

1 Wstęp do matematyki

Definicja 1: Relacja częściowego porządku

Mówimy, że relacja \sim w zbiorze X jest relacją częściowego porządku jeśli jest

1. zwrotna $x \sim x$
2. przechodnia $x \sim y \text{ i } y \sim z \Rightarrow x \sim z$
3. antysymetryczna $x \sim y \text{ i } y \sim x \Rightarrow y = x$

Definicja 2: Zbiór liniowo uporządkowany

Niech relacja \leq częściowo porządkuje zbiór X . Jeśli każde dwa elementy tego zbioru są porównywalne to relację \leq nazwiemy relacją linowego porządku na X .

Relacje niewiększości na zbiorze liczb całkowitych czy wymiernych są porządkami liniowymi.

Niech A będzie zbiorem liniowo uporządkowanym, wtedy:

1. w A istnienie co najwyżej jeden element minimalny,
2. jeśli w A istnieje element minimalny, to jest też elementem najmniejszym w A .

Definicja 3: Zbiór dobrze uporządkowany

Niech zbiór X będzie liniowo uporządkowany przez relację \leq . Liniowy porządek jest dobry jeśli

$$\forall \emptyset \neq A \subseteq X \text{ istnieje element najmniejszy}$$

Niech X będzie zbiorem dobrze uporządkowanym przez relację \leq , wtedy:

1. każdy właściwy odcinek początkowy w zbiorze X jest jest wyznaczony przez pewien element,
2. żaden ciąg o wyrazach w zbiorze X nie jest ciągiem malejącym w sensie porządku \leq ,
3. dla każdego jego elementu, z wyjątkiem największego, istnieje dokładnie jeden bezpośredni następnik.

Definicja 4: Izomorfizm porządkowy

Niech (X, \leq) oraz (Y, \preceq) będą porządkami częściowymi. Funkcja $h: X \rightarrow Y$ jest izomorfizmem porządkowym X, Y jeśli

1. h jest wzajemnie jednoznaczna (istnieje bijekcja)
2. $\forall x,y \in X \quad x \leq y \Leftrightarrow h(x) \preceq h(y)$

Definicja 5: Równoliczność

Zbiory A i B są równoliczne, gdy istnieje bijekcja (funkcja różnowartościowa i “na”) między zbiorami A i B .

Mówimy, że zbiór A ma mniej elementów niż zbiór B , jeśli $|A| \leq |B|$ oraz zbiory A i B nie są równoliczne. Piszemy wówczas $|A| < |B|$.

Definicja 6: Łańcuch

Jeśli podzbiór L zbioru X , częściowo uporządkowanego przez relację \leq jest liniowo uporządkowany przez relację $\leq|_L$, to mówimy, że L jest łańcuchem w zbiorze X .

Definicja 7: Element największy w zbiorze A

Niech X będzie zbiorem częściowo uporządkowanym przez relację \leq i niech $A \subseteq X$ oraz $a \in X$. Mówimy, że a jest elementem największym w zbiorze A , jeśli

$$a \in A \text{ i } \forall_{x \in A} x \leq a$$

to znaczy a jest elementem zbioru A mniejszym od wszystkich innych (różnych od a) elementów tego zbioru.

Definicja 8: Element maksymalny

Niech X będzie zbiorem częściowo uporządkowanym przez relację \leq i niech $A \subseteq X$ oraz $a \in X$. Mówimy, że a jest elementem maksymalnym w zbiorze A , jeśli

$$a \in A \text{ i } \forall_{x \in A} a \leq x \implies a = x$$

Definicja 9: Ograniczenie górne

Niech X będzie zbiorem częściowo uporządkowanym przez relację \leq i niech $A \subseteq X$ oraz $a \in X$. Mówimy, że a jest ograniczeniem górnym zbioru A , jeśli

$$\forall_{x \in A} x \leq a$$

Przy czym a nie musi należeć do zbioru X

Twierdzenie 1: Lemat Kuratowskiego Zorna

Niech X będzie niepustym zbiorem częściowo uporządkowanym przez relację \preceq . Załóżmy, że każdy łańcuch w zbiorze X ma ograniczenie górne w X . Wtedy w zbiorze X istnieje element maksymalny

Przykład zastosowań: Każda przestrzeń liniowa ma bazę. Każdy ideał właściwy I w pierścieniu P jest zawarty w pewnym ideale maksymalnym.

Przykład 1: Zbiory przeliczalne i nieprzeliczalne

Przeliczalne: zbiór parzystych liczb naturalnych, zbiór wszystkich liczb naturalnych, zbiór liczb pierwszych, zbiór liczb całkowitych, wymiernych.

Nieprzeliczalne: zbiór liczb rzeczywistych, przedział $[0, 1]$

Twierdzenie 2: Twierdzenie Cantora

Moc zbioru X jest mniejsza od mocy zbioru potęgowego $\mathcal{P}(X)$.

Pozwala to odpowiedzieć na pytanie czy istnieje zbiór o największej mocy. Załóżmy, że X jest takim zbiorem, ale wtedy $|X| < |\mathcal{P}(X)|$, sprzeczność.

Czy każdy zbiór nieprzeliczalny jest równoliczny ze zbiorem wszystkich liczb rzeczywistych? Nie, $|\mathbb{R}| < |\mathcal{P}(\mathbb{R})|$.

Twierdzenie 3: Twierdzenie Cantora-Bernsteina

Dla dowolnych zbiorów A i B , jeśli $|A| \leq |B|$ oraz $|B| \leq |A|$, to $|A| = |B|$ (A jest równoliczne B).

Własności obrazu i przeciwobrazu zbioru względem funkcji

Niech dana będzie funkcja $f: X \rightarrow Y$. Dla wszystkich podzbiorów $A, A_1, A_2 \subseteq X$ oraz $B, B_1, B_2 \subseteq Y$ zachodzą następujące własności:

- obraz jest podzbiorem przeciwdziedziny, a przeciwobraz dziedziny, $f[A] \subseteq Y$ oraz $f^{-1}[B] \subseteq X$;
- działania brania obrazu i przeciwobrazu związane są ze sobą następującymi relacjami: $f[f^{-1}[B]] \subseteq B$ (równość dla funkcji „na”),
- $f^{-1}[f[A]] \supseteq A$ (równość dla funkcji różnowartościowej),

Twierdzenie 4: Zachowanie operacji obrazu i przeciwobrazu

1. $A \subseteq B \implies f[A] \subseteq f[B]$
2. $A \subseteq B \implies f^{-1}[A] \subseteq f^{-1}[B]$
3. $f[A \cup B] = f[A] \cup f[B]$
4. $f[A \cap B] \subseteq f[A] \cap f[B]$
5. $f[A] \setminus f[B] \subseteq f[A \setminus B]$
6. $f^{-1}[A \cup B] = f^{-1}[A] \cup f^{-1}[B]$
7. $f^{-1}[A \cap B] = f^{-1}[A] \cap f^{-1}[B]$
8. $f^{-1}[A \setminus B] = f^{-1}[A] \setminus f^{-1}[B]$

2 Analiza Matematyczna

Definicja 10: Ciąg liczb rzeczywistych

Ciąg jest to funkcja określona na zbiorze liczb naturalnych. Ciągi o wyrazach rzeczywistych, czyli funkcje

$$a: \mathbb{N} \rightarrow \mathbb{R}$$

Definicja 11: Zbieżność ciągu liczb rzeczywistych

Ciąg (a_n) liczb rzeczywistych jest zbieżny do granicy $g \in \mathbb{R}$ (inaczej: ma granicę $g \in \mathbb{R}$) wtedy i tylko wtedy, gdy dla każdej liczby $\varepsilon > 0$ istnieje $n_\varepsilon \in \mathbb{N}$ takie, że dla wszystkich numerów $m > n_\varepsilon$ zachodzi nierówność

$$|a_m - g| < \varepsilon$$

Definicja 12: Warunek Cauchy'ego dla liczb rzeczywistych

Mówimy, że ciąg $(x_n) \subset X$ spełnia warunek Cauchy'ego wtedy i tylko wtedy, gdy dla każdej liczby $\varepsilon > 0$ istnieje $n_0 \in \mathbb{N}$ takie, że dla wszystkich $n, m > n_0$ jest

$$d(x_n, x_m) < \varepsilon$$

Twierdzenie 5: Warunek Cauchy'ego

Ciąg $(a_n) \subset \mathbb{R}$ jest zbieżny wtedy i tylko wtedy, gdy spełnia warunek Cauchy'ego.

Wynika z tego że \mathbb{R} jest przestrzenią zupełną.

Dowód: Ustalmy $\varepsilon > 0$. Wiemy, że ciąg Cauchy'ego jest ograniczony. Z twierdzenia Cauchy'ego-Weierstrassa (każdy ciąg ograniczony ma podciąg zbieżny) wynika, że możemy wybrać podciąg zbieżny do jakiejś granicy g . Istnieje wtedy takie n że $|a_{n_k} - g| < \frac{\varepsilon}{2}$ oraz $|a_n - a_{n_k}| < \frac{\varepsilon}{2}$ z czego wynika

$$|a_n - g| < |a_n - a_{n_k}| + |a_{n_k} - g| < \varepsilon$$

a więc ciąg jest zbieżny.

Definicja 13: Szereg liczbowy

Dla danego nieskończonego ciągu liczb rzeczywistych (a_n) definiuje się N -tą sumę częściową ciągu (a_n) bądź sumę częściową szeregu wzorem

$$s_N = \sum_{n=0}^N a_n = a_0 + a_1 + a_2 + \cdots + a_N.$$

Szeregiem nazywa się ciąg (s_N) sum częściowych. Formalnie szereg należy więc traktować jako parę uporządkowaną $((a_n), (s_N))$.

Sumą szeregu nazywa się liczbę $S = \lim_{N \rightarrow \infty} s_N$, o ile granica ta istnieje i jest właściwa. W przeciwnym przypadku szereg nie ma sumy. Szereg, który ma sumę nazywa się zbieżnym, który jej nie ma – rozbieżnym.

Definicja 14: Zbieżność i rozbieżność szeregu

O szeregu $\sum a_n$ mówi się, że jest zbieżny, jeżeli ciąg (s_N) sum częściowych ma skończoną granicę, wtedy

$$\sum_{n=0}^{\infty} a_n = \lim_{N \rightarrow \infty} s_N = \lim_{N \rightarrow \infty} \sum_{n=0}^N a_n.$$

Jeżeli granica (s_N) jest nieskończona lub nie istnieje, to szereg nazywa się rozbieżnym.

Szereg $\sum a_n$ jest zbieżny wtedy i tylko wtedy, gdy ciąg (s_N) sum częściowych spełnia warunek Cauchy'ego. Zmiana (również opuszczenie, czy dodanie) skończenie wielu wyrazów szeregu zbieżnego nie wpływa na jego zbieżność.

Definicja 15: Zbieżność bezwzględna

Szereg $\sum a_n$ nazywa się zbieżnym bezwzględnie/absolutnie, jeśli

$$\sum |a_n| < \infty.$$

Definicja 16: Zbieżność względna

Szereg $\sum a_n$ nazywa się zbieżnym względnie/nieabsolutnie, jeśli jest zbieżny, ale nie bezwzględnie, tzn.

$$\sum a_n < \infty, \quad \text{lecz} \quad \sum |a_n| = \infty.$$

Definicja 17: Ciągłość

Funkcja $f: \mathbb{R} \rightarrow \mathbb{R}$ jest ciągła jeśli spełnia

$$\forall \varepsilon \exists \delta \forall x \forall y \quad |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon$$

Definicja 18: Jednostajna ciągłość

Funkcja $f: \mathbb{R} \rightarrow \mathbb{R}$ jest jednostajnie ciągła jeśli spełnia

$$\forall \varepsilon \exists \delta \forall x \forall y \quad |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon$$

Różnica między tymi definicjami polega na tym, że przy ciągłości dobieramy deltę do każdego x osobno, zaś w jednostajnej ciągłości bierzemy jedną deltę która ma działać dla każdego x . Przykład funkcji ciągłej, ale nie jednostajnie ciągłej to np. $f(x) = \frac{1}{x}$ na $(0, 1]$.

Definicja 19: Pochodna jednowymiarowa w punkcie

Niech $f(x)$ będzie funkcją rzeczywistą zmiennej rzeczywistej x określoną w otoczeniu punktu x_0 . Pochodną funkcji $f(x)$ w punkcie x_0 nazywamy granicę (o ile istnieje):

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h},$$

Definicja 20: Pochodna kierunkowa

Niech $f: \mathbb{R}^n \supset A \rightarrow \mathbb{R}^m$ wzdłuż wektora jednostkowego $u = (u_1, \dots, u_n) \in \mathbb{R}^n$ w punkcie $x = (x_1, \dots, x_n) \in A$ nazywamy granicę

$$\frac{\partial f(x)}{\partial u} = \lim_{h \rightarrow 0} \frac{f(x + hu) - f(x)}{h},$$

zakładając, że ta granica istnieje.

Definicja 21: Różniczka zupełna

Niech $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$. f jest różniczkowalna w punkcie x wtedy i tylko wtedy, gdy istnieje przekształcenie liniowe $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ takie, że:

$$\lim_{\|h\| \rightarrow 0} \frac{\|f(x+h) - f(x) - A_x h\|}{\|h\|} = 0,$$

Przekształcenie A_x nazywa się różniczką, pochodną lub różniczką zupełną f w punkcie x .

Definicja 22: Pochodna cząstkowa

Niech $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$. Pochodną cząstkową funkcji f względem jej i -tej współrzędnej x_i nazywa się wartość granicy

$$\lim_{h \rightarrow 0} \frac{f(x_1, \dots, x_{i-1}, x_i + h, x_{i+1}, \dots, x_n) - f(x_1, \dots, x_n)}{h},$$

o ile istnieje i jest skończona.

Pochodne cząstkowe a różniczka zupełna

Niech $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ i $r > 0$. Jeśli wszystkie pochodne cząstkowe $\frac{\partial f}{\partial x_i}$ istnieją na całej kuli $B(a, r) \subseteq \Omega$ i są ciągłe w punkcie a , to f jest różniczkowalna w punkcie $a \in \Omega$. Zachodzi wtedy wzór

$$Df(a)h = \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(a), \quad h = (h_1, \dots, h_n) \in \mathbb{R}^n$$

Twierdzenie 6: Twierdzenie Rolle'a

Niech f będzie ciągłą funkcją rzeczywistą określoną na przedziale domkniętym $[a, b]$, różniczkowalną na przedziale otwartym (a, b) . Wówczas jeżeli

$$f(a) = f(b),$$

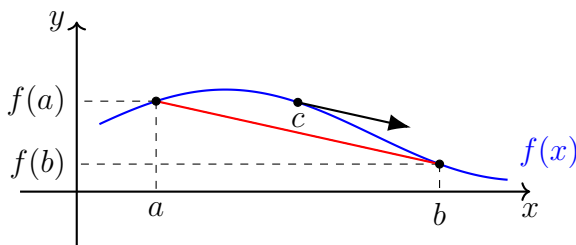
to istnieje taki punkt c należący do przedziału otwartego (a, b) , że $f'(c) = 0$.

Twierdzenie 7: Twierdzenie Lagrange'a (o wartości średniej)

Jeśli dana funkcja $f: [a, b] \rightarrow \mathbb{R}$ jest ciągła na $[a, b]$, różniczkowalna w przedziale (a, b) , to istnieje taki punkt $c \in (a, b)$, że:

$$\frac{f(b) - f(a)}{b - a} = f'(c).$$

Intuicja za tym twierdzeniem jest taka, że jeśli połączymy punkty $(a, f(a))$ i $(b, f(b))$ prostą to istnieje taki punkt c , że styczna do funkcji f w tym punkcie będzie równoległa do prostej łączącej punkty brzegowe.



Definicja 23: Szereg potęgowy

Szeregiem potęgowym o środku w punkcie $z_0 \in \mathbb{C}$ i współczynnikach $a_n \in \mathbb{C}$ nazywamy szereg

$$\sum_{n=0}^{\infty} a_n (z - z_0)^n.$$

Definicja 24: Przedział zbieżności szeregu

Niech (a_n) będzie dowolnym ciągiem liczb zespolonych i niech

$$\frac{1}{R} = \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|}.$$

Wtedy szereg potęgowy $S(z)$ jest, dla każdego $\rho < R$, zbieżny bezwzględnie i jednostajnie w kole $D_\rho = \{z \in \mathbb{C} : |z - z_0| \leq \rho\}$, oraz rozbieżny w punktach zbioru $\{z \in \mathbb{C} : |z - z_0| > R\}$.

Ciągłość sumy szeregu w kole zbieżności Suma $S(z)$ szeregu potęgowego jest funkcją ciągłą wewnątrz koła $\{z \in \mathbb{C} : |z| < R\}$, gdzie liczba R jest promieniem zbieżności. (Jeśli $\frac{1}{R} = 0$, to $S(\cdot)$ jest funkcją ciągłą na całej płaszczyźnie \mathbb{C} .)

Szereg możemy całkować i różniczkować wyraz po wyrazie jeśli dla pewnego x_0 jest on zbieżny oraz jeśli ciąg pochodnych jest jednostajnie zbieżny.

Twierdzenie 8: Różniczkowanie szeregu potęgowego wyraz po wyrazie

Założmy, że $R > 0$ jest promieniem zbieżności szeregu potęgowego

$$S(z) = \sum_{n=0}^{\infty} a_n z^n.$$

Wtedy funkcja S ma pochodną w każdym punkcie $z \in \{w \in \mathbb{C} : |w| < R\}$ i zachodzi wzór

$$S'(z) = \sum_{n=1}^{\infty} n a_n z^{n-1}.$$

Definicja 25: Maksimum lokalne

Jeśli istnieje otoczenie otwarte U punktu x_0 takie, że dla każdego $x \in U$

$$f(x) \leq f(x_0),$$

czyli nie występują w okolicy punktu x_0 wartości większe od $f(x_0)$ ani nieporównywalne, choć mogą występować wartości równe.

Definicja 26: Minimum lokalne

Jeśli istnieje otoczenie otwarte U punktu x_0 takie, że dla każdego $x \in U$

$$f(x) \geq f(x_0),$$

czyli nie występują w okolicy punktu x_0 wartości mniejsze od $f(x_0)$ ani nieporównywalne, choć mogą istnieć wartości równe.

Twierdzenie 9: Twierdzenie Fermata

Niech $f: [a, b] \rightarrow \mathbb{R}$ będzie funkcją ciągłą oraz różniczkowalną w przedziale (a, b) . Warunkiem koniecznym istnienia ekstremów lokalnych różniczkowalnych funkcji f w pewnym punkcie $x_0 \in (a, b)$ jest

$$f'(x_0) = 0.$$

Geometrycznie oznacza to, że styczna do wykresu funkcji jest w tym punkcie prostą poziomą.

Twierdzenie 10: Warunek konieczny istnienia ekstremum wielu zmiennych

Zerowanie się wszystkich pochodnych cząstkowych w punkcie.

Definicja 27: Hesjan

Niech $f: \mathbb{R}^n \rightarrow \mathbb{R}$ będzie funkcją biorącą jako argument wektor $x \in \mathbb{R}^n$ i produkującą jako wartość skalar $f(x) \in \mathbb{R}$. Jeśli wszystkie drugie pochodne f istnieją i są ciągłe na dziedzinie, wtedy hesjan H funkcji f jest macierzą kwadratową $n \times n$, zwykle zapisywaną:

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Hesjan jest macierzą symetryczną. Ciągłość drugich pochodnych implikuje, że kolejność różniczkowania nie ma znaczenia (tw. Schwarz’a).

Twierdzenie 11: Warunek dostateczny istnienia ekstremum wielu zmiennych

- 1. Jeśli hesjan jest dodatnio określony (równoważnie ma wszystkie wartości własne dodatnie) w a , to funkcja f ma minimum lokalne w a .
- 2. Jeśli hesjan jest ujemnie określony (równoważnie ma wszystkie wartości własne ujemne) w a , to funkcja f ma maksimum lokalne w a .

Definicja 28: Kontrakcja lub przekształcenie zwężające

[PS2] Niech (X, ϱ) będzie przestrzenią metryczną i niech $T: X \rightarrow X$. Mówimy, że T jest zwężające (albo inaczej: jest kontrakcją) wtedy i tylko wtedy, gdy istnieje stała $\lambda \in (0, 1)$ taka, że

$$\varrho(T(x), T(y)) \leq \lambda \varrho(x, y)$$

dla wszystkich $x, y \in X$.

Definicja 29: Punkt stały

[PS2] Punkt $x \in X$ nazywa się punktem stałym odwzorowania $T: X \rightarrow X$ wtedy i tylko wtedy, gdy $T(x) = x$.

Twierdzenie 12: Twierdzenie Banacha o punkcie stałym

[PS2] Jeśli (X, ϱ) jest przestrzenią metryczną zupełną, zaś $T: X \rightarrow X$ jest kontrakcją, to T ma dokładnie jeden punkt stały $x \in X$.

Note: dowód tego twierdzenia jest wyjątkowo przyjazny.

Najpierw sprawdzimy jednoznaczność znalezionej punktu stałego. Załóżmy, że istnieją dwa takie punkty. Mamy więc $T(x) = x$ oraz $T(y) = y$.

$$\varrho(x, y) = \varrho(T(x), T(y)) \leq \lambda \varrho(x, y)$$

Dla $\varrho(x, y) > 0$ nierówność jest spełniona dla $\lambda \geq 1$. Założyliśmy jednak, że $\lambda < 1$. To oznacza że $\varrho(x, y) = 0$. Z definicji metryki wiemy więc że $x = y$.

Pozostaje wykazać istnienie punktu stałego. Niech $x_0 \in X$ będzie dowolnym punktem. Rozpatrzmy zdefiniowany rekurencyjnie ciąg $x_{n+1} = T(x_n)$, gdzie $n = 0, 1, 2, \dots$. Ponieważ T jest kontrakcją, więc

$$\varrho(x_{n+1}, x_n) = \varrho(T(x_n), T(x_{n-1})) \leq \lambda \varrho(x_n, x_{n-1}) \leq \dots \leq \lambda^n \varrho(x_1, x_0)$$

dla pewnej liczby $\lambda \in (0, 1)$. Jeśli $m > n$, to na mocy nierówności trójkąta

$$\varrho(x_m, x_n) \leq \varrho(x_m, x_{m-1}) + \dots + \varrho(x_{n+1}, x_n) \leq \sum_{j=n}^{m-1} \lambda^j \varrho(x_1, x_0) \leq \sum_{j=n}^{\infty} \lambda^j \varrho(x_1, x_0) = \lambda^n \frac{\varrho(x_1, x_0)}{1 - \lambda} = C \lambda^n,$$

gdzie stała C nie zależy od n . Zatem ciąg (x_n) spełnia warunek Cauchy'ego, a więc jest zbieżny, gdyż przestrzeń (X, ϱ) jest zupełna. Niech $x = \lim_{n \rightarrow \infty} x_n$. Wobec ciągłości T ,

$$T(x) = T\left(\lim_{n \rightarrow \infty} x_n\right) = \lim_{n \rightarrow \infty} T(x_n) = \lim_{n \rightarrow \infty} x_{n+1} = x.$$

Twierdzenie 13: Twierdzenie o funkcji odwrotnej

Niech Ω będzie zbiorem otwartym w \mathbb{R}^n i $f \in C^1(\Omega, \mathbb{R}^n)$. Załóżmy, że dla pewnego $a \in \Omega$ różniczka $Df(a) \in L(\mathbb{R}^n, \mathbb{R}^n)$ jest izomorfizmem liniowym. Istnieją wówczas liczba $\delta > 0$ i zbiór otwarty $V \subset \mathbb{R}^n$ takie, że

1. $f: B(a, \delta) \rightarrow V$ jest bijekcją;
2. przekształcenie $g = f^{-1}: V \rightarrow B(a, \delta)$ jest klasy C^1 na V ;
3. jeśli $y = f(x)$ i $x \in B(a, \delta)$, to $Dg(y) = (Df(x))^{-1}$.

