

Р. ГАБАСОВ
Ф. М. КИРИЛОВА

ОСНОВЫ ДИНАМИЧЕСКОГО ПРОГРАММИРОВАНИЯ

Р. ГАБАСОВ • Ф. М. КИРИЛОВА

ОСНОВЫ
ДИНАМИЧЕСКОГО
ПРОГРАММИРОВАНИЯ

Р. ГАБАСОВ, Ф. М. КИРИЛЛОВА

**ОСНОВЫ
ДИНАМИЧЕСКОГО
ПРОГРАММИРОВАНИЯ**

**Издагелъство БГУ им. В. И. Ленина
Минск 1975**

Дано подробное изложение одного из основных методов теории оптимальных процессов. Принцип оптимальности, функциональные уравнения Беллмана поясняются вначале на простейшем случае многоэтапных детерминированных процессов. Постепенно изучаемые модели усложняются, включая непрерывные, стохастические и игровые задачи. Обсуждается связь динамического программирования с принципом максимума Понтрягина и другими необходимыми условиями оптимальности. Большое внимание уделяется вычислительным аспектам динамического программирования. Изучаются различные современные методы преодоления «проклятия размерности». Отдельно рассматривается задача аналитического конструирования регуляторов, для решения которой динамическое программирование оказалось чрезвычайно эффективным. Изучается связь задачи аналитического конструирования регуляторов с задачей стабилизации движений. Исследуются обобщения динамического программирования (метод Кротова, динамическое программирование высокого порядка, дифференциальное программирование).

Рецензент — доктор технических наук
профессор В. Ф. КРОТОВ



Scan AAW

ОГЛАВЛЕНИЕ

Предисловие	5
Глава I. Дискретные процессы	7
§ 1. Простейшая задача минимизации	—
§ 2. Задача минимизации с дополнительными ограниче- ниями	9
§ 3. Обобщение вида минимизируемой функции	15
§ 4. Задачи оптимизации дискретных процессов	19
§ 5. Инвариантное погружение	24
§ 6. Принцип оптимальности	26
§ 7. Уравнение Беллмана	39
§ 8. Применение динамического программирования к зада- чам терминального управления	47
§ 9. Задачи оптимизации типа Лагранжа	69
§ 10. Минимизация максимального отклонения	80
§ 11. Линейные системы с квадратичными критериями ка- чества	84
§ 12. Приложение к задаче стабилизации	93
§ 13. Применение динамического программирования для решения некоторых комбинаторных задач	96
Глава II. Непрерывные процессы	104
§ 1. Постановка основных задач оптимизации.	—
§ 2. Инвариантное погружение	110
§ 3. Принцип оптимальности	111
§ 4. Дифференциальные уравнения Беллмана	113
§ 5. Проблема обоснования динамического программиро- вания	125
§ 6. Достаточные условия оптимальности	130
§ 7. Минимизация квадратичных функционалов вдоль траек- торий линейных систем	135
§ 8. Проблема стабилизации	144
§ 9. Оптимизация систем с инвариантной нормой	148
Глава III. Вычислительные методы	152
§ 1. Стандартная процедура	—
§ 2. Использование штрафных функций	167
§ 3. Аппроксимация функций Беллмана	172
§ 4. Метод переменного шага	175
§ 5. Использование специальной структуры задачи	181
§ 6. Последовательные приближения в пространстве функций Беллмана	188

§ 7. Последовательные приближения в пространстве управлений	190
§ 8. Использование множеств достижимости	193
§ 9. Искусственное понижение размерности задачи	196
§ 10. Введение искусственных ограничений	200
Глава IV. Стохастические процессы. Адаптация. Игры.	203
§ 1. Инвариантное погружение. Принцип оптимальности.	203
§ 2. Оптимизация дискретных систем	205
§ 3. Оптимальное управление марковской цепью	210
§ 4. Управление в условиях запаздывания информации.	212
§ 5. Оптимизация непрерывных стохастических систем	215
§ 6. Процессы управления с адаптацией.	222
§ 7. Процессы управления в конфликтных ситуациях.	224
Глава V. Связь динамического программирования с другими методами	240
§ 1. Принцип максимума Понтрягина	240
§ 2. Принцип оптимальности Кротова	246
§ 3. Об одном аналоге функции Беллмана.	251
§ 4. Динамическое программирование особых режимов.	254
Литература	261

ПРЕДИСЛОВИЕ

Динамическое программирование — один из наиболее мощных современных методов оптимизации. С задачами принятия рациональных решений, выбора наилучших вариантов, оптимального управления встречаются специалисты разного профиля. Интерес к подобным задачам растет неуклонно.

Среди методов оптимизации динамическое программирование занимает особое положение. Этот метод исключительно привлекателен благодаря простоте и ясности своего основного принципа — принципа оптимальности. Сфера приложения принципа оптимальности чрезвычайно широка, круг задач, к которым он может быть применен, до настоящего времени еще полностью не очерчен. Динамическое программирование с самого начала выступает как средство практического решения задач оптимизации.

Возникновение динамического программирования связывают с именем американского ученого Р. Беллмана, который в начале пятидесятых годов применил к ряду конкретных задач прием, названный впоследствии принципом оптимальности. Основной областью приложения принципа оптимальности являются многошаговые процессы, т. е. процессы, развивающиеся во времени, что дало основание назвать новый метод оптимизации «динамическим». Указанием на динамичность этот метод отличался от линейного и математического программирования, исходная постановка основных задач которых имела статический характер.

Принцип оптимальности Беллмана представляет собой принцип последовательного анализа процесса, протекающего во времени. Идея последовательного анализа широко использовалась и до возникновения динамического программирования. Четко она была сформулирована Вальдом при исследовании задач математической статистики. Для непрерывных систем принцип оптимальности естественно приводит к дифференциальному уравнению Беллмана. Видимо, подобными уравнениями независимо пользовался и Р. Айзекс [1] в своих исследованиях по дифференциальным играм. Однако наиболее полная формулировка идей динамического программирования, их развитие и пропаганда принадлежат Р. Беллману.

Трудно дать четкое определение динамическому программированию. Укажем лишь на три характерные его особенности. Кроме принципа оптимальности, основного приема исследования, большую роль в аппарате динамического программирования играет идея погружения конкретной задачи оптимизации в семейство аналогичных задач. Этот прием восходит к работам В. А. Амбарцумяна по задачам математической физики. Им пользовался С. Чандрасекар, и, наконец, он был сформулирован Р. Беллманом в виде принципа ин-

вариантного погружения и применен к широкому кругу задач (не обязательно оптимизационного типа).

Третьей особенностью динамического программирования, выделяющей его среди других методов оптимизации, является форма конечного результата. Применение принципа оптимальности и принципа погружения в многошаговых, дискретных процессах приводит к рекуррентно-функциональным уравнениям относительно минимального значения критерия качества. Полученные уравнения позволяют последовательно выписать оптимальные управления для исходной задачи. Выигрыш здесь состоит в том, что задача вычисления управления для всего процесса разбивается на ряд более простых задач вычисления управления для отдельных этапов процесса.

В многочисленных работах Р. Беллмана и его последователей динамическое программирование распространено на самые разнообразные задачи (см., например, [2—7, 20, 24, 25]). Как правило, во всех работах получаются рекуррентные соотношения, исследуются некоторые их свойства. Но основное внимание исследователей по динамическому программированию уделяется преодолению «проклятия размерности». Этот термин, введенный Р. Беллманом, связан со следующим обстоятельством. Сведение исходной задачи оптимизации к рекуррентному уравнению требует для решения последнего на каждом этапе запоминать функцию, размерность аргумента которой равна размерности состояния изучаемого процесса. Поскольку речь идет о рабочей информации, для ее хранения в ЭЦВМ используется только оперативная память. Размеры последней даже в современных ЭЦВМ не настолько велики, чтобы преодолеть трудности табулирования функций трех-четырёх аргументов.

При написании книги авторы стремились дать четкое изложение основных принципов динамического программирования, привести классы задач, при решении которых динамическое программирование проявило себя наиболее эффективным средством, изложить современные результаты по развитию и общению динамического программирования.

Основой книги послужили лекции, которые читались авторами на факультете прикладной математики Белгосуниверситета и имели целью познакомить студентов с одним из самых общих современных методов оптимизации. Авторы старались не превращать книгу в некое подобие справочника, уделяя внимание принципиальным вопросам метода. Для единообразия и удобства изложения основные идеи динамического программирования поясняются на примере задач управления. Переложение идей, методов и результатов на задачи, которые имеют другую словесную формулировку, не представляет собой трудности, если для изучаемых задач составлена математическая модель.

Авторы выражают глубокую благодарность И. В. Гайшуну, составившему § 7 гл. IV, Л. Е. Забелло, написавшему пункты по задаче аналитического конструирования регуляторов и стабилизации, и сотрудникам, принявшим участие в оформлении работы.

Принцип оптимальности — основной момент динамического программирования. Прежде чем привести его формулировку, остановимся на решении ряда задач оптимизации без явного использования этого принципа. Вначале ознакомимся с принципом инвариантного погружения. Постепенно усложняя модели, естественно, придем к многошаговым процессам, для которых принципы оптимальности и инвариантного погружения можно сформулировать весьма прозрачно и в достаточно общем виде [2, 3, 5, 7].

§ 1. Простейшая задача минимизации

Начнем с задачи, для которой результат очевиден, но которая удобна для первой иллюстрации основных приемов динамического программирования.

1. Постановка задачи. Метод решения. Пусть R_n — n -мерное евклидово пространство, состоящее из n -мерных векторов $x = \{x_1, \dots, x_n\}$ с координатами x_1, \dots, x_n ; $i = \overline{1, n}$ — символ, означающий, что индекс i принимает все значения: 1, 2, ..., n . Предположим, что в пространстве R_n задано множество G , $G = G_1 \times G_2 \times \dots \times G_n$, представляющее собой прямое произведение множеств G_i , $i = \overline{1, n}$, числовой оси, и скалярная функция $f(x)$, $x = \{x_1, \dots, x_n\} \in G$, специальной структуры

$$f(x) = \sum_{i=1}^n f_i(x_i), \quad (1)$$

где $f_i(x_i)$, $i = \overline{1, n}$ — скалярные функции, определенные на множествах G_i .

Задача. Среди векторов $x \in G$ найти такой x^0 , на котором функция (1) достигает минимума:

$$f(x^0) = \min f(x), \quad x \in G. \quad (2)$$

Решение. Первый этап. Рассмотрим семейство задач

$$\min \sum_{i=1}^s f_i(x_i), \quad x_i \in G_i, \quad i = \overline{1, s}, \quad s = 1, 2, \dots, n. \quad (3)$$

Ясно, что среди задач (3) находится и задача (2), которая получается из (3) при $s=n$. В этом смысле задача (2) вложена в семейство задач (3) или, другими словами, к задаче (2) применен принцип инвариантного погружения.

Введем функцию

$$B_s = \min \sum_{i=1}^s f_i(x_i), \quad x_i \in G_i, \quad i = \overline{1, s}, \quad (4)$$

физический смысл которой следует из самого определения: B_s — минимальное значение функции $\sum_{i=1}^s f_i(x_i)$ при фиксированном значении параметра s .

Назовем B_s функцией Беллмана.

Составление уравнения, которому удовлетворяет функция Беллмана, представляет собой второй этап применения динамического программирования к задаче оптимизации. На этом этапе в дальнейшем для сложных задач будет применяться принцип оптимальности, но для рассматриваемой частной задачи можно воспользоваться очевидным свойством операции «min»:

$$\min_{\alpha \in A, \beta \in B} [f_1(\alpha) + f_2(\beta)] = \min_{\alpha \in A} f_1(\alpha) + \min_{\beta \in B} f_2(\beta). \quad (5)$$

Из определения (4) функции Беллмана и свойства (5) имеем

$$\begin{aligned} B_{s+1} &= \min_{\substack{x_i \in G_i \\ i=\overline{1, s+1}}} \sum_{i=1}^{s+1} f_i(x_i) = \min_{\substack{x_i \in G_i \\ i=\overline{1, s+1}}} [f_{s+1}(x_{s+1}) + \sum_{i=1}^s f_i(x_i)] = \\ &= \min_{\alpha \in G_{s+1}} [f_{s+1}(\alpha) + \min_{\substack{x_i \in G_i \\ i=\overline{1, s}}} \sum_{i=1}^s f_i(x_i)] = \min_{\alpha \in G_{s+1}} f_{s+1}(\alpha) + B_s. \end{aligned}$$

Таким образом, функция Беллмана в задаче (2) удовлетворяет рекуррентно-функциональному уравнению

$$B_{s+1} = B_s + \min_{\alpha \in G_{s+1}} f_{s+1}(\alpha), \quad s = \overline{1, n-1}. \quad (6)$$

Для решения рекуррентных уравнений нужно знать начальные условия. В нашем случае начальное условие получается из (4):

$$B_1 = \min_{\alpha \in G_1} f_1(\alpha). \quad (7)$$

Решая уравнение (6) при $\overline{1, n}$ условии (7), найдем последовательность чисел $B_s, s = \overline{1, n}$. Значение B_n есть минимум в задаче (2). Компоненты x_1^0, \dots, x_n^0 вектора x^0 , доставляющего минимум в задаче (2), находятся, согласно (6), (7), поэтапно:

$$f_s(x_s^0) = \min_{\alpha \in G_s} f_s(\alpha), \quad \alpha \in G_s, \quad s = \overline{1, n}. \quad (8)$$

Уравнение (6) назовем уравнением Беллмана.

2. Анализ результата. Сравнение задачи (2) с задачей (8) показывает, что исходная задача минимизации скалярной функции $f(x)$ от n переменных сведена к n задачам минимизации скалярных функций $f_i(\alpha)$ от одной переменной. В этом основная особенность метода динамического программирования при решении задач оптимизации. Следует обратить внимание на то, что информация о множествах G_i и функциях $f_i(x_i)$ имеет самый общий вид. Множества G_i и функции $f_i(x_i)$ могут быть заданы произвольным способом: аналитически, графически, таблично, алгоритмически и т. п.

З а м е ч а н и е. Не всегда в задачах оптимизации минимизируемая функция имеет вид (1). Однако дополнительными преобразованиями иногда удается исходную функцию привести к виду (1). Например, если дана функция

$$\hat{f}(x) = f_1(x_1) f_2(x_2) \dots f_n(x_n),$$

то после логарифмирования получим функцию вида (1).

§ 2. Задача минимизации с дополнительными ограничениями

Задача предыдущего параграфа усложняется введением дополнительных ограничений, наложенных на совокупность компонент x_1, \dots, x_n . Хотя принципиально схема рассуждения при этом не меняется, результат становится менее тривиальным. Задачи, рассматриваемые в данном

параграфе, имеют наглядную экономическую интерпретацию. Фактически речь будет идти о таком распределении заданных ресурсов между несколькими производствами, при котором себестоимость минимальна.

1. Постановка задачи. Метод решения. Пусть, кроме функции $f(x)$ и множества G , определенных в § 1, задана функция

$$g(x) = \sum_{i=1}^n g_i(x_i), \quad x \in G.$$

Задача. Среди элементов $x \in G$, удовлетворяющих условию

$$g(x) \leq a \quad (-\infty < a < \infty), \quad (1)$$

найти такой x^0 , на котором функция $f(x)$ достигает минимума:

$$f(x^0) = \min f(x), \quad x \in G, \quad g(x) \leq a. \quad (2)$$

Решение. Наряду с поставленной задачей рассмотрим семейство задач

$$\min \sum_{i=1}^s f_i(x_i), \quad x_i \in G_i, \quad \sum_{i=1}^s g_i(x_i) \leq y, \quad (3)$$

где $s = \overline{1, n}$; $y \in Y_s$, Y_s — множество чисел y , при которых множество $\{x_i : x_i \in G_i, \sum_{i=1}^s g_i(x_i) \leq y, i = \overline{1, s}\}$ непусто.

Поскольку при $s = n$, $y = a$ из семейства (3) выделяется исходная задача (2), введение задач (3) реализует принцип инвариантного погружения.

На решениях задач (3) определим функцию Беллмана.

$$B_s(y) = \min \sum_{i=1}^s f_i(x_i), \quad x_i \in G_i, \quad \sum_{i=1}^s g_i(x_i) \leq y. \quad (4)$$

Составим уравнение для функции Беллмана. При этом можно опять обойтись без использования принципа оптимальности, если заметить, что

$$\min_{\alpha \in A, \beta \in B(\alpha)} [f_1(\alpha) + f_2(\alpha, \beta)] = \min_{\alpha \in A} [f_1(\alpha) + \min_{\beta \in B(\alpha)} f_2(\alpha, \beta)]. \quad (5)$$

Из определения функции $B_s(y)$ и свойства (5) следует:

$$\begin{aligned} B_{s+1}(y) &= \min_{x_i \in G_{i, \overline{1, s+1}}} \sum_{i=1}^{s+1} f_i(x_i) = \\ &\quad \sum_{i=1}^{s+1} g_i(x_i) \leq y \\ &= \min_{x_{s+1} \in G_{s+1}} [f_{s+1}(x_{s+1}) + \min_{x_i \in G_{i, \overline{1, s}}} \sum_{i=1}^s f_i(x_i)] \\ &\quad \sum_{i=1}^s g_i(x_i) \leq y - g_{s+1}(x_{s+1}). \end{aligned} \quad (6)$$

Второе слагаемое в квадратных скобках представляет, согласно определению (4), значение функции $B_s(y)$ в точке $y - g_{s+1}(x_{s+1})$. Поэтому из (6) получается рекуррентно-функциональное уравнение Беллмана для задачи (2):

$$B_{s+1}(y) = \min [f_{s+1}(\alpha) + B_s(y - g_{s+1}(\alpha))], \quad \alpha \in G_{s+1}. \quad (7)$$

Начальное условие следует из (4):

$$B_1(y) = \min f_1(\alpha), \quad \alpha \in G_1, \quad g_1(\alpha) \leq y. \quad (8)$$

Из (7), (8) вычисляем последовательность функций $B_s(y)$, $s=1, n$. Значение $B_n(a)$ равно, очевидно, минимуму $f(x^0)$ в задаче (2).

Компоненты x_1^0, \dots, x_n^0 вектора x^0 , доставляющего минимум в задаче (2), вычисляются, согласно (7), (8), следующим образом. После того как найдены функции $B_1(y), \dots, B_{n-1}(y)$, находим число x_n^0 из соотношения

$$\begin{aligned} f_n(x_n^0) + B_{n-1}(a - g_n(x_n^0)) &= \min [f_n(\alpha) + \\ &\quad + B_{n-1}(a - g_n(\alpha))], \quad \alpha \in G_n. \end{aligned} \quad (9)$$

Другими словами, x_n^0 — число, на котором достигается минимум функции, стоящей в квадратных скобках в (9).

Далее находим число x_{n-1}^0 . Из (2), (7) следует, что x_{n-1}^0 удовлетворяет соотношению

$$f_{n-1}(x_{n-1}^0) + B_{n-2}(a - g_n(x_n^0) - g_{n-1}(x_{n-1}^0)) = \\ = \min [f_{n-1}(\alpha) + B_{n-2}(a - g_n(x_n^0) - g_{n-1}(\alpha))], \quad \alpha \in G_{n-1}.$$

Продолжая этот процесс, для вычисления x_k^0 , $n \geq k > 1$, получаем соотношение

$$f_k(x_k^0) + B_{k-1}(a - \sum_{i=n}^k g_i(x_i^0)) = \min [f_k(\alpha) + B_{k-1}(a - \\ - \sum_{i=n}^{k+1} g_i(x_i^0) - g_k(\alpha))], \quad \alpha \in G_k.$$

Значение x_1^0 , согласно (2), (8), удовлетворяет соотношению

$$f_1(x_1^0) = \min f_1(\alpha), \quad \alpha \in G_1, \quad g_1(\alpha) \leq a - \sum_{i=n}^2 g_i(x_i^0).$$

Полученные таким образом числа x_1^0, \dots, x_n^0 суть компоненты вектора x^0 — решения задачи (2).

2. Анализ результата. Хотя формулы (7), (8) внешне мало отличаются от формул (6), (7) из § 1, следует заметить, что между результатами п. 1 из § 2 и § 1 имеется существенное отличие. Функция Беллмана B_s (см. (4), § 1) зависит лишь от дискретного аргумента s , новая функция (4) зависит дополнительно и от непрерывного аргумента y . Согласно (7), при вычислении функции $B_{s+1}(y)$ требуется знать функцию $B_s(y)$ в точках $y = g_{s+1}(\alpha)$, $\alpha \in G_{s+1}$. Поэтому при фактическом решении задачи, когда точное значение $y = g_{s+1}(\alpha)$ становится известным лишь после вычисления $B_n(a)$, для того, чтобы продолжить процесс вычислений, требуется на $s+1$ этапе хранить в памяти всю функцию $B_s(y)$. Так как y — скаляр, это не очень сложная задача, в более сложном случае (см. п. 5) хранение в памяти ЭЦВМ функций $B_s(y)$ представляет уже серьезную проблему.

3. Обратное уравнение Беллмана. Уравнение Беллмана (7) назовем уравнением в прямом времени, поскольку вычисление последовательности функций $B_s(y)$, согласно ему, происходит в направлении возрастания дискретного аргумента. Цель настоящего пункта — получить уравнение Беллмана в обратном времени (обратное уравнение Беллмана).

Задачу (2) вложим в семейство задач

$$\min \sum_{i=s}^n f_i(x_i), \quad x_i \in G_i, \quad \sum_{i=s}^n g_i(x_i) \leq y, \\ s = n, n-1, \dots, 1.$$

Функцию Беллмана определим равенством

$$B^s(y) = \min \sum_{i=s}^n f_i(x_i), \quad x_i \in G_i, \quad \sum_{i=s}^n g_i(x_i) \leq y, \quad i = \overline{s, n}. \quad (10)$$

Нетрудно убедиться, как и в п. 1, что справедливы равенства

$$\begin{aligned} B^{s-1}(y) &= \min_{x_i \in G_i, i=\overline{s-1, n}} \sum_{i=s-1}^n f_i(x_i) = \\ &= \min_{x_{s-1} \in G_{s-1}} \left[f_{s-1}(x_{s-1}) + \min_{x_i \in G_i, i=\overline{s, n}} \sum_{i=s}^n f_i(x_i) \right]. \\ &\quad \sum_{i=s}^n g_i(x_i) \leq y - g_{s-1}(x_{s-1}) \end{aligned}$$

Поэтому, согласно определению (10), функция Беллмана $B^s(y)$ удовлетворяет уравнению

$$B^{s-1}(y) = \min [f_{s-1}(\alpha) + B^s(y - g_{s-1}(\alpha))], \quad \alpha \in G_{s-1}, \quad (11)$$

с начальным условием

$$B^n(y) = \min f_n(\alpha), \quad \alpha \in G_n, \quad g_n(\alpha) \leq y. \quad (12)$$

Вычисление функций $B^s(y)$ по уравнениям (11), (12) идет по мере убывания аргумента s (справа налево). Поэтому уравнение (11) с начальным условием (12) назовем уравнением Беллмана в обратном времени (обратное уравнение Беллмана).

Понятно, что число $B^1(\alpha)$ равно минимальному значению функции $f(x)$ в задаче (2).

Значения x_1^0, \dots, x_n^0 , доставляющие минимум в (2), находятся следующим образом. Функции $B^n(y), \dots, B^2(y)$

считаются известными (они использовались уже для вычисления значения $f(x^0)$).

Число x_1^0 — первая компонента вектора x^0 , доставляющего решение задачи (2), удовлетворяет, согласно (2), (11), соотношению

$$f_1(x_1^0) + B^2(a - g_1(x_1^0)) = \min [f_1^2(\alpha) + B(a - g_1(\alpha))], \\ \alpha \in G_1.$$

Аналогично, числа x_k^0 , $2 \leq k < n$, определяются из соотношений

$$f_k(x_k^0) + B^{k+1}(a - \sum_{i=1}^k g_i(x_i^0)) = \min [f_k(\alpha) + B^{k+1}(a - \\ - \sum_{i=1}^{k-1} g_i(x_i^0) - g_k(\alpha))], \quad \alpha \in G_k, \quad k = 2, \dots, n-1.$$

Наконец, последняя компонента x_n^0 вектора x^0 удовлетворяет соотношению

$$f_n(x_n^0) = \min f_n(\alpha), \quad \alpha \in G_n, \quad g_n(\alpha) \leq a - \sum_{i=1}^{n-1} g_i(x_i^0).$$

4. Учет ограничения типа равенства. Выше рассматривалась задача с ограничениями (1) типа неравенства. Аналогично исследуется случай, когда ограничение имеет вид

$$g(x) = \sum_{i=1}^n g_i(x_i) = a. \quad (13)$$

Для этого во всех выкладках, связанных с функцией $g(x)$, достаточно знак \leq заменить на знак $=$. Результат опять будет иметь вид (7), (8), только в (8) знак \leq нужно заменить на знак $=$.

Если в задаче оптимизации присутствуют ограничения обоих типов (1) и (13), то изменения опять же очевидны.

5. Задача минимизации со многими ограничениями. Задача п. 1 допускает очевидное обобщение, когда вместо одного ограничения (1) учитываются несколько ограничений того же типа:

$$\sum_{i=1}^n g_i^j(x_i) \leq a_j, \quad j = \overline{1, m}. \quad (14)$$

Пусть функцию (2) требуется минимизировать на векторах $x = \{x_1, \dots, x_n\}$, принимающих значения из множества G (см. § 1) и удовлетворяющих ограничениям (14).

Рассуждения п. 1 полностью переносятся на данную задачу. Изменения состоят в следующем. В (3) вместо числа y будет стоять m -вектор y , а под $g_s(\alpha)$ понимается m -мерная функция, составленная из функций $g_s^j(\alpha)$, $j = \overline{1, m}$. В уравнении Беллмана (7) аргумент y функции $B_{s+1}(y)$ теперь уже m -вектор.

Таким образом, в случае (14) размерность непрерывного аргумента функции Беллмана равна числу ограничений. Это обстоятельство затрудняет реальное решение задачи (2), если $m \geq 3$, ибо объем оперативной памяти современных ЭЦВМ не позволяет табулировать функции многих ($m \geq 3$) переменных.

§ 3. Обобщение вида минимизируемой функции

Предельная простота результата § 1 объясняется структурой минимизируемой функции (1) § 1. При ее усложнении результат усложняется, и, если у минимизируемой функции не удастся выявить специальную структуру, с помощью динамического программирования упростить исходную задачу минимизации нельзя.

1. Постановка задачи и ее решение. Пусть минимизируемая функция определена на множестве G (см. § 1) и имеет вид

$$f(x) = f(x_1, \dots, x_n) = \sum_{i=1}^{n-k} f_i(x_i, \dots, x_{i+k}), \quad 0 \leq k \leq n-1. \quad (1)$$

Задача. Найти такой вектор $x^0 \in G$, что

$$f(x^0) = \min f(x), \quad x \in G. \quad (2)$$

Решение. Наряду с задачей (2) рассмотрим семейство задач

$$F_s(x_1, \dots, x_{s-k}, y_1, \dots, y_k) = \sum_{i=1}^{s-k} f_i(x_i, \dots, x_{i+k}) \rightarrow \min \left. \begin{array}{l} x_i \in G_i, \quad i = 1, \overline{s-k}; \quad s = k+1, k+2, \dots, n, \end{array} \right\} \quad (3)$$

где переменные $x_{\overline{s-k+1}}, \dots, x_s$ считаются фиксированными: $x_{s-k+i} = y_i, \quad i = \overline{1, k}$.

Из определения (3) следует, что

$f(x^0) = \min F_n(x_1^0, \dots, x_{n-k}^0, y_1, \dots, y_k), y_i \in G_{n-k+i} \quad i = \overline{1, k}$,
 где $x_1^0(y), \dots, x_{n-k}^0(y)$ — решение задачи (3) при $s=n$.
 Можно считать поэтому, что задача (2) погружена в семейство задач (3).

Для задач (3) введем функцию Беллмана

$$B_s(y_1, \dots, y_k) = \min F_s(x_1, \dots, x_{s-k}, y_1, \dots, y_k), x_i \in G_i, \\ i = \overline{1, s-k}. \quad (4)$$

Переходя к составлению уравнения относительно функции B_s , запишем по определению

$$B_{s+1}(y_1, \dots, y_k) = \min_{x_i \in G_i, i=\overline{1, s+1-k}} \sum_{i=1}^{s+1-k} f_i(x_i, \dots, \\ x_{i+k}) \left| \begin{array}{l} x_{s-k+i+1} = y_i \\ i=\overline{1, k} \end{array} \right. = \min_{x_{s+1-k} \in G_{s+1-k}} [f_{s+1-k}(x_{s+1-k}, y_1, \dots, y_k) + \\ + \min_{x_i \in G_i, i=\overline{1, s-k}} \sum_{i=1}^{s-k} f_i(x_i, \dots, x_{i+k}) \left| \begin{array}{l} x_{s-k+i} = y_{i-1} \\ i=\overline{2, k} \end{array} \right. \right]. \quad (5)$$

Поскольку в квадратных скобках во втором слагаемом по переменной x_{s-k+i} не берется операция \min , эту переменную вместе с x_{s-k+i} , $i=\overline{2, k}$, можно считать фиксированной. Тогда по определению (4) второе слагаемое в (5) равно $B_s(x_{s-k+1}, y_1, \dots, y_{k-1})$.

Таким образом, из (5) получается уравнение Беллмана в задаче (2):

$$B_{s+1}(y_1, \dots, y_k) = \min [f_{s+1-k}(\alpha, y_1, \dots, y_k) + \\ + B_s(\alpha, y_1, \dots, y_{k-1})], \alpha \in G_{s+1-k} \quad (6)$$

с начальным условием

$$B_{k+1}(y_1, \dots, y_k) = \min f_1(\alpha, y_1, \dots, y_k), \alpha \in G_1. \quad (7)$$

Решение исходной задачи (2) получается из (6), (7) по следующей схеме: находится последовательность функций $B_{k+1}(y_1, \dots, y_k), B_{k+2}(y_1, \dots, y_k), \dots, B_n(y_1, \dots, y_k)$.

Величина

$$\min B_n(y_1, \dots, y_k), y_i \in G_{n-k+i}, i = \overline{1, k}, \quad (8)$$

равна $f(x^0)$.

Поскольку функции B_{k+1}, \dots, B_n можно считать известными, из (8) прежде всего можно найти числа $x_{n-k+1}^0, \dots, x_n^0$, ибо

$$B_n(x_{n-k+1}^0, \dots, x_n^0) = \min B_n(y_1, \dots, y_k), y_i \in G_{n-k+i}, i = \overline{1, k}.$$

Далее из уравнения (6) при $s=n-1$ можно найти число x_{n-k}^0 :

$$\begin{aligned} f_{n-k}(x_{n-k}^0, x_{n-k+1}^0, \dots, x_n^0) + B_{n-1}(x_{n-k}^0, x_{n-k+1}^0, \dots, x_{n-1}^0) = \\ = \min [f_{n-k}(\alpha, x_{n-k+1}^0, \dots, x_n^0) + B_{n-1}(\alpha, x_{n-k+1}^0, \dots, x_{n-1}^0)], \\ \alpha \in G_{n-k}. \end{aligned}$$

Аналогично для вычисления x_{n-k-l}^0 , $1 \leq l < n-k-1$, получаем соотношение

$$\begin{aligned} f_{n-k-l}(x_{n-k-l}^0, x_{n-k-l+1}^0, \dots, x_{n-l}^0) + B_{n-l-1}(x_{n-k-l}^0, \\ x_{n-k-l+1}^0, \dots, x_{n-l-1}^0) = \min [f_{n-k-l}(\alpha, x_{n-k-l+1}^0, \dots, x_{n-l}^0) + \\ + B_{n-l-1}(\alpha, x_{n-k-l+1}^0, \dots, x_{n-l-1}^0)], \alpha \in G_{n-k-l}. \end{aligned}$$

Наконец, первая компонента x_1^0 вектора x^0 , доставляющего решение задачи (2), находится из соотношения

$$f_1(x_1^0, x_2^0, \dots, x_{k+1}^0) = \min f_1(\alpha, x_2^0, \dots, x_{k+1}^0), \alpha \in G_1.$$

2. Анализ результата. Исходная задача (2) является задачей минимизации скалярной функции по n переменным. С помощью динамического программирования она сведена к $(n-k)$ задачам (6), (7) минимизации по одной переменной и к задаче (8) минимизации по k переменным. Таким образом, параметр k , характеризующий сложность структуры минимизируемой функции (1), определяет «степень разложения» (декомпозиции) задачи минимизации. При $k=0$ получается результат § 1. При

$k=n-1$, когда о структуре минимизируемой функции нет дополнительной информации, нужно решить задачи (7), (8), которые, очевидно, эквивалентны исходной задаче (2).

3. Обратное уравнение Беллмана. Получим уравнение Беллмана в обратном времени для задачи (2).

Наряду с задачей (2) рассмотрим задачи

$$F^s(y_1, \dots, y_k, x_{k+s}, \dots, x_n) = \sum_{i=s}^{n-k} f_i(x_i, \dots, x_{i+k}) \rightarrow \min, \left. \begin{array}{l} x_i \in G_i, \quad i = \overline{k+s, n}, \quad s = \overline{1, n-k}, \end{array} \right\} \quad (9)$$

где переменные x_s, \dots, x_{s+k-1} считаются фиксированными: $x_{s+i-1} = y_i, \quad i = \overline{1, k}$.

Из (9) следует, что

$$f(x^0) = \min F^1(y_1, \dots, y_k, x_{k+1}^0(y_1, \dots, y_k), \dots, x_n^0(y_1, \dots, y_k)), \\ y_i \in G_i, \quad i = \overline{1, k},$$

где $x_{k+1}^0(y), \dots, x_n^0(y)$ — решение задачи (9) при $s=1$.

Для задачи (9) введем в рассмотрение функцию Беллмана

$$B^s(y_1, \dots, y_k) = \min_{i = \overline{k+s, n}} F^s(y_1, \dots, y_k, x_{k+s}, \dots, x_n), \quad x_i \in G_i,$$

Составим рекуррентное соотношение для функций B^s . По определению,

$$\begin{aligned} B^{s-1}(y_1, \dots, y_k) &= \min_{\substack{x_i \in G_i, \\ i = \overline{k+s-1, n}}} \sum_{i=s-1}^{n-k} f_i(x_i, \dots, x_{i+k}) \left| \begin{array}{l} x_{s+i-2} = y_i \\ i = \overline{1, k} \end{array} \right. = \\ &= \min_{x_{k+s-1} \in G_{k+s-1}} [f_{s-1}(y_1, \dots, y_k, x_{k+s-1}) + \\ &+ \min_{\substack{x_i \in G_i, \\ i = \overline{k+s, n}}} \sum_{i=s}^{n-k} f_i(x_i, \dots, x_{i+k}) \left| \begin{array}{l} x_{s+i-1} = y_{i+1}, \\ i = \overline{1, k-1} \end{array} \right. \quad]. \quad (10) \end{aligned}$$

Так как второе слагаемое в квадратных скобках равняется $B^s(y_2, \dots, y_k, x_{k+s-1})$, из (10) непосредственно получаем обратное уравнение Беллмана

$$B^{s-1}(y_1, \dots, y_k) = \min [f_{s-1}(y_1, \dots, y_k, \alpha) + B^s(y_2, \dots, y_k, \alpha)], \alpha \in G_{k+s-1} \quad (11)$$

с начальным условием

$$B^{n-k}(y_1, \dots, y_k) = \min f_{n-k}(y_1, \dots, y_k, \alpha), \alpha \in G_n. \quad (12)$$

Из (11) и (12) находим последовательность функций

$$B^{n-k}(y_1, \dots, y_k), B^{n-k-1}(y_1, \dots, y_k), \dots, B^1(y_1, \dots, y_k).$$

Компоненты x_1^0, \dots, x_k^0 вектора $x^0 = \{x_1^0, \dots, x_n^0\}$, являющегося решением исходной задачи, находятся из соотношения

$$B^1(x_1^0, \dots, x_k^0) = \min B^1(y_1, \dots, y_k), y_i \in G_i, i = \overline{1, k}.$$

Далее из уравнения (11) при $s=2$ определяем x_{k+1}^0 :

$$\begin{aligned} f_1(x_1^0, \dots, x_k^0, x_{k+1}^0) + B^2(x_2^0, \dots, x_k^0, x_{k+1}^0) = \\ = \min f_1(x_1^0, \dots, x_k^0, \alpha) + B^2(x_2^0, \dots, x_k^0, \alpha), \alpha \in G_{k+1}. \end{aligned}$$

Для вычисления остальных компонент x_{k+l}^0 ($1 < l < n-k$) получаем соотношение

$$\begin{aligned} f_l(x_l^0, \dots, x_{k+l-1}^0, x_{k+l}^0) + B^{l+1}(x_{l+1}^0, \dots, x_{k+l-1}^0, x_{k+l}^0) = \\ = \min f_l(x_l^0, \dots, x_{k+l-1}^0, \alpha) + B^{l+1}(x_{l+1}^0, \dots, x_{k+l-1}^0, \alpha), \\ \alpha \in G_{k+l}. \end{aligned}$$

Наконец, последняя компонента x_n^0 вектора x^0 находится следующим образом:

$$f_{n-k}(x_{n-k}^0, \dots, x_{n-1}^0, x_n^0) = \min f_{n-k}(x_{n-k}^0, \dots, x_{n-1}^0, \alpha), \alpha \in G_n.$$

§ 4. Задачи оптимизации дискретных процессов

Задачи, рассмотренные в § 1—3, по своей постановке не являются динамическими, они не связаны непосредственно с процессами. Только специальная структура функции $f(x)$ и $g(x)$ позволила с помощью принципа инвариантного погружения придать им «динамический

характер» и включить их в некоторое подобие процессов. С данного параграфа начинаем изучать оптимизацию процессов. Процессы будут участвовать в постановке задач с самого начала. Задачи, рассмотренные в § 1—3, суть частные случаи новых задач (см. п. 3).

Пусть имеется некоторый объект (система), состояние которого меняется со временем в дискретные моменты $t=t_0, t_0+1, t_0+2, \dots$, образуя некоторый процесс. Будем считать, что в момент t состояние объекта полностью описывается n -вектором $x(t) = \{x_1(t), \dots, x_n(t)\}$. Совокупность значений $x(t), t \geq t_0$, характеризует процесс изменения объекта со временем. Допустим, что процесс $x(t), t \geq t_0$, определяется как собственной динамикой объекта, так и внешними воздействиями $u(t) = \{u_1(t), \dots, u_r(t)\}$, а общая зависимость выражается уравнением

$$x(t+1) = f(x(t), u(t), t), \quad t \geq t_0, \quad x(t_0) = x_0, \quad (1)$$

где t_0 — начальный момент времени; x_0 — начальное состояние процесса (объекта, системы); $f(x, u, t)$ — n -мерная функция, определенная в области значений своих аргументов.

Функцию $u(t), t \geq t_0$, назовем управлением. Выбирая то или иное управление $u(t_0), u(t_0+1), \dots, u(t), \dots$, получаем в силу (1) соответствующее течение процесса $x(t_0), x(t_0+1), \dots, x(t+1), \dots$

Управление $u(t)$ называется допустимым, если в каждый момент времени $t, t \geq t_0$, оно принимает значение из заданного множества $U(t)$ r -мерного пространства:

$$u(t) \in U(t), \quad t \geq t_0. \quad (2)$$

Процесс $x(t), t \geq t_0$, соответствующий допустимому управлению, называется допустимым процессом. Иногда, имея в виду процесс $x(t), t \geq t_0$, порожденный уравнением (1), говорят «движение системы», «траектория системы».

Рассмотрим процесс $x(t), t \in T = [t_0, t_1]$, длительностью $t_1 - t_0$, соответствующий управлению $u(t), t \in [t_0, t_1 - 1]$. Качество процесса $x(t), t \geq t_0$, оценим величиной

$$I(u) = \sum_{t=t_0}^{t_1-1} f_0(x(t), u(t), t), \quad (3)$$

которая называется критерием качества.

Простейшая задача оптимизации многошагового процесса (1) состоит в минимизации величины (3) по всем допустимым управлениям, т. е. в построении такого допустимого управления $u^0(t)$, $t \in [t_0, t_1-1]$, называемого оптимальным, при котором

$$I(u^0) = \min I(u),$$

где минимум справа берется по всем допустимым управлениям $u = \{u(t) \in U(t), t \in [t_0, t_1-1]\}$.

Траектория $x^0(t)$, $t \geq t_0$, соответствующая управлению $u^0(t)$, называется оптимальной*.

Решение $u^0(t)$, $t \geq t_0$, задачи зависит от начальных условий: момента t_0 и состояния x_0 . Поэтому в случае необходимости будем говорить, что управление $u^0(t)$, $t \geq t_0$, оптимально относительно t_0, x_0 .

Сформулированная задача оптимизации допускает обобщения в нескольких направлениях.

а) Если длительность процесса $t_1 - t_0$ не фиксирована, имеем задачу оптимизации с нефиксированной продолжительностью процесса.

б) Пусть на концы траектории $x(t)$ наложены ограничения. Если левый конец $x(t_0)$ траектории не закреплен, а принимает значения из заданного множества G_0 n -мерного пространства, $x(t_0) \in G_0$, получаем задачу оптимизации с подвижным левым концом. Если условие накладывается на правый конец $x(t_1)$ траектории, $x(t_1) \in G_1$, G_1 — множество n -мерного пространства, получаем задачу с подвижным правым концом.

Если множества G_0, G_1 являются точками, имеем задачу с закрепленными концами.

Если $G_1 = R_n$, т. е. G_1 совпадает с n -мерным пространством R_n и на правый конец $x(t_1)$ траектории по существу не накладывается условий, задачу оптимизации называют задачей со свободным правым концом.

В задачах с ограничениями на состояния процессов под допустимыми управлениями будем понимать управления, удовлетворяющие ограничениям (2), и такие, что траектории, ими порожденные, удовлетворяют присутствующим в задаче ограничениям.

в) Ограничения могут накладываться на течение процесса: $x(t) \in G(t)$, $t \in [t_0+1, t_1-1]$, где $G(t)$ — задан-

* Всюду в работе предполагается, что рассматриваемые задачи оптимизации имеют решения.

ные множества n -мерного пространства. В этом случае задача называется задачей с ф а з о в ы м и о г р а н и ч е н и я м и.

г) Если ограничения накладываются на совокупность управлений и траекторий

$$\{u(t), x(t)\} \in H(t) \quad (4)$$

($H(t)$ — заданное множество $n+r$ -мерного пространства), то говорят о задаче со с м е ш а н н ы м и о г р а н и ч е н и я м и. Допустимым считается такое управление $u(t)$, $t \geq t_0$, которое в совокупности с соответствующей ему траекторией $x(t)$, $t \geq t_0$, системы (1) удовлетворяет ограничению (4).

д) Критерий (3) является дискретным аналогом критерия задачи Лагранжа из вариационного исчисления, поэтому задачи с подобным критерием называют задачами т и п а Л а г р а н ж а. Если критерий качества имеет вид

$$I(u) = \varphi(x(t_1)), \quad (5)$$

где $\varphi(x)$ — скалярная функция, задача оптимизации называется задачей т и п а М а й е р а (задачей т е р м и н а л ь н о г о у п р а в л е н и я, или задачей управления к о н е ч н ы м с о с т о я н и е м).

Если критерий качества имеет вид

$$I(u) = \varphi(x(t_1)) + \sum_{t=t_0}^{t_1-1} f_0(x(t), u(t), t),$$

то задача оптимизации называется задачей т и п а Б о л ь ц а.

Если в (3) $f_0(x, u, t) \equiv 1$ и число t_1 не фиксировано, то задача оптимизации с критерием $I = t_1 - t_0$ называется задачей б ы с т р о д е й с т в и я (понятно, что в такой задаче концы траектории не могут быть свободными).

Вводя дополнительные переменные, можно во многих случаях перейти от задачи с одним типом критерия качества к задаче с другим типом. Например, задача типа Лагранжа сводится к задаче типа Майера, если ввести

дополнительную переменную $x_0(t) = \sum_{\tau=t_0}^{t-1} f_0(x(\tau), u(\tau), \tau)$,

которая, очевидно, удовлетворяет уравнению

$$x_0(t+1) = x_0(t) + f_0(x(t), u(t), t) \quad (6)$$

того же типа, что и уравнение (1). Добавляя уравнение (6) к системе (1), вместо критерия (3) получим частный случай критерия (5):

$$I(u) = x_0(t_1).$$

Можно, наоборот, от задачи типа Майера перейти к задаче типа Лагранжа. Введем функцию

$$f_0(x(t), u(t), t) = \varphi(f(x(t), u(t), t)) - \varphi(x(t)).$$

Ясно, что

$$\sum_{t=t_0}^{t_1-1} f_0(x(t), u(t), t) = \varphi(x(t_1)) - \varphi(x(t_0)).$$

Член $\varphi(x(t_0))$ не зависит от управления, поэтому можно считать, что задача типа Майера сведена к задаче типа Лагранжа.

Аналогичным путем можно показать «взаимозаменяемость» задач типа Майера и Больца.

Ограничения на управление и траекторию не всегда задаются в виде (2), (4). Они могут иметь вид

$$\sum_{t=t_0}^{t_1-1} g(x(t), u(t), t) \leq a, (=a). \quad (7)$$

Ограничение (7) введением дополнительной координаты сводится к добавлению в (1) уравнения

$$x_{n+1}(t+1) = x_{n+1}(t) + g(x(t), u(t), t)$$

и к наложению ограничения на правый конец траектории: $x_{n+1}(t_1) \leq a, (=a)$.

Задачи, рассмотренные в § 1—3, можно сформулировать в терминах многошаговых процессов. Для определенности рассмотрим задачу из § 2. Напомним, что требовалось минимизировать функцию

$$\sum_{i=1}^n f_i(x_i) \quad (8)$$

при ограничениях

$$\sum_{i=1}^n g_i(x_i) \leq a, \quad x_i \in G_i, \quad i = \overline{1, n}. \quad (9)$$

Для формулировки этой задачи в терминах процессов управления введем скалярный процесс $x(t)$, описываемый уравнением

$$x(t+1) = f(x(t), u(t), t), \quad x(0) = 0, \quad (10)$$

где функция $f(x, u, t)$ задана равенством $f(x, u, t) = g_t(u) + x$. Далее положим $U(t) = G_t$, $f_0(x, u, t) = f_t(u)$.

В новых обозначениях ограничения (9) принимают вид

$$x(n+1) \leq a, \quad u(t) \in G_t, \quad t \in [1, n]; \quad (11)$$

минимизация функции (8) эквивалентна минимизации критерия качества $I(u) = \sum_{t=1}^n f_0(x, u, t)$ вдоль траекторий уравнения (10), стесненных ограничениями (11). Таким образом, задача § 2 свелась к частному случаю задачи типа Лагранжа с фиксированной продолжительностью процесса, с закрепленным левым и подвижным правым концами траектории.

§ 5. Инвариантное погружение

При решении частных задач в § 1—3 принцип инвариантного погружения использовался каждый раз по-иному. После введения в § 4 общей модели дискретных процессов, с помощью которой можно поставить обширный класс задач оптимизации, сформулируем принцип инвариантного погружения применительно к дискретным процессам управления.

1. Прямое погружение многошагового процесса. Основная черта прямого погружения состоит в том, что вводится семейство процессов $\{x(t)\}_{\tau, y}$, которые исходят из точки y в момент τ . Пусть процесс $x(t)$, $t \in [t_0, t_1]$, описывается уравнением

$$x(t+1) = f(x(t), u(t), t), \quad x(t_0) = x_0. \quad (1)$$

Если выбрать управление $u(t)$, $t \in [t_0, t_1-1]$, то при фиксированном начальном условии x_0 уравнение (1) задает единственный процесс $x(t)$, $t \in [t_0, t_1]$. Наряду с процессом $x(t)$, $t \in [t_0, t_1]$, рассмотрим семейство $\{x(t)\}_{\tau, y}$ процессов, описываемых уравнением

$$x(t+1) = f(x(t), u(t), t), \quad x(\tau) = y, \quad \tau \leq t \leq t_1 - 1, \quad (2)$$

где τ — произвольное число $t_0 \leq \tau \leq t_1 - 1$; y — произвольный n -вектор.

Введенное семейство процессов $\{x(t)\}_{\tau, y}$ зависит от параметров: скаляра τ и вектора y (рис. 1). При частном

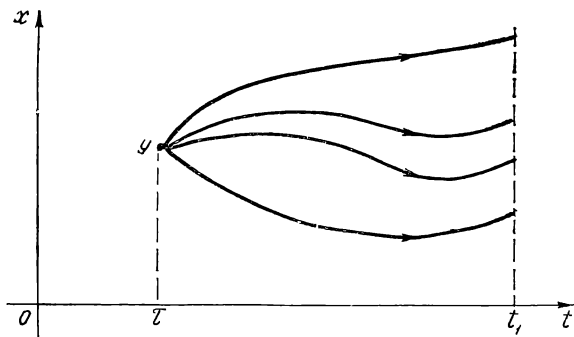


Рис. 1

значении этих параметров $y = x_0$, $\tau = t_0$ из семейства (2) выделяется процесс (1).

Переход от уравнения (1) к уравнению (2) (или, что то же самое, от $x(t)$, $t \in [t_0, t_1]$ к $\{x(t)\}_{\tau, y}$) назовем

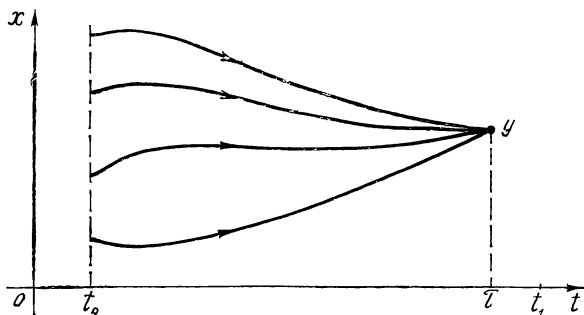


Рис. 2

принципом инвариантного погружения многошагового процесса $x(t)$, $t \in [t_0, t_1]$.

2. Обратное погружение многошагового процесса.

В некоторых случаях оказывается полезным и обратное погружение многошагового процесса. В этом случае рассматривается семейство процессов $\{x(t)\}_{\tau, y}$, которые в момент τ приходят в точку y (рис. 2).

Пусть задан многошаговый процесс

$$x(t+1) = f(x(t), u(t), t), \quad x(t_1) = x_1, \quad t_0 \leq t \leq t_1 - 1. \quad (3)$$

Математически процедура обратного погружения процесса (3) состоит во введении семейства процессов $\{x(t)\}_{\tau, y}$, $t_0 \leq t \leq \tau$, зависящего от скаляра τ , $t_0 + 1 \leq \tau \leq t_1$, и n -вектора y . Элементы семейства $\{x(t)\}_{\tau, y}$ описываются уравнениями

$$x(t+1) = f(x(t), u(t), t), \quad x(\tau) = y, \quad t_0 \leq t \leq \tau - 1. \quad (4)$$

Нетрудно видеть, что при $\tau = t_1$, $y = x_1$ из семейства $\{x(t)\}_{\tau, y}$ выделяется процесс (3).

Двумя описанными способами не исчерпываются все виды погружения многошаговых процессов. Например, при прямом погружении процесса (1) иногда удобнее ввести параметр τ иначе, чем в п. 1. Вместо семейства (2) введем семейство $\{x(t)\}_{\tau, y}$, описываемое уравнением

$$x(t+1) = f(x(t), u(t), t), \quad x(t_1 - \tau) = y, \\ t \in [t_1 - \tau, t_1 - 1]. \quad (5)$$

При $\tau = t_1 - t_0$, $y = x_0$ из семейства (5) выделяется процесс (1), поэтому переход от (1) к (5) также реализует принцип инвариантного погружения.

Аналогичную модификацию допускает и метод обратного погружения, описанный в п. 2.

§ 6. Принцип оптимальности

После рассмотрения в § 1—3 конкретных реализаций принципа оптимальности Беллмана и введения в § 4 общей модели дискретного процесса можно сформулировать и доказать принцип оптимальности для ряда важных критериев качества. Оказывается (см. п. 4), что, несмотря на внешнюю простоту, вопрос требует внимательного отношения.

1. Принцип оптимальности как необходимое условие. Рассмотрим дискретный процесс, описываемый уравнением

$$x(t+1) = f(x(t), u(t), t), \quad x(t_0) = x_0, \quad t \in [t_0, t_1 - 1]. \quad (1)$$

Основная черта задания процесса $x(t)$ уравнением (1) состоит в том, что для любого вектора y , числа τ ,

$\tau \in [t_0, t_1 - 1]$, выбор начального условия $x(\tau) = y$ и функции $u(t)$, $t \in [\tau, t_1 - 1]$, однозначно определяет течение процесса $x(t)$ на промежутке $[\tau, t_1]$. В силу этого значение $x(t)$ процесса называется состоянием процесса в момент t .

Другое важное свойство процесса, заданного уравнением (1), заключается в том, что при любом $\tau \in [t_0 + 1, t_1 - 1]$ состояния $x(t)$, $t \in [t_0, \tau]$, не зависят от выбора управления на отрезке $[\tau, t_1 - 1]$. В последующих рассуждениях именно эти два свойства процессов оказываются решающими при обосновании принципа оптимальности.

Поскольку при любом τ , $\tau \geq t_0$, каждый рассматриваемый процесс $x(t)$, $t \geq \tau$, однозначно определяется состоянием $x(\tau)$ и управлением $u(t)$, $t \geq \tau$, при оценке этого процесса значение критерия качества будет функцией этих параметров:

$$I = I(x(\tau), u[\tau, t_1 - 1]), \quad (2)$$

где символом $u[\tau, t_1 - 1]$ обозначено множество $\{u(\tau), u(\tau + 1), \dots, u(t_1 - 1)\}$.

Пусть управление $u(t)$ процессом $x(t)$ определено на отрезке $[t_0, t_1 - 1]$. Разделим промежуток $[t_0, t_1]$ на два периода: первый $[t_0, \tau - 1]$ и второй $[\tau, t_1]$, где τ — любое натуральное число из $[t_0 + 1, t_1 - 1]$. Управление $u[t_0, \tau - 1]$ в первом периоде приводит процесс в состояние $x(\tau)$, которое является начальным для управления $u[\tau, t_1 - 1]$ во втором периоде. Предположим, что критерий качества (2) определен на любом заключительном подпроцессе исходного процесса. Тогда качество управления в первом периоде зависит от начального состояния $x(t_0)$ и управления $u[t_0, \tau - 1]$, а во втором — оценивается величиной $I(x(\tau), u[\tau, t_1 - 1])$. Для всего процесса значение критерия качества равно $I(x(t_0), u[t_0, t_1 - 1])$.

Допустим, что между вычисленными значениями критерия качества существует связь

$$I(x(t_0), u[t_0, t_1 - 1]) = \gamma(x(t_0), u[t_0, \tau - 1], I(x(\tau), u[\tau, t_1 - 1])), \quad (3)$$

где функция $\gamma(\alpha, \beta, \gamma)$ такова, что при любых α, β, γ , $-\infty < \alpha, \beta < \infty$,

$$\nu(\alpha, \beta, \bar{\gamma}) < \nu(\alpha, \beta, \gamma), \quad (4)$$

как только $\bar{\gamma} < \gamma$.

Будем считать, что на управление $u[t_0, t_1-1]$ и процесс $x[t_0, t_1]$ наложены ограничения, при которых существуют допустимые управления и соответствующие им допустимые траектории.

Пусть рассматривается задача минимизации критерия (2) на траекториях системы (1), удовлетворяющих наложенным ограничениям, и пусть эта задача имеет решение. Тогда при $t_1 - t_0 \geq 2$ справедливо следующее утверждение.

Принцип оптимальности Беллмана. Для оптимальности управления необходимо, чтобы при любой длительности периодов управление во втором периоде было оптимально относительно состояния, в котором оказался процесс после управления в первом периоде.

Доказательство. Допустим, что утверждение неверно. Тогда найдется такое натуральное число τ , $\tau \in [t_0+1, t_1-1]$, что оптимальное управление $u^0[t_0, t_1-1]$ во втором периоде $[\tau, t_1-1]$ не оптимально относительно состояния $x^0(\tau)$; это значит, что существует функция $u[\tau, t_1-1]$, составляющая вместе с функцией $u^0[t_0, \tau-1]$ допустимое управление $u[t_0, t_1-1]$, определенное на отрезке $[t_0, t_1-1]$, и такая, что

$$I(x^0(\tau), u[\tau, t_1-1]) < I(x^0(\tau), u^0[\tau, t_1-1]). \quad (5)$$

В силу (3) — (5) нетрудно подсчитать, что

$$\begin{aligned} I(x_0, u[t_0, t_1-1]) &= \nu(x_0, u^0[t_0, \tau-1], I(x^0(\tau), \\ &u[\tau, t_1-1])) < \nu(x_0, u^0[t_0, \tau-1], I(x^0(\tau), \\ &u^0[\tau, t_1-1])) = I(x_0, u^0[t_0, t_1-1]). \end{aligned}$$

Сравнивая крайние члены полученного неравенства, убеждаемся, что оно противоречит оптимальности управления $u_0[t_0, t_1-1]$. Утверждение доказано.

З а м е ч а н и е. Свойство (4) критерия качества использовано в доказательстве лишь при тех значениях α, β , которые определяются заданным начальным состоянием x_0 и оптимальным управлением $u^0[t_0, \tau-1]$.

2. Частные виды критериев качества. При доказательстве принципа оптимальности существенно используется условие (4), которому должен удовлетворять критерий

качества. Распространенные критерии качества, описанные в § 4, удовлетворяют условию (4).

Рассмотрим, например, критерий качества в задаче типа Лагранжа:

$$I = \sum_{t=t_0}^{t_1-1} f_0(x(t), u(t), t).$$

Очевидно, что при любом $\tau \in [t_0+1, t_1-1]$ справедлива формула

$$\begin{aligned} I(x_0, u | t_0, t_1-1] &= \sum_{t=t_0}^{t_1-1} f_0(x(t), u(t), t) = \\ &= \sum_{t=t_0}^{\tau-1} f_0(x(t), u(t), t) + \sum_{t=\tau}^{t_1-1} f_0(x(t), u(t), t) = \\ &= \sum_{t=t_0}^{\tau-1} f_0(x(t), u(t), t) + I(x(\tau), u[\tau, t_1-1]), \end{aligned}$$

т. е. функция $v(\alpha, \beta, \gamma)$ из (3) имеет вид

$$v(\alpha, \beta, \gamma) = v_1(\alpha, \beta) + \gamma,$$

для которого условие (4) выполняется.

Пусть задан критерий задачи типа Майера

$$I = \varphi(x(t_1)).$$

Нетрудно сообразить, что при любом $\tau, \tau \in [t_0+1, t_1-1]$, справедливы равенства

$$I(x_0, u[t_0, t_1-1]) = \varphi(x(t_1)) = I(x(\tau), u[\tau, t_1-1]),$$

из которых следует, что $v(\alpha, \beta, \gamma) = \gamma$.

Условие (4) для этой функции выполняется.

Рассмотрим еще критерий задачи Больца.

$$I = \varphi(x(t_1)) + \sum_{t=t_0}^{t_1-1} f_0(x(t), u(t), t).$$

Имеем

$$\begin{aligned} I(x_0, u[t_0, t_1-1]) &= \varphi(x(t_1)) + \sum_{t=t_0}^{t_1-1} f_0(x(t), u(t), t) = \\ &= \varphi(x(t_1)) + \sum_{t=t_0}^{\tau-1} f_0(x(t), u(t), t) + \sum_{t=\tau}^{t_1-1} f_0(x(t), u(t), t) = \\ &= \sum_{t=t_0}^{\tau-1} f_0(x(t), u(t), t) + I(x(\tau), u[\tau, t_1-1]). \end{aligned}$$

В этом случае функция $v(\alpha, \beta, \gamma)$ имеет вид $v(\alpha, \beta, \gamma) = v_1(\alpha, \beta) + \gamma$. Ясно, что и здесь условие (4) выполняется.

3. Другие модели дискретных процессов. В качестве модели для дискретных процессов при доказательстве принципа оптимальности было взято уравнение

$$x(t+1) = f(x(t), u(t), t). \quad (6)$$

Это уравнение не является наиболее общим из уравнений, встречающихся в приложениях. Например, вполне возможны процессы, описываемые уравнением

$$x(t+1) = f(x(t), x(t-h), u(t), t), \quad (7)$$

где h — некоторое натуральное число, $h > 0$.

В этом случае нельзя сказать, что значение $x(t)$ процесса будет его состоянием в момент t , ибо, зная лишь $x(t)$ и управление $u[t, t_1-1]$, невозможно однозначно продолжить процесс вправо от точки t . Поэтому теряет смысл и определение (2) критерия качества. Неверным будет и принцип оптимальности, если при его формулировке под состоянием понимать вектор $x(t)$. Истинным состоянием в момент t , $t > h$, процесса, заданного уравнением (7), является совокупность векторов

$$X(t) = \{x(t), x(t-1), \dots, x(t-h)\}. \quad (8)$$

Если под состоянием понимать вектор (8), тогда принцип оптимальности сохранится и для процессов, заданных уравнением (7).

Уравнение с запаздывающим аргументом (7) легко сводится к уравнению типа (6), если ввести вектор (8) и составить для него уравнение. Действительно нетрудно проверить, что $n(h+1)$ -мерный процесс $X(t)$ удовлетворяет уравнению

$$X(t+1) = F(X(t), u(t), t),$$

где $n(h+1)$ -мерная функция $F(X, u, t)$ имеет вид

$$F(X, u, t) = \begin{Bmatrix} f(x^1, x^{h+1}, u, t) \\ x^1 \\ \vdots \\ x^h \end{Bmatrix}, \quad X = \begin{Bmatrix} x^1 \\ \vdots \\ x^{h+1} \end{Bmatrix}.$$

К уравнению (6) дополнительными преобразованиями можно свести много других уравнений. Однако эта процедура осуществима не всегда. Например, серьезные трудности возникают при исследовании уравнения

$$x(t+1) = f(x(t), x(t-h), x(t+h), u(t), t), h \geq 1.$$

4. Контрпримеры. Хотя приведенные в п. 1 условия для критерия качества являются лишь достаточными, нарушение их может привести к тому, что в конкретной задаче принцип оптимальности Беллмана не будет выполняться.

Примерами критериев качества, не удовлетворяющих условиям п. 1, являются следующие:

$$I = \sum_{t=t_0}^{t_1-k} f_0(x(t), x(t+1), \dots, x(t+k), t), k \geq 1; \quad (9)$$

$$I = \max_{t_0+1 \leq t \leq t_1} \varphi(x(t), t). \quad (10)$$

В случае (9) при $t_0 \leq \tau \leq t_1 - k$ имеем:

$$\begin{aligned} I(x_0, u[t_0, t_1-1]) &= \sum_{t=t_0}^{t_1-k} f_0(x(t), \dots, x(t+k), t) = \\ &= \sum_{t=t_0}^{\tau} f_0(x(t), \dots, x(t+k), t) + \sum_{t=\tau+1}^{t_1-k} f_0(x(t), \dots, x(t+k), t) = \\ &= \sum_{t=t_0}^{\tau} f_0(x(t), \dots, x(t+k), t) + I(x(\tau), u[\tau, t_1-1]), \end{aligned}$$

т. е.

$$\begin{aligned} I(x_0, u[t_0, t_1-1]) &= v_1(x_0, u[t_0, \tau-1], \\ &u[\tau, \tau+k-1], I(x(\tau), u[\tau, t_1-1])). \end{aligned}$$

Предположение (3) не выполняется.

Если критерий качества имеет вид (10), то

$$\begin{aligned} I(x_0, u[t_0, t_1-1]) &= \max_{t_0+1 \leq t \leq t_1} \varphi(x(t), t) = \\ &= \max_{t_0+1 \leq t \leq \tau} [\max_{t_0+1 \leq t \leq \tau} \varphi(x(t), t), \max_{\tau+1 \leq t \leq t_1} \varphi(x(t), t)] = \\ &= \max [I(x_0, u[t_0, \tau-1]), I(x(\tau), u[\tau, t_1-1])] = \\ &= v(x_0, u[t_0, \tau-1], I(x(\tau), u[\tau, t_1-1])). \end{aligned}$$

В этом случае выполняется предложение (3), но функция $v(\alpha, \beta, \gamma)$ не удовлетворяет условию (4).

Покажем теперь, что оба условия (3), (4) существенны для выполнения принципа оптимальности.

Пример 1. $x(t+1) = u(t)$, $x(0) = 2$, $t \in [0, 2]$, $|u(t)| \leq 1$,
 $t \in [0; 1]$, $-1 \leq u(2) \leq 0$, $I(u) = \sum_{t=0}^1 x(t) x(t+1) x(t+2)$.

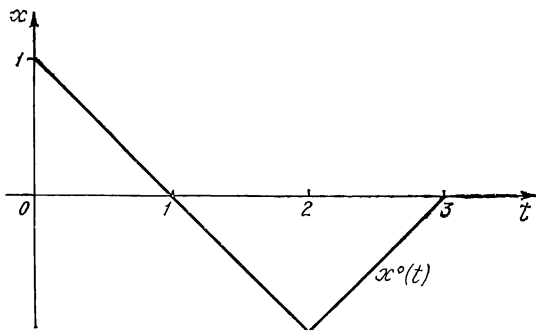


Рис. 3

Подсчитаем оптимальное управление в этой задаче:

$$I(u) = x(0)x(1)x(2) + x(1)x(2)x(3) = x(0)u(0)u(1) + \\ + u(0)u(1)u(2) = u(0)u(1)[x(0) + u(2)].$$

Управление $u^0(0) = 1$, $u^0(1) = -1$, $u^0(2) = 0$, очевидно, будет оптимальным.

Рассмотрим два периода процесса: $[0, 1]$, $[1, 2]$. В конце первого периода $x^0(1) = 1$. Для процесса, начинающегося из этой точки, минимум критерия

$$I = \sum_{t=1}^1 x(t)x(t+1)x(t+2) = x(1)x(2)x(3) = u(1)u(2)$$

достигается на управлении $u(1) = 1$, $u(2) = -1$, которое не совпадает с оптимальным управлением $u^0(t)$ во втором периоде.

Пример 2. $x(t+1) = x(t) + u(t)$, $x(0) = 1$, $|u(t)| \leq 1$,
 $t \in [0, 2]$, $I(u) = \max_{0 \leq t \leq 3} |x(t)|$.

Нетрудно видеть, что функция $u^0(0) = -1$, $u^0(1) = -1$, $u^0(2) = 1$ является оптимальным управлением. Для подпроцесса, начинающегося в момент $t=1$ из точки $x^0(1) = 0$, управление $u^0(t)$, $t=1, 2$, не является оптимальным (рис. 3). Для точки $x^0(1) = 0$ оптимальным будет управление $u(t) \equiv 0$, $t=1, 2$.

Как и в п. 3, дополнительные преобразования позволяют иногда «спасти» принцип оптимальности. Например,

в случае критерия (9) достаточно вместо n -вектора $x(t)$ ввести $(k+1)$ n -мерный вектор

$$X(t) = \{x(t), x(t+1), \dots, x(t+k)\}.$$

Тогда уравнение (1), записанное для $x(t)$, перейдет в следующее уравнение для $X'(t) = \{x^1(t), \dots, x^{k+1}(t)\}$

$$X(t+1) = F(X(t), v(t), t), \quad (11)$$

$$\text{где } F(X, v, t) = \begin{Bmatrix} x^2 \\ x^3 \\ \vdots \\ x^{k+1} \\ f(x^{k+1}, v, t+k) \end{Bmatrix}, \quad v(t) = u(t+k).$$

Для новой системы критерий качества (9) имеет вид

$$I = \sum_{t=t_0}^{t_1-k} F_0(X(t), t), \quad (12)$$

где $F_0(X, t) = f_0(x^1, x^2, \dots, x^{k+1}, t)$.

Ясно, что для системы (11) и критерия качества (12) принцип оптимальности сохраняется. Но дополнительные преобразования не всегда приводят к цели. В частности, не очевиден путь, ведущий от критерия (10) к новой формулировке, в которой принцип оптимальности справедлив. К счастью, задачи оптимизации с критерием (10) все же можно решить методами динамического программирования (см. § 10). Однако при этом принцип оптимальности выступает не как необходимое, а как достаточное условие оптимальности.

5. Обсуждение принципа оптимальности. При выводе рекуррентно-функциональных уравнений Беллмана для задач, рассмотренных в § 1—3, мы специально избегали использования принципа оптимальности. В каждом из исследованных случаев соответствующие уравнения Беллмана получались путем эквивалентных и весьма простых преобразований, не требующих дополнительных фактов. В сущности эти преобразования и содержали принцип оптимальности или, другими словами, именно они явля-

лись конкретной реализацией принципа оптимальности. В § 1—3 не вставал вопрос о справедливости принципа оптимальности, ибо его обоснованием служили строгие математические преобразования. Однако принцип оптимальности является общим утверждением, и его обоснование представляет интерес, независимо от конкретных задач. Как показано в п. 4, даже для процессов, описываемых уравнением (1), принцип оптимальности как необходимое условие оптимальности не всегда верен. Естественно задаться вопросом об отношении принципа оптимальности к достаточным условиям оптимальности. Ответ на этот вопрос играет принципиальную роль при использовании динамического программирования для решения конкретных задач. Действительно, при применении принципа оптимальности как необходимого условия получаем соотношения, процедуры, ведущие к управлениям не обязательно оптимальным. Поскольку эффективность динамического программирования ожидается именно в методах построения алгоритмов, ведущих к оптимальным управлениям, в основу динамического программирования должен быть положен принцип оптимальности в форме достаточного условия. С этой точки зрения вопрос освещен в литературе слабо. Это объясняется тем, что в рассмотренных методом динамического программирования задачах оптимизации принцип оптимальности лишь провозглашался, но по существу не использовался в общепринятой форме. Каждый раз при выводе уравнений Беллмана не буквально выписывались соотношения, следующие из принципа оптимальности, а осуществлялись преобразования, ведущие к уравнению Беллмана, и законность их следовала не из принципа оптимальности, а из математических (как правило, очень простых) фактов. Преобразования, применяемые для получения уравнений Беллмана, в большинстве случаев являются, как и в § 1—3, эквивалентными в том смысле, что соотношения, к которым они ведут, необходимы и достаточны для оптимальности управлений.

6. Принцип оптимальности как достаточное условие. Рассмотрим достаточные аспекты принципа оптимальности. Будем считать, что дискретный процесс описывается уравнением (1). Предположим далее, что критерий качества удовлетворяет условию (3) и ослабленному условию (4)

$$\nu(\alpha, \beta, \gamma) \leq \nu(\alpha, \beta, \bar{\gamma}) \quad (13)$$

при $\gamma \leq \bar{\gamma}$.

Говорят, что управление $u^* [t_0, \tau - 1]$, составляющее с $u^0[x^*(\tau), \tau, t_1 - 1]$ допустимое управление, оптимально в первом периоде, если выполняется неравенство

$$\begin{aligned} \nu(x_0, u^* [t_0, \tau - 1], I(x^*(\tau), u^0[x^*(\tau), \tau, t_1 - 1])) &\leq \\ &\leq \nu(x_0, u [t_0, \tau - 1], I(x(\tau), u^0[x(\tau), \tau, t_1 - 1])) \end{aligned} \quad (14)$$

для функций $u[t_0, \tau - 1]$, составляющих с $u^0[x(\tau), \tau, t_1 - 1]$ допустимое управление, где $u^0[x, \tau, t_1 - 1]$ — управление, оптимальное относительно x .

Теорема. Если управление во втором периоде оптимально относительно состояния, возникшего в результате управления в первом периоде, а управление в первом периоде оптимально, то управление на всем промежутке $[t_0, t_1 - 1]$ будет оптимальным.

