Reformulations of nonlinear binary optimization problems

Yves Crama
HEC Management School, University of Liège, Belgium

Koper, May 2018

Outline

Definitions

Pseudo-Boolean functions

A pseudo-Boolean function is a mapping $f : \{0, 1\}^n \to \mathbb{R}$, that is, a real-valued function of $0 - 1$ variables.

Definitions

Pseudo-Boolean functions

A pseudo-Boolean function is a mapping $f : \{0, 1\}^n \to \mathbb{R}$, that is, a real-valued function of $0 - 1$ variables.

Multilinear polynomials

Every pseudo-Boolean function can be represented – in a unique way – as a *multilinear polynomial* in its variables. (Note: $x_k^2 = x_k$.)

Definitions

Pseudo-Boolean functions

A pseudo-Boolean function is a mapping $f : \{0, 1\}^n \to \mathbb{R}$, that is, a real-valued function of $0 - 1$ variables.

Multilinear polynomials

Every pseudo-Boolean function can be represented – in a unique way – as a *multilinear polynomial* in its variables. (Note: $x_k^2 = x_k$.)

Example:

$f = 4 - 9x_1 - 5x_2 - 2x_3 + 13x_1x_2 + 13x_1x_3 + 6x_2x_3x_4 - 13x_1x_2x_3x_4$

Multilinear optimization in binary variables

$$\text{(MOB)} \quad \min_{x \in \{0,1\}^n} f(x) = \sum_{S \in 2^{[n]}} a_S \prod_{k \in S} x_k$$

Multilinear optimization in binary variables

$$\text{(MOB)} \quad \min_{x \in \{0,1\}^n} f(x) = \sum_{S \in 2^{[n]}} a_S \prod_{k \in S} x_k$$

Complexity

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

Multilinear optimization in binary variables

$$\text{(MOB)} \quad \min_{x \in \{0,1\}^n} f(x) = \sum_{S \in 2^{[n]}} a_S \prod_{k \in S} x_k$$

Complexity

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

The quadratic case has attracted much attention:

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

Outline

Quadratization

Observations

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

Quadratization

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.

Quadratization

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.

Quadratization

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

Quadratization

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

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

The y-variables are called *auxiliary* variables.

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

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

The y-variables are called *auxiliary* variables.

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

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

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

The y-variables are called *auxiliary* variables.

- $\min\{f(x) \mid x \in \{0, 1\}^n\} = \min\{g(x, y) \mid (x, y) \in \{0, 1\}^{n+m}\}$.
- Akin to linearization procedures for MOB.

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

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

The y-variables are called *auxiliary* variables.

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

Existence

Existence of quadratizations (Rosenberg 1975)

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 $f(x)$.

- Idea: in each term $\prod_{i \in A} x_i$ of f, with $\{1,2\} \subseteq A$, replace the product $x_1 x_2$ by y:

$$t(x,y) = \left(\prod_{i \in A \setminus \{1,2\}} x_i\right) y \; + \; M(x_1 x_2 - 2x_1 y - 2x_2 y + 3y).$$

Example

Example:

$\min_x f = 4 - 9x_1 - 5x_2 - 2x_3 + 13x_1x_2 + 13x_1x_3 + 6x_2x_3x_4 - 13x_1x_2x_3x_4$

Substitute x_3x_4:

$\min_{x,y} \quad 4 - 9x_1 - 5x_2 - 2x_3 + 13x_1x_2 + 13x_1x_3 + 6x_2y_{34} - 13x_1x_2y_{34} +$
$M\,(x_3x_4 - 2x_3y_{34} - 2x_4y_{34} + 3y_{34})$

Example

Example:

$\min_x f = 4 - 9x_1 - 5x_2 - 2x_3 + 13x_1x_2 + 13x_1x_3 + 6x_2x_3x_4 - 13x_1x_2x_3x_4$

Substitute x_3x_4:

$\min_{x,y} \quad 4 - 9x_1 - 5x_2 - 2x_3 + 13x_1x_2 + 13x_1x_3 + 6x_2y_{34} - 13x_1x_2y_{34} + M(x_3x_4 - 2x_3y_{34} - 2x_4y_{34} + 3y_{34})$

In a minimizer (x, y),

- if $x_3 = x_4 = 1$, then $M(1 - y_{34}) = 0$ and $y_{34} = x_3x_4 = 1$;
- if $x_3 = 0$, then $M(-2x_4y_{34} + 3y_{34}) = 0$ and $y_{34} = x_3x_4 = 0$.

