УДК 62-50 (575.2) (04)

ПОСТРОЕНИЕ АЛЬТЕРНАТИВНОЙ ТЕОРИИ ДУАЛЬНОГО УПРАВЛЕНИЯ

В.П. Живоглядов

Рассматривается методология построения теории управления, альтернативной классической теории дуального управления, показано использование многокритериальной оптимизации для синтеза систем дуального управления с регулярными и рандомизированными стратегиями.

Ключевые слова: дуальное управление; многокритериальная оптимизация; синтез алгоритмов дуального управления; управление мультипликативным объектом.

Постановка задачи. Разработчик теории дуального управления А.А. Фельдбаум внес фундаментальный вклад в развитие теории автоматического управления. Первые его работы по теории дуального управления были опубликованы в 1960–1961 гг. [1, 2]. Систематически основы теории дуального управления изложены в работе [3]. В дальнейшем они были развиты в работах учеников и последователей (см., например, [4–7]). Теоретически было установлено, что существует класс не приводимых к разомкнутым систем, в которых возможны процессы активного накопления информации, то есть класс систем активно-адаптивного управления в условиях неполной информации об объекте. При этом управляющие воздействия носят двойственный дуальный характер. С одной стороны, они направлены на приведение объекта в требуемое состояние (направляющая функция) в текущий момент времени. С другой – они должны обеспечить лучшее изучение характеристик объекта и соответственно лучшее управление в будущем (изучающая функция). Применены байесов подход, методы теории статистических решений и динамического программирования. Оптимальными считаются стратегии управления, обеспечивающие минимум полного риска – математического ожидания суммы удельных функций потерь.

В рамках принятой формализованной модели доказано, что оптимальной стратегией управления является регулярная неслучайная стратегия. Хотя теория управления с активным изучением возникла в технической литературе, идея компромисса между осторожностью и зондированием произвела впечатление на экономистов, появились попытки применения этого подхода к проблеме макроэкономической стабилизации. В обзоре литературы [7] отмечается большой перечень исследований численных задач, решений функциональных уравнений и при этом лишь немногие простые примеры были решены. Часто рассматривались линейно-квадратические гауссовские задачи. Предпринимались попытки найти простые неоптимальные решения.

Что же препятствует широкому применению данной теории, несмотря на всю привлекательность идей дуального управления? Главное препятствие – чрезвычайные аналитические и вычислительные трудности даже для простых моделей объектов.

В данной работе предпринята попытка заложить основы альтернативной теории дуального управления, в которой используются другие формализованные модели и которая позволяет находить решения в явном виде задач, не поддающихся решению в классической теории дуального управления. Показано, что утверждение о том, что оптимальная стратегия управления регулярная неслучайная, определяется не спецификой активно-адаптивного управления, а принятой в классической теории дуального управления формализованной моделью.

Рассмотрим предлагаемую методологию на примере одномерных объектов без памяти и с дискретным временем. Возможно ее распространение на многомерные объекты с памятью.

Пусть модели объекта и наблюдений имеют вид:

$$f(q[s], u[s], \mu) = 0,\ y[s] = H(q[s], h[s])\ ,\ s = 1,2, \ldots, n, \quad (1)$$

где s – дискретное время; n задано; q[s] – выход объекта, u[s] – управление; μ – неизвестный случайный параметр; {h[s]} – последовательность независимых случайных величин – помех измерения.

Функции f(.) и H(.) – известные взаимно однозначные такие, что известны обратные функции:

$\mu = f_\mu(q[s], u[s]), u[s] = f_u(q[s], \mu), q[s] = f_q(u[s], \mu),$ (2)

$h[s] = H(q[s], y[s])$, например, $h[s] = y[s] - q[s]$. (3)

Применяется байесов подход. Плотности вероятности случайных факторов $P(\mu)$ и $P(h[s])$ известны. Цель управления состоит в том, чтобы обеспечить совпадение или близость выхода объекта $q[s]$ и задающего воздействия $q^\circ[s]$. Ошибку управления обозначим как $e[s] = q^\circ[s] - q[s]$.

