

Partially Optimal Cubic Subspace Clustering

Research Project Machine Learning

Volodymyr Drobitko

Technische Universität Dresden

21.07.2025

- 1 Introduction
- 1 Partial Optimality for Cubic Clique Partition Problem**
- 3 Cubic Subspace Instance Construction
- 4 Experiments and Evaluation
- 5 Research Results

Partial Optimality for Cubic Clique Partition Problem

Extended cost function $c: \binom{S}{3} \cup \binom{S}{2} \cup \emptyset \rightarrow \mathbb{R}$

Partial Optimality for Cubic Clique Partition Problem

Extended cost function $c: \binom{S}{3} \cup \binom{S}{2} \cup \emptyset \rightarrow \mathbb{R}$

Instance of the extended cubic clique partition problem:

$$\min_{y: \binom{S}{2} \rightarrow \{0,1\}} \sum_{\{a,b,c\} \in \binom{S}{3}} c_{\{a,b,c\}} y_{\{a,b\}} y_{\{b,c\}} y_{\{a,c\}} + c_{\{a,b\}} y_{\{a,b\}} + c_{\emptyset}$$

Partial Optimality for Cubic Clique Partition Problem

Extended cost function $c: \binom{S}{3} \cup \binom{S}{2} \cup \emptyset \rightarrow \mathbb{R}$

Instance of the extended cubic clique partition problem:

$$\min_{y: \binom{S}{2} \rightarrow \{0,1\}} \sum_{\{a,b,c\} \in \binom{S}{3}} c_{\{a,b,c\}} y_{\{a,b\}} y_{\{b,c\}} y_{\{a,c\}} + c_{\{a,b\}} y_{\{a,b\}} + c_{\emptyset}$$

Construct **Improving Maps** for the clustering y

Partial Optimality for Cubic Clique Partition Problem

Extended cost function $c: \binom{S}{3} \cup \binom{S}{2} \cup \emptyset \rightarrow \mathbb{R}$

Instance of the extended cubic clique partition problem:

$$\min_{y: \binom{S}{2} \rightarrow \{0,1\}} \sum_{\{a,b,c\} \in \binom{S}{3}} c_{\{a,b,c\}} y_{\{a,b\}} y_{\{b,c\}} y_{\{a,c\}} + c_{\{a,b\}} y_{\{a,b\}} + c_{\emptyset}$$

Construct **Improving Maps** for the clustering y

→ **Partial Optimality Conditions:**

Partial Optimality for Cubic Clique Partition Problem

Extended cost function $c: \binom{S}{3} \cup \binom{S}{2} \cup \emptyset \rightarrow \mathbb{R}$

Instance of the extended cubic clique partition problem:

$$\min_{y: \binom{S}{2} \rightarrow \{0,1\}} \sum_{\{a,b,c\} \in \binom{S}{3}} c_{\{a,b,c\}} y_{\{a,b\}} y_{\{b,c\}} y_{\{a,c\}} + c_{\{a,b\}} y_{\{a,b\}} + c_{\emptyset}$$

Construct **Improving Maps** for the clustering y

→ **Partial Optimality Conditions:**

- ① Subproblem-CUT-condition (cut subset from its complement)

Partial Optimality for Cubic Clique Partition Problem

Extended cost function $c: \binom{S}{3} \cup \binom{S}{2} \cup \emptyset \rightarrow \mathbb{R}$

Instance of the extended cubic clique partition problem:

$$\min_{y: \binom{S}{2} \rightarrow \{0,1\}} \sum_{\{a,b,c\} \in \binom{S}{3}} c_{\{a,b,c\}} y_{\{a,b\}} y_{\{b,c\}} y_{\{a,c\}} + c_{\{a,b\}} y_{\{a,b\}} + c_{\emptyset}$$

Construct **Improving Maps** for the clustering y

→ **Partial Optimality Conditions:**

- 1 Subproblem-CUT-condition (cut subset from its complement)
- 2 CUT-conditions (cut pairs and triples)

Partial Optimality for Cubic Clique Partition Problem

Extended cost function $c: \binom{S}{3} \cup \binom{S}{2} \cup \emptyset \rightarrow \mathbb{R}$

Instance of the extended cubic clique partition problem:

$$\min_{y: \binom{S}{2} \rightarrow \{0,1\}} \sum_{\{a,b,c\} \in \binom{S}{3}} c_{\{a,b,c\}} y_{\{a,b\}} y_{\{b,c\}} y_{\{a,c\}} + c_{\{a,b\}} y_{\{a,b\}} + c_{\emptyset}$$

