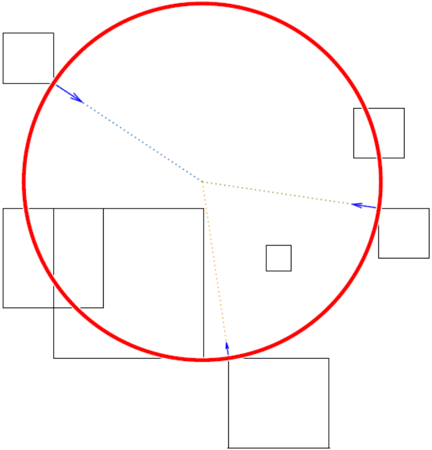
Solving nonlinear minmax location problems with minimal time functions by means of proximal point methods via conjugate duality

Sorin-Mihai Grad

joint work with Oleg Wilfer (Chemnitz UT)

Games, Dynamics & Optimization – GDO2019, Cluj-Napoca

Find the disc with minimal radius that intersects some given sets.



Outline

- Preliminaries
- Minmax location problems
- Duality investigations
- Numerical experiments

- X - Hilbert space
- $f : X \to \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$
- $\emptyset \neq U \subseteq X$
- $\emptyset \neq \Omega \subseteq X$
- *indicator function*: $\delta_U : X \to \overline{\mathbb{R}}$, $\delta_U(x) = 0$ if $x \in U$ and $\delta_U(x) = +\infty$ otherwise
- *support function*: $\sigma_U : X \to \overline{\mathbb{R}}$, $\sigma_U(y) = \sup_{x \in U} y^\top x$
- *gauge*: $\gamma_U : X \to \overline{\mathbb{R}}$, $\gamma_U(x) = \inf\{t > 0 : x \in tU\}$
- *generalized minimal-time function*: $\mathcal{T}^U_{\Omega,f} : X \to \mathbb{R}$, $\mathcal{T}^U_{\Omega,f}(x) := \inf \{\gamma_U(x - y - z) + f(y) : y \in X, z \in \Omega\}$

Theorem. (parallel splitting) [Bauschke & Combettes, 2009] *Let $f_i : \mathbb{R}^d \to \overline{\mathbb{R}}$ be proper convex lsc, $i = 1, \dots, n$. When the problem*

$$(P^{DR}) \quad \min_{x \in X} \left\{ \sum_{i=1}^{n} f_i(x) \right\}$$

has at least one solution, $\operatorname{dom} f_1 \cap \cap_{i=2}^{n} \operatorname{int} \operatorname{dom} f_i \neq \emptyset$, $(\mu_k)_{k \in \mathbb{N}}$ *is a sequence in* $[0, 2]$ *s.t.* $\sum_{k \in \mathbb{N}} \mu_k(2 - \mu_k) = +\infty$, *let* $\nu > 0$, *and* $(x_{i,0})_{i=1}^{n} \in X \times \ldots \times X$, *setting*

$$(\forall k \in \mathbb{N}) \quad \left| \begin{array}{l} r_k = \frac{1}{n} \sum_{i=1}^{n} x_{i,k}, \\ y_{i,k} = \operatorname{prox}_{\nu f_i} x_{i,k}, \ i = 1, \dots, n, \\ q_k = \frac{1}{n} \sum_{i=1}^{n} y_{i,k}, \\ x_{i,k+1} = x_{i,k} + \mu_k(2q_k - r_k - y_{i,k}), \ i = 1, \dots, n, \end{array} \right.$$

then $(r_k)_{k \in \mathbb{N}}$ *converges weakly to an optimal solution to* (P^{DR}).

