Lower bounds certification for multivariate real functions using SDP

Joint Work with B. Werner, S. Gaubert and X. Allamigeon

Victor MAGRON

LIX/INRIA, École Polytechnique

LIX PhD Seminar Friday 18 th January

Two Problems

$K \subset \mathbb{R}^n$: a compact set
$f : K \to \mathbb{R}$: a real multivariate function

Two challenging problems:

1. $\inf_{x \in K} f(x)$ when f is a multivariate polynomial of degree d
   Number of variables n is large, no sparsity $\Longrightarrow$ very hard to solve using Interval Arithmetic

Example:

$K := [0, 1]^n$, random numbers $(r_i)_{1 \leq i \leq n}$:

$f_d := \left(\frac{1}{n} \sum_{i=1}^{n} \frac{4}{r_i^2} x_i (r_i - x_i)\right)^{\lceil d/2 \rceil}$, the range of f_d is $[0, 1]$

2. $\inf_{x \in K} f(x)$ when f is a multivariate real function involving transcendental univariate functions

Contents

0. Solving Polynomial Problems using Sum of Squares (SOS) and Semidefinite Programming (SDP)
1. Lower bounds of multivariate polynomial with large number of variables
2. Lower bounds of transcendental multivariate functions

SOS and SDP Relaxations

Polynomial Optimization Problem (POP):

Let $f, g_1, \cdots, g_m \in \mathbb{R}[X_1, \cdots, X_n]$
$K_{pop} := \{x \in \mathbb{R}^n : g_1(x) \geq 0, \cdots, g_m(x) \geq 0\}$ is the feasible set
General POP: compute $f^*_{pop} = \inf_{x \in K_{pop}} f(x)$

Example:

$f := 10 - x_1^2 - x_2^2$, $g_1 := 1 - x_1^2 - x_2^2$
$K_{pop} := \{x \in \mathbb{R}^2 : g_1(x) \geq 0\}$ is the feasible set

SOS and SDP Relaxations

Convexify the problem:

$f^*_{pop} = \inf_{x \in K_{pop}} f_{pop}(x) = \inf_{\mu \in \mathcal{P}(K_{pop})} \int f_{pop} \, d\mu$, where $\mathcal{P}(K_{pop})$ is the set of all probability measures μ supported on the set K_{pop}.

Equivalent formulation:

$f^*_{pop} = \min \{ L(f) : L : \mathbb{R}[X] \to \mathbb{R}$ linear, $L(1) = 1$ and each $\mathcal{L}_{g_j}$ is SDP $\}$, with $g_0 = 1$, $\mathcal{L}_{g_0}, \cdots, \mathcal{L}_{g_m}$ defined by:

$$\begin{array}{rcccl} \mathcal{L}_{g_j} & : & \mathbb{R}[X] \times \mathbb{R}[X] & \to & \mathbb{R} \\ & & (p, q) & \mapsto & L(p \cdot q \cdot g_j) \end{array}$$

SOS and SDP Relaxations: Lasserre Hierarchy

- $\mathcal{B} := (X^\alpha)_{\alpha \in \mathbb{N}^n}$: the monomial basis and $y_\alpha = L(X^\alpha)$, this identifies L with the infinite series $y = (y_\alpha)_{\alpha \in \mathbb{N}^n}$
- Infinite moment matrix M:
  $M(y)_{u,v} := L(u \cdot v), \; u, \, v \in \mathcal{B}$
- Localizing matrix $M(g_j y)$:
  $M(g_j y)_{u,v} := L(u \cdot v \cdot g_j), \; u, \, v \in \mathcal{B}$
- $k \geq k_0 := \max\{\lceil \deg f_{pop} \rceil / 2, \lceil \deg g_0 / 2 \rceil, \cdots, \lceil \deg g_m / 2 \rceil\}$
  Index $M(y)$ and $M(g_j y)$ with elements in $\mathcal{B}$ of degree at most k, it gives the semidefinite relaxations hierarchy:

$$Q_k : \begin{cases} \inf_y L(f) & = \int f_\alpha \, x^\alpha \, d\mu(x) = \sum_\alpha f_\alpha \, y_\alpha \\ M_{k - \lceil \deg g_j / 2 \rceil}(g_j y) & \succcurlyeq 0, \qquad 0 \leq j \leq m, \\ y_1 & = 1 \end{cases}$$

SOS and SDP Relaxations

Convergence Theorem [Lasserre]:

The sequence $\inf(Q_k)_{k \geq k_0}$ is non-decreasing and under the SOS assumption converges to f^*_{pop}.

