Прикладная математика и механика. Том 33, 1969

ИГРОВАЯ ЗАДАЧА О КОРРЕКЦИИ ДВИЖЕНИЯ

Н. Н. Красовский

(Свердловск)

Рассматривается задача о коррекции движения с целью минимизировать рассогласование избранных координат в заданный момент времени ϑ. Задача осложнена отсутствием полной информации о текущем фазовом состоянии $x\,[t]$ управляемого объекта. Это создает игровую ситуацию, трактуемую как дифференциальная игра двух сближающихся движений. Уравнения движения рассматриваются в линейном приближении. Статья относится к кругу проблем, изученных, например, в работах [1-4].

§ 1. Постановка задачи. Рассмотрим движение $x\,[t]$, описываемое в линейном приближении уравнением

$$dx/dt = A\,(t)\,x + B(t)\,u \tag{1.1}$$

Здесь $x = \{x_i\}$ — n-мерный фазовый вектор объекта, отсчитываемый от заданного движения $x^\circ\,[t] \equiv 0$; r — мерный вектор u описывает управляющие воздействия. Реализации $u\,[t]$ допустимого управления u стеснены условием

$$u\,[t] \in U \tag{1.2}$$

где U есть некоторое ограниченное выпуклое замкнутое множество в r-мерном векторном пространстве. Процесс рассматривается на отрезке $0 \leqslant t \leqslant \vartheta$. Отклонение движения $x[t]$ от заданного движения $x^\circ\,[t] \equiv 0$ оценивается величиной

$$\gamma\,[t] = \|\{x\,[t]\}_m\| \tag{1.3}$$

где символ $\{x\}_m$ обозначает вектор, составленный из первых m координат вектора x, символ $\|q\|$ есть знак евклидовой нормы вектора q.

Цель управления — добиться возможно меньшего значения величины $\gamma\,[\vartheta]$. Задача осложнена невозможностью точного измерения текущих значений $x[t]$. Примем, что информация о фазовых состояниях $x[t]$ $(0 \leqslant t \leqslant \vartheta)$ доставляется [5] наблюдаемым сигналом $w\,[\tau]$ $(0 \leqslant \tau \leqslant t)$, который связан с фазовым вектором $x\,[\tau]$ соотношением

$$w\,[\tau] = P\,(\tau)\,x\,[\tau] + f\,[\tau] \tag{1.4}$$

Здесь $w\,[\tau]$ — k-мерная вектор-функция, $f\,[\tau]$ — погрешность измерения. Реализация $f\,[\tau]$ неизвестна, но известно неравенство

$$\varkappa\,[f\,[\tau], t] \leqslant \nu \qquad (0 \leqslant \tau \leqslant t) \tag{1.5}$$

стесняющее интенсивность ([5], стр. 273) $\varkappa$ возможных реализаций $f\,[\tau]$. Будем предполагать, что величина $\varkappa\,[f, t]$ имеет смысл какой-нибудь нормы вектор-функции $f\,[\tau]$ $(0 \leqslant \tau \leqslant t)$. Кроме того, примем, что оговорена область $\Gamma[0]$ возможных начальных значений фазового вектора $x[0]$ и имеется возможность измерять значения вырабатываемых управляющих воздействий $u[\tau]$.

Задача состоит в определении такого управления u, которое приводило бы движение $x[t]$ к возможно меньшему значению величины $\gamma[\vartheta]$ (1.3) в самых неблагоприятных случаях помехи $f[\tau]$. Уточним задачу.

Пусть в какой-то момент времени t $(0 \leqslant t < \vartheta)$ на деле реализовалось фазовое состояние $x[t]$, которое, однако, в органах, вырабатывающих управление $u[t]$, неизвестно. На основании известного сигнала $\{w[\tau]\ (0 \leqslant \tau \leqslant t),\ u[\tau]\ (0 \leqslant \tau < t)\}$ можно составить представление о той области $\Gamma[t]$ фазового пространства $\{x\}$, в которой может содержаться $x[t]$. (Оценка области $\Gamma[t]$ осуществляется некоторой операцией $\varphi^{\circ}[\{w,\ u\},\ t]$, которая будет охарактеризована ниже.) В соответствии с принципом обратной связи реализуемое в данный момент t управление $u[t]$ следует выбирать исходя из оценки области $\Gamma[t]$. При этом удобно было бы строить управление в форме функциональной зависимости

$$u[t] = u(t,\ \Gamma[t]) \tag{1.6}$$

Однако, чтобы охватить разрывные законы управления u, характерные для игровых задач, примем (см., например, [6]), что найденная к моменту t оценка области $\Gamma[t]$ определяет не одно значение $u[t]$ (1.6), но примем, что она задает целое множество $U(t,\ \Gamma[t])$ возможных значений $u[t]$. Таким образом, закон управления системой (1.1) (стратегия U) будет описываться контингенцией

$$u[t] \in U(t,\ \Gamma[t]) \tag{1.7}$$

Следовательно, чтобы задать какой-то закон управления u системой (1.1) (т. е., чтобы выбрать какую-то стратегию U), надлежит задать систему множеств $U(t,\ \Gamma)$, которые были бы определены для всех t из отрезка $[0,\ \vartheta)$ и для всех областей $\Gamma = \Gamma[t]$, которые могут объявиться по ходу процесса. Итак, стратегию U будем отождествлять с системой множеств $U(t,\ \Gamma)$. При этом движением $x[t]$ $(t_0 \leqslant t \leqslant \vartheta)$ системы (1.1) при выборе некоторой стратегии U будем называть всякую абсолютно непрерывную вектор-функцию $x[t]$, которая будет удовлетворять уравнению (1.1) при $u = u[t]$ из (1.7) при всех почти значениях t из отрезка $[t_0,\ \vartheta]$. Стратегию U будем называть допустимой, если при ее выборе любое исходное состояние $x[t_0] = x_0$ (и $\Gamma[t_0]$) будет порождать движение $x[t]$ $(t_0 \leqslant t \leqslant \vartheta)$ системы (1.1).

