Primal-dual optimization methods for solving optimal control problems with saddle-point structure

Pavel Dvurechensky[1]

PreMoLab/MIPT

21.06.13

[1] joint work with Yurii Nesterov

Problem statement

We have two objects with motion given by equations

$$\frac{dx(t)}{dt} = A_x(t)x(t) + B(t)u(t), \quad \frac{dy(t)}{dt} = A_y(t)y(t) + C(t)v(t),$$

$(x(0), y(0)) = (x_0, y_0)$. $x(t) \in \mathbb{R}^n$, $y(t) \in \mathbb{R}^m$, $u(t) \in P \subset \mathbb{R}^p$, $v(t) \in Q \subset \mathbb{R}^q$, $t \in [0, \theta]$.
P, Q closed convex sets.

Problem statement

We have two objects with motion given by equations

$$\frac{dx(t)}{dt} = A_x(t)x(t) + B(t)u(t), \quad \frac{dy(t)}{dt} = A_y(t)y(t) + C(t)v(t),$$

$(x(0), y(0)) = (x_0, y_0)$. $x(t) \in \mathbb{R}^n$, $y(t) \in \mathbb{R}^m$, $u(t) \in P \subset \mathbb{R}^p$, $v(t) \in Q \subset \mathbb{R}^q$, $t \in [0, \theta]$.
P, Q closed convex sets.
Performance index:

$$F(u, v) + \Phi(x, y) = \int_0^\theta \tilde{F}(\tau, u(\tau), v(\tau))d\tau + \Phi(x(\theta), y(\theta)). \qquad (1)$$

Problem statement

We have two objects with motion given by equations

$$\frac{dx(t)}{dt} = A_x(t)x(t) + B(t)u(t), \quad \frac{dy(t)}{dt} = A_y(t)y(t) + C(t)v(t),$$

$(x(0), y(0)) = (x_0, y_0)$. $x(t) \in \mathbb{R}^n$, $y(t) \in \mathbb{R}^m$, $u(t) \in P \subset \mathbb{R}^p$,
$v(t) \in Q \subset \mathbb{R}^q$, $t \in [0, \theta]$.
P, Q closed convex sets.
Performance index:

$$F(u, v) + \Phi(x, y) = \int_0^\theta \tilde{F}(\tau, u(\tau), v(\tau))d\tau + \Phi(x(\theta), y(\theta)). \quad (1)$$

We are looking for $u(t) \in L_2([0, \theta], P)$, $v(t) \in L_2([0, \theta], Q)$ s.t. (u, v) is the saddle point of (1).

Problem statement

Let us introduce operators
$\mathcal{B} : L_2([0,\theta], P) \to \mathbb{R}^n, \quad \mathcal{C} : L_2([0,\theta], Q) \to \mathbb{R}^m$:

$$x(\theta) = V_x(\theta, 0)x_0 + \int_0^\theta V_x(\theta,\tau)B(\tau)u(\tau)d\tau = x_0 + \mathcal{B}u,$$

$$y(\theta) = V_y(\theta, 0)y_0 + \int_0^\theta V_y(\theta,\tau)C(\tau)v(\tau)d\tau = y_0 + \mathcal{C}v.$$

Problem statement

Let us introduce operators
$\mathcal{B} : L_2([0,\theta], P) \to \mathbb{R}^n, \quad \mathcal{C} : L_2([0,\theta], Q) \to \mathbb{R}^m$:

$$x(\theta) = V_x(\theta,0)x_0 + \int_0^\theta V_x(\theta,\tau)B(\tau)u(\tau)d\tau = x_0 + \mathcal{B}u,$$

$$y(\theta) = V_y(\theta,0)y_0 + \int_0^\theta V_y(\theta,\tau)C(\tau)v(\tau)d\tau = y_0 + \mathcal{C}v.$$

We assume that

Problem statement

Let us introduce operators
$\mathcal{B} : L_2([0,\theta],P) \to \mathbb{R}^n, \quad \mathcal{C} : L_2([0,\theta],Q) \to \mathbb{R}^m$:

$$x(\theta) = V_x(\theta,0)x_0 + \int_0^\theta V_x(\theta,\tau)B(\tau)u(\tau)d\tau = x_0 + \mathcal{B}u,$$

$$y(\theta) = V_y(\theta,0)y_0 + \int_0^\theta V_y(\theta,\tau)C(\tau)v(\tau)d\tau = y_0 + \mathcal{C}v.$$

We assume that

- **A1** The sets P, Q are bounded.

