Certified Global Optimization using Max-Plus based Templates

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

Third year PhD Victor MAGRON

LIX/CMAP INRIA, École Polytechnique

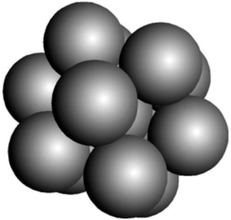
LAAS 2013 Monday May 27 th

The Kepler Conjecture

Kepler Conjecture (1611):

The maximal density of sphere packings in 3D-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



Contents

The Kepler Conjecture

Inequalities issued from Flyspeck non-linear part involve:

1. Multivariate Polynomials:
   $x_1x_4(-x_1+x_2+x_3-x_4+x_5+x_6)+x_2x_5(x_1-x_2+x_3+x_4-x_5+x_6)+x_3x_6(x_1+x_2-x_3+x_4+x_5-x_6)-x_2(x_3x_4+x_1x_6)-x_5(x_1x_3+x_4x_6)$
2. Semi-Algebraic functions algebra $\mathcal{A}$: composition of polynomials with $|\cdot|, \sqrt{}, +, -, \times, /, \sup, \inf, \cdots$
3. Transcendental functions $\mathcal{T}$: composition of semi-algebraic functions with $\arctan, \exp, \sin, +, -, \times, \cdots$

Lemma from Flyspeck (inequality ID 6096597438)

$$\forall x \in [3, 64], 2\pi - 2x \arcsin(\cos(0.797) \sin(\pi/x)) + 0.0331x - 2.097 \geq 0$$

Global Optimization Problems: Examples from the Literature

- *H3*: $$\min_{\mathbf{x}\in[0,1]^3} -\sum_{i=1}^{4} c_i \exp\left[-\sum_{j=1}^{3} a_{ij}(x_j - p_{ij})^2\right]$$
- *MC*: $$\min_{\substack{x_1\in[-3,3] \\ x_2\in[-1.5,4]}} \sin(x_1 + x_2) + (x_1 - x_2)^2 - 0.5x_2 + 2.5x_1 + 1$$
- *SBT*: $$\min_{\mathbf{x}\in[-10,10]^n} \prod_{i=1}^{n} \left(\sum_{j=1}^{5} j\cos((j+1)x_i + j)\right)$$
- *SWF*: $$\min_{\mathbf{x}\in[1,500]^n} -\sum_{i=1}^{n} (x_i + \epsilon x_{i+1})\sin(\sqrt{x_i}) \quad (\epsilon \in \{0,1\})$$

Global Optimization Problems: a Framework

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

1. f is underestimated by a semialgebraic function f_{sa}
2. We reduce the problem $f_{\text{sa}}^* := \inf_{\mathbf{x} \in K} f_{\text{sa}}(\mathbf{x})$ to a polynomial optimization problem in a lifted space K_{pop} (with lifting variables $\mathbf{z}$)
3. We solve the POP problem $f_{\text{pop}}^* := \inf_{(\mathbf{x},\mathbf{z}) \in K_{\text{pop}}} f_{\text{pop}}(\mathbf{x}, \mathbf{z})$ using a hierarchy of SDP relaxations by Lasserre

If the relaxations are accurate enough, $f^* \geq f_{\text{sa}}^* \geq f_{\text{pop}}^* \geq 0$.

Contents

Semialgebraic Optimization Problems

- Polynomial Optimization Problem (POP):
  $p^* := \min_{\mathbf{x} \in K} p(\mathbf{x})$ with K the compact set of constraints:

$$K = \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \geq 0, \cdots, g_m(\mathbf{x}) \geq 0\}$$

- Let $\Sigma[\mathbf{x}]$ be the cone of Sum-of-Squares (SOS) and consider the restriction $\Sigma_d[\mathbf{x}]$ to polynomials of degree at most $2d$:

$$\Sigma_d[\mathbf{x}] = \left\{ \sum_i q_i(\mathbf{x})^2, \text{ with } q_i \in \mathbb{R}_d[\mathbf{x}] \right\}$$

- Let $g_0 := 1$ and $M(\mathbf{g})$ be the quadratic module generated by $g_1, \cdots, g_m$:

$$M(\mathbf{g}) = \left\{ \sum_{j=0}^{m} \sigma_j(\mathbf{x}) g_j(\mathbf{x}), \text{ with } \sigma_j \in \Sigma[\mathbf{x}] \right\}$$

- Certificates for positive polynomials: Sum-of-Squares

Semialgebraic Optimization Problems

Proposition (Putinar)

Suppose $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$. $p(\mathbf{x}) - p^* > 0$ on $K \implies (p(\mathbf{x}) - p^*) \in M(\mathbf{g})$

- But the search space for $\sigma_0, \cdots, \sigma_m$ is infinite so consider the truncated module $M_d(\mathbf{g})$:

$$M_d(\mathbf{g}) = \left\{ \sum_{j=0}^{m} \sigma_j(\mathbf{x}) g_j(\mathbf{x}), \text{ with } \sigma_j \in \Sigma[\mathbf{x}], (\sigma_j g_j) \in \mathbb{R}_{2d}[\mathbf{x}] \right\}$$

- $M_0(\mathbf{g}) \subset M_1(\mathbf{g}) \subset M_2(\mathbf{g}) \subset \cdots \subset M(\mathbf{g})$
- Hence, we consider the hierarchy of SOS *relaxation* programs: $\mu_k := \sup_{\mu, \sigma_0, \cdots, \sigma_m} \left\{ \mu : (p(\mathbf{x}) - \mu) \in M_k(\mathbf{g}) \right\}$

Classical Approach: Taylor + SOS

Lasserre Hierarchy Convergence:

- Let $k \geq k_0 := \max\{\lceil \deg f/2 \rceil, \lceil \deg g_1/2 \rceil, \cdots, \lceil \deg g_m/2 \rceil\}$.
- The sequence $\inf(\mu_k)_{k \geq k_0}$ is non-decreasing. Under a certain (moderate) assumption, it converges to p^*.

