

**УПРАВЛЕНИЕ В СТОХАСТИЧЕСКИХ СИСТЕМАХ
И В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ**

УДК 681.514.5

**ОПЕРАТИВНОЕ АБСОЛЮТНО ОПТИМАЛЬНОЕ ДИНАМИЧЕСКОЕ
УПРАВЛЕНИЕ СОСТОЯНИЕМ СТОХАСТИЧЕСКОГО
ДИФФЕРЕНЦИАЛЬНОГО ОБЪЕКТА ПО ЕГО ВЫХОДУ**

© 2023 г. Е. А. Руденко^{а,*}

^аМАИ (национальный исследовательский ун-т), Москва, Россия

*e-mail: rudenkoevg@yandex.ru

Поступила в редакцию 08.10.2022 г.

После доработки 10.10.2022 г.

Принята к публикации 05.12.2022 г.

Рассматривается задача синтеза оптимального в среднем закона управления динамическим объектом, который подвержен действию случайных возмущений, если переменные его состояния измеряются частично или со случайными погрешностями. Используя метод апостериорных достаточных координат, описана сложность построения известного интервально-оптимального регулятора Мортенсена и 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 объекта и сводит процедуру нахождения зависимости управления от достаточных координат к интегрированию в прямом времени уравнения типа Фоккера–Планка–Колмогорова и к решению задачи параметрического нелинейного программирования. Применение полученного алгоритма демонстрируется на примере линейно-квадратично-гауссовской задачи, в результате решения которой сформулирована новая оперативная версия известной теоремы разделения. Она представляет стохастическое устройство управления как соединение линейного фильтра Калмана–Бьюси и линейного оперативно-оптимального позиционного регулятора. Последний отличается от традиционного интервально-оптимального регулятора известностью своего коэффициента усиления и не требует решения в обратном времени соответствующего матричного уравнения Риккати.

DOI: 10.31857/S0002338823020166, EDN: JDIVKJ

Введение. Как известно, решением задачи синтеза управления объектом, оптимального в среднем и на определенном *интервале* времени, в случае отсутствия точных измерений всех переменных его состояния является динамическое преобразование измерений в управление [1–6]. Оно состоит из инерционного стохастического фильтра Стратоновича, который накапливает информацию об измерениях, преобразуя их в апостериорную плотность вероятности неизмеряемого вектора состояния объекта, и безынерционного детерминированного регулятора Мортенсена, который в каждый момент времени вычисляет управление по сечению (мгновенному виду) апостериорной плотности в этот момент времени. На такой закон управления никакие ограничения вроде объема используемой памяти и скорости обработки измерений не накладываются. Поэтому будем называть подобные устройства управления *абсолютно оптимальными*, подчеркивая этим их способность обеспечивать достижение глобального экстремума соответствующего критерия качества управления. Эффективность же других устройств, учитывающих какие-либо ограничения, например конечномерность [7], будет заведомо хуже, но при их лучшей реализуемости.

Однако процедура *построения* интервально-оптимального регулятора Мортенсена весьма сложна из-за необходимости находить функционал от апостериорной плотности. В этом случае метод динамического программирования приводит к необходимости решать весьма сложное уравнение Беллмана–Мортенсена в вариационных производных Фреше. Только замена апостериорной плотности вероятности соответствующими ей достаточными координатами (статистиками) неизмеряемого состояния объекта в виде всех его условных моментов, квазимоментов или кумулянтов приводит к поиску функции Беллмана из уравнения в частных производных. Хотя

эта функция и имеет бесконечное количество аргументов, но, ограничиваясь достаточными координатами только нескольких младших порядков, можно найти некоторое приближение к абсолютно оптимальному регулятору.

Таким способом в частном случае *линейно-квадратично-гауссовской* (ЛКГ) задачи неограниченного управления удалось доказать теорему (принцип) разделения [3, 4, 6, 8], согласно которой оптимальное устройство управления Стратоновича–Мортенсена распадается на два легко получаемых и реализуемых блока. Первым из них является линейный стохастический фильтр Калмана–Бьюси, оптимальный в смысле минимума среднего квадрата ошибки оценивания в каждый момент времени. Синтез фильтра сводится к решению независимого от измерений *прямого* (в прямом времени) дифференциального уравнения Риккати для матрицы ковариаций этой ошибки, которое учитывает только параметры управляемой системы. Вторым представляет собой линейный позиционный регулятор, оптимальный в смысле детерминированной версии квадратичного критерия качества управления. Построение этого регулятора требует решения своего *обратного* (в обратном времени) уравнения Риккати, которое учитывает еще и весовые матрицы критерия. Но нарушение любого из четырех довольно жестких условий этой теоремы приводит к необходимости все же решать уравнение Беллмана.

В работе предлагается более простая процедура синтеза абсолютно оптимального безынерционного 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 критерием, но зависящим от времени. Это позволяет учитывать информацию только о прошлом и текущем состоянии объекта, оперативно меняя закон управления при чем-либо вызванных изменениях его будущего поведения. 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 условий задачи. Подобный эффект от оперативного критерия в нелинейном детерминированном случае продемонстрирован в [9]. Кроме того, процедура синтеза стохастического оперативно-оптимального конечномерного устройства управления, который учитывает требования к скорости обработки измерений, а потому не является абсолютно оптимальным, описана в [7].

1. Постановки задачи. Рассмотрим две задачи управления динамическими объектами в предположении марковости и диффузионности (отсутствия скачков) их векторов состояния, для описания которых будем использовать стохастические дифференциальные уравнения Ито.

Пусть $t \geq 0$ – время, X_t – n -мерный не измеряемый вектор состояния объекта управления, Y_t – m -мерный вектор его измеряемого выхода, U_t – l -мерный вектор кусочно-непрерывного управления из в общем случае ограниченной области Ω , V_t – k -мерный вектор непрерывных возмущений в виде центрированного и нормированного белого шума.

1.1. Управление по неполным точным измерениям. Эта задача возникает в случае, когда имеет место *парная* марковская модель объекта управления в виде системы из двух взаимозависимых уравнений для его общего вектора состояния (X_t, Y_t) :

$$\begin{aligned} \dot{X}_t &= a(t, X_t, Y_t, U_t) + B(t, X_t, Y_t, U_t)V_t, & \begin{bmatrix} X_0 \\ Y_0 \end{bmatrix} &\sim p_0(x, y). \\ \dot{Y}_t &= c(t, X_t, Y_t, U_t) + D(t, Y_t, U_t)V_t, \end{aligned} \quad (1.1)$$

Здесь и далее производные случайных функций по времени t будем понимать как отношение их стохастического дифференциала Ито вроде dX_t к дифференциалу времени, например $\dot{X}_t = dX_t/dt$. В этом смысле белый шум является производной $V_t = dW_t/dt$ соответствующего стандартного винеровского процесса W_t .

В (1.1) плотность распределения вероятности $p_0(x, y)$ начальных условий X_0, Y_0 известна не полностью, а лишь с точностью до условной плотности вероятности $\rho_0(x|y)$ начального значения неизмеряемого вектора X_0 , тогда как маргинальная (частная) плотность вероятности $q_0(y)$ измеряемого вектора Y_0 может быть произвольной:

$$p_0(x, y) = \rho_0(x|y)q_0(y) \quad \forall q_0(y). \quad (1.2)$$

Требуется найти не упреждающую зависимость управления U_t объектом (1.1), (1.2) от всех предыдущих измерений $Y_0^t = \{Y_\tau | \tau \in [0, t)\}$ и выполненных управлений $U_0^t = \{U_\tau | \tau \in [0, t)\}$ как таковой их функционал:

$$U_t = \vartheta(t, Y_0^t, U_0^t), \tag{1.3}$$

который обеспечивает минимум некоторому критерию качества управления состоянием объекта. В зависимости от объема располагаемой информации о процессе управления будем различать два вида таких критериев.

