Machine Learning II

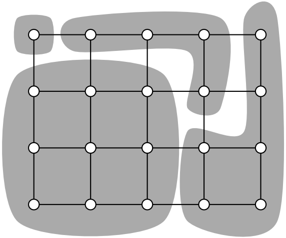
Bjoern Andres

Machine Learning for Computer Vision
TU Dresden

https://mlcv.cs.tu-dresden.de/courses/26-summer/ml2/

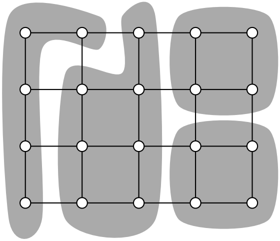
Summer Term 2026

Clustering



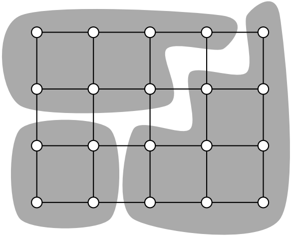
Clustering of a graph $G = (V, E)$

Clustering



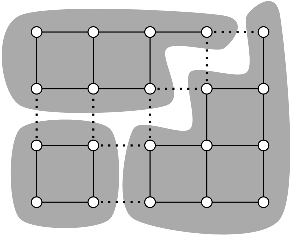
Clustering of a graph $G = (V, E)$

Clustering



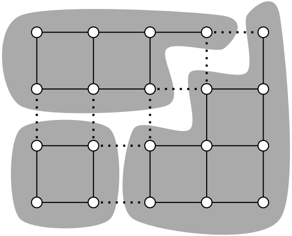
Clustering of a graph $G = (V, E)$

Clustering



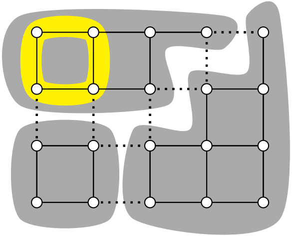
Clustering of a graph $G = (V, E)$

Clustering



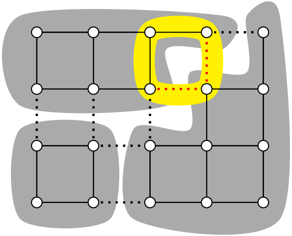
Multicut of a graph $G = (V, E)$

Clustering



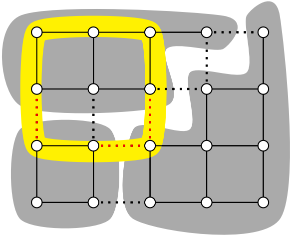
Multicut of a graph $G = (V, E)$

Clustering



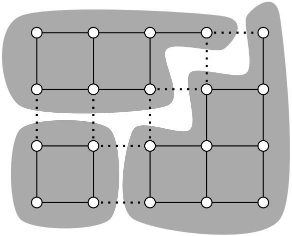
Multicut of a graph $G = (V, E)$

Clustering



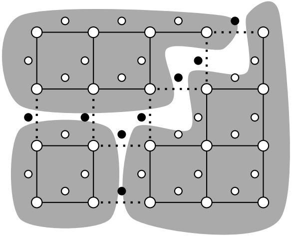
Multicut of a graph $G = (V, E)$

Clustering



Multicut of a graph $G = (V, E)$

Clustering



Multicut of a graph $G = (V, E)$

Clustering

Definition 1. Let $G = (V, E)$ be any graph.

- A subgraph $G' = (U, E')$ of G is called a **cluster** of G iff G' is non-empty, node-induced (i.e. $E' = E \cap \binom{U}{2}$) and connected.
- A partition Π of the node set V is called a **clustering** of G iff for every $U \in \Pi$ the subgraph $(U, E \cap \binom{U}{2})$ of G induced by U is connected (and thus a component of G).
- Let C_G denote the set of all clusterings of G.

Clustering

Definition 2. Let $G = (V, E)$ be any graph.

