Lecture 11: Average Sub-Gradient Method as a Version of Integral Sliding Mode Control

Plan of presentation

- Where it was published
- Model description and problem setting
- Accepted assumptions
- Problem formulation
- Examples of loss-functions
- Desired regime
- Functional convergence
- Main theorem on ASG robust controller

Where it was published

1. Poznyak, A.S.; Nazin A.V.; Alazki H., Integral Sliding Mode Convex Optimization in Uncertain Lagrangian Systems Driven by PMDC Motors: Averaged Subgradient Approach. *IEEE Transactions on Automatic Control*, 2021, 66(9), 4267–4273.
2. Alexander Nazin, Hussain Alazki, and Alexander Poznyak, Robust Tracking as Constrained Optimization by Uncertain Dynamic Plant: Mirror Descent Method and ASG – Version of Integral Sliding Mode Control. *Mathematics 2023, MDPI* (to be published).
3. A.V. Nazin and A. S. Poznyak, Non-quadratic proxy functions in Mirror Descent Method applied to designing of robust controllers for nonlinear dynamic systems with uncertainty. *Computational Mathematics and Mathematical Physics*, Springer (to be published).

Model description and problem setting

- Here we will deal with the construction of a feedback, which designing is very close to the ISM approach, together with the, so-called, **Averaged Sub-Gradient (ASG)** Technique.
- Consider the dynamic model of a Lagrangian mechanical system with n-degrees of freedom in the standard form given by the following set of differential equations:

$$\boxed{\begin{array}{c} D(q(t))\ddot{q}(t) + C(q(t), \dot{q}(t))\dot{q}(t) + G(q(t)) \\ = \tau(t) + \vartheta(t), \end{array}} \tag{1}$$

where $q(t), \dot{q}(t) \in R^n$ are the state vectors (generalized coordinates and their velocities, $t \geq 0$), $\tau(t) \in R^n$ is a vector of external torques (control) acting to the mechanical system, and $\vartheta(t) \in R^n$ is the disturbance (or uncertainty) vector.

Model of tracking error

If we wish to resolve the tracking problem for the given nominal trajectory $q^*(t)$, then we can represent the dynamics of the controlled plant **in deviation coordinates** $\delta(t) := q(t) - q^*(t)$ as follows

$$\boxed{\tilde{D}(\delta(t))\ddot{\delta}(t) = \tau(t) + \vartheta(t) - \tilde{C}\left(\delta(t), \dot{\delta}(t)\right)\dot{\delta}(t) - \tilde{G}(\delta(t))} \tag{2}$$

with

$$\boxed{\begin{gathered} \tilde{D}(\delta) := D(\delta + q^*), \\ \tilde{C}\left(\delta, \dot{\delta}\right) := C\left(\delta + q^*, \dot{\delta} + \dot{q}^*\right), \\ \tilde{G}(\delta) := G(\delta + q^*). \end{gathered}}$$

Notice that the deviation dynamics (2) may be represented as (omitting the time-argument)

$$\boxed{\ddot{\delta} = \tilde{D}^{-1}(\delta)\,\tau + \tilde{D}^{-1}(\delta)\,\xi,} \tag{3}$$

or, equivalently, as

$$\boxed{\left.\begin{array}{c} \delta_1 := \delta, \\ \\ \dot{\delta}_1 = \delta_2, \\ \dot{\delta}_2 = \tilde{D}^{-1}(\delta_1)\,\tau + \tilde{D}^{-1}(\delta_1)\,\xi. \end{array}\right\}} \tag{4}$$

Accepted assumptions

A1. The vector of generalized coordinate $q(t)$ and its derivative $\dot{q}(t)$ are measurable on-line during the process.

A2. The matrix $D\left(q\right)$ is supposed to be known and invertible (the usual property of any mechanical system).

