Математические структуры и моделирование
2002, вып. 10, с. 53–64

УДК 517.11+519.92

ИГРОВОЙ ПОДХОД К РЕШЕНИЮ ЗАДАЧИ ВЫБОРА ОПТИМАЛЬНОГО УПРАВЛЕНИЯ АППАРАТОМ, ДВИГАЮЩИМСЯ К ГРУППЕ ЦЕЛЕЙ

А.К. Гуц

The game theory approach to optimal conrol of apparatus that is moving to some group of aims is given. The well-known results of Krasovsky are used.

Рассмотрим задачу движения аппарата A, который принимает сигналы от N целей $M_i,\ i = 1, ..., N$, определяет их координаты и старается выйти к этой группе целей так, чтобы быть как можно ближе к привилегированной цели P. При этом будем считать, что координаты каждой цели M_i «плавают», т.е. могут меняться в пределах некоторой ограниченной области V_i. Аппарат принимает сигнал от каждой цели не постоянно, а в определенные отрезки времени. Следовательно, в исходный момент начала движения $t = t_0$ не все цели фиксируются: число зафиксированных целей N_t меняется в пределах $0 \le N_t \le N$ при $t \in [t_0, T]$. Момент T – это время, к которому аппарат должен максимально сблизиться с приоритетной группой целей. Приоритет π_i цели M_i – это функция от $|M_i - P|$; он тем больше, чем ближе цель к точке P.

Почему нас интересует выход у группе целей, ближайших к P? Потому, что нацеливание на конкретную цель в условиях, когда может так случится, что уже после начала движения аппарата цель на отрезке времени $[t_0, T]$ вообще может не «заговорить», или «замолчать» тогда, когда до нее оставалось совсем немного, или станет недоступной в последний момент по ряду особых причин, способно сорвать задачу выхода к цели P. Иначе говоря, помимо основной задачи – выхода к P должна ставиться и дополнительная задача: при срыве основной задачи выйти к другим приоритетным целям. Надо двигаться к P так, чтобы всегда была возможность начать движение к ближайшей к P другой приоритетной цели. Ясно, это связано с тем, что аппарат располагает ограниченными ресурсами топлива, и с тем, что его вылет не должен быть бесцельным и безрезультативным. Наконец отметим, что движение не по прямой к цели P ценно еще тем, что позволяет более точно вычислять ее координаты.

E-mail: guts@univer.omsk.su
Омский государственный университет

Работа поддержана грантом РФФИ (проект 01-01-00303)

1. Постановка игровой задачи

Дадим математическую формулировку описанной во введении задачи управляемого движения аппарата. Используем подход, разработанный в теории дифференциальных игр [1].

Движения аппарата описываем системой дифференциальных уравнений

$$\begin{cases} \dot{\mathbf{y}} = \mathbf{f}(t, \mathbf{y}, u), \\ \mathbf{y}(t_0) = \mathbf{y}_0, \end{cases} \tag{1}$$

где $u = u(t)$ – управление аппаратом, выбираемое в момент времени t. Для нас это «игрок» A. «Игрок» M – это координаты целей M_i ($i = 1, ...N$). С учетом того, что эти координаты «плавают», примем, что цель M_i описывается уравнениями

$$\begin{cases} \dot{\mathbf{z}}_i = v_i(t)), \\ \mathbf{z}_i(t_0) = \mathbf{z}_{i0}, \\ \ (i = 1, , , , N), \end{cases} \tag{2}$$

где $v_i = v_i(t)$ – управление, задающее «плавание» координат цели M_i. Поскольку при $t = t_0$ не все цели могут фиксироваться, то не все уравнения (2) учитываются на некотором отрезке времени $[t_0, t_1]$. Иначе говоря, начальные условия надо понимать как $\mathbf{z}_i(t_i) = \mathbf{z}_{i0}$, где t_i – момент времени, когда впервые фиксируются координаты $\mathbf{z}_{i0}$ цели M_i. Управление $v_i(t)$ для «игрока» A всегда является неизвестной функцией.

Кроме того, если цель M_i «выключается» и «включается», то игроку A известны только координаты $\mathbf{z}_i(\theta_i(t))$, где

$$\theta_i(t) = \begin{cases} t, t \in [\alpha_{ij}, \beta_{ij}] \\ \beta_{ij}, t \in (\beta_{ij}, \alpha_{i,j+1}), \end{cases}$$

$$t_0 \le \alpha_{i0} \le \beta_{i0} \le ... \le \alpha_{ij} \le \beta_{ij} \le ... \le \alpha_{in_i} \le \beta_{in_i} = T.$$

Это условие – не что иное, как условие информированности игрока A о движении игрока M. На интервале «молчания» $(\beta_{ij}, \alpha_{i,j+1})$ цели M_i ее координаты берутся те, что были определены в момент $t = \beta_{ij}$.

