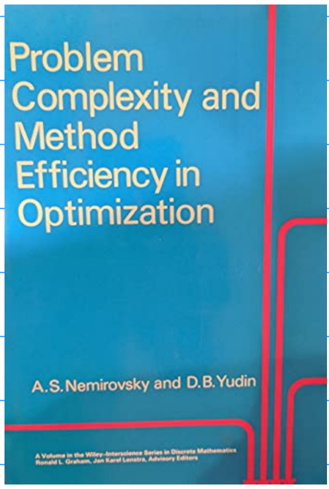
Complexity in continuous optimization (4/6)

February 13, 2025

Today: From blackbox complexity to computational complexity

•



Mathematical Programming 39 (1987) 117-129 117
North-Holland

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

Advances in Mathematics 397 (2022) 108119

Contents lists available at ScienceDirect

Advances in Mathematics

www.elsevier.com/locate/aim

Complexity aspects of local minima and related notions ☆

Amir Ali Ahmadi[a,*], Jeffrey Zhang[b]

Mathematical Programming
https://doi.org/10.1007/s10107-024-02139-3

FULL LENGTH PAPER

Series A

The computational complexity of finding stationary points in non-convex optimization

Alexandros Hollender[1] · Manolis Zampetakis[2]

Received: 24 January 2024 / Accepted: 12 August 2024

① Motivation

Setup: $\displaystyle \underset{x\in\mathbb{R}^n}{\text{minimize}}\ f(x)$, $f:\mathbb{R}^n\to\mathbb{R}$ nonconvex (C^1, C^2, C^3)

Blackbox complexity: Given an oracle (e.g. $x\mapsto(f(x),\nabla f(x))$)
how many oracle calls do we need to reach an ε-approximate solution as a function of ε?

e.g. $\|\nabla f(x)\|\le\varepsilon$

$\begin{cases}\|\nabla f(x)\|\le\varepsilon\\ \lambda_{min}(\nabla^2 f(x))\ge-\varepsilon\end{cases}$

$\|x-x^*\|\le\varepsilon$
where $x^*\in\underset{x}{\operatorname{argmin}}\ f(x)$

Computational complexity (for this lecture)

For this nonconvex problem,

① How hard is it to check that a point is a local minimum or a stationary point?

② How hard is it to determine whether the problem has a local minimum or a stationary point of some kind?

x^* local minimum of f if $\exists\,\delta>0$, $f(x^*)\le f(x)\ \forall x\in\mathbb{R}^n$, $\|x-x^*\|\le\delta$

NB: Since f is nonconvex, it can have local but non-global minima

x^* first-order stationary point (aka first-order critical point)

if $\|\nabla f(x^*)\| = 0$

x^* second-order stationary point if $\begin{cases} \|\nabla f(x^*)\| = 0 \\ \lambda_{min}(\nabla^2 f(x^*)) \geq 0 \end{cases}$

NB: • For a convex function, the three concepts are the same (and they also coincide with the notion of global minimum)

• For nonconvex functions,

x^* local minimum $\Rightarrow$ x^* 2^{nd}-order stationary

$\Rightarrow$ x^* 1^{st}-order stationary

(2) Local minima: hard (Murty & Kabadi, 1987)

→ Decision problem: Subset sum problem

Given $d_0, d_1, \ldots, d_n$ integer values, is there a solution to

$$\sum_{j=1}^{n} d_j y_j = d_0 \quad \text{with } y_j \in \{0,1\} \; \forall j = 1..n$$

→ Known to be NP-complete

→ With the same $d_0, d_1, \ldots, d_n$, consider the optimization problem

Nonconvex quadratic program (QP)
$$\begin{cases} \underset{y \in \mathbb{R}^n}{\text{minimize}} & \left(\sum_{j=1}^{m} d_j y_j - d_0\right)^2 + \underbrace{\sum_{j=1}^{m} y_j(1-y_j)}_{\text{Nonconvex quadratic}} \\ \text{s.t.} & 0 \leq y \leq 1 \iff 0 \leq y_j \leq 1 \ \forall j = 1..m \quad \text{linear constraints} \end{cases}$$

The Subset sum problem has a solution if and only if the optimal value of (QP) is 0.

