Eliminating Switching Components in Binary Matrices

Norbert Hantos and Péter Balázs

Let $A = (a_{ij})_{m \times n}$ be a binary matrix of size $m \times n$. The indices $1 \leq i_1 < i_2 \leq m$ and $1 \leq j_1 < j_2 \leq n$ form a *switching component* in A, if either $a_{i_1 j_1} = a_{i_2 j_2} = 1$ and $a_{i_1 j_2} = a_{i_2 j_1} = 0$, or $a_{i_1 j_1} = a_{i_2 j_2} = 0$ and $a_{i_1 j_2} = a_{i_2 j_1} = 1$. In other words, a switching component is a 2×2 sub-matrix of A with exactly two 1-s in its diagonal and two 0-s in its antidiagonal, or vice versa.

Switching components play an important role in image reconstruction and lossless image compression. The absence of switching components in the matrix is a necessary and sufficient condition for the unique reconstruction of the matrix from its horizontal and vertical projections, i.e., from the row and column sums of the matrix. Therefore, in that case the binary image represented by the binary matrix can be stored in a (lossless) compressed form by those two projections. However, if the matrix contains switching components, there is still a chance to reconstruct the matrix uniquely, if properly chosen elements of the matrix are stored as well. One can store, e.g., the positions of 0-s which need to be inverted into 1-s in order to make the matrix switching component free. These positions are called 0-1 *flips*. Then, the aim is to find the minimal number of flips needed to achieve uniqueness. Unfortunately, the problem is generally NP-complete, thus there is no efficient exact algorithm to solve it, unless P = NP.

Switching components are also important in biogeography, where matrices represent the presence or absence of certain species (rows) on certain locations (columns). Here, the so-called nestedness is a relevant measurment of the matrix, which has a strong connection to the 0-1 flips.

In this paper we show that the minimal number of 0-1 flips can be found by determining the proper ordering of the columns regarding a certain filling function, rather than searching through matrix elements and switching components. Based on theoretical results, we develop two deterministic, polynomial-time heuristics to minimize the number of 0-1 flips. We compare those algorithms to another well-known ones in the literature, both on artifical random binary matrices and real-life biogeographical matrices. We conclude that the algorithms searching for proper column permutations perform better, both in the number of 0-1 flips and running time, especially on sparse matrices.