Symmetries in AM/GM based Optimization

Philippe Moustrou, Helen Naumann, Cordian Riener, Thorsten Theobald, and Hugues Verdure (arXiv:2102.12913)

MEGA, Tromsø, Norway, June 2021

Optimization and Non-negativity

Signomial: Exponential sum supported on finite $\mathcal{T} \subseteq \mathbb{R}^n$:

$$f = \sum_{\alpha \in \mathcal{T}} c_\alpha e^{\langle \mathbf{x}, \alpha \rangle}, \quad c_\alpha \in \mathbb{R} \text{ for all } \alpha \in \mathcal{T}.$$

Signomial optimization problem

$$f^* = \inf\{f(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^n\} = \sup\{\lambda \in \mathbb{R} : f - \lambda \geq 0\}$$

- Deciding non-negativity NP-hard in general $\rightarrow$ find suitable certificate of nonnegativity.
- Sparse approaches building upon earlier work by Reznick based on
  - SONC (Iliman, de Wolff, 2016) or
  - SAGE (Chandrasekaran, Shah, 2016)

Main idea

Symmetric problem: invariant under some group action

1. We prove a symmetry-adapted decomposition theorem for SAGE exponentials.
2. This decomposition reduces the size of a Relative Entropy Program deciding non-negativity.
3. Computational experiments:
   - Strong reductions of size and computation time
   - Cases where symmetry-adapted computation succeeds when conventional SAGE computation fails

Sums of Arithmetic-Geometric Exponentials

AGE exponentials

- Arithmetic-geometric exponential: For $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$ finite: $\sum_{\alpha \in \mathcal{A}} c_\alpha \, \mathrm{e}^{\langle \mathbf{x}, \alpha \rangle} + c_\beta \, \mathrm{e}^{\langle \mathbf{x}, \beta \rangle}$ with $c_\alpha \geq 0$ for all $\alpha \in \mathcal{A}$
- Non-negativity can be certified via arithmetic-geometric inequality
- SAGE cone: Sums of AGE exponentials
- Optimization approach: $\mathcal{T} := \mathcal{A} \cup \mathcal{B}$ with disjoint sets $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$, corresponding to positive coefficients c_α, $\mathcal{B} \subseteq \mathbb{R}^n$ corresponding to negative coefficients c_β

$$\rightarrow f = \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \mathbf{x}, \alpha \rangle} + \sum_{\beta \in \mathcal{B}} c_\beta e^{\langle \mathbf{x}, \beta \rangle}.$$

Example: Exponential version of Motzkin polynomial

Example

$f = 1 + e^{2x+4y} + e^{4x+2y} - 3\,e^{2x+2y}$ with

$A = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \begin{pmatrix} 4 \\ 2 \end{pmatrix} \right\}, \beta = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \in \operatorname{int} \operatorname{conv} A.$

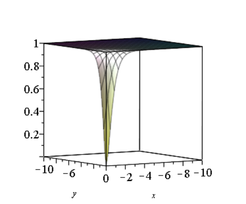


Figure 1: Motzkin

$\rightarrow$ is an AGE exponential.

Some results on the SAGE Cone

- Chandrasekaran, Shah: Deciding membership in the SAGE cone can be done efficiently via Relative Entropy Programming
- Iliman, de Wolff: For a subclass, efficient optimization via Geometric Programming is possible
- Murray, Chandrasekaran, Wierman (for polynomials also observed by Wang): Sparsity of the support of a SAGE exponential is preserved in any decomposition
- Murray, Chandrasekaran, Wierman: Constrained SAGE approximation efficiently computable using REPs and the support function
- Murray, N., Theobald: Sublinear circuits for constrained SAGE

Relative Entropy Characterization

- Relative Entropy Function: $D : \mathbb{R}^{\mathcal{A}}_{>0} \times \mathbb{R}^{\mathcal{A}}_{>0} \to \mathbb{R}$ with

$$D(\nu, \gamma) \;=\; \sum_{\alpha \in \mathcal{A}} \nu_\alpha \ln\left(\frac{\nu_\alpha}{\gamma_\alpha}\right)$$

for $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$ finite

Theorem (Murray, Chandrasekaran, Wierman)

