AMEFSS 2021

The Game Theoretic Approach of a Stochastic Linear Quadratic Tracking Problem

Vasile Drăgan
Institute of Mathematics "Simion Stoilow" of the Romanian Academy,
P.O.Box 1-764, RO-014700, Bucharest, Romania
and the Academy of the Romanian Scientists
e-mail: *Vasile.Dragan@imar.ro*
Ivan G. Ivanov
Faculty of Economics and Business Administration,
Sofia University "St. Kl. Ohridski",
Sofia 1113, Bulgaria
e-mail: *i_ivanov@feb.uni-sofia.bg*
Ioan-Lucian Popa
Department of Computing, Mathematics and Electronics,
"1 Decembrie 1918" University of Alba Iulia, Alba Iulia, 510009, Romania
e-mail: *lucian.popa@uab.ro*

June 5, 2021

1 The problem

Consider the controlled system having the state space representation described by:

$$\begin{aligned}
&dx(t) = [A_0(\eta_t)x(t) + B_1(\eta_t)u_1(t) + B_2(\eta_t)u_2(t)]dt + A_1(\eta_t)x(t)dw(t),\\
&x(t_0) = x_0, && \text{(1a)}\\
&z_k(t) = C_k(\eta_t)x(t), \quad k = 1, 2, && \text{(1b)}
\end{aligned}$$

$t \in [t_0, t_f] \subset [0, \infty)$, where $x(t) \in \mathbb{R}^n$ is the state vector at instance time t and $u_k(t) \in \mathbb{R}^{m_k}$, $k = 1, 2$, are the vectors of control parameters. In (1), $\{w(t)\}_{t\geq 0}$, is a 1-dimensional standard Wiener process defined on a given probability space $(\Omega, \mathcal{F}, \mathcal{P})$ and $\{\eta_t\}_{t\geq 0}$ is a standard right continuous Markov process taking values in a finite set $\mathcal{N} = \{1, 2, \dots, N\}$ and having the transition semigroup $\mathbb{P}(t) = e^{Qt}$, $t \geq 0$. The elements q_{ij} of the generator matrix $Q \in \mathbb{R}^{N\times N}$

satisfy

$$q_{ij} \geq 0 \ \text{ if } \ i \neq j, \tag{2a}$$

$$\sum_{l=1}^{N} q_{il} = 0, \tag{2b}$$

for all $(i,j) \in \mathcal{N} \times \mathcal{N}$. Throughout the paper, we assume that $\{w(t)\}_{t\geq 0}$ and $\{\eta_t\}_{t\geq 0}$ are independent stochastic processes. Assume that $A_k(i) \in \mathbb{R}^{n\times n}$, $k = 0,1$, $(B_j(i),\ C_j(i)) \in \mathbb{R}^{n\times m_j} \times \mathbb{R}^{n_{z_j}\times n}$, $j = 1,2$ are given matrices.

The set $\mathcal{U}_k$ of the admissible controls available to the decision makers $\mathcal{P}_k$, $k = 1,2$, consists of the stochastic processes in an affine state feedback form, i.e.,

$$u_k(t) = F_k(t,\eta_t)x(t) + \varphi_k(t,\eta_t) \tag{3}$$

where $t \to F_k(t,i) : [t_0,t_f] \to \mathbb{R}^{m_k\times n}$ and $t \to \varphi_k(t,i) : [t_0,t_f] \to \mathbb{R}^n$ are arbitrary continuous functions. Roughly speaking, the aim of the decision maker $\mathcal{P}_k$ is to find a control law (or an admissible strategy) of type (3), $\tilde{u}_k(\cdot) \in \mathcal{U}_k$ which minimizes the deviation of the signal $z_k(\cdot)$ from a given reference signal $r_k(\cdot)$, when the other decision maker $\mathcal{P}_l$ wants to minimize the deviation of the signal $z_l(\cdot)$, $(l \neq k)$ from another reference signal $r_l(\cdot)$.