Доказательство. Пусть управление $u[t_0, t_1 - 1]$, составленное из управлений $u[t_0, \tau - 1]$ и $u[\tau, t_1 - 1]$, допустимо, а управление $u[\tau, t_1 - 1]$ оптимально относительно $x(\tau)$, и управление $u[t_0, \tau - 1]$ оптимально в первом периоде. Докажем, что $u[t_0, t_1 - 1]$ — оптимальное управление.

Пусть $\tilde{u}[t_0, t_1 - 1]$ — любое другое допустимое управление, составленное из $\tilde{u}[t_0, \tau - 1]$ и $\tilde{u}[\tau, t_1 - 1]$.

Используя свойство (3) и условия (13), (14), нетрудно получить неравенства

$$\begin{aligned} I(x_0, u[t_0, t_1 - 1]) &= \nu(x_0, u[t_0, \tau - 1], I(x(\tau), \\ u[x(\tau), \tau, t_1 - 1])) &\leq \nu(x_0, \tilde{u}[t_0, \tau - 1], I(\tilde{x}(\tau), \\ u[\tilde{x}(\tau), \tau, t_1 - 1])) &\leq \nu(x_0, \tilde{u}[t_0, \tau - 1], I(\tilde{x}(\tau), \\ \tilde{u}[\tilde{x}(\tau), \tau, t_1 - 1])) &= I(x_0, \tilde{u}[t_0, t_1 - 1]), \end{aligned}$$

из которых следует, что проверяемое управление $u[t_0, t_1 - 1]$ оптимально.

Если проанализировать способы получения уравнений Беллмана § 1—3, нетрудно убедиться, что всюду принцип оптимальности используется именно в сформулированном виде, а не как необходимое условие оптимальности.

Доказательство достаточности принципа оптимальности проведено при некоторых предположениях относительно структуры критерия качества. Именно подобная структура (3) позволяет получить характерные уравнения Беллмана.

В общем случае критерий качества обладает свойством

$$I(x_0, u[t_0, t_1 - 1]) = \lambda(x_0, u[t_0, \tau - 1], u[\tau, t_1 - 1])$$

и для него можно сформулировать достаточное условие оптимальности типа: если управления $u[t_0, \tau - 1]$ и $u[\tau, t_1 - 1]$ минимизируют функцию λ , то управление $u[t_0, t_1 - 1]$ оптимально. Но такое утверждение тривиально и бесполезно для исследования задач оптимизации. По сравнению с ним принцип оптимальности, сформулированный в данном пункте, позволяет эффективно исследовать широкий круг задач оптимизации.

З а м е ч а н и я. 1. Очевидно, управление, оптимальное в первом периоде, совпадает в этом периоде с оптимальным управлением на всем промежутке. Для критерия качества произвольного типа оптимальность в первом периоде можно понимать в таком смысле.

Если предположения относительно структуры критерия качества не выполнены, то принцип оптимальности в достаточном аспекте, вообще говоря, не верен. Так, в примере 1 управление $u(0)=1$, $u(1)=1$, $u(2)=-1$ удовлетворяет условиям принципа оптимальности, но не является оптимальным.

2. Если условие (4) заменить на условие (13), то принцип оптимальности Беллмана как необходимое условие оптимальности справедлив в следующей ослабленной форме.

Пусть $u^0[t_0, t_1 - 1]$ — оптимальное управление, тогда для любого τ найдется такое оптимальное управление $\bar{u}^0[t_0, t_1 - 1] = \{u^0[t_0, \tau - 1], \bar{u}^0[\tau, t_1 - 1]\}$, что управление $u^0[\tau, t_1 - 1]$ оптимально относительно состояния, в котором при $t = \tau$ оказался процесс под воздействием управления $u^0[t_0, \tau - 1]$.

7. Обратный принцип оптимальности. Принцип оптимальности, рассмотренный в пп. 1—6, связан с прямым погружением дискретного процесса в семейство аналогичных процессов. Поэтому естественно называть его **прямым принципом оптимальности** в отличие от **обратного принципа оптимальности**, связанного с обратным погружением (см. § 5).

Рассмотрим дискретный процесс $x(t)$, $t \in [t_0, t_1]$, заданный уравнением

$$x(t+1) = f(x(t), u(t), t), \quad x(t_1) = x^1. \quad (15)$$

Будем считать, что при выбранном управлении $u(t)$, $t \in$

$[t_0, t_1-1]$, рекуррентное уравнение (15) порождает единственную траекторию $x(t)$, $t \in [t_0, t_1]$.

Критерием качества дискретного процесса, порожденного вектором x^1 и управлением $u[t_0, t_1-1]$, назовем величину

$$I = I(u[t_0, t_1-1], x^1), \quad (16)$$

где функция $I(u, x^1)$ определена на конечных состояниях процесса и управлениях, приводящих траектории системы (15) в эти состояния.

Разобьем промежуток времени $[t_0, t_1-1]$ на два периода $[t_0, \tau-1]$ и $[\tau, t_1-1]$. Предположим, что при любом τ , $t_0+1 \leq \tau \leq t_1-1$, критерий качества (16) определен на подпроцессе $x(t)$, $t \in [t_0, \tau]$, протекающем в первом периоде и

$$I(u[t_0, t_1-1], x^1) = \mu(I(u[t_0, \tau-1], x(\tau)), \\ u[\tau, t_1-1], x^1). \quad (17)$$

Относительно функции $\mu(\alpha, \beta, \gamma)$ предположим, что при любых β, γ выполняется неравенство

$$\mu(\alpha, \beta, \gamma) < \mu(\bar{\alpha}, \beta, \gamma), \quad (18)$$

как только $\alpha < \bar{\alpha}$.

Будем говорить, что допустимое управление $u[t_0, \tau-1]$ оптимально относительно конечного состояния x , если выполняется неравенство

$$I(u[t_0, \tau-1], x) \leq I(\tilde{u}[t_0, \tau-1], x) \quad (19)$$

для любого допустимого управления $\tilde{u}[t_0, \tau-1]$, причем $x(\tau) = \tilde{x}(\tau) = x$. При сделанных предположениях справедливо следующее утверждение.

Обратный принцип оптимальности. Для оптимальности управления необходимо, чтобы при любом разбиении промежутка $[t_0, t_1-1]$ управление в первом периоде было оптимально относительно начального состояния второго периода.

Доказательство. Допустим противное: при некотором τ , $t_0 \leq \tau \leq t_1-1$, для оптимального управления $u[t_0, \tau-1]$ неравенство (19) не имеет места, т. е. существу-

ет управление $\tilde{u}[t_0, \tau-1]$, допустимое в совокупности с $u[\tau, t_1-1]$, для которого

$$I(u[t_0, \tau-1], x(\tau)) > I(\tilde{u}[t_0, \tau-1], x(\tau)). \quad (20)$$

Приведем элементарные вычисления с учетом (17), (18), (20):

$$\begin{aligned} I(u[t_0, t_1-1], x^1) &= \mu(I(u[t_0, \tau-1], x(\tau)), u[\tau, \\ t_1-1], x^1) &> \mu(I(\tilde{u}[t_0, \tau-1], x(\tau)), u[\tau, t_1-1], x^1) = \\ &= I(\tilde{u}[t_0, t_1-1], x^1). \end{aligned} \quad (21)$$

Здесь $\tilde{u}[t_0, t_1-1]$ — допустимое управление, составленное из управлений $\tilde{u}[t_0, \tau-1], u[\tau, t_1-1]$. Сравнивая крайние члены в (21), видим, что получилось противоречие с предположением об оптимальности управления $u[t_0, t_1-1]$. Утверждение доказано.

Примерами конкретных критериев качества, удовлетворяющих условиям (17), (18), являются

$$\begin{aligned} I &= \varphi(x(t_0)); \\ I &= \sum_{t=t_0}^{t_1-1} f_0(x(t), u(t), t); \\ I &= \varphi(x(t_0)) + \sum_{t=t_0}^{t_1-1} f_0(x(t), u(t), t). \end{aligned}$$

По аналогии с п. 4 нетрудно показать, что условия (17), (18) существенны для справедливости обратного принципа оптимальности.

8. Обратный принцип оптимальности как достаточное условие. Сформулированный обратный принцип оптимальности — необходимое условие оптимальности. Достаточный аспект этого принципа вполне аналогичен прямому его варианту.

Пусть дискретный процесс описывается уравнением (15). Предположим, что критерий качества удовлетворяет условию (17) и ослабленному условию (18).

Будем говорить, что управление $u^*[\tau, t_1-1]$, составляющее с $u^0[t_0, \tau-1, x^*(\tau)]$ допустимое управление, опти-

мально во втором периоде, если выполняется неравенство

$$\mu(I(u^0[t_0, \tau-1, x^*(\tau)], x^*(\tau)), u^*[\tau, t_1-1], x^1) \leq \\ \leq \mu(I(u^0[t_0, \tau-1, x(\tau)], x(\tau), u[\tau, t_1-1], x^1) \quad (22)$$

для всех функций $u[\tau, t_1-1]$, составляющих с $u^0[t_0, \tau-1, x(\tau)]$ допустимое управление. Здесь $u^0[t_0, \tau-1, x]$ — управление, оптимальное относительно конечного состояния x .

Теорема. Если управление в первом периоде оптимально относительно начального состояния второго периода, а управление во втором периоде оптимально, то управление на всем промежутке $[t_0, t_1-1]$ будет оптимальным.

Доказательство. Пусть управление $u[t_0, t_1-1]$, составленное из управлений $u[t_0, \tau-1]$ и $u[\tau, t_1-1]$, допустимо, управление $u[t_0, \tau-1]$ оптимально относительно конечного состояния $x(\tau)$, а управление $u[\tau, t_1-1]$ оптимально во втором периоде. Докажем, что $u[t_0, t_1-1]$ — оптимальное управление.

Пусть $\tilde{u}[t_0, t_1-1]$ — произвольное допустимое управление, составленное из $\tilde{u}[t_0, \tau-1]$, $\tilde{u}[\tau, t_1-1]$. Тогда из свойства (17), условия (22) и ослабленного условия (18) следует

$$I(u[t_0, t_1-1], x^1) = \mu(I(u^0[t_0, \tau-1, x(\tau)], x(\tau)), \\ u[\tau, t_1-1], x^1) \leq \mu(I(u^0[t_0, \tau-1, \tilde{x}(\tau)], \tilde{x}(\tau)), \tilde{u}[\tau, t_1-1], \\ x^1) \leq \mu(I(\tilde{u}[t_0, \tau-1], \tilde{x}(\tau)), \tilde{u}[\tau, t_1-1], x^1) = \\ = I(\tilde{u}[t_0, t_1-1], x^1).$$

Таким образом, управление $u[t_0, t_1-1]$ оптимально.

§ 7. Уравнение Беллмана

Различные варианты принципа оптимальности, доказанные в предыдущем параграфе, позволяют получать соответствующие уравнения Беллмана. В теории динамического программирования эти уравнения играют разнообразную роль. Помимо того, что они получаются из принципа оптимальности, уравнения Беллмана интерес-

ны как самостоятельный объект исследования, поскольку с помощью их формулируются весьма общие достаточные условия оптимальности.

1. Функция Беллмана. Пусть исследуется некоторая задача оптимизации дискретного процесса, заданного уравнением

$$x(t+1) = f(x(t), u(t), t), \quad t \in [t_0, t_1 - 1]. \quad (1)$$

Не будем выписывать всех ограничений, наложенных на управления и траектории, поскольку понятие допустимых управлений и допустимых траекторий аккумулирует всю соответствующую информацию. Дискретный процесс (1) вложим прямо в семейство процессов

$$x(t+1) = f(x(t), u(t), t), \quad x(\tau) = y, \quad t \in [\tau, t_1 - 1], \\ \tau \in [t_0, t_1 - 1]. \quad (2)$$

Обозначим через $F(\tau)$ множество таких n -векторов y , для которых при данном τ «сужение» исходной задачи на систему (2) имеет решение. «Сужение» исходной задачи состоит в следующем.

Изменяется промежуток времени, в течение которого исследуется процесс. Левый конец траектории в новой задаче считается закрепленным. Других изменений не делается: хотя в исходной задаче левый конец траектории, возможно, свободен или подвижен, а в системе (2) он уже закреплен, это не противоречит тому, что процесс (1) погружен в семейство (2), ибо, рассматривая семейство (2) при $\tau = t_0$ и выбирая вектор y из соответствующего множества ограничений, получаем исходную задачу. При выбранных значениях τ , $y \in F(\tau)$ и допустимом управлении $u[\tau, t_1 - 1]$ значение критерия качества полностью определяется

$$I = I(\tau, y, u[\tau, t_1 - 1]). \quad (3)$$

Минимальное значение функционала (3) на допустимых управлениях называется функцией Беллмана $B(\tau, y)$. Таким образом,

$$B(\tau, y) = \min I(\tau, y, u[\tau, t_1 - 1]), \quad (4)$$

где минимум берется по всем допустимым управлениям.

Функция Беллмана при сделанных предположениях определена при всех $\tau \in [t_0, t_1 - 1]$, $y \in F(\tau)$.

2. Обратное уравнение Беллмана. Предположим, что критерий качества (3) удовлетворяет условиям (3), (4) § 6. Тогда, как доказано в § 6, справедлив прямой принцип оптимальности, согласно которому при любом θ , $\tau \leq \theta \leq t_1 - 1$, для оптимального управления $u[\tau, t_1 - 1]$ и оптимальной траектории $x[\tau, t_1]$ выполняется соотношение

$$I(\theta, x(\theta), u[\theta, t_1 - 1]) = \min I(\theta, x(\theta), \tilde{u}[\theta, t_1 - 1]), \quad (5)$$

где минимум берется по всем управлениям $\tilde{u}[\theta, t_1 - 1]$, составляющим с $u[\tau, \theta - 1]$ допустимое управление.

Согласно определению (4), равенство (5) можно записать в форме

$$I(\theta, x(\theta), u[\theta, t_1 - 1]) = B(\theta, x(\theta)).$$

По предположению, выполняется условие (3) § 6. Поэтому

$$I(\tau, y, u[\tau, t_1 - 1]) = v(\tau, y, u[\tau, \theta - 1]),$$

$$I(\theta, x(\theta), u[\theta, t_1 - 1]).$$

Используя (4), можем получить равенство

$$B(\tau, y) = v(\tau, y, u[\tau, \theta - 1], B(\theta, x(\theta))), \quad (6)$$

которое справедливо в предположении, что $u[\tau, \theta - 1]$ — часть оптимального управления.

Пусть $\bar{u}[\tau, \theta - 1]$ — некоторое управление; $\bar{x}[\tau, \theta]$ — траектория системы (2), соответствующая управлению $\bar{u}[\tau, \theta - 1]$. Обозначим через $\bar{u}[\theta, t_1 - 1]$ управление, которое с $\bar{u}[\tau, \theta - 1]$ составляет допустимое управление, такое, что

$$\begin{aligned} I(\theta, \bar{x}(\theta), \bar{u}[\theta, t_1 - 1]) &= \min I(\theta, \bar{x}(\theta), \tilde{u}[\theta, t_1 - 1]) = \\ &= B(\theta, \bar{x}(\theta)), \end{aligned}$$

где минимум берется по всем $\tilde{u}[\theta, t_1 - 1]$, составляющим с $\bar{u}[\tau, \theta - 1]$ допустимое управление.

Подсчитаем значение критерия качества для управления $\bar{u}[\tau, t_1 - 1]$, составленного из $\bar{u}[\tau, \theta - 1]$ и $\bar{u}[\theta, t_1 - 1]$,

$$I(\tau, y, \bar{u}[\tau, t_1 - 1]) = v(\tau, y, \bar{u}[\tau, \theta - 1], I(\theta, \bar{x}(\theta), \bar{u}[\theta, t_1 - 1])) = v(\tau, y, \bar{u}[\tau, \theta - 1], B(\theta, \bar{x}(\theta))). \quad (7)$$

Поскольку, по определению,

$$I(\tau, y, u[\tau, t_1 - 1]) \leq I(\tau, y, \bar{u}[\tau, t_1 - 1]) \quad (8)$$

и значение правой части этого неравенства в силу (7) полностью определяется выбором $\bar{u}[\tau, \theta - 1]$ и равно значению левой части (8) при $\bar{u}[\tau, \theta - 1] = u[\tau, \theta - 1]$, то из (6) — (8) следует уравнение

$$B(\tau, y) = \min_{\bar{u}[\tau, \theta - 1]} v(\tau, y, \bar{u}[\tau, \theta - 1], B(\theta, \bar{x}(\theta))), \quad (9)$$

где минимум берется по всем $\bar{u}[\tau, \theta - 1]$, составляющим с $\bar{u}[\theta, t_1 - 1]$ допустимое управление.

Полученное уравнение (9) для функции Беллмана назовем **о б р а т н ы м** уравнением Беллмана.

З а м е ч а н и е. Уравнение (9) получено с помощью прямого погружения и прямого принципа оптимальности. Так как вычисление функции Беллмана по первому аргументу идет справа налево, назовем уравнение (9) **обратным** уравнением Беллмана.

Если положить $\theta = \tau + 1$, из (2) имеем

$$\bar{x}(\theta) = f(y, \bar{u}(\tau), \tau)$$

и уравнение Беллмана (9) принимает вид

$$B(\tau, y) = \min v(\tau, y, u, B(\tau + 1, f(y, u, \tau))), \quad (10)$$

где минимизация проводится по всем r -векторам u , при которых

$$f(y, u, \tau) \in F(\tau + 1).$$

Уравнение Беллмана в форме (10) наиболее простое, оно удобно для вычислений, его частные случаи, соответствующие определенному выбору функции v , имеют наибольшее распространение в динамическом программировании дискретных процессов.

Уравнение (10), как и (9), — рекуррентно-функциональное, и для его решения нужно задать начальные условия. Эти условия получаются в конкретных задачах из краевых условий, наложенных на правый конец траектории системы (1). В § 8 этот вопрос будет рассмотрен детально.

3. Другой вывод обратного уравнения Беллмана. Обратное уравнение Беллмана (10) в п. 2 получено как уравнение, которому удовлетворяет оптимальное управление. Возникает вопрос: если допустимое управление получено из уравнения Беллмана, будет ли оно оптимальным? Ответ на него дается в п. 4. А здесь для вывода уравнения Беллмана воспользуемся достаточным аспектом принципа оптимальности, рассмотренным в § 6.

Пусть допустимое управление $u[\tau, t_1-1]$ построено по рецептам принципа оптимальности как достаточного условия.

Тогда выполняются соотношения

$$I(\theta, x(\theta), u[\theta, t_1-1]) = \min I(\theta, x(\theta), \bar{u}[\theta, t_1-1]), \quad (11)$$

где минимум берется по всем $\bar{u}[\theta, t_1-1]$, составляющим с $u[\tau, \theta-1]$ допустимое управление:

$$\begin{aligned} & \nu(\tau, y, u[\tau, \theta-1], I(\theta, x(\theta), u[\theta, t_1-1])) = \\ & = \min \nu(\tau, y, \bar{u}[\tau, \theta-1], I(\theta, \bar{x}(\theta), \bar{u}[\theta, t_1-1])). \end{aligned} \quad (12)$$

Здесь 1) управление $\bar{u}[\theta, t_1-1]$ в совокупности с $\bar{u}[\tau, \theta-1]$ допустимо и оптимально относительно $\bar{x}(\theta)$; 2) минимум берется по всем $\bar{u}[\tau, \theta-1]$, которые в совокупности с соответствующими $\bar{u}[\theta, t_1-1]$ допустимы.

Согласно достаточному аспекту принципа оптимальности, допустимое управление, удовлетворяющее условиям (11), (12), оптимально. Но в силу определения (4) функции Беллмана соотношение (12) представляет собой уравнение Беллмана (9). Таким образом, доказано, что любое допустимое управление, построенное согласно достаточному принципу оптимальности, удовлетворяет уравнению Беллмана.

4. Уравнение Беллмана и достаточные условия оптимальности. Рассуждения, которые в п. 3 привели к доказательству достаточных условий оптимальности, основаны на принципе оптимальности в его достаточном аспек-

те и обоснованы постольку, поскольку обоснован последний результат.

Оказывается, что построение достаточных условий оптимальности можно базировать непосредственно на уравнении Беллмана, не связывая этот вопрос с принципом оптимальности. При этом подходе принципу оптимальности отводится роль эвристического соображения, ведущего к уравнению Беллмана.

Итак, пусть имеется уравнение Беллмана (10), в котором функция $v(\alpha, \beta, \gamma, \delta)$ удовлетворяет условию: при каждом наборе α, β, γ выполняется неравенство

$$v(\alpha, \beta, \gamma, \bar{\delta}) \leq v(\alpha, \beta, \gamma, \delta), \quad (13)$$

как только $\bar{\delta} \leq \delta$.

Будем считать, что функция $B(t_1, y)$, $y \in F(t_1)$ определена по условиям задачи независимо от уравнения (10). Из рекуррентного уравнения (10), меняя параметр τ от t_1-1 до t_0 , получим последовательность функций $B(t_1-1, y)$, $B(t_1-2, y), \dots, B(t_0, y)$ и последовательность управлений $u(t_1-1, y)$, $u(t_1-2, y), \dots, u(t_0, y)$.

Предположим для определенности, что левый конец траектории $x(t_0)$ фиксирован $x(t_0) = x_0$.

Подсчитаем $B(t_0, x_0)$ и $u^0(t)$ по следующему правилу:

$$u^0(t_0) = u(t_0, x_0), \quad u^0(t_0 + 1) = u(t_0 + 1, f(x_0, u^0(t_0, t_0))), \dots,$$

$$u^0(t) = u^0(t, x^0(t)), \dots, \quad u^0(t_1 - 1) = u^0(t_1 - 1, x^0(t_1 - 1)).$$

где $x^0(t)$ — состояние в момент t процесса, порожденно-го в силу (1) управлением $u^0(t_0), \dots, u^0(t-1)$ и начальным условием $x^0(t_0) = x_0$.

Для скалярной функции $v(\alpha, \beta, \gamma, \delta)$ введем понятие суперпозиции вдоль функций $\alpha(t)$, $\beta(t)$, $\gamma(t)$. По определению, суперпозиция

$$\int_{t=t_0}^{t_1} v(\alpha(t), \beta(t), \gamma(t), v(t+1)) = v(\alpha(t_0), \beta(t_0), \gamma(t_0),$$

$$v(\alpha(t_0 + 1), \beta(t_0 + 1), \gamma(t_0 + 1), v(\alpha(t_0 + 2), \dots, v(\alpha(t_1), \beta(t_1), \gamma(t_1), v(t_1 + 1)) \dots))),$$

где $v(t_1 + 1)$ — некоторое фиксированное число.

Докажем, что $u^0(t)$, $t \in [t_0, t_1-1]$, — оптимальное управление, доставляющее минимум критерию

$$I = \int_{t=t_0}^{t_1-1} v(t, x(t), u(t), v(t+1)), v(t_1) = B(t_1, x(t_1)), \quad (14)$$

и этот минимум равен $B(t_0, x_0)$.

Действительно, из уравнения Беллмана (10) вдоль управления $u^0(t)$ получаем:

$$\begin{aligned} B(t_0, x_0) &= v(t_0, x_0, u^0(t_0), B(t_0+1, x^0(t_0+1))) = \\ &= v(t_0, x^0(t_0), u^0(t_0), v(t_0+1, x^0(t_0+1), u^0(t_0+1), \\ &\quad B(t_0+2, x^0(t_0+2)))) = \dots = v(t_0, x^0(t_0), \quad (15) \\ &\quad u^0(t_0), v(t_0+1, \dots, v(t_1-1, x^0(t_1-1), u^0(t_1-1), \\ &\quad B(t_1, x^0(t_1)))) = \int_{t=t_0}^{t_1-1} v(t, x^0(t), u^0(t), v(t+1)). \end{aligned}$$

С другой стороны, для любого допустимого управления

$$u(t_0), u(t_0+1), \dots, u(t_1-1)$$

и соответствующей ему траектории $x(t)$, $t \in [t_0, t_1]$ из (10) получаем неравенства:

$$\begin{aligned} B(t_0, x_0) &\leq v(t_0, x(t_0), u(t_0), B(t_0+1, x(t_0+1))); \\ B(t_0+1, x(t_0+1)) &\leq v(t_0+1, x(t_0+1), u(t_0+1), \\ &\quad B(t_0+2, x(t_0+1))); \\ B(t_1-1, x(t_1-1)) &\leq v(t_1-1, x(t_1-1), u(t_1-1), \\ &\quad B(t_1, x(t_1))). \end{aligned}$$

В силу свойства (13) функции $v(\alpha, \beta, \gamma, \delta)$ отсюда следует, что

$$B(t_0, x_0) \leq \int_{t=t_0}^{t_1-1} v(t, x(t), u(t), v(t+1)).$$

Сравнивая это неравенство с соотношением (15), убеждаемся, что сформулированное утверждение доказано.

З а м е ч а н и е. Легко показать, что функционал (14) приводит к критериям качества (3), (5) § 4 в случае, когда функция $\gamma(\alpha, \beta, \gamma, \delta)$ строится по этим критериям.

5. Прямое уравнение Беллмана. Вернемся к исходной формулировке задачи оптимизации, поставленной в п. 1. Однако вместо прямого погружения теперь осуществим обратное погружение (см. п. 2 § 5) процесса в семейство

$$x(t+1)=f(x(t), u(t), t), \quad x(\tau)=y, \quad t \in [t_0, \tau-1], \\ \tau \in [t_0, t_1-1].$$

Обозначим через $F^-(\tau)$ множество n -векторов y , при котором для данного τ сужение исходной задачи на систему имеет решение. Относительно функции $f(x, u, t)$ будем предполагать выполненными условия п. 1 § 6.

Пусть критерий качества (3) удовлетворяет условиям (17), (18) § 6. Тогда, используя обратный принцип оптимальности, для функции Беллмана $B^-(\tau, y) = \min I(u[t_0, \tau-1], \tau, y)$, нетрудно получить прямое уравнение Беллмана:

$$B^-(\tau, y) = \min_u \mu(B^-(\tau-1, f^{-1}(y, u, \tau-1)), u, \tau, y), \quad (16)$$

где минимум берется по всем r -векторам u , при которых вектор $f^{-1}(y, u, \tau-1) \in F^-(\tau-1)$.

Здесь символом $f^{-1}(y, u, \tau)$ обозначен оператор, который совокупности y, u, τ ставит в соответствие вектор x такой, что $y=f(x, u, \tau)$.

При сделанных предположениях каждое оптимальное управление удовлетворяет прямому уравнению Беллмана (16).

Если условие (18) § 6, наложенное на функцию μ , заменить ослабленным, то, как в п. 3, можно доказать, что допустимое управление, построенное по рецептам обратного достаточного принципа оптимальности, удовлетворяет прямому уравнению Беллмана.

Обратной суперпозицией функции $\mu(\alpha, \beta, \gamma, \delta)$ вдоль функций $\beta(t), \gamma(t), \delta(t), t \in [t_0, t_1]$ назовем величину

$$\int_{t=t_1}^{t_0} \mu(\mu(t-1), \beta(t-1), \gamma(t), \delta(t)) = \mu(\dots \mu(\mu(t_0-1),$$

$$\beta(t_0-1), \gamma(t_0), \delta(t_0)), \dots), \beta((t_1-1), \gamma(t_1), \delta(t_1)),$$

где $\mu(t_0-1)$ — некоторое фиксированное число.

Пусть функция $B^-(t_0, y)$, $y \in F^-(t_0)$ определена по исходной задаче непосредственно. По уравнению (16) вычислим функции

$$B^-(t_0 + 1, y), B^-(t_0 + 2, y), \dots, B^-(t_1, y)$$

и соответствующие им управления

$$u(t_0, y), u(t_0 + 1, y), \dots, u(t_1 - 1, y).$$

Если правый конец $x(t)$ траектории по условиям задачи оптимизации подвижен $x(t_1) \in G(t_1)$, найдем вектор $x^0(t_1)$:

$$B^-(t_1, x^0(t_1)) = \min B^-(t_1, y), y \in G(t_1). \quad (17)$$

Подсчитаем управление $u^0(t)$, $t \in [t_0, t_1 - 1]$, по схеме

$$u^0(t_1 - 1) = u(t_1 - 1, x^0(t_1)), u^0(t_1 - 2) = u(t_1 - 2, x^0(t_1 - 1)), \dots, u^0(t_0) = u(t_0, x^0(t_0 + 1)), \quad (18)$$

где точка $x^0(t)$ строится по $x^0(t+1)$ так, чтобы

$$x^0(t+1) = f(x^0(t), u^0(t), t).$$

Как и в п. 4, нетрудно показать, что если функция $\mu(\alpha, \beta, \gamma, \delta)$ такова, что при $\alpha \leq \bar{\alpha}$ выполняется неравенство $\mu(\alpha, \beta, \gamma, \delta) \leq \mu(\bar{\alpha}, \beta, \gamma, \delta)$, то на управлении (18) достигает минимума критерий качества

$$I = \int_{t=t_1}^{t_0+1} \mu(\mu(t-1), u(t-1), t, x(t)), \quad (19)$$

где $\mu(t_0) = B^-(t_0, x(t_0))$.

Таким образом, доказано, что каждое управление, полученное из уравнения (16), является оптимальным в смысле критерия (19), определенного на траектории системы (1).

§ 8. Применение динамического программирования к задачам терминального управления

Исходя из принципов оптимальности и инвариантного погружения, получим уравнения Беллмана для нескольких типичных задач оптимального управления конечным состоянием дискретных процессов.

1. Простейшая задача терминального управления.
Пусть требуется решить задачу

$$x(t+1) = f(x(t), u(t), t), \quad x(t_0) = x_0, \quad t \in [t_0, t_1 - 1]; \quad (1)$$

$$u(t) \in U(t); \quad (2)$$

$$I(u) = \varphi(x(t_1)) \rightarrow \min. \quad (3)$$

Погрузим процесс $x(t)$ в семейство процессов

$$x(t+1) = f(x(t), u(t), t), \quad x(\tau) = y, \quad t \in [\tau, t_1 - 1]. \quad (4)$$

На введенном семействе определим критерий качества (3). При фиксированных y, τ минимальное значение критерия качества назовем функцией Беллмана $B(\tau, y)$ в задаче (1) — (3):

$$B(\tau, y) = \min \varphi(x(t_1)), \quad (5)$$

где $x(t_1)$ — конечное состояние процесса из семейства (4) и минимум берется по всем допустимым управлениям.

Следуя схеме § 7, получим уравнение для функции Беллмана.

Пусть $x^0(t), t \in [\tau, t_1], \quad x^0(\tau) = y$ — оптимальный процесс. Отрезок $[\tau, t_1]$ разобьем на два периода $[\tau]$ и $[\tau+1, t_1]$, отнеся к первому периоду лишь один момент τ . Согласно принципу оптимальности (см. § 6), для оптимальности всего процесса необходимо, чтобы управление во втором периоде было оптимальным относительно состояния, возникшего после первого периода. Если в первом периоде управление $u(t)$ равно $u(\tau) = v$, то состояние процесса к началу второго периода в силу (4) будет таким: $x(\tau+1) = f(y, v, \tau)$. Поэтому оптимальное управление во втором периоде доставит критерию качества значение $B(\tau+1, f(y, v, \tau))$. В силу специальной структуры критерия качества оно равно значению критерия качества на управлении, составленном из v , и оптимального управления во втором периоде. Значение $B(\tau+1, f(y, v, \tau))$ зависит от выбора управления v в первом периоде. Если $v \in U(\tau)$ выбрать так, чтобы минимизировать $B(\tau+1, f(y, v, \tau))$, то результат в силу достаточности принципа оптимальности даст $B(\tau, y)$, ибо при этом управление будет оптимальным на всем от-

резке $[\tau, t_1 - 1]$. Таким образом, функция Беллмана удовлетворяет уравнению

$$B(\tau, y) = \min_{v \in U(\tau)} B(\tau + 1, f(y, v, \tau)). \quad (6)$$

Начальные условия для уравнения (6) получаются из (3):

$$B(t_1, y) = \varphi(y). \quad (7)$$

Рекуррентное уравнение (6) с начальным условием (7) назовем обратным уравнением Беллмана в простейшей задаче терминального управления. Оптимальный закон управления $u^0(x, t)$ удовлетворяет условию

$$B(t + 1, f(x, u^0(x, t), t)) = \min_{u \in U(t)} B(t + 1, f(x, u, t)),$$

$$t = t_0, t_0 + 1, \dots, t_1 - 1.$$

Оптимальный закон управления $u^0(x, t)$ в отличие от оптимального управления $u^0(t)$ зависит от текущих момента времени и состояния. Ясно, что $u^0(t) = u^0(x^0(t), t)$.

З а м е ч а н и е. Иногда для получения уравнения Беллмана при погружении процесса (1) в семейство параметр τ вводят несколько иначе, чем в (4), а именно: полагают

$$x(t + 1) = f(x(t), u(t), t), x(t_1 - \tau) = y, t \in [t_1 - \tau, t_1 - 1]. \quad (8)$$

Пусть $B(\tau, y) = \min \varphi(x(t_1))$ вдоль допустимых траекторий семейства (8). Если $x^0(t)$, $t_1 - \tau \leq t \leq t_1$, — оптимальный процесс, то, рассматривая опять два периода $[t_1 - \tau]$ и $[t_1 - \tau + 1, t_1 - 1]$, получаем уравнение Беллмана в задаче (1) — (3):

$$B(\tau, y) = \min_{v \in U(t_1 - \tau)} B(\tau + 1, f(y, v, t_1 - \tau));$$

$$B(0, y) = \varphi(y).$$

Оптимальный закон управления $u^0(x, t)$ удовлетворяет условиям

$$B(t_1 - t - 1, f(x, u^0(x, t), t)) = \min_{u \in U(t)} B(t_1 - t - 1, f(x, u, t)),$$

$$t = t_1 - 1, \dots, t_0.$$

2. Прямое уравнение Беллмана. Для задачи

$$x(t + 1) = f(x(t), u(t), t), x(t_1) = x_1, u(t) \in U(t),$$

$$t \in [t_0, t_1 - 1]; \quad (9)$$

$$I(u) = \varphi(x(t_0)) \rightarrow \min$$

естественным является обратное погружение процесса (9), введенное в п. 2 § 5:

$$x(t+1) = f(x(t), u(t), t), \quad x(\tau) = y, \quad t_0 \leq t \leq \tau - 1, \\ \tau \in [t_0 + 1, t_1].$$

Определим функцию Беллмана

$$B^-(\tau, y) = \min \varphi(x(t_0)), \quad u[t_0, \tau - 1], \quad x(\tau) = y.$$

Тогда (см. п. 5 § 7)

$$B^-(\tau, y) = \min_{\substack{u[t_0, \tau-1] \\ x(\tau)=y}} \varphi(x(t_0)) = \min_{u(\tau-1)} \min_{\substack{x(\tau-1)=f^{-1}(y, u(\tau-1), \tau-1) \\ u[t_0, \tau-2]}} \varphi(x(t_0)) = \\ = \min_{u \in U(\tau-1)} B^-(\tau-1, f^{-1}(y, u, \tau-1));$$

$$B^-(t_0, y) = \varphi(y).$$

Таким образом, получилось прямое уравнение Беллмана:

$$B^-(\tau, y) = \min_{v \in U(\tau-1)} B^-(\tau-1, f^{-1}(y, v, \tau-1));$$

$$B^-(t_0, y) = \varphi(y).$$

3. Процессы с нефиксированной продолжительностью.

До сих пор при оптимизации дискретных процессов предполагалось, что отрезок $T = [t_0, t_1]$, на котором развивается процесс, фиксирован. Для приложений представляет интерес и случай, когда этот отрезок не фиксирован, а его концы могут выбираться из заданных множеств $T_0, T_1, t_0 \in T_0, t_1 \in T_1$. Критерий качества

$$I = I(t_0, x_0, u[t_0, t_1 - 1], t_1) \quad (10)$$

является теперь функцией и величин t_0, t_1 , по которым производится дополнительная минимизация. Естественным методом решения задачи при указанных условиях является следующий. Задаются некоторые значения $t, t_1, t_0 \in T_0, t_1 \in T_1$. Решается задача оптимизации процесса с фиксированной продолжительностью $[t_0, t_1]$. Функция Беллмана и оптимальный закон управления будут, очевидно, зависеть от t_0, t_1 : $B(\tau, y, t_0, t_1), u(y, \tau, t_0, t_1)$. Величина $B(t_0, x_0, t_0, t_1)$ равна, по доказанному ранее, минимальному значению критерия (10) при фиксированных t_0, t_1 . Поэтому оптимальная продолжительность про-

цесса T^0 , определяемая оптимальным начальным моментом t_1^0 и оптимальным конечным моментом t_1^0 , находится из условия

$$B(t_0^0, x_0, t_0^0, t_1^0) = \min \min B(t_0, x_0, t_0, t_1), \quad t_0 \in T_0, \quad t_1 \in T_1,$$

где $B(t_0^0, x_0, t_0^0, t_1^0)$ — минимальное значение критерия качества в рассматриваемой задаче.

Оптимальное управление $u^0(t)$, $t \in [t_0^0, t_1^0 - 1]$ находится по обычной схеме:

$$u^0(t) = u(x^0(t), t, t_0^0, t_1^0), \quad t \in [t_0^0, t_1^0 - 1],$$

где $x^0(t)$, $t \in T^0$ — оптимальная траектория.

З а м е ч а н и е. При решении задачи по описанной схеме число аргументов функции Беллмана возрастает на два, что существенно усложняет практическую осуществимость вычислений (см. гл. III). В следующих пунктах предлагается другой подход к задачам оптимизации с нефиксированной продолжительностью процесса.

4. Процессы с нефиксированным конечным моментом. В задаче п. 1 сохраним все предположения, но заменим критерий качества и вместо (3) рассмотрим

$$I(u, t_1) = \varphi(x(t_1), t_1).$$

Будем предполагать, что конечный момент t_1 не фиксирован, а может выбираться из отрезка $T_1 = [\theta_1, \theta_2]$ $\theta_0 \leq \theta_1 < \theta_2 < \infty$.

Задача оптимального управления теперь состоит в нахождении момента t_1^0 , $t_1^0 \in T_1$, и допустимого управления $u^0(t)$, $t \in [t_0, t_1^0 - 1]$, таких, что

$$I(u^0, t_1^0) = \min I(u, t_1), \quad u(t) \in U(t), \quad t \in [t_0, t_1 - 1], \quad t_1 \in T_1$$

Для решения поставленной задачи погрузим процесс (1) в семейство

$$x(t+1) = f(x(t), u(t), t), \quad x(\tau) = y, \quad t \in [\tau, t_1 - 1] \quad (11)$$

и введем функцию Беллмана

$$B(\tau, y) = \min \varphi(x(t_1), t_1), \quad (12)$$

где минимум берется по допустимым управлениям $u(t)$, $t \in [\tau, t_1 - 1]$, и моментам t_1 из отрезка $[\max(\tau, \theta_1), \theta_2]$.

Найдем уравнение Беллмана. Рассмотрим сначала случай $t_0 \leq \tau < \theta_1$. Справедливы равенства:

$$\begin{aligned}
B(\tau, y) &= \min_{t_1 \in [\theta_1, \theta_2]} \min_{u \in [\tau, t_1-1]} \varphi(x(t_1), t_1) = \\
&= \min_{u(\tau) \in U(\tau)} \min_{t_1 \in [\theta_1, \theta_2]} \min_{\substack{u(\tau+1, t_1-1) \\ x(\tau+1)=f(y, u(\tau), \tau)}} \varphi(x(t_1), t_1) = \\
&= \min_{u \in U(\tau)} B(\tau+1, f(y, u, \tau)). \quad (13)
\end{aligned}$$

Рассмотрим вторую возможность $\theta_1 \leq \tau < \theta_2 - 1$. В этом случае в силу (11), (12) имеем:

$$\begin{aligned}
B(\tau, y) &= \min_{t_1 \in [\tau, \theta_2]} \min_{u \in [\tau, t_1-1]} \varphi(x(t_1), t_1) = \min \{ \varphi(x(\tau), \tau), \\
&\min_{t_1 \in [\tau+1, \theta_2]} \min_{u(\tau) \in U(\tau)} \min_{\substack{u(\tau+1, t_1-1) \\ x(\tau+1)=f(y, u(\tau), \tau)}} \varphi(x(t_1), t_1) \} = \\
&= \min_{u \in U(\tau)} \min \{ \varphi(y, \tau), B(\tau+1, \\
&\quad f(y, u, \tau)) \}. \quad (14)
\end{aligned}$$

Наконец, при $\tau = \theta_2 - 1$ получаем

$$\begin{aligned}
B(\theta_2 - 1, y) &= \min_{u \in U(\theta_2-1)} \min \{ \varphi(y, \theta_2 - 1), \\
&\varphi(f(y, u, \theta_2 - 1), \theta_2) \}. \quad (15)
\end{aligned}$$

Из соотношений (13)—(15) следует уравнение Беллмана:

$$B(\tau, y) = \begin{cases} \min_{u \in U(\tau)} B(\tau+1, f(y, u, \tau)), & \text{если } t_0 \leq \tau < \theta_1 - 1 \\ \min_{u \in U(\tau)} \min \{ \varphi(y, \tau), \\ B(\tau+1, f(y, u, \tau)) \}, & \text{если } \theta_1 \leq \tau < \theta_2 - 1 \end{cases} \quad (16)$$

с начальным условием (см. (15))

$$\begin{aligned}
B(\theta_2 - 1, y) &= \min \{ \varphi(y, \theta_2 - 1), \min_{u \in U(\theta_2-1)} \varphi(f(y, u, \\
&\theta_2 - 1), \theta_2) \}. \quad (17)
\end{aligned}$$

С помощью (16), (17) исходная задача решается по следующей схеме. Пусть $\tau = \theta_2 - 2, \theta_2 - 3, \dots, t_0$. Из (16) с использованием (17) находим функции

$$B(\theta_2 - 2, y), B(\theta_2 - 3, y), \dots, B(t_0, y).$$

Отмечая на каждом шаге функции, на которых в (16), (17) достигается минимум по u , получаем законы управления:

$$u(y, \theta_2 - 1), u(y, \theta_2 - 2), \dots, u(y, t_0).$$

Значение $B(t_0, x_0)$ равно минимальному значению критерия качества. Вектор $u^0(t_0) = u^0(x_0, t_0)$ представляет собой значение оптимального управления в момент $t = t_0$. В остальные моменты оптимальное управление находится по формуле

$$u^0(t) = u(x^0(t), t),$$

где $x^0(t)$ — состояние в момент t процесса (1), порожденного начальным состоянием x_0 и управлением $u^0(t_0), \dots, u^0(t-1)$.

Оптимальный конечный момент t_1^0 , характеризующий оптимальную продолжительность процесса, равен первому моменту t_1 , при котором

$$B(t_1, x^0(t_1)) < B(t_1 + 1, x^0(t_1 + 1)).$$

5. Процессы с нефиксированным начальным моментом.

Рассмотрим опять процессы нефиксированной продолжительности, но конечный момент t_1 зафиксируем, а начальный t_0 выберем из множества $T_0 = [v_1, v_2]$, — $-\infty < v_1 \leq v_2 < t_1$.

Поскольку на значения критерия качества влияет выбор начального момента t_0 , теперь в задаче п. 1 с критерием качества $I(u, t_0) = \varphi(x(t_1))$ наряду с поиском оптимального управления $u^0(t)$ будем искать и оптимальный начальный момент t_0^0 :

$$I(u^0, t_0^0) = \min_{t_0 \in T_0} \min_{u[t_0, t_1-1]} I(u, t_0). \quad (18)$$

После прямого погружения исходного процесса введем функцию

$$B(\tau, y) = \min \varphi(x(t_1)), \quad \tau \in T_0,$$

где минимум берется по допустимым управлениям $u[t_0, t_1-1]$ и t_0 из $[v_1, v_2]$.

Нетрудно видеть, что при $v_1 \leq \tau \leq v_2 - 1$

$$\begin{aligned}
B(\tau, y) &= \min_{t_0 \in [\tau, v_2]} \min_{\substack{u[t_0, t_1-1] \\ x(t_0)=y}} \varphi(x(t_1)) = \\
&= \min \{ \min_{\substack{u[\tau, t_1-1] \\ x(\tau)=y}} \varphi(x(t_1)), B(\tau+1, y) \}. \quad (19)
\end{aligned}$$

Если $\tau = v_2$, то

$$B(v_2, y) = \min_{\substack{u[v_2, t_1-1] \\ x(v_2)=y}} \varphi(x(t_1)). \quad (20)$$

Для каждого $\tau, \tau \in [v_1, v_2-1]$, введем еще одну функцию Беллмана

$$B^1(\theta, z, \tau) = \min_{\substack{u[\theta, t_1-1] \\ x(\theta)=z}} \varphi(x(t_1)), \quad u[\theta, t_1-1], \quad x(\theta)=z, \quad \theta \in [\tau, t_1-1].$$

Для этих функций имеем

$$\begin{aligned}
B^1(\theta, z, \tau) &= \min_{\substack{u[\theta, t_1-1] \\ x(\theta)=z}} \varphi(x(t_1)) = \min_{u(\theta)} \min_{\substack{u[\theta+1, t_1-1] \\ x(\theta+1)=f(z, u(\theta), \theta)}} \varphi(x(t_1)) = \\
&= \min_{u \in U(\theta)} B^1(\theta+1, f(z, u, \theta), \tau), \quad \theta \in [\tau, t_1-1], \quad (21)
\end{aligned}$$

где $B^1(t_1, z, \tau) = \varphi(z)$.

Поскольку

$$\min_{\substack{u[\tau, t_1-1] \\ x(\tau)=y}} \varphi(x(t_1)) = B^1(\tau, y, \tau),$$

то, объединяя выражения (19) — (21), получаем систему уравнений Беллмана и граничных условий к ним:

$$B(\tau, y) = \min \{ B^1(\tau, y, \tau), B(\tau+1, y) \}, \quad v_1 \leq \tau \leq v_2-1; \quad (22)$$

$$B(v_2, y) = B^1(v_2, y, v_2); \quad (23)$$

$$\begin{aligned}
B^1(\theta, z, \tau) &= \min_{u \in U(\theta)} B^1(\theta+1, f(z, u, \theta), \tau), \quad \theta \in [\tau, \\
&\quad t_1-1]; \quad (24)
\end{aligned}$$

$$B^1(t_1, z, \tau) = \varphi(z). \quad (25)$$

Процедура вычисления решения задачи (18) состоит в следующем. Положим $\tau = v_2-1$, $\tau = v_2$. Из (25), (24) найдем последовательности функций

$$\begin{aligned}
&B^1(t_1, z, v_2-1), B^1(t_1-1, z, v_2-1), \dots, \\
&B^1(v_2, z, v_2-1), B^1(v_2-1, z, v_2-1);
\end{aligned}$$

$$B^1(t_1, z, v_2), B^1(t_1 - 1, z, v_2), \dots, B^1(v_2, z, v_2)$$

и соответствующих им в силу (24) законов управления

$$u(t_1 - 1, z, v_2 - 1), u(t_1 - 2, z, v_2 - 1), \dots,$$

$$u(v_2 - 1, z, v_2 - 1);$$

$$u(t_1 - 1, z, v_2), u(t_1 - 2, z, v_2), \dots, u(v_2, z, v_2).$$

Из (23) находим $B(v_2, y)$, а из (22) функцию $B(v_2 - 1, y)$. Следующий шаг является начальным для однотипных вычислений.

При $\tau = v_2 - 2$ из (25), (24) определяем функции

$$B^1(t_1 - 1, z, v_2 - 2), B^1(t_1 - 2, z, v_2 - 2), \dots,$$

$$B^1(v_2 - 1, z, v_2 - 2), B^1(v_2 - 2, z, v_2 - 2)$$

и законы управления

$$u(t_1 - 1, z, v_2 - 2), u(t_1 - 2, z, v_2 - 2), \dots,$$

$$u(v_2 - 2, z, v_2 - 2).$$

Из (22) находим $B(v_2 - 2, y)$. Продолжая этот процесс, получаем функции

$$B(v_2, y), B(v_2 - 1, y), \dots, B(v_1, y) \quad (26)$$

и законы управления

$$u(t_1 - 1, y, \tau), \dots, u(\tau, y, \tau), \tau \in [v_1, v_2 - 1]. \quad (27)$$

Полагая в (26) $y = x_0$, найдем среди полученных чисел наименьшее $B(v^0, x_0)$:

$$B(v^0, x_0) = \min B(v, x_0), v \in [v_1, v_2].$$

Число $t_0^0 = v^0$ есть оптимальный начальный момент; $B(v^0, x_0)$ — минимальное значение критерия качества (18). Оптимальное управление $u^0(t)$, $t \in [t_0^0, t_1 - 1]$ находится из (27): $u^0(t) = u(t, x^0(t), t_0^0)$.

З а м е ч а н и е. Описанная схема решения задачи оптимизации с нефиксированным начальным моментом лишь формально отличается от общей схемы решения, предложенной в п. 3.

Принципиальное изменение схемы п. 3 достигается в задачах оптимизации с нефиксированным начальным моментом, если критерий качества имеет вид $I(u, t_0) = \varphi(x(t_0), t_0)$. В этом случае, используя обратное погружение, нетрудно получить аналог результата из п. 4 в терминах прямого уравнения Беллмана относительно функции только $n+1$ переменных (а не $n+2$, как в данном пункте).

6. Случай подвижного правого конца. Пусть задача п. 1 дополнена ограничением $x(t_1) \in G_1$, где G_1 — некоторое множество n -мерного пространства.

Используя прямое погружение из п. 1, введем функцию Беллмана

$$B(\tau, y) = \min \varphi(x(t_1)), \quad (28)$$

где \min берется по допустимым управлениям, приводящим в момент $t=t_1$ траектории $x(t)$, $t \in [\tau, t_1]$, системы (4) на множество G_1 .

Обозначим через $F(\tau)$, $\tau \in [t_0, t_1-1]$, $F(t_1) = G_1$ множество точек, начиная из которых $(x(\tau) = y)$ допустимая траектория $x(t)$, $t \in [\tau, t_1]$, попадает в момент $t=t_1$ на множество G_1 . Через $\partial^0 F(\tau)$ обозначим точки пограничного слоя множества $F(\tau)$, отнеся к ним точки y , для которых существует век-

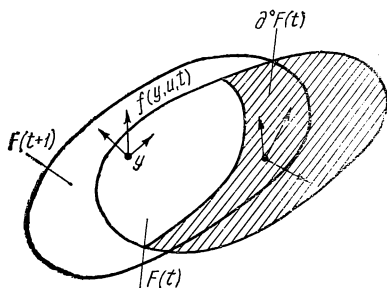


Рис. 4

тор $u \in U(\tau)$ такой, что (рис. 4) $f(y, u, \tau) \in F(\tau+1)$.

Для всех τ , $\tau \in [t_0, t_1-1]$ и $y, y \in \partial^0 F(\tau)$, построим подмножество $U_1(y, \tau)$ множества $U(\tau)$ по правилу: вектор $u \in U(\tau)$ отнесем к $U_1(y, \tau)$, если $f(y, u, \tau) \in F(\tau+1)$.

Ясно, что функция Беллмана $B(\tau, y)$ определена лишь на множестве $F(\tau)$.

Пусть $y \in \partial^0 F(\tau)$. Тогда, рассуждая, как в п. 1, получим

$$B(\tau, y) = \min B(\tau+1, f(y, u, \tau)), u \in U(\tau); \quad (29)$$

$$B(t_1, y) = \varphi(y), y \in G_1. \quad (30)$$

Оптимальный закон управления $u^0(y, \tau)$ находится из условия

$$B(\tau+1, f(y, u^0(y, \tau), \tau)) = \min B(\tau+1, f(y, u, \tau)), u \in U(\tau).$$

Если $y \in \partial^0 F(\tau)$, то уравнение Беллмана имеет вид

$$B(\tau, y) = \min B(\tau+1, f(y, u, \tau)), u \in U_1(y, \tau). \quad (31)$$

Действительно, пусть $y \in \partial^0 F(\tau)$, тогда

$$B(\tau, y) = \min_{\substack{u[\tau, t_1-1], \\ x(t_1) \in G_1}} \varphi(x(t_1)) = \min_{u(\tau) \in U_1(y, \tau)} \varphi(x(t_1))$$

$$\min_{\substack{u[\tau+1, t_1-1], \\ x(t_1) \in G_1}} \varphi(x(t_1)) = \min_{u \in U_1(y, \tau)} B(\tau+1, f(y, u, \tau)).$$

Оптимальный закон $u(y, \tau)$ при $y \in \partial^0 F(\tau)$ определяется из условия

$$B(\tau+1, f(y, u^0(y, \tau), \tau)) = \min_{u \in U_1(y, \tau)} B(\tau+1, f(y, u, \tau)).$$

Таким образом, из (29)–(31) следует, что для задачи оптимизации данного пункта уравнение Беллмана определено при $\tau \in [t_0, t_1-1]$, $y \in F(\tau)$ и имеет вид

$$B(\tau, y) = \begin{cases} \min_{u \in U(\tau)} B(\tau+1, f(y, u, \tau)), & \text{если } y \in \partial^0 F(\tau), \\ \min_{u \in U_1(y, \tau)} B(\tau+1, f(y, u, \tau)), & \text{если } y \in \partial^0 F(\tau), \end{cases}$$

с граничным условием $B(t_1, y) = \varphi(y)$.

Множества $F(\tau)$, $\tau \in [t_0, t_1]$, строятся по формулам $F(t_1) = G_1$, $F(\tau) = \{y: f(y, u, \tau) \in F(\tau+1) \text{ при некотором } u \in U(\tau)\}$.

З а м е ч а н и е. Иногда множество $F(\tau)$ называют множеством управляемости в момент $t = \tau$.

7. Случай подвижного левого конца. Пусть в задаче п. 1 левый конец является подвижным, т. е. $x(t_0) \in G_0$, где G_0 — заданное множество n -мерного пространства.

Введем ту же функцию Беллмана, что в п. 1. Вычислим ее согласно уравнению (6) и условию (2). В результате получим функцию $B(t_0, y)$. После этого находим левый конец $x_0^0 = x^0(t_0)$ оптимальной траектории $x_0(t)$ из условия

$$B(t_0, x_0^0) = \min_{y \in G_0} B(t_0, y).$$

Оптимальное управление $u^0(t)$ находим по формуле $u^0(t) = u(x^0(t), t)$, при использовании которой вместо x_0 берется вектор x_0^0 .

8. Учет фазовых ограничений. Рассмотрим задачу п. 1 при дополнительных ограничениях

$$x(t) \in G(t), \quad t \in [t_0+1, t_1-1]. \quad (32)$$

Сохраняя прямое погружение, введем функцию Беллмана

$$B(\tau, y) = \min \varphi(x(t_1)), \quad (33)$$

где минимум берется по всем допустимым управлениям, т. е. по управлениям $u[\tau, t_1-1]$, которые удовлетворяют ограничениям (2), а траектории $x[\tau, t_1]$, $x(\tau)=y$, ими порожденные, — ограничениям (32).

Пусть $F(\tau)$, $\tau \in [t_0, t_1-1]$, $F(t_1)=R_n$, — множество векторов $x \in G(\tau)$, таких, что существует допустимая траектория $x(t)$ $t \geq \tau$, $x(\tau)=x$, системы, (4), определенная при $t \in [\tau, t_1]$. Через $\partial^0 F(\tau)$ обозначим пограничный слой множества $F(\tau)$, т. е. точки $x \in F(\tau)$, для которых существуют такие векторы $u \in U(\tau)$, что $f(x, u, \tau) \in F(\tau+1)$ (рис. 5). Множество $U_1(y, \tau)$ определяется, как в п. 5.

Ясно, что функция

$B(\tau, y)$ определена лишь в точках множества $F(\tau)$.

Выведем уравнение Беллмана для функции (33). Будем различать два случая: а) $y \in \partial^0 F(\tau)$; б) $y \in \partial^0 F(\tau)$.

В случае а) имеем:

$$\begin{aligned} B(\tau, y) &= \min_{\substack{u[\tau, t_1-1], \\ x(\tau)=y \\ x(t) \in G(t), t \in [\tau+1, t_1-1]}} \varphi(x(t_1)) = \\ &= \min_{u(\tau) \in U(\tau)} [\min_{\substack{u[\tau+1, t_1-1], \\ x(\tau+1)=f(y, u(\tau), \tau) \\ x(t) \in G(t), t \in [\tau+2, t_1-1]}} \varphi(x(t_1))]. \end{aligned} \quad (34)$$

Поскольку выражение в квадратных скобках, по определению, есть $B(\tau+1, f(y, u, \tau))$, из (34) получается уравнение для функции Беллмана:

$$B(\tau, y) = \min_{u \in U(\tau)} B(\tau+1, f(y, u, \tau)). \quad (35)$$

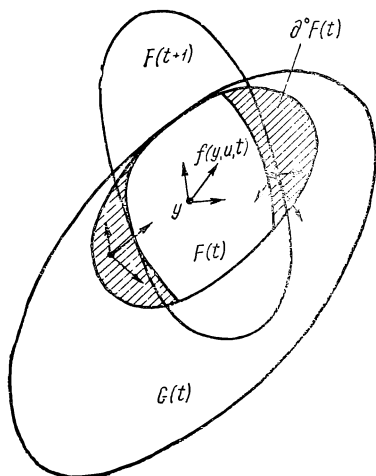


Рис. 5

В случае б):

$$B(\tau, y) = \min_{\substack{u[\tau, t_1-1], x(\tau)=y \\ x(t) \in G(t), t \in [\tau+1, t_1-1]}} \varphi(x(t_1)) = \\ = \min_{u(\tau) \in U_1(y, \tau)} [\min_{\substack{u[\tau+1, t_1-1], x(\tau+1)=f(y, u(\tau), \tau) \\ x(t) \in G(t), t \in [\tau+2, t_1-1]}} \varphi(x(t_1))].$$

Нетрудно заметить, что выражение в квадратных скобках есть $B(\tau+1, f(y, u(\tau), \tau))$. Поэтому уравнение Беллмана принимает вид

$$B(\tau, y) = \min_{u \in U_1(y, \tau)} B(\tau+1, f(y, u, \tau)). \quad (36)$$

Граничное условие получается из определения (33):

$$B(t_1, y) = \varphi(y). \quad (37)$$

Ясно, что $F(t_1) = R_n$, $F(t_1-1) = G(t_1-1)$, остальные множества $F(t)$, $t_0 \leq t \leq t_1-2$, получаются последовательно по формуле $F(\tau) = \{y: f(y, u, \tau) \in F(\tau+1), y \in G(\tau), \text{ при некотором } u \in U(\tau)\}$.

Построение оптимального управления проводится по обычной схеме с использованием уравнений (35), (36) и граничного условия (37).

9. Смешанные ограничения. Еще раз вернемся к задаче п. 1 и добавим ограничение $\{x(t), u(t)\} \in H(t)$, $t \in [t_0, t_1-1]$, где $H(t)$ — заданное множество $n+r$ -мерного пространства. Сохраним тип вложения и введем функцию Беллмана

$$B(\tau, y) = \min_{\substack{u[\tau, t_1-1], x(\tau)=y \\ \{x(t), u(t)\} \in H(t) \\ t \in [\tau, t_1-1]}} \varphi(x(t_1)), \quad (38)$$

которая, как обычно, имеет смысл минимума критерия качества на допустимых управлениях задачи терминального управления семейством (4).

Пусть $F(\tau)$ — множество n -векторов y , из которых выходят допустимые траектории $x(t)$, $x(\tau)=y$, определенные на $[\tau, t_1-1]$. Обозначим через $\partial^0 F(\tau)$ пограничный слой множества $F(\tau)$, отнеся к нему те точки $y \in F(\tau)$, для которых существуют такие элементы

$u \in U(\tau)$, что $f(y, u, \tau) \in F(\tau+1)$. Для моментов τ , $\tau \in [t_0, t_1-1]$, и точек $y, y \in F(\tau)$ определим подмно-

жество $U_1(y, \tau)$ множества $U(\tau)$. Вектор $u \in U(\tau)$ принадлежит множеству $U_1(y, \tau)$, если $f(y, u, \tau) \in F(\tau + 1)$.

Функция $B(\tau, y)$ определена, очевидно, лишь на множестве $F(\tau)$. Вид уравнения для $B(\tau, y)$ зависит от того, лежит точка y в пограничном слое множества $F(\tau)$ или нет.

Пусть $y \in F(\tau)$, $y \in \partial^0 F(\tau)$. Тогда

$$B(\tau, y) = \min_{\substack{u(t) \in U(t), x(\tau)=y \\ \{x(t), u(t)\} \in H(t), t \in [\tau, t_1-1]}} \varphi(x(t_1)) = \min_{u \in U(\tau), \{y, u\} \in H(\tau)} B(\tau + 1, f(y, u, \tau)). \quad (39)$$

Если $y \in F(\tau)$, $y \in \partial^0 F(\tau)$, то

$$B(\tau, y) = \min_{\substack{u(t) \in U(t), x(\tau)=y \\ \{x(t), u(t)\} \in H(t), t \in [\tau, t_1-1]}} \varphi(x(t_1)) = \min_{u \in U_1(y, \tau), \{y, u\} \in H(\tau)} B(\tau + 1, f(y, u, \tau)). \quad (40)$$

Граничное условие для $B(\tau, y)$ следует из определения (38):

$$B(t_1, y) = \varphi(y). \quad (41)$$

Множества $F(t)$, $t \in [t_0, t_1]$, строятся индуктивно:

$$F(t_1) = R_n;$$

$$F(t_1-1) = \{y : \{y, u\} \in H(t_1-1) \text{ при некотором } u \in U(t_1-1)\}. \quad (42)$$

$$F(t) = \{y : f(y, u, t) \in F(t+1) \text{ при некотором } u \in U(t) \\ \{y, u\} \in H(t)\}. \quad (43)$$

Процедура вычисления решения задачи оптимизации данного пункта состоит в следующем: по формуле (41) находится функция $B(t_1, y)$, определенная на всем n -мерном пространстве R_n , а из (42) — множество $F(t_1-1)$, по нему и множеству $H(t_1-1)$ строятся множества $\partial^0 F(t_1-1)$, $U_1(y, t_1-1)$. В точках $y \in F(t_1-1)$, $y \in \partial^0 F(t_1-1)$ функция $B(t_1-1, y)$ вычисляется из уравнения

$$B(\tau, y) = \min_{u \in U(\tau), \{y, u\} \in H(\tau)} B(\tau + 1, f(y, u, \tau)), \quad (44)$$

следующего из (39). Если $y \in F(t_1-1)$, $y \in \partial^0 F(t_1-1)$, то значение $B(t_1-1, y)$ подсчитывается из уравнения

$$B(\tau, y) = \min_{u \in U_1(y, \tau), \{y, u\} \in H(\tau)} B(\tau + 1, f(y, u, \tau)), \quad (45)$$

следующего из (40).

Одновременно определяется оптимальный закон управления $u(y, t_1 - 1)$, как вектор, на котором достигается минимум в (44) или в (45). Далее вычисления $F(\tau)$, $U_1(y, \tau)$, $B(\tau, y)$, $u(y, \tau)$, $\tau \in [t_0, t_1 - 2]$, идут индуктивно по формулам (43), (44), если $y \in F(\tau)$, $y \in \partial^0 F(\tau)$, и формулам (43), (45), если $y \in F(\tau)$, $y \in \partial^0 F(\tau)$. В результате получаются функции

$$B(t_1, y), B(t_1 - 1, y), \dots, B(t_0, y)$$

и законы управления

$$u(t_1 - 1, y), u(t_1 - 2, y), \dots, u(t_0, y).$$

Число $B(t_0, x_0)$ равно минимальному значению критерия качества, оптимальное управление $u^0(t)$ строится по формуле

$$u^0(t) = u(x^0(t), t), \quad t \in [t_0, t_1 - 1],$$

где $x^0(t)$ — оптимальная траектория.

10. Оптимизация двумя типами управлений. Обобщим модель управляемого дискретного процесса (1). Будем считать, что уравнения движения имеют вид

$$x(t + 1) = f(x(t), u(t), v(s(t)), t), \quad t \in [t_0, t_1 - 1],$$

$$x(t_0) = x_0, \quad (46)$$

где новый символ $v(s)$ означает управление, способное изменять свое значение лишь в моменты $s = s_0, s_0 + \Delta, \dots$; Δ — натуральное число, $\Delta > 1$.

Через $s(t)$ обозначим функцию, определенную формулой

$$s(t) = s_0 + \left\lceil \frac{t - s_0}{\Delta} \right\rceil \Delta.$$

Если $s(t_0) < t_0$, будем считать, что в (46) аргумент $v(s)$ принимает заданное значение $v(s(t_0))$ для всех t из промежутка $[t_0, s(t_0) + \Delta - 1]$. В моменты $s(t_0) + \Delta, s(t_0) + 2\Delta, \dots$ значения $v(s)$ можно выбрать из множества $V(s)$: $v(s) \in V(s)$.

Если $s(t_0)=t_0$, то управление $v(s)$ можно выбрать в моменты $t_0, t_0+\Delta, \dots$. Для упрощения изложения будем считать, что $s(t_0)=t_0$.

Рассмотрим задачу терминального управления

$$x(t+1) = f(x(t), u(t), v(s(t)), t), \quad x(t_0) = x_0, \\ t \in [t_0, t_1 - 1]; \quad (47)$$

$$u(t) \in U(t), \quad v(s) \in V(s), \quad s \in [t_0, s(t_1 - 1)]; \quad (48)$$

$$I(u, v) = \varphi(x(t_1)) \rightarrow \min. \quad (49)$$

Пусть σ — переменная, принимающая значения $t_0, t_0 + \Delta, \dots, t_0 + l_1 \Delta$, где $l_1 = \left\lfloor \frac{t_1 - t_0 - 1}{\Delta} \right\rfloor$. Погрузим процесс (46) в семейство

$$x(t+1) = f(x(t), u(t), v(s(t)), t), \quad x(\sigma) = y, \quad t \in [\sigma, t_1 - 1] \quad (50)$$

и введем функцию Беллмана

$$B(\sigma, y) = \min \varphi(x(t_1)), \quad (51)$$

где минимум берется вдоль траекторий системы (50), порожденных управлениями $u(t) \in U(t), t \in [\sigma, t_1 - 1], v(s) \in V(s), s \in [\sigma, s(t_1 - 1)]$.

При $t_0 \leq \sigma \leq s(t_1 - 1) - \Delta$ очевидны следующие соотношения для функции (51):

$$B(\sigma, y) = \min_{u[\sigma, t_1 - 1]} \min_{\substack{v[\sigma, s(t_1 - 1)] \\ x(\sigma) = y}} \varphi(x(t_1)) = \\ = \min_{v(\sigma)} \min_{u[\sigma, \sigma + \Delta - 1]} \min_{\substack{u[\sigma + \Delta, t_1 - 1], \\ x(\sigma + \Delta) = x(\sigma + \Delta, y, u[\sigma, \sigma + \Delta - 1], v(\sigma))}} \min_{v[\sigma + \Delta, s(t_1 - 1)]} \varphi(x(t_1)), \quad (52)$$

где символ $x(\sigma + \Delta, y, u[\sigma, \sigma + \Delta - 1], v(\sigma))$ означает состояние в момент $t = \sigma + \Delta$ процесса (50).

В силу определения (51) имеем

$$B(\sigma, y) = \min_{v \in V(\sigma)} \{ \min_{\substack{u(t) \in U(t), \\ x(\sigma) = y}} B(\sigma + \Delta, x(\sigma + \Delta, y, \\ u[\sigma, \sigma + \Delta - 1], v)) \}. \quad (53)$$

Для $\sigma = s(t_1 - 1)$ из (51) получаем

$$B(s(t_1 - 1), y) = \min_{u[s(t_1 - 1), t_1 - 1]} \min_{\substack{v(s(t_1 - 1)) \\ x(s(t_1 - 1)) = y}} \varphi(x(t_1)). \quad (54)$$

Если в (53) рассмотреть вторую операцию взятия минимума

$$\min_{u[\sigma, \sigma+\Delta-1], x(\sigma)=y} B(\sigma+\Delta, x(\sigma+\Delta, y, u[\sigma, \sigma+\Delta-1], v)),$$

легко заметить, что при фиксированном $v(\sigma)$ речь идет о новой задаче терминального управления. Поэтому можно ввести еще одну функцию Беллмана.

Пусть $t_0 \leq \sigma \leq s(t_1-1) - \Delta$. Введем вторую функцию Беллмана по формуле

$$B^1(\tau, z, \sigma, v) = \min_{\substack{u[\tau, \sigma+\Delta-1] \\ x(\tau)=z}} B(\sigma+\Delta, x(\sigma+\Delta, z, u[\sigma, \sigma+\Delta-1], v)). \quad (55)$$

Выше (п. 1) показано, что функция типа (55) удовлетворяет уравнению Беллмана:

$$B^1(\tau, z, \sigma, v) = \min_{u \in U(\tau)} B^1(\tau+1, f(z, u, v, \tau), \sigma, v), \quad \tau \in [\sigma, \sigma+\Delta-1], \quad (56)$$

с граничным условием

$$B^1(\sigma+\Delta, z, \sigma, v) = B(\sigma+\Delta, z), \quad \sigma \in [t_0, s(t_1-1) - \Delta]. \quad (57)$$

Минимальное значение величины, стоящей в фигурных скобках в (53), равно $B^1(\sigma, y, \sigma, v)$. Поэтому, подставляя это значение в (53), получаем соотношение

$$B(\sigma, y) = \min_{v \in V(\sigma)} B^1(\sigma, y, \sigma, v), \quad (58)$$

справедливое для всех σ таких, что $t_0 \leq \sigma \leq s(t_1-1) - \Delta$.

Если $\sigma = s(t_1-1)$, обратимся к равенству (54). Задачу

$$\min_{\substack{u[s(t_1-1), t_1-1] \\ x(s(t_1-1))=y}} \varphi(x(t_1)) \quad (59)$$

интерпретируем как задачу терминального управления и для нее введем функцию Беллмана

$$B^1(\tau, z, s(t_1-1), v) = \min_{\substack{u[\tau, t_1-1] \\ x(\tau)=z}} \varphi(x(t_1)).$$

Обычным путем получаем уравнение

$$B^1(\tau, z, s(t_1 - 1), v) = \min_{u \in U(\tau)} B^1(\tau + 1, f(z, u, v, \tau), \\ s(t_1 - 1), v), \tau \in [s(t_1 - 1), t_1 - 1], \quad (60)$$

с граничным условием

$$B^1(t_1 - 1, z, s(t_1 - 1), v) = \varphi(z). \quad (61)$$

Величина (59) равна

$$B^1(s(t_1 - 1), y, s(t_1 - 1), v).$$

Поэтому для функции (54) получаем представление

$$B(s(t_1 - 1), y) = \min_{v \in V(s(t_1 - 1))} B^1(s(t_1 - 1), y, s(t_1 - 1), v).$$

Объединяя уравнения (58), (56), (60) и граничные условия (57), (61), приходим окончательно к следующей системе уравнений Беллмана для исходной задачи (47) — (49) оптимизации по двум типам управлений:

$$B(\sigma, y) = \min_{v \in V(\sigma)} B^1(\sigma, y, \sigma, v), \sigma \in [t_0, s(t_1 - 1) - \Delta]; \quad (62)$$

$$B(s(t_1 - 1), y) = \min_{v \in V(s(t_1 - 1))} B^1(s(t_1 - 1), y, s(t_1 - 1), v); \quad (63)$$

$$B^1(\tau, z, \sigma, v) = \min_{u \in U(\tau)} B^1(\tau + 1, f(z, u, v, \tau), \theta, v),$$

$$\tau \in [\sigma, \sigma + \Delta - 1], \sigma \in [t_0, s(t_1 - 1) - \Delta]; \quad (64)$$

$$B^1(\tau, z, s(t_1 - 1), v) = \min_{u \in U(\tau)} B^1(\tau + 1, f(z, u, v, \tau), s(t_1 - 1), v),$$

$$\tau \in [s(t_1 - 1), t_1 - 1], \quad (65)$$

с граничными условиями

$$B^1(\sigma + \Delta, z, \sigma, v) = B(\sigma + \Delta, z), \sigma \in [t_0, s(t_1 - 1) - \Delta]; \quad (66)$$

$$B^1(t_1 - 1, z, s(t_1 - 1), v) = \varphi(z). \quad (67)$$

Процедура вычисления минимального значения критерия качества и оптимальных управлений в задаче (47) — (49) такова. Исходя из граничного условия (67), по уравнению (65) вычисляем последовательно функции

$$B^1(t_1 - 1, z, s(t_1 - 1), v), B^1(t_1 - 2, z, s(t_1 - 1), v), \dots, B^1(s(t_1 - 1), z, s(t_1 - 1), v)$$

и законы управления

$$u(z, t_1 - 1), u(z, t_1 - 2), \dots, u(z, s(t_1 - 1)),$$

на которых достигается минимум в правой части уравнения (65).

