UBC CPSC 536N: Sparse Approximations Winter 2013

Lecture 16 — March 6, 2013

Prof. Nick Harvey *Scribe: Zachary Drudi*

Recall from last time that our goal is to prove the following theorem. Similar results were first announced by Goemans (unpublished, 2012).

Theorem 0.1. Let $G = (V, E)$ be a connected graph, let $n = |V|$ and assume $|E| \geq 3$. Let P be the spanning tree polytope of G, and let $x \in P$. For $e \in E$, let $w_e \in \mathbb{R}^{n-1}$ be such that

$$\sum_{e \in E} x_e w_e w_e^T = I$$

Then there exists a set of edges F with $|F| \geq \frac{n}{2}$ such that

$$\lambda_{\max}\Big(\sum_{e \in F} w_e w_e^T\Big) \leq 35,$$

and (V, F) is acyclic.

Corollary 0.2. Let L_x be the weighted Laplacian of the fractional spanning tree x. There exists $F \subset E$ with $|F| \geq \frac{n}{2}$ such that F is acyclic and $L_F \preceq 35 \cdot L_x$, where L_F is the Laplacian of the forest F.

We give an algorithm to produce this thin forest F below. We will prove its correctness, thus proving the theorem.

Algorithm 0.1 Thin Forest Algorithm

Require: A graph $G = (V, E)$.
Ensure: A forest $F \subset E$ with $\lambda_{\max}\big(\sum_{e \in F} w_e w_e^T\big) \leq 35$ and $|F| \geq \frac{n}{2}$.

$A \leftarrow 0$
$F \leftarrow \emptyset$
$u \leftarrow u_0 := 20$
$\delta := \frac{20}{n-1}$
$\Phi^u(A) := \mathrm{trace}((uI - A)^{-1}) = \sum_{i=1}^{n-1} (u - \lambda_i)^{-1}$, where $\{\lambda_i\}_{i=1}^{n-1}$ are the eigenvalues of A.
for $j = 1 \ldots n/2$ **do**
  Find a good edge e
  $F \leftarrow F \cup \{e\}$
  $A \leftarrow A + w_e w_e^T$
  $u \leftarrow u + \delta$
end for

Given $F \subset E$, we say an edge $e \in E$ is *good* if:

(a) $F \cup \{e\}$ is acyclic and $e \notin F$

(b) $\lambda_{\max}(A + w_e w_e^T) < u + \delta$

(c) $\Phi^{u+\delta}(A + w_e w_e^T) \leq \Phi^u(A)$

We will show that the for loop maintains the following invariants:

(a) F is acyclic

(b) $\lambda_{\max}(A) < u$

(c) $\Phi^u(A) \leq \frac{1}{\delta}$

We will need the Sherman-Morrison Formula to prove the correctness of the algorithm.

Theorem (Sherman-Morrison Formula). Let B be an $n \times n$ nonsingular matrix, and $a, b \in \mathbb{R}^n$. Then $(B + ab^T)^{-1}$ exists iff $1 \neq -b^T B^{-1} a$, and in that case

$$(B + ab^T)^{-1} \;=\; B^{-1} - \frac{B^{-1} a b^T B^{-1}}{1 + b^T B^{-1} a}.$$

1 Correctness of the Algorithm

Assume $\lambda_{\max}(A) < u$. Define

$$\begin{aligned} M &:= ((u+\delta)I - A)^{-1} \\ N &:= \frac{M^2}{\Phi^u(A) - \Phi^{u+\delta}(A)} + M \end{aligned}$$

Observation 1: $(u+\delta)I - A$ is invertible.

Observation 2: $\Phi^{u+\delta}(A) < \Phi^u(A)$

Observation 3: M is a positive definite matrix.

Observation 4: $M \prec N$, i.e. $v^T M v < v^T N v$ for all non-zero $v \in \mathbb{R}^{n-1}$.

Proof. (1) As $\lambda_{\max}(A) < u$, in particular $u+\delta$ is not an eigenvalue of A. Therefore $(u+\delta)I - A$ has trivial kernel, and so is invertible.

(2) Since $\lambda_i < u$ and $\delta > 0$, we have $0 < u - \lambda_i < u + \delta - \lambda_i$.
Thus $\Phi^{u+\delta}(A) = \sum_i \big((u+\delta) - \lambda_i\big)^{-1} < \sum_i (u - \lambda_i)^{-1} = \Phi^u(A)$

(3) Every eigenvalue of M is strictly postive, so M is positive definite.

(4) Consider $N - M = \frac{M^2}{\Phi^u(A) - \Phi^{u+\delta}(A)}$
As the denominator is positive by (2), and M is positive definite, we have that $N - M$ is positive definite, and so $M \prec N$.

□

Lemma 1.1. Suppose $\lambda_{\max} < u$. Let $v \in \mathbb{R}^{n-1}$ and $t > 0$ be arbitrary. If $v^T N v \leq \frac{1}{t}$, then $\Phi^{u+\delta}(A + tvv^T) \leq \Phi^u(A)$ and $\lambda_{\max}(A + tvv^T) < u + \delta$.

