MODERN METHODS FOR NONCONVEX OPTIMIZATION PROBLEMS

Alexander S. Strekalovsky

Russian Academy of Sciences,
Siberian Branch,
Institute for System Dynamics and Control Theory

e-mail: strekal@icc.ru

Workshop JBHU2010

Bayonne, France
October 25–27, 2010

Definition

Let set $X \subset I\!R^n$ be convex. Then $f\colon X \to I\!R$ is a d.c. function on X, when

$$\exists\, g, h \in \mathrm{Conv}\,(X)\colon \qquad f(x) = g(x) - h(x) \quad x \in X.$$

$DC(X)$ is a linear space of d.c. functions on X.
$\mathrm{Conv}\,(X)$ is a convex cone of convex functions on X.

1) $DC(X) = \mathrm{lin}\,(\mathrm{Conv}\,(X))$.

$$\begin{cases} \mathcal{K}(X) \text{ is a convex cone of functions} \\ \mathrm{Conv}\,(X) \subset \mathcal{K}(X) \subset DC(X), \\ \mathcal{K}(X) - \mathcal{K}(X) = DC(X). \end{cases}$$

2) $C^2(\Omega) \subset DC(\Omega)$, where Ω is a open convex set.

3) $\mathrm{cl}\,(DC(X)) = C(X)$, if X is a convex compact.

4) $DC(X)$ is a closed w.r.t. the following operations:

$$\sum_1^m \lambda_i f_i(x); \quad \max_i f_i(x); \quad \min_i f_i(x); \quad |f(x)|; \quad \prod_1^n f_i(x);$$

$$f^+(x) = \max\{0, f(x)\}; \quad f^-(x) = \min\{0, f(x)\}.$$

$$\left.\begin{array}{r} f_0(x) = g_0(x) - h_0(x) \downarrow \min, \quad x \in S, \\ f_i(x) = g_i(x) - h_i(x) \leqslant 0, \quad i \in I = \{1, ..., m\}. \end{array}\right\} \quad (\mathcal{P})$$

$$g_i, h_i \in CONVEX(\mathbb{R}^n), \;\; i \in \{0, 1, ..., m\},$$

$S \subset \mathbb{R}^n$ is a closed convex set.

$$f, g, h \in CONVEX(I\!R^n)$$

a) D.C. minimization:

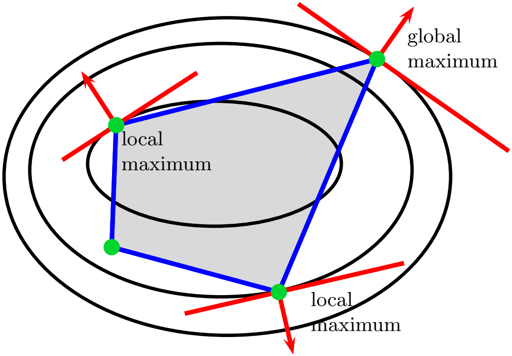
$$g(x) - h(x) \downarrow \min, \quad x \in D. \qquad (DC)$$

b) D.C. constrained problem:

$$\left.\begin{array}{c} f(x) \downarrow \min, \quad x \in S, \\ g(x) - h(x) \leqslant 0. \end{array}\right\} \qquad (DCC)$$

Note:
i) If $g \equiv 0$ in (DC), then (DC) is convex maximization problem;
ii) If $g \equiv 0$ in (DCC), then (DCC) is reverse-convex problem.

1. Linearization of basic nonconvexities of the problem under scrutiny and, consequently, reduction of the problem to a family of (partially) linearized problems.
2. Application of convex optimization methods for solving linearized problems and, as a consequence, “within” special local search methods.
3. Construction of “good” approximations (resolving sets) of level surfaces/epigraph boundaries of convex functions.



Methodology for Solving Nonconvex Optimization Problems

1. Never apply convex optimization methods DIRECTLY.
2. Exact classification of the problem under scrutiny.
3. Application of special (for the class of problem under scrutiny) local search (LS) methods, or (problem) specific methods.
4. Application of global search strategies specialized for the class of nonconvex problems.
5. Construction of pertinent approximations of level surfaces with the aid of the experience obtained during solving similar nonconvex problems.
6. Application of convex optimization methods for solving linearized problems and within the framework of special LS methods.

1. Special local search methods. Convergence theorems.
2. Global Optimality Conditions (GOC).
3. Global search algorithms.
4. Convergence theorems of global search algorithms.
5. Testing.

A.S. Strekalovsky, *Elements of Nonconvex Optimization* (Nauka, Novosibirsk, 2003) [in Russian].

Problem Formulation

$$\Phi(x) = \Phi(u,v,w) \downarrow \min, \quad x = (u,v,w) \in D, \quad x \in I\!R^n, \tag{1}$$

where $\Phi(\cdot)$ is partly convex w.r.t. variables $u \in I\!R^m$, $v \in I\!R^p$ and $w \in I\!R^q$ separately.

Step 0. Initialization. $(u^0, v^0, w^0) \in D, k := 0.$

Step 1. Solve the problem:

$$\Phi(u, v^k, w^k) \downarrow \min_u, \; u \in U_k \triangleq \{u \in I\!R^m \mid (u, v^k, w^k) \in D\} \neq \varnothing. \qquad (\mathcal{P}_u^k)$$

Let $\bar{u}^k \in Sol(\mathcal{P}_u^k)$.

Step 2. Find a global solution $\bar{v}^k \in Sol(\mathcal{P}_v^k)$:

$$\Phi(\bar{u}^k, v, w^k) \downarrow \min_v, \; v \in V_k \triangleq \{v \in I\!R^p \mid (\bar{u}^k, v, w^k) \in D\} \neq \varnothing. \qquad (\mathcal{P}_v^k)$$

Step 3. Solve the following problem ($\bar{w}^k \in Sol(\mathcal{P}_w^k)$):

$$\Phi(\bar{u}^k, \bar{v}^k, w) \downarrow \min_w, \; w \in W_k \triangleq \{w \in I\!R^q \mid (\bar{u}^k, \bar{v}^k, w) \in D\} \neq \varnothing. \qquad (\mathcal{P}_w^k)$$

Step 4. $u^{k+1} := \bar{u}^k$, $v^{k+1} := \bar{v}^k$, $w^{k+1} := \bar{w}^k$, $k := k+1$. Go to Step 1.