A3. The uncertain term

$$\xi\left(t\right) := \vartheta\left(t\right) - \tilde{C}\left(\delta\left(t\right), \dot{\delta}\left(t\right)\right)\dot{\delta}\left(t\right) - \tilde{G}\left(\delta\left(t\right)\right) \tag{5}$$

is admitted to be unknown and unmeasurable, but is bounded as

$$\left\|\xi\left(t\right)\right\| \leq c + c_0\left\|\delta\left(t\right)\right\| + c_1\left\|\dot{\delta}\left(t\right)\right\|, c, c_0, c_1 \geq 0. \tag{6}$$

Accepted assumptions

A4. The loss function $F : R^n \rightarrow R^1$, characterizing the quality of a controlled process, is assumed to be unknown, convex (not obligatory, strongly convex), differentiable for almost all $\delta \in R^n$ (the Radamacher theorem) and its sub-gradient $a(\delta)$ is supposed to be measurable and bounded at any point δ_1, that is,

$$\|a(\delta(t))\| \leq d_g < \infty,$$

and the reaction $a(\delta)$ is available for any argument $\delta \in R^n$.

A5. The minimum of the loss function $F(\delta)$ exists, namely,

$$F^* = \min_{\delta \in \mathbb{R}^n} F(\delta) > -\infty.$$

Problem formulation

Problem

Under the assumptions A1-A3 we need to design a control strategy $\tau(t)$ *as a feedback* $\tau(\delta(\cdot))$, *which provides the* **functional convergence** *of the cost function* $F(\delta(t))$ *to its minimum value* F^*, *in the presence of uncertainties* $\xi(t)$, *that is, to guarantee*

$$\boxed{F(\delta(t)) \underset{t\longrightarrow\infty}{\longrightarrow} \inf_{\delta\in\mathbb{R}^n} F(\delta) = F^*,} \tag{7}$$

supposing that the current **sub-gradient** $a(\delta(t))$ *of the convex function* $F(\delta)$, *to be optimized, is available on-line.*

Examples of loss-functions

The convex (not obligatory strongly) loss function $F : \mathbb{R}^n \to \mathbb{R}^1$ defines the quality of control actions $\{\tau(t)\}_{t\geq 0}$ in the point $\delta(t)$. For example, the following two functions belong to the considered class of the convex loss functions to be optimized:

1. $$F(\delta) = \sum_{i=1}^{n} |\delta_i|, \;\; a_i(\delta) = \operatorname{sign}(\delta_i),$$

2. $$F(\delta) = \sum_{i=1}^{n} |\delta_i|_{\varepsilon}^{+}, \quad |z|_{\varepsilon}^{+} := \begin{cases} z - \varepsilon & \text{if} \quad z \geq \varepsilon \\ -z - \varepsilon & \text{if} \quad z \leq -\varepsilon \\ 0 & \text{if} \quad |z| < \varepsilon \end{cases},$$

$$a_i(\delta_i) = \begin{cases} 1 & \text{if} \quad \delta_i \geq \varepsilon \\ -1 & \text{if} \quad \delta_i \leq -\varepsilon \\ (-1,1) & \text{if} \quad |\delta_i| < \varepsilon \end{cases} = \operatorname{sign}(|\delta| - \varepsilon).$$

In both these examples $F^* = F(0) = 0$.

Define the vector function $s(t) \in \mathbb{R}^n$, which from now on and throughout this lecture will be referred to as *"sliding variable"*:

$$\left.\begin{array}{c} s(t) = \dot{\delta}(t) + \dfrac{\delta(t) + \eta}{t + \theta} + \tilde{G}(t), \ \eta = \mathrm{const} \in \mathbb{R}^n, \\ \\ \tilde{G}(t) := \dfrac{1}{t+\theta} \int\limits_{\tau = t_0}^{t} a(\delta(\tau))\, d\tau, \ \theta > 0, \\ \\ a(\delta_1(\tau)) \in \partial F(\delta_1(\tau)) \end{array}\right\} \quad (8)$$

Here $\delta(t) := q(t) - q^*(t) \in \mathbb{R}^n$, η is a constant vector and $\tilde{G}(t)$ is the averaged subgradient (ASG) of the function $F(\delta(t))$ (7). *Note that the sliding variable $s(t)$ contains the integral term which is physically measurable.*

