Global Optimization of binary polynomial programs

Sourour Elloumi, Amélie Lambert, Arnaud Lazare

CEDRIC - CNAM
UMA - ENSTA

PGMO Days
November 14^{th}, 2017

Plan

Introduction

Introduction

We are interested in solving the following problem:

$$(P)\begin{cases} \min f(x) \\ \text{s.t.} \\ \quad x_i \in \{0,1\} \qquad i = 1, ..., n \end{cases}$$

where f is a polynomial with n binary variables. (P) is NP-hard whenever $deg(f) \geq 2$.

Difficulties come from

- Non-convexity of f
- Integer variables

Introduction

Only a few efficient algorithms exist to find the global minimum of a mixed-integer non-convex polynomial:

- Some are based on a hierarchy of relaxations (Moment/SOS) [Lasserre, GloptiPoly, 2003]
- Others use reformulation/relaxation techniques before using a Branch and Bound

Introduction

Only a few efficient algorithms exist to find the global minimum of a mixed-integer non-convex polynomial:

- Some are based on a hierarchy of relaxations (Moment/SOS) [Lasserre, GloptiPoly, 2003]
- Others use reformulation/relaxation techniques before using a Branch and Bound

Convex reformulation

- The continuous relaxation of (P) is not necessarily convex
- We want an equivalent reformulation of (P) such as its continuous relaxation is convex
- Possibly lift (P) in a higher dimensional space
- Idea: Choose the best reformulation among a family of "convex" reformulations (continous relaxation bound)
- 2 equivalent convex reformulations

Convex quadratic reformulation

Example

Let us consider the following program:

$$(Ex) \begin{cases} \min 5x_4x_3x_2 - 8x_1 \\ \text{s.t.} \\ \quad x_i \in \{0,1\} \qquad\qquad \text{i=1,...,4} \end{cases}$$

We want to find a quadratic reformulation of (Ex) in a higher dimensional space

Example - Quadratization

$$(Ex)\begin{cases} \min 5x_4x_3x_2 - 8x_1 \\ \text{s.t.} \\ \quad x_i \in \{0,1\} \qquad \text{i=1,...,4} \end{cases}$$

$$(Exquad)\begin{cases} \min 5x_5x_2 - 8x_1 \\ \text{s.t.} \\ x_5 \leq x_4 \\ x_5 \leq x_3 \\ x_5 \geq x_3 + x_4 - 1 \\ x_5 \geq 0 \\ x_1, x_2, x_3, x_4, x_5 \in \{0,1\} \end{cases}$$

- 5 variables (1 new variable for each monomial of degree 3)
- Up to $\alpha - 2$ new variables for each monomial of degree α
- 4 constraints ($g_5(x) \geq 0$)

Quadratic reformulation

$\longrightarrow$ We found a quadratic reformulation (QP) of (P) in a higher dimension

$\longrightarrow$ $(QP) \iff (P)$

$$
(QP) \begin{cases} \min f(x) = \sum_i \sum_j q_{ij} x_i x_j + \sum_i c_i x_i \\ \text{s.t.} \\ g_j(x) \geq 0 \qquad j = n+1, ..., m \\ x_i \in \{0,1\} \qquad i = 1, ..., m \end{cases}
$$

Quadratic convex reformulation

$\longrightarrow$ We want to find a convexification using the identity $x_i^2 = x_i$

$\longrightarrow$ We want to compute $\sigma \in \mathbb{R}^n$ to find a convex reformulation f_σ of f

$$(QP_\sigma) \begin{cases} \min f_\sigma(x) = \sum_i \sum_j q_{ij} x_i x_j + \sum_i c_i x_i + \sum_i \sigma_i (x_i^2 - x_i) \\ \text{s.t.} \\ g_j(x) \geq 0 \qquad j = n+1, ..., m \\ x_i \in \{0, 1\} \qquad i = 1, ..., m \end{cases}$$

$\longrightarrow$ How to compute a "good" σ ?

