Spectral Graph Theory Lecture 14

More Effective Resistance

Daniel A. Spielman October 15, 2018

14.1 Introduction

My plan for this lecture is to teach too much:

1. The Matrix Tree Theorem.
2. Effective Resistance / Leverage Scores, and the probability an edge appears in a random spanning tree.
3. Estimating effective resistances quickly.
4. Rayleigh's Monotonicity Theorem.

14.2 Effective Resistance and Energy Dissipation

In the last lecture we saw two ways of defining effective resistance. I will define it one more way, but skip the proof. If a current f flows through a resistor of resistance R, the amount of energy that is dissipated as heat is proportional to Rf^2. If the potential difference across the resistor is v, then $f = v/R$, and the energy dissipation is

$$Rf^2 = v^2/R = wv^2,$$

where w is the weight of the edge. We can define the effective resistance between vertices a and b to be the minimum of the total energy dissipation when we flow one unit of current from a to b. You could compute this by evaluating the Laplacian quadratic form on the vector of voltages induced by this flow.

14.3 Determinants

To begin, we review some facts about determinants of matrices and characteristic polynomials. We first recall the Leibniz formula for the determinant of a square matrix $\boldsymbol{A}$:

$$\det(\boldsymbol{A}) = \sum_{\pi} \left(\mathsf{sgn}(\pi) \prod_{i=1}^{n} \boldsymbol{A}(i, \pi(i)) \right), \tag{14.1}$$

where the sum is over all permutations π of $\{1, \ldots, n\}$.

Also recall that the determinant is multiplicative, so for square matrices $\boldsymbol{A}$ and $\boldsymbol{B}$

$$\det(\boldsymbol{A}\boldsymbol{B}) = \det(\boldsymbol{A}) \det(\boldsymbol{B}). \tag{14.2}$$

Elementary row operations do not change the determinant. If the columns of $\boldsymbol{A}$ are the vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$, then for every c

$$\det\left(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\right) = \det\left(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n + c\boldsymbol{a}_1\right).$$

This fact gives us two ways of computing the determinant. The first comes from the fact that we can apply elementary row operations to transform $\boldsymbol{A}$ into an upper triangular matrix, and (14.1) tells us that the determinant of an upper triangular matrix is the product of its diagonal entries.

The second comes from the observation that the determinant is the volume of the parallelepiped with axes $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$: the polytope whose corners are the origin and $\sum_{i \in S} \boldsymbol{a}_i$ for every $S \subseteq \{1, \ldots, n\}$. Let

$$\boldsymbol{\Pi}_{\boldsymbol{a}_1}$$

be the symmetric projection orthogonal to $\boldsymbol{a}_1$. As this projection amounts to subtracting off a multiple of $\boldsymbol{a}_1$ and elementary row operations do not change the determinant,

$$\det\left(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\right) = \det\left(\boldsymbol{a}_1, \boldsymbol{\Pi}_{\boldsymbol{a}_1}\boldsymbol{a}_2, \ldots, \boldsymbol{\Pi}_{\boldsymbol{a}_1}\boldsymbol{a}_n\right).$$

The volume of this parallelepiped is $\|\boldsymbol{a}_1\|$ times the volume of the parallelepiped formed by the vectors $\boldsymbol{\Pi}_{\boldsymbol{a}_1}\boldsymbol{a}_2, \ldots, \boldsymbol{\Pi}_{\boldsymbol{a}_1}\boldsymbol{a}_n$. I would like to write this as a determinant, but must first deal with the fact that these are $n-1$ vectors in an n dimensional space. The way we first learn to handle this is to project them into an $n-1$ dimensional space where we can take the determinant. Instead, we will employ other elementary symmetric functions of the eigenvalues.

