Optimization of quadratic forms and t-norm forms on interval domains and computational complexity

Michal Černý[1], Milan Hladík[2] and Vladik Kreinovich[3]

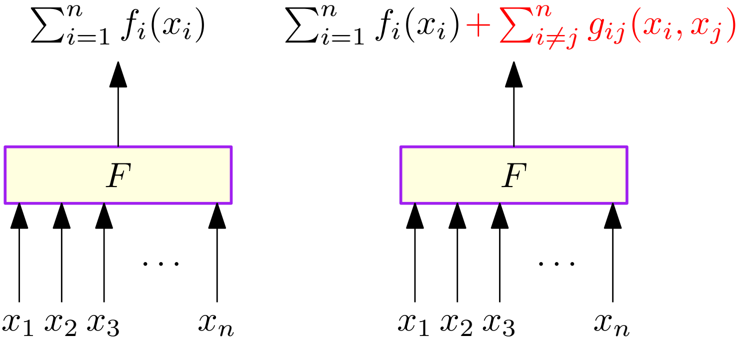
[1] University of Economics, Prague, Czech Republic

[2] Charles University, Prague, Czech Republic

[3] University of Texas at El Paso, Texas, USA

7th World Conference on Soft Computing, Baku, 2018

Problem formulation



separable case **quadratic interactions**

Model for imprecision of inputs.

- Instead of x_i we can observe only bounds $\underline{x}_i, \overline{x}_i$ s.t.

$$\underline{x}_i \leqslant x_i \leqslant \overline{x}_i, \quad i = 1, \dots, n.$$

Problem formulation (contd.)

Model for imprecision of inputs.

- Instead of x_i we can observe only bounds $\underline{x}_i, \overline{x}_i$ s.t.

$$\underline{x}_i \leqslant x_i \leqslant \overline{x}_i, \quad i = 1, \ldots, n.$$

The main question.

- Given $F : \mathbb{R}^n \to \mathbb{R}$, can we find bounds on $F(x_1, \ldots, x_n)$ given the observable bounds $[\underline{x}_1, \overline{x}_1], \ldots, [\underline{x}_n, \overline{x}_n]$?

More formally. The task is to compute

$$\overline{F} = \max\{F(x_1, \ldots, x_n) \mid x_i \in [\underline{x}_i, \overline{x}_i], i = 1, \ldots, n\},$$
$$\underline{F} = \min\{F(x_1, \ldots, x_n) \mid x_i \in [\underline{x}_i, \overline{x}_i], i = 1, \ldots, n\}.$$

Some well-known results

- **Theorem (the general case).** For a general function F, the bounds $\underline{F}, \overline{F}$ are nonrecursive.

Some well-known results

- **Theorem (the general case).** For a general function F, the bounds $\underline{F}, \overline{F}$ are nonrecursive.
- **Observation (the separable case).** For the separable case

$$F(x_1, \ldots, x_n) = \sum_{i=1}^{n} f_i(x_i),$$

the bounds reduce to

$$\underline{F} = \sum_{i=1}^{n} \underline{f}_i(x_i), \quad \overline{F} = \sum_{i=1}^{n} \overline{f}_i(x_i),$$

where $\underline{f}_i = \min_{\underline{x}_i \leqslant \xi \leqslant \overline{x}_i} f_i(\xi)$ and $\overline{f}_i = \max_{\underline{x}_i \leqslant \xi \leqslant \overline{x}_i} f_i(\xi)$.

The case with quadratic interactions

The case with quadratic interactions

- **The general form:** $F = \sum_i f_i(x_i) + \sum_{j \neq i} g_{ij}(x_i, x_j)$

The case with quadratic interactions

- **The general form:** $F = \sum_i f_i(x_i) + \sum_{j \neq i} g_{ij}(x_i, x_j)$
- **A natural example — a quadratic form:**

$$F = \sum_i q_{ii} x_i^2 + \sum_{j \neq i} q_{ij} x_i x_j = x^T Q x,$$

where $Q = (q_{ij})_{i,j=1,\ldots,n}$ and $x = (x_1, \ldots, x_n)^T$.