PSD based convexification

Theorem (MIQCR, 2012)

$$
(PSD_D)\begin{cases}
\min f(X,x) = \sum_i \sum_j q_{ij} X_{ij} + \sum_i c_i x_i \\
s.t. \\
g_j(x) \geq 0 \qquad j = n+1, ..., m \\
X_{ii} = x_i \qquad i = 1, ..., m \\
X - xx^T \succeq 0 \\
X \in S^m, x \in \mathbb{R}^m
\end{cases} \quad (1)
$$

Possible values of σ_i are the optimal values of the dual variables associated with constraint (1)*.*

$\longrightarrow Q + diag(\sigma) \succeq 0$

Summary of the quadratization method

- Find a quadratization (QP) of (P). Moving from a n-dimensional space to a m-dimensional space
- Solve (PSD_D) to compute values of σ and find a convexification of (QP)
- Solve (QP_σ) using a Branch and Bound. At each node we solve a convex quadratic relaxation

Convex polynomial reformulation

Idea - Equivalent polynomial reformulation

$\longrightarrow$ We want to reformulate f in a convex polynomial with the same degree

$\longrightarrow$ Idea: Find a perturbation of the Hessian matrix $H(x)$ to make it positive semidefinite (PSD) for all $x \in [0,1]^n$

$\longrightarrow$ Method:

- Let $\lambda \in \mathbb{R}^n$ and $f_\lambda(x) = f(x) - \sum_i \lambda_i(x_i^2 - x_i)$
- $f(x) = f_\lambda(x)$, $\forall x \in \{0,1\}^n$
- We search for λ such as f_λ is convex over $[0,1]^n$

Equivalent polynomial reformulation

$\longrightarrow$ We obtain the following convex reformulation of (P):

$$(P) \iff (P_\lambda) \begin{cases} \min f(x) - \sum_i \lambda_i (x_i^2 - x_i) \\ \text{s.t.} \\ x_i \in \{0,1\} \qquad i = 1, \ldots, n \end{cases}$$

$\longrightarrow$ Both problems are equivalent and have the same degree and the same variables

$\longrightarrow$ We want to compute λ such as f_λ is convex

Hessian matrix and approximation

- Goal: Find λ such as f_λ is convex
- Let $H_\lambda(x)$ be the Hessian matrix of f_λ, $H_\lambda(x) = H(x) - diag(\lambda)$
- We want λ such as $\forall x \in [0,1]^n,\ H_\lambda(x) \succeq 0$
- Method
    1. Include $H(x)$ in an interval matrix C
    2. Find λ so that $\forall M \in C,\ M - diag(\lambda) \succeq 0$

Interval matrices

Definition (Interval matrix)

An interval matrix $\mathcal{C}$ is a matrix whose elements are interval numbers.

$\longrightarrow$ Possibility to bound each entry of the Hessian matrix

$$\forall x \in [0,1]^n, \quad \underline{c_{ij}} \leq h_{ij}(x) = \frac{\partial^2 f}{\partial x_i x_j} \leq \overline{c_{ij}}$$

$\longrightarrow$ We want to compute an interval matrix $C : \forall i,j, \; C_{ij} = [\underline{c_{ij}}, \overline{c_{ij}}]$

$\longrightarrow$ Remark : It may be difficult to compute tight bounds

Approximation of the Hessian matrix

Let f be a binary polynomial defined by $f(x) = \sum_i c_i \prod_j x_{i_j}$. The following inequalities hold

$$\underline{c_{ij}} = \sum_{\substack{\text{monomial k such as} \\ x_i \wedge x_j \in k \\ c_k < 0}} c_k \leq \frac{\partial^2 f}{\partial x_i x_j} \leq \sum_{\substack{\text{monomial k such as} \\ x_i \wedge x_j \in k \\ c_k > 0}} c_k = \overline{c_{ij}}$$

Example

Let f be the polynomial
$f : (x_1, x_2, x_3, x_4) \mapsto 2x_1 - 2x_2x_3x_4 + 3x_2x_3 - 3x_1x_2x_3x_4$

$$H(x) = \begin{pmatrix} 0 & -3x_3x_4 & -3x_2x_4 & -3x_2x_3 \\ -3x_3x_4 & 0 & 3 - 3x_1x_4 - 2x_4 & -2x_3 - 3x_1x_3 \\ -3x_2x_4 & 3 - 3x_1x_4 - 2x_4 & 0 & -2x_2 - 3x_1x_2 \\ -3x_2x_3 & -2x_3 - 3x_1x_3 & -2x_2 - 3x_1x_2 & 0 \end{pmatrix}$$

$$M = \begin{pmatrix} [0,0] & [-3,0] & [-3,0] & [-3,0] \\ \text{[-3,0]} & [0,0] & [-2,3] & [-5,0] \\ \text{[-3,0]} & [-2,3] & [0,0] & [-5,0] \\ \text{[-3,0]} & [-5,0] & [-5,0] & [0,0] \end{pmatrix}$$

