\
& \vdots & & \vdots \\
& = & p_{\underrightarrow{iB}} \sum_{n=0}^{\infty} r_i^n \quad = \quad \dfrac{p_{\underrightarrow{iB}}}{1-r_i} & = \quad \dfrac{p_{\underrightarrow{iB}}}{p_{\underrightarrow{i}}}.
\end{array} \tag{13}
$$

So, $\lambda_{AB}(i) = \lambda_A(i)p_{iB} = \lambda_A(i)u_B(i)$ holds for all interior vertices i. Since, for boundary node C, $C \neq A$, $\lambda_{AB}(C) = \lambda_A(C) = 0$, it holds for all vertices i, $i \neq A$. □

If we substitute A for i in $\lambda_{AB}(i) = \lambda_A(i)u_B(i) = \left(\frac{1}{\sigma_A}\right)\sigma_i u_A(i)u_B(i)$, we get $\lambda_{AB}(A) = 0$, which is obviously untrue. We may, however, calculate $\lambda_{AB}(A)$ directly. Every random walk from A to any node on the boundary has one leaving from node A, so $\lambda_A(A) = 1$. $\lambda_{AB}(A)$, however, counts only those random walk paths that end at B. Hence, it will equal one times the probability of getting from A to B, which is, as discussed in remark 4.2,

$$
\lambda_{AB}(A) = \sum_{j\in N(A)} P_{Aj}p_{jB} = \sum_{j\in N(A)} \left[\frac{(1/r_{Aj})u_B(j)}{\sigma_A}\right]. \tag{14}
$$

We can now make the calculation

$$
E[X_{AB}] = \sum_{\text{all vertices } i} \lambda_{AB}(i) =
$$

$$
= \sum_{\substack{\text{all vertices } i,\\ i\neq A}} \left(\frac{1}{\sigma_A}\right)\sigma_i u_A(i)u_B(i) \;+ \sum_{j\in N(A)} \left[\frac{(1/r_{Aj})u_B(j)}{\sigma_A}\right] \tag{15}
$$

$$
= \left(\frac{1}{\sigma_A}\right)\left[\sum_{\substack{\text{all vertices } i,\\ i\neq A}} \sigma_i u_A(i)u_B(i) \;+ \sum_{j\in N(A)} (1/r_{Aj})u_B(j)\right].
$$

For our special model where all resistances are one, this becomes

$$
E[X_{AB}] = \left(\frac{1}{d_A}\right)\left[\sum_{\substack{\text{all vertices } i,\\ i\neq A}} d_i u_A(i)u_B(i) \;+ \sum_{j\in N(A)} u_B(j)\right]. \tag{16}
$$

where d_i is the degree of vertex i. Note that $\sum \sigma_i u_A(i) u_B(i)$ or, alternatively, $\sum d_i u_A(i) u_B(i)$ has the form of a weighted inner product. We shall refer this sum as the random walk inner product over vertices.

Remark 5.5. Before launching into a formal definition, we should call attention to a small abuse of notation. Up to this point, we have use $u_A(i)$ and $u_B(i)$ to denote the the solutions to the Dirichlet problem on the vertices of network where for boundary node C, in the first case, $u_A(C) = 1$, if $A = C$, and $u_A(C) = 0$ otherwise and, in the second case, $u_B(C) = 1$, if $B = C$, and $u_B(C) = 0$ otherwise. This is to say the boundary values are the basis vectors e_A and e_B, respectively. We will want to denote sometimes use the notation $u_{e_A}(i)$ and $u_{e_B}(i)$ for $u_A(i)$ and $u_B(i)$, because this notation also allows us to easily denote solutions to the Dirichlet problem with an arbitrary vector, ϕ, the boundary condition, as $u_\phi(i)$. This alternative notation should not cause any confusion.