Из соотношения (63) находим $B(s(t_1 - 1), y)$ и закон управления $v(y, s(t_1 - 1))$, на котором достигается минимум в правой части (63). Далее полагаем $\sigma = s(t_1 - 1) - \Delta$ и из соотношения (66) находим функцию $B^1(s(t_1 - 1), z, s(t_1 - 1) - \Delta, v) \equiv B(s(t_1 - 1), z)$. Это позволяет по формуле (64) вычислить функции

$$B^1(s(t_1 - 1) - 1, z, s(t_1 - 1) - \Delta, v), B^1(s(t_1 - 1) - 2, z, s(t_1 - 1) - \Delta, v), \dots, B^1(s(t_1 - 1) - \Delta, z, s(t_1 - 1) - \Delta, v)$$

и законы управления

$$u(z, s(t_1 - 1) - 1), u(z, s(t_1 - 1) - 2), \dots, u(z, s(t_1 - 1) - \Delta),$$

на которых в (65) достигается минимум.

Из (62), $\sigma = s(t_1 - 1) - \Delta$, находим функцию $B(s(t_1 - 1) - \Delta, y)$ и закон управления $v(y, s(t_1 - 1) - \Delta)$.

Повторяя вычисления для $\sigma = s(t_1 - 1) - 2\Delta$ и т. д., через конечное число шагов найдем функцию $B(t_0, y)$ и законы управления $u(z, t), t \in [t_0, t_1 - 1], v(y, s), s \in [t_0, s(t_1 - 1)]$.

Число $B(t_0, x_0)$ — минимальное значение критерия качества в задаче (47) — (49), векторы $u^0(t_0) = u(x_0, t_0)$, $v^0(t_0) = v(x_0, t_0)$ — значения оптимальных управлений $u^0(t)$, $v^0(s)$ в момент $t = t_0$, $s = t_0$. Оптимальные управления в остальные моменты можно найти по формулам

$$u^0(t) = u(x^0(t), t), v^0(s) = v(x^0(s), s),$$

где $x^0(t)$ — состояние в момент t оптимального процесса, вычисленного по (46) на управлениях $u^0(\tau)$, $\tau \in [t_0, t - 1]$, $v^0(\sigma)$, $\sigma \in [t_0, s(t - 1)]$.

З а м е ч а н и е. Сравнивая полученные результаты для каждого из двух типов управлений, нетрудно заметить, что с увеличением Δ эффективность применения динамического программирования уменьшается. При $\Delta = t_1 - t_0$ (задача оптимизации по параметрам) при вычислении управления $v(s)$ не достигается никакого эффекта. Это еще раз подчеркивает то обстоятельство, что динамическое программирование наиболее эффективно для процессов, которые по существу и постоянно подвергаются воздействиям в ходе их развития.

11. Оптимизация двумя типами управлений (продолжение). Рассмотрим еще раз два типа управлений. Наряду с управлением $u(t)$, которое может менять свое значение в те же моменты времени, что и дискретный процесс $x(t)$, введем управление $v(s)$, значения которого могут меняться несколько раз между моментами t и $t+1$. В этом пункте будем считать, что $l = \frac{1}{\Delta}$ — целое число, $\Delta > 0$, равное количеству возможных изменений $v(s)$ на единичном промежутке. Параметр s принимает значения $t_0 + \Delta, t_0 + 2\Delta, \dots$

Пусть требуется решить задачу

$$x(t+1) = \sum_{s=t}^{t+1-\Delta} f(x(t), u(t), v(s), s, t), \quad x(t_0) = x_0, \quad t \in [t_0, t_1 - 1]; \quad (68)$$

$$u(t) \in U(t), \quad v(s) \in V(s); \quad (69)$$

$$I(u) = \varphi(x(t_1)) \rightarrow \min. \quad (70)$$

Погрузим процесс (68) в семейство

$$x(t+1) = \sum_{s=t}^{t+1-\Delta} f(x(t), u(t), v(s), s, t), \quad x(\tau) = y, \\ t \in [\tau, t_1 - 1]$$

и на траекториях семейства определим функцию Беллмана

$$B(\tau, y) = \min_{u[\tau, t_1-1]} \min_{v[\tau, t_1-\Delta]} \varphi(x(t_1)).$$

Ясно, что

$$B(\tau, y) = \min_{u[\tau, t_1-1]} \min_{v[\tau, t_1-\Delta]} \varphi(x(t_1)) =$$

$$= \min_{u \in U(\tau)} \min_{v[\tau, \tau+1-\Delta]} B(\tau+1, \sum_{s=\tau}^{\tau+1-\Delta} f(y, u, v(s), s, \tau)),$$

$$\tau \in [t_0, t_1 - 1]. \quad (71)$$

Рассмотрим отдельно задачу

$$\min_{v[\tau, \tau+1-\Delta]} B(\tau+1, \sum_{s=\tau}^{\tau+1-\Delta} f(y, u, v(s), s, \tau)), \quad \tau \in [t_0, t_1 - 1]. \quad (72)$$

Для ее решения введем фиктивный процесс $\tilde{x}(s)$, подчиненный уравнению

$$\tilde{x}(s + \Delta) = \tilde{x}(s) + f(y, u, v(s), s, \tau), \quad \tilde{x}(\tau) = 0, \quad s \in [\tau, \tau + 1 - \Delta], \quad v(s) \in V(s). \quad (73)$$

На траекториях этого процесса будем минимизировать критерий качества

$$\tilde{I}(v) = B(\tau+1, \tilde{x}(\tau+1)) \rightarrow \min. \quad (74)$$

Поскольку $\tilde{x}(\tau+1) = \sum_{s=\tau}^{\tau+1-\Delta} f(y, u, v(s), s, \tau)$, задача (73),

(74) эквивалентна задаче (72). Функцию Беллмана для задачи (73), (74) обозначим через $B^1(\sigma, z, \tau, u)$:

$$B^1(\sigma, z, \tau, u) = \min_{\substack{v[\sigma, \tau+1-\Delta] \\ \tilde{x}(\sigma)=z}} B(\tau+1, \tilde{x}(\tau+1)), \quad \sigma = \tau, \\ \tau + \Delta, \dots, \tau + 1 - \Delta.$$

Эта функция удовлетворяет (см. п. 1) уравнению

$$B^1(\sigma, z, \tau, u) = \min_{v \in V(\sigma)} B^1(\sigma + \Delta, z + f(y, u, v, \sigma, \tau), \tau, u),$$

$$\sigma \in [\tau, \tau + 1 - \Delta],$$

с граничным условием $B^1(\tau+1, z, \tau, u) = B(\tau+1, z)$. Минимальное значение критерия качества (74), а значит, минимальное значение величины (72), равно $B^1(\tau, 0, \tau, u)$. Поэтому соотношение (71) принимает вид

$$B(\tau, y) = \min_{u \in U(\tau)} B^1(\tau, 0, \tau, u), \tau \in [t_0, t_1 - 1].$$

При $\tau = t_1$ из (70) получается формула

$$B(t_1, y) = \varphi(y).$$

Объединяя полученные результаты, приходим к следующему выводу: оптимальные процессы в задаче (68) — (70) управляются системой уравнений Беллмана

$$B(\tau, y) = \min_{u \in U(\tau)} B^1(\tau, 0, \tau, u), \tau = t_0, t_0 + 1, \dots, t_1 - 1; \quad (75)$$

$$B^1(\sigma, z, \tau, u) = \min_{v \in V(\sigma)} B^1(\sigma + \Delta, z + f(y, u, v, \sigma, \tau),$$

$$\tau, u), \sigma = \tau, \tau + \Delta, \dots, \tau + 1 - \Delta, \quad (76)$$

с граничными условиями

$$B(t_1, y) = \varphi(y); \quad (77)$$

$$B^1(\tau + 1, z, \tau, u) = B(\tau + 1, z), \tau = t_0, t_0 + 1, \dots, t_1 - 1. \quad (78)$$

Процедура получения решения задачи (68) — (70) из уравнений Беллмана такова. Из (77) находим функцию

$$B(t_1, y) \equiv \varphi(y).$$

Полагая $\tau = t_1 - 1$, из (78) определяем функцию

$$B^1(t_1, z, t_1 - 1, u) \equiv B(t_1, z).$$

Тогда из уравнения (76) можно найти функции

$$B^1(t_1 - \Delta, z, t_1 - 1, u), B^1(t_1 - 2\Delta, z, t_1 - 1, u), \\ \dots, B^1(t_1 - 1, z, t_1 - 1, u)$$

и законы управления

$$v(z, t_1 - \Delta), v(z, t_1 - 2\Delta), \dots, v(z, t_1 - 1).$$

При $\tau = t_1 - 1$ из (75) определяем функцию $B(t_1 - 1, y)$ и закон управления $u(y, t_1 - 1)$.

Из (78) вычисляем функцию

$$B^1(t_1 - 1, z, t_1 - 2, u) \equiv B(t_1 - 1, z).$$

Используя ее в уравнении (76), находим функции

$$B^1(t_1 - 1 - \Delta, z, t_1 - 2, u), B^1(t_1 - 1 - 2\Delta, z, t_1 - 2, u), \dots, B^1(t_1 - 2, z, t_1 - 2, u)$$

и законы управления

$$v(z, t_1 - 1 - \Delta), v(z, t_1 - 1 - 2\Delta), \dots, v(z, t_1 - 2).$$

Продолжая вычисления по этой схеме, через конечное число шагов получим функцию $B(t_0, y)$ и законы управления $v(z, s)$, $s = t_0, t_0 + \Delta, \dots, t_1 - \Delta$; $u(y, t)$, $t = t_0, t_0 + 1, \dots, t_1 - 1$.

Число $B(t_0, x_0)$ равно минимальному значению критерия качества в задаче (68) — (70). Векторы $v(x_0, t_0)$, $u(x_0, t_0)$ — значения оптимальных управлений $v^0(s)$, $u^0(t)$ в момент $t = t_0$.

Значения оптимальных управлений в остальные моменты

$$u^0(t) = u(x^0(t), t), v^0(s) = v(\tilde{x}^0(s), s), s \in [t, t + 1 - \Delta],$$

где $x^0(t)$ — оптимальная траектория системы (68); $\tilde{x}^0(s)$ — оптимальная траектория системы (73) при $\tau = t$, $u = u^0(t)$, $y = x^0(t)$.

З а м е ч а н и е. В задаче данного пункта управление $v^0(s)$ находится из уравнения (73), которое по структуре несравненно проще исходного уравнения (68), поскольку линейно по вектору состояния.

§ 9. Задачи оптимизации типа Лагранжа

Задачи типа Лагранжа, как отмечалось в § 8, можно свести к задачам терминального управления, однако следует специально остановиться на этих чрезвычайно распространенных задачах, отличающихся вместе с тем рядом особенностей, свойственных только им.

1. Случай свободного правого конца. Рассмотрим задачу

$$x(t + 1) = f(x(t), u(t), t), t \in [t_0, t_1 - 1], x(t_0) = x_0; \quad (1)$$

$$I(u) = \sum_{t=t_0}^{t_1-1} f_0(x(t), u(t), t) \rightarrow \min u(t) \in U(t). \quad (2)$$

Процесс (1) погрузим в семейство

$$x(t+1) = f(x(t), u(t), t), \quad x(\tau) = y, \quad t \in [\tau, t_1 - 1], \\ t_0 \leq \tau \leq t_1 - 1,$$

и определим функцию Беллмана

$$B(\tau, y) = \min_{\substack{u \in U(\tau, t_1-1), \\ x(\tau)=y}} \sum_{t=\tau}^{t_1-1} f_0(x(t), u(t), t). \quad (3)$$

Поскольку

$$B(\tau, y) = \min_{u(\tau) \in U(\tau)} \min_{\substack{u \in U(\tau+1, t_1-1) \\ x(\tau+1)=f(y, u(\tau), \tau)}} [f_0(x(\tau), u(\tau), \tau) + \\ + \sum_{t=\tau+1}^{t_1-1} f_0(x(t), u(t), t)] = \min_{u(\tau) \in U(\tau)} [f_0(y, u(\tau), \tau) + \\ + \min_{\substack{u \in U(\tau+1, t_1-1) \\ x(\tau+1)=f(y, u(\tau), \tau)}} \sum_{t=\tau+1}^{t_1-1} f_0(x(t), u(t), t),$$

то из (3) получаем обратное уравнение Беллмана для задачи (1), (2):

$$B(\tau, y) = \min_{u \in U(\tau)} [f_0(y, u, \tau) + B(\tau+1, f(y, u, \tau))]. \quad (4)$$

Граничное условие следует из (3):

$$B(t_1 - 1, y) = \min f_0(y, u, t_1 - 1), \quad u \in U(t_1 - 1). \quad (5)$$

Вычислим оптимальное управление. Исходя из (5), через уравнение (4) находим последовательность функций

$$B(t_1 - 1, y), B(t_1 - 2, y), \dots, B(t_0, y).$$

Попутно определяем оптимальные законы управления

$$u(y, t_1 - 1), u(y, t_1 - 2), \dots, u(y, t_0)$$

из соотношений

$$f_0(y, u(y, t_1 - 1), t_1 - 1) = \min f_0(y, u, t_1 - 1), \\ u \in U(t_1 - 1);$$

$$\begin{aligned} & f_0(y, u(y, \tau), \tau) + B(\tau + 1, f(y, u(y, \tau), \tau)) = \\ & = \min \{f_0(y, u, \tau) + B(\tau + 1, f(y, u, \tau))\}, u \in U(\tau), \\ & \tau \in [t_0, t_1 - 2]. \end{aligned}$$

Число $B(t_0, x_0)$ равно минимальному значению критерия качества в задаче (1), (2). Вектор $u^0(t_0) = u(x_0, t_0)$ — значение оптимального управления $u^0(t)$ в момент t_0 . В остальные моменты оптимальное управление находим по формуле

$$u^0(t) = u(x^0(t), t), t \in [t_0 + 1, t_1 - 1], \quad (6)$$

где $x^0(t)$ — оптимальная траектория системы (1). В силу рекуррентной зависимости (1) формула (6) вместе с уравнением (1) позволяет последовательно вычислить оптимальное управление и оптимальную траекторию.

2. Прямое уравнение Беллмана. Несколько изменим задачу (1), (2), а именно: вместо условия на левый конец $x(t) = x_0$ наложим ограничение на правый конец

$$x(t_1) = x^1.$$

Погрузим процесс (1) в семейство (см. п. 2 § 5)

$$\begin{aligned} x(t+1) &= f(x(t), u(t), t), x(\tau) = y, t \in [t_0, \\ & \tau - 1], \tau \in [t_0 + 1, t_1], \end{aligned}$$

и введем функцию

$$B^-(\tau, y) = \min_{\substack{u [t_0, \tau-1], \\ x(\tau)=y}} \sum_{t=t_0}^{\tau-1} f_0(x(t), u(t), t). \quad (7)$$

Критерию качества типа Лагранжа характерна симметричность относительно начальных и конечных состояний процесса. Он удовлетворяет условиям (см. п. 6), при которых обратный принцип оптимальности справедлив в качестве и необходимого, и достаточного условия оптимальности.

Получим прямое уравнение Беллмана. Предположим, что при заданных u, t для каждого n -вектора y существует единственный n -вектор x такой, что $y = f(x, u, t)$. Иначе говоря, будем считать, что определен оператор

$$x = f^{-1}(y, u, t).$$

Из (7) имеем

$$\begin{aligned}
 B^-(\tau, y) &= \min_{u \in [t_0, \tau-1], x(\tau)=y} \sum_{t=t_0}^{\tau-1} f_0(x(t), u(t), t) = \\
 &= \min_{u(\tau-1) \in U(\tau-1)} \{f_0(x(\tau-1), u(\tau-1), \tau-1) + \\
 &+ \min_{\substack{u \in [t_0, \tau-2] \\ x(\tau-1)=f^{-1}(y, u(\tau-1), \tau-1)}} \sum_{t=t_0}^{\tau-2} f_0(x(t), u(t), t)\}. \quad (8)
 \end{aligned}$$

В фигурных скобках второе слагаемое равно, очевидно, величине $B^-(\tau-1, f^{-1}(y, u(\tau-1), \tau-1))$. Поэтому из (8) получаем прямое уравнение Беллмана для задачи (1), (2):

$$\begin{aligned}
 B^-(\tau, y) &= \min_{u \in U(\tau-1)} \{f_0(f^{-1}(y, u, \tau-1), u, \tau-1) + \\
 &+ B^-(\tau-1, f^{-1}(y, u, \tau-1))\}, \quad (9) \\
 &\tau \in [t_0 + 2, t_1].
 \end{aligned}$$

Граничное условие следует из (7):

$$B^-(t_0 + 1, y) = \min_{u \in U(t_0)} f_0(f^{-1}(y, u, t_0), u, t_0). \quad (10)$$

Из соотношений (9), (10) найдем решение задачи данного пункта.

Вычислив функцию $B^-(t_0 + 1, y)$ из (10), подставим ее в (9) и найдем последовательно функции

$$B^-(t_0 + 2, y), B^-(t_0 + 3, y), \dots, B^-(t_1, y).$$

Попутно вычислим и законы управления

$$u(y, t_0), u(y, t_0 + 1), \dots, u(y, t_1 - 1)$$

из соотношений

$$\begin{aligned}
 f_0(f^{-1}(y, u(y, t_0), t_0), u(y, t_0), t_0) &= \min_{u \in U(t_0)} f_0(f^{-1}(y, \\
 &u, t_0), u, t_0);
 \end{aligned}$$

$$\begin{aligned}
 &f_0(f^{-1}(y, u(y, \tau-1), \tau-1), u(y, \tau-1), \tau-1) + \\
 &+ B^-(\tau-1, f^{-1}(y, u(y, \tau-1), \tau-1))) = \min_{u \in U(\tau-1)} \{f_0(f^{-1}(y,
 \end{aligned}$$

$$u, \tau - 1), u, \tau - 1) + B^-(\tau - 1, f^{-1}(y, u, \tau - 1))\},$$

$$\tau \in [t_0 + 1, t_1 - 1].$$

Число $B^-(t_1, x^1)$ равно минимальному значению критерия качества (2). Вектор $u^0(t_1 - 1) = u(x^1, t_1 - 1)$ есть значение оптимального управления в момент $t = t_1 - 1$. По известным векторам $u^0(t_1 - 1)$ и x^1 однозначно найдется вектор $x^0(t_1 - 1) = f^{-1}(x^1, u^0(t_1 - 1), t_1 - 1)$.

Поэтому можно вычислить и $u^0(t_1 - 2) = u(x^0(t_1 - 1), t_1 - 2)$. Продолжая этот процесс, найдем все оптимальное управление $u^0(t)$ и оптимальную траекторию $x^0(t)$.

З а м е ч а н и е. Задачи данного и предыдущего пунктов симметричны относительно критерия качества и ограничений на управление. Однако уравнение процесса и ограничения на состояние в обоих случаях несимметричны. Если бы дискретный процесс в данном пункте описывался уравнением

$$x(t - 1) = f^-(x(t), u(t), t),$$

то прямое уравнение Беллмана заметно бы упростилось. В этом смысле можно сказать, что простота каждого из уравнений Беллмана (прямого или обратного) связана с той или иной несимметричностью задачи.

3. Случай подвижных концов. Рассмотрим задачу

$$x(t + 1) = f(x(t), u(t), t), t \in [t_0, t_1 - 1], u(t) \in U(t). \quad (11)$$

$$x(t_0) \in G_0, x(t_1) \in G_1, I(u) = \sum_{t=t_0}^{t_1-1} f_0(x(t), u(t), t) \rightarrow \min. \quad (12)$$

По сравнению с задачей п. 2 поставленная задача симметрична и относительно ограничений на состояние. Осуществим прямое погружение процесса (11) в семейство

$$x(t + 1) = f(x(t), u(t), t), x(\tau) = y, t \in [\tau, t_1 - 1].$$

Пусть

$$B(\tau, y) = \min_{\substack{u[\tau, t_1-1], \\ x(t_1) \in G_1}} x(\tau) = y \sum_{t=\tau}^{t_1-1} f_0(x(t), u(t), t). \quad (13)$$

Обозначим через $F(\tau)$ множество точек x , из которых допустимые траектории $x(t), x(\tau) = x$ попадают на множество G_1 . Пусть $\partial^0 F(\tau)$ — пограничный слой множества $F(\tau)$, состоящий из таких $x \in F(\tau)$, для которых найдется вектор $u \in U(\tau)$ такой, что $\dot{f}(x, u, \tau)$

$\bar{\tau}) \in F(\tau + 1)$. При $y \in \partial^0 F(\tau)$ множество векторов $u \in U(\tau)$, при которых $f(y, u, \tau) \in F(\tau + 1)$, обозначим через $U_1(y, \tau)$.

Если $x \in F(\tau)$, $x \in \partial^0 F(\tau)$, то, очевидно:

$$B(\tau, y) = \min_{u(\tau) \in U(\tau)} \min_{\substack{u[\tau+1, t_1-1] \\ x(\tau+1)=f(y, u(\tau), \tau), x(t_1) \in G_1}} [f_0(y, u(\tau), \tau) + \\ + \sum_{t=\tau+1}^{t_1-1} f_0(x(t), u(t), t)] = \min_{u \in U(\tau)} [f_0(y, u, \tau) + \\ + B(\tau + 1, f(y, u, \tau))].$$

При $y \in \partial^0 F(\tau)$ из (13) имеем

$$B(\tau, y) = \min_{u(\tau) \in U_1(y, \tau)} \min_{\substack{u[\tau+1, t_1-1] \\ x(\tau+1)=f(y, u(\tau), \tau), x(t_1) \in G_1}} [f_0(y, u(\tau), \tau) + \\ + \sum_{t=\tau+1}^{t_1-1} f_0(x(t), u(t), t)] = \min_{u \in U_1(y, \tau)} [f_0(y, u, \tau) + \\ + B(\tau + 1, f(y, u, \tau))].$$

Граничное условие получаем из (13):

$$B(t_1 - 1, y) = \min_{u \in U(t_1 - 1), f(y, u, t_1 - 1) \in G_1} f_0(y, u, t_1 - 1).$$

Таким образом, уравнение Беллмана имеет вид

$$B(\tau, y) = \begin{cases} \min_{u \in U(\tau)} [f_0(y, u, \tau) + B(\tau + 1, f(y, u, \tau))], & \text{если} \\ & y \in F(\tau), y \in \partial^0 F(\tau); \\ \min_{u \in U_1(y, \tau)} [f_0(y, u, \tau) + B(\tau + 1, f(y, u, \tau))], & \text{если} \\ & y \in \partial^0 F(\tau) \end{cases} \quad (14)$$

и ему соответствует граничное условие

$$B(t_1 - 1, y) = \min_{u \in U(t_1 - 1), f(y, u, t_1 - 1) \in G_1} f_0(y, u, t_1 - 1). \quad (15)$$

Множества $F(\tau)$, участвующие в (14), строятся индуктивно: $F(t_1) = G_1$, $F(\tau) = \{y, f(y, u, \tau) \in F(\tau + 1) \text{ при некотором } u \in U(\tau)\}$.

Из уравнения Беллмана решение задачи (11), (12) получается следующим образом. Исходя из (15) через уравнение (14) вычислим функции

$$B(t_1 - 1, y), B(t_1 - 2, y), \dots, B(t_0, y)$$

и оптимальные законы управления

$$u(y, t_1 - 1), u(y, t_1 - 2), \dots, u(y, t_0),$$

для чего фиксируем векторы u , на которых в (14) и (15) достигается минимум.

Находим вектор x_0^0 из условия

$$B(t_0, x_0^0) = \min B(t_0, y), y \in G_0.$$

Ясно, что x_0^0 — левый конец оптимальной траектории $x^0(t_0) = x_0^0$. По смыслу закона $u(y, t_0)$, вектор $u(x_0^0, t_0)$ — значение оптимального управления $u^0(t)$ в момент $t = t_0$. Далее полагаем $u^0(t) = u^0(x^0(t), t)$, $t \in [t_0 + 1, t_1 - 1]$, и с помощью уравнения (11) полностью строим оптимальное управление и оптимальную траекторию.

З а м е ч а н и е. В силу симметричности ограничений на крайние значения $x(t_0)$, $x(t_1)$ процесс (11) допускает естественное обратное погружение в семейство

$$x(t + 1) = f(x(t), u(t), t), x(\tau) = y, t \in [t_0, \tau - 1].$$

Схема получения прямого управления Беллмана ничем не отличается от аналогичной схемы, описанной в п. 2. Изменения касаются только вычисления правого конца оптимальной траектории.

4. Симметричная задача оптимизации. Рассмотрим процесс $x(t)$, $t \in [t_0, t_1]$, развитие которого в положительном направлении оси времени идет согласно уравнению

$$x(t + 1) = f(x(t), u(t), t), t \geq t_0, \quad (16)$$

а развитие в обратном направлении по уравнению

$$x(t - 1) = f^-(x(t), u(t), t), t \leq t_1. \quad (17)$$

Предположим, что r -мерные управления $u(t)$ выбираются из непустого множества $U(x, t)$, если рассматривается модель (16), или из множества $U^-(x, t)$ в случае (17): $u(t) \in U(x, t)$, $u(t) \in U^-(x, t)$.

Пусть, кроме того, наложены ограничения на состояния: $x(t) \in G(t)$, $t \in [t_0, t_1]$.

При этих условиях требуется минимизировать критерий качества

$$I = \sum_{t=t_0}^{t_1} f_0(x(t), u(t), t). \quad (18)$$

Решение задачи получим с помощью прямого и обратного уравнений Беллмана.

Обратное уравнение Беллмана связано с прямым погружением. Поэтому рассмотрим процесс (16) и погружим его в семейство

$$x(t+1) = f(x(t), u(t), t), \quad x(\tau) = y, \quad t \in [\tau, t_1], \\ \tau \in [t_0, t_1].$$

Определим функцию Беллмана равенством

$$B(\tau, y) = \min \sum_{t=\tau}^{t_1} f_0(x(t), u(t), t), \quad (19)$$

где минимум берется по всем допустимым управлениям $u[\tau, t_1]$, порождающим допустимые траектории $x(t)$, $t \in [\tau, t_1 + 1]$.

Пусть $F(\tau)$ — множество управляемости в рассматриваемой задаче, т. е. множество векторов y , при которых задача (19) имеет решение. Через $\partial^0 F(\tau)$ обозначим пограничный слой множества $F(\tau)$ (см. определение в п. 3), для каждого вектора $y \in F(\tau)$ выделим из $U(y, \tau)$ подмножество $U_1(y, \tau)$, отнеся к последнему такие векторы $u \in U(y, \tau)$, что $f(y, u, \tau) \in F(\tau + 1)$.

Для функции (19) при $y \in F(\tau)$, $y \in \partial^0 F(\tau)$ имеем

$$B(\tau, y) = \min_{u[\tau, t_1]} \sum_{t=\tau}^{t_1} f_0(x(t), u(t), t) = \min_{u \in U(y, \tau)} \{f_0(y, u, \tau) + \\ + B(\tau + 1, f(y, u, \tau))\}. \quad (20)$$

Если $y \in \partial^0 F(\tau)$, то обычным путем получаем

$$B(\tau, y) = \min_{u[\tau, t_1]} \sum_{t=\tau}^{t_1} f_0(x(t), u(t), t) = \min_{u(\tau) \in U_1(y, \tau)} \{f_0(y, \\ u, \tau) + B(\tau + 1, f(y, u, \tau))\}. \quad (21)$$

Далее из (19) при $y \in G(t_1)$ следует

$$B(t_1, y) = \min_{u \in U(y, t_1)} f_0(y, u, t_1). \quad (22)$$

Таким образом, в силу (20) — (22) для задачи данного пункта обратное управление Беллмана и граничное условие имеют вид

$$B(\tau, y) = \begin{cases} \min_{u \in U(y, \tau)} \{f_0(y, u, \tau) + B(\tau + 1, f(y, u, \tau))\}, \\ y \in F(\tau), y \in \bar{\partial}^0 F(\tau); \\ \min_{u \in U_1(y, \tau)} \{f_0(y, u, \tau) + B(\tau + 1, f(y, u, \tau))\}, \\ y \in \partial^0 F(\tau); \end{cases} \quad (23)$$

$$B(t_1, y) = \min_{u \in U(y, t_1)} f_0(y, u, t_1), y \in G(t_1). \quad (24)$$

Множества $F(\tau)$, $\tau \in [t_0, t_1]$ строятся по формулам $F(t_1) = G(t_1)$, $F(\tau) = \{y : f(y, u, \tau) \in F(\tau + 1) \text{ при некотором } u \in U(y, \tau)\}$.

Решение задачи оптимизации из (23), (24) получаем таким образом. При минимизации правых частей (23), (24) последовательно находим скалярные

$$B(t_1, y), B(t_1 - 1, y), \dots, B(t_0, y)$$

и векторные

$$u(y, t_1), u(y, t_1 - 1), \dots, u(y, t_0),$$

функции. Вектор x_0^0 , удовлетворяющий соотношению

$$B(t_0, x_0^0) = \min B(t_0, y), y \in G(t_0),$$

есть начальный вектор оптимальной траектории $x^0(t)$, $x^0(t_0) = x_0^0$; $B(t_0, x_0^0)$ — минимальное значение критерия качества; $u^0(t) = u(x^0(t), t)$, $t \in [t_0, t_1]$, — оптимальное управление.

Прямое уравнение Беллмана. Для описания процесса $x(t)$ используем уравнение (17). Погрузим процесс (17) в семейство

$$\begin{aligned} x(t - 1) &= f^-(x(t), u(t), t), x(\tau) = y, t \in [t_0, \tau], \\ &\tau \in [t_0, t_1] \end{aligned} \quad (25)$$

и введем функцию Беллмана

$$B^-(\tau, y) = \min_{u \in [t_0, \tau]} \sum_{t=t_0}^{\tau} f_0(x(t), u(t), t), \quad (26)$$

как минимальное значение критерия качества на допустимых управлениях семейства (25).

Обозначим через $F^-(\tau)$ множество достижимости в рассматриваемой задаче, отнеся к нему все векторы y , при которых задача (26) имеет решение. Пусть $\partial^0 F^-(\tau)$ — пограничный слой множества $F^-(\tau)$: $y \in \partial^0 F^-(\tau)$, если $y \in F^-(\tau)$, и существует такой вектор

$v \in U^-(y, \tau)$, что $f^-(y, v, \tau) \in F^-(\tau - 1)$.

В множестве $U^-(y, \tau)$ выделим подмножество $U_1^-(y, \tau)$ векторов u таких, что $f^-(y, u, \tau) \in F^-(\tau - 1)$.

Если $y \in F^-(\tau)$, $y \in \partial^0 F^-(\tau)$, $\tau \in [t_0 + 1, t_1]$, то

$$\begin{aligned} B^-(\tau, y) &= \min_{u \in [t_0, \tau]} \sum_{t=t_0}^{\tau} f_0(x(t), u(t), t) = \\ &= \min_{u \in U^-(y, \tau)} \{f_0(y, u, \tau) + B^-(\tau - 1, f^-(y, u, \tau))\}. \end{aligned} \quad (27)$$

Если $y \in \partial^0 F^-(\tau)$, $\tau \in [t_0 + 1, t_1]$, то, очевидно,

$$\begin{aligned} B^-(\tau, y) &= \min_{u(\tau) \in U_1^-(y, \tau)} \min_{\substack{u \in [t_0, \tau-1] \\ x(\tau-1) = f^-(y, u(\tau), \tau)}} \{f_0(y, u, (\tau), \tau) + \\ &+ \sum_{t=t_0}^{\tau-1} f_0(x(t), u(t), t)\} = \min_{u \in U_1^-(y, \tau)} \{f_0(y, u, \tau) + \\ &+ B^-(\tau - 1, f^-(y, u, \tau))\}. \end{aligned} \quad (28)$$

При $\tau = t_0$, $y \in G(t_0)$ из (26) получаем

$$B^-(t_0, y) = \min_{u \in U^-(y, t_0)} f_0(y, u, t_0). \quad (29)$$

Таким образом, в силу (27) — (29) для исследуемой задачи оптимизации справедливо прямое уравнение Беллмана

$$B^-(\tau, y) = \begin{cases} \min_{u \in U^-(y, \tau)} \{f_0(y, u, \tau) + B^-(\tau - 1, f^-(y, u, \tau))\}, & y \in F^-(\tau), y \in \partial^0 F^-(\tau); \\ \min_{\substack{u \in U_1^-(y, \tau) \\ y \in \partial^0 F^-(\tau), \tau \in [t_0 + 1, t_1]}} \{f_0(y, u, \tau) + B^-(\tau - 1, f^-(y, u, \tau))\}, & \end{cases} \quad (30)$$

с граничным условием

$$B^-(t_0, y) = \min f_0(y, u, t_0), u \in U^-(y, t_0), y \in G(t_0). \quad (31)$$

Полученное уравнение и граничное условие симметричны с обратным уравнением Беллмана (23) и граничным условием (24). Множества $F^-(\tau)$, участвующие в записи уравнения (30), строятся по формулам

$$F^-(t_0) = G(t_0), F^-(\tau) = \{y : f^-(y, u, \tau) \in F^-(\tau - 1)\}$$

при некотором $u \in U^-(y, \tau)$, $\tau \in [t_0 + 1, t_1]$.

Процедура решения исходной задачи оптимизации такова.

Из (31) вычислим вектор $u(y, t_0)$, $y \in G(t_0)$:

$$f_0(y, u(y, t_0), t_0) = \min_{u \in U^-(y, t_0)} f_0(y, u, t_0).$$

Используя вычисленную левую часть из (31), по формуле (30) находим функции

$$B^-(t_0 + 1, y), B^-(t_0 + 2, y), \dots, B^-(t_1, y)$$

и законы управления

$$u(y, t_0 + 1), u(y, t_0 + 2), \dots, u(y, t_1), \quad (32)$$

как векторы, на которых правые части в (30) достигают минимума.

По известной функции $B^-(t_1, y)$ вычисляем $x^0(t_1)$:

$$B^-(t_1, x^0(t_1)) = \min_{y \in G(t_1)} B^-(t_1, y).$$

Вектор $x^0(t_1)$ есть правый конец оптимальной траектории. По нему из (32) определяем вектор $u(x^0(t_1), t_1)$, равный значению оптимального управления $u^0(t)$ в момент $t = t_1$. Далее из уравнения (17) находим вектор $x^0(t_1 - 1) = f^-(x^0(t_1), u^0(t_1), t_1)$, а по нему из (32) вычисляем вектор $u^0(t_1 - 1) = u(x^0(t_1 - 1), t_1 - 1)$. Процесс продолжается до тех пор, пока не будет найдено все оптимальное управление $u^0(t)$, $t \in [t_0, t_1]$.

Минимальное значение критерия качества (18) равно $B^-(t_1, x^0(t_1))$.

§ 10. Минимизация максимального уклонения

Задачи на $\min \max$, к которым относится распространенная в приложениях задача минимизации максимального уклонения, существенно отличается по структуре критерия от задач, рассмотренных в § 8—9. Однако использование принципа оптимальности в его достаточном аспекте позволяет решить и эти трудные задачи.

1. Постановка задачи. Рассмотрим n -мерный дискретный процесс $x(t)$, заданный уравнением

$$x(t+1) = f(x(t), u(t), t), \quad t \in [t_0, t_1 - 1], \quad (1)$$

и r -мерными управлениями $u(t)$, принимающими значение из заданного множества $U(t)$:

$$u(t) \in U(t), \quad t \in [t_0, t_1 - 1]. \quad (2)$$

Качество процесса $x(t)$ (а значит, и управления $u(t)$) оценим критерием

$$I = \max_{t \in [t_0, t_1]} \varphi(x(t), t), \quad (3)$$

где $\varphi(x, t)$ — известная скалярная функция.

Требуется среди допустимых управлений (последовательностей $u(t_0), u(t_0 + 1), \dots, u(t_1 - 1)$, удовлетворяющих ограничению (2)) найти такое $u^0(t)$, $t \in [t_0, t_1 - 1]$, на котором критерий качества (3) достигает минимума:

$$\max_{t \in [t_0, t_1]} \varphi(x^0(t), t) = \min_{u[t_0, t_1 - 1]} \max_{t \in [t_0, t_1]} \varphi(x(t), t).$$

Здесь $x^0(t)$, $t \in [t_0, t_1]$ — оптимальная траектория, порожденная управлением $u^0(t)$; $x(t)$, $t \in [t_0, t_1]$ — траектория, соответствующая в силу (1) допустимому управлению $u[t_0, t_1 - 1]$.

Критерий качества (3), как и критерий качества в задаче типа Лагранжа, симметричен.

Критерий качества (3) (см. п. 6) по своей структуре близок к критериям качества в задачах типа Майера, Лагранжа, Больца, но отличается от последних по существу, в силу чего для (3) принцип оптимальности как необходимое условие оптимальности, вообще говоря, не-

справедлив. Упомянутое сходство критериев состоит в том, что для всех их выполняется равенство

$$I(t_0, x_0, u[t_0, t_1 - 1]) = v(t_0, x_0, u[t_0, \tau - 1]), \\ I(\tau, x(\tau), u[\tau, t_1 - 1])$$

при некоторой (для каждого критерия своей) функции $v(\alpha, \beta, \gamma, \delta)$. Отличается же критерий (3) от остальных тем, что соответствующая ему функция v не является строго монотонно возрастающей по четвертому аргументу. Однако функция $v(\alpha, \beta, \gamma, \delta)$ в случае (3) монотонно зависит от δ . Действительно, имеем:

$$I(t_0, x_0, u[t_0, t_1 - 1]) = \max_{t \in [t_0, t_1]} \varphi(x(t), t) = \\ = \max_{t \in [t_0, \tau]} \{\max_{t \in [t_0, \tau]} \varphi(x(\tau), t), \max_{t \in [\tau, t_1]} \varphi(x(t), t)\} = \\ = \max\{I(t_0, x_0, u[t_0, \tau - 1]), I(\tau, x(\tau), u[\tau, t_1 - 1])\} \leq \\ \leq \max\{I(t_0, x_0, u[t_0, \tau - 1]), I(\tau, x(\tau), \bar{u}[\tau, t_1 - 1])\}, \quad (4) \\ \text{если} \quad \delta = I(\tau, x(\tau), u[\tau, t_1 - 1]) \leq I(\tau, x(\tau), \bar{u}[\tau, \\ t_1 - 1]) = \bar{\delta}.$$

Таким образом, все условия п. 6 выполнены, и поэтому можно воспользоваться прямым принципом оптимальности как достаточным условием оптимальности.

2. Обратное уравнение Беллмана. Осуществим прямое погружение процесса (1) в семейство

$$x(t+1) = f(x(t), u(t), t), \quad x(\tau) = y, \quad t \in [\tau, t_1 - 1] \quad (5)$$

и введем функцию Беллмана

$$B(\tau, y) = \min_{u[\tau, t_1 - 1]} \max_{t \in [\tau, t_1]} \varphi(x(t), t), \quad (6)$$

определенную на всех допустимых траекториях семейства (5).

Схема применения прямого принципа оптимальности как достаточного условия оптимальности состоит в следующем. Промежуток $[\tau, t_1 - 1]$, $\tau < t_1 - 1$, разбиваем на два периода τ и $[\tau + 1, t_1 - 1]$. В первом периоде выбираем некоторое допустимое управление $u(\tau) \in U(\tau)$.

Оно переводит процесс (5) из состояния $x(\tau) = y$ в состояние $x(\tau + 1) = f(y, u(\tau), \tau)$. Далее, управление $u(t)$ во втором периоде $[\tau + 1, t_1 - 1]$ выбираем оптимальным относительно состояния $x(\tau + 1)$, т. е. из условия

$$\begin{aligned} I(\tau + 1, x(\tau + 1), u[x(\tau + 1), \tau + 1, t_1 - 1]) &= \\ &= \min I(\tau + 1, x(\tau + 1), u[\tau + 1, t_1 - 1]) = \\ &= B(\tau + 1, x(\tau + 1)). \end{aligned}$$

Значение критерия качества (3) на всем допустимом управлении $u[\tau, t_1 - 1]$, очевидно, равно

$$\begin{aligned} I(\tau, x(\tau), u[\tau, t_1 - 1]) &= \max_{t \in [\tau, t_1 - 1]} \varphi(x(t), t) = \\ &= \max \{ \varphi(x(\tau), \tau), \max_{t \in [\tau + 1, t_1 - 1]} \varphi(x(t), t) \} = \max \{ \varphi(x(\tau), \tau), \\ &B(\tau + 1, f(y, u(\tau), \tau)) \}. \end{aligned} \quad (7)$$

Принцип оптимальности гласит (см. п. 6 § 6): если допустимое управление $u(\tau) \in U(\tau)$ в первом периоде выбрать так, чтобы величина (7) была минимальной, то управление $u[\tau, t_1 - 1]$ будет оптимальным, т. е.

$$\begin{aligned} I(\tau, x(\tau), u[\tau, t_1 - 1]) &= \min_{\bar{u}[\tau, t_1 - 1]} I(\tau, x(\tau), \bar{u}[\tau, t_1 - 1]) = \\ &= B(\tau, x(\tau)). \end{aligned} \quad (8)$$

Проведя в (7) указанную операцию минимизации по $u(\tau)$ и используя (8), имеем обратное уравнение Белмана:

$$\begin{aligned} B(\tau, y) &= \min_{u \in U(\tau)} \max \{ \varphi(y, \tau), B(\tau + 1, f(y, u, \tau)) \}, \\ &\tau \in [t_0, t_1 - 2]. \end{aligned} \quad (9)$$

Граничное условие получим из (6):

$$\begin{aligned} B(t_1 - 1, y) &= \max \{ \varphi(y, t_1 - 1), \min_{u \in U(t_1 - 1)} \varphi(f(y, u, \\ &t_1 - 1), t_1) \}. \end{aligned} \quad (10)$$

Для вычисления решения задачи (1) — (3) поступаем по обычным правилам. Из (9), (10) находим функции

$$B(t_1 - 1, y), B(t_1 - 2, y), \dots, B(t_0, y)$$

и оптимальные законы управления

$$u(y, t_1 - 1), u(y, t_1 - 2), \dots, u(y, t_0).$$

Определяем x_0^0 из соотношения

$$B(t_0, x_0^0) = \min_{y \in R_n} B(t_0, y).$$

Ясно, что x_0^0 — левый конец оптимальной траектории; число $B(t_0, x_0^0)$ равно минимальному значению критерия качества (3); $u(x_0^0, t_0)$ — оптимальное управление $u^0(t)$ в момент $t = t_0$. В остальные моменты выполняется равенство $u^0(t) = u(x^0(t), t)$, $t \in [t_0 + 1, t_1 - 1]$.

3. Прямое уравнение Беллмана. Как следует из (4) (с очевидными изменениями), критерий качества (3) удовлетворяет условиям, при которых справедлив обратный принцип оптимальности как достаточное условие оптимальности (см. п. 6 § 8). Поэтому, используя обратное погружение, можно получить для задачи (1) — (3) прямое уравнение Беллмана. Ради упрощения записей вместо уравнения (1), как и в § 9, используем эквивалентное уравнение

$$x(t-1) = f^-(x(t), u(t), t), \quad t \in [t_0 + 1, t_1], \quad u(t) \in U^-(t), \quad (11)$$

которое устанавливает закон развития процесса $x(t)$ в направлении убывания времени. Введем семейство процессов (обратное погружение):

$$x(t-1) = f^-(x(t), u(t), t), \quad x(\tau) = y, \quad t \in [t_0 + 1, \tau]$$

и на его элементах определим функцию Беллмана

$$B^-(\tau, y) = \min_{u[t_0+1, \tau]} \max_{t \in [t_0, \tau], x(\tau)=y} \varphi(x(t), t). \quad (12)$$

Следуя обратному принципу оптимальности (п. 6), выберем некоторое допустимое управление $u(\tau) \in U^-(\tau)$ во втором периоде, состоящем из одного момента τ . Оно переведет процесс $x(t)$ из состояния $x(\tau) = y$ в состояние $x(\tau-1) = f^-(y, u(\tau), \tau)$. Управление в первом периоде $[t_0 + 1, \tau - 1]$ выберем оптимальным относительно состояния $x(\tau-1)$:

$$I(u[t_0 + 1, \tau - 1], x(\tau - 1)) = \min I(u[t_0 + 1, \tau - 1], x(\tau - 1)) = B^-(\tau - 1, x(\tau - 1)).$$

Тогда значение критерия качества вдоль построенного управления $u[t_0+1, \tau]$ можно записать в виде

$$I(u[t_0+1, \tau], x(\tau)) = \max \{ \varphi(x(\tau), \tau), \max_{t \in [t_0, \tau-1]} \varphi(x(t), t) \} = \\ = \max \{ \varphi(x(\tau), \tau), B^-(\tau-1, x(\tau-1)) \}. \quad (13)$$

Если теперь вектор $u(\tau) \in U^-(\tau)$ выбрать так, чтобы величина (13) стала минимальной, то, согласно обратному принципу оптимальности, управление $\{u[t_0+1, \tau-1], u(\tau)\}$ будет оптимальным, т. е. соотношение (13) принимает вид

$$B^-(\tau, y) = \min_{u \in U^-(\tau)} \max \{ \varphi(y, \tau), B^-(\tau-1, f^-(y, u, \tau)) \}, \\ \tau \in [t_0+2, t_1]. \quad (14)$$

Уравнение (14) и есть прямое уравнение Беллмана в задаче (11), (2), (3). Граничное условие для (14) получаем из (12):

$$B^-(t_0+1, y) = \max \{ \varphi(y, t_0+1), \min_{u \in U^-(t_0+1)} \varphi(f^-(y, u, t_0+1), t_0) \}. \quad (15)$$

Процедура вычисления оптимального управления в задаче (11), (2), (3) такова. Из (15), (14) находим функции

$$B^-(t_0+1, y), B^-(t_0+2, y), \dots, B^-(t_1, y)$$

и одновременно оптимальные законы управления

$$u(y, t_0+1), u(y, t_0+2), \dots, u(y, t_1). \quad (16)$$

Вычисляем вектор $x^0(t_1)$ из условия

$$B^-(t_1, x^0(t_1)) = \min_{y \in R_n} B^-(t_1, y).$$

Ясно, что $x^0(t_1)$ — правый конец оптимальной траектории. По нему из (11) и (16) легко восстановить оптимальное управление $u^0(t) = u(x^0(t), t)$, $t \in [t_0+1, t_1]$.

§ 11. Линейные системы с квадратичными критериями качества

До сих пор для решения уравнений Беллмана использовались рекуррентные процедуры, которые требовали табулирования функций многих переменных. Рассмотрим

рим задачу, в которой этих трудностей можно избежать и получить решение в явном виде.

1. Задача аналитического конструирования регулятора [15, 22]. Рассмотрим задачу

$$x(t+1) = A(t)x(t) + B(t)u(t), \quad x(t_0) = x_0, \\ t \in [t_0, t_1 - 1]; \quad (1)$$

$$I(u) = \sum_{t=t_0}^{t_1-1} [x'(t)M(t)x(t) + u'(t)R(t)u(t)] \rightarrow \min, \quad (2)$$

где * $M(t) \geq 0$, $R(t) > 0$ — симметрические матрицы, $A(t)$, $B(t)$ имеют размеры, соответствующие n -вектору x и r -вектору u . Ограничения на управление и состояние не накладываются.

Обратное уравнение Беллмана и соответствующее граничное условие (см. § 9) в данном случае имеют вид

$$B(\tau, y) = \min_v [y' M(\tau) y + v' R(\tau) v + B(\tau + 1, \\ A(\tau) y + B(\tau) v)]; \quad (3)$$

$$B(t_1 - 1, y) = y' M(t_1 - 1) y. \quad (4)$$

Специфика рассматриваемой задачи позволяет решение уравнения (3) свести к простым операциям.

Будем искать функцию $B(\tau, y)$ в форме

$$B(\tau, y) = y' L(\tau) y \quad (5)$$

с неизвестной матрицей $L(\tau)$. Из граничного условия (4) получаем

$$y' L(t_1 - 1) y = y' M(t_1 - 1) y.$$

Поскольку y — произвольный n -вектор,

$$L(t_1 - 1) = M(t_1 - 1). \quad (6)$$

Подставим функцию (5) в уравнение (3):

$$y' L(\tau) y = \min_v [y' M(\tau) y + v' R(\tau) v + [A(\tau) y + B(\tau) v]' \times \\ \times L(\tau + 1) [A(\tau) y + B(\tau) v]]. \quad (7)$$

* Напомним: матрица M неотрицательна: $M \geq 0$, если $x' M x \geq 0$ для всех $x \in R_n$; пишут $R > 0$, если $x' R x > 0$ при $x \neq 0$.

Минимум в квадратной скобке достигается, очевидно, в точке $u(y, \tau)$, доставляющей минимум выражению

$$\begin{aligned} \Phi(v, \tau) = & v' R(\tau) v + y' A'(\tau) L(\tau + 1) B(\tau) v + \\ & + v' B'(\tau) L(\tau + 1) A(\tau) y + v' B'(\tau) L(\tau + 1) B(\tau) v. \end{aligned} \quad (8)$$

Если $L(\tau) \geq 0$ (это доказывается ниже), то в силу того, что $R(\tau) > 0$, выражение (8) представляет строго выпуклую функцию относительно v с единственной экстремальной точкой $u(y, \tau)$. Поэтому для вычисления $u(y, \tau)$ достаточно решить уравнение

$$\frac{\partial \Phi(u, (y, \tau), \tau)}{\partial v} = 0,$$

которое линейно и имеет вид

$$\begin{aligned} 2[R(\tau) + B'(\tau) L(\tau + 1) B(\tau)] u(y, \tau) + \\ + 2B'(\tau) L(\tau + 1) A(\tau) y = 0. \end{aligned}$$

Отсюда, поскольку матрица коэффициентов при $u(y, \tau)$ неособая, получаем

$$\begin{aligned} u(y, \tau) = -[R(\tau) + B'(\tau) L(\tau + 1) B(\tau)]^{-1} B'(\tau) L(\tau + \\ + 1) A(\tau) y. \end{aligned} \quad (9)$$

Подставляя это значение в (7) и приравнявая коэффициенты квадратичных форм (y — произвольный вектор!), получаем уравнение для $L(\tau)$:

$$\begin{aligned} L(\tau) = & M(\tau) + [R(\tau) + B'(\tau) L(\tau + 1) B(\tau)]^{-1} B'(\tau) L(\tau + \\ & + 1) A(\tau)]' R(\tau) [R(\tau) + B'(\tau) L(\tau + 1) B(\tau)]^{-1} B'(\tau) L(\tau + \\ & + 1) A(\tau)] + [A(\tau)] - B(\tau) [R(\tau) + B'(\tau) L(\tau + 1) \times \\ & \times B(\tau)]^{-1} B'(\tau) L(\tau + 1) A(\tau)]' L(\tau + 1) [A(\tau) - \\ & - B(\tau) [R(\tau) + B'(\tau) L(\tau + 1) B(\tau)]^{-1} B'(\tau) L(\tau + 1) A(\tau)]. \end{aligned} \quad (10)$$

Уравнение (10) — рекуррентное, по структуре напоминает известное дифференциальное уравнение Риккати и поэтому называется рекуррентным уравнением Риккати. При начальном условии (6) уравнение (10) имеет единственное решение.

Таким образом, решение уравнения Беллмана относительно скалярной функции $B(\tau, y)$ от $n+1$ переменных сведено к более простой задаче решения уравнения Риккати относительно $n \times n$ -матричной функции $L(\tau)$ от одной переменной. Оптимальный закон управления задается формулой (9) и представляет линейную форму по состоянию. Минимальное значение критерия качества (2) равно $x_0' L(t_0) x_0$.

Для полного обоснования метода решения нужно доказать, что в силу сделанных предположений матрица $L(\tau)$ неотрицательна.

При $\tau = t_1 - 1$ этот вывод следует из (6). Докажем по индукции, что если $L(\tau + 1) \geq 0$, то и $L(\tau) \geq 0$. Умножая обе части уравнения (10) слева на y' , справа на y , получаем, что справа стоит определенно-положительная форма. Значит, $L(\tau) > 0$. Утверждение доказано.

2. Обобщение. Добавим в критерий (2) терминальный член $\varphi(x(t_1)) = x' K(t_1) x$ и рассмотрим задачу минимизации критерия

$$I(u) = x' K(t_1) x + \sum_{t=t_0}^{t_1-1} [x'(t) M(t) x(t) + u'(t) R(t) u(t)].$$

Если повторить рассуждения п. 1, нетрудно убедиться, что функция Беллмана в новой задаче имеет вид

$$B(\tau, y) = y' L(\tau) y,$$

где $n \times n$ — матричная функция $L(\tau)$ удовлетворяет уравнению Риккати (1) с начальным условием

$$L(t_1) = K(t_1).$$

3. Стационарная задача аналитического конструирования регуляторов. Рассмотрим частный случай задачи (1), (2):

$$x(t+1) = Ax(t) + Bu(t), \quad x(0) = x_0, \quad t \in [0, t_1 - 1]; \quad (11)$$

$$I(u) = \sum_{t=0}^{t_1-1} [x'(t) Mx(t) + u'(t) Ru(t)] \rightarrow \min, \quad (12)$$

в которой параметры, представленные матрицами A , B , M , R , не зависят от времени.

В этом случае функция Беллмана имеет вид

$$B(\tau, y) = y' L(\tau) y,$$

где $L(\tau)$ — решение уравнения

$$L(\tau) = M + [[R + B'L(\tau + 1)B]^{-1} B'L(\tau + 1)A]' R [[R + B'L(\tau + 1)B]^{-1} B'L(\tau + 1)A] + [A - B[R + B'L(\tau + 1)B]^{-1} B'L(\tau + 1)A]' L(\tau + 1) [A - B[R + B'L(\tau + 1)B]^{-1} B'L(\tau + 1)A]; \quad (13)$$

$$L(t_1 - 1) = M. \quad (14)$$

Минимальное значение критерия качества (12)

$$I(u^0) = x_0' L(0) x_0.$$

Пусть в задаче (11), (12) процессы продолжимы неограниченно вправо и момент $t_1 = \infty$. Рассмотрим стационарную задачу аналитического конструирования регуляторов:

$$x(t + 1) = Ax(t) + Bu(t), \quad x(0) = x_0, \quad t \geq 0; \quad (15)$$

$$I(u) = \sum_{t=0}^{\infty} [x'(t) Mx(t) + u'(t) Ru(t)] \rightarrow \min, \quad (16)$$

$$M > 0, \quad R > 0.$$

Управления $u(t)$, $t \geq 0$, при которых значения (16) ограничены, назовем допустимыми. Будем считать, что в задаче (15), (16) имеются допустимые управления.

Основываясь на особенностях задачи (15), (16) (матрицы A , B , M , R постоянны, $t_1 = \infty$), можно функцию Беллмана ввести более простым путем.

Положим

$$B(y) = \min_{u \in [0, \infty)} \sum_{t=0}^{\infty} [x'(t) Mx(t) + u'(t) Ru(t)], \quad (17)$$

где \min берется по всем допустимым управлениям вдоль траекторий, исходящих в момент $t=0$ из точки $x(0) = y$.

Ясно, что

$$B(y) = \min_{u(0)} [y' My + u'(0) Ru(0) + \min_{u \in [1, \infty)} \sum_{t=1}^{\infty} [x'(t) Mx(t) +$$

$$+ u'(t) Ru(t)] = \min_{u(0)} [y' My + u'(0) Ru(0) + \\ + \min_{u(1, \infty)} \sum_{t=0}^{\infty} [x'(t+1) Mx(t+1) + u'(t+1) Ru(t+1)]]]. \quad (18)$$

Введем новые переменные: $v(t) = u(t+1)$, $z(t) = x(t+1)$, которые в силу (15) удовлетворяют уравнению

$$\left. \begin{aligned} z(t+1) &= Az(t) + Bv(t), \quad t \geq 0; \\ z(0) &= x(1) = Ay + Bu(0). \end{aligned} \right\} \quad (19)$$

Второе выражение в квадратных скобках последнего члена (18) представляет собой задачу минимизации критерия (16) вдоль траекторий системы (19). Поэтому оно, по определению (17), равно $B(Ay + Bu(0))$.

Таким образом, из (18) следует уравнение Беллмана для задачи (15), (16):

$$B(y) = \min_v [y' My + v' Rv + B(Ay + Bv)]. \quad (20)$$

Полученное уравнение кажется более простым (ср., например, с (3)), так как здесь не присутствует параметр τ , однако решение его не сводится к элементарным операциям, соответствующим рекуррентному уравнению (3).

Решение уравнения (20) будем искать в виде

$$B(y) = y' Ny.$$

Подставив эту функцию в (20), получим

$$y' Ny = y' My + \min_v [v' Rv + (Ay + Bv)' N(Ay + Bv)]. \quad (21)$$

По определению функции Беллмана и свойствам задачи ($M > 0$, $R > 0$), матрица N должна быть положительной: $N > 0$. Поэтому при решении уравнения (21) будем искать лишь такие варианты, для которых $N > 0$. В этом случае минимум в (21) достигается на единственном элементе

$$u(y) = -(R + B'NB)^{-1} B'NAy, \quad (22)$$

который получается методом, описанным в п. 1.

Подставим значение (22) в (21), приравняем элементы квадратичных форм. В результате получим уравнение для матрицы N :

$$N = M + [(R + B'NB)^{-1} B'NA]' R [(R + B'NB)^{-1} B'NA] + \\ + [A - B(R + B'NB)^{-1} B'NA]' N [A - B(R + \\ + B'NB)^{-1} B'NA]. \quad (23)$$

Это уравнение нелинейно относительно N , и отыскание положительного ($N > 0$) решения — задача трудная. Если она решена, то $I(u^0) = x_0' N x_0$.

4. Использование предельного перехода [22]. Укажем способ решения задачи (15), (16) с помощью решения задачи (11), (12). Попутно получим решение для (23).

Предварительно «обратим» время в уравнении (13), решение которого $L(\tau)$ зависит, очевидно, от момента t_1 : $L(\tau) = L(\tau, t_1)$. Для этого введем матричную функцию

$$P(t, t_1) = L(t_1 - t, t_1). \quad (24)$$

Уравнения и начальные условия для $P(t, t_1)$ получаются из (13), (14):

$$P(t_1 - t, t_1) = M + [R + B'P(t_1 - (t + 1), t_1) B]^{-1} B'P(t_1 - (t + 1), t_1) A]' R [R + B' \times \\ \times P(t_1 - (t + 1), t_1) B]^{-1} B'P(t_1 - (t + 1), t_1) B] + \\ + [A - B[R + B'P(t_1 - (t + 1), t_1) B]^{-1} B'P(t_1 - \\ - (t + 1), t_1) A]' P(t_1 - (t + 1), t_1) [A - B[R + \\ + B'P(t_1 - (t + 1), t_1) B]^{-1} B' \times \\ \times P(t_1 - (t + 1), t_1) A], P(1, t_1) = M.$$

В силу (24) и результатов п. 3 минимальное значение критерия качества в задаче (11), (12): $I(u^0) = x_0' P(t_1, t_1) x_0$.

Связь между задачами (15), (16) и (11), (12) устанавливается следующим утверждением.

Теорема. Для того чтобы задача (15), (16) имела решение для любого x_0 , необходимо и достаточно, чтобы существовала такая постоянная матрица N , что

$$\lim_{t_1 \rightarrow +\infty} P(t_1, t_1) = N. \quad (25)$$

Доказательство. Необходимость. Пусть найден закон управления u^0 , который доставляет критерию качества $\sum_{t=0}^{\infty} \Phi(x, u)$, $\Phi(x, u) = x'(t) Mx(t) + u'(t) Ru(t)$, минимальное значение. Построим последовательность положительных чисел $t_1 < t_2 < \dots < t_n < \dots$ и последовательность квадратичных форм

$$\sum_{t=0}^{t_1} \Phi(x, u^1), \dots, \sum_{t=0}^{t_n} \Phi(x, u^n), \quad (26)$$

где u^i — оптимальный закон управления, доставляющий минимальное значение критерию $I(x_0, t_i) = \sum_{t=0}^{t_i} \Phi(x, u)$.

Эта последовательность является монотонно возрастающей

$$\sum_{t=0}^{t_i} \Phi(x, u^i) \leq \sum_{t=0}^{t_i} \Phi(x, u^{i+1}) \leq \sum_{t=0}^{t_{i+1}} \Phi(x, u^{i+1}).$$

Кроме того, последовательность (26) ограничена при каждом x_0 , так как

$$\sum_{t=0}^{t_i} \Phi(x, u^i) \leq \sum_{t=0}^{t_i} \Phi(x, u^0) \leq \sum_{t=0}^{\infty} \Phi(x, u^0) \text{ при любом } t_i.$$

Значит, последовательность квадратичных форм

$$x'_0 P(t_1, t_1) x_0, \dots, x'_0 P(t_n, t_n) x_0, \dots$$

сходится при любых x_0 к некоторой квадратичной форме $x'_0 N x_0$ с постоянными коэффициентами. Необходимость доказана.

Достаточность. Пусть выполнено условие (25). Значит,

$$\lim_{t_i \rightarrow +\infty} \sum_{t=0}^{t_i} \Phi(x, u^i) = \lim_{t_i \rightarrow +\infty} x'_0 P(t_i, t_i) x_0 = x'_0 N x_0 < +\infty.$$

Утверждается, что закон управления

$$u = \lim_{t_1 \rightarrow +\infty} [R + B'P(t_1 - (t + 1), t_1) B']^{-1} B'P(t_1 - (t + 1), t_1) Ax \quad (27)$$

оптимален для задачи (15), (16). Заметим, что в силу стационарности задачи выполняется соотношение

$$\lim_{t_1 \rightarrow +\infty} P(t_1 - (t + 1), t_1) = \lim_{t_1 \rightarrow +\infty} P(t_1, t_1).$$

Сначала покажем, что

$$\lim_{t_i \rightarrow +\infty} \sum_{t=0}^{t_i} \Phi(x, u^t) = \sum_{t=0}^{\infty} \Phi(x, u). \quad (28)$$

Действительно, для каждого фиксированного $s < t_1$ имеем:

$$\sum_{t=0}^{t_i} \Phi(x, u^t) \geq \sum_{t=0}^s \Phi(x, u^t).$$

При этом

$$\lim_{t_i \rightarrow +\infty} \sum_{t=0}^{t_i} \Phi(x, u^t) \geq \sum_{t=0}^s \Phi(x, u).$$

Переходя к пределу при $s \rightarrow +\infty$, получаем

$$\lim_{i \rightarrow +\infty} \sum_{t=0}^{t_i} \Phi(x, u^t) \geq \sum_{t=0}^{\infty} \Phi(x, u). \quad (29)$$

С другой стороны,

$$\sum_{t=0}^{t_i} \Phi(x, u^t) \leq \sum_{t=0}^{t_i} \Phi(x, u),$$

т. е.

$$\lim_{i \rightarrow +\infty} \sum_{t=0}^{t_i} \Phi(x, u^t) \leq \sum_{t=0}^{\infty} \Phi(x, u).$$

Теперь из (29) следует (28).

Допустим, что закон управления (27) не является оптимальным. Тогда существует управление \bar{u} , для которого

$$\sum_{t=0}^{+\infty} \Phi(x, \bar{u}) < \sum_{t=0}^{+\infty} \Phi(x, u), \quad (30)$$

но

$$\sum_{t=0}^{t_i} \Phi(x, \bar{u}) \geq \sum_{t=0}^{t_i} \Phi(x, u^i).$$

Следовательно,

$$\sum_{t=0}^{+\infty} \Phi(x, \bar{u}) \geq \sum_{t=0}^{+\infty} \Phi(x, u),$$

что противоречит (30). Значит, u — оптимальный закон управления. Теорема доказана.

§ 12. Приложение к задаче стабилизации

Результаты § 11 можно использовать для решения проблемы стабилизации дискретных процессов.

1. Стабилизация линейных систем. Рассмотрим систему

$$x(t+1) = Ax(t), \quad x — n\text{-вектор}, \quad (1)$$

с постоянной матрицей A , среди собственных чисел которой имеются числа с модулями, большими единицы. В этом случае система (1) неустойчива и найдутся начальные состояния x_0 , для которых $\|x(t)\| \rightarrow \infty$ при $t \rightarrow \infty$.

Допустим, что система (1) имеет входное устройство, через которое можно ею управлять, и общее уравнение системы и входного устройства таково:

$$x(t+1) = Ax(t) + Bu(t), \quad (2)$$

где $u(t)$ — r -мерное управление.

Проблема линейной стабилизации системы (1) состоит в указании такой $r \times n$ -матрицы Q , что управление $u(x) = Qx$ стабилизирует систему (2), т. е. все собствен-

ные числа матрицы $A+BQ$ по модулю будут меньше единицы.

Для решения проблемы стабилизации обратимся к задаче (15), (16) § 11. Пусть $R>0$, $M>0$ — заданные симметрические матрицы. Решим задачу (15), (16) с данными матрицами R и M . Согласно (22) § 11, решение имеет вид

$$u(x) = Qx, \quad (3)$$

где $Q = -(R + B'NB)^{-1} B'NA$.

Подставим функцию (3) в уравнение (2), и пусть $x(t)$ — решение этого уравнения с некоторым начальным условием. Исследуем поведение функции Беллмана вдоль $x(t)$:

$$\begin{aligned} B(x(t)) = & \sum_{\tau=t}^{\infty} [x'(\tau) Mx(\tau) + u'(\tau) Ru(\tau)] = x'(t) Mx(t) + \\ & + u'(t) Ru(t) + \sum_{\tau=t+1}^{\infty} [x'(\tau) Mx(\tau) + u'(\tau) Ru(\tau)]. \end{aligned}$$

После подстановки $u(t)=u(x(t))$ из (3) получаем

$$B(x(t+1)) - B(x(t)) = -x'(t) [M + Q'RQ] x(t).$$

По доказанному в § 11, функция $B(x)$ является определенно-положительной формой, матрица $M+Q'RQ$ определенно-положительна в силу того, что $R, M>0$. Следовательно, выполнены все условия дискретного аналога теоремы Ляпунова об асимптотической устойчивости [21] и поэтому $\|x(t)\| \rightarrow 0$, $t \rightarrow \infty$ для любых начальных возмущений.

Построенный закон управления (3) не только стабилизирует систему (1), но одновременно и доставляет минимальное значение функции

$$I(u) = \sum_{t=0}^{\infty} [x'(t) Mx(t) + u'(t) Ru(t)],$$

т. е. переходный процесс в системе (2) обладает определенными экстремальными свойствами. Выбирая те или иные матрицы \bar{R} и M , можно получить различные переходные процессы.

2. О стабилизации нелинейных систем. Обобщая проблему линейной стабилизации, задачу в общем случае можно сформулировать следующим образом. Пусть дана система

$$\begin{aligned}x(t+1) &= f(x(t), u(t), t), \quad t \geq 0, \\f(0, 0, t) &= 0,\end{aligned}\tag{4}$$

нулевое состояние $x=0$ которой неустойчиво при $u(t)=0$. Требуется найти закон управления $u = u(x, t)$, $u(0, t) = 0$, при котором нулевое состояние системы

$$x(t+1) = f(x(t), u(x(t), t), t), \quad t \geq 0\tag{5}$$

асимптотически устойчиво.

Задаче стабилизации системы (4) поставим в соответствие следующую задачу оптимизации:

$$x(t+1) = f(x(t), u(t), t), \quad x(0) = x_0;$$

$$I(u) = \sum_{t=0}^{\infty} f_0(x(t), u(t), t) \rightarrow \min,$$

где $f_0(x, u, t)$ — скалярная функция, определенно-положительная в смысле Ляпунова по аргументу x .

Считая, что задача оптимизации имеет решение для любых x_0 , используем прямое погружение системы (5) и введем функцию Беллмана:

$$B(\tau, y) = \min_{u \in [\tau, \infty)} \sum_{t=\tau}^{\infty} f_0(x(t), u(t), t).\tag{6}$$

Уравнение для $B(\tau, y)$ получается обычным путем (см. § 9) и имеет вид

$$B(\tau, y) = \min_v [f_0(y, v, \tau) + B(\tau+1, f(y, v, \tau))].\tag{7}$$

Пусть $B(\tau, y)$ — решение уравнения (7), удовлетворяющее условию для каждого y : $B(\tau, y) \rightarrow 0$ при $\tau \rightarrow \infty$. При каждом τ , $\tau \geq 0$, функция $B(\tau, y)$ в силу (6) положительна: $B(\tau, y) \geq C(y) > 0$, $y \neq 0$.

Подсчитаем поведение ее вдоль оптимальной траектории, порожденной управлением $u(x, t)$:

$$\begin{aligned}& f_0(x, u(x, t), t) + B(\tau+1, f(x, u(x, t), t)) = \\&= \min_v [f_0(x, v, t) + B(\tau+1, f(x, v, t))].\end{aligned}\tag{8}$$

Из (6) имеем

$$B(t, x(t)) = \sum_{\tau=t}^{\infty} f_0(x(\tau), u(x(\tau), \tau), \tau) = \sum_{\tau=t+1}^{\infty} f_0(x(\tau), \\ u(x(\tau), \tau), \tau) + f_0(x(t), u(x(t), t), t) = B(t+1, \\ x(t+1)) + f_0(x(t), u(x(t), t), t).$$

Таким образом, полное приращение функции $B(t, x)$,

$$B(t+1, x(t+1)) - B(t, x(t)) = -f_0(x(t), u(x(t), t), t),$$

вдоль продолжимых вправо траекторий системы

$$x(t+1) = g(x(t), t), \quad g(x, t) = f(x, u(x, t), t) \quad (9)$$

есть функция определенно-отрицательная. Отсюда заключаем (в силу дискретного аналога теоремы Ляпунова об асимптотической устойчивости [21], что управление $u = u(x, t)$, вычисленное в силу (8), стабилизирует систему (5) до асимптотически устойчивой системы (9).

§ 13. Применение динамического программирования для решения некоторых комбинаторных задач

Как указывалось, область применения метода динамического программирования обширна и охватывает такие задачи, которые, на первый взгляд, невозможно решить математически. Приведем варианты задач, при решении которых существенно используются особенности динамического программирования [3—7, 9].

1. Задача о фальшивой монете. Пользуясь рычажными весами без гирь, среди s одинаковых монет найти одну более тяжелую.

Количество s монет, представленных к анализу, назовем начальным состоянием процесса. Процедура взвешивания преобразует это состояние следующим образом. Из данных монет выбираются две группы по u , $1 \leq u \leq \frac{s}{2}$, монет и кладутся на рычажные весы. Если рычажные весы останутся в равновесии, фальшивая монета находится среди оставшихся $s - 2u$ монет. В противном случае для анализа остается группа из u монет.

