On Local Minima of Cubic Polynomials

Amir Ali Ahmadi
Princeton, ORFE
Affiliated member of PACM, COS, MAE, CSML

Joint work with

Jeffrey Zhang
Princeton, ORFE

CRM-DIMACS Workshop on Mixed-Integer Nonlinear Programming
October 2019, Montreal

PRINCETON UNIVERSITY ORFE

Deciding local minimality

Consider the optimization problem

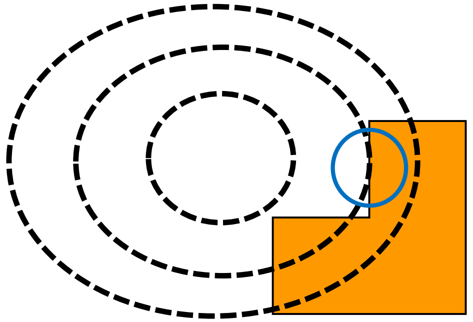
$$\min_{x \in \mathbb{R}^n} f(x)$$
$$x \in \Omega$$

Given a point $\bar{x}$, decide if it is a local minimum.

Why local minima?

- Global minima are often intractable
- Recent interest in local minima, particularly in machine learning applications
- Existing notions that local minima are "easier to find" or are sufficient for applications
- Formal understanding of local minima is desirable

Local minima



A point $\bar{x}$ is a **local minimum** of

$$\min_{x \in \mathbb{R}^n} f(x)$$
$$x \in \Omega$$

if there exists a ball of radius $\epsilon > 0$ such that $p(\bar{x}) \leq p(x)$ for all $x \in B_\epsilon(\bar{x}) \cap \Omega$.

$\bar{x}$ is a **strict local minimum** if $p(\bar{x}) < p(x)$ for all $x \in B_\epsilon(\bar{x}) \cap \Omega \backslash \bar{x}$.

Our focus: polynomial optimization problems

f is a polynomial, Ω is defined by polynomial inequalities.

$$\min_{x \in \mathbb{R}^n} p(x)$$
$$q_i(x) \geq 0, i = 1, \ldots, m$$

PRINCETON UNIVERSITY ORFE

Known tractable cases

Unconstrained quadratic optimization

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} x^T Q x + c^T x$$

$\bar{x}$ is a local minimum if and only if

$$Q\bar{x} + c = 0$$
$$Q \succcurlyeq 0$$

$\bar{x}$ is a strict local minimum if and only if

$$Q\bar{x} + b = 0$$
$$Q \succ 0$$

Compute n leading principal minors

Check coefficients of characteristic polynomial

Linear Programming

$$\min_{x \in \mathbb{R}^n} c^T x$$
$$Ax = b$$
$$x \geq 0$$

$\bar{x}$ is a local minimum if and only if it is optimal.

$\bar{x}$ is a strict local minimum if and only it is the unique optimal solution.

$A\bar{x} = b, \bar{x} \geq 0$, and $c^T\bar{x}$ is attainable in the dual

Check if $\bar{x}$ is optimal. If it is, add $c^T x = c^T \bar{x}$ as a constraint, and solve sequence of LPs

Known intractable cases

Unconstrained quartic optimization	Quadratic programming
$\min_{x \in \mathbb{R}^n} p(x)$ p is a quartic polynomial	$\min_{x \in \mathbb{R}^n} \frac{1}{2} x^T Q x + c^T x$ $Ax \geq b$

A matrix M is copositive if $x^T M x \geq 0, \forall x \geq 0$

A matrix M is copositive if and only if 0 is a local minimum of

$$\begin{bmatrix} x_1^2 \\ \dots \\ x_n^2 \end{bmatrix}^T M \begin{bmatrix} x_1^2 \\ \dots \\ x_n^2 \end{bmatrix}^T$$

or of

$$\min_{x \in \mathbb{R}^n} x^T M x$$
$$x \geq 0$$

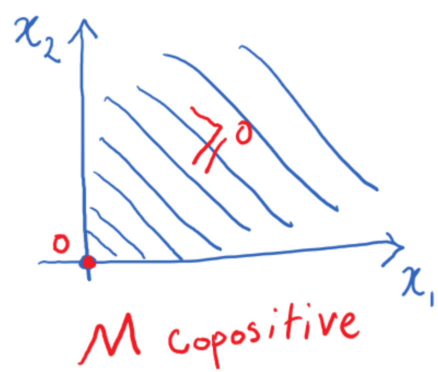
SOME NP-COMPLETE PROBLEMS IN QUADRATIC AND NONLINEAR PROGRAMMING

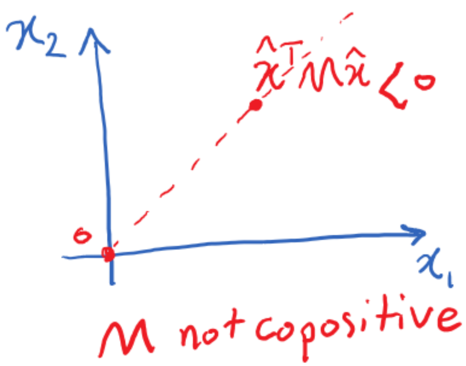
Katta G. MURTY*

Department of Industrial and Operations Engineering, The University of Michigan, 1205 Beal Avenue, Ann Arbor, MI 48109-2117, USA

Santosh N. KABADI**

Faculty of Administration, University of New Brunswick, Fredericton, NB, Canada E3B 5A6





PRINCETON UNIVERSITY ORFE

Strict local minimality also NP-hard.

Summary of prior literature

Unconstrained quadratic optimization	Linear Programming
$\min_{x \in \mathbb{R}^n} \frac{1}{2} x^T Q x + c^T x$	$\min_{x \in \mathbb{R}^n} c^T x$ $Ax = b$ $x \geq 0$

Poly-time (both for local min and strict local min)

Unconstrained quartic optimization	Quadratic programming
$\min_{x \in \mathbb{R}^n} p(x)$ p is a quartic polynomial	$\min_{x \in \mathbb{R}^n} \frac{1}{2} x^T Q x + c^T x$ $Ax \geq b$

NP-hard (both for local min and strict local min)

Open cases?

Unconstrained
cubic minimization

Omar Khayyam (1048-1131)

PRINCETON UNIVERSITY ORFE

Outline

- **Part I: Testing local minimality of a given point for a cubic polynomial**

- **Part 2:** Finding a local minimum of a cubic polynomial

Classical optimality conditions

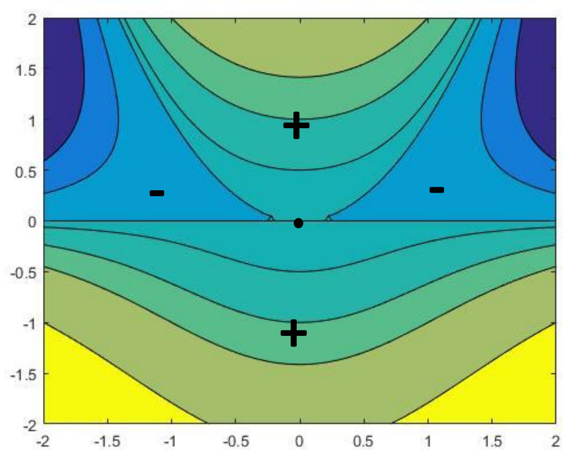
First Order Necessary Condition (FONC)	Second Order Necessary Condition (SONC)
$\bar{x}$ is a local minimum $\Rightarrow \nabla p(\bar{x}) = 0$	$\bar{x}$ is a local minimum $\Rightarrow \nabla^2 p(\bar{x}) \succcurlyeq 0$

Second Order Sufficient Condition (SOSC):
FONC + $\nabla^2 p(\bar{x}) \succ 0 \Rightarrow \bar{x}$ is a (strict) local minimum

$\bar{x}$ **is a local minimum** $\Rightarrow$ **no descent directions at** $\bar{x}$
A direction d is a descent direction for p at $\bar{x}$ if for some $\alpha^* > 0$, $p(\bar{x} + \alpha d) < p(\bar{x})$ for all $\alpha \in (0, \alpha^*)$

Unlike quadratics, not sufficient for cubic polynomials

$$p(x_1, x_2) = x_2^2 - x_1^2 x_2$$



Necessary and sufficient condition for local minima

Theorem (Third Order Condition, TOC)

Let p be a cubic polynomial and suppose $\bar{x}$ satisfies FONC and SONC. Then $\bar{x}$ is a local minimum of p if and only if

$$d \in N(\nabla^2 p(\bar{x})) \Rightarrow \nabla p_3(d) = 0$$

Moreover, this condition can be checked in polynomial time.

