Complexity of Deciding Convexity in Polynomial Optimization

Amir Ali Ahmadi

Joint work with:
Alex Olshevsky, Pablo A. Parrilo, and John N. Tsitsiklis

Laboratory for Information and Decision Systems
Massachusetts Institute of Technology

SIAM Conference on Optimization
May 18, 2011
Darmstadt, Germany
Convexity

Rockafellar, ’93:

“In fact the great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity.”

But how easy is it to distinguish between convexity and nonconvexity?

This talk:

-- Given a multivariate polynomial, can we efficiently **decide** if it is convex?

-- Given a basic semialgebraic set, can we efficiently **decide** if it is a **convex set**?
Convexity in optimization

- Global optimization
  -- Minimizing polynomials is **NP-hard** for degree $\geq 4$
  -- But if polynomial is known to be convex, even simple gradient descent methods can find a global min

(often we check convexity based on “simple rules” from calculus of convex functions)

**Applications:** convex envelopes, convex data fitting, defining norms
Complexity of deciding convexity

- Input to the problem: an ordered list of the coefficients (all rational)
- Degree $d$ odd: trivial
- $d=2$, i.e., $p(x)=x^TQx+q^Tx+c$: check if $Q$ is PSD
- $d=4$, first interesting case
  - Question of N. Z. Shor:
    - “What is the complexity of deciding convexity of a multivariate polynomial of degree four?”

(appeared on a list of seven open problems in complexity of numerical optimization in 1992, [Pardalos, Vavasis])

Our main result: problem is strongly NP-hard
Agenda for the rest of the talk

1. Idea of the proof
2. Complexity of deciding variants of convexity
   -- (strong, strict, pseudo, quasi)-convexity
NP-hardness of deciding convexity of quartics

**Thm:** Deciding convexity of quartic forms is strongly NP-hard.

- Reduction from problem of deciding “nonnegativity of biquadratic forms”

**Biquadratic form:**

\[ b(x; y) = \sum_{i \leq j, k \leq l} \alpha_{ijkl} x_i x_j y_k y_l \]

- Can write any biquadratic form as \( y^T A(x) y \), where \( A(x) \) is a matrix whose entries are quadratic forms

**Example:** \( y^T A(x) y \), with \( A(x) = \begin{bmatrix} x_1^2 + 2x_2^2 & -x_1x_2 & -x_1x_3 \\ -x_1x_2 & x_2^2 + 2x_3^2 & -x_2x_3 \\ -x_1x_3 & -x_2x_3 & x_3^2 + 2x_1^2 \end{bmatrix} \)
Sequence of reductions

STABLE SET

Minimizing an indefinite quadratic form over the simplex

Nonnegativity of biquadratic forms

Convexity of quartic forms

\[
\frac{1}{\alpha(G)} = \min \quad x^T (A + I) x \\
\sum x_i = 1 \\
x_i \geq 0
\]

[Motzkin, Straus]

[Gurvits], [Ling, Nie, Qi, Ye]

Our work
The Hessian structure

- **Biquadratic form**: a form of the type $y^T A(x) y$, where $A(x)$ is a matrix whose entries are quadratic forms.

- **Example**: $y^T A(x) y$, with $A(x) = \begin{bmatrix} x_1^2 + 2x_2^2 & -x_1x_2 & -x_1x_3 \\ -x_1x_2 & x_2^2 + 2x_3^2 & -x_2x_3 \\ -x_1x_3 & -x_2x_3 & x_3^2 + 2x_1^2 \end{bmatrix}$

- **Biquadratic Hessian form**: Special biquadratic form where $A(x)$ is a valid Hessian.

- $A(x)$ above is *not* a valid Hessian:

  $\frac{\partial A_{1,1}(x)}{\partial x_3} = 0 \neq -x_3 = \frac{\partial A_{1,3}(x)}{\partial x_1}$
From biquadratic forms to biquadratic Hessian forms

- We give a constructive procedure to go from any biquadratic form \( y^T A(x) y \) to a biquadratic Hessian form \( z^T H(x, y) z \) by doubling the number of variables, such that:

\[
y^T A(x) y \text{ psd } \iff z^T H(x, y) z \text{ psd}
\]

- In fact, we construct the polynomial \( f(x,y) \) that has \( H(x,y) \) as its Hessian directly

- Let’s see this construction...
The main reduction

**Thm:** Given any biquadratic form $b(x; y)$,

Let $[C(x, y)]_{ij} := \frac{\partial b(x; y)}{\partial x_i \partial y_j}$. Let $\gamma := \max |\text{coeff}(C(x, y))|$.

