Analysis of Stochastic Probing Methods for Estimating the Trace of Functions of Sparse Symmetric Matrices

Andreas Frommer, Michele Rinelli, Marcel Schweitzer

Abstract

Estimating the trace of an implicitly given matrix $B \in \mathbb{R}^{n \times n}$,

$$\operatorname{tr}(B) = \sum_{i=1}^{n} [B]_{ii}, \tag{1}$$

is an important task in many areas of applied mathematics and computer science. In many of these applications, we have $B = f(A)$, where $A \in \mathbb{R}^{n \times n}$ is a large and sparse (or structured) matrix. A common practice is to approximate (1) with an estimator of the form

$$\operatorname{tr}(B) \approx \sum_{k=1}^{N} \boldsymbol{v}_k^T B \boldsymbol{v}_k, \tag{2}$$

for suitably crafted vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N$. With this approach, approximating (1) relies on matrix-vector products or quadratic forms with B, which are, e.g., performed by applying a polyomial (or rational) Krylov subspace method or a Chebyshev expansion for approximating $f(A)\boldsymbol{v}$ or $\boldsymbol{v}^T f(A)\boldsymbol{v}$, avoiding the often prohibitive tasks of forming B or computing the eigenvalues of A.

Prominent examples are *stochastic estimators*, including Hutchinson's method [5], based on choosing random vectors in (2), and recent variants based on low-rank approximations, Hutch++ [6] and XTrace [2], which work especially well if a fast decay is present in the singular values of B.

When $B = f(A)$ with sparse, symmetric A, a popular other class of methods are based on *probing* [4, 7]. This approach requires the computation of a distance-d coloring of the graph $\mathcal{G}(A)$ associated with A, which is a feasible task only under suitable assumptions; see [4]. The *probing estimator* is obtained by using *probing vectors* in (2), i.e., vectors associated with each color whose entries are 0 or 1 depending on the coloring pattern. In [4], the authors show that the error of the probing approximation is bounded by $n \cdot \eta_d$, where η_d decays with a rate that depends on how regular f is over the spectrum of A. The numerical experiments in [4] prove that $O(n)$ bounds are the best we can achieve with this method.

We consider a *stochastic probing* approach, given by the combination of probing techniques with stochastic estimators. The nonzero entries of the *stochastic probing vectors* are the same as the deterministic counterparts, but filled with ± 1 with a uniform distribution (Rademacher entries). This allows to average more than one vector per color,

with an improvement on the convergence related to Hutchinson's estimator. Although this combination is algorithmically quite straightforward and has already been used before by practitioners [1], a detailed analysis was lacking.

In [3], we show for which matrix functions f and matrices A the standard deviation of the stochastic probing estimator can be bounded by quantities of the form $\sqrt{n} \cdot \eta_d$, where η_d has the same asymptotic behavior as the deterministic case. This significantly improves on the linear scaling with the size of the error in the deterministic case, even if just one stochastic vector is associated to any color. As a by-product of our analysis, we refined classical results on sign patterns in the entries of $f(A)$.

Our theoretical findings are illustrated and confirmed by a variety of numerical experiments, where we observed the scaling of the error with the size and compared the performance with other known estimators, indicating when stochastic probing can be the method of choice.

References

[1] E. Aune, D. P. Simpson, and J. Eidsvik. Parameter estimation in high dimensional Gaussian distributions. *Stat. Comput.*, 24:247–263, 2014.

[2] E. N. Epperly, J. A. Tropp, and R. J. Webber. XTrace: making the most of every sample in stochastic trace estimation. *SIAM J. Matrix Anal. Appl.*, 45(1):1–23, 2024.

[3] A. Frommer, M. Rinelli, and M. Schweitzer. Analysis of stochastic probing methods for estimating the trace of functions of sparse symmetric matrices. *Math. Comp.*, Published online, 2024.

[4] A. Frommer, C. Schimmel, and M. Schweitzer. Analysis of probing techniques for sparse approximation and trace estimation of decaying matrix functions. *SIAM J. Matrix Anal. Appl.*, 42(3):1290–1318, 2021.

[5] M. F. Hutchinson. A stochastic estimator of the trace of the influence matrix for Laplacian smoothing splines. *Commun. Stat. Simul.*, 19(2):433–450, 1990.

[6] R. A. Meyer, C. Musco, C. Musco, and D. P. Woodruff. Hutch++: Optimal stochastic trace estimation. In *Symposium on Simplicity in Algorithms (SOSA)*, 142–155. SIAM, 2021.

[7] J. M. Tang and Y. Saad. A probing method for computing the diagonal of a matrix inverse. *Numer. Linear Algebra Appl.*, 19(3):485–501, 2012.