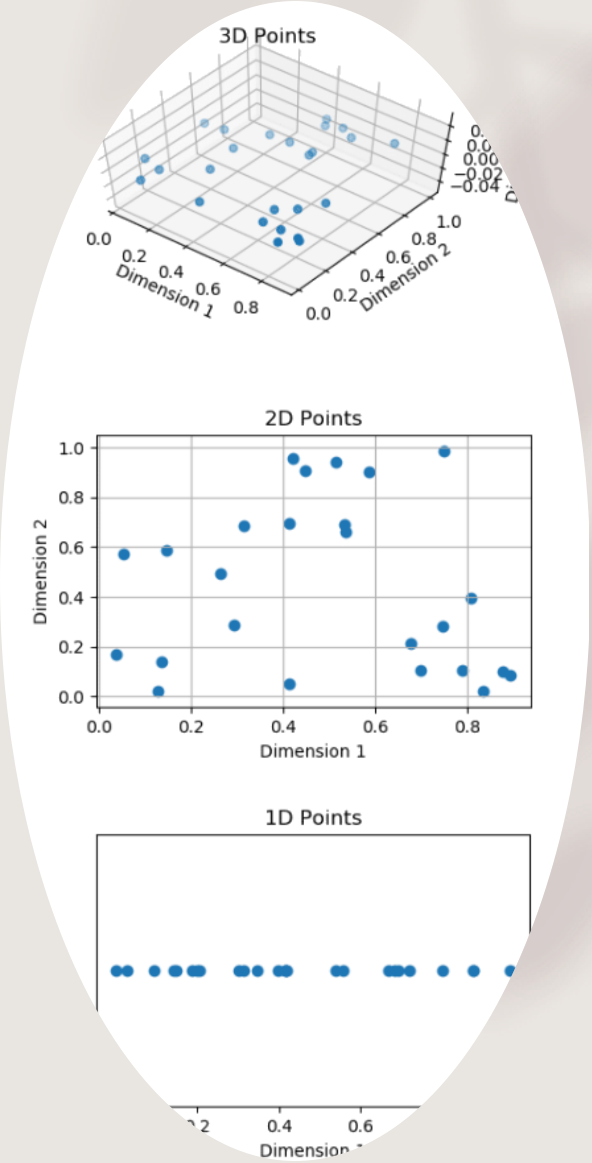
Redukcja wymiarowości

Wprowadzenie



Plan na dzisiaj

- Motywacja - klątwa wymiarowości
- Przegląd technik redukcji wymiarowości
- PCA
- SVD

Klątwa wymiarowości

Im więcej wymiarów mają nasze dane tym więcej danych potrzebujemy do zachowania tej samej wydajności modelu. I ten wzrost ma charakter wykładniczy.

Wymiarów ?

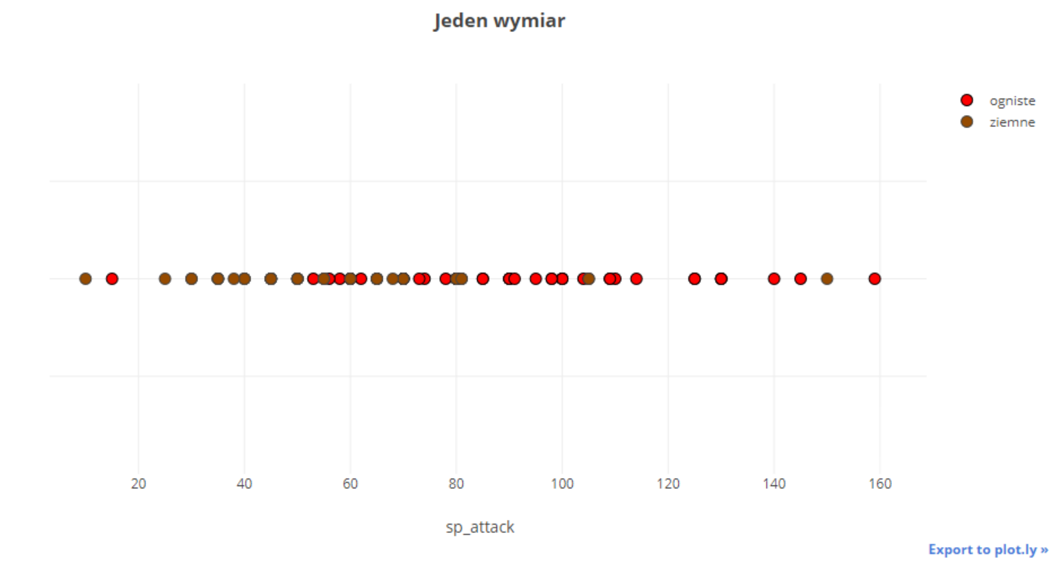
Wymiary
(za https://miroslawmamczur.pl/)

Popatrzmy na dwa typy pokemonów:
ogniste i ziemne

Wymiary

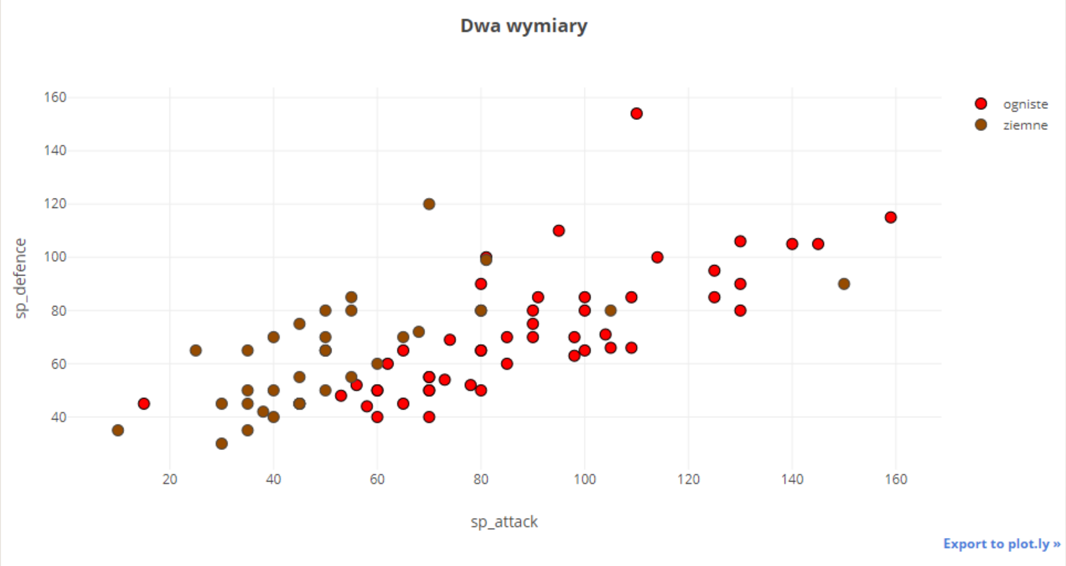
(za https://miroslawmamczur.pl/)

Nanieśmy na wykres ich siłę



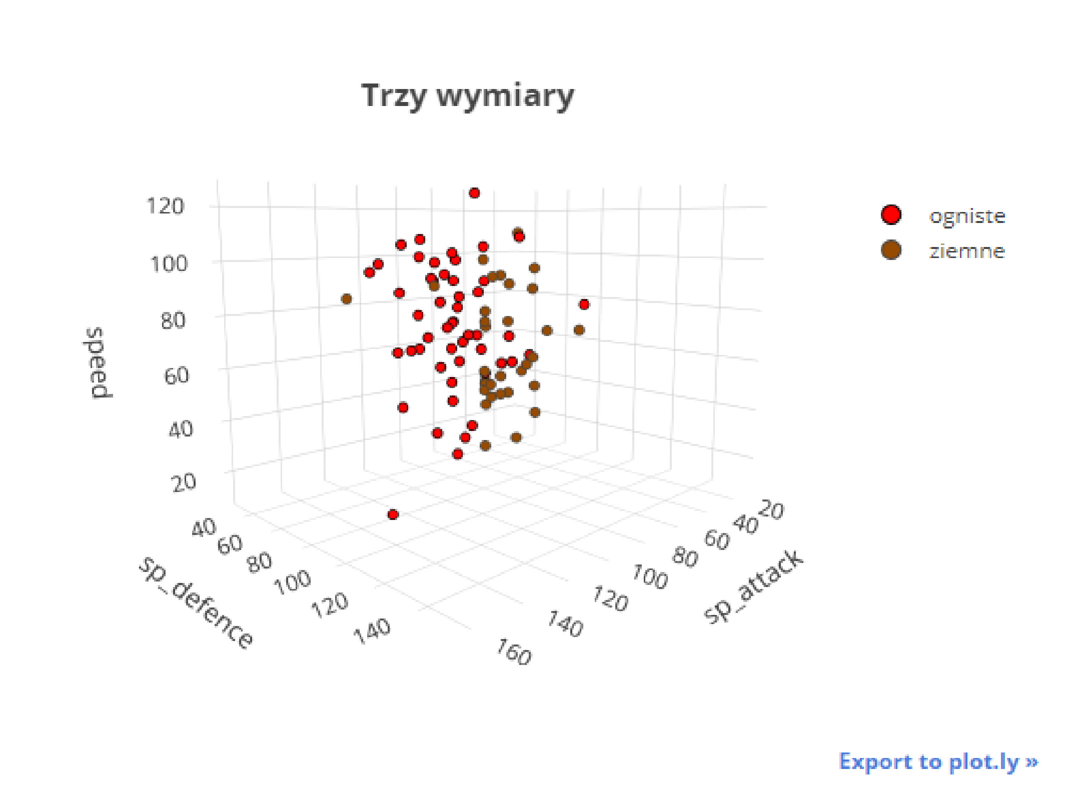
Wymiary
(za https://miroslawmamczur.pl/)

Nanieśmy na wykres ich siłę i wytrzymałość



Wymiary
(za https://miroslawmamczur.pl/)

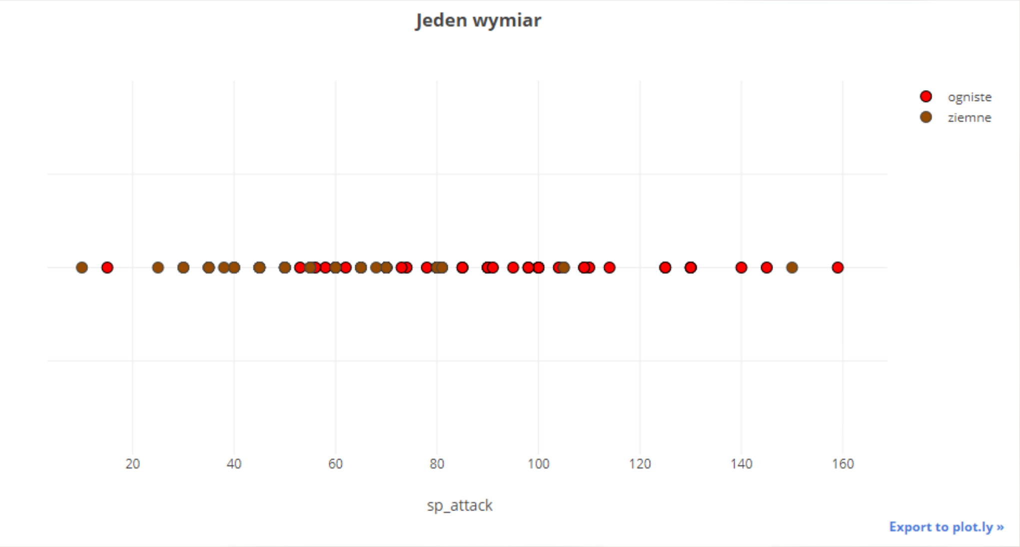
Nanieśmy na wykres ich siłę, wytrzymałość i szybkość

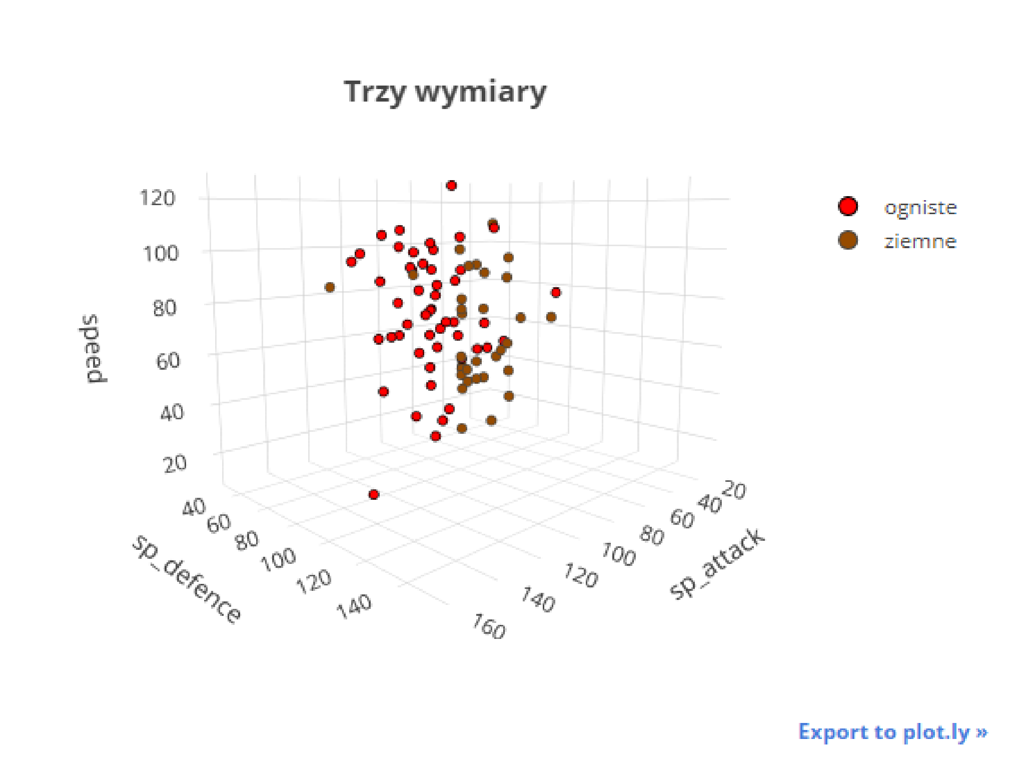


Wymiary

(za https://miroslawmamczur.pl/)

Odległość pomiędzy poszczególnymi pokemonami zwiększa się razem ze wzrostem liczby wymiarów. Oznacza to, że w wyższych wymiarach próbki będą coraz bardziej od siebie oddalone i modelowi będzie coraz trudniej poprawnie je aproksymować. Dla zachowania swojej dokładności będzie potrzebował coraz większej liczby próbek, żeby wypełnić nimi dostępną przestrzeń wartości.



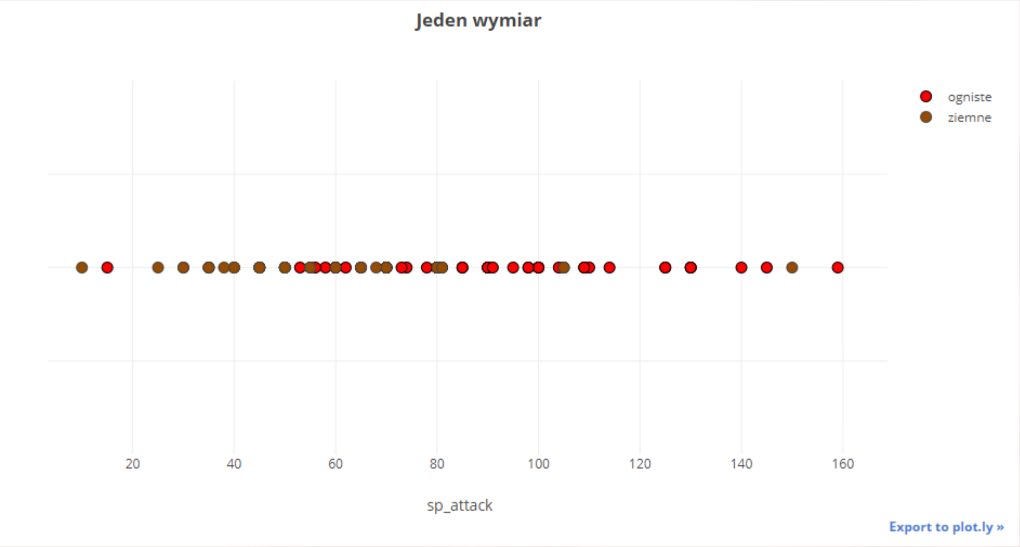


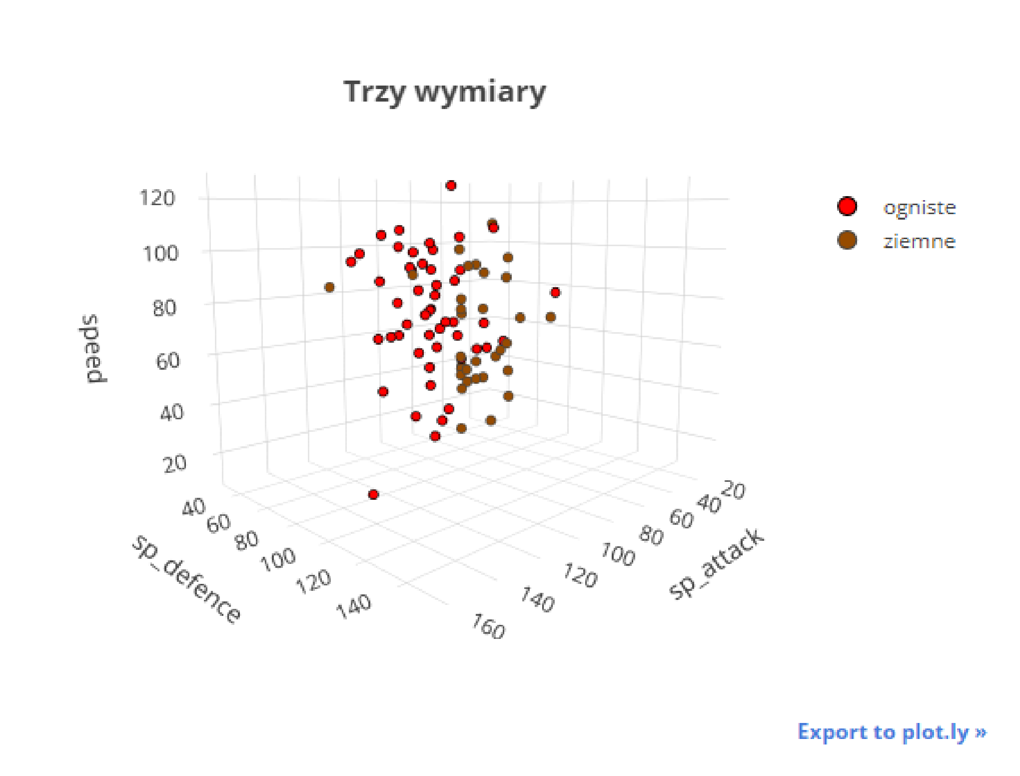
Klątwa wymiarowości

(rozrzedzenie danych)

W przypadku dwóch punktów wybranych w jednostkowym kwadracie (dwa wymiary) średnia odległość między tymi punktami wyniesie 0.52, w jednostkowym sześcianie (trzy wymiary) 0.62, a w przypadku jednostkowego 1.000.000-wymiarowego hipersześcianu średnia odległość pomiędzy nimi wyniesie 408.25.

