УДК 004.9:(005.94:78)

СИНТЕЗ АЛГОРИТМОВ КОМПЕНСАЦИИ ВОЗМУЩЕНИЙ В СИСТЕМАХ ДУАЛЬНОГО УПРАВЛЕНИЯ

В.П. Живоглядов

Рассматривается новый общий подход к синтезу алгоритмов дуального управления, названный “дуальное управление 2.0”. Предложенная методология позволяет осуществлять аналитический синтез оптимальных регулярных и рандомизированных алгоритмов управления объектами со случайными параметрами в задачах, не поддающихся аналитическому решению в рамках классической теории дуального управления. Получены конкретные алгоритмы.

Ключевые слова: система дуального управления; дуальное управление 2.0; синтез алгоритмов; многокритериальная оптимизация; рандомизированная стратегия; компенсация возмущений.

Введение и постановка задачи. Предложенная в 1960 г. А.А. Фельдбаумом [1, 2] теория дуального управления (ТДУ) явилась крупным вкладом в развитие современной теории автоматического управления. ТДУ основана на использовании байесова подхода, методов теории статистических решений и динамического программирования. Выделен класс систем оптимального управления с активным накоплением информации. При этом управляющие воздействия носят двойственный дуальный характер: приведение объекта в требуемое состояние в текущий момент времени и обеспечение лучшего управления в будущем за счет активного изучения характеристик объекта. Хотя теория дуального управления с активным изучением объекта возникла в технической сфере, идея оптимального компромисса между текущими интересами и зондированием оказалась привлекательной для специалистов в области макроэкономики и организационного управления, информационных систем управления, математики, прикладных экономических исследований, стратегическом ценообразовании, разработки стратегии ценообразования при внедрении новых продуктов и др. [3–6]. В обзоре литературы [6] отмечается большой перечень исследований численных задач, причем лишь немногие простые примеры были решены. Во многих случаях эффективное использование функциональных уравнений динамического программирования пока не представляется возможным. Применение численных методов динамического программирования сталкивается с огромными трудностями в связи с так называемым проклятием размерности.

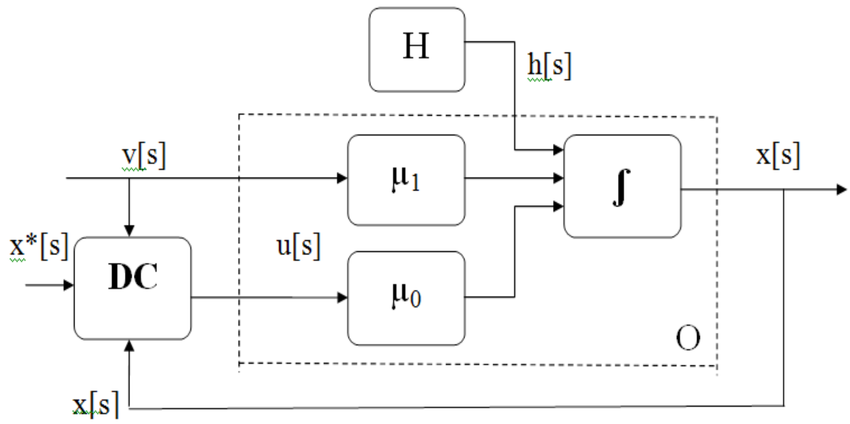

Рисунок 1 – Структура системы управления

Нужны более эффективные способы борьбы с этим явлением.

В данной работе используется и развивается новый конструктивный подход к синтезу алгоритмов дуального управления. Предложенная в [7, 8] методология дуального управления 2.0 позволяет преодолеть многие трудности, которые встречаются в классической ТДУ. Рассмотрим задачу управления объектом, модель которого описывается системой уравнений S(μ,s) с вектором неизвестных параметров $\mu=(\mu_1, \mu_2)^T$:

$$S(\mu,s):\quad \begin{cases} x[s]-f_x(x[s-1])-\mu_1 f_u(u[s])-\mu_2 v[s]-h[s]=0, \\ s-1,2,...n. \end{cases} \tag{1}$$

Структура системы дуального управления представлена на рисунке 1. Объект **O** включает усилители μ_1 и μ_2 и интегрирующее звено ∫ с обратной связью; **H** – источник внешних случайных помех h[s]. Дуальный контроллер **DC** на основе измерений выхода объекта x[s], внешних контролируемых возмущений v [s] и известного задания x*[s] вырабатывает управляющие воздействия u[s], где s – дискретное время, s = 1,2, …, n. Функции f_u и f_x – известные и взаимно однозначные.