Definition

The point $(\bar{u}, \bar{v}, \bar{w})$ is a critical point in the problem (1), if the following conditions are fulfilled:

$$\Phi(\bar{u}, \bar{v}, \bar{w}) \leqslant \Phi(u, \bar{v}, \bar{w}) \quad \forall u \in U(\bar{v}, \bar{w}),$$
$$\Phi(\bar{u}, \bar{v}, \bar{w}) \leqslant \Phi(\bar{u}, v, \bar{w}) \quad \forall v \in V(\bar{u}, \bar{w}),$$
$$\Phi(\bar{u}, \bar{v}, \bar{w}) \leqslant \Phi(\bar{u}, \bar{v}, w) \quad \forall w \in W(\bar{u}, \bar{v}).$$

Example

$$D = \{(u, v, w) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^q \mid Au + Bv + Cw \leqslant d\}.$$

Problem Formulation

$$g(x) - h(x) \downarrow \min, \quad x \in D. \qquad (\mathcal{P})$$

The basic element of the local search method is solving the (linearized at a current feasible point $x^s \in D$) convex problem

$$J_s(x) = g(x) - \langle h'_s, x \rangle \downarrow \min, \quad x \in D, \qquad (\mathcal{PL}_s)$$

where $h'_s = h'(x^s) \in \partial h(x^s)$.

Next point $x^{s+1} \in D$ is constructed as an approximate solution of the problem $(\mathcal{P}L_s)$.

THEOREM 1

Let $F = g - h$ be a bounded below function on D and function $h(\cdot)$ be convex on D.
Then the sequence $\{x^s\}$, generated by the rule

$$g(x^{s+1}) - \langle h'(x^s), x^{s+1}\rangle \leqslant \inf_x \{g(x) - \langle h'(x^s), x\rangle \mid x \in D\} + \delta_s, \qquad (2)$$

satisfies the following condition:

$$\lim_{s\to\infty} [\mathcal{V}(PL_s) - J_s(x^{s+1})] = 0. \qquad (3)$$

At the same time any accumulation point x_* of sequence $\{x^s\}$ is a solution of the problem

$$J_*(x) = g(x) - \langle h'(x_*), x\rangle \downarrow \min, \quad x \in D, \qquad (\mathcal{PL}(x_*))$$

where $h'(x_*) \in \partial h(x_*)$.
If $h(\cdot)$ is strongly convex, then $x^s \to x_* \in D$. Besides,

$$\|x^s - x^{s+1}\|^2 \leqslant \frac{2}{\mu}(F(x^s) - F(x^{s+1}) + \delta_s),$$

where $\mu > 0$ is a strong convexity constant of function $h(\cdot)$.

Problem Formulation

$$f_0(x) \downarrow \min, \quad x \in S,$$
$$g(x) - h(x) \leqslant 0.$$

Procedure 1 constructes $x(y) \in S$ from a predetermined point $y \in S,\ F(y) = g(y) - h(y) \leqslant 0$:

$$F(x(y)) = 0, \;\; f_0(x(y)) \leqslant f_0(y).$$

Procedure 2 consists in sequential solution of the linearized problems

$$g(x) - \langle h'(u), x\rangle \downarrow \min, \quad x \in S, \quad f_0(x) \leqslant \zeta, \qquad (\mathcal{LQ}(u,\xi))$$

where $h'(u) \in \partial h(u)$.

$$(H_0): \qquad \exists v \in S, \quad g(v) - h(v) > 0: \;\; f_0(v) < f_0^* \triangleq \mathcal{V}(P). \qquad (4)$$

Problem Formulation

$$\left.\begin{array}{r} f_0(x) = g_0(x) - h_0(x) \downarrow \min, \quad x \in S, \\ f_i(x) = g_i(x) - h_i(x) \leqslant 0, \quad i \in I = \{1, ..., m\}. \end{array}\right\} \qquad (\mathcal{P})$$

Linearized Problem 1

$$\left.\begin{array}{r} g_0(x) - \langle h_0'(x^s), x\rangle \downarrow \min, \quad x \in S, \\ g_i(x) - \langle h_i'(x^s), x - x^s\rangle - h_i(x^s) \leqslant 0, \quad i \in I, \end{array}\right\} \qquad (\mathcal{PL}_s^1)$$

where $h_j'(x^s) \in \partial h_j(x^s), \quad j = 0, 1, 2, \ldots, m.$

Linearized Problem 2

$$\left.\begin{array}{r} g_j(x) - \langle h_j'(x^s), x\rangle \downarrow \min, \quad x \in S, \\ g_i(x) - \langle h_i'(x^s), x - x^s\rangle - h_i(x^s) \leqslant 0, \quad i \neq j, \\ g_0(x) - \langle h_0'(x^s), x - x^s\rangle - h_0(x^s) \leqslant \zeta, \end{array}\right\} \qquad (\mathcal{PL}_s^2)$$

where $h_j'(x^s) \in \partial h_j(x^s), \quad j = 0, 1, 2, \ldots, m.$

General procedure of global search consists of the two parts:
a) local search;
b) procedure of escaping from a critical point, which is based on the global optimality conditions (GOC), with the following inclusion of the local search.

$$g(x) - h(x) \downarrow \min, \quad x \in D. \qquad (\mathcal{P})$$

THEOREM 2. (GOC)

Let a feasible point $z \in D$ be a global solution to Problem $(\mathcal{P})$ $(z \in Sol(\mathcal{P}))$. Then

$$(\mathcal{E}): \quad \begin{array}{c} \forall (y, \beta): \quad y \in D, \quad \beta - h(y) = \zeta \triangleq g(z) - h(z), \\ g(y) \leqslant \beta \leqslant \sup(g, D), \\ g(x) - \beta \geqslant \langle h'(y), x - y \rangle, \ \forall x \in D, \end{array} \qquad (5)$$

where $h'(y) \in \partial h(y)$.

If, in addition, the following regularity condition holds

$$(\mathcal{H}): \qquad \exists v \in D: \quad g(v) - h(v) > \zeta, \qquad (6)$$

then conditions $(\mathcal{E})$ turns out to be sufficient for the point z being a global solution to Problem $(\mathcal{P})$.

Suppose, the 3-tuple $(\hat{y}, \hat{\beta}, \hat{x})$, such that $(\hat{y}, \hat{\beta})$: $h(\hat{y}) = \hat{\beta} - \zeta$, $\zeta := f(z)$ and $\hat{x} \in D$, violates the GOC $(\mathcal{E})$, i.e.

$$g(\hat{x}) < \hat{\beta} + \langle h'(\hat{y}), \hat{x} - \hat{y} \rangle,$$

Then from convexity of $h(\cdot)$ it follows that

$$f(\hat{x}) = g(\hat{x}) - h(\hat{x}) < h(\hat{y}) + \zeta - h(\hat{y}) = f(z)$$

or $f(\hat{x}) < f(z)$. Therefore, $\hat{x} \in D$ is «better» than z.
And so, overhauling «perturbation parameters» (y, β) in $(\mathcal{E})$ and solving linearized problems (see GOC)

$$g(x) - \langle h'(y), x \rangle \downarrow \min, \; x \in D, \tag{7}$$

we obtain a family of initial points $x(y, \beta)$ for the local search methods.

$$g(x) - h(x) \downarrow \min, \quad x \in D. \qquad (\mathcal{P})$$

1) Find a critical point z with special local search algorithm, $\zeta \triangleq g(z) - h(z)$.