Example

Example:

$\min_x f = 4 - 9x_1 - 5x_2 - 2x_3 + 13x_1x_2 + 13x_1x_3 + 6x_2x_3x_4 - 13x_1x_2x_3x_4$

Substitute x_3x_4:

$\min_{x,y} \quad 4 - 9x_1 - 5x_2 - 2x_3 + 13x_1x_2 + 13x_1x_3 + 6x_2y_{34} - 13x_1x_2y_{34} + M\left(x_3x_4 - 2x_3y_{34} - 2x_4y_{34} + 3y_{34}\right)$

Example

Example:

$\min_x f = 4 - 9x_1 - 5x_2 - 2x_3 + 13x_1x_2 + 13x_1x_3 + 6x_2x_3x_4 - 13x_1x_2x_3x_4$

Substitute x_3x_4:

$\min_{x,y} \quad 4 - 9x_1 - 5x_2 - 2x_3 + 13x_1x_2 + 13x_1x_3 + 6x_2y_{34} - 13x_1x_2y_{34} + M(x_3x_4 - 2x_3y_{34} - 2x_4y_{34} + 3y_{34})$

Substitute x_1x_2:

$\min \quad 4 - 9x_1 - 5x_2 - 2x_3 + 13y_{12} + 13x_1x_3 + 6x_2y_{34} - 13y_{12}y_{34} + M(x_3x_4 - 2x_3y_{34} - 2x_4y_{34} + 3y_{34}) + M(x_1x_2 - 2x_1y_{12} - 2x_2y_{12} + 3y_{12})$

Existence

Existence of quadratizations (Rosenberg 1975)

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 $f(x)$.

- Idea: in each term $\prod_{i \in A} x_i$ of f, with $\{1,2\} \subseteq A$, replace the product $x_1 x_2$ by y:
  $$t(x,y) = \left(\prod_{i \in A \setminus \{1,2\}} x_i \right) y \; + \; M(x_1 x_2 - 2x_1 y - 2x_2 y + 3y).$$
- Fix x. In every minimizer of $t(x,y)$, $y = x_1 x_2$ and $t(x,y) = \prod_{i \in A} x_i$.
- Potential drawbacks: introduces many auxiliary variables, 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?
- Easier question: What if f is a single monomial?
- Can we chararacterize all quadratizations of f?
- How many variables are needed in a quadratization?
- etc.

Refs: Boros and Gruber (2011); Buchheim and Rinaldi (2007); 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.

Outline

Focus of the presentation:

Outline

Focus of the presentation:

- lower and upper bounds on size of quadratizations

Outline

Focus of the presentation:

- lower and upper bounds on size of quadratizations
- the case of symmetric functions

Outline

Focus of the presentation:

- lower and upper bounds on size of quadratizations
- the case of symmetric functions

M. Anthony, E. Boros, Y. Crama and M. Gruber, Quadratization of symmetric pseudo-Boolean functions, *Discrete Applied Mathematics* 203 (2016) 1–12.

M. Anthony, E. Boros, Y. Crama and M. Gruber, Quadratic reformulations of nonlinear binary optimization problems *Mathematical Programming* 162 (2017) 115-144.

E. Boros, Y. Crama and E. Rodrìguez-Heck, Compact quadratizations for pseudo-Boolean functions, Working paper, 2018.

Outline

General question

- How many auxiliary variables are needed in general?