Problem statement

Let us introduce operators
$\mathcal{B} : L_2([0,\theta], P) \to \mathbb{R}^n, \quad \mathcal{C} : L_2([0,\theta], Q) \to \mathbb{R}^m$:

$$x(\theta) = V_x(\theta,0)x_0 + \int_0^{\theta} V_x(\theta,\tau)B(\tau)u(\tau)d\tau = x_0 + \mathcal{B}u,$$

$$y(\theta) = V_y(\theta,0)y_0 + \int_0^{\theta} V_y(\theta,\tau)C(\tau)v(\tau)d\tau = y_0 + \mathcal{C}v.$$

We assume that

- **A1** The sets P, Q are bounded.
- **A2** In performance index $F(u,v) + \Phi(x,y)$ the functional $F(\cdot,v)$ is convex for any fixed v, $F(u,\cdot)$ is concave for any fixed u, $\Phi(\cdot,y)$ is convex for any fixed y, $\Phi(x,\cdot)$ is concave for any fixed x.

Problem statement

Let us introduce operators
$\mathcal{B} : L_2([0,\theta], P) \to \mathbb{R}^n, \quad \mathcal{C} : L_2([0,\theta], Q) \to \mathbb{R}^m$:

$$x(\theta) = V_x(\theta,0)x_0 + \int_0^\theta V_x(\theta,\tau)B(\tau)u(\tau)d\tau = x_0 + \mathcal{B}u,$$

$$y(\theta) = V_y(\theta,0)y_0 + \int_0^\theta V_y(\theta,\tau)C(\tau)v(\tau)d\tau = y_0 + \mathcal{C}v.$$

We assume that

- **A1** The sets P, Q are bounded.
- **A2** In performance index $F(u,v) + \Phi(x,y)$ the functional $F(\cdot, v)$ is convex for any fixed v, $F(u,\cdot)$ is concave for any fixed u, $\Phi(\cdot, y)$ is convex for any fixed y, $\Phi(x,\cdot)$ is concave for any fixed x.
- **A3** Operators $\mathcal{B}, \mathcal{C}$ are bounded.

Problem statement

Let us introduce operators
$\mathcal{B} : L_2([0,\theta], P) \to \mathbb{R}^n, \quad \mathcal{C} : L_2([0,\theta], Q) \to \mathbb{R}^m$:

$$x(\theta) = V_x(\theta,0)x_0 + \int_0^\theta V_x(\theta,\tau)B(\tau)u(\tau)d\tau = x_0 + \mathcal{B}u,$$

$$y(\theta) = V_y(\theta,0)y_0 + \int_0^\theta V_y(\theta,\tau)C(\tau)v(\tau)d\tau = y_0 + \mathcal{C}v.$$

We assume that

- **A1** The sets P, Q are bounded.
- **A2** In performance index $F(u,v) + \Phi(x,y)$ the functional $F(\cdot,v)$ is convex for any fixed v, $F(u,\cdot)$ is concave for any fixed u, $\Phi(\cdot,y)$ is convex for any fixed y, $\Phi(x,\cdot)$ is concave for any fixed x.
- **A3** Operators $\mathcal{B}, \mathcal{C}$ are bounded.

Hence $x(\theta), y(\theta)$ are bounded and we can assume that $x(\theta) \in X, y(\theta) \in Y$, where X, Y are closed convex bounded.

