Екзаменаційна робота з математичного аналізу

Студента 2 курсу групи МП-21 Захарова Дмитра Олеговича

9 червня 2023 р.

1 Питання 1.

1.1 Означення локального та абсолютного екстремумів функції декількох змінних. Необхідна умова локального екстремуму.

Отже, почнемо, з яким об'єктом ми маємо справу. Нехай ми маємо деяку область $G \subset \mathbb{R}^m$ і на цій множині визначена функція $f : G \to \mathbb{R}$. Також, розглядаємо деяку внутрішню точку $\boldsymbol{a} \in G$. Тепер, наведемо означення, коли точка $\boldsymbol{a}$ є точкою локального максимуму або мінімуму.

Означення 1: Точка локального максимуму або мінімуму

Точка $\boldsymbol{a}$ є **точкою локального максимуму** функції f (з умовами, наведеними вище), якщо

$$\exists \mathcal{U}_\delta(\boldsymbol{a}) \subset G \; \forall \boldsymbol{x} \in \mathcal{U}_\delta(\boldsymbol{a}) : f(\boldsymbol{x}) \le f(\boldsymbol{a})$$

Відповідно, точка $\boldsymbol{a}$ є **точкою локального мінімуму** функції f, якщо

$$\exists \mathcal{U}_\delta(\boldsymbol{a}) \subset G \; \forall \boldsymbol{x} \in \mathcal{U}_\delta(\boldsymbol{a}) : f(\boldsymbol{x}) \ge f(\boldsymbol{a})$$

Також, розглянемо випадок **строгого** максимуму або мінімуму.

Означення 2: Точка строгого локального максимуму або мінімуму

Точка $\boldsymbol{a}$ є **точкою строгого локального максимуму** функції f, якщо

$$\exists \mathcal{U}_\delta(\boldsymbol{a}) \subset G \; \forall \boldsymbol{x} \in \mathring{\mathcal{U}}_\delta(\boldsymbol{a}) : f(\boldsymbol{x}) < f(\boldsymbol{a}),$$

де $\mathring{\mathcal{U}}_\delta(\boldsymbol{a}) = \mathcal{U}_\delta(\boldsymbol{a}) \setminus \{\boldsymbol{a}\}$. Для локального **мінімуму** означення аналогічне, тільки замість умови $f(\boldsymbol{x}) < f(\boldsymbol{a})$ маємо умову $f(\boldsymbol{x}) > f(\boldsymbol{a})$

Нарешті, для точки **глобального** мінімуму або максимуму маємо

Означення 3: Точка глобального максимуму або мінімуму

Точка $\boldsymbol{a}$ є **точкою глобального максимуму** якщо

$$\forall \boldsymbol{x} \in G : f(\boldsymbol{x}) \le f(\boldsymbol{a})$$

Відповідно, точка **строгого** глобального максимуму це

$$\forall \boldsymbol{x} \in G \setminus \{\boldsymbol{a}\} : f(\boldsymbol{x}) < f(\boldsymbol{a})$$

Аналогічні означення для **точок (строгого) глобального мінімуму**, проте нерівності треба замінити на $f(\boldsymbol{x}) \ge f(\boldsymbol{a}), f(\boldsymbol{x}) > f(\boldsymbol{a})$.

Точка $\boldsymbol{a} \subset G$ є точкою **локального екстремуму**, якщо точка $\boldsymbol{a}$ або є точкою локального максимуму, або точкою локального мінімуму.

Точка $\boldsymbol{a} \subset G$ є точкою **абсолютного екстремуму**, якщо точка $\boldsymbol{a}$ або є точкою глобального максимуму, або точкою глобального мінімуму.

Приклад 1: Про мінімуми і максимуми

Нехай, наприклад, маємо функцію $f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$ і $G = \mathbb{R}^3$. Тоді точка $\boldsymbol{a} = [0, 0, 0]^\top$ є як точкою локального мінімуму, так і точкою глобального мінімуму. Тому, ця точка є як локальним, так і абсолютним екстремумом.

Тепер, сформулюємо і доведемо необхідну умову локального екстремуму.

Теорема 1: Необхідна умова локального екстремуму.

Нехай маємо множину $G \subset \mathbb{R}^m$, функцію $f : G \to \mathbb{R}$ і нам відомо, що $\boldsymbol{a} \subset G$ є точкою локального екстремуму. Якщо f має в цій точці $\{f'_{x_k}(\boldsymbol{a})\}_{k=1}^m \subset \mathbb{R}$, тоді $\forall k \in \{1, \ldots, m\}$ маємо $f'_{x_k}(\boldsymbol{a}) = 0$ (або, аналогічно, $\operatorname{grad} f(\boldsymbol{a}) = \mathbf{0}$).

Доведення. Виділимо деякий дельта-окіл $\mathcal{U}_\delta(\boldsymbol{a}) \subset G$. Також візьмемо довільний $k \in \{1, \ldots, m\}$ і будемо розглядати поведінку функції якщо рухатись від точки $\boldsymbol{a}$ "паралельно" напрямку x_k. Введемо функцію

$$\varphi_k(t) := f(a_1, \ldots, a_{k-1}, t, a_{k+1}, \ldots, a_m), \tag{1}$$

де t не виходить за наш обраний дельта-окіл $\mathcal{U}_\delta(\boldsymbol{a})$, тобто, $t \in (a_k - \delta, a_k + \delta)$. Помітимо, що $\varphi'_k(a_k) = f'_{x_k}(\boldsymbol{a})$. Застосовуємо **теорему Ферма**: функція f визначена на $(a_k - \delta, a_k + \delta)$ і в точці a_k вона приймає або найбільше, або найменше значення (за умовою, оскільки точка $\boldsymbol{a}$ є точкою локального екстремуму) і має в цій точці скінченну похідну, тоді $\varphi'_k(a_k) = 0$, тому і $f'_{x_k}(\boldsymbol{a}) = 0$. Теорему доведено.

