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*.

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.

Contents. In this part of the course, we construct improving maps for the clique partition problem, an inference problem for clustering.

References.

- Stein D., Di Gregorio S. and Andres B. Partial Optimality in Cubic Correlation Clustering. ICML 2023
- Lange J.-H., Andres B. and Swoboda P. Combinatorial persistency criteria for multicut and max-cut. CVPR 2019
- Lange J.-H., Karrenbauer A. and Andres B. Partial Optimality and Fast Lower Bounds for Weighted Correlation Clustering. ICML 2018
- Alush, A. and Goldberger, J. Ensemble segmentation using efficient integer linear programming. TPAMI, 34(10):1966–1977, 2012

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

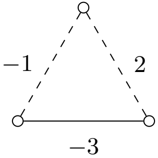
$$Y_A := \left\{ y \colon \binom{A}{2} \to \{0,1\} \;\middle|\; \forall a \in A \;\forall b \in A \setminus \{a\} \;\forall c \in A \setminus \{a,b\} \colon \right.$$
$$\left. y_{ab} + y_{bc} - 1 \leq y_{ac} \right\} \qquad (1)$$

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

$$\partial U := \left\{ \{u, a\} \in \binom{A}{2} \;\middle|\; u \in U \wedge a \notin U \right\} \quad . \tag{3}$$

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 $\forall y \in Y_A\ \forall \{a,b\} \in \binom{A}{2}$:

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

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 $A \neq \emptyset$ 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. □

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 $y \in Y_A$:

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

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. $\square$

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

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

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.

Contents: In this part of the course, we discuss partial optimality in the graphical model inference problem.

References:

- E. Boros, P. L. Hammer, X. Sun: Network flows and minimization of quadratic pseudo-Boolean functions. RUTCOR Research Report 17-1991
- E. Boros, P. L. Hammer: Pseudo-Boolean optimization. Discrete Applied Mathematics 123(1–3): 155–225 (2002)
- E. Boros, P. L. Hammer, R. Sun, G. Tavares: A max-flow approach to improved lower bounds for quadratic unconstrained binary optimization (QUBO). Discrete Optimization 5(2): 501–529 (2008)

Definition 4. For any $n \in \mathbb{N}$, any $d \in \{0, \dots, n\}$, let

$$J_{nd} := \bigcup_{m=0}^{d} \tbinom{\{1,\dots,n\}}{d} \qquad C_{nd} := \mathbb{R}^{J_{nd}} \tag{29}$$

and call any $c \in C_{nd}$ an n-variate **multi-linear polynomial form** of degree at most d.

Example. For $n = d = 2$, we have

$$\begin{aligned} J_{22} &= \bigcup_{m=0}^{2} \tbinom{\{1,2\}}{m} \\ &= \tbinom{\{1,2\}}{0} \cup \tbinom{\{1,2\}}{1} \cup \tbinom{\{1,2\}}{2} \\ &= \{\varnothing\} \cup \{\{1\},\{2\}\} \cup \{\{1,2\}\} \\ &= \{\varnothing, \{1\}, \{2\}, \{1,2\}\} \end{aligned}$$

Definition 5. For any $f\colon A \to B$ and any $n \in \mathbb{N}$, f is called an n-**variate pseudo-Boolean function (PBF)** iff $A = \{0,1\}^n$ and $B \subseteq \mathbb{R}$. For any $f\colon A \to B$, f is called a PBF iff f is an n-variate PBF for some $n \in \mathbb{N}$.

Definition 6. For any $n \in \mathbb{N}$, any $d \in \{0, \dots, n\}$ and any $c \in C_{nd}$, the function f_c defined below is called the **PBF defined by** c.

$$f_c\colon \quad \{0,1\}^n \to \mathbb{R}\colon \quad x \mapsto \sum_{m=0}^{d} \sum_{J \in \binom{\{1,\dots,n\}}{m}} c_J \prod_{j \in J} x_j \tag{30}$$

Example. For any $c \in C_{22}$, f_c is such that for all $x \in \{0,1\}^2$:

$$f_c(x_1, x_2) = c_\varnothing + c_{\{1\}} x_1 + c_{\{2\}} x_2 + c_{\{1,2\}} x_1 x_2 \quad .$$