Обычно требуется минимизация *постоянного* во времени критерия оптимальности, задаваемого как среднее значение полученной на всем заданном отрезке времени $t \in [0, T]$ суммы случайных интегральных и терминальных потерь от управления:

$$J[\vartheta(\cdot)] = M \left[\int_0^T \mu(\tau, X_\tau, Y_\tau, U_\tau) d\tau + v(X_T, Y_T) \right] \rightarrow \min \quad \forall q_0(y). \tag{1.4}$$

Здесь M – оператор математического ожидания, конечный момент времени T будем считать, для простоты, фиксированным, а функции потерь являются неотрицательными $\mu(t, x, y, u) \geq 0$, $v(x, y) \geq 0$. Далее такой критерий будем называть *интервальным* (И-критерий).

Однако получаемое с его помощью управление (1.3) хоть и является не упреждающим, но в каждый текущий момент времени t требует информации о процессе функционирования объекта и в будущем, на интервале времени $(t, T]$. Это проявляется в использовании при решении задачи предварительного знания функций его уравнений (1.1) *на всем интервале* управления $[0, T]$, которое наблюдается при применении соответствующих стохастических версий как принципа максимума Понтрягина, так и метода динамического программирования Беллмана. Более того, функции потерь критерия (1.4) штрафуют и возможное будущее поведение объекта управления. Таким образом, использование критерия (1.4) требует полной информации как об уравнениях объекта, так и функциях потерь критерия оптимальности на всем плановом отрезке времени управления $[0, T]$.

В отличие от такой традиционной постановки задачи управления в последнее время усилился интерес к оптимизации *переменного* во времени критерия [7, 9]

$$I_t[\vartheta(\cdot)] = M \left[\int_0^t \varphi(\tau, X_\tau, Y_\tau, U_\tau) d\tau + \psi(t, X_t, Y_t) \right] \rightarrow \min \quad \forall q_0(y), \tag{1.5}$$

с функциями потерь $\varphi(t, x, y, u) \geq 0$, $\psi(t, x, y) \geq 0$, который будем называть *оперативным* (О-критерий). В отличие от И-критерия (1.4) он уже *не штрафует* неизвестное будущее поведение объекта и, как показало исследование его детерминированного аналога [9], получаемое с его помощью управление *не использует информации* об этом будущем. В результате внезапно возникшие в уравнениях объекта изменения *не требуют пересчета* всего закона оптимального управления им, а вид функций потерь $\varphi(\cdot)$, $\psi(\cdot)$ этого критерия можно, по мере необходимости, изменять со временем без какого-либо их влияния на предыдущее управление. Применение такого критерия в стохастической задаче синтеза *конечномерного* динамического устройства управления продемонстрировано в [7].

1.2. Управление по измерениям со случайными погрешностями. Такая более популярная задача возникает, если имеет место *скрытая* марковская модель объекта управления, когда измеряемый выход Y_t не влияет на неизмеряемое состояние X_t . В таком случае система уравнений (1.1) распадается на независящее от измерений Y_t уравнение состояния объекта с полностью известным начальным условием

$$\dot{X}_t = a(t, X_t, U_t) + B(t, X_t, U_t)V_t, \quad X_0 \sim p_0(x) \tag{1.6}$$

и зависимое только от состояния X_t и управления U_t уравнение измерителя этого состояния

$$\dot{Y}_t = c(t, X_t, U_t) + D(t, U_t)V_t, \quad Y_0 = 0, \tag{1.7}$$

причем его нулевое начальное условие общности измерений не ограничивает.

Тогда для синтеза закона управления (1.3) только состоянием объекта (1.6) И-критерий (1.4) не должен штрафовать сами измерения, поэтому он принимает вид

$$J[\vartheta(\cdot)] = M \left[\int_0^T \mu(\tau, X_\tau, U_\tau) d\tau + v(X_T) \right] \rightarrow \min. \quad (1.8)$$

Аналогично изменяется и О-критерий (1.5):

$$I_1[\vartheta(\cdot)] = M \left[\int_0^t \varphi(\tau, X_\tau, U_\tau) d\tau + \psi(t, X_t) \right] \rightarrow \min. \quad (1.9)$$

Формально уравнения (1.6), (1.7) и критерии (1.8), (1.9) отличаются от системы уравнений (1.1) и критериев (1.4), (1.5) лишь независимостью определяющих их функций от переменной выхода y . Поэтому и весь алгоритм решения этой задачи может быть получен из результатов решения предыдущей простым удалением из них этой переменной.

2. Апостериорная плотность и достаточные координаты. Синтез абсолютно оптимального закона управления основан на использовании наиболее полной информации о возможных значениях случайного вектора X_t , которая получена в результате всех произведенных к моменту времени t измерений Y_0^t и выполненных управлений U_0^t . Эта информация содержится в случайных значениях $\rho(t, x | Y_0^t, U_0^t)$ функционала условной плотности вероятности $\rho(t, x | y_0^t, u_0^t)$, определенной в результате регистрации измерений $Y_0^t = y_0^t$ и управлений $U_0^t = u_0^t$. Называемая поэтому *апостериорной плотностью вероятности* (АПВ), случайная функция

$$P(t, x) = \rho(t, x | Y_0^t, U_0^t)$$

удовлетворяет сложному стохастическому *интегродифференциальному уравнению* (ИДУ) Стратоновича—Кушнера. Последнее является уравнением состояния *абсолютно оптимального фильтра* (АОФ) Стратоновича, который представляет собой динамическое устройство с распределенными параметрами. Поэтому практическое использование АПВ как функции его состояния весьма затруднительно.

Выход состоит в замене АПВ вектором зависящих только от времени t статистик S_t неизмеряемого вектора X_t в виде числовых характеристик АПВ, называемых *достаточными координатами* (ДК) [1]. Ими могут быть бесконечные последовательности соответствующих условных моментов, квазимоментов или кумулянтов [10, 11]. Каждая из них, являясь функционалом измерений и управлений, определяется по АПВ с помощью определенной производящей вектор-функции $\xi(x)$ как ее апостериорное среднее:

$$S_t = \hat{\xi}(t, Y_0^t, U_0^t) = M[\xi(X_t) | Y_0^t, U_0^t] = \int \xi(x) P(t, x) dx. \quad (2.1)$$

Например, производящая функция начальных моментов содержит все степени элементов усредняемого вектора $\xi(x) = [x_i, x_i x_j, x_i x_j x_k, \dots]_{i,j,k,\dots=1,\dots,n}$. Далее символом “ $\hat{\ }^{\wedge}$ ” над именами и других функций будем отмечать результаты их *апостериорного усреднения*:

$$\hat{\eta}(t, Y_0^t, U_0^t) = M[\eta(t, x, Y_t, U_t) | Y_0^t, U_0^t] = \int \eta(t, x, Y_t, U_t) P(t, x) dx, \quad (2.2)$$

а подобные интегралы считать определенными, берущимися по всему евклидову пространству соответствующей размерности:

$$\int \alpha(x) dx \triangleq \int_{\mathbb{R}^n} \alpha(x) dx.$$

Восстановление АПВ по ДК выполняется по формуле

$$\rho(t, x | Y_0^t, U_0^t) = \pi(t, x, S_t), \quad (2.3)$$

где в качестве функции $\pi(\cdot)$ выступает функциональный ряд Эджворта или Грама—Шарлье.

Второе преимущество ДК, кроме их зависимости только от времени, состоит в возможности “урезания” бесконечных последовательностей квазимоментов или кумулянтов волевым обнулением (отбрасыванием) старших из них, начиная с некоторого порядка $L = 3, 4, \dots$. Это позволяет

получать приближенные решения задач, улучшая точность за счет увеличения порядка L . Недостатком такой процедуры является факториально быстрый рост размерности урезанного вектора $S_t^{(L)}$, определяемый по формуле [10]

$$\dim S_t^{(L)} = n + \frac{n(n+1)}{2} + \frac{n(n+1)(n+2)}{3!} + \dots + \frac{n(n+1)\dots(n+L-1)}{L!} = \frac{(n+L)!}{n!L!} - 1,$$

где n – размерность вектора X_t .

Приведем уравнения, которым удовлетворяют ДК в двух рассматриваемых случаях.