(w obliczeniach przyjęto, że punkty losujemy z rozkładu płaskiego, czyli że są rozmieszczone równomiernie)



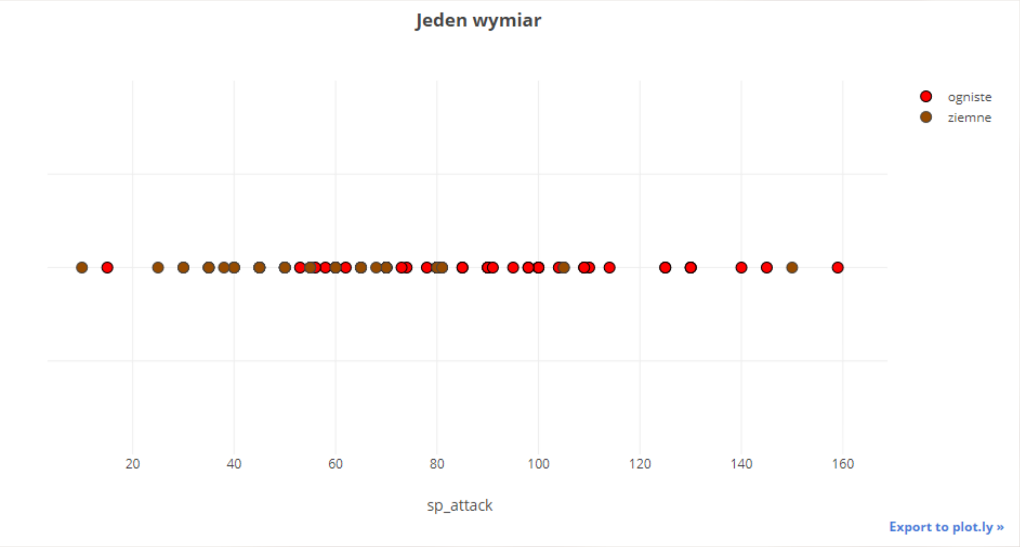


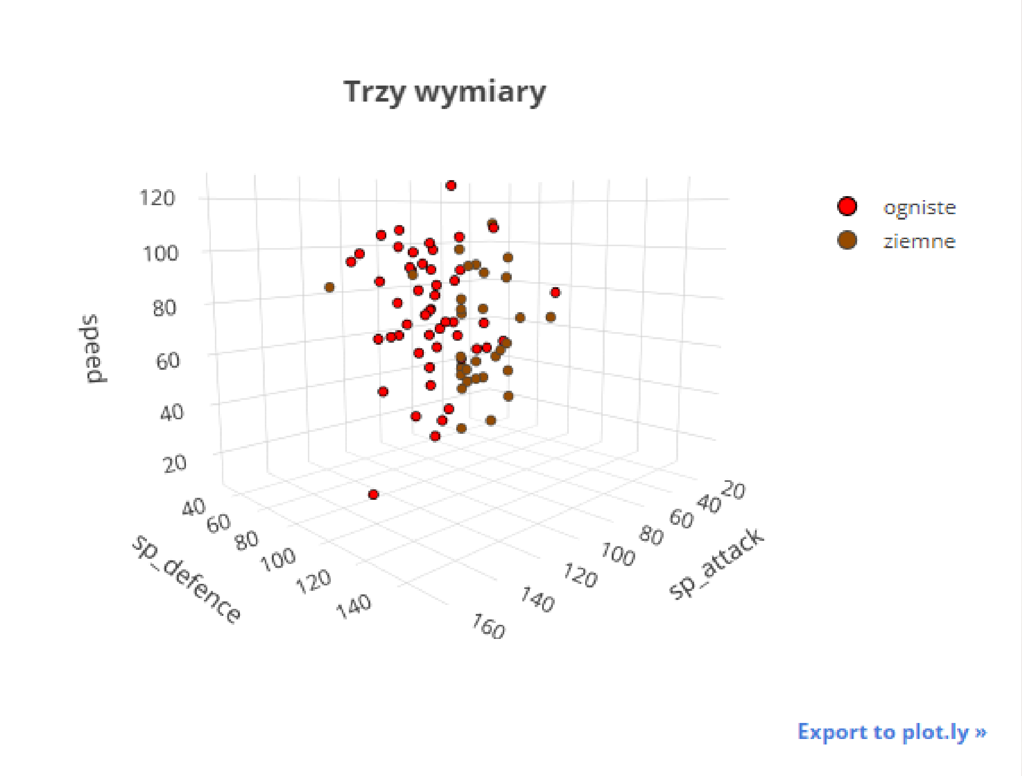
Klątwa wymiarowości

(przetrenowanie modelu)

Rozrzedzenie danych to nie jedyny problem wielowymiarowych zbiorów.

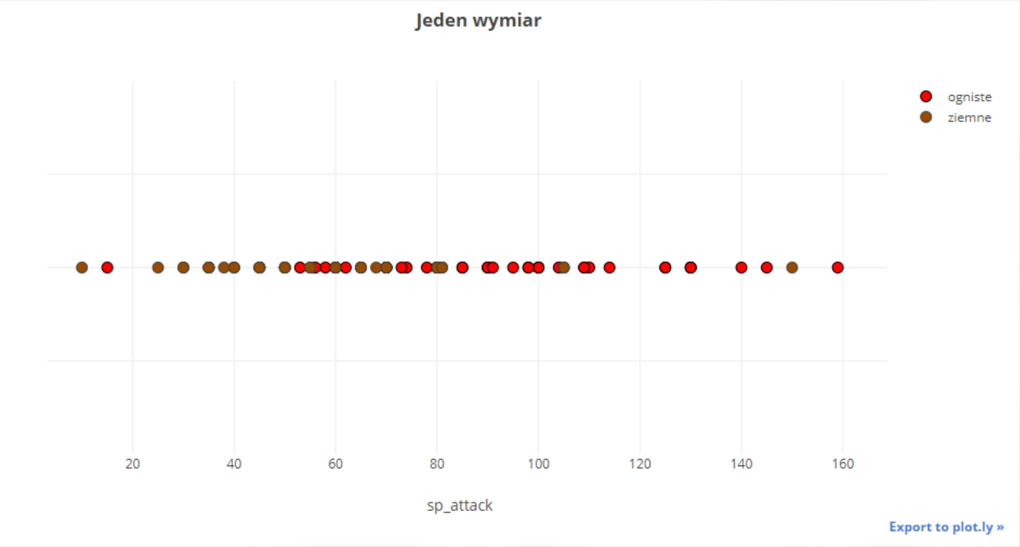
Nowe próbki będą znajdowały się daleko od przykładów uczących, przez co wyliczane prognozy będą znacznie mniej pewne niż uzyskane w przestrzeni o mniejszej liczbie wymiarów, ponieważ będą one bazować na znacznie większych ekstrapolacjach. Innymi słowy, wraz ze wzrostem liczby wymiarów zwiększa się ryzyko **przetrenowania modelu**.

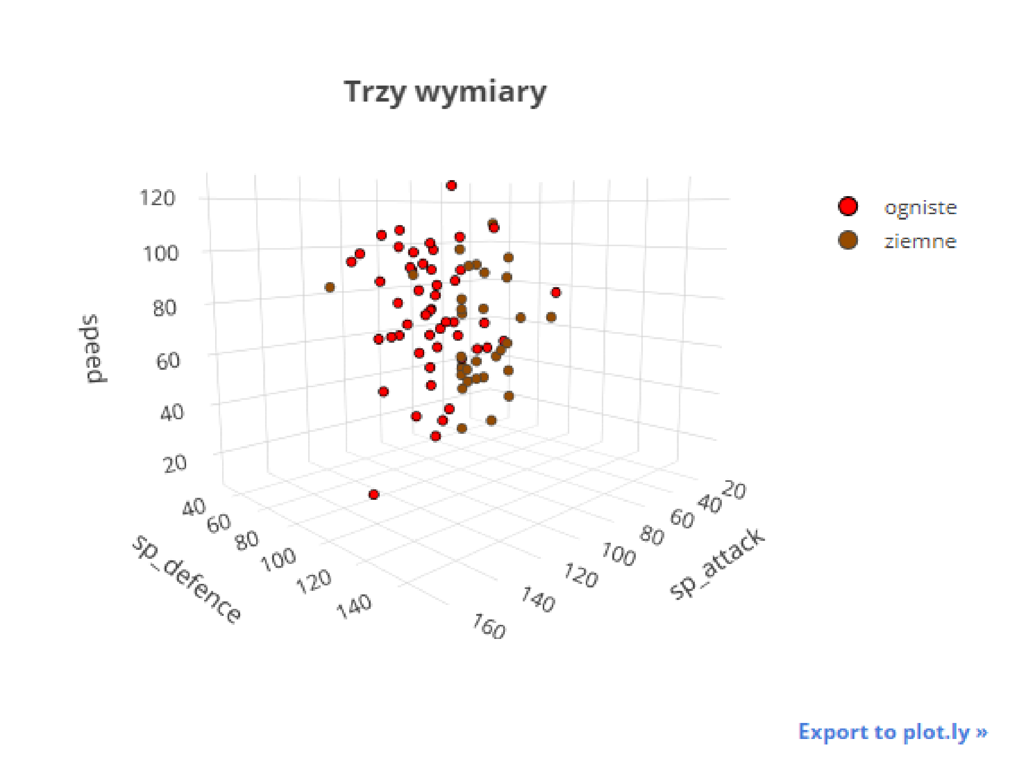




Klątwa wymiarowości

Rozwiązanie ?

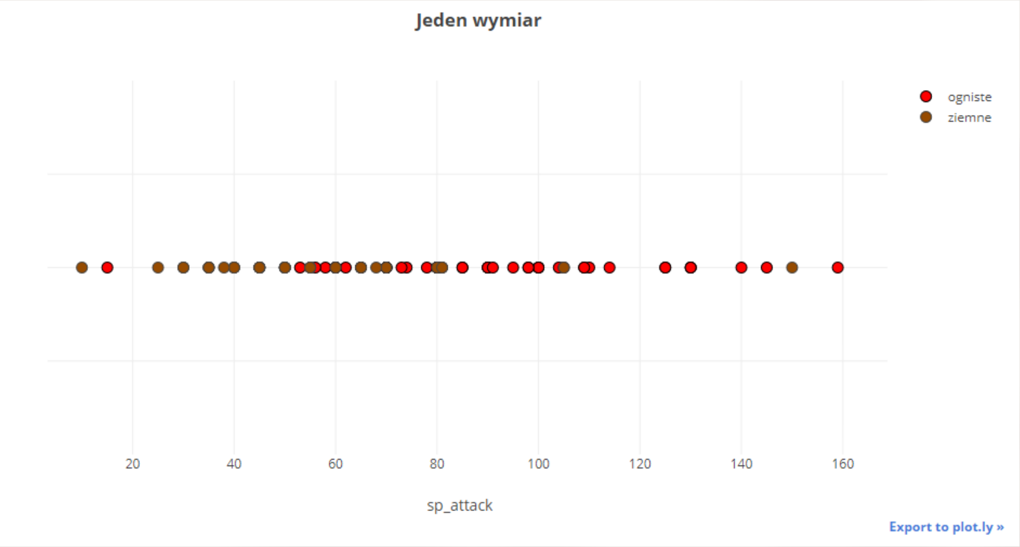


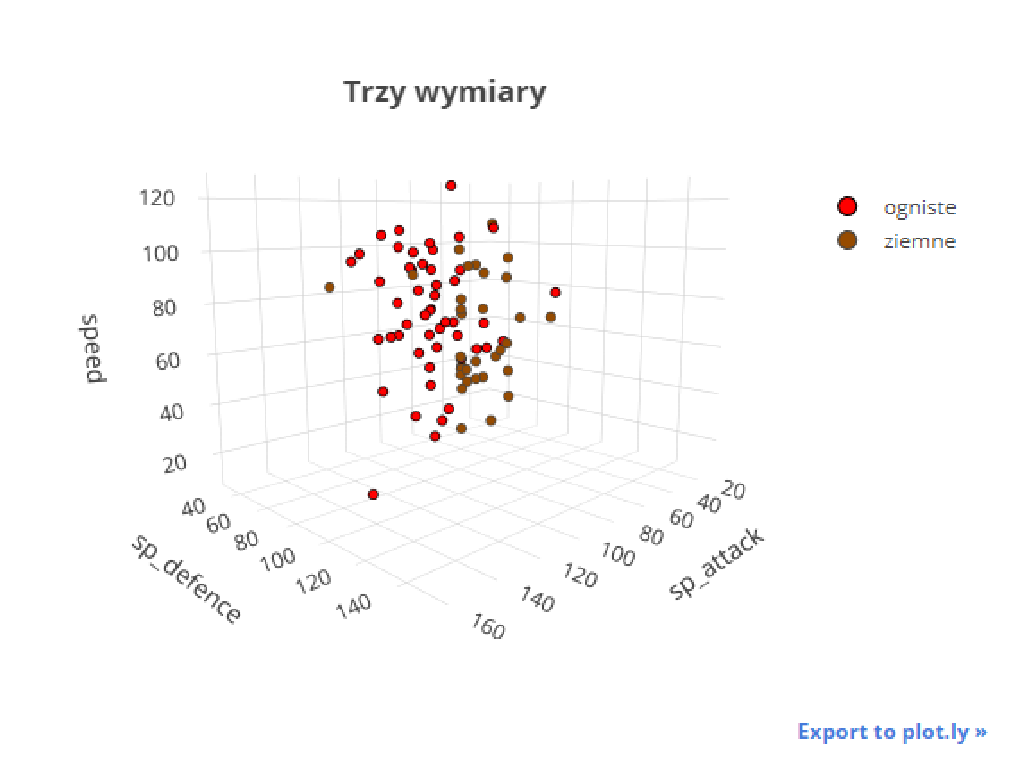


Klątwa wymiarowości

Rozwiązanie

A może dodać więcej danych?



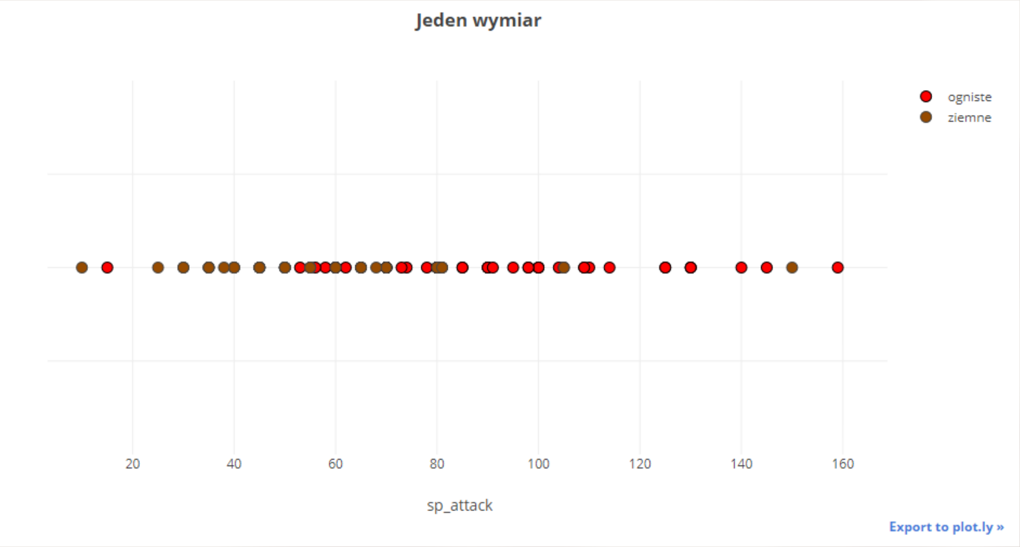


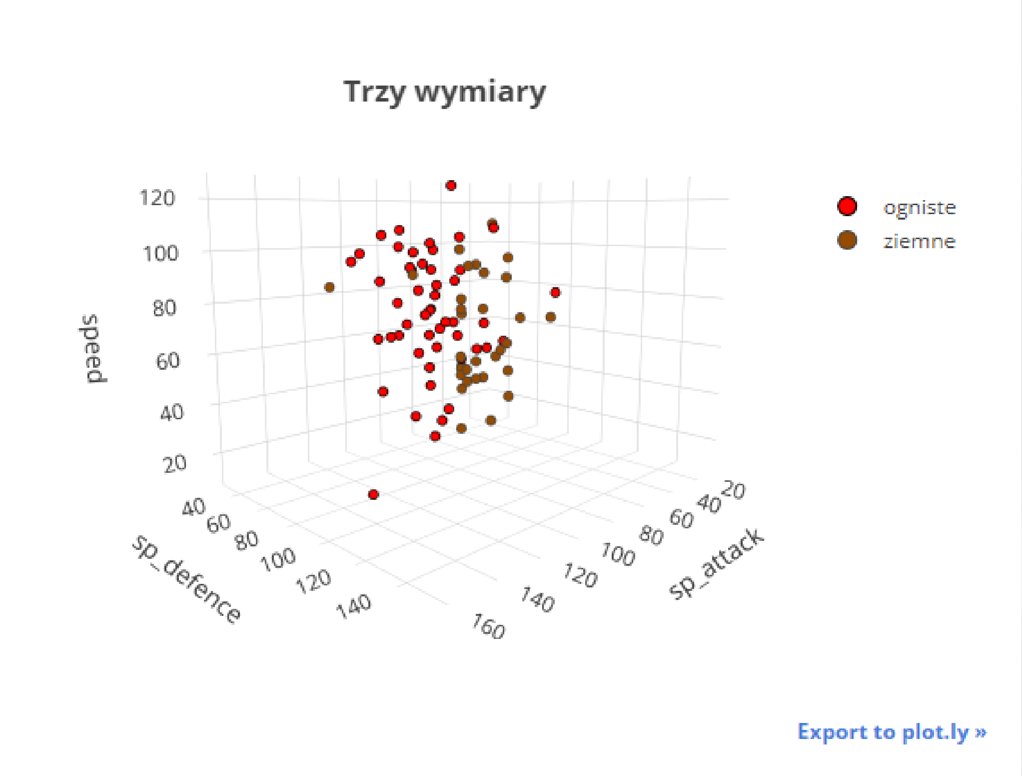
Klątwa wymiarowości

Rozwiązanie

A może dodać więcej danych?

