Machine Learning II

Jannik Irmai, Jannik Presberger, David Stein, Bjoern Andres

Machine Learning for Computer Vision
TU Dresden

https://mlcv.cs.tu-dresden.de/courses/25-summer/ml2/

Summer Term 2025

Contents.

- In this part of the course, we discuss a technique for solving combinatorial optimization problems **partially** and **efficiently**: the construction of **improving maps**.
- We concentrate on three NP-hard problems that arise as inference problems in the field of machine learning: the clique partition problem for clustering, the linear ordering problem, and the graphical model inference problem.

Definition 1. Let $Y \neq \varnothing$ finite, $\varphi \colon Y \to \mathbb{R}$ and $\sigma \colon Y \to Y$. We call σ **improving** for the problem $\min\{\varphi(y) \mid y \in Y\}$ iff $\varphi \circ \sigma \leq \varphi$.

Lemma 1. Let $Y \neq \varnothing$ finite and $\varphi \colon Y \to \mathbb{R}$. Let $\sigma \colon Y \to Y$ improving for the problem $\min\{\varphi(y) \mid y \in Y\}$. If $Q \subseteq Y$ and $\sigma(Y) \subseteq Q$, there exists a solution y^* such that $y^* \in Q$.

Proof. A solution y' exists because Y is non-empty and finite. $y^* := \sigma(y')$ is also a solution because σ is improving. Moreover, $y^* \in Q$ because $\sigma(Y) \subseteq Q$. $\square$

Corollary 1. Let $S \neq \varnothing$ finite, $Y \subseteq \{0,1\}^S$ and $\varphi\colon Y \to \mathbb{R}$. Let $s \in S$ and $q \in \{0,1\}$. If $\sigma\colon Y \to Y$ is improving for the problem $\min\{\varphi(y) \mid y \in Y\}$ such that $\forall y \in Y\colon \sigma(y)_s = q$, there exists a solution y^* such that $y^*_s = q$.

Remark 1. If we can construct such an improving map, we can fix the variable y^*_s to q without compromising optimality.

Definition 2. For any $A \neq \varnothing$ finite, any $c\colon \binom{A}{2} \to \mathbb{R}$,

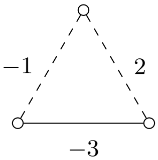
$$Y_A := \left\{ y\colon \binom{A}{2} \to \{0,1\} \;\middle|\; \forall a \in A \; \forall b \in A \setminus \{a\} \; \forall c \in A \setminus \{a,b\} \colon \right.$$
$$\left. y_{ab} + y_{bc} - 1 \leq y_{ac} \right\} \qquad (1)$$

and $\varphi_c\colon Y_A \to \mathbb{R}\colon y \mapsto \langle c, y \rangle$,

$$\min\{\varphi_c(y) \mid y \in Y_A\} \qquad (2)$$

is called the instance of the **(clique) partition problem** wrt. A and c, which we abbreviate as $\mathrm{CPP}(A, c)$.

Example 1.



For any set A and any $U \subseteq A$, we write

$$\partial U := \Big\{ \{u, a\} \in \tbinom{A}{2} \;\Big|\; u \in U \wedge a \notin U \Big\} \quad . \tag{3}$$

Definition 3. Let $A \neq \varnothing$ finite and $U \subseteq A$.

- The **elementary cut map** wrt. U is the $\sigma_U \colon Y_A \to Y_A$ such that $\forall y \in Y_A\ \forall \{a,b\} \in \binom{A}{2}$:

$$\sigma_U(y)_{ab} = \begin{cases} 0 & \text{if } \{a,b\} \in \partial U \\ y_{ab} & \text{otherwise} \end{cases} . \tag{4}$$

- The **elementary join map** wrt. U is the $\sigma'_U \colon Y_A \to Y_A$ such that $\forall y \in Y_A\ \forall \{a,b\} \in \binom{A}{2}$:

$$\sigma'_U(y)_{ab} = \begin{cases} 1 & \text{if } \{a,b\} \in \binom{U}{2} \\ 1 & \text{if } a \in U \wedge \exists u \in U \colon y_{ub} = 1 \\ 1 & \text{if } b \in U \wedge \exists u \in U \colon y_{ua} = 1 \\ 1 & \text{if } (\exists u \in U \colon y_{ua} = 1) \wedge \\ & \quad (\exists u \in U \colon y_{ub} = 1) \\ y_{ab} & \text{otherwise} \end{cases} . \tag{5}$$

Remark 2. σ_U is well-defined, i.e. $\sigma_U(Y_A) \subseteq Y_A$. σ'_U is well-defined. $\sigma'_U \circ \sigma_U$ is well-defined.