General question

- How many auxiliary variables are needed in general?
- Worst-case bound based on termwise quadratizations:

General question

- How many auxiliary variables are needed in general?
- Worst-case bound based on termwise quadratizations:

Observation

Every term of the form $a\prod_{i=1}^{n} x_i$ can be quadratized using $n-2$ auxiliary variables (Rosenberg 1975), and even $\left\lfloor \frac{n-1}{2} \right\rfloor$ auxiliary variables (Ishikawa 2011).

General question

- How many auxiliary variables are needed in general?
- Worst-case bound based on termwise quadratizations:

Observation

Every term of the form $a\prod_{i=1}^{n} x_i$ can be quadratized using $n-2$ auxiliary variables (Rosenberg 1975), and even $\left\lfloor \frac{n-1}{2} \right\rfloor$ auxiliary variables (Ishikawa 2011).

So:

Ishikawa (2011)

Every n-variable pBF has a quadratization involving $\left\lfloor \frac{n-1}{2} \right\rfloor 2^n$ auxiliary variables.

General question

- How many auxiliary variables are needed in general?
- Worst-case bound based on termwise quadratizations:

Observation

Every term of the form $a\prod_{i=1}^{n} x_i$ can be quadratized using $n-2$ auxiliary variables (Rosenberg 1975), and even $\lfloor \frac{n-1}{2} \rfloor$ auxiliary variables (Ishikawa 2011).

So:

Ishikawa (2011)

Every n-variable pBF has a quadratization involving $\lfloor \frac{n-1}{2} \rfloor 2^n$ auxiliary variables.

- Best known bound, until recently.
- Digression...

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 $|x| = \sum_{j=1}^{n} x_j$ (number of ones) of x.

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

Examples:

- Monomials: $a \prod_{i=1}^{n} x_i = a\, x_1 \ldots x_n$.
- At least k-out-of-n function: takes value 1 if and only if $|x| \geq k$.
- Parity function: takes value 1 if and only if $|x|$ is even.

Upper bounds for symmetric functions

- How many auxiliary variables are needed to quadratize a symmetric function?

Upper bounds for symmetric functions

- How many auxiliary variables are needed to quadratize a symmetric function?

- Negative monomial: $N_n(x) = -\prod_{i=1}^{n} x_i = -x_1 \dots x_n$.
- (Freedman and Drineas 2005) $N_n(x) = \min_y \, (n - 1 - \sum_{i=1}^{n} x_i)y$.

Upper bounds for symmetric functions

- How many auxiliary variables are needed to quadratize a symmetric function?

- Negative monomial: $N_n(x) = -\prod_{i=1}^{n} x_i = -x_1 \dots x_n$.
- (Freedman and Drineas 2005) $N_n(x) = \min_y \left(n - 1 - \sum_{i=1}^{n} x_i\right) y$.
- Positive monomial: $P_n(x) = \prod_{i=1}^{n} x_i = x_1 \dots x_n$.

Upper bounds for symmetric functions

- How many auxiliary variables are needed to quadratize a symmetric function?

- Negative monomial: $N_n(x) = -\prod_{i=1}^n x_i = -x_1 \dots x_n$.
- (Freedman and Drineas 2005) $N_n(x) = \min_y \, (n - 1 - \sum_{i=1}^n x_i)y$.
- Positive monomial: $P_n(x) = \prod_{i=1}^n x_i = x_1 \dots x_n$.
- (Ishikawa 2011) P_n can be quadratized using $\lfloor \frac{n-1}{2} \rfloor$ auxiliary variables.

Upper bounds for symmetric functions

- How many auxiliary variables are needed to quadratize a symmetric function?

- Negative monomial: $N_n(x) = -\prod_{i=1}^{n} x_i = -x_1 \dots x_n$.
- (Freedman and Drineas 2005) $N_n(x) = \min_y \, (n - 1 - \sum_{i=1}^{n} x_i)y$.
- Positive monomial: $P_n(x) = \prod_{i=1}^{n} x_i = x_1 \dots x_n$.
- (Ishikawa 2011) P_n can be quadratized using $\left\lfloor \frac{n-1}{2} \right\rfloor$ auxiliary variables.
- (Fix 2011) $n - 1$ variables suffice for any symmetric function.

