0368-4283: Space-Bounded Computation 12/6/2018 – Lecture 13

Solving Laplacian systems space-efficiently

Amnon Ta-Shma and Dean Doron

1 The Moore-Penrose pseudo-inverse

Definition 1. *Let L be any matrix. The Moore-Penrose pseudo-inverse of L, $L^\star$, is the unique matrix satisfying all of the following:*

1. $LL^\star L = L$.
2. $L^\star L L^\star = L^\star$.
3. $LL^\star$ *is Hermitian.*
4. $L^\star L$ *is Hermitian.*

The SVD gives us a way to compute the pseudo-inverse.

Lemma 2. *Let $L = U\Sigma V^\dagger$ be the SVD of L. Then, $L^\star = V\Sigma^\star U^\dagger$, where $\Sigma^\star$ is obtained by taking the reciprocal of each non-zero element on the diagonal, leaving the zeros in place.*

When L is normal and has non-negative eigenvalues (e.g., when G is regular), the SVD coincides with the spectral decomposition. Thus:

Lemma 3. *Assume that L is Hermitian with eigenvalues $0 = \lambda_n \le \ldots \le \lambda_1$. Then, $L^\star$ has the same eigenvectors as L and eigenvalues $\lambda_n^\star \le \ldots \le \lambda_1^\star$, where $\lambda_i^\star = \lambda_i^{-1}$ if $\lambda_i \neq 0$ and 0 otherwise.*

2 The Laplacian

Let M be a transition matrix of an undirected graph G over n vertices (that is, $M = AD^{-1}$ where A is the adjacency matrix of G and D is its degree diagonal matrix). The *normalized Laplacian* of G is

$$L \;=\; I - M,$$

and is an important operator in many areas of science. We make a few simple observations:

1. M is similar to the Hermitian matrix $D^{-/12}MD^{1/2} = D^{-1/2}AD^{-/2}$, and so L has real eigenvalues satisfying $0 = \lambda_n \le \ldots \le \lambda_1 \le 2$, i.e., L is PSD.
2. G is connected if $\lambda_{n-1} = \gamma(G) > 0$, where $\gamma(G)$ is the spectral gap of G.

We would like to solve the linear system $Lx = b$ space-efficiently, and present a solution due to Murtagh, Reingold, Sidford and Vadhan [2]. As L is not invertible (0 is an eigenvalue), we need to work with a generalized inverse. For undirected graphs, the Moore-Penrose pseudo-inverse is a natural choice.

In fact, we can (and will) assume w.l.o.g. that G is regular. Indeed,

Lemma 4. *Suppose G is an undirected graph with adjacency matrix A, degree matrix D, and transition matrix $M = AD^{-1}$. Let E a diagonal matrix of self-loops to make the graph d-regular. Then, Let $L = I - M$ be the original Laplacian and $L' = I - \frac{1}{d}(A+E)$ the new Laplacian. Then*

$$L^\star = \frac{1}{d} D (L')^\star$$

Proof. First, $LD = (I-M)D = D - A = D + E - (A+E) = dI - (A+E) = d(I - \frac{1}{d}(A+E)) = dL'$. Now we check that $\frac{1}{d}D(L')^*$ satisfies the four conditions above (and notice that L' and its pseudo inverse are normal). □

This way, we can also make sure that G is aperiodic.

We now need to justify that the pseudo-inverse indeed helps us solving linear equations.

Claim 5. *If the linear system $Lx = b$ has a solution then a solution is given by $x = L^\star b$.*

Proof. Let x_0 be a solution, so $Lx_0 = b$. Then $LL^*b = LL^*Lx_0 = Lx_0 = b$. Hence, $L^\star b$ is a solution to the system. □

In fact, a full characterization of the solutions of $Lx = b$ is known, but we will not get into the details.

3 A PSD partial order

Definition 6. *Let A be a symmetric real matrix. The following is equivalent:*

- *For every $v \in \mathbb{R}^n$ it holds that $v^\dagger A v \geq 0$,*
- *All the eigenvalues of A are non-negative.*

If any condition happens we say A is positive semi-definite, and denote $A \geq 0$. We write $A \leq B$ if $B - A \geq 0$.

Given a d-regular graph G (we assume w.l.o.g. that d is a power of two), we want to approximate the entries of $L^\star$. We use the spectral approximation of matrices.