SDP relaxations:

Many solvers (e.g. Sedumi, SDPA) solve the pair of (standard form) semidefinite programs:

$$(SDP) \left\{ \begin{array}{lll} \mathcal{P}: & \min\limits_{y} & \sum\limits_{\alpha} c_\alpha y_\alpha \\ & \text{subject to} & \sum\limits_{\alpha} F_\alpha \, y_\alpha - F_0 \succcurlyeq 0 \\ \mathcal{D}: & \max\limits_{Y} & \text{Trace}\,(F_0 \, Y) \\ & \text{subject to} & \text{Trace}\,(F_\alpha \, Y) = c_\alpha \end{array} \right.$$

SDP relaxation Q_k at order $k \geq \max_j \{\lceil \deg f_{pop}/2 \rceil, \lceil \deg g_j/2 \rceil\}$:

- $\mathcal{O}(n^{2k})$ moment variables
- linear matrix inequalities (LMIs) of size $\mathcal{O}(n^k)$

polynomial in n, exponential in k

On our example:

$K := [0,1]^n$, random numbers $(r_i)_{1 \leq i \leq n}$:

$$f_d := \Big(\frac{1}{n} \sum_{i=1}^{n} \frac{4}{r_i^2} x_i (r_i - x_i)\Big)^{\lceil d/2 \rceil}$$

$\deg g_j = 1$, $k \geq d \implies$ at least $\mathcal{O}(n^{2d})$ moment variables with LMIs of size $\mathcal{O}(n^d)$!!

Large-scale POP

Multivariate Taylor-Models Underestimators:

- $f : K \to \mathbb{R}$ is a multivariate polynomial
- Consider a minimizer guess x_c obtained by heuristics
- Let q_{x_c} be the quadratic form defined by:

$$\begin{aligned} q_{x_c} : K &\longrightarrow \mathbb{R} \\ x &\longmapsto f(x_c) + \mathcal{D}_f(x_c)\,(x - x_c) \\ &\quad + \frac{1}{2}(x - x_c)^T \mathcal{D}_f^2(x_c)\,(x - x_c) + \lambda(x - x_c)^2 \end{aligned}$$

with $\lambda := \min_{x \in K} \{\lambda_{\min}(\mathcal{D}_f^2(x) - \mathcal{D}_f^2(x_c))\}$

Theorem:

$\forall x \in K, f(x) \geq q_{x_c}$, that is q_{x_c} understimates f on K.
q_{x_c} is called a quadratic cut.

How to compute λ? How to compute a lower bound of f?

- $\lambda := \min_{x \in K} \{\lambda_{\min}(\mathcal{D}_f^2(x) - \mathcal{D}_f^2(x_c))\}$
- Bound the Hessian difference on K by POP (using SDP relaxations) to get $\bar{\mathcal{D}}_f^2$:
- Define the symmetric matrix B containing the bounds on the entries of $\bar{\mathcal{D}}_f^2$.
- Let $\mathcal{S}^n$ be the set of diagonal matrices of sign.
  $\mathcal{S}^n := \{\text{diag}\,(s_1, \cdots, s_n),\ s_1 = \pm 1, \cdots\ s_n = \pm 1\}$
  $\lambda := \lambda_{\min}(\bar{\mathcal{D}}_f^2 - \mathcal{D}_f^2(x_c)\bar{I})$: minimal eigenvalue of an interval matrix

Robut Optimization with Reduced Vertex Set [Calafiore, Dabbene]

The robust interval SDP problem $\lambda_{\min}(\bar{\mathcal{D}}_f^2 - \mathcal{D}_f^2(x_c)\bar{I})$ is equivalent to the following SDP in the single variable $t \in \mathbb{R}$:

$$\begin{cases} \min & -t \\ \text{s.t.} & -t\, I - \mathcal{D}_f^2(x_c) - S\, B\, S \succeq 0,\ S = \text{diag}\,(1,\ \tilde{S}),\ \forall \tilde{S} \in \mathcal{S}^{n-1} \end{cases}$$

