Московский Государственный Университет им. М.В. Ломоносова

Механико-математический факультет

Кафедра общих проблем управления

Дипломная работа

О явном решении одной задачи стохастического оптимального управления

Выполнил студент 6 курса 632 группы

отделения математики

Шангин Владимир Викторович

Научный руководитель

Д.ф.-м.н. член-корр. РАН Зеликин М.И.

Москва 2018

Содержание

Введение

В данной работе изучается линейно-квадратичная задача стохастического оптимального управления с параметром. С помощью стохастического принципа максимума найдена пара процессов, удовлетворяющая необходимому условию оптимальности. Затем доказано, что она действительно минимизирует функционал. В приложении приведены результаты компьютерной симуляции.

Постановка задачи стохастического управления

Пусть $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathrm{P})$ – фильтрованное вероятностное пространство, удовлетворяющее обычным условиям (см. [2], [3]), на котором задан одномерный стандартный винеровский процесс $W(t)$. Будем считать, что $\{\mathcal{F}_t\}_{t\geq 0}$ - пополненная фильтрация, порожденная $W(t)$.

Рассмотрим одномерную управляемую систему, задаваемую стохастическим дифференциальным уравнением (СДУ):

(1) $$\begin{cases} dX(t) = b\big(t, X(t), u(t)\big)dt + \sigma\big(t, X(t), u(t)\big)dW(t) \\ X_0 = a \end{cases}, \ t \in [0, T], \ T < \infty$$

и функционал, определенный на решениях это системы:

(2) $$\mathcal{J}(u(\cdot)) = \boldsymbol{E}[\int_0^T z\left(t, X(t), u(t)\right)dt + h(X(T))].$$

Здесь $u : [0, T] \times \Omega \rightarrow U -$ управляющий процесс, $b : [0, T] \times \mathbb{R} \times U \rightarrow \mathbb{R}$, $\sigma : [0, T] \times \mathbb{R} \times U \rightarrow \mathbb{R}$, $z : [0, T] \times \mathbb{R} \times U \rightarrow \mathbb{R}$, $h : \mathbb{R} \rightarrow \mathbb{R}, a -$ константа, $U -$ множество допустимых значений управления.

Допустимое управление $u(\cdot)$ – это процесс удовлетворяющий следующим свойствам:

1. $u(\cdot)$ адаптирован к фильтрации $\{\mathcal{F}_t\}_{t\geq 0}$.
2. $\sup\limits_{t\in[0,T]} \boldsymbol{E}|u(t)|^{\boldsymbol{n}} < \infty \ \forall \, n \in \mathbb{N}$.
3. Существует единственное решение $X(\cdot)$ управляемой системы.
4. Корректно определено значение функционала $\mathcal{J}$.

Обозначим за $\mathcal{U}$ множество допустимых управлений.

Задача стохастического оптимального управления:

Найти (если существует) управляющий процесс $\bar{u}(\cdot) \in \mathcal{U}$, такой что

$$\mathcal{J}\big(\bar{u}(\cdot)\big) = \inf_{u(\cdot)\in\mathcal{U}} \mathcal{J}(u(\cdot)).$$

$\bar{u}(\cdot)$ называется оптимальным управлением. Обозначим за $\bar{X}(\cdot)$ - соответствующее решение СДУ. Пара $(\bar{X}(\cdot), \bar{u}(\cdot))$ – решение задачи (1) - (2).

Стохастический принцип максимума

Теорема (Стохастический принцип максимума, [1, Theorem 3]).