Пусть реализовались неизвестное состояние $x[t_0] = x_0$ и известная оценка $\Gamma[t_0]$. Выбранная допустимая стратегия U (1.7) и реализация помехи $f[t]$ $(t \geqslant t_0)$ в (1.4) определят движения (1.1) $x[t]$ $(t \geqslant t_0)$. Обозначим символом $\gamma(\vartheta \mid t_0,\ \Gamma[t_0],\ U,\ f)$ реализующуюся при этом величину $\gamma[\vartheta]$ (1.3). Теперь задачу можно сформулировать следующим образом.

Задача 1.1. Среди допустимых стратегий U надлежит найти оптимальную стратегию U°, которая обеспечивала бы неравенство

$$\gamma(\vartheta \mid t_0,\ \Gamma[t_0],\ U^{\circ},\ f) \leqslant \sup_{x_0} \operatorname*{sur}_{f[t]} \inf_{x[t]} \gamma(\vartheta \mid t_0,\ \Gamma[t_0],\ U,\ f)$$
$$(x_0 \in \Gamma[t_0],\ \varkappa[f[\tau],\ t] \leqslant \nu,\ \tau \leqslant t,\ t_0 \leqslant t \leqslant \vartheta) \tag{1.8}$$

для каждой возможной исходной оценки $\Gamma[t_0]$.

(Единственность движений $x[t]$ $(t_0 \leqslant t \leqslant \vartheta)$, отвечающих исходному состоянию $x[t_0] = x_0$, не предполагается. Этим объясняется значок $\inf\limits_{x[t]}$ в правой части неравенства (1.8). В левой части (1.8) имеется в виду любое движение $x[t]$, отвечающее указанным там данным.)

§ 2. **Операции наблюдения.** Содержание задачи 1.1 зависит от выбора тех операций φ° $[\{w, u\}, t]$, которые определяют оценку областей Γ $[t]$. Можно было бы вообще поставить вопрос о выборе наилучших операций φ° из какого-либо широкого класса. Однако ограничимся довольно узким классом этих операций, но зато таких, которые можно описать более или менее эффективно. Эти операции строятся следующим образом.

По условиям задачи известна априорная оценка области Γ [0] возможных начальных фазовых состояний x [0]. Однако удобнее работать не с этой областью Γ [0], но с областью $G_0^{\cdot}[0]$, которая является отображением области Γ [0] в m-мерное пространство $\{q\}$ в соответствии с равенством

$$q = \{X\ [\vartheta, 0]\ x\}_m \qquad (x \in \Gamma\,[0]) \tag{2.1}$$

где X $[t, t_0]$ — фундаментальная матрица решений для уравнения

$$dx/dt = A\ (t)\ x \tag{2.2}$$

Очевидно, $G_0^{\cdot}[0]$ — это совокупность тех точек $q = \{x\ (\vartheta)\}_m$, в которые система (2.2) переводится к моменту ϑ из состояний $x\ (0) \in \Gamma[0]$. В дальнейшем примем, что область $G_0^{\cdot}[0]$ имеет форму параллелепипеда

$$\alpha_i^\circ \leqslant q_i \leqslant \beta_i^\circ \qquad (i = 1, \ldots, m) \tag{2.3}$$

(Если для исходной области Γ [0] это не так, то можно заключить Γ [0] в большую область Γ_* [0], для которой уже требуемое условие будет выполняться. В то же время высказанное условие облегчает счет.)

Пусть далее к моменту t на отрезке $0 \leqslant \tau \leqslant t$ реализовался сигнал w $[\tau]$ (1.4) и на интервале $0 \leqslant \tau < t$ работало управление u $[\tau]$. Построим операцию φ $[\{w\ [\tau],\ u\ [\tau]\}, t]$, которая восстанавливает вектор

$$q = \{X\ [\vartheta, t]\ x\ [t]\}_m \tag{2.4}$$

по сигналу $\{w\ [\tau],\ u\ [\tau]\}$ и делается это оптимально в том смысле, что каждая из координат q_i вектора q определяется с наименьшей возможной ошибкой ω_i $[t]$. (Построение таких оптимальных операций наблюдения описано в книге ([5], стр. 279—291). Будем предполагать, что условия разрешимости задачи, указанные там, выполнены.) Итак, пусть операция φ $[\{w, u\}, t]$ дает результат $q = q$ $[t]$. Это означает, что в данный момент t действительно реализовавшееся состояние x $[t]$ таково, что из этого состояния система (2.2) может перейти к моменту ϑ лишь в состояние $\{x\ (\vartheta)\}_m = q$ из некоторой области G $[t]$, которая описывается неравенствами

$$|\ q_i - q_i\ [t]\ | \leqslant \omega_i\ [t] \qquad (i=1, \ldots, m) \tag{2.5}$$

Эти неравенства уже дают некоторую оценку области Γ $[t]$. Однако следует еще учесть аналогичные результаты предыдущих оценок областей G $[t_*]$ при $t_* < t$, а также априорную оценку G [0] (2.3). Это осуществляется следующим образом. Предположим сначала, что на отрезке $[0, t]$ никаких оценок области Γ $[t_*]$ для $0 < t_* < t$ не производилось и к моменту t имеется лишь априорная оценка Γ [0]. Поскольку на отрезке $0 \leqslant \tau < t$ работало управление u $[\tau]$, известная априори область Γ [0]

трансформируется в некоторую область Г $[0, t]$ и эта трансформация описывается линейным преобразованием

$$x_* = X[t, 0]x + \int_0^t X[t, \tau]B(\tau)u(\tau)\,d\tau \tag{2.6}$$

$$(x_* \in \Gamma[0, t],\ x \in \Gamma[0])$$

Область Г $[0, t]$ опять удобно оценивать ее образом $G_0[0, t]$ в пространстве $\{q\}$ при линейном преобразовании

$$q_* = \{X[\vartheta, t]\,x_*\}_m, \qquad x_* \in \Gamma[0, t] \tag{2.7}$$

Из (2.1) и (2.7) вытекает, что область $G_0[0, t]$ есть область $G[0]$, смещенная на вектор

$$\Delta^u(0, t) = \left\{\int_0^t X[\vartheta, \tau]B(\tau)u(\tau)\,d\tau\right\}_m$$

