Pergamon

J. Appl. Maths Mechs, Vol. 58, No. 4, pp. 575–581, 1994
Copyright © 1995 Elsevier Science Ltd
Printed in Great Britain. All rights reserved
0021–8928/94 $24.00+0.00

0021–8928(94)00084-0

STRATEGIES OF MINIMAX AIMING IN THE DIRECTION OF THE QUASIGRADIENT†

G. G. GARNYSHEVA and A. I. SUBBOTIN

Ekaterinburg

(*Received* 15 *April* 1993)

A differential game with fixed end time is considered. In the general case, when the utility function is not assumed to be smooth, quasigradients should replace gradients in the construction of ε-optimal strategies. Constructions determining quasigradients are described.

In control theory and the theory of differential games one distinguishes two versions of the problem of feedback control. The first requires the determination of a strategy that will guarantee the solution of the problem for a fixed initial state of the system. In the other version one has to construct a *universal* strategy, to guarantee the solution for a whole domain of initial states. The construction of such strategies has been considered in several publications (see, e.g. [1–4]).

The present paper continues these investigations, using certain ideas due to Krasovskii (e.g. [1, 2]). The constructions will invoke, in addition, elements of non-smooth analysis and the theory of generalized (minimax and viscosity) solutions of first-order partial differential equations [5–13]. Our main result is a construction reminiscent of the usual definition of optimal strategy in classical dynamic programming when the utility function of the differential game is smooth, except that the gradient of the utility function (which may not exist) will be replaced by the quasigradient.

1. Let the motion of the controlled system be described by the equation

$$\dot{x}(t) = f(t, x(t), u(t), \upsilon(t)), \quad t_0 \le t \le \theta \tag{1.1}$$

where $x(t) \in \mathbf{R}^n$ is the phase state of the system at time t.

In the theory of differential games $u(t)$ and $\upsilon(t)$ are the controls of the first (minimizing) player and the second (maximizing) player. In control problems with a guaranteed result the first player tries to maintain the performance of the control process whatever the noise $\upsilon(t)$, which is "chosen" by a fictitious second player.

Let us consider a differential game in which the payoff functional is defined by the equation

$$\gamma(x(\cdot), u(\cdot), \upsilon(\cdot)) = \sigma(x(\theta)) - \int_{t_0}^{\theta} g(t, x(t), u(t), \upsilon(t))dt \tag{1.2}$$

It is assumed that

$$\begin{aligned} &\min_{u\in P} \max_{\upsilon\in Q}[\langle s, f(t,x,u,\upsilon)\rangle - g(t,x,u,\upsilon)] = \\ &= \max_{\upsilon\in Q} \min_{u\in P}[\langle s, f(t,x,u,\upsilon)\rangle - g(t,x,u,\upsilon)] = H(t,x,s) \end{aligned} \tag{1.3}$$

†*Prikl. Mat. Mekh.* Vol. 58, No. 4, pp. 5–11, 1994.

where $\langle s, f\rangle$ is the scalar product of the vectors $s, f \in \mathbf{R}^n$. It is also assumed that the functions f, g and σ are jointly continuous in all their variables. The function σ satisfies the condition

$$|\sigma(x)| \leq K_0(1+\|x\|), \quad \forall x \in R^n$$

where K_0 is some positive number. The functions f and g satisfy a Lipschitz condition in x

$$|H(t,x,s)-H(t,y,s)| \leq \lambda\|x-y\|(1+\|s\|) \quad x,y \in R^n \tag{1.4}$$

These conditions, which make the discussion easier, may be weakened.

Suppose that the first player has chosen some positional strategy U and some partition

$$\Delta=\{t_i : i=0,\ldots,m+1\}, \quad t_0<t_1<\ldots<t_{m+1}=\theta$$

of the time interval $[t_0, \theta]$. The strategy U will be identified here with an arbitrary function

$$[0,\theta]\times R^n \ni (t,x) \to U(t,x) \in P$$

We emphasize that $U(t,x)$ need not be continuous.