Upper bounds for symmetric functions

- How many auxiliary variables are needed to quadratize a symmetric function?

- Negative monomial: $N_n(x) = -\prod_{i=1}^n x_i = -x_1 \dots x_n$.
- (Freedman and Drineas 2005) $N_n(x) = \min_y \, (n - 1 - \sum_{i=1}^n x_i)y$.
- Positive monomial: $P_n(x) = \prod_{i=1}^n x_i = x_1 \dots x_n$.
- (Ishikawa 2011) P_n can be quadratized using $\lfloor \frac{n-1}{2} \rfloor$ auxiliary variables.
- (Fix 2011) $n - 1$ variables suffice for any symmetric function.

- Based on ad hoc arguments.
- In *DAM* (2016), we proposed a generic approach based on a general representation theorem for discrete functions.

- Recent improvements: Boros, Crama, Rodrìguez-Heck (2018).

- Recent improvements: Boros, Crama, Rodrìguez-Heck (2018).
- Assume: $k \geq \frac{n}{2}$.

- Recent improvements: Boros, Crama, Rodrìguez-Heck (2018).
- Assume: $k \geq \frac{n}{2}$.

Function	**Lower Bound**	**Upper Bound**	**Previous UB**
Symmetric	$\Omega(\sqrt{n})$	$2\lceil\sqrt{n+1}\rceil$	$n-2$
Positive monomial	$\lceil log(n)\rceil - 1$	$\lceil log(n)\rceil - 1$	$\lfloor\frac{n-1}{2}\rfloor$
k-out-of-n	$\lceil log(k)\rceil - 1$	$\lceil log(k)\rceil$	$\lceil\frac{n}{2}\rceil$
Parity	$\lceil log(n)\rceil - 1$	$\lceil log(n)\rceil - 1$	$\lfloor\frac{n-1}{2}\rfloor$

Basic idea:

Positive monomial: upper bound

Assume that $n = 2^k$. Then,

$$g(x,y) = (\sum_{i=1}^{n} x_i - \sum_{j=0}^{k-1} 2^j y_j)^2$$

is a quadratization of the positive monomial $P_n(x) = \prod_{i=1}^{n} x_i$ using $k = log(n)$ auxiliary variables.

Proof. (Sketch.) For all (x,y), $g(x,y) \geq 0$ and $0 \leq \sum_{j=0}^{k-1} 2^j y_j \leq n-1$.
If $\sum_{i=1}^{n} x_i < n$, then one can make $g(x,y) = 0$.
If $\sum_{i=1}^{n} x_i = n$, then $\min_y g(x,y) = 1$.

Refinement:

Positive monomial: improved upper bound

Assume that $n = 2^k$. Then,

$$g(x,y) = \frac{1}{2}\left(\sum_{i=1}^{n} x_i - \sum_{j=1}^{k-1} 2^j y_j\right)\left(\sum_{i=1}^{n} x_i - \sum_{j=1}^{k-1} 2^j y_j - 1\right)$$

is a quadratization of the positive monomial $P_n(x) = \prod_{i=1}^{n} x_i$ using $k-1 = log(n) - 1$ auxiliary variables.

We also have a matching lower bound.

Positive monomial: lower bound

Every quadratization of the positive monomial $P_n(x) = \prod_{i=1}^{n} x_i$ must use at least $log(n) - 1$ auxiliary variables.

Positive monomial: lower bound

Every quadratization of the positive monomial $P_n(x) = \prod_{i=1}^{n} x_i$ must use at least $log(n) - 1$ auxiliary variables.

Proof. (Sketch.) For a quadratization $g(x, y)$, define $r(x) = \prod_{y \in \{0,1\}^m} g(x, y)$.