- For any $M \subseteq E$, M is called a **multicut** of G iff for every cycle (V_C, E_C) of G: $|E_C \cap M| \neq 1$.
- Let M_G denote the set of all multicuts of G.

Lemma 1. Let $G = (V, E)$ be any graph.

- For any clustering Π of G, the set $\{ab \in E \mid \forall U \in \Pi \colon a \notin U \vee b \notin U\}$ $=: M_\Pi$ of those edges that straddle distinct clusters is a multicut of G.
- For any clustering Π of G, the multicut M_Π is said to be **induced** by Π.
- The map $\Pi \mapsto M_\Pi$ from clusterings to induced multicuts is a **bijection** from C_G to M_G.

Clustering

Remark 1. The characteristic function $y \in \{0,1\}^E$ of a multicut $y^{-1}(1)$ makes explicit for every edge $ab = e \in E$ whether the incident nodes a and b belong to the same clusters, $y_e = 0$, or distinct clusters, $y_e = 1$.

Lemma 2. For any $y \in \{0,1\}^E$, the set $y^{-1}(1)$ is a multicut of G iff the following inequalities are satisfied:

$$\forall (V_C, E_C) \in \mathrm{cycles}(G)\ \forall e \in E_C\colon \quad y_e \leq \sum_{e' \in E_C \setminus \{e\}} y_{e'} \tag{1}$$

Learning and inferring clusterings:

- Instead of the problem of learning and inferring clusterings, we consider the problem of learning and inferring multicuts. By Lemma 1, this is w.l.o.g..
- We reduce the problem of learning and inferring multicuts to the problem of learning and inferring decisions, by defining constrained data (S, X, x, Y) with

$$S = E \tag{2}$$

$$\mathcal{Y} = \left\{ y \in \{0,1\}^E \;\middle|\; \forall (V_C, E_C) \in \text{cycles}(G) \; \forall e \in E_C \colon y_e \leq \sum_{e' \in E_C \setminus \{e\}} y_{e'} \right\} . \tag{3}$$

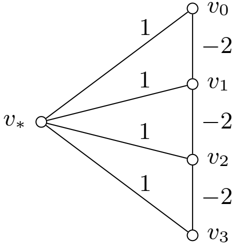
- As inference problem, we obtain the **(minimum cost) multicut problem**

$$\min_{y \in \{0,1\}^E} \quad \sum_{e \in E} \underbrace{(-f_\theta(x_e))}_{=: c_e} y_e \tag{4}$$

$$\text{subject to} \quad \forall (V_C, E_C) \in \text{cycles}(G) \; \forall e \in E_C \colon \; y_e \leq \sum_{e' \in E_C \setminus \{e\}} y_{e'} \; . \tag{5}$$

Theorem 1. The minimum cost multicut problem is NP-hard.

Proof. (Sketch) By reduction of **maximum independent set**:

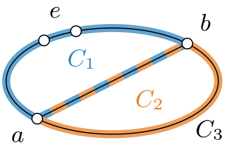


□

Clustering

Theorem 2 (Chopra and Rao 1993)**.** In the minimum cost multicut problem, the inequalities (1) are redundant for chordal cycles.

Proof. (Sketch) Consider a cycle C_3 with a chord ab.



By induction, (1) holds for C_1 and C_2. For any $e \in C_1 \setminus \{ab\}$:

$$\begin{aligned} y_e &\leq \sum_{e' \in C_1 \setminus \{e\}} y_{e'} = y_{ab} + \sum_{e' \in C_1 \setminus \{e, ab\}} y_{e'} \\ &\leq \sum_{e' \in C_2 \setminus \{ab\}} y_{e'} + \sum_{e' \in C_1 \setminus \{e, ab\}} y_{e'} = \sum_{e' \in C_3 \setminus \{e\}} y_{e'} \ . \end{aligned}$$

For any $e \in C_2 \setminus \{ab\}$, the argument is analogous. ☐

Corollary 1 (Chopra and Rao 1993)**.** The multicut problem for a complete graph is isomorphic to the clique partition problem for the node set.