Применяя инвариантное погружение вместо исходной задачи, рассмотрим задачу с произвольным числом x монет. Тогда описанная процедура взвешивания на этапе t приводит к следующему уравнению:

$$x(t+1) = x(t) - \begin{cases} 2u(t), & \text{если весы в равновесии;} \\ x(t) - u(t), & \text{в противном случае,} \end{cases} \quad (1)$$

где $x(1) = x$; $x(t)$ — количество монет на этапе t , $t = 1, 2, \dots$, взвешивания; $u(t)$ — количество монет в группе для взвешивания на этапе t .

Уравнение (1) — пример математической модели дискретного процесса, заданной аналитическим способом не полностью.

Величину $u(t)$ можно трактовать как управление. Из постановки задачи следует, что должны выполняться ограничения

$$1 \leq u(t) \leq \frac{x(t)}{2}. \quad (2)$$

Задачу о фальшивой монете теперь можно сформулировать как задачу о быстродействии*: найти управление $u(t)$, $t \geq 1$, стесненное ограничениями (2), которое за наименьшее время t_1^0 переводит траекторию $x(t)$ системы (1) из состояния s в состояние $x(t_1^0) = 1$.

Предположим, что при любом x сформулированная задача имеет решение. Обозначим через $B(x)$ — функцию Беллмана — минимальное число взвешиваний в самом неблагоприятном случае для обнаружения фальшивой монеты среди заданных x монет.

Составим уравнение Беллмана. Имея x монет, сделаем одно взвешивание, выбрав две группы по u монет. В зависимости от положения рычажных весов состояние x при этом перейдет в одно из следующих состояний:

$$x - 2u \quad (\text{весы в равновесии}); \quad (3)$$

$$u \quad (\text{нет равновесия}). \quad (4)$$

* В действительности, это не полная модель исходной задачи, поскольку в ней не учтена неопределенность в выборе элементов групп для взвешивания. При составлении уравнения Беллмана эта неопределенность учитывается.

В случае (3) для выделения фальшивой монеты останется $B(x-2u)$, в случае (4) — $B(u)$ взвешиваний. Таким образом, для x монет при сделанном выборе u монет потребуется $1+B(x-2u)$ или $1+B(u)$ взвешиваний. В самом неблагоприятном случае число взвешиваний

$$1 + \max \{B(x-2u), B(u)\}. \quad (5)$$

Если выбрать число u , $1 \leq u \leq \frac{x}{2}$, так, чтобы выражение (5) было минимальным, очевидно, результат будет равен $B(x)$:

$$B(x) = 1 + \min_{1 \leq u \leq \frac{x}{2}} \max \{B(x-2u), B(u)\}. \quad (6)$$

Для полученного уравнения Беллмана граничные условия имеют вид

$$B(0) = 0, B(1) = 0, B(2) = 1, B(3) = 1. \quad (7)$$

Используя соотношения (6), (7), нетрудно подсчитать числа $B(4)$, $B(5)$, ..., равные минимальным количествам взвешиваний при 4, 5, ... заданных монетах.

По своему определению функция $B(x)$ является неубывающей по x . Поэтому в (6) минимум достигается при тех значениях u , при которых разность между $x-2u$ и u минимальна, т. е.

$$\left. \begin{aligned} u &= \left\lfloor \frac{x}{3} \right\rfloor, \text{ если } x = 3m \text{ или } x = 3m + 1; \\ u &= \left\lfloor \frac{x}{3} \right\rfloor + 1, \text{ если } x = 3m + 2, m = 1, 2, \dots \end{aligned} \right\} \quad (8)$$

Соотношения (8) доставляют оптимальное правило действий в общем случае. Например, пусть $s=23$. Тогда $m=7$, $23=3 \cdot 7+2$ и при первом взвешивании выделяем две группы по 8 монет. Если фальшивая монета оказалась среди 8 монет, выделяем две группы по 3 монеты. Если же фальшивая монета осталась среди 7 монет, выделяем две группы по 2 монеты и т. д.

2. Задача о замене оборудования. Для производства некоторого изделия намечено в течение n лет использо-

вать оборудование одного типа, первоначальная стоимость которого равна α и которое к началу намеченного периода проработало t лет. В силу естественных причин доход, полученный от оборудования, со временем меняется и равен $f(t)$, если оборудованием пользоваться в период от t до $t+1$ года. Найти план замены оборудования, при котором доход за n лет будет максимальным.

Погрузим сформулированную задачу в семейство аналогичных задач, в которых вместо данного числа n фигурирует произвольная длительность s использования оборудования, и эксплуатация начинается не с заданного оборудования, а с оборудования, использованного произвольное число t лет.

Пусть $B(s, t)$ — функция Беллмана — максимальный доход, который можно получить за s лет, если в начале эксплуатируется оборудование, которому t лет.

Получим уравнение Беллмана. В начале t -го года (от t до $t+1$) имеются две возможности: 1) не менять оборудование, 2) купить новое. В первом случае доход $B(s, t)$ складывается из двух частей: а) за год от t до $t+1$ получим доход $f(t)$; б) за оставшееся время, равное $s-1$ году, максимальный доход от оборудования, использованного $t+1$ год, равен $B(s-1, t+1)$. Итого:

$$f(t) + B(s-1, t+1). \quad (9)$$

Во втором случае доход $B(s, t)$ складывается из следующих частей: а) годовой доход от нового оборудования $f(0)$ за вычетом его стоимости α ; б) максимальный доход $B(s-1, 1)$ за $s-1$ год от оборудования, которое проработало один год. Итого:

$$f(0) - \alpha + B(s-1, 1). \quad (10)$$

Сравнивая (9) и (10), получаем выражение для $B(s, t)$:

$$B(s, t) = \max \{f(t) + B(s-1, t+1), f(0) - \alpha + B(s-1, 1)\}. \quad (11)$$

Если максимум в правой части достигается на первом выражении, то оборудование сохраняется, в противном случае покупается новое оборудование в начале t -го года.

Граничное условие для решения уравнения (11) следует из определения функции Беллмана

$$B(1, t) = \max \{f(t), f(0) - \alpha\}. \quad (12)$$

Решая уравнение (11) с граничным условием (12), получаем последовательность функций: $B(1, t)$, $B(2, t)$, ..., $B(n, t)$. Значение $B(n, m)$ равно максимальному доходу, который можно получить на заданном оборудовании в исходной задаче. Чтобы определить план замены оборудования, поступаем следующим образом. Из (11) находим

$$B(n, m) = \max \{f(m) + B(n-1, m+1), f(0) - \alpha + B(n-1, 1)\}.$$

Если максимум достигается на выражении $f(m) + B(n-1, m+1)$, то первый год оборудование не меняем, в противном случае сразу покупаем новое оборудование.

Найдем план действий на второй год. Если в течение первого года работали на старом оборудовании, то из (11) находим

$$B(n-1, m+1) = \max \{f(m+1) + B(n-2, m+2), f(0) - \alpha + B(n-2, 1)\}.$$

При

$$f(m+1) + B(n-2, m+2) > f(0) - \alpha + B(n-2, 1)$$

оборудование сохраняется и на второй год работы, в противном случае в начале второго года нужно его заменить на новое.

Если уже в течение первого года работало новое оборудование, то для нахождения плана действий на второй год из (11) вычисляем

$$B(n-1, 1) = \max \{f(1) + B(n-2, 2), f(0) - \alpha + B(n-2, 1)\}.$$

При

$$f(1) + B(n-2, 2) > f(0) - \alpha + B(n-2, 1) \quad (13)$$

оборудование, использованное в течение первого года, сохраняется. Если вместо (11) выполняется неравенство противоположного смысла, то оборудование в начале второго года заменяется на новое.

Продолжая описанный процесс, можно полностью найти оптимальный план замены оборудования на весь период из n лет.

3. Оптимальные пути в сетях. Заданы n точек a_1, \dots, a_n . Расходы по перемещению из точки a_i в точку a_j равны f_{ij} ($f_{ii} = 0$), $i, j = \overline{1, n}$. Найти путь из точки a_k в точку a_m , минимальный по затратам.

В данной задаче начальная и конечная точки заданы, причем промежуточными могут быть все остальные $n-2$ точки. Погрузим задачу в семейство задач, в которых начальная точка произвольная a_s и число промежуточных точек не превосходит t .

Пусть $B(s, t)$ — функция Беллмана, равная минимальным затратам по перемещению из точки a_s в точку a_m не более чем через t промежуточных точек.

Составим уравнение Беллмана. Сделаем пробный шаг: из точки a_s перейдем в точку a_j с затратами f_{sj} . Минимальные расходы перемещения из точки a_j в точку a_m с заходом в не более чем $t-1$ точку равны $B(j, t-1)$. Поэтому

$$B(s, t) = \min_{j=\overline{1, n}, j \neq s} \{f_{sj} + B(j, t-1)\}. \quad (14)$$

Граничные условия имеют вид

$$B(s, 0) = f_{sm}. \quad (15)$$

Решая уравнение (14) при условии (15), получим последовательность функций

$$B(s, 0), B(s, 1), \dots, B(s, n-2).$$

Ясно, что число $B(k, n-2)$ равно минимальным затратам по перемещению из точки a_k в точку a_m . Для определения оптимального маршрута поступаем следующим образом. Из (14) находим

$$B(k, n-2) = \min_j \{f_{kj} + B(j, n-3)\} = f_{kj^0} + B(j^0, n-3)$$

Отсюда следует, что из точки a_k нужно перейти в точку a_{j^0} . Далее, из (14) получаем

$$B(j^0, n-3) = \min_j \{f_{j^0j} + B(j, n-4)\} = f_{j^0j^1} + B(j^1, n-4).$$

Таким образом, следующей точкой на пути будет a_j^1 . Процесс продолжается до полного нахождения оптимального пути из a_k в a_m .

4. Задача о коммивояжере. Имеется n городов a_1, \dots, a_n ; расстояние между городами a_i и a_j равно a_{ij} . Найти кратчайший маршрут, начинающийся и оканчивающийся в городе a_1 и проходящий по одному разу через все остальные города. .

Воспользуемся сначала принципом инвариантного погружения: вместо фиксированного числа n городов рассмотрим произвольное число s , $2 \leq s \leq n$, городов, причем группу из s городов будем составлять из любых данных городов a_{i_1}, \dots, a_{i_s} . Далее, вместо фиксированного начального города a_1 рассмотрим произвольный город a_i из заданной группы a_1, \dots, a_n .

Функцию Беллмана $B_s(i/i_1, \dots, i_s)$ определим как длину кратчайшего маршрута из a_i в a_1 , проходящего через города $a_{i_1}, a_{i_2}, \dots, a_{i_s}$.

Для составления уравнения Беллмана сделаем пробное перемещение из города a_i в город a_{i_k} , $k = 1, 2, \dots, s$. По определению функции Беллмана, длина кратчайшего маршрута из города a_{i_k} в город a_1 , проходящего через города $a_{i_1}, \dots, a_{i_{k-1}}, a_{i_{k+1}}, \dots, a_{i_s}$, равна $B_{s-1}(i_k/i_1, \dots, i_{k-1}, i_{k+1}, \dots, i_s)$. Поэтому длина маршрута из a_i в a_1 не меньше, чем

$$a_{ii_k} + B_{s-1}(i_k/i_1, \dots, i_{k-1}, i_{k+1}, \dots, i_s). \quad (16)$$

Следовательно, минимальная длина маршрута из a_i в a_1

$$B_s(i/i_1, \dots, i_s) = \min_{1 \leq k \leq s} [a_{ii_k} + B_{s-1}(i_k/i_1, \dots, i_{k-1}, i_{k+1}, \dots, i_s)] \quad (17)$$

и определяется из (16) перебором всех пробных шагов.

Граничное условие для решения уравнения Беллмана (17) получается из определения функции Беллмана при $s=1$:

$$B_1(i/i_1) = a_{ii_1} + a_{i_1 1}. \quad (18)$$

Используя (18), из уравнения (17) легко подсчитать последовательность функций: $B_1(i/i_1)$, $B_2(i/i_1, i_2)$, \dots , $B_{n-2}(i/i_1, \dots, i_{n-2})$, $B_{n-1}(i/i_1, \dots, i_{n-1})$.

Число $B_{n-2}(1/2, 3, \dots, n)$ равно длине кратчайшего маршрута, решающего поставленную в начале пункта задачу о коммивояжере.

Оптимальный маршрут получается по следующей схеме. Из уравнения (17) находим

$$B_{n-1}(1/2, 3, \dots, n) = \min_{2 \leq i \leq n} [a_{1i} + B_{n-2}(i/2, 3, \dots, \\ i-1, i+1, \dots, n)] = a_{1i^0} + B_{n-2}(i^0/2, 3, \dots, \\ i^0-1, i^0+1, \dots, n).$$

Отсюда следует, что первым городом за a_1 на оптимальном маршруте является город a_{i^0} . Далее из (17) находим

$$B_{n-2}(i^0/2, 3, \dots, i^0-1, i^0+1, \dots, n) = \\ = \min_{i=2, 3, \dots, i^0-1, i^0+1, \dots, n} [a_{i^0 i}^0 + B_{n-3}(i/2, 3, \dots, i-1, i+1, \\ \dots, i^0-1, i^0+1, \dots, n)] = a_{i^0 i^1}^0 + B_{n-3}(i^1/2, 3, \dots, i^1-1, \\ i^1+1, \dots, i^0-1, i^0+1, \dots, n).$$

Полученное равенство означает, что вслед за городом a_{i^0} на кратчайшем маршруте лежит город $a_{i^1}^1$.

Продолжая вычисления аналогичным образом, найдем оптимальный маршрут коммивояжера.

В непрерывных процессах время (независимый аргумент) изменяется непрерывно. Наиболее распространенной моделью для описания непрерывных процессов являются дифференциальные уравнения. Непрерывные и дискретные процессы, как правило, тесно связаны: непрерывные процессы часто получаются предельным переходом из дискретных, а дискретные — из непрерывных путем квантования независимого аргумента. Поэтому изучение каждого типа件 полезно для понимания и исследования свойств другого типа процессов.

§ 1. Постановка основных задач оптимизации

Задачи вариационного исчисления и оптимального управления непрерывными процессами по постановке мало отличаются от аналогичных задач для дискретных систем. Изменения по форме вызваны типом уравнений, задающих процессы. В методах же решения можно обнаружить много общих идей. Но по трудности решения, по внутренней сложности между задачами оптимизации двух типов процессов огромная разница [2, 9, 11, 12, 13, 17].

1. Задачи вариационного исчисления [12]. Рассмотрим кривые, заданные в n -мерном пространстве функциями $x(t)$, $t \in [t_0, t_1]$, класса $C^{(1)}$ и удовлетворяющие краевым условиям:

$$x(t_0) = x_0, \quad x(t_1) = x^1, \quad (1)$$

где x_0, x^1 — фиксированные n -векторы.

В простейшей задаче вариационного исчисления требуется среди введенных кривых (называемых допустимыми) выбрать ту $x^0(t)$, вдоль которой функционал

$$I(x) = \int_{t_0}^{t_1} f_0(x(t), \dot{x}(t), t) dt \quad (2)$$

достигает минимального значения.

Если допустимые кривые удовлетворяют дифференциальному уравнению $\Phi(x, \dot{x}, t) = 0$, то имеем задачу Лагранжа.

Пусть на правый конец кривых $x(t)$ ограничений не накладывается. Задача Майера связана с минимизацией функционала $I(x) = \varphi(x(t_1))$ при соблюдении условий задачи Лагранжа.

Задачей Больца называется задача минимизации функционала

$$I(x) = \varphi(x(t_1)) + \int_{t_0}^{t_1} f_0(x, \dot{x}, t) dt$$

на кривых, удовлетворяющих условиям задачи Майера. В перечисленных задачах могут существовать другие ограничения на концах, что приводит к задачам с подвижными, свободными и концами.

Как и задачи § 1—3 гл. I, сформулированные выше задачи можно решить методом динамического программирования еще до введения основных понятий. Рассмотрим, например, простейшую задачу вариационного исчисления.

Вместо класса допустимых кривых, определенных фиксированными условиями (1), введем кривые $x(t)$, $t \in [\tau, t_1]$, класса $C^{(1)}$ такие, что $x(\tau) = y$, $x(t_1) = x^1$. На этих семействах кривых, зависящих от τ и y , определим функцию

$$B(\tau, y) = \min_{x[\tau, t_1]} \int_{\tau}^{t_1} f_0(x(t), \dot{x}(t), t) dt. \quad (3)$$

Ясно, что справедливы равенства $(\tau < t_1, \tau + \Delta\tau \leq t_1)$:

$$\begin{aligned} B(\tau, y) = \min_{\substack{x[\tau, \tau+\Delta\tau] \\ x(\tau)=y}} \min_{\substack{x[\tau+\Delta\tau, t_1] \\ x(\tau+\Delta\tau)}} \left[\int_{\tau}^{\tau+\Delta\tau} f_0(x(t), \dot{x}(t), t) dt + \right. \\ \left. + \int_{\tau+\Delta\tau}^{t_1} f_0(x(t), \dot{x}(t), t) dt \right] = \min_{\substack{x[\tau, \tau+\Delta\tau] \\ x(\tau)=y}} \left[\int_{\tau}^{\tau+\Delta\tau} f_0(x(t), \dot{x}(t), t) dt + \right. \end{aligned}$$

$$+ \min_{\substack{x[\tau+\Delta\tau, t_1] \\ x(\tau+\Delta\tau)}} \int_{\tau+\Delta\tau}^{t_1} f_0(x(t), \dot{x}(t), t) dt]. \quad (4)$$

Здесь символ $x[\tau + \Delta\tau, t_1]$ означает кривую $x(t)$, $t \in [\tau + \Delta\tau, t_1]$, исходящую из точки $x(\tau + \Delta\tau)$, которая представляет правый конец кривой $x(t)$, $t \in [\tau, \tau + \Delta\tau]$.

В (4) минимизация производится по кривым $x[\tau, \tau + \Delta\tau]$ и $x[\tau + \Delta\tau, t_1]$, составляющим кривую $x[\tau, t_1]$ класса $C^{(1)}$.

Рассмотрим второй член, стоящий в квадратных скобках. В нем в силу определения (3) нетрудно узнать функцию $B(\tau + 1, x(\tau + \Delta\tau))$. Таким образом, функция $B(\tau, y)$ удовлетворяет соотношению

$$B(\tau, y) = \min_{x[\tau, \tau+\Delta\tau]} \left[\int_{\tau}^{\tau+\Delta\tau} f_0(x(t), \dot{x}(t), t) dt + B(\tau + \Delta\tau, x(\tau + \Delta\tau)) \right]. \quad (5)$$

Из (3) получается и граничное условие

$$B(t_1, x^1) = 0. \quad (6)$$

На этом этапе заканчивалась процедура динамического программирования в дискретном случае, которая приводила к эффективному результату. Для непрерывных систем по соотношениям (5), (6) трудно найти решение задачи. Минимизация функционала (2) по кривым, заданным на отрезке $[t_0, t_1]$, свелась к минимизации правой части (5) по кривым, определенным на отрезке $[\tau, \tau + \Delta\tau]$. Эти две задачи с принципиальной точки зрения (по трудности решения) эквивалентны.

Поэтому сделаем еще один шаг. Предположим, что функция $B(\tau, y)$ имеет непрерывные производные по обоим аргументам. Тогда, считая, что $f_0(x, \dot{x}, t)$ — непрерывная функция, получаем:

$$\int_{\tau}^{\tau+\Delta\tau} f_0(x, \dot{x}, t) dt = f_0(x(\tau), \dot{x}(\tau), \tau) \Delta\tau + o(\Delta\tau),$$

$$B(\tau + \Delta\tau, x(\tau + \Delta\tau)) = B(\tau, x(\tau)) + \frac{\partial B(\tau, x(\tau))}{\partial \tau} \Delta\tau + \\ + \frac{\partial B'(\tau, x(\tau))}{\partial y} \dot{x}(\tau) \Delta\tau + o(\Delta\tau).$$

Подставим эти значения в (5):

$$B(\tau, y) = \min_{x[\tau, \tau + \Delta\tau]} [f_0(y, \dot{x}(\tau), \tau) \Delta\tau + B(\tau, y) + \\ + \frac{\partial B(\tau, y)}{\partial \tau} \Delta\tau + \frac{\partial B'(\tau, y)}{\partial y} \dot{x}(\tau) \Delta\tau + o(\Delta\tau)].$$

Уберем в обеих частях член $B(\tau, y)$ и разделим на $\Delta\tau > 0$. Пусть $\Delta\tau \rightarrow 0$. Поскольку символ $\min_{x[\tau, \tau + \Delta\tau]}$ означает операцию минимизации по допустимым кривым $x[\tau, \tau + \Delta\tau]$, а в правой части характеристики этих кривых содержатся лишь в $\dot{x}(\tau)$ и $o(\Delta\tau)$, после предельного перехода ($o(\Delta\tau)/\Delta\tau \rightarrow 0$) естественно в (5) использовать операцию $\min_{\dot{x}(\tau)}$. В результате приходим к дифференциально-функциональному уравнению Беллмана:

$$-\frac{\partial B(\tau, y)}{\partial \tau} = \min_v [f_0(y, v, \tau) + \frac{\partial B'(\tau, y)}{\partial y} v] \quad (7)$$

с граничным условием $B(t_1, x^1) = 0$.

Таков окончательный результат динамического программирования для простейшей задачи вариационного исчисления. В основных чертах этот результат сохранится и для других сформулированных выше задач вариационного исчисления. Но на этом останавливаться не будем.

З а м е ч а н и е. Поскольку в простейшей задаче вариационного исчисления не накладываются ограничения на значения производной допустимых кривых, то в (7) область изменения параметра v открыта. Следовательно, наряду с (7) вдоль оптимальной кривой $x(t)$ выполняются равенства:

$$\frac{\partial f_0(x(\tau), \dot{x}(\tau), \tau)}{\partial \dot{x}} + \frac{\partial B(\tau, x(\tau))}{\partial y} = 0; \quad (8)$$

$$-\frac{\partial B(\tau, x(\tau))}{\partial \tau} = f_0(x(\tau), \dot{x}(\tau), \tau) + \frac{\partial B'(\tau, x(\tau))}{\partial y} \dot{x}(\tau). \quad (9)$$

Равенство (9) означает

$$\left. \frac{dB(y, \tau)}{dt} \right|_{x(\tau)} = -f_0(x(\tau), \dot{x}(\tau), \tau),$$

т. е.

$$B(t_1, x(t_1)) - B(\tau, y) = - \int_{\tau}^{t_1} f_0(x(t), \dot{x}(t), t) dt, \quad y = x(\tau).$$

Используя граничные условия (6), находим (формально)

$$\frac{\partial B(\tau, x(\tau))}{\partial y} = \int_{\tau}^{t_1} \frac{\partial f_0(x(t), \dot{x}(t), t)}{\partial y} dt.$$

Подставив это выражение в (8), получим классическое уравнение Эйлера в интегральной форме:

$$\frac{\partial f_0(x(\tau), \dot{x}(\tau), \tau)}{\partial \dot{x}} + \int_{\tau}^{t_1} \frac{\partial f_0(x(t), \dot{x}(t), t)}{\partial y} dt = 0,$$

которое после дифференцирования по τ принимает обычный вид

$$\frac{\partial f_0(x, \dot{x}, t)}{\partial x} - \frac{d}{dt} \frac{\partial f_0(x, \dot{x}, t)}{\partial \dot{x}} = 0.$$

2. Задачи оптимального управления [11, 17]. Пусть $x(t), t \geq t_0$, — процесс, полученный в результате воздействия управления $u(t), t \geq t_0$, на систему

$$\frac{dx}{dt} = f(x, u, t), \quad x(t_0) = x_0, \quad (10)$$

где x — n -вектор; u — r -вектор; t — скаляр.

Относительно функции $f(x, u, t)$ всегда будем предполагать выполненными условия, которые обеспечивают существование и единственность решения $x(t), t \geq \tau$, уравнения (10) при заданных начальном условии $x(\tau) = y$ и кусочно-непрерывном управлении $u(t), t \geq \tau$.

Управление $u(t), t \geq t_0$, и соответствующая ему траектория называются допустимыми, если $u(t)$ — r -мерная кусочно-непрерывная функция со значениями из множества $U: u(t) \in U, t \geq t_0$, а траектория удовлетворяет всем ограничениям задачи.

Упомянутые ограничения могут накладываться на концы траектории и на отдельные участки. В соответствии с этим говорят о задаче: а) со свободным правым концом, если ограничений на $x(t_1)$ нет; б) с подвиж-

н ы м правым концом, если $x(t_1) \in G$, где G — множество n -мерного пространства; в) с з а к р е п л е н н ы м правым концом, если $x(t_1) = x^1$ (x^1 — фиксированный n -вектор).

В зависимости от ограничений на промежуток $[t_0, t_1]$ различают задачи: а) с ф и к с и р о в а н н о й продолжительностью процесса, если t_0, t_1 — заданные фиксированные числа; б) с н е ф и к с и р о в а н н о й продолжительностью процесса, если хотя бы одно из чисел t_0, t_1 может принимать значения из некоторого заданного промежутка.

Каждый процесс $x(t), t \geq t_0$, порожденный допустимым управлением, можно оценить различными функционалами, критериями качества. Если

$$I(u) = \varphi(x(t_1)),$$

говорят о задаче т е р м и н а л ь н о г о управления (задаче оптимального управления т и п а М а й е р а); при

$$I(u) = \int_{t_0}^{t_1} f_0(x(t), u(t), t) dt$$

— о задаче оптимального управления т и п а Л а г р а н ж а; если критерий качества имеет вид

$$I(u) = \varphi(x(t_1)) + \int_{t_0}^{t_1} f_0(x(t), u(t), t) dt,$$

то задачу оптимизации принято называть задачей оптимального управления т и п а Б о л ь ц а.

3. Эквивалентность задач оптимального управления. Задачи вариационного исчисления можно формулировать в терминах теории оптимального управления. Каждая из трех задач оптимального управления сводится к другим с помощью несложных преобразований. Так, например, задача Больца сводится к задаче терминального управления введением дополнительной координаты

$$\dot{x}_{n+1} = \frac{\partial \varphi'(x)}{\partial x} f(x, u, t) + f_0(x, u, t), \quad x_{n+1}(t_0) = \varphi(x_0).$$

Действительно,

$$I(u) = \int_{t_0}^{t_1} f_0(x(t), u(t), t) dt + \varphi(x(t_1)) =$$

$$\begin{aligned}
&= \varphi(x_0) + \int_{t_0}^{t_1} \left\{ \frac{\partial \varphi'(x(t))}{\partial x} f(x(t), u(t), t) + f_0(x(t), u(t), t) \right\} dt = \\
&= x_{n+1}(t_0) + \int_{t_0}^{t_1} \dot{x}_{n+1}(t) dt = x_{n+1}(t_1).
\end{aligned}$$

Аналогично устанавливается связь в остальных случаях.

Допустимые управления и соответствующие им траектории называются оптимальными в конкретной задаче, если они минимизируют критерий качества этой задачи.

§ 2. Инвариантное погружение

Как и в дискретном случае (см. § 5 гл. I), будем различать два типа инвариантного погружения.

Будем говорить, что осуществлено прямое погружение процесса, если вместо фиксированного процесса

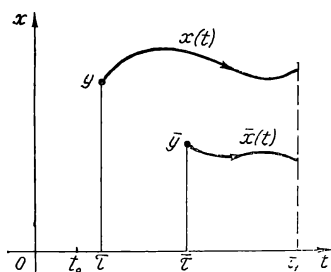


Рис. 6

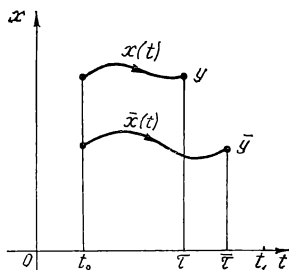


Рис. 7

рассматривается семейство, заданное уравнением (рис. 6)

$$\frac{dx(t)}{dt} = f(x(t), u(t), t), \quad x(\tau) = y, \quad \tau \leq t \leq t_1,$$

и параметрами τ, y , где τ — произвольное число; y — произвольный n -вектор.

Если вместо конкретного процесса $x(t)$, $t \in [t_0, t_1]$, рассматривается семейство (рис. 7)

$$\frac{dx(t)}{dt} = f(x(t), u(t), t), \quad x(\tau) = y, \quad t_0 \leq t \leq \tau,$$

будем говорить, что осуществлено обратное погружение.

§ 3. Принцип оптимальности

В описание принципа оптимальности [3], сформулированного в § 6 гл. I для дискретных процессов, следует внести некоторые изменения, вызванные новым содержанием понятия допустимого управления в непрерывных процессах. При определении допустимых управлений теперь четко оговариваются не только ограничения, накладываемые на значения управлений и траекторий, но и классы функций, из которых выбираются управления и траектории. В дискретном случае последнее обстоятельство было не существенно, ибо оно непосредственно связано лишь с непрерывностью изменения аргумента.

Пусть непрерывный процесс $x(t)$ задан на отрезке $[t_0, t_1]$, $t_1 > t_0$, уравнением (10) § 1 и порождается некоторым управлением $u(t)$, $t \in [t_0, t_1]$.

Отрезок $[t_0, t_1]$ разобьем на два периода $[t_0, \tau)$, $[\tau, t_1]$. Управление $u[\tau, t_1]$ во втором периоде назовем допустимым, если оно вместе с управлением $u[t_0, \tau)$ первого периода составляет допустимое управление $u[t_0, t_1]$. Тонкость в определении состоит в том, что из рассмотрения исключаются все управления $u[\tau, t_1]$, которые в совокупности с $u[t_0, \tau)$ не составляют функцию $u[t_0, t_1]$ заданного класса, хотя остальные требования могут быть и выполненными. Аналогичные изменения следует внести и в определение оптимального управления в первом периоде.

С учетом указанных изменений принцип оптимальности (см. § 6 гл. I) переносится на непрерывные процессы. То, что изменения необходимы, видно из следующего примера.

Пример. В классе непрерывных функций с кусочно-непрерывными производными найти управление $u(t)$, $|\dot{u}(t)| \leq 1$, $t \in [0, 1]$, на котором достигает минимума функционал

$$I(u) = \int_0^1 \dot{x}(t) \ddot{x}(t) dt,$$

где $x(t)$ — непрерывный процесс, заданный уравнением $\dot{x} = u$, $x(0) = 1$. Ясно, что

$$I(u) = \int_0^1 u(t) \dot{u}(t) dt = \frac{1}{2} \int_0^1 \frac{d}{dt} u^2(t) dt = \frac{1}{2} (u^2(1) - u^2(0)).$$

Поэтому управление $u^0(t) = 1 - t$ оптимально. Рассмотрим это управление во втором периоде $[1/2, 1]$. В начале этого периода процесс находится в состоянии $x^0(1/2) = 11/8$. Если искать оптимальное управление для подпроцесса $x(t)$, $t \in [1/2, 1]$, $x(1/2) = 11/8$, то критерий качества

$$I = \int_{1/2}^1 \dot{x} \dot{x} dt = 1/2 [u^2(1) - u^2(1/2)]$$

достигает минимального значения на управлении $u^*(t) = 2 - 2t$, $t \in [1/2, 1]$, а управление $u^0(t)$, $t \in [1/2, 1]$ заведомо не оптимально.

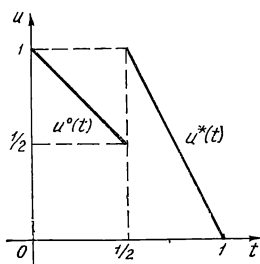


Рис. 8

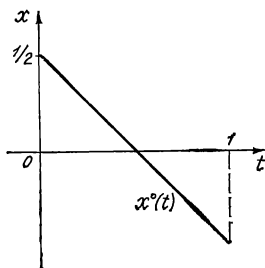


Рис. 9

Полученное противоречие с принципом оптимальности как необходимым условием объясняется тем, что совокупность функций $u^0 [0, 1/2]$ и $u^*[1/2, 1]$ не составляет непрерывную функцию, а значит, и допустимое управление (рис. 8).

Существенными остаются и предположения, сделанные в § 6 гл. I при формулировке принципа оптимальности. Действительно, в примере 2 (см. § 6 гл. I) для непрерывных процессов (допустимые управления — кусочно-непрерывные функции):

$$\dot{x} = u, \quad |u| \leq 1, \quad x(0) = 1/2, \quad I(u) = \max_{0 \leq t \leq 1} |x(t)|.$$

Нетрудно подсчитать, что на оптимальном управлении $u^0(t) = -1$ принцип оптимальности не выполняется для периодов $[0, 1/2]$, $[1/2, 1]$ (рис. 9). Для начального состояния $x(1/2) = 0$ второго периода $[1/2, 1]$ оптимальным является управление $u^*(t) \equiv 0$, а управление $u^0 [1/2, 1] \equiv -1$, очевидно, не оптимально.

Полностью сохраняются замечания, сделанные в п. 6 § 1 относительно достаточности принципа оптимальности.

Как и в дискретном случае, ниже принцип оптимальности будет использован скорее как наводящее соображение для вывода уравнений Беллмана. Сам вывод каждый раз будет опираться существенно лишь на прин-

цип инвариантного погружения и эквивалентные математические преобразования, связанные со свойствами операций \min .

§ 4. Дифференциальные уравнения Беллмана

Непрерывные процессы $x(t)$ отличаются от дискретных $x(t)$ тем, что параметр (аргумент) t меняется непрерывно, принимая все значения из промежутка числовой оси, на котором задан процесс. Это новое свойство позволяет сделать переход от рекуррентно-функциональных уравнений к дифференциально-функциональным уравнениям Беллмана (для краткости в дальнейшем слово «функциональное» будем опускать). В данном параграфе этот переход осуществляется для ряда типичных задач оптимизации.

1. Задача быстродействия для стационарных систем. Пусть непрерывный процесс $x(t)$, $t \geq 0$, описывается дифференциальным уравнением

$$\frac{dx}{dt} = f(x, u), \quad (1)$$

правая часть которого не зависит от времени t . Вектор x является n -мерным, u — r -мерным.

Допустимыми назовем управления $u(t)$, $t \geq 0$, представляющие r -мерные кусочно-непрерывные функции со значениями из ограниченного множества U : $u(t) \in U$, $t \geq 0$. В точках разрыва ϑ функцию $u(t)$ определим как предел справа, через $u(\vartheta-)$ обозначим предел слева.

Задача. Заданы два n -вектора x_0, x^1 . Требуется с помощью допустимого управления за минимальное время перевести траекторию $x(t)$ системы (1) из состояния x_0 ($x(0) = x_0$) в состояние x^1 ($x(t_1) = x^1$).

Критерием качества в задаче является время перехода $t_1 = t(x_0, x^1)$, зависящее от векторов x_0, x^1 и управления $u[0, t_1]$. Таким образом,

$$I = I(x_0, u[0, t_1], x^1) = t(x_0, x^1). \quad (2)$$

Пусть $u[0, \tau)$, $u[\tau, t_1)$, $t_0 < \tau < t_1$, — части допустимого управления $u[0, t_1]$, переводящего траекторию $x(t)$ из точки x_0 в x^1 , $x^1 \neq x_0$. Тогда, очевидно, выполняется равенство

$$I(x_0, u[0, t_1], x^1) = \tau + I(x(\tau), u[\tau, t_1], x^1), \quad (3)$$

т. е. критерий качества удовлетворяет условию (3) § 6 гл. I. Из представления (3) следует, что выполняется и условие (4) § 6 гл. I. Поэтому для критерия (3) принцип оптимальности справедлив и как необходимое и как достаточное условие оптимальности.

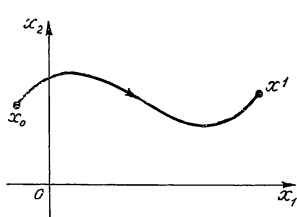


Рис. 10

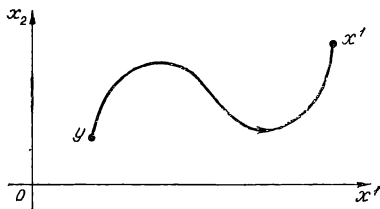


Рис. 11

Погрузим процесс (1) (рис. 10), соответствующий начальному состоянию x_0 ($x_0(0) = x_0$), в семейство

$$\frac{dx}{dt} = f(x, u), \quad x(0) = y, \quad (4)$$

зависящее от n -вектора y (рис. 11). На семействе (4) введем функцию Беллмана

$$B(y) = \min t(y, x^1), \quad (5)$$

где минимум берется по всем допустимым управлениям, переводящим траекторию $x(t)$ системы из состояния y в состояние x^1 . Функция $B(y)$ определена лишь на множестве управляемости системы (1) (относительно точки x^1 для момента $t=0$), состоящем из таких точек y , для которых $t(y, x^1) < \infty$.

Пусть $y, y \neq x^1$, — внутренняя точка множества управляемости A . Тогда найдется настолько малое число Δ , $\Delta > 0$, что ни одна траектория $x(t)$ системы (4) на отрезке $[0, \Delta]$ не выйдет из множества управляемости и не достигнет точки x^1 (рис. 12).

При выбранном y и любых Δt , $0 < \Delta t < \Delta$, очевидны соотношения

$$\begin{aligned} B(y) = \min_{u(t) \in U, t \geq 0} t(y, x^1) &= \min_{u(t) \in U, t \in [0, \Delta t]} \{ \Delta t + \\ &+ \min_{\substack{u(t) \in U, t \geq \Delta t \\ x(\Delta t) = x(\Delta t, y, u[0, \Delta t])}} t(x(\Delta t), x^1) \}. \end{aligned} \quad (6)$$

При сделанных предположениях все члены в (6) имеют смысл. Вычислим второе слагаемое в фигурных скобках. Для этого рассмотрим функции $y(t) = x(t + \Delta t)$, $v(t) = u(t + \Delta t)$. Непосредственной подстановкой убеждаемся, что пара $y(t)$, $v(t)$ удовлетворяет уравнению (4) с начальным условием $y(0) = x(\Delta t)$ и что функция

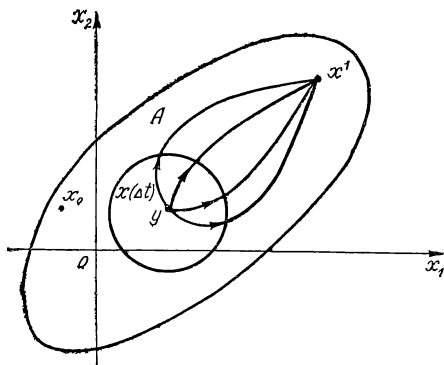


Рис. 12

$v(t), t \geq 0$, — допустимое управление. Поэтому в силу определения (5) имеем

$$\min_{\substack{u(t) \in U, t \geq \Delta t \\ x(\Delta t) = x(\Delta t, y, u[0, \Delta t])}} t(x(\Delta t), x^1) = \min_{\substack{v(t) \in U, t \geq 0 \\ y(0) = x(\Delta t)}} t(x(\Delta t), x^1) = B(x(\Delta t)).$$

Подставив полученное выражение в (6), получим соотношение

$$B(y) = \Delta t + \min_{u(t) \in U, t \in [0, \Delta t]} B(x(\Delta t)), \quad (7)$$

которое эквивалентно рекуррентно-функциональному уравнению Беллмана, полученному в гл. I для дискретных процессов с $\Delta t = 1$.

Теперь воспользуемся тем, что процесс $x(t)$ непрерывен, т. е. величину Δt можно выбрать сколь угодно малой. Чтобы дальнейшие операции были законны, предположим, что функция $B(y)$ дифференцируема. Тогда, с одной стороны, из (4) имеем

$$x(\Delta t) = x(0) + f(x(0), u(0)) \Delta t + o(\Delta t) = y + f(y, u(0)) \Delta t + o(\Delta t),$$

с другой,

$$B(x(\Delta t)) = B(x(0)) + \frac{\partial B'(x(0))}{\partial y} [x(\Delta t) - x(0)] + \\ + o(\|x(\Delta t) - x(0)\|).$$

Объединяя эти разложения, приходим к соотношению

$$B(x(\Delta t)) = B(y) + \frac{\partial B'(y)}{\partial y} f(y, u(0)) \Delta t + o(\Delta t).$$

Подставим это разложение в (7), уберем в результате член $B(y)$, остаток разделим на Δt , $\Delta t > 0$, и устремим Δt к нулю. Поскольку $o(\Delta t)/\Delta t \rightarrow 0$ при $\Delta t \rightarrow 0$, окончательно получим уравнение

$$\min_{u \in U} \frac{\partial B'(y)}{\partial y} f(y, u) = -1, \quad (8)$$

которое называется дифференциальным уравнением Беллмана для задачи быстрогодействия в стационарной системе (1).

Уравнение (8) представляет собой дифференциальное уравнение в частных производных. Граничное условие для него получается из (5):

$$B(x^1) = 0. \quad (9)$$

Попутно с выводом уравнения Беллмана мы доказали, что оптимальный закон управления $u(y)$ в задаче быстрогодействия удовлетворяет соотношению

$$\frac{\partial B'(y)}{\partial y} f(y, u(y)) = \min_{u \in U} \frac{\partial B'(y)}{\partial y} f(y, u).$$

Если оптимальный закон управления известен, то оптимальное управление $u^0(t)$ и траектория $x^0(t)$ находятся из (1) и равенства $u^0(t) = u(x^0(t))$.

Основным достоинством полученного результата является то, что использованная схема рассуждений приводит к оптимальному закону управления $u(x)$. В технике управления, имеющие форму $u = u(x)$, называются управлениями типа обратной связи. Если такое управление получено для некоторой задачи, к объекту управления (1) (рис. 13) можно добавить обратную связь, реализующую закон $u = u(x)$ (рис. 14), и получить замк-

н у т у ю систему управления (рис. 15), которая функционирует в оптимальном режиме без участия человека, т. е. в ней при любом возмущении начального состояния x_0 автоматически возбуждается оптимальный переходный процесс $x(t)$.

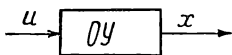


Рис. 13

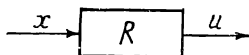


Рис. 14

Получение управлений типа обратной связи — основное достоинство динамического программирования.

Недостатком метода динамического программирования в рассмотренной задаче является то, что исходная задача оптимизации сведена к решению сложного уравнения в частных производных (8) со специфическим граничным условием (9), причем это уравнение играет

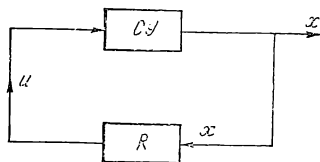


Рис. 15

роль лишь необходимого условия оптимальности, и его вывод нельзя считать вполне строгим, так как использованное в нем предположение дифференцируемости до решения задачи, как правило, нельзя проверить.

З а м е ч а н и я. 1. При выводе уравнения (8) мы следовали в основном классической схеме получения дифференциальных уравнений Беллмана. В этом доказательстве некоторое сомнение может вызвать законность перестановки операций \lim и \min при переходе от соотношения (7) к уравнению (8). Существует другая схема рассуждений (см. п. 3), в котором такого сомнения не возникает.

2. Функция Беллмана (5) зависит лишь от вектора состояния. Это следствие как стационарности системы (1), так и вида критерия качества, ограничений, наложенных на управление и состояние. Задачи оптимизации, для которых можно ввести функцию Беллмана, не зависящую от времени, назовем стационарными.

2. Прямое уравнение Беллмана. Опять рассмотрим задачу быстрогодействия, сформулированную в п. 1, однако теперь используем обратное погружение, т. е. вместо траектории $x(t)$ системы (1), приходящей из x_0 в точку x^1 (рис. 16), введем семейство траекторий

$$\frac{dx}{dt} = f(x, u), \quad x(0) = x_0, \quad x(t_1) = y, \quad (10)$$

зависящих от n -вектора y . При фиксированном y траектории семейства за конечное время допустимым управлением переводятся из x_0 в y (рис. 17). На траекториях семейства (10) определим функцию

$$B^-(y) = \min t(x_0, y), \quad (11)$$

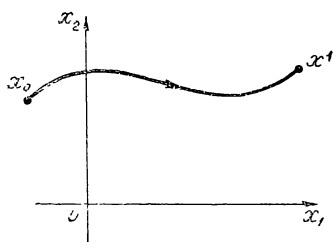


Рис. 16

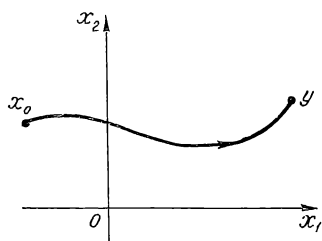


Рис. 17

где минимум берется по допустимым управлениям $u(t)$, $t \in [0, t(x_0, y)]$, переводящим траектории из x_0 в y .

Множество точек y , для которых $t(x_0, y) < \infty$, называется множеством достижимости системы (1) (относительно точки x_0 для момента $t=0$). Функция $B^-(y)$ определена лишь на последнем множестве.

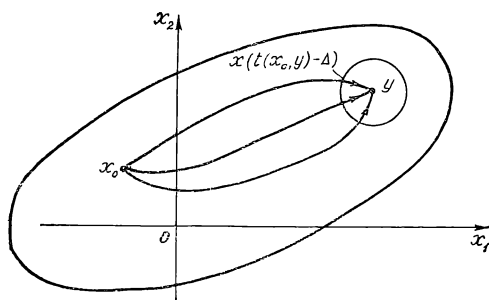


Рис. 18

В силу свойства (3) критерий быстродействия удовлетворяет условиям, при которых справедлив обратный принцип оптимальности (см. § 6 гл. I).

Пусть $y, y \neq x_0$, — внутренняя точка множества достижимости. Тогда найдется столь малое число $\Delta, \Delta > 0$, что ни одна траектория системы (10) на отрезке $[t(x_0,$

$y) - \Delta, t(x_0, y)]$ не выйдет из множества достижимости B и не пройдет через точку x_0 (рис. 18).

При выбранном y и любых $\Delta t, 0 < \Delta t < \Delta$, справедливы соотношения

$$B^-(y) = \min_t t(x_0, y) = \Delta t + \min_{\substack{u(t) \in U \\ t \in [t(x_0, y) - \Delta t, t(x_0, y)]}} \min_{\substack{t(x_0, y) - \Delta t \\ t(x_0, y)}} t(x_0, y) \\ x(t(x_0, y) - \Delta t) = \Delta t + \min_{\substack{u(t) \in U \\ t \in [t(x_0, y) - \Delta t, t(x_0, y)]}} B^-(x(t(x_0, y) - \Delta t)). \quad (12)$$

Считая функцию $B^-(y)$ дифференцируемой, запишем разложения:

$$B^-(x(t(x_0, y) - \Delta t)) = B^-(x(t(x_0, y))) + \frac{\partial B^-}{\partial y} [x(t(x_0, y) - \Delta t) - x(t(x_0, y))] + o(\Delta x); \\ x(t(x_0, y) - \Delta t) = x(t(x_0, y)) - f(x(t(x_0, y)), u(t^-(x_0, y))) \Delta t + o(\Delta t).$$

Подставим их в (12) и, учитывая, что $x(t(x_0, y)) = y$, разделим на Δt . При $\Delta t \rightarrow 0$ получим прямое уравнение Беллмана:

$$\min_{u \in U} \frac{\partial B^-}{\partial y} f(y, u) = 1. \quad (13)$$

Граничное условие следует из определения (11):

$$B^-(x_0) = 0. \quad (14)$$

Прямое уравнение Беллмана (13) отличается от обратного (8) лишь знаком в правой части. Граничные условия задаются в разных точках. Полученная антисимметрия объясняется антисимметричностью критерия качества задачи. В отличие от дискретного аналога $x(t+1) = f(x(t), u(t))$ дифференциальное уравнение (1) не имеет явно выраженной несимметричности к течению времени, и поэтому уравнения (8) и (13) отличаются лишь знаком.

Использование уравнения (13), (14) приводит снова к оптимальным законам управления.

3. Простейшая задача терминального управления.
Пусть требуется решить задачу

$$\frac{dx}{dt} = f(x, u, t), \quad x(t_0) = x_0, \quad t \in [t_0, t_1], \quad (15)$$

$$u(t) \in U, \quad I(u) = \varphi(x(t_1)) \rightarrow \min, \quad (16)$$

где x — n -вектор; u — r -вектор; t_0, t_1, x_0 — заданные параметры; $f(x, u, t)$, $\varphi(x)$ — непрерывные функции. В остальном сохраним предположения п. 1.

Каждый процесс $x(t), t \in [t_0, t_1]$, соответствующий в силу (15) некоторому допустимому управлению $u(t), t \in [t_0, t_1]$, погрузим в семейство

$$\frac{dx}{dt} = f(x, u, t), \quad x(\tau) = y, \quad t \in [\tau, t_1]. \quad (17)$$

На решениях семейства (17) определим функцию

$$B(\tau, y) = \min \varphi(x(t_1)), \quad (18)$$

где минимум берется по всем допустимым управлениям $u(t) \in U, t \in [\tau, t_1]$. Пусть $u(t), t \in [\tau, t_1]$, — управление, на котором правая часть из (18) достигла минимума; $x(t), t \in [\tau, t_1]$, — соответствующая траектория системы (17). Из свойства критерия (16) следует, что при любом $\Delta \tau, \Delta \tau > 0, \tau + \Delta \tau < t_1$, выполняется равенство

$$B(\tau, y) \equiv B(\tau + \Delta \tau, x(\tau + \Delta \tau)). \quad (19)$$

Наряду с управлением $u(t), t \in [\tau, t_1]$, рассмотрим управление $\tilde{u}(t), t \in [\tau, t_1]$, построенное по правилу

$$\tilde{u}(t) = \begin{cases} v, & t \in [\tau, \tau + \Delta \tau), \\ \bar{u}(t), & t \in [\tau + \Delta \tau, t_1), \end{cases} \quad (20)$$

где v — некоторый вектор из U , а управление $\bar{u}(t), t \in [\tau + \Delta \tau, t_1]$, оптимально относительно точки $\tilde{x}(\tau + \Delta \tau)$, в которую попала траектория $x(t)$ системы (17) с управлением $\tilde{u}(t) = v, t \in [\tau, \tau + \Delta \tau)$. Управление во втором периоде доставляет критерию качества (16) значение $B(\tau + \Delta \tau, \tilde{x}(\tau + \Delta \tau))$. В силу специфики критерия качества (16) это число равно значению критерия ка-

чества (16) на управлении (20). А поскольку управление (20), вообще говоря, отлично от оптимального, то при всех $\Delta \tau > 0$, $\tau + \Delta \tau < t_1$, $v \in U$, выполняется неравенство

$$B(\tau + \Delta \tau, \tilde{x}(\tau + \Delta \tau)) \geq B(\tau, y). \quad (21)$$

Из тождества (19) вдоль оптимального управления получаем равенство

$$\left. \frac{d_+ B(\tau, y)}{dt} \right|_{(17), u(\tau)} = 0, \quad (22)$$

где слева стоит полная правосторонняя производная по времени вдоль оптимальной траектории $x(\tau)$ системы (17).

Из неравенства (21) имеем

$$\left. \frac{d_+ B(\tau, y)}{dt} \right|_{(17), v} = \lim_{\Delta \tau \rightarrow 0, \Delta \tau > 0} \frac{B(\tau + \Delta \tau, \tilde{x}(\tau + \Delta \tau)) - B(\tau, y)}{\Delta \tau} \geq 0. \quad (23)$$

Объединяя соотношения (22) и (23), убеждаемся, что вдоль оптимального управления $u(\tau)$ выполняется уравнение

$$\left. \frac{d_+ B(\tau, y)}{dt} \right|_{(17), u(\tau)} = \min_{v \in U} \left. \frac{d_+ B(\tau, y)}{dt} \right|_{(17), v} = 0. \quad (24)$$

Уравнение (24) — неявная форма дифференциального уравнения Беллмана для простейшей задачи терминального управления

$$-\frac{\partial B(\tau, y)}{\partial \tau} = \min_{u \in U} \frac{\partial B'(\tau, y)}{\partial y} f(y, u, \tau), \quad (25)$$

которое получается из (24), если предположить, что функция $B(\tau, y)$ имеет непрерывные производные по аргументам, и воспользоваться известной формулой

$$\left. \frac{d_+ B(\tau, y)}{dt} \right|_{(17), v} = \frac{\partial B(\tau, y)}{\partial \tau} + \frac{\partial B'(\tau, y)}{\partial y} f(y, v, \tau). \quad (26)$$

Граничное условие для уравнения (25) следует из определения (18): $B(t_1, y) = \varphi(y)$.

Попутно с установлением соотношения (25) для функции Беллмана предыдущие рассуждения привели к соотношению

$$\left. \frac{d_+ B(\tau, y)}{dt} \right|_{(17), u(y, \tau)} = \min_{v \in U} \left. \frac{d_+ B(\tau, y)}{dt} \right|_{(17), v},$$

или, в подробной записи:

$$\frac{\partial B'(\tau, y)}{\partial y} f(y, u(y, \tau), \tau) = \min_{u \in U} \frac{\partial B'(\tau, y)}{\partial y} f(y, u, \tau),$$

которым удовлетворяет оптимальный закон управления в простейшей задаче терминального управления.

4. Задача оптимального управления типа Лагранжа. Рассмотрим задачу

$$\left. \begin{aligned} \frac{dx}{dt} &= f(x, u, t), \quad x(t_0) = x_0, \quad x(t_1) = x^1, \\ t &\in [t_0, t_1], \quad u(t) \in U, \quad I(u) = \int_{t_0}^{t_1} f_0(x, u, t) dt, \end{aligned} \right\} \quad (27)$$

в которой относительно входящих в нее функций сохраним предположения предыдущих пунктов, считая дополнительно, что функция $f_0(x, u, t)$ того же типа, что и компоненты функции $f(x, u, t)$.

При каждом допустимом управлении процесс (27) погрузим в семейство

$$\frac{dx}{dt} = f(x, u, t), \quad x(\tau) = y, \quad x(t_1) = x^1, \quad t \in [\tau, t_1], \quad (28)$$

и введем функцию Беллмана

$$B(\tau, y) = \min_{u} \int_{\tau}^{t_1} f_0(x, u, t) dt, \quad (29)$$

где минимум берется по всем допустимым управлениям.

Пусть $y, y \neq x^1$, — внутренняя точка множества управляемости системы (28) относительно точки x^1 для момента $t = \tau$, а управление $u(t)$, $t \in [\tau, t_1]$, доставляет минимум правой части в (29); $x(t)$, $t \in [\tau, t_1]$, — соответствующая траектория системы (28). Найдется столь малое число $\Delta, \Delta > 0$, что ни одна траектория системы (28) на промежутке $[\tau, \tau + \Delta]$ не выйдет из множества управляемости для момента $\tau + \Delta$ и не достигнет в момент $\tau + \Delta$ точки x^1 .

Из (29) следует, что вдоль оптимального управления $u(t)$, $t \in [\tau, t_1]$, при любом $\Delta \tau > 0$, $\Delta \tau < \Delta$, выполняется равенство

$$\begin{aligned} B(\tau, y) &= \int_{\tau}^{t_1} f_0(x(t), u(t), t) dt = \int_{\tau}^{\tau+\Delta\tau} f_0(x(t), u(t), t) dt + \\ &+ \int_{\tau+\Delta\tau}^{t_1} f_0(x(t), u(t), t) dt = \int_{\tau}^{\tau+\Delta\tau} f_0(x(t), u(t), t) dt + \\ &+ B(\tau + \Delta\tau, x(\tau + \Delta\tau)). \end{aligned} \quad (30)$$

Введем еще одно допустимое управление

$$\tilde{u}(t) = \begin{cases} v, & t \in [\tau, \tau + \Delta\tau), \\ \tilde{u}(t), & t \in [\tau + \Delta\tau, t_1), \end{cases} \quad (31)$$

где $v \in U$, а управление $\tilde{u}(t)$, $t \in [\tau + \Delta\tau, t_1)$, оптимально относительно состояния $\tilde{x}(\tau + \Delta\tau)$, возникшего в системе (28) после управления в первом периоде $\tilde{u}(t) = v$, $t \in [\tau, \tau + \Delta\tau)$. По построению значение критерия качества на управлении (31) равно

$$I(\tilde{u}) = \int_{\tau}^{\tau+\Delta\tau} f_0(\tilde{x}(t), v, t) dt + B(\tau + \Delta\tau, \tilde{x}(\tau + \Delta\tau))$$

и при любых указанных $\Delta\tau$, v удовлетворяет неравенству

$$B(\tau, y) \leq \int_{\tau}^{\tau+\Delta\tau} f_0(\tilde{x}(t), v, t) dt + B(\tau + \Delta\tau, \tilde{x}(\tau + \Delta\tau)). \quad (32)$$

Из (30) вдоль оптимального управления следует соотношение

$$\left. \frac{d_+ B(\tau, y)}{dt} \right|_{(28), u(\tau)} = -f_0(y, u(\tau), \tau). \quad (33)$$

С другой стороны, из (32) при любых $v \in U$ получаем

$$\frac{d_+ B(\tau, y)}{dt} = \lim_{\Delta\tau > 0, \Delta\tau \rightarrow 0} \frac{B(\tau + \Delta\tau, \tilde{x}(\tau + \Delta\tau)) - B(\tau, y)}{\Delta\tau} \geqslant \geqslant -f_0(y, v, \tau). \quad (34)$$

Объединяя (33), (34), приходим к уравнению

$$\begin{aligned} & \frac{d_+ B(\tau, y)}{dt} \Big|_{(28), u(y, \tau)} + f_0(y, u(y, \tau), \tau) = \\ & = \min_{v \in U} \left\{ \frac{d_+ B(\tau, y)}{dt} \Big|_{(28), v} + f_0(y, v, \tau) \right\} = 0, \end{aligned} \quad (35)$$

которое является неявной формой дифференциального уравнения Беллмана в задаче оптимизации типа Лагранжа.

Если предположить, что функция $B(\tau, y)$ имеет непрерывные производные, то, заменяя операции дифференцирования по времени, согласно (26), получим из (35) уравнение Беллмана в явном виде:

$$\min_{u \in U} \left\{ \frac{\partial B(\tau, y)}{\partial t} + \frac{\partial B'(\tau, y)}{\partial y} f(y, u, \tau) + f_0(y, u, \tau) \right\} = 0, \quad (36)$$

причем оптимальный закон управления $u = u(y, \tau)$ при каждом τ , y должен равняться вектору, на котором в (36) достигается минимум.

Граничное условие для уравнения (35) следует из определения (29): $B(t_1, x^1) = 0$.

З а м е ч а н и я. 1. Если задачу оптимизации типа Лагранжа заменить задачей типа Больца, т. е. в критерий качества (27) добавить слагаемое $\varphi(x(t_1))$, а правый конец траектории считать свободным, то приведенные рассуждения почти не меняются, а результат будет отличаться лишь тем, что вместо граничного условия (36) появится условие $B(t_1, y) = \varphi(y)$.

2. Если изучаемая система стационарна, функция $f_0(x, u, t)$ не зависит от t и конечный момент t_1 не фиксирован, задача оптимального управления является стационарной.

5. О других задачах оптимального управления. В пп. 1—4 исследовано несколько типичных задач оптимизации. Рассмотренные случаи представляют лишь малую часть всех возможных задач оптимизации. Даже в гл. I изучен более широкий круг задач. Мы ограничились приведенными задачами из следующих соображений: 1) техника получения дифференциального уравнения Беллмана по известному рекуррентному уравнению

Беллмана везде одна и та же, если считать, что рассматриваются внутренние точки множеств управляемости или достижимости (в зависимости от типа получаемого управления); 2) с усложнением задачи оптимизации предположение о гладкости функции Беллмана становится, как правило, все менее обоснованным и ценность дифференциальных уравнений Беллмана для решения задачи уменьшается; 3) из-за трудностей аналитического решения дифференциальных уравнений Беллмана в общем случае при практической оптимизации непрерывных процессов переходят к их дискретным аналогам, для которых соответствующие уравнения Беллмана достаточно подробно исследованы в гл. I.

§ 5. Проблема обоснования динамического программирования

Дифференциальные уравнения Беллмана в § 4 получались в качестве необходимых условий оптимальности. Но вывод всегда сопровождался дополнительным предположением о дифференцируемости функции качества, которое трудно проверить заранее, до решения задачи. Поэтому остается сомнение в законности использования уравнений, т. е. возникает проблема обоснования динамического программирования [8].

Функция Беллмана $B(\tau, y)$ может оказаться недифференцируемой уже в простейших задачах.

Пример 1. Рассмотрим задачу быстрогодействия для системы

$$\dot{x}_1 = u_1, \quad \dot{x}_2 = u_2, \quad |u_1| \leq 1, \quad |u_2| \leq 2.$$

Пусть начальное состояние имеет вид $x_1(0) = y_1$, $x_2(0) = y_2$ и требуется за минимальное время перевести траекторию в начало координат. Очевидно, что время быстрогодействия равно

$$B(y) = \max \left\{ |y_1|, \frac{|y_2|}{2} \right\} \quad (1)$$

и достигается на управлении $u_1(y) = -\text{sign } y_1$, $u_2(y) = -2 \text{ sign } y_2$,

$$\text{где } \text{sign } \alpha = \begin{cases} 1, & \alpha > 0, \\ -1, & \alpha < 0, \\ 0, & \alpha = 0. \end{cases}$$

Функция Беллмана (1) недифференцируема вдоль прямых

$$y_1 = \pm \frac{y_2}{2}. \quad (2)$$

Линии уровня её и оптимальные траектории изображены на рис. 19.

Контрпример. Хотя в предыдущем примере функция Беллмана недифференцируема на прямых (2), но вдоль любой оптимальной траектории, не начинающейся на них, она дифференцируема. Это позволяет законным образом пользоваться уравнением Беллмана почти всюду на плоскости, а исключительные траектории получить предельным переходом. В общем случае ситуация значительно сложнее.

Пример 2. Рассмотрим задачу быстрогодействия для системы

$$\ddot{x} = u, \quad x(0) = x, \quad \dot{x}(0) = y,$$

$$|u| \leq 1, \quad x(t_1) = \dot{x}(t_1) = 0. \quad (3)$$

Функцию Беллмана построим с помощью одной из первых конструкций, предложенных в теории оптимальных процессов.

Уравнение (3) эквивалентно системе

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = u. \quad (4)$$

При $u \equiv 1$ фазовый портрет системы имеет вид, изображенный на рис. 20, где каждая кривая получается смещением вдоль оси ox_1 параболы $x_1 = \frac{1}{2} x_2^2$,

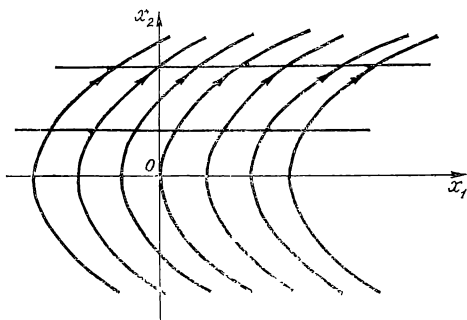


Рис. 20

Построенные кривые обладают таким свойством: если на плоскости $\{x_1, x_2\}$ провести две горизонтальные прямые, то фазовая точка $x(t)$ системы (4) пройдет отрезки кривых между прямыми за одно и то же время. Действительно, из второго уравнения (4) ($u \equiv 1$) по-

лучаем $dx_2=dt$, т. е. время движения вдоль двух дуг траекторий, концы которых имеют равные ординаты, одинаково.

При $u(t) \equiv -1$ фазовый портрет (рис. 21) получается смещением вдоль оси ox_1 параболы $x_1 = -\frac{1}{2}x_2^2$. (Полученные кривые об-

ладают тем же свойством, которое отмечено для кривых, изображенных на рис. 20.)

Из каждого семейства выделим кривую, проходящую через начало координат. Докажем, что кривая AOB (рис. 22) является линией переключения оптимального управления, т. е. оптимальный закон управления имеет вид:

$$u^0(x) = \begin{cases} +1, & \text{если } x \text{ ниже } AOB. \\ -1, & \text{если } x \text{ выше } AOB. \end{cases} \quad (5)$$

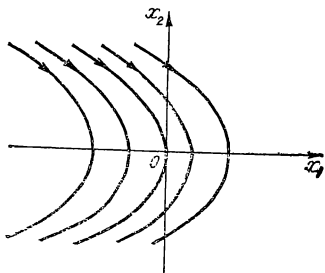


Рис. 21

Предварительно покажем, что оптимальное управление не может переключаться в верхней полуплоскости со значения -1 на значение $+1$, в нижней — с $+1$ на -1 .

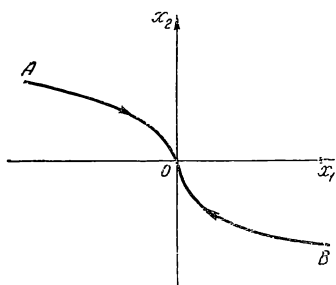


Рис. 22

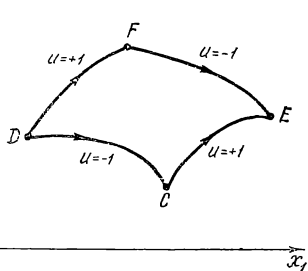


Рис. 23

Пусть это не так, и в точке C верхней полуплоскости (рис. 23) произошло переключение с -1 на $+1$. На территории, соответствующей этому управлению и проходящей через точку C , отметим две точки D и E , достаточно близкие к C . Из точки D выпустим траекторию с управлением $u \equiv +1$ и отметим траекторию, приходящую в точку E с управлением $u \equiv -1$. Из вида фазовых портретов (см. рис. 20, 21) следует, что построенные траектории пересекутся в точке F . Время движения Δt из D в E через F меньше времени движения $\Delta \bar{t}$ из D в E через C , ибо ординаты x_2, x_2 сравниваемых кривых удовлетворяют неравенству $x_2(x_1) > \bar{x}_2(x_1)$. Поэтому

$$\Delta t(DFE) = t(E) - t(D) = \int_{DFE} dt = \int_{DFE} \frac{dx_1}{x_2} < \int_{DCE} \frac{dx_1}{x_2} =$$

$$= \int_{DCE} dt = \bar{t}(E) - \bar{t}(D) = \Delta \bar{t}(DCE).$$

Теперь уже видно, что из D в E траектория DFE приходит за кратчайшее время по сравнению с другими допустимыми траекториями. Аналогично исследуется случай, когда точка C находится в нижней полуплоскости.

Доказанного факта достаточно, чтобы убедиться, что на рис. 22 изображена линия переключения оптимального управления.

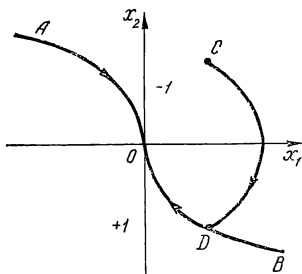


Рис. 24

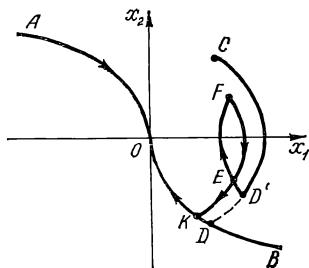


Рис. 25

Пусть C — произвольная точка (рис. 24). Траектория, соответствующая закону (5) и ведущая из C в начало координат, есть CDO .

Допустим, что траектория CDO неоптимальна. Возможны три основных варианта замены закона (5). Рассмотрим их.

а) Переключение управления с -1 на $+1$ произошло до точки D (рис. 25). Выше оси Ox_1 это переключение произойти не может в силу только что доказанного факта. Пусть оно произошло в точке D' . После точки D' траектория движется с управлением $u = +1$. Она может двигаться как до точки E , так и до точки F , а потом идти с управлением $u = -1$ до линии AOB . Траектория $D'EK$ неоптимальна, так как нарушено правило переключения оптимального управления. Траектория $D'FK$ неоптимальна, так как содержит петлю. Нетрудно проверить*, что пересход в точке D' на другое допустимое управление, отличное от $u = +1$, не меняет вывода, ибо все допустимые траектории будут выше линии $D'DO$ и поэтому приходят в точку O за большее время.

б) Возможен вариант нарушения закона (5), изображенный на рис. 26. Здесь переключение управления с -1 на $+1$ произошло в точке D'' после точки D . Траектория $CD''EO$ не может быть оптимальной, ибо включает участок FK , время движения вдоль которого равно времени движения вдоль DO траектории CDO . Опять ясно.

* Если читатель знаком с принципом максимума Понтрягина, то значения $|u| < 1$ не следует и проверять, так как оптимальное управление может быть только релейным и принимать значения ± 1 .

что введение после точки D управлений $u(t)$, $|u| \leq 1$, лишь ухудшает ситуацию.

в) При третьем варианте изменения закона (5) траектория движения изображена на рис. 27. На участке CE управление $u=+1$, далее, до точки D'' , $u=-1$. Построенная траектория $CED''O$ неоптимальна, ибо участок FK по времени движения эквивалентен участку CD траектории CDO .

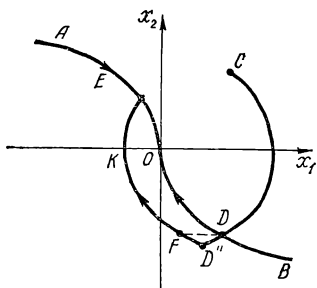


Рис. 26

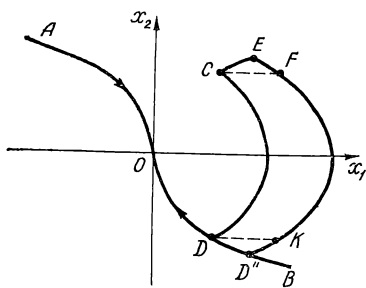


Рис. 27

Комбинация рассмотренных вариантов, а также применение управлений $u(t)$, $|u| < 1$, могут лишь «ухудшить» траекторию движения. Таким образом, закон управления (5) оптимален.

Зная закон оптимального управления, нетрудно вычислить функцию Беллмана $B(y)$. Если точка y находится ниже кривой AOB , то, полагая $u=+1$, вычисляем время движения до кривой AOB , потом, полагая $u=-1$, вычисляем время движения до точки O . Суммарное время и есть значение $B(y)$. Аналогично поступаем, если точка y лежит выше кривой AOB . Элементарный подсчет приводит к следующему результату:

$$B(y) = \begin{cases} -y_2 + 2 \sqrt{-y_1 + \frac{1}{2} y_2^2}, & \text{если } y \text{ ниже } AOB; \\ y_2 + 2 \sqrt{y_1 + \frac{1}{2} y_2^2}, & \text{если } y \text{ выше } AOB. \end{cases} \quad (6)$$

На линии AOB оба выражения дают один результат. Линия уровня функции $B(y)$ (множество точек с одним временем быстрого действия) изображена на рис. 28.

Из (6) нетрудно подсчитать, что производные $\frac{\partial B}{\partial y_1}$, $\frac{\partial B}{\partial y_2}$ равны в каждой точке линии AOB . Поскольку, по доказанному, каждая оптимальная траектория включает в себя участок линии AOB , то нет траектории, вдоль которой функция $B(y)$ дифференцируема.

Приведенные примеры показывают, что вывод уравнения Беллмана (см. § 4) нельзя в общем случае считать обоснованным. Вопрос о том, будет ли в конкретной

задаче функция Беллмана дифференцируемой, не решено до сих пор. Неизвестны и эффективные условия, при которых функция принадлежит к более широким классам (например, к классу непрерывных кусочно-гладких функций).

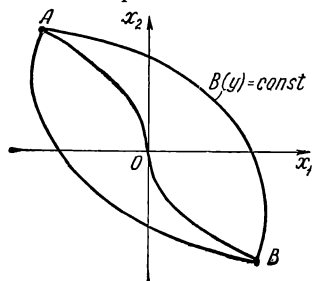


Рис. 28

§ 6. Достаточные условия оптимальности

Дифференциальное уравнение Беллмана в первое время рассматривалось как необходимое условие оптимальности. В

связи с недифференцируемостью в общем случае функции Беллмана и появлением мощного необходимого условия оптимальности, выраженного в принципе максимума Понтрягина, этот взгляд постепенно менялся, и в настоящее время уравнения Беллмана, как правило, участвуют лишь в формулировке достаточных условий оптимальности.