Definition 5.6. *Let ϕ and ψ be two vectors representing the boundary conditions for two solutions, $u_\phi(i)$ and $u_\psi(i)$, to the Dirichlet problem for resistance network G on all vertices i in G. Also, let r_{ij} be the resistance on edge $\{i, j\}$ and $\sigma_i = \sum_{k \in N(i)} 1/r_{ik}$. The random walk inner product over vertices of ϕ and ψ is*

$$\langle \phi, \psi \rangle_G = \sum_{\substack{\text{all vertices } i \\ \text{in } G}} \sigma_i u_\phi(i) u_\psi(i). \tag{17}$$

For our special model where all resistances are one, this may be written

$$\langle \phi, \psi \rangle_G = \sum_{\substack{\text{all vertices } i \\ \text{in } G}} d_i u_\phi(i) u_\psi(i). \tag{18}$$

where d_i is the degree of vertex i. We shall often omit the subscript G where it is obvious and, also, often use $\langle A, B \rangle_G$ to mean $\langle e_A, e_B \rangle_G$.

It can be readily verified that this construction has all the properties of an inner product: it is symmetric $[\langle \phi, \psi \rangle = \langle \psi, \phi \rangle]$, positive $[\langle \phi, \phi \rangle > 0$ if $\phi \neq 0]$, definite $[\langle \phi, \phi \rangle = 0$ impies $\phi = 0]$. It is less obvious that it is bilinear, but this follows since the solutions to the Dirichlet problem are linear in their boundary conditions, that is, $u_{a\phi + b\psi}(i) = a u_\phi(i) + b u_\psi(i)$ for any real numbers a, b.

For the case of a simple random walk on a network G where resistances are one, given boundary nodes $A \neq B$,

$$\langle A, B \rangle_G = d_A E[X_{AB}] - \sum_{j \in N(A)} u_B(j) \tag{19}$$

affords a straightforward interpretation. As discussed earlier, $(1/d_A) \sum_{j \in N(A)} u_B(j)$ is the expected number of leavings from A. So, subtracting them, means $E[X_{AB}] - (1/d_A) \sum_{j \in N(A)} u_B(j)$ is the expected number of steps on a random walk from A to B, just forgetting to count the first step. The first term in the product of probabilities for paths summed in this expectation is $1/d_A$ as every path still starts by leaving A. Thus, multiplying by d_A, pushes A out of the random walk entirely and gives us the expected total number of steps if we started d_A random walks each beginning at each of the neighbors of A. In other words,

$$\langle A, B \rangle_G = \sum_{j \in N(A)} E[X_{jB}] \tag{20}$$

The interpretation of $\langle A, A\rangle_G$ works the same way except in that case the inner product sum includes the term $d_A u_A(A) u_A(A)$ $[= d_A]$ for vertex A, so equation (19) becomes

$$\langle A, A\rangle_G - d_A = d_A E[X_{AA}] - \sum_{j\in N(A)} u_A(j), \tag{21}$$

and, hence,

$$\langle A, A\rangle_G - d_A = \sum_{j\in N(A)} E[X_{jA}]. \tag{22}$$

A similar interpretation is possible for the network model in which the resistances vary, but it is much less straight forward.

Example 5.7. Let's calculate this inner product for a simple network of a line of n unit resistors connecting boundary nodes A and B. Labeling the vertices

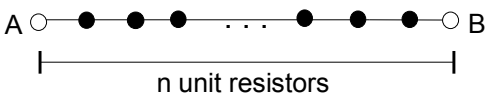

FIGURE 2. A linear network of unit resistors.

as $A = 0, 1, 2, \cdots n-1, n = B$, the solutions to the Dirichlet problem are linear: $u_A(i) = (n-i)/n$ and $u_B(i) = i/n$. So,

$$\begin{aligned}
\langle B, B\rangle &= 1u_B(0)^2 + \sum_{i=1}^{n-1} 2u_B(i)^2 + 1u_B(n)^2 \\
&= 1 + 2\sum_{i=1}^{n-1} \frac{i^2}{n^2} \\
&= 1 + 2\left[\frac{n(n-1)(2n-1)}{6n^2}\right] \\
&= \frac{3n}{3n} + \frac{2n^2-3n+1}{3n} = \frac{2}{3}n + \frac{1}{3n}
\end{aligned}$$