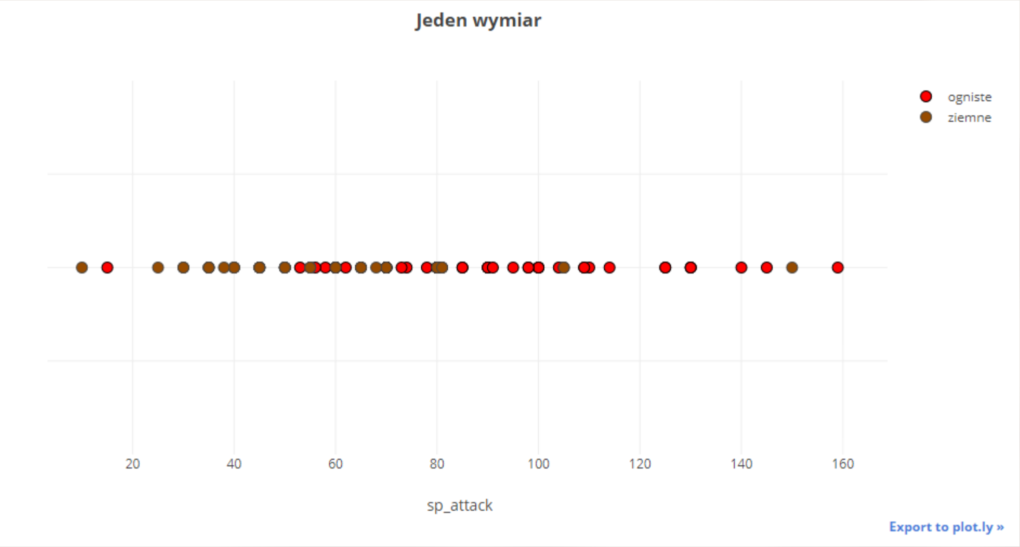
Niestety liczba przykładów potrzebnych do osiągnięcia zakładanej gęstości wzrast wykładniczo wraz z liczbą wymiarów. Oznacza to, że przy 100 wymiarach (a nie są to duże liczby) w celu wyuczenia próbek znajdujących się średnio w odległości 0.1 od siebie wymagana byłaby liczba przykładów przekraczająca liczbę atomów znajdujących się w obserwowalnym Wszechświecie (10**80).

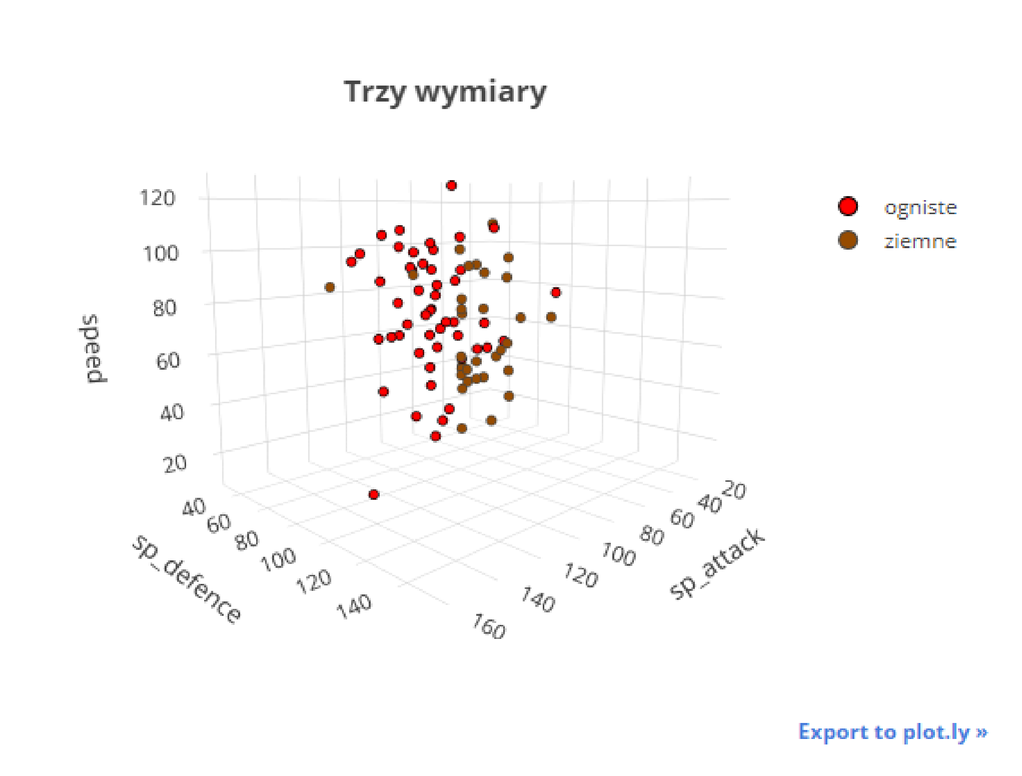




Klątwa wymiarowości

Rozwiązanie ?

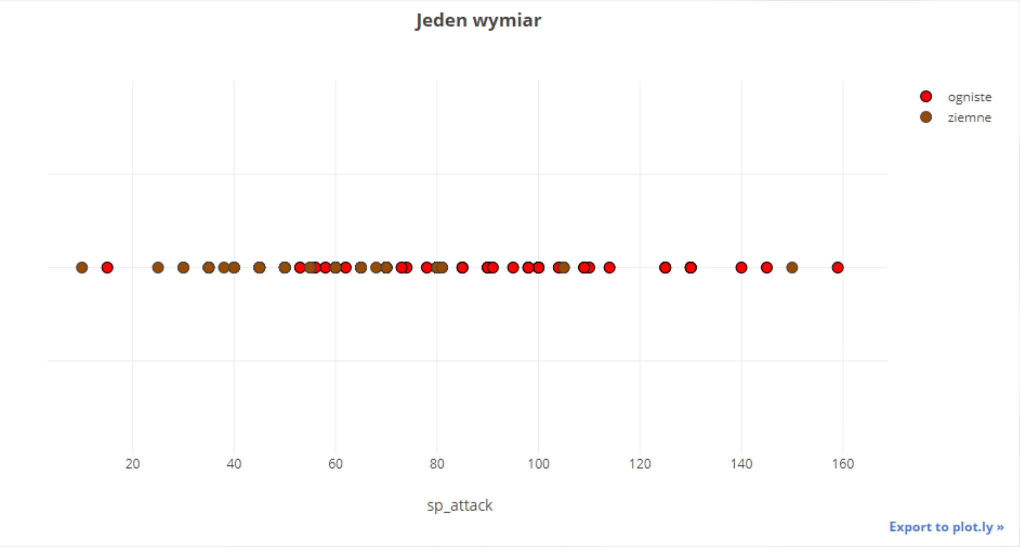


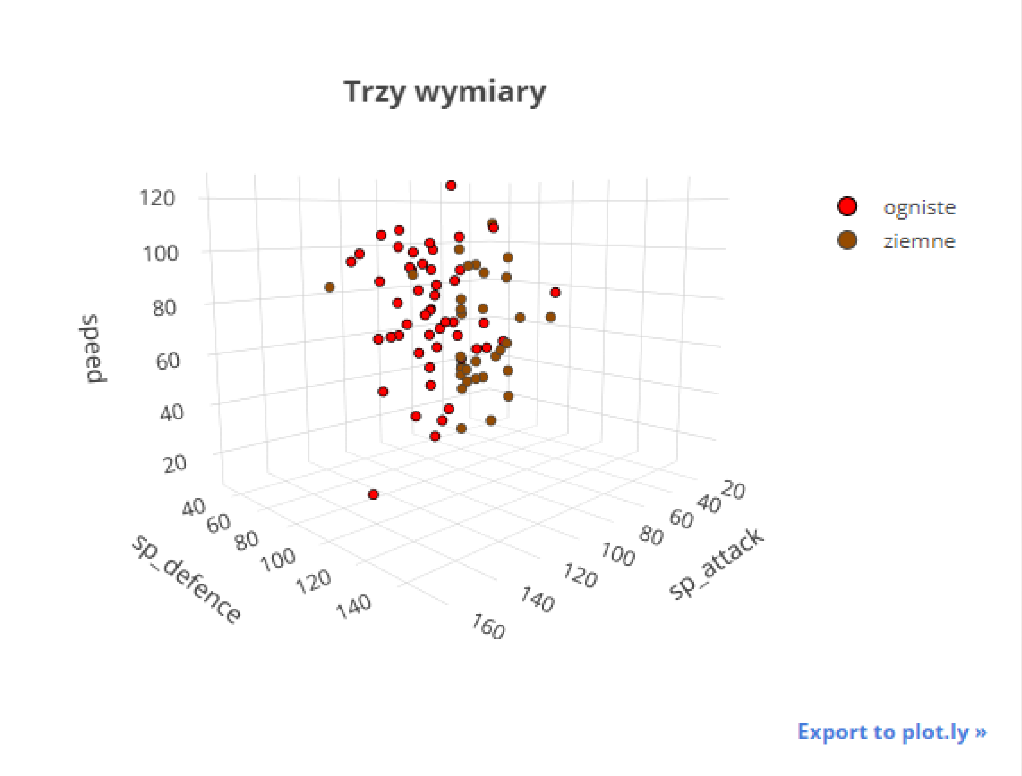


Klątwa wymiarowości

Rozwiązanie

To może zmniejszyć liczbę wymiarów?





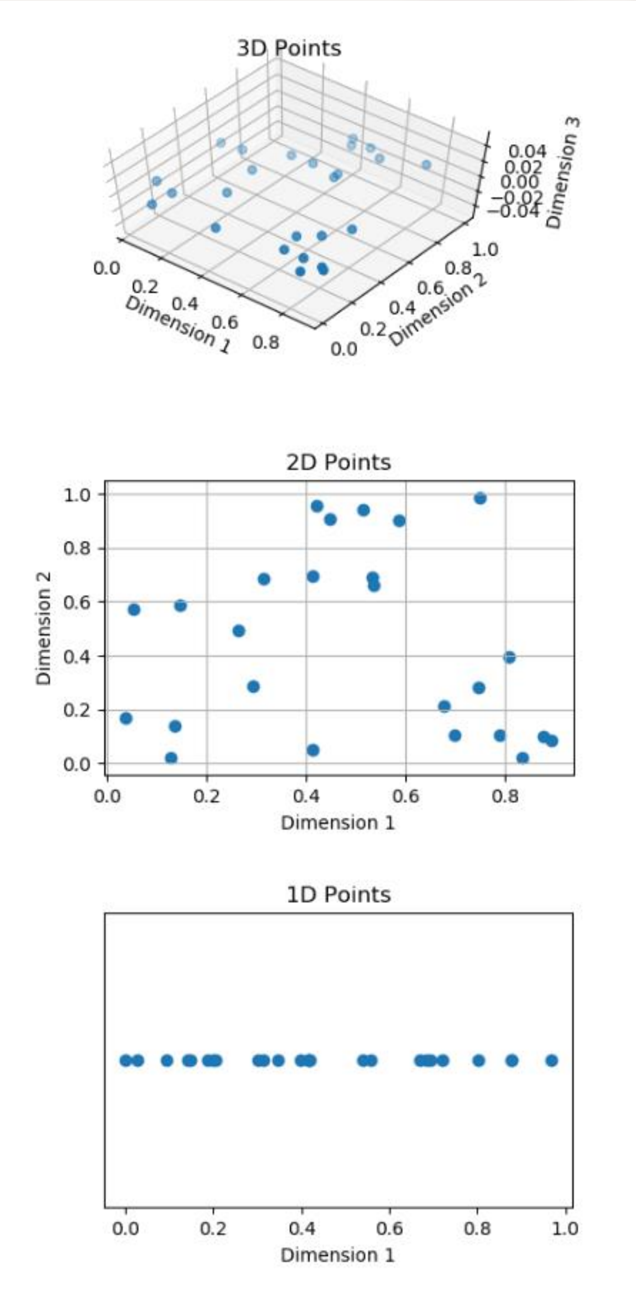
Klątwa wymiarowości

Rozwiązanie

To może zmniejszyć liczbę wymiarów?

Czyli przejść jeszcze raz tą samą drogą, którą tu doszliśmy tylko w przeciwnym kierunku.

Tylko jak ?



Klątwa wymiarowości

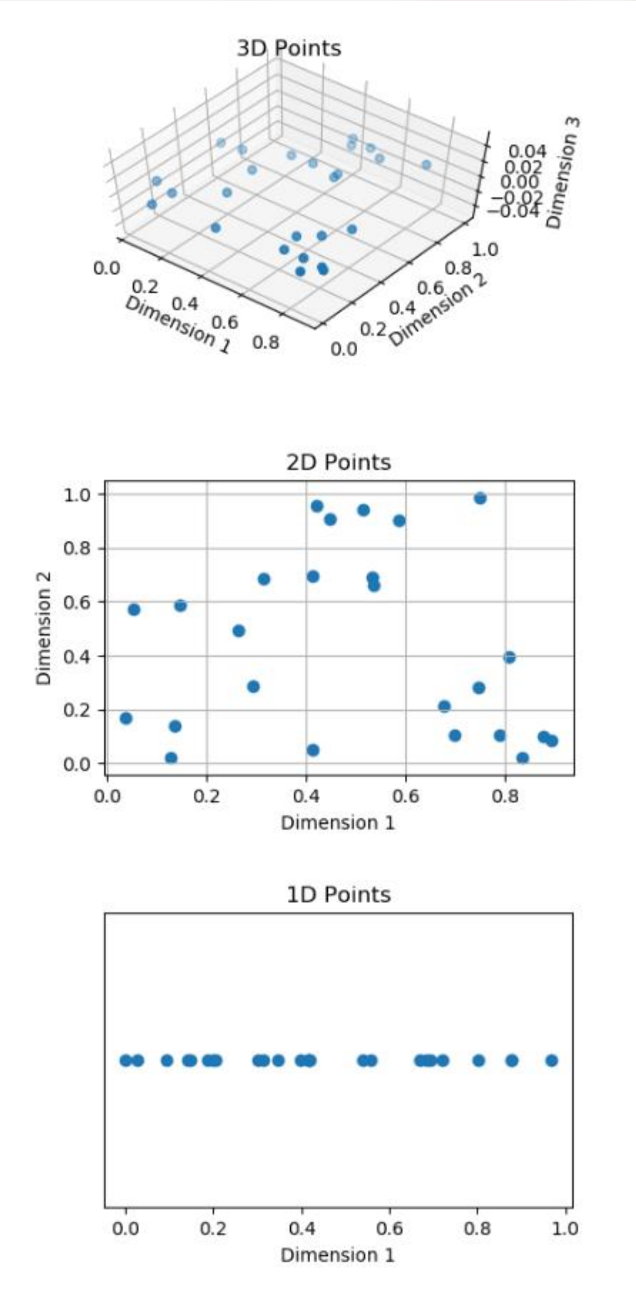
Rozwiązanie

To może zmniejszyć liczbę wymiarów?