Since this problem is an optimization problem with two objective function, its solution can be viewed as an equilibrium strategy for a non-cooperative differential game with two players. For a rigourous mathematical setting of this optimization problem, we introduce the following cost functions:

$$\begin{aligned} J_k(x_0;u_1(\cdot),u_2(\cdot)) =& \mathbb{E}[(z_k(t_f)-\zeta_k)^T G_k(\eta_{t_f})(z_k(t_f)-\zeta_k)] \\ &+ \mathbb{E}[\int_{t_0}^{t_f} \{(z_k(t)-r_k(t))^T M_k(\eta_t)(z_k(t)-r_k(t)) \\ &+ u_1^T(t)R_{k1}(\eta_t)u_1(t) + u_2^T(t)R_{k2}(\eta_t)u_2(t)\}dt], \quad k=1,2. \end{aligned} \tag{4}$$

Here $r_k(\cdot) : [t_0,t_f] \to \mathbb{R}^{n_{z_k}}$, $k = 1,2$ is the reference which have to be tracked by the signal $z_k(\cdot)$, and $\zeta_k \in \mathbb{R}^{n_{z_k}}$, $k = 1,2$ is the target of the final value $z_k(t_f)$. Throughout this work $\mathbb{E}[\cdot]$ stands for the mathematical expectation. The weights matrices involved in (4) are satisfying the assumption:

H1)

(a) $(M_k(i),\ R_{k1}(i),\ R_{k2}(i)) \in \mathcal{S}_{n_{z_k}} \times \mathcal{S}_{m_1} \times \mathcal{S}_{m_2}$, $i \in \mathcal{N}$ are given matrices;

(b) $M_k(i) \geq 0$, $R_{kl}(i) \geq 0$, $R_{kk}(i) > 0$, for all $k,l = 1,2$, $l \neq k$, $G_k(i) \geq 0$, $i \in \mathcal{N}$.

Here and in the sequel, $\mathcal{S}_p \subset \mathbb{R}^{p\times p}$ denotes the subspace of symmetric matrices of size $p \times p$, $p \geq 1$.

Definition 1.1 *We say that the pair of admissible strategies* $(\tilde{u}_1(\cdot),\tilde{u}_2(\cdot)) \in \mathcal{U}_1 \times \mathcal{U}_2$ *achieves a Nash equilibrium for the differential game described by the controlled system (1), the performance criteria (4) and the admissible strategies of type (3) if*

$$J_1(x_0;\tilde{u}_1(\cdot),\tilde{u}_2(\cdot)) \leq J_1(x_0;u_1(\cdot),\tilde{u}_2(\cdot)), \ \ \textit{for all} \ \ u_1 \in \mathcal{U}_1 \tag{5}$$

and

$$J_2(x_0; \tilde{u}_1(\cdot), \tilde{u}_2(\cdot)) \le J_2(x_0; \tilde{u}_1(\cdot), u_2(\cdot)), \quad \textit{for all} \;\; u_2 \in \mathcal{U}_2. \tag{6}$$

In the next section we shall derive explicit formulae of a Nash equilibrium strategy $(\tilde{u}_1(\cdot), \tilde{u}_2(\cdot))$ for the LQ differential game described by (1), (3) and (4).

Remark 1.1 We shall see that for the computation of the gain matrices of a Nash equilibrium strategy we need to know a priori the whole reference signal $r_k(\cdot)$. This may done difficult the implementation of the solution of a tracking problem on long intervals of time, because it requires big memories. That is way, in applications the signal which must be tracked need to be as simpler as possible, to be easy to memorize.

2 The main results

In this section we derive explicit formulae for the Nash equilibrium strategy $(\tilde{u}_1(\cdot), u_2(\cdot))$. To this end we consider the following terminal value problems (TVPs)

$$\begin{aligned}
&\dot{X}_1(t,i) + (A_0(i) - S_2(i)X_2(t,i))^T X_1(i) + X_1(t,i)(A_0(i) - S_2(i)X_2(t,i)) \\
&\;+ A_1^T(i)X_1(t,i)A_1(i) - X_1(t,i)S_1(i)X_1(t,i) + X_2(t,i)S_{12}(i)X_2(t,i) \\
&\;+ \sum_{j=1}^{N} q_{ij}X_1(t,j) + C_1^T(i)M_1(i)C_1(i) = 0 &&\text{(7a)}\\
&\dot{X}_2(t,i) + (A_0(i) - S_1(i)X_1(t,i))^T X_2(t,i) + X_2(t,i)(A_0(i) - S_1(i)X_1(t,i)) \\
&\;+ A_1^T(i)X_2(t,i)A_1(i) - X_2(t,i)S_2(i)X_2(t,i) + X_1(t,i)S_{21}(i)X_1(t,i) \\
&\;+ \sum_{j=1}^{N} q_{ij}X_2(t,j) + C_2^T(i)M_2(i)C_2(i) = 0 &&\text{(7b)}\\
&X_k(t_f,i) = C_k^T(i)G_k(i)C_k(i) &&\text{(7c)}\\
&F_k(t,i) = -R_{kk}^{-1}(i)B_k^T(i)X_k(t,i) &&\text{(7d)}
\end{aligned}$$

