1 ОСНОВНІ ПОНЯТТЯ ВАРІАЦІЙНОГО ЧИСЛЕННЯ

План:

1.1 Функціонал. Варіація функціонала.
1.2 Інтегральні функціонали та їх диференційовність.
1.3 Екстремум функціонала.
1.4 Основні леми варіаційного числення.
1.5 Найпростіша задача варіаційного числення. Рівняння Ейлера.
1.6 Окремі випадки найпростішої задачі варіаційного числення.

Ключові поняття та терміни: інтегральний функціонал, область визначення функціонала, допустима функція, варіаційне числення, лінійний функціонал, неперервний функціонал, варіація функції, близькість функцій, варіація функціонала, сильний (слабкий) екстремум функціонала, абсолютний екстремум, найпростіша задача варіаційного числення, рівняння Ейлера, екстремаль.

1.1 Функціонал. Варіація функціонала

Означення 1.1 Нехай задана деяка функціональна множина M. **Функціоналом** J на множині M називають відображення даної функціональної множини у множину $\mathbb{R}$ дійсних чисел. Функціональну множину M називають **областю визначення функціонала** J, а функції з цієї множини – **допустимими функціями**.

Наведемо приклади функціоналів.

1. На множині M всіх функцій. визначених на $[0;1]$ можна задати функціонал $J(y)=y(0)$.

2. На множині M неперервних та диференційовних на $[a;b]$ функцій можна задати функціонал $J(y)=\int\limits_a^b \sqrt{1+y'^2}\,dx$, що визначає довжину кривої $y(x),\ a\le x\le b$.

Означення 1.2 **Варіаційним численням** називають математичну дисципліну, що вивчає умови, за яких функціонали досягають своїх найбільших та найменших значень.

Далі будемо розглядати функціонали, задані на нормованому просторі $C_{[a;b]}$ функцій $y(x)$, неперервних на $[a;b]$ з нормою $\|y\|_C=\max\limits_{[a;b]}|y(x)|$ та нормованому просторі $C^1_{[a;b]}$ неперервних та диференційовних на $[a;b]$ з нормою $\|y\|_{C^1}=\max\limits_{[a;b]}|y(x)|+\max\limits_{[a;b]}|y'(x)|$.

При дослідженні функціоналів вводять поняття, аналогічні відповідним поняттям для функцій: неперервність, диференційовність, екстремум тощо.

Означення 1.3 Функціонал $J(y)$, визначений на множині M, називають **лінійним**, якщо $\forall y_1, y_2 \in M, \forall \alpha_1, \alpha_2 \in R \;\; J(\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 J(y_1) + \alpha_2 J(y_2)$.

Наприклад, функціонал $J(y) = \int\limits_a^b y(x)\,dx$ є лінійним, оскільки маємо:

$$J(\alpha_1 y_1 + \alpha_2 y_2) = \int\limits_a^b (\alpha_1 y_1 + \alpha_2 y_2)\,dx = \alpha_1 \int\limits_a^b y_1 dx + \alpha_2 \int\limits_a^b y_2 dx = \alpha_1 J(y_1) + \alpha_2 J(y_2).$$

Функції з нормованого простору, що входять до області визначення деякого функціонала, будемо називати її точками. У нормованому просторі можна ввести поняття околу точки та неперервності функціонала.

Означення 1.4 Число $\|y_1 - y_2\|$ називають **відстанню між точками** y_1 та y_2 **нормованого простору** M.

Означення 1.5 Нехай $\varepsilon \in R, \varepsilon > 0$. ε**-околом** точки y_0 нормованого простору M називають множину $O_\varepsilon(y_0) = \{y \in N : \|y - y_0\| < \varepsilon\}$.

Якщо у області визначення функціонала розглядається кілька норм, то кожній з них відповідає власний ε-окіл фіксованої точки y_0 з області визначення, наприклад, для функцій з простору $C^1_{[a;b]}$ можна розглядати норму простору $C_{[a;b]}$.

Означення 1.6 **Сильним** ε**-околом** точки $y_0 \in C^1_{[a;b]}$ називають множину функцій

$$\left\{ y \in C^1_{[a;b]} : \max_{[a;b]} \left(|y - y_0| \right) = \|y - y_0\|_C < \varepsilon \right\}.$$

Слабким ε**-околом** цієї ж точки називають множину функцій

$$\left\{ y \in C^1_{[a;b]} : \max_{[a;b]} |y - y_0| + \max_{[a;b]} |y' - y_0'| = \|y - y_0\|_{C^1} < \varepsilon \right\}.$$

З означення 1.6 випливає, що до сильного ε-околу точки $y_0 \in C^1_{[a;b]}$ належать усі точки даного простору, відстань яких від точки y_0, обчислена за нормою простору $C_{[a;b]}$, є меншою ε. Слабкий ε-окіл точки y_0 утворюють точки простору $C^1_{[a;b]}$, відстань яких від точки y_0, обчислена за нормою простору $C^1{}_{[a;b]}$, є меншою за ε.

Таким чином, сильний ε-окіл визначається з точки зору близькості функцій за нормою простору $C_{[a;b]}$, слабкий – за нормою простору $C^1_{[a;b]}$. У першому випадку говорять про ε- **близькість нульового порядку та** ε**-окіл нульового порядку**, у другому – про ε- **близькість першого порядку та** ε**-окіл першого порядку**.

Аналогічно можна визначити ε-близькість довільного, k-го порядку як близькість за нормою простору $C^k_{[a;b]}: \|y\|_{C^k} = \sum_{m=1}^{k} \max\left|y^{(k)}(x)\right|$.

З означень 1.5 та 1.6 випливає, що функція з слабкого ε-околу точки y_0 завжди знаходиться і в її сильному ε-околі.

Означення 1.7 Функціонал $J(y)$, визначений на нормованому просторі M, називають **неперервним** у точці $y_0 \in M$, якщо для довільного $\delta > 0$ існує такий ε-окіл точки y_0, що для довільного y з цього ε-околу виконується нерівність $\left|J(y) - J(y_0)\right| < \delta$.

Означення 1.8 У нормованому функціональному просторі M виберемо деяку функцію $y_0(x)$. Різницю $\delta y(x) = y(x) - y_0(x)$ між довільною функцією $y(x) \in M$ та $y_0(x)$ називають **варіацією функції** $y_0(x)$.

Розглянемо приріст функціонала $J(y)$, визначеного на нормованому просторі $C^1_{[a;b]}$, у точці y, який відповідає варіації його аргументу δy:

$$\Delta J(y) = J(y + \delta y) - J(y). \tag{1.1}$$

Нехай його можна представити у вигляді:

$$\Delta J(y) = J(y + \delta y) - J(y) = J_1(y, \delta y) + o(\delta y), \tag{1.2}$$

де $J_1(y, \delta y)$ – функціонал, лінійний відносно δy, а $o(\delta y)$ – функціонал, що є нескінченно малою більш високого порядку у порівнянні з $\|\delta y\|$ при $\|\delta y\| \to 0$. Тут маємо на увазі норму у просторі $C^1_{[a;b]}$.

Означення 1.9 Якщо приріст функціонала $J(y)$ можна представити у вигляді (1.2), то цей функціонал називають **диференційовним** у точці y, а лінійний відносно δy функціонал $J_1(y, \delta y)$ називають **диференціалом Фреше** або **сильним диференціалом**.

Поняття диференційовності функціонала за змістом є аналогічним поняттю диференційовності функції.

Означення 1.10 **Варіацією (першою варіацією)** $\delta J(y, \delta y)$ функціонала $J(y)$ у точці y називають границю

$$\delta J(y, \delta y) = \lim_{\alpha \to 0} \frac{J(y + \alpha \delta y) - J(y)}{\alpha} = \frac{d\left[J(y + \alpha \delta y)\right]}{d\alpha}\Big|_{\alpha=0}. \tag{1.3}$$

Ця границя є функціоналом, який кожній варіації δy при фіксованому y ставить у відповідність деяке число δJ.

Означення 1.11 Якщо функціонал $\delta J(y, \delta y)$ (1.3) є лінійним відносно δy, то його називають **диференціалом Гато** або **слабким диференціалом**.

Теорема 1.1 Якщо функціонал $J(y)$ є диференційовним у точці y, то його диференціал Гато у цій точці існує та співпадає з диференціалом Фреше.

Доведення. Доведемо, що з існування диференціала Фреше функціонала $J(y)$ у точці y випливає існування у цій точці функціонала Гато, причому ці диференціали співпадають.

Виберемо деяку варіацію δy у точці y та обчислимо границю:

$$\delta J(y,\delta y)=\lim_{\alpha\to 0}\frac{J(y+\alpha\cdot\delta y)-J(y)}{\alpha}=\lim_{\alpha\to 0}\frac{J_1(y,\alpha\cdot\delta y)+o(\alpha\cdot\delta y)}{\alpha}=$$

$$=\lim_{\alpha\to 0}\frac{\alpha\cdot J_1(y,\delta y)}{\alpha}+0=J_1(y,\delta y).$$

Тут друга рівність виконується, оскільки функціонал $J(y)$ є диференційовним, третя рівність має місце в силу лінійності функціонала J_1 відносно δy. При цьому отримуємо $\lim\limits_{\alpha\to 0}\frac{o(\alpha\cdot\delta y)}{\alpha}=0$, оскільки

$$\lim_{\alpha\to 0}\left|\frac{o(\alpha\cdot\delta y)}{\alpha}\right|=\|\delta y\|\cdot\lim_{\alpha\to 0}\frac{|o(\alpha\cdot\delta y)|}{\|\alpha\cdot\delta y\|}=\|\delta y\|\cdot 0=0.$$

Доведена рівність свідчить про те, що перша варіація $\delta J(y,\delta y)$ диференційовного функціонала є функціоналом, лінійним відносно варіації функції δy, тобто є диференціалом Гато, який співпадає з диференціалом Фреше.

Твердження, обернене даній теоремі, є невірним: диференціал Гато може існувати і у недиференційовного функціонала.

Доведена теорема визначає коректність наступного означення варіації функціонала, що є альтернативним означенню 1.10.

Означення 1.12 Головну, лінійну по відношенню до варіації функції δy, частину приросту диференційовного функціонала $J(y)$, називають його **варіацією (першою варіацією)** $\delta J(y,\delta y)$.

1.2 Інтегральні функціонали та їх диференційовність

Основним об'єктом дослідження у варіаційному численні є функціонали, задані у вигляді інтегралів. Їх називають **інтегральними функціоналами**. Прикладами інтегральних функціоналів є:

$$J_1=\int_a^b F(x,y,y')dx,\; J_2=\int_a^b F\left(x,y,y',\ldots,y^{(n)}\right)dx,$$

$$J_3=\iint_D F\left(x,y,z(x,y),z'_x(x,y),z'_y(x,y)\right)dxdy.$$

З'ясуємо достатні умови існування сильного та слабкого диференціалів у інтегрального функціонала

$$J=\int_a^b F\left(x,y,y'\right)dx\,, \tag{1.4}$$

заданого на нормованому просторі $C^1_{[a;b]}$.

Нехай F є двічі диференційовною функцією своїх аргументів. Знайдемо приріст функціонала (1.4) на деякій функції $y(x)\in C^1_{[a;b]}$:

$$\Delta J=J\left(y+\delta y\right)-J\left(y\right)=\int_a^b\left[F\left(x,y+\delta y,y'+\delta y'\right)-F\left(x,y,y'\right)\right]dx\,.$$

Застосувавши до підінтегральної функції формулу Тейлора, отримуємо наступний вираз для приросту функціоналу:

$$\Delta J=\int_a^b\left[F'_y\left(x,y,y'\right)\delta y+F'_{y'}\left(x,y,y'\right)\delta y'\right]+\frac{1}{2}\int_a^b F''_{yy}\left(x,y+\theta\delta y,y'+\theta\delta y'\right)\cdot\left(\delta y\right)^2dx+$$

$$+\int_a^b F''_{y'y}\left(x,y+\theta\delta y,y'+\theta\delta y'\right)\cdot\delta y\cdot\delta y'dx+\frac{1}{2}\int_a^b F''_{y'y'}\left(x,y+\theta\delta y,y'+\theta\delta y'\right)\left(\delta y'\right)^2dx,$$

де $\theta\in\left[0;1\right]$.

Оцінимо останні три доданки у правій частині цієї рівності.

$$\left|\frac{1}{2}\int_a^b F''_{yy}\left(x,y+\theta\delta y,y'+\theta\delta y'\right)\cdot\left(\delta y\right)^2dx+\int_a^b F''_{y'y}\left(x,y+\theta\delta y,y'+\theta\delta y'\right)\cdot\delta y\cdot\delta y'dx+\right.$$

$$\left.+\frac{1}{2}\int_a^b F''_{y'y'}\left(x,y+\theta\delta y,y'+\theta\delta y'\right)\left(\delta y'\right)^2dx\right|\le\frac{A}{2}\int_a^b\left[\left(\delta y\right)^2+2\left|\delta y\right|\cdot\left|\delta y'\right|+\left(\delta y'\right)^2\right]dx\le$$

$$\le 2A\cdot\left(b-a\right)\left\|\delta y\right\|^2=o\left(\delta y\right).$$

У цих нерівностях $\left\|\delta y\right\|$ є нормою простору $C^1_{[a;b]}$, $A=\max\limits_{[a;b]}\left\{F''_{yy},F''_{yy'},F''_{y'y'}\right\}$.

Таким чином, отримуємо:

$$\Delta J=\int_a^b\left[F'_y\left(x,y,y'\right)\delta y+F'_{y'}\left(x,y,y'\right)\delta y'\right]dx+o\left(\delta y\right). \tag{1.5}$$

Виконаємо інтегрування частинами другого доданка у підінтегральному виразі в (1.5), враховуючи при цьому, що $\delta y'dx=d\left(\delta y\right)$. Отримаємо:

$$\int_a^b F'_y\left(x,y,y'\right)\cdot\delta y\cdot dx+\int_a^b F'_{y'}\left(x,y,y'\right)\cdot\delta y'\cdot dx=\int_a^b F'_y\left(x,y,y'\right)\cdot\delta y\cdot dx+$$

$$+F'_{y'}\left(x,y,y'\right)\cdot\delta y\Big|_a^b-\int_a^b\frac{d}{dx}\left[F'_{y'}\left(x,y,y'\right)\right]\cdot\delta y\cdot\delta x.$$

Підставивши цей результат у (1.5), маємо

$$\Delta J=\int_a^b\left[F'_y\left(x,y,y'\right)-\frac{d}{dx}\left(F'_{y'}\left(x,y,y'\right)\right)\right]\cdot\delta y\cdot dx+F'_{y'}\left(x,y,y'\right)\cdot\delta y\Big|_a^b+o\left(\delta y\right).$$

Перші два доданки у правій частині цієї рівності є функціоналом, лінійним відносно варіації функції δy, $o(\delta y)$ – нескінченно мала вищого порядку по відношенню до δy при $\delta y \to 0$.

Таким чином, диференційовність інтегрального функціонала (1.4) доведено і знайдено його диференціал Фреше. Цей диференціал можна записати у вигляді:

$$J_1(y,\delta y) = \int_a^b \left[F'_y(x,y,y') - \frac{d}{dx}\left(F'_{y'}(x,y,y')\right)\right]\cdot \delta y\, dx + F'_{y'}(x,y,y')\cdot \delta y\Big|_a^b .$$

При цьому для існування диференціала Фреше функціонала (1.4) достатньо, щоб його підінтегральна функція $F(x,y,y')$ була двічі неперервно диференційовною функцією своїх аргументів.

Для існування диференціала Гато достатніми є умови неперервності функції $F(x,y,y')$ та її частинних похідних F'_y та $F'_{y'}$.

Знайдемо за означенням варіацію функціонала (1.4):

$$\delta J(y,\delta y) = \frac{d}{d\alpha}\left[\int_a^b F(x, y+\alpha\delta y, y'+\alpha\delta y')\right]dx\Bigg|_{\alpha=0} = \int_a^b \left(F'_y\delta y + F'_{y'}\delta y'\right)dx .$$

Знайдена варіація функціонала (1.4) (диференціал Гато) співпадає з отриманим раніше диференціалом Фреше, тобто

$$\delta J(y,\delta y) = \int_a^b \left[F'_y(x,y,y') - \frac{d}{dx}\left(F'_{y'}(x,y,y')\right)\right]\cdot \delta y\, dx + F'_{y'}(x,y,y')\cdot \delta y\Big|_a^b .$$

1.3 Екстремум функціонала

<u>Означення 1.13</u> Функціонал $J(y)$, визначений на просторі $C^1_{[a;b]}$, досягає **сильного (слабкого) мінімуму** на функції $y_0(x)\in C^1_{[a;b]}$, якщо знайдеться такий сильний (слабкий) ε-окіл функції $y_0(x)$, що для довільної функції $y(x)$ з цього ε-околу виконується нерівність $J(y)\ge J(y_0)$. Якщо дана нерівність є строгою, то мінімум називають **строгим**.

Аналогічним чином визначають **сильний (слабкий) максимум**. У цьому випадку повинна виконуватися нерівність $J(y)\le J(y_0)$. Сильні (слабкі) мінімуми та максимуми об’єднують під загальною назвою **сильний (слабкий) екстремум**.

Функцію $y_0(x)$, що надає сильний (слабкий) екстремум функціоналу $J(y)$, називають **точкою сильного (слабкого) екстремуму** даного функціонала.

<u>Означення 1.14</u> Функціонал $J(y)$, визначений на деякій функціональній множині M, досягає на функції $y_0\in M$ **глобального або абсолютного**

мінімуму, якщо $J(y_0) \le J(y)\ \forall y \in M$. Для **глобального (абсолютного) максимуму** у останній нерівності знак змінюється на протилежний.

Оскільки будь-яка точка з слабкого ε-околу функції $y_0(x)$ одночасно належить її сильному ε-околу, то будь-який сильний екстремум функціонала одночасно є також його слабким екстремумом, проте слабкий екстремум функціонала не обов'язково є його сильним екстремумом. Це пояснюється тим, що функції, близькі за своїми значеннями (такі, що належать сильному ε-околу), можуть мати значні відхилення у значеннях їх похідних, що, в свою чергу, може вплинути на значення функціонала. З цього випливає, що будь-яка необхідна умова слабкого екстремуму функціонала є одночасно необхідною умовою його сильного екстремуму, а будь-яка достатня умова сильного екстремуму функціонала є одночасно достатньою умовою його слабкого екстремуму.

<u>Теорема 1.2</u> (**Необхідна умова слабкого екстремуму функціонала**). Якщо функціонал $J(y)$ досягає слабкого екстремуму у внутрішній точці $y_0(x)$ своєї області визначення, причому у цій точці існує його варіація, то вона дорівнює нулю при $y = y_0(x)$.

Доведення. Нехай функціонал $J(y)$ досягає слабкого екстремуму у внутрішній точці $y_0(x)$ своєї області визначення. При фіксованих $y_0(x)$ та δy будемо обчислювати його значення на кривих, що мають вигляд $y = y_0(x) + \alpha \cdot \delta y$. На цій множині функцій функціонал можна розглядати як функцію аргументу α: $g(\alpha) = J(y_0 + \alpha \cdot \delta y)$.

За умовою ця функція досягає екстремуму при $y = y_0(x)$, тобто при $\alpha = 0$. Тому при $\alpha = 0$ повинна виконуватися необхідна умова екстремуму для функції $g(\alpha)$:

$$g'(0) = \frac{d\left[J(y + \alpha\delta y)\right]}{d\alpha}\Big|_{\alpha=0} = 0.$$

Враховуючи означення 1.10 варіації функціонала $J(y)$, згідно з яким

$$\delta J(y, \delta y) = \frac{d\left[J(y + \alpha\delta y)\right]}{d\alpha}\Big|_{\alpha=0},$$

отримуємо при $y = y_0(x)$ рівність $\delta J(y_0) = 0$, що й потрібно було довести.

1.4 Основні леми варіаційного числення

Сформулюємо та доведемо леми, що у подальшому будуть використані для отримання необхідних умов екстремуму інтегральних функціоналів.

<u>Лема 1.1</u> (**Лема Лагранжа**). Якщо функція $f(x)$ є неперервною на $[a; b]$

та для довільної нескінченно диференційовної функції $\eta(x)$, такої, що на межах відрізка $[a;b]$ $\eta(a)=\eta(b)=0$, виконується рівність

$$\int_a^b f(x)\cdot\eta(x)dx=0, \tag{1.6}$$

то $f(x)\equiv 0$ на відрізку $[a;b]$.

Доведення. Нехай $x_0\in[a;b]$ та $f(x_0)\neq 0$. Для визначеності приймемо $f(x_0)>0$.Оскільки $f(x)$ є неперервною, то можна вибрати деякий проміжок $(c;d)\subset(a;b)$, що містить точку x_0, на якому $f(x)>0$.

Розглянемо функцію

$$\varphi(x)=\begin{cases} e^{-1/x}, x>0, \\ 0, x\le 0. \end{cases}$$

Можна показати, що дана функція має похідні довільного порядку у кожній точці числової прямої, тому функція $\eta(x)=\varphi(x-c)\cdot\varphi(d-x)$ також є нескінченно диференційовною і при цьому не дорівнює нулю лише на $(c;d)$, тобто для неї виконана умова $\eta(a)=\eta(b)=0$. Для вибраної таким чином $\eta(x)$ отримуємо:

$$\int_a^b f(x)\cdot\eta(x)dx=\int_c^d f(x)\cdot\eta(x)dx>0,$$

оскільки тут підінтегральна функція є неперервною та додатною на проміжку інтегрування $(c;d)$. Отримане протиріччя з (1.11) доводить лему.

Лема 1.1 може бути узагальненою і на випадок функцій багатьох змінних. Наприклад, у двовимірному випадку є справедливим наступне твердження:

Лема 1.2 Якщо функція $f(x,y)$ двох дійсних змінних є неперервною у обмеженій області G декартової площини з межею L, та для довільної функції $\eta(x,y)$, нескінченно диференційовної у G, неперервної у $G\bigcup L$, що дорівнює нулю на L, виконується рівність $\iint\limits_G f(x,y)\cdot\eta(x,y)dxdy=0$, то $f(x,y)\equiv 0$ у області G.

Лема 1.3 (**Лема Дюбуа – Раймона**). Нехай функції $f(x)$ та $g(x)$ є неперервними на відрізку $[a;b]$ та для довільної нескінченно диференційовної на $[a;b]$ функції $\eta(x)$, такої, що $\eta(a)=\eta(b)=0$, виконується рівність

$$\int_a^b\left[f(x)\cdot\eta'(x)+g(x)\cdot\eta(x)\right]dx=0.$$

Тоді функція $f(x)$ є неперервно диференційовною на $[a;b]$ і при цьому

$$f'(x)-g(x)\equiv 0, x\in[a;b]. \tag{1.7}$$

Доведення. Оскільки функція $g(x)$ є неперервною на $[a;b]$, то вона має на цьому відрізку первісну $G(x)$, що визначається з точністю до довільної сталої, тобто $G(x)=G_0(x)+C$, де $G_0(x)$ – деяка фіксована первісна функції $g(x)$. Виберемо сталу C таким чином, щоб виконувалася рівність $\int\limits_a^b \left[f(x)-G(x)\right]dx=0$.

Підставивши сюди вираз для $G(x)$, знаходимо значення сталої C:

$$\int\limits_a^b \left[f(x)-G(x)\right]dx=0 \Rightarrow \int\limits_a^b \left[f(x)-G_0(x)\right]dx-\int\limits_a^b Cdx=0 \Rightarrow C=$$

$$=\frac{1}{b-a}\int\limits_a^b \left[f(x)-G_0(x)\right]dx.$$

Для довільної функції $\eta(x)$, що задовольняє умовам леми, за формулою інтегрування частинами отримуємо:

$$\int\limits_a^b g(x)\cdot\eta(x)dx=\int\limits_a^b G'(x)\cdot\eta(x)dx=\left[G(x)\cdot\eta(x)\right]\Big|_a^b-\int\limits_a^b G(x)\cdot\eta'(x)dx=$$

$$=-\int\limits_a^b G(x)\cdot\eta'(x)dx.$$

Звідси випливає:

$$\int\limits_a^b \left[f(x)\cdot\eta'(x)+g(x)\cdot\eta(x)\right]dx=0 \Rightarrow \int\limits_a^b \left[f(x)-G(x)\right]\cdot\eta'(x)dx=0.$$

Виберемо довільну функцію $\eta(x)$, що задовольняє умовам леми. Введемо позначення $C_\eta=\frac{1}{b-a}\int\limits_a^b \eta(x)dx$. Функція $\xi(x)=\int\limits_a^x \left[\eta(t)-C_\eta\right]dt$ є нескінченно диференційовною і при цьому $\xi(a)=\xi(b)=0$. Згідно умов леми для неї виконується рівність (1.13), тому виконується і рівність (1.16). Отже, з врахуванням співвідношення $\xi'(x)=\eta(x)-C_\eta$, отримуємо:

$$\int\limits_a^b \left[f(x)-G(x)\right]\cdot\eta(x)dx-C_\eta\cdot\int\limits_a^b \left[f(x)-G(x)\right]dx=0.$$

Оскільки другий інтеграл у лівій частині останньої рівності дорівнює нулю, то для довільної $\eta(x)$ виконується рівність:

$$\int\limits_a^b \left[f(x)-G(x)\right]\cdot\eta(x)dx=0.$$

Використовуючи лему 1.1, отримаємо $f(x)-G(x)\equiv 0,\ x\in[a;b]$. Функція $G(x)$ є диференційовною і $G'(x)=g(x)$. Тому $f(x)$ також є диференційовною та $f'(x) = g(x)$. З останньої рівності та неперервності функції $g(x)$ випливає твердження леми.

1.5 Найпростіша задача варіаційного числення. Рівняння Ейлера

Розглянемо задачу про знаходження екстремуму функціонала

$$J(y) = \int_a^b F(x, y, y')dx, \qquad (1.8)$$

визначеного на множині функцій $y(x) \in C^1_{[a;b]}$, що задовольняють умовам:

$$y(a) = y_1, \quad y(b) = y_2. \qquad (1.9)$$

У функціоналі (1.8) $F(x, y, y')$ будемо вважати двічі неперервно диференційовною функцією своїх аргументів.

Сформульовану задачу називають **найпростішою задачею варіаційного числення.**

Як показано у п. 1.2, варіацію функціонала (1.8) можна записати у вигляді:

$$\delta J(y, \delta y) = \int_a^b \left(F'_{y'} \delta y + F'_{y'} \delta y' \right) dx. \qquad (1.10)$$

Тут δy – варіація допустимої функції $y(x)$, $\delta y'$ – похідна цієї варіації.

Оскільки умови (1.9) фіксують значення допустимих функцій на межах відрізка інтегрування $[a; b]$, тому варіація δy функції $y(x)$ у цих точках повинна дорівнювати нулю: $\delta y|_{x=a} = \delta y|_{x=b} = 0$.

Необхідна умова екстремуму у найпростішій задачі варіаційного числення визначається наступною теоремою.

<u>Теорема 1.3</u> Для того, щоб функція $y_0(x)$ надавала слабкий екстремум функціоналу (1.8), необхідно, щоб вона задовольняла диференціальному рівнянню

$$F_y - \frac{d}{dx}\left(F_{y'}\right) = 0. \qquad (1.11)$$

Це рівняння називають **рівнянням Ейлера** для функціонала (1.8). У загальному випадку воно є звичайним диференціальним рівнянням другого порядку відносно функції $y(x)$.

Доведення. Використавши необхідну умову слабкого екстремуму диференційовного функціонала – рівність нулю його варіації (теорема 1.2), а також вираз (1.10) для варіації інтегрального функціонала, отримаємо:

$$\delta J(y, \delta y) = \int_a^b \left(f'_{y'} \delta y + f'_{y'} \delta y' \right) dx = 0. \qquad (1.12)$$

Дане співвідношення виконується для довільної варіації $\delta y \in C^1_{[a;b]}$, що задовольняє крайовим умовам $\delta y|_{x=a} = \delta y|_{x=b} = 0$, у тому числі воно виконується і для довільної, нескінченне число разів диференційовної функції δy, що задовольняє цим умовам. Тому, згідно з лемою Дюбуа – Раймона 1.3 для довільного $x \in [a; b]$ виконується рівність (1.11), тобто крива $y_0(x)$ що надає слабкий екстремум функціоналу (1.8), повинна задовольняти рівнянню Ейлера для даного функціонала.

Рівняння Ейлера (1.11) відображає також необхідну умову сильного екстремуму для функціонала (1.8).

Означення 1.15 Розв'язки рівняння Ейлера для даного функціонала називають **екстремалями** цього функціонала.

Оскільки умова (1.11) є необхідною, то точки екстремуму функціонала слід шукати серед його екстремалей.

Приклад. Знайти екстремалі функціонала $J(y) = \int_0^{\pi/2} \left(y'^2 - y^2 \right) dx$, що задовольняють крайові умови $y(0) = 0,\ y(\pi/2) = 1$ [2].

Розв'язання. Оскільки $F(x, y, y') = y'^2 - y^2$, то $F_y = -2y$, $F_{y'} = 2y'$, $\frac{d}{dx} F_{y'} = 2y''$. Рівняння Ейлера має вигляд $y'' + y = 0$. Його загальний розв'язок $y(x) = C_1 \cos x + C_2 \sin x$. З граничних умов обчислюємо $C_1 = 0,\quad C_2 = 1$. Отже, функція $y(x) = \sin x$ єдина допустима екстремаль.

1.6 Окремі випадки найпростішої задачі варіаційного числення

Розглянемо розв'язання рівняння Ейлера для деяких типів підінтегральних виразів у функціоналі (1.8).

1. Нехай F не залежить від y', тобто $F = F(x, y)$. У цьому випадку рівняння Ейлера набуває вигляду $F_y = 0$, тобто є скінченним рівнянням відносно $y(x)$. Розв'язки цього рівняння, якщо вони існують, можуть не задовольняти заданим крайовим умовам для функції $y(x)$.