Signomial f belongs to $C_{\mathrm{SAGE}}(\mathcal{A}, \mathcal{B})$ iff for every $\beta \in \mathcal{B}$ there exist $c^{(\beta)} \in \mathbb{R}^{\mathcal{A}}_{+}$ and $\nu^{(\beta)} \in \mathbb{R}^{\mathcal{A}}_{+}$ such that

$$\begin{aligned}
\sum_{\alpha \in \mathcal{A}} \nu^{(\beta)}_\alpha (\alpha - \beta) \;&=\; 0 && \text{for every } \beta \in \mathcal{B}, \\
D(\nu^{(\beta)}, e \cdot c^{(\beta)}) \;&\leq\; c_\beta && \text{for every } \beta \in \mathcal{B}, \\
\sum_{\beta \in \mathcal{B}} c^{(\beta)}_\alpha \;&\leq\; c_\alpha && \text{for every } \alpha \in \mathcal{A}.
\end{aligned}$$

Symmetry reductions in AM/GM based Optimization

Some Notation

- G finite group acting linearly on $\mathbb{R}^n$ on the left
- Orbit of set $\mathcal{S} \subseteq \mathbb{R}^n$ of exponent vectors under G:

$$G \cdot \mathcal{S} = \{\sigma(s) : s \in \mathcal{S},\, \sigma \in G\}.$$

- Set of Orbit Representatives $\hat{\mathcal{S}} \subseteq \mathcal{S}$ for $\mathcal{S}$: inclusion-minimal set with $(G \cdot \hat{\mathcal{S}}) = \mathcal{S}$.
- Stabilizer of β: $\operatorname{Stab} \beta := \{\sigma \in G \;:\; \sigma(\beta) = \beta\}$
- Left Quotient Space

$$G / \operatorname{Stab}(\beta) = \{\{h \in G : \exists\, \sigma \in \operatorname{Stab}(\beta) \text{ with } h = g \cdot \sigma\} : g \in G\}$$

Main results(1)

Reminder: $f = \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \mathbf{x}, \alpha \rangle} + \sum_{\beta \in \mathcal{B}} c_\beta e^{\langle \mathbf{x}, \beta \rangle} \in \mathbb{R}^{\mathcal{T}}$ with $c_{|\mathcal{A}} \in \mathbb{R}_+^{\mathcal{A}}$, $c_{|\mathcal{B}} \in -\mathbb{R}_+^{\mathcal{B}}$ with orbit representatives $\hat{\mathcal{B}}$ of $\mathcal{B}$.

Theorem (Moustrou, N., Riener, Theobald, Verdure)

A G-symmetric $f \in C_{\mathrm{SAGE}}(\mathcal{A}, \mathcal{B})$ iff for every $\hat{\beta} \in \hat{\mathcal{B}}$, there exists an AGE exponential $h_{\hat{\beta}} \in C_{\mathrm{AGE}}(\mathcal{A}, \hat{\beta})$ such that

$$f = \sum_{\hat{\beta} \in \hat{\mathcal{B}}} \sum_{\rho \in G / \mathrm{Stab}(\hat{\beta})} \rho h_{\hat{\beta}}$$

and $h_{\hat{\beta}}$ invariant under the action of $\mathrm{Stab}(\hat{\beta})$

Idea of proof

- Clearly, signomial $f = \sum_{\hat{\beta} \in \hat{\mathcal{B}}} \sum_{\rho \in G / \text{Stab}(\hat{\beta})} \rho h_{\hat{\beta}}$ non-negative
- $f \in C_{\text{SAGE}}(\mathcal{A}, \mathcal{B}) \Rightarrow$ for every $\beta \in \mathcal{B}$ exists $f_\beta \in C_{\text{AGE}}(\mathcal{A}, \beta)$: $f = \sum_{\beta \in \mathcal{B}} f_\beta$.
- f G-symmetric implies:

$$f = \frac{1}{|G|} \sum_{\sigma \in G} \sigma f = \frac{1}{|G|} \sum_{\sigma \in G} \sum_{\beta \in \mathcal{B}} \sigma f_\beta.$$