$N(\nabla^2 p(\bar{x}))$ is the null space of Hessian at $\bar{x}$

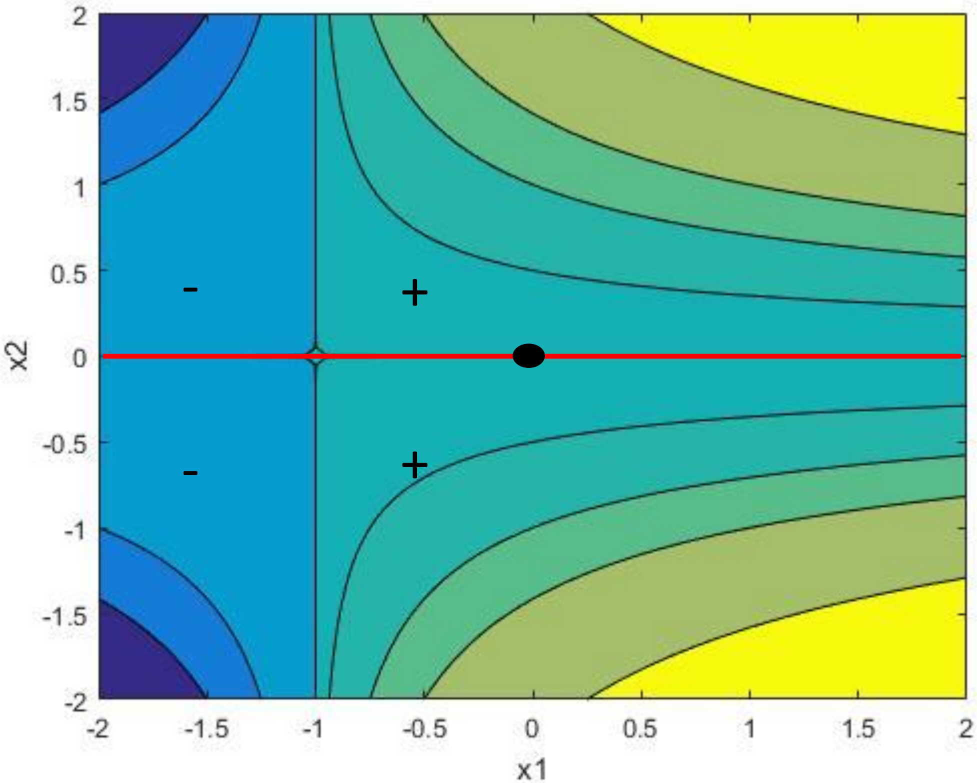
p_3 is the cubic component of p

Example: origin a local minimum

$$p(x_1, x_2) = x_2^2 + x_1 x_2^2$$

$$\nabla p(0,0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \nabla^2 p(0,0) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$$

$$d \in N(\nabla^2 p(\bar{x})) \Rightarrow \nabla p_3(d) = 0?$$



$$\nabla p_3(x_1, x_2) = \begin{bmatrix} 2x_2^2 \\ 2x_1 x_2 \end{bmatrix}$$

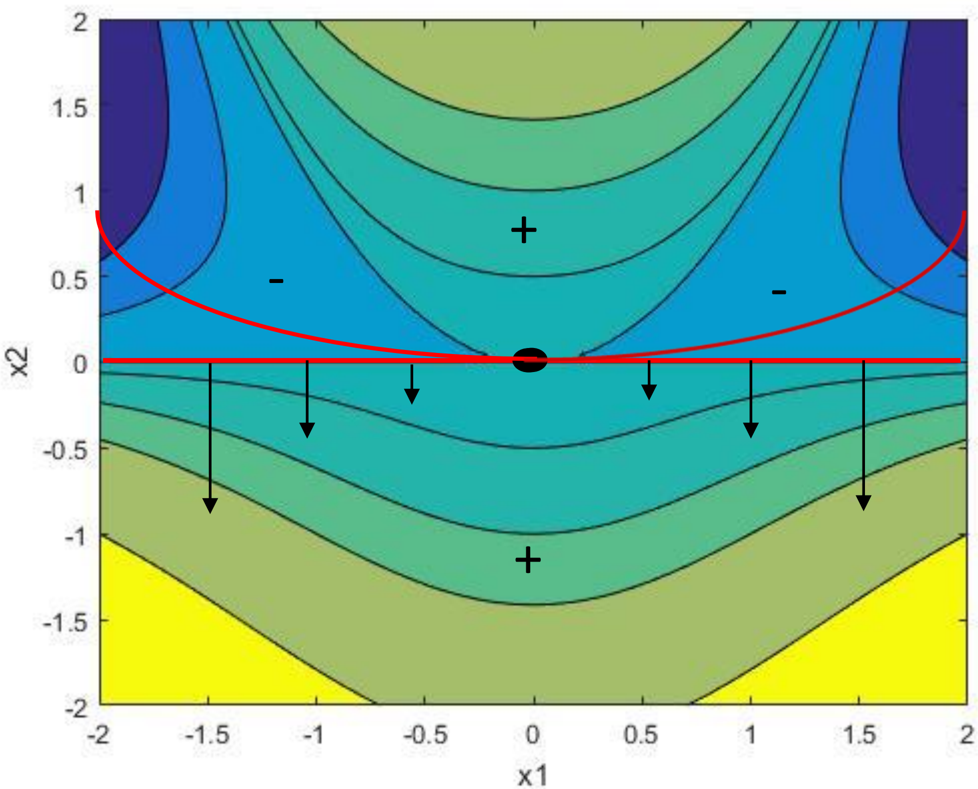
$$\nabla p_3(\alpha, 0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Example: origin not a local minimum

$$p(x_1, x_2) = x_2^2 - x_1^2 x_2$$

$$\nabla p(0,0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \nabla^2 p(0,0) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$$

$$d \in N(\nabla^2 p(\bar{x})) \Rightarrow \nabla p_3(d) = 0?$$



$$\nabla p_3(x_1, x_2) = \begin{bmatrix} -2x_1 x_2 \\ -x_1^2 \end{bmatrix}$$

$$\nabla p_3(\alpha, 0) = \begin{bmatrix} 0 \\ -\alpha^2 \end{bmatrix}$$

A more natural condition?

$$d \in N(\nabla^2 p(\bar{x})) = 0 \Rightarrow p_3(d) = 0$$

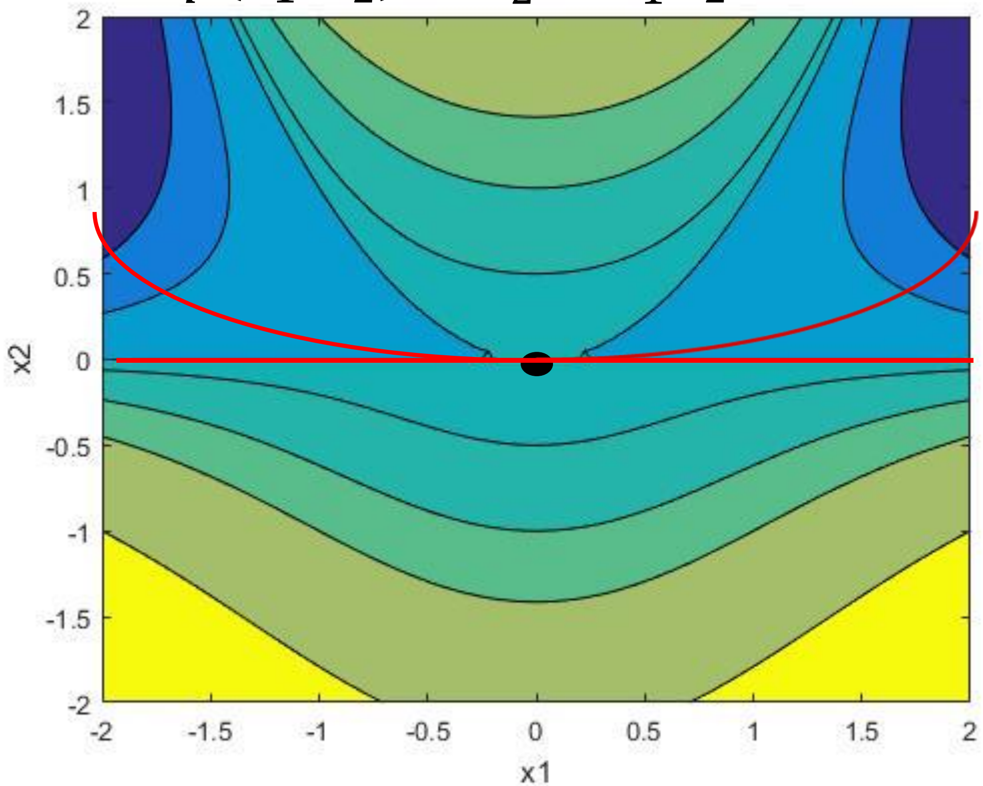
$\nabla p_3(d) = 0$ → Euler's identity:

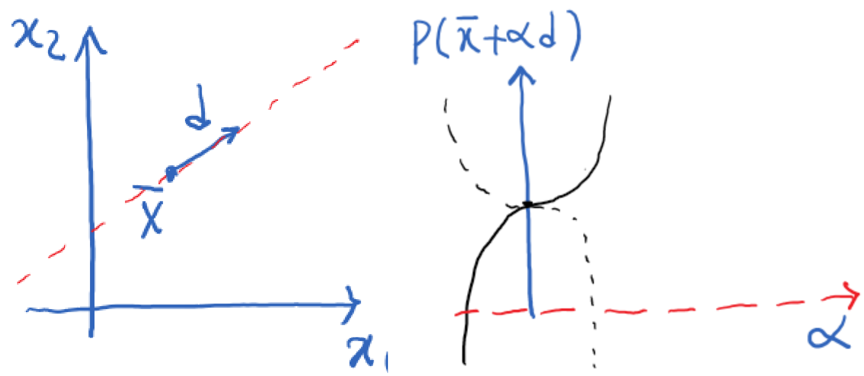
$p_3(d) = \frac{1}{3} d^T \nabla p_3(d)$

Necessary for C^3 functions (where p_3 would be the cubic component of the Taylor expansion).

