5.1. МЕТОД ДИНАМІЧНОГО ПРОГРАМУВАННЯ РОЗВ’ЯЗУВАННЯ ЗАДАЧ ОПТИМАЛЬНОГО КЕРУВАННЯ

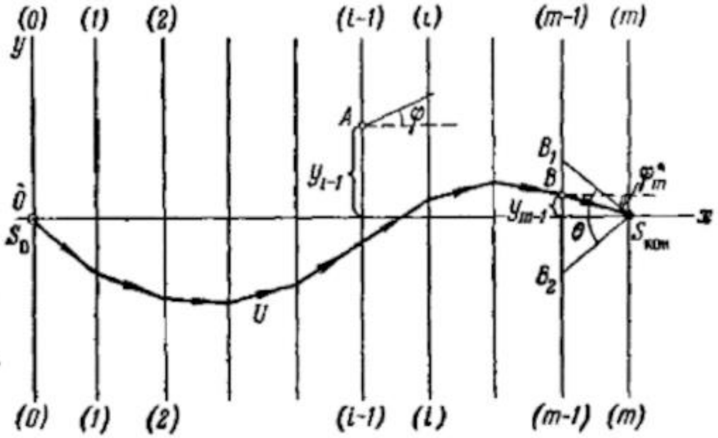
$$U_i^*(y_{i-1}) = \varphi_i^*(y_{i-1})$$

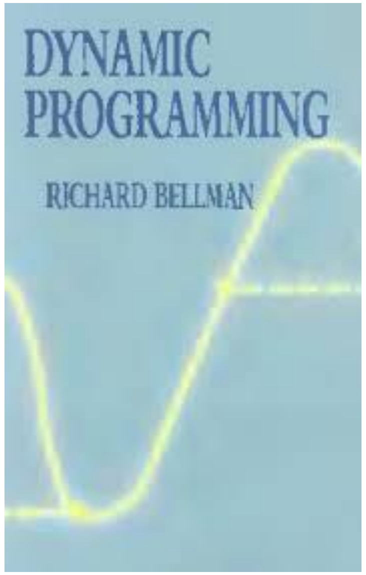
$$\varphi_m^* = \varphi_m^*(y_{m-1})$$

$$T_m^* = T_m^*(y_{m-1})$$

$$\varphi_m^*(y_{m-1}^{(1)}); \quad \varphi_m^*(y_{m-1}^{(2)}); \ \dots$$

$$T_m^*(y_{m-1}^{(1)}); \quad T_m^*(y_{m-1}^{(2)}); \ \dots$$





Розглянемо задачу оптимального керування:
знайти керування та траєкторії, на яких функціонал

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

досягає свого екстремального (мінімального) значення для системи

$$x'(t) = f(x, u, t), \tag{5.2}$$

де

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

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

Тут X – фазовий простір, U – простір керувань, $t \in [t_0, t_1]$.

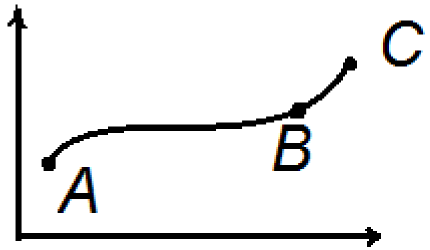
У задачі (5.1) – (5.4) моменти часу t_0, t_1 у загальному випадку вважаються невідомими й підлягають визначенню.

Ці моменти після їх визначення будемо позначати через t_0^0, t_1^0.

Метод динамічного програмування є наслідком принципу оптимальності, який був сформульований Р.Белманом. Принцип оптимальності справедливий для досить широкого класу задач оптимального керування, але не для всіх.

Для задачі (5.1) – (5.4) **принцип оптимальності** може бути сформульований таким чином:

якщо деяка траєкторія $\boldsymbol{AC}$ керованої системи (5.2) є оптимальною траєкторією задачі (5.1) – (5.4), то траєкторія $\boldsymbol{BC}$ також буде оптимальною при будь-якому виборі точки $\boldsymbol{B}$ на оптимальній траєкторії $\boldsymbol{AC}$.



Наведемо інше формулювання принципу оптимальності.

Нехай $\boldsymbol{u^0(t), x^0(t), t_0 \le t \le t_1}$ – розв'язок задачі (5.1) – (5.4),

де $\boldsymbol{u^0(t)}$ – оптимальне керування,

$\boldsymbol{x^0(t)}$ – оптимальна траєкторія,

і нехай $\boldsymbol{t'}$ – довільний фіксований момент часу, $\boldsymbol{t' \in [t_0, t_1]}$.