To begin with, we establish a trivial partial optimality condition for the CPP:

Lemma 2. Let $A \neq \varnothing$ finite and $c\colon \binom{A}{2} \to \mathbb{R}$. If there exists $U \subseteq A$ such that

$$\forall \{a, b\} \in \partial U\colon \quad 0 \leq c_{ab} \ , \tag{6}$$

there exists a solution y^* to $\mathrm{CPP}(A, c)$ such that

$$\forall \{a, b\} \in \partial U\colon \quad y^*_{ab} = 0 \ . \tag{7}$$

Proof. For any $y \in Y_A$, $\sigma_U(y)$ satisfies (7). Moreover, σ_U is improving for $\mathrm{CPP}(A, c)$ because for any $y \in Y_A$ and $y' := \sigma_U(y)$:

$$
\begin{aligned}
\varphi_c(y') - \varphi_c(y) &= \sum_{\{a,b\} \in \binom{A}{2}} c_{ab}\, y'_{ab} - \sum_{\{a,b\} \in \binom{A}{2}} c_{ab}\, y_{ab} && (8)\\
&= \sum_{\{a,b\} \in \binom{A}{2}} c_{ab}(y'_{ab} - y_{ab}) && (9)\\
&= \sum_{\{a,b\} \in \partial U} c_{ab}(0 - y_{ab}) && (10)\\
&= -\sum_{\{a,b\} \in \partial U} c_{ab}\, y_{ab} && (11)\\
&\overset{(6)}{\leq} 0 \ . && (12)
\end{aligned}
$$

The assertion follows by Lemma 1. $\square$

For any $r \in \mathbb{R}$, we write

$$[r]_+ := \begin{cases} r & \text{if } r \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

$$[r]_- := \begin{cases} 0 & \text{if } r \geq 0 \\ -r & \text{otherwise} \end{cases} \quad . \tag{14}$$

Next, we establish a less trivial partial optimality condition for the CPP:

Proposition 1. Let $A \neq \varnothing$ finite and $c\colon \binom{A}{2} \to \mathbb{R}$. If there exist $U \subseteq A$ and $\{u, v\} \in \partial U$ such that

$$\sum_{\{a,b\} \in \partial U \setminus \{\{u,v\}\}} [c_{ab}]_- \leq c_{uv} \quad , \tag{15}$$

there exists a solution y^* to $\mathrm{CPP}(A, c)$ such that $y^*_{uv} = 0$.

Proof. Let $\xi \colon Y_A \to Y_A$ such that for all $y \in Y_A$:

$$\xi(y) = \begin{cases} y & \text{if } y_{uv} = 0 \\ \sigma_U(y) & \text{otherwise} \end{cases} . \tag{16}$$

For any $y \in Y_A$ and $y' := \xi(y)$, we have $y'_{uv} = 0$.

Moreover, ξ is improving for $\mathrm{CPP}(A, c)$ because for all $y \in Y_A$ and $y' := \xi(y)$, the following holds: If $y_{ab} = 0$ then $\varphi_c(y') - \varphi_c(y) = \varphi_c(y) - \varphi_c(y) = 0 \leq 0$. Otherwise:

$$\begin{aligned}
\varphi_c(y') - \varphi_c(y) &= \sum_{\{a,b\} \in \binom{A}{2}} c_{ab}(y'_{ab} - y_{ab}) && (17)\\
&= c_{uv}(0 - 1) + \sum_{\{a,b\} \in \partial U \setminus \{\{u,v\}\}} c_{ab}(0 - y_{ab}) && (18)\\
&= -c_{uv} - \sum_{\{a,b\} \in \partial U \setminus \{\{u,v\}\}} c_{ab}\, y_{ab} && (19)\\
&\leq -c_{uv} + \sum_{\{a,b\} \in \partial U \setminus \{\{u,v\}\}} [c_{ab}]_- && (20)\\
&\overset{(15)}{\leq} 0 \ . && (21)
\end{aligned}$$

