УДК В161.1
М33

МЕТОД РЕШЕНИЯ ОБОБЩЕННОГО УРАВНЕНИЯ СТРАТОНОВИЧА-БЕЛЛМАНА

Кудинов А.Н., Катулев А.Н., Малевинский М.Ф.
Кафедра математического моделирования

Выведено дифференциальное уравнение в частных производных, описывающее оптимальное движение управляемой немарковской системы. Предложен метод решения уравнения с использованием обобщенного ряда Котельникова, проведено численное исследование и установлено, что метод обеспечивает высокую точность решения.

The partial differential equation for describing optimum movement of non-Markov system is derived. The method of a solution of the equation based on the generalized Kotelnikov series is proposed. By using numerical examination is established, that the method ensures high accuracy of solution.

Введение. Известно, что задачи управления стохастическими динамическими объектами (процессами) при неполной информации к настоящему времени ставятся и разрешаются для марковских процессов, то есть когда движение объекта задается с помощью стохастического дифференциального уравнения, в котором случайные возмущения – помехи представляются непрерывными случайными процессами диффузионного или скачкообразного типа. Естественным при этом является описание плотностей распределения вероятностей таких процессов уравнениями Колмогорова и сведение задачи выработки оптимального управления к решению уравнения с частными производными второго порядка типа уравнения Стратоновича-Беллмана для критериального функционала.

При управлении объектом в этом случае исключается зависимость текущего состояния от его поведения на некотором отрезке времени до фиксированного предшествующего – начального момента, хотя соответствующая исходная информация для управления доставляется, в соответствии с теоремой разделения или стохастической эквивалентности, линеаризованными фильтрами типа Камана-Бьюси или байесовскими с конечной памятью.

В связи с этим возникает актуальная задача разработки метода и алгоритма управления стохастическим объектом без сведения его к марковскому, то есть с учетом его поведения на некотором отрезке времени, непосредственно предшествующему текущему моменту, принимаемому за начальный. Соответствующий управляемый процесс классифицируется как процесс с последействием.

Цель статьи: получение обобщенного уравнения Стратоновича-Беллмана на основе обобщенного обратного – первого уравнения Колмогорова, содержащего частные производные по фазовой координате любых порядков и переменные коэффициенты, и метода его решения при выработке оптимального управления динамическим немарковским объектом (процессом с последействием) по неполным данным.

В известной литературе [1-14], в том числе и указанных в ней используемых работ, постановок и решений таких задач нами не обнаружено.

1. Постановка задачи. Воспользуемся формулой для условных вероятностей

$$P(x,s,z,t)=\int\limits_X P(x,s,y,u,z,t)P(x,s,y,u)dy, \tag{1}$$

где X – множество допустимых состояний объекта,

$(z,\,t)\sim z_t$ – состояние, в которое переходит объект в момент времени t из состояния $(y,u)\sim y_u$ на момент u при условии предшествования ему состояния $(x,s)\sim x_s$ на момент s, $t>u>s$,

$P(x,s,y,u,z,t)=P(z_t/y_u,x_s)$ – условная плотность вероятности состояния z_t при гипотезах x_s, y_u,

$P(x,s,y,u)=P(y_u/x_s)$ – условная плотность вероятности состояния y_u при гипотезе x_s,

$P(x,s,z,t)=P(z_t/x_s)$ – условная плотность вероятности состояния z_t при гипотезе x_s как состояния объекта на момент s.

Из приведенной формулы, как известно, следует уравнение Колмогорова-Чепмена для марковских прцессов – процессов без последействия.

Ниже будем рассматривать одномерный процесс с последействием, и, опираясь на выписанную формулу, выпишем сначала обобщенное обратное уравнение Колмогорова. Вид последнего непосредственно следует из традиционных преобразований условной плотности вероятности $P(z_t/y_u,x_s)$, если положить промежуточный момент u в силу его произвольности как $u=s+\Delta t$, где Δt-мало, затем разложить функцию

$$P(z_t/y_{s+\Delta t},x_s)=P(z_t/y_{s+\Delta t}(x_s))$$

в ряд Тейлора по $y(x_s)$ и учесть, что

$$\int\limits_X P(y_u/x_s)dy=1.$$

При этом предполагаем, что плотность $P(z_t/y_{s+\Delta t}(x_s))$ дифференцируема до любого порядка и её производные всех порядков ограничены. Тогда имеем

$$P(z_t/y_{s+\Delta t}(x_s))=P(z_t/x_s,s+\Delta t)+\frac{\partial P(z_t/x_s,s+\Delta t)}{\partial x_s}(y(x_s)-x_s)+$$

$$+\frac{1}{2}\frac{\partial^2 P(z_t/x_s,s+\Delta t)}{\partial x_s^2}(y(x_s)-x_s)^2+\frac{1}{6}\frac{\partial^3 P(z_t/x_s,s+\Delta t)}{\partial x_s^3}(y(x_s)-x_s)^3+\cdots$$

