Прикладная математика и механика. Том 34, 1970

О ТОЧНОСТИ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ КОНЕЧНЫМ СОСТОЯНИЕМ

Е. М. Мошков

(Москва)

Рассматривается уравнение Беллмана в частных производных, связанное с синтезом стохастически оптимального управления конечным состоянием линейной системы. Получены приближенные формулы и оценки решения, основанные на решении уравнения Беллмана для детерминированного варианта задачи. Предложен метод численного решения. В одномерном случае задача сведена к интегральному уравнению первого рода и при некоторых дополнительных предположениях получена конечная формула решения.

1. Исходные уравнения. В работе [1] и ряде других рассматривалась задача синтеза ограниченного стохастически оптимального управления конечным состоянием линейной системы, описываемой уравнениями [1]

$$x^{\bullet} = Q(t)\,u + \xi(t) \qquad (0 < t \leqslant T, \quad u(t) \in U(t)) \tag{1.1}$$

Здесь t — время, x, u и $\xi(t)$ — векторы координат, управления и случайных возмущений, имеющие соответственно порядки n, m и n; $Q(t)$ — заданная матрица коэффициентов, $U(t)$ — замкнутое множество допустимых значений вектора управления, $[0, T]$ — заданный интервал времени. Информация о текущем состоянии системы дается вектором $y(t) = G(t)\,x(t) + h(t)$, где $G(t)$ — заданная матрица, $h(t)$ — вектор случайных ошибок получения информации. Предполагалось, что воздействия ξ и h представляют собой белые шумы и имеют вместе с вектором $x(0)$ нормальный закон распределения. Требовалось найти оптимальный оператор управления $u(t) = u\{y(\tau),\ u(\tau);\ 0 \leqslant \tau < t\}$, минимизирующий оценку точности $S = M\omega\,[x(T)]$ — математическое ожидание заданной скалярной функции $\omega\,[x(T)]$ вектора конечного состояния системы.

Показано, что оптимальное управление представляется в виде функции $u(t, z(t))$, где $z(t) = M\,[x(t) \diagdown y(\tau),\ u(\tau); 0 \leqslant \tau < t]$ —условное математическое ожидание вектора $x(t)$ при известной реализации $y(\tau)$, $u(\tau)$ в интервале $[0, t)$. Функция $u(t, z)$ определяется вместе с соответствующей апостериорной оценкой $S(t, z(t)) = M\,\{\omega\,[x(T)] \diagdown y(\tau),\ u(\tau);\ 0 \leqslant \tau < t\}$ из нелинейного уравнения Беллмана в частных производных второго порядка

$$-S_t = \min_u (S_z,\ Q(t)\,u) + {}^1\!/_2\, \mathrm{sp}\,(S_{zz} R(t)) \qquad (0 \leqslant t < T) \tag{1.2}$$

---

[1] Заметим, что система более общего вида

$$x_1^{\bullet} = A(t) x_1 + Q_1(t)\,u + \xi_1(t)$$

приводится к виду (1.1) заменой переменной $x = L(T) L^{-1}(t)\, x_1$, где $L(t)$ — фундаментальная матрица однородной системы.

с граничным условием при $t = T$

$$S(T, z) = \Psi(z) = \int \omega(z + x)\, q(T, x)\, dx \tag{1.3}$$

$$q(T, x) = [(2\pi)^n \,|C(T)|]^{-1/2} \exp {}^1/_2 (-C^{-1}(T)\, x, x)$$

Здесь z — n-мерный вектор, S_t — частная производная функции Беллмана $S(t, z)$ по t, S_z и S_{zz} — вектор первых частных производных и матрица вторых частных производных функции $S(t, z)$ по составляющим вектора z, $(.,.)$ — скалярное произведение векторов, sp — след матрицы, $R(t)$ и $C(T)$ — ограниченные неотрицательно определенные квадратные матрицы, $|C(T)|$ — определитель матрицы $C(T)$, $q(T, x)$ — плотность нормального распределения случайного вектора с корреляционной матрицей $C(T)$. Интегрирование в формуле (1.3) проводится по всему пространству векторной переменной интегрирования x (в дальнейшем это относится ко всем случаям, когда пределы интегрирования не указаны).

Матрицы $R(t)$ и $C(T)$ определяются формулами

$$R(t) = B^{\bullet}(T, t), \quad B(T, t) =$$

$$= M\,[(x(T) - z(t))(x(T) - z(t))' \setminus u(\tau) = 0, \ \tau > t] \tag{1.4}$$

$$C(T) = B(T, T) = M\,[(x(T) - z(T))(x(T) - z(T))']$$

где точка означает дифференцирование по времени t, а штрих — операцию транспонирования. Они выражаются через коэффициенты системы и характеристики закона распределения процессов $\xi(t)$ и $h(t)$ и вектора $x(0)$. При этом, если $x(0) - Mx(0) = \xi(t) - M\xi(t) \equiv 0$, то $C(T) = R(t) \equiv 0$. В случае отсутствия ошибок получения информации матрица $R(t)$ совпадает с интенсивностью белого шума $\xi(t)$ в [2].