Lemma 4. Every PBF has a unique multi-linear polynomial form. More precisely,

$$\forall n \in \mathbb{N} \quad \forall f : \{0,1\}^n \to \mathbb{R} \quad \exists_1 c \in C_{nn} \quad f = f_c \ . \tag{31}$$

Example. For $n = d = 2$ and any $f : \{0,1\}^2 \to \mathbb{R}$, the existence of a $c \in C_{22}$ such that $f = f_c$ means

$$\forall x \in \{0,1\}^2 \colon \quad f(x_1, x_2) = c_\varnothing + c_{\{1\}} x_1 + c_{\{2\}} x_2 + c_{\{1,2\}} x_1 x_2 \ .$$

Explicitly,

$$\begin{aligned} f(0,0) &= c_\varnothing \\ f(1,0) &= c_\varnothing + c_{\{1\}} \\ f(0,1) &= c_\varnothing \qquad\quad + c_{\{2\}} \\ f(1,1) &= c_\varnothing + c_{\{1\}} + c_{\{2\}} + c_{\{1,2\}} \ . \end{aligned}$$

In this example, a suitable c exists and is defined uniquely by f.

Proof. For any $J \subseteq \{1, \dots, n\}$, let $x^J \in \{0,1\}^n$ such that for all $j \in \{1, \dots, n\}$:

$$x_j^J = \begin{cases} 1 & \text{if } j \in J \\ 0 & \text{otherwise} \end{cases} \quad .$$

Now,

$$\forall x \in \{0,1\}^n \colon \quad f(x) = \sum_{J \subseteq \{1,\dots,n\}} c_J \prod_{j \in J} x_j$$

is written equivalently as

$$f(x^\varnothing) = c_\varnothing$$
$$\forall J \neq \varnothing \colon \quad f(x^J) = c_J + \sum_{J' \subset J} c_{J'} \quad .$$

Thus, c is defined uniquely (by induction over the cardinality of J). $\square$

Definition 7. For any $n \in \mathbb{N}$ and any $d \in \{0, \dots, n\}$, let

$$F_{nd} := \{f : \{0,1\}^n \to \mathbb{R} \mid \exists c \in C_{nd} : \; f = f_c\} \tag{32}$$

and call any $f \in F_{nd}$ an n-**variate PBF of degree at most** d. In addition, call any $f \in F_{n2}$ a **quadratic PBF (QPBF)**.

Remark 3. For any $n \in \mathbb{N}$, F_{nn} is the set of all n-variate PBFs (by Lemma 4).

Definition 8.

- For any $n \in \mathbb{N}$ and any $f : \{0,1\}^n \to \mathbb{R}$, call

$$\min \{f(x) \mid x \in \{0,1\}^n\} \tag{33}$$

  the instance of the **pseudo-boolean optimization (PBO)** problem wrt. f.

- For any $n \in \mathbb{N}$ and any $f \in F_{n2}$, call

$$\min \{f(x) \mid x \in \{0,1\}^n\} \tag{34}$$

  the instance of the **quadratic pseudo-boolean optimization (QPBO)** problem wrt. f.

Is QPBO less complex than PBO?

Definition 9. For any $n \in \mathbb{N}$ and any $c \in C_{nn}$, define the **size** of c as

$$\mathrm{size}(c) := \sum_{J \subseteq \{1,\ldots,n\}:\ c_J \neq 0} |J| \ . \tag{35}$$

Lemma 5. For any $x, y, z \in \{0, 1\}$:

$$z = xy \quad \Leftrightarrow \quad xy - 2xz - 2yz + 3z = 0 \ , \tag{36}$$

$$z \neq xy \quad \Leftrightarrow \quad xy - 2xz - 2yz + 3z > 0 \ . \tag{37}$$

Proof. By verifying equivalence for all eight cases. □

Algorithm 1 (Boros and Hammer 2001).