Definition 7. *Let X, Y be real, symmetric matrices. We say that $X \approx_\varepsilon Y$ if $e^{-\varepsilon} \cdot X \leq Y \leq e^{\varepsilon} \cdot X$.*

The above notion of approximation is seemingly weaker than approximation in the ℓ_∞-norm. However, for the purpose of approximating a solution to a Laplacian system, it suffices:

Claim 8. *Assume $\widetilde{L^\star} \approx_\varepsilon L^\star$ and $b \in \mathsf{Image}(L)$. Then, $\left\| L^\star b - \widetilde{L^\star} b \right\| \leq \varepsilon \left\| L^\star b \right\|$.*

Proof. See Appendix D (or Lemma 8.3) of the MRSV paper. □

We record some useful facts about the spectral approximation. We start with expected facts:

Claim 9. *Let X, Y, W, Z be real, symmetric, positive semi-definite matrices. Then:*

1. *If* $X \approx_\varepsilon Y$ *then* $Y \approx_\varepsilon X$.
2. *If* $X \approx_\varepsilon Y$ *and* c *is a non-negative real number then* $cX \approx_\varepsilon cY$.
3. *If* $X \approx_\varepsilon Y$ *and* $W \approx_\varepsilon Z$ *then* $X + W \approx_\varepsilon Y + Z$.
4. *If* $X \approx_{\varepsilon_1} Y$ *and* $Y \approx_{\varepsilon_2} Z$ *then* $X \approx_{\varepsilon_1+\varepsilon_2} Z$.
5. *If* $X \approx_\varepsilon Y$ *then* $I \otimes X \approx_\varepsilon I \otimes Y$.

All the above items are straightforward.

We now record two other facts:

1. If $X \approx_\varepsilon Y$ and M is any matrix then $M^\dagger X M \approx_\varepsilon M^\dagger Y M$.
2. If $X \approx_\varepsilon Y$ and $\ker(X) = \ker(Y)$ then $X^\star \approx_\varepsilon Y^\star$.

The first inequality is straight forward but quite useful, and does not hold with, e.g., operator norm approximation. Let us prove the second one.

Claim 10. *If* $X \approx_\varepsilon Y$ *and* $\ker(X) = \ker(Y)$ *then* $X^\star \approx_\varepsilon Y^\star$.

Proof. Let $v \in \mathbb{R}^n$ and write $v = a + b$ where a belongs to $\ker(X) = \ker(Y)$ and b belongs to the kernel's orthogonal complement. As $\ker(X^\star) = \ker(X)$, it holds that $v^\dagger X^\star v = b^\dagger X^\star b$ and likewise for Y. Thus, we can restrict ourselves to the invertible subspace, assume w.l.o.g. that $X, Y > 0$ and prove that $X \approx_\varepsilon Y$ implies $X^{-1} \approx_\varepsilon Y^{-1}$. In fact, we will show that for general $A, B > 0$, $A \le B$ implies $B^{-1} \le A^{-1}$ and the claim will follow by first taking $B = e^\varepsilon X$ and $A = Y$ and then taking $B = Y$ and $A = e^{-\varepsilon} X$.

Let $A, B > 0$ such that $A \le B$. As $B > 0$ it has a unique positive definite square root. Let $M = \sqrt{B}^{-1} A \sqrt{B}^{-1}$, and so $M^{-1} = \sqrt{B} A^{-1} \sqrt{B}$. Note that $M \le I$ if and only if $M^{-1} \ge I$. Now, $A \le B$ implies $M \le I$ by multiplying both sides by $\sqrt{B}^{-1}$, and thus $M^{-1} \ge I$. Multiplying again both sides by $\sqrt{B}^{-1}$ gives $A^{-1} \ge B^{-1}$, as required. □

4 A second look at derandomized squaring

4.1 Expnders

Even if you take a very good expander G it will not be the case that $G \approx J$, because a non-zero number cannot be multiplicatively close to zero. However, with the Laplacian things are different:

Claim 11. *Let* G *be a* d *regular, undirected graph with transition matrix* M. *Let* $L = I - M$. *Then,* $L \approx_{\lambda(G)} I - J$.

Proof. We can express $M = \frac{1}{d} \sum_i \lambda_i |v_i\rangle\langle v_i|$ with $\lambda_1 = d$ and $|v_1\rangle\langle v_1| = J$. Hence,

$$L = I - M = I - J - C$$

where $C = \sum_{i \ge 2} \lambda_i |v_i\rangle\langle v_i|$, $\|C\| \le \lambda(G)$ and C, I, J commute. For every $i \ge 2$, $e^{-\lambda} \le 1 - \lambda_i \le 1$, because $(1 - \lambda)e^\lambda \ge 1$. □

Next, we turn to derandomized squaring.