Итак, из априорной оценки области Г [0] вытекает, что в момент t возможно лишь такое фазовое состояние $x[t]$, что из этого состояния система (2.2) может перейти к моменту ϑ лишь в состояние $\{x(\vartheta)\}_m = q$ из области $G_0[0, t]$. Но, вспоминая характеристику области $G[t]$, данную выше (см. стр. 338) и учитывая теперь и результат операции $\varphi[\{w, u\}, t]$, можно утверждать, что в момент t реализовалось такое состояние $x[t]$, из которого система (2.2) обязательно перейдет в состояние $q = \{x(\vartheta)\}_m$, лежащее в пересечении областей $G_0[0, t]$ и $G[t]$. Рассуждая аналогичным образом и учитывая результаты $G[t_*]$ операций $\varphi[\{w, u\}, t_*]$ для всех $t_* \leqslant t$, придем к выводу, что в момент t реализуется такое состояние, из которого система (2.2) может перейти лишь в состояние $q = \{x(\vartheta)\}_m$, лежащее в области

$$G^\circ[t] = \bigcap_{0 \leqslant t_* \leqslant t} G[t_*, t] \cap G_0[0, t] \tag{2.8}$$

где $G[t_*, t]$ — область в пространстве $\{q\}$, получающаяся из области $G[t_*]$ смещением на вектор

$$\Delta^u(t_*, t) = \left\{\int_{t_*}^t X[\vartheta, t_*]B(\tau)u(\tau)\,d\tau\right\}_m \tag{2.9}$$

Построение области $G^\circ[t]$ и будем именовать операцией $\varphi^\circ[\{w, u\}, t]$. Области $G^\circ[t]$ дадут нужную характеристику областей Г $[t]$. В соответствии с этим множества $U(t, \Gamma[t])$, определяющие стратегию U, будем в дальнейшем сопоставлять областям $G^\circ[t]$, т. е. контингенции (1.7) будут задаваться в форме

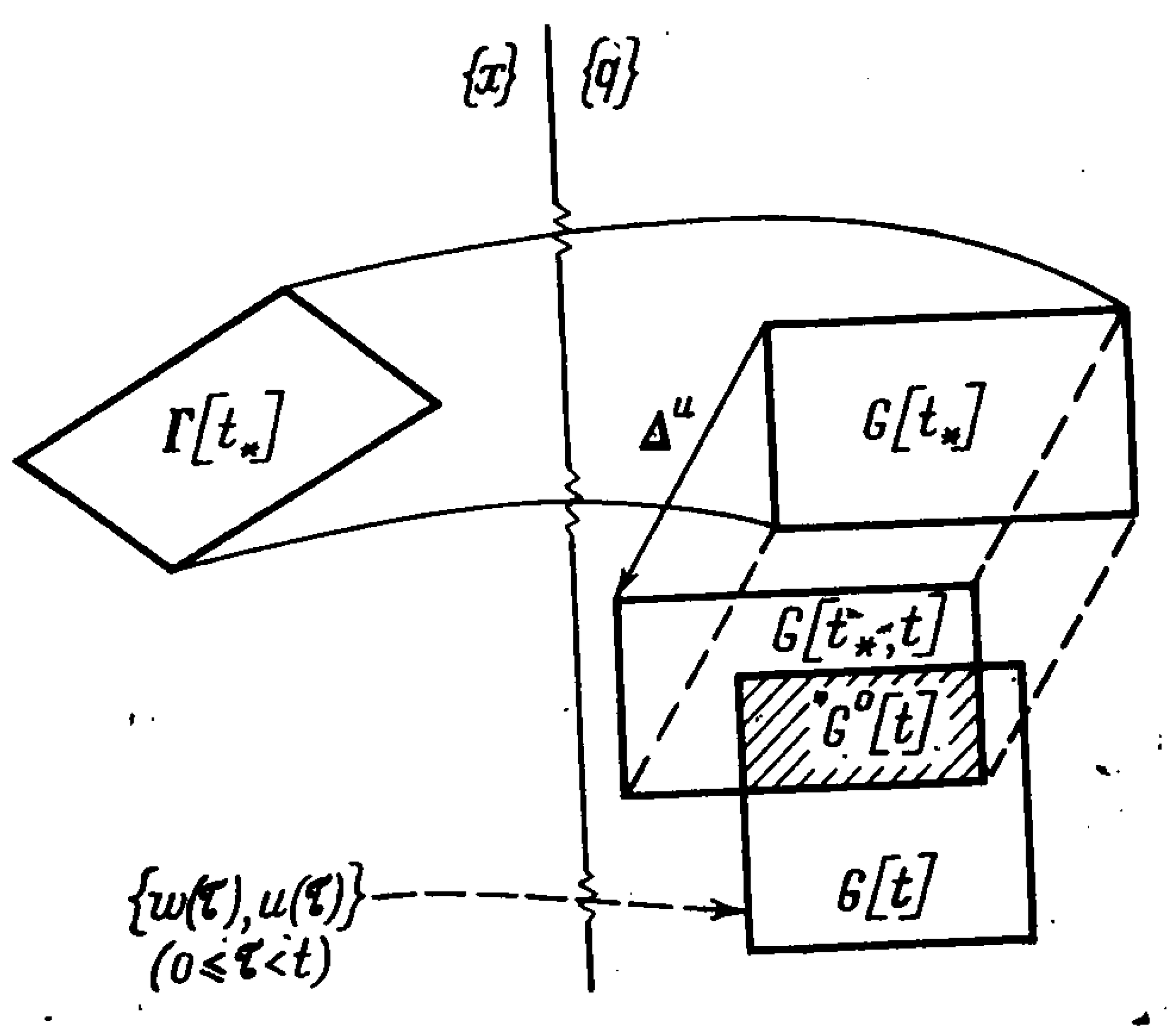
$$u[t] \in U(t, G^\circ[t]) \qquad (2.10$$

Заметим в заключение этого параграфа, что области $G^\circ[t]$ суть параллелепипеды

$$\alpha_i\ [t] \leqslant q_i \leqslant \beta_i\ [t] \qquad (i=1,\ldots n) \tag{2.11}$$

ибо параллелепипедами будут все области $G\ [t_*, t]$, которые согласно (2.8) определяют область $G^\circ\ [t]$ (фиг. 1).

