Machine Learning I

Lucas Fabian Naumann, David Stein, Bjoern Andres

Machine Learning for Computer Vision
TU Dresden

https://mlcv.cs.tu-dresden.de/courses/25-winter/ml1/

Winter Term 2025/2026

Contents: In this part of the course, we discuss one approach to pseudo-boolean optimization.

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 1

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

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

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

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

Definition 2

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 3

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.

$$f_c\colon \quad \{0,1\}^n \to \mathbb{R}\colon \quad x \mapsto \sum_{m=0}^{d} \sum_{J \in \binom{\{1,\dots,n\}}{m}} c_J \prod_{j \in J} x_j \tag{2}$$

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

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

Lemma 4

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

$$\forall n \in \mathbb{N} \quad \forall f : \{0,1\}^n \to \mathbb{R} \quad \exists_1 c \in C_{nn} \quad f = f_c \ . \tag{3}$$

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

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

Explicitly,

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

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

Proof.

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

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

Now,

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

is written equivalently as

$$\begin{aligned} f(x^\emptyset) &= c_\emptyset \\ \forall J \neq \emptyset: \quad f(x^J) &= c_J + \sum_{J' \subset J} c_{J'} \ . \end{aligned}$$

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

Definition 5

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

$$F_{nd} := \{f : \{0,1\}^n \to \mathbb{R} \mid \exists c \in C_{nd} : \; f = f_c\} \tag{4}$$

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

Definition 6

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

$$\min \{f(x) \mid x \in \{0,1\}^n\} \tag{5}$$

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

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

$$\min \{f(x) \mid x \in \{0,1\}^n\} \tag{6}$$

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

Is QPBO less complex than PBO?

Definition 7

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

$$\text{size}(c) := \sum_{J \subseteq \{1,\dots,n\}\colon\, c_J \neq 0} |J| \ . \tag{7}$$

Lemma 8

For any $x, y, z \in \{0, 1\}$*:*

$$z = xy \quad \Leftrightarrow \quad xy - 2xz - 2yz + 3z = 0 \ , \tag{8}$$

$$z \neq xy \quad \Leftrightarrow \quad xy - 2xz - 2yz + 3z > 0 \ . \tag{9}$$

Proof.

By verifying equivalence for all eight cases. □

Algorithm 1 (Boros and Hammer 2001)

Input: $c \in C_{nn}$
Output: $c' \in C_{n2}$
$M := 1 + 2\sum_{J \subseteq \{1,\ldots,n\}} |c_J|$
$m := n$
$c^m := c$
while *there exists a* $J \subseteq \{1,\ldots,n\}$ *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 $\{j,k\} \subseteq J' \subseteq \{1,\ldots,n\}$ *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 9

- *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* $\forall \hat{x} \in \mathbb{R}^n$:

$$\hat{x} \in \operatorname*{argmin}_{x \in \{0,1\}^n} f_c(x)$$
$$\Rightarrow \quad \exists \hat{x}' \in \{0,1\}^m \left(\hat{x}'_{\{1,\ldots,n\}} = \hat{x}_{\{1,\ldots,n\}} \wedge \hat{x}' \in \operatorname*{argmin}_{x' \in \{0,1\}^m} f_{c'}(x') \right) \quad . \tag{10}$$

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 8 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. □

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

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

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

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

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

Example. For $n = d = 2$,

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

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

$$\forall x \in \{0,1\}^n \qquad f_c(x) := \sum_{(J^1, J^0) \in J^+_{nd}} c_{J^1 J^0} \prod_{j \in J^1} x_j \prod_{j' \in J^0} (1 - x'_j) \qquad (11)$$

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 $\forall x \in \{0,1\}^2$:

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

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

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

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

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

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

Lemma 13

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 14

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 $x \in \{0,1\}^n$,

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

are the only elements of J for which

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

Thus,

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

□

Remark. 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 15

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

$$C^+_{nd}(f) := \left\{ c \in C^+_{nd} \mid f_c = f \right\} \quad . \tag{12}$$

Remark. 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 16

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

Proof.

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

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

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

Therefore, the second sum is non-negative.

Thus,

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

□

Definition 17

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

$$\begin{aligned}\forall (J^1, J^0) \in J \colon \qquad & (J^1 \cap S = \emptyset \quad \wedge \quad J^0 \cap S = \emptyset) \\ & \vee\ (\exists j \in J^1 \cap S \quad y_j = 0) \\ & \vee\ (\exists j \in J^0 \cap S \quad y_j = 1)\ . \end{aligned} \tag{13}$$

Theorem 18 (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

$$\forall j \in S: \quad x^*_j = y_j \quad . \tag{14}$$

Proof.

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

$$\sigma_{Sy}(x)_j = \begin{cases} y_j & \text{if } j \in S \\ x_j & \text{otherwise} \end{cases} . \tag{15}$$

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

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

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

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

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

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

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

Lemma 19

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* $c' \in C^+_{n2}(f)$*:*

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

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

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

Proof.

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

Definition 20 (Complementation)

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

Corollary 21 (of Lemma 19)

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

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

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
  - $c_{\emptyset\emptyset}$ is a lower bound on the minimum of f
  - partial optimality holds at any contractor of c
- For any PBF, a quadratic posiform with maximum floor dual bound $c_{\emptyset\emptyset}$ can be found by solving a linear program.