- $\min_{\mathbf{x} \in [4,6.3504]^6} \Delta(\mathbf{x}) = x_1 x_4(-x_1 + x_2 + x_3 - x_4 + x_5 + x_6) + x_2 x_5(x_1 - x_2 + x_3 + x_4 - x_5 + x_6) + x_3 x_6(x_1 + x_2 - x_3 + x_4 + x_5 - x_6) - x_2(x_3 x_4 + x_1 x_6) - x_5(x_1 x_3 + x_4 x_6) = \mu_2 = 128$
- $\Delta(\mathbf{x}) - \mu_2 = \sigma_0(\mathbf{x}) + \sum_{j=1}^{6} \sigma_j(\mathbf{x})(6.3504 - x_j)(x_j - 4)$ with $\sigma_0 \in \Sigma_2[\mathbf{x}], \sigma_j \in \Sigma_1[\mathbf{x}]$
- Also works for Semialgebraic functions via *lifting* variables:

$$f_{sa}^* := \min_{\mathbf{x} \in [4,6.3504]^6} f_{sa}(\mathbf{x}) = \frac{\partial_4 \Delta \mathbf{x}}{\sqrt{4 x_1 \Delta \mathbf{x}}}$$

Semialgebraic Optimization Problems: examples

b.s.a.l. lemma [Lasserre, Putinar] :

Let $\mathcal{A}$ be the semi-algebraic functions algebra obtained by composition of polynomials with $|\cdot|$, $(\cdot)^{\frac{1}{p}} (p \in \mathbb{N}_0), +, -, \times, /, \sup, \inf$. Then every well-defined $f_{sa} \in \mathcal{A}$ has a basic semi-algebraic lifting.

Example from Flyspeck:

$$z_1 := \sqrt{4x_1 \Delta \mathbf{x}},\ m_1 = \inf_{\mathbf{x} \in [4, 6.3504]^6} z_1(\mathbf{x}),\ M_1 = \sup_{\mathbf{x} \in [4, 6.3504]^6} z_1(\mathbf{x}).$$

$$K := \{(\mathbf{x}, \mathbf{z}) \in \mathbb{R}^8 : \mathbf{x} \in [4, 6.3504]^6, h_1(\mathbf{x}, \mathbf{z}) \geq 0, \cdots, h_7(\mathbf{x}, \mathbf{z}) \geq 0\}$$

$$\begin{aligned} h_1 &:= z_1 - m_1 & h_4 &:= -z_1^2 + 4x_1 \Delta \mathbf{x} & h_7 &:= -z_2 z_1 + \partial_4 \Delta \mathbf{x} \\ h_2 &:= M_1 - z_1 & h_5 &:= z_1 & & \\ h_3 &:= z_1^2 - 4x_1 \Delta \mathbf{x} & h_6 &:= z_2 z_1 - \partial_4 \Delta \mathbf{x} & & \end{aligned}$$

$p^* := \inf_{(\mathbf{x}, \mathbf{z}) \in K} z_2 = f_{\text{sa}}^*$. We obtain $\mu_2 = -0.618$ and $\mu_3 = -0.445$.

Taylor Approximation of Transcendental Functions

SWF: $\min_{\mathbf{x} \in [1,500]^n} f(\mathbf{x}) = -\sum_{i=1}^{n} (x_i + x_{i+1}) \sin(\sqrt{x_i})$

Classical idea: approximate $\sin(\sqrt{\cdot})$ by a degree-d Taylor

Polynomial f_d, solve $\min_{\mathbf{x} \in [1,500]^n} -\sum_{i=1}^{n} (x_i + x_{i+1}) f_d(x_i)$ (POP)

Issues:

- Lack of accuracy if d is not large enough $\implies$ expensive Branch and Bound
- POP may involve many lifting variables :depends on semialgebraic and univariate transcendental components of f
- No free lunch: solving POP with Sum-of-Squares of degree $2k$ involves $O(n^{2k})$ variables

SWF with $n = 10, d = 4$: takes already $38\, min$ to certify a lower bound of $-430n$

Contents

Max-Plus Estimators

Goals:

- Reduce the $O(n^{2k})$ polynomial dependency: decrease the number of lifting variables
- Reduce the $O(n^{2k})$ exponential dependency: use low degree approximations
- Reduce the Branch and Bound iterations: refine the approximations with an adaptive iterative algorithm