Пусть в задаче стохастического управления функции b, σ, z, h - дважды непрерывно дифференцируемы по x, непрерывны вместе с $b_x, b_{xx}, \sigma_x, \sigma_{xx}, z_x, z_{xx}, h_x, h_{xx}$ по (x, u). Пусть также $b_x, b_{xx}, \sigma_x, \sigma_{xx}, z_{xx}, h_{xx}$ – ограничены, а b, σ, z_x, h_x – ограничены выражением $K(1 + |x| + |u|)$. Если $(\bar{X}(\cdot), \bar{u}(\cdot))$ – решение задачи стохастического оптимального управления, то

1. Найдутся пары процессов

(3) $$\begin{cases} \big(p(\cdot), q(\cdot)\big) \in L^2_{\mathcal{F}}(0, T, \mathbb{R}) \times L^2_{\mathcal{F}}(0, T, \mathbb{R}) \\ \big(P(\cdot), Q(\cdot)\big) \in L^2_{\mathcal{F}}(0, T, \mathbb{R}) \times L^2_{\mathcal{F}}(0, T, \mathbb{R}), \end{cases}$$

являющиеся единственными адаптированными решениями соответственно (4) – (5):

(4) $$\begin{cases} dp(t) = -\big(\bar{b}_x p(t) + \bar{\sigma}_x q(t) - \bar{z}_x\big)dt + q(t)dW(t) \\ p(T) = -h_x\big(\bar{X}(T)\big), \end{cases}$$

где $\bar{b}_x \overset{\text{def}}{=} b_x(t, \bar{X}(t), \bar{u}(t))$ и аналогично для остальных.

(5)

$$\begin{cases} dP(t) = -\big(2\bar{b}_x P(t) + \bar{\sigma}_x^2 P(t) + 2\bar{\sigma}_x Q(t) + \bar{H}_{xx}\big)dt + q(t)dW(t) \\ P(T) = -h_{xx}\big(\bar{X}(T)\big), \end{cases}$$

где $H = H(t, x, u, p, q) \overset{\text{def}}{=} pb(t, x, u) + q\sigma(t, x, u) - z(t, x, u)$, $\bar{H} \overset{\text{def}}{=} H\big(t, \bar{X}(t), \bar{u}(t), p(t), q(t)\big)$

2. Выполняется *условие максимума* по u:

$$\mathcal{H}(t, \bar{X}(t), \bar{u}(t)) = \max_{u \in U} \mathcal{H}(t, \bar{X}(t), u), \forall t \in [0, T], \text{ P} - \text{п. н.},$$

где $\mathcal{H}$- функция, соответствующая набору $(\bar{X}(\cdot), \bar{u}(\cdot), p(\cdot), q(\cdot), P(\cdot), Q(\cdot))$ определяется следующим выражением:

$$\mathcal{H}(t, x, u) \overset{\text{def}}{=} p(t)b(t, x, u) + q(t)\sigma(t, x, u) - z(t, x, u) + \frac{1}{2}\sigma(t, x, u)^2 P(t) - \sigma(t, x, u)\sigma(t, \bar{X}(t), \bar{u}(t))P(t)$$

Применение стохастического принципа максимума

Рассмотрим следующую линейно-квадратичную задачу:

$$(6)\quad \begin{cases} dX(t) = \big(X(t) + u(t)\big)dt + \big(X(t) + \alpha \cdot u(t)\big)dW(t) \\ X_0 = 1,\ t \in [0,1] \end{cases}, U = \mathbb{R}, \alpha > \frac{1+\sqrt{3}}{2}$$

$$(7)\quad \mathcal{J}\big(u(\cdot)\big) = \boldsymbol{E}\left[\frac{1}{2}\int_0^1 (X(t)^2 - 2X(t)u(t) - u(t)^2)\, dt + \frac{1}{2}X(1)^2\right] \to inf.$$

Предположим $(\bar{X}(\cdot), \bar{u}(\cdot))$ – оптимальная пара в задаче (6) - (7) (которую нужно определить). Запишем необходимое условие оптимальности (стохастический принцип максимума):

Существует единственная пара согласованных процессов $(p(\cdot), q(\cdot))$, которая удовлетворяет *обратному* (поскольку задано *терминальное* значение) СДУ, (сопряженному уравнению первого порядка)

$$(8)\quad \begin{cases} dp(t) = -(p(t) + q(t) - \bar{X}(t) + \bar{u}(t))\, dt + q(t)dW(t) \\ p(1) = -\bar{X}(1) \end{cases}, t \in [0,1]$$

