



POLITECHNIKA WARSZAWSKA
WYDZIAŁ MATEMATYKI I NAUK INFORMACYJNYCH



OPRACOWANIE PYTAŃ NA EGZAMIN USTNY
NA KIERUNKU MATEMATYKA

LICENCJAT

AUTORZY:
WOJCIECH STOKOWIEC
BARTŁOMIEJ ZABŁOCKI

WARSZAWA, GRUDZIEŃ 2014

Spis treści

1. Kryteria zbieżności szeregów o wyrazach rzeczywistych.	7
1.1. Podstawowe informacje	7
1.1.1. Definicja szeregu liczbowego i jego sumy	7
1.1.2. Warunek konieczny zbieżności	7
1.2. Kryteria zbieżności	8
1.2.1. Kryterium porównawcze	8
1.2.2. Graniczne kryterium porównawcze	8
1.2.3. Kryterium całkowite	9
1.2.4. Kryterium d'Alemberta (kryterium ilorazowe)	9
1.2.5. Kryterium Cauchy'ego	10
1.2.6. Kryterium Leibniza	11
2. Definicja ekstremum lokalnego funkcji wielu zmiennych. Warunki konieczne i dostateczne do istnienia ekstremum lokalnego.	13
2.1. Ekstrema lokalne	13
2.2. Warunek konieczny	13
2.3. Kryterium Sylvestra	14
2.4. Warunek wystarczający	15
3. Miara Lebesgue'a, funkcje mierzalne, całka Lebesgue'a, tw. Fubiniego	19
3.1. Funkcje mierzalne	20
3.2. Całka Lebesgue'a	21
3.2.1. Całka z funkcji charakterystycznej zbioru	21
3.2.2. Całka z funkcji prostej	21
3.2.3. Całka z nieujemnej funkcji mierzalnej	21
3.2.4. Całka z funkcji mierzalnej	22
3.3. Twierdzenie Fubiniego	22
4. Całka powierzchniowa. Klasyczne twierdzenie Stokesa. Twierdzenie Greena-Gaussa-Ostrogradskiego	25
4.1. Całka krzywoliniowa	25
4.1.1. Całka krzywoliniowa nieorientowana	25
4.1.2. Całka krzywoliniowa zorientowana	27
4.1.3. Pole zachowawcze	28
4.1.4. Twierdzenie Greena	31
4.1.5. Całka Riemanna w \mathbb{R}^n (całki wielokrotne)	32
4.2. Całka powierzchniowa	33
4.3. Twierdzenie Stokesa	35
4.4. Twierdzenie Gaussa-Ostrogradskiego	36

5. Twierdzenia i wzory całkowe Cauchy’ego.	37
5.1. Podstawowe informacje	37
5.1.1. Funkcje holomorficzne	37
5.1.2. Twierdzenie podstawowe Cauchy’ego	37
5.1.3. Punkty regularne i osobliwe	38
5.2. Twierdzenia	38
5.2.1. Twierdzenie o residuach	38
5.2.2. Wzór całkowy Cauchy’ego	40
5.2.3. Inne zastosowania	40
6. Metody całkowania układu liniowych równań różniczkowych zwyczajnych 1 rzędu.	43
6.1. Układy równań różniczkowych 1 rzędu	43
6.2. Rozwiązanie układu jednorodnego o stałych współczynnikach	46
6.3. Metoda bezpośrednia	47
6.4. Metoda sprowadzania układu równań do równania rzędu wyższego	49
7. Słabe i mocne prawa wielkich liczb	51
7.1. Słabe prawa wielkich liczb	51
7.1.1. Zbieżność względem prawdopodobieństwa	51
7.1.2. Słabe prawa wielkich liczb	52
7.2. Mocne prawa wielkich liczb	53
7.2.1. Zbieżność prawie wszędzie	53
7.2.2. Mocne prawa wielkich liczb	53
7.3. Zastosowania	54
7.3.1. Metoda Monte Carlo obliczania całek	54
8. Centralne Twierdzenie Graniczne rachunku prawdopodobieństwa	55
8.1. Zbieżność względem rozkładu	55
8.1.1. Słaba zbieżność rozkładów	55
8.1.2. Zbieżność względem rozkładu	55
8.2. CTG rachunku prawdopodobieństwa	56
8.2.1. Twierdzenie de Moivre’a-Laplace’a	56
8.2.2. Twierdzenie Lindeberga-Levy’ego	56
8.2.3. Twierdzenie Lapunowa	57
8.2.4. Twierdzenie Lindeberga	57
9. Metody estymacji nieznanymi parametrów rozkładu zmiennych losowych	59
9.1. Podstawowe definicje rachunku prawdopodobieństwa	59
9.2. Model statystyczny i statystyki dostateczne	59
9.3. Estymatory	60
9.3.1. Kryteria oceny jakości estymatorów	61
9.3.2. Metody wyznaczania estymatorów	61
10. Metody numeryczne rozwiązywania układów równań liniowych	65
11. Przestrzenie liniowe, bazy, homomorfizmy przestrzeni liniowych i ich reprezentacje macierzowe	67
11.1. Przestrzenie wektorowe	67
11.2. Układy wektorów	68
11.3. Baza i wymiar przestrzeni wektorowej	69

11.4. Odwzorowanie liniowe	71
12.Podstawowe struktury algebraiczne - grupy, pierścienie, ciała, kraty.	73
12.1. Grupy	73
12.2. Pierścienie	76
12.3. Ciała	77
12.4. Kraty	77
13.Podstawowe konstrukcje algebraiczne - podalgebry, produkty, obrazy homomorficzne	79
14.Relacja równoważności	81
14.1. Relacja równoważności	81
14.2. Faktoryzacja odwzorowań	82
15.Przeliczalność i nieprzeliczalność	83
15.1. Konstrukcja von Neumanna liczb naturalnych	83
15.2. Teoria mocy	83
15.2.1. Zbiory przeliczalne	83
15.2.2. Zbiory nieprzeliczalne	85
15.2.3. Zbiory mocy continuum	85
15.3. Nierówności dla liczb kardynalnych	86
16.Liczba chromatyczna grafu, twierdzenie Brooksa, twierdzenie o 4-kolorach	87
16.1. Definicja grafu	87
16.2. Kolorowanie grafu	88

Rozdział 1

Kryteria zbieżności szeregów o wyrazach rzeczywistych.

1.1. Podstawowe informacje

1.1.1. Definicja szeregu liczbowego i jego sumy

Definicja 1.1.1. (Szereg liczbowy) Mając dany ciąg liczb $a_1, a_2, \dots, a_n, \dots$, gdzie $a_i \in \mathbb{R}$ wyrażenie $\sum_{n=1}^{\infty} a_n = a_1 + a_2 + \dots + a_n + \dots$ nazywamy *szeregiem nieskończonym*, lub krótko *szeregiem* o wyrazach $a_1, a_2, \dots, a_n, \dots$

Definicja 1.1.2. (Suma szeregu) Ciąg $s_1 = a_1, s_2 = a_1 + a_2, \dots, s_n = a_1 + \dots + a_n, \dots$ nazywamy *sumami cząstkowymi* szeregu. Jeżeli $\{s_n\}$ jest zbieżny do s , to mówimy, że szereg jest *zbieżny* i ma sumę s . Gdy $\{s_n\}$ nie jest zbieżny to szereg nazywamy *rozbieżnym*. Szereg rozbieżny nie ma sumy.

PRZYKŁAD 1.1.1. (Szereg geometryczny)

$$\sum_{n=0}^{\infty} aq^n = a + aq + aq^2 + \dots + aq^n + \dots$$

$a, q \in \mathbb{R}$, zbieżny dla $|q| < 1$, rozbieżny dla $|q| \geq 1, a \neq 0$

PRZYKŁAD 1.1.2. (Szereg harmoniczny rzędu $\alpha \in \mathbb{R}$)

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha} = 1 + \frac{1}{2^\alpha} + \frac{1}{3^\alpha} + \dots + \frac{1}{n^\alpha} + \dots$$

zbieżny dla $\alpha > 1$, rozbieżny dla $\alpha \leq 1$

1.1.2. Warunek konieczny zbieżności

Definicja 1.1.3. (Warunek konieczny)

$$\text{Szereg } \sum_{n=1}^{\infty} a_n \text{ jest zbieżny} \Rightarrow \lim_{n \rightarrow \infty} a_n = 0$$

PRZYKŁAD 1.1.3.

$$\sum_{n=1}^{\infty} \frac{1}{2^n} = \frac{1}{2} + \frac{1}{4} + \dots + \frac{1}{2^n} + \dots = 1$$

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} + \dots = \infty$$

Euler obliczył, że $s_{1000000} = 14,39$

1.2. Kryteria zbieżności**1.2.1. Kryterium porównawcze**

Definicja 1.2.1. (Kryterium porównawcze)

$$\exists m \forall (n \geq m) \quad 0 \leq a_n \leq b_n \Rightarrow$$

Jeżeli $\sum_{n=1}^{\infty} b_n$ jest zbieżny, to $\sum_{n=1}^{\infty} a_n$ jest zbieżny

Jeżeli $\sum_{n=1}^{\infty} a_n$ jest rozbieżny, to $\sum_{n=1}^{\infty} b_n$ jest rozbieżny

PRZYKŁAD 1.2.1.

$$0 \leq \sum_{n=1}^{\infty} \frac{(\sin(n))^2}{n\sqrt{n}} \leq \sum_{n=1}^{\infty} \frac{1}{n^{\frac{3}{2}}}$$

$$a_n \in \mathbb{Z}, \quad 0 \leq a_n \leq 9, \quad 0 \leq \sum_{n=1}^{\infty} \frac{a_n}{10^n} \leq \sum_{n=1}^{\infty} \frac{9}{10^n}$$

$$\alpha \leq 1, \quad 0 \leq \sum_{n=1}^{\infty} \frac{1}{n} \leq \sum_{n=1}^{\infty} \frac{1}{n^{\alpha}} \text{ zatem } \sum_{n=1}^{\infty} \frac{1}{n^{\alpha}} \text{ rozbieżny}$$

1.2.2. Graniczne kryterium porównawcze

Definicja 1.2.2. (Graniczne kryterium porównawcze)

$$\text{Jeżeli } \exists m \forall (n \geq m) \quad b_n \geq 0 \text{ oraz } \lim_{n \rightarrow \infty} \frac{a_n}{b_n} = g > 0 \Rightarrow$$

$$\sum_{n=1}^{\infty} a_n \text{ jest zbieżny} \Leftrightarrow \sum_{n=1}^{\infty} b_n \text{ jest zbieżny}$$

PRZYKŁAD 1.2.2. Graniczne kryterium porównawcze

$$\sum_{n=1}^{\infty} \frac{\sin(\frac{1}{n})}{\sqrt[3]{n}} \rightarrow a_n = \frac{\sin(\frac{1}{n})}{\sqrt[3]{n}}, \quad b_n = \frac{\frac{1}{n}}{\sqrt[3]{n}} = \frac{1}{n^{\frac{4}{3}}} \geq 0$$

$$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = \lim_{n \rightarrow \infty} \frac{\sin(\frac{1}{n})}{\frac{1}{n}} = 1 > 0$$

Ponieważ $\sum_{n=1}^{\infty} \frac{1}{n^{\frac{4}{3}}}$ jest zbieżny (bo harmoniczny rzędu $\alpha = \frac{4}{3} > 1$)

zatem też $\sum_{n=1}^{\infty} \frac{\sin(\frac{1}{n})}{\sqrt[3]{n}}$ jest zbieżny.

1.2.3. Kryterium całkowe

Definicja 1.2.3. (Kryterium całkowe) Niech $f : [n_0, \infty) \rightarrow \mathbb{R}$, $n_0 \in \mathbb{N}$, jeśli funkcja f jest:

- $\forall (x \geq n_0) \quad f(x) \geq 0$ (nieujemna dla $x \geq n_0$)
- f jest nierosnąca na $[n_0, \infty)$
- $\forall [a, b] \subset [n_0, \infty) \quad f \in R[a, b]$ (f jest całkowalna w sensie Riemanna na $[a, b]$)

to wówczas:

$$\int_{n_0}^{\infty} f(x) dx \text{ jest zbieżna} \Leftrightarrow \sum_{n=n_0}^{\infty} f(n) \text{ jest zbieżny}$$

PRZYKŁAD 1.2.3. Kryterium całkowe

$$\sum_{n=1}^{\infty} \frac{2^{-\sqrt{n}}}{\sqrt{n}} \rightarrow f(x) = \frac{2^{-\sqrt{x}}}{\sqrt{x}} > 0, \quad \forall x \geq 1$$

Ponadto funkcja jest malejąca dla $x \geq 1$, oraz ciągła w $[1, \infty)$ zatem $\forall [a, b] \subset [1, \infty) \quad f \in R[a, b]$

$$\int_1^{\infty} f(x) dx = \int_1^{\infty} \frac{1}{\sqrt{x} \cdot 2^{\sqrt{x}}} dx = \frac{1}{\ln(2)}, \text{ zatem szereg jest zbieżny}$$

PRZYKŁAD 1.2.4. Kryterium całkowe

$$\sum_{n=2}^{\infty} \frac{1}{n \sqrt[3]{\ln(n)}} \rightarrow f(x) = \frac{1}{x \sqrt[3]{\ln(x)}} > 0, \quad \forall x > 1$$

Ponadto funkcja jest malejąca dla $x > 1$, oraz ciągła w $(1, \infty)$ zatem $\forall [a, b] \subset (1, \infty) \quad f \in R[a, b]$

$$\int_2^{\infty} f(x) dx = \int_2^{\infty} \frac{1}{x} (\ln(x))^{-\frac{1}{3}} dx = \infty, \text{ zatem szereg jest rozbieżny}$$

1.2.4. Kryterium d'Alemberta (kryterium ilorazowe)

Definicja 1.2.4. (Zbieżność bezwzględna)

Mówimy, że szereg $\sum_{n=1}^{\infty} a_n$ jest *zbieżny bezwzględnie* \Leftrightarrow

$$\sum_{n=1}^{\infty} |a_n| \text{ jest zbieżny.}$$

Definicja 1.2.5. (Zbieżność warunkowa)

Szereg $\sum_{n=1}^{\infty} a_n$ jest *warunkowo zbieżny* \Leftrightarrow

jest on zbieżny i nie jest zbieżny bezwzględnie

Definicja 1.2.6. (Kryterium d'Alemberta)

Jeżeli $\exists m \forall (n \geq m) \quad a_n \neq 0$ oraz $\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = g \Rightarrow$

Jeżeli $g < 1$ to szereg $\sum_{n=1}^{\infty} a_n$ jest zbieżny bezwzględnie

Jeżeli $g > 1$ to szereg $\sum_{n=1}^{\infty} a_n$ jest rozbieżny

PRZYKŁAD 1.2.5. Kryterium d'Alemberta

$$\sum_{n=1}^{\infty} \frac{2n-1}{2^n}, \quad \left| \frac{a_{n+1}}{a_n} \right| = \left| \frac{2n+1}{2^{n+1}} \cdot \frac{2^n}{2n-1} \right| = \frac{2n+1}{2n-1} \cdot \frac{1}{2} \rightarrow \frac{1}{2} < 1$$

$$\sum_{n=1}^{\infty} \frac{2^n n!}{n^n}, \quad \left| \frac{a_{n+1}}{a_n} \right| = \left| \frac{2^{n+1} (n+1)!}{(n+1)^{n+1}} \cdot \frac{n^n}{2^n n!} \right| = 2 \cdot \left(\frac{1}{1 + \frac{1}{n}} \right)^n \rightarrow \frac{2}{e} < 1$$

1.2.5. Kryterium Cauchy'ego**Definicja 1.2.7. (Kryterium Cauchy'ego)**

Jeżeli istnieje (właściwa lub niewłaściwa) granica $\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} = g$

Jeżeli $g < 1$ to szereg $\sum_{n=1}^{\infty} a_n$ jest zbieżny bezwzględnie

Jeżeli $g > 1$ to szereg $\sum_{n=1}^{\infty} a_n$ jest rozbieżny

PRZYKŁAD 1.2.6. Kryterium Cauchy'ego

$$\sum_{n=1}^{\infty} \left(1 - \frac{1}{n}\right)^{n^2}, \quad \sqrt[n]{|a_n|} = \left(1 - \frac{1}{n}\right)^n = \left[\left(1 + \frac{-1}{n}\right)^{-n}\right]^{-1} \rightarrow \frac{1}{e} < 1$$

$$\sum_{n=1}^{\infty} \left(n^{\frac{1}{n}} - \frac{1}{1000}\right)^n, \quad \sqrt[n]{|a_n|} = \left(n^{\frac{1}{n}} - \frac{1}{1000}\right) \rightarrow \left(1 - \frac{1}{1000}\right) < 1$$

1.2.6. Kryterium Leibniza

Definicja 1.2.8. (Szereg naprzemienny) Szereg liczbowy $\sum_{n=1}^{\infty} a_n$ nazywamy *naprzemiennym* $\Leftrightarrow \forall n \quad a_n \cdot a_{n+1} < 0$ (są to szeregi postaci $\sum_{n=1}^{\infty} (-1)^n b_n$ bądź $\sum_{n=1}^{\infty} (-1)^{n-1} b_n$ gdzie $\forall n \quad b_n > 0$)

Jeżeli $\sum_{n=1}^{\infty} a_n$ jest naprzemienny, oraz:

- $\lim_{n \rightarrow \infty} |a_n| = 0$
- $\exists n_0 \quad \forall (n \geq n_0) \quad |a_{n+1}| \leq |a_n|$

to wówczas $\sum_{n=1}^{\infty} a_n$ jest zbieżny

PRZYKŁAD 1.2.7. Kryterium Leibniza

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{\sqrt[3]{n}} \rightarrow \sum_{n=1}^{\infty} |a_n| = \sum_{n=1}^{\infty} \frac{1}{\sqrt[3]{n}} \text{ nie jest zbieżny bezwzględnie}$$

$$\lim_{n \rightarrow \infty} |a_n| = 0 \text{ pierwszy warunek kryterium Leibniza spełniony}$$

$$\forall n \in \mathbb{N} \quad \frac{1}{\sqrt[3]{n+1}} \leq \frac{1}{\sqrt[3]{n}} \text{ zatem szereg jest zbieżny warunkowo}$$

Rozdział 2

Definicja ekstremum lokalnego funkcji wielu zmiennych. Warunki konieczne i dostateczne do istnienia ekstremum lokalnego.

2.1. Ekstrema lokalne

O ekstremach lokalnych można najogólniej mówić dla dowolnych funkcji, których dziedziną jest przestrzeń topologiczną, a przeciwdziedzina zbiorem co najmniej częściowo uporządkowanym. Z praktycznego punktu widzenia najczęstszym przypadkiem z jakim mamy do czynienia są funkcje postaci $f : \mathbb{R}^n \supset X \rightarrow \mathbb{R}$.

Definicja 2.1.1. Niech $f : \mathbb{R}^n \supset X \rightarrow \mathbb{R}$. Mówimy, że f ma w punkcie $\mathbf{x} = (x_1, \dots, x_n) \in X$ **maksimum lokalne** (odpowiednio **minimum lokalne**) \Leftrightarrow

1. $\mathbf{x} \in \text{Int}(X)$ gdzie $\text{Int}(x)$ oznacza wnętrze zbioru X , czyli $\exists_{r>0} B(\mathbf{x}, r) \subset X$,
2. $\exists_{\delta>0} \forall_{\mathbf{y} \in B(\mathbf{x}, \delta)} f(\mathbf{y}) \leq f(\mathbf{x})$ (dla minimum odpowiednio $\exists_{\delta>0} \forall_{\mathbf{y} \in B(\mathbf{x}, \delta)} f(\mathbf{y}) \geq f(\mathbf{x})$).

Mówimy, że f ma w \mathbf{x} **właściwe maksimum lokalne** (odpowiednio **minimum lokalne**) \Leftrightarrow

1. $\mathbf{x} \in \text{Int}(X)$, czyli $\exists_{r>0} B(\mathbf{x}, r) \subset X$,
2. $\exists_{\delta>0} \forall_{\mathbf{y} \in B(\mathbf{x}, \delta)} \mathbf{x} \neq \mathbf{y} \Rightarrow f(\mathbf{y}) < f(\mathbf{x})$ (dla minimum odpowiednio $\exists_{\delta>0} \forall_{\mathbf{y} \in B(\mathbf{x}, \delta)} f(\mathbf{y}) > f(\mathbf{x})$).

Mówimy, że f ma **ekstremum lokalne** w $\mathbf{x} \Leftrightarrow f$ ma maksimum lub minimum lokalne w \mathbf{x}

2.2. Warunek konieczny

Twierdzenie 2.2.1. (Warunek konieczny istnienia ekstremum lokalnego) Jeśli funkcja $f : \mathbb{R}^n \supset X \rightarrow \mathbb{R}$ ma w punkcie $\mathbf{x} = (x_1, \dots, x_n) \in X$ ekstremum lokalne, wówczas:

1. $\exists_{1 \leq i \leq n} \frac{\partial f}{\partial x_i}(\mathbf{x})$ nie istnieje, lub

2. $\forall_{1 \leq i \leq n} \frac{\partial f}{\partial x_i}(\mathbf{x}) = 0$ (czyli gradient $\nabla f(\mathbf{x}) = \left[\frac{\partial f}{\partial x_1}(\mathbf{x}), \dots, \frac{\partial f}{\partial x_n}(\mathbf{x}) \right] = [0, \dots, 0] = \mathbf{0}$, w skrócie $\nabla f(\mathbf{x}) = \mathbf{0}$ lub inaczej \mathbf{x} jest **punktem stacjonarnym** funkcji f).

Uwaga. Jeśli $\mathbf{x} = (x_1, \dots, x_n) \in \text{Int}(X)$ oraz $f : \mathbb{R}^n \supset X \rightarrow \mathbb{R}$ to f nie ma ekstremum lokalnego w $\mathbf{x} \Leftrightarrow$

$$\forall_{\delta > 0} \exists_{\mathbf{p}, \mathbf{q} \in B(\mathbf{x}, \delta)} f(\mathbf{p}) < f(\mathbf{x}) < f(\mathbf{q})$$

a to w przestrzeni metrycznej \mathbb{R}^n jest równoważne temu, że istnieją ciągi $(\mathbf{p}^{(k)})_{k=1}^\infty, (\mathbf{q}^{(k)})_{k=1}^\infty$ elementów zbioru X takie, że:

$$\lim_{n \rightarrow \infty} \mathbf{p}^{(k)} = \lim_{n \rightarrow \infty} \mathbf{q}^{(k)} = \mathbf{x}, \text{ oraz } \forall_k f(\mathbf{p}^{(k)}) < f(\mathbf{x}) < f(\mathbf{q}^{(k)}).$$

PRZYKŁAD 2.2.1. (Na to, że warunek konieczny nie jest wystarczający)

Weźmy funkcję $f(x_1, x_2) = x_1^2 - x_2^2$ i punkt $\mathbf{x} = (0, 0)$. Mamy następujące pochodne cząstkowe:

$$\frac{\partial f}{\partial x_1} = 2x_1, \quad \frac{\partial f}{\partial x_2} = 2x_2, \quad \mathbf{x} = (0, 0) \Rightarrow \frac{\partial f}{\partial x_1}(\mathbf{x}) = 0 = \frac{\partial f}{\partial x_2}(\mathbf{x})$$

zatem warunek konieczny jest spełniony, ale f nie ma ekstremum w punkcie $\mathbf{x} = (0, 0)$. Weźmy bowiem ciągi:

$$\mathbf{p}^{(k)} = \left(\frac{1}{k}, 0 \right) \xrightarrow{k \rightarrow \infty} (0, 0) = \mathbf{x}, \quad f(\mathbf{p}^{(k)}) = \frac{1}{k^2} > f(\mathbf{x}) = 0$$

$$\mathbf{q}^{(k)} = \left(0, \frac{1}{k} \right) \xrightarrow{k \rightarrow \infty} (0, 0) = \mathbf{x}, \quad f(\mathbf{q}^{(k)}) = -\frac{1}{k^2} < f(\mathbf{x}) = 0$$

zatem na mocy powyższej uwagi funkcja $f(x_1, x_2) = x_1^2 - x_2^2$ nie ma ekstremum w punkcie $\mathbf{x} = (0, 0)$ pomimo, iż spełniony jest warunek konieczny.

2.3. Kryterium Sylvestra

Definicja 2.3.1. Niech $A_{n \times n}$ będzie macierzą o współczynnikach rzeczywistych. Wówczas:

1. Mówimy, że A jest macierzą **dodatnio określoną** $\Leftrightarrow A$ jest macierzą symetryczną (tzn. $A = A^T$) i dla każdego niezerowego wektora $\mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0} = (0, \dots, 0)$ zachodzi $\mathbf{x}^T A \mathbf{x} > 0$, czyli po wymnożeniu $\sum_{i,j=1}^n a_{ij} x_i x_j > 0$ (równoważna definicja: A jest dodatnio określona $\Leftrightarrow A$ ma wszystkie wartości własne > 0).
2. Mówimy, że A jest macierzą **ujemnie określoną** $\Leftrightarrow A$ jest macierzą symetryczną (tzn. $A = A^T$) i dla każdego niezerowego wektora $\mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0} = (0, \dots, 0)$ zachodzi $\mathbf{x}^T A \mathbf{x} < 0$, czyli po wymnożeniu $\sum_{i,j=1}^n a_{ij} x_i x_j < 0$ (równoważna definicja: A jest ujemnie określona $\Leftrightarrow A$ ma wszystkie wartości własne < 0).
3. Mówimy, że A jest macierzą **nieokreśloną** $\Leftrightarrow A$ jest macierzą symetryczną (tzn. $A = A^T$) i istnieją takie wektory $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, że zachodzi $\mathbf{x}^T A \mathbf{x} > 0$ i $\mathbf{y}^T A \mathbf{y} < 0$, czyli po wymnożeniu $\sum_{i,j=1}^n a_{ij} x_i x_j > 0$ i $\sum_{i,j=1}^n a_{ij} y_i y_j < 0$.

Twierdzenie 2.3.1. (Kryterium Sylvestra)

Niech $A_{n \times n}$ będzie macierzą symetryczną (tzn. $A = A^T$) o współczynnikach rzeczywistych. Wówczas:

1. A jest dodatnio określona $\Leftrightarrow \forall_{1 \leq k \leq n} \det(A_k) > 0$.

2. A jest ujemnie określona $\Leftrightarrow \forall_{1 \leq k \leq n} (-1)^k \det(A_k) > 0$ (czyli gdy k jest parzyste to $\det(A_k) > 0$, zaś gdy k nieparzyste to $\det(A_k) < 0$).
3. A jest nieokreślona $\Leftrightarrow \forall_{1 \leq k \leq n} \det(A_k) \neq 0$ i jednocześnie A nie jest ani dodatnio, ani ujemnie określona.

PRZYKŁAD 2.3.1. Dla macierzy $A_{2 \times 2}$ ($n = 2$) mamy z kryterium Sylvestra:

1. A jest dodatnio określona $\Leftrightarrow a_{11} > 0$ i $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0$.
2. A jest ujemnie określona $\Leftrightarrow a_{11} < 0$ i $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0$.
3. A jest nieokreślona $\Leftrightarrow \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} < 0$.