Тоді розв'язок задачі (5.1) – (5.4) для $\boldsymbol{t \ge t'}$ визначається фіксованим значенням $\boldsymbol{x^0(t')}$ і не залежить від $\boldsymbol{u^0(t), x^0(t)}$ для $\boldsymbol{t < t'}$,

тобто

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

Для задачі (5.1)–(5.4) принцип оптимальності Белмана доводиться на основі властивості адитивності визначеного інтеграла.

Доведення принципу оптимальності можна провести наступним чином.
Нехай:

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

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

Тут t^* – довільна точка з $[t_0^0, t_1^0]$.

Розглянемо задачу (5.1)-(5.4) за умови, що: $t_0^0 = t^*,\ x(t^*) = x^0(t^*)$.

Розв’язок цієї задачі позначимо через $\tilde{u}(t), \tilde{x}(t), \quad t^* \le t \le t_1^0$.

Припустимо, всупереч принципу оптимальності, що цей розв’язок не співпадає з $u^0(t), x^0(t)$ при $t > t^*$.

Тоді

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

Побудуємо допустиме керування для задачі (5.1)-(5.4) у вигляді кусково-неперервної функції

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

Відповідна цьому керуванню траєкторія буде мати вигляд:

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

Для розв'язку $u_*(t), x_*(t)$ задачі (5.1)-(5.4) будемо мати

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

Остання нерівність вказує на те, що розв'язок $u^0(t), x^0(t)$ не є оптимальним, оскільки $u_*(t)$ дає менше значення функціоналу Q.

Протиріччя доводить справедливість принципу оптимальності.

Різницеве Рівняння Белмана для дискретних систем

Наведемо дискретний аналог задачі оптимального керування (5.1)-(5.4).

Розіб'ємо заданий інтервал часу рівномірно точками: $t_0^0 = t_0, t_1, t_2, \ldots t_k, \ldots, t_1^0 = t_N$.

Позначимо
підінтервали часу через $\Delta t_k = \Delta t = t_{k+1} - t_k$,

стан системи в моменти часу t_k через $x(t_k) = x_k, k = \overline{0, N}$,

і керування відповідно $u(t_k) = u_k, k = \overline{0, N-1}$.

Тоді дискретний аналог функціоналу (5.1) буде мати вигляд:

$$Q = Q(x_0, t_0, \{u_k\}_{k=0}^{k=N-1}) = \sum_{k=0}^{N-1} G(x_k, u_k, t_k)\Delta t_k + \Phi(x_N) =$$

$$= \sum_{k=0}^{N-1} F_0(x_k, u_k, t_k) + \Phi(x_N) \longrightarrow \inf \qquad (5.5)$$

Дискретний аналог для системи (5.2) отримаємо наступним чином:

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

звідки

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

або

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

Множини (5.3), (5.4) у випадку дискретного часу будуть мати вигляд, відповідно:

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

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

Для постановки задачі оптимального керування дискретною системою (5.6) припускається, що множини (5.7), (5.8) непорожні та обмежені. Задача оптимального керування (5.5) – (5.8) має сенс лише в тому випадку, коли з точок множини $\Omega_0(X)$ можна перейти в точки множини $\Omega_N(X)$ через точки множин $\Omega_k(X)$, $k=\overline{1, N-1}$.

Визначення 5.1. Множина $\Omega_N(X)$ називається **досяжною** з точок $x_k \in \Omega_k(X)$, $k=\overline{0, N-1}$, якщо існують такі допустимі керування $\{u_j\}$, $j=\overline{k, N-1}$, що відповідна їм згідно з рівнянням (5.6) траєкторія $\{x_j\}$, $j=\overline{k, N}$ з початковою точкою x_k з'єднує цю точку з деякою точкою множини $\Omega_N(X)$.

Якщо множина початкових значень $\Omega_0(X)$ складається не з одного елементу, то задача (5.5) – (5.8) розбивається на дві задачі:

а) знаходження допустимих керувань, які доставляють мінімум функціонала (5.5) при фіксованому значенні $x_0 \in \Omega_0(X)$, тобто

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

б) знаходження мінімуму $Q(x_0, t_0)$ як функції змінної x_0 на множині $\Omega_0(X)$, тобто

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

Для фіксованого моменту часу $t_k, k = \overline{0, N-1}$ введемо деяку функцію $S_k(x_k, t_k)$, яку будемо називати **функцією Белмана**, у вигляді:

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

де $u_j^0(x_k)$, $k = \overline{j, N-1}$, – послідовність керувань, що відповідає оптимальному руху системи (5.6) з деякої точки $x_k \in \Omega_k(X)$, взятої в момент t_k, у точки множини $\Omega_N(X)$.

Виокремо у формулі (5.9) перший член.