Lemma 12. *Let G be a d-regular, undirected, aperiodic graph on n vertices with transition matrix M and H be a regular graph over d vertices. Then,*

$$I - M^2 \approx_{\lambda(H)} I - M_{G Ⓢ H},$$

where $M_{G Ⓢ H}$ is the transition matrix of $G Ⓢ H$.

Proof. Using the notations of previous lectures, $M_{G Ⓢ H} = P\dot{M}\tilde{H}\dot{M}L$. Using Claim 9, we can conclude the following:

$$\begin{aligned}
I - P\dot{M}\tilde{H}\dot{M}L &= P\dot{M}I\dot{M}L - P\dot{M}\tilde{H}\dot{M}L \\
&= P\dot{M}(I \otimes (I - H))\dot{M}L \\
&= d(\dot{M}L)^\dagger (I \otimes (I - H))\dot{M}L \\
&\approx_{\lambda(H)} d(\dot{M}L)^\dagger (I \otimes (I - J))\dot{M}L \\
&= P\dot{M}(I - \tilde{J})\dot{M}L \\
&= P\dot{M}I\dot{M}L - P\dot{M}\tilde{J}\dot{M}L \\
&= I - M^2.
\end{aligned}$$

□

5 Approximating $L^\star$ space-efficiently

We prove:

Theorem 13 ([2])**.** *Given an undirected graph G over n vertices with a Laplacian L and $\varepsilon > 0$, there exists a deterministic algorithm that computes $\widetilde{L^\star}$, where $\widetilde{L^\star} \approx_\varepsilon L^\star$, running in space $O(\log n \log\log(n/\varepsilon))$.*

We remark that with a *probabilistic* algorithm, one can do it in space $O(\log(n/\varepsilon))$ [1].

Our main tool, together with a clever identity, will be our well-known *derandomized squaring*. We shall see that the derandomized squaring does not only improves a graph's connectivity but can also replace the original squaring when it comes to approximating a solution to a Laplacian system!

We will prove the theorem for constant error ε.

Lemma 14. *Given an undirected graph G over n vertices with a Laplacian L, there exists a deterministic algorithm that computes $\widetilde{L^\star}$, where $\widetilde{L^\star} \approx_{1/2} L^\star$, running in space $O(\log n \log\log n)$.*

Proof. Assume w.l.o.g. that G is d-regular, with $d/2$ self-loops on every vertex. Fix $k = O(\log n)$, $c = \text{poly}(\log n)$. Define the sequence of graphs $G_0 = G, \ldots, G_k$ where

$$G_i = G_{i-1} Ⓢ H_i$$

and every H_i is a c-regular expander over $d \cdot c^{i-1}$ vertices with $k \cdot \lambda(H_i) \le \frac{1}{4}$. There are easy explicit H_i like that.

Let $M_0, \ldots, M_k$ be the transition matrices of $G_0 = G, \ldots, G_k$. Since we take $k = O(\log n)$ we have $\lambda(M_k) \leq \frac{1}{4}$.

We define

$$W_i = (I + M_0) \cdot \ldots \cdot (I + M_i)(I - J)(I + M_i) \cdot \ldots \cdot (I + M_0).$$

for all $i \in \{0, \ldots, k-1\}$.

Let

$$\widetilde{L_0}^{\star} = \frac{1}{2}(I - J) + \left(\sum_{i=0}^{k-1} \frac{1}{2^{i+2}} W_i \right) + \frac{1}{2^{k+1}} W_k.$$

The space complexity follows from the following two claims (and composition of space-bounded computations). First, iterated derandomized squaring can be computed in small space:

Claim 15. *Every coordinate of every M_i can be computed in space $O(\log n + k \log c) = O(\log n \log\log n)$.*

Second, it is not difficult to verify that:

Claim 16. *Every coordinate of the multiplication of k matrices of dimension n can be computed in space $O(\log n \cdot \log k)$.*

Together, this shows that $\widetilde{L_0}^{\star}$ can be computed in $O(\log n \cdot \log\log n)$ space.

In the next section we turn to the correctness proof. □

6 Correctness

6.1 Expressing the Laplacian of M with that of M^2

First, let us see a nice matrix identity, following [3], that expresses the pseudo-inverse of the Laplacian of G via the pseudo-inverse of the Laplacian of G^2.

