UMassAmherst

Manning College of Information & Computer Sciences

Sublinear Time Eigenvalue Approximation via Random Sampling

Rajarshi Bhattacharjee*, Gregory Dexter†, Petros Drineas†, Cameron Musco* and Archan Ray*

{rbhattacharj, cmusco, ray}@cs.umass.edu, {gdexter, pdrineas}@purdue.edu

*Manning College of Information and Computer Sciences, UMass Amherst

†Department of Computer Science, Purdue University

PURDUE UNIVERSITY

Problem Description

Given: A **symmetric** matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ in the **bounded entry model** *i.e* $\|\mathbf{A}\|_\infty \leq 1$ [1]. $\mathbf{A}$ can be **indefinite**.

Exact Eigenvalues: SVD, power methods, etc. require reading the **full matrix** and have time complexity close to $O(n^\omega)$.

Can **approximate** top k largest magnitude eigenvalues using $\tilde{O}(k)$ matrix vector multiplications with A (power method, Krylov subspace methods, etc.) $\tilde{O}(n^2\cdot k)$ time for dense matrices.

Goal: Approximate the spectrum in sublinear *i.e.* $o(n^2)$ time for dense matrices.

Bounded entry assumption: Otherwise, a single pair *i.e.* $\mathbf{A}_{ij}$ and $\mathbf{A}_{ji}$ can be arbitrarily large and dominate the top eigenvalues. Finding this pair takes $\Omega(n^2)$ time.

Approximation using Uniform Sampling

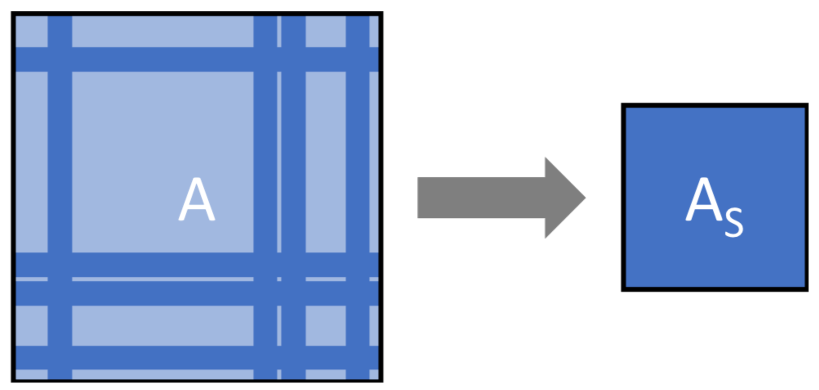
Theorem 1

There is an algorithm that reads $\tilde{O}(\frac{\log^6 n}{\epsilon^6})$ entries of a symmetric $\mathbf{A}$ with $\|\mathbf{A}\|_\infty \leq 1$ and outputs $\tilde{\lambda}_1, \tilde{\lambda}_2, \cdots, \tilde{\lambda}_n$ such that, $\forall i \in [n]$: $|\lambda_i - \tilde{\lambda}_i| \leq \epsilon n$.

$\mathbf{A}_S$: **Random principal submatrix** of $\mathbf{A}$ where each row/column is included independently with probability $\frac{s}{n}$ $(s = \frac{c\log^3 n}{\epsilon^3})$.

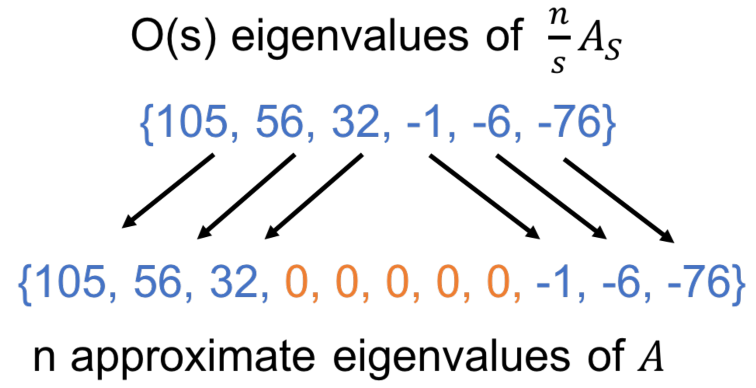
Compute all eigenvalues of $\frac{n}{s}\mathbf{A}_S$: use these to approximate $\lambda_i(\mathbf{A})$.

Need to **align** the eignevalues correctly.



Aligning Eigenvalues

$\mathbf{A}_S$ has only $O(s)$ eigenvalues but $\mathbf{A}$ has n eigenvalues.



Lower Bounds

General lower bound: of $O\left(\frac{1}{\epsilon^2}\right)$ **total** samples to distinguish an all zeros matrix from a matrix with a $O(\epsilon n \times \epsilon n)$ block of ones.

For principal submatrices – need at least $\tilde{O}\left(\frac{1}{\epsilon^2}\times\frac{1}{\epsilon^2}\right)$ samples. [2].

Proof Techniques – Uniform Sampling

Key Proof Idea: Split $\mathbf{A}$ into its **outlying** and **middle** eigenvalues analyze each component separately.