The case with quadratic interactions

- **The general form:** $F = \sum_i f_i(x_i) + \sum_{j \neq i} g_{ij}(x_i, x_j)$
- **A natural example — a quadratic form:**

$$F = \sum_i q_{ii} x_i^2 + \sum_{j \neq i} q_{ij} x_i x_j = x^T Q x,$$

where $Q = (q_{ij})_{i,j=1,\ldots,n}$ and $x = (x_1, \ldots, x_n)^T$.

- **Classical result:** If Q is psd, then $\underline{F}$ is **computable in polynomial time** (via IPMs for convex quadratic programming).

The case with quadratic interactions

- **The general form:** $F = \sum_i f_i(x_i) + \sum_{j \neq i} g_{ij}(x_i, x_j)$
- **A natural example — a quadratic form:**

$$F = \sum_i q_{ii} x_i^2 + \sum_{j \neq i} q_{ij} x_i x_j = x^T Q x,$$

  where $Q = (q_{ij})_{i,j=1,\dots,n}$ and $x = (x_1, \dots, x_n)^T$.
- **Classical result:** If Q is psd, then $\underline{F}$ is **computable in polynomial time** (via IPMs for convex quadratic programming).
- **On the contrary:** For Q psd, computation of $\overline{F}$ is an **NP-hard problem**.

The case with quadratic interactions

- **The general form:** $F = \sum_i f_i(x_i) + \sum_{j \neq i} g_{ij}(x_i, x_j)$
- **A natural example — a quadratic form:**

$$F = \sum_i q_{ii} x_i^2 + \sum_{j \neq i} q_{ij} x_i x_j = x^T Q x,$$

  where $Q = (q_{ij})_{i,j=1,\ldots,n}$ and $x = (x_1, \ldots, x_n)^T$.
- **Classical result:** If Q is psd, then $\underline{F}$ is **computable in polynomial time** (via IPMs for convex quadratic programming).
- **On the contrary:** For Q psd, computation of $\overline{F}$ is an **NP-hard problem**.

Our general goal: Inspect further complexity-theoretic results.

Sparse quadratic forms

Definition. The quadratic form $F(x) = x^T Qx$ is sparse is there are "many" zeros in the matrix Q.

Sparse quadratic forms

Definition. The quadratic form $F(x) = x^T Qx$ is sparse is there are "many" zeros in the matrix Q.

Problem. How "sparse" must the matrix Q be to make the values $\underline{F}, \overline{F}$ polynomially computable?

Sparse quadratic forms

Definition. The quadratic form $F(x) = x^T Qx$ is sparse is there are "many" zeros in the matrix Q.

Problem. How "sparse" must the matrix Q be to make the values $\underline{F}, \overline{F}$ polynomially computable?

Solution.

- **Theorem 1.** If there are at most $O(\log n)$ nonzero off-diagonal entries in Q, then both bounds $\underline{F}, \overline{F}$ are computable in polynomial time.

Sparse quadratic forms

Definition. The quadratic form $F(x) = x^T Qx$ is sparse is there are "many" zeros in the matrix Q.

Problem. How "sparse" must the matrix Q be to make the values $\underline{F}, \overline{F}$ polynomially computable?

Solution.

- **Theorem 1.** If there are at most $O(\log n)$ nonzero off-diagonal entries in Q, then both bounds $\underline{F}, \overline{F}$ are computable in polynomial time.
- **Theorem 2.** But: Even if the matrix Q is psd and there are $\Omega(n^\varepsilon)$ non-zero off-diagonal entries in Q, for an arbitrarily small $\varepsilon > 0$, then computation of $\overline{F}$ is NP-hard.

(In)approximability

Let us refine the results:

(In)approximability

Let us refine the results:

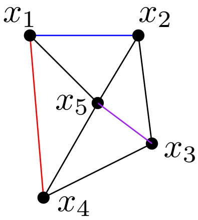
- **Corollary (absolute approximation).** When we have $\Omega(n^{\varepsilon})$ non-zero off-diagonal entries, then $\overline{F}$ is inapproximable with an arbitrarily large absolute error.
- **Relative approximation.** A result by Nesterov (1998) implies that the problem can be approximated with some "reasonable" relative error by semidefinite relaxation.

The quadratic form as a graph

The Q-graph of a quadratic form $x^T Qx$:

- Assume that Q is triangular (without loss of generality).
- Define the Q-graph as follows:
  - vertices: variables $x_1, \ldots, x_n$,
  - edges: $\{x_i, x_j\}$ is an edge if $i \neq j$ and $q_{ij} \neq 0$,
  - weights of edges: the weight of the edge is q_{ij}.

