УДК 519.2

ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ ЛИНЕЙНОЙ ДИСКРЕТНОЙ СИСТЕМОЙ ПО ВЕРОЯТНОСТНЫМ КРИТЕРИЯМ

В.М. Азанов

Московский авиационный институт (Научный исследовательский университет)

Россия, 125993, Москва, Волоколамское шоссе, 4

E-mail: azanov59@gmail.com

Ключевые слова: стохастическое оптимальное управление, дискретная система, вероятностные критерии качества

Аннотация: Рассматривается задача оптимального управления линейной дискретной стохастической системой. Критерием оптимальности служит вероятность попадания первой координаты системы в заданную окрестность нуля за время, не превышающее заданную величину. Задача сводится к эквивалентной задаче стохастического оптимального управления с вероятностным терминальным критерием. Последняя аналитически решается методом динамического программирования. Приводятся достаточные условия, при которых найденное оптимальное управление оказывается оптимальным и по квантильному критерию.

1. Введение

Задачи оптимального управления по вероятностным критериям качества составляют предмет изучения специального раздела теории стохастического оптимального управления. К числу вероятностных критериев относятся функционал вероятности и функционал квантили. Функционал вероятности представляет собой вероятность непревышения некоторым точностным функционалом заданного допустимого уровня. Сам точностный функционал при этом характеризует точность системы управления, но зависит от траектории стохастической системы. Примером точностного функционала служит терминальный промах системы наведения. В постановке задачи оптимального управления с критерием качества в форме функционала вероятности обычно требуется этот функционал максимизировать.

Функционал квантили является, в некотором смысле, обратной характеристикой по отношению к функционалу вероятности. Физический смысл функционала квантили в том, что он, будучи верхней доверительной границей для точностного функционала, по сути характеризует гарантированную по вероятности точность системы управления. Задача оптимального управления с критерием в форме функционала квантили обычно трактуется как задача минимизации.

Практическая важность задач оптимального управления с вероятностными критериями исторически признавалась авторитетными специалистами в теории управления. Среди публикаций по данной тематике можно выделить [1–4]. При этом наиболее значимые результаты по развитию общей теории данного класса задач можно найти в [1], где подробно исследованы задачи оптимального управления дискретными стохастическими системами с вероятностными критериями и терминальным точностным функционалом. Можно отметить также статью [5], в которой исследован вопрос об эквивалентности задач оптимального управления с функционалами вероятности и квантили.

В настоящей статье впервые рассматривается задача оптимального управления с функционалом вероятности с нефиксированным, но ограниченным сверху моментом окончания. Это обстоятельство позволяет интерпретировать рассматриваемую задачу управления, как задачу наведения. Рассматриваемая ниже постановка схожа с непрерывной задачей управления материальной точкой [3], однако в [3] на время окончания процесса управления никаких ограничений не накладывалось.

2. Постановка задачи

Рассмотрим задачу управления с учетом случайных воздействий в дискретном времени $k = \overline{0, N+1}$. Динамика системы описывается системой рекуррентных уравнений

$$(1) \qquad \begin{cases} x^1_{k+1} = x^1_k + x^2_k h, \\ x^2_{k+1} = x^2_k + x^3_k h, \\ \dots \\ x^n_{k+1} = x^n_k + u_k h + \xi_k. \end{cases}$$

Здесь x^i_k - i-ая координата фазового вектора $x_k = (x^1_k, ..., x^n_k)^T$ в k-ый момент времени, u_k - управление, h - параметр, возникающий при дискретизации непрерывной системы, $n \geqslant 2$, $N \geqslant n-1$. Функции распределения независимых случайных величин ξ_k непрерывно дифференцируемы. Плотности распределений ξ_k являются четными и унимодальными функциями. ξ_k одинаково распределены. Начальные условия $x^1_0, ..., x^n_0$ для системы (1) детерминированы.

Структура (1) возникает при дискретизации линейной непрерывной системы, записанной в канонической форме управляемости, где случайная помеха моделирует ошибки в канале управления.

Функционал вероятности определяется выражением

$$(2) \qquad P_\varphi(u(\cdot)) = P\left(\min_{k \in \{0,...,N+1\}} |x^1_k| \leqslant \varphi \right),$$

где P - вероятность, $\varphi \in R^1$ - скалярный параметр, $u(\cdot) = (u_0(\cdot), ..., u_N(\cdot))$ - управление. Стратегия управления на k-ом шаге ищется, как функция $u_k(x_0, ..., x_k)$.