Приклад. Знайти екстремалі функціонала $J(y) = \int_a^b y^3 dx$, що задовольняють крайові умови $y(a) = y_1, y(b) = y_2$.

Розв'язання. Для заданого функціонала $F = y^3, F_y = 3y^2, F_{y'} = 0$. Отримуємо рівняння Ейлера $3y^2 = 0$, звідки знаходимо рівняння екстремалі $y = 0$. Якщо хоча б одне з чисел y_1 або y_2 є відмінним від нуля, то у просторі $C^1_{[a;b]}$ функціонал не має екстремалей, які б задовольняли заданим крайовим умовам.

2. Підінтегральний вираз функціонала лінійно залежить від y', тобто $F(x, y, y') = P(x, y) + Q(x, y) \cdot y'$. У цьому випадку рівняння Ейлера набуває вигляду $P_y + Q_y \cdot y' - Q_x - Q_y \cdot y' = 0$, або $P_y = Q_x$. Це рівняння є скінченним, а не диференціальним, тому його розв'язок може і не задовольняти заданим крайовим умовам. Якщо вираз $P(x, y)dx + Q(x, y)dy$ є повним диференціалом, то рівняння $P_y = Q_x$ є тотожністю. У цьому випадку будь-яка функція $y(x) \in C^1_{[a;b]}$ є розв'язком даного рівняння, тобто екстремаллю функціонала.

Приклад. Знайти екстремалі функціонала $J(y) = \int\limits_a^b \left(y^2 + yy' \right) dx$, що задовольняють крайовим умовам $y(a) = y_1$, $y(b) = y_2$.

Розв'язання. Рівняння Ейлера для даного функціоналу набуває вигляду $2y = 0$, тобто екстремаллю є $y = 0$. Якщо хоча б одне з чисел y_1 або y_2 не дорівнює нулю, то задача не має розв'язків.

3. Підінтегральний вираз функціонала (1.8) залежить лише від y', тобто маємо $F = F(y')$. Оскільки $F_y = 0$, то з рівняння Ейлера для даного функціонала $\frac{d}{dx}\left(F_{y'} \right) = 0$ отримуємо $F_{y'} = const$ – скінченне рівняння відносно y', всі розв'язки якого можна записати у вигляді $y' = C_1$, де C_1 – довільна стала. Таким чином, у даному випадку екстремалями функціонала є сімейство прямих $y = C_1 x + C_2$. Сталі C_1 та C_2 визначаються з заданих крайових умов.

Приклад. Знайти екстремалі функціонала $J(y) = \int\limits_0^1 \left(y'^2 + 2\sqrt{y'} \right) dx$, що задовольняють крайові умови $y(0) = 0$, $y(1) = 2$.

Розв'язання. Оскільки підінтегральний вираз заданого функціонала залежить лише від похідної y', то екстремалі даного функціонала є лінійними функціями: $y = C_1 x + C_2$. $y(0) = C_2 = 0,\quad y(1) = C_1 + C_2 = 2 \Rightarrow C_1 = 2$. Отримуємо екстремаль $y = 2x$.

4. Підінтегральний вираз функціонала (1.8) не залежить від y, тобто $F = F(x, y')$. Отримуємо рівняння Ейлера $\frac{d}{dx}\left(F_{y'} \right) = 0$, або $F_{y'} = C$, де C – стала інтегрування. Це звичайне диференціальне рівняння першого порядку – перший інтеграл рівняння Ейлера.

Приклад. Знайти екстремалі функціонала $J(y) = \int\limits_0^1 \left(y'^2 + 2xy' \right) dx$, якщо $y(0) = 0$, $y(1) = 2$.

Розв'язання. Підінтегральний вираз у заданому функціоналі не залежить від y, тому замість рівняння Ейлера можемо використати його перший інтеграл:

$$F_{y'} = C \Rightarrow 2y' + 2x = C.$$

Маємо диференціальне рівняння першого порядку, з якого знаходимо:

$$y' = \frac{C}{2} - x,\ \ y = \frac{Cx}{2} - \frac{x^2}{2} + C_1 = C_2 x - \frac{x^2}{2} + C_1.$$

Тут $C_2 = \frac{C}{2}$.

Використовуючи крайові умови, знаходимо значення сталих C_1 та C_2:

$$y(0) = C_1 = 0,\ y(1) = C_2 - \frac{1}{2} = 2 \Rightarrow C_2 = \frac{5}{2}.$$

Таким чином, отримали екстремаль $y = \frac{5}{2}x - \frac{x^2}{2}$.

5. Підінтегральний вираз функціонала (1.8) не залежить явно від x. У цьому випадку $F = F(y, y')$. Рівняння Ейлера набуває вигляду:

$$F_y - F_{y'y} \cdot y' - F_{y'y'} \cdot y'' = 0.$$

Помноживши обидві частини цієї рівності на y', можемо записати її у формі:

$$\frac{d}{dx}\left(F - y' \cdot F_{y'}\right) = 0.$$

Звідси отримуємо для даного випадку перший інтеграл рівняння Ейлера у вигляді $F - y' \cdot F_{y'} = C$.

Приклад. Знайти екстремалі функціонала $J(y) = \int\limits_0^a \frac{\sqrt{1 + y'^2}}{\sqrt{y}} dx$.

Розв'язання. Оскільки підінтегральний вираз функціонала не залежить явно від змінної x, то для знаходження екстремалей доцільно використати перший інтеграл рівняння Ейлера $F - y' \cdot F_{y'} = C$. Для даного функціонала $f = \frac{\sqrt{1 + y'^2}}{\sqrt{y}}$, звідси $f_{y'} = \frac{y'}{\sqrt{y\left(1 + y'^2\right)}}$. Отримуємо звичайне диференціальне рівняння першого порядку:

$$\frac{\sqrt{1 + y'^2}}{\sqrt{y}} - \frac{y'^2}{\sqrt{y\left(1 + y'^2\right)}} = C.$$

Приводячи дроби у лівій частині даної рівності до спільного знаменника, після перетворень отримаємо:

$$\frac{1}{\sqrt{y\left(1 + y'^2\right)}} = C \Rightarrow y\left(1 + y'^2\right) = C_1,\ C_1 = \frac{1}{C^2}.$$

Будемо шукати розв'язок даного диференціального рівняння у параметричній формі:

$$y' = \operatorname{ctg} t,\; y = \frac{C_1}{1+\operatorname{ctg}^2 t} = C_1 \sin^2 t,\; dy = C_1 \sin 2t\, dt,$$

$$dx = \frac{dy}{\operatorname{ctg} t} = \frac{C_1 \sin 2t dt}{\operatorname{ctg} t} = 2C_1 \sin^2 t\, dt.$$

Інтегруючи, знаходимо:

$$x = 2C_1 \int \sin^2 t dt = C_1\left(t - \frac{1}{2}\sin 2t\right) + C_2 .$$

Таким чином, отримали рівняння екстремалі варіаційної задачі у параметричній формі:

$$\begin{cases} x = C_1\left(t - \dfrac{1}{2}\sin 2t\right) + C_2, \\ y = C_1 \sin^2 t. \end{cases}$$

Це рівняння сімейства циклоїд.

До дослідження на екстремум функціонала з цього прикладу зводиться задача про брахістохрону, змістовна постановка якої була наведена у вступі.

Зауважимо, що у випадку коли підінтегральний вираз фунціонала є квадратичним відносно величин y та y', то рівняння Ейлера є лінійним диференціальним рівнянням другого порядку, при тому, що його перший інтеграл є нелінійним рівнянням першого порядку. Тому для знаходження екстремалей тут здебільшого доцільно використовувати безпосередньо рівняння Ейлера.

Приклад. Знайти екстремалі функціонала $J(y) = \int_0^1 \left(yy' + y'^2\right)dx$, якщо $y(0) = 0$, $y(1) = 1$.

Розв'язання. Підінтегральний вираз даного фунціонала є квадратичним відносно y та y', тому, незважаючи на те, що він не залежить явно від x, доцільно безпосередньо використати рівняння Ейлера замість його першого інтеграла. Отримаємо:

$$F = yy' + y'^2, F_y = y', F_{y'} = y + 2y', \frac{d}{dx}\left(F_{y'}\right) = y' + 2y''.$$

Рівняння Ейлера для функціонала $J(y)$ має вигляд $-2y'' = 0$, звідки отримуємо $y = C_1 x + C_2$. З крайових умов знаходимо значення сталих C_1 та C_2:

$$y(0) = C_2 = 0, y(1) = C_1 = 1.$$

Таким чином, шукана екстремаль має вигляд $y = x$.

Питання для самоперевірки

1. Сформулювати означення функціонала.
2. Навести приклади функціоналів.
3. Які функціонали є лінійними? Навести приклади.
4. Сформулювати означення неперервного функціонала.
5. Розкрити зміст поняття нормованого простору. Навести приклади.
6. Визначити норми у просторах $C_{[a;b]}$ та $C^1_{[a;b]}$.
7. Які функції утворюють сильний та слабкий ε- околи точки нормованого простору?
8. Що означає поняття близькості функцій нульового, першого, k-го порядків?
9. Що називають відстанню між точками нормованого простору? Навести приклади.
10. Поняття диференційовності функціонала.
11. Навести означення диференціала Фреше.
12. Навести означення диференціала Гато.
13. Надати означення варіації функціонала.
14. Для яких функціоналів диференціали Фреше та Гато співпадають?
15. Навести приклади інтегральних функціоналів.
16. Записати варіацію інтегральних функціоналів виду $J = \int\limits_a^b f\left(x, y, y'\right)dx$.
17. Пояснити зміст поняття сильного та слабкого екстремумів функціонала.
18. Надати означення глобального екстремуму функціонала.
19. Сформулювати та довести необхідну умову екстремуму функціонала.
20. Сформулювати та довести лему Лагранжа. Навести її узагальнення на випадок простору функцій двох змінних.
21. Сформулювати та довести лему Дюбуа – Раймона.
22. Охарактеризувати область визначення інтегрального функціонала виду $J\left(y\right) = \int\limits_a^b f\left(x, y, y'\right)dx$.
23. Виходячи з необхідної умови екстремуму функціонала, отримати рівняння Ейлера.
24. Надати означення екстремалі функціонала.
25. Пояснити, для якого типу екстремуму рівняння Ейлера виражає його необхідну умову.
26. У яких випадках рівняння Ейлера допускає безпосередній перехід до його першого інтегралу? Навести приклади перших інтегралів рівняння Ейлера.
27. Навести приклади функціоналів, для яких екстремалей не існує.
28. Чи має зміст варіаційна задача для функціонала, підінтегральний вираз у якому є повним диференціалом деякої функції?

Практичні завдання

Відшукати екстремалі наступних функціоналів [1,2]:

1. $J(y)=\int\limits_{-1}^{0}\left(12xy-y'^2\right)dx\to extr,\quad y(-1)=1,\ y(0)=0.$

2. $J(y)=\int\limits_{0}^{\pi}\left(4y\cos x+y'^2-y^2\right)dx\to extr,\quad y(0)=0,\ y(\pi)=0.$

3. $J(y)=\int\limits_{0}^{\pi/2}\left(4y\sin x+y^2-y'^2\right)dx\to extr;\ y(0)=y(\pi/2)=0.$

4. $J(y)=\int\limits_{1}^{2}\left(xy^2-y\right)dx\to extr,\quad y(1)=\frac{1}{2},\ y(2)=\frac{1}{2}.$

5. $J(y)=\int\limits_{0}^{1}\left(2xy+(x^2+e^y)y'\right)dx\to extr,\quad y(0)=0,\ y(1)=1.$

6. $J(y)=\int\limits_{1}^{2}\left(y'^2+2y'\right)dx\to extr,\quad y(1)=0,\ y(2)=1.$

7. $J(y)=\int\limits_{0}^{1}\left(xy'+y'^2\right)dx\to extr,\quad y(0)=0,\ y(1)=1.$

8. $J(y)=\int\limits_{0}^{1}\left(y'^2+xy\right)dx\to extr;\ y(0)=y(1)=0.$

9. $J(y)=\int\limits_{0}^{\pi/2}\left(y'^2-y^2-xy\right)dx\to extr;\ y(0)=y(\pi/2)=0.$

10. $J(y)=\int\limits_{0}^{1}\left(y'^2+yy'+12xy\right)dx\to extr;\ y(0)=y(1)=0.$

11. $J(y)=\int\limits_{1}^{e}xy'^2dx\to extr;\ y(1)=0,\ y(e)=1.$

12. $J(y)=\int\limits_{1}^{2}x^2y'^2dx\to extr;\ y(1)=3,\ y(2)=1.$

13. $J(y)=2\pi\int\limits_{x_0}^{x_1}y\sqrt{1+y'^2}\,dx\to extr,\quad y(x_0)=y_0,\ y(x_1)=y_1.$

14. $J(y)=\int\limits_{0}^{1/2}\frac{\sqrt{1+y'^2}}{y}dx\to extr;\ y(0)=1,\ y(1/2)=\sqrt{3}/2.$

15. $J(y)=\frac{1}{\sqrt{2g}}\int\limits_{x_1}^{x_2}\frac{\sqrt{1+y'^2}}{\sqrt{y}}\,dx\to extr,\quad y(x_1)=y_1,\ y(x_2)=y_2.$

2 УЗАГАЛЬНЕННЯ НАЙПРОСТІШОЇ ЗАДАЧІ ВАРІАЦІЙНОГО ЧИСЛЕННЯ

План:

2.1 Необхідна умова екстремуму для функціоналів, що залежать від кількох функцій.
2.2 Функціонали, що залежать від похідних вищих порядків.
2.3 Функціонали, що залежать від функцій кількох змінних.
2.4 Канонічний вигляд рівнянь Ейлера.

Ключові поняття та терміни: система рівнянь Ейлера, рівняння Ейлера-Пуассона, рівняння Ейлера-Остроградського, канонічна форма рівнянь Ейлера.

2.1 Необхідна умова екстремуму для функціоналів, що залежать від кількох функцій

Розглянемо випадок, коли функціонал залежить від кількох функцій однієї незалежної змінної, тобто він має вигляд:

$$J\left(y_1(x), y_2(x), \ldots, y_n(x)\right) = \int_a^b F\left(x, y_1, y_2, \ldots, y_n, y_1', y_2', \ldots, y_n'\right) dx, \qquad (2.1)$$

де F – двічі неперервно диференційовна функція своїх аргументів x, y_i, $i = 1, 2, \ldots, n$. Нехай функціонал (2.1) визначений на множині вектор-функцій $Y = \left(y_1, y_2, \ldots, y_n\right)$, де функції $y_i(x)$, $i = 1, 2, \ldots, n$ належать простору $C^1_{[a;b]}$ та задовольняють заданим крайовим умовам:

$$\begin{cases} y_1(a) = y_{11},\ y_2(a) = y_{21}, \ldots, y_n(a) = y_{n1}; \\ y_1(b) = y_{12},\ y_2(b) = y_{22}, \ldots, y_n(b) = y_{n2}. \end{cases} \qquad (2.2)$$

Допустимі варіації δy_i, $i = 1, 2, \ldots, n$, для функцій $y_i(x)$, $i = 1, 2, \ldots, n$, повинні також належати функціональному простору $C^1_{[a;b]}$. Вони також дорівнюють нулю на межах відрізка інтегрування, оскільки у точках $x = a$ та $x = b$ функції $y_i(x)$, $i = 1, 2, \ldots, n$ мають фіксовані значення.

Необхідна умова слабкого екстремуму функціонала (2.1) визначається наступною теоремою.

<u>Теорема 2.1</u> Якщо функціонал (2.1) досягає слабкого екстремуму на вектор-функції $Y(x) = \left(y_1(x), y_2(x), \ldots, y_n(x)\right)$, то функції $y_i(x)$, $i = 1, 2, \ldots, n$, є розв'язками системи диференціальних рівнянь Ейлера:

$$F_{y_i} - \frac{d}{dx}\left(F_{y_i'}\right) = 0, \quad i = 1, 2, \ldots, n. \qquad (2.3)$$

Доведення. Для довільних допустимих варіацій $\delta y_1, \delta y_2, \ldots, \delta y_n$ розглянемо функцію $\varphi(\alpha_1, \alpha_2, \ldots, \alpha_n) = I\left(y_1 + \alpha_1 \delta y_1, y_2 + \alpha_2 \delta y_2, \ldots, y_n + \alpha_n \delta y_n\right)$.

Якщо на вектор-функції $Y(x)=(y_1(x), y_2(x), \ldots, y_n(x))$ функціонал (2.1) досягає екстремуму, то функція $\varphi(\alpha_1, \alpha_2, \ldots, \alpha_n)$ повинна мати екстремум при $\alpha_1=\alpha_2=\ldots=\alpha_n=0$, тобто при нульових значеннях $\alpha_i, i=1,2,\ldots,n$ повинна виконуватися необхідна умова екстремуму даної функції – $\frac{\partial\varphi}{\partial\alpha_i}=0, i=1,2,\ldots,n$.

Таким чином, при нульових значеннях $\alpha_i, i=1,2,\ldots,n$ отримуємо:

$$\frac{\partial\varphi}{\partial\alpha_i}=\int_a^b \left(F_{y_i}\cdot\delta y_i+F_{y_i'}\cdot\delta y_i'\right)dx=0, i=1,2,\ldots,n.$$

Застосувавши до цих рівностей лему Дюбуа – Раймона (лему 1.3), отримуємо систему рівнянь Ейлера (2.3) для функціонала (2.1).

Означення 2.1 Будь-який розв’язок $Y(x)=(y_1(x), y_2(x), \ldots, y_n(x))$ системи рівнянь Ейлера для функціонала (2.1) називають **екстремаллю** цього функціонала.

Система рівнянь Ейлера (2.3) для функціонала (2.1) є системою n звичайних диференціальних рівнянь першого порядку, тому її розв’язки – екстремалі $Y(x)$ містять n довільних сталих $C_1, C_2, \ldots, C_n$. Для їх визначення використовують крайові умови (2.2).

Приклад. Для функціонала $J(y_1(x), y_2(x))=\int_0^{\pi/2}\left(y_1'^2+y_2'^2+2y_1y_2\right)dx$ знайти екстремалі, що задовольняють крайові умови $y_1(0)=0,\ y_2(0)=0,$ $y_1\left(\frac{\pi}{2}\right)=1,\ y_2\left(\frac{\pi}{2}\right)=-1$ [3].

Розв’язання. Побудуємо систему рівнянь Ейлера для заданого функціонала:

$$F=y_1'^2+y_2'^2+2y_1y_2,\ F_{y_1}=2y_2, F_{y_2}=2y_1,\ f_{y_1'}=2y_1',\ F_{y_2'}=2y_2',$$

$$\frac{d}{dx}\left(F_{y_1'}\right)=2y_1'',\ \frac{d}{dx}\left(F_{y_2'}\right)=2y_2''.$$

Підставивши у (2.3), отримуємо систему:

$$\begin{cases} y_2-y_1''=0, \\ y_1-y_2''=0. \end{cases}$$

Розв’яжемо цю систему, звівши її до одного рівняння. Для цього двічі продиференціюємо перше рівняння системи та за допомогою другого рівняння виключимо з нього функцію y_2. Отримаємо рівняння $y_1^{(4)}-y_1=0$. Корені його характеристичного рівняння $k^4-1=0$ $k_{1,2}=\pm1, k_{3,4}=\pm i$. За цими коренями знаходимо функцію $y_1(x)$:

$$y_1(x)=C_1\operatorname{ch}x+C_2\operatorname{sh}x+C_3\cos x+C_4\sin x.$$

З першого рівняння системи $y_2 = y_1''$ знаходимо y_2:

$$y_2(x) = C_1 \operatorname{ch} x + C_2 \operatorname{sh} x - C_3 \cos x - C_4 \sin x.$$

Для знаходження сталих інтегрування C_1, C_2, C_3, C_4 використаємо крайові умови:

$$\begin{cases} y_1(0) = C_1 + C_3 = 0, \\ y_2(0) = C_1 - C_3 = 0, \\ y_1\left(\dfrac{\pi}{2}\right) = C_1 \operatorname{ch}\dfrac{\pi}{2} + C_2 \operatorname{sh}\dfrac{\pi}{2} + C_4 = 1, \\ y_2\left(\dfrac{\pi}{2}\right) = C_1 \operatorname{ch}\dfrac{\pi}{2} + C_2 \operatorname{sh}\dfrac{\pi}{2} - C_4 = -1. \end{cases}$$

З отриманої системи лінійних алгебраїчних рівнянь знаходимо $C_1 = C_2 = C_3 = 0$, $C_4 = 1$. Таким чином, отримано рівняння екстремалі заданого функціонала $Y(x) = (y_1(x), y_2(x))$, де $y_1 = \sin x$, $y_2 = -\sin x$.

2.2 Функціонали, що залежать від похідних вищих порядків

Нехай функція $F\left(y, y', y'', \ldots, y^{(n)}\right)$ є неперервно диференційовною $n+2$ рази за своїми аргументами функцією. Розглянемо функціонал

$$J(y) = \int_a^b F\left(y, y', y'', \ldots, y^{(n)}\right) dx, \tag{2.4}$$

визначений на множині n разів неперервно диференційовних на $[a; b]$ функцій $y(x) \in C^n_{[a;b]}$, що задовольняють заданим крайовим умовам:

$$\begin{cases} y(a) = y_{10}, y'(a) = y_{11}, \ldots, y^{(n-1)}(a) = y_{1,n-1}, \\ y(b) = y_{20}, y'(b) = y_{21}, \ldots, y^{(n-1)}(b) = y_{2,n-1}. \end{cases} \tag{2.5}$$

Оскільки будь-яка допустима для даної варіаційної задачі функція $y(x) + \delta y(x)$ повинна задовольняти умови (2.5), то допустимою варіацією функції $y(x)$ є довільна функція $\delta y \in C^n_{[a;b]}$, що задовольняє однорідні крайові умови:

$$\begin{cases} \delta y(a) = \delta y'(a) = \ldots = \delta y^{(n-1)}(a) = 0, \\ \delta y(b) = \delta y'(b) = \ldots = \delta y^{(n-1)}(b) = 0. \end{cases} \tag{2.6}$$

Визначимо необхідну умову, якій повинна задовольняти функція $y(x)$ з даної множини, що надає екстремум функціоналу (2.4).

Теорема 2.2 Якщо функціонал (2.4), визначений на множині функцій з $C^n_{[a;b]}$, що задовольняють крайові умови (2.5), досягає екстремуму на деякій функції $y(x) \in C^{2n}_{[a;b]}$, то ця функція є розв'язком **рівняння Ейлера – Пуассона**:

$$F_y - \frac{d}{dx}\left(F_{y'}\right) + \frac{d^2}{dx^2}\left(F_{y''}\right) - \ldots + (-1)^n \frac{d^n}{dx^n}\left(F_{y^{(n)}}\right) = 0. \tag{2.7}$$

Доведення. Нехай функція $y(x)$ надає екстремум функціоналу (2.4). Вибравши довільно варіацію δy, що задовольняє крайові умови (2.6), та зафіксувавши її, розглянемо функцію

$$\varphi(\alpha) = J\left(y + \alpha \cdot \delta y\right) = \int_a^b F\left(x, y + \alpha \cdot \delta y, y' + \alpha \cdot \delta y', \ldots, y^{(n)} + \alpha \cdot \delta y^{(n)}\right) dx.$$

Оскільки функція $y(x)$ надає екстремум функціоналу $J(y)$, то $\varphi(\alpha)$ досягає екстремуму при $\alpha = 0$, тому $\varphi'(0) = 0$, тобто отримуємо рівність

$$\varphi'(0) = \int_a^b \left(F_y \delta y + F_{y'} \delta y' + \ldots + F_{y^{(n)}} \delta y^{(n)}\right) dx = 0. \tag{2.8}$$

Нехай $y(x) \in C^{2n}_{[a;b]}$. У цьому випадку, використавши інтегрування частинами та крайові умови (2.6), отримуємо:

$$\int_a^b F_{y'} \delta y' dx = -\int_a^b \left(\frac{d\left(F_{y'}\right)}{dx}\right) \delta y dx;$$

$$\int_a^b F_{y''} \delta y'' dx = \int_a^b \left(\frac{d^2\left(F_{y''}\right)}{dx^2}\right) \delta y dx;$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\int_a^b F_{y^{(n)}} \delta y^{(n)} dx = (-1)^n \int_a^b \left(\frac{d^n\left(F_{y^{(n)}}\right)}{dx^n}\right) \delta y dx.$$

Підставивши ці співвідношення у (2.8), отримуємо:

$$\int_a^b \left(F_y - \frac{d\left(F_{y'}\right)}{dx} + \frac{d^2\left(F_{y''}\right)}{dx^2} - \ldots + (-1)^n \frac{d^n\left(F_{y^{(n)}}\right)}{dx^n}\right) \delta y dx = 0.$$

Остання рівність виконується для довільної нескінченно диференційованої функції $\delta y(x)$, тому, за лемою Лагранжа, отримуємо рівняння (2.7).

Його розв'язки називаються **екстремалями** функціонала (2.4).

Рівняння Ейлера – Пуассона (2.7) є звичайним диференціальним рівнянням порядку $2n$. У його загальному розв'язку присутні $2n$ довільних сталих, які визначаються з крайових умов (2.5).

Приклад. Знайти екстремалі функціонала $J(y)=\int_a^b\left((y''')^2+y^2-2yx^3\right)dx$.

Розв'язання. Запишемо для даного функціонала рівняння Ейлера – Пуассона, для чого знайдемо необхідні похідні.

$$F_y=2y-2x^3;\ F_{y'}=F_{y''}=0;\ F_{y'''}=2y''';\ \frac{d\left(F_{y'''}\right)}{dx^3}=2y^{(6)}.$$

Підставивши їх у рівняння (2.7), отримаємо:

$$y^{(6)}-y=-x^3.$$

Коренями характеристичного рівняння $k^6-1=0$ для отриманого лінійного неоднорідного диференціального рівняння з сталими коефіцієнтами є

$$k_{1,2}=\pm1,\ k_{3,4}=\frac{-1\pm\sqrt{3}i}{2},\ k_{5,6}=\frac{1\pm\sqrt{3}i}{2}.$$

За коренями характеристичного рівняння запишемо розв'язок однорідного рівняння $y^{(6)}-y=0$:

$$y_0=C_1e^x+C_2e^{-x}+e^{-\frac{x}{2}}\left(C_3\cos\frac{\sqrt{3}}{2}x+C_4\sin\frac{\sqrt{3}}{2}x\right)+$$
$$+e^{\frac{x}{2}}\left(C_5\cos\frac{\sqrt{3}}{2}x+C_6\sin\frac{\sqrt{3}}{2}x\right).$$

Частинний розв'язок отриманого неоднорідного рівняння Ейлера – Пуассона знаходимо за виглядом правої частини: $\tilde{y}=Ax^3+Bx^2+Cx+D$. Невизначені коефіцієнти у цьому розв'язку знаходимо шляхом підстановки його у неоднорідне диференціальне рівняння. При цьому отримуємо: $A=1$, решта коефіцієнтів дорівнює нулю. Тому $\tilde{y}=x^3$, а екстремалі заданого функціонала мають вигляд $y=y_0+\tilde{y}$.

2.3 Функціонали, що залежать від функцій кількох змінних

Знайдемо необхідну умову екстремуму функціонала, визначеного на множині функцій, що залежать від кількох змінних на прикладі функціонала, областю визначення якого є множина функцій виду $z=z(x,y)$, залежних від 2 змінних. Далі будемо розглядати функціонали $J(z)$, визначені на функціях $z=z(x,y)$, двічі неперервно диференційовних у деякій області $D\subset\mathbb{R}^2$, обмеженій замкненою кривою L, тобто функціонали виду

$$J(z)=\iint_D F\left(x,y,z(x,y),z_x(x,y),z_y(x,y)\right)dxdy. \qquad (2.9)$$

У (2.9) F є двічі неперервно диференційованою функцією своїх аргументів x та y.

Розглянемо задачу про знаходження екстремуму функціонала (2.9), визначеного на множині функцій $z = z(x, y)$, які на межі L області D набувають заданих значень, тобто

$$z(x, y)\Big|_L = \varphi(x, y), \qquad (2.10)$$

де задана функція $\varphi(x, y)$ визначена на кривій L.

Допустимими варіаціями для $z = z(x, y)$ у даному випадку є функції $\delta z(x, y)$, двічі диференційовані у області D, що дорівнюють нулю на межі L даної області. Запишемо необхідну умову екстремуму функціонала $J(z)$:

$$\delta J = \frac{\partial}{\partial \alpha}\big(J(z + \alpha \cdot \delta z)\big)\Big|_{\alpha=0} = 0. \qquad (2.11)$$

Вираз для варіації функціонала (2.9) набуває вигляду:

$$\delta J = \frac{\partial}{\partial \alpha}\big(J(z + \alpha \cdot \delta z)\big)\Big|_{\alpha=0} = \iint_D \big(F_z \delta z + F_p \delta p + F_q \delta q\big) dxdy,$$

де $p = z_x, q = z_y$.

Оскільки $\frac{\partial}{\partial x}\big(F_p \delta z\big) = \frac{\partial F_p}{\partial x}\delta z + F_p \delta p$, $\frac{\partial}{\partial y}\big(F_q \delta z\big) = \frac{\partial F_q}{\partial y}\delta z + F_q \delta q$, отримуємо:

$$\iint_D \big(F_p \delta p + F_q \delta q\big) dxdy = \iint_D \left(\frac{\partial}{\partial x}\big(F_p \delta z\big) + \frac{\partial}{\partial y}\big(F_q \delta z\big)\right) dxdy -$$

$$- \iint_D \left(\frac{\partial F_p}{\partial x}\delta z + \frac{\partial F_q}{\partial y}\delta z\right) dxdy.$$

До першого з інтегралів у правій частині останньої рівності застосуємо формулу Гріна, згідно з якою

$$\iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy = \int_L P dx + Q dy.$$

Отримаємо:

$$\iint_D \left(\frac{\partial}{\partial x}\big(F_p \delta z\big) + \frac{\partial}{\partial y}\big(F_q \delta z\big)\right) dxdy = \int_L F_p \delta z\, dy - F_q \delta z\, dx = \int_L \delta z\big(F_p dy - F_q dx\big) = 0,$$

оскільки на контурі L, де значення допустимих функцій $z(x, y)$ є фіксованими, $\delta z = 0$.