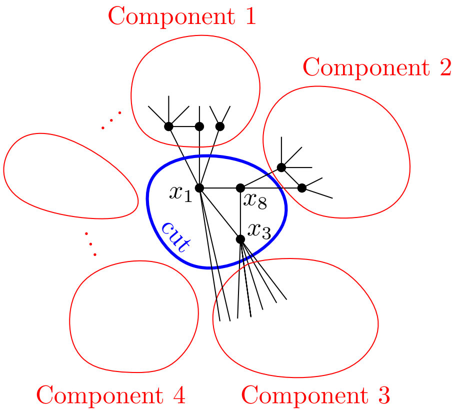
$$Q = \begin{bmatrix} q_{11} & 2 & 0 & 1 & 1 \\ & q_{22} & 1 & 0 & 3 \\ & & q_{33} & 1 & 1 \\ & & & q_{44} & 2 \\ & & & & q_{55} \end{bmatrix}$$



Sunflower graph

Special shapes of the Q-graph...

- **Sunflower graph G:** There exists a cut C of vertex size $O(\log n)$ such that $G \setminus C$ has components of vertex size $O(\log n)$.
- **Theorem.** If the Q-graph is a sunflower graph, then $\overline{F}$ is computable in polynomial time.



Further special forms of Q-graphs

Further polynomially solvable cases:

- Q-graph is a tree

Further special forms of Q-graphs

Further polynomially solvable cases:

- Q-graph is a tree
- Q-graph is planar, with $O(\log n)$ faces

Further special forms of Q-graphs

Further polynomially solvable cases:

- Q-graph is a tree
- Q-graph is planar, with $O(\log n)$ faces
- Q-graph is a bipartite graph, where one of the partites has size $O(\log n)$

Further special forms of Q-graphs

Further polynomially solvable cases:

- Q-graph is a tree
- Q-graph is planar, with $O(\log n)$ faces
- Q-graph is a bipartite graph, where one of the partites has size $O(\log n)$
- (few negative coefficients) there exists a cut C of size $O(\log n)$, such that all variables incident with negative coefficients are in C

Research problems

Research problem.

- Many properties of the Q-graph, needed for poly-time algorithms, are defined by nonconstructive existential quantification.

Research problem.

- Many properties of the Q-graph, needed for poly-time algorithms, are defined by nonconstructive existential quantification.
  - "There exists a cut which is small and the removal of which leaves only small components."
  - "There exists a drawing of the graph in $\mathbb{R}^2$ where the number of faces is small".

Research problem.

- Many properties of the Q-graph, needed for poly-time algorithms, are defined by nonconstructive existential quantification.
  - "There exists a cut which is small and the removal of which leaves only small components."
  - "There exists a drawing of the graph in $\mathbb{R}^2$ where the number of faces is small".
- Is it possible to verify the existential property in polynomial time, at least on some classes of Q-graphs?

More general forms of quadratic interactions

To recall: The general form is

$$F = \sum_i f_i(x_i) + \sum_{j \neq i} g_{ij}(x_i, x_j).$$

More general forms of quadratic interactions

To recall: The general form is

$$F = \sum_i f_i(x_i) + \sum_{j \neq i} g_{ij}(x_i, x_j).$$

We generalize the quadratic interaction terms:

- So far we considered the quadratic terms $g_{ij}(x_i, x_j) = q_{ij} x_i x_j$ (i.e., multiplication).
- Now, let us generalize the multiplication term to be a t-norm form:

$$F = \sum_i f_i(x_i) + \sum_{j \neq i} t_{ij}(x_i, x_j).$$

What is a t-norm?

To recall: Definition. A t-norm is a function $t : \mathbb{R}^2 \to \mathbb{R}$ generalizing multiplication in the following way:

- (commutativity) $t(a, b) = t(b, a)$,
- (monotonicity) $a \leqslant a', b \leqslant b' \Rightarrow t(a, b) \leqslant t(a', b')$;
- (associativity) $t(a, t(b, c)) = t((a, b), c)$;
- (1 is identity) $t(a, 1) = a$.

Comment.

- t-norm indeed generalizes multiplication.
- It is also known from logic as an extension of the AND-connective for other arguments than $\{0, 1\}^2$.