§ 3. Задача коррекции как игра сближения. Рассматриваемую задачу 1.1 можно трактовать как игру на минимакс рассогласования координат двух управляемых движений. Поясним это. Пусть в некоторый момент времени t стала известна область $G^\circ\ [t]$ (2.8), причем реализовалось неизвестное состояние $x\ [t]$. Обозначим символом $x_*\ [t]$ фазовое состояние из области $\Gamma\ [t]$, удовлетворяющее условию



Фиг. 1

$$\{X[\vartheta,t]\ x_*\ [t]\}_m = q_*\ [t] \tag{3.1}$$

где $q_*\ [t]$ — центр области $G^\circ\ [t]$, т. е.

$$q_{*i}\,[t] = {}^1/_2\,(\beta_i\,[t] + \alpha_i\,[t]) \qquad (i=1,\ldots,m) \tag{3.2}$$

Запишем уравнение движения для реализации $x_*\ [t]$ в виде

$$dx_*/dt = A\ (t)\ x_* + B\ (t)\ u - v \tag{3.3}$$

где v — некоторое фиктивное управление. Примем, что в рассматриваемой системе (1.1) в момент t реализуется известное состояние $x = x_*[t]$, но под действием «управления» v это состояние может смещаться неизвестным образом, оставаясь, однако, при $\tau + 0$ в пределах области $\Gamma\ [\tau]$ $(\tau \geqslant t)$. (Чтобы получить, например, действительно реализующееся состояние $x\ [t] = x_*\ [t+0]$, достаточно положить

$$v\ [\tau] = (x_*[t] - x\ [t])\,\delta\ (\tau - t)$$

где $\delta\ (\tau)$ — дельта-функция.) Далее вектор $x_*\ [t]$ представим в форме разности

$$x_*\ [t] = y\ [t] - z\ [t] \tag{3.4}$$

полагая, что движение $y\ [t]$ описывается уравнением

$$dy/dt = A\ (t)\ y + B\ (t)\ u \tag{3.5}$$

а движение $z\ [t]$ — уравнением

$$dz/dt = A\ (t)\ z + v \tag{3.6}$$

Если рассогласование между движениями $y\ [t]$ и $z\ [t]$ оценивать величиной

$$\gamma_*\ [t] = \|\ \{y\ [t] - z\ [t+0]\}_m\,\| \tag{3.7}$$

то исходную задачу 1.1 можно трактовать как задачу о выборе оптимального управления u, которое обеспечивает «преследователю» $y\ [t]$ (3.5) наилучшее возможное сближение $\gamma_*\ [\vartheta]$ с «преследуемым» $z\ [t]$ (3.6) при самом неблагоприятном поведении последнего. При этом управление u стеснено условием (1.2), а управление v стеснено тем ограничением, что вызываемые им изменения вектора $z\ [t]$ все время сохраняют вектор $x_*\ [t+0] = y\ [t] - z\ [t+0]$ в областях $\Gamma\ [t]$ $(0 \leqslant t \leqslant \vartheta)$. Итак, исходная задача действительно сводится к некоторой игровой задаче о сближении двух движений.

§ 4. **Экстремальная конструкция.** Будем решать задачу о минимаксе величины $\gamma_*\ [\vartheta]$ (3.7), опираясь на правило экстремального прицеливания [5,6]. Для этого надлежит рассмотреть области достижимости (см. определение, например в [5], стр. 331) $G^{(1)}\ (t, \vartheta, y)$ и $G^{(2)}\ (t, \vartheta, z)$ для движений y и z.

Область достижимости $G^{(1)}\ (t, \vartheta, y)$ для движения y (3.5) из состояния $y\ [t] = y$ к моменту ϑ при ограничении (1.2) строится обычным образом и описывается неравенством (см. например, [5])

$$\rho^{(1)}\ (t,\ l) + l' \{X\ [\vartheta,\ t]\ y\}_m - l'q \geqslant 0 \tag{4.1}$$

которому удовлетворяют точки q из области $G^{(1)}\ (t, \vartheta, y)$ (и только такие точки) при всех значениях m-мерного вектора l. Здесь

$$\rho^{(1)}(t,\ l) = \max_{u(\tau)\in U} \left[l' \left\{ \int_t^{v} X\ [\vartheta,\ \tau]\ B\ (\tau)\ u\ (\tau)\ d\tau \right\}_m \right] \tag{4.2}$$

верхний индекс штрих означает транспонирование.

Построение областей $G^{(2)}\ (t, \vartheta, z)$ определяется областями $G^{\circ}\ [t]$ (2.8), (2.11). В самом деле, пусть реализовалось состояние $z = z\ [t]$. Согласно (3.4), $z\ [t] = y\ [t] - x_*\ [t]$. Если бы при $t \leqslant \tau \leqslant \vartheta$ управление $v\ (\tau)$ было нулем, то система (3.6) к моменту ϑ оказалась бы в состоянии (см. (3.1))

$$\begin{aligned} \{z\ (\vartheta)\}_m = q = \{X\ [\vartheta,\ t]\ (y[t] - x_*[t])\}_m &= \\ = \{X\ [\vartheta,\ t]\ y\ [t]\}_m - q_*[t] & \end{aligned} \tag{4.3}$$

Однако под действием управления $v\ (\tau)$ система (3.6) (по смыслу этого управления $v\ (\tau)$) может оказаться в любом состоянии $\{z\ (\vartheta + 0)\}_m = q$, которое удовлетворяет неравенствам

$$|\ q_i - \{X[\vartheta,\ t]\ y\ [t]\}_{mi} + q_{*i}\ [t]\ | \leqslant {}^1/_2\ (\beta_i\ [t] - \alpha_i\ [t]) \tag{4.4}$$
$$(i = 1, \ldots, m)$$

