Global optimality conditions in non-convex optimization

Panos M. Pardalos
Distinguished Professor Center For Applied Optimization,
Industrial and Systems Engineering, University of Florida,
Florida, USA.
National Research University Higher School of Economics,
Laboratory of Algorithms and Technologies for Network Analysis,
Nizhny Novgorod, Russia

Outline

Introduction

Greek mathematicians solved optimally some problems related to their geometrical studies.

- Euclid considered the minimal distance between a point and a line.
- Heron proved that light travels between two points through the path with shortest length when reflecting from a mirror.

Optimality in Nature

- Fermat's principle (principle of least time)
- Hamilton's principle (principle of stationary action)
- Maupertuis' principle (principle of least action)

Optimality in Biology:

- Optimality Principles in Biology (R. Rosen, New York: Plenum Press, 1967).
- Example: What determines the radius of the aorta? The human aortic radius is about 1.5cm (minimize the power dissipated through blood flow)

1951: H.W. Kuhn and A.W. Tucker, Optimality conditions for nonlinear problems.
F. John in 1948 and W. Karush in 1939 had presented similar conditions

Complexity of Kuhn-Tucker Conditions

Consider the following quadratic problem:

$$\begin{aligned} \min \quad & f(x) = c^T x + \frac{1}{2} x^T Q x \\ st. \quad & x \geq 0, \end{aligned}$$

where Q is an arbitrary $n \times n$ symmetric matrix, $x \in R^n$. The KKT optimality conditions for this problem become so-called linear complementarity problem $(LCP(Q, c))$

Complexity of Kuhn-Tucker Conditions

Linear complementarity problem $LCP(Q, c)$ is formulated as follows.
Find $x \in R^n$ (or prove that no such an x exists) such that:

$$Qx + c \geq 0, \ \ x \geq 0$$
$$x^T(Qx + c) = 0.$$

Theorem

Theorem (Horst, Pardalos, Thoai, 1994 - [5])

The problem $LCP(Q,c)$ is NP-hard.

Proof.

Consider the following $LCP(Q,c)$ problem in R^{n+3} defined by

$$Q_{(n+3)\times(n+3)} = \begin{pmatrix} -I_n & e_n & -e_n & 0_n \\ e_n^T & -1 & -1 & -1 \\ -e_n^T & -1 & -1 & -1 \\ 0_n^T & -1 & -1 & -1 \end{pmatrix}, \ c_{n+3}^T = (a_1, \ldots, a_n, -b, b, 0),$$

where a_i, $i = 1, \ldots, n$, and b are positive integers, I_n is the $n \times n$-unit matrix and the vectors $e_n \in R^n$, $0_n \in R^n$ are defined by

$$e_n^T = (1, 1, \ldots, 1), \ 0_n^T = (0, 0, \ldots, 0).$$

Theorem (Continue)

Proof.

Consider the following knapsack problem. Find a feasible solution to the system

$$\sum_{i=1}^{n} a_i x_i = b, \;\; x_i \in \{0, 1\} \; (i = 1, \dots, n).$$

This problem is known to be NP-complete. We will show that $LCP(Q, c)$ is solvable iff the associated knapsack problem is solvable.
If x solves the knapsack problem, then $y = (a_1 x_1, \dots, a_n x_n, 0, 0, 0)^T$ solves $LCP(Q, c)$.

Conversely, assume y solves the considered $LCP(Q, c)$. This implies that $\sum_{i=1}^{n} y_i = b$ and $0 \leq y_i \leq a_i$. Finally, if $y_i < a_i$, then $y^T(Qy + c) = 0$ enforces $y_i = 0$. Hence, $x = (\frac{y_1}{a_1}, \dots, \frac{y_n}{a_n})$ solves the knapsack problem. □

Complexity of local minimization

Consider following quadratic problem:

$$\begin{aligned} \min \quad & f(x) \\ s.t. \quad & Ax \geq b, \; x \geq 0 \end{aligned}$$

where $f(x)$ is an indefinite quadratic function. We showed, that the problem of checking local optimality for a feasible point and the problem of checking whether a local minimum is strict are NP-hard.