Подставив правую часть этого разложения в формулу для условной вероятности состояния z_t, получим

$$P(z_t/x_s)=P(z_t/x_s,s+\Delta t)+\int\limits_X\frac{\partial P(z_t/x_s,s+\Delta t)}{\partial x_s}(y(x_s)-x_s)P(y_{s+\Delta t}(x_s)/x_s)dy+$$

$$+\frac{1}{2}\int\limits_X \frac{\partial^2 P(z_t/x_s, s+\Delta t)}{\partial x_s^2}(y(x_s)-x_s)^2 P(y_{s+\Delta t}/x_s)dy+$$

$$+\frac{1}{6}\int\limits_X \frac{\partial^3 P(z_t/x_s, s+\Delta t)}{\partial x_s^3}(y(x_s)-x_s)^3 P(y_{s+\Delta t}/x_s)dy+\cdots$$

Отсюда после переноса первого слагаемого правой части в левую и деления обеих частей на Δt получаем соотношение

$$\frac{P(z_t/x_s)-P(z_t/x_s, s+\Delta t)}{\Delta t} = \frac{1}{\Delta t}\frac{\partial P(z_t/x_s, s+\Delta t)}{\partial x_s}\int\limits_X (y(x_s)-x_s)P(y_{s+\Delta t}(x_s)/x_s)dy+$$

$$+\frac{1}{2\Delta t}\frac{\partial^2 P(z_t/x_s, s+\Delta t)}{\partial x_s^2}\int\limits_X (y(x_s)-x_s)^2 P(y_{s+\Delta t}/x_s)dy+$$

$$+\frac{1}{6\Delta t}\frac{\partial^3 P(z_t/x_s, s+\Delta t)}{\partial x_s^3}\int\limits_X (y(x_s)-x_s)^3 P(y_{s+\Delta t}/x_s)dy+\cdots,$$

которое в предположении существования пределов всех его членов вида

$$\frac{1}{l!\Delta t}\int\limits_X (y(x_s)-x_s)^l P(y_{s+\Delta t}/x_s)dy$$

при $\Delta t \to 0$ очевидным образом преобразуется в обобщенное обратное уравнение Колмогорова с переменными коэффициентами

$$-\frac{\partial P(z_t/x_s)}{\partial s} = A_1(x_s)\frac{\partial P(z_t/x_s)}{\partial x_s}+\frac{1}{2}A_2(x_s)\frac{\partial^2 P(z_t/x_s)}{\partial x_s^2}+\frac{1}{6}A_3(x_s)\frac{\partial^3 P(z_t/x_s)}{\partial x_s^3}+\cdots \quad (2)$$

Коэффициенты могут быть вычислены по заданному стохастическому уравнению, описывающему движение управляемого объекта [8]. Так, если уравнение движения задается в виде

$$\dot z(t) = [c(t)+\vartheta(t)]\varphi(z(t))+b(t)u(t)+n(t), \quad (3)$$

где $c(t)$, $b(t)$ – детерминированные функции времени, $\varphi(t)$ – нелинейная функция, представимая отрезком ряда Тейлора, $\vartheta(t)$ – случайный коррелированный шум, возникающий в ходе движения объекта, $u(t)$ – в общем случае кусочно-непрерывная функция из допустимого множества U, подлежащая вычислению в каждый момент времени s поступления наблюденной реализации процесса z_s, $s<t$, $n(t)$ – входное коррелированное случайное возмущение, которое в частном случае можно аппроксимировать белым шумом, то коэффициенты, как функции фазовой координаты и управления, определяются следующими выражениями

$$A_1(x_s, u(t)) = \lim_{\Delta t\to 0} M[\Delta z/x_s],$$

$$A_2(x_s, u(t)) = \lim_{\Delta t \to 0} M[\Delta z \Delta z / x_s],$$

$$A_3(x_s, u(t)) = \lim_{\Delta t \to 0} M[\Delta z \Delta z \Delta z / x_s],$$

$$\Delta z = \int_t^{t+\Delta t} [c(\tau) + \vartheta(\tau)] \varphi(z(\tau)) d\tau + \int_t^{t+\Delta t} b(\tau) u(\tau)\, d\tau + \int_t^{t+\Delta t} n(\tau)\, d\tau$$

и пусть

$$\varphi(z(\tau)) = \varphi(x_s) + \frac{\partial \varphi(x_s)}{\partial z} \int_t^T dz,$$

где dz выражается непосредственно из уравнения движения объекта.