"Third Order Necessary Condition" (TONC)

$$p(x_1, x_2) = x_2^2 - x_1^2 x_2$$





Not sufficient for local optimality, even for cubics

Guarantees no descent directions for cubic polynomials

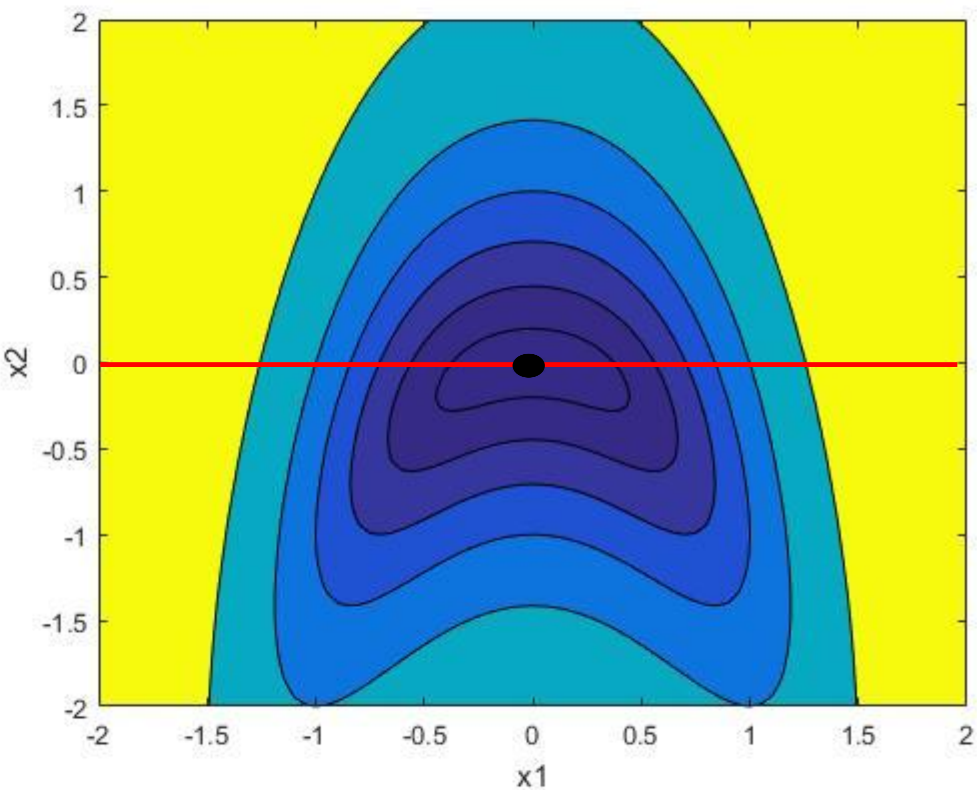
Does not guarantee no parabolas of descent

Is TOC necessary for general functions?

$$d \in N(\nabla^2 p(\bar{x})) \Rightarrow \nabla p_3(d) = 0?$$

Easy to see not sufficient for higher degree polynomials (e.g., $p(x) = x^5$), but is it necessary? No!

$$p(x_1, x_2) = x_1^4 + (x_1^2 + x_2)^2$$



$$\nabla p(0,0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\nabla^2 p(0,0) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$$

$$\nabla p_3 = \begin{pmatrix} 2x_2^2 \\ 4x_1 x_2 \end{pmatrix}$$

$$\nabla p_3(0,1) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

A characterization of local minima for cubics

Theorem (Third Order Condition, TOC)

Let p be a cubic polynomial and suppose $\bar{x}$ satisfies FONC and SONC. Then $\bar{x}$ is a local minimum of p if and only if

$$d \in N(\nabla^2 p(\bar{x})) \Rightarrow \nabla p_3(d) = 0$$

Moreover, this condition can be checked in polynomial time.

Proof of characterization of local minima (1/3)

Taylor expansion of cubic polynomials

$$p(x+\lambda v) = p(x) + \lambda \nabla p(x)^T v + \frac{1}{2}\lambda^2 v^T \nabla^2 p(x) v + \lambda^3 p_3(v)$$

Suppose $\bar{x}$ satisfies FONC, SONC. We show $\bar{x}$ is a local min iff TOC.

For any unit vectors d in the null space of $\nabla^2 p(\bar{x})$ and z in the range of $\nabla^2 p(\bar{x})$,

$$p(\bar{x} + \alpha d + \beta z) = p(\bar{x}) + \nabla p(\bar{x})^T(\alpha d + \beta z) \longrightarrow 0$$

$$+\frac{1}{2}(\alpha d + \beta z)^T \nabla^2 p(\bar{x})(\alpha d + \beta z) \longrightarrow \beta^2 z^T \nabla^2 p(\bar{x}) z$$

$$+p_3(\alpha d + \beta z)$$

$$p_3(\alpha d + \beta z) = p_3(\alpha d) + \beta \nabla p_3(\alpha d)^T z + \frac{1}{2}\beta^2 z^T \nabla^2 p_3(\alpha d) z + \beta^3 p_3(z)$$

($0 \longleftarrow p_3(\alpha d)$)

$$p(\bar{x} + \alpha d + \beta z) - p(\bar{x}) =$$

$$\frac{1}{2}\beta^2 z^T \nabla^2 p(\bar{x}) z + \alpha^2 \beta \nabla p_3(d)^T z + \frac{1}{2}\alpha\beta^2 z^T \nabla^2 p_3(d) z + \beta^3 p_3(z)$$

Proof of characterization (2/3) [Sufficiency]

$$p(\bar{x} + \alpha d + \beta z) - p(\bar{x}) =$$

$$\frac{1}{2}\beta^2 z^T \nabla^2 p(\bar{x}) z + \alpha^2 \beta \nabla p_3(d)^T z + \frac{1}{2}\alpha\beta^2 z^T \nabla^2 p_3(d) z + \beta^3 p_3(z)$$

$\nabla p_3(d) = 0 \; \forall \; d \in N(\nabla^2 p(\bar{x})) \Rightarrow$ local minimum

$$\frac{1}{2}\beta^2 z^T \nabla^2 p(\bar{x}) z + \frac{1}{2}\alpha\beta^2 z^T \nabla^2 p_3(d) z + \beta^3 p_3(z)$$

$$\beta^2 \left(\frac{1}{2} z^T \nabla^2 p(\bar{x}) z + \frac{1}{2}\alpha z^T \nabla^2 p_3(d) z + \beta p_3(z)\right)$$

$\geq$ smallest nonzero eigenvalue

Upper bounded in abs. value

Therefore, $\exists \alpha^*, \beta^*$ such that if $\alpha < \alpha^*, \beta < \beta^*$, this expression is nonnegative

Proof of characterization (3/3) [Necessity]

$$p(\bar{x} + \alpha d + \beta z) - p(\bar{x}) =$$

$$\frac{1}{2}\beta^2 z^T \nabla^2 p(\bar{x}) z + \alpha^2 \beta \nabla p_3(d)^T z + \frac{1}{2}\alpha\beta^2 z^T \nabla^2 p_3(d) z + \beta^3 p_3(z)$$

local minimum $\Rightarrow \nabla p_3(d) = 0 \; \forall d \in N\left(\nabla^2 p(\bar{x})\right)$

Otherwise, for the sake of contradiction pick $\hat{d} \in N\left(\nabla^2 p(\bar{x})\right)$ such that $\nabla p_3(\hat{d}) \neq 0$ and pick $\hat{z} = -\nabla p_3(\hat{d})$

Pick a sequence $\beta_i \to 0, \alpha_i \propto \beta_i^{\frac{1}{3}}$

$$\frac{1}{2}\beta^2 \hat{z}^T \nabla^2 p(\bar{x}) \hat{z} + \alpha^2 \beta \nabla p_3\left(\hat{d}\right)^T \hat{z} + \frac{1}{2}\alpha\beta^2 \hat{z}^T \nabla^2 p_3\left(\hat{d}\right) \hat{z} + \beta^3 p_3(\hat{z})$$

$$\frac{1}{2}\beta_i^2 \hat{z}^T \nabla^2 p(\bar{x}) \hat{z} + \beta_i^{5/3} \nabla p_3(\hat{d})^T \hat{z} + \frac{1}{2}\beta_i^{7/3} \hat{z}^T \nabla^2 p_3(\hat{d}) \hat{z} + \beta_i^3 p_3(\hat{z})$$

< 0

Characterization of strict local minima

Proposition

$\bar{x}$ is a strict local minimum of a cubic polynomial p if and only if

$$\nabla p(\bar{x}) = 0 \text{ (FONC)}$$

$$\nabla^2 p(\bar{x}) \succ 0 \text{ (SOSC)}$$

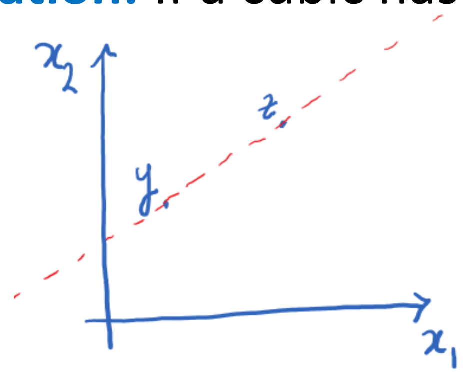
(Note: SOSC is not necessary in general: $p(x) = x^4$.)

