Estimation of Spectral Gaps for Sparse Symmetric Matrices

Michele Benzi, Michele Rinelli, Igor Simunec

Abstract

Identifying the gap between two consecutive eigenvalues of a real symmetric matrix $A \in \mathbb{R}^{n \times n}$ is an important task that is often encountered in applications, such as in electronic structure computations. For instance, in Kohn-Sham Density Functional Theory [5] one has to determine the spectral projector P_{μ} associated with the *Fermi level* (or *chemical potential*) $\mu \in \mathbb{R}$, which corresponds to the occupied states of a system described by a discrete Hamiltonian A. In the case of insulators at zero electronic temperature, there is a gap separating the first k eigenvalues of A from the rest of the spectrum, and the Fermi level μ lies inside the gap between the k-th and (k + 1)-th eigenvalue of A, where k is the number of electrons in the system. In this setting, the spectral projector $P_{\mu} = h_{\mu}(A)$ is often approximated using polynomials or rational functions that approximate the step function $h_{\mu}(\lambda)$, which takes the value 1 for $\lambda < \mu$ and 0 for $\lambda > \mu$. In order to use this approach, one first needs to compute a value of μ that lies in the gap between λ_k and λ_{k+1} , where we denote the eigenvalues of A as $\lambda_1, \ldots, \lambda_n$ in nondecreasing order.

Instead of looking for the gap between λ_k and λ_{k+1} for a specific k, in this talk we tackle this problem from a different perspective, and aim to find all gaps in the spectrum of A that are larger than a certain threshold. Since in practical applications the gap between λ_k and λ_{k+1} is often relatively large, we expect that this approach will provide useful results even when one needs to find a single, specific gap.

Let us denote by $n_e(\mu)$ the number of eigenvalues of A that are strictly smaller than μ . Assuming that μ does not coincide with an eigenvalue of A, we have $n_e(\mu) = \operatorname{rank}(P_{\mu}) = \operatorname{tr}(P_{\mu})$. Therefore, $n_e(\mu)$ can be estimated by estimating $\operatorname{tr}(P_{\mu})$ with Hutchinson's stochastic trace estimator [4] combined with the Lanczos algorithm [7, Algorithm 6.15]. Given s random Gaussian vectors $\{\boldsymbol{x}_i\}_{i=1}^s$, Hutchinson's stochastic trace estimator approximates $\operatorname{tr}(P_{\mu})$ with

$$\operatorname{tr}^{\operatorname{H}}_{s}(P_{\mu}) := rac{1}{s} \sum_{i=1}^{s} oldsymbol{x}_{i}^{T} P_{\mu} oldsymbol{x}_{i}$$

Since $P_{\mu} = h_{\mu}(A)$, each quadratic form can be approximated using the Lanczos algorithm. Let $V_m^{(i)} \in \mathbb{R}^{n \times m}$ be an orthonormal basis of the Krylov subspace

$$\mathcal{K}_m(A, \boldsymbol{x}_i) = \operatorname{span}\{\boldsymbol{x}_i, A\boldsymbol{x}_i, \dots, \boldsymbol{A}^{m-1}\boldsymbol{x}_i\},\$$

constructed with the Lanczos algorithm, and let the tridiagonal matrix $T_m^{(i)} := V_m^{(i)T} A V_m^{(i)}$ be the projection of A onto $\mathcal{K}_m(A, \boldsymbol{x}_i)$. We can approximate $\boldsymbol{x}_i^T P_{\mu} \boldsymbol{x}_i$ with

$$\psi_m^{(i)}(\mu) := \| \boldsymbol{x} \|_2^2 \boldsymbol{e}_1^T h_\mu(T_m^{(i)}) \boldsymbol{e}_1, \qquad j = 1, \dots, N_f,$$

so we obtain the trace approximations

$$\operatorname{tr}(P_{\mu}) \approx \frac{1}{s} \sum_{i=1}^{s} \psi_{m}^{(i)}(\mu)$$

If we use the same vectors $\{\boldsymbol{x}_i\}_{i=1}^s$ for different μ , these trace approximations can be computed simultaneously for many different μ by using the same Krylov subspaces $\mathcal{K}_m(A, \boldsymbol{x}_i)$, with a cost that

is only slightly higher than for a single value of μ . This approach has already been used in literature on related problems, such as the estimation of the number of eigenvalues of A in an interval or the estimation of spectral densities [3, 6]. In this talk we will focus on thoroughly analyzing this method for the detection of spectral gaps, with the goal of determining how to choose the parameters sand m in order to minimize the computational cost and ensure that all gaps with relative width above a given threshold $\theta \in (0, 1)$ are found (up to a failure probability δ).