$\langle A, A\rangle$ is the same, since it is the same sum done in reverse order, and

$$\begin{aligned}
\langle A, B\rangle = \langle B, A\rangle &= 1u_A(0)u_B(0) + \sum_{i=1}^{n-1} 2u_A(i)u_B(i) + 1u_A(n)u_B(n) \\
&= 2\sum_{i=1}^{n-1} \frac{(n-i)i}{n^2} \\
&= 2\left[n\frac{n(n-1)}{2n^2} - \frac{n(n-1)(2n-1)}{6n^2}\right] \\
&= \frac{3n^2-3n}{3n} - \frac{2n^2-3n+1}{3n} = \frac{1}{3}n - \frac{1}{3n}.
\end{aligned}$$

Since, in this example, the degree of A is one, $\langle A, B\rangle$ is expected number of steps in a random walk starting at A counting only those paths that end at B. Similarly, $\langle A, A\rangle - 1$ is expected number of steps in a random walk starting at A counting only those paths that end at A.

Note also in this example

$$\langle A, A\rangle + \langle A, B\rangle + \langle B, A\rangle + \langle B, B\rangle = 2n,$$

which, along with a number other calculations, motivates the following theorem.

Theorem 5.8. *Let G be a resistance network, let a random walk be defined on that network using either model. Let ∂V be the set of boundary nodes of G.*

$$\sum_{A,B\in\partial V} \langle A, B\rangle_G = \sum_{\substack{all\ vertices\ i \\ in\ G}} \sigma_i = 2\left(\sum_{\substack{all\ edges \\ \{i,j\}\ in\ G}} \frac{1}{r_{ij}}\right).$$

This may not seem especially interesting as stated, but in the special model where resistances are one, the above sum of inner products equals twice the number of edges in the network. This portends good things for the inverse problem to come.

Proof.

$$\sum_{A,B\in\partial V} \langle A, B\rangle_G = \sum_{A,B\in\partial V} \sum_{\substack{\text{all vertices } i \\ \text{in } G}} \sigma_i u_A(i) u_B(i)$$

$$= \sum_{\substack{\text{all vertices } i \\ \text{in } G}} \sigma_i \sum_{A\in\partial V} u_A(i) \sum_{B\in\partial V} u_B(i)$$

Since a random walk beginning at some vertex i must eventually walk out at some boundary node, $\sum_{A\in\partial V} u_A(i) = \sum_{B\in\partial V} u_B(i) = 1$, so

$$\sum_{A,B\in\partial V} \langle A, B\rangle_G = \sum_{\substack{\text{all vertices } i \\ \text{in } G}} \sigma_i.$$

Since this last term sums $(1/r_{ij})$ on each end of each edge $\{i, j\}$, it is equal to

$$2\left(\sum_{\substack{\text{all edges} \\ \{i,j\}\ \text{in } G}} \frac{1}{r_{ij}}\right).$$

□

Example 5.9. As an simple application of theorem 5.8, imagine we have resistance network with only one boundary node A. If we model this network using unit resistances, we have from equation (21)

$$\langle A, A\rangle_G - d_A = d_A E[X_{AA}] - \sum_{j\in N(A)} u_A(j).$$

The solution to the Dirichlet problem in this instance is easy: $u_A(i) = 1$ for all i. Hence, by theorem 5.8,

$$\langle A, A\rangle_G = d_A E[X_{AA}] = 2 \times (\text{the number of edges in } G).$$

5.4. **The edge inner product.** We can go through the same line of thinking that we went though for vertices in section 5.3 and apply it to edges, that is, use probability arguments to calculate $\kappa_{AB}(i,j)$, and then calculate

$$E[X_{AB}] = \sum_{\text{all edges } \{i,j\}} [\kappa_{AB}(i,j) + \kappa_{AB}(j,i)].$$

The result will involve yet another inner product.

