Прикладная математика и механика. Том 36, 1972

УДК 62—50

ПРИБЛИЖЕННЫЙ МЕТОД СИНТЕЗА ОПТИМАЛЬНОГО УПРАВЛЕНИЯ СИСТЕМОЙ, ПОДВЕРЖЕННОЙ СЛУЧАЙНЫМ ВОЗМУЩЕНИЯМ

А. И. Солянник, Ф. Л. Черноусько

(Киев, Москва)

Предлагается приближенный метод построения синтеза оптимального управления динамической системой при наличии внешних случайных возмущений и ошибок измерений. Поставленная задача синтеза сводится, как известно, к решению нелинейного уравнения в частных производных параболического типа (уравнения Беллмана), точные решения которого известны лишь в немногих случаях. Предполагается, что либо внешние возмущения, действующие на систему, достаточно малы, либо погрешности измерений велики. При этих условиях уравнение Беллмана содержит малый параметр при старших производных, и решение строится путем разложения по малому параметру. Показано, что приближенный синтез оптимального управления для системы с возмущениями может быть построен в явном виде, если известно решение соответствующей задачи для системы без возмущений. Доказаны оценки погрешности метода. Приведены примеры.

1. Постановка задачи. Рассмотрим задачу оптимального управления конечным состоянием системы, движение которой описывается уравнением

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

Здесь t — время, x — n-мерный вектор фазовых координат, u — m-мерный вектор управляющих функций, $b(t, u)$ — заданная вектор-функция, $\xi(t)$ — s-мерная вектор-функция случайных возмущений, действующих на систему. Матрицы $A(t)$, $C(t)$ — заданные функции времени и имеют размерность $n \times n$, $n \times s$ соответственно. Предполагается, что вектор $\xi(t)$ в каждый момент t распределен по нормальному закону с нулевым математическим ожиданием и корреляционной матрицей $G(t)$ и не коррелирован по времени, т. е. является белым шумом. Интенсивность внешних возмущений далее предполагается малой, т. е. $G = \varepsilon G_*$, где ε — малый параметр ($0 \leqslant \varepsilon \ll 1$), а $G_*(t)$ — матрица с ограниченными элементами. На вектор управляющих функций наложено ограничение

$$u(t) \in U \tag{1.2}$$

где U — заданное ограниченное замкнутое множество. Предполагается, что в любой момент t возможно точное измерение фазового вектора системы $x(t)$.

Требуется найти управление u в зависимости от времени и текущего фазового вектора $x(t)$, минимизирующее при условии (1.2) математическое ожидание скалярной функции $F(x(T))$ фазовых координат в конечный момент времени T.

Введем функцию Беллмана $S(t, x)$, равную оптимальному значению минимизируемого функционала при условии, что процесс начинается в момент t при фазовом векторе $x(t) = x$. Тогда поставленная задача оптимального управления сводится, как известно [1], к решению нелинейного параболического уравнения

$$S_t - H(S_x, x, t) + {}^1/_2\varepsilon\,\mathrm{Sp}(LS_{xx}) = 0 \tag{1.3}$$

с начальным условием

$$S(T, x) = F(x) \tag{1.4}$$

Здесь S_t — частная производная по времени, S_x — вектор первых частных производных, S_{xx} — матрица вторых частных производных функции S по компонентам вектора x, Sp означает след матрицы, и, кроме того, введены обозначения

$$H(S_x, x, t) = -\min\nolimits_{u\in U}(S_x, Ax + b(t, u)) \tag{1.5}$$
$$L(t) = CG_*C'$$

где скобками обозначено скалярное произведение векторов.

Будем предполагать для общности, что момент T окончания процесса не фиксирован и определяется условием $h(T, x(T)) = 0$, а минимизируется математическое ожидание значения функции $F(T, x(T))$ в конце процесса. Здесь $h(t, x)$, $F(t, x)$ — заданные функции, причем в случае фиксированного T имеем $h = T - t$. При сделанных предположениях функция S удовлетворяет вместо (1.4) условию

$$S(t, x) = F(t, x) \quad \text{при } h(t, x) = 0 \tag{1.6}$$

Для определенности будем считать, что $h > 0$ в начале процесса. Решение задачи Коши (1.3), (1.6) нужно построить в области $h \geqslant 0$.

Таким образом, задача синтеза оптимального управления в случае малой интенсивности внешних возмущений сводится к решению задачи Коши для нелинейного параболического уравнения (1.3) с малым параметром при старших производных и с начальными условиями (1.4) или (1.6). Оптимальное управление определяется после нахождения функции S из условия достижения минимума в (1.5).

Отметим, что точно к такой же математической задаче приводится и задача оптимального управления системой (1.1), (1.2) при наличии больших погрешностей измерений.

Пусть процесс измерения задан уравнением

$$y(t) = Q(t)\,x(t) + \eta(t) \tag{1.7}$$

Здесь y — l-мерный вектор результатов измерений, $\eta(t)$ — вектор-функция ошибок измерений, которая так же, как и $\xi(t)$, является белым шумом с корреляционной матрицей $B(t)$.