Этими неравенствами и определим область $G^{(2)}\ (t, \vartheta, z\ [t])$. Важно заметить, что реализующиеся по ходу процесса оценки (4.4) определяют области $G^{(2)}\ (t, \vartheta, z\ [t])$ так, что выполняются следующие основные условия:

$$G^{(2)}\ (t_*, \vartheta, z\ [t_*]) \supset G^{(2)}\ (t, \vartheta, z\ [t]) \quad (t_* < t) \tag{4.5}$$

$$\{x\ [\vartheta + 0]\}_m = \{y[\vartheta]\}_m - \{z\ [\vartheta + 0]\}_m, \quad \{z\ [\vartheta + 0]\}_m \in G^{(2)}(\vartheta, \vartheta, z\ [\vartheta]) \tag{4.6}$$

Соотношения (4.5) и (4.6) и позволяют трактовать $G^{(2)}(t, \vartheta, z)$ как область достижимости для движения z, к которой применимы рассуждения [5,6], характерные для схемы экстремального прицеливания.

Справедливость (4.6) следует сразу из определения области $G^{(2)}$ (4.4) по смыслу величин α_i $[t]$, β_i $[t]$ (2.11) и вектора q_* $[t]$ (3.1). Проверим справедливость вложения (4.5). Из (4.4) следует, что $G^{(2)}$ $(t_*, \vartheta, z$ $[t_*])$ есть область, симметричная области G° $[t_*]$ относительно точки $q = 0$ и смещенная еще поступательно на вектор

$$p\ [t_*] = \{X\ [\vartheta,\ t_*]\ y\ [t_*]\}_m \tag{4.7}$$

а $G^{(2)}$ $(t, \vartheta, z$ $[t])$ $(t > t_*)$ есть область, симметричная области $G^\circ[t]$ относительно точки $q = 0$ и смещенная еще поступательно на вектор

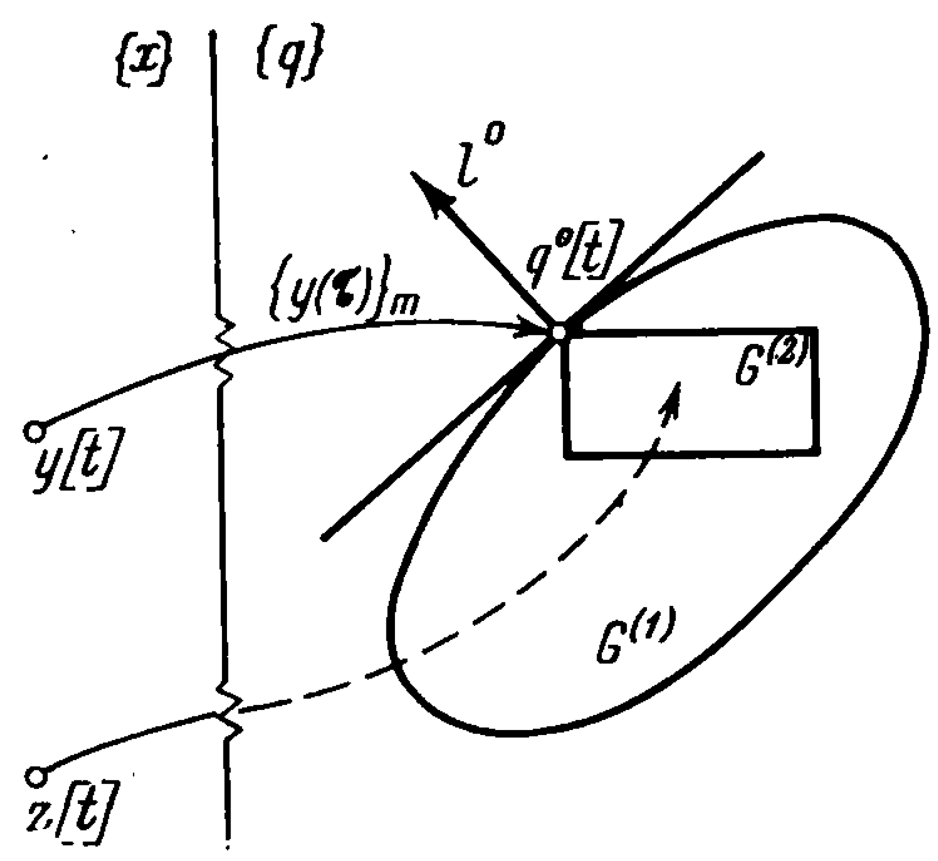
$$p\,[t] = \{X\,[\vartheta,\,t]\,y\,[t]\}_m = \left\{X\,[\vartheta,\,t_*]\,y\,[t_*] + \int\limits_{t_*}^{t} X\,[\vartheta,\,\tau]\,B\,(\tau)\,u\,(\tau)\,d\tau\right\}_m \tag{4.8}$$

Но согласно (2.8) при $t > t_*$ область G° $[t]$ составляет часть области G° $[t_*]$, смещенной на вектор Δ^u (t_*, t) (2.9). Отсюда согласно (4.7), (4.8) и вытекает справедливость соотношения (4.5).

Опишем теперь схему экстремального прицеливания в приложении к данному случаю. Пусть в момент t реализовались значения $y\,[t]$ и z $[t]$, информацию о которых несут известные параметры $q_*[t]$ (3.1) и α_i $[t]$, β_i $[t]$ (2.11) области G° $[t]$. Полагая на время реализовавшиеся векторы y $[t]$ и z $[t]$ известными, построим область достижимости $G^{(2)}(t, \vartheta, z[t])$ и наименьшую содержащую ее ε-окрестность $G^{(1)}$ $(t, \vartheta, y$ $[t], \varepsilon^\circ)$ области достижимости $G^{(1)}$ $(t, \vartheta, y$ $[t])$. Если окажется для $0 \leqslant t < \vartheta$, что всякий раз при $\varepsilon^\circ > 0$ границы этих областей будут пересекаться лишь по множествам Q° $[t]$, каждое из которых будет лежать целиком в единственной (при каждом t) гиперплоскости

$$l^{\circ\prime}\ [t]\ q = \rho\ [t] \tag{4.9}$$

то будет иметь место грубый случай [6]. Тогда оптимальное управление $u = u$ $[t]$, разрешающее задачу, будет формироваться из условия прицеливания [6] движения y $[t]$ в любую точку q° $[t] \in Q^\circ$ $[t]$ (фиг. 2).

