Quadratization of symmetric pseudo-Boolean functions

Yves Crama
with Martin Anthony, Endre Boros and Aritanan Gruber

HEC Management School, University of Liège

Liblice, April 2013

Outline

1. Pseudo-Boolean functions
2. Quadratization
3. Symmetric functions
4. About negative monomials

Objectives

Focus:

- basic facts about pseudo-Boolean minimization
- quadratization techniques
- the case of symmetric functions
- more about negative monomials.

Definitions

Pseudo-Boolean functions

A pseudo-Boolean function is a mapping $f : \{0, 1\}^n \rightarrow \mathbb{R}$

Multilinear polynomials

Every pseudo-Boolean function can be represented – in a unique way – by a *multilinear polynomial* in its variables.

Example:

$f = 4 - 9x_1 - 5x_2 - 2x_3 + 13x_1x_2 + 13x_1x_3 + 6x_2x_3 - 13x_1x_2x_3$

Pseudo-Boolean optimization

Complexity:

PB optimization

Given a multilinear polynomial f of degree at least 2, it is NP-hard to find the minimum of f.

- Many applications: MAX CUT, MAX SAT, computer vision, etc.
- When f is quadratic and has no positive quadratic terms, then f is submodular and its minimization reduces to minimum cost flow.

Quadratic optimization

The quadratic case has attracted most of the attention:

- many examples arise in this form: MAX CUT, MAX 2SAT, simple computer vision models,...
- higher-degree cases can be efficiently reduced to the quadratic case, and this leads to good optimization algorithms.

Observations

- Say $g(x, y)$, $(x, y) \in \{0, 1\}^{n+m}$, is a quadratic function.
- Then, for all $x \in \{0, 1\}^n$,

$$f(x) = \min\{g(x, y) \mid y \in \{0, 1\}^m\}$$

  is a pseudo-Boolean function.
- $f(x)$ may be quadratic, or not.
- $\min\{f(x) \mid x \in \{0, 1\}^n\} = \min\{g(x, y) \mid (x, y) \in \{0, 1\}^{n+m}\}$.
- Conversely...

Quadratization

Quadratization

The quadratic function $g(x, y)$, $(x, y) \in \{0, 1\}^{n+m}$ is an *m-quadratization* of the pseudo-Boolean function $f(x)$, $x \in \{0, 1\}^n$, if

$$f(x) = \min\{g(x, y) \mid y \in \{0, 1\}^m\} \quad \text{for all } x \in \{0, 1\}^n.$$

- $\min\{f(x) \mid x \in \{0, 1\}^n\} = \min\{g(x, y) \mid (x, y) \in \{0, 1\}^{n+m}\}$.
- Does every function f have a quadratization?

Existence

Existence of quadratizations

Given the multilinear expression of a pseudo-Boolean function $f(x), x \in \{0,1\}^n$, one can find in polynomial time a quadratization $g(x,y)$ of $f(x)$.

- Due to Rosenberg (1975).
- Idea: replace the term $\prod_{i \in A} x_i$ of f, with $\{1,2\} \subseteq A$, by
  $t(x,y) = \left(\prod_{i \in A \setminus \{1,2\}} x_i \right) y + M(x_1x_2 - 2x_1y - 2x_2y + 3y)$.
- Fix x. In every minimizer of $t(x,y)$, $y = x_1x_2$ and $t(x,y) = \prod_{i \in A} x_i$.
- Drawbacks: introduces many additional variables, many positive quadratic terms, big M.

Questions arising...

- Many quadratization procedures proposed in recent years. Which ones are “best”? Small number of variables, of positive terms, good properties with respect to persistencies, submodularity?
- Can we chararacterize all quadratizations of f?
- Easier question: What if f is a single monomial?
- How many variables are needed in a quadratization?
- etc.

Refs: Boros and Gruber (2011); Fix, Gruber, Boros and Zabih (2011): Freedman and Drineas (2005); Ishikawa (2011); Kolmogorov and Zabih (2004); Ramalingam et al. (2011); Rosenberg (1975); Rother et al. (2009); Živný, Cohen and Jeavons (2009); etc.

The case of symmetric functions

Symmetric functions

A pseudo-Boolean function f is *symmetric* if the value of $f(x)$ depends only on the Hamming weight $\mathrm{wt}(x) = \sum_{j=1}^{n} x_j$ (number of ones) of x.

