Structure-Aware Global Mixed-Integer Nonlinear Optimisation

Ksenia Bestuzheva

The London School of Economics and Political Science
25 February 2026

Global Optimisation: Problem Statement

Optimisation problem:

$$\begin{aligned} &\min f(\mathbf{x}) && \text{(objective function)} \\ &\text{s.t. } g_k(\mathbf{x}) \leq 0 \ \forall k \in \mathcal{C}, && \text{(constraints)} \\ &\quad \mathbf{x}_i \in \mathbb{R} \ \forall i \in \mathcal{I}_C, \ \mathbf{x}_i \in \mathbb{Z} \ \forall i \in \mathcal{I}_I, \ \mathbf{x}_i \in \{0,1\} \ \forall i \in \mathcal{I}_B. && \text{(continuous, integer, binary variables)} \end{aligned}$$

- f, g differentiable
- Constraint qualifications satisfied (no degenerate behaviour of constraint gradients)

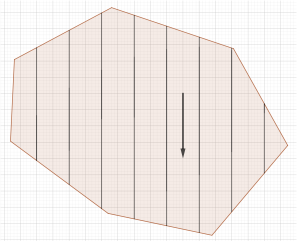
My work is concerned with problems that are **nonlinear** (including **nonconvex**) and/or **mixed-integer**.

Global optimisation seeks global optima (optimal for the entire feasible set) rather than local optima.

Main Challenges

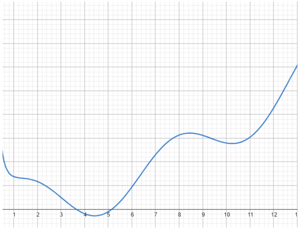
Mixed-integer programs:

- Cannot follow a continuous improving path
- Exponential number of combinations of discrete variable values



Nonconvex nonlinear programs:

- Local optima that are not global
- Difficult to certify infeasibility



How to **prove optimality**, or compute how far from optimal a solution is?
How to **find good feasible solutions** fast?

Optimality Conditions

More commonly employed in continuous optimisation.

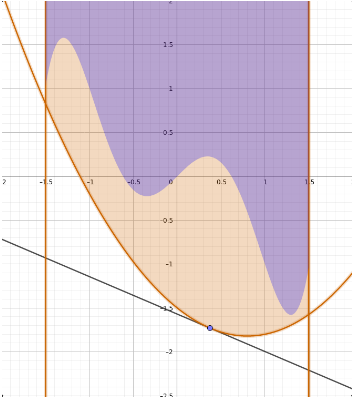
First-order conditions:
Karush-Kuhn-Tucker conditions are satisfied at a feasible point x^* if there exists a nonnegative multiplier λ such that:

$$-\nabla f(x^*) = \sum_{k \in \mathcal{C}} \lambda_k \nabla g_k(x^*)$$
$$\lambda_i g_i(x^*) = 0 \ \forall i \in \mathcal{C}$$

KKT conditions are necessary for local optimality and, if the problem is convex, sufficient for global optimality.

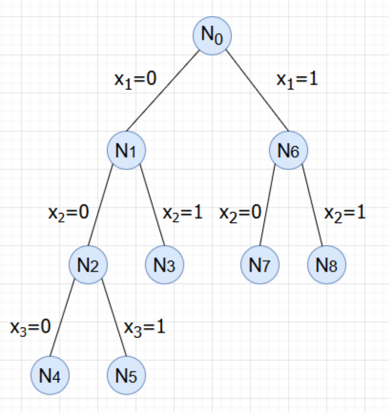
Optimality conditions for various problem classes are an active research area.