Lemma 5.10. *Given boundary nodes A and B possibly equal, let $\lambda_A(i)$ and $\kappa_{AB}(i,j)$ are as defined in section 5.1, namely, $\lambda_A(i)$ is the expected number of leavings from vertex i starting at boundary node A and ending the next time the boundary is reached, and $\kappa_{AB}(i,j)$ is the expected number of crossings of edge $\{i,j\}$ from vertex i to j in such random walks counting only crossings for walks that end at boundary node B. Moreover, if $u_B(i) = u_{e_B}(i)$ is the solution for all vertices i to the Dirichet problem on the resistance network satisfying the boundary conditions:*

$$u_B(B) = 1$$

while, for any other boundary node $C, C \neq B$

$$u_B(C) = 0$$

then, for all vertices i and j, where $j \in N(i)$,

$$\kappa_{AB}(i,j) = \lambda_A(i) P_{ij} u_B(j)$$

where P_{ij} is the probability of a random walk at vertex i crossing to j.

Proof. The argument is very similar to that for lemma 5.4. Our definitions will be a little different, however. For boundary nodes A and B possibly equal, and vertex i, that is either an interior vertex or $i = A$, and for vertex j neighboring i, let $p_{\underrightarrow{Ai}}$ be the probability of random walk going 'directly' from A to i, in a crossing sense this time, that is, without crossing from i to j. Let $p_{\underrightarrow{j}}$ be the probability of random walk going 'directly' from i to the boundary, that is, not returning to j by crossing from i anywhere in the process. Similarly, let $p_{\underrightarrow{iB}}$ be the probability of random walk going 'directly' from i to B, again, not crossing from i to j along the way. Finally, let r_{ij} be the probability of random walk returning to j by crossing from i. Note that if $i = A$, $r_{ij} = 0$.

As before, if a random walk is at i it can go 'directly' to the boundary or it can return to i, by crossing edge $\{i,j\}$, so $p_{\underrightarrow{j}} + r_{ij} = 1$ or $p_{\underrightarrow{j}} = 1 - r_{ij}$. Notice, also, on a random walk from A to the boundary, every time it leaves i, it has probability P_{ij} of crossing to j. So, on such a random walk, the expected number of crossings from i to j is $\lambda_A(i)P_{ij}$. We can expand this as we did $\lambda_A(i)$ in the proof of lemma 5.4,

$$\begin{aligned}
\lambda_A(i)P_{ij} = \quad & 1\times \quad p_{\underrightarrow{Ai}}\, P_{ij}\, p_{\underrightarrow{j}} && \text{(one crossing from } i \text{ to } j) \\
+ \quad & 2\times \quad p_{\underrightarrow{Ai}}\, P_{ij} r_{ij}\, p_{\underrightarrow{j}} && \text{(two crossings from } i \text{ to } j) \\
+ \quad & 3\times \quad p_{\underrightarrow{Ai}}\, P_{ij} r_{ij}^2\, p_{\underrightarrow{j}} && \text{(three crossings from } i \text{ to } j) \\
& \vdots && \vdots
\end{aligned}$$

Of course, we can make similar calculation for $\kappa_{AB}(i,j)$ in which we only count crossings on paths that end at B. There we would replace the $p_{\underrightarrow{j}}$ on the end of

each term with a $\underset{\rightarrow}{p_{jB}}$. Hence, $\kappa_{AB}(i,j) = \lambda_A(i)P_{ij}\left(\underset{\rightarrow}{p_{jB}}/\underset{\rightarrow}{p_j}\right)$. Finally, we can calculate p_{jB}, the probability a random walk going from j to B:

$$\begin{aligned}
p_{jB} = \quad & \underset{\rightarrow}{p_{jB}} && \text{(go directly to } B\text{)} \\
+ \quad & r_{ij}\,\underset{\rightarrow}{p_{jB}} && \text{(return to } j \text{ by crossing from } i\text{, once, then directly to } B\text{)} \\
+ \quad & r_{ij}^2\,\underset{\rightarrow}{p_{jB}} && \text{(return to } j \text{ by crossing from } i\text{, twice, then directly to } B\text{)} \\
\vdots \quad & \qquad\vdots \\
= \quad & \underset{\rightarrow}{p_{jB}} \sum_{n=0}^{\infty} r_{ij}^n \quad = \quad \frac{\underset{\rightarrow}{p_{jB}}}{1-r_{ij}} \quad = \quad \frac{\underset{\rightarrow}{p_{jB}}}{\underset{\rightarrow}{p_j}}.
\end{aligned}$$