The assertion follows by Lemma 1. $\square$

Next, we establish a non-trivial partial optimality condition for the CPP:

Lemma 3. Let $A \neq \varnothing$ finite and $c \colon \binom{A}{2} \to \mathbb{R}$. If there exist $U \subseteq A$ such that

$$\sum_{\{u,a\} \in \partial U} [c_{ua}]_{-} \leq \min_{\{s,t\} \in \binom{U}{2}} \min_{\substack{y \in Y_U | \\ y_{st}=0}} \sum_{\{u,v\} \in \binom{U}{2}} (-c_{uv})(1 - y_{uv}) \ , \qquad (22)$$

there exists a solution y^* to $\mathrm{CPP}(A, c)$ such that $\forall \{u,v\} \in \binom{U}{2} \colon y^*_{uv} = 1.$

Proof. Let $\xi\colon Y_A \to Y_A$ such that for all $y \in Y_A$:

$$\xi(y) := \begin{cases} (\sigma'_U \circ \sigma_U)(y) & \text{if } \exists \{u,v\} \in \binom{U}{2}\colon y_{uv} = 0 \\ y & \text{otherwise} \end{cases} \quad . \tag{23}$$

For any $y \in Y_A$, $y' := \xi(y)$ and all $\{u,v\} \in \binom{U}{2}$, we have $y'_{uv} = 1$.

Moreover, ξ is improving because for all $y \in Y_A$ and $y' := \xi(y)$, the following condition holds: If $\forall \{u,v\} \in \binom{U}{2}\colon y_{uv} = 1$ then $\varphi_c(y') - \varphi_c(y) = \varphi_c(y) - \varphi_c(y) = 0 \leq 0$. Otherwise:

$$\varphi_c(y') - \varphi_c(y) = \sum_{\{u,a\} \in \partial U} c_{ua}(0 - y_{ua}) + \sum_{\{u,v\} \in \binom{U}{2}} c_{uv}(1 - y_{uv}) \tag{24}$$

$$\leq \sum_{\{u,a\} \in \partial U} [c_{ua}]_- + \max_{\{s,t\} \in \binom{U}{2}} \max_{\substack{y \in Y_U | \\ y_{st} = 0}} \sum_{\{u,v\} \in \binom{U}{2}} c_{uv}(1 - y_{uv}) \tag{25}$$

$$\leq \sum_{\{u,a\} \in \partial U} [c_{ua}]_- - \min_{\{s,t\} \in \binom{U}{2}} \min_{\substack{y \in Y_U | \\ y_{st} = 0}} \sum_{\{u,v\} \in \binom{U}{2}} (-c_{uv})(1 - y_{uv}) \tag{26}$$

$$\overset{(22)}{\leq} 0 \ . \tag{27}$$

The assertion follows by Lemma 1. □

Even if set $U \subseteq A$ is given, Condition (22) of Lemma 3 cannot be checked efficiently: In general, the calculation of

$$\min_{\{s,t\} \in \binom{U}{2}} \min_{\substack{y \in Y_U | \\ y_{st}=0}} \sum_{\{u,v\} \in \binom{U}{2}} (-c_{uv})(1 - y_{uv}) \tag{28}$$

requires solving CPPs with the additional constraint $y_{st} = 0$.

However, in the special case where $\forall \{u,v\} \in \binom{U}{2}$: $c_{uv} \leq 0$, these problems become minimum st-cut problems that can be solved efficiently.

Hence, an idea toward applying Lemma 3 algorithmically is to work in two steps:

1. to heuristically search for a set U such that
   - inside U, all costs are non-positive
   - on the boundary of U, the sum of the negative costs is large.
2. to efficiently test (22) from Lemma 3 for these sets U.