Маємо

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

Далі візьмемо $j = k+1$

і для керувань $u_k^0(x_k)$, під дією яких система (5.6) переходить у точку

$$x_{k+1}: \quad x_{k+1} = F_0(x_k, u_k^0(x_k), t_k),$$

розглянемо функцію Белмана

$$S_{k+1}(x_{k+1}, t_{k+1}) = \min_{\{u_k\}_{j=k+1}^{j=N-1}} \{ \sum_{j=k+1}^{N-1} F_0(x_j, u_j, t_j) + \Phi(x_N) \} = \sum_{j=k+1}^{N-1} F_0(x_j, u_j^0(x_{k+1}), t_j) + \Phi(x_N), \tag{5.10}$$

де $u_j^0(x_{k+1})$, $j = \overline{k+1, N-1}$ – послідовність керувань, які відповідають оптимальному руху системи (5.6) із вказаної точки $x_{k+1} \in \Omega_{k+1}(X)$ у точки множини $\Omega_N(X)$.

З принципу оптимальності Белмана випливає, що розв'язок задачі (5.5)-(5.8) на проміжку $[t_k, t_N]$ збігається з розв'язком відповідної задачі на $[t_{k+1}, t_N]$, якщо перехід від x_k до x_{k+1} здійснено згідно з оптимальним керуванням $u_k^0(x_k)$ для системи керування (5.6).

Звідси будуть збігатися керування:

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

Отже, враховуючи це та формулу (5.10), вираз для функції $S_k(x_k, t_k)$ можна записати у вигляді:

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

Оскільки $x_{k+1} = F_0(x_k, u_k^0(x_k), t_k)$, остаточно отримаємо:

$$S_k(x_k, t_k) = \min_{u_k \in \Omega_k(U)} \{F_0(x_k, u_k, t_k) + S_{k+1}(F(x_k, u_k, t_k), t_{k+1})\}, \quad k = \overline{0, N-1}. \tag{5.11}$$

При цьому $\qquad S_N(x_N, t_N) = \Phi(x_N).$

Рівняння (5.11) називається **різницевим рівнянням Белмана.**

5.2. АЛГОРИТМ МЕТОДУ ДИНАМІЧНОГО ПРОГРАМУВАННЯ ДЛЯ ДИСКРЕТНИХ СИСТЕМ

Алгоритм методу динамічного програмування розв'язування задачі вигляду (5.5) – (5.8) для дискретних систем керування складається з двох частин:

знаходження керувань як функцій від станів системи (**прямий хід**)

та обчислення оптимальних керувань і оптимальної траєкторії (**зворотний хід**).

<u>А: Прямий хід.</u>

Крок 1.

Покладемо в рівнянні Белмана (5.11) $k = N-1$ і розв'яжемо 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 всіх точок множини $\Omega_{N-1}(X)$, з яких досяжна множина $\Omega_N(X)$,

тобто для точок $x_N = F(x_{N-1}, u_{N-1}(x_{N-1}), t_{N-1}) \in \Omega_N(X)$.

Знаходимо $u^0_{N-1}(x_{N-1})$ як функцію точок $x_{N-1} \in \Omega_{N-1}(X)$.

Крок 2. Для $k = N-2$ розв'яжемо 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 всіх $x_{N-2} \in \Omega_{N-2}(X)$, з яких досяжна множина $\Omega_N(X)$.

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

……………………….

Продовжуємо далі процес, поки не дійдемо до $k = 0$.

………………………..

Крок N**.** Для $k = 0$ розв'яжемо 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 всіх $x_0 \in \Omega_0(X)$, з яких досяжна множина $\Omega_N(X)$.

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

В: Зворотній хід.

Якщо множина $\Omega_0(X)$ складається більш ніж з одного елементу, то потрібно розв'язати задачу:

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

Знайшовши x_0^0, отримаємо оптимальне керування $u_0^0(x_0^0) = u_0^0$ у момент часу $t = t_0$.

Крок 1. Підставимо знайдені оптимальні x_0^0, $u_0^0(x_0^0)$ у рівняння (5.6):

$$x_1^0 = F(x_0^0, u_0^0, t_0).$$

Знайшли x_1^0 у момент $t = t_1$.

Підставляючи значення x_1^0 у функцію $u_1^0(x_1^0)$, отриману на прямому ході алгоритму, знаходимо оптимальне керування $u_1^0(x_1^0) = u_1^0$.

Продовжуємо цей процес.

……………………………

Крок N. Аналогічно знаходимо керування $u_{N-1}^0(x_{N-1}^0) = u_{N-1}^0$ і точку $x_N^0 = F(x_{N-1}^0, u_{N-1}^0, t_{N-1})$.

Таким чином, знайшли $\{u_j^0\}, j = \overline{0, N-1}, \{x_j^0\}, j = \overline{0, N}$ – оптимальне керування та оптимальну траєкторію для задачі (5.5)-(5.8).