14.4 Characteristic Polynomials

Recall that the characteristic polynomial of a matrix $\boldsymbol{A}$ is

$$\det(x\boldsymbol{I} - \boldsymbol{A}).$$

I will write this as

$$\sum_{k=0}^{n} x^{n-k}(-1)^k \sigma_k(\boldsymbol{A}),$$

where $\sigma_k(\boldsymbol{A})$ is the kth elementary symmetric function of the eigenvalues of $\boldsymbol{A}$, counted with algebraic multiplicity:

$$\sigma_k(\boldsymbol{A}) = \sum_{|S|=k} \prod_{i \in S} \lambda_i.$$

Thus, $\sigma_1(\boldsymbol{A})$ is the trace and $\sigma_n(\boldsymbol{A})$ is the determinant. From this formula, we know that these functions are invariant under similarity transformations.

In Exercise 3 from Lecture 2, you were asked to prove that

$$\sigma_k(\boldsymbol{A}) = \sum_{|S|=k} \det(\boldsymbol{A}(S,S)). \tag{14.3}$$

This follows from applying the Leibnitz formula (14.1) to $\det(x\boldsymbol{I} - \boldsymbol{A})$.

If we return to the vectors $\boldsymbol{\Pi}_{\boldsymbol{a}_1}\boldsymbol{a}_2, \dots, \boldsymbol{\Pi}_{\boldsymbol{a}_1}\boldsymbol{a}_n$ from the previous section, we see that the volume of their parallelepiped may be written

$$\sigma_{n-1}\left(\boldsymbol{0}_n, \boldsymbol{\Pi}_{\boldsymbol{a}_1}\boldsymbol{a}_2, \dots, \boldsymbol{\Pi}_{\boldsymbol{a}_1}\boldsymbol{a}_n\right),$$

as this will be the product of the $n-1$ nonzero eigenvalues of this matrix.

Recall that the matrices $\boldsymbol{B}\boldsymbol{B}^T$ and $\boldsymbol{B}^T\boldsymbol{B}$ have the same eigenvalues, up to some zero eigenvalues if they are rectangular. So,

$$\sigma_k(\boldsymbol{B}\boldsymbol{B}^T) = \sigma_k(\boldsymbol{B}^T\boldsymbol{B}).$$

This gives us one other way of computing the absolute value of the product of the nonzero eigenvalues of the matrix

$$\left(\boldsymbol{\Pi}_{\boldsymbol{a}_1}\boldsymbol{a}_2, \dots, \boldsymbol{\Pi}_{\boldsymbol{a}_1}\boldsymbol{a}_n\right).$$

We can instead compute their square by computing the determinant of the square matrix

$$\begin{pmatrix} \boldsymbol{\Pi}_{\boldsymbol{a}_1}\boldsymbol{a}_2 \\ \vdots \\ \boldsymbol{\Pi}_{\boldsymbol{a}_1}\boldsymbol{a}_n \end{pmatrix} \left(\boldsymbol{\Pi}_{\boldsymbol{a}_1}\boldsymbol{a}_2, \dots, \boldsymbol{\Pi}_{\boldsymbol{a}_1}\boldsymbol{a}_n\right).$$

When $\boldsymbol{B}$ is a singular matrix of rank k, $\sigma_k(\boldsymbol{B})$ acts as the determinant of $\boldsymbol{B}$ restricted to its span. Thus, there are situations in which σ_k is multiplicative. For example, if $\boldsymbol{A}$ and $\boldsymbol{B}$ both have rank k and the range of $\boldsymbol{A}$ is orthogonal to the nullspace of $\boldsymbol{B}$, then

$$\sigma_k(\boldsymbol{B}\boldsymbol{A}) = \sigma_k(\boldsymbol{B})\sigma_k(\boldsymbol{A}). \tag{14.4}$$

We will use this identity in the case that $\boldsymbol{A}$ and $\boldsymbol{B}$ are symmetric and have the same nullspace.

14.5 The Matrix Tree Theorem

We will state a slight variant of the standard Matrix-Tree Theorem. Recall that a spanning tree of a graph is a subgraph that is a tree.

Theorem 14.5.1. *Let $G = (V, E, w)$ be a connected, weighted graph. Then*

$$\sigma_{n-1}(\boldsymbol{L}_G) = n \sum_{\textit{spanning trees } T} \prod_{e \in T} w_e.$$

Thus, the eigenvalues allow us to count the sum over spanning trees of the product of the weights of edges in those trees. When all the edge weights are 1, we just count the number of spanning trees in G.