Без потери общности можно рассматривать более простую модель:

$$S(\mu,s):\quad \begin{cases} x[s]-f_x(x[s-1])-\mu_1 u[s]-\mu_2 v[s]-h[s]=0, \\ s-1,2,...n. \end{cases} \tag{2}$$

Применяется байесов подход. Плотности вероятности случайных факторов *P(μ)* и *P(h[s])* известны. Считаем, что параметры μ и помехи h[s] подчиняются нормальному закону распределения вероятностей, то есть плотности вероятности имеют вид

$$P(\mu)\sim N\left(m[1],\ Q^{-1}[1]\ \right) \qquad P(h[s])\sim N\left(0,\ \sigma^2\right), \tag{3}$$

где $m[1]$ – априорное математическое ожидание μ, а $Q^{-1}[1]$ – априорная ковариационная матрица; σ^2 - дисперсия помехи *h[s], E{ h[s]}* = 0;

E{.} – знак математического ожидания. Апостериорную плотность вероятности перед s-м тактом управления обозначим $P_s(\mu)\sim N(m[s],Q^{-1}[s]\ )$.

Требуется найти такую последовательность стратегий – вектор $\Gamma = \left(\Gamma[1],\ \dots,\ \Gamma[s],\ \dots,\Gamma[n]\right)^T$, чтобы максимизировать векторный критерий успешности управления $M = (M[1]\dots M[s]\dots M[n])^{\mathrm{T}}$ при некоторых ограничениях *G* на область изменения управлений *u[s]. Здесь M[s]= E{e[s]}* – локальный критерий для s-го момента времени; e[s] – специально выбранная целевая функция. Стратегию управления

$$\Gamma[s]\ =P(u[s]\,|\,x^*[s],I[s],v[s]) \tag{4}$$

предполагаем для общности случайной (рандомизированной), заданной в виде условной плотности вероятности *P(.)*, зависящей от задания *x*[s]*, измеряемого возмущения *v[s]* и информации $I[s]=I_s$, накопленной к *s*-му моменту времени и содержащейся во множестве измерений входов и выхода объекта

$$(u^{\to}[s-1],v^{\to}[s-1],x^{\to}[s-1])\ . \tag{5}$$

Использованы обозначения вида

$$\begin{aligned} u^{\to}[s]&=\left(u[1],\dots,u[s]\right)^T, \\ v^{\to}[s]&=\left(v[1],\dots,v[s]\right),^T \\ x^{\to}[s]&=\left(x[1],\dots,x[s]\right)^T. \end{aligned} \tag{6}$$

Стрелками сверху отмечены временные векторы.

Г[s]* – локально оптимальная стратегия управления в *s*-й момент. Если стратегия регулярная, то *Г[s]* превращается в дельта-функцию, локально оптимальное управление *u*[s]* зависит от I_s, v[s] и задания *x*[s]*.

Методология дуального управления 2.0. Предложенная в [7, 8] и развитая в данной работе методология, названная “дуальное управление 2.0”, принципиально отличается от принятой в классической теории дуального управления. Она включает следующие этапы.

Этап 1. Исследование задачи оптимального детерминированного управления с заданной целевой функцией $W[s] = W(x[s], x^*[s]) \sim \min\limits_x$.

Здесь $x^*[s]$ – задание. Результат этапа: получение зависимости между параметрами μ и управляющими воздействиями *u* при задании целевой функцией *W* и отсутствии помех *h*.

Этап 2. Решение задачи оценки параметров μ в рамках байесового подхода. Результат: получение апостериорной плотности вероятности $P(\mu|I[s])$.