Twierdzenie 14: Twierdzenie o funkcji uwikłanej

Niech Ω będzie zbiorem otwartym w $\mathbb{R}^n \times \mathbb{R}^m$ i niech $(\mathbf{a}, \mathbf{b}) \in \Omega$. Załóżmy, że $F \in C^1(\Omega, \mathbb{R}^m)$ i $F(\mathbf{a}, \mathbf{b}) = 0$. Niech wreszcie

$$\det D_y F(\mathbf{a}, \mathbf{b}) \neq 0.$$

Istnieją wówczas zbiory otwarte $U \subset \mathbb{R}^n$ i $V \subset \mathbb{R}^m$ oraz funkcja $h \in C^1(U, \mathbb{R}^m)$ takie, że $\mathbf{a} \in U$, $\mathbf{b} \in V$, zaś warunek

$$F(\mathbf{x}, \mathbf{y}) = 0, \quad (\mathbf{x}, \mathbf{y}) \in U \times V \subset \Omega$$

zachodzi wtedy i tylko wtedy, gdy $\mathbf{y} = h(\mathbf{x})$ dla pewnego $\mathbf{x} \in U$. Ponadto,

$$Dh(\mathbf{x}) = -(D_y F(\mathbf{x}, h(\mathbf{x})))^{-1} \cdot D_x F(\mathbf{x}, h(\mathbf{x})).$$

Nazwa twierdzenia ma następujący sens: w małym otoczeniu takiego punktu $(\mathbf{a}, \mathbf{b}) \in \mathbb{R}^{n+m}$, w którym spełnione są założenia, równanie $F(\mathbf{x}, h(\mathbf{x})) = 0$ definiuje funkcję $\mathbf{y} = h(\mathbf{x})$ w sposób uwikłany.

Definicja 30: Rozmaitość różniczkowa

Zbiór $M \subset \mathbb{R}^{n+m}$ nazywamy zanurzoną rozmaitością n -wymiarową klasy C^1 wtedy i tylko wtedy, gdy dla każdego punktu $p \in M$ istnieje kula $B(p, r)$ w \mathbb{R}^{n+m} , n -wymiarowa podprzestrzeń liniowa $P = \text{span}\{e_1, e_2, \dots, e_n\} \subset \mathbb{R}^{n+m}$, zbiór U otwarty w \mathbb{R}^n i funkcja $\varphi \in C^1(U, P^\perp)$ takie, że

$$M \cap B(p, r) = \text{wykres } \varphi \cap B(p, r),$$

gdzie

$$\text{wykres } \varphi = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{n+m} = P \oplus P^\perp : \mathbf{x} \in U, \mathbf{y} = \varphi(\mathbf{x})\}.$$

Mówiąc krótko i potocznie, zanurzona rozmaitość n -wymiarowa klasy C^1 w \mathbb{R}^{n+m} to zbiór, który lokalnie, w otoczeniu każdego swojego punktu, jest wykresem pewnej funkcji klasy C^1 wybranych n zmiennych. Liczbę n nazywamy kowymiarem rozmaitości $M \subset \mathbb{R}^{n+m}$.

Definicja 31: Podział przedziału

Podziałem P przedziału $[a, b]$ nazwiemy każdy skończony ciąg punktów (x_0, x_1, \dots, x_n) taki, że

$$a = x_0 \leq x_1 \leq \dots \leq x_n = b.$$

Piszemy $\Delta x_i = x_i - x_{i-1}$, $i = 1, \dots, n$. Zbiór wszystkich podziałów odcinka $[a, b]$ będziemy oznaczać literą \mathcal{P} .

Definicja 32: Sumy całkowite Riemanna

Niech $f: [a, b] \rightarrow \mathbb{R}$ będzie funkcją ograniczoną, a $P = (x_0, x_1, \dots, x_n) \in \mathcal{P}$ - ustalonym podziałem $[a, b]$. Sumy

$$G(P, f) = \sum_{i=1}^n \sup_{[x_{i-1}, x_i]} f \cdot \Delta x_i, \quad D(P, f) = \sum_{i=1}^n \inf_{[x_{i-1}, x_i]} f \cdot \Delta x_i$$

nazywamy odpowiednio górną i dolną sumą Riemanna funkcji f dla podziału P .

Suma górna jest sumą pól prostokątów, których podstawy są odcinkami podziału P , a wysokości dobrane zostały tak, by suma prostokątów przykryła (jak najoszczędniej) cały zbiór $Z = \{(x, y) : 0 \leq y \leq f(x), x \in [a, b]\}$ punktów pod wykresem f . Obliczając zaś sumę dolną, dobieramy wysokości tak, aby jak najlepiej przybliżyć zbiór Z od wewnątrz.

Definicja 33: Górna i dolna całka Riemanna

Niech $f: [a, b] \rightarrow \mathbb{R}$ będzie funkcją ograniczoną. Liczby

$$\inf_{P \in \mathcal{P}} G(P, f), \quad \sup_{P \in \mathcal{P}} D(P, f)$$

nazywamy odpowiednio górną i dolną całką Riemanna funkcji f na odcinku $[a, b]$.

W swojej interpretacji geometrycznej na płaszczyźnie całka to operator przypisujący danej rzeczywistej funkcji ograniczonej określonej na przedziale (rzeczywistym) pewną liczbę rzeczywistą, którą można rozumieć jako pole powierzchni między jej wykresem a osią odciętych (pole zorientowane: jego znak zależy od znaku wartości funkcji).

Definicja 34: Całka nieoznaczona

Niech $f: P \rightarrow \mathbb{R}$, gdzie $P \subset \mathbb{R}$ jest dowolnym przedziałem. Całką nieoznaczoną funkcji f nazywamy rodzinę wszystkich funkcji pierwotnych funkcji f .

Definicja 35: Całka oznaczona

Niech $f: [a, b] \rightarrow \mathbb{R}$ będzie funkcją ciągłą. Całką oznaczoną funkcji f na przedziale $[a, b]$ nazywamy liczbę

$$\int_a^b f(x) dx = F(b) - F(a),$$

gdzie F jest dowolną funkcją pierwotną funkcji f .

Twierdzenie 15: Zasadnicze twierdzenie rachunku różniczkowego i całkowego

Załóżmy, że f jest funkcją całkowaną w sensie Riemanna na przedziale $[a, b]$. Wtedy funkcja $F: [a, b] \rightarrow \mathbb{R}$

$$F(x) = \int_a^x f(t) dt$$

jest ciągła na $[a, b]$. Dodatkowo jeśli f jest ciągła w x_0 to F jest różniczkowalna i $F'(x_0) = f(x_0)$.

Definicja 36: Objętość d -wymiarowej kostki

Niech d będzie ustaloną dodatnią liczbą całkowitą. d -wymiarową objętością d -wymiarowego przedziału $P = [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_d, b_d]$, gdzie $b_i \geq a_i$, nazywana jest liczba

$$\text{vol}_d(P) = (b_1 - a_1) \cdot (b_2 - a_2) \cdot \dots \cdot (b_d - a_d).$$

Definicja 37: Miara Lebesgue'a

Dla dowolnego zbioru $A \subseteq \mathbb{R}^d$ można skonstruować miarę zewnętrzną $\lambda^*(A)$ wyznaczoną przez funkcję vol_d , nazywaną miarą zewnętrzną Lebesgue'a:

$$\inf \left\{ \sum_{B \in \mathcal{C}} \text{vol}_d(B) : \mathcal{C} \text{ jest przeliczalnym zbiorem przedziałów, których suma pokrywa } A \right\}.$$

Definicja 38: Warunek Caratheodory'ego

Niech λ^* będzie miarą zewnętrzną na X . Powiemy, że $A \subset X$ spełnia warunek Caratheodory'ego jeżeli dla każdego zbioru $S \subseteq X$ zachodzi:

$$\lambda^*(S) = \lambda^*(S \cap A) + \lambda^*(S \setminus A).$$

Twierdzenie 16: Twierdzenie Caratheodory'ego

Niech λ^* będzie miarą zewnętrzną na X . Rodzina $\mathcal{F} \subset 2^X$ wszystkich zbiorów $A \subset X$, spełniających warunek Carathéodory'ego, jest σ -ciałem. Funkcja λ^* obcięta do tego σ -ciała

$$\lambda = \lambda^*|_{\mathcal{F}}: \mathcal{F} \rightarrow [0, \infty]$$

(tj. zbiorów spełniających warunek Caratheodory'ego) jest miarą.

Z twierdzenia Carathéodory'ego wynika, że λ^* obcięta do rodziny zbiorów spełniających warunek Carathéodory'ego jest miarą zupełną – miara ta nazywana jest **miarą Lebesgue'a** w przestrzeni \mathbb{R}^d .

Konstrukcja całki Lebesgue'a

1. Jedyną rozsądną możliwością przypisania wartości całce z funkcji charakterystycznej χ_A zbioru mierzalnego A jest miara tego zbioru:

$$\int \chi_A := \mu(A)$$

2. Jeżeli f jest nieujemną funkcją prostą (kombinacją liniową funkcji charakterystycznych), to całkę Lebesgue'a tej funkcji definiuje się wzorem

$$\int f = \int \sum_{i=1}^n a_i \chi_{A_i} := \sum_{i=1}^n a_i \int \chi_{A_i} = \sum_{i=1}^n a_i \mu_{A_i}$$

3. Całkę Lebesgue'a nieujemnej funkcji mierzalnej g określa się jako

$$\int g := \sup \left\{ \int f : f \text{ jest nieujemną funkcją prostą taką, że } f \leq g \right\}$$

4. Całkę z funkcji mierzalnej f na zbiorze mierzalnym $E \in \mathcal{F}$ określa się jako

$$\int_E f := \int f \chi_E$$

gdzie χ_E oznacza funkcję charakterystyczną zbioru E .

Definicja 39: Miara powierzchniowa

Zdefiniujemy m -wymiarową miarę powierzchniową na rozmaitości $M \subset \mathbb{R}^n$. Przypuśćmy, że dla pewnego zbioru otwartego $U \subset \mathbb{R}^n$ część wspólna $U \cap M$ jest równa $\psi(V)$, gdzie V jest zbiorem otwartym w \mathbb{R}^m . Zakładamy, że przekształcenie

$$\psi: \mathbb{R}^m \supset V \rightarrow \psi(V) = M \cap U \subset \mathbb{R}^n$$

jest homeomorfizmem na obraz, klasy C^1 , a jego różniczka $D\psi(x)$ ma maksymalny rząd m (tzn. jest monomorfizmem liniowym) w każdym punkcie $x \in V$. $M \cap U = \psi(V)$, gdzie parametryzacja ψ jest różnowartościowa i jej różniczka ma rząd m , miara powierzchniowa σ_m podzbiorów $M \cap U$ określona jest wzorem

$$\sigma_m(\psi(A)) = \int_A \sqrt{\det(D\psi(x)^T D\psi(x))} d\lambda_m(x) \quad \text{dla } A \subset V \in \mathbb{R}^m, A \in \mathcal{L}(\mathbb{R}^m).$$

Intuicja: W małej skali przekształcenie klasy C^1 jest z dobrym przybliżeniem liniowe. Dlatego miarę gładkiej, m -wymiarowej powierzchni przybliżamy za pomocą sumy m -wymiarowych objętości równoległoscianów, będących obrazami (pod działaniem różniczki parametryzacji) małych, m -wymiarowych kostek, zawartych w dziedzinie przekształcenia.

Twierdzenie 17: Fubini

Niech $f: \mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ będzie funkcją całkowalną (lub mierzalną w sensie Lebesgue'a i nieujemną). Wówczas: Dla λ_n -prawie wszystkich $x \in \mathbb{R}^n$ i λ_m -prawie wszystkich $y \in \mathbb{R}^m$ funkcje $f_x(y) := f(x, y)$ oraz $f_Y(x) := f(x, y)$ są mierzalne odpowiednio względem (\mathbb{R}^m) i (\mathbb{R}^n) ; Funkcje

$$\mathbb{R}^n \ni x \mapsto \int_{\mathbb{R}^m} f(x, y) d\lambda_m(y) \in \mathbb{R}, \quad \mathbb{R}^m \ni y \mapsto \int_{\mathbb{R}^n} f(x, y) d\lambda_n(x) \in \mathbb{R},$$

są mierzalne odpowiednio względem σ -ciał (\mathbb{R}^n) i (\mathbb{R}^m) ;

$$\begin{aligned} \int_{\mathbb{R}^{n+m}} f d\lambda_{n+m} &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^m} f(x, y) d\lambda_m(y) \right) d\lambda_n(x) \\ &= \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^n} f(x, y) d\lambda_n(x) \right) d\lambda_m(y). \end{aligned}$$

Całkowanie przez podstawianie dla funkcji wielu zmiennych

Niech $\Omega \subset \mathbb{R}^n$ będzie zbiorem otwartym, a $\Phi: \Omega \rightarrow \Phi(\Omega) \subset \mathbb{R}^n$ dyfeomorfizmem klasy C^1 zbioru Ω na $\Phi(\Omega)$. Załóżmy, że f jest funkcją całkowalną (lub mierzalną i nieujemną) względem miary Lebesgue'a λ_n na $\Phi(\Omega)$. Wtedy $(f \circ \Phi) \cdot |\det D\Phi|$ jest całkowalna (odpowiednio, mierzalna i nieujemna) na zbiorze Ω i zachodzi równość

$$\int_{\Phi(\Omega)} f d\lambda_n = \int_{\Omega} (f \circ \Phi) \cdot |\det D\Phi| d\lambda_n.$$

Twierdzenie 18: Lebesgue'a o zmajoryzowanym przejściu pod znakiem całki

Załóżmy, że funkcje $f_j, f: X \rightarrow \mathbb{R}, j = 1, 2, \dots$, są mierzalne i $|f_j| \leq g$, gdzie $g: X \rightarrow [0, \infty]$ jest funkcją całkowalną. Jeśli $f_j(x) \rightarrow f(x)$ prawie wszędzie w X , to

$$\lim_{j \rightarrow \infty} \int_X |f_j - f| d\mu = 0, \quad \lim_{j \rightarrow \infty} \int_X f_j d\mu = \int_X f d\mu.$$

Istotność wspólnej ograniczoności

Rozważmy odcinek $(0, 1)$ wyposażony w miarę Lebesgue'a λ . Dla liczby naturalnej $n \in \mathbb{N}$ zdefiniujmy funkcję $f_n: (0, 1) \rightarrow \mathbb{R}$ przez

$$f_n(x) = \begin{cases} n & \text{gdy } x \in \left(0, \frac{1}{n}\right), \\ 0 & \text{gdy } x \in \left(\frac{1}{n}, 1\right). \end{cases}$$

Wtedy $f_n(x) \rightarrow 0$ dla $x \in (0, 1)$, natomiast

$$\int f_n d\lambda = n\lambda\left(\left(0, \frac{1}{n}\right)\right) = n \cdot \frac{1}{n} = 1 \not\rightarrow 0 = \int 0 d\lambda.$$

A więc nie można pominąć założenia o wspólnym ograniczeniu tych funkcji.

Twierdzenie 19: Twierdzenie Greena

Załóżmy, że $\Omega \subset \mathbb{R}^2$ jest obszarem ograniczonym z brzegiem klasy C^1 . Niech $f, g \in C^1(U)$, gdzie $U \subset \mathbb{R}^2$ jest otwarty i $\bar{\Omega} \subset U$. Wreszcie, niech brzeg $\partial\Omega$ zbioru Ω ma naturalną orientację. Wówczas

$$\begin{aligned} \int_{\partial\Omega} g dy &= \int_{\Omega} \frac{\partial g}{\partial x} d\lambda_2, \\ - \int_{\partial\Omega} f dx &= \int_{\Omega} \frac{\partial f}{\partial y} d\lambda_2. \end{aligned}$$

3 Geometria z Algebrą Liniową

3.1 Układy równań liniowych, operacje na macierzach.

Definicja 40: Układ równań liniowych

Układ m równań liniowych o zmiennych x_1, x_2, \dots, x_n i o współczynnikach rzeczywistych, to ciąg m równań liniowych o współczynnikach rzeczywistych postaci

$$U : \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases} \quad (1)$$

Rozwiązanie powyższego układu równań liniowych to ciąg elementów (s_1, \dots, s_n) ze zbioru \mathbb{R} , który jest rozwiązaniem każdego z m równań liniowych tego układu.

Twierdzenie 20

W układzie równań dodanie do równania innego równania pomnożonego przez liczbę (1), przemnożenie równania przez liczbę (2) lub zamiana równań miejscami (3) nie zmieniają rozwiązań równania. Operacje (1), (2) i (3) nazywamy operacjami elementarnymi.

Układ równań jak w definicji 1 można przepisać jako macierz rozmiaru $m \times (n + 1)$

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right] \quad (\heartsuit)$$

i nazywamy ją macierzą (lub macierzą rozszerzoną) układu równań. Na macierzy, analogicznie jak w twierdzeniu 1, możemy zdefiniować operacje elementarne na wierszach i kolumnach macierzy.

Mówimy, że macierz A jest w postaci schodkowej jeśli:

- każdy wiersz zerowy znajduje się poniżej każdego wiersza niezerowego (czyli jeśli i -ty wiersz tej macierzy jest niezerowy, to j -ty wiersz jest zerowy tylko gdy $j > i$),
- dla każdego $i > 1$ pierwszy, licząc od lewej, niezerowy wyraz w i -tym wierszu znajduje się w kolumnie stojącej na prawo od pierwszego niezerowego wyrazu $(i-1)$ -szego wiersza.

Pierwsza macierz jest w postaci schodkowej, druga nie jest.

$$\left[\begin{array}{cccc} 2 & 1 & 3 & 7 \\ 0 & 2 & 0 & 4 \\ 0 & 0 & 0 & 0 \end{array} \right], \left[\begin{array}{cccc} 0 & 0 & 3 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right].$$

Mówimy, że macierz jest w zredukowanej postaci schodkowej (w postaci zredukowanej), jeśli jest w postaci schodkowej oraz w każdym niezerowym wierszu pierwszy niezerowy wyraz wynosi 1 i jest jedynym niezerowym wyrazem w swojej kolumnie.

Twierdzenie 21

Za pomocą operacji elementarnych (1) i (2) każdą macierz można sprowadzić do postaci schodkowej, a za pomocą operacji (1), (2) i (3) do postaci zredukowanej.

Metoda rozwiązywania układów równań liniowych nazywana metodą eliminacji Gaussa polega na sprowadzeniu macierzy (rozszerzonej) układu równań (♡) do postaci schodkowej zredukowanej za pomocą operacji elementarnych.

Aby znaleźć rozwiązanie układu równań U wystarczy macierz tego układu (♡) sprowadzić do zredukowanej postaci schodkowej. Jeśli otrzymana macierz nie zawiera wiersza postaci $0 \dots 0 1$ to można z niej odczytać rozwiązanie ogólne układu U . Jeżeli takowy wiersz istnieje to układ jest sprzeczny.

Twierdzenie 22

Jeśli (s_1, s_2, \dots, s_n) oraz $(s'_1, s'_2, \dots, s'_n)$ są rozwiązaniami układu równań

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases}, \tag{2}$$

to ciąg

$$(s_1 - s'_1, s_2 - s'_2, \dots, s_n - s'_n)$$

jest rozwiązaniem układu równań:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = 0 \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = 0 \end{cases}. \tag{3}$$

Twierdzenie 23

Założmy, że (s_1, \dots, s_n) jest rozwiązaniem układu równań (2). Wówczas każde rozwiązanie układu (2) jest postaci:

$$(s_1 + u_1, \dots, s_n + u_n),$$

gdzie (u_1, \dots, u_n) jest rozwiązaniem układu (3).

3.2 Ciała i liczby zespolone

Definicja 41: Ciało

Mówimy, że zbiór K zawierający co najmniej dwa elementy jest ciałem jeśli

- i) zadane są odwzorowania:

$K \times K \rightarrow K \quad (x, y) \mapsto x + y$ zwane działaniem dodawania

$K \times K \rightarrow K \quad (x, y) \mapsto x \cdot y$ zwane działaniem mnożenia
- ii) wyróżnione są dwa elementy zbioru K :

element zerowy, oznaczany 0 ,

element jedynekowy, oznaczany 1 ,
- iii) spełnione są następujące warunki, zwane aksjomatami ciała. Dla każdych $a, b, c \in K$

1. $a + (b + c) = (a + b) + c$ łączność dodawania

2. $a + b = b + a$ przemienność dodawania

3. $0 + a = a + 0 = a$ 0 jest elementem neutralnym dodawania

4. $\forall a \in K \exists p \in K \quad a + p = 0$ istnienie elementu odwrotnego w dodawaniu

5. $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ łączność mnożenia

6. $a \cdot b = b \cdot a$ przemienność mnożenia

7. $1 \cdot a = a \cdot 1 = a$ 1 jest elementem neutralnym mnożenia

8. $\forall a \neq 0 \in K \exists q \in K \quad a \cdot q = 1$ istnienie elementu odwrotnego w mnożeniu

9. $(a + b) \cdot c = a \cdot c + b \cdot c$ rozdzielność mnożenia względem dodawania.

Definicja 42: Podciało

Podzbiór $L \subset K$ nazywamy podciałem ciała K , jeśli $0, 1 \in L$ oraz dla każdych $a, b \in L$ zachodzi:
 $a + b \in L$; $a \cdot b \in L$; $-a \in L$; $a^{-1} \in L$ (gdy $a \neq 0$).

Twierdzenie 24: Podciało ciałem

Podciało ciała K (z działaniami takimi jak w K) jest ciałem.

Przykład 2: Różne ciała

- i) Zbiór liczb rzeczywistych \mathbb{R} (ze zwykłymi działaniami $+$ i \cdot oraz liczbami 0 i 1 jako elementami wyróżnionymi) jest ciałem.
- ii) Niech Q oznacza zbiór wszystkich liczb wymiernych i niech

$$Q(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$$

Wtedy \mathbb{Q} oraz $Q(\sqrt{2})$ są podciałami ciała \mathbb{R} , więc są ciałami.

- iii) Ciałem, zwanym \mathbb{Z}_p , jest dla liczby pierwszej p zbiór $\{0, 1, 2, \dots, p-1\}$ z działaniami zadanymi wzorami
 $(a, b) \mapsto (a + b) \bmod p$ dodawanie
 $(a, b) \mapsto (a \cdot b) \bmod p$ mnożenie
i z liczbami $0, 1$ jako elementami zerowymi i jedynkowymi.

Definicja 43: Ciało liczb zespolonych

Ciało liczb zespolonych, oznaczane przez \mathbb{C} , to ciało, którego elementami są wszystkie uporządkowane pary liczb rzeczywistych, w którym działania są określone wzorami

$$(a, b) + (c, d) = (a + c, b + d)$$

$$(a, b) \cdot (c, d) = (ac - bd, ad + bc)$$

a elementami neutralnymi są $(0, 0)$ dla dodawania oraz $(1, 0)$ dla mnożenia.
Dla każdego $a \in \mathbb{R}$ liczbę zespoloną $(a, 0)$ oznaczamy przez a , a liczbę $(0, 1)$ przez i

Definicja 44: Re i Im

Dla $z = a + bi$

- i) Liczbę a nazywamy częścią rzeczywistą liczby z i oznaczamy $\operatorname{Re} z$, liczbę b nazywamy częścią urojoną liczby z i oznaczamy $\operatorname{Im} z$.
- ii) Liczbę $a^2 + b^2$ nazywamy modułem liczby z i oznaczamy $|z|$.
- iii) Liczbę $a - bi$ nazywamy liczbą sprzężoną do z i oznaczamy \bar{z} .

Twierdzenie 25: Własności modułu i sprzężenia

Dla dowolnych $z, z + 1, z_2 \in \mathbb{C}$ mamy

- i) $z + \bar{z} = 2\operatorname{Re} z$, $z \cdot \bar{z} = |z|^2$

- ii) $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}, \overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$
- iii) $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$
- iv) $|z_1 + z_2| \leq |z_1| + |z_2|, ||z_1| - |z_2|| \leq |z_1 - z_2|$

Definicja 45: Postać trygonometryczna

Postacią trygonometryczną liczby zespolonej $z = a + bi$ nazywamy zapis dla $z \neq 0$ postaci

$$z = |z|(\cos(\phi) + i \sin(\phi))$$

gdzie $\phi \in \text{Arg } z$.

Twierdzenie 26: Mnożenie postaci trygonometrycznych

Zachodzą poniższe stwierdzenia:

- i) Dla $z_1 = |z_1|(\cos(\phi_1) + i \sin(\phi_1)), z_2 = |z_2|(\cos(\phi_2) + i \sin(\phi_2))$ zachodzi

$$z_1 z_2 = |z_1| |z_2| (\cos(\phi_1 + \phi_2) + i \sin(\phi_1 + \phi_2))$$

- ii) Dla dowolnych $z, w \in \mathbb{C}$ zachodzi $\text{Arg}(z \cdot w) = \text{Arg } z + \text{Arg } w$

- iii) (wzór de Moivre’a) $(|z|(\cos(\phi) + i \sin(\phi)))^n = |z|^n (\cos(n\phi) + i \sin(n\phi))$

Definicja 46: Pierwiastek z jedności

1. Pierwiastkiem stopnia n z jedności będziemy nazywać każdą liczbę zespoloną z taką, że $z^n = 1$.
2. Wszystkie pierwiastki stopnia n z jedności przedstawiają się jako $\omega^k = \cos\left(\frac{2k\pi}{n}\right) + i \sin\left(\frac{2k\pi}{n}\right)$ dla $k \in \{0, 1, \dots, n - 1\}$.
3. Pierwiastek stopnia n z jedności nazywamy pierwotnym jeśli nie jest pierwiastkiem z jedności stopnia mniejszego niż n .

