

УДК 519.857.6

ЧИСЛЕННАЯ РЕАЛИЗАЦИЯ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ ВЫХОДОМ СТОХАСТИЧЕСКОЙ ДИФФЕРЕНЦИАЛЬНОЙ СИСТЕМЫ ПО КВАДРАТИЧЕСКОМУ КРИТЕРИЮ

А.В. Босов

Институт проблем информатики ФИЦ ИУ РАН

Россия, 119333, Москва, Вавилова ул., 44/2

E-mail: avbosov@frccsc.ru

А.И. Стефанович

Институт проблем информатики ФИЦ ИУ РАН

Россия, 119333, Москва, Вавилова ул., 44/2

E-mail: astefanovich@frccsc.ru

Ключевые слова: стохастическое дифференциальное уравнение, оптимальное управление, динамическое программирование, функция Беллмана, уравнение Риккати, линейные уравнения параболического типа.

Аннотация: В докладе представлены решение задачи управления по квадратическому критерию качества для диффузионного процесса Ито и линейного управляемого выхода и его численная реализация. Оптимальное управление определяется функцией Беллмана, параметризованной коэффициентами, описываемыми дифференциальными уравнениями. Численная реализация обеспечивается приближенным решением этих уравнений сеточными методами. Расчеты иллюстрируются модельным примером, основанным на простой дифференциальной модели для показателя RTT сетевого протокола TCP.

1. Введение

Принципиальные классические результаты по управлению стохастическими динамическими системами получены более 40 лет назад (см., например, работу [1]). Прежде всего, это касается задач с линейными моделями и квадратическим критерием качества. Несмотря на множество последующих достижений как в части моделирования, так и в части представления целей и применяемых методов исследования, привлекательные свойства линейных моделей и квадратического критерия не исчерпаны и эксплуатируются в данной работе. Основные усилия по развитию теории направлены на то, чтобы сделать модель нелинейной. Для этого предполагается, что стохастическая динамическая система описывается двумя уравнениями: нелинейным дифференциальным уравнением Ито без управления задается состояние, управляемый выход описывается линейным уравнением. Для решения задачи используется

метод динамического программирования, решается уравнение Беллмана [2], техническую проблему при этом составляет поиск верной параметризации функции Беллмана. Подходящая структура этой функции обеспечивает конечные соотношения для оптимального решения, параметры которых определяются решениями конкретных дифференциальных уравнений (обыкновенных и в частных производных). Решение этих уравнений и модельный пример составляют содержание данного доклада.

2. Постановка задачи управления выходом

Для простоты далее будем предполагать рассматриваемые случайные функции скалярными. Состояние y_t описывается нелинейным стохастическим дифференциальным уравнением Ито

$$(1) \quad dy_t = A_t(y_t) dt + \Sigma_t(y_t) dv_t, \quad y_0 = Y,$$

где v_t — стандартный винеровский процесс, Y — случайная величина с конечным вторым моментом, функции A_t, Σ_t удовлетворяют условиям Ито, обеспечивающим существование единственного решения [2].

С состоянием y_t связан выход z_t :

$$(2) \quad dz_t = a_t y_t dt + b_t z_t dt + c_t u_t dt + \sigma_t dw_t, \quad z_0 = Z,$$

где w_t — независимый от v_t, Y, Z стандартный винеровский процесс, Z — случайная величина с конечным вторым моментом, u_t — допустимое управление. Функции a_t, b_t, c_t, σ_t предполагаются ограниченными, процесс управления — допустимым неупреждающим [2], что обеспечивает существование решения уравнения (2) для любого допустимого управления.

Используется целевой функционал следующего вида:

$$(3) \quad J(U_0^T) = E \left\{ \int_0^T (S_t(s_t y_t - g_t z_t - h_t u_t)^2 + G_t z_t^2 + H_t u_t^2) dt + \right. \\ \left. + S_T(s_T y_T - g_T z_T)^2 + G_T z_T^2 \right\},$$

$$U_0^T = \{u_t, 0 \leq t \leq T\},$$

где S_t, G_t, H_t — неотрицательные функции.