1. Простейшие достаточные условия оптимальности. Предположим, что $B(t, x)$ — гладкое решение уравнения Беллмана:

$$-\frac{\partial B(t, x)}{\partial t} = \min_{u \in U} \frac{\partial B'(t, x)}{\partial x} f(x, u, t) \quad (1)$$

с граничным условием

$$B(t_1, x) = \varphi(x). \quad (2)$$

Пусть функция $u(x, t)$, найденная из условия

$$\frac{\partial B'(t, x)}{\partial x} f(x, u(x, t), t) = \min_{u \in U} \frac{\partial B'(t, x)}{\partial x} f(x, u, t), \quad (3)$$

порождает единственную траекторию $x(t)$ уравнения

$$\dot{x} = f(x, u(x, t), t), \quad x(t_0) = x_0,$$

определенную на отрезке $[t_0, t_1]$, вдоль которой функция $u(t) = u(x(t), t)$ кусочно-непрерывна. Тогда $u(t) = u(x(t), t)$ — оптимальное управление в задаче

$$\dot{x} = f(x, u, t), \quad x(t_0) = x_0, \quad t \in [t_0, t_1], \quad u(t) \in U, \quad (4)$$

$$I(u) = \varphi(x(t_1)) \rightarrow \min.$$

Доказательство. Для управления $u(t) = u(x(t), t)$ в каждой точке t его непрерывности в силу (1), (3) можно записать

$$\frac{dB(t, x(t))}{dt} = \frac{\partial B(t, x(t))}{\partial t} + \frac{\partial B'(t, x(t))}{\partial x} f(x(t), u(t), t) = 0.$$

Интегрируя это тождество вдоль $x(t)$, получаем

$$B(t_1, x(t_1)) - B(t_0, x(t_0)) = 0$$

или, с учетом (2),

$$\varphi(x(t_1)) = B(t_0, x(t_0)) = B(t_0, x_0). \quad (5)$$

Пусть $\bar{u}(t)$ — другое допустимое управление и $\bar{x}(t)$ — соответствующая траектория системы (4). Для них в силу (3)

$$\begin{aligned} \frac{dB(t, \bar{x}(t))}{dt} &= \frac{\partial B(t, \bar{x}(t))}{\partial t} + \frac{\partial B'(t, \bar{x}(t))}{\partial x} f(\bar{x}(t), \bar{u}(t), t) \geq \\ &\geq \frac{\partial B(t, \bar{x}(t))}{\partial t} + \min_{u \in U} \left\{ \frac{\partial B'(t, \bar{x}(t))}{\partial x} f(\bar{x}(t), u, t) \right\} = 0. \end{aligned}$$

Интегрируя вдоль $\bar{x}(t)$, получаем

$$B(t_1, \bar{x}(t_1)) - B(t_0, \bar{x}(t_0)) \geq 0,$$

что в силу (2) эквивалентно неравенству

$$\varphi(\bar{x}(t_1)) \geq B(t_0, \bar{x}(t_0)) = B(t_0, x_0).$$

Сравнивая полученное неравенство с (5), убеждаемся, что $\varphi(x(t_1)) \leq \varphi(\bar{x}(t_1))$, т. е. управление $u(t)$ оптимально.

Аналогично формулируется и доказывается достаточное условие оптимальности в задаче типа Лагранжа.

Пусть $B(t, x)$ — гладкое решение уравнения Беллмана

$$\begin{aligned} -\frac{\partial B(t, x)}{\partial t} &= \min_{u \in U} \left[\frac{\partial B'(t, x)}{\partial x} f(x, u, t) + f_0(x, u, t) \right]; \\ B(t_1, x) &= 0, \end{aligned}$$

закон управления $u = u(x, t)$ удовлетворяет условию

$$\begin{aligned} & \frac{\partial B'(t, x)}{\partial x} f(x, u(x, t), t) + f_0(x, u(x, t), t) = \\ & = \min_{u \in U} \left[\frac{\partial B'(t, x)}{\partial x} f(x, u, t) + f_0(x, u, t) \right] \end{aligned}$$

и, кроме того: а) система

$$\frac{dx}{dt} = f(x, u(x, t), t), \quad x(t_0) = x_0$$

имеет единственное продолжимое на $[t_0, t_1]$ решение $x(t)$; б) вдоль $x(t)$ функция $u(t) = u(x(t), t)$ кусочно-непрерывна. Тогда $u(t) = u(x(t), t)$ — оптимальное управление.

2. Достаточные условия оптимальности в терминах кусочно-гладких функций Беллмана. Предположим, что непрерывная функция $B(t, x)$ обладает непрерывными производными всюду, кроме многообразия M пространства $\{x, t\}$. Если в точках, не лежащих на M , функция $B(t, x)$ удовлетворяет уравнению Беллмана и каждая траектория, порожденная допустимым законом управления, пересекает многообразие M лишь в конечном числе точек, то утверждение п. 1 остается справедливым. Доказательство проходит по той же схеме для участков траектории, лежащих между двумя соседними пересечениями многообразия M . Затем результаты складываются.

Неудобство полученного достаточного условия оптимальности состоит в трудности проверки характера пересечения траектории с многообразием. Как показал В. Г. Болтянский [8], проверку характера пересечения траекторий и многообразия M не нужно проводить, если M — кусочно-гладкое многообразие. Примерами кусочно-гладких многообразий являются: гладкая линия на плоскости, гладкая поверхность в пространстве. В общем случае кусочно-гладкое многообразие определяется следующим образом.

Пусть K — замкнутый ограниченный s -мерный выпуклый многогранник в векторном пространстве той же размерности. Предположим, что на некотором открытом множестве, содержащем многогранник K , задана n -мерная непрерывно дифференцируемая функция

$$\varphi(\xi_1, \dots, \xi_s) = (\varphi_1(\xi_1, \dots, \xi_s), \dots, \varphi_n(\xi_1, \dots, \xi_s)),$$

обладающая тем свойством, что ранг функциональной матрицы

$$\left(\frac{\partial \varphi_i}{\partial \xi_j} \right) (i = 1, \dots, n; j = 1, \dots, s)$$

в каждой точке многогранника K равен s . Если $\varphi(\xi)$ взаимно однозначно отображает K в n -мерное векторное пространство, то образ $L = \varphi(K)$ будем называть криволинейным s -мерным многогранником в n -мерном векторном пространстве.

Пусть теперь D — некоторое открытое множество в R^n . Всякое множество $M \subset D$, которое может быть представлено в виде объединения конечного или бесконечного числа криволинейных многогранников, причем с каждым замкнутым ограниченным множеством из D , пересекается лишь конечное число этих многогранников, будем называть кусочно-гладким множеством в D . Кусочно-гладкое множество называют k -мерным, если максимальная размерность составляющих его криволинейных многогранников равна k .

З а м е ч а н и е. Анализируя доказательство достаточных условий оптимальности, нетрудно заключить, что существование и непрерывность производных функций $B(t, x)$ фактически не используются. Важно только, чтобы функция $\frac{d_+ B(t, x(t))}{dt}$ вдоль допустимых законов управления была суммируемой.

Таким образом, если непрерывная функция $B(t, x)$ удовлетворяет уравнению

$$\min_{u \in U} \frac{d_+ B(t, x)}{dt} = 0, \quad B(t_1, x) = \varphi(x),$$

и вдоль каждого допустимого закона управления $u(x, t)$ функция $\frac{d_+ B(t, x)}{dt}$ определена и суммируема на $[t_0, t_1]$, то закон $u(x, t)$, удовлетворяющий условию

$$\left. \frac{d_+ B(t, x)}{dt} \right|_{u(x, t)} = \min_{u \in U} \left. \frac{d_+ B(t, x)}{dt} \right|_u,$$

является оптимальным.

3. Примеры. Непосредственное и элементарное использование результатов пп. 2, 3 проводится по следу-

ющей схеме. Выписывается уравнение Беллмана. Исходя из граничных условий, находятся функции, удовлетворяющие уравнению в различных областях пространства $\{x, t\}$, где производные функции непрерывны. Если удастся непрерывно «склеить» полученные функции, то результирующая функция, как правило, и есть требуемая функция. Для проверки этого нужно убедиться, что множество M , где разрывны производные функции, является кусочно-гладким или для всех $\{x, t\}$ функция $\frac{d_+ B(t, x)}{dt}$ суммируема.

В общем случае построение функции Беллмана — задача чрезвычайно сложная, поскольку она сводится к решению нелинейных уравнений в частных производных со специфическими граничными условиями.

Приведем элементарные примеры построения функции Беллмана.

Пример 1. Пусть

$$\dot{x} = u + x, \quad x(0) = x_0, \quad t \geq 0, \quad I(u) = \int_0^{\infty} (x^2 + u^2) dt.$$

Запишем обратное уравнение Беллмана:

$$\min_u \left[-\frac{\partial B(x)}{\partial x} (x + u) + x^2 + u^2 \right] = 0. \quad (6)$$

Решение $B(x)$ будем искать в виде квадратичной функции $B(x) = Kx^2$. Тогда из (6) следует, что $\min_u [2Kx(x+u) + x^2 + u^2] = 0$. Оче-

видно, что левая часть достигает минимума при $u = u(x) = -Kx$. Поэтому уравнение Беллмана принимает вид $(K^2 - 2K - 1)x^2 = 0$. Отсюда находим значение параметра $K = 1 + \sqrt{2}$. Второй корень уравнения отбрасываем, так как функция Беллмана с этим значением параметра противоречит ее физическому смыслу.

Таким образом, $B(x) = (1 + \sqrt{2})x^2$, $u(x) = -(1 + \sqrt{2})x$, $I(u^0) = (1 + \sqrt{2})x_0^2$, $x^0(t) = x_0 e^{-\sqrt{2}t}$, $u^0(t) = -(1 + \sqrt{2})x_0 e^{-\sqrt{2}t}$.

Пример 2. $\ddot{x} = u$, $x(0) = x_{10}$, $\dot{x}(0) = x_{20}$, $|u| \leq 1$, $I(u) = t_1$, $x(t_1) = 0$, $\dot{x}(t_1) = 0$.

Уравнение Беллмана в переменных $x_1 = x$, $x_2 = \dot{x}$ имеет вид (см. § 4 гл. II)

$$\frac{\partial B}{\partial x_1} x_2 + \min_{|u| \leq 1} \frac{\partial B}{\partial x_2} u = -1, \quad B(0) = 0.$$

В области $\frac{\partial B}{\partial x_2} > 0$ плоскости $\{x_1, x_2\}$ это уравнение таково:

$$\frac{\partial B}{\partial x_1} x_2 - \frac{\partial B}{\partial x_2} = -1. \quad (7)$$

Будем искать функцию $B(x) \equiv B_1(x_1, x_2)$ в области $\frac{\partial B}{\partial x_2} > 0$ в виде

$$B_1(x_1, x_2) = a_1 x_1^2 + a_2 x_1 x_2 + a_3 x_2^2 + a_4 x_1 + a_5 x_2. \quad (8)$$

После подстановки этой функции в уравнение (7) и приравнивания коэффициентов при одинаковых степенях получаем

$$B_1(x_1, x_2) = a(x_1 + \frac{x_2^2}{2}) + x_2, \quad (9)$$

где a — произвольное число.

Аналогично строим функцию $B_2(x) \equiv B(x)$ для области $\frac{\partial B}{\partial x_2} < 0$:

$$B_2(x) = a(x_1 - \frac{x_2^2}{2}) - x_2. \quad (10)$$

Таким образом, функция (9) определена в области

$$\frac{\partial B_1(x)}{\partial x_2} = ax_2 + 1 > 0, \quad (11)$$

а функция (10) — в области

$$\frac{\partial B_2(x)}{\partial x_2} = -ax_2 - 1 < 0. \quad (12)$$

Области (11) и (12) совпадают. Это означает, что функция Беллмана не может иметь вид (8).

«Угадать» класс функций, в котором уравнение Беллмана имеет решение, — задача трудная. Поэтому эффективнее решать уравнение методом характеристик.

§ 7. Минимизация квадратичных функционалов вдоль траекторий линейных систем

Как показывают примеры § 6, даже в простейших случаях решение уравнений Беллмана представляет серьезную трудность. Среди исключений из общего правила видное место занимает линейная задача об аналитическом конструировании регуляторов [9, 15, 22]. Частный вариант задачи уже рассматривался в примере 1 п. 3 § 6.

1. Постановка задачи. Пусть решается задача

$$\dot{x}(t) = A(t)x(t) + B(t)u, \quad x(t_0) = x_0, \quad t \in [t_0, t_1], \quad (1)$$

$$I(u) = \int_{t_0}^{t_1} [x'(t) M(t) x(t) + u'(t) R(t) u(t)] dt \rightarrow \min, \quad (2)$$

где x — n -вектор; u — r -вектор; $A(t)$, $B(t)$, $M(t)$, $R(t)$ — непрерывные матрицы, причем $M(t) \geq 0$, $R(t) > 0$ — симметрические матрицы.

Ограничения на управление $u(t)$ не накладываются, но неявно они имеются. Действительно, выражение

$$\int_{t_0}^{t_1} x' M(t) x dt \quad (3)$$

есть среднеквадратичное отклонение траектории от траектории $x(t) \equiv 0$, $t \geq 0$. Во многих практических задачах желательно, чтобы величина (3) была минимальной. Если в задаче расходы на управление существенны, то они зависят от величины

$$\int_{t_0}^{t_1} u' u dt. \quad (4)$$

Вводя «весовую» матрицу $R(t) > 0$, проблему минимизации (3) с учетом ограничений на (4) можно свести к задаче минимизации функционала (2).

2. Существование оптимального управления. При любой ограниченной непрерывной функции $u(t)$, $t \in [t_0, t_1]$, решение $x(t)$ уравнения (1) существует и

$$0 \leq \int_{t_0}^{t_1} [x'(t) M(t) x(t) + u'(t) R(t) u(t)] dt < \infty. \quad (5)$$

Поэтому можно указать последовательность управлений $u^1(t)$, $u^2(t)$, ..., вдоль которых

$$I(u^1) > I(u^2) > \dots \rightarrow \inf_u I(u). \quad (6)$$

Покажем, что существует ограниченное непрерывное управление $u^0(t)$ такое, что $I(u^0) = \inf I(u)$.

Обозначим через $L_2[t_0, t_1]$ пространство r -мерных функций, суммируемых с квадратом на $[t_0, t_1]$.

Теорема. В задаче (1), (2) существует оптимальное управление, ограниченное в $L_2[t_0, t_1]$.

Доказательство. Пусть u^1, \dots, u^n, \dots — минимизирующая последовательность допустимых управлений. Из (5), (6) следует, что последовательность функций $\{u^n(t)\}$, $t \in [t_0, t_1]$, $n = 1, 2, \dots$, равномерно ограничена в пространстве $L_2[t_0, t_1]$. Следовательно, множество $\{u^n(t)\}$, $t \in [t_0, t_1]$, $n = 1, 2, \dots$, слабо компактно в пространстве $L_2[t_0, t_1]$, т. е. из последовательности $\{u^n(t)\}$ можно выбрать подпоследовательность $\{u^n(t)\}$ такую, что для любой функции $l(t) \in L_2[t_0, t_1]$ выполняется равенство

$$\lim_{n \rightarrow +\infty} \int_{t_0}^{t_1} l'(t) u^n(t) dt = \int_{t_0}^{t_1} l'(t) u(t) dt, \quad (7)$$

где $u(t)$ — некоторая функция из $L_2[t_0, t_1]$.

Пусть $x(t)$ — соответствующее $u(t)$ решение уравнения (1). В силу формулы Коши для решений уравнения (1) и равенства (7) получаем при любом $t \in [t_0, t_1]$ соотношения

$$\begin{aligned} x(t) &= F(t, t_0) x_0 + \int_{t_0}^t F(t, \tau) B(\tau) u(\tau) d\tau = F(t, t_0) x_0 + \\ &+ \lim_{n \rightarrow +\infty} \int_{t_0}^t F(t, \tau) B(\tau) u^n(\tau) d\tau = \lim_{n \rightarrow +\infty} x^n(t). \end{aligned}$$

Здесь $F(t, \tau)$, $t \in [t_0, t_1]$ — решение матричного дифференциального уравнения

$$\begin{aligned} \frac{\partial F(t, \tau)}{\partial \tau} &= -F(t, \tau) A(\tau), \quad F(t, t-0) = E, \\ F(t, \tau) &\equiv 0, \quad \tau \geq t + 0; \end{aligned}$$

$x^n(t)$ — траектория системы (1), соответствующая управлению $u^n(t)$, $t \geq t_0$.

Осталось показать, что

$$\begin{aligned} &\int_{t_0}^{t_1} [x'(t) M(t) x(t) + u'(t) R(t) u(t)] dt - \\ &- \lim_{n \rightarrow +\infty} \int_{t_0}^{t_1} [(x^n(t))' M(t) x^n(t) + (u^n(t))' R(t) u^n(t)] dt \leq 0, \end{aligned} \quad (8)$$

ибо из (6) следует обратное неравенство.

Из теоремы Лебега вытекает, что

$$\int_{t_0}^{t_1} x'(t) M(t) x(t) dt = \lim_{n \rightarrow \infty} \int_{t_0}^{t_1} (x^n(t))' M(t) x^n(t) dt.$$

Поэтому для доказательства неравенства (8) достаточно доказать, что

$$\int_{t_0}^{t_1} [u'(t) R(t) u(t)] dt \leq \lim_{n \rightarrow \infty} \int_{t_0}^{t_1} [(u^n(t))' R(t) u^n(t)] dt. \quad (9)$$

Воспользуемся теоремой Мазура о слабом пределе. Рассмотрим любую сходящуюся подпоследовательность последовательности $\left\{ \int_{t_0}^{t_1} [(u^n(t))' R(t) u^n(t)] dt \right\}, n=1, 2, \dots,$

и для этой подпоследовательности докажем (не переходя к новым обозначениям), что

$$\int_{t_0}^{t_1} [u'(t) R(t) u(t)] dt \leq \lim_{n \rightarrow \infty} \int_{t_0}^{t_1} [(u^n(t))' R(t) u^n(t)] dt.$$

Этим, очевидно, будет доказано неравенство (9).

Для любого $\varepsilon > 0$ укажем такое натуральное число N , что при всех $k \geq N$

$$\int_{t_0}^{t_1} [(u^k(t))' R(t) u^k(t)] dt \leq \lim_{n \rightarrow \infty} \int_{t_0}^{t_1} [(u^n(t))' R(t) u^n(t)] dt + \varepsilon. \quad (10)$$

По теореме Мазура найдется последовательность $v^m(t)$, $m=1, 2, \dots$, выпуклых комбинаций элементов $u^k(t)$, $k \geq N$, которая сходится к $u(t)$ по норме $L_2[t_0, t_1]$. Функции $v^m(t)$, $t \in [t_0, t_1]$, можно представить в виде

$$v^m(t) = \sum_{i=N}^{p(m)} \alpha_i u^i(t), \quad m = 1, 2, \dots, \quad (11)$$

где $\alpha_i \geq 0$, $\sum_{i=N}^{p(m)} \alpha_i = 1$.

В силу соотношений (10), (11) и выпуклости функции $\varphi(u) = u' R u$ имеем

$$\begin{aligned}
& \int_{t_0}^{t_1} [(v^m(t))' R v^m(t)] dt = \\
& = \int_{t_0}^{t_1} \left\{ \left[\sum_{i=N}^{p(m)} \alpha_i u^i(t) \right]' R \left[\sum_{i=N}^{p(m)} \alpha_i u^i(t) \right] \right\} dt \leq \\
& \leq \int_{t_0}^{t_1} \left\{ \sum_{i=N}^{p(m)} \alpha_i [(u^i(t))' R u^i(t)] \right\} dt = \\
& = \sum_{i=N}^{p(m)} \alpha_i \int_{t_0}^{t_1} [(u^i(t))' R u^i(t)] dt \leq \lim_{n \rightarrow \infty} \int_{t_0}^{t_1} [(u^n(t))' R u^n(t)] dt + \varepsilon.
\end{aligned}$$

Итак, для всех $m=1, 2, \dots$ справедливо неравенство

$$\int_{t_0}^{t_1} [(v^m(t))' R v^m(t)] dt \leq \lim_{n \rightarrow \infty} \int_{t_0}^{t_1} [(u^n(t))' R u^n(t)] dt + \varepsilon. \quad (12)$$

Последовательность $\{v^m(t)\}$, $m=1, 2, \dots$, сходится к функции $u(t)$, $t_0 \leq t \leq t_1$, по норме пространства $L_2[t_0, t_1]$. Следовательно (переходя в случае необходимости к подпоследовательности), можно считать, что последовательность $\{v^m(t)\}$, $m=1, 2, \dots$, сходится к $u(t)$ почти всюду на $[t_0, t_1]$.

Переходя в (12) к пределу, в силу леммы Фату получаем неравенство

$$\int_{t_0}^{t_1} u'(t) R u(t) dt \leq \lim_{n \rightarrow \infty} \int_{t_0}^{t_1} (u^n(t))' R u^n(t) dt + \varepsilon,$$

которое и завершает доказательство теоремы.

З а м е ч а н и е. Оптимальное управление $u^0(t)$, $t \in [t_0, t_1]$, найдено в классе функций из $L_2[t_0, t_1]$. Однако, как следует из п. 3, оно всегда может быть выбрано непрерывным.

3. Уравнение Беллмана. Поскольку в классе функций из $L_2[t_0, t_1]$ оптимальное управление $u^0(t)$ существует, оно удовлетворяет уравнению Эйлера

$$\frac{\delta I(u^0)}{\delta u(t)} = 0,$$

или, в подробной записи,

$$B'(t) \int_t^{t_1} F'(\tau, t) M(\tau) x^0(\tau) d\tau + R(t) u^0(t) = 0.$$

Функционал $I(u)$ является строго выпуклым, что следует из его вида и того, что $R(t) > 0$. Поэтому минимум $I(u)$ достигается на единственной функции $u^0(t)$:

$$\begin{aligned} u^0(t) &= -R^{-1}(t) B'(t) \int_t^{t_1} F'(\tau, t) M(\tau) x^0(\tau) d\tau = \\ &= -R^{-1}(t) B'(t) L(t) x^0(t), \end{aligned}$$

где $L(t)$ — решение уравнения (15) (см. ниже).

Таким образом, решение линейно по x_0 и непрерывно по t , так что в результате его подстановки в (2) получаем определенно-положительную квадратичную форму относительно x_0 .

Если процесс (1) погрузить в семейство

$$\dot{x} = A(t)x + B(t)u, \quad x(\tau) = y, \quad t \in [\tau, t_1],$$

и ввести функцию Беллмана

$$B(\tau, y) = \min_{u[\tau, t_1]} \int_{\tau}^{t_1} [x'(t) M(t)x(t) + u'(t) R(t)u(t)] dt,$$

то в силу доказанного очевидно, что $B(\tau, y)$ — определенно-положительная квадратичная форма по y с гладкими коэффициентами относительно τ . Таким образом, мы показали, что применение динамического программирования в рассмотренной задаче обосновано.

Уравнение Беллмана в задаче п. 1 имеет вид

$$\begin{aligned} -\frac{\partial B(t, x)}{\partial t} &= \min_u \left[\frac{\partial B'(t, x)}{\partial x} \{A(t)x + B(t)u\} + \right. \\ &\quad \left. + x' M(t)x + u' R(t)u \right], \\ B(t_1, x) &= 0. \end{aligned} \tag{13}$$

Нетрудно прямой подстановкой проверить, что уравнение (13) имеет решение

$$B(t, x) = x' L(t) x, \tag{14}$$

где $L(t)$ — решение уравнения Риккати

$$\frac{dL}{dt} = -L(t)A(t) - A'(t)L(t) + L(t)B(t)R^{-1}(t)B'(t)L(t) - M(t), \quad L(t_1) = 0. \quad (15)$$

Уравнение (15) имеет решение $L(t)$, продолжимое на $[t_0, t_1]$. Так как правая часть уравнения (15) удовлетворяет условию локального существования и единственности решения, достаточно проверить, что решение $L(t)$, начинающееся в точке $L(t_1) = 0$, продолжимо на $[t_0, t_1]$.

По определению величины $B(\tau, y)$ выполняются неравенства

$$0 < B(\tau, y) = y' L(\tau) y = \min_{u \in [\tau, t_1]} \int_{\tau}^{t_1} [x'(t) M(t) x(t) + u'(t) R(t) u(t)] dt \leq \int_{\tau}^{t_1} \bar{x}'(t) M(t) \bar{x}(t) dt \leq l = \text{const},$$

где $\bar{x}(t)$ — решение уравнения (1) при $u(t) = 0$.

Поскольку $\|L(\tau)\| = \max_y \frac{B(\tau, y)}{\|y\|}$, то из последнего неравенства следует равномерная ограниченность решения $L(t)$, а следовательно, и продолжимость этого решения на весь отрезок.

Таким образом, функция (14) — решение уравнения Беллмана, и в силу теоремы из § 6 закон управления

$$u(x, t) = -R^{-1}(t)B(t)L(t)x, \quad (16)$$

на котором достигается минимум в правой части (13), является оптимальным.

Квадратичность функции Беллмана (14) и линейность оптимального закона (16) относительно состояния x — главные особенности задачи (1), (2).

4. Процессы на бесконечном интервале. Существование допустимого закона управления. Рассмотрим задачу

$$\dot{x} = Ax + bu, \quad x(0) = x_0, \quad (17)$$

$$I(u) = \int_0^{\infty} (x'(t) M x(t) + R u^2(t)) dt \rightarrow \min, \quad (18)$$

где x — n -вектор; u — скаляр; A, M — постоянные матри-

цы; $M > 0$; b — постоянный вектор; R — положительное число.

Закон $u = u(x)$ назовем допустимым, если он порождает единственное решение уравнения $\dot{x} = Ax + bu(x)$, $x(0) = x_0$, продолжимое на ось $[0, \infty)$, вдоль которого функционал (18) ограничен. Найдем условия существования допустимого управления. Пусть

$$\lambda_1, \dots, \lambda_n \quad (19)$$

собственные числа матрицы A . Вычислим векторы

$$b, Ab, \dots, A^{n-1}b. \quad (20)$$

Пусть среди векторов (20) первые k линейно независимы, но $-A^k b = \alpha_k b + \dots + \alpha_1 A^{k-1} b$. Введем полином $\varphi(\lambda) = \lambda^k + \alpha_1 \lambda^{k-1} + \dots + \alpha_k$.

Теорема. Для того чтобы в задаче (17), (18) существовало допустимое управление, необходимо и достаточно, чтобы все числа из (19), которые имеют неотрицательные действительные части, были корнями уравнения

$$\varphi(\lambda) = 0. \quad (21)$$

Доказательство. Необходимость. Не нарушая общности, можем предполагать, что A и b имеют вид

$$A = \begin{Bmatrix} A_1 & A_3 \\ \vdots & \vdots \\ 0 & A_2 \end{Bmatrix}, \quad b = \begin{Bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{Bmatrix}, \quad A_1 = \begin{Bmatrix} 0 & 0 & 0 & \dots & 0 & -\alpha_k \\ 1 & 0 & 0 & \dots & 0 & -\alpha_{k-1} \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -\alpha_1 \end{Bmatrix} \quad (22)$$

Пусть задача (17), (18) имеет решение, но существует число λ_s из (19), $\operatorname{Re} \lambda_s \geq 0$, которое не является корнем уравнения (21). Тогда λ_s является собственным значением матрицы A_2 , так как

$$\begin{aligned} \det |\lambda E - A| &= \det |\lambda E_1 - A_1| \cdot \det |\lambda E_2 - A_2| = \\ &= \varphi(\lambda) \det |\lambda E_2 - A_2|. \end{aligned} \quad (23)$$

Таким образом, в силу (22) система (17) содержит часть, которая не является асимптотически устойчивой и на которую не действует управление. Отсюда следует, что функционал (18) не ограничен сверху, ибо $M > 0$, $R > 0$. Необходимость доказана.

Достаточность. Пусть условие теоремы выполняется, а задача (17), (18) не имеет допустимого управления для некоторого начального состояния x_0 . Поскольку все собственные числа матрицы A с неотрицательными действительными частями находятся среди корней уравнения (21), то матрица A_2 асимптотически устойчива.

Рассмотрим систему $\dot{y} = A_1 y + b_1 u$, где $b'_1 = \{1, 0, \dots, 0\}$. Будем искать такой закон управления

$$u = p' y, \quad (24)$$

чтобы матрица $A_1 + b_1 p'$ была асимптотически устойчивой. Поскольку

$$\begin{aligned} \det |\lambda E_1 - A_1 - b_1 p'| &= \lambda^k + \lambda^{k-1} (\alpha_1 - p_1) + \\ &+ \dots + \lambda^{k-m} (\alpha_m - p_m - \sum_{i=1}^{m-1} \alpha_{m-i} p_i) + (\alpha_k - p_k - \\ &- \sum_{i=1}^{k-1} \alpha_{k-i} p_i), \end{aligned}$$

то для любых чисел μ_1, \dots, μ_k , $\operatorname{Re} \mu_i < 0$, можно подобрать вектор p такой, чтобы эти числа стали собственными для матрицы $A_1 + b_1 p'$.

Считая, что вектор y представляет первые k компонентов вектора x , делаем вывод, что при выборе закона (24) система (17) в силу (23) асимптотически устойчива, т. е.

$$\|x(t)\| \leq \alpha e^{-\beta t}, \quad \beta > 0, \quad t \rightarrow \infty. \quad (25)$$

Подставляя управление $u(t) = p'[x(t)]^k$ в (18), нетрудно в силу (25) подсчитать, что функционал сходится, т. е. управление $u(t)$ допустимо. Теорема доказана.

З а м е ч а н и е. Элементарный анализ теоремы показывает, что в ее формулировке слова «допустимое управление» можно заменить на «допустимый закон управления».

5. Существование оптимального закона управления [18]. Пусть $K(t_1, t) = L(t)$, где $L(t)$, $t \in [t_0, t_1]$, — решение уравнения Риккати (15).

Теорема. Для того чтобы в задаче (17), (18) существовал линейный оптимальный закон управления, необ-

ходимо и достаточно, чтобы существовал конечный предел

$$\lim_{t_1 \rightarrow +\infty} K(t_1, 0) = K. \quad (26)$$

Оптимальный закон управления имеет вид

$$u(x) = -R^{-1} b' K x. \quad (27)$$

Доказательство проводится по схеме доказательства аналогичного утверждения из п. 4 § 11 гл. I.

6. Эффективное построение оптимального закона. Из приведенного утверждения следует, что для построения оптимального закона управления в задаче (17), (18) нужно осуществить следующие операции: 1) решить вспомогательную задачу (1), (2), построенную по параметрам задачи (17), (18) при помощи уравнения Риккати (15); 2) найти предел функции (26) при $t_1 \rightarrow \infty$. Если предел не существует, задача (17), (18) не имеет решения. Если предел существует, формула (27) доставляет оптимальный закон управления в задаче (17), (18). На практике в качестве предельного значения функции $K(t_1, t_0)$ берут ее установившееся значение.

§ 8. Проблема стабилизации

Результаты решения задачи аналитического конструирования регуляторов имеют большое значение при стабилизации неустойчивых систем.

1. Общая постановка проблемы стабилизации. Задачу в общем случае можно сформулировать следующим образом.

Дана система

$$\dot{x} = f(x, u, t); \quad t \geq t_0, \quad f(0, 0, t) = 0, \quad (1)$$

тривиальное движение $x(t) \equiv 0$ которой при $u(t) \equiv 0$ не является асимптотически устойчивым. Требуется найти закон управления $u = u(x, t)$, $u(0, t) = 0$, при котором система

$$\dot{x}(t) = f(x(t), u(x(t), t), t), \quad t \geq t_0,$$

асимптотически устойчива.

2. Схема применения динамического программирования. Задаче стабилизации системы (1) поставим в соответствие следующую задачу оптимизации:

$$\begin{aligned} \dot{x} &= f(x, u, t), \quad x(t_0) = x_0, \\ I(u) &= \int_0^{\infty} f_0(x(t), u(t), t) dt \rightarrow \min, \end{aligned} \quad (2)$$

где $f_0(x, u, t)$ — скалярная функция, определенно-положительная в смысле Ляпунова по аргументу x .

Используя прямое погружение системы (2), введем функцию Беллмана:

$$B(\tau, y) = \min_{v[\tau, +\infty)} \int_{\tau}^{+\infty} f_0(y(t), v(t), t) dt. \quad (3)$$

Уравнение для $B(\tau, y)$ получается обычным путем и имеет вид

$$-\frac{\partial B(\tau, y)}{\partial \tau} = \min_v [f_0(y, v, \tau) + \frac{\partial B'(\tau, y)}{\partial y} f(y, v, t)]. \quad (4)$$

Пусть $B(\tau, y)$ — непрерывное решение уравнения (4), удовлетворяющее для каждого y условиям:

$$B(\tau, y) \rightarrow 0 \text{ при } \tau \rightarrow \infty, \quad B(\tau, y) \geq C(y) > 0, \quad y \neq 0.$$

Изучим поведение функции $B(\tau, y)$ вдоль оптимальной траектории, порожденной законом управления $u(x, t)$, найденным из соотношения

$$\begin{aligned} f_0(x, u(x, t), t) + \frac{\partial B'(\tau, x)}{\partial x} f(x, u(x, t), t) &= \\ &= \min_v [f_0(x, v, t) + \frac{\partial B'(\tau, x)}{\partial x} f(x, v, t)]. \end{aligned} \quad (5)$$

Из (3) при любом $\Delta t, \Delta t > 0$, имеем

$$\begin{aligned} B(t, x(t)) &= \int_t^{+\infty} f_0(x(\tau), u(x(\tau), \tau), \tau) d\tau = \\ &= \int_{t+\Delta t}^{+\infty} f_0(x(\tau), u(x(\tau), \tau), \tau) d\tau + \end{aligned}$$

$$+ \int_t^{t+\Delta t} f_0(x(\tau), u(x(\tau), \tau), \tau) d\tau = B(t + \Delta t, x(t + \Delta t)) + \\ + \int_t^{t+\Delta t} f_0(x(\tau), u(x(\tau), \tau), \tau) d\tau.$$

Поэтому $\left. \frac{d_+ B(x, t)}{dt} \right|_{(2), u(t, x)} = -f_0(x, u(x, t), t).$

Таким образом, полная производная по времени функции $B(t, x)$ вдоль траектории системы

$$\dot{x} = g(x, t), \quad g(x, t) = f(x, u(x, t), t) \quad (6)$$

есть функция определенно-отрицательная. Отсюда заключаем (в силу теоремы Ляпунова об асимптотической устойчивости [16]), что закон управления $u=u(x, t)$, вычисленный в силу (5), стабилизирует систему (1) до асимптотически устойчивой системы (6).

3. Решение проблемы стабилизации для линейных систем. Результаты § 7 можно непосредственно использовать для решения проблемы стабилизации линейных систем.

Рассмотрим систему

$$\dot{x} = Ax \quad (7)$$

с матрицей A , не все собственные числа которой имеют отрицательные действительные части. В этом случае система (7) не является асимптотически устойчивой и найдутся начальные состояния x_0 , из которых движения $x(t)$ не стремятся к состоянию $x=0$.

Допустим, что система имеет входное устройство, через которое можно ею управлять, и общее уравнение системы и входного устройства

$$\dot{x} = Ax + Bu, \quad (8)$$

где u — r -мерный вектор управления.

Проблема линейной стабилизации системы состоит в построении $(r \times n)$ -матрицы Q такой, что закон управления

$$u(x) = Qx \quad (9)$$

стабилизирует систему (7), т. е. система

$$\dot{x} = (A + BQ)x, \quad (10)$$

полученная из (8) и (9), асимптотически устойчива. Критерий стабилизируемости содержится в теореме п. 4 § 7.

Для эффективного построения закона (9), решающего проблему стабилизации, обратимся к задаче (17), (18) из § 7. Пусть $R > 0$, $M > 0$ — симметрические матрицы, выбор которых выходит за рамки рассматриваемой задачи. Решим задачу (17), (18) § 7 с данными матрицами R и M (см. п. 4 § 7). Согласно (17) § 7, решение имеет вид

$$u(x) = Qx, \quad (11)$$

где $Q = -R^{-1}B'K$.

Подставим функцию (11) в уравнение (8), и пусть $x(t)$ — решение полученного уравнения с некоторым начальным условием. Исследуем поведение функции Белмана вдоль $x(t)$:

$$\begin{aligned} B(x(t)) &= \int_t^{+\infty} [x'(\tau) M x(\tau) + u'(x) R u(x)] d\tau = \\ &= \int_t^{t+\Delta t} [x'(\tau) M x(\tau) + u'(x) R u(x)] d\tau + \\ &+ \int_{t+\Delta t}^{+\infty} [x'(\tau) M x(\tau) + u'(x) R u(x)] d\tau. \end{aligned}$$

Таким образом, для $\Delta t > 0$

$$\begin{aligned} B(x(t + \Delta t)) - B(x(t)) &= - \int_t^{t+\Delta t} [x'(\tau) M x(\tau) + \\ &+ u'(x) R u(x)] d\tau \end{aligned}$$

или

$$\left. \frac{dB(x)}{dt} \right|_{(10)} = -x'(M + KBR^{-1}B'K)x.$$

Поскольку функции $-x'(M + KBR^{-1}B'K)x$ — определено-положительные формы, то, следовательно, выполнены

все условия теоремы Ляпунова об асимптотической устойчивости [16] и поэтому $\|x(t)\| \rightarrow 0$, $t \rightarrow +\infty$, для любых начальных возмущений.

Построенный закон управления (11) не только стабилизирует систему, но одновременно доставляет минимум функционалу

$$I(u) = \int_0^{+\infty} [x' M x + u' R u] dt,$$

т. е. переходный процесс в системе (10) обладает определенными экстремальными свойствами. Выбирая те или иные матрицы R и M , можно получить различные переходные процессы.

§ 9. Оптимизация систем с инвариантной нормой

Второй класс систем, при оптимизации которых динамическое программирование применяется эффективно, представляют системы с инвариантной нормой.

1. Системы с инвариантной нормой. Рассмотрим задачу быстрогодействия для системы

$$\dot{x} = f(x) + bu, \quad x(0) = x_0, \quad x(t_1) = 0, \quad \|u\|^2 = (u', u) \leq 1, \quad (1)$$

где x, u — n -векторы; b — число; $b > 0$.

Система (1) называется системой с инвариантной нормой, если при $u(t) \equiv 0$ вдоль любого решения $x(t)$ системы выполняется условие

$$x'(t)x(t) \equiv \text{const}, \quad (2)$$

где постоянная зависит лишь от x_0 .

Необходимым и достаточным условием выполнения условия (2) является тождество

$$\frac{d}{dt} x' x = 2x' f(x) \equiv 0. \quad (3)$$

Обратное уравнение Беллмана для задачи быстрогодействия в системе (1) имеет вид (см. § 4)

$$\min_{\|u\| \leq 1} \frac{\partial B'(x)}{\partial x} (f(x) + bu) = -1, \quad B(0) = 0. \quad (4)$$

После взятия операции \min из (4) получаем

$$\frac{\partial B'(x)}{\partial x} f(x) - b \parallel \frac{\partial B(x)}{\partial x} \parallel = -1, \quad B(0) = 0. \quad (5)$$

Решением уравнения (5) является функция

$$B(x) = \frac{\parallel x \parallel}{b}. \quad (6)$$

Действительно, эта функция имеет непрерывные производные $\frac{\partial B}{\partial x} = \frac{x}{\parallel x \parallel b}$, и поэтому уравнение (5) сводится к тождеству

$$\frac{x' f(x)}{\parallel x \parallel b} - b \parallel \frac{x}{\parallel x \parallel b} \parallel = \frac{x' f(x)}{\parallel x \parallel b} - 1 = -1,$$

так как в силу (3) $x' f(x) = 0$. Поскольку для функции (6) выполнено и граничное условие $B(0) = 0$, то все условия теоремы из п. 1 § 6 выполнены и поэтому закон управления

$$\begin{aligned} u^0(x) &= -\frac{\partial B(x)}{\partial x} / \parallel \frac{\partial B(x)}{\partial x} \parallel = \\ &= -\frac{x}{\parallel x \parallel}, \end{aligned} \quad (7)$$

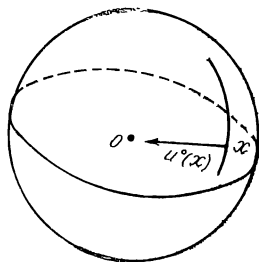


Рис. 29

доставляющий минимум в задаче (4), является оптимальным. Время быстрогодействия из точки x_0 равно

$$B(x_0) = \frac{\parallel x_0 \parallel}{b}.$$

Таким образом, если при выключенном управлении траектории системы (1) движутся по поверхности сферы (2), оптимальным по быстродействию будет управление (7), направленное в центр сферы (рис. 29) и равное по норме единице. Для всех точек сферы фиксированного радиуса R время быстрогодействия одинаково и равно R/b .

2. Обобщение. Обобщением условия (3) является неравенство

$$x' f(x) = h(\parallel x \parallel) < b \parallel x \parallel. \quad (8)$$

Если сохранить остальные предположения п. 1, то решением уравнения (4) при условии (8) будет функция

$$B(x) = \int_0^{\|x\|} \frac{y \, dy}{by - h(y)}.$$

Действительно, подставляя ее в уравнение (4), получаем

$$\begin{aligned} \frac{\|x\|}{b\|x\| - h(\|x\|)} \cdot \frac{x'f(x)}{\|x\|} - b \frac{\|x\|}{b\|x\| - h(\|x\|)} = \\ = \frac{h(\|x\|) - b\|x\|}{b\|x\| - h(\|x\|)} = -1. \end{aligned}$$

Граничное условие (4) выполняется тривиально:
 $B(0) = 0$.

В силу теоремы из п. 1 § 6 закон управления (7) опять является оптимальным.

3. Системы с инвариантной обобщенной нормой. Условие (2) можно заменить на $C(x) = \text{const}$, где $C(x)$ — определено-положительная функция. Другими словами, рассматривается инвариантная функция $C(x)$ вместо инвариантной нормы $\|x\|$. Рассмотрим частный случай, когда $C(x) = x'Cx$ — определено-положительная квадратичная форма.

Пусть $\lambda_1 > 0, \dots, \lambda_n > 0$ — собственные значения матрицы C . Существует неособое преобразование $x = Py$, такое, что матрица $Q = P'CP$ имеет диагональный вид:

$$Q = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}.$$

Поэтому после еще одного преобразования

$$z_1 = \frac{1}{\sqrt{\lambda_1}} y_1, \dots, z_n = \frac{1}{\sqrt{\lambda_n}} y_n$$

получаем систему с инвариантной нормой, ибо $z'(t)z(t) = \text{const}$.

Пример. Вращение тела около неподвижной точки описывается уравнениями:

$$\left. \begin{aligned} I_1 \dot{x}_1 &= (I_2 - I_3) x_2 x_3 + b u_1 \\ I_2 \dot{x}_2 &= (I_3 - I_1) x_3 x_1 + b u_2 \\ I_3 \dot{x}_3 &= (I_1 - I_2) x_1 x_2 + b u_3, \quad u_1^2 + u_2^2 + u_3^2 \leq 1 \end{aligned} \right\}, \quad (9)$$

где x_1, x_2, x_3 — угловые скорости тела; I_1, I_2, I_3 — моменты инерции относительно осей x_1, x_2, x_3 ; u_1, u_2, u_3 — управляющие вращающие моменты относительно этих осей.

При $u_1 = u_2 = u_3 = 0$ в системе (9) инвариантной является функция $C(x) = I_1^2 x_1^2 + I_2^2 x_2^2 + I_3^2 x_3^2$. Поэтому, используя приведенные выше преобразования, получаем следующий закон оптимального управления:

$$u^0(z) = - \frac{z}{\|z\|}, \quad z_1 = \frac{1}{I_1} x_1, \quad z_2 = \frac{1}{I_2} x_2, \quad z_3 = \frac{1}{I_3} x_3.$$

При этом управлении система (9) с начальными угловыми скоростями x_{10}, x_{20}, x_{30} успокаивается за минимальное время $\|z_0\|/b$.

Вычислительные аспекты динамического программирования всегда привлекали внимание специалистов, занятых решением практических задач оптимизации. Для задач невысокой размерности стандартные методы динамического программирования могут быть реализованы на ЭЦВМ. В более сложных случаях возможности оперативной памяти современных вычислительных машин затрудняют решение, однако исключительная ценность результата (абсолютный минимум), гарантируемого методом, с одной стороны, а с другой — быстрое усовершенствование ЭЦВМ заставляют исследователей вновь и вновь обращаться к динамическому программированию, искать пути расширения его возможностей. К настоящему времени значительно расширился арсенал вычислительных методов, основанных на динамическом программировании. Опубликован ряд работ [5—10, 24], в которых приведены решения задач, еще недавно недоступных методу динамического программирования. По мере дальнейшего развития ЭЦВМ удельный вес подобных задач будет, несомненно, расти.

§ 1. Стандартная процедура

Все модификации вычислительных схем динамического программирования исходят из стандартной процедуры и имеют целью преодолеть те или иные ее недостатки. Поэтому естественно ознакомление с вычислительными методами начать со стандартной процедуры.

1. Дискретизация динамических систем. В общем случае стандартная процедура динамического программирования предполагает запись уравнений процесса в дискретной форме. Уравнения процесса в виде

$$x(t+1) = f(x(t), u(t), t) \quad (1)$$

являются наиболее удобными. Многие дискретные процессы могут быть записаны в этой форме или преобразованы к ним с помощью несложных операций (см. гл. I).

Квантование по времени непрерывных систем во многих случаях тоже приводит к уравнениям (1). Например, с уравнением

$$\dot{y} = f(y, v, \tau), \quad \tau \in [t_0, t_1], \quad (2)$$

поступают следующим образом. Отрезок $[t_0, t_1]$ разбивают на N частей точками $t_0 = \tau_0 < \tau_1 < \dots < \tau_N = t_1$. Полагая при $\tau_i < \tau < \tau_{i+1}$

$$\Delta \tau_i = \tau_{i+1} - \tau_i, \quad \dot{y}(\tau) \approx \frac{y(\tau_{i+1}) - y(\tau_i)}{\Delta \tau_i}, \quad (3)$$

от уравнения (2) приходят к рекуррентному соотношению

$$y(\tau_{i+1}) = y(\tau_i) + \Delta \tau_i f(y(\tau_i), v(\tau_i), \tau_i), \quad (4)$$

которое можно записать в форме (1), если ввести обозначения:

$$x(t) = y(\tau_i), \quad u(t) = v(\tau_i).$$

Аналогичный переход можно сделать и для уравнения с запаздывающим аргументом

$$\dot{y}(\tau) = f(y(\tau), y(\tau - h), u(\tau), \tau).$$

С помощью замены (3), (4) сначала получают рекуррентное соотношение

$$x(t+1) = x(t) + \Delta \tau_i f(x(t), x(t-k), u(t), t),$$

где k таково, что $\tau_k \leq \tau - h < \tau_{k+1}$. Затем, вводя вектор $X(t) = \{x(t), x(t-1), \dots, x(t-k)\}$, получают уравнение в форме (1).

Замена (3) является простейшей формой дискретизации непрерывных систем. Можно использовать более точные аппроксимации производных.

Процедура перехода от непрерывных систем к дискретным естественна при решении задач на цифровых вычислительных устройствах дискретного действия. При достаточно мелком разбиении отрезка $[t_0, t_1]$ решения уравнения (1) сохраняют основные свойства решений (2).

Форма (4) рекуррентных уравнений специфична при дискретизации непрерывных систем и специально используется в некоторых вычислительных методах.

2. Формальное описание процедуры. Рассмотрим простейшую задачу терминального управления

$$x(t+1) = f(x(t), u(t), t), \quad x(t_0) = x_0, \quad t \in [t_0, t_1],$$

$$I(u) = \varphi(x(t_1)) \rightarrow \min, \quad u(t) \in U(t).$$

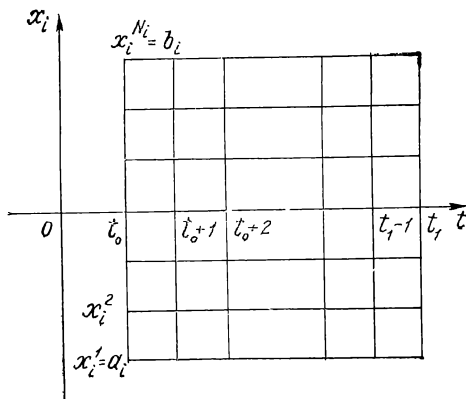


Рис. 30

Обратное уравнение Беллмана для нее имеет вид (см. § 8 гл. I):

$$B(t, x) = \min_{u \in U(t)} B(t+1, f(x, u, t)), \quad B(t_1, x) = \varphi(x). \quad (5)$$

Каждую координату x_i вектора x квантуем по уровню, т. е. будем считать, что x_i может принимать значения

$$x_i^1, x_i^2, \dots, x_i^{N_i}. \quad (6)$$

Числа (6) можно считать упорядоченными. Как правило, они выбираются следующим образом. Пусть $[a_i, b_i]$ — отрезок возможных изменений координаты x_i . Выбираем число $\Delta x_i > 0$ и полагаем

$$x_i^k = a_i + (k-1) \Delta x_i, \quad k = 1, 2, \dots, N_i, \quad N_i = \frac{b_i - a_i}{\Delta x_i} + 1.$$

В результате квантования в пространстве $\{x, t\}$ образуется фазовая сетка (рис. 30). Перед началом

процедуры функция $\varphi(x)$ вычисляется в узлах сетки с абсциссой t_1 и заносится в оперативную память. Таким образом, перед первым шагом процедуры имеем картину, изображенную на рис. 31, где в узлах, обведенных кружком, функция Беллмана $B(t_1, x)$ известна. Теперь можно приступить к подсчету по формуле (5) значений $B(t_1-1, x)$ в узлах фазовой сетки с абсциссой t_1-1 . В общем случае, когда нет никакой специальной информа-

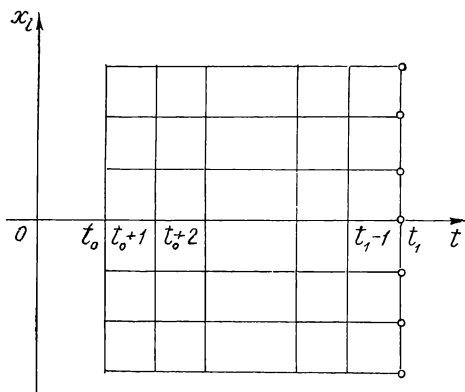


Рис. 31

ции об оптимальном управлении, поступают следующим образом. Путем квантования координат управления множество значений $U(t_1-1)$ допустимых управлений покрывается сеткой (рис. 32). Пусть $u^1, \dots, u^{M(t_1-1)}$ — узлы сетки, лежащие в $U(t_1-1)^*$. Для вычисления значения функции $B(t_1-1, x)$ в узле $x^k = \{x_1^k, \dots, x_n^k\}$ с абсциссой t_1-1 (рис. 33) вычисляем векторы

$$f(x^k, u^1, t_1-1), f(x^k, u^2, t_1-1), \dots, f(x^k, u^{M(t_1-1)}, t_1-1). \quad (7)$$

В общем случае не все векторы (7) попадут в узлы фазовой сетки. Поэтому значения функции $B(t_1, x)$ в точках (7) находятся с помощью интерполирования известных значений в узлах с абсциссой t_1 . Среди вычисленных

* Узлы u^1, u^2, \dots зависят от t , однако для упрощения обозначений эта зависимость указана лишь для последнего узла. Аналогичное соглашение принято и в дальнейшем.

значений $B(t_1, x)$ находится минимальное, которое в силу (5) равно $B(t_1-1, x^k)$ и поэтому записывается в оперативную память. Точка $u^l = u(x^k, t_1-1)$, на которой достигается минимум, также запоминается. Подобным образом перебираются все узлы сетки с абсциссой t_1-1 . В результате в оперативной памяти появится запись всех значений $B(t_1-1, x)$ в этих узлах, будут известны и значения $u(x, t_1-1)$. Значения $B(t_1, x)$ могут быть убраны из оперативной памяти.

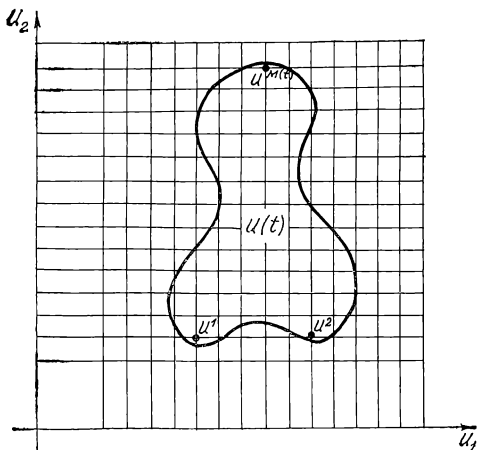


Рис. 32

Следующие шаги процедуры в точности повторяют первый, и процесс продолжается вычислением функций $B(t, x)$, $u(x, t)$ в узлах фазовой сетки в направлении убывания аргумента t . Через t_1-t_0 шагов получатся функции $B(t_0, x)$, $u(x, t_0)$ в узлах фазовой сетки с абсциссой $t=t_0$.

Исходную задачу оптимизации можно считать решенной, ибо $B(t_0, x_0)$ — минимальное значение критерия качества, а оптимальное управление и траектория находятся по записанным функциям $u(x, t_0)$, $u(x, t_0+1)$, ..., $u(x, t_1-1)$ элементарно:

$$u^0(t_0) = u(x_0, t_0), \quad x^0(t_0+1) = f(x_0, u^0(t_0), t_0);$$

$$u^0(t_0+1) = u(x^0(t_0+1), t_0+1), \quad x^0(t_0+2) = \\ = f(x^0(t_0+1), u^0(t_0+1), t_0+1); \dots;$$

$$u^0(t_1 - 1) = u(x^0(t_1 - 1), t_1 - 1),$$

$$x^0(t_1) = f(x^0(t_1 - 1), u^0(t_1 - 1), t_1 - 1).$$

3. Анализ процедуры. Проанализируем полученный результат. Отметим сначала достоинства: а) процедура не использует никаких аналитических свойств функций $f(x, u, t)$, $\varphi(x)$ и множества U — они могут быть заданы любым способом; б) управление $u^0(t)$, $t \in [t_0, t_1 - 1]$, доставляет критерию качества абсолютный минимум (а не относительный); в) оптимальное управление по-

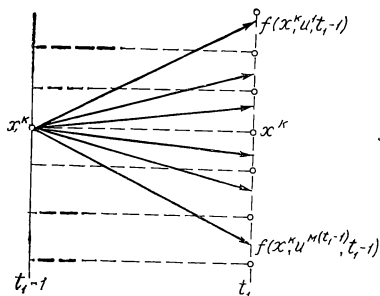


Рис. 33

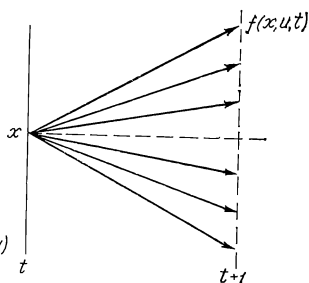


Рис. 34

лучилось не для одной точки x_0 , а для всех узлов фазовой сетки с абсциссами $t \geq t_0$, что позволяет проанализировать чувствительность решения к изменению начальных условий.

Одним из недостатков процедуры являются чрезмерные требования к объему оперативной памяти. На каждом шаге нужно хранить в памяти функции $B(t, x)$, $B(t+1, x)$, $u(x, t), \dots, u(x, t_1 - 1)$ в узлах фазовой сетки с фиксированной абсциссой t и все значения $u(x, \tau)$ при $\tau \geq t$. В совокупности число запомненных точек к моменту $t_1 - k$ не более

$$(2 + rk) \prod_{i=1}^n N_i.$$

Поскольку числа N_i , $100 \leq N_i \leq 150$, можно считать обычными, то уже для $n=4$ получаем число, превосходящее $100^4 = 10^8$, что с точки зрения объема оперативной памяти ЭЦВМ очень велико. Беллман этот недостаток процедуры назвал «проклятием размерности». Другой

недостаток — опасность расширяющейся сетки. Этот недостаток состоит в следующем. Нетрудно построить процессы, для которых в каждой точке $\{x, t\}$ совокупность векторов (7) имеет вид, изображенный на рис. 34, когда концы векторов попадают и выше и ниже горизонтальной линии. В такой ситуации начальная фазовая сетка может оказаться недостаточной для завершения

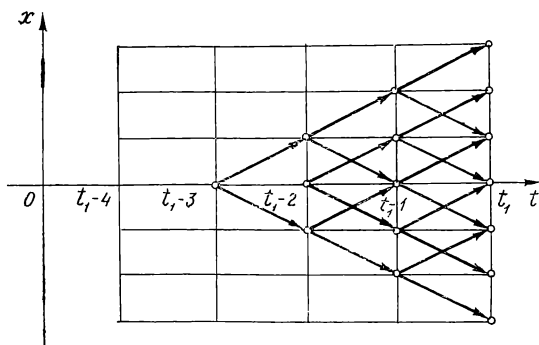


Рис. 35

процесса вычислений, и вычисления функции $B(t, x)$ оборвутся на шаге t_k из-за того, что в памяти нет значений $B(t_k+1, x)$ (рис. 35). Во всех узлах, не обведенных кружком, значения $B(t, x)$ и $u(x, t)$ нельзя вычислить, если фазовая сетка в момент $t=t_i$ недостаточно велика. Увеличение сетки ведет к увеличению необходимой памяти.

4. О количестве вычислений. Простейшим способом решения задачи является, как известно, перебор всех допустимых управлений, который гарантирует нахождение абсолютного минимума. Сколько же операций требуется для решения задачи п. 2 методом перебора? Графически решение задачи методом перебора выглядит так, как изображено на рис. 36. Из точки x_0 откладывают $M(t_0)$ векторов, ведущих в точки

$$f(x_0, u^1, t_0), \dots, f(x_0, u^{M(t_0)}, t_0). \quad (8)$$

Из каждой полученной точки с абсциссой $t=t_0+1$ откладывают $M(t_0+1)$ векторов, ведущих в $f(x, u^1, t_0+1), \dots, f(x, u^{M(t_0+1)}, t_0+1)$, где x — одна из точек (8). Этот процесс продолжается до прямой $t=t_1-1$. Откла-

дывая из концов x векторов, дошедших до прямой $t=t_1-1$, пучок из $M(t_1-1)$ векторов, ведущих в

$$f(x, u^1, t_1-1), \dots, f(x, u^{M(t_1-1)}, t_1-1),$$

получаем все точки на прямой t_1 , в которых нужно вычислить значения функции $\varphi(x)$ и выбрать среди них минимальное. Ясно, что количество точек, в которых

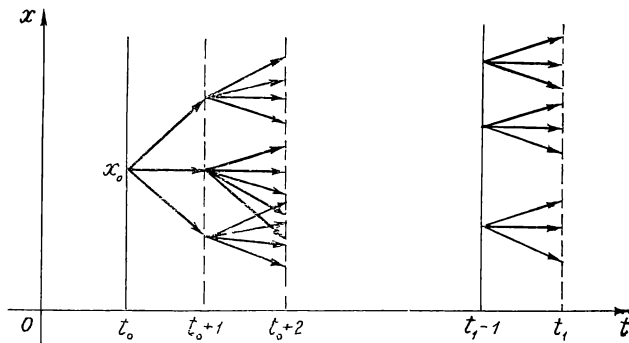


Рис. 36

вычисляется функция $\varphi(x)$, а следовательно, и общее число вычислений равно $\prod_{t=t_0}^{t_1-1} M(t)$.

Теперь на простом примере покажем, что перебор может оказаться лучшим методом вычисления абсолютного минимума в задаче п. 2. Нетрудно подобрать числовые параметры, при которых все допустимые траектории имеют вид, изображенный на рис. 37.

В методе перебора требуется 8 вычислений функции $\varphi(x)$. Если применять стандартную процедуру, то уже на первом шаге при вычислении функции $B(t, x)$ нужно тоже 8 вычислений функции $\varphi(x)$. К этим вычислениям добавляются дополнительные при нахождении функций $B(1, x)$, $B(0, x)$. Таким образом, в данном примере стандартная процедура приводит к большему числу операций. Это и понятно, ибо стандартная процедура доставляет дополнительную информацию об оптимальных процессах при других начальных условиях, а не только при данном x_0 . Однако в исходной постановке задачи нет требований о получении этой дополнительной информации. Если бы они были, стандартная процедура

оказалась бы и в данном примере оптимальной по числу операций вычисления. Но заранее сформулировать дополнительные требования указанного типа бывает трудно. В приведенном примере метод перебора оптимален, если нет специальных данных о свойствах функций $f(x, u, t)$, $\varphi(x)$ и множества $u(t)$.

Специфика примера, приведенного на рис. 37, состоит в том, что траектории системы не «перемешиваются»:

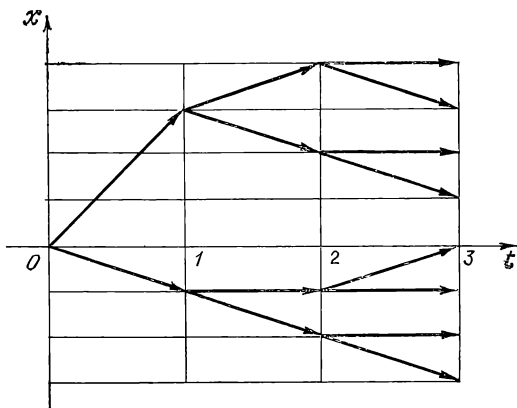


Рис. 37

каждая точка множества достижимости получается одним допустимым управлением. Однако такая ситуация нетипична в теории управления. Как правило, траектории многих систем «сильно перемешаны»: большинство точек достижимости может быть реализовано с помощью допустимых управлений. В подобных случаях стандартная процедура может оказаться значительно выгоднее метода перебора. Почти крайний вариант изображен на рис. 38. Не зная о специальной структуре системы, при применении метода перебора в этом примере нужно проверить $3^4=81$ допустимое управление. С другой стороны, в стандартной процедуре, если выбрана сетка, указанная на рисунке, вычисление оптимального управления потребует на каждом шаге 9, а всего $9 \times 4=36$ операций. С увеличением числа шагов t_1 выигрыш растет очень быстро.

То обстоятельство, что чем больше перемешиваемость траекторий, тем лучше работает стандартная процедура,

можно объяснить и таким образом. Метод перебора в начальной постановке оперирует со значениями, которые принимает в разные моменты времени допустимое управление, т. е., другими словами, он задан в пространстве управлений. В отличие от метода перебора стандартная процедура проводится в пространстве состояний, которое зависит от конкретной системы. И здесь имеется возможность выигрыша. Кроме того, поскольку

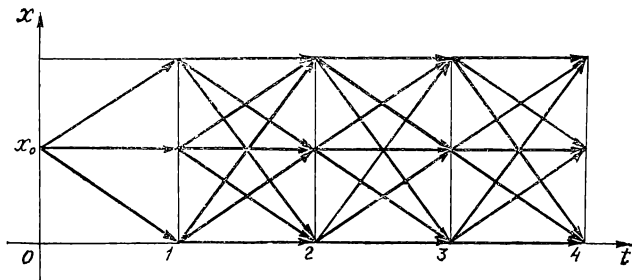


Рис. 38

каждое состояние может быть получено различными допустимыми управлениями, то исключение из рассмотрения одного состояния ведет к исключению и всех управлений, породивших его. Поэтому, хотя стандартная процедура и есть вариант метода перебора, но, заданная в пространстве состояний, она может дать огромную экономию в числе операций.

Описанный эффект просто иллюстрируется на следующем примере. Пусть требуется найти минимум функции $\varphi(u) = u_1 + u_2 + \dots + u_n$ при ограничении $\psi(u) = \psi(u_1 + \dots + u_n) \leq 1$. На первый взгляд это — многомерная задача, но легко сообразить, что введение новой скалярной переменной $x = u_1 + \dots + u_n$ сводит ее к одномерной задаче: $x \rightarrow \min, \psi(x) \leq 1$. В стандартной процедуре реализуется неявно замена такого типа, причем существенно используется структура системы управления.

5. Задача Лагранжа. Опишем стандартную процедуру для задачи типа Лагранжа. Для этой задачи

$$x(t+1) = f(x(t), u(t), t), \quad x(t_0) = x_0, \quad x(t_1) = x^1,$$

$$u(t) \in U(t), \quad I(u) = \sum_{t=t_0}^{t_1-1} f_0(x(t), u(t), t) \rightarrow \min,$$

обратное уравнение Беллмана (см. § 9 гл. I) имеет вид

$$B(t, x) = \min_{u \in U(t)} [f_0(x, u, t) + B(t+1, f(x, u, t))],$$

если $x \in F(t), \quad x \notin \partial^0 F(t), \quad (9)$

$$B(t, x) = \min_{u \in U_1(x, t)} [f_0(x, u, t) + B(t+1, f(x, u, t))],$$

если $x \in \partial^0 F(t), \quad (10)$

$$B(t_1, x^1) = 0. \quad (11)$$

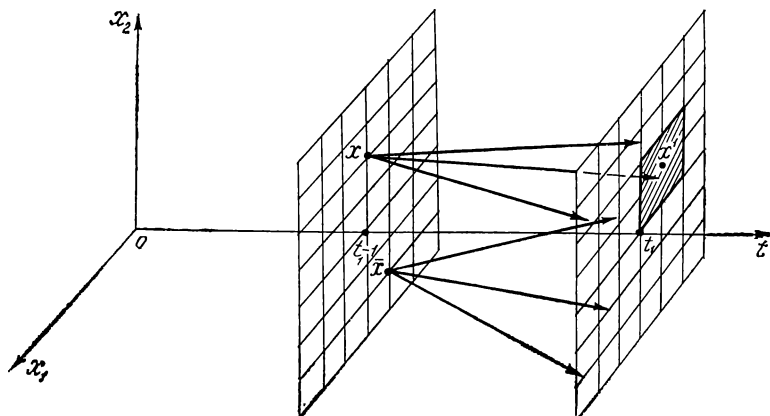


Рис. 39

Пусть фазовая сетка построена (см. п. 2). Точки x , удовлетворяющие уравнению $f(x, u, t_1-1) = x^1$ при некотором $u \in U(t_1-1)$, составляют множество $F(t_1-1)$. Функция $B(t_1-1, x)$ определена лишь для точек $F(t_1-1)$.

Прежде чем начать вычисления, определим множества $F^0(t)$ и понятие « $x \in F^0(t)$ ». При $t=t_1-1$ множество $F^0(t)$ определим как множество узлов x фазовой сетки с $t=t_1-1$, таких, что при некотором векторе $u^i, i=1, \dots, M(t_1-1)$, вектор $f(x, u^i, t_1-1)$ попадает в одну из клеток, лежащих в сечении $t=t_1$ фазовой сетки и содержащих узел x^1 . На рис. 39 точка $x \in F^0(t_1-1)$, но $x \notin F(t_1-1)$, если $M(t_1-1)=3$. Допустим, что множество $F^0(\tau)$ построено. Определим множество $F^0(\tau-1)$. Узел x сечения $t=\tau-1$ отнесем к множеству $F^0(\tau-1)$: $x \in F^0(\tau-1)$, если существ-

вует вектор u^i , $i=1, \dots, M(\tau-1)$, такой, что вектор $f(x, u^i, \tau-1)$ попадает в одну из клеток сечения $t=\tau$ фазовой сетки, которая содержит хотя бы один узел из $F^0(\tau)$. Такой же смысл придается верхнему индексу ⁰ в обозначениях других множеств.

В сечении $t=t_1$ фазовой сетки функция $B(t_1, x)$ определена в одной точке, и в силу (11) $B(t_1, x^1)=0$. Рассмотрим сечение $t=t_1-1$. Для каждого узла $x \in F^0(t_1-1)$ вычисляем значение функции $B(t_1-1, x)$ по формулам (9) или (10) в зависимости от того, не принадлежит или принадлежит точка x пограничному слою множества $F^0(t_1-1)$. Пограничный слой $\partial^0 F^0(t)$ и множества $U_1(x, t)$ определяются так же, как в § 8 гл. I, с заменой множества F на F^0 . Поскольку при фиксированном $x \in F^0(t_1-1)$ векторы

$$f(x, u^1, t_1-1), \dots, f(x, u^{M(t_1-1)}, t_1-1)$$

не все, возможно, являются узлами сечения $t=t_1$ фазовой сетки, то при вычислении значений $B(t_1-1, f(x, u^i, t_1-1))$ используется интерполяция и экстраполяция в пределах одной клетки. Далее переходим к моменту $t=t_1-2$. В узле x сечения $t=t_1-2$ фазовой сетки вычисляются значения

$$f(x, u^1, t_1-2), \dots, f(x, u^{M(t_1-2)}, t_1-2). \quad (12)$$

Если полученные векторы принадлежат множеству $F^0(t_1-1)$, то по формуле (9) вычисляются $B(t_1-2, x)$, $u(x, t_1-2)$ и вместе с $x \in F^0(t_1-2)$ запоминаются. Если некоторые из векторов (12) не принадлежат множеству $F^0(t_1-1)$, то значения $B(t_1-2, x)$, $u(x, t_1-2)$ вычисляются по формуле (10) и вместе с $x \in \partial^0 F^0(t_1-2)$ запоминаются. Если ни один из векторов (12) не принадлежит множеству $F^0(t_1-1)$, точка x исключается. Таким образом перебираются все узлы фазовой сетки в сечении $t=t_1-2$. После этого значения $B(t_1-1, x)$ можно убрать из памяти и перейти к следующему шагу, который в точности повторяет только что описанный. Окончательная картина изображена на рис. 40. Заштрихованное множество — область определения функций $B(t, x)$, $u(x, t)$. На рисунке исходной фазовой сетки оказалось достаточно для завершения процесса вычислений. В других случаях из-за эффекта расширяющейся сетки для

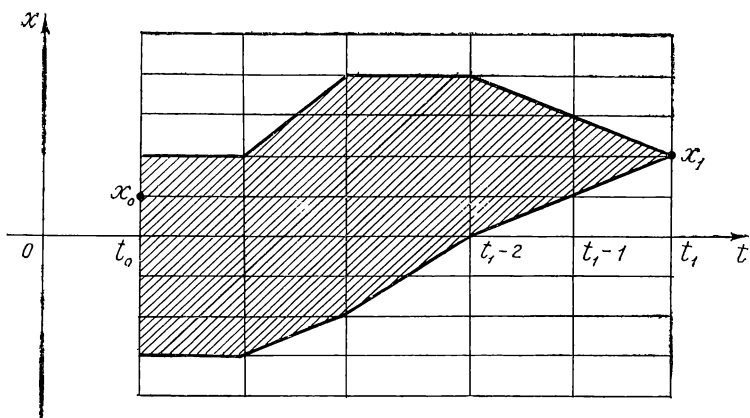


Рис. 40

завершения процесса может потребоваться увеличение исходной сетки.

6. Стандартная процедура для прямого уравнения Беллмана. Рассмотрим задачу оптимизации (см. § 9 гл. I):

$$x(t-1) = f^-(x(t), u(t), t), \quad x(t_1) = x^1, \quad t \in [t_0 + 1, t_1], \quad (13)$$

$$u(t) \in U^-(t), \quad I(u) = \varphi(x(t_0)) \rightarrow \min. \quad (14)$$

Прямое уравнение Беллмана в этой задаче

$$B^-(\tau, y) = \min_{u \in U^-(\tau)} B^-(\tau-1, f^-(y, u, \tau)), \quad (15)$$

$$B^-(t_0, y) = \varphi(y).$$

Прежде чем начать вычисления, строим фазовую сетку в пространстве $\{x, t\}$ и находим значения $u^1, \dots, u^{M(t)}$ построением сетки для множества $U^-(t)$ (см. п. 1).

