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**?

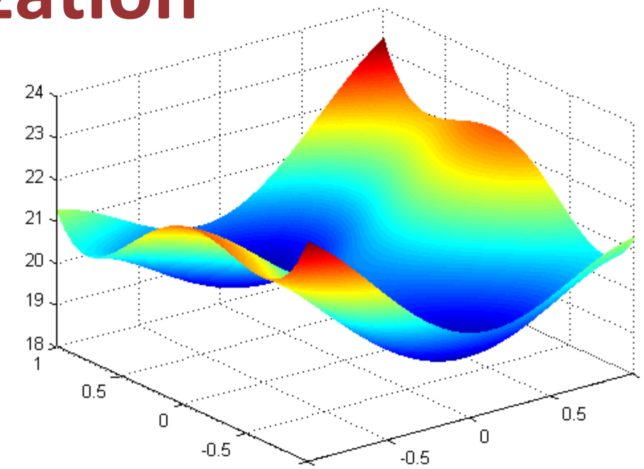
Convexity in optimization

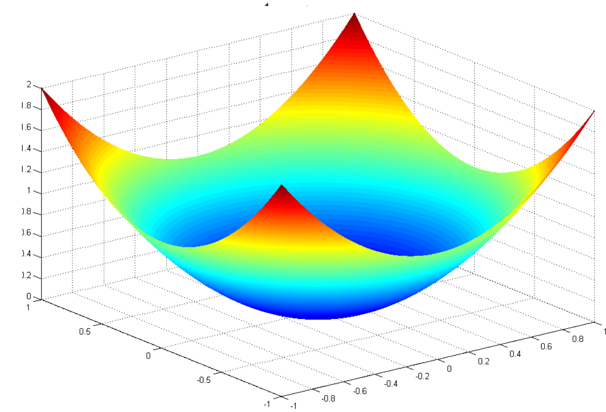
- Global optimization

  -- Minimizing polynomials is **NP-hard for degree ≥ 4**

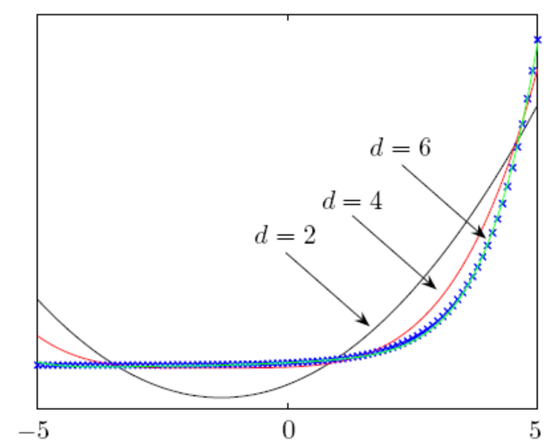
  -- But if polynomial is known to be convex, even simple gradient descent methods can find a global min

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





Applications: convex envelopes, convex data fitting, defining norms



[Magnani, Lall, Boyd]

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

[Motzkin, Straus] $\frac{1}{\alpha(G)} = \min_{\substack{\sum x_i = 1 \\ x_i \geq 0}} x^T(A+I)x$

Minimizing an indefinite quadratic form over the simplex

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

Nonnegativity of biquadratic forms

Our work

Convexity of quartic forms

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:

$$y^T A(x) y \;\; \text{psd} \Leftrightarrow z^T H(x,y) z \;\; \text{psd}$$

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

Let

$$f(x,y) := b(x;y) + \underbrace{\frac{n^2\gamma}{2}\Big(\sum_{i=1}^{n} x_i^4 + \sum_{i=1}^{n} y_i^4 + \sum_{\substack{i,j=1,\ldots,n \\ i<j}} x_i^2 x_j^2 + \sum_{\substack{i,j=1,\ldots,n \\ i<j}} y_i^2 y_j^2\Big)}_{g(x,y)}$$

