УДК 519.246.2: 681.518.22

ОПЕРАТИВНО-ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ СТОХАСТИЧЕСКОЙ СИСТЕМОЙ ПО ЕЕ ВЫХОДУ

Е.А. Руденко
Московский авиационный институт (национальный исследовательский университет)
Россия, 125871, Москва, Волоколамское ш., 4
E-mail: rudenkoevg@yandex.ru

Ключевые слова: дифференциальная стохастическая система, неточные измерения вектора состояния, переменный критерий оптимальности, фильтры Стратоновича и Калмана, безынерционный регулятор состояния

Аннотация: рассматривается задача синтеза оптимального в среднем закона управления дифференциальным объектом, если переменные его состояния измеряются со случайными погрешностями. Используя метод апостериорных достаточных координат, вместо известного интервально-оптимального регулятора Мортенсена получен существенно более простой алгоритм нахождения его оперативно-оптимального аналога. Новый регулятор не требует решения в обратном времени соответствующего уравнения Беллмана, так как оптимален в смысле переменного во времени критерия. Это позволяет не учитывать информацию о будущем поведении системы, сводя процедуру нахождения зависимости управления от достаточных координат к интегрированию в прямом времени уравнения Фоккера–Планка–Колмогорова и к решению задачи параметрического нелинейного программирования. Применение полученного алгоритма демонстрируется на примере линейно-квадратично-гауссовской задачи, в результате чего сформулирована новая оперативная версия известной теоремы разделения.

1. Введение

Известным решением задачи синтеза управления, оптимального в среднем и на определенном *интервале* времени, является динамическое преобразование неточных измерений в управление [1-3]. Оно состоит из инерционного стохастического фильтра Стратоновича, который накапливает информацию об измерениях, преобразуя их в апостериорную плотность вероятности вектора состояния объекта, и безынерционного детерминированного регулятора Мортенсена, который в каждый момент времени вычисляет управление по виду апостериорной плотности. Однако для этого необходимо решать весьма сложное уравнение Беллмана–Мортенсена в вариационных производных Фреше. Только замена апостериорной плотности достаточными координатами приводит к решению уравнения Беллмана в частных производных.

В данной работе предлагается более простая процедура синтеза регулятора, который тоже использует апостериорные достаточные координаты, но оптимален в несколько другом смысле. Традиционный *интервальный* критерий оптимальности, обеспечивающий управление объектом на всем заранее заданном интервале времени, а потому от времени не зависящий, заменяется похожим *оперативным* критерием, зависящим от времени. Это позволяет учитывать информацию только о прошлом и текущем состоянии объекта и не пересчитывать закон управления при чем-либо вызванных изменениях уравнения объекта в будущем. Кроме того, переменный критерий оптимальности существенно упрощает процедуру синтеза регулятора, позволяя отказаться от решения уравнения Беллмана. В частности, в доказанной ниже оперативной теореме разделения коэффициент усиления соответствующего линейного позиционного регулятора оказывается известным из критерия оптимальности.

2. Постановка задачи синтеза оперативного управления

Пусть $t \geq 0$ – время, $\boldsymbol{X_t}$ – *n*-мерный вектор состояния объекта управления, $\boldsymbol{Y_t}$ – *m*-мерный вектор его измеряемого выхода, $\boldsymbol{U_t}$ – $\boldsymbol{l}$-мерный вектор управления из в общем случае ограниченной области , V_t – *k*-мерный вектор гауссовского белого шума. Известны уравнения Ито состояния объекта и измерителя этого состояния

(1) $$\dot{X}_t = a(t, X_t, U_t) + B(t, X_t, U_t)\, V_t, X_0 \sim p_0(x),$$
$$\dot{Y}_t = c(t, X_t, U_t) + D(t, U_t)\, V_t, Y_0 = 0.$$

Требуется найти не упреждающую зависимость управления U_t системой (1) от всех предыдущих измерений и уже выполненных управлений как их функционал

(2) $$U_t = \vartheta(t, Y_0^t, U_0^t)\,, Y_0^t = \{Y \;|\; \in [0, t)\}, U_0^t = \{U \;|\; \in [0, t)\},$$

который обеспечивает минимум некоторому критерию качества управления.

Обычно минимизируют постоянный во времени критерий, определяющий потери на управление на всем заданном отрезке $t \in [0, T]$ времени управления

(3) $$J[\vartheta(\cdot)] = \mathrm{M}\left[\int_0^T (\;, X\;, U\;) d\; + \;(X_T)\right] \to \min.$$

Будем называть его *интервальным* (И-критерий). Однако получаемое с его помощью управление (2) в каждый момент времени *t* требует информации о функционировании системы и в будущем, на интервале времени $(t, T]$.