Let $S(t_0, x_0, U, \Delta)$ denote the set whose elements are the triples (controllable processes) $(x(\cdot), u(\cdot), \upsilon(\cdot))$ such that $\upsilon(\cdot):[t_0, \theta]\to Q$ is an arbitrary measurable function, $u(\cdot)$ a piecewise constant control of the form

$$u(t)=U(t_i, x(t_i)), \quad t_i \leq t < t_{i+1}, \quad i=0,1,2,\ldots,m$$

and $x(\cdot):[t_0, \theta]\to R^n$ is an absolutely continuous function satisfying Eq. (1.1) and the condition $x(t_0)=x_0$. A similar definition yields the set $S(t_0, x_0, V, \Delta)$, where V: $[0, \theta]\times R^n \to Q$ is the strategy of the second player.

The first (second) player aims to guarantee a minimum (maximum) value of the payoff functional. Optimal outcomes for the first and second players, respectively, will be the quantities $\Gamma^*(t_0, x_0)$ and $\Gamma_*(t_0, x_0)$ defined as follows:

$$\Gamma^*(t_0,x_0) := \inf_{U,\Delta}\Gamma_1(t_0,x_0,U,\Delta) \quad \Gamma_1(t_0,x_0,U,\Delta) := \sup\gamma(S(t_0,x_0,U,\Delta)$$

$$\Gamma_*(t_0,x_0) := \sup_{V,\Delta}\Gamma_2(t_0,x_0,V,\Delta) \quad \Gamma_2(t_0,x_0,V,\Delta) := \inf\gamma(S(t_0,x_0,V,\Delta)$$

where

$$\inf\rho(A) := \inf_{\alpha\in A}\rho(\alpha), \quad \sup\rho(A) := \sup_{\alpha\in A}\rho(\alpha)$$

We know that under the above conditions the optimal outcomes for the first and second players are identical. The quantity $\mathrm{Val}(t_0, x_0) := \Gamma^*(t_0, x_0)=\Gamma_*(t_0, x_0)$ is called the value of the differential game (1.1), (1.2). The value of a game depends on the initial position. We may therefore define the utility function $(t_0, x_0)\to \mathrm{Val}(t_0, x_0): [0, \theta]\times R^n \to R$.

These definitions presuppose the formalization of differential games proposed in [1, 14]. Although the formal notions of strategy and definitions of value in the theory of differential games may differ from one author to another, these formalizations turn out to be equivalent, in the sense that different definitions lead to the same utility function.

We know that the utility function is u-stable [14]. In order to formulate this property, we will need some notation. We define

$$E(t,x,\upsilon) := \mathrm{co}\{(f(t,x,u,\upsilon),\ g(t,x,u,\upsilon)) \in R^n\times R : u\in P\} \tag{1.5}$$

Let $S(t_0, x_0, z_0, \upsilon)$ denote the set of trajectories $x(\cdot), z(\cdot)): [t_0, \theta] \to R^n \times R$ of the differential inclusion

$$(\dot{x}(t), \dot{z}(t)) \in E(t, x(t), \upsilon), \quad t_0 \le t \le \theta \tag{1.6}$$

that satisfy the initial condition $(x(t_0), z(t_0)) = (x_0, z_0)$. A function $\rho(t, x)$: $[0, \theta] \times R^n \to R$ is said to be u-stable if, for any point $(t_0, x_0, z_0) \in \operatorname{gr}\rho := \{(t, x, \rho(t, x)): (t, x) \in [0, \theta] \times R^n\}$ and any control $\upsilon \in Q$, a trajectory $(x(\cdot), z(\cdot)) \in S(t_0, x_0, z_0, \upsilon)$ exists such that $z(t) \geqslant \rho(t, x(t))$ whenever $t \in [t_0, \theta]$. Note that u-stability implies that the over-graph of ρ is weakly invariant with respect to the differential inclusions (1.6). Now, the theory of generalized (minimax and viscosity) solutions of Hamilton–Jacobi equations considers what are known as upper and lower solutions (see, e.g. [6–11]). We know that u-stable functions are upper solutions of the Isaacs–Bellman equation

$$\partial\rho / \partial t + H(t, x, D_x\rho) = 0$$

According to the method of dynamic programming (see, e.g. [9]), if the utility function is differentiable, the players will have optimal strategies of the form

$$U_0(t, x) = u_0(t, x, s(t, x)), \quad V_0(t, x) = \upsilon_0(t, x, s(t, x))$$

where

$$u_0(t, x, s) \in \operatorname*{Arg\,min}_{u \in P}\{\max_{\upsilon \in Q}[\langle f(t, x, u, \upsilon), s\rangle - g(t, x, u, \upsilon)]\} \tag{1.7}$$

$$\upsilon_0(t, x, s) \in \operatorname*{Arg\,max}_{\upsilon \in Q}\{\min_{u \in P}[\langle f(t, x, u, \upsilon), s\rangle - g(t, x, u, \upsilon)]\} \tag{1.8}$$

