THE LYAPUNOV functions method and applications
P. Borne and V. Matrosov (editors)
J.C. Baltzer AG, Scientific Publishing Co. © IMACS, 1990
pp. 117-120

APPLICATION OF THE LYAPUNOV FUNCTION METHOD TO A LINEAR DIFFERENTIAL GAME WITH INTEGRAL–TERMINAL PAYOFF

V.L. TUROVA
Institute of Mathematics and Mechanics, Ural Branch of Academy of Sciences of the USSR
S. Kovalevskaya 16, 620219 Sverdlovsk, USSR

This paper describes a closed-loop ε-optimal strategy, constructed with the help of some Lyapunov function. As an example, the problem of transfering of two-link pendulum to the upper unstable equilibrium state is presented.

1. INTRODUCTION

The linear differential game under consideration is of the form

$$\dot{x}=Ax+Bu+Cv,\quad t\in[t_o,\theta],\, x\in R^n,\, u\in R^p,\, v\in R^q \quad (1)$$

Here x is the phase vector, u is a control parameter of the first player, v - of the second one, A, B, C are the constant matrices, t_o,θ are fixed time moments. No a priori restrictions on the control parameters u and v are imposed. The payoff functional (quality index of process $\{x(\cdot),u(\cdot),v(\cdot)\}$) is

$$\gamma(x(\cdot),u(\cdot),v(\cdot))= \| x(\theta) \| + \int_{t_o}^{\theta} [\langle\Phi(t)u(t),u(t)\rangle-\langle\Psi(t)v(t),v(t)\rangle]dt \quad (2)$$

Here $\langle\Phi(t)u,u\rangle$, $\langle\Psi(t)v,v\rangle$ are positive-definite quadratic forms, $t_*\in[t_o,\theta]$ is an initial moment of control process. First player is interested to minimize the value of γ , the aim of the second player is opposite.

Note, that in contrust with linear-quadratic game, whose solution is defined for some perhaps small time interval $[t,\theta]$, the differential game (1)-(2) has a value for any position $(t_*,x_*)\in[t_o,\theta]\times R^n$. In papers [1,2] the analitical representation of value function and formulas for optimal strategies of the players were obtained using the stochastic programming maxmin method. But for the practical use strategy $u^o(t,x,\varepsilon)$ from [1,2] is not always convenient, because the value of the game exponentially depends on the duration of the game. In the present paper we describe a modernized strategy $U^o(t,x,\varepsilon)$, constructed with the help of Lyapunov function.

The method of constructing of optimal guaranteeing strategy based on the aiming to the motion of some auxiliary model is well known in the theory of differential games [3]. The model is closely connected with the optimal guaranteed result function (value of game) and is being choosen so that the value function does not increase along the model motion. Then the optimal strategy is constructed on the principle of maximal displacement of original system phase vector in the direction of disagreement vector of the auxiliary and original systems motion. In this way the strategy $u^o(t,x,\varepsilon)$ from [1,2] is formed. Since the evaluation of disagreement vector magnitude depends on duration of the game exponentially, the same is true for the difference of the result guaranteed by strategy $u^o(t,x,\varepsilon)$ and the value of the game. The method of constructing of strategy $U^o(t,x,\varepsilon)$ provides the stabilization of disagreement magnitude of initial and auxiliary systems motions.

2. THE MODERNIZED STRATEGY

It will be convenient to increase the dimension of the phase vector $x(t)$ from n to $n+1$. The original differential game is equivalent to the differential game with dynamics

$$\begin{aligned} \dot{x} &= Ax + Bu + Cv \\ \dot{x}_{n+1} &= \langle\Phi(t)u,u\rangle - \langle\Psi(t)v,v\rangle \end{aligned} \quad (3)$$

and terminal payoff function

$$\sigma(x,x_{n+1}) = \|x\| + x_{n+1}$$

Value function $\varrho(t,x,x_{n+1})$ of game (3) is connected with the value function $\rho(t,x)$ of game (1)-(2) as follows:

$$\varrho(t,x,x_{n+1}) = \rho(t,x) + x_{n+1}$$

The structure of model motion equations is identical with that of the equation (3)