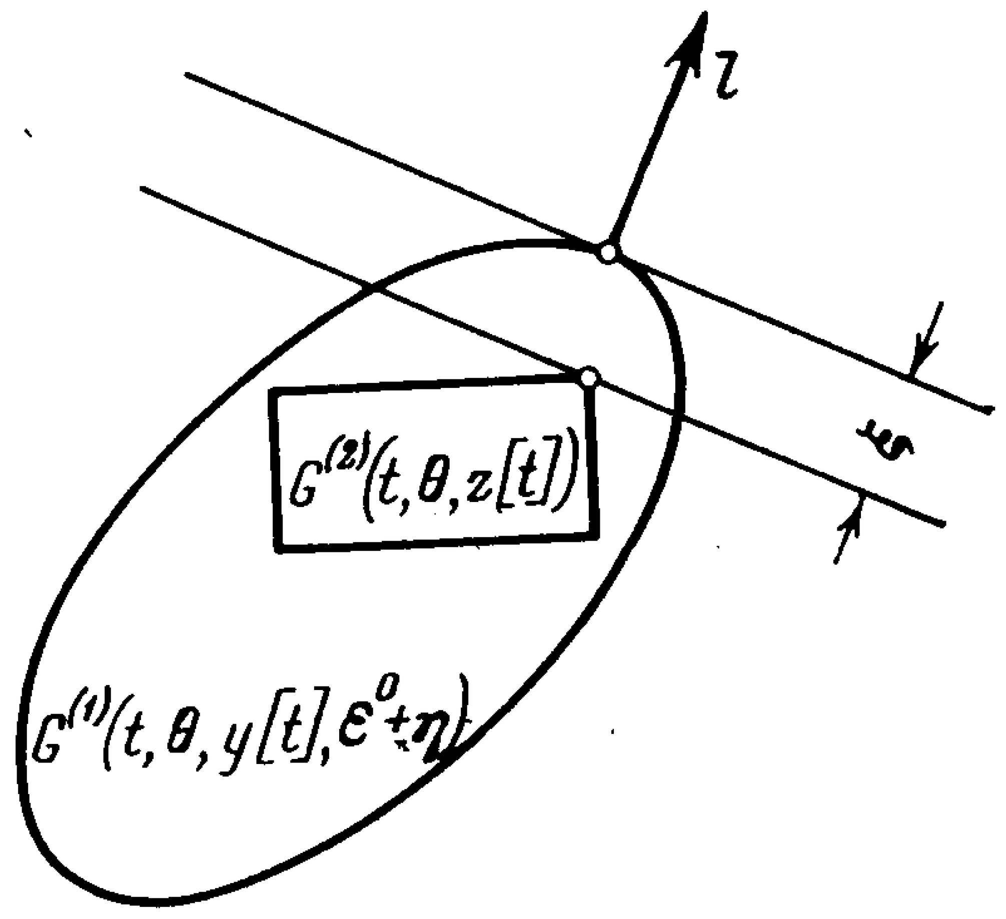


Фиг. 2

Важно заметить, что это управление будет определяться известными величинами α_i $[t]$, β_i $[t]$. В самом деле, прицеливающее управление u $[t] = u_* \in U$ определяется здесь условием максимума

$$l^{\circ\prime}\,\{X\,[\vartheta,\,t]\,B\,(t)\,u_*\}_m = \max_{u \in U}\,[l^{\circ\prime}\,\{X\,[\vartheta,\,t]\,B\,(t)\,u\}_m] \tag{4.10}$$

причем единичный m-мерный вектор l° есть внешняя нормаль к области $G^{(1)}$ $(t, \vartheta, y$ $[t]$, ε° $[t])$ в точке q° $[t]$, где граница этой области соприкасается с областью $G^{(2)}(t, \vartheta, z[t])$. Но этот вектор $l^\circ[t]$ определяется лишь взаимным расположением данных областей одна относительно другой. Вследствие однотипности систем (3.5) и (3.6) и вследствие равенства $x_*[t] = y[t] - z[t]$ это взаимное расположение $G^{(1)}$ и $G^{(2)}$ определяется целиком параметрами $\alpha_i[t]$, $\beta_i[t]$ и $\rho^{(1)}(l, t)$ областей $G^{(1)}$ и $G^{(2)}$, которые, следовательно, и определяют вектор l°, а с ним — и оптимальное управление $u^\circ[t] = u_*$ (4.10). (Справедливость этого утверждения видна и непосредственно из соотношений (4.1), (4.4), которые определяют области $G^{(1)}$ и $G^{(2)}$.)



Фиг. 3

Итак, в грубом случае согласно [6] задача 1.1 имеет решение, причем оптимальная стратегия U° задается множествами $U^\circ(t, \alpha_1[t], \ldots, \alpha_m[t], \beta_1[t], \ldots, \beta_m[t])$, складывающимися из всех векторов $u_* \in U$, которые удовлетворяют условию (4.10). Эта стратегия обеспечивает результат коррекции $\gamma[\vartheta]$, не худший, чем $\varepsilon^\circ[0]$, т. е.

$$\gamma[\vartheta] \leqslant \varepsilon^\circ[0] \tag{4.11}$$

Следует заметить, что грубый случай для рассматриваемой игровой задачи сближения, которая получается из исходной задачи 1.1 о коррекции, составляет, пожалуй, исключение. Поэтому в следующем параграфе рассмотрим несколько более естественный регуляризируемый критический случай.