2.4. Warunek wystarczający

Definicja 2.4.1. Niech $f : \mathbb{R}^n \supset X \rightarrow \mathbb{R}$ będzie funkcją klasy co najmniej C^2 w pewnym otoczeniu punktu $\mathbf{x}^{(0)} \in X$. Wówczas **macierzą Hessego** funkcji f w punkcie $\mathbf{x}^{(0)}$ nazywamy macierz:

$$H(\mathbf{x}^{(0)}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}^{(0)}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{x}^{(0)}) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}^{(0)}) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{x}^{(0)}) & \frac{\partial^2 f}{\partial x_2^2}(\mathbf{x}^{(0)}) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\mathbf{x}^{(0)}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}^{(0)}) & \frac{\partial^2 f}{\partial x_n \partial x_2}(\mathbf{x}^{(0)}) & \dots & \frac{\partial^2 f}{\partial x_n^2}(\mathbf{x}^{(0)}) \end{bmatrix}$$

Jest to macierz kwadratowa drugich pochodnych cząstkowych funkcji f . Wyznacznik podmacierzy stopnia $1 \leq k \leq n$ macierzy Hessego postaci:

$$H_k = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}^{(0)}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{x}^{(0)}) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_k}(\mathbf{x}^{(0)}) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{x}^{(0)}) & \frac{\partial^2 f}{\partial x_2^2}(\mathbf{x}^{(0)}) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_k}(\mathbf{x}^{(0)}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_k \partial x_1}(\mathbf{x}^{(0)}) & \frac{\partial^2 f}{\partial x_k \partial x_2}(\mathbf{x}^{(0)}) & \dots & \frac{\partial^2 f}{\partial x_k^2}(\mathbf{x}^{(0)}) \end{vmatrix}$$

nazywamy **hesjanem**.

Twierdzenie 2.4.1. (Warunek wystarczający/dostateczny istnienia ekstremum lokalnego) Niech $f : \mathbb{R}^n \supset X \rightarrow \mathbb{R}$ będzie funkcją klasy co najmniej C^2 w pewnym otoczeniu punktu $\mathbf{x}^{(0)} \in X$ i niech $\frac{\partial f}{\partial x_1}(\mathbf{x}^{(0)}) = 0, \dots, \frac{\partial f}{\partial x_n}(\mathbf{x}^{(0)}) = 0$ (czyli $\nabla f(\mathbf{x}^{(0)}) = \mathbf{0}$). Wówczas:

1. Jeżeli $\forall_{1 \leq k \leq n} H_k > 0$, to f ma właściwe minimum lokalne w punkcie $\mathbf{x}^{(0)}$.
2. Jeżeli $\forall_{1 \leq k \leq n} (-1)^k H_k > 0$, to f ma właściwe maksimum lokalne w punkcie $\mathbf{x}^{(0)}$.
3. Jeżeli $\forall_{1 \leq k \leq n} H_k \neq 0$, ale $\neg(\forall_k H_k > 0 \vee \forall_k (-1)^k H_k > 0)$, to f nie ma ekstremum lokalnego w punkcie $\mathbf{x}^{(0)}$.

Uwaga. Jeżeli $\exists_{1 \leq k \leq n} H_k = 0$, to istnienie bądź nieistnienie ekstremum lokalnego funkcji f w punkcie $\mathbf{x}^{(0)}$ należy badać z definicji.

Twierdzenie 2.4.2. (Szczególny przypadek powyższego twierdzenia - warunek wystarczający/dostateczny istnienia ekstremum lokalnego dla funkcji dwóch zmiennych rzeczywistych) Niech $f : \mathbb{R}^2 \supset X \rightarrow \mathbb{R}$ będzie funkcją klasy co najmniej C^2 w pewnym otoczeniu punktu $\mathbf{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}) \in \text{Int}(X)$ i niech $\frac{\partial f}{\partial x_1}(\mathbf{x}^{(0)}) = \frac{\partial f}{\partial x_2}(\mathbf{x}^{(0)}) = 0$ (czyli $\nabla f(\mathbf{x}^{(0)}) = \mathbf{0}$) oraz

$$H = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}^{(0)}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{x}^{(0)}) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{x}^{(0)}) & \frac{\partial^2 f}{\partial x_2^2}(\mathbf{x}^{(0)}) \end{vmatrix}$$

niech będzie hesjanem funkcji f . Wówczas:

1. Jeżeli $H > 0$, to f ma ekstremum lokalne w punkcie $\mathbf{x}^{(0)} = (x_1^{(0)}, x_2^{(0)})$, przy czym jeśli $\frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}^{(0)}) > 0$ to f ma właściwe minimum lokalne.
2. Jeżeli $H > 0$, to f ma ekstremum lokalne w punkcie $\mathbf{x}^{(0)} = (x_1^{(0)}, x_2^{(0)})$, przy czym jeśli $\frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}^{(0)}) < 0$ to f ma właściwe maksimum lokalne.
3. Jeżeli $H < 0$, to f nie ma ekstremum lokalnego w punkcie $\mathbf{x}^{(0)} = (x_1^{(0)}, x_2^{(0)})$.

Uwaga. Jeżeli $H = 0$, to istnienie bądź nieistnienie ekstremum lokalnego funkcji f w punkcie $\mathbf{x}^{(0)}$ należy badać z definicji.

PRZYKŁAD 2.4.1. (Dla funkcji dwóch zmiennych)

1. $f(x, y) = x^2 + y^2$, $f : \mathbb{R}^2 \supset X \rightarrow \mathbb{R}$ może mieć ekstremum tylko w punktach stacjonarnych, więc musimy je wyznaczyć. Mamy:

$$\begin{cases} f'_x = 2x \\ f'_y = 2y \end{cases} \wedge \begin{cases} f'_x = 0 \\ f'_y = 0 \end{cases} \Rightarrow \begin{cases} 2x = 0 \Leftrightarrow x = 0 \\ 2y = 0 \Leftrightarrow y = 0 \end{cases}$$

Zatem f ma jeden punkt stacjonarny: $(0, 0)$ podejrzany o istnienie ekstremum. Obliczmy pochodne cząstkowe: $f''_{xx} = 2$, $f''_{yy} = 2$, $f''_{xy} = 0 = f''_{yx}$. Hesjan ma zatem postać:

$$H(x, y) = \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = 4. \text{ Stąd mamy:}$$

$$\begin{cases} H(0, 0) = 4 > 0 \\ f''_{xx}(0, 0) = 2 > 0 \end{cases} \Rightarrow f \text{ ma właściwe minimum lokalne w } (0, 0).$$

2. $f(x, y) = 2x^3 + xy^2 + 5x^2 + y^2$, $f : \mathbb{R}^2 \supset X \rightarrow \mathbb{R}$ może mieć ekstremum tylko w punktach stacjonarnych, więc musimy je wyznaczyć. Mamy:

$$\begin{cases} f'_x = 6x^2 + y^2 + 10x \\ f'_y = 2xy + 2y \\ y = 0 \end{cases} \wedge \begin{cases} f'_x = 0 \\ f'_y = 0 \end{cases} \Rightarrow \begin{cases} 6x^2 + y^2 + 10x = 0 \\ 2y(x + 1) = 0 \Leftrightarrow y = 0 \vee x = -1 \end{cases}$$

$$\begin{cases} 6x^2 + 10x = 0 \Rightarrow x = 0 \vee x = -\frac{5}{3} \end{cases} \vee \begin{cases} x = -1 \\ y^2 - 4 = 0 \Rightarrow y = 2 \vee y = -2 \end{cases}$$

Zatem f ma cztery punkty stacjonarne: $P_1\left(-\frac{5}{3}, 0\right)$, $P_2(0, 0)$, $P_3(-1, 2)$, $P_4(-1, -2)$ podejrzane o istnienie ekstremum. Obliczmy pochodne cząstkowe: $f''_{xx} = 12x + 10$, $f''_{yy} = 2x + 2$, $f''_{xy} = 2y = f''_{yx}$. Hesjan ma zatem postać: $H(x, y) = \begin{vmatrix} 12x + 10 & 2y \\ 2y & 2x + 2 \end{vmatrix}$. Stąd mamy:

$$\begin{cases} H\left(-\frac{5}{3}, 0\right) = \begin{vmatrix} -20 & 0 \\ 0 & -2 \end{vmatrix} > 0 \\ f''_{xx}\left(-\frac{5}{3}, 0\right) = -20 < 0 \end{cases} \Rightarrow f \text{ ma właściwe maksimum lokalne w } \left(-\frac{5}{3}, 0\right),$$

$$\begin{cases} H(0, 0) = \begin{vmatrix} 10 & 0 \\ 0 & 2 \end{vmatrix} > 0 \\ f''_{xx}(0, 0) = 10 > 0 \end{cases} \Rightarrow f \text{ ma właściwe minimum lokalne w } (0, 0),$$

$$\begin{cases} H(-1, 2) = \begin{vmatrix} -2 & 4 \\ 4 & 0 \end{vmatrix} < 0 \\ H(-1, -2) = \begin{vmatrix} -2 & -4 \\ -4 & 0 \end{vmatrix} < 0 \end{cases} \Rightarrow f \text{ nie ma ekstremum w punktach } (-1, 2) \text{ i } (-1, -2).$$

Rozdział 3

Miara Lebesgue’a, funkcje mierzalne, całka Lebesgue’a, tw. Fubiniego

Definicja 3.0.1. Niech Ω będzie dowolnym niepustym zbiorem i niech \mathcal{F} będzie niepustą rodziną podzbiorów zbioru Ω . \mathcal{F} nazywamy σ -ciałem podzbiorów zbioru Ω , jeżeli spełnia następujące warunki:

1. $\emptyset \in \mathcal{F}$, gdzie symbolem \emptyset oznaczamy zbiór pusty,
2. jeżeli $A \in \mathcal{F}$, to $\Omega \setminus A \in \mathcal{F}$,
3. jeżeli $A_1, A_2, \dots \in \mathcal{F}$, to $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

PRZYKŁAD 3.0.1. Zbiór wszystkich podzbiorów zbioru X jest σ -algebrą. W tym przypadku stosujemy oznaczenie $\mathcal{A} = 2^X$.

PRZYKŁAD 3.0.2. $\mathcal{A} = \{\emptyset, X\}$ jest σ -algebrą, zwaną σ -algebrą trywialną.

Twierdzenie 3.0.1. Niech Ω będzie dowolnym niepustym zbiorem i niech \mathcal{A} będzie niepustą rodziną podzbiorów zbioru Ω , wtedy istnieje najmniejsze σ -ciało podzbiorów zbioru Ω zawierający \mathcal{A} . Takie σ -ciało nazywać będziemy σ -ciałem generowanym przez rodzinę \mathcal{A} .

Definicja 3.0.2. Niech (X, ρ) będzie przestrzenią metryczną, a Γ będzie rodziną zbiorów otwartych (lub domkniętych). Wtedy σ -algebrę generowaną przez rodzinę Γ , oznaczaną $\mathcal{B}(X)$, nazywamy σ -algebrą zbiorów borelowskich, a jej elementy *zbiorami borelowskimi*.

Definicja 3.0.3. Niech X będzie niepustym zbiorem, a \mathcal{A} ustaloną σ -algebrą podzbiorów zbioru X . Funkcję $\mu : \mathcal{A} \rightarrow [0, \infty]$ spełniającą następujące warunki

1. $\mu(\emptyset) = 0$,
2. dla dowolnego ciągu $(A_k)_{k \in \mathbb{N}}$ zbiorów z \mathcal{A} takiego, że $A_i \cap A_j = \emptyset$, dla $i \neq j$, zachodzi równość

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k), \quad (3.1)$$

nazywamy *miarą*.

PRZYKŁAD 3.0.3. Niech X będzie dowolnym zbiorem i $\mathcal{A} = 2^X$. Miarę μ na \mathcal{A} definiujemy następująco:

$$\mu(A) = \begin{cases} |A|, & \text{jeżeli } A \text{ jest zbiorem skończonym,} \\ \infty, & \text{w przeciwnym przypadku.} \end{cases}$$

Taką miarę nazywamy *miarą liczącą*.

Twierdzenie 3.0.2. Niech $X = \mathbf{R}^n$ i niech \mathcal{A} będzie σ -algebrą zbiorów borelowskich w X . Wtedy istnieje dokładnie jedna miara λ_n określona na \mathcal{A} taka, że dla dowolnego prostokąta

$$P = [a_1, b_1] \times \dots \times [a_n, b_n]$$

mamy

$$\lambda_n(P) = (b_1 - a_1) \cdot \dots \cdot (b_n - a_n).$$

Definicja 3.0.4. Miarę zdefiniowaną w twierdzeniu 3.0.2 nazywamy *miarą Lebesgue'a na zbiorach borelowskich*.

Definicja 3.0.5. Parę (X, \mathcal{A}) taką, że \mathcal{A} jest σ -ciałem podzbiorów zbioru X nazywać będziemy *przestrzenią mierzalną*.

Definicja 3.0.6. Trójkę (X, \mathcal{A}, μ) taką, że \mathcal{A} jest σ -ciałem podzbiorów zbioru X , a μ miarą określoną na \mathcal{A} nazywać będziemy *przestrzenią mierzalną z miarą*.

Definicja 3.0.7. Niech (X, \mathcal{A}, μ) przestrzenią mierzalną z miarą. Ciąg zbiorów mierzalnych $(A_n)_{n \in \mathbb{N}}$ nazywamy *wstępującym*, gdy $A_n \subset A_{n+1}$ dla $n \in \mathbb{N}$, *zstępującym*, gdy $A_{n+1} \subset A_n$ dla $n \in \mathbb{N}$.

Twierdzenie 3.0.3 (Ciągłość miary). Jeżeli $(A_n)_{n \in \mathbb{N}}$ jest wstępującym ciągiem zbiorów mierzalnych, to

$$\mu \left(\bigcup_{n=1}^{\infty} A_n \right) = \lim_{n \rightarrow \infty} \mu(A_n).$$

Jeżeli $(A_n)_{n \in \mathbb{N}}$ jest zstępującym ciągiem zbiorów mierzalnych i $\mu(A_1) < \infty$, to

$$\mu \left(\bigcap_{n=1}^{\infty} A_n \right) = \lim_{n \rightarrow \infty} \mu(A_n).$$

3.1. Funkcje mierzalne

Poprzez $\bar{\mathbb{R}}$ będziemy oznaczali zbiór liczb rzeczywistych uzupełniony o dwa elementy: $-\infty, +\infty$.

Definicja 3.1.1. Niech (X, \mathcal{F}) oraz (Y, \mathcal{E}) będą dwoma przestrzeniami mierzalnymi. Odwzorowanie $f : X \rightarrow Y$ nazywamy $(\mathcal{F}, \mathcal{E})$ -mierzalnym, jeżeli przeciwobraz każdego zbioru $E \in \mathcal{E}$ względem f należy do \mathcal{F} , tj.

$$f^{-1}(E) := \{x \in X \mid f(x) \in E\} \in \mathcal{F}, \quad \forall E \in \mathcal{E}.$$

Uwaga. W dalszej części wywodu mówić będziemy o funkcjach, dla których przestrzeń mierzalna (Y, \mathcal{E}) jest postaci $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, gdzie \mathbb{R} oznacza rozszerzony zbiór liczb rzeczywistych, a $\mathcal{B}(\mathbb{R})$ σ -ciało zbiorów borelowskich.

Twierdzenie 3.1.1. Jeżeli X jest przestrzenią metryczną, a \mathcal{A} jest σ -algebrą zbiorów borelowskich, to dowolna funkcja ciągła $f : X \rightarrow \mathbb{R}$ jest mierzalna.

3.2. Całka Lebesgue'a

3.2.1. Całka z funkcji charakterystycznej zbioru

Definicja 3.2.1. Funkcją charakterystyczną zbioru $A \subset X$ nazywamy funkcję $\chi_A : X \rightarrow \mathbb{R}$ określoną wzorem:

$$\chi_A(x) = \begin{cases} 1, & \text{dla } x \in A, \\ 0, & \text{dla } x \notin A. \end{cases} \quad (3.2)$$

Definicja 3.2.2. Niech χ_A będzie funkcją charakterystyczną zbioru $A \subset X$. Wtedy całkę funkcji χ_A względem miary μ definiujemy jako:

$$\int_X \chi_A(x) d\mu = \mu(A). \quad (3.3)$$

3.2.2. Całka z funkcji prostej

Definicja 3.2.3. Funkcją prostą nazywamy funkcję o skończonym zbiorze wartości.

Uwaga. Każdą funkcję prostą f można przedstawić jako kombinację liniową funkcji charakterystycznych:

$$f(x) = \sum_{i=1}^n a_i \chi_{A_i}(x), \quad \text{gdzie } A_i = \{x \in X : f(x) = a_i\}. \quad (3.4)$$

Definicja 3.2.4. Niech f_n będzie funkcją prostą, nieujemną i mierzalną określoną na zbiorze X . Wtedy całką funkcji f_n względem miary μ definiujemy jako:

$$\int_X f(x) d\mu = \sum_{i=1}^n a_i \mu(A_i). \quad (3.5)$$

3.2.3. Całka z nieujemnej funkcji mierzalnej

Definicja 3.2.5. Niech f będzie nieujemną funkcją mierzalną, a S rodziną funkcji prostych mierzalnych. Wtedy całkę funkcji f względem miary μ definiujemy jako:

$$\int_X f d\mu = \sup_{s \in S} \left\{ \int_X s d\mu : 0 \leq s \leq f \right\}. \quad (3.6)$$

Twierdzenie 3.2.1. Niech f będzie nieujemną funkcją mierzalną. Wtedy istnieje niemalejący ciąg f_n funkcji prostych, nieujemnych i mierzalnych taki, że

$$\forall_{x \in X} \lim_{n \rightarrow \infty} f_n(x) = f(x).$$

Twierdzenie 3.2.2. Niech f będzie nieujemną funkcją mierzalną, a f_n ciągiem nieujemnych mierzalnych funkcji prostych zbieżnych punktowo do f . Wtedy:

$$\int_X f d\mu = \lim_{n \rightarrow \infty} \int_X f_n d\mu. \quad (3.7)$$

3.2.4. Całka z funkcji mierzalnej

Definicja 3.2.6. Mówimy, że funkcja mierzalna $f : X \rightarrow \mathbb{R}$ jest całkowalna sensie Lebesgue'a, jeżeli

$$\int_X |f| d\mu < \infty. \quad (3.8)$$

Wtedy całkę funkcji f względem miary μ definiujemy jako:

$$\int_X f d\mu = \int_X f^+ d\mu - \int_X f^- d\mu, \quad (3.9)$$

gdzie $f = f^+ - f^-$ jest rozkładem funkcji f na $f^+ = \max(f, 0)$ oraz $f^- = -\min(f, 0)$.

Twierdzenie 3.2.3. Niech $f : [a, b] \rightarrow \mathbb{R}$ będzie ograniczoną funkcją całkowalną w sensie Riemanna. Wtedy f jest λ -mierzalna i obie całki są równe:

$$\int_a^b f(x) dx = \int_{[a, b]} f d\lambda. \quad (3.10)$$

3.3. Twierdzenie Fubiniego

Twierdzenie 3.3.1. Niech dane będą dwie przestrzenie mierzalne z miarami $(X_1, \mathcal{A}_1, \mu_1)$ oraz $(X_2, \mathcal{A}_2, \mu_2)$. Najmniejsze σ -ciało zawierające rodzinę wszystkich zbiorów postaci $A_1 \times A_2$, gdzie $A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2$ nazywamy σ -ciałem produktowym σ -ciał \mathcal{A}_1 i \mathcal{A}_2 i oznaczamy poprzez $\mathcal{A}_1 \otimes \mathcal{A}_2$.

Twierdzenie 3.3.2. Niech $(X_1, \mathcal{A}_1, \mu_1)$ oraz $(X_2, \mathcal{A}_2, \mu_2)$ będą dwoma σ -skończonymi przestrzeniami mierzalnymi z miarą. Na σ -ciele $\mathcal{A}_1 \otimes \mathcal{A}_2$ istnieje tylko jedna miara $\mu_1 \otimes \mu_2$ spełniająca dla każdego $A_1 \in \mathcal{A}_1$ oraz $A_2 \in \mathcal{A}_2$ warunek

$$\mu_1 \otimes \mu_2(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2). \quad (3.11)$$

Twierdzenie 3.3.3. Niech $(X_1, \mathcal{A}_1, \mu_1)$ oraz $(X_2, \mathcal{A}_2, \mu_2)$ będą dwoma σ -skończonymi przestrzeniami mierzalnymi z miarą. Niech funkcja $f : X_1 \times X_2 \rightarrow \mathbb{R}$ będzie $\mathcal{A}_1 \otimes \mathcal{A}_2$ mierzalna, oraz $\mu_1 \otimes \mu_2$ całkowalna, wtedy funkcje:

$$I : x_1 \mapsto \int_{X_2} f(x_1, x_2) d\mu_2(x_2), \quad J : x_2 \mapsto \int_{X_1} f(x_1, x_2) d\mu_1(x_1),$$

są odpowiednio \mathcal{A}_1 oraz \mathcal{A}_2 mierzalne, oraz:

$$\begin{aligned}\int_{X_1 \times X_2} f d\mu_1 \otimes \mu_2 &= \int_{X_1} \left(\int_{X_2} f(x_1, x_2) d\mu_2(x_2) \right) d\mu_1(x_1) \\ &= \int_{X_2} \left(\int_{X_1} f(x_1, x_2) d\mu_1(x_1) \right) d\mu_2(x_2).\end{aligned}$$

Powyższe twierdzenie uogólnia się w sposób bezpośredni na wielowymiarowe przestrzenie produktowe.

Rozdział 4

Całka powierzchniowa. Klasyczne twierdzenie Stokesa. Twierdzenie Greena-Gaussa-Ostrogradskiego

4.1. Całka krzywoliniowa

4.1.1. Całka krzywoliniowa nieorientowana

Definicja 4.1.1. (Dyfeomorfizm)

Niech X i Y będą przestrzeniami unormowanymi oraz niech U będzie niepustym, otwartym podzbiorem X . Przekształcenie $h : U \rightarrow Y$ nazywamy **dyfeomorfizmem**, gdy spełnia następujące warunki:

1. $h(U)$ jest otwartym podzbiorem Y .
2. h jest funkcją różnowartościową.
3. h i h^{-1} (h^{-1} rozumiana jako funkcja określona na $h(U)$) są klasy co najmniej C^1 .

Z powyższej definicji wynika natychmiast, że każdy dyfeomorfizm jest homeomorfizmem.

Definicja 4.1.2. **Przedstawieniem** klasy C^r krzywej w \mathbb{R}^n nazywamy dowolną funkcję $\alpha : I \rightarrow \mathbb{R}^n$ ($I \subset \mathbb{R}$ przedział niekoniecznie skończony). Obraz $\alpha(I)$ nazywamy **krzywą** w \mathbb{R}^n , zaś wektor $\alpha'(t) = (\alpha'_1(t), \dots, \alpha'_n(t))$ nazywamy **wektorem stycznym** do krzywej. Przedstawienie to nazywamy **regularnym**, jeśli $(\forall t \in I) \alpha'(t) \neq \mathbf{0} = (0, \dots, 0)$.

Definicja 4.1.3. Jeśli $\alpha : I \rightarrow \mathbb{R}^n$ jest przedstawieniem klasy C^r krzywej w \mathbb{R}^n takim, że:

1. $I = [a, b]$ gdzie $a, b \in \mathbb{R}$,
2. α jest przedstawieniem regularnym,
3. α jest iniekcją (1:1, nie posiada samoprzecięć),

to zbiór $\alpha(I) = L \subset \mathbb{R}^n$ nazywamy **łukiem** a przedstawienie α o powyższych własnościach nazywamy **parametryzacją** klasy C^r łuku L .

Twierdzenie 4.1.1. Jeśli $\alpha : [a, b] \rightarrow \mathbb{R}^n$, $\beta : [c, d] \rightarrow \mathbb{R}^n$ są dwiema parametryzacjami klasy C^r tego samego łuku L , to istnieje dyfeomorfizm $h : [a, b] \rightarrow [c, d]$ klasy C^r taki, że $\alpha(t) = \beta(h(t))$, $t \in [a, b]$.

Twierdzenie 4.1.2. Niech $\alpha : [a, b] \rightarrow \mathbb{R}^n$ będzie parametryzacją klasy C^1 łuku L . Wówczas długością łuku L nazywamy całkę:

$$d(L) = \int_a^b \|\alpha'(t)\| dt$$

Definicja 4.1.4. Niech $\alpha : [a, b] \rightarrow \mathbb{R}^n$ będzie parametryzacją klasy C^1 łuku L . Wtedy:

$$h(t) = \int_a^t \|\alpha'(\tau)\| d\tau$$

jest funkcją długości łuku od $\alpha(a)$ do $\alpha(t)$ i parametryzacja $\beta : [0, d(L)] \rightarrow \mathbb{R}^n$ zadana przez $\beta(s) = \alpha(h^{-1}(s))$ nazywa się **parametryzacją po długości łuku L** .

Definicja 4.1.5. (Całka krzywoliniowa pierwszego rodzaju)

1. Niech $\alpha : [a, b] \rightarrow \mathbb{R}^n$ będzie parametryzacją klasy C^1 łuku L oraz niech funkcja $f : L \rightarrow \mathbb{R}$ będzie funkcją ciągłą (czyli będzie funkcją klasy C^0 lub inaczej **polem skalarnym** klasy C^0). Wtedy:

$$\int_L f ds := \int_a^b f(\alpha(t)) \|\alpha'(t)\| dt$$

nazywamy **całką krzywoliniową pierwszego rodzaju** po łuku L z funkcji f (inaczej **całką krzywoliniową niezorientowaną**).

2. Jeżeli $K = L_1 \cup L_2 \cup \dots \cup L_N$ jest łańcuchem łuków (tzn. każde dwa mogą mieć co najwyżej skończoną ilość punktów wspólnych) oraz $f : K \rightarrow \mathbb{R}$ jest funkcją ciągłą, to definicja całki krzywoliniowej z f po K wyraża się wzorem

$$\int_K f ds := \sum_{k=1}^N \int_{L_k} f ds$$

Uwaga. Długość łuku jest całką krzywoliniową pierwszego rodzaju z funkcji $f \equiv 1$.

Twierdzenie 4.1.3.

$$(\forall \lambda, \mu \in \mathbb{R} \quad \forall f, g \in C^0) \quad \int_L \lambda f + \mu g ds = \lambda \int_L f ds + \mu \int_L g ds$$

$$\left| \int_L f ds \right| \leq \sup_{x \in L} |f(x)| d(L)$$

Uwaga. Całka krzywoliniowa pierwszego rodzaju (czyli całka krzywoliniowa niezorientowana) nie zależy od kierunku przebiegu parametru po zadanym łuku.