Max-Plus Estimators

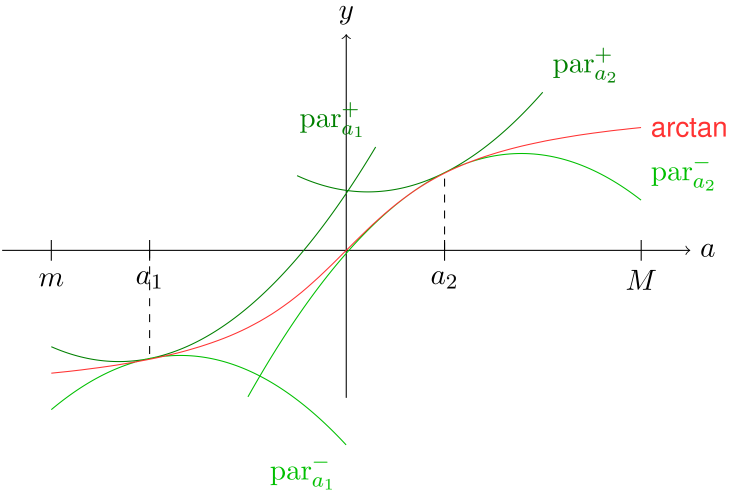
- Let $\hat{f} \in \mathcal{T}$ be a transcendental univariate elementary function such as $\arctan$, $\exp$ defined on a real interval I.
- Convexity/semi-convexity properties and monotonicity of $\hat{f}$
- $\hat{f}$ is semi-convex: there exists a constant $c_j > 0$ s.t. $a \mapsto \hat{f}(a) + c_j/2(a - a_j)^2$ is convex
- By convexity: $\forall a \in I, \hat{f}(a) \geq -c_j/2(a - a_j)^2 + \hat{f}'(a_j)(a - a_j) + \hat{f}(a_j) = \mathrm{par}^-_{a_j}(a)$
- $\forall j, \hat{f} \geq \mathrm{par}^-_{a_j} \implies \hat{f} \geq \max_j \{\mathrm{par}^-_{a_j}\}$ **Max-Plus underestimator**

Example with $\arctan$:

- $\hat{f}'(a_j) = \dfrac{1}{1 + a_j^2}, \quad c_j = \sup_{a \in I}\{-\hat{f}''(a)\}$ (always work)
- c_j depends on a_j and the curvature variations of $\arctan$ on the considered interval I

Max-Plus Estimators

Example with arctan:



Max-Plus Estimators

- $l := -\frac{\pi}{2} + 1.6294 - 0.2213\,(\sqrt{x_2} + \sqrt{x_3} + \sqrt{x_5} + \sqrt{x_6} - 8.0) + 0.913\,(\sqrt{x_4} - 2.52) + 0.728\,(\sqrt{x_1} - 2.0)$

Lemma$_{9922699028}$ from Flyspeck:

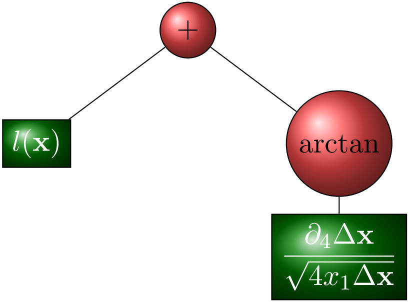
$$\forall \mathbf{x} \in [4, 6.3504]^6, \arctan\left(\frac{\partial_4 \Delta \mathbf{x}}{\sqrt{4x_1 \Delta \mathbf{x}}}\right) + l(\mathbf{x}) \geq 0$$

- Using semialgebraic optimization methods:
  $\forall x \in [4, 6.3504]^6,\ m \leq \frac{\partial_4 \Delta \mathbf{x}}{\sqrt{4x_1 \Delta \mathbf{x}}} \leq M$
- Using the $\arctan$ properties with two points $a_1, a_2 \in [m, M]$:
  $\forall \mathbf{x} \in [4, 6.3504]^6, \arctan\left(\frac{\partial_4 \Delta \mathbf{x}}{\sqrt{4x_1 \Delta \mathbf{x}}}\right) \geq \max_{j \in \{1,2\}} \{\mathrm{par}^-_{a_j}\left(\frac{\partial_4 \Delta \mathbf{x}}{\sqrt{4x_1 \Delta \mathbf{x}}}\right)\}$

Semialgebraic Max-Plus Algorithm

Abstract syntax tree representations of multivariate transcendental function:

- leaves are semialgebraic functions of $\mathcal{A}$
- nodes are univariate transcendental functions of $\mathcal{T}$ or binary operations



Semialgebraic Max-Plus Algorithm

Recursive Algorithm `samp_approx`:

Input: tree t, box K, SDP relaxation order k, control points sequence $s = \mathbf{x}^1, \ldots, \mathbf{x}^p \in K$

Output: lower bound m, upper bound M, lower tree t^-, upper tree t^+

if $t \in \mathcal{A}$ **then**
  $t^- := t$, $t^+ := t$
else if $r := \mathtt{root}(t) \in \mathcal{T}$ parent of the single child c **then**
  $m_c, M_c, c^-, c^+ := \mathtt{samp\_approx}(c, K, k, s)$
  $\mathrm{par}^-, \mathrm{par}^+ := \mathtt{build\_par}(r, m_c, M_c, s)$
  $t^-, t^+ := \mathtt{compose}(\mathrm{par}^-, \mathrm{par}^+, c^-, c^+)$
else if $\mathtt{bop} := \mathtt{root}\ (t)$ is a binary operation parent of two children c_1 and c_2 **then**
  $m_{c_i}, M_{c_i}, c_i^-, c_i^+ := \mathtt{samp\_approx}(c_i, K, k, s)$ for $i \in \{1, 2\}$
  $t^-, t^+ := \mathtt{compose\_bop}(c_1^-, c_1^+, c_2^-, c_2^+)$
end
return $\min(t^-, k), \max(t^+, k), t^-, t^+$

