ТЕМА 5 МЕТОД ДИНАМИЧСКОГО ПРОГРАММИРОВАНИЯ БЕЛЛМАНА ДЛЯ СИСТЕМ УПРАВЛЕНИЯ С НЕПРЕРЫВНЫМ И ДИСКРЕТНЫМ ВРЕМЕНЕМ

5.1. Метод динамического программирования решения задач оптимального управления

Рассмотрим задачу оптимального управления: найти управления и траектории, на которых функционал

$$Q = \int_{t_0}^{t_1} G(x,u,t)dt + \Phi(x(t_1)) \tag{5.1}$$

достигает своего экстремального (минимального) значения для системы

$$\dot{x}(t) = f(x,u,t), \tag{5.2}$$

где

$$x(t) = (x_1(t),\ldots,x_n(t))^T \in \Omega_t(X) \subseteq X\,, \tag{5.3}$$

$$u(t) = (u_1(t),\ldots,u_r(t))^T \in \Omega_t(U) \subseteq U\,. \tag{5.4}$$

Здесь X – фазовое пространство, U – пространство управляющих воздействий, $t \in [t_0, t_1]$.

В задачах (5.1) - (5.4) моменты времени t_0, t_1 в общем случае считаются неизвестными и подлежат определению. Эти моменты после их определения будем обозначать через t_0^0, t_1^0.

Метод динамического программирования является следствием принципа оптимальности, который был сформулирован Р.Белманом [2]. Принцип оптимальности справедлив для достаточно широкого класса задач оптимального управления, но не для всех. Для задачи (5.1) – (5.4) принцип оптимальности может быть сформулирован следующим образом: если некоторая траектория AC управляемой системы (5.2) является оптимальной траекторией задачи (5.1) – (5.4), то траектория BC также будет оптимальной при любом выборе точки B на оптимальной траектории AC.

Приведем другую формулировку принципа оптимальности. Пусть $u^0(t), x^0(t),\ t_0 \le t \le t_1$ – решение задачи (5.1) – (5.4), где $u^0(t)$ – оптимальное управление, $x^0(t)$ – оптимальная траектория, и пусть t' – произвольный фиксированный момент времени, $t' \in [t_0, t_1]$. Тогда решение задачи (5.1) – (5.4) для $t \geq t'$ определяется фиксированным значением $x^0(t')$ и не зависит от $u^0(t), x^0(t)$ для $t < t'$, то есть

$$\inf_{u \in \Omega(U(x^0(t')))} \{\int_{t'}^{t_1} G(x,u,t)dt + \Phi(x(t_1))\} = \int_{t'}^{t_1} G(x^0,u^0,t)dt + \Phi(x^0(t_1)).$$

Для задачи (5.1) – (5.4) принцип оптимальности Беллмана следует из свойства аддитивности определенного интеграла [4].

Доказательство принципа оптимальности можно провести следующим образом. Пусть:

$$Q^0 = \inf_{\substack{u(t) \in \Omega_t(U) \\ x(t) \in \Omega_t(X) \\ t_0^0 \leq t \leq t_1^0}} Q(u) = \int_{t_0^0}^{t_1^0} G(x^0,u^0,t)dt + \Phi(x^0(t_1^0)) =$$

$$= \int_{t_0^0}^{t^*} G(x^0,u^0,t)dt + \int_{t^*}^{t_1^0} G(x^0,u^0,t)dt + \Phi(x^0(t_1^0)) = Q_1^0 + Q_2^0.$$

Здесь t^* – произвольная точка с $[t_0^0, t_1^0]$.

Рассмотрим задачу (5.1) - (5.4) при условии, что:

$$t_0^0 = t^*, x(t^*) = x^0(t^*).$$

Решение этой задачи обозначим через $\tilde{u}(t), \tilde{x}(t), t^* \leq t \leq t_1^0$.

Допустим, вопреки принципу оптимальности, это решение не совпадает с $u^0(t), x^0(t)$ при $t > t^*$. Тогда

$$\tilde{Q} = \int_{t^*}^{\tilde{t}_1} G(\tilde{x},\tilde{u},t)dt + \Phi(\tilde{x}(\tilde{t}_1)) < Q_2^0.$$

Построим допустимое управление для задачи (5.1) – (5.4) в виде кусочно-непрерывной функции

$$u_*(t) = \begin{cases} u^0(t), & t_0^0 \leq t < t^*, \\ \tilde{u}(t), & t^* \leq t < t_1^0. \end{cases}$$

Соответствующая этому управлению траектория будет иметь вид:

$$x_*(t)=\begin{cases} x^0(t), & t_0^0 \le t < t^*, \\ \tilde{x}(t), & t^* \le t \le t_1^0. \end{cases}$$

Для решения $u_*(t), x_*(t)$ задачи (5.1) – (5.4) будем иметь

$$Q(u_*)=Q_1^0+\tilde{Q}<Q_1^0+Q_2^0=Q^0.$$

Последнее неравенство означает, что решение $u^0(t), x^0(t)$ не является оптимальным, поскольку $u_*(t)$ дает меньшее значение функционала Q.

Полученное противоречие доказывает справедливость принципа оптимальности.

5.1.1. Уравнение Беллмана для непрерывных систем в интегральной форме

Приведем дискретный аналог задачи оптимального управления (5.1) – (5.4). Разобьем заданный интервал времени равномерно точками: $t_0^0=t_0,t_1,t_2,\ldots t_k,\ldots,t_1^0=t_N$.