Поэтому будем считать все коэффициенты известными непрерывными и удовлетворяющими условию Липшица функциями.

Теперь введем критерий качества управления динамическим объектом в виде условного относительно z_s среднего риска

$$J(z_s, s) = \min_{u \in U} M[\int_s^T f(t, z_t, u(t)) dt + \Phi(z_T, T) / z_s = x_s] =$$

$$= \int_X [\int_s^T f(t, z_t, u(t)) dt + \Phi(z_T, T) / z_s = x_s] P(z_t / x_s) dz, \tag{4}$$

где M - обозначение математического ожидания, z_s – реализация процесса на непосредственно предшествующем моменту t отрезке времени или это условие выработки оптимального управления по наилучшему переводу объекта в состояние z_t.

Продифференцируем этот критерий по s, получим

$$\frac{\partial J(z_s, s)}{\partial s} = \min_{u \in U} \{ - \int_X f(s, z_s, u(s)) P(z_s / z_s) dz +$$

$$+ \int_X [\int_s^T f(t, z_t, u(t)) dt + \Phi(z_T, T) / z_s = x_s] \frac{\partial P(z_t / x_s)}{\partial s} dz\}$$

или, после очевидного результата вычисления первого интеграла и подстановки во второй выражения для производной условной плотности из обобщенного уравнения Колмогорова, имеем следующее выражение для критерия

$$-\frac{\partial J(z_s, s)}{\partial s} = \min_{u \in U} \{ f(s, z_s, u(s)) + A_1(z_s, u) \frac{\partial J(z_s, s)}{\partial x_s} + \frac{1}{2} A_2(x_s, u) \frac{\partial^2 J(z_s, s)}{\partial x_s^2} +$$

$$+ \frac{1}{6} A_3(x_s, u) \frac{\partial^3 J(z_s, s)}{\partial x_s^3} + \dots \}. \tag{5}$$

Это обобщенное уравнение Стратоновича–Беллмана, искомое оптимальное управление по которому немарковским объектом должно отыскиваться с учетом граничного условия

$$J(z_T, T) = \Phi(z_T, T).$$

Итак, задача, подлежащая решению, состоит в том, чтобы найти оптимальную функцию управления $u(t) \in U$ динамическим немарковским объектом, описываемым стохастическим уравнением типа (3), при которой обеспечивалось бы оптимальное движение объекта на отрезке $[t,T]$ и минимальное среднее значение критерия (4) как функционала от управления и наблюденных состояний объекта на некотором отрезке времени, непосредственно предшествующем отрезку $[t,T]$.

Близкими к такой задаче являются задачи оптимального управления марковскими объектами при неполной обратной связи [6].

2. Метод выработки оптимального управления. В результате определения оптимального управления $u^*(t)$ на интервале времени $[s,T]$ уравнение (5) запишется в виде

$$-\frac{\partial J(z_s, s)}{\partial s} = f(s, z_s, u^*(s)) + \sum_{i=1}^{L} \frac{1}{i!} A_i(z_s, u^*) \frac{\partial^i J(z_s, s)}{\partial z_s^i} \tag{6}$$

с граничным условием

$$J(z_s, T) = \Phi(z_s, T). \tag{7}$$

В основу решения уравнения (6) положим идею интерполяционного метода Канторовича и представим искомое решение обобщенным рядом Котельникова

$$J_N(z_s, s) = \sum_{k=-N}^{N} a_k(s) \left(\frac{\sin \alpha (z_s - k\pi/\alpha)}{\alpha (z_s - k\pi/\alpha)} \right)^n, \tag{8}$$

где $a_k(s)$ – неизвестные коэффициенты, подлежащие вычислению; n – заданная степень функции отсчетов ряда; π/α – допустимый интервал времени между замерами, соответствующий граничной частоте спектра функции $J(z_s, s)$.

На интервале изменения аргумента z_s выберем точки $z_{sk} = k\pi/\alpha$, $k = \overline{-N, N}$, затем потребуем, чтобы в этих точках после подстановки приближенного решения (8) в уравнение (6) последнее обращалось в равенство.

В результате получим систему обыкновенных дифференциальных уравнений $2 \times N + 1$ порядка относительно коэффициентов $a_k(s)$:

$$\dot{a}_k(s) = -f(s, z_{sk}, u^*(s)) - \sum_{i=1}^{L} \frac{1}{i!} A_i(z_{sk}, u^*) \frac{\partial^i J(z_{sk}, s)}{\partial z_s^i}, k = \overline{-N, N}, \tag{9}$$

с граничным условием

$$a_k(T) = \Phi(z_{sk}, T),$$

интегрирование которой может быть выполнено численно с использованием стандартных методов.