2.1. **Неполные измерения.** Если объект управления определяется соотношениями (1.1), (1.2), то АПВ является решением следующего стохастического ИДУ в частных производных [4, 10, 11]:

$$\frac{\partial P(t, x)}{\partial t} = K_x^{Y_t, U_t} [P] + \left[(c - \hat{c})^T P - \nabla_x^T (\Sigma P) \right] R^{-1} (\dot{Y}_t - \hat{c}) \quad (2.4)$$

с известным начальным условием

$$P(0, x) = \rho_0(x | Y_0) \quad (2.5)$$

и с естественными нулевыми граничными условиями на бесконечности для нее самой $\lim_{|x| \rightarrow \infty} P(t, x) = 0$ и для вектора потока ее вероятности $\Pi = aP - 0.5[\nabla_x^T(QP)]^T$. Здесь $K_x^{y,u}$ – прямой производящий дифференциальный оператор Фоккера–Планка–Колмогорова (ФПК) управляемого диффузионного процесса X_t , в обозначении которого верхние индексы подчеркивают зависимости его коэффициентов от соответствующих переменных, а нижний – переменную, по которой им осуществляется дифференцирование:

$$K_x^{y,u} [P(t, x)] = -\nabla_x^T [a(t, x, y, u)P(t, x)] + 0.5 \text{tr} [\nabla_x \nabla_x^T [Q(t, x, y, u)P(t, x)]],$$

тогда как $Q = BB^T$, $R = DD^T$, $\Sigma = BD^T$ – матрицы условных интенсивностей собственных и взаимных возмущений элементов X_t , Y_t вектора состояния объекта (1.1). При этом в уравнении (2.4) функции $a(\cdot)$, $c(\cdot)$, $Q(\cdot)$, $\Sigma(\cdot)$ зависят от аргументов (t, x, Y_t, U_t) , а не зависят от переменной x только функции $\hat{c} = \hat{c}(t, Y_t^1, U_t^1)$, $R = R(t, Y_t, U_t)$.

Для нахождения соответствующего уравнения для вектора ДК, определяющего состояние АОФ Стратоновича, продифференцируем по времени его связь с АПВ (2.1) и подставим в нее (2.4). Получаем

$$\dot{S}_t = \int \xi(x) \left\{ K_x^{Y_t, U_t} [P] + \left[(c - \hat{c})^T P - \nabla_x^T (\Sigma P) \right] R^{-1} (\dot{Y}_t - \hat{c}) \right\} dx$$

или, учитывая независимость множителя $R^{-1}(\dot{Y}_t - \hat{c})$ от переменной интегрирования,

$$\dot{S}_t = \int \xi K_x^{Y_t, U_t} [P] dx + \left[\int [\xi c^T - \xi \hat{c}^T] P dx - \int \xi \nabla_x^T (\Sigma P) dx \right] R^{-1} (\dot{Y}_t - \hat{c}).$$

Избавляясь здесь от частных производных АПВ по x вычислением первого и третьего интеграла по частям, с учетом упомянутых выше нулевых граничных условий для АПВ находим

$$\dot{S}_t = \int K_x^{*Y_t, U_t} [\xi] P dx + [(\widehat{\xi c^T} - \widehat{\xi \hat{c}^T}) + \int (\xi_x \Sigma) P dx] R^{-1} (\dot{Y}_t - \hat{c}), \quad (2.6)$$

где $K_x^{*y,u}$ – сопряженный к оператору $K_x^{y,u}$ обратный производящий оператор процесса X_t :

$$K_x^{*y,u} = a^T \nabla_x + 0.5 \text{tr} [Q \nabla_x \nabla_x^T],$$

причем результат его действия на вектор-функцию $\xi(x) = [\xi^{(i)}(x)]_{i=1,2,\dots}$ следует понимать поэлементно как $K_x^{*y,u} [\xi] = [K_x^{*y,u} [\xi^{(i)}]]_{i=1,2,\dots}$, а $\xi_x = (\nabla_x \xi^T)^T$ – матрица Якоби первых частных производных вектор-функции $\xi(x)$. Поэлементная запись выражения (2.6) имеет вид

$$\dot{S}_t^{(i)} = \int K_x^{*y,u} [\xi^{(i)}] P dx + \left[(\widehat{\xi^{(i)} c^T} - \widehat{\xi^{(i)} \hat{c}^T}) + \int [\xi_x^{(i)T} \Sigma] P dx \right] R^{-1} (\dot{Y}_t - \hat{c}), \quad i = 1, 2, \dots,$$

и при отсутствии управления она совпадает с формулой для стохастического дифференциала апостериорного среднего $\hat{\xi}^{(i)}(\cdot)$, непосредственно полученной в [10, с. 401].

Наконец, заменяя АПВ функцией $\pi(\cdot)$ ее восстановления (2.3) по ДК, получаем, что апостериорное среднее (2.2) из функционала прошлых измерений и управлений превращается в обычную функцию

$$\hat{\eta}(t, y, u, s) = \int \eta(t, x, y, u) \pi(t, x, s) dx, \quad (2.7)$$

а соотношение (2.6) принимает вид искомого уравнения состояния АОФ [4]:

$$\dot{S}_t = h(t, Y_t, U_t, S_t) + \Theta(t, Y_t, U_t, S_t) [\dot{Y}_t - \hat{c}(t, Y_t, U_t, S_t)]. \quad (2.8)$$

Здесь функция смещения $h(\cdot)$ и коэффициент усиления измерения $\Theta(\cdot)$ вычисляются по использующим исходные данные (1.1) формулам

$$h(t, y, u, s) = \int K_x^{*y, u} [\xi(x)] \pi(t, x, s) dx, \\ \Theta(t, y, u, s) = \left[(\widehat{\xi c^T} - \hat{\xi} \hat{c}^T) + \int (\xi_x \Sigma) \pi(t, x, s) dx \right] R^{-1}(t, y, u).$$

Начальным же условием для уравнения (2.8) является, согласно (2.1), (2.5), случайная величина

$$S_0 = \hat{\xi}(Y_0) = \int \xi(x) p_0(x | Y_0) dx. \quad (2.9)$$

В результате уравнение состояния АОФ из интегродифференциального (2.4) преобразовано в дифференциальное (2.8), что существенно облегчает решение задачи управления.

2.2. Стохастические измерения. Пусть теперь объект управления определяется уравнениями (1.6), (1.7). Тогда из-за независимости их функций $a(\cdot)$, $B(\cdot)$, $c(\cdot)$, $D(\cdot)$, $p_0(\cdot)$ от вектора измерений Y_t уравнение состояния АОФ (2.8) и его начальное условие (2.9) упрощаются соответственно:

$$\dot{S}_t = h(t, U_t, S_t) + \Theta(t, U_t, S_t) [\dot{Y}_t - \hat{c}(t, U_t, S_t)], \quad S_0 = \hat{\xi}_0 = \int \xi(x) p_0(x) dx, \quad (2.10)$$

причем структурные функции этого уравнения определяются по формулам

$$h(t, u, s) = \int K_x^{*u} [\xi(x)] \pi(t, x, s) dx, \\ \Theta(t, u, s) = \left[(\widehat{\xi c^T} - \hat{\xi} \hat{c}^T) + \int (\xi_x \Sigma) \pi(t, x, s) dx \right] R^{-1}(t, u), \\ \hat{\eta}(t, u, s) = \int \eta(t, x, u) \pi(t, x, s) dx.$$

3. Интервально-оптимальный регулятор. Представим основанные на применении ДК известные решения обеих задач оптимального управления с И-критериями [4].