Обозначим подинтервалы времени через $\Delta t_k=\Delta t=t_{k+1}-t_k$, состояние системы в момент времени t_k через $x(t_k)=x_k, k=\overline{0,N}$, и управления в соответствии $u(t_k)=u_k, k=\overline{0,N-1}$.

Тогда дискретный аналог функционала (5.1) будет иметь вид:

$$\begin{gathered} Q=Q(x_0,t_0,\{u_k\}_{k=0}^{k=N-1})=\sum_{k=0}^{N-1}G(x_k,u_k,t_k)\Delta t_k+\Phi(x_N)= \\ =\sum_{k=0}^{N-1}F_0(x_k,u_k,t_k)+\Phi(x_N)\longrightarrow\inf \end{gathered} \tag{5.5}$$

Дискретный аналог для системы (5.2) получим следующим образом:

$$\frac{x_{k+1}-x_k}{\Delta t_k} = f(x_k, u_k, t_k)\text{, откуда}$$

$$x_{k+1} = x_k + f(x_k, u_k, t_k)\Delta t_k = F(x_k, u_k, t_k),\ k = \overline{0, N-1}, \qquad (5.6)$$

Множества (5.3), (5.4) в случае дискретного времени будут выглядеть соответственно:

$$x_k \in \Omega_k(X),\ \ k = \overline{0, N}\ , \qquad (5.7)$$

$$u_k \in \Omega_k(U),\ \ k = \overline{0, N-1}\,. \qquad (5.8)$$

Для постановки задачи оптимального управления дискретной системой (5.6) предполагается, что множества (5.7), (5.8) непустые и ограничены. Задача оптимального управления (5.5) – (5.8) имеет смысл только в том случае, когда из точек множества $\Omega_0(X)$ можно перейти в точки множества $\Omega_N(X)$ через точки множеств $\Omega_k(X)$, $k = \overline{1, N-1}$.

Определение 5.1. Множество $\Omega_N(X)$ называется **достижимым** из точек $x_k \in \Omega_k(X)$, $k = \overline{0, N-1}$, если существуют такие допустимые управления $\{u_j\}$, $j = \overline{k, N-1}$, что соответствующая им, согласно уравнению (5.6), траектория $\{x_j\}$, $j = \overline{k, N}$ с начальной точкой x_k соединяет эту точку с некоторой точкой множества $\Omega_N(X)$.

Если множество начальных значений $\Omega_0(X)$ состоит не из одного элемента, то задача (5.5) - (5.8) разбивается на две задачи:

а) нахождения допустимых управлений, на которых достигается минимум функционала (5.5) при фиксированном значении $x_0 \in \Omega_0(X)$, то есть

$$\min_{\{u_k\}_{k=0}^{k=N-1}} Q(x_0, t_0, \{u_k\}_{k=0}^{k=N-1}) = Q(x_0, t_0)\,;$$

б) нахождения минимума $Q(x_0, t_0)$ как функции переменной x_0 на множестве $\Omega_0(X)$, то есть

$$Q^0 = \min_{x_0 \in \Omega_0(X)} Q(x_0, t_0) = Q^0(t_0)\,.$$

Для фиксированного момента времени $t_k, k=\overline{0,N-1}$ введем некоторую функцию $S_k(x_k,t_k)$, которую будем называть **функцией Беллмана** в виде:

$$S_k(x_k,t_k)=\min_{\{u_j\}_{j=k}^{j=N-1}}\{\sum_{j=k}^{N-1}F_0(x_j,u_j,t_j)+\Phi(x_N)\}=$$

$$=\sum_{j=k}^{N-1}F_0(x_j,u_j^0(x_k),t_j)+\Phi(x_N)\ , \qquad (5.9)$$

где $u_j^0(x_k)$, $k=\overline{j,N-1}$, – последовательность управлений, что соответствует оптимальному движению системы (5.6) с некоторой точки $x_k\in\Omega_k(X)$, взятой в момент t_k, в точки множества $\Omega_N(X)$.

Выделим в формуле (5.9) первый член. Имеем

$$S_k(x_k,t_k)=F_0(x_k,u_k^0(x_k),t_k)+\sum_{j=k+1}^{N-1}F_0(x_j,u_j^0(x_k),t_j)+\Phi(x_N)\ .$$

Далее возьмем $j=k+1$ и для управлений $u_k^0(x_k)$, под действием которых система (5.6) переходит в точку x_{k+1}: $x_{k+1}=F_0(x_k,u_k^0(x_k),t_k)$, рассмотрим функцию Беллмана

$$S_{k+1}(x_{k+1},t_{k+1})=\min_{\{u_k\}_{j=k+1}^{j=N-1}}\{\sum_{j=k+1}^{N-1}F_0(x_j,u_j,t_j)+\Phi(x_N)\}=$$

$$=\sum_{j=k+1}^{N-1}F_0(x_j,u_j^0(x_{k+1}),t_j)+\Phi(x_N), \qquad (5.10)$$

где $u_j^0(x_{k+1})$, $j=\overline{k+1,N-1}$ – последовательность управлений, которые соответствуют оптимальному движения системы (5.6) с указанной точки $x_{k+1}\in\Omega_{k+1}(X)$ в точки множества $\Omega_N(X)$.