3. Оптимальное решение

Решение задачи поиска u_t^* — допустимого управления, доставляющего минимум квадратическому функционалу $J(U_0^T)$, составляют следующие соотношения. Функция Беллмана $V_t(y, z)$ может быть представлена в виде

$$(4) \quad V_t(y, z) = \alpha_t z^2 + \beta_t(y) z + \gamma_t(y),$$

а оптимальное управление

$$(5) \quad u_t^* = u_t^*(y, z) = -\frac{1}{2} (S_t h_t^2 + H_t)^{-1} (c_t (2\alpha_t z + \beta_t(y)) + 2S_t (s_t y - g_t z) h_t),$$

$$y = y_t, \quad z = z_t.$$

Коэффициенты $\alpha_t, \beta_t(y), \gamma_t(y)$ задаются следующими дифференциальными уравнениями:

$$(6) \quad \frac{\partial \alpha_t}{\partial t} + 2\alpha_t \left(b_t - (S_t h_t^2 + H_t)^{-1} c_t S_t h_t g_t \right) + \left(S_t - (S_t h_t^2 + H_t)^{-1} S_t^2 h_t^2 \right) g_t^2 + \\ + G_t - (S_t h_t^2 + H_t)^{-1} c_t^2 \alpha_t^2 = 0,$$

$$\alpha_T = S_T g_T^2 + G_T,$$

$$(7) \quad \frac{\partial \beta_t(y)}{\partial t} + \frac{1}{2} \Sigma_t^2(y) \frac{\partial^2 \beta_t(y)}{\partial y^2} + A_t(y) \frac{\partial \beta_t(y)}{\partial y} + \\ + 2\alpha_t \left(a_t + (S_t h_t^2 + H_t)^{-1} c_t S_t h_t s_t \right) y + \beta_t(y) \left(b_t - (S_t h_t^2 + H_t)^{-1} c_t S_t h_t g_t \right) - \\ - 2 \left(S_t - (S_t h_t^2 + H_t)^{-1} S_t^2 h_t^2 \right) s_t g_t y - (S_t h_t^2 + H_t)^{-1} c_t^2 \alpha_t \beta_t(y) = 0,$$

$$\beta_T(y) = -2S_T s_T g_T y,$$

$$(8) \quad \frac{\partial \gamma_t(y)}{\partial t} + \frac{1}{2} \Sigma_t^2(y) \frac{\partial^2 \gamma_t(y)}{\partial y^2} + \sigma_t^2 \alpha_t + A_t(y) \frac{\partial \gamma_t(y)}{\partial y} + \\ + \beta_t(y) \left(a_t + (S_t h_t^2 + H_t)^{-1} c_t S_t h_t s_t \right) y + \left(S_t - (S_t h_t^2 + H_t)^{-1} S_t^2 h_t^2 \right) s_t^2 y^2 - \\ - \frac{1}{4} (S_t h_t^2 + H_t)^{-1} c_t^2 \beta_t^2(y) = 0,$$

$$\gamma_T(y) = S_T s_T^2 y^2.$$

Уравнение (6) является уравнением Риккати и имеет единственное неотрицательное решение для всех $0 \leq t \leq T$, т.к. предполагается $S_t h_t^2 + H_t > 0$. Уравнения (7) и (8) являются линейными уравнениями в частных производных второго порядка, относятся к параболическому типу, поскольку $\Sigma_t^2(y) > 0$. Предполагается, что данные уравнения имеют на рассматриваемом интервале $0 \leq t \leq T$ хотя бы одно ограниченное решение, которое и определяет оптимальное решение рассматриваемой задачи оптимизации. Подробное обсуждение решения имеется в работе [3].

4. Модельный эксперимент и приближенное численное решение

Практическая реализация оптимального управления (5), конечно, возможна только приближенно в результате численного интегрирования уравнений (6) и (7), а при необходимости и (8). В качестве модельного примера из потенциальной прикладной области выбрана простая модель для показателя RTT (Round-Trip Time) сетевого протокола TCP (Transmission Control Protocol), предложенная в [4]:

$$(9) \quad dy_t = (1 - 0,1y_t) dt + 0,5\sqrt{y_t} dv_t, \quad y_0 = Y \sim N(15, 9).$$

Здесь $N(M, D)$ — нормальное распределение со средним M и дисперсией D .