Proof. (Sketch) In a complete graph $G = (V, E)$, the chordless cycles are precisely the triangles.

The inequalities (1) for all triangles are written equivalenly as

$$\forall a \in V \; \forall b \in V \setminus \{a\} \; \forall c \in V \setminus \{a, b\} \colon \quad y_{ac} \leq y_{ab} + y_{bc} \; .$$

Substituting $1 - y'$ for y (i.e. swapping the roles of 0 and 1), yields the transitivity constraints of the clique partition problem:

$$\forall a \in V \; \forall b \in V \setminus \{a\} \; \forall c \in V \setminus \{a, b\} \colon \quad y'_{ab} + y'_{bc} - 1 \leq y'_{ac} \; .$$

□

Definition 3. For any graph $G = (V, E)$ and any $c \in \mathbb{R}^E$, the **cycle relaxation** of the minimum cost multicut problem is

$$\min_{y \in \mathbb{R}^E} \quad \langle c, y \rangle \tag{6}$$

$$\text{subject to} \quad \forall (V_C, E_C) \in \text{cycles}(G)\ \forall e \in E_C \colon\ y_e \leq \sum_{e' \in E_C \setminus \{e\}} y_{e'} \tag{7}$$

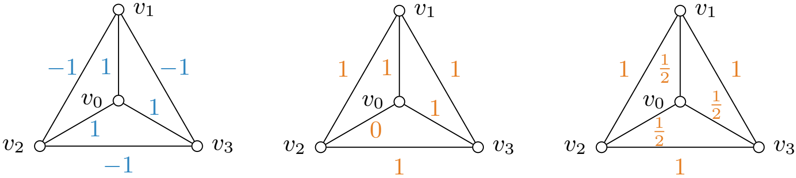
$$\forall e \in E \colon \quad 0 \leq y_e \leq 1\ . \tag{8}$$

Remark 2. The cycle relaxation of the minimum cost multicut problem

- can be solved by the simplex algorithm
- yields a lower bound on the minimum cost
- can be solved efficiently
- is not tight (see below)

Clustering

Example 1. Instance of the multicut problem for an odd wheel:



Instance

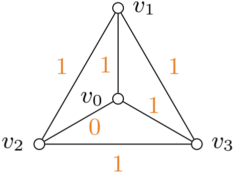
Solution
$\langle c, y \rangle = -1$

Solution to relaxation
$\langle c, y \rangle = -\frac{3}{2}$

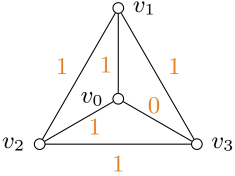
Remark 3. Approaches to working with LP relaxations that are not tight:

- branching
- cutting planes

Example 2. Branching on $y_{v_0 v_2}$

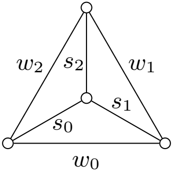


Solution to relaxation with $y_{v_0 v_2} = 0$
$\langle c, y \rangle = -1$



Solution to relaxation with $y_{v_0 v_2} = 1$
$\langle c, y \rangle = -1$

Example 3. An **odd wheel inequality** as a **Chvátal-Gomory cutting plane**:



From the cycle relaxation of the multicut problem, the inequalities

$$\begin{aligned} y_{w_0} &\leq y_{s_0} + y_{s_1} & \qquad y_{w_0} &\leq 1 \\ y_{w_1} &\leq y_{s_1} + y_{s_2} & \qquad y_{w_1} &\leq 1 \\ y_{w_2} &\leq y_{s_2} + y_{s_0} & \qquad y_{w_2} &\leq 1 \end{aligned}$$

together imply

$$\sum_j y_{w_j} - \sum_j y_{s_j} \leq \frac{3}{2}$$

For any integral feasible solution y, the lhs. is an integer. Thus:

$$\sum_j y_{w_j} - \sum_j y_{s_j} \leq \left\lfloor \frac{3}{2} \right\rfloor \ .$$