Существует единственная пара согласованных процессов $(P(\cdot), Q(\cdot))$, которая удовлетворяет обратному СДУ (сопряженному уравнению второго порядка)

$$(9)\quad \begin{cases} dP_t = -(3P(t) + 2Q(t) - 1)\, dt + Q(t)dW(t) \\ P(1) = -1 \end{cases}, t \in [0,1]$$

Очевидно, что пара $(P(t), Q(t)) = \left(-\frac{4}{3}e^{-3(t-1)} + \frac{1}{3}, 0\right)$ – единственное согласованное решение (9).

$\mathcal{H}$- функция принимает вид:

$$\mathcal{H}(t, \bar{X}(t), u) = \left(\frac{\alpha^2}{2}P(t) + \frac{1}{2}\right)u^2 + (p(t) + \alpha q(t) + \bar{X}(t) - \alpha^2\bar{u}(t)P(t))u + r,$$

где r – не зависит от u. Так как $P(t)$ – возрастающая функция и $P(1) = -1$, то $\frac{\alpha^2}{2}P(t) + \frac{1}{2} < 0\ \forall t \in [0,1]$. Следовательно, $\mathcal{H}$- функция является вогнутой по u (при $t \in [0,1]$). Необходимое условие максимума $\left.\frac{\partial \mathcal{H}}{\partial u}\right|_{u=\bar{u}(t)} = 0$ приводит к следующему уравнению:

$$2\left(\frac{\alpha^2}{2}P(t) + \frac{1}{2}\right)\bar{u}(t) + (p_t + \alpha q_t + \bar{X}(t) - \alpha^2\bar{u}(t)P(t)) = 0,$$

откуда находим выражение для управления, претендующего на роль оптимального:

$$\bar{u}(t) = -(p(t) + \alpha q(t) + \bar{X}(t)).$$

Подставляя полученное значение в (6), (8) и объединяя эти уравнения вместе, получаем систему СДУ с прямым и обратным уравнениями:

$$(10)\ \begin{cases} d\bar{X}(t) = -\big(p(t) + \alpha q(t)\big)dt - \big(\alpha p(t) + \alpha^2 q(t) + (\alpha-1)\bar{X}(t)\big)dW(t) \\ dp_t = ((\alpha-1)q(t) + 2\bar{X}(t))\,dt + q(t)dW(t) \\ \bar{X}(0) = 1,\ p(1) = -\bar{X}(1),\ t \in [0,1] \end{cases}$$

Попробуем решить данную систему уравнений. Терминальное условие для процесса $p(\cdot)$ подсказывает, что разумно предположить выполнение следующих равенств:

$$p(t) = f(t)\bar{X}(t) + g(t), f(1) = -1, g(1) = 0, \forall t \in [0,T], \mathrm{P} - \text{п. н.},$$

где $f(\cdot)$ и $g(\cdot)$ – некоторые детерминированные функции времени.

По формуле Ито имеем:

$$dp(t) = (\dot{f}(t)\bar{X}(t) + \dot{g}(t))dt + f(t)d\bar{X}(t) =$$

$$= \left(\dot{f}(t)\bar{X}(t) + \dot{g}(t) - f(t)^2\bar{X}(t) - f(t)g(t) - \alpha f(t)q(t)\right)dt - f(t)\big(\alpha(f(t)\bar{X}(t) + g(t)) + \alpha^2 q(t) + (\alpha-1)\bar{X}(t)\big)dW(t)\,.$$

Сопоставляя полученное выражение со вторым уравнением системы (10), получаем:

$$(11)\ \begin{cases} \big(\dot{f}(t)\bar{X}(t) + \dot{g}(t) - f(t)^2\bar{X}(t) - f(t)g(t) - \alpha f(t)q(t)\big) = (\alpha-1)q(t) + 2\bar{X}(t) \\ -f(t)\left(\alpha\big(f(t)\bar{X}(t) + g(t)\big) + \alpha^2 q(t) + (\alpha-1)\bar{X}(t)\right) = q(t) \end{cases}$$

Из второго уравнения (11) находим $q(t) = -\frac{\alpha f(t)^2\bar{X}(t) + \alpha f(t)g(t) + (\alpha-1)f(t)\bar{X}(t)}{\alpha^2 f(t) + 1}$.