Fakt 4.1.1. Niech $L \subset \mathbb{R}^n$ będzie łukiem klasy C^1 oraz $f : L \rightarrow \mathbb{R}$ niech będzie funkcją ciągłą, wtedy $\int_L f \, ds$ nie zależy od parametryzacji tego łuku.

Uwaga. Całka krzywoliniowa pierwszego rodzaju (czyli całka krzywoliniowa nieorientowana) zależy od całego łuku, a nie tylko od jego końców, jak to ma czasem miejsce w całkach krzywoliniowych drugiego rodzaju (czyli całkach krzywoliniowych zorientowanych).

4.1.2. Całka krzywoliniowa zorientowana

Definicja 4.1.6. (Całka krzywoliniowa drugiego rodzaju)

1. Niech $\Omega \subset \mathbb{R}^n$ będzie zbiorem otwartym. **Polem wektorowym** klasy C^r na zbiorze Ω nazywamy dowolną funkcję $v : \Omega \rightarrow \mathbb{R}^n$ klasy C^r . Wykres takiego odwzorowania wyobrażamy sobie, jako zaczepiony w punkcie $\mathbf{x} = (x_1, \dots, x_n) \in \Omega$ wektor $v(\mathbf{x}) = (v_1(\mathbf{x}), \dots, v_n(\mathbf{x}))$ (gdzie odwzorowanie $v_i : \mathbb{R}^n \supset \Omega \rightarrow \mathbb{R}$ jest postaci $v_i(\mathbf{x}) = y_i$ jeśli $v(\mathbf{x}) = v((x_1, \dots, x_n)) = (y_1, \dots, y_n)$ dla $i = 1, \dots, n$).
2. Niech $\alpha : [a, b] \rightarrow \Omega \subset \mathbb{R}^n$ będzie C^1 parametryzacją łuku L oraz $v : \Omega \rightarrow \mathbb{R}^n$ niech będzie ciągłym polem wektorowym. Wtedy wyrażenie

$$\begin{aligned} \int_L v(\mathbf{x}) d\mathbf{x} &= \int_L v_1(\mathbf{x}) dx_1 + v_2(\mathbf{x}) dx_2 + \dots + v_n(\mathbf{x}) dx_n = \int_{\alpha} v(\mathbf{x}) d\mathbf{x} := \\ &:= \int_a^b \langle v(\alpha(t)), \alpha'(t) \rangle dt \end{aligned}$$

nazywamy **całką krzywoliniową drugiego rodzaju (zorientowaną)** z pola wektorowego v po łuku L .

3. Jeżeli $K = L_1 \cup L_2 \cup \dots \cup L_N$ jest łańcuchem łuków to całkę po K definiujemy jako sumę całek po kolejnych łukach

$$\int_K v(\mathbf{x}) d\mathbf{x} := \sum_{k=1}^N \int_{L_k} v(\mathbf{x}) d\mathbf{x}$$

Uwaga. 1. Całkę krzywoliniową zorientowaną oznacza się czasami inaczej np.

$$\int_L \langle v(\mathbf{x}), d\mathbf{x} \rangle = \int_L v(\mathbf{x}) \cdot d\mathbf{x}$$

2. Całkę krzywoliniową zorientowaną można wyrazić za pomocą całki nieorientowanej, mamy bowiem

$$\int_L v(\mathbf{x}) d\mathbf{x} = \int_a^b \langle v(\alpha(t)), \alpha'(t) \rangle dt = \int_a^b \underbrace{\left\langle v(\alpha(t)), \frac{\alpha'(t)}{\|\alpha'(t)\|} \right\rangle}_{\text{funkcja skalarna}} \cdot \underbrace{\|\alpha'(t)\|}_{\text{elem. dl. łuku}} dt$$

Twierdzenie 4.1.4. Niech $\alpha : [a, b] \rightarrow \mathbb{R}^n$, $\beta : [c, d] \rightarrow \mathbb{R}^n$ będą dwiema parametryzacjami klasy C^1 tego samego łuku L , oraz niech h będzie C^1 dyfeomorfizmem takim, że $(\forall t \in [a, b]) \quad \alpha(t) = \beta(h(t))$ wtedy dla dowolnego ciągłego pola wektorowego na L mamy:

$$\int_{\alpha} v(\mathbf{x}) d\mathbf{x} = \begin{cases} \int_{\beta} v(\mathbf{x}) d\mathbf{x} & \text{jeżeli } h'(t) > 0 \\ - \int_{\beta} v(\mathbf{x}) d\mathbf{x} & \text{jeżeli } h'(t) < 0 \end{cases}$$

czyli znak całki zależy od kierunku przebiegu parametru po łuku L .

Twierdzenie 4.1.5. Niech $v, w : \mathbb{R}^n \supset \Omega \rightarrow \mathbb{R}^n$ będą C^0 polami wektorowymi, oraz niech $\alpha : [a, b] \rightarrow \mathbb{R}^n$, $\beta : [c, d] \rightarrow \mathbb{R}^n$ będą C^1 parametryzacjami łuków L_1, L_2 , ponadto niech $\lambda, \mu \in \mathbb{R}$ wtedy:

1. $\int_{\alpha+\beta} v(\mathbf{x}) d\mathbf{x} = \int_{\alpha} v(\mathbf{x}) d\mathbf{x} + \int_{\beta} v(\mathbf{x}) d\mathbf{x}$ gdy $\alpha(b) = \beta(c)$
(gdzie $\alpha + \beta$ jest łukiem $L_1 \cup L_2$)
2. $\int_{\alpha} (\lambda v(\mathbf{x}) + \mu w(\mathbf{x})) d\mathbf{x} = \lambda \int_{\alpha} v(\mathbf{x}) d\mathbf{x} + \mu \int_{\alpha} w(\mathbf{x}) d\mathbf{x}$
3. $\left| \int_{\alpha} v(\mathbf{x}) d\mathbf{x} \right| \leq \left(\sup_{x \in L_1} |v(\mathbf{x})| \right) \cdot \int_a^b \|\alpha'(t)\| dt = \left(\sup_{x \in L_1} |v(\mathbf{x})| \right) \cdot d(L)$

Fakt 4.1.2. Pole obszaru którego brzeg jest krzywą zamkniętą postaci $K = L_1 \cup L_2 \cup \dots \cup L_N$ (łańcuchem łuków) wynosi (ostatni wzór to tzw. **wzór Leibniza**):

$$P = \oint_K x dy = - \oint_K y dx = \frac{1}{2} \left(\oint_K (x dy - y dx) \right)$$

Zatem np. w przypadku wzoru Leibniza liczymy całkę zorientowaną po krzywej zamkniętej K z pola wektorowego $v(x, y) = (-y, x)$ w kierunku dodatnim tzn. przeciwnie do ruchu wskazówek zegara (tak jak zakresłany jest dodatni kąt).

4.1.3. Pole zachowawcze

Definicja 4.1.7. Ciągłe pole wektorowe $v : \mathbb{R}^n \supset \Omega \rightarrow \mathbb{R}^n$ nazywamy **zachowawczym** (lub **zupełnym**) jeżeli całka zorientowana $\int_K v(\mathbf{x}) d\mathbf{x}$ po krzywej $K \subset \Omega$ będącej łańcuchem

$K = L_1 \cup L_2 \cup \dots \cup L_N$ łuków klasy C^m zależy tylko od punktów końcowych krzywej K a nie od punktów tej krzywej. Jeżeli końce krzywej K oznaczymy jako $\mathbf{x}_a, \mathbf{x}_b \in \mathbb{R}^n$ wówczas całkę zorientowaną po krzywej K oznaczamy również w następujący sposób

$$\int_{\mathbf{x}_a}^{\mathbf{x}_b} v(\mathbf{x}) d\mathbf{x}$$

Uwaga. Czasami spotyka się równoważną definicję pola zachowawczego, mianowicie: Ciągłe pole wektorowe $v : \mathbb{R}^n \supset \Omega \rightarrow \mathbb{R}^n$ nazywamy **zachowawczym** jeżeli:

$$(\forall K \subset \Omega) \quad \oint_K v(\mathbf{x}) d\mathbf{x} = 0$$

gdzie krzywa $K = L_1 \cup L_2 \cup \dots \cup L_N$ jest łańcuchem łuków klasy C^n .

Definicja 4.1.8. Pole wektorowe v nazywamy **gradientowym** jeśli istnieje funkcja $U : \mathbb{R}^n \supset \Omega \rightarrow \mathbb{R}$ klasy $C^1(\Omega)$ taka, że

$$v(\mathbf{x}) = \text{grad } U(\mathbf{x}) = \nabla U(\mathbf{x}) = \left(\frac{\partial U(\mathbf{x})}{\partial x_1}, \dots, \frac{\partial U(\mathbf{x})}{\partial x_n} \right)$$

Innymi słowy v jest gradientowe jeśli istnieje pole skalarne U (zwane **potencjałem** pola wektorowego v) takie, że gradient U równa się polu v (zauważmy tu, że gradient jest operatorem przyporządkowującym polu skalarne pole wektorowe). W związku z powyższym mówimy też, że pole wektorowe v jest gradientowe jeśli posiada potencjał.

Twierdzenie 4.1.6. Ciągłe pole wektorowe v na zbiorze otwartym $\Omega \subset \mathbb{R}^n$ jest zachowawcze \Leftrightarrow gdy v jest gradientowym polem wektorowym.

Uwaga. Powyższe tw. w klasycznych podręcznikach analizy matematycznej można znaleźć w następującej wersji:

Ciągłe pole wektorowe v na zbiorze otwartym $\Omega \subset \mathbb{R}^n$ jest zachowawcze \Leftrightarrow gdy $v_1 dx_1 + \dots + v_n dx_n$ jest różniczką zupełną pewnej funkcji U , tzn. gdy

$$v_1(\mathbf{x}) = \frac{\partial U(\mathbf{x})}{\partial x_1}, \dots, v_n(\mathbf{x}) = \frac{\partial U(\mathbf{x})}{\partial x_n}$$

Jak widać jest to definicja pola gradientowego.

Fakt 4.1.3. Pole wektorowe v klasy C^1 posiada potencjał \Rightarrow zachodzą następujące warunki (tzw. **warunki całkowalności pola** lub **warunki zgodności**):

$$\forall (i, k = 1, \dots, n) \quad \frac{\partial v_i(\mathbf{x})}{\partial x_k} = \frac{\partial v_k(\mathbf{x})}{\partial x_i}$$

Uwaga. Zauważmy, że mamy tu implikację \Rightarrow czyli samo spełnienie warunków całkowalności pola v w Ω nie wystarcza aby stwierdzić, że v jest polem gradientowym.

Definicja 4.1.9. (Zbiór gwiaździsty i jednospójny)

1. Mówimy, że zbiór otwarty $\Omega \subset \mathbb{R}^n$ jest **gwiaździsty** jeżeli

$$\exists(\mathbf{x}_0 \in \Omega) \quad \forall(\mathbf{x} \in \Omega) \quad [\mathbf{x}_0, \mathbf{x}] = \{\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0) : t \in [0, 1]\} \subset \Omega$$

Zauważmy, że wg powyższej definicji zbiór gwiaździsty jest uogólnieniem zbioru wypukłego.

2. Mówimy, że zbiór otwarty jest **jednospójny** jeżeli dla dowolnej krzywej K zamkniętej, bez samoprzecięć zawartej w Ω istnieje punkt $\mathbf{x} \in \Omega$ taki, że krzywą tą możemy w sposób ciągły przekształcić w \mathbf{x} nie opuszczając tego zbioru. Innymi słowy jest to zbiór/przestrzeń łukowo spójna o trywialnej grupie podstawowej.

Twierdzenie 4.1.7. Niech $\Omega \subset \mathbb{R}^n$ będzie zbiorem gwiaździstym, oraz niech $v : \Omega \rightarrow \mathbb{R}^n$ będzie C^1 polem wektorowym spełniającym warunki zgodności $\Rightarrow v$ posiada w Ω potencjał.

Uwaga. 1. Twierdzenie powyższe jest prawdziwe dla zbiorów jednospójnych jednakże dowód w tym przypadku jest bardziej skomplikowany.

2. Twierdzenie powyższe uogólnia się łatwo na obszar który jest C^2 - dyfeomorficzny ze zbiorem gwiaździstym (np. na rozcięty pierścień w \mathbb{R}^2).

Znajdowanie potencjału pola wektorowego

Przypadek \mathbb{R}^2

Niech $\Omega = (a, b) \times (c, d) \subset \mathbb{R}^2$ (zauważmy, że jest to zbiór gwiaździsty) i niech $v(\mathbf{x}) = v(x_1, x_2) = (v_1(x_1, x_2), v_2(x_1, x_2))$ będzie polem wektorowym klasy C^1 spełniającym warunki zgodności tzn.

$$\frac{\partial v_1(\mathbf{x})}{\partial x_2} = \frac{\partial v_2(\mathbf{x})}{\partial x_1}$$

Wówczas spełnione są założenia twierdzenia i pole v ma potencjał. Zgodnie z definicją szukamy więc funkcji $U(\mathbf{x}) = U(x_1, x_2)$ takiej, że:

$$(*) \quad \frac{\partial U(\mathbf{x})}{\partial x_1} = v_1(\mathbf{x}) = v_1(x_1, x_2) \quad \frac{\partial U(\mathbf{x})}{\partial x_2} = v_2(\mathbf{x}) = v_2(x_1, x_2)$$

Całkując np. pierwsze z równań (*) po x_1 otrzymujemy

$$U(x_1, x_2) = \int v_1(x_1, x_2) dx_1 + f(x_2)$$

Musimy znaleźć $f(x_2)$ w tym celu różniczkujemy otrzymane U po x_2 i otrzymujemy

$$\frac{\partial U(x_1, x_2)}{\partial x_2} = \frac{\partial}{\partial x_2} \left(\int v_1(x_1, x_2) dx_1 \right) + f'(x_2)$$

a następnie wstawiamy to do drugiego z równań (*) i mamy

$$v_2(x_1, x_2) - \frac{\partial}{\partial x_2} \left(\int v_1(x_1, x_2) dx_1 \right) = f'(x_2)$$

Ponieważ lewa strona ostatniego równania nie zależy od x_1 mamy bowiem

$$\frac{\partial}{\partial x_1} \left(v_2(x_1, x_2) - \frac{\partial}{\partial x_2} \left(\int v_1(x_1, x_2) dx_1 \right) \right) = \frac{\partial}{\partial x_1} (f'(x_2)) = 0$$

więc możemy wyznaczyć $f(x_2)$ całkując po x_2

$$f(x_2) = \int \left(v_2(x_1, x_2) - \frac{\partial}{\partial x_2} \left(\int v_1(x_1, x_2) dx_1 \right) \right) dx_2$$

skąd wstawiając do $U(x_1, x_2)$ otrzymujemy szukany potencjał.

Przypadek \mathbb{R}^3

Niech $\Omega \subset \mathbb{R}^3$ będzie prostopadłością (zauważmy, że jest to zbiór gwiaździsty) i niech $v(\mathbf{x}) = (v_1(\mathbf{x}), v_2(\mathbf{x}), v_3(\mathbf{x}))$ będzie polem wektorowym klasy C^1 spełniającym warunki zgodności tzn.

$$\frac{\partial v_1(\mathbf{x})}{\partial x_2} = \frac{\partial v_2(\mathbf{x})}{\partial x_1} \quad \frac{\partial v_1(\mathbf{x})}{\partial x_3} = \frac{\partial v_3(\mathbf{x})}{\partial x_1} \quad \frac{\partial v_2(\mathbf{x})}{\partial x_3} = \frac{\partial v_3(\mathbf{x})}{\partial x_2}$$

Wówczas spełnione są założenia twierdzenia i pole v ma potencjał. Zgodnie z definicją szukamy więc funkcji $U(\mathbf{x})$ takiej, że:

$$(*) \quad \frac{\partial U(\mathbf{x})}{\partial x_1} = v_1(\mathbf{x}) \quad \frac{\partial U(\mathbf{x})}{\partial x_2} = v_2(\mathbf{x}) \quad \frac{\partial U(\mathbf{x})}{\partial x_3} = v_3(\mathbf{x})$$

Całkując np. pierwsze z równań $(*)$ po x_1 otrzymujemy

$$U(\mathbf{x}) = \int v_1(\mathbf{x}) dx_1 + g(x_2, x_3)$$

Aby znaleźć $g(x_2, x_3)$ postępujemy dalej analogicznie jak w przypadku \mathbb{R}^2 powyżej (nasze $g(x_2, x_3)$ to po prostu $U(x_2, x_3)$ z przypadku \mathbb{R}^2).

4.1.4. Twierdzenie Greena

Twierdzenie 4.1.8. Niech $\Omega \subset \mathbb{R}^2$ będzie obszarem normalnym względem osi x oraz y . Niech $v(x, y) = (P(x, y), Q(x, y))$ będzie C^1 polem wektorowym takim, że pochodne tego pola są ciągle aż do brzegu obszaru $\partial\Omega$ włącznie (tzn. $v \in C^1(\overline{\Omega})$). Wtedy zachodzi tzw. **wzór Greena**:

$$\int_{\Omega} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dxdy = \oint_{\partial\Omega} P dx + Q dy$$

Uwaga. 1. Łatwo zauważyć, że wzór Greena zachodzi dla dowolnego zbioru otwartego $\Omega \subset \mathbb{R}^2$, którego domknięcie $\overline{\Omega}$ jest skończoną sumą domkniętych obszarów spełniających założenia twierdzenia.

2. Wzór Greena jest prawdziwy dla dowolnego ograniczonego zbioru otwartego Ω z brzegiem klasy C^1 (tzn. że lokalnie jest wykresem funkcji klasy C^1) oraz dowolnego pola wektorowego $v \in C^1(\overline{\Omega})$.

Wniosek 4.1.1. Niech $\Omega \subset \mathbb{R}^2$ będzie zbiorem otwartym, ograniczonym dla którego prawdziwy jest wzór Greena. Wtedy zachodzi tzw. **wzór Leibniza**:

$$|\Omega| = \frac{1}{2} \oint_{\partial\Omega} (xdy - ydx) = \oint_{\partial\Omega} xdy = - \oint_{\partial\Omega} ydx$$

wystarczy bowiem przyjąć $v(x, y) = \frac{1}{2}(-y, x)$ wtedy $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = \frac{1}{2} + \frac{1}{2} = 1$ i z wzoru Greena mamy $|\Omega| = \int_{\Omega} 1 dxdy = \frac{1}{2} \oint_{\partial\Omega} (xdy - ydx)$.

4.1.5. Całka Riemanna w \mathbb{R}^n (całki wielokrotne)

- Definicja 4.1.10.** 1. Niech $P_k \subset \mathbb{R}^n$ będzie przedziałem ograniczonym o końcach $a_k \leq b_k$ $k = 1, \dots, n$. Zbiór $I = P_1 \times \dots \times P_n$ nazywamy **n-wymiarowym prostopadłościanem**. Prostopadłościan nazywamy otwartym (domkniętym) jeśli wszystkie przedziały P_k , $k = 1, \dots, n$ są otwarte (domknięte). Wnętrzem prostopadłościanu I nazywamy prostopadłościan otwarty $\dot{I} = (a_1, b_1) \times \dots \times (a_n, b_n)$
2. Liczbę $V(I) = (b_1 - a_1) \cdot (b_2 - a_2) \cdot \dots \cdot (b_n - a_n)$ nazywamy **objętością prostopadłościanu I** (zauważmy, że $V(I) = V(\dot{I})$ czyli, że brzeg I ma zerową objętość).

Fakt 4.1.4. Jeśli mamy skończoną ilość prostopadłościanów $I_1, \dots, I_m \subset \mathbb{R}^n$ to istnieją prostopadłościany $J_1, \dots, J_p \subset \mathbb{R}^n$ o rozłącznych wnętrzach takie, że $J_1 \cup \dots \cup J_p = I_1 \cup \dots \cup I_m$.

Definicja 4.1.11. Niech A będzie dowolnym zbiorem, zaś B jego podzbiorem, $B \subseteq A$. **Funkcją charakterystyczną** zbioru B lub **indykatorem** B nazywamy funkcję $f : A \rightarrow \{0, 1\}$ określoną następująco:

$$f(x) := \begin{cases} 1 & \text{jeżeli } x \in B \\ 0 & \text{jeżeli } x \notin B \end{cases}$$

Na przykład funkcja Dirichleta jest funkcją charakterystyczną zbioru \mathbb{Q} .

Definicja 4.1.12. 1. Niech $I_1, \dots, I_m \subset \mathbb{R}^n$ będą prostopadłościanami. Funkcję $\phi : \mathbb{R}^n \ni \mathbf{x} \rightarrow \phi(\mathbf{x}) \in \mathbb{R}$ postaci:

$$\phi(\mathbf{x}) = \phi = \sum_{i=1}^m \alpha_i X_{I_i}(\mathbf{x}) \quad \text{gdzie } \alpha_i \in \mathbb{R}, \quad X_{I_i}(\mathbf{x}) = \begin{cases} 1 & \text{jeżeli } \mathbf{x} \in I_i \\ 0 & \text{jeżeli } \mathbf{x} \notin I_i \end{cases}$$

nazywamy **funkcją prostą**. Jeśli $I_1, \dots, I_m \subset I$ - prostopadłościan to mówimy, że ϕ jest funkcją prostą w I . Zbiór funkcji prostych w I oznaczamy $P(I)$.

2. Funkcje proste $\phi_1, \phi_2 \in P(I)$ nazywają się równoważnymi ($\phi_1 \sim \phi_2$) jeśli różnią się co najwyżej na brzegach definiujących je prostopadłościanów.

3. Jeśli $\phi \in P(I)$ i $\phi = \sum_{i=1}^m \alpha_i X_{I_i}(\mathbf{x})$ to:

$$\int_I \phi(\mathbf{x}) d\mathbf{x} := \sum_{i=1}^m \alpha_i V(I_i)$$

Uwaga. Każda funkcja prosta $\phi \in P(I)$ posiada przedstawienie rozłączne tzn. istnieje taki skończony układ prostopadłościanów $J_1, \dots, J_p \subset \mathbb{R}^n$, $(\forall i \neq k) J_i \cap J_k = \emptyset$, że:

$$\phi = \sum_{i=1}^p \alpha_i X_{J_i}(\mathbf{x})$$

Fakt 4.1.5. Jeśli $\phi_1, \phi_2 \in P(I)$ i $\phi_1 \sim \phi_2$ to $\int_I \phi_1(\mathbf{x}) d\mathbf{x} = \int_I \phi_2(\mathbf{x}) d\mathbf{x}$.

4.2. Całka powierzchniowa

Definicja 4.2.1. Zbiór $M \subset \mathbb{R}^3$ nazywamy **gładkim płatem** (C^r - płatem), jeśli M jest obrazem pewnej funkcji gładkiej (klasy C^r) $\phi : U \rightarrow \mathbb{R}^3$ zdefiniowanej na zbiorze otwartym $U \subset \mathbb{R}^2$ takiej, że:

1. ϕ jest funkcją różnowartościową (brak samo przecięć powierzchni).
2. Wektory $\frac{\partial \phi}{\partial u_1} = \partial_{u_1} \phi$, $\frac{\partial \phi}{\partial u_2} = \partial_{u_2} \phi$ są w każdym punkcie zbioru M liniowo niezależne (czyli zbiór M jest dwuwymiarowy).
3. ϕ^{-1} jest funkcją ciągłą (czyli wykluczamy samo przecięcia na brzegu).

Każdą funkcję ϕ spełniającą punkty 1÷3 nazywa się **parametryzacją płata** M .

PRZYKŁAD 4.2.1. (Parametryzacje)

1. Niech $f : \mathbb{R}^2 \supset U \rightarrow \mathbb{R}$ gdzie U jest zbiorem otwartym, będzie funkcją gładką (klasy C^r). Określmy $\phi : U \rightarrow \mathbb{R}^3$ wzorem $\phi(u_1, u_2) = (u_1, u_2, f(u_1, u_2)) \in \mathbb{R}^3$. Jest to tzw. **parametryzacja po wykresie**. ϕ istotnie jest parametryzacją bowiem:

- ϕ jest gładka, gdyż jej współrzędne są funkcjami gładkimi,
- ϕ jest różnowartościowa (spójrzmy na dwie pierwsze współrzędne - będą zawsze różne dla różnego argumentu),
- wektory styczne do M w punkcie $\phi(u_1, u_2)$ są postaci:

$$\phi_{u_1} = \left(1, 0, f_{u_1} = \frac{\partial f}{\partial u_1}\right) \quad \phi_{u_2} = \left(0, 1, f_{u_2} = \frac{\partial f}{\partial u_2}\right)$$

widać, że są one zawsze liniowo niezależne,

- ϕ^{-1} jest ciągła, bo mamy tu rzutowanie wykresu funkcji na płaszczyznę a z topologii wiadomo, że rzutowania są odwzorowaniami ciągłymi.

Twierdzenie 4.2.1. Niech $\phi : U \rightarrow \mathbb{R}^3$ oraz $\tau : V \rightarrow \mathbb{R}^3$ będą dwiema parametryzacjami C^r tego samego płata M . Wtedy istnieje C^r dyfeomorfizm $h : U \rightarrow V$ taki, że $\phi = \tau \circ h$ (h jest transformacją parametrów).

Definicja 4.2.2. Jeżeli $U \subset \mathbb{R}^2$ jest zbiorem otwartym i ograniczonym oraz $\phi : U \rightarrow \mathbb{R}^3$ jest parametryzacją płata M , to liczbę:

$$A(M) = A(\phi(U)) = \int_U \|\partial_{u_1} \phi \times \partial_{u_2} \phi\| du_1 du_2$$

nazywamy **pojemnością powierzchni** płata M .

Uwaga. Zauważmy, że (jesteśmy w przestrzeni Hilberta):

$$\|a \times b\|^2 = \|a\|^2 \cdot \|b\|^2 \cdot \sin^2 \angle(a, b) = \|a\|^2 \cdot \|b\|^2 \cdot (1 - \cos^2 \angle(a, b)) = \|a\|^2 \cdot \|b\|^2 - \langle a, b \rangle^2$$

gdzie $\langle a, b \rangle$ jest iloczynem skalarnym punktów a i b . Ostatnie wyrażenie, jest równe wyznacznikowi z **macierzy Grama** tzn.:

$$\|a\|^2 \cdot \|b\|^2 - \langle a, b \rangle^2 = \det \begin{bmatrix} \langle a, a \rangle & \langle a, b \rangle \\ \langle b, a \rangle & \langle b, b \rangle \end{bmatrix} = \begin{vmatrix} \langle a, a \rangle & \langle a, b \rangle \\ \langle b, a \rangle & \langle b, b \rangle \end{vmatrix}$$

zatem możemy całkę pola powierzchni płata równoważnie przedstawić jako:

$$A(\phi(U)) = \int_U \sqrt{g(\mathbf{u})} du_1 du_2, \quad g(\mathbf{u}) = \det[\langle \partial_{u_i} \phi, \partial_{u_j} \phi \rangle]_{i,j=1,2} = \begin{vmatrix} \langle \partial_{u_1} \phi, \partial_{u_1} \phi \rangle & \langle \partial_{u_1} \phi, \partial_{u_2} \phi \rangle \\ \langle \partial_{u_2} \phi, \partial_{u_1} \phi \rangle & \langle \partial_{u_2} \phi, \partial_{u_2} \phi \rangle \end{vmatrix}$$

Definicja 4.2.3. (Całka powierzchniowa pierwszego rodzaju (niezorientowana))

Niech M będzie gładkim płatem i niech $\phi : U \rightarrow \mathbb{R}^3$ będzie parametryzacją tego płata. Ponadto niech $f : M \rightarrow \mathbb{R}$ będzie funkcją ciągłą. Wówczas:

$$\int_M f(\mathbf{x}) dS(\mathbf{x}) := \int_U f(\phi(\mathbf{u})) \sqrt{g(\mathbf{u})} d\mathbf{u}$$

nazywa się **całką powierzchniową pierwszego rodzaju** z funkcji f po płacie M (**całką niezorientowaną**).