3.1. Неполные измерения. Сначала рассмотрим задачу (1.1)–(1.4). Согласно (2.1), вектор S_t информативно эквивалентен всем прошлым измерениям Y_0^t и управлениям U_0^t , поэтому текущее управление U_t можно искать не как их инерционный функционал (1.3), а как позиционную версию регулятора Мортенсена в виде безынерционной зависимости управления от последнего измерения и текущего вектора ДК:

$$U_t = u(t, Y_t, S_t). \quad (3.1)$$

Тогда, исключая переменную U_t из уравнений (1.1), (2.8), получим замкнутую систему из трех уравнений Ито:

$$\dot{X}_t = a^u(t, X_t, Y_t, S_t) + B^u(t, X_t, Y_t, S_t) V_t, \\ \dot{Y}_t = c^u(t, X_t, Y_t, S_t) + D^u(t, Y_t, S_t) V_t, \\ \dot{S}_t = e^u(t, X_t, Y_t, S_t) + F^u(t, Y_t, S_t) V_t. \quad (3.2)$$

Здесь функции сноса и диффузии третьего уравнения имеют вид

$$e^u(t, x, y, s) = h^u(t, y, s) + \Theta^u(t, y, s)[c^u(t, x, y, s) - \hat{c}^u(t, y, s)],$$

$$F^u(t, y, s) = \Theta^u(t, y, s)D^u(t, y, s),$$

а верхним индексом “ u ” отмечены сложные функции своих аргументов, содержащие функцию выхода $u(\cdot)$ регулятора (3.1), например

$$a^u(t, x, y, s) = a(t, x, y, u(t, y, s)). \tag{3.3}$$

Из (3.2) следует, что случайный процесс (X_t, Y_t, S_t) является марковским со следующей вектор-функцией сноса $\omega(\cdot)$ и матричной функцией диффузии $\Upsilon(\cdot)$:

$$\omega^u(t, x, y, s) = \begin{bmatrix} a^u(t, x, y, s) \\ c^u(t, x, y, s) \\ e^u(t, x, y, s) \end{bmatrix}, \quad \Upsilon^u(t, x, y, s) = \begin{bmatrix} B^u(t, x, y, s) \\ D^u(t, y, s) \\ F^u(t, y, s) \end{bmatrix}.$$

Его совместная плотность вероятности $r(t, x, y, s)$ удовлетворяет дифференциальному уравнению ФПК:

$$\frac{\partial r(t, x, y, s)}{\partial t} = K_{xys}^u[r] \triangleq -\nabla^T(\omega^u r) + 0.5\text{tr}[\nabla\nabla^T(\Upsilon^u r)], \quad t \geq 0, \tag{3.4}$$

где $\nabla = (\nabla_x^T, \nabla_y^T, \nabla_s^T)^T$ – оператор общего градиента, или соответствующему интегродифференциальному тождеству ФПК [12, 13]:

$$\frac{d}{dt} \iiint \eta r dx dy ds = \iiint \left(\frac{\partial \eta}{\partial t} + K_{xys}^{*u}[\eta] \right) r dx dy ds \quad \forall \eta(t, x, y, s) \in \mathbb{C}^{1,2,2,2}, \tag{3.5}$$

в котором сопряженный к K_{xys}^u обратный производящий оператор K_{xys}^{*u} этого же процесса определим более подробным выражением:

$$K_{xys}^{*u}[\eta] = a^{uT}\eta_x + c^{uT}\eta_y + e^{uT}\eta_s + 0.5\text{tr}[Q^u \eta_{xx}] + 0.5\text{tr}[2\Sigma^{uT}\eta_{xy} + R^u\eta_{yy}] + \\ + 0.5\text{tr}[2\Theta^u\Sigma^{uT}\eta_{xs} + 2\Theta^u R^u\eta_{ys} + \Theta^u R^u\Theta^{uT}\eta_{ss}]. \tag{3.6}$$

Начальное значение совместной плотности $r(\cdot)$, согласно (1.2), (2.9), задано частично:

$$r(0, x, y, s) = \rho_0(x|y)q_0(y)\delta[s - \hat{\xi}(y)], \tag{3.7}$$

где $\delta(\cdot)$ – функция Дирака, а плотность $q_0(y)$ является произвольной.

Подобным соотношениям, но с другим начальным условием удовлетворяет и характеризующая переход этого процесса за интервал времени $[t, \theta]$ из состояния $(X_t = x, Y_t = y, S_t = s)$ в состояние $(X_\theta = \tilde{x}, Y_\theta = \tilde{y}, S_\theta = \tilde{s})$ плотность вероятности перехода $p(\theta, \tilde{x}, \tilde{y}, \tilde{s} | t, x, y, s)$. Она позволяет определить функцию будущих *априорных* средних потерь $w(t, x, y, s)$ как ту часть минимизируемого И-критерия общих потерь (1.4), которые будут получены за оставшееся время управления $\theta \in [t, T]$ при старте из любого текущего состояния $(X_t = x, Y_t = y, S_t = s)$:

$$w(t, x, y, s) = \int_t^T d\theta \iiint \mu^u(\theta, \tilde{x}, \tilde{y}, \tilde{s}) p(\theta, \tilde{x}, \tilde{y}, \tilde{s} | t, x, y, s) d\tilde{x} d\tilde{y} d\tilde{s} + \iiint v(\tilde{x}, \tilde{y}) p(T, \tilde{x}, \tilde{y}, \tilde{s} | t, x, y, s) d\tilde{x} d\tilde{y} d\tilde{s}.$$

Однако эта функция зависит от значения x неизмеряемого вектора X_t . Не зависящим от него является ее *апостериорное* среднее:

$$W(t, y, s) = \hat{w}(t, y, s) = \int w(t, x, y, s) \pi(t, x, s) dx,$$

которое удовлетворяет следующему уравнению Беллмана [4]:

$$-\frac{\partial W(t, y, s)}{\partial t} = \min_{u \in \Omega} \left\{ \hat{\mu} + \hat{c}^T W_y + h^T W_s + 0.5\text{tr} \left(\begin{bmatrix} W_{yy} & W_{ys} \\ W_{sy} & W_{ss} \end{bmatrix} \begin{bmatrix} R & R\Theta^T \\ \Theta R & \Theta R\Theta^T \end{bmatrix} \right) \right\}, \tag{3.8}$$

с конечным условием в виде *апостериорного* среднего терминального члена критерия

$$W(T, y, s) = \hat{v}(T, y, s) = \int v(x, y) \pi(T, x, s) dx \quad (3.9)$$

и следующими коэффициентами:

$$\begin{aligned} \hat{c} &= \hat{c}(t, y, u, s) = \int c(t, x, y, u) \pi(t, x, s) dx, & R &= D(t, y, u) D^T(t, y, u), \\ \hat{\mu} &= \hat{\mu}(t, y, u, s) = \int \mu(t, x, y, u) \pi(t, x, s) dx, & h &= h(t, y, u, s), & \Theta &= \Theta(t, y, u, s). \end{aligned}$$

Однако, как обычно, процедура применения уравнения Беллмана (3.8) требует выполнения двух весьма сложных операций. Во-первых, необходимо найти частный минимум по переменной u функции $f(\cdot)$ его правой части:

$$v(t, y, s, W) = \arg \min_{u \in \Omega} f(t, y, s, u, W) \quad \forall t, y, s, W.$$

Во-вторых, следует решить в обратном времени, учитывая конечное условие (3.9), полученное в результате уравнение в частных производных, найдя тем самым функцию Беллмана $W(t, y, s)$. Наконец, подставляя ее в функцию частного минимума, получим искомую функцию безынерционной версии регулятора Мортенсена:

$$u(t, y, s) = v[t, y, s; W(t, y, s)].$$

Подчеркнем, что уравнение АОФ Стратоновича (2.8) при синтезе этого регулятора решать не требуется. Оно используется только на следующем этапе реализации оптимального устройства управления, поставляя информацию регулятору Мортенсена (3.1) для получения управления U_t и учитывая это управление в своих структурных функциях $h(\cdot)$, $\Theta(\cdot)$, $\hat{c}(\cdot)$ для соответствующей коррекции своего состояния S_t .