$i \in \mathcal{N}$, $k = 1, 2$,

$$\begin{aligned}
&\dot{\Psi}_1(t,i) + (A_0(i) - S_1(i)X_1(t,i) - S_2(i)X_2(t,i))^T \Psi_1(t,i) \\
&\;+ \sum_{j=1}^{N} q_{ij}\Psi_1(t,j) + (X_1(t,i)S_2(i) - X_2(t,i)S_{12}(i))\Psi_2(t,i) + C_1^T(i)M_1(i)r_1(t) = 0 &&\text{(8a)}\\
&\dot{\Psi}_2(t,i) + (X_2(t,i)S_1(i) - X_1(t,i)S_{21}(i))\Psi_1(t,i) + (A_0(i) - S_1(i)X_1(t,i) \\
&\;- S_2(i)X_2(t,i))^T\Psi_2(t,i) + \sum_{j=1}^{N} q_{ij}\Psi_2(t,j) + C_2^T(i)M_2(i)r_2(t) = 0 &&\text{(8b)}\\
&\Psi_k(t_f,i) = C_k^T(i)G_k(i)\zeta_k &&\text{(8c)}\\
&\varphi_k(t,i) = -R_{kk}^{-1}(i)B_k^T(i)\Psi_k(t,i) &&\text{(8d)}
\end{aligned}$$

$i \in \mathcal{N}$, $k = 1, 2$. In (7) and (8) we have denoted

$$S_j(i) \triangleq B_j(i)R_{jj}^{-1}(i)B_j^T(i)$$

$$S_{jk}(i) \stackrel{\triangle}{=} B_k(i)R_{kk}^{-1}(i)R_{jk}(i)R_{kk}^{-1}(i)B_k^T(i)$$

$j = 1, 2,\ k = 3 - j$. Also we associate the following TVP:

$$\dot{\mu}_k(t,i) + \sum_{j=1}^{N} q_{ij}\mu_k(t,j) + \tilde{h}_k(t,i) = 0 \tag{9a}$$

$$\mu_k(t_f,i) = \zeta_k^T G_k(i)\zeta_k \tag{9b}$$

$$\begin{aligned}\tilde{h}_1(t,i) &= r_1^T(t)M_1(i)r_1(t) + \tilde{\Psi}_2^T(t,i)S_{12}(i)\tilde{\Psi}_2(t,i)\\ &\quad - 2\tilde{\Psi}_1^T(t,i)S_2(i)\tilde{\Psi}_2(t,i) - \tilde{\Psi}_1^T(t,i)S_1(i)\tilde{\Psi}_1(t,i)\end{aligned} \tag{9c}$$

$$\begin{aligned}\tilde{h}_2(t,i) &= r_2^T(t)M_2(i)r_2(t) + \tilde{\Psi}_1^T(t,i)S_{21}(i)\tilde{\Psi}_1(t,i)\\ &\quad - 2\tilde{\Psi}_2^T(t,i)S_1(i)\tilde{\Psi}_1(t,i) - \tilde{\Psi}_2^T(t,i)S_2(i)\tilde{\Psi}_2(t,i).\end{aligned} \tag{9d}$$

In (9), $(\tilde{\Psi}_1(\cdot,i), \tilde{\Psi}_2(\cdot,i))$, $i \in \mathcal{N}$ is the solution of the TVP (8).

One obtains:

Theorem 2.1 *Assume:*

(a) the assumption ***H1)*** *is fulfilled;*

(b) the solution $(\tilde{\tilde{X}}_1(\cdot,i), \tilde{\tilde{X}}_2(\cdot,i))$, $i \in \mathcal{N}$, *of the TVP (7a)-(7c) is defined on the whole interval* $[t_0, t_f]$.

We set

$$\tilde{\tilde{u}}_j(t) = -R_{jj}^{-1}(\eta_t)B_j^T(\eta_t)(\tilde{\tilde{X}}_j(t,\eta_t)\tilde{\tilde{x}}(t) + \tilde{\tilde{\Psi}}_j(t,\eta_t)), \tag{10}$$