Transform problem

We get the problem:

$$\min_{u \in \mathcal{U}} \left[\max_{v \in \mathcal{V}} \{ F(u,v) + \Phi(x,y) : y = y_0 + \mathcal{C}v \} : x = x_0 + \mathcal{B}u \right]$$

Transform problem

We get the problem:

$$\min_{u\in\mathcal{U}}\left[\max_{v\in\mathcal{V}}\{F(u,v)+\Phi(x,y): y=y_0+\mathcal{C}v\} : x=x_0+\mathcal{B}u\right]$$

Adjoint problem:

$$\min_{\lambda}\max_{\mu}\left\{\min_{u\in\mathcal{U}}\max_{v\in\mathcal{V}}[F(u,v)-\langle\mu,\mathcal{B}u\rangle+\langle\lambda,\mathcal{C}v\rangle]+\right.$$

$$\left.+\min_{x\in X}\max_{y\in Y}[\Phi(x,y)+\langle\mu,x\rangle-\langle\lambda,y\rangle]-\langle\mu,x_0\rangle+\langle\lambda,y_0\rangle\right\}=$$

$$=\min_{\lambda}\max_{\mu}\psi(\lambda,\mu)$$

Finite-dimensional VI problem

Denote $z = (\lambda, \mu)$. The problem is to find z^* s.t.

$$\langle g(z), z - z^* \rangle \geq 0 \quad \forall z \in \mathbb{R}^n \times \mathbb{R}^m,$$

where $g(z) = (\psi'_\lambda(\lambda, \mu), -\psi'_\mu(\lambda, \mu))$ – bounded and monotone ($\langle g(z_1) - g(z_2), z_1 - z_2 \rangle \geq 0$) operator.

Finite-dimensional VI problem

Denote $z = (\lambda, \mu)$. The problem is to find z^* s.t.

$$\langle g(z), z - z^* \rangle \geq 0 \quad \forall z \in \mathbb{R}^n \times \mathbb{R}^m,$$

where $g(z) = (\psi'_\lambda(\lambda, \mu), -\psi'_\mu(\lambda, \mu))$ – bounded and monotone $(\langle g(z_1) - g(z_2), z_1 - z_2 \rangle \geq 0)$ operator.

Some auxiliary objects

- Prox-function $d(z)$ – strongly convex with parameter $\sigma > 0$ ($d''(z) - \sigma I \succeq 0$)
- $D : d(z^*) \leq D$
- $\mathcal{F}_D = \{z \in S : d(z) \leq D\}$
- $\pi_\beta(s) = \arg\min_{z \in S}\{-\langle s, z \rangle + \beta d(z)\}$
- for the sequences $\lambda_i \geq 0, z_i \in S, g_i, i = 0, \ldots, k$

$$\delta_k(D) = \max_z \left\{ \sum_{i=0}^{k} \lambda_i \langle g_i, z_i - z \rangle : \quad z \in \mathcal{F}_D \right\}$$

- sequence $\hat{\beta}_i$: $\quad \hat{\beta}_0 = \hat{\beta}_1 = 1, \quad \hat{\beta}_{i+1} = \hat{\beta}_i + \frac{1}{\hat{\beta}_i}$. $(\hat{\beta}_k \sim \sqrt{2k})$

Simple Dual Averages Method [Yu. Nesterov, 2009]

Initialization $s_0 = 0, z_0, \gamma > 0$
Step $k \geq 0$

1. $g_k = g(z_k)$. $\quad s_{k+1} = s_k + g_k$.
2. $\beta_{k+1} = \gamma \hat{\beta}_{k+1}$. $\quad z_{k+1} = \pi_{\beta_{k+1}}(-s_{k+1})$.

Simple Dual Averages Method [Yu. Nesterov, 2009]

Initialization $s_0 = 0, z_0, \gamma > 0$
Step $k \geq 0$

1. $g_k = g(z_k)$. $\quad s_{k+1} = s_k + g_k$.
2. $\beta_{k+1} = \gamma \hat{\beta}_{k+1}$. $\quad z_{k+1} = \pi_{\beta_{k+1}}(-s_{k+1})$.

Theorem [Yu. Nesterov, 2009]

Assume that $\|g_k\| \leq L, k \geq 0$. Then

Simple Dual Averages Method [Yu. Nesterov, 2009]

Initialization $s_0 = 0, z_0, \gamma > 0$
Step $k \geq 0$

1. $g_k = g(z_k)$. $\quad s_{k+1} = s_k + g_k$.
2. $\beta_{k+1} = \gamma \hat{\beta}_{k+1}$. $\quad z_{k+1} = \pi_{\beta_{k+1}}(-s_{k+1})$.