Пусть в начальный момент времени t_0 случайный вектор $x(t_0)$ распределен по нормальному закону с математическим ожиданием x_0 и корреляционной матрицей D_0.

Тогда при обработке результатов измерений по методу максимума правдоподобия в силу линейности уравнений (1.1), (1.7) апостериорное распределение вектора $x(t)$ нормально, а уравнения, описывающие изменение вектора математического ожидания $z(t)$ и корреляционной матрицы $D(t)$ этого распределения, имеют следующий вид [2]:

$$\begin{gathered} dz / dt = Az + b + DQ'B^{-1}(y - Qz), \quad z(t_0) = x_0 \\ dD / dt = AD + DA' - DQ'B^{-1}QD + CGC', \; D(t_0) = D_0 \end{gathered} \tag{1.8}$$

Здесь опущена явная зависимость функций от времени, штрих означает транспонированную, а степень минус единица — обратную матрицу. Так как уравнение (1.8) для D не зависит от $b(t, u)$, то матрица $D(t)$ может быть определена заранее вне зависимости от результатов измерений и закона управления. Далее $D(t)$ считаем известной функцией времени. Снова рассмотрим задачу определения оптимального управления в зависимости от времени и данных измерений, минимизирующего математическое ожидание функции $F(x(T))$ в конечный момент T. Введем функцию Беллмана $S(t, z)$, равную оптимальному значению минимизируемого функционала при условии, что процесс начинается в момент t и в этот момент известно $z(t) = z_0$. Тогда поставленная выше задача оптимального управления сводится к решению задачи Коши для уравнения Беллмана [1]

$$S_t + \min_{u\in U}(S_z, Az + b) + {}^1\!/_2 \operatorname{Sp}(DQ'B^{-1}QDS_{zz}) = 0 \tag{1.9}$$

$$S(T, z) = [(2\pi)^n \det D(T)]^{-1/2} \int_{-\infty}^{+\infty} F(x) \exp\left[-\frac{1}{2}(D^{-1}(T)(z - x), (z - x))\right] dx$$

Если погрешности измерений велики, т. е. $B^{-1} = \varepsilon B_*$, где B_* — матрица с ограниченными элементами, ε — малый параметр, то задача (1.9) будет эквивалентна задаче (1.3), (1.4).

2. Метод малого параметра. Перейдем к построению приближенного решения задачи (1.3) (1.6) методом малого параметра.

Предположим, что функции H, L, F и S достаточно гладкие, и будем искать решение задачи (1.3), (1.6) в виде регулярного разложения по степеням малого параметра

$$S(t, x) = S^\circ(t, x) + \varepsilon S^1(t, x) + \ldots \tag{2.1}$$

а функцию H из (1.5) представим в виде

$$H(S_x, x, t) = H(S_x^\circ, x, t) + \varepsilon(\nabla H(S_x^\circ, x, t), S_x^1(t, x)) + \ldots \tag{2.2}$$

Здесь и далее ∇H — вектор частных производных функции H по компонентам вектора S_x.

Подставим разложения (2.1), (2.2) в (1.3), (1.6) и, ограничиваясь членами первого порядка по ε, найдем уравнения и начальные значения для функций S°, S^1

$$S_t^{1\circ} - H(S_x^\circ, x, t) = 0, \quad S^\circ(t, x) = F(t, x) \quad \text{при} \quad h(t, x) = 0 \tag{2.3}$$

$$\begin{gathered} S_t^1 - (\nabla H(S_x^\circ, x, t), S_x^1) + {}^1\!/_2 \operatorname{Sp}(LS_{xx}^\circ) = 0 \\ S^1(t, x) = 0 \quad \text{при} \quad h(t, x) = 0 \end{gathered} \tag{2.4}$$

Аналогичные уравнения можно выписать и для высших приближений.

Из соотношений (1.5), (2.3) следует, что H — функция Гамильтона, а $S^\circ(t, x)$ — функция Беллмана для задачи оптимального управления при отсутствии случайных возмущений. Будем предполагать, что эта задача решена, т. е. найден синтез оптимального управления $u^\circ(t, x)$, функция $S^\circ(t, x)$ для детерминированной системы и соответствующее поле оптимальных траекторий

$$x = \varphi(t, a) \tag{2.5}$$

где $\varphi(t, a)$ — вектор-функция, a — n-мерный вектор произвольных постоянных. Предположим, что равенство (2.5) можно разрешить относительно a в некоторой области фазового пространства и получить зависимость

$$a = \psi(t, x) \tag{2.6}$$

Уравнение первого приближения (2.4) есть линейное неоднородное уравнение в частных производных первого порядка. Система уравнений, определяющих его характеристики, имеет вид

$$dx / dt = -\nabla H(S_x^\circ, x, t), \qquad dS^1 / dt = -{}^1/_2 \operatorname{Sp}(LS_{xx}^\circ) \tag{2.7}$$