Examples of t-norms

- $t(a, b) = ab$ (product t-norm),

Examples of t-norms

- $t(a, b) = ab$ (product t-norm),
- $t(a, b) = \min\{a, b\}$ (minimum t-norm),

Examples of t-norms

- $t(a, b) = ab$ (product t-norm),
- $t(a, b) = \min\{a, b\}$ (minimum t-norm),
- $t(a, b) = \max\{a + b - 1, 0\}$ (Lukasiewicz t-norm),

Examples of t-norms

- $t(a, b) = ab$ (product t-norm),
- $t(a, b) = \min\{a, b\}$ (minimum t-norm),
- $t(a, b) = \max\{a + b - 1, 0\}$ (Lukasiewicz t-norm),
- $t(a, b) = \begin{cases} \min\{a, b\} & \text{if } a + b > 1, \\ 0 & \text{otherwise,} \end{cases}$ (nilpotent t-norm),

t-norms: Examples

Examples of t-norms

- $t(a,b) = ab$ (product t-norm),
- $t(a,b) = \min\{a,b\}$ (minimum t-norm),
- $t(a,b) = \max\{a+b-1, 0\}$ (Lukasiewicz t-norm),
- $t(a,b) = \begin{cases} \min\{a,b\} & \text{if } a+b>1, \\ 0 & \text{otherwise,} \end{cases}$ (nilpotent t-norm),
- $t(a,b) = \begin{cases} 0 & \text{if } a=b=0, \\ \frac{ab}{a+b-ab} & \text{otherwise} \end{cases}$ (Hamacher t-norm),

Examples of t-norms

- $t(a,b) = ab$ (product t-norm),
- $t(a,b) = \min\{a,b\}$ (minimum t-norm),
- $t(a,b) = \max\{a+b-1,0\}$ (Lukasiewicz t-norm),
- $t(a,b) = \begin{cases} \min\{a,b\} & \text{if } a+b>1, \\ 0 & \text{otherwise,} \end{cases}$ (nilpotent t-norm),
- $t(a,b) = \begin{cases} 0 & \text{if } a=b=0, \\ \frac{ab}{a+b-ab} & \text{otherwise} \end{cases}$ (Hamacher t-norm),
- $t(a,b) = \begin{cases} \min\{a,b\} & \text{if } \max\{a,b\}=1, \\ 0 & \text{otherwise} \end{cases}$ (drastic t-norm).

The main (negative) result on quadratic interactions with t-norms.

- **Theorem.** Let t_{ij} be arbitrary fixed Lipschitz continuous t-norms. For the t-norm form

$$F(x_1, \ldots, x_n) = \sum_{i=1}^{n} f_i(x_i) + \sum_{j \neq i} t_{ij}(x_i, x_j)$$

with interval data $\underline{x}_i \leqslant x_i \leqslant \overline{x}_i$, it is NP-hard to evaluate the upper bound

$$\overline{F} = \max\{F(\xi_1, \ldots, \xi_n) \mid \underline{x}_i \leqslant \xi_i \leqslant \overline{x}_i,\ i = 1, \ldots, n\}.$$

The main result on t-norm forms

The main (negative) result on quadratic interactions with t-norms.

- **Theorem.** Let t_{ij} be arbitrary fixed Lipschitz continuous t-norms. For the t-norm form

$$F(x_1, \ldots, x_n) = \sum_{i=1}^{n} f_i(x_i) + \sum_{j \neq i} t_{ij}(x_i, x_j)$$

with interval data $\underline{x}_i \leqslant x_i \leqslant \overline{x}_i$, it is NP-hard to evaluate the upper bound

$$\overline{F} = \max\{F(\xi_1, \ldots, \xi_n) \mid \underline{x}_i \leqslant \xi_i \leqslant \overline{x}_i,\ i = 1, \ldots, n\}.$$

- **Remark.** To avoid confusion: This is indeed a nontrivial result. The fact that the special choice $t_{ij}(x_i, x_j) = q_{ij} x_i x_j$ is NP-hard does not imply NP-hardness for other choices of t_{ij}.
  - *Say, for example, that the Lipschitz continuous t-norm is chosen by an opponent. However (s)he makes the choice, we must always be able to construct a reduction from an NP-hard problem to the case formulated by the opponent.*

And the last slide. . .

And the last slide. . .

Thanks for your attention.

Michal, Milan and Vladik