Problem Set 7
CME 305

Dan Johnson
dansj

March 7, 2019

Problem 1

First, we show that $\vec{p}$ is unique. We write the PageRank expression in vector form.

$$\vec{p} = \frac{1-\alpha}{|V|}\vec{1} + \alpha W \vec{p}$$

$$(I - \alpha A D^{-1})DD^{-1}\vec{p} = \frac{1-\alpha}{|V|}\vec{1}$$

$$(D - \alpha A)D^{-1}\vec{p} = \frac{1-\alpha}{|V|}\vec{1}$$

If we can show that $(D - \alpha A)$ is invertible, then we can multiply both sides by $D(D-\alpha A)^{-1}$, which will show that $\vec{p}$ has one solution and is therefore unique. $(D - \alpha A)$ is nonsingular and invertible if it is strictly diagonally dominant. A matrix B is strictly diagonally dominant if $|B_{ii}| > \sum_{j \neq i}|B_{ij}|, \forall i$. In our case, we need to show that $|D_{ii}| > \sum_{j \neq i}|-\alpha A_{ij}|, \forall i$. By definition each $D_{ii} = \sum_{j \neq i}A_{ij}$.

$$|D_{ii}| > \sum_{j \neq i}|-\alpha A_{ij}|$$

$$\left|\sum_{j \neq i}A_{ij}\right| > \sum_{j \neq i}|-\alpha A_{ij}|$$

$$\sum_{j \neq i}A_{ij} > \sum_{j \neq i}\alpha A_{ij}$$

$\alpha \in (0,1)$, so this is clearly true. Therefore, $(D - \alpha A)$ is invertible which means that $\vec{p}$ has one solution and is therefore unique. ∎

Now we look to compute $\vec{p}$. We construct a new graph $G' = (V', E')$ so that we can solve a Laplacian system. Let $V' = V \cup \{v\}$. Let $E' = \{e$ with weight $\alpha w_e | \forall e \in$

$E\} \cup \{\{u, v\}$ with weight $(1 - \alpha) \deg_G(u) | \forall u \in V\}$. We now construct D', A', and $\mathcal{L}'$.

$$D' = \begin{pmatrix} \alpha D + (1 - \alpha)D & \vec{0} \\ \vec{0}^T & (1 - \alpha)\vec{1}^T D \vec{1} \end{pmatrix}$$

$$A' = \begin{pmatrix} \alpha A & (1 - \alpha)D\vec{1} \\ (1 - \alpha)(D\vec{1})^T & 0 \end{pmatrix}$$

$$\mathcal{L}' = \begin{pmatrix} D - \alpha A & -(1 - \alpha)D\vec{1} \\ -(1 - \alpha)(D\vec{1})^T & (1 - \alpha)\vec{1}^T D \vec{1} \end{pmatrix}$$

We now want to solve a Laplacian system that computes $\vec{p}$. We note that G' is fully connected by construction so $\mathcal{L}'\vec{1} = 0\vec{1}$. We construct $\vec{x}'$ and $\vec{b}'$ so that solving the Laplacian system $\mathcal{L}'\vec{x}' = \vec{b}'$ solves for $\vec{p}$.

$$\vec{x}' = \begin{pmatrix} D^{-1}\vec{p} \\ 1 \end{pmatrix}$$

$$\vec{b}' = \begin{pmatrix} \frac{1-\alpha}{|V|}\vec{1} - (1 - \alpha)D\vec{1} \\ -(1 - \alpha)(D\vec{1})^T D^{-1}\vec{p} + (1 - \alpha)\vec{1}^T D\vec{1} \end{pmatrix}$$

$$\vec{b}' = \begin{pmatrix} \frac{1-\alpha}{|V|}\vec{1} + (\alpha - 1)D\vec{1} \\ (\alpha - 1)\vec{1}^T\vec{p} + (1 - \alpha)\vec{1}^T D\vec{1} \end{pmatrix}$$

$$\vec{b}' = \begin{pmatrix} \frac{1-\alpha}{|V|}\vec{1} + (\alpha - 1)D\vec{1} \\ \alpha - 1 + (1 - \alpha)\vec{1}^T D\vec{1} \end{pmatrix}$$