1.2 Класифікація квадратичних форм. Критерій Сільвестера.

Для початку розглянемо означення квадратичної форми.

Означення 4: Квадратична форма

Вираз $Q(\boldsymbol{x}) = \sum_{i,j=1}^m \alpha_{i,j} x_i x_j$ називають квадратичною формою.

Також зручно розглянути ще одну форму запису квадратичної форми. Помітимо, що:

$$Q(x_1, \ldots, x_m) = [x_1, \ldots, x_m] \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & \ldots & \alpha_{1,m} \\ \alpha_{2,1} & \alpha_{2,2} & \ldots & \alpha_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m,1} & \alpha_{m,2} & \ldots & \alpha_{m,m} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \tag{2}$$

Матриця $\boldsymbol{A} = \{\alpha_{i,j}\}_{i,j=1}^m$ називається матрицею квадратичної форми. Запис 2 можна записати більш лаконічно, а саме $Q(\boldsymbol{x}) = \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}$, де $\boldsymbol{x} = [x_1, \dots, x_m]^\top$.

Тепер, розглянемо класифікацію квадратичних форм.

Означення 5: Класифікація квадратичних форм

Нехай маємо квадратичну форму $Q(\boldsymbol{x}) = \sum_{i,j=1}^m \alpha_{ij}x_ix_j$.

- Q є **додатно визначеною**, якщо $\forall \boldsymbol{x} \in \mathbb{R}^m : Q(\boldsymbol{x}) \geq 0$.
- Q є **строго додатно визначеною**, якщо $\forall \boldsymbol{x} \in \mathbb{R}^m \setminus \{\mathbf{0}\} : Q(\boldsymbol{x}) > 0$.
- Q є **від'ємно визначеною**, якщо $\forall \boldsymbol{x} \in \mathbb{R}^m : Q(\boldsymbol{x}) \leq 0$.
- Q є **строго від'ємно визначеною**, якщо $\forall \boldsymbol{x} \in \mathbb{R}^m : Q(\boldsymbol{x}) < 0$.
- Q є **знакозмінною**, якщо $\exists \boldsymbol{x}^-, \boldsymbol{x}^+ \in \mathbb{R}^m : \{Q(\boldsymbol{x}^-) < 0 \wedge Q(\boldsymbol{x}^+) > 0\}$.
- Q є **напівдодатно визначеною**, якщо $\forall \boldsymbol{x} \in \mathbb{R}^m : Q(\boldsymbol{x}) \geq 0$ і $\exists \boldsymbol{x}_0 \neq \mathbf{0} : Q(\boldsymbol{x}_0) = 0$.
- Q є **напіввід'ємно визначеною**, якщо $\forall \boldsymbol{x} \in \mathbb{R}^m : Q(\boldsymbol{x}) \leq 0$ і $\exists \boldsymbol{x}_0 \neq \mathbf{0} : Q(\boldsymbol{x}_0) = 0$.

Наведемо деякі приклади.

Приклад 2: Про класифікацію квадратичних форм

Розглянемо, наприклад, $Q_1(\boldsymbol{x}) = \sum_{k=1}^m x_k^2$. Ця квадратична форма є як *додатно визначеною*, так і *строго додатно визначеною*, оскільки для будь-якого $k \in \{1, \dots, m\}$ маємо $x_k^2 \geq 0$ (а для $x_k \neq 0$ маємо $x_k^2 > 0$) і тому $Q_1(\boldsymbol{x}) \geq 0$ для кожного $\boldsymbol{x} \in \mathbb{R}^m$ і $Q_1(\boldsymbol{x}) > 0$ для всіх $\boldsymbol{x} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$.

Якщо, відповідно, розглянути $Q_2(\boldsymbol{x}) := -Q_1(\boldsymbol{x})$. Ця форма є як *від'ємно визначеною*, так і *строго від'ємно визначеною*. Менш тривіальним прикладом є, наприклад, $Q_3(x_1, x_2) = -10x_1^2 + 4x_1x_2 - x_2^2$. Ця форма є *строго від'ємно визначеною*, оскільки яку б пару (x_1, x_2) ми не обирали, значення $Q_3(x_1, x_2)$ буде від'ємним (ми це доведемо в наступному прикладі).

Якщо ж розглянути, наприклад, $Q_4(\boldsymbol{x}) = \sum_{k=1}^m (x_k - a_k)^2$ де $a_k \neq 0$ для будь-якого $k \in \{1, \dots, m\}$, то ця форма вже буде *напівдодатно визначеною*, оскільки для будь-якого $\boldsymbol{x} \neq [a_1, \dots, a_m]^\top$ маємо $Q_4(\boldsymbol{x}) > 0$, проте для $\boldsymbol{x} = [a_1, \dots, a_m]^\top \neq \mathbf{0}$ в нас $Q_4(\boldsymbol{x}) = 0$.

Проте, нам потрібен якийсь зручний "інструмент", який дозволяє виявляти класифікацію квадратичної форми. Одним з цих інструментів є **критерій Сільвестера**.

