On Convexity of Polynomials over a Box

Georgina Hall
Decision Sciences, INSEAD

Joint work with
Amir Ali Ahmadi
ORFE, Princeton University

Convexity over a box

- A **box B** is a set of the form:

$$B = \{x \in \mathbb{R}^n \mid l_i \leq x_i \leq u_i, i = 1, \dots, n\}$$

where $l_1, \dots, l_n, u_1, \dots, u_n \in \mathbb{R}$ with $l_i \leq u_i$.

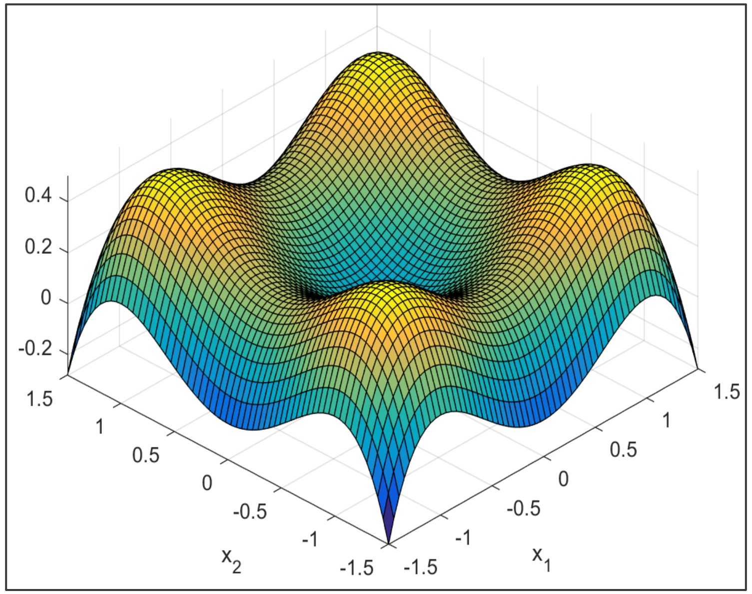
- A **function f is convex over B** if

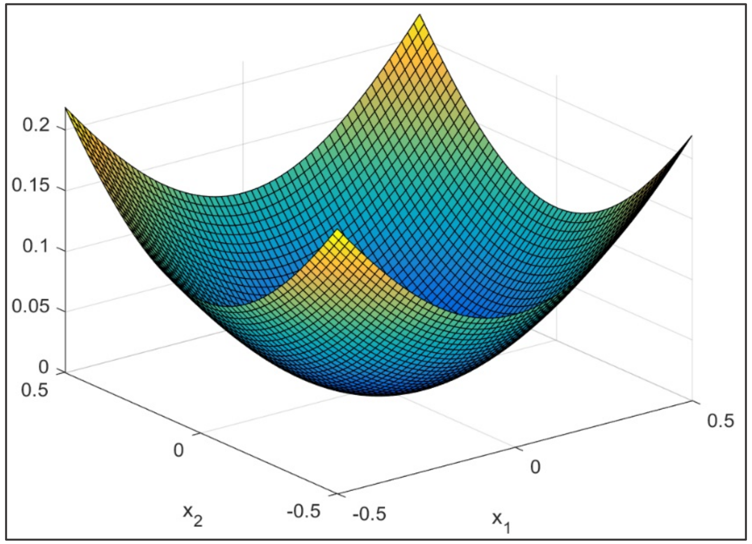
$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda) f(y)$$

for any $x, y \in B$ and $\lambda \in [0,1]$.

- If **B is full dimensional** (i.e., $l_i < u_i, i = 1, \dots, n$), this is equivalent to

$$\nabla^2 f(x) \succcurlyeq 0, \forall x \in B.$$





Complexity questions

Goal: study the complexity of testing convexity of a function over a box

- Restrict ourselves to **polynomial functions.**
- Related work:

Problem 6. N.Z. Shor proposed the question: Given a degree-4 polynomial of n variables, what is the complexity of determining whether this polynomial describes a convex function?

Theorem [Ahmadi, Olshevsky, Parrilo, Tsitsiklis]
It is strongly NP-hard to test (global) convexity of polynomials of degree 4.

NP-hardness of Deciding Convexity of Quartic Polynomials and Related Problems

Amir Ali Ahmadi, Alex Olshevsky, Pablo A. Parrilo, and John N. Tsitsiklis *†

- One may hope that adding the restriction to a box could make things easier.