Theorem [Yu. Nesterov, 2009]

Assume that $\|g_k\| \leq L, k \geq 0$. Then

1. $\delta_k(D) \leq \hat{\beta}_{k+1} \left(\gamma D + \frac{L^2}{2\sigma\gamma} \right)$

Simple Dual Averages Method [Yu. Nesterov, 2009]

Initialization $s_0 = 0, z_0, \gamma > 0$
Step $k \geq 0$

1. $g_k = g(z_k)$. $\quad s_{k+1} = s_k + g_k$.
2. $\beta_{k+1} = \gamma \hat{\beta}_{k+1}$. $\quad z_{k+1} = \pi_{\beta_{k+1}}(-s_{k+1})$.

Theorem [Yu. Nesterov, 2009]

Assume that $\|g_k\| \leq L, k \geq 0$. Then

1. $\delta_k(D) \leq \hat{\beta}_{k+1}\left(\gamma D + \frac{L^2}{2\sigma\gamma}\right)$
2. If the solution z^* exists, then $\|z_k - z^*\|^2 \leq \frac{2}{\sigma} d(z^*) + \frac{L^2}{\sigma^2\gamma^2}$

Equivalent problem

- Denote

$$\phi(u,x,v,y) = \min_{\lambda} \max_{\mu}$$
$$\{F(u,v) + \Phi(x,y) + \langle \mu, x - x_0 - \mathcal{B}u \rangle + \langle \lambda, y_0 + \mathcal{C}v - y \rangle :$$
$$d_\lambda(\lambda) \le D_\lambda, d_\mu(\mu) \le D_\mu\}$$

- Then our problem is equivalent to $\min\limits_{u \in \mathcal{U}, x \in X} \max\limits_{v \in \mathcal{V}, y \in Y} \phi(u,x,v,y)$.

Equivalent problem

- Denote

$$\phi(u,x,v,y) = \min_{\lambda} \max_{\mu}$$
$$\{F(u,v) + \Phi(x,y) + \langle \mu, x - x_0 - \mathcal{B}u \rangle + \langle \lambda, y_0 + \mathcal{C}v - y \rangle :$$
$$d_\lambda(\lambda) \le D_\lambda, d_\mu(\mu) \le D_\mu\}$$

- Then our problem is equivalent to $\min\limits_{u \in \mathcal{U}, x \in X} \max\limits_{v \in \mathcal{V}, y \in Y} \phi(u,x,v,y)$.
- Denote

$$\xi(u,x) = \max_{v \in \mathcal{V}, y \in Y} \phi(u,x,v,y)$$

$$\eta(v,y) = \min_{u \in \mathcal{U}, x \in X} \phi(u,x,v,y)$$

Equivalent problem

- Denote

$$\phi(u,x,v,y) = \min_{\lambda} \max_{\mu}$$
$$\{F(u,v) + \Phi(x,y) + \langle \mu, x - x_0 - \mathcal{B}u\rangle + \langle \lambda, y_0 + \mathcal{C}v - y\rangle :$$
$$d_\lambda(\lambda) \le D_\lambda, d_\mu(\mu) \le D_\mu\}$$

- Then our problem is equivalent to $\min\limits_{u\in\mathcal{U},x\in X} \max\limits_{v\in\mathcal{V},y\in Y} \phi(u,x,v,y)$.
- Denote

$$\xi(u,x) = \max_{v\in\mathcal{V},y\in Y} \phi(u,x,v,y)$$
$$\eta(v,y) = \min_{u\in\mathcal{U},x\in X} \phi(u,x,v,y)$$

- Then
$\xi(u,x) \ge \phi(u^*,x^*,v^*,y^*) \ge \eta(v,y) \quad \forall u \in \mathcal{U}, v \in \mathcal{V}, x \in X, y \in Y$

Main result

Denote

$$\hat{u}_{k+1} = \frac{1}{k+1}\sum_{i=0}^{k} u_i, \quad \hat{v}_{k+1} = \frac{1}{k+1}\sum_{i=0}^{k} v_i,$$
$$\hat{x}_{k+1} = \frac{1}{k+1}\sum_{i=0}^{k} x_i, \quad \hat{y}_{k+1} = \frac{1}{k+1}\sum_{i=0}^{k} y_i$$

(u_i, v_i, x_i, y_i) s.-p. defining $\psi(\lambda_i, \mu_i)$, where (λ_i, μ_i) is generated by SDA.