$$\dot{y} = Ay + B\bar{u} + Cv$$
$$\dot{y}_{n+1} = \langle\Phi(t)\bar{u},\bar{u}\rangle - \langle\Psi(t)v,v\rangle \qquad (4)$$

and differs only in the way of forming of first player control, which is selected according to the condition of nonincreasing of function ϱ along the motion of system (4). Let us introduce the disagreement vector for the motion of systems (3),(4):

$$(s(t),s_{n+1}(t))=(x(t)-y(t),x_{n+1}(t)-y_{n+1}(t))$$

The dynamics of disagreement vector is described by equations

$$\dot{s} = As + B(u-\bar{u})$$
$$\dot{s}_{n+1} = \langle\Phi(t)u,u\rangle - \langle\Phi(t)\bar{u},\bar{u}\rangle \qquad (5)$$

The modernized strategy $U^o(t,x,\varepsilon)$ consists of two terms : the vector $\tilde{u}(t,x,\varepsilon)$ and the stabilization component $u_{st}(t) = Ps(t)$. Matrix P is selected so that system (5) is asymptotically stable (we assume that the system $\dot{s} = As + Bu$ is stabilizing). Since the system (5) becomes asymptotically stable, we can find Lyapunov function $\lambda(z,z_{n+1}) = \langle\Lambda z,z\rangle + z_{n+1}^2$, whose derivative along the system $\dot{z}= (A+BP)z$, $\dot{z}_{n+1}= 0$ trajectory will be negatively definite. Vector $\tilde{u}(t,x,\varepsilon)$ is defined by condition of maximal displacement of system (5) phase vector in the direction of function λ antigradient vector, that provides nonincreasing of λ along the system (5) motion and thus nondivergence of system (3),(4) motions.

After the above given informal description of the strategy $U^o(t,x,\varepsilon)$ we give the exact definition for the case of discrete scheme of control. Let us define matrices P ,Λ the above way, fix $\varepsilon > 0$ and consider an equistep partition t_i of the interval $[t_*,\theta]$ with the step Δ . Suppose that at the moment t_i the game position is $(t_i,x(t_i),x_{n+1}(t_i))$. Denote $S_i^\varepsilon=$

$$\{(w,w_{n+1}):\langle\Lambda(x(t_i)-w),x(t_i)-w\rangle+w_{n+1}^2\le\varepsilon^2\}$$

Consider the motion of system (4) emanating from the state

$$(w^o(t_i),w^o_{n+1}(t_i))=$$
$$=(w^o(t_i,x(t_i),\varepsilon),w^o_{n+1}(t_i,x(t_i),\varepsilon)) ,$$

satisfying the relation

$$\varrho(t_i,w^o(t_i),w^o_{n+1}(t_i))= \min_{(w,w_{n+1})\in S_i^\varepsilon} \varrho(t_i,w,w_{n+1})$$

Let us find vector $\tilde{u}(t_i,x(t_i),\varepsilon)$ by means of condition

$$\min_u[\langle\Lambda(x(t_i)-w^o(t_i)),Bu\rangle+ w^o_{n+1}(t_i)\langle\Phi(t_i)u,u\rangle]=$$
$$=\langle\Lambda(x(t_i)-w^o(t_i)),B\tilde{u}(t_i,x(t_i),\varepsilon)\rangle+$$
$$+w^o_{n+1}(t_i)\langle\Phi(t_i)\tilde{u}(t_i,x(t_i),\varepsilon),\tilde{u}(\ldots)\rangle$$

We get

$$u(t_i,x(t_i),\varepsilon)= -\frac{1}{2}\,\Phi^{-1}(t_i)B'\Lambda(x(t_i)-$$
$$-w^o(t_i))/w^o_{n+1}(t_i)$$

Since ϱ is a sum of the form $\rho+x_{n+1}$, the control does not depend on x_{n+1}. Formula for $\tilde{u}(t_i,x(t_i),\varepsilon)$ is analogous to that for $u^o(t,x,\varepsilon)$ from [1,2]. Put

$$U^o(t_i,x(t_i),\varepsilon)=\tilde{u}(t_i,x(t_i),\varepsilon) + u_{st}(t_i),$$

where

$$u_{st}(t_i)= P(x(t_i)-w^o(t_i))$$

Since P is the constant matrix, and

the following evaluation is fulfilled

$$\| x(t_i)-w^o(t_i) \| \le K(\varepsilon^2+ d(\Delta))^{1/2},$$

where K is a constant and $d(\Delta)\to 0$ for $\Delta\to 0$, we can make the stabilization component magnitude be of order ε by means of decreasing of discrete scheme step Δ .