2) Choose number $\beta \in [\beta_-, \beta_+]$, $\beta_- = \inf(g, D)$, $\beta_+ = \sup(g, D)$.

3) Construct an approximation

$$A(\beta) = \{v^1, ..., v^N \mid \quad h(v^i) = \beta - \zeta, \quad i = 1, ..., N, \quad N = N(\beta)\},$$

of level surface of function $h(\cdot)$.

4) Beginning from each point v^i of the approximation $A(\beta)$ find a point u^i by means of a special local search algorithm.

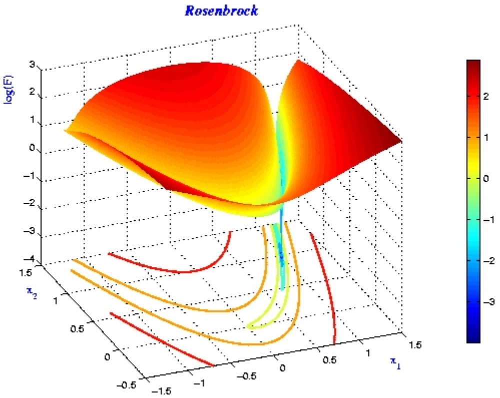
5) Verify the inequality from GOC:

$$g(u^i) - \beta \geqslant \langle h'(v^i), u^i - v^i \rangle \quad \forall i = 1, 2, \dots, N.$$

$$(w^i \mapsto v^i)$$

Minimization of Rosenbrock's function

$$(\mathcal{P}): \quad F(x) = \sum_{i=1}^{n-1} \left[(x_i - 1)^2 + N(x_{i+1} - x_i^2)^2 \right] \downarrow \min, \quad x \in \mathbb{R}^n.$$



$$F(x) = G(x) - H(x), \tag{8}$$

$$G(x) = \sum_{i=1}^{n-1} [(x_i - 1)^2 + \frac{N}{2}(1 + x_{i+1})^4 + \frac{3N}{2} x_i^4 + N x_i^2 x_{i+1}^2 + N x_i^2 + N x_{i+1}^2],$$

$$H(x) = \frac{N}{2} \sum_{i=1}^{n-1} ((1 + x_{i+1})^2 + x_i^2)^2.$$

$$G(x) = \sum_{i=1}^{n-1} g(x_i, x_{i+1}), \quad H(x) = \sum_{i=1}^{n-1} h(x_i, x_{i+1}). \tag{9}$$

The functions $G(\cdot)$ and $H(\cdot)$ are convex.

n	Special Local Search Method				BFGS Method			
	$\|z - x_*\|$	$F(z)$	It	$Time$	$\|z - x_*\|$	$F(z)$	It	$Time$
5	1.12	0.64	46	1.25	2.01	3.93	75	0.04
5	1.12	0.64	44	1.25	2.01	3.93	65	0.04
5	1.12	0.64	44	1.35	2.01	3.93	90	0.06
5	$1.3 \cdot 10^{-3}$	$5.7 \cdot 10^{-5}$	56	1.30	$3.1 \cdot 10^{-5}$	$5.2 \cdot 10^{-9}$	61	0.02
10	1.22	0.81	45	1.30	1.99	3.98	67	0.55
10	1.22	0.81	44	1.30	1.99	3.98	75	0.55
10	$4.1 \cdot 10^{-4}$	$8.5 \cdot 10^{-6}$	49	1.40	1.99	3.98	71	0.50
10	0.024	0.0045	59	1.60	$1.4 \cdot 10^{-5}$	$3.33 \cdot 10^{-9}$	129	0.81
50	1.31	0.76	46	7.20	1.87	4.20	85	2.85
50	1.31	0.76	48	7.35	1.87	4.20	93	2.90
50	0.0035	$1.2 \cdot 10^{-5}$	48	7.40	1.87	4.20	88	2.85
50	$1.8 \cdot 10^{-4}$	$3.1 \cdot 10^{-6}$	57	8.25	$1.3 \cdot 10^{-6}$	$2.6 \cdot 10^{-10}$	136	3.15
100	1.46	4.56	51	14.20	1.73	0.83	112	10.00
100	1.46	4.56	50	13.90	1.73	0.83	120	10.30
100	0.0012	$4.5 \cdot 10^{-5}$	53	15.25	1.73	0.83	115	10.00
100	$3.8 \cdot 10^{-4}$	$5.5 \cdot 10^{-7}$	61	17.10	$1.8 \cdot 10^{-7}$	$8.2 \cdot 10^{-11}$	145	12.00

$$F(x) = \sum_{i=1}^{n-1} \Big[(x_i - 1)^2 + N(x_{i+1} - x_i^2)^2 \Big] \downarrow \min, \quad x \in I\!R^n. \qquad (\mathcal{P})$$

1) Find a critical point z with the special local search algorithm, $\zeta \triangleq F(z) = G(z) - H(z)$.

2) Choose number $\beta \in [\beta_-, \beta_+]$, $\quad \beta_- = \inf\,(G, I\!R^n)$, $\quad \beta_+ = \sup\,(G, I\!R^n)$.

3) Construct an approximation

$$A(\beta) = \{v^1, ..., v^N \mid \quad H(v^i) = \beta - \zeta, \quad i = 1, ..., N, \quad N = N(\beta)\},$$

of level surface of function $H(\cdot)$.

4) Beginning from each point v^i of the approximation $A(\beta)$ find a point u^i by means of a special local search algorithm.