Задача нахождения оптимального управления, максимизирующего вероятностный критерий оптимальности имеет вид

$$(3) \qquad P_\varphi(u(\cdot)) \to \max_{u(\cdot)}.$$

Данную задачу можно интерпретировать, как задачу наведения объекта в заданную окрестность нуля за время, не превышающее фиксированную величину.

3. Эквивалентная задача с терминальным точностным функционалом

Для использования методологии [1], вводится новая координата y_k, динамика которой описывается следующим соотношением.

$$y_{k+1} = \min\{y_k, |x_k^1 + x_k^2 h|\}, \tag{4}$$

$y_0 = |x_0^1|$. Отметим, что система (1), (4) является нелинейной.

Эквивалентная задача имеет вид.

$$P(y_{N+1} \leqslant \varphi) \to \max_{u(\cdot)} \tag{5}$$

(5) относится к классу задач с фиксированным моментом окончания, поэтому решение исходной существует [1] в классе марковских стратегий $u_k = u_k(x_k)$. В [1] также обосновано, что для решения эквивалентной задачи (5) можно применить метод динамического программирования.

В соответствии с алгоритмом динамического программирования функция выигрыша определяется следующим выражением

$$W_k^{\varphi}(x, y) = \sup_{u_k(\cdot),\dots,u_N(\cdot)} P(y_{N+1} \leqslant \varphi | x_k = x, y_k = y),$$

С граничными условиями

$$W_{N+1}^{\varphi}(x_{N+1}, y_{N+1}) = \begin{cases} 1, & \text{при } y_{N+1} \leqslant \varphi, \\ 0, & \text{при } y_{N+1} > \varphi. \end{cases} \tag{6}$$

Тогда для $k = N$ функция выигрыша определяется в результате решения конечномерной задачи

$$W_N^{\varphi}(x_N, y_N) = \max_{u_N} M[W_{N+1}^{\varphi}(x_{N+1}, y_{N+1}) | x_N, y_N].$$

$M[\cdot]$ - математическое ожидание. В правой части последнего выражения аргументы функции $W_{N+1}^{\varphi}(x_{N+1}, y_{N+1})$ преобразуются в соответствии с (1), x_N, y_N фиксированы. С учетом (6) получим

$$W_N^{\varphi}(x_N, y_N) = \max_{u_N} P(\min\{y_N, |x_N^1 + x_N^2 h|\} \leqslant \varphi).$$

Максимизируемая функция в правой части этого выражения не зависит от управления u_N. Поэтому

$$W_N^{\varphi}(x_N, y_N) = \begin{cases} 1, & \text{при } \min\{y_N, |x_N^1 + x_N^2 h|\} \leqslant \varphi, \\ 0, & \text{при } \min\{y_N, |x_N^1 + x_N^2 h|\} > \varphi. \end{cases} \tag{7}$$

Ввиду того, что управление вместе со случайной ошибкой содержится только в соотношении для n-ой координаты фазового вектора, управление будет любым на последних $n-1$ шагах. Функция выигрыша на $k=N-n+1$-ом шаге имеет вид

$$W^{\varphi}_{N-n+1}(x_{N-n+1},y_{N-n+1})=\max_{u_{N-n+1}}P(\min\{y_{N-n+1},|x^1_{N-n+1}+ \\ +x^2_{N-n+1}h|,|x^1_{N-n+1}+2x^2_{N-n+1}h+x^3_{N-n+1}h^2|,...,|\sum_{i=0}^{n-1}C^i_{n-1}h^ix^{i+1}_{N-n+1}|, \\ ,|\sum_{i=0}^{n-1}C^i_nh^ix^{i+1}_{N-n+1}+h^nu_{N-n+1}+h^{n-1}\xi_{N-n+1}|\}\leqslant\varphi), \tag{8}$$