Input: $c \in C_{nn}$
Output: $c' \in C_{n2}$
$M := 1 + 2\sum_{J \subseteq \{1,\dots,n\}} |c_J|$
$m := n$
$c^m := c$
while there exists a $J \subseteq \{1,\dots,n\}$ such that $|J| > 2$ and $c^m_J \neq 0$
  Choose $j, k \in J$ such that $j \neq k$
  $c^{m+1} := c^m$
  $c^{m+1}_{\{j,k\}} := c^{m+1}_{\{j,k\}} + M$
  $c^{m+1}_{\{j,m+1\}} := -2M$
  $c^{m+1}_{\{k,m+1\}} := -2M$
  $c^{m+1}_{\{m+1\}} := 3M$
  for all $\{j,k\} \subseteq J' \subseteq \{1,\dots,n\}$ such that $c^{m+1}_{J'} \neq 0$
    $c^{m+1}_{J' - \{j,k\} \cup \{m+1\}} := c^{m+1}_{J'}$
    $c^{m+1}_{J'} := 0$
  $m := m + 1$
$c' := c^m$

Theorem 1.

- Algorithm 1 terminates in polynomial time in $\mathrm{size}(c)$.
- $\mathrm{size}(c')$ is polynomially bounded by $\mathrm{size}(c)$.
- The multi-linear quadratic form c' is such that $\forall \hat{x} \in \mathbb{R}^n$:

$$\begin{aligned} & \hat{x} \in \operatorname*{argmin}_{x \in \{0,1\}^n} f_c(x) \\ \Rightarrow \quad & \exists \hat{x}' \in \{0,1\}^m \left(\hat{x}'_{\{1,\ldots,n\}} = \hat{x}_{\{1,\ldots,n\}} \wedge \hat{x}' \in \operatorname*{argmin}_{x' \in \{0,1\}^m} f_{c'}(x') \right) \quad . \end{aligned} \tag{38}$$

Proof. The algorithm replaces the occurrence of $x_j x_k$ by x_{m+1} and adds the form $M(x_j x_k - 2x_j x_{m+1} - 2x_k x_{m+1} + 3x_{m+1})$.

- If $x_{m+1} = x_j x_k$,

$$f^{m+1}(x_1, \ldots, x_{m+1}) = f^m(x_1, \ldots, x_n) \leq \max_{x' \in \{0,1\}^n} f^m(x') < M/2 \ .$$

- If $x_{m+1} \neq x_j x_k$,

$$f^{m+1}(x_1, \ldots, x_{m+1}) \geq M/2$$

(by Lemma 5 and by definition of M).

For every iteration m,

$$|\{J \subseteq \{1, \ldots, n\} | |J| > 2 \wedge c_J^{m+1} \neq 0\}| < |\{J \subseteq \{1, \ldots, n\} | |J| > 2 \wedge c_J^{m} \neq 0\}|$$

which proves the complexity claims. $\square$

Summary:

- Every PBF has a unique multi-linear polynomial form.
- PBO is polynomially reducible to QPBO.

Definition 10. For any $n \in \mathbb{N}$ and any $d \in \{0, \dots, n\}$, let

$$K^+_{nd} := \{(K^1, K^0) \mid K^1, K^0 \subseteq \{1, \dots, n\} \wedge K^1 \cap K^0 = \varnothing \wedge |K^1| + |K^0| = d\}$$

$$J^+_{nd} := \bigcup_{m=0}^{d} K^+_{nm}$$

$$C^+_{nd} := \{c : J^+_{nd} \to \mathbb{R} \mid \forall j \in J^+_{nd} \setminus \{(\varnothing, \varnothing)\} : \ 0 \leq c_j\}$$

and call any $c \in C^+_{nd}$ an n-variate **posiform** of degree at most d.

Example. For $n = d = 2$,

$$\begin{aligned} J^+_{22} = &\ \{\ (\varnothing, \varnothing)\ \} \\ &\cup \{\ (\{1\}, \varnothing),\ (\varnothing, \{1\}),\ (\{2\}, \varnothing),\ (\varnothing, \{2\})\ \} \\ &\cup \{\ (\{1,2\}, \varnothing),\ (\{1\}, \{2\}),\ (\{2\}, \{1\}),\ (\varnothing, \{1,2\})\ \} \end{aligned}$$

Definition 11. For any $n \in \mathbb{N}$, any $d \in \{0, \dots, n\}$ and any $c \in C^+_{nd}$, $f_c : \{0,1\}^n \to \mathbb{R}$ such that

$$\forall x \in \{0,1\}^n \qquad f_c(x) := \sum_{(J^1,J^0) \in J^+_{nd}} c_{J^1 J^0} \prod_{j \in J^1} x_j \prod_{j' \in J^0} (1 - x'_j) \tag{39}$$

is called the **PBF defined by** c.