Для уравнений типа (1.2), (1.3) в работе [3] доказана теорема существования и единственности решения. В данной статье рассматриваются вопросы получения конечных формул приближенного и точного решений задачи Коши (1.2), (1.3), которые были бы удобны для практического применения. Более подробно исследуется случай, когда множество $U(t)$ представляет собой параллелепипед

$$|u^{(i)}(t)| \leqslant l^{(i)}(t) \qquad (i = 1, \ldots, m) \tag{1.5}$$

ограничивающий составляющие $u^{(i)}(t)$ вектора управления. Минимизация по параметру u в уравнении (1.2) в этом случае проводится в явной форме. Если система описывается скалярным уравнением

$$x^{\bullet} = u + \xi \qquad (0 < t \leqslant T) \tag{1.6}$$

а множество $U(t)$ представляет собой отрезок $|u(t)| \leqslant l(t)$ за вычетом быть может некоторых внутренних участков, то уравнения Беллмана принимают вид:

$$-S_t = -l(t)\,|S_x| + {}^1/_2 R(t)\, S_{xx} \qquad (0 \leqslant t < T) \tag{1.7}$$

$$S(T, x) = \Psi(x) \tag{1.8}$$

(для удобства дальнейшего изложения в уравнениях (1.7), (1.8) по сравнению с (1.2), (1.3), изменено обозначение фазовой переменной).

Заметим, что система, у которой $n = 1$, $m \neq 1$, приводится к виду (1.6) введением скалярного управляющего параметра $u_1 = Q(t)\, u - u_0(t)$, где $2u_0(t) = \max_u Q(t)\, u + \min_u Q(t)\, u$ и нового возмущения $\xi_1(t) = \xi(t) + u_0(t)$. При этом

$$2l(t) = \max_u Q(t)\, u - \min_u Q(t)\, u$$

2. Решение уравнения Беллмана для системы с полной информацией. Случай с полной информацией, когда начальные условия и возмущения у системы являются заданными, можно рассматривать как частный случай системы с неполной информацией, положив в уравнениях начальное отклонение $x(0) - Mx(0)$ и возмущение $\xi(t)$ тождественно равными нулю [1]. Матрицы $R(t)$ и $C(t)$ при этом равны нулю, и уравнения (1.1) — (1.3) принимают вид

$$x^{\cdot} = Q(t)u \tag{2.1}$$

$$-S_t^{\circ} = \min_u (S_x^{\circ}, Q(t)\, u) \qquad (0 \leqslant t < T,\ u \in U(t)) \tag{2.2}$$

$$S^{\circ}(T, x) = \omega(x) \tag{2.3}$$

Чтобы отличить детерминированный вариант задачи от стохастического, апостериорная оценка $S^{\circ}(t, x)$ в детерминированной задаче снабжается индексом «нолик сверху». Рассмотрим несколько важных для дальнейшего частных случаев, когда апостериорная оценка $S^{\circ}(t, x)$ находится на основе уравнений движения (2.1).

1°. Пусть векторы координат и управления одномерны. Уравнения Беллмана в этом случае записываются следующим образом:

$$-S_t^{\circ} = -l(t)\, |S_x^{\circ}|, \qquad S^{\circ}(T, x) = \omega(x) \tag{2.4}$$

Предположим, что функция $\omega(x)$ кусочно-непрерывна, имеет единственный минимум в некоторой точке x° и является при $x \geqslant x^{\circ}$ неубывающей, а при $x \leqslant x^{\circ}$ — невозрастающей. Из определения апостериорной оценки

$$S^{\circ}(t, x) = \min_u \omega\, [x(T) \diagdown x(t) = x] \tag{2.5}$$

и уравнений (2.1) следует, что функция $S^{\circ}(t, x)$ в этом случае выражается равенствами

$$\begin{gathered} S^{\circ}(t, x) = \omega\, [x - b(t, T)\operatorname{sgn}(x - x^{\circ})] \qquad (|x - x^{\circ}| \geqslant b(t, T)) \\ S^{\circ}(t, x) = \omega\, [x^{\circ}] \qquad (|x - x^{\circ}| \leqslant b(t, T)), \qquad b(t, T) = \int_t^T l(\tau)\, d\tau \end{gathered} \tag{2.6}$$

---

[1] Система, у которой $\xi(t) = M\,\xi(t) \neq 0$, приводится к системе без возмущений переходом к новым координатам

$$x_1 = x + \int_t^T \xi(\tau)\, d\tau$$

Подставляя функцию (2.6) в уравнения (2.4), можно убедиться в том, что эта функция служит решением задачи Коши (2.4) в обычном смысле [4], если $\omega(x)$ еще и непрерывно дифференцируема.

2°. Пусть функция $\omega(x)$ задана равенствами

$$\omega(x) = 1 \quad (x \notin D), \qquad \omega(x) = 0 \quad (x \in D) \tag{2.7}$$

где D — заданная замкнутая область в пространстве n-мерных векторов. При таком выборе функции $\omega(x)$ оптимизируется попадание вектора $x(T)$ в заданную область D. Определим область $D(t, T)$ как область тех значений вектора x в момент t, из которых система (2.1) в момент T может быть переведена в область D. Очевидно, что необходимым и достаточным условием оптимальности управления $u(\tau, x)$, $t \leqslant \tau \leqslant T$, является требование, чтобы это управление обеспечивало перевод системы в момент T в область D для всех $x(t) \in D(t, T)$. Отсюда следует, что апостериорная оценка, соответствующая оптимальному управлению, выражается равенствами

$$S^\circ(t, x) = 1 \quad (x \notin D(t, T)), \qquad S^\circ(t, x) = 0 \quad (x \in D(t, T)) \tag{2.8}$$

Функция (2.8) ввиду ее разрывности не может служить решением задачи Коши (2.2), (2.3) в обычном смысле. В одномерном случае ее можно рассматривать как решение в следующем обобщенном смысле: к функции (2.8) сходится последовательность решений (2.6), соответствующая последовательности $\omega_1(x), \omega_2(x), \ldots$ дифференцируемых униминимальных функций, которая сходится к функции (2.7). (Можно ожидать, что это утверждение остается справедливым также и в многомерном случае.) В отличие от обычного определения обобщенного решения [4] в данном определении отсутствует требование, чтобы сходимость была равномерной.

