Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 12, n. 1, 2026.

Trabalho apresentado no XLIV CNMAC, Fundação Getulio Vargas - Rio de Janeiro - RJ, 2025.

Clusterização Espectral Utilizando a Matriz Laplaciana sem Sinal Normalizada

Jean J. D. A. Maria[1] Carla S. Oliveira[2]
ENCE/IBGE, Rio de Janeiro, RJ

João D. G. S. Junior[3]
Colégio Pedro II, Rio de Janeiro, RJ

Clusters são sistemas organizados que consistem em conjuntos de elementos interconectados e que compartilham características ou funções semelhantes, presentes em diversas áreas como a tecnologia, a indústria, a biologia e o meio social. Assim, compreender a estrutura, as interações e as propriedades desses agrupamentos é fundamental para otimizar o desempenho dos sistemas e promover a eficiência em suas respectivas áreas. O crescente interesse da comunidade científica pelo estudo dos clusters reflete a importância de desvendar os mecanismos que regem essas interações e que possibilitam o aprimoramento das estratégias de gerenciamento e desenvolvimento de sistemas complexos.

Para obter os clusters, é necessário modelar as redes a serem trabalhadas. Isso pode ser feito por meio da Teoria dos Grafos. Nesse contexto, a clusterização espectral surge como uma abordagem fundamentada na análise espectral de matrizes associadas ao grafo, permitindo a identificação de agrupamentos.

Os dados são representados por um grafo $G = (V(G), E(G))$, onde $V(G)$ é o conjunto de vértices, que representam os elementos individuais, e $E(G)$ é o conjunto de arestas, que indica as relações entre esses elementos. A estrutura do grafo pode ser representada por diversas matrizes como por exemplo a matriz de adjacência, $A(G)$, cuja entrada a_{ij} assume o valor 1 se existir uma aresta entre os vértices v_i e v_j, e 0 caso contrário, e a matriz diagonal dos graus, $D(G)$, onde o elemento diagonal $d(v_i)$ representa o número de arestas incidentes em v_i. Para um aprofundamento teórico sobre grafos, recomendamos a leitura de [1].

Utilizando as matrizes $A(G)$ e $D(G)$, definem-se, respectivamente, a matriz Laplaciana e a matriz Laplaciana sem sinal como:

$$L(G) = D(G) - A(G) \tag{1}$$

$$Q(G) = D(G) + A(G). \tag{2}$$

Essas matrizes, cujo aprofundamento pode ser feito em [2], desempenham um papel importante quando se fala em clusterização espectral, pois seus autovalores e autovetores fornecem informações essenciais para a identificação de comunidades, conforme discutido em [3].

A partir dessas definições, tem-se a matriz Laplaciana normalizada:

$$L_{\text{sym}} = I - D^{-\frac{1}{2}} A D^{-\frac{1}{2}}, \tag{3}$$

[1]jeanjaimedutra@gmail.com
[2]carla.olivera@ibge.gov.br
[3]joao.junior.2@cp2.edu.br

que também é utilizada em algoritmos de clusterização, como apresentado em [4]. Por outro lado, a matriz Laplaciana sem sinal normalizada é definida como:

$$Q_{\text{sym}} = I + D^{-\frac{1}{2}} A D^{-\frac{1}{2}}, \tag{4}$$

não sendo explorada na literatura, e sua aplicação em técnicas de clusterização permanece pouco investigada.

Neste contexto, o objetivo deste trabalho é investigar a utilização da matriz Laplaciana sem sinal normalizada para a obtenção de uma clusterização, analisando o seu processo e a obtenção dos resultados. Ao final do estudo, será realizada uma comparação entre os resultados obtidos com essa abordagem e os resultados de um método já estabelecido que também emprega a matriz Laplaciana sem sinal. A comparação tomará como base a modularidade e o Rand Index.

Referências

[1] S. Jurkiewicz. **Grafos – Uma Introdução**. Rio de Janeiro: OBMEP, 2009. ISBN: 978-65-250-0689-5.

[2] N. M. M. de Abreu, R. R. Del-Vecchio, C. T. M. Vinagre e D. Stevanovic. **Introdução à Teoria Espectral dos Grafos com Aplicações**. São Carlos, SP: SBMAC, 2007. ISBN: 978-85-86883-65-1.

[3] J. M. S. Costa. **Algoritmos Espectrais de Agrupamento em Redes Sociais de Coautoria**. Dissertação de Mestrado. CEFET/RJ, Rio de Janeiro, 2014.

[4] D. O. Cardoso, J. D. G. da Silva Junior, C. S. Oliveira et al. **Greedy Recursive Spectral Bisection for Modularity-Bound Hierarchical Divisive Community Detection**. Em: *Statistics and Computing* 34 (2024), p. 145. DOI: 10.1007/s11222-024-10451-3.