Desired dynamics

Define the desired ASG dynamics as

$$\boxed{s(t) = \dot{s}(t) = 0, t \geq t_0,} \tag{9}$$

which corresponds exactly to the situation when the sliding variable $s(t)$ is equal to zero for all $t \geq t_0$. Below we will show why the dynamic (9) is called a desired. Since

$$\left.\boxed{\begin{array}{c} (t+\theta)\, s(t) = (t+\theta)\, \dot{\delta}(t) + \delta(t) + \eta = \zeta(t), \\ \dot{\zeta}(t) = -a(\delta(t)), \ \zeta(t_0) = 0, \end{array}}\right\} \tag{10}$$

in the desired regime (9) we have

$$\boxed{\left.\begin{array}{c} (t+\theta)\, \dot{\delta}(t) + \delta(t) + \eta = \zeta(t), \quad t \geq t_0 \geq 0, \\ t_0 \text{ is the moment when the desired dynamics may begin.} \end{array}\right\}} \tag{11}$$

Why this regime is referred to as "desired" ?

Lemma

For the variable $\delta(t)$*, satisfying the ideal dynamics (9), with any* $\theta > 0$ *and* η*, for all* $t \geq t_0 \geq 0$ *the following inequality is guaranteed:*

$$\boxed{F(\delta(t)) - F^* \leq \frac{\Phi(t_0)}{t+\theta} \underset{t\to\infty}{\to} 0,} \tag{12}$$

where

$$\boxed{\Phi(t_0) = \Phi(\delta(t_0), \theta, \eta) := (t_0 + \theta) F(\delta(t_0)) - F^* + \frac{1}{2}\|\delta^* - \eta\|^2,} \tag{13}$$

and

$$\boxed{\begin{array}{c} \delta^* \in \mathrm{Arg}_{\inf} \inf_{\delta \in \mathbb{R}^n} F(\delta) \\ (\delta^* \textit{ may be not unique}). \end{array}} \tag{14}$$

Proof of Lemma on Functional convergence (1)

Proof.

Defining $\mu(t) := t + \theta$ we have

$$\frac{d}{dt}\left[\frac{1}{2}\|\zeta(t)\|^2 - \zeta^{\intercal}(t)\,\delta^*\right] = \dot{\zeta}^{\intercal}(t)\left(\zeta(t) - \delta^*\right) =$$
$$-a^{\intercal}(\delta(t))\left[\mu(t)\,\dot{\delta}(t) + \delta(t) + \eta - \delta^*\right] =$$
$$-a^{\intercal}(\delta(t))\left(\delta(t) - \delta^*\right) - a^{\intercal}(\delta(t))\left(\mu(t)\,\dot{\delta}(t) + \eta\right).$$

Using the inequality $(\delta - \delta^*)^T a(\delta) \geq F(\delta) - F^*$, valid for convex (not obligatory stongly convex) functions in the first term on the right side, and applying the identity $a^T(\delta(t))\,\dot{\delta}(t) = \frac{d}{dt}\left[F(\delta(t)) - F^*\right]$, we get

$$\frac{d}{dt}\left[\frac{1}{2}\|\zeta(t)\|^2 - \zeta^{\intercal}(t)\,\delta^*\right] \leq -\left[F(\delta(t)) - F^*\right]$$
$$-\mu(t)\,\frac{d}{dt}\left[F(\delta(t)) - F^*\right] - a^T(\delta(t))\,\eta.$$

Proof of Lemma on Functional convergence (2)

Proof.

Then, integrating the last inequality in the interval $[t_0, t]$ and applying the formula of integration by parts, we derive

$$\int_{\tau=t_0}^{t} \left[F\left(\delta\left(\tau\right)\right) - F^*\right] d\tau \leq \frac{1}{2}\left(\left\|\zeta\left(t_0\right)\right\|^2 - \left\|\zeta\left(t\right)\right\|^2\right) + \left(\zeta\left(t\right) - \zeta\left(t_0\right)\right)^T \delta^* - \left(\mu\left(t\right)\left[F\left(\delta\left(t\right)\right) - F^*\right]\right)_{t_0}^{t} + \int_{\tau=t_0}^{t} \left[F\left(\delta\left(\tau\right)\right) - F^*\right] \dot{\mu}\left(\tau\right) d\tau - \left[\int_{\tau=t_0}^{t} a^{\intercal}\left(\delta\left(\tau\right)\right) d\tau\right] \eta.$$

□

Proof of Lemma on Functional convergence (3)

Proof.