Качество управления характеризуется статистическим векторным критерием – индексом успешности управления:

$$M = (M[1], \dots, M[s], \dots, M[n])^T \sim \max, \quad (4)$$

где $M[s]$ – локальный индекс успеха для момента времени s.

Требуется найти вектор оптимальных по критерию M стратегий управления

$$\Gamma = (\Gamma[1], \dots, \Gamma[s], \dots, \Gamma[n])^T, \quad (5)$$

где $\Gamma[s] = P(u[s] \mid I[s], q^\circ[s])$ – условная плотность вероятности управления $u[s]$ при фиксированной информации $I[s]$, накопленной к s-му моменту времени и содержащейся в множестве наблюдений $\{\vec{u}[s-1], \vec{y}[s-1]\}$,

где $\vec{u}[s-1] = \left(u[1], \dots, u[s-1]\right)^T$,

$$\vec{y}[s-1] = \left(y[1], \dots, y[s-1]\right)^T. \quad (6)$$

Другими словами, решение будем искать в общем случае в классе рандомизированных стратегий управления.

Общая методология исследования. Описанная выше многокритериальная задача не полностью определена. От того, как она будет доопределена, зависит выбор метода решения. Методология, принятая в данной работе, существенно отличается от принятой в классической теории дуального управления. В отличие от традиционной теории дуального управления качество управления в s-й момент времени будем оценивать значениями соответствующих условных плотностей вероятности. Локальный индекс успеха для момента времени s:

$$M[s] = \int\limits_{\Omega(u[s],\ \vec{u}[s-1], \vec{y}[s-1], \mu)} F\left(e[s]\right) P(u[s], \vec{u}[s-1], \vec{y}[s-1], \mu) d\Omega,$$

$$M[s] = \int\limits_{\Omega(\vec{u}[s-1], \vec{y}[s-1])} J[s] P(\vec{u}[s-1], \vec{y}[s-1]) d\Omega, \quad (7)$$

где $F(e[s])$ – целевая дельта-функция,

$$J[s] = \int\limits_{\Omega(u[s],\mu)} F\left(e[s]\right) P(u[s], \mu \mid I[s], q^\circ[s]) d\Omega,$$

$$J[s] = \int\limits_{\Omega(u[s],\ \mu)} F\left(e[s]\right) P(u[s] \mid I[s], \mu, q^\circ[s]) P(\mu \mid I[s]) d\Omega. \quad (8)$$

Учтем, что стратегии $\Gamma[s]$ при заданных $I[s]$, $q^\circ[s]$ не зависят от μ, получим:

$$\Gamma[s] = P(u[s] \| I[s], \mu, q^\circ[s]) = P(u[s] \mid I[s], q^\circ[s]). \quad (9)$$

Выбрав целевую функцию в виде дельта-функции $F(e[s]) = \delta(\mu - f_\mu(q^\circ[s], u[s]))$ и обозначив апостериорную плотность вероятности неизвестного параметра $P(\mu \mid I[s]) = P_s(\mu)$, получим:

$$J[s] = \int\limits_{\Omega(u[s])} \alpha[s] \Gamma[s]\, d\Omega, \quad (10)$$

где $\alpha[s] = \int\limits_{\Omega(\mu)} \delta(\mu - f_\mu(q^\circ[s], u[s]))\, P_s(\mu) d\Omega \geq 0.$ (11)

Вектор оптимальных стратегий $\Gamma = (\Gamma[1], \dots, \Gamma[s], \dots, \Gamma[n])^T$ находим, максимизируя векторный критерий $J = (J[1], \dots, J[s], \dots, J[n])^T$, начиная с последнего такта n. Локально оптимальную стратегию $\Gamma^*[s]$ определяем из условия максимума $J[s]$. Можно показать, используя (10) – (11) и теорему о среднем значении, что локально оптимальная стратегия $\Gamma^*[s]$ регулярная: $\Gamma^*[s] = \delta(u[s] - u^*[s])$. Для получения конкретных решений необходимо доопределить постановку задачи. Проведем это в несколько этапов:

- конкретизации для нормального закона распределения вероятностей;
- синтеза локально оптимального управления мультипликативным объектом;
- многокритериальной оптимизации управления,
- синтеза алгоритмов дуального управления при наличии вероятностных ограничений.