5) Verify the inequality from GOC $(w^i \mapsto v^i)$:

$$g(u^i) - \beta \geqslant \langle h'(v^i), u^i - v^i \rangle \quad \forall i = 1, 2, \ldots, N.$$

n	$F(x^0)$	$\|\nabla F(x^0)\|$	$\|z - x_*\|$	$F(z)$	PL	$Time$
5	$1.2 \cdot 10^6$	408815	$1.1 \cdot 10^{-5}$	$1.6 \cdot 10^{-9}$	41	4.21
5	458810	196853	$1.2 \cdot 10^{-5}$	$7.5 \cdot 10^{-10}$	36	4.11
5	$2.5 \cdot 10^6$	617951	$1.2 \cdot 10^{-5}$	$1.7 \cdot 10^{-9}$	38	4.15
10	38084	24647	$3.2 \cdot 10^{-5}$	$2.0 \cdot 10^{-9}$	85	9.60
10	7940	8543	$3.1 \cdot 10^{-5}$	$1.8 \cdot 10^{-9}$	82	9.55
10	11101	11388	$1.5 \cdot 10^{-5}$	$1.8 \cdot 10^{-9}$	76	9.43
50	$8.1 \cdot 10^7$	$2.9 \cdot 10^6$	$3.4 \cdot 10^{-5}$	$2.5 \cdot 10^{-9}$	410	45.25
50	686176	256999	$1.7 \cdot 10^{-5}$	$3.7 \cdot 10^{-9}$	386	45.01
50	102509	32805.5	$2.1 \cdot 10^{-5}$	$4.2 \cdot 10^{-9}$	431	46.60
100	181996	44044.5	$1.6 \cdot 10^{-5}$	$2.5 \cdot 10^{-9}$	753	1 : 14.00
100	$2.1 \cdot 10^7$	$4.8 \cdot 10^6$	$2.2 \cdot 10^{-5}$	$2.6 \cdot 10^{-9}$	781	1 : 15.00
100	152183	41207	$2.4 \cdot 10^{-6}$	$6.8 \cdot 10^{-10}$	812	1 : 17.00
125	$4,4 \cdot 10^7$	$5,7 \cdot 10^6$	$3 \cdot 10^{-6}$	$1,5 \cdot 10^{-9}$	1044	1 : 50.10
125	$7,8 \cdot 10^7$	$4,1 \cdot 10^6$	$6,7 \cdot 10^{-6}$	$6,3 \cdot 10^{-9}$	1012	1 : 49.00
125	$3,6 \cdot 10^7$	$7,4 \cdot 10^6$	$4,6 \cdot 10^{-6}$	$2,1 \cdot 10^{-10}$	1101	1 : 51.00
151	$3.0 \cdot 10^8$	$7,2 \cdot 10^7$	$2,2 \cdot 10^{-5}$	$5,1 \cdot 10^{-9}$	1361	2 : 26.10
151	$5,8 \cdot 10^8$	$8,6 \cdot 10^7$	$3,8 \cdot 10^{-5}$	$3,1 \cdot 10^{-10}$	1355	2 : 25.20
151	$8,8 \cdot 10^8$	$8,1 \cdot 10^7$	$2,2 \cdot 10^{-5}$	$5,1 \cdot 10^{-9}$	1394	2 : 28.00

$$\|\nabla F(z)\| < 10^{-3}$$

For all dimensions the number of iteration of global search equals $It = 2$.

1. Hierarchical problems and variational inequalities

a) *Linear optimistic bilevel problems*

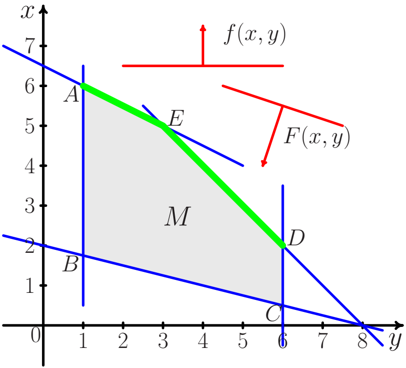
$$\left.\begin{array}{c} F(x,y) \triangleq \langle c,x\rangle + \langle d,y\rangle \downarrow \min\limits_{x}, \\ (x,y) \in X \triangleq \{x \in \mathbb{R}^m \mid Ax \leqslant b\}, \\ y \in Y_*(x) \triangleq \operatorname*{Argmin}\limits_{y}\{\langle d^1, y\rangle \mid y \in Y(x)\}, \\ Y(x) \triangleq \{y \in \mathbb{R}^n \mid A_1 x + B_1 y \leqslant b^1\} \end{array}\right\} \qquad (\mathcal{BP})$$

Problem with d.c. equation-constraint

$$\left.\begin{array}{c} \langle c,x\rangle + \langle d,y\rangle \downarrow \min\limits_{x,y,v}, \\ (x,y,v) \in S \triangleq \{Ax \leqslant b,\;\; A_1x + B_1 y \leqslant b^1,\;\; vB_1 = -d^1,\;\; v \geqslant 0\}, \\ \langle d^1, y\rangle = \langle A_1 x - b^1, v\rangle. \end{array}\right\}$$

$$\begin{cases} F(x,y) = 3x + y \downarrow \min\limits_{x,y}, \\ 1 \leqslant y \leqslant 6, \;\; x \in X_*(y) = Sol(\mathcal{P}_L), \end{cases}$$

$$(\mathcal{P}_L) \quad \begin{cases} f(x) = -x \downarrow \min\limits_{x}, \\ x + y \leqslant 8, \\ 4x + y \geqslant 8, \\ 2x + y \leqslant 13. \end{cases}$$



1. Hierarchical problems and variational inequalities

b) *Non-linear optimistic bilevel problems*

$$\left.\begin{array}{c} F(x,y) \triangleq \frac{1}{2}\langle x, Cx\rangle + \langle c, x\rangle + \frac{1}{2}\langle y, C_1 y\rangle + \langle c_1, y\rangle \downarrow \min\limits_{x,y}, \\ (x,y) \in X \triangleq \{(x,y) \in I\!R^m \times I\!R^n \mid Ax + By \leqslant a, \;\; x \geqslant 0\}, \\ y \in Y_*(x) \triangleq \operatorname*{Argmin}\limits_{y}\{\langle d, y\rangle \mid \; y \in Y(x)\}, \\ Y(x) \triangleq \{y \in I\!R^n \mid A_1 x + B_1 y \leqslant b, \;\; y \geqslant 0\} \end{array}\right\} \qquad (\mathcal{BP})$$

Problem with bilinear equation-constraint

$$\left.\begin{array}{c} F(x,y) \downarrow \min\limits_{x,y,v}, \quad Ax + By \leqslant b, \\ D_1 y + d_1 + xQ_1 + vB_1 = 0, \\ r(x,y,v) \triangleq \langle v, b_1 - A_1 x - B_1 y\rangle = 0 \\ v \geqslant 0, \qquad A_1 x + B_1 y \leqslant b_1. \end{array}\right\} \qquad (\mathcal{P})$$

Problem with d.c. objective function

$$\left.\begin{array}{c} \Phi(x,y,v) \triangleq F(x,y) + \mu r(x,y,v) \downarrow \min\limits_{x,y,v}, \\ (x,y,v) \in D \triangleq \{(x,y,v) \mid Ax + By \leqslant b, \;\; v \geqslant 0, \\ D_1 y + d_1 + xQ_1 + vB_1 = 0, \;\; A_1 x + B_1 y \leqslant b_1\}, \end{array}\right\} \qquad (\mathcal{P}(\mu))$$

where $\mu > 0$ is a penalty parameter, $r(x,y,v) \geqslant 0, \;\; \forall (x,y,v) \in D$.