Example. For any $c \in C^+_{22}$, $f_c : \{0,1\}^2 \to \mathbb{R}$ is such that $\forall x \in \{0,1\}^2$:

$$\begin{aligned} f(x) = \quad & c_{\varnothing\varnothing} \\ & + c_{\{1\}\varnothing} x_1 + c_{\varnothing\{1\}}(1 - x_1) + c_{\{2\}\varnothing} x_2 + c_{\varnothing\{2\}}(1 - x_2) \\ & + c_{\{1,2\}\varnothing} x_1 x_2 + c_{\{1\}\{2\}} x_1 (1 - x_2) + c_{\{2\}\{1\}} (1 - x_1) x_2 \\ & \quad + c_{\varnothing\{1,2\}} (1 - x_1)(1 - x_2) \ . \end{aligned}$$

Definition 12. For any $n \in \mathbb{N}$ and any $f : \{0,1\}^n \to \mathbb{R}$, the posiform defined by

$$\forall x \in \{0,1\}^n: \quad K_x^1 := \{j \in \{1,\dots,n\} | x_j = 1\}$$
$$K_x^0 := \{j \in \{1,\dots,n\} | x_j = 0\}$$

and

$$J := \{(\varnothing, \varnothing)\} \cup \bigcup_{x \in \{0,1\}^n} \{(K_x^1, K_x^0)\}$$

and $c : J \to \mathbb{R}$ such that

$$c_{\varnothing\varnothing} := \min_{x \in \{0,1\}^n} f(x)$$
$$\forall x \in \{0,1\}^n \quad c_{K_x^1 K_x^0} := f(x) - c_{\varnothing\varnothing}$$

is called **min-term posiform** of f.

Lemma 6. For any $n \in \mathbb{N}$ and any $f : \{0,1\}^n \to \mathbb{R}$, the min-term posiform c of f is such that $f_c = f$.

Corollary 2. For any $n \in \mathbb{N}$ and any $f : \{0,1\}^n \to \mathbb{R}$, there exists a posiform $c \in C^+_{nn}$ such that $f_c = f$.

Proof. Let $n \in \mathbb{N}$ and $f : \{0,1\}^n \to \mathbb{R}$. Moreover, let $c : J \to \mathbb{R}$ the min-term posiform of f.

c is a posiform (by definition).

Let $g : \{0,1\}^n \to \mathbb{R}$ be the PBF defined by this posiform.

Then, for any $x \in \{0,1\}^n$,

$$(J^1, J^0) \in \{(\varnothing, \varnothing), (K_x^1, K_x^0)\} \subseteq J$$

are the only elements of J for which

$$0 \neq \prod_{j \in J^1} x_j \prod_{j' \in J^0} (1 - x'_j) = 1 \enspace .$$

Thus,

$$\begin{aligned} \forall x \in \{0,1\}^n \qquad g(x) &= c_{\varnothing\varnothing} + c_{K_x^1 K_x^0} \\ &= c_{\varnothing\varnothing} + f(x) - c_{\varnothing\varnothing} \qquad \text{(by definition of } c\text{)} \\ &= f(x) \enspace . \end{aligned}$$

□

Remark 4. Unlike multi-linear polynomial forms, posiforms of PBFs need not be unique, e.g., $x_1 = x_1 x_2 + x_1(1 - x_2)$.

Definition 13. For any $n \in \mathbb{N}$, any $f : \{0,1\}^n \to \mathbb{R}$ and any $d \in \{0, \dots, n\}$, let

$$C^+_{nd}(f) := \left\{ c \in C^+_{nd} \mid f_c = f \right\} \quad . \tag{40}$$

Remark 5. For any $n \in \mathbb{N}$ and any $f : \{0,1\}^n \to \mathbb{R}$, $C^+_{nn}(f)$ contains at least the min-term posiform of f.

