The generating function approach to mixed-integer nonlinear optimization

Matthias Köppe

UC Davis, Mathematics

based on joint work with
V. Baldoni, N. Berline, B. Dutra, J. A. De Loera, M. Vergne, Y. Zhou

(and earlier joint work with J. A. De Loera, R. Hemmecke, R. Weismantel)
Maximization of a non-negative function $f$ over a feasible region $F$ is “just” the limit case of power-$p$ integration (summation):

$$\lim_{p \to \infty} \|f\|_{p,F} = \|f\|_{\infty,F}$$

With
- $f \in \mathbb{Q}[x_1, \ldots, x_d]$ a non-negative polynomial, $g := f^p$,
- $F = P$ a polytope or $F = P \cap \mathbb{Z}^d$ or $F = P \cap (\mathbb{Z}^{d_1} \times \mathbb{R}^{d_2})$

this leads to studying:

**Problem of exact integration**

Given a polytope $P \subseteq \mathbb{R}^n$ and a polynomial $g \in \mathbb{Q}[x_1, \ldots, x_n]$, compute the integral

$$\int_P g(x) \, dx.$$  

**Problem of exact summation**

Given a polytope $P \subseteq \mathbb{R}^d$ and a polynomial $g \in \mathbb{Q}[x_1, \ldots, x_d]$, compute the discrete sum

$$\sum_{x \in P \cap \mathbb{Z}^d} g(x)$$
Summation of polynomial densities with the Euler differential operator

\[ S(z) = z^0 + z^1 + z^2 + z^3 + z^4 = \frac{1}{1 - z} - \frac{z^5}{1 - z} \]

Apply differential operator:

\[ \left( z \frac{d}{dz} \right) S(z) = 1z^1 + 2z^2 + 3z^3 + 4z^4 = \frac{1}{(1 - z)^2} - \frac{-4z^5 + 5z^4}{(1 - z)^2} \]

Apply differential operator again:

\[ \left( z \frac{d}{dz} \right) \left( z \frac{d}{dz} \right) S(z) = 1z^1 + 4z^2 + 9z^3 + 16z^4 = \frac{z + z^2}{(1 - z)^3} - \frac{25z^5 - 39z^6 + 16z^7}{(1 - z)^3} \]
Generating functions for MINLP
Consider the (discrete) optimization problem $\max \{ f(x) : x \in F \}$ where $f$ is non-negative on $F$ and the feasible region $F$ is finite.

**Approximation properties of $\ell_p$ norms (Hölder)**

\[ F = \{ x^1, \ldots, x^N \} \quad \text{and} \quad f = \begin{pmatrix} f(x^1) \\ \vdots \\ f(x^N) \end{pmatrix} \in \mathbb{R}^N. \]

**Dual bound** (upper bound on maximum value):

\[ \| f \|_{\infty} \leq \| f \|_p =: U_p \]

**Primal bound** (lower bound on maximum value):

\[ L_p := N^{-1/p} \| f \|_p = \frac{\| f \|_p}{\| 1 \|_p} \leq \| f \|_{\infty} \]

(both bounds converge to $\| f \|_{\infty}$, oblivious to properties of $f$)

Continuous case: Use properties of $f$, e.g., a Lipschitz constant.
Fix $d$.

- Linear $f$ can be optimized in \textit{polynomial time} \hfill \text{(Lenstra, 1983)}
- Convex polynomial $f$ can be minimized in \textit{polynomial time} \hfill \text{(Khachiyan–Porkolab, 2000)}
- Optimizing an arbitrary degree-4 polynomial $f$ for $d = 2$ is \textbf{NP-hard}

\textbf{Theorem (Fully Polynomial-Time Approximation Scheme, FPTAS)}

For every $\epsilon > 0$, there exists an algorithm $A_\epsilon$ for \textit{non-negative} $f$ with running time polynomial in the input size and $1/\epsilon$, which computes an approximation $x_\epsilon \in P \cap \mathbb{Z}^d$ with
\[
|f(x_\epsilon) - f(x_{\text{max}})| \leq \epsilon f(x_{\text{max}}).
\]

\textbf{“Weak” ("Range-relative") FPTAS for Integer Optimization}

\ldots for arbitrary $f$ \ldots computes a solution $x_\epsilon \in P \cap \mathbb{Z}^d$ with
\[
|f(x_\epsilon) - f(x_{\text{max}})| \leq \epsilon |f(x_{\text{max}}) - f(x_{\text{min}})|.
\]

Via discretization: FPTAS for \textit{fixed number} of integer and \textit{continuous} variables.
Inspired by the preprint:
- C. Buchheim and C. D’Ambrosio, Box-constrained mixed-integer polynomial optimization using separable underestimators, 2013,
we consider separable polynomial functions over box domains.