Relaxations



- Relaxations are sets containing the feasible set
- Usually convex or linear
- Yield **lower bounds** on the objective
- Tighter relaxation $\Rightarrow$ better bound
- Relaxations are often iteratively strengthened by adding **cutting planes** (valid linear inequalities)

Branch-and-Bound



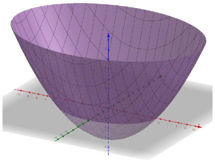
(Figure created with the tool `https://app.diagrams.net`)

- Global optimisation approach for mixed-integer and/or nonlinear nonconvex problems
- Recursively **divides the feasible set**, subproblems form a tree
- Feasible solutions $\rightarrow$ upper bounds, relaxations $\rightarrow$ lower bounds
- Primal heuristics to find good feasible solutions early
- Continues until gap between upper and lower bounds is closed or below tolerance

On Local Optima of Box-Constrained Quadratic Programs

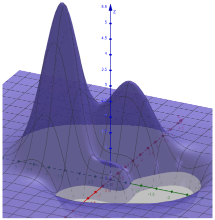
Generalised Convexity: Motivation

- Nonconvexity $\rightarrow$ local optima and non-optimal KKT points, optimality/infeasibility difficult to prove



convex

or



nonconvex

- Some nonconvex problems share certain properties with convex problems
- Empirical example: for some optimal power flow and pooling problems, solvers tend to converge to the global optimum
- Research on generalised convexity is concerned with leveraging such similarities
- **Goal:** tackle nonconvexity by understanding what makes many nonconvex problems challenging

Invex Functions and Problems

Definition

Function f is invex if: $\exists \eta \mid \eta(x, y)\nabla f(x) \leq f(y) - f(x)$ for all $x, y \in X$
(Hanson, 1981)

Martin, 1985: relaxed Hanson's conditions, definition of KKT-invex problems:

Definition

An optimisation problem is KKT-invex if there exists a function η such that:

$$g(x) \leq 0,\ g(u) \leq 0 \ \Rightarrow$$
$$\begin{cases} f(x) - f(u) - \nabla f(u) \cdot \eta(x, u) \geq 0, \\ \text{for } i \in \mathcal{C},\ \text{if } g_i(u) = 0,\ \text{then } \ -\nabla g_i(u) \cdot \eta(x, u) \geq 0. \end{cases}$$

Every KKT point is a global optimum if and only if the problem is KKT-invex.

Boundary-Invexity

New approach to KKT-invexity, based on a special subclass of stationary points of subproblems that indicate a violation of KKT-invexity.
(Bestuzheva, Hijazi, 2019)

- Boundary-invexity is sufficient for KKT-invexity of problems in $\mathbb{R}^2$ and a weaker version of boundary-invexity is necessary for KKT-invexity of problems in $\mathbb{R}^n$
- **Can be verified algorithmically**
- Used to show that optimal power problems for one line with realistic parameters are KKT-invex

Further Shift of Perspective: Optima-Invexity

Definition

A problem is optima-invex if it has a unique connected locally optimal subset or a unique connected set of unbounded directions.

KKT-invexity	**Optima-invexity**
Allows multiple basins of attraction (if optima are equivalent)	Does not allow multiple basins of attraction
Does not allow non-optimal KKT points	Allows some non-optimal KKT points

- Focus on local optima/unbounded directions and their basins of attraction
  - In practice, convergence to non-optimal KKT points tends to not be an issue (Lee et al., 2016)
  - Optima-invexity serves as a basis for studying the relation between non-optimal KKT points and local optima

Conditions for Optima-Invexity

Conjecture

If an optimisation problem does not have non-optimal KKT points, then it is optima-invex.

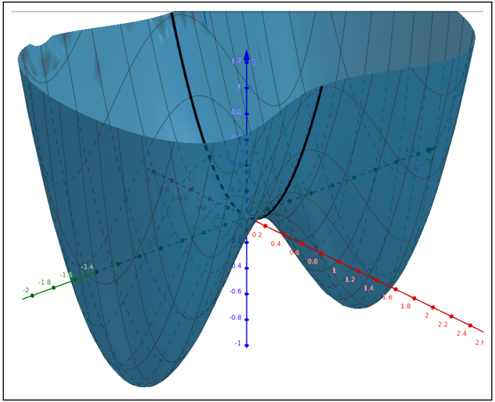
That is, the absence of non-optimal KKT points implies the absence of multiple distinct local optima or sets of unbounded directions.

Proof Approach - Unconstrained Case

Step 1: Assume the contrary: there exist distinct local optima $\mathbf{x}^{1*}$, $\mathbf{x}^{2*}$.