3. Результаты исследования метода. Для проверки точности предложенного метода решения обобщенного уравнения Стратоновича – Беллмана (6) зададим дифференциальное уравнение (3) в виде

$$\dot{z}_t = Signz_t + b(t)u(t) + n(t),$$

где $n(t)$- белый шум с единичной интенсивностью, а $b(t)$и $f(s, z_s, u*(s))$ положим равными нулю.

При этом уравнение (6) примет вид

$$\frac{\partial J(z_s, s)}{\partial s} = -\frac{\partial J(z, s)}{\partial z} - \frac{1}{2}\frac{\partial^2 J(z, s)}{\partial z^2}, \tag{10}$$

а его точное установившееся решение –

$$J(z_s) = \exp\{-2|z_s|\}.$$

Аппроксимируем решение (10) обобщенным рядом Котельникова

$$J_4(z_s) = \sum_{k=0}^{4} a_k \left(\frac{\sin(2\pi z_s - k\pi)}{2\pi z_s - k\pi}\right)^2, \tag{11}$$

где a_k, k=0,1,2,3,4, - неизвестные постоянные коэффициенты, подлежащие вычислению, и в качестве точек z_k возьмем точки 0.5; 1; 1.5; 2. Тогда коэффициенты определяются из решения следующей системы алгебраических уравнений

$$a_0 - \frac{1}{3}\pi^2 a_1 + a_2 + \frac{1}{4}a_3 + \frac{1}{9}a_4 = 0$$

$$\frac{1}{4}a_0 + a_1 - \frac{1}{3}\pi^2 a_2 + a_3 + \frac{1}{4}a_4 = 0$$

$$\frac{1}{9}a_0 + \frac{1}{4}a_1 + a_2 - \frac{1}{3}\pi^2 a_3 + a_4 = 0$$

$$\frac{1}{16}a_0 + \frac{1}{9}a_1 + \frac{1}{4}a_2 + a_3 - \frac{1}{3}\pi^2 a_4 = 0$$

$$0.242a_0 + 0.463a_1 + 0.463a_2 + 0.242a_3 + 0.018a_4 = 0.5.$$

Заметим, что в этой системе последнее уравнение составлено из условия нормировки

$$\int_0^{1.5} J_4(z)dz = 0.5.$$

В таблице приведены приближенное и точное решения уравнения (10).

z_s	0	0,25	0,5	0,75	1,0	1,25	1,5	1,75	2,0
$J_4(z_s)$	0,8461	0,4928	0,3379	0,2718	0,2187	0,1815	0,1478	0,1183	0,0890
$J(z_s)$	1,00	0,6065	0,3679	0,2231	0,1353	0,0821	0,0498	0,0302	0,0183

Видно, что предложенный метод обеспечивает получение искомого решения уравнения (6) с погрешностью не более 15%.

Список литературы

[1] Пупков К.А., Егупов Н.Д., Макаренков А.М., Трофимов А.И. Теория и компьютерные методы исследования стохастических систем. М., Физматлит, 2003. 400с.

[2] Кляцкин В.И. Динамика стохастических систем. М., Физматлит, 2003.

[3] Беллман Р. Динамическое программирование. М., ИЛ, 1960.

[4] Ховард Р.А. Динамическое программирование и марковские процессы. М., сов. Радио, 1964.

[5] Прохоров Ю.В., Розанов Ю.А. Теория вероятностей. Основные понятия, предельные теоремы, случайные процессы. М., Наука, 1967.

[6] Параев Ю.И. Введение в статистическую динамику процессов управления и фильтрации. М., сов. Радио, 1976.

[7] Хазен Э.М. Методы оптимальных статистических решений и задачи оптимального управления. М., сов. Радио, 1968.

[8] Евланов Л.Г., Константинов В.М. Системы со случайными параметрами. М., Наука, 1976.

[9] Крылов Н.В. Управляемые процессы диффузионного типа. М., Наука, 1977.

[10] Гихман И.И., Скороход А.В. Введение в теорию случайных процессов. М., Наука, 1977.

[11] Андреева Е.А., Колмановский В.Б., Шайхет Л.Е. Управление системами с последействием. М., Наука, 1991.

[12] Ярлыков М.С., Миронов М.А. Марковская теория оценивания случайных процессов. М., Радио и связь, 1993.

[13] Красовский А.А., Буков В.Н., Шендрик В.С. Универсальные алгоритмы оптимального управления непрерывными процессами. М., Наука, 1977.

[14] Пугачев В.Н., Лифшиц Н.А. Вероятностный анализ систем автоматического управления. М., Советское радио, 1963.