Поэтому рассмотрим переменный во времени критерий

(4) $$I_t[\vartheta(\cdot)] = \mathrm{M}\left[\int_0^t (\;, X\;, U\;) d\; + \;(t, X_t)\right] \to \min,$$

который назовем *оперативным* (О-критерий). В отличие от (3) он уже не штрафует будущее поведение системы и, как показало исследование его детерминированного аналога [4], соответствующий регулятор не использует информации об этом.

3. Достаточные координаты и уравнение фильтра

Наиболее качественный синтез оптимального закона управления по выходу основан на использовании *апостериорной плотности вероятности* (АПВ) $(t, x|Y_0^t, U_0^t)$, определяющей состояние нелинейного фильтра Стратоновича. Переход к вектору S_t *достаточных координат* (ДК) [1, 2] дает $(t, x|Y_0^t, U_0^t) = (t, x, S_t)$.

Для рассматриваемого объекта управления (1), (2) уравнение состояния фильтра Стратоновича в форме ДК и его начальное условие имеют вид [1]

(5) $$\dot{S}_t = h(t, U_t, S_t) + (t, U_t, S_t)[\dot{Y}_t - \hat{c}(t, U_t, S_t)], S_0 = \bar{\ }_0 = \int (x) p_0(x) dx,$$

причем структурные функции этого уравнения определяются по формулам

$$h(t, u, s) = \int K_x^{*u}[\;(x)]\;(t, x, s) dx,$$

$$(t, u, s) = \left[(\;\widehat{\;c}^{\mathrm{T}} - \hat{\ }\hat{c}^{\mathrm{T}}) + \int (\;_x\;)\;(t, x, s) dx\right] R^{-1}(t, u),$$

где $K_x^{*u} = a^{\mathrm{T}} \nabla_x + 0.5\,\mathrm{tr}[Q\, \nabla_x \nabla_x]$ – обратный производящий оператор процесса X_t, $Q = BB^{\mathrm{T}}$, $R = DD^{\mathrm{T}}$, $= BD^{\mathrm{T}}$ – матрицы условных интенсивностей белых возмущений векторов X_t, Y_t, а символ ^ над именем функции означает ее апостериорное среднее.

4. Оперативно-оптимальный регулятор

С помощью ДК управление U_t можно искать не как инерционный функционал (2), а в виде безынерционной зависимости $U_t = u(t, S_t)$. Тогда, исключая переменную U_t из уравнений (1), (5), получим замкнутую систему из двух уравнений Ито для марковской пары (X_t, S_t). Ее плотность вероятности $r(t,x,s)$ удовлетворяет уравнению *Фоккера–Планка–Колмогорова* (ФПК) или соответствующему ему тождеству [5]

$$\frac{d}{dt}\iint \quad r\,dxds = \iint \left(\frac{\partial}{\partial t} + K_{xs}^{*u}[\quad]\right) r\,dxds \ \forall \quad (t,x,s) \in \mathbb{C}^{1,2,2}$$

с начальным условием $r(0,x,s) = p_0(x) \quad (s - \bar{\ }_0)$ и с сопряженным оператором

$$K_{xs}^{*u}[\quad] = a^{u\mathrm{T}} \quad _x + e^{u\mathrm{T}} \quad _s + 0.5\,\mathrm{tr}[Q^u \quad _{xx}] +$$
$$+0.5\,\mathrm{tr}[2 \quad ^u \quad ^{u\mathrm{T}} \quad _{xs} + \quad ^u R^u \quad ^{u\mathrm{T}} \quad _{ss}].$$

Здесь $e^u(t,x,s) = h^u(t,s) + \quad ^u(t,s)[c^u(t,x,s) - \hat{c}^u(t,s)], F^u(t,s) = \quad ^u(t,s)D^u(t,s)$, а верхним индексом u отмечены сложные функции, например $a^u(t,x,s) = a(t,x,u(t,s))$.

Потребуем от искомого закона управления оптимальности в смысле О-критерия (4), который в результате использования плотности $r(t,x,s)$ принимает вид

$$I_t = \int_0^t \langle \quad ^u(\quad, x, s), r(\quad, x, s)\rangle d \quad + \langle \quad (t,x), r(t,x,s)\rangle,$$

где $\langle \quad , r\rangle = \iint \quad (t,x,s) r(t,x,s)\,dxds$. Для его минимизации используем метод припасовывания, состоящий в оптимизации приращения этого критерия за каждый сколь угодно малый промежуток времени $[t, t + \quad t]$, $\quad t \downarrow 0$. Последнее, как показано в [5], сводится к минимизации скорости изменения этого функционала в любой момент времени $\dot{I}_t \to \min, \forall t \geq 0$, что известно как *условие локальной оптимальности*. Заметим, что вследствие неотрицательности самого критерия $I_t \geq 0$ эта скорость ограничена снизу $\dot{I}_t > -\infty$, так что ее минимум существует.