Таким чином, маємо:

$$\iint_D \big(F_p \delta p + F_q \delta q\big) dxdy = -\iint_D \left(\frac{\partial F_p}{\partial x} + \frac{\partial F_q}{\partial y}\right) \delta z\, dxdy.$$

Необхідна умова екстремуму функціонала (2.9) набуває вигляду:

$$\delta J = \iint_D \left(F_z - \frac{\partial F_p}{\partial x} - \frac{\partial F_q}{\partial y}\right) \delta z\, dxdy = 0. \qquad (2.12)$$

Оскільки перший співмножник під знаком інтеграла у (2.12) є неперервним, а варіація δz є довільною функцією, що приймає нульове значення на контурі L, то згідно леми 1.2, повинна виконуватися рівність

$$F_z - \frac{\partial F_p}{\partial x} - \frac{\partial F_q}{\partial y} = 0. \qquad (2.13)$$

Отримане диференціальне рівняння у частинних похідних називається **рівнянням Ейлера – Остроградського**, а будь-який диференційований розв'язок цього рівняння – **екстремаллю функціонала** (2.9).

Приклад. Записати рівняння Ейлера – Остроградського для функціонала Діріхле $J(z) = \iint\limits_D \left(z_x{}^2 + z_y{}^2\right)dxdy$.

Розв'язання. Оскільки $p = z_x$, $q = z_y$, то $F(x, y, z, p, q) = p^2 + q^2$. Звідси:

$$F_z = 0, F_p = 2p = 2\frac{\partial z}{\partial x}, F_q = 2q = 2\frac{\partial z}{\partial y}, \frac{\partial F_p}{\partial x} = 2\frac{\partial^2 z}{\partial x^2}, \frac{\partial F_q}{\partial y} = 2\frac{\partial^2 z}{\partial y^2}.$$

Рівняння Ейлера – Остроградського для функціонала Діріхле набуває вигляду:

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = 0.$$

Отримали рівняння Лапласа для функції $z(x, y)$. Таким чином, екстремалями функціонала Діріхле є гармонічні функції.

2.4 Канонічний вигляд рівнянь Ейлера

Функціоналу

$$J(y_1, y_2, \ldots, y_n) = \int_a^b F(x, y_1, y_2, \ldots, y_n, y_1', y_2', \ldots, y_n')dx, \qquad (2.14)$$

де F – двічі неперервно диференційована функція, відповідає система рівнянь Ейлера:

$$F_{y_i} - \frac{d\left(F_{y_i''}\right)}{dx} = 0, \quad i = 1, 2, \ldots, n. \qquad (2.15)$$

У загальному випадку система (2.15) є системою, що складається з n звичайних диференціальних рівнянь другого порядку. Введемо нові змінні $z_i = y_i'$, $i = 1, 2, \ldots, n$. Тоді $F_{y_i'} = F_{z_i}$. Отримуємо систему, що складається з $2n$ звичайних диференціальних рівнянь першого порядку, еквівалентну системі (2.15):

$$\begin{cases} F_{y_i} - \dfrac{d\left(F_{z_i}\right)}{dx} = 0, i = 1, 2, \ldots, n; \\ \dfrac{dy_i}{dx} = z_i, i = 1, 2, \ldots, n. \end{cases} \qquad (2.16)$$

Отримана система не є нормальною системою диференціальних рівнянь, тобто вона не розв'язана відносно перших похідних. Нехай $p_i = F_{y_i'}, i = 1, 2, \ldots, n$.

Визначник матриці Гессе (матриці частинних похідних другого порядку) функції F за змінними y_i' співпадає з якобіаном

$$\frac{D\left(p_1,p_2,\ldots,p_n\right)}{D\left(y_1',y_2',\cdots,y_n'\right)}=\begin{vmatrix}\frac{\partial p_1}{\partial y_1'} & \frac{\partial p_1}{\partial y_2'} & \cdots\frac{\partial p_1}{\partial y_n'}\\ \frac{\partial p_2}{\partial y_1'} & \frac{\partial p_2}{\partial y_2'} & \cdots\frac{\partial p_2}{\partial y_n'}\\ \cdots & \cdots & \cdots\\ \frac{\partial p_n}{\partial y_1'} & \frac{\partial p_n}{\partial y_2'} & \cdots\frac{\partial p_n}{\partial y_n'}\end{vmatrix}.$$

Якщо цей визначник не дорівнює нулю, то згідно з теоремою про існування оберненої функції, система рівнянь $p_i=F_{y_i'},i=1,2,\ldots,n$, визначає сукупність змінних y_i' як функцію введених змінних p_i:

$$y_i'=h_i\left(x,y_1,y_2,\ldots y_n,p_1,p_2,\ldots,p_n\right),\, i=1,2,\ldots,n. \qquad (2.17)$$

Розглянемо функцію

$$H=-f\left(x,y_1,y_2,\ldots,y_n,h_1,h_2,\ldots,h_n\right)+\sum_{i=1}^{n}h_i p_i. \qquad (2.18)$$

Функцію (2.18) називають **функцією Гамільтона** функціонала (2.14), а змінні $x,y_1,\ldots,y_n,p_1,\ldots,p_n$ – **канонічними змінними** даного функціонала.

Приклад. Побудувати функцію Гамільтона для функціонала

$$J\left(y_1,y_2\right)=\int_a^b\left(3xy_1y_1'^2+2e^{y_1}y_2'^2+y_1y_2\right)dx.$$

Розв’язання. Маємо $p_1=F_{y_1'}=6xy_1y_1',\, p_2=F_{y_2'}=4e^{y_1}y_2'$. Тоді $y_1'=\frac{p_1}{6xy_1}=h_1$, $y_2'=\frac{p_2}{4e^{y_1}}=h_2$. Підставивши отримані вирази у (2.18), знаходимо функцію Гамільтона для заданого функціонала:

$$H=-\left(3xy_1\cdot\frac{p_1^2}{36x^2y_1^2}+2e^{y_1}\cdot\frac{p_2^2}{16e^{2y_1}}+y_1y_2\right)+\frac{p_1^2}{6xy_1}+\frac{p_2^2}{4e^{y_1}}=\frac{p_1^2}{12xy_1}+\frac{p_2^2}{8e^{y_1}}-y_1y_2.$$

З означення функції Гамільтона (2.18) випливає, що її повний диференціал

$$\begin{gathered}dH=-dF+\sum_{i=1}^{n}p_idy_i'+\sum_{i=1}^{n}y_i'dp_i=-F_xdx-\sum_{i=1}^{n}F_{y_i}dy_i-\sum_{i=1}^{n}F_{y_i'}dy_i'+\\ +\sum_{i=1}^{n}p_idy_i'+\sum_{i=1}^{n}y_i'dp_i.\end{gathered} \qquad (2.19)$$

Оскільки $p_i=F_{y_i'},i=1,\ldots,n$, то третій та четвертий доданки у правій частині рівності (2.19) взаємно знищуються. Тому отримуємо наступний вираз для dH:

$$dH=-F_xdx-\sum_{i=1}^{n}F_{y_i}dy_i+\sum_{i=1}^{n}y_i'dp_i=\frac{\partial H}{\partial x}dx+\sum_{i=1}^{n}\frac{\partial H}{\partial y_i}dy_i+\sum_{i=1}^{n}\frac{\partial H}{\partial p_i}dp_i.$$

З останньої рівності випливає: $\dfrac{\partial H}{\partial x}=-\dfrac{\partial F}{\partial x},\dfrac{\partial H}{\partial y_i}=-\dfrac{\partial F}{\partial y_i},\dfrac{\partial H}{\partial p_i}=y_i', i=1,2,\dots,n$. Звідси знаходимо, що $F_{y_i}=-H_{y_i}, y_i'=H_{p_i}$. Оскільки $F_{z_i}=F_{y_i'}=p_i, z_i=y_i'$, то з системи (2.16) отримуємо нормальну систему звичайних диференціальних рівнянь, еквівалентну системі (2.15):

$$\begin{cases}\dfrac{dy_i}{dx}=\dfrac{\partial H}{\partial p_i}, i=1,2,\dots,n;\\[2mm] \dfrac{dp_i}{dx}=-\dfrac{\partial H}{\partial y_i}, i=1,2,\dots,n\end{cases}\tag{2.20}$$

Система (2.20) називається **канонічною формою рівнянь Ейлера** функціонала (2.14).

Перший інтеграл системи звичайних диференціальних рівнянь є функцією, що зберігає стале значення вздовж кожної інтегральної кривої даної системи. Для того, щоб деяка диференційована функція була першим інтегралом системи звичайних диференціальних рівнянь, необхідно та достатньо, щоб повна похідна цієї функції з врахуванням даної системи диференціальних рівнянь тотожно дорівнювала нулю. В окремих випадках можна вказати перші інтеграли системи (2.20), які є також першими інтегралами системи (2.15).

Нехай підінтегральний вираз у функціоналі (2.14) не залежить явно від x, тобто $F=F\left(y_1,y_2,...,y_n,y_1',y_2',...,y_n'\right)$. Тоді повна похідна функції Гамільтона набуває вигляду:

$$\frac{dH}{dx}=\sum_{i=1}^{n}H_{y_i}\frac{dy_i}{dx}+\sum_{i=1}^{n}H_{p_i}\frac{dp_i}{dx}.\tag{2.21}$$

Підставивши вирази для похідних H_{y_i} та H_{p_i} з системи (2.20) у рівність (2.21), отримуємо, що $\dfrac{dH}{dx}\equiv 0$, тобто функція Гамільтона є сталою вздовж кожної екстремалі $\left(y_1,y_2,\dots,y_n\right)$ функціонала (2.14). Таким чином, якщо підінтегральний вираз F функціонала (2.14) не залежить явно від x, то його функція Гамільтона є першим інтегралом системи рівнянь Ейлера (2.15).

З'ясуємо, за яких умов задана функція $\Phi\left(y_1,y_2,...,y_n,p_1,p_2,...,p_n\right)$, що не залежить явно від x є першим інтегралом системи (2.15). Обчислимо повну похідну функції Φ по незалежній змінній x, з врахуванням системи (2.15). Отримаємо:

$$\frac{d\Phi}{dx}=\sum_{i=1}^{n}\left(\Phi_{y_i}\frac{dy_i}{dx}+\Phi_{p_i}\frac{dp_i}{dx}\right)=\sum_{i=1}^{n}\left(\Phi_{y_i}H_{p_i}-\Phi_{p_i}H_{y_i}\right).$$

Вираз $\left[\Phi,H\right]=\sum\limits_{i=1}^{n}\left(\Phi_{y_i}H_{p_i}-\Phi_{p_i}H_{y_i}\right)$ називають **дужкою Пуассона** функцій Φ та H. Таким чином, $\dfrac{d\Phi}{dx}=\left[\Phi,H\right]$. Ми довели наступне твердження.

Теорема 2.3 Диференційована функція $\Phi\left(y_1, y_2, ..., y_n, p_1, p_2, ..., p_n\right)$ є першим інтегралом для канонічної форми рівнянь Ейлера (2.20) тоді і лише тоді, коли $\left[\Phi, H\right] = 0$.

Питання для самоперевірки

1. Отримати рівняння Ейлера для функціонала, що залежить від кількох функцій однієї змінної.
2. Яким умовам повинна задовольняти варіація інтегрального функціонала, що залежить від похідних вищих порядків?
3. Отримати рівняння Ейлера – Пуассона.
4. Які функції є екстремалями функціонала $J\left(y\right) = \int\limits_a^b f\left(x, y, y'\right) dx$, якщо його підінтегральна функція $f\left(x, y, y'\right) = f\left(y'\right)$?
5. Отримати рівняння Ейлера – Остроградського для функціоналів, що залежать від функції двох змінних.
6. Для якого функціонала рівняння Ейлера – Остроградського є рівнянням Лапласа?
7. Надати означення функції Гамільтона.
8. Навести систему рівнянь Ейлера у канонічному вигляді.

Практичні завдання

Відшукати екстремалі наступних функціоналів [1,2]:

1. $\int\limits_1^2 \left(y_1'^2 + y_2^2 + y_2'^2\right) dx \to extr;\ y_1(1) = 1,\ y_1(2) = 2,\ y_2(1) = 0,\ y_2(2) = 1.$
2. $\int\limits_0^\pi \left(2y_1 y_2 - 2y_1^2 + y_1'^2 - y_2'^2\right) dx \to extr;\ y_1(0) = 0,\ y_1(\pi) = 1,\ y_2(0) = 0,$ $y_2(\pi) = 1.$
3. $\int\limits_0^{\pi/4} \left(2y_2 - 4y_1^2 + y_1'^2 - y_2'^2\right) dx \to extr;\ y_1(0) = 0,\ y_1(\pi/4) = 1,\ y_2(0) = 0,$ $y_2(\pi/4) = 1.$
4. $\int\limits_{-1}^1 \left(2xy_1 - y_1'^2 + y_2'^3/3\right) dx \to extr;\ y_1(1) = 0,\ y_1(-1) = 2,\ y_2(1) = 1,$ $y_2(-1) = -1.$
5. $\int\limits_0^1 \left(y_1'^2 + y_2'^2 - 2y_1 y_2\right) dx \to extr;\ y_1(0) = 0,\ y_1(1) = sh\ 1,\ y_2(0) = 0,$ $y_2(1) = -sh\ 1.$

6. $\int_0^1 \left(y_1^2+y_2^2+2y_1'y_2'\right)dx \to extr;\ y_1(0)=y_2(0)=0,\ y_1(1)=y_2(1)=sh\ 1.$

7. $\int_a^b \left(2y_1\cos x+2y_2^2+2y_1'y_2'+y_1'^2+y_2'^2\right)dx \to extr.$

8. $\int_0^1 \left(y_1'y_2'+y_1y_2\right)dx \to extr;\ y_1(0)=y_2(0)=1,\ y_1(1)=e,\ y_2(1)=1/e.$

9. $\int_0^{\pi/2} \left(y_1'y_2'-y_1y_2\right)dx \to extr;\ y_1(0)=y_2(0)=0,\ y_1(\pi/2)=1,\ y_2(\pi/2)=-1.$

10. $\int_0^{\pi/2} \left(y_1'^2+y_3'^2+2y_1y_2+2y_2y_3+2y_2^2\right)dx \to extr;\ y_1(0)=y_3(0)=1,\ y_2(0)=-1,$
$y_1(\pi/2)=\pi/2,\ y_2\left(\pi/2\right)=0,\ y_3\left(\pi/2\right)=-\pi/2.$

11. $\int_0^1 e^{-x}y''^2dx \to extr;\ y(0)=0,\ y'(0)=1,\ y(1)=e,\ y'(1)=2e.$

12. $\int_0^1 e^{-x}y''^2dx \to extr;\ y(0)=1,\ y'(0)=1,\ y(1)=y'(1)=e.$

13. $\int_0^1 \left(x+1\right)y''^2dx \to extr;\ y(0)=y'(0)=0,\ y(1)=1,\ y'(1)=2.$

14. $\int_0^1 \left(x+1\right)y''^2dx \to extr;\ y(0)=0,\ y(1)=\ln 2,\ y'(0)=1,\ y'(1)=1/2.$

15. $\int_{x_0}^{x_1} \left(16y^2-y''^2+x^2\right)dx \to extr.$

16. $\int_0^1 \left(y''^2+y'^2\right)dx \to extr;\ y(0)=0,\ y'(0)=1,\ y(1)=sh\ 1,\ y'(1)=ch\ 1.$

17. $\int_0^1 \left(y''^2+y'^2\right)dx \to extr;\ y(0)=1,\ y'(0)=0,\ y(1)=ch\ 1,\ y'(1)=sh\ 1.$

18. $\int_{-1}^0 \left(240y-y'''^2\right)dx \to extr;\ y(-1)=1, y'(-1)=-4,5, y''(-1)=16,$
$y(0)=0, y'(0)=0, y''(0)=0.$

19. Написати рівняння Ейлера-Остроградського для функціонала

$$J\left[z(x,y)\right]=\iint_D\left[\left(\frac{\partial^2 z}{\partial x^2}\right)^2+\left(\frac{\partial^2 z}{\partial y^2}\right)^2+2\left(\frac{\partial^2 z}{\partial x\partial y}\right)^2-2zf(x,y)\right]dxdy.$$

3 ВАРІАЦІЙНІ ЗАДАЧІ З РУХОМИМИ МЕЖАМИ

План:

3.1 Природні крайові умови.

3.2 Загальна варіаційна задача з рухомими межами.

3.3 Умови трансверсальності.

Ключові поняття та терміни: природні крайові умови, умови трансверсальності.

3.1 Природні крайові умови

Вище були розглянуті варіаційні задачі, у яких допустимі функції були визначені на фіксованому відрізку $[a;b]$, при цьому були зафіксовані граничні точки допустимих кривих – точки $A(a,y(a))$ та $B(b,y(b))$. Розглянемо варіаційні задачі, у яких екстремум функціонала визначається на множині функцій, значення яких на межах відрізка $[a;b]$ не є фіксованими.

Нехай потрібно дослідити на екстремум функціонал

$$J(y)=\int_a^b F(x,y,y')dx, \qquad (3.1)$$

визначений на множині неперервно диференційованих функцій $y(x)$. На відміну від розглянутих раніше задач тут відсутні обмеження у вигляді крайових умов, що фіксують значення допустимих функцій на межах відрізку інтегрування $[a;b]$. З геометричної точки зору дана задача полягає у знаходженні кривої, що є графіком функції, кінці якого розташовані на вертикальних прямих $x=a$ та $x=b$, для якої відповідне значення функціонала (3.1) є екстремальним. Варіаційні задачі такого типу називають **задачами з рухомими межами**. У таких задачах варіація δy допустимої функції $y(x)$ уже не зобов'язана задовольняти нульовим крайовим умовам, як це мало місце для найпростішої варіаційної задачі з фіксованими межами, а може приймати довільні значення на кінцях відрізка $[a;b]$.

Нехай підінтегральний вираз у (3.1) є двічі неперервно диференційованою функцією своїх аргументів. Як було показано раніше, перша варіація функціонала (2.1) має вигляд:

$$\delta J(y,\delta y)=\int_a^b\left(\frac{\partial F}{\partial y}\delta y+\frac{\partial F}{\partial y'}\delta y'\right)dx \qquad (3.2)$$

З необхідної умови екстремуму функціонала (3.1) – рівності нулю його варіації випливає, що $\int_a^b\left(\frac{\partial F}{\partial y}\delta y+\frac{\partial F}{\partial y'}\delta y'\right)dx=0$.

Ця рівність виконується і для варіацій δy, що приймають нульові значення на кінцях відрізку $[a;b]$, тому функція $y(x)$, на якій функціонал (3.1) досягає екстремуму, повинна задовольняти рівнянню Ейлера:

$$\frac{\partial F}{\partial y}-\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)=0, \tag{3.3}$$

тобто вона повинна бути екстремаллю даного функціонала.

Після інтегрування частинами варіація функціонала (3.2) приймає вигляд:

$$\delta J(y,\delta y)=\int_a^b\left[\frac{\partial F}{\partial y}-\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right]\delta y dx+\left[\frac{\partial F}{\partial y'}\delta y\right]_a^b. \tag{3.4}$$

Інтеграл у правій частині (3.4) на екстремалі $y(x)$ дорівнює нулю, тому з рівності нулю варіації функціонала (3.4) випливає, що $\left[\frac{\partial F}{\partial y'}\delta y\right]_a^b=0$. Звідси отримуємо:

$$\left.\frac{\partial F}{\partial y'}\right|_{x=b}\delta y(b)-\left.\frac{\partial F}{\partial y'}\right|_{x=a}\delta y(a)=0.$$

Оскільки $\delta y(b)$ та $\delta y(a)$ у задачі з рухомими межами можуть приймати довільні значення, то остання рівність виконується у випадку, коли

$$\left.\frac{\partial F}{\partial y'}\right|_{x=b}=\left.\frac{\partial F}{\partial y'}\right|_{x=a}=0. \tag{3.5}$$

Таким чином, криві, на яких досягається екстремум функціонала (3.1) у задачі з рухомими межами, повинні задовольняти рівнянню Ейлера (3.3) та додатковим умовам (3.5) на межах відрізку інтегрування $[a;b]$. Умови (3.5) називають **природними крайовими умовами.**

Для функціонала (3.1) можна поставити також змішану задачу, у якій один кінець графіка шуканої екстремалі є фіксованим, тобто задано крайову умову виду $y(a)=y_1$, а інший кінець вільно переміщується вздовж прямої $x=b$. У таких випадках на вільному кінці кривої повинна виконуватися відповідна природна крайова умова.

Приклад. Знайти екстремалі функціонала $J(y)=\int_0^{\pi/4}\left(y'^2-y^2+4y\cos x\right)dx$ у варіаційній задачі з правою рухомою межею, якщо $y(0)=0$.

Розв’язання. Складемо рівняння Ейлера (2.3) для даного функціоналу. Знаходимо відповідні похідні:

$$\frac{\partial F}{\partial y}=-2y+4\cos x;\frac{\partial F}{\partial y'}=2y';\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)=2y''.$$

Рівняння Ейлера набуває вигляду: $y''+y=2\cos x$. Загальним розв’язком цього рівняння є $y=C_1\cos x+C_2\sin x+x\sin x$.

Природна крайова умова у точці $x=\frac{\pi}{4}$ має вигляд $\left.\frac{\partial F}{\partial y'}\right|_{x=\frac{\pi}{4}}=0$ або $\left.(2y')\right|_{x=\frac{\pi}{4}}=0$. З рівності $y(0)=0$ випливає, що $C_1=0$. Тоді $y'=C_2\cos x+\sin x+x\cos x$. Тоді $y'\left(\frac{\pi}{4}\right)=\frac{C_2}{\sqrt{2}}+\frac{1}{\sqrt{2}}+\frac{\pi}{4\sqrt{2}}=0$. Звідси знаходимо $C_2=-\frac{\pi}{4}-1$. Таким чином, шукана екстремаль даної варіаційної задачі має вигляд:

$$y=x\sin x-\left(\frac{\pi}{4}+1\right)\sin x.$$

3.2 Загальна варіаційна задача з рухомими межами

Розглянемо задачу дослідження на екстремум функціонала

$$J(y,a,b)=\int_a^b F(x,y,y')dx, \tag{3.6}$$

у випадку, коли на кінці допустимих кривих $y(x)$ не накладено ніяких обмежень, тобто вони є вільними. У цьому випадку область визначення допустимих функцій не є фіксованою, а змінюється у залежності від вибору функції. Якщо функція $y(x)$ є неперервно диференційовною на відрізку інтегрування $[a;b]$, причому на його межах $x=a$ та $x=b$ існують відповідні односторонні похідні, то цю функцію можна продовжити на більший відрізок $[a_0;b_0]$ так, що продовжена функція $\tilde{y}(x)$ буде неперервно диференційовною на $[a_0;b_0]$.

Варіаційну задачу з рухомими межами можна сформулювати наступним чином. Потрібно знайти екстремум функціонала (3.6) у просторі $C^1_{[a_0;b_0]}$ неперервно диференційовних на відрізку $[a_0;b_0]$ функцій при значеннях параметрів $a_0\le a\le b\le b_0$.

Норма простору $C^1_{[a_0;b_0]}$ не враховує відстані між межами графіків допустимих функцій, тому для даної варіаційної задачі близькість функцій визначимо за допомогою відстані ρ, яка для довільних функцій $y(x)$ з межами $A(a;y(a)),B(b;y(b))$ та $\tilde{y}(x)$ з межами $\tilde{A}(\tilde{a};y(\tilde{a}))$, $\tilde{B}(\tilde{b};y(\tilde{b}))$ задається рівністю

$$\rho(y,\tilde{y})=\max_T\left(\left|\tilde{y}(x)-y(x)\right|+\left|\tilde{y}'(x)-y'(x)\right|\right)+\left|A\tilde{A}\right|+\left|B\tilde{B}\right|,$$

де $T=[a;b]\cap\left[\tilde{a};\tilde{b}\right]$.

Величина приросту функціонала залежить не тільки від варіації функції $\delta y=\tilde{y}(x)-y(x)$, але й від варіації рухомих меж $\delta a=\tilde{a}-a$ та $\delta b=\tilde{b}-b$.

Для фіксованих варіацій $\delta y,\delta a,\delta b$ визначимо функцію $\varphi(\alpha)=$

$= J(y+\alpha\delta y, a+\alpha\delta a, b+\alpha\delta b)$. Якщо функція $y(x)$ є точкою екстремуму функціонала (3.6), то функція $\varphi(\alpha)$ досягатиме екстремуму при $\alpha = 0$. Якщо при цьому $\varphi(\alpha)$ диференційовна у точці $\alpha = 0$, то $\varphi'(0) = 0$.

Нехай підінтегральний вираз $F(x,y,y')$ у функціоналі (3.6) є двічі неперервно диференційовною функцією всіх своїх аргументів, а функція $y(x)$, графік якої обмежений точками $A(a; y(a))$ та $B(b; y(b))$, надає екстремум функціоналу (3.6). Функція $\varphi(\alpha)$ має вигляд:

$$\varphi(\alpha) = \int_{a+\alpha\delta a}^{b+\alpha\delta b} F(x, y+\alpha\delta y, y'+\alpha\delta y')dx.$$

Знайдемо $\varphi'(\alpha)$, використовуючи формулу диференціювання інтегралу за параметром, враховуючи, що від параметру у даному випадку залежать і межі інтегрування. Отримаємо:

$$\varphi'(\alpha) = \int_{a+\alpha\delta a}^{b+\alpha\delta b} F(x, y+\alpha\delta y, y'+\alpha\delta y')dx + F(x, y+\alpha\delta y, y'+\alpha\delta y')\Big|_{x=b+\alpha\delta b} \cdot \delta b - \\ - F(x, y+\alpha\delta y, y'+\alpha\delta y')\Big|_{x=a+\alpha\delta a} \cdot \delta a.$$

Звідси знаходимо:

$$\varphi'(0) = \int_a^b \left[\frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial y'}\delta y'\right]dx + F(x,y,y')\Big|_{x=b} \cdot \delta b - F(x,y,y')\Big|_{x=a} \cdot \delta a. \quad (3.7)$$

Перетворимо інтеграл у правій частині (3.7), виконавши інтегрування за частинами. Отримаємо:

$$\int_a^b \left[\frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial y'}\delta y'\right]dx = \int_a^b \left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right]\delta y dx + \frac{\partial F}{\partial y'}\delta y\Big|_a^b. \quad (3.8)$$

Нехай

$$\delta y_b = \tilde{y}(b+\delta b) - y(b) = \tilde{y}(b+\delta b) - \tilde{y}(b) + \tilde{y}(b) - y(b) \approx \tilde{y}'(b)\delta b + \delta y(b).$$

Тоді для лівої межі відрізку інтегрування отримуємо:

$$\delta y_a = \tilde{y}(a+\delta a) - y(a) \approx \tilde{y}'(a)\delta a + \delta y(a).$$

З врахуванням цих наближених виразів (3.8) можна записати у вигляді:

$$\int_a^b \left[\frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial y'}\delta y'\right]dx = \int_a^b \left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right]\delta y dx + \\ + \frac{\partial F}{\partial y'}\Big|_{x=b} \cdot (\delta y_b - \tilde{y}'(b)\delta b) - \frac{\partial F}{\partial y'}\Big|_{x=a} \cdot (\delta y_a - \tilde{y}'(a)\delta a).$$

Підставивши цей вираз у (3.7), отримаємо:

$$\varphi'(0) = \int_a^b \left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)\right]\delta y dx + \frac{\partial F}{\partial y'}\Big|_{x=b} \delta y_b + \\ + \left(F - y'\frac{\partial F}{\partial y'}\right)\Big|_{x=b} \delta b - \frac{\partial F}{\partial y'}\Big|_{x=a} \delta y_a - \left(F - y'\frac{\partial F}{\partial y'}\right)\Big|_{x=a} \delta a. \quad (3.9)$$

Отримали вираз, що є лінійним відносно варіацій $\delta y, \delta a, \delta y_a, \delta b, \delta y_b$.

Вираз $\varphi'(0) = \delta J(y, \delta y, \delta a, \delta y_a, \delta b, \delta y_b)$, що є лінійним відносно варіацій функції $y(x)$ та меж інтегрування, називають **варіацією функціонала** $J(y,a,b)$ (3.6) **у задачі з рухомими межами.**

Формула (3.9) для варіації функціонала у задачі з рухомими межами містить як окремі випадки:

– формулу для варіації функціонала у найпростішій задачі варіаційного числення (при $\delta a = \delta b = \delta y_a = \delta y_b = 0$);

– формулу для варіації функціонала, заданого на кривих, межі яких можуть переміщуватися вздовж вертикальних прямих ($\delta a = \delta b = 0$);

– формулу для варіації функціонала, заданого на кривих, межі яких можуть переміщуватися вздовж горизонтальних прямих ($\delta y_a = \delta y_b = 0$).

3.3 Умови трансверсальності

Нехай функціонал (3.6) задано на неперервно диференційовних функціях, межі графіків яких розміщені на фіксованих кривих $y = f(x)$ та $y = g(x)$, визначених на відрізку $[a_0; b_0]$. Прикладом варіаційної задачі з такою областю визначення функціонала є задача про знаходження відстані між двома кривими.