We first prove this in the case that G is just a tree.

Lemma 14.5.2. *Let $G = (V, E, w)$ be a weighted tree. Then,*

$$\sigma_{n-1}(\boldsymbol{L}_G) = n \prod_{e \in E} w_e.$$

Proof. For $a \in V$, let $S_a = V - \{a\}$. We know from (14.3)

$$\sigma_{n-1}(\boldsymbol{L}_G) = \sum_{a \in V} \det(\boldsymbol{L}_G(S_a, S_a).$$

We will prove that for every $a \in V$,

$$\det(\boldsymbol{L}_G(S_a, S_a)) = \prod_{e \in E} w_e.$$

Write $\boldsymbol{L}_G = \boldsymbol{U}^T \boldsymbol{W} \boldsymbol{U}$, where $\boldsymbol{U}$ is the signed edge-vertex adjacency matrix and $\boldsymbol{W}$ is the diagonal matrix of edge weights. Write $\boldsymbol{B} = \boldsymbol{W}^{1/2} \boldsymbol{U}$, so

$$\boldsymbol{L}_G(S_a, S_a) = \boldsymbol{B}(:, S_a)^T \boldsymbol{B}(:, S_a),$$

and

$$\det(\boldsymbol{L}_G(S_a, S_a)) = \det(\boldsymbol{B}(:, S_a))^2,$$

where we note that $\boldsymbol{B}(:, S_a)$ is square because a tree has $n-1$ edges and so $\boldsymbol{B}$ has $n-1$ rows.

To see what is going on, first consider the case in which G is a weighted path and a is the first vertex. Then,

$$\boldsymbol{U} = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & -1 \end{pmatrix}, \quad \text{and} \quad \boldsymbol{B}(:, S_1) = \begin{pmatrix} -\sqrt{w_1} & 0 & \cdots & 0 \\ \sqrt{w_2} & -\sqrt{w_2} & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & -\sqrt{w_{n-1}} \end{pmatrix}.$$

We see that $\boldsymbol{B}(:, S_1)$ is a lower-triangular matrix, and thus its determinant is the product of its diagonal entries, $-\sqrt{w_i}$.

To see that the same happens for every tree, renumber the vertices (permute the columns) so that a comes first, and that the other vertices are ordered by increasing distance from 1, breaking ties arbitrarily. This permutations can change the sign of the determinant, but we do not care because we are going to square it. For every vertex $c \neq 1$, the tree now has exactly one edge (b, c) with $b < c$. Put such an edge in position $c - 1$ in the ordering, and let w_c indicate its weight. Now, when we remove the first column to form $\boldsymbol{B}(:, S_1)$, we produce a lower triangular matrix with the entry $-\sqrt{w_c}$ on the cth diagonal. So, its determinant is the product of these terms and

$$\det(\boldsymbol{B}(:, S_a))^2 = \prod_{c=2}^{n} w_c.$$

□

Proof of Theorem 14.5.1 . As in the previous lemma, let $\boldsymbol{L}_G = \boldsymbol{U}^T \boldsymbol{W} \boldsymbol{U}$ and $\boldsymbol{B} = \boldsymbol{W}^{1/2} \boldsymbol{U}$. So,

$$\begin{aligned}
\sigma_{n-1}(\boldsymbol{L}_G) &= \sigma_{n-1}(\boldsymbol{B}^T \boldsymbol{B}) \\
&= \sigma_{n-1}(\boldsymbol{B}\boldsymbol{B}^T) \\
&= \sum_{|S|=n-1, S \subseteq E} \sigma_{n-1}(\boldsymbol{B}(S,:)\boldsymbol{B}(S,:)^T) \quad \text{(by (14.3))} \\
&= \sum_{|S|=n-1, S \subseteq E} \sigma_{n-1}(\boldsymbol{B}(S,:)^T \boldsymbol{B}(S,:)) \\
&= \sum_{|S|=n-1, S \subseteq E} \sigma_{n-1}(\boldsymbol{L}_{G_S}),
\end{aligned}$$