Consider the following generalized location problem

$$(P^S_{h,\mathcal{T}}) \qquad \inf_{x\in S} \max_{1\le i\le n} \left\{ h_i\left(\mathcal{T}^{C_i}_{\Omega_i,f_i}(x)\right) + a_i \right\},$$

where

- $\emptyset \neq S \subseteq X$ is closed and convex, $n \geq 2$

for $i = 1, \ldots, n$ $(n \geq 2)$:

- $a_i \in \mathbb{R}_+ = [0, +\infty)$ (set-up costs)
- $C_i \subseteq X$ is closed and convex with $0 \in \operatorname{int} C_i$
- $\emptyset \neq \Omega_i \subseteq X$ is convex and compact
- $f_i : X \to \overline{\mathbb{R}}$ is proper convex lsc
- $h_i : \mathbb{R} \to \overline{\mathbb{R}}$ with $h_i(x) \begin{cases} \in \mathbb{R}_+, & \text{if } x \in \mathbb{R}_+ \\ = +\infty, & \text{otherwise} \end{cases}$ is convex lsc and increasing on $\mathbb{R}_+$

- take $f_i = \delta_{L_i}$, where $\emptyset \neq L_i \subseteq X$, is closed and convex
- $h_i(x) = x + \delta_{\mathbb{R}_+}(x)$, $x \in \mathbb{R}$, $i = 1, \ldots, n$
- $(P^S_{h,\mathcal{T}})$ can be equivalently written as

$$(P^S_{G,\mathcal{T}}) \qquad \inf_{\substack{x\in S,\ t\in\mathbb{R},\\ \inf\{\lambda_i>0:(x-\lambda_i C_i)\cap(\Omega_i+L_i)\neq\emptyset\}+a_i\leq t, i=1,\ldots,n}} t$$

Geometric interpretation

$(P^S_{G,\mathcal{T}})$: determine a point $\bar{x} \in S$ and the smallest $\bar{t} > 0$ s.t.

$$(\bar{x} - (\bar{t} - a_i)C_i) \cap (\Omega_i + L_i) \neq \emptyset, \ i = 1, \ldots, n$$

- approach useful when the target sets are hard to handle, but can be split into Minkowski sums of two simpler sets Ω_i and L_i, $i = 1, \ldots, n$, (e.g.: rounded rectangles = sums of rectangles and circles)
- $(P^S_{G,\mathcal{T}})$ is a generalization of the *Sylvester problem* (find the smallest circle that encloses finitely many given points)

- take $f_i = \gamma_{G_i}$, where $\emptyset \neq G_i \subseteq X$, is closed and convex
- take $h_i(x) = x + \delta_{\mathbb{R}_+}(x)$, $x \in \mathbb{R}$, $a_i = 0$, $i = 1, \ldots, n$
- $(P^S_{h,\mathcal{T}})$ can be equivalently written as

$$(P^S_{E,\mathcal{T}}) \quad \inf_{\substack{x\in X, z_i\in\Omega_i,\ \alpha_i,\ \beta_i,\ t>0,\ \alpha_i+\beta_i\leq t, \\ (x+\alpha C_i)\cap(z_i+\beta_i G_i)\neq\emptyset,\ i=1,\ldots,n}} t$$

Economic interpretation

- cities: $1, \ldots, n$

for $i = 1, \ldots, n$

- G_i: demand of i for product P (produced by W)
- Ω_i: characterization of the budget of i
- C_i: characterization of the importance of i for W

$(P^S_{E,\mathcal{T}})$: determine a location $\overline{x} \in S$ for a production facility s.t. the total demand for P can be satisfied in the shortest time $\bar{t} > 0$

The dual problem to $(P^S_{h,\mathcal{T}})$ we obtain (by means of a "multicomposed" approach (cf. [**G**, Wanka & Wilfer, 2017]) is

$$(D^S_{h,\mathcal{T}}) \quad \sup_{\substack{\lambda_i,\ z_i^*\geq 0,\ w_i^*\in X, \sum_{i=1}^n \lambda_i \leq 1, \\ \gamma_{C_i^0}(w_i^*)\leq z_i^*, i=1,\dots,n}} \left\{ -\sigma_S\left(-\sum_{i=1}^n w_i^*\right) - \sum_{i=1}^n \left[(\lambda_i h_i)^*(z_i^*) - \lambda_i a_i + (z_i^* f_i)^*(w_i^*) + \sigma_{\Omega_i}(w_i^*)\right] \right\}$$

Theorem. (strong duality) *Between* $(P^S_{h,\mathcal{T}})$ *and* $(D^S_{h,\mathcal{T}})$ *holds strong duality, i.e. their optimal objective values coincide* $(v(P^S_{h,\mathcal{T}}) = v(D^S_{h,\mathcal{T}}))$ *and the dual problem has an optimal solution* $(\bar{\lambda}, \bar{z}^*, \bar{w}) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+ \times X^n$.