Учитывая обозначение (1.5) и известное равенство $S_x'^\circ = -p$, где p — вектор сопряженных переменных для детерминированной системы, заметим, что первое уравнение (2.7) определяет, согласно принципу максимума, оптимальные траектории детерминированной системы. Общее решение этого уравнения задается равенством (2.5), а система первых интегралов — равенством (2.6). Тогда из (2.7), (2.5), (2.6) вытекает, что решение задачи Коши (2.4) для функции S^1 выражается следующим образом:

$$S^1(t, x) = \frac{1}{2} \int\limits_t^{T(x)} \operatorname{Sp}[L(\tau) S_{\zeta\zeta}^\circ(\tau, \zeta)]\, d\tau \tag{2.8}$$

Здесь $T(x)$ — корень уравнения $h(T, x) = 0$, а матрицу $S_{\zeta\zeta}^\circ$ следует брать при

$$\zeta = \varphi(\tau, \psi(t, x)) \tag{2.9}$$

Синтез оптимального управления в нулевом и первом приближении u°, u^1 получается из условий

$$(S_x^\circ, b(t, u^\circ(t, x))) = \min\nolimits_{u \in U} (S_x^\circ, b(t, u)) \tag{2.10}$$
$$(S_x^\circ + \varepsilon S_x^1, b(t, u^1(t, x))) = \min\nolimits_{u \in U} (S_x^\circ + \varepsilon S_x^1, b(t, u))$$

Равенства (2.1), (2.8) — (2.10) определяют в явном виде приближенное решение уравнения Беллмана и синтез оптимального управления для задачи со случайными возмущениями, если известно решение задачи синтеза для системы без возмущений. Построенное решение несправедливо там, где преобразование (2.5) необратимо. Отметим также, что вблизи поверхностей разрыва функции Беллмана S° или её производных необходимо к построенному выше решению, регулярному по ε, добавить нерегулярную часть решения типа пограничного слоя.

3. Оценки погрешности приближенного решения. Рассмотрим оценки погрешности приближенного решения задачи Коши (1.3), (1.4) для случая, когда момент T окончания процесса фиксирован, а минимизируется математическое ожидание значения функции $F(x(T))$ в конце процесса. В этом случае уравнения и начальные условия для функции Беллмана, а также функции нулевого и первого приближений (2.3), (2.4) имеют вид

$$S_t + (Ax, S_x) + M(t, S_x) + {}^1/_2\varepsilon\,\mathrm{Sp}\,(LS_{xx}) = 0, \quad S(T, x) = F(x) \tag{3.1}$$

$$S_t^\circ + (Ax, S_x^\circ) + M(t, S_x^\circ) = 0, \quad S^\circ(T, x) = F(x) \tag{3.2}$$

$$S_t^1 + (Ax, S_x^1) + (\nabla M(t, S_x^\circ), S_x^1) + {}^1/_2\,\mathrm{Sp}\,(LS_{xx}^\circ) = 0, \quad S^1(T, x) = 0 \tag{3.3}$$

Здесь и далее введены обозначения для функции M, вектора ее первых производных и матрицы вторых производных

$$M(t, S_x) = \min\nolimits_{u \in U}(S_x, b(t, u)), \quad \nabla M = \partial M / \partial S_x \tag{3.4}$$
$$N = \partial^2 M / dS_x^2$$

Предположим, что для всех $t \in [t_0, T]$, $u \in U$ и $x \in R_n$, где R_n — n-мерное евклидово пространство с нормой $\|x\| = (x_1^2 + \ldots + x_n^2)^{1/2}$, выполнены следующие условия.

1. Матрица L положительно определена.

2. Компоненты вектора b и элементы матриц A, L ограничены по модулю.

3. Функция F непрерывна и ограничена по модулю вместе со своими производными до третьего порядка включительно.

4. Функция M из (3.4) непрерывна и ограничена по модулю вместе со своими производными по S_x до второго порядка включительно.

5. Решения задач (3.1) — (3.3) существуют, единственны, непрерывны и ограничены по модулю вместе со своими производными до третьего порядка включительно.

Дальнейшие оценки опираются на следующую лемму, непосредственно вытекающую (после замены $t' = T - t$) из теоремы 10 работы [3].

Лемма. Пусть непрерывная и ограниченная функция $v(t, x)$ служит решением следующей задачи Коши:

$$\sum_{i,j=1}^{n} \alpha_{ij}(t, x)\frac{\partial^2 v}{\partial x_i \partial x_j} + \sum_{i=1}^{n} \beta_i(t, x)\frac{\partial v}{\partial x_i} + \gamma(t, x)v + \frac{\partial v}{\partial t} = \delta(t, x)$$
$$v(T, x) = v_0(x)$$

причем для всех $t \in [t_0, T]$, $x \in R_n$ матрица α_{ij} положительно определена, имеем $|v_0(x)| \leqslant k_1$, $|\delta(t, x)| \leqslant k_2$, $|\gamma(t, x)| \leqslant k_3$, а коэффициенты α_{ij}, β_i $(i, j = 1, \ldots, n)$ удовлетворяют условиям

$$|\alpha_{ij}(t, x)| \leqslant k_4(\|x\|^2 + 1), \quad |\beta_i(t, x)| \leqslant k_5(\|x\|^2 + 1)^{1/2}$$

где $k_1, \ldots, k_5$ — неотрицательные константы.