Область $D(t, T)$ описывается неравенством $T^\circ(t, x) \leqslant T - t$, где $T^\circ(t,x)$ — минимальное время, необходимое для перевода системы в область D из положения x в момент t. Построение области $D(t, T)$, связанное с решением задачи оптимального быстродействия [2], в многомерном случае является достаточно сложной задачей. Пусть множество $U(t)$ и область D есть параллелепипеды, заданные соответственно неравенствами (1.5) и неравенствами

$$|x^{(i)} - a_0^{(i)}| \leqslant d_0^{(i)} \qquad (i = 1, \ldots, n) \tag{2.9}$$

где $a_0^{(i)}$, $d_0^{(i)}$ — заданные величины. В этом случае сравнительно просто определяется прямоугольная область

$$|x^{(i)} - a_0^{(i)}| \leqslant d^{(i)}(t) \qquad (i = 1, \ldots, n)$$

описанная вокруг области $D(t, T)$. Она обозначается в дальнейшем $D^*(t, T)$. Размеры $d^{(i)}(t)$ $(i = 1, \ldots, n)$ выражаются формулой

$$d^{(i)}(t) = d_0^{(i)} + \max_u \{x^{(j)}(T) \diagdown x^{(i)}(t) = 0\} = d_0^{(i)} + \sum_{j=1}^{m} \int_t^T |Q_{(j)}^{(i)}(\tau)| \, l^{(j)}(\tau)\, d\tau \tag{2.10}$$

Функция, определяемая равенствами

$$S^*(t, x) = 1 \quad (x \notin D^*(t, T)), \qquad S^*(t, x) = 0 \quad (x \in D^*(t, T)) \tag{2.11}$$

служит оценкой снизу для апостериорной оценки (2.8).

В случае произвольных D и $U(t)$ их всегда можно заключить в параллелепипеды D_1 и $U_1(t)$ и определить по формулам (2.10) соответствующую им область $D_1^*(t, T)$. Эта область содержит в себе область $D^*(t, T)$. Поэтому функция $S_1^*(t, x)$, определяемая равенствами (2.11), где область $D^*(t, T)$ заменяется областью $D_1^*(t, T)$, служит оценкой снизу для функций $S^*(t, x)$ и $S^\circ(t, x)$

$$S_1^*(t, x) \leqslant S^*(t, x) \leqslant S^\circ(t, x) \tag{2.12}$$

3°. Пусть функция $\omega(x)$ достигает минимума в некоторой точке x°. В этом случае для апостериорной оценки $S^\circ(t, x)$ выполняется условие

$$S^\circ(t, x) = \omega(x^\circ) \quad (x \in D_{x^\circ}(t, T)) \tag{2.13}$$

Здесь $D_{x^\circ}(t, T)$ — область значений координат x в момент t, из которых система в момент T может быть приведена в точку x°. Равенство (2.13) следует из определения апостериорной оценки (2.5).

3. Оценки и приближенные формулы решения уравнения Беллмана для стохастической системы. В этом разделе исследуется несколько вариантов приближенного решения задачи Коши (1.2), (1.3), каждый из которых является оценкой снизу для точного решения. Рассмотрим функцию, определяемую формулой

$$F_1(t, z) = \int S_1(t, x)\, P(t, T, x - z)\, dx = \int S_1(t, x + z)\, P(t, T, x)\, dx \tag{3.1}$$

Здесь $S_1(t, x)$ — решение уравнения (2.2) с условием (1.3), $P(t, T, x - z)$ — фундаментальное решение [5] уравнения

$$-P_t = {}^1/_2\, \mathrm{sp}\,(P_{zz} R(t)) \tag{3.2}$$

соответствующего системе (1.1) при отсутствии управления. Функция $P(t, T, x)$ выражается [6] формулами:

$$P(t, T, x) = (2\pi)^{-n/2} \,|\, K_1(t, T)\,|^{-1/2} \exp\left(-{}^1/_2\, K_1^{-1}(t, T)\, x, x\right) \tag{3.3}$$

$$K_1(t, T) = \int_t^T R(\tau)\, d\tau = B(T, t) - C(T) \tag{3.4}$$

Она равна плотности нормального распределения вектора $z(T) - z(t)$ в предположении, что $u(\tau) = 0$ при $\tau \geqslant t$. Это следует из равенств (1.4) и некоррелированности векторов $x(T) - z(T)$ и $z(T) - z(t)$. Функция $S_1(t, x)$ представляет собой апостериорную оценку при оптимальном управлении системой с полной информацией в предположении, что оптимизируемая функция конечного состояния задана формулой (1.3). Функция $F_1(t, z(t))$ по построению равна апостериорной оценке для случая оптимального управления системой, которая в дополнение к имевшейся информации $y(\tau)$, $u(\tau)$ $(0 \leqslant \tau < t)$ в момент t получает всю последующую информацию $y(\tau)$ $(t \leqslant \tau < T)$. При оптимальном управлении дополнительная информация не может ухудшить точность управления (увеличить оценку точности), поэтому функция $F_1(t, z)$ будет оценкой снизу для функции Беллмана

$$F_1(t, z) \leqslant S(t, z) \tag{3.5}$$

В двух следующих случаях разность $F_1(t, z) - S(t, z)$ обращается в нуль:

1) если равна нулю при всех t матрица $R(t)$ (это соответствует отсутствию измерений или отсутствию возмущений и начальных отклонений);

2) если система не управляется (т. е. множество $U(t)$ состоит из одной точки $u = 0$).