Solving the previous SDP is expensive because the dimension of $\mathcal{S}^n$ grows exponentially. Instead, we can underestimate λ:

- Write $\bar{\mathcal{D}_f^2} - \mathcal{D}_f^2(x_c)\bar{I} := \bar{X} + \bar{Y}$ with
  $\bar{X}_{ij} := [\frac{a_{ij} + b_{ij}}{2}, \frac{a_{ij} + b_{ij}}{2}]$ and $\bar{Y}_{ij} := [-\frac{b_{ij} - a_{ij}}{2}, \frac{b_{ij} - a_{ij}}{2}]$
- $\lambda_{\min}(\bar{X} + \bar{Y}) \geq \lambda_{\min}(\bar{X}) + \lambda_{\min}(\bar{Y}) = \lambda_{\min}(\bar{X}) - \lambda_{\max}(-\bar{Y})$
- $\lambda_{\max}(\bar{Y}) \leq \max_i \sum_j \frac{b_{ij} - a_{ij}}{2}$

Computing a lower bound of $\lambda_{\min}(\bar{X})$ is easier because $\bar{X}$ is a real matrix. We can do it again by SDP: $\begin{cases} \min & -t \\ \text{s.t.} & -t\, I - \bar{X} \succeq 0 \end{cases}$

... and how to compute a lower bound of the polynomial f?

Large-scale POP
Computing lower bounds

Input: f, box K, SDP relaxation order k, control points sequence $s = (x_1) \in K$, n_{cuts} (final number of quadratic cuts)

Output: lower bound m of f

$cuts := 1$

while $cuts \leq n_{cuts}$ **do**

  For $c \in \{1, \dots, \#s\}$: compute λ using robust SDP or $\lambda_{\min}$ approximation and compute q_{x_c}

  $f_p := \max_{1 \leq c \leq p} q_{x_c}$, $K_{pop} := \{x \in K : z \geq q_{x_1}(x), \cdots, z \geq q_{x_p}(x)\}$

  Compute a lower bound m of f_p by POP at the SDP relaxation order k: $m \leq \inf_{x \in K_{pop}} z$

  $x_{opt} :=$ `guess_argmin` (f_p): a minimizer candidate for f_p

  $s := s \cup \{x_{opt}\}$

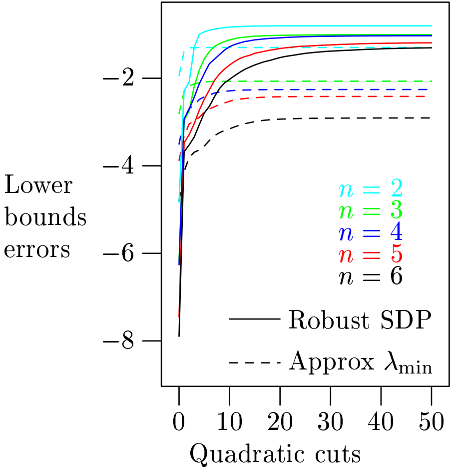
  $cuts := cuts + 1$

done

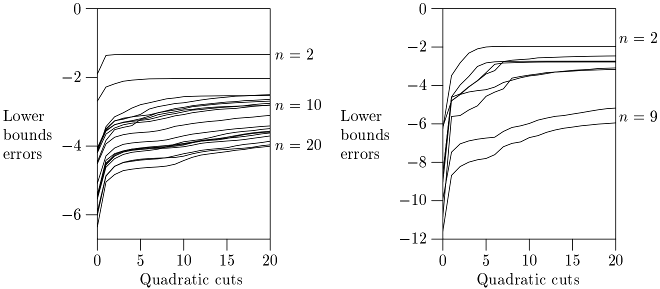
$$K := [0,1]^n$$

$$f_6 := \Big(\frac{1}{n}\sum_{i=1}^{n}\frac{4}{r_i^2}x_i(r_i - x_i)\Big)^3$$

- We compare the quality of the successive lower bounds (previous algorithm) with different λ underestimators
- $\lambda_{\text{robust}} \geq \lambda_{\text{approx}} \implies$ Better quadratic approximations when using the Robust SDP approach



- When n is large, Robust SDP approach is too expensive. It becomes impossible to compute λ and the quadratic cuts q_{x_c}.