A.S. Strekalovsky, A.V. Orlov
Bimatrix games and bilinear programming.
FizMatLit, Moscow, 2007 [in Russian].

Strekalovsky A.S., Orlov A.V., Malyshev A.V
On numerical search for optimistic solutions in bilevel problems.
Journal of Global Optimization. 2010. V. 48. No. 1. P. 159–172.

A.V. Orlov
Numerical solution of bilinear programming problems.
Computational Mathematics and Mathematical Physics, V.48, No.2, P.225–241, 2008.

A.V. Orlov, A.S. Strekalovsky
Numerical search for equilibria in bimatrix games.
Computational Mathematics and Mathematical Physics, V.45, No.6, P.947–960, 2005.

1a. Pessimistic bilevel problems

$$\left.\begin{array}{c} W(x,\varepsilon) \triangleq \sup\limits_{y}\{F(x,y) \mid y \in Y_*(x,\varepsilon)\} \downarrow \min\limits_{x}, \quad x \in X, \\ Y_*(x,\varepsilon) \triangleq \{y \in Y(x) \mid \langle d, y\rangle \leqslant \inf\limits_{z}[\langle d, z\rangle \mid z \in Y(x)] + \varepsilon\}, \end{array}\right\} (\mathcal{BP}(\varepsilon))$$

$F(x,y) \triangleq \langle c, x\rangle + \langle x, Cx\rangle + \langle c_1, y\rangle - \langle y, C_1 y\rangle, \quad C, C_1 \geqslant 0,$
$X \triangleq \{x \in I\!R^m_+ \mid Ax \leqslant a\}, \ \ Y(x) \triangleq \{y \in I\!R^n_+ \mid A_1 x + B_1 y \leqslant a_1\},$
$A \in I\!R^{p\times m}, \ \ A_1 \in I\!R^{q\times m}, \ \ B_1 \in I\!R^{q\times n}.$

$$\left.\begin{array}{c} F(x,y) \downarrow \min\limits_{x,y}, \quad x \in X, \\ y \in Y_{**}(x,\delta,\nu) \triangleq \{y \in Y(x) \mid \langle d + \nu(C_1 y - c_1), y\rangle \leqslant \\ \leqslant \inf\limits_{z}[\langle d + \nu(C_1 z - c_1), z\rangle \mid z \in Y(x)] + \delta\}, \end{array}\right\} (\mathcal{BP}_o(\delta,\nu))$$

where $\nu > 0$ is a penalty parameter

$(\mathcal{H})$: set X is bounded, set $Y(x)$ is such that $\exists Y$ — compact: $Y(x) \subseteq Y \quad \forall x \in X$.

THEOREM 3

Assume that $(\mathcal{H})$ holds and sequences $\{\tau_k\}, \{\delta_k\}, \{\nu_k\}$ are such that

$$\tau_k \downarrow 0, \quad \delta_k \downarrow 0, \quad \nu_k \downarrow 0, \quad \frac{\delta_k}{\nu_k} \downarrow 0.$$

Then any limit point (x_g, y_g) of sequence $\{(x^k, y^k)\}$ of τ_k-solutions to problems $(\mathcal{BP}_o(\delta_k,\nu_k))$ is a pessimistic solution to problem $(\mathcal{BP}(0))$.

Solving the auxiliary optimistic bilevel problem

$$\left.\begin{array}{c} F(x,y) \triangleq \langle c, x\rangle + \langle x, Cx\rangle + \langle c_1, y\rangle - \langle y, C_1 y\rangle \downarrow \min\limits_{x,y}, \\ Ax \leqslant a, \;\; x \geqslant 0, \\ y \in Y_{**}(x) \triangleq \operatorname*{Argmin}\limits_{y}\{\langle d - \nu c_1, y\rangle + \nu\langle C_1 y, y\rangle \mid \\ A_1 x + B_1 y \leqslant a_1, \;\; y \geqslant 0\}. \end{array}\right\} \qquad (\mathcal{BP}_o(0,\nu))$$

$$\left.\begin{array}{c} F(x,y) \downarrow \min\limits_{x,y,v}, \\ Ax \leqslant a, \;\; x \geqslant 0, \\ A_1 x + B_1 y \leqslant a_1, \;\; y \geqslant 0, \\ d - \nu c_1 + 2\nu C_1 y + vB_1 \geqslant 0, \;\; u \geqslant 0, \\ h(x,y,v) \triangleq \langle d - \nu c_1 + 2\nu C_1 y, y\rangle + \langle a_1 - A_1 x, v\rangle = 0. \end{array}\right\} \qquad (\mathcal{P}(\nu))$$

$$\left.\begin{array}{c} \Phi(x,y,v) \triangleq F(x,y) + \mu h(x,y,v) \downarrow \min\limits_{x,y,v}, \\ D \triangleq \{Ax \leqslant a, \;\; x \geqslant 0, \\ A_1 x + B_1 y \leqslant a_1, \;\; y \geqslant 0, \\ d - \nu c_1 + 2\nu C_1 y + vB_1 \geqslant 0, \;\; v \geqslant 0\}. \end{array}\right\} \qquad (\mathcal{P}(\mu,\nu))$$

where $\mu > 0$ is a penalty parameter.

Global search for pessimistic solutions in random generated bilevel problems

№	$Name$	$LocSol$	Loc_X	St_X	T_X	Loc_V	St_V	T_V
1	$2 \times 4\_2$	3	32	2	0.45	5	2	**0.08**
2	$2 \times 4\_3$	3	33	2	0.32	9	2	**0.10**
3	$2 \times 4\_4$	2	34	2	0.44	6	2	**0.09**
4	$2 \times 4\_6$	3	28	2	0.34	7	1	**0.09**
5	$5 \times 10\_1$	24	138	4	17.84	42	3	**6.70**
6	$5 \times 10\_2$	30	90	4	14.32	77	5	**13.48**
7	$5 \times 10\_3$	30	93	3	11.63	47	3	**7.57**
8	$5 \times 10\_4$	30	105	5	16.38	75	5	**13.11**
9	$5 \times 10\_5$	31	90	3	12.50	46	2	**8.23**
10	$10 \times 20\_1$	896	365	5	6:23.00	94	2	**1:39.74**
11	$10 \times 20\_2$	1016	288	7	5:00.41	159	4	**3:14.98**
12	$10 \times 20\_3$	1020	325	3	**4:29.05**	273	6	4:36.09
13	$10 \times 20\_4$	1023	320	2	11:06.42	162	4	**1:45.24**
14	$10 \times 20\_5$	1022	184	2	5:32.34	59	2	**1:57.09**
15	$20 \times 40\_2$	1048575	839	2	2:10:38.96	353	2	**1:02:27.62**
16	$20 \times 40\_3$	1048575	961	3	1:27:36.88	394	4	**39:59.01**
17	$20 \times 40\_4$	1047552	1127	5	2:28:50.65	595	4	**58:20.47**