$s(t, x) = D_x \mathrm{Val}(t, x)$, i.e. the strategies U_0 and V_0 are superpositions of the functions (pre-strategies) $u_0(t, x, s)$, $\upsilon_0(t, x, s)$ and the gradient $s(t, x)$ of the utility function with respect to x.

Here and below

$$\operatorname*{Arg\,max}_{z \in Z} h(z) := \{z^0 \in Z: h(z^0) \geqslant h(z) \forall z \in Z\}$$

$$\operatorname*{Arg\,min}_{z \in Z} h(z) := \{z_0 \in Z: h(z_0) \leqslant h(z) \forall z \in Z\}$$

2. In the general case, when the utility function is not differentiable, one can modify strategies of the above kind, replacing the gradient $s(t, x)$ by a "quasigradient" $s_\alpha(t, x)$ defined as follows.

Let $\rho(t, x)$ be a u-stable lower semicontinuous function satisfying the equality $\rho(\theta, x) = \sigma(x)$. This is the case, in particular, for the utility function Val. It can be shown that under the above assumptions ρ satisfies the estimate

$$\rho(t, x) \ge -K(1 + \|x\|), \quad \forall (t, x) \in [0, \theta] \times R^n \tag{2.1}$$

Consider the following transformation of ρ (α denotes a positive parameter)

$$\rho_\alpha(t, x) := \min_{y \in R^n}[\rho(t, y) + w_\alpha(t, x, y)] \tag{2.2}$$

$$w_\alpha(t, x, y) := \alpha^{-1}(e^{-\lambda t} - \alpha)\sqrt{\alpha^4 + \|x - y\|^2} \tag{2.3}$$

Choose a number $\alpha > 0$ so small that

$$L(\alpha) := \alpha^{-1}(e^{-\lambda\theta} - \alpha) > K \tag{2.4}$$

Note that

$$w_\alpha(t,x,y) > L(\alpha)\|x-y\| \tag{2.5}$$

for any $t\in[0, \theta]$, $x\in R^n$, $y\in R^n$. Taking into account that the function $y\to\rho(t, y)+w_\alpha(t, x, y)$ is lower semicontinuous, we see that the minimum in (2.2) is achieved.

Choose an arbitrary

$$y_\alpha(t,x)\in \operatorname{Arg}\min_{y\in R^n}[\rho(t,y)+w_\alpha(t,x,y)] \tag{2.6}$$

We assert that

$$\|y_\alpha(t,x)-x\|\le \frac{\rho(t,x)+K(1+\|x\|)+\alpha}{(L(\alpha)-K)} \tag{2.7}$$

Indeed, it follows from (2.2), (2.3) that

$$\rho_\alpha(t,x)\le\rho(t,x)+w_\alpha(t,x,x)\le\rho(t,x)+\alpha \tag{2.8}$$

In order to simplify the notation, we put $\eta:=y_\alpha(t, x)$. By (2.2), (2.1) and (2.5), we have

$$\rho_\alpha(t,x)=\rho(t,\eta)+w_\alpha(t,x,\eta)\ge -K(1+\|\eta\|)+w_\alpha(t,x,\eta)> -K(1+\|\eta\|)+L(\alpha)\|x-\eta\|=$$
$$=(L(\alpha)-K)\|x-\eta\|+K\|x-\eta\|-K\|\eta\|-K\ge(L(\alpha)-K)\|x-\eta\|-K(1+\|x\|)$$

Together with the previous estimate, this implies that

$$(L(\alpha)-K)\|x-\eta\|-K(1+\|x\|)\le\rho_\alpha(t,x)\le\rho(t,x)+\alpha$$

Thus, we have proved (2.7).