PRZYKŁAD 4.2.2. Całki powierzchniowe pierwszego rodzaju opisują wielkości fizyczne typu:

- masa o zadanej funkcją f gęstości na powierzchni zadanego płata M ,
- ładunek elektryczny o zadanej funkcją f gęstości rozłożony na powierzchni zadanego płata M ,
- środek ciężkości płata i geometryczny środek ciężkości płata.

Twierdzenie 4.2.2. Całka powierzchniowa pierwszego rodzaju nie zależy od parametryzacji płata M .

Uwaga. Jeśli $\phi : U \rightarrow \mathbb{R}^3$ jest gładką parametryzacją płata M to w dowolnym punkcie $\phi(\mathbf{u}) \in M$ możemy zdefiniować wektor normalny $n(\phi(\mathbf{u})) = \frac{\partial_{u_1} \phi \times \partial_{u_2} \phi}{\|\partial_{u_1} \phi \times \partial_{u_2} \phi\|}$ do płata M . Jeżeli $\psi : V \rightarrow \mathbb{R}^3$ jest inną parametryzacją M , wówczas $\partial_{u_1} \phi \times \partial_{u_2} \phi = \det(\nabla h) \cdot \partial_{v_1} \psi \times \partial_{v_2} \psi$ (gdzie $\nabla h = Dh = \begin{bmatrix} \nabla h_1 \\ \nabla h_2 \end{bmatrix}$ jest macierzą Jacobiego) co oznacza, że $n(\phi(\mathbf{u})) = \pm n(\psi(h(\mathbf{u})))$ w zależności od tego czy $\det(\nabla h) > 0$, czy $\det(\nabla h) < 0$. Z tego, że h jest dyfeomorfizmem wynika, że $\det(\nabla h)$ ma stały znak w całym obszarze M . Stąd mamy następującą definicję:

Definicja 4.2.4. (Orientacja)

1. Dwie parametryzacje płata M nazywają się **jednakowo (zgodnie) zorientowane**, jeżeli dla transformacji parametrów h zachodzi $\det(\nabla h) > 0$ w U . W przeciwnym przypadku ($\det(\nabla h) < 0$) ϕ i ψ nazywają się **przeciwnie zorientowane**. Wszystkie parametryzacje M rozbijają się więc na dwie rozłączne klasy (bo nie może zachodzić $\det(\nabla h) = 0$, gdyż wtedy nie istniała by funkcja odwrotna).
2. Niech n będzie ciągłym polem wektorowym kierunku normalnego na płacie M . Wówczas płat M nazywa się **dwustronnym**, jeżeli $\forall p \in M$ i dla każdej krzywej zamkniętej na M ,

nie przecinającej brzegu M zachodzi: biegnąc po krzywej z punktu p wektorem normalnym $n(p)$ wracamy do p z tym samym kierunku normalnym pola n .

Uwaga. **Płat zorientowany** posiada pole wektorowe kierunków normalnych:

$$n(\phi(\mathbf{u})) = \pm \frac{\partial_{u_1}\phi \times \partial_{u_2}\phi}{\|\partial_{u_1}\phi \times \partial_{u_2}\phi\|}$$

gdzie wybieramy „+” jeśli ϕ jest zorientowana zgodnie z ustaloną parametryzacją (dodatnio zorientowana) oraz „−” jeśli ϕ jest zorientowana przeciwnie do ustalonej parametryzacji.

Definicja 4.2.5. (Całka powierzchniowa drugiego rodzaju (zorientowana))

Niech M będzie dwustronnym płatem zorientowanym, oraz niech $n : M \rightarrow \mathbb{R}^3$ będzie dodatnim polem kierunków normalnych. Ponadto niech $v : M \rightarrow \mathbb{R}^3$ będzie ciągłym polem wektorowym. Wtedy wyrażenie:

$$\int_M v \cdot dS := \int_M \langle v(\mathbf{x}), n(\mathbf{x}) \rangle dS = \pm \int_U \langle v \circ \phi, \partial_{u_1}\phi \times \partial_{u_2}\phi \rangle d\mathbf{u}$$

(gdzie znak wybieramy zgodnie z orientacją ϕ w stosunku do U) nazywamy **całką powierzchniową drugiego rodzaju (zorientowaną)** pola v po płacie M .

Uwaga. (Inne oznaczenie)

1. Całkę powierzchniową drugiego rodzaju nazywamy też **strumieniem pola wektorowego** v przez powierzchnię M .
2. Zauważmy, że:

$$\begin{aligned} \partial_{u_1}\phi \times \partial_{u_2}\phi &= \begin{pmatrix} \left| \begin{matrix} \partial_{u_1}\phi_2 & \partial_{u_2}\phi_2 \\ \partial_{u_1}\phi_3 & \partial_{u_2}\phi_3 \end{matrix} \right|, & \left| \begin{matrix} \partial_{u_1}\phi_3 & \partial_{u_2}\phi_3 \\ \partial_{u_1}\phi_1 & \partial_{u_2}\phi_1 \end{matrix} \right|, & \left| \begin{matrix} \partial_{u_1}\phi_1 & \partial_{u_2}\phi_1 \\ \partial_{u_1}\phi_2 & \partial_{u_2}\phi_2 \end{matrix} \right| \end{pmatrix} = \\ &= \begin{pmatrix} \frac{\partial(y,z)}{\partial u_1 \partial u_2}, & \frac{\partial(z,x)}{\partial u_1 \partial u_2}, & \frac{\partial(x,y)}{\partial u_1 \partial u_2} \end{pmatrix} \end{aligned}$$

Stąd wynika alternatywny zapis całki powierzchniowej zorientowanej:

$$\int_M v_1 dydz + v_2 dzdx + v_3 dx dy$$

3. Bezpośrednio z definicji wynika, że $\int_M V \cdot dS$ zależy od parametryzacji płata M . Gdy dwie parametryzacje są przeciwnie zorientowane to otrzymujemy przeciwne wyniki.

PRZYKŁAD 4.2.3. 1. Całka powierzchniowa zorientowana przedstawia (np. w prawie Gaussa dla elektryczności, magnetyzmu) strumień Φ natężenia pola elektrycznego (magnetycznego) danego przez pole wektorowe \vec{E} , przenikającego przez powierzchnię S tzn.: $\int_S \vec{E} \cdot d\vec{S}$

4.3. Twierdzenie Stokesa

Definicja 4.3.1. Niech $v : \Omega \rightarrow \mathbb{R}^n$ będzie polem wektorowym klasy co najmniej C^1 , gdzie $\Omega \subset \mathbb{R}^n$ jest zbiorem otwartym. Wówczas:

1. **Dywergencją** pola wektorowego v nazywamy wyrażenie skalarne:

$$\operatorname{div} v(x_1, \dots, x_n) = \frac{\partial v_1}{\partial x_1} + \dots + \frac{\partial v_n}{\partial x_n} = \operatorname{trace}(Dv)$$

czyli jest to ślad macierzy Jacobiego pola wektorowego v .

2. **Rotacją** pola wektorowego v nazywamy wyrażenie wektorowe:

$$\operatorname{rot} v(x_1, \dots, x_n) = \left(\frac{\partial v_i}{\partial x_j} - \frac{\partial v_j}{\partial x_i} \right)_{i \neq j, i, j=1, \dots, n}$$

Na przykład dla $n = 3$ mamy wzór na rotację:

$$\operatorname{rot} v(x_1, x_2, x_3) = \left(\frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3}, \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1}, \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \right)$$

Pole v nazywamy **bezwirowym**, jeśli $\operatorname{rot} v = 0$.

Twierdzenie 4.3.1. (Wzór całkowy Stokesa)

Niech $M \subset \mathbb{R}^3$ będzie zorientowanym płatem z kawałkami gładkim brzegiem ∂M . Niech $\phi : \bar{U} \rightarrow M$ (\bar{U} to domknięcie U) będzie parametryzacją tego płata taką, że $\phi(\partial U) = \partial M$. Ponadto zakładamy, że dodatni kierunek obiegu ∂U odpowiada dodatniemu kierunkowi ∂M związanemu z orientacją M . Niech v będzie C^1 polem wektorowym zdefiniowanym w otoczeniu M . Wówczas ma miejsce następujący **wzór całkowy Stokesa**:

$$\int_M \operatorname{rot} v \cdot dS = \int_{\partial M} v \cdot dx$$

czyli strumień pola rotacji przez płat M jest równy cyrkulacji pola v po brzegu tego płata (po lewej mamy całkę powierzchniową zorientowaną a po lewej całkę krzywoliniową zorientowaną).

Wniosek 4.3.1. (Wzór Greena)

Wzór Greena (łączy całkę podwójną z całką krzywoliniową), czyli:

$$\int_{\Omega} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = \int_{\partial \Omega} P dx + Q dy$$

jest szczególnym przypadkiem wzoru Stokesa.

4.4. Twierdzenie Gaussa-Ostrogradskiego

Twierdzenie 4.4.1. (Wzór całkowy Gaussa - Ostrogradskiego)

Niech $V \subset \mathbb{R}^3$ będzie obszarem ograniczonym w \mathbb{R}^3 normalnym w każdym kierunku osi współrzędnych. Niech ∂V będzie kawałkami gładkim płatem zorientowanym na zewnątrz V . Wtedy dla dowolnego pola wektorowego $v : \bar{V} \rightarrow \mathbb{R}^3$ klasy C^1 zachodzi **wzór całkowy Gaussa - Ostrogradskiego**:

$$\int_V \operatorname{div} v \, d^3(x, y, z) = \int_{\partial V} v \cdot dS = \int_{\partial V} \langle v, n \rangle \, dS = \int_{\partial V} v_1 \, dydz + v_2 \, dzdx + v_3 \, dxdy$$

Uwaga. Wzór Gaussa - Ostrogradskiego zachodzi dla dowolnego obszaru $V \subset \mathbb{R}^n$ ograniczonego z kawałkami gładkim brzegiem. Jest on wnioskiem z ogólnego twierdzenia Stokesa.

Rozdział 5

Twierdzenia i wzory całkowe Cauchy'ego.

5.1. Podstawowe informacje

5.1.1. Funkcje holomorficzne

Definicja 5.1.1. Funkcja holomorficzna Mówimy, że funkcja $f : \mathbb{C} \rightarrow \mathbb{C}$ jest *holomorficzna* w punkcie $z \in \mathbb{C}$, jeżeli jest ona określona w pewnym otoczeniu tego punktu i ma pochodną w każdym punkcie tego otoczenia. Mówimy, że funkcja f jest *holomorficzna w zbiorze* $D \subseteq \mathbb{C}$ jeżeli jest holomorficzna w każdym punkcie tego zbioru.

Definicja 5.1.2. Równania Cauchy'ego - Riemanna Funkcja $f : D \rightarrow \mathbb{C}$ której rozkład na część rzeczywistą i urojoną jest postaci $f(x + iy) = u(x, y) + iv(x, y)$ jest holomorficzna w $D \Leftrightarrow$ gdy u i v są różniczkowalne i spełniają równania różniczkowe Cauchy'ego - Riemanna:

$$\begin{cases} u_x = v_y \\ -u_y = v_x \end{cases}$$

Wtedy $f'(z) = u_x + iv_x = v_y - iu_y = u_x - iu_y = v_y + iv_x$

PRZYKŁAD 5.1.1. (Funkcja kwadratowa)

$$f(z) = z^2 = (x^2 - y^2) + 2xyi \rightarrow u_x = 2x = v_y, -u_y = 2y = v_x$$

PRZYKŁAD 5.1.2. (Inwersja)

$$f(z) = \frac{1}{z} = \frac{x}{x^2 + y^2} - \frac{iy}{x^2 + y^2} \rightarrow u_x = \frac{y^2 - x^2}{(x^2 + y^2)^2} = v_y \quad -u_y = \frac{2xy}{(x^2 + y^2)^2} = v_x$$

PRZYKŁAD 5.1.3. (Funkcja wykładnicza)

$$f(z) = e^z = e^x \cos(y) + ie^x \sin(y) \rightarrow u_x = e^x \cos(y) = v_y, -u_y = e^x \sin(y) = v_x$$

5.1.2. Twierdzenie podstawowe Cauchy'ego

Definicja 5.1.3. Twierdzenie całkowe Cauchy'ego Niech $G \subseteq \mathbb{C}$ będzie obszarem jedno-spójnym, $f : G \rightarrow \mathbb{C}$ będzie holomorficzna, oraz niech γ będzie kawałkami gładką krzywą

zamkniętą leżącą w obszarze G . Wówczas:

$$\oint_{\gamma} f(z) dz = 0$$

Uwaga. Twierdzenia tego używa się w dowodach wielu twierdzeń z analizy zespolonej m.in. twierdzenia o residuach czy wzoru całkowego Cauchy'ego

5.1.3. Punkty regularne i osobliwe

Definicja 5.1.4. (Punkt regularny)

Mówimy, że z_0 jest **punktem regularnym** funkcji $f(z)$ jeśli jest ona holomorficzną w tym punkcie.

Definicja 5.1.5. (Punkt osobliwy)

Punkt z_0 nazywamy **punktem osobliwym, odosobnionym (izolowanym)** jeśli istnieje sąsiedztwo $0 < |z - z_0| < R$ tego punktu w którym funkcja $f(z)$ jest holomorficzną.

PRZYKŁAD 5.1.4. $f(z) = \frac{1}{1-z}$ punkt $z = 1$ jest punktem osobliwym odosobnionym, pozostałe punkty \mathbb{C} są regularne

5.2. Twierdzenia

5.2.1. Twierdzenie o residuach

Definicja 5.2.1. (Definicja residuum)

Niech f holomorficzną w $G_{z_0} = \{z : 0 < |z - z_0| < \varepsilon\}$ i γ dowolny dodatnio skierowany okrąg $|z - z_0| = r < \varepsilon$. Wówczas

$$a_{-1} = \frac{1}{2\pi i} \oint_{\gamma} f(z) dz$$

gdzie a_{-1} jest współczynnikiem szeregu Laurenta postaci

$$f(z) = \sum_{n=-\infty}^{\infty} a_n z^n$$

Liczbę a_{-1} nazywamy **residuum** funkcji $f(z)$ w punkcie z_0 i oznaczmy przez $Res_{z_0} f(z)$.

Twierdzenie 5.2.1. Jeżeli $f : G \rightarrow \mathbb{C}$ jest holomorficzną w obszarze $G \subseteq \mathbb{C}$ z wyjątkiem co najwyżej skończonej liczby punktów $z_1, z_2, \dots, z_n \in G$ zaś γ jest krzywą zamkniętą, kawałkami gładką, dodatnio zorientowaną, leżącą w tym obszarze oraz zawierającą wskazane punkty w swoim wnętrzu to:

$$\oint_{\gamma} f(z) dz = 2\pi i \sum_{k=1}^n \left(\nu_{\gamma}(z_k) \cdot Res_{z_k} f \right)$$

gdzie $\nu_{\gamma}(z_k)$ to indeks punktu z_0 względem krzywej γ (intuicyjnie jest to ilość okrążeń krzywej γ dookoła punktu z_0).

Twierdzenie 5.2.2. Jak liczyć residua z funkcji holomorficzych w $G \setminus \{z_0, \dots, z_n\}$? Niech f będzie holomorficzną w G oraz $z_0 \in G$ wówczas:

$$\operatorname{Res}_{z_0} \left(\frac{f(z)}{z - z_0} \right) = f(z_0)$$

PRZYKŁAD 5.2.1. Obliczyć całkę $\int_C \frac{dz}{z^2(z+2i)}$ gdzie C jest okręgiem $|z + 2i| = 1$ dodatnio zorientowanym

$$\oint_{|z+2i|=1} \frac{dz}{z^2(z+2i)} = 2\pi i \cdot \nu_\gamma(-2i) \cdot \operatorname{Res}_{-2i} \left(\frac{1}{z^2(z+2i)} \right)$$

$$\oint_{|z+2i|=1} \frac{dz}{z^2(z+2i)} = 2\pi i \cdot 1 \cdot \frac{1}{(-2i)^2} = -\frac{\pi i}{2}$$

PRZYKŁAD 5.2.2.

$$\oint_{|z|=3} \frac{e^z}{z^2+2z} dz = \frac{1}{2} \oint_{|z|=3} \frac{e^z}{z} dz - \frac{1}{2} \oint_{|z|=3} \frac{e^z}{z+2} dz =$$

$$= 2\pi i \cdot \left[\frac{1}{2} \operatorname{Res}_0 \left(\frac{e^z}{z} \right) - \frac{1}{2} \operatorname{Res}_{-2} \left(\frac{e^z}{z+2} \right) \right] = 2\pi i \cdot \left(\frac{1}{2} - \frac{e^{-2}}{2} \right)$$

zatem ostatecznie mamy:

$$\oint_{|z|=3} \frac{e^z}{z^2+2z} dz = \pi i \cdot (1 - e^{-2})$$

Definicja 5.2.2. (Klasyfikacja punktów osobliwych)

Niech dane będzie rozwinięcie funkcji $f(z)$ w szereg Laurenta w sąsiedztwie punktu osobliwego, odizolowanego z_0 :

$$f(z) = \sum_{n=-\infty}^{\infty} a_n z^n = \sum_{n=0}^{\infty} a_n (z - z_0)^n + \sum_{n=1}^{\infty} \frac{a_{-n}}{(z - z_0)^n}$$

- Jeśli część osobliwa rozwinięcia "znika" to punkt z_0 nazywamy **pozornie osobliwym**.
- Jeśli część osobliwa zawiera skończoną liczbę składników to istnieje takie $k \in \mathbb{N}$, że $a_{-k} \neq 0$ oraz $a_{-n} = 0$ dla $n > k$ wtedy z_0 nazywamy **biegunem k-krotnym**
- Jeśli część osobliwa zawiera nieskończenie wiele składników to punkt z_0 nazywamy **istotnie osobliwym**.

Twierdzenie 5.2.3. (Residuum w biegunie k-krotnym)

Jeśli z_0 jest biegunem k-krotnym funkcji $f(z)$ to residuum w punkcie z_0 funkcji $f(z)$ wynosi:

$$\operatorname{Res}_{z_0} f(z) = \frac{1}{(k-1)!} \lim_{z \rightarrow z_0} \left[\frac{d^{k-1}}{dz^{k-1}} \left((z - z_0)^k f(z) \right) \right]$$

Twierdzenie 5.2.4. (Odpowiedniość między biegunami a zerami funkcji)

- Jeśli z_0 jest biegunem k -krotnym funkcji $f(z)$ to dla funkcji:

$$\begin{cases} \frac{1}{f(z)}, & z \neq z_0 \\ 0, & z = z_0 \end{cases}$$

jest on k -krotnym zerem.

- Jeśli z_0 jest k -krotnym zerem funkcji $f(z)$ to jest on biegunem k -krotnym funkcji $\frac{1}{f(z)}$ (bardzo pomocne przy liczeniu residuów w biegunach)

5.2.2. Wzór całkowy Cauchy'ego

Twierdzenie 5.2.5. Niech $f : G \rightarrow \mathbb{C}$ będzie holomorficzną w obszarze $G \subseteq \mathbb{C}$ oraz niech $\{z : |z - z_0| \leq r\} \subset G$. Wtedy dla każdego z takiego, że $|z - z_0| < r$ mamy

$$f(z) = \frac{1}{2\pi i} \oint_{|w-z_0|=r} \frac{f(w)}{w-z} dw$$

zatem funkcja holomorficzna zdefiniowana na dysku jest całkowicie zdeterminowana przez wartości, które przyjmuje na brzegu tego dysku.

PRZYKŁAD 5.2.3.

$$\begin{aligned} \oint_{|z|=2} \frac{z^2}{z^2 + 2z + 2} dz &= \oint_{C_1} \frac{\frac{z^2}{z-z_2}}{z-z_1} dz + \oint_{C_2} \frac{\frac{z^2}{z-z_1}}{z-z_2} dz = \\ &= 2\pi i \left(\frac{z_1^2}{z_1 - z_2} + \frac{z_2^2}{z_2 - z_1} \right) = 2\pi i(-2) = -4\pi i \end{aligned}$$

gdzie $z_1 = -1 + i$, $z_2 = -1 - i$ są punktami osobliwymi funkcji $\frac{z^2}{z^2+2z+2}$

5.2.3. Inne zastosowania

Twierdzenie 5.2.6. Niech funkcja $f(z) = \frac{p(z)}{q(z)}$ będzie zespoloną funkcją wymierną gdzie $p(z)$ i $q(z)$ są wielomianami o współczynnikach rzeczywistych, ponadto niech $\forall x \in \mathbb{R}$, $q(x) \neq 0$ oraz $\deg q \geq 2 + \deg p$ wtedy:

$$\int_{-\infty}^{\infty} f(x) dx = 2\pi i \cdot \left[\sum_{\operatorname{Im}(z_k) > 0} \left(\operatorname{Res}_{z_k} f \right) \right]$$

PRZYKŁAD 5.2.4.

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^2} = 2\pi i \cdot \left[\sum_{\operatorname{Im}(z_k) > 0} \left(\operatorname{Res}_{z_k} \left(\frac{1}{1+z^2} \right) \right) \right]$$

Stąd miejsca zerowe $1 + z^2$ (a zatem bieguny $\frac{1}{1+z^2}$) wynoszą $z_1 = i$, $z_2 = -i$ zaś residuum wynosi

$$\operatorname{Res}_i \left(\frac{1}{1+z^2} \right) = \operatorname{Res}_i \left(\frac{1}{(z+i)(z-i)} \right) = \frac{1}{i+i} = \frac{1}{2i}$$

więc w "prosty" sposób otrzymujemy znany wzór:

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^2} = 2\pi i \cdot \frac{1}{2i} = \pi$$

Rozdział 6

Metody całkowania układu liniowych równań różniczkowych zwyczajnych 1 rzędu.

6.1. Układy równań różniczkowych 1 rzędu

Definicja 6.1.1. Układem normalnym n równań różniczkowych 1 rzędu o n funkcjach niewiadomych x_1, \dots, x_n nazywamy układ postaci:

$$(\text{URN}) \left\{ \begin{array}{l} \frac{dx_1}{dt} = f_1(t, x_1, \dots, x_n) \\ \frac{dx_2}{dt} = f_2(t, x_1, \dots, x_n) \\ \vdots \\ \frac{dx_n}{dt} = f_n(t, x_1, \dots, x_n) \end{array} \right.$$

gdzie: $t \in (a, b) = I$, t - zmienna niezależna
 x_1, x_2, \dots, x_n - zmienne zależne
 $D = (a, b) \times D_1 \times \dots \times D_n \subset \mathbb{R}^{n+1}$, $D_i \subset \mathbb{R}$
 $f_i : \mathbb{R}^{n+1} \supset D \rightarrow \mathbb{R}$ - dane

Warunki początkowe:

$$(\text{WP}) \left\{ \begin{array}{l} x_1(t_0) = \hat{x}_1 \\ x_2(t_0) = \hat{x}_2 \\ \vdots \\ x_n(t_0) = \hat{x}_n \end{array} \right.$$

gdzie: $t_0 \in (a, b) = I$, $\hat{x}_1, \dots, \hat{x}_n$ - dane stałe. W zapisie macierzowym oznaczając:

$$x = x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix}, \quad f(t, x) = f(t, x_1, x_2, \dots, x_n) = \begin{bmatrix} f_1(t, x_1, x_2, \dots, x_n) \\ f_2(t, x_1, x_2, \dots, x_n) \\ \vdots \\ f_n(t, x_1, x_2, \dots, x_n) \end{bmatrix}$$

$$\dot{x} = x'(t) = \frac{dx}{dt} = \begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \vdots \\ \frac{dx_n}{dt} \end{bmatrix}, \quad \hat{x} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \vdots \\ \hat{x}_n \end{bmatrix}$$

zagadnienie Cauchy'ego układu równań (URN) przyjmuje postać:

$$\begin{cases} \dot{x} = f(t, x) \\ x(t_0) = \hat{x} \end{cases}$$

Definicja 6.1.2. Rozwiązaniem (URN) na przedziale I nazywamy funkcję wektorową $x = x(t)$, różniczkowalną na tym przedziale i spełniającą następujące warunki:

1. $\forall t \in I \ (t, x_1(t), x_2(t), \dots, x_n(t)) \in D$,
2. $\forall t \in I \ \dot{x}(t) = f(t, x(t))$

Definicja 6.1.3. Zagadnieniem Cauchy'ego dla układu (URN) nazywamy wyznaczenie takiego rozwiązania $x = x(t)$ tego układu, które spełnia warunek początkowy (WP).

Definicja 6.1.4. (Ogólna definicja warunku Lipschitza)

Niech (X, ϱ) , (Y, σ) będą przestrzeniami metrycznymi. Mówimy, że $f : X \rightarrow Y$ spełnia **warunek Lipschitza** \Leftrightarrow

$$\exists L > 0 \ \forall x_1, x_2 \in X \quad \sigma(f(x_1), f(x_2)) \leq L \cdot \varrho(x_1, x_2)$$

Najmniejszą wartość L (o ile istnieje) dla której nierówność powyższa jest prawdziwa nazywamy **stałą Lipschitza**.

Definicja 6.1.5. (Definicja warunku Lipschitza dla $f(x, t)$)

Mówimy, że funkcja wektorowa $f(t, x) = f(t, x_1, x_2, \dots, x_n)$ spełnia **warunek Lipschitza** względem zmiennych x_1, x_2, \dots, x_n na zbiorze $D \subset \mathbb{R}^{n+1}$, jeżeli istnieje taka stała dodatnia L , zwana **stałą Lipschitza**, że dla każdych dwóch punktów $(t, x_1^{(1)}, x_2^{(1)}, \dots, x_n^{(1)})$, $(t, x_1^{(2)}, x_2^{(2)}, \dots, x_n^{(2)}) \in D$ spełnione są nierówności:

$$\left| f_i(t, x_1^{(1)}, x_2^{(1)}, \dots, x_n^{(1)}) - f_i(t, x_1^{(2)}, x_2^{(2)}, \dots, x_n^{(2)}) \right| \leq L \cdot \sum_{k=1}^n |x_k^{(2)} - x_k^{(1)}| \quad \text{dla } i = 1, 2, \dots, n$$

Twierdzenie 6.1.1. (Picarda - Lindelofa)

Jeżeli funkcje $f_i = f_i(t, x_1, x_2, \dots, x_n)$, $i = 1, 2, \dots, n$ spełniają warunki:

- 1) są ciągle na obszarze $D \subset \mathbb{R}^{n+1}$ (f_i są funkcjami $n + 1$ zmiennych),
- 2) spełniają warunek Lipschitza względem zmiennych x_1, x_2, \dots, x_n na obszarze D ,

to wówczas dla każdego punktu $(t_0, \hat{x}_1, \hat{x}_2, \dots, \hat{x}_n) = (t_0, \hat{x}) \in D$ istnieje $h > 0$ takie, że na przedziale $[t_0 - h, t_0 + h]$ układ (URN) ma dokładnie jedno rozwiązanie klasy C^1 spełniające warunek początkowy (WP). Jest to rozwiązanie lokalne.