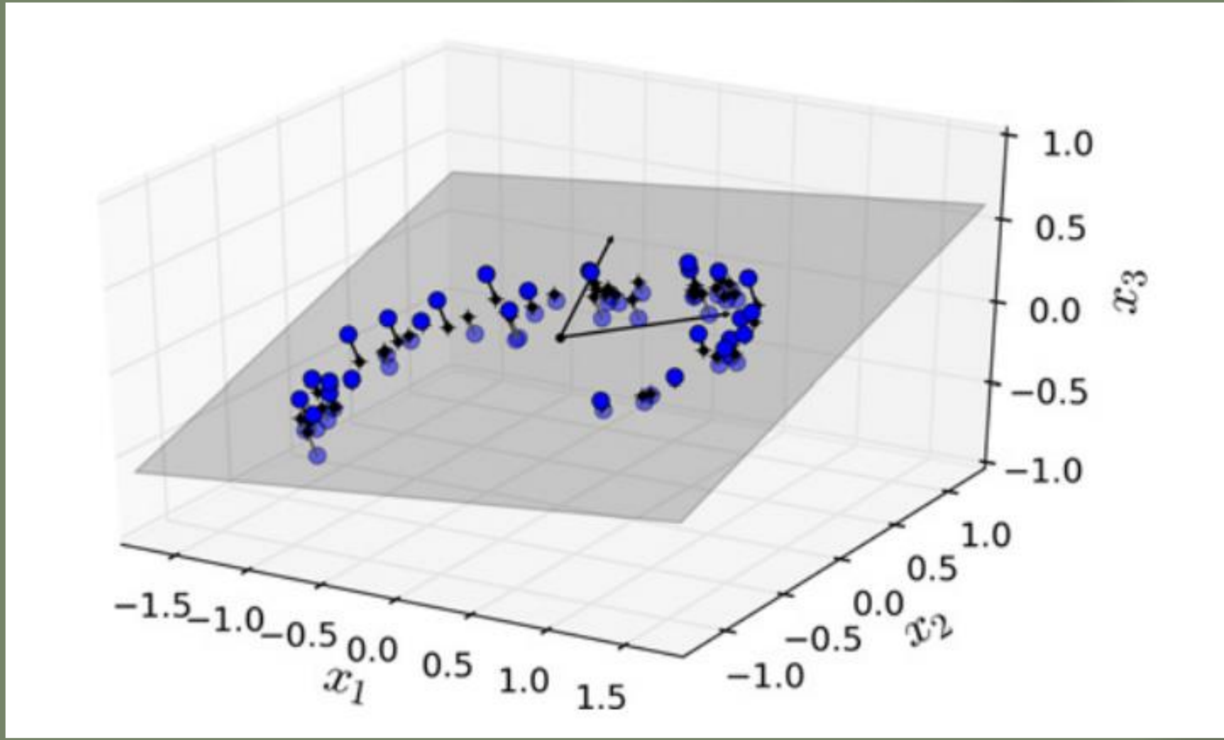
Czyli przejść jeszcze raz tą samą drogą, którą tu doszliśmy tylko w przeciwnym kierunku.

Tylko jak ?

Istnieje wiele technik



Techniki redukcji wymiarowości



Techniki redukcji wymiarowości

- PCA (*ang. Principle Component Analysis*)
- SVD (*ang. Single Value Decomposition*)
- ICA (*ang. Independent Component Analysis*)
- Rzutowanie losowe (*ang. Random projections*)
- Isomap (*ang. Isometric mapping*)
- MDS (*ang. MultiDimensional Scaling*)
- LDA (*ang. Linear Discriminant Analysis*)
- GDA (*ang. Generalized Discriminant Analysis*)
- Autoenkodery (*ang. Autoencoders*)
- Metody rozmaitościowe (*ang. Mainfold learning algorithms*):
  - FA (*ang. Factor Analysis*)
  - LLA (*ang. Locally Linear Embedding*)
  - T-SNA (*ang. t-Distributed stochastic neighbour embedding*)
- ...

Techniki redukcji wymiarowości

Redukcja wymiarów zawsze prowadzi do pozbycia się części informacji (sygnału). Podczas wyboru metody redukcji wymiarów jednym z kluczowych czynników branych pod uwagę jest to, jakie informacje chcemy zachować w sygnale.

Techniki redukcji wymiarowości

- **PCA (*ang. Principle Component Analysis*)**
- **SVD (*ang. Single Value Decomposition*)**
- ICA (*ang. Independent Component Analysis*)
- Rzutowanie losowe (*ang. Random projections*)
- Isomap (*ang. Isometric mapping*)
- MDS (*ang. MultiDimensional Scaling*)
- LDA (*ang. Linear Discriminant Analysis*)
- GDA (*ang. Generalized Discriminant Analysis*)
- Autoenkodery (*ang. Autoencoders*)
- Metody rozmaitościowe (*ang. Mainfold learning algorithms*):
  - FA (*ang. Factor Analysis*)
  - LLA (*ang. Locally Linear Embedding*)
  - T-SNA (*ang. t-Distributed stochastic neighbour embedding*)
- ...

Redukcja wymiarowości - zastosowania

Redukcja liczby wymiarów pozwala na:

- przyśpieszenie procesu uczenia
- wyeliminowanie szumu i nieistotnych szczegółów
- zwizualizowanie zależności w wielowymiarowych zbiorach danych

PCA

Analiza składowych głównych

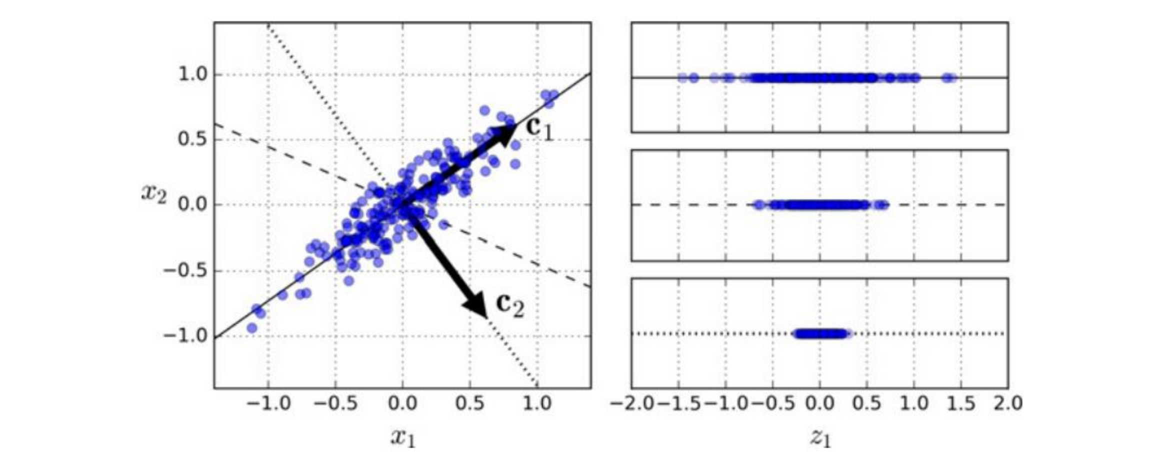
(Principle Component Analysis)

PCA

W analizie składowych głównych możemy wyróżnić dwa podstawowe etapy:

1. Poszukiwanie składowych głównych

Poszukiwanie takich osi (kierunków), które tłumaczą jak największą część zmienności sygnału (wariancje), te osie nazywamy właśnie **składowymi głównymi**.



1. Poszukiwanie składowych głównych

PCA

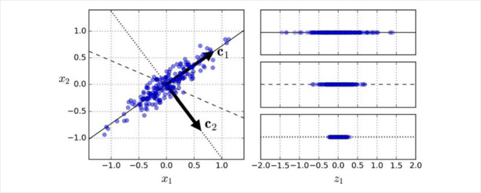
W analizie składowych głównych możemy wyróżnić dwa podstawowe etapy:

1. Poszukiwanie składowych głównych

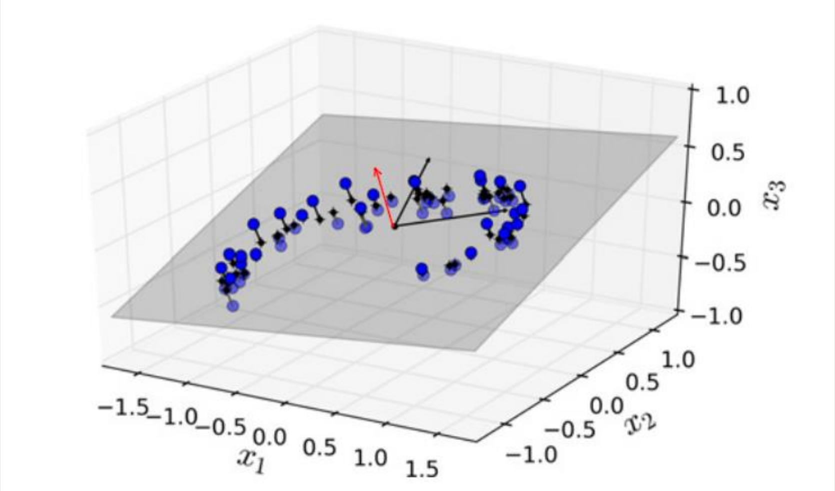
Poszukiwanie takich osi (kierunków), które tłumaczą jak największą część zmienności sygnału (wariancje), te osie nazywamy właśnie **składowymi głównymi**.

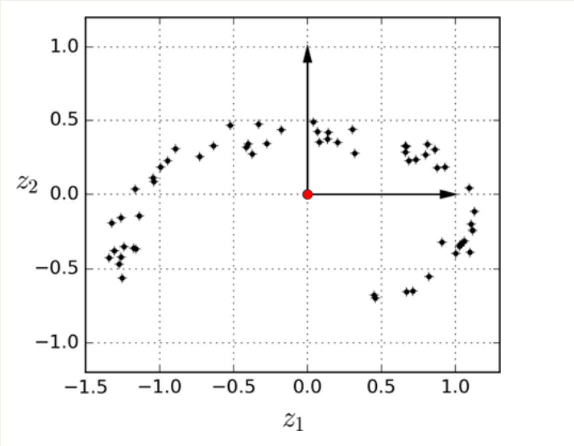
2. Rzutowanie danych na hiperprzestrzeń rozpiętą przez wybrane składowe główne

Rzutowanie zbioru danych na przestrzeń rozpiętą przez wybrane, "najsilniejsze" składowe główne.



1. Poszukiwanie składowych głównych





2. Rzutowanie

PCA

Ale jak znaleźć składowe główne ?

Tym zajmiemy się później. Teraz spójrzmy na przykład.

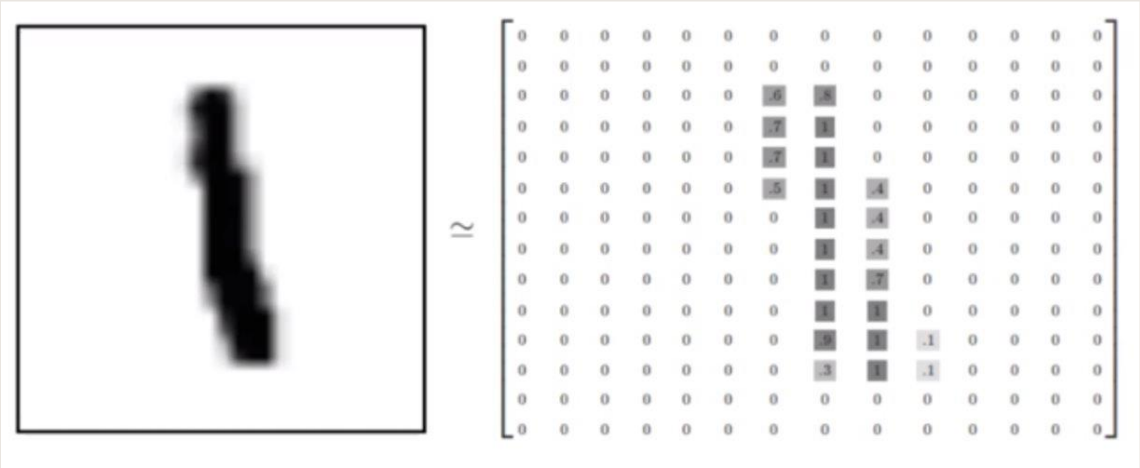
PCA

Dobrym przykładem na początek jest zbiór danych MNIST. Jest to zbiór czarno-białych ilustracji przedstawiających odręcznie zapisane cyfry.

W modelach uczenia maszynowego działających na obrazach najczęściej każdy piksel to cecha. Obrazki mają rozmiar 28x28 co daje 784 cech (czyli 784 wymiary), przy czym większość pikseli jest białych i takie piskele nie niosą ze sobą żadnej informacji. Poza tym piksele czarne są ze sobą silnie skorelowane. Widzimy, że taki sygnał ma ogromny potencjał do redukcji.

Zajmiemy się samym procesem redukcji wymiarów. Nie będziemy budowali modelu na zredukowanych danych.

Ale najpierw powiedzmy trochę więcej o tym jak są reprezentowane obrazy w modelach uczenia maszynowego.

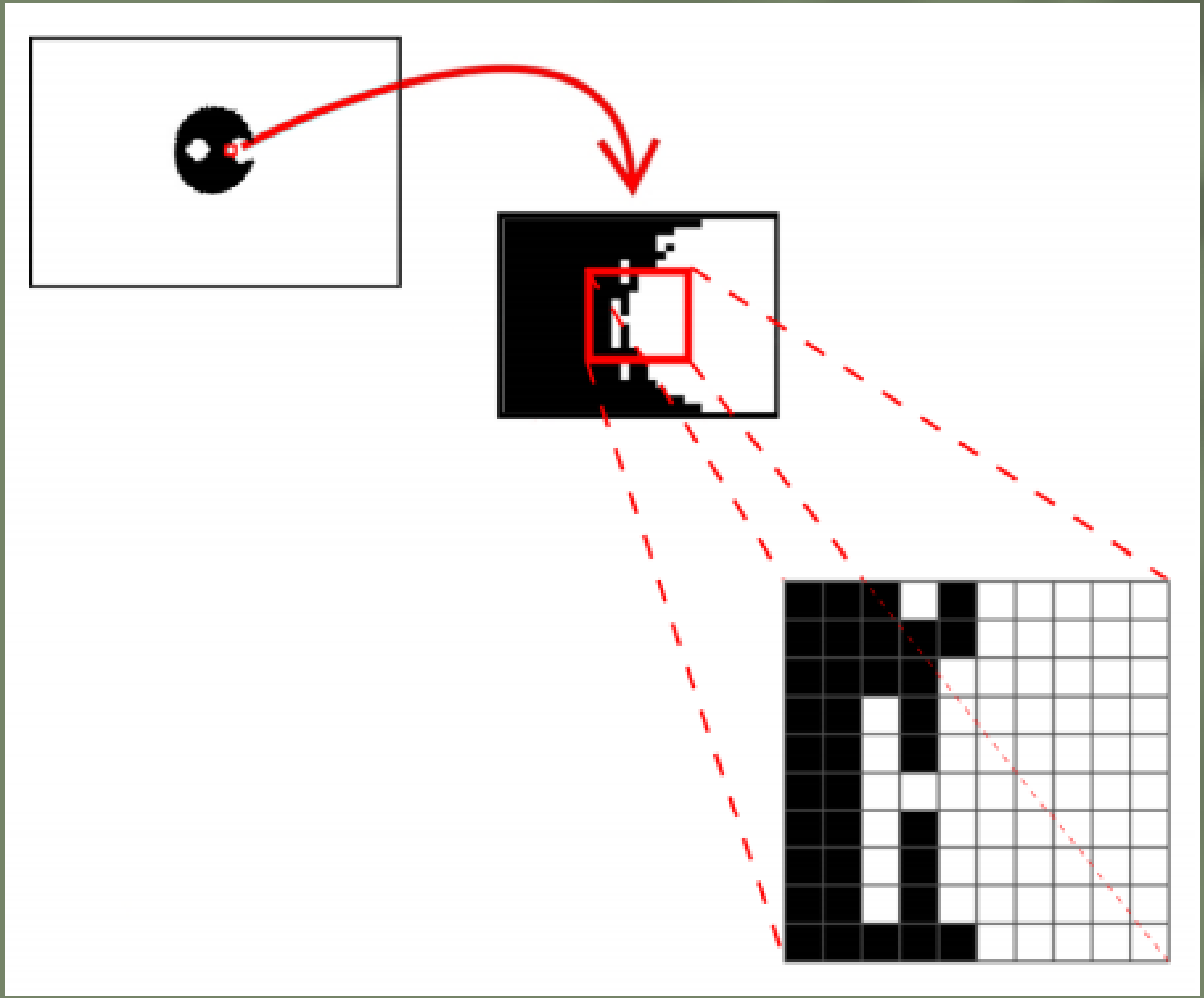


Dodatki

Obraz reprezentatywny

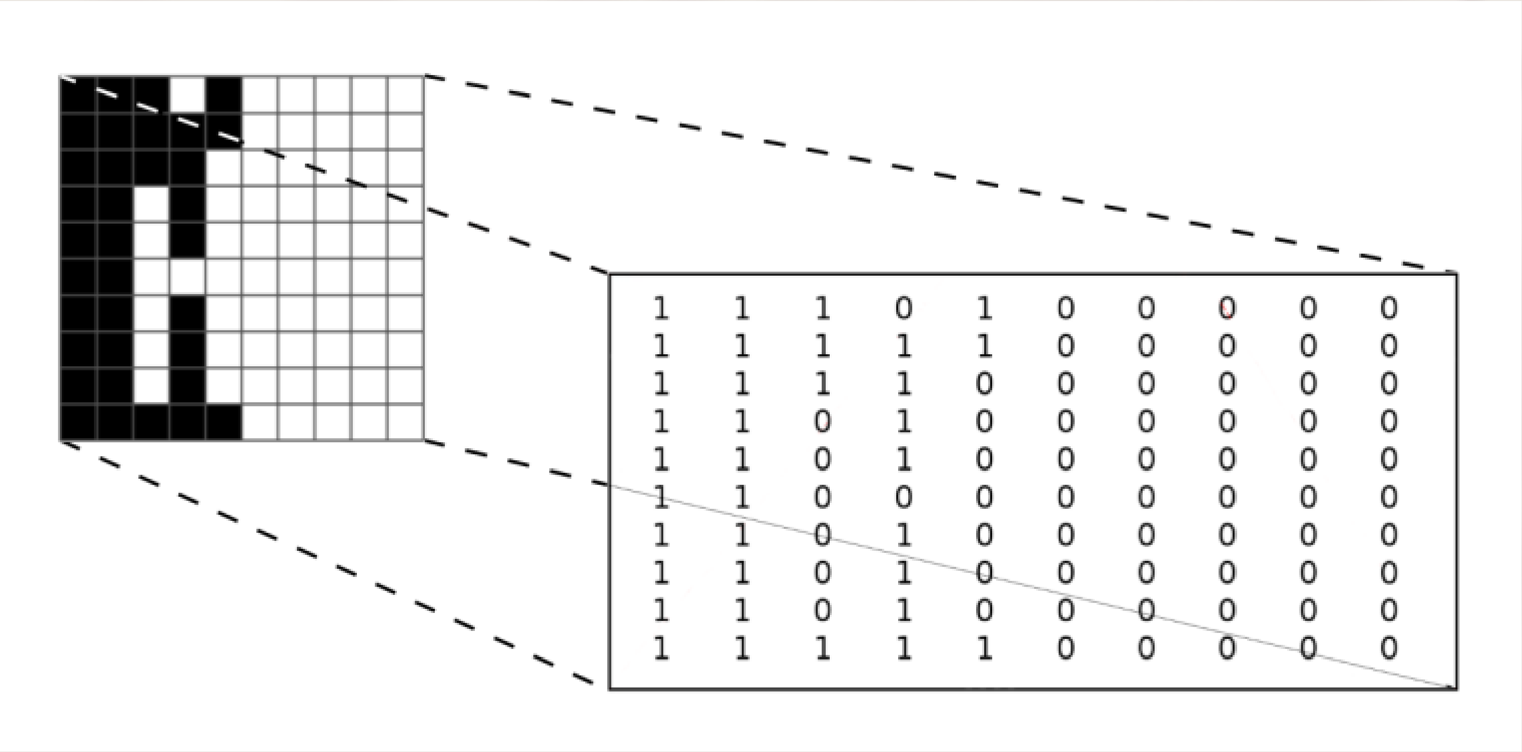
Obraz źrenicy oka w bliskiej podczerwieni