where by G_S we mean the graph containing just the edges in S. As S contains $n-1$ edges, this graph is either disconnected or a tree. If it is disconnected, then its Laplacian has at least two zero eigenvalues and $\sigma_{n-1}(\boldsymbol{L}_{G_S}) = 0$. If it is a tree, we apply the previous lemma. Thus, the sum equals

$$\sum_{\text{spanning trees } T \subseteq E} \sigma_{n-1}(\boldsymbol{L}_{G_T}) = n \sum_{\text{spanning trees } T} \prod_{e \in T} w_e.$$

□

14.6 Leverage Scores and Marginal Probabilities

The *leverage score* of an edge, written ℓ_e is defined to be $w_e \mathrm{R}_{\text{eff}}(e)$. That is, the weight of the edge times the effective resistance between its endpoints. The leverage score serves as a measure of how important the edge is. For example, if removing an edge disconnects the graph, then $\mathrm{R}_{\text{eff}}(e) = 1/w_e$, as all current flowing between its endpoints must use the edge itself, and $\ell_e = 1$.

Consider sampling a random spanning tree with probability proportional to the product of the weights of its edges. We will now show that the probability that edge e appears in the tree is exactly its leverage score.

Theorem 14.6.1. *If we choose a spanning tree T with probability proportional to the product of its edge weights, then for every edge e*

$$Pr\left[e \in T\right] = \ell_e.$$

For simplicity, you might want to begin by thinking about the case where all edges have weight 1.

Recall that the effective resistance of edge $e = (a, b)$ is

$$(\delta_a - \delta_b)^T \boldsymbol{L}_G^+ (\delta_a - \delta_b),$$

and so

$$\ell_{a,b} = w_{a,b} (\delta_a - \delta_b)^T \boldsymbol{L}_G^+ (\delta_a - \delta_b).$$

We can write a matrix $\boldsymbol{\Gamma}$ that has all these terms on its diagonal by letting $\boldsymbol{U}$ be the edge-vertex adjacency matrix, $\boldsymbol{W}$ be the diagonal edge weight matrix, $\boldsymbol{B} = \boldsymbol{W}^{1/2}\boldsymbol{U}$, and setting

$$\boldsymbol{\Gamma} = \boldsymbol{B}\boldsymbol{L}_G^{+}\boldsymbol{B}^T.$$

The rows and columns of $\boldsymbol{\Gamma}$ are indexed by edges, and for each edge e,

$$\boldsymbol{\Gamma}(e,e) = \ell_e.$$

For off-diagonal entries corresponding to edges (a,b) and (c,d), we have

$$\boldsymbol{\Gamma}((a,b),(c,d)) = \sqrt{w_{a,b}}\sqrt{w_{c,d}}(\boldsymbol{\delta}_a - \boldsymbol{\delta}_b)^T\boldsymbol{L}_G^{+}(\boldsymbol{\delta}_c - \boldsymbol{\delta}_d).$$

Claim 14.6.2. *The matrix $\boldsymbol{\Gamma}$ is a symmetric projection matrix and has trace $n-1$.*

Proof. The matrix $\boldsymbol{\Gamma}$ is clearly symmetric. To show that it is a projection, it suffices to show that all of its eigenvalues are 0 or 1. This is true because, excluding the zero eigenvalues, $\boldsymbol{\Gamma}$ has the same eigenvalues as

$$\boldsymbol{L}_G^{+}\boldsymbol{B}^T\boldsymbol{B} = \boldsymbol{L}_G^{+}\boldsymbol{L}_G = \boldsymbol{\Pi},$$

where $\boldsymbol{\Pi}$ is the projection orthogonal to the all 1 vector. As $\boldsymbol{\Pi}$ has $n-1$ eigenvalues that are 1, so does $\boldsymbol{\Gamma}$. $\square$

As the trace of $\boldsymbol{\Gamma}$ is $n-1$, so is the sum of the leverage scores:

$$\sum_e \ell_e = n-1.$$

This is a good sanity check on Theorem 14.6.1: every spanning tree has $n-1$ edges, and thus the probabilities that each edge is in the tree must sum to $n-1$.