Теорема 2: Критерій Сільвестера

Нехай матриця квадратичної форми Q має вигляд

$$\boldsymbol{A} = \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & \dots & \alpha_{1,m} \\ \alpha_{2,1} & \alpha_{2,2} & \dots & \alpha_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m,1} & \alpha_{m,2} & \dots & \alpha_{m,m} \end{bmatrix}$$

Якщо позначити $\Delta_k := \det \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & \dots & \alpha_{1,k} \\ \alpha_{2,1} & \alpha_{2,2} & \dots & \alpha_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{k,1} & \alpha_{k,2} & \dots & \alpha_{k,k} \end{bmatrix}$, $k \in \{1, \dots, m\}$, то

- Щоб Q була строго додатно визначеною, необхідно і достатньо, щоб $\Delta_k > 0\ \forall k \in \{1, \dots, m\}$.
- Щоб Q була строго від'ємно визначеною, необхідно і достатньо, щоб $(-1)^k \Delta_k > 0\ \forall k \in \{1, \dots, m\}$.

Доведення цієї теореми було наведено у курсі лінійної алгебри, тому не будемо його виписувати. Але ось приклад наведемо.

Приклад 3: Застосування критерію Сільвестера

З минулого прикладу, ми наводили $Q(x_1, x_2) = -10x_1^2 + 4x_1x_2 - x_2^2$ як приклад *строго від'ємно визначеної* форми. Дійсно, випишемо матрицю цієї форми:

$$\boldsymbol{A} = \begin{bmatrix} -10 & 2 \\ 2 & -1 \end{bmatrix}$$

І згідно *теоремі Сільвестера* маємо $\Delta_1 = -10 < 0, \Delta_2 = 10 - 4 = 6 > 0$, тому і отримуємо, що вона є строго від'ємно визначеною.

1.3 Достатні умови локального екстремуму функції декількох змінних.

По-перше, відповімо на питання, чому саме нам важливі квадратичні форми в контексті дослідження функції на екстремум.

Знову розглядаємо $G \subset \mathbb{R}^m, f : G \to \mathbb{R}$ і нехай $\boldsymbol{a} \in G$ є внутрішньою точкою. Оскільки нам для розглядання цікаві лише стаціонарні точки (бо в нестаціонарних точках локального екстремуму точно немає), то також будемо вважати, що $\boldsymbol{a}$ є стаціонарною.

Тепер в деякому околі $\mathcal{U}_\delta(\boldsymbol{a})$ можемо розкласти нашу функцію f у ряд Тейлора:

$$\forall \boldsymbol{x} \in \mathcal{U}_\delta(\boldsymbol{a}) : f(\boldsymbol{x}) = f(\boldsymbol{a}) + df(\boldsymbol{a}) + \frac{1}{2}d^2 f(\boldsymbol{a}) + \overline{o}(\|\boldsymbol{x} - \boldsymbol{a}\|^2), \ \boldsymbol{x} \to \boldsymbol{a} \tag{3}$$

де $df(\boldsymbol{a}) = \sum_{k=1}^m f'_{x_k}(\boldsymbol{a}) \cdot (x_k - a_k)$. Оскільки точка $\boldsymbol{a}$ є стаціонарною, то бачимо, що $df(\boldsymbol{a}) = 0$. Тому, рівняння 3 можемо переписати як (якщо позначити $\boldsymbol{\delta x} := \boldsymbol{x} - \boldsymbol{a}$):

$$f(\boldsymbol{x}) - f(\boldsymbol{a}) = \frac{1}{2}d^2 f(\boldsymbol{a}) + \overline{o}(\|\boldsymbol{\delta x}\|^2), \ \boldsymbol{\delta x} \to \boldsymbol{0} \tag{4}$$

А тепер розглянемо вираз для другого диференціалу:

$$d^2 f(\boldsymbol{a}) = \sum_{i,j=1}^m f''_{x_i x_j}(\boldsymbol{a}) \delta x_i \delta x_j$$

І можемо побачити, що це є квадратичною формою відносно $(\delta x_1, \delta x_2, \ldots, \delta x_m)$, оскільки $d^2 f(\boldsymbol{a}) = (\boldsymbol{\delta x})^\top \cdot \boldsymbol{H}_f(\boldsymbol{a}) \cdot (\boldsymbol{\delta x})$ де $\boldsymbol{H}_f(\boldsymbol{x}) = \{f''_{x_i x_j}\}_{i,j=1}^m$ це матриця Гессе.

І ось бачимо, що в нас з'являється квадратична форма. Більш того, вже зрозуміло, для чого нам потрібна класифікація квадратичних форм: знову розглядаємо рівняння 4. Якщо говорити нестрого (проте, це допоможе нам інтуїтивно зрозуміти процес), то вираз $\overline{o}(\|\boldsymbol{\delta x}\|^2)$ є дуже малою величиною і тому головним чином різниця $f(\boldsymbol{x}) - f(\boldsymbol{a})$ контролюється саме квадратичною формою $d^2 f(\boldsymbol{a})$, і якщо, наприклад, ця форма є додатно визначеною, то в нас $f(\boldsymbol{x}) > f(\boldsymbol{a})$ для будь-яких точок $\boldsymbol{x}$ з околу $\mathcal{U}_\delta(\boldsymbol{a})$, що відповідає означенню локального мінімума. А для аналізу знаковизначенності квадратичної форми $d^2 f(\boldsymbol{a})$ нам потрібно аналізувати її матрицю $H_f(\boldsymbol{a})$, для чого ми можемо застосувати *критерій Сільвестера.*

Проте, звичайно, цей факт треба доводити більш строго. Отже, сформулюємо **достатні умови локального екстремуму**.