Задачей игрока A является минимизация функционала

$$J(t_0, \mathbf{y}_0, \mathbf{z}_{i0}, u, v) = \sum_{i=1}^{N} e^{\pi_i(T)} |\mathbf{z}_i(T) - \mathbf{y}(T)|, \tag{3}$$

где $\pi_i(T) = \pi_i(|\mathbf{z}_i(T) - P|)$.

Если бы в момент начала движения аппарата A были известны, зафиксированы координаты всех N целей, то мы имели бы стандартную дифференциальную игру (1),(2), (3). Но в нашем случае число зафиксированных целей меняется, точнее, нарастает с течении времени. Поэтому мы имеем конечную последовательность игр. Новая игра начинается в момент $t = t_k, t_0 < t_k < T$, когда к числу N_{k-1} ранее зафиксированных целей $M_1, ..., M_{k-1}$ добавляется еще одна цель – M_k. Следовательно, при $t < t_{k-1}$ имеем игру

$$\begin{cases} \dot{\mathbf{y}} = \mathbf{f}(t, \mathbf{y}, u), \\ \mathbf{y}(t_{k-1}) = \mathbf{y}_{k-1}, \end{cases} \qquad \text{(i=1,...,k-1)} \qquad \begin{cases} \dot{\mathbf{z}}_i = v_i, \\ \mathbf{z}_i(t_{k-1}) = \mathbf{z}_{i,k-1}, \end{cases}$$

$$J_{k-1} = \sum_{i=1}^{N_{k-1}} e^{\pi_i(T)} |\mathbf{z}_i - \mathbf{y}(T)|,$$

а при $t \geq t_k$ имеем игру

$$\begin{cases} \dot{\mathbf{y}} = \mathbf{f}(t, y, u), \\ \mathbf{y}(t_k) = \mathbf{y}_k, \end{cases} \qquad \begin{cases} \dot{\mathbf{z}}_i = v_i, \\ \mathbf{z}_i(t_k) = \mathbf{z}_{ik}, \end{cases}$$

$$(i = 1, ..., k)$$

$$J_k = \sum_{i=1}^{N_k} e^{\pi_i(T)} |\mathbf{z}_i - \mathbf{y}(T)|.$$

Значение $\mathbf{y}_k$ – координаты аппарата на момент $t = t_k$ из предыдущей игры. Это же верно для $\mathbf{z}_{ik}$ при $i < k$, а $\mathbf{z}_{ik}$ – это впервые зафиксированные координаты цели M_k при $t = t_k$.

Очевидно $J_k \geq J_{k-1}$.

По существу не всегда при приеме сигналов от новой цели нужно начинать новую игру. Это следует делать лишь в том случае, когда на новую цель есть время, а кроме того, если новая цель имеет очень высокий приоритет π_k.

2. Метод Красовского решения задачи сближения

Наша задача будет решена, если будет найден способ выбора управления $u(t)$, который в момент t задается на основе полученной информации о координатах $\mathbf{y}(t), \mathbf{z}_i(t)$. Оптимальное управление $\overset{\circ}{u}(t)$ обладает свойством

$$J(t_0, \mathbf{y}_0, \mathbf{z}_{i0}, \overset{\circ}{u}, \widetilde{v}) \leq \min_{u(t)} \sup_{\widetilde{v}} \inf_{A(t),\ M_i(t)} J(t_0, \mathbf{y}_0, \mathbf{z}_{i0}, u, \widetilde{v}) \tag{4}$$

для любых управлений $\widetilde{v}_i$ и данных $t_0, \mathbf{y}_0, \mathbf{z}_{i0}$. Ясно, что важно оценить функционал $J(t_0, \mathbf{y}_0, \mathbf{z}_{i0}, u, \widetilde{v})$, опираясь на данные $t_0, \mathbf{y}_0, \mathbf{z}_{i0}$. Это можно сделать в так называемом регулярном случае линейной игры с помощью гипотетического рассогласования [1, с.105]. *Гипотетическое рассогласование* – это величина $\varepsilon^0(t_0, \mathbf{y}_0, \mathbf{z}_{i0})$, обладающая свойством

$$\sup_{v(t)} J(t_0, \mathbf{y}_0, \mathbf{z}_{i0}, \overset{\circ}{u}, v) \geq \varepsilon^0(t_0, \mathbf{y}_0, \mathbf{z}_{i0}),$$

т.е. она оценивает наименьшее значение функционала (3) снизу. Более точно и более значимо то, что $\varepsilon^0(t_0, \mathbf{y}_0, \mathbf{z}_{i0})$ оценивает снизу оптимальный результат игры (1), (2), (3), (4). Заманчиво организовать управление $u(t)$ так, чтобы эта нижняя оценка достигалась.