Lemma 17. *If $L = I - M$ is the normalized Laplacian of an undirected, connected, regular, aperiodic graph on n vertices then*

$$(I - M)^{\star} = L^{\star} \;=\; \frac{1}{2} \left(I - J + (I + M)(I - M^2)^{\star}(I + M) \right)$$

Proof. The matrices $I, M, J, U - M^2$ are symmetric and commute. Furthermore, we saw that also $(I - M^2)^*$ commute with them. So L^* commutes with L and the share the same orthonormal basis. It is enough to compare eigenvalues.

Let v be an eigenvector of $L = I - M$ with eigenvalue λ. Because of $(I - M^2)^*$ we have a different behaviour with $\lambda = 0$ and $\lambda \neq 0$. For $\lambda = 0$ we get v is all the one vector, $Jv = Iv = Mv = v$ and $(I - M^2)^* v = 0$. So the RHS on v is 0. Clearly $(I - M)^* v$ so an equality holds.

Let v be an eigenvector of L with eigenvalue $\lambda \neq 0$, so $v \perp \mathbf{1}$. Thus, $(I - J)v = v$, $(I + M)v = (1 + \lambda)v$ and $(I - M^2)^{\star} v = \frac{1}{1 - \lambda^2} v$. Altogether,

$$\frac{1}{2} \left(I - J + (I + M)(I - M^2)^{\star}(I + M) \right) v = \frac{1}{2} v + \frac{1}{2} \frac{(1 + \lambda)^2}{1 - \lambda^2} v = \frac{1}{1 - \lambda} v = L^{\star} v.$$

□

Applying the above identity k times:

$$\begin{aligned}(I-M)^\star = L^\star &= \frac{1}{2}(I-J) + \frac{1}{2}(I+M)(I-M^2)^\star(I+M)\\ &= \frac{1}{2}(I-J) + \frac{1}{4}(I+M)(I-J)(I+M) + \frac{1}{4}(I+M)(I+M^2)(I-M^4)^\star(I+M^2)(I+M)\\ &= \dots\\ &= \frac{1}{2}(I-J) + \left(\sum_{i=0}^{k-1}\frac{1}{2^{i+2}}P_i\right) + \frac{1}{2^{k+1}}(I+M)\cdot\ldots\cdot(I+M^{2^k})(I-M^{2^{k+1}})^\star(I+M^{2^k})\cdot\ldots\cdot(I+M)\end{aligned}$$

where for all $i \in \{0,\dots,k-1\}$,

$$P_i = (I+M)\cdot\ldots\cdot(I+M^{2^i})(I-J)(I+M^{2^i})\cdot\ldots\cdot(I+M).$$

To summarize:

Lemma 18 ([3])**.** *Let $M_0,\dots,M_k$ be symmetric matrices such that $L_i = I-M_i \geq 0$, $L_i = I-M_{i-1}^2$ (i.e., $M_i = M_{i-1}^2$). Then,*

$$(I-M_0)^\star = \frac{1}{2}(I-J) + \left(\sum_{i=0}^{k-1}\frac{1}{2^{i+2}}W_i\right) + \frac{1}{2^{k+1}}R_k.$$

where $W_i = (I+M_0)\cdot\ldots\cdot(I+M_i)(I-J)(I+M_i)\cdot\ldots\cdot(I+M_0)$ and $R_k = (I+M_0)(I+M_1)\dots(I+M_k)(I-M_{k+1})(I+M_k)\dots(I+M_0)$.

6.2 Approximating the Laplacian of M with approximations of the Laplacian of M^2

Let us recap. We have a sequence of graphs $G_0 = G,\dots,G_k$ where

$$G_i = G_{i-1} \circledS H_i$$

and every H_i is a c-regular expander over $d\cdot c^{i-1}$ vertices with $\lambda(H_i) \leq \varepsilon = \frac{1}{4k}$. We let M_i denote the transition matrix of G_i and $L_i = I - M_i$. In Lemma 12 we have seen that

$$L_{i+1} = I - M_{i+1} = I - M_{G_i \circledS H_i} \approx_{\lambda(H_i)} I - M_i^2,$$

i.e., $L_{i+1} \approx_\varepsilon I - M_i^2$.