Since $\dot{\mu}_\tau = 1$, the above inequality becomes

$$\left.\begin{array}{c}
\mu(t)\left[F(\delta(t)) - F^*\right] \le \mu(t_0)\left[F(\delta(t_0)) - F^*\right] + \\
\frac{1}{2}\left(\|\zeta(t_0)\|^2 - \|\zeta(t)\|^2\right) + \left(\zeta(t) - \zeta(t_0)\right)^\intercal \delta^* + \zeta^\intercal(t)\,\eta = \\
(t_0 + \theta)\left[F(\delta(t_0)) - F^*\right] + \left(\frac{1}{2}\|\zeta(t_0)\|^2 - \zeta^\intercal(t_0)\,\delta^*\right) + \\
\frac{1}{2}\|\delta^* - \eta\|^2 - \frac{1}{2}\underbrace{\left[\|\zeta(t)\|^2 - 2\,\zeta^\intercal(t)(\delta^* - \eta) + \|\delta^* - \eta\|^2\right]}_{\|\zeta(t) - (\delta^* - \eta)\|^2} \\
\le (t_0 + \theta)\left[F(\delta(t_0)) - F^*\right] - \frac{1}{2}\|\zeta(t) - (\delta^* - \eta)\|^2 + \\
\left(\frac{1}{2}\|\zeta(t_0)\|^2 - \zeta^\intercal(t_0)\,\delta^*\right) + \frac{1}{2}\|\delta^* - \eta\|^2 \le \Phi_{t_0},
\end{array}\right\} \quad (15)$$

from which we obtain (13). Lemma is proved. □

Important comment (1)

Remark

The parameter η will be chosen below in such a way that the desired optimization regime starts from the beginning of the process, namely, when, $t_0 = 0$.

Important comment (2)

Corollary

In the partial case when

$$\delta^* = 0, \;\; t_0 = 0 \text{ and } F^* = 0$$

the formula (13) becomes

$$\boxed{\Phi(t_0) = \Phi(\delta(t_0), \theta, \eta) := \theta F(\delta(0)) + \frac{1}{2}\|\eta\|^2 .} \qquad (16)$$

Main theorem on ASG robust controller

Theorem

Under assumptions 1-5 the ISM robust controller

$$\left.\begin{array}{c} \tau(t) = \tilde{D}(\delta(t))\left[-k_t \mathrm{SIGN}(s(t)) + u_{comp}(t)\right], \\ u_{comp}(t) = -p_t^{reali}, \\ k_t = \left\|\tilde{D}^{-1}(\delta(t))\right\|\left(c + c_0\|\delta(t)\| + c_1\left\|\dot{\delta}(t)\right\|\right) + \rho_0, \ \rho_0 > 0, \end{array}\right\} \quad (17)$$

where

$$p_t^{reali} := \frac{1}{t+\theta}\left(\dot{\delta}(t) - \frac{\delta(t)+\eta}{t+\theta} - \tilde{G}(t) + a(\delta(t))\right) \quad (18)$$

with

$$\eta = -\theta\delta_{2,0} - \delta_{1,0} \quad (19)$$

guarantees the functional convergence (12) from $t_0 = 0$.

Proof of Main Theorem (1)

Proof.

In view of the assumption A2 we have

$$\left.\begin{array}{c}\delta\left(t\right):=q\left(t\right)-q^{*}\left(t\right),\ \dot{\delta}\left(t\right)=\dot{q}\left(t\right)-\dot{q}^{*}\left(t\right),\\ \ddot{\delta}\left(t\right)=\tilde{D}^{-1}\left(\delta\left(t\right)\right)\tau\left(t\right)+\tilde{D}^{-1}\left(\delta\left(t\right)\right)\xi\left(t\right).\end{array}\right\}$$