Proof. Only need to show $\bar{x}$ strict local min $\Rightarrow \nabla^2 p(\bar{x}) \succ 0$.
Otherwise for the sake of contradiction pick $d \in N\big(\nabla^2 p(\bar{x})\big)$.

$$p(\bar{x} + \alpha d) = p(\bar{x}) + \alpha \nabla p(\bar{x})^T d + \frac{1}{2}\alpha^2 d^T \nabla^2 p(\bar{x}) d + \alpha^3 p_3(d).$$

Observation: If a cubic has a strict local min, then that is the **unique** local min.

Proof.



Checking local minimality in polynomial time

- Input: $p, \bar{x}$
- Compute gradient and Hessian of p at $\bar{x}$
- Check FONC and SONC
- Compute a basis $\{v_1, v_2, \dots, v_k\}$ for null space of $\nabla^2 p(\bar{x})$ (solving linear systems)

$$\nabla^2 p(\bar{x}) v_1 = 0, \quad v_1(i) = 0 \quad \Big| \quad \nabla^2 p(\bar{x}) v_2 = 0, \quad v_1^T v_2 = 0, \quad v_2(i) = 0 \quad \Big| \quad \cdots$$

- Compute gradient of p_3, evaluated on the null space of $\nabla^2 p(\bar{x})$

$$\nabla p_3(\bar{x}) = \begin{pmatrix} \frac{\partial p_3}{\partial x_1}(\bar{x}) \\ \cdots \\ \frac{\partial p_3}{\partial x_n}(\bar{x}) \end{pmatrix} \Rightarrow \begin{pmatrix} \frac{\partial p_3}{\partial x_1}(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k) \\ \cdots \\ \frac{\partial p_3}{\partial x_n}(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k) \end{pmatrix} \longleftarrow \begin{pmatrix} g_1(\alpha_1, \dots, a_k) \\ \cdots \\ g_n(\alpha_1, \dots, a_k) \end{pmatrix}$$

- All coefficients of all g_i must be zero

For strict local minima, check FONC and SOSC (leading n principal minors must be positive)

Outline

- **Part I:** Testing local minimality of a given point for a cubic polynomial

- **Part 2: Finding a local minimum of a cubic polynomial**

Finding local minima

Given a cubic polynomial p, can we efficiently find a local minimum of p?

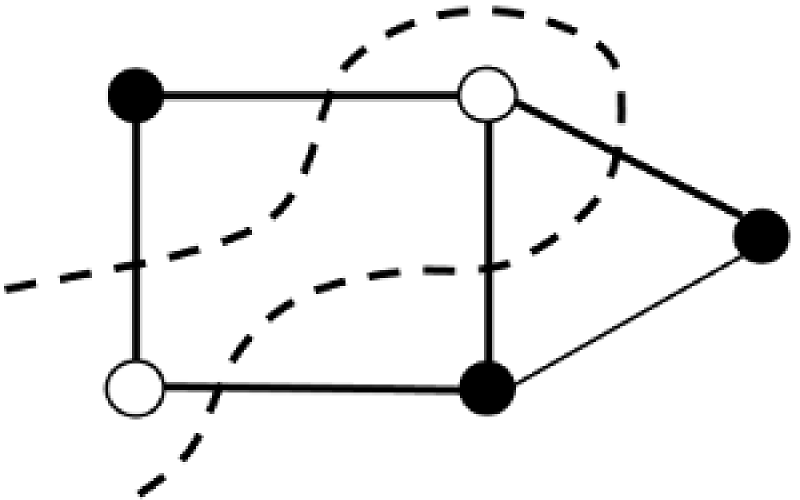
Let's start with a "simpler" question. Can we efficiently find a critical point of p?

Unfortunately…

Theorem
Deciding if a cubic polynomial has a critical point is strongly NP-hard.

Reduction from MAXCUT

Given a graph $G = (V, E)$, partition the vertices into two sets such that as many edges as possible are between vertices in opposite sets



MAXCUT (decision version)

Is there a cut of size k?

⇩

Quadratic satisfiability

$$\frac{1}{4}\sum_{(i,j)\in E}(1 - x_i x_j) = k \qquad 1 - x_i^2 = 0, i = 1, \dots, n$$

⇩

Critical points of a cubic polynomial

$$p(x_1, \dots, x_n, y_0, y_1, \dots, y_n) = y_o\left(\frac{1}{4}\sum_{(i,j)\in E}(1 - x_i x_j) - k\right) + \sum_{i=1}^{n} y_i(1 - x_i^2)$$

Critical points of a cubic polynomial

$$p(x,y) = y_o\left(\frac{1}{4}\sum_{(i,j)\in E}(1 - x_i x_j) - k\right) + \sum_{i=1}^{n} y_i(1 - x_i^2)$$

$$\nabla p(x,y) = \begin{bmatrix} \frac{dp}{dx_i} \\ \frac{dp}{dy_0} \\ \frac{dp}{dy_i} \end{bmatrix} = \begin{bmatrix} -\frac{y_0}{4}\left(\sum_{(i,j)\in E} x_j\right) - 2x_i y_i \\ \frac{1}{4}\sum_{(i,j)\in E}(1 - x_i x_j) - k \\ 1 - x_i^2 \end{bmatrix}$$

Any cut of size k $\Rightarrow$ critical point $(x = \text{cut}, y = 0)$

Any critical point $\Rightarrow$ cut of size k $(x \Rightarrow \text{cut})$

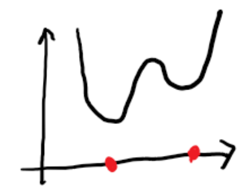
But this doesn't necessarily mean finding local minima is NP-hard.
First some geometry…

Some geometric properties of local minima

Theorem

The local minima of any cubic polynomial p form a convex set.

Not true for quartics:



Lemma

If $\bar{x}$ is a local minimum of a cubic polynomial p and $d \in N(\nabla^2 p(\bar{x}))$, then for any α,

$$\nabla p(\bar{x}+\alpha d) = 0$$

Proof (of theorem).

Let x and y be local minima. Note p is constant on the line between x and y.

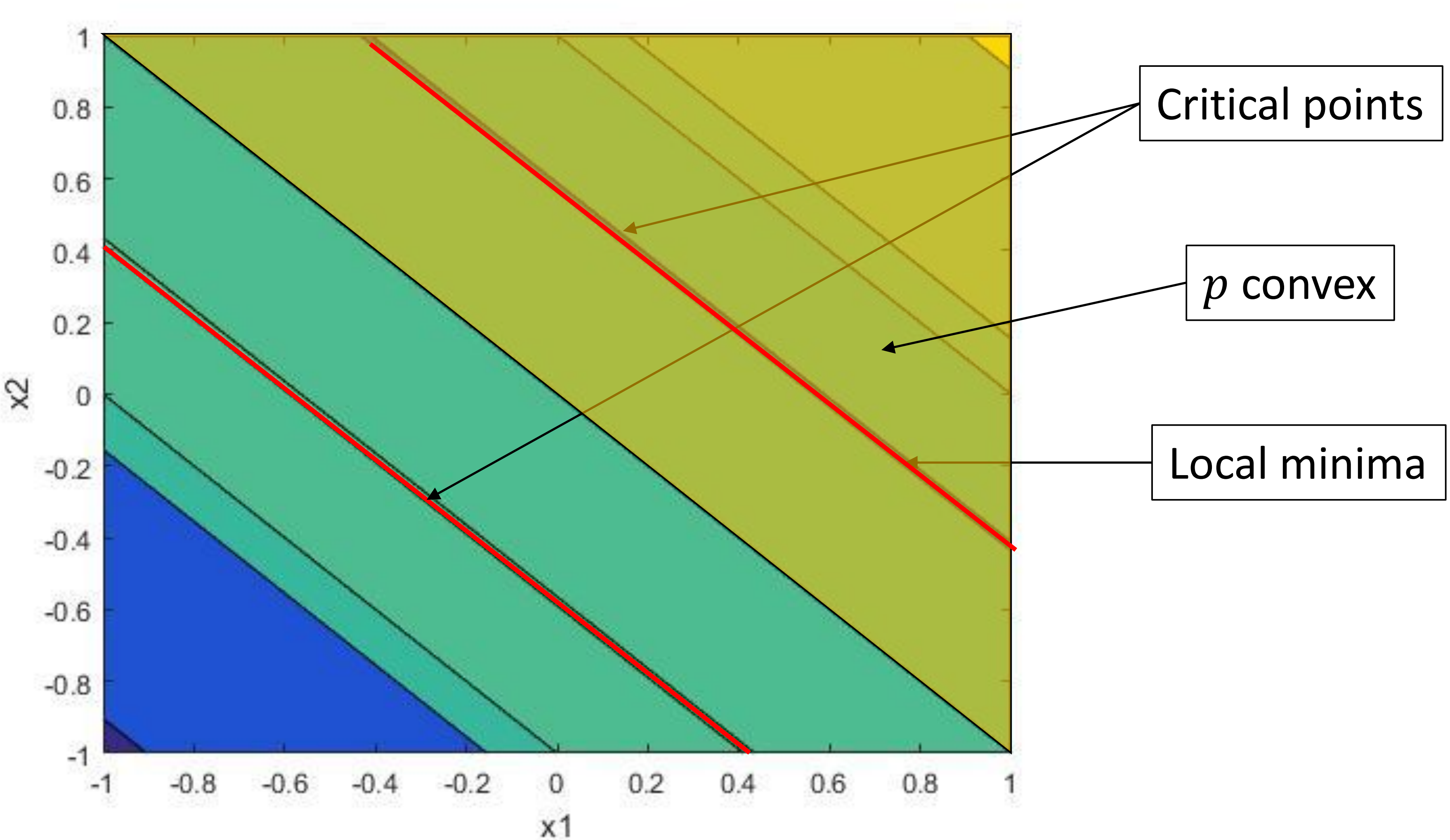
SD

$$= N((1-\alpha)\nabla^2 p(x) + \alpha \nabla^2 p(y))$$
$$= N\left(\nabla^2 p(x)\right) \cap N\left(\nabla^2 p(y)\right) \subseteq N\left(\nabla^2 p(x)\right)$$