3.2. С т о х а с т и ч е с к и е и з м е р е н и я. Если же имеет место более простая задача управления (1.6)–(1.8), то из-за отсутствия в ней влияния измерения Y_t на состояние X_t объекта и на значение И-критерия J можно, в отличие от (3.1), искать управление без непосредственного учета измерения как функцию только ДК:

$$U_t = u(t, S_t). \quad (3.10)$$

Тогда, учитывая соответствующее уравнение АОФ (2.10), вместо трех уравнений Ито (3.2) удастся ограничиться рассмотрением только двух:

$$\begin{aligned} \dot{X}_t &= a^u(t, X_t, S_t) + B^u(t, X_t, S_t) V_t, & X_0 &\sim p_0(x), \\ \dot{S}_t &= e^u(t, X_t, S_t) + F^u(t, S_t) V_t, & S_0 &= \hat{\xi}_0, \end{aligned} \quad (3.11)$$

с такими функциями второго из них:

$$\begin{aligned} e^u(t, x, s) &= h^u(t, s) + \Theta^u(t, s) [c^u(t, x, s) - \hat{c}^u(t, s)], \\ F^u(t, s) &= \Theta^u(t, s) D^u(t, s). \end{aligned}$$

Поэтому процесс (X_t, S_t) марковский, а его плотность вероятности $r(t, x, s)$ удовлетворяет аналогичному (3.5) тождеству ФПК:

$$\frac{d}{dt} \iint \eta r dx ds = \iint \left(\frac{\partial \eta}{\partial t} + K_{xs}^{*u}[\eta] \right) r dx ds \quad \forall \eta(t, x, s) \in \mathbb{C}^{1,2,2} \quad (3.12)$$

с полностью заданным начальным условием $r(0, x, s) = \rho_0(x) \delta(s - \bar{\xi}_0)$ и с оператором

$$K_{xs}^{*u}[\eta] = a^{uT} \eta_x + e^{uT} \eta_s + 0.5 \text{tr}[Q^u \eta_{xx}] + 0.5 \text{tr}[2\Theta^u \Sigma^{uT} \eta_{xs} + \Theta^u R^u \Theta^{uT} \eta_{ss}].$$

В результате уравнение Беллмана (3.8) и его конечное условие (3.9) тоже принимают более простой вид

$$-\frac{\partial W(t, s)}{\partial t} = \min_{u \in \Omega} \left\{ \hat{\mu} + h^T W_s + 0.5 \text{tr}[W_{ss} \Theta R \Theta^T] \right\}, \quad W(T, s) = \hat{v}(T, s). \quad (3.13)$$

Однако процедура его решения остается по-прежнему весьма сложной. Аналитически оно решено только в весьма частном случае ЛКГ-задачи неограниченного управления, которая будет рассмотрена в разд. 5.

4. Оперативно-оптимальный регулятор. Приступим теперь к решению двух новых задач синтеза аналогов регулятора Мортенсена, оптимальных в смысле изменяющихся во времени О-критериев качества (1.5) или (1.9).

4.1. *Неполные измерения.* Снова рассмотрим *парную* марковскую модель объекта управления (1.1) с начальными условиями (1.2):

$$\begin{aligned} \dot{X}_t &= a(t, X_t, Y_t, U_t) + B(t, X_t, Y_t, U_t)V_t, & \begin{bmatrix} X_0 \\ Y_0 \end{bmatrix} &\sim \rho_0(x|y)q_0(y), \\ \dot{Y}_t &= c(t, X_t, Y_t, U_t) + D(t, Y_t, U_t)V_t, \end{aligned}$$

для которой уравнение АОФ имеет вид (2.8) с начальным условием (2.9). Для этого объекта также будем искать уравнение безынерционного регулятора (3.1):

$$U_t = u(t, Y_t, S_t), \tag{4.1}$$

но от последнего потребуем оптимальности в смысле О-критерия (1.5), который в результате подстановки в него выражения (4.1) принимает вид

$$I_t = M \left[\int_0^t \varphi''(\tau, X_\tau, Y_\tau, S_\tau) d\tau + \psi(t, X_t, Y_t) \right]. \tag{4.2}$$

В таком случае вместо метода динамического программирования приходится использовать метод припасовывания, состоящий в последовательной оптимизации сначала стартового значения критерия

$$I_0 = M[\psi(0, X_0, Y_0)] = \int \psi(0, x, y) \rho_0(x|y) q_0(y) dx dy,$$

которое в данном случае от выбора управления не зависит, а затем и приращения этого критерия за каждый сколь угодно малый промежуток времени $[t, t + \Delta t]$, $\Delta t \downarrow 0$. Последнее, как показано в [7, 13], сводится к минимизации скорости изменения этого функционала в любой момент времени:

$$\dot{I}_t \rightarrow \min_{u(\cdot)}, \quad \forall t \geq 0, \tag{4.3}$$

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.

Для нахождения зависимости \dot{I}_t от искомой функции регулятора $u(\cdot)$ представим критерий (4.2) через плотность вероятности $r(t, x, y, s)$ случайных аргументов усредняемых в нем функций потерь:

$$I_t = \int_0^t \langle \varphi''(\tau, x, y, s), r(\tau, x, y, s) \rangle d\tau + \langle \psi(t, x, y), r(t, x, y, s) \rangle. \tag{4.4}$$

Здесь для краткости символом $\langle \eta, r \rangle$ обозначен интеграл от произведения функций $\eta(\cdot)$ и $r(\cdot)$, который в данном случае имеет смысл математического ожидания (среднего значения) первой из них:

$$\langle \eta, r \rangle = \iiint \eta(t, x, y, s) r(t, x, y, s) dx dy ds = M[\eta(t, X_t, Y_t, S_t)]. \tag{4.5}$$

Дифференцируя равенство (4.4) по времени, имеем $\dot{I}_t = \langle \varphi'' , r \rangle + d \langle \psi, r \rangle / dt$ или, используя во втором слагаемом этой суммы тождество ФПК (3.5), находим вид производной критерия:

$$\dot{I}_t = \left\langle \varphi'' + \frac{\partial \psi}{\partial t} + K_{xy^s}^{*u}[\psi], r \right\rangle.$$

Искомая функция входит в это выражение только в виде ее сечения $u(t, \cdot)$ при фиксированном t . Так как слагаемое $\langle \partial \psi / \partial t, r \rangle$ от него не зависит, то, согласно (4.3), достаточно минимизировать функционал

$$\Xi_t[u(t, \cdot)] = \langle \varphi^u + K_{xys}^{*u}[\psi], r \rangle \rightarrow \min_{u(t, \cdot)}, \quad \forall t \geq 0.$$

Учитывая в нем вид (3.6) оператора K_{xys}^{*u} и независимость дифференцируемой им функции текущих потерь $\psi(t, x, y)$ от переменной s , получаем явное выражение

$$\Xi_t[u(t, \cdot)] = \langle \varphi^u + a^{uT} \psi_x + c^{uT} \psi_y + 0.5 \text{tr} [Q^u \psi_{xx}] + 0.5 \text{tr} [2\Sigma^{uT} \psi_{xy} + R^u \psi_{yy}], r \rangle. \quad (4.6)$$

Так как искомая функция $u(t, y, s)$ от одной из трех переменных интегрирования не зависит, то этот функционал можно упростить, заменив совместную плотность вероятности $r(\cdot)$ случайного процесса (X_t, Y_t, S_t) произведением маргинальной (частной) плотности $q(\cdot)$ случайных величин Y_t, S_t на соответствующую условную плотность $\rho(\cdot)$:

$$r(t, x, y, s) = q(t, y, s) \rho(t, x | y, s), \quad q(t, y, s) = \int r(t, x, y, s) dx. \quad (4.7)$$

Действительно, представляя тройной интеграл совместного среднего (4.5) в виде повторного, состоящего из внутреннего интеграла усреднения по условной плотности и внешнего интеграла усреднения по маргинальной плотности

$$\langle \eta, r \rangle = \iint \left[\int \eta(t, x, y, s) \rho(t, x | y, s) dx \right] q(t, y, s) dy ds,$$

получаем, что функционал (4.6) принимает вид маргинального среднего

$$\Xi_t[u(t, \cdot)] = \iint \zeta[t, y, z; u(t, y, s)] q(t, y, s) dy ds$$

от следующей подинтегральной функции условного среднего:

$$\zeta(t, y, s; u) = \int (\varphi + a^T \psi_x + c^T \psi_y + 0.5 \text{tr} [Q \psi_{xx}] + 0.5 \text{tr} [2\Sigma^T \psi_{xy} + R \psi_{yy}]) \rho dx. \quad (4.8)$$

Тогда из свойства монотонности операции интегрирования и неотрицательности плотности вероятности $q(\cdot) \geq 0$ следует [7, 13], что для отыскания минимума функционала (4.6) достаточно найти минимум его маргинально усредняемой функции условного среднего $\zeta(\cdot)$ по одному из ее аргументов при любых значениях других:

$$u(t, y, s) = \arg \min_{v \in \Omega \subset \mathbb{R}^l} \zeta(t, y, s; v), \quad \forall t, y, s. \quad (4.9)$$

Однако минимизируемая здесь функция (4.8) требует знания условной плотности вероятности $\rho(\cdot)$. Хотя она формулами (4.7) и выражается через совместную плотность $r(\cdot)$, которая определяется как решение задачи Коши (3.4), (3.7) для уравнения ФПК или соответствующего ему тождества (3.5), но частичная неопределенность начальных условий (1.2) объекта управления (1.1) делает нахождение плотности $r(\cdot)$ в общем случае *невозможным*.