Из принципа оптимальности Беллмана следует, что решение задачи (5.5) – (5.8) на интервале $[t_k,t_N]$ совпадает с решением соответствующей задачи на $[t_{k+1},t_N]$, если переход от x_k к x_{k+1} осуществлено согласно с оптимальным управлением $u_k^0(x_k)$ для системы управления (5.6). Отсюда получим, что управления будут совпадать:

$$u_k^0(x_k) = u_j^0(x_{k+1}), \ j = \overline{k+1, N-1},$$

Итак, учитывая это и формулу (5.10), выражение для функции $S_k(x_k, t_k)$ можно записать в виде:

$$S_k(x_k, t_k) = F_0(x_k, u_k^0(x_k), t_k) + S_{k+1}(x_{k+1}, t_{k+1}).$$

Поскольку $x_{k+1} = F_0(x_k, u_k^0(x_k), t_k)$, окончательно получим:

$$S_k(x_k, t_k) = \min_{u_k \in \Omega_k(U)} \{F_0(x_k, u_k, t_k) + S_{k+1}(F(x_k, u_k, t_k), t_{k+1})\} \quad (5.11)$$

для всех $k = \overline{0, N-1}$.

При этом $S_N(x_N, t_N) = \Phi(x_N)$.

Уравнение (5.11) называется **разностным уравнением Беллмана.**

Полученное уравнение Беллмана лежит в основе **метода динамического программирования.**

5.2. Алгоритм метода динамического программирования для дискретных систем

Алгоритм метода динамического программирования решения задачи вида (5.5) – (5.8) для дискретных систем управления состоит из двух частей: нахождение управлений как функций от состояний системы (прямой ход) и вычисления оптимальных управлений и оптимальной траектории (обратный ход).

<u>А: Прямой ход.</u>

Шаг 1. Положим в уравнении Беллмана (5.11) $k = N-1$ и решим задачу

$$S_{N-1}(x_{N-1}, t_{N-1}) =$$
$$= \min_{u_{N-1} \in \Omega_{N-1}(U)} \{F_0(x_{N-1}, u_{N-1}, t_{N-1}) + \Phi(F(x_{N-1}, u_{N-1}, t_{N-1}))\}$$

для всех точек множества $\Omega_{N-1}(X)$, из которых достижимо (можно перевести систему) множество $\Omega_N(X)$, то есть для точек $x_N = F(x_{N-1}, u_{N-1}(x_{N-1}), t_{N-1}) \in \Omega_N(X)$.

Находим $u^0_{N-1}(x_{N-1})$ как функцию точек $x_{N-1} \in \Omega_{N-1}(X)$.

Шаг 2. Для $k = N-2$ решим задачу

$$S_{N-2}(x_{N-2}, t_{N-2}) =$$
$$= \min_{u_{N-2} \in \Omega_{N-2}(U)} \{F_0(x_{N-2}, u_{N-2}, t_{N-2}) + S_{N-1}(x_{N-1}, t_{N-1})\} =$$
$$= \min_{u_{N-2} \in \Omega_{N-2}(U)} \{F_0(x_{N-2}, u_{N-2}, t_{N-2}) +$$
$$+ S_{N-1}(F(x_{N-2}, u_{N-2}, t_{N-2}), t_{N-1})\}$$

для всех $x_{N-2} \in \Omega_{N-2}(X)$, из которых достижимо множество $\Omega_N(X)$.

Отсюда находим $u^0_{N-2}(x_{N-2})$ для $x_{N-2} \in \Omega_{N-2}(X)$.

Продолжаем дальше процесс, пока не дойдем до $k = 0$.

Шаг N. Для $k = 0$ решим задачу

$$S_0(x_0, t_0) = \min_{u_0 \in \Omega_0(U)} \{F_0(x_0, u_0, t_0) + S_1(F(x_0, u_0, t_0), t_1)\}$$

для всех $x_0 \in \Omega_0(X)$, из которых достижимо множество $\Omega_N(X)$.

Получим $u^0_0(x_0), x_0 \in \Omega_0(X)$.

В: Обратный ход.

Если множество $\Omega_0(X)$ состоит более чем из одного элемента, то нужно решить задачу:

$$\min_{x_0 \in \Omega_0(X)} S_0(x_0, t_0) = S_0(x^0_0, t_0) = Q^0(t_0).$$

Найдя x^0_0, получим оптимальное управление $u^0_0(x^0_0) = u^0_0$ в момент времени $t = t_0$.

Шаг 1. Подставим найденные оптимальные x^0_0, $u^0_0(x^0_0)$ в уравнение (5.6):

$$x^0_1 = F(x^0_0, u^0_0, t_0).$$

Таким образом, найдем x^0_1 в момент $t = t_1$. Подставляя значения x^0_1 в функцию $u^0_1(x^0_1)$, полученную на прямом ходе алгоритма, находим оптимальное управление $u^0_1(x^0_1) = u^0_1$.

Продолжаем этот процесс.

$\vdots$

Шаг N. Аналогично находим управления $u_{N-1}^0(x_{N-1}^0) = u_{N-1}^0$ и точку $x_N^0 = F(x_{N-1}^0, u_{N-1}^0, t_{N-1})$.

Таким образом, нашли $\{u_j^0\}, j = \overline{0, N-1}, \{x_j^0\}, j = \overline{0, N}$ – оптимальное управление и оптимальную траекторию для задачи (5.5) – (5.8).