Our theorem

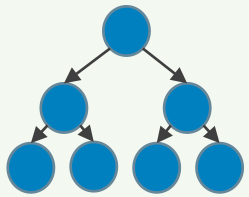
Theorem [Ahmadi, H.]

It is strongly NP-hard to test convexity of polynomials of degree 3 over a box.

Why are we interested in convexity over a box?

Detecting

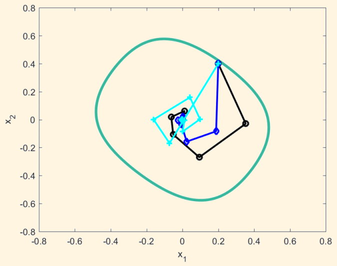
- Nonconvex optimization: **branch-and-bound**



- **Prior work:**
  - Sufficient conditions for convexity [Orban et al.], [Grant et al.]
  - In practice, BARON, CVX, Gurobi check convexity of quadratics and computationally tractable sufficient conditions for convexity

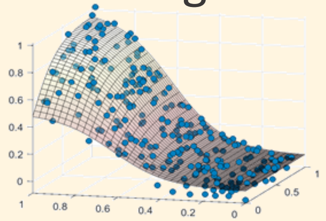
Imposing

- **Control theory:** convex Lyapunov functions



[Ahmadi and Jungers]
[Chesi and Hung]

- **Statistics:** convex regression



Proof of the theorem

Theorem [Ahmadi, H.]

It is strongly NP-hard to test convexity of polynomials of degree 3 over a box.

How to prove this?

In general:

Generic instance I of a known NP-hard problem → Construct J from I (**Reduction**) → Instance J of problem we are interested in

Question: What to do a reduction from?

Idea: A cubic polynomial f is convex over a (full-dimensional) box B if and only if $\nabla^2 f(x) \succcurlyeq 0, \forall x \in B$

$\nabla^2 f(x)$ is a matrix with entries **affine in $\boldsymbol{x}$**

Theorem [Nemirovski]:

Let $L(x)$ be a matrix with entries affine in x.

It is (weakly) NP-hard to test whether $L(x) \succcurlyeq 0$ for all x in a full-dimensional box B.

Are we done?

INSEAD
The Business School for the World®

No!

Issue 1: We want to show strong NP-hardness. Nemirovski's result shows weak NP-hardness.

Issue 2: Not every affine polynomial matrix is a valid Hessian!

Example: $L(x_1, x_2) = \begin{pmatrix} 10 & 2x_1 + 1 \\ 2x_1 + 1 & 10 \end{pmatrix}$. We have $\frac{\partial L_{11}(x)}{\partial x_2} \neq \frac{\partial L_{12}(x)}{\partial x_1}$.

Dealing with Issue 1 (1/3)

Reminder: weak vs strong NP-hardness

- Distinction only concerns problems where input is numerical
- **Max(I)**: largest number in magnitude that appears in the input of instance I (numerator or denominator)
- **Length(I)**: number of bits it takes to write down input of instance I

Strong	**Weak**
• There are instances I that are hard with Max(I)$\leq p$(Length(I)) (p is a polynomial)	• The instances that are hard may contain numbers of large magnitude (e.g., 2^n).
• No pseudo-polynomial algorithm possible	• Pseudo-polynomial algorithms possible
• Examples: SAT, MAX-CUT	• Examples: KNAPSACK, PARTITION

Dealing with Issue 1 (2/3)

Theorem [Nemirovski]: INTERVAL-PSDNESS

Let $L(x)$ be a matrix with entries affine in x.

It is (weakly) NP-hard to test whether $L(x) \succcurlyeq 0$ for all x in a full-dimensional box B.

Why weakly NP-hard?

PA ON:

Input: $a \in \mathbb{R}^n$ such that $\|a\|_2 \leq 0.1$

Test: does there exist $t \in \{-1,1\}^n$ such that $\sum_i a_i t_i = 0$?

REDUCTION

INTERVAL PSDNESS

Construct: $C = (I \ldots)^{-1}$,

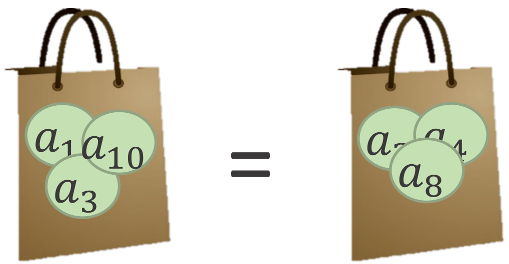
$\mu = n - d^{-2}(a)$, where $d(a) =$ smallest cd of a.