- in general one has $v(P^S_{h,\mathcal{T}}) \geq v(D^S_{h,\mathcal{T}})$
- the constraint qualification usually needed for strong duality is fulfilled due of the hypotheses
- one can also derive corresponding necessary and sufficient optimality conditions $\Rightarrow$ in some special cases the optimal solution $\bar{x}$ to $(P^S_{h,\mathcal{T}})$ can be characterized as

$$\overline{x} = \frac{1}{\sum\limits_{i \in \overline{I}} \beta_i \|\overline{w}_i^*\|} \sum_{i \in \overline{I}} \beta_i \|\overline{w}_i^*\| p_i$$

Rewrite the location problem

$$(P_G, \mathcal{T}) \qquad \inf_{x \in X} \max_{1 \leq i \leq n} \left\{ \mathcal{T}^{C_i}_{\Omega_i, \delta_{L_i}}(x) \right\},$$

where $C_i,\ L_i \subseteq X$ are closed and convex sets with $0 \in \operatorname{int} C_i$ and $\Omega_i \subseteq X$ are convex and compact sets, $i = 1, \ldots, n$, as follows

$$(P_G, \mathcal{T})$$

$$\inf_{\substack{t \geq 0,\ x, y_i, z_i \in X, \\ i=1,\ldots,n}} \left\{ t + \sum_{i=1}^{n} \left[\delta_{\operatorname{epi} \gamma_{C_i}}(x - y_i - z_i, t) + \delta_{\Omega_i}(y_i) + \delta_{L_i}(z_i) \right] \right\}$$

The dual problem to $(P_{G,\,\mathcal{T}})$ can be rewritten as

$$(D_{G,\,\mathcal{T}})$$
$$-\inf_{w_i^* \in X,\ i=1,\dots,n} \left\{ \sum_{i=1}^n \left[\sigma_{L_i}(w^*) + \sigma_{\Omega_i}(w^*) \right] + \delta_F(w^*) + \delta_E(w^*) \right\},$$

where $E = \left\{ w^* = (w_1^*, \dots, w_n^*) \in X \times \ldots \times X : \sum_{i=1}^n w_i^* = 0 \right\}$ and
$F = \left\{ w^* = (w_1^*, \dots, w_n^*) \in X \times \ldots \times X : \sum_{i=1}^n \gamma_{C_i^0}(w_i^*) \le 1 \right\}$

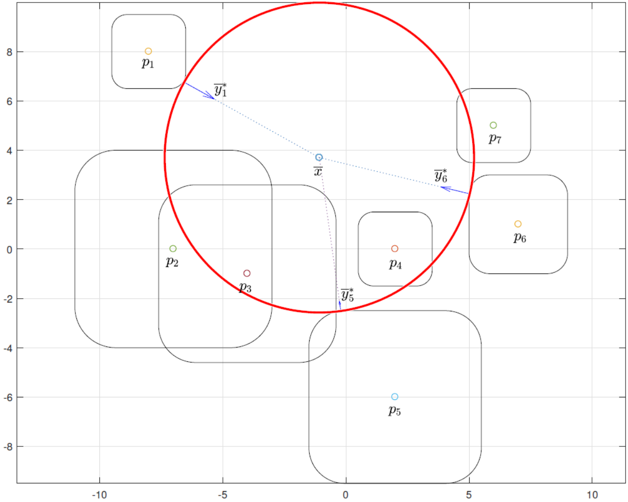
Theorem. (epi-projection) *Let* $\gamma_C : X^n \to \mathbb{R}$ *be defined by* $\gamma_C(x_1, \dots, x_n) := \max_{1 \le i \le n}\{\|x_i\|/w_i\}$. *It holds*