Якщо на деякій функції $y(x)$ функціонал (3.6) досягає екстремуму, то ця функція є точкою екстремуму серед всіх функцій, графіки яких мають з графіком $y(x)$ спільні межі. Звідси випливає, що функція $y(x)$ повинна задовольняти рівнянню Ейлера для функціонала (3.6), тобто повинна бути його екстремаллю. Тому у загальній формулі (3.9) для варіації функціонала (3.6) інтеграл у правій частині дорівнює нулю. Необхідна умова екстремуму функціонала (3.6) набуває вигляду:

$$\left.\frac{\partial F}{\partial y'}\right|_{x=b} \delta y_b + \left.\left(F - y'\frac{\partial F}{\partial y'}\right)\right|_{x=b} \delta b - \left.\frac{\partial F}{\partial y'}\right|_{x=a} \delta y_a - \left.\left(F - y'\frac{\partial F}{\partial y'}\right)\right|_{x=a} \delta a = 0. \quad (3.10)$$

Оскільки кінці графіків допустимих функцій знаходяться на фіксованих кривих $y = f(x)$ та $y = g(x)$, то варіації $\delta y_b = g'(b)\delta b$, $\delta y_a = f'(a)\delta a$. Тому умова (3.10) набуває вигляду:

$$\left.\left(F + (g' - y')\frac{\partial F}{\partial y'}\right)\right|_{x=b} \delta b - \left.\left(F + (f' - y')\frac{\partial F}{\partial y'}\right)\right|_{x=a} \delta a = 0.$$

Оскільки варіації δa та δb є незалежними, то отримуємо рівності:

$$\begin{cases} \left.\left(F + (g' - y')\dfrac{\partial F}{\partial y'}\right)\right|_{x=b} = 0; \\ \left.\left(F + (f' - y')\dfrac{\partial F}{\partial y'}\right)\right|_{x=a} = 0. \end{cases} \quad (3.11)$$

Крайові умови (3.11) називають **умовами трансверсальності**. Криву $y(x)$, що задовольняє умовам трансверсальності (3.11), називають **кривою, трансверсальною до кривих $y = g(x)$ та $y = f(x)$.**

Визначимо геометричний зміст умов трансверсальності для функціоналів виду $J(y) = \int_a^b A(x,y)\sqrt{1+y'^2}dx$, де $A(x,y)$ – неперервна двічі диференційовна функція своїх аргументів. Дослідимо цей функціонал на екстремум на множині неперервно диференційовних функцій, кінці графіків яких знаходяться на неперервно диференційовних кривих $y = f(x)$ та $y = g(x)$. Маємо:

$$F = A(x,y)\sqrt{1+y'^2}\ ;\ \frac{\partial F}{\partial y'} = A(x,y)\cdot\frac{y'}{\sqrt{1+y'^2}}\ ;$$

$$F + \left(g' - y'\right)\frac{\partial F}{\partial y'} = A(x,y)\frac{1+g'y'}{\sqrt{1+y'^2}}\ .$$

Звідси випливає, що для функціоналів даного типу умови трансверсальності можна записати у вигляді $y'(b) = -\dfrac{1}{g'(b)}$ та, відповідно, $y'(a) = -\dfrac{1}{f'(a)}$. Таким чином, умови трансверсальності екстремалей даного функціоналу до кривих $y = f(x)$ та $y = g(x)$ виражають умови їх ортогональності до цих кривих на відповідних межах відрізка інтегрування.

Розглянемо випадок, коли межі екстремалей функціонала (3.6) можуть переміщуватись вздовж кривих, заданих у неявній формі, тобто $h_1(a, y_a) = 0;\ h_2(b, y_b) = 0$.

Для кривої $h_1(x,y) = 0$ отримуємо $dy = -\dfrac{\partial h_1/\partial x}{\partial h_1/\partial y}dx$, звідки $\delta y_a = -\left.\dfrac{\partial h_1/\partial x}{\partial h_1/\partial y}\right|_{x=a}\delta a$. Аналогічно отримуємо $\delta y_b = -\left.\dfrac{\partial h_2/\partial x}{\partial h_2/\partial y}\right|_{x=b}\delta b$.

Підставляючи у (3.10), з врахуванням незалежності варіацій δa та δb отримаємо умови трансверсальності для даного випадку:

$$\begin{cases} \left.\left(\dfrac{\partial h_2}{\partial y}\left(F - y'\dfrac{\partial F}{\partial y'}\right) - \dfrac{\partial h_2}{\partial x}\dfrac{\partial F}{\partial y'}\right)\right|_{x=b} = 0; \\ \left.\left(\dfrac{\partial h_1}{\partial y}\left(F - y'\dfrac{\partial F}{\partial y'}\right) - \dfrac{\partial h_1}{\partial x}\dfrac{\partial F}{\partial y'}\right)\right|_{x=a} = 0. \end{cases} \tag{3.12}$$

Нехай кінці екстремалей можуть переміщуватися вздовж горизонтальних прямих $y_a = y_1,\ y_b = y_2$. Тоді $\delta y_a = \delta y_b = 0$ і з рівності (3.10) отримуємо умови трансверсальності для даного випадку у вигляді:

$$\left.\left(F - y'\frac{\partial F}{\partial y'}\right)\right|_{x=a} = \left.\left(F - y'\frac{\partial F}{\partial y'}\right)\right|_{x=b} = 0. \qquad (3.13)$$

У випадку, коли кінці екстремалей можуть переміщуватись вздовж вертикальних прямих $x = a$ та $x = b$, отримуємо $\delta a = 0$, $\delta b = 0$. Тоді з (3.10) випливають природні крайові умови, що вже були отримані у попередньому параграфі для даного випадку:

$$\left.\frac{\partial F}{\partial y'}\right|_{x=b} = \left.\frac{\partial F}{\partial y'}\right|_{x=a} = 0. \qquad (3.14)$$

Приклад 1. Знайти криву, що з'єднує точку $A(0;0)$ з прямою $y = 1$, для якої функціонал $J(y) = \int_0^b \left(xy' + y'^2\right)dx$ може досягати екстремуму.

Розв'язання. Знайдемо сімейство екстремалей даного функціонала. Для цього складемо та розв'яжемо рівняння Ейлера.

$$F = xy' + y'^2; \frac{\partial F}{\partial y} = 0; \frac{\partial F}{\partial y'} = x + 2y'; \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = 1 + 2y''.$$

Рівняння Ейлера має вигляд: $1 + 2y'' = 0$. Звідси знаходимо рівняння сімейства екстремалей функціонала: $y = -\frac{x^2}{4} + C_1 x + C_2$. З крайової умови $y(0) = 0$ знаходимо $C_2 = 0$. На правій межі інтегрування $y = b$ виконується умова трансверсальності (3.13): $\left.\left(F - y'\frac{\partial F}{\partial y'}\right)\right|_{x=b} = 0.$

При $x = b$ $y' = -\frac{b}{2} + C_1$, $F = b\left(-\frac{b}{2} + C_1\right) + \left(-\frac{b}{2} + C_1\right)^2$, $\frac{\partial F}{\partial y'} = b + 2\left(-\frac{b}{2} + C_1\right)$.

Підставляючи ці вирази у умову трансверсальності при $x = b$ та враховуючи, що $y(b) = 1$, отримаємо систему алгебраїчних рівнянь відносно невідомих C_1 та b:

$$\begin{cases} b\left(C_1 - \frac{b}{2}\right) + \left(C_1 - \frac{b}{2}\right)^2 - \left(C_1 - \frac{b}{2}\right)\left(b - b + 2C_1\right) = 0; \\ y(b) = -\frac{b^2}{4} + C_1 b = 1. \end{cases}$$

Розв'язуючи цю систему з врахуванням того, що $b > 0$, отримуємо $C_1 = 1$; $b = 2$. Таким чином, отримуємо рівняння екстремалі даної варіаційної задачі у вигляді $y = -\frac{x^2}{4} + x$.

Приклад 2. Знайти відстань між параболою $y = x^2$ та прямою $x - y = 5$ [2].

Розв'язання. Задача зводиться до знаходження екстремального значення функціонала

$$J[y(x)]=\int_{x_0}^{x_1}\sqrt{1+y'^2}dx \to extr \qquad (3.15)$$

при умові, що лівий кінець екстремалі може рухатися по кривій $y=x^2$, а правий – по прямій $y=x-5$. Таким чином, маємо $f(x)=x^2$, $g(x)=x-5$. Загальний розв'язок рівняння Ейлера – $y=C_1x+C_2$, де C_1 і C_2 – сталі, які треба визначити. Умови трансверсальності (3.11) мають вигляд

$$\left[\sqrt{1+y'^2}+(2x-y')\frac{y'}{\sqrt{1+y'^2}}\right]\Bigg|_{x=x_0}=0,$$

$$\left[\sqrt{1+y'^2}+(1-y')\frac{y'}{\sqrt{1+y'^2}}\right]\Bigg|_{x=x_1}=0.$$

Рівняння

$$y(x_0,C_1,C_2)=f(x_0),$$
$$y(x_1,C_1,C_2)=g(x_1)$$

набувають вигляду

$$C_1x_0+C_2=x_0^2,$$
$$C_1x_1+C_2=x_1-5.$$

Враховуючи, що $y'=C_1$, отримуємо систему чотирьох рівнянь з чотирьома невідомими C_1,C_2,x_0,x_1:

$$\begin{cases}\sqrt{1+C_1^2}+(2x_0-C_1)\dfrac{C_1}{\sqrt{1+C_1^2}}=0,\\ \sqrt{1+C_1^2}+(1-C_1)\dfrac{C_1}{\sqrt{1+C_1^2}}=0,\\ C_1x_0+C_2=x_0^2,\\ C_1x_1+C_2=x_1-5.\end{cases} \qquad (3.16)$$

Розв'язком системи (3.16) є: $C_1=-1$, $C_2=3/4$, $x_0=1/2$, $x_1=23/8$.

Отже, $y=-x+3/4$ і $J[y(x)]=\int_{1/2}^{23/8}\sqrt{1+(-1)^2}\,dx=\sqrt{2}\,x\Big|_{1/2}^{23/8}=\dfrac{19\sqrt{2}}{8}$.

Рівняння екстремалі $y=-x+3/4$. Відстань між параболою та прямою дорівнює $19\sqrt{2}/8$.

Питання для самоперевірки:

1. Сформулювати природні крайові умови для функціонала $J(y(x))$.

2. Надати геометричну інтерпретацію природних крайових умов.
3. Сформулювати загальну варіаційну задачу з рухомими межами.
4. Які крайові умови називають умовами трансверсальності?
5. Вкажіть етапи реалізації алгоритму розв'язування варіаційної задачі з рухомими межами.
6. Отримати умови трансверсальності для випадку, коли межі допустимих кривих можуть рухатися вздовж кривих, рівняння яких задано у неявній формі.
7. Сформулювати крайові умови для випадку, коли межі допустимих кривих можуть переміщуватися вздовж вертикальних прямих.
8. У чому полягає геометричний зміст умов трансверсальності для функціоналів виду $J(y) = \int_a^b A(x,y)\sqrt{1+y'^2}dx$?

Практичні завдання

1. Знайти найкоротшу відстань від точки $A(-1,3)$ до прямої $y = 1-3x$ [1].
2. Знайти найкоротшу відстань від точки $A(-1,5)$ до параболи $y^2 = x$ [1].
3. Знайти найкоротшу відстань між колом $x^2+y^2=1$ та прямою $x+y=4$ [1].
4. Знайти найкоротшу відстань від точки $A(1,0)$ до еліпсу $4x^2+9y^2=36$ [1].
5. Дослідити на екстремум функціонал $J[y(x)] = \int_0^{x_1} y^2 y'^2 dx$ при умові, що один кінець допустимих кривих є зафіксованим у точці $(0,0)$, а другий кінець рухається по прямій $y = \frac{1}{2}(3-x)$.
6. Дослідити на екстремум функціонал
$$J[y(x)] = \int_0^{x_1} \frac{\sqrt{1+y'^2}}{y} dx \to extr; \quad y(0)=0, \quad y_1 = x_1 - 5.$$
7. Знайти криву, на якій реалізується екстремум функціонала $J[y(x)] = \int_0^2 \left(xy' + y'^2\right)dx$ при умові, що лівий кінець її знаходиться у точці $A(0,1)$, а правий – на прямій $x = 2$.
8. Серед ліній, що з'єднують точку $O\ (0,0)$ з кривою $y^3 = 2-x$ знайти ту, яка дає мінімум функціонала $J[y(x)] = \frac{1}{2}\int_0^{x_1}\left(y'^2 - y^2\right)dx$.

4 ЕКСТРЕМАЛІ З КУТОВИМИ ТОЧКАМИ

Ключові поняття та терміни: кусочно гладка функція, кутова точка, умови Вейєрштрасса – Ердмана.

У варіаційному численні зустрічаються задачі, у яких область визначення функціонала не може обмежуватися неперервно диференційовними функціями. Безпосередньо з постановки задачі може випливати, що похідна функції, на якій функціонал досягає екстремуму, має розрив першого типу у деякій внутрішній точці $x = c$ проміжку інтегрування. У цій точці лівостороння похідна екстремалі не дорівнює її правосторонній похідній. У той же час і при відсутності додаткових умов для варіаційної задачі її функціонал може не мати екстремумів серед неперервно диференційовних функцій, а розв'язок задачі може мати кутові точки, у яких похідна екстремалі має розрив.

Знайдемо умови, яким повинні задовольняти точки екстремуму функціонала

$$J(y) = \int_a^b F(x, y, y')dx, \tag{4.1}$$

якщо його областю визначення є множина кусочно гладких на $[a;b]$ функцій. По-перше, кожна гладка ділянка функції, що надає екстремум функціоналу (4.1), повинна задовольняти рівнянню Ейлера для цього функціонала. По-друге, виходячи з необхідної умови екстремуму функціонала та розглядаючи при цьому кутову точку як рухому межу, можна зробити висновок про те, що у кожній кутовій точці $x = c$ повинні виконуватися умови:

$$\begin{cases} \left.\left(F - y'\dfrac{\partial F}{\partial y'}\right)\right|_{x=c-0} = \left.\left(F - y'\dfrac{\partial F}{\partial y'}\right)\right|_{x=c+0}; \\ \left.\dfrac{\partial F}{\partial y'}\right|_{x=c-0} = \left.\dfrac{\partial F}{\partial y'}\right|_{x=c+0}. \end{cases} \tag{4.2}$$

Тут запис $x = c - 0$ означає лівосторонню границю відповідного виразу у точці $x = c$, а запис $x = c + 0$ – відповідно правосторонню границю.

Умови (4.2), що повинні виконуватися у кутовій точці функціонала (4.1), називають **умовами Вейєрштрасса – Ердмана**.

Таким чином, алгоритм знаходження екстремалей функціонала (4.1) з кутовими точками складається з наступних кроків.

1. Для заданого функціонала записуємо умови Вейєрштрасса – Ердмана. Якщо з них випливає умова неперервності похідної екстремалі $y'(c-0) = y'(c+0)$, то екстремалі з кутовими точками відсутні.
2. У іншому разі складаємо рівняння Ейлера та знаходимо його загальні розв'язки на відрізках $[a;c]$ та $[c;b]$ – відповідно $y_1 = y_1(x, C_1, C_2)$ та $y_2 = y_2(x, C_3, C_4)$.

3. Визначаємо сталі інтегрування C_1, C_2, C_3, C_4 з заданих крайових умов $y_1(a) = y_A$, $y_2(b) = y_B$, умови неперервності екстремалі у кутовій точці $y_1(c) = y_2(c)$ та умов Вейєрштрасса – Ердмана. У результаті отримуємо екстремаль:

$$y = \begin{cases} y_1(x), x \in [a;c]; \\ y_2(x), x \in [c;b]. \end{cases}$$

Приклад 1. Знайти екстремалі з кутовою точкою для функціонала $J(y) = \int_0^4 y'^2 \left(y' - 2\right)^2 dx$, якщо $y(0) = 0,\ y(4) = 4$.

Розв'язання. Запишемо умови Вейєрштрасса–Ердмана для даного функціонала.

$$\frac{\partial F}{\partial y'} = 2y'\left(y' - 2\right)^2 + 2y'^2(y' - 2) = 4y'\left(y' - 1\right)\left(y' - 2\right);$$

$$F - y'\frac{\partial F}{\partial y'} = y'^2\left(y' - 2\right)^2 - 4y'^2\left(y' - 1\right)\left(y' - 2\right) = y'^2\left(y' - 2\right)\left(2 - 3y'\right).$$

Умови Вейєрштрасса – Ердмана набувають вигляду:

$$\begin{cases} y'^2\left(y' - 2\right)\left(2 - 3y'\right)\Big|_{x=c-0} = y'^2\left(y' - 2\right)\left(2 - 3y'\right)\Big|_{x=c+0}; \\ 4y'\left(y' - 1\right)\left(y' - 2\right)\Big|_{x=c-0} = 4y'\left(y' - 1\right)\left(y' - 2\right)\Big|_{x=c+0}. \end{cases}$$

Ці рівності можуть виконуватися у одному з трьох випадків:

1. $y'\big|_{x=c-0} = y'\big|_{x=c+0}$. У цьому випадку похідна екстремалі у точці $x = c$ є неперервною, тому кутові точки відсутні. Оскільки інтегральний вираз у функціоналі залежить тільки від y', то його екстремаль має вигляд $y = C_1 x + C_2$. Сталі інтегрування знаходимо з крайових умов: $y(0) = C_2 = 0$, $y(4) = 4C_1 = 4 \Rightarrow C_1 = 1$. Отже, рівняння екстремалі на $[0;4]$ має вигляд $y = x$.

2. $y'\big|_{x=c-0} = 0; y'\big|_{x=c+0} = 2$. При $x \in [0;c]$ рівняння екстремалі має вигляд $y_1(x) = C_1 x + C_2$, при $x \in [c;4]$ рівняння екстремалі запишемо у вигляді $y_2(x) = C_3 x + C_4$. З крайової умови при $x = 0$ знаходимо, що $y_1(0) = C_2 = 0$. Маємо $y_1(x) = C_1 x$. Для прямої $y_2(x)$ отримуємо: $y_2(4) = 4C_3 + C_4 = 4 \Rightarrow C_4 = 4\left(1 - C_3\right)$; $y_2 = C_3 x + 4(1 - C_3)$; $y'\big|_{x=c-0} = y_1' = C_1 = 0$; $y'\big|_{x=c+0} = C_3 = 2$. Підставивши знайдені значення сталих інтегрування у рівняння екстремалей, отримаємо $y_1 \equiv 0,\ y_2 = 2x - 4$. Знайдемо кутову точку c: $y_1(c) = y_2(c) \Rightarrow 0 = 2c - 4 \Rightarrow c = 2$. Екстремаль з кутовою точкою має вигляд:

$$y(x) = \begin{cases} 0,\ x \in [0;2]; \\ 2x - 4,\ x \in [2;4]. \end{cases}$$

3. $y'\big|_{x=c-0} = 2,\ y'\big|_{x=c+0} = 0.$ Аналогічно попередньому випадку при $x \in [0;c]$ рівняння екстремалі запишемо у вигляді $y_1(x) = C_1 x + C_2$, при $x \in [c;4]$ маємо

екстремаль $y_2(x) = C_3x + C_4$. Знайдемо значення сталих інтегрування. $y_1(0) = 0 \Rightarrow C_2 = 0 \Rightarrow y_1 = C_1x$. З умови $y_2(4) = 4$ знаходимо $C_4 = 4(1 - C_3)$. Тоді $y_2(x) = C_3x + 4(1 - C_3)$. $y'\big|_{x=c-0} = y_1' = C_1 = 2$. Тому $y_1 = 2x$.

Для прямої $y_2(x)$ $y_2'\big|_{x=c+0} = C_3 = 0 \Rightarrow y_2 \equiv 4$. Використовуючи умову неперервності екстремалі при $x = c$, знайдемо кутову точку: $y_1(c) = y_2(c) \Rightarrow 2c = 4 \Rightarrow c = 2$. Отримали екстремаль з кутовою точкою:

$$y = \begin{cases} 2x, x \in [0;2]; \\ 4, x \in [2;4]. \end{cases}$$

На неперервно диференційовній екстремалі $y = x$ значення функціонала $J(y)$ дорівнює $\int_0^4 dx = 4$, на екстремалях з кутовою точкою $J(y) = 0$. Оскільки підінтегральна функція для даного функціоналу є невід’ємною, то $J(y) \geq 0$, тому мінімум функціонала досягається на екстремалях з кутовою точкою.

Приклад 2. Відшукати ламані екстремалі функціонала

$$J[y(x)] = \int_0^2 \left(y'^4 - 6y'^2\right)dx, \quad y(0) = 0, \quad y(2) = 0,$$

у припущенні, що y' може мати одну точку злому, яка відповідає абсцисі $x = c$ [1].

Розв’язання. Оскільки $F_{y'y'} = 12y'^2 - 12$ може перетворюватися в нуль, то наявність зломів екстремалі є можливою. Підінтегральна функція залежить лише від y', тому екстремалями є прямі $y = C_1x + C_2$.

Нехай

$$y_- = mx + n \quad (0 \leq x < c), \quad y_+ = px + q \quad (c \leq x \leq 2).$$

З граничних умов знаходимо $n = 0$, $q = -2p$, так що

$$y_- = mx, \quad y_+ = p(x - 2). \tag{4.3}$$

Умова неперервності екстремалі дає

$$mc = p(c - 2). \tag{4.4}$$

Умови Вейєрштрасса–Ердмана набувають вигляду:

$$F_{y'} = 4y'^3 - 12y',$$

$$F - y' \cdot F_{y'} = -3y'^4 + 6y'^2.$$

Оскільки $y'_- = m$, $y'_+ = p$, отримаємо

$$\begin{cases} 4m^3 - 12m = 4p^3 - 12p, \\ -3m^4 + 6m^2 = -3p^4 + 6p^2, \end{cases}$$

або

$$\begin{cases} (m-p)\,\left(m^2+mp+p^2-3\right)=0, \\ \left(m^2-p^2\right)\,\left(m^2+p^2-2\right)=0. \end{cases} \tag{4.5}$$

Друге рівняння в (4.5) відразу дає $m=p$ або $m=-p$ або $m^2+p^2-2=0$. Розв'язок $m=p$ повинен бути відкинутий: при ньому екстремаль має неперервну похідну, а із умови (4.4) отримуємо, що $m=0$, тобто екстремаль – відрізок осі Ox.

Таким чином, розв'язання системи (4.5) зводиться до розв'язання наступних систем рівнянь:

$$\begin{cases} m=-p, \\ m^2+mp+p^2=3 \end{cases} \tag{4.6}$$

і

$$\begin{cases} m^2+p^2=2, \\ m^2+mp+p^2=3. \end{cases} \tag{4.7}$$

Розв'язок системи (4.6): $m=\sqrt{3}$, $p=-\sqrt{3}$ і $m=-\sqrt{3}$, $p=\sqrt{3}$. Розв'язок системи (4.7) $m=p$ повинен бути відкинутий. Отже, $m=-p$ і умова неперервності (4.4) дає $c=1$.

Ламані екстремалі з однією точкою злому мають вигляд:

$$y_1(x)=\begin{cases} \sqrt{3}\,x, & 0\le x<1, \\ -\sqrt{3}(x-2), & 1\le x\le 2 \end{cases} \quad \text{і} \quad y_2(x)=\begin{cases} -\sqrt{3}\,x, & 0\le x<1, \\ \sqrt{3}(x-2), & 1\le x\le 2. \end{cases}$$

$$J\left[y_1(x)\right]=J\left[y_2(x)\right]=-18.$$

Питання для самоперевірки

1. Які криві називають екстремалями з кутовими точками?
2. Сформулювати умови Вейєрштрасса – Ердмана.
3. Які етапи передбачає алгоритм знаходження екстремалей з кутовими точками?

Практичні завдання

Знайти ламані екстремалі (якщо вони існують) для функціоналів [1-3]:

1. $J\left[y(x)\right]=\int\limits_0^4\left(y'^4-2y'^2\right)dx\to extr,\quad y(0)=0,\quad y(4)=2.$

2. $J\left[y(x)\right]=\int\limits_{x_0}^{x_1}\left(y'^2+2xy-y^2\right)dx\to extr,\quad y(x_0)=y_0,\quad y(x_1)=y_1.$

3. $J[y(x)]=\int\limits_0^2 y'^2\left(y'-1\right)^2 dx \to extr, \quad y(0)=0, \quad y(2)=1.$

4. $J[y(x)]=\int\limits_0^4 \left(y'-1\right)^2\left(y'+1\right)^2 dx \to extr, \quad y(0)=0, \quad y(4)=2.$

5. $J[y(x)]=\int\limits_{x_0}^{x_1} \left(y'^2+4x^2y-3y^2\right)dx \to extr, \quad y(x_0)=y_0, \quad y(x_1)=y_1.$

6. $J[y(x)]=\int\limits_{-1}^{1} y^2\left(1-y'\right)^2 dx \to extr, \quad y(-1)=0, \quad y(1)=1.$

7. $J[y(x)]=\int\limits_0^{\pi/4} \left(y'^2-6yy'-16y^2\right)dx \to extr, \quad y(0)=0, \quad y\left(\frac{\pi}{4}\right)=1.$

8. $J[y(x)]=\int\limits_3^4 \frac{\left(yy'\right)^2}{4+y^2} dx \to extr, \quad y(3)=0, \quad y(4)=\sqrt{5}.$

9. $J[y(x)]=\int\limits_0^a \left(3y'^2-y^2\right)dx \to extr, \quad y(0)=0, \quad y(a)=b.$

5 ДОСТАТНІ УМОВИ ЕКСТРЕМУМУ ФУНКЦІОНАЛА

План:
5.1 Поле екстремалей.
5.2 Достатня умова Вейєрштрасса.
5.3 Достатні умови Лежандра.

Ключові поняття та терміни: поле екстремалей, нахил поля у точці, центральне поле екстремалей, умова Якобі, функція Вейєрштрасса, умова Лежандра, достатні умови слабкого екстремуму, достатні умови сильного екстремуму

5.1 Поле екстремалей

Якщо на координатній площині xOy через кожну точку деякої області D проходить одна і тільки одна крива сімейства $y = y(x, C)$, то говорять, що дане сімейство кривих утворює у області D **поле (власне поле)**. Кутовий коефіцієнт $p(x, y)$ дотичної до кривої сімейства $y = y(x, C)$, що проходить через точку (x, y), називають **нахилом поля** у точці (x, y). Наприклад, всередині круга $x^2 + y^2 \le a^2$ прямі $y = 2x + C$ утворюють поле з нахилом $p(x, y) = 2$. Сімейство парабол всередині цього ж круга поля не утворює, оскільки всередині цієї області дані параболи перетинаються.

Якщо всі криві сімейства $y = y(x, C)$ проходять через деяку точку (x_0, y_0) області D, тобто утворюють пучок кривих з центром у даній точці, то вони не утворюють власного поля у цій області. Проте, якщо ці криві покривають область D повністю, перетинаючись у ній тільки в центрі пучка, то кажуть, що сімейство кривих $y = y(x, C)$ утворює **центральне поле** з центром у точці (x_0, y_0). Наприклад, пучок синусоїд $y = C\sin x$ при $0 \le x \le a, a < \pi$ утворює центральне поле. Той же пучок при $0 \le x \le a, a > \pi$ поля не утворює.

Якщо власне або центральне поле утворене сімейством екстремалей деякої варіаційної задачі, то його називають **полем екстремалей**.

Нехай крива $y = y(x)$ є екстремаллю варіаційної задачі про екстремум функціонала

$$J(y) = \int_a^b F(x, y, y')dx \tag{5.1}$$

і при цьому $y(a) = y_0$, $y(b) = y_1$. Екстремаль $y = y(x)$, яка проходить через точки $A(a, y_0)$ та $B(b, y_1)$, входить до поля екстремалей, якщо сімейство екстремалей даного функціонала, що утворює поле, при деякому значенні $C = C_0$ містить екстремаль $y = y(x)$. При цьому дана екстремаль не повинна міститися на межі області D, у якій сімейство екстремалей утворює поле. Якщо пучок екстремалей з центром у точці A утворює поле екстремалей у околі екстремалі $y = y(x)$, що проходить через цю точку, то маємо центральне поле, яке містить дану екстремаль $y(x)$. За параметр сімейства екстремалей у даному випадку можна взяти кутовий коефіцієнт дотичної до кривих пучка у точці A.

Приклад. Для функціоналу $J(y) = \int_0^a (y'^2 - y^2)dx$, де $y(0) = y(a) = 0$ перевірити, чи входить дуга екстремалі $y = 0$ у центральне поле екстремалей з центром у точці $O(0,0)$, якщо $a < \pi$.

Розв'язання. Рівняння Ейлера для даного функціонала має вигляд $y'' + y = 0$. Звідси знаходимо його розв'язок $y = C_1 \cos x + C_2 \sin x$. З крайових умов отримуємо $y(0) = 0 \Rightarrow C_1 = 0,\ y = C_2 \sin x.$

Якщо $a < \pi$, то $y(a) = C_2 \sin a = 0 \Rightarrow C_2 = 0$, $y = 0$ – екстремаль, що входить до центрального поля екстремалей, утвореного кривими пучка $y = C_2 \sin x$.

Якщо $a \geq \pi$, то сімейство екстремалей $y = C_2 \sin x$ поля не утворює, оскільки на відрізку інтегрування $\left[0;a\right]$ ці криві перетинаються у точках, відмінних від центру пучка.

Можна довести умову (наприклад, [3]), яка дає змогу встановити входження екстремалі функціонала до складу поля:

<u>Теорема 5.1</u> Нехай $\left[a;b\right]$ – відрізок інтегрування функціонала (5.1). Якщо нетривіальний розв'язок $u(x)$ рівняння

$$\left(\frac{\partial^2 F}{\partial y^2} - \frac{d}{dx}\left(\frac{\partial^2 F}{\partial y \partial y'}\right)\right) \cdot u - \frac{d}{dx}\left(\frac{\partial^2 F}{\partial y' \partial y'} u'\right) = 0, \tag{5.2}$$

який дорівнює нулю при $x = a$, не обертається в нуль у якій-небудь іншій точці відрізка $\left[a;b\right]$, то екстремаль входить до центрального поля екстремалей з центром у точці $(a; y_0)$.