That is, there is a function $k : \{0, 1, \ldots, n\} \to \mathbb{R}$ such that $f(x) = k(w)$ where $w = \mathrm{wt}(x)$.

Examples

- Negative monomial: $N_n(x) = -\prod_{i=1}^n x_i = -x_1 \ldots x_n$.
- (Freedman and Drineas 2005) $N_n(x) = \min_y \, [n-1-\sum_{i=1}^n x_i]y$.
- Positive monomial: $P_n(x) = \prod_{i=1}^n x_i = x_1 \ldots x_n$.
- $P_n(x) = -x_1 \ldots x_{n-1}\overline{x}_n + P_{n-1}(x)$: so P_n can be quadratized using $n-2$ additional variables.
- (Ishikawa 2011) P_n can be quadratized using $\left\lfloor \frac{n-1}{2} \right\rfloor$ additional variables.
- How many variables are needed for other symmetric functions?
- (Fix 2011) $n-1$ variables suffice.

We propose a generic approach.

A representation theorem

Let $[a]^- = \min(a, 0)$.

Theorem: Representation of symmetric functions

For all $0 < \epsilon_i \leq 1$, $i = 0, \ldots n$, every symmetric pseudo-Boolean function $f : \{0, 1\}^n \to \mathbb{R}$ can be uniquely represented in the form

$$f(x) = \sum_{i=0}^{n} \alpha_i \left[i - \epsilon_i - \sum_{j=1}^{n} x_j \right]^- .$$

- Idea: $\left[i - \epsilon_i - \sum_{j=1}^n x_j \right]^-$ reflects whether $\sum_{j=1}^n x_j$ is larger than i.
- System of linear equations: $\alpha_0, \ldots, \alpha_n$ can be efficiently computed.

Example: Negative monomials

- Let $N_n(x) = -\prod_{i=1}^{n} x_i$.
- Then: $N_n = \left[n - 1 - \sum_{i=1}^{n} x_i\right]^-$.
- Note: $[a]^- = \min(a, 0) = \min_{y \in \{0,1\}} ay$.
- So: $N_n = \min_y \left[n - 1 - \sum_{i=1}^{n} x_i\right] y$. (Freedman and Drineas 2005).

Quadratization: first attempt

- More generally, using $[a]^- = \min(a, 0) = \min_{y \in \{0,1\}} ay$
- for symmetric f:

$$f(x) = \sum_{i=0}^{n} \alpha_i \left[i - \epsilon_i - \sum_{j=1}^{n} x_j \right]^- = \min_y \sum_{i=0}^{n} \alpha_i \left(i - \epsilon_i - \sum_{j=1}^{n} x_j \right) y_i.$$

- Well, not quite: $-[a]^- = -\min(a, 0) \neq \min_{y \in \{0,1\}} (-ay)$.
- Modify the representation of f to cancel negative coefficients.

Some identities

Define

- $E(l) = \frac{l(l-1)}{2} + 2l + 1 + \sum_{i=-1}^{n-1} [i-l]^-$,
- $E'(l) = \frac{l(l-1)}{2} + 2\sum_{\substack{i=2: \\ i\ even}}^{n} \left[i - \frac{1}{2} - l\right]^-$,
- $E''(l) = \frac{l(l+1)}{2} + 2\sum_{\substack{i=1: \\ i\ odd}}^{n} \left[i - \frac{1}{2} - l\right]^-$.

Lemma.

For all $l = 0, \ldots, n$, $E(l) = E'(l) = E''(l) = 0$.

Proof. Follows from the representation theorem applied to $e(x) = \sum_{i<j} x_i x_j = w(w-1)/2$, where w is the Hamming weight of x.

Quadratization: second attempt

- for symmetric f:

$$f(x) = \sum_{i=0}^{n} \alpha_i \left[i - \epsilon_i - \sum_{j=1}^{n} x_j \right]^{-} = \min_{y} \sum_{i=0}^{n} \alpha_i \left(i - \epsilon_i - \sum_{j=1}^{n} x_j \right) y_i .$$

- Well, not quite: $-[a]^{-} = -\min(a,0) \neq \min_{y\in\{0,1\}}(-ay)$.
- Modify first the representation of f by adding one of $E(l)$, $E'(l)$, $E'(l)$ so as to cancel negative coefficients.

