Algebra liniowa i Macierze

Metody Numeryczne

dr inż. Grzegorz Fotyga

Politechnika Gdańska
Wydział Elektroniki, Telekomunikacji i Informatyki
Katedra Inżynierii Mikrofalowej i Antenowej

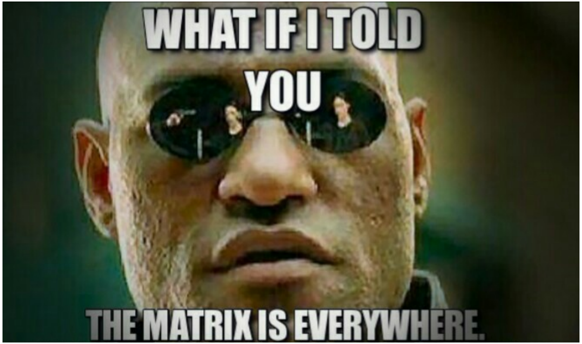
24 lutego 2018

Plan prezentacji

1. Podstawowe założenia
2. Macierze
3. Działania na Macierzach
4. Fraktale
5. Operacje na wektorach

Podstawowe założenia (1)

Duża część metod przedstawianych na wykładzie będzie wykorzystywała takie operacje, jak: **konstruowanie**, **modyfikowanie**, **rozwiązywanie** układów równań liniowych (równań macierzowych - ang. **matrix** equations)



Podstawowe założenia (1b)

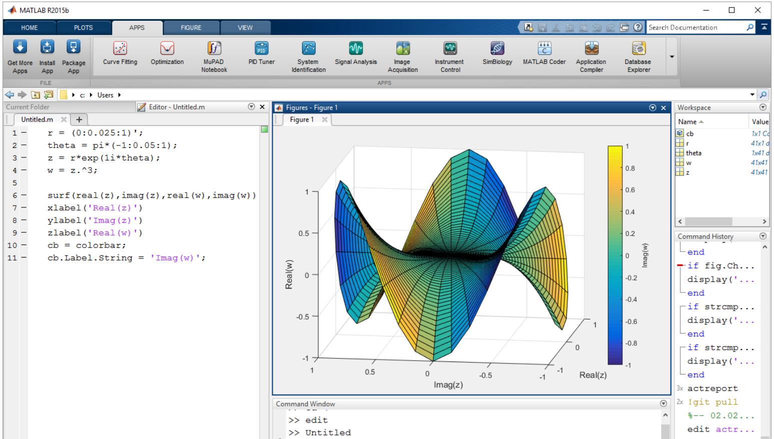
Duża część metod przedstawianych na wykładzie będzie wykorzystywała takie operacje, jak: **konstruowanie**, **modyfikowanie**, **rozwiązywanie układów równań liniowych** (równań macierzowych)

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

$$Ax = b$$

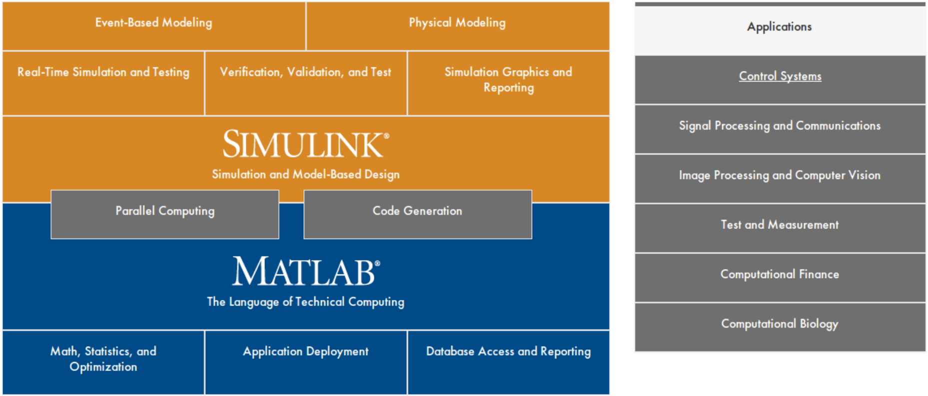
Podstawowe założenia (2)

- Na slajdach będą zamieszczane fragmenty kodów w języku **Matlab** (MATrix LABoratory) - program komputerowy będący interaktywnym środowiskiem do wykonywania obliczeń naukowych i inżynierskich oraz do tworzenia symulacji komputerowych (https://www.mathworks.com/products/pfo.html)