Semialgebraic Max-Plus Algorithm

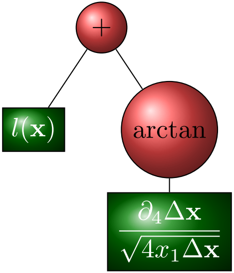
Iterative Algorithm `samp_optim`:

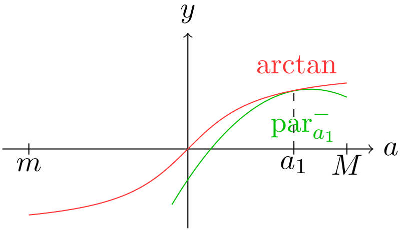
Input: tree t, box K, $iter_{\max}$ (optional argument)
Output: lower bound m, feasible solution $\mathbf{x}_{opt}$

$s := [\, \texttt{argmin}\,(\texttt{randeval}\ t)\,]$ ▷ $s \in K$
$iter := 0$
$m := -\infty$
while $iter \leq iter_{\max}$ **do**
    Choose an SDP relaxation order k
    $m,\ M,\ t^-,\ t^+ := \texttt{samp\_approx}\,(t,\ K,\ k,\ s)$
    $\mathbf{x}_{opt} := \texttt{guess\_argmin}\,(t^-)$ ▷ $t^-(\mathbf{x}_{opt}) \simeq m$
    $s := s \cup \{\mathbf{x}_{opt}\}$
    $iter := iter + 1$
done
return $m,\ \mathbf{x}_{opt}$

Semialgebraic Max-Plus Algorithm

`samp_optim` First iteration:



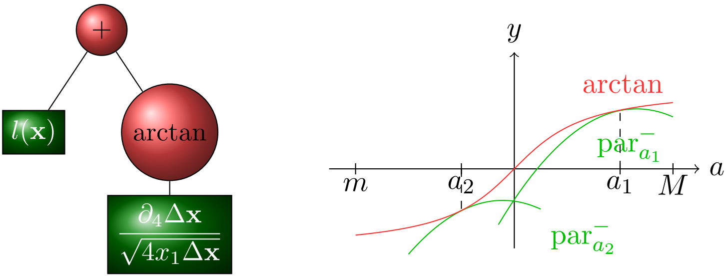


1. Evaluate f with `randeval` and obtain a minimizer guess $\mathbf{x}^1_{opt}$.

   Compute $a_1 := \frac{\partial_4 \Delta \mathbf{x}}{\sqrt{4 x_1 \Delta \mathbf{x}}} (\mathbf{x}^1_{opt}) = f_{\mathsf{sa}}(\mathbf{x}^1_{opt}) = 0.84460$
2. Get the equation of $\mathrm{par}^-_{a_1}$ with $\mathtt{build}_{\mathtt{par}}$
3. Compute $m_1 \leq \min_{\mathbf{x} \in [4, 6.3504]} (l(\mathbf{x}) + \mathrm{par}^-_{a_1}(f_{\mathsf{sa}}(\mathbf{x})))$

Semialgebraic Max-Plus Algorithm

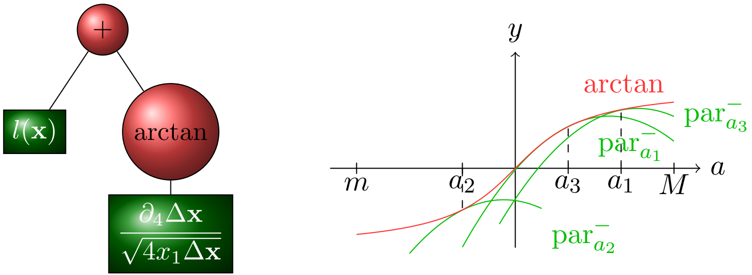
samp_optim Second iteration:



1. For $k = 3, m_1 = -0.746 < 0$, obtain a new minimizer $\mathbf{x}^2_{opt}$.
2. Compute $a_2 := f_{\text{sa}}(\mathbf{x}^2_{opt}) = -0.374$ and $\text{par}^-_{a_2}$
3. Compute $m_2 \leq \min_{\mathbf{x} \in [4, 6.3504]} (l(\mathbf{x}) + \max_{i \in \{1,2\}} \{\text{par}^-_{a_i}(f_{\text{sa}}(\mathbf{x}))\})$

Semialgebraic Max-Plus Algorithm

samp_optim Third iteration:



1. For $k = 3, m_2 = -0.112 < 0$, obtain a new minimizer $\mathbf{x}_{opt}^3$.
2. Compute $a_3 := f_{\mathsf{sa}}(\mathbf{x}_{opt}^3) = 0.357$ and $\mathrm{par}_{a_3}^-$
3. Compute $m_3 \leq \min_{\mathbf{x} \in [4, 6.3504]} (l(\mathbf{x}) + \max_{i \in \{1,2,3\}} \{\mathrm{par}_{a_i}^-(f_{\mathsf{sa}}(\mathbf{x}))\})$