PRINCETON UNIVERSITY ORFE

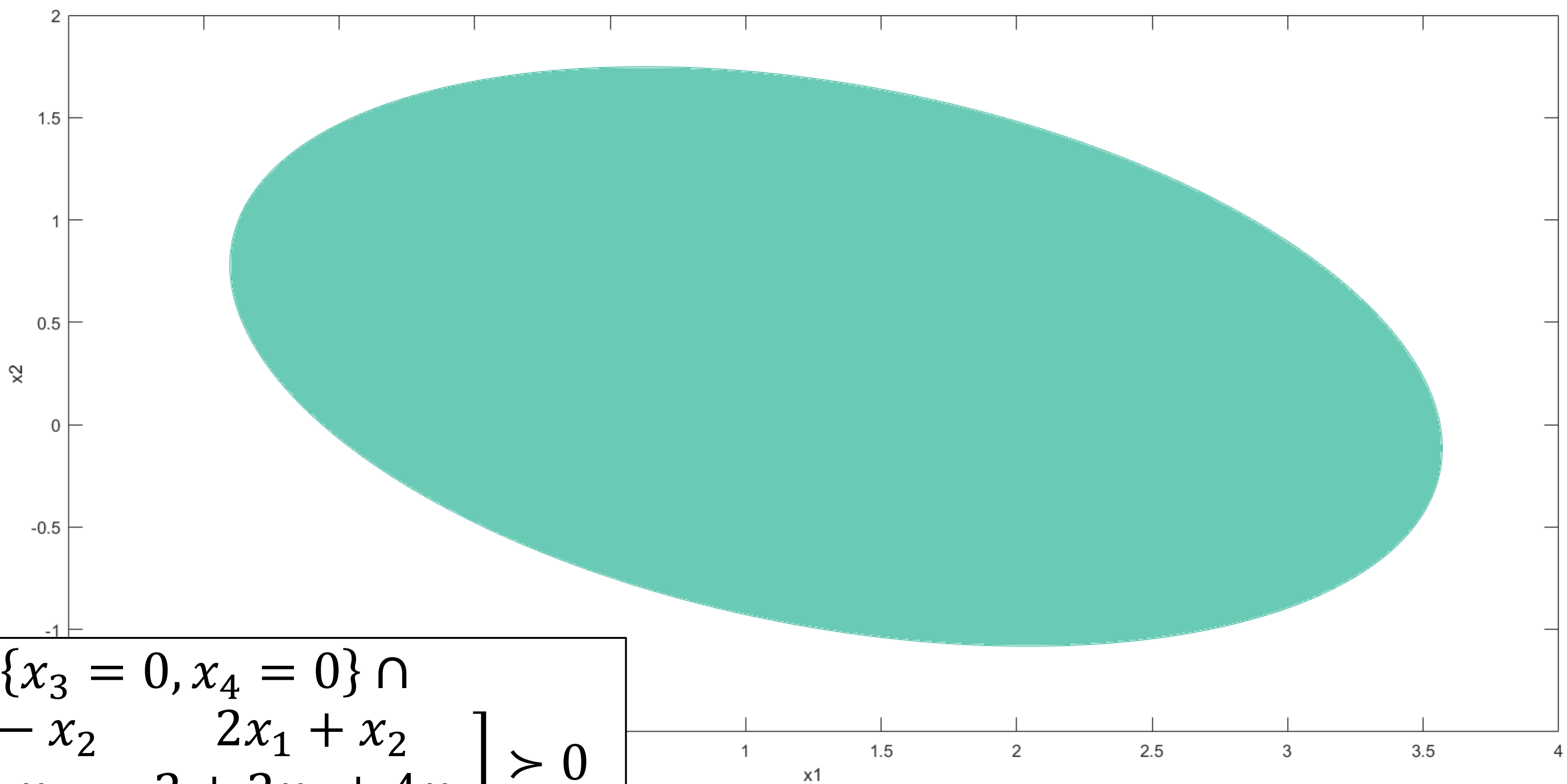
Convexity of the set of local minima

$$p(x_1, x_2) = x_1^3 + 3x_1^2x_2 + 3x_1x_2^2 + x_2^3 - x_1 - x_2$$



Set of local minima not necessarily polyhedral

$$p(x_1, x_2, x_3, x_4) = \frac{1}{2} x_1^2 x_3^2 + 2x_1 x_3 x_4 + \frac{1}{2} x_1 x_4^2 - \frac{1}{2} x_2 x_3^2 + x_2 x_3 x_4 + 2x_2 x_4^2 + x_3^2 + x_4^2$$



$$\{x_3 = 0, x_4 = 0\} \cap \begin{bmatrix} 2 + x_1 - x_2 & 2x_1 + x_2 \\ 2x_1 + x_2 & 2 + 2x_1 + 4x_2 \end{bmatrix} \succ 0$$

PRINCETON UNIVERSITY ORFE

Convexity region

Definition (Convexity region)

The convexity region of a polynomial p is the set

$$\{x \in \mathbb{R}^n \mid \nabla^2 p(x) \succcurlyeq 0\}$$

- The convexity region of a cubic polynomial is a spectrahedron (we call it a **"CH-spectrahedron"**)
- Not every spectrahedron is a CH-spectrahedron: e.g., $\begin{bmatrix} x_1+x_2 & 1 \\ 1 & x_2 \end{bmatrix} \succcurlyeq 0$

$A_0 + \sum_{i=1}^n x_i A_i \succcurlyeq 0$ gives a CH-spectrahedron

$$\Updownarrow$$

$$\frac{1}{6}\nabla^2\left[x^T\left(\sum_{i=1}^n x_i A_i\right)x\right] = \sum_{i=1}^n x_i A_i.$$

- Any spectrahedron $\{x \in \mathbb{R}^n \mid \sum_{i=1}^n A_i x_i + Q \succcurlyeq 0\}$ with A_i in $\mathbb{R}^{m\times m}$ is the shadow of a CH-spectrahedron in dimension $n+m$.

A "convex" optimization problem

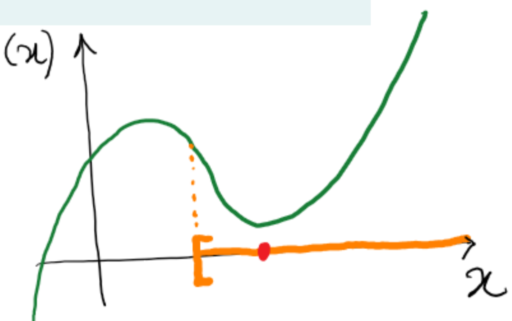
Theorem

If a cubic polynomial p *has a local minimum, the solution set of the following optimization problem is the closure of its local minima.*

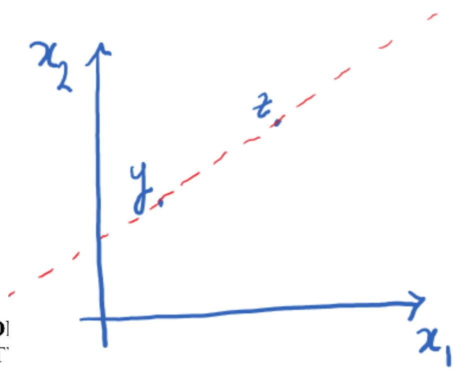
$$\min_{x} p(x)$$
$$\nabla^2 p(x) \succcurlyeq 0$$

In particular, the optimal value of this "convex" problem gives the value of p *at any local minimum.*

Very rough intuition:



