# Combinatorial Optimization in Computer Vision

## Chapter 10: Introduction to Pseudo-Boolean Optimization

WS 2011/12

Ulrich Schlickewei
Computer Vision and Pattern Recognition Group
Technische Universität München

# Plan for Today

Pseudo-Boolean Functions

- Definition
- Different Representations

Pseudo-Boolean Optimization

- Lower Bound
- Equivalent Continuous Optimization Problem
- Reduction to Quadratic Pseudo-Boolean Optimization

# Pseudo-Boolean Functions

**Definition**: Let $V = \{1, \ldots, n\}$ be a finite set.

a) A set function on $V$ is a real-valued function

$$f : \mathcal{P}(V) \to \mathbb{R}$$

defined on the power set of $V$.

b) A pseudo-Boolean function on $V$ is a real-valued function

$$f : \{0, 1\}^n \to \mathbb{R}.$$

# Set Function = Pseudo-Boolean Function

Since subsets $S \subset V$ can be represented by their indicator vectors $\mathbf{1}_S \in \{0,1\}^n$ with $\big((\mathbf{1}_S)_i = 1\big) \iff \big(i \in S\big)$ , there is a natural bijection between pseudo-Boolean functions and set functions. We will switch between the two notions with no further comment.

There are (at least) three ways of representing pseudo-Boolean functions:

- table
- multilinear polynomial
- posiform.

# Representation by Table

Let $V = \{1, 2, 3\}$. The most unefficient way of representing a pseudo-Boolean function on $V$ is by specifying its value on all possible subsets:

| $S$ | $f(S)$ |
|---|---:|
| $\emptyset$ | 5 |
| $\{1\}$ | -1 |
| $\{2\}$ | 2 |
| $\{3\}$ | -3 |
| $\{1, 2\}$ | 1 |
| $\{1, 3\}$ | 15 |
| $\{2, 3\}$ | -7 |
| $\{1, 2, 3\}$ | 2 |

# Representation by Multilinear Polynomial

**Prop**.: Every pseudo-Boolean function

$$f : \{0,1\}^n \to \mathbb{R}.$$

has a unique representation as a multilinear polynomial

$$f(x_1, \ldots, x_n) = \sum_{S \subset V} c_S \prod_{j \in S} x_j .$$

(A multivariate polynomial is multilinear if in each monomial $c_{i_1 \ldots i_d} x_{i_1} \cdot \ldots \cdot x_{i_d}$ each variable $x_i$ appears at most in degree one. This is equivalent to say that if all variables but $x_i$ are fixed, then the polynomial is linear in $x_i$ .)

# Degree and Size

Let $f : \{0,1\}^n \to \mathbb{R}$ be a pseudo-Boolean function represented by a multilinear polynomial

$$f(x_1, \ldots, x_n) = \sum_{S \subset V} c_S \prod_{j \in S} x_j .$$

- The degree of $f$ is defined as the degree of its multilinear polynomial representation.
- The size of $f$ is defined as

$$\mathrm{size}(f) = \sum_{S \in V : c_S \neq 0} |S| .$$

# Representation by a Posiform I

Let $L = \{1, \dots, n, \overline{1}, \dots, \overline{n}\}$. (Sometimes, this is called the set of literals.)

For a vector $x = (x_1, \dots, x_n) \in \{0,1\}^{|V|}$, we define the corresponding literal $u(x) \in \{0,1\}^{|L|}$ by

$$(u(x))_l = \begin{cases} x_l & \text{if } l \in \{1, \dots, n\} \\ \overline{x_l} = 1 - x_l & \text{if } l \in \{\overline{1}, \dots, \overline{n}\} \end{cases} .$$

# Representation by a Posiform II

Then any pseudo-Boolean function $f : \{0,1\}^n \to \mathbb{R}$ can be represented by a posiform

$$\phi(x) = \sum_{T \subset L} a_T \prod_{l \in T} u(x)_l,$$

where $a_T \geq 0$ whenever $T \neq \emptyset$.

In contrast to the multilinear polynomial representation, the posiform representation is not unique. For example

$$\begin{aligned} x_1 \overline{x_2} &= (1 - \overline{x_1})(1 - x_2) \\ &= 1 - \overline{x_1} - x_2 + \overline{x_1} x_2 \\ &= 1 - (1 - x_1) - (1 - \overline{x_2}) + \overline{x_1} x_2 \\ &= -1 + x_1 + \overline{x_2} + \overline{x_1} x_2. \end{aligned}$$

# Example

Binary image segmentation:

- $V$ = set of pixels
- $x_i$ = labelling of pixel $i$ .

Define

$$\phi(x) = \sum_{i \in V} c_i x_i + \overline{c_i x_i} + \sum_{i \sim j} c_{ij} (x_i \overline{x_j} + \overline{x_i} x_j).$$

Then binary image segmentation can be formulated as minimizing the quadratic posiform $\phi$ .

We will see later, that quadratic pseudo-Boolean optimization is intimately related to graph cuts.