Semialgebraic Max-Plus Algorithm

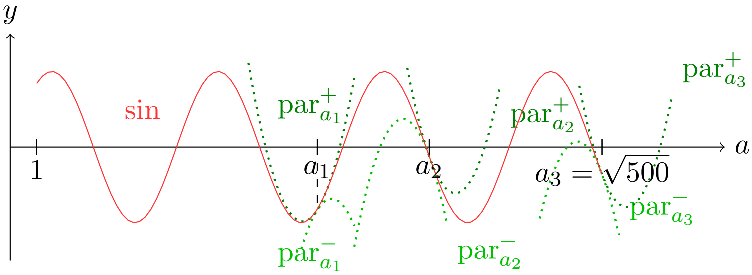
- For $k = 3, m_3 = -0.0333 < 0$, obtain a new minimizer $\mathbf{x}_{opt}^4$ and iterate again...

Theorem: Convergence of Semialgebraic underestimators

Let f be a multivariate transcendental function that can be represented by such syntaxic abstract trees.
Let $(\mathbf{x}_{opt}^p)_{p\in\mathbb{N}}$ be a sequence of control points obtained to define the hierarchy of underestimators in the algorithm `samp_optim` and $\mathbf{x}^*$ be an accumulation point of $(\mathbf{x}_{opt}^p)_{p\in\mathbb{N}}$.
Then, $\mathbf{x}^*$ is a global minimizer of f on K.

Max-Plus Based Templates Approach

Example with $\sin$:



Semialgebraic Max-Plus Algorithm

SWF: $\min_{\mathbf{x} \in [1,500]^n} - \sum_{i=1}^{n} (x_i + x_{i+1}) \sin(\sqrt{x_i}) \quad (\epsilon = 1)$

- Use one lifting variable y_i to represent $x_i \mapsto \sqrt{x_i}$ and one lifting variable z_i to represent $x_i \mapsto \sin(x_i)$

$$\begin{cases} \min_{\mathbf{x} \in [1,500]^n, \mathbf{y} \in [1,\sqrt{500}]^n, \mathbf{z} \in [-1,1]^n} & - \sum_{i=1}^{n} (x_i + x_{i+1}) z_i \\ \text{s.t.} & z_i \leq \text{par}^{+}_{a_{ji}}(y_i), j \in \{1,2,3\} \\ & y_i^2 = x_i \end{cases}$$

- POP with $n + 2n$ variables ($n_{\text{lifting}} = 2n$ variables), with Sum-of-Squares of degree $2d$: $O((3n)^{2d})$ complexity

Templates

With `samp_optim`: the number of lifting variables is not bounded
Remedy: select some subcomponents of f and compute estimators involving less lifting variables

- Let t be such a subcomponent and $\mathbf{x}^j$ be a control point and suppose that t is twice differentiable.
- Define the interval matrix $\tilde{D}$ enclosing all the entries of $(\mathcal{D}^2(t)(\mathbf{x}) - \mathcal{D}^2(t)(\mathbf{x}^j))$ for $\mathbf{x} \in K$
- Define the quadratic form $q_{\mathbf{x}^j,\lambda} : \mathbf{x} \mapsto t(\mathbf{x}^j) + \mathcal{D}(t)(\mathbf{x}^j)\,(\mathbf{x} - \mathbf{x}^j) + \frac{1}{2}(\mathbf{x} - \mathbf{x}^j)^T \mathcal{D}^2(t)(\mathbf{x}^j)\,(\mathbf{x} - \mathbf{x}^j) + \frac{1}{2}\lambda \parallel \mathbf{x} - \mathbf{x}^j \parallel_2^2$

Templates

- Lower bound of $\min_{\mathbf{x} \in K} \{\lambda_{\min}(\mathcal{D}^2(t)(\mathbf{x}) - \mathcal{D}^2(t)(\mathbf{x}_j))\}$: $\lambda_{\min}(\tilde{D})$

  $\lambda^- := \lambda_{\min}(\tilde{D})$: minimal eigenvalue of an interval matrix
- For each interval $\tilde{D}_{ij} = [m_{ij}, M_{ij}]$, define the symmetric matrix entry $B_{ij} := \max\{\mid m_{ij} \mid, \mid M_{ij} \mid\}$
- Let $\mathcal{S}^n$ be the set of diagonal matrices of sign.