Поэтому функцию $F_1(t, z)$ можно рассматривать в качестве приближенного решения задачи Коши (1.2), (1.3), если мала матрица $R(t)$ или малы размеры множества $U(t)$.

Пусть множество $U(t)$ представляет собой замкнутую область, а система (2.1) является управляемой [7] в интервале $[T - \delta, T]$ при любом как угодно малом $\delta > 0$. Тогда разность между точным и приближенным решениями стремится к нулю также и в случае неограниченного расширения области $U(t)$

$$F_1(t, T)\big|_{r(U)=\infty} = \lim_{r(U)\to\infty} F_1(t, z) = S(t, z)\big|_{r(U)=\infty} \tag{3.6}$$

Здесь $r(U)$ — радиус шара с центром в точке $u = 0$, целиком умещающегося в области $U(t)$ при всех t из интервала $[T - \delta, T]$ для некоторого $\delta > 0$.

В самом деле, если управление неограничено, то при сделанных предположениях для оптимальности управления необходимо и достаточно выполнения равенства $z(T) = z^\circ$, где z° есть точка минимума функции $\Psi(z)$ в равенстве (1.3). Априорная и апостериорная оценки S и $S(t, z)$ при этом достигают минимума

$$\min_u S\big|_{r(U)\to\infty} = \lim_{r(U)\to\infty} S(t, z) = \Psi(z^\circ) \qquad (0 \leqslant t < T) \tag{3.7}$$

С другой стороны, используя формулы (3.1), (3.3) и (2.13), где $\omega(x)$ заменяется на $\Psi(x)$, а также неограниченность расширения области $D_{z^\circ}(t, T)$ с расширением области U, легко доказать равенство

$$F_1(t, z)\big|_{r(U)\to\infty} = \lim_{r(U)\to\infty} F_1(t, z) = \Psi(z^\circ) \qquad (0 \leqslant t < T) \tag{3.8}$$

Из сравнения формул (3.7) и (3.8) следует требуемый результат.

Рассмотрим функцию $F_2(t, z)$, которая определяется формулами (3.1), (3.3), где функция $S_1(t, x)$ заменяется функцией $S^\circ(t, x)$ (уравнения (2.2), (2.3)), а матрица $K_1(t, T)$ заменяется матрицей

$$K_2(t, T) = C(T) + \int_t^T R(\tau)\, d\tau = B(T, t) \tag{3.9}$$

Функция $F_2(t, z)$ будет оценкой снизу для $F_1(t, z)$.

В самом деле, пусть система в момент t получает полную в смысле критерия $S = M\omega\,[x\,(T)]$ информацию $y_1 = x\,(T) \setminus u\,(\tau) = 0\,(\tau > t)$. Функция, определяемая равенствами (3.3), (3.9), (1.4), равна плотности нормального распределения вектора $y_1 - z\,(t)$. Отсюда следует, что функция $F_2\,(t,\ z\,(t))$ есть апостериорная оценка для данной гипотетической системы при оптимальном управлении. Поэтому функция $F_2\,(t,\ z)$ будет оценкой снизу для функций $F_1\,(t,\ z)$ и $S\,(t,\ z)$.

Можно показать точно так же, как это было сделано в случае с функцией $F_1\,(t,\ z)$, что разность $S\,(t,\ z) - F_2\,(t,\ z)$ стремится к нулю, если устремить к нулю размеры множества $U\,(t)$ или дисперсии случайных воздействий.

В случае (2.7) оценками снизу для $F_2\,(t,\ z)$, очевидно, будут функции

$$F^*\,(t,\ z) = \int S^*\,(t,\ x)\,P\,(t,\ T,\ z - x)\,dx, \quad F_1^*\,(t,\ z) = \\ = \int S_1^*\,(t,\ x)\,P\,(t,\ T,\ z - x)\,dx \tag{3.10}$$

Здесь функция $P\,(t,\ T,\ x)$ задается также, как при определении $F_2\,(t,\ z)$ (формулы (3.3), (3.9)).Функции $S^*\,(t, x)$ и $S_1^*\,(t,\ x)$ находятся по формуле (2.11), они служат оценками снизу (2.12) для $S^\circ\,(t,\ x)$. Таким образом, выполняется цепочка неравенств

$$F_1^*\,(t,\ z)) \leqslant F^*\,(t,\ z) \leqslant F_2\,(t,\ z) \leqslant F_1\,(t,\ z) \leqslant S\,(t,\ z)$$

Оценки $F_1^*\,(t,\ z)$, $F^*\,(t,\ z)$, $F_2\,(t,\ z)$ более грубые по сравнению с $F_1\,(t, z)$ однако они проще вычисляются. В частности, функция $F_1^*\,(t,\ z)$ выражается конечными формулами.