And so,

$$\kappa_{AB}(i,j) = \lambda_A(i)P_{ij}p_{jB} = \lambda_A(i)P_{ij}u_B(j) \tag{23}$$

holds for interior vertex i or if $i = A$ with neighboring vertex j. If i is a boundary node, and $i \neq A$ then $\kappa_{AB}(i,j) = 0$, but $\lambda_A(i) = 0$, so equation (23) holds for all neighboring vertices i and j. □

With this, we can calculate the expected number of steps in a random walk starting at A counting only those walks that end at B,

$$\begin{aligned}
E[X_{AB}] &= \sum_{\text{all edges } \{i,j\}} [\kappa_{AB}(i,j) + \kappa_{AB}(j,i)] \\
&= \sum_{\text{all edges } \{i,j\}} [\lambda_A(i)P_{ij}u_B(j) + \lambda_A(j)P_{ji}u_B(i)] \qquad (24) \\
&= \sum_{\text{all edges } \{i,j\}} \left[\left(\frac{1}{\sigma_A}\right)\sigma_i u_A(i)\left(\frac{1/r_{ij}}{\sigma_i}\right)u_B(j) + \left(\frac{1}{\sigma_A}\right)\sigma_j u_A(j)\left(\frac{1/r_{ij}}{\sigma_j}\right)u_B(i)\right] \\
&= \left(\frac{1}{\sigma_A}\right)\sum_{\text{all edges } \{i,j\}} \left(\frac{1}{r_{ij}}\right)[u_A(i)u_B(j) + u_A(j)u_B(i)]\,.
\end{aligned}$$

In our model with unit resistances, this is

$$E[X_{AB}] = \left(\frac{1}{d_A}\right)\sum_{\text{all edges } \{i,j\}} [u_A(i)u_B(j) + u_A(j)u_B(i)]\,. \tag{25}$$

Once again, $E[X_{AB}]$ involves a sum with the form of an inner product, in this case, over edges instead of vertices.

Definition 5.11. *Let ϕ and ψ be two vectors representing the boundary conditions for two solutions, $u_\phi(i)$ and $u_\psi(i)$, to the Dirichlet problem for resistance network G on all vertices i in G, and r_{ij} be the resistance on edge $\{i,j\}$. The random walk inner product over edges of ϕ and ψ is*

$$\langle\widetilde{\phi,\psi}\rangle_G = \sum_{\textit{all edges } \{i,j\}} \left(\frac{1}{r_{ij}}\right)[u_\phi(i)u_\psi(j) + u_\phi(j)u_\psi(i)]\,. \tag{26}$$

(Again, this can be readily shown to have all the properties of an inner product.)

This time the interpretation of the inner product is simple and straightforward regardless of the random walk model we use. Let A and B be boundary nodes of resistance network G. For the general model,

$$\widetilde{\langle A,B\rangle}_G \ [= \widetilde{\langle e_A, e_B\rangle}_G] = \sigma_A E[X_{AB}] \tag{27}$$

For the unit resistor model,

$$\widetilde{\langle A,B\rangle}_G = d_A E[X_{AB}] \tag{28}$$

where d_A is the degree of boundary node A.