Let $D_xw(t, x, y)$ and $D_yw(t, x, y)$ be the gradients of w_α with respect to x and y. We have $D_xw(t, x, y)=-D_yw(t, x, y)$. Define

$$s_\alpha(t,x):=D_xw_\alpha(t,x,y_\alpha(t,x))=-D_yw_\alpha(t,x,y_\alpha(t,x)) \tag{2.9}$$

If the function $y\to\rho(t, y)$ is continuously differentiable in some neighbourhood of x containing the point $y_\alpha(t, x)$, then

$$D_y\rho(t,y_\alpha(t,x))+D_yw_\alpha(t,x,y_\alpha(t,x))=0$$

Therefore $s_\alpha(t, x):=D_y\rho(t, y_\alpha(t, x))$. Noting that $\lim_{\alpha\downarrow 0}y_\alpha(t, x)=x$, we obtain

$$\lim_{\alpha\to 0}s_\alpha(t,x)=D_x\rho(t,x)$$

For that reason we may call $s_\alpha(t, x)$ the quasigradient of ρ with respect to x.

Define a strategy of the first player by

$$U_\alpha(t,x)=u_0(t,x,s_\alpha(t,x)) \tag{2.10}$$

where the function u_0 (pre-strategy) is defined by (1.7).

Theorem. For any compact set $D\subset[0, \theta]\times R^n$ and any positive number ε $\alpha>0$ and $\beta>0$ exist such that

$$\Gamma_1(t_0, x_0, U_\alpha, \Delta) \le \rho(t_0, x_0) + \varepsilon \tag{2.11}$$

for all $(t_0, x_0) \in D$ and any partitions with $\text{diam}(\Delta) := \max\{t_{i+1} - t_i : i = 0, 1, \ldots, m\} \leq \beta$. In particular, if $\rho = \text{Val}$, the strategy U_α is ε-optimal and universal.

Proof. Inequality (2.8) holds, as does the following estimate for any bounded set $M \subset R^n$ a number $\nu(\alpha)$ such that

$$\rho_\alpha(\theta, x) \ge \sigma(x) - \nu(\alpha), \quad \lim_{\alpha \downarrow 0} \nu(\alpha) = 0 \tag{2.12}$$

This follows from (2.7) and the continuity of σ.

Using condition (1.4), one can verify that a function w_α of the form (2.3) satisfies the inequality

$$\partial w_\alpha / \partial t + H(t, x, D_x w_\alpha) - H(t, y, -D_y w_\alpha) \le 0 \tag{2.13}$$

Inequalities of this type play an important role in the theory of viscosity solutions of Hamilton–Jacobi equations (see, e.g. condition A4 in [8]).

Choose the parameter $\alpha \in (0, 2e^{-\lambda\theta}]$ so that inequality (2.4) holds. Let $(x(\cdot), u(\cdot), \upsilon(\cdot)) \in S(t_0, x_0, U_\alpha, \Delta)$. We assert that

$$\begin{aligned} &\rho_\alpha(t_{i+1}, x(t_{i+1})) - \int_{t_i}^{t_{i+1}} g(\tau, x(\tau), u(\tau), \upsilon(\tau)) d\tau \le \\ &\le \rho_\alpha(t_i, x(t_i)) + h_1(t_{i+1} - t_i)(t_{i+1} - t_i) \end{aligned} \tag{2.14}$$

Here and below $h_i(\delta) \to 0$ as $\delta \to 0$.

It follows from this estimate that

$$\rho_\alpha(\theta, x(\theta)) - \int_{t_0}^{\theta} g(\tau, x(\tau), u(\tau), \upsilon(\tau)) d\tau \le \rho_\alpha(t_0, x_0) + h_2(\beta) \tag{2.15}$$

Combining this estimate with (2.8) and (2.12), we conclude that

$$\gamma(x(\cdot), u(\cdot), \upsilon(\cdot)) = \sigma(x(\theta)) - \int_{t_0}^{\theta} g(\tau, x(\tau), u(\tau), \upsilon(\tau)) d\tau \le \rho(t_0, x_0) + h_2(\beta) + \alpha + \nu(\alpha) \tag{2.16}$$

as required.