Intel Core 2 Duo 1.6GHz CPU

2. Linear complementarity problem

Well-known complementarity problem, which consists in finding a pair of vectors (x, w) satisfying the following condition, have been solved similarly:

$$(LCP): \quad \left. \begin{array}{c} Mx + q = w, \;\; \langle x, w \rangle = 0, \\ x \geqslant 0, \;\; w \geqslant 0, \end{array} \right\}$$

where $x, w \in I\!R^n$, and vector $q \in I\!R^n$ and the real sign-indefinite $(n \times n)$-matrix M are given.

$$(P): \quad \left. \begin{array}{c} F(x) = \langle x, Mx + q \rangle \downarrow \min\limits_x, \\ x \geqslant 0, \;\; Mx + q \geqslant 0. \end{array} \right\}$$

Numerical experiment

Dim: up to 400

Time: 18 min 46 sec (Intel Pentium 4 3.0GHz CPU)

Mazurkevich E. O., Petrova E. G., and Strekalovsky A. S. *On the Numerical Solution of the Linear Complementarity Problem* // Computational Mathematics and Mathematical Physics. 2009, Vol. 49, No. 8, pp. 1318-1331.

3. Problems of financial and medical diagnostics

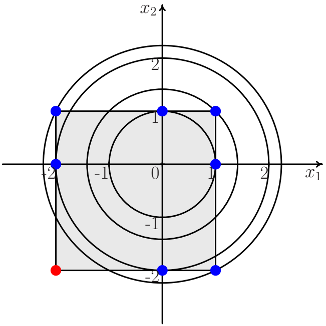
Similar problems have found deverse applications and are known as generalized separability problems. For instance, if two set of points $\mathcal{A}$ and $\mathcal{B}$ are characterized by matrix $\mathcal{A} = [a^1, \ldots, a^M]$, $B = [b^1, \ldots, b^N]$, $a^j, b^j \in I\!R^n$, then the polyhedral separability problem can be reduced to minimization of the non-convex non-differentiable function of error

$$F(V,\Gamma) = F_1(V,\Gamma) + F_2(V,\Gamma), \tag{10}$$

$$\left.\begin{aligned} F_1(V,\Gamma) &= \frac{1}{M}\sum_{i=1}^{M} \max\{0, \max_{1\leqslant p\leqslant P}(\langle a^i, v^p\rangle - \gamma_p + 1)\}, \\ F_2(V,\Gamma) &= \frac{1}{N}\sum_{j=1}^{N} \max\{0; \min_{1\leqslant p\leqslant P}(-\langle b^j, v^p\rangle + \gamma_p + 1)\}. \end{aligned}\right\} \tag{11}$$

A.S. Strekalovsky, A.V.Orlov, *A new approach to nonconvex optimization* // Numerical Methods and Programming, V.8, P.160-176, 2007. (http://num-meth.srcc.msu.su/english/index.html)

$$\|x(1)\|^2 \uparrow \max,$$
$$\dot{x}_i(t) = u_i(t), \quad t \in]0,1[,$$
$$x_i(0) = 0, \quad i = 1, \ldots, n,$$
$$-2 \leqslant u_i(t) \leqslant 1 \quad \overset{\circ}{\forall} t \in]0,1[.$$
$$X(t_1) = \Pi \triangleq \left\{ x \in I\!R^n \mid -2 \leqslant x_i \leqslant 1,\ i = 1, \ldots, n \right\}.$$



2^n is a number of local maxima
$(3^n - 1)$ is a number of processes satisfying Pontryagin's Maximum Principle.

UNIQUE global maximum!!!

I. D.C. mixed minimization:

$$J_0(u) \downarrow \min_u, \quad u(\cdot) \in \mathscr{U},$$

$$J_0(u) = g_1(x(t_1)) - h_1(x(t_1)) + \int_T \Big[g(x(t),u(t),t) - h(x(t),u(t),t) \Big] dt,$$

$g_1(\cdot)$, $h_1(\cdot)$, $g(\cdot)$, $h(\cdot)$ — convex w.r.t. $x \in I\!R^n$.

II. D.C. integral-terminal constraints:

$$J_i(u) \leqslant 0, \; i = 1, \ldots, m,$$

$$J_i(u) = g_i^1(x(t_1)) - h_i^1(x(t_1)) + \int_T \Big[g_i(x(t),u(t),t) - h_i(x(t),u(t),t) \Big] dt,$$

$g_i^1(\cdot)$, $h_i^1(\cdot)$, $g_i(\cdot)$, $h_i(\cdot)$ — convex w.r.t. $x(\cdot)$.

III. Non-linear control systems:

$$\dot{x} = f(x(t),u(t),t), \quad x(t_0) = x^0 \in R^n,$$

$$f_j(x(t),u(t),t) = g_j(x(t),u(t),t) - h_j(x(t),u(t),t), \; j = 1, \ldots, n.$$

$$(\mathcal{P}): \quad J(u) \triangleq F_1(x(t_1)) + \int\limits_T [F(x(t),t)] \downarrow \min_{x,u}, \tag{12}$$

$$\dot{x}(t) = f(x(t),u(t),t) \quad \overset{\circ}{\forall} t \in T \triangleq [t_0,t_1], \quad x(t_0) = x_0, \tag{13}$$

$$u \in \mathscr{U} = \{\, u(\cdot) \in L^r_\infty(T) \mid u(t) \in U \quad \overset{\circ}{\forall} t \in T \,\}. \tag{14}$$

$$F_1(x) = g_1(x) - h_1(x) \quad \forall x \in I\!R^n,$$
$$F(x,t) = g(x,t) - h(x,t) \quad \forall x \in I\!R^n, \quad t \in T.$$

Assumptions:

- $g_1(\cdot)$, $h_1(\cdot)$, $x \mapsto g(x,t)\colon I\!R^n \to I\!R$ и $x \mapsto h(x,t)\colon I\!R^n \to I\!R$, $t \in T$ are convex functions on $I\!R^n$;
- $t \mapsto g(x,t)$ и $t \mapsto h(x,t)$ are continuous functions;
- $f\colon I\!R^n \times I\!R^r \times I\!R \to I\!R^n$ is continuous vector-function w.r.t. variables $x \in I\!R^n, \quad u \in I\!R^r, \quad t \in [t_0,t_1]$;
- for all $(x,y) \in I\!R^n \times I\!R^n$ и $(u,t) \in I\!R^r \times [t_0,t_1]$ $\|f(x,u,t) - f(y,u,t)\| \leqslant L(t)\|x-y\|$, where $L(t) > 0$ and $L(\cdot) \in L_1(T)$;
- there exist continuous derivatives $\dfrac{\partial f_i(x,u,t)}{\partial x_j} \quad \forall (x,u,t) \in \Omega$;
- $U \subset I\!R^r$ is compact set.