Scaled Gerschgorin Theorem

$\longrightarrow$ Idea : Find a perturbation of the diagonal of $H(x)$ using a vector λ

Theorem (Scaled Gerschgorin)

Let C be an interval matrix, we define $\lambda \in \mathbb{R}^n$ by

$$\lambda_i = \min \left[0, \left(\underline{c_{ii}} - \sum_{j \neq i} \max(|\underline{c_{ij}}|, |\bar{c_{ij}}|)\right)\right] \quad \forall i \in \{1, ..., n\}$$

Then $C - diag(\lambda)$ is positive semidefinite.

$\longrightarrow$ $H(x) - diag(\lambda)$ is positive semidefinite $\forall x \in [0, 1]^n$.

Reformulated problem

$\longrightarrow$ We obtain the following "convex" equivalent problem (P_λ)

$$(P_\lambda)\begin{cases} \min f(x) - \sum_i \lambda_i(x_i^2 - x_i) \\ \text{s.t.} \\ x_i \in \{0,1\} \qquad i = 1,\ldots,n \end{cases}$$

$\longrightarrow$ Where $\lambda_i = \min\left[0, \left(\underline{c_{ii}} - \sum_{j\neq i} \max(|\underline{c_{ij}}|, |\bar{c_{ij}}|)\right)\right]$

$\longrightarrow$ In our example, $\lambda_i = (-9, -13, -13, -9)$

Example

⟶ Let f be the polynomial previously defined by
$f : (x_1, x_2, x_3, x_4) \mapsto 2x_1 - 2x_2x_3x_4 + 3x_2x_3 - 3x_1x_2x_3x_4$.
It's minimum over $\{0, 1\}$ is 0

⟶ Convex quadratic reformulation:

- 7 variables
- 12 constraints
- Continous relaxation -1.75

⟶ Convex polynomial reformulation:

- 4 variables
- 0 constraint
- Continous relaxation -1.85

Numerical results

Quadratization algorithm

$\longrightarrow$ Results depend on the way we quadratize the objective function

$\longrightarrow$ In each monomial, we sort out the variables in increasing order

$\longrightarrow$ We replace each product of variables by a new variable using the following rule:

- for a monomial of degree 4: $\alpha x_1 x_2 x_3 x_4 \longrightarrow \alpha x_5 x_6$
  with $x_5 = x_1 x_2$ and $x_6 = x_3 x_4$ (first appearance of the product)
- for a monomial of degree 3: $\alpha x_1 x_2 x_3 \longrightarrow \alpha x_5 x_3$
  with $x_5 = x_1 x_2$ (first appearance of the product)

Results - Convex quadratic reformulation

$\longrightarrow$ Uniform coefficients in $[-1, 1]$

$\longrightarrow$ Uniform degrees in $\{1, .., 4\}$

$\longrightarrow$ Solvers: CSDP, Cplex

Var/Mon	Nb var	Nb cons	SDP	B&B nodes	B&B	Gap
10/50	37	108	0.1s	30	2.6s	112%
10/100	49	156	0.1s	0	2.3s	29%
20/50	84	256	0.3s	333	2.5s	49%
20/200	133	452	1.1s	0	2.5s	57%
20/400	181	644	1.8s	3747	2.5s	82%
50/500	386	1344	5.3s	224777	147.1s	72%

Var/Mon	Nb var	Nb cons	SDP	B&B nodes	B&B	Gap
	145	493	1.6s	38147	26.6s	67%

Results - Convex polynomial reformulation

Var/Mon	B&B nodes	B&B	Gap
10/50	2	1s	220%
10/100	11	0.8s	62%
20/50	2	1s	91%
20/200	111	1.5s	111%
20/400	464	4.1s	115%
50/500	3112	37.9s	105%

Var/Mon	B&B nodes	B&B	Gap
	617	7.7s	117%
	38147	26.6s	67%

Conclusion

Conclusion

- We have presented two different convex reformulations
- Polynomial reformulation: First method dealing with direct convex reformulation of a polynomial (without changing the degree)

$\longrightarrow$ Many improvements are possible

- Quadratic reformulation : Add cuts to the SDP, find a better quadratization,...
- Polynomial reformulation : More precise computation of λ, interval matrix,...