Введем обозначения

$$U_k(x_k,y_k)=\min\{y_k,|x^1_k+hx^2_k|,...,|\sum_{i=0}^{n-1}C^i_{n-1}h^ix^{i+1}_k|\},$$

$$V_k(x_k,u_k)=\sum_{i=0}^{n-1}C^i_nh^ix^{i+1}_k+h^nu_k.$$

Задача оптимизации на $k=\overline{0,N-n+1}$ шаге примет вид

$$P(\min\{U_k(x_k,y_k),|V_k(x_k,u_k)+\xi_kh^{n-1}|\}\leqslant\varphi)\to\max_{u_k}, \tag{9}$$

Лемма 1. *Пусть ξ_k имеет распределение, плотность которого четная и унимодальная функция. Тогда задача (9) эквивалентна задаче*

$$|V_k(x_k,u_k)|\to\min_{u_k}, \tag{10}$$

Доказательство леммы 1.

$$P(\min\{U_k(x_k,y_k),|V_k(x_k,u_k)+\xi_kh^{n-1}|\}\leqslant\varphi)= \\ =1-P(\min\{U_k(x_k,y_k),|V_k(x_k,u_k)+\xi_kh^{n-1}|\}>\varphi)= \\ =1-P(\{U_k(x_k,y_k)>\varphi\}\{|V_k(x_k,u_k)+\xi_kh^{n-1}|>\varphi\}).$$

Отметим, что $U_k(x_k,y_k)$ - детерминированная функция, поэтому рассмотрим случай, когда $U_k(x_k,y_k)>\varphi$. Тогда правая часть последнего выражения равна

$$1-P(|V_k(x_k,u_k)+\xi_kh^{n-1}|>\varphi)= \\ =P(-\varphi-V_k(x_k,u_k)\leqslant\xi_kh^{n-1}\leqslant\varphi-V_k(x_k,u_k))=\int\limits_{-\varphi-V_k(x_k,u_k)}^{\varphi-V_k(x_k,u_k)}f_\xi(t)dt,$$

где $f_\xi(t)$ - плотность распределения случайной величины ξ_kh^{n-1}. Отметим, что последний функционал вероятности относится к так называемому случаю аддитивной функции потерь, описанному в [6]. Запишем детерминированный эквивалент, полученный в [6]

$$|M[\xi_kh^{n-1}]+V_k(x_k,u_k)|\to\min_{u_k}.$$

Последняя задача эквивалентна задаче (9), если плотность распределения случайной величины $\xi_k h^{n-1}$ симметрична и унимодальна относительно $M[\xi_k h^{n-1}]$. Но поскольку по условию леммы $f_\xi(t)$ - четная и унимодальная функция, имеем

$$|V_k(x_k, u_k)| \to \min_{u_k}.$$

Лемма 1 доказана.

Задача (10) является детерминированным эквивалентом задачи стохастического программирования (9). Построение детерминированных эквивалентов для задач стохастического программирования с вероятностными критериями отражено в [6]. Оптимальное управление на k-ом шаге легко находится из (10):

$$u_k = -\frac{1}{h^n}\sum_{i=0}^{n-1} C_n^i h^i x_k^{i+1}, \tag{11}$$

а функция выигрыша запишется в виде

$$W_k^\varphi(x_k, y_k) = \begin{cases} 1, & \text{при } U_k(x_k, y_k) \leqslant \varphi, \\ p(\varphi) \sum\limits_{j=0}^{N-(n+k)} (1 - p(\varphi))^j, & \text{при } U_k(x_k, y_k) > \varphi. \end{cases}$$

где

$$p(\varphi) = P(|\xi_0 h^{n-1}| \leqslant \varphi).$$

Функция выигрыша на $k = 0$ шаге запишется

$$W_0^\varphi(x_0) = \begin{cases} 1, & U_0(x_0) \leqslant \varphi, \\ p(\varphi) \sum\limits_{j=0}^{N-n} (1 - p(\varphi))^j, & U_0(x_0) > \varphi. \end{cases}$$

Далее будем называть ее функцией оптимального выигрыша, поскольку в соответствии с методом динамического программирования

$$W_0^\varphi(x_0) = \max_{u(\cdot)} P_\varphi(u(\cdot)).$$

Отметим, что $W_0^\varphi(x_0)$ и $U_0(x_0)$ не зависят от y_0 в силу начальных условий $y_0 = |x_0^1|$.