- The degree of $r(x)$, $deg(r)$, is at most 2^{m+1}, by definition.
- For every x with $|x| < n$, there is $y \in \{0, 1\}^m$ such that $g(x, y) = 0$. So, $r(x) = 0$.
- When $|x| = n$, $g(x, y) \geq 1$ for all $y \in \{0, 1\}^m$, and hence $r(x) \geq 1$.
- It follows that $deg(r) = n$
- We get: $deg(r) = n \leq 2^{m+1}$.

- Similar ideas and bounds extend to k-out-of-n (“exact” or “at least”) and to parity functions.

- Similar ideas and bounds extend to k-out-of-n (“exact” or “at least”) and to parity functions.
- For symmetric functions:

General symmetric functions: Lower bound

- Some symmetric functions have no m-quadratization using less than $\sqrt{n}$ auxiliary variables. (Anthony, Boros, Crama, Gruber 2017).
- Every symmetric function has a quadratization using at most $2\lceil\sqrt{n+1}\rceil$ auxiliary variables. (Boros, Crama, Rodrìguez-Heck 2018.)

- Similar ideas and bounds extend to k-out-of-n (“exact” or “at least”) and to parity functions.
- For symmetric functions:

General symmetric functions: Lower bound

- Some symmetric functions have no m-quadratization using less than $\sqrt{n}$ auxiliary variables. (Anthony, Boros, Crama, Gruber 2017).
- Every symmetric function has a quadratization using at most $2\lceil\sqrt{n+1}\rceil$ auxiliary variables. (Boros, Crama, Rodrìguez-Heck 2018.)

- **Note**: Proofs rely on techniques developed for the analysis of the size of threshold circuits, and of slicing or covering of the vertices of the hypercube by hyperplanes (work by Alon, Füredi, Linial, Radhakrishnan, Saks, etc.)

Outline

Upper bound

- Worst-case bound based on termwise quadratizations:

Corollary

Every function f of n variables has a quadratization involving at most $O(log(n)\,2^n)$ auxiliary variables.

Upper bound

- Worst-case bound based on termwise quadratizations:

Corollary

Every function f of n variables has a quadratization involving at most $O(log(n)\,2^n)$ auxiliary variables.

- We can prove:

Theorem: upper bound (*Math. Prog.* (2017))

Every function f of n variables has a quadratization involving at most $O(2^{n/2})$ auxiliary variables.

Pairwise cover

Based on a construction using small *pairwise covers:*

Pairwise cover

A hypergraph $\mathcal{H}$ is a *pairwise cover* of $\{1, \ldots, n\}$ if, for every $S \subseteq \{1, \ldots, n\}$ with $|S| \geq 3$, there are sets $A, B \in \mathcal{H}$ such that $|A| < |S|$, $|B| < |S|$ and $A \cup B = S$.

- Pairwise covers are (almost) identical to so-called 2-*bases* investigated by Erdös, Füredi and Katona (2006), Frein, Lévêque and Sebö (2008), Ellis and Sudakov (2011).

Pairwise cover

Based on a construction using small *pairwise covers:*

Pairwise cover

A hypergraph $\mathcal{H}$ is a *pairwise cover* of $\{1, \ldots, n\}$ if, for every $S \subseteq \{1, \ldots, n\}$ with $|S| \geq 3$, there are sets $A, B \in \mathcal{H}$ such that $|A| < |S|$, $|B| < |S|$ and $A \cup B = S$.

- Pairwise covers are (almost) identical to so-called 2-*bases* investigated by Erdös, Füredi and Katona (2006), Frein, Lévêque and Sebö (2008), Ellis and Sudakov (2011).
- $\mathcal{P}(even) =$ all subsets of even integers in $\{1, \ldots, n\}$.
- $\mathcal{P}(odd) =$ all subsets of odd integers in $\{1, \ldots, n\}$.
- $\mathcal{H} = \mathcal{P}(even) \cup \mathcal{P}(odd)$ is a “small” pairwise cover with size $O(2^{n/2})$.