Дану умову входження екстремалі до центрального поля екстремалей називають **умовою Якобі**, а рівняння (5.2) – **рівнянням Якобі**.

Приклад. Встановити, чи входить до центрального поля екстремалей екстремаль функціонала $J(y) = \int_0^a (y'^2 + y^2 + 2x)dx$, якщо $y(0) = 0,\ y(a) = 0$.

Розв'язання. Знайдемо екстремаль даного функціонала, що задовольняє крайовим умовам. Рівняння Ейлера має вигляд $y'' - y = 0$. Його загальним розв'язком є сімейство екстремалей $y = C_1 \operatorname{ch} x + C_2 \operatorname{sh} x$. Для знаходження сталих C_1 та C_2 використаємо краєві умови. $y(0) = 0 \Rightarrow C_1 = 0$, $y(a) = 0 \Rightarrow C_2 \operatorname{sh} a = 0 \Rightarrow C_2 = 0$ при $a \neq 0$. Отримали екстремаль $y = 0$.

Запишемо рівняння Якобі. Знайдемо необхідні для цього частинні похідні:

$$\frac{\partial F}{\partial y}=2y,\quad \frac{\partial^2 F}{\partial y^2}=2,\quad \frac{\partial^2 F}{\partial y\partial y'}=0,\quad \frac{\partial F}{\partial y'}=2y',\quad \frac{\partial^2 F}{\partial y'^2}=2.$$

Підставляючи у (5.2), отримуємо рівняння Якобі $2u-2u''=0$. Тоді $u=C_1\text{ch}\,x+C_2\text{sh}\,x$. Оскільки $u(0)=0$, то $C_1=0, u=C_2\text{sh}\,x$. Криві пучка $u=C_2\text{sh}\,x$ перетинають $[0;a]$ лише у точці $(0;0)$ при $C_2\neq 0$, $a>0$. Тому умова Якобі виконана, екстремаль $y=0$ входить до центрального поля екстремалей з центром у точці $(0;0)$.

5.2 Достатня умова Вейєрштрасса

Дослідження достатніх умов існування екстремуму функціонала є більш складними, проте, у деякій мірі аналогічними дослідженню на екстремум функцій багатьох змінних.

Розглянемо функціонал

$$J\left[y(x)\right]=\int_{x_0}^{x_1}F(x,y,y')dx, \tag{5.3}$$

де допустимі функції задовольняють крайовим умовам $y(x_1)=y_1, y(x_2)=y_2$. Нехай функція $y=y(x)$ є екстремаллю, що входить до центрального поля екстремалей, тобто для неї виконана умова Якобі. Нахил дотичної до графіка цієї функції у точках з координатами $(x;y)$, тобто нахил поля, має значення $\frac{dy}{dx}=p(x,y)$. На цій екстремалі підінтегральна функція у (5.3) має вигляд $F(x,y,p)$. Зміна значення величини $y'=\frac{dy}{dx}$ спричиняє зміну значення підінтегрального виразу $F(x,y,y')$, при цьому здійснюється перехід на наступну допустиму екстремаль. У цьому випадку частинний приріст функції $F(x,y,y')$ за змінною y' позначимо $\Delta_{y'}F(x,y,y')$. Цей приріст можна представити у вигляді:

$$\Delta_{y'}F(x,y,p)=F(x,y,y')-F(x,y,p)=(y'-p)\left.\frac{\partial F}{\partial y'}\right|_{y'=p}+\beta\cdot\Delta y',$$

де $\Delta y'$ – приріст похідної y', при $\beta\to 0$ $\Delta y'\to 0$.

Позначимо $\beta\cdot\Delta y'=E(x,y,p,y')$. Тоді з попередньої рівності отримаємо

$$E(x,y,p,y')=F(x,y,y')-F(x,y,p)-(y'-p)\left.\frac{\partial F}{\partial y'}\right|_{y'=p}. \tag{5.4}$$

Введену таким чином функцію $E(x,y,p,y')$ називають **функцією Вейєрштрасса** для функціонала (5.3). Для визначення знаку приросту $\Delta J\left[y(x)\right]$

цього функціонала при переході від екстремалі γ з рівнянням $y=y(x)$ до деякої близької допустимої екстремалі Γ, запишемо його у вигляді:

$$\Delta J\left[y(x)\right]=\int\limits_{\Gamma}F(x,y,y')dx-\int\limits_{\gamma}F(x,y,y')dx.$$

Тут криволінійні інтеграли виражають значення функціонала (5.3), взяті відповідно вздовж кривих Γ та γ. Якщо тепер перейти від кривої γ до кривої Γ, що проходить через нерухомі точки з координатами $(x_1;y_1)$ та $\left(x_2;y_2\right)$, то функцію $F(x,y,y')$ можна представити за допомогою ряду Тейлора:

$$F(x,y,y')=F(x,y,p)+(y'-p)\left.\frac{\partial F}{\partial y'}\right|_{y'=p}+\ldots,$$

тому приріст функціонала з точністю до відкинутих членів ряду Тейлора можна представити у вигляді:

$$\Delta J\left[y\left(x\right)\right]=\int\limits_{\Gamma}\left[F(x,y,y')-F(x,y,p)-(y'-p)\left.\frac{\partial F}{\partial y'}\right|_{y'=p}\right]dx.$$

Використовуючи позначення функції Вейєрштрасса та переходячи від криволінійного до визначеного інтеграла, отримаємо:

$$\Delta J\left[y(x)\right]=\int\limits_{x_0}^{x_1}E(x,y,p,y')dx.$$

Очевидно, що достатньою умовою мінімуму функціонала (5.3) на кривій γ буде умова невід'ємності функції Вейєрштрасса $E(x,y,p,y')$, оскільки якщо $E(x,y,p,y')\geq 0$, то й $\Delta J\geq 0$. Відповідно, достатня умова максимуму виражатиметься нерівністю $E\leq 0$, оскільки у цьому випадку $\Delta J\leq 0$. При цьому для слабкого мінімуму достатньо, щоб нерівність $E(x,y,p,y')\geq 0$ ($E(x,y,p,y')\leq 0$ для слабкого максимуму) виконувалася для всіх значень x, y, близьких до значень x,y на досліджуваній екстремалі, та для значень y', близьких до $p(x,y)$ на тій же екстремалі, тобто для кривих, близьких до даної екстремалі за нормою простору $C^1_{[a;b]}$. Для сильного мінімуму (максимуму) ця ж нерівність повинна виконуватися для кривих, близьких до даної екстремалі за нормою простору $C_{[a;b]}$, тобто для кривих, що є близькими за ординатами, при довільних y'.

Таким чином, достатніми для досягнення функціоналом $J(y)$ слабкого екстремуму на екстремалі γ з рівнянням $y=y(x)$ є наступні умови:

1. Крива γ є екстремаллю, що задовольняє заданим крайовим умовам.
2. Екстремаль γ входить до центрального поля екстремалей, тобто виконується умова Якобі.

3. Функція Вейєрштрасса $E(x, y, p, y')$ не змінює знаку у всіх точках (x, y), близьких до відповідних точок кривої γ та для близьких до $p(x, y)$ значень y'. У випадку мінімуму $E \geq 0$, для максимуму $E \leq 0$.

Для випадку сильного екстремуму функція Вейєрштрасса $E(x, y, p, y')$ не змінює знаку у всіх точках (x, y), близьких до відповідних точок кривої γ та для довільних значень y'.

Сформульовану достатню умову екстремуму функціонала називають **достатньою умовою Вейєрштрасса**.

Приклад. Дослідити на екстремум функціонал $J\left(y(x)\right) = \int\limits_0^2 \left(xy' + y'^2\right) dx,$ якщо $y(0) = 1,\ y(2) = 0$.

Розв'язання. Знайдемо екстремаль функціонала, що задовольняє крайовим умовам. Рівняння Ейлера для даного функціонала має вигляд $y'' = -\dfrac{1}{2}$. Звідси отримуємо рівняння сімейства екстремалей $y = -\dfrac{x^2}{4} + C_1 x + C_2$. $y(0) = C_2 = 1$, $y(2) = -1 + 2C_1 + 1 = 0$; $C_1 = 0$. Отримали рівняння екстремалі $y = -\dfrac{x^2}{4} + 1$. Перевіримо виконання умови Якобі. Запишемо рівняння Якобі. $F = xy' + y'^2$, $\dfrac{\partial^2 F}{\partial y^2} = \dfrac{\partial^2 F}{\partial y \partial y'} = 0$, $\dfrac{\partial^2 F}{\partial y' \partial y'} = 2$. Підставивши отримані частинні похідні у (5.2), знаходимо $\dfrac{d}{dx}\left(2u'\right) = 0$, звідки отримуємо: $u = C_1 x + C_2$. $u(0) = 0 \Rightarrow C_2 = 0$, тому $u = C_1 x \neq 0$ на $\left(0;2\right]$ при $C_1 \neq 0$. Умова Якобі виконується, тому екстремаль входить до центрального поля.

Запишемо для даного функціоналу функцію Вейєрштрасса. $F(x, y, p) = xp + p^2$. $F'_p = x + 2p$. Функція (5.4) набуває вигляду:

$$E(x, y, p, y') = xy' + y'^2 - xp - p^2 - (y' - p)(x + 2p) = (y' - p)^2 \geq 0.$$

При цьому знак функції Вейєрштрасса зберігається для довільних значень y'. На екстремалі $y = -\dfrac{x^2}{4} + 1$ функціонал досягає сильного мінімуму.

5.3 Достатня умова Лежандра

Дослідження знаку функції Вейєрштрасса пов'язане з деякими труднощами. Можна показати, що її знак співпадає з знаком виразу $\dfrac{\partial^2 F}{\partial \left(y'\right)^2}$, тобто зі знаком другої похідної функції $F(x, y, y')$ за змінною y' на екстремалі

$y(x)$. Для цього допустимо, що функція $F = F(x, y, y')$ є тричі диференційовною за змінною y'. Тоді за формулою Тейлора отримуємо:

$$F(x, y, y') = F(x, y, p) + (y' - p)\frac{\partial F}{\partial y'}\bigg|_{y'=p} + \frac{(y' - p)^2}{2}\frac{\partial^2 F}{\partial (y')^2}\bigg|_{y'=q},$$

де величина q знаходиться між p та y'. Тоді, за означенням функції Вейєрштрасса $E(x, y, p, y')$ можна записати рівність

$$E(x, y, p, y') = \frac{(y' - p)^2}{2}\frac{\partial^2 F}{\partial (y')^2}\bigg|_{y'=q}.$$

Звідси випливає, що функція $E(x, y, p, y')$ зберігає свій знак, якщо зберігає знак величина $\frac{\partial^2 F}{\partial (y')^2}\bigg|_{y'=p} = F_{y'y'}(x, y, q)$.

При дослідженні на мінімум умову $E(x, y, p, y') \ge 0$ можна замінити умовою $F_{y'y'}(x, y, q) \ge 0$ на екстремалі $y = y(x)$, а при дослідженні на максимум умова $E(x, y, p, y') \le 0$ може бути заміненою умовою $F_{y'y'}(x, y, q) \le 0$ на екстремалі $y = y(x)$. Такі умови називаються **умовами Лежандра**.

Приклад 1. Дослідити на екстремум функціонал $J = \int\limits_0^{1/2} \left(y'^2 + 2yy' - 16y^2 \right) dx$, якщо $y(0) = y(1/2) = 0$.

Розв'язання. Рівняння Ейлера для даного функціонала має вигляд $y'' + 16y = 0$. Його загальний розв'язок $y = C_1 \cos 4x + C_2 \sin 4x$. $y(0) = C_1 = 0$, $y(1/2) = C_2 \sin 2 = 0$, тому $C_2 = 0$. Отримуємо екстремаль $y = 0$, для якої перевіримо виконання умови Якобі. Запишемо для даного функціонала рівняння Якобі $u'' + 16u = 0$.

Звідси $u = A_1 \cos 4x + A_2 \sin 4x$. $u(0) = 0 \Rightarrow A_1 = 0$. $u = A_2 \sin 4x \ne 0$ при $A_2 \ne 0$, $x \in (0; 1/2]$. Отже, умова Якобі виконується. $F = y'^2 + 2yy' - 16y^2$, $F_{y'y'} = 2 > 0$ для довільного y', тому на екстремалі $y = 0$ функціонал досягає сильного мінімуму.

Приклад 2. Дослідити на екстремум функціонал

$J\left[y(x) \right] = \int\limits_0^1 \left(y'^3 - \alpha y' \right) dx \to extr$, $y(0) = 0,\ y(1) = -2$, ($\alpha -$ будь-яке дійсне число).

Розв'язання. Оскільки підінтегральна функція залежить лише від y', то екстремалями є прямі $y = C_1 x + C_2$. Екстремаллю, що задовольняє граничні умови, буде пряма $y = -2x$, яка належить до центрального поля екстремалей $y = Cx$. На екстремалі $y = -2x$ нахил поля $p = -2$. Знайдемо $F_{y'y'} = 6y'$. На екстремалі $y=-2x$ маємо $F_{y'y'}=-12<0$, тобто на лінії $y=-2x$ досягається слабкий максимум функціонала. Для довільних значень y' знак виразу $F_{y'y'}=6y'$ не зберігається, отже, достатня умова сильного максимуму не виконується.

Функція Вейєрштрасса $E(x,y,p,y')$ має вигляд

$$E(x,y,p,y')=\left(y'-p\right)^2\left(y'+2p\right).$$

Для довільних значень y' знак виразу $E(x,y,p,y')$ не зберігається, отже, достатня умова сильного максимуму не виконується. Для значень y', близьких до $p=-2$, функція $E(x,y,p,y')\le 0$, отже, на екстремалі $y=-2x$ досягається слабкий максимум.

На екстремалі $y=-2x$ функціонал досягає слабкого максимуму.

Питання для самоперевірки

1. Надати означення власного поля, утвореного сімейством кривих.
2. Що називають нахилом поля у точці?
3. Сформулювати означення центрального поля.
4. Надати означення поля екстремалей.
5. Записати рівняння Якобі.
6. Сформулювати умову Якобі включення екстремалей до центрального поля.
7. Записати функцію Вейєрштрасса.
8. Сформулювати достатню умову Вейєрштрасса для слабкого мінімуму функціонала.
9. Сформулювати достатню умову Вейєрштрасса для сильного мінімуму функціонала.
10. Навести формулювання достатньої умови Вейєрштрасса для сильного та слабкого максимумів функціонала.
11. Сформулювати достатню умову Лежандра для дослідження функціоналу на мінімум та максимум.

Практичні завдання

Дослідити на екстремум функціонали [1,3]:

1. $J\left[y(x)\right]=\int\limits_0^2\left(xy'+y'^2\right)\,dx\to extr,\quad y(0)=1,\quad y(2)=0.$

2. $J\left[y(x)\right]=\int\limits_0^a\left(y'^2+2yy'-16y^2\right)dx\to extr,\ a>0,\quad y(0)=0,\quad y(a)=0.$

3. $J[y(x)] = \int\limits_1^2 y'\left(1+x^2 y'\right)\, dx \to extr, \quad y(1)=3, \quad y(2)=5$.

4. $J[y(x)] = \int\limits_0^{\pi/4} \left(4y^2 - y'^2 + 8y\right)\, dx \to extr, \quad y(0)=-1, \quad y\left(\pi/4\right)=0$.

5. $J[y(x)] = \int\limits_1^2 \left(x^2 y'^2 + 12y^2\right)\, dx \to extr, \quad y(1)=1, \quad y(2)=8$.

6. $J[y(x)] = \int\limits_0^1 (1+x)\mathrm{y}'^2\, dx \to extr, \quad y(0)=0, \quad y(1)=1$.

7. $J[y(x)] = \int\limits_0^a \frac{dx}{y'} \to extr, \quad y(0)=0, \quad y(a)=b, \quad a>0, b>0$.

8. $J[y(x)] = \int\limits_1^2 \left(xy'^4 - 2yy'^3\right)\, dx \to extr, \quad y(1)=0, \quad y(2)=1$.

9. $J[y(x)] = \int\limits_0^a \left(1 - e^{-y'^2}\right)\, dx \to extr, \quad y(0)=0, \quad y(a)=b, \quad a>0$.

10. $J[y(x)] = \int\limits_0^a \frac{\sqrt{1+y'^2}}{\sqrt{y}}\, dx \to extr, \quad y(0)=0, \quad y(a)=y_1$.

11. $J[y(x)] = \int\limits_{-1}^2 y'\left(1+x^2 y'\right)\, dx \to extr, \quad y(-1)=1, \quad y(2)=4$.

6 ВАРІАЦІЙНІ ЗАДАЧІ НА УМОВНИЙ ЕКСТРЕМУМ

План:
6.1 Основні типи варіаційних задач на умовний екстремум.
6.2 Необхідні умови екстремуму функціонала у задачі Лагранжа.
6.3 Ізопериметрична задача.

Ключові поняття та терміни: цільовий функціонал, голономні зв'язки, диференціальні зв'язки, ізопериметричні зв'язки, задача Лагранжа, функція Лагранжа, множники Лагранжа, ізопериметрична задача.

6.1 Основні типи варіаційних задач на умовний екстремум

Варіаційні задачі, що були розглянуті раніше, характеризувалися тим, що їх розв'язки повинні були задовольняти крайовим умовам, проте у багатьох задачах, пов'язаних з дослідженням функціоналів на екстремум, на допустимі функції накладаються і інші додаткові умови, які називають зв'язками.

Нехай потрібно знайти екстремум функціонала

$$J(\bar{y}) = \int_a^b F(x, \bar{y}, \bar{y}')dx \qquad (6.1)$$

на множині неперервно диференційовних вектор-функцій $\bar{y}(x) = \left(y_1(x), y_2(x), ..., y_n(x)\right)$, що задовольняють крайовим умовам

$$y_1(a) = y_{11}, y_2(a) = y_{21}, ..., y_n(a) = y_{n1},\ y_1(b) = y_{12}, y_2(b) = y_{22}, ..., y_n(b) = y_{n2}, \qquad (6.2)$$

а також умовам зв'язків. Умови зв'язків можуть задовольняти як диференціальним співвідношенням

$$g_i(x, y_1, y_2, ..., y_n, y_1', y_2', ..., y_n') = 0; i = 1, 2, ..., k; k < n, \qquad (6.3)$$

так і інтегральним рівностям

$$\int_a^b h_j(x, y_1, y_2, ..., y_n, y_1', y_2', ..., y_n')dx = L_j,\ j = 1, 2, ..., s. \qquad (6.4)$$

При цьому похідні у виразах (6.3) та (6.4) можуть бути відсутніми. Для зв'язків (6.3) вважаємо, що функції $g_i, i = 1, 2, ..., k$ є неперервно диференційовними за аргументами $y_1', y_2', ..., y_n'$, а ранг матриці Якобі за цими змінними є максимальним та дорівнює k. Вважається також, що функції F, g_i, h_j є двічі диференційованими, а крайові умови (6.2) узгоджені з умовами зв'язку.

Якщо у співвідношеннях (6.3) функції g_i не залежать від похідних, то такі зв'язки називають **фазовими обмеженнями або голономними зв'язками**.

Функціонал (6.1) називають **цільовим функціоналом**, диференціальні співвідношення (6.3) – **диференціальними зв'язками**, а співвідношення (6.4) – **інтегральними або ізопериметричними зв'язками**.

Дану задачу, сформульовану у загальному вигляді, називають **варіаційною задачею на умовний екстремум**. Окремий випадок цієї задачі, коли на допустиму вектор-функцію накладені лише зв'язки виду (6.3), називають **задачею Лагранжа**. Варіаційну задачу на умовний екстремум з інтегральними зв'язками (6.4) називають **ізопериметричною задачею**.

Ізопериметричну задачу (6.1), (6.2), (6.4) можна звести до задачі Лагранжа, ввівши нові функції

$$\varphi_i(x) = \int_a^x h_i(t, y_1, y_2, ..., y_n, y_1', y_2', ..., y_n')dt, i = 1, 2, ..., s. \tag{6.5}$$

Тоді замість інтегральних співвідношень (6.4) отримаємо диференціальні зв'язки:

$$\varphi_i'(x) = h_i(x, y_1, y_2, ..., y_n, y_1', y_2', ..., y_n') = 0; i = 1, 2, ..., s. \tag{6.6}$$

6.2 Необхідні умови екстремуму функціонала у задачі Лагранжа

Визначимо необхідні умови екстремуму функціонала (6.1) з крайовими умовами (6.2) та рівняннями зв'язків (6.3). Для цього побудуємо допоміжний функціонал

$$\tilde{J}(y_1, y_2, ..., y_n) = \int_a^b \left(F + \sum_{j=1}^k \lambda_j(x) g_j \right) dx. \tag{6.7}$$

Вираз $F + \sum_{j=1}^k \lambda_j(x) g_j$ називають **функцією Лагранжа** даної варіаційної задачі, коефіцієнти $\lambda_j(x)$ – **множниками Лагранжа**.

Розглянемо найпростіший випадок варіаційної задачі Лагранжа, коли функціонал (6.1) залежить від двох невідомих функцій $y(x)$ та $z(x)$, задане одне фазове обмеження, тобто розв'язуємо варіаційну задачу про екстремум функціонала

$$J(y, z) = \int_a^b F(x, y, z, y', z')dx \tag{6.8}$$

за наявності крайових умов

$$y(a) = y_a, y(b) = y_b, z(a) = z_a, z(b) = z_b \tag{6.9}$$

та обмеження

$$g(x, y, z) = 0. \tag{6.10}$$

Окремим випадком цієї задачі є задача про геодезичні лінії, де потрібно на заданій поверхні знайти криву найменшої довжини, що з'єднує задані точки. У задачі (6.8) – (6.10) допустимими функціями є пари функцій $y(x)$ та $z(x)$, які задають просторову криву, розташовану на поверхні $g(x, y, z) = 0$. Для такої задачі можна сформулювати наступну теорему.

<u>Теорема 6.1</u> Якщо пара функцій $y^*(x), z^*(x)$, що належать простору $C^1_{[a;b]}$, є

розв'язком варіаційної задачі (6.8) – (6.10), причому $\left(\frac{\partial g}{\partial y}\right)^2+\left(\frac{\partial g}{\partial z}\right)^2 \neq 0$, то існує така функція $\lambda(x)$, що пара функцій $y^*(x), z^*(x)$ є екстремаллю функціонала

$$\tilde{J}(y,z)=\int_a^b \left[F(x,y,z,y',z')+\lambda(x)\cdot g(x,y,z)\right]dx. \qquad (6.11)$$

Доведення. Розглянемо випадок, коли рівняння зв'язку (6.10) є розв'язним відносно однієї з змінних, наприклад, z, тобто з нього можна отримати $z=\varphi(x,y)$. Це рівняння дозволяє звести дану задачу до варіаційної задачі з однією невідомою функцією $y^*(x)$, що є розв'язком задачі про екстремум функціонала

$$\int_a^b F\left(x,y,\varphi,\frac{\partial\varphi}{\partial x}+\frac{\partial\varphi}{\partial y}y'\right)dx,$$

де $y(a)=y_a,\ y(b)=y_b$. Тому функція $y^*(x)$задовольняє рівнянню Ейлера для функціонала цієї задачі. Дане рівняння після нескладних перетворень можна записати у вигляді:

$$\frac{\partial F}{\partial y}-\left(\frac{d}{dx}\left(\frac{\partial F}{\partial z'}\right)-\frac{\partial F}{\partial z}\right)\frac{\partial\varphi}{\partial y}-\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)=0.$$

З цього рівняння, враховуючи, що $\frac{\partial\varphi}{\partial y}=-\frac{g'_y}{g'_z}$, отримуємо:

$$\frac{\partial F}{\partial y}+\left(\frac{d}{dx}\left(\frac{\partial F}{\partial z'}\right)-\frac{\partial F}{\partial z}\right)\frac{g'_y}{g'_z}-\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)=0. \qquad (6.12)$$

У рівнянні (6.12) зробимо заміну:

$$\lambda(x)=\frac{1}{g'_z}\left(\frac{d}{dx}\left(\frac{\partial F}{\partial z'}\right)-\frac{\partial F}{\partial z}\right). \qquad (6.13)$$

З врахуванням (6.13) рівняння (6.12) приймає вигляд:

$$\frac{\partial F}{\partial y}+\lambda(x)\frac{\partial g}{\partial y}-\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right)=0, \qquad (6.14)$$

а рівність (6.13) можна записати у вигляді:

$$\frac{\partial F}{\partial z}+\lambda\frac{\partial g}{\partial z}-\frac{d}{dx}\left(\frac{\partial F}{\partial z'}\right)=0. \qquad (6.15)$$

Рівняння (6.14) та (6.15) являють собою рівняння Ейлера для допоміжного функціонала (6.11). Таким чином, якщо функції $y^*(x), z^*(x)$ є розв'язком варіаційної задачі (6.8) – (6.10), то вони задовольняють системі рівнянь Ейлера – Лагранжа для допоміжного функціонала (6.11).

Дану теорему можна узагальнити для варіаційної задачі (6.1) – (6.3) для функціонала, що залежить від n функцій $y_1(x), y_2(x),...,y_n(x)$ за наявності диференціальних зв'язків.

Теорема 6.2 Якщо вектор-функція $y^* = (y_1^*(x), y_2^*(x), ..., y_n^*(x))$ з простору $C^1_{[a;b]}$ є розв'язком варіаційної задачі (6.1) – (6.3), то існують такі функції $\lambda_1(x), \lambda_2(x), ..., \lambda_n(x)$, що y^* є екстремаллю допоміжного функціонала (6.7).

Алгоритм застосування необхідних умов екстремуму функціонала у варіаційній задачі (6.1) – (6.3) на умовний екстремум складається з наступних кроків.

1. Для заданої варіаційної задачі записуємо функцію Лагранжа

$$L = F + \sum_{j=1}^{k} \lambda_j(x) g_j. \tag{6.16}$$

2. Складаємо систему з рівнянь Ейлера – Лагранжа для допоміжного функціонала (6.7) та умов зв'язку (6.3):

$$\begin{cases} \dfrac{\partial L}{\partial y_i} - \dfrac{d}{dx}\left(\dfrac{\partial L}{\partial y_i'}\right) = 0; i = 1, 2, ..., n; \\ g_j(x, y_1, y_2, ..., y_n, y_1', y_2', ..., y_n') = 0, j = 1, 2, ..., k. \end{cases} \tag{6.17}$$

3. Знаходимо загальний розв'язок системи (6.17) $y_i(x, C_1, C_2, ..., C_{2n})$, $i = 1,\ 2,\ ...,\ n$, $\lambda_j(x)$, $j = 1, 2, ..., k$.

4. Визначаємо значення сталих $C_1, C_2, ..., C_{2n}$ з крайових умов (6.2) та записуємо вираз для отриманої екстремалі $y(x) = y_1(x), y_2(x), ..., y_n(x))$.

Приклад. Знайти екстремаль функціонала $J(y_1, y_2) = \int_0^1 \left(y_1'^2 + y_2'^2\right) dx$, що задовольняє крайовим умовам $y_1(0) = 2,\ y_2(0) = 0,\ y_1(1) = 2\text{ch}1,\ y_2(1) = 2\text{sh}1$ та диференціальному зв'язку $y_1' - y_2 = 0$.

Розв'язання. Функція Лагранжа даної варіаційної задачі має вигляд (6.16):

$$L = y_1'^2 + y_2'^2 + \lambda(y_1' - y_2).$$

Запишемо систему (6.17).

$$\frac{\partial L}{\partial y_1} = 0; \frac{\partial L}{\partial y_1'} = 2y_1' + \lambda; \frac{d}{dx}\left(\frac{\partial L}{\partial y_1'}\right) = 2y_1'' + \lambda'; \quad \frac{\partial L}{\partial y_2} = -\lambda; \frac{\partial L}{\partial y_2'} = 2y_2'; \frac{d}{dx}\left(\frac{\partial L}{\partial y_1'}\right) = 2y_2''.$$

Підставляючи ці вирази, а також рівняння зв'язку $g\left(y_1, y_2\right) = y_1' - y_2$ у (6.17), отримуємо систему

$$\begin{cases} 2y_1'' + \lambda' = 0; \\ -\lambda - 2y_2'' = 0; \\ y_1' - y_2 = 0. \end{cases}$$

Розв'язавши цю систему, після виключення λ отримуємо:

$$\left[\begin{array}{l} y_1(x) = C_1 x + C_2 + C_3 e^{-x} + C_4 e^x; \\ y_2(x) = C_1 - C_3 e^{-x} + C_4 e^x. \end{array} \right.$$

Значення сталих інтегрування визначаємо з крайових умов:

$$\begin{cases} y_1(0) = C_2 + C_3 + C_4 = 2; \\ y_2(0) = C_1 - C_3 + C_4 = 0; \\ y_1(1) = C_1 + C_2 + C_3 e^{-1} + C_4 e = e + e^{-1}; \\ y_2(1) = C_1 - C_3 e^{-1} + C_4 e = e - e^{-1}. \end{cases} \Rightarrow \begin{cases} C_1 = 0; \\ C_2 = 0; \\ C_3 = 1; \\ C_4 = 1. \end{cases}$$

Таким чином, екстремаль даної варіаційної задачі на умовний екстремум визначається співвідношеннями:

$$\begin{cases} y_1 = e^{-x} + e^{x}; \\ y_2 = -e^{-x} + e^{x}. \end{cases}$$

Варіаційну задачу Лагранжа отримуємо при знаходженні ліній найменшої довжини (геодезичних ліній), що належать заданій поверхні $g(x, y, z) = 0$ та проходять через задані точки (x_1, y_1, z_1) та (x_2, y_2, z_2). Якщо розв'язок даної задачі шукати у вигляді пари функцій $y = y(x), z = z(x)$, то отримаємо задачу дослідження на мінімум функціонала $J(y, z)2 = \int_{x_0}^{x_1} \sqrt{1 + y'^2 + z'^2}\,dx$ за умови $g(x, y, z) = 0$ та крайовими умовами $y(x_1) = y_1$, $z(x_1) = z_1$, $y(x_2) = y_2$, $z(x_2) = z_2$, розв'язок якої знаходиться згідно викладеного вище алгоритму розв'язування варіаційної задачі Лагранжа.