4. Численное решение уравнения Беллмана. Пусть функция $\omega\,(x)$ задана равенствами (2.7). Апостериорная оценка $S\,(t,\ z\,(t))$ в этом случае представляет собой условную вероятность того, что точка $x\,(T)$ при оптимальном управлении окажется за пределами области D, и поэтому для функции $S\,(t,\ z)$ выполняются неравенства

$$F_1^*\,(t,\ z) \leqslant S\,(t,\ z) \leqslant 1 \tag{4.1}$$

Определим в координатах z прямоугольную область $\Omega\ (\varepsilon,\ t)$ по формулам

$$|\,z^{(i)} - a_0^{(i)}\,| \leqslant d_2^{(i)}\,(\varepsilon,\ t) \qquad (i = 1, \ldots, n)$$

$$d_2^{(i)}\,(\varepsilon,\ t) = d_1^{(i)}\,(t) + \gamma\,(\varepsilon)\,\sqrt{[K_2\,(t,\ T)]_{(i)}^{(i)}} \tag{4.2}$$

Здесь $\gamma\ (\varepsilon)$ — функция обратная функции

$$^1/_2 - \Phi\,(\gamma) = {}^1/_2 - (2\pi)^{-1/2}\int\limits_0^\gamma e^{-\frac{x^2}{2}}\,dx \tag{4.3}$$

ε — заданное достаточно малое положительное число, $d_1^{(i)}\,(t)$ ($i = 1, \ldots, n$) — размеры области $D_1^*\,(t,\ T)$, определяемые по формулам (2.10); $[K_2\,(t,\ T)]_{(i)}^{(i)}$ — соответствующий диагональный элемент матрицы (3.9). Из равенств (2.11) и (3.10) видно, что функция $F_1^*\,(t,\ z)$ всюду не превосходит единицы, а вне области $\Omega\ (\varepsilon,\ t)$ отличается от единицы меньше, чем на ε (размеры области $\Omega\ (\varepsilon,\ t)$ определялись именно из этого условия).

Отсюда с учетом неравенств (4.1) следует, что и функция $S(t, z)$ при $z \notin \Omega(\varepsilon, t)$ отличается от единицы меньше, чем на ε. Исходя из этого факта, можно предложить следующий численный метод решения уравнений (1.2), (1.3).

Выбирается достаточно малое число ε и решение полагается равным единице на границе области $\Omega(\varepsilon, t)$. В результате исходная задача Коши (1.2), (1.3), заменяется первой краевой задачей [5] для уравнения (1.2) с условиями (1.3) и

$$S(t, z)\big|_{z\in\Gamma(\varepsilon, t)} = 1 \tag{4.4}$$

Здесь $\Gamma(\varepsilon, t)$ — граница области $\Omega(\varepsilon, t)$. Ввиду непрерывной зависимости решения уравнения (1.2) от граничных условий [3] погрешность от такой замены будет иметь порядок величины ε. Область $\Omega(\varepsilon, t)$ в равенстве (4.4) можно заменить более широкой областью $\Omega(\varepsilon, 0)$ (см. формулы (4.2), (3.9), (2.10)). В этом случае численное решение задачи можно вести по явной разностной схеме [8] с сеткой узлов, независящей от номера k слоя $t_k = T - k\Delta$, $k = 0, 1, \ldots, N$. Решение еще более упрощается в случае (1.5).

Рассмотренный метод, очевидно, применим и в случае произвольной функции $\omega(x)$, если она имеет конечный предел $\omega(\infty)$ при $|x| \to \infty$. Вместо заданной области D (или области, описанной вокруг D) в этом случае в формулах, определяющих размеры области $\Omega(\varepsilon, t)$, будут фигурировать размеры прямоугольной области, за пределами которой функция $\omega(x)$ отличается от предельного значения $\omega(\infty)$ меньше, чем на ε. Граничное значение функции $S(t, x)$ в равенстве (4.4) полагается равным $\omega(\infty)$.

5. Конечная формула решения уравнения Беллмана для одномерного случая. Рассмотрим уравнения (1.7), (1.8), соответствующие одномерной системе (1.6). Предположим, что коэффициенты $l(t)$ и $R(t)$ непрерывны по t и положительны в интервале $[0, T]$

$$l(t) > 0, \qquad R(t) > 0 \qquad (0 \leqslant t \leqslant T) \tag{5.1}$$

а функция $\Psi(x)$ является четной, неубывающей при $x \geqslant 0$, непрерывно дифференцируемой и ограниченной

$$\Psi(-x) = \Psi(x), \quad \Psi'(x) \geqslant 0 \qquad (x \geqslant 0) \tag{5.2}$$

Штрихом здесь и далее обозначается производная по x. В этом случае условия существования и единственности решения задачи Коши (1.7), (1.8) будут заведомо выполненными [3]. Докажем, что решение выражается равенствами

$$S(t, x) = s(t, x) \quad (x \geqslant 0) \tag{5.3}$$
$$S(t, x) = s(t, -x) \quad (x \leqslant 0)$$

Здесь $s(t, x)$ — решение линейного параболического уравнения

$$-s_t = -l(t)s_x + {}^1/_2 R(t)s_{xx} \quad (x > 0, \ 0 \leqslant t < T) \tag{5.4}$$

удовлетворяющее условиям

$$s(T, x) = \Psi(x) \quad (x \geqslant 0) \tag{5.5}$$

$$s_x(t, x)\big|_{x=0} = 0 \quad (0 \leqslant t \leqslant T) \tag{5.6}$$

(Это уравнение получается из уравнения (1.7) заменой $|S_x|$ на s_x.)