  $\mathcal{S}^n := \{\text{diag}\,(s_1, \cdots, s_n),\ s_1 = \pm 1, \cdots s_n = \pm 1\}$

Robut Optimization with Reduced Vertex Set [Calafiore, Dabbene]

The robust interval SDP problem $\lambda_{\min}(\tilde{D})$ is equivalent to the following Semidefinite Program (SDP) in the single variable $t \in \mathbb{R}$:

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

Templates

Previous Algorithm:

Input: tree t, box K, SDP relaxation order k, control points sequence $s = \mathbf{x}^1, \ldots, \mathbf{x}^p \in K$

Output: lower bound m, upper bound M, lower tree t^-, upper tree t^+

if $t \in \mathcal{A}$ **then**
  $t^- := t, t^+ := t$
else if $r := \texttt{root}(t) \in \mathcal{T}$ parent of the single child c **then**
  $m_c, M_c, c^-, c^+ := \texttt{samp\_approx}(c, K, k, s)$
  $\text{par}^-, \text{par}^+ := \texttt{build\_par}(r, m_c, M_c, s)$
  $t^-, t^+ := \texttt{compose}(\text{par}^-, \text{par}^+, c^-, c^+)$
else if $\texttt{bop} := \texttt{root}\ (t)$ is a binary operation parent of two children c_1 and c_2 **then**
  $m_{c_i}, M_{c_i}, c_i^-, c_i^+ := \texttt{samp\_approx}(c_i, K, k, s)$ for $i \in \{1, 2\}$
  $t^-, t^+ := \texttt{compose\_bop}(c_1^-, c_1^+, c_2^-, c_2^+)$
end
return $\min(t^-, k), \max(t^+, k), t^-, t^+$

Templates

Input: tree t, box K, SDP relaxation order k, control points sequence $s = \mathbf{x}^1, \ldots, \mathbf{x}^p \in K$

Output: lower bound m, upper bound M, lower tree t^-, upper tree t^+

if $t \in \mathcal{A}$ **then**
    $t^-, t^+ := t, t$
else if $r := \texttt{root}(t) \in \mathcal{T}$ parent of the single child c **then**
    $m_c, M_c, c^-, c^+ := \texttt{template\_optim}(c, K, k, s)$
    $\text{par}^-, \text{par}^+ := \texttt{build\_par}(r, m_c, M_c, s)$
    $t^-, t^+ := \texttt{compose}(\text{par}^-, \text{par}^+, c^-, c^+)$
else if $\texttt{bop} := \texttt{root}\ (t)$ is a binary operation parent of two children c_1 and c_2 **then**
    $m_{c_i}, M_{c_i}, c_i^-, c_i^+ := \texttt{template\_optim}(c_i, K, k, s)$ for $i \in \{1, 2\}$
    $t^-, t^+ := \texttt{compose\_bop}(c_1^-, c_1^+, c_2^-, c_2^+)$
end
$t_2^-, t_2^+ := \texttt{build\_template}(t, K, k, s, t^-, t^+)$
return $\texttt{min\_sa}(t_2^-, k), \texttt{max\_sa}(t_2^+, k), t_2^-, t_2^+$

Templates

$\texttt{build}_{\texttt{template}}$ builds quadratic forms by solving SDP problems.

Input: tree t, box K, SDP relaxation order k, control points sequence $s = \mathbf{x}^1, \ldots, \mathbf{x}^p \in K$, lower/upper semialgebraic estimator t^-, t^+

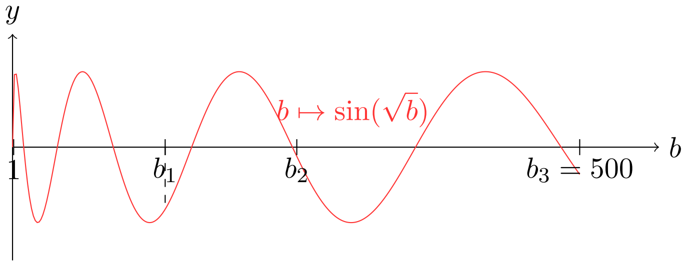
if the number of lifting variables exceeds $n_{\text{lifting}}^{\max}$ **then**
    for $\mathbf{x}^j \in s$ **do**
        Compute the interval matrix $\tilde{D}^j$
        $\lambda^- := \lambda_{\min}(\tilde{D}^j) \quad q_j^- := q_{\mathbf{x}^j, \lambda^-}$
        $\lambda^+ := \lambda_{\max}(\tilde{D}^j) \quad q_j^+ := q_{\mathbf{x}^j, \lambda^+}$
    done
    return $\max_{1 \leq j \leq p} \{q_j^-\}$, $\min_{1 \leq j \leq p} \{q_j^+\}$
else
    return t^-, t^+
end

Templates

When t is univariate, $\lambda^- = -c_j$ (the semi-convexity constant)

SWF: $$\min_{\mathbf{x} \in [1,500]^n} - \sum_{i=1}^{n} (x_i + x_{i+1}) \sin(\sqrt{x_i})$$

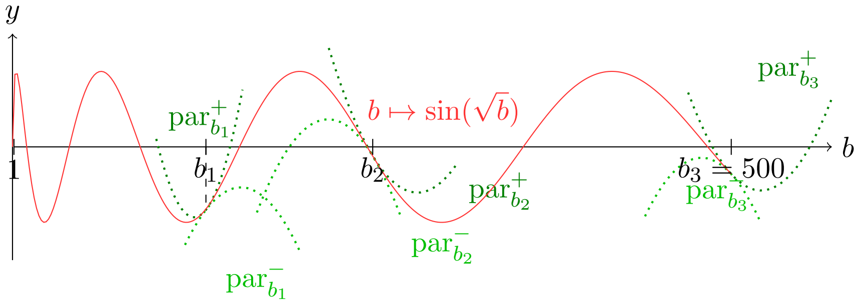
- Consider the univariate function $b \mapsto \sin(\sqrt{b})$ on $I = [1, 500]$