(a) $d = 4$

(b) $d = 6$

- Bottleneck: computation of the $n(n+1)$ bounds of the Hessian entries $\bar{\mathcal{D}}_f^2 - \mathcal{D}_f^2(x_c)\bar{I}$ (multivariate polynomial of degree $d-2$)

Bounding multivariate transcendental functions

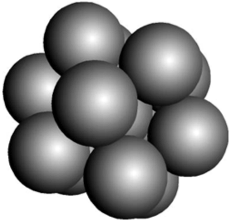
- Now, consider a semialgebraic compact set $K \subset \mathbb{R}^n$ and $f : K \to \mathbb{R}$ a multivariate transcendental function
- We want to compute a precise lower bound of f. The previous approach only gives a hierarchy of coarse bounds

Motivations?
How to approach the univariate transcendental functions involved in f?

Kepler Conjecture (1611):

The maximal density of sphere packings in 3-space is $\frac{\pi}{\sqrt{18}}$

- It corresponds to the way people would intuitively stack oranges, as a pyramid shape
- The proof of T. Hales (1998) consists of thousands of non-linear inequalities
- Many recent efforts have been done to give a formal proof of these inequalities: Flyspeck Project
- Motivation: get positivity certificates and check them with Proof assistants like COQ



Inequalities issued from Flyspeck non-linear part involve:

1. Semi-Algebraic functions algebra $\mathcal{A}$: composition of polynomials with $|\cdot|$, $(\cdot)^{\frac{1}{p}} (p \in \mathbb{N}_0), +, -, \times, /, \sup, \inf$
2. Transcendental functions $\mathcal{T}$: composition of semi-algebraic functions with arctan, arcos, arcsin, exp, log, $|\cdot|$, $(\cdot)^{\frac{1}{p}} (p \in \mathbb{N}_0), +, -, \times, /, \sup, \inf$

Lemma$_{9922699028}$ from Flyspeck

$K := [4, 6.3504]^3 \times [6.3504, 8] \times [4, 6.3504]^2 \qquad P, Q \in \mathbb{R}[X]$

$$\forall x \in K, -\frac{\pi}{2} + \mathit{arctan} \frac{P(x)}{\sqrt{Q(x)}} + 1.6294 - 0.2213 (\sqrt{x_2} + \sqrt{x_3} + \sqrt{x_5} + \sqrt{x_6} - 8.0) + 0.913 (\sqrt{x_4} - 2.52) + 0.728 (\sqrt{x_1} - 2.0) \geq 0.$$

Tight inequality: global optimum $\simeq 1.7 \times 10^{-4}$

Given K a compact set, and f a transcendental function, bound from below $f^* = \inf_{x \in K} f(x)$ and prove $f^* \geq 0$

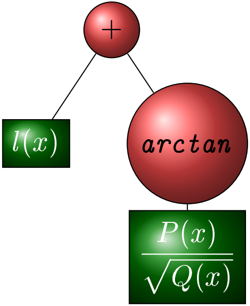
1. f is approximated by a semi-algebraic function f_{sa}
2. Reduce the problem $\inf_{x \in K} f_{sa}(x)$ to a polynomial optimization problem (POP) in a lifted space K_{pop}
3. Solve classically the POP problem $\inf_{x \in K_{pop}} f_{pop}(x)$ using a sparse variant hierarchy of SDP relaxations by Lasserre

$$f^* \geq f^*_{sa} \geq f^*_{pop} \underbrace{\geq 0}_{\text{If the relaxations are accurate enough}}$$

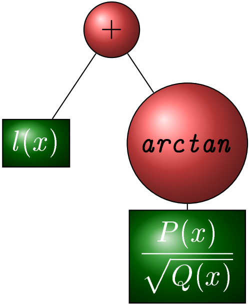
Bounding multivariate transcendental functions