Construct **Improving Maps** for the clustering y

→ **Partial Optimality Conditions:**

- 1 Subproblem-CUT-condition (cut subset from its complement)
- 2 CUT-conditions (cut pairs and triples)
- 3 JOIN-conditions (join subsets, pairs and triples)

Partial Optimality for Cubic Clique Partition Problem

Extended cost function $c: \binom{S}{3} \cup \binom{S}{2} \cup \emptyset \rightarrow \mathbb{R}$

Instance of the extended cubic clique partition problem:

$$\min_{y: \binom{S}{2} \rightarrow \{0,1\}} \sum_{\{a,b,c\} \in \binom{S}{3}} c_{\{a,b,c\}} y_{\{a,b\}} y_{\{b,c\}} y_{\{a,c\}} + c_{\{a,b\}} y_{\{a,b\}} + c_{\emptyset}$$

Construct **Improving Maps** for the clustering y

→ **Partial Optimality Conditions:**

- 1 Subproblem-CUT-condition (cut subset from its complement)
- 2 CUT-conditions (cut pairs and triples)
- 3 JOIN-conditions (join subsets, pairs and triples)

CUT-conditions can be applied simultaneously

JOIN-conditions cannot be applied simultaneously!

Partial Optimality for Cubic Clique Partition Problem

Extended cost function $c: \binom{S}{3} \cup \binom{S}{2} \cup \emptyset \rightarrow \mathbb{R}$

Instance of the extended cubic clique partition problem:

$$\min_{y: \binom{S}{2} \rightarrow \{0,1\}} \sum_{\{a,b,c\} \in \binom{S}{3}} c_{\{a,b,c\}} y_{\{a,b\}} y_{\{b,c\}} y_{\{a,c\}} + c_{\{a,b\}} y_{\{a,b\}} + c_{\emptyset}$$

Construct **Improving Maps** for the clustering y

→ **Partial Optimality Conditions:**

- 1 Subproblem-CUT-condition (cut subset from its complement)
- 2 CUT-conditions (cut pairs and triples)
- 3 JOIN-conditions (join subsets, pairs and triples)

CUT-conditions can be applied simultaneously

JOIN-conditions cannot be applied simultaneously!

Apply partial optimality conditions → solve subproblems

Partial Optimality Algorithm

Partial Optimality Algorithm:

Input: clustering y without fixed labels

while condition applied **do**

 apply subproblem-CUT-condition exhaustively

 apply one of JOIN-conditions (in effective order)

end while

apply CUT-conditions exhaustively

Output: partially optimal clustering y with some fixed labels

Partial Optimality Algorithm

Partial Optimality Algorithm:

Input: clustering y without fixed labels

while condition applied **do**

 apply subproblem-CUT-condition exhaustively

 apply one of JOIN-conditions (in effective order)

end while

apply CUT-conditions exhaustively

Output: partially optimal clustering y with some fixed labels

Reduction to subproblems:

- 1 Subproblem-CUT-condition: fix CUT labels for element pairs from different sample subsets; solve each subset as an independent problem and accumulate the results in c_\emptyset ;

Partial Optimality Algorithm

Partial Optimality Algorithm:

Input: clustering y without fixed labels

while condition applied **do**

 apply subproblem-CUT-condition exhaustively

 apply one of JOIN-conditions (in effective order)

end while

apply CUT-conditions exhaustively

Output: partially optimal clustering y with some fixed labels

Reduction to subproblems:

- 1 Subproblem-CUT-condition: fix CUT labels for element pairs from different sample subsets; solve each subset as an independent problem and accumulate the results in c_\emptyset ;
- 2 JOIN-Conditions: fix JOIN labels for elements of the sample subset; add the join-cost to c_\emptyset ; solve the problem where the subset is considered as one sample;

Program Structure

TODO

Class Diagram Algorithm implementation in ClusteringProblem

Features: ClusteringProblem is generally defined for all types of Cubic Clique Partition Problem (not necessarily points), cost function + sparse costs!, label computation, cut triples, logs joins and cuts! (add screenshots)

Subproblem-CUT and JOIN-Subset

Subproblem-CUT: cut sample subsets R_1, R_2, \dots, R_k that are only connected via non-negative costs (applied if $k > 1$)

Subproblem-CUT and JOIN-Subset

Subproblem-CUT: cut sample subsets R_1, R_2, \dots, R_k that are only connected via non-negative costs (applied if $k > 1$)

JOIN-Subset: join sample subset R with only non-positive costs if its worst bipartition joining cost is less than or equal to the reward of joining R with \bar{R} (applied if $|R| > 1$)

JOIN-conditions

Overview of the other join-conditions (with pictures)

CUT-conditions

Overview of the cut-conditions (with pictures)