Take: $B = [-1,1]^n$ and $L(x) = \begin{pmatrix} C & x \\ x^T & \mu \end{pmatrix}$.

Test: Is $L(x) \succcurlyeq 0 \; \forall x \in B$?

Show: No to PARTITION ⇔ Yes to INTERVAL PSDNESS

Weakly NP-hard



Operation that can make the numbers in the instance blow up

Example: $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ \vdots & \ddots & \ddots & 0 \\ -1 & -1 & -1 & 1 \end{pmatrix}$ but one of the entries of A^{-1} is 2^{n-2}!

Dealing with Issue 1 (3/3)

INSEAD
The Business School for the World®

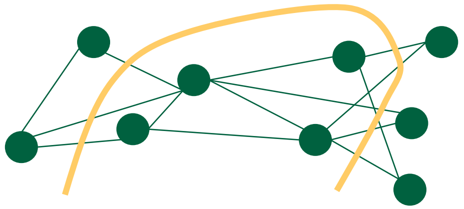
Theorem [Ahmadi, H.]: INTERVAL-PSDNESS
Let $L(x)$ be a matrix with entries affine in x.
It is **strongly** NP-hard to test whether $L(x) \succcurlyeq 0$ for all x in a full-dimensional box B.

MAX-CUT:

Input: simple graph G=(V,E) with $|V| = n$ and adj. matrix A, and a positive integer $k \leq n^2$

Test: does there exist a cut in the graph of size greater or equal to k?

Strongly NP-hard



REDUCTION

Preserves strong NP-hardness

INTERVAL PSDNESS

Construct: $\alpha = \frac{1}{(n+1)^3}$, $C = 4\alpha(I_n + \alpha A)$

$$\mu = \frac{n}{4\alpha} + k - 1 - \frac{1}{4} e^T A e$$

Take: $B = [-1,1]^n$ and $L(x) = \begin{pmatrix} C & x \\ x^T & \mu \end{pmatrix}$.

Test: Is $L(x) \succcurlyeq 0 \ \forall x \in B$?

Show: No to MAX-CUT $\Leftrightarrow$ Yes to INTERVAL PSDNESS

Taylor series of $4\alpha(I - \alpha A)^{-1}$ truncated at the first term

Scaling needed so that $(I_n - \alpha A)^{-1} \approx I_n + \alpha A$

Dealing with Issue 2

Theorem [Ahmadi, H.] CONV3BOX
It is strongly NP-hard to test convexity of polynomials of degree 3 over a box.

Proof: Reduction from INTERVAL PSDNESS

INTERVAL PSDNESS
Input: $L(x), \hat{B}$
Test: Is $L(x) \succcurlyeq 0, \ \forall x \in \hat{B}$?

Problem: How to construct a cubic polynomial f from $L(x)$?
Idea: Want $\nabla^2 f(x) = L(x)$.
Issue: Not all $L(x)$ are valid Hessians!

Key ideas for the construction of f:

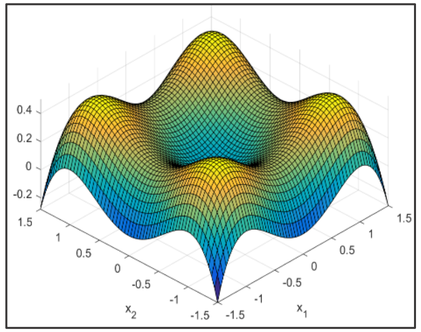
- Start with $f(x,y) = \frac{1}{2} y^T L(x) y$
- For $\nabla^2 f(x,y)$ to be able to be psd when $L(x) \succcurlyeq 0$, we need to have a nonzero diagonal: add $\frac{\alpha}{2} x^T x$ to $f(x,y)$.
- $L(x)$ and $H(y)$ do not depend on the same variable: what if $\exists (x,y)$ s.t. $L(x) = 0$ but $H(y)$ is not? The matrix cannot be psd: add $\frac{\eta}{2} y^T y$ to $f(x,y)$.