§ 5. Критический случай. Пусть к некоторому моменту $t = t_0 \geqslant 0$ реализуется такая ситуация, в которой граница $H^{(1)}(t_0, \vartheta, y[t_0]; \varepsilon^\circ[t_0])$ $(\varepsilon^\circ > 0)$ области $G^{(1)}(t_0, \vartheta, y[t_0]; \varepsilon^\circ[t_0])$ имеет с областью $G^{(2)}(t_0, \vartheta, z[t_0])$ пересечение $Q^\circ[t_0]$, не укладывающееся на единственную гиперплоскость вида (4.9). Построим область $G^{(1)}(t_0, \vartheta, y[t_0]; \varepsilon^\circ[t_0] + \eta)$, где $\eta > 0$ — малая постоянная. По смыслу величины ε° заключаем, что область $G^{(2)}(t_0, \vartheta, z[t_0])$ лежит строго внутри области $G^{(1)}(t_0, \vartheta, y[t_0]; \varepsilon^\circ[t_0] + \eta)$. Выбор управления $u[t]$ при $t \geqslant t_0$ будет подчинен стремлению сохранить вложение

$$G^{(2)}(t, \vartheta, z[t]; \varepsilon[t]) \subset G^{(1)}(t, \vartheta, y[t]; \varepsilon^\circ[t_0] + \eta[t]) \tag{5.1}$$

при $\varepsilon[t] > 0$ и при $\eta[t] = \eta = \text{const}$ или, в крайнем случае, при условии предельно медленного возрастания переменной $\eta[t]$ с ростом времени t. Искомое управление u будем конструировать следующим образом.

Пусть в какой-то момент $t = t_* \geqslant t_0$ сложилось состояние $\{t_*, y[t_*] = y_*, z[t_*] = z_*\}$ (характеризуемое доступными измерению величинами $q_*[t_*]$, $\alpha_i[t_*]$, $\beta_i[t_*]$ $(i = 1, \ldots, m)$). Предполагаем, что при этом выполнено условие (5.1) с каким-то реализовавшимся к данному моменту времени $t = t_*$ значением $\eta[t_*]$. Выберем какой-нибудь m-вектор l внешней нормали к области $G^{(1)}(t_*, \vartheta, y_*;\ \varepsilon^\circ [t_0] + \eta [t_*])$. Обозначим символом $\xi (l, t_*, y_*, G^{(2)}(t_*, \vartheta, z_*)) = \xi (l, t_* \{\alpha_i [t_*], \beta_i [t_*]\})$ — расстояние между гиперплоскостями $L^{(1)}(l)$ и $L^{(2)}(l)$, которые ортогональны к l и касаются областей $G^{(1)}$ и $G^{(2)}$ (фиг. 3).

Составим функцию

$$\lambda (t, y, G^{(2)}(t, \vartheta, z)) = \int\limits_{\|l\| \leqslant 1} \varphi [\zeta (l, t, y, G^{(2)})]\, d\zeta \tag{5.2}$$

где $\varphi [\xi] \geqslant 0$ — гладкая, монотонная при $\xi > 0$ функция порядка ξ^{-n} при $\xi \to 0$; интеграл берется по сфере $\|l\| \leqslant 1$. Предположим, что при $t_* \leqslant t \leqslant t_* + \Delta t$ величина $y(t)$ изменяется в соответствии с уравнением (3.5), где $u = u(t)$ — некоторая измеримая функция, непрерывная справа в точке $t = t_*$; область $G^{(2)}(t, \vartheta, z(t))$ пусть изменяется в соответствии с результатами измерения сигнала $w(\tau)$ (см. выше § 4); величину $\eta(t)$ будем полагать на отрезке $[t_*, t_* + \Delta t]$ постоянной. Тогда, рассматривая изменение функции $\lambda(t) = \lambda(t, y(t), G^{(2)}(t, \vartheta, z(t)))$ на малом отрезке $[t_*, t_* + \Delta t]$, можно стандартными в теории принципа максимума [7] рассуждениями вывести оценку

$$\left(\frac{d\lambda}{dt}\right)_+^{(b)} \leqslant -s'(t_*, G^\circ [t_*])\, B(t_*)\, u(t_*) + \chi \tag{5.3}$$

характеризующую влияние управления u на изменение функции $\lambda(t)$.

Здесь символ $(d\lambda / dt)_+^{(b)}$ означает правое верхнее производное число функции $\lambda(t)$ при $t = t_*$; n-мерный вектор s определен равенством

$$s(t, G^\circ [t]) = \int\limits_{\|l\| \leqslant 1} \frac{d\varphi}{d\xi} \psi (t_*, l)\, d\zeta \tag{5.4}$$

причем n — вектор-функция $\varphi(t, l)$ будет решением уравнения

$$\frac{d\psi}{dt} = -A'(t)\psi \tag{5.5}$$

при краевом условии

$$\psi' (\vartheta, l)) = \{l', 0\}$$

Слагаемое χ в правой части (5.3) в дальнейшем при построении управления $u[t_*]$ роли играть не будет. Выражение χ, впрочем, важно знать для оценки величины $(d\lambda / dt_+^{(b)})$. Однако в данной статье автор не может дать подобной априорной оценки, которая была бы эффективна в более или менее общем случае. Поэтому изучать член χ здесь не будем.

Первое слагаемое в правой части (5.3) достигает минимума при управлении $u(t_*) = u_* \in U$, которое удовлетворяет следующему рандомизированному условию максимума:

$$s'(t_*, G^\circ [t_*])\, B(t_*)\, u_* = \max_{u \in U} s'(t_*, G^\circ [t_*])\, B(t_*)\, u \tag{5.6}$$