We also obtain another formula for the leverage score. As a symmetric projection is its own square,

$$\boldsymbol{\Gamma}(e,e) = \boldsymbol{\Gamma}(e,:)\boldsymbol{\Gamma}(e,:)^T = \|\boldsymbol{\Gamma}(e,:)\|^2.$$

This is the formula I introduced in Section 14.2. If we flow 1 unit from a to b, the potential difference between c and d is $(\boldsymbol{\delta}_a - \boldsymbol{\delta}_b)^T\boldsymbol{L}_G^{+}(\boldsymbol{\delta}_c - \boldsymbol{\delta}_d)$. If we plug these potentials into the Laplacian quadratic form, we obtain the effective resistance. Thus this formula says

$$w_{a,b}\mathrm{R_{eff}}_{a,b} = w_{a,b}\sum_{(c,d)\in E} w_{c,d}\left((\boldsymbol{\delta}_a - \boldsymbol{\delta}_b)^T\boldsymbol{L}_G^{+}(\boldsymbol{\delta}_c - \boldsymbol{\delta}_d)\right)^2.$$

Proof of Theorem 14.6.1. Let $\mathrm{Span}(G)$ denote the set of spanning trees of G. For an edge e,

$$\begin{aligned}\Pr_T[e \in T] &= \sum_{T\in\mathrm{Span}(G):e\in T} \frac{\sigma_{n-1}(\boldsymbol{L}_{G_T})}{\sigma_{n-1}(\boldsymbol{L}_G)} \\ &= \sum_{T\in\mathrm{Span}(G):e\in T} \sigma_{n-1}(\boldsymbol{L}_{G_T})\sigma_{n-1}(\boldsymbol{L}_G^{+}) \\ &= \sum_{T\in\mathrm{Span}(G):e\in T} \sigma_{n-1}(\boldsymbol{L}_{G_T}\boldsymbol{L}_G^{+}),\end{aligned}$$

by (14.4). Recalling that the subsets of $n-1$ edges that are not spanning trees contribute 0 allows us to re-write this sum as

$$\sum_{|S|=n-1, e \in S} \sigma_{n-1}(\boldsymbol{L}_{G_S} \boldsymbol{L}_G^+).$$

To evaluate the terms in the sum, we compute

$$\begin{aligned}
\sigma_{n-1}(\boldsymbol{L}_{G_S} \boldsymbol{L}_G^+) &= \sigma_{n-1}(\boldsymbol{B}(:,S)\boldsymbol{B}(:,S)^T \boldsymbol{L}_G^+) \\
&= \sigma_{n-1}(\boldsymbol{B}(:,S)^T \boldsymbol{L}_G^+ \boldsymbol{B}(:,S)) \\
&= \sigma_{n-1}(\boldsymbol{\Gamma}(S,S)) \\
&= \sigma_{n-1}(\boldsymbol{\Gamma}(S,:)\boldsymbol{\Gamma}(:,S)).
\end{aligned}$$

Let $\boldsymbol{\gamma}_e = \boldsymbol{\Gamma}(e,:)$ and let $\boldsymbol{\Pi}_{\boldsymbol{\gamma}_e}$ denote the projection orthogonal to $\boldsymbol{\gamma}_e$. As $e \in S$, we have

$$\sigma_{n-1}(\boldsymbol{\Gamma}(S,:)\boldsymbol{\Gamma}(:,S)) = \|\boldsymbol{\gamma}_e\|^2 \sigma_{n-2}(\boldsymbol{\Gamma}(S,:)\boldsymbol{\Pi}_{\boldsymbol{\gamma}_e}\boldsymbol{\Gamma}(:,S)) = \|\boldsymbol{\gamma}_e\|^2 \sigma_{n-2}((\boldsymbol{\Gamma}\boldsymbol{\Pi}_{\boldsymbol{\gamma}_e}\boldsymbol{\Gamma})(S,S)).$$