- Idea: Group all σf_β that have same "possibly negative" term:

$$h_{\hat{\beta}} = \frac{1}{|G|} \sum_{\sigma \in G} \sigma f_{\sigma^{-1} \hat{\beta}}$$

Main results(2)

Theorem (Moustrou, N., Riener, Theobald, Verdure)

G-symmetric $f \in C_{\mathrm{SAGE}}(\mathcal{A}, \mathcal{B})$ iff for every $\beta \in \hat{\mathcal{B}}$ there exist $c^{(\beta)} \in \mathbb{R}_+^{\mathcal{A}}, \nu^{(\beta)} \in \mathbb{R}_+^{\mathcal{A}}$, invariant under the action of Stab(β), s. t.

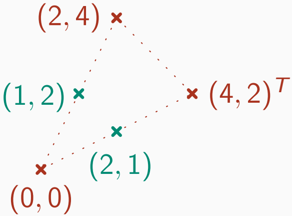
$$\begin{aligned} \sum_{\alpha \in \mathcal{A}} \nu_\alpha^{(\hat{\beta})}(\alpha - \beta) &= 0 \quad \text{for every } \beta \in \hat{\mathcal{B}}, \\ D(\nu^{(\beta)}, e \cdot c^{(\beta)}) &\leq c_\beta \quad \text{for every } \beta \in \hat{\mathcal{B}}, \\ \sum_{\beta \in \hat{\mathcal{B}}} \sum_{\sigma \in \mathrm{Stab}\,(\beta) \backslash G} c_{\sigma(\alpha)}^{(\beta)} &\leq c_\alpha \quad \text{for every } \alpha \in \mathcal{A}. \end{aligned}$$

$\rightarrow$ Stab$(\beta)\backslash G$: right quotient space

Example: Comparison of Standard and Symmetric Method

$$f = 1 + \mathrm{e}^{2x+4y} + \mathrm{e}^{4x+2y} + \delta(\mathrm{e}^{x+2y} + \mathrm{e}^{2x+y})$$

$\rightarrow$ has optimal value $\delta = -\sqrt{2}$,



- Standard REP : $|\mathcal{B}| \cdot n + |\mathcal{B}| + |\mathcal{A}| = 9$ ineq., $|\mathcal{B}| \cdot |\mathcal{A}| = 6$ var.
- Symmetric REP: $|\hat{\mathcal{B}}| \cdot n + |\hat{\mathcal{B}}| + |\mathcal{A}| = 6$ ineq., $|\hat{\mathcal{B}}| \cdot |\mathcal{A}| = 3$ var.

The Case of the Symmetric Group

Stabilization of Symmetric REP

Symmetric group: $\mathcal{S}_n$

$\rightarrow$ $\mathrm{wt}(\alpha)$ weight of $\alpha \in \mathbb{R}^n$: number of non-zero coordinates

Theorem (Moustrou, N., Riener, Theobald, Verdure)

Let $w \in \mathbb{N}$ be fixed, $\mathcal{A}, \mathcal{B} \subseteq \mathbb{R}^n$ with sets of orbit representatives $\hat{\mathcal{A}}, \hat{\mathcal{B}}$, and

$$\max_{\hat{\gamma} \in \hat{\mathcal{A}} \cup \hat{\mathcal{B}}} \mathrm{wt}(\hat{\gamma}) \leqslant w.$$

$\Rightarrow$ For all $n \geqslant 2w$, $f \in \mathcal{C}_{\mathsf{SAGE}}(\mathcal{A}, \mathcal{B})$ $\mathcal{S}_n$-invariant: Symm. REP has

$$\#\text{coeff.} \leqslant |\hat{\mathcal{A}}| + |\hat{\mathcal{B}}| + |\hat{\mathcal{B}}|(w+1) \quad \text{and} \quad \#\text{var.} \leqslant 2|\hat{\mathcal{B}}||\hat{\mathcal{A}}|u(w),$$

where $u(w) = \displaystyle\sum_{i=0}^{w} \binom{w}{i}^2 i!\,.$

Exemplary Comparison of the numbers of variables and constraints

- Examples: Function with exponents in $\mathcal{A}, \mathcal{B} \subseteq \mathbb{R}^n$ and orbit representatives $\hat{\mathcal{A}}, \hat{\mathcal{B}}$
- Extremal situations: Orbits either very large or very small

