4. Метод динамічного програмування у неперервних задачах

4.1 Функція Беллмана

Розглянемо задачу оптимального керування:

$$I = \int_{t_0}^{t_1} f_0\left(t, \bar{x}, \bar{u}\right) dt + g\left(t_1, \bar{x}\left(t_1\right)\right) \to \min, \tag{4.1}$$

$$\frac{d\bar{x}}{dt} = \bar{f}\left(t, \bar{x}, \bar{u}\right), \tag{4.2}$$

$$\bar{x}\left(t_0\right) = \bar{x}_0, \bar{u}\left(t\right) \in U, t \in \left[t_0, t_1\right]. \tag{4.3}$$

У цих рівностях $\bar{f} = \left(f_1\left(t, \bar{x}, \bar{u}\right), f_2\left(t, \bar{x}, \bar{u}\right), ..., f_n\left(t, \bar{x}, \bar{u}\right)\right)$, t_0 – фіксований момент початку керування, t_1 – момент його завершення, який потрібно визначити. Необхідно також визначити керування $\bar{u}$ та фазовий вектор $\bar{x}$, для яких функціонал якості (3.1) досягає мінімуму.

Нехай $M \subset \mathbb{R}^{n+1}$ – замкнена множина, якій належить точка $\left(t_1, \bar{x}\left(t_1\right)\right)$. Позначимо $U_{s,\bar{y}}$ множину допустимих керувань, що переміщують об'єкт керування з положення $\left(s, \bar{y}\right)$ у множину M.

Функцію $B\left(s, \bar{y}\right) = \min\limits_{\bar{u} \in U_{s,\bar{y}}} I\left(t_1, \bar{x}\left(t_1\right)\right)$ називають **функцією Беллмана** задачі оптимального керування (4.1)-(4.3). Якщо $U_{s,\bar{y}} = \varnothing$, то вважають, що $B\left(s, \bar{y}\right) = \infty$.

<u>Теорема 4.1.</u> Для функції Беллмана $B\left(s, \bar{y}\right)$ виконуються наступні властивості:

1) $B\left(s, \bar{y}\right) = g\left(s, \bar{y}\right)$ при $\left(s, \bar{y}\right) \in M$;

2) якщо $\bar{u}(t) \in U_{t_0,\bar{x}_0}$ а $\bar{x}(t)$ – фазова траєкторія що відповідає керуванню $\bar{u}(t)$, то функція Беллмана $B(t,\bar{x}(t))$ є неспадною на відрізку $[t_0, t_1]$ вздовж цієї траєкторії;

3) функція Беллмана є сталою на оптимальній траєкторії $\bar{x}^*(t)$.

Ці властивості функції Беллмана є необхідними умовами оптимальності керування.

<u>Теорема 4.2 (Беллмана).</u> Для того, щоб керування $\bar{u}^*(t) \in U_{t_0,\bar{x}_0}$ та відповідна йому фазова траєкторія $\bar{x}^*(t)$ були оптимальними, необхідно та достатньо, щоб існувала функція Беллмана $B(s,\bar{y}) : \mathbb{R}^{n+1} \to \mathbb{R}$, така, що задовольняє умови:

1) $B(t_1,\bar{x}(t_1)) = g(t_1,\bar{x}(t_1))$;

2) якщо $\bar{u}(t) \in U_{t_0,\bar{x}_0}$, а $\bar{x}(t)$ – фазова траєкторія, що відповідає керуванню $\bar{u}(t)$, то функція $B(t,\bar{x}(t))$ є скінченною та неспадною на $[t_0, t_1]$;

3) $B(t,\bar{x}^*(t)) = \mathrm{const}$, якщо $t \in [t_0, t_1]$.

Нехай функція Беллмана $B(t,\bar{x}(t))$ є неперервно диференційованою. Тоді вона задовольняє нелінійне диференціальне рівняння з частинними похідними:

$$\min_{\bar{u}\in U}\left(\frac{\partial B}{\partial t} + \sum_{i=1}^{n}\frac{\partial B}{\partial x_i}\cdot f_i(t,\bar{x},\bar{u}) + f_0(t,\bar{x},\bar{u})\right) = 0. \tag{4.4}$$

Крайова умова для цього рівняння має вигляд:

$$B(t_1,\bar{x}(t_1)) = g(t_1,\bar{x}(t_1)). \tag{4.5}$$

Рівняння (3.4) називають **рівнянням Беллмана**, а метод розв'язання задач оптимального керування з допомогою функції Беллмана називають **методом динамічного програмування**.