If $f \geq 0$ on $F$, if $p = \lceil (1 + \frac{1}{\epsilon}) \log N \rceil$,
then $L_p$ is a $(1 - \epsilon)$-approximation, $U_p$ a $(1 + \epsilon)$-approximation to $f_{\text{max}}$.
- Consider the box domains $[-M, M]^20$ for $M = 10, M = 100, M = 1000, M = 10000$.
- Draw ten random homogeneous degree-4 separable polynomials $f$ in 20 variables.
- Apply a range-approximating shifting trick to make $f + s \geq 0$.

Plot the actual approximation error of primal (lower) bounds $L_p$ for different prescribed $\epsilon$ values.

The actual approximation error for the dual (upper) bounds $U_p$ is much lower.
Even with $\epsilon = 1$ the percent error is $< 1\%$ for $M \geq 100$.
The $[-10, 10]^20$ box starts at $6\%$. 
Computational experiments – computation time of the bounds
Ongoing experiments by Ph.D. student Brandon Dutra, with J.A. De Loera

The table below show the average time in seconds needed to compute the upper and lower bounds.

<table>
<thead>
<tr>
<th>$M$</th>
<th>$\epsilon = 0.9$</th>
<th>$\epsilon = 0.7$</th>
<th>$\epsilon = 0.5$</th>
<th>$\epsilon = 0.3$</th>
<th>$\epsilon = 0.1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.14</td>
<td>0.20</td>
<td>0.48</td>
<td>1.46</td>
<td>30.34</td>
</tr>
<tr>
<td>100</td>
<td>0.96</td>
<td>1.34</td>
<td>2.60</td>
<td>7.54</td>
<td>271.76</td>
</tr>
<tr>
<td>1000</td>
<td>2.68</td>
<td>3.88</td>
<td>7.40</td>
<td>26.66</td>
<td>1426.20</td>
</tr>
<tr>
<td>10000</td>
<td>6.14</td>
<td>10.46</td>
<td>21.16</td>
<td>82.14</td>
<td>4182.74</td>
</tr>
</tbody>
</table>

(Preliminary implementation in LattE integrale.)


http://www.math.ucdavis.edu/~latte/
Intermediate sums and mixed-integer optimization

The idea to use intermediate sums appeared first in Barvinok (2006), for the computation of the top $k$ Ehrhart coefficients of a rational simplex in varying dimension. We take them to the generating-function (Laplace-transform) level and use them for mixed-integer optimization.

**Theorem (SL version of the Khovanskii–Pukhlikov theorem)**

Let $L \subseteq V$ be a rational subspace. There exists a unique valuation $S^L$ which to every rational polyhedron $P \subseteq V$ associates a meromorphic function with rational coefficients $S^L(P) \in \mathcal{M}(V^*)$ so that the following properties hold:

1. If $P$ contains a line, then $S^L(P) = 0$.
2. 

$$S^L(P)(\xi) = \sum_{y \in \Lambda \setminus V/L} \int_{P \cap (y + L)} e^{\langle \xi, x \rangle} dm_L(x),$$

for every $\xi \in V^*$ such that the above sum converges.
3. For every point $s \in \Lambda + L$, we have

$$S^L(s + P)(\xi) = e^{\langle \xi, s \rangle} S^L(P)(\xi).$$
Fix a non-negative integer $k_0$. There exists a polynomial time algorithm for the following problem. Given the following input:

1. A simple polytope $P \subset \mathbb{R}^d$, represented by its vertices, rational vectors $s_1, \ldots, s_{d+1} \in \mathbb{Q}^d$ in binary encoding,
2. A subspace $L \subseteq \mathbb{Q}^d$ of codimension $k_0$, represented by $d - k_0$ linearly independent vectors $b_1, \ldots, b_{d-k_0} \in \mathbb{Q}^d$ in binary encoding,

compute the rational data such that we have the following equality of meromorphic functions:

$$S^L(P)(\xi) = \sum_{n \in \mathbb{N}} \alpha^{(n)} \left( e^{\langle \xi, s^{(n)} \rangle} \prod_{i=1}^{k_0} T(z_i^{(n)}, \langle \xi, w_i^{(n)} \rangle) \right) \frac{1}{\prod_{i=1}^{d} \langle \xi, w_i^{(n)} \rangle}.$$ 

From this, we can extract intermediate sums of powers of linear forms $g(x) = \langle \ell, x \rangle^M$ using series expansions in $O(1)$ variables.