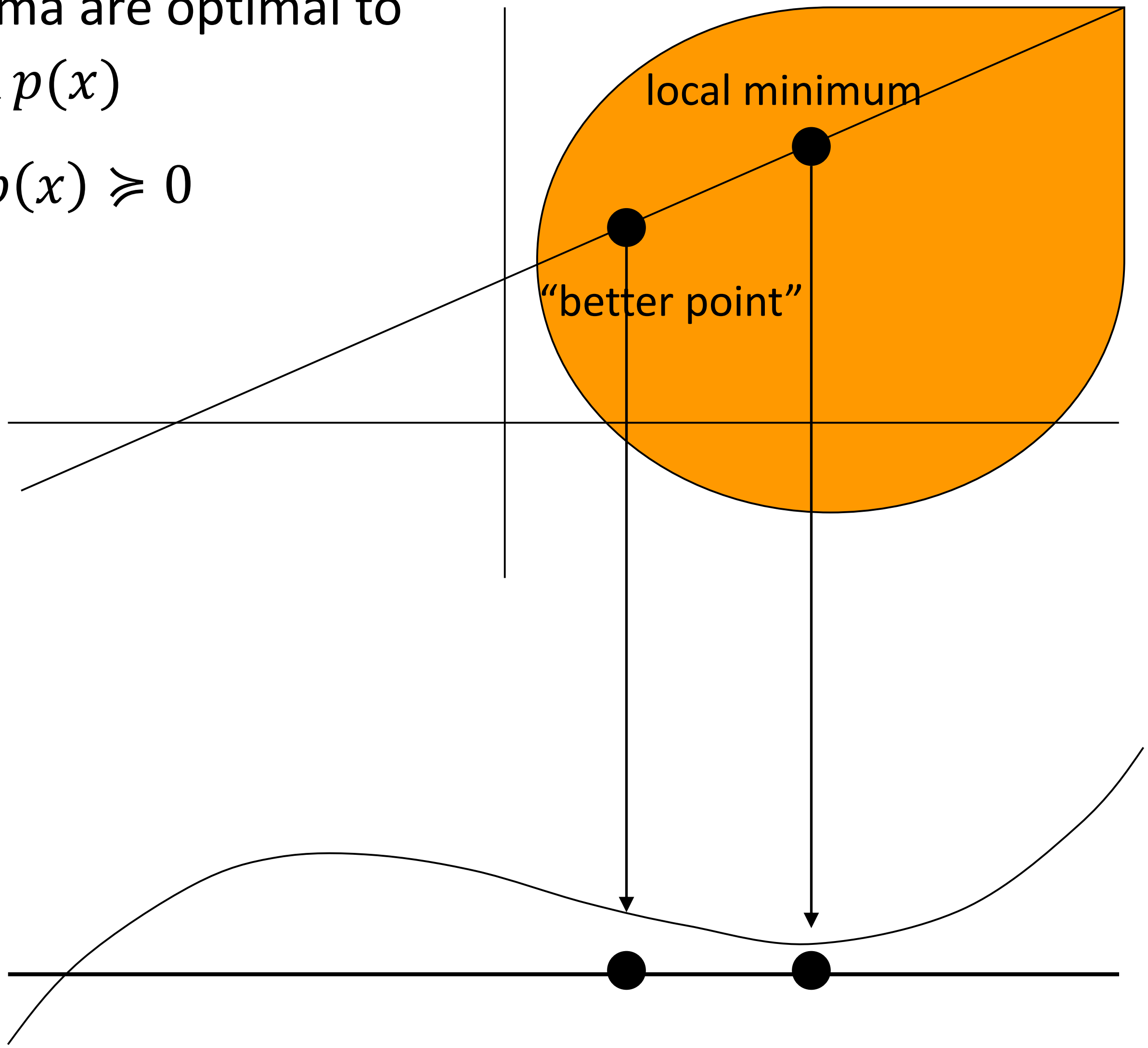
Note: the value of p at local minima must be the same.



A "convex" optimization problem proof (1/2)

Local minima are optimal to

$$\min_x p(x)$$
$$\nabla^2 p(x) \succcurlyeq 0$$

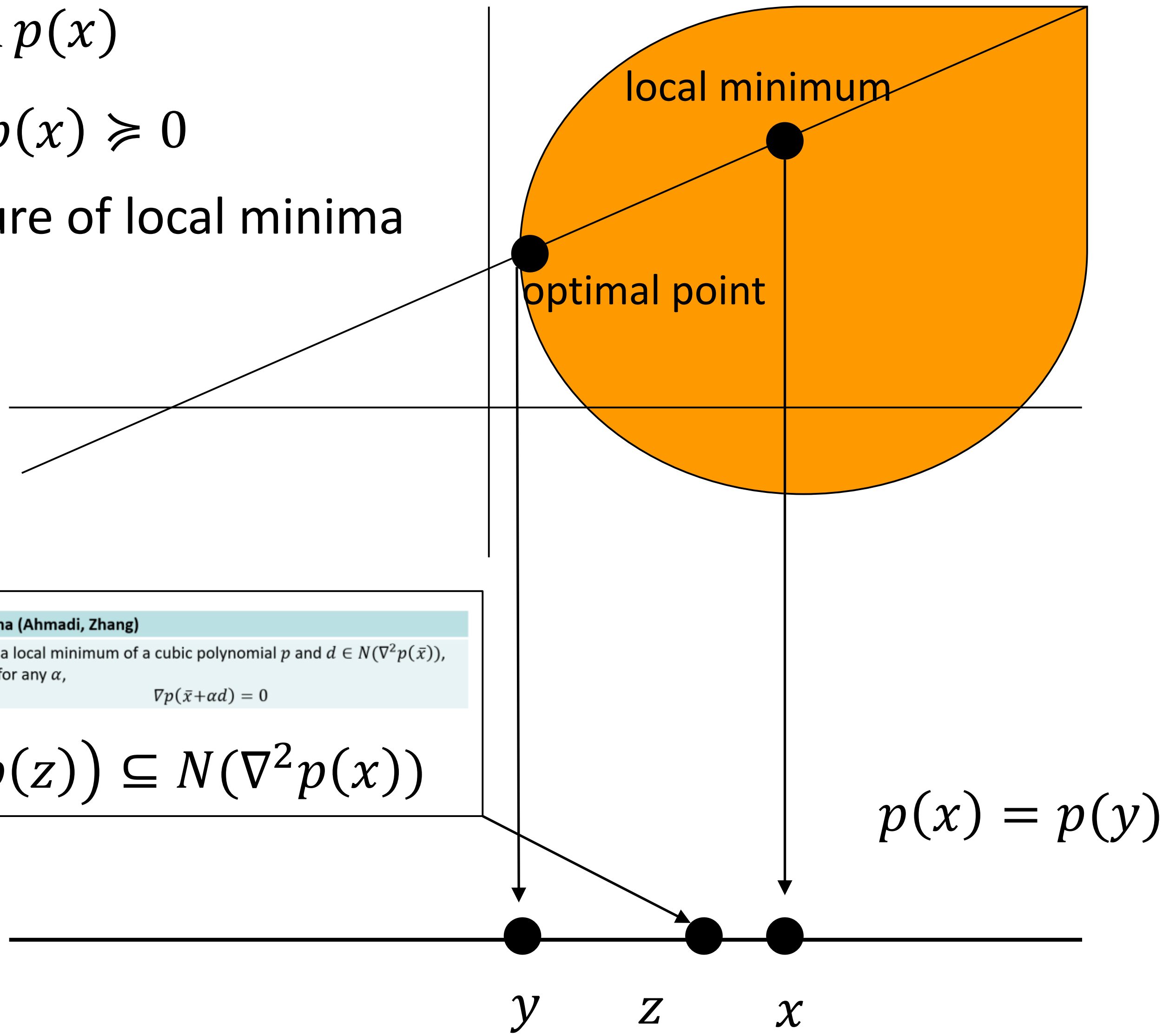


A "convex" optimization problem proof (2/2)

Solutions to

$$\min_x p(x)$$
$$\nabla^2 p(x) \succcurlyeq 0$$

are in closure of local minima



FONC: ✓

SONC: ✓

TOC: $N\left(\nabla^2 p(z)\right) \subseteq N(\nabla^2 p(x))$

Lemma (Ahmadi, Zhang)

If $\bar{x}$ is a local minimum of a cubic polynomial p and $d \in N(\nabla^2 p(\bar{x}))$, then for any α,

$$\nabla p(\bar{x}+\alpha d) = 0$$

Sum of squares polynomials

Sum of squares polynomials

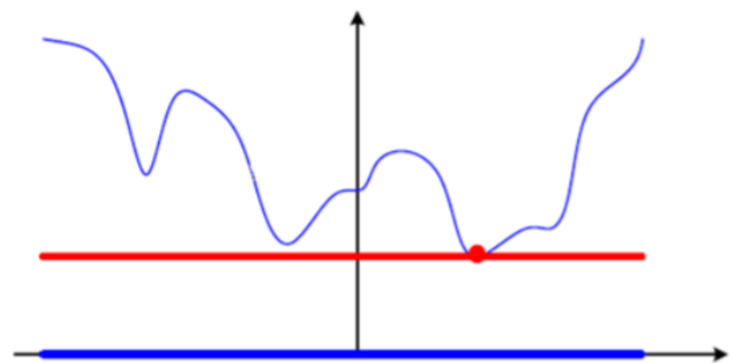
- A polynomial p is a *sum of squares* (sos) if it can be written as

$$p(x) = \sum q_i^2(x)$$

- Any sos polynomial is nonnegative
- Imposing that a polynomial is sos is a semidefinite constraint
- A matrix of polynomials $M(x)$ is an *sos-matrix* if the polynomial $y^T M(x) y$ is sos, or equivalently if $M(x) = R(x)R(x)^T$

Sum of squares relaxations

Find lower bounds on the optimal value of a polynomial optimization problem



$$\min_x f(x) \not= \max_\gamma \gamma$$
$$\geq$$

$f(x) - \gamma$ is a ~~nonnegative polynomial~~

sos

Sum of squares relaxations for constrained problems

Lasserre hierarchy:

$$\max_{\gamma, \sigma_i \, sos} \quad \gamma \qquad \leq \quad \min_{x \in \mathbb{R}^n} p(x)$$

$$p(x) - \gamma = \sigma_0(x) + \sum_{i=1}^{m} q_i(x)\sigma_i(x) \qquad\qquad q_i(x) \geq 0, i = 1, \dots, m$$

Putinar's Psatz

For σ_i of fixed degree, this is an SDP of size polynomial in data

As $\deg(\sigma_i) \to \infty$, the optimal value of the sos program will converge to the true optimal value (under a mild assumption)

Sos relaxation

$$\max_{\sigma(x),S(x)} \gamma \qquad\qquad \leq \min_x p(x)$$

$$p(x) - \gamma = \sigma(x) + Tr(\nabla^2 p(x) S(x)) \qquad \nabla^2 p(x) \succcurlyeq 0$$

σ is sos

S is an sos-matrix

Theorem

If p has a local minimum, the first level of this sos relaxation (i.e., when $\deg(\sigma) = \deg(S) = 2$) is tight.