Main result

Denote

$$\hat{u}_{k+1} = \frac{1}{k+1}\sum_{i=0}^{k} u_i, \quad \hat{v}_{k+1} = \frac{1}{k+1}\sum_{i=0}^{k} v_i,$$

$$\hat{x}_{k+1} = \frac{1}{k+1}\sum_{i=0}^{k} x_i, \quad \hat{y}_{k+1} = \frac{1}{k+1}\sum_{i=0}^{k} y_i$$

(u_i, v_i, x_i, y_i) s.-p. defining $\psi(\lambda_i, \mu_i)$, where (λ_i, μ_i) is generated by SDA.

Theorem 1

$$\xi(\hat{u}_{k+1}, \hat{x}_{k+1}) - \eta(\hat{v}_{k+1}, \hat{y}_{k+1}) \leq \frac{1}{k+1}\delta_k(D) = O\left(\frac{1}{\sqrt{k}}\right)$$

Main result

Denote

$$\hat{u}_{k+1} = \frac{1}{k+1}\sum_{i=0}^{k} u_i, \quad \hat{v}_{k+1} = \frac{1}{k+1}\sum_{i=0}^{k} v_i,$$

$$\hat{x}_{k+1} = \frac{1}{k+1}\sum_{i=0}^{k} x_i, \quad \hat{y}_{k+1} = \frac{1}{k+1}\sum_{i=0}^{k} y_i$$

(u_i, v_i, x_i, y_i) s.-p. defining $\psi(\lambda_i, \mu_i)$, where (λ_i, μ_i) is generated by SDA.

Theorem 1

$$\xi(\hat{u}_{k+1}, \hat{x}_{k+1}) - \eta(\hat{v}_{k+1}, \hat{y}_{k+1}) \leq \frac{1}{k+1}\delta_k(D) = O\left(\frac{1}{\sqrt{k}}\right)$$

Theorem 2

$$\|x_0 + \mathcal{B}\hat{u}_{k+1} - \hat{x}_{k+1}\| = O\left(\frac{1}{\sqrt{k}}\right), \quad \|y_0 + \mathcal{C}\hat{v}_{k+1} - \hat{y}_{k+1}\| = O\left(\frac{1}{\sqrt{k}}\right)$$

Numerical example

We have two objects with motion given by equations

$$\frac{dx(t)}{dt} = \begin{pmatrix} 1-t \\ 1 \end{pmatrix} u(t), \quad \frac{dy(t)}{dt} = \begin{pmatrix} 1-t \\ 1 \end{pmatrix} v(t), \quad u(t) \in P, v(t) \in Q$$

$$t \in [0,1], \quad n = 2, \quad m = 2,\ P = Q = [-1,1].$$

Numerical example

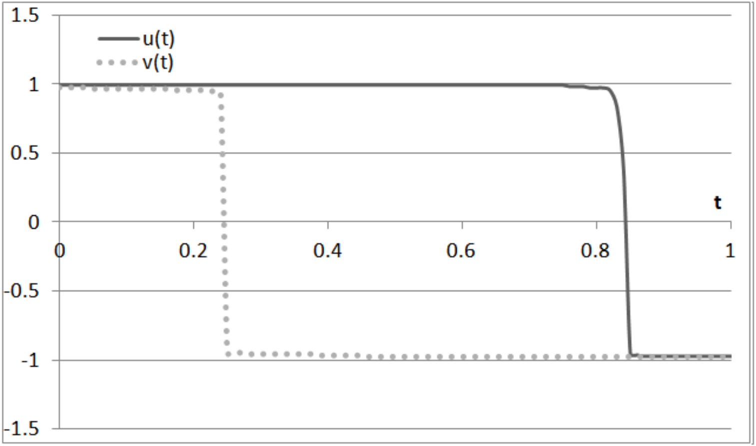
We have two objects with motion given by equations

