

Biomathematics/Biomatemática

The Max-Out Min-In Problem: A Tool for Data Analysis in Biomathematics

Jorge Orestes Cerdeira¹, Maria João Martins²,
Marcos Raydan³

¹ Centro de Matemática e Aplicações (CMA) and Departamento de Matemática, Faculdade de Ciências e Tecnologia, Universidade NOVA de Lisboa, Portugal.

² Centro de Estudos Florestais, Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda 1349-017, Lisboa, Portugal.

³ Centro de Matemática e Aplicações, FCT, UNL, Caparica 2829-516, Portugal.

Consider a graph with vertex set V and non-negative weights on the edges. For every subset of vertices S , define $\phi(S)$ to be the sum of the weights of edges with one vertex in S and the other in $V \setminus S$, minus the sum of the weights of the edges with both vertices in S . We consider the problem of finding $S \subseteq V$ for which $\phi(S)$ is maximized. We call this combinatorial optimization problem the max-out min-in problem (MOMIP).

In this talk I will *i*) present a quadratic unconstrained binary optimization formulation for MOMIP; *ii*) prove that the problem is NP-hard; and *iii*) illustrate the applicability of MOMIP in the selection of variables in exploratory data analysis and in the identification of clusters in the context of cluster analysis, and report results on data on gene expressions of patients having different types of tumor.