Struktura czarno-białego (binarnego) obrazu



Metody reprezentacji obrazu

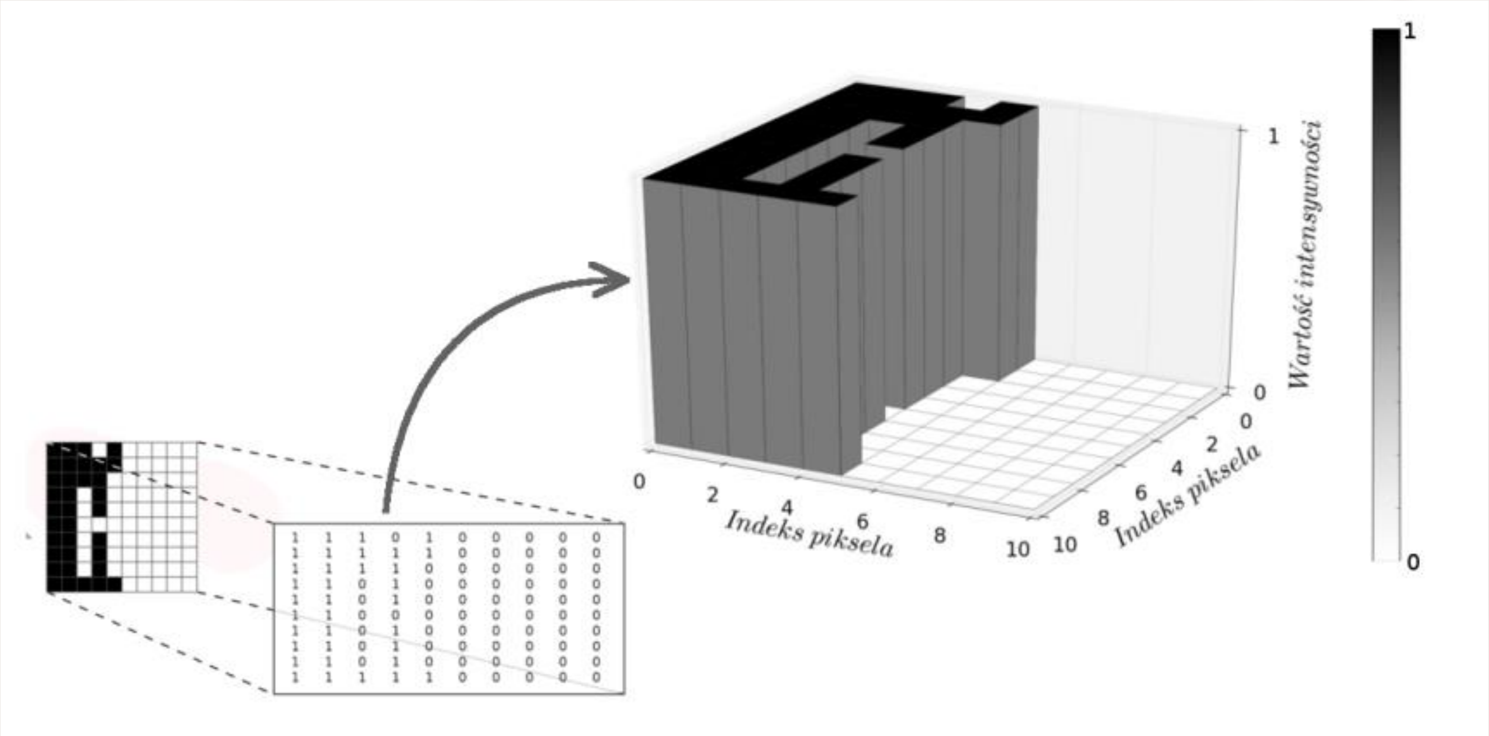
Macierzowa reprezentacja obrazu



Metody reprezentacji obrazu

Funkcyjna reprezentacja obrazu

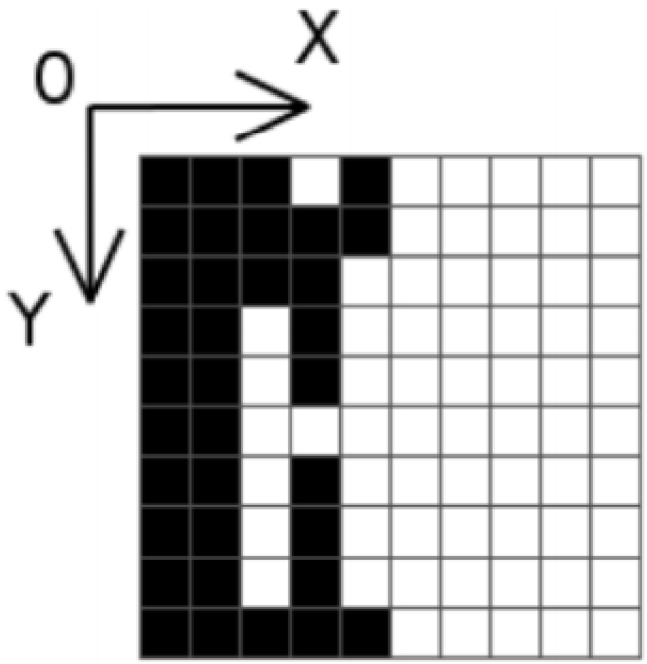
W ujęciu algebraicznym macierz może reprezentować odwzorowanie (funkcję). Tym samym o obrazie możemy myśleć jako o funkcji (odwzorowaniu).



Metody reprezentacji obrazu

Teoriomnogościowa reprezentacja obrazu

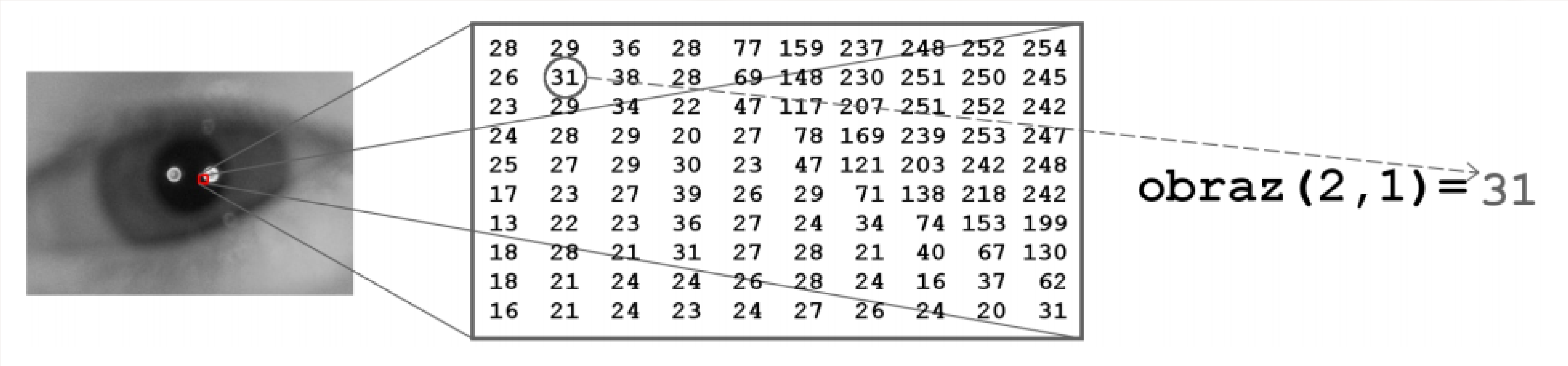
Obraz binarny jako zbiór tych punktów (x, y), których wartość intensywności wynosi 1.



$$
\begin{aligned}
A = \{ \\
&(0,0), (1,0), (2,0), (4,0), \\
&(0,1), (1,1), (2,1), (3,1), (4,1), \\
&(0,2), (1,2), (2,2), (3,2), \\
&(0,3), (1,3), (3,3), \\
&(0,4), (1,4), (3,4), \\
&(0,5), (1,5), \\
&(0,6), (1,6), (3,6), \\
&(0,7), (1,7), (3,7), \\
&(0,8), (1,8), (3,8), \\
&(0,9), (1,9), (2,9), (3,9), (4,9) \\
\}
\end{aligned}
$$

Uogólnienie na obrazy w skali szarości

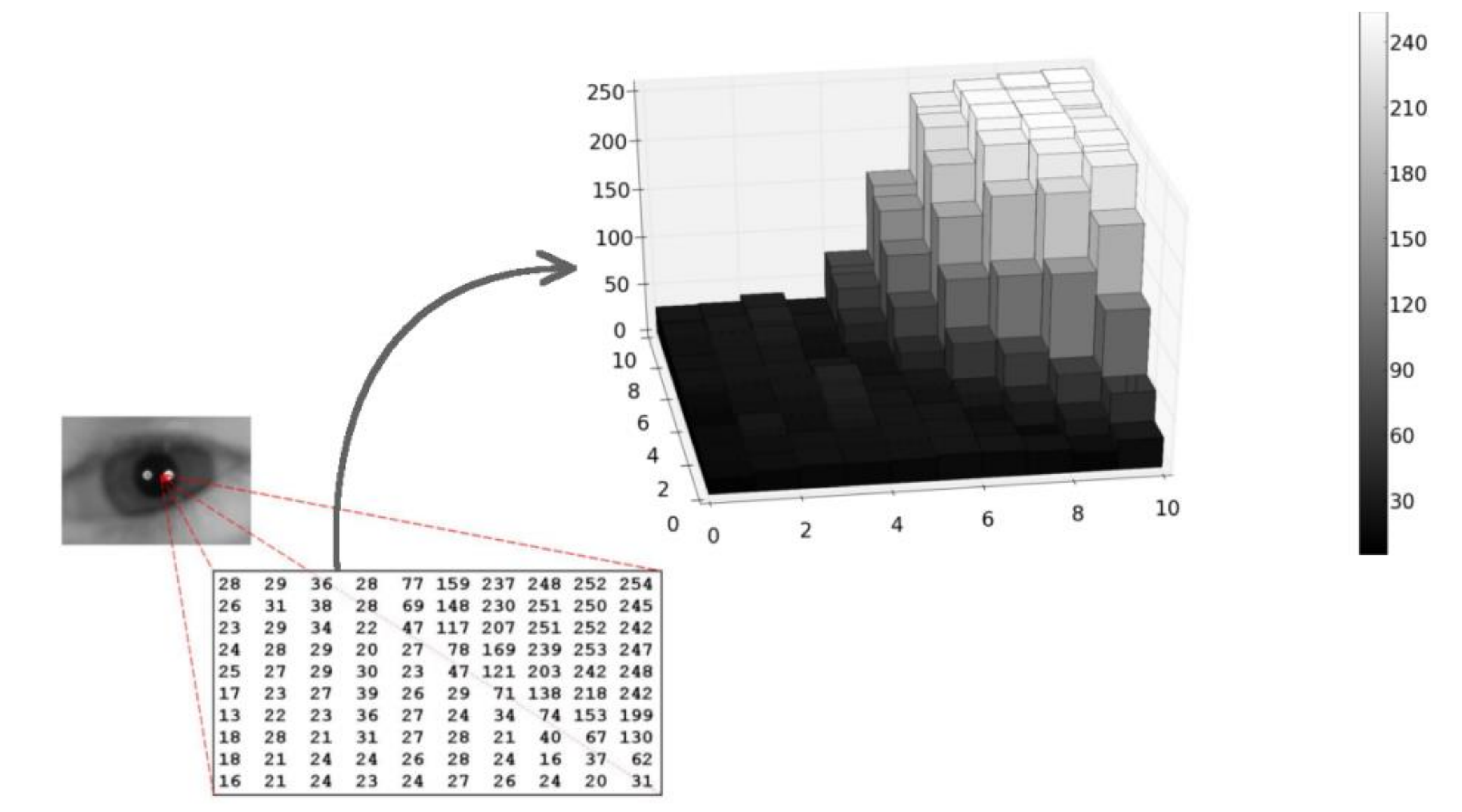
Reprezentacja macierzowa



Macierz jako reprezentacja cyforwego obrazu w skali szarości oraz skalar jako reprezentacja piksela w skali szarości.