При этих предположениях

$$|v(t, x)| \leqslant [k_1 + k_2(T - t)], \exp[k_3(T - t), \quad t_0 \leqslant t \leqslant T, \; x \in R_n$$

Для дальнейших выкладок понадобятся следующие формулы для функции M, справедливость которых следует из условия 4:

$$M(t,\ S_x) = M(t,\ S_x^\circ) + (\nabla M(t,\ G_1),\ S_x - S_x^\circ) \tag{3.5}$$

$$M(t,\ S_x) = M(t,\ S_x^\circ) + (\nabla M(t, S_x^\circ),\ S_x - S_x^\circ) + (N(t,\ G_2)(S_x - S_x^\circ),\ S_x - S_x^\circ) \tag{3.6}$$

$$\nabla M(t,\ S_x) = \nabla M(t,\ S_x^\circ) + N(t, G_3)(S_x - S_x^\circ) \tag{3.7}$$

$$G_i = S_x^\circ + \theta_i(S_x - S_x^\circ), \qquad 0 < \theta_i < 1, \quad i=1,2,3$$

Для удобства записи условимся обозначать буквой C различные постоянные, не зависящие от ε.

Теорема 1. Если выполнены условия 1—5, то для $t \in [t_0, T]$, $x \in R_n$ имеем

$$|S(t,\ x) - S^\circ(t,\ x)| \leqslant C\varepsilon, \qquad \| S_x(t,\ x) - S_x^\circ(t,\ x) \| \leqslant C\varepsilon \tag{3.8}$$

Доказательство. Обозначим $v = S - S^\circ$. Из (3.1), (3.2) и разложения (3.5) находим, что $v(t,\ x)$ удовлетворяет уравнению

$$v_t + (Ax + \nabla M, v_x) + {}^1/_2\varepsilon\,\mathrm{Sp}\,(Lv_{xx}) = -{}^1/_2\varepsilon\,\mathrm{Sp}\,(LS_{xx}^\circ)$$

и начальному условию $v(T,\ x) = 0$. В силу 1—5 условия леммы выполнены, причем $|{}^1/_2\ \varepsilon\mathrm{Sp}\ (L\overset{\circ}{S}_{xx})| \leqslant C\varepsilon$. Тогда из леммы сразу следует первая оценка (3.8).

Для доказательства второй оценки (3.8) продифференцируем (3.1), (3.2) по компонентам вектора x, и, используя (3.7) получим после преобразований систему уравнений и начальных условий

$$(v_x)_t + (A' + S_{xx}N)v_x + v_{xx}[Ax + \nabla M(t, S_x^\circ)] + \tag{3.9}$$
$$+ {}^1/_2\varepsilon\,[\mathrm{Sp}\,(Lv_{xx})]_x = -{}^1/_2\varepsilon\,[\mathrm{Sp}\,(LS_{xx}^\circ)]_x$$
$$v_x(T, x) = 0$$

Обозначим $v_i^*(t) = \sup |\partial v(\tau,\ x) / \partial x_i|$ при $\tau \in [T - t,\ T]$, $x \in R_n$ и, применяя к (3.9) лемму, для каждой из компонент вектора v_x найдем

$$v_i^*(t) \leqslant C_1\varepsilon + C_2(T - t)\sum_{j \neq i} v_j^*(t), \qquad i = 1, \ldots, n \tag{3.10}$$

Здесь C_1, C_2 — константы, независящие от ε. Разобьем интервал времени $[t_0,\ T]$ на $N \gg 1$ равных интервалов длиной d такой, что $Nd = T - t_0$, $(n-1)C_2d > 1$ и просуммируем неравенства (3.10). Тогда для $t \in [T - d,\ T]$ получим

$$\|v_x(t, x)\| \leqslant \sum_{i=1}^{n} v_i^*(T - d) \leqslant \frac{nC_1\varepsilon}{1 - (n-1)C_2d} = C\varepsilon \tag{3.11}$$

Рассмотрим теперь интервал $[T - 2d,\ T - d]$. Так как в момент времени $T - d$ выполнено неравенство $\| v_x(T - d,\ x) \| \leqslant C\varepsilon$, то в силу леммы оценки типа (3.10), (3.11) имеют место и для $t \in [T - 2d,\ T - d]$. Рассуждая аналогичным образом, устанавливаем, что для всех $t \in [t_0,\ T]$, $x \in R_n$ будет $\| v_x(t,\ x) \| \leqslant C\varepsilon$, что и требовалось доказать.

Из теоремы 1 следует, что точное решение уравнения Беллмана отличается от функции нулевого приближения на величину порядка ε. Следующая теорема является оценкой погрешности первого приближения.