В [1] введена теоретическая конструкция *экстремального управления* $u_e(t)$, которое в так называемом регулярном случае обладает свойством

$$J(t_0, \mathbf{y}_0, \mathbf{z}_{i0}, u_e, v) \le \varepsilon^0(t_0, \mathbf{y}_0, \mathbf{z}_{i0}),$$

для начальной позиции $(t_0, \mathbf{y}_0, \mathbf{z}_{i0})$ и любого управления $v(t)$ (см. [1, с.153], теорема 16.1). Причем для экстремальных управлений $u_e(t), v_e(t)$

$$J(t_0, \mathbf{y}_0, \mathbf{z}_{i0}, u_e, v_e) = \varepsilon^0(t_0, \mathbf{y}_0, \mathbf{z}_{i0}).$$

Поэтому экстремальное управления $u_e(t)$ в регулярном случае дает решение минимальной задачи (4) и следовательно, мы имеем оптимальное управление (см. [1, с.155], теорема 16.3). Предполагается, что при этом игра линейная, т.е.

$$\begin{cases} \dot{\mathbf{y}}_i = a(t)\mathbf{y} + u + \mathbf{g}(t), \\ \dot{\mathbf{z}}_i = b(t)\mathbf{z} + v + \mathbf{h}(t), \end{cases}$$

а управления $u(t), v(t)$ подчинены некоторым ограничениям $\mathcal{U}_t, \mathcal{V}_t$ (в каждый момент времени t) математического характера. Возвращаясь к нашей игре (1), (2), (3), (4), примем, что она линейная и приоритеты π_i уже определены, и для простоты положим, что $\pi_i = const$. Тогда игру (1), (2), (3) заменим на следующую: берем N экземпляров уравнения (1) – это игрок Y причем $i-$й экземпляр умножается на $\exp(\pi_i)$. Получаем игру

$$\begin{cases} \dot{Y} = A(t)Y + U(t) + F(t), \\ Y(t_0) = Y_0, \end{cases} \tag{5}$$

где

$$Y_1 = e^{\pi_1} y_1, \quad Y_2 = e^{\pi_1} y_2, \quad Y_3 = e^{\pi_1} y_3,$$

$$Y_4 = e^{\pi_2} y_1, \quad Y_5 = e^{\pi_2} y_2, \quad Y_6 = e^{\pi_2} y_3,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$Y_{3N-2}4 = e^{\pi_N} y_1, \quad Y_{3N-1} = e^{\pi_N} y_2, \quad Y_{3N} = e^{\pi_N} y_3.$$

Игрок Z – это игра для N целей:

$$e^{\pi_i} \dot{z}_t = v_i e^{\pi_i} \ (i = 1, ..., N)$$

или

$$\begin{cases} \dot{Z} = V(t), \\ Z(t_0) = Z_0, \end{cases} \tag{6}$$

где

$$Z_1 = e^{\pi_1} z_{11}, \quad Z_2 = e^{\pi_1} z_{12}, \quad Z_3 = e^{\pi_1} z_{13},$$

$$Z_4 = e^{\pi_2} z_{21}, \quad Z_5 = e^{\pi_2} z_{22}, \quad Z_6 = e^{\pi_2} z_{23},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$Z_{3N-2} = e^{\pi_N} z_{N1}, \quad Z_{3N-1} = e^{\pi_N} z_{N2} \quad Z_{3N} = e^{\pi_N} z_{N3}.$$

И наконец,

$$J(t_0, y_0, z_{i0}, u_e, v) \equiv J(t_0, Y_0, Z_{i0}, U, V) = |Y(t) - Z(t)|. \tag{7}$$

Иначе говоря, имеем игру в $3N$-мерном пространстве.

Далее, справедлива формула [1]

$$\begin{aligned} \varepsilon^0(t_0, Y_0, Z_0) = \max_{|\vec{l}|=1} \{ \rho^{(2)}(t_0, T, \vec{l}) - \rho^{(1)}(t_0, T, \vec{l}) + \\ + \vec{l}^{\,tr} [\underset{\circ}{Z}(t_0, Z_0, T) - \underset{\circ}{Y}(t_0, Y_0, T)] \}, \end{aligned} \tag{8}$$

при условии, что правая часть этой формулы положительна; в противном случае – $\varepsilon^0(t_0, Y_0, Z_0) = 0$.

$$\vec{l} = \begin{pmatrix} l_1 \\ \cdot \\ \cdot \\ \cdot \\ l_{3N} \end{pmatrix}, \qquad \vec{l}^{\,tr} = (l_1, ..., l_{3N}).$$

Здесь

$$\rho^{(1)}(t_0, T, \vec{l}) = \max_{U(t) \in \mathcal{U}_t} \int_{t_0}^{T} \vec{l}^{\,tr} K^{(1)}(T, t) U(t) dt,$$

$$\rho^{(2)}(t_0, T, \vec{l}) = \max_{V(t) \in \mathcal{V}_t} \int_{t_0}^{T} \vec{l}^{\,tr} K^{(2)}(T, t) V(t) dt,$$