$\Rightarrow$ Finding the optimal value of (QP) is NP-hard.

↳ From this example, the authors conclude that in general computing a global minimum of a nonconvex problem is hard.

Other questions in the paper: Answer through copositivity.

Copositivity

- Given $D \in \mathbb{Z}^{n \times n}$, checking $D \succeq 0$ $\left(\iff x^T D x \geq 0 \ \forall x \in \mathbb{R}^n\right)$

corresponds to the decision problem:

Does there exist $x \in \mathbb{R}^n$ such that $x^T D x < 0$?

$\Rightarrow$ Can answer in polynomial time

(e.g. Gaussian pivots, $O(\underset{\text{ops}}{n^3})$)

$\Rightarrow$ Implies that checking $\nabla^2 f(x) \succeq 0$ for some f and x can be done in polynomial time
(hence checking that a point is 2^{nd}-order stationary point can be done in polynomial time)

- $D \in \mathbb{Z}^{n\times n}$ is copositive if $x^T D x \geq 0 \quad \forall x \geq 0$

Decision pb: Does there exist x such that $x \geq 0$ and $x^T D x < 0$?

Unless D is $\succeq 0$, NP-complete

$$\Rightarrow \quad \underset{x\in\mathbb{R}^n}{\text{minimize}} \; x^T D x \quad \text{s.t.} \quad x \geq 0 \quad \Big| \text{ NP-hard}$$

$\hookrightarrow$ Can use copositive matrices to reduce unconstrained nonconvex optimization problems to nonconvex QPs

* Consider

$$(P) \quad \underset{z\in\mathbb{R}^n}{\text{minimize}} \; h(z) = \begin{bmatrix} z_1^2 \\ \vdots \\ z_n^2 \end{bmatrix}^T D \begin{bmatrix} z_1^2 \\ \vdots \\ z_n^2 \end{bmatrix}$$

with $D \in \mathbb{Z}^{n\times n}$, $D = D^T$

$\rightarrow$ Quartic polynomial (polynomial of degree 4 in the entries of z)

with the change of variables $x = z^2$ ($x_i = z_i^2 \ \forall i = 1..n$), the problem (P) is reformulated as

$$\left.\underset{x \in \mathbb{R}^n}{\text{minimize}} \ x^T D x \quad \text{s.t. } x \geq 0 \ \right]$$ Back to the copositivity problem

$\rightarrow$ From there, conclude that unconstrained nonconvex problems are NP-hard

Takeaway

- For general nonconvex optimization problems, checking that a point is a local minimum is NP-hard
- Tool: Matrix copositivity problem

(3) Polynomial problems

$\rightarrow$ Idea: what if we restrict ourselves to (nonconvex) polynomial functions?

Motivation: Taylor expansion

$f \ C^1 \Rightarrow f(x+s) \simeq f(x) + \nabla f(x)^T s$) 1st order polynomial

$f \ C^2 \Rightarrow f(x+s) \simeq f(x) + \nabla f(x)^T s + \frac{1}{2} s^T \nabla^2 f(x) s$) 2nd order polynomial

$f \; C^p \Rightarrow f(x+s) \simeq f(x) + \nabla f(x)^T s + \dots + \underbrace{\nabla^p f(x)}_{\text{pth order derivative}} \underbrace{[s, s, \dots, s]}_{p \text{ times}}$

Decision problems

① Given a nonconvex polynomial f and a point $x \in \mathbb{R}^n$,

is x a $\begin{cases} \text{local minimum of } f? \\ \text{second-order stationary point?} \\ \text{first-order stationary point?} \end{cases}$

② Given a nonconvex polynomial f,

does f possess $\begin{cases} \text{a local minimum?} \\ \text{a } 2^{nd} \text{ order-stationary point?} \\ \text{a } 1^{st} \text{ order-stationary point?} \end{cases}$

Answers (Ahmadi & Zhang 2022)