Теорема 3: Достатні умови локального екстремуму функції декількох змінних

Нехай $G \subset \mathbb{R}^m$, $f : G \to \mathbb{R}$, маємо внутрішню точку $\boldsymbol{a} \in G$ і нехай знайдеться деякий дельта-окіл точки $\boldsymbol{a}$ в якому існують неперервні часткові похідні f до другого порядку включно, тобто

$$\exists \mathcal{U}_\delta(\boldsymbol{a}) \subset G : f(\boldsymbol{x}) \in \mathcal{C}^2(\mathcal{U}_\delta(\boldsymbol{a}))$$

Нехай точка $\boldsymbol{a}$ є стаціонарною. Тоді

1. Якщо $d^2 f(\boldsymbol{a})$ є строгододатно визначеною квадратичною формою, тоді $\boldsymbol{a}$ є точкою строгого локального мінімуму.
2. Якщо $d^2 f(\boldsymbol{a})$ є строговід'ємно визначеною квадратичною формою, тоді $\boldsymbol{a}$ є точкою строгого локального максимуму.
3. Якщо $d^2 f(\boldsymbol{a})$ є знакозмінною квадратичною формою, то в точці $\boldsymbol{a}$ екстремуму немає.

Доведення. Як ми вже наводили на початку відповіді на це питання, маємо

$$f(\boldsymbol{x}) - f(\boldsymbol{a}) = \frac{1}{2} d^2 f(\boldsymbol{a}) + \overline{o}(\|\boldsymbol{\delta x}\|^2),\ \boldsymbol{\delta x} \to \boldsymbol{0} \tag{5}$$

де $d^2 f(\boldsymbol{a}) = \sum_{i,j=1}^m f''_{x_i x_j}(\boldsymbol{a}) \delta x_i \delta x_j$ є квадратичною форма відносно $\boldsymbol{\delta x} := \boldsymbol{x} - \boldsymbol{a}$.

Тепер, розпишемо рівняння 5:

$$f(\boldsymbol{x}) - f(\boldsymbol{a}) = \frac{1}{2} \sum_{i,j=1}^m f''_{x_i x_j}(\boldsymbol{a}) \delta x_i \delta x_j + \|\boldsymbol{\delta x}\|^2 \overline{o}(1),\ \boldsymbol{\delta x} \to \boldsymbol{0} \tag{6}$$

І трошки перетворимо рівняння 6:

$$f(\boldsymbol{x}) - f(\boldsymbol{a}) = \left(\sum_{i,j=1}^m f''_{x_i x_j}(\boldsymbol{a}) \cdot \frac{\delta x_i}{\|\boldsymbol{\delta x}\|} \cdot \frac{\delta x_j}{\|\boldsymbol{\delta x}\|} + \overline{o}(1) \right) \cdot \frac{\|\boldsymbol{\delta x}\|^2}{2},\ \boldsymbol{\delta x} \to \boldsymbol{0} \tag{7}$$

Помітимо, що тепер в нас фігурують вирази виду $\delta x_i / \|\boldsymbol{\delta x}\|$, що є компонентами вектора $[\delta x_1/\|\boldsymbol{\delta x}\|, \dots, \delta x_m/\|\boldsymbol{\delta x}\|]^\top$. Позначимо $\boldsymbol{\xi} := \boldsymbol{\delta x}/\|\boldsymbol{\delta x}\|$. Особливість такого вектора це те, що він лежить на одиничній сфері $\mathcal{S}_1^m = \{\boldsymbol{x} \in \mathbb{R}^m : \|\boldsymbol{x}\| = 1\}$. Тоді, рівняння 7 можемо переписати як:

$$f(\boldsymbol{x}) - f(\boldsymbol{a}) = (Q(\boldsymbol{\xi}) + \overline{o}(1)) \frac{\|\boldsymbol{\delta x}^2\|}{2},\ \boldsymbol{\delta x} \to \boldsymbol{0} \tag{8}$$

де ми позначили $Q(\boldsymbol{\xi}) = \sum_{i,j=1}^{m} f''_{x_i x_j}(\boldsymbol{a})\xi_i\xi_j$. Помітимо, що Q є неперервною на всьому $\mathbb{R}^m$, а також $Q(\boldsymbol{\xi})$ приймає найбільше і найменше значення на $\mathcal{S}_1^m$, тобто

$$\exists \boldsymbol{\xi}_{\min}, \boldsymbol{\xi}_{\max} \in \mathcal{S}_1^m : Q(\boldsymbol{\xi}_{\min}) = \min_{\boldsymbol{\xi}\in\mathcal{S}_1^m} Q(\boldsymbol{\xi}),\ Q(\boldsymbol{\xi}_{\max}) = \max_{\boldsymbol{\xi}\in\mathcal{S}_1^m} Q(\boldsymbol{\xi}) \tag{9}$$

Перший факт випливає з неперервності функцій виду $\psi_k(\boldsymbol{\xi}) = \xi_k$, тому і функція $\sum_{i,j=1}^{m} f''_{x_i x_j}(\boldsymbol{a})\psi_i(\boldsymbol{\xi})\psi_j(\boldsymbol{\xi})$ є неперервною як суми і добутки неперервних функцій. Другий факт випливає з другої теореми Вейєрштраса: маємо неперервну функцію на компакті $\mathcal{S}_1^m \subset \mathbb{R}^m$.