Definicja 6.1.6. (Układ równań różniczkowych liniowych)

Układem n równań różniczkowych **liniowych** 1 rzędu nazywamy układ postaci:

$$(\text{URL}) \begin{cases} \frac{dx_1}{dt} = a_{11}(t)x_1 + a_{12}(t)x_2 + \dots + a_{1n}(t)x_n + f_1(t) \\ \frac{dx_2}{dt} = a_{21}(t)x_1 + a_{22}(t)x_2 + \dots + a_{2n}(t)x_n + f_2(t) \\ \vdots \\ \frac{dx_n}{dt} = a_{n1}(t)x_1 + a_{n2}(t)x_2 + \dots + a_{nn}(t)x_n + f_n(t) \end{cases}$$

gdzie a_{ij} , $i, j = 1, 2, \dots, n$ oraz f_i , $i = 1, 2, \dots, n$ są danymi funkcjami, ciągłymi na pewnym przedziale $I \subset \mathbb{R}$. Oznaczając (i biorąc pod uwagę poprzednie oznaczenia):

$$A(t) = \begin{bmatrix} a_{11}(t) & a_{12}(t) & \dots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \dots & a_{2n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}(t) & a_{n2}(t) & \dots & a_{nn}(t) \end{bmatrix}, \quad f(t) = \begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix}$$

możemy zagadnienie Cauchy'ego dla układu (URL) zapisać w postaci macierzowej:

$$\begin{cases} \dot{x} = A(t) \cdot x + f(t) \\ x(t_0) = \hat{x} \end{cases}$$

Jeżeli $f_i(t) \equiv 0$ na I dla $i = 1, 2, \dots, n$ to układ (URL) nazywamy **układem jednorodnym**, w przeciwnym przypadku układ nazywamy **niejednorodnym**.

Twierdzenie 6.1.2. Jeżeli funkcje a_{ij} , $i, j = 1, 2, \dots, n$ oraz f_i , $i = 1, 2, \dots, n$ są ciągłe na przedziale I , to przez każdy punkt $(t_0, \hat{x}_1, \hat{x}_2, \dots, \hat{x}_n) \in I \times \mathbb{R}^n$ przechodzi **jedynie wysycone** rozwiązanie (URL) określone na całym I (**rozwiązanie wysycone** to rozwiązanie szczególne, określone na pewnym przedziale, które nie daje się rozszerzyć do rozwiązania na żadnym większym przedziale).

Definicja 6.1.7. Niech będzie danych n funkcji wektorowych $\bar{x}_1(t), \bar{x}_2(t), \dots, \bar{x}_n(t)$, które są rozwiązaniami układu jednorodnego $\dot{x} = A(t) \cdot x$. **Wrońskianem** nazywamy wyznacznik macierzy $W(t)$:

$$|W(t)| = \det \begin{bmatrix} x_1^{(1)}(t) & x_1^{(2)}(t) & \dots & x_1^{(n)}(t) \\ x_2^{(1)}(t) & x_2^{(2)}(t) & \dots & x_2^{(n)}(t) \\ \vdots & \vdots & \ddots & \vdots \\ x_n^{(1)}(t) & x_n^{(2)}(t) & \dots & x_n^{(n)}(t) \end{bmatrix} = \begin{vmatrix} x_1^{(1)}(t) & x_1^{(2)}(t) & \dots & x_1^{(n)}(t) \\ x_2^{(1)}(t) & x_2^{(2)}(t) & \dots & x_2^{(n)}(t) \\ \vdots & \vdots & \ddots & \vdots \\ x_n^{(1)}(t) & x_n^{(2)}(t) & \dots & x_n^{(n)}(t) \end{vmatrix}$$

Twierdzenie 6.1.3. Jeżeli funkcje wektorowe $\bar{x}_1(t), \bar{x}_2(t), \dots, \bar{x}_n(t)$ są całkami szczególnymi układu jednorodnego $\frac{d\bar{x}}{dt} = f(t)\bar{x}$, $t \in I$ i $\det W(t) \neq 0$ dla $t \in I$ to funkcja wektorowa $\bar{x}(t) = c_1\bar{x}_1(t) + c_2\bar{x}_2(t) + \dots + c_n\bar{x}_n(t)$, $c_i \in \mathbb{R}$, $i = 1, 2, \dots, n$ jest dowolnym rozwiązaniem równania jednorodnego.

Twierdzenie 6.1.4. Przestrzeń rozwiązań równania jednorodnego jest skończenie wymiarowa o wymiarze n .

Twierdzenie 6.1.5. Jeżeli $W(t)$ jest **macierzą fundamentalną** równania jednorodnego to całka ogólna tego układu na przedziale I ma postać:

$$\bar{x}_J(t) = W(t) \cdot c, \quad c = [c_1 \dots c_n]^T$$

$W(t_0) = E$, gdzie E - macierz jednostkowa (jest to własność macierzy fundamentalnej)

Twierdzenie 6.1.6. Całka ogólna $x(t) = \bar{x}(t)$ (wektor) układu liniowego niejednorodnego $\dot{x} = A(t) \cdot x + f(t)$ jest sumą dowolnej całki szczególnej $x_S(t)$ tego układu oraz całki ogólnej $x_J(t)$ odpowiadającego mu układu liniowego jednorodnego tzn.

$$x(t) = x_J(t) + x_S(t) \quad (\bar{x}(t) = \bar{x}_J(t) + \bar{x}_S(t))$$

Twierdzenie 6.1.7. Całka ogólna układu liniowego niejednorodnego ma postać:

$$x(t) = \underbrace{W(t) \cdot c}_{x_J(t)} + \underbrace{W(t) \int_{t_0}^t W^{-1}(s) f(s) ds}_{x_S(t)}$$

gdzie $f(s)$ - wektor. Zagadnienie Cauchy'ego:

$$\begin{aligned} x(t) &= W(t) \cdot \underbrace{W^{-1}(t_0) \cdot \hat{x}}_c + W(t) \int_{t_0}^t W^{-1}(s) f(s) ds \\ \hat{x} = x(t_0) &= W(t_0) \cdot c \quad \Rightarrow \quad c = W^{-1}(t_0) \cdot \hat{x} \\ x(t) &= W(t - t_0) \cdot \hat{x} + \int_{t_0}^t W(t - s) f(s) ds \end{aligned}$$

6.2. Rozwiązanie układu jednorodnego o stałych współczynnikach

Definicja 6.2.1. (Układ równań różniczkowych liniowych o stałych współczynnikach)
Układem n równań różniczkowych **liniowych o stałych współczynnikach** 1 rzędu nazywamy układ postaci:

$$\begin{cases} \frac{dx_1}{dt} = a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n + f_1(t) \\ \frac{dx_2}{dt} = a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n + f_2(t) \\ \vdots \\ \frac{dx_n}{dt} = a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n + f_n(t) \end{cases}$$

gdzie a_{ij} , $i, j = 1, 2, \dots, n$ są danymi liczbami, zaś f_i , $i = 1, 2, \dots, n$ są danymi funkcjami, ciągłymi na pewnym przedziale $I \subset \mathbb{R}$. Oznaczając (i biorąc pod uwagę poprzednie oznaczenia):

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}, \quad f(t) = \begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix}$$

możemy zagadnienie Cauchy'ego dla tego układu zapisać w postaci macierzowej:

$$\begin{cases} \dot{x} = A \cdot x + f(t) \\ x(t_0) = \hat{x} \end{cases}$$

Jeżeli $f_i(t) \equiv 0$ na I dla $i = 1, 2, \dots, n$ to układ nazywamy **układem jednorodnym**, w przeciwnym przypadku układ nazywamy **niejednorodnym**.

Definicja 6.2.2. Normą macierzy $A_{n \times n}$ nazywamy liczbę:

$$\|A\|_2 = \left(\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2 \right)^{\frac{1}{2}}$$

Normę tę (są też inne) nazywamy **normą Frobeniusa**.

6.3. Metoda bezpośrednia

PRZYKŁAD 6.3.1. Rozwiążemy zagadnienie Cauchy'ego dla układu równań:

$$(\text{URL}) \begin{cases} \frac{dx}{dt} = x + y + t \\ \frac{dy}{dt} = 4x + y + t^2 \end{cases} \quad (\text{WP}) \begin{cases} x(0) = 1 \\ y(0) = 1 \end{cases} \quad \hat{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad f(t) = \begin{bmatrix} t \\ t^2 \end{bmatrix}$$

Jest to układ równań różniczkowych 1 rzędu, liniowych o stałych współczynnikach, niejednorodny. Najpierw rozwiązujemy układ jednorodny. W tym celu znajdujemy pierwiastki równania charakterystycznego macierzy (czyli znajdujemy jej wartości własne λ z równania postaci $(A - \lambda E) \cdot x = \mathbf{0}$):

$$\begin{cases} \dot{x} = x + y \\ \dot{y} = 4x + y \end{cases} \Rightarrow A = \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix} \Rightarrow (\text{RCH}) \begin{vmatrix} 1 - \lambda & 1 \\ 4 & 1 - \lambda \end{vmatrix} = 0 \Rightarrow (1 - \lambda)^2 = 4 \Rightarrow \begin{cases} \lambda_1 = -1 \\ \lambda_2 = 3 \end{cases}$$

Obie wartości własne są rzeczywiste o krotności 1. Wyznaczamy podprzestrzeń X_1 odpowiadającą wartości własnej $\lambda_1 = -1$ (krotność 1):

$$X_1 = \{x^{(1)} \in \mathbb{R}^2 : (A + 1 \cdot E) \cdot x^{(1)} = \mathbf{0}\} \Rightarrow A + 1 \cdot E = \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix}$$

$$x^{(1)} = \begin{bmatrix} x \\ y \end{bmatrix} \Rightarrow (A + 1 \cdot E) \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{cases} 2x + y = 0 \\ 4x + 2y = 0 \end{cases}$$

Jedno z równań jest zależne, więc mamy $x = C_1$, $y = -2C_1$ i zbiór X_1 ma postać:

$$X_1 = \left\{ C_1 \cdot \begin{bmatrix} 1 \\ -2 \end{bmatrix} : C_1 \in \mathbb{R} \right\}, \quad \begin{bmatrix} 1 \\ -2 \end{bmatrix} e^{-t}$$

Teraz wyznaczamy podprzestrzeń X_2 odpowiadającą wartości własnej $\lambda_2 = 3$ (krotność 1):

$$X_2 = \{x^{(2)} \in \mathbb{R}^2 : (A - 3 \cdot E) \cdot x^{(1)} = \mathbf{0}\} \Rightarrow A - 3 \cdot E = \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix} - \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 4 & -2 \end{bmatrix}$$

$$x^{(2)} = \begin{bmatrix} x \\ y \end{bmatrix} \Rightarrow (A - 3 \cdot E) \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} -2 & 1 \\ 4 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{cases} -2x + y = 0 \\ 4x - 2y = 0 \end{cases}$$

Jedno z równań jest zależne, bowiem: $\begin{vmatrix} -2 & 1 \\ 4 & -2 \end{vmatrix} = 0$, więc mamy $x = C_2$, $y = 2C_2$ i zbiór X_2 ma postać:

$$X_2 = \left\{ C_2 \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} : C_2 \in \mathbb{R} \right\}, \quad \begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{3t}$$

Zatem mamy macierz fundamentalną (sprawdzenie $\det \neq 0$):

$$W(t) = \begin{bmatrix} 1 \cdot e^{-t} & 1 \cdot e^{3t} \\ -2 \cdot e^{-t} & 2 \cdot e^{3t} \end{bmatrix} \Rightarrow \begin{vmatrix} e^{-t} & e^{3t} \\ -2e^{-t} & 2e^{3t} \end{vmatrix} = 4e^{2t} \neq 0$$

Znajdziemy całkę ogólną (URL). W tym celu musimy wyznaczyć macierz odwrotną do macierzy $W(0)$:

$$W(0) = \begin{bmatrix} 1 & 1 \\ -2 & 2 \end{bmatrix} \Rightarrow \begin{vmatrix} 1 & 1 \\ -2 & 2 \end{vmatrix} = 4 \neq 0$$

Możemy posłużyć się gotowym wzorem otrzymanym w następujący sposób:

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \Rightarrow M^D = \begin{bmatrix} (-1)^{1+1}d & (-1)^{1+2}c \\ (-1)^{2+1}b & (-1)^{2+2}a \end{bmatrix} = \begin{bmatrix} d & -c \\ -b & a \end{bmatrix} \Rightarrow (M^D)^T = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$|M| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \Rightarrow M^{-1} = \frac{1}{|M|} \cdot (M^D)^T = \frac{1}{ad - bc} \cdot \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Zatem w naszym przypadku mamy:

$$W(0) = \begin{bmatrix} 1 & 1 \\ -2 & 2 \end{bmatrix} \Rightarrow W^{-1}(0) = \frac{1}{4} \cdot \begin{bmatrix} 2 & -1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} \end{bmatrix}$$

$$W_0(t) = W(t) \cdot W^{-1}(0) = \begin{bmatrix} e^{-t} & e^{3t} \\ -2e^{-t} & 2e^{3t} \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2} & -\frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} \end{bmatrix} = \begin{bmatrix} \frac{1}{2}e^{-t} + \frac{1}{2}e^{3t} & -\frac{1}{4}e^{-t} + \frac{1}{4}e^{3t} \\ -e^{-t} + e^{3t} & \frac{1}{2}e^{-t} + \frac{1}{2}e^{3t} \end{bmatrix}$$

Całka ogólna (URL) ma postać:

$$x(t) = \underbrace{W_0(t) \cdot \hat{x}}_{x_J(t)} + \underbrace{\int_{t_0=0}^t W_0(t-s)f(s)ds}_{x_S(t)}$$

Biorąc pod uwagę, że:

$$W_0(t-s) = \begin{bmatrix} \frac{1}{2}e^{-(t-s)} + \frac{1}{2}e^{3(t-s)} & -\frac{1}{4}e^{-(t-s)} + \frac{1}{4}e^{3(t-s)} \\ -e^{-(t-s)} + e^{3(t-s)} & \frac{1}{2}e^{-(t-s)} + \frac{1}{2}e^{3(t-s)} \end{bmatrix} \quad \hat{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad f(s) = \begin{bmatrix} s \\ s^2 \end{bmatrix}$$

otrzymujemy ostatecznie $x(t) = x_J(t) + x_S(t)$ gdzie:

$$x_J(t) = \begin{bmatrix} \frac{1}{2}e^{-t} + \frac{1}{2}e^{3t} & -\frac{1}{4}e^{-t} + \frac{1}{4}e^{3t} \\ -e^{-t} + e^{3t} & \frac{1}{2}e^{-t} + \frac{1}{2}e^{3t} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$x_S(t) = \int_0^t \begin{bmatrix} \frac{1}{2}e^{-(t-s)} + \frac{1}{2}e^{3(t-s)} & -\frac{1}{4}e^{-(t-s)} + \frac{1}{4}e^{3(t-s)} \\ -e^{-(t-s)} + e^{3(t-s)} & \frac{1}{2}e^{-(t-s)} + \frac{1}{2}e^{3(t-s)} \end{bmatrix} \cdot \begin{bmatrix} s \\ s^2 \end{bmatrix} ds$$

Oczywiście ostatnią całkę w $x_S(t)$ należy wyliczyć ze znanych wzorów, ale to już za dużo roboty jak na to opracowanie.

6.4. Metoda sprowadzania układu równań do równania rzędu wyższego

PRZYKŁAD 6.4.1. Rozwiążemy zagadnienie Cauchy'ego dla układu równań:

$$\begin{aligned}
 (\text{URL}) \quad & \begin{cases} \frac{dx}{dt} = x + y + t & (1) \\ \frac{dy}{dt} = 4x + y + t^2 & (2) \end{cases} \quad (\text{WP}) \quad \begin{cases} x(0) = 1 \\ y(0) = 1 \end{cases} \\
 (2) - (1) : & \quad \frac{dy}{dt} - \frac{dx}{dt} = 3x + t^2 - t \Rightarrow \frac{dy}{dt} = \frac{dx}{dt} + 3x + t^2 - t \\
 \frac{d}{dt}(1) : & \quad \frac{d^2x}{dt^2} = \frac{dx}{dt} + \frac{dy}{dt} + 1 \\
 \begin{cases} \frac{dy}{dt} = \frac{dx}{dt} + 3x + t^2 - t \\ \frac{d^2x}{dt^2} = \frac{dx}{dt} + \frac{dy}{dt} + 1 \end{cases} & \Rightarrow \frac{d^2x}{dt^2} - 2\frac{dx}{dt} - 3x = t^2 - t + 1 \quad (*) \\
 (\text{RCH}) : & \quad \begin{aligned} \ddot{x} - 2\dot{x} - 3x &= 0 \\ \lambda^2 - 2\lambda - 3 &= 0 \\ \Delta = 16 \quad \sqrt{\Delta} &= 4 \\ \lambda_1 = -1 \quad \lambda_2 &= 3 \\ x_1(t) = e^{-t} \quad x_2(t) &= e^{3t} \\ x_{\text{RORJ}}(t) &= C_1 e^{-t} + C_2 e^{3t} \end{aligned}
 \end{aligned}$$

Rozwiązanie szczególne równania (*) znajdziemy metodą przewidywań:

$$\begin{aligned}
 \begin{cases} x_S(t) = At^2 + Bt + C \\ \dot{x}_S(t) = 2At + B \\ \ddot{x}_S(t) = 2A \end{cases} & \Rightarrow (*) \Rightarrow 2A - 4At - 2B - 3At^2 - 3Bt - 3C = t^2 - t + 1 \\
 \begin{cases} -3A = 1 \Rightarrow A = -\frac{1}{3} \\ -4A - 3B = -1 \Rightarrow B = \frac{1}{3}(4 \cdot \frac{1}{3} + 1) = \frac{7}{9} \\ 2A - 2B - 3C = 1 \Rightarrow C = \frac{1}{3}(-2 \cdot \frac{1}{3} - 2 \cdot \frac{7}{9} - 1) = -\frac{29}{27} \end{cases} & \Rightarrow \begin{cases} A = -\frac{1}{3} \\ B = \frac{7}{9} \\ C = -\frac{29}{27} \end{cases}
 \end{aligned}$$

Zatem $x_S(t) = -\frac{1}{3}t^2 + \frac{7}{9}t - \frac{29}{27}$ i całka ogólna ma postać:

$$x(t) = x_{\text{RORJ}}(t) + x_S(t) = C_1 e^{-t} + C_2 e^{3t} - \frac{1}{3}t^2 + \frac{7}{9}t - \frac{29}{27}$$

Rozwiązanie układu równań zaś ma postać :

$$\begin{cases} x(t) = C_1 e^{-t} + C_2 e^{3t} - \frac{1}{3}t^2 + \frac{7}{9}t - \frac{29}{27} \\ y(t) = \frac{dx}{dt} - x - t \end{cases}$$

Ponieważ $\frac{dx}{dt} = -C_1 e^{-t} + 3C_2 e^{3t} - \frac{2}{3}t + \frac{7}{9}$, więc $y(t) = -2C_1 e^{-t} + 2C_2 e^{3t} + \frac{1}{3}t^2 - \frac{22}{9}t + \frac{50}{27}$ i mamy rozwiązanie ogólne układu w postaci:

$$\begin{cases} x(t) = C_1 e^{-t} + C_2 e^{3t} - \frac{1}{3}t^2 + \frac{7}{9}t - \frac{29}{27} \\ y(t) = -2C_1 e^{-t} + 2C_2 e^{3t} + \frac{1}{3}t^2 - \frac{22}{9}t + \frac{50}{27} \end{cases}$$

Uwzględniając warunek początkowy (zagadnienie Cauchy'ego) mamy $x(0) = y(0) = 1 \Rightarrow C_1 = \frac{5}{4}, C_2 = \frac{89}{108}$.

Rozdział 7

Słabe i mocne prawa wielkich liczb

7.1. Słabe prawa wielkich liczb

Omówimy najpierw serię twierdzeń znanych pod nazwą słabych praw wielkich liczb. Nazwa ta związana jest z typem zbieżności ciągów zmiennych losowych występujących w tych twierdzeniach, zwanym zbieżnością według prawdopodobieństwa.

7.1.1. Zbieżność względem prawdopodobieństwa

Definicja 7.1.1. Mówimy, że ciąg zmiennych losowych (X_n) określonych na tej samej przestrzeni probabilistycznej $(\Omega, \mathcal{F}, \mathbf{P})$ jest zbieżny według prawdopodobieństwa do zmiennej losowej X (oznaczenie: $X_n \xrightarrow{P} X$), jeśli dla dowolnego $\epsilon > 0$ zachodzi

$$\lim_{n \rightarrow \infty} \mathbf{P}\{\omega : |X_n(\omega) - X(\omega)| < \epsilon\} = 1. \quad (7.1)$$

PRZYKŁAD 7.1.1. Aby zrozumieć dlaczego tak zdefiniowana zbieżność nie jest zbyt mocna, rozważmy następujący przykład: $\Omega = [0, 1]$, \mathbf{P} jest unormowana miarą Lebesgue'a na Ω , a ciąg zmiennych losowych X_n jest określony wzorem:

$$X_n(\omega) = \mathbf{1}_{(0, \frac{1}{n})}(\{n\pi + \omega\}),$$

gdzie $\{a\}$ oznacza ułamkową część liczby a . Łatwo zauważyć, że X_n przyjmuje wartość 1 na odcinku o długości $\frac{1}{n}$, a na pozostałej części odcinka $[0, 1]$ przyjmuje wartość zero. Otrzymujemy więc:

$$\mathbf{P}\{\omega : |X_n - 0| < \epsilon\} \geq 1 - \frac{1}{n} \xrightarrow{n \rightarrow \infty} 1.$$

Stąd wynika, że $X_n \xrightarrow{P} 0$. Jeżeli jednak ustalimy $\omega \in \Omega$, to wiadomo, że dla dowolnego $n \in \mathbb{N}$ istnieje $k > n$, dla którego

$$\{k\pi + \omega\} \in (0, \frac{1}{k}), \text{ więc } X_k(\omega) = 1.$$

Oznacza to, że ciąg $X_n(\omega), n \in \mathbb{N}$ jest ciągiem rozbieżnym dla każdego ustalonego ω .

7.1.2. Słabe prawa wielkich liczb

Twierdzenie 7.1.1 (Słabe prawo wielkich liczb Bernoulliego). Niech S_n będzie zmienną losową oznaczającą liczbę sukcesów w n próbach Bernoulliego z prawdopodobieństwem sukcesu w pojedynczej próbie równym p . Wtedy dla każdego $\epsilon > 0$ zachodzi

$$\lim_{n \rightarrow \infty} \mathbf{P} \left\{ \omega : \left| \frac{S_n}{n} - p \right| < \epsilon \right\} = 1.$$

Twierdzenie 7.1.2 (Słabe prawo wielkich liczb Czebyszewa). Jeśli (X_n) jest ciągiem parami niezależnych zmiennych losowych, dla których dla każdego n istnieje skończona wariancja $\text{Var} X_n$, przy czym dla pewnego c $\text{Var} X_n \leq c < \infty$ dla wszystkich n , to dla każdego $\epsilon > 0$ zachodzi

$$\lim_{n \rightarrow \infty} \mathbf{P} \left\{ \omega : \left| \frac{1}{n} \sum_{k=1}^n (X_k - \mathbf{E} X_k) \right| < \epsilon \right\} = 1.$$

Twierdzenie 7.1.3 (Słabe prawo wielkich liczb Markowa). Jeśli (X_n) jest ciągiem zmiennych losowych takich, że

$$\lim_{n \rightarrow \infty} \frac{1}{n^2} \text{Var} \left(\sum_{k=1}^n X_k \right) = 0, \quad (7.2)$$

to dla każdego $\epsilon > 0$ zachodzi

$$\lim_{n \rightarrow \infty} \mathbf{P} \left\{ \omega : \left| \frac{1}{n} \sum_{k=1}^n (X_k - \mathbf{E} X_k) \right| < \epsilon \right\} = 1.$$

Powyższe twierdzenie nie zakłada identyczności rozkładów tylko wymaga, aby istniały i były skończone wariancje rozważanych zmiennych. Podamy teraz inną postać prawa wielkich liczb, w którym rezygnujemy z tego założenia kosztem założenia o jednakowych rozkładach zmiennych losowych.

Twierdzenie 7.1.4 (Słabe prawo wielkich liczb Chinczyna). Jeśli (X_n) jest ciągiem niezależnych zmiennych losowych o jednakowych rozkładach i skończonej wartości oczekiwanej $m = \mathbf{E} X_i$, to dla każdego $\epsilon > 0$ zachodzi

$$\lim_{n \rightarrow \infty} \mathbf{P} \left\{ \omega : \left| \frac{1}{n} \sum_{k=1}^n (X_k - m) \right| < \epsilon \right\} = 1.$$

7.2. Mocne prawa wielkich liczb

Innym typem zbieżności zmiennych losowych jest *zbieżność z prawdopodobieństwem 1*, którą nazywamy również *zbieżnością prawie na pewno* lub *zbieżnością prawie wszędzie*.

7.2.1. Zbieżność prawie wszędzie

Definicja 7.2.1. Mówimy, że ciąg zmiennych losowych (X_n) określonych na tej samej przestrzeni probabilistycznej $(\Omega, \mathcal{F}, \mathbf{P})$ jest prawie na pewno zbieżny do zmiennej losowej X (oznaczenie: $X_n \xrightarrow{p.n.} X$), jeśli

$$\mathbf{P}\{\omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\} = 1. \quad (7.3)$$

Uwaga. Niech (X_n) będzie ciągiem zmiennych losowych określonych na tej samej przestrzeni probabilistycznej $(\Omega, \mathcal{F}, \mathbf{P})$. Wtedy

$$X_n \xrightarrow{p.n.} X \implies X_n \xrightarrow{P} X.$$

Widzimy zatem, iż zbieżność z prawdopodobieństwem 1 jest rzeczywiście "mocniejsza" od zbieżności według prawdopodobieństwa.