Доказательство. Пусть $S(t, x)$ есть решение уравнений (1.7), (1.8). Тогда функция $S(t, -x)$ также служит решением, в чем легко убедиться простой подстановкой этой функции в указанные уравнения. Ввиду единственности решения отсюда следует, что функция $S(t, x)$ является четной по x: $S(t, -x) = S(t, x)$. Из полученного равенства и непрерывной дифференцируемости $S(t, x)$ по x (это требование входит в определение понятия решения [4]) следует, что $S_x(t, x)|_{x=0} = 0$ при всех t из интервала $[0, T]$. Учитывая неотрицательность $S_x(T, x) = \Psi'(x)$ при $x \geqslant 0$, можно предположить, что $S_x(t, x) \geqslant 0$ при всех t из интервала $[0, T]$ и $x \geqslant 0$. В связи с этим рассмотрим краевую задачу для линейного параболического уравнения (5.4) с условиями (5.5), (5.6). Эта краевая задача имеет единственное решение [5]. Дифференцируя равенства (5.4), (5.5) по x, приходим к заключению, что производная $s_x(t, x)$ удовлетворяет тому же самому уравнению (5.4) и условиям

$$s_x(T, x) = \Psi'(x) \geqslant 0 \quad (x \geqslant 0), \qquad s_x(t, x)|_{x=0} = 0$$

Отсюда на основании принципа максимума для линейного параболического уравнения [5] следует, что $s_x(t, x) \geqslant 0$ при $x \geqslant 0$. Таким образом, $s_x(t, x) = |s_x(t, x)|$ при $x \geqslant 0$, и поэтому функция $s(t, x)$ удовлетворяет уравнениям (1.7), (1.8) при $x > 0$. Функция, определяемая равенствами (5.3), служит решением задачи Коши (1.7), (1.8), так как она удовлетворяет условию (1.8) при всех x и уравнению (1.7) при $x \neq 0$, а вследствие непрерывности этой функции и ее производных S_x, S_{xx}, S_t при всех x и t удовлетворяет уравнению (1.7) также и при $x = 0$.

Таким образом, исходная задача (1.7), (1.8), сводится к решению второй краевой задачи [5] для линейного уравнения (5.4) с условиями (5.5), (5.6). Покажем, что эта задача в свою очередь может быть сведена к решению некоторого интегрального уравнения первого рода.

Продолжим некоторым образом функцию $s(T, x)$ в равенстве (5.5) на отрицательные значения x с сохранением непрерывности

$$s(T, x) = \Psi(x) \quad (x \geqslant 0), \qquad s(T, x) = \Psi_1(-x) \quad (x \leqslant 0) \tag{5.7}$$

Здесь $\Psi_1(x)$ — пока неизвестная дифференцируемая функция, удовлетворяющая условию

$$\Psi_1(0) = \Psi(0) \tag{5.8}$$

Ищем решение краевой задачи (5.4) — (5.6) в виде решения задачи Коши (5.4), (5.7) в полосе $-\infty \leqslant x < \infty$, $0 \leqslant t < T$ по известной [5] формуле

$$s(t, x) = \int_{-\infty}^{+\infty} p(t, T, x-y) \quad s(T, y)\,dy = \int_0^\infty [\Psi(y)\,p(t, T, x-y) + \\ + \Psi_1(y)\,p(t, T, x+y)]\,dy \tag{5.9}$$

Здесь $p(t, T, x-y)$ — фундаментальное решение уравнения (5.4), определяемое по формулам

$$p(t, \tau, x-y) = [2\pi t_1(t, \tau)]^{-1/2} \exp \frac{-[x - b(t, \tau) - y]^2}{2t_1(t, \tau)} \tag{5.10}$$

$$t_1(t, \tau) = \int_t^\tau R(s)\,ds, \quad b(t, \tau) = \int_t^\tau l(s)\,ds \tag{5.11}$$

Формула (5.10) может быть получена, например, следующим образом. Посредством замены переменных

$$t_1 = t_1(t, T) \qquad x_1 = x - b(t, T), \qquad s_1(t_1, x_1) = s[t(t_1), x(t_1, x_1)]$$

уравнение (5.4) сводится к обычному уравнению теплопроводности, фундаментальное решение которого выражается известной формулой

$$p_1(t_1 - \tau_1, x_1 - y_1) = [2\pi(t_1 - \tau_1)]^{-1/2} \exp \frac{-[x_1 - y_1]^2}{2(t_1 - \tau_1)}$$

Заметим, что обратное преобразование $t(t_1)$, $x(t_1, x_1)$ существует и единственно ввиду предположения о положительности $R(t)$ и $l(t)$. Возвращаясь к прежним переменным, получим формулу (5.10).

Дифференцируя равенство (5.9) по x и используя условие (5.6), получим уравнение первого рода [9] относительно производной $\Psi_1'(x)$

$$\int_0^\infty \left[\Psi'(x) \exp\left(-\frac{2b(t, T)}{t_1(t, T)} x\right) - \Psi_1'(x)\right] \exp \frac{-[x^2 - 2b(t, T)x]}{2t_1(t, T)} dx = 0 \qquad (5.12)$$

$$(0 \leqslant t \leqslant T)$$

Вследствие равенства (5.8) функция $\Psi_1(x)$ выражается формулой

$$\Psi_1(x) = \Psi(0) + \int_0^x \Psi_1'(y)\, dy \qquad (x \geqslant 0) \qquad (5.13)$$

В тех случаях, когда уравнение (5.12) имеет достаточно гладкое решение, решение задачи Коши (1.7), (1.8) может быть найдено по формулам (5.3), (5.9), (5.13).