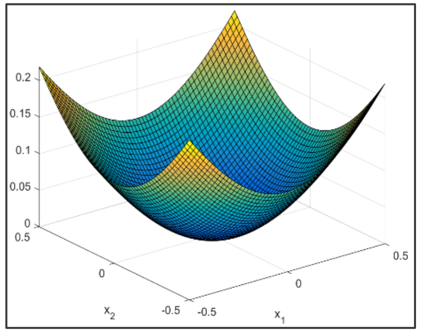
$$\nabla^2 f(x,y) = \begin{bmatrix} \alpha I_n & \frac{1}{2} H(y) \\ \frac{1}{2} H(y)^T & L(x) + \eta I_{n+1} \end{bmatrix}$$

$$\Rightarrow f(x) = \frac{1}{2} y^T L(x) y + \frac{\alpha}{2} x^T x + \frac{\eta}{2} y^T y, \qquad B = [-1,1]^{2n+1}$$

Summary

- Interested in **testing convexity of a polynomial over a box**.





- Showed that **strongly NP-hard** to test convexity of **cubics** over a box.

- Can be extended to give a **complete characterization** of the complexity of testing convexity of a polynomial (of any degree) over a box.

Thank you for listening

Questions?

Want to learn more?

https://scholar.princeton.edu/ghall

Dealing with Issue 1 (4/5)

In more detail: No to MAX-CUT ⇒ Yes to INTERVAL PSDNESS

No cut in G of size $\geq k$ ⇔ $[\max_{x\in\{-1,1\}^n} \frac{1}{4}\sum_{i,j} A_{ij}(1 - x_i x_j)] \leq k-1$

Size of largest cut in G

⇕

Convex

$[\max_{x\in\{-1,1\}^n} \frac{1}{4} x^T((n+1)^3 I_n - A)x] \leq \frac{n(n+1)^3}{4} - \frac{1}{4}e^T Ae + k - 1 := \mu$ ⇔ $[\max_{x\in\{-1,1\}^n} -\frac{1}{4}x^T Ax] \leq -\frac{1}{4}e^T Ae + k - 1$

⇕ $\alpha = (n+1)^3$

$[\max_{x\in[-1,1]^n} \frac{1}{4}x^T(\alpha I_n - A)x] \leq \mu$ ⇔ $\frac{1}{4}x^T(\alpha I_n - A)x \leq \mu, \forall x \in [-1,1]^n$

Approximation $C^{-1} \approx \left(\frac{1}{4}(\alpha I - A)\right)$ ⇓

$L(x) = \begin{pmatrix} C & x \\ x^T & \mu + \frac{1}{4} \end{pmatrix} \succcurlyeq 0, \forall x \in [-1,1]^n$ ⇐ $x^T C^{-1} x \leq \mu + \frac{1}{4}, \forall x \in [-1,1]^n$

Schur complement

Approximation error

Corollary

Completely classifies the complexity of testing convexity of a polynomial f of degree d over a box for any integer $d \geq 1$.

d

- $d = 1$: f is always convex
- $d = 2$: $\nabla^2 f(x)$ constant
- $d = 3$: Previous theorem (strongly NP-hard)
- $d = 4$ and above: Strongly NP-hard

Strongly NP-hard

Proof sketch:

- $g(x_1, \dots, x_n)$ =cubic polynomial for which testing convexity over a box B is hard
- $f(x_1, \dots, x_n, x_{n+1}) = g(x_1, \dots, x_n) + x_{n+1}^d$
- $\tilde{B} = B \times [0,1]$

We have $\nabla^2 f(x, x_{n+1}) = \begin{pmatrix} \nabla^2 g(x) & 0 \\ 0 & d(d-1)x_{n+1}^{d-2} \end{pmatrix}$

$\Rightarrow \nabla^2 f(x, x_{n+1}) \succcurlyeq 0$ on $\tilde{B} \Leftrightarrow \nabla^2 g(x) \succcurlyeq 0$ on B

Dealing with Issue 1 (5/5)

For converse: Yes to MAX-CUT ⇒ No to INTERVAL PSDNESS

There is a cut of size $\geq k$:

Let $\hat{x}_i = \begin{cases} 1 \text{ if node } i \text{ on one side of cut} \\ -1 \text{ if node } i \text{ on other side of cut} \end{cases}$ ⇒ Similar steps to previously ⇒ $\hat{x}^T C^{-1} \hat{x} \geq \mu + \frac{3}{4} > \mu + \frac{1}{4}$