7.2.2. Mocne prawa wielkich liczb

Pierwszy wariant silnego prawa wielkich został sformułowany i wykazany przez Emila Borela w kontekście schematu Bernoulliego

Twierdzenie 7.2.1 (Mocne prawo wielkich liczb Borela). Niech S_n będzie zmienną losową oznaczającą liczbę sukcesów w n próbach Bernoulliego z prawdopodobieństwem sukcesu w pojedynczej próbie równym p . Wtedy dla każdego $\epsilon > 0$ zachodzi

$$\mathbf{P}\left\{\omega : \left| \lim_{n \rightarrow \infty} \frac{S_n}{n} = p \right| \right\} = 1.$$

Twierdzenie 7.2.2 (Pierwsze prawo wielkich liczb Kołmogorowa). Jeśli (X_n) jest ciągiem zmiennych losowych o skończonych wariancjach, oraz

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \text{Var} X_n < \infty,$$

to

$$\mathbf{P}\left\{\omega : \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n (X_k - \mathbf{E}X_k) = 0\right\} = 1.$$

Twierdzenie 7.2.3 (Drugie prawo wielkich liczb Kołmogorowa). Jeśli (X_n) jest ciągiem niezależnych zmiennych losowych o jednakowych rozkładach i skończonej wartości oczekiwanej $m = \mathbf{E}X_k$, to

$$\mathbf{P}\left\{\omega : \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n X_k = m\right\} = 1.$$

7.3. Zastosowania

7.3.1. Metoda Monte Carlo obliczania całek

Twierdzenie 7.3.1. Niech $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ będzie funkcją borelowską, a X zmienną losową o wartościach w \mathbb{R}^n . Wtedy

$$\mathbf{E}\varphi(X) = \int_{\mathbb{R}^n} \varphi(x) \mu_X(dx),$$

przy czym równość należy rozumieć tak: jeśli całka po jednej stronie istnieje, to istnieje także całka po drugiej stronie i są one równe.

Wniosek 7.3.1. Jeśli zmienna losowa X o wartościach w \mathbb{R}^n ma rozkład ciągły o gęstości g , a funkcja $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ jest borelowska, to

$$\mathbf{E}\varphi(X) = \int_{\mathbb{R}^n} \varphi(x) g(x) dx.$$

Twierdzenie 7.3.2. Niech X_1, X_2, \dots, X_n będzie ciągiem niezależnych zmiennych losowych o jednakowych rozkładach i skończonej wartości oczekiwanej, dodatkowo niech powyższe zmienne losowe przyjmują wartości w $(0, 1)$ i mają funkcję gęstości g . Niech f będzie funkcją rzeczywistą taką, że $\int_0^1 f(x) dx$ istnieje i jest skończona. Wtedy

$$\frac{1}{n} \sum_{k=1}^n \frac{f(X_k)}{g(X_k)} \xrightarrow{p.n.} \int_0^1 f(x) dx. \quad (7.4)$$

Dowód. Na mocy drugiego prawa wielkich liczb Kołmogorowa i wniosku 7.3.1 otrzymujemy:

$$\frac{1}{n} \sum_{k=1}^n \frac{f(X_k)}{g(X_k)} \xrightarrow{p.n.} \mathbf{E} \left(\frac{f(X_k)}{g(X_k)} \right) = \int_0^1 \frac{f(x)}{g(x)} g(x) dx = \int_0^1 f(x) dx. \quad (7.5)$$

■

Rozdział 8

Centralne Twierdzenie Graniczne rachunku prawdopodobieństwa

8.1. Zbieżność względem rozkładu

8.1.1. Słaba zbieżność rozkładów

Twierdzenie 8.1.1. Niech $(\mu_n)_{n=1}^\infty$ będzie ciągiem rozkładów prawdopodobieństwa na $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Powiemy, że jest on *słabo zbieżny* do rozkładu μ (co będziemy oznaczać poprzez $\mu_n \xrightarrow{\text{sl}} \mu$ lub $\mu_n \Rightarrow \mu$), jeśli dla dowolnej funkcji ciągłej i ograniczonej $f : \mathbb{R} \rightarrow \mathbb{R}$ zachodzi

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} f d\mu_n = \int_{\mathbb{R}} f d\mu.$$

Uwaga. Nazwa *słaba zbieżność* nawiązuje tutaj do ogólnych pojęć analizy funkcjonalnej. Każda funkcja ograniczona i ciągła f na prostej definiuje wzorem $\mu \rightarrow \int f d\mu$ funkcjonal na przestrzeni miar na \mathbb{R} . Żądamy więc, by dla każdego takiego funkcjonału zachodziła zbieżność jego wartości na ciągu $(\mu_n)_{n=1}^\infty$ do wartości na mierze μ . Tego rodzaju zbieżność nazywa się słabą, aby odróżnić ją od mocnej zbieżności, wyznaczonej przez normę.

8.1.2. Zbieżność względem rozkładu

Definicja 8.1.1 (Zbieżność względem rozkładu). Niech X_1, X_2, X_3, \dots będzie ciągiem zmiennych losowych o rozkładach $\mu_1, \mu_2, \mu_3, \dots$. Mówimy, że ciąg (X_n) jest *zbieżny według rozkładu* do zmiennej losowej X (oznaczenie $X_n \xrightarrow{d} X$) jeśli $\mu_n \Rightarrow \mu$.

Zbieżność względem rozkładu przekłada się na bardzo ciekawą własność dystrybuant rozwiązań zmiennych losowych.

Twierdzenie 8.1.2. Niech X_1, X_2, X_3, \dots będzie ciągiem zmiennych losowych, a ciąg F_1, F_2, F_3, \dots ciągiem odpowiadającym im dystrybuant. Wówczas $X_n \xrightarrow{d} X$ wtedy i tylko wtedy, gdy

$$\lim_{n \rightarrow \infty} F_n(x) = F(x)$$

w każdym punkcie ciągłości dystrybuanty granicznej F .

8.2. CTG rachunku prawdopodobieństwa

8.2.1. Twierdzenie de Moivre'a-Laplace'a

Twierdzenie 8.2.1. Niech S_n będzie liczbą sukcesów w n próbach Bernoulliego z prawdopodobieństwem sukcesu p . Wówczas

$$\frac{S_n - \mathbb{E}S_n}{\sqrt{\text{Var}(S_n)}} = \frac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow{d} N(0, 1).$$

Równoważnie, dla dowolnego $t \in \mathbb{R}$ zachodzi

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\frac{S_n - np}{\sqrt{np(1-p)}} < t \right) = \Phi(t),$$

gdzie Φ jest dystrybuantą standardowego rozkładu normalnego $N(0, 1)$.

8.2.2. Twierdzenie Lindeberga-Levy'ego

Twierdzenie 8.2.2. Niech (X_n) będzie ciągiem niezależnych zmiennych losowych o tym samym rozkładzie i parametrach $\mathbb{E}X_i = \mu$, $0 < \text{Var}X_i = \sigma^2 < \infty$. Połóżmy $S_n = X_1 + X_2 + \dots + X_n$, wówczas

$$\frac{S_n - \mathbb{E}S_n}{\sqrt{\text{Var}(S_n)}} = \frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} N(0, 1).$$

Równoważnie, dla dowolnego $t \in \mathbb{R}$ zachodzi

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\frac{S_n - n\mu}{\sigma\sqrt{n}} < t \right) = \Phi(t),$$

gdzie Φ jest dystrybuantą standardowego rozkładu normalnego $N(0, 1)$.

Nierówność Berry-Essena

Twierdzenie 8.2.3. Niech (X_n) będzie ciągiem niezależnych zmiennych losowych o tym samym rozkładzie i niezerowej wariancji, ponadto $\mathbb{E}|X_n|^3 < \infty$. Połóżmy $S_n = X_1 + X_2 + \dots + X_n$, wówczas:

$$\sup_{x \in \mathbb{R}} \left| \mathbf{P} \left(\frac{S_n - \mathbb{E}S_n}{\sqrt{\text{Var}S_n}} < x \right) - \Phi(x) \right| \leq C \frac{\mathbb{E}|X_1 - \mathbb{E}X_1|^3}{\sigma^3 \sqrt{n}}, \quad (8.1)$$

gdzie $\sigma = \sqrt{\text{Var}X_1}$, a $\frac{1}{\sqrt{2\pi}} \leq C < 0,8$.

Poniżej przedstawione zostaną twierdzenia uogólniające twierdzenie Lindeberga-Levy'ego na zmienne losowe o różnych rozkładach.

8.2.3. Twierdzenie Lapunowa

Definicja 8.2.1 (Warunek Lapunowa). Ciąg niezależnych zmiennych losowych (X_n) spełnia warunek Lapunowa, jeśli dla wszystkich $k \in \mathbb{N}$ i dla pewnego $\delta > 0$ zachodzi

$$\mathbb{E}|X_k|^{2+\delta} < \infty,$$

oraz

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^{2+\delta}} \sum_{k=1}^n \mathbb{E}|X_k - \mathbb{E}X_k|^{2+\delta} = 0,$$

gdzie

$$s_n^2 = \sum_{k=1}^n \text{Var}X_k.$$

Twierdzenie 8.2.4. Niech ciąg niezależnych zmiennych losowych (X_n) spełnia warunek Lapunowa, a S_n i s_n są zdefiniowane tak, jak poprzednio. Wówczas

$$\frac{S_n - \mathbb{E}S_n}{s_n} \xrightarrow{d} N(0, 1).$$

8.2.4. Twierdzenie Lindeberga

Definicja 8.2.2 (Warunek Lindeberga). Ciąg niezależnych zmiennych losowych (X_n) spełnia warunek Lindeberga, jeśli dla każdego $\epsilon > 0$ zachodzi

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^2} \sum_{k=1}^n \mathbb{E}[(X_k - \mathbb{E}X_k)^2 \cdot \mathbf{1}_{\{|X_k - \mu_k| > \epsilon s_n\}}] = 0,$$

gdzie

$$s_n^2 = \sum_{k=1}^n \text{Var}X_k.$$

Stwierdzenie. Jeśli ciąg zmiennych losowych spełnia warunek Lapunowa, to spełnia również warunek Lindeberga. Implikacja odwrotna nie jest prawdziwa.

Twierdzenie 8.2.5. Niech ciąg niezależnych zmiennych losowych (X_n) spełnia warunek Lindeberga, a S_n i s_n są zdefiniowane tak, jak poprzednio. Wówczas

$$\frac{S_n - \mathbb{E}S_n}{s_n} \xrightarrow{d} N(0, 1).$$

Rozdział 9

Metody estymacji nieznanymi parametrów rozkładu zmiennych losowych

9.1. Podstawowe definicje rachunku prawdopodobieństwa

Definicja 9.1.1. Niech (Ω, \mathcal{F}) oraz (E, \mathcal{E}) będą dwoma przestrzeniami mierzalnymi. Odwzorowanie $f : \Omega \rightarrow E$ nazywamy $(\mathcal{F}, \mathcal{E})$ -*mierzalnym*, jeżeli przeciwobraz każdego zbioru $A \in \mathcal{E}$ względem f należy do \mathcal{F} , tj.

$$\{\omega \in \Omega \mid f(\omega) \in A\} \in \mathcal{F}, \quad \forall A \in \mathcal{E}.$$

Uwaga. Jeżeli σ -ciała \mathcal{F} oraz \mathcal{E} są jasne z kontekstu to mówić będziemy po prostu o odwzorowaniach mierzalnych.

Definicja 9.1.2. Niech (Ω, \mathcal{F}, P) będzie przestrzenią probabilistyczną, a (E, \mathcal{E}) przestrzenią mierzalną. $(\mathcal{F}, \mathcal{E})$ -mierzalne odwzorowanie $X : \Omega \rightarrow E$ nazywamy *elementem losowym o wartościach w E* , bądź *zmienną losową o wartościach w E* .

9.2. Model statystyczny i statystyki dostateczne

Definicja 9.2.1. Zbiór wartości elementu losowego X będziemy oznaczali przez \mathcal{X} i nazywali *przestrzenią próby*.

Definicja 9.2.2. Niech $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ będzie rodziną rozkładów prawdopodobieństwa określonych na przestrzeni próby \mathcal{X} , indeksowaną parametrem θ . Parę $(\mathcal{X}, \{P_\theta : \theta \in \Theta\})$ nazywamy *modelem statystycznym*. Zbiór \mathcal{X} nazywamy *przestrzenią obserwacji*, zaś Θ nazywamy *przestrzenią parametrów*.

Definicja 9.2.3. Niech \mathcal{X} będzie przestrzenią obserwacji, a $(\mathcal{T}, \mathcal{U})$ będzie przestrzenią mierzalną. Wówczas *statystyką* nazywamy mierzalną funkcję $T : \mathcal{X} \rightarrow \mathcal{T}$.

Definicja 9.2.4. T jest *statystyką dostateczną* dla rodziny rozkładów $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ (dla parametru θ), jeżeli dla każdej wartości t rozkład warunkowy $P_\theta(\cdot | T = t)$ nie zależy od θ .

PRZYKŁAD 9.2.1. Niech $X = (X_1, \dots, X_n)$ będzie próbą z populacji o rozkładzie Poissona opisanym funkcją prawdopodobieństwa

$$p_\lambda(x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \dots,$$

gdzie $\lambda > 0$ jest parametrem. Wówczas statystyka

$$T(X) = \sum_{k=1}^n X_k$$

jest dostateczna dla parametru λ .

Twierdzenie 9.2.1 (Twierdzenie o faktoryzacji). Statystyka T jest dostateczna dla rodziny rozkładów $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ wtedy i tylko wtedy, gdy funkcję gęstości próby X można przedstawić w postaci

$$p_\theta(x) = g_\theta(T(x))h(x),$$

gdzie funkcja h nie zależy od parametru θ , a funkcja g zależy od x tylko poprzez wartości funkcji T .

Definicja 9.2.5. Statystykę dostateczną S nazywamy *minimalną statystyką dostateczną* jeżeli dla każdej statystyki dostatecznej T istnieje funkcja H taka, że $S = H(T)$.

PRZYKŁAD 9.2.2. Niech $X = (X_1, \dots, X_n)$ będzie próbą z populacji o rozkładzie Poissona z parametrem λ . Wówczas, statystyka $T(X) = \sum_{k=1}^n X_k$ jest minimalną statystyką dostateczną dla parametru λ .

Definicja 9.2.6. Statystyka T jest *zupełna* dla rodziny rozkładów $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ (dla parametru θ), gdy z warunku

$$\forall \theta \in \Theta \quad \mathbb{E}_\theta[h(T)] = 0$$

wynika, że $h \equiv 0$ prawie wszędzie względem \mathcal{P} .

Twierdzenie 9.2.2. Jeżeli statystyka T jest dostateczna i zupełna dla rodziny rozkładów $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$, to T jest minimalną statystyką dostateczną dla rodziny \mathcal{P} .

9.3. Estymatory

Definicja 9.3.1. Niech $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ będzie rodziną rozkładów określonych na przestrzeni próby \mathcal{X} , gdzie P_θ opisuje wielowymiarowy rozkład łączny wszystkich obserwacji w próbie X . *Estymatorem parametru θ* nazywamy statystykę $\hat{\theta}(X) : \mathcal{X} \rightarrow \Theta$.

Definicja 9.3.2. Niech $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ będzie rodziną rozkładów określonych na przestrzeni próby \mathcal{X} , gdzie P_θ opisuje wielowymiarowy rozkład łączny wszystkich obserwacji w próbie X . Dodatkowo, niech $g(\theta) : X \times \Theta \rightarrow (\mathcal{T}, \mathcal{S})$ będzie funkcją mierzalną. *Estymatorem funkcji parametrycznej funkcji $g(\theta)$* nazywamy statystykę $\hat{g}(X) : \mathcal{X} \rightarrow \mathcal{T}$.

9.3.1. Kryteria oceny jakości estymatorów

Zgodność

Definicja 9.3.3. Estymator $T_n = T(X_1, \dots, X_n)$ wielkości $g(\theta)$ jest *zgodny*, jeśli jest stochastycznie zbieżny do szacowanej funkcji parametru, tj.

$$\forall_{\epsilon > 0} \lim_{n \rightarrow \infty} P(|T_n - g(\theta)| < \epsilon) = 1.$$

Nieobciążoność

Definicja 9.3.4. Estymator $T = T(X_1, \dots, X_n)$ funkcji $g(\theta)$ jest *nieobciążony*, jeśli

$$\mathbb{E}_\theta(T(X_1, \dots, X_n)) = g(\theta), \quad \forall \theta \in \Theta.$$

Uwaga. Estymator nieobciążony nie zawsze istnieje.

PRZYKŁAD 9.3.1. Niech $X = (X_1, \dots, X_n)$ będzie próbą z populacji o rozkładzie ze skończoną wartością oczekiwaną μ . Wówczas statystyka

$$\hat{\mu}(X) = \frac{1}{n} \sum_{k=1}^n X_k$$

jest nieobciążonym estymatorem wartości oczekiwanej μ .

PRZYKŁAD 9.3.2. Niech $X = (X_1, \dots, X_n)$, $n > 1$ będzie próbą z populacji o rozkładzie ze skończoną i niezerową wariancją σ^2 . Wówczas statystyka

$$T(X) = \frac{1}{n-1} \sum_{k=1}^n (X_k - \hat{\mu}(X))^2$$

jest nieobciążonym estymatorem wariancji σ^2 .

9.3.2. Metody wyznaczania estymatorów

Metoda największej wiarygodności

W poniższej sekcji rozważamy model statystyczny $(\mathcal{X}, \{P_\theta : \theta \in \Theta\})$ taki, że wszystkie rozkłady z rodziny \mathcal{P} mają gęstości względem tej samej miary. Przez $p_\theta(x)$ oznaczamy gęstość rozkładu P_θ ¹

Definicja 9.3.5. Dla ustalonego $x \in \mathcal{X}$ wielkość

$$L(\theta; x) = p_\theta(x), \quad \theta \in \Theta,$$

nazywamy wiarygodnością θ , gdy zaobserwowano x .

¹Należy pamiętać, że dla rozkładów dyskretnych funkcję gęstości definiuje się przy użyciu funkcji masy prawdopodobieństwa.

Definicja 9.3.6. Estymatorem największej wiarygodności parametru θ nazywamy statystykę $\hat{\theta} : \mathcal{X} \rightarrow \Theta$, której wartości $\hat{\theta}(x)$, $x \in \mathcal{X}$ spełniają warunek:

$$L(\hat{\theta}(x); x) = \sup_{\theta \in \Theta} L(\theta; x), \quad x \in \mathcal{X}.$$

Zamiast "estymator największej wiarygodności" będziemy pisali krótko ENW, a zamiast "estymator największej wiarygodności parametru θ " będziemy pisali ENW(θ).

Uwaga. Dla danego parametru θ ENW może nie istnieć lub może być wyznaczony niejednoznacznie.

Uwaga. Przyjmujemy, że estymatorem największej wiarygodności funkcji parametrycznej $g(\theta)$ jest statystyka $g(\hat{\theta}(X))$, gdzie $\hat{\theta}(X)$ jest ENW parametru θ .

Zazwyczaj, podczas wyznaczania ENW, wygodniej jest operować funkcją $\log L$, niż funkcją L . Stosujemy wówczas oznaczenie $\ell(\theta) = \log L(\theta)$.

PRZYKŁAD 9.3.3. Niech $X = X_1, \dots, X_n$ będzie próbą z rozkładu wykładniczego o parametrze θ . Wówczas funkcja wiarygodności zadana jest wzorem

$$L(\theta; X) = \prod_{k=1}^n (\theta \exp(-\theta X_k)),$$

zatem

$$\ell(\theta; X) = n \log \theta - \theta \sum_{i=1}^n X_i.$$

Następnie przyrównując pochodną funkcji ℓ do zera

$$\ell'(\theta; X) = \frac{n}{\theta} - \sum_{i=1}^n X_i = 0.$$

otrzymujemy ENW(θ) równy $\hat{\theta} = 1/\bar{X}$, gdzie \bar{X} oznacza średnią z próby X .

Metoda momentów

Metoda momentów polega na przyrównaniu k pierwszych momentów empirycznych z k pierwszymi momentami (zwykłymi, bądź centralnymi) rozkładu teoretycznego. Z powstałych w ten sposób równań wyliczamy szukane parametry. Układamy tyle równań, ile jest niewiadomych parametrów.

PRZYKŁAD 9.3.4. Niech $X = X_1, \dots, X_n$ będzie próbą z rozkładu o ciągłej funkcji gęstości f_θ , $\theta \in \mathbb{R}^k$, to EMM znajdujemy rozwiązując poniższy układ równań

$$\begin{cases} \hat{\mu} = \mu(\theta), & \text{gdzie } \mu(\theta) = \int_{\mathbb{R}} x f_\theta(x) dx \\ \vdots \\ \hat{m}_k = m_k, & \text{gdzie } m_k = \int_{\mathbb{R}} (x - \mu(\theta))^k f_\theta(x) dx, \end{cases} \quad (9.1)$$

przy czym $\hat{\mu}$ to średnia z próby X a m_k oraz \hat{m}_k oznaczają odpowiednio k -ty moment teoretyczny i empiryczny.

PRZYKŁAD 9.3.5. Niech $X = X_1, \dots, X_n$ będzie próbą z rozkładu wykładniczego o parametrze θ . Z własności rozkładu wykładniczego wiemy, iż $\mu(\theta) = \frac{1}{\theta}$. Interesujący nas układ równań przedstawia się zatem następująco:

$$\frac{1}{\theta} = \frac{1}{n} \sum_{i=1}^n X_i.$$

Estymator parametru θ otrzymany metodą momentów jest zatem równy $\hat{\theta} = 1/\bar{X}$, gdzie \bar{X} oznacza średnią z próby X .

Rozdział 10

Metody numeryczne rozwiązywania układów równań liniowych

Rozdział 11

Przestrzenie liniowe, bazy, homomorfizmy przestrzeni liniowych i ich reprezentacje macierzowe

11.1. Przestrzenie wektorowe

Definicja 11.1.1. Czwórka $(V, +, \mathbb{K}, \cdot)$ nazywa się *przestrzenią wektorową nad \mathbb{K}* , jeśli V jest zbiorem, $+$ działaniem dwuargumentowym w zbiorze V , \mathbb{K} jest ciałem, a $\cdot : \mathbb{K} \times V \rightarrow V$ jest działaniem \mathbb{K} na V oraz spełnione są następujące warunki:

1. $(V, +)$ jest grupą abelową,
2. $a \cdot (b \cdot \mathbf{v}) = (ab) \cdot \mathbf{v}$,
3. $1 \cdot \mathbf{v} = \mathbf{v}$, gdzie 1 jest *jedynką* ciała \mathbb{K} ,
4. $(a + b) \cdot \mathbf{v} = a \cdot \mathbf{v} + b \cdot \mathbf{v}$,
5. $a \cdot (\mathbf{u} + \mathbf{v}) = a \cdot \mathbf{u} + a \cdot \mathbf{v}$,

dla dowolnych $a, b \in \mathbb{K}$, $\mathbf{u}, \mathbf{v} \in V$.

Uwaga. Elementy zbioru V będziemy nazywali *wektorami* przestrzeni wektorowej $(V, +, \mathbb{K}, \cdot)$, elementy ciała \mathbb{K} - *skalarami*. Poza przypadkami, gdy mogłoby to prowadzić do nieporozumień przestrzeń wektorową $(V, +, \mathbb{K}, \cdot)$ i zbiór V będziemy oznaczać tym samym symbolem \mathbf{V} . Element neutralny grupy $(V, +)$ nazywamy *wektorem zerowym* i oznaczamy symbolem \mathbf{o} .

PRZYKŁAD 11.1.1. Dowolne ciało \mathbf{K} może być traktowane jako przestrzeń wektorowa nad \mathbb{K} .

PRZYKŁAD 11.1.2. Jeśli w produkcji kartezjańskim

$$\mathbb{K}^n = \{(x_1, \dots, x_n) : x_i \in \mathbb{K} \text{ dla } i = 1, \dots, n\}$$

określimy dodawanie i mnożenie przez skalary wzorami

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n), \quad (11.1)$$

$$a \cdot (x_1, \dots, x_n) = (ax_1, \dots, ax_n), \quad (11.2)$$

to otrzymamy przestrzeń wektorową nad \mathbf{K}

Definicja 11.1.2. Przestrzeń wektorowa $(W, +, \mathbb{K}, \cdot)$ nazywa się *podprzestrzenią* przestrzeni wektorowej $(V, +, \mathbb{K}, \cdot)$, jeśli grupa $(W, +)$ jest podgrupą grupy $(V, +)$, a działanie \cdot ciała \mathbb{K} na zbiór W jest indukowane przez działanie \cdot ciała \mathbb{K} na zbiór V .

11.2. Układy wektorów

Definicja 11.2.1. *Układem wektorów* w przestrzeni wektorowej \mathbf{V} będziemy nazywali dowolny skończony ciąg $\mathcal{A} = (\mathbf{v}_1, \dots, \mathbf{v}_m)$ wektorów przestrzeni \mathbf{V} .

Definicja 11.2.2. Jeśli $\mathcal{A} = (\mathbf{v}_1, \dots, \mathbf{v}_m)$ i $1 \leq i_1 < \dots < i_k \leq m$, to układ $(\mathbf{v}_{i_1}, \dots, \mathbf{v}_{i_k})$ będziemy nazywali *podukładem* układu \mathcal{A} .

Definicja 11.2.3. Niech $\mathcal{A} = (\mathbf{v}_1, \dots, \mathbf{v}_m)$ będzie układem wektorów przestrzeni wektorowej \mathbf{V} na ciałem \mathbb{K} . Każde wyrażenie postaci

$$x_1 \cdot \mathbf{v}_1 + \dots + x_m \cdot \mathbf{v}_m, \quad (11.3)$$

gdzie $x_1, \dots, x_m \in \mathbf{K}$ będziemy nazywali *kombinacją liniową układu \mathcal{A}* lub *kombinacją liniową wektorów $\mathbf{v}_1, \dots, \mathbf{v}_m$* . Skalary x_1, \dots, x_m nazywają się współczynnikami kombinacji liniowej 11.3.

Uwaga. Zbiór wszystkich kombinacji liniowych układu $\mathcal{A} = (\mathbf{v}_1, \dots, \mathbf{v}_m)$ oznaczamy przez

$$\mathcal{L}(\mathcal{A}) = \{x_1 \cdot \mathbf{v}_1 + \dots + x_m \cdot \mathbf{v}_m, \quad x_i \in \mathbb{K}, \quad i = 1, \dots, m\}.$$

Twierdzenie 11.2.1. Jeśli $\mathcal{A} = (\mathbf{v}_1, \dots, \mathbf{v}_m)$ jest układem wektorów z przestrzeni wektorowej \mathbf{V} nad ciałem \mathbb{K} , to $\mathcal{L}(\mathcal{A})$ jest podprzestrzenią przestrzeni \mathbf{V} .

Definicja 11.2.4. Niech $\mathcal{A} = (\mathbf{v}_1, \dots, \mathbf{v}_m)$ będzie układem wektorów z przestrzeni wektorowej \mathbf{V} nad ciałem \mathbb{K} . Podprzestrzeń $\mathcal{L}(\mathcal{A})$ przestrzeni \mathbf{V} nazywamy *przestrzenią generowaną przez układ \mathcal{A}* , a układ \mathcal{A} nazywamy *układem generatorów* podprzestrzeni $\mathcal{L}(\mathcal{A})$.