6.3 Ізопериметрична задача

Розглянемо добуток двох величин x та y сума яких дорівнює сталій величині l, тобто $x + y = l$. Відомо, що цей добуток має максимальне значення, якщо співмножники є рівними між собою: $x = y$. Інакше можна сказати, що серед усіх прямокутників, що мають сталий периметр, найбільшу площу має квадрат. Якщо перейти від випадку многокутника зі сталим периметром до більш загальної задачі про фігуру, обмежену замкненою лінією, отримаємо відому ізопериметричну задачу: серед всіх замкнених кривих довжини l знайти таку, що обмежує максимальну площу. Очевидно, що цю площу можна виразити за допомогою формули $S = \int_a^b y(x)dx$ з додатковою умовою

$$\int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2} = l.$$

Дану задачу можна узагальнити наступним чином. Нехай задані тричі диференційовані за змінними x, y та y' функції $F = F(x, y, y')$ та $G = G(x, y, y')$, де $x \in [a; b]$. Потрібно серед всіх допустимих функцій $y = y(x)$ знайти ту, на якій інтеграл $\int_a^b G(x, y, y')dx$ приймає задане значення K, тобто

$$\int_a^b G(x,y,y')dx = K, \tag{6.18}$$

і при цьому функціонал

$$J(y(x)) = \int_a^b F(x,y,y')dx \tag{6.19}$$

досягає екстремального значення.

Для розв'язання цієї задачі застосовують метод Лагранжа, який полягає у наступному. На основі виразів (6.18) та (6.19) будуємо допоміжний функціонал

$$H(y(x)) = \int_a^b F(x,y,y')dx + \lambda\int_a^b G(x,y,y')dx = \int_a^b \left[F(x,y,y') + \lambda G(x,y,y')\right]dx.$$

Введемо позначення для підінтегрального виразу:

$$L(x,y,y',\lambda) = F(x,y,y') + \lambda G(x,y,y'), \tag{6.20}$$

де $x \in [a;b]$. Функція (6.20) являє собою функцію Лагранжа ізопериметричної задачі. Множник Лагранжа λ визначається у процесі розв'язання задачі.

Оскільки необхідною умовою екстремуму функціонала $H(y(x))$ є рівність нулю його варіації, то задача про умовний екстремум функціонала (6.19) зводиться до задачі про безумовний екстремум функціонала $H(y(x))$.

<u>Теорема 6.3</u> Якщо функція $y = y(x)$ реалізує екстремум функціонала (6.19) за умови (6.18), то існує такий сталий параметр λ, що функція $y(x)$ є екстремаллю для функціонала

$$H(y(x)) = \int_a^b L(x,y,y',\lambda)dx, \tag{6.21}$$

де $L(x,y,y',\lambda)$ – функція Лагранжа (6.20).

Доведення. Необхідна умова екстремуму функціонала $H(y(x))$ має вигляд: $\delta H[y(x)] = 0$. Тому, якщо функція $y = y(x)$ реалізує екстремум функціонала $H[y(x)]$, то необхідно, щоб вона задовольняла рівняння Ейлера:

$$\frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) = 0.$$

Підставляючи сюди вираз (6.20), отримаємо:

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) + \lambda\left[\frac{\partial G}{\partial y} - \frac{d}{dx}\left(\frac{\partial G}{\partial y'}\right)\right] = 0. \tag{6.22}$$

<u>Наслідок</u>. Якщо функції F та G не залежать у явному вигляді від аргументу x, тобто $F = F(y,y')$, $G = G(y,y')$, то перший інтеграл для рівняння (6.22) має вигляд

$$F + \lambda G - y'\frac{\partial F}{\partial y'} - \lambda\frac{\partial G}{\partial y'}y' = C.$$

Дану теорему можна узагальнити для варіаційної задачі (6.1), (6.2), (6.4) на умовний екстремум з s інтегральними зв'язками для функціонала, що залежить від n функцій $y_1(x), y_2(x),..., y_n(x)$ у вигляді наступної теореми.

Теорема 6.4 Якщо функції $y_1(x), y_2(x),..., y_n(x)$ надають екстремум функціоналу $J(\bar{y})$ у варіаційній задачі (6.1), (6.2), (6.4), то вони є екстремалями функціонала

$$H(y_1, y_2,..., y_n) = \int_a^b \left(F + \sum_{j=1}^k \lambda_j h_j \right) dx. \qquad (6.23)$$

Приклад. Визначити екстремалі ізопериметричної задачі [1]:

$$J\left[y(x), z(x)\right] = \int_0^1 \left(y'^2 + z'^2 - 4xz' - 4z\right) dx, \;\; y(0)=0, \;\; z(0)=0, \;\; y(1)=1, \;\; z(1)=1,$$

$$\int_0^1 \left(y'^2 - xy' - z'^2\right) dx = 2. \qquad (6.24)$$

Розв'язання. Складемо допоміжний функціонал

$$\Phi = \int_0^1 \left[y'^2 + z'^2 - 4xz' - 4z + \lambda\left(y'^2 - xy' - z'^2\right)\right] dx$$

та запишемо для нього систему рівнянь Ейлера

$$-\frac{d}{dx}\left(2y' + 2\lambda y' - \lambda x\right) = 0,$$

$$-4 - \frac{d}{dx}\left(2z' - 4x - 2\lambda z'\right) = 0,$$

розв'язуючи яку, знаходимо

$$y(x) = \frac{\lambda x^2 + 2C_1 x}{4(1+\lambda)} + C_2,$$

$$z(x) = \frac{C_3 x}{2(1-\lambda)} + C_4.$$

З крайових умов отримуємо:

$$C_1 = \frac{3\lambda + 4}{2}, \;\; C_2 = 0, \;\; C_3 = 2(1-\lambda), \;\; C_4 = 0,$$

так що

$$y(x) = \frac{\lambda x^2 + (3\lambda + 4)x}{4(1+\lambda)},$$

$$z(x) = x.$$

Для знаходження λ використовуємо ізопериметричну умову (6.24). Оскільки $y'(x) = \dfrac{2\lambda x + 3\lambda + 4}{4(1+\lambda)}$, а $z'(x) = 1$, то знаходимо

$$\int_0^1 \left[\frac{(2\lambda x+3\lambda+4)^2}{16(1+\lambda)^2} - \frac{(2\lambda x+3\lambda+4)x}{4(1+\lambda)} - 1 \right] dx = 2,$$

звідки маємо рівняння для визначення λ:

$$\frac{1}{3}\left(23\lambda^2+46\lambda+24\right)=48\left(\lambda^2+2\lambda+1\right).$$

Звідси $\lambda_1=-10/11$ і $\lambda_2=-12/11$. Підстановкою у (6.24) переконуємось, що $\lambda_2=-12/11$ ізопериметричну умову не задовольняє, а $\lambda_1=-10/11$ задовольняє.

Шукана екстремаль визначається рівняннями

$$\begin{cases} y(x)=\dfrac{7x-5x^2}{2}, \\ z(x)=x. \end{cases}$$

Питання для самоперевірки

1. Навести приклади варіаційних задач на умовний екстремум.
2. Сформулювати варіаційну задачу Лагранжа.
3. Навести формулювання ізопериметричної задачі.
4. Яким чином ізопериметричну задачу можна звести до задачі Лагранжа?
5. Навести класифікацію типів зв’язків у варіаційних задачах на умовний екстремум.
6. Надати означення функції Лагранжа для варіаційних задач з диференціальними та інтегральними типами зв’язків.
7. Навести алгоритм розв’язання варіаційної задачі Лагранжа.
8. Алгоритм розв’язання ізопериметричної задачі.

Практичні завдання

1. Знайти екстремаль задачі Лагранжа: $J[y,z]=\int_0^1 \left(y'^2+z\right) dx \to extr$, $y(0)=1$, $z(0)=1$, $y(1)=e$, $z(1)=e^2+1$, якщо $z-y^2-x=0$.

2. Знайти екстремаль задачі Лагранжа: $J[y,z]=\int_0^1 \left(y'^2+z^2\right) dx \to extr$, $y(0)=0$, $z(0)=1$, $y(1)=1$, $z(1)=\sqrt{2}$, якщо $y-z^2+1=0$.

3. Знайти екстремаль задачі Лагранжа: $J[y,z]=\int_0^{\pi/2} \left(y-z'^2\right) dx \to extr$, $y(0)=3$, $z(0)=1$, $y(\pi/2)=2$, $z(\pi/2)=0$, якщо $y-z^2-2=0$.

4. Знайти екстремаль задачі Лагранжа: $J[y,z]=\int\limits_0^1\left(y'^2+z^2+x\right)dx\to extr$,

$y(0)=0,\ z(0)=1,\ y(1)=e,\quad z(1)=\sqrt{e^2+1}$, якщо $y^2-z^2+1=0$.

5. Відшукати найкоротшу відстань між точками $A(0,0,3)$ і $B(1,1,-1)$ на поверхні $2x+2y+z-3=0$.
6. Відшукати найкоротшу відстань між точками $A(1,0,-1)$ і $B(0,-1,1)$ на поверхні $x+y+z=0$.
7. Відшукати найкоротшу відстань між точками $A(-5,1,1)$ і $B(1,1,-1)$ на поверхні $x+2y+3z=0$.
8. Знайти найкоротшу відстань між двома точками $A(x_0,y_0,z_0)$ і $B(x_1,y_1,z_1)$ на поверхні $\phi(x,y,z)=0$ (у загальному вигляді).
9. Знайти екстремаль ізопериметричної задачі [1,2]:

9.1 $J[y]=\int\limits_0^1 y'^2dx\to extr$, $y(0)=1,\ y(1)=6,\ \int\limits_0^1 ydx=3$.

9.2 $J[y]=\int\limits_0^1\left(y'^2+x^2\right)dx\to extr$, $y(0)=0,\ y(1)=0,\ \int\limits_0^1 y^2dx=2$.

9.3 $J[y]=\int\limits_0^1 y'^2dx\to extr$, $y(0)=0,\ y(1)=\dfrac{1}{4},\ \int\limits_0^1\left(y-y'^2\right)dx=\dfrac{1}{12}$.

9.4 $J[y]=\int\limits_0^1 y'^2dx\to extr;\ \int\limits_0^1 xy\ dx=0,\ y(0)=-4,\ y(1)=4.$

9.5 $J[y]=\int\limits_1^2 x^3y'^2dx\to extr;\ \int\limits_1^2 y\ dx=2,\ y(1)=4,\ y(2)=1.$

9.6 $J[y]=\int\limits_0^1 x\left(y_1-y_2\right)dx\to extr;\ \int\limits_0^1 y_1'y_2'dx=-4/5,\ y_1(0)=y_2(0)=y_2(1)=0,$ $y_1(1)=2.$

9.7 $J[y]=\int\limits_0^1\left(y_1+y_2\right)dx\to extr;\ \int\limits_0^1 y_1'y_2'dx=0,\ y_1(0)=y_2(0)=0,\ y_1(1)=1,$ $y_2(1)=-3.$

9.8 $J[y]=\int\limits_0^1 y'^2dx\to extr;\ \int\limits_0^1 y\ dx=1,\ \int\limits_0^1 xy\ dx=0,\ y(0)=y(1)=0.$

9.9 $J[y]=\int\limits_0^1 y_1'y_2'dx\to extr;\ \int\limits_0^1 xy_1dx=\int\limits_0^1 xy_2dx=0,\ y_1(0)=y_1(1)=y_2(0)=0,$ $y_2(1)=1.$

9.10 $J[y]=\int\limits_0^1 y'^2dx\to extr;\ \int\limits_0^1 y\ dx=-3/2,\ \int\limits_0^1 xy\ dx=-2,\ y(0)=2,\ y(1)=-14.$

7 ПРЯМІ МЕТОДИ ВАРІАЦІЙНОГО ЧИСЛЕННЯ

План:

7.1 Сутність прямих методів дослідження функціоналів на екстремум.
7.2 Метод Рітца.
7.3 Скінченно-різницевий метод Ейлера.
7.4 Метод Канторовича.

Ключові поняття та терміни: наближений розв'язок, координатні функції, невизначені коефіцієнти, точність наближеного розв'язку.

7.1 Сутність прямих методів дослідження функціоналів на екстремум

Диференціальні рівняння, отримані у процесі розв'язування варіаційних задач, інтегруються у скінченому вигляді лише у окремих випадках. У зв'язку з цим виникає необхідність у побудові наближених методів розв'язування варіаційних задач, якими є прямі методи, що не зводять розв'язання варіаційної задачі до розв'язання диференціальних рівнянь.

Розглянемо задачу про знаходження мінімуму деякого функціонала $J(y)$, визначеного на певній множині M допустимих кривих. Для того, щоб ця задача мала сенс, слід допустити, що у класі M існують криві, для яких функціонал $J(y) < +\infty$ і при цьому $\inf J(y) = \alpha > -\infty$. У цьому випадку існує така послідовність кривих y_1, $y_2, \ldots$, $y_n, \ldots$, що $\lim\limits_{n\to\infty} J(y_n) = \alpha$. Таку послідовність називають **мінімізуючою послідовністю**.

Якщо для мінімізуючої послідовності $\{y_n\}$ існує гранична крива y_0, тобто $\lim\limits_{n\to\infty} y_n = y_0$, і при цьому $\lim\limits_{n\to\infty} J(y_n) = J(y_0)$, то $J(y_0) = \alpha$ і крива y_0 є розв'язком даної задачі. Отже, отримання наближеного розв'язку варіаційної задачі прямим методом передбачає побудову мінімізуючої послідовності $\{y_n\}$, доведення існування у цієї послідовності граничної кривої y_0 та доведення можливості граничного переходу $\lim\limits_{n\to\infty} J(y_n) = J(y_0)$. Члени мінімізуючої послідовності $\{y_n\}$ можна розглядати як наближені розв'язки варіаційної задачі. Прямі методи розрізняються між собою способами побудови мінімізуючої послідовності.

Основна ідея **прямих методів** полягає у тому, що варіаційна задача розглядається як гранична для деякої задачі про екстремум функції скінченної кількості змінних. Допустима функція $y(x)$ деякої варіаційної задачі може бути представленою у вигляді степеневого ряду $y(x) = \sum\limits_{n=0}^{\infty} a_n x^n$, ряду Фур'є

$y(x)=\frac{a_0}{2}+\sum_{n=1}^{\infty}\left(a_n\cos nx+b_n\sin nx\right)$, або, у загальному випадку, функціонального ряду виду

$$y(x)=\sum_{n=0}^{\infty}a_n\varphi_n(x), \tag{7.1}$$

де $\{\varphi_n(x)\}$ – деяка наперед вибрана система функцій.

Для визначення функції $y(x)$, заданої за допомогою такого ряду, достатньо визначити всі його коефіцієнти. Значення функціонала $J(y)$ у даному випадку визначається коефіцієнтами a_n даного функціонального ряду. Таким чином, функціонал можна розглядати як функцію нескінченної кількості змінних a_n, $n=1,2,...$.

Задача дослідження функціонала на екстремум являє собою задачу дослідження на екстремум функції нескінченної кількості змінних. При застосуванні прямих методів дана задача наближено замінюється задачею про екстремум функції скінченної кількості змінних шляхом заміни у функціоналі $J(y)$ ряду (7.1) скінченною комбінацією перших N функцій $\varphi_k(x)$, $k=1,2,...,N$. Прикладами прямих методів варіаційного числення є метод Рітца та скінченно-різницевий метод Ейлера.

7.2 Метод Рітца

Основна ідея **методу Рітца** полягає у тому, що значення функціонала

$$J(y)=\int_a^b F(x,y,y')dx \tag{7.2}$$

розглядається не на довільних допустимих кривих, а тільки на всіх можливих лінійних комбінаціях виду

$$y_n=\sum_{i=1}^{n}a_i\varphi_i(x), \tag{7.3}$$

де $a_i, i=1,2,...,n$ – невідомі сталі коефіцієнти, $\varphi_1,\varphi_2,...,\varphi_n$ – перші n функцій з деякої наперед вибраної функціональної послідовності. Така послідовність повинна бути повною, тобто будь-яка допустима функція може бути апроксимованою при відповідному виборі коефіцієнтів a_i, $i=1,2,...,n$, лінійною комбінацією виду (7.3), тобто повинна виконуватися рівність:

$$\lim_{n\to\infty}\left\| y(x)-\sum_{i=1}^{n}a_i\varphi_i(x)\right\|=0.$$

У цій рівності розглядається норма функціонального простору $C^1_{[a;b]}$ неперервно диференційовних на відрізку $[a;b]$ функцій.

Для того, щоб функція (7.3), що є наближенням точного розв'язку варіаційної задачі, була допустимою, необхідно, щоб вона задовольняла заданим крайовим умовам. Якщо крайові умови варіаційної задачі є однорідними, то функції $\varphi_i(x), i=1,2,...,n,$ вибирають так, щоб вони задовольняли цим крайовим умовам. Наприклад, для варіаційної задачі про екстремум функціонала (7.2) при крайових умовах $y(a)=y(b)=0$ функції $\varphi_i(x)$ вибирають так, щоб $\varphi_i(a)=\varphi_i(b)=0, i=1, 2,..., n.$ Наприклад, можна взяти $\varphi_i(x)=\sin\dfrac{i\pi(x-a)}{b-a}$, $i=1, 2,..., n$, або $\varphi_i(x)=\left(x-a\right)\left(x-b\right)^i$, $i=1, 2,..., n.$

Якщо крайові умови є неоднорідними, наприклад, $y(a)=y_a\neq 0; y(b)=y_b\neq 0$, то наближений розв'язок варіаційної задачі шукають у вигляді:

$$y_n(x)=\varphi_0(x)+\sum_{i=1}^{n}a_i\varphi_i(x)\,, \tag{7.4}$$

де функція $\varphi_0(x)$ задовольняє заданим крайовим умовам, тобто $\varphi_0(a)=y_a, \varphi_0(b)=y_b$, а функції $\varphi_i(x), i=1, 2,..., n,$ задовольняють однорідним крайовим умовам. Наприклад, можна вибрати $\varphi_0(x)=\dfrac{y_b-y_a}{b-a}\left(x-a\right)+y_a$.

На лінійних комбінаціях виду (7.3) або (7.4) функціонал $J(y)$ перетворюється у функцію $f\left(a_1,a_2,...,a_n\right)$ коефіцієнтів $a_1,a_2,...,a_n$. Ці коефіцієнти вибираються так, щоб функція f досягала екстремуму. При цьому вони визначаються за допомогою необхідної умови екстремуму, тобто як розв'язки системи алгебраїчних рівнянь

$$\frac{\partial f}{\partial a_i}=0, i=1, 2,..., n. \tag{7.5}$$

Підставляючи знайдені з системи (7.5) значення коефіцієнтів a_i у апроксимацію (7.3) або (7.4), отримуємо наближений розв'язок варіаційної задачі.

Функції $\varphi_i(x), i=1, 2,..., n,$ називають **базисними (координатними, пробними) функціями**.

7.3 Скінченно-різницевий метод Ейлера

Основна ідея **скінченно-різницевого методу Ейлера** полягає у тому, що допустимі криві, на яких заданий функціонал (7.2), апроксимуються ламаними, складеними з заданої кількості n відрізків. Потім розглядаються значення функціонала на таких апроксимаціях.

Розглянемо найпростішу варіаційну задачу про знаходження екстремуму функціонала (7.2) з крайовими умовами $y(a)=y_a, y(b)=y_b$. Значення функціонала розглядаються на ламаних, складених з n відрізків з заданими абсцисами вершин: $x_0 = a, x_1 = x_0 + \Delta x, x_2 = x_1 + \Delta x, \ldots, x_n = x_0 + n\Delta x = b$, де $\Delta x = \dfrac{b-a}{n}$.

На таких ламаних функціонал (7.2) перетворюється у функцію $g(y_1, y_2, ..., y_{n-1})$ ординат $y_1 = y(x_1), y_2 = y(x_2), ..., y_{n-1} = y(x_{n-1})$ вершин ламаної:

$$g(y_1, y_2, ..., y_{n-1}) = \sum_{k=0}^{n-1} F\left(x_k, y_k, \frac{y_{k+1} - y_k}{\Delta x}\right)\Delta x. \qquad (7.6)$$

Функція $g(y_1, y_2, ..., y_{n-1})$ є наближеним значенням інтегралу $\int\limits_{x_0}^{x_n} F(x, y, y')dx$, обчисленого за формулою прямокутників. Похідна y' при цьому замінюється різницевим співвідношенням $y'(x_k) \approx \dfrac{y_{k+1} - y_k}{\Delta x}$.

Ординати $y_1, y_2, ..., y_{n-1}$ вибираються так, щоб функція $g(y_1, y_2, ..., y_{n-1})$ досягала екстремуму. Вони визначаються з системи рівнянь

$$\frac{\partial g}{\partial y_i} = 0, i = 1, 2, ..., n-1. \qquad (7.7)$$

Приклад. Знайти наближений розв'язок задачі:

$$J(y) = \int\limits_0^1 (y'^2 + y^2 + 2xy)dx \to extr, \quad y(0) = 0, y(1) = 0.$$

Розв'язання. Скінченно-різницевий метод Ейлера.

Перше наближення: $n = 2,\ \Delta x = 0,5.$ Апроксимуємо екстремаль ламаною з координатами вершин $(0;0),(0,5;y_1),(1;0)$, де y_1 – шукана ордината. Маємо

$$\Phi(y_1) = \left(\left(\frac{y_1 - y_0}{\Delta x}\right)^2 + y_0^2 + 2x_0 y_0 + \left(\frac{y_2 - y_1}{\Delta x}\right)^2 + y_1^2 + 2x_1 y_1\right)\Delta x$$

та після підстановки $x_0 = 0;\ x_1 = 0,5;\ x_2 = 1;\ y_0 = 0;\ y_2 = 0$ отримаємо:

$$\Phi(y_1) = \left(\left(\frac{y_1}{0,5}\right)^2 + \left(-\frac{y_1}{0,5}\right)^2 + y_1^2 + y_1\right)\cdot 0,5 = \left(\frac{y_1^2}{0,25} + \frac{y_1^2}{0,25} + y_1^2 + y_1\right)\cdot 0,5 =$$

$$= \frac{9}{2}y_1^2 + \frac{1}{2}y_1 \to extr.$$

Обчислюємо похідну:

$$\frac{d\Phi(y_1)}{dy_1} = 9y_1 + \frac{1}{2} = 0,$$

отже,

$$y_1^* = -\frac{1}{18} \approx -0.05556.$$

Друге наближення: $n=5,\ \Delta x=0,2$. Апроксимуємо екстремаль ламаною з координатами вершин $(0;0),\ (0,2;y_1),\ (0,4;y_2),\ (0,6;y_3),\ (0,8;y_4),\ (1;0)$, де y_1,y_2,y_3,y_4 – невідомі. Маємо:

$$\Phi(y_1,y_2,y_3,y_4)=\left(\left(\frac{y_1-y_0}{\Delta x}\right)^2+y_0^2+2x_0y_0+\left(\frac{y_2-y_1}{\Delta x}\right)^2+y_1^2+2x_1y_1+\right.$$

$$\left.+\left(\frac{y_3-y_2}{\Delta x}\right)^2+y_2^2+2x_2y_2+\left(\frac{y_4-y_3}{\Delta x}\right)^2+y_3^2+2x_3y_3+\left(\frac{y_5-y_4}{\Delta x}\right)^2+y_4^2+2x_4y_4\right)\Delta x=$$

$$=\left(\left(\frac{y_1}{0,2}\right)^2+\left(\frac{y_2-y_1}{0,2}\right)^2+y_1^2+0,4y_1+\left(\frac{y_3-y_2}{0,2}\right)^2+y_2^2+0,8y_2+\left(\frac{y_4-y_3}{0,2}\right)^2+y_3^2+1,2y_3+\right.$$

$$\left.+\left(\frac{y_4}{0,2}\right)^2+y_4^2+1,6y_4\right)\cdot 0,2=$$

$$=\left(\frac{y_1^2}{0,04}+\frac{(y_2-y_1)^2}{0,04}+y_1^2+0,4y_1+\frac{(y_3-y_2)^2}{0,04}+y_2^2+0,8y_2+\frac{(y_4-y_3)^2}{0,04}+y_3^2+\right.$$

$$\left.+1,2y_3+\frac{y_4^2}{0,04}+y_4^2+1,6y_4\right)\cdot 0,2=$$

$$=\left(51\left[y_1^2+y_2^2+y_3^2+y_4^2\right]-50\left[y_1y_2+y_2y_3+y_3y_4\right]+0,4\left[y_1+2y_2+3y_3+4y_4\right]\right)\cdot 0,2=$$

$$=\frac{51}{5}\left[y_1^2+y_2^2+y_3^2+y_4^2\right]-10\left[y_1y_2+y_2y_3+y_3y_4\right]+\frac{2}{25}\left[y_1+2y_2+3y_3+4y_4\right]\to extr.$$

Для дослідження функції $\Phi(y_1,y_2,y_3,y_4)$ на екстремум, прирівнюємо до нуля частинні похідні та отримуємо систему рівнянь:

$$\begin{cases}\dfrac{\partial\Phi}{\partial y_1}=\dfrac{102}{5}y_1-10y_2+\dfrac{2}{25}=0,\\ \dfrac{\partial\Phi}{\partial y_2}=-10y_1+\dfrac{102}{5}y_2-10y_3+\dfrac{4}{25}=0,\\ \dfrac{\partial\Phi}{\partial y_3}=-10y_2+\dfrac{102}{5}y_3-10y_4+\dfrac{6}{25}=0,\\ \dfrac{\partial\Phi}{\partial y_4}=-10y_3+\dfrac{102}{5}y_4+\dfrac{8}{25}=0.\end{cases}$$

Звідси $y_1^*\approx-0,0285944,\ y_2^*\approx-0,0503325,\ y_3^*\approx-0,058084,\ y_4^*\approx-0,0441588.$

Метод Рітца.

Для знаходження функції, яка буде наближеним розв’язком варіаційної задачі, можна обрати многочлени

$$y_n(x) = x(1-x)(\alpha_1 + \alpha_2 x + ... + \alpha_n x^{n-1}).$$

У такому разі $\varphi_0(x) \equiv 0$ і координати функції набувають вигляду

$$\varphi_1(x) = x - x^2, ..., \varphi_n(x) \equiv x^n - x^{n-1}.$$

Перше наближення: $y_1(x) = \alpha_1 x(1-x).$ Після підстановки $y_1(x)$ у вираз для функціонала та інтегрування отримуємо задачу:

$$\Phi\left(\alpha_1\right) = \int_0^1 \left(\alpha_1^2(1-2x)^2 + \alpha_1^2 x^2(1-x)^2 + 2\alpha_1 x^2(1-x)\right)dx =$$

$$= \left[\alpha_1^2\left(x - 2x^2 + \frac{4}{3}x^3 + \frac{1}{3}x^3 - \frac{1}{2}x^4 + \frac{1}{5}x^5\right) + \alpha_1\left(\frac{2}{3}x^3 - \frac{1}{2}x^4\right)\right]\Bigg|_0^1 =$$

$$= \alpha_1^2\left(1 - 2 + \frac{4}{3} + \frac{1}{3} - \frac{1}{2} + \frac{1}{5}\right) + \alpha_1\left(\frac{2}{3} - \frac{1}{2}\right) = \frac{11}{30}\alpha_1^2 + \frac{1}{6}\alpha_1 \to extr.$$

Необхідною умовою екстремуму є рівність

$$\frac{d\Phi}{d\alpha_1} = \frac{11}{15}\alpha_1 + \frac{1}{6} = 0\,.$$

Звідси знаходимо

$$\alpha_1^* = -\frac{15}{66} \approx -0{,}227\,.$$

Отже, наближений вираз для екстремалі має вигляд:

$$y_1^*(x) = -\frac{15}{66}x(1-x) \approx -0{,}227x(1-x).$$

Друге наближення: $y_2(x) = x(1-x)(\alpha_1 + \alpha_2 x).$ Після підстановки $y_2(x)$ у вираз для функціонала обчислюємо інтеграл

$$\Phi(\alpha_1, \alpha_2) =$$

$$= \int_0^1 \left[\left(-3\alpha_2 x^2 - 2\alpha_1 x + 2\alpha_2 x + \alpha_1\right)^2 + x^2(1-x)^2(\alpha_1 + \alpha_2 x)^2 + 2x^2(1-x)(\alpha_1 + \alpha_2 x)\right]dx =$$

$$= \frac{1}{7}\alpha_2^2 + \frac{11}{30}\alpha_2\alpha_1 + \frac{11}{30}\alpha_1^2 + \frac{1}{10}\alpha_2 + \frac{1}{6}\alpha_1 \to extr.$$

Для визначення можливого екстремуму функції $\Phi(\alpha_1, \alpha_2)$ досліджуємо систему рівнянь:

$$\begin{cases} \dfrac{\partial\Phi}{\partial\alpha_1} = \dfrac{11}{30}\alpha_2 + \dfrac{11}{15}\alpha_1 + \dfrac{1}{6} = 0, \\ \dfrac{\partial\Phi}{\partial\alpha_2} = \dfrac{2}{7}\alpha_2 + \dfrac{11}{30}\alpha_1 + \dfrac{1}{10} = 0. \end{cases}$$

Звідси отримуємо:

$$\begin{cases} \alpha_1^* = -\dfrac{69}{473} \approx -0{,}1459, \\ \alpha_2^* = -\dfrac{7}{43} = -0{,}1628. \end{cases}$$

Отже, наближений вираз для екстремалі набуває вигляду:

$$y_2^*(x) \approx x(1-x)(-0{,}1459 - 0{,}1628x).$$

Точний розв'язок. Рівняння Ейлера має вигляд: $y'' - y = x$.
Екстремаль, що задовольняє граничні умови:

$$y(x) = \frac{sh(x)}{sh(1)} - x.$$

Результати, отримані різними методами, представлені у таблиці 7.1.