$\mathcal{U}_t, \mathcal{V}_t$ – ограничения на выбор управлений, $K^{(1)}(T, t)$ – фундаментальная матрица решений уравнения

$$\begin{cases} \dfrac{dY}{d\tau} = A(t) Y, \\ Y|_{\tau=s} = E, \end{cases}$$

E – единичная матрица и $K^{(2)}(T, t)$ – фундаментальная матрица решений уравнения

$$\begin{cases} \dfrac{dZ}{d\tau} = 0, \\ Z|_{\tau=s} = E, \end{cases}$$

$\underset{\circ}{Y}(t_0, Y_0, t)$ – решение уравнения

$$\begin{cases} \dot{Y} = A(t) Y + F(t), \\ Y(t_0) = Y_0, \end{cases}$$

и, наконец, $\underset{\circ}{Z}(t_0, Z_0, t)$ – решение уравнения

$$\begin{cases} \dot{Z} = 0, \\ Z(t_0) = Z_0. \end{cases}$$

Задача *регулярная*, если в (8) вектор $\vec{l}$, реализующий максимум, единственен. В нашем случае это выполняется, поскольку траектория второго игрока Z является точкой ([1, с.115] и определение регулярности 13.1).

Если взять любой момент времени $t = t_*$ и начать игру (1),(2), (3), исходя из положения $(t_*, \mathbf{y}_*, \mathbf{z}_{i*})$, где $\mathbf{y}_* = \mathbf{y}(t_*), \mathbf{z}_{i*} = \mathbf{z}_i(t_*)$, или игру (5),(6), то при выборе экстремального управления $v_e(t)$ (соотв.: $V_e(t)$), $t_* < t < T$

$$J(t_*, \mathbf{y}_*, \mathbf{z}_{i*}, u(t), v_e) \geq \varepsilon^0(t_*, \mathbf{y}_*, \mathbf{z}_{i*})$$

(или соответственно

$$J(t_*, Y_*, Z_*, U(t), V_e) \geq \varepsilon^0(t_*, Y_*, Z_*)),$$

т.е. на промежутке времени $[t_*, T]$ гипотетическое рассогласование – это оценка снизу для функционала J.

Пример 1.

Для простого движения аппарата – $a(t) \equiv 0,\ \mathbf{g}(t) \equiv 0$ и значит $A(t) \equiv 0, F(t) \equiv 0$. Пусть задача плоская и имеется только три цели, т.е. $N = 3$.

Тогда

$$K^{(1)}(T,t) = E, \quad K^{(2)}(T,t) = E,$$

$$\underset{\circ}{Y}\,(t_0, Y_0, t) = Y_0, \quad \underset{\circ}{Z}\,(t_0, Z_0, t) = Z_0,$$

$$\rho^{(1)}(t_0, T, \vec{l}) = \max_{U(t) \in \mathcal{U}_t} \int_{t_0}^{T} \sum_{i=1}^{6} l_i U_i(t) dt,$$

$$\rho^{(2)}(t_0, T, \vec{l}) = \max_{V(t) \in \mathcal{V}_t} \int_{t_0}^{T} \sum_{i=1}^{6} l_i V_i(t) dt,$$

и поэтому

$$\varepsilon^0(t_0, Y_0, Z_0) = \max_{|\vec{l}|=1} \{ \max_{V(t) \in \mathcal{V}_t} \int_{t_0}^{T} \sum_{i=1}^{6} l_i V_i(t) dt - \max_{U(t) \in \mathcal{U}_t} \int_{t_0}^{T} \sum_{i=1}^{6} l_i U_i(t) dt + \sum_{i=1}^{6} l_i (Z_{i0} - Y_{i0}) \}$$

или

$$\varepsilon^0(t_0, Y_0, Z_0) = \max_{|\vec{l}|=1} \{ \max_{V(t) \in \mathcal{V}_t} \int_{t_0}^{T} \sum_{i=1}^{6} l_i V_i(t) dt - \max_{U(t) \in \mathcal{U}_t} \int_{t_0}^{T} \sum_{i=1}^{6} l_i U_i(t) dt +$$

$$+ e^{\pi_1} [l_1(z_{110} - y_{10}) + l_2(z_{120} - y_{20})] + e^{\pi_2} [l_3(z_{210} - y_{10}) + l_4(z_{220} - y_{20})] +$$

$$+ e^{\pi_3} [l_5(z_{310} - y_{10}) + l_6(z_{320} - y_{20})] \}.$$

Если $V(t) \equiv 0$, т.е. «плавания» координат целей нет, то $\rho^{(2)} \equiv 0$.