Подставляя выражение для $q(t)$ в первое уравнение (11), получаем, что функции $f(\cdot)$ и $g(\cdot)$ должны удовлетворять следующим задачам Коши:

(12) $$\begin{cases}\dot f(t)=\frac{(-2\alpha^2+2\alpha+1)f(t)^2+(\alpha^2+2\alpha-1)f(t)+2}{\alpha^2 f(t)+1} \\ f(1)=-1\end{cases},t\in[0,1]$$

(13) $$\begin{cases}\dot g(t)=\frac{(-\alpha^2+\alpha+1)f(t)}{\alpha^2 f(t)+1}g(t) \\ g(1)=0\end{cases},t\in[0,1]$$

Покажем, что уравнение (12) имеет *глобальное* решение на [0,1].

$\frac{1+\sqrt{3}}{2}$ – больший корень уравнения $-2\alpha^2+2\alpha+1=0$. Нули числителя и знаменателя правой части (12) имеет соответственно вид $f_{+,-}=\frac{1-2\alpha-\alpha^2\pm\sqrt{\alpha^4+4\alpha^3+18\alpha^2-20\alpha-7}}{2(1+2\alpha-2\alpha^2)}$ – действительные, $f_0=-\frac{1}{\alpha^2}$.

Интегрируя уравнение, получаем, что оно эквивалентно следующему:

$$ALn|f-f_+|+BLn|f-f_-|=t+C,$$

где $A=\frac{-2-4\alpha+2\alpha^3+\alpha^4-\alpha^2(-3+\sqrt{-7-20\alpha+18\alpha^2+4\alpha^3+\alpha^4})}{2(-1-2\alpha+2\alpha^2)\sqrt{-7-20\alpha+18\alpha^2+4\alpha^3+\alpha^4}}, B=\frac{2+4\alpha-2\alpha^3-\alpha^4-\alpha^2(3+\sqrt{-7-20\alpha+18\alpha^2+4\alpha^3+\alpha^4})}{2(-1-2\alpha+2\alpha^2)\sqrt{-7-20\alpha+18\alpha^2+4\alpha^3+\alpha^4}}$, C находится из терминального условия $f(1)=-1$. Несложно показать, что при всех допустимых значениях параметра $A\le 0, B<0$. Это следует из того, что $\alpha^2\sqrt{-7-20\alpha+18\alpha^2+4\alpha^3+\alpha^4}\ge -2-4\alpha+3\alpha^2+2\alpha^3+\alpha^4$, так как это эквивалентно $4(1+\alpha-\alpha^2)^2(2\alpha^2-2\alpha-1)\ge 0, \alpha>\frac{1+\sqrt{3}}{2}$. Из этого же неравенства следует, что $f_+\le -\frac{1}{\alpha^2}$. К тому же легко заметить, что $f_+>-1$ (это сводится к неравенству $-8(2+\alpha+2\alpha^2-3\alpha^3+3\alpha^4)<0$, которое, очевидно, выполнено при допустимых α).