Step 2: consider basins of attraction A_1, A_2 of $\mathbf{x}^{1*}$, $\mathbf{x}^{2*}$ respectively, and the shared boundary δA.
Prove that $\delta A \not\subset A_i$ for $i = 1, 2$.

Step 3: consider $x^B \in \delta A$. Two possibilities:
- x^B is a non-optimal stationary point or
- x^B belongs to a basin of attraction of a non-optimal stationary point.

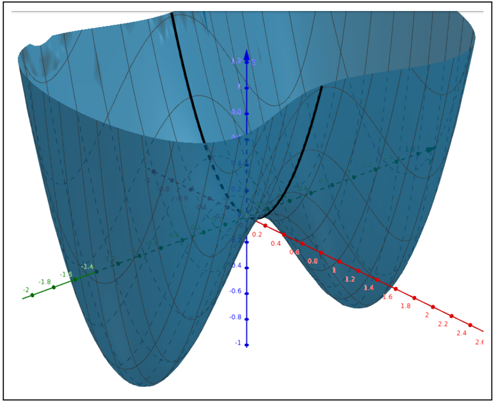


Ridges

Definition

A **ridge** is a hypersurface separating different basins of attraction w.r.t. a given algorithm.

Important for algorithms: ridges tell us how to partition spaces.



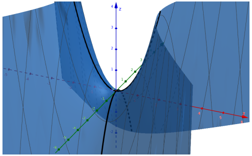
Optima-Invexity of Unconstrained Quadratic Problems

Nonconvex quadratic function (that has been diagonalised): $f(x) = \lambda_1 x_1^2 + \lambda_2 x_2^2$.

Unbounded directions of f: $\mathcal{U} = \left\{ x \mid \sum_{i \in \mathcal{I}^-} \lambda_i x_i^2 + \sum_{i \in \mathcal{I}^+} \lambda_i x_i^2 < 0 \right\}$

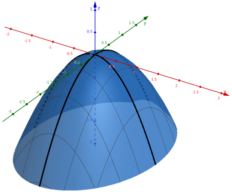
Case 1: one negative eigenvalue $\Rightarrow$ not optima-invex.

- $\mathcal{S}^+$: $(n-1)$-dimensional subspace spanned by e_i, $i \in \mathcal{I}^+$
- $\mathcal{U} \cap \mathcal{S}^+ = \emptyset$, therefore $\mathcal{U}$ contains two disjoint cones



Case 2: more than one negative eigenvalue $\Rightarrow$ optima-invex.

- Consider the subspace $\mathcal{S}^-$ spanned by e_i, $i \in \mathcal{I}^-$
- There is always a continuous path between any $x^{(1)}, x^{(2)} \in \mathcal{U}$
  - Project into $\mathcal{S}^-$ $\rightarrow$ rotate in $\mathcal{S}^-$ $\rightarrow$ lift

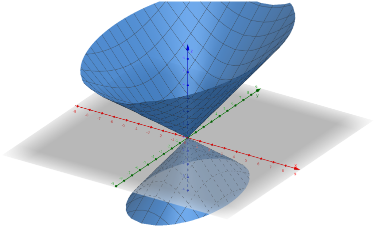


Unconstrained Case: Examples in $\mathbb{R}^3$

$f(x_1, x_2, x_3) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \lambda_3 x_3^2$

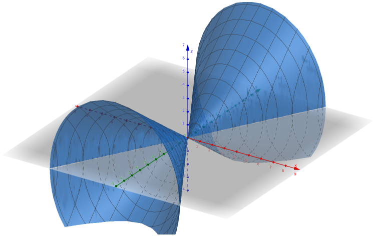
Case 1: one negative eigenvalue ($\lambda_3 < 0$).

- Unbounded rays lie in the two disjoint cones
- $\mathcal{S}^+$ is the (x_1, x_2)-plane
- Not optima-invex



Case 2: more than one negative eigenvalue ($\lambda_1, \lambda_3 < 0$).