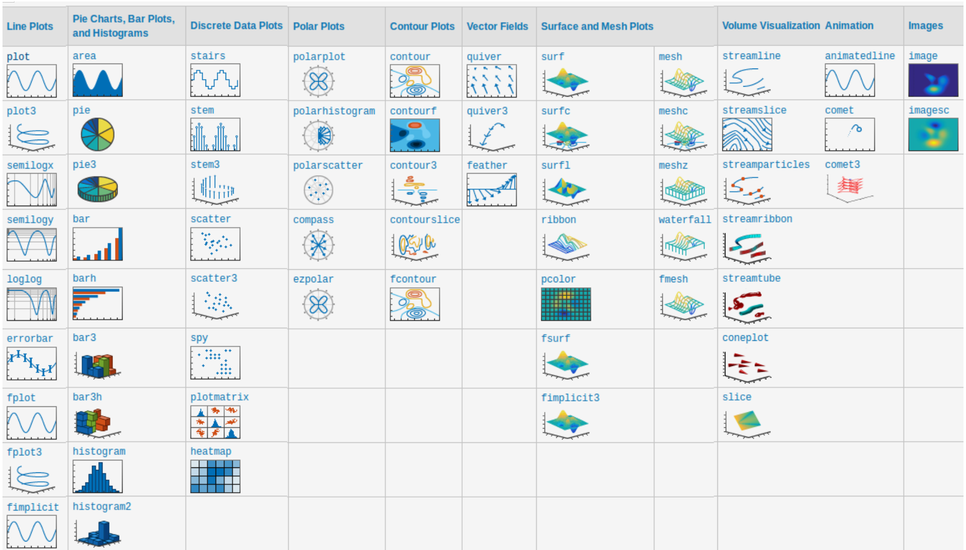
Podstawowe założenia (3)

- Przeznaczony głównie do numerycznych obliczeń **macierzowych**. Obecnie cechuje go duża liczba **funkcji bibliotecznych** oraz duże możliwości rozbudowy przez użytkownika za pomocą pisania własnych funkcji. Posiada on swój język programowania, co umożliwia pisanie w pełni funkcjonalnych programów działających w środowisku Matlaba



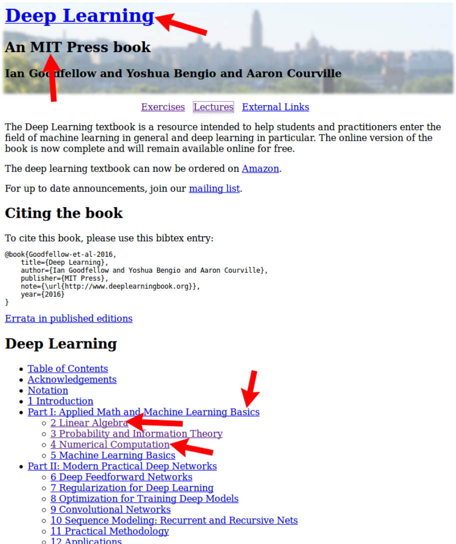
Podstawowe założenia (4)

- Rozbudowane narzędzia do **wizualizacji danych**.



Podstawowe założenia (5)

Podręcznik do **Deep learning** (http://www.deeplearningbook.org/)



Deep Learning

An MIT Press book

Ian Goodfellow and Yoshua Bengio and Aaron Courville

Exercises Lectures External Links

The Deep Learning textbook is a resource intended to help students and practitioners enter the field of machine learning in general and deep learning in particular. The online version of the book is now complete and will remain available online for free.

The deep learning textbook can now be ordered on Amazon.

For up to date announcements, join our mailing list.

Citing the book

To cite this book, please use this bibtex entry:

```
@book{Goodfellow-et-al-2016,
    title={Deep Learning},
    author={Ian Goodfellow and Yoshua Bengio and Aaron Courville},
    publisher={MIT Press},
    note={\url{http://www.deeplearningbook.org}},
    year={2016}
}
```

Errata in published editions

Deep Learning

- Table of Contents
- Acknowledgements
- Notation
- 1 Introduction
- Part I: Applied Math and Machine Learning Basics
  - 2 Linear Algebra
  - 3 Probability and Information Theory
  - 4 Numerical Computation
  - 5 Machine Learning Basics
- Part II: Modern Practical Deep Networks
  - 6 Deep Feedforward Networks
  - 7 Regularization for Deep Learning
  - 8 Optimization for Training Deep Models
  - 9 Convolutional Networks
  - 10 Sequence Modeling: Recurrent and Recursive Nets
  - 11 Practical Methodology
  - 12 Applications