Замечание 5.1. К достоинствам метода динамического программирования можно отнести сведение исходной задачи большой размерности к последовательному решению однотипных задач меньшей размерности. Таким образом, вместо одновременного нахождения всех $N \cdot r$ неизвестных управлений для задачи оптимального управления (5.5) – (5.8) последовательно решаем N задач условной минимизации по $u_k, k = \overline{0, N-1}$ и каждая из этих задач имеет r неизвестных.

Замечание 5.2. Метод динамического программирования всегда дает решение задачи синтеза оптимального управления, которая заключается в нахождении оптимального управления как функции фазовых координат системы. В частности, для дискретных систем синтезирующие управления получаем на прямом ходе алгоритма.

5.3. Уравнение Беллмана для непрерывных систем управления

5.3.1. Уравнение Беллмана для непрерывных систем в интегральной форме

Рассмотрим задачу оптимального управления с фиксированным временем и свободным правым концом траектории. Нужно найти минимум функционала

$$Q = Q(u) = \int_{t_0}^{t_1} G(x, u, t)dt + \Phi(x(t_1)) \tag{5.12}$$

для системы

$$\dot{x}(t) = f(x, u, t) \tag{5.13}$$

с начальным условием $x(t_0) = x_0$ и с ограничениями на управление

$$u = u(t) \in \Omega_t(U) \tag{5.14}$$

и на траектории

$$x = x(t) \in \Omega_t(X)\,, \tag{5.15}$$

для всех $t_0 \le t \le t_1$.

Считаем, что моменты времени t_0, t_1 фиксированные, функции $G(x,u,t), \Phi(x(t))$, вектор-функции $f(x,u,t)$ – непрерывные по переменным x,u и кусочно-непрерывные по t на интервале $[t_0,t_1]$.

Кроме того, для функции $f(x,u,t)$ выполняются условия Липшица по переменной управления, то есть для произвольных w,v из множества (5.14) выполняется условие:

$$\left| f(x,w,t) - f(x,v,t) \right| \le \alpha \left| w - v \right|,$$

где $\alpha > 0$ – некоторая постоянная величина.

Предположим, что $u(t)$ – кусочно-непрерывная функция переменной t на интервале $[t_0,t_1]$.

Возьмем для произвольного фиксированного момента времени $t \in [t_0,t_1]$ некоторую точку $x = x(t) \in \Omega_t(X)$. Для этих t и $x(t)$, которые возьмем за t_0 и $x(t_0)$ соответственно, рассмотрим задачу (5.12) – (5.15). Решение этой задачи запишем как $u^0(\tau,x)$, $x^0(\tau)$, $t_0 \le \tau \le t_1$. Минимум соответствующего функционала для данного решения обозначим через $S(x,t)$.

$$S(x,t) = \min_{\substack{u(\tau)\in\Omega_\tau(U)\\ t\le\tau\le t_1\\ x=x(t)\in\Omega_t(X)}} \{\int_t^{t_1} G(x(\tau),u(\tau),\tau)d\tau + \Phi(x(t_1))\} =$$

$$= \int_t^{t_1} G(x^0(\tau),u^0(\tau,x),\tau)d\tau + \Phi(x^0(t_1))\,.$$

Возьмем на интервале $[t,t_1]$ произвольный момент времени $t+\Delta t$ и точку $x(t+\Delta t) \in \Omega_{t+\Delta t}(X)$. Рассмотрим задачу (5.12) – (5.15) для этих $t+\Delta t$, $x(t+\Delta t)$, которые возьмем t_0 та $x(t_0)$ соответственно. Отметим, что эта задача отличается от предыдущей лишь исходными данными.

Минимум соответствующего функционала обозначим через $S(x(t+\Delta t),t+\Delta t)$.

$$S(x(t+\Delta t),t+\Delta t)=$$

$$=\min_{\substack{u(\tau)\in\Omega_\tau(U)\\ t+\Delta t\le\tau\le t_1\\ x(t+\Delta t)\in\Omega_{t+\Delta t}(X)}}\{\int_{t+\Delta t}^{t_1}G(x(\tau),u(\tau),\tau)d\tau+\Phi(x(t_1))\}=$$

$$=\int_{t+\Delta t}^{t_1}G(\tilde{x}(\tau),\tilde{u}(\tau,x(t+\Delta t),\tau)d\tau+\Phi(\tilde{x}(t_1)),$$

где $\tilde{u}(\tau,x(t+\Delta t))$, $\tilde{x}(\tau)$ – решение задачи (5.12) – (5.15) на интервале $[t+\Delta t,t_1]$.

Выберем за точку $x(t+\Delta t)$ то состояние системы (5.13), в который она попадает в момент $t+\Delta t$, двигаясь из точки $x(t)$ по оптимальной траектории $x^0(\tau)$, то есть за состояние $x(t+\Delta t)$ возьмем состояние $x^0(t+\Delta t)$. Тогда, согласно принципу оптимальности, решения приведенных выше двух задач совпадают на интервале $t+\Delta t\le\tau\le t\le t_1$. Поэтому

$$S(x(t+\Delta t),t+\Delta t)=\int_{t+\Delta t}^{t_1}G(x^0(\tau),u^0(\tau,x),\tau)d\tau+\Phi(x^0(t_1)).$$

Вернемся к значению функционала $S(x,t)$. Используя свойство аддитивности интеграла можем записать

$$S(x,t)=\int_{t}^{t+\Delta t}G(x^0(\tau),u^0(\tau,x),\tau)d\tau+$$

$$+\int_{t+\Delta t}^{t_1}G(x^0(\tau),u^0(\tau,x),\tau)d\tau+\Phi(x^0(t_1)).$$

Отсюда

$$S(x,t)=\int_{t}^{t+\Delta t}G(x^0(\tau),u^0(\tau,x),\tau)d\tau+S(x^0(t+\Delta t),t+\Delta t),$$

где траєктория $x^0(t+\Delta t)$ системы (5.13) полученная под действием управления $u^0(\tau)$.