Теорема 2. Для функции $v^1(t, x) = S - S^\circ - \varepsilon S^1$ при условиях 1—5 имеют место оценки

$$|v^1(t, x)| \leqslant C\varepsilon^2, \qquad \| v_x^1(t, x) \| \leqslant C\varepsilon^2 \tag{3.12}$$

Доказательство. Воспользуемся разложением (3.6) и из (3.1) — (3.3) найдем

$$v_t^1 + (Ax + \nabla M, v_x^1) + {}^1/_2\, \varepsilon \mathrm{Sp}\,(Lv_{xx}^1) = -{}^1/_2 \varepsilon^2 \mathrm{Sp}\,(LS_{xx}^1) - \\ - (N(S_x - S_x^\circ), S_x - S_x^\circ), \quad v^1(T, x) = 0 \tag{3.13}$$

Из условий 4—5 и теоремы 1 следует, что правая часть в уравнении (3.13) — величина порядка ε^2. Отсюда по лемме получаем первую оценку (3.12). Вторая оценка (3.12) доказывается аналогично второй оценке (3.8).

Перейдем теперь к обоснованию формул приближенного синтеза (2.10) оптимального управления для системы с возмущениями.

Обозначим через $Z^0(t, x)$, $Z^1(t, x)$ математические ожидания значения функции $F(x(T))$ при условии, что в момент времени t система находится в состоянии x, а в качестве управления в (1.1) применяются управления соответственно нулевого и первого приближений $u^\circ(t, x)$, $u^1(t, x)$.

Теорема 3. Для введенных выше функций справедливы оценки

$$|S(t, x) - Z^\circ(t, x)| \leqslant C\varepsilon, \qquad |S(t, x) - Z^1(t, x)| \leqslant C\varepsilon^2 \tag{3.14}$$

Доказательство. Функции Z°, Z^1 являются (см. [4]) решениями следующих задач Коши:

$$Z_t^\circ + (Ax + b(t, u^\circ), Z_x^\circ) + {}^1/_2 \varepsilon\, \mathrm{Sp}\,(LZ_{xx}^\circ) = 0, \quad Z^\circ(T, x) = F(x) \tag{3.15}$$
$$Z_t^1 + (Ax + b(t, u^1), Z_x^1) + {}^1/_2 \varepsilon\, \mathrm{Sp}\,(LZ_{xx}^1) = 0, \quad Z^1(T, x) = F(x)$$

Для сокращения записи обозначим $w^\circ = Z^\circ - S^\circ$, $w^1 = Z^1 - S^\circ - \varepsilon S^1$. Из соотношений (3.2), (3.3), (3.15), с учетом (3.4), (2.10) найдем

$$w_t^\circ + (Ax + b(t, u^\circ), w_x^\circ) + {}^1/_2 \varepsilon\, \mathrm{Sp}\,(Lw_{xx}^\circ) = -{}^1/_2 \varepsilon\, \mathrm{Sp}\,(LS_{xx}^\circ) \\ w^\circ(T, x) = 0 \\ w_t^1 + (Ax + b(t, u^1), w_x^1) + {}^1/_2 \varepsilon\, \mathrm{Sp}\,(Lw_{xx}^1) = \\ = M(t, S_x^\circ) + \varepsilon(\nabla M(t, S_x^\circ), S_x^1) - M(t, S_x^\circ + \varepsilon S_x^1) - \\ - {}^1/_2 \varepsilon\, \mathrm{Sp}\,(LS_{xx}^1), \quad w^1(T, x) = 0 \tag{3.16}$$

Из свойств функции $M(t, S_x)$ и условий, 2,5 следует, что правая часть уравнения во второй задаче (3.16) — величина порядка ε^2. Последовательно применяя лемму к задачам (3.16), устанавливаем, что $|w^\circ(t, x)| \leqslant C\varepsilon$, $|w^1(t, x)| \leqslant C\varepsilon^2$. Отсюда с помощью теорем 1, 2 получаем

$$|S(t, x) - Z^\circ(t, x)| = |S - S^\circ + S^\circ - Z^\circ| \leqslant |S - S^\circ| + |Z^\circ - S^\circ| \leqslant C\varepsilon \\ |S(t, x) - Z^1(t, x)| = |S - S^\circ - \varepsilon S^1 - w^1| \leqslant |S - S^\circ - \varepsilon S^1| + |w^1| \leqslant \\ \leqslant C\varepsilon^2$$

что и требовалось доказать.

4. Преобразование приближенного решения. Приведем решение (2.8) — (2.10) к более удобному для применения виду. Введем параметрическое представление терминальной поверхности

$$t = T = \beta_0(a), \quad x(T) = \beta(a); \quad h(\beta_0(a), \beta(a)) \equiv 0 \tag{4.1}$$

Здесь a — n-мерный вектор, β_0 — скалярная функция, β — вектор-функция. Эти функции должны удовлетворять тождеству $h \equiv 0$ в (4.1). Векторный параметр a сохраняется постоянным вдоль оптимальных траекторий детерминированной системы (см. (2.5), (2.6)). Тогда из соотношений (2.3), (2.6), (4.1) вытекает, что

$$S^\circ(t, x) = F(\beta_0(a), \beta(a)) = F[\beta_0(\psi(t, x)), \beta(\psi(t, x))] \tag{4.2}$$