Twierdzenie 27: Pierwiastkowanie zespolonych

Zachodzą dwa niezależne fakty:

- i) Pierwiastek $\omega^k = \cos\left(\frac{2k\pi}{n}\right) + i \sin\left(\frac{2k\pi}{n}\right)$ dla $1 \leq k \leq n$ stopnia n z jedności jest pierwotny wtedy i tylko wtedy, gdy k i n są względnie pierwsze.
- ii) Dla z w standardowej postaci zespolonej pierwiastkiem stopnia n z z są liczby

$$\sqrt[n]{|z|} \left(\cos\left(\frac{\phi + 2k\pi}{n}\right) + i \sin\left(\frac{\phi + 2k\pi}{n}\right) \right)$$

dla $k \in \{0, 1, \dots, n - 1\}$.

Definicja 47: Ciało algebraicznie domknięte

Jeśli każdy wielomian stopnia większego od zera o współczynnikach z ciała K ma w ciele K pierwiastek, to K nazywamy ciałem algebraicznie domkniętym.

Twierdzenie 28: Zasadnicze Twierdzenie Algebry

\mathbb{C} jest ciałem algebraicznie domkniętym.

3.3 Przestrzenie liniowe**Definicja 48: Przestrzeń liniowa nad ciałem K**

Przestrzenią liniową nad ciałem $(K, +, \cdot, 0, 1)$ nazywamy zbiór V , wraz z:

- odwzorowaniem: $\oplus : V \times V \rightarrow V$, zwanym dodawaniem wektorów,
- odwzorowaniem: $\otimes : K \times V \rightarrow V$, zwanym mnożeniem wektora przez skalar,
- wyróżnionym elementem Θ w V , zwanym wektorem zerowym,

przy czym spełnione są następujące aksjomaty przestrzeni liniowej:

1. $\alpha \oplus (\beta \oplus \gamma) = (\alpha \oplus \beta) \oplus \gamma \quad \forall \alpha, \beta, \gamma \in V$ (łączność dodawania wektorów),
2. $\alpha \oplus \beta = \beta \oplus \alpha \quad \forall \alpha, \beta \in V$ (przemienność dodawania wektorów),
3. $\alpha \oplus \Theta = \alpha \quad \forall \alpha \in V$ (Θ jest elementem neutralnym dla \oplus),
4. $\alpha \oplus \gamma = \Theta$ dla wszystkich $\forall \alpha \in V \quad \exists \gamma \in V$ (istnienie wektora przeciwnego),
5. $1 \otimes \alpha = \alpha \quad \forall \alpha \in V$ (mnożenie wektora przez 1),
6. $(a \cdot b) \otimes \alpha = a \otimes (b \otimes \alpha) \quad \forall \alpha, \beta, \gamma \in V$ (zgodność \cdot z \otimes),
7. $(a + b) \otimes \alpha = (a \otimes \alpha) \oplus (b \otimes \alpha) \quad \forall \alpha \in V, \forall a, b \in K$ (rozdzielność \otimes względem $+$),
8. $a \otimes (\alpha \oplus \beta) = (a \otimes \alpha) \oplus (a \otimes \beta) \quad \forall \alpha, \beta \in V, \forall a \in K$ (rozdzielność \otimes względem \oplus).

Elementy przestrzeni liniowej V nazywamy wektorami.

Przykłady przestrzeni liniowych:

- Przestrzeń liniowa macierzy $m \times n$ o wyrazach z ciała K - $M_{m \times n}(K)$.
- Przestrzeń liniowa wielomianów o współczynnikach w ciele K - $K[x]$.
- Przestrzeń liniowa ciągów nieskończonych o wyrazach z ciała K - K^∞ .

Definicja 49: Podprzestrzeń przestrzeni liniowej

Niepusty podzbiór $W \subset V$ nazywamy podprzestrzenią przestrzeni liniowej V jeśli dla każdego $\alpha, \beta \in W$ oraz każdego $a \in K$ zachodzi:

- i) $\alpha + \beta \in W$,
- ii) $a \cdot \alpha \in W$.

W każdej przestrzeni liniowej V podzbiór $\{0\}$, złożony tylko z wektora zerowego, jest jej podprzestrzenią. Nazywamy ją podprzestrzenią zerową. Mówimy, że przestrzeń liniowa V jest przestrzenią zerową, jeśli składa się tylko z wektora zerowego.

Przykłady podprzestrzeni liniowych:

- zbiór wielomianów nad ciałem K stopnia co najwyżej m jest podprzestrzenią przestrzeni $K[x]$.
- ciągi zbieżne, ciągi ograniczone, ciągi mające skończenie wiele niezerowych wyrazów są podprzestrzeniami przestrzeni K^∞ .

Definicja 50: Kombinacja liniowa

Niech V będzie przestrzenią liniową nad ciałem K . Kombinacją liniową układu wektorów $\alpha_1, \dots, \alpha_k$ o współczynnikach $a_1, \dots, a_k \in K$ nazywamy wektor:

$$\beta = a_1\alpha_1 + \dots + a_k\alpha_k = \sum_{i=1}^k a_i\alpha_i.$$

Definicja 51: Podprzestrzeń rozpięta na układzie wektorów

Niech V będzie przestrzenią liniową nad ciałem K i niech $\alpha_1, \dots, \alpha_k \in V$. Wówczas przez $\text{lin}(\alpha_1, \dots, \alpha_k)$ oznaczamy zbiór wszystkich kombinacji liniowych wektorów $\alpha_1, \dots, \alpha_k$.

Definicja 52: Układ wektorów rozpinający podprzestrzeń

Niech $\alpha_1, \dots, \alpha_k$ będzie układem wektorów w V . Wówczas podprzestrzeń liniową $\text{lin}(\alpha_1, \dots, \alpha_k)$ nazywamy przestrzenią rozpiętą na układzie $\alpha_1, \dots, \alpha_k$. Mówimy, że układ $\alpha_1, \dots, \alpha_k$ rozpina przestrzeń V , jeśli $V = \text{lin}(\alpha_1, \dots, \alpha_k)$, to znaczy każdy wektor z V jest kombinacją liniową wektorów $\alpha_1, \dots, \alpha_k$.

3.4 Liniowa zależność. Baza

4 Algebra

4.1 Podstawowe struktury algebraiczne: grupy, pierścienie, ciała

Definicja 53: Grupa

Zbiór G z określonym na nim dwuargumentowym działaniem nazywa się grupą, jeżeli ma on następujące własności

- $\forall x, y, z \in G \quad (x \circ y) \circ z = x \circ (y \circ z)$
- $\forall x \in G \quad x \circ 1 = 1 \circ x = x$
- $\forall x \in G \quad x \circ x^{-1} = x^{-1} \circ x = 1$

Definicja 54: Pierścień

Mówimy, że $(R, +, \cdot)$ jest **pierścieniem przemiennym z 1**, jeśli:

- $(R, +)$ jest grupa abelowa z elementem neutralnym 0
- \cdot jest działaniem łącznym w R z elementem neutralnym $1 \neq 0$
- $a(b + c) = ab + ac, (b + c)a = ba + ca$ dla $a, b, c \in R$.

w ogólnej definicji pierścienia grupa nie musi być abelowa, a do R nie musi należeć element 1. Na wykładzie zajmowaliśmy się tylko pierścieniami przemiennymi z 1.

Ciało jest specjalnym pierścieniem. Najpierw wprowadźmy pomocnicze definicje:

Definicja 55: Dzielnik zera

Dzielnik zera w pierścieniu R to taki element $a \in R$, że istnieje $0 \neq b \in R$, dla którego $ab = 0$.

Definicja 56: Dziedzina

Pierścień R , który nie zawiera żadnych dzielników zera, nazywany jest **dziedziną całkowitości** lub **dziedziną**.

Definicja 57: Ciało

Skończoną dziedzinę nazywamy **ciałem**. *To chyba nie jest prawda?*

Alternatywnie można powiedzieć, że ciało jest pierścieniem (przemiennym z 1 – dalej będziemy to pomijać i oznaczać literą R), w którym każdy element jest odwracalny.

Definicja 58: Element odwracalny

Element $a \in R$ jest **odwracalny**, jeśli istnieje taki $b \in R$, że $ab = 1$.

Definicja 59: Ciało po raz drugi

Ciałem nazywamy pierścień, w którym każdy niezerowy element jest odwracalny.

Alternatywnie można wypisać 9 aksjomatów ciała.

Intuicja: mamy jakiś zbiór elementów. Chcemy badać relacje między nimi, więc wprowadza-

my jakieś działanie dwuargumentowe. Np. mamy liczby, które chcemy do siebie dodawać (bez tego są trochę bezużyteczne). W skrócie: grupa - dodawanie/mnożenie, pierścień - i dodawanie, i mnożenie, ciało - pierścień, w którym możemy też dzielić. Żeby to wszystko miało sens, wprowadzamy kilka pomocniczych warunków, np. istnienie elementu neutralnego, odwrotnego.

4.2 Grupy, pierścienie, ciała - przykłady

Przykład 3: Grupy

1. Grupa \mathbb{Z} liczb całkowitych z dodawaniem - jest to przykład grupy **cyklicznej**, czyli generowanej przez 1 element: wszystkie elementy powstają przez dodanie (lub odejmowanie) 1 pewną liczbę razy.
2. Niech \mathbb{K} będzie ciałem. Symbolem $GL(n, \mathbb{K})$ oznaczamy grupę macierzy odwracalnych $n \times n$ o współczynnikach z \mathbb{K} z mnożeniem. Macierze o wyznaczniku 1 stanowią podgrupę oznaczaną symbolem $SL(n, K)$. Jest to przykład grupy nieabelowej: mnożenie nie jest przemienne ($AB \neq BA$).
3. Grupa dihedralna — podgrupa D_{2n} przekształceń zachowujących n -kąt foremny o środku symetrii w początku układu współrzędnych.

$$D_{2n} = \{1, \rho, \dots, \rho^{n-1}, \varepsilon, \rho\varepsilon, \dots, \rho^{n-1}\varepsilon\}$$

gdzie ρ jest obrotem o $\frac{1}{n}$ kąta pełnego, a ε symetrią osiową.

4. Niech X będzie zbiorem. Symbolem S_X oznaczamy grupę bijekcji zbioru X z działaniem składania jako mnożeniem i identycznością jako elementem neutralnym. Nazywamy ją grupą permutacji zbioru X . Jeżeli X jest zbiorem n -elementowym, to taką grupę oznaczamy symbolem S_n .

Przykład 4: Pierścienie

1. Pierścień \mathbb{Z} z operacjami dodawania (element neutralny: 0) i mnożenia (element neutralny: 1).
2. Niech R będzie pierścieniem. Przez $R[x]$ oznaczamy pierścień wielomianów o współczynnikach w pierścieniu R . Jego elementy mają postać

$$p = \sum_{k=0}^n a_k x^k$$

dla pewnych $n \in \mathbb{N}$, przy czym $x^0 = 1$. Działania: zwykłe dodawanie i mnożenie wielomianów, elementy neutralne: 0 i 1. Definicję można rozszerzyć na wielomiany wielu zmiennych, np. $R[x, y]$.

3. Pierścień szeregów formalnych $R[[x]]$, którego elementami są szeregi potęgowe (bez względu na to, czy są zbieżne). Dodawanie „po współrzędnych”, mnożenie – iloczyn Cauchy’ego:

$$\left(\sum_{i=0}^{\infty} a_i x^i \right) \cdot \left(\sum_{j=0}^{\infty} b_j x^j \right) = \sum_{k=0}^{\infty} \left(\sum_{l=0}^k a_l b_{k-l} \right) x^k.$$

4. Pierścień funkcji ciągłych z dodawaniem $(f + g)(x) = f(x) + g(x)$ i mnożeniem $(f \cdot g)(x) = f(x) \cdot g(x)$ dla każdego x .

Przykład 5: Ciała

- 1. Ciało \mathbb{R} liczb rzeczywistych, \mathbb{C} liczb zespolonych.
- 2. Ciało \mathbb{Q} liczb wymiernych jako (**ciało ułamków** liczb całkowitych): ciało to powstaje przez wprowadzenie relacji równoważności

$$(a,b) \sim (c,d) \iff ad = bc,$$

w tym przypadku dla liczb całkowitych a, b, c, d . Jej klasy abstrakcji oznaczamy przez $[a,b] = \frac{a}{b}$ (np. $\frac{2}{4} = \frac{1}{2}$ - to tylko skracanie ułamków, nic strasznego) wprowadzamy standardowe dodawanie i mnożenie ułamków i tak otrzymujemy ciało.

- 3. Ciało funkcji wymiernych – ciało ułamków, gdzie a, b, c, d przy powyższej notacji są wielomianami nad pierścieniem R . Dodawanie i mnożenie jak zawsze przy funkcjach wymiernych.
- 4. Ciała skończone \mathbb{F}_p dla $n, p \in \mathbb{N}$, gdzie p – liczba pierwsza (gdyby nie była pierwsza, jej dzielniki byłyby dzielnikami zera – nie może to być ciało). Dla $n = 1$ mamy po prostu $\mathbb{F}_p = \{0, 1, \dots, p - 1\}$. **Charakterystyką** ciała nazywamy liczbę p taką, że $1 + \dots + 1 = 0$ (dodawanie p razy). W przypadku ciał nieskończonych przyjmujemy, że $p = 0$.

Definicja 60: Homomorfizm grup

Przekształcenie $\varphi : G \rightarrow H$ spełniające

$$\forall_{g_1, g_2 \in G} \varphi(g_1 \circ g_2) = \varphi(g_1)\varphi(g_2)$$

Przykład 6: Homomorfizmów grup

- 1. $\varphi : \mathbb{Z} \rightarrow \mathbb{Z}_n$
- 2. Każdy element $g \in G$ wyznacza pewien automorfizm $\phi_g : G \rightarrow G$ zadany wzorem $\phi_g(x) = gxg^{-1}$. Nazywamy go **automorfizmem wewnętrznym** grupy G

Twierdzenie 29: Twierdzenie Lagrange’a

Jeżeli G jest grupą skończoną i $H \leq G$ to $|G| = |H| \cdot |G : H|$
Wniosek: rząd elementu jest dzielnikiem rzędu grupy. Każda grupa rzędu p (gdzie p jest liczbą pierwszą) jest izomorficzna z Z_p .

Twierdzenie 30: Twierdzenie Cauchy’ego

Jeżeli G jest grupą skończoną i liczba pierwsza p dzieli rząd grupy G , to w G istnieje element rzędu p .

Definicja 61: Działanie grupy na zbiorze

Działaniem grupy G na zbiorze X nazywamy homomorfizm $\phi : G \rightarrow S_X$. Równoważnie możemy definiować działanie grupy na zbiorze X jako funkcję dwuargumentową $G \times X \rightarrow X$ t.je $(g, x) \rightarrow g \cdot x$, która ma następujące własności:

- $(g \cdot h) \cdot x = g \cdot (h \cdot x)$
- $e \cdot x = x$

5 Wstęp do Informatyki

Definicja 62: Problem Algorytmiczny

PROBLEM ALGORYTMICZNY to formalnie zdefiniowane zadanie, które polega na znalezieniu rozwiązania pewnego zagadnienia przy pomocy kroków wykonywanych przez algorytm. Problem taki opisuje zestaw danych wejściowych, wymagania dotyczące rozwiązania oraz cel, który algorytm ma osiągnąć.

Kluczowe elementy problemu algorytmicznego:

- Dane wejściowe** – określony zbiór wartości, na których ma działać algorytm.
- Dane wyjściowe** – oczekiwane rezultaty przetworzenia danych wejściowych przez algorytm.
- Warunki i ograniczenia** – reguły, które muszą być spełnione podczas przetwarzania danych (np. czas wykonania, pamięć, wymagania co do poprawności wyniku).
- Cel** – jasne określenie, co ma być osiągnięte w wyniku działania algorytmu.

Przykładem problemu algorytmicznego jest np. problem sortowania zbioru liczb. Zestawem danych wejściowych jest nieuporządkowany zbiór liczb, a celem jest ich uporządkowanie w określonej kolejności (np. rosnącej).

Definicja 63: Funkcja rekurencyjna

FUNKCJA REKURENCYJNA to funkcja, która odwołuje się do samej siebie w celu rozwiązania podproblemu, będącego uproszczoną wersją problemu wyjściowego. Kluczowymi elementami funkcji rekurencyjnej są:

- Warunek bazowy (warunek stopu)** – określa sytuację, w której funkcja nie wywołuje już samej siebie i zwraca konkretny wynik.
- Rekurencyjne wywołanie** – przypadek, w którym funkcja wywołuje samą siebie z innymi argumentami, przybliżając się do warunku bazowego.

Rekurencja jest szczególnie użyteczna do rozwiązywania problemów, które można podzielić na mniejsze, podobne do siebie podproblemy. Przykładem jest opisana poniżej funkcja liczenia silni:

```
def silnia(n):  
    if n == 0:  
        return 1 # warunek bazowy  
    else:  
        return n * silnia(n - 1) # wywołanie rekurencyjne
```

Definicja 64: Dziel i Zwyciężaj

Metoda projektowania algorytmów w informatyce, prowadząca do bardzo efektywnych rozwiązań, polegająca na dzieleniu problemów rekurencyjnie na dwa lub więcej mniejszych

podproblemów tego samego (lub podobnego) typu, tak długo, aż fragmenty staną się wystarczająco proste do bezpośredniego rozwiązania. Z kolei rozwiązania otrzymane dla podproblemów scala się, uzyskując rozwiązanie całego zadania.

Przykłady algorytmów korzystające z tej metody:

- MergeSort
- QuickSort
- Wyszukiwanie Binarne

Definicja 65: Lista

LISTA to linearna struktura danych, która składa się z sekwencji elementów uporządkowanych według ich pozycji w liście. W odróżnieniu od tablic, lista może dynamicznie zmieniać swój rozmiar. Każdy element listy składa się z dwóch części:

- **Wartość** – przechowywana informacja.
- **Odnosińnik (wskaźnik)** – odniesienie do kolejnego elementu na liście (lub wartość NULL w przypadku ostatniego elementu).

Listy dzielą się na różne rodzaje, w zależności od sposobu organizacji elementów:

- LISTA JEDNOKIERUNKOWA – każdy element wskazuje na kolejny element, ale nie ma odniesienia do poprzedniego.
- LISTA DWUKIERUNKOWA – każdy element ma odniesienie zarówno do kolejnego, jak i poprzedniego elementu.
- LISTA CYKLICZNA – ostatni element listy wskazuje na pierwszy element, tworząc cykl.

Definicja 66: Stos

STOS to linearna struktura danych, w której elementy są dodawane i usuwane według zasady *LIFO* (ang. *Last In, First Out*), co oznacza, że ostatni element, który został dodany do stosu, jest pierwszym, który zostanie usunięty.

Stos składa się z dwóch podstawowych operacji:

- **push** – dodanie nowego elementu na szczyt stosu.
- **pop** – usunięcie elementu ze szczytu stosu.

Tak jak stos talerzy – aby wyjąć talerz, musimy zdjąć najpierw ten, który jest na samej górze.

Definicja 67: Kolejka

KOLEJKA to linearna struktura danych, która działa zgodnie z zasadą *FIFO* (ang. *First In, First Out*), co oznacza, że pierwszy element, który zostanie dodany, jest również pierwszym, który zostanie usunięty. Kolejkę można porównać do rzeczywistej kolejki, na przykład osób czekających w sklepie – osoba, która przychodzi jako pierwsza, zostaje obsłużona jako pierwsza.

Kolejka składa się z dwóch podstawowych operacji:

- **enqueue** – dodanie nowego elementu na koniec kolejki.
- **dequeue** – usunięcie elementu z początku kolejki.

Według tej samej reguły działają produkty spożywcze w sklepach.

Definicja 68: Drzewo binarnych wyszukiwań (BST)

DRZEWO BINARNYCH WYSZUKIWAŃ to struktura danych w formie drzewa binarnego, która spełnia następujące warunki:

- Każdy węzeł drzewa zawiera wartość, a także odniesienia do dwóch potomków: lewego i prawego.
- Wartość w lewym poddrzewie jest mniejsza od wartości w węźle głównym.
- Wartość w prawym poddrzewie jest większa od wartości w węźle głównym.
- Zarówno lewe, jak i prawe poddrzewo są również drzewami binarnymi wyszukiwania.

6 Topologia

Definicja 69: Przestrzeń topologiczna

Przestrzeń topologiczna to para (X, \mathcal{T}) , gdzie X jest zbiorem, a \mathcal{T} rodziną podzbiorów X , spełniającą następujące warunki:

- 1. $\emptyset, X \in \mathcal{T}$,
- 2. suma dowolnie wielu elementów \mathcal{T} należy do \mathcal{T} ,
- 3. przecięcie skończenie wielu elementów \mathcal{T} należy do \mathcal{T} .

Definicja 70: Metryka

Metryką na zbiorze X nazywa się funkcję $d : X \times X \rightarrow \mathbb{R}$ spełniającą następujące warunki:

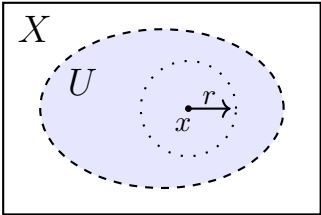
- 1. $d(x, y) = 0 \iff x = y$,
- 2. $\forall x, y \in X \quad d(x, y) = d(y, x)$,
- 3. $\forall x, y, z \in X \quad d(x, y) \leq d(x, z) + d(z, y)$.

Parę (X, d) nazywamy przestrzenią metryczną.

Definicja 71: Zbiór otwarty w przestrzeni metrycznej

Niech (X, d) będzie przestrzenią metryczną. Zbiór $U \subset X$ jest otwarty, jeśli dla każdego x istnieje promień t.ż. kula o promieniu r i środku x zawiera się w U

$$\forall_{x \in U} \quad \exists_{r > 0} \quad B(x, r) \subset U$$



Definicja 72: Topologia w przestrzeni metrycznej

Rodzinę $\mathcal{T}(d)$ wszystkich zbiorów otwartych w (X, d) nazywamy topologią tej przestrzeni metrycznej albo topologią generowaną przez metrykę d .

Definicja 73: Przestrzeń metryzowalna

Jeśli dla przestrzeni (X, \mathcal{T}) można określić metrykę d na X taką, że $\mathcal{T} = \mathcal{T}(d)$, to przestrzeń topologiczna (X, \mathcal{T}) jest metryzowalna.

Czy każda przestrzeń topologiczna jest metryzowalna?
Rozważmy topologię antydyskretną, czyli taką do której należy tylko cała przestrzeń i zbiór pusty: $\mathcal{T}_a = \{\emptyset, X\}$.
Topologia antydyskretna, jeżeli zawiera conajmniej dwa punkty, jest niemetryzowalna. Wiemy, że dopełnienie $X \setminus F$ zbioru skończonego F w przestrzeni metrycznej (X, d) jest otwarte. Załóżmy więc że ta topologia jest metryzowalna. Wówczas, dla dowolnego $x \in X$, $X \setminus \{x\}$ powinno być zbiorem otwartym, ale $X \setminus \{x\} \notin \mathcal{T}_a$.

Definicja 74: Ciągłość w przestrzeniach metrycznych

- (1) $\forall \varepsilon > 0 \exists \delta > 0 \forall x \in X \quad d_X(a, x) < \delta \implies d_Y(f(a), f(x)) < \varepsilon.$
- (2) $f(B_X(a, \delta)) \subseteq B_Y(f(a), \varepsilon)$
- (3) $B_X(a, \delta) \subseteq f^{-1}(B_Y(f(a), \varepsilon))$

Ciągłość przekształcenia $f : X \rightarrow Y$ przestrzeni metrycznych (X, d_X) w przestrzeni metryczną (Y, d_Y) jest równoważna warunkowi, że jeśli $x_n \rightarrow x_0$, to $f(x_n) \rightarrow f(x_0)$.

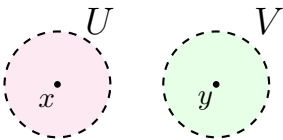
Definicja 75: Ciągłość w przestrzeniach topologicznych

Przekształcenie $f : X \rightarrow Y$ przestrzeni topologicznej (X, \mathcal{T}_X) w (Y, \mathcal{T}_Y) jest ciągłe, jeśli dla każdego $U \in \mathcal{T}_Y$, $f^{-1}(U) \in \mathcal{T}_X$.

Definicja 76: Przestrzeń zwarta

Przestrzeń topologiczna (X, \mathcal{T}) jest zwarta, jeśli jest przestrzenią Hausdorffa i z każdego otwartego pokrycia tej przestrzeni można wybrać pokrycie skończone.