Let $\mathbf{A}$: $\mathbf{A} = \mathbf{A}_o + \mathbf{A}_m$ where $\mathbf{A}_o = \mathbf{V}_o\mathbf{\Lambda}_o\mathbf{V}_o^T$ and $\mathbf{A}_m = \mathbf{V}_m\mathbf{\Lambda}_m\mathbf{V}_m^T$ where $\mathbf{\Lambda}_o$, $\mathbf{\Lambda}_m$ are diagonal matrices, with eigenvalues of $\mathbf{A}$ with magnitude $\geq \epsilon n$ and $< \epsilon n$ on their diagonal respectively.

$\frac{n}{s}\cdot\mathbf{A}_S = \mathbf{S}^T\mathbf{A}S = \mathbf{S}^T\mathbf{A}_o\mathbf{S} + \mathbf{S}^T\mathbf{A}_m\mathbf{S}$.

Key Proof Idea: Since $\mathbf{A}$ has bounded entries, the **outlying** eigenvectors of $\mathbf{A}$, $\mathbf{V}_o$ are all **incoherent** i.e. their mass is spread out – $\|[\mathbf{V}_o]_{i,:}\|_2 \leq \frac{1}{\epsilon^2 n}$. So uniform sampling approximately preserves eigenvalues of $\mathbf{A}_o$. Thus, non-zero eigenvalues of $\mathbf{S}^T\mathbf{A}_o\mathbf{S}$ approximate eigenvalues of $\mathbf{A}$ up to $\pm\epsilon n$ error.

Use incoherence of $\mathbf{A}_o$ to argue $\mathbf{A}_m = \mathbf{A} - \mathbf{A}_o$ is entrywise bounded and thus, $\|\mathbf{S}^T\mathbf{A}_m\mathbf{S}\|_2 \leq \epsilon n$ using known spectral norm bounds. Finally combine the above using **Weyl's inequality** $\|\frac{n}{s}\cdot\mathbf{A}_S - \mathbf{S}^T\mathbf{A}_o\mathbf{S}\|_2 \leq \|\mathbf{S}^T\mathbf{A}_m\mathbf{S}\|_2 \leq \epsilon n$.

Approximation using Sparsity Sampling

Theorem 2

There is an algorithm that reads $\tilde{O}(\frac{\log^{16} n}{\epsilon^{16}})$ entries of a symmetric $\mathbf{A}$ with $\|\mathbf{A}\|_\infty \leq 1$ and outputs $\tilde{\lambda}_1, \cdots, \tilde{\lambda}_n$ such that, $\forall i \in [n]$: $|\lambda_i - \tilde{\lambda}_i| \leq \epsilon\sqrt{\text{nnz}(\mathbf{A})}$.

Key Step: **Let $\mathbf{A}'$ be equal to $\mathbf{A}$ but with $\mathbf{A}'_{ij} = 0$** $\sqrt{\text{nnz}(\mathbf{A}_i)\,\text{nnz}(\mathbf{A}_j)} \leq \frac{\epsilon\sqrt{\text{nnz}(\mathbf{A})}}{c\log n}$

$\mathbf{A}_S$: Random principal submatrix of $\mathbf{A}'$ where each i^{th} row/column is included independently with probability $p_i \geq \min\left(1, \frac{s\,\text{nnz}(\mathbf{A}_i)}{\text{nnz}(\mathbf{A})}\right)$ $(s \approx \frac{c\log^8 n}{\epsilon^8})$.

Let $\mathbf{D}$ be a diagonal matrix with $\mathbf{D}_{i,i} = \frac{1}{\sqrt{p_i}}$. Compute all eigenvalues of $\mathbf{D}\mathbf{A}_S\mathbf{D}$: use these to approximate $\lambda_i(\mathbf{A})$.

Key Idea: Zeroing out entries ensures that after sampling and scaling, no entries are scaled up by too much. Can show $\|\mathbf{A}' - \mathbf{A}\|_2 \leq \epsilon\sqrt{\text{nnz}(\mathbf{A})}$; can be thought of generalization of **Girshgorin theorem**. Allows us to extend our uniform sampling proof.

Extensions to ℓ_2 sampling – sample with $p_i \geq \min\left(1, \frac{s\|\mathbf{A}_i\|_2^2}{\|\mathbf{A}\|_F^2}\right)$, zero out and scale appropriately to get $\pm\epsilon\|\mathbf{A}\|_F$ error without bounded entry assumption.

Open Questions

Obtain tight $\tilde{O}(\frac{1}{\epsilon^2})$ query complexity for computing $\pm\epsilon n$ approximation. Requires going **beyond principal submatrix sampling**.

How to estimate **bulk spectral properties** like Schatten norm using $o(n^2)$ queries.

References

[1] Balcan, M.-F., Y. Li, D. P. Woodruff, et al. Testing matrix rank, optimally. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 727–746. SIAM, 2019.

[2] Bakshi, A., N. Chepurko, R. Jayaram. Testing positive semi-definiteness via random submatrices. *arXiv preprint arXiv:2005.06441*, 2020.