$$\frac{dx(t)}{dt} = \begin{pmatrix} 1-t \\ 1 \end{pmatrix} u(t), \quad \frac{dy(t)}{dt} = \begin{pmatrix} 1-t \\ 1 \end{pmatrix} v(t), \quad u(t) \in P, v(t) \in Q$$

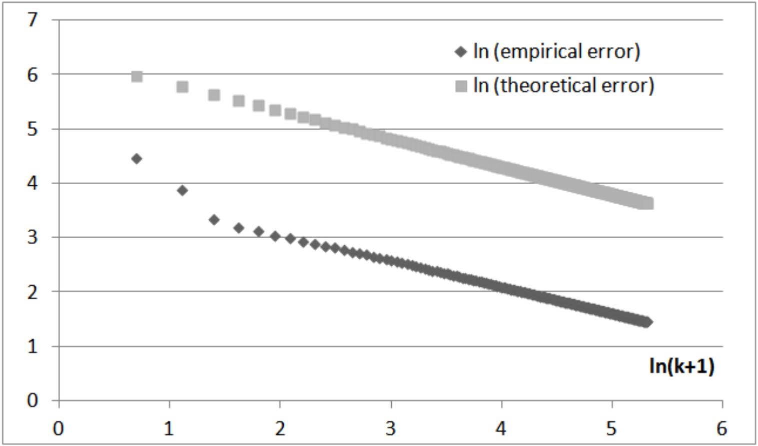
$t \in [0,1], \quad n = 2, \quad m = 2,\ P = Q = [-1,1].$

$$J(u,v) = \frac{1}{2}\|x(1) - y(1)\|^2 - \|y(1) - a\|^2$$

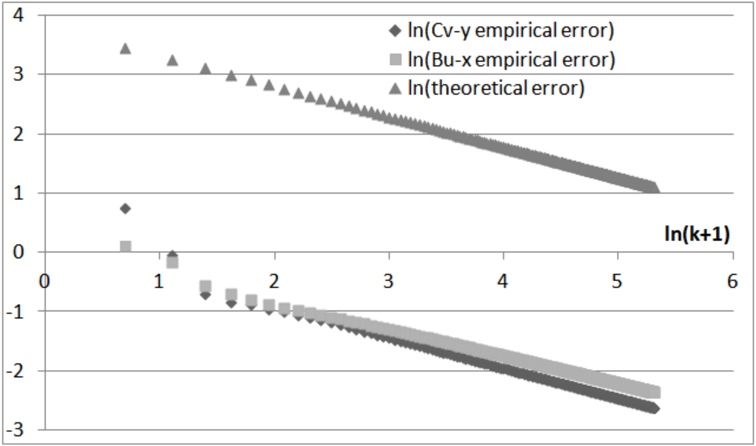
Controls



Error of funcional value



Error of equality consraint



Conclusion

- We consider convex-concave saddle-point optimal control problem (differential game)

Conclusion

- We consider convex-concave saddle-point optimal control problem (differential game)
- We propose an algorithm for finding an approximation to its saddle-point

Conclusion

- We consider convex-concave saddle-point optimal control problem (differential game)
- We propose an algorithm for finding an approximation to its saddle-point
- We estimate the complexity of the proposed algorithm

Conclusion

- We consider convex-concave saddle-point optimal control problem (differential game)
- We propose an algorithm for finding an approximation to its saddle-point
- We estimate the complexity of the proposed algorithm
- We perform numerical experiments and show that they are in consistency with theoretical estimates of complexity

Conclusion

- We consider convex-concave saddle-point optimal control problem (differential game)
- We propose an algorithm for finding an approximation to its saddle-point
- We estimate the complexity of the proposed algorithm
- We perform numerical experiments and show that they are in consistency with theoretical estimates of complexity

Thank you for your attention!