Przestrzeń Hausdorffa: taka przestrzeń w której dla dowolnych dwóch różnych punktów x, y można wskazać ich rozłączne otoczenia U, V tzn. takie rozłączne zbiory otwarte tej przestrzeni, które spełniałyby $x \in U, y \in V$.



Twierdzenie 31: Warunki zwartości dla przestrzeni metryzowalnych

Dla przestrzeni metryzowalnej (X, \mathcal{T}) następujące warunki są równoważne:

- (1) Z każdego otwartego pokrycia przestrzeni X można wybrać pokrycie skończone,
- (2) Z każdego ciągu punktów w X można wybrać podciąg zbieżny w tej przestrzeni,
- (3) każdy zstępujący ciąg niepustych zbiorów domkniętych w X ma niepuste przecięcie.

Definicja 77: Ciąg Cauchy’ego

Mówimy, że ciąg (x_n) w przestrzeni metrycznej (X, d) jest ciągiem Cauchy’ego jeśli

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n, m > n_0 \quad d(x_n, x_m) < \varepsilon.$$

Definicja 78: Przestrzeń zupełna

Przestrzeń metryczna (X, d) nazywa się zupełna wtedy i tylko wtedy, gdy każdy ciąg Cauchy’ego $(x_n) \subset X$ jest zbieżny w X .

Przykłady przestrzeni zupełnych:

- Przestrzenie euklidesowe (\mathbb{R}^n, d_e) ,
- Domknięte podzbiory przestrzeni zupełnych.

Kwadrat leksykograficzny jest przestrzenią zwartą, która nie jest metryzowalna.

Definicja 79: Zbiór brzegowy

Zbiór A w przestrzeni topologicznej (X, \mathcal{T}) jest brzegowy, jeśli ma puste wnętrze.

Twierdzenie 32: Twierdzenie Baire'a

W przestrzeni metrycznej zupełnej (X, d) , przeliczalna suma domkniętych zbiorów brzegowych jest zbiorem brzegowym.

Istotność założeń w twierdzeniu Baire'a:

- **domkniętość:** samo założenie, że zbiory w są brzegowe nie wystarcza: prosta euklidesowa jest sumą dwóch zbiorów brzegowych – zbioru liczb wymiernych i zbioru liczb niewymiernych.
- **zupełność:** rozpatrzmy przestrzeń liczb wymiernych (\mathbb{Q}, d_e) z metryką euklidesową $d_e(x, y) = |x - y|$ i ustawmy liczby wymierne w ciąg q_1, q_2, \dots . Zbiory $A_i = \{q_i\}$ są domknięte i brzegowe w (\mathbb{Q}, d_e) , ale ich suma $\bigcup_{i=1}^{\infty} A_i = \mathbb{Q}$ nie jest brzegowa w przestrzeni \mathbb{Q} .

Twierdzenie 33: Warunek Cantora

Przestrzeń (X, d) jest zupełna wtedy i tylko wtedy, gdy spełnia następujący warunek Cantora: każdy ciąg niepustych zbiorów domkniętych o średnicach dążących do zera ma niepuste przecięcie.

Zupełność wynikająca ze zwartości

Ponieważ warunek Cantora jest słabszy niż (3) warunek zwartości dla przestrzeni metryzowalnych, więc dla każdej metryki d generującej topologię przestrzeni zwartej (X, \mathcal{T}) , przestrzeń metryczna (X, d) jest zupełna.

Definicja 80: Całkowita ograniczoność

Zbiór w przestrzeni metrycznej (X, d) jest całkowicie ograniczony, jeśli dla każdego $\varepsilon > 0$ można go pokryć skończenie wieloma zbiorami o średnicach $\leq \varepsilon$.

Twierdzenie 34: Zupełność + całkowita ograniczoność = zwartość

Przestrzeń metryczna (X, d) jest zwarta wtedy i tylko wtedy, gdy jest zupełna i całkowicie ograniczona.

Definicja 81: Spójność

Przestrzeń topologiczna (X, \mathcal{T}) jest spójna, jeśli zbioru X nie można rozłożyć na sumę dwóch rozłącznych, niepustych zbiorów domkniętych (równoważnie - otwartych). Zbiór $S \subset X$ jest spójny, jeśli podprzestrzeń (S, \mathcal{T}_S) przestrzeni (X, \mathcal{T}) jest spójna.

Twierdzenie 35: Spójność

Zbiór S w przestrzeni jest spójny \iff dla każdych niepustych zbiorów A, B takich, że $S = A \cup B$, mamy $A \cap B \neq \emptyset$ lub $A \cap B \neq \emptyset$.

Przekształcenia ciągłe przeprowadzają zbiory spójne na zbiory spójne.

Definicja 82: Droga

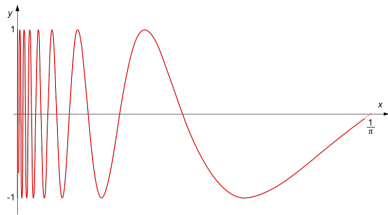
W przestrzeni topologicznej (X, \mathcal{T}) drogą łączącą punkty $a, b \in X$ nazywamy przekształcenie ciągłe $f : [0, 1] \rightarrow X$ takie, że $f(0) = a$, $f(1) = b$.

Definicja 83: Łukowa spójność

Przestrzeń topologiczna (X, \mathcal{T}) jest łukowo spójna, jeśli każdą parę punktów z X można połączyć drogą.

Łukowa spójność implikuje spójność, ale nie na odwrót.

Sinusoida warszawska jest spójna, ale nie łukowo spójna. Jest to zbiór $T = \{(t, \sin \frac{1}{t}) : t \in (0, 1]\} \cup \{0\} \times [-1, 1]$. $S = \{(t, \sin \frac{1}{t}) : t \in (0, 1]\}$ jest obrazem przedziału $[0, 1]$ przy przekształceniu ciągłym $t \rightarrow (t, \sin \frac{1}{t})$, jest więc zbiorem spójnym, oraz $T = \overline{S}$.



Definicja 84: Homeomorfizm

Homeomorfizmem nazywamy przekształcenie $f : X \rightarrow Y$ przestrzeni topologicznej (X, \mathcal{T}_X) w (Y, \mathcal{T}_Y) , które jest różnowartościowe, $f(X) = Y$ oraz takie, że f i f^{-1} są ciągłe.

Definicja 85: Przestrzenie topologiczne homeomorficzne

Przestrzenie (X, \mathcal{T}_X) i (Y, \mathcal{T}_Y) nazwiemy homeomorficznymi, jeśli istnieje homeomorfizm $f : X \rightarrow Y$.

Założenie ciągłości funkcji odwrotnej w powyższej definicji jest konieczne, ponieważ istnieją nieciągłe funkcje odwrotne do ciągłych bijekcji. Niech S^1 będzie okręgiem jednostkowym oraz niech $f : [0, 2\pi) \rightarrow S^1$ będzie funkcją daną wzorem $f(\varphi) = (\cos \varphi, \sin \varphi)$ ($\varphi \in [0, 2\pi)$). Funkcja f jest ciągła i bijektywna. Jednak jej funkcja odwrotna nie jest ciągła w punkcie $(1, 0)$, gdyż $f^{-1}((1, 0)) = 0$, ale obraz żadnego otwartego łuku otaczającego punkt $(1, 0)$ nie jest zawarty w otoczeniu $[0, \frac{1}{2})$ punktu $\varphi = 0$.

Twierdzenie 36: Zwartość przestrzeni i ciągła bijekcja a homeomorfizm

Ciągłe i różnowartościowe przekształcenie przestrzeni zwartej na przestrzeń Hausdorffa jest homeomorfizmem.

7 Równania Różniczkowe Zwyczajne

Twierdzenie 37: Picarda - Lindelofa

Niech funkcja $f(t, x) : \mathbb{R}^{m+1} \rightarrow \mathbb{R}^m$ będzie ciągła w zbiorze $Q = \{(t, x) : |t - t_0| \leq a, |x - x_0| \leq b\}$. Zakładamy ponadto, że $\sup_{(t,x) \in Q} |f(t, x)| = M$ oraz f spełnia w zbiorze Q warunek Lipschitza względem zmiennej x

$$|f(t, x_1) - f(t, x_2)| \leq L|x_1 - x_2|,$$

dla pewnej stałej L . Wtedy zagadnienie Cauchy’ego

$$\dot{x} = f(t, x), \quad x(t_0) = x_0,$$

ma jednoznaczne rozwiązanie na przedziale $|t - t_0| \leq \alpha$, gdzie $\alpha < \min(a, \frac{b}{M}, \frac{1}{L})$.

Twierdzenie 38: Peano

Niech funkcja $f(t, x) : \mathbb{R}^{m+1} \rightarrow \mathbb{R}^m$ będzie ciągła w zbiorze $Q = \{(t, x) : t \in [t_0, t_0 + a], |x - x_0| \leq b\}$. Zakładamy także, że $\sup_{(t,x) \in Q} |f(t, x)| = M$. Wtedy zagadnienie Cauchy’ego

$$\dot{x} = f(t, x), \quad x(t_0) = x_0$$

ma rozwiązanie na przedziale $[t_0, t_0 + \alpha]$, gdzie $\alpha = \min(a, \frac{b}{M})$.

Definicja 86: Rozwiązania stabilne w sensie Lapunowa

Niech dany będzie układ równań autonomicznych z funkcją $f : Q \rightarrow \mathbb{R}^m$, gdzie Q jest otwartym zbiorem w \mathbb{R}^m , a f jest funkcją klasy C^1 . Niech $\bar{x}(t)$ będzie rozwiązaniem jednego z tych układów w przedziale $[0, +\infty)$. Mówimy, że rozwiązanie $\bar{x}(t)$ jest **stabilne w sensie Lapunowa** dla $t \rightarrow +\infty$, jeśli dla każdego $\epsilon > 0$ istnieje takie $t_0 \geq 0$ oraz $\eta > 0$, że każde rozwiązanie $x(t)$, takie że

$$|x(t_0) - \bar{x}(t_0)| < \eta,$$

spełnia dla $t > t_0$ warunek

$$|x(t) - \bar{x}(t)| < \epsilon.$$

Jeśli dodatkowo

$$\lim_{t \rightarrow +\infty} |x(t) - \bar{x}(t)| = 0,$$

to mówimy, że rozwiązanie $\bar{x}(t)$ jest **asymptotycznie stabilne**.

Definicja 87: Całka pierwsza

Niech będzie dane równanie

$$\dot{x} = f(t, x)$$

z prawą stroną ciągłą w pewnym zbiorze otwartym $Q \subset \mathbb{R}^{m+1}$. Funkcję $U(t, x)$ określoną w otwartym zbiorze $Q_0 \subset Q$ nazywamy całką pierwszą równania+, jeśli jest ona stała na każdej krzywej całkowej tego równania.

Definicja 88: Równaniem różniczkowe zwyczajne rzędu n

$$F\left(t, x, \dot{x}, \ddot{x}, \dots, x^{(n)}\right) = 0$$

W równaniu tym t jest zmienną niezależną, a x zmienną zależną. Zmienną zależną $x(t)$ oraz funkcje F traktujemy jako funkcje wektorowe o wartościach w przestrzeni \mathbb{R}^m . Rozwiąza-

niem równania nazywamy funkcje $\varphi(t)$ klasy C^n , która podstawiona do równania w miejsce x , odpowiednio φ' w miejsce \dot{x} , ..., $\varphi^{(n)}$ w miejsce $x^{(n)}$, zmienia to równanie w tożsamość.

Definicja 89: Krzywa całkowa

Wykres funkcji $\varphi(t)$ w przestrzeni \mathbb{R}^{m+1} zmiennych (t, x) nazywamy krzywą całkową równania

Rozważmy ponownie równanie autonomiczne

$$\dot{x} = f(x),$$

którego prawa strona jest funkcją klasy C^1 w pewnym zbiorze otwartym $Q \subset \mathbb{R}^m$. Z założenia tego wynika, że równanie (5.7) uzupełnione warunkiem początkowym

$$x(0) = p, \quad p \in Q,$$

ma jednoznaczne rozwiązanie w pewnym przedziale $(-a, a)$.

Niech funkcja $\varphi(t; p)$, jako funkcja argumentu t , będzie rozwiązaniem zagadnienia początkowego, określonym w zbiorze otwartym $\Omega \subset (-a, a) \times Q$. Funkcja ta spełnia warunki:

1. $\varphi(0; p) = p$,
2. $\varphi(t; p)$ jest ciągła na Ω ,
3. $\varphi(t + \tau; p) = \varphi(t; \varphi(\tau; p))$ na Ω .

Rozważmy teraz przestrzeń fazową M równania i niech $f(x)$ będzie funkcją klasy C^1 na M . Wtedy dla każdego warunku początkowego $p \in M$ mamy rozwiązanie $\varphi(t; p) \subset M$. Załóżmy w dalszym ciągu, że rozwiązanie $\varphi(t; p)$ może być przedłużone na całą prostą $(-\infty, \infty)$ z zachowaniem warunku $\varphi(t; p) \subset M$. Z założeń na temat funkcji $f(x)$ oraz z tw. Picarda wynika, że istnieje rodzina dyfeomorfizmów parametryzowana zmienną t

$$g^t : M \rightarrow M,$$

wyznaczona równością $g^t(p) = \varphi(t; p)$.

Definicja 90: Potok

Para (M, g^t) , gdzie M jest przestrzenią fazową, a g^t , $t \in (-\infty, \infty)$, jest rodziną dyfeomorfizmów M , spełniającą warunki:

1. $g^t : M \rightarrow M$,
2. $g^0 = \text{id}$ oraz $g^{t+s} = g^t \circ g^s$ są różniczkowalnymi przekształceniami M w M ,
3. $(g^t)^{-1} = g^{-t}$.

Definicja 91: Trajektoria/ orbita punktu

w potoku (M, g^t) nazywamy zbiór wartości odwzorowania $g^t(p)$, $t \in \mathbb{R}$, tzn. zbiór

$$\{g^t(p) : t \in \mathbb{R}\}.$$

Portrety fazowe

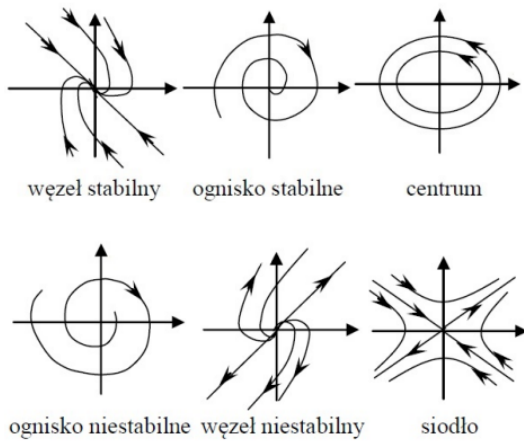
Klasyfikację punktów krytycznych układów autonomicznych ograniczymy do badania dwuwymiarowych układów liniowych o stałych współczynnikach

$$\dot{x} = Ax$$

z macierzą

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Jak łatwo zauważyć, punkt $x = 0$ jest punktem krytycznym równania. Jeśli macierz A jest nieosobliwa ($\det A \neq 0$), to układ nazywa się **prostym**.



Analizę zachowania się rozwiązań układu w otoczeniu punktu krytycznego rozpoczniemy od przypadku układów prostych. W celu znalezienia rozwiązań rozpatrzmy wielomian charakterystyczny macierzy A

$$p(\lambda) = \lambda^2 - (\operatorname{tr} A)\lambda + \det A = 0,$$

gdzie $\operatorname{tr} A = a_{11} + a_{22}$ jest śladem macierzy A . Znajdujemy pierwiastki wielomianu charakterystycznego

$$\lambda_1 = \frac{1}{2} \left(\operatorname{tr} A + \sqrt{\Delta} \right), \quad \lambda_2 = \frac{1}{2} \left(\operatorname{tr} A - \sqrt{\Delta} \right),$$

gdzie $\Delta = (\operatorname{tr} A)^2 - 4 \det A$. Znajomość pierwiastków wielomianu charakterystycznego pozwala znaleźć postać kanoniczną macierzy A . W zależności od rozwiązania układu dostajemy różne portrety fazowe.

8 Rachunek Prawdopodobieństwa

Punkt pierwszy

Model doświadczenia losowego

Matematyczny model doświadczenia losowego to trójka:

$$(\Omega, \mathcal{F}, \mathbb{P}),$$

gdzie \mathcal{F} to σ -ciało podzbiorów zdarzeń elementarnych Ω , spełniające:

$$\emptyset \in \mathcal{F}$$

$$A \in \mathcal{F} \Rightarrow A' \in \mathcal{F}$$

$$A_1, A_2, \dots \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$$

Twierdzenie 39: Aksjomaty prawdopodobieństwa

$$\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$$

$$\mathbb{P}(\Omega) = 1$$

dla parami rozłącznych A_1, A_2, \dots

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

Własności prawdopodobieństwa

Dla każdego $A \in \mathcal{F}$

1. $\mathbb{P}(\emptyset) = 0$
2. $\mathbb{P}(A') = 1 - \mathbb{P}(A)$
3. $A \subseteq B \Rightarrow \mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$
4. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
5. $\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mathbb{P}(A_i)$
6. $\mathbb{P}(A) \leq 1$.

Definicja 92: Prawdopodobieństwo klasyczne

Założmy, że Ω jest zbiorem skończonym, $\mathcal{F} = 2^{\Omega}$ i wszystkie zdarzenia jednoelementowe są jednakowo prawdopodobne. Wówczas każdy zbiór jednoelementowy ma prawdopodobieństwo $\frac{1}{|\Omega|}$, a w konsekwencji, dla dowolnego $A \in \mathcal{F}$

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|}$$

Definicja 93: Prawdopodobieństwo geometryczne

Założmy, że

$$\Omega \in \mathcal{B}(\mathbb{R}^d), 0 < |\Omega| < \infty$$

(tu $|\cdot|$ oznacza miarę Lebesgue'a w \mathbb{R}^d)

$$\mathcal{F} = \mathcal{B}(\Omega)$$

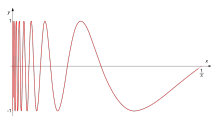
Miara probabilistyczna \mathbb{P} będzie zadana przez

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|}$$

Wówczas trójka $(\Omega, \mathcal{F}, \mathbb{P})$ jest przestrzenią probabilistyczną. Przestrzeń tę wykorzystujemy do modelowania doświadczenia polegającego na losowaniu punktu ze zbioru Ω .

8.0.1 Paradoks Bertranda

Z okręgu o promieniu 1 wylosowano cięciwę AB. Jakie jest prawdopodobieństwo tego, że będzie ona dłuższa niż bok trójkąta równobocznego wpisanego w ten okrąg?



1) Wylosowanie cięciwy możemy utożsamić z wylosowaniem miary kąta środkowego $\alpha = \angle AOB \in [0, 2\pi)$. $\Omega = [0, 2\pi)$, $\mathcal{F} = \mathcal{B}(\Omega)$. Cięciwa spełnia warunki zadania wtedy i tylko wtedy, gdy $\alpha \in (2\pi/3, 4\pi/3)$, a zatem szukane prawdopodobieństwo wynosi

$$\mathbb{P}((2\pi/3, 4\pi/3)) = \frac{|(2\pi/3, 4\pi/3)|}{|[0, 2\pi)|} = \frac{1}{3}$$

2) Wylosowanie cięciwy można utożsamić z wylosowaniem jej środka. $\Omega = B(0, 1)$, $\mathcal{F} = \mathcal{B}(\Omega)$. Cięciwa będzie spełniała zadane warunki wtedy i tylko wtedy, gdy jej środek będzie leżał wewnątrz koła o promieniu $\frac{1}{2}$ współśrodkowego z danym okręgiem, zatem szukane prawdopodobieństwo wynosi

$$\mathbb{P}\left(\left[0, \frac{1}{2}\right)\right) = \frac{|B(0, \frac{1}{2})|}{|B(0, 1)|} = \frac{1}{4}$$

3) Bierzemy pod uwagę położenie środka cięciwy, lecz tym razem patrzymy na jego odległość od środka okręgu. $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}(\Omega)$. Cięciwa będzie spełniała warunki zadania jeśli jej środek będzie odległy od środka okręgu o mniej niż $\frac{1}{2}$. Zatem szukane prawdopodobieństwo wynosi

$$\mathbb{P}\left(\left[0, \frac{1}{2}\right)\right) = \frac{|[0, \frac{1}{2}]|}{|[0, 1]|} = \frac{1}{2}$$

Punkt drugi**Definicja 94: Prawdopodobieństwo warunkowe**

Założmy, że $(\Omega, \mathcal{F}, \mathbb{P})$ jest przestrzenią probabilistyczną oraz $A, B \in \mathcal{F}$. Jeśli $\mathbb{P}(B) > 0$, to definiujemy prawdopodobieństwo zajścia zdarzenia A pod warunkiem zdarzenia B :

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

Przykład 7

Rzucamy dwukrotnie kostką, wiemy że suma oczek jest równa 8. Jakie jest prawdopodobieństwo, że w pierwszym rzucie kostką uzyskaliśmy dwójkę lub trójkę? Niech A_2 - na pierwszym miejscu dwójka, A_3 - na pierwszym miejscu trójka, B - suma 8.

$$\begin{aligned}\mathbb{P}(A_2 \cap B) &= \mathbb{P}(A_3 \cap B) = \frac{1}{36} & \mathbb{P}(B) &= \frac{5}{36} \\ \mathbb{P}(A_2 \cup A_3|B) &= \frac{2}{5}\end{aligned}$$

Definicja 95: Prawdopodobieństwo całkowite

Niech $(\Omega, \mathcal{F}, \mathbb{P})$ będzie ustaloną przestrzenią probabilistyczną oraz załóżmy, że $(B_i)_{i \in I}$ jest rozbiciem*. Wówczas dla dowolnego zdarzenia A zachodzi wzór:

$$\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(A|B_i)\mathbb{P}(B_i)$$

*Rozbiciem nazywamy dowolną co najwyżej przeliczalną rodzinę $(B_i)_{i \in I}$ parami rozłącznych zdarzeń, której sumą jest zbiór Ω .

Przykład 8

W urnie I znajdują się dwie kule białe i jedna czarna. Losujemy kulę i odkładamy. Następnie losujemy drugą kulę. Jakie jest prawdopodobieństwo tego, że kula wyciągnięta w drugim losowaniu jest biała?
Rozwiązanie: H_1 -wylosowano kulę białą, H_2 - wylosowano kulę czarną, $H_1 \cap H_2 = \emptyset$ oraz $H_1 \cup H_2 = \Omega$. Oznaczmy interesujące nas zdarzenie przez A .

$$\mathbb{P}(A) = \mathbb{P}(A|H_1)\mathbb{P}(H_1) + \mathbb{P}(A|H_2)\mathbb{P}(H_2)$$

Z warunków zadania wynika wprost, że $\mathbb{P}(H_1) = \frac{2}{3}$, $\mathbb{P}(H_2) = \frac{1}{3}$, $\mathbb{P}(A|H_1) = \frac{1}{2}$, $\mathbb{P}(A|H_2) = 1$. Wobec tego, $\mathbb{P}(A) = \frac{2}{3}$.

Twierdzenie 40: Wzór Bayesa

dla $j \in I$

$$\mathbb{P}(B_j|A) = \frac{\mathbb{P}(A|B_j)\mathbb{P}(B_j)}{\sum_{i \in I} \mathbb{P}(A|B_i)\mathbb{P}(B_i)}$$

o ile $\mathbb{P}(A) > 0$

Punkt trzeci

Definicja 96: Niezależność zdarzeń

Założmy, że $(\Omega, \mathcal{F}, \mathbb{P})$ jest przestrzenią probabilistyczną oraz $A, B \in \mathcal{F}$. Mówimy, że zdarzenia A i B są niezależne, jeśli

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

WAŻNE: Mówimy, że zdarzenia $(A_i)_{i \in I}$ są niezależne jeśli dla KAŻDEGO podzioru zbioru indeksów $i_1, \dots, i_k \in I$ zachodzi równość

$$\mathbb{P}(A_{i_1} \cap \dots \cap A_{i_k}) = \mathbb{P}(A_{i_1}) \cdot \dots \cdot \mathbb{P}(A_{i_k})$$

Definicja 97: Schemat Bernoulliego

Próba Bernoulliego nazywamy doświadczenie, w którym możliwe są wyłącznie dwa wyniki; jeden z nich interpretujemy jako sukces - 1 , drugi jako porażkę - 0. Schematem Bernoulliego nazywamy ciąg niezależnych powtórzeń tej samej próby Bernoulliego.