The error of Hutchinson's estimator can be bounded, for instance, with [2, Theorem 1], which states that

$$\mathbb{P}\left(|\operatorname{tr}(P_{\mu}) - \operatorname{tr}_{s}^{\mathrm{H}}(P_{\mu})| \geq \varepsilon\right) \leq \delta \quad \text{if} \quad s \geq \frac{4}{\varepsilon^{2}} \left(\|P_{\mu}\|_{F}^{2} + \varepsilon \|P_{\mu}\|_{2}\right) \log(2/\delta).$$

However, we have $||P_{\mu}||_{F}^{2} = n_{e}(\mu)$, and $n_{e}(\mu) = O(n)$ when μ is near the middle of the spectrum, so to achieve any fixed absolute accuracy ε we would have to take s = O(n), which becomes unfeasible as n grows. This means that with this approach it is prohibitively expensive to try and find a value of μ in the gap between λ_{k} and λ_{k+1} by requiring that $\operatorname{tr}_{s}^{\mathrm{H}}(P_{\mu}) \approx n_{e}(\mu) = k$. Instead, we use a different point of view.

If we consider $\operatorname{tr}(P_{\mu})$ as a function of μ , it is a nondecreasing and piecewise constant function, with a jump of height 1 whenever μ coincides with an eigenvalue of A. A similar property holds for $\operatorname{tr}_{s}^{\mathrm{H}}(P_{\mu})$, with the difference that the jumps have random heights, each with expected value 1. In particular, $\operatorname{tr}_{s}^{\mathrm{H}}(P_{\mu})$ is constant for all $\mu \in [a, b]$ if the interval [a, b] contains no eigenvalues of A. This observation can be exploited to find gaps in the spectrum of A, by looking for intervals in which the Lanczos approximation $\frac{1}{s} \sum_{i=1}^{s} \psi_{m}^{(i)}(\mu)$ is almost constant in μ and has a small error. For instance, if for a constant C and a small $\varepsilon > 0$ we can show that

$$\operatorname{tr}_{s}^{\mathrm{H}}(P_{\mu}) \in [C - \varepsilon, C + \varepsilon] \quad \text{for all } \mu \in [a, b],$$

then we can conclude that either $\operatorname{tr}_{s}^{\mathrm{H}}(P_{\mu})$ is constant in the interval [a, b] and hence [a, b] is a gap in the spectrum of A, or all jumps in $\operatorname{tr}_{s}^{\mathrm{H}}(P_{\mu})$ associated with eigenvalues in [a, b] have heights smaller than 2ε . If ε is small enough, the latter event has a small chance of occurring.

In order to make the argument outlined above rigorous, we obtain a bound on the probability of having small jumps in $\operatorname{tr}_{s}^{\mathrm{H}}(P_{\mu})$, as well as a posteriori error bounds and estimates for the Lanczos approximation of the quadratic forms $\boldsymbol{x}_{i}^{T}P_{\mu}\boldsymbol{x}_{i}$. By combining these bounds, we will show that for a given budget of matrix-vector products with A, the best way to allocate it is to set s = 1, i.e., use a single random vector for Hutchinson's estimator and concentrate all the computational effort on the Lanczos algorithm. We also obtain an a priori bound on the accuracy of the Lanczos approximation that depends on the relative gap width θ , which will allow us to predict how many Lanczos iterations are needed to ensure that all gaps larger than a given width are detected.

The theoretical analysis is complemented by a detailed computational discussion, leading to an algorithm that is able to detect gaps efficiently and reliably. The effectiveness of the proposed method will be showcased with several numerical examples. Further details can be found in the preprint [1].

References

 M. Benzi, M. Rinelli, and I. Simunec. Estimation of spectral gaps for sparse symmetric matrices, arXiv:2410.15349, 2024.

- [2] A. Cortinovis and D. Kressner. On randomized trace estimates for indefinite matrices with an application to determinants. *Found. Comput. Math.*, 22(3):875–903, 2022.
- [3] E. Di Napoli, E. Polizzi, and Y. Saad. Efficient estimation of eigenvalue counts in an interval. Numer. Linear Algebra Appl., 23(4):674–692, 2016.
- [4] M. F. Hutchinson. A stochastic estimator of the trace of the influence matrix for Laplacian smoothing splines. Comm. Statist. Simulation Comput., 18(3):1059–1076, 1989.
- [5] L. Lin, J. Lu, and L. Ying. Numerical methods for Kohn-Sham density functional theory. Acta Numer., 28:405–539, 2019.
- [6] L. Lin, Y. Saad, and C. Yang. Approximating spectral densities of large matrices. SIAM Rev., 58(1):34–65, 2016.
- [7] Y. Saad. Iterative Methods for Sparse Linear Systems. SIAM, Philadelphia, PA, second edition, 2003.