- **MIT**
- Part I: Applied Math and Machine Learning **Basics**
- 2 **Linear Algebra**
- 4 **Numerical** Computation

Macierze (1)

1. Notacja macierzowa znacznie upraszcza wykonywanie operacji na układach równań liniowych,
2. Notacja macierzowa może być wprost stosowana implementacji algorytmów.
3. W Matlabie każda zmienna jest traktowana jak macierz.
4. **Macierz** o m wierszach i n kolumnach.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

5. **Rozmiar** macierzy: $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{A} \in \mathbb{C}^{k \times l}$ itp.
6. $\mathbb{R}$ - zbiór liczb rzeczywistych, $\mathbb{C}$ - zbiór liczb zespolonych
7. wektor o długości n: $\mathbf{a} \in \mathbb{R}^n$, skalar: a

Macierze (2)

1. **Macierz** o m wierszach i n kolumnach.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

2. i-ty wiersz (A_i^W):

$$\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix}$$

3. j-ta kolumna (A_j^K):

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

Macierze (3)

1. Macierz **kwadratowa**: n wierszy i n kolumn:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

2. Macierz **diagonalna** (Matlab: $diag(vect)$): macierz kwadratowa z elementami niezerowymi jedynie na diagonali ($a_{ij} = 0 \iff i \neq j$) :

$$\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

Macierze (4)

1. Macierz **wstęgowa**: niezerowe elementy jedynie na głównej diagonali i kilku sąsiednich przekątnych **(dyskretyzacja równań różniczkowych!)**

$$\begin{bmatrix} a_{11} & a_{12} & 0 & \cdots & 0 & 0 \\ a_{21} & a_{22} & a_{23} & \cdots & 0 & 0 \\ 0 & a_{32} & a_{33} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & 0 & \ddots & a_{n,n-1} & a_{n,n} \end{bmatrix}$$

2. Macierz **trójkątna**: macierz kwadratowa z elementami zerowymi powyżej (lub poniżej diagonali) (**faktoryzacja**!)

$$\begin{bmatrix} a_{11} & a_{1,2} & \cdots & a_{1,n} \\ 0 & a_{22} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

Macierze (5)

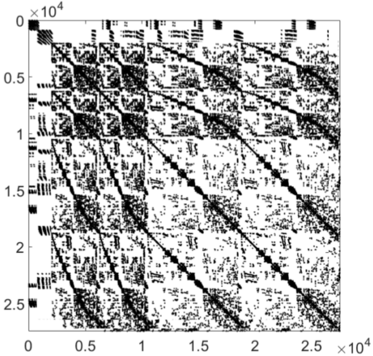
1. Macierz **jednostkowa** (oznaczana przez $\mathbf{I}$, Matlab: *eye*(n)): macierz diagonalna z wszystkimi elementami o wartości jeden.

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

2. Macierz **symetryczna**: macierz kwadratowa, dla której zachodzi zależność: $a_{ij} = a_{ji}$ (osobne procedury dla macierzy symetrycznych i niesymetrycznych)

Macierze (6)

Macierze **rzadkie** (Matlab - format *sparse*, macierze gęste - format *full*)



- w wielu problemach występują macierze, w których nie więcej, niż 1% elementów ma wartość niezerową
- Można je opisać w formacie **rzadkim**: (nr_wiersza, nr_kolumny, wartość_niezerowa)
- pozwala to na znaczną **oszczędność** pamięci RAM i na **przyspieszanie** obliczeń.

Działania na Macierzach (1)

1. **Dodawanie** macierzy **A** i **B** tylko jeżeli mają ten sam rozmiar $(n \times m)$:

$$\mathbf{C} = \mathbf{A} + \mathbf{B} \tag{1}$$

gdzie $c_{ij} = a_{ij} + b_{ij}$

2. **Odejmowanie** macierzy **A** i **B** tylko jeżeli mają ten sam rozmiar $(n \times m)$:

$$\mathbf{C} = \mathbf{A} - \mathbf{B} = \mathbf{A} + (-1)\mathbf{B} \tag{2}$$

gdzie $c_{ij} = a_{ij} - b_{ij}$

3. **Mnożenie** macierzy **A** przez skalar γ:

$$\mathbf{C} = \gamma\mathbf{A} \tag{3}$$

gdzie $c_{ij} = \gamma a_{ij}$

Działania na Macierzach (2)