Then

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

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

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

Observation on the Hessian of a biquadratic form

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

$$[A(x)]_{ij} := \frac{\partial b(x;y)}{\partial y_i \partial y_j} \qquad [B(y)]_{ij} := \frac{\partial b(x;y)}{\partial x_i \partial x_j}$$

$$\frac{1}{2} y^T A(x) y = b(x;y) \qquad \frac{1}{2} x^T B(y) x = b(x;y)$$

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

$$H_b(x,y) = \begin{bmatrix} B(y) & C(x,y) \\ C^T(x,y) & A(x) \end{bmatrix}$$

Proof of correctness of the reduction

Start with $b(x;y)$,

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

$$f(x,y) := b(x;y) + \underbrace{\frac{n^2\gamma}{2}\Big(\sum_{i=1}^{n} x_i^4 + \sum_{i=1}^{n} y_i^4 + \sum_{\substack{i,j=1,\ldots,n\\ i<j}} x_i^2 x_j^2 + \sum_{\substack{i,j=1,\ldots,n\\ i<j}} y_i^2 y_j^2\Big)}_{g(x,y)}$$

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

Claim: $b(x;y)$ psd $\Leftrightarrow z^T H(x,y) z$ psd

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

Reduction on an instance

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

$$\begin{bmatrix} \boxed{y_1^2+2y_3^2}+24x_1^2+4x_2^2+4x_3^2 & \boxed{-y_1y_2}+8x_1x_2 & \boxed{-y_1y_3}+8x_1x_3 & 2x_1y_1-x_2y_2-x_3y_3 & -x_2y_1 & 4x_1y_3-x_3y_1 \\ \boxed{-y_1y_2}+8x_1x_2 & \boxed{2y_1^2+y_2^2}+24x_2^2+4x_1^2+4x_3^2 & \boxed{-y_2y_3}+8x_2x_3 & -x_1y_2+4x_2y_1 & -x_1y_1+2x_2y_2-x_3y_3 & -x_3y_2 \\ \boxed{-y_1y_3}+8x_1x_3 & \boxed{-y_2y_3}+8x_2x_3 & \boxed{2y_2^2+y_3^2}+24x_3^2+4x_1^2+4x_2^2 & -x_1y_3 & -x_2y_3+4x_3y_2 & -x_1y_1-x_2y_2+2x_3y_3 \\ 2x_1y_1-x_2y_2-x_3y_3 & -x_1y_2+4x_2y_1 & -x_1y_3 & \boxed{x_1^2+2x_2^2}+24y_1^2+4y_2^2+4y_3^2 & \boxed{-x_1x_2}+8y_1y_2 & \boxed{-x_1x_3}+8y_1y_3 \\ -x_2y_1 & -x_1y_1+2x_2y_2-x_3y_3 & -x_2y_3+4x_3y_2 & \boxed{-x_1x_2}+8y_1y_2 & \boxed{x_2^2+2x_3^2}+24y_2^2+4y_1^2+4y_3^2 & \boxed{-x_2x_3}+8y_2y_3 \\ 4x_1y_3-x_3y_1 & -x_3y_2 & -x_1y_1-x_2y_2+2x_3y_3 & \boxed{-x_1x_3}+8y_1y_3 & \boxed{-x_2x_3}+8y_2y_3 & \boxed{2x_1^2+x_3^2}+24y_3^2+4y_1^2+4y_2^2 \end{bmatrix}$$

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

$\Downarrow$

- Strict convexity
  - "curve strictly below the line"
  - Guarantees uniqueness of optimal solution

$\Downarrow$

- Convexity

$\Downarrow$

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

$\Downarrow$

- Quasiconvexity
  - "Convexity of sublevel sets"
  - Deciding convexity of basic semialgebraic sets

Summary of complexity results

property vs. degree	1	2	odd ≥ 3	even ≥ 4
strong convexity	no	P	no	strongly NP-hard
strict convexity	no	P	no	strongly NP-hard
convexity	yes	P	no	strongly NP-hard
pseudoconvexity	yes	P	P	strongly NP-hard
quasiconvexity	yes	P	P	strongly NP-hard