Из начального условия (2.3) следует, что на терминальной поверхности производные функции S° можно представить в виде

$$S_t^\circ = F_t + \lambda h_t, \quad S_x^\circ = F_x + \lambda h_x \quad (h = 0) \tag{4.3}$$

где $\lambda = \lambda(a)$ — некоторая функция от a. Подставляя равенства (4.3) в уравнение (2.3), получим

$$F_t + \lambda h_t - H(F_x + \lambda h_x, x, t) = 0 \tag{4.4}$$

Если в уравнение (4.4) подставить H из (1.5) и заменить t, x по формулам (4.1), то оно превратится в нелинейное алгебраическое уравнение для определения функции $\lambda(a)$.

Обозначим через $X(t)$ фундаментальную матрицу однородной системы, соответствующей (1.1). Эта матрица определяется условиями

$$dX / dt = A(t)X, \quad X(t_1) = E \tag{4.5}$$

где E — единичная матрица, t_1 — постоянная.

Как уже отмечалось, вектор S_x° вдоль оптимальных траекторий детерминированной системы с точностью до знака равен сопряженному вектору. Следовательно, при фиксированном a зависимость S_x° от t определяется матрицей $X'^{-1}(t)$. Учитывая еще условие (4.3), имеющее место в конце процесса ($t = T$), получим при всех t, x

$$S_x^\circ = X'^{-1}(t)q(a), \quad q(a) = X'(\beta_0(a))(F_x + \lambda h_x), \quad t = \beta_0(a), \; x = \beta(a) \tag{4.6}$$

В правую часть второго равенства (4.6) нужно подставить функцию $\lambda(a)$, а вместо аргументов t, x функций F, h — их выражения (4.1). Соотношения (4.2), (4.4), (4.6) определяют функции S°, S_x° в виде зависимости от t, a. Управление u° после этого определяется из первого условия (2.10).

В дальнейшем для функций от аргументов t, a будем сохранять прежние обозначения S°, S^1, u°, u^1.

Определив указанным способом управление $u^\circ(t, a)$, подставим его в систему (1.1) при $\xi = 0$ и запишем решение этой системы, удовлетворяющее условиям (4.1)

$$x = X(t)\Big[X^{-1}(\beta_0(a))\beta(a) + \int\limits_{\beta_0(a)}^{t} X^{-1}(\tau)\, b(\tau, u^\circ(\tau, a))\, d\tau\Big] \equiv \varphi(t, a) \qquad (4.7)$$

Соотношение (4.7) задает конкретный вид функции φ из (2.5), определяющей поле оптимальных траекторий детерминированной задачи. Разрешая (4.7) относительно a, получим функцию $\psi(t, x)$ в (2.6).

Завершив построение нулевого приближения, перейдем к первому приближению. Введем матрицу

$$\Phi(t, a) = \|\partial\varphi_i(t, a)/\partial a_j\|, \qquad i, j = 1, \ldots, n \qquad (4.8)$$

Тогда решение (2.8), (2.9) для функции S^1 и ее производных с учетом равенств (4.6), (4.8) примет вид

$$S^1(t, a) = \frac{1}{2} \int\limits_{t}^{\beta_0(a)} \mathrm{Sp}\left[L(\tau) X'^{-1}(\tau) \frac{\partial q(a)}{\partial a} \Phi^{-1}(\tau, a)\right] d\tau \qquad (4.9)$$

$$S_x^1 = \Phi^{-1}(t, a) S_a^1$$

Здесь $\partial q / \partial a$ — матрица частных производных вектора q из (4.6) по компонентам вектора a.

Таким образом, функция Беллмана в переменных t, a для нулевого и первого приближений определена равенствами (2.1), (4.2), (4.9) и соотношениями (4.1), (4.4) — (4.8) для функций β_0, β, λ, X, q, φ, Φ. Синтез оптимального управления в нулевом и первом приближении в этих же переменных задан равенствами (2.10), в которые нужно подставить S_x°, S_x^1 из (4.6), (4.9).

5. Частный случай и примеры. 1. Конкретизируем решение п. 2,4 для случая, когда функция b из (1.1) и ограничение (1.2) имеют вид

$$b(t, u) = K(t)\, u + c(t), \qquad \|u\| \leqslant k \qquad (5.1)$$

Здесь $K(t)$ — матрица размера $n \times m$, $c(t)$ — n-мерная вектор-функция, k — положительная постоянная, ограничивающая модуль вектора u. Из условий (2.10), (2.11) для случая (5.1) вытекает

$$u^\circ(t, x) = -kK'(t) S_x^\circ(t, x) \| K'(t) S_x^\circ(t, x)\|^{-1}$$
$$u^1(t, x) = -kK'(t)(S_x^\circ + \varepsilon S_x^1) \| K'(S_x^\circ + \varepsilon S_x^1)\|^{-1} \qquad (5.2)$$