Итак,

$$S(x,t)=\int_t^{t+\Delta t} G(x^0(\tau),u^0(\tau,x),\tau)d\tau+S(x^0(t+\Delta t,u^0(\tau)),t+\Delta t) .$$

Это равенство выполняется только для оптимального управления $u^0(\tau)$. Если брать другие управления из множества допустимых управлений согласно (5.14), то правая часть последнего равенства может только увеличиться.

Итак, получим **уравнение Беллмана в интегральной форме**

$$* \qquad S(x,t)=$$

$$=\min_{\substack{u(\tau)\in\Omega_\tau(U)\\ t\le\tau\le t+\Delta t\le t_1\\ x=x(t)\in\Omega_t(X)}}\{\int_t^{t+\Delta t} G(x(\tau),u(\tau,x),\tau)d\tau+S(x(t+\Delta t,u(\tau,x)),t+\Delta t)\}.$$

Обозначим это уравнение через **∗**.

Рассмотрим задачу с закрепленными концами траекторий и свободным временем для автономной системы и запишем для этой задачи уравнение Беллмана.

Постановка задачи заключается в следующем: нужно найти управления и траектории, на которых функционал

$$Q=Q(u)=\int_{t_0}^{t_1} G(x,u)dt+\Phi(x(t_1)) \tag{5.16}$$

достигает своего минимального значения для системы

$$\dot{x}=f(x,u) \tag{5.17}$$

с закрепленными концами $x(t_0)=x_0, x(t_1)=x_1$, и ограничениями на управление и траекторию

$$u=u(t)\in\Omega_t(U)\,,\ x=x(t)\in\Omega_t(X)\,,\ t\in[t_0,t_1]. \tag{5.18}$$

Заметим, что, поскольку моменты времени не фиксированы, то минимум функционала (5.16) будет функцией только начального состояния x_0.

Обозначим это значение через $S(x_0)$:

$$S(x_0) = \min_{\substack{u(\tau)\in\Omega_\tau(U)\\ t_0\le\tau\le t_1\\ x=x(t)\in\Omega_t(X)}} \{\int_{t_0}^{t_1} G(x,u)d\tau + \Phi(x(t_1))\}.$$

Аналогично предыдущей задаче получим:

$$S(x_0) = \min_{\substack{u(\tau)\in\Omega_\tau(U)\\ t_0\le\tau\le t_0+\Delta t\\ x=x(t)\in\Omega_t(X)}} \int_{t_0}^{t_0+\Delta t} G(x,u)d\tau +$$

$$+ \min_{\substack{u(\tau)\in\Omega_\tau(U)\\ t_0+\Delta t\le\tau\le t_1\\ x=x(t)\in\Omega_t(X)}} \{ \int_{t_0+\Delta t}^{t_1} G(x,u)d\tau + \Phi(x(t_1))\}.$$

Из принципа оптимальности имеем

$$S(x_0) = \min_{\substack{u(\tau)\in\Omega_\tau(U)\\ t_0\le\tau\le t_0+\Delta t\\ x=x(t)\in\Omega_t(X)}} \left\{ \int_{t_0}^{t_0+\Delta t} G(x,u)d\tau + S(x(t_0+\Delta t, u(\tau))) \right\}.$$

Поскольку это применимо к любой точки x фазовой траектории, то уравнение Беллмана в интегральной форме для задачи со свободным временем для автономной системы (которое обозначим через **) бет иметь вид:

$$** \qquad S(x) = \min_{\substack{u(\tau)\in\Omega_\tau(U)\\ t_0\le t\le\tau\le t+\Delta t\le t_1\\ x=x(t)\in\Omega_t(X)}} \left\{ \int_{t}^{t+\Delta t} G(x,u)d\tau + S(x(t+\Delta t, u(\tau))) \right\}.$$

Рассмотрим задачу быстродействия для автономной системы, которая заключается в нахождении минимума функционала

$$T = t_1 - t_0 \qquad (5.19)$$

для системы

$$\dot{x} = f(x,u) \qquad (5.20)$$

с закрепленными концами $x(t_0) = x_0, x(t_1) = x_1$ и ограничениями на управление и траекторию

$$u = u(t) \in \Omega_t(U), \ x = x(t) \in \Omega_t(X), \ t \in [t_0, t_1]. \quad (5.21)$$

Отметим, что при $G(x,u) \equiv 1, \Phi(x) \equiv 0$ в (5.16) эта задача является частным случаем предыдущей задачи (5.16) – (5.18) и функция $S(x)$ будет иметь смысл минимального времени достижения системой (5.20) точки x_1 из начальной точки x. Обозначим это минимальное время через $T(x)$. Тогда из уравнения ** получим **уравнение Беллмана в интегральной форме** для задачи быстродействия для автономной системы (это уравнение обозначим через *****):

$$*** \qquad T(x) = \min_{\substack{u(\tau)\in\Omega_\tau(U) \\ t_0 \le t \le \tau \le t+\Delta t \le t_1 \\ x = x(t)\in\Omega_t(X)}} \left\{\Delta t + T(x(t+\Delta t, u(\tau)))\right\}.$$