I. Liczba prób jest skończona i wynosi n . Wówczas $\Omega = \{0, 1\}^n$, $\mathcal{F} = 2^\Omega$, a \mathbb{P} bierzemy miarę jednoznacznie zadaną przez warunek:

$$P(\{(a_1, a_2, \dots, a_n)\}) = p^k(1 - p)^{n-k}$$

gdzie k oznacza liczbę jedynek w ciągu (a_1, a_2, \dots, a_n) .

II. Gdy liczba prób jest nieskończona, jako przestrzeń probabilistyczną należy wziąć nieskończony produkt przestrzeni $(\{0, 1\}, 2^{\{0,1\}}, (1 - p)\delta_0 + p\delta_1)$.

Prawdopodobieństwo zajścia dokładnie k sukcesów w schemacie Bernoulliego n prób z prawdopodobieństwem sukcesu w pojedynczej próbie równym p wynosi $\binom{n}{k}p^k(1 - p)^{n-k}$.

Twierdzenie 41: Twierdzenie Poissona

Jeśli $n \rightarrow \infty$, $p_n \rightarrow 0$, $np_n \rightarrow \lambda > 0$, to

$$\binom{n}{k}p_n^k(1 - p_n)^{n-k} \rightarrow e^{-\lambda} \frac{\lambda^k}{k!}$$

gdys $n \rightarrow \infty$.

Punkt czwarty

Definicja 98: Zmienna losowa

Odwzorowanie $X : \Omega \rightarrow \mathbb{R}^n$ nazywamy zmienną losową o wartościach w \mathbb{R}^n , jeśli dla każdego $A \in \mathcal{B}(\mathbb{R}^n)$ zbiór $X^{-1}(A) \in \mathcal{F}$.

Niech X będzie zmienną losową o wartościach w \mathbb{R}^n , a $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ funkcją borelowską. Wtedy $f(X)$ jest zmienną losową o wartościach w \mathbb{R}^m .

Definicja 99: Rozkład

Rozkładem zmiennej losowej X nazywamy miarę \mathbb{P}_X na przestrzeni $(\Omega, \mathcal{F}, \mathbb{P})$ t.ż. dla dowolnego $A \in \mathcal{B}(\mathbb{R})$ zachodzi $\mathbb{P}_X(A) = \mathbb{P}(X \in A) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in A\})$.

Definicja 100: Gęstość

Niech \mathbb{P} rozkład prawdopodobieństwa na \mathbb{R}^n . Gęstością nazywamy dowolną funkcję borelowską $f : \mathbb{R}^n \rightarrow \mathbb{R}_+$ całkowalną w sensie Lebesgue’a t.ż. dla dowolnego zbioru $A \in \mathcal{B}(\mathbb{R}^n)$

$$\mathbb{P}(A \in X) = \mathbb{P}_X(A) = \int_A f(x) \, dx$$

Rozkład, który ma gęstość, będziemy nazywać rozkładem ciągłym.

Rozkład \mathbb{P} na \mathbb{R}^n nazywamy dyskretnym, jeśli istnieje zbiór przeliczalny $S \subseteq \mathbb{R}^n$, dla którego $\mathbb{P}(S) = 1$.

Definicja 101: Dystrybuanta

Niech $(\Omega, \mathcal{F}, \mathbb{P})$, a X jest n wymiarową zmienną losową. Dystrybuantą tej zmiennej (bądź dystrybuantą rozkładu tej zmiennej) nazywamy funkcję $F_X : \mathbb{R}^n \rightarrow [0, 1]$ daną wzorem

$$F_X(t_1, \dots, t_n) = \mathbb{P}(X_1 \leq t_1, \dots, X_n \leq t_n)$$

Dystrybuanta zależy tylko od rozkładu

Twierdzenie 42: Niezależność zmiennych losowych

Dla zmiennych losowych X_1, X_2, \dots, X_n takich że X_i przyjmuje wartości w \mathbb{R}^{d_i} zmienne są niezależne:

1. jeśli dla każdego k , dowolnych parami różnych i_1, \dots, i_k i dowolnych $B_1 \in \mathbb{R}^{d_{i_1}}, B_2 \in \mathbb{R}^{d_{i_2}}, \dots, B_k \in \mathbb{R}^{d_{i_k}}$ zachodzi

$$P(X_{i_1} \in B_1, X_{i_2} \in B_2, \dots, X_{i_n} \in B_n) = P(X_{i_1} \in B_1) \cdot \dots \cdot P(X_{i_n} \in B_n).$$

2. σ -ciała generowane przez te zmienne losowe są niezależne*
3. dla dowolnych $B_1 \in \mathbb{R}^{d_1}, B_2 \in \mathbb{R}^{d_2}, \dots, B_k \in \mathbb{R}^{d_n}$ zdarzenia $\{X_1 \in B_1\}, \dots, \{X_n \in B_n\}$ są niezależne
4. $\mathbb{P}_{(X_1, \dots, X_n)} = \mathbb{P}_{X_1} \otimes \dots \otimes \mathbb{P}_{X_n}$
5. dla wszystkich t_1, t_2, \dots, t_n mamy

$$F_{(X_1, X_2, \dots, X_n)}(t_1, t_2, \dots, t_n) = F_{X_1}(t_1) \cdot F_{X_2}(t_2) \cdot \dots \cdot F_{X_n}(t_n)$$

6. dla rozkładów z gęstościami g_1, \dots, g_n

$$g(x_1, x_2, \dots, x_n) = g_1(x_1)g_2(x_2) \dots g_n(x_n)$$

$$*\sigma(X) := \{\{X \in B\} : B \in \mathcal{B}(\mathbb{R}^n)\}$$

Model probabilistyczny ciągu niezależnych doświadczeń to trójka:

$$(\Omega_1 \times \dots \times \Omega_n, \mathcal{F}_1 \otimes \dots \otimes \mathcal{F}_n, \mathbb{P}_1 \otimes \dots \otimes \mathbb{P}_n)$$

Definicja 102: Wartość oczekiwana

Założmy, że X jest zmienną losową określoną na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, \mathbb{P})$. Wartością oczekiwaną (średnią) zmiennej X nazywamy liczbę

$$\mathbb{E}[X] = \int_{\Omega} X d\mathbb{P} = \int_{\mathbb{R}} x d\mathbb{P}_X,$$

o ile całki istnieją. W przypadku gdy zmienna X jest dyskretna i jest skoncentrowana na zbiorze S , wartość oczekiwana wyraża się przez sumę

$$\mathbb{E}[X] = \sum_{x \in S} x \mathbb{P}(X = x).$$

Założmy, że $f : \mathbb{R} \rightarrow \mathbb{R}$ jest funkcją borelowską (co gwarantuje, że $f(X)$ będzie funkcją mierzalną na $(\Omega, \mathcal{F}, \mathbb{P})$). Wówczas

$$\mathbb{E}[f(X)] = \int_{\Omega} f(X) d\mathbb{P} = \int_{\mathbb{R}} f(x) d\mathbb{P}_X,$$

o ile całki po prawej stronie istnieją. W przypadku gdy zmienna X jest dyskretna i skoncentrowana na zbiorze przeliczalnym S , mamy

$$\mathbb{E}[f(X)] = \sum_{x \in S} f(x) \mathbb{P}(X = x).$$

Definicja 103: Wariancja

Dla danej zmiennej losowej X posiadającej skończoną wartość oczekiwaną $\mathbb{E}[X]$, definiujemy jej wariancję wzorem

$$\text{Var}(X) = \mathbb{E}(X - \mathbb{E}[X])^2.$$

Twierdzenie 43: Nierówność Czebyszewa

Niech X będzie nieujemną zmienną losową. Wtedy dla każdego $\varepsilon > 0$ mamy

$$\mathbb{P}(X \geq \varepsilon) \leq \frac{\mathbb{E}[X]}{\varepsilon}$$

Dowód:

$$\varepsilon \mathbb{1}_{X \geq \varepsilon} \leq X \mathbb{1}_{X \geq \varepsilon} \leq X$$

bierzemy średnią i otrzymujemy tezę

Twierdzenie 44: Suma zmiennych losowych

Założmy, że X_1 i X_2 są niezależnymi, jednowymiarowymi zmiennymi losowymi o rozkładach z gęstościami g_1 oraz g_2 , odpowiednio. Wówczas zmienna $X_1 + X_2$ ma rozkład z gęstością

$$g_1 * g_2(x) = \int_{\mathbb{R}} g_1(x - y)g_2(y) dy.$$

gdzie $g_1 * g_2$ to spłot gęstości g_1 i g_2 .

Dowód na gęstościach:

$$\begin{aligned} P(X_1 + X_2 \in B) &= P_{(X_1, X_2)}(\{(x, y) : x + y \in B\}) \\ &= \iint_{\{(x, y) : x + y \in B\}} g_{(X_1, X_2)}(x, y) dx dy \\ &= \iint_{\{(x, y) : x + y \in B\}} g_1(x)g_2(y) dx dy \\ &= \int_{\mathbb{R}^2} \mathbb{1}_B(x + y)g_1(x)g_2(y) dx dy \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathbb{1}_B(x + y)g_1(x) dx \right) g_2(y) dy \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathbb{1}_B(z)g_1(z - y) dz \right) g_2(y) dy \\ &= \int_B \left(\int_{\mathbb{R}} g_1(z - y)g_2(y) dy \right) dz \\ &= \int_B (g_1 * g_2)(z) dz. \end{aligned}$$

Warto również pamiętać, że jeśli choć jedna z niezależnych zmiennych losowych X_1, \dots, X_n ma gęstość to suma $X_1 + \dots + X_n$ ma gęstość. Wynika to z własności spłotu.

Rozkłady dyskretne

Rozkład dwupunktowy:

Jeśli $P(X = a) = p$ i $P(X = b) = 1 - p = q$, $p \in (0, 1)$.

Parametry: $a, b \in \mathbb{R}$, $a \neq b$, $p \in (0, 1)$. Momenty: $\mathbb{E}[X] = pa + qb$, $\text{Var}(X) = pq(a - b)^2$.

Rozkład Bernoulliego (dwumianowy):

$$P(X = k) = B(k, n, p) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad k = 0, 1, \dots, n, \quad p \in [0, 1],$$

$B(n, p)$. Jest to rozkład liczby sukcesów w n doświadczeniach Bernoulliego, gdzie szansa sukcesu w pojedynczym doświadczeniu wynosi p .

Parametry: $n \in \mathbb{N}$, $p \in [0, 1]$. Momenty: $\mathbb{E}[X] = np$, $\text{Var}(X) = npq$.

Rozkład Poissona:

$$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k = 0, 1, 2, \dots, \lambda > 0$$

$\text{Pois}(\lambda)$. Jest to rozkład graniczny dla ciągu rozkładów Bernoulliego $B(n, p_n)$, gdy $n \rightarrow \infty$, $p_n \rightarrow 0$ i $np_n \rightarrow \lambda$.

Parametr: $\lambda \in (0, \infty)$. Momenty: $\mathbb{E}[X] = \lambda$, $\text{Var}(X) = \lambda$.

Rozkład geometryczny:

$$P(X = k) = (1 - p)^{k-1} p, \quad p \in (0, 1), \quad k = 1, 2, \dots$$

Jest to rozkład czasu oczekiwania na pierwszy sukces w ciągu doświadczeń Bernoulliego.

Parametr: $p \in (0, 1)$. Momenty: $\mathbb{E}[X] = \frac{1}{p}$, $\text{Var}(X) = \frac{1-p}{p^2}$.

Rozkłady ciągłe

Rozkład jednostajny na $[a, b]$:

$$g(x) = \frac{1}{b - a} \mathbf{1}_{[a, b]}(x)$$

Oznaczenie: $\mathcal{U}[a, b]$

Momenty: $\mathbb{E}[X] = \frac{a+b}{2}$, $\text{Var}(X) = \frac{(b-a)^2}{12}$.

Rozkład wykładniczy:

Ma gęstość

$$g(x) = \lambda e^{-\lambda x} \mathbf{1}_{(0, \infty)}(x)$$

Parametr: $\lambda \in (0, \infty)$. Momenty: $\mathbb{E}[X] = \frac{1}{\lambda}$, $\text{Var}(X) = \frac{1}{\lambda^2}$

Rozkład normalny (Gaussa):

Jednowymiarowy standardowy rozkład normalny, $\mathcal{N}(0, 1)$, ma gęstość

$$g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Jeśli zmienna losowa X ma rozkład $\mathcal{N}(0, 1)$, to zmienna losowa $Y = \sigma X + \mu$, gdzie $\sigma > 0$, ma rozkład $\mathcal{N}(\mu, \sigma^2)$ o gęstości

$$g_Y(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Momenty: $\mathbb{E}[X] = \mu$, $\text{Var}(X) = \sigma^2$.

Punkt siódmy

Definicja 104: Zbieżności zmiennych losowych

Założmy, że $(X_n)_{n \geq 1}$ jest ciągiem zmiennych losowych. Mówimy, że:

(i) $X_n \xrightarrow{p.n.} X$ zbiega prawie na pewno jeśli

$$\mathbb{P}(\lim_{n \rightarrow \infty} X_n = X) = 1$$

(ii) $X_n \xrightarrow{L^p} X$ zbiega w L^p jeśli $X_1, X_2, \dots \in L^p$ oraz

$$\lim_{n \rightarrow \infty} \|X_n - X\|_p = 0$$

gdzie $\|\xi\|_p = (E|\xi|^p)^{1/p}$ dla $p < \infty$, $\|\xi\|_\infty = \text{ess sup}|\xi|$.

(iii) $X_n \xrightarrow{\mathbb{P}} X$ zbiega według prawdopodobieństwa, jeśli dla każdego $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0$$

Zależności między zbieżnościami

- 1. Jeśli $X_n \xrightarrow{p.n.} X$, to $X_n \xrightarrow{\mathbb{P}} X$. Implikacja przeciwna nie zachodzi.
- 2. Jeśli $X_n \xrightarrow{L^p} X$, to $X_n \xrightarrow{\mathbb{P}} X$. Implikacja przeciwna nie zachodzi.
- 3. Jeśli $p < q$ oraz $X_n \xrightarrow{L^q} X$, to $X_n \xrightarrow{L^p} X$.

Warto tutaj podać przykłady zmiennych dla których implikacja przeciwna nie zachodzi

Przykład 9: Przykład zmiennej losowej zbieżnej p.n. ale nie w L^p

Weźmy $\Omega = (0, 1]$, zbiory $A_n = (0, \frac{1}{n}]$ i $X_n = 2^n \mathbb{1}_{A_n}$. Wtedy $X_n(\omega) = 0$ dla $\omega > \frac{1}{n}$. Zatem $X_n \rightarrow 0$ p.n. (czyli również wdg. \mathbb{P}). Mamy jednak $\mathbb{E}|X_n|^p = 2^{np} \cdot \frac{1}{n} \rightarrow \infty$ czyli X_n nie jest zbieżny w L^p dla żadnego p .

Przykład 10: Przykład zmiennej zbieżnej wdg. \mathbb{P} ale nie p.n

Założmy, że przestrzeń probabilistyczna to przedział $[0, 1]$ wraz ze swoimi podzbiórami borelowskimi oraz miarą Lebesgue’a. Niech

$$\begin{aligned} X_1 &= \mathbb{1}_{[0,1]}, \\ X_2 &= \mathbb{1}_{[0,1/2)}, \quad X_3 = \mathbb{1}_{[1/2,1)}, \\ X_4 &= \mathbb{1}_{[0,1/4)}, \quad X_5 = \mathbb{1}_{[1/4,1/2)}, \quad X_6 = \mathbb{1}_{[1/2,3/4)}, \quad X_7 = \mathbb{1}_{[3/4,1)}, \\ &\vdots \end{aligned}$$

Wówczas ciąg $(X_n)_{n \geq 0}$ zbiega do 0 wdg. p-stwa: dla dowolnego ε , $\mathbb{P}(|X_n - 0| > \varepsilon)$ jest potęgą dwójki z coraz mniejszym całkowitym wykładnikiem. Z drugiej strony, dla dowolnego $\omega \in [0, 1)$, liczbowy ciąg $(X_n(\omega))_{n \geq 1}$ nie jest zbieżny; jest to ciąg zawierający nieskończenie wiele zer oraz nieskończenie wiele jedynek.

Punkt ósmy

Twierdzenie 45: Twierdzenie de Moivre’a-Laplace’a

Niech S_n będzie zmienną o rozkładzie Bernoulliego $B(n, p)$. Dla dostatecznie dużych n możemy przybliżać tą zmienną rozkładem normalnym

$$\mathbb{P} \left(a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b \right) \approx \Phi(b) - \Phi(a)$$

gdzie $\Phi(x)$ jest dystrybuantą rozkładu normalnego.

Twierdzenie 46: Prawa Wielkich Liczb

Założmy, że X_1, X_2, \dots jest ciągiem niezależnych zmiennych losowych.

1. Mocne Prawo Wielkich Liczb:

Jeśli zachodzi warunek

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

mamy:

$$\frac{1}{n} \sum_{i=1}^n (X_i - \mathbb{E}(X_i)) \xrightarrow[n \rightarrow \infty]{p.n} 0$$

2. Słabe Prawo Wielkich Liczb:

Jeśli zachodzi warunek

$$\sum_{i=1}^n \frac{\text{Var} X_i}{n^2} \xrightarrow[n \rightarrow \infty]{} 0$$

mamy:

$$\frac{1}{n} \sum_{i=1}^n (X_i - \mathbb{E}(X_i)) \xrightarrow[n \rightarrow \infty]{\text{wdg. } \mathbb{P}} 0$$

9 Matematyka Obliczeniowa

Uwaga

Rozdział ten będzie się różnił od pozostałych. Wynika to ze specyfiki przedmiotu, który ma więcej zastosowań od ścisłych definicji.

Punkt pierwszy

Numeryczny rozkład macierzy - trójkątno-trójkątny LU.

Jest to rozkład dla kwadratowej macierzy A o rozmiarach $n \times n$. Metoda ta bywa czasem niewystarczająca i należy wtedy zrobić rozkład PLU, który zostanie omówiony później.

Rozkład LU dla macierzy A . Piszemy: $A = \text{Id}_n \cdot A$, gdzie Id_n jest macierzą identycznościową o rozmiarze $n \times n$. Następnie robimy typową eliminację Gaussa dla macierzy A , przy czym:

1. nie zamieniamy wierszy macierzy,
2. w miejscu zera a_{ij} macierzy Id_n piszemy liczbę c , którą odjęliśmy, by od i -tego wiersza odjąć j -ty wiersz pomnożony przez c – powyższą operację należy robić kolumnami, z góry na dół (względem macierzy A przekształcającej się w macierz U).

Najlepiej zrozumieć to na przykładzie.

Przykład 11

$$\begin{aligned}
 A &= \begin{bmatrix} 2 & 0 & 3 \\ 4 & -5 & 7 \\ -2 & -15 & 1 \end{bmatrix} \\
 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 0 & 3 \\ 4 & -5 & 7 \\ -2 & -15 & 1 \end{bmatrix} \quad \{W_2 - 2W_1\} \\
 &= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 0 & 3 \\ 0 & -5 & 1 \\ -2 & -15 & 1 \end{bmatrix} \quad \{W_3 - (-1)W_1\} \\
 &= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 0 & 3 \\ 0 & -5 & 1 \\ 0 & -15 & 4 \end{bmatrix} \quad \{W_3 - 3W_2\} \\
 &= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 0 & 3 \\ 0 & -5 & 1 \\ 0 & 0 & 1 \end{bmatrix} \\
 &= L \cdot U.
 \end{aligned}$$

Naturalnym pytaniem jest, dlaczego ta metoda w niektórych przypadkach może nie zadziałać? Otóż dzieje się tak w momencie, gdy bez zamiany wierszy nie możemy macierzy zeszkodkować. Wtedy należy zastosować rozkład PLU.

Rozkład PLU dla macierzy A . Piszemy: $A = \text{Id}_n \cdot A$, gdzie Id_n jest macierzą identycznościową o rozmiarze $n \times n$. Następnie robimy typową eliminację Gaussa dla macierzy A , przy czym:

1. nie zamieniamy wierszy macierzy
2. w miejscu zera a_{ij} macierzy Id_n piszemy liczbę c , którą odjęliśmy, by od i -tego wiersza odjąć j -ty wiersz pomnożony przez c – powyższą operację należy robić kolumnami, z góry na dół (względem macierzy A przekształcającej się w macierz U),
3. w przypadku, gdy konieczne do zeszkodkowania dalej macierzy jest zamiana wierszy, tworzymy macierz permutacji $P_{i \leftrightarrow j}$, która jest macierzą identycznościową z zamienionym i -tym z j -tym wierszem – stąd wynika, że $P_{i \leftrightarrow j}^2 = \text{Id}_n$.

Przykład 12

Ponownie posłużymy się przykładem. Pamiętać należy, że $P_{i \leftrightarrow j}M$ zamienia i -ty z j -tym wierszem macierzy M , zaś $MP_{i \leftrightarrow j}$ zamienia i -tą z j -tą kolumną macierzy M . Jako ćwiczenie dla czytelnika zostawiamy sprawdzenie, że jeśli macierz M jest dolnotrójkątna, to macierz $N = PMP$, gdzie P jest macierzą permutacji, też będzie dolnotrójkątna. Kluczowa w rachunku niżej jest przedostatnia równość, która zapewnia, że kolejne trzy macierze w dużych nawiasach są odpowiednio postaci P (permutacja), L (dolnotrójkątna) i U (górnortrójkątna). Mamy:

$$\begin{aligned}
 A &= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix} \\
 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 2 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix} \quad \{W_2 - 2W_1\} \\
 &= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix} \quad \{W_3 - (-1)W_1\} \\
 &= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \{\text{problem: nie da się zrobić } W_3 - tW_2 \text{ dla żadnego } t \in \mathbb{R}\} \\
 &= P_{2 \leftrightarrow 3}^2 \cdot \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \cdot P_{2 \leftrightarrow 3}^2 \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \\
 &= \left(P_{2 \leftrightarrow 3} \right) \cdot \left(P_{2 \leftrightarrow 3} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \cdot P_{2 \leftrightarrow 3} \right) \cdot \left(P_{2 \leftrightarrow 3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \right) \\
 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
 &= P \cdot L \cdot U.
 \end{aligned}$$

Zastosowania do rozwiązywania układów równań algebraicznych liniowych - LU.

Mając rozkład macierzy $A = LU$ możemy przystąpić do rozwiązywania przykładowego równania $A\mathbf{x} = \mathbf{b}$, gdzie macierz A ma wymiary $n \times n$, natomiast \mathbf{x} oraz \mathbf{b} są wektorami wymiaru $n \times 1$.

Aby rozwiązać ten (de facto) układ równań, należy wykonać następujące kroki:

1. Rozwiązać równanie $L\mathbf{y} = \mathbf{b}$,
2. Rozwiązać równanie $U\mathbf{x} = \mathbf{y}$.

Dlaczego tak obliczony \mathbf{x} spełnia układ równań? Mamy bowiem: $A\mathbf{x} = LU\mathbf{x} = L\mathbf{y} = \mathbf{b}$. Dzięki rozkładowi $A = LU$, rozwiązywanie każdego z powyższych dwóch równań można wykonać, dokonując schodkowania macierzy co najwyżej $(n - 1)$ razy. W pierwszym kroku, jeśli oznaczymy $\mathbf{x} = (x_1, x_2, \dots, x_n)$, to w pierwszym równaniu mamy bowiem już dany x_1 , natomiast każde kolejne x_i wyznaczamy na podstawie wierszy nad aktualnym. W drugim kroku, to w ostatnim równaniu mamy dany x_n , natomiast każde poprzednie x_i wyznaczamy na podstawie wierszy pod aktualnym.

Zainteresowany czytelnik może spróbować rozwiązać równanie $A\mathbf{x} = \mathbf{b}$, gdzie macierz A jest taka sama jak w pierwszym przykładzie (w rozkładzie $A = LU$), zaś $\mathbf{b} = (11, 28, 8)$. Rozwiązaniem powinien być wektor $\mathbf{x} = (4, -1, 1)$, a po drodze powinno nam wyjść $\mathbf{y} = (11, 6, 1)$.

W przypadku rozkładu $A = PLU$, zauważmy, że równanie $A\mathbf{x} = \mathbf{b}$ jest równoważne $PLU\mathbf{x} = \mathbf{b}$, czyli $P^2LU\mathbf{x} = P\mathbf{b}$, czyli $LU\mathbf{x} = P\mathbf{b}$, pamiętając o tym że $P^2 = \text{Id}_n$. Wtedy "zerowym krokiem" jest policzenie $P_{i \leftrightarrow j}\mathbf{b}$, które jest po prostu zamienieniem b_i z b_j , gdzie $\mathbf{b} = (b_1, b_2, \dots, b_n)$.

Koszt, własności numeryczne - LU.