Этап 3. Интеграция решений первых двух задач управления и идентификации с выбором специальной целевой функции *e[s]* и введением локального для s-го момента критерия *J[s]* – индекса успешности управления. Функции *J[s]* формируются путем интеграции специальным образом апостериорных плотностей вероятности неизвестных параметров $P_s(\mu) = P(\mu | I_s)$ и условий оптимальности управления при полной информации о характеристиках объекта. Решение локальной задачи оптимизации по локальному критерию *M[s]:*

$$M[s] = E\{J[s]\} = \int\limits_{\Omega(u^{\rightarrow}[s-1], v^{\rightarrow}[s-1], x^{\rightarrow}[s-1])} J[s] P\left(u^{\rightarrow}[s-1], \text{v}^{\rightarrow}[s-1], x^{\rightarrow}[s-1]\right) d\Omega. \quad (7)$$

Результат этапа: получение алгоритма пассивно адаптивного управления с регулярной стратегией, то есть нахождение локально оптимального управления u*[s]:

$$\Gamma[s] = P(u[s] \mid I[s], x*[s]) = \delta(u[s] - u*[s]). \quad (8)$$

Этап 4. Введение в рассмотрение векторного критерия успешности управления $M = (M[1] \ldots M[s] \ldots M[n])^{\mathbf{T}}$ с последующим выбором алгоритма многокритериальной оптимизации, учет ограничений G на область изменения управлений типа u[s] ϵ Ω(u[s]) и вероятностных ограничений вида E{u[s]} = u*[s]. Результат: получение регулярных или рандомизированных стратегий активно адаптивного управления. Заметим, что в классической теории дуального управления оптимальные стратегии – регулярные.

Синтез локально оптимальных алгоритмов управления с пассивным накоплением информации. Рассмотрим применение общей методологии к исследованию конкретной задачи управления объектом, описываемым моделью (2).

Исследование задачи оптимального детерминированного управления предусматривает получение зависимости между параметрами μ и управляющими воздействиями *u* при задании целевой функцией *W* и отсутствии помех *h*, то есть нахождение корня системы уравнений

$$\mu* = root\{S(\mu, s) | h[s] = 0, x[s] = \arg\min W\}. \quad (9)$$

Пусть целевая функция квадратическая $W[s] = (x[s] - x^*[s])^2$, помеха h отсутствует, вектор μ двумерный. Тогда

$$\mu* = root\{S(\mu, s, i) | x[s] = x*[s])\}. \quad (10)$$

Система $S(\mu, s, i)$ для нахождения μ* включает два уравнения для s-го и i-го моментов времени, i < s.

$S(\mu, s, i)$:

$$\begin{cases} x[s] - f_x(x[s-1]) - \mu_1 u[s] - \mu_2 v[s] = 0 \\ x[i] - f_x(x[i-1]) - \mu_1 u[i] - \mu_2 v[i] = 0 \end{cases} \quad (11)$$

Введем обозначения:

$$X(s,i) = (x[s] - f_x(x[s-1]), x[i] - f_x(x[i-1]))^T,$$

$$X^*(s,i) = (x*[s] - f_x(x[s-1]), x[i] - f_x(x[i-1]))^T.$$

С учетом этих обозначений из (11) находим:

$$\mu^* = U^{-1}(s,i) X^*(s,i). \quad (12)$$

Введем локальный индекс успешности управления J[s] в s-й момент:

$$J[s] = \int\limits_{\Omega(u[s],\mu)} e[s]P(u[s],\mu \mid x^*[s], I[s], v[s])d\Omega\,, \tag{13}$$

Учитывая, что стратегии *Г[s]* при заданных *I[s]* не зависят от μ, получим:

$$\Gamma[s] = P(u[s] \| x^*[s], I[s], v[s], \mu) = P\left(u[s] \| x^*[s], I[s], v[s]\right). \tag{14}$$

Преобразовав (13) при $e[s] = \delta(\mu - \mu^*)$, получим:

$$J[s] = \int\limits_{\Omega(u[s],\,\mu)} e[s]P\left(u[s] \| x^*[s], I[s], v[s]\right)P(\mu \mid I[s])d\Omega,$$

$$J[s] = \int\limits_{\Omega(u[s])} \alpha[s]\Gamma[s]\, d\Omega, \tag{15}$$

$$\alpha[s] = \int\limits_{\Omega(\mu)} \delta(\mu - \mu^*)\, P(\mu \mid I[s])\, d\Omega \sim max. \tag{16}$$