<u>Теорема 4.3 (достатня умова оптимальності у динамічному програмуванні)</u>. Якщо функція Беллмана $B\left(t,\overline{x}\left(t\right)\right)$ є розв'язком рівняння Беллмана (3.4), таким, що $B\left(t_1,\overline{x}\left(t_1\right)\right)=g\left(t_1,\overline{x}\left(t_1\right)\right)$, $\overline{u}^*\left(t\right)\in U_{t_0,\overline{x}_0}$, $\overline{x}^*\left(t\right)$ – фазова траєкторія, що відповідає керуванню $\overline{u}^*\left(t\right)$, для якої виконується рівність

$$\frac{\partial B\left(t,\overline{x}^*\right)}{\partial t}+\sum_{i=1}^{n}\frac{\partial B}{\partial x_i}\cdot f_i\left(t,\overline{x}^*,\overline{u}^*\right)+f_0\left(t,\overline{x}^*,\overline{u}^*\right)=0, \qquad (4.6)$$

то $\overline{u}^*\left(t\right)$ є оптимальним керуванням для задачі (4.1) – (4.3).

<u>Приклад 4.1.</u> Для задачі оптимальної швидкодії $t_1-t_0\to\min$, $\ddot{x}=u$, $|u|\le 1$, $x(0)=1$, $x\left(t_1\right)=\dot{x}\left(t_1\right)=0$ виразити оптимальне керування через функцію Беллмана, що не містить керування у явному вигляді.

Розв'язання. Перейдемо у моделі об'єкта керування до системи двох рівнянь першого порядку:

$$\begin{cases}\dot{x}=y,\\ \dot{y}=u,\end{cases}$$

де $x(0)=y(0)=1$, $x\left(t_1\right)=y\left(t_1\right)=0$, $|u|\le 1$.

Функція Беллмана залежить від трьох змінних: t, x, y: $B = B(t, x, y)$. Запишемо рівняння Беллмана. Оскільки $I = \int_0^{t_1} dt \to \min$, то $f_0 = 1$, $f_1 = y$, $f_2 = u$. Рівняння Беллмана набуває вигляду:

$$\min_{|u|\le 1}\left(\frac{\partial B}{\partial t} + \frac{\partial B}{\partial x} y + \frac{\partial B}{\partial y} u + 1\right) = 0.$$

Оскільки перші два доданки від u не залежать, то останнє рівняння набуває вигляду:

$$\frac{\partial B}{\partial t} + \frac{\partial B}{\partial x} y + \min_{|u|\le 1}\left(\frac{\partial B}{\partial y} u\right) = -1.$$

Значення $\min_{|u|\le 1}\left(\frac{\partial B}{\partial y} u\right)$ дорівнює:

$$\min_{|u|\le 1}\left(\frac{\partial B}{\partial y} u\right) = \begin{cases} -\frac{\partial B}{\partial y}, \frac{\partial B}{\partial y} > 0, u = u^* = -1, \\ \frac{\partial B}{\partial y}, \frac{\partial B}{\partial y} \le 0, u = u^* = -1. \end{cases}$$

Отже, $\min_{|u|\le 1}\left(\frac{\partial B}{\partial y} u\right) = -\left|\frac{\partial B}{\partial y}\right|$. Отримали рівняння Беллмана, що не містить керування u у явному вигляді:

$$\frac{\partial B}{\partial t} + \frac{\partial B}{\partial x} y - \left|\frac{\partial B}{\partial y}\right| = -1.$$

Розв’язавши це рівняння, знаходимо у явному вигляді функцію Беллмана $B = B(t, x, y)$. Далі знаходимо області у просторі $\mathbb{R}^3$, для яких $\dfrac{\partial B}{\partial y}$ зберігає знак, звідки для кожної точки $(t, x, y) \in \mathbb{R}^3$ знаходимо оптимальне керування $u^*(t, x, y)$. Підставивши його у систему диференціальних рівнянь $\begin{cases} \dot{x} = y, \\ \dot{y} = u, \end{cases}$ знаходимо оптимальну фазову траєкторію $\left(x^*(t), y^*(t)\right)$.