Let D is set of all feasible processes for the problem $(\mathcal{P})$. Suppose $(x^s(\cdot), u^s(\cdot)) \in D$.

The linearized problem $(\mathcal{P}L_s)$

$$I_s(x(\cdot), u(\cdot)) = g_1(x(t_1)) - \langle h_1'(x^s(t_1)), x(t_1)\rangle + \\ + \int\limits_T \Big[g(x(t), t) - \langle h'(x^s(t), t), x(t)\rangle\Big]\, dt \downarrow \min, \quad (x(\cdot), u(\cdot)) \in D, \qquad (15)$$

where $h_1'(x^s(t_1)) \in \partial h_1(x^s(t_1))$, $h'(x^s(t), t) \in \partial h(x^s(t), t)$, $t \in T$.

The following process $(x^{s+1}(\cdot), u^{s+1}(\cdot))$ is constructed as process satisfying PMP:

$$\begin{gathered} \dot{\psi}^s(t) = -A(t)^\top \psi^s(t) + g'(x^{s+1}(t), t) - h'(x^s(t), t), \quad t \in T, \\ \psi^s(t_1) = h_1'(x^s(t_1)) - g_1'(x^{s+1}(t_1)), \\ g_1'(x^{s+1}(t_1)) \in \partial g_1(x^{s+1}(t_1)), \quad g'(x^{s+1}(t), t) \in \partial g(x^{s+1}(t), t),\ t \in T, \end{gathered} \qquad (16)$$

$$\delta_s > 0, \quad s = 0, 1, 2..., \quad \sum_{s=0}^{\infty} \delta_s < +\infty, \qquad (17)$$

$$\langle \psi^s(t), b(u^{s+1}(t), t)\rangle + \frac{\delta_s}{t_1 - t_0} \geqslant \sup_{v \in U} \langle \psi^s(t), b(v, t)\rangle \quad \overset{\circ}{\forall} t \in T. \qquad (18)$$

THEOREM 4

Let the sequence of processes $\{x^s(\cdot), u^s(\cdot)\}$ be constructed by the rule (16)–(18). Then the following condition is satisfied:

$$\lim_{s\to\infty} \sup_v \{\langle \psi^s(t), b(v,t) - b(u^s(t),t)\rangle \mid v \in U\} = 0 \qquad \overset{\circ}{\forall} t \in T,$$

Furthermore, the numerical sequence $\{J(x^s(\cdot), u^s(\cdot))\}$ converges.
The numerical sequence $\{I_s(x^s(\cdot), u^s(\cdot))\}$ also converges and the limited equation is fulfilled, as follows,

$$\lim_{s\to\infty} (I_s(x^s(\cdot), u^s(\cdot)) - \mathcal{V}(PL_s)) = 0.$$

THEOREM 5

Let functions $h_1(\cdot)$ and $h(\cdot, t)$, $t \in T$ be strongly convex. Then the state sequence $\{x^s(\cdot)\}$, which is generated by the rule (16)–(18), converge, as follows,

$$x^s(t_1) \to x_1^* \text{ in } I\!R^n,$$

$$x^s(\cdot) \to x_*(\cdot) \text{ in } L_2(T).$$

Let us introduce the notation

$$\zeta := J(w) = g_1(z(t_1)) - h_1(z(t_1)) + \int_T [g(z(t),t) - h(z(t),t)]dt, \tag{19}$$

where $z(t) = x(t,w),\ t \in [t_0, t_1],\ w(\cdot) \in \mathcal{U}.$

THEOREM 6. (*Necessary GOC*)

Assume that the process $(z(\cdot), w(\cdot)),\ z(t) = x(t,w),\ t \in T,\ w(\cdot) \in \mathcal{U}$, is globally optimal in Problem $(\mathcal{P})$–(12)–(14).
Then for each triplet $(y(\cdot), p, \beta)$ such that $y(\cdot) \in L^n_\infty(T),\ p \in I\!R^n,\ \ \beta \in I\!R$, satisfying the equality

$$h_1(p) + \int_T h(y(t),t)dt = \beta - \zeta,$$

the following variational inequality holds $(h'(y(t),t) \in \partial h(y(t),t),\ h_1'(p) \in \partial h_1(p))$

$$g_1(x(t_1,u)) - \langle h_1'(p), x(t_1,u)\rangle + \int_T [g(x(t,u),t) - \langle h'(y(t),t), x(t,u)\rangle]dt \geqslant$$

$$\geqslant \beta - \langle h_1'(p), p\rangle - \int_T \langle h'(y(t),t), y(t)\rangle dt \ \ \forall u(\cdot) \in \mathcal{U}.$$

Consider the assumption $(\mathcal{H})$:
for a feasible control $\tilde{u}(\cdot) \in \mathscr{U}$ the following inequality holds

$$J(\tilde{u}) = F_1(\tilde{x}(t_1)) + \int\limits_T F(\tilde{x}(t), t)\, dt > \zeta, \tag{20}$$

where $\tilde{x}(t) = x(t, \tilde{u})$, $t \in T$ and $\zeta = J(w)$.

THEOREM 7. (*Necessary and sufficient GOC*)

Suppose, in Problem $(\mathcal{P})$–(12)–(14) the assumption $(\mathcal{H})$–(20) is satisfied. Hence for a process $(z(\cdot), w(\cdot))$, to be a globally optimal, it is necessary and sufficient that

$(\mathcal{E})$: the following variational inequality would hold for each perturbation pair $(y(\cdot), \beta)$, where $y(t)$ is absolutely continuous, $h_1(y(t_1)) + \int\limits_T h(y(t), t)dt = \beta - \zeta$, $\beta \in I\!R$:

$$g_1(y(t_1)) + \int\limits_T g(y(t), t)\, dt \leqslant \beta \leqslant \sup_u \{g_1(x(t_1, u)) + \int\limits_T g(x(t, u), t)\, dt \mid u \in \mathscr{U}\},$$