Uogólnienie na obrazy w skali szarości

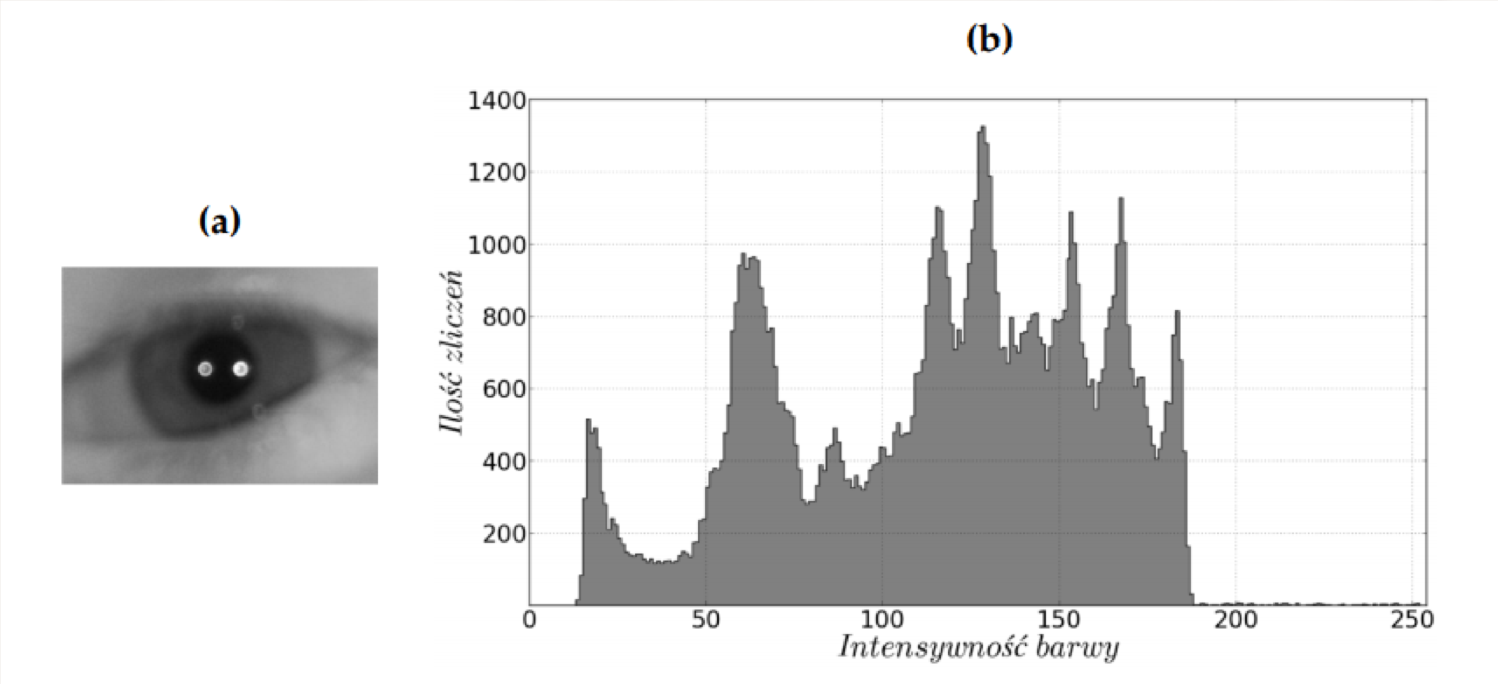
Reprezentacja funkcyjna



Funkcja jako reprezentacja cyfrowego obrazu w skali szarości

Uogólnienie na obrazy w skali szarości

Probabilistyczna reprezentacja obrazu

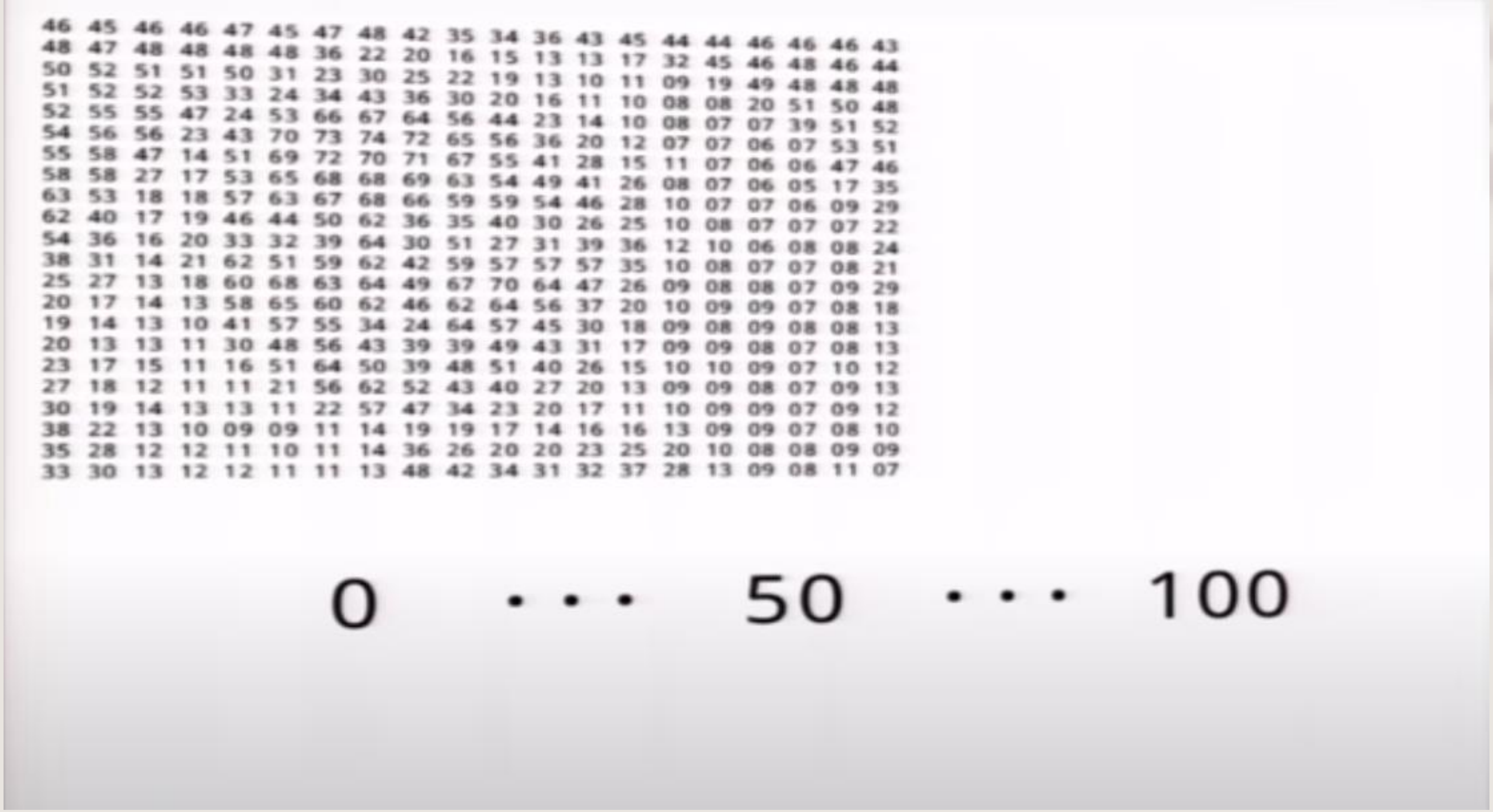


a. Obraz w skali szarości b. Histogram

Mapa kolorów

46	45	46	46	47	45	47	48	42	35	34	36	43	45	44	44	46	46	46	43
48	47	48	48	48	48	36	22	20	16	15	13	13	17	32	45	46	48	46	44
50	52	51	51	50	31	23	30	25	22	19	13	10	11	09	19	49	48	48	48
51	52	52	53	33	24	34	43	36	30	20	16	11	10	08	08	20	51	50	48
52	55	55	47	24	53	66	67	64	56	44	23	14	10	08	07	07	39	51	52
54	56	56	23	43	70	73	74	72	65	56	36	20	12	07	07	06	07	53	51
55	58	47	14	51	69	72	70	71	67	55	41	28	15	11	07	06	06	47	46
58	58	27	17	53	65	68	68	69	63	54	49	41	26	08	07	06	05	17	35
63	53	18	18	57	63	67	68	66	59	59	54	46	28	10	07	07	06	09	29
62	40	17	19	46	44	50	62	36	35	40	30	26	25	10	08	07	07	07	22
54	36	16	20	33	32	39	64	30	51	27	31	39	36	12	10	06	08	08	24
38	31	14	21	62	51	59	62	42	59	57	57	57	35	10	08	07	07	08	21
25	27	13	18	60	68	63	64	49	67	70	64	47	26	09	08	08	07	09	29
20	17	14	13	58	65	60	62	46	62	64	56	37	20	10	09	09	07	08	18
19	14	13	10	41	57	55	34	24	64	57	45	30	18	09	08	09	08	08	13
20	13	13	11	30	48	56	43	39	39	49	43	31	17	09	09	08	07	08	13
23	17	15	11	16	51	64	50	39	48	51	40	26	15	10	10	09	07	10	12
27	18	12	11	11	21	56	62	52	43	40	27	20	13	09	09	08	07	09	13
30	19	14	13	13	11	22	57	47	34	23	20	17	11	10	09	09	07	09	12
38	22	13	10	09	09	11	14	19	19	17	14	16	16	13	09	09	07	08	10
35	28	12	12	11	10	11	14	36	26	20	20	23	25	20	10	08	08	09	09
33	30	13	12	12	11	11	13	48	42	34	31	32	37	28	13	09	08	11	07

Mapa kolorów



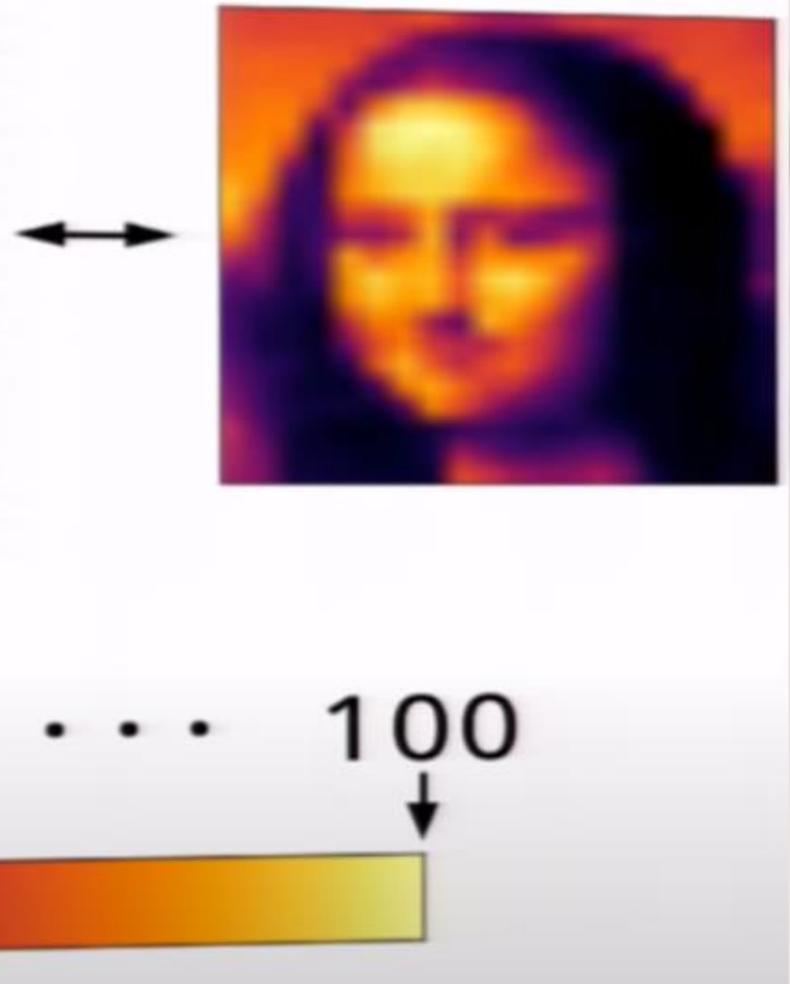
46	45	46	46	47	45	47	48	42	35	34	36	43	45	44	44	46	46	46	43
48	47	48	48	48	48	36	22	20	16	15	13	13	17	32	45	46	48	46	44
50	52	51	51	50	31	23	30	25	22	19	13	10	11	09	19	49	48	48	48
51	52	52	53	33	24	34	43	36	30	20	16	11	10	08	08	20	51	50	48
52	55	55	47	24	53	66	67	64	56	44	23	14	10	08	07	07	39	51	52
54	56	56	23	43	70	73	74	72	65	56	36	20	12	07	07	06	07	53	51
55	58	47	14	51	69	72	70	71	67	55	41	28	15	11	07	06	06	47	46
58	58	27	17	53	65	68	68	69	63	54	49	41	26	08	07	06	05	17	35
63	53	18	18	57	63	67	68	66	59	59	54	46	28	10	07	07	06	09	29
62	40	17	19	46	44	50	62	36	35	40	30	26	25	10	08	07	07	07	22
54	36	16	20	33	32	39	64	30	51	27	31	39	36	12	10	06	08	08	24
38	31	14	21	62	51	59	62	42	59	57	57	57	35	10	08	07	07	08	21
25	27	13	18	60	68	63	64	49	67	70	64	47	26	09	08	08	07	09	29
20	17	14	13	58	65	60	62	46	62	64	56	37	20	10	09	09	07	08	18
19	14	13	10	41	57	55	34	24	64	57	45	30	18	09	08	09	08	08	13
20	13	13	11	30	48	56	43	39	39	49	43	31	17	09	09	08	07	08	13
23	17	15	11	16	51	64	50	39	48	51	40	26	15	10	10	09	07	10	12
27	18	12	11	11	21	56	62	52	43	40	27	20	13	09	09	08	07	09	13
30	19	14	13	13	11	22	57	47	34	23	20	17	11	10	09	09	07	09	12
38	22	13	10	09	09	11	14	19	19	17	14	16	16	13	09	09	07	08	10
35	28	12	12	11	10	11	14	36	26	20	20	23	25	20	10	08	08	09	09
33	30	13	12	12	11	11	13	48	42	34	31	32	37	28	13	09	08	11	07

0 · · · 50 · · · 100

Mapa kolorów

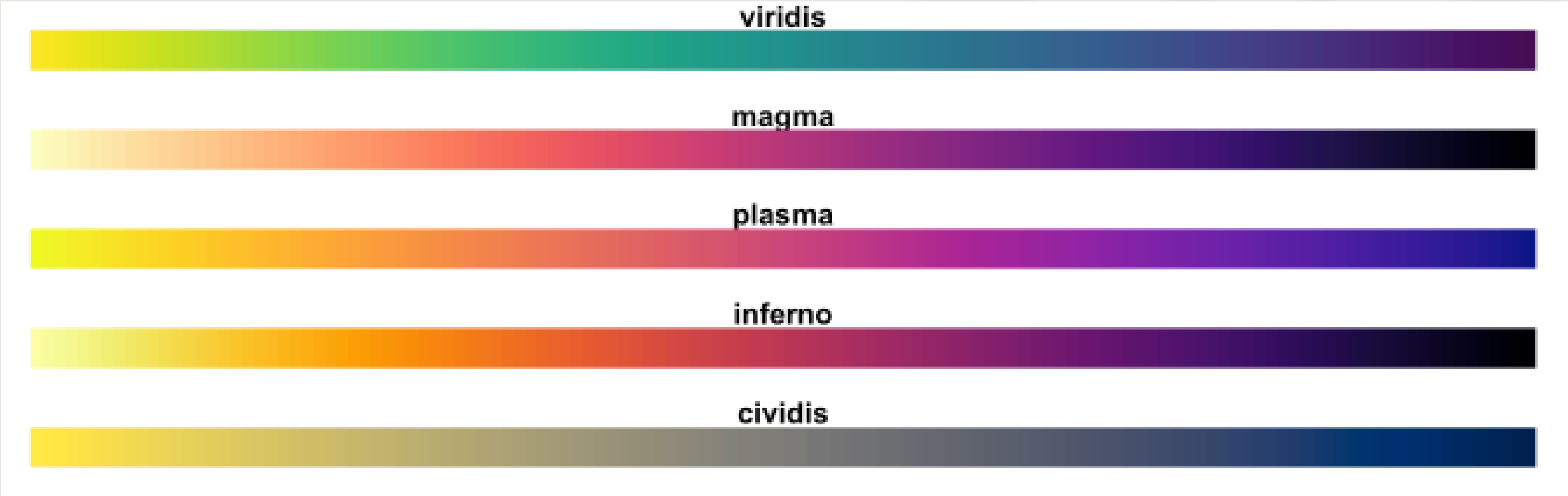
46	45	46	46	47	45	47	48	42	35	34	36	43	45	44	44	46	46	46	43
48	47	48	48	48	48	36	22	20	16	15	13	13	17	32	45	46	48	46	44
50	52	51	51	50	31	23	30	25	22	19	13	10	11	09	19	49	48	48	48
51	52	52	53	33	24	34	43	36	30	20	16	11	10	08	08	20	51	50	48
52	55	55	47	24	53	66	67	64	56	44	23	14	10	08	07	07	39	51	52
54	56	56	23	43	70	73	74	72	65	56	36	20	12	07	07	06	07	53	51
55	58	47	14	51	69	72	70	71	67	55	41	28	15	11	07	06	06	47	46
58	58	27	17	53	65	68	68	69	63	54	49	41	26	08	07	06	05	17	35
63	53	18	18	57	63	67	68	66	59	59	54	46	28	10	07	07	06	09	29
62	40	17	19	46	44	50	62	36	35	40	30	26	25	10	08	07	07	07	22
54	36	16	20	33	32	39	64	30	51	27	31	39	36	12	10	06	08	08	24
38	31	14	21	62	51	59	62	42	59	57	57	57	35	10	08	07	07	08	21
25	27	13	18	60	68	63	64	49	67	70	64	47	26	09	08	08	07	09	29
20	17	14	13	58	65	60	62	46	62	64	56	37	20	10	09	09	07	08	18
19	14	13	10	41	57	55	34	24	64	57	45	30	18	09	08	09	08	08	13
20	13	13	11	30	48	56	43	39	39	49	43	31	17	09	09	08	07	08	13
23	17	15	11	16	51	64	50	39	48	51	40	26	15	10	10	09	07	10	12
27	18	12	11	11	21	56	62	52	43	40	27	20	13	09	09	08	07	09	13
30	19	14	13	13	11	22	57	47	34	23	20	17	11	10	09	09	07	09	12
38	22	13	10	09	09	11	14	19	19	17	14	16	16	13	09	09	07	08	10
35	28	12	12	11	10	11	14	36	26	20	20	23	25	20	10	08	08	09	09
33	30	13	12	12	11	11	13	48	42	34	31	32	37	28	13	09	08	11	07

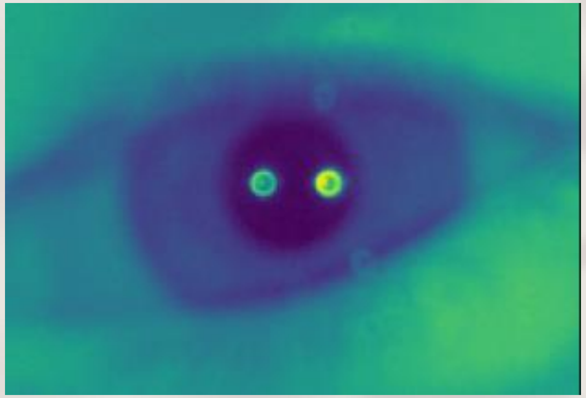
Mapa kolorów

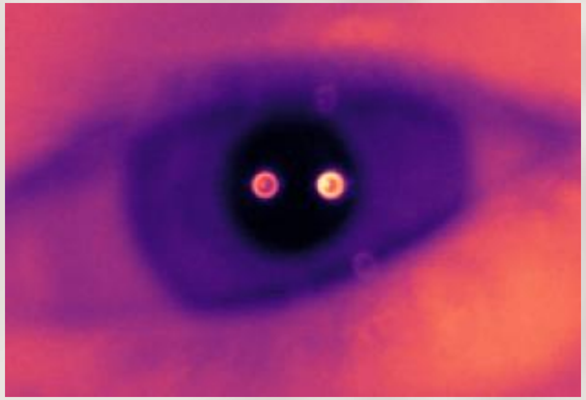


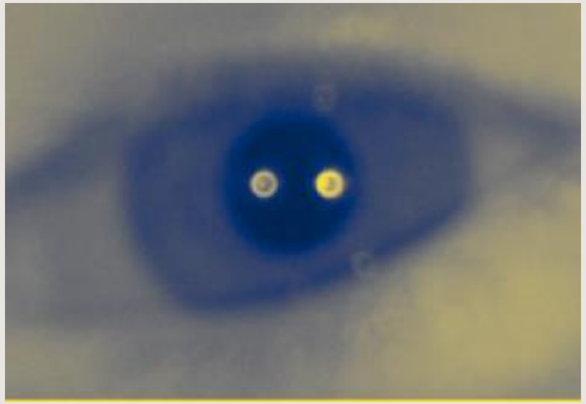
46	45	46	46	47	45	47	48	42	35	34	36	43	45	44	44	46	46	46	43
48	47	48	48	48	48	36	22	20	16	15	13	13	17	32	45	46	48	46	44
50	52	51	51	50	31	23	30	25	22	19	13	10	11	09	19	49	48	48	48
51	52	52	53	33	24	34	43	36	30	20	16	11	10	08	08	20	51	50	48
52	55	55	47	24	53	66	67	64	56	44	23	14	10	08	07	07	39	51	52
54	56	56	23	43	70	73	74	72	65	56	36	20	12	07	07	06	07	53	51
55	58	47	14	51	69	72	70	71	67	55	41	28	15	11	07	06	06	47	46
58	58	27	17	53	65	68	68	69	63	54	49	41	26	08	07	06	05	17	35
63	53	18	18	57	63	67	68	66	59	59	54	46	28	10	07	07	06	09	29
62	40	17	19	46	44	50	62	36	35	40	30	26	25	10	08	07	07	07	22
54	36	16	20	33	32	39	64	30	51	27	31	39	36	12	10	06	08	08	24
38	31	14	21	62	51	59	62	42	59	57	57	57	35	10	08	07	07	08	21
25	27	13	18	60	68	63	64	49	67	70	64	47	26	09	08	08	07	09	29
20	17	14	13	58	65	60	62	46	62	64	56	37	20	10	09	09	07	08	18
19	14	13	10	41	57	55	34	24	64	57	45	30	18	09	08	09	08	08	13
20	13	13	11	30	48	56	43	39	39	49	43	31	17	09	09	08	07	08	13
23	17	15	11	16	51	64	50	39	48	51	40	26	15	10	10	09	07	10	12
27	18	12	11	11	21	56	62	52	43	40	27	20	13	09	09	08	07	09	13
30	19	14	13	13	11	22	57	47	34	23	20	17	11	10	09	09	07	09	12
38	22	13	10	09	09	11	14	19	19	17	14	16	16	13	09	09	07	08	10
35	28	12	12	11	10	11	14	36	26	20	20	23	25	20	10	08	08	09	09
33	30	13	12	12	11	11	13	48	42	34	31	32	37	28	13	09	08	11	07