		Standard method		Symmetric method	
$\|\mathcal{B}\|$	$\|\mathcal{A}\|$	# Var.	# Coeff.	# Var.	# Coeff.
1	$n!$	$2n! + 3$	$n! + n + 2$	5	4
$n!$	n	$2(n+1)n! + 1$	$(n+1)(n!+1)$	$2n+3$	$n+3$
$n!$	$n!$	$2(n!+1)n! + 1$	$n!(n+2)+1$	$2n!+3$	$n+3$
n	n	$2n(n+1)+1$	$(n+1)^2$	7	5

Table 1: Comparison of the parameters when $\hat{\mathcal{A}} = \{0, \hat{\alpha}\}$ and $\hat{\mathcal{B}} = \{\hat{\beta}\}$.

Numerical Experiments

Abbreviations

- Underlying group: Symmetric group $\mathcal{S}_n$
- Standard method: Classical REP without exploitation of symmetries
- Symmetric version (including combinatorial techniques to further reduce size)
- dim: Dimension of the exponential sum
- V_n:= Number of auxiliary variables ν in REP
- C_n:= Number of coefficients in REP
- t:= Running time of REP (including preliminary computations)

Example (1)

	Standard method			Symmetric method		
dim	V_n	C_n	t	V_n	C_n	t
2	13	9	0.0185	7	5	0.0311
3	49	28	0.0454	9	6	0.0264
4	241	125	0.1701	11	7	0.0318
5	1441	726	0.8433	13	8	0.0384
6	10081	5047	5.843	15	9	0.0458
7	80641	40328	66.67	17	10	0.0538
8	725761	362889	2211	19	11	0.0626
9	7257601	3628810	–	21	12	0.0835
10	79833601	39916811	–	23	13	–

Table 2: Exponential sum with $\mathcal{B} = \mathcal{S}_n \cdot \{(1, 2, \ldots, n)^T\}$ and $\mathcal{A} = \mathcal{S}_n \cdot \{(n^2, 0, \ldots, 0)^T\}$

Example (2)

	Standard method			Symmetric method		
dim	V_n	C_n	t	V_n	C_n	t
2	13	9	0.0323	7	5	0.0465
3	85	31	0.0603	15	6	0.0569
4	1201	145	–	51	7	0.1301
5	29041	841	–	243	8	0.6215
6	1038241	5761	–	1443	9	–

Table 3: Exponential sum with two orbits of maximal size

Outlook

- This paper also includes:
  - Symmetric decomposition theorem also works for the constrained situation
- Follow-up work will include:
  - Minimizers of symmetric SAGE exponentials
  - The symmetric SAGE cone and the non-negativity cone

Thank you for your attention!

Some Literature

- J. B. Lasserre, "*Global optimization with polynomials and the problem of moments*", SIAM Journal of Optimization, 2011.
- C. Riener, T. Theobald, L. Jansson-Andrén, J. B. Lasserre, "*Exploiting symmetries in SDP-relaxations for polynomial optimization*" Mathematics of Operations Research, 2013.
- S. Iliman, T. de Wolff, "*Amoebas, Nonnegative Polynomials and Sums of Squares Supported on Circuits*", Research in the Mathematical Sciences, 2016.
- V. Chandrasekaran, P. Shah, "*Relative Entropy Relaxations for Signomial Optimization*", SIAM Journal of Optimization, 2016.
- R. Murray, V. Chandrasekaran, A. Wierman, "*Newton Polytopes and Relative Entropy Optimization*", Found. Comput. Math., 2021.
- R. Murray, V. Chandrasekaran, A. Wierman, "*Signomial and Polynomial Optimization via Relative Entropy and partial Dualization*", Math. Program. Comput., 2021.
- R. Murray, H. Naumann, T. Theobald "*Sublinear Circuits and the Constrained Signomial Nonnegativity Problem*", arXiv:2006.06811.