Example: Positive monomials

- Let $P_n(x) = \prod_{i=1}^n x_i$.
- Then: $P_n = -2\left[n - \frac{1}{2} - \sum_{i=1}^n x_i\right]^-$.
- For even n: add $E'(\sum_{i=1}^n x_i)$ to P_n, leading to

$$\begin{aligned} P_n &= \sum_{1\le i<j\le n} x_i x_j + \sum_{\substack{i=2:\\ i \text{ even}}}^{n-2} 2\left[i - \frac{1}{2} - \sum_{j=1}^n x_j\right]^- \\ &= \sum_{1\le i<j\le n} x_i x_j + \min_y \sum_{\substack{i=2:\\ i \text{ even}}}^{n-2} 2y_i(i - \frac{1}{2} - \sum_{j=1}^n x_j). \end{aligned}$$

Quadratization of symmetric functions

Theorem: Positive monomials (Ishikawa 2011)

The positive monomial P_n has a $\left\lfloor \frac{n-1}{2} \right\rfloor$-quadratization.

Theorem: k-out-of-n function

The k-out-of-n function has a $\left\lceil \frac{n}{2} \right\rceil$-quadratization.

Theorem: Parity function

The parity function has a $\left\lfloor \frac{n}{2} \right\rfloor$-quadratization, and its complement has a $\left\lfloor \frac{n-1}{2} \right\rfloor$-quadratization.

Theorem: General symmetric functions (Fix 2011)

Every symmetric pseudo-Boolean function has an $(n-1)$-quadratization.

About negative monomials

Remember

- Negative monomial: $N_n(x) = -\prod_{i=1}^n x_i = -x_1 \dots x_n$.
- (Freedman and Drineas 2005) $N_n(x) = \min_y \, [n - 1 - \sum_{i=1}^n x_i]y$.
- Question: can we characterize the quadratizations of N_n?

- Seems surprisingly difficult.
- Note: many quadratizations are somehow "reducible", e.g., by fixing variables, or are identical, up to replacing y by $\bar{y}$.

About negative monomials

We can establish:

Theorem: 1-Quadratizations of negative monomials

Up to a permutation of the x-variables, and up to a switch of the y-variable, the only irreducible 1-quadratizations of N_n are

$$s_n = [n - 1 - \sum_{i=1}^{n} x_i]y,$$

$$s_n^+ = (n-2)x_n y - \sum_{i=1}^{n-1} x_i(y - \bar{x}_n).$$

Proof. Very long...

Conclusions

- Structure and properties of quadratizations are poorly understood.
- Many intriguing questions and conjectures, much computational and theoretical work to be done.

Some references

E. Boros and A. Gruber, On quadratization of pseudo-Boolean functions, Working paper, 2011.

A. Fix, A. Gruber, E. Boros and R. Zabih, A graph cut algorithm for higher-order Markov random fields, Proceedings of the 2011 IEEE International Conference on Computer Vision (ICCV), pages 1020-1027.

D. Freedman and P. Drineas, Energy minimization via graph cuts: Settling what is possible, in: *IEEE Conference on Computer Vision and Pattern Recognition* (2) (2005) pp. 939-946.

H. Ishikawa, Transformation of general binary MRF minimization to the first-order case, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33(6) (2011) 1234-1249.

V. Kolmogorov and C. Rother, Minimizing non-submodular functions with graph cuts - A review, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29 (2007) 1274-1279.

Some references

V. Kolmogorov and R. Zabih, What energy functions can be minimized via graph cuts? *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26(2) (2004) 147-159.

S. Ramalingam, Ch. Russell, L. Ladický and Ph.H.S. Torr, Efficient Minimization of Higher Order Submodular Functions using Monotonic Boolean Functions, arXiv:1109.2304v1, 2011.

I.G. Rosenberg, Reduction of bivalent maximization to the quadratic case, *Cahiers du Centre d'Etudes de Recherche Opérationnelle* 17 (1975), 71–74.

C. Rother, P. Kohli, W. Feng and J. Jia, Minimizing sparse higher order energy functions of discrete variables, in: *IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 1382-1389.

C. Rother, V. Kolmogorov, V. Lempitsky and M. Szummer, Optimizing binary MRFs via extended roof duality, in: *IEEE Conference on Computer Vision and Pattern Recognition*, 2007.

S. Živný, D.A. Cohen and P.G. Jeavons, The expressive power of binary submodular functions, *Discrete Applied Mathematics* 157 (2009) 3347-3358.