Пусть управление $U(t)$ удовлетворяет условию

$$\mathcal{U}_t : \sum_{i=1}^{6} [U_i]^2(t) \leq \delta^2.$$

Тогда

$$\rho^{(1)}(t_0,T,\vec{l}) = \max_{U(t)\in\mathcal{U}_t} \int\limits_{t_0}^{T} \sum_{i=1}^{6} l_i U_i(t)dt \le \int\limits_{t_0}^{T} [\max_{U(t)\in\mathcal{U}_t} \sum_{i=1}^{6} l_i U_i(t)]dt =$$

$$\le \delta \int\limits_{t_0}^{T} |\vec{l}|dt = \delta|\vec{l}|(T-t_0).$$

Максимум в $\rho^{(1)}$, как легко видеть, реализуется для управления $u_i(t) = \delta l_i/|\vec{l}|$, и, значит,

$$\rho^{(1)}(t_0,T,\vec{l}) = \delta|\vec{l}|(T-t_0).$$

Но тогда

$$\varepsilon^0(t_0,\mathbf{y}_0,\mathbf{z}_{i0}) = -\delta(T-t_0) + \max_{|\vec{l}|=1}[e^{\pi_1}[l_1(z_{110}-y_{10}) + l_2(z_{120}-y_{20})]+$$

$$+e^{\pi_2}[l_3(z_{210}-y_{10}) + l_4(z_{220}-y_{20})] + e^{\pi_3}[l_5(z_{310}-y_{10}) + l_6(z_{320}-y_{20})]\}.$$

Из геометрических соображений видно, что

$$\varepsilon^0(t_0,\mathbf{y}_0,\mathbf{z}_{i0}) = -\delta(T-t_0) + [e^{\pi_1}[l_1^0(z_{110}-y_{10}) + l_2^0(z_{120}-y_{20})]+$$

$$+e^{\pi_2}[l_3^0(z_{210}-y_{10}) + l_4^0(z_{220}-y_{20})] + e^{\pi_3}[l_5^0(z_{310}-y_{10}) + l_6^0(z_{320}-y_{20})]\},$$

где

$$l_1^0 = \frac{e^{\pi_1}[l_1^0(z_{110}-y_{10})}{\alpha}, \quad l_2^0 = \frac{e^{\pi_1}[l_1^0(z_{120}-y_{20})}{\alpha}$$

$$l_3^0 = \frac{e^{\pi_2}[l_1^0(z_{210}-y_{10})}{\alpha}, \quad l_4^0 = \frac{e^{\pi_2}[l_1^0(z_{220}-y_{20})}{\alpha}$$

$$l_5^0 = \frac{e^{\pi_3}[l_1^0(z_{310}-y_{10})}{\alpha}, \quad l_6^0 = \frac{e^{\pi_3}[l_1^0(z_{320}-y_{20})}{\alpha}$$

и

$$\alpha = [e^{2\pi_1}(z_{110}-y_{10})^2 + e^{2\pi_1}(z_{120}-y_{20})^2+$$

$$+e^{2\pi_2}(z_{210}-y_{10})^2 + e^{2\pi_2}(z_{220}-y_{20})^2 + e^{2\pi_3}(z_{310}-y_{10})^2 + e^{2\pi_3}(z_{320}-y_{20})^2]^{1/2}.$$

Поэтому

$$\varepsilon^0(t_0,\mathbf{y}_0,\mathbf{z}_{i0}) = -\delta(T-t_0) + \alpha. \tag{9}$$

Гипотетическое рассогласование (9), как мы помним, дает оценку снизу для функционала J в момент времени $t = t_0$, отвечающего моменту начала движения аппарата.

Найдем экстремальное управление аппаратом. В соответствии с [1, с.117] для этого решаем систему уравнений

$$\begin{cases} \dfrac{dS^{(1)}}{dt} = 0, \\ S^{(1)}(T) = \vec{l}^{\,0}, \end{cases}$$

Ясно $S^{(1)} = \vec{l}^{\,0}$. Тогда

$$[S^{(1)}]^{tr}(t_0)U_e = \max_{U\in\mathcal{U}_{t_0}} [S^{(1)}]^{tr}(t_0)U$$

или

$$\sum_{i=1}^{6} l_i^0 U_{ie} = \max_{U\in\mathcal{U}_{t_0}} \sum_{i=1}^{6} l_i^0 U_i.$$

Вспомним, что уравнения для игры игрока Y – это одно двумерное уравнение, взятое трижды с весом $\exp(\pi_j)$, $j = 1, 2, 3$. Поэтому должны учитываться соотношения:

$$U_1 = e^{\pi_1 - \pi_2} U_3 = e^{\pi_1 - \pi_3} U_5, \quad U_2 = e^{\pi_1 - \pi_2} U_4 = e^{\pi_1 - \pi_3} U_6,$$

$$U_1^2 + U_2^2 \le \delta^2.$$

Тогда

$$[S^{(1)}]^{tr}(t_0) U_e = \max_{U \in \mathcal{U}_{t_0}} \{(l_1^0 + e^{\pi_2 - \pi_1} l_3^0 + e^{\pi_3 - \pi_1} l_5^0) U_1 + (l_2^0 + e^{\pi_2 - \pi_1} l_4^0 + e^{\pi_3 - \pi_1} l_6^0) U_2\}. \quad (10)$$