Случайные параметры $\mu=(\mu_1, \mu_2)^T$ и помехи измерений *h* подчиняются нормальному закону. Для модели объекта, линейной по параметрам, апостериорная плотность вероятности также гауссовская [9]:

$$P_s(\mu) = (2\pi)^{-1}\sqrt{|Q[s]|}exp\left\{-\frac{1}{2}(\mu - m[s])^T Q[s](\mu - m[s])\right\}, \tag{17}$$

где $m[s] = (m_1[s], m_2[s])^T$ – математическое ожидание параметра μ на основе информации I_s;

$Q_s^{-1} = Q^{-1}[s]$ – ковариационная матрица апостериорного распределения.

$$m[s] = Q_s^{-1}(\sigma^{-2}\begin{pmatrix} u[s-1] \\ v[s-1] \end{pmatrix} y[s-1] + Q_{s-1}m[s-1]), \tag{18}$$

$$Q_s = Q_{s-1} + \sigma^{-2}\begin{pmatrix} u[s-1] \\ v[s-1] \end{pmatrix}\begin{pmatrix} u[s-1] \\ v[s-1] \end{pmatrix}^T, \tag{19}$$

$$y[s] = x[s] - f_x(x[s-1]). \tag{20}$$

Подставив (18) в (17), после интегрирования получим:

$$J[s] = \sqrt{|Q[s]|(2\pi)^{-2}}exp\{-\frac{1}{2}(\mu^* - m[s])^T Q[s](\mu^* - m[s])\},$$

$$J[s] = (2\pi)^{-1}\sqrt{|Q[s]|}exp\left\{-\frac{1}{2}(U^{-1}(s,i)X^*(s,i) - m[s])^T,\right.$$

$$\left. Q[s](U^{-1}(s,i)X^*(s,i) - m[s])\right\}. \tag{21}$$

Из (19) видно, что максимумы по *u[s]* показателей *J[s]* и соответственно *α[s]* достигаются при выполнении условия

$$\mu^* - m[s] = U^{-1}(s,i)X^*(s,i) - m[s] = 0. \tag{22}$$

Отсюда получаем:

$$X^*(s,i) - U(s,i)m[s] = 0\,. \tag{23}$$

Или в развернутом виде

$S^*(m[s], s, i)$:

$$\begin{cases} x*[s] - f_x(x[s-1]) - m_1[s]u[s] - m_2[s]v[s] = 0 \\ x[i] - f_x(x[i-1]) - m_1[s]u[i] - m_2[s]v[i] = 0 \end{cases}. \tag{24}$$

Управление входит только в первое уравнение и не зависит от переменных в *i*-е моменты времени. Из (24) находим

$$u[s] = u^*[s] = (m_1[s])^{-1}(x^*[s] - f_x(x[s-1]) - m_2[s]v[s]). \tag{25}$$

При этом *max J[s] = J*[s]* равны

$$J^*[s] = \sqrt{|Q[s]|(2л)^{-2}} = (2л)^{-1}\sqrt{|Q[s]|}, \tag{26}$$

где $|Q[s]|$ – определитель матрицы *Q[s]*.

В формулах (25) и (26) *u*[s]* и *J*[s]* – локально оптимальные значения управления и критерия. По структуре алгоритм (25) эквивалентен алгоритму управления в условиях полной определенности.

Синтез алгоритмов дуального управления с активным накоплением информации. Рассмотрим далее задачу многокритериальной оптимизации по вектору критериев М с ограничениями *G* двух типов (27) и (28):

$$u[s] \in \Omega_u[s] : u^*[s] - \Delta[s] \le u[s] \le u^*[s] + \Delta[s], \tag{27}$$

$$E\{u[s]\} = \int\limits_{\Omega(u)} u[s]\,Г[s]\,d\Omega = u^*[s]. \tag{28}$$

Ограничение (27) определяет допустимую область $\Omega_u[s]$ изменения *u[s]* в окрестности локально оптимального управления *u*[s]*. Стохастическое ограничение (28) означает, что управления *u[s]* в моменты (1,..., *s*,..., *n*-1) должны быть в среднем равны локально оптимальным значениям *u*[s]*.