Let

$$f(x, y) := b(x; y) + \frac{n^2 \gamma}{2} \left( \sum_{i=1}^{n} x_i^4 + \sum_{i=1}^{n} y_i^4 + \sum_{i,j=1,...,n, i<j} x_i^2 x_j^2 + \sum_{i,j=1,...,n, i<j} y_i^2 y_j^2 \right)$$

Then

$$b(x; y) \text{ psd} \iff f(x, y) \text{ convex}$$

$$H(x, y) = H_b(x, y) + H_g(x, y)$$

$$b(x; y) \text{ psd} \iff z^T H(x, y) z \text{ psd}$$
Observation on the Hessian of a biquadratic form

\[ b(x; y) = \sum_{i \leq j, \, k \leq l} \alpha_{ijkl} x_i x_j y_k y_l \]

\[
\begin{align*}
[A(x)]_{ij} &:= \frac{\partial b(x; y)}{\partial y_i \partial y_j} & [B(y)]_{ij} &:= \frac{\partial b(x; y)}{\partial x_i \partial x_j} \\
\frac{1}{2} y^T A(x) y &= b(x; y) & \frac{1}{2} x^T B(y) x &= b(x; y) \\
[C(x, y)]_{ij} &:= \frac{\partial b(x; y)}{\partial x_i \partial y_j}
\end{align*}
\]

\[ H_b(x, y) = \begin{bmatrix} B(y) & C(x, y) \\ C^T(x, y) & A(x) \end{bmatrix} \]
Proof of correctness of the reduction

Start with $b(x; y)$,

Let $[C(x, y)]_{ij} := \frac{\partial b(x; y)}{\partial x_i \partial y_j}$

Let $\gamma := \max |\text{coeff}(C(x, y))|$

$f(x, y) := b(x; y) + \frac{n^2 \gamma}{2} \left( \sum_{i=1}^{n} x_i^4 + \sum_{i=1}^{n} y_i^4 + \sum_{i<j} x_i^2 x_j^2 + \sum_{i<j} y_i^2 y_j^2 \right)$

$H(x, y) = H_b(x, y) + H_g(x, y)$

Claim: $b(x; y) \text{ psd } \iff z^T H(x, y) z \text{ psd}$

$$H(x, y) = \begin{bmatrix} B(y) & C(x, y) \\ C^T(x, y) & A(x) \end{bmatrix} + \frac{n^2 \gamma}{2} \begin{bmatrix} H_{11}^g(x) & 0 \\ 0 & H_{22}^g(y) \end{bmatrix}$$
Reduction on an instance

$$A(x) = \begin{bmatrix}
  x_1^2 + 2x_2^2 & -x_1x_2 & -x_1x_3 \\
  -x_1x_2 & x_2^2 + 2x_3^2 & -x_2x_3 \\
  -x_1x_3 & -x_2x_3 & x_3^2 + 2x_1^2 
\end{bmatrix}$$

A 6x6 **Hessian** with quadratic form entries
Other notions of interest in optimization

- **Strong convexity**
  - “Hessian uniformly bounded away from zero”
  - Appears e.g. in convergence analysis of Newton-type methods

- **Strict convexity**
  - “curve strictly below the line”
  - Guarantees uniqueness of optimal solution

- **Convexity**

- **Pseudoconvexity**
  - “Relaxation of first order characterization of convexity”
  - Any point where gradient vanishes is a global min

- **Quasiconvexity**
  - “Convexity of sublevel sets”
  - Deciding convexity of basic semialgebraic sets
# Summary of complexity results

<table>
<thead>
<tr>
<th>property vs. degree</th>
<th>1</th>
<th>2</th>
<th>odd $\geq 3$</th>
<th>even $\geq 4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>strong convexity</td>
<td>no</td>
<td>P</td>
<td>no</td>
<td>strongly NP-hard</td>
</tr>
<tr>
<td>strict convexity</td>
<td>no</td>
<td>P</td>
<td>no</td>
<td>strongly NP-hard</td>
</tr>
<tr>
<td><strong>convexity</strong></td>
<td>yes</td>
<td>P</td>
<td>no</td>
<td><strong>strongly NP-hard</strong></td>
</tr>
<tr>
<td>pseudoconvexity</td>
<td>yes</td>
<td>P</td>
<td>P</td>
<td>strongly NP-hard</td>
</tr>
<tr>
<td>quasiconvexity</td>
<td>yes</td>
<td>P</td>
<td>P</td>
<td>strongly NP-hard</td>
</tr>
</tbody>
</table>
A multivariate polynomial \( p(x) = p(x_1, \ldots, x_n) \) is quasiconvex if all its sublevel sets

\[
S_\alpha := \{ x \in \mathbb{R}^n \mid p(x) \leq \alpha \}
\]

are convex.