Claim 19. *$L_i^\star = (I-M_i)^\star \approx_\varepsilon \frac{1}{2}(I-J+(I+M_i)L_{i+1}^\star(I+M_i))$, for every $i \in \{0,\dots,k-1\}$.*

Proof. As we noted before $L_i = (I-M_i), I+M_i, (I-M_i)^\star, (I-M_i^2)^\star, I, J$ commute, and therefore the two matrices on the both sides of the claim share the same orthonormal eigenvector basis. Furthermore, they share the same kernel. By Claim 10, $L_{i+1}^\star \approx_\varepsilon (I-M_i^2)^\star$. By Claim 9, we can multiply both sides by $I+M_i$ and get

$$(I+M_i)L_{i+1}^\star(I+M_i) \approx_\varepsilon (I+M_i)(I-M_i^2)^\star(I+M_i).$$

Again by Claim 9, we can multiply both sides by $\frac{1}{2}$ and add $\frac{1}{2}(I-J) \geq 0$, so

$$\frac{1}{2}(I-J+(I+M_i)L_{i+1}^\star(I+M_i)) \approx_\varepsilon \frac{1}{2}(I-J+(I+M_i)(I-M_i^2)^\star(I+M_i)) = (I-M_i)^\star = L_i^\star,$$

where we have used Lemma 17. $\square$

Lemma 20 ([3]). *Let $\varepsilon, \delta > 0$ and let $M_0, \ldots, M_k$ be symmetric matrices such that $L_i = I - M_i \geq 0$, $L_i \approx_\varepsilon I - M_{i-1}^2$ and $L_k \approx_\delta I - J$. Then,*

$$L_0^\star \approx_{k\varepsilon+\delta} \frac{1}{2}(I-J) + \left(\sum_{i=0}^{k-1} \frac{1}{2^{i+2}} W_i\right) + \frac{1}{2^{k+1}} W_k.$$

where $W_i = (I+M_0) \cdot \ldots \cdot (I+M_i)(I-J)(I+M_i) \cdot \ldots \cdot (I+M_0)$..

Proof.

$$\begin{aligned}
L_0^\star &= (I-M_0)^\star \approx_\varepsilon \frac{1}{2}(I - J + (I+M_0)L_1^\star(I+M_0)) \\
&\approx_\varepsilon \frac{1}{2}(I-J) + \frac{1}{4}(I+M_0)(I-J)(I+M_0) + \frac{1}{4}(I+M_0)(I+M_1)L_2^\star(I+M_1)(I+M_0) \\
&\approx_\varepsilon \ldots \\
&\approx_\varepsilon \frac{1}{2}(I-J) + \left(\sum_{i=0}^{k-1} \frac{1}{2^{i+2}} W_i\right) + \frac{1}{2^{k+1}} R_k
\end{aligned}$$

where $R_k = (I+M_0) \cdot \ldots \cdot (I+M_{k-1})(I-M_k)^\star(I+M_{k-1}) \cdot \ldots \cdot (I+M_0)$. Finally, $L_k \approx_\delta I - J$ and this implies $R_k \approx_\delta W_k$. □

6.3 Obtaining a constant error approximation

Note that if we want a constant error approximation, in order to use Lemma 20 we need each $\varepsilon = \frac{1}{O(\log n)}$ so that $k\varepsilon$ is constant for $k = O(\log n)$ (and we need $k = O(\log n)$ to reach a good enough approximation of $I - J$). This, in turn, implies that we need to take c, the degree of the expanders, to be poly-logarithmic in n.

7 Construction for low ε

It is left to show that by investing extra $O(\log n \log\log(n/\varepsilon))$, we can boost our $\frac{1}{2}$-approximation to an ε-approximation. This follows from the following nice lemma:

Lemma 21. *Let A, P be real symmetric and invertible matrices of dimension n so that there exists ε for which $(1-\varepsilon)P^{-1} \leq A \leq P^{-1}$. Then, $P_k = \sum_{i=0}^{k} P(I-AP)^i$ satisfies*

$$\left(1 - \frac{\varepsilon^{k+1}}{1-\varepsilon}\right) A^{-1} \leq P_k \leq A^{-1}$$

For the proof of the above lemma, and how to use it in order to get Theorem 13, see [2].

References

[1] Dean Doron, François Le Gall, and Amnon Ta-Shma. Probabilistic logarithmic-space algorithms for laplacian solvers. In *LIPIcs-Leibniz International Proceedings in Informatics*, volume 81. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017.

[2] Jack Murtagh, Omer Reingold, Aaron Sidford, and Salil Vadhan. Derandomization beyond connectivity: Undirected laplacian systems in nearly logarithmic space. *arXiv preprint arXiv:1708.04634*, 2017.

[3] Richard Peng and Daniel A Spielman. An efficient parallel solver for sdd linear systems. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pages 333–342. ACM, 2014.