3-satisfiability problem

Consider the 3-satisfiability (3-SAT) problem: given a set of Boolean variables $x_1, \ldots, x_n$ and given a Boolean expression S (in conjunctive normal form) with exactly 3 literals per clause,

$$S = \bigwedge_{i=1}^{m} (\bigvee_{j=1}^{3} l_{ij}), \; l_{ij} \in \{x_i, \bar{x}_i | i = 1, \ldots, n\}$$

is there a truth assignment for the variables x_i which makes S true?

The 3-SAT problem is known to be NP-complete.

Construction of indefinite quadratic problem instances

For each instance of a 3-SAT problem we construct an instance of an optimization problem in the real variables $x_0, \ldots, x_n$. For each clause in S we associate a linear inequality in a following way:

$$\text{If } l_{ij} = x_k, \text{ we retain } x_k$$
$$\text{If } l_{ij} = \bar{x}_k \text{ we use } 1 - x_k$$

We add an additional variable x_0 and require that the corresponding sum is greater than or equal to $\frac{3}{2}$. Thus, we associate to S a system of linear inequalities $A_s x \geq (3/2 + c)$. Let $D(S) \subset R^{n+1}$ be a feasible set of points satisfying these constraints.
With a given instance of the 3-SAT problem we associate the following indefinite quadratic problem:

$$\min_{x \in D(S)} f(x) = -\sum_{i=1}^{n} (x_i - (1/2 - x_0))(x_i - (1/2 + x_0)).$$

Complexity of local minimization

Theorem (Pardalos, Schnitger, 1988 - [6])

S is satisfiable iff $x^* = (0, 1/2, \ldots, 1/2)^T$ is not a strict minimum
S is satisfiable iff $x^* = (0, 1/2, \ldots, 1/2)^T$ is not a local minimum

Corollary

For a quadratic indefinite problem the problem of checking local optimality for a feasible point and the problem of checking whether a local minimum is strict are NP-hard.

Complexity of checking convexity of a function

- The role of convexity in modern day mathematical programming has proven to be fundamental
- The great watershed in optimization is not between linearity and nonlinearity, but convexity and nonconvexity (R. Rockafellar)
- The tractability of a problem is often assessed by whether the problem has some sort of underlying convexity.
- Can we decide in an efficient manner if a given optimization problem is convex?

Complexity of checking convexity of a function

One of seven open problems in complexity theory for numerical optimization (Pardalos, Vavasis, 1992 - [7]):

Given a degree-4 polynomial in n variables, what is the complexity of determining whether this polynomial describes a convex function?

Theorem

Theorem (Ahmadi et al., 2011)

Deciding convexity of degree four polynomials is strongly NP-hard. This is true even when the polynomials are restricted to be homogeneous (all terms with nonzero coefficients have the same total degree).

Corollary (Ahmadi et al., 2011)

It is NP-hard to check convexity of polynomials of any fixed even degree $d \geq 4$.

Theorem

Theorem (Ahmadi et al., 2011)

It is NP-hard to decide strong convexity of polynomials of any fixed even degree $d = 4$.

Theorem (Ahmadi et al., 2011)

It is NP-hard to decide strict convexity of polynomials of any fixed even degree $d = 4$.

Theorem

Theorem (Ahmadi et al., 2011)

For any fixed odd degree d, the quasi-convexity of polynomials of degree d can be checked in polynomial time.

Corollary (Ahmadi et al., 2011)

For any fixed odd degree d, the pseudoconvexity of polynomials of degree d can be checked in polynomial time.

Theorem

Theorem (Ahmadi et al., 2011)

It is NP-hard to check quasiconvexity/pseudoconvexity of degree four polynomials. This is true even when the polynomials are restricted to be homogeneous.

Corollary (Ahmadi et al., 2011)

It is NP-hard to decide quasiconvexity of polynomials of any fixed even degree $d \geq 4$.

The complexity results described above can be summarized in the following table [1]:

Property versus 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

A yes (no) entry means that the question is trivial for that particular entry because the answer is always yes (no) independent of the input. By P, we mean that the problem can be solved in polynomial time

Linear Complementarity problem formulated as MIP

Theorem (Pardalos, Rosen, 1988)

The following mixed integer linear program

$$\begin{aligned} & \max\ \alpha \\ \text{s.t.}\quad & 0 \le My + \alpha q \le z, \\ & 0 \le y \le e - z,\ z \in \{0,1\}^n,\ 0 \le \alpha \le 1, \end{aligned}$$

associated to the LCP(M,q) with $q \neq 0$ has an optimal solution (y^*, z^*, α^*) satisfying $\alpha^* \ge 0$.
The LCP has a solution, and $x^* = \frac{y^*}{\alpha^*}$ is a solution, if and only if $\alpha^* > 0$.