Proof. We will apply the Sherman-Morrison formula with $B = (u+\delta)I - A$, $a = -tv$ and $b = v$. This is justified because, assuming $v \neq 0$,

$$-b^T B^{-1} a \;=\; t \cdot v^T M v \;<\; t \cdot v^T N v \;\leq\; 1,$$

by Observation 4 and the hypothesis of the lemma. So

$$\begin{aligned}
\Phi^{u+\delta}(A + tvv^T) &= \operatorname{trace}[((u+\delta)I - A - tvv^T)^{-1}] \\
&= \operatorname{trace}\left(M - \frac{M(-tv)v^T M}{1 + v^T M(-tv)}\right) \\
&= \operatorname{trace}(M) + \operatorname{trace}\left(\frac{tMvv^T M}{1 - tv^T M v}\right) \\
&= \Phi^{u+\delta}(A) + \frac{t \cdot \operatorname{trace}(v^T M M v)}{1 - tv^T M v} \\
&= \Phi^{u+\delta}(A) + \frac{v^T M^2 v}{\frac{1}{t} - v^T M v} \\
&= \Phi^{u}(A) - (\Phi^{u}(A) - \Phi^{u+\delta}(A)) + \frac{v^T M^2 v}{\frac{1}{t} - v^T M v},
\end{aligned}$$

where the fourth equality follows by applying the identity $\operatorname{trace}(AB) = \operatorname{trace}(BA)$.

So

$$\begin{aligned}
& \Phi^{u+\delta}(A + tvv^T) \leq \Phi^{u}(A) \\
\iff & -(\Phi^{u}(A) - \Phi^{u+\delta}(A)) + \frac{v^T M^2 v}{\frac{1}{t} - v^T M v} \leq 0 \\
\iff & \frac{v^T M^2 v}{\Phi^{u}(A) - \Phi^{u+\delta}(A)} + v^T M v \leq \frac{1}{t} \\
\iff & v^T N v \leq \frac{1}{t}
\end{aligned}$$

To establish the bound on $\lambda_{\max}(A + tvv^T)$, we collect some facts about the spectral norm. Define

$$\|B\| := \max\{\|Bx\| \mid \|x\| \leq 1\}$$

We establish the triangle inequality for $\|\cdot\|$:

$$\begin{aligned}
\|B + C\| &= \max\{\|(B+C)x\| \mid \|x\| \leq 1\} \\
&\leq \max\{\|Bx\| + \|Cx\| \mid \|x\| \leq 1\} \\
&\leq \max\{\|Bx\| \mid \|x\| \leq 1\} + \max\{\|Cx\| \mid \|x\| \leq 1\} \\
&= \|B\| + \|C\|
\end{aligned}$$

Now, since $\|\cdot\|$ satisfies the triangle inequality, we get the inequality $|\,\|B\| - \|C\|\,| \leq \|B - C\|$, from which continuity of $\|\cdot\|$ follows.

Consider $\phi(t) := \|A + tvv^T\|$. For a symmetric, positive definite matrix B, $\|B\| = |\lambda_{\max}(B)| = \lambda_{\max}(B)$, so $\phi(t) = \lambda_{\max}(A + tvv^T)$. Furthermore, since ϕ is a composition of the continuous functions $t \to A + tvv^T$ and $\|\cdot\|$, ϕ is continuous.

If $\phi(t_0) > u + \delta$ for some $v^T N v \le \frac{1}{t_0}$, then since $\phi(0) = \lambda_{\max}(A) < u < u + \delta$, there exists some $0 < t_1 < t_0$ such that $\phi(t_1) = u + \delta$. So $(u+\delta)I - (A + t_1 vv^T)$ is not an invertible matrix. However, by the Sherman-Morrison formula we have

$$\big((u+\delta)I - (A + t_1 vv^T)\big)^{-1} = M - \frac{M(-t_1 v)v^T M}{1 + v^T M(-t_1 v)}$$

The denominator is non-zero since $t_1 v^T M v < t_0 v^T N v \le 1$, so in particular $(u+\delta)I - (A + t_1 vv^T)$ is invertible. This is a contradiction.

$\square$

Lemma 1.2. $\operatorname{trace}(N) \le \frac{2}{\delta}$

Proof.

$$\begin{aligned}
\operatorname{trace}(N) &= \operatorname{trace}(M) + \operatorname{trace}\Big(\frac{M^2}{\Phi^u(A) - \Phi^{u+\delta}(A)}\Big) \\
&= \Phi^{u+\delta}(A) + \frac{\sum_{i=1}^{n-1}((u+\delta) - \lambda_i)^{-2}}{\sum_{i=1}^{n-1}\left[(u-\lambda_i)^{-1} - ((u+\delta) - \lambda_i)^{-1}\right]} \\
&\overset{(1)}{=} \Phi^{u+\delta}(A) + \frac{\sum_{i=1}^{n-1}((u+\delta) - \lambda_i)^{-2}}{\delta \sum_{i=1}^{n-1}\left[(u+\delta-\lambda_i)(u-\lambda_i)\right]^{-1}} \\
&\overset{(2)}{\le} \Phi^{u}(A) + \frac{\sum_{i=1}^{n-1}(u+\delta - \lambda_i)^{-2}}{\delta \sum_{i=1}^{n-1}\left[(u+\delta-\lambda_i)(u+\delta-\lambda_i)\right]^{-1}} \\
&= \Phi^u(A) + \frac{1}{\delta} \\
&\le \frac{2}{\delta}
\end{aligned}$$