Примем, что момент T окончания процесса фиксирован, так что $h = T - t$. В формулах (4.1) можно положить

$$\beta_0(a) = T, \qquad x(T) = \beta(a) = a \qquad (5.3)$$

т. е. в качестве параметра a принять конечное значение фазового вектора. Функцию F в (1.6) считаем не зависящей от t, т. е. $F = F(x)$. Кроме того, положим $t_1 = T$ в соотношении (4.5), так что $X(T) = E$. Тогда равенства (4.2), (4.6), (4.7), (4.9) с учетом (5.1) — (5.3) примут следующий вид:

$$S^\circ = F(a), \quad S_x^\circ = X'^{-1}(t)\, q(a), \quad q(a) = \partial F(a) / \partial a$$

$$x = \varphi(t, a) = X(t)\left\{a + \int\limits_{T}^{t} X^{-1}(\tau)\left[c(\tau) - \frac{kK(\tau) K'(\tau) X'^{-1}(\tau)\, q(a)}{\| K'(\tau) X'^{-1}(\tau)\, q(a)\|}\right] d\tau\right\} \qquad (5.4)$$

Приближенное решение полностью определено формулами (2.1), (5.4), (5.2), (4.8).

2. Пусть выполнены предположения п. 4 и, кроме того, функция $F(x)$ линейна по x, т. е. $F(x) = (r, x)$, где r — постоянный вектор. Тогда из соотношений (5.4), (5.2) вытекает

$$q(a) = r, \qquad S^\circ(t, a) = (r, a), \qquad S^1 \equiv 0$$
$$u^1 = u^\circ = -kK'(t)X'^{-1}(t)\, r \,\|K'(t)\, X'^{-1}(t)\, r\|^{-1} \tag{5.5}$$

Кроме того, из (5.4) следует, что функция $\varphi(t, a)$ оказывается линейной по a, так что между x и a имеется линейная связь. Решение (5.5) показывает, что оптимальное управление и функционал для задачи с возмущениями совпадает с решением детерминированной задачи. Этот результат очевиден: здесь задача Коши (1.3), (1.4) допускает точное решение $S(t, x)$, являющееся линейной функцией фазовых координат. Поэтому последний член уравнения (1.3), содержащий вторые производные и обусловленный случайными возмущениями, тождественно равен нулю для этого решения.

3. Рассмотрим простейшую систему, для которой уравнение (1.1) скалярно и имеет вид

$$dx / dt = u + \xi, \qquad |u| \leqslant k, \qquad k > 0 \tag{5.6}$$

Здесь ξ — белый шум постоянной малой интенсивности εg, k и g — постоянные, $\varepsilon \ll 1$. Требуется найти закон управления, доставляющий минимум математическому ожиданию функционала $F = x^2 / 2$ в заданный конечный момент T. Уравнение Беллмана (1.3) и начальное условие (1.4) для системы (5.6) при сделанных предположениях примут вид

$$S_t - k|S_x| + {}^1\!/_2 \varepsilon g S_{xx} = 0, \qquad S(T, x) = x^2 / 2 \tag{5.7}$$

Все матрицы и векторы здесь превращаются в скаляры. Функции, входящие в соотношения (5.1), (5.2), (5.4), для примера (5.6) примут вид

$$K(t) = X(t) = 1, \quad L(t) = g, \quad c(t) = 0, \quad S_x{}^\circ = q(a) = a \tag{5.8}$$

Используя соотношения (5.8), получим из равенств (5.4), (4.8) $x = \varphi(t, a) = a - k(t - T)\operatorname{sign} a$, $\Phi(t, a) = 1 \quad (a \neq 0)$

$$u^0(t, a) = u^1(t, a) = -k \operatorname{sign} a, \quad S^\circ(t, a) = a^2 / 2 \tag{5.9}$$
$$S^1(t, a) = \frac{1}{2} g (T - t)$$

Равенства (5.9) справедливы при $a \neq 0$. Случай $a = x(T) = 0$ отвечает попаданию в нуль, которое для детерминированной системы можно осуществить, очевидно, из области $|x| \leqslant k(T - t)$. В этой области оптимальное управление и траектории детерминированной системы неединственны, а функция Беллмана $S^\circ = 0$ при $a = 0$. Отметим, что при $a = 0$, $|x| \leqslant k(T - t)$ преобразование $x = \varphi(t, a)$ не является взаимно-однозначным, и во всей области $|x| \leqslant k(T - t)$ имеем тождественно $a = 0$. Из определения (4.8) формально следует, что $\Phi^{-1} = 0$, и тогда по формуле (4.9) найдем $S^1(t, a) = 0$ при $a = 0$. Переходя к переменным t, x, получим, согласно (5.9) и с учетом сделанных замечаний, приближенное решение $S^\circ + \varepsilon S^1$ задачи (5.7) в виде

$$S(t, x) = {}^1\!/_2 [|x| + k(t - T)]^2 + {}^1\!/_2 \varepsilon g (T - t), \quad |x| > k(T - t), \quad a \neq 0$$
$$S(t, x) = 0, \quad |x| < k(T - t), \quad a = 0 \quad (t \leqslant T) \tag{5.10}$$

Приближенное решение (5.10) имеет разрыв на прямых $|x| = k(T - t)$; для того, чтобы получить гладкое решение, нужно добавить к (5.10) решение типа пограничного слоя вблизи этих прямых.