As $\boldsymbol{\gamma}_e$ is in the span on $\boldsymbol{\Gamma}$, the matrix $\boldsymbol{\Gamma}\boldsymbol{\Pi}_{\boldsymbol{\gamma}_e}\boldsymbol{\Gamma}$ is a symmetric projection onto an $n-2$ dimensional space, and so

$$\sigma_{n-2}(\boldsymbol{\Gamma}\boldsymbol{\Pi}_{\boldsymbol{\gamma}_e}\boldsymbol{\Gamma}) = 1.$$

To exploit this identity, we return to our summation:

$$\begin{aligned}
\sum_{|S|=n-1, e \in S} \sigma_{n-1}(\boldsymbol{L}_{G_S} \boldsymbol{L}_G^+) &= \sum_{|S|=n-1, e \in S} \|\boldsymbol{\gamma}_e\|^2 \sigma_{n-2}((\boldsymbol{\Gamma}\boldsymbol{\Pi}_{\boldsymbol{\gamma}_e}\boldsymbol{\Gamma})(S,S)) \\
&= \|\boldsymbol{\gamma}_e\|^2 \sum_{|S|=n-1, e \in S} \sigma_{n-2}((\boldsymbol{\Gamma}\boldsymbol{\Pi}_{\boldsymbol{\gamma}_e}\boldsymbol{\Gamma})(S,S)) \\
&= \|\boldsymbol{\gamma}_e\|^2 \sigma_{n-2}(\boldsymbol{\Gamma}\boldsymbol{\Pi}_{\boldsymbol{\gamma}_e}\boldsymbol{\Gamma}) \\
&= \|\boldsymbol{\gamma}_e\|^2 \\
&= \ell_e.
\end{aligned}$$

□

14.7 Quickly estimating effective resistances

We can compute $\mathrm{R_{eff}}(a,b)$ by solving a system of equations in $\boldsymbol{L}$. We know how to solve such systems of linear equations to high accuracy in time nearly linear in the number of nonzero entries in $\boldsymbol{L}$ [**?**]. But, what if we want to know the effective resistance of every edge or between every pair of vertices?

We will see that we can do this by solving on $O(\log n)$ systems of equations in $\boldsymbol{L}$. The reason is that the effective resistances are the squares of Euclidean distances:

$$\mathrm{R_{eff}}(a,b) = \left\| \boldsymbol{L}_G^{+/2}(\boldsymbol{\delta}_a - \boldsymbol{\delta}_b) \right\|^2 = \left\| \boldsymbol{L}_G^{+/2}\boldsymbol{\delta}_a - \boldsymbol{L}_G^{+/2}\boldsymbol{\delta}_a \right\|^2.$$

The reason is that we can exploit the Johnson-Lindenstrauss Theorem [**?**].

Theorem 14.7.1. *[Johnson Lindenstrauss] Let $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ be vectors in an m dimensional vector space. Let $\boldsymbol{R}$ be a d-by-m matrix of independent Gaussian random variables of variance $1/d$. If*

$$d \geq \frac{8}{\delta^2} \ln(n/\epsilon),$$

then with probability at most ϵ for all $i \neq j$,

$$1 - \delta \leq \frac{\|\boldsymbol{R}\boldsymbol{v}_i - \boldsymbol{R}\boldsymbol{v}_j\|}{\|\boldsymbol{v}_i - \boldsymbol{v}_j\|} \leq 1 + \delta.$$

That is, the distances between all pairs of vectors $\boldsymbol{R}\boldsymbol{v}_i$ are approximately the same as between the vectors $\boldsymbol{v}_i$.

One can prove this by using tail bounds on χ-square random variables. I'll include a proof in the Appendix.

Here's one way we could try to use this. If we want to estimate all effective resistances to within error δ, with probability at least $1 - \epsilon$, we set

$$d = \left\lceil \frac{8}{\delta^2} \ln(n/\epsilon) \right\rceil,$$

choose $\boldsymbol{R}$ to be a d-by-n matrix of independent random Gaussians, and then compute

$$\boldsymbol{R}\boldsymbol{L}^{+/2}.$$

This requires solving d systems of linear equations in $\boldsymbol{L}^{1/2}$.