В сечении $t=t_0$ фазовой сетки вычисляем $B^-(t_0, y) = \varphi(y)$ и запоминаем. Далее рассматриваем сечение $t=t_0+1$ фазовой сетки. Для каждого узла этого сечения строим векторы

$$f^-(y, u^1, t_0+1), \dots, f^-(y, u^{M(t_0+1)}, t_0+1) \quad (16)$$

и подсчитываем на них значения функции $B^-(t_0, y)$. Среди полученных значений выбираем минимальное, которое

в силу (15) равно $B^-(t_0+1, y)$. Если не все векторы (16) попали в узлы сечения $t=t_0$ фазовой сетки, то для вычисления $B^-(t_0+1, y)$ применяем интерполирование. Вектор u^l , при котором функция $B^-(t_0, f(y, u^l, t_0))$ достигла минимума, запоминаем, ибо $u^l = u(y, t_0+1)$. Перебирая все узлы сечения $t=t_0+1$ фазовой сетки, находим функцию $B^-(t_0+1, y)$, которую записываем вместо $B^-(t_0, y)$, и закон управления $u(y, t_0+1)$. Для оперативных вычислений достаточно знать лишь значения $B^-(t, y)$. Далее процесс вычислений повторяется и приводит (если не оборвется из-за опасности расширяемости сетки) к функции $B^-(t_1, y)$ и закону $u(y, t_1)$, определенным на узлах сечения $t=t_1$ фазовой сетки. Если точка x^1 лежит в области определения построенной функции $B^-(t_1, y)$, то величина $B^-(t_1, x^1)$ — минимальное значение критерия качества в задаче (13), (14), вектор $u^0(t_1) = u(x^1, t_1)$ — значение оптимального управления в момент $t=t_1$. Остальные значения оптимального управления находятся по построенному оптимальному закону и уравнению (13).

З а м е ч а н и е. Задача (13), (14) несимметрична по краевым условиям и критерию. Она выбрана нами лишь для того, чтобы упростить объяснение сущности стандартной процедуры в приложении к прямому уравнению Беллмана. В симметричных задачах можно эффективно пользоваться как прямым, так и обратным уравнением Беллмана. Нет строго обоснованных критериев для выбора того или иного уравнения. Однако для «борьбы» с опасностью расширяющейся сетки и получения устойчивости счета важно учитывать характер развития изучаемого процесса при увеличении времени.

7. Учет фазовых ограничений. Вернемся к задаче терминального управления (см. п. 2), но предположим дополнительно, что имеются ограничения $x(t) \in G(t)$, $t \in [t_0+1, t_1-1]$. Через $G^0(\tau)$ обозначим множество узлов сечения $t=\tau$ фазовой сетки, принадлежащей множеству $G(\tau)$ (рис. 41; кружками обозначены узлы из $G^0(t)$). Будем считать, что сечения фазовой сетки настолько мелки, что любая компонента множества $G(t)$ содержит хотя бы один узел сетки.

Обратное уравнение Беллмана в этом случае имеет вид (см. § 8 гл. I):

$$B(t, x) = \min_{u \in U(t)} B(t+1, f(x, u, t)),$$

если

$$\bar{x} \in \partial^0 F(t), \quad x \in F(t), \quad (17)$$

$$B(t, x) = \min_{u \in U(x, t)} B(t+1, f(x, u, t)), \text{ если } x \in \partial^0 F(t). \quad (18)$$

При $t=t_1$ функция $B(t, x)$ табулируется в точках множества $F^0(t_1)$, которое строится с учетом ограничения $G(t_1-1)$ (рис. 42; точки отмечены кружками):

$$F^0(t_1) = \{y : y = f(x, u, t_1 - 1), u \in U^0(t_1 - 1), x \in G^0(t_1 - 1)\}.$$

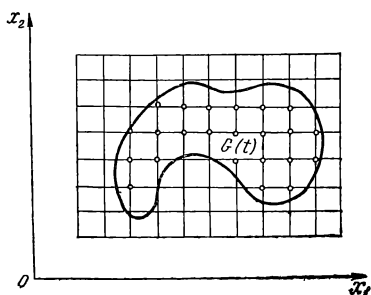


Рис. 41

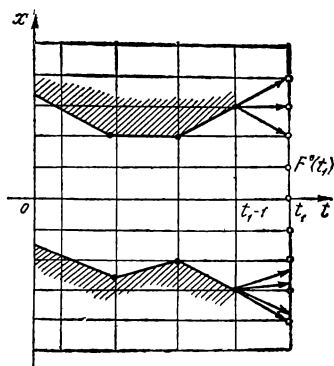


Рис. 42

По определению множества $F^0(t)$, $F^0(t_1-1) = G^0(t_1-1)$, $\partial^0 F(t_1-1)$ — пустое множество.

Функции $B(t_1-1, x)$, $u(x, t_1-1)$ вычисляются в точках $F^0(t_1-1)$ по обычным правилам (см. п. 2) и запоминаются.

Далее переходим к построению множеств $F^0(t_1-2)$, $\partial^0 F^0(t_1-2)$ (рис. 43). По определению

$$F^0(t_1-2) = \{x : y = f(x, u, t_1-2), y \in F^0(t_1-1), \text{ при некотором } u \in U^0(t_1-2); x \in G^0(t_1-2)\}^0.$$

К пограничному слою $\partial^0 F^0(t_1-2)$ относятся те точки x из $F^0(t_1-2)$, для которых существуют такие

$$u \in U^0(t_1-2), \text{ что } f(x, u, t_1-2) \in \bar{F}^0(t_1-1).$$

Для точек $x \in \partial^0 F^0(t_1-2)$ функция $B(t_1-2, x)$ строится по формуле (17). Для $x \in \partial^0 F^0(t_1-2)$ используется формула (18). Функция $u(x, t_1-2)$ запоминается, а функция $B(t_1-2, x)$ записывается на место функции $B(t_1-1, x)$.

По аналогии с $F^0(t_1-2)$, $\partial^0 F^0(t_1-2)$ строятся множества $F^0(t)$, $\partial^0 F^0(t)$, $t < t_1-2$, а на них вычисляются функции $B(t, x)$, $u(x, t)$ (рис. 44; множество $\partial^0 F^0(t)$ дважды заштриховано).

На примере этой задачи обнаруживается одно из основных достоинств стандартной процедуры динамического программирования: чем больше ограничений в задаче, чем они жестче, тем эффективнее процедура. Метод перебора, описанный в п. 3, в задачах с фазовыми ограничениями становится очень трудоемким из-за того,

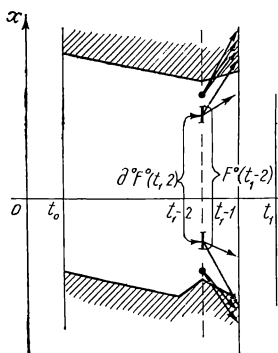


Рис. 43

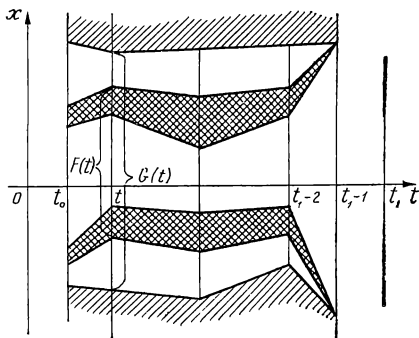


Рис. 44

что ограничения на координаты очень трудно записать через ограничения на управления. Отмеченное неожиданное свойство динамического программирования наводит на мысль о введении искусственных ограничений для приближенного решения задачи (см. § 10). Объем табулирования зависит от «жесткости» ограничений $G(t)$ и может быть сведен до уровня объема оперативной памяти современных ЭЦВМ.

З а м е ч а н и е. Для дискретных процессов (4), полученных с помощью квантования по времени из непрерывных, с уменьшением $\Delta \tau_l$ (с увеличением точности аппроксимации) пограничный слой $\partial^0 F(t)$ уменьшается. Однако при этом не обязательно, чтобы $F(t) \rightarrow G(t)$.

§ 2. Использование штрафных функций

Как видно из результатов § 2 гл. I, введение дополнительных ограничений, как правило, повышает размерность аргумента функции Беллмана. Из классической

теории условного минимума известно, что решение задач типа рассмотренных в § 2 гл. I можно в определенном смысле свести к решению на безусловный минимум, если ввести множители Лагранжа. В последние годы идея такой замены интенсивно развивается в рамках теории штрафных функций. Использование ее существенно понижает размерность задачи в смысле динамического программирования.

1. Множители Лагранжа в выпуклом программировании. Пусть рассматривается задача

$$\left. \begin{aligned} f(x) &\rightarrow \min, \\ g(x) &\leq 0, \quad x \geq 0, \end{aligned} \right\} \quad (1)$$

где $f(x)$ — скалярная выпуклая функция n -векторного аргумента x ; $g(x)$ — m -мерная функция с выпуклыми компонентами; символ \leq обозначает покомпонентные неравенства.

Предположим, что в задаче (1) ограничения удовлетворяют условию регулярности: $\bar{g}(x) < 0$ для некоторого $\bar{x} \geq 0$.

Теорема Куна—Таккера утверждает: для того чтобы вектор $x^0 \geq 0$ был решением задачи (1), необходимо и достаточно, чтобы существовал m -вектор $\lambda \geq 0$ такой, что x^0 — решение задачи

$$f(x) + \lambda' g(x) \rightarrow \min, \quad x \geq 0. \quad (2)$$

Таким образом, задача на условный минимум (1) сведена к задаче на безусловный минимум (2).

Как следует из результатов § 1, 2 гл. I, размерность задачи (1) равна m (если выполняются дополнительные ограничения на структуру функций $f(x)$, $g(x)$), а задачи (2) — нулю. Следовательно, введение множителя λ (множитель Лагранжа) резко уменьшает размерность исследуемой задачи. В этом состоит основная цель введения множителей Лагранжа при решении задач оптимизации с помощью динамического программирования.

В общем случае сведение задачи на условный минимум к задаче на безусловный минимум невозможно (т. е. для произвольных функций $f(x)$, $g(x)$ теорема Куна—Таккера неверна). Однако некоторое обобщение множителей Лагранжа возможно и в общем случае. В ос-

нове этого обобщения лежит следующая интерпретация множителей Лагранжа. Если вектор $\lambda \geq 0$ задан и минимизируется выражение (2), то слагаемое $\lambda' g(x)$ можно интерпретировать как штраф за нарушение ограничения $g(x) \leq 0$, а λ — как относительный штраф (штраф за единицу нарушения). Ясно, что чем больше вектор λ , тем меньше «желание» нарушить ограничение. В задаче выпуклого программирования по теореме Куна—Таккера существует конечный относительный штраф λ , для которого при минимизации функции (2) ограничение в (1) не нарушается. Невозможность распространения теоремы Куна—Таккера на общий случай означает, что имеются задачи типа (1), в которых, как бы ни был высок относительный штраф λ , остается «выгода» от нарушения ограничения в (1).

2. Штрафные функции. Развитием только что приведенной интерпретации множителей Лагранжа является метод штрафных функций, в котором штраф возрастает не обязательно линейно с нарушением ограничения. Функция штрафа может задаваться произвольным образом. Например, если рассматривается задача

$$f(x) \rightarrow \min, g(x) = 0, \quad (3)$$

то вводится скалярная функция $\nu(g)$ такая, что $\nu(g) > 0$ при $g \neq 0$, и вместо задачи (3) рассматривается задача

$$f(x) + \lambda \nu(g(x)) \rightarrow \min, \lambda > 0. \quad (4)$$

При весьма общих условиях можно показать, что если $\lambda \rightarrow \infty$, то решение $x(\lambda)$ задачи обладает свойством $g(x(\lambda)) \rightarrow 0$.

Таким образом, выбирая достаточно большие числа λ , можно, решая задачу (4), сколько угодно точно удовлетворить ограничениям задачи (3). Кроме того, если $x(\lambda)$ — решение задачи (4) при фиксированном λ , то оно является решением следующей задачи на условный минимум: $f(x) \rightarrow \min, \nu(g(x)) = \nu(g(x(\lambda)))$.

Действительно, если найдется вектор \bar{x} такой, что

$$f(\bar{x}) < f(x(\lambda)), \quad \nu(g(\bar{x})) = \nu(g(x(\lambda))),$$

то

$$f(\bar{x}) + \lambda \nu(g(\bar{x})) < f(x(\lambda)) + \lambda \nu(g(x(\lambda))),$$

что противоречит предположению об оптимальности вектора $x(\lambda)$.

Если ограничения имеют вид $g(x) \leq 0$, то введением свободных переменных x_{n+1}, \dots, x_{n+m} их можно свести к ограничениям типа равенства

$$g_i(x) + x_{n+i}^2 = 0, \quad i = \overline{1, m},$$

или ввести штрафную функцию $\mu(g)$ со свойством $\mu(g) > 0$ при $g > 0$ и вместо задачи $f(x) \rightarrow \min, g(x) \leq 0$, рассмотреть задачу

$$f(x) + \lambda \mu(g(x)) \rightarrow \min. \quad (5)$$

Решение $x(\lambda)$ задачи (5) при любом $\lambda, \lambda \geq 0$, обладает следующими свойствами:

а) вектор $x(\lambda)$ доставляет минимум в задаче $f(x) \rightarrow \min, \mu(g(x)) \leq \mu(g(x(\lambda)))$; действительно, если при некотором \bar{x} выполняются соотношения $f(\bar{x}) < f(x(\lambda))$, $\mu(g(\bar{x})) \leq \mu(g(x(\lambda)))$, то при $\lambda \geq 0$ будет выполняться и неравенство

$$f(\bar{x}) + \lambda \mu(g(\bar{x})) < f(x(\lambda)) + \lambda \mu(g(x(\lambda))),$$

что противоречит определению вектора $x(\lambda)$;

б) если $\bar{\lambda} > \lambda$ и $x(\bar{\lambda}), x(\lambda)$ — соответствующие решения задачи (5), то $g(x(\bar{\lambda})) \leq g(x(\lambda))$; действительно, по определению элементов $x(\bar{\lambda}), x(\lambda)$:

$$f(x(\lambda)) + \lambda g(x(\lambda)) \leq f(x(\bar{\lambda})) + \lambda g(x(\bar{\lambda})),$$

$$f(x(\bar{\lambda})) + \bar{\lambda} g(x(\bar{\lambda})) \leq f(x(\lambda)) + \bar{\lambda} g(x(\lambda));$$

складывая эти неравенства, получаем

$$(\bar{\lambda} - \lambda) g(x(\bar{\lambda})) \leq (\bar{\lambda} - \lambda) g(x(\lambda)),$$

откуда следует доказываемое свойство.

В методе штрафных функций не обязательно освобождаться от всех ограничений, а, учитывая то, что динамическое программирование с успехом справляется с задачами, размерность которых 2, можно оставлять од-

но или два ограничения. Это тем более целесообразно, что введение штрафных функций связано с использованием больших чисел λ , а это может вызвать неудобства при работе на ЭЦВМ.

Выше метод штрафных функций пояснялся на примере простых задач минимизации функций конечного числа переменных. Однако метод легко распространяется и на задачи оптимизации дискретных систем управления. При этом уравнения движения $x(t+1) = f(x(t), u(t), t)$ легко интерпретируются как ограничения типа равенства, а ограничения $u(t) \in U(t)$ можно учесть введением штрафной функции $v(u)$ такой, что $v(u) = 0, u \in U(t)$

и $v(u) > 0$ при $u \notin U(t)$. Распространение метода штрафных функций на непрерывные процессы не вызывает принципиальных трудностей.

3. Понижение размерности задачи с помощью множителей Лагранжа. Хотя метод штрафных функций сводит исходную задачу на условный минимум к новой задаче на безусловный и в этом смысле аналогичен методу множителей Лагранжа в выпуклом программировании, однако эффективность его с точки зрения динамического программирования, с точки зрения понижения размерности исходной задачи может значительно уступать методу множителей Лагранжа. Введение нелинейных штрафов, как правило, затрудняет получение уравнений Беллмана. Если с помощью некоторых операций все же удастся получить уравнение для функции Беллмана, размерность вспомогательной задачи может снова повыситься. Этого недостатка лишен метод множителей Лагранжа.

Рассмотрим задачу:

$$\sum_{t=0}^{t_1-1} f_0(x(t), t) \rightarrow \min, \sum_{t=0}^{t_1-1} g(x(t), t) \leq 0, \\ \sum_{t=0}^{t_1-1} h(x(t), t) \leq 0, \quad (6)$$

где f_0 — скалярная функция; $g(x, t)$ и $h(x, t)$ — m - и k -векторные функции, x — n -вектор. Если известно (например, применима теорема Куна—Таккера), что при некотором $\lambda > 0$ решение $x(\lambda)$ задачи

$$\sum_{t=0}^{t_1-1} [f_0(x(t), t) + \lambda' g(x(t), t)] \rightarrow \min, \quad \sum_{t=0}^{t_1-1} h(x(t), t) \leq 0, \quad (7)$$

удовлетворяет ограничениям задачи (6), то $x(\lambda)$ есть решение задачи (6). Поэтому вместо решения задачи (6) можно перейти к решению задачи (7). С точки зрения приближенных вычислений к задаче (7) можно перейти и тогда, когда при $\lambda \rightarrow \bar{\lambda} > 0$ решение $x(\lambda)$ асимптотически удовлетворяет ограничениям задачи (6).

Размерность задачи (6) равна $m+k+1$ (см. § 2 гл. I). Уравнение Беллмана для задачи (7) имеет вид (см. § 2 гл. I):

$$B(t, c) = \min_y [f_0(y, t) + \lambda' g(y, t) + B(t, c - h(y, t))],$$

$$B(t_1 - 1, c) = c, \quad (8)$$

где c — k -вектор.

Таким образом, введение m -мерного множителя Лагранжа λ понижает размерность задачи (6) с $m+k+1$ до $k+1$. Процедура решения такова. Выбирается некоторый вектор $\lambda \geq 0$, каким-нибудь методом решается уравнение (8). Проверяется, удовлетворяет ли $x(\lambda)$ ограничениям задачи (6). Если ограничения удовлетворяются с требуемой точностью, процесс заканчивается.

В противном случае выбирается новое $\bar{\lambda} > \lambda$, причем его значение зависит от предыдущего решения. По предположению найдется конечное $\bar{\lambda}$, при котором решение задачи (7) является и решением задачи (6). Организуя последовательные приближения при выборе $\bar{\lambda}$, находим решение исходной задачи (6).

§ 3. Аппроксимация функций Беллмана

В некоторых задачах структура функции Беллмана не очень сложна и допускает хорошую аппроксимацию заданными классами функций. Для определенности будем говорить о степенных полиномах. Естественна мысль вместо табулирования функций Беллмана запоминать на каждом шаге коэффициенты разложения по полиномам.

Это может существенно сэкономить необходимую оперативную память ЭЦВМ при решении задач оптимизации.

1. Аппроксимация полиномами. Пусть исследуется задача

$$\begin{aligned} x(t+1) &= f(x(t), u(t), t), \quad x(t_0) = x_0; \quad t \in [t_0, t_1 - 1], \\ u(t) &\in U(t), \quad I(u) = \varphi(x(t_1)) \rightarrow \min. \end{aligned} \quad (1)$$

Уравнение Беллмана имеет вид (см. § 8 гл. I):

$$B(t, x) = \min_{u \in U(t)} B(t+1, f(x, u, t)), \quad B(t_1, x) = \varphi(x).$$

Размерность задачи совпадает с размерностью вектора $\{x, t\}$, т. е. равна $n+1$. Процедура решения такова. Задается область $H(t_1)$ в n -мерном пространстве с таким расчетом, чтобы правый конец оптимальной траектории $x^0(t_1)$ содержался в ней. Последующие вычисления могут привести к необходимости изменить или расширить эту область. Функция $B(t_1, x) = \varphi(x)$ в области $H(t_1)$ аппроксимируется полиномами

$$\begin{aligned} B(t_1, x) = \varphi(x) \approx \sum_{k=1}^m \sum_{i_1, \dots, i_k=1}^n c_{i_1 \dots i_k}(t_1) x_{i_1} \dots x_{i_k} + \\ + c_0(t_1). \end{aligned} \quad (2)$$

Значения коэффициентов

$$c_0(t_1), c_{i_1 \dots i_k}(t_1), \quad i_1, \dots, i_k = \overline{1, n}, \quad (3)$$

запоминаются.

Переходим к вычислению функции $B(t_1-1, x)$. Для этого с помощью разложения (2) вычисляем значения $B(t_1, f(x, u^1, t_1-1)), \dots, B(t_1, f(x, u^{M(t_1-1)}, t_1-1))$. (4)

Среди чисел (4) находим минимальное:

$$\begin{aligned} B(t_1-1, x) &= \min B(t_1, f(x, u^i, t_1-1)), \\ i &= 1, \dots, M(t_1-1). \end{aligned}$$

Задавая опять множество $H(t_1-1)$ и вычисляя по описанной схеме в точках этого множества значения $B(t_1-1, x)$, определяем коэффициенты разложения:

$$B(t_1 - 1, x) \approx \sum_{k=1}^m \sum_{i_1, \dots, i_k=1}^n c_{i_1 \dots i_k} (t_1 - 1) x_{i_1} \dots x_{i_k} + c_0(t_1 - 1),$$

которые запоминаем наряду с коэффициентами (3).

Продолжая этот процесс, получаем последовательность коэффициентов

$$c_0(t), c_{i_1 \dots i_k}(t), t = t_1 - 1, \dots, t_0. \quad (5)$$

С помощью запомненных коэффициентов (5) нетрудно построить решение исходной задачи. Вычислим через коэффициенты (5) значение функции $B(t_0, x)$ в точке $x = x_0$. Оно, как известно, равно минимальному значению критерия качества. Далее находим вектор $u^l = u(x_0, t_0)$, на котором достигается минимум в совокупности

$$B(t_0 + 1, f(x_0, u^1, t_0)), \dots, B(t_0 + 1, f(x_0, u^{M(t_0+1)}, t_0)),$$

при вычислении которой используются коэффициенты (5) с $t = t_0 + 1$. Подставляя вектор $u^0(t_0)$ в (1), находим вектор $x^0(t_0 + 1)$. Продолжая этот процесс, через конечное число шагов с помощью коэффициентов (5) найдем оптимальное управление $u^0(t)$ и оптимальную траекторию $x^0(t)$, $t \in [t_0, t_1]$. Если полиномы хорошо аппроксимируют функцию $B(t, x)$, то объем памяти на хранение коэффициентов (5) может оказаться значительно меньшим, чем в стандартной процедуре.

2. Аппроксимация законов управления. В некоторых задачах (наряду с функцией Беллмана) и оптимальные управления $u^0(x, t)$ допускают хорошую аппроксимацию в специальных классах функций. Тогда (если для определенности ограничиться полиномами) можно поступить следующим образом.

Аппроксимируем функцию $B(t_1, x) = \varphi(x)$:

$$B(t_1, x) \approx \sum_{k=1}^m \sum_{i_1, \dots, i_k=1}^n c_{i_1 \dots i_k}(t_1) x_{i_1} \dots x_{i_k} + c_0(t_1).$$

Вычисляем значения (4) и отмечаем векторы u^l , на которых достигается минимум. Получаем $B(t_1 - 1, x)$ и $u(x, t_1 - 1) = u^l$.

Повторяя эти вычисления для нескольких точек, находим коэффициенты разложений

$$B(t_1 - 1, x) = \sum_{k=1}^m \sum_{i_1, \dots, i_k=1}^n c_{i_1 \dots i_k}(t_1 - 1) x_{i_1} \dots x_{i_k} + c_0(t_1 - 1),$$

$$u^0(x, t_1 - 1) = \sum_{m=1}^s \sum_{i_1, \dots, i_m=1}^n d_{i_1 \dots i_m}(t_1 - 1) x_{i_1} \dots x_{i_m} + d_0(t_1 - 1).$$

Коэффициенты $d_{i_1 \dots i_m}(t_1 - 1)$, $d_0(t_1 - 1)$, $c_{i_1 \dots i_k}(t_1 - 1)$, $c_0(t_1 - 1)$ запоминаются, коэффициенты $c_{i_1 \dots i_k}(t_1)$, $c_0(t_1)$ можно убрать из оперативной памяти.

Продолжая процесс вычислений, получаем коэффициенты

$$d_0(t), d_{i_1 \dots i_m}(t), t \in [t_0, t_1 - 1], \quad (6)$$

которые полностью задают оптимальный закон управления

$$u^0(x, t) = \sum_{m=1}^s \sum_{i_1, \dots, i_m=1}^n d_{i_1 \dots i_m}(t) x_{i_1} \dots x_{i_m} + d_0(t).$$

Таким образом, если есть возможность аппроксимации законов управления, то в оперативной памяти в каждый момент можно хранить коэффициенты разложения функции Беллмана лишь для одного момента времени. Коэффициенты (6) разложения закона управления не участвуют в оперативной работе и поэтому могут храниться в другом месте.

§ 4. Метод переменного шага

Анализ стандартной процедуры показывает, что расход оперативной памяти можно существенно понизить, если удачно ввести сетку и «предугадать» возможное течение процесса. Одним из эффективных приемов эконо-

номии оперативной памяти является разумный выбор [24] на каждом этапе вычислений величины шага по аргументу t . Понятно, что такой путь возможен только для системы вида (4) § 1.

1. Использование предварительных оценок. Объем оперативной памяти, необходимый в стандартной процедуре, зависит от числа узлов фазовой сетки. Уменьшению числа узлов препятствует расширение сетки, из-за чего стандартная процедура может остановиться. Но, с

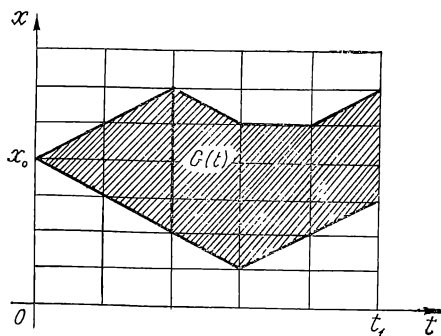


Рис. 45

другой стороны, ясно, что при решении конкретной задачи с фиксированными начальными условиями многие узлы фазовой сетки оказываются лишними. Действительно, рассмотрим картину, изображенную на рис. 45. На рисунке $G(t)$ — множество достижимости системы,

$$G(t) = \{x : x = x(t), x(0) = x_0, u(t) \in U(t)\}.$$

Для вычисления оптимального управления из точки x_0 достаточно узлов фазовой сетки, окаймляющих множества $G(t)$, $t \in [1, t_f]$. Построение множеств достижимости — задача трудоемкая. Однако при изучении конкретных задач предварительные исследования могут дать оценки для множеств $G(t)$. Поэтому, строя фазовую сетку по полученным оценкам, можно существенно сэкономить объем оперативной памяти при решении задачи оптимизации.

2. Описание метода. Опасность расширяющейся сетки связана с тем, что множество достижимости $G(t)$ с ростом t может увеличиваться. Для преодоления этой

трудности можно использовать метод переменного шага, который применим для дискретных систем

$$x(t + \Delta t) = x(t) + f(x(t), u(t), t) \Delta t, \quad x(t_0) = x_0, \quad (1)$$

полученных квантованием по времени из непрерывных.

Если в стандартной процедуре вычисления для (1) идут с постоянным шагом Δt , то в методе переменного шага очередной шаг выбирается в зависимости от текущего состояния.

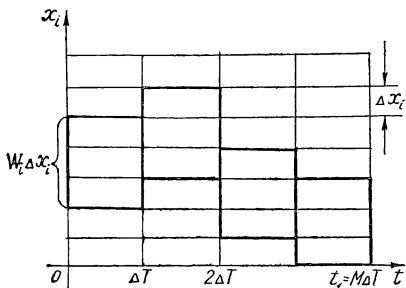


Рис. 46

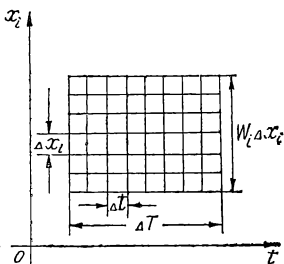


Рис. 47

Опишем вычислительную схему метода. В методе принципиальными являются два момента: 1) построение блоков; 2) использование переменного шага.

Пусть $\beta_i^- \leq x_i \leq \beta_i^+$, $i = \overline{1, n}$, — область возможных изменений координат состояния. Переменные состояния квантуем по уровню $x_i = \beta_i^- + j_i \Delta x_i$, $j_i = 0, 1, \dots, N_i$, $N_i \Delta x_i = \beta_i^+ - \beta_i^-$, $i = \overline{1, n}$.

Пусть множество U содержит конечное число элементов

$$U = \{u^1, \dots, u^M\}$$

или получено из заданного множества U квантованием по уровню. Ось времени квантуем на части $\Delta T: t_1 = M \Delta T$. В пространстве $X \times T$ вводим блоки (рис. 46). Каждый блок состоит из W_i приращений Δx_i вдоль оси Ox_i и ΔT секунд вдоль Ot . Обычно $W_i = 0 \div 5$, $\Delta T \gg \Delta t_{\min}$. В методе переменного шага различают два случая: 1) вычисления внутри блока; 2) переход между блоками.

1) Вычисления в блоке. Пусть текущее состояние $\{x, t\}$ находится внутри блока. Число ΔT делится

на s частей ($s=5 \div 15$): $s \Delta t = \Delta T$ (рис. 47). Оптимальное управление вычисляется в каждом узле блока. Сначала для каждого $u^k \in U$ находится число

$$\Delta t_{\min}^k = \min_{i=\overline{1, n}} \Delta t_{ik}, \quad \Delta t_{ik} = \left| \frac{\Delta x_i}{f_i(x, u^k, t)} \right|.$$

Далее вычисляется вектор

$$x^k = x + f(x, u^k, t) \Delta t_{\min}^k. \quad (2)$$

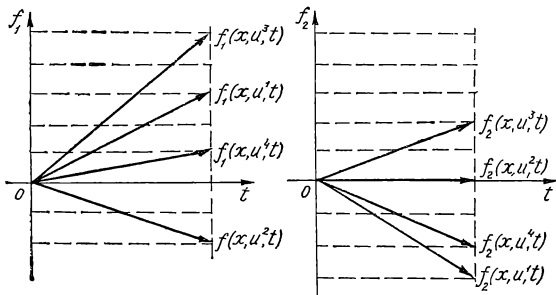


Рис. 48

Значение $B(t + \Delta t_{\min}^k, x^k)$ определяется по вычисленным ранее значениям с помощью интерполирования и экстраполирования.

Функция Беллмана в точке $\{x, t\}$ вычисляется по формуле

$$B(t, x) = \min_k B(t + \Delta t_{\min}^k, x^k).$$

Если \bar{k} — индекс, на котором достигается минимум, то оптимальное управление в точке $\{x, t\}$ равно $u^0(x, t) = u^{\bar{k}}$. Поясним описанную процедуру для $n=2, k=4$.

Пусть система (1) такова, что координаты функции $f(x, u, t)$ имеют вид, изображенный на рис. 48. Числа Δt_{ik} находятся так, как это показано на рис. 49. Видно, что $\Delta t_{\min}^1 = \Delta t_{21}$, $\Delta t_{\min}^2 = \Delta t_{12}$, $\Delta t_{\min}^3 = \Delta t_{13}$, $\Delta t_{\min}^4 = \Delta t_{24}$. Точки $x^k = \{x_1^k, x_2^k\}$ изображены на рис. 50.

Из изложенного видно, что в отличие от стандартной процедуры, где x^k вычисляется по формуле

$$x^k = x + f(x, u^k, t) \Delta t,$$

в методе переменного шага применяется формула (2), в которой шаг Δt^k зависит от u^k .

Метод переменного шага обладает рядом преимуществ по сравнению со стандартной процедурой. Преж-

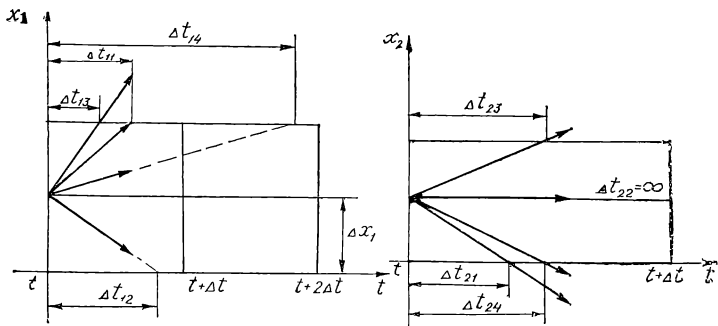


Рис. 49

де всего он гарантирует, что очередная точка x^k лежит в заданной малой окрестности исходной точки x (рис. 51), а это увеличивает точность экстраполяции и интер-

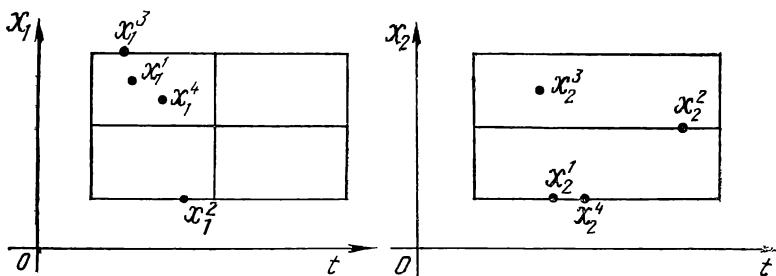


Рис. 50

поляции. Далее, метод в некоторой степени уменьшает опасность расширяющейся сетки, хотя для этого приходится прибегать к дополнительным операциям интерполирования и экстраполирования. Удачное выделение блоков существенно экономит память и время ЭЦВМ, но требует хорошего предварительного исследования задачи с целью определения примерного течения оптимального процесса. В этом отношении он близок методу искусственных ограничений (§ 10), где около допустимых

траекторий строятся области ограничений. Если оптимальная траектория проходит по построенной области, то, очевидно, можно ограничиться блоками, покрывающими область.

2) Переход между блоками. Рассмотрим на плоском случае два вида

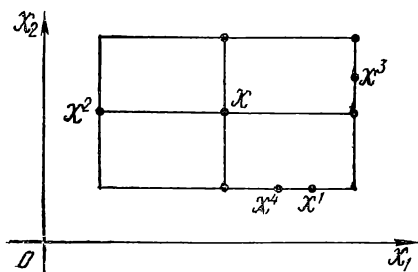


Рис. 51

перехода между блоками, изображенными на рис. 52, 53. Будем считать, что вычисления во внутренних точках блока 1 уже проведены. Нужно вычислить функцию Беллмана на границе между блоками и в узлах блока 2. Для узла $\{x, t\}$ (рис. 52) ситуация может оказаться такой, как на рис. 54 (крестиками обозначены узлы, в которых функция Беллмана

неизвестна, но ее значение при вычислении $B(t, x)$ получено экстраполяцией значений в узлах, помеченных кружком, где функция Беллмана уже известна; стрел-

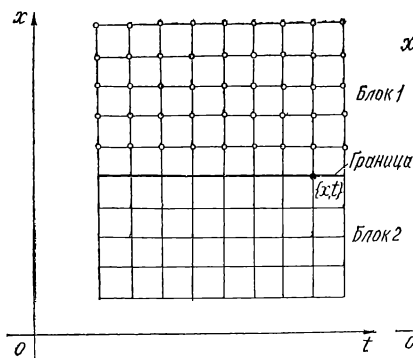


Рис. 52

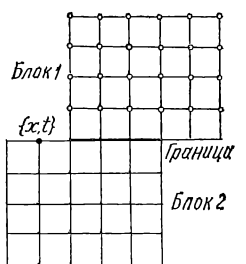


Рис. 53

ками обозначены векторы $f(x, u^k, t)$). При удачном делении блоков количество экстраполяций для точек типа α, β (см. рис. 54) будет незначительным. Для остальных узлов границы и для точек внутри блока 2 ситуация будет, как на рис. 55.

Рассмотрим теперь переход между блоками, изображенными на рис. 53. Новым моментом здесь является использование экстраполяции не только по переменным

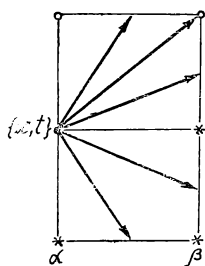


Рис. 54

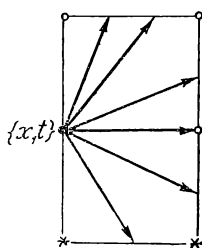


Рис. 55

состояния, но и вдоль оси времени. Например, значение функции Беллмана в узле $\{x, t\}$ получается экстраполяцией значений функции Беллмана в узлах, лежащих справа от узла $\{x, t\}$.

§ 5. Использование специальной структуры задачи

Иногда удачное использование особенностей задачи позволяет резко уменьшить размерность задачи с точки зрения динамического программирования.

1. Специальная задача терминального управления. Рассмотрим задачу

$$x(t+1) = A(t)x + b(t, u), \quad x(t_0) = x_0, \quad t \in [t_0, t_1 - 1], \quad (1)$$

$$u(t) \in U(t), \quad I(u) = \varphi([x(t_1)]^m) \rightarrow \min, \quad (2)$$

где x — n -вектор; u — r -вектор; $[x]^m$ — m -вектор ($m < n$), составленный из m определенных компонент вектора x ; $A(t)$, $b(t, u)$ — произвольные матрицы соответствующих размерностей.

Уравнение Беллмана, записанное для задачи (1), (2) стандартным методом, имеет размерность, равную n (см. § 8 гл. I). Покажем, что размерность задачи (1), (2) можно понизить до m .

По формуле Коши решение линейного по x уравнения (1) можно представить в виде

$$x(t) = F(t, t_0)x_0 + \sum_{\tau=t_0}^{t-1} F(t, \tau+1)b(u(\tau), \tau),$$

где матричная функция $F(t, \tau)$ удовлетворяет уравнению

$$F(t+1, \tau) = A(t)F(t, \tau), \quad t \geq \tau,$$

с граничным условием $F(\tau, \tau) = E$. Поэтому

$$[x(t_1)]^m = [F(t_1, t_0)x_0]^m + \sum_{\tau=t_0}^{t_1-1} [F(t_1, \tau+1)b(u(\tau), \tau)]^m.$$

Задача (1), (2) эквивалентна задаче

$$\varphi([F(t_1, t_0)x_0]^m + \sum_{s=t_0}^{t_1-1} [F(t_1, s+1)b(u(s), s)]^m) \rightarrow \min_{u[t_0, t_1-1]} \quad (3)$$

Погрузим задачу (3) в семейство

$$\varphi(y + \sum_{s=\tau}^{t_1-1} [F(t_1, s+1)b(u(s), s)]^m) \rightarrow \min_{u[\tau, t_1-1]}$$

где y — m -вектор, и введем функцию Беллмана

$$B(\tau, y) = \min_{u[\tau, t_1-1]} \varphi(y + \sum_{s=\tau}^{t_1-1} [F(t_1, s+1)b(u(s), s)]^m). \quad (4)$$

Имеем

$$B(\tau, y) = \min_{u(\tau) \in U(\tau)} \min_{u[\tau+1, t_1-1]} \varphi(y + [F(t_1, \tau+1)b(u(\tau), \tau)]^m + \sum_{s=\tau+1}^{t_1-1} [F(t_1, s+1)b(u(s), s)]^m).$$

Поэтому функция (4) удовлетворяет уравнениям

$$B(t, y) = \min_{u \in U(t)} B(t+1, y + [F(t_1, t+1)b(u, t)]^m),$$

$$B(t_1-1, y) = \min_{u \in U(t_1-1)} \varphi(y + [F(t_1, t_1)b(u, t_1-1)]^m).$$

Вычислив отсюда функцию $B(t_0, y)$ и положив $y = [F(t_1, t_0)x_0]^m$, получим минимальное значение критерия качества. Оптимальный закон управления $u^0(x, t)$ удовлетворяет условию

$$u^0(x, t) = u^*([F(t_1, t)x]^m, t),$$

где $u^*(y, t)$ находится из соотношения

$$B(t+1, y + [F(t_1, t+1) b(u^*(y, t), t)]^m) = \\ = \min_{u \in U(t)} B(t+1, y + [F(t_1, t+1) b(u, t)]^m).$$

2. Непрерывные процессы. Описанный прием можно использовать и при оптимизации непрерывных систем. Пусть, например, рассматривается задача

$$\dot{x} = A(t)x + b(u, t), \quad x(t_0) = x_0, \quad t \in [t_0, t_1], \quad (5)$$

$$u(t) \in U, \quad I(u) = \varphi([x(t_1)]^m) \rightarrow \min. \quad (6)$$

Запишем решение системы (5) с помощью формулы Коши

$$x(t) = F(t)x_0 + \int_{t_0}^t F(t)F^{-1}(\tau)b(u, \tau)d\tau,$$

где $n \times n$ -матричная функция $F(t)$ удовлетворяет уравнению

$$\frac{dF(t)}{dt} = A(t)F(t); \quad F(t_0) = E.$$

Задача (5), (6) эквивалентна задаче

$$\varphi([F(t_1)x_0]^m) + \int_{t_0}^{t_1} [F(t_1)F^{-1}(\tau)b(u, \tau)]^m d\tau \rightarrow \min. \quad (7)$$

Погружая задачу (7) в семейство

$$\varphi(y + \int_{\tau}^{t_1} [F(t_1)F^{-1}(s)b(u, s)]^m ds) \rightarrow \min, \quad y - m\text{-вектор},$$

для функции Беллмана

$$B(\tau, y) = \min_{u \in [U, t_1-1]} \varphi(y + \int_{\tau}^{t_1} [F(t_1)F^{-1}(s)b(u, s)]^m ds)$$

получаем соотношение

$$B(\tau, y) = \min_{u \in [\tau, \tau+\Delta\tau]} \min_{u \in [\tau+\Delta\tau, t_1]} \varphi(y + \int_{\tau}^{\tau+\Delta\tau} [F(t_1)F^{-1}(s)b(u, s)]^m ds)$$

$$s)]^m ds + \int_{\tau+\Delta\tau}^{t_1} [F(t_1) F^{-1}(s) b(u, s)]^m ds = \\ = \min_{u[\tau, \tau+\Delta\tau]} B(\tau + \Delta\tau, y + \int_{\tau}^{\tau+\Delta\tau} [F(t_1) F^{-1}(s) b(u, s)]^m ds),$$

Предполагая функцию $B(\tau, y)$ гладкой по аргументам, выводим уравнение

$$-\frac{\partial B(t, y)}{\partial t} = \min_{u \in U} \frac{\partial B'(t, y)}{\partial y} [F(t_1) F^{-1}(t) b(u, t)]^m,$$

$$B(t_1, y) = \varphi(y).$$

3. Система с запаздыванием. Прием, описанный в пп. 1, 2, особенно эффективен при оптимизации систем с запаздыванием. Предварительно получим уравнение Беллмана стандартным путем (см. § 4).

Рассмотрим задачу

$$\left. \begin{aligned} \dot{x}(t) &= f(x(t), x(t-h), u(t), t), \quad t \in [0, t_1], \\ x_0(\cdot) &= \{x(t) = \Phi(t), \quad -h \leq t \leq 0, x(0) = x_0\}, \end{aligned} \right\} \quad (8)$$

$$u(t) \in U, \quad I(u) = \varphi(x(t_1)) \rightarrow \min. \quad (9)$$

Через $x_t(\cdot)$ будем обозначать отрезок траектории:

$$x_t(\cdot) = \{x(t + \vartheta), \quad -h \leq \vartheta < 0\},$$

через $y(\cdot)$ — гладкую функцию $y(t)$, заданную на множестве $[-h, 0]$.

Погрузим процесс (8) в семейство

$$\dot{x}(t) = f(x(t), x(t-h), u(t), t), \quad t \in [\tau, t_1],$$

$$x_\tau(\cdot) = y(\cdot), \quad x(\tau) = y,$$

и введем функцию Беллмана

$$B(\tau, y, y(\cdot)) = \min_{u[\tau, t_1]} \varphi(x(t_1)), \quad (10)$$

$$x_\tau(\cdot) = y(\cdot), \quad x(\tau) = y.$$

Нетрудно видеть, что справедливо соотношение

$$\begin{aligned} B(\tau, y, y(\cdot)) &= \min_{u[\tau, \tau+\Delta\tau]} \min_{\substack{u[\tau+\Delta\tau, t_1] \\ x_{\tau+\Delta\tau}(\cdot)=x(\tau+\Delta\tau)}} \varphi(x(t_1)) = \\ &= \min_{\substack{u(s) \in U \\ s \in [\tau, \tau+\Delta\tau]}} B(\tau + \Delta\tau, x(\tau + \Delta\tau), x_{\tau+\Delta\tau}(\cdot)). \end{aligned} \quad (11)$$

Предположим, что функция $B(\tau, y, y(\cdot))$ имеет непрерывные первые вариации по аргументам. Тогда

$$\begin{aligned} B(\tau + \Delta\tau, x(\tau + \Delta\tau), x_{\tau+\Delta\tau}(\cdot)) &= B(\tau, y, y(\cdot)) + \\ &+ \int_{-h}^0 \frac{\delta B'(\tau, y, y(\cdot))}{\delta y(\vartheta)} [x(\tau + \Delta\tau + \vartheta) - y(\vartheta)] d\vartheta + \\ &+ \frac{\partial B'(\tau, y, y(\cdot))}{\partial y} [x(\tau + \Delta\tau) - y] + \frac{\partial B(\tau, y, y(\cdot))}{\partial \tau} \Delta\tau + \\ &+ o(\|x_{\tau+\Delta\tau}(\cdot) - y(\cdot)\|, \|x(\tau + \Delta\tau) - y\|, |\Delta\tau|). \end{aligned}$$

Из уравнения (8) получаем:

$$\begin{aligned} x(\tau + \Delta\tau) &= y + f(y, y(-h), u(\tau), \tau) \Delta\tau + o(|\Delta\tau|), \\ x(\tau + \Delta\tau + \vartheta) &= y(\vartheta + \Delta\tau), \quad -h \leq \vartheta < -\Delta\tau, \\ x(\tau + \Delta\tau + \vartheta) &= y + f(y, y(-h), u(\tau), \tau) (\vartheta + \Delta\tau) + \\ &+ o(|\vartheta + \Delta\tau|), \quad -\Delta\tau \leq \vartheta < 0. \end{aligned}$$

Поэтому

$$\begin{aligned} &\int_{-h}^0 \frac{\delta B'(\tau, y, y(\cdot))}{\delta y(\vartheta)} [x(\tau + \Delta\tau + \vartheta) - y(\vartheta)] d\vartheta = \\ &= \int_{-h}^{-\Delta\tau} \frac{\delta B'(\tau, y, y(\cdot))}{\delta y(\vartheta)} \dot{y}(\vartheta) d\vartheta \cdot \Delta\tau + \int_{-\Delta\tau}^0 \frac{\delta B'(\tau, y, y(\cdot))}{\delta y(\vartheta)} f(y, \\ &y(-h), u(\tau), \tau) \cdot (\vartheta + \Delta\tau) d\vartheta + o(|\Delta\tau|) = \\ &= \int_{-h}^0 \frac{\delta B'(\tau, y, y(\cdot))}{\delta y(\vartheta)} \dot{y}(\vartheta) d\vartheta \cdot \Delta\tau + o(|\Delta\tau|). \end{aligned}$$

Из (11) после деления на $\Delta\tau$ и предельного перехода ($\Delta\tau \rightarrow 0$) получим уравнение Беллмана

$$-\frac{\partial B(t, x, x(\cdot))}{\partial t} = \min_{u \in U} \frac{\partial B'(t, x, x(\cdot))}{\partial x} \cdot f(x, x(-h), u, t) + \\ + \int_{-h}^0 \frac{\delta B'(t, x, x(\cdot))}{\delta x(\vartheta)} \dot{x}(\vartheta) d\vartheta. \quad (12)$$

Граничное условие получается из (10):

$$B(t_1, x, x(\cdot)) = \varphi(x). \quad (13)$$

Если решить уравнение (12) при условии (13), то минимальное значение критерия качества для задачи (8) — (9)

$$I(u^0) = B(0, x_0, x_0(\cdot)).$$

Оптимальное управление u^0 получается как функция величин $x(\cdot)$, x , t из соотношения

$$\frac{\partial B'(t, x, x(\cdot))}{\partial x} f(x, x(-h), u^0(x, x(\cdot), t), t) = \\ = \min_{u \in U} \frac{\partial B'(t, x, x(\cdot))}{\partial x} f(x, x(-h), u, t).$$

Использование уравнения (12) для численного решения задачи крайне затруднительно, так как одним из аргументов функции Беллмана является отрезок функции $x(\cdot)$ и табулирование в этом случае практически неосуществимо.

Если система (8) линейна,

$$\dot{x}(t) = A(t)x(t) + C(t)x(t-h) + b(u(t), t),$$

и критерий качества квадратичен,

$$I(u) = x'(t_1) S x(t_1) + \int_0^{t_1} [x'(t) M(t) x(t) + \\ + u'(t) R(t) u(t)] dt, \\ S \geq 0, M \geq 0, R > 0,$$

задача упрощается и из (12) получаются (см. § 7 гл. II) уравнения для параметров функционала

$$B(t, x, x(\cdot)) = x' L(t) x + x' \int_{-h}^0 L(t, \vartheta) x(\vartheta) d\vartheta +$$

$$+ \int_{-h}^0 x'(\vartheta) L_1(t, \vartheta) x(\vartheta) d\vartheta + \int_{-h}^0 \int_{-h}^0 x'(\vartheta_1) L(t, \vartheta_1, \vartheta_2) x(\vartheta_2) d\vartheta_1 d\vartheta_2.$$

Другой путь упрощения связан с приемом, использованным в пп. 1, 2. Можно проверить, что функция

$$x(t) = F(t, t_0) x_0 + \int_{t_0-h}^{t_0} F(t, \tau + h) C(\tau + h) \Phi(\tau) d\tau + \\ + \int_{t_0}^t F(t, \tau) b(u(\tau), \tau) d\tau$$

является решением уравнения

$$\dot{x}(t) = A(t)x(t) + C(t)x(t-h) + b(u(t), t). \quad (14)$$

Если ввести обозначение

$$y = [F(t_1, t_0) x_0 + \int_{t_0-h}^{t_0} F(t_1, \tau + h) C(\tau + h) \Phi(\tau) d\tau]^m,$$

то задача минимизации функционала

$$I(u) = \varphi([x(t_1)]^m) \quad (15)$$

сведется к задаче минимизации выражения

$$\varphi(y + \int_{t_0}^{t_1} [F(t_1, \tau) b(u(\tau), \tau)]^m d\tau,$$

которая уже рассмотрена в п. 1. Таким образом, для случая (14), (15) размерность задачи с точки зрения динамического программирования понижается от бесконечности до m .

4. Дискретные процессы с запаздыванием. Задача

$$x(t+1) = A(t)x(t) + C(t)x(t-h) + b(u(t), t), \\ t \in [t_0, t_1 - 1],$$

$$x(t) = \Phi(t), t \in [t_0 - h, t_0], u(t) \in U(t),$$

где x — n -вектор, $I(u) = \varphi(x(t_1)) \rightarrow \min$, с точки зрения динамического программирования имеет размерность $n(h+1)$ (см. § 6 гл. I). Применение схемы п. 1 позволяет снизить размерность этой задачи до n .

§ 6. Последовательные приближения в пространстве функций Беллмана

В некоторых задачах, например в стационарных задачах (§ 4 гл. II), уравнения Беллмана не имеют рекуррентного вида и поэтому стандартная процедура, описанная в § 1, к ним неприменима. Одним из методов решения уравнений вида $T(x) = 0$ является метод последовательных приближений.

1. Задача быстрогодействия в стационарных системах. Рассмотрим задачу

$$\left. \begin{aligned} x(t+1) &= f(x(t), u(t)), \quad x(0) = x_0, \quad t \geq 0, \\ u(t) &\in U, \quad x(t_1) = 0, \quad t_1 \rightarrow \min. \end{aligned} \right\} \quad (1)$$

Будем для простоты рассуждений считать, что для каждой точки x_0 существует допустимое управление, переводящее эту точку вдоль $x(t)$ в начало координат (система (1) управляема в целом). Уравнение Беллмана для точек x , достаточно удаленных от начала координат, имеет вид (см. § 4 гл. II)

$$B(x) = 1 + \min_{u \in U} B(f(x, u)), \quad B(0) = 0. \quad (2)$$

Метод последовательных приближений в пространстве функций Беллмана состоит в следующем: задается нулевое приближение $B_0(x)$, а последующие приближения $B_k(x)$, $k=1, 2, \dots$, вычисляются по формулам

$$B_k(x) = 1 + \min_{u \in U} B_{k-1}(f(x, u)), \quad k = 1, 2, \dots$$

Если нулевое приближение угадано достаточно точно (получено из достаточно тщательного предварительного исследования, опыта), можно ожидать, что $B_k(x) \rightarrow B(x)$ при $k \rightarrow \infty$, где $B(x)$ удовлетворяет уравнению (2). Точные условия сходимости метода пока неизвестны.

2. Общий случай. Описанная схема может быть применена к задачам более общим, чем задача, рассмотренная в п. 1.

Пусть уравнение Беллмана некоторой задачи оптимизации таково:

$$B(x) = \min_{u \in U} [h(x, u) + B(f(x, u))], \quad B(y) = a. \quad (3)$$

Аналогично предыдущему строим последовательные приближения $B_k(x)$ по формулам

$$B_k(x) = \min_{u \in U} [h(x, u) + B_{k-1}(f(x, u))], \quad k = 1, 2, \dots, \quad (4)$$

где $B_0(x)$ — начальное приближение.

Пусть $h(x, u) \geq 0$ и для каждого x найдется вектор \bar{u} такой, что $f(x, \bar{u}) = x$. Тогда из (4) получаются неравенства

$$B_k(x) \leq h(x, \bar{u}) + B_{k-1}(f(x, \bar{u})) \leq B_{k-1}(x),$$

т. е. если последовательность функций $B_k(x)$ сходится, то сходится монотонно, причем $B_k(x) - B_{k-1}(x) \leq h(x, \bar{u})$.

3. Использование вспомогательных задач. Для задачи быстрогодействия из п. 1 можно сформулировать другую задачу, у которой уравнение Беллмана рекуррентно, и из решения последней получить решение исходной задачи.

Рассмотрим вспомогательную задачу

$$\left. \begin{aligned} x(t+1) &= f(x(t), u(t)), \quad x(0) = x_0, \quad t \in [0, t_1 - 1], \\ u(t) &\in U, \quad I(u) = \|x(t_1)\| \rightarrow \min. \end{aligned} \right\} \quad (5)$$

Уравнение Беллмана здесь имеет вид

$$B(t, x) = \min_{u \in U} B(t+1, f(x, u)), \quad B(t_1, x) = \|x\|,$$

и может быть решено по стандартной процедуре. Чтобы решить задачу быстрогодействия, достаточно заметить, что время оптимального быстрогодействия $t_1^0(x_0)$ — это наименьшее t_1 , при котором $\min \|x(t_1)\| = 0$. Отсюда следует метод вычисления управления, оптимального по быстродействию. Решая задачу (5) при различных t_1 , находим наименьшее t_1^0 , при котором $B(0, x_0) = 0$. Оптимальное управление для задачи (5) с $t_1 = t_1^0$ будет оптимальным управлением в задаче (1).

Аналогично уравнение (3) можно интерпретировать как обратное уравнение Беллмана в задаче оптимизации типа Лагранжа

$$\left. \begin{aligned} x(t+1) &= f(x(t), u(t)), \quad x(0) = x_0, \quad t \geq 0, \\ u(t) &\in U, \quad I(u) = \sum_{t=0}^{t_1} h(x(t), u(t)) \rightarrow \min, \quad t_1 \geq 0, \end{aligned} \right\} \quad (6)$$

с нефиксированным конечным моментом. Выбирая некоторое число $t_1, t_1 \geq 0$, решаем задачу (6) с фиксированным моментом. Для этой задачи функция Беллмана $B_{t_1}(t, x)$ удовлетворяет рекуррентному уравнению (см. § 9 гл. I). Находим число t_1^0 так, чтобы $B_{t_1^0}(0, x_0) = \min_{t_1 \geq 0} B_{t_1}(0, x_0)$. Тогда функция $B_{t_1^0}(0, x)$ будет решением уравнения (3).

§ 7. Последовательные приближения в пространстве управлений

При анализе стандартной процедуры нетрудно заметить, что наряду с функцией Беллмана постоянно используется закон управления. Это наводит на мысль положить в основу последовательных приближений закон управления. Переход к управлениям разумен и с той точки зрения, что во многих задачах интуиция, опыт работы скорее подскажут достаточно точный вид оптимального управления, чем функции Беллмана.

1. Задача быстродействия. Опять рассмотрим задачу

$$x(t+1) = f(x(t), u(t)), \quad x(0) = x_0, \quad t \geq 0, \\ u(t) \in U, \quad x(t_1) = 0, \quad t_1 \rightarrow \min.$$

Обозначим через G множество $G = \{x : f(x, u) = 0 \text{ при некотором } u \in U\}$.

Для решения уравнения Беллмана

$$B(x) = 1 + \min_{u \in U} B(f(x, u)), \quad x \in G,$$

с граничным условием $B(x) = 1, x \in G$, поступим следующим образом.

Зададим нулевое приближение закона управления $u_0 = u_0(x)$, который переводит точку x_0 на множество G . Время, за которое траектория $x(t)$ системы с законом $u_0(x)$ из точки x попадает на множество G , обозначим через $B_0(x)$. Следующее приближение $u_1(x)$ к закону управления находим из соотношения

$$B_0(f(x, u_1(x))) = \min_{u \in U} B_0(f(x, u)).$$

Далее ищем функцию $B_1(x)$ из уравнения

$$B_1(x) = 1 + B_1(f(x, u_1(x))).$$

Если приближения $u_k(x)$ и $B_k(x)$ построены, закон управления $u_{k+1}(x)$ и функцию $B_{k+1}(x)$ находим из соотношений:

$$B_k(f(x, u_{k+1}(x))) = \min_{u \in U} B_k(f(x, u)),$$

$$B_{k+1}(x) = 1 + B_{k+1}(f(x, u_{k+1}(x))).$$

Если нулевое приближение угадано достаточно точно, можно ожидать, что последовательности функций $u_k(x)$, $B_k(x)$ сходятся к $u^0(x)$, $B(x)$. Точные условия сходимости метода неизвестны.

2. Общий случай. Метод приближений в пространстве управлений для решения уравнения

$$B(x) = \min_{u \in U} [h(x, u) + B(f(x, u))] \quad (1)$$

состоит в следующем. Задается нулевое приближение закона управления

$$u_0 = u_0(x). \quad (2)$$

Определяется функция $B_0(x)$ как решение уравнения

$$B_0(x) = h(x, u_0(x)) + B_0(f(x, u_0(x))). \quad (3)$$

Допустим, что вычисления доведены до $u_k(x)$, $B_k(x)$. Приближения $u_{k+1}(x)$, $B_{k+1}(x)$ находятся из соотношений

$$h(x, u_{k+1}(x)) + B_k(f(x, u_{k+1}(x))) = \min_{u \in U} [h(x, u) + B_k(f(x, u))], \quad (4)$$

$$B_{k+1}(x) = h(x, u_{k+1}(x)) + B_{k+1}(f(x, u_{k+1}(x))). \quad (5)$$

Уравнению (1) можно поставить в соответствие задачу

$$y(\tau + 1) = f(y(\tau), v(\tau)), \quad y(0) = x, \quad \tau \geq 0, \quad (6)$$

$$v(\tau) \in U, \quad I(v, t_1) = \sum_{\tau=0}^{t_1} h(y, v) \rightarrow \min, \quad t_1 \geq 0, \quad (7)$$

функция Беллмана $B(x) = \min_{v, t_1} \sum_{\tau=0}^{t_1} h(y, v)$ которой удовлетворяет уравнению (1).

Пусть нулевое приближение (2) закона управления является допустимым в задаче (6), (7), т. е. $I(u_0) < \infty$. Функция $B_0(x)$, удовлетворяющая уравнению (3), имеет четкий физический смысл: $B_0(x) = I(u_0)$. Отсюда при $h(x, u) \geq 0$ получаем $B_0(x) \geq 0$. Из (4), (5) следует

$$B_k(x) = h(x, u_k(x)) + B_k(f(x, u_k(x))) \geq h(x, u_{k+1}(x)) + B_k(f(x, u_{k+1}(x))).$$

Таким образом, функция $B_k(x)$ удовлетворяет неравенству

$$C(x) \geq h(x, u_{k+1}(x)) + C(f(x, u_{k+1}(x))),$$

а функция $B_{k+1}(x)$ — уравнению

$$C(x) = h(x, u_{k+1}(x)) + C(f(x, u_{k+1}(x))).$$

Следовательно, если выполняется теорема о конечных неравенствах [3], то $B_{k+1}(x) \leq B_k(x)$. По физическому смыслу величина $B_k(x)$ равна значению критерия качества (7) на управлении $u_k(x)$ и поэтому не равна числу $B(x)$ — минимальному значению критерия (7).

Таким образом, выполнение некоторых условий приближения в пространстве управлений привело к монотонно убывающей последовательности функций, ограниченных снизу:

$$B_0(x) \geq B_1(x) \geq \dots \geq B_k(x) \geq \dots \geq B(x).$$

3. Квазилинеаризация. Идея метода последовательных приближений в пространстве управлений может быть использована при решении задач, не связанных с оптимизацией. Классическим примером является уравнение Риккати

$$\dot{x} + x^2 + a(t)x + b(t) = 0, \quad x(0) = c. \quad (8)$$

Это уравнение нелинейно и не может быть решено в квадратурах при произвольных $a(t)$, $b(t)$.

К в а з и л и н е а р и з а ц и я уравнения (8) состоит в замене нелинейного члена x^2 выражением

$$x^2 = \max_u (2ux - u^2) = - \min_u (u^2 - 2ux),$$

в правую часть которого переменная x входит линейно.

Вместо (8) рассмотрим уравнение

$$\dot{x} = \min_u [u^2 - 2ux - a(t)x - b(t)].$$

Отсюда следует, что решение уравнения (8) удовлетворяет дифференциальному неравенству

$$\dot{x} \leq u^2 - 2ux - a(t)x - b(t), \quad x(0) = c. \quad (9)$$

Пусть $y(t, u)$ — решение дифференциального уравнения

$$\dot{y} = u^2 - 2ux - a(t)x - b(t), \quad y(0) = c. \quad (10)$$

Известно, что $y(t)$ можно записать в квадратурах.

Из теории дифференциальных неравенств следует, что решения уравнения (10) и дифференциального неравенства (9) связаны соотношением $x(t) \leq y(t, u)$. При $u(t) = x(t)$ неравенство (9) переходит в уравнение (8), поэтому

$$x(t) = \min_u y(t, u).$$

Метод решения нелинейного уравнения (8) с помощью решений линейного уравнения (10) есть реализация общего метода к в а з и л и н е а р и з а ц и и, восходящего к методу приближений в пространстве управлений.

§ 8. Использование множеств достижимости

При постановке задач терминального управления явно участвуют множества достижимости. Поэтому может оказаться целесообразным непосредственно строить эти множества. Идея заманчива в тех случаях, когда критерий качества зависит явно лишь от небольшого числа координат.

1. Простейшая задача терминального управления.

Пусть имеется задача

$$x(t+1) = f(x(t), u(t), t), \quad x(t_0) = x_0, \quad t \in [t_0, t_1 - 1], \\ u(t) \in U(t), \quad I(u) = \varphi(x(t_1)) \rightarrow \min.$$

Множеством достижимости $F^-(\tau)$ в момент τ называется множество, которое пробегает конец $x(\tau)$ траектории $x(t)$, $t \in [t_0, \tau]$, когда управление $u(t)$, $t \in [t_0, \tau - 1]$, пробегает все множество $U(t)$, $t \in [t_0, \tau - 1]$ (рис. 56).

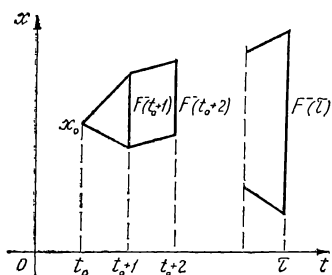


Рис. 56

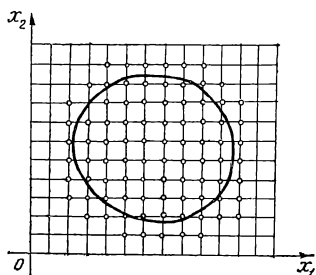


Рис. 57.

Из определения множества достижимости следует правило его построения:

$$F^-(t_0) = x_0, F^-(t+1) = \{y : y = f(x, u, t), x \in F^-(t), u \in U(t)\}, t \in [t_0, t_1 - 1]. \quad (1)$$

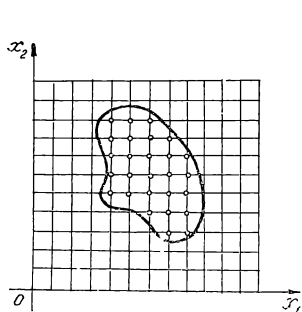


Рис. 58

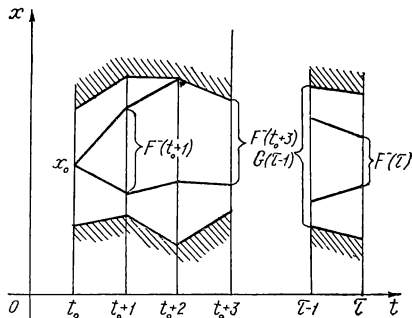


Рис. 59

При реализации на вычислительных устройствах пространство векторов x квантуется с требуемой точностью, а множество $F^-(t)$ в каждый момент аппроксимируется с «избытком» (рис. 57) или с «недостатком» (рис. 58) узлами фазовой сетки.

Объем памяти на каждом этапе t зависит от числа узлов, занятых аппроксимацией множества $F^-(t)$.

В системах с «сильным перемешиванием» (см. §1) он невелик. Проводя вычисления по рекуррентной формуле (1), получаем множество $F^-(t_1)$. Среди точек этого множества нужно найти точку $x^0(t_1)$, на которой достигается минимум функции $\varphi(x)$. Если последняя зависит от небольшого числа переменных (например, во многих задачах это число равно единице), точку $x^0(t_1)$ можно найти перебором по множеству $F^-(t_1)$. В других случаях можно воспользоваться свойствами функции $\varphi(x)$ и применить дополнительные приемы.

Допустим, что точка $x^0(t_1)$ найдена. Остальные точки оптимальной траектории $x^0(t)$ и оптимальное управление находятся по следующему правилу. Вычисляются $x^0(t_1-1)$, $u^0(t_1-1)$ из условий

$$\begin{aligned} f(x^0(t_1-1), u^0(t_1-1), t_1-1) &= x^0(t_1), \\ x^0(t_1-1) &\in F^-(t_1-1). \end{aligned} \quad (2)$$

Пусть значения $x^0(t)$, $u^0(t)$ известны. Векторы $x^0(t-1)$, $u^0(t-1)$ находятся так, чтобы

$$\begin{aligned} f(x^0(t-1), u^0(t-1), t-1) &= x^0(t), \\ x^0(t-1) &\in F^-(t-1). \end{aligned} \quad (3)$$

Задачи (2), (3) всегда разрешимы. Через t_1-t_0 шагов оптимальное управление и траектория будут определены на всем отрезке $[t_0, t_1-1]$. Если при каждом t , $t \in [t_0+1, t_1]$, задача (3) однозначно разрешима относительно $u^0(t-1)$, то оптимальное управление единственно.

З а м е ч а н и е. Описанная процедура никак не учитывает структуру множеств $F^-(t)$. В действительности же в реальных системах структура этих множеств может оказаться весьма специальной и поэтому для описания их не обязательно знать и, следовательно, хранить в памяти все точки из множества. Например, каждое выпуклое множество можно восстановить по совокупности крайних точек.

2. Учет ограничений. Рассмотрим задачу

$$\begin{aligned} x(t+1) &= f(x(t), u(t), t), \quad x(t_0) = x_0, \quad t \in [t_0, t_1-1], \\ u(t) &\in U(t), \quad x(t) \in G(t), \quad I(u) = \varphi(x(t_1)) \rightarrow \min. \end{aligned}$$

Множеством достижимости $F^-(\tau)$ назовем множество, которое пробегает конец $x(\tau)$ траектории $x(t)$, $t \in [t_0, \tau]$, порожденной управлением $u(t) \in U(t)$, $t \in [t_0, \tau-1]$,

и удовлетворяющей ограничению $x(t) \in G(t)$, $t \in [t_0, \tau]$ (рис. 59).

Ясно, что множества $F^-(t)$ строятся по формулам

$$F^-(t_0) = x_0, F^-(t+1) = \{y: y = f(x, u, t), \\ x \in F^-(t), u \in U(t), y \in G(t+1)\}.$$

Наличие ограничений может лишь уменьшить объем памяти, требуемой для аппроксимации множеств $F^-(t)$ узлами фазовой сетки, причем чем «жестче» ограничения, тем лучше результат.

Построив множество $F^-(t_1)$, найдем точку $x^0(t_1)$, на которой достигается минимум функции $\varphi(x)$. Восстановление остальных точек оптимальной траектории и оптимального управления проводится по формулам (2), (3).

§ 9. Искусственное понижение размерности задачи

До сих пор описывались методы решения, которые были направлены на получение абсолютного минимума. Теперь рассмотрим ряд приемов, позволяющих решать задачи высокой размерности. При этом сознательно допускается возможность получения не абсолютного минимума, а только относительного. В свете современных представлений о существовании методов решения задач оптимизации улучшение каким-то образом заданных исходных управлений есть основная цель проблемы оптимизации. Поэтому содержание этого и следующего параграфов может представить серьезный практический интерес.

1. Вспомогательная задача. Предварительно рассмотрим задачу терминального управления, в которой множество значений допустимых управлений зависит от вектора фазового состояния:

$$\left. \begin{aligned} x(t+1) &= f(x(t), u(t), t), \quad x(t_0) = x_0, \quad t \in [t_0, t_1-1], \\ u(t) &\in U(x, t), \quad I(u) = \varphi(x(t_1)) \rightarrow \min. \end{aligned} \right\} \quad (1)$$

Если множества $U(x, t)$ при всех x и $t \in [t_0, t_1-1]$ не пусты, то уравнение Беллмана для задачи (1) имеет следующий вид:

$$B(\tau, y) = \min_{u \in U(y, \tau)} B(\tau+1, f(y, u, \tau)), \quad B(t_1, y) = \varphi(y),$$

в чем легко убедиться, используя рассуждения, приведенные при выводе уравнения Беллмана для простейшей задачи терминального управления.

Для исследования общего случая необходимы некоторые дополнительные конструкции.

Введем в рассмотрение множества $F(t)$, $t \in [t_0, t_1 - 1]$, точек фазового пространства. К множеству $F(t_1 - 1)$ отнесем те точки x , для которых $U(x, t_1 - 1)$ непусто. Для остальных моментов времени множества $F(t)$ строятся рекуррентным образом: $F(\tau) = \{x : f(x, u, \tau) \in F(\tau + 1) \text{ при некотором } u \in U(x, \tau)\}$.

Как и в п. 6 § 8 гл. I, в множествах $F(t)$ выделим пограничный слой $\partial^0 F(t)$ и для точек пограничного слоя введем подмножества $U_1(x, t) \subset U(x, t)$.