# From Tables to Posiforms I

Consider the pseudo-Boolean function defined by the table.

| $S$ | $f(S)$ | term |
|---|---|---|
| $\emptyset$ | 5 | $\overline{x_1 x_2 x_3}$ |
| $\{1\}$ | -1 | $x_1 \overline{x_2 x_3}$ |
| $\{2\}$ | 2 | $\overline{x_1} x_2 \overline{x_3}$ |
| $\{3\}$ | -3 | $\overline{x_1 x_2} x_3$ |
| $\{1, 2\}$ | 1 | $x_1 x_2 \overline{x_3}$ |
| $\{1, 3\}$ | 15 | $x_1 \overline{x_2} x_3$ |
| $\{2, 3\}$ | -7 | $\overline{x_1} x_2 x_3$ |
| $\{1, 2, 3\}$ | 2 | $x_1 x_2 x_3$ |

This is represented by

$$5\overline{x_1 x_2 x_3} - x_1 \overline{x_2 x_3} + 2\overline{x_1} x_2 \overline{x_3} - 3\overline{x_1 x_2} x_3$$

$$+ x_1 x_2 \overline{x_3} + 15 x_1 \overline{x_2} x_3 - 7\overline{x_1} x_2 x_3 + 2 x_1 x_2 x_3.$$

# From Tables to Posiforms II

The expression

$$5\overline{x_1}\,\overline{x_2}\,\overline{x_3} - x_1\overline{x_2}\,\overline{x_3} + 2\overline{x_1}x_2\overline{x_3} - 3\overline{x_1}\,\overline{x_2}x_3$$
$$+ x_1x_2\overline{x_3} + 15x_1\overline{x_2}x_3 - 7\overline{x_1}x_2x_3 + 2x_1x_2x_3.$$

is not a posiform because of negative coefficients in degree terms. However, it can be transformed to a posiform using expressions like

$$\begin{aligned}
-x_1\overline{x_2}\,\overline{x_3} &= -(1-\overline{x_1})\overline{x_2}\,\overline{x_3} \\
&= \overline{x_1}\,\overline{x_2}\,\overline{x_3} - \overline{x_2}\,\overline{x_3} \\
&= \overline{x_1}\,\overline{x_2}\,\overline{x_3} - (1-x_2)\overline{x_3} \\
&= \overline{x_1}\,\overline{x_2}\,\overline{x_3} + x_2\overline{x_3} - \overline{x_3} \\
&= \overline{x_1}\,\overline{x_2}\,\overline{x_3} + x_2\overline{x_3} + x_3 - 1.
\end{aligned}$$

# From Posiforms to Multilinear Polynomials

In order to pass from a posiform representation to a multilinear polynomial, replace each occurence of $\overline{x_i}$ by $1 - x_i$ and multiply out.

# Plan for Today

Pseudo-Boolean Functions

- Definition
- Different Representations

Pseudo-Boolean Optimization

- Lower Bound
- Equivalent Continuous Optimization Problem
- Reduction to Quadratic Pseudo-Boolean Optimization

# Pseudo-Boolean Optimization

Consider a pseudo-Boolean function $f : \{0,1\}^n \to \mathbb{R}$ and the optimization problem

$$\min_{x \in \{0,1\}^n} f(x).$$

It is easy to see that this problem is NP-hard in general. For example, the Boolean satisfiability problem SAT can be formulated as a posiform minimization problem.

An even easier way to the the NP-hardness is by noting, that in general, each evaluation of $f$ can have exponential cost in $n$.

# General Results

We now start our analysis of pseudo-Boolean optimization with three general results:

- We derive a lower bound on minimization problems induced by the posiform representation.
- We formulate an equivalent continuous optimization problem.
- We note that every pseudo-Boolean optimization problem can be reduced to a quadratic one.

# Lower Bound

Assume that the pseudo-Boolean function $f$ is represented by a posiform

$$\phi(x) = \sum_{T \subset L} a_T \prod_{l \in T} u(x)_l,$$

with $a_T \geq 0$ for $T \neq \emptyset$ . Let $C(\phi) = a_\emptyset$ be the constant coeff.

**Proposition**:

a) We have $a_\emptyset \leq \min_{x \in \{0,1\}^n} f(x).$

b) There exists a representation of $f$ by a posiform $\phi^*$ with

$$C(\phi^*) = \min_{x \in \{0,1\}^n} f(x).$$

# Equivalent Continuous Optimization Problem

**Prop.**: Assume that $f$ is represented by a multilinear polynomial

$$f(x_1, \dots, x_n) = \sum_{S \subset V} c_S \prod_{j \in S} x_j.$$

Let $r \in [0,1]^n$. Then there exist $x, y \in \{0,1\}^n$ such that

$$f(x) \leq f(r) \leq f(y).$$

Furthermore, such vectors can be generated in $\mathcal{O}(\mathrm{size}(f))$.

# Derivatives of Pseudo-Boolean Functions

For the proof of the Proposition we need some further notation.

**Def.**: The i-th derivative of $f$ is defined as

$$\frac{\partial f}{\partial x_i}(x) := f(x_1, \dots, \underbrace{1}_{i\text{-th position}}, \dots, x_n) - f(x_1, \dots, \underbrace{0}_{i\text{-th position}}, \dots, x_n).$$

# Pseudo-Boolean Optimization $\subset$ Continuous Optimization

An immediate corollary of the previous proposition is

**Corollary**: For any pseudo-Boolean function $f$

$$\min_{x \in \{0,1\}^n} f(x) = \min_{r \in [0,1]^n} f(x).$$

# Reduction to Quadratic Pseudo-Boolean Optimization

**Theorem** (Rosenberg): For every pseudo-Boolean function $f : \{0,1\}^n \to \mathbb{R}$ there exists a quadratic (i.e. degree 2) pseudo-Boolean function $g : \{0,1\}^m \to \mathbb{R}$ such that

$$\min_{x \in \{0,1\}^n} f(x) = \min_{y \in \{0,1\}^m} g(y).$$

Furthermore, $g$ can be constructed in polynomial time and the optimal $x$ can be read off the optimal $y$ .

# Summary

- Pseudo-Boolean functions are functions defined on subsets of finite sets.
- They admit (at least) three different representations:
  - table
  - multilinear polynomial
  - posiform
- Pseudo-Boolean optimization is NP-hard in general.
- A lower bound can be determined using the posiform representation.
- There exists an equivalent continuous optimization problem.
- Pseudo-Boolean optimization can be reduced to quadratic pseudo-Boolean optimization.