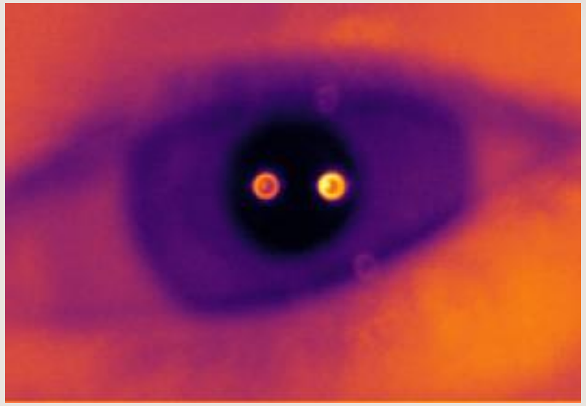
Mapy kolorów

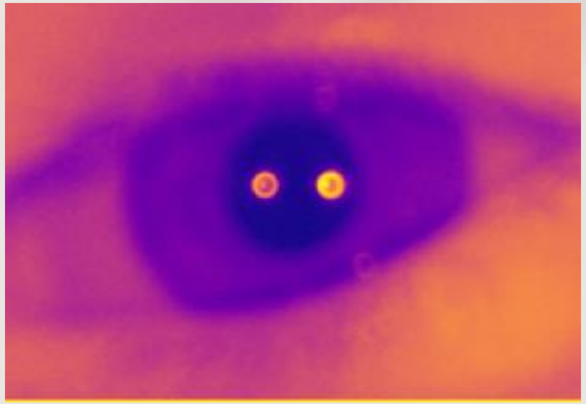












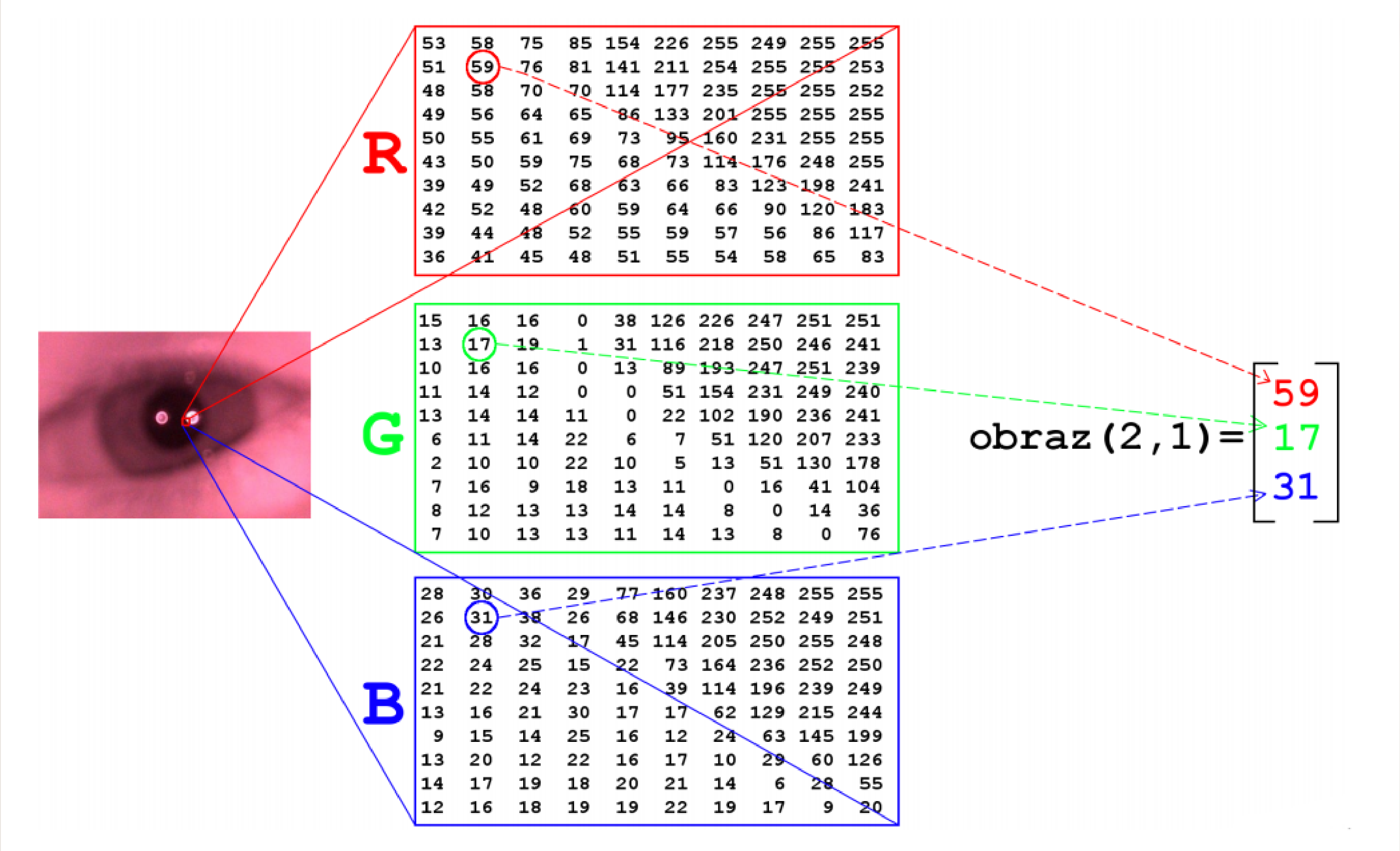
Uogólnienie na obrazy w skali szarości

Reprezentacja teoriomnogościowa może zostać uogólniona poprzez zastosowanie:

- trójwymiarowych zbiorów
- zbioru trójwymiarowych wektorów
- rodziny zbiorów dwuwymiarowych

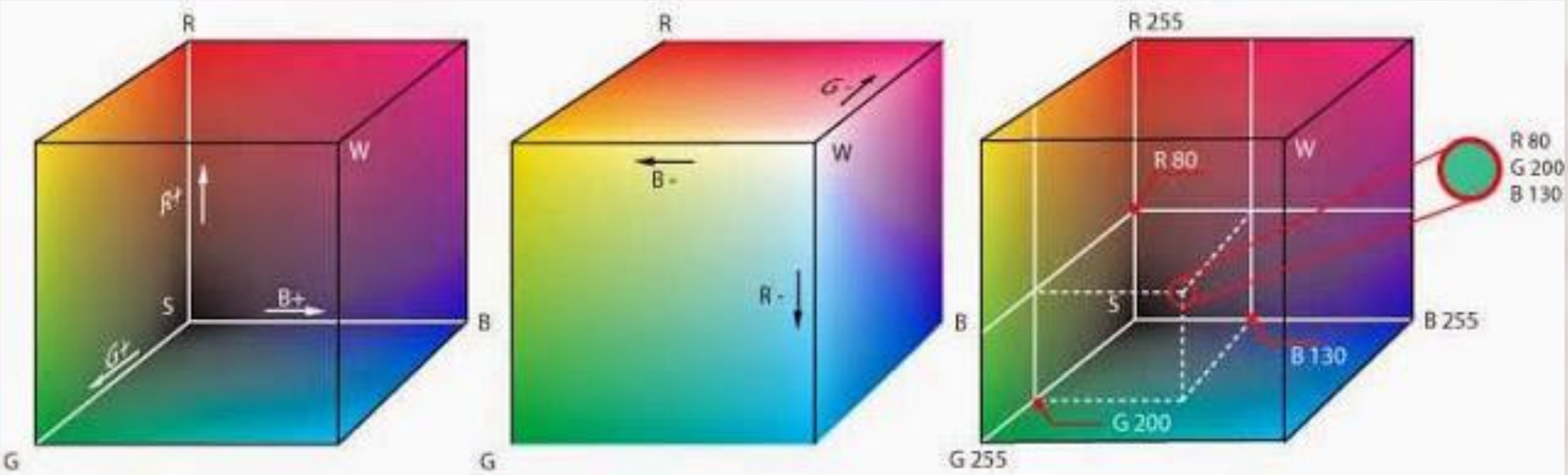
Uogólnienie na obrazy kolorowe

Macierzowa reprezentacja obrazu

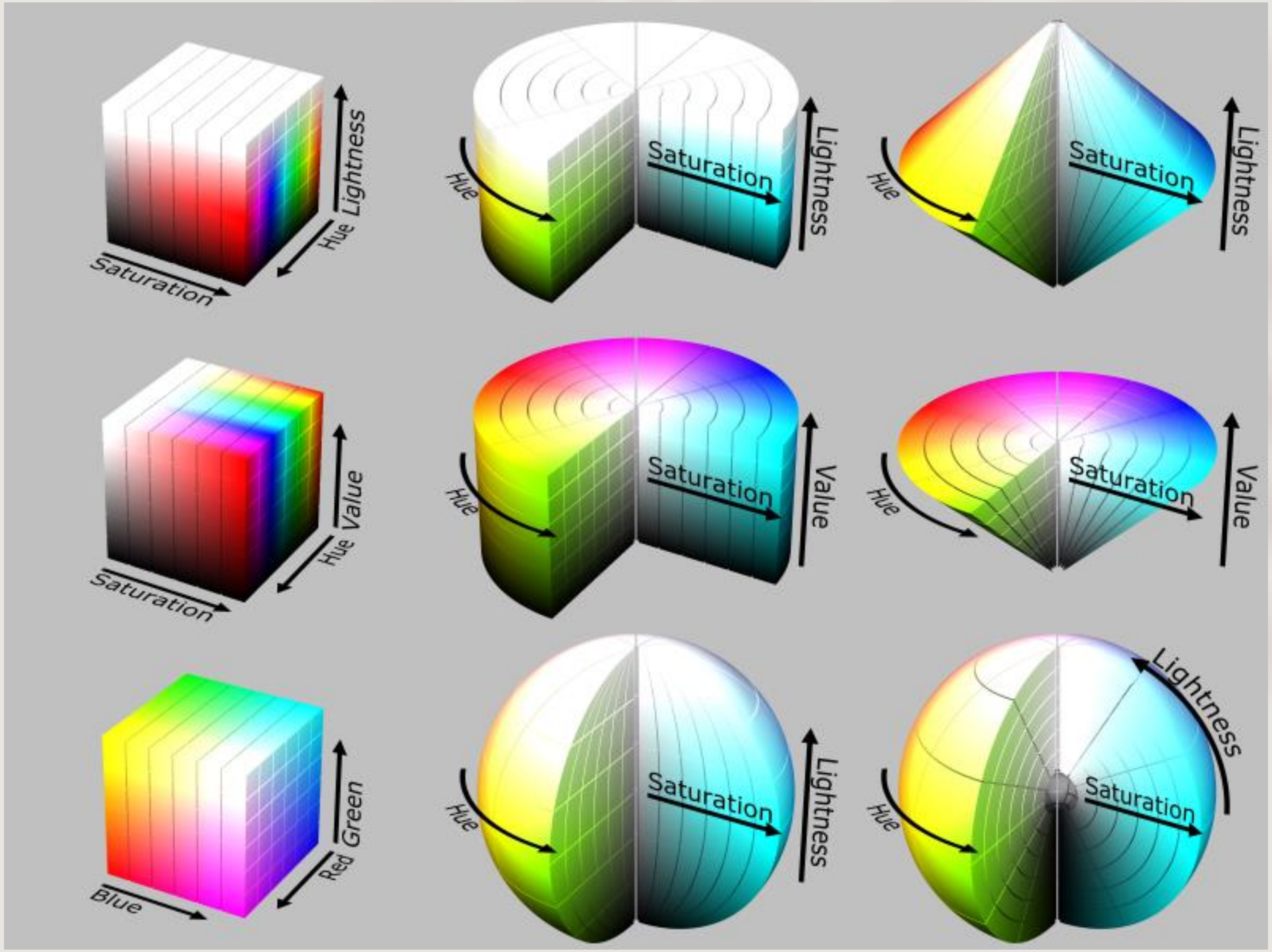


Trzy macierze jako reprezentacja RGB cyfrowego obrazu kolorowego oraz trójwymiarowy wektor jako reprezentacja kolorowego piksela.

Przestrzenie barw



Przestrzenie kolorów



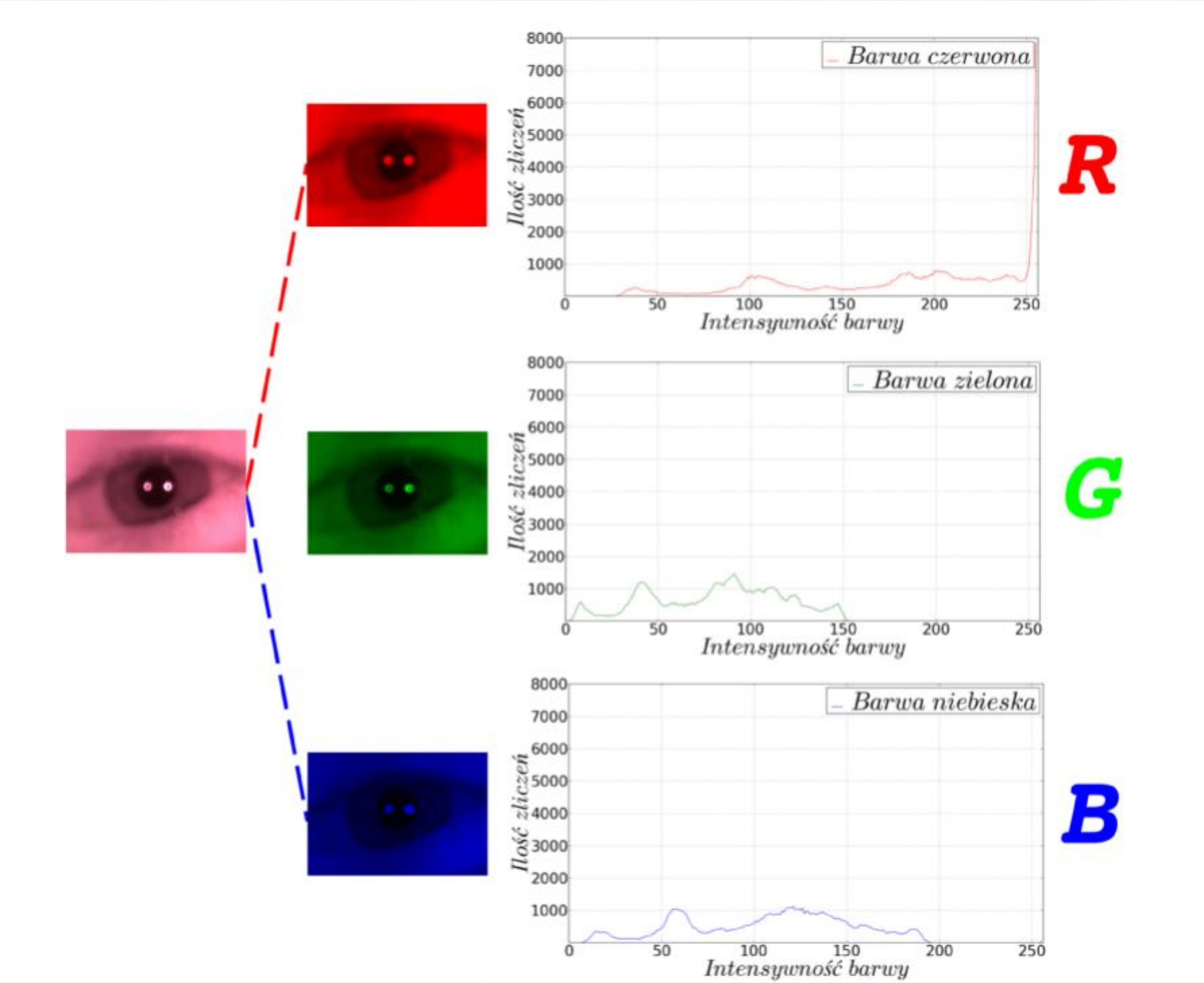
Uogólnienie na obrazy kolorowe

Funkcyjna reprezentacja obrazu zostaje uogólniona poprzez rozpatrywanie funkcji wektorowych.