Lemma 7.

$$\forall n \in \mathbb{N} \quad \forall f : \{0,1\}^n \to \mathbb{R} \quad \forall c \in C^+_{nn}(f) \quad \forall x \in \{0,1\}^n : \quad c_{\varnothing\varnothing} \leq f(x) \ .$$

Proof. By definition, we have, for all $x \in \{0,1\}^n$,

$$\begin{aligned} f(x) &= \sum_{m=0}^{d} \sum_{(K^1,K^0) \in K^+_{nm}} c_{K^1 K^0} \prod_{j \in K^1} x_j \prod_{j' \in K^0} (1 - x'_j) \\ &= c_{\varnothing\varnothing} + \sum_{m=1}^{d} \sum_{(K^1,K^0) \in K^+_{nm}} c_{K^1 K^0} \prod_{j \in K^1} x_j \prod_{j' \in K^0} (1 - x'_j) \ , \end{aligned}$$

and all coefficients $c_{K^1 K^0}$ in the second sum are non-negative.

Therefore, the second sum is non-negative.

Thus,

$$\forall x \in \{0,1\}^n \qquad f(x) \geq c_{\varnothing\varnothing} \ .$$

$\square$

Definition 14. For any posiform $c : J \to \mathbb{R}$, a pair (S, y) such that $S \subseteq \{1, \ldots, n\}$ and $y : S \to \{0, 1\}$ is called a **contractor** of c iff

$$\begin{aligned} \forall (J^1, J^0) \in J : \qquad & (J^1 \cap S = \varnothing \quad \wedge \quad J^0 \cap S = \varnothing) \\ & \vee\ (\exists j \in J^1 \cap S \quad y_j = 0) \\ & \vee\ (\exists j \in J^0 \cap S \quad y_j = 1)\ . \end{aligned} \tag{41}$$

Theorem 2 (partial optimality)**.** For any $n \in \mathbb{N}$, any $f : \{0,1\}^n \to \mathbb{R}$, any posiform $c \in C^+_{nn}(f)$ and any contractor (S, y) of c, there exists a solution x^* to the problem $\min\{f(x) \mid x \in \{0,1\}^n\}$ such that

$$\forall j \in S: \quad x^*_j = y_j \ . \tag{42}$$

Proof. Let $\sigma_{Sy} : \{0,1\}^n \to \{0,1\}^n$ such that $\forall x \in \{0,1\}^n \ \forall j \in \{1,\dots,n\}$:

$$\sigma_{Sy}(x)_j = \begin{cases} y_j & \text{if } j \in S \\ x_j & \text{otherwise} \end{cases} . \tag{43}$$

Let $J^{\bar{S}} := \{(J^1, J^0) \in J^+_{nn} \mid J^1 \cap S = J^0 \cap S = \varnothing\}$ and $J^S := J \setminus J^{\bar{S}}$.
Now, $\forall x \in \{0,1\}^n$:

$$f(x) = \underbrace{\sum_{(J^1,J^0)\in J^S} c_{J^1 J^0} \prod_{j\in J^1} x_j \prod_{j'\in J^0} (1 - x'_j)}_{=: f^S(x)} + \underbrace{\sum_{(J^1,J^0)\in J^{\bar{S}}} c_{J^1 J^0} \prod_{j\in J^1} x_j \prod_{j'\in J^0} (1 - x'_j)}_{=: f^{\bar{S}}(x)} .$$

Furthermore, $\forall x \in \{0,1\}^n$:

$$\begin{aligned} f^S(\sigma_{Sy}(x)) &= 0 && \text{(by definition)} \\ 0 &\le f^S(x) && \text{(because } (\varnothing,\varnothing) \notin J^S\text{)} \\ f^{\bar{S}}(\sigma_{Sy}(x)) &= f^{\bar{S}}(x) && \text{(by definition)} . \end{aligned}$$

Adding the lhs. and rhs. shows that σ_{Sy} is improving for the problem $\min\{f(x) \mid x \in \{0,1\}^n\}$. □

Summary:

- Every PBF has a posiform
- The posiform of a PBF need not be unique
- For every PBF f and every posiform c of f
  - $c_{\varnothing\varnothing}$ is a lower bound on the minimum of f
  - partial optimality holds at any contractor of c

For any $n \in \mathbb{N}$, consider n-variate **quadratic** forms, i.e.