Выход для (9) задается уравнением

$$(10) \quad dz_t = 0,1y_t dt - z_t dt + u_t dt + dw_t, \quad z_0 = Z \sim N(9, 9),$$

целевой функционал

$$(11) \quad J(U_0^T) = E \left\{ \int_0^T ((y_t - z_t)^2 + z_t^2 + u_t^2) dt + (y_T - z_T)^2 + z_T^2 \right\}.$$

Моделировалось $N = 1000$ траекторий y_t, z_t^*, u_t^* для $T = 50$ (через z_t^* обозначен выход (10), рассчитанный для $u_t = u_t^*$). Кроме того, моделировались траектории $z_t^{\text{prog}}, u_t^{\text{prog}}$ для наилучшего программного управления $u_t^{\text{prog}} = E\{u_t^*\}$ и z_t^0, u_t^0 для неуправляемого выхода (10), т.е. для $u_t = 0$. В каждом случае путем осреднения по пучку траекторий оценивались величины $J(U_0^t)$. Соответственно, в результате можно увидеть не только конечные значения целевой функции для разных вариантов управления $J((U^*)_0^T), J((U^{\text{prog}})_0^T), J((U^0)_0^T)$, но и их формирование в динамике.

Для определения оптимального управления численно решались уравнения для α_t и $\beta_t(y)$. Сначала было получено численное решение уравнения Риккати (6) неявным методом Эйлера. Уравнение (7) решалось методом конечных разностей с использованием явной и неявной численных схем для двух видов граничных условий: Дирихле $\beta_t(y) = 0$, т.е. условие поглощения, и Неймана $\frac{\partial \beta_t(y)}{\partial y} = 0$, т.е. условия отражения, для всех граничных точек. Наличие четырех вариантов интегрирования уравнения (7) позволило подтвердить предположение об устойчивости решения $\beta_t(y)$ в отношении граничных условий, физические основания для задания которых отсутствуют. Шаги интегрирования δ_t, δ_y при этом выбирались разными, например, явная схема обсчитывалась для $\delta_t = 0,0002$ и $\delta_y = 0,1$, окончательный расчет для неявной схемы выполнен с шагами $\delta_t = 0,001$ (в т.ч. для приближенного вычисления α_t) и $\delta_y = 0,01$. Результаты расчетов представлены на рис. 1, в частности, приведены примеры траекторий для выхода и управлений $z_t^*, u_t^*, z_t^{\text{prog}}, u_t^{\text{prog}}$ и z_t^0, u_t^0 .

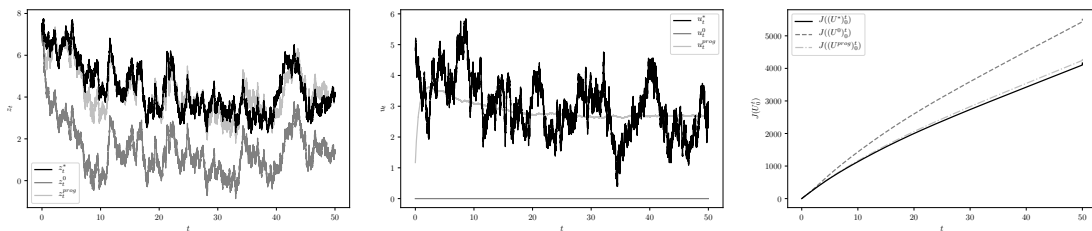


Рис. 1: Выборочные траектории и динамика целевого функционала

Результаты демонстрируют ожидаемый проигрыш неуправляемой системы z_t^0, u_t^0 , высокое качество программного управления и выхода $z_t^{\text{prog}}, u_t^{\text{prog}}$, объясняемое выбором параметров рассматриваемого модельного примера, сделанного с для представления целевого функционала для разных стратегий управления в одном масштабе.

5. Заключение

Представленный результат имеет определенные перспективы как теоретического развития, так и поиска приложений. Но приоритетного изучения требует то обстоятельство, что численная реализация полученного решения сопровождалась определенными вычислительными трудностями. Так, отсутствие физически обоснованных граничных условий потребовало значительных усилий для анализа численного решения на границе. Кроме того, очевидное обобщение постановки на многомерный случай вычислительные проблемы еще более усугубит. По этой причине ближайшим актуальным развитием задачи должны стать более действенные инструменты в части численной реализации полученных точных решений.

Список литературы

1. Athans M. Editorial on the LQG Problem // IEEE Trans. Automat. Control, 1971. Vol. 16, No. 6. P. 528-552.
2. Fleming W.H., Rishel R.W. Deterministic and Stochastic Optimal Control. New York: Springer, 1975.
3. Босов А.В., Стефанович А.И. Управление выходом стохастической дифференциальной системы по квадратичному критерию. I. Оптимальное решение методом динамического программирования // Информатика и ее применения. 2018. Т. 12, Вып. 3. С. 99-106.
4. Bohacek S., Rozovskii B. A diffusion model of roundtrip time // Computational Statistics & Data Analysis. 2004. Vol. 45, No. 1. P. 25-50.