$$\mathrm{Pr}_{\mathrm{epi}\,\gamma_C}(x_1, \dots, x_n, \xi) = \begin{cases} (x_1, \dots, x_n), & \text{if } \max\limits_{1 \le i \le n} \left\{ \frac{1}{w_i}\|x_i\| \right\} \le \xi, \\ (0, \dots, 0, 0), & \text{if } \xi < 0 \text{ and } \sum\limits_{i=1}^{n} w_i\|x_i\| \le -\xi, \\ (\overline{y}_1, \dots, \overline{y}_n, \overline{\theta}), & \text{otherwise,} \end{cases}$$

where for $i = 1, \dots, n$ *one has*

$$\overline{y}_i = x_i - \frac{\max\{\|x_i\| - (\overline{\kappa} + \xi)w_i, 0\}}{\|x_i\|} x_i, \text{ and } \overline{\theta} = \frac{\sum_{i=k+1}^{n} w_i^2 \tau_i + \xi}{\sum_{i=k+1}^{n} w_i^2 + 1}$$

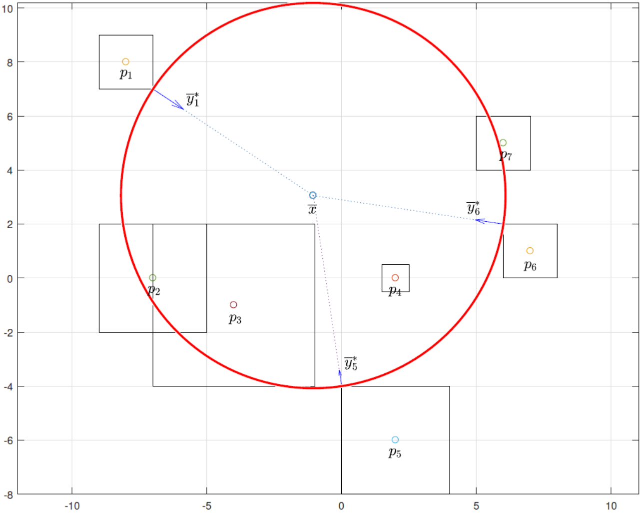
with $\overline{\kappa} = (\sum_{i=k+1}^{n} w_i^2 \tau_i - \xi \sum_{i=k+1}^{n} w_i^2)/(\sum_{i=k+1}^{n} w_i^2 + 1)$ *and* $k \in \{0, 1, \dots, n-1\}$ *is the unique integer such that* $\tau_k + \xi \le \overline{\kappa} \le \tau_{k+1} + \xi$, *where the values* $\tau_0, \dots, \tau_n$ *are defined by* $\tau_0 := 0$ *and* $\tau_i := \|x_i\|/w_i$, $i = 1, \dots, n$, *and in ascending order.*


p_1
p_2
p_3
p_4
p_5
p_6
p_7
$\overline{y}_1^*$
$\overline{y}_5^*$
$\overline{y}_6^*$
$\overline{x}$
8
6
4
2
0
-2
-4
-6
-8
-10
-5
0
5
10

Let $d = 2$, $p_1 = (-8, 8)^T$, $p_2 = (-7, 0)^T$, $p_3 = (-4, -1)^T$, $p_4 = (2, 0)^T$, $p_5 = (2, -6)^T$, $p_6 = (7, 1)^T$, $p_7 = (6, 5)^T$, $c_1 = 1,\ c_2 = 2,\ c_3 = 3,\ c_4 = 0.5,\ c_5 = 2,\ c_6 = 1,\ c_7 = 1,\ b_1 = 0.5,\ b_2 = 2,\ b_3 = 0.6,\ b_4 = 1,\ b_5 = 1.5,\ b_6 = 1,\ b_7 = 0.5$, $\Omega_i = \{x \in \mathbb{R}^2 : \|x - p_i\|_\infty \leq c_i\}$, $L_i = \{x \in \mathbb{R}^2 : \|x\| \leq b_i\}$ and $\gamma_{C_i} = \|\cdot\|$, $i = 1, \ldots, 7$.