Используя тождество ФПК, находим вид производной критерия $\dot{I}_t = \langle \quad ^u + \partial \quad /\partial t + K_{xs}^{*u}[\quad], r\rangle$. Искомая функция входит в это выражение только в виде ее сечения $u(t,\cdot)$ при фиксированном *t*. Слагаемое $\langle \partial \quad /\partial t, r\rangle$ от него не зависит, поэтому достаточно минимизировать функционал $\quad _t[u(t,\cdot)] = \langle \quad ^u + K_{xs}^{*u}[\quad], r\rangle$. Учитывая в нем вид оператора K_{xs}^{*u} и независимость дифференцируемой им функции текущих потерь $\quad (t,x)$ от переменной *s*, имеем $\quad _t[u(t,\cdot)] = \left\langle \quad ^u + a^{u\mathrm{T}} \quad _x + 0.5\,\mathrm{tr}[Q^u \quad _{xx}], r\right\rangle$. Заменив здесь совместную плотность вероятности $r(\cdot)$ произведением маргинальной на условную $r(t,x,s) = q(t,s) \quad (t,x|s), q(t,s) = \int r(t,x,s)\,dx$, получаем функционал маргинального среднего $\quad _t[u(t,\cdot)] = \int \quad [t,s;u(t,s)]\,q(t,s)ds$ от функции условного среднего

(6) $$\quad (t,s;\,u) = \int(\quad + a^{\mathrm{T}} \quad _x + 0.5\,\mathrm{tr}\,[Q \quad _{xx}]) \quad dx.$$

Тогда из свойства монотонности операции интегрирования и неотрицательности плотности вероятности $q(\cdot) \geq 0$ следует [5], что для отыскания минимума функционала $\quad _t[\cdot]$ достаточно найти минимум его маргинально усредняемой функции условного среднего $\quad (\cdot)$ по одному из ее аргументов при любых значениях других

(7) $$u(t,s) = \operatorname*{argmin}_{v \in \quad \subset \mathbb{R}^l} \quad (t,s;v), \forall t,s.$$

При этом функция $\quad (\cdot)$ требует знания условной плотности вероятности $\quad (\cdot)$. Она определяется по получаемой решением задачи Коши для тождества ФПК совместной плотности $r(\cdot)$ как отношение $\quad (t,x|s) = r(t,x,s)/\int r(t,x,s)\,dx$.

5. Два решения ЛКГ-задачи

Пусть уравнения объекта (1) и измерителя (2) *линейные*

$$\dot{X}_t = A(t)X_t + K(t)U_t + B(t)V_t, \dot{Y}_t = C(t)X_t + M(t)U_t + D(t)dV_t,$$

начальное состояние *гауссовское* $X_0 \sim N(x||m_0^x, D_0^x)$, управление не ограничено $U_t \in \quad = \mathbb{R}^l$, а функции критериев (3), (4) *квадратические*

(8) $\quad = \quad = 0.5\,[x^{\mathrm{T}} \quad (t)\,x + u^{\mathrm{T}} \quad (t)\,u], \quad = 0.5\,x^{\mathrm{T}} \quad x, \quad = 0.5\,x^{\mathrm{T}} \quad (t)\,x.$

В этом случае АПВ является гауссовской $\quad (t, x, S_t) = N(x||\hat{X}_t, P_t)$, так что вектор ДК состоит только из двух компонент $S_t = (\hat{X}_t, P_t)$. При этом случайная оценка $\hat{X}_t$ состояния X_t определяется уравнением фильтра Калмана–Бьюси, а детерминированная матрица ковариаций ошибки оценивания $P_t = \mathrm{cov}(X_t - \hat{X}_t)$ находится из не зависящего ни от оценки $\hat{X}_t$, ни от управления U_t уравнения Риккати. Это делает P_t заранее известной функцией времени, что позволяет исключить ее из состава ДК, считая $S_t = \hat{X}_t$, и искать управление как функцию оценки $U_t = u(t, \hat{X}_t)$.

Для нее при квадратическом И-критерии (3), (8) известен результат, полученный аналитическим решением соответствующего уравнения Беллмана.