Основна складність таких задач полягає у розв’язанні рівняння Беллмана.

У багатьох задачах оптимального керування рівняння Беллмана застосовують у формі, еквівалентній (3.4):

$$\max_{\bar{u} \in U} \left(\frac{\partial B}{\partial t} + \sum_{i=1}^{n} \frac{\partial B}{\partial x_i} \cdot f_i(t, \bar{x}, \bar{u}) - f_0(t, \bar{x}, \bar{u}) \right) = 0. \qquad (3.7)$$

При цьому повинна виконуватися умова

$$B\left(t_1, \bar{x}(t_1)\right) = -g\left(t_1, \bar{x}(t_1)\right), \qquad (3.8)$$

Функція Беллмана, знайдена з рівняння (3.8), відрізняється від розв’язку рівняння (3.4) лише знаком.

4.2 Знаходження оптимального керування з повним зворотним зв’язком

При розв’язанні задачі з повним зворотним зв’язком для знаходження оптимального керування використовують інформацію про час t і всі координати фазового вектора $\bar{x}$. Оптимальне керування $\bar{u}^*\left(t, \bar{x}^*\right)$, що належить множині U допустимих керувань з повним зворотним зв'язком, надає мінімум функціоналу якості

$$I = \int_{t_0}^{t_1} f_0\left(t, \overline{x}, \overline{u}\right) dt + g\left(t_1, \overline{x}\left(t_1\right)\right) \tag{4.9}$$

за умов:

$$\frac{d\overline{x}}{dt} = \overline{f}\left(t, \overline{x}, \overline{u}\right), \tag{4.10}$$

$$\overline{x}\left(0\right) = \overline{x}_0 = \left(x_{10}, x_{20}, ..., x_{n0}\right), \tag{4.11}$$

$$h_i\left(t_1, \overline{x}\left(t_1\right)\right) = 0, i = 1, 2, ..., l, l \le n + 1. \tag{4.12}$$

Достатні умови оптимальності у задачі (4.9) – (4.12) надаються наступною теоремою.

<u>Теорема 3.4.</u> Якщо існує функція Беллмана $B = B\left(t, \overline{x}\right)$, що задовольняє рівняння Беллмана

$$\max_{\overline{u} \in U} \left(\frac{\partial B}{\partial t} + \sum_{i=1}^{n} \frac{\partial B}{\partial x_i} \cdot f_i\left(t, \overline{x}, \overline{u}\right) - f_0\left(t, \overline{x}, \overline{u}\right) \right) = 0 \tag{4.13}$$

та крайову умову

$$B\left(t_1, \overline{x}\left(t_1\right)\right) = -g\left(t_1, \overline{x}\left(t_1\right)\right), \tag{4.14}$$

а керування $\overline{u}^*\left(t, \overline{x}\right) \in U$ таке, що на ньому вираз $\sum_{i=1}^{n} \frac{\partial B}{\partial x_i} f_i\left(t, \overline{x}, \overline{u}\right) - f_0\left(t, \overline{x}, \overline{u}\right)$ досягає максимуму на множині U, то $\overline{u}^*\left(t, \overline{x}\right)$ є оптимальним керуванням з повним зворотним зв'язком і при цьому $\min I = -B\left(t_0, \overline{x}_0\right)$.

Алгоритм знаходження оптимального керування з повним зворотним зв'язком складається з наступних етапів.

1. Записати рівняння Беллмана (4.13) з крайовою умовою (4.14).

2. Визначити структуру оптимального керування $\bar{u}^*(t,\bar{x})$ з повним зворотним зв'язком шляхом знаходження максимуму по $\bar{u}$ виразу $\sum_{i=1}^{n}\frac{\partial B}{\partial x_i} f_i(t,\bar{x},\bar{u}) - f_0(t,\bar{x},\bar{u})$. При цьому $\bar{u}^*$ звичайно виражається через $\frac{\partial B}{\partial x_i}$.

3. Підставити отримане у п.2 керування у рівняння Беллмана та розв'язати це рівняння.

4. За знайденою функцією Беллмана $B(t,\bar{x})$ та її похідним знаходимо оптимальне керування $\bar{u}^*(t,\bar{x})$.