So, now we have our Laplacian system $\mathcal{L}'\vec{x}' = \vec{b}'$. The system is solved in $O(\mathcal{T})$ time and gives us back some $\vec{x}''$. Since $\mathcal{L}'\vec{1} = 0\vec{1}$, we notice that our target solution $\vec{x}' = \vec{x}'' + c\vec{1}$. So, we solve for c knowing that the last component of $\vec{x}'$ is 1 and using the last component of the solution to the Laplacian $\vec{x}''$.

$$1 = [\vec{x}'']_{n+1} + c$$

We then get $\vec{x}'$ easily. The first n components of this vector are $D^{-1}\vec{p}$. So, we get $\vec{p} = D\vec{x}'$. So, the algorithm constructs G' and then $\mathcal{L}$ and $\vec{b}'$ all in $O(|V| + |E|)$ time. It then solves this Laplacian system in $O(\mathcal{T})$ time and gets $\vec{p}$ in $O(|V| + |E|)$ time. Therefore, the total run time of this algorithm is $O(|V| + |E| + \mathcal{T})$. ∎

Problem 2

$$M = D^{-1/2}\tilde{W}D^{1/2}$$
$$= D^{-1/2}\frac{1}{2}(I + AD^{-1})D^{1/2}$$
$$= \frac{1}{2}(D^{-1/2}D^{1/2} + D^{-1/2}AD^{-1/2})$$
$$= \frac{1}{2}(I + D^{-1/2}AD^{-1/2})$$

This matrix is clearly symmetric because I, $D^{-1/2}$, and A are symmetric. We now use the

normalized Laplacian

$$M = \frac{1}{2}(I + D^{-1/2}AD^{-1/2})$$
$$= \frac{1}{2}(I + I - N(G))$$
$$= I - \frac{1}{2}N(G)$$

we find the relationship between eigenvalues of N and M.

$$N(G)\vec{v}_i = \lambda_i(N(G))\vec{v}_i$$
$$\frac{1}{2}N(G)\vec{v}_i = \frac{1}{2}\lambda_i(N(G))\vec{v}_i$$
$$\vec{v}_i - \frac{1}{2}N(G)\vec{v}_i = \vec{v}_i - \frac{1}{2}\lambda_i(N(G))\vec{v}_i$$
$$\left(I - \frac{1}{2}N(G)\right)\vec{v}_i = \left(1 - \frac{1}{2}\lambda_i(N(G))\right)\vec{v}_i$$
$$M\vec{v}_i = \left(1 - \frac{1}{2}\lambda_i(N(G))\right)\vec{v}_i$$
$$\lambda_i(M) = 1 - \frac{1}{2}\lambda_i(N)$$

The largest eigenvalue for M is the smallest for N. From the notes, we know that $\lambda_1(N) = 0$, so $\lambda_{\max}(M) = 1$. The second smallest eigenvalue of N is $\lambda_2(N)$, so the second largest eigenvalue of M is $\lambda_{n-1}(M) = 1 - \frac{1}{2}\lambda_2(N)$. ∎

Problem 3

First, we want to get $\tilde{W}^k$ into an easier form. Let $M = D^{-1/2}\tilde{W}D^{1/2}$ like in problem 2.

$$\tilde{W} = D^{1/2}MD^{-1/2}$$
$$\tilde{W}^k = D^{1/2}M^kD^{-1/2}$$

Let $x = p - \frac{1}{||d||_1}d$.

$$\left|\left|\tilde{W}^k p - \frac{1}{||d||_1}d\right|\right|_1 = \left|\left|\tilde{W}^k p - \frac{1}{||d||_1}\tilde{W}^k d\right|\right|_1$$
$$= \left|\left|\tilde{W}^k x\right|\right|_1$$

why does $\tilde{W}^k d = d$?

$$\tilde{W}d = \frac{1}{2}(I + AD^{-1})d$$
$$= \frac{1}{2}(d + A\vec{1})$$
$$= \frac{1}{2}(d + d)$$
$$= d$$
$$\downarrow$$
$$\tilde{W}^k d = d$$

We now look back at the LHS.