5.3.2. Уравнение Беллмана в дифференциальной форме для непрерывных систем

Для задач (5.12)–(5.15), (5.16)–(5.18), (5.19)–(5.21) запишем уравнения Беллмана в дифференциальной форме. Для этого воспользуемся полученными выше уравнениями Беллмана в интегральной форме. Дополнительно к приведенным в этих задачах условиям будем считать:

а) для задачи (5.12)–(5.15): управление $u(t)$ непрерывные по t, функция $S(x,t)$ имеет непрерывные частные производные $\dfrac{\partial S(x,t)}{\partial x} = grad_x^T S(x,t) = (\dfrac{\partial S(x,t)}{\partial x_1}, \ldots, \dfrac{\partial S(x,t)}{\partial x_n})$, $\dfrac{\partial S(x,t)}{\partial t}$;

б) для задачи (5.16)–(5.18): управление $u(t)$ непрерывные по t, функция $S(x)$ имеет непрерывные частные производные $\dfrac{\partial S(x)}{\partial x} = grad_x^T S(x) = (\dfrac{\partial S(x)}{\partial x_1}, \ldots, \dfrac{\partial S(x)}{\partial x_n})$;

в) для задачи (5.19)–(5.21): управление $u(t)$ непрерывные по t, функция $T(x)$ имеет непрерывные частные производные $\dfrac{\partial T(x)}{\partial x} = grad_x^T T(x) = (\dfrac{\partial T(x)}{\partial x_1}, \ldots, \dfrac{\partial T(x)}{\partial x_n})$.

С учетом этих предположений, разложив уравнения Беллмана в интегральной форме ***,**,***** в ряды Тейлора и пренебрегая членами второго порядка и выше, можно записать эти уравнения в виде:
для задачи с фиксированным временем и свободным правым концом

$$S(x,t)=\min_{\substack{u(\tau)\in\Omega_\tau(U)\\ t_0\le t\le\tau\le t+\Delta t\le t_1\\ x=x(t)\in\Omega_t(X)}}\{G(x(\tau),u(\tau),\tau)\Delta t+S(x,t)+$$

$$+\frac{\partial S(x,t)}{\partial t}\Delta t+grad_x^T S(x,t)(x(t+\Delta t,u(\tau))-x)+o(\Delta t)\};$$

для задачи с закрепленными концами и свободным временем

$$S(x)=\min_{\substack{u(\tau)\in\Omega_\tau(U)\\ t_0\le t\le\tau\le t+\Delta t\le t_1\\ x=x(t)\in\Omega_t(X)}}\{G(x(\tau),u(\tau))\Delta t+S(x)+$$

$$+grad_x^T S(x)(x(t+\Delta t,u(\tau))-x)+o(\Delta t)\};$$

для задачи быстродействия

$$T(x)=\min_{\substack{u(\tau)\in\Omega_\tau(U)\\ t_0\le t\le\tau\le t+\Delta t\le t_1\\ x=x(t)\in\Omega_t(X)}}\{\Delta t+T(x)+$$

$$+grad_x^T T(x)(x(t+\Delta t,u(\tau))-x)+o(\Delta t)\}.$$

Здесь τ – некоторое фиксированное значение: $t_0\le t\le\tau\le t+\Delta t\le t_1$, $o(\Delta t)$ – бесконечно малая более высокого порядка малости, чем Δt. Заметим, что величины $S(x,t)$, $S(x)$, $T(x)$ слева и справа взаимно уничтожаются. При переходе к пределу при $\Delta t\to 0$ имеем:

$$\tau\to t, u(\tau)\to u(t), x(\tau)\to x(t), x(t+\Delta t,u(\tau))\to x(t),$$

$$\frac{x(t+\Delta t,u(\tau))-x(t)}{\Delta t}\to\dot{x}(t).$$

Таким образом, получим уравнение:

$$-\frac{\partial S(x,t)}{\partial t}=\min_{\substack{u(t)\in\Omega_t(U)\\ t_0\le t\le t_1}}\{G(x,u,t)+\mathrm{grad}_x^T S(x,t)\dot{x}(t)\},$$

$$\min_{\substack{u(t)\in\Omega_t(U)\\ t_0\le t\le t_1}} \{G(x,u)+\mathrm{grad}_x^T S(x)\dot{x}(t)\}=0\,,$$

$$\min_{\substack{u(t)\in\Omega_t(U)\\ t_0\le t\le t_1}} \{\mathrm{grad}_x^T T(x)\dot{x}(t)\}=-1\,.$$

Эти уравнения должны выполняться в каждой точке оптимальной траектории системы $\dot{x}=f(x,u,t)$ или $\dot{x}=f(x,u)$.

В результате **уравнение Беллмана в дифференциальной форме** примет вид (обозначим их символами + в соответствии с *****):

$$+\qquad \begin{aligned} &-\frac{\partial S(x,t)}{\partial t}=\min_{\substack{u(t)\in\Omega_t(U)\\ t_0\le t\le t_1}} \{G(x,u,t)+\mathrm{grad}_x^T S(x,t)\dot{x}(t)\},\\ &S(x,t_1)=\Phi(x(t_1)). \end{aligned}$$

$$++\qquad \begin{aligned} &\min_{\substack{u(t)\in\Omega_t(U)\\ t_0\le t\le t_1}} \{G(x,u)+\mathrm{grad}_x^T S(x)\dot{x}(t)\}=0,\\ &S(x_1)=\Phi(x_1). \end{aligned}$$

$$+++\qquad \begin{aligned} &\min_{\substack{u(t)\in\Omega_t(U)\\ t_0\le t\le t_1}} \{\mathrm{grad}_x^T T(x)\dot{x}(t)\}=-1,\\ &T(x_1)=0. \end{aligned}$$

Начальные условия, которые записаны для каждого из уравнений Беллмана, следуют из вида функций $S(x,t)$, $S(x)$, $T(x)$ соответственно. Отметим, что начальные условия для уравнений Беллмана +, ++, +++ задаются на правом конце интервала (в момент времени t_1).