Если проинтегрировать по частям левую часть уравнения (5.12), то получится интегральное уравнение для определения непосредственно функции $\Psi_1(x)$

$$\int_0^\infty K(x, t)\, \Psi_1(x)\, dx = f(t) \qquad (0 \leqslant t \leqslant T)$$

Ядро $K(x, t)$ и свободный член $f(t)$ интегрального уравнения выражаются формулами

$$K(x, t) = \frac{x - b(t, T)}{t_1(t, T)} \exp\left(-\frac{x^2 - 2b(t, T)x}{2t_1(t, T)}\right)$$

$$f(t) = \int_0^\infty \Psi(x) \frac{x + b(t, T)}{t_1(t, T)} \exp\left(-\frac{x^2 + 2b(t, T)x}{2t_1(t, T)}\right) dx$$

1°. Пусть отношение коэффициентов уравнения Беллмана (1.7) равно постоянной величине

$$\frac{l(t)}{R(t)} = \alpha(t) = \alpha = \mathrm{const} \qquad (5.14)$$

В этом случае решение интегрального уравнения (5.12) выражается конечной формулой

$$\Psi_1'(x) = e^{-2\alpha x} \Psi'(x) \qquad (5.15)$$

Действительно, как видно из равенств (5.11), в случае (5.14) $b(t, T)\, t_1^{-1}(t, T) = \alpha$, и поэтому уравнение (5.12) при подстановке (5.15) обращается в тождество. Подставляя выражение (5.15) в формулу (5.13), находим

$$\Psi_1(x) = e^{-2\alpha x}\Psi(x) + 2\alpha\int_0^x e^{-2\alpha y}\Psi(y)\,dy \tag{5.16}$$

С учетом полученного равенства формула (5.9) приводится к виду

$$s(t, x) = \int_0^\infty g_0(t, x; T, y)\, s(T, y)\,dy, \quad s(T, y) = \Psi(y) \tag{5.17}$$

$$\begin{aligned} g_0(t, x; T, y) = {} & p(t, T, x - y) + e^{-2\alpha y} p(t, T, x + y) - \\ & - 2\alpha e^{-2\alpha y}\left[\Phi\left(\frac{x + y - b(t, T)}{\sqrt{t_1(t, T)}}\right) - {}^1/_2\right] \end{aligned} \tag{5.18}$$

Здесь $\Phi(x)$ — интеграл вероятности (4.3).

Формулы (5.17), (5.18) выводятся следующим образом. Подставляя выражение (5.16) функции $\Psi_1(x)$ в формулу (5.9) получаем

$$\begin{aligned} s(t, x) = {} & \int_0^\infty [p(t, T, x - y) + e^{-2\alpha y} p(t, T, x + y)]\,\Psi(y)\,dy + \\ & + 2\alpha\int_0^\infty p(t, T, x + y)\int_0^y e^{-2\alpha z}\Psi(z)\,dz\,dy \end{aligned} \tag{5.19}$$

Используя тождество

$$p(t, T, x + y) = \frac{d}{dy}\left[\Phi\left(\frac{x + y - b(t, T)}{\sqrt{t_1(t, T)}}\right) - {}^1/_2\right]$$

имеем

$$\begin{aligned} \int_0^\infty p(t, T, x + y)\int_0^y e^{-2\alpha z}\Psi(z)\,dz\,dy = {} & \left[\Phi\left(\frac{x + y - b(t, T)}{\sqrt{t_1(t, T)}}\right) - {}^1/_2\right]\int_0^y e^{-2\alpha z}\,\Psi(z)\,dz\,\Big|_{y=0}^{y=\infty} - \\ & - \int_0^\infty \left[\Phi\left(\frac{x + y - b(t, T)}{\sqrt{t_1(t, T)}}\right) - {}^1/_2\right] e^{-2\alpha y}\Psi(y)\,dy \end{aligned}$$

Первое слагаемое при подстановке пределов обращается в нуль, в результате чего формула (5.19) с учетом обозначения (5.18) приводится к виду (5.17).

В некоторых случаях функцию $g_0(t, x; T, y)$ удобнее определять по формуле

$$g_0(t, x; T, y) = -\frac{\partial}{\partial y}\int_\infty^x [p(t, T, z - y) - e^{-2\alpha y} p(t, T, z + y)]\,dz \tag{5.20}$$

которая дает выражение, тождественно равное (5.18).

Хотя формулы (5.17), (5.18), были получены в предположении непрерывной дифференцируемости и ограниченности функции $\Psi(x)$, они применимы и в случае, когда функция $\Psi(x)$ имеет разрывы первого рода, а также в случае ее неограниченности на бесконечности при условии, что она возрастает не сильнее, чем функция

$$\exp\left[{}^1/_2\left(\int_0^T R(\tau)\,d\tau-\varepsilon\right)^{-1}x^2\right]$$

где $\varepsilon>0$. Действительно, в указанных случаях функция (5.17), (5.18) сохраняет достаточную гладкость и по построению удовлетворяет уравнению (5.4) и условиям (5.5), (5.6). Полученные формулы точного решения могут быть использованы при анализе точности различных приближенных методов решения уравнения Беллмана путем сравнения приближенного и точного решений при $\alpha(t)=\mathrm{const}$.