При застосуванні цього алгоритму використовують і іншу форму рівняння Беллмана. Його записують у вигляді (4.4) з крайовою умовою (4.5). Функція Беллмана у цьому випадку відрізнятиметься від розв'язку задачі (4.13), (4.14) лише знаком. При цьому $\min I = B(t_0,\bar{x}_0)$.

<u>**Приклад 3.2.**</u> Модель об'єкта керування має вигляд: $\dot{x} = u$, $t \in [0;1]$, Функціонал якості $I = \frac{1}{2}\int_0^1 u^2 dt + \frac{1}{2}x^2(1) \to \min$. Знайти оптимальне керування $\bar{u}^*(t,\bar{x})$.

Розв'язання. В умовах даного прикладу $f = u$, $f_0 = \frac{1}{2}u^2$, $g(x) = \frac{1}{2}x^2$. Розв'язуємо задачу Больца. Запишемо для неї рівняння Беллмана (3.13) та крайову умову (3.14).

$$\max_{\bar{u}\in U}\left(\frac{\partial B}{\partial t} + \frac{\partial B}{\partial x}u - \frac{1}{2}u^2\right) = 0,\ B(1,x(1)) = -\frac{1}{2}x^2(1).$$

Знаходимо максимум квадратного тричлена відносно змінної u. Оскільки коефіцієнт при u^2 є від'ємним, то точкою його абсолютного максимуму є абсциса вершини параболи $u^* = \frac{\partial B}{\partial x}$. Підставимо знайдене значення u^* у рівняння Беллмана:

$$\frac{\partial B}{\partial t} + \left(\frac{\partial B}{\partial x}\right)^2 - \frac{1}{2}\left(\frac{\partial B}{\partial x}\right)^2 = 0,\ \frac{\partial B}{\partial t} + \frac{1}{2}\left(\frac{\partial B}{\partial x}\right)^2 = 0,\ B(1,x) = -\frac{x^2}{2}.$$

Шукаємо розв'язок цього рівняння у вигляді $B(t,x) = k(t)x^2$. Підставивши цей вираз у рівняння Беллмана, отримуємо:

$$\frac{dk}{dt}x^2 + \frac{1}{2}k^2 \cdot 4x^2 = 0,\ \frac{dk}{dt} = -2k^2,\ k(t) = \frac{1}{2t+C},$$

$$B(1,x) = k(1)\cdot x^2 = -\frac{1}{2}x^2 \Rightarrow k(1) = -\frac{1}{2} \Rightarrow \frac{1}{2+C} = -\frac{1}{2} \Rightarrow C = -4.$$

$$B(t,x) = \frac{x^2}{2(t-2)},\ u^* = \frac{\partial B}{\partial x} = \frac{x}{t-2}.$$

Покажемо, що оптимальне керування $u^* = \frac{x}{t-2}$ з повним зворотним зв'язком породжує оптимальне програмне керування $u^*(t)$ та оптимальну фазову траєкторію $x^*(t)$ для будь-якого початкового стану $x(0) = x_0$. Отримуємо:

$$\dot{x} = u^* = \frac{x}{t-2} \Rightarrow x = C|t-2| = C(2-t),\ t \in [0;1],$$

$$x(0) = 2C = x_0 \Rightarrow C = \frac{x_0}{2},\ x^*(t) = \frac{x_0}{2}(2-t),\ u^* = \frac{x^*}{t-2} = -\frac{x_0}{2}.$$

Визначимо оптимальне програмне керування $u^*(t)$ та оптимальну фазову траєкторію $x^*(t)$, використавши принцип максимуму. Функція Гамільтона має вигляд: $H = \psi u - \frac{u^2}{2}$, $u^* = \psi$, $\dot{x} = \psi, \dot{\psi} = 0 \Rightarrow \psi = A = \text{const}$. $x = At + B$, $x(0) = B = x_0$, $x = At + x_0$.

Умова трансверсальності має вигляд:

$$\delta g(t_1) - H(t_1)\delta t_1 + \psi \cdot \delta x = 0.$$

Оскільки $t_1 = 1 = \text{const}$, $\delta g(t_1) = x(1)\delta x_1$, то з останнього рівняння випливає, що $x(1) + \psi(1) = 0$, тобто $A + x_0 + A = 0$, $A = -\frac{x_0}{2}$, $x^* = x_0 - \frac{x_0}{2}t$, $u^* = \psi = A = -\frac{x_0}{2}$.