Поэтому при неопределенности плотности вероятности $q_0(y)$ начального измерения Y_t следует решать, вместо уравнения ФПК, полностью определенную задачу Коши для ИДУ относительно условной плотности $\rho(\cdot)$ [13]. Оно получено декомпозицией уравнения ФПК (3.4) на независимое уравнение для плотности $\rho(\cdot)$ и зависящее от нее уравнение для маргинальной плотности $q(\cdot)$. В рассматриваемом здесь случае это уравнение и его начальное условие имеют вид

$$\frac{\partial \rho(t, x | y, s)}{\partial t} = -\nabla_x^T (a^u \rho) + 0.5 \text{tr} [\nabla_x \nabla_x^T (Q^u \rho)] - L_{ys}^{*u}[\rho], \quad \rho(0, x | y, s) = \rho_0(x | y), \quad (4.10)$$

причем в коэффициенты его оператора

$$L_{ys}^{*u}[\rho] = \rho_y^T \bar{c}^u + \rho_s^T \bar{e}^u + 0.5 \text{tr} [R^u \rho_{yy}] + 0.5 \text{tr} [2\Theta R^u \rho_{ys} + \Theta R^u \Theta^T \rho_{ss}],$$

входят интегральные функции условного среднего \bar{c}^u, \bar{e}^u , обозначенные, в отличие от (2.7), чертой сверху:

$$\bar{\eta}(t, y, u, s) = \int \eta(t, x, y, u) \rho(t, x | y, s) dx.$$

Таким образом, синтез оперативного-оптимального регулятора (4.1) сведен к задаче параметрического нелинейного программирования (4.9), но минимизируемая в ней функция $\zeta(\cdot)$ с течением

времени изменяется не только из-за нестационарности исходных функций объекта управления (1.1) и оперативного критерия (1.5), но и благодаря изменению условной плотности $\rho(\cdot)$. Динамика последней определяется задачей Коши (4.10), которая функцию $u(\cdot)$ использует. Оперативность этой процедуры, в отличие от решения уравнения Беллмана (3.8), состоит в неиспользовании информации о будущем поведении объекта.

4.2. Стохастические измерения. Теперь рассмотрим *скрытую* марковскую модель объекта управления (1.6), (1.7):

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

Снова будем искать закон управления (3.10):

$$U_t = u(t, S_t),$$

но оптимальный в смысле О-критерия (1.9)

$$I_t[\vartheta(\cdot)] = \mathbb{M} \left[\int_0^t \varphi(\tau, X_\tau, U_\tau) d\tau + \psi(t, X_t) \right] \rightarrow \min.$$

В этом случае уравнение состояния АОФ имеет вид (2.10), а тождество ФПК для совместной плотности вероятности $r(t, x, s)$ – (3.12).

Тогда, повторяя рассуждения из разд. 4.1, из (4.8), (4.9) получаем, что построение О-регулятора также сводится к параметрической минимизации:

$$u(t, s) = \arg \min_{v \in \Omega \subset \mathbb{R}^l} \zeta(t, s; v), \quad \forall t, s, \tag{4.11}$$

но более простой функции условного среднего:

$$\zeta(t, s; u) = \int (\varphi + a^T \psi_x + 0.5 \operatorname{tr}[Q \psi_{xx}]) \rho dx. \tag{4.12}$$

Здесь условная плотность $\rho(\cdot)$ определяется либо по получаемой решением задачи Коши для тождества ФПК (3.12) совместной плотности $r(\cdot)$ как отношение

$$\rho(t, x | s) = r(t, x, s) / \int r(t, x, s) dx,$$

либо находится из своего уравнения, которое является частным случаем (4.10) и имеет вид [13]

$$\frac{\partial \rho(t, x | s)}{\partial t} = -\nabla_x^T (a^u \rho) + 0.5 \operatorname{tr}[\nabla_x \nabla_x^T (Q^u \rho)] - L_s^{*u}[\rho], \quad \rho(0, x | s) = p_0(x), \tag{4.13}$$

$$L_s^{*u}[\rho] = \rho_s^T e^{-u} + 0.5 \operatorname{tr}[\Theta R^u \Theta^T \rho_{ss}].$$

5. Пример ЛКГ-задачи. Ограничимся более простым вариантом случайных измерений. Пусть уравнения объекта (1.6) и измерителя (1.7) линейные:

$$\dot{X}_t = A(t)X_t + K(t)U_t + B(t)V_t, \quad X_0 \sim N(x \| m_0^x, D_0^x), \tag{5.1}$$

$$\dot{Y}_t = C(t)X_t + M(t)U_t + D(t)dV_t, \quad Y_0 = 0, \tag{5.2}$$

начальное состояние объекта X_0 гауссовское с плотностью вероятности нормального закона распределения $N(\cdot)$ при математическом ожидании m_0^x и ковариации D_0^x , управление не ограничено $U_t \in \Omega = \mathbb{R}^l$, а критерии его оптимальности (1.8), (1.9) являются квадратическими:

$$J = \frac{1}{2} \mathbb{M} \left\{ \int_0^T [X_\tau^T X(\tau) X_\tau + U_\tau^T \Phi(\tau) U_\tau] d\tau + X_T^T \Pi X_T \right\} \rightarrow \min, \tag{5.3}$$

$$I_t = \frac{1}{2} \mathbb{M} \left\{ \int_0^t [X_\tau^T X(\tau) X_\tau + U_\tau^T \Phi(\tau) U_\tau] d\tau + X_t^T \Upsilon(t) X_t \right\} \rightarrow \min \tag{5.4}$$

с весовыми матрицами $\Phi(t), \Upsilon(t) > 0, X(t), \Pi \geq 0$.

Исходные соотношения этой задачи отличаются от общих выражений линейностью функций сноса объекта и измерителя $a(t, x, u) = A(t)x + K(t)u$, $c(t, x, u) = C(t)x + M(t)u$, зависимостью интенсивностей шумов только от времени $Q(t, x, u) = Q(t)$, $R(t, u) = R(t)$, $S(t, x, u) = S(t)$, гауссовостью начальной плотности вероятности $p_0(x) = N(x \| m_0^x, D_0^x)$, а также квадратичностью интегрантов и терминантов обоих критериев:

$$\mu(t, x, u) = \varphi(t, x, u) = 0.5[x^T X(t)x + u^T \Phi(t)u], \quad v(x) = 0.5x^T \Pi x, \quad \psi(t, x) = 0.5x^T \Upsilon(t)x.$$

Далее очевидные зависимости параметров системы и критериев от времени t будем опускать.