⇓

$\exists\, \hat{x} \in [-1,1]^n$ s.t. $L(\hat{x}) \nsucceq 0$

Corollary [Ahmadi, H.]: Let n be an integer and let $\hat{q}_{ij}, \bar{q}_{ij}$ be rational numbers with $\hat{q}_{ij} \leq \bar{q}_{ij}$ and $\hat{q}_{ij} = \hat{q}_{ji}$ and $\bar{q}_{ij} = \bar{q}_{ji}$ for all $i = 1, \ldots, n$ and $j = 1, \ldots, n$.
It is **strongly** NP-hard to test whether all symmetric matrices with entries in $[\hat{q}_{ij}; \bar{q}_{ij}]$ are positive semidefinite.

- Initial problem studied by Nemirovski
- Of independent interest in robust control

Dealing with Issue 2 (2/3)

INSEAD
The Business School for the World®

Show NO to INTERVAL PSDNESS ⇒ NO to CONV3BOX.

This is equivalent to:

$$\exists \bar{x} \in [-1,1]^n \text{ s.t. } L(\bar{x}) \not\succcurlyeq 0 \Rightarrow \exists\, \hat{x}, \hat{y} \in [-1,1]^{2n+1}, z \text{ s.t. } z^T \nabla^2 f(\hat{x}, \hat{y}) z < 0$$

Need to leverage extra structure of $L(x)$: $L(x) = \begin{pmatrix} C & x \\ x^T & \mu + \frac{1}{4} \end{pmatrix}$

Candidates: $\hat{x} = \bar{x}, \quad \hat{y} = 0, \quad \boldsymbol{z} = \begin{pmatrix} \boldsymbol{0} \\ -\boldsymbol{C}^{-1}\bar{\boldsymbol{x}} \\ \boldsymbol{1} \end{pmatrix}$

$$z^T \nabla^2 f(\hat{x}, \hat{y}) z = \begin{pmatrix} 0 \\ -C^{-1}\bar{x} \\ 1 \end{pmatrix}^T \begin{pmatrix} \alpha I_n & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C} + \eta I_n & \bar{\boldsymbol{x}} \\ \boldsymbol{0} & \bar{\boldsymbol{x}}^T & \boldsymbol{\mu} + \frac{\boldsymbol{1}}{\boldsymbol{4}} + \eta \end{pmatrix} \begin{pmatrix} 0 \\ -C^{-1}\bar{x} \\ 1 \end{pmatrix} = \mu + \frac{1}{4} - \bar{x}^T C^{-1} \bar{x} + \eta(1 + \|C^{-1}\bar{x}\|_2^2)$$

< 0 as $L(\bar{x}) \not\succcurlyeq 0$

Appropriately scaled so that $z^T \nabla^2 f(\hat{x}, \hat{y}) z$ remains <0.

Dealing with Issue 2 (3/3)

INSEAD

The Business School for the World®

Show YES to INTERVAL PSDNESS ⇒ YES to CONV3BOX.

This is equivalent to:

$$L(x) \succcurlyeq 0 \ \forall x \in [-1,1]^n \Rightarrow \nabla^2 f(x,y) = \begin{bmatrix} \alpha I_n & \frac{1}{2} H(y) \\ \frac{1}{2} H(y)^T & L(x) + \eta I_{n+1} \end{bmatrix} \succcurlyeq 0, \forall (x,y) \in [-1,1]^{2n+1}$$

But…

$$\nabla^2 f(x,y) \succcurlyeq 0, \forall (x,y) \in [-1,1]^{2n+1} \quad \overset{\text{Schur}}{\Longleftrightarrow} \quad \boldsymbol{L(x)} + \boldsymbol{\eta I_{n+1}} - \frac{\mathbf{1}}{\mathbf{4\alpha}} \boldsymbol{H(y)^T H(y)} \succcurlyeq 0, \ \forall (x,y) \in [-1,1]^{2n+1}$$

$\succcurlyeq \mathbf{0}$ $\forall \boldsymbol{x} \in [\mathbf{-1}, \mathbf{1}]^{\boldsymbol{n}}$ **(Assumption)**

$\boldsymbol{\alpha}$ **chosen large enough so that** $\succcurlyeq \mathbf{0} \ \forall \boldsymbol{y} \in [\mathbf{-1}, \mathbf{1}]^{\boldsymbol{n+1}}$