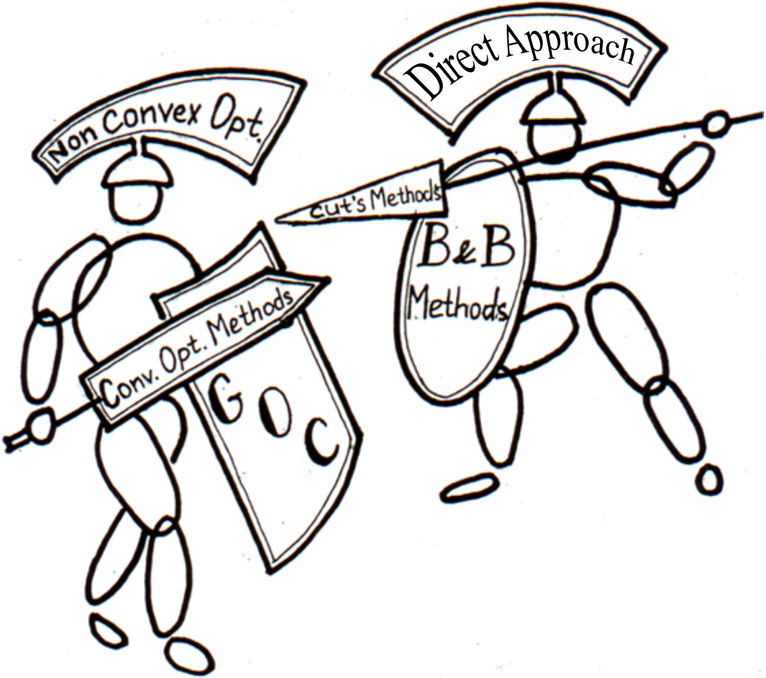
$$\mathcal{V}(\mathcal{P}L(y(\cdot))) \geqslant N(y, \beta) := \beta - \langle h_1'(y(t_1)), y(t_1)\rangle - \int\limits_T \langle h'(y(t), t), y(t)\rangle\, dt,$$

where $h_1'(y(t_1)) \in \partial h_1(y(t_1))$, $h'(y(t), t) \in \partial h(y(t), t)$.

n	r	PMP	J_0	Set $\mathcal{R}_1$			Set $\mathcal{R}_2$		
				J_*	St	Time	J_*	St	Time
6	**6**	**27**	7.23	**-11.09**	10	1:26.24	-14.57	12	1:35.47
6	6	18	3.47	-19.43	11	1:39.27	-19.43	9	1:31.16
6	4	12	4.15	-12.10	10	1:15.86	-12.10	8	1:12.11
6	5	8	-2.39	-14.75	4	1:54.14	-14.75	5	1:37.19
10	9	162	-3.18	-21.89	15	2:19.32	-21.89	17	2:29.62
10	8	108	-2.75	-22.69	14	2:23.21	-22.69	16	2:32.15
10	9	72	3.16	-25.36	15	2:54.56	-25.36	15	2:47.29
10	8	72	-5.11	-28.03	16	2:43.10	-28.03	14	2:25.72
14	12	972	-6.57	-33.29	21	4:59.32	-33.29	20	4:51.10
14	13	648	-5.95	-32.36	20	5:18.12	-32.36	19	5:14.18
14	12	648	-3.12	-35.02	21	6:12.45	-35.02	23	6:57.92
14	10	288	-4.59	-33.02	17	5:37.11	-33.02	16	5:31.16
20	**18**	**7776**	-2.57	**-43.98**	25	10:32.15	-48.72	27	12:27.11
20	17	7776	-9.18	-48.07	28	11:10.45	-48.07	29	12:03.44
20	**18**	**17496**	-6.36	-56.19	37	18:53.19	**-52.33**	36	18:05.84
20	16	7776	-5.18	-47.14	25	11:24.47	-47.14	25	11:49.17
20	20	11664	3.12	-52.73	26	12:52.11	-52.73	29	13:41.62
20	20	59048	-3.95	-63.19	49	17:19.54	-63.19	47	16:58.12

№	Generation				n	r	PMP	$\mathcal{V}(\mathcal{P})$	J_0	J_*	St	Time (min:sec)
	m_1	m_2	m_3	m_4								
1	1	1	0	0	4	4	6	-15.35	-3.92	-15.35	2	2:14.47
2	0	1	1	0	4	3	4	9.33	33.19	9.33	3	2:23.16
3	1	0	0	1	4	4	6	-29.03	-7.42	-29.03	3	2:07.14
4	2	1	0	0	6	6	18	-33.88	-10.77	-33.88	6	4:58.49
5	0	2	0	1	6	6	8	4.14	19.3	-4.14	4	4:21.24
6	3	1	1	0	10	9	108	-46.26	-17.65	-46.26	8	4:24.38
7	2	1	2	0	10	8	72	-21.58	-5.39	-21.58	9	5:07.62
8	1	1	1	2	10	9	48	-30.20	-10.18	-30.20	10	5:52.82
9	0	2	2	1	10	8	32	8.16	20.51	8.16	9	6:45.51
10	3	2	2	0	14	12	432	-36.93	3.74	-36.93	11	7:35.73
11	3	1	1	2	14	13	432	-67.26	-21.95	-67.26	12	7:42.89
12	2	2	2	1	14	12	288	-28.90	-4.69	-28.90	14	10:05.21
13	1	1	4	1	14	10	192	-1.25	15.40	-1.25	11	7:53.27
14	3	2	2	3	20	18	3456	**-68.43**	-36.15	**-56.13**	14	11:32.18
15	3	2	3	2	20	17	3456	51.78	11.82	-51.78	14	12:05.36
16	2	5	2	1	20	18	2304	-19.36	-1.46	-19.36	13	14:31.19
17	3	2	4	1	20	16	3456	**-35.13**	-9.29	**-33.29**	11	12:21.77
18	3	3	0	4	20	20	3456	-88.05	-17.33	-88.05	16	17:46.23
19	7	1	1	1	20	19	17496	-130.88	-41.19	-130.88	17	18:37.11

- A.S. Strekalovsky, M.V. Yanulevich
  Global Search in the Optimal Control Problem with a Terminal Objective Functional Represented as the Difference of Two Convex Functions.
  Journal of Computational Mathematics and Mathematical Physics.
  Vol. 48. No. 7, 2008. pp. 1119–1132.
- A.S. Strekalovsky, E.V. Sharankhaeva
  Global Search in a Non-Convex Optimal Control Problem.
  Journal of Computational Mathematics and Mathematical Physics,
  Vol. 45. No. 10, 2005. pp. 1719–1734.
- A.S. Strekalovsky
  On Global Maximum of a Convex Terminal Functionals in Optimal Control Problems.
  Journal of Global Optimazation. No. 7, 1995. pp. 75–91.
- A.S. Strekalovsky
  On Extremal Problems on Complements of Convex Sets.
  Cybernetics and System Analysis, No. 1, 1993. pp. 113–126.


Non Convex Opt.
Direct Approach
Cut's Methods
B&B
Methods
Conv. Opt. Methods
G
O
C

THANK YOU!