От управления *u[n]* зависит только критерий *J[n]*. Поэтому оптимальное управление $u^{opt}[n]$ в последний момент времени *n* совпадает с локально оптимальным управлением *u*[n]*. Перейдем к моменту времени *(n-1)*.

Значения локально оптимального управления и критерия получаем из (25), (26) при *s* = *n*-1.

$$u[s] = u^*[s] = (m_1[s])^{-1}(x^*[s] - f_x(x[s-1]) - m_2[s]v[s]), \tag{29}$$

$$J^*[s] = \sqrt{|Q[s]|(2л)^{-2}} = (2л)^{-1}\sqrt{|Q[s]|}. \tag{30}$$

Оптимальное по двум критериям управление $u^{opt}[n-1]$ находим из условия

$$(J[n-1], J^*[n\,|\,n-1])^T \sim \max_{u[n-1]}. \tag{31}$$

Вычислим оценку *J*[n|n-1]* критерия *J*[n]* на основе информации, накопленной к *(n-1)*-му такту:

$$J^*[n\,|\,n-1] = \int\limits_{\Omega(y[n-1])} J^*[n]\,P\big(y[n-1]\,|\,I[n-1],\ u[n-1]\big)d\Omega$$

Видно, что *J*[n]* не зависит от *x[n-1]*. Поэтому *J*[n|n-1] = J*[n]*. Аналогично

$$J^*[n|s] = J^*[n]. \tag{32}$$

При этом значение локального критерия равно:

$$\max J[n-1] = J^*[n-1] = (2л)^{-1}\sqrt{|Q[n-1]|}. \tag{33}$$

Оптимальное по двум критериям управление $u^{opt}[n-1]$ находим из условия (24), применяя метод уступок.

В системе возможно активное накопление информации, если *J*[n|s]* зависит от управлений u[s]. Пусть $|Q[n]|$ – выпуклая вниз функция *u[s], s =1,...,n-1.* Тогда все оптимальные по векторным критериям типа (31) и (36) управления *u[s], s =1,...,n-1,* лежат на границах допустимых областей (27). Вычислим Q[s+1]:

$$Q[s+1] = Q[s] + \sigma^{-2}\begin{pmatrix} u^2[s] & u[s]v[s] \\ u[s]v[s] & v^2[s] \end{pmatrix} = Q[s] + \sigma^{-2}\overline{\overline{U[s]}};$$

$$Q[s+1] = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix} + \sigma^{-2}\begin{pmatrix} u^2[s] & u[s]v[s] \\ u[s]v[s] & v^2[s] \end{pmatrix};$$

$$Q[s+1]=\begin{pmatrix} q_{11}+\sigma^{-2}u^2[s] & q_{12}+\sigma^{-2}u[s]v[s] \\ q_{21}+\sigma^{-2}u[s]v[s] & q_{22}+\sigma^{-2}v^2[s] \end{pmatrix}.$$

Определитель ковариационной матрицы равен

$$\left|Q[s+1]\right| = a_2u^2[s]-a_1u[s]v[s]+a_0 \qquad a_2>0\,. \tag{34}$$

Проанализируем вид функций:

$\left|Q[s+1]\right| = f_q(u[s])$ и $\left|Q[s+i]\right| = f_q(u[s])$,

$Q[s+i]=Q[s]+\sigma^{-2}\sum_{j=1}^{i}\overline{\overline{U[s+j]}}=\theta(s,i)+\sigma^{-2}\overline{\overline{U[s]}}$, i=1,2,…n-s.

$$\left|Q[s+i]\right| = a_{2i}u^2[s]-a_{1i}u[s]v[s]+a_{0i}\,,\quad a_{2i}>0 \tag{35}$$

Все критерии *J*[s+i]* $J^*[s+i]$ $J^*[s+i]$ являются выпуклыми вниз функциями u[s]. Как показано выше, от *u[s]* зависит не только критерий *J[s]*, но и все *J*[i]* для последующих моментов времени *i = s+1, s+2, …, n.*

$$\left(J[s],\ J[s+1\,|\,s],\ \ldots, J[n\,|\,s]\right)^T \sim \max_{u[s]} \tag{36}$$

Применим метод уступок для любого s.