Замечание 5.3. Уравнение Беллмана для дискретных и непрерывных систем является необходимым и достаточным условием оптимальности [4, 16], что дает полное обоснование метода динамического программирования.

Замечание 5.4. Для непрерывных систем, описываемых дифференциальными уравнениями, применение соответствующих уравнений Беллмана, которые являются нелинейными уравнениями в частных производных, требует дополнительных исследований [16].

5.4. Метод динамического программирования для задачи построения оптимального регулятора линейных систем управления

Приведем пример использования метода динамического программирования для синтеза оптимального регулятора линейных систем управления. Пусть объект управления описывается уравнениями:

$$\frac{dx}{dt} = A(t)x + B(t)u, \tag{5.22}$$

где $x(t) = (x_1, \ldots, x_n)^T$ – вектор состояния системы, $A(t)$ – $n \times n$ матрица, $u(t) = (u_1, \ldots, u_r)^T$ – вектор управлений, $B(t)$ – матрица размерности $n \times r$, $t \in [t_0, t_1]$, n, r - заданные челые числа.

Задано исходное состояние $x(t_0) = x_0$, время t_1 – фиксировано, состояние на конце интервала $x(t_1)$ – свободное.

Задача состоит в том, чтобы для системы (5.22) найти управления и траекторию, на которых функционал

$$Q(x,u) = \frac{1}{2}\int_{t_0}^{t_1} (x^T Q(t)x + u^T R(t)u)dt + \frac{1}{2} x^T(t_1)Fx(t_1) \tag{5.23}$$

достигает своего минимального значения.

Здесь $Q(t), R(t), F$ – симметричные положительно-определенные матрицы.

Задача оптимального управления линейной системой (5.22) с минимизацией квадратичного функционала (5.23) в теории управления называется задачей **аналитического конструирования оптимального регулятора** для линейной системы.

Решим эту задачу с помощью уравнения Беллмана в дифференциальной форме, которое для данной задачи приобретает вид:

$$-\frac{\partial S(x,t)}{\partial t} =$$

$$= \min_{u(t)}\{\frac{1}{2}x^T Q(t)x + \frac{1}{2}u^T R(t)u + \mathrm{grad}_x^T S(x,t)(A(t)x + B(t)u)\},$$

где $grad_x^T S(x,t) = \frac{\partial S^T(x,t)}{\partial x} = (\frac{\partial S(x,t)}{\partial x_1}, \dots, \frac{\partial S(x,t)}{\partial x_n})$.

Найдем управления из необходимых условий экстремума:

$$\frac{\partial S(x,t)}{\partial u} = R(t)u + B^T(t)\frac{\partial S(x,t)}{\partial x} = 0,$$

$$R(t)u = -B^T(t)\frac{\partial S(x,t)}{\partial x}.$$

Отсюда оптимальное управление будет иметь вид:

$$u^0 = -R^{-1}(t)B^T(t)\frac{\partial S(x,t)}{\partial x}.$$

Подставим это управление в уравнение Беллмана для $S(x,t)$. Имеем

$$-\frac{\partial S(x,t)}{\partial t} = \frac{1}{2}x^T Q(t)x +$$

$$+\frac{1}{2}(\frac{\partial S(x,t)}{\partial x})^T B(t)R^{-1}(t)R(t)R^{-1}(t)B^T(t)\frac{\partial S(x,t)}{\partial x} +$$

$$+(\frac{\partial S(x,t)}{\partial x})^T (A(t)x - B(t)R^{-1}(t)B^T(t)\frac{\partial S(x,t)}{\partial x})\},$$

Откуда

$$-\frac{\partial S(x,t)}{\partial t}=\frac{1}{2}x^T Q(t)x-$$
$$-\frac{1}{2}(\frac{\partial S(x,t)}{\partial x})^T B(t)R^{-1}(t)B^T(t)\frac{\partial S(x,t)}{\partial x}+$$
$$+(\frac{\partial S(x,t)}{\partial x})^T A(t)x,$$
$$S(x,t_1)=\frac{1}{2}x^T(t_1)Fx(t_1).$$

Решение этого уравнения – функцию $S(x,t)$ – будем искать в виде квадратичной формы $S(x,t)=\frac{1}{2}x^T P(t)x$, где $P(t)$ – симметричная матрица, которая подлежит определению. Найдем производные по времени и по состояниями этой функции.