$$\left\|\tilde{W}^k x\right\|_1 \le \sqrt{n}\left\|\tilde{W}^k x\right\|_2 = \sqrt{n}\left\|D^{1/2}M^k D^{-1/2}x\right\|_2 \le \sqrt{n}\left\|D^{1/2}\right\|_2\left\|M^k D^{-1/2}x\right\|_2$$

$$\left\|\tilde{W}^k p - \frac{1}{||d||_1}d\right\|_1 \le \sqrt{n}\sqrt{d_{\max}}\left\|M^k D^{-1/2}x\right\|_2$$

From problem 2, we know the largest eigenvalue of M is 1.

$$\begin{aligned}
M\sqrt{d} &= \frac{1}{2}(I + D^{-1/2}AD^{-1/2})\sqrt{d} \\
&= \frac{1}{2}(\sqrt{d} + D^{-1/2}A\vec{1}) \\
&= \frac{1}{2}(\sqrt{d} + D^{-1/2}d) \\
&= \frac{1}{2}(\sqrt{d} + \sqrt{d}) \\
&= \sqrt{d}
\end{aligned}$$

So, we know that $\sqrt{d}$ is the eigenvector of the largest eigenvalue. If $\sqrt{d} \perp D^{-1/2}x$, then we can bound the 2 norm with the second largest eigenvalue.

$$\begin{aligned}
\sqrt{d}^T(D^{-1/2}x) &= \vec{1}^T\left(p - \frac{1}{||d||_1}d\right) \\
&= 1 - 1 \\
&= 0
\end{aligned}$$

The first term is zero because the sum of probabilities is 1 and the second term is 1 by definition. So, we can bound $\left\|M^k D^{-1/2}x\right\|_2$ using the second largest eigenvalue: $1 - \frac{1}{2}\lambda_2(N)$.

$$\left\|\tilde{W}^k p - \frac{1}{||d||_1}d\right\|_1 \le \sqrt{n}\sqrt{d_{\max}}\left\|M^k D^{-1/2}x\right\|_2 \le \sqrt{n}\sqrt{d_{\max}}\left\|(1 - \frac{1}{2}\lambda_2(N))^k D^{-1/2}x\right\|_2$$

$$\left\|\tilde{W}^k p - \frac{1}{||d||_1}d\right\|_1 \le \sqrt{n}\sqrt{d_{\max}}(1 - \frac{1}{2}\lambda_2(N))^k\left\|D^{-1/2}\right\|_2\left\|x\right\|_2$$

Bound $\left\|D^{-1/2}\right\|_2$, $||x||_2$, and $(1 - \frac{1}{2}\lambda_2(N))^k$.

$$\left\|D^{-1/2}\right\|_2 \le \frac{1}{\sqrt{d_{\min}}}$$

$$||x||_2 = \left\|p - \frac{1}{||d||_1}d\right\|_2 \le 1 + 1 = 2$$

$$(1 - \frac{1}{2}\lambda_2(N))^k \le \exp\left(-\frac{k}{2}\lambda_2(N)\right)$$

Now we can finally bound it.

$$\left\| \tilde{W}^k p - \frac{1}{||d||_1} d \right\|_1 \leq \sqrt{n} \sqrt{d_{\max}} \exp\left(-\frac{k}{2}\lambda_2(N)\right) \frac{1}{\sqrt{d_{\min}}} 2 \leq \epsilon$$

$$\left\| \tilde{W}^k p - \frac{1}{||d||_1} d \right\|_1 \leq \sqrt{n} \sqrt{d_{\max}} \exp\left(-\frac{c}{2}\log(\frac{nd_{\max}}{d_{\min}})\right) \frac{1}{\sqrt{d_{\min}}} 2 \leq \epsilon$$

$$\left\| \tilde{W}^k p - \frac{1}{||d||_1} d \right\|_1 \leq \sqrt{n} \sqrt{d_{\max}} \exp(c) \frac{\sqrt{d_{\min}}}{\sqrt{n}\sqrt{d_{\max}}} \frac{1}{\sqrt{d_{\min}}} 2 \leq \epsilon$$

$$\left\| \tilde{W}^k p - \frac{1}{||d||_1} d \right\|_1 \leq \exp(c) 2 \leq \epsilon$$

$$\text{Suppose } c = \log(1/4)$$

$$\exp(\log(1/4))2 = 1/2 \leq \epsilon$$

So, we can clearly pick a c such that the LHS is less than or equal to ϵ. We can always pick a c such that this is true given a ϵ.