The LCP is equivalent to the linear integer feasibility problem

Motzkin-Strauss formulation for maximum clique

Let A be the adjacency matrix of a graph G. We write $u \sim v$ if u and v are distinct and adjacent, and $u \not\sim v$ otherwise. Consider the Motzkin - Strauss formulation (also called the Motskin-Strauss QP) of the Maximum Clique Problem:

$$\begin{aligned} \max \quad & x^T A x/2 \\ s.t. \quad & e^T x = 1 \\ & x \geq 0 \end{aligned} \tag{P}$$

The simplex of feasible solutions of P we denote by $\triangle_n$.

Necessary Optimality Conditions

First Order (KKT) Conditions:

$$\begin{aligned} -Ax + \lambda e - \mu &= 0 \\ x^T e &= 1 \\ x &\geq 0 \\ \mu &\geq 0 \\ \mu_u x_u &= 0 \ \forall u \end{aligned} \tag{1}$$

Second Order (KKT) Conditions:

$$\text{Let } Z(x) = \{u | x_u = 0\}. \\ -y^T A y \geq 0, \ \forall y \text{ s.t.} y^T e = 0, \ y_u = 0 \ \forall u \in Z(x). \tag{2}$$

Necessary Optimality Conditions

- Since the constraints of P are simplex constraints, regularity conditions always hold, and hence both (1) and (2) are necessary for a solution to be locally (globally) optimal.
- The following chain of inclusions takes place:

$$\begin{aligned}\text{global optimal solutions} &\subset \text{locally optimal solutions}\\ &\subset \text{second order points}\\ &\subset \text{first order points.}\end{aligned}$$

A key semidefiniteness lemma

Lemma (L. E. Gibbons, D. W. Hearn, P. M. Pardalos, M. V. Ramana, 1997)

Let $G = (V, E)$ be a graph whose adjacency matrix is A. Let $S \subset V$, $T = V \backslash S$, and define the polyhedral cone:

$$Y = \{y | y^T e = 0, y_v \geq 0 \ \forall v \in T\}.$$

Then the following are equivalent.

A key semidefiniteness lemma (Continue)

Lemma

(1) A is negative semidefinite over the cone Y, i.e., $yTAy \leq 0\ \forall y \in Y$.
(2) If $u, v, w \in V, u \nsim v, u \nsim w, v \sim w$, then either $u,\ v \in T$ or $u,\ w \in T$.
(3) The node sets S, T can be partitioned as:

$$S = S_1 \cup S_2 \cup \cdots \cup S_l, \quad T = T_0 \cup T_1 \cup \cdots \cup T_l,$$

such that
(a) Each S_i is nonempty.
(b) $S_i \cup T_i$ is an independent set for every $i = 1, \ldots, l$.
(c) For all $i \neq j$, every vertex in $S_i \cup T_i$ is adjacent to every vertex in S_j.
(d) Every node of T_0 is adjacent to every node in S.

First order solutions

Theorem

Let $x \in \triangle_n$, *and* $S(x) = \{u | x_u > 0\}$. *Then* x *is a first order point if and only if the vector*

$$\mu = (x^T A x)e - Ax$$

satisfies $\mu \geq 0$ *and* $\mu_u = 0 \ \forall u \in S(x)$.

Second order solutions

Theorem

Let x be a feasible solution to P, and let H be the induced subgraph of G indexed on $S(x)$, the support of x. Then x is a second order point if and only if the following hold:
(1) H is a complete l-partite graph (for some l), with the partition $S(x) = S_1 \cup \ldots s_l$, with each S_i being nonempty. U Si, with each Si being nonempty.
(2) $\sum_{u \in S_j} x_u = 1/l \; \forall j = 1, \ldots, l$.
(3) $Ax \leq (1 - 1/l)e$
Also, if x is a second order point, then $f(x)$ is $\frac{1}{2}(1 - 1/l)$, and $\lambda = 1 - 1/l$, where λ is the Lagrange multiplier associated with the constraint $e^T x = 1$ in P.