Thus, it remains to prove (2.14). To simplify the notation, let us put

$$\tau = t_i, \quad \tau + \delta = t_{i+1}, \quad \xi = x(t_i)$$
$$s^* = s_\alpha(\tau, \xi), \quad u^* = U_\alpha(\tau, \xi) = u_0(\tau, \xi, s^*)$$

We have to prove that

$$\rho_\alpha(\tau + \delta, x(\tau + \delta)) - \int_{\tau}^{\tau+\delta} g(t, x(t), u^*, \upsilon(t)) dt - h_1(\delta)\delta \le \rho_\alpha(\tau, \xi) \tag{2.17}$$

Put

$$f^* = \frac{1}{\delta} \int_{\tau}^{\tau+\delta} f(t, x(t), u^*, \upsilon(t)) dt, \; g^* = \frac{1}{\delta} \int_{\tau}^{\tau+\delta} g(t, x(t), u^*, \upsilon(t)) dt \tag{2.18}$$

Define $\eta = y_\alpha(\tau, \xi)$, $\upsilon_* = \upsilon_0(\tau, \eta, s^*)$. By (2.2) and (2.6), we have

$$\rho_\alpha(\tau, \xi) = \rho(\tau, \eta) + w_\alpha(\tau, \xi, \eta) \tag{2.19}$$

The u-stability property of ρ implies the existence of a vector $(f_*, g_*) \in R^n \times R$ such that

$$\operatorname{dist}((f_*, g_*), E(\tau, \eta, \upsilon_*)) \le h_1(\delta), \quad \rho(\tau, \eta) + g_*\delta \ge \rho(\tau+\delta, \eta + f_*\delta) \tag{2.20}$$

It therefore follows from (2.19) and (2.20) that

$$\rho_\alpha(\tau, \xi) \ge \rho(\tau+\delta, \eta + f_*\delta) + w_\alpha(\tau, \xi, \eta) - g_*\delta \ge \rho_\alpha(\tau+\delta, \xi + f^*\delta) - r$$
$$r := w_\alpha(\tau+\delta, \xi + f^*\delta, \eta + f_*\delta) - w_\alpha(\tau, \xi, \eta) + g_*\delta$$

The last estimate follows from (2.2). Since $\xi + f^*\delta = x(\tau+\delta)$, we obtain

$$\rho_\alpha(\tau, \xi) \ge \rho_\alpha(\tau+\delta, x(\tau+\delta)) - r \tag{2.21}$$

Let us estimate r. Since w_α is differentiable, we have

$$r - g_*\delta = [(\partial w_\alpha / \partial t)(\tau, \xi, \eta) + \langle s^*, f^*\rangle - \langle s^*, f_*\rangle]\delta + h_3(\delta)\delta$$

where $s^* := s_\alpha(\tau, \xi) := D_x w_\alpha(t, \xi, \eta) = -D_y w_\alpha(t, \xi, \eta)$ (see (2.6) and (2.9)).

Recall that $u^* = U_\alpha(\tau, \xi) = u_0(\tau, \xi, s^*)$. By (1.7), we have $\langle s^*, f(\tau, \xi, u^*, \upsilon)\rangle - g(\tau, \xi, u^*, \upsilon) \le H(\tau, \xi, s^*)$ for any $\upsilon \in Q$. Therefore

$$\langle s^*, f^*\rangle - g^* \le H(\tau, \xi, s^*) + h_3(\delta)$$

Recall, moreover, that $\upsilon_* = \upsilon_0(\tau, \eta, s^*)$. It therefore follows from (1.8) that

$$\langle s^*, f(\tau, \eta, u, \upsilon_*)\rangle - g(\tau, \eta, u, \upsilon_*) \ge H(\tau, \eta, s^*), \quad \forall u \in P$$

Consequently

$$\langle s^*, f_*\rangle - g_* \ge H(\tau, \eta, s^*) - h_4(\delta)$$

Thus, we obtain

$$\left| \begin{array}{l} (\partial w_\alpha / \partial t)(\tau, \xi, \eta) + \langle s^*, f^*\rangle - \langle s^*, f_*\rangle \le \\ \le (\partial w_\alpha / \partial t)(\tau, \xi, \eta) + H(\tau, \xi, s^*) - H(\tau, \eta, s^*) + g^* - g_* + h_5(\delta) \end{array} \right.$$

Taking (2.13) into account, we obtain

$$(\partial w_\alpha / \partial t)(\tau, \xi, \eta) + \langle s^*, f^*\rangle - \langle s^*, f_*\rangle \le g^* - g_* + h_5(\delta)$$

Thus, $r \le g^*\delta + h_5(\delta)$. Substituting this bound into (2.21), we obtain

$$\rho_\alpha(\tau, \xi) \ge \rho_\alpha(\tau+\delta, x(\tau+\delta)) - (g^* + h_5(\delta))\delta$$

(recall that g^* was defined by (2.18)). We have thus proved (2.17).