Złożoność obliczeniowa rozkładu $A = LU$ wynosi $O(n^3)$. Algorytm może być czuły na błędy zaokrągleń, w szczególności, gdy wartości na przekątnej macierzy A są bliskie zeru.

Złożoność obliczeniowa rozwiązania równania $A\mathbf{x} = \mathbf{b}$ wynosi $O(n^2)$. Dobrze uwarunkowane macierze A dają dokładne rozwiązania równania $A\mathbf{x} = \mathbf{b}$.

Numeryczny rozkład macierzy - ortogonalno-trójkątny QR.

Rozkład QR dla macierzy A . Macierz A może być prostokątna.

Macierz Q ma być ortogonalna, zaś macierz R - górnotrójkątna (tak jak U w rozkładzie LU). Jeśli macierz A ma wymiary $m \times n$, to macierz Q ma wymiary $m \times m$, zaś macierz R - $m \times n$.

Przypomnijmy, że macierz Q jest ortogonalna, jeśli $Q^T = Q^{-1}$. Własnością macierzy ortogonalnej jest to, że kolumny macierzy A , które oznaczать będziemy jako $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ traktowane jako wektory przestrzeni \mathbb{R}^n tworzą bazę ortonormalną, czyli że są prostopadłe oraz dla każdego $i = 1, 2, \dots, n$ mamy $\|\mathbf{a}_i\| = 1$.

Zacznijmy więc dekompozycję. Będziemy oznaczać przez $\mathbf{q}_1, \dots, \mathbf{q}_n$ kolumny macierzy Q , zaś przez $\mathbf{a}_1, \dots, \mathbf{a}_n$ - kolumny macierzy A . Definiujemy $\mathbf{q}_1 = \mathbf{a}_1 / \|\mathbf{a}_1\|$. Oczywiście wtedy $\|\mathbf{q}_1\| = 1$. Następnie definiujemy każde \mathbf{p}_i , tak jak w ortogonalizacji Grama-Schmidta, tzn. $\mathbf{p}_1 = \mathbf{a}_1$ oraz dla $k \geq 2$:

$$\mathbf{p}_k = \mathbf{a}_k - \langle \mathbf{a}_k, \mathbf{q}_{k-1} \rangle \cdot \mathbf{q}_{k-1} - \langle \mathbf{a}_k, \mathbf{q}_{k-2} \rangle \cdot \mathbf{q}_{k-2} - \dots - \langle \mathbf{a}_k, \mathbf{q}_1 \rangle \cdot \mathbf{q}_1.$$

Zaś \mathbf{q}_k definiujemy jako unormowany wektor \mathbf{p}_k , czyli $\mathbf{q}_k = \mathbf{p}_k / \|\mathbf{p}_k\|$.

Macierz R jest zaś macierz, której współczynnikami na przekątnej są $\|\mathbf{p}_1\|, \|\mathbf{p}_2\|, \dots, \|\mathbf{p}_n\|$, oczywiście odpowiednio w polach $r_{1,1}, r_{2,2}, \dots, r_{n,n}$. Współczynniki $a_{i,j}$ dla $i > j$ to oczywiście zera - macierz ma być górnotrójkątna. Współczynniki $a_{i,j}$ dla $i < j$ jest liczba $\langle \mathbf{a}_j, \mathbf{q}_i \rangle$.

Posłużmy się przykładem, by lepiej tę ideę zrozumieć.

Przykład 13

Weźmy

$$A = \begin{bmatrix} 3 & 13 & -13 \\ -6 & -12 & 5 \\ 2 & 18 & 10 \end{bmatrix}.$$

Mamy $\mathbf{a}_1 = (3, -6, 2)$, zatem $\|\mathbf{a}_1\| = \sqrt{3^2 + (-6)^2 + 2^2} = \sqrt{49} = 7$. Możemy więc uzupełnić pierwszą kolumnę macierzy Q wektorem $\mathbf{a}_1 / \|\mathbf{a}_1\|$, zaś pole $r_{1,1}$ liczbą $\|\mathbf{p}_1\| = \|\mathbf{a}_1\| = 7$. Widzimy, że

$$A = \begin{bmatrix} 3 & 13 & -13 \\ -6 & -12 & 5 \\ 2 & 18 & 10 \end{bmatrix} = \begin{bmatrix} 3/7 & \square & \square \\ -6/7 & \square & \square \\ 2/7 & \square & \square \end{bmatrix} \cdot \begin{bmatrix} 7 & \square & \square \\ 0 & \square & \square \\ 0 & 0 & \square \end{bmatrix} = Q \cdot R.$$

Liczmy teraz

$$\begin{aligned} \mathbf{p}_2 &= \mathbf{a}_2 - \langle \mathbf{a}_2, \mathbf{q}_1 \rangle \cdot \mathbf{q}_1 = \\ &= (13, -12, 18) - \left(\frac{13 \cdot 3/7 + (-12) \cdot (-6/7) + 18 \cdot 2/7}{7} \right) \cdot (3/7, -6/7, 2/7) = \\ &= (13, -12, 18) - \frac{21}{7} \cdot (3/7, -6/7, 2/7) = \\ &= (13, -12, 18) - (9, -18, 6) = (4, 6, 12). \end{aligned}$$

Norma wektora $(4, 6, 12)$ wynosi 14, zatem

$$\mathbf{q}_2 = \mathbf{p}_2 / \|\mathbf{p}_2\| = \frac{1}{14}(4, 6, 12) = (2/7, 3/7, 6/7).$$

Uzupełniamy więc

$$A = \begin{bmatrix} 3 & 13 & -13 \\ -6 & -12 & 5 \\ 2 & 18 & 10 \end{bmatrix} = \begin{bmatrix} 3/7 & 2/7 & \square \\ -6/7 & 3/7 & \square \\ 2/7 & 6/7 & \square \end{bmatrix} \cdot \begin{bmatrix} 7 & 21 & \square \\ 0 & 14 & \square \\ 0 & 0 & \square \end{bmatrix} = Q \cdot R.$$

W końcu liczymy

$$\begin{aligned} \mathbf{p}_3 &= \mathbf{a}_3 - \langle \mathbf{a}_3, \mathbf{q}_1 \rangle \cdot \mathbf{q}_1 - \langle \mathbf{a}_3, \mathbf{q}_2 \rangle \cdot \mathbf{q}_2 \\ &= (-13, 5, 10) - \left(\frac{-13 \cdot 3/7 + 5 \cdot (-6/7) + 10 \cdot 2/7}{1} \right) \cdot \mathbf{q}_1 - \left(\frac{-13 \cdot 2/7 + 5 \cdot 3/7 + 10 \cdot 6/7}{1} \right) \cdot \mathbf{q}_2 \\ &= (-13, 5, 10) - (-7) \cdot (3/7, -6/7, 2/7) - 7 \cdot (2/7, 3/7, 6/7) \\ &= (-13, 5, 10) - (-3, 6, -2) - (2, 3, 6) = (-12, -4, 6). \end{aligned}$$

Norma wektora $(-12, -4, 6)$ wynosi 14, zatem

$$\mathbf{q}_3 = \mathbf{p}_3 / \|\mathbf{p}_3\| = \frac{1}{14}(-12, -4, 6) = (-6/7, -2/7, 3/7).$$

Ostateczny rozkład:

$$A = \begin{bmatrix} 3 & 13 & -13 \\ -6 & -12 & 5 \\ 2 & 18 & 10 \end{bmatrix} = \begin{bmatrix} 3/7 & 2/7 & -6/7 \\ -6/7 & 3/7 & -2/7 \\ 2/7 & 6/7 & 3/7 \end{bmatrix} \cdot \begin{bmatrix} 7 & 21 & -7 \\ 0 & 14 & 7 \\ 0 & 0 & 14 \end{bmatrix} = Q \cdot R$$

Zastosowania do rozwiązywania układów równań algebraicznych liniowych - QR.

Uporaliśmy się z tym, jak rozkładać odwracalną kwadratową macierz A na iloczyn macierzy QR . W jaki sposób można teraz rozwiązać równanie $A\mathbf{x} = \mathbf{b}$? Ponownie, A jest macierzą $n \times n$, zaś \mathbf{x} oraz \mathbf{b} są wektorami wymiaru $n \times 1$.

Przekształcamy równoważnie:

$$\begin{aligned} A\mathbf{x} &= \mathbf{b} \\ QR\mathbf{x} &= \mathbf{b} \\ Q^{-1}QR\mathbf{x} &= Q^{-1}\mathbf{b} \\ R\mathbf{x} &= Q^T\mathbf{b} \end{aligned}$$

W ostatnim przekształceniu skorzystaliśmy z tego, że $Q^{-1}Q = \text{Id}_m$ oraz z faktu, że skoro macierz Q jest ortogonalna, to $Q^{-1} = Q^T$.

Aby teraz rozwiązać to równanie (znaleźć niewiadomą \mathbf{x}), to zauważmy, że $Q^T\mathbf{b}$ to zwykłe przemnożenie macierzy przez wektor. Lewa strona zaś łatwo daje się rozłożyć jak poprzednio (to jest jak macierz U w rozkładzie $A = LU$). Przypomnijmy: w ostatnim równaniu mamy dany x_n , natomiast każde poprzednie x_i wyznaczamy na podstawie wierszy pod aktualnym.

Koszt, własności numeryczne - QR.

Złożoność obliczeniowa rozkładu $A = QR$ wynosi $O(n^3)$. Algorytm jest przydatny w rozwiązywaniu problemów najmniejszych kwadratów. Algorytm QR jest stabilniejszy od LU dla gorzej uwarunkowanych macierzy. Złożoność obliczeniowa rozwiązania równania $A\mathbf{x} = \mathbf{b}$ wynosi $O(n^2)$.

Punkt drugi

Pisząc ten punkt dużo wiedzy czerpałem z prezentacji prof. Macieja Paszyńskiego (prof. z AGH) z wykładów o rachunku macierzowym.

Normy wektorowe i macierzowe oraz ich własności.

Przypomnijmy pojęcie normy najpierw dla wektorów. Później przytoczymy pojęcie normy dla macierzy.

Definicja 105: Norma wektorowa

Norma wektorowa to funkcja $\|\cdot\|: \mathbb{R}^n \rightarrow \mathbb{R}$, spełniająca trzy warunki:

Pierwszy warunek: $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$,

Drugi warunek (jednorodność): $\|t \cdot \mathbf{x}\| = |t| \cdot \|\mathbf{x}\|$,

Trzeci warunek (nierówność trójkąta): $\|\mathbf{x}\| + \|\mathbf{y}\| \geq \|\mathbf{x} + \mathbf{y}\|$,
przy czym $t \in \mathbb{R}$ oraz $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$

Najważniejsze normy na \mathbb{R}^n . Przyjmujemy $\mathbf{x} = (x_1, x_2, \dots, x_n)$.

$$\begin{aligned}\|\mathbf{x}\|_1 &= \sum_{i=1}^n |x_i| \\ \|\mathbf{x}\|_\infty &= \max_{1 \leq i \leq n} |x_i| \\ \|\mathbf{x}\|_2 &= \sqrt{\sum_{i=1}^n x_i^2}\end{aligned}$$

Normy wektorowe są używane do określania, czy ciąg wektorów zbiega do rozwiązania. Norma jest miarą odległości między kolejnymi iteracjami. Normy umożliwiają mierzenie różnicy między przybliżonym a dokładnym rozwiązaniem.

Definicja 106: Norma macierzowa

Norma macierzowa to funkcja $\|\cdot\|: \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$, spełniająca trzy warunki:

Pierwszy warunek: $\|A\| = 0 \iff$ macierz A jest macierzą zerową,

Drugi warunek (jednorodność): $\|t \cdot A\| = |t| \cdot \|A\|$,

Trzeci warunek (nierówność trójkąta): $\|A\| + \|B\| \geq \|A + B\|$,
przy czym $t \in \mathbb{R}$ oraz $A, B \in \mathbb{R}^{m \times n}$

Najważniejsze normy na $\mathbb{R}^{m \times n}$. Przyjmujemy $A = [a_{i,j}]$. Macierz A ma m wierszy i n kolumn, czyli $i \in \{1, 2, \dots, m\}$ oraz $j \in \{1, 2, \dots, n\}$.

Norma macierzowa jest indukowana przez normę wektorową. Dla $p \in [1, \infty) \cup \{\infty\}$ mamy:

$$\|A\|_p = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p} = \max_{\|\mathbf{x}\|_p=1} \|A\mathbf{x}\|_p.$$

Zauważmy, że:

$$\begin{aligned}\|A\|_1 &= \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{i,j}| \quad (\text{maksymalna suma wartości bezwzględnych kolumn}) \\ \|A\|_\infty &= \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{i,j}| \quad (\text{maksymalna suma wartości bezwzględnych wierszy}) \\ \|A\|_2 &= \sqrt{\lambda_{\max}(A^T A)} \quad (\text{liczba } \lambda_{\max} \text{ to największa wartość własna macierzy } A^T A).\end{aligned}$$

Uwaga. Liczba λ_{\max} jest nieujemna, gdyż macierz $A^T A$ jest dodatnio półokreślona.

Normy macierzowe pomagają ocenić, jak wrażliwy jest algorytm na perturbacje danych wejściowych. Normy są używane do określania liczby, zwanej współczynnikiem uwarunkowania macierzy, co jest miarą tego, jak bardzo rozwiązanie układu równań liniowych może się zmieniać w zależności od małych zmian w danych wejściowych. Dalej zostanie to bardziej omówione.

Wrażliwość numerycznych rozw. układu równań liniowych na zaburzenia danych.

Rozważmy układ równań liniowych w postaci $A\mathbf{x} = \mathbf{b}$. Analiza polega na oszacowaniu, jak bardzo zmienia się rozwiązanie \mathbf{x} podanego układu równań przy małej perturbacji macierzy A .

Definicja 107: Współczynnik uwarunkowania macierzy

Współczynnik uwarunkowania dla kwadratowej i odwracalnej macierzy A dla p -tej normy, gdzie $p \in [1, \infty) \cup \{\infty\}$

$$\text{cond}_p(A) = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p} : \min_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p} = \|A\|_p \cdot \|A^{-1}\|_p.$$

Intuicyjnie, jest to iloraz "maksymalnego wydłużenia pewnego wektora przez macierz" przez "minimalne wydłużenie pewnego wektora przez macierz".

Jeśli liczba $\text{cond}_p(A)$ jest postaci $1 + \varepsilon$, to układ jest dobrze uwarunkowany i zaburzenia danych mają niewielki wpływ na rozwiązanie. Jeśli liczba $\text{cond}_p(A)$ jest bardzo duża, to układ jest źle uwarunkowany i nawet małe błędy w danych mogą prowadzić do dużych błędów w rozwiązaniu.

Punkt trzeci

Interpolacja wielomianowa.

Interpolacja wielomianowa to jedna z podstawowych metod numerycznych, która polega na znalezieniu wielomianu, który dokładnie przechodzi przez zadany zbiór punktów danych. Jest to przydatne w przypadku, gdy dysponujemy ograniczonym zbiorem danych i chcemy uzyskać przybliżone wartości funkcji pomiędzy tymi punktami.

Definicja 108: Interpolacja wielomianowa

Dla danego zbioru $n + 1$ punktów $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$, interpolacja wielomianowa polega na znalezieniu wielomianu $P(x)$ stopnia n , który spełnia warunek:

$$P(x_i) = y_i \quad \text{dla } i = 0, 1, \dots, n$$

Istnieje kilka metod interpolacji wielomianowej, które umożliwiają obliczenie takiego wielomianu:

Definicja 109: Interpolacja Lagrange’a

Wielomian interpolacyjny Lagrange’a można zapisać w postaci:

$$P(x) = \sum_{i=0}^n y_i \cdot l_i(x)$$

gdzie funkcje bazowe $l_i(x)$ są dane wzorem:

$$l_i(x) = \prod_{\substack{0 \leq j \leq n \\ j \neq i}} \frac{x - x_j}{x_i - x_j}$$

Wielomian Lagrange’a jest łatwy do zapisania, ale trudniejszy do modyfikacji, gdy dodajemy nowe punkty.

Zanim przejdziemy do interpolacji Newtona, wprowadźmy definicję ilorazów różnicowych.

Definicja 110: Iloraz różnicowy

Ilorazem różnicowym rzędu n nazywamy wyrażenie definiowane rekurencyjnie:

$$f[x_0, x_1, \dots, x_n] = \frac{f[x_1, x_2, \dots, x_n] - f[x_0, x_1, \dots, x_{n-1}]}{x_n - x_0},$$
$$f[x_0, x_1] = \frac{f(x_1) - f(x_0)}{x_1 - x_0}.$$

Definicja 111: Interpolacja Newtona

Wielomian interpolacyjny Newtona można zapisać w formie:

$$P(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \dots + a_n(x - x_0)(x - x_1) \dots (x - x_{n-1})$$

Współczynniki a_i można obliczyć korzystając z ilorazów różnicowych, mianowicie

$$a_0 = f(x_0), a_n = f[x_0, x_1, \dots, x_n]$$

Metoda Newtona jest bardziej elastyczna niż metoda Lagrange’a, zwłaszcza gdy dodajemy nowe punkty.

Przykład 14

Zastosujmy interpolację Newtona oraz Lagrange’a dla funkcji

$$f(x) = \frac{e^{(x^2-x)/2}}{-x^2+2x+1},$$

dla $n = 2$ w punktach $x_0 = 0, x_1 = 1, x_2 = 2$.

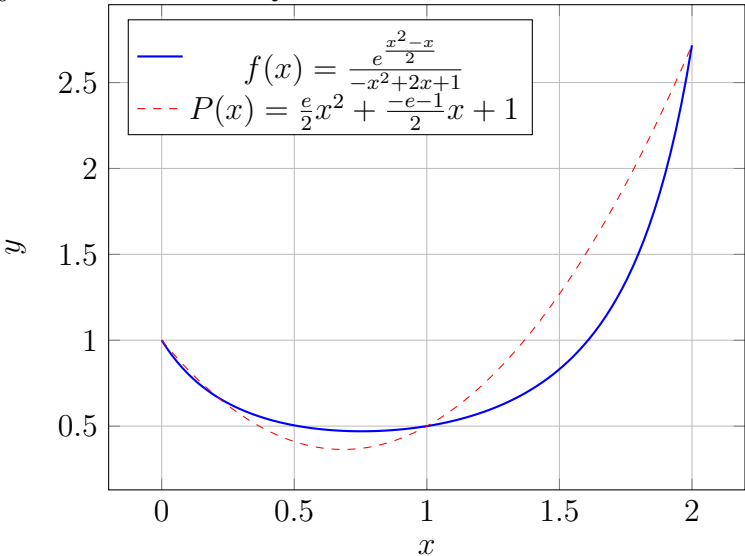
Podstawiając do funkcji f otrzymujemy wartości $y_0 = 1, y_1 = \frac{1}{2}, y_2 = e$. Wielomian interpolacyjny Lagrange’a stopnia 2 ma więc postać:

$$P_L(x) = 1 \cdot \frac{(x - 1)(x - 2)}{(0 - 1)(0 - 2)} + \frac{1}{2} \cdot \frac{(x - 0)(x - 2)}{(1 - 0)(1 - 2)} + e \cdot \frac{(x - 0)(x - 1)}{(2 - 0)(2 - 1)} = \frac{e}{2}x^2 + \frac{-e - 1}{2}x + 1.$$

Wielomian interpolacyjny Newtona jest postaci $P_N(x) = a_0 + a_1x + a_2x(x - 1)$. Wyliczmy współczynniki a_0, a_1, a_2 . Mamy $a_0 = y_0 = 1, a_1 = \frac{y_1 - y_0}{x_1 - x_0} = -\frac{1}{2}, a_2 = \frac{\frac{y_2 - y_1}{x_2 - x_1} - \frac{y_1 - y_0}{x_1 - x_0}}{x_2 - x_0} = \frac{e}{2}$, więc

$$P_N(x) = 1 - \frac{1}{2}x + \frac{e}{2}x(x - 1) = \frac{e}{2}x^2 + \frac{-e - 1}{2}x + 1.$$

Nie jest przypadkiem to, że wielomian wyszedł tej samej postaci. Zadanie interpolacyjne ma jednoznaczne rozwiązanie.



Wzór na resztę interpolacyjną i jego zastosowania

Reszta interpolacyjna w interpolacji wielomianowej opisuje błąd, który powstaje przy aproksymacji funkcji przez wielomian interpolacyjny. W przypadku interpolacji wielomianowej funkcji $f(x)$ w punktach x_0, x_1, \dots, x_n , reszta interpolacyjna $R(x)$ jest definiowana jako różnica między rzeczywistą wartością funkcji $f(x)$ a wartością wielomianu interpolacyjnego $P(x)$.

Definicja 112: Reszta interpolacyjna

Jeśli funkcja $f(x)$ jest $n + 1$ razy różniczkowalna na przedziale zawierającym punkty x_0, x_1, \dots, x_n oraz punkt x , to reszta interpolacyjna $R(x)$ dla interpolacji wielomianowej ma postać:

$$R(x) = f(x) - P(x) = \frac{f^{(n+1)}(\xi)}{(n + 1)!} \prod_{i=0}^n (x - x_i),$$

dla pewnego $\xi \in [x_0, x_n]$

Wzór ten pokazuje, że błąd interpolacji zależy od:

- 1. **Pochodnej rzędu $n + 1$** funkcji $f(x)$ – co oznacza, że im funkcja jest "gładsza", tym mniejszy błąd interpolacji.
- 2. **Odległości punktu x od punktów interpolacji** x_0, x_1, \dots, x_n .
- 3. **Stopnia wielomianu n** – wraz ze wzrostem stopnia wielomianu, może wzrosnąć liczba pochodnych wyższych rzędów, co może prowadzić do większego błędu, szczególnie gdy funkcja nie jest wystarczająco gładka.

W niektórych przypadkach interesuje nas maksymalny błąd interpolacji na całym przedziale. Wzór na resztę interpolacyjną można użyć do oszacowania tego błędu, jeśli potrafimy oszacować maksymalną wartość $f^{(n+1)}(\xi)$ na danym przedziale. Ocena błędu interpolacji pozwala na ustalenie, czy wyniki obliczeń są wystarczająco dokładne.

Punkt czwarty

Aproksymacja w przestrzeniach unitarnych.

Definicja 113: Przestrzeń unitarna

Przestrzeń unitarna to przestrzeń liniowa wyposażona w iloczyn skalarny, który umożliwia definiowanie normy.

Aproksymacja funkcji w przestrzeniach unitarnych polega na znalezieniu funkcji z pewnej podprzestrzeni, która jest jak najbardziej zbliżona do aproksymowanej funkcji w sensie normy tej przestrzeni. Rozważmy przestrzeń unormowaną $\mathcal{C}[a, b]$ funkcji ciągłych na odcinku $[a, b]$ z p -tą normą Höldera, która dla funkcji f i przybliżenia g jest opisana wzorem

$$||f - g||_p = \left(\int_a^b |f(x) - g(x)|^p \right)^{\frac{1}{p}}$$

Najbardziej użyteczne są przypadki $p = 2$ (aproksymacja średniokwadratowa) oraz $p \rightarrow \infty$ (aproksymacja jednostajna).

Aproksymacja jednostajna.

Aproksymacja jednostajna (zwana również aproksymacją Czebyszewa) polega na minimalizacji maksymalnej odległości (maksymalnego błędu) między funkcją aproksymowaną $f(x)$ a funkcją aproksymującą $g(x)$ na całym przedziale $[a, b]$.

Definicja 114: Błąd aproksymacji jednostajnej

Błędem aproksymacji jednostajnej na przedziale $[a, b]$ między funkcjami $f, g: [a, b] \rightarrow \mathbb{R}$ nazywamy liczbę

$$E(f, g) = \max_{x \in [a, b]} |f(x) - g(x)|$$

Celem aproksymacji jednostajnej jest znalezienie funkcji $g(x)$ z pewnej klasy funkcji (np. wielomianów) tak, aby $E(f, g)$ było jak najmniejsze.

Twierdzenie 47: Twierdzenie o aproksymacji jednostajnej wielomianem

Twierdzenie o aproksymacji jednostajnej (tzw. twierdzenie Czebyszewa) mówi, że dla każdej ciągłej funkcji $f(x)$ na przedziale $[a, b]$ istnieje wielomian $P_n(x)$ stopnia co najwyżej n , który minimalizuje błąd jednostajny. Ten minimalny błąd jest osiągany, gdy błąd aproksymacji w $n + 2$ punktach ma tę samą wartość i na przemian zmienia znak.

Definicja 115: Alternans

Zbiór tych $n + 2$ punktów z powyższego twierdzenia nazywamy alternansem.