Pairwise covers and universal sets

We prove:

Theorem: From pairwise cover to quadratization

If there exists a pairwise cover of $\{1, \ldots, n\}$ of size m, then every pseudo-Boolean function has an m-quadratization.

Pairwise covers and universal sets

We prove:

Theorem: From pairwise cover to quadratization

If there exists a pairwise cover of $\{1, \ldots, n\}$ of size m, then every pseudo-Boolean function has an m-quadratization.

- Idea of the proof: write $\prod_{i \in S} x_i = (\prod_{j \in A} x_j)(\prod_{k \in B} x_k)$; substitute y_A for $\prod_{j \in A} x_j$ and y_B for $\prod_{k \in B} x_k$.

Pairwise covers and universal sets

We prove:

Theorem: From pairwise cover to quadratization

If there exists a pairwise cover of $\{1, \ldots, n\}$ of size m, then every pseudo-Boolean function has an m-quadratization.

- Idea of the proof: write $\prod_{i \in S} x_i = (\prod_{j \in A} x_j)(\prod_{k \in B} x_k)$; substitute y_A for $\prod_{j \in A} x_j$ and y_B for $\prod_{k \in B} x_k$.
- There are pairwise covers with size $O(2^{n/2})$.

Pairwise covers and universal sets

We prove:

Theorem: From pairwise cover to quadratization

If there exists a pairwise cover of $\{1, \dots, n\}$ of size m, then every pseudo-Boolean function has an m-quadratization.

- Idea of the proof: write $\prod_{i \in S} x_i = (\prod_{j \in A} x_j)(\prod_{k \in B} x_k)$; substitute y_A for $\prod_{j \in A} x_j$ and y_B for $\prod_{k \in B} x_k$.
- There are pairwise covers with size $O(2^{n/2})$.
- Hence, every pseudo-Boolean function has a quadratization with $O(2^{n/2})$ auxiliary variables.

By-products and corollaries

Similarly:

Fixed-degree functions

For every fixed d, every pseudo-Boolean function of degree d has a quadratizations with $O(n^{d/2})$ auxiliary variables.

Outline

Lower bound

- Any good lower bound on the number of auxiliary variables?

Lower bound

- Any good lower bound on the number of auxiliary variables?

Theorem: lower bound (*Math. Prog.* (2017))

There are pseudo-Boolean functions of n variables for which every quadratization must involve at least $\Omega(2^{n/2})$ auxiliary variables.

Lower bound

- Any good lower bound on the number of auxiliary variables?

Theorem: lower bound (*Math. Prog.* (2017))

There are pseudo-Boolean functions of n variables for which every quadratization must involve at least $\Omega(2^{n/2})$ auxiliary variables.

- This lower bound matches the $O(2^{n/2})$ upper bound.
- Non constructive proof based on dimensionality argument: if too few auxiliary variables, then we cannot generate the whole vector space of pseudo-Boolean functions.

Proof

Main idea: dimensionality argument.

- Suppose $f(x)$ has an m-quadratization $g(x, y)$: for all $x \in \{0,1\}^n$,

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

- $g(x, y)$ has $O(n^2 + m^2)$ coefficients.
- But the vector space of pseudo-Boolean functions on n variables has dimension 2^n.
- So, if m is too small, we cannot generate the whole vector space.
- It follows that if m additional variables suffice to quadratize any pseudo-Boolean function in n variables, then m is $\Omega(2^{n/2})$.

Must be refined, since the relation (1) is not linear.

By-products and corollaries

Almost all functions...

Almost all pBFs need at least $\Omega(2^{n/2})$ auxiliary variables to be quadratized.

By-products and corollaries

Almost all functions...

Almost all pBFs need at least $\Omega(2^{n/2})$ auxiliary variables to be quadratized.

- But no good lower bound for any specific function...

By-products and corollaries

Almost all functions...

Almost all pBFs need at least $\Omega(2^{n/2})$ auxiliary variables to be quadratized.

- But no good lower bound for any specific function...
- We have seen earlier that pBFs of degree d can be quadratized using $O(n^{d/2})$ auxiliary variables.