Стратегию U°, которая определит управление u, зададим множествами $U^\circ (t_*, G^\circ [t_*])$, состоящими из всех тех векторов u_* из U, которые удовлетворяют условию (5.6). При достаточно широких условиях эта стратегия U° будет допустимой и, более того, оценка вида (5.3) имеет место вдоль действительных реализаций движений $y\ [t]$, $z\ [t]$ при $u\ [t_*] \in U^\circ\ (t_*,\ G^\circ\ [t_*])$. Если при этом все время при $t \geqslant t_0$ для $\eta\ [t] = \eta\ [t_0] = \mathrm{const}$ окажется справедливым неравенство

$$\left(\frac{d\lambda}{dt}\right)_+^{(b)} \leqslant -s'(t,\ G^\circ[t])\,B(t)\,u[t] + \chi < \mu \tag{5.7}$$

$$(u\ [t] \in U^\circ\ (t,\ G^\circ\ [t]),\ \mu = \mathrm{const})$$

то реализация $\lambda\ [t]$ остается ограниченной для всех $t \geqslant t_0$. Однако по построению функции λ (5.2) отсюда вытекает, что для всех $t \geqslant t_0$ будет справедливо вложение (5.1) при $\eta\ [t] = \eta\ [t_0]$. Но при $t = \vartheta$ это вложение означает, что

$$\gamma\ [\vartheta] \leqslant \varepsilon^\circ\ [t_0] + \eta\ [t_0] \tag{5.8}$$

Случай, когда выполняется условие (5.7), назовем регуляризируемым критическим случаем. Итак, в регуляризируемом критическом случае строится стратегия U°, которая обеспечивает результат коррекции $\gamma\ [\vartheta]$, не худший, чем (5.8).

Если выполнить неравенство (5.7) не удается, то можно остановиться на той же самой стратегии U°, которая была описана выше в этом параграфе, однако для сохранения вложения (5.1) тогда надлежит изменять величину $\eta[t]$. При изменении величины $\eta\ [t]$ в правой части неравенства (5.7) появляется дополнительное слагаемое ν вида

$$\nu = \left(\frac{d\eta}{dt}\right)_{(H)+} \int\limits_{\|l\| \leqslant 1} \frac{d\varphi}{d\xi}\, d\zeta \tag{5.9}$$

и величину правого нижнего производного числа $(d\eta\ /\ dt)_{(H)+}$ можно выбирать теперь из того условия, чтобы удовлетворить, например, неравенству $(d\lambda\ /\ dt)_+^{(b)} \leqslant 0$. Если оценка нужного значения $(d\eta\ /\ dt)_{(H)+}$ по реализующейся позиции $\{t,\ G\ [t]\}$ затруднительна из-за трудностей с априорной оценкой χ, то возможен подбор значений $(d\eta\ /\ dt)_{(H)+}$ в каждый момент времени t, например, из условий

$$\left(\frac{d\lambda}{dt}\right)_+^{(b)} + \left(\frac{d\lambda}{dt}\right)_{(H)-} \leqslant 0 \tag{5.10}$$

где левое нижнее производное число $(d\lambda/\ dt)_{(H)-}$ функции $\lambda\ (t)$ извлекается из известной уже к данному моменту реализации $\lambda\ (t)$ в прошлом. (Разумеется, реально данный способ управления должен осуществляться в дискретной схеме.) Следует заметить, что последний способ управления требует оценки его устойчивости, ибо при нем производится операция дифференцирования.

§ 6. Сравнение с минимаксной стратегией. Задача о коррекции, аналогичная той, какая послужила исходным пунктом для задачи 1.1, рассмотрена в статье [4]. Однако способ коррекции, описанный в статье [4], отличается от того способа управления, который описан в предлагаемой статье и который отвечает задаче 1.1. (В статье [4] описана дискретная по времени схема учета результатов измерения сигнала $w\ (\tau)$, но это не существенно, так как та же схема переносится и на непрерывный случай.) Суть различия в способах управления состоит в следующем. Если говорить на языке систем (3.5), (3.6), то можно сказать, что управление $u\,[t]$ в статье [4] выбиралось в текущий момент t из условия прицеливания движения $y\,[t]$ в ту точку $q = q_0$ из области достижимости $G^{(1)}(t, \vartheta, y\,[t])$, в которой достигается минимакс

$$\min_q \max_p \| p - q \| = \varepsilon_0\,[t] \qquad (6.1)$$

по всем q из $G^{(1)}(t, \vartheta, y\,[t])$ и p из $G^{(2)}(t, \vartheta, z\,[t])$ и при этом обязательно обеспечивался результат коррекции

$$\gamma\,[\vartheta] \leqslant \varepsilon_0\,[t] \qquad (6.2)$$

В предлагаемой статье управление $u\,[t]$ выбиралось из условия прицеливания движения $y\,[t]$ в ту точку $q = q_0$ из области $G^{(1)}(t, \vartheta, y\,[t])$, которая отвечает максимину

$$\max_p \min_q \| p - q \| = \varepsilon^\circ\,[t] \qquad (6.3)$$

по всем q из $G^{(1)}(t, \vartheta, y\,[t])$ и p из $G^{(2)}(t, \vartheta, z\,[t])$. Так как $\varepsilon^\circ\,[t] \leqslant \varepsilon_0[t]$, то второй способ управления в грубом случае или хотя бы в регуляризуемом случае, когда выполнены неравенства (4.11) или (5.8), должен давать, вообще говоря, лучший результат. Следует, впрочем, иметь в виду, что неравенство (6.2) обеспечивается всегда, а неравенства (4.11) или (5.8) лишь при сделанных выше оговорках.

Поступила 28 X 1968

Уральский государственный университет

ЛИТЕРАТУРА

1. Рясин В. А. Оптимальная одноразовая коррекция в модельной задаче. Теория вероятностей и ее применение, 1966, т. 11, вып. 4.
2. Шелементьев Г. С. Об оптимальном сочетании управления и наблюдения. ПММ, 1968, т. 32, вып. 2.
3. Черноусько Ф. Л. Минимаксная задача одноразовой коррекции при погрешностях измерений. ПММ, 1968, т. 32, вып. 4.
4. Шелементьев Г. С. Об одной задаче коррекции движения. ПММ, 1969, т. 33, вып. 2.
5. Красовский Н. Н. Теория управления движением. М., «Наука», 1968.
6. Красовский Н. Н. О дифференциальной игре на сближение. Докл. АН СССР, 1968, т. 182, № 6.
7. Понтрягин Л. С., Болтянский В. Г., Гамкрелидзе Р. В., Мищенко Е. Ф. Математическая теория оптимальных процессов. М., Физматгиз, 1961.