	$\varepsilon = 10^{-4}$		$\varepsilon = 10^{-8}$	
	primal	dual	primal	dual
CPU	0.3786	0.1174	0.7640	0.2973
NI	541	330	1106	830

Performance evaluation for 7 sets in $\mathbb{R}^2$
(CPU = CPU Time in seconds; NI = Number of iterations, ε =distance from the optimal value of the problem)


10
8
6
4
2
0
-2
-4
-6
-8
-10
-5
0
5
10
p_1
p_2
p_3
p_4
p_5
p_6
p_7
$\overline{x}$
$\overline{y}_1^*$
$\overline{y}_5^*$
$\overline{y}_6^*$

Let $d = 2,\ p_1 = (-8,8)^T,\ p_2 = (-7,0)^T,\ p_3 = (-4,-1)^T,\ p_4 = (2,0)^T,\ p_5 = (2,-6)^T,\ p_6 = (7,1)^T,\ p_7 = (6,5)^T,\ c_1 = 1,\ c_2 = 2,\ c_3 = 3,\ c_4 = 0.5,\ c_5 = 2,\ c_6 = 1,\ c_7 = 1,\ \Omega_i = \{x \in \mathbb{R}^2 : \|x - p_i\|_\infty \leq c_i\},\ L_i = \{0_{\mathbb{R}^2}\},\ \gamma_{C_i} = \|\cdot\|,\ i = 1,\ldots,7.$
We compare our methods with the subgradient methods of [Mordukhovich & Nam, 2014] and [Nam, An & Salinas, 2015]

	primal	dual	subgrad.(1)	subgrad.(2)
CPU	0.1904	0.0871	0.0416	1.2782
NI	399	181	918	70752

Performance evaluation for 7 sets in $\mathbb{R}^2$ with $\varepsilon = 10^{-4}$

	primal	dual	subgrad.(1)	subgrad. (2)
CPU	0.3377	0.1608	0.7016	-
NI	730	453	37854	500000+

Performance evaluation for 7 sets in $\mathbb{R}^2$ with $\varepsilon = 10^{-8}$

We compare our methods with the subgradient methods of [Mordukhovich & Nam, 2014] and [Nam, An & Salinas, 2015]

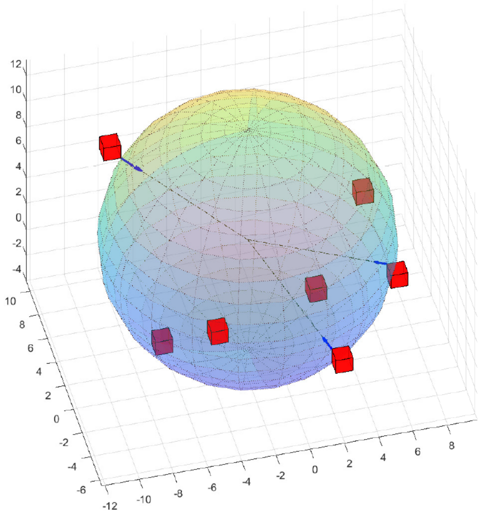
	primal	dual	subgrad.(1)	subgrad.(2)
CPU Time in sec.	5.6477	0.4292	-	27.1555
Number of It.	2421	735	500000+	383782

Performance evaluation for 50 sets in $\mathbb{R}^2$ with $\varepsilon = 10^{-4}$

	primal	dual	subgrad.(1)	subgrad.(2)
CPU Time in sec.	16.1011	3.6020	-	32.2530
Number of It.	6983	7207	500000+	436138

Performance evaluation for 50 sets in $\mathbb{R}^2$ with $\varepsilon = 10^{-8}$

(CPU = CPU Time in seconds; NI = Number of iterations)