С помощью рассуждений, принципиально не отличающихся от приведенных в п. 6 § 8 гл. I, получается уравнение Беллмана для задачи (1) в общем случае:

$$B(\tau, y) = \begin{cases} \min_{u \in U(y, \tau)} B(\tau + 1, f(y, u, \tau)), & y \in \partial^0 F(\tau), y \in F(\tau), \\ \min_{u \in U_1(y, \tau)} B(\tau + 1, f(y, u, \tau)), & y \in \partial^0 F(\tau), \end{cases}$$

$$B(t_1, y) = \varphi(y).$$

2. Понижение размерности простейшей задачи терминального управления. Рассмотрим задачу

$$\left. \begin{aligned} x(t+1) &= f(x(t), u(t), t), \quad x(t_0) = x_0, \quad t \in [t_0, t_1 - 1], \\ u(t) &\in U(t), \quad I(u) = \varphi(x(t_1)) \rightarrow \min, \end{aligned} \right\} \quad (2)$$

где x — n -вектор; u — r -вектор.

Предположим, что $r \geq m$. Обозначим через ω_m некоторый набор из m координат n -вектора. Через $U_{\omega_m}(x, y, t)$, $t \in [t_0, t_1 - 1]$, обозначим множество векторов $u \in U(t)$, удовлетворяющих равенству

$$[x]^m = [f(y, u, t)]^m,$$

где $[x]^m$ — m -вектор, составленный из ω_m .

Метод искусственного понижения размерности задачи состоит в следующем. Задается нулевое приближение процесса $x_0(t)$ и исходная комбинация $\omega_m(0)$ из m координат n -вектора x . Находится первое приближение $u^1(t)$, $x^1(t)$ как решение задачи

$$\left. \begin{aligned} x(t+1) &= f(x(t), u(t), t), \quad x(t_0) = x_0, \\ u(t) &\in U_{\omega_m(0)}(x_0(t+1), x, t), \quad I(u) = \varphi(x(t_1)) \rightarrow \min \end{aligned} \right\} \quad (3)$$

(см. п. 1). Из определения множества $U_{\omega_m(0)}(x, y, t)$ следует, что в задаче (3) m выбранных координат допустимых траекторий $x(t)$ фиксированы и совпадают с соответствующими координатами траектории $x_0(t)$. С точки зрения динамического программирования такая задача имеет размерность $n-m$. Обратное уравнение Беллмана для нее:

$$B(\tau, [y]^{n-m}) = \begin{cases} \min_{u \in U(y, \tau)} B(\tau+1, [f(y, u, \tau)]^{n-m}), \\ \quad y \in \partial^0 F(\tau), \quad y \in F(\tau), \\ \min_{u \in U_1(y, \tau)} B(\tau+1, [f(y, u, \tau)]^{n-m}), \\ \quad y \in \partial^0 F(\tau), \end{cases} \quad (4)$$

$$B(t_1, [y]^{n-m}) = \varphi([y]^{n-m}, [x_0(t_1)]^m).$$

Здесь

$$y = \{[y]^{n-m}, [x_0(\tau)]^m\}, \quad U(y, \tau) = U_{\omega_m(0)}(x_0(\tau+1), y, \tau),$$

Решив уравнение (4), получим функцию $[x(t)]^{n-m}$, которая в совокупности с $[x_0(t)]^m$ образует функцию $x^1(t) = \{[x(t)]^{n-m}, [x_0(t)]^m\}$. Функцию $x^1(t)$ принимаем за первое приближение оптимальной траектории.

Допустим, что построено k -е приближение $x^k(t)$. Выбираем k -ю комбинацию $\omega_m(k)$ из m координат n -вектора x . Траектория $x^{k+1}(t)$ — $(k+1)$ -е приближение оптимальной траектории — ищется как решение задачи

$$x(t+1) = f(x(t), u(t), t), \quad x(t_0) = x_0,$$

$$u(t) \in U_{\omega_m(k)}(x^k(t+1), x, t), \quad I(u) = \varphi(x(t_1)) \rightarrow \min.$$

Эта задача опять имеет с точки зрения динамического программирования размерность, равную $n-m$. Процесс вычисления продолжается до тех пор, пока выигрыш от очередного приближения не станет меньшим наперед заданного числа. Понятно, что процесс вычисления ведет к последовательности монотонно убывающих значений критерия качества.

З а м е ч а н и е. В описанном методе выбор нулевого приближения $x_0(t)$ и последовательных комбинаций $\omega_m(0)$, $\omega_m(1)$, ... находится в распоряжении исследователя. Если нет особых соображений, то при выборе $\omega_m(0)$, $\omega_m(1)$, ... можно пользоваться случайным механизмом.

Метод не гарантирует достижения абсолютного минимума, но в ряде практических задач привел к очень хорошим результатам. Выделение множества U_{ω_m} по уравнению (2) имеет в каждой задаче содержательное толкование.

3. Учет ограничений. Понятно, что метод искусственного понижения размерности задачи можно распространить на другие типы задач оптимизации. Введение дополнительных ограничений не скажется существенно на его эффективности.

4. Искусственное увеличение размерности вектора управления. Для применения метода п. 1 существенно условие $r > m$, которое, вообще говоря, предполагает, что размерность r вектора управления u является достаточно большой. На практике же типичным является случай, когда, напротив, $r \ll n$. Приведем модификацию метода из п. 1, в которой размерность вектора u , как правило, не имеет значения.

Вместо вспомогательной задачи (3) вводят новую:

$$x(t+l) = g(x(t), v(t), t), \quad x(t_0) = x_0, \quad t = t_0,$$

$$t_0 + l, \dots, t_1 - l,$$

$$u(t) \in U(t), \quad I(u) = \varphi(x(t_1)) \rightarrow \min,$$

$$v(t) = \{u(t), u(t+1), \dots, u(t+(l-1))\},$$

где $g(x(t), v(t), t) = \overbrace{f(f(f \dots f(x(t), u(t), t), u(t+1), t+1), u(t+2), t+2), \dots, u(t+(l-1)), t+(l-1))}^l$.

Новая задача удовлетворяет всем ограничениям задачи (2) и может быть решена по схеме п. 2.

5. Нулевое приближение. Построение нулевого приближения $x_0(t)$, как правило, является нестандартной процедурой. В конкретных задачах функция $x_0(t)$ может быть известна перед решением задачи оптимизации как инструкция поведения, выработанная в результате практических действий. В рамках динамического программирования для построения нулевого приближения можно рекомендовать использование стандартной проце-

дуры с грубой сеткой. Постепенное уменьшение размеров сетки может привести к достаточно хорошему нулевому приближению для исходной задачи.

§ 10. Введение искусственных ограничений

Как было замечено в § 1, при использовании стандартной процедуры возможности динамического программирования в направлении необходимого объема

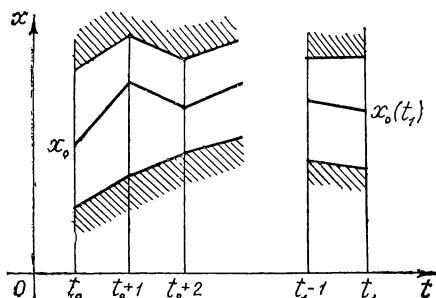


Рис. 60

оперативной памяти растут, если имеются фазовые ограничения. Возникает мысль намеренно вводить фазовые ограничения. При этом, понятно, сужается множество допустимых траекторий и нельзя, следовательно, гарантировать достижения абсолютного минимума. Однако в ряде случаев получение и относительного минимума является очень важным.

1. Простейшая задача терминального управления. Пусть требуется решить задачу

$$\left. \begin{aligned} x(t+1) &= f(x(t), u(t), t), \quad x(t_0) = x_0, \quad t \in [t_0, t_1 - 1], \\ u(t) &\in U(t), \quad I(u) = \varphi(x(t_1)) \rightarrow \min. \end{aligned} \right\} \quad (1)$$

Возьмем допустимое управление $u_0(t) \in U(t)$, вычислим в силу (1) траекторию $x_0(t)$, которую возьмем за нулевое приближение. Построим множества $G_0(t)$ как некоторые окрестности точек $x_0(t)$ (рис. 60). Вместо задачи (1) по методу § 1 решим вспомогательную задачу

$$\left. \begin{aligned} x(t+1) &= f(x(t), u(t), t), \quad x(t_0) = x_0, \quad t \in [t_0, t_1 - 1], \\ u(t) &\in U(t), \quad x(t) \in G_0(t), \quad I(u) = \varphi(x(t_1)) \rightarrow \min. \end{aligned} \right\} \quad (2)$$

Решение этой задачи — траекторию $x_1(t)$ — примем за первое приближение к оптимальной траектории задачи (1). По траектории $x_1(t)$ строим множества $G_1(t)$ и опять методом § 1 решаем вспомогательную задачу (2), где вместо $G_0(t)$ берем $G_1(t)$.

Описанный процесс продолжаем до тех пор, пока последовательные приближения не станут достаточно близкими.

Стандартная процедура, примененная к вспомогательной задаче (2), осуществима полностью (процесс вычислений не остановится из-за расширяющейся сетки) в силу того, что множества $G_k(t)$ строятся около допустимых траекторий.

2. Общий случай. Метод искусственных ограничений можно применить к задаче

$$\left. \begin{aligned} x(t+1) &= f(x(t), u(t), t), \quad t \in [t_0, t_1 - 1], \\ u(t) &\in U(t), \quad x(t) \in G(t), \quad I(u) = \varphi(x(t_1)) \rightarrow \min, \end{aligned} \right\} \quad (3)$$

если естественные ограничения на фазовые координаты не достаточно «жестки» и не позволяют осуществить стандартную процедуру.

Для начала процесса вычислений опять нужно задать нулевое приближение $x_0(t)$, которое должно быть допустимой траекторией в задаче (3) (т. е. удовлетворять всем ограничениям задачи). Построение $x_0(t)$ — нестандартный этап. Остальные этапы полностью повторяют этапы метода из п. 1 с тем изменением, что искусственные ограничения $G_k(t)$ около очередных приближений $x_k(t)$ строятся так, чтобы

$$G_k(t) \subset G(t).$$

Если $G_k(t) = x_k(t)$ при всех $t \in [t_0, t_1]$ за исключением одной точки $t = \tau$, в которой $G_k(\tau) = G(\tau)$, то метод искусственных ограничений называется методом локальных вариаций.

3. Искусственные ограничения в других методах. Как видно из описания вычислительных методов, приведенных в § 1—9, присутствие ограничений на фазовые координаты, как правило, делает метод проще реализуемым. Поэтому разумно во всех подобных случаях вводить искусственные ограничения. Например, искусственные ограничения существенно упрощают вычисления, основанные на множествах достижимости.

Рассмотрим задачу

$$x(t+1) = f(x(t), u(t), t), \quad t \in [t_0, t_1 - 1],$$

$$u(t) \in U(t), \quad x(t) \in G(t), \quad I(u) = \varphi(x(t_1)) \rightarrow \min.$$

Пусть $x_0(t)$ — некоторая допустимая траектория задачи. Исходя из возможностей конкретной машины, зададим множество G_x узлов с «центром» в точке x . Множества достижимости строим по формулам:

$$\Phi(t_0) = G(t_0), \quad \Phi(t+1) = \{x : x = f(y, u, t), \quad y \in \Phi(t),$$

$$u(t) \in U(t), \quad x \in G_{x_0(t+1)}\}.$$

Наличие допустимой траектории $x_0(t)$ гарантирует, что множество $\Phi(t_1)$ непустое. В нем найдем точку $x_1(t_1)$, на которой функция $\varphi(x)$ достигает минимума. По точке $x_1(t_1)$ восстанавливаем траекторию $x_1(t)$, которую принимаем за первое приближение. Процесс вычислений продолжается с заменой $x_0(t)$ на $x_1(t)$.

ГЛАВА IV ● СТОХАСТИЧЕСКИЕ ПРОЦЕССЫ. АДАПТАЦИЯ. ИГРЫ

Первым естественным обобщением детерминированных процессов являются стохастические процессы [3, 5, 7, 9, 23, 25]. Причины введения вероятностных моделей могут быть самые разнообразные. Следующий шаг на пути обобщения математических моделей — процессы с адаптацией [4, 7, 9]. Завершающим звеном этой цепи могут служить процессы управления в конфликтных ситуациях [1, 9]. Ниже все эти процессы исследуются с помощью динамического программирования.

§ 1. Инвариантное погружение. Принцип оптимальности

При изучении задач оптимизации стохастических процессов основные понятия динамического программирования приобретают новое содержание.

1. Инвариантное погружение. Будем рассматривать процессы, которые можно описать уравнениями

$$x(t+1) = f(x(t), u(t), w(t), t),$$

$$dx(t) = f(x(t), u(t), t)dt + g(x(t), \xi(t), t)d\xi(t),$$

где новые по сравнению с гл. I, II символы $w(t)$, $\xi(t)$ характеризуют действие случайных возмущений.

Поскольку теперь процессы $x(t)$ являются случайными, для сравнения их по качеству следует ввести вероятностные характеристики. Наиболее удобной и распространенной характеристикой является математическое ожидание. Зачастую задачи оптимизации стохастических систем формально отличаются от детерминированных аналогов лишь наличием в критериях качества операции математического ожидания. При постановке стохастических задач принимаются различные предположения о характере поступающей информации.

Например, состояния могут точно замеряться в текущий момент, точные измерения могут поступать с запаздыванием, измерения искажаются ошибками и т. п. В этих условиях не всегда сразу по доступным измерениям можно вычислить критерий качества. Требуется из имеющихся данных извлечь нужную информацию, переработать их. В первом приближении можно считать, что последняя задача является самостоятельной и не относится к проблеме управления.

Совокупность величин, доступных измерению к моменту t , обработка которых по принятому правилу позволяет для любого управления вычислить однозначно критерий качества, называется состоянием стохастической системы в момент t (информационное состояние).

Основная цель динамического программирования при оптимизации стохастических процессов остается прежней — получить уравнения для функций Беллмана. Поэтому неизбежно вместо одного конкретного процесса с условиями, заданными в начальный момент, следует рассмотреть семейство процессов при условиях, типичных в любой момент времени. Этот переход, называемый инвариантным погружением, зависит от конкретной задачи, но формально остается таким же, как в детерминированном случае, если заменить соответственно понятие состояния.

2. Принцип оптимальности. Принцип оптимальности для стохастических процессов формально не отличается от детерминированного аналога (см. § 6 гл. I, § 3 гл. II). Однако содержание его значительно сложнее. Как необходимое условие оптимальности он снова выполняется не всегда. В дальнейшем чаще будем пользоваться им как достаточным условием оптимальности. При получении уравнений Беллмана принцип оптимальности важен как способ рассуждения, а успех его определяется свойствами применяемых математических операций. Следует иметь в виду, что в стохастических процессах возможности принципа оптимальности могут быть ограничены из-за двойственной роли управления в некоторых задачах. Если на результаты измерений оказывает влияние управление в предыдущие моменты времени, то не исключено, что критерий качества, подсчитанный в данный момент, будет зависеть от прошлого управления. Таким

образом, с одной стороны, управление участвует как средство для минимизации, с другой — как орудие для получения информации. Это усложняет задачу выбора управления в текущий момент и делает принцип оптимальности, вообще говоря, несправедливым.

§ 2. Оптимизация дискретных систем

Использование динамического программирования для оптимизации стохастических процессов начнем с простейшей задачи терминального управления для дискретной системы, подверженной действиям случайной последовательности векторов [3, 5, 9, 25].

1. Постановка задачи. Пусть процесс описывается уравнением

$$\begin{aligned} x(t+1) &= f(x(t), u(t), w(t), t), \\ x(t_0) &= x_0, \quad t \geq t_0, \end{aligned} \quad (1)$$

где x — n -вектор; u — r -вектор управления; w — q -вектор внешних воздействий; t — дискретное время.

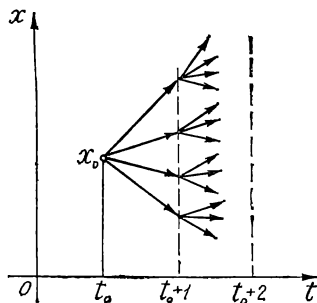


Рис. 61

Сохраняя за остальными обозначениями предположения гл. I, поясним смысл внешних воздействий $w(t)$. Будем считать, что в каждый момент времени t внешнее воздействие $w(t)$ есть случайный вектор с известным распределением $F(t)$; значения $w(t)$, $w(\bar{t})$ в разные моменты t , \bar{t} представляют собой независимые случайные векторы.

Присутствие случайного воздействия усложняет течение процесса. При выбранном управлении $u(t_0)$ из начального вектора x_0 может исходить не определенная траектория, а одна из совокупности, соответствующей различным реализациям вектора $w(t_0)$ (рис. 61). Таким образом, процесс в момент $t=t_0+1$ характеризуется случайным вектором $x(t_0+1)$, распределение которого определяется функцией $F(t_0)$ и преобразованием $x(t_0+1) = f(x_0, u(t_0), w(t_0), t_0)$. Продолжая следить за процессом $x(t)$, при выбранном управлении $u(t)$, $t \geq t_0$, к моменту $t=t_1$ получим случайный вектор $x(t_1)$ с функ-

цией распределения, определяемой $F(t_0), \dots, F(t_1-1)$ и преобразованием (1).

Качество описанного процесса оценим величиной

$$I(u) = M \varphi(x(t_1)), \quad (2)$$

где символ M означает операцию математического ожидания от случайной скалярной величины $\varphi(x(t_1))$, вычисленной по случайному вектору $x(t_1)$.

Поскольку при известном векторе x_0 для любого управления $u(t)$, $t \geq t_0$, значение критерия качества (2) может быть однозначно подсчитано, то в соответствии с § 1 вектор x_0 называется начальным состоянием. При оптимизации стохастических процессов возможны и другие определения начального состояния, которые зависят от постановки конкретных задач, от типа информации.

Задачу теперь можно описать следующим образом:

$$x(t+1) = f(x(t), u(t), w(t), t), x(t_0) = x_0, \quad (3)$$

$$t \in [t_0, t_1 - 1], u(t) \in U(t), I(u) = M \varphi(x(t_1)) \rightarrow \min. \quad (4)$$

При решении задачи оптимизации стохастической системы возможны два соглашения о типе управления процессом: а) в течение процесса измерение реализовавшихся значений вектора $x(t)$ недоступно; б) в каждый момент t известна реализация процесса $x(t)$ (третий случай, когда значение $x(t)$ известно со случайной ошибкой, сводится ко второму и поэтому здесь не рассматривается). Решение в случае а) получается в виде программы управления, в случае б) — в виде управления типа обратной связи. Динамическое программирование более подходит к исследованию случая б).

В данном параграфе будем считать, что реализации процесса $x(t)$ известны в текущий момент и управление строится по типу обратной связи $u(x, t)$.

2. Уравнение Беллмана. По аналогии с детерминированным случаем осуществим прямое погружение системы (3) в семейство

$$x(t+1) = f(x(t), u(t), w(t), t), x(\tau) = y, t \in [\tau, t_1 - 1], \quad (5)$$

где τ , y — фиксированные число и n -вектор. Каждый элемент семейства есть случайный процесс, для оценки качества которого введем функцию Беллмана

$$B(\tau, y) = \min_{\substack{u[x(\tau), x(t_1-1)] \\ x(\tau)=y}} M \varphi(x(t_1)). \quad (6)$$

Здесь минимум берется вдоль допустимых законов управления для процесса (5), а математическое ожидание — по распределениям случайных векторов $w(\tau), \dots, w(t_1-1)$.

Получим уравнение для функции $B(\tau, y)$. Из формулы полной вероятности следует, что

$$\begin{aligned} B(\tau, y) &= \min_{\substack{u[x(\tau), x(t_1-1)] \\ x(\tau)=y}} M \varphi(x(t_1)) = \\ &= \min_{u(x(\tau)) \in U(\tau)} \min_{\substack{u[x(\tau+1), x(t_1-1)] \\ x(\tau+1)=f(y, u(\tau), w(\tau), \tau)}} M [M_{w(\tau)}^{\varphi(x(t_1))}], \end{aligned} \quad (7)$$

где символы M и $M_{w(\tau)}^{\varphi(x(t_1))}$ означают математическое ожидание по $w(\tau), \dots, w(t_1-1)$ и условное математическое ожидание по $w(\tau+1), \dots, w(t_1-1)$ при условии, что значение $w(\tau)$ известно. По предположению, случайные векторы $w(t), t \in [t_0, t_1-1]$, независимы, поэтому условная вероятность равна полной вероятности:

$$M_{w(\tau)}^{\varphi(x(t_1))} = M_{w(\tau+1, t_1-1)}^{\varphi(x(t_1))}.$$

Кроме того, при сделанном предположении относительно типа управления операции $\min_{u[x(\tau+1), x(t_1-1)]}$ и $M_{w(\tau)}$ перестановочны:

$$\min_{u[x(\tau+1), x(t_1-1)]} M(\dots) = M \min_{w(\tau) u[x(\tau+1), x(t_1-1)]} (\dots).$$

Поэтому из (7) получаем

$$B(\tau, y) = \min_{u(x(\tau)) \in U(\tau)} M \left[\min_{\substack{u[x(\tau+1), x(t_1-1)] \\ x(\tau+1)=f(y, u(\tau), w(\tau), \tau)}} M_{w(\tau)}^{\varphi(x(t_1))} \right].$$

В силу определения (6) в квадратных скобках получилось выражение $B(\tau+1, f(y, u(\tau), w(\tau), \tau))$. Таким образом, обратное уравнение Беллмана в задаче (3), (4) принимает вид

$$B(t, x) = \min_{u \in U(t)} M B(t+1, f(x, u, w(t), t)). \quad (8)$$

Граничное условие получается из (6):

$$B(t_1, x) = \varphi(x). \quad (9)$$

Минимальное значение критерия качества — $B(t_0, x_0)$.

Оптимальный закон управления $u(x, t)$ удовлетворяет соотношению

$$\begin{aligned} MB(t+1, f(x, u(x, t), w(t), t)) = \\ = \min_{u \in U(t)} M B(t+1, f(x, u, w(t), t)). \end{aligned}$$

Процесс управления протекает следующим образом. По известному состоянию x_0 находится вектор $u^0(t_0) = u(x_0, t_0)$. Этот вектор подставляется в уравнение (3), и измеряется реализовавшееся значение вектора $x^0(t_0+1)$. Далее находится вектор $u^0(t_0+1) = u(x^0(t_0+1), t_0+1)$, измеряется $x^0(t_0+2)$ и т. д. Последовательность векторов $u^0(t)$, $t \in [t_0, t_1-1]$, — оптимальное управление; $x^0(t)$, $t \in [t_0, t_1]$, — оптимальная траектория. Понятно, что если реализовать оптимальный закон несколько раз при одинаковых условиях, то каждый раз будем получать, вообще говоря, разные $u^0(t)$, $x^0(t)$.

З а м е ч а н и е. Нетрудно проверить, что рассуждения, которые привели к уравнению Беллмана, корректны и тогда, когда распределение вектора $w(t)$ зависит от $t, x(t), u(t)$.

3. Стандартная процедура решения уравнения Беллмана. Численное решение уравнения (8) с граничным условием (9) мало отличается от решения уравнения Беллмана в детерминированном случае. Кроме величин x, u квантуется по уровню и вектор w . При этом исходное распределение вероятностей заменяется на дискретное и считается, что случайный вектор $w(t)$ принимает значения $w^1, \dots, w^{L(t)}$ с вероятностями $p_1(t), \dots, p_{L(t)}(t)$. Тогда уравнение (8) принимает вид

$$B(t, x) = \min_{u \in U(t)} \sum_{i=1}^{L(t)} p_i(t) B(t+1, f(x, u, w^i, t)).$$

Значения $B(t, x)$ на сечении $t=t_1$ фазовой сетки заданы соотношением (9) (рис. 62). Для вычисления значений $B(t_1-1, x)$ выбираем один из векторов $u^1, \dots, u^{L(t_1-1)}$, полученных квантованием управления $u(t_1-1)$. По выбранному вектору u^j вычисляем значение

$$p_1 \varphi(f(x, u^j, w^1, t_1 - 1)) + \dots + \\ + p_{L(t_1-1)} \varphi(f(x, u^j, w^{L(t_1-1)}, t_1 - 1)), \quad (10)$$

используя в случае необходимости интерполирование функции $\varphi(x)$.

Вычисляя выражение (10) для каждого u^j , $j=1, \dots, M(t_1-1)$, выбираем наименьшее. Оно равно $B(t_1-1, x)$, а вектор u^k , на котором достигается минимум, есть значение оптимального закона управления $u^0(x, t_1-1)$.

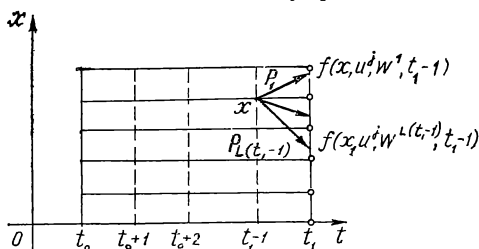


Рис. 62

Меняя x по сечению $t=t_1$ фазовой сетки, получаем функции $B(t_1-1, x)$, $u(x, t_1-1)$. После этого переходим на шаг влево, где проводим аналогичные вычисления. Через t_1-t_0 шагов построим функции $B(t_0, x)$, $u^0(x, t_1-1)$, $u^0(x, t_1-2)$, ..., $u^0(x, t_0)$, которые полностью решают задачу.

Ясно, что для решения уравнения (8) могут быть обобщены вычислительные алгоритмы, описанные в гл. III.

4. Линейные системы. Квадратичный критерий [9]. Рассмотрим частный случай задачи (3), (4):

$$x(t+1) = A(t)x(t) + B(t)u(t) + C(t)w(t), \\ x(t_0) = x_0, \quad t \in [t_0, t_1 - 1], \quad (11)$$

$$I(u) = M\{x(t_1) S x(t_1) + \sum_{t=t_0}^{t_1-1} [x'(t) N(t) x(t) + u'(t) R(t) u(t)]\}, \quad (12)$$

$S \geq 0$, $N(t) > 0$, $R(t) > 0$ — симметрические матрицы.

По аналогии с детерминированным случаем осуществляется прямое погружение системы (11) в семейство

$$x(t+1) = A(t)x(t) + B(t)u(t) + C(t)w(t), \quad x(\tau) = y, \\ t \in [\tau, t_1 - 1], \quad (13)$$

где τ, y — фиксированные число и n -вектор. Уравнение Беллмана имеет вид (п. 2)

$$B(t, x) = x'N(t)x + \min_u [u'(t)R(t)u(t) + \\ + M_{w(t)} B(t+1, A(t)x + B(t)u(t) + C(t)w(t))], \quad (14)$$

$$B(t_1, x) = x'Sx. \quad (15)$$

§ 3. Оптимальное управление марковской цепью

Вместо чисто случайной последовательности, рассмотренной в § 1, можно ввести марковские процессы $w(t)$, в которых значения $w(t)$ и $w(t)$ стохастически связаны [20].

1. Постановка задачи. Пусть $w(t), t \in [t_0, t_1 - 1]$, — марковская цепь, т. е. в каждый момент t вектор $w(t)$ может принимать одно из значений w^1, \dots, w^L с вероятностями $p_1(t), \dots, p_L(t)$ и вероятность перехода из состояния w^k в момент t в состояние w^i в момент $t+1$ равна $p_{ki}(t)$.

Рассмотрим задачу

$$x(t+1) = f(x(t), u(t), w(t+1), t), \quad x(t_0) = x_0, \quad (1)$$

$$t \in [t_0, t_1 - 1], \quad u(t) \in U(t), \quad I(u) = M \varphi(x(t_1)) \rightarrow \min, \quad (2)$$

где $w(t)$ — марковская цепь, а математическое ожидание берется по совокупности случайных векторов $w(t_0+1), \dots, w(t_1)$.

2. Уравнение Беллмана. Как и в § 2, будем искать оптимальное управление типа обратной связи в предположении, что в каждый момент t известны точные значения реализовавшихся векторов $x(t)$ и $w(t)$.

Схема рассуждений остается такой же, как в § 2. Процесс (1) и марковскую цепь $w(t)$ погружаем в семейство

$$x(t+1) = f(x(t), u(t), w(t+1), t), \quad x(\tau) = y, \\ w(\tau) = z = w^k, \quad t \in [\tau, t_1 - 1]. \quad (3)$$

На элементах процессов (3) определим функцию Беллмана:

$$B(\tau, y, z) = \min_{\substack{u[x(\tau), x(t_1-1)] \\ x(\tau)=y}} M_{w(\tau)=z} \varphi(x(t_1)). \quad (4)$$

Воспользовавшись формулой полной вероятности, получим уравнение Беллмана

$$\begin{aligned} B(\tau, y, z) &= \min_{\substack{u(\tau) \in U(\tau) \\ x(\tau)=y, w(\tau)=z}} \min_{\substack{u[x(\tau+1), x(t_1-1)] \\ x(\tau+1)=f(y, u(\tau), w(\tau+1), \tau)}} M_{w(\tau+1)} [M_{w(\tau+1)} \varphi(x(t_1))] = \\ &= \min_{u(\tau) \in U(\tau)} M_{w(\tau+1)} \min_{\substack{u[x(\tau+1), x(t_1-1)] \\ x(\tau+1)=f(y, u(\tau), w(\tau+1), \tau)}} M_{w(\tau+1)} \varphi(x(t_1)). \end{aligned}$$

В силу определения функции (4) заключаем:

$$\begin{aligned} B(t, x, w^k) &= \min_{u \in U(t)} M_{w(t+1)} B(t+1, f(x, u, w(t+1), t), w(t+1)) = \\ &= \min_{u \in U(t)} \sum_{i=1}^L p_i(t+1) B(t+1, f(y, u, w^i, t), w^i) = \\ &= \min_{u \in U(t)} \sum_{i=1}^L p_{ki}(t) B(t+1, f(y, u, w^i, t), w^i), \quad (5) \end{aligned}$$

где $B(t_1, x, z) = \varphi(x)$.

З а м е ч а н и е. Из приведенного вывода уравнения Беллмана видно, что вероятности переходов $p_{ki}(t)$ можно считать функциями $x(t), u(t)$.

3. Стандартная процедура. По сравнению с § 2 в исследуемой задаче (1), (2) размерность аргументов функции Беллмана возросла. Это связано с усложнением задачи по существу. Принципиально же вычислительная схема решения уравнения (5) не отличается от схем, описанных в гл. III и в § 2. Вычисления начинаются с момента $t=t_1$. В узлах сечения $t=t_1$ сетки в пространстве $\{x, w, t\}$ записываются значения $B(t_1, x, z) = \varphi(x)$. Далее рассматриваем узлы сечения $t=t_1-1$ (рис. 63). Фиксируем вектор u^j из множества

$$u^1, \dots, u^{M(t_1-1)}. \quad (6)$$

Вычисляем

$$\sum_{i=1}^L p_{ki}(t_1-1) B(t_1, f(x, u^j, w^i, t_1-1), w^i). \quad (7)$$

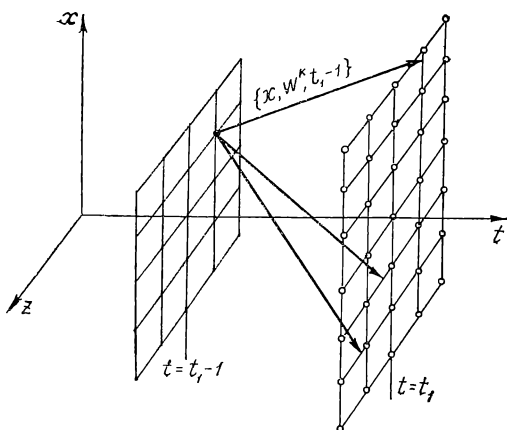


Рис. 63

Перебирая все векторы (6) и вычисляя для них (7), найдем среди результатов наименьшее число. Оно равно, по определению (5), значению $B(t_1-1, x, w^k)$. Дальнейшая процедура вычислений стандартна. Она приводит к функции $B(t_0, x, z)$ и оптимальным законам управления $u(t_0, x, z), \dots, u(t_1-1, x, z)$.

§ 4. Управление в условиях запаздывания информации

До сих пор при оптимизации систем предполагалось, что текущие состояния процессов доступны точному измерению. Поскольку операция измерения, как правило, занимает время, то разумно считать, что в момент, когда к системе прилагается управление, о ее состоянии нет полной информации. В данном параграфе решаются задачи оптимизации в одной из таких ситуаций.

1. Постановка задачи. Рассмотрим задачу

$$\left. \begin{aligned} \dot{x}(t+1) &= f(x(t), u(t), w(t), t), x(t_0) = x_0, t \in [t_0, t_1-1], \\ u(t) &\in U(t), I(u) = M \varphi(x(t_1)) \rightarrow \min. \end{aligned} \right\} (1)$$

Здесь $w(t), t \in [t_0, t_1-1]$, — независимые случайные q -векторы с известными вероятностными характеристиками.

Примем соглашение об условиях управления. Пусть в момент $t=t_0$ состояние $x(t_0)$ системы известно и до

момента $t=t_0+l$ информация о $x(t)$ отсутствует, а в момент $t=t_0+l+1$ известно $x(t_0+1)$. В произвольный момент t известно состояние $x(t-l)$. Кроме того, предположим, что в произвольный момент t в системе хранятся точные значения векторов $u(t-l), \dots, u(t-1)$. На основании этой информации в момент t принимается решение о выборе управления $u(t)$. Поэтому управление $u(t)$ в момент t есть некоторая функция $u(x(t-l), u(t-l), \dots, u(t-1), t)$, зависящая от доступной информации.

2. Уравнение Беллмана. Процесс (1) погрузим в семейство

$$\begin{aligned} x(t+1) &= f(x(t), u(t), w(t), t), \quad x(\tau-l) = y, \\ t &\in [\tau, t_1-1], \quad \tau \geq t_0+l, \\ u(\tau-l) &= u^1, \dots, u(\tau-1) = u^l, \end{aligned}$$

где y, u^1, \dots, u^l — некоторые векторы. Поскольку знание величин τ, y, u^1, \dots, u^l позволяет в момент τ для каждого управления $u(t), t \in [\tau, t_1-1]$, подсчитать величину $M\varphi(x(t_1))$, то всю совокупность y, u^1, \dots, u^l, τ , следуя соглашению, принятому в § 1, назовем состоянием процесса в момент τ .

Введем функцию Беллмана

$$B(y, u^1, \dots, u^l, \tau) = \min_{u[\tau, t_1-1]} M \varphi(x(t_1)). \quad (2)$$

Сделаем пробный шаг: в момент $t=\tau$ выберем некоторое управление $u \in U(\tau)$ и проанализируем получившуюся ситуацию. После выбора управления $u \in U(\tau)$ по принятому соглашению нам станут известны $x(\tau-l+1), u(\tau-l+1) = u^2, \dots, u(\tau-1) = u^l, u(\tau) = u$, по которым можно подсчитать для каждого управления $u(t) \in U(t), t \in [\tau+1, t_1-1]$, величину $M\varphi(x(t_1))$. Среди $u(t) \in U(t), t \in [\tau+1, t_1-1]$, выберем управление $u^0(t)$, минимизирующее величину $M\varphi(x(t_1))$. В силу (2) имеем

$$\min_{u[\tau+1, t_1-1]} M\varphi(x(t_1)) = B(x(\tau-l+1), u^2, \dots, u^l, u, \tau+1). \quad (3)$$

Величина (3) зависит от вероятности реализации значения $x(\tau-l+1)$ и пробного шага u , который теперь

выберем так, чтобы число $M_{x(\tau-l+1)} B(x(\tau-l+1), u^2, \dots, u^l, u, \tau+1)$ стало минимальным. Согласно (2) и достаточности принципа оптимальности, получаем

$$B(y, u^1, \dots, u^l, \tau) = \min_{u \in U(\tau)} M_{w(\tau-l)} B(f(y, u^1, w(\tau-l), \tau-l), u^2, \dots, u^l, u, \tau+1).$$

Пусть $t=t_1$. Тогда

$$B(y, u^1, \dots, u^l, t_1) = M_{\substack{w[t_1-l, \dots, t_1-1] \\ x(t_1-l)=y, u(t_1-l)=u^1, \dots, u(t_1-1)=u^l}} \varphi(x(t_1)).$$

Таким образом, для решения задачи (1) имеем уравнение Беллмана

$$B(x, u^1, \dots, u^l, t) = \min_{u \in U(t)} M_{w(t-l)} B(f(x, u^1, w(t-l), t-l), u^2, \dots, u^l, u, t+1) \quad (4)$$

с граничным условием

$$B(x, u^1, \dots, u^l, t_1) = M_{\substack{w[t_1-l, \dots, t_1-1] \\ x(t_1-l)=x \\ u(t_1-l)=u^1, \dots, u(t_1-1)=u^l}} \varphi(x(t_1)). \quad (5)$$

Процедура решения уравнения (4) с условием (5) такова. Выбираются n -вектор x и r -векторы u^1, \dots, u^l (в узлах соответствующих сеток). Вычисляются случайные векторы

$$\begin{aligned} x^1 &= f(x, u^1, w(t_1-l), t_1-l), \dots, \\ x^l &= x(t_1) = f(x^{l-1}, u^l, w(t_1-1), t_1-1), \end{aligned} \quad (6)$$

распределение которых определяется распределением случайных векторов $w(t_1-l), \dots, w(t_1-1)$. Усредняя значение $\varphi(x(t_1))$ по этим векторам, получаем $B(x, u^1, \dots, u^l, t_1)$. Пусть значение $B(x, u^1, \dots, u^l, t+1)$ известно в узлах сеток для переменных x, u^1, \dots, u^l . Для вычисления значения функции $B(x, u^1, \dots, u^l, t)$ в точке x, u^1, \dots, u^l, t подсчитываем случайный вектор $f(x, u^1, w(t-l), t-l)$ и вычисляем значение функции

$$B(f(x, u^1, w(t-l), t-l), u^2, \dots, u^l, u, t+1), \quad (7)$$

где u — некоторый вектор из $U(t)$. Из-за вектора $w(t-l)$

величина (7) — случайное число. Находим среднее значение

$$M_{w(t_1-l)} B(f(x, u^1, w(t-l), t-l, u^i, \dots, u^l, u, t+1)).$$

Минимизируя его по $u \in U(t)$, получаем значение $B(x, u^1, \dots, u^l, t)$ и оптимальный закон управления $u = u(x, u^1, \dots, u^l, t)$.

Процесс продолжаем до $t = t_0 + l$, когда станут известны функции $B(x, u^1, \dots, u^l, t_0 + l)$, $u(x, u^1, \dots, u^l, t_0 + l)$. Минимальное значение критерия качества:

$$\begin{aligned} \min M \varphi(x(t_1)) &= \min_{u^i \in U(i-1)} B(x_0, u^1, \dots, u^l, t_0 + l) = \\ &= B(x_0, u^{01}, \dots, u^{0l}, t_0 + l). \end{aligned}$$

Оптимальное управление в моменты $t = t_0, \dots, t_0 + l - 1$:

$$u^0(t_0) = u^{01}, \dots, u^0(t_0 + l - 1) = u^{0l}. \quad (8)$$

Тип управления в рассматриваемой задаче можно трактовать как промежуточный между управлением по типу обратной связи и программным управлением. При $l=0$ (нет задержки в передаче информации, измерения происходят мгновенно) из (4), (5) получается результат § 2, где оптимальное управление имеет вид $u^0(x, t)$. При $l=t_1-t_0$ (в процессе управления измерения не производятся или на обработку информации уходит времени больше, чем длительность процесса) уравнение (4) становится излишним, остается лишь краевое условие (5). Оптимальное управление получается из (8) как функция времени.

§ 5. Оптимизация непрерывных стохастических систем

При оптимизации непрерывных стохастических систем принципиально идеи динамического программирования остаются такими же, как в дискретном случае. Однако резко возрастают аналитические трудности описания и анализа задач. Аппарат непрерывных стохастических процессов [14] несравненно сложнее тех элементарных фактов теории вероятностей, которых было достаточно в § 1—2. Поэтому последующее изложение § 7 будет несколько формальным.

1. Стохастическое уравнение. Формула Ито. Случайный (скалярный) процесс $\alpha(t), t \in [t_0, t_1]$, называется винеровским (или процессом броуновского движения), если для любых $t_0 = \tau_0 < \tau_1 < \dots < \tau_k = t_1$ случайные величины $\alpha(\tau_1) - \alpha(\tau_0), \dots, \alpha(\tau_k) - \alpha(\tau_{k-1})$ независимы и имеют нормальное распределение, математическое ожидание и дисперсия которого

$$M[\alpha(t+h) - \alpha(t)] = 0, D[\alpha(t+h) - \alpha(t)] = h.$$

Рассмотрим разностное скалярное стохастическое уравнение

$$x(t + \Delta t) - x(t) = f(x(t), t) \Delta t + g(x(t), t)[\alpha(t + \Delta t) - \alpha(t)], \quad (1)$$

где $\alpha(t)$ — винеровский процесс. Пусть задано начальное условие $x(t_0) = x_0$ ($x(t_0)$ может быть детерминированной или случайной величиной, не зависящей от $\alpha(t)$). Тогда $x(t)$, определенное уравнением (1) и начальным условием $x(t_0)$ в момент $t_0 + k \Delta t (k \geq 1)$, можно представить в виде

$$\begin{aligned} x(t_0 + k \Delta t) = & x(t_0) + \sum_{i=0}^{k-1} f(x(t_0 + i \Delta t), t_0 + i \Delta t) \Delta t + \\ & + \sum_{i=0}^{k-1} g(x(t_0 + i \Delta t), t_0 + i \Delta t) [\alpha(t_0 + (i+1) \Delta t) - \\ & - \alpha(t_0 + i \Delta t)]. \end{aligned} \quad (2)$$

Предположим, что задана дважды непрерывно дифференцируемая по x и непрерывно дифференцируемая по t функция $v(x, t)$. Найдем формулу для приращения $M[v(x(t_0 + k \Delta t), t_0 + k \Delta t) - v(x_0, t_0)]$ при условии, что Δt мало.

Разлагая разность $v(x(t + \Delta t), t + \Delta t) - v(x(t), t)$ (t — число вида $t_0 + s \Delta t$) по формуле Тейлора и учитывая, что $x(t)$ удовлетворяет уравнению (1), получаем:

$$\begin{aligned} v(x(t + \Delta t), t + \Delta t) - v(x(t), t) = & \frac{\partial v(x(t), t)}{\partial t} \Delta t + \\ & + \frac{\partial v(x(t), t)}{\partial x} (f(x(t), t) \Delta t + g(x(t), t) [\alpha(t + \Delta t) - \alpha(t)]) + \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{2} \cdot \frac{\partial^2 v(x(t), t)}{\partial x^2} (f(x(t), t) \Delta t + g(x(t), t) [\alpha(t + \Delta t) - \\
& - \alpha(t)])^2 + o(\Delta t) = \left(\frac{\partial v(x(t), t)}{\partial t} + \frac{\partial v(x(t), t)}{\partial x} f(x(t), t) + \right. \\
& + \frac{1}{2} \cdot \frac{\partial^2 v(x(t), t)}{\partial x^2} g^2(x(t), t) \frac{[\alpha(t + \Delta t) - \alpha(t)]^2}{\Delta t} \Big) \Delta t + \\
& + \frac{\partial v(x(t), t)}{\partial x} g(x(t), t) [\alpha(t + \Delta t) - \alpha(t)] + o(\Delta t).
\end{aligned}$$

Так же, как при переходе от (1) к (2), получаем

$$\begin{aligned}
& v(x(t_0 + k \Delta t), t_0 + k \Delta t) = v(x_0, t_0) + \\
& + \sum_{i=0}^{k-1} \left\{ \frac{\partial v(x(t_0 + i \Delta t), t_0 + i \Delta t)}{\partial t} + \frac{\partial v(x(t_0 + i \Delta t), t_0 + i \Delta t)}{\partial x} \times \right. \\
& \times f(x(t_0 + i \Delta t), t_0 + i \Delta t) + \frac{1}{2} \cdot \frac{\partial^2 v(x(t_0 + i \Delta t), t_0 + i \Delta t)}{\partial x^2} \times \\
& \times g^2(x(t_0 + i \Delta t), t_0 + i \Delta t) \frac{[\alpha(t_0 + (i+1) \Delta t) - \alpha(t_0 + i \Delta t)]^2}{\Delta t} \Big\} \Delta t + \\
& + \sum_{i=0}^{k-1} \frac{\partial v(x(t_0 + i \Delta t), t_0 + i \Delta t)}{\partial x} g(x(t_0 + i \Delta t), t_0 + i \Delta t) \\
& \times [\alpha(t_0 + (i+1) \Delta t) - \alpha(t_0 + i \Delta t)] + o(\Delta t). \quad (3)
\end{aligned}$$

Как следует из (1), значение процесса $x(t)$ в момент t зависит от разности $\alpha(t) - \alpha(t - \Delta t)$. Так как приращения процесса $\alpha(t)$ на непересекающихся интервалах независимы, то при любом фиксированном t случайные величины $x(t)$ и $\alpha(t + \Delta t) - \alpha(t)$ независимы. Поэтому

$$\begin{aligned}
& M\{\beta(x(t), t) [\alpha(t + \Delta t) - \alpha(t)]^s\} = \\
& = M\beta(x(t), t) M[\alpha(t + \Delta t) - \alpha(t)]^s
\end{aligned}$$

при любой непрерывной функции $\beta(x, t)$.

Если теперь учесть, что $M[\alpha(t + \Delta t) - \alpha(t)] = 0$, $M[\alpha(t + \Delta t) - \alpha(t)]^2 = \Delta t$ при любых $t, \Delta t$, и применить к формуле (3) операцию математического ожидания, то для приращения $M[v(x(t), t) - v(x_0, t_0)]$ получим формулу

$$\begin{aligned}
M[v(x(t_0 + k \Delta t), t_0 + k \Delta t) - v(x_0, t_0)] = \\
= \sum_{i=0}^{k-1} M \left\{ \frac{\partial v(x(t_0 + i \Delta t), t_0 + i \Delta t)}{\partial t} + \right. \\
+ \frac{\partial v(x(t_0 + i \Delta t), t_0 + i \Delta t)}{\partial x} f(x(t_0 + i \Delta t), t_0 + i \Delta t) + \\
+ \frac{1}{2} \cdot \frac{\partial^2 v(x(t_0 + i \Delta t), t_0 + i \Delta t)}{\partial x^2} g^2(x(t_0 + i \Delta t), \\
\left. t_0 + i \Delta t) \right\} \Delta t + o(\Delta t). \quad (4)
\end{aligned}$$

Пусть задан отрезок $[t_0, t_1]$. В качестве Δt возьмем величину $\Delta t = \frac{t_1 - t_0}{k}$. Тогда на сумму (4) можно смотреть как на интегральную сумму Римана. При $\Delta t \rightarrow 0$ имеем

$$M(v(\tilde{x}(t), t) - v(x_0, t_0)) = \int_{t_0}^{t_1} ML v(\tilde{x}(s), s) ds, \quad (5)$$

где

$$\begin{aligned}
L v(x, t) = \frac{\partial v(x, t)}{\partial t} + \frac{\partial v(x, t)}{\partial x} f(x, t) + \\
+ \frac{1}{2} \cdot \frac{\partial^2 v(x, t)}{\partial x^2} g^2(x, t),
\end{aligned}$$

а под $\tilde{x}(t)$ понимается предел в среднем квадратичном последовательности решений $x(t)$ разностного уравнения (1), т. е.

$$\lim_{\Delta t \rightarrow 0} M(x(t) - \tilde{x}(t))^2 = 0.$$

При $\Delta t \rightarrow 0$ разностное уравнение (1) формально переходит в выражение

$$dx(t) = f(x(t), t) dt + g(x(t), t) d\alpha(t), \quad (6)$$

которое называется уравнением Ито. Под решением $\tilde{x}(t)$ уравнения Ито понимается предел в среднем

квадратичном полигонов, построенных по решениям разностного уравнения (1).

Формулу (5) можно обобщить на стохастические системы Ито:

$$dx(t) = f(x(t), t)dt + \sum_{i=1}^k g_i(x(t), t) d\alpha_i(t), \quad (7)$$

где $x(t), f, g_i$ — n -мерные вектор-функции; $\alpha_i(t)$ — независимые скалярные винеровские процессы.

Систему (7) можно понимать как предельный случай системы разностных уравнений

$$x(t + \Delta t) - x(t) = f(x(t), t) \Delta t + \sum_{i=1}^k g_i(x(t), t) [\alpha_i(t + \Delta t) - \alpha_i(t)].$$

Легко подсчитать, что для любой скалярной функции $v(x, t)$, у которой производные $\frac{\partial v}{\partial t}, \frac{\partial v}{\partial x_i}, \frac{\partial^2 v}{\partial x_i \partial x_j}$ существуют и непрерывны, выполняется соотношение (5), если под $Lv(x, t)$ понимать величину

$$Lv(x, t) = \frac{\partial v(x, t)}{\partial t} + f'(x, t) \frac{\partial v(x, t)}{\partial x} + \\ + \frac{1}{2} \sum_{i=1}^k g'_i(x, t) \frac{\partial^2 v(x, t)}{\partial x^2} g_i(x, t). \quad (8)$$

Формула (5) получена из (3) с помощью усреднения и последующего предельного перехода ($\Delta t \rightarrow 0$). Если же перейти к пределу в самом выражении (3), то формально получим

$$v(x(t), t) - v(x_0, t_0) = \int_{t_0}^t Lv(x(\tau), \tau) d\tau + \\ + \int_{t_0}^t \frac{\partial v(x(\tau), \tau)}{\partial x} g(x(\tau), \tau) d\alpha(\tau). \quad (9)$$

Выражение (9) называется формулой Ито. Она отличается от обычной формулы

$$v(x(t), t) - v(x_0, t_0) = \int_{t_0}^t \frac{\partial v(x(\tau), \tau)}{\partial t} d\tau + \\ + \int_{t_0}^t \frac{\partial v(x(\tau), \tau)}{\partial x} g(x(\tau), \tau) d\alpha(\tau) \quad (10)$$

для уравнения (7), где под $\alpha(t)$ понимается неслучайная функция ограничений вариации, тем, что под первым интегралом в (9) стоит оператор L вместо d/dt , стоящего в (10). Как видно из (8), значения операторов отличаются членом $\frac{1}{2} \sum g^2 \frac{\partial^2 v}{\partial x^2}$.

2. Оптимизация нелинейных систем. Пусть стохастический процесс $x(t)$ описывается уравнением

$$dx(t) = f(x, u, t) dt + \sum_{i=1}^k g_i(x, u, t) d\alpha_i(t), \quad x(t_0) = x_0, \quad (11)$$

где $f(x, u, t), g_i(x, u, t)$ — непрерывные функции n -вектора x , r -вектора u и скаляра t ; $\alpha_i(t)$, $i = \overline{1, k}$ — независимые винеровские процессы.

Закон управления $u = u(x, t) \in U$ назовем допустимым, если при его подстановке в (11) уравнение имеет решение $x(t)$.

Среди допустимых законов управлений требуется найти такой $u^0(x, t)$, при котором минимален функционал

$$I(u) = \int_{t_0}^{t_1} M f_0(x, u, t) dt.$$

Погрузим процесс (11) в семейство

$$dx(t) = f(x, u, t) dt + \sum_{i=1}^k g_i(x, u, t) d\alpha_i(t), \quad t \in [\tau, t_1], \\ x(\tau) = y,$$

и определим функцию Беллмана

$$B(\tau, y) = \min_{\substack{u[x(\tau), x(t_1)] \\ x(\tau) = y}} \int_{\tau}^{t_1} M f_0(x, u, t) dt. \quad (12)$$

Чтобы получить уравнение Беллмана, рассмотрим управление

$$u(x, t) = \begin{cases} v \in U, & \tau \leq t < \tau + \Delta\tau, \\ u^0(x(t), t), & \tau + \Delta\tau \leq t \leq t_1, \end{cases}$$

где $u^0(x(t), t)$ — оптимальный закон управления на $[\tau + \Delta\tau, t_1]$. Вдоль него выполняется неравенство

$$\begin{aligned} B(\tau, y) &\leq \int_{\tau}^{\tau+\Delta\tau} Mf_0(x, v, t) dt + \\ &+ \min_{u[x(\tau+\Delta\tau), x(t_1)]} \int_{\tau+\Delta\tau}^{t_1} Mf_0(x, u, t) dt = \\ &= \int_{\tau}^{\tau+\Delta\tau} Mf_0(x, v, t) dt + B(\tau + \Delta\tau, x(\tau + \Delta\tau)). \end{aligned} \quad (13)$$

С другой стороны, для оптимального закона $u^0(x(t), t)$, $t \in [\tau, t_1]$,

$$B(\tau, y) = \int_{\tau}^{\tau+\Delta\tau} Mf_0(x, u^0(x, t), t) dt + B(\tau + \Delta\tau, x^0(\tau + \Delta\tau)). \quad (14)$$

Считая функцию $B(\tau, y)$ достаточно гладкой, из (13) при $\Delta\tau \rightarrow 0$ с учетом формулы Ито (9) получаем неравенство

$$-LB(\tau, y) \Big|_{(11), v} \leq f_0(y, v, \tau),$$

а из (14) — равенство

$$-LB(\tau, y) \Big|_{(11), u^0} = f_0(y, u^0(y, \tau), \tau).$$

Таким образом, уравнение Беллмана принимает вид

$$\min_{u \in U} [LB(\tau, y) \Big|_{(11), u} + f_0(y, u, \tau)] = 0. \quad (15)$$

Граничное условие следует из (12): $B(t_1, y) = 0$.

По внешним признакам уравнение (15) мало отличается от уравнения Беллмана для детерминированных процессов (см. гл. II). Однако теперь оно представляет собой уравнение в частных производных уже второго порядка.

Как и в детерминированном случае, с помощью уравнения (15) можно сформулировать достаточные условия оптимальности, исследовать линейные системы с квадратичным критерием качества.

§ 6. Процессы управления с адаптацией

В § 2 рассматривались задачи управления стохастическими системами, на которые действуют независимые случайные возмущения с известными характеристиками. Размерность задачи с точки зрения динамического программирования по сравнению с детерминированным случаем не увеличилась. В задаче § 3, где внешние воздействия составляли марковскую цепь, размерность задачи увеличилась. Задачи с марковской цепью весьма близки к задачам с адаптацией в том смысле, который вкладывают в последние на современном этапе развития теории управления. Рассмотрим простейшую задачу с адаптацией, когда вероятностные характеристики внешнего воздействия неизвестны, но уточняются по мере развития процесса управления.

1. Постановка задачи. Пусть внешнее воздействие $w(t)$ представляет собой последовательность независимых случайных величин, принимающих в каждый момент лишь два значения:

$$w(t) = \begin{cases} 1 & \text{с вероятностью } p, \\ 0 & \text{с вероятностью } 1 - p. \end{cases}$$

Будем считать, что значение величины p не зависит от времени, но неизвестно, а по ходу процесса могут измеряться реализовавшиеся значения $w(t)$. В таком случае можно в начале процесса управления высказать гипотезу о распределении случайного числа p . Первая же реализация $w(1)$ позволяет уточнить исходную гипотезу. Воспользуемся простейшим правилом уточнения гипотез — правилом Байеса. Пусть к моменту t случайная величина $w(t)$ приняла m раз значение, равное 1; $F(p)$ — априорная функция распределения величины p к началу процесса управления; $F_m(t, p)$ — функция распределения, полученная после обработки поступивших данных. Согласно правилу Байеса,

$$dF_m(t, p) = \frac{p^m (1-p)^{t-m} dF(p)}{\int_0^1 p^m (1-p)^{t-m} dF(p)},$$

где $dF(p)$ означает вероятность того, что величина p принимает значения из множества $(p, p+dp)$.

Поэтому вероятность $p_m(\tau)$ реализации в момент $t = \tau$ значения $w(\tau) = 1$, если известно, что m раз до момента t оказалось $w = 1$, равна

$$p_m(\tau) = \frac{\int_0^1 p^{m+1} (1-p)^{\tau-m} dF(p)}{\int_0^1 p^m (1-p)^{\tau-m} dF(p)}. \quad (1)$$

Рассмотрим задачу

$$\left. \begin{aligned} x(t+1) &= f(x(t), u(t), w(t), t), \quad x(t_0) = x_0, \\ t &\in [t_0, t_1 - 1], \\ u(t) &\in U(t), \quad I(u) = M \varphi(x(t_1)) \rightarrow \min, \end{aligned} \right\} \quad (2)$$

где $w(t)$ — случайное воздействие описанного выше типа.

Будем считать, что в каждый момент t известно состояние $x(t)$ и реализовавшееся значение $w(t)$. Информация о $w(t)$ полностью заключена в исходном распределении $F(p)$ и числе m реализовавшихся к моменту t верхних значений $w(t)$.

2. Уравнение Беллмана. Процесс (2) погрузим в семейство

$$x(t+1) = f(x(t), u(t), w(t), t), \quad x(\tau) = y, \quad t \in [\tau, t_1 - 1], \quad (3)$$

и будем считать, что к моменту $t = \tau$ реализовалось m единиц. Введем функцию

$$B(\tau, y, m) = \min_{u \in [\tau, t_1 - 1]} M \varphi(x(t_1)). \quad (4)$$

Для получения уравнения Беллмана применим традиционное рассуждение. После пробного шага $u(\tau) = u \in U(\tau)$ в момент $t = \tau$ возможны два исхода: а) с вероятностью $p_m(\tau)$ система (3) окажется в положении $f(y, u, I, \tau)$, а число единичных значений w дойдет до $m+1$; б) с вероятностью $1 - p_m(\tau)$ система (3) окажется в поло-

жении $f(y, u, 0, \tau)$, а число единичных значений w останется прежним. Допустим, что с момента $t = \tau + 1$ управление оптимально. В случае а) на подпроцессе $x(t)$, $t \in [\tau + 1, t_1 - 1]$, критерий качества равен числу $B(\tau + 1, f(y, u, 1, \tau), m + 1)$; в случае б) — числу $B(\tau + 1, f(y, u, 0, \tau), m)$. Среднее значение критерия качества после пробного шага равно величине

$$p_m(\tau) B(\tau + 1, f(y, u, 1, \tau), m + 1) + (1 - p_m(\tau)) B(\tau + 1, f(y, u, 0, \tau), m) \quad (5)$$

и зависит от величины $u = u(\tau)$ на пробном шаге. Минимизируя (5) по $u \in U(\tau)$, получим минимальное значение критерия качества для всего процесса. Уравнение Беллмана примет вид

$$B(\tau, y, m) = \min_{u \in U(\tau)} [p_m(\tau) B(\tau + 1, f(y, u, 1, \tau), m + 1) + (1 - p_m(\tau)) B(\tau + 1, f(y, u, 0, \tau), m)]. \quad (6)$$

Краевое условие следует из (4):

$$B(t_1, y, m) = \varphi(y). \quad (7)$$

Решение уравнения (6) осуществляется по следующей схеме. В момент $t = t_1$ функция $B(t_1, y, m)$ известна (см. (7)). Полагаем $t = t_1 - 1$. При фиксированных y, m , $0 \leq m \leq t_1 - t_0 - 1$, вычисляем $p_m(t_1 - 1)$ по формуле (1). Из (6) находим $B(t_1 - 1, y, m)$. Перебирая все узлы сетки для y и числа m , $0 \leq m \leq t_1 - t_0 - 1$, определяем функцию $B(t_1 - 1, y, m)$. Аналогичным образом вычисления продолжаются до момента $t = t_0$. Величина $B(t_0, x_0, 0)$ равна минимальному значению критерия качества. Оптимальный закон управления получается попутно из элементов $u^0 \in U(t)$, на которых в правой части (6) достигается минимум.

§ 7. Процессы управления в конфликтных ситуациях

Внешние воздействия можно приписать другому участнику, который влияет на ход процесса управления, но интересы которого не совпадают с нашими. В процессах управления, учитывающих подобные ситуации конфликтного типа, большое значение имеет степень информиро-

ванности участников о действиях противника [1]. Здесь возможны два случая: а) один из участников знает очередной ход другого; б) участники делают ходы одновременно и не могут определенно знать о ходах противника.

1. Игра против природы. Пусть процесс описывается уравнением

$$x(t+1) = f(x(t), u(t), w(t), t), x(t_0) = x_0, \quad (1)$$

где $u(t)$ — управление участника 1; $w(t)$ — управление участника 2.

Будем считать, что имеются ограничения на управления участников: $u(t) \in U(t)$, $w(t) \in W(t)$, критерий качества задан в виде

$$I(u, w) = \varphi(x(t_1)). \quad (2)$$

Процесс игры протекает таким образом. Для каждого состояния x , реализовавшегося в момент t , участник 1 выбирает управление $u(x, t) \in U(t)$; участник 2, зная $u(x, t)$, выбирает управление $w(x, t) \in W(t)$. Подстановка этих значений в (1) дает состояние $x(t+1)$. В состоянии $x(t+1)$ процедура повторяется: сначала выбирается управление $u(x(t+1), t+1)$, затем $w(x(t+1), t+1)$. Цель участника 1 — доставить критерию качества (2) минимальное значение, цель участника 2 — добиться, чтобы критерий качества (2) достиг максимального значения.

Аналитически описанную процедуру выбора управлений можно записать так:

$$I(u^0, w^0) = \min_{u[x(t_0)]} \max_{w[x(t_0)]} \min_{u[x(t_0+1)]} \max_{w[x(t_0+1)]} \dots \\ \dots \min_{u[x(t_1-1)]} \max_{w[x(t_1-1)]} \varphi(x(t_1)).$$

2. Уравнение Беллмана. Погрузим процесс (1) в семейство

$x(t+1) = f(x(t), u(t), w(t), t)$, $x(\tau) = y$, $t \in [\tau, t_1 - 1]$,
на котором определим функцию Беллмана:

$$B(\tau, y) = \min_{\substack{u[y] \\ x(\tau)=y}} \max_{w[y]} \min_{u[x(\tau+1)]} \max_{w[x(\tau+1)]} \dots \\ \dots \min_{u[x(t_1-1)]} \max_{w[x(t_1-1)]} \varphi(x(t_1)). \quad (3)$$

Для функции $B(\tau, x)$ очевидны равенства

$$B(\tau, y) = \min_{u(\tau) \in U(\tau)} \max_{w(\tau) \in W(\tau)} [\min_{\substack{u[x(\tau+1)] \\ x(\tau+1)=f(y, u(\tau), w(\tau), \tau)}} \max_{w[x(\tau+1)]} \dots \\ \dots \min_{u[x(t_1-1)]} \max_{w[x(t_1-1)]} \varphi(x)(t_1)] = \min_{u \in U(\tau)} \max_{w \in W(\tau)} B(\tau+1, \\ f(y, u, w, \tau)).$$

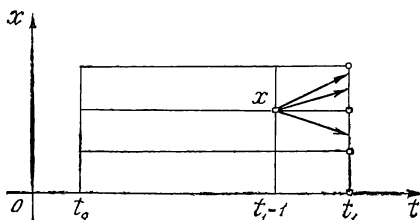


Рис. 64

Таким образом, получая граничные условия из (3), приходим к уравнению Беллмана

$$B(t, x) = \min_{u \in U(t)} \max_{w \in W(t)} B(t+1, f(x, u, w, t)), \quad (4)$$

$$B(t_1, x) = \varphi(x). \quad (5)$$

Решая уравнение (4) справа налево, найдем $B(t_0, x)$ и оптимальные законы управлений участников $u(x, t)$, $w(x, t)$:

$$B(t+1, f(x, u(x, t), w(x, t), t)) = \\ = \min_{u \in U(t)} \max_{w \in W(t)} B(t+1, f(x, u, w, t)).$$

Значение $B(t_0, x_0)$ равно $I(u^0, w^0)$. Оптимальное управление $u^0(t)$ и траектория $x^0(t)$ при оптимальном поведении второго участника находятся из соотношений: $u^0(t_0) = u(x_0, t_0)$, $w^0(t_0) = w(x_0, t_0)$, $x^0(t_0+1) = f(x_0, u^0(t_0), w^0(t_0), t_0), \dots$, $u^0(t) = u(x^0(t), t)$, $x^0(t+1) = f(x^0(t), u^0(t), w^0(t), t)$, \dots .

Уравнение (4) принципиально не отличается от уравнений Беллмана, рассмотренных в гл. I, поэтому вычислительные алгоритмы, описанные в гл. III, нетрудно приспособить для решения уравнения (4). Например, для вычисления значений $B(t_1-1, x)$ (значения $B(t_1, x)$ из-

вестны в силу (5)) наряду с управлением u квантуем по уровню и управление $\omega: \omega^1, \dots, \omega^{L(t)}$, где $L(t)$ — число узлов, содержащихся в $W(t)$. Для каждого узла $\{x, t_1 - 1\}$ фазовой сетки (рис. 64) составляем $M(t_1 - 1) \times L(t_1 - 1)$ -матрицу A с элементами a_{ij} :

$$a_{ij} = B(t_1, f(x, u^i, \omega^j, t_1 - 1)).$$

В построенной матрице находим минимаксную точку $\{i_0, j_0\}$ $a_{i_0 j_0} = \min_i \max_j a_{ij}$. Управления $u(x, t_1 - 1) = u^{i_0}$, $\omega(x, t_1 - 1) = \omega^{j_0}$ запоминаем вместе со значением $B(t_1 - 1, x) = B(t_1, f(x, u^{i_0}, \omega^{j_0}, t_1 - 1))$. Остальные операции такие же, как в § 1 гл. III.

3. Игра неравноинформированных участников. Игру, рассмотренную в п. 1, можно считать игрой, когда участник 2 (природа) более информирован, чем участник 1. Не меняя цели игроков, изменим предположение об информированности. Пусть участник 1 на каждом шаге знает очередной ход участника 2 (иначе говоря, на каждом шаге участник 2 выбирает свое управление первым). Оставим прежним и соглашение о типе используемых управлений и уравнениях процесса. Тогда задача сводится к нахождению таких законов управления $u = u(x, t)$, $\omega = \omega(x, t)$, что

$$I(u^0, \omega^0) = \max_{\omega[x(t_0)]} \min_{u[x(t_0)]} \dots \max_{\omega[x(t_1-1)]} \min_{u[x(t_1-1)]} \varphi(x(t_1)).$$

Нетрудно видеть, что рассуждения п. 2 приводят в данном случае к уравнению Беллмана:

$$B(t, x) = \max_{\omega \in W(t)} \min_{u \in U(t)} B(t+1, f(x, u, \omega, t)),$$

$$B(t_1, x) = \varphi(x),$$

которое решается так же, как уравнение (4).

4. Игра равноинформированных участников. Как и в теории матричных игр, ситуация серьезно усложняется, если участники делают ходы одновременно, не зная об очередных ходах противника.

Пусть множества $U(t)$, $W(t)$, $t \in [t_0, t_1 - 1]$, состоят из конечного набора векторов. В начале каждого хода участники не сами делают выбор управления, а поручают его случайному механизму. Каждый игрок может влиять на работу этого механизма в том смысле, что

может задать закон распределения вероятностей, который имеет случайный вектор управления, указанный на очередном ходе. Будем считать, что значения случайных векторов $u(t)$, $w(t)$, $t \in [t_0, t_1-1]$, независимы. Игра по указанному правилу приведет процесс $x(t)$ к моменту $t=t_1$ в состояние $x(t_1)$, которое, очевидно, является случайным вектором. Одним из возможных способов оценки поведения участников является такой:

$$I(p^0, q^0) = \min_{p(x(t_0))} \max_{q(x(t_0))} \min_{p(x(t_0+1))} \max_{q(x(t_0+1))} \dots \\ \dots \min_{p(x(t_1-1))} \max_{q(x(t_1-1))} M \varphi(x(t_1)). \quad (6)$$

Здесь $p(x) = \{p^1(x), \dots, p^{K(t)}(x)\}$ — распределение вероятностей вектора u в момент t , когда процесс находится в состоянии x ; $q(x) = \{q^1(x), \dots, q^{L(t)}(x)\}$ — распределение вероятностей вектора w в момент t , когда процесс находится в состоянии x ; символ M означает математическое ожидание, вычисленное по всем распределениям $p(x(t_0)), \dots, p(x(t_1-1)), q(x(t_0)), \dots, q(x(t_1-1))$.

В силу теоремы о минимаксе из теории матричных игр операции \min и \max в пределах одного шага перестановочны. Поэтому

$$I(p^0, q^0) = \max_{q(x(t_0))} \min_{p(x(t_0))} \min_{p(x(t_0+1))} \max_{q(x(t_0+1))} \dots \\ \dots \min_{p(x(t_1-1))} \max_{q(x(t_1-1))} M \varphi(x(t_1)). \quad (7)$$

Погрузив процесс (1) в семейство процессов и воспользовавшись равенствами (6), (7), приходим к следующему уравнению Беллмана:

$$B(t, x) = \min_{\substack{0 \leq p_i \leq 1 \\ \sum p_i = 1}} \max_{\substack{0 \leq q_j \leq 1 \\ \sum q_j = 1}} \left[\sum_{i,j=1}^{K(t), L(t)} p_i q_j B(t+1, f(x, u^i, w^j, t)) \right] = \\ = \max_{0 \leq q_i \leq 1} \min_{0 \leq p_i \leq 1} [\dots], \quad B(t_1, x) = \varphi(x).$$

По структуре полученное уравнение напоминает уравнение Беллмана в простейшей задаче терминального управления. Поэтому и методы его решения очевидным образом следуют из процедур, описанных в гл. III.

5. Дифференциальные игры. Если непрерывные процессы управления рассматриваются в конфликтных ситуациях, а уравнения движения являются дифференциальными, принято говорить о дифференциальных играх. Дифференциальные игры можно рассматривать как предельный случай дискретных игр, изученных в пп. 1—3.

От уравнения $\dot{x} = f(x, u, v, t)$, $x(t_0) = x_0$, квантованием по времени (см. § 1 гл. I) переходим к уравнению (1), в терминах которого задача ставится вполне определено. Затем в уравнении Беллмана осуществляем предельный переход при $\Delta t \rightarrow 0$.

Для примера рассмотрим непрерывный аналог задачи п. 1. Переходя к дискретному аргументу t , получаем

$$x(t + \Delta t) = x(t) + \Delta t f(x(t), u(t), v(t), t), \\ t = t_0, t_0 + \Delta t, \dots, t_1 - \Delta t.$$

При условиях п. 1 уравнение Беллмана таково:

$$B(t, x) = \min_{u \in U(t)} \max_{v \in V(t)} B(t + \Delta t, x + \Delta t f(x, u, v, t)), \\ B(t_1, x) = \varphi(x). \quad (8)$$

Предположим, что функция $B(t, x)$ дифференцируема по аргументам. Тогда

$$B(t + \Delta t, x + \Delta t f(x, u, v, t)) = B(t, x) + \\ + \frac{\partial B(t, x)}{\partial t} \Delta t + \frac{\partial B(t, x)}{\partial x} f(x, u, v, t) \Delta t + o(\Delta t).$$

Поэтому, подставив это выражение в (8), поделив обе части на Δt и устремив $\Delta t \rightarrow 0$, получим дифференциальное уравнение для функции Беллмана:

$$-\frac{\partial B(t, x)}{\partial t} = \min_{u \in U(t)} \max_{v \in V(t)} \frac{\partial B(t, x)}{\partial x} f(x, u, v, t), \quad (9) \\ B(t_1, x) = \varphi(x).$$

С аналитической точки зрения уравнение (9) представляет еще более сложный объект, чем дифференциальное уравнение Беллмана для процессов управления с одним участником (см. § 4 гл. II).

Уравнение (9) можно использовать для получения достаточных условий оптимальности для законов управления участниками. Ограничимся простым утверждением.