- $\forall b \in I, \hat{f}(b) \geq -c_j/2(b - b_j)^2 + \hat{f}'(b_j)(b - b_j) + \hat{f}(b_j) = \text{par}^-_{b_j}(b)$

Templates

- $\forall j, \hat{f} \geq \text{par}^-_{b_j} \implies \hat{f} \geq \max_j \left\{ \text{par}^-_{b_j} \right\}$: **Max-Plus underestimator**
- $\forall j, \hat{f} \leq \text{par}^+_{b_j} \implies \hat{f} \leq \min_j \left\{ \text{par}^+_{b_j} \right\}$: **Max-Plus overestimator**



Templates based on Max-plus Estimators for $b \mapsto \sin(\sqrt{b})$:

$$\max_{j \in \{1,2,3\}} \{\text{par}^-_{b_j}(x_i)\} \leq \sin\sqrt{x_i} \leq \min_{j \in \{1,2,3\}} \{\text{par}^+_{b_j}(x_i)\}$$

Templates

- Use a lifting variable z_i to represent $x_i \mapsto \sin(\sqrt{x_i})$
- For each i, pick points b_{ji}
- With 3 points b_{ji}, we solve the POP:

$$\begin{cases} \min\limits_{\mathbf{x}\in[1,500]^n, \mathbf{z}\in[-1,1]^n} & -\sum\limits_{i=1}^{n}(x_i + x_{i+1})z_i \\ \text{s.t.} & z_i \leq \text{par}^+_{b_{ji}}(x_i), j \in \{1, 2, 3\} \end{cases}$$

- POP with $n + n$ variables ($n_{\text{lifting}} = n$ variables), with Sum-of-Squares of degree $2d$: $O((2n)^{2d})$ complexity
- Taylor approximations: templates with n variables ($n_{\text{lifting}} = 0$ variables)

Benchmarks

$$\min_{\mathbf{x} \in [1,500]^n} f(\mathbf{x}) = -\sum_{i=1}^{n} (x_i + \epsilon x_{i+1}) \sin(\sqrt{x_i})$$

n	lower bound	n_{lifting}	#boxes	time
$10 (\epsilon = 0)$	$-430n$	$2n$	16	$40\ s$
$10 (\epsilon = 0)$	$-430n$	0	827	$177\ s$
$1000 (\epsilon = 1)$	$-967n$	$2n$	1	$543\ s$
$1000 (\epsilon = 1)$	$-968n$	n	1	$272\ s$

Benchmarks

- SOS of degree $2k$, $\#s$: `template_optim` iterations
- when $\#s = 0$, $n_{\text{lifting}} = 0$: interval arithmetic + SOS

Problem	n	lower bound	k	$\#s$	n_{lifting}	#boxes	time
H3	3	-3.863	2	3	4	99	$101\,s$
				0	0	1096	$247\,s$
MC	2	-1.92	1	2	1	17	$1.8\,s$
				0	0	92	$7.6\,s$
ML	10	-0.966	1	1	6	8	$8.2\,s$
				0	0	8	$6.6\,s$
PP	10	-46	1	3	2	135	$89\,s$
				0	0	3133	$115\,s$

Benchmarks

- $n = 6$ variables, SOS of degree $2k = 4$
- $n_{\mathcal{T}}$ univariate transcendental functions, #boxes sub-problems

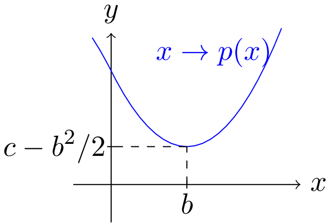
Inequality id	$n_{\mathcal{T}}$	n_{lifting}	#boxes	time
9922699028	1	9	47	$241\,s$
9922699028	1	3	39	$190\,s$
3318775219	1	9	338	$26\,min$
7726998381	3	15	70	$43\,min$
7394240696	3	15	351	$1.8\,h$
4652969746_1	6	15	81	$1.3\,h$
OXLZLEZ 6346351218_2_0	6	24	200	$5.7\,h$

Contents

Certification Framework: who does what?

Polynomial Optimization (POP): $\min_{x \in \mathbb{R}} p(x) = 1/2x^2 - bx + c$

1. A program written in OCaml/C provides the SOS decomposition: $1/2(x-b)^2$
2. A program written in Coq checks: $\forall x \in \mathbb{R}, p(x) = 1/2(x-b)^2 + c - b^2/2$



- Sceptical approach: obtain *certificates* of positivity with efficient oracles and check them formally

Coq tactics: `field`, `interval`

Formal proofs for lower bounds of POP:

- The oracle returns floating point certificate: $\mu, \sigma_0, \cdots, \sigma_m$
- Check equality of polynomials: $f(\mathbf{x}) - \mu = \sum_{i=0}^{m} \sigma_i(\mathbf{x}) g_i(\mathbf{x})$

  with the Coq `field` tactic.

