Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Deterministic Graph Spectral Sparsification

Fabricio A. Mendoza Granada[1]
Facultad Politécnica, UNA, Asunción, Paraguay
Sergio Mercado[2]
Facultad Politécnica, UNA, Asunción, Paraguay
Marcos Villagra[3]
Facultad Politécnica, UNA, Asunción, Paraguay

An important technique in data analysis is *principal component analysis* or PCA. Given a covariance matrix S, in PCA we need to compute the eigenvector associated to a greatest eigenvalue of S in order to determine the direction of the so-called principal components [3]. It is well know that computation of eigenvalues of general matrices is expensive, and therefore, several authors use techniques of numerical approximation [5]. Furthermore, computations are more efficient whenever the matrices are sparse.

A technique for constructing *sparse matrices* (matrices with many zeros) is *spectral sparsification*, which is a technique from algebraic graph theory [2]. Given a square and symmetric matrix M which is *dense* (a matrix with few zeros), spectral sparsification approximates M by a sparse matrix $\tilde{M}$ under some well-defined criteria and with no loss of the spectral properties of M. Many other notions of approximation were introduced before, however, we will use an idea of similarity between matrices called *spectral sparsification* introduced by Spielman and Teng [2]. Two matrices are similar if their respective quadratic Laplacian forms are. More formally, let $G = (V, E)$ be a undirected and weighted graph and let $A = (A_{ab})$ in $\mathbb{R}^{n\times n}$ be its adjacency matrix. We define the *Laplacian matrix* $L = (L_{ab})$ of G as

$$L_{ab} = \begin{cases} -A_{ab} & if \quad a \neq b \\ \\ \sum_b A_{ab} & if \quad a = b. \end{cases} \tag{1}$$

The quadratic Laplacian form of L is define as

$$x^T L x = \sum_{(u,v)\in E} (x(u) - x(v))^2, \tag{2}$$

where $x \in \mathbb{R}^V = \{x : \mathbb{V} \to \mathbb{R}\}$. We say a graph $\tilde{G}$ is a σ-spectral sparsifier of G if for all $x \in \mathbb{R}^V$ it holds that

$$\frac{1}{\sigma} x^T \tilde{L} x \leq x^T L x \leq \sigma x^T \tilde{L} x, \tag{3}$$

[1]fabromendoza95@gmail.com
[2]sergiomscmat14@gmail.com
[3]mvillagra@pol.una.py

where $\tilde{L}$ is the Laplacian of $\tilde{G}$.

Spielman and Teng [2] showed that all weighted graphs have a spectral sparsifier with $O(n \log_2^c n)$ edges and can be computed in time $O(m \log_2^c m)$, where n and m are the the number of vertices and edges, respectively. This result has been continuously improved until recently where it was shown that any graph G has a spectral sparsifier of linear size that can be computed in almost linear time, thus, improving all previous results [4].

Intuitively, the fact that $\tilde{G}$ is a σ-spectral sparsifier of G is equivalently to delete edges from G and get zeros in the adjacency matrix of G. With this idea it is possible to reduce the cost of computing eigenvalues without losing any spectral property, that is, the eigenvalues are approximated by some factor σ.

Most of spectral sparsifier methods, including the one of [4], are probabilistic. The currently best deterministic algorithm for spectral sparsification is due to Zouzias [1].

In this work we propose to find a new deterministic method for finding spectral sparsifiers. To that end, we will study several restrictions to the adjacency matrix in order to decrease the number of deleted edges and improve the execution time of the algorithm of Zouzias [1]. This method could be used as a preprocessing step before any other application that requires the computation of eigenvalues, for example, clustering, PCA, etc.

Acknowledgement. S.M. is supported by FEEI- CONACyT-PROCIENCIA research grant PINV15-706 "COMIDENCO". F.A.M.G and M.V. are supported by Conacyt research grant PINV15-208.

References

[1] A. Zouzias. A matrix hyperbolic cosine algorithm and applications. In *Proceedings of the 39th International Colloquium on Automata, Languages, and Programming (ICALP)*, 2012.

[2] D. A. Spielman, and S.H.Teng. *Spectral sparsification of graphs.* SIAM Journal on Computing, 40(4) : 981-1025, 2011.

[3] D. Peña, Análisis de datos Multivariantes, S.A. McGraw-Hil/ INTERAMERICANA DE ESPAÑA. ISBN: 8448136101, 2002.

[4] Y. T. Lee, and H. Sun. *Constructing linear-sized spectral sparsification in almost-linear time.* Foundations of Computer Science (FOCS), 56th Annual Symposium on. IEEE, 2015.

[5] V. Simoncini and D. B. Szyld, Recent computational developments in Krylov subspace methods for linear systems, Numer. Linear Algebra Appl., 14:1-59, 2007.