Sublinear Time Eigenvalue Approximation via Random Sampling

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

†Manning College of Information and Computer Sciences, UMass Amherst
‡Department of Computer Science, Purdue University

Problem Description

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

Exact Eigenvalues: SVD, power methods, etc. require reading the full matrix and have time complexity close to \( O(n^3) \).

Can approximate top \( k \) largest magnitude eigenvalues using \( O(k) \) matrix vector multiplications with \( A \) (power method, Krylov subspace methods, etc.) \( O(n^2 \cdot k) \) 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. \( A_{ij} \) and \( A_{ji} \) 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 \( \tilde{O}(\frac{kn^{10} \log n}{\epsilon^6}) \) entries of a symmetric \( A \) with \( \| A \|_\infty \leq 1 \) and outputs \( \tilde{\lambda}_1, \tilde{\lambda}_2, \ldots, \tilde{\lambda}_n \) such that, \( \forall i \in [n]: |\lambda_i - \tilde{\lambda}_i| \leq \epsilon n \).

Key Step: Let \( A' \) be equal to \( A \) but with \( A'_{ij} = 0 \) and \( A_{ij} \) is included independently with probability \( p_i \). Compute all eigenvalues of \( A' \) and outputs eigenvalues of \( A \).

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

Extensions to \( \ell_2 \) sampling – sample with \( p_i \geq \min \{ 1, \frac{\| A \|_F}{\sqrt{\text{nnz}(A)}} \} \), zero out and scale appropriately to get \( \pm \epsilon\| A \|_F \) error without bounded entry assumption.

Aligning Eigenvalues

\( A_S \) has only \( O(s) \) eigenvalues but \( A \) has \( n \) eigenvalues.

\[
O(s) \text{ eigenvalues of } \tilde{A}_S = \begin{cases} \frac{n}{s} A_S & (105, 56, 32, -1, -6, -76) \\
(105, 56, 32, 0, 0, 0, 0, -1, -6, -76) & n \text{ approximate eigenvalues of } A
\end{cases}
\]

Lower Bounds

General lower bound: of \( O(\frac{1}{\epsilon}) \) total samples to distinguish an all zeros matrix from a \( O(en \times en) \) block of ones.

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

Proof Techniques – Uniform Sampling

Key Proof Idea: Split \( A \) into its outlying and middle eigenvalues analyze each component separately.

Let \( A = A_0 + A_m \) where \( A_0 = V_0 A_0 V_0^T \) and \( A_m = V_m A_m V_m^T \) where \( A_0, A_m \) are diagonal matrices, with eigenvalues of \( A \) with magnitude \( \geq en \) and \( < en \) on their diagonal respectively.

\[
\frac{n}{s} A_S = S^T A S = S^T A_0 S + S^T A_m S.
\]

Key Proof Idea: Since \( A \) has bounded entries, the outlying eigenvectors of \( A_0, V_0 \) are all incoherent i.e. their mass is spread out and \( \| V_0 \|_2 \leq \frac{1}{\epsilon^2} \). So uniform sampling approximately preserves eigenvalues of \( A_0 \). Thus, non-zero eigenvalues of \( S^T A_0 S \) approximate eigenvalues of \( A \) up to \( \pm \epsilon \) error.

Use incoherence of \( A_0 \) to argue \( A_m = A - A_0 \) is entrywise bounded and thus, \( \| S^T A_m S \|_2 \leq en \) using known spectral norm bounds. Finally combine the above using Weyl’s inequality \( \| S^T A_S - S^T A_0 S \|_2 \leq |\tilde{\lambda}_i - \lambda_i| \leq \epsilon n \).

Approximation using Sparsity Sampling

Theorem 2

There is an algorithm that reads \( \tilde{O}(\frac{kn^{10} \log n}{\epsilon^6}) \) entries of a symmetric \( A \) with \( \| A \|_\infty \leq 1 \) and outputs \( \tilde{\lambda}_1, \ldots, \tilde{\lambda}_n \) such that, \( \forall i \in [n]: |\lambda_i - \tilde{\lambda}_i| \leq \epsilon \sqrt{\text{nnz}(A)} \).

Key Step: Let \( A' \) be equal to \( A \) but with \( A'_{ij} = 0 \) and \( A_{ij} \) is included independently with probability \( p_i \). Compute all eigenvalues of \( DA_D \); use these to approximate \( \lambda_i(A) \).

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

Obtain tight \( \tilde{O}(\frac{1}{\epsilon}) \) 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