Coq tactics: `field`, `interval`

- The equality test often fails. Two workarounds:
  1. Rounding and Projection of the certificate (Peyrl and Parillo, Kaltofen) until we get the equality
  2. Bound $f(\mathbf{x}) - \mu - \sum_{i=0}^{m} \sigma_i(\mathbf{x}) g_i(\mathbf{x}) = \sum_{\alpha \in \mathcal{F}} \epsilon_\alpha \mathbf{x}^\alpha$ on K with the Coq `interval` tactic
     $\mathcal{F}$ is support of the SOS certificate, hence can be reduced by exploiting the system properties. Hope that ϵ_α is not too large!
- In both cases, the initial lower bound is decreased to achieve the formal certification.

Formal proofs for Max-Plus estimators: certify rigorous estimators for univariate transcendental functions

Polynomial Underestimators of Semialgebraic functions using SDP

- Let t be a semialgebraic leaf of the abstract syntaxic tree of f
- Let $\mathbf{x}^j \in K$ a control point
- Let λ denote the Lebesgue measure distributed on K

Consider the following optimization problem with optimal solution h_d^*:

$$\begin{cases} \min_{h \in \mathbb{R}_d[\mathbf{x}]} & \int_K (t - h) d\lambda \\ \text{s.t.} & t - h \geq 0 \text{ on } K \\ & h(\mathbf{x}^j) = t(\mathbf{x}^j), \quad h'(\mathbf{x}^j) = t'(\mathbf{x}^j) \end{cases}$$

Idea: provide a sequence of degree-d polynomial underestimators $(h_{dk}) \subset \mathbb{R}_d[\mathbf{x}], k \in \mathbb{N}$ such that $\| t - h_{dk} \|_1 \to \| t - h_d^* \|_1$ for the L_1 norm on K

Polynomial Underestimators of Semialgebraic functions using SDP

- There exist lifting variables $z_1, \cdots, z_p$ and polynomials $g_j \in \mathbb{R}[\mathbf{x}, \mathbf{z}], j = 1, \cdots, m$ defining the semialgebraic set: $K_{\text{pop}} := \{(\mathbf{x}, \mathbf{z}) \in \mathbb{R}^{n+p} : \mathbf{x} \in K, g_1(\mathbf{x}, \mathbf{z}) \geq 0, \cdots, g_m(\mathbf{x}, \mathbf{z}) \geq 0\}$ such that $\Psi_t := \{(\mathbf{x}, t(\mathbf{x})) : \mathbf{x} \in K\} = \{(\mathbf{x}, z_p) : (\mathbf{x}, \mathbf{z}) \in K_{\text{pop}}\}$
- Then we can rewrite the previous optimization problem:

$$\begin{cases} \min\limits_{h \in \mathbb{R}_d[\mathbf{x}]} & \displaystyle\int_{K_{\text{pop}}} (z_p - h) d\lambda \\ \text{s.t.} & z_p - h(\mathbf{x}) \geq 0 \text{ on } K_{\text{pop}} \\ & h(\mathbf{x}^j) = t(\mathbf{x}^j), \quad h'(\mathbf{x}^j) = t'(\mathbf{x}^j) \end{cases}$$

Polynomial Underestimators of Semialgebraic functions using SDP

- $K := \{\mathbf{x} \in \mathbb{R}^n : f_1(\mathbf{x}) \geq 0, \cdots, f_{2n}(\mathbf{x}) \geq 0\}$
- Let $g_0 := 1$ and $\omega_0 := \deg(g_0), \cdots, \omega_m := \deg(g_m)$
- For $k \geq k_0 = \max\{\lceil d/2 \rceil, \lceil \omega_1/2 \rceil, \cdots, \lceil \omega_m/2 \rceil\}$, introduce the following SDP relaxation F_{dk}:

$$\begin{cases} \max\limits_{h \in \mathbb{R}_d[\mathbf{x}], \sigma, \phi} & \displaystyle\int_K h d\lambda \\ \text{s.t.} & \forall \mathbf{x}, \mathbf{z}, z_p - h(\mathbf{x}) = \displaystyle\sum_{i=0}^{m} \sigma_i(\mathbf{x}, \mathbf{z}) g_i(\mathbf{x}) + \sum_{i=1}^{2n} \phi_i(\mathbf{x}, \mathbf{z}) f_i(\mathbf{x}) \\ & h(\mathbf{x}^j) = t(\mathbf{x}^j), \quad h'(\mathbf{x}^j) = t'(\mathbf{x}^j) \\ & \sigma_i \in \Sigma_{k-\omega_i}[\mathbf{x}, \mathbf{z}], \quad \phi_i \in \Sigma_{k-1}[\mathbf{x}, \mathbf{z}] \end{cases}$$

The optimal solutions h_{dk} of F_{dk} satisfy $\| t - h_{dk} \|_1 \to \| t - h_d^* \|_1$ for the L_1 norm on K

Exploiting System Properties

- Templates preserve system properties: Sparsity / Symmetries
- Implementation in OCaml of the sparse variant of SDP relaxations (Kojima) for POP and semialgebraic underestimators
- Reducing the size of SDP input data has a positive domino effect:
  1. on the global optimization oracle to decrease the $O(n^{2d})$ complexity
  2. to check SOS with `field` and `interval` Coq tactics

Xavier Allamigeon, Stéphane Gaubert, Victor Magron, and Benjamin Werner.
Certification of bounds of non-linear functions : the templates method, 2013.
To appear in the Proceedings of Conferences on Intelligent Computer Mathematics, CICM 2013 Calculemus, Bath.