We obtain (1) by using the identity $\frac{1}{a} - \frac{1}{a+b} = \frac{b}{a(a+b)}$ with $a = u - \lambda_i$ and $b = \delta$. For (2), we use Observation (2), and for the fraction, observe we are decreasing the denominator and thus increasing the fraction. $\square$

Now randomly pick an edge e with probability $\frac{x_e}{n-1}$. ($x \in P$, so $\sum_{e\in E} x_e = n-1$). Using the following two claims, we will establish that with non-zero probability e is a good edge.

Claim 1.3. $\Pr[w_e^T N w_e > 1] \le \frac{1}{10}$

Proof. Applying Markov's inequality, we have

$$\begin{aligned}
\Pr[w_e^T N w_e > 1] &\le \mathrm{E}[w_e^T N w_e] \\
&= \sum_{e\in E} \frac{x_e}{n-1} w_e^T N w_e \\
&= \frac{1}{n-1} \sum_{e\in E} x_e \cdot \mathrm{trace}(w_e^T N w_e) \\
&= \frac{1}{n-1} \sum_{e\in E} x_e \cdot \mathrm{trace}(N w_e w_e^T) \\
&= \frac{1}{n-1} \mathrm{trace}\Big(N \big(\textstyle\sum_{e\in E} x_e w_e w_e^T\big)\Big) \\
&= \frac{1}{n-1} \mathrm{trace}(N) \\
&\le \frac{1}{n-1} \frac{2}{\delta} \\
&= \frac{1}{n-1} \frac{2(n-1)}{20} \\
&= \frac{1}{10}
\end{aligned}$$

$\square$

Applying Lemma 1.1 with $t = 1$, if $w_e^T N w_e \le 1$ then $\lambda_{\max}(A + w_e w_e^T) < u + \delta$ and $\Phi^{u+\delta}(A + w_e w_e^T) \le \Phi^u(A)$, or in other words e satisfies conditions (b) and (c) of being a good edge. So $\Pr[e$ violates condition (b) or (c)$] \le \frac{1}{10}$.
We next check condition (a), that $F \cup \{e\}$ is acyclic and $e \notin F$.

Claim 1.4. $\Pr[F \cup \{e\}$ contains a cycle or $e \in F] \le \frac{3}{4}$

Proof. Let $C_1, C_2, \ldots, C_k \subset V$ be the components of F at iteration j. Initially there are n components, and each iteration decreases the number of components by 1. Since $j \le \frac{n}{2}$, $k \ge \frac{n}{2}$. Recall that $E[C_i]$ denotes the set of edges with both endpoints in C_i.

Let $R = \{e \in E : e \notin \bigcup_i^k E[C_i]\} = E \setminus \bigcup_i^k E[C_i]$.

Note: $e \in R \iff F \cup \{e\}$ is acyclic, and $e \notin F$. Since $x \in P$, the spanning tree polytope,

$$x(E[C_i]) \le |C_i| - 1 \qquad \forall i = 1, \ldots, k$$

$$\implies x\Big(\bigcup_{i=1}^k E[C_i]\Big) = \sum_{i=1}^k x(E[C_i]) \le \sum_{i=1}^k (|C_i| - 1) = n - k \le \frac{n}{2},$$

so $x(R) = x(E) - x(\bigcup_i E[C_i]) \ge n - 1 - \frac{n}{2} = \frac{n}{2} - 1$.

Thus

$$
\begin{aligned}
&\Pr[F \cup \{e\} \text{ is acyclic and } e \notin F] \\
&= \Pr[e \in R] \\
&= \frac{1}{n-1} x(R) \\
&\geq \frac{\frac{n}{2} - 1}{n-1} \\
&= \frac{1}{2}\left(\frac{n-1}{n-1} - \frac{1}{n-1}\right) \\
&= \frac{1}{2}(1 - \frac{1}{n-1}) \\
&\geq \frac{1}{4},
\end{aligned}
$$

assuming $n \geq 3$.

$\square$

So applying a union bound we get

$$
\begin{aligned}
&\Pr[e \text{ is not good}] \\
&= \Pr[e \text{ violates (a), (b), or (c)}] \\
&\leq \Pr[e \text{ violates (a)}] + \Pr[e \text{ violates (b) or (c)}] \\
&\leq \frac{3}{4} + \frac{1}{10} \\
&= \frac{17}{20},
\end{aligned}
$$

thus $\Pr[e \text{ is good}] \geq \frac{3}{20}$. In particular, there exists a good edge.