Definicja 11.2.5. Niech \mathbf{V} będzie przestrzenią wektorową na ciałem \mathbb{K} i niech $\mathcal{A} = (\mathbf{v}_1, \dots, \mathbf{v}_m)$ będzie układem wektorów w przestrzeni \mathbf{V} . Mówimy, że układ \mathcal{A} jest *liniowo niezależny* (lub, że wektory $\mathbf{v}_1, \dots, \mathbf{v}_m$ są *liniowo niezależne*), jeśli z równości

$$x_1 \cdot \mathbf{v}_1 + \dots + x_m \cdot \mathbf{v}_m = \mathbf{0}$$

wynika, że $x_1 = 0, \dots, x_m = 0$. Jeżeli układ \mathcal{A} nie jest liniowo niezależny, to mówimy, że jest *liniowo zależny* (lub, że wektory $\mathbf{v}_1, \dots, \mathbf{v}_m$ są *liniowo zależne*).

Uwaga. Układ wektorów $\mathcal{A} = (\mathbf{v}_1, \dots, \mathbf{v}_m)$ w przestrzeni \mathbf{V} jest liniowo zależny wtedy i tylko wtedy, gdy co najmniej jeden z wektorów \mathbf{v}_i można zapisać jako kombinację liniową pozostałych.

Uwaga. W wielu sytuacjach wygodnie jest traktować niepuste układy wektorów jako macierze jednowierszowe. Jeśli traktujemy układ $\mathcal{A} = (\mathbf{u}_1, \dots, \mathbf{u}_m)$ jako macierz jednowierszową, to dla dowolnej macierzy

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$$

$$\mathcal{A} \cdot A = (\mathbf{u}_1, \dots, \mathbf{u}_m) \cdot \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} = (\mathbf{v}_1, \dots, \mathbf{v}_n),$$
$$\begin{aligned} \mathbf{v}_1 &= a_{11}\mathbf{u}_1 + \dots + a_{m1}\mathbf{u}_m = \mathcal{A} \cdot A^{(1)} \\ \dots & \\ \mathbf{v}_n &= a_{1n}\mathbf{u}_1 + \dots + a_{mn}\mathbf{u}_m = \mathcal{A} \cdot A^{(n)}. \end{aligned}$$

W szczególności wektor $\mathbf{v} = x_1 \mathbf{u}_1 + \dots + x_m \mathbf{u}_m$ może być przedstawiony w postaci iloczynu

$$\mathbf{v} = (\mathbf{u}_1, \dots, \mathbf{u}_m) \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = \mathcal{A}X, \quad \text{gdzie } X = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}$$

$$\mathcal{L}(\mathcal{A}) = \{\mathcal{A}X : X \in \mathbb{K}_m^1\}.$$

11.3. Baza i wymiar przestrzeni wektorowej

- jest liniowo niezależny,
- $\mathcal{L}(\mathcal{B}) = \mathbf{V}$.

Jako bazę przestrzeni zerowej przyjmujemy układ pusty.

PRZYKŁAD 11.3.1. Wielomiany $1, x, \dots, x^n$ tworzą bazę przestrzeni $\mathbb{K}_n[x]$.

Jeśli \mathbf{V} jest skończenie wymiarową przestrzenią wektorową, to liczbę elementów bazy przestrzeni \mathbf{V} nazywamy *wymiarem* przestrzeni \mathbf{V} i oznaczamy symbolem $\dim \mathbf{V}$.

1. \mathcal{A} jest bazą przestrzeni V ,
2. \mathcal{A} jest liniowo niezależny,

3. $\mathcal{L}(\mathcal{A}) = \mathbf{V}$.

Twierdzenie 11.3.2. Jeśli A jest macierzą typu $m \times n$ nad \mathbb{K} , to

1. rz A równy jest wymiarowi podprzestrzeni \mathbb{K}_m^1 generowanej przez kolumny macierzy A ,
2. rz A równy jest wymiarowi podprzestrzeni \mathbb{K}_1^n generowanej przez wiersze macierzy A .

Definicja 11.3.2. Niech $\mathcal{B} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ będzie bazą przestrzeni liniowej \mathbf{V} nad ciałem \mathbb{K} , oraz niech $\mathbf{v} \in \mathbf{V}$. Skalary $x_1, \dots, x_n \in \mathbb{K}$ nazywamy współrzędnymi wektora \mathbf{v} w bazie \mathcal{B} , gdy

$$\mathbf{v} = x_1 \cdot \mathbf{v}_1 + \dots + x_n \cdot \mathbf{v}_n.$$

Jednokolumnową macierz utworzoną ze współrzędnych wektora \mathbf{v} w bazie \mathcal{B} będziemy oznaczali symbolem $M_{\mathcal{B}}(\mathbf{v})$, tzn.

$$M_{\mathcal{B}}(\mathbf{v}) = \begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix}.$$

Nietrudno zauważyć, że

$$X = M_{\mathcal{B}}(\mathbf{v}) \iff \mathbf{v} = \mathcal{B}X.$$

Definicja 11.3.3. Niech $\mathcal{A} = (\mathbf{u}_1, \dots, \mathbf{u}_k)$ będzie układem wektorów w \mathbf{V} nad ciałem \mathbb{K} , wtedy *współrzędnymi układu \mathcal{A} w bazie \mathcal{B}* , będziemy nazywać k -kolumnową macierz nad \mathbb{K} , oznaczaną poprzez $M_{\mathcal{B}}(\mathcal{A})$, taką, że $(M_{\mathcal{B}}(\mathcal{A}))^{(j)} = M_{\mathcal{B}}(\mathbf{u}_j)$ dla $j = 1, \dots, k$. Jeśli więc

$$\begin{aligned} \mathbf{u}_1 &= a_{11}\mathbf{v}_1 + \dots + a_{n1}\mathbf{v}_n, \\ \dots\dots\dots, & \quad \text{to} \quad M_{\mathcal{B}}(\mathcal{A}) = A = \begin{bmatrix} a_{11} & \dots & a_{1k} \\ \dots\dots\dots \\ a_{n1} & \dots & a_{nk} \end{bmatrix}. \\ \mathbf{u}_k &= a_{1k}\mathbf{v}_1 + \dots + a_{nk}\mathbf{v}_n, \end{aligned}$$

Podobnie, jak poprzednio mamy

$$A = M_{\mathcal{B}}(\mathcal{A}) \iff \mathcal{A} = \mathcal{B}A.$$

Twierdzenie 11.3.3. Jeśli $\mathcal{A} = (\mathbf{u}_1, \dots, \mathbf{u}_n)$ i $\mathcal{B} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ są bazami przestrzeni wektorowej \mathbf{V} , to macierze

$$A = M_{\mathcal{B}}(\mathcal{A}) \quad \text{i} \quad B = M_{\mathcal{A}}(\mathcal{B})$$

są względem siebie odwrotne.

Dowód. $\mathcal{A} = \mathcal{B}A$ i $\mathcal{B} = \mathcal{A}B$. Stąd $\mathcal{B} = \mathcal{B}AB$, a stąd $I = AB$. ■

Definicja 11.3.4. Jeśli $\mathcal{A} = (\mathbf{u}_1, \dots, \mathbf{u}_n)$ i $\mathcal{B} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ są bazami przestrzeni wektorowej \mathbf{V} , to macierze $A = M_{\mathcal{B}}(\mathcal{A})$ i $B = M_{\mathcal{A}}(\mathcal{B})$ będziemy nazywali *macierzami zmiany bazy*.

Twierdzenie 11.3.4. Jeśli \mathcal{A} i \mathcal{B} są bazami przestrzeni wektorowej \mathbf{V} i $\mathbf{v} \in \mathbf{V}$, to

$$M_{\mathcal{B}}(\mathbf{v}) = M_{\mathcal{B}}(\mathcal{A})M_{\mathcal{A}}(\mathbf{v}) \quad \text{i} \quad M_{\mathcal{A}}(\mathbf{v}) = M_{\mathcal{A}}(\mathcal{B})M_{\mathcal{B}}(\mathbf{v}).$$

11.4. Odwzorowanie liniowe

Definicja 11.4.1. Niech $(\mathbf{V}, +)$ i (\mathbf{V}', \oplus) będą przestrzeniami wektorowymi na tym samym ciałem \mathbb{K} . Odwzorowanie $F : \mathbf{V} \rightarrow \mathbf{V}'$ nazywa się *odwzorowaniem liniowym* (lub *homomorfizmem przestrzeni wektorowych*), jeśli jest addytywne i jednorodne, tzn.:

Addytywne $F(\mathbf{u} + \mathbf{v}) = F(\mathbf{u}) \oplus F(\mathbf{v})$,

Jednorodne $F(a\mathbf{v}) = aF(\mathbf{v})$

dla dowolnych $\mathbf{u}, \mathbf{v} \in \mathbf{V}$, $a \in \mathbb{K}$.

Uwaga. Zbiór wszystkich odwzorowań liniowych z przestrzeni \mathbf{V} do przestrzeni \mathbf{V}' będzie oznaczany symbolem $\text{hom}(\mathbf{V}, \mathbf{V}')$. Odwzorowania linowe z przestrzeni \mathbf{V} w siebie będziemy nazywać *operatorami liniowymi* (lub krótko, *operatorami*) na przestrzeni \mathbf{V} . Zbiór wszystkich operatorów liniowych na \mathbf{V} będziemy oznaczali symbolem $\text{op}(\mathbf{V})$.

PRZYKŁAD 11.4.1. Jeśli \mathbf{V} jest przestrzenią wektorową, to odwzorowanie $I_{\mathbf{V}} : \mathbf{V} \rightarrow \mathbf{V}$ zdefiniowane przez $I_{\mathbf{V}}(\mathbf{v}) = \mathbf{v}$ dla każdego $\mathbf{v} \in \mathbf{V}$, jest operatorem liniowym na \mathbf{V} . Odwzorowanie $I_{\mathbf{V}}$ będziemy nazywali operatorem *identycznościowym* na przestrzeni \mathbf{V} .

PRZYKŁAD 11.4.2. Niech $V = \mathbb{R}[x]$ i niech odwzorowania $F : V \rightarrow V$, $G : V \rightarrow V$ będą zdefiniowane wzorami

$$(Ff)(x) = \frac{df}{dx}(x) \quad \text{ i } \quad (Gf)(x) = \int_0^x f(t)dt$$

dla $f(x) \in \mathbb{R}[x]$, gdzie $\mathbb{R}[x]$ oznacza pierścień wielomianów nad ciałem liczb rzeczywistych. Ze znanych właściwości pochodnych i całek wynika, że F i G są operatorami liniowymi na przestrzeni V .

Twierdzenie 11.4.1. Niech \mathbf{V} i \mathbf{V}' będą przestrzeniami liniowymi nad \mathcal{K} i niech F będzie odwzorowaniem z \mathbf{V} do \mathbf{V}' . F jest odwzorowaniem liniowym wtedy i tylko wtedy, gdy

$$F(a\mathbf{u} + \mathbf{v}) = aF(\mathbf{u}) + F(\mathbf{v})$$

dla dowolnych $\mathbf{u}, \mathbf{v} \in \mathbf{V}$, $a \in \mathbb{K}$.

Twierdzenie 11.4.2. Niech $\mathbf{V}, \mathbf{V}', \mathbf{V}''$ będą przestrzeniami wektorowymi nad \mathbb{K} . Wówczas

1. Jeśli $F \in \text{hom}(\mathbf{V}, \mathbf{V}')$ i $F' \in \text{hom}(\mathbf{V}', \mathbf{V}'')$, to $F' \circ F \in \text{hom}(\mathbf{V}, \mathbf{V}'')$.
2. Jeśli $F \in \text{hom}(\mathbf{V}, \mathbf{V}')$ i F jest odwzorowaniem odwracalnym, to $F^{-1} \in \text{hom}(\mathbf{V}', \mathbf{V})$.

Definicja 11.4.2. Niech $F \in \text{hom}(\mathbf{V}, \mathbf{V}')$. Zbiór

$$F^{-1}(\mathbf{0}) = \{\mathbf{v} \in \mathbf{V} : F\mathbf{v} = \mathbf{0}\}$$

nazywamy *jądrem* odwzorowania F i oznaczamy symbolem $\ker F$. Zbiór

$$F(\mathbf{V}) = \{\mathbf{v}' \in \mathbf{V}' : (\exists \mathbf{v} \in \mathbf{V}) \mathbf{v}' = F\mathbf{v}\}$$

nazywamy *obrazem* odwzorowania F i oznaczamy $\text{im} F$.

Definicja 11.4.3. Niech $F \in \text{hom}(\mathbf{V}, \mathbf{V}')$ i niech będą ustalone bazy \mathcal{B} i \mathcal{B}' przestrzeni \mathbf{V} i \mathbf{V}' . Macierz $M_{\mathcal{B}'}(F(\mathcal{B}))$ będziemy nazywali *macierzą odwzorowania F w bazach \mathcal{B} i \mathcal{B}'* i oznaczali symbolem $M_{\mathcal{B}'}^{\mathcal{B}}(F)$.

Twierdzenie 11.4.3. Niech $F \in \text{hom}(\mathbf{V}, \mathbf{V}')$ i niech będą ustalone bazy \mathcal{B} i \mathcal{B}' przestrzeni \mathbf{V} i \mathbf{V}' i niech $A \in \mathbb{K}_m^n$. Wówczas $A = M_{\mathcal{B}'}^{\mathcal{B}}(F)$ wtedy i tylko wtedy, gdy dla każdego $\mathbf{v} \in \mathbf{V}$

$$M_{\mathcal{B}'}(F\mathbf{v}) = AM_{\mathcal{B}}(\mathbf{v}). \quad (11.4)$$

Wniosek 11.4.1. Jeśli \mathcal{A} jest układem wektorów w przestrzeni V , to

$$M_{\mathcal{B}'}(F(\mathcal{A})) = M_{\mathcal{B}'}^{\mathcal{B}}(F)M_{\mathcal{B}}(F(\mathcal{A})).$$

Twierdzenie 11.4.4. Niech $\mathbf{V}, \mathbf{V}', \mathbf{V}''$ będą przestrzeniami wektorowymi i niech $\mathcal{B}, \mathcal{B}', \mathcal{B}''$ będą bazami w przestrzeniach $\mathbf{V}, \mathbf{V}', \mathbf{V}''$. Jeśli $F \in \text{hom}(\mathbf{V}, \mathbf{V}')$ i $F' \in \text{hom}(\mathbf{V}', \mathbf{V}'')$, to

$$M_{\mathcal{B}''}^{\mathcal{B}}(F' \circ F) = M_{\mathcal{B}''}^{\mathcal{B}'}(F') \cdot M_{\mathcal{B}'}^{\mathcal{B}}(F).$$

Rozdział 12

Podstawowe struktury algebraiczne - grupy, pierścienie, ciała, kraty.

12.1. Grupy

Niech będzie dany niepusty zbiór G wraz z wewnętrznym działaniem dwuargumentowym $\cdot : G \times G \rightarrow G$.

Definicja 12.1.1. **Półgrupą** nazywamy parę (G, \cdot) , czyli zbiór G wraz z działaniem $\cdot : G \times G \rightarrow G$ spełniającym następujący warunek:

- (1) $\forall_{a,b,c \in G} (ab)c = a(bc)$ **łączność**.

Jeżeli dodatkowo półgrupa spełnia warunek:

- (2) $\exists_{e \in G} \forall_{a \in G} ea = ae = a$ istnienie elementu **neutralnego** e ,

wówczas nazywamy ją **monoidem** lub **półgrupą z jedyneką**. Półgrupa lub monoid, które spełniają dodatkowo warunek:

- (3) $\forall_{a,b \in G} ab = ba$ **przemienność**,

nazywamy **przemiennymi**.

PRZYKŁAD 12.1.1. (1) Niech Ω będzie dowolnym zbiorem, a $M(\Omega)$ rodziną wszystkich przekształceń zbioru Ω (wszystkich odwzorowań Ω w Ω). $(M(\Omega), \circ, e_\Omega)$ jest monoidem, gdzie \circ jest działaniem składania przekształceń. Ogólnie monoid ten nie jest przemienny.

- (2) Niech Ω będzie dowolnym zbiorem, a $P(\Omega)$ zbiorem wszystkich jego podzbiorów. Ponieważ $A \cap (B \cap C) = (A \cap B) \cap C$ oraz $A \cup (B \cup C) = (A \cup B) \cup C$, więc mamy w $P(\Omega)$ dwa naturalne działania łączne. Ponadto $\emptyset \cup A = A$ i $A \cap \Omega = A$. Mamy więc dwa przemienne monoidy $(P(\Omega), \cup, \emptyset)$ i $(P(\Omega), \cap, \Omega)$.

Definicja 12.1.2. **Grupą** nazywamy parę (G, \cdot) , czyli zbiór G wraz z działaniem $\cdot : G \times G \rightarrow G$ spełniającym następujące warunki 1 ÷ 3:

- (1) $\forall_{a,b,c \in G} (ab)c = a(bc)$ **łączność**,

- (2) $\exists_{e \in G} \forall_{a \in G} ea = ae = a$ istnienie elementu **neutralnego** e ,

(3) $\forall a \in G \exists b \in G ab = ba = e$ istnienie elementu **odwrotnego do a** oznaczanego a^{-1} .

Jeżeli dodatkowo działanie \cdot spełnia warunek:

(4) $\forall a, b \in G ab = ba$ **przemienność**,

wówczas grupę (G, \cdot) nazywamy grupą przemianą lub **abelową**.

PRZYKŁAD 12.1.2. (1) Abelowe: $(\{0\}, +)$, $(\mathbb{Z}, +)$, $(\mathbb{Q} \setminus \{0\}, \cdot)$

(2) (S_n, \circ) - grupa symetryczna stopnia n (**twierdzenie Cayleya** - każdą grupę można utożsamiać z pewną podgrupą permutacji zbioru składającego się z elementów tej grupy), $(GL_n(\mathbb{R}), \cdot)$ - zbiór wszystkich macierzy kwadratowych stopnia n o wyrazach rzeczywistych i wyznaczniku różnym od zera, z działaniem mnożenia macierzy, $\pi_1(X, x_0)$

Definicja 12.1.3. Niech G będzie grupą, niepusty podzbiór $H \subseteq G$ nazywamy **podgrupą** grupy G ze względu na działanie określone na G jeśli:

(1) $\forall a, b \in H ab \in H$,

(2) $\forall a \in H a^{-1} \in H$.

W przypadku gdy G jest skończona, wówczas wystarczy warunek $\forall a, b \in H ab \in H$.

PRZYKŁAD 12.1.3. Podgrupa trywialna (e, \cdot) , podgrupa niewłaściwa (G, \cdot) , $(SL_n(\mathbb{R}), \cdot)$ - zbiór wszystkich macierzy kwadratowych stopnia n o wyrazach rzeczywistych i wyznaczniku równym 1 - jest to podgrupa grupy $(GL_n(\mathbb{R}), \cdot)$.

Definicja 12.1.4. Niech G będzie dowolną grupą, zaś H dowolną jej podgrupą. Podzbiory grupy G dane dla dowolnego $a \in G$ jako:

$$aH = \{ah : h \in H\}, \quad Ha = \{ha : h \in H\}$$

nazywamy odpowiednio **warstwami lewostronnymi** i **prawostronnymi** grupy G względem H wyznaczonymi przez element $a \in G$. Jeżeli $aH = Ha$ wówczas mówi się po prostu o **warstwach** (obustronnych). Zbiór wszystkich warstw lewostronnych oznaczamy przez $G/H = \{aH : a \in G\}$, zaś zbiór warstw prawostronnych przez $G \setminus H = \{Ha : a \in G\}$ (nie mylić z oznaczeniem różnicy zbiorów). W przypadku zbioru warstw obustronnych używamy po prostu pierwszego oznaczenia.

PRZYKŁAD 12.1.4. (1) Weźmy grupę S_3 permutacji zbioru 3 elementowego postaci $S_3 = \{e, \tau_1, \tau_2, \tau_3, \sigma_1, \sigma_2\}$, gdzie:

$$\begin{aligned} e &= (1), & \tau_1 &= (1\ 2), \\ \sigma_1 &= (1\ 2\ 3), & \tau_2 &= (1\ 3), \\ \sigma_2 &= (1\ 3\ 2), & \tau_3 &= (2\ 3). \end{aligned}$$

Ma ona cztery podgrupy właściwe (tzn. różne od podgrupy trywialnej i niej samej), wszystkie cykliczne. Trzy (izomorficzne) rzędu 2, są to: $\langle \tau_1 \rangle$, $\langle \tau_2 \rangle$, $\langle \tau_3 \rangle$, oraz jedną rzędu 3, jest to $\langle \sigma_1 \rangle = \langle \sigma_2 \rangle$. Znajdziemy warstwy lewostronne i prawostronne grupy S_3 względem podgrupy $\langle \sigma_1 \rangle = \{e, \sigma_1, \sigma_2\}$. Mamy:

$$\begin{aligned} e\langle \sigma_1 \rangle &= \{e, \sigma_1, \sigma_2\}, & \langle \sigma_1 \rangle e &= \{e, \sigma_1, \sigma_2\}, & \tau_1 \langle \sigma_1 \rangle &= \{\tau_1, \tau_2, \tau_3\}, & \langle \sigma_1 \rangle \tau_1 &= \{\tau_1, \tau_3, \tau_2\}, \\ \tau_2 \langle \sigma_1 \rangle &= \{\tau_2, \tau_3, \tau_1\}, & \langle \sigma_1 \rangle \tau_2 &= \{\tau_2, \tau_1, \tau_3\}, & \tau_3 \langle \sigma_1 \rangle &= \{\tau_3, \tau_1, \tau_2\}, & \langle \sigma_1 \rangle \tau_3 &= \{\tau_3, \tau_2, \tau_1\}, \\ \sigma_1 \langle \sigma_1 \rangle &= \{\sigma_1, \sigma_2, e\}, & \langle \sigma_1 \rangle \sigma_1 &= \{\sigma_1, \sigma_2, e\}, & \sigma_2 \langle \sigma_1 \rangle &= \{\sigma_2, e, \sigma_1\}, & \langle \sigma_1 \rangle \sigma_2 &= \{\sigma_2, \sigma_1, e\}. \end{aligned}$$

Widzimy, że:

$$\begin{aligned} e\langle\sigma_1\rangle &= \langle\sigma_1\rangle e, & \tau_1\langle\sigma_1\rangle &= \langle\sigma_1\rangle\tau_1, \\ \tau_2\langle\sigma_1\rangle &= \langle\sigma_1\rangle\tau_2, & \tau_3\langle\sigma_1\rangle &= \langle\sigma_1\rangle\tau_3, \\ \sigma_1\langle\sigma_1\rangle &= \langle\sigma_1\rangle\sigma_1, & \sigma_2\langle\sigma_1\rangle &= \langle\sigma_1\rangle\sigma_2. \end{aligned}$$

Nie jest to jednak regułą, że $aH = Ha$. Dla przykładu znajdźmy jeszcze warstwy lewostronne i prawostronne grupy S_3 względem podgrupy $\langle\tau_1\rangle = \{e, \tau_1\}$:

$$\begin{aligned} e\langle\tau_1\rangle &= \{e, \tau_1\}, & \langle\tau_1\rangle e &= \{e, \tau_1\}, & \tau_1\langle\tau_1\rangle &= \{\tau_1, e\}, & \langle\tau_1\rangle\tau_1 &= \{\tau_1, e\}, \\ \tau_2\langle\tau_1\rangle &= \{\tau_2, \sigma_2\}, & \langle\tau_1\rangle\tau_2 &= \{\tau_2, \sigma_1\}, & \tau_3\langle\tau_1\rangle &= \{\tau_3, \sigma_1\}, & \langle\tau_1\rangle\tau_3 &= \{\tau_3, \sigma_2\}, \\ \sigma_1\langle\tau_1\rangle &= \{\sigma_1, \tau_3\}, & \langle\tau_1\rangle\sigma_1 &= \{\sigma_1, \tau_2\}, & \sigma_2\langle\tau_1\rangle &= \{\sigma_2, \tau_2\}, & \langle\tau_1\rangle\sigma_2 &= \{\sigma_2, \tau_3\}. \end{aligned}$$

Widzimy, że w tym przypadku mamy:

$$\begin{aligned} e\langle\tau_1\rangle &= \langle\tau_1\rangle e, & \tau_1\langle\tau_1\rangle &= \langle\tau_1\rangle\tau_1, \\ \tau_2\langle\tau_1\rangle &\neq \langle\tau_1\rangle\tau_2, & \tau_3\langle\tau_1\rangle &\neq \langle\tau_1\rangle\tau_3, \\ \sigma_1\langle\tau_1\rangle &\neq \langle\tau_1\rangle\sigma_1, & \sigma_2\langle\tau_1\rangle &\neq \langle\tau_1\rangle\sigma_2, \end{aligned}$$

zatem nie zawsze $aH = Ha$. Oczywiście jest natomiast, że jeśli G jest abelowa to oczywiście $aH = Ha$ dla dowolnej jej podgrupy.

- (2) Wyznamy podział na warstwy lewostronne grupy \mathbb{Z}_8 względem podgrupy $H = \{0, 4\}$. Ponieważ \mathbb{Z}_8 jest abelowa, więc warstwy prawostronne będą takie same jak lewostronne. Mamy następujące 4 warstwy lewostronne:

$$\begin{aligned} 0 + H &= [0]_H = \{0, 4\} = H, & 1 + H &= [1]_H = \{1, 5\}, \\ 2 + H &= [2]_H = \{2, 6\}, & 3 + H &= [3]_H = \{3, 7\}. \end{aligned}$$

Definicja 12.1.5. Podgrupę N grupy G nazywa się **podgrupą normalną**, jeśli wszystkie jej warstwy lewostronne równają się odpowiadającym im warstwom prawostronnym, tzn. gdy $\forall g \in G \quad gN = Ng$. Fakt ten oznacza się symbolem $N \trianglelefteq G$. Jeżeli N jest podgrupą właściwą G (tzn. jest różna od podgrupy trywialnej i samej G) wówczas piszemy $N \triangleleft G$.

PRZYKŁAD 12.1.5. Oczywiście każda grupa ma zawsze co najmniej dwie podgrupy normalne tzn. podgrupę trywialną $\{e\} \trianglelefteq G$ i samą siebie $G \trianglelefteq G$. Jak widać z poprzedniego przykładu $\langle\sigma_1\rangle \triangleleft S_3$ jest podgrupą normalną w S_3 . Natomiast $\langle\tau_1\rangle$ nie jest podgrupą normalną w S_3 .

Definicja 12.1.6. Jeśli $H \trianglelefteq G$ to w zbiorze warstw G/H można wprowadzić **działanie indukowane**. Niech $xH = [x]_H$, $yH = [y]_H$ będą dwiema warstwami, zdefiniujmy działanie:

$$xH \cdot yH = (x \cdot y)H, \text{ czyli } [x]_H \cdot [y]_H = [x \cdot y]_H$$

wówczas zbiór warstw G/H z tak określonym działaniem tworzy grupę $(G/H, \cdot)$, zwaną **grupą ilorazową**.