Lets calculate the edge inner product for the simple line of n unit resistors connecting boundary nodes A and B just as we did for the vertex inner product in example 5.7. Remember, we labeled $A = 0, 1, 2, \cdots n-1, n = B$, the solutions to the Dirichlet problem were $u_A(i) = (n-i)/n$ and $u_B(i) = i/n$. Thus,

$$\widetilde{\langle A,A\rangle} = \widetilde{\langle B,B\rangle} = \sum_{i=0}^{n-1} 2\left[\frac{i(i+1)}{n^2}\right] = 2\left[\frac{n(n-1)}{2n^2} + \frac{n(n-1)(2n-1)}{6n^2}\right]$$

$$= \frac{3(n-1)}{3n} + \frac{(n-1)(2n-1)}{3n} = \frac{2(n-1)(n+1)}{3n} = \frac{2}{3}n - \frac{2}{3n}$$

and

$$\widetilde{\langle A,B\rangle} = \widetilde{\langle B,A\rangle} = \sum_{i=0}^{n-1}\left[\frac{i(n-i-1)}{n^2} + \frac{(i+1)(n-i)}{n^2}\right] = \sum_{i=0}^{n-1}\left[\frac{n+2i(n-1)-2i^2}{n^2}\right]$$

$$= \left[\frac{n^2}{n^2} + \frac{n(n-1)^2}{n^2} - \frac{n(n-1)(2n-1)}{3n^2}\right] = \frac{1}{3n}\left[3n + 3n^2 - 6n + 3 - 2n^2 + 3n - 1\right]$$

$$= \frac{1}{3n}\left[n^2 + 2\right] = \frac{1}{3}n + \frac{2}{3n}$$

Notice that this is slightly different from the values calculated for the vertex inner product. We will expore the relationship between the two inner products in the next section. Now, we must prove the edge inner product version of theorem 5.8.

Theorem 5.12. *Let G be a resistance network, let a random walk be defined on that network using either model. Let ∂V be the set of boundary nodes of G.*

$$\sum_{A,B\in\partial V} \widetilde{\langle A,B\rangle}_G = \sum_{\substack{\text{all vertices } i \\ \text{in } G}} \sigma_i = 2\left(\sum_{\substack{\text{all edges} \\ \{i,j\} \text{ in } G}} \frac{1}{r_{ij}}\right).$$

Proof.

$$\sum_{A,B\in\partial V} \widetilde{\langle A,B\rangle}_G = \sum_{A,B\in\partial V} \sum_{\substack{\text{all edges} \\ \{i,j\} \text{ in } G}} \left(\frac{1}{r_{ij}}\right)\left[u_A(i)u_B(j) + u_A(j)u_B(i)\right]$$

$$= \sum_{\substack{\text{all edges} \\ \{i,j\} \text{ in } G}} \left(\frac{1}{r_{ij}}\right)\left[\sum_{A\in\partial V} u_A(i) \sum_{B\in\partial V} u_B(j) + \sum_{A\in\partial V} u_A(j) \sum_{B\in\partial V} u_B(i)\right]$$

Since a random walk beginning at some vertex i must eventually walk out at some boundary node, $\sum_{A\in\partial V} u_A(i) = \sum_{B\in\partial V} u_B(i) = 1$, so

$$\sum_{A,B\in\partial V} \widetilde{\langle A,B\rangle}_G = \sum_{\substack{\text{all edges}\\ \{i,j\} \text{ in } G}} 2\left(\frac{1}{r_{ij}}\right)$$

Since this last term sums $(1/r_{ij})$ on each end of each edge $\{i,j\}$, we can apply each $(1/r_{ij})$ to each vertex the edge contacts, so

$$\sum_{A,B\in\partial V} \widetilde{\langle A,B\rangle}_G = \sum_{\substack{\text{all vertices } i\\ \text{in } G}} \sum_{j\in N(i)} \frac{1}{r_{ij}} = \sum_{\substack{\text{all vertices } i\\ \text{in } G}} \sigma_i$$

□

6. Relationships between the vertex inner product, the edge inner product, and the Dirichlet-Neumann map

two [...]

atop
abottom

References

[1] McCormick, Megan. "Metric Recoverability" 2005.
http://www.math.washington.edu/ reu/papers/2005/megan/metric6.pdf