Тепер починаємо розглядати випадки з нашої теореми.

Пункт 1. $d^2 f(\boldsymbol{a})$ є строгододатно визначена квадратична форма.

Знову розглядаємо вираз 8. Маємо $d^2 f(\boldsymbol{a}) = Q(\boldsymbol{\xi}) \geq Q(\boldsymbol{\xi}_{\min})$. Оскільки в нас $\|\boldsymbol{\xi}_{\min}\| = 1$, то з означення строгододатно визначеної квадратичної форми, маємо $d^2 f(\boldsymbol{a}) = Q(\boldsymbol{\xi}) \geq Q(\boldsymbol{\xi}_{\min}) =: \rho > 0$. Тому, з 8 маємо

$$f(\boldsymbol{x}) - f(\boldsymbol{a}) \geq (\rho + \overline{o}(1))\frac{\|\boldsymbol{\delta x}\|^2}{2},\ \boldsymbol{\delta x} \to \boldsymbol{0}$$

Помітимо, що

$$(\exists \delta' > 0 : 0 < \delta' < \delta)\ (\forall \boldsymbol{x} : \|\boldsymbol{\delta x}\| < \delta') : \{|\overline{o}(1)| < \rho/2\}$$

Отже, тепер будемо розглядати довільне $\boldsymbol{x} \in \mathcal{U}_{\delta'}(\boldsymbol{a})$. Тоді маємо:

$$f(\boldsymbol{x}) - f(\boldsymbol{a}) \geq \left(\rho - \frac{\rho}{2}\right)\frac{\|\boldsymbol{\delta x}\|^2}{2}$$

Отже остаточно отримуємо:

$$\forall \boldsymbol{x} \in \mathcal{U}_{\delta'}(\boldsymbol{a}) \setminus \{\boldsymbol{a}\} : f(\boldsymbol{x}) - f(\boldsymbol{a}) \geq \frac{\rho}{4}\|\boldsymbol{\delta x}\|^2 > 0$$

Пункт 2. $d^2 f(\boldsymbol{a})$ є строговід'ємно визначеною квадратичною формою. Цей пункт доводиться аналогічно першому, зміниться тільки знак нерівності.

Пункт 3. Якщо маємо знакозмінну квадратичну форму, то за означенням:

$$\exists \boldsymbol{x}^-, \boldsymbol{x}^+ \in \mathbb{R}^m : \{Q(\boldsymbol{x}^-) < 0 \land Q(\boldsymbol{x}^+) > 0\}$$

Знову розглядаємо вираз 8. Візьмемо такий $\boldsymbol{x}$, що $\boldsymbol{x}^+ = \boldsymbol{x} - \boldsymbol{a}$. Тоді $\boldsymbol{\xi}^+ = \frac{\boldsymbol{x}^+}{\|\boldsymbol{x}^+\|}$. Помітимо, що $Q(\boldsymbol{\xi}^+) = \frac{1}{\|\boldsymbol{x}^+\|^2}Q(\boldsymbol{x}^+) > 0$ (в нас $\boldsymbol{x}^+ \neq \boldsymbol{0}$).

Перепишемо рівняння 8, змінюючи $\boldsymbol{x}$ як $\boldsymbol{a}+\theta\boldsymbol{x}^+$, тоді отримаємо:

$$f(\boldsymbol{a}+\theta\boldsymbol{x}^+)-f(\boldsymbol{a})=\big(Q(\boldsymbol{\xi}^+)+\overline{o}(1)\big)\frac{\|\boldsymbol{x}^+\|^2\theta^2}{2},\ \theta\to 0^+$$

Якщо позначити $Q(\boldsymbol{\xi}^+):=\rho^+>0$, то за аналогією пункта 1, отримаємо

$$(\exists\delta^+>0:0<\delta^+<\delta)\ (\forall\theta:0<\theta<\delta^+):\{|\overline{o}(1)|<\rho^+/2\}$$

Тоді якщо розглядати довільні $0<\theta<\delta^+$, то отримаємо

$$\forall\theta\in\overset{\circ}{\mathcal{U}}_{\delta^+}(0):f(\boldsymbol{a}+\theta\boldsymbol{x}^+)-f(\boldsymbol{a})\geq\frac{\rho^+}{4}\|\boldsymbol{x}^+\|^2\theta^2>0$$

Аналогічно можна довести, позначивши $Q(\boldsymbol{x}^-):=-\rho^-$, $\rho^->0$ що

$$(\exists\delta^->0:0<\delta^-<\delta):(\forall\theta\in(0,\delta^-)):f(\boldsymbol{a}+\theta\boldsymbol{x}^-)-f(\boldsymbol{a})\leq-\frac{\rho^-}{4}\|\boldsymbol{x}^-\|^2\theta^2<0$$

Отже, якщо взяти $\widetilde{\delta}:=\min\{\delta^+,\delta^-\}$, то отримаємо, що в околі $\mathcal{U}_{\widetilde{\delta}}(\boldsymbol{a})$ ми маємо як $f(\boldsymbol{x})<f(\boldsymbol{a})$, так і $f(\boldsymbol{x})>f(\boldsymbol{a})$. Отже, екстремуму в нас дійсно немає.