5.1. **Л и н е й н ы й ф и л ь т р К а л м а н а – Б ь ю с и.** Из-за линейности уравнений (5.1), (5.2) и гауссовости как начального состояния X_0 , так и белого возмущения V_t , АПВ (2.3) является гауссовской $\pi(t, x, S_t) = N(x \| \hat{X}_t, P_t)$, так что вектор кумулянтных или квазиомментных ДК состоит только из двух компонент $S_t = (\hat{X}_t, P_t)$. При этом случайная оценка \hat{X}_t состояния X_t определяется управляемым стохастическим уравнением Ито:

$$\dot{\hat{X}}_t = A\hat{X}_t + KU_t + (P_t C^T + \Sigma)R^{-1}[\dot{Y}_t - (C\hat{X}_t + MU_t)], \quad \hat{X}_0 = m_0^x, \quad (5.5)$$

а детерминированная матрица ковариаций ошибки оценивания $P_t = \text{cov}(X_t - \hat{X}_t)$ находится из не зависящего ни от оценки \hat{X}_t , ни от управления U_t прямого уравнения Риккати:

$$\dot{P}_t = AP_t + P_t A^T + Q - [P_t C^T + \Sigma]R^{-1}[CP_t + \Sigma^T], \quad P_t = D_0^x. \quad (5.6)$$

Последнее обстоятельство делает матрицу P_t заранее известной функцией времени. Это позволяет исключить ее из состава ДК, считая теперь $S_t = \hat{X}_t$, и искать управление вместо (3.10) как функцию только от оценки

$$U_t = u(t, \hat{X}_t). \quad (5.7)$$

При этом уравнением состояния фильтра (2.10) становится только уравнение для оценки (5.5), так что далее

$$h(t, u, \hat{x}) = A\hat{x} + Ku, \quad \Theta(t, u, \hat{x}) = \Theta(t) = (P_t C^T + \Sigma)R^{-1}.$$

5.2. **И н т е р в а л ь н а я т е о р е м а р а з д е л е н и я.** Приведем известный результат [4, 8]. Оптимизируем функцию (5.7) по квадратическому И-критерию (5.3), для чего воспользуемся уравнением Беллмана (3.13). Оно теперь принимает вид

$$-\frac{\partial W(t, \hat{x})}{\partial t} = \min_{u \in \mathbb{R}^l} \{ \hat{\mu} + h^T W_{\hat{x}} + 0.5 \text{tr}[W_{\hat{x}\hat{x}} \Theta R \Theta^T] \}, \quad W(T, \hat{x}) = \hat{v}(T, \hat{x}), \quad (5.8)$$

в котором

$$\hat{\mu}(t, u, \hat{x}) = 0.5 \int (x^T X x + u^T \Phi u) N(x \| \hat{x}, P_t) dx = 0.5 (\text{tr}[X(P_t + \hat{x}\hat{x}^T)] + u^T \Phi u),$$

$$\hat{v}(T, \hat{x}) = 0.5 \int x^T \Pi x N(x \| \hat{x}, P_T) dx = 0.5 \text{tr}[\Pi(P_T + \hat{x}\hat{x}^T)].$$

В уравнении (5.8) минимизируемая по переменной $u \in \mathbb{R}^l$ функция является квадратической:

$$f(t, \hat{x}, u, W) = 0.5 u^T \Phi u + u^T K^T W_{\hat{x}} + \text{invar}(u)$$

и (так как $\Phi > 0$) достигает минимума в точке

$$u = -\Phi^{-1} K^T W_{\hat{x}}. \quad (5.9)$$

Подставляя это в (5.8), имеем обратную задачу Коши для уравнения в частных производных:

$$-\frac{\partial W(t, \hat{x})}{\partial t} = 0.5 (\text{tr}[X(P_t + \hat{x}\hat{x}^T)] + W_{\hat{x}}^T K \Phi^{-1} K^T W_{\hat{x}}) + (A\hat{x} - K \Phi^{-1} K^T W_{\hat{x}})^T W_{\hat{x}} + 0.5 \text{tr}[W_{\hat{x}\hat{x}} \Theta R \Theta^T], \quad (5.10)$$

$$W(T, \hat{x}) = 0.5 (\text{tr}[\Pi P_T] + \hat{x}^T \Pi \hat{x}).$$

Покажем, что его решением является смещенная по времени квадратическая форма

$$W(t, \hat{x}) = 0.5[\sigma(t) + \hat{x}^T L(t) \hat{x}],$$

а формула (5.9) принимает вид

$$u = -\Phi^{-1} K^T L \hat{x}. \tag{5.11}$$

Действительно, подставляя производные $\partial W / \partial t = 0.5[\dot{\sigma} + \hat{x}^T \dot{L} \hat{x}]$, $W_{\dot{x}} = L \hat{x}$, $W_{\hat{x}\hat{x}} = L$ в равенство (5.10) и приводя в нем подобные члены, получаем два алгебраических тождества:

$$\begin{aligned} -\dot{\sigma} - \hat{x}^T \dot{L} \hat{x} &= \text{tr}[X P_t] + \hat{x}^T (X - L^T K \Phi^{-1} K^T L) \hat{x} + 2 \hat{x}^T (A^T L) \hat{x} + \text{tr}[L \Theta R \Theta^T], \\ \sigma(T) + \hat{x}^T L(T) \hat{x} &= \text{tr}[\Pi P_T] + \hat{x}^T \Pi \hat{x}, \end{aligned}$$

которые справедливы при любых значениях переменной \hat{x} . Приравнявая в них свободные члены, находим уравнение для величины смещения $\sigma(t)$:

$$-\dot{\sigma} = \text{tr}[X P_t] + \text{tr}[L \Theta R \Theta^T], \quad \sigma(T) = \text{tr}[\Pi P_T],$$

а из равенства квадратичных форм, представив одну из них в симметрическом виде $2 \hat{x}^T (A^T L) \hat{x} = \hat{x}^T (A^T L) \hat{x} + \hat{x}^T (A^T L)^T \hat{x}$, имеем известное обратное уравнение Риккати для матрицы коэффициентов $L(t)$:

$$-\dot{L} = A^T L + L A + X - L^T K \Phi^{-1} K^T L, \quad L(T) = \Pi. \tag{5.12}$$

В результате из (5.7), (5.11) следует, что интервально-оптимальным в рассматриваемой ЛКГ-задаче является линейный регулятор

$$U_t = -F(t) \hat{X}_t, \quad F(t) = \Phi^{-1}(t) K^T(t) L(t), \tag{5.13}$$

часть $L(t)$ матрицы коэффициентов которого $F(t)$ вычисляется заранее путем решения в обратном времени уравнения (5.12). Для этого на всем плановом отрезке управления $t \in [0, T]$ должны быть известны матрицы $A(t)$, $K(t)$ уравнения состояния объекта (5.1) и весовые матрицы $\Phi(t)$, $X(t)$ квадратичного критерия (5.3), а также его терминальная матрица Π .

Отметим, что матрица усиления $F(t)$ оценки \hat{X}_t в (5.13) совпадает с таковой у позиционного регулятора $u_t = u(t, x_t)$, оптимального в *детерминированной* линейно-квадратической задаче:

$$\dot{x}_t = A x_t + K u_t, \quad x_0 = m_0^x, \tag{5.14}$$

с соответствующим (5.3) интервальным критерием оптимальности

$$J = \frac{1}{2} \left\{ \int_0^T [x_\tau^T X(\tau) x_\tau + u_\tau^T \Phi(\tau) u_\tau] d\tau + x_T^T \Pi x_T \right\} \rightarrow \min,$$

так как в этом случае $u_t = -F(t) x_t$. Это обстоятельство и позволило называть инерционный закон стохастического управления (5.5), (5.13) термином “разделенный”. Итак, справедливо следующее утверждение.

Т е о р е м а 1 [8]. В интервальной ЛКГ-задаче (5.1)–(5.3) абсолютно оптимальное устройство управления (1.3) разделяется на линейный фильтр Калмана–Бьюси (5.5) с параметром, оперативно получаемым по (5.6), и интервально-оптимальный линейный детерминированный регулятор (5.13), настраиваемый на весь отрезок управления по (5.12).