Convexity \( \Rightarrow \) Quasiconvexity

(converse fails)

- Deciding quasiconvexity of polynomials of even degree 4 or larger is strongly NP-hard
- Quasiconvexity of odd degree polynomials can be decided in polynomial time
Quasiconvexity of even degree forms

**Lemma:** A homogeneous polynomial $p(x)$ of even degree $d$ is quasiconvex if and only if it is convex.

**Proof:**

- A homogeneous quasiconvex polynomial is nonnegative.
- The unit sublevel sets of $p(x)$ and $p(x)^{1/d}$ are the same convex set.
- $p(x)^{1/d}$ is the Minkowski norm defined by this convex set and hence a convex function.
- A convex nonnegative function raised to a power $d$ larger than one remains convex.

**Corollaries:**

- Deciding quasiconvexity is NP-hard.
- Deciding convexity of basic semialgebraic sets is NP-hard.
**Thm:** The sublevel sets of a quasiconvex polynomial $p(x)$ of odd degree are halfspaces.

**Proof:**
- Show super level sets must also be convex sets
- Only convex set whose complement is also convex is a halfspace

**Thm:** A polynomial $p(x)$ of odd degree $d$ is quasiconvex iff it can be written as

$$p(x) = h(\xi^T x)$$

$\xi \in \mathbb{R}^n$, $h(t)$ monotonic univariate polynomial of degree $d$.

This representation can be checked in polynomial time.
What can we do?

One possibility: natural relaxation based on sum of squares

**Defn.** ([Helton, Nie]): A polynomial $p(x) := p(x_1, \ldots, x_n)$ is **sos-convex** if its Hessian factors as

$$H(x) = M^T(x)M(x)$$

for a possibly nonsquare polynomial matrix $M(x)$.

- $p(x)$ sos-convex $\Rightarrow$ $p(x)$ convex (obvious)
- Deciding sos-convexity: a **semidefinite program (SDP)**
Gap between convexity and sos-convexity

A convex form that is not sos-convex:

\[ p(x) = x_1^4 + x_2^4 + x_3^4 + x_4^4 + x_5^4 + x_6^4 + 2(x_1^2 x_2^2 + x_1^2 x_3^2 + x_2^2 x_3^2 + x_4^2 x_5^2 + x_4^2 x_6^2 + x_5^2 x_6^2) + \frac{1}{2}(x_1^2 x_4^2 + x_2^2 x_5^2 + x_3^2 x_6^2) + x_1^2 x_6^2 + x_2^2 x_4^2 + x_3^2 x_5^2 - (x_1 x_2 x_4 x_5 + x_1 x_3 x_4 x_6 + x_2 x_3 x_5 x_6) \]

convex=sos-convex?

<table>
<thead>
<tr>
<th>n,d</th>
<th>2</th>
<th>4</th>
<th>≥6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>2</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>3</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>≥4</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>n,d</th>
<th>2</th>
<th>4</th>
<th>≥6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>2</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>3</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>≥4</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

[Ahmadi, Parrilo, ‘10]

[Ahmadi, Blekherman, Parrilo, ‘10]
Messages to take home...

<table>
<thead>
<tr>
<th>property vs. degree</th>
<th>1</th>
<th>2</th>
<th>odd $\geq 3$</th>
<th>even $\geq 4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>strong convexity</td>
<td>no</td>
<td>P</td>
<td>no</td>
<td>strongly NP-hard</td>
</tr>
<tr>
<td>strict convexity</td>
<td>no</td>
<td>P</td>
<td>no</td>
<td>strongly NP-hard</td>
</tr>
<tr>
<td>convexity</td>
<td>yes</td>
<td>P</td>
<td>no</td>
<td>strongly NP-hard</td>
</tr>
<tr>
<td>pseudoconvexity</td>
<td>yes</td>
<td>P</td>
<td>P</td>
<td>strongly NP-hard</td>
</tr>
<tr>
<td>quasiconvexity</td>
<td>yes</td>
<td>P</td>
<td>P</td>
<td>strongly NP-hard</td>
</tr>
</tbody>
</table>

- **SOS-Convexity**: a powerful SDP relaxation for convexity

<table>
<thead>
<tr>
<th>convex=sos-convex?</th>
<th>Polynomials</th>
<th>Forms</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n,d$</td>
<td>$2$</td>
<td>$4$</td>
</tr>
<tr>
<td>1</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>2</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>3</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>$\geq 4$</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>
Thank you for your attention!
Questions?

Want to know more?
http://aaa.lids.mit.edu/