Конкретизация для нормального закона распределения вероятностей. Пусть случайный параметр и помеха измерений подчиняются нормальному закону: $P(\mu) = N(\mu_0, D_0)$, $P(h[s]) = N(0, \sigma^2)$ и $P_s(\mu) = N(m[s], D[s])$. Тогда из (11) получаем:

$$\alpha[s] = \left(\sqrt{2\pi D[s]}\right)^{-1}$$

$$\exp\left\{-\frac{(m[s] - f_\mu(q^\circ[s], u[s]))^2}{2D[s]}\right\} \geq 0. \quad (12)$$

Максимумы по $u[s]$ (12) и соответственно (10) достигаются при

$$m[s] - f_\mu(q^\circ[s], u[s]) = 0, \quad (13)$$

$$u[s] = u^*[s] = f_u(q^\circ[s], m[s]). \quad (14)$$

При этом максимальные значения $\alpha[s]$ и $J[s]$ равны:

$$\max J[s] = J^*[s] = \alpha^*[s] = \left(\sqrt{2\pi D[s]}\right)^{-1}. \quad (15)$$

Заметим, что они не зависят от $m[s]$ и соответственно от $(y[1], \dots, y[s-1])^T$.

Синтез локально оптимального управления. Рассмотрим задачу дуального управления мультипликативным объектом с моделью:

$$q[s]=f_q(u[s],\mu)=\mu\, u[s],\; y[s]=q[s]+h[s], \quad (16)$$

$$P(\mu)=N(\mu_0, D_0),\; P(h[s])=N(0,\sigma^2).$$

Коэффициент μ не известен. В рамках классической теории дуального управления задача не имеет точного аналитического решения. На ней испытывались различные приближенные методы и вычислительные алгоритмы [4–6]. В рамках предложенной методики решение находится легко. Из (12) получаем локально оптимальное управление:

$$u^*[s]=f_u(q^\circ[s], m[s])=q^\circ[s]\,/\,m[s]. \quad (17)$$

Предполагаем, что m[s] не равны нулю, $m[s]\neq 0$.

Значение критерия определяется формулой (15):

$$J^*[s]=\left(\sqrt{2\pi D[s]}\right)^{-1}.$$

Статистики m[s] и D[s] вычисляются по формулам:

$$m[s]=(m[s-1]+r_{s-1}u[s-1]y[s-1])(1+r_{s-1}(u[s-1])^2)^{-1}, \quad (18)$$

$$D[s]=D[s-1]\,(1+r_{s-1}(u[s-1])^2)^{-1} \quad (19)$$

или

$$\left(D[s]\right)^{-1}=\left(D[s-1]\right)^{-1}+\sigma^{-2}\left(u[s-1]\right)^2= =\left(D_0\right)^{-1}+\sigma^{-2}\sum_{i=1}^{s-1}\left(u[i]\right)^2, \quad (20)$$

где $r_{s-1}=D[s-1]/\sigma^{-2}$

По структуре алгоритм (17) эквивалентен алгоритму управления в условиях полной определенности. Меняя целевую функцию F(e[s]) можно получить другие структуры. Так, выбрав ее в виде дельта-функции

$$F\left(e[s]\right)=\delta(\mu - (2u)^{-1}[q^\circ+\sqrt{(q^\circ)^2-4(u)^2D}\,]), \quad (21)$$

получим:

$$u^*[s]=\frac{q^\circ[s]\,m[s]}{\left(m[s]\right)^2+D[s]}. \quad (22)$$

Это так называемое осторожное управление с пассивным накоплением информации при квадратической функции потерь и отсутствии эффекта дуальности [4–6].