Имеем
$$\frac{\partial S(x,t)}{\partial x}=P(t)x, \frac{\partial S(x,t)}{\partial t}=\frac{1}{2}x^T\frac{dP(t)}{dt}x.$$

Подставим эти выражения в уравнение Беллмана для $S(x,t)$ и получим:
$$-\frac{1}{2}x^T\frac{dP(t)}{dt}x=\frac{1}{2}x^T Q(t)x-\frac{1}{2}x^T P(t)B(t)R^{-1}(t)B^T(t)P(t)x+$$
$$+x^T P(t)A(t)x.$$

Это будет выполняться для произвольных значений x тогда и только тогда, когда матрица $P(t)$ удовлетворяет уравнению

$$-\frac{1}{2}\frac{dP(t)}{dt}=\frac{1}{2}Q(t)-\frac{1}{2}P(t)B(t)R^{-1}(t)B^T(t)P(t)+P(t)A(t).$$

Воспользуемся тем, что для матриц справедливо: $CD=\frac{1}{2}CD+\frac{1}{2}C^T D$.

Окончательно получим

$$\frac{dP(t)}{dt} = -P(t)A(t) - A^T(t)P(t) - Q(t) + \\ + P(t)B(t)R^{-1}(t)B^T(t)P(t). \tag{5.24}$$

$$P(t_1) = F. \tag{5.25}$$

Дифференциальное уравнение (5.24) называются **матричным уравнением Риккати**.

Таким образом, матричная функция $P(t)$, $t \in [t_0, t_1]$ является решением задачи Коши (5.24), (5.25) с обратным направлением изменения аргумента t.

Решив эту задачу Коши, найдем $P(t)$, а значит и функцию $S(x,t)$. Тогда оптимальное управление получим в виде

$$u^0 = -R^{-1}(t)B^T(t)P(t)x$$

и система управления примет вид

$$\frac{dx}{dt} = \left[A(t) - B(t)R^{-1}(t)B^T(t)P(t)\right]x.$$

Вместе с начальным условием $x(t_0) = x_0$ имеем задачу Коши решив которую получим оптимальную траекторию.

Таким образом, используя метод динамического программирования, для линейной нестационарной системы (5.22) нашли оптимальное управление и оптимальную траекторию на которых функционал (5.23) достигает своего минимального значения.

ЛИТЕРАТУРА

Основная

1. Понтрягин Л.С., Болтянский В.Г., Гамкрелидзе Р.В., Мищенко Е.Ф. Математическая теория оптимальных процессов. – М., 1983.

2. Беллман Р. Динамическое программирование. – М., 1960.

3. Брайсон А., Хо Ю-ши. Прикладная теория оптимального управления. Оптимизация, оценка и управление. – М., 1972.

4. Бублик Б.Н., Кириченко Н.Ф. Основы теории управления. – К., 1975.

5. Васильев Ф.П. Численные методы решения экстремальных задач. – М., 1980.

6. Зубов В.И. Лекции по теории управления. – М., 1975.

7. Моисеев Н.Н. Элементы теории оптимальных систем. – М., 1975.

8. Острем К. Введение в стохастическую теорию оптимального управления. – М., 1973.

9. Атанс М., Фалб П. Оптимальное управление. – М., 1968.

10. Флеминг У., Ришел Р. Оптимальное управление детерминированными и стохастическими системами. – М., 1978.

Дополнительная

11. Алексеев В.М., Тихомиров В.М., Фомин С.В. Оптимальное управление. – М., 1979.

12. Андреев Ю.Н. Управление конечномерными линейными обьектами. – М., 1976.

13. Аоки М. Оптимизация стохастических систем. – М., 1971.

14. Балакришнан А.В. Теория фильтрации Калмана. – М., 1984.

15. Бейко И.В., Бублик Б.Н., Зинько П.Н. Методы и алгоритмы решения задач оптимизации. – К., 1983.

16. Болтянский В.Г. Математические методы оптимального управления. – М., 1969.

17. Бублик Б.Н., Гаращенко Ф.Г., Кириченко Н.Ф. Структурно-параметрическая оптимизация и устойчивость динамики пучков. – К., 1985.

18. Бублик Б.Н., Данилов В.Я., Наконечный А.Г. Некоторые задачи наблюдения и управления в линейных системах / Учеб. пособие. – К., 1988.

19. Габасов Р., Кириллова Ф. Качественная теория оптимальных процессов. – М., 1971.

20. Гноенский Л.С., Каменский Г.А., Эльсгольц Л.Э. Математические основы теории управляемых систем. – М., 1969.

21. Зайцев Г.Ф., Костюк В.И., Чинаев П.И. Основы автоматического управления и регулирования. – К., 1977.

22. Иванов В.А., Фалдин Н.В. Теория оптимальных систем автома тического регулирования. – М., 1981.

23. Калман Р., Фалб П., Арбиб М. Очерки по математической теории систем. – М., 1971.

24. Кротов В.Ф., Гурман В.И. Методы и задачи оптимального управления. – М., 1973.

25. Ли Э.Б., Маркус Л. Основы теории оптимального управления. – М., 1972.

26. Матвеев Н.М. Методы интегрирования обыкновенных дифференциальных уравнений. – Минск, 1974.

27. Моисеев Н.Н. Математические задачи системного анализа. – М., 1981.

28. Пропой А.И. Элементы теории дискретных оптимальных процессов. – М., 1973.

29. Растригин Л.А. Современные принципы управления сложными сис темами. – М., 1980.

30. Сейдж Э.П., Уайт Ч.С. III. Оптимальное управление системами. – М., 1982.

31. Федоренко Р.П. Приближенное решение задач оптимального управ ления. – М., 1978.

32. Цыпкин Я.З. Основы теории автоматических систем. – М., 1977.

33. Эльсгольц Л.Э. Дифференциальные уравнения и вариационное исчисление. – М., 1969.