The following theorem is valid. Let $\Gamma(t,x,U^o(\cdot,\cdot,\varepsilon))$ be a result guaranteed by strategy U^o in position (t,x). Then for any $\varepsilon > 0$

$$\Gamma(t,x,U^o(\cdot,\cdot,\varepsilon)) \le \rho(t,x) + \beta\varepsilon \left[1+\int_t^\theta \|\Phi(\tau)\| d\tau \right]$$

Here $\beta > 0$ is a constant.

That is, the difference between the result guaranteed by strategy U^o and the value of game depends linearly on duration of the game.

The high speed algorithm of calculation of strategy U^o values is presented in [4]. The strategy U^o is good for unstable systems. Below we consider an example which is solved using the standard programme elaborated on the basis of this algorithm.

3. EXAMPLE.

It is requied to transfer the two-link pendulum to the upper unstable equilibrium state (fig. 1).

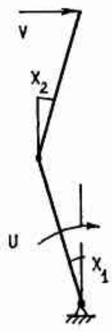

Figure 1

The control parameter u is a torque applied to the lower link. The upper link is subjected to the effect of the force disturbunce v (the wind). The pendulum motion is described by the following equations provided the pendulum rods have the same mass equal 1 and the same length equal 10 :

$$\dot{x}_1 = x_3$$

$$\dot{x}_2 = x_4$$

$$\dot{x}_3 = 2.523\ x_1 - 1.261\ x_2 + 0.0171u - 0.086v$$

$$x_4 = -3.784\ x_1 + 3.363\ x_2 - 0.0257u + 0.429v$$

(we assume that oscilations are small).

The quality index is of the form

$$\int_{t_*}^{\theta} (10^{-4}u^2(\tau) - 10^2 v^2(\tau)) d\tau + \|x(\theta)\|$$

The simulation results of pendulum phase trajectories for $\theta = 3$, $t_* = 0$, $x_* = (0.2 , -0.2 , 0 , 0)$ are depicted on fig.2(a,b).

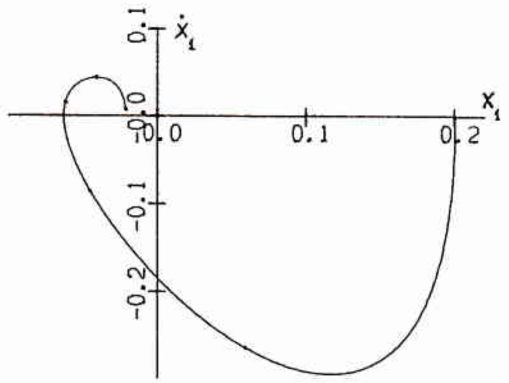

Figure 2a

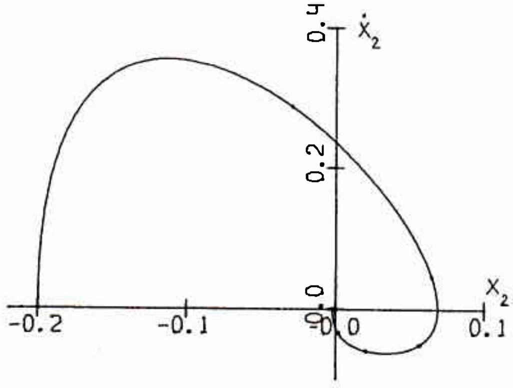

Figure 2b

The first player uses the strategy $U^{o}(t,x,\varepsilon)$ with $\varepsilon=0.01$. The strategy of extremal aiming in direction of the value function gradient is used by the second player. Realisations of first and second controls are presented on fig.2 (c-d).