Таблиця 7.1.

Метод розв'язку	x_i						
	0	0.2	0.4	0.5	0.6	0.8	1
Перше наближення методу Ейлера	0			–0.05556			0
Друге наближення методу Ейлера	0	–0.02859	–0.05033		–0.058084	–0.0441588	0
Перше наближення методу Рітца	0	–0.03632	–0.05448	–0.05675	–0.05448	–0.03632	0
Друге наближення методу Рітца	0	–0.02855	–0.05064	–0.05681	–0.05845	–0.04418	0
Точний розв'язок	0	–0.02868	–0.0509	–0.0566	–0.05826	–0.0443	0

Порівняння значень ординат для розв'язків, отриманих з використанням других наближень методів Ейлера і Рітца, з точним розв'язком показує, що вони збігаються з точністю до $\varepsilon = 0{,}001$.

7.4 Метод Канторовича

При застосуванні методу Рітца до функціоналів $J(u)$, що залежать від функцій кількох змінних $u(x_1, x_2, ..., x_n)$, спочатку вибирають систему базисних (координатних) функцій $U_1(x_1, x_2, ..., x_n)$, $U_2(x_1, x_2, ..., x_n), \ldots$, $U_m(x_1, x_2, ..., x_n)$, …, потім шукають наближений розв'язок варіаційної задачі у вигляді лінійної комбінації $u_m = \sum_{k=1}^{m} \alpha_k U_k(x_1, x_2, ..., x_n)$, де коефіцієнти α_k є сталими. При застосуванні методу Канторовича також вибирають аналогічну систему базисних функцій $U_1(x_1, x_2, ..., x_n)$, $U_2(x_1, x_2, ..., x_n), \ldots$, $U_m(x_1, x_2, ..., x_n)$, …, проте наближений розв'язок будують у вигляді

$$u_m = \sum_{k=1}^{m} \alpha_k(x_i) U_k(x_1, x_2, ..., x_n) \qquad (7.8)$$

де коефіцієнти $\alpha_k(x_i)$ є невідомими функціями однієї з незалежних змінних x_i. Функціонал $J(u)$ на множині функцій, що мають вигляд (7.8), залежить від m функцій $\alpha_1(x_i)$, $\alpha_2(x_i)$, …, $\alpha_m(x_i)$ однієї незалежної змінної x_i. Їх вибирають так, щоб функціонал $J(u_m) = J(\alpha_1(x_i), \alpha_2(x_i), ..., \alpha_m(x_i))$ досягав екстремуму.

Якщо можливо здійснити граничний перехід $u = \lim_{m\to\infty} u_m$, то за умови повноти вибраної системи базисних функцій та наявності екстремуму у даного функціонала можна отримати точний розв'язок варіаційної задачі. Без здійснення граничного переходу отримуємо наближений розв'язок варіаційної задачі у вигляді (7.8). Цей розв'язок у загальному випадку є точнішим, ніж розв'язок, отриманий у результаті застосування методу Рітца з тією ж кількістю m базисних функцій. Це пояснюється тим, що клас функцій виду (7.8) є ширшим, ніж аналогічна множина функцій з сталими коефіцієнтами α_k, $k = 1, ..., m$. Тому серед функцій виду (7.8) можна підібрати кращі апроксимації розв'язку варіаційної задачі порівняно з множиною лінійних комбінацій базисних функцій з сталими коефіцієнтами.

Приклад. Знайти функцію $u(x, y)$, що надає екстремум функціоналу $J[u] = \iint_D \left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 - 2u \right] dxdy$ та дорівнює нулю на межі області D, обмеженої прямими $y = \pm \frac{x}{\sqrt{3}}$ та $x = 1$.

Розв'язання. При побудові наближеного розв'язку (7.8) даної варіаційної задачі обмежимось однією базисною функцією. Перше наближення розв'язку будемо шукати у вигляді $u_1(x, y) = \alpha(x) U_1(x, y)$, де $U_1(x, y) = y^2 - \left(\frac{x}{\sqrt{3}}\right)^2$. При такому виборі базисної функції $u_1(x, y) = 0$ на прямих $y = \pm \frac{x}{\sqrt{3}}$. Крайові умови будуть виконуватися на прямій $x = 1$, якщо $\alpha(1) = 0$. Підставивши вираз для $u_1(x, y)$ замість $u(x, y)$ у функціонал $J(u)$ після інтегрування за змінною y отримуємо функціонал, що залежить від функції однієї змінної $\alpha(x)$:

$$J(\alpha(x)) = \frac{8\sqrt{3}}{405} \int_0^1 \left(2x^5 \alpha'^2 + 10x^4 \alpha \cdot \alpha' + 30x^3 \alpha^2 + 15x^2 \alpha\right) dx.$$

Рівняння Ейлера для цього функціонала має вигляд:

$$x^2 \alpha'' + 5x\alpha' - 5\alpha = \frac{15}{4}.$$

Загальний розв’язок цього рівняння має вигляд: $\alpha(x) = C_1 x + C_2 x^{-5} - \dfrac{3}{4}$. Оскільки при $x = 0$ розв’язок задачі має бути обмеженим, то $C_2 = 0$. Сталу C_1 знаходимо з умови $\alpha(1) = 0$, звідки $C_1 = \dfrac{3}{4}$. Таким чином, отримуємо наближений розв’язок варіаційної задачі у вигляді:

$$u_1(x, y) = \frac{3}{4}(x-1)\left(y^2 - \left(\frac{x}{\sqrt{3}}\right)^2\right).$$

Питання для самоперевірки

1. Які методи називають прямими методами варіаційного числення?
2. Яку послідовність називають мінімізуючою?
3. З яких етапів складається отримання наближеного розв’язку варіаційної задачі прямим методом?
4. У чому полягає сутність прямих методів варіаційного числення?
5. Сформулювати основну ідею методу Рітца.
6. Які функції називають базисними?
7. Які вимоги ставляться до базисних функцій у методі Рітца?
8. У чому полягає основна ідея скінченно-різницевого методу Ейлера?
9. Для яких функціоналів використання методу Ейлера є найбільш ефективним?
10. У чому полягає основна відмінність методу Канторовича від методу Рітца?
11. Для розв’язання яких варіаційних задач застосовують метод Канторовича?
12. Чому застосування методу Канторовича у загальному випадку дозволяє отримати більш точну апроксимацію розв’язку варіаційної задачі у порівняння з методом Рітца?

Практичні завдання

1. Методом Ейлера знайти наближений розв’язок задачі ($n-$число ланок ламаної).

1.1. $J[y] = \int_0^1 (y'^2 + 2y)dx \to extr$, $y(0) = 0, y(1) = 0;\ n = 2,\ n = 4$.

1.2. $J[y] = \int_1^2 (xy'^2 - \dfrac{x^2-1}{x} y^2 - 2x^2 y)dx \to extr$, $y(1) = 0, y(2) = 0;\ n = 2,\ n = 3$.

1.3. $J[y] = \int_0^2 (y^2 + 2xy + y'^2)dx \to extr$, $y(0) = 0, y(2) = 0;\ n = 2,\ n = 3$.

1.4. $J[y] = \int_0^1 (y'^2 + 1)dx \to extr$,

a) $y(0)=0, y(1)=0$;

b) $y(0)=0, y(1)=1$; $n=2,\ n=4$

2. Методом Рітца знайти наближений розв'язок задачі (вказано вид функцій $y_1(x)$ та $y_2(x)$).

2.1. $J[y]=\int\limits_0^1 (y'^2+2y)dx \to extr$,

$$y(0)=0, y(1)=0;\quad y_1=\alpha_1 x(1-x),\ y_2=\alpha_1 x(1-x)+\alpha_2 x^2(1-x).$$

2.2. $J[y]=\int\limits_0^1 (y'^2+2xy-x^2y^2)dx \to extr$,

$$y(0)=0, y(1)=0;\quad y_1=\alpha_1 x(1-x),\ y_2=\alpha_1 x(1-x)+\alpha_2 x^2(1-x).$$

2.3. $J[y]=\int\limits_{-1}^1 (y'^2-\left(1+x^2\right)y-2y)dx \to extr$,

$$y(-1)=0, y(1)=0;\quad y_1=\alpha_1(1-x^2),\ y_2=\alpha_1(1-x^2)+\alpha_2(1-x^4).$$

2.4. $J[y]=\int\limits_0^1 (y'^2-y^2+2xy)dx \to extr$, $y(0)=0, y(1)=0$; $y_1=\alpha_1(x-x^2)$.

2.5. $J[y]=\int\limits_0^2 \left(y'^2+y^2+2xy\right)dx \to extr$, $y(0)=y(2)=0$;

$$y_1=\alpha_1 x(1-x),\ y_2=\alpha_1 x(1-x)+\alpha_2 x^2(1-x).$$

8 ВСТУП ДО ЧИСЕЛЬНИХ МЕТОДІВ ОПТИМІЗАЦІЇ ФУНКЦІЙ ОДНІЄЇ ТА КІЛЬКОХ ЗМІННИХ

План:

8.1 Постановка задачі одновимірної оптимізації.
8.2 Метод дихотомії.
8.3 Метод золотого перерізу.
8.4 Метод Фібоначчі.
8.5 Загальна характеристика методів багатовимірної оптимізації.
8.6 Градієнтні методи.

Ключові поняття та терміни: унімодальна та мультимодальна функція, інтервал невизначеності, числа Фібоначчі.

8.1 Постановка задачі одновимірної оптимізації

Задачі одновимірної оптимізації – це задачі пошуку екстремуму функції однієї змінної. Такі задачі виникають під час дослідження процесів, що залежать від однієї скалярної величини. Крім того, вони є складовою частиною багатьох ітераційних методів розв’язування більш складних оптимізаційних задач пошуку екстремумів функцій багатьох змінних.

Задачі мінімізації функції однієї змінної методами диференціального числення детально розглядаються у курсі математичного аналізу, проте такі методи мають обмежене застосування і не завжди є зручними для комп’ютерної реалізації.

Далі розглянемо методи оптимізації функції однієї змінної, що не потребують неперервної диференційовності функції та дозволяють ефективну реалізацію на комп’ютері. Відомі методи відшукання екстремуму нелінійної функції залежно від використаної інформації для пошуку наближеного значення екстремуму можна поділити на три групи:

1. Методи прямого пошуку (методи нульового порядку), які використовують тільки значення функції.
2. Методи першого порядку, які додатково використовують і значення першої похідної.
3. Методи другого порядку, які використовують другі похідні.

Розглянемо задачу дослідження на мінімум функції $f(x)$, $x \in [a;b] \subset \mathbb{R}$ (дослідження на максимум функції $f(x)$ еквівалентно дослідженню на мінімум функції $-f(x)$). Оскільки точний локальний мінімум $-f(x)$ на $[a;b]$ невідомий, то назвемо цей проміжок інтервалом невизначеності.

Нехай x^* – точка, у якій $f(x)$ досягає мінімуму на відрізку $[a;b]$. Функцію $f(x)$, $x \in \mathbb{R}$, називають **унімодальною** на відрізку $[a;b]$, якщо вона є неперервною на $[a;b]$ і для всіх точок x_1 та x_2, $a \le x_1 < x_2 \le b$, таких, що для $x_1 > x^*$ $f(x_1) < f(x_2)$, а для $x_2 < x^*$ $f(x_1) > f(x_2)$. Функції, що не є унімодальними на $[a;b]$, називають **мультимодальними** на цьому відрізку. З означення унімодальної функції випливає, що вона має єдиний мінімум на відрізку $[a;b]$, причому до нього вона монотонно спадає, а після нього монотонно зростає.

Розглянемо далі застосування методів прямого пошуку для мінімізації функції однієї змінної на заданому проміжку, що передбачають використання лише значень цієї функції. Розрізняють дві групи методів прямого пошуку. **Методи пасивного пошуку** передбачають, що точки $x_k \in [a;b]$, $k = 1, 2, ..., N$, вибирають одномоментно до обчислення функції у цих точках. Більш ефективними з точки зору можливості досягнення заданої точності знаходження точки мінімуму є **методи послідовного пошуку.** При застосуванні цієї групи методів точки x_k визначаються послідовно, при цьому для вибору наступної точки використовують значення функції, обчислені у попередніх точках. Вибір наступної точки x_k та обчислення значення функції $f(x_k)$ називають кроком послідовного пошуку. Після N обчислень значень функції знаходять інтервал довжиною l_N, у якому гарантовано знаходиться шукана точка мінімуму x^*. Цей інтервал називають **інтервалом невизначеності**. Умовою закінчення обчислень при застосуванні методів прямого пошуку є виконання нерівності $l_N < \varepsilon$, де ε – задана максимальна допустима довжина інтервалу невизначеності.

При застосуванні методів прямого пошуку використовують **процедуру виключення відрізка**. Вона полягає у наступному. Нехай на відрізку $[a;b]$ розташовані точки c та d так, що $a < c < d < b$ і обчислено значення унімодальної на $[a;b]$ функції: $f(c)$ та $f(d)$. Якщо $f(c) < f(d)$, то $x^* \in [a;d]$, а відрізок $[d;b]$ можна виключити з подальшого розгляду. При $f(c) \geq f(d)$ робимо висновок, що $x^* \in [c;b]$, а відрізок $[a;c]$ надалі можна не розглядати. Отже, внаслідок застосування процедури виключення відрізка отримуємо новий відрізок, вкладений у попередній. Цей новий відрізок містить шукану точку мінімуму. Методи послідовного пошуку використовують процедуру виключення відрізка на кожному черговому кроці такого пошуку. Далі розглянемо основні такі методи.

8.2 Метод дихотомії

Найпростішим методом мінімізації функції однієї змінної, що не потребує обчислення похідної, є **метод дихотомії або ділення відрізка навпіл**. Розглянемо цей метод за умови, що функція $f(x)$ є унімодальною на $[a;b]$. Знайдемо розв’язок задачі одновимірної оптимізації функції $f(x)$ на $[0;1]$

методом дихотомії (за допомогою заміни $x=(b-a)t+a$ довільний відрізок $[a;b]$ можна відобразити на $[0;1]$).

Нехай відомо, що на k-му кроці послідовного пошуку $x^* \in [a_k; b_k] \subseteq [0;1]$ (на першому кроці при $k=1$ $a_1=0$, $b_1=1$). На відрізку $[a_k; b_k]$ довжина якого дорівнює l_k, вибираємо дві точки: $x_{k1}=\frac{a_k+b_k}{2}-\delta$ та $x_{k2}=\frac{a_k+b_k}{2}+\delta$, де δ – вибране достатньо мале число, $0<\delta<1$. Точки x_{k1} та x_{k2} розташовані на відрізку $[a_k; b_k]$ симетрично відносно його середини $x=\frac{a_k+b_k}{2}$. Після вибору цих точок обчислюють значення $f(x_{k1})$ та $f(x_{k2})$ і виконують процедуру виключення відрізка, порівнюючи ці значення між собою. Внаслідок цього отримуємо новий відрізок $[a_{k+1}; b_{k+1}] \subset [a_k; b_k]$. Якщо довжина l_{k+1} нового відрізка не перевищує заданого ε, то обчислення зупиняють і приймають $x^*=\frac{a_{k+1}+b_{k+1}}{2}$. У іншому випадку переходять до наступного, $(k+1)$-го кроку.

Оскільки $l_{k+1}=\frac{l_k}{2}+\delta$ або $l_{k+1}-2\delta=\left(l_k-2\delta\right)/2$, то $l_k-2\delta=\frac{l_1-2\delta}{2^{k-1}}$. З цієї рівності отримуємо наступну формулу для визначення довжини l_k відрізка $[a_k; b_k]$, який отримуємо на k-му кроці застосування методу дихотомії:

$$l_k=\frac{l_1-2\delta}{2^{k-1}}+2\delta. \tag{8.1}$$

З рівності (8.1) випливає, що $l_k \to 2\delta$ при $k\to\infty$, але при цьому $l_k>2\delta$. Тому виконання нерівності $l_{k+1}<\varepsilon$, що свідчить про досягнення заданої точності знаходження точки мінімуму x^*, можливе лише за умови вибору $2\delta<\varepsilon$.

Отже, метод дихотомії полягає у послідовній побудові на кожному k-му кроці пошуку точок

$$x_{k1}=\frac{a_k+b_k}{2}-\delta \text{ та } x_{k2}=\frac{a_k+b_k}{2}+\delta, \tag{8.2}$$

що є симетричними відносно середини відрізка $[a_k; b_k]$ з довжиною l_k. Після виконання k-го кроку буде отримано відрізок $[a_{k+1}; b_{k+1}]$ та обчислено $N=2k$ значень функції $f(x)$. Використовуючи формулу (8.1) довжини інтервалу невизначеності $[a_k; b_k]$ і прийнявши у ній $l_1=1$, отримуємо:

$$l_N=l_{k+1}=\frac{1-2\delta}{2^{k+1}}+2\delta=\frac{1-2\delta}{2^{N/2}}+2\delta.$$

Зазначимо, що після виключення відрізка на k-му кроці алгоритму методу дихотомії точки x_{k1} та x_{k2} належать новому відрізку $[a_{k+1}; b_{k+1}]$, причому одна з них є внутрішньою для цього відрізка. Проте значення функції $f(x)$, обчислене у цій точці, у методі дихотомії не використовують для виключення відрізка на наступному кроці, а здійснюють обчислення у двох нових точках.

Алгоритм методу.

1. Обчислити x_{k1}, x_{k2} за формулами (8.2) (для першої ітерації $k=1$).

2. Обчислити значення $f(x_{k1}), f(x_{k2})$ і порівняти їх. Якщо $f(x_{k1}) \le f(x_{k2})$, то перейти до відрізку $[a_{k+1} = a_k; b_{k+1} = x_{k2}]$; при $f(x_{k1}) > f(x_{k2})$ перейти до відрізку $[a_{k+1} = x_{k1}; b_{k+1} = b_k]$.

3. Знайти точність за формулою $\varepsilon_{k+1} = b_{k+1} - a_{k+1}$, де k – номер ітерації. При умові $\varepsilon_{k+1} > \varepsilon$ повернутися до кроку 1 і виконати наступну $(k+1)$-у ітерацію; якщо $\varepsilon_{k+1} < \varepsilon$, то завершити пошук .

4. Покласти $x^* = \dfrac{a_{k+1} + b_{k+1}}{2}$ та обчислити значення функції $f^* = f(x^*)$.

Приклад. Розв'язати задачу $y = x + \dfrac{2}{x^2} \to \min$, $x \in [1;2]$ з точністю $\varepsilon = 0{,}1$.

Розв'язання. Нехай $\delta = 0{,}02$. Розрахунки наведено у таблиці 8.1.

Таблиця 8.1.

Номер ітерації, k	a_k	b_k	$\varepsilon_k = b_k - a_k$	x_{k1}	x_{k2}	$f(x_{k1})$	$f(x_{k2})$	Порівняння
1	1	2	1	1,48	1,52	2,39308	2,38565	$f(x_{11}) > f(x_{12})$
2	1,48	2	0,52	1,72	1,76	2,39604	2,40566	$f(x_{21}) < f(x_{22})$
3	1,48	1,76	0,28	1,6	1,64	2,38125	2,38361	$f(x_{31}) < f(x_{32})$
4	1,48	1,64	0,16	1,54	1,58	2,38331	2,38115	$f(x_{41}) > f(x_{42})$
5	1,54	1,64	0,1	1,57	1,61	2,38139	2,38158	$f(x_{51}) < f(x_{52})$
6	1,57	1,64	0,07	0,07<0,1 – точність досягнуто				

Звідки $x^* = \dfrac{1{,}57 + 1{,}64}{2} \approx 1{,}605$, $f^* = f\left(x^* = 1{,}605\right) \approx 2{,}3814$.

Існують методи послідовного пошуку, у яких на кожному k-му кроці, починаючи з $k=2$, обчислюють лише одне нове значення функції у точці, що належить відрізку $[a_{k+1}; b_{k+1}]$. Це значення, разом з вже обчисленим на попередньому кроці значенням функції у внутрішній точці відрізка $\left[a_k;b_k\right]$ використовують при виключенні відрізка на наступному кроці послідовного пошуку. Одним з таких методів є метод золотого перерізу.

8.3 Метод золотого перерізу

Розглянемо метод мінімізації унімодальної функції, що дає змогу розв'язати задачу пошуку мінімуму з заданою точністю при меншій кількості обчислень значень функції, який називають методом золотого перерізу.

Золотим перерізом відрізка називають поділ цього відрізка на дві нерівні частини так, щоб відношення довжини всього відрізка до довжини його більшої частини дорівнювала відношенню довжини його більшої частини до довжини меншої.

Розглянемо k-ий крок послідовного пошуку. Щоб виключити відрізок на цьому кроці, потрібно поділити відрізок $\left[a_k;b_k\right]$ двома внутрішніми точками x_{k1} та x_{k2} $(x_{k1}<x_{k2})$ на три частини. Ці точки виберемо симетрично відносно середини відрізка $\left[a_k;b_k\right]$ так, щоб кожна з них здійснювала золотий переріз цього відрізка. У цьому випадку всередині відрізка $\left[a_{k+1};b_{k+1}\right]$ знаходитиметься одна з точок x_{k1} чи x_{k2}, а інша точка буде одним з кінців цього відрізка. При цьому внутрішня точка відрізка $\left[a_{k+1};b_{k+1}\right]$ здійснюватиме його золотий переріз, що випливає з рівності довжин відрізків $\left[a_k;x_{k1}\right]$ та $\left[x_{k2};b_k\right]$. Таким чином, на $(k+1)$-му кроці у одній з точок, x_{k1} чи x_{k2}, обчислювати значення функції $f(x)$ не потрібно. При цьому відношення $\dfrac{l_k}{l_{k+1}}$ довжин відрізків зберігається на кожному кроці пошуку:

$$\frac{l_k}{l_{k+1}}=\frac{l_{k+1}}{l_{k+2}}=\tau=\text{const}. \tag{8.3}$$

Число τ називають **відношенням золотого перерізу**.

Метод послідовного пошуку, у якому на k-му кроці кожна з симетрично вибраних на відрізку $\left[a_k;b_k\right]$ точок x_{k1}, x_{k2} здійснює золотий переріз цього відрізка, називають **методом золотого перерізу**. При застосуванні цього методу кожне виключення відрізка зменшує відрізок, що залишився у τ разів.

Знайдемо величину τ відношення золотого перерізу. Оскільки точки x_{k1} та x_{k2} $(x_{k1}<x_{k2})$ вибрані симетрично відносно середини відрізка $\left[a_k;b_k\right]$, то виконується рівність:

$$b_k-x_{k2}=x_{k1}-a_k=l_k-l_{k+1}.$$

Для визначеності вважатимемо, що на k-му кроці вибрано відрізок $\left[a_k;x_{k2}\right]$. Тоді на $(k+1)$-му кроці однією з точок поділу (правою) буде точка x_{k1}. Тому довжина l_{k+2} відрізка, що вибирається на $(k+1)$-му кроці дорівнює довжині відрізка $\left[a_k; x_{k1}\right]$ і виконується рівність $l_{k+2} = l_k - l_{k+1}$. Підставивши цей вираз для l_{k+2} у формулу (8.3), отримуємо:

$$\frac{l_k}{l_{k+1}} = \frac{l_{k+1}}{l_k - l_{k+1}} = \tau .$$

Звідси знаходимо, що $\tau = \dfrac{1}{\tau - 1}$ або $\tau^2 - \tau - 1 = 0$. Це квадратне рівняння має єдиний додатний корінь $\tau = \dfrac{1+\sqrt{5}}{2} \approx 1,618034$.

Нехай відрізком мінімізації унімодальної функції $f(x)$ є $[0;1]$, тобто $a_1 = 0$, $b_1 = 1$, $l_1 = 1$. На першому кроці послідовного пошуку ($k = 1$) на відрізку $[0;1]$ виберемо дві точки: $x_{11} = a_1 + \left(1 - \dfrac{1}{\tau}\right) b_1 = 1 - \dfrac{1}{\tau}$ та $x_{12} = a_1 + \dfrac{b_1}{\tau} = \dfrac{1}{\tau}$, що здійснюють золотий переріз відрізка $[0;1]$. Обчислимо значення функції $f(x)$ у цих точках і виконаємо процедуру виключення відрізка. Якщо $f(x_{11}) < f(x_{12})$, то вибираємо відрізок $[a_1; x_{12}]$, тобто приймаємо $a_2 = a_1 = 0$, $b_2 = x_{12}$, у іншому випадку виберемо відрізок $[x_{11}; b_1]$, тобто $a_2 = x_{11}$, $b_2 = b_1 = 1$. При цьому у першому випадку приймемо $\tilde{x}_2 = x_{11}$, у другому випадку – $\tilde{x}_2 = x_{12}$. Точка $\tilde{x}_2$ – це одна з точок, що здійснюють золотий переріз відрізка $[a_2; b_2]$, вона є меншою у першому випадку та більшою у другому. Якщо довжина отриманого відрізка є більшою, ніж задана допустима довжина ε інтервалу невизначеності, то переходимо до другого кроку алгоритму, на якому одна з точок, x_{21} чи x_{22} – це точка $\tilde{x}_2$, а другу можна знайти як $a_2 + b_2 - \tilde{x}_2$. На цьому кроці обчислюємо лише одне значення функції $f(x)$ у точці, що є симетричною $\tilde{x}_2$ відносно середини відрізка $[a_2; b_2]$. Якщо довжина l_2 цього відрізка виявилася меншою за ε, то пошук завершено. За наближене значення точки мінімуму x^* приймають $\dfrac{a_2 + b_2}{2}$.

Нехай на k-му кроці, $k \geq 2$, вибрано відрізок $[a_k; b_k]$ і у ньому вибрано точку $\tilde{x}_k$, що здійснює золотий переріз цього відрізка. Значення $f(\tilde{x}_k)$ вже обчислене на попередньому кроці. Другу точку $\bar{x}_k$ золотого перерізу знайдемо за формулою: $\bar{x}_k = a_k + b_k - \tilde{x}_k$.

Обчислимо у цій точці значення функції $f(x)$. Якщо $\bar{x}_k < \tilde{x}_k$, то $x_{k1} = \bar{x}_k$ і $x_{k2} = \tilde{x}_k$, інакше $x_{k1} = \tilde{x}_k$ і $x_{k2} = \bar{x}_k$.

Нехай для визначеності приймемо $\bar{x}_k < \tilde{x}_k$ і $x_{k1} = \bar{x}_k$, $x_{k2} = \tilde{x}_k$. Якщо $f(x_{k1}) < f(x_{k2})$, то вибираємо відрізок $[a_k; x_{k2}]$, тобто приймаємо, що $a_{k+1} = a_k$, $b_{k+1} = x_{k2}$, $\tilde{x}_{k+1} = x_{k1}$. При $f(x_{k1}) > f(x_{k2})$ вибираємо відрізок $[x_{k1}; b_k]$, при цьому $a_{k+1} = x_{k1}$, $b_{k+1} = b_k$, $\tilde{x}_{k+1} = x_{k2}$. Довжину l_{k+1} нового відрізка $[a_{k+1}; b_{k+1}]$

порівнюємо з величиною ε, якщо $l_{k+1} \geq \varepsilon$, то продовжуємо пошук, При $l_{k+1} < \varepsilon$ пошук закінчується за наближене значення x^* точки мінімуму приймаємо $\frac{a_k + b_k}{2}$.

Згідно з алгоритмом методу золотого перерізу, на першому кроці значення функції обчислюють у двох точках, а на кожному з наступних кроків обчислюють лише одне значення функції. Тому після k кроків алгоритма значення функції буде обчислено у $N = k + 1$ точках. Після кожного кроку інтервал невизначеності зменшується у τ разів, тому для довжини l_{k+1} відрізка $\left[a_{k+1}; b_{k+1}\right]$ отримуємо, що $l_{k+1} = \frac{l_1}{\tau^k} = \frac{1}{\tau^{N-1}}$.

Алгоритм методу.

1. На першій ітерації $k = 1$ для початкового відрізку $[a; b]$ прийняти $a_1 = a, b_1 = b$.

2. Обчислити x_{i1}, x_{i2} для $k = 1$ за формулами:

$$x_{k1} = \frac{3-\sqrt{5}}{2}\left(b_1 - a_1\right) + a_1,\ x_{k2} = \frac{\sqrt{5}-1}{2}\left(b_1 - a_1\right) + a_1.$$

3. Обчислити значення функції $f\left(x_{11}\right), f\left(x_{12}\right)$ і порівняти їх. Якщо

$f\left(x_{11}\right) < f\left(x_{12}\right)$, то $a_2 = a_1, b_2 = x_{12}$; $x_{21} = \frac{3-\sqrt{5}}{2}\left(b_2 - a_2\right) + a_2$, $x_{22} = x_{11}$ якщо

$f\left(x_{11}\right) > f\left(x_{12}\right)$, то $a_2 = x_{11}, b_2 = b_1$, $x_{21} = x_{12}$, $x_{22} = \frac{\sqrt{5}-1}{2}\left(b_2 - a_2\right) + a_2$.

4. Перевірити умову закінчення пошуку: $\varepsilon_1 = b_2 - a_2$. При умові $\varepsilon_1 > \varepsilon$ для нового відрізку $\left[a_2; b_2\right]$ повернутися до кроку 2 і виконати наступну $(k+1)$-*у* ітерацію; якщо $\varepsilon_1 < \varepsilon$, то завершити пошук.