1. W wyniku **mnożenia** macierzy $\mathbf{A} \in \mathbb{R}^{n \times m}$ i $\mathbf{B} \in \mathbb{R}^{k \times l}$ powstaje macierz $\mathbf{C} \in \mathbb{R}^{n \times l}$, przy czym m musi być równe k (liczba kolumn macierzy **A** musi być równa liczbie wierszy macierzy **B**).
$$\mathbf{C} = \mathbf{AB} \tag{4}$$
2. gdzie element c_{ij} jest wynikiem **iloczynu skalarnego** (o tym szerzej za moment)
$$\mathbf{c}_{ij} = \mathbf{A}_i^W \mathbf{B}_j^K = \sum_{l=1}^{m} a_{il} b_{lj} \tag{5}$$
3. Macierz Odwrotna do **A**:
$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{I} \tag{6}$$

Działania na Macierzach (3)

1. **Przykład** mnożenia macierzy:

$$\begin{bmatrix} 1 & 2 & 7 \\ \underline{\mathbf{8}} & \underline{\mathbf{1}} & \underline{\mathbf{2}} \\ 9 & 5 & 3 \end{bmatrix} \begin{bmatrix} 6 & \underline{\mathbf{2}} \\ 5 & \underline{\mathbf{1}} \\ 2 & \underline{\mathbf{1}} \end{bmatrix} = \begin{bmatrix} 30 & 11 \\ 57 & \underline{\mathbf{19}} \\ 85 & 26 \end{bmatrix}$$

2. Mnożenie macierzy w Matlabie:

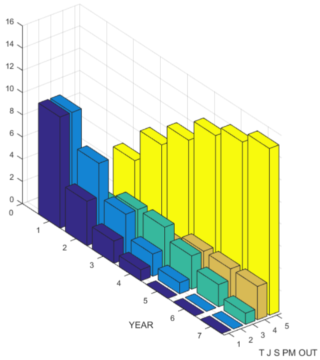
```
A = [1 2 7; 8 1 2; 9 5 3];
B = [6 2; 5 1; 2 1];
C = A * B
```

3. Dodawanie i mnożenie macierzy **A** i **B** (**C**) ma następujące **własności**:
   - $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
   - $\mathbf{AB} \neq \mathbf{BA}$
   - Dodawanie jest łączne: $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$
   - Mnożenie jest łączne: $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
   - Mnożenie jest rozdzielne względem dodawania: $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
   - $\mathbf{AI} = \mathbf{IA} = \mathbf{A}$, gdzie **I** to macierz jednostkowa. Odpowiednik $a \cdot 1 = 1 \cdot a = a$
   - $\mathbf{A} + \mathbf{0} = \mathbf{0} + \mathbf{A} = \mathbf{A}$, gdzie **0** to macierz zer. Odpowiednik $a + 0 = 0 + a = a$

Działania na Macierzach (4) - Przykład

Macierze Marcova

	tester	junior	senior	project menager	OUT
tester	0.1	0	0	0	0
junior	0.5	0.1	0	0	0
senior	0	0.7	0.7	0	0
project menager	0	0	0.3	0.8	0
OUT	0.4	0.2	0.1	0.2	1



Działania na Macierzach (5)

1. **Transpozycja** macierzy **A** oznaczona jest przez $\mathbf{A}^T$ (w Matlabie: A' lub $A.'$)
2. *zamiana* wierszy i kolumn, w taki sposób, że element macierzy oryginalnej a_{ij} po transpozycji ma wartość a_{ji}.
3. Dla macierzy **kwadratowej** jest to zamiana elementów leżących **symetrycznie** względem **diagonali**.
4. Dla macierzy **niekwadratowej** poddanej transpozycji zamianie ulega liczba kolumn i wierszy.
5. **Przykłady:**

$$\begin{bmatrix} 1 & 2 & 7 \\ 8 & 1 & 2 \\ 9 & 5 & 3 \end{bmatrix}^T = \begin{bmatrix} 1 & 8 & 9 \\ 2 & 1 & 5 \\ 7 & 2 & 7 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 2 & 7 \\ 8 & 1 & 2 \end{bmatrix}^T = \begin{bmatrix} 1 & 8 \\ 2 & 1 \\ 7 & 2 \end{bmatrix}$$

Działania na Macierzach (6)

Własności transpozycji:

- $(\mathbf{A}^T)^T = \mathbf{A}$
- $(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$
- Transpozycja $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$
- **Symetryzacja** kwadratowej macierzy $\mathbf{A} \in \mathbb{R}^{n \times n}$:

```
A = rand(3,3);
A_sym = (A + A')/2;
>> A =
    0.32529    0.25397    0.10845
    0.47230    0.73288    0.46246
    0.57019    0.41153    0.34496
>> A_sym =
    0.32529    0.36313    0.33932
    0.36313    0.73288    0.43699
    0.33932    0.43699    0.34496
```

Działania na Macierzach (7)

Własności transpozycji:

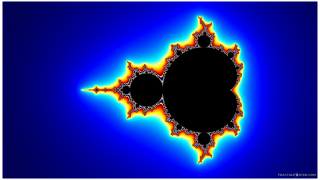
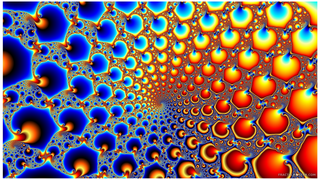
- Sprawdzenie, czy kwadratowa macierz $\mathbf{A} \in \mathbb{R}^{n \times n}$ jest symetryczna:

```
max(max(abs( A - A')))
    ans =  0.66819
issymmetric(A)
     ans = 0
max(max(abs( A_sym - A_sym')))
     ans = 0
issymmetric(A_sym)
     ans = 1
max(max(abs( B_sym - B_sym')))
     ans = 9.9920e-16
issymmetric(B_sym)
     ans = 0
```

Działania na Macierzach - dygresja: Fraktale (1)

Fraktal - zbiór, który posiada poniższe cechy:

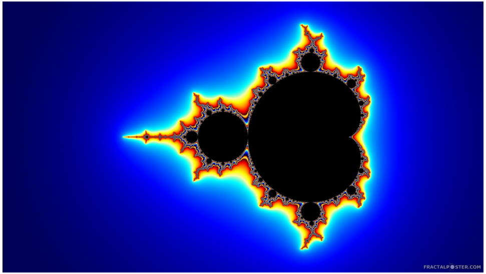
- **Samo-podobny** tzn. że jego części (po dowolnym zbliżeniu) są podobne do całości,
- ukazujący **subtelne detale** nawet w wielokrotnym powiększeniu,
- struktura fraktala nie daje się łatwo opisać w języku tradycyjnej geometrii euklidesowej,
- definiuje się go za pomocą prostej **zależności rekurencyjnej**.
- ...

Działania na Macierzach - dygresja: Fraktale (2)

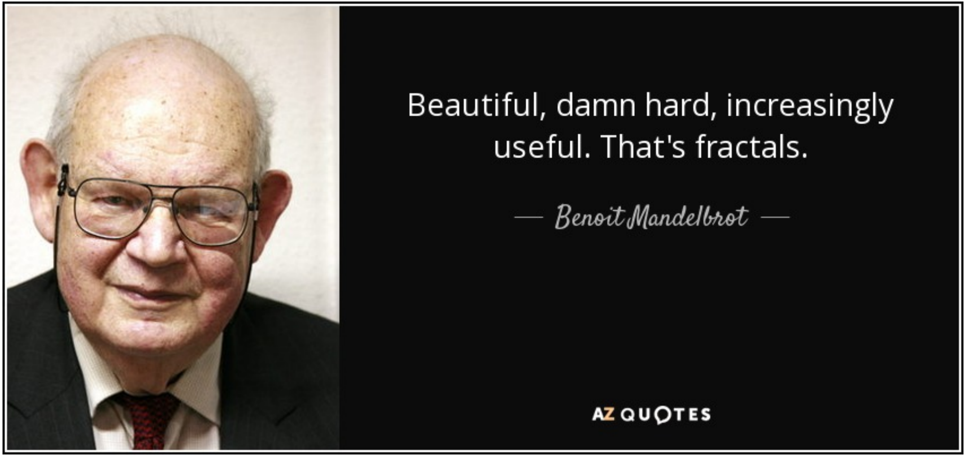
Benoit Mandelbrot urodził się w 1924r. w Warszawie. W wieku dwunastu lat wraz z rodziną wyemigrowali do Francji. Ukończył studia na École Polytechnique, a następnie wyjechał do Stanów Zjednoczonych na Kalifornijski Instytut Technologiczny. W 1958 roku rozpoczął współpracę z IBM. Fraktal **Mandelbrota**:

- dla każdego punktu należącego do płaszczyzny zespolonej $p \in \mathbb{C}$
- tworzymy zależność rekurencyjną: $z_0 = 0$, $z_{i+1} = z_i^2 + p$,
- Zbiór Mandelbrota tworzą punkty, dla których powyższa zależność nie dąży do nieskończoności. W praktyce: $\forall_{n \in \mathbb{N}} |z_n| < 2$