Задача (10) сводится к нахождению условного экстремума функции

$$\beta = a U_1 + B U_2,$$

$$U_1^2 + U_2^2 = \delta^2.$$

Отсюда имеем

$$(U_{1e}, U_{2e}) = \left(\frac{l_1^0 + e^{\pi_2 - \pi_1} l_3^0 + e^{\pi_3 - \pi_1} l_5^0}{\gamma}, \frac{l_2^0 + e^{\pi_2 - \pi_1} l_4^0 + e^{\pi_3 - \pi_1} l_6^0}{\gamma} \right),$$

$$\gamma = [(l_1^0 + e^{\pi_2 - \pi_1} l_3^0 + e^{\pi_3 - \pi_1} l_5^0)^2 + (l_2^0 + e^{\pi_2 - \pi_1} l_4^0 + e^{\pi_3 - \pi_1} l_6^0)^2]^{1/2}.$$

Выбор этого управления, как следует из теоремы 16.3 из [1], дает искомую оптимальную минимизирующую функционал J стратегию u.

Пример 2.

Для случая трех целей в трехмерном пространстве рассмотрим следующую игру. Игрок A – это аппарат:

$$\begin{cases} \dot{Y} = A(t)Y + U(t) + F(t), Y \in \mathbb{R}^9, \\ Y(t_0) = Y_0, \end{cases}$$

где

$$Y_1 = Y_4 = Y_7 = y_1, \quad Y_2 = Y_5 = Y_8 = y_2, \quad Y_3 = Y_6 = Y_9 = y_3,$$

а игрок M – это система вида

$$\begin{cases} \dot{Z} = V(t), Z \in \mathbb{R}^9, \\ Z(t_0) = Z_0, \end{cases}$$

где

$$Z_1 = z_{11}, \quad Z_2 = z_{22}, \quad Z_3 = z_{33},$$

$$Z_4 = z_{12}, \quad Z_5 = z_{13}, \quad Z_6 = z_{21}, Z_7 = z_{23}, \quad Z_8 = z_{31} \quad Z_9 = z_{32}$$

и аналогичные равенства для $V(t), v_i(t)$.

Функционал берется в виде

$$I(t, Y, Z) = ||\{Y(T)\}_3 - \{Z(T)\}_3||,$$

$$||x|| = \max_{1 \le i \le 3} |x_i|,$$

где $\{Y\}_3$ означает, что берутся первые три координаты вектора Y. Следовательно, минимаксная задача

$$\min_{u(t)} \sup_{V(t)} \sup_{Y(t),Z(t)} I(t,Y,Z) \tag{11}$$

ставит целью отыскание оптимального управления $\overset{\circ}{u}$, которое к моменту $t=T$ приводит аппарат в точку $\{\overset{\circ}{Y}(T)\}_3$, лежащую в вершине куба

$$\prod_{i=1}^{3}\{x_i : |Z_i(T)-x_i| = j\},$$

где

$$j = ||\{\overset{\circ}{Y}(T)\}_3 - \{Z(T)\}_3||.$$

Рассматриваем простое движение аппарата, т.е. $A(t)=F(t)=0$. Предполагаем также, что «плавание» целей происходит так, что $v_{ii}(t)=0$, т.е. $\{V(t)\}_3=0$. Это означает, что «плавание» координат цели M_i $(i=1,2,3)$ происходит в плоскости с постоянной i-й координатой. Следует сразу сказать, что такое ограничение является обременительным и искусственным.

Для гипотетического рассогласования в позиции (t_*, Y_*, Z_{i*}) имеем ([1, с.131]):

$$\varepsilon^0(t_*,Y_*,Z_{i*}) = \max_{|\vec{l}|=1}\{\rho^{(2)}(t_*,T,\vec{l}) - \rho^{(1)}(t_*,T,\vec{l}) + \vec{l}^{\,tr}(|\{Z(t_*,Z_{i*},T)\}_3 - \{\underset{\circ}{Y}(t_*,Z_{i*},T)\}_3|),$$

где

$$\rho^{(1)}(t_*,T,\vec{l}) = \max_{U(t)\in\mathcal{U}_t}\int_{t_*}^{T}\sum_{i=1}^{3} l_i U_i(t)dt,$$

$$\rho^{(2)}(t_*,T,\vec{l}) = \max_{V(t)\in\mathcal{V}_t}\int_{t_*}^{T}\sum_{i=1}^{3} l_i V_i(t)dt \equiv 0.$$