2 Питання 2.

2.1 Формула Гаусса-Остроградського

Теорема 4: Формула Гаусса-Остроградського

Нехай G є областю в $\mathbb{R}^3$, $\overline{\Omega}\subset G$, де Ω є циліндричною відносно кожної з осей і $\partial\Omega$ є кусково-гладкою регулярною поверхнею. Розглядаємо векторне поле

$$\boldsymbol{F}(x,y,z)=[F_x(x,y,z),F_y(x,y,z),F_z(x,y,z)]^\top\in\mathcal{C}^1(G).$$

Тоді справедливо:

$$\oiint_{\partial\Omega}F_xdydz+F_ydxdz+F_zdxdy=\iiint_\Omega\operatorname{div}\boldsymbol{F}(x,y,z)dV$$

Доведення. Спочатку доведемо, що

$$\oiint_{\partial\Omega}F_z(x,y,z)dxdy=\iiint_\Omega\frac{\partial F_z}{\partial z}(x,y,z)dV\tag{10}$$

Використовуємо циліндричність по вісі Oz, тобто наша область Ω може бути записана як

$$\Omega = \{(x,y,z) \in \mathbb{R}^3 : (x,y) \in E - \text{область в } \mathbb{R}^2,\ E \in \mathcal{J}(\mathbb{R}^2),\ \varphi_-(x,y) \le z \le \varphi_+(x,y)\},$$

тоді

$$\iiint_\Omega \frac{\partial F_z(x,y,z)}{\partial z} dV = \iint_E dxdy \int_{\varphi_-(x,y)}^{\varphi_+(x,y)} \frac{\partial F_z(x,y,z)}{\partial z} dz \tag{11}$$

Далі помітимо, що

$$\int_{\varphi_-(x,y)}^{\varphi_+(x,y)} \frac{\partial F_z(x,y,z)}{\partial z} dz = F_z(x,y,\varphi_+(x,y)) - F_z(x,y,\varphi_-(x,y)) \tag{12}$$

Отже, підставляючи у 11, маємо:

$$\iiint_\Omega \frac{\partial F_z(x,y,z)}{\partial z} dV = \iint_E F_z(x,y,\varphi_+(x,y)) - F_z(x,y,\varphi_-(x,y)) dxdy = \\ \iint_E F_z(x,y,\varphi_+(x,y)) dxdy - \iint_E F_z(x,y,\varphi_-(x,y)) dxdy =: \mathcal{I}_+ - \mathcal{I}_-$$

Доведемо, що отримана різниця $\mathcal{I}_+ - \mathcal{I}_-$ і є $\oiint_{\partial\Omega} F_z(x,y,z)dxdy$. Дійсно, ми можемо розбити поверхню на дві межі: нижню $\partial\Omega^-$ та верхню $\partial\Omega^+$ (також, можливо, вона ще з складається з межі $\partial\Omega_\parallel$, що є частиною циліндра, твірні якого паралельні Oz, проте у цьому випадку інтеграл по цій поверхні буде нулем, оскільки векторне поле буде просто перпендикулярним до вектора нормалі). В такому разі:

$$\oiint_{\partial\Omega} F_z(x,y,z)dxdy = \iint_{\partial\Omega^-} F_z(x,y,z)dxdy + \iint_{\partial\Omega^+} F_z(x,y,z)dxdy$$

І тепер розглянемо окремо кожен з інтегралів. Наприклад, $\iint_{\partial\Omega^+} F_z(x,y,z)dxdy$. Параметризуємо цю поверхню як $\boldsymbol{r}(x,y) = [x,y,\varphi_+(x,y)]^\top$ де $(x,y) \in E$, і тоді, згідно теоремі про обрахування поверхневого інтегралу через параметризацію поверхні

$$\iint_{\partial\Omega^+} F_z(x,y,z)dxdy = \iint_E [0,0,F_z(x,y,\varphi_+(x,y))]^\top \cdot (\boldsymbol{r}'_x \times \boldsymbol{r}'_y) dxdy$$

Часткові похідні $\boldsymbol{r}'_x = [1,0,\partial\varphi_+/\partial x]^\top, \boldsymbol{r}'_y = [0,1,\partial\varphi_+/\partial y]^\top$, в такому разі векторний добуток $\boldsymbol{r}'_x \times \boldsymbol{r}'_y = [-\partial\varphi_+/\partial x, -\partial\varphi_+/\partial y, 1]$ і в такому разі наш скалярний добуток на вектор $[0,0,F_z(x,y,\varphi_+(x,y))]$ це просто $F_z(x,y,\varphi_+(x,y))$. Тому остаточно:

$$\iint_{\partial\Omega^+} F_z(x,y,z)dxdy = \iint_E F_z(x,y,\varphi_+(x,y))dxdy$$

Так само ми можемо показати, що $\iint_{\partial\Omega^-} F_z(x,y,z)dxdy = -\iint_E F_z(x,y,\varphi_-(x,y))dxdy$, але тут мінус виникає через те, що в нас параметризація в цей раз не є узгодженою. Отже, ми показали, що 10 виконується. Аналогічним чином, можна показати:

$$\oiint_{\partial\Omega} F_y(x,y,z)dxdz = \iiint_\Omega \frac{\partial F_y}{\partial y}(x,y,z)dV \tag{13}$$

$$\oiint_{\partial\Omega} F_x(x,y,z)dydz = \iiint_\Omega \frac{\partial F_x}{\partial x}(x,y,z)dV \tag{14}$$

Якщо скласти 10, 13, 14, то отримаємо:

$$\oiint_{\partial\Omega} F_x dydz + F_y dxdz + F_z dxdy = \iiint_\Omega \left(\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}\right) dV,$$

що і потрібно було довести.

