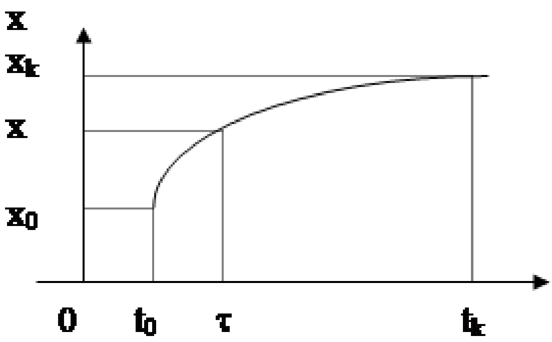
Лекція 13.

Метод динамічного програмування.

Метод динамічного програмування, запропонований на початку 50-х років американським вченим Р. Беллманом, передбачає вирішення поставленої загальної задачі оптимального управління дещо іншим, ніж в принципі максимуму, способом. Грунтується динамічне програмування на принципі оптимальності Беллмана, який можна сформулювати так:

Оптимальне керування визначається тільки кінцевою метою управління та станом системи в даний момент часу, незалежно від того, яким чином система прийшла в цей стан.



Тобто якщо траєкторія x = f (t), tÎ [$t_{0,}$ $t_{k]}$ є оптимальною, то будь-яка її частина x = f (t),

tÎ [t, $t_{k],}$ що приводить до кінцевого стану X_k є також оптимальною траєкторією.

Отже, об'єкт описується звичайною системою диференціальних рівнянь з жорстко заданими початковими умовами

$$\dot{x}=f(x,u),\ x(t_0)=x_0,\ u(t)\in U$$

Потрібно знайти рівняння u = $u_{опт,}$ що забезпечують мінімум функціоналу $I=\int_{t_0}^{t_k} f_0(x,u)dt+g_0[x(t_k)]$. Введемо в розгляд деяку функцію $S\{x(\tau)\}=\min_{u\in U}\left\{\int_{\tau}^{t_k} f_0(x,u)dt+g_0[x(t_k)]\right\}$, (13.1)

певну на оптимальної траєкторії x $_{опт}$ (t), t £ t £ t $_{k.}$ Відмітимо, що $S\{x(\tau)\}$ не залежить від управління u (t), а тільки від x (t) - початку інтервалу оптимізації. Розіб'ємо відрізок часу [t, t $_{k]}$ на два [t, t + D], [t + D, t $_{k]}$ де D- як завгодно мала величина. Тоді (1) представимо у вигляді $$S\{x(\tau)\}=\min_{u\in U}\left\{\int_{\tau}^{\tau+\Delta} f_0(x,u)dt+\int_{\tau+\Delta}^{t_k} f_0(x,u)dt+g_0[x(t_k)]\right\}$$.

Відповідно до принципу оптимальності частина

$$\int_{\tau+\Delta}^{t_k} f_0(x,u)dt+g_0[x(t_k)]$$

приймає на оптимальної траєкторії мінімальне значення, отже, вона може бути позначена як

$S\{x(\tau+\Delta)\}$ отримали:

$$S\{x(\tau)\}=\min_{u\in U}\left\{\int_{\tau}^{\tau+\Delta} f_0(x,u)dt+S[x(\tau+\Delta)]\right\} \quad (13.2)$$

Так як D- як завгодно мала величина, то (метод прямокутників)

$$\int_{\tau}^{\tau+\Delta} f_0(x,u)dt=f_0(x,u)\cdot\Delta$$

$$x(\tau+\Delta)=x(\tau)+\left(\frac{\partial x}{\partial t}\right)_{t=\tau}\cdot\Delta$$

(Ряд Тейлора з першим членом). Врахуємо, що

$$\left(\frac{\partial x}{\partial t}\right)_{t=\tau}=\left(\frac{dx}{dt}\right)_{t=\tau}=f[x(\tau),u(\tau)]$$

тоді

$$S[x(\tau+\Delta)]=S\big[x(\tau)+f[x(\tau),u(\tau)]\Delta\big]\approx S[x(\tau)]+\frac{\partial S}{\partial x}\Big|_{x=x(\tau)}\cdot f[x(\tau),u(\tau)]\cdot\Delta$$

Таким чином,

$$S\{x(\tau)\}=\min_{u\in U}\left\{f_0(x,u)\Delta+S[x(\tau)]+\left(\frac{\partial S}{\partial x}\right)_{x=x(\tau)}\cdot f[x(\tau),u(\tau)]\cdot\Delta\right\}$$