Аналогично примеру 1

$$\underset{\circ}{Y}(t_*,Z_{i*},T) = Y_*, \qquad \underset{\circ}{Z}(t_*,Z_{i*},T) = \begin{pmatrix} Z_{1*}\\ Z_{2*}\\ Z_{3*}\end{pmatrix} = Z_*$$

– постоянные решения. Далее

$$\rho^{(1)}(t_*,T,\vec{l}) = \max_{U(t)\in\mathcal{U}_t}\sum_{i=1}^{3} l_i \left(\int_{t_*}^{T} U_i(t)dt\right). \tag{12}$$

Так как

$$\rho^{(1)}(t_*,T,\vec{l}) \le \int_{t_*}^{T}\left(\max_{U(t)\in\mathcal{U}_t}\sum_{i=1}^{3} l_i U_i(t)\right)dt,$$

то при ограничении на управление

$$\mathcal{U}_t : U_1^2+U_2^2+U_3^2 \le \delta^2$$

находим, что для всякого $t \in [t_*, T]$

$$\max_{U(t)\in\mathcal{U}_t} \sum_{i=1}^{3} l_i U_i(t) = \delta|\vec{l}|.$$

Следовательно,

$$\rho^{(1)}(t_*, T, \vec{l}) \le \delta|\vec{l}|(T - t_*).$$

Это неравенство показывает, что управление $\tilde{U}_i(t) = \delta l_i/|\vec{l}|$ реализует максимум (12), поскольку

$$\sum_{i=1}^{3} l_i \int\limits_{t_*}^{T} \tilde{U}_i(t)dt = \sum_{i=1}^{3} l_i \int\limits_{t_*}^{T} \delta \frac{l_i}{|\vec{l}|}\, dt = \delta \sum_{i=1}^{3} \frac{l_i^2}{|\vec{l}|}(T - t_*) = \delta|\vec{l}|(T - t_*).$$

Таким образом,

$$\rho^{(1)}(t_*, T, \vec{l}) = \delta|\vec{l}|(T - t_*)$$

и

$$\varepsilon^0(t_*, Y_*, Z_{i*}) = \max_{|\vec{l}|=1}\{\vec{l}^{\,tr}(Z_* - Y_*) - \delta|\vec{l}|(T - t_*)\} =$$

$$-\delta|\vec{l}|(T - t_*) + \max_{|\vec{l}|=1} \sum_{i=1}^{3} l_i(Z_{i*} - Y_{i*}). \tag{13}$$

Откуда вектор $\overset{\circ}{\vec{l}}$, реализующий максимум (13), равен

$$\overset{\circ}{\vec{l}} = \frac{Z_{i*} - Y_{i*}}{|Z_{i*} - Y_{i*}|}$$

и

$$\varepsilon^0(t_*, Y_*, Z_{i*}) = -\delta(T - t_*) + |Z_{i*} - Y_{i*}|.$$

Рассматриваемая задача является регулярной, поскольку вектор $\overset{\circ}{\vec{l}}$, реализующий максимум (13), (при $\varepsilon^0(t_*, Y_*, Z_{i*}) > 0$) единственен. Поэтому экстремальное управление, находимое ниже, будет оптимальным, т.е. реализующим минимаксную задачу (11).

Для экстремального управления U_e имеем

$$[S^{(1)}]^{tr}(t_*)U_e = \max_{U\in\mathcal{U}_{t_*}} [S^{(1)}]^{tr}(t_*)U, \tag{14}$$

где $S^{(1)}(t)$ – 3-мерный вектор-функция, являющаяся решением задачи

$$\begin{cases} \dfrac{dS^{(1)}}{dt} = 0, \\ S^{(1)}(T) = \vec{l}^{\,0}, \end{cases}$$

откуда $S^{(1)}(t) = \vec{l}^{\,0}$ и для (14) получаем

$$[S^{(1)}]^{tr}(t_*)U_e = \max_{U\in\mathcal{U}_{t_*}} \sum_{i=1}^{3} l_i^0 U_i(t_*).$$

Значит, $U_e = \delta\vec{l}^{\,0}$ или $U_{ie} = \delta l_i^0$ $(i = 1, 2, 3)$.

Из примеров 1 и 2 видно, что наличие более сложного управления V_i у целей привело бы к ρ^2, отличному от нуля, и, следовательно, было бы затруднительно вычислять $\varepsilon^0(t_*, Y_*, Z_{i*})$, а также проверять условия определения регулярности рассматриваемой игры.

3. Решение задачи сближения с помощью динамического программирования

Вместо задачи (1),(2), (3) можно решать эквивалентную задачу (5),(6), (7), в которой речь идет о сближении двух объектов в $3N$-мерном пространстве.