For the Lyapunov function $V\left(s\right)=\frac{1}{2}s^{\intercal}s$ we have

$$\left.\begin{array}{c}\dot{V}\left(s\left(t\right)\right)=s^{\intercal}\left(t\right)\dot{s}\left(t\right)=\\ s^{\intercal}\left(t\right)\left(\ddot{\delta}\left(t\right)+\frac{\dot{\delta}(t)}{t+\theta}-\frac{\delta(t)+\eta}{(t+\theta)^{2}}-\frac{1}{t+\theta}\tilde{G}\left(t\right)+\frac{1}{t+\theta}a\left(\delta\left(t\right)\right)\right)=\\ s^{\intercal}\left(t\right)\left(\tilde{D}^{-1}\left(\delta\left(t\right)\right)\tau\left(t\right)+\tilde{D}^{-1}\left(\delta\left(t\right)\right)\xi\left(t\right)\right)+\\ s^{\intercal}\left(t\right)\left(\frac{1}{t+\theta}\left(\dot{\delta}\left(t\right)-\frac{\delta(t)+\eta}{t+\theta}-\tilde{G}\left(t\right)+a\left(\delta\left(t\right)\right)\right)\right)=\\ s^{\intercal}\left(t\right)p_{t}^{\text{reali}}+s^{\intercal}\left(t\right)\tilde{D}^{-1}\left(\delta\left(t\right)\right)\tau\left(t\right)+s^{\intercal}\left(t\right)\tilde{D}^{-1}\left(\delta\left(t\right)\right)\xi\left(t\right).\end{array}\right\}\quad(20)$$

□

Proof of Main Theorem (2)

Proof.

Selecting τ as in (17) for the second term in (20) we get

$$\left.\begin{aligned} \dot{V}(s_t) &= -k_t s^{\intercal}(t)\,\mathrm{SIGN}(s(t)) + s^{\intercal}(t)\,\tilde{D}^{-1}(\delta(t))\,\xi(t) \\ &\leq -k_t \sum_{i=1}^{n} |s_i(t)| + \|s(t)\| \left\|\tilde{D}^{-1}(\delta(t))\right\| \|\xi(t)\| \end{aligned}\right\} \quad (21)$$

□

Proof of Main Theorem (3)

Proof.

Taking into account that $\sum_{i=1}^{n} |s_i(t)| \geq \|s(t)\|$ and, in view of (6) and (21), we derive

$$\dot{V}(s(t)) \leq -k_t \|s(t)\| + \|s(t)\| \left\|\tilde{D}^{-1}(\delta(t))\right\| \left(c + c_0 \|\delta(t)\| + c_1 \left\|\dot{\delta}(t)\right\|\right)$$
$$= -\rho_0 \|s(t)\| = -\sqrt{2}\rho_0 \sqrt{V(s(t))},$$

implying $2\left(\sqrt{V(s(t))} - \sqrt{V(s(t_0))}\right) \leq -\sqrt{2}\rho_0 t$ and

$$0 \leq \sqrt{V(s(t))} \leq \sqrt{V(s(t_0))} - \frac{\rho_0}{\sqrt{2}} t,$$

which leads to the conclusion that for all $t \geq t_{reach} := \frac{1}{\rho_0}\sqrt{2V(s_{t_0})} = \frac{\|s_{t_0}\|}{\rho_0}$ we have that $V(s(t)) = 0$ and $s(t) = 0$. □

Proof of Main Theorem (4)

Proof.

To make the reaching time $t_{reach} = 0$ it is sufficient to gurantee that $s_{t_0=0} = 0$. But since by (10)

$$(t+\theta)\, s(t) = (t+\theta)\, \dot{\delta}(t) + \delta(t) + \eta = \zeta(t),$$
$$(t_0+\theta)\, s(t_0) = (t_0+\theta)\, \dot{\delta}(t_0) + \delta(t_0) + \eta = \zeta(t_0)$$
$$s_{t_0} = \dot{\delta}_{t_0} + \frac{\delta_{t_0} + \eta}{t_0 + \theta},$$

we need to fulfill the condition $s_{t_0=0} = \dot{\delta}_{t_0=0} + \frac{\delta_{t_0=0} + \eta}{\theta} = 0$, which is possible if take η as in (19), providing

$$t_{reach} = \frac{\| s_{t_0=0} \|}{\rho_0} = 0.$$

Theorem is proven. □