Заметим, что если глобальное решение существует, то из условий знакопостоянства правой части уравнения необходимо чтобы оно монотонно возрастало и $f(t)<-1$ при $t\in[0,1]$). Зафиксируем допустимое α и рассмотрим функцию $l(t,x)=ALn(f_+-x)+BLn(f_--x)-t-C, t\in[0,1], x\le -1$. Заметим, что если $t\in[0,1)$, то $l(t,-1)=1-t>0$. Далее, $\lim_{x\to-\infty} l(0,x)=-\infty$, поэтому $\exists\, c<-1$ т.ч. $\forall t\in[0,1)$ выполнено $l(t,c)<0$. Поэтому $\forall t\in[0,1)\ \exists!\, x=f(t)\in(c,-1)$ т.ч. $l(t,f(t))=0$. Так как $l(1,-1)=0$, то $f(1)=-1$. Таким образом, искомое решение существует. Очевидно, что уравнение (13) имеет единственное решение $g\equiv 0$.

Далее, $\bar{u}(t) = -\big(p(t) + \alpha q(t) + \bar{X}(t)\big) = -\frac{1+(\alpha+1)f(t)}{1+\alpha^2 f(t)}\bar{X}(t), \forall t \in [0,T], \mathrm{P} -$ п. н., $\bar{u}(\cdot)$ — управление по принципу обратной связи. Таким образом, если решение $(\bar{X}(\cdot), \bar{u}(\cdot))$ задачи (6) - (7) существует, то оптимальное управление имеет найденный вид. Тогда оптимальная траектория удовлетворят СДУ:

$$d\bar{X}(t) = v(t)\bar{X}(t)dt + y(t)\bar{X}(t)dW(t)\ ,\ \bar{X}(0) = 1,$$

где $v(t),\ y(t)$ – некоторые детерминированные функции (зависящие еще от параметра α), решение которого имеет вид (проверяется непосредственным применением формулы Ито):

$$\bar{X}(t) = e^{\int_0^t \left(v(s) - \frac{1}{2}y(s)^2\right)ds + \int_0^t y(s)dW(s)}\ .$$

Покажем, что пара $(\bar{X}(\cdot), \bar{u}(\cdot))$ на самом деле является оптимальной. Пусть $(X(\cdot),\ u(\cdot))$ – произвольная допустимая пара процессов. Применим формулу Ито к $f(\cdot)X(\cdot)^2$:

$d(f(t)X(t)^2) = \Big(\frac{(-2\alpha^2+2\alpha+1)f(s)^2+(\alpha^2+2\alpha-1)f(s)+2}{\alpha^2 f(s)+1}X(t)^2 + 2f(t)X(t)\big(X(t) + u(t)\big) + f(t)(X(t) + \alpha u(t))^2\Big)\,dt + \{\dots\}dW(t)$ или в интегральном виде:

$$f(t)X(t)^2 = f(0) + \int_0^t \Big(\frac{(-2\alpha^2+2\alpha+1)f(s)^2+(\alpha^2+2\alpha-1)f(s)+2}{\alpha^2 f(s)+1}X(s)^2 + 2f(s)X(s)\big(X(s) + u(s)\big) + f(s)(X(s) + \alpha u(s))^2\Big)\,ds + \int_0^t \{\dots\}dW(s)\ .$$

Положим $t = 1$ и возьмем математическое ожидание обеих частей, получаем

$$\boldsymbol{E}{X_1}^2 = -f(0) - \boldsymbol{E}\left[\int_0^1 \Big(\frac{(-2\alpha^2+2\alpha+1)f(s)^2+(\alpha^2+2\alpha-1)f(s)+2}{\alpha^2 f(s)+1}{X_s}^2 + 2f(s)X(s)\big(X(s) + u(s)\big) + f(s)(X(s) + \alpha u(s))^2\Big)\,ds\right].$$

Подставим в выражение для минимизируемого функционала (7):

$$\mathcal{J} = -\frac{1}{2}f(0) - \frac{1}{2}\boldsymbol{E}\left[\int_0^1 (1 + \alpha^2 f(s))\left(u(s) + \frac{1 + (\alpha + 1)f(s)}{1 + \alpha^2 f(s)}X(s)\right)^2 ds\right]$$