By-products and corollaries

Almost all functions...

Almost all pBFs need at least $\Omega(2^{n/2})$ auxiliary variables to be quadratized.

- But no good lower bound for any specific function...
- We have seen earlier that pBFs of degree d can be quadratized using $O(n^{d/2})$ auxiliary variables.

Fixed-degree functions

For every fixed d, there are degree-d pseudo-Boolean functions of n variables for which any quadratization must involve at least $\Omega(n^{d/2})$ auxiliary variables.

By-products and corollaries

Almost all functions...

Almost all pBFs need at least $\Omega(2^{n/2})$ auxiliary variables to be quadratized.

- But no good lower bound for any specific function...
- We have seen earlier that pBFs of degree d can be quadratized using $O(n^{d/2})$ auxiliary variables.

Fixed-degree functions

For every fixed d, there are degree-d pseudo-Boolean functions of n variables for which any quadratization must involve at least $\Omega(n^{d/2})$ auxiliary variables.

Comments and interpretation

- All quadratization procedures considered here can be implemented in time polynomial in n and t, where t is the number of terms of the function.

Comments and interpretation

- All quadratization procedures considered here can be implemented in time polynomial in n and t, where t is the number of terms of the function.
- All termwise procedures require $\Omega(log(n)\,t)$ auxiliary variables (compare with t auxiliary variables for linearization).

Comments and interpretation

- All quadratization procedures considered here can be implemented in time polynomial in n and t, where t is the number of terms of the function.
- All termwise procedures require $\Omega(log(n)\, t)$ auxiliary variables (compare with t auxiliary variables for linearization).
- Since almost all functions contain $\Omega(2^n)$ terms, termwise quadratization usually requires $\Omega(log(n)\, 2^n)$ auxiliary variables.

Comments and interpretation

- All quadratization procedures considered here can be implemented in time polynomial in n and t, where t is the number of terms of the function.
- All termwise procedures require $\Omega(log(n)\, t)$ auxiliary variables (compare with t auxiliary variables for linearization).
- Since almost all functions contain $\Omega(2^n)$ terms, termwise quadratization usually requires $\Omega(log(n)\, 2^n)$ auxiliary variables.
- Our results improve this from $\Omega(log(n)\, 2^n)$ to $\Omega(2^{n/2})$ auxiliary variables (or from $\Omega(log(n)\, n^d)$ to $\Omega(n^{d/2})$ for degree-d functions).

Comments and interpretation

- All quadratization procedures considered here can be implemented in time polynomial in n and t, where t is the number of terms of the function.
- All termwise procedures require $\Omega(log(n)\, t)$ auxiliary variables (compare with t auxiliary variables for linearization).
- Since almost all functions contain $\Omega(2^n)$ terms, termwise quadratization usually requires $\Omega(log(n)\, 2^n)$ auxiliary variables.
- Our results improve this from $\Omega(log(n)\, 2^n)$ to $\Omega(2^{n/2})$ auxiliary variables (or from $\Omega(log(n)\, n^d)$ to $\Omega(n^{d/2})$ for degree-d functions).
- Our quadratization procedure based on pairwise covers probably yields small quadratizations as a function of t as well, but we have no generic bounds in this case.

Outline

Conclusions

- Tight lower and upper bound for special classes of functions (e.g., for positive monomials).
- Tight lower and upper bounds on the number of auxiliary variables required for arbitrary and for fixed-degree functions.
- Structure and properties of quadratizations are poorly understood.
- Computational tests show that pairwise covers yield good quadratizations and efficient approaches for certain classes of nonlinear problems, but linearization remains competitive in most cases.
- Many intriguing questions and conjectures, much computational and theoretical work to be done.

Additional references

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

C. Buchheim and G. Rinaldi, Efficient reduction of polynomial zero-one optimization to the quadratic case, *SIAM Journal on Optimization* 18 (2007) 1398-1413.

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.

Additional 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.