$$c \leq \log\left(\frac{\epsilon}{2}\right)$$

$$\text{Example: } c = \log\left(\frac{\epsilon}{2}\right) - 1$$

Therefore, we have shown that k is of the given form. ∎

Problem 4

a. We think about this problem using effective resistance. We know what the commute time from a to b is from lemma 3 in lecture 14.

$$C_{ab} = ||\vec{d}||_1 (\vec{1}_a - \vec{1}_b)^T \mathcal{L}^\dagger (\vec{1}_a - \vec{1}_b)$$

We also know the effective resistance between a and b from the same lecture.

$$\tilde{R}_{ab} = \sum_{e \in E} r_e f_e^2 = (\vec{1}_a - \vec{1}_b)^T \mathcal{L}^\dagger (\vec{1}_a - \vec{1}_b)$$

So, we know that the commute time between a and b in terms of the effective resistance between a and b.

$$C_{ab} = \left(\sum_{v \in V} \deg(v)\right) \tilde{R}_{ab}$$

We know that the size of the min cut separating a and b is 1, so the only path between a and b is the edge $\{a, b\}$. The effective resistance $\tilde{R}_{ab}$ is therefore just the resistance between a and b which is 1. We also know that $\sum_{v \in V} \deg(v) = 2m$ because every edge is counted twice, once at each endpoint.

$$C_{ab} = (2m)(1) = 2m$$

b. Again we have the commute time in terms of effective resistance.

$$C_{ab} = \left(\sum_{v \in V} \deg(v)\right) \tilde{R}_{ab} = 2m\tilde{R}_{ab}$$

we want to show that this is upper-bounded by $2md(a,b)$.

$$C_{ab} \leq 2md(a,b)$$
$$2m\tilde{R}_{ab} \leq 2md(a,b)$$
$$\tilde{R}_{ab} \leq d(a,b)$$

When we find the effective resistance, we have to consider all paths between a and b. If all paths use the same edge, the edges are in series and the resistance is added normally. If the edges are used in some of the paths, the edges are in parallel, so the effective resistance is defined as $\frac{1}{\tilde{R}_{ab}} = \frac{1}{R_1} + \frac{1}{R_2}$. In our case, $R = 1$ for all edges. This means that the largest effective resistance is if all the edges between a and b are in series or, equivalently, there is just one path between the two nodes. In this case, every edge in the path would sum to $d(a,b)$, which means $C_{ab} = 2md(a,b)$. Now if we have any other path between the two nodes, they will be in parallel.

$$\frac{1}{\tilde{R}_{ab}} = \frac{1}{R_1} + \frac{1}{R_2}$$

Clearly, $\tilde{R}_{ab}$ is largest when the effective resistance of the second path is infinitely large which can be interpreted as no path existing. Since adding any path lowers the effective resistance, the commute time must be upper-bounded by the case of a single shortest path.

$$C_{ab} \leq 2md(a,b)$$

Problem 5

a. If G is a tree, there is only one path between a and b. Otherwise, this induces a cycle and G is not a tree. As argued in problem 4 part b, in the case where there is only one path between a and b, the effective resistance is the sum of all the weights of this path. With each weight being one, the effective resistance $\tilde{R}_{ab}$ becomes $d(a,b)$.

$$C_{ab} = \left(\sum_{v \in V} \deg(v) \right) \tilde{R}_{ab} = 2md(a,b)$$

b. First, we note the relationship between hit time and commute time.

$$C_{ab} = H_{ab} + H_{ba}$$

I state that the max commute time in a graph is larger than the max hit time. We easily prove this by taking the nodes from the max hit time and see that the commute time for these nodes is the max hit time plus the time it takes to get back to the first node. In the lecture notes, we found an upper

bound on the cover time.

$$Cover(G) \leq \left(\max_{a,b \in V} H_{ab} \right) O(\log n)$$
$$\leq \left(\max_{a,b \in V} C_{ab} \right) O(\log n)$$
$$\leq 2md(a,b)O(\log n)$$
$$\leq 2mnO(\log n)$$
$$\leq O(mn \log n)$$

In the third line, we use the upper bound found in problem 4 part b. In the forth line we note that $d(a,b) < n$.