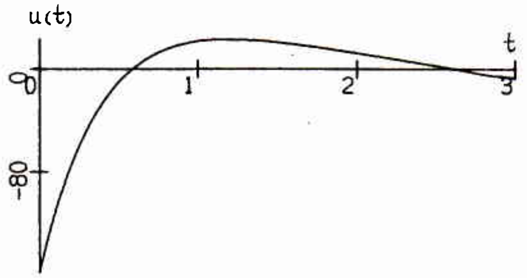

Figure 2c

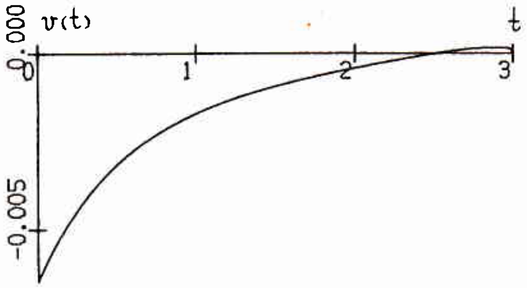

Figure 2d

Fig.3 demonstrates the corresponding results provided the strategy $u^{o}(t,x,\varepsilon)$ from [1,2] is applied.

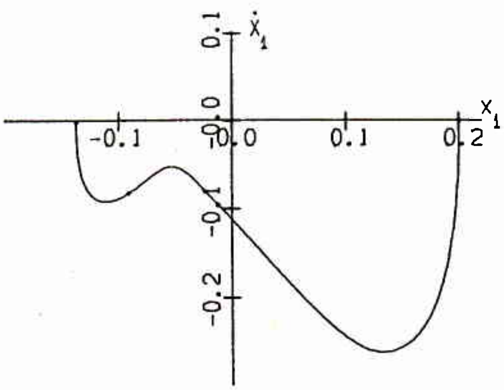

Figure 3a

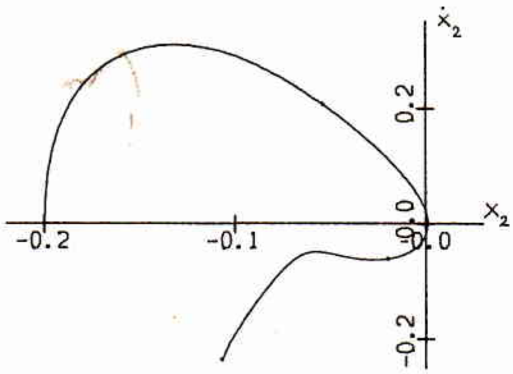

Figure 3b

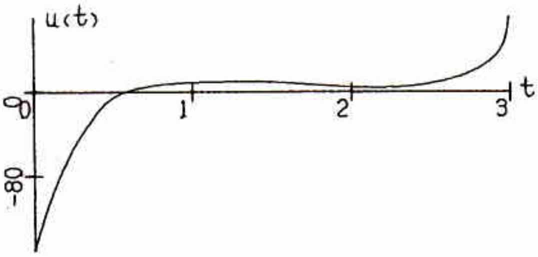

Figure 3c

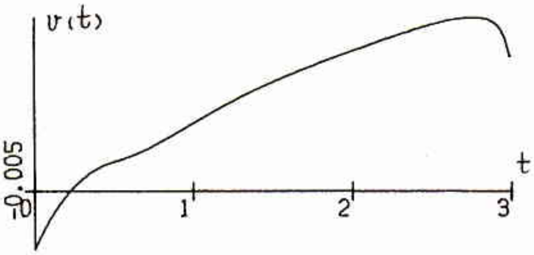

Figure 3d

REFERENCES

[1] Krasovskii, N.N., Tretyakov, V.E., Izv. Akad. Nauk SSSR. Tehn. Kibernetica 2(1983) 6-23.

[2] Krasovskii, N.N., Tretyakov, V.E., Probl. of Control and Inform.Theory 2(1983) 79-95.

[3] Krasovskii, N.N., Subbotin, A.I., Closed-Loop Differential Games (USSR, Moscow, 1974).

[4] Turova ,V.L., Izv. Akad. Nauk SSSR. Tehn. Kibernetika 4(1984) 148-155.