Sums of powers of linear forms generalize separable polynomial functions.
Powers of linear forms are enough: The polynomial Waring problem

Theorem (Alexander–Hirschowitz, 1995)

A generic homogeneous polynomial of degree $M$ in $n$ variables is expressible as the sum of

$$r(M, n) = \left\lceil \frac{(n+M-1)}{M} \right\rceil$$

$M$-th powers of linear forms, with the exception of the cases $r(3, 5) = 8$, $r(4, 3) = 6$, $r(4, 4) = 10$, $r(4, 5) = 15$, and $M = 2$, where $r(2, n) = n$. (Non-constructive.)

Theorem (Carlini–Catalisano–Geramita, 2011)

Minimal, constructive solution for monomials $x^M$, $M_1 \leq \cdots \leq M_n$ with $\prod_{i=2}^n (M_i + 1)$, involving roots of unity.

Effective (constructive) version?


Simple (suboptimal) rational constructions

$$x^M = \frac{1}{|M|!} \sum_{0 \leq p_i \leq M_i} \alpha_p (p_1 x_1 + \cdots + p_n x_n)^{|M|}$$

with $\alpha_p = (-1)^{|M|-(p_1+\cdots+p_n)} \binom{M_1}{p_1} \cdots \binom{M_n}{p_n}$
Three sources of computational difficulty

1. Size of the decomposition into powers of linear forms (especially for $g = f^p$ with large $p$)
   - Precompute the decomposition of $g(x) = (f(x) + s)^p$ as a function of the shift $s$
   - Numerical linear algebra techniques to compute decompositions.
   - Approximative decompositions into powers of linear forms?
   - Use different functions $g(x) = h(f(x))$?
   - For separable functions $f$ over boxes, all cross terms can be removed.
   - Results for near-separable objective functions?

2. Arithmetic complexity (determinants) of vertex cones
   - Number of integer variables will have to be small enough.
   - Use discretization-free intermediate summation.
   - Relax integrality in some discrete directions.

3. Number of vertex cones
   - Assume number of constraints $= \text{dimension} + O(1)$.
   - Use “restriction formulations” with a penalty method (next slide)?
Yannakakis’ theorem:

- Let $P \subseteq \mathbb{R}^d$ be a polytope of dimension $\geq 1$.
- Consider extended formulations $Q \subseteq \mathbb{R}^d \times \mathbb{R}^m$, i.e., $\{ x \in \mathbb{R}^d : \exists y : (x, y) \in Q \} = P$.
- The extension complexity (smallest number of facets of any extended formulation) is equal to the non-negative rank of any slack matrix of $P$.
- $Q$ can be made bounded at the cost of 1 extra facet.

Dualize:

- Consider “restriction formulations” $R \subseteq \mathbb{R}^d \times \mathbb{R}^m$, i.e., $\{ x \in \mathbb{R}^d : (x, 0) \in R \} = P$.
- The “restriction–vertex-count complexity” (smallest number of vertices and rays of any restriction formulation) is equal to the non-negative rank of any slack matrix of $P$.
- $R$ can be made bounded at the cost of 1 extra vertex.

Enter penalty methods:

\[
\begin{align*}
\max & \quad f(x) \\
\text{s.t.} & \quad x \in P
\end{align*} \quad \rightarrow \quad \begin{align*}
\max & \quad f(x) \\
\text{s.t.} & \quad (x, z) \in R \\
& \quad z = 0
\end{align*} \quad \rightarrow \quad \begin{align*}
\max & \quad f(x) - \sum_i h_i(z_i) \\
\text{s.t.} & \quad (x, z) \in R
\end{align*}
\]
Computation of the highest coefficients of weighted
Ehrhart quasi-polynomials of rational polyhedra.
Foundations of Computational Mathematics,

Intermediate sums on polyhedra II: Bidegree and Poisson
summation formula.

V. Baldoni, N. Berline, M. Köppe, and M. Vergne.
Intermediate sums on polyhedra: Computation and real
Ehrhart theory.

J. A. De Loera, B. E. Dutra, M. Köppe, S. Moreinis,
G. Pinto, and J. Wu.
Software for exact integration of polynomials over
polyhedra.
Computational Geometry: Theory and Applications,