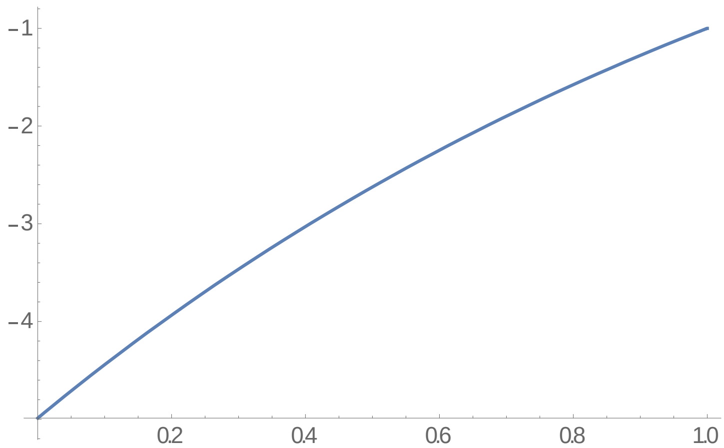
Так как $\alpha^2 f(t) + 1 < 0$ при $t \in [0,1]$, то $\bar{u}(\cdot) = -\frac{1+(\alpha+1)f(t)}{1+\alpha^2 f(t)}\bar{X}(\cdot)$ действительно доставляет минимум функционалу.

Приложение

Пусть $\alpha = 2$, тогда имеем:

$$\begin{cases} \dot{f}(t) = \dfrac{-3f(t)^2 + 7f(t) + 2}{4f(t) + 1} \\ f(1) = -1 \end{cases}, t \in [0,1]$$

График $f(t)$ приведен ниже:

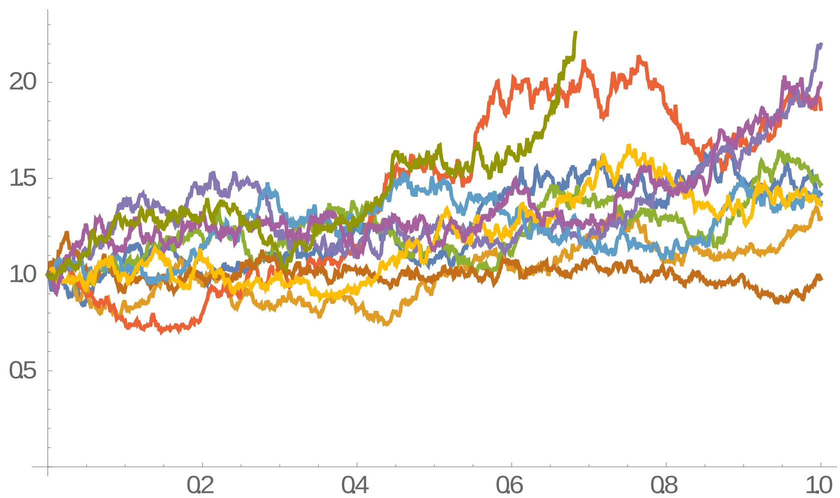


$$\bar{u}(t) = -\frac{3f(t) + 1}{4f(t) + 1}\bar{X}(t), \forall t \in [0, T], \mathrm{P} - \text{п. н}$$

Приведем компьютерную симуляцию (10 выборочных траекторий) процесса, являющегося решением СДУ:

$$d\bar{X}(t) = v(t)\bar{X}(t)dt + y(t)\bar{X}(t)dW(t)\ ,\ \bar{X}(0) = 1,$$

где $v(t) = \frac{f(t)}{4f(t)+1}, y(t) = -\frac{2f(t)+1}{4f(t)+1}$.



Список литературы

[1] Peng, S. A general stochastic maximum principle for optimal control problems, SIAM J. Control & Optim., 28 (1990), 966-979

[2] Karatzas I., Shreve S.E. Brownian Motion and Stochastic Calculus, 2nd ed. Springer, 1991.

[3] Жакод Ж., Ширяев А.Н. Предельные теоремы для случайных процессов. Т. 1, 2. М.: Физматлит, 1994.