- Unbounded rays outside of the cone
- $\mathcal{S}^-$ is the (x_1, x_3)-plane
- Optima-invex



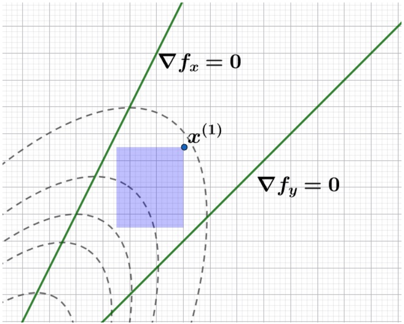
Optima-Invexity of Box-Constrained Concave Quadratic Problems

Consider the problem: $\min f(x)$ s.t. $x \in \mathcal{B}$, where $\mathcal{B} = \underline{x} \leq x \leq \overline{x}$.
Let $p_i = \{x \mid \nabla f_i(x) = 0\}$

Optima-invexity sufficient condition 1:
if $\mathcal{B} \cap p_i = \emptyset \ \forall i$, then the problem is optima-invex.

Proof outline: use the fact that f increases in the feasible direction along all edges at any local optimum.

Algorithmic use: can be applied directly to prove optima-invexity of some problems.



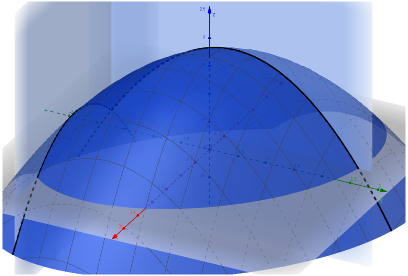
Optima-Invexity of Box-Constrained Concave Quadratic Problems

Optima-invexity sufficient condition 2:

The problem is optima-invex if the following conditions hold:

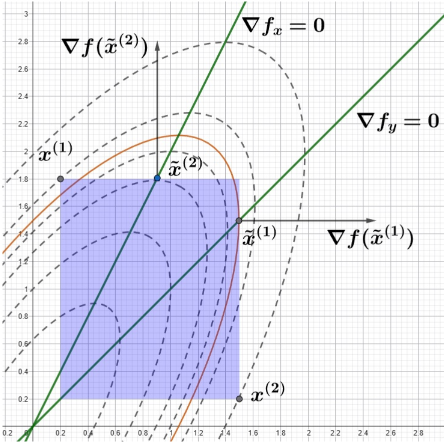
1) for each i, j either
   a) $\text{int}(\mathcal{B}) \cap p_i \cap p_j = \emptyset$ or
   b) $\mathcal{B} \cap p_i \cap p_j \neq \emptyset$ for each feasible fixing of x_k, $k \neq i, j$

2) the edges of $\mathcal{B}$ do not contain non-optimal KKT points.



Sufficient Condition 2: Proof Outline

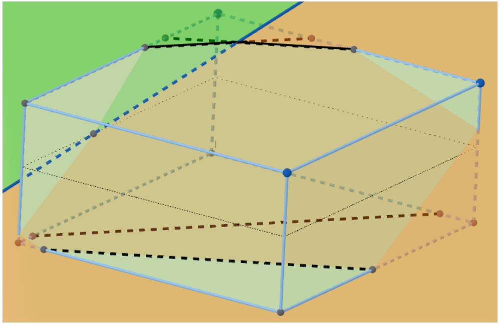
- Local optima x^{*1} and x^{*2}: vertices of $\mathcal{B}$ due to concavity
- Suppose that $x_i^{*1} \neq x_i^{*2}$ $\forall i = 1, \ldots, m$
- Consider a path between x^{*1} and x^{*2} along edges $e_1, \ldots, e_m$
- The path must cross all $p_1, \ldots, p_m$
- Choose the order of $e_1, \ldots, e_m$ so that is matches the order of crossing $p_1, \ldots, p_m$ (using (1a) and (1b))
- An edge e_k exists that crosses p_k: the crossing point is KKT



(Note: each p_i is crossed an odd number of times. The question is: is p_i crossed an even or odd number of times before p_j is first crossed?)