$\tilde{\tilde{x}}(\cdot)$ *being the solution of the IVP obtained substituting (10) in (1). Under these conditions* $(\tilde{\tilde{u}}_1(\cdot), \tilde{\tilde{u}}_2(\cdot))$ *is an equilibrium strategy for the differential game described by the controlled system (1), the performance criteria (4) and the admissible strategies of type (3). The optimal values of the performance criteria (4) are given by*

$$J_k(x_0; \tilde{\tilde{u}}_1(\cdot), \tilde{\tilde{u}}_2(\cdot)) = x_0^T\mathbb{E}[\tilde{\tilde{X}}_k(t_0,\eta_{t_0})]x_0 - 2x_0^T\mathbb{E}[\tilde{\tilde{\Psi}}_k(t_o,\eta_{t_0})] + \mathbb{E}[\tilde{\tilde{\mu}}_k(t_0,\eta_{t_0})],$$

$k = 1, 2$, $\tilde{\tilde{\Psi}}_k(\cdot,i)$, $\tilde{\tilde{\mu}}_k(\cdot,i)$ *being the solutions of the TVPs (8) and (9), respectively.*

3 Numerical experiments

We consider a numerical example for the stochastic system:

$$\begin{aligned}&dx(t) = [A_0x(t) + B_1u_1(t) + B_2u_2(t)]dt + A_1x(t)dw(t),\\ &x(t_0) = x_0,\\ &z_k(t) = C_kx(t), \quad k = 1, 2\end{aligned}$$

with coefficient matrices:

A_0=[1 0 3 0; -4 2 0 -10; -14 8.5 -2.5 0; 0 -2 0 -10];, $A_0 \in \mathbb{R}^{4\times4}$

A_1=[0 2 0. -1; 0 0 -3 1.5; -1.45 0.6 -2 0; 0 -3 0 5]; $A_1 \in \mathbb{R}^{4\times4}$

$B_1 = [1\ 0; 2\ 5; -1\ 4; 2\ 6]; \quad B_2 = [0\ 1; -1.5\ \ -3; 2\ \ -4; 2\ 0];, B_1, B_2 \in \mathbb{R}^{4\times 2}$.
$C_1 = [1.0\ \ -0.25\ 0.75\ \ -0.5; 0.25\ 0.75\ \ -0.5\ \ -0.25]$;
$C_2 = [1.2\ \ -0.5\ 0.75\ \ -0.045; 0.5\ 0.15\ 0.35\ 0.55];\ ,\ C_1, C_2 \in \mathbb{R}^{4\times 4}$.

The weight matrices for the performance criteria are of the form:

$G_1 = [0.2778\ -0.094; -0.094\ \ 0.166];,\ G_2 = [0.2778\ \ -0.133; -0.133\ \ 0.31];,\ G_1, G_2 \in \mathbb{R}^{2\times 2}$.
$M_1 = [3.2\ \ 0.5; 0.5\ \ 2.5]; \quad M2 = [0.75\ \ 0.05; 0.05\ \ 0.95];,\ M_1, M_2 \in \mathbb{R}^{2\times 2}$.
$R_{11} = [0.8\ \ 0.3; 0.3\ \ 1.5];, \quad R_{22} = [0.95\ \ 0.65; 0.6\ \ 1.25];, \quad R_{12} = [0.6\ \ \ -0.3; -0.3\ \ 1.2];,$
$R_{21} = [0.8\ \ \ -0.2; \quad -0.21.0]; \ \ R_{ij} \in \mathbb{R}^{2\times 2}$.
Moreover, $r_k(t) = 0, k = 1, 2$ and the targets are $\zeta_1 = [0.3; 0.8]\,,\ \zeta_2 = [0.6; 0.9]$, and $t \in [0, 1]$.
For computational execution we use the Euler discretization method as follow:

$$\tilde{X}_k(jh) = \tilde{X}_k((j+1)h) + h\mathcal{R}_k((j+1)h, \tilde{X}_1((j+1)h), \tilde{X}_2((j+1)h)\,)$$

with $k = 1, 2, j = N-1, ..., 1, 0$ and $\tilde{X}_k(Nh) = C_k^T G_k C_k\,, k = 1, 2\,, t \in [0, 1]\,, h = 0.1(N = 10)$; $\mathcal{R}_k(.,.,.)$ is a Riccati operator defined by (7). Finally , we obtain $J_1(x_0; \tilde{\tilde{u}}_1(\cdot), \tilde{\tilde{u}}_2(\cdot)) = 0.4393\,, J_2(x_0; \tilde{\tilde{u}}_1(\cdot), \tilde{\tilde{u}}_2(\cdot)) = 1.3199$ for $x_0 = [0.4; 0.01; 0.2; 0.25]\,$.

References

[1] V. Dragan, I. G. Ivanov,I.-L. Popa, A Game Theoretic Setting of a Stochastic Linear Quadratic Tracking Problem, under preparation, 2021.