But, that is not quite the same as solving systems in $\boldsymbol{L}$. To turn this into a problem of solving systems in $\boldsymbol{L}$, we exploit a slightly different formula for effective resistances. As before, write $\boldsymbol{L} = \boldsymbol{U}^T \boldsymbol{W} \boldsymbol{U}$. We then have

$$\boldsymbol{L}^+ \boldsymbol{U}^T \boldsymbol{W}^{1/2} \boldsymbol{W}^{1/2} \boldsymbol{U} \boldsymbol{L}^+ = \boldsymbol{L}^+ \boldsymbol{L} \boldsymbol{L}^+ = \boldsymbol{L}^+.$$

So,

$$\left\| \boldsymbol{W}^{1/2} \boldsymbol{U} \boldsymbol{L}^+ (\boldsymbol{\delta}_a - \boldsymbol{\delta}_b) \right\|^2 = \mathrm{R_{eff}}(a,b).$$

Now, we let $\boldsymbol{R}$ be a d-by-$|E|$ matrix of random Gaussians of variance $1/d$, and compute

$$\boldsymbol{R} \boldsymbol{W}^{1/2} \boldsymbol{U} \boldsymbol{L}^+ = (\boldsymbol{R} \boldsymbol{W}^{1/2} \boldsymbol{U}) \boldsymbol{L}^+.$$

This requires solving d systems of linear equations in $\boldsymbol{L}$. We then set

$$\boldsymbol{v}_a = (\boldsymbol{R} \boldsymbol{W}^{1/2} \boldsymbol{U}) \boldsymbol{L}^+ \boldsymbol{\delta}_a.$$

Each of these is a vector in d dimensions, and with high probability $\|\boldsymbol{v}_a - \boldsymbol{v}_b\|^2$ is a good approximation of $\mathrm{R_{eff}}(a,b)$.

14.8 Monotonicity

Rayleigh's Monotonicity Principle tells us that if we alter the spring network by decreasing some of the spring constants, then the effective spring constant between s and t will not increase. In terms of effective resistance, this says that if we increase the resistance of some resistors then the effective resistance can not decrease. This sounds obvious. But, it is in fact a very special property of linear elements like springs and resistors.

Theorem 14.8.1. *Let $G = (V, E, w)$ be a weighted graph and let $\widehat{G} = (V, E, \widehat{w})$ be another weighted graph with the same edges and such that*

$$\widehat{w}_{a,b} \leq w_{a,b}$$

for all $(a, b) \in E$. For vertices s and t, let $c_{s,t}$ be the effective spring constant between s and t in G and let $\widehat{c}_{s,t}$ be the analogous quantity in $\widehat{G}$. Then,

$$\widehat{c}_{s,t} \leq c_{s,t}.$$

Proof. Let $\boldsymbol{x}$ be the vector of minimum energy in G such that $\boldsymbol{x}(s) = 0$ and $\boldsymbol{x}(t) = 1$. Then, the energy of $\boldsymbol{x}$ in $\widehat{G}$ is no greater:

$$\frac{1}{2} \sum_{(a,b) \in E} \widehat{w}_{a,b} (\boldsymbol{x}(a) - \boldsymbol{x}(b))^2 \leq \frac{1}{2} \sum_{(a,b) \in E} w_{a,b} (\boldsymbol{x}(a) - \boldsymbol{x}(b))^2 = c_{s,t}.$$

So, the minimum energy of a vector $\boldsymbol{x}$ in $\widehat{G}$ such that $\boldsymbol{x}(s) = 0$ and $\boldsymbol{x}(t) = 1$ will be at most $c_{s,t}$, and so $\widehat{c}_{s,t} \leq c_{s,t}$. $\square$

While this principle seems very simple and intuitively obvious, it turns out to fail in just slightly more complicated situations.

14.9 Notes

A Proof of Johnson Lindenstrauss