Przykład 15

Rozważmy problem aproksymacji jednostajnej funkcji $f: [0, 1] \rightarrow \mathbb{R}$ określoną wzorem $f(x) = e^x$ wielomianem stopnia co najwyżej pierwszego. Przyjmijmy, że jest nim wielomian $P(x) = ax + b$. Potrzebne będą nam trzy punkty alternansu, którymi są $0, c$ oraz 1 , gdzie $c \in (0, 1)$ musi zostać wyznaczone. Chcemy wyznaczyć minimalny błąd jednostajny t , taki że

$$f(0) - P(0) = t, \quad f(c) - P(c) = -t, \quad f(1) - P(1) = t.$$

Niżej pod rozwiązaniem znajduje się rysunek. Może warto na niego zerknąć. Pierwsze równanie orzeka, że $1 - b = t$, zaś trzecie, że $e - a - b = t$. Stąd wnioskujemy, że

$$a = e - 1 \quad \text{oraz} \quad b = 1 - t.$$

Policzmy pochodną funkcji $f - P$. Mamy

$$(f - P)'(x) = (e^x - (e - 1)x - b)' = e^x - (e - 1).$$

Pochodna ta zeruje się w punkcie $x_0 = \ln(e - 1)$. Druga pochodna funkcji $f - P$ wynosi $e^x > 0$, więc w punkcie x_0 otrzymujemy najmniejszą wartość tego wyrażenia. Stąd wynika, że $c = x_0 = \ln(e - 1)$ oraz

$$-t = f(c) - P(c) = e^{\ln(e-1)} - (e - 1) \ln(e - 1) - (1 - t),$$

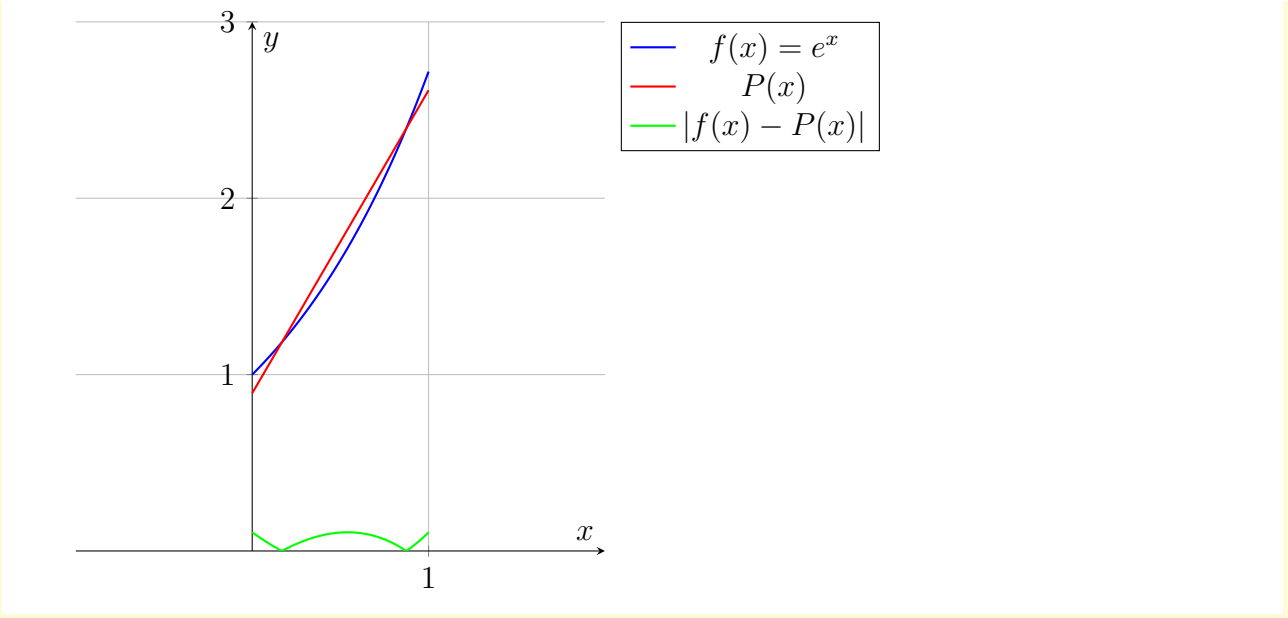
czyli

$$t = \frac{(e - 1) \ln(e - 1) - e + 2}{2},$$

zaś wielomian określony jest wzorem

$$P(x) = (e - 1)x - \frac{(e - 1) \ln(e - 1) - e}{2}.$$

Alternansem jest zbiór $\{0, \ln(e - 1), 1\}$.



W praktyce aproksymacja jednostajna znajduje zastosowanie w sytuacjach, gdy równomierną dokładność na całym przedziale jest kluczowa, np. w problemach optymalizacji, numerycznym rozwiązywaniu równań różniczkowych, gdzie krytyczne jest unikanie dużych błędów lokalnych.

Punkt piąty

Kwadratury interpolacyjne dla numerycznego całkowania funkcji jednej zmiennej.

Definicja 116: Kwadratura

Dana jest funkcja ciągła $f: [a, b] \rightarrow \mathbb{R}$. Załóżmy, że mamy n punktów x_0, x_1, \dots, x_{n-1} , w których znamy wartości funkcji f . Kwadraturą funkcji f na przedziale $[a, b]$ nazywamy wyrażenie:

$$Q(f) = \sum_{i=0}^{n-1} w_i \cdot f(x_i) \approx \int_a^b f(x) \, dx,$$

gdzie w_i są współczynnikami (wagami) kwadratury, które zależą od wyboru węzłów (punktów).

Przykład 16

Kwadratura trapezów. Dla funkcji f na przedziale całkowania $[a, b]$ określamy

$$Q(f) = \frac{f(a) + f(b)}{2} \cdot (b - a) = \frac{b - a}{2} \cdot f(a) + \frac{b - a}{2} \cdot f(b).$$

Przykład 17

Kwadratura Simpsona. Oparta na trzech węzłach. Oznaczmy $m = \frac{a+b}{2}$.

$$Q(f) = (b - a) \cdot \left(\frac{1}{6} f(a) + \frac{2}{3} f(m) + \frac{1}{6} f(b) \right).$$

Definicja 117: Kwadratura interpolacyjna

Założmy, że na przedziale $[a, b]$ funkcja f ma wielomian interpolacyjny P stopnia n . Kwadraturą interpolacyjną nazywamy całkę z wielomianu P .

Zauważmy, że w przypadku interpolacji Lagrange’a, jeśli przyjmiemy, że wielomianem interpolacyjnym funkcji f jest wielomian

$$P(x) = \sum_{i=0}^n y_i \cdot l_i(x) = \sum_{i=0}^n f(x_i) \cdot l_i(x),$$

to kwadraturą interpolacyjną jest

$$Q(f) = \int_a^b P(x) \, dx = \int_a^b \sum_{i=0}^n f(x_i) \cdot l_i(x) \, dx = \sum_{i=0}^n f(x_i) \int_a^b l_i(x) \, dx = \sum_{i=0}^n f(x_i) \cdot w_i,$$

gdzie w_i (liczby zależne od parametrów a i b) są współczynnikami kwadratury.

Definicja 118: Rząd kwadratury

Przypuśćmy, że $Q(f)$ jest kwadraturą funkcji f . Rzędem kwadratury nazywamy liczbę r , taką że dla każdego wielomianu $W(x)$ stopnia ściśle mniejszego od r zachodzi

$$\int_a^b f(x) \, dx = Q(W)$$

oraz istnieje wielomian $V(x)$ stopnia r , dla którego

$$\int_a^b f(x) \, dx \neq Q(V).$$

Dla kwadratury interpolacyjnej rzędu r o n węzłach zachodzą nierówności

$$n \leq r \leq 2n.$$

Kwadratury złożone dla numerycznego całkowania funkcji jednej zmiennej.

Definicja 119: Kwadratura złożona

Kwadratura złożona polega na podzieleniu przedziału całkowania $[a, b]$ na n przedziałów i zastosowaniu kwadratury dla każdego z nich.

Przykład 18

Złożona kwadratura trapezów. Niech

$$x_i = a + i \cdot \frac{b - a}{n} \quad \text{dla} \quad i \in \{0, 1, \dots, n\}$$

Dla funkcji f na przedziale całkowania $[a, b]$ określamy

$$Q(f) = \frac{b - a}{n} \cdot \left(\frac{1}{2} f(x_0) + \sum_{i=1}^{n-1} f(x_i) + \frac{1}{2} f(x_n) \right).$$

Zbieżność kwadratur złożonych.

Kwadratura złożona zbiega szybciej do dobrego oszacowania całki z funkcji. Można zastosować metodę zwaną adaptacją polegającą na podzieleniu przedziału najgorzej szacującego całkę na mniejsze przedziały i ponownym zastosowaniu metody kwadratury.

Punkt szósty

Metody numerycznego rozwiązywania równań nieliniowych skalarnych. Szybkość i warunki zbieżności tych metod.

Metoda 1: Styczne Newtona

Przyjmijmy założenia

1.

$f : [a, b] \rightarrow \mathbb{R}, f \in C^2$
2.

f ma dokładnie jedno miejsce zerowe na $[a, b]$
3.

Funkcja ma różne znaki na krańcach przedziału, tj. $f(a) \cdot f(b) < 0$
4.

Pierwsza i druga pochodna funkcji mają stały znak w tym przedziale.

to przybliżony wynik jest wyrażony wzorem

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

Zalety

- szybko zbieżna
- błąd wyznaczonego miejsca zerowego maleje kwadratowo wraz z ilością iteracji

Wady

- jest zbieżna lokalnie, co oznacza, że może zdarzyć się, iż wyznaczona wartość oddala się od rzeczywistego miejsca zerowego
- potencjalne problemy z wyznaczaniem pierwiastków wielokrotnych
- konieczność obliczania pierwszej i drugiej pochodnej funkcji

Metoda 2: Iteracja prostej

Służy ona do znajdowania rozwiązań równań postaci $x = f(x)$, czyli ma na celu znajdowanie punktów stałych danej funkcji, sama metoda jest zdefiniowana przez proces iteracyjny:

$$x_{n+1} = f(x_n)$$

- x^* jest punktem przyciągającym (atraktorem) procesu iteracyjnego wtedy i tylko wtedy, gdy $|f'(x^*)| < 1$
- Jeśli $g'(x^*) = 0, g''(x^*) < \infty$ to proces iteracji prostej jest kwadratowo zbieżny

Metoda 3: Siecznych

Przyjmijmy założenia

1.

$f : [a, b] \rightarrow \mathbb{R}, f \in C^1$
2.

$$\begin{cases} x_0 = a \\ x_1 = b \\ x_{n+1} = \frac{f(x_n)x_{n-1} - f(x_{n-1})x_n}{f(x_n) - f(x_{n-1})} \end{cases}$$
3.

Dla każdego n zachodzi $f(x_n)f(x_{n-1}) < 0$

to wtedy metoda zbiega do miejsca zerowego $f(x)$

Zalety

- nie wymaga obliczania pochodnej funkcji
- jest efektywniejsza od metody stycznych (Newtona)

Wady

- jest zbieżna lokalnie, co oznacza, że może zdarzyć się, iż wyznaczona wartość oddala się od rzeczywistego miejsca zerowego

10 Statystyczna Analiza Danych

10.1 Pojęcia wstępne

Definicja 120: Próbką

Próbką (próbą losową) nazywamy ciąg niezależnych zmiennych losowych X_1, \dots, X_n o jednakowym rozkładzie.

Definicja 121: Dystrybuanta empiryczna

Niech X_1, \dots, X_n będzie próbką z rozkładu o dystrybuancie F . Funkcję

$$\hat{F}(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}(X_i \leq x)$$

nazywamy **dystrybuantą empiryczną**.

Definicja 122: Model statystyczny

Model statystyczny określamy przez podanie rodziny $\{\mathbb{P}_\theta; \theta \in \Theta\}$ rozkładów prawdopodobieństwa na przestrzeni próbkowej Ω oraz zmiennej losowej $X : \Omega \rightarrow \mathcal{X}$, którą traktujemy jako obserwację. Zbiór \mathcal{X} nazywamy przestrzenią obserwacji, zaś Θ nazywamy przestrzenią parametrów.

Definicja 123: Przestrzeń statystyczna

Trójka (Ω, F, \mathcal{P}) , gdzie \mathcal{P} to rodzina miar probabilistycznych, $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$, a Θ to przestrzeń parametrów.

Definicja 124: Statystyka

Mierzalną funkcję $T : \mathcal{X} \rightarrow \mathcal{T}$ określoną na przestrzeni obserwacji \mathcal{X} nazywamy **statystyką** o wartościach w przestrzeni \mathcal{T} .

Inaczej: statystyka – zmienna losowa $T : \Omega \rightarrow \mathbb{R}$ i funkcja t taka, że $T(\omega) = t(Y_1(\omega), \dots, Y_n(\omega))$.

Przykład 19: Przykłady statystyk

Średnia z próby:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i,$$

Wariancja z próby:

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

10.2 Estymatory

Intuicja: rozkłady prawdopodobieństwa często zależą od pewnych parametrów. Zwykle przyjmujemy, że jakieś zjawisko jest zgodne z jakimś rozkładem (najczęściej normalnym). Problem w tym, że z góry nie znamy ani średniej, ani wariancji (w przypadkach innych rozkładów mogą to być też inne parametry). Chcemy je **estymować**, czyli przybliżyć. W tym celu badamy odpowiednią próbkę i wykorzystujemy wybraną statystykę, żeby z jakimś prawdopodobieństwem błędu wybrać odpowiednie parametry. Np. jeśli średnia z próby wynosi 0 (albo np. 0,01, to

spodziewamy się, że wartość oczekiwana też będzie równa 0.

Definicja 125: Estymator

Estymatorem parametru $\theta \in \Theta$ nazywamy dowolną statystykę $T = T(X)$ o wartościach w zbiorze Θ .
Oznaczenie: $T =: \hat{\Theta}$.

Definicja 126: Estymator nieobciążony

Estymator nieobciążony – estymator $\theta_n = \theta_n(X_1, \dots, X_n)$ parametru θ z próby losowej X_1, \dots, X_n , taki, że

$$\mathbb{E}[\theta_n] = \theta.$$

Definicja 127: Obciążenie estymatora

Obciążenie estymatora:

$$b(\theta_n) = \mathbb{E}[\theta_n] - \theta$$

Intuicja: liczymy, jak bardzo estymator różni się od realnej wartości.

Definicja 128: Estymator efektywny

Spośród zbioru wszystkich nieobciążonych estymatorów $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_r$ najefektywniejszym nazywamy estymator o najmniejszej wariancji.

Definicja 129: Wiarygodność

Wiarygodność jest to funkcja $\mathcal{L}: \Theta \rightarrow \mathbb{R}$ dana wzorem

$$\mathcal{L}(\theta) = f_{\theta}(x),$$

gdzie f_{θ} jest łączną gęstością obserwacji.

Wiarygodność jest właściwie tym samym, co gęstość prawdopodobieństwa, ale rozważana jako funkcja parametru θ , przy ustalonych wartościach obserwacji $x = X(\omega)$.
Uwaga praktyczna: zwykle f_{θ} jest iloczynem gęstości obserwacji (gdy są one niezależne).

Definicja 130: Estymator największej wiarygodności

Estymatorem największej wiarygodności (MLE) parametru θ nazywamy estymator $\hat{\theta}$, dla którego

$$f_{\hat{\theta}(x)}(x) = \sup_{\theta \in \Theta} f_{\theta}(x).$$

Intuicja: optymalizujemy funkcję wiarygodności (zwykle iloczyn gęstości).

Definicja 131: Estymator zgodny

Estymator θ_n parametru θ jest **zgodny**, jeśli dla każdego $\epsilon > 0$

$$\lim_{n \rightarrow \infty} P(|\theta_n - \theta| \geq \epsilon) = 0.$$

Interpretacja: dla wystarczająco dużych licznosci próby estymator przyjmuje z dużym prawdopodobieństwem wartości bliskie estymowanemu parametrowi θ .

Definicja 132: Estymator mocno zgodny

Estymator θ_n jest **mocno zgodny**, jeśli

$$P\left(\lim_{n \rightarrow \infty} \theta_n = \theta\right) = 1.$$

Interpretacja: z prawdopodobieństwem 1 realizacje estymatora dążą, przy $n \rightarrow \infty$, do estymowanego parametru.

Zgodny - zbieżność według prawdopodobieństwa, mocno zgodny - zbieżność prawie na pewno. Mocno zgodny zawsze jest zgodny.

Twierdzenie 48: (Gliwienko-Cantelli)

Jeżeli X_1, \dots, X_n, \dots jest próbką z rozkładu o dystrybuancie F to

$$\sup_{-\infty < x < \infty} |\hat{F}_n(x) - F(x)| \xrightarrow{p.n.} 0 \quad (n \rightarrow \infty).$$

10.3 Przedziały ufności**Definicja 133: Przedział ufności**

Niech $\theta \in \Theta$ oraz niech X_1, \dots, X_n będzie próbką. **Przedziałem ufności na poziomie ufności** $1 - \alpha \in (0, 1)$ nazywamy przedział (θ_1, θ_2) spełniający warunki

- (i) $\theta_1 = \theta_1(X_1, \dots, X_n)$ oraz $\theta_2 = \theta_2(X_1, \dots, X_n)$ – funkcje próbki X_1, \dots, X_n ;
- (ii) $\mathbb{P}(\theta_1(X_1, \dots, X_n) < \theta < \theta_2(X_1, \dots, X_n)) = 1 - \alpha$.

Najbardziej dokładna estymacja przedziałowa to ta, która daje najkrótszy przedział ufności $I_n = \theta_2(X_1, \dots, X_n) - \theta_1(X_1, \dots, X_n)$.

W następnych przykładach opisujemy przedziały ufności dla próbek z rozkładu normalnego $\mathcal{N}(\mu, \sigma)$, gdzie μ jest wartością oczekiwaną, a σ **odchyleniem standardowym**, czyli σ^2 jest wariancją.

Przypomnienie: \bar{X}_n - średnia z próby.

Przykład 20: Przedział ufności: zmany σ , nie znamy μ , estymujemy μ

$$\left(\bar{X}_n - q_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X}_n + q_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right),$$

gdzie $q_{1-\frac{\alpha}{2}}$ jest kwantylem standardowego rozkładu normalnego $\mathcal{N}(0, 1)$ rzędu $1 - \frac{\alpha}{2}$.

Definicja 134: Rozkład t -studenta

Jeśli $X_0, X_1, \dots, X_n \sim N(0, 1)$ są niezależnymi zmiennymi losowymi, to zmienna $T = \frac{X_0}{\sqrt{\sum_{i=1}^n X_i^2/n}}$ ma **rozkład t -Studenta** $t(n)$.

Liczbę n nazywamy liczbą **stopni swobody**.

Przykład 21: Przedział ufności: nie zmany ani σ , ani μ , estymujemy μ

$$\left(\bar{X}_n - t_{1-\frac{\alpha}{2}, n-1} \frac{S_n}{\sqrt{n-1}}, \bar{X}_n + t_{1-\frac{\alpha}{2}, n-1} \frac{S_n}{\sqrt{n-1}}\right),$$

gdzie $t_{1-\frac{\alpha}{2}, n-1}$ jest kwantylem rozkładu t -studenta o $n - 1$ stopniach swobody rzędu $1 - \frac{\alpha}{2}$.

Definicja 135: Rozkład chi-kwadrat

Niech $X_1, \dots, X_n \sim \mathcal{N}(0, 1)$ będą niezależne. Rozkład zmiennej $Q = \sum_{i=1}^n X_i^2$ nosi nazwę rozkładu **chi-kwadrat**, $\chi^2(n)$.

Liczbę n nazywamy liczbą **stopni swobody**.

Przypomnienie: S_n^2 - wariancja z próby.

Przykład 22: Przedział ufności: nie znamy ani σ , ani μ , estymujemy σ^2 (a nie σ !)

$$\left(\frac{nS_n^2}{\chi_{1-\frac{\alpha}{2}, n-1}^2}, \frac{nS_n^2}{\chi_{\frac{\alpha}{2}, n-1}^2} \right),$$

gdzie $\chi_{\frac{\alpha}{2}, n-1}^2$ jest kwantylem rozkładu χ^2 o $n - 1$ stopniach swobody rzędu $\frac{\alpha}{2}$.

10.4 Testowanie hipotez statystycznych

Definicja 136: Hipoteza statystyczna

Hipotezy statystyczne – stwierdzenia postaci $\theta \in \Theta_0$ oraz $\theta \in \Theta_1$, nazywane odpowiednio **hipotezą zerową** (H_0) i **hipotezą alternatywną** (H_1).

Intuicja: mamy przypuszczenia co do rozkładu danej populacji (jego postaci funkcyjnej lub wartości parametrów). Prawdziwość przypuszczenia sprawdzamy na podstawie wyników próby losowej. Na przykład mamy próbkę z rozkładu z jakimś parametrem i hipotezą jest dokładna wartość tego parametru lub przedział, do którego należy. Albo przypuszczamy, że rozkład jest normalny, a wcale nie jesteśmy tego pewni.

Definicja 137: Hipoteza prosta

Hipoteza prosta – hipoteza H_i , dla której odpowiadający jej zbiór $\Theta_i (i = 0, 1)$ składa się z tylko jednego elementu.

Definicja 138: Hipoteza złożona

Hipoteza złożona – hipoteza H_i , dla której odpowiadający jej zbiór $\Theta_i (i = 0, 1)$ ma więcej elementów.

Przykład hipotezy prostej: średnia wynosi albo 1, albo 2. Przykład hipotezy złożonej: próbka pochodzi z rozkładu normalnego (dopuszczamy różne parametry).

Definicja 139: Hipoteza parametryczna

Hipoteza parametryczna – hipoteza dotycząca wartości parametru rozkładu.

Definicja 140: Test statystyczny

Test statystyczny – reguła postępowania, która każdej próbie przyporządkowuje decyzję przyjęcia lub odrzucenia hipotezy.

Definicja 141: Błędy

Błąd I rodzaju – odrzucenie prawdziwej hipotezy zerowej. **Błąd II rodzaju** – przyjęcie fałszywej hipotezy zerowej (odrzućenie prawdziwej hipotezy alternatywnej).

Definicja 142: Obszar krytyczny, obszar przyjęć

Niech Z_n będzie **statystyką testową** (statystyką, którą wykorzystujemy w celu badania hipotezy). **Zbiór (obszar) krytyczny** (odrzuceń H_0) W i **zbiór (obszar) przyjęć** W' (nieodrzuceń H_0) spełniają warunki:

- (i) $W \cap W' = \emptyset, W \cup W' = \mathbb{R}$ (rozłączny podział \mathbb{R}),
- (ii) jeżeli $Z_n(X_1, \dots, X_n) \in W$ to H_0 odrzucamy,
- (iii) jeżeli $Z_n(X_1, \dots, X_n) \in W'$, to H_0 przyjmujemy (nie ma podstaw do odrzucenia H_0).

Definicja 143: Poziom istotności

Poziom istotności – prawdopodobieństwo

$$\alpha(W) := \mathbb{P}(Z_n \in W \mid H_0)$$

popęśnienia błędu I rodzaju.

Dodatkowo oznaczamy

$$\beta(W) := P(Z_n \in W' \mid H_1) = P(Z_n \in \mathbb{R} \setminus W \mid H_1).$$

Interpretacja: przyjmujemy hipotezę zerową i sprawdzamy, co się dzieje – jakie jest prawdopodobieństwo, że w takim przypadku trafiamy do obszaru krytycznego?

Definicja 144: Moc testu

Moc testu – prawdopodobieństwo $\mathbb{P}(Z_n \in W \mid H_1) = 1 - \beta(W)$.

Definicja 145: Test najmocniejszy

Testem najmocniejszym nazywamy test o tak dobranym obszarze krytycznym W , aby zminimalizować $\beta(W)$ przy ustalonym z góry poziomie istotności (prawdopodobieństwie błędu I rodzaju) $\alpha = \alpha(W)$.

Twierdzenie 49: Lemat Neymanna-Pearsona

Rozważmy hipotezy proste H_0 i H_1 takie, że obserwowana statystyka testowa Z pochodzi z rozkładów o gęstości, odpowiednio, f_1 i f_2 .

Najmocniejszy test osiągniemy dla obszaru krytycznego W , który spełnia warunek

$$W = \left\{ z \in \mathbb{R} : \frac{f_2(z)}{f_1(z)} \geq k \right\}$$

dla dodatniej liczby k spełniającej warunek

$$\int_W f_1(z) dt = \alpha.$$

Uwaga: $\operatorname{argmax}_x f$ jest zbiorem tych argumentów, dla których funkcja $f(x)$ przyjmuje największą wartość (globalnie).