Теорема 1 [1, 6]. *В интервальной ЛКГ-задаче оптимальное устройство управления (2) разделяется на фильтр Калмана–Бьюси с матрицей ковариации, получаемой решением прямой задачи Коши для уравнения Риккати, и интервально-оптимальный линейный детерминированный регулятор, настраиваемый на весь отрезок времени управления $t \in [0, T]$ решением обратной задачи Коши для своего уравнения Риккати.*

Если же оптимизировать квадратический О-критерий (4), (8), то достаточно найти частный минимум (7) условного среднего (6), который в этом случае имеет вид

$$(t, \hat{x};\, u) = \int (0.5\,(x^{\mathrm{T}} \quad x + u^{\mathrm{T}} \quad u) + (Ax + Ku)^{\mathrm{T}} \quad x + 0.5\,\mathrm{tr}\,[Q \quad]) \quad (t, x|\hat{x})\,dx,$$

причем от переменной управления u в нем зависят только два слагаемых

$$(t, \hat{x};\, u) = 0.5\,u^{\mathrm{T}} \quad u + u^{\mathrm{T}} K^{\mathrm{T}} \quad \bar{x}(t, \hat{x}) + \mathrm{invar}(u), \bar{x}(t, \hat{x}) = \int x \quad (t, x|\hat{x}) dx.$$

Так как $\quad > 0$, то минимум этой функции достигается на $u(t, \hat{x}) = - \quad ^{-1} K^{\mathrm{T}} \quad \bar{x}(t, \hat{x})$.

При этом совместная плотность гауссовская $r(t, x, \hat{x}) = N(x, \hat{x}||m_t^x, m_t^{\hat{x}}, D_t^x, D_t^{\hat{x}}, D_t^{x\hat{x}})$. Тогда по теореме о нормальной корреляции гауссова и условная плотность $\quad (t, x|\hat{x}) = N\,[x||\bar{x}(t, \hat{x}), \quad (t)]$, а ее среднее и ковариация находятся по формулам $\bar{x}(t, \hat{x}) = m_t^x + D_t^{x\hat{x}} D_t^{\hat{x}\,\oplus} (\hat{x} - m_t^{\hat{x}})$, $\quad (t) = D_t^x - D_t^{x\hat{x}} D_t^{\hat{x}\,\oplus} D_t^{\hat{x}x}$. Используя здесь известные свойства несмещенности оценки $\quad [\hat{X}_t] = \quad [X_t]$ и ее ортогональности к ошибке оценивания $\mathrm{cov}(X_t - \hat{X}_t, \hat{X}_t) = 0$, имеем $m_t^{\hat{x}} = m_t^x$, $D_t^{x\hat{x}} = D_t^{\hat{x}}$. В результате $\bar{x}(t, \hat{x}) = \hat{x}$, $\quad (t) = D_t^x - D_t^{\hat{x}}$ и искомая минималь принимает требуемый вид $u(t, \hat{x}) = - \quad ^{-1} K^{\mathrm{T}} \quad \hat{x}$.

Теорема 2. *В оперативной ЛКГ-задаче оптимальное устройство управления (2) разделяется на линейный фильтр Калмана–Бьюси и линейный детерминированный регулятор с известной из условий задачи матрицей усиления* $\quad ^{-1}(t) K^{\mathrm{T}}(t) \quad (t)$.

6. Заключение

Решена новая задача синтеза оперативно получаемого темпе со временем закона оптимального в среднем управления нелинейным стохастическим дифференциальным объектом по его измеряемому выходу. Она обобщает задачу синтеза локально-оптимального управления детерминированным объектом по точным измерениям его состояния [4]. При этом также используется переменный во времени критерий оптимальности, который минимизирует потери, накопленные к текущему моменту времени и не учитывает их будущие значения. Это позволяет не учитывать и возможные в будущем изменения параметров и структуры управляемого объекта.

Процедура синтеза продемонстрирована на примере ЛКГ-задачи, в результате чего получена оперативная версия известной теоремы разделения.

Список литературы

1. Параев Ю.И. Введение в статистическую динамику процессов управления и фильтрации. М.: Сов. Радио, 1976.
2. Стратонович Р.Л. К теории оптимального управления. Достаточные координаты // Автоматика и тлемеханика. 1962. № 7. С. 910-917.
3. Mortensen R.E. Stochastic Optimal Control with Noisy Observations // Int. J. Control. 1966. Vol. 4, No. 5. P. 455–466.
4. Верба B.C., Меркулов В.И., Руденко Е.А. Линейно-кубическое локально-оптимальное управление линейными системами и его применение для наведения летательных аппаратов // Изв. РАН. ТиСУ. 2020. № 5. С. 129-141.
5. Руденко Е.А. Оптимальная структура непрерывного нелинейного фильтра Пугачева пониженного порядка // Изв. РАН. ТиСУ. 2013. № 6. С. 25-51
6. Wonham W.M. On the Separation Theorem of Stochastic Control // SIAM J. Control. 1968. Vol. 6, No. 2. P. 312-326.