Так як S {x (t)} не залежить від управління u (t), то вона може бути винесена за знак мінімуму, тоді

$$\min_{u\in U}\left\{f_0(x,u)+\left(\frac{\partial S}{\partial x}\right)_{x=x(\tau)} f[x(\tau),u(\tau)]\right\}=0 \quad (13.3)$$

Дане рівняння називається функціональним рівнянням Беллмана, що дозволяє визначати оптимальне рівняння $u_{опт}(t)$. Очевидно, мінімум вираження в квадратних дужках досягається або за умови, що його похідна по управлінню u дорівнює нулю, або сам вираз дорівнює нулю при будь-яких u, тобто

$$\begin{cases} f_0(x,u)+\frac{\partial S}{\partial x}f(x,u)=0, \\ \frac{\partial f_0(x,u)}{\partial u}+\frac{\partial S}{\partial x}\cdot\frac{\partial f(x,u)}{\partial u}=0 \end{cases}$$

виключимо з даної системи $\frac{\partial S}{\partial x}$ і отримаємо $f_0(x,u)\frac{\partial f}{\partial u}=f(x,u)\frac{\partial f_0}{\partial u}$. (13.4)

З рішення даного рівняння і визначається оптимальне рівняння

$u = u_{опт}$

<u>Приклад:</u> дан електродвигун постійного струму з незалежним збудженням $T\dot{\omega}+\omega=ku_y$ де w - кутова швидкість обертання;

u_y - керуюча напруга.

Потрібно визначити оптимальний закон керування u_y^*, Що мінімізує функціонал $I=\int_0^{t_k}(\rho\omega^2+\alpha u_y^2)dt$.

Для даної задачі очевидно $f(x,u)=-\frac{1}{T}\omega+\frac{ku_y}{T}$, $f_0(x,u)=\beta\omega^2+\alpha u_y^2$

Складемо рівняння (13.4)

$$f_0(x,u)\frac{\partial f}{\partial u}=f(x,u)\frac{\partial f_0}{\partial u}\Leftrightarrow(\rho\omega^2+\alpha u_y^2)\frac{K}{T}=2\alpha u_y(-\frac{1}{T}\omega+\frac{K}{T}u_y)$$

$$\rho\frac{K}{T}\omega^2+\frac{\alpha K}{T}u_y^2=-\frac{2\alpha}{T}u_y\omega+\frac{2\alpha K}{T}u_y^2;$$

$$\frac{\alpha K}{T}u_y^2-\frac{2\alpha}{T}u_y\omega-\frac{\beta K}{T}\omega^2=0;$$

$$u_y^2-\frac{2}{K}\omega u_y-\frac{\beta}{\alpha}\omega^2=0$$

отримуємо:

$$u_y=\left(+\frac{1}{K}\pm\sqrt{\frac{1}{K^2}+\frac{\beta}{\alpha}}\right)\omega$$

Як бачимо, управління виходить у вигляді зворотного зв'язку за станом об'єкта. Це одне з основних переваг методу динамічного програмування в порівнянні з принципом максимуму. Однак, для багатовимірних систем і обмежень рішення рівняння (13.4) в явному вигляді отримати складно. У цьому випадку доводиться застосовувати чисельне інтегрування (13.4) в реальному часі.

2. Рівняння Белмана задає необхідну умову мінімуму функціоналу. Якщо функція S(x) є неперервно-диференційована за всіма своїми аргументами, то вона задовольняє рівнянню Белмана.

Для задачі максимальної швидкодії рівняння Белмана має вигляд:

$$-1=\min_{\mathbf{u}\in U}\frac{dS}{d\mathbf{x}}\cdot\mathbf{f}(\mathbf{x},\mathbf{u})$$ . . (13.5)

При t ®¥ на оптимальне керування накладається додаткова умова

асимптотичної стійкості. Якщо функція G > 0 та S [x1 ,...,xn] >0 для всіх x1 ,..., xn ,

то система асимптотично стійка. В цьому випадку функція

[S x ,...,x] є функцією Ляпунова (другий метод Ляпунова) і для асимптотичної

стійкості при t ®¥ та G > 0 достатньо додатньовизначеності функції

[S x ,...,x] . При цьому крайова умова виконується автоматично.

Відмітимо, якщо функціонал має загальний вигляд:

$$I = \int_0^T G(\mathbf{x},\mathbf{u})dt + g(\mathbf{x}(T)),$$

то крайова умова записується так:

S(x(T)) = g(x(T)) (13.6)

Для неавтономного випадку (функції f,G,u явно залежать від t)

рівняння Белмана перепишеться:

$$-\frac{\partial S}{\partial t} = \min_{\mathbf{u}\in U}\left[G(\mathbf{x},\mathbf{u},t) + \frac{\partial S}{\partial \mathbf{x}}\cdot \mathbf{f}(\mathbf{x},\mathbf{u},t)\right]. \quad (13.7)$$

Припустимо, що вдалося знайти вираз керування, при якому вираз в фігурних дужках (або (13.7) досягає мінімуму

$$\mathbf{u} = r(\mathbf{x},t,\mathbf{S}_\mathbf{x}), \quad (13.8)$$

де Sx - вектор з компонентами $\left(\frac{\partial S}{\partial x_1}\frac{\partial S}{\partial x_2}\cdots\frac{\partial S}{\partial x_n}\right)$.