- Dla $p = 0.2 + 0.2i$: $z_0 = 0$
- $z_1 = 0.2 + 0.2i$
- $z_2 = 0.2 + 0.28i$
- $z_3 = 0.1616 + 0.312i$
- $z_4 = 0.128 + 0.301i$ ...
- $|z_n| < 2$

Działania na Macierzach - dygresja: Fraktale (3)



http://www.azquotes.com/picture-quotes/quote-beautiful-damn-hard-increasingly-useful-that-s-fractals-benoit-mandelbrot-94-54-66.jpg

Operacje na wektorach (1)

1. Rozpatrzmy dwa wektory **a** i **b** $\in \mathbb{R}^3$:

$$\mathbf{a} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}, \ \mathbf{b} = \begin{bmatrix} 5 \\ 4 \\ 7 \end{bmatrix}$$

2. **Iloczyn skalarny (ang. *dot product*)** (!):

$$\mathbf{a}^T \cdot \mathbf{b} = \mathbf{b}^T \cdot \mathbf{a} = \sum_{i=1}^{3}(a_i b_i) = \begin{bmatrix} 1, 3, 5 \end{bmatrix} \begin{bmatrix} 5 \\ 4 \\ 7 \end{bmatrix} = 1 \cdot 5 + 3 \cdot 4 + 5 \cdot 7 = 52$$

3. Iloczyn diadyczny (ang. *outer product*):

$$\mathbf{ab}^T = (\mathbf{ba}^T)^T = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} \begin{bmatrix} 5, 4, 7, \end{bmatrix} = \begin{bmatrix} 1 \cdot 5 & 1 \cdot 4 & 1 \cdot 7 \\ 3 \cdot 5 & 3 \cdot 4 & 3 \cdot 7 \\ 5 \cdot 5 & 5 \cdot 4 & 5 \cdot 7 \end{bmatrix} = \begin{bmatrix} 5 & 4 & 7 \\ 15 & 12 & 21 \\ 25 & 20 & 35 \end{bmatrix}$$

Operacje na wektorach (2)

1. **Norma** wektora, czyli jego **długość** (jedno z zastosowań iloczynu skalarnego):

$$\mathbf{a} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}, \|\mathbf{a}\| = \sqrt{a_1^2 + a_2^2 + a_3^3} = \sqrt{\mathbf{a}^T\mathbf{a}}$$

2. **Kąt** miedzy dwoma wektorami:

$$\mathbf{a}^T\mathbf{b} = \|\mathbf{a}\|\|\mathbf{b}\|cos(\theta_{ab})$$

$$\theta_{ab} = arccos\frac{\mathbf{a}^T\mathbf{b}}{\|\mathbf{a}\|\|\mathbf{b}\|}$$

3. Wektory **jednostkowe** (!)

$$\|\mathbf{i}\| = 1$$

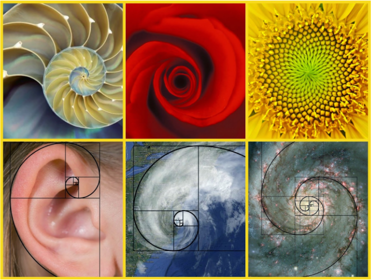
4. Wektory **ortogonalne** (Ortogonalizacja Grama Schmidta!)

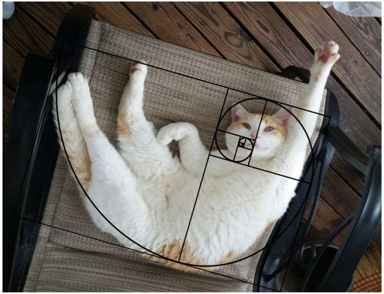
$$\mathbf{a}^T\mathbf{b} = 0$$

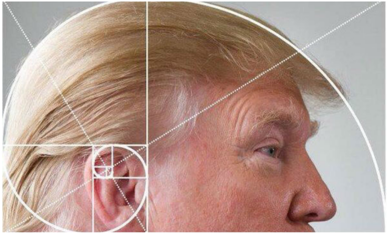
Działania na Macierzach - dygresja 2

Ciąg Fibanacciego:

$$f(n) = \begin{cases} 0 & n = 0 \\ 1 & n = 1 \\ f(n-1) + f(n-2) & \text{otherwise} \end{cases}$$







www.cems.uvm.edu, http://www.livetradingnews.com, https://www.iizcat.com

Dziękuję