Многокритериальная оптимизация: метод уступок. Рассматриваемая задача синтеза дуального управления – это специфическая многокритериальная задача. От управления u[n] зависит только критерий J[n]. Поэтому оптимальное управление $u^\circ[n]$ в последний момент времени n совпадает с локально оптимальным управлением u*[n].

При этом

$$\max J[n]=J^*[n]= =(\sqrt{2\pi D[n]})^{-1}=(\sqrt{2\pi})^{-1}(\sqrt{D[n]})^{-1},$$

$$(D[n])^{-1}=(D_0)^{-1}+\sigma^{-2}\sum_{i=1}^{n-1}\left(u[i]\right)^2,$$

$$\frac{1}{D[n]}=\frac{1}{D_0}+\sigma^{-2}\sum_{i=1}^{n-1}\left(u[i]\right)^2. \quad (23)$$

Перейдем к моменту времени (n-1). Из (17) получаем локально оптимальное управление:

$$u^*[n-1]=f_u(q^\circ[n-1], m[n-1])=q^\circ[n-1]\,/\,m[n-1]. \quad (24)$$

При этом значение локального критерия равно:

$$\max J[n-1]=J^*[n-1]= =(\sqrt{2\pi D[n-1]})^{-1}.$$

Оптимальное управление $u^\circ[n-1]$ находим из условия:

$$(J[n-1], J^*[n\,|\,n-1])^T\sim\max_{u[n-1]}. \quad (25)$$

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

$$J^*[n\,|\,n-1]=\int_{\Omega(y[n-1])}J^*[n]\,P\left(y[n-1]\,|\,I[n-1],u[n-1]\right)d\Omega.$$

Заметим, что J*[n] не зависит от y[n-1].

Поэтому J*[n| n-1] = J*[n].

Функция зависимости J*[n] от u[n-1] выпуклая вниз монотонно и неограниченно возрастающая с ростом $(u[n-1])^2$.

Для решения оптимизационной задачи (25) применим метод уступок по отношению к локально оптимальному значению управления

$$u^*[n-1]-\Delta[n-1]\le u[n-1]\le u^*[n-1]+\Delta[n-1]$$

и выбирем Δ [n-1] в зависимости от уровня текущей неопределенности:

$\Delta[n-1]=K\sqrt{D[n-1]}$, где K – коэффициент пропорциональности.

Оптимальное дуальное управление в (n-1)-й момент лежит на границе допустимой области и равно:

$$u^\circ[n-1]=(\mathrm{sign}\,u^*[n-1])(\,|u^*[n-1]|+\Delta[n-1]). \quad (26)$$

Заметим, что J[n] зависит от всего вектора $(u[1],\dots,u[n])^T$, причем, чем больше значения каждого $(u[i])^2$, тем выше индекс успешности управления в n-й момент времени. От выбора величины управления u[s] зависят значения всех критериев от J[s] до J[n], то есть значения вектора $J_s=(J[s],\dots,J[n])^T$. Характер зависимости для любого s однотипный.

$$\left(J[s],\,J[s+1\,|\,s],\,\dots,J[n\,|\,s]\right)^T\sim\max_{u[s]}, \quad (27)$$

$$J[n\,|\,s]=\int_{\Omega(y[s],\dots,y[n-1])}J^*[n]\,P\left(y[s],\dots,y[n-1]\,|\,I[s]\right)d\Omega=J^*[n]= =\left(\sqrt{2\pi}\right)^{-1}\left(\sqrt{D[n]}\right)^{-1},$$

$$(D[n])^{-1} = (D_0)^{-1} + \sigma^{-2} \sum_{i=1}^{n-1} \left(u[i]\right)^2.$$

Применим метод уступок для любого s.

$$u^*[s] - \Delta[s] \le u[s] \le u^*[s] + \Delta[s]. \qquad (28)$$

Естественно выбирать величины Δ [s] уступок в зависимости от уровня текущей неопределенности: $\Delta[s] = K\sqrt{D[s]}$.