Підставляючи останній вираз в (13.7) отримаємо нелінійне рівняння в частинних похідних першого порядку:

$$-\frac{\partial S}{\partial t} = G(\mathbf{x},\mathbf{u}(\mathbf{x},t,\mathbf{S}_\mathbf{x}),t) + \sum_{i=1}^{n}\frac{\partial S}{\partial x_i}f_i(\mathbf{x},\mathbf{u}(\mathbf{x},t,\mathbf{S}_\mathbf{x}),t). \quad (13.9)$$

Числовий розв’язок цього рівняння з крайовими умовами (13.6) є більш складною задачею чим розв’язок крайової задачі принципу максимуму, бо там мова йшла про крайову задачу для звичайних диференціальних рівнянь, а тут про крайову задачу для рівнянь в частинних похідних. Збільшення складності розв’язку є зрозумілим, тому що в методі динамічного програмування шукається оптимальна стратегія, а не програма.

Для числового розв’язку рівняння (13.9) відомі різні методи, такі як різницеві методи, метод характеристик, метод прямих… Однак є спеціальний метод наближеного обчислення цього рівняння. Суть методу полягає в заміні диференціальних рівнянь (системою диференціально-різнецевих рівнянь, а інтегралу – сумою, таким чином знаходиться оптимальне дискретне керування для дискретних систем.

Приклад.

Розглянемо об'єкт управління, що описується системою диференціальних рівнянь:

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = -x_1 + u. \end{cases}$$

Необхідно знайти управління $u(x_1, x_2)$ таке, щоб мінімізувати Функціонал

$$I = \int_{t_0}^{t_1} (q_{11} x_1^{\ 2} + u^2) dt + \alpha_1 x_1^{\ 2}(t_1) + \alpha_2 x_2^{\ 2}(t_1),$$

(де q_{11}, α_1, α_2 – задані числа (вагові коефіцієнти)) при рухах системи, що збурені початковими відхиленнями. Керування обмежене $|u(t)| \le 1.$.

<u>Розв'язання.</u>

Згідно з функціональне рівняння Белмана запишеться:

$$\min_{|u(t)| \le 1} \left[q_{11} x_1^{\ 2} + u^2 + \frac{\partial S}{\partial x_1} x_2 + \frac{\partial S}{\partial x_2}(-x_1 + u) \right] = 0$$

з крайовими умовами

$$S(x_1(t_1), x_2(t_1)) = \alpha_1 x_1^{\ 2}(t_1) + \alpha_2 x_2^{\ 2}(t_1).$$

Оптимальне керування:

$$\frac{\partial \left[q_{11} x_1^{\ 2} + u^2 + \frac{\partial S}{\partial x_1} x_2 + \frac{\partial S}{\partial x_2}(-x_1 + u) \right]}{\partial u} = 2u + \frac{\partial S}{\partial x_2} = 0 \rightarrow u = -\frac{1}{2} \frac{\partial S}{\partial x_2},$$

тоді:

$$u = \begin{cases} -\frac{1}{2}\frac{\partial S}{\partial x_2}, \text{ якщо } \left|\frac{1}{2}\frac{\partial S}{\partial x_2}\right| < 1, \\ 1, \text{ якщо } -\frac{1}{2}\frac{\partial S}{\partial x_2} > 1, \\ -1, \text{ якщо } -\frac{1}{2}\frac{\partial S}{\partial x_2} < -1 \end{cases}$$

Це співвідношення разом з рівнянням:

$$q_{11}{x_1}^2 + u^2 + \frac{\partial S(x_1, x_2)}{\partial x_1} x_2 + \frac{\partial S(x_1, x_2)}{\partial x_2}(-x_1 + u) = 0$$

та крайовими умовами створюють крайову задачу методу динамічного програмування.

Відмітимо, що існує зв'язок з принципом максимуму. Якщо покласти в $-\frac{\partial S}{\partial x_i}$ - та зробити деякі математичні перетворення, то отримаємо рівняння принципу максимуму , де- $-\frac{\partial S}{\partial x_i}$ є $\psi_i(t)$ $(i = \overline{1, n+1})$ при однакових початкових умовах.