5.3. О п е р а т и в н а я т е о р е м а р а з д е л е н и я. Если же оптимизировать функцию (5.7) по квадратическому О-критерию (5.4), то достаточно найти частный минимум (4.11) функции условного среднего (4.12). В данном случае она имеет вид

$$\zeta(t, \hat{x}; u) = \int (0.5(x^T X x + u^T \Phi u) + (A x + K u)^T \Upsilon x + 0.5 \text{tr}[Q \Upsilon]) \rho(t, x | \hat{x}) dx,$$

в котором от переменной $u \in \mathbb{R}^l$ зависят только два слагаемых:

$$\zeta(t, \hat{x}; u) = 0.5 u^T \Phi u + u^T K^T \Upsilon \bar{x}(t, \hat{x}) + \text{invar}(u),$$

где $\bar{x}(t, \hat{x}) = \int x \rho(t, x | \hat{x}) dx$ – условное среднее. Так как и здесь $\Phi > 0$, то минимум этой квадратической функции достигается в точке

$$u(t, \hat{x}) = -\Phi^{-1} K^T \Upsilon \bar{x}(t, \hat{x}). \quad (5.15)$$

При этом из линейности уравнений объекта (5.1), измерителя (5.2) и фильтра (5.5), а также из гауссовости их возмущения и начальных условий следует, что совместная плотность вероятности $r(t, x, \hat{x})$ является гауссовской:

$$r(t, x, \hat{x}) = N(x, \hat{x} | m_t^x, m_t^{\hat{x}}, D_t^x, D_t^{\hat{x}}, D_t^{x\hat{x}}),$$

а ее параметры могут быть определены из уравнений метода моментов Пугачева–Дункана. Тогда по теореме о нормальной корреляции [14] гауссова и условная плотность

$$\rho(t, x | \hat{x}) = N[x | \bar{x}(t, \hat{x}), \Gamma(t)],$$

а ее среднее и ковариация находятся по известным формулам

$$\bar{x}(t, \hat{x}) = m_t^x + D_t^{x\hat{x}} D_t^{\hat{x}\otimes} (\hat{x} - m_t^{\hat{x}}), \quad \Gamma(t) = D_t^x - D_t^{x\hat{x}} D_t^{\hat{x}\otimes} D_t^{\hat{x}\otimes},$$

где \otimes – символ псевдообращения матрицы по Муру–Пенроузу. Используя здесь известные свойства несмещенности оценки $M[\hat{X}_t] = M[X_t]$ и ее ортогональности к ошибке оценивания $\text{cov}(X_t - \hat{X}_t, \hat{X}_t) = 0$ [13, 15], имеем равенства $m_t^{\hat{x}} = m_t^x$, $D_t^{x\hat{x}} = D_t^{\hat{x}}$. В результате предыдущие общие соотношения существенно упрощаются:

$$\bar{x}(t, \hat{x}) = \hat{x}, \quad \Gamma(t) = D_t^x - D_t^{\hat{x}},$$

формула (5.15) принимает вид $u(t, \hat{x}) = -\Phi^{-1} K^T \Upsilon \hat{x}$, а из (5.7) окончательно получаем

$$U_t = -G(t) \hat{X}_t, \quad G(t) = \Phi^{-1}(t) K^T(t) \Upsilon(t). \quad (5.16)$$

Таким образом оперативно-оптимальным в ЛКГ-задаче тоже является линейный регулятор, но матрица его коэффициентов полностью известна из исходных данных. Действительно, сравнивая полученное соотношение (5.16) с (5.13), видим, что место определяемой из уравнения Риккати (5.12) матрицы $L(t)$ здесь заняла заданная весовая матрица $\Upsilon(t)$ терминального члена О-критерия (5.4). В этом состоит *принципиальное отличие* данного оперативно-оптимального случая от рассмотренного выше интервально-оптимального, когда для синтеза регулятора еще нужно решать и обратное уравнение Риккати (5.12), используя для этого дополнительную информацию о будущем.

Остается отметить, что в детерминированной задаче при управлении линейным объектом (5.14) по соответствующему (5.4) О-критерию

$$I_t = \frac{1}{2} \left\{ \int_0^t [x_\tau^T X(\tau) x_\tau + u_\tau^T \Phi(\tau) u_\tau] d\tau + x_t^T \Upsilon(t) x_t \right\} \rightarrow \min$$

оптимальным является подобный (5.16) закон позиционного управления $u_t = -G(t)x_t$ [9]. Итак, доказано следующее утверждение.

Т е о р е м а 2. В оперативной ЛКГ-задаче (5.1), (5.2), (5.4) абсолютно оптимальное устройство управления (1.3) разделяется на линейный фильтр Калмана–Бьюси (5.5), (5.6) и оперативно-оптимальный линейный детерминированный регулятор (5.16) с известной матрицей усиления.

Заключение. Поставлена и решена новая задача синтеза оперативно получаемого закона оптимального в среднем управления нелинейным стохастическим динамическим объектом по его измеряемому выходу. Она обобщает задачу синтеза локально-оптимального управления детерминированным объектом по измерениям его состояния [9]. При этом также применяется переменный во времени критерий оптимальности, который минимизирует потери, накопленные к текущему моменту времени, и не учитывает их будущие значения. Это позволяет не принимать во внимание и возможные в будущем изменения параметров и структуры управляемого объекта.

Для получения управления используется абсолютно вся статистическая информация о его случайном состоянии, которую дают проведенные измерения. Сосредоточенная в известном векторе апостериорных достаточных координат, получаемых с помощью инерционного нели-

нейного фильтра Стратоновича, она позволяет искать управление в виде функции этих координат. Следовательно, задача оптимизации управления сводится к синтезу такого безынерционно-го регулятора. Но построение традиционного интервально-оптимального регулятора, который обеспечивает минимизацию всех потерь на заданном плановом отрезке времени как прошлых так и будущих, требует априорного синтеза регулятора путем решения в обратном времени уравнения Беллмана.

Показано, что синтез предлагаемого регулятора может быть выполнен оперативно, в темпе со временем, решением задачи Коши для уравнения типа ФПК и задачи нелинейного программирования. Процедура синтеза продемонстрирована на примере ЛКГ-задачи, в результате чего доказана новая оперативная версия известной теоремы разделения.

СПИСОК ЛИТЕРАТУРЫ

1. *Стратонович Р.Л.* К теории оптимального управления. Достаточные координаты // *АиТ.* 1962. № 7. С. 910–917.
2. *Mortensen R.E.* Stochastic Optimal Control with Noisy Observations // *Int. J. Control.* 1966. V. 4. № 5. P. 455–466.
3. *Davis M.H.A., Varaiya P.P.* Dynamic Programming Conditions for Partially Observable Stochastic Systems // *SIAM J. Control.* 1973. V. 11. № 2. P. 226–262.
4. *Параев Ю.И.* Введение в статистическую динамику процессов управления и фильтрации. М.: Сов. радио, 1976.
5. *Benes V.E., Karatzas I.* On the Relation of Zakai's and Mortensen's Equations // *SIAM J. Control and Optimization.* 1983. V. 21. № 3. P. 472–489.
6. *Bensoussan A.* Stochastic Control of Partially Observable Systems. Cambridge: Cambridge University Press, 1992.
7. *Руденко Е.А.* Оперативно-оптимальный конечномерный динамический регулятор состояния стохастического дифференциального объекта по его выходу. I. Общий нелинейный случай // *Изв. РАН. ТиСУ.* 2022. № 5. С. 23–39.
8. *Wonham W.M.* On the Separation Theorem of Stochastic Control // *SIAM J. Control.* 1968. V. 6. № 2. P. 312–326.
9. *Верба В.С., 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.
10. *Пугачев В.С., Синицын И.Н.* Стохастические дифференциальные системы. Анализ и фильтрация. М.: Наука, 1985.
11. *Синицын И.Н.* Фильтры Калмана и Пугачева. М.: Логос, 2007.
12. *Пантелеев А.В., Руденко Е.А., Бортакровский А.С.* Нелинейные системы управления: описание, анализ и синтез. М.: Вузовская книга, 2008.
13. *Руденко Е.А.* Оптимальная структура непрерывного нелинейного фильтра Пугачева пониженного порядка // *Изв. РАН. ТиСУ.* 2013. № 6. С. 25–51.
14. *Ширяев А.Н.* Вероятность. М.: Наука, 1980.
15. *Браммер К., Зиффлинг Г.* Фильтр Калмана–Бьюси / Пер. с англ. М.: Наука, 1982.