4. Задача с квантильным критерием

Квантильный критерий определяется выражением [1]

$$\Phi_\alpha(u) = \min\{\varphi | P_\varphi(u) \geqslant \alpha\}, \tag{12}$$

где $\alpha \in (0, 1)$ - заданная доверительная вероятность. Функционал квантили (12) представляет собой гарантированный с заданной вероятностью уровень указанного выше точностного функционала, т.е. верхнюю доверительную границу для нее.

Рассмотрим задачу минимизации квантильного критерия:

$$\Phi_\alpha(u) \to \min_{u(\cdot)}. \tag{13}$$

Ввиду наличия вероятностного ограничения $P_\varphi(u) \geqslant \alpha$, задачу (13) решить методом динамического программирования нельзя. Подход, использованный в [1] для преодоления этого затруднения в случае дискретных стохастических систем, позволяет получить лишь приближенные (так называемые гарантирующие) решения квантильной задачи в классе управлений, зависящих от всех прошлых состояний системы. При этом вопрос о существовании оптимальных управлений, зависящих только от текущего состояния, остается открытым. В предыдущем разделе было получено аналитическое решение задачи оптимального управления с вероятностным критерием. В [5] предложен метод преобразования этого решения в решение задачи (13). Функция оптимального выигрыша имеет вид

$$F(\varphi) = \begin{cases} 1, & U_0(x_0) \leqslant \varphi, \\ p(\varphi)\sum\limits_{j=0}^{N-n}(1-p(\varphi))^j, & U_0(x_0) > \varphi, \end{cases} \tag{14}$$

где

$$p(\varphi) = P(|\xi_0 h^{n-1}| \leqslant \varphi).$$

Для выполнения условий эквивалентности задач (3) и (13) достаточно [5], чтобы функция $p(\varphi)$ строго возрастала по аргументу.

Теорема 1. *Пусть $p(\varphi)$ строго возрастает по φ. Тогда задача оптимального управления с критерием в форме функционала вероятности (3) эквивалентна задаче оптимального управления с квантильным критерием (13), причем управление, решающее задачу (3), является оптимальным и в задаче (13).*

Доказательство теоремы 1. Воспользуемся теоремой 1. Для справедливости её условий, учитывая что $\alpha \in (0,1)$, достаточно проверить, что функция, стоящая во второй ветви (14) строго возрастает. Имеем:

$$p(\varphi)\sum_{j=0}^{N-n}(1-p(\varphi))^j = p(\varphi)\frac{(1-p(\varphi))^{N-n}-1}{1-p(\varphi)-1} = 1-(1-p(\varphi))^{N-n},$$

Правая часть этого выражения с учетом $p(\varphi) \in (0,1)$ строго возрастает по φ как суперпозиция строго возрастающих функций ($p(\varphi)$ строго возрастает по условию теоремы).

Теорема 1 доказана.

5. Заключение

Рассмотрена задача оптимального управления линейной дискретной стохастической системой по вероятностному критерию с нефиксированным, но ограниченным сверху временем окончания. Путем расширения фазового вектора получена эквивалентная задача оптимального управления более высокой размерности, но с фиксированным моментом окончания. Решение эквивалентной задачи получено в аналитической форме с использованием метода динамического программирования. Приведены достаточные условия, позволяющие получить из этого решения решение аналогичной задачи, но с квантильным критерием качества.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (12-08-00453а), (14-07-00089а).

Список литературы

1. Малышев В.В., Кибзун А.И. Анализ и синтез высокоточного управления летательными аппаратами. М.: Машиностроение, 1987. С. 39-50.
2. Зубов В.И. Лекции по теории управления. М., Наука, 1975. 495 с.
3. Афанасьев В.Н., Колмановский В.Б., Носов В.Р. Математическая теория конструирования систем управления. М.: Высшая школа, 2003.
4. Красовский Н.Н. Об оптимальном регулировании при случайных возмущениях // Прикл. математика и механика. 1960. Т. 24, Вып. 1. С. 64-79.
5. Кан Ю.С. Оптимизация управления по квантильному критерию // Автоматика и телемеханика. 2001. № 5.
6. Вишняков Б.В., Кибзун А.И. Детерминированные эквиваленты для задач стохастического программирования с вероятностными критериями. // Автоматика и телемеханика. 2006. № 6.