Algorithmic Use of Sufficient Condition 2

- More expensive to directly apply than sufficient condition 1
- Provides guidelines on how to divide $\mathcal{B}$ to direct the search towards new local optima
- To check property 1: solve linear systems or programs
- To check property 2: solve nonlinear programming subproblems
- If property 1 is violated, partition the domain (e.g. in a branch-and-bound approach)



Practical Uses of Optima-Invexity

- Optimality proofs
  - Challenging in general, may be feasible in special cases
  - Optimality in subproblems: e.g. node subproblem in branch-and-bound / subproblem in a decomposition
- Heuristics (standalone or within branch-and-bound)
  - Use invexity-breaking points to guide the search towards new local optima
  - Can be applied to problems for which there is yet no formal proof of sufficiency

Constraint-Wise Algorithm

First heuristic prototype:

- Rather than compute non-optimal KKT points, consider each relevant constraint and variable separately
- Solve computationally cheap subproblems
- Use solutions to construct a grid
- Begin a new local search in each cell of the grid

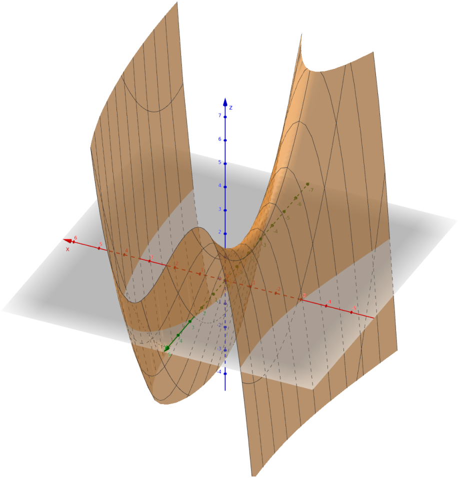
Computationally fast and easy to implement, but may result in suboptimal grids.

Example Problem

ex14_1_1 from MINLPLib
(from Floudas et al., Handbook of Test Problems in Local and Global Optimization)

Formulation:

$$\min\ z \text{ s.t.}$$
$$2y^2 + 4xy - 42x + 4x^3 - 14 <= z$$
$$-2y^2 - 4xy + 42x - 4x^3 + 14 <= z$$
$$2x^2 + 4xy - 26y + 4y^3 - 22 <= z$$
$$-2x^2 - 4xy + 26y - 4y^3 + 22 <= z$$
$$(x, y) \in [-5, 5] \times [-5, 5]$$



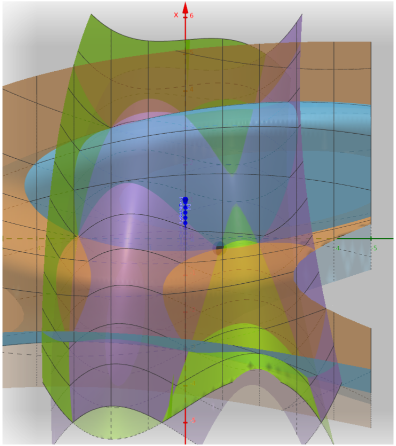
Multiple local optima (with the same value)

Example Problem

ex14_1_1 from MINLPLib
(from Floudas et al., Handbook of Test Problems in Local and Global Optimization)

Formulation:

$$
\begin{aligned}
&\min\ z \text{ s.t.} \\
&2y^2 + 4xy - 42x + 4x^3 - 14 <= z \\
&-2y^2 - 4xy + 42x - 4x^3 + 14 <= z \\
&2x^2 + 4xy - 26y + 4y^3 - 22 <= z \\
&-2x^2 - 4xy + 26y - 4y^3 + 22 <= z \\
&(x, y) \in [-5, 5] \times [-5, 5]
\end{aligned}
$$



Multiple local optima (with the same value)

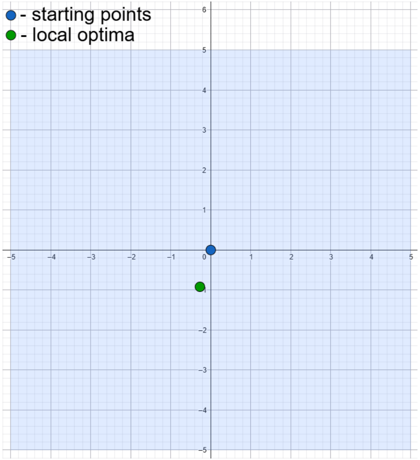
Algorithm Applied to Example

Grid: $[-5,5] \times [-5,5]$
Starting point: (0,0,5)
NLP solution: (-0.27,-0.92,0)
Active constraints: (1), (2), (3), (4)

Constraint (1):

- Convex in y, nonconvex in x $\Rightarrow$ focus on x
- Fix $y = 5$, find local optimum (x^*, z^*) of: $\max z$ s.t. $g_1(x, 5, z) = 0$
- Add separation line: $x = x^* = -1.36$

Similarly find separation lines for remaining constraints.