Дуальное управление s-й момент принимаем в виде:

$$u^\circ[s] = (\text{sign } u^*[s])(|u^*[s]| + \Delta[s]), \qquad (29)$$

где $u^*[s] = f_u(q^\circ[s], m[s]) = q^\circ[s] / m[s]$.

Синтез алгоритмов дуального управления при наличии вероятностных ограничений. Решение будем искать в классе случайных стратегий Г[s]. Доопределим исходную задачу – введем дополнительное требование:

$$E\{u[s]\} = u^*[s],\ s = 1,2, \ldots, n. \qquad (30)$$

$$\int_{\Omega(u[s])} u[s] P\left(u[s \mid I[s]]\right) d\Omega = u^*[s].$$

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.

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

$$J[n \mid s] = \int_{\Omega(y[s],\ldots,y[n-1])} J^*[n] P\left(y[s],\ldots,y[n-1] \mid I[s]\right) d\Omega =$$

$$= J^*[n] = \left(\sqrt{2\pi D[n]}\right)^{-1},$$

$$(D[n])^{-1} = (D_0)^{-1} + \sigma^{-2} \sum_{i=1}^{n-1} \left(u[i]\right)^2.$$

Применим метод уступок для любого s.

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

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

$$\Gamma^\circ[s] = P(u[s] \mid I[s], q^\circ[s]),$$

Г°[s] = δ(u[s] – (u*[s] +Δ [s])) с вероятностью 0,5, (31)

Г°[s] = δ(u[s] – (u*[s] – Δ [s])) с вероятностью 0,5.

Перепишем (31) в более простом виде:

u[s] = u*[s] +Δ [s] с вероятностью 0,5, (32)

u[s] = u*[s] – Δ [s] с вероятностью 0,5.

Такой же вид Г°[s] имеет и для многих других моделей объектов, для которых условие выпуклости соблюдается, например, для модели объекта

$$q[s] = f_q(u[s], \mu) = \mu\,(u[s])^3.$$

В противном случае оптимальная стратегия Г°[s] регулярная, например, для модели $q[s] = f_q\left(u[s],\ \mu\right) = \mu\sqrt{u[s]}$.

Таким образом, существует класс задач дуального управления, в которых смешанные (рандомизированные) стратегии лучше регулярных.

Предложенная формализованная модель обладает рядом существенных достоинств по сравнению с классической теорией дуального правления. Она позволяет решать не решаемые ранее задачи оптимального управления с активным накоплением информации. Впервые в теории дуального управления формализованная процедура синтеза привела к неочевидному неожиданному результату – получению оптимальной рандомизированной стратегии управления. Фактически речь идет о построении основ новой альтернативной теории дуального правления.

В последние годы многие крупные фирмы создают и внедряют системы менеджмента знаний, реализующие три группы процессов: извлечение знаний, накопление знаний, доставку знаний для их практического использования. Менеджмент знаний встраивается непосредственно в системы организационного управления. Можно предположить, что идеи дуального управления окажутся полезными при создании формализованной теории менеджмента знаний в организационном управлении.

Литература и источники в Интернет

1. *Фельдбаум А.А.* Теория дуального управления. I, II // Автоматика и телемеханика. 1960. 21(9), 21(11).
2. *Фельдбаум А.А.* Теория дуального управления. III, IV // Автоматика и телемеханика. 1961. 22(1), 22(2).
3. *Фельдбаум А.А.* Основы теории оптимальных автоматических систем. М.-Л.: Физматгиз, 1963. 552 с.
4. *Живоглядов В.П.* Автоматические системы с накоплением информации. Фрунзе: Илим, 1966.
5. *Живоглядов В.П.* Адаптация в автоматизированных системах управления технологическими процессами. Фрунзе: Илим, 1974.
6. *Filatov N. M., and Unbenhauen H.* Adaptive Dual Control: Theory and Applications. Vol. 302. Lecture Notes in Control and Information Sciences. New York: Springer-Verlag, 2004 .
7. *Morozov C.* Bayesian Dual Control: Review of the Literature. URL: http://www.wavelet3000.org/images/litreview.pdf 24.04.2008.