Пусть $A(t) = F(t) = 0$ $|U(t)| \le \varepsilon,\ |V(t)| \le \delta$.

Рассмотрим разбиение $0 = t_0 < t_1 < ... < t_m = T$ отрезка $[0, T]$.

Интегрируя (5), имеем

$$Y_{k+1} - Y_k = \int\limits_{t_k}^{t_{k+1}} U(t)dt = \widehat{U}_k(t_{k+1} - t_k),$$

где

$$Y_k = Y(t_k), \quad \widehat{U}_k = \frac{\int\limits_{t_k}^{t_{k+1}} U(t)dt}{t_{k+1} - t_k}.$$

Аналогично для (6)

$$Z_{k+1} = Z_k + \widehat{V}_k(t_{k+1} - t_k),$$

где

$$Z_k = Z(t_k), \quad \widehat{V}_k = \frac{\int\limits_{t_k}^{t_{k+1}} V(t)dt}{t_{k+1} - t_k}.$$

Легко видеть, что задача сближения (5),(6), (7) сводится к дискретной многошаговой задаче

$$\begin{cases} Y_{k+1} = Y_k + \widehat{U}_k(t_{k+1} - t_k), \\ Z_{k+1} = Z_k + \widehat{V}_k(t_{k+1} - t_k), \\ \quad (k = 0, 1, ..., m-1) \\ J = |Y_m - Z_m|. \end{cases} \tag{15}$$

Соответствующая минимаксная задача имеет вид

$$\min_{\widehat{U}_k} \max_{\widehat{V}_k} |Y_m - Z_m|.$$

Задачи (5),(6), (7) и (18) эквивалентны в следующем смысле. Управления $\widehat{U}_k, \widehat{V}_k$ строятся по формулам, приведенным выше из управлений $U(t), V(t)$. Обратно, если даны управления $\widehat{U}_k, \widehat{V}_k$, то поставим им в соответствие управления

$$U(t) = \widehat{U}_k \text{ при } t \in (t_k, t_{k+1}], \quad U(t_0) = \widehat{U}_0,$$

$$V(t) = \widehat{V}_k \text{ при } t \in (t_k, t_{k+1}], \quad V(t_0) = \widehat{V}_0.$$

Ясно, что управлениям $\widehat{U}_k, \widehat{V}_k$ можно поставить в соответствия и другие управления $U(t), V(t)$. Но нетрудно понять, что оптимальной стратегии $\overset{\circ}{U}(t)$ и управлению $V(t)$ в первой игре отвечает оптимальная стратегия $\overset{\circ}{U}_k$ и управление V_k во второй задаче, и наоборот. Соответствующие оптимальные значения функционалов в обеих задачах при этом равны друг другу.

Многошаговая игра (18) решается методом динамического программирования. Вводится функция Беллмана $s_k(X_k)$, $X_k = Y_k - Z_k$, равная минимальному гарантированному значению функционала J для игры, начинающейся в момент t_k при значении X_k. Эта функция удовлетворяет рекуррентному соотношению

$$s_k(X_k) = \min_{\widehat{U}_k} \max_{\widehat{V}_k} s_{k+1}(X_{k+1}) \ (k = 0, ..., m-1) \tag{16}$$

$$s_k(X) = |X|, \tag{17}$$

$$|\widehat{U}_k| \leq \varepsilon, \quad |\widehat{V}_k| \leq \delta$$

и игре

$$\begin{cases} X_{k+1} = X_k + \widehat{U}_k(t_{k+1} - t_k), \\ \quad (k = 0, 1, ..., m-1) \\ J = |X_m|. \end{cases} \tag{18}$$

Известно, единственным решением для (18), (17) будет

$$f_{k,j} = \delta(T - t_j) - \varepsilon(T - t_{j+1}),$$

$$f_{k,m} = |X| - (\varepsilon - \delta)(T - t_k),$$

$$j = k, ..., m-1.$$

Оптимальная стратегия игрока Y имеет вид

$$\overset{\circ}{U}_k = \begin{cases} -\varepsilon \frac{X_k}{|X_k|}, & X_k \neq 0, \\ 0, & X_k = 0, \end{cases}$$

а минимальное гарантированное значение функционала равно

$$\overset{\circ}{J} = \max\{|Y_0 - Z_0| - (\varepsilon - \delta)T, \max_{0 \leq j \leq m-1} [\delta(T - t_j) - \varepsilon(T - t_{j+1}]\}.$$

Следовательно, игроку Y (аппарату) надо двигаться в точку, где в последний раз наблюдался игрок Z с координатами $Z_k = (e^{\pi_1} Z_1, ..., e^{\pi_N} Z_N)$, а при достижении этой точки до следующего момента наблюдения t_{k+1} игрок Y должен ждать – приняв $U = 0$.

Литература

1. Красовский Н.Н. *Игровые задачи о смене движений.* М.: Наука, 1970.