2°. Если $\alpha(t)=l(t)/R(t)$ — кусочно-постоянная функция, принимающая значения $\alpha_1,\ldots,\alpha_N$ на интервалах $[0,t_1]$, $(t_1,t_2]$, ..., $(t_{N-1}\ T]$, то решение $s(t,x)$ на каждом из интервалов $[t_{k-1},t_k]$ можно получить последовательно для $k=N,N-1,\ldots,1$ по формулам (5.17), (5.18), где T и α заменяются соответственно на t_k и α_k.

Решение и в этом случае можно представить в виде, аналогичном (5.17)

$$s(t,x)=\int_0^\infty g_{N-k}(t,x;T,y)\,s(T,y)\qquad(t_{k-1}\leqslant t\leqslant t_k)\tag{5.21}$$

если положить

$$g_{N-k}(t,x;T,y)=\int_0^\infty g_0(t,x;t_k,z)\,g_{N-k-1}(t_k,z;T,y)\,dz\tag{5.22}$$

Функция $s(t,x)$, определяемая равенствами (5.21), (5.22), и ее производные по x до второго порядка включительно непрерывны при всех x и t из интервала $0\leqslant t\leqslant T$. Производная $s_t\ (t,x)$ при $t=t_k\ (k=1,\ldots,N)$ имеет разрывы первого рода (испытывает скачки). Поэтому формулы (5.3), (5.21), (5.22) следует рассматривать как обобщенное решение [4] уравнений (1.7), (1.8).

Пример. Предположим, что ошибки измерения отсутствуют, интенсивность белого шума ξ в уравнении (1.6) постоянна и равна R, управление одномерно, ограничение $l(t)$ постоянно и равно l, а функция $\omega(x)$ задана равенствами

$$\omega(x)=0\quad(|x|\leqslant d_0),\qquad\omega(x)=1\quad(|x|>d_0)$$

В этом случае в уравнениях (1.7), (1.8) $\Psi(x)=\omega(x)$, $R(t)=R$, $l(t)=l$. Таким образом, $\alpha(t)=l/R=\mathrm{const}$ и решение выражается формулами (5.3), (5.17), (5.20), (5.10), (5.11). В результате вычислений получаем

$$S(t,x)=\int_{d_0}^\infty g_0(t,x;T,y)\,dy=\int_{d_0}^\infty\frac{\partial}{\partial y}\int_\infty^x[p(t,T,z-y)-e^{-2\alpha y}p(t,T,z+y)]\,dz\,dy=$$

$$=\Phi\left(\frac{x-d_0-l(T-t)}{\sqrt{R(T-t)}}\right)+\frac{1}{2}-e^{-2\frac{l}{R}d_0}\left[\Phi\left(\frac{x+d_0-l(T-t)}{\sqrt{R(T-t)}}\right)-\frac{1}{2}\right]\ (x\geqslant 0)\tag{5.23}$$

Приближенное решение (3.1) — (3.3) в этом случае в соответствии с равенствами (2.6) приводится к виду

$$F_1(t, x) = F_2(t, x) = 1 - \int\limits_{-d_0 - l(T-t)}^{d_0 + l(T-t)} p(t, T, x - y)\, dy =$$

$$= 1 - \Phi\left(\frac{x + d_0 + l(T-t)}{\sqrt{R(T-t)}}\right) + \Phi\left(\frac{x - d_0 - l(T-t)}{\sqrt{R(T-t)}}\right) \qquad (5.24)$$

Из сравнения точного (5.23) и приближенного (5.24) решений видно, что приближенное решение действительно является оценкой снизу для точного решения, и что разность между ними стремится к нулю в случае $l \to 0$ или $l \to \infty$, а также в случае $R \to 0$ или $R \to \infty$.

Поступила 24 VI 1969

ЛИТЕРАТУРА

1. Богуславский И. А. О статистически оптимальном управлении конечным состоянием. Автоматика и телемеханика, 1966, № 5.
2. Фельдбаум А. А. Основы теории оптимальных автоматических систем, изд. 2, М., «Наука», 1966, стр. 248—254.
3. Fleming W. H. Some Markovian optimisation problems. J. Math. and Mech., 1963, vol. 12, No. 1, pp. 131—140.
4. Бабич В. М., Капилевич М. Б., Михлин С. Г., Натансон Г. И., Риз П. М., Слободецкий Л. Н., Смирнов М. М. Линейные уравнения математической физики. М., «Наука», 1964, стр. 18.
5. Ильин А. М., Калашников А. С., Олейник О. А. Линейные уравнения второго порядка параболического типа. Усп. матем. н., 1962, т. 17, вып. 3 (105).
6. Эйдельман С. Д. Параболические системы. М., «Наука», 1964, стр. 15—16.
7. Красовский Н. Н. Теория управления движением. М., «Наука», 1968, стр. 138.
8. Березин И. С., Жидков Н. П. Методы вычислений. М., Физматгиз, 1960, т. 2. стр. 490—497.
9. Забрейко П. П., Кошелев А. И., Красносельский М. А., Михлин С. Г., Раковщик Л. С., Стеценко В. Я. Интегральные уравнения. М., «Наука», 1968, стр. 154—158.