Teoriomnogościowa reprezentacja obrazu zostaje uogólniona poprzez zwiększenie wymiarowości zagadnienia.

Uogólnienie na obrazy kolorowe

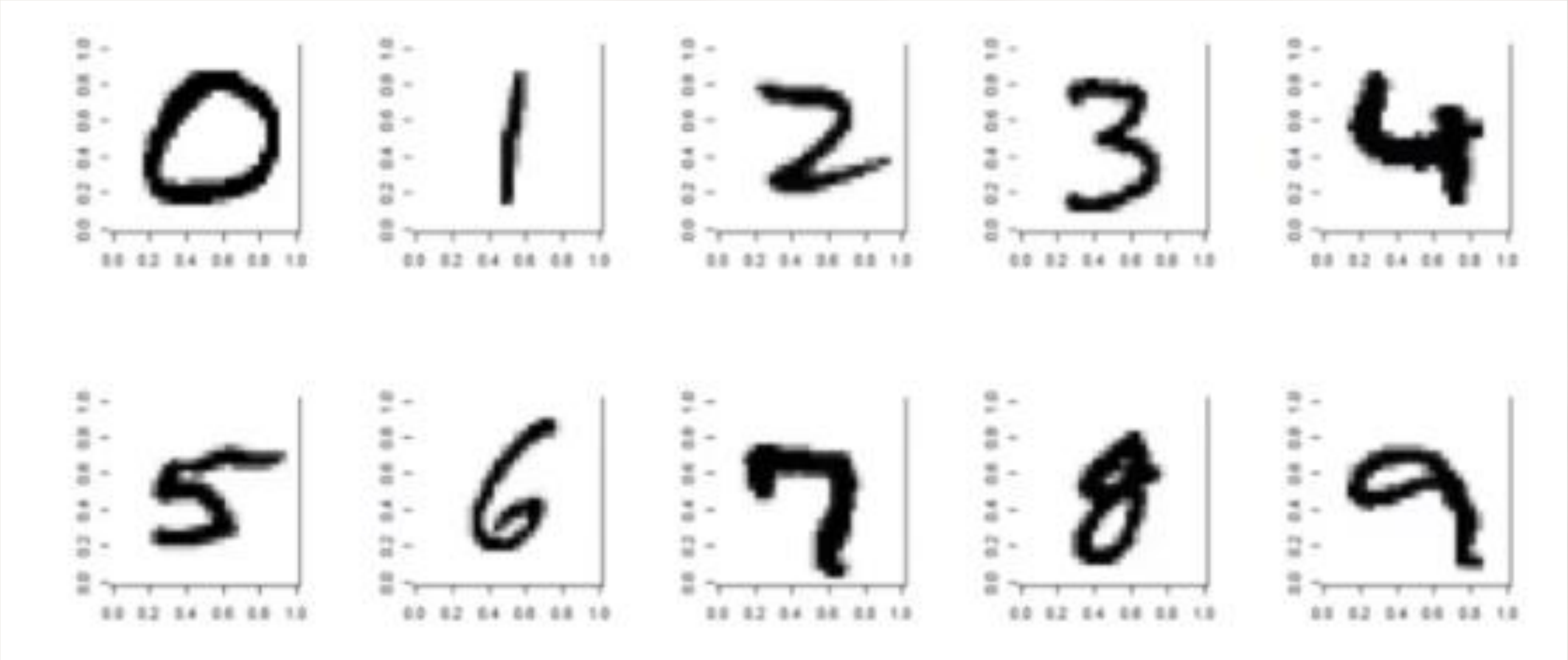
Probabilistyczna reprezentacja obrazu



Histogramy trzech podstawowych składowych obrazu (RGB)

MNIST

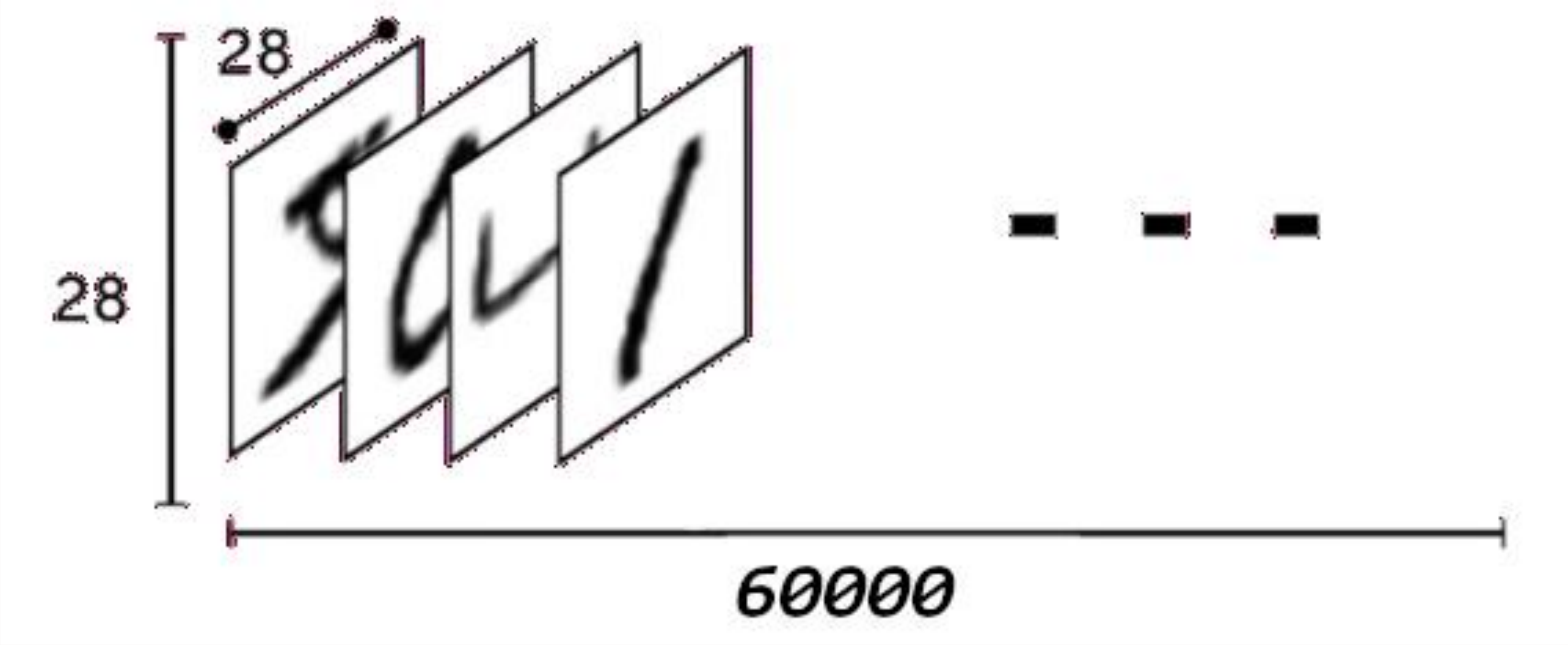
MNIST



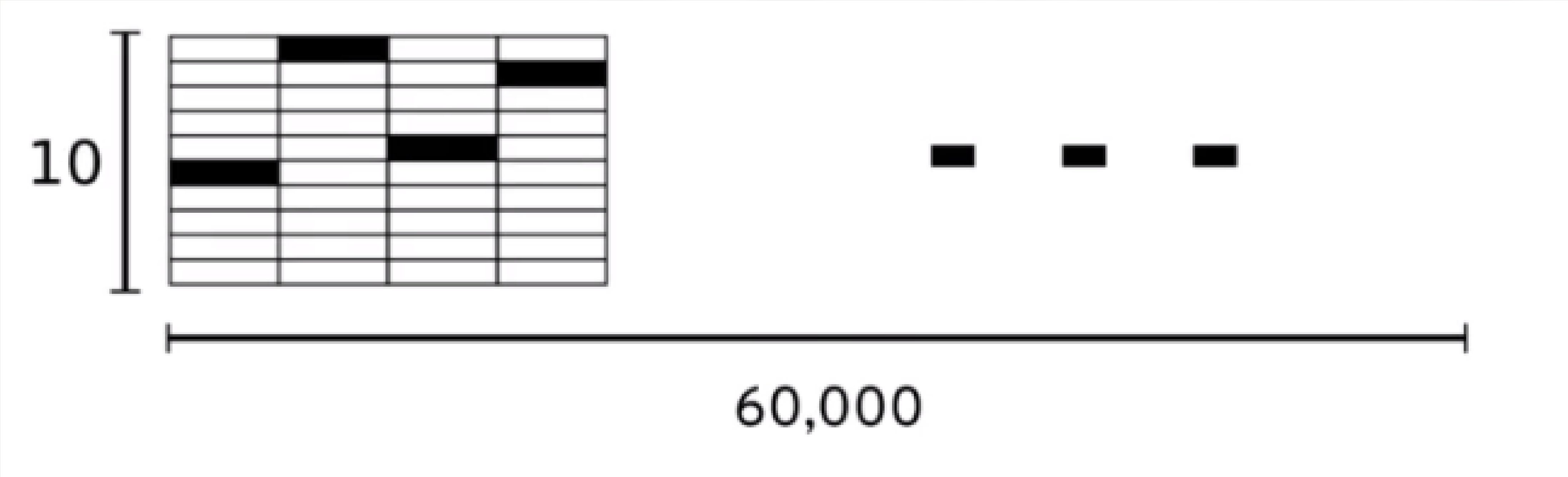
MNIST

$$\simeq \begin{bmatrix}
0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&.6&.8&0&0&0&0&0&0\\
0&0&0&0&0&0&.7&1&0&0&0&0&0&0\\
0&0&0&0&0&0&.7&1&0&0&0&0&0&0\\
0&0&0&0&0&0&.5&1&.4&0&0&0&0&0\\
0&0&0&0&0&0&0&1&.4&0&0&0&0&0\\
0&0&0&0&0&0&0&1&.4&0&0&0&0&0\\
0&0&0&0&0&0&0&1&.7&0&0&0&0&0\\
0&0&0&0&0&0&0&1&1&0&0&0&0&0\\
0&0&0&0&0&0&0&.9&1&.1&0&0&0&0\\
0&0&0&0&0&0&0&.3&1&.1&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0
\end{bmatrix}$$

MNIST - feature



MNIST - targets



Technikalia

Macierz

Na dane z którymi pracujemy możemy patrzeć jak na macierz. Wystarczy spojrzeć na dowolny z omawianych dataframów, żenby się o tym przekonać.

Ale czym jest macierz ?

To już trochę trudniejsz pytanie, bo na macierz możemy patrzeć na wiele różnych sposobów. W wielu zagadnieniach patrzy się na macierz z kilku stron jednocześnie, korzystając z różnych interpretacji macierzy w ramach tego samego wywodu.

	Work	Price	Salary
Amsterdam	1714	65.6	49.0
Athens	1792	53.8	30.4
Bogota	2152	37.9	11.5
Bombay	2052	30.3	5.3
Brussels	1708	73.8	50.5
BuenosAires	1971	56.1	12.5
Caracas	2041	61.0	10.9
Chicago	1924	73.9	61.9
Copenhagen	1717	91.3	62.9

$$= \begin{bmatrix} 1714 & 65.6 & 49.0 \\ 1792 & 53.8 & 30.4 \\ 2152 & 37.9 & 11.5 \\ 2052 & 30.3 & 5.3 \\ 1708 & 73.8 & 50.5 \\ 1971 & 56.1 & 12.5 \\ \cdots & & \end{bmatrix}$$

Macierz

Skupmy się na razie na jednej z możliwych interpretacji macierzy. Macierz jako reprezentacja jakiejś transformacji liniowej.

Dla danego wektora v operację:

$$Av$$

możemy potraktować jako nowy wektor, który powstał w wyniku zadziałania macierzy A na wektor v. Macierz A jest w tym przykładzie pewną transformacją.

$$Av = \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} x \\ y \end{bmatrix}$$

$$Av = x\begin{bmatrix} a \\ c \end{bmatrix} + y\begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} ax+by \\ cx+dy \end{bmatrix}$$

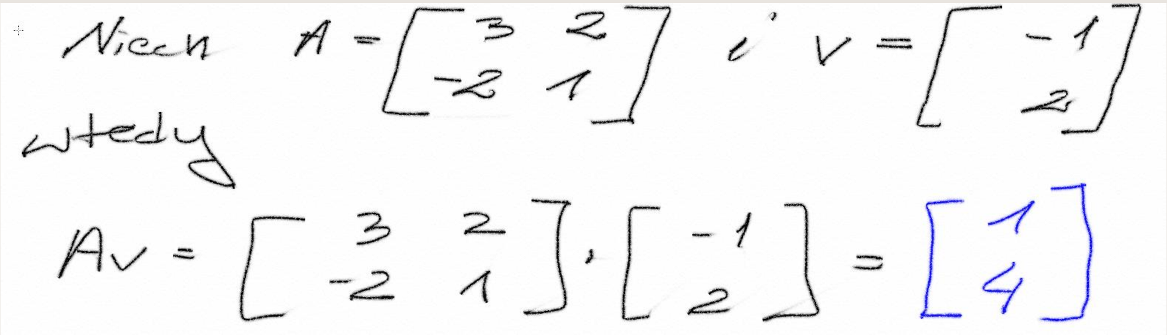
Macierz

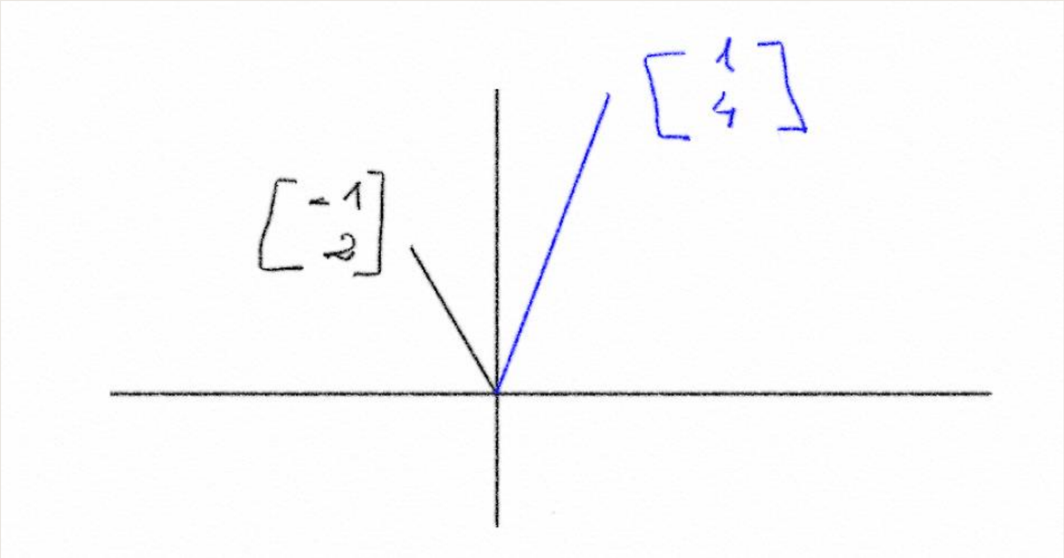
Skupmy się na jednej z możliwych interpretacji macierzy. Macierz jako reprezentacja jakiejś transformacji liniowej.

Dla danego wektora v operację:

$$Av$$

możemy potraktować jako nowy wektor, który powstał w wyniku zadziałania macierzy A na wektor v. Macierz A jest w tym przykładzie pewną transformacją.



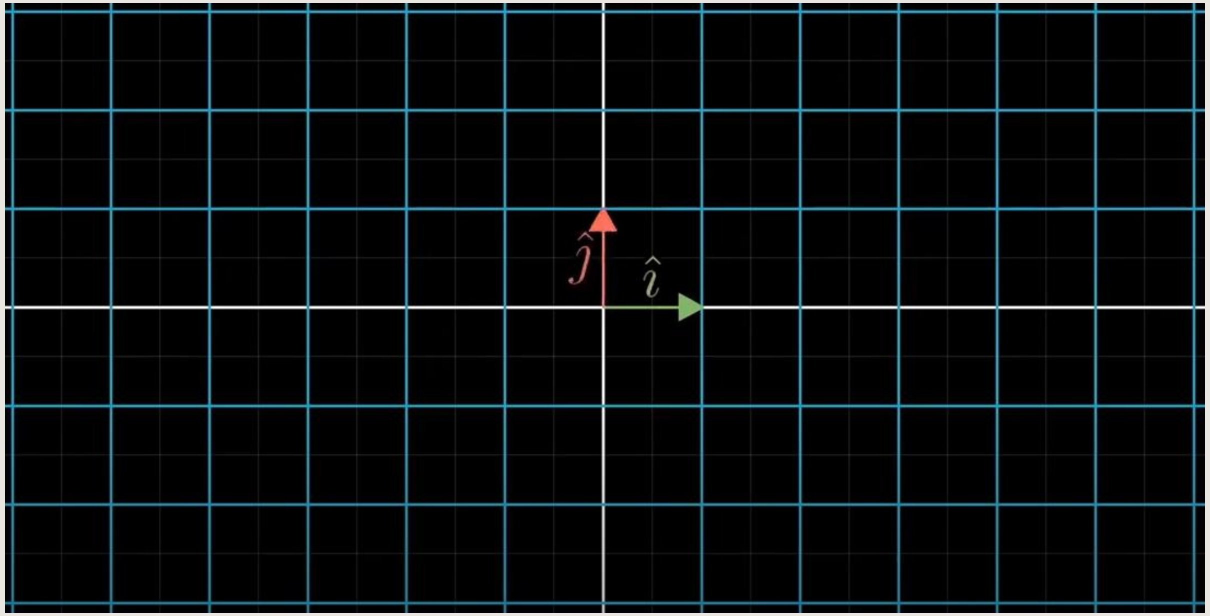


Macierz

Na tą samą macierz możemy spojrzeć w trochę inny sposób.