Proof.

Produce an algebraic identity that attains the best possible value.

For any local minimum $\bar{x}$,

$$p(x) - p^* = \frac{1}{3}(x-\bar{x})^T \nabla^2 p(\bar{x})(x-\bar{x}) + Tr\left(\nabla^2 p(x)\left(\frac{1}{6}(x-\bar{x})(x-\bar{x})^T\right)\right)$$

Value at local min

$\sigma(x)$
sos

$S(x)$
sos-matrix

How to extract a local min itself?

Idea: Find the zeros of

$$\sigma(x) + Tr(\nabla^2 p(x) S(x))$$

Solve:

$$\begin{aligned} \min_x \quad & 0 \\ & \nabla^2 p(x) \succcurlyeq 0 \\ & \sigma(x) = 0 \\ & Tr\big(\nabla^2 p(x) S(x)\big) = 0 \end{aligned}$$

Nonlinear constraints…

or are they?

Recovering a local minimum

$\nabla^2 p(x) \succcurlyeq 0$ is a linear matrix inequality

σ is an sos quadratic, so the solutions to $\sigma(x) = 0$ can be found by solving a system of linear equations

$$\begin{aligned} \min_x \quad & 0 \\ & \nabla^2 p(x) \succcurlyeq 0 \quad ✓ \\ & \sigma(x) = 0 \quad ✓ \\ & Tr\left(\nabla^2 p(x) S(x)\right) = 0 \end{aligned}$$

$Tr\left(\nabla^2 p(x) S(x)\right) = 0$? (cubic equation)

Observation:

Since S is a quadratic sos matrix, $S(x) = R(x)R(x)^T$, where $R(x)$ is affine

$Tr\left(\nabla^2 p(x) S(x)\right) = 0 \Leftrightarrow \nabla^2 p(x) R(x) = 0$ (quadratic equation)

$R_i(x) \in N\left(\nabla^2 p(x)\right), \forall i$

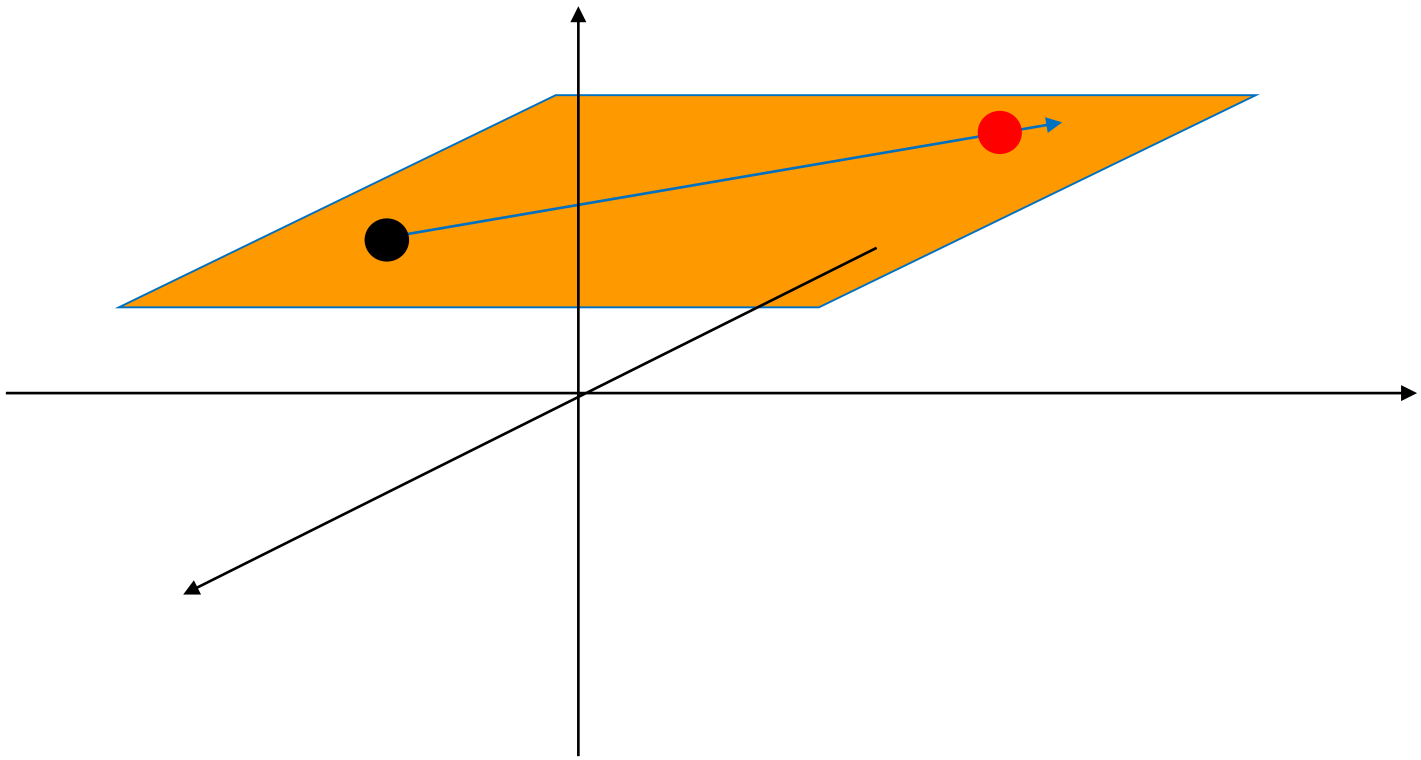
More geometry…

Relative Interior

Definition (Relative Interior)

The relative interior of a nonempty convex set S is the set

$$\{x \in S \mid \forall y \in S, \exists \alpha > 1, y + \alpha(y - x) \in S\}$$



More geometry

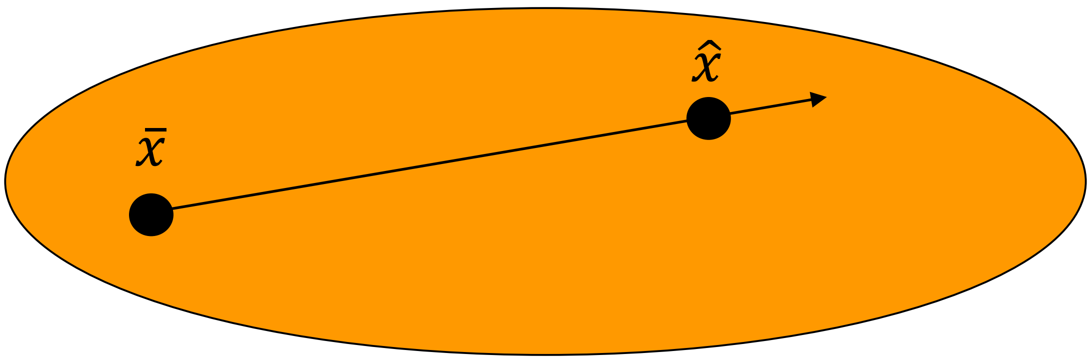
Lemma

Let $\bar{x}$ be a local minimum of a cubic polynomial p. Then for any $x \in \mathbb{R}^n$ and $d \in N(\nabla^2 p(\bar{x}))$, $d^T \nabla^2 p(x) d = 0$.

Lemma

Let $\bar{x}$ be a local minimum of a cubic polynomial p. Then for any $\hat{x}$ in the relative interior of the convexity region of p, $N(\nabla^2 p(\hat{x})) = N\left(\nabla^2 p(\bar{x})\right)$.

Proof (of second lemma).



Convex combination of PSD matrices: $N\left(\nabla^2 p(\hat{x})\right) \subseteq N\left(\nabla^2 p(\bar{x})\right)$

First Lemma + $\nabla^2 p(\hat{x}) \succcurlyeq 0$: $N\left(\nabla^2 p(\bar{x})\right) \subseteq N\left(\nabla^2 p(\hat{x})\right)$

Rewriting the cubic equation

What does this buy us?

Goal: Impose $Tr\left(\nabla^2 p(x)S(x)\right) = 0$ ✓

- Find any point $\hat{x}$ in the relative interior of the convexity region
- Find a basis $\{v_1, v_2, \ldots, v_k\}$ for $N(\nabla^2 p(\hat{x}))$
- Decompose $S(x) = R(x)R(x)^T$
- Impose $R_i(x) \in N\left(\nabla^2 p(\hat{x})\right) \forall i$ as $R_i(x) = \sum_{j=1}^{k} \alpha_j v_j \;\; \forall i$ (linear constraint!)

  For any x such that $\nabla^2 p(x) \succcurlyeq 0$, this is equivalent to imposing $R_i(x) \in N\left(\nabla^2 p(x)\right) \forall i \Leftrightarrow Tr\left(\nabla^2 p(x)S(x)\right) = 0$

PRINCETON UNIVERSITY ORFE

An SDP!

$$\min_{x} \quad 0$$

$$\nabla^2 p(x) \succcurlyeq 0$$

$$\sigma(x) = 0$$

$$Tr\left(\nabla^2 p(x) S(x)\right) = 0$$

Rewritable as an SDP!