Locally optimal solutions

Theorem

Let $x \in \triangle_n$. Then x is a local maximum of P if and only if there exists an integer l such that, with

$$\lambda := 1 - 1/l, \quad S := \{u | x_u > 0\}, \quad T := \{u | x_u = 0, (Ax)_u = \lambda\},$$

we have
(1) $Ax \leq \lambda e$.
(2) $(Ax)_u = \lambda \; \forall u \in S$ *(3) There exist partitions of* S *and* T,

$$S = S_1 \cup \cdots \cup S_l, \quad T = T_1 \cup \cdots \cup T_l,$$

such that
(a) $S_i \neq \emptyset \; \forall i = 1, \ldots, l$.
(b) $S_i \cup T_i$ *is independent for every* i.
(c) $i \neq j, \; u \in S_i \cup T_i, v \in S_j \Rightarrow u \sim v$.

Minimizing the rank of a matrix

Consider a problem of minimizing the rank of a matrix:

$$\begin{aligned} \min \quad & f(A) := \text{rank of } A \\ s.t. \quad & A \in C, \end{aligned}$$

where C is a subset of $M_{m,n}(R)$(the vector space of m by n real matrices).

Theorem (Hiriart-Urruty, 2011)

Every admissible point in the above problem is a local minimizer.

Comment: The result is somehow strange: the various generalized subdifferentials, including Clarkes one, all coincide and, of course, contain the zero element.

Min-max optimality conditions

Consider a problem:

$$\min_{x \in X} \max_{i \in I} f_i(x),$$

where X is a convex region id a d-dimensional Euclidian space R^n, I is a finite index set, and the $f_i(x)$ are continuous functions over X. A subset Z of X is called an extreme subset of X if

$$\left.\begin{array}{r} x, y \in X \\ \lambda x + (1-\lambda) y \in Z \text{ for some } 0 < \lambda < l \end{array}\right\} \Rightarrow \; x, y \in Y$$

Min-max optimality conditions

Theorem (Du, Pardalos, 1993)

Let $f(x,y)$ be a continuous function on $X \times Y$ where X is a polytope in R^m and Y is a compact set in R^n. Let $g(x) = \max_{y \in Y} f(x,y)$. If (x,y) is concave with respect to x, then the minimum value of $g(x)$ over X is achieved at some point $\hat{x}$ satisfying the following condition:
(*) There exists an extreme subset Z of X such that $\hat{x} \in Z$ and the set $I(\hat{x}) = \{y | g(\hat{x}) = f(\hat{x}, y)\}$ is maximal over Z.

General localization theorem

Theorem (Georgiev, Pardalos, Chinchuluun, 2008 - [4])

Suppose that Y is a compact topological space, $X \subset R^n$ is a compact polyhedron, $f : X \times Y \to R$ is a continuous function and $f(\cdot, y)$ is concave for every $y \in Y$. Define $g(x) = \max_{y \in Y} f(x, y)$. Then the minimum of g over X is attained at some generalized vertex $\hat{x}$.

Optimality Conditions for Lipschitz functions

Consider following notations:

- E - Banach space with dual E^*;
- $\|\cdot\|$ and $\|\cdot\|_*$ - the norms in E and E^* respectively;
- B_E and B_{E^*} - closed unit balls of E and E^* respectively;
- D - closed and nonempty subset of E.

Clarke's subdifferential

Consider a locally Lipschitz function $f : E \to \mathbb{R}$. The Clarke's subdifferential is given by (Clarke, 1990 - [2]):

$$\partial f(x) = \left\{ x^* \in E^* : \langle x^*, h \rangle \leq f^\circ(x; h) \ \forall h \in E \right\},$$

where

$$f^\circ(x; h) = \limsup_{u \to x, t \downarrow 0} \frac{f(u + th) - f(u)}{t}$$

is the Clarke directional derivative of f at x in the direction h.

Clarke's regularity

If the regular directional derivative of f at x in direction h

$$f'(x;h) = \lim_{t \downarrow 0} \frac{f(x+th) - f(x)}{t}$$

exists and $f'(x;h) = f^{\circ}(x;h)$ for every $h \in E$, then f is called *regular* in sense of Clarke.