Все оптимальные по векторным критериям типа (24) управления *u[s], s =1,...,n-1,* лежат на границах допустимых областей

$$u^*[s]-\Delta[s]\le u[s]\ \le\ u^*[s]+\Delta[s]\,.$$

Перейдем к синтезу алгоритмов дуального управления при наличии вероятностных ограничений. Стохастическое ограничение (28) *E{u[s]} = u*[s], s = 1,2, …, n* предполагает возможность применения рандомизированной стратегии. Решение будем искать в классе случайных стратегий *Γ[s].*

В полученной таким образом задаче стохастической оптимизации все критерии *J*[s], s* = 1,2, …, *n,* $J^*[s], s=1,2,\ \ldots, n,\ J^*[s], s=1,2,\ \ldots, n,$ – это выпуклые вниз функции управлений $(u[1],\ \ldots\ ,u[s-1])^T$. Поэтому оптимальная стратегия $\Gamma^{opt}[s]$ смешанная (рандомизированная):

$$\Gamma^{opt}[s]=\begin{pmatrix} \delta(u[s]-u^*[s]-\Delta[s]) & \textit{с вероятностью } 0{,}5 \\ \delta(u[s]-u^*[s]+\Delta[s]) & \textit{с вероятностью } 0{,}5 \end{pmatrix}. \tag{37}$$

Перепишем (37) в более простом виде:

$$u^{opt}[s]=\begin{pmatrix} u^*[s]+\Delta[s] & \textit{с вероятностью } 0{,}5 \\ u^*[s]-\Delta[s] & \textit{с вероятностью } 0{,}5 \end{pmatrix}. \tag{38}$$

Предложенная методология нового общего подхода к синтезу алгоритмов дуального управления, названная “дуальное управление 2.0”, позволила осуществлять аналитический синтез оптимальных регулярных и рандомизированных алгоритмов управления объектами со случайными параметрами. Рассмотренная задача не поддается аналитическому решению в рамках классической теории дуального управления. Предложенный подход открывает широкие возможности проектирования оптимальных алгоритмов активно-адаптивного управления.

Литература

1. *Фельдбаум А.А.* Теория дуального управления. I, II, III, IV / А.А. Фельдбаум // Автоматика и телемеханика. 1960. № 21 (9), 21 (11). 1961. № 22 (1), 22 (2).
2. *Фельдбаум А.А.* Основы теории оптимальных автоматических систем / А.А. Фельдбаум. М.; Л.: Физматгиз, 1963. 552 с.
3. *Живоглядов В.П.* Адаптация в автоматизированных системах управления технологическими процессами / В.П. Живоглядов. Фрунзе: Илим, 1974.
4. *Filatov N.M., and Unbenhauen H.* Adaptive Dual Control: Theory and Applications. Vol. 302 of Lecture Notes in Control and Information Sciences. New York: Springer-Verlag, 2004.

5. *Åström K.J. and Helmersson A.* Dual control of an integrator with unknown gain (1986) // Computers & Mathematics with Applications. Vol. 12, Issue 6, Part A, June 1986, Pages 653–662.
6. *Morozov S.* Bayesian Dual Control: Review of the Literature. http://www.wavelet3000.org/images/litreview.pdf *24.04.2008*
7. *Живоглядов В.П.* Построение альтернативной теории дуального управления / В.П. Живоглядов // Вестник КРСУ. 2012. Т. 12. № 10.
8. *Живоглядов В.П.* Многокритериальное дуальное управление экстремальными объектами / В.П. Живоглядов // Известия НАН КР. 2012. № 3.
9. *Кашьяп Р.Л.* Построение динамических стохастических моделей по экспериментальным данным / Р.Л. Кашьяп, А.П. Рао; пер. с англ. М.: Наука; Физматгиз, 1983. 384 с.