2.2 Означення дивергенції та її фізичний зміст

Отже, визначимо формально, що таке дивергенція.

Означення 6: Дивергенція векторного поля

Нехай в деякій області $\mathcal{D}$ задано диференційоване векторне поле

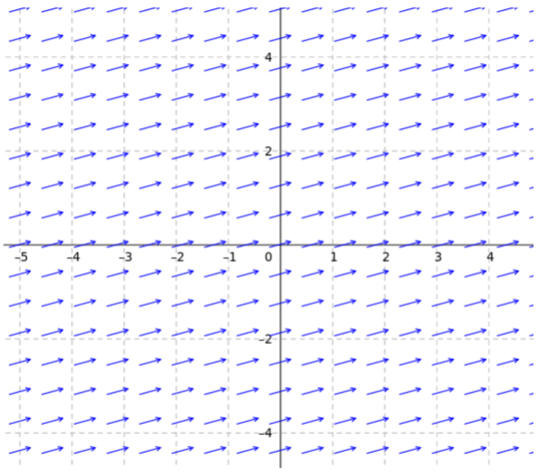
$$\boldsymbol{F}(\boldsymbol{r}) = [F_x(\boldsymbol{r}), F_y(\boldsymbol{r}), F_z(\boldsymbol{r})]^\top$$

Тоді дивергенцією $\operatorname{div}\boldsymbol{F}$ називають скалярне поле

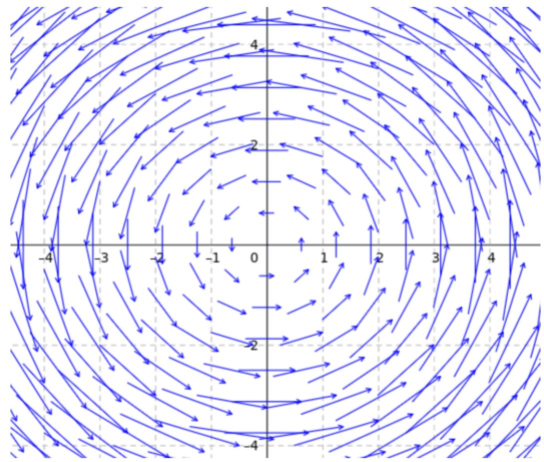
$$\operatorname{div}\boldsymbol{F} = \frac{\partial F_x(\boldsymbol{r})}{\partial x} + \frac{\partial F_y(\boldsymbol{r})}{\partial y} + \frac{\partial F_z(\boldsymbol{r})}{\partial z}$$

Фізичний зміст дивергенції наступний: нехай маємо деякий потік, наприклад, якоїсь рідини, або розглядаємо лінії електричного/магнітного полів. Дивергенція характеризує густину джерела поля, тобто характеризує поглинання або, навпаки, виділення поля в деякій точці. При цьому модуль показує інтенсивність процесу, а знак показує на поглинання (тоді знак мінус) або на виділення (тоді знак плюс).

Краще це проілюструвати малюнком в $\mathbb{R}^2$. Наприклад, нехай маємо 2 векторних поля: векторне поле, яке є константним вектором $\boldsymbol{F} \equiv \boldsymbol{\nu} \in \mathbb{R}^2$, а також векторне поле $\boldsymbol{F}(x,y) = \alpha[-y,x]^\top, \alpha \in \mathbb{R}$. Розглянемо рисунок 1.



(а) $\boldsymbol{F} \equiv \boldsymbol{\nu}$



(б) $\boldsymbol{F}(x, y) = \alpha\{-y, x\}$

Рис. 1: Векторні поля з нулевою дивергенцією

Якщо порахувати дивергенції суто формально, то маємо $\operatorname{div} \boldsymbol{\nu} = 0$, $\operatorname{div} \alpha[-y, x]^\top = \frac{\partial(-y)}{\partial x} + \frac{\partial x}{\partial y} = 0$.

Проте, з самого малюнко інтуїтивно можна помітити, що в будь-яку точку скільки втікає якогось "матеріалу", стільки з неї цього "матеріалу" і витікає. Якщо ж, навпроти, розглянути поле $\boldsymbol{F} = -\beta[x, y]^\top, \beta > 0$ (див. рис. 2), то видно, як в кожній точці "втікає" рідини більше, ніж "витікає", тому дивергенція має бути від'ємною.