- any **multi-linear polynomial form** $c \in C_{n2}$, and f_c, i.e. for all $x \in \{0,1\}^n$:

$$f_c(x) = c_\varnothing + \sum_{j \in \{1,\dots,n\}} c_{\{j\}} x_j + \sum_{\{j,k\} \in \binom{\{1,\dots,n\}}{2}} c_{\{j,k\}} x_j x_k$$

- any **posiform** $c' \in C^+_{n2}$, and f'_c, i.e. for all $x \in \{0,1\}^n$:

$$\begin{aligned} f'_{c'}(x) = c'_{\varnothing\varnothing} &+ \sum_{j \in \{1,\dots,n\}} \left(c'_{\{j\}\varnothing} x_j + c'_{\varnothing\{j\}} (1 - x_j) \right) \\ &+ \sum_{\{j,k\} \in \binom{\{1,\dots,n\}}{2}} \Big(c'_{\{j,k\}\varnothing} x_j x_k + c'_{\{j\}\{k\}} x_j (1 - x_k) \\ &\qquad + c'_{\{k\}\{j\}} x_k (1 - x_j) + c'_{\varnothing\{j,k\}} (1 - x_j)(1 - x_k) \Big) \end{aligned}$$

Lemma 8. For any $n \in \mathbb{N}$, any QPBF $f : \{0,1\}^n \to \mathbb{R}$, the $c \in C_{n2}$ such that $f_c = f$ and any $c' \in C^+_{n2}(f)$:

$$c_\varnothing = c'_{\varnothing\varnothing} + \sum_{j=1}^{n} c'_{\varnothing\{j\}} + \sum_{\{j,k\} \in \binom{\{1,\dots,n\}}{2}} c'_{\varnothing\{j,k\}}$$

$$\forall j \in \{1,\dots,n\}: \qquad c_{\{j\}} = c'_{\{j\}\varnothing} - c'_{\varnothing\{j\}} + \sum_{k \in \{1,\dots,n\} \setminus \{j\}} \left(c'_{\{j\}\{k\}} - c'_{\varnothing\{j,k\}} \right)$$

$$\forall \{j,k\} \in \tbinom{\{1,\dots,n\}}{2}: \quad c_{\{j,k\}} = c'_{\{j,k\}\varnothing} + c'_{\varnothing\{j,k\}} - c'_{\{j\}\{k\}} - c'_{\{k\}\{j\}}$$

Proof. Expansion of the posiform c' yields a quadratic multi-linear polynomial form. Comparison with c yields the conditions stated in the Lemma. □

Definition 15 (Complementation)**.** For any $n \in \mathbb{N}$ and any QPBF $f : \{0,1\}^n \to \mathbb{R}$, the real number $\max \{c'_{\varnothing\varnothing} \mid c' \in C^+_{n2}(f)\}$ is called the **floor dual** of f.

Corollary 3 (of Lemma 8)**.** For any $n \in \mathbb{N}$ and any QPBF $f : \{0,1\}^n \to \mathbb{R}$, the floor dual is the value of an optimal solution to the linear program

$$
\begin{aligned}
\max_{c' : J^+_{n2} \to \mathbb{R}} \quad & c_\varnothing - \sum_{j=1}^{n} c'_{\varnothing\{j\}} - \sum_{\{j,k\} \in \binom{\{1,\ldots,n\}}{2}} c'_{\varnothing\{j,k\}} \\
\text{subject to} \quad & \forall j \in \{1,\ldots,n\}: \quad c_{\{j\}} = c'_{\{j\}\varnothing} - c'_{\varnothing\{j\}} + \sum_{k \in \{1,\ldots,n\}-\{j\}} \left(c'_{\{j\}\{k\}} - c'_{\varnothing\{j,k\}}\right) \\
& \forall \{j,k\} \in \tbinom{\{1,\ldots,n\}}{2}: \quad c_{\{j,k\}} = c'_{\{j,k\}\varnothing} + c'_{\varnothing\{j,k\}} - c'_{\{j\}\{k\}} - c'_{\{k\}\{j\}} \\
& \forall J \in J^+_{n2} - \{(\varnothing,\varnothing)\}: \quad 0 \le c'_J \ .
\end{aligned}
$$

Summary:

- For any PBF, a quadratic posiform with maximum floor dual bound $c_{\varnothing\varnothing}$ can be found by solving a linear program.