Quasiconvexity

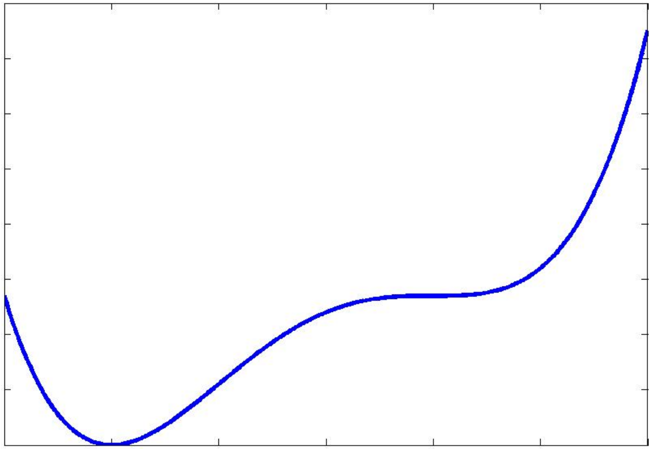
A multivariate polynomial $p(x)=p(x_1,\ldots,x_n)$ is *quasiconvex* if all its sublevel sets

$$\mathcal{S}_\alpha := \{x \in \mathbb{R}^n \mid p(x) \leq \alpha\}$$

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:

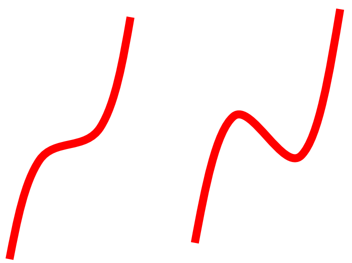
-- Deciding quasiconvexity is NP-hard

-- Deciding convexity of basic semialgebraic sets is NP-hard

Quasiconvexity of odd degree polynomials

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:

$$
\begin{aligned}
p(x) = {} & x_1^4 + x_2^4 + x_3^4 + x_4^4 + x_5^4 + x_6^4 \\
& + 2(x_1^2x_2^2 + x_1^2x_3^2 + x_2^2x_3^2 + x_4^2x_5^2 + x_4^2x_6^2 + x_5^2x_6^2) \\
& + \frac{1}{2}(x_1^2x_4^2 + x_2^2x_5^2 + x_3^2x_6^2) \\
& + x_1^2x_6^2 + x_2^2x_4^2 + x_3^2x_5^2 \\
& - (x_1x_2x_4x_5 + x_1x_3x_4x_6 + x_2x_3x_5x_6)
\end{aligned}
$$

[Ahmadi, Parrilo, ‘10]

convex=sos-convex?

Polynomials

n,d	2	4	≥6
1	yes	yes	yes
2	yes	yes	no
3	yes	no	no
≥4	yes	no	no

Forms

n,d	2	4	≥6
1	yes	yes	yes
2	yes	yes	yes
3	yes	yes	no
≥4	yes	no	no

[Ahmadi, Parrilo, ‘10]

[Ahmadi, Blekherman, Parrilo, ‘10]

Messages to take home…

property vs. degree	1	2	odd ≥ 3	even ≥ 4
strong convexity	no	P	no	strongly NP-hard
strict convexity	no	P	no	strongly NP-hard
convexity	yes	P	no	strongly NP-hard
pseudoconvexity	yes	P	P	strongly NP-hard
quasiconvexity	yes	P	P	strongly NP-hard

- SOS-Convexity: a powerful SDP relaxation for convexity

convex=sos-convex?

Polynomials

n,d	2	4	≥6
1	yes	yes	yes
2	yes	yes	no
3	yes	no	no
≥4	yes	no	no

Forms

n,d	2	4	≥6
1	yes	yes	yes
2	yes	yes	yes
3	yes	yes	no
≥4	yes	no	no

Thank you for your attention!

Questions?

Want to know more?

http://aaa.lids.mit.edu/