Пусть $B(t, x)$ — гладкое решение уравнения (9); $u(x, t)$, $v(x, t)$ — законы управления, удовлетворяющие условиям

$$\begin{aligned} & -\frac{\partial' B(t, x)}{\partial x} f(x, u(x, t), v(x, t), t) = \\ & = \min_{u \in U(t)} \max_{v \in V(t)} \frac{\partial' B(t, x)}{\partial x}, f(x, u, v, t) \end{aligned}$$

и такие, что уравнение $\dot{x} = f(x, u(x, t), v(x, t), t)$, $x(t_0) = x_0$, имеет единственное решение, продолжимое на отрезок $[t_0, t_1]$. Тогда $u(x, t)$ и $v(x, t)$ — оптимальные законы управления для первого и второго участников (соответственно).

6. Линейные системы. Квадратичный критерий [9].

Пусть система управления имеет вид

$$\dot{x} = A(t)x + B(t)u + C(t)v, \quad t \in [t_0, t_1], \quad (10)$$

где x — n -вектор; u — r -вектор; v — q -вектор; A , B , C — матрицы соответствующих размерностей.

Качество процессов управления оценим критерием

$$I(u, v) = \int_{t_0}^{t_1} (x' M(t)x + u' R(t)u + v' Q(t)v) dt, \quad (11)$$

где симметрические матрицы M , R и Q удовлетворяют условиям: $M(t) \geq 0$, $R(t) > 0$, $Q(t) < 0$ (на управления $u(t)$, $v(t)$ ограничения не накладываются).

Под дифференциальной игрой для системы (10) с критерием (11) будем понимать предельный вариант дискретной игры, описанной в п. 1, с естественной заменой критерия качества.

Уравнение Беллмана для сформулированной дифференциальной игры можно получить непосредственно по схеме п. 1 предельным переходом или из (9) путем сведения критерия (11) к критерию (2). Оно таково:

$$\begin{aligned} & -\frac{\partial B(t, x)}{\partial t} = \min_u \max_v \left[\frac{\partial B'(t, x)}{\partial x} A(t)x + \right. \\ & \left. + \frac{\partial B'(t, x)}{\partial x} B(t)u + \frac{\partial B'(t, x)}{\partial x} C(t)v + x' M(t)x + \right. \end{aligned}$$

$$+ u'R(t)u + v'Q(t)v], \quad (12)$$

$$B(t_1, x) = 0. \quad (13)$$

В силу сделанных предположений функция, стоящая в квадратных скобках, строго выпукла по u и строго вогнута по v . Поэтому она имеет седловую точку и для нее справедлива теорема о минимаксе, т. е. в (12) операции \min \max можно переставить. К полученному таким образом уравнению можно прийти предельным переходом от игры, рассмотренной в п. 2 применительно к системе (10) и критерию (11). Это наводит на мысль, что в исследуемом случае введение смешанных стратегий излишне.

После взятия операций \min \max в правой части выражения (12) приходим к уравнению

$$\begin{aligned} -\frac{\partial B(t, x)}{\partial t} = \frac{\partial B'(t, x)}{\partial x} A(t)x + x'M(t)x + \\ + \min_u \left\{ \frac{\partial B'(t, x)}{\partial x} B(t)u + u'R(t)u \right\} + \max_v \left\{ \frac{\partial B'(t, x)}{\partial x} C(t)v + \right. \\ \left. + v'Q(t)v \right\}. \end{aligned} \quad (14)$$

Будем искать решение в виде $B(t, x) = x'L(t)x$. Подстановка $B(t, x)$ в (13) и (14) дает следующее дифференциальное уравнение Риккати:

$$\begin{aligned} \dot{L}(t) = -L(t)A(t) - A'(t)L(t) + L(t)B(t)R^{-1}(t)B'(t)L(t) + \\ + L(t)C(t)Q^{-1}(t)C'(t)L(t) - M(t) \end{aligned}$$

с начальным условием $L(t_1) = 0$. Как и в § 7 гл. II, не трудно убедиться, что уравнение имеет единственное решение $\bar{L}(t)$, определенное на отрезке $[t_0, t_1]$.

Оптимальные законы управления участников получаются при минимизации правой части уравнения (12)

$$u(x, t) = -R^{-1}(t)B'(t)L(t)x,$$

$$v(x, t) = -Q^{-1}(t)C'(t)L(t)x.$$

Покажем, что полученные законы управления являются оптимальными для участников в том смысле, что отказ от них какого-либо участника позволяет против-

нику получить определенный выигрыш. Это равносильно доказательству того, что закон управления (u^0, v^0) является седловой точкой функционала $I(u, v)$, т. е.

$$I(u^0, v) \leq I(u^0, v^0) \leq I(u, v^0). \quad (15)$$

Поскольку $I(u^0, v^0) = B(t_0, x_0)$, то для доказательства левой части неравенства (15) достаточно показать, что

$$\begin{aligned} \left. \frac{dB(t, x)}{dt} \right|_{u^0, v} &= \frac{\partial B(t, x)}{\partial t} + \frac{\partial B'(t, x)}{\partial x} A(t, x) x + \\ &+ x' M(t) x + \frac{\partial B'(t, x)}{\partial x} B(t) u^0 + u^{0'} R(t) u^0 + \\ &+ \frac{\partial B'(t, x)}{\partial x} C(t) v + v' Q(t) v \leq 0. \end{aligned} \quad (16)$$

Но последнее неравенство — очевидное следствие уравнения (14).

Правая часть неравенства (15) доказывается аналогично. Единственность седловой точки следует из строгой выпуклости по u и строгой вогнутости по v функционала $I(u, v)$.

7. Построение управлений для неравноинформированных участников. Пусть процесс описывается уравнением $x(t+1) = f(x(t), u(t), v(t), t)$, $x(t_0) = x_0$, $t \in [t_0, t_1-1]$. (17)

На управления $u(t)$, $v(t)$ участников наложены ограничения

$$u(t) \in U(t), \quad v(t) \in V(t). \quad (18)$$

Свое поведение участники оценивают критерием качества

$$I(u, v) = \varphi(x(t_1)), \quad (19)$$

причем участник 1 (с управлением $u(t)$) стремится минимизировать критерий (19), участник 2 — добиться его максимального значения.

Будем считать, что перед очередным шагом t участник 1 знает, какое управление $v=v(y, t)$ на этот шаг сделал участник 2. Построим оптимальный закон управления для участника 1.

Допустим, что до конца процесса управления остался один шаг (т. е. $t=t_1-1$) и процесс $x(t)$ находится в

состоянии y : $x(t_1-1)=y$. По предположению, участник 1 знает значение $v(t_1-1)=v$. Выбор допустимого вектора $u \in U(t_1-1)$ приведет процесс в состояние $x(t_1)=f(y, u, v, t_1-1)$, которое доставит критерию качества (19) значение $\varphi(f(y, u, v, t_1-1))$. Оптимальный закон $u(y, v, t_1-1)$ должен, очевидно, удовлетворять условию

$$\begin{aligned} \varphi(f(y, u(y, v, t_1-1), v, t_1-1)) = \\ = \min_{u \in U(t_1-1)} \varphi(f(y, u, v, t_1-1)). \end{aligned}$$

Положим

$$B(t_1-1, y) = \min_{u \in U(t_1-1)} \varphi(f(y, u, v, t_1-1)). \quad (20)$$

Физический смысл функции (20) состоит в следующем: $B(t_1-1, y)$ равно минимальному значению критерия качества (19) в задаче (17) — (19), если за шаг до окончания процесса ($t=t_1-1$) процесс находился в состоянии y : $x(t_1-1)=y$.

Предположим, что известно минимальное значение $B(\tau+1, y)$ критерия (19) для процесса, находящегося в момент $t=\tau+1$ в состоянии y : $x(\tau+1)=y$. Найдем значения функции $B(\tau, y)$. Пусть в момент $t=\tau$ процесс (1) находится в состоянии y : $x(\tau)=y$. По предположению, в этот момент участник 1 знает, какое управление $v(\tau)=v$ выбрал участник 2. Для каждого своего выбора $u(\tau)=u \in U(\tau)$ участник 1 может подсчитать значение $f(y, u, v, \tau)$, а следовательно, знать, в какое состояние перейдет процесс в момент $t=\tau+1$. Указанный выбор управления $u(\tau)$ доставит критерию качества (19) значение $B(\tau+1, f(y, u, v, \tau))$. Понятно, что оптимальный закон управления $u(y, v, \tau)$ для участника 1 должен удовлетворять соотношению

$$\begin{aligned} B(\tau+1, f(y, u(y, v, \tau), v, \tau)) = \min_{u \in U(\tau)} B(\tau+1, \\ f(y, u, v, \tau)), \end{aligned} \quad (21)$$

и он доставит критерию качества (19) значение $B(\tau, y)$.

Таким образом, получилось уравнение Беллмана

$$B(\tau, y) = \min_{u \in U(\tau)} B(\tau+1, f(y, u, v, \tau)) \quad (22)$$

с граничным условием (20).

Уравнение (22) с условием (20) решается по обычным схемам гл. III. Размерность аргументов функции $B(\tau, y)$ остается прежней, но увеличивается размерность аргументов закона управления. Поэтому иногда может оказаться целесообразным запоминать лишь значения функции $B(\tau, y)$, а закон управления подсчитывать по соотношению (21).

Построим оптимальный закон управления для участника 2. Начнем опять с момента $t=t_1-1$ и предположим, что процесс (17) находится в состоянии y : $x(t_1-1)=y$. Участник 2 знает, что о его выборе $v(t_1-1)=v$ противник будет знать и будет стремиться, используя это, минимизировать критерий (19). На каждый выбор $v \in V(t_1-1)$ участник 1 ответит выбором $u(t_1-1)=u(y, v, t_1-1)$. Это приведет процесс (17) в состояние $f(y, u(y, v, t_1-1), v, t_1-1)$ и доставит критерию качества (19) значение

$$\begin{aligned} \varphi(f(y, u(y, v, t_1-1), v, t_1-1)) = \\ = \min_{u \in U(t_1-1)} \varphi(f(y, u, v, t_1-1)), \end{aligned}$$

зависящее от выбора $v \in V(t_1-1)$. Ясно, что оптимальный закон управления $v(y, t_1-1)$ должен удовлетворять соотношению

$$\begin{aligned} \varphi(f(y, u(y, v(y, t_1-1), t_1-1), v(y, t_1-1), t_1-1)) = \\ = \max_{v \in V(t_1-1)} \varphi(f(y, u(y, v, t_1-1), v, t_1-1)) = \\ = \max_{v \in V(t_1-1)} \min_{u \in U(t_1-1)} \varphi(f(y, u, v, t_1-1)). \quad (23) \end{aligned}$$

Нетрудно понять, что величина (23) равна значению критерия качества при условии, что процесс находится в момент $t=t_1-1$ в состоянии y и участники ведут себя наилучшим образом. Пусть

$$B^0(t_1-1, y) = \max_{v \in V(t_1-1)} \min_{u \in U(t_1-1)} \varphi(f(y, u, v, t_1-1)). \quad (24)$$

Предположим, что функция $B^0(\tau+1, y)$ уже построена. Вычислим функцию $B^0(\tau, y)$. Итак, процесс в момент $t=\tau$ находится в состоянии $x(\tau)=y$, и перед участником 2 стоит задача — выбрать в этот момент оптимальное управление, причем он знает, что его выбор

будет известен противнику. Пусть сделан некоторый выбор $v(\tau) = v$, противник ответит управлением $u(\tau) = u$, и процесс в момент $t = \tau + 1$ перейдет в состояние $f(y, u, v, \tau)$, на котором критерий качества (19) при дальнейшем оптимальном поведении участников имеет значение $B^0(\tau + 1, f(y, u, v, \tau))$. Ясно, что оптимальный закон управления $u(y, v, \tau)$ участника 1 должен удовлетворять соотношению

$$B^0(\tau + 1, f(y, u(y, v, \tau), v, \tau)) = \min_{u \in U(\tau)} B^0(\tau + 1, f(y, u, v, \tau)), \quad (25)$$

а оптимальный закон управления $v(y, \tau)$ участника 2 — равенству

$$\begin{aligned} B^0(\tau + 1, f(y, u(y, v(y, \tau), \tau), v(y, \tau), \tau)) = \\ = \max_{v \in V(\tau)} B^0(\tau + 1, f(y, u(y, v, \tau), v, \tau)) = \\ = \max_{v \in V(\tau)} \min_{u \in U(\tau)} B^0(\tau + 1, f(y, u, v, \tau)). \end{aligned} \quad (26)$$

Поскольку построенные законы управления дают оптимальное значение $B^0(\tau, y)$ критерию качества (19), то приходим к следующему уравнению Беллмана:

$$B^0(\tau, y) = \max_{v \in V(\tau)} \min_{u \in U(\tau)} B^0(\tau + 1, f(y, u, v, \tau)) \quad (27)$$

с граничным условием (24).

Решение уравнения (27) при условии (24) мало отличается от рассмотренных ранее случаев. Закон управления участника 1 находится по соотношению (25), которое не отличается от (21). Для участника 2 при вычислении оптимального закона следует пользоваться соотношением (26).

Рассмотрим, наконец, последний случай. Построим оптимальный закон управления для участника 2, который знает, что его очередной выбор управления будет известен противнику, но допускает, что противник мог сделать не наилучший выбор.

Пусть процесс находится в состоянии y и осталось сделать один шаг, т. е. $x(t_1 - 1) = y$. Участник 2 делает выбор $v(t_1 - 1) = v$, но, не зная о $u(t_1 - 1) = u$, не может оценить свое поведение, ибо значение $f(y, u, v, t_1 - 1)$, а значит, и $\varphi(f(y, u, v, t_1 - 1))$ он не может подсчитать. Если у участника 2 нет никакой информации о $u(t_1 - 1)$

и он не в состоянии по предыдущим ситуациям сделать прогноз о $u(t_1-1)$, то для него разумнее всего сделать такой выбор $v(t_1-1)$, который гарантирует ему выигрыш. Ясно, что гарантирующее управление $v(t_1-1)$ совпадает с управлением, подсчитанным из (24). Аналогичные рассуждения показывают, что гарантирующие управления участника 2 совпадают с оптимальными в задаче, когда оба участника действуют оптимально. Нетрудно понять, что при таком выборе закона управления участник 2 не только гарантирует себе определенный выигрыш, который при оптимальном поведении участника 1 не может быть большим, но и в состоянии использовать ошибки участника 1, если тот уклонится от оптимального поведения. Действительно, пусть в состоянии $x(\tau)=y$ участник 1 выбрал управление $u(\tau)=u$, которое отличается от оптимального $u^0=u(y, v^0, \tau)$, а участник 2 выбрал управление $v^0(\tau)=v^0$ в расчете, что участник 1 сделал оптимальный выбор. Тогда процесс в момент $t = \tau + 1$ перейдет в состояние $f(y, u, v^0, \tau)$, для которого

$$B^0(\tau + 1, f(y, u, v^0, \tau)) > B^0(\tau + 1, f(y, u^0, v^0, \tau)). \quad (28)$$

При оптимальном поведении участников после $\tau + 1$ критерий качества (19) принимает значение $B^0(\tau + 1, f(y, u, v^0, \tau))$, если $u \neq u^0$, и значение $B^0(\tau + 1, f(y, u^0, v^0, \tau))$ при $u = u^0$. Неравенство (28) означает, что при указанном выше законе управления участнику 2 удастся использовать ошибки участника 1.

З а м е ч а н и е. Участник 1 использует все ошибки участника 2. Это видно из закона управления $u(y, v, \tau)$. Поскольку отказ от построенных законов $u(y, v, \tau)$, $v(y, \tau)$ не выгоден обоим участникам, то говорят, что эти законы образуют седловую точку для критерия качества (19).

Сохраним постановку задачи, но изменим условия информированности: будем считать, что в каждый момент $t = \tau$ участник 2 при выборе управления $v = v(\tau)$ знает, какой выбор сделал противник. Возможны три случая: а) участник 2 строит закон управления для каждого поведения участника 1; б) участники строят оптимальные законы управления; в) участник 1 строит гарантирующие управления. Каждый из этих случаев исследуется по схеме, описанной выше. Следует только в рассуждениях поменять номера участников, а также операции \min на \max в соответствии с целями участников.

8. Задача преследования. Рассмотрим процесс

$$x(t+1) = f(x(t), u(t), v(t)), \quad x(0) = x_0, \quad t \geq 0, \quad (29)$$

управляемый двумя участниками посредством управлений $u(t)$, $v(t)$, стесненных ограничениями

$$u(t) \in U, \quad v(t) \in V. \quad (30)$$

Пусть в пространстве векторов x задано множество M . Цель участника 1 — перевести за минимальное время процесс на M . Участник 2 препятствует этому и стремится максимально затянуть время встречи процесса с множеством M .

Будем считать сначала, что участник 1 при выборе управления на очередном шаге t знает, какой выбор $v = v(y)$ сделан участником 2. Кроме того, будем считать, что для любого вектора x_0 и любого допустимого управления $v(t)$ можно найти управление $u(t)$, переводящее процесс на множество M .

Нетрудно проверить, что сформулированная задача содержит в себе задачу преследования, в которой задаются два процесса $y(t)$, $z(t)$, описываемые уравнениями

$$y(t+1) = f^1(y(t), u(t)), \quad y(0) = y_0;$$

$$z(t+1) = f^2(z(t), v(t)), \quad z(0) = z_0,$$

и управления $u(t)$, $v(t)$ из (30) выбираются так, чтобы встреча $y(t_1) = z(t_1)$ произошла для участника 1 как можно скорее, а для участника 2 по возможности позднее.

Вводя новый вектор $x = \{y, z\}$, нетрудно задачу преследования свести к задаче, поставленной в начале пункта. Задачу будем решать по следующей схеме. Пусть $F^-(v)$, $v \in V$, — множество векторов y таких, что при некоторых $u \in U$ выполняется соотношение

$$f(y, u, v) \in M. \quad (31)$$

Ясно, что из каждой точки множества $F^-(v)$ участник 1 может за один шаг перевести процесс на множество M . Поэтому время быстрогодействия $B(y)$ удовлетворяет тождеству

$$B(y) \equiv 1, \quad y \in F^-(v), \quad (32)$$

и оптимальный закон $u(y, v)$ находится из соотношения (31).

Пусть процесс находится в точке $y : x(0) = y$, $y \in F(v)$. По известному вектору $v \in V$ участник 1 выбирает вектор $u \in U$ и вычисляет вектор $f(y, u, v)$, являющийся состоянием процесса в следующий момент. Минимальное время попадания процесса из точки y на множество M обозначим $B(y)$. Тогда $B(f(y, u, v))$ — минимальное время попадания из точки $f(y, u, v)$ на множество M . Меняя $u \in U$, можно добиться, чтобы это время было минимальным, т. е. найти $u(y, v)$ так, чтобы

$$B(f(y, u(y, v), v)) = \min_{u \in U} B(f(y, u, v)). \quad (33)$$

Тогда минимальное время движения $B(y)$ из y до M удовлетворяет соотношению

$$B(y) = 1 + \min_{u \in U} B(f(y, u, v)),$$

которое является уравнением Беллмана для рассматриваемой задачи.

Процедура решения уравнения (33) при граничном условии (32) такова. Для разных $v \in V$ строятся множества $F^-(v)$. На этих множествах функция $B(y)$ равна единице. Для каждой пары $v^1, v^2 \in V$ строится множество $F^-(v^1, v^2)$ векторов y так, чтобы $f(y, u, v^2) \in F^-(v^1)$ при некотором $u \in U$. На множествах $F^-(v^1, v^2)$ полагаем $B(y) = 2$. Допустим, что при фиксированных $v^1, \dots, v^k \in V$ множество $F^-(v^1, \dots, v^k)$ построено и $B(y) = k$, если $y \in F^-(v^1, \dots, v^k)$. Множество $F^-(v^1, \dots, v^{k+1})$ строится из векторов y таких, что $f(y, u, v^{k+1}) \in F^-(v^1, \dots, v^k)$ при некотором $u \in U$. Полагаем $B(y) = k+1$, если $y \in F^-(v^1, \dots, v^{k+1})$.

В силу сделанных предположений найдется число L , $L < \infty$, такое, что при любых $v^1, \dots, v^L \in V$ множества $F^-(v^1, \dots, v^L)$ содержат точку x_0 . Минимальное L^0 , при котором это свойство выполняется, и есть, очевидно, оценка сверху для времени попадания процесса из точки y на множество M .

Процесс управления для участника 1 протекает следующим образом. По заданному вектору x_0 и известному значению $v(0)$ для некоторого $u \in U$ вычисляется век-

тор $f(x_0, u, v(0))$. Затем отыскивается $u(x_0, v(0))$, при котором минимально число $L(0)$, удовлетворяющее соотношению

$$f(x_0, u(x_0, v(0)), v(0)) \in F^-(v^{L(0)}, \dots, v^1, v(0)), \quad (34)$$

$$v^1, \dots, v^{L(0)} \in V.$$

Если вектор $u(x_0, v(0))$ найден, то в силу (29) вычисляем вектор $x^0(1) = f(x_0, u(x_0, v(0)), v(0))$ и определяем вектор $u(x^0(1), v(1))$ при известном значении $v(1)$. Для этого повторяем процедуру, проделанную при отыскании вектора $u(x_0, v(0))$. Через конечное число шагов процесс попадет на множество M .

Оптимальный закон управления для участника 2 определяется из соотношения (34). Поскольку число $L(0)$ зависит от $v(0)$, то выбирается такое $v(x_0)$, при котором $L(0)$ максимально. Остальные построения очевидны.

ГЛАВА V ● СВЯЗЬ ДИНАМИЧЕСКОГО ПРОГРАММИРОВАНИЯ С ДРУГИМИ МЕТОДАМИ

Наряду с динамическим программированием существует ряд других методов решения задач оптимизации, в той или иной степени связанных с идеями динамического программирования. Наиболее близким к достаточным условиям динамического программирования является метод Кротова [13], который как частный случай включает достаточные условия динамического программирования. Метод Кротова позволяет исследовать и особые управления, для которых уравнения Беллмана неэффективны. В связи с изучением особых управлений авторами были предложены дифференциальные уравнения высокого порядка для функций качества. При исследовании оптимальных процессов эффективной является и функция Розоноэра [19], введенная при исследовании связи между динамическим программированием и принципом максимума Понтрягина. Исходя из функции Розоноэра, можно построить метод решения задач оптимизации, аналогичный динамическому программированию, который можно считать локальным вариантом классического.

§ 1. Принцип максимума Понтрягина

В современных работах динамическое программирование используется в основном для формулировки достаточных условий оптимальности. Для формулировки необходимых условий [11, 17] привлекается, как правило, принцип максимума Понтрягина — другой мощный метод исследования задач оптимизации. Принцип максимума к настоящему времени доказан для весьма широкого класса задач без тех стеснительных предположений, при которых получались в гл. II необходимые условия оптимальности из динамического программирования. В свя-

зи с развитием обоих методов возникает вопрос о взаимоотношениях между ними. Этот вопрос был решен Л. И. Розоноэром [19]. Оказалось, что между двумя методами существует тесная связь, аналогичная той, которая была известна в классическом вариационном исчислении [12] между уравнениями Эйлера—Гамильтона и Гамильтона — Якоби. Принцип максимума явился существенным обобщением условий стационарности (уравнения Эйлера — Гамильтона и Вейерштрасса), а уравнения Беллмана — уравнений Гамильтона — Якоби и достаточных условий оптимальности теории поля.

1. Получение принципа максимума из уравнения Беллмана. Вернемся к задаче терминального управления в дискретных системах

$$x(t+1) = f(x(t), u(t), t), \quad x(t_0) = x_0, \quad t \in [t_0, t_1 - 1], \quad (1)$$

$$u(t) \in U(t), \quad I(u) = \varphi(x(t_1)) \rightarrow \min.$$

Для нее уравнение Беллмана

$$B(t, x) = \min_{u \in U(t)} B(t+1, f(x, u, t)), \quad (2)$$

$$B(t_1, x) = \varphi(x), \quad (3)$$

доставляет необходимое и достаточное условия.

Из уравнения (2) можно получить необходимые условия оптимальности, которые лишены некоторых недостатков, присущих динамическому программированию.

Предположим, что функция $B(t, x)$ дифференцируема по аргументам, а множество

$$f(x, U(t), t) = \{y : y = f(x, u, t), u \in U(t)\} \quad (4)$$

выпукло при каждом x и $t \in [t_0, t_1 - 1]$. Пусть $u^0(t), x^0(t)$ — оптимальные управление и траектория. Согласно (2), имеем

$$B(t, x^0(t)) = B(t+1, f(x^0(t), u^0(t), t)) \leq B(t+1, f(x^0(t), u, t)) \quad (5)$$

для любого вектора $u \in U(t)$. Построим вектор u_ε так, чтобы

$$f(x^0(t), u_\varepsilon, t) = (1 - \varepsilon) f(x^0(t), u^0(t), t) + \varepsilon f(x^0(t), u, t). \quad (6)$$

Геометрически точка $f(x^0(t), u_\varepsilon, t)$ лежит на отрезке,

соединяющем точки $f(x^0(t), u^0(t), t)$ и $f(x^0(t), u, t)$ (рис. 65). В силу выпуклости множества (4) точка u_ε при любых ε , $0 \leq \varepsilon \leq 1$, может быть выбрана из множества $U(t)$. Подставив вектор u_ε в (5), получим

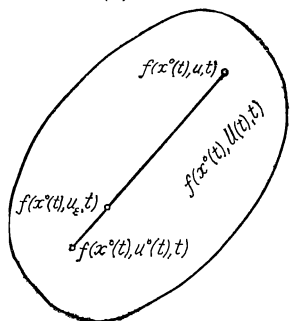


Рис. 65

$$\begin{aligned}
 0 \leq B(t+1, f(x^0(t), u_\varepsilon, t)) - \\
 - B(t+1, f(x^0(t), u^0(t), t)) = \\
 = \frac{\partial B'(t+1, f(x^0(t), u^0(t), t))}{\partial x} \times \\
 \times (f(x^0(t), u_\varepsilon, t) - f(x^0(t), \\
 u^0(t), t)) + o(\|f(x^0(t), u_\varepsilon, t) - \\
 - f(x^0(t), u^0(t), t)\|). \quad (7)
 \end{aligned}$$

Поскольку из (6) ясно, что

$$\begin{aligned}
 f(x^0(t), u_\varepsilon, t) - f(x^0(t), u^0(t), t) = \\
 = \varepsilon (f(x^0(t), u, t) - f(x^0(t), u^0(t), t)),
 \end{aligned}$$

то из (7) следует

$$\begin{aligned}
 \frac{\partial B'(t+1, x^0(t+1))}{\partial x} f(x^0(t), u^0(t), t) = \\
 = \min_{u \in U(t)} \frac{\partial B'(t+1, x^0(t+1))}{\partial x} f(x^0(t), u, t). \quad (8)
 \end{aligned}$$

Введем функции

$$\psi^0(t) = - \frac{\partial B(t+1, x^0(t+1))}{\partial x}, \quad H(x, \psi, u, t) = \psi' f(x, u, t), \quad (9)$$

которые назовем импульсом и гамильтонианом системы (1).

В новых обозначениях соотношение (8), которому удовлетворяет оптимальное управление, принимает вид

$$H(x^0(t), \psi^0(t), u^0(t), t) = \max_{u \in U(t)} H(x^0(t), \psi^0(t), u, t).$$

Найдем уравнение для $\psi^0(t)$. Из (3) имеем

$$\psi^0(t_1 - 1) = - \frac{\partial \varphi(x^0(t_1))}{\partial x}.$$

Далее, если к произвольной точке $x \neq x^0(t)$ приложить в момент t управление $u^0(t)$, то

$$B(t, x) \leq B(t+1, f(x, u^0(t), t)), \quad (10)$$

$$B(t, x^0(t)) = B(t+1, f(x^0(t), u^0(t), t)), \quad (11)$$

т. е. функция $C(x, t) = B(t, x) - B(t+1, f(x, u^0(t), t))$ в точке $x = x^0(t)$ достигает максимума по x . Поэтому

$$\begin{aligned} \frac{\partial C'(x, t)}{\partial x} \Big|_{x=x^0(t)} &= \frac{\partial B'(t, x^0(t))}{\partial x} - \\ - \frac{\partial B'(t+1, x^0(t+1))}{\partial x} - \frac{\partial f(x^0(t), u^0(t), t)}{\partial x} &= 0. \end{aligned}$$

Это равенство в обозначениях (9) принимает вид

$$\psi^0(t-1) = \left[\frac{\partial f(x^0(t), u^0(t), t)}{\partial x} \right]' \psi^0(t) = \frac{\partial H(x^0(t), \psi^0(t), u^0(t), t)}{\partial x}.$$

Таким образом, при сделанных предположениях справедлива

Теорема. Вдоль оптимального управления $u^0(t)$ выполняется принцип максимума

$$H(x^0(t), \psi^0(t), u^0(t), t) = \max_{u \in U(t)} H(x^0(t), \psi^0(t), u, t),$$

где $x^0(t)$ — оптимальная траектория, а $\psi^0(t)$ — решение уравнения

$$\dot{\psi}^0(t-1) = \frac{\partial H(x^0(t), \psi^0(t), u^0(t), t)}{\partial x},$$

$$\dot{\psi}^0(t_1-1) = - \frac{\partial \varphi(x^0(t_1))}{\partial x}.$$

Полученные условия оптимальности называются необходимыми условиями первого порядка, так как в них используются лишь первые производные $\frac{\partial B}{\partial x}$, $\frac{\partial f}{\partial x}$ вдоль оптимальных управлений и траекторий.

2. Связь с динамическим программированием. Теорема п. 1 представляет собой дискретный аналог принципа максимума Понтрягина, а первое соотношение из (9) выражает связь между динамическим программированием

и принципом максимума. Геометрически связь между двумя результатами можно выразить следующим образом: направление $\psi^0(t)$, вдоль которого максимизируется гамильтониан системы, представляет антиградиент функции Беллмана в точке оптимальной траектории (рис. 66).

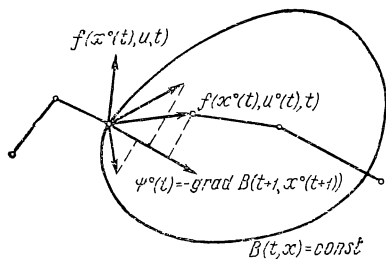


Рис. 66

Принцип максимума в п. 1 получен в предположении, что $B(t, x)$ — дифференцируемая функция. Хотя это предположение во многих задачах, как правило, не выполняется, принцип максимума остается справедливым.

3. Случай непрерывных процессов. Рассмотрим задачу

$$\dot{x} = f(x, u, t), \quad x(t_0) = x_0, \quad t \in [t_0, t_1],$$

$$u(t) \in U, \quad I(u) = \varphi(x(t_1)) \rightarrow \min.$$

Дифференциальное уравнение Беллмана здесь имеет вид

$$-\frac{\partial B(t, x)}{\partial t} = \min_{u \in U} \frac{\partial B'(t, x)}{\partial x} f(x, u, t), \quad (12)$$

$$B(t_1, x) = \varphi(x).$$

Оптимальный закон управления $u(x, t)$ удовлетворяет соотношению

$$\frac{\partial B'(t, x)}{\partial x} f(x, u(x, t), t) = \min_{u \in U} \frac{\partial B'(t, x)}{\partial x} f(x, u, t). \quad (13)$$

Если ввести обозначения

$$\psi^0(t) = -\frac{\partial B(x^0(t), t)}{\partial x}, \quad H(x, \psi, u, t) = \psi' f(x, u, t),$$

то свойство (12) можно записать в новой форме

$$H(x^0(t), \psi^0(t), u^0(t), t) = \max_{u \in U} H(x^0(t), \psi^0(t), u, t). \quad (14)$$

Получим уравнение для функции $\psi^0(t)$. Для этого предположим, что функция $B(t, x)$ дважды дифференцируема по x .

Во-первых, из (12) ясно, что

$$\psi^0(t_1) = - \frac{\partial B(t_1, x^0(t_1))}{\partial x}. \quad (15)$$

Во-вторых, из рассуждений, которые привели к соотношениям (10), (11), следует, что функция

$$C(x, t) = \frac{\partial B'(t, x)}{\partial x} f(x, u^0(t), t) + \frac{\partial B(t, x)}{\partial t}$$

в точке $x=x^0(t)$ достигает минимума. Следовательно,

$$0 = \left. \frac{\partial C(x, t)}{\partial x} \right|_{x=x^0(t)} = \frac{\partial^2 B(t, x^0(t))}{\partial x^2} f(x^0(t), u^0(t), t) + \\ + \frac{\partial f'(x^0(t), u^0(t), t)}{\partial x} \frac{\partial B(t, x^0(t))}{\partial x} + \frac{\partial^2 B(t, x^0(t))}{\partial x \partial t}. \quad (16)$$

Но

$$\dot{\psi}^0(t) = - \frac{d}{dt} \frac{\partial B(t, x^0(t))}{\partial x} = - \frac{\partial^2 B(t, x^0(t))}{\partial x^2} f(x^0(t), u^0(t), t) - \\ - \frac{\partial^2 B(t, x^0(t))}{\partial x \partial t}.$$

Поэтому из (16) получаем уравнение

$$\dot{\psi}^0(t) = - \frac{\partial f'(x^0(t), u^0(t), t)}{\partial x} \psi^0(t) = \\ = - \frac{\partial H(x^0(t), \psi^0(t), u^0(t), t)}{\partial x}. \quad (17)$$

Свойство (14), выполняющееся вдоль $u^0(t)$, $x^0(t)$ и $\psi^0(t)$, удовлетворяющих уравнениям процесса и (15), (17), и есть принцип максимума Понтрягина. Приведенный вывод не есть доказательство принципа максимума, который справедлив и без предположения о гладкости функции Беллмана. Для нас главным было установление связи

$$\psi^0(t) = - \frac{\partial B(t, x^0(t))}{\partial x}. \quad (18)$$

Соотношение (18) имеет такой же смысл, как аналогичное соотношение в дискретном случае (см. п. 2).

§ 2. Принцип оптимальности Кротова

Этот принцип сначала был сформулирован для непрерывных систем, затем перенесен на дискретные системы [13].

1. Непрерывные процессы. Рассмотрим процесс управления, описываемый уравнением

$$\dot{x} = f(x, u, t), \quad t \in [t_0, t_1], \quad (1)$$

где x — n -вектор; u — r -вектор.

Пусть на течение процесса $x(t)$ и поведение управления $u(t)$ наложено ограничение

$$\{x(t), u(t)\} \in G(t), \quad t \in [t_0, t_1], \quad (2)$$

где $G(t)$ — множество $n+r$ -мерного пространства.

Совокупность пар $\{x(t), u(t)\}$, состоящих из абсолютно-непрерывных функций $x(t)$ и кусочно-непрерывных $u(t)$, удовлетворяющих уравнению (1) и ограничению (2), обозначим через D . На множестве D требуется найти последовательность $\{x^s(t), u^s(t)\}$, вдоль которой функционал

$$I(x, u) = \int_{t_0}^{t_1} f_0(x, u, t) dt + g(x(t_0), x(t_1)) \quad (3)$$

стремится к нижней грани

$$I(x^s, u^s) \rightarrow \inf_D I(x, u). \quad (4)$$

Сформулированная задача охватывает большую часть задач гл. II. Она имеет смысл как в случае, когда оптимальное управление существует, так и тогда, когда его в заданном классе нет.

Пусть $K(x, t)$ — функция, определенная на множестве значений x, t и непрерывная вместе с производными по аргументам. Введем функции:

$$\begin{aligned} \Phi(x, u, t) = & \frac{\partial K'(x, t)}{\partial x} f(x, u, t) - f_0(x, u, t) + \\ & + \frac{\partial K(x, t)}{\partial t}, \end{aligned} \quad (5)$$

$$Q(x_0, x_1) = g(x_0, x_1) + K(x_1, t_1) - K(x_0, t_0), \quad (6)$$

$$\mu(t) = \sup_{\{x, u\} \in G(t)} \Phi(x, u, t), \quad \nu = \inf_{x_0 \in G_x(t_0), x_1 \in G_x(t_1)} Q(x_0, x_1).$$

Теорема (принцип оптимальности Кротова). Для того чтобы последовательность $\{x^s(t), u^s(t)\} \in D$ была решением задачи (1)–(4), достаточно существования такой функции $K(x, t)$, что: 1) $\Phi(x^s(t), u^s(t), t) \rightarrow \mu(t)$, $t \in (t_0, t_1)$ (по мере); 2) $Q(x^s(t_0), x^s(t_1)) \rightarrow \nu$; 3) $\mu(t)$ — кусочно-непрерывная функция; 4) $\Phi(x^s(t), u^s(t), t) > Q > -\infty$, $s < \infty$, $t \in (t_0, t_1)$.

Здесь G_x — проекция множества G на пространство векторов x .

Доказательство. Погрузим множество D во множество E пар $\{x(t), u(t)\}$ из кусочно-непрерывных функций $x(t), u(t)$, удовлетворяющих ограничению (2). На множестве E определим функционал

$$L(x, u) = Q(x(t_0), x(t_1)) - \int_{t_0}^{t_1} \Phi(x(t), u(t), t) dt. \quad (7)$$

Если $\{x(t), u(t)\} \in D$, то в силу (5)

$$\Phi(x, u, t) = \frac{dK(x, t)}{dt} \Big|_{(1)} - f_0(x, u, t).$$

Поэтому из (3), (5) и (7) получаем

$$L(x, u) = I(x, u), \text{ если } \{x, u\} \in D. \quad (8)$$

Пусть $l = \inf L(x, u)$, $\{x, u\} \in E$. Из (6) ясно, что

$$\begin{aligned} l &= \inf_{x_0 \in G_x(t_0), x_1 \in G_x(t_1)} Q(x_0, x_1) - \sup_E \int_{t_0}^{t_1} \Phi(x(t), u(t), t) dt = \\ &= \nu - \int_{t_0}^{t_1} \mu(t) dt. \end{aligned} \quad (9)$$

Покажем, что вдоль $\{x^s(t), u^s(t)\} \in D$ выполняется соотношение $L(x^s, u^s) \rightarrow l$, $s \rightarrow \infty$. Имеем

$$\begin{aligned} \lim_{s \rightarrow \infty} L(x^s, u^s) &= \lim_{s \rightarrow \infty} Q(x^s(t_0), x^s(t_1)) - \\ &- \lim_{s \rightarrow \infty} \int_{t_0}^{t_1} \Phi(x^s(t), u^s(t), t) dt. \end{aligned}$$

Меняя операции предельного перехода и интегрирования (теорема Лебега), с учетом условий 1), 2) и (9) получим

$$\lim_{s \rightarrow \infty} L(x^s, u^s) = \nu - \int_{t_0}^{t_1} \mu(t) dt = l.$$

Осталось проверить свойство (4). В силу (8) вдоль последовательности $\{x^s, u^s\}$ выполняется равенство $L(x^s, u^s) = I(x^s, u^s)$. Таким образом,

$$I(x^s, u^s) \rightarrow l, s \rightarrow \infty. \quad (10)$$

По определению l и множества E , выполняется неравенство

$$l \leq \inf_D I(x, u). \quad (11)$$

Сравнивая (10) с (11), получаем соотношение

$$I(x^s, u^s) \rightarrow \inf_D I(x, u),$$

которое доказывает теорему.

2. Связь с динамическим программированием. Пусть ограничение (2) имеет вид

$$u(t) \in U, t \in [t_0, t_1], x(t_0) = x_0, \quad (12)$$

и $g(x(t_0), x(t_1)) = g(x(t_1))$. Введем функцию $P(x, t) = \sup \Phi(x, u, t), u \in U$. Допустим, что удалось найти такую функцию $K(x, t)$, что

$$P(x, t) \equiv c(t), t \in [t_0, t_1], Q(x_0, x_1) = c_1(x_0), \quad (13)$$

т. е. $P(x, t)$ не зависит от x , а $Q(x_0, x_1)$ от x_1 . В этом случае $\mu(t) = \sup_x P(x, t) = c(t)$. Предположим, что задача

оптимизации из п. 1 имеет решение $\{x^0(t), u^0(t)\} \in D$. Тогда условия теоремы п. 1 принимают форму:

$$1) \Phi(x^0(t), u^0(t), t) = \mu(t); \quad (14)$$

$$2) Q(x^0(t_0), x^0(t_1)) = \nu. \quad (15)$$

Остальные условия выполняются автоматически.

В подробной записи условия (14), (15) для частного случая (12), (13) имеют вид

$$c(t) = \frac{\partial K(x^0(t), t)}{\partial t} + \frac{\partial K'(x^0(t), t)}{\partial x} f(x^0(t), u^0(t), t) - \\ - f_0(x^0(t), u^0(t), t) = \sup_{u \in U} \left[\frac{\partial K(x^0(t), t)}{\partial t} + \right. \\ \left. + \frac{\partial K'(x^0(t), t)}{\partial x} f(x^0(t), u, t) - f_0(x^0(t), u, t) \right], \quad (16) \\ K(x, t_1) = -g(x).$$

Сравнивая полученные уравнения с уравнением Беллмана (§ 4 гл. II), нетрудно заметить, что при $K(x, t) = -B(t, x)$, $c(t) = 0$, уравнение (16) переходит в уравнение Беллмана для задачи типа Больца.

Таким образом, уравнение Беллмана получается из теоремы п. 1 при специальном задании функции $K(x, t)$.

3. «Другой формализм» Кротова. При выборе функции $K(x, t)$ так, чтобы выполнялось условие (13), метод Кротова близок к методу динамического программирования. Требование (13) означает, что из общего уравнения

$$\sup_x \sup_u \Phi(x, u, t) = \Phi(x^0(t), u^0(t), t) = \mu(t), \quad (17)$$

следующего из условия 1) теоремы п. 1, искусственно исключена одна операция \sup_x , именно та, из-за которой теорема п. 1 отличается от теоремы, доставляющей достаточные условия оптимальности динамического программирования.

Имеется и вторая возможность — исключить из (17) операцию \sup_u . Иными словами, можно попытаться так построить функцию $K(x, t)$, чтобы функция

$$\sup_x \Phi(x, u, t) \quad (18)$$

не зависела от u . Подобная свобода выбора позволяет иногда обойти трудности, с которыми не удастся справиться в рамках динамического программирования. Типичную ситуацию такого рода доставляют особые управления и скользящие режимы. Не останавливаясь подробно на деталях «другого формализма» Кротова, приведем иллюстративные примеры к теореме п. 1.

Пример 1. $\dot{x} = u$, $|u| \leq 1$, $x(0) = x(1) = 0$, $I(x, u) = \int_0^1 (x^2 - u) dt$, где x, u — скаляры.

Функция $\Phi(x, u, t)$ в данном случае имеет вид

$$\Phi = \frac{\partial K}{\partial x} u - x^2 + u + \frac{\partial K}{\partial t}. \quad (19)$$

Функцию Q составлять не нужно, так как рассматривается задача с закрепленными концами.

Если выбрать функцию $K(x, t)$ так, чтобы

$$\frac{\partial K(x, t)}{\partial x} = -1, \quad (20)$$

то функция Φ не будет зависеть от u . Из (20) получаем

$$K(x, t) = -x + c(t), \quad (21)$$

где $c(t)$ — произвольная гладкая функция.

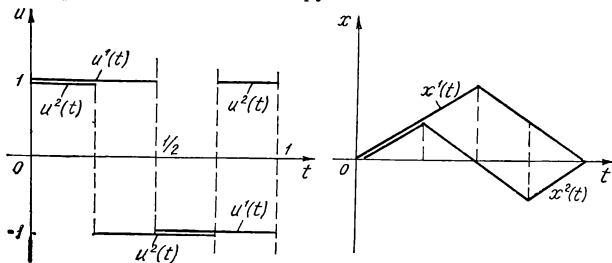


Рис. 67

В случае (21) выражение (19) принимает вид

$$\Phi = -x^2 + \frac{dc}{dt}.$$

Максимум этой функции по x достигается в точке $x^* = 0$ при любом t . Функцию $x^*(t) = 0$ проверяем на принадлежность множеству D .

Для этого подставляем ее в уравнение движения $\dot{x} = u$, откуда $\dot{x}^*(t) = u^*(t) = 0$. Пара $\{x^*(t), u^*(t)\}$ удовлетворяет всем условиям теоремы п. 1 и потому является решением задачи. Управление $u(t) = 0$ называется с о б ы м.

Пример 2. $\dot{x} = u$, $|u| \leq 1$, $x(0) = x(1) = 0$, $I(x, u) = \int_0^1 (x^2 - u^2) dt$. В этом случае

$$\Phi = \frac{\partial K}{\partial x} u - x^2 + u^2 + \frac{\partial K}{\partial t}.$$

Возьмем функцию $K(x, t) \equiv 0$. Для нее $\Phi = -x^2 + u^2$. Максимум этой функции достигается при $x^* = 0$, $u^* = \pm 1$. Поэтому $x^*(t) = 0$, $u^*(t) = \text{релейная функция } |u^*(t)| = 1$.

Но пара $\{x^*(t), u^*(t)\}$ не является элементом множества D , ибо $\dot{x}^*(t) = 0 \neq u^*(t) = \pm 1$. Однако можно построить последовательность пар $\{x^s(t), u^s(t)\}$ из $u^s(t) = \text{sign} \cos \pi s t$, $s = 1, 2, \dots$ и соответствующих траекторий $x^s(t)$ (рис. 67) системы $\dot{x} = u$, вдоль которых $\Phi(x^s(t), u^s(t), t) \rightarrow \mu(t) = 1$. Все условия теоремы п. 1 выполнены, построенная последовательность решает задачу. Совокупность таких управлений и соответствующих им траекторий называется оптимальным скользющим режимом.

§ 3. Об одном аналоге функции Беллмана

Как известно (см. гл. II), функция Беллмана определяется с помощью оптимальных управлений, построенных для задач оптимизации семейства процессов, в которое погружается исходная задача. При таком определении аналитические свойства функции Беллмана становятся «хуже» аналитических свойств функций, участвующих в формулировке задачи оптимизации. Другой тип функций был введен Л. И. Розоноэром [19]. Функции опять определялись вдоль решений семейства задач, но их аналитические свойства были такими же, как у исходных функций. В отличие от функций Беллмана использование новых функций приводит лишь к локальным условиям оптимальности.

1. Функция Розоноэра для дискретных процессов. Рассмотрим задачу терминального управления для дискретных систем

$$\left. \begin{aligned} x(t+1) &= f(x(t), u(t), t), \quad x(t_0) = x_0, \quad t \in [t_0, t_1-1], \\ u(t) &\in U(t), \quad I(u) = \varphi(x(t_1)) \rightarrow \min. \end{aligned} \right\} \quad (1)$$

Пусть $u^0(t)$ — оптимальное управление в задаче (1); $x^0(t)$ — соответствующая ему траектория. Введем функцию $R(t, x)$, $t \in [t_0, t_1]$. По определению,

$$R(\tau, y) \equiv \varphi(x(y, \tau, t_1)), \quad (2)$$

где $x(y, \tau, t)$ — решение уравнения

$$x(t+1) = g(x(t), t) \equiv f(x(t), u^0(t), t), \quad (3)$$

отвечающее начальному состоянию (y, τ) .

Из соотношений (2) — (3) следует, что функция $R(t, x)$ удовлетворяет рекуррентному уравнению

$$R(t, x) = R(t+1, f(x, u^0(t), t)) \quad (4)$$

с граничным условием

$$R(t_1, x) = \varphi(x). \quad (5)$$

Кроме того, если управление $u^0(t)$ оптимально, то

$$R(t, x^0(t)) = \min_{u \in U(t)} R(t+1, f(x^0(t), u, t)) \quad (6)$$

для всех $t \in [t_0, t_1-1]$.

Таким образом, условие (6) является необходимым условием оптимальности управления $u^0(t)$. Для того чтобы воспользоваться этим условием, из уравнения (4) — (5) нужно найти функции $R(t_1, x)$, $R(t_1-1, x)$, ..., $R(t_0, x)$.

2. Функция Розоноэра для непрерывных процессов. Рассмотрим задачу

$$\begin{aligned} \dot{x} &= f(x, u, t), \quad x(t_0) = x_0, \quad t \in [t_0, t_1], \\ u(t) &\in U, \quad I(u) = \varphi(x(t_1)) \rightarrow \min. \end{aligned}$$

Пусть $u^0(t)$ — оптимальное управление в этой задаче, а $x^0(t)$ — соответствующая ему оптимальная траектория. По аналогии с дискретным случаем введем функцию $R(t, x)$, $t \in [t_0, t_1]$:

$$R(\tau, y) \equiv \varphi(x(y, \tau, t_1)),$$

где $x(y, \tau, t)$ — траектория уравнения

$$\dot{x} = g(x, t) \equiv f(x, u^0(t), t), \quad (8)$$

отвечающая начальному состоянию (y, τ) .

Используя теорему о дифференцируемости решений обыкновенного дифференциального уравнения по начальным данным, нетрудно доказать, что если функции $f(x, u, t)$ и $\varphi(x)$ непрерывно дифференцируемы по x , то функция $R(t, x)$ обладает частными производными

$$\frac{\partial R(t, x)}{\partial x}, \quad \frac{\partial R(t, x)}{\partial t}, \quad \text{причем производные первого типа}$$

непрерывны по обоим аргументам, а функция $\frac{\partial R(t, x)}{\partial t}$ непрерывна по x и кусочно-непрерывна по t .

В силу того, что функция $R(t, x)$ постоянна вдоль любой траектории системы (8), справедливо тождество

$$\frac{\partial R'(t, x)}{\partial x} f(x, u^0(t), t) + \frac{\partial R(t, x)}{\partial t} \equiv 0, \quad t \in [t_0, t_1]. \quad (9)$$

Кроме того, из определения функции $R(t, x)$ следует

$$R(t_1, x) = \varphi(x). \quad (10)$$

Если управление $u^0(t)$ оптимально, то справедливо соотношение

$$\left. \frac{dR(t, x^0(t))}{dt} \right|_{(8) u^0(t)} = \min_{u \in U} \left. \frac{dR(t, x^0(t))}{dt} \right|_{(8) u} \quad (11)$$

или

$$\frac{\partial R'(t, x^0(t))}{\partial x} f(x^0(t), u^0(t), t) = \min_{u \in U} \frac{\partial R'(t, x^0(t))}{\partial x} f(x^0(t), u, t). \quad (12)$$

Условие (11)—(12) является необходимым условием оптимальности управления $u^0(t)$. Функцию $R(t, x)$ можно найти, решив уравнение в частных производных (9) с краевым условием (10).

3. Связь с функцией Беллмана. Из определения функции $R(t, x)$ следует, что она удовлетворяет соотношениям

$$R(t, x) \geq B(t, x), \quad R(t, x^0(t)) = B(t, x^0(t)), \quad (13)$$

где $B(t, x)$ — функция Беллмана в задаче терминального управления.

Если предположить, что функции $R(t, x)$ и $B(t, x)$ непрерывно дифференцируемы по x (функция $R(t, x)$ будет обладать этим свойством, если $f(x, u, t)$ и $\varphi(x)$ непрерывно дифференцируемы по x), то из соотношений (13) следует

$$\frac{\partial R(t, x^0(t))}{\partial x} = \frac{\partial B(t, x^0(t))}{\partial x}.$$

Функция $R(t, x)$, очевидно, зависит от начальных условий конкретной задачи. Если начальные условия не фиксировать, получим семейство функций $R(t, x, t_0, x_0)$, причем для каждой пары x_0, t_0 выполняются соотношения (13).

Условия (13) означают, что функция Беллмана является огибающей (снизу) семейства функций $R(t, x, t_0, x_0)$ Розоноэра. Взаимное расположение поверхностей $y = B(t, x)$, $y = R(t, x, t_0, x_0)$ (для непрерывного случая) по-

казано на рис. 68, где $y=R(t, x)$ — одна из поверхностей семейства функций Розоноэра с параметрами t_0, x_0 ; $y=B(t, x)$ — поверхность, соответствующая функции Беллмана. Поверхности соприкасаются по оптимальной траектории $x^0(t)$, исходящей из точки x_0 : $x^0(t_0)=x_0$.

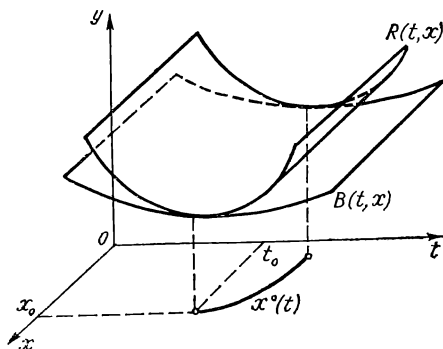


Рис. 68

Итак, функция Розоноэра $R(t, x)$ удовлетворяет уравнению, аналогичному уравнению Беллмана, но аналитически более простому. Краевые условия одинаковы. Гладкость функции можно установить априорно, по условиям задачи. Она того же порядка, что и гладкость рассматриваемой системы. Функция $R(t, x)$ (см. (11), (12)) доставляет необходимые условия оптимальности и строится вдоль определенных траекторий. При разных траекториях получается семейство функций $R(t, x)$, для которых функция Беллмана $B(t, x)$ задает огибающую поверхность.

§ 4. Динамическое программирование особых режимов

Если вспомнить вывод дифференциального уравнения Беллмана (см. § 4 гл. II), нетрудно заметить, что оно получается как уравнение первого порядка из рекуррентного уравнения (см. § 7 гл. I), аккумулирующего необходимые и достаточные свойства оптимального управления. Вполне возможно, что вдоль некоторых управлений информация, заключенная в первой производной, окажется недостаточной для характеристики оптимального управления. Этот случай наиболее четко реализуется на особых управлениях [10, 11].

1. Уравнение Беллмана и особые управления. В задаче оптимизации

$$\left. \begin{aligned} \dot{x} &= f(x, u, t), \quad x(0) = x_0, \quad t \in [0, t_1], \\ u(t) &\in U, \quad I(u) = \varphi(x(t_1)) \rightarrow \min, \\ x &= \{x_1, \dots, x_n\}, \quad u = \{u_1, \dots, u_r\}, \end{aligned} \right\} \quad (1)$$

уравнение Беллмана (см. § 4 гл. II) имеет вид

$$-\frac{\partial B(t, x)}{\partial t} = \min_{u \in U} \frac{\partial B'(t, x)}{\partial x} f(x, u, t), \quad B(t_1, x) = \varphi(x). \quad (2)$$

Оптимальное управление $u^0(x, t)$ находится из условия

$$\frac{\partial B'(t, x)}{\partial x} f(x, u^0(x, t), t) = \min_{u \in U} \frac{\partial B'(t, x)}{\partial x} f(x, u, t). \quad (3)$$

В регулярных случаях оптимальное управление из условия (3) определяется однозначно. Однако при решении широкого круга задач возникает положение, когда правая часть из (3) не зависит от управления на некотором множестве Ω элементов x, t . Подобные случаи называют особыми. Типичным примером особого случая является ситуация, связанная со скользящим режимом.

Допустим, что задача (1) не имеет решения в классе измеримых функций $u(t)$ ($t \in T$) со значениями в ограниченном множестве U . Перейдем к вспомогательной задаче, которая всегда имеет решение:

$$\left. \begin{aligned} \dot{y} &= \sum_{i=1}^{n+1} \alpha_i f(y, v_i(t), t), \quad y(0) = x_0, \\ \alpha_i(t) &\geq 0, \quad \sum_{i=1}^{n+1} \alpha_i(t) = 1, \quad v_i(t) \in U \quad (i = 1, \dots, n+1), \\ I(\alpha, v) &= \varphi(y(t_1)) \rightarrow \min_{\alpha, v}, \quad y = \{y_1, \dots, y_n\}. \end{aligned} \right\} \quad (4)$$

Для задачи (4) уравнение Беллмана имеет вид

$$-\frac{\partial \tilde{B}(t, y)}{\partial t} = \min_{\alpha_i \geq 0, \sum \alpha_i = 1, v_i \in U} \frac{\partial \tilde{B}'(t, y)}{\partial y} \sum_{i=1}^{n+1} \alpha_i f(y, v_i, t), \quad (5)$$

$$\tilde{B}(t_1, y) = \varphi(y).$$

Если сначала взять минимум по v , то получившееся выражение не будет зависеть от параметров α_i . Это значит, что во вспомогательной задаче управление $\{v, \alpha\}$ является особым по второй группе управляющих параметров. Окончательный вид уравнения такой:

$$-\frac{\partial \tilde{B}(t, y)}{\partial t} = \min_{v \in U} \frac{\partial \tilde{B}'(t, y)}{\partial y} f(y, v, t), \quad \tilde{B}(t_1, y) = \varphi(y),$$

что с точностью до обозначений совпадает с уравнением (3).

Отсюда следует, что уравнения Беллмана для задач, не имеющих решения, и их модификаций, обладающих решением, совпадают, а условия, из которых во второй задаче находятся оптимальные управления, не зависят от группы управляющих параметров. Таким образом, переход к вспомогательной задаче всегда приводит к появлению особого управления и уравнение Беллмана для его вычисления становится неэффективным.

2. Дифференциальные уравнения для функции Беллмана в особом случае [10]. Вложим задачу (1) в семейство задач

$$\frac{dx}{dt} = f(x, u, \tau), \quad x(t) = x, \quad t \leq \tau \leq t_1, \quad u(\tau) \in U, \quad (6)$$

$$I(u) = \varphi(x(t_1)) \rightarrow \min_u,$$

зависящих от n -вектора x и скаляра t .

Пусть для этих задач

$$B(t, x) = \min_{u[t, t_1], x(t)=x} I(u).$$

Согласно принципу оптимальности,

$$B(t, x) \leq B(\bar{t}, \bar{x}), \quad (7)$$

где $\bar{x} = x(\bar{t})$ — положение в момент $\tau = \bar{t} = t + \Delta t$ точки $x(t)$, движущейся по траектории уравнения (6) с допустимым управлением $u(\tau) = v, t \leq \tau \leq \bar{t}$. Неравенство (7) означает, что любое приращение функции $B(t, x)$ вдоль допустимой траектории системы (6) неотрицательно:

$$\Delta B(t, x) |_{[6], v \in U} \geq 0. \quad (8)$$

Разделим обе части этого неравенства на Δt и устремим Δt к нулю. Если предел в левой части существует, то из (8) получаем

$$\left. \frac{dB(t, x)}{dt} \right|_{(6), v \in U} \geq 0. \quad (9)$$

Может оказаться, что на некотором множестве $\Omega_1 = \Omega$ элементов x, t условие (9) выполняется как равенство. Тогда из (8) легко следует (если соответствующий предел существует) неравенство

$$\left. \frac{d^2 B(t, x)}{dt^2} \right|_{(6), v \in U} \geq 0. \quad (10)$$

Пусть аналогично на некотором множестве Ω_k выполняются равенства

$$\left. \frac{d^m B(t, x)}{dt^m} \right|_{(6), v \in U} = 0 \quad (m = 1, \dots, k). \quad (11)$$

Тогда из (8) имеем неравенство

$$\left. \frac{d^{k+1} B(t, x)}{dt^{k+1}} \right|_{(6), v \in U} \geq 0, \quad \{x, t\} \in \Omega_k. \quad (12)$$

С другой стороны, если точка \bar{x} лежит на оптимальной траектории $\bar{x} = x^0(t)$, то, согласно принципу оптимальности, $B(t, x) = B(t, \bar{x})$, т. е. приращение $\Delta B(t, x)$ вдоль оптимальной траектории тождественно равно нулю: $\Delta B(t, x) |_{(6), u^0(t)} \equiv 0$. Это означает, что для оптимального управления u^0 выполняются равенства

$$\left. \frac{d^m B(t, x)}{dt^m} \right|_{(6), u^0} = 0 \quad (m = 1, 2, \dots). \quad (13)$$

Объединяя (9) с (13), при $k=1$ получаем

$$\min_{u \in U} \left. \frac{dB(t, x)}{dt} \right|_{(6), u} = 0. \quad (14)$$

Аналогично из (12), (13) с учетом условия (11) следует

$$\min_{u \in U} \left. \frac{d^k B(t, x)}{dt^k} \right|_{(6), u} = 0, \quad \{x, t\} \in \Omega_{k-1} \quad (15)$$

$$(k = 2, 3, \dots).$$

При явной записи оператора d^k/dt^k через условия задачи из уравнений (14), (15) получаются уравнения в частных производных. Например, при $k=1$ имеем

$$-\frac{\partial B}{\partial t} = \min_{u \in U} 2 \frac{\partial B'(t, x)}{\partial x} f(x, u, t),$$

что совпадает с уравнением Беллмана (2). Пусть $k=2$ и выполняется (11), $m=1$. Тогда на множестве Ω_1 функция $B(t, x)$ в силу (15) при $k=2$ удовлетворяет уравнению

$$-\frac{\partial^2 B}{\partial t^2} = \min_{u \in U} \left[2 \frac{\partial^2 B'}{\partial x \partial t} f(x, u, t) + \right. \\ \left. + f'(x, u, t) \frac{\partial^2 B}{\partial x^2} f(x, u, t) + \frac{\partial B'}{\partial x} \cdot \frac{d}{dt} f(x, u, t) \right]_{(6), u}.$$

Граничные условия для этих уравнений следуют из определения функции $B(t, x)$ и условий сопряжения на множествах Ω_k . Оптимальные управления также получаются из (12), (13). При $k=1$ имеем (3), при $k=2$ функция $u^0(x, t)$ на Ω_1 удовлетворяет условию

$$2 \frac{\partial^2 B'(t, x)}{\partial x \partial t} f(x, u^0(x, t), t) + f'(x, u^0(x, t), t) \times \\ \times \frac{\partial^2 B(t, x)}{\partial x^2} f(x, u^0(x, t), t) + \frac{\partial B'(t, x)}{\partial x} \frac{d}{dt} f(x, u, t) \Big|_{(6), u=u^0(x, t)} = \\ = \min_{u \in U} \left[2 \frac{\partial^2 B(t, x)}{\partial x \partial t} f(x, u, t) + \right. \\ \left. + f'(x, u, t) \frac{\partial^2 B(t, x)}{\partial x^2} f(x, u, t) + \right. \\ \left. + \frac{\partial B'(t, x)}{\partial x} \cdot \frac{d}{dt} f(x, u, t) \right]_{(6), u}.$$

Пример 1. $\dot{x} = u$, $\dot{y} = xu$, $\varphi(x, y) = y$, $T = [0, 1]$, $U = \{u : |u| \leq 1\}$, x, y, u — скаляры.

Уравнение Беллмана

$$-\frac{\partial B(t, x, y)}{\partial t} = \min_{|u| \leq 1} \left[\frac{\partial B(t, x, y)}{\partial x} u + \frac{\partial B(t, x, y)}{\partial y} xu \right], \\ B(1, x, y) = y.$$

Взяв операцию \min , получаем уравнение частных производных:

$$\frac{\partial B(t, x, y)}{\partial t} = \left| \frac{\partial B(t, x, y)}{\partial x} + \frac{\partial B(t, x, y)}{\partial y} x \right|, B(1, x, y) = y.$$

Нетрудно проверить, что решением этого уравнения является функция

$$B(t, x, y) = \begin{cases} y - \frac{x^2}{2} \text{ при } t + |x| \leq 1 & (0 \leq t \leq 1), \\ y - |x|(1-t) + \frac{(1-t)^2}{2} \text{ при } t + |x| \geq 1 & (0 \leq t \leq 1). \end{cases}$$

Выражение, подлежащее минимизации, согласно (3), имеет вид

$$\frac{\partial B'(t, x)}{\partial x} f(x, u, t) = \begin{cases} 0 & \text{при } t + |x| < 1 \quad (0 \leq t \leq 1), \\ u(t + |x| - 1) \operatorname{sign} x & \text{при } t + |x| \geq 1 \\ & (0 \leq t \leq 1). \end{cases}$$

Таким образом, в области $\Omega = \{x, t: t + |x| \leq 1, 0 \leq t \leq 1\}$ условие (3) неэффективно.

Из (15) ($k=2$) следует, что оптимальное управление $u^0(x, t)$ должно минимизировать выражение

$$\begin{cases} 0 & \text{при } t + |x| < 1 \quad (0 \leq t \leq 1), \\ 2u + u^2 & \text{при } t + |x| = 1 \quad (0 \leq t \leq 1). \end{cases} \quad (16)$$

Оптимальное управление для области $t + |x| > 1 \quad (0 \leq t \leq 1)$

$$u^0(x, t) = -\operatorname{sign} x.$$

Минимизация выражения (16) доопределяет управление на множестве $t + |x| = 1 \quad (0 \leq t \leq 1)$: $u^0(x, t) = -\operatorname{sign} x$. Доопределение управления $u^0(x, t)$ на внутренние точки множества Ω не вызывает затруднений: оно должно быть лишь согласовано со значениями на границе. В данном примере продолжение $u^0(x, t)$ на $\operatorname{int} \Omega$ можно осуществить бесчисленным количеством способов. Положим для определенности

$$u^0(x, t) = \begin{cases} -\operatorname{sign} x, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

Доопределение же $u^0(x, t)$ на граничные точки множества Ω может встретить серьезные трудности.

Пример 2. $\dot{x} = u$, $\dot{y} = x^2$, $\varphi(x, y) = y$, $T = [0, 1]$, $U = \{u: |u| \leq 1\}$, x, y, u — скаляры.

Уравнение Беллмана (2) приводит к следующему уравнению в частных производных:

$$\frac{\partial B(t, x, y)}{\partial t} = \left| \frac{\partial B(t, x, y)}{\partial x} \right| - \frac{\partial B(t, x, y)}{\partial y} x^2, B(1, x, y) = y,$$

решением которого является функция

$$B(t, x, y) = \begin{cases} y + \frac{|x|^3}{3}, & t + |x| \leq 1 \quad (0 \leq t \leq 1), \\ y + x^2(1-t) - (1-t)^2|x| + \frac{(1-t)^3}{3}, & t + |x| \geq 1 \\ & (0 \leq t \leq 1). \end{cases}$$

Далее,

$$\frac{\partial B'(t, x)}{\partial x} f(x, u, t) = \begin{cases} |x|xu + x^2, & t + |x| \leq 1 \quad (0 \leq t \leq 1), \\ x^2 + 2ux(1-t) - u(1-t)^2 \operatorname{sign} x, & t + |x| \geq 1 \quad (0 \leq t \leq 1). \end{cases}$$

Множество Ω имеет вид $x=0$. Вне множества Ω оптимальное управление $u^0(x, t) = -\operatorname{sign} x$. Хотя каждое из продолжений

$$u^0(x, t) = -1, \quad \{x, t\} \in \Omega,$$

$$u^0(x, t) = 1, \quad \{x, t\} \in \Omega,$$

удовлетворяет условию (3), ни одно из них не приемлемо. Очевидное в данном примере продолжение $u^0(x, t) = 0, x=0$, получается из (15) при $k=3$.

Теперь несколько изменим условия примера 2. Вместо приведенного множества U рассмотрим $U = \{u: |u|=1\}$. Это изменение никак не скажется на уравнении Беллмана. Попытка доопределить $u^0(x, t)$ на Ω с помощью (3) не приводит к цели. Это один из признаков наличия в задаче скользящего режима, который обязывает перейти к вспомогательной задаче (4).

Следовательно, трудности при использовании уравнения Беллмана (и других уравнений, следующих из принципа оптимальности) для получения достаточных условий оптимальности связаны не только с решением уравнения в частных производных, но и с нахождением оптимального управления в особых случаях.

ЛИТЕРАТУРА

1. Айзекс Р. Дифференциальные игры. «Мир», 1967.
2. Арис Р. Дискретное динамическое программирование. «Мир», 1969.
3. Беллман Р. Динамическое программирование. ИЛ, 1960.
4. Беллман Р. Процессы регулирования с адаптацией. «Наука», 1964.
5. Беллман Р., Дрейфус С. Прикладные задачи динамического программирования. «Наука», 1965.
6. Беллман Р., Калаба Р. Квазилинеаризация и нелинейные краевые задачи. «Мир», 1968.
7. Беллман Р., Калаба Р. Динамическое программирование и современная теория управления. «Наука», 1969.
8. Болтянский В. Г. Достаточные условия оптимальности и обоснование метода динамического программирования. Изв. АН СССР, сер. матем., т. XXVIII, № 3, 1961.
9. Брайсон А., Хо Ю-ши. Прикладная теория оптимального управления. «Мир», 1972.
10. Габасов Р., Кириллова Ф. М. Метод динамического программирования в теории оптимальных особых управлений. «Автоматика и телемеханика», № 8, 1970.
11. Габасов Р., Кириллова Ф. М. Качественная теория оптимальных процессов. «Наука», 1971.
12. Гельфанд И. М., Фомин С. В. Вариационное исчисление. Физматгиз, 1961.
13. Кротов В. Ф., Букреев В. З., Гурман В. И. Новые методы вариационного исчисления в динамике полета. «Машиностроение», 1969.
14. Леви П. Стохастические процессы и броуновское движение. «Наука», 1972.
15. Летов А. М. Аналитическое конструирование регуляторов, I—III. «Автоматика и телемеханика», № 4—6, 1960.
16. Ляпунов А. М. Общая задача об устойчивости движения. Гостехиздат, 1950.
17. Понtryгин Л. С., Болтянский В. Г., Гамкрелидзе Р. В., Мищенко Е. Ф. Математическая теория оптимальных процессов. «Наука», 1969.
18. Репин Ю. М., Третьяков В. Е. Решение задачи об аналитическом конструировании регуляторов на электронных моделирующих устройствах. «Автоматика и телемеханика», № 6, 1963.
19. Розоноэр Л. И. Принцип максимума в теории оптимальных систем, I—III. «Автоматика и телемеханика», № 10—12, 1959.

20. Х о в а р д Р. А. Динамическое программирование и марковские процессы. «Советское радио», 1964.
21. Х а л а н а й А., В е к с л е р Д. Качественная теория импульсных систем. «Мир», 1971.
22. K a l m a n R. E. Contributions to the theory of optimal control. Proc. Symp. on Ordinary Differential Equations. Mexico City, 1959.
23. K a l m a n R. E. Control of Randomly Varying Linear Dynamic Systems. Proc. Symp. Appl. Mathem., v. 13. 1962.
24. L a r s o n R. E. State Increment Dynamic Programming. New York, American: Elsevier, 1967.
25. N e m h a u s e r G. L. Introduction to Dynamic Programming. New York, Willey, 1966.

- Г 12 **Габасов Р., Кириллова Ф. М.**
Основы динамического программирования. Мн.,
Изд-во БГУ, 1975.
264 с. с ил.

В книге излагается один из основных методов теории оптимальных процессов. Принцип оптимальности, функциональные уравнения Беллмана рассматриваются сначала на простейших моделях многоэтапных детерминированных процессов, а затем на более сложных (включая непрерывные, стохастические и игровые задачи). Большое внимание уделяется вычислительным аспектам динамического программирования. Рассматривается задача аналитического конструирования регуляторов и связь ее с задачей стабилизации движения.

Книга будет полезна научным работникам, аспирантам и студентами вузов, специализирующимся в области прикладной математики.

Рафаил Габасов
Фаина Михайловна Кириллова

ОСНОВЫ ДИНАМИЧЕСКОГО ПРОГРАММИРОВАНИЯ

Редактор *А. И. Калинин*
Редакторы издательства *З. И. Лакизо, Т. М. Белая*
Обложка *Л. Г. Медведевой*
Художественный редактор *В. Т. Лапицкий*
Технический редактор *В. П. Безбородова*
Корректоры *В. М. Кизил, Л. П. Морозова, Г. Н. Новиков*

АТ 06155 Сдано в набор 15/XI 1973 г. Подписано к печати 21/V 1975 г. Формат 84X108^{1/32}. Бумага типогр № 1. Физ-печ л 8,25 Усл. печ л. 13,86. Уч.-изд. л. 14,3. Тираж 4000 экз Заказ № 286. Цена 1 р. 59 к.

Издательство Белорусского государственного университета
им. В. И. Ленина. Минск, ул. Кирова, 24.

Типография Издательства БГУ им. В. И. Ленина.
Минск, ул. Кирова, 24.