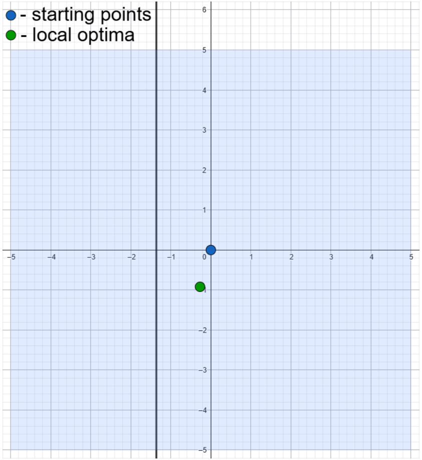
Algorithm Applied to Example

Grid: $[-5, 5] \times [-5, 5]$
Starting point: (0,0,5)
NLP solution: (-0.27,-0.92,0)
Active constraints: (1), (2), (3), (4)

Constraint (1):

- Convex in y, nonconvex in x $\Rightarrow$ focus on x
- Fix $y = 5$, find local optimum (x^*, z^*) of: $\max z$ s.t. $g_1(x, 5, z) = 0$
- Add separation line: $x = x^* = -1.36$

Similarly find separation lines for remaining constraints.



Algorithm Applied to Example

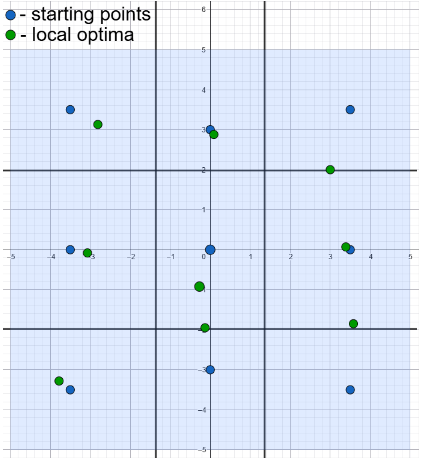
Grid: $[-5, 5] \times [-5, 5]$
Starting point: (0,0,5)
NLP solution: (-0.27,-0.92,0)
Active constraints: (1), (2), (3), (4)

Constraint (1):

- Convex in y, nonconvex in x $\Rightarrow$ focus on x
- Fix $y = 5$, find local optimum (x^*, z^*) of: $\max z$ s.t. $g_1(x, 5, z) = 0$
- Add separation line: $x = x^* = -1.36$

Similarly find separation lines for remaining constraints.

Result: 9 cells, 9 local optima found.



Summary and Conclusions

- **New approach to nonconvex optimisation**: understand when local optima occur $\rightarrow$ develop efficient algorithms
- **New notion**: optima-invexity
- **Conjecture**: absence of non-optimal KKT points $\Rightarrow$ optima-invexity
- **Sufficient conditions** for optima-invexity of unconstrained quadratic programs
- **Sufficient conditions** for optima-invexity of box-constrained concave quadratic programs
- Invexity-guided **heuristic algorithm** prototype

References

- Hanson, On sufficiency of the Kuhn-Tucker conditions, *Journal of Mathematical Analysis and Applications*, 80 (1981), pp. 545–550.
- Martin, The essence of invexity, *Journal of optimization Theory and Applications*, 47 (1985), pp. 65–76.
- Bestuzheva, Hijazi. Invex optimization revisited. *Journal of Global Optimization*, 74(4) (2019), pp. 753–782.
- Lee et al. Gradient descent only converges to minimizers. *Conference on Learning Theory* (2016).
- Floudas et al. Handbook of test problems in local and global optimization. *Springer Science & Business Media*, vol. 33 (2013).

All figures were created by myself with the use of the tool (unless specified otherwise) Geogebra.