Theorem

The relative interior of the feasible set of this SDP is the set of local minima of p.

Two steps require a point in the relative interior of a set

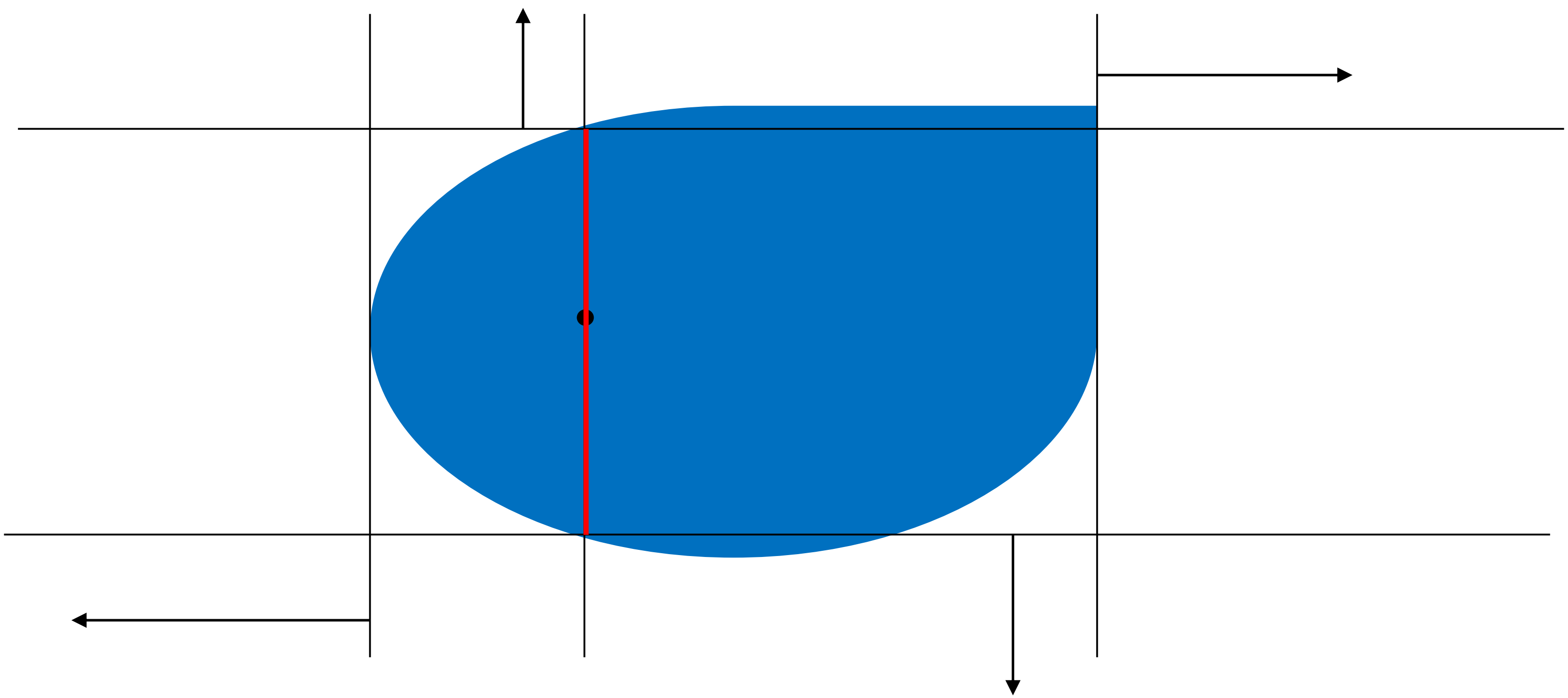
How can we get a point in the relative interior of a set?

Finding a point in the relative interior

Definition (Relative Interior)

The relative interior of a nonempty convex set S is the set

$$\{x \in S \mid \forall y \in S, \exists \alpha > 0, x + \alpha(y - x) \in S\}$$



Algorithm for finding a local minimum

- Find an sos-certified lower bound for value at any local minimum
- Find any point in the relative interior of the convexity region
- Find a basis for the null space of the Hessian of any local minimum
- Find relative interior solution of equivalent SDP

$$\begin{array}{ll} \max_{\gamma\in\mathbb{R},\sigma(x)\in\mathbb{R},S(x)\in\mathbb{R}^{n\times n}} & \gamma \\ \textit{subject to} & p(x)-\gamma=\sigma(x)+\mathrm{Tr}(S(x)\nabla^2 p(x)), \\ & S(x) \textit{ is sos}, \\ & \sigma(x) \textit{ is sos and has degree 2}, \\ & S_{ij}(x) \textit{ has degree 2}, \forall\ i,j\in\{1,\ldots,n\}. \end{array}$$

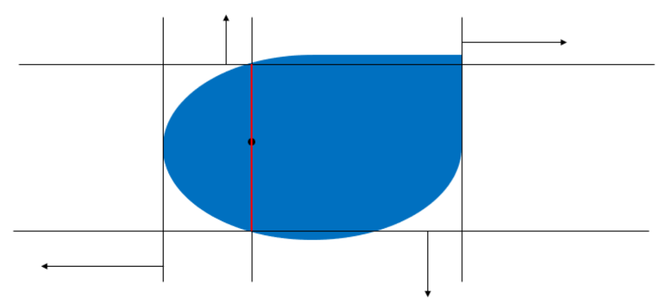
$$\begin{array}{ll} \min_{x\in\mathbb{R}^n} & 0 \\ \textit{subject to} & \nabla^2 p(x)\succeq 0, \\ & \sigma(x)=0, \\ & \mathrm{Tr}(S(x)\nabla^2 p(x))=0. \end{array}$$

Overall result

Theorem

Deciding if a cubic polynomial p has a local minimum, and finding one if it does, can be done in polynomially many calls to an SDP blackbox, Choleskly decompositions, and linear system solves of polynomial size.

Can be used to recover solutions



Why the blackbox assumption?

Local minima can be irrational:

$$p(x) = x^3 - 6x$$

$x = \sqrt{2}$ is the unique local minimum

Even if there are rational local minima, they can all have size exponential in the input:

$p(x) = y^T A(x) y$, where

$$A(x) = \begin{pmatrix} x_1 & 2 & & & & & \\ 2 & 1 & & & & & \\ & & x_2 & x_1 & \cdots & & \\ & & x_1 & 1 & & & \\ & & \cdots & & \ddots & & \cdots \\ & & & & & x_n & x_{n-1} \\ & & & & \cdots & x_{n-1} & 1 \end{pmatrix}$$

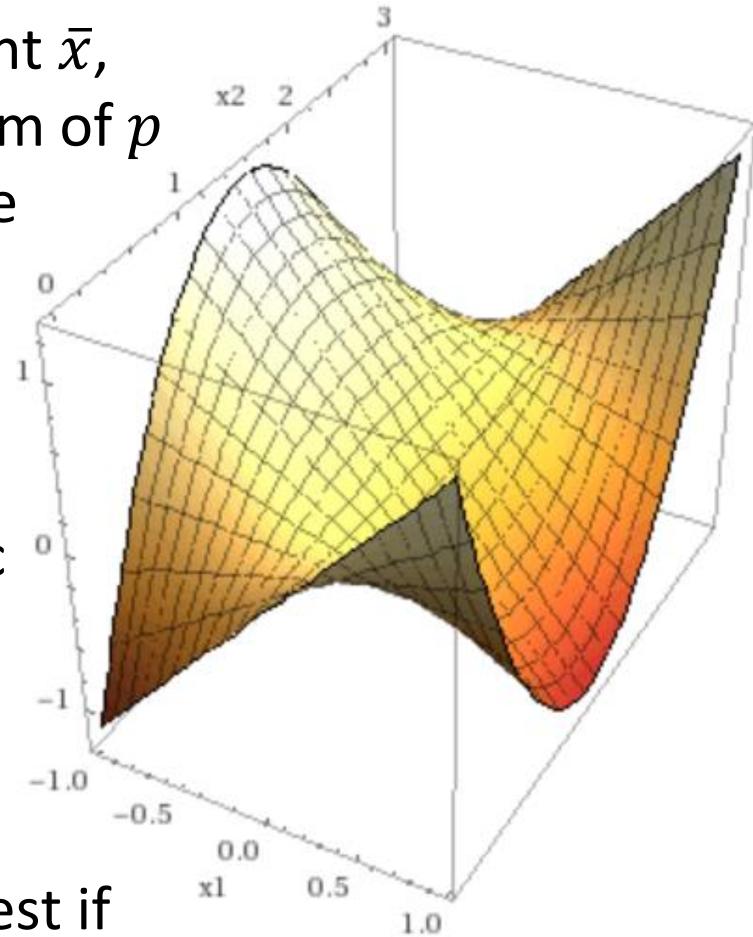
Local minima:

$\{y = 0\} \cap \{A(x) \succ 0\}$

$x_1 > 4, x_2 > 16, \ldots, x_n > 2^{2^n}$

Summary

- Given a cubic polynomial p and a point $\bar{x}$, checking whether $\bar{x}$ is a local minimum of p can be done in polynomial time in the Turing model



- It is strongly NP-hard to test if a cubic polynomial has a critical point

- Given a cubic polynomial p, we can test if there is a local minimum by solving polynomially many SDPs of polynomial size

Thank you!

Want to know more?

aaa.princeton.edu
princeton.edu/~jeffz