Optimality Conditions

Suppose the following locally Lipschitz functions are given:

$$g_i : E \to \mathbb{R}, \ \ i \in I \cup \{0\},$$
$$h_j : E \to \mathbb{R}, \ \ j \in J.$$

where

$$I = \{1, 2, ..., n\}$$
$$J = \{1, 2, ..., m\}.$$

Optimality Conditions

Consider the following minimization problem:

$$\begin{cases} \text{minimize} & g_0(x) \\ \text{subject to} & x \in D,\ g_i(x) \le 0,\ h_j(x) = 0\ (i \in I, j \in J). \end{cases} \tag{1}$$

Optimality Conditions

Denote $N_D(x)$ the generalized normal cone in the sense of Clarke at $x \in D$ (Clarke, 1990):

$$N_D(x) = \{x^* \in E^* : \exists t > 0, x^* \in t\partial d(x, D)\},$$

where $d(., D$ is the distance function:

$$d(x, D) = \inf\{\|x - y\| : y \in D\}.$$

When D is convex, then $N_D(x)$ coincides with the classical normal cone in the sense of convex analysis $N(D, x)$ [2].

Optimality Conditions

Theorem (Georgiev, Chinchuluun, Pardalos, 2011 - [3])

Consider the following minimization problem:

$$\begin{cases} \text{minimize } f(x) \\ \text{subject to } x \in D, \end{cases} \tag{2}$$

where $f : D \to \mathbb{R}$ is a locally Lipschitz function, D is a nonempty closed convex subset of a Banach space E. Denote the level set of f at c on D by

$$D_c(f) = \{y \in D : f(y) = c\}.$$

Optimality Conditions

Theorem

a) **Sufficient condition for a global minimum**.
If the following qualification condition holds:

$$\partial f(y) \cap N(D, y) = \emptyset \quad \forall y \in D_{f(z)}(f), \tag{3}$$

then the condition

$$-\partial f(y) \subset N(D, y) \ \ \forall y \in D_{f(z)}(f) \tag{4}$$

is sufficient for $z \in D$ to be a global solution to Problem (2).
b) **Necessary condition for a global minimum**.
If $-f$ is regular in sense of Clarke, then the condition (4) is necessary for $z \in D$ to be a global solution to Problem (2).

Special Case

The function f is concave and the space is finite dimensional and Euclidean. Then the condition (4) was refined (Strekalovsky, 1987) by replacing the set $D_{f(z)}(f)$ with the following set:

$$D_{f(z)}^{concave}(f) = \{y \in D : f(y) = f(z),\ y \text{ is an extreme point of } D\}.$$

This refinement reflects the fact that the minimum of a concave function over a convex set is attained at an extreme point.

Special Case

The function f is a maximum of concave functions:

$$f(x) = \max_{y \in Y} g(x, y),$$

where $g(., y)$ is concave for every $y \in Y$, Y being a compact topological space and

$$g : X \times Y \to \mathbb{R}$$

is continuous.

Open Question

Is it possible to refine the condition (4) in sense to replace the set $D_{f(z)}(f)$ with the following set

$$D_{f(z)}^{max,conc}(f) = \{y \in D : f(y) = f(z),\ y \text{ is a generalized vertex of } D\}?$$

The motivation for this question is the following result (Georgiev, Chinchuluun, Pardalos, 2008 - [4]): the global minimum of f over D is attained at some generalized vertex of D.

References

A.A. Ahmadi, A. Olshevsky, P.A. Parrilo, and J.N. Tsitsiklis.
Np-hardness of deciding convexity of quartic polynomials and related problems.
Mathematical Programming, pages 1–24, 2011.

F.H. Clarke.
Optimization and nonsmooth analysis, volume 5.
Society for Industrial Mathematics, 1990.

P. Georgiev, A. Chinchuluun, and P. Pardalos.
Optimality conditions of first order for global minima of locally lipschitz functions.
Optimization, 60, 1(2):277–282, 2011.

P.G. Georgiev, P.M. Pardalos, and A. Chinchuluun.
Localization of minimax points.
Journal of Global Optimization, 40(1):489–494, 2008.

R. Horst, P.M. Pardalos, and N.V. Thoai.
Introduction to global optimization.
Springer, 2000.

P.M. Pardalos and G. Schnitger.
Checking local optimality in constrained quadratic programming is np-hard.
Operations Research Letters, 7(1):33–35, 1988.

P.M. Pardalos and S.A. Vavasis.
Open questions in complexity theory for numerical optimization.
Mathematical Programming, 57(1):337–339, 1992.