General Framework

- The first step is to build the abstract syntax tree from an inequality, where leaves are semi-algebraic functions and nodes are univariate transcendental functions (arctan, exp, ...) or basic operations ($+$, $\times$, $-$, $/$).
- With $l := -\frac{\pi}{2} + 1.6294 - 0.2213\,(\sqrt{x_2} + \sqrt{x_3} + \sqrt{x_5} + \sqrt{x_6} - 8.0) + 0.913\,(\sqrt{x_4} - 2.52) + 0.728\,(\sqrt{x_1} - 2.0)$, the tree of the example is:



- Let $t \in \mathcal{T}$ be a transcendental univariate elementary function such as $\mathtt{arctan}$, $\mathtt{exp}$, ..., defined on a real interval I. Let $d \in \mathbb{N}$ given.
- Minimax: Best uniform degree d polynomial approximation $\hat{t}$: solution of $||\epsilon||_\infty := \min_{p \in \mathbb{R}_d[X]} ||t - p||_\infty$
- Existence and uniqueness of $\hat{t}$
- Remez algorithm implementation in Sollya: computes $\hat{t}$ for each univariate transcendental function involved in the Flyspeck inequalities with given $I := [a, b]$ and $d \in \mathbb{N}$
- Also computes a certified upper bound of $||\epsilon||_\infty$ related to the minimax polynomial



Two kinds of semialgebraic leaves:

- multivariate functions: $\frac{P(x)}{\sqrt{Q(x)}}$: we can get the bounds by POP using lifting variables
- sum of univariate functions: $l := -\frac{\pi}{2} + 1.6294 - 0.2213\,(\sqrt{x_2} + \sqrt{x_3} + \sqrt{x_5} + \sqrt{x_6} - 8.0) + 0.913\,(\sqrt{x_4} - 2.52) + 0.728\,(\sqrt{x_1} - 2.0)$: we can approximate $\sqrt{\cdot}$ by a minimax polynomial with Sollya

$$\frac{P(x)}{\sqrt{Q(x)}} \text{ on } K := [4, 6.3504]^3 \times [6.3504, 8] \times [4, 6.3504]^2 ?$$

Lifting procedure by POP:

1. Get bounds of $P(x)$ by POP
2. Get bounds of $Q(x)$ by POP and $\sqrt{Q(x)}$ by interval arithmetic
3. Lifting variable representing $\sqrt{Q(x)}$: $q \in I_q$
4. Coarse bounds of $\dfrac{P(x)}{\sqrt{Q(x)}}$ by interval arithmetic: interval I_z
5. Lifting variable representing $\dfrac{P(x)}{\sqrt{Q(x)}}$: $z \in I_z$
6. Lifting space:
   $K_{\text{pop}} := \{(x, q, z) \in K \times I_q \times I_z \, : \, q^2 = Q(x), zq = P(x)\}$
7. Solving $\inf_{(x,q,z) \in K_{\text{pop}}} z$ by POP gives a lower bound of $\dfrac{P(x)}{\sqrt{Q(x)}}$

8. We get an interval I enclosing $\dfrac{P(x)}{\sqrt{Q(x)}}$ from POP.
9. Minimax polynomials for the univariate real functions of f:

t	d	Upper bound of $\|\|\epsilon\|\|_\infty$
$\arctan$ on I	5	2.01×10^{-4}
$\sqrt{}$ on $[4, 6.3504]$	4	2.50×10^{-5}
$\sqrt{}$ on $[6.3504, 8]$	2	9.34×10^{-8}

10. We obtain a minimax polynomial for $\arctan$. With the minimax polynomial for $\sqrt{}$: we can approach l by $\hat{l}$.

Again by using POP:

- Lifting variable representing $\sqrt{Q(x)}$: $q \in I_q$
- Lifting variable representing $\dfrac{P(x)}{\sqrt{Q(x)}}$: $z \in I_z$
- Lifting space:
  $K_{\text{pop}} := \{(x, q, z) \in K \times I_q \times I_z \ : \ q^2 = Q(x), zq = P(x), \}$
- Solving $\inf_{(x,q,z) \in K_{\text{pop}}} \hat{l}(x) + \widehat{\arctan}(z)$ by POP gives a lower bound of $\hat{f} := \hat{l}(x) + \widehat{\arctan}(z)$
- Finally, Subtract the minimax errors $||\epsilon||_\infty$ to $\hat{f}$ gives a lower bound of f

End

Thanks for your attention! Questions?