If the payoff functional does not contain an integral term, the role of the penalty function in the transformation (2.2) may be assigned to the function

$$w_\alpha(x, y) = \|x - y\|^2 / (2\alpha)$$

In that case the equality

$$\rho_\alpha(t,x) := \min_{y\in R^n}[\rho(t,y)+\|x-y\|^2/(2\alpha)]$$

is, in fact, a well-known transform, used in convex analysis and sometimes called the Yosida–Moreau transform (see, e.g. [13]). Equality (2.2) may be viewed as a modification of this transform.

An analogous construction of ε-optimal universal strategies may be proposed for time-optimal control.

The construction proposed above is a fairly compact way of proving the existence of an ε-saddle point in a differential game. However, it is not very suitable for computer algorithmization. Nevertheless, one can modify the penalty function w_ε, in particular, adjusting it to ensure stability with respect to information or computational errors.

Note that the above solution is based on a "smoothing" transform of the utility function. In that connection we must mention [15], in which a non-traditional approach to the synthesis of strategies is proposed, based on the use of the concept of the analytical centre of a convex set and internal smoothed realizations of the controls.

This research was supported financially by the Russian Fund for Fundamental Research (93-011-16032).

REFERENCES

1. KRASOVSKII N. N., *Control of a Dynamical System*. Nauka, Moscow, 1985.
2. KRASOVSKII N. N., Differential games. Approximate and formal models. *Mat. Sbornik* **107**, 4, 541–471, 1978.
3. KONONENKO A. F., On equilibrium positional strategies in non-antagonistic differential games. *Dokl. Akad. Nauk SSSR* **231**, 2, 285–288, 1976.
4. SUBBOTINA N. N., Universal optimal strategies in positional differential games. *Differents. Uravn.* **19**, 11, 1890–1896, 1983.
5. SUBBOTIN A. I., Generalization of the fundamental equation of the theory of differential games. *Dokl. Akad. Nauk SSSR* **254**, 2, 293–297, 1980.
6. CRANDALL M. G. and LIONS P.-L., Viscosity solutions of Hamilton–Jacobi equations. *Trans. Amer. Math. Soc.* **277**, 1, 1–42, 1983.
7. CRANDALL M. G., EVANS L. C. and LIONS P. -L., Some properties of viscosity solutions of Hamilton–Jacobi equations. *Trans. Amer. Math. Soc.* **282**, 2, 487–502, 1984.
8. CRANDALL M. G., ISHII H. and LIONS P. -L., Uniqueness of viscosity solutions of Hamilton–Jacobi equations. *J. Math. Soc. Japan* **39**, 4, 581–596, 1987.
9. ISAACS R., *Differential Games*. Wiley, New York, 1965.
10. SUBBOTIN A. I., *Minimax Inequalities and Hamilton–Jacobi Equations*. Nauka, Moscow, 1991.
11. SUBBOTIN A. I., Generalized characteristics of first-order partial differential equations. Preprint CRM-1848, Montreal, 1993.
12. CLARKE F. H., *Optimization and Nonsmooth Analysis*. John Wiley, New York, 1983.
13. AUBIN J.-P., *Non-linear Analysis and its Economic Applications*. Mir, Moscow, 1988.
14. KRASOVSKII N. N. and SUBBOTIN A. I., *Positional Differential Games*. Nauka, Moscow, 1974.
15. SONNEVEND G., Constructing feedback control in differential games by use of "central" trajectories. Schwespunktprogramm DFG. Anwendungsbezogen Optimierung und Steuerung. Institut für angewandte Mathematik und Statistik. Universität Würzburg, 1992, Report No. 385.

Translated by D.L.