Рассматриваемый пример относится к числу немногих случаев, когда можно получить точное решение уравнения Беллмана, которое затем можно сравнить с приближенным решением (5.10). Это представляет интерес, так как для данного примера не выполнены условия теорем п. 3, и вопрос об оценке погрешности решения (5.10) остается открытым. Заметим, что решение задачи (5.7) будет четной функцией от x, ограничимся поэтому областью $x \geqslant 0$. Следуя работе [5], получим точное решение задачи (5.7) в виде

$$S(t, x) = \int_0^\infty \left[\frac{1}{2} y^2 p(t, x - y) + \psi(y)\, p(t, x + y)\right] dy$$

$$p(t, x \pm y) = \frac{1}{\sqrt{2\pi\varepsilon g(T-t)}} \exp\left\{-\frac{[x-(T-t)k \pm y]^2}{2\varepsilon g(T-t)}\right\} \tag{5.11}$$

$$\psi(y) = \frac{1}{4\alpha^2}(1 - e^{-2\alpha y}) - \frac{ye^{-2\alpha y}}{2\alpha}, \quad \alpha = \frac{k}{\varepsilon g} = \text{const}, \quad x \geqslant 0$$

При помощи замены $x = a + k(T - t)$ и некоторых преобразований решение (5.11) приводится к виду

$$S(t, x) = \frac{a^2 + \varepsilon g(T-t)}{2} + \frac{1}{\sqrt{2\pi\varepsilon g(T-t)}} \int_0^\infty \left[\left(\frac{1}{4\alpha^2} - \frac{y^2}{2}\right) - \right. \tag{5.12}$$

$$\left. - \left(\frac{1}{4\alpha^2} + \frac{y}{2\alpha}\right)e^{-2\alpha y}\right] \exp\left[-\frac{(a+y)^2}{2\varepsilon g(T-t)}\right] dy, \quad \alpha = \frac{k}{\varepsilon g}$$

$$a = x - k(T-t), \quad x \geqslant 0$$

Получим теперь асимптотическое представление решения (5.12) при $\varepsilon \to 0$. Для этого оценим по методу Лапласа [6] интегралы, входящие в (5.12), поочередно для различных областей изменения x, t. Окончательно получим

$$S(t, x) = \frac{1}{2}[x + k(t-T)]^2 + {}^1/_2\varepsilon g(T-t) + \alpha_1, \quad x > k(T-t)$$

$$S(t, x) = {}^1/_4\varepsilon g(T-t) + \frac{\varepsilon^2 g^2}{8K^2} + \alpha_2, \quad x = k(T-t) \tag{5.13}$$

$$S(t, x) = \frac{\varepsilon^2 g^2}{4K^2} + \alpha_3, \quad 0 \leqslant x < k(T-t)$$

Здесь величины α_1, α_2, α_3 стремятся к нулю при $\varepsilon \to 0$ и фиксированных x, t быстрее, чем любая степень ε.

Учитывая свойство четности функции S и сравнивая приближенное решение (5.10) с асимптотическим представлением точного решения (5.13), убеждаемся, что они совпадают с точностью до величин порядка ε^2 всюду, кроме прямых $|x| = k(T - t)$, на которых приближенное решение (5.10) имеет разрыв. Пример показывает возможность применения метода малого параметра и в тех случаях, когда не выполнены условия теорем п. 3.[1]

Авторы выражают глубокую благодарность Г. И. Эскину за ценное обсуждение.

Поступила 18 XI 1971

ЛИТЕРАТУРА

1. Фельдбаум А. А. Основы теории оптимальных автоматических систем. М., Физматгиз, 1963.
2. Kalman R. E., Bucy R. S. New results in linear filtering and prediction theory. Trans. ASME, Ser. D, J. Basic Engng., 1961, vol. 83, No. 1.
3. Ильин А. М., Калашников А. С., Олейник О. А. Линейные уравнения второго порядка параболического типа. Успехи матем. наук, 1962, т. 17, вып. 3.
4. Дынкин Е. Б. Марковские процессы. М., Физматгиз, 1963.
5. Мошков Е. М. О точности оптимального управления конечным состоянием. ПММ, 1970, т. 34, вып. 3.
6. Лаврентьев М. А., Шабат Б. В. Методы теории функций комплексного переменного. М., Физматгиз, 1958.
7. Колмановский В. Б., Черноусько Ф. Л. Задачи оптимального управления при неполной информации. Тр. зимней школы по математическому программированию и смежным вопросам. М., 1971, вып. 1.
8. Fleming W. H. Stochastic control for small noise intensities SIAM, J. Control, 1971, vol. 9, No 3.

[1] Метод малого параметра для уравнения Беллмана, разработанный выше, был ранее предложен в [7]. Другой подход к аналогичным задачам содержится в недавно опубликованной статье [8].