5. Покласти $x^* = \frac{a_{k+1} + b_{k+1}}{2}$ та обчислити значення функції $f^* = f\left(x^*\right)$.

Приклад. Розв'язати задачу $y = x + \frac{2}{x^2} \to \min$, $x \in [1; 2]$ з точністю $\varepsilon = 0{,}1$.

Розв'язання. Розрахунки наведено у таблиці 8.2.

Таблиця 8.2.

Номер ітерації, k	$\left[a_k; b_k\right]$	$\varepsilon_k = b_k - a_k$	x_{k1}	x_{k2}	$f\left(x_{k1}\right)$	$f\left(x_{k2}\right)$	Порівняння
1	[1;2]	1	1,382	1,618	2,42916	2,38196	$f\left(x_{11}\right) > f\left(x_{12}\right)$
2	[1,382; 2]	0,618	1,618	1,7639	2,38196	2,40671	$f\left(x_{21}\right) < f\left(x_{22}\right)$

Продовження таблиці 8.2.

Номер ітерації, k	$[a_k;b_k]$	$\varepsilon_k = b_k - a_k$	x_{k1}	x_{k2}	$f(x_{k1})$	$f(x_{k2})$	Порівняння
3	[1,382; 1,7639]	0,382	1,528	1,618	2,38462	2,38196	$f(x_{31}) > f(x_{32})$
4	[1,528; 1,7639]	0,236	1,618	1,674	2,38196	2,38765	$f(x_{41}) < f(x_{42})$
5	[1,528; 1,674]	0,146	1,584	1,618	2,38111	2,38196	$f(x_{51}) < f(x_{52})$
6	[1,528; 1,618]	0,09<0,1	0,09<0,1 – точність досягнуто				

Звідки $x^* = \dfrac{1,528+1,618}{2} \approx 1,573$, $f^* = f\left(x^* = 1,573\right) \approx 2,3813$.

8.4 Метод Фібоначчі

Нехай при пошуку точки $x^* \in [0;1]$, у якій унімодальна на цьому відрізку функція $f(x)$ досягає мінімуму, можна обчислити значення $f(x)$ лише у двох точках. Тоді більш ефективним є метод дихотомії, оскільки він дозволяє зменшити інтервал невизначеності майже вдвічі, а метод золотого перерізу – лише у $\dfrac{\sqrt{5}+1}{2}$ разів. У той же час, якщо можна обчислювати значення функції $f(x)$ велику кількість разів, метод золотого перерізу є ефективнішим, ніж метод дихотомії.

При довільній заданій загальній кількості обчислень $N > 2$ значення функції $f(x)$ можна побудувати метод, що є більш ефективним у порівнянні з методами дихотомії та золотого перерізу, який потребує виконання $N-1$ кроків. Він поєднує у собі переваги симетричного розташування нових точок на відрізку $[a_k;b_k]$ відносно його середини, що має місце у методах дихотомії та золотого перерізу, з можливістю зміни на кожному кроці відношення $\dfrac{l_k}{l_{k+1}}$ довжин попереднього та нового відрізків невизначеності. У випадку вибору нових внутрішніх точок симетрично відносно середини попереднього інтервалу невизначеності для трьох послідовних кроків методу золотого перерізу виконується співвідношення $l_{k-1} = l_k + l_{k+1}$, $k = 2, 3, \ldots$.

Нехай потрібно, маючи можливість обчислити значення функції $f(x)$ у точках $x_k \in [0;1]$, $k = 1, 2, \ldots, N$, знайти точку x^* мінімуму цієї функції, звузивши інтервал невизначеності до найменшої довжини. На останньому, $(N-1)$-му кроці маємо відрізок $[a_{N-1}, b_{N-1}]$ довжиною l_{N-1} з двома внутрішніми точками x_{N-1} та x_N, розташованими симетрично середині відрізка на достатньо малій відстані 2δ між ними. У цих точках знайдені значення функції $f\left(x_{N-1}\right)$ та $f\left(x_N\right)$. Приймемо для визначеності, що $f\left(x_N\right)< f\left(x_{N-1}\right)$. Тоді для нового відрізка $\left[a_N, b_N\right]$, що має довжину $l_N=\dfrac{l_{N-1}}{2}+\delta$, внутрішньою буде точка x_N, а точка x_{N-1} співпаде з однією з його меж. Тут при виборі наближеного значення $x^*=x_N$ довжина інтервалу невизначеності дорівнюватиме поки що невідомій довжині l_N відрізка $\left[a_N, b_N\right]$. Через l_N можна виразити довжину $l_{N-1}=2l_N-2\delta$ відрізка $\left[a_{N-1}, b_{N-1}\right]$. Далі, у відповідності з співвідношеннями між довжинами відрізків $l_{k-1}=l_k+l_{k+1}$, отримуємо:

$$l_{N-2}=l_{N-1}+l_N=3l_N-2\delta,\ l_{N-3}=l_{N-2}+l_{N-1}=5l_N-4\delta, \ldots .$$

У загальному вигляді маємо:

$$l_{N-k}=F_{k+2}l_N-2F_k\delta,\ k=0,\,1,\,\ldots,\,N-1. \tag{8.4}$$

Тут коефіцієнти F_m визначаються рекурентним співвідношенням:

$$F_m=F_{m-1}+F_{m-2},\ m=3,\ldots,N-1,\ F_1=F_2=1. \tag{8.5}$$

Оскільки при $k=N-1$ довжина $l_{N-k}=l_1=1$ відрізка $\left[0;1\right]$ є відомою, то з формули (8.4) можна знайти довжину інтервалу невизначеності:

$$l_N=\frac{l_1}{F_{N+1}}+2\delta\frac{F_{N-1}}{F_{N+1}}. \tag{8.6}$$

Існує алгоритм методу прямого пошуку, що задовольняє співвідношення (8.6). Всі коефіцієнти F_m належать множині натуральних чисел, їх називають **числами Фібоначчі** ($F_1=1, F_2=1, F_3=2, F_4=3, F_5=5, F_6=8, F_7=13,\ldots$). Метод, що використовує числа Фібоначчі для вибору довжин відрізків l_k і, відповідно, точок x_k, $k=1,\,2,\,\ldots,\,\mathrm{N}$, у яких обчислюють значення мінімізованої функції $f\left(x\right)$, називають **методом Фібоначчі** або методом оптимального послідовного пошуку.

Якщо на першому кроці пошуку інтервал невизначеності має довжину l_1, то у відповідності з (8.4) та (8.6) довжина l_2 нового відрізка $\left[a_2, b_2\right]$ дорівнює:

$$l_2=F_Nl_N-2\delta F_{N-2}=\frac{F_N}{F_{N+1}}l_1+2\delta\frac{F_NF_{N-1}-F_NF_{N-2}}{F_{N+1}}=\frac{F_N}{F_{N+1}}l_1+\left(-1\right)^{N+1}\frac{2\delta}{F_{N+1}}.$$

Будемо вважати δ достатньо малим так, що

$$\frac{l_1}{l_2}=\frac{F_{N+1}}{F_N}. \tag{8.7}$$

У цьому випадку виконання процедури виключення відрізка на останньому $\left(N-1\right)$-му кроці пошуку приводить до співпадання внутрішніх точок x_{N-1} та x_N. При $l_1=1$ з (8.7) знаходимо $l_2=\dfrac{F_N}{F_{N+1}}$. Таким чином, враховуючи (8.5), маємо, що на першому кроці вибір точок, симетричних відносно середини відрізка $[0;1]$, можна визначити за формулами:

$$x_1 = l_2 = \frac{F_N}{F_{N+1}},\ x_2 = 1 - l_2 = 1 - \frac{F_N}{F_{N+1}} = \frac{F_{N-1}}{F_{N+1}},\ x_2 < x_1,$$

причому відстань між ними дорівнюватиме $d_1 = x_1 - x_2 = \frac{F_N}{F_{N+1}} - \frac{F_{N-1}}{F_{N+1}} = \frac{F_{N-2}}{F_{N+1}}$.

Після виконання на цьому кроці процедури виключення відрізка одна з точок, x_1 чи x_2, буде граничною точкою нового відрізку $[a_2;a_1]$, а друга – його внутрішньою точкою. Позначимо її x_2'. Інша внутрішня точка на цьому відрізку повинна бути вибраною симетрично точці x_2' відносно його середини. Аналогічно здійснюється вибір другої внутрішньої точки нового відрізка на всіх наступних кроках пошуку.

На i-му кроці у відповідності з рівністю (8.4), у якій потрібно прийняти $k = N - i$, та рівністю (8.6) довжина відрізка $[a_i;b_i]$ дорівнює $l_i = \frac{F_{N+2-i}}{F_{N+1}}$. На цьому кроці довжина відрізка $[a_i;b_i]$ зменшується у $\frac{l_i}{l_{i+1}} = \frac{F_{N+2-i}}{F_{N+1-i}}$ разів. Якщо внутрішні точки на цьому відрізку позначити α_i та β_i, то для їх знаходження маємо формули:

$$\alpha_i = a_i + \frac{F_{N-i}}{F_{N+1}},\ \beta_i = a_i + \frac{F_{N+1-i}}{F_{N+1}},\ \alpha_i < \beta_i,\ i = 1, \ldots, N-1.$$

Отже, застосування методу Фібоначчі передбачає апріорне визначення потрібної кількості N обчислень значень функції $f(x)$ (або кількості кроків пошуку). Величина N потрібна для здійснення першого кроку алгоритму для вибору на цьому кроці точок x_1 та x_2 поділу відрізка $[a_1;b_1]$.

Алгоритм методу.

1. Обрати число обчислень n, використовуючи формулу $F_{n+2} \ge \frac{b-a}{\varepsilon}$, де $[a;b]$ – початковий відрізок, ε – точність.

2. Покласти $a_1 = a, b_1 = b$ та обчислити значення точок x_{k1}, x_{k2} при $k = 1$ за формулами:

$$x_{k1} = a_k + (b_k - a_k)\frac{F_{n+1-k}}{F_{n+3-k}},\ x_{k2} = a_k + (b_k - a_k)\frac{F_{n+2-k}}{F_{n+3-k}}.$$

3. Обчислити значення функції $f(x_{k1}), f(x_{k2})$ при $k = 1$ і порівняти їх. Якщо $f(x_{11}) < f(x_{12})$, то $a_2 = a_1, b_2 = x_{12}$; $x_{21} = x_{k1}$ при $k = 2$, $x_{22} = x_{11}$;
якщо $f(x_{11}) > f(x_{12})$, то $a_2 = x_{11}, b_2 = b_1$, $x_{21} = x_{12}$, $x_{22} = x_{k2}$ при $k = 2$.

4. Крок 3 повторити $k = n$ раз.

5. При $k = n$ точки x_{k1}, x_{k2} співпадають і відповідають середині відрізка, тому для того, щоб забезпечити подальше скорочення відрізка, точка останнього обчислення функції переміщується вправо на величину $\delta > 0, \delta << \varepsilon$.

6. Покласти $x^* = \dfrac{a_{k+1} + b_{k+1}}{2}$ та обчислити значення функції $f^* = f\left(x^*\right)$.

Приклад. Розв'язати задачу $y = x + \dfrac{2}{x^2} \to \min$, $x \in [1;2]$ з точністю $\varepsilon = 0{,}1$.

Розв'язання. Оскільки $F_{n+2} \geq \dfrac{2-1}{0{,}1} = 10$, то $n = 5$. Розрахунки наведено у таблиці 8.3.

Таблиця 8.3.

Номер ітерації, k	$[a_k;b_k]$	x_{k1}	x_{k2}	$f(x_{k1})$	$f(x_{k2})$	Порівняння
1	[1;2]	1,3846	1,6154	2,42783	2,38182	$f(x_{11}) > f(x_{12})$
2	[1,3846; 2]	1,6154	1,7692	2,38182	2,40816	$f(x_{21}) < f(x_{22})$
3	[1,3846; 1,7692]	1,5384	1,6154	2,38346	2,38182	$f(x_{31}) > f(x_{32})$
4	[1,5384; 1,7692]	1,6154	1,6923	2,38182	2,39065	$f(x_{41}) < f(x_{42})$
5	[1,5384; 1,6923]	1,6154	1,6154	2,38182		

Оскільки $x_{51} = x_{52}$, для забезпечення заданої точності переміщуємо точку останнього обчислення функції вправо на величину константи $\delta = 0{,}02$: $x_{52} = 1{,}6154 + 0{,}02 = 1{,}6354$. Обчислимо $f\left(x_{52}\right) = f\left(1{,}6354\right) \approx 2{,}3831$. Оскільки $f\left(x_{51}\right) \approx 2{,}38182 < f\left(x_{52}\right) = f\left(1{,}6354\right) \approx 2{,}3831$, то остаточний відрізок $[a_6;b_6] =$ = [1,5384; 1,6354]. За точку мінімуму приймають середину відрізку
$x^* = \dfrac{1{,}5384 + 1{,}6354}{2} \approx 1{,}5869$, $f^* = f\left(x^* = 1{,}5869\right) \approx 2{,}3811$. Точність
$\varepsilon = 1{,}6354 - 1{,}5384 = 0{,}097 < 0{,}1$.

8.5 Загальна характеристика методів багатовимірної оптимізації

Розглянемо задачу безумовної оптимізації функції $f(x)$, заданої на точках x n-вимірного евклідового простору E^n. Усі методи пошуку оптимізаційної задачі $f(x) \to \min$ приводять до побудови деякої послідовності точок x^0, $x^1, \ldots, x^k, \ldots$ з простору E^n. Елементи цієї послідовності повинні відповідати умові монотонності $f\left(x^0\right) > f\left(x^1\right) > \ldots > f\left(x^k\right) > \ldots$. Такі послідовності точок

$\{x^k\}$ називають **релаксаційними**, а методи їх побудови – **методами спуску**. У цих методах точки послідовності $\{x^k\}$ обчислюються за формулою $x^{k+1} = x^k + \alpha_k v^k$, де $\alpha_k \geq 0$ – довжина кроку у напрямку вектора $v^k \in E^n$, $k = 0, 1, \ldots$.Відмінність методів спуску полягає у способах вибору v^k та α_k. Проте для виконання умови монотонності послідовності $\{x^k\}$ на вибір довжини кроку α_k повинні накладатися певні обмеження. Наприклад, для мінімізації квадратичної функції $f(x) = (Dx, x) + (b, x)$ (тут і надалі позначатимемо символом (a,b) скалярний добуток векторів $a \in E^n$ та $b \in E^n$) , де $x \in E^n$, $b \in E^n$, D – симетрична додатно визначена матриця, що забезпечує опуклість вниз функції $f(x)$, маємо:

$$f\left(x^{k+1}\right) = f\left(x^k + \alpha \cdot v^k\right) = \left(D\left(x^k + \alpha \cdot v^k\right), x^k + \alpha \cdot v^k\right) + \left(b, x^k + \alpha \cdot v^k\right) =$$
$$= \left(Dv^k, v^k\right)\alpha^2 + \left(2Dx^k + b, v^k\right)\alpha + \left(Dx^k + b, x^k\right) = \frac{1}{2}\left(f''\left(x^k\right)v^k, v^k\right)\alpha^2 +$$
$$+\left(f'\left(x^k\right), v^k\right)\alpha + f\left(x^k\right).$$

Тут $f''\left(x^k\right)$ – матриця других частинних похідних (матриця Гессе), а $f'\left(x^k\right)$ – вектор – градієнт функції $f(x)$ у точці x^k. З останньої рівності знаходимо:

$$f\left(x^{k+1}\right) - f\left(x^k\right) = \left(Dv^k, v^k\right)\alpha^2 + \left(2Dx^k + b, v^k\right)\alpha = \frac{1}{2}\left(f''\left(x^k\right)v^k, v^k\right)\alpha^2 +$$
$$+\left(f'\left(x^k\right), v^k\right)\alpha.$$

Умова монотонності $f\left(x^{k+1}\right) < f\left(x^k\right)$ виконується для таких значень α, що

$$0 < \alpha < -\frac{\left(2Dx^k + b, v^k\right)}{\left(Dv^k, v^k\right)} = -2\frac{\left(f'\left(x^k\right), v^k\right)}{\left(f''\left(x^k\right)v^k, v^k\right)}.$$

Звідси випливає, що $\left(f'\left(x^k\right), v^k\right) < 0$. Це означає, що вектор v^k, який вказує напрямок руху від точки x^k до точки x^{k+1}, має розміщуватися у тому самому півпросторі з E^n, що і вектор $\left(-f'\left(x^k\right)\right)$ та визначається гіперплощиною з вектором нормалі $f'\left(x^k\right)$.

Зазначимо, що функція $f(x)$ максимально зменшується у напрямку v^k, якщо вибрати

$$\alpha_k = \alpha_{\min} = -\frac{\left(2Dx^k + b, v^k\right)}{2\left(Dv^k, v^k\right)} = \frac{-\left(f'\left(x^k\right), v^k\right)}{\left(f''\left(x^k\right)v^k, v^k\right)}.$$

Важливою характеристикою методів спуску є швидкість їх збіжності. Кажуть, що має місце **лінійна збіжність** або **збіжність зі швидкістю геометричної прогресії**, якщо $\left\| x^{k+1} - x^* \right\| \le q \left\| x^k - x^* \right\|$, де x^* – точка мінімуму функції $f(x)$, $0 < q < 1$ – деяка константа. Швидкість збіжності є **надлінійною**, якщо $\left\| x^{k+1} - x^* \right\| \le q_k \left\| x^k - x^* \right\|$, де $q_k \to 0$ при $k \to \infty$. Швидкість збіжності є **квадратичною**, якщо $\left\| x^{k+1} - x^* \right\| \le c \left\| x^k - x^* \right\|^2$, $c \ge 0$.

8.6 Градієнтні методи

Нагадаємо, що градієнт функції $f(x)$ у точці x^k є вектором, компонентами якого є частинні похідні $f(x)$ по відповідним змінним. Градієнт функції $f(x)$ у точці x^k будемо позначати $f'\left(x^k\right)$ або $\nabla f\left(x^k\right)$.

Вектор градієнта функції $f(x)$ у точці x^k вказує напрямок найшвидшого зростання значення цієї функції з точки x^k, крім того, він ортогональний лінії (поверхні) рівня функції $f(x)$, яка проходить через цю точку.

Вектор $\left(-\nabla f\left(x^k\right)\right)$ – антиградієнт, направлений у бік найшвидшого зменшення функції значення функції $f(x)$ у точці x^k. Якщо вибрати вектор v^k як антиградієнт цієї функції, то схема ітераційного процесу у методах спуску набуває вигляду:

$$x^{k+1} = x^k - \alpha_k \nabla f\left(x^k\right),\ \alpha_k > 0,\ k = 0,\ 1,\ \ldots, \tag{8.8}$$

де k – номер ітерації.

Усі ітераційні процеси оптимізації функції $f(x)$, в яких напрямок руху на кожному кроці k збігається з антиградієнтом або градієнтом функції $f(x)$ у точці x^k, називають **градієнтними методами**. Оскільки для їх застосування необхідним є обчислення перших похідних $f(x)$, то ці методи належать до оптимізаційних методів першого порядку. Вони відрізняються між собою лише способом вибору кроку α_k. Серед способів вибору α_k найпоширенішими є два: з подрібленням кроку та оптимальним кроком.

У першому випадку у разі переходу від точки x^k до точки x^{k+1} для вибору α_k перевіряють умову

$$f\left(x^k - \alpha_k \nabla f\left(x^k\right)\right) - f\left(x^k\right) \le -\varepsilon \cdot \alpha_k \left\| f\left(x^k\right) \right\|^2, \tag{8.9}$$

де $\varepsilon \in (0;1)$ – довільно вибрана стала, однакова для всіх ітерацій. Умова (8.9) означає зменшення значення функції $f(x)$ у точці x^{k+1} у порівнянні з її значенням у точці x^k.

У другому випадку при переході з точки x^k до точки x^{k+1} функція мінімізується по α. Градієнтні методи з оптимальним кроком α називають **методами найшвидшого градієнтного спуску**.

Розглянемо градієнтні методи з дробленням кроку. Здійснення ітераційного процесу (8.8) вимагає формулювання правил вибору кроку α_k, за якого зменшується значення функції $f(x)$, тобто виконується нерівність:

$$f\left(x^k - \alpha_k \nabla f\left(x^k\right)\right) \le f\left(x^k\right). \tag{8.10}$$

Вибір малого кроку α_k може привести до надзвичайно великої кількості ітерацій, необхідних для досягнення точки мінімуму функції $f(x)$. Вибір великого значення цього кроку може привести до порушення умови (8.10) або до коливання наближень x^k відносно шуканої точки мінімуму.

Величина кроку α_k повинна вибиратися з умови виконання вимоги (8.9). Ітераційний процес (8.8) з вибором кроку α_k, що задовольняє умову (8.9), виконують за наступним алгоритмом.

1. Вибирають $\alpha > 0$.
2. На k-й ітерації перевіряють умову (8.9) з $\alpha_k = \alpha$. Якщо умова виконується, то переходять до $(k+1)$-ї ітерації. Інакше α зменшується (дробиться) доти, доки умова (8.9) не буде виконана, після чого переходять до $(k+1)$-ї ітерації

Перевірка умови (8.9) на кожному кроці є досить трудомісткою процедурою, тому можна використати процедуру (8.8) з постійним на всіх ітераціях кроком $\alpha_k = \alpha = \dfrac{1-\varepsilon}{R}$, де ε – те ж саме, що і у формулі (8.9), R – константа, така, що нерівність Ліпшиця

$$\left\|\nabla f(x) - \nabla f(y)\right\| \le R\|x - y\| \tag{8.11}$$

виконується для довільних x та y з E^n.

У методах **найшвидшого градієнтного спуску** α_k вибирають на кожній ітерації за умови

$$f\left(x^k - \alpha_k \nabla f\left(x^k\right)\right) = \min_{\alpha \ge 0} f\left(x^k - \alpha \nabla f\left(x^k\right)\right). \tag{8.12}$$

Розглянемо схему методу найшвидшого градієнтного спуску на $(k+1)$-й ітерації.

1. Для знайденої точки x^k розв'язують задачу одновимірної оптимізації (8.12) відносно α.
2. Знайдене α є кроком α_k для $(k+1)$-ї ітерації, що виконується за схемою (8.8).

Зазначимо, що оскільки α_k вибирають за умови мінімізації функції

$$\varphi(\alpha) = f\left(x^k - \alpha \nabla f\left(x^k\right)\right),$$

то для вибраного значення α_k повинна виконуватися рівність $\varphi'\left(\alpha_k\right) = -\nabla f\left(x^{k+1}\right)\cdot\nabla f\left(x^k\right) = 0$, тобто напрямки спуску для двох послідовних ітерацій повинні бути взаємно ортогональними.

Збіжність методів градієнтного спуску визначається теоремою, яку наведемо тут без доведення.

Теорема 8.1 Якщо функція $f\left(x\right)$ є обмеженою знизу, її градієнт задовольняє умову Ліпшиця (8.11) і крок α_k вибирають одним з описаних вище способів, то при довільному початковому наближенні x^0 для послідовності точок $\left\{x^k\right\}$, що будується за формулою (8.8), виконується умова $\lim\limits_{k\to\infty}\left\|\nabla f\left(k\right)\right\| = 0$.

При практичній реалізації ітераційного процесу його завершують при виконанні умови $\left\|\nabla f\left(x^k\right)\right\| \le \delta$, де $\delta > 0$ – деяке задане число, що характеризує точність отриманого розв'язку.

Питання для самоперевірки

1. Які задачі відносяться до задач одновимірної оптимізації? Навести приклади.
2. Надати означення унімодальної функції та навести приклади таких функцій.
3. Сформулювати необхідні та достатні умови екстремуму диференційовної функції однієї змінної.
4. Навести класифікацію методів пошуку екстремумів функцій.
5. Який інтервал називають інтервалом невизначеності?
6. У чому полягає основна відмінність між методами пасивного та послідовного пошуку?
7. У чому полягає процедура виключення відрізка?
8. Охарактеризувати етапи процесу послідовного пошуку при застосуванні методу дихотомії.
9. Які етапи входять до алгоритму методу золотого перерізу?
10. Що називають золотим перерізом?
11. Вкажіть переваги методу золотого перерізу у порівнянні з методом дихотомії. У яких випадках більш ефективним є метод дихотомії?
12. Наведіть рекурентне співвідношення, яке визначає послідовність чисел Фібоначчі. Вкажіть кілька перших членів цієї послідовності.
13. Які кроки входять до алгоритму методу Фібоначчі?
14. Що можна віднести до переваг та недоліків методу Фібоначчі?
15. Які послідовності називають релаксаційними?
16. Які методи називають методами спуску?
17. Вкажіть види збіжності методів спуску за її швидкістю.
18. Надати означення градієнта функції кількох змінних та матриці Гессе цієї функції.

19. Які методи багатовимірної оптимізації відносять до градієнтних?
20. Яка умова перевіряється при застосуванні градієнтних методів з подрібленням кроку?
21. Яка умова перевіряється при виконанні ітерацій у методах найшвидшого градієнтного спуску?

Практичні завдання

Розв'язати задачу з точністю $\varepsilon = 0{,}1$.

1. $y = x^2 - 3\sin x \to \min$, $x \in [1;2]$.
2. $y = x^4 + x + 1 \to \min$, $x \in [-1;0]$.
3. $y = e^x + \dfrac{2}{x} \to \min$, $x \in [0{,}5;1{,}5]$.
4. $y = x^2 + e^{-x} \to \min$, $x \in [-1;1{,}5]$.
5. $y = x\sin x \to \min$, $x \in [-6;-4]$.
6. $y = x + \dfrac{1}{x^3} \to \min$, $x \in [1;2]$.
7. $y = -x + 2\cos x \to \min$, $x \in [-3;-2]$.
8. $y = \dfrac{x^2}{2-x} \to \min$, $x \in [-0{,}5;0{,}5]$.
9. $y = x^2 - 2x \to \min$, $x \in [0{,}5;1{,}5]$.
10. $y = x^3 + e^{-x} \to \min$, $x \in [0;1]$

РЕКОМЕНДОВАНА ЛІТЕРАТУРА

Основна

1. Аттетков А. В., Галкин С. В., Зарубин В. С. Методы оптимизации : учеб. для вузов. Москва : МГТУ им. Н.Э. Баумана, 2003. 440 с.
2. Банди Б. Методы оптимизации. Вводный курс. Москва : Радио и связь, 1988. 128 с.
3. Буслаев В. С. Вариационное исчисление. Ленинград : Изд-во Ленинградского университета, 1980. 288 с.
4. Ванько В. И., Ермошина О. В., Кувыркин Г. Н. Вариационное исчисление и оптимальное управление. Москва : Изд-во МГТУ им. Н.Э. Баумана, 2006. 488 с.
5. Жалдак М. І., Тривус Ю. В. Основи теорії і методів оптимізації. Черкаси : Брама-Україна, 2005. 608 с.
6. Карташов А. П., Рождественский Б. Л. Обыкновенные дифференциальные уравнения и основы вариационного исчисления. Москва : Наука, 1986. 238с.
7. Клименко В. І., Панасенко Є. В., Стреляєв Ю. М., Ткаченко І. Г. Варіаційне числення та методи оптимізації : навчальний посібник для студентів освітньо-кваліфікаційного рівня «бакалавр», напряму підготовки «Математика». Запоріжжя : ЗНУ, 2015. 84 с.
8. Краснов М. Л., Макаренко Г. И., Киселев А. И. Вариационное исчисление : задачи и упражнения. Москва : Наука, 1973. 190 с.
9. Моклячук М. П. Варіаційне числення: екстремальні задачі. Київ : Либідь, 1994. 328 с.
10. Моклячук М. П. Збірник задач з варіаційного числення та методів оптимізації. Київ : ВПЦ «Київський університет», 2014. 255 с.
11. Пантелеев А. В., Летова Т. А. Методы оптимизации в примерах и задачах. Москва : Высшая школа, 2002. 543 с.
12. Эльсгольц Л. Э. Дифференциальные уравнения и вариационное исчисление. Москва : Наука, 1969. 424 с.

Додаткова

1. Бердичевский В. Л. Вариационные принципы механики сплошной среды. Москва : Наука, 1983. 448 с.
2. Галеев Э.М., Тихомиров В.Н. Краткий курс теории экстремальных задач. Москва : Изд-во Моск. ун-та, 1989. 204 с.
3. Коша А. Вариационное исчисление. Москва : Высшая школа, 1983. 280 с.
4. Ректорис К. Вариационные методы в математической физике. Москва : Мир, 1985. 590 с.
5. Спасский Р. А. Классическое вариационное исчисление и вариационные принципы механики. Севастополь : СевНТУ, 2004. 176 с.
6. Янг Л. Лекции по вариационному исчислению и оптимальному управлению. Москва : Мир, 1974. 488с.

ВИКОРИСТАНА ЛІТЕРАТУРА

1. Краснов М. Л., Макаренко Г. И., Киселев А. И. Вариационное исчисление: задачи и упражнения. Москва : Наука, 1973. 190 с.
2. Моклячук М. П. Варіаційне числення: екстремальні задачі. Київ : Либідь, 1994. 328 с.
3. Эльсгольц Л. Э. Дифференциальные уравнения и вариационное исчисление. Москва : Наука, 1969. 424 с.

Навчальне видання
(українською мовою)

Клименко Михайло Іванович
Швидка Світлана Петрівна
Кондрат'єва Наталія Олександрівна

ВАРІАЦІЙНЕ ЧИСЛЕННЯ ТА МЕТОДИ ОПТИМІЗАЦІЇ

Навчальний посібник
для здобувачів ступеня вищої освіти магістра
спеціальності «Прикладна математика»
освітньо-професійної програми «Прикладна математика»

Рецензент *С. М. Гребенюк*
Відповідальний за випуск *В. З. Грищак*
Коректор *С. П. Швидка*