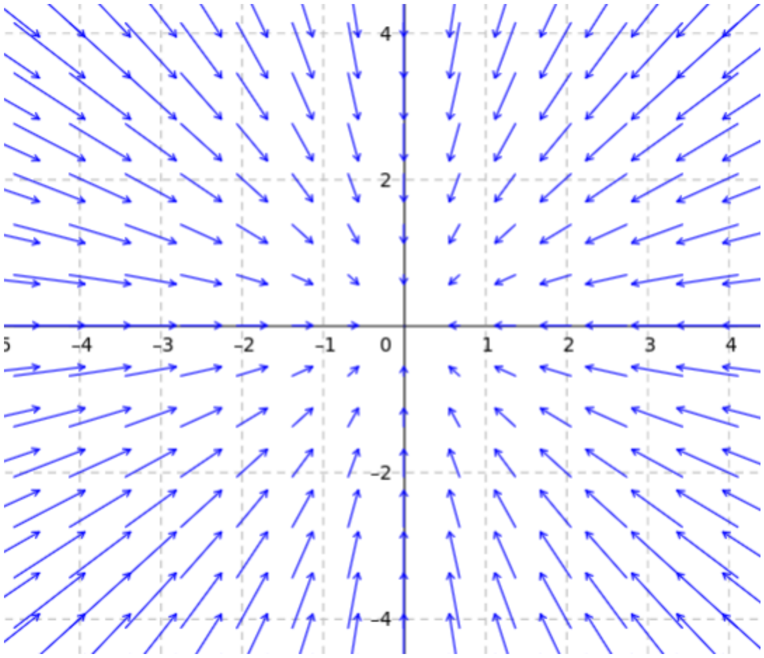


Рис. 2: Векторне поле $\boldsymbol{F} = -\beta[x, y]^\top$

Якщо це порахувати, то дійсно:

$$\operatorname{div}(-\beta[x, y]^\top) = -\beta\left(\frac{\partial x}{\partial x} + \frac{\partial y}{\partial y}\right) = -2\beta < 0$$

Дивергенція дуже часто використовується у фізиці. Наприклад, з курсу шкільної фізики можна згадати, що магнітне поле не має джерел і воно завжди замкнуте (на відміну від електричного). Математично, це описується як $\operatorname{div} \boldsymbol{B} \equiv 0$ і це є одним з рівнянь Максвелла.

Наведемо тепер строго геометричний зміст дивергенції.

Теорема 5: Геометричний зміст дивергенції

Нехай $\boldsymbol{F} : G \to \mathbb{R}^3$ є неперервно диференційованим векторним полем в області $G \subset \mathbb{R}^3$, яка обмежена кусоково-гладкою поверхнею $\partial G, \Omega \subset G$, $\partial\Omega^+$ це поверхня $\partial\Omega$, орієнтована за допомогою вибору зовнішньої нормалі $\boldsymbol{\nu}$, деяка точка $\boldsymbol{a} \in \Omega$ і нарешті $d(\Omega)$ є діаметром області Ω. Тоді

$$\operatorname{div} \boldsymbol{F}(\boldsymbol{a}) = \lim_{d(\Omega)\to 0} \frac{1}{\mu(\Omega)} \iint_{\partial\Omega^+} \langle \boldsymbol{F}, \boldsymbol{\nu} \rangle dS$$

Доведення. За теоремою Гаусса-Остроградського, маємо

$$\iint_{\partial\Omega^+} \langle \boldsymbol{F}, \boldsymbol{\nu} \rangle dS = \iiint_{\Omega} \operatorname{div} \boldsymbol{F} dV$$

Тоді, за теоремою про середнє маємо:

$$\exists \boldsymbol{A} \in \Omega : \iiint_{\Omega} \operatorname{div} \boldsymbol{F} dV = \operatorname{div} \boldsymbol{F}(\boldsymbol{A}) \cdot \mu(\Omega)$$

Звідси:

$$\operatorname{div} \boldsymbol{F}(\boldsymbol{A}) = \frac{1}{\mu(\Omega)} \iint_{\partial\Omega^+} \langle \boldsymbol{F}, \boldsymbol{\nu} \rangle dS$$

І якщо перейти до границі при $d(\Omega) \to 0$, то отримуємо

$$\operatorname{div} \boldsymbol{F}(\boldsymbol{a}) = \lim_{d(\Omega)\to 0} \frac{1}{\mu(\Omega)} \iint_{\partial\Omega^+} \langle \boldsymbol{F}, \boldsymbol{\nu} \rangle dS$$

Отже, якщо дуже нестрого, то бачимо, що потік $\iint_{\partial\Omega^+} \langle \boldsymbol{F}, \boldsymbol{\nu} \rangle dS$ приблизно дорівнює $\operatorname{div} \boldsymbol{F}(\boldsymbol{a})\Delta V$ де ΔV це дуже маленький кусок об'єму, що узгоджується з тою інтуїцією, що ми наводили на початку цього питання.

2.3 Незалежність дивергенції від вибору ортонормованої системи координат

Доведемо наступне твердження:

Теорема 6: Незалежність дивергенції від вибору ортонормованої системи координат

Нехай ми маємо векторне поле $\boldsymbol{F}(x, y, z) = [F_x(x, y, z), F_y(x, y, z), F_z(x, y, z)]^\top$ у початковому базисі (x, y, z). Якщо перейти до нового ортонормованого базису, скажімо, (u, v, w), то

$$\operatorname{div} \boldsymbol{F}(x, y, z) = \operatorname{div} \boldsymbol{F}(x(u, v, w), y(u, v, w), z(u, v, w))$$

Доведення. Цей факт одразу випливає з **геометричного змісту дивергенцій**, тобто

$$\operatorname{div} \boldsymbol{F}(\boldsymbol{a}) = \lim_{d(\Omega)\to 0} \frac{1}{\mu(\Omega)} \iint_{\partial\Omega^+} \langle \boldsymbol{F}, \boldsymbol{\nu} \rangle dS$$

В цьому виразі ані $\mu(\Omega)$ не змінюється (оскільки в нас немає ніякого стискання або розтягування простору), ані інтеграл по поверхні. В такому разі, і значення $\operatorname{div} \boldsymbol{F}(\boldsymbol{a})$ теж не змінюється.