For ① (f, x)

x 1^{st} order stationary point $\in P$

x 2^{nd} order ———— $\in P$

x local minimum

$\in$ NP-hard if degree of $f \geq 4$
(Murty & Kabadi)

$\in P$ if degree ≤ 2

(Ahmadi & Zhang) $\in P$ if degree $= 3$

For (2) f,

- f has a 1^{st} order stationary point
  $\in P$ if degree of $f \leq 2$
  $\in$ NP-hard otherwise

- f has a 2nd order stationary point / local minimum
  $\in P$ if degree ≤ 2
  $\in$ NP-hard if degree ≥ 4
  $\in$ "SDP" if degree $= 3$
  (can find the answer by solving a finite number of SDPs)

Proof sketches

(a) If $f : \mathbb{R}^n \to \mathbb{R}$ is a polynomial of degree $d \geq 3$, then it is strongly NP-hard to decide whether f has a first-order stationary point

$\longrightarrow$ Reduction to MAXCUT

MAXCUT: G undirected, unweighted graph with m vertices

Goal: Find a cut of size k

Let A be the adjacency matrix of G.
$A \in \mathbb{R}^{m \times m}$

G has a cut of size $k \iff \begin{cases} q_0(x) = \frac{1}{4} \sum_{i=1}^{m} \sum_{j=1}^{m} A_{ij}(1 - x_i x_j) - k = 0 \\ q_i(x) = x_i^2 - 1 = 0 \quad i = 1 \dots m \end{cases}$

has a solution

Define $p(\underbrace{x}_{m},\underbrace{y}_{m+1},\underbrace{z}_{1}) = \sum_{i=0}^{m} y_i q_i(x) + z^d$

p has degree d and $2m+2$ variables.

p has a 1st order critical point

$\Leftrightarrow \begin{cases} q_0(x) = 0 \\ q_i(x) = 0 \quad i = 1..m \end{cases}$ has a solution

$$\nabla p(\bar{x},\bar{y},\bar{z}) = 0 \Leftrightarrow \begin{cases} \frac{\partial p}{\partial x_1}(\ ) = 0 \\ \vdots \\ \frac{\partial p}{\partial x_m}(\ ) = 0 \\ \frac{\partial p}{\partial y_0}(\bar{x},\bar{y},\bar{z}) = q_0(x) = 0 \\ \vdots \\ \frac{\partial p}{\partial y_i}(\bar{x},\bar{y},\bar{z}) = q_i(x) = 0 \\ \vdots \\ \frac{\partial p}{\partial z}(\bar{x},\bar{y},\bar{z}) = 0 \end{cases}$$

(b) If f is a polynomial with degree $d \geq 4$, it is strongly NP-hard whether f has a 2nd order stationary point.

$\rightarrow$ Again reduction to MAXCUT
(with the same notations as before)

Polynomial

$$p(x,y,z,w) = \sum_{i=0}^{m}\left(y_i^2 q_i(x) - z_i^2 q_i(x)\right) + \underbrace{w^d}_{\text{degree } d}$$

2nd order condition $\nabla^2 p(x,y,z,w) \succeq 0$

$\frac{\partial^2}{\partial y_0^2} p(x,y,z,w)$

only holds if

$$2\begin{bmatrix} q_0(x) & & 0 \\ & \ddots & \\ 0 & & q_m(x)\end{bmatrix} = \nabla^2_{yy} p(x,y,z,w) \succeq 0$$

$$2\begin{bmatrix} -q_0(x) & & 0 \\ & \ddots & \\ 0 & & -q_m(x)\end{bmatrix} = \nabla^2_{zz} p(x,y,z,w) \succeq 0$$

p has a 2nd order stationary point $\Rightarrow$ $\begin{cases} q_0(x) = 0 \\ q_i(x) = 0 \quad i = 1..m \end{cases}$ has a solution.

($\Leftarrow$) If $q_0(x) = \dots = q_m(x) = 0$,

then $(x, 0_{\mathbb{R}^{m+1}}, 0_{\mathbb{R}^{m+1}}, 0)$ is a second-order stationary point for p.