PRZYKŁAD 12.1.6. W nawiązaniu do omawianych wcześniej przykładów mamy, że:

- (1) $S_3/\langle\sigma_1\rangle$ jest grupą ilorazową, dwuelementową, izomorficzną z grupami $S_3/\langle\sigma_1\rangle \simeq \langle\tau_1\rangle \simeq \langle\tau_2\rangle \simeq \langle\tau_3\rangle \simeq \mathbb{Z}_2$.
- (2) $\mathbb{Z}_8/\{0, 4\}$ jest grupą ilorazową izomorficzną z \mathbb{Z}_4 . Działanie na warstwach w $\mathbb{Z}_8/\{0, 4\}$ wygląda następująco:

+	[0]	[1]	[2]	[3]
[0]	[0]	[1]	[2]	[3]
[1]	[1]	[2]	[3]	[0]
[2]	[2]	[3]	[0]	[1]
[3]	[3]	[0]	[1]	[2]

czyli analogicznie do tabelki działania $+$ w \mathbb{Z}_4 .

Definicja 12.1.7. Odwzorowanie $f : G \rightarrow G'$ grupy $(G, *)$ w grupę (G', \circ) nazywamy **homomorfizmem**, jeśli:

$$\forall a, b \in G \quad f(a * b) = f(a) \circ f(b).$$

Jądrem homomorfizmu f nazywamy zbiór:

$$\text{Ker } f = \{g \in G : f(g) = e'\}$$

gdzie e' jest elementem neutralnym (jedynką) grupy (G', \circ) . Mamy następujące określenia poszczególnych rodzajów homomorfizmów:

- **Endomorfizm** - jest to homomorfizm grupy w tę samą grupę,
- **Epimorfizm** - jest to homomorfizm „na” będący surjekcją,
- **Monomorfizm** - jest to homomorfizm różnowartościowy będący injekcją,
- **Izomorfizm** - jest to homomorfizm różnowartościowy i „na” będący bijekcją,
- **Automorfizm** - jest to izomorfizm grupy w tę samą grupę, czyli izomorfizm będący endomorfizmem.

PRZYKŁAD 12.1.7. • Epimorfizm $f(\pi) = \varepsilon_\pi$ grup $(S_n, \circ) \rightarrow (C_2, \cdot)$ przypisujący permutacji jej znak, gdzie C_2 jest grupą cykliczną rzędu 2 postaci $C_2 = \langle -1 \rangle = (\{-1, 1\}, \cdot)$ (zauważmy że $\text{Ker } f = A_n$, gdzie A_n jest zbiorem permutacji parzystych, nazywany **grupą alternującą**),

- Epimorfizm $f(A) = \det(A)$ grup $(GL_n(\mathbb{R}), \cdot) \rightarrow (\mathbb{R}^*, \cdot)$ przypisujący macierzom kwadratowym stopnia n o wyrazach rzeczywistych i niezerowym wyznaczniku ich wyznacznik, gdzie \mathbb{R}^* jest zbiorem elementów odwracalnych w \mathbb{R} czyli $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ (zauważmy, że $\text{Ker } f = SL_n(\mathbb{R})$),
- Izomorfizm $f(x) = e^x$ grup $(\mathbb{R}, +) \rightarrow (\mathbb{R}^+, \cdot)$,
- Izomorfizm $f(k) = nk$ grup $(\mathbb{Z}, +) \rightarrow (n\mathbb{Z}, +)$, gdzie $n\mathbb{Z} = \{nk : k \in \mathbb{Z}\}$ jest podgrupą właściwą grupy \mathbb{Z} .

12.2. Pierścienie

Definicja 12.2.1. Niech R będzie pewnym niepustym zbiorem, w którym określone są dwa działania dwuargumentowe, które nazywać będziemy dodawaniem $+$ i mnożeniem \cdot spełniające następujące aksjomaty 1 ÷ 3:

- (1) $(R, +)$ jest grupą abelową,
- (2) (R, \cdot) jest grupą półgrupą (tzn. działanie \cdot jest łączne),
- (3) $\forall a, b, c \in R \quad (a + b)c = ac + bc \wedge c(a + b) = ca + cb$.

Wówczas system algebraiczny $(R, +, \cdot, 0)$ nazywamy **pierścieniem**. Grupę $(R, +)$ nazywamy grupą **addytywną** pierścienia, zaś półgrupę (R, \cdot) nazywamy półgrupą **multiplikatywną** pierścienia. Jeśli (R, \cdot) jest monoidem (tzn. ma jedynkę) to mówimy, że $(R, +, \cdot, 0, 1)$ jest **pierścieniem z jedynką**. Pierścień nazywamy **przemiennym** jeśli:

$$(4) \quad \forall_{a,b \in R} \quad ab = ba.$$

PRZYKŁAD 12.2.1. $(\mathbb{Z}, +, \cdot)$ jest pierścieniem liczb całkowitych ze zwykłymi działaniami dodawania i mnożenia (przemienny z jedynką). Analogicznie pierścieniami przemiennymi z jedynką są $(\mathbb{Q}, +, \cdot)$ i $(\mathbb{R}, +, \cdot)$ oraz pierścień $(\mathbb{Z}_m, +, \cdot)$ liczb całkowitych modulo m . Przykładem pierścienia nieprzemiennego z jedynką (dla $n > 1$) jest $(M_n(\mathbb{R}), +, \cdot)$ pierścień macierzy kwadratowych stopnia n , w którym jedynką jest macierz jednostkowa E .

12.3. Ciała

Definicja 12.3.1. **Ciałem** nazywamy pierścień przemienny K z jedynką $1 \neq 0$, w którym każdy niezerowy element jest odwracalny. Grupę $K^* = K \setminus \{0\} = U(K)$ elementów odwracalnych ciała nazywamy **grupą multiplikatywną ciała**.

PRZYKŁAD 12.3.1. Ciała klasyczne, nieskończone: $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$, $(\mathbb{C}, +, \cdot)$. Przykład ciała skończonego: F_m - pierścień $(\mathbb{Z}_m, +, \cdot)$ liczb całkowitych modulo m , dla m będącego liczbą pierwszą.

12.4. Kraty

Definicja 12.4.1. **Kratą** nazywamy strukturę matematyczną, którą można opisywać albo algebraicznie, albo w sensie częściowych porządków (jako **poset**).

Z algebraicznego punktu widzenia kraty to struktura algebraiczna (A, \wedge, \vee) , gdzie A jest niepustym zbiorem, zaś \wedge i \vee są działaniami dwuargumentowymi $A \times A \rightarrow A$ spełniającymi dla dowolnych $x, y, z \in A$ następujące warunki:

1. $x \wedge x = x$ $x \vee x = x$
2. $(x \wedge y) \wedge z = x \wedge (y \wedge z)$ $(x \vee y) \vee z = x \vee (y \vee z)$ łączyność
3. $x \wedge y = y \wedge x$ $x \vee y = y \vee x$ przemienność
4. $(x \wedge y) \vee y = y$ $(x \vee y) \wedge y = y$ absorpcja

Aksjomat 1 podaje się tradycyjnie w definicji kraty, ale nie jest to konieczne, bowiem wynika on z aksjomatu 4. W każdej kratce spełniona jest równoważność: $x \vee y = y \Leftrightarrow x \wedge y = x$. Relacja \leq zdefiniowana za pomocą równoważności:

$$x \leq y \Leftrightarrow x \vee y = y$$

jest relacją częściowego porządku (tzn. jest zwrotna, antysymetryczna i przechodnia), w którym każda para x, y ma kres górny i dolny:

$$\sup(x, y) = x \vee y \quad \inf(x, y) = x \wedge y$$

Krata w sensie porządków to niepusty, częściowy porządek (A, \leq) w którym każda para x, y ma kres górny $\sup(x, y)$ i dolny $\inf(x, y)$. Jeśli zdefiniujemy:

$$x \vee y := \sup(x, y) \quad x \wedge y := \inf(x, y)$$

to dostaniemy kratę w sensie algebraicznym (tzn. jako strukturę algebraiczną).

PRZYKŁAD 12.4.1. (1) Kratą jest algebr Boole'a tzn. struktura algebraiczna $\mathbb{B} := (B, \wedge, \vee, \neg, 0, 1)$ w której \wedge i \vee są działaniami dwuargumentowymi, \neg jest operacją jednoargumentową, a 0 i 1 są wyróżnionymi, różnymi elementami zbioru B . Stosuje się też notację $\mathbb{B} := (B, \cap, \cup, \sim, 0, 1)$ dla algebry Boole'a podzbiorów ustalonego zbioru.

(2) „Pięciokąt” lub krata N_5 to krata pięciu elementów a, b, c, d, e spełniających relacje:

- $\forall_{x \in N_5} \quad c \leq x \leq a$
- $d \wedge b = e \wedge b = c$
- $d \vee b = e \vee b = a$

(3) „Diamant” lub krata M_3 to krata pięciu elementów a, b, c, d, e spełniających relacje:

- $\forall_{x \in M_3} \quad c \leq x \leq a$
- $x \wedge y = c$ dla każdych $x \neq y$ w zbiorze $\{b, d, e\}$
- $x \vee y = a$ dla każdych $x \neq y$ w zbiorze $\{b, d, e\}$

Rozdział 13

Podstawowe konstrukcje algebraiczne - podalgebry, produkty, obrazy homomorficzne

Rozdział 14

Relacja równoważności

14.1. Relacja równoważności

Definicja 14.1.1. Jeśli dane są dwa zbiory X i Y , to każdy podzbiór $R \subset X \times Y$ nazywamy **relacją dwuargumentową** (binarną) między elementami zbiorów X i Y .

Jeżeli dla pary uporządkowanej (x, y) takiej, że: $x \in X$, $y \in Y$ mamy $(x, y) \in R$, to piszemy również xRy i mówimy, że x jest w relacji z y .

Przykłady

1. Dla każdej funkcji $f : X \rightarrow Y$ można rozpatrzyć jej *wykres*: jest to podzbiór

$$\Gamma(f) = \{(x, y) \mid x \in X, y = f(x)\} \subset X \times Y, \quad (14.1)$$

a więc pewna relacja pomiędzy elementami zbiorów X i Y .

Uwaga. Nie każda relacja R jest wykresem pewnego odwzorowania $f : X \rightarrow Y$.

Warunek konieczny i wystarczający jest następujący: dla każdego $x \in X$ istnieje dokładnie jeden element $y \in Y$ taki, że xRy .

2. $R = \{(x, y) \mid x \in \mathbb{N}, y \in \mathbb{N}, x + y \text{ jest liczbą parzystą}\}.$

Definicja 14.1.2. Relacja dwuargumentowa \sim w zbiorze X jest **relacją równoważności**, jeśli dla dowolnych $x, x', x'' \in X$ spełnione są następujące warunki:

1. $x \sim x$ (*zwrotność*);
2. $x \sim x' \implies x' \sim x$ (*symetria*);
3. $x \sim x' \text{ i } x' \sim x'' \implies x \sim x''$ (*przechodność*).

Zapis $a \not\sim b$ oznacza, że elementy $a, b \in X$ nie są w relacji \sim .

Definicja 14.1.3. Dla danego $x \in X$ podzbiór

$$[x]_{\sim} = \{x' \in X \mid x' \sim x\} \subset X \quad (14.2)$$

złożony z wszystkich elementów równoważnych z x nazywamy **klasą równoważności** (lub **abstrakcją**) elementu x . Jeżeli relacja równoważności znana jest z kontekstu, pisze się zwykle po prostu $[x]$.

Każdy element $x' \in [x]_{\sim}$ nazywamy **reprezentantem** klasy $[x]_{\sim}$.

Stwierdzenie. Zbiór klas równoważności względem relacji \sim stanowi rozkład zbioru X w tym sensie, że różne klasy abstrakcji są rozłączne i X jest ich sumą.

Stwierdzenie. Jeśli dany jest rozkład zbioru X na parami rozłączne podzbiory C_{α} , to istnieje taka relacja równoważności w X , że C_{α} są jej klasami.

14.2. Faktoryzacja odwzorowań

Definicja 14.2.1. Zbiór klas równoważności dla relacji równoważności \sim w zbiorze X oznaczamy przez X/\sim i nazywamy **zbiorem ilorazowym** zbioru X względem relacji \sim .

Definicja 14.2.2. Niech X będzie zbiorem, na którym określono relację równoważności \sim . Wtedy odwzorowanie $p : X \rightarrow X/\sim$ zadane wzorem

$$p(x) = [x]_{\sim} \quad (14.3)$$

nazywamy **odwzorowaniem kanonicznym** (lub **rzutem kanonicznym**) zbioru X na zbiór ilorazowy X/\sim .

Uwaga. Odwzorowanie kanoniczne jest *suriekcją*.

Dla danych zbiorów X i Y oraz odwzorowania $f : X \rightarrow Y$ definiujemy teraz następującą relację R_f w zbiorze X :

$$xR_fx' \Leftrightarrow f(x) = f(x'), \quad x, x' \in X. \quad (14.4)$$

Relacja z równania 14.4 jest relacją równoważności o następujących klasach abstrakcji:

$$[x]_{R_f} = \{x' \in X : f(x') = f(x)\}.$$

Odwzorowanie $f : X \rightarrow Y$ indukuje odwzorowanie $\bar{f} : X/R_f \rightarrow Y$, określone wzorem:

$$\bar{f}([x]_{R_f}) = f(x) \quad (14.5)$$

lub, co na jedno wychodzi,

$$(\bar{f} \circ p)(x) = f(x), \quad (14.6)$$

gdzie p jest odwzorowaniem kanonicznym (14.3) zbioru X na zbiór ilorazowy X/R_f .

Uwaga. Odwzorowanie \bar{f} jest iniekcją.

Diagram przemienny

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ & \searrow p \quad \nearrow \bar{f} & \\ & X/R_f & \end{array}$$

ilustruje *faktoryzację* (rozkład) odwzorowania f na złożenie suriekcji p i iniekcji \bar{f}

$$f = \bar{f}p.$$

Rozdział 15

Przeliczalność i nieprzeliczalność

15.1. Konstrukcja von Neumanna liczb naturalnych

Aksjomat 15.1.1 (Aksjomat nieskończoności). Istnieje zbiór X o następujących właściwościach:

- $\emptyset \in X$
- $\forall y \in X (S(y) \in X)$

gdzie $S(y)$ jest następnikiem porządkowym zbioru y , tj. $S(y) = y \cup \{y\}$. Zbiór spełniający powyższe właściwości nazywać będziemy *zbiorem induktywnym*.

Lemat 15.1.0.1. Niech $\mathcal{P} = \{Y \subset X : Y \text{ jest zbiorem induktywnym}\}$, wtedy zbiór $\cap_{P \in \mathcal{P}} P$ jest zbiorem induktywnym.

Twierdzenie 15.1.1. Istnieje najmniejszy, pod względem inkluzji, zbiór induktywny.

Definicja 15.1.1. Najmniejszy pod względem inkluzji zbiór induktywny nazywamy *zbiorem liczb naturalnych* i oznaczamy przez \mathbb{N} . Korzystając z induktywności \mathbb{N} :

- $\emptyset \in \mathbb{N}$ oznaczamy jako 0,
- $S(\emptyset) = \{\emptyset\}$ oznaczamy jako 1,
- $S(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\}$ oznaczamy jako 2,
- i tak dalej ...

Elementy tego zbioru nazywamy *liczbami naturalnymi*.

15.2. Teoria mocy

15.2.1. Zbiory przeliczalne

Definicja 15.2.1. Zbiory A i B nazywamy *równolicznymi*, gdy istnieje bijekcja $f : A \rightarrow B$. Równoliczność zbiorów oznaczamy przez $A \sim B$.

Twierdzenie 15.2.1. Dla dowolnych zbiorów X, Y, Z zachodzą następujące wzory:

1. $X \sim X$,
2. $X \sim Y \implies Y \sim X$,
3. $(X \sim Y) \wedge (Y \sim Z) \implies X \sim Z$.

Uwaga. Z twierdzenia 15.2.1 wynika, iż relacja równoliczności jest relacją równoważności.

Definicja 15.2.2. Zbiór A nazywamy *skończonym*, gdy $A \sim n$, dla $n \in \mathbb{N}$.

Definicja 15.2.3. Zbiór A nazywamy *nieskończonym*, gdy A nie jest skończony.

Definicja 15.2.4. *Zbiorami przeliczalnymi* nazywamy zbiory skończone lub równoliczne ze zbiorem \mathbb{N} wszystkich liczb naturalnych.

Lemat 15.2.1.1. Własności zbiorów przeliczalnych:

- podzbiór przeliczalnego zbioru jest przeliczalny,
- suma dowolnej skończonej ilości zbiorów przeliczalnych, jest zbiorem przeliczalnym,
- iloczyn kartezjański skończonej ilości zbiorów przeliczalnych jest przeliczalny.
- dla każdej przeliczalnej rodziny indeksowanej zbiorów $(A_n)_{n \in \mathbb{N}}$, takiej, że A_n jest zbiorem przeliczalnym dla każdego $n \in \mathbb{N}$, suma uogólniona $\cup_{n \in \mathbb{N}} A_n$ jest zbiorem przeliczalnym.
- zbiór wszystkich ciągów skończonych o wyrazach należących do ustalonego zbioru przeliczalnego jest zbiorem przeliczalnym.

PRZYKŁAD 15.2.1. Zbiór wszystkich liczb całkowitych ujemnych \mathbb{Z}^- jest zbiorem przeliczalnym, ponieważ istnieje bijekcja $f : \mathbb{N} \rightarrow \mathbb{Z}^-$ określona wzorem $f(n) = -n$.

PRZYKŁAD 15.2.2. Zbiór wszystkich liczb całkowitych \mathbb{Z} , jako suma skończonej ilości zbiorów przeliczalnych $\mathbb{Z}^- \cup \{0\} \cup \mathbb{Z}^+$, jest zbiorem przeliczalnym.

PRZYKŁAD 15.2.3. Zbiór liczb pierwszych, jako podzbiór zbioru liczb naturalnych, jest przeliczalny.

PRZYKŁAD 15.2.4. Zbiór \mathbb{Q} wszystkich liczb wymiernych jest zbiorem przeliczalnym.

Dowód. Niech $\mathbb{Z}^* = \mathbb{Z}/\{0\}$. Zbiór $\mathbb{Z} \times \mathbb{Z}^*$, jako iloczyn kartezjański zbiorów przeliczalnych, jest zbiorem przeliczalnym. Istnieje zatem bijekcja $f : \mathbb{N} \rightarrow \mathbb{Z} \times \mathbb{Z}^*$.

Każda liczba wymierna da się przedstawić w postaci $\frac{m}{n}$, gdzie $m \in \mathbb{Z}$, $n \in \mathbb{Z}^*$. Niech funkcja $g : \mathbb{Z} \times \mathbb{Z}^* \rightarrow \mathbb{Q}$ dana będzie następującym wzorem: $g(m, n) = \frac{m}{n}$ dla każdego elementu $(m, n) \in \mathbb{Z} \times \mathbb{Z}^*$.

Niech $h = g \circ f$. Wówczas $h : \mathbb{N} \rightarrow \mathbb{Q}$, jako złożenie dwóch bijekcji, jest bijekcją. Wnioskujemy stąd, że zbiór liczb wymiernych jest przeliczalny. ■

15.2.2. Zbiory nieprzeliczalne

Definicja 15.2.5. Zbiór nazywamy *nieprzeliczalnym*, gdy nie jest przeliczalny.

Twierdzenie 15.2.2 (Cantora). Zbiór liczb rzeczywistych nie jest przeliczalny.

PRZYKŁAD 15.2.5. Zbiór $2^{\mathbb{N}}$ jest zbiorem nieprzeliczalnym.

PRZYKŁAD 15.2.6. Zbiór $\mathbb{N}^{\mathbb{N}}$ jest zbiorem nieprzeliczalnym.

PRZYKŁAD 15.2.7. Zbiór $\mathbb{Q}^* = \{x \in \mathbb{R} : x \notin \mathbb{Q}\}$ liczby niewymiernych jest nieprzeliczalny.

Dowód. Gdyby zbiór \mathbb{Q}^* był przeliczalny, to zbiór liczb rzeczywistych $\mathbb{R} = \mathbb{Q} \cup \mathbb{Q}^*$, jako suma dwóch zbiorów przeliczalnych, również byłby zbiorem przeliczalnym, co przeczy twierdzeniu 15.2.2. ■

15.2.3. Zbiory mocy continuum

Definicja 15.2.6. Mówimy, że zbiór jest *mocy continuum*, gdy jest równoliczny z \mathbb{R} .

Twierdzenie 15.2.3. Jeżeli $A \subset \mathbb{R}$ i A zawiera pewien przedział otwarty, to A jest mocy continuum.

Twierdzenie 15.2.4. Jeżeli $B \subset A$ jest przeliczalnym podzbiorem zbioru A mocy continuum, to $A \setminus B$ jest mocy continuum.

Twierdzenie 15.2.5. Jeżeli B jest przeliczalnym, a A jest mocy continuum, to $A \cup B$ jest mocy continuum.

PRZYKŁAD 15.2.8. Zbiór wszystkich funkcji $f : \mathbb{N} \rightarrow \{0, 1\}$, tj. zbiór wszystkich ciągów $(a_n)_{n \in \mathbb{N}}$, takich że $a_n \in \{0, 1\}$ dla każdego $n \in \mathbb{N}$, jest mocy continuum.

PRZYKŁAD 15.2.9. Zbiór $2^{\mathbb{N}}$ jest mocy continuum.

PRZYKŁAD 15.2.10. Każdy przedział otwarty $(a, b) \subset \mathbb{R}$, gdzie $a < b$, jest mocy continuum.

PRZYKŁAD 15.2.11. Przedział otwarty $(-\frac{1}{2}\pi, \frac{1}{2}\pi) \subset \mathbb{R}$ jest mocy continuum.

Dowód. Funkcja $f : (-\frac{1}{2}\pi, \frac{1}{2}\pi) \rightarrow \mathbb{R}$ określona wzorem $f(x) = \operatorname{tg}(x)$ jest różnowartościowa i przekształca swoją dziedzinę na \mathbb{R} , ustala więc równoliczność tych zbiorów. ■

15.3. Nierówności dla liczb kardynalnych

Definicja 15.3.1. Niech $|A| = n$ oraz $|B| = m$. Przyjmujemy, że liczba kardynalna n jest nie większa od liczby kardynalnej m , wtedy i tylko wtedy, kiedy istnieje iniekcja $f : A \rightarrow B$, co oznaczamy poprzez $A \leq_m B$.

Definicja 15.3.2. Jeżeli $A \leq_m B$ i nieprawda, że $A \sim_m B$, to mówimy, że liczba kardynalna zbioru A jest *mniejsza* od liczby kardynalnej zbioru B , co oznaczamy poprzez $A <_m B$.

Twierdzenie 15.3.1. Następujące warunki są równoważne:

- Dla dowolnych zbiorów A, B zachodzi $A \leq_m B$ i $B \leq_m A$, to $A \sim_m B$.
- Dla dowolnych zbiorów A, B zachodzi $A \leq_m B$ i $B \subset A$, to $A \sim_m B$.
- Dla dowolnych zbiorów A, B, C zachodzi $A <_m B$ i $B <_m C$, to $A <_m C$.

Twierdzenie 15.3.2 (Cantora - Bernsteina). Jeżeli $A \leq_m B$ i $B \leq_m A$ to $A \sim_m B$.

Twierdzenie 15.3.3 (Cantora). $A <_m \mathcal{P}(A)$.

Twierdzenie 15.3.4. Nie istnieje zbiór wszystkich zbiorów.

Twierdzenie 15.3.5. $\mathbb{N} <_m \mathcal{P}(\mathbb{N}) \sim_m 2^{\mathbb{N}} \sim_m \mathbb{R}$.

Hipoteza continuum: czy istnieje taki zbiór, którego moc dałoby się ulokować pomiędzy mocą zbioru liczb naturalnych a mocą continuum. Czyli, czy istnieje A takie, że $\mathbb{N} <_m A <_m \mathbb{R}$?

Rozdział 16

Liczba chromatyczna grafu, twierdzenie Brooksa, twierdzenie o 4-kolorach

16.1. Definicja grafu

Definicja 16.1.1. Grafem prostym nazywamy parę uporządkowaną $\mathbf{G} = (V, E)$ taką, że:

- V jest niepustym zbiorem
- E jest rodziną dwuelementowych podzbiorów zbioru wierzchołków V :

$$E \subseteq \{\{u, v\} : u, v \in V, u \neq v\},$$

dalej zwanych **krawędziami**.

Definicja 16.1.2. Grafem skierowanym nazywamy parę uporządkowaną $\mathbf{G} = (V, E)$ taką, że:

- V jest niepustym zbiorem
- E jest zbiorem uporządkowanych par różnych wierzchołków ze zbioru V : $E \subseteq V \times V$,
dalej zwanych **krawędziami**.

Definicja 16.1.3 (Krawędzie incydentne). Jeśli istnieje krawędź vw to mówimy, że v i w są sąsiadami; oraz że krawędź vw jest incydentna do v (w).

Definicja 16.1.4 (Stopień wierzchołka). Stopień wierzchołka v w grafie \mathbf{G} to liczba krawędzi incydentnych z v . Stopień wierzchołka v oznaczany jest jako $\deg(v)$.

Przykłady

- **Graf spójny** to graf, w którym dla każdej pary wierzchołków istnieje ścieżka, która je łączy.
- **Graf pełny** to graf, w którym dla każdej pary węzłów istnieje krawędź je łącząca. Równoważnie nazywany kliką.

- **Graf pusty** to graf bez krawędzi. Równoważnie nazywany antykliką.
- Graf planarny, to graf, który można narysować na płaszczyźnie tak, by krzywe obrazujące krawędzie grafu nie przecinały się ze sobą.

16.2. Kolorowanie grafu

Definicja 16.2.1 (Kolorowanie grafu). Kolorowanie grafu $\mathbf{G} = (E, V)$ to funkcja $c : V \rightarrow N$ taka, że $c(v) \neq c(w)$ ilekroć vw jest krawędzią grafu \mathbf{G} .

Definicja 16.2.2 (Liczba chromatyczna grafu). Liczba chromatyczna grafu $\chi(\mathbf{G})$ to najmniejsza liczba barw, którymi można pokolorować graf \mathbf{G} .

Definicja 16.2.3. Graf k -kolorowalny (k -barwny) to graf dający się pokolorować k barwami.

Twierdzenie 16.2.1 (Tw. Brooksa). Niech $\mathbf{G} = (E, V)$ będzie spójnym grafem o największym stopniu wierzchołka równym d .

- Jeżeli \mathbf{G} jest grafem pełnym lub składa się z pojedynczego cyklu o nieparzystej liczbie krawędzi, to $\chi(G) = d + 1$.
- We wszystkich pozostałych przypadkach wystarcza $\chi(G) < d$.

Twierdzenie 16.2.2 (Twierdzenie o czterech barwach). Jeżeli \mathbf{G} jest grafem planarnym, to $\chi(G) < 4$.