Przykłady testów dla hipotez:

1. $H_0 : \theta = \theta_0,$
2. $H_1 : \theta \neq \theta_0$

pokrywają się z przykładami przedziałów ufności wypisanych wyżej (przy czym zwykle podajemy obszary krytyczne, czyli dopełnienia przedziałów ufności). Jeśli zaś hipotezą alternatywną jest $\theta > \theta_0$, zamiast obustronnego przedziału bierzemy tylko $(-\infty, -q_{1-\alpha}]$, gdzie q jest kwantylem odpowiedniego rozkładu. Ważne jest, że w tym przypadku bierzemy α zamiast $\frac{\alpha}{2}$.

10.5 Model liniowy Gaussa (regresja liniowa)

Przykład 23: Model regresji dla jednej zmiennej objaśniającej

Model bazuje na n obserwacjach $(x_1, y_1), \dots, (x_n, y_n)$, które nazywamy **danymi treningowymi**.

Model regresji liniowej zakłada, że dla każdego i zachodzi zależność

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i,$$

gdzie x_i są deterministyczne (obserwowane), a zmienne ε_i są niezależne o takim samym rozkładzie o zerowej wartości oczekiwanej i skończonej wariancji.

Definicja 146: W

modelu regresji liniowej

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

Y nazywamy **zmienną objaśnianą**, X **zmienną objaśniającą**, a ε błędem.

Interpretacja: mamy zestaw n danych. Badamy kolejne obserwacje x_j , $j > n$, ale y_j możemy już nie mieć. Chcemy je oszacować. Zakładamy, że zależność między nimi jest liniowa (po prostu taki mamy model, moglibyśmy też mieć inny). Czyli musimy „wyznaczyć” β_0 i β_1 , aby dostać wzór na y_j .

Oczywiście to jest statystyka, tutaj nic nie jest pewne, więc parametry estymujemy, a nie wyznaczamy. Są na to wzory.

Definicja 147: Resztowa suma kwadratów, RSS

Resztowa suma kwadratów jest liczbą daną wzorem

$$RSS = \sum_{i=1}^n \left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i\right)^2,$$

gdzie (x_i, y_i) są danymi treningowymi dla $i = 1, \dots, n$, a $\hat{\beta}_0$ i $\hat{\beta}_1$ są estymatorami β_0 i β_1 odpowiednio.

Wzory otrzymujemy poprzez minimalizację błędu mierzonego przez RSS .

Twierdzenie 50: Estymacja parametrów dla regresji z jedną zmienną objaśnianą

Minimalizacja RSS prowadzi do wzorów na estymatory $\hat{\beta}_0$ i $\hat{\beta}_1$:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2},$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x},$$

gdzie $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ oraz $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ (średnie arytmetyczne).

Przykład 24: Regresja liniowa z wieloma zmiennymi objaśniającymi

Dane treningowe są postaci (x_i, y_i) dla $i = 1, \dots, n$, przy czym $x_i = (x_{i,1}, \dots, x_{i,p})^T$ jest wektorem wejściowych wartości p zmiennych objaśniających. Model regresji liniowej zakłada, że

$$y_i = \beta_0 + \beta_1 x_{i,1} + \dots + \beta_p x_{i,p} + \varepsilon_i$$

W postaci macierzowej: niech X będzie macierzą $n \times (p + 1)$ taką, że

$$X = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{bmatrix}$$

oraz niech $y = (y_1, \dots, y_n)^T$ będzie wektorem wartości zmiennej objaśnianej, a $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)^T$ – wektorem błędów (niezależnych o jednakowym rozkładzie o zerowej wartości oczekiwanej i o skończonej wariancji, tzn. o macierzy $\sigma^2 \cdot I_n$ dla $\sigma^2 < \infty$ i macierzy identycznościowej I_n). Wtedy możemy zapisać model regresji liniowej w postaci

$$y = X\beta + \varepsilon.$$

Dodatkowo zakładamy, że wszystkie wyrazy macierzy X są deterministyczne (obserwowane) oraz, że X ma pełny rząd, czyli $\text{rank}(X) = p + 1$.

Twierdzenie 51: Estymacja parametrów dla regresji z jedną zmienną objaśnianą

Błąd RSS w modelu regresji liniowej z wieloma niewiadomymi jest minimalizowany dla estymatora $\hat{\beta}$ takiego, że

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

przy oznaczeniach podanych powyżej.

Definicja 148: Macierz daszkowa

Predykcją \hat{y} nazywamy

$$\hat{y} = X\hat{\beta} = X (X^T X)^{-1} X^T y.$$

Macierz $X (X^T X)^{-1} X^T$ nazywamy **macierzą daszkową**.

Twierdzenie 52: Twierdzenie Gaussa-Markowa

Dla błędów ε_i nieskorelowanych ($\mathbb{E}[\varepsilon_i \varepsilon_j] = 0$ dla $i \neq j$) i **homoskedastycznych** (czyli $\mathbb{E}[\varepsilon_i^2] = \sigma^2 < \infty$), estymator $\hat{\beta}$ jest liniowym, nieobciążonym estymatorem o najmniejszej wariancji parametru β .

Definicja 149: Całkowita suma kwadratów (TSS)

Całkowita suma kwadratów TSS dana jest wzorem

$$TSS = \sum (y_i - \bar{y})^2.$$

Definicja 150: Współczynnik determinacji (statystyka R^2)

Współczynnikiem determinacji nazywamy statystykę

$$R^2 = \frac{TSS - RSS}{TSS} = 1 - \frac{RSS}{TSS}.$$

Interpretacja: statystyka R^2 mierzy proporcję zmienności y , którą można wyjaśnić, używając modelu \hat{y} .

11 Funkcje Analityczne

Oznaczenia

- Przez Ω, U, V będziemy oznaczać zbiory otwarty w \mathbb{C} .
- $M(z) = \begin{bmatrix} \operatorname{Re}\{z\} & -\operatorname{Im}\{z\} \\ \operatorname{Im}\{z\} & \operatorname{Re}\{z\} \end{bmatrix}$.
- $J = M(i) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.
- Jako $D'(a, r)$ oznaczamy nakłute koło o środku a i promieniu r .

Rozwiązania właściwe

1

Definicja 151: Pochodna zespolona

Niech $f: \Omega \rightarrow \mathbb{C}$. Dla $z_0 \in \Omega$ definiujemy pochodną zespoloną f w punkcie z_0 jako

$$f'(z_0) = \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0}.$$

Definicja 152: Funkcja holomorficzna

Jeśli $f'(z_0)$ istnieje dla dowolnego $z_0 \in \Omega$, to mówimy, że funkcja f jest holomorficzna na Ω . Oznaczamy zbiór wszystkich funkcji holomorficznnych na Ω przez $H(\Omega)$.

Jeśli $f = u + iv$ oraz $z = x + iy$ to

$$Df(z_0) = \begin{bmatrix} \frac{\partial u}{\partial x}(z_0) & \frac{\partial u}{\partial y}(z_0) \\ \frac{\partial v}{\partial x}(z_0) & \frac{\partial v}{\partial y}(z_0) \end{bmatrix}.$$

Mamy też pochodne Wirtingera (są ponoć ważne, ale to zależy kogo spytać):

$$\frac{\partial f}{\partial z}(z_0) := \frac{1}{2} \left[\frac{\partial f}{\partial x}(z_0) - i \cdot \frac{\partial f}{\partial y}(z_0) \right] \text{ oraz } \frac{\partial f}{\partial \bar{z}}(z_0) := \frac{1}{2} \left[\frac{\partial f}{\partial x}(z_0) + i \cdot \frac{\partial f}{\partial y}(z_0) \right].$$

Twierdzenie 53: Równania Cauchy’ego-Riemanna

Niech $z = x + iy$ i niech $f: \Omega \rightarrow \mathbb{C}$ będzie różniczkowalny w $z_0 = x_0 + iy_0$. Wtedy w punkcie (x_0, y_0) funkcje $u = \operatorname{Re} f$ i $v = \operatorname{Im} f$ spełniają równania Cauchy’ego-Riemanna:

$$\begin{cases} \frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \\ \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \end{cases}.$$

Możemy w ładny sposób połączyć różniczkowalność z pochodnymi Wirtingera: równanie Cauchy’ego-Riemanna odpowiadają warunkowi:

$$\frac{\partial f}{\partial \bar{z}} = 0.$$

Definicja 153: Macierz zespolona

Mówimy, że macierz A jest zespolona, jeśli zachodzi $AJ = JA$.

Twierdzenie 54

Niech $f: \Omega \rightarrow \mathbb{C}$ będzie klasy C^1 oraz niech $z_0 \in \Omega$. Wówczas $f'(z_0)$ istnieje wtedy, i tylko wtedy, gdy $Df(z_0)$ jest zespolona.

Macierz różniczki $Df(z_0)$ jest macierzą przekształceniem konforemnym (zachowującym orientację oraz kąty), $Df(z_0) = M(f'(z_0))$ oraz $\det Df(z_0) = |f'(z_0)|$.

Twierdzenie 55: Holomorficzność a analityczność

Jeśli f jest holomorficzna to jest analityczna.

Twierdzenie 56: Holomorficzność a harmoniczność

Jeśli f jest holomorficzna to f jest harmoniczna.

2

Definicja 154: Podstawowe funkcje elementarne

- funkcja wykładnicza $\exp: \mathbb{C} \rightarrow \mathbb{C}$:

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}.$$

- funkcje trygonometryczne:

$$\sin(z) = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!}$$

$$\cos(z) = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!}$$

Niech U będzie obszarem. Powiemy, że w U istnieje ciągła gałąź logarytmu, jeśli istnieje ciągła funkcja $g: U \rightarrow \mathbb{C}$ taka, że $e^{g(z)} = z$ w U . Okazuje się, że nie zawsze możemy taką funkcję stworzyć. Spróbujmy to zrobić dla $U = \mathbb{C} \setminus \{0\}$. Możemy wziąć za $g(1)$ dowolną z liczb $2\pi i k, k \in \mathbb{Z}$, przykładowo 0. Jak popatrzymy na wartości logarytmu na okręgu jednostkowym to okazuje się, że po wykonaniu jednego obrotu wartość w 1 powinna nam wyjść $2\pi i$. Czyli mamy problem, który rozwiązuje nam następujące twierdzenie:

Twierdzenie 57:

Założmy, że U jest jednospójna (mówiąc po ludzku - nie ma dziur). Wtedy istnieje taka funkcja $g: U \rightarrow \mathbb{C}$, taka że $e^{g(z)} = z$ w U .

Definicja 155:

Homografie to funkcje przekształcające sferę Riemanna w nią samą (bijektywnie) postaci

$$\left\{ f: \hat{\mathbb{C}} \rightarrow \hat{\mathbb{C}}: f(z) = \frac{az + b}{cz + d} \quad a, b, c, d \in \mathbb{C}, ad - bc \neq 0 \right\}.$$

Są one złożeniem inwersji, przesunięcia i jednokładności.

Możemy przyporządkować każdej homografii macierz 2 na 2 w następujący sposób:

$$f(z) = \frac{az + b}{cz + d} \text{ przyporządkowujemy macierz } \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Warto zwrócić uwagę, że to przyporządkowanie nie jest jednoznaczne, ale możemy to poprawić wymagając żeby ta macierz była o wyznaczniku 1.

Twierdzenie 58:

Złożenie dwóch homografii odpowiada mnożeniu odpowiadających im macierzy.

Możemy stąd wyciągnąć wniosek, że homografie tworzą grupę.

Twierdzenie 59:

Homografie przenoszą uogólnione okręgi (czyli proste i okręgi) na uogólnione okręgi.

3

Definicja 156: Całka z funkcji zespolonej

Niech $f: U \rightarrow \mathbb{C}$ będzie ciągła, a $\gamma \subset U$ będzie krzywą kawałkami \mathcal{C}^1 , parametryzowaną przez $t \mapsto \gamma(t)$, $t \in [0, 1]$. Wtedy definiujemy

$$\int_{\gamma} f(z) dz = \int_0^1 f(\gamma(t)) \gamma'(t) dt.$$

Równoważnie, dla miłośników form różniczkowych całka określona może zostać przez

$$\int_{\gamma} f(z) dz = \int_{\gamma} (u dx - v dy) + i \int_{\gamma} (v dx + u dy).$$

A dla masochistów przez sumy Riemanna.

Definicja 157: Homotopia ścieżek

Niech $\gamma, \lambda: [0, 1] \rightarrow \mathbb{C}$ będą parametryzacjami krzywych. Powiemy, że γ i λ są homotopijne, jeżeli istnieje będzie ciągłe przekształcenie $H: [0, 1] \times [0, 1] \rightarrow \mathbb{C}$, takie że dla każdego $t \in [0, 1]$ zachodzi $H(t, 0) = \gamma(t)$ i $H(t, 1) = \lambda(t)$.

Twierdzenie 60: Twierdzenie Cauchy'ego

Niech U będzie otwarty i jednospójny, $f \in H(U)$, a $\gamma, \lambda \subset U$ będą homotopijnymi krzywymi kawałkami \mathcal{C}^1 , dodatnio zorientowanymi. Wtedy

$$\int_{\gamma} f(z) dz = \int_{\lambda} f(z) dz.$$

Twierdzenie 61: Kolejne twierdzenie Cauchy'ego

Niech U będzie otwarty, a $f \in H(U)$. Wtedy dla $\gamma \subset U$ będącej zamkniętą krzywą bez samoprzecięć, kawałkami \mathcal{C}^1 , zachodzi

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

Twierdzenie 62: Wzór całkowy Cauchy'ego

Niech $U \subset \mathbb{C}$ będzie otwarty, a $\gamma \subset U$ będzie krzywą zamkniętą, bez samoprzebieć, dodatnio zorientowaną. Wtedy dla $f \in H(U)$ i a należącego do obszaru ograniczonego przez γ , zachodzi

$$f(a) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(z)}{z-a} dz.$$

Przykład 25

Funkcja holomorficzna jest analityczna. Aby to zobaczyć, niech $z_0 \in U$ i R będzie takie, że $D = D(z_0, R) \subset U$. Obierzmy punkt $z \in D$, wtedy dla $s \in \partial D$ zachodzi $|s - z_0| > |z - z_0|$, dzięki temu możemy pozwolić sobie na przejście

$$\begin{aligned} f(z) &= \frac{1}{2\pi i} \int_D \frac{f(s)}{s-z} ds = \frac{1}{2\pi i} \int_D \frac{f(s)}{s-z_0+z_0-z} ds \\ &= \frac{1}{2\pi i} \int_D \frac{f(s)}{s-z_0} \cdot \frac{1}{1-\frac{z-z_0}{s-z_0}} ds \\ &= \frac{1}{2\pi i} \int_D \frac{f(s)}{s-z_0} \cdot \sum_{k=0}^{\infty} \left(\frac{z-z_0}{s-z_0} \right)^k ds \\ &= \sum_{k=0}^{\infty} \left(\frac{k!}{2\pi i} \cdot \int_D \frac{f(s)}{(s-z_0)^{k+1}} ds \right) \cdot \frac{(z-z_0)^k}{k!}. \end{aligned}$$

Dzięki jednoznaczności rozwinięcia w szereg Taylora, dostajemy że zachodzi

$$f^{(k)}(z) = \frac{k!}{2\pi i} \cdot \int_{\gamma} \frac{f(s)}{(s-z)^{k+1}} ds.$$

Definicja 158: Indeks punktu względem krzywej

Niech $\gamma: [0, 1] \rightarrow \mathbb{C} \setminus \{p\}$ będzie parametryzacją krzywej zamkniętej $\gamma \subset \mathbb{C}$. Indeks punktu p względem γ , nazwiemy wyrażenie

$$\text{ind}_p \gamma = \frac{1}{2\pi i} \int_{\gamma} \frac{1}{z-p} dz.$$

Indeks krzywej interpretowany może być, jako liczba nawinięć krzywej γ , wokół punktu p . Istotnie bowiem będzie on zawsze liczbą całkowitą, jako że $2\pi i \cdot \text{ind}_p \gamma = \log(\gamma(1)) - \log(\gamma(0))$, tym samym $2\pi i \cdot \text{ind}_p \gamma$ jest więc przyrostem argumentu podniesienia $\tilde{\gamma}$, który jest całkowitą wielokrotnością $2\pi i$

Twierdzenie 63

Jeżeli $\gamma, \lambda: [0, 1] \rightarrow \mathbb{C} \setminus \{p\}$ będą zamkniętymi krzywymi w $\mathbb{C} \setminus \{p\}$, to $\text{ind}_p \gamma = \text{ind}_p \lambda$ wtedy i tylko wtedy, gdy γ i λ są homotopijne w $\mathbb{C} \setminus \{p\}$. W szczególności każda krzywa zamknięta wokół punktu 0 homotopijna jest z pewnym nawinięciem okręgu, to znaczy krzywą parametryzowaną przez $t \mapsto e^{2\pi i n t}$ dla pewnego $n \in \mathbb{Z}$.

4

Twierdzenie 64: Twierdzenie Morery

Załóżmy, że $f: \Omega \rightarrow \mathbb{C}$ jest ciągła oraz dla każdego domkniętego trójkąta $\triangle \subseteq \Omega$ zachodzi

$$\int_{\partial \triangle} f(z) dz = 0,$$

wówczas $f \in H(\Omega)$.

Definicja 159: Niemal jednostajna zbieżność

Rozważmy ciąg $\{f_n\}$ funkcji ciągłych $f_n: \Omega \rightarrow \mathbb{C}$. Mówimy, że ciąg funkcji f_n zbiega do funkcji f niemal jednostajnie, oznaczamy $f_n \xrightarrow{n.j.} f$, gdy dla dowolnego zwartego $K \subseteq \Omega$, $f_n \rightrightarrows f$.

Twierdzenie 65: Twierdzenie Weierstrassa

Niech $\{f_n\}$ będzie taki, że $f_n \in H(\Omega)$, dla dowolnego $n \geq 1$. Jeśli $f_n \xrightarrow{n.j.} f$, to $f \in H(\Omega)$. Dodatkowo, $f'_n \xrightarrow{n.j.} f'$.

5

Twierdzenie 66: Zasada maksimum

Niech U będzie spójna i niech $f: U \rightarrow \mathbb{C}$ będzie holomorficzna. Wtedy jeśli $|f|$ osiąga maksimum w U to f jest stała.

Twierdzenie 67: Liouville

Niech $f: \mathbb{C} \rightarrow \mathbb{C}$ będzie holomorficzna i ograniczona. Wtedy f jest stała.

Twierdzenie 68: Zasadnicze twierdzenie algebry

Założmy, że wielomian $f(z) = a_n z^n + \dots + a_0 \neq 0$ dla każdego $z \in \mathbb{C}$. Wtedy funkcja $\frac{1}{f(z)}$ jest dobrze określoną funkcją holomorficzną, która jest ograniczona, ponieważ wielomian poza pewnym dyskiem jest dowolnie duży. W takim razie z zasady maksimum ten wielomian jest stały (czyli każdy niestały wielomian ma pierwiastek zespolony).

6

Twierdzenie 69: Twierdzenie Laurenta

Funkcja holomorficzna w pierścieniu $P = \{z \in \mathbb{C} : r < |z - z_0| < R\}$ rozwija się w tym pierścieniu w szereg Laurenta:

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n, \quad \forall z \in P.$$

Definicja 160: Osobliwości

- Jeśli $f \in H(\Omega \setminus \{z_0\})$, to mówimy, że f ma osobliwość w z_0 .
- Jeśli istnieje $F \in H(\Omega)$ t.ż. $f(z) = F(z)$, dla $z \neq z_0$, to mówimy, że f ma osobliwość usuwalną w z_0 .
- Jeśli istnieją $c_1, \dots, c_m \in \mathbb{C}$, $m \in \mathbb{Z}_{>0}$ oraz $c_m \neq 0$, t. że

$$f(z) = \sum_{k=1}^m \frac{c_k}{(z - z_0)^k}$$

ma usuwalną osobliwość w z_0 , to mówimy, że f ma biegun rzędu m w z_0 .

- Jeśli osobliwość w z_0 nie jest ani usuwalna, ani nie jest biegunem, to mówimy, że f ma istotną osobliwość w z_0 .

Twierdzenie 70: Twierdzenie Riemanna o osobliwości usuwalnej

Dla $f \in H(\Omega \setminus z_0)$, gdzie $z_0 \in \Omega$, NWSR:

1. osobliwość w z_0 jest usuwalna,
2. istnieje $F \in C(\Omega)$ t.ż. $F(z) = f(z)$, dla $z \neq z_0$,
3. istnieje $r > 0$ t.ż. f jest ograniczona na $D'(z_0, r)$,
4. $\lim_{z \rightarrow z_0} (z - z_0)f(z) = 0$.

Twierdzenie 71: Twierdzenie Casoratiego-Weierstrassa

Założmy, że $f \in H(\Omega \setminus z_0)$ oraz f ma osobliwość istotną w z_0 . Wówczas, dla dowolnego $r > 0$ t. że $D(z_0, r) \subseteq \Omega$, zbiór $f(D'(z_0, r))$ jest gęsty w \mathbb{C} .

Twierdzenie 72: Twierdzenie o residuach

Założmy, że $f \in M(\Omega)$ (f jest funkcją meromorficzną). Niech A będzie zbiorem biegunów f . Założmy, że Γ jest cyklem w $\Omega \setminus A$ spełniającym

$$\operatorname{ind}_{\Gamma}(\alpha) = 0 \text{ dla dowolnego } \alpha \notin \Omega.$$

Wówczas

$$\frac{1}{2\pi i} \int_{\Gamma} f(z) dz = \sum_{a \in A} \operatorname{Res}(f; a) \operatorname{ind}_{\Gamma} a.$$

Przykład 26: Zastosowanie twierdzenia o residuach

Przykłady dokładne znajdują się w notatkach Politarczyka, wykład 11, jakby ktoś chciał je przepisać lub przeczytać. Ale to twierdzenie przydaje się do rozważania funkcji typu $\int_{-\infty}^{\infty} R(x) dx$, które słabo się liczy, do fajnych całek na dziedzinie zespolonej.

7

Definicja 161: Funkcja meromorficzna

Funkcja $f: U \rightarrow \mathbb{C}$ nazywana będzie meromorficzną, jeżeli jest holomorficzna na U , poza zbiorem punktów $\{z_1, z_2, \dots\}$, który nie zawiera punktu skupienia i każde z_i jest biegunem f .

Twierdzenie 73: Zasada argumentu

Niech $U \subset \mathbb{C}$ będzie obszarem jednospójnym, $f: U \rightarrow \mathbb{C}$ meromorficzna, a $\gamma \subset U$ będzie dodatnio zorientowaną krzywą zamkniętą bez samoprzecięć, taką że $f(z) \neq 0$ dla $z \in \gamma$. Oznaczmy przez $Z(\gamma)$ liczbę zer (z krotnościami) funkcji f zawartych w obszarze ograniczonym przez γ , a $P(\gamma)$ liczbę biegunów (z krotnościami) w tymże obszarze. Wtedy zachodzi

$$\frac{1}{2\pi i} \int_{\gamma} \frac{f'(z)}{f(z)} dz = Z(\gamma) - P(\gamma).$$

Twierdzenie 74: Twierdzenie Rouchégo

Niech U będzie jak wyżej, $f, g \in H(U)$. Dla $\gamma \subset U$, będącej dodatnio zorientowaną krzywą zamkniętą bez samoprzecięć, jeżeli dla $z \in \gamma$ zachodzi $|f(z) - g(z)| < |f(z)| + |g(z)|$ to f i g mają tyle samo zer (z krotnościami) w obszarze ograniczonym przez γ . W szczególności, jeżeli $|g(z)| < |f(z)|$ na γ , to f i $f + g$ mają tyle samo zer w obszarze ograniczonym przez γ .

Przykład 27: Zasadnicze twierdzenie algebry przez twierdzenie Rouchégo

Niech $w(z) = z^n + a_{n-1}z^{n-1} + \dots + a_0$ będzie wielomianem zespolonym, pokażemy że f musi mieć pierwiastek. Oczywiście $f(z) = z^n$ ma pierwiastek n krotny w $z = 0$, wobec tego jeżeli wszystkie a_i są zerowe to teza zachodzi. Jeżeli $g(z) = a_{n-1}z^{n-1} + \dots + a_0$ jest niezerowa, to

$$\left| \frac{f(z)}{g(z)} \right| = \left| \frac{z^n}{a_{n-1}z^{n-1} + \dots + a_0} \right| \rightarrow \infty.$$

Istnieć więc będzie dysk $D = D(0, R)$, że dla $z \in D$ zachodzi $|f(z)/g(z)| > 1$. W szczególności, dla $z \in \partial D$ mamy $|f(z)| > |g(z)|$, wobec tego na mocy twierdzenia Rouchégo f i $f + g = w$ mają tyle samo zer w D , czyli n .