(c) For f polynomial of degree 4, it is strongly NP-hard to decide whether it has a local minimum

$\longrightarrow$ Reduction to the stable set problem

Given G graph with m vertices

Find a set of pairwise not-adjacent vertices

G has a stable of size r

$\iff$ the polynomial $p_{A,r-0.5}$ has no local minima

where A is the adjacency matrix of G

and $\forall \rho > 0$, $p_{A,\rho}(x) \doteq \begin{bmatrix} x_1^2 \\ \vdots \\ x_n^2 \end{bmatrix}^T (\rho A + \rho I - J) \begin{bmatrix} x_1^2 \\ \vdots \\ x_n^2 \end{bmatrix}$

matrix of all ones

$\rightarrow$ Hardness results for quartic polynomials

$\rightarrow$ Cubic polynomials are a special case

(4) Cubic polynomials

$p : \mathbb{R}^n \to \mathbb{R}$ polynomial of degree 3

p is C^3 and we know that

$$\underbrace{\bar{x} \text{ local minimum of } p \Rightarrow \begin{cases} \|\nabla p(\bar{x})\| = 0 \\ \nabla^2 p(\bar{x}) \succeq 0 \end{cases}}_{\in P}$$

Existing result (applies to every C^3 function)

$$\bar{x} \text{ local minimum of } p \Rightarrow \begin{cases} \|\nabla p(\bar{x})\| = 0 \\ \nabla^2 p(\bar{x}) \succeq 0 \\ \nabla^3 p(\bar{x})[d,d,d] = 0 \quad \forall d \text{ such that } \nabla^2 p(\bar{x}) d = 0 \end{cases}$$

Third-order stationary point

The 3rd order derivative vanishes in the null space of the Hessian

$\ominus$ Third-order stationarity does not characterize local minima

(Ex) $n=2$, $p(x) = x_2^2 - x_1^2 x_2$

$$\bar{x} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \nabla p(x) = \begin{pmatrix} -2x_1 x_2 \\ 2x_2 - x_1^2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\nabla^2 p(x) = \begin{pmatrix} -2x_2 & -2x_1 \\ -2x_1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$$

$$p\begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0$$

$$\underbrace{\nabla^3 p(x)[d,d,d] = 0}_{\frac{\partial^3}{\partial x_1^3} p(\,) = 0} \quad \forall \underbrace{d, \nabla^2 p(x) d = 0}_{d = \begin{pmatrix} \alpha \\ 0 \end{pmatrix}}$$

$$p\begin{pmatrix} 1 \\ 1/2 \end{pmatrix} = \frac{1}{4} - \frac{1}{2} = -\frac{1}{4}$$

Key result

$\bar{x}$ is a local minimum of p

$$\iff \begin{cases} \|\nabla p(\bar{x})\| = 0 \\ \nabla^2 p(\bar{x}) \succeq 0 \\ \nabla p_3(\bar{x}) = 0 \quad \forall d \text{ such that } \nabla^2 p(\bar{x}) d = 0 \end{cases}$$

p_3 is the homogeneous part of degree 3 of p

(Ex) $_{n=3}$ $p(x) = x_1^3 + x_2^2 x_3 + x_1 x_2 + 2x_3$

$$p_3(x) = x_1^3 + x_2^2 x_3$$

As a result, can check that a point is a local minimum of p in polynomial time

(in particular because you can compute a rational basis of the null space of $\nabla^2 p(\bar{x})$ when p has integer coefficients in polynomial time)

Procedure

→ Check $\underbrace{\nabla p(\bar{x})}_{=0}$, $\nabla^2 p(x) \succeq 0$

→ Compute $(v_1, \dots, v_\ell)$ basis for null$(\nabla^2_p(x))$

→ Check if $g(\lambda) = \nabla p_3\left(\sum_{i=1}^{\ell} \lambda_i v_i\right)$ has zero coefficients.

