

# Sublinear Time Eigenvalue Approximation via Random Sampling

Tuesday, July 11, 2023 4:45 PM (20 minutes)

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

**Abstract:** We study the problem of approximating the eigenspectrum of a symmetric matrix  $\mathbf{A} \in \mathbb{R}^{n \times n}$  with bounded entries (i.e.,  $\|\mathbf{A}\|_\infty \leq 1$ ). We present a simple sublinear time algorithm that approximates all eigenvalues of  $\mathbf{A}$  up to additive error  $\pm \epsilon n$  using those of a randomly sampled  $\tilde{O}\left(\frac{\log^3 n}{\epsilon^3}\right) \times \tilde{O}\left(\frac{\log^3 n}{\epsilon^3}\right)$  principal submatrix. Our result can be viewed as a concentration bound on the complete eigenspectrum of a random submatrix, significantly extending known bounds on just the singular values (the magnitudes of the eigenvalues). We give improved error bounds of  $\pm \epsilon \sqrt{nnz(\mathbf{A})}$  and  $\pm \epsilon \|\mathbf{A}\|_F$  when the rows of  $\mathbf{A}$  can be sampled with probabilities proportional to their sparsities or their squared  $\ell_2$  norms respectively. Here  $nnz(\mathbf{A})$  is the number of non-zero entries in  $\mathbf{A}$  and  $\|\mathbf{A}\|_F$  is its Frobenius norm. Even for the strictly easier problems of approximating the singular values or testing the existence of large negative eigenvalues (Bakshi, Chepurko, and Jayaram, FOCS '20), our results are the first that take advantage of non-uniform sampling to give improved error bounds. From a technical perspective, our results require several new eigenvalue concentration and perturbation bounds for matrices with bounded entries. Our non-uniform sampling bounds require a new algorithmic approach, which judiciously zeroes out entries of a randomly sampled submatrix to reduce variance, before computing the eigenvalues of that submatrix as estimates for those of  $\mathbf{A}$ . We complement our theoretical results with numerical simulations, which demonstrate the effectiveness of our algorithms in practice.

**Presenter:** RAY, Archan

**Session Classification:** Track A-4