12
10
8
6
4
2
0
-2
-4
10
8
6
4
2
0
-2
-4
-6
-12
-10
-8
-6
-4
-2
0
2
4
6
8

Let $d = 3$, $p_1 = (-8, 8, 8)^T$, $p_2 = (-7, 0, 0)^T$, $p_3 = (-4, -1, 1)^T$, $p_4 = (2, 0, 2)^T$, $p_5 = (2, -6, 2)^T$, $p_6 = (7, 1, 1)^T$, $p_7 = (6, 5, 4)^T$, $c_1 = 0.5$, $\Omega_i = \{x \in \mathbb{R}^3 : \|x - p_i\|_\infty \leq c_i\}$, $L_i = \{0_{\mathbb{R}^3}\}$, $\gamma_{C_i} = \|\cdot\|$, $i = 1, \ldots, 7$.

We compare our methods with the accelerated log-exponential smoothing technique of [An, Giles, Nam & Rector, 2017]

	$\varepsilon = 10^{-4}$			$\varepsilon = 10^{-8}$		
	primal	dual	log-exp	primal	dual	log-exp
CPU	0.1871	0.0992	6.9425	0.4234	0.2042	23.6893
NI	357	192	2340	955	523	9983

Performance evaluation for 7 sets in $\mathbb{R}^3$

(CPU = CPU Time in seconds; NI = Number of iterations)

Let $\Omega_i = \{x \in \mathbb{R}^d : \|x - p_i\|_\infty \leq c_i\}$, $L_i = \{0_{\mathbb{R}^d}\}$ and $\gamma_{C_i} = \|\cdot\|$, $i = 1, \ldots, n$, $\varepsilon = 10^{-6}$ (p_i and c_i random, $i = 1, \ldots, n$)

We compare our dual method with the accelerated log-exponential smoothing technique of [An, Giles, Nam & Rector, 2017]

	dual	log-exp
CPU	0.2889	55.4856
NI	1167	32265

Performance evaluation for 10 sets in $\mathbb{R}^{10}$, $\varepsilon = 10^{-6}$
(CPU = CPU Time in seconds; NI = Number of iterations)

	dual	log-exp
CPU	7.6268	70.3653
NI	1956	44173

Performance evaluation for 50 sets in $\mathbb{R}^{50}$, $\varepsilon = 10^{-6}$
(CPU = CPU Time in seconds; NI = Number of iterations)

We compare our dual method with the accelerated log-exponential smoothing technique of [An, Giles, Nam & Rector, 2017]

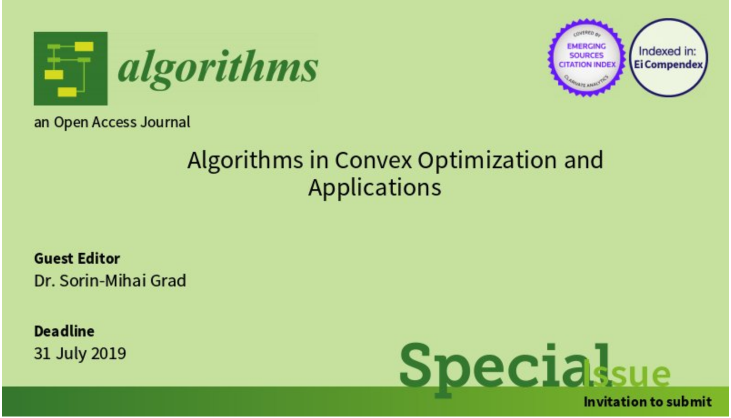
	dual	log-exp
CPU	104.5634	145.2422
NI	3003	69163

Performance evaluation for 100 sets in $\mathbb{R}^{100}$, $\varepsilon = 10^{-6}$
(CPU = CPU Time in seconds; NI = Number of iterations)

	dual	log-exp
CPU	5328.3671	7026.1593
NI	4017	691412

Performance evaluation for 100 sets in $\mathbb{R}^{1000}$, $\varepsilon = 10^{-6}$
(CPU = CPU Time in seconds; NI = Number of iterations)

www.mdpi.com/journal/algorithms/special_issues



- (Stochastic) Proximal point methods & Splitting techniques
- Convex & constrained optimization problems // Game theory
- Image processing // Machine learning // Location theory