For the example

$$p_3(x) = -x_1^2 x_2$$

$$\nabla p_3(x) = \begin{bmatrix} -2x_1x_2 \\ -x_1^2 \end{bmatrix} =$$

$$\text{null}\left(\nabla^2 p\begin{pmatrix}0\\0\end{pmatrix}\right) = \text{vect}\left(\begin{pmatrix}1\\0\end{pmatrix}\right)$$

$$\nabla p_3\begin{pmatrix}\alpha\\0\end{pmatrix} = \begin{bmatrix} 0 \\ -\alpha^2 \end{bmatrix}$$

⇒ checking 1st order stationarity / 2nd order stationarity / local optimality is polynomial for cubic polynomials

⇒ Deciding whether a cubic polynomial has a 1st order stationary point is NP-hard

$\Rightarrow$ Deciding whether the polynomial has a 2^{nd} order stationary point or a local minimum can be done via solving SDPs.

For second-order stationarity

If the problem is in ~~P~~ NP, then so is the SDP feasibility problem.

SDP Feasibility Problem: Given $A_0, \ldots, A_m$ $m+1$ symmetric matrices $n \times n$ with rational entries, does there exist $x \in \mathbb{R}^m$ such that

$$A_0 + \sum_{i=1}^{m} x_i A_i \succeq 0$$

One form of SDP:

SDP in standard form

$$\begin{aligned} \underset{X \in \mathbb{R}^{n \times n}}{\text{minimize}} \quad & \operatorname{trace}(CX) \\ \text{s.t.} \quad & \operatorname{trace}(A_i X) = b_i \quad i = 1..m \\ & X \succeq 0 \end{aligned}$$

C, A_i symmetric matrices

$b \in \mathbb{R}^m$

Dual of SDP:

$$\begin{aligned} \underset{\substack{y \in \mathbb{R}^m \\ S \in \mathbb{R}^{n \times n}}}{\text{maximize}} \quad & b^T y \\ \text{s.t.} \quad & -C + \sum_{i=1}^{m} y_i A_i \succeq 0 \\ & S \succeq 0 \end{aligned}$$

$\rightarrow$ Approach: Reformulate the decision problem using a semidefinite program

(using sum-of-squares polynomials and sum-of-squares matrices)

p has a 2^{nd} order stationary point

$\Rightarrow$ A certain SDP has a finite optimal value, and that value is the value of p at every 2^{nd} order stationary point

(the set of 2^{nd} order stationary points is a convex set)

Thm] For every cubic polynomial $p: \mathbb{R}^n \to \mathbb{R}$, p can be written as

$$p(x) = c + b^T x + \frac{1}{2} x^T Q x + \frac{1}{6} \sum_{i=1}^{n} x^T x_i H_i x$$

for $b \in \mathbb{R}^n$, $Q \in \mathbb{R}^{n \times n}$, $H_i \in \mathbb{R}^{n \times n}$ $\forall i$

Then the set of second-order stationary points of p is

$$\left\{ x \in \mathbb{R}^n \mid \exists\, Y \in \mathbb{R}^{n \times n},\ Y = Y^T,\ \exists\, z \in \mathbb{R} \text{ such that} \right.$$

$$\frac{1}{2} \operatorname{trace}(QY) + b^T x + \frac{z}{2} = 0$$

$$\frac{1}{2} \operatorname{trace}(H_i Y) + e_i^T Q x + b_i = 0$$

$$\begin{bmatrix} \sum y_i H_i + Q & \sum \operatorname{trace}(H_i Y) e_i + Qy \\ \sum \operatorname{trace}(H_i Y) e_i + Qy & z \end{bmatrix} = 0$$

$$\left. \begin{bmatrix} Y & x \\ x^T & 1 \end{bmatrix} \succeq 0 \right\}$$

Summary : → For checking / finding stationary points / local minima, we have hardness results

⇒ Case of cubic polynomials not fully understood

→ Partly explains the success of blackbox complexity

→ Recent results on complexity of finding approximate 1^{st} order stationary points

$\|\nabla f(x)\| \leq \varepsilon$ where f nonconvex $C^{1,1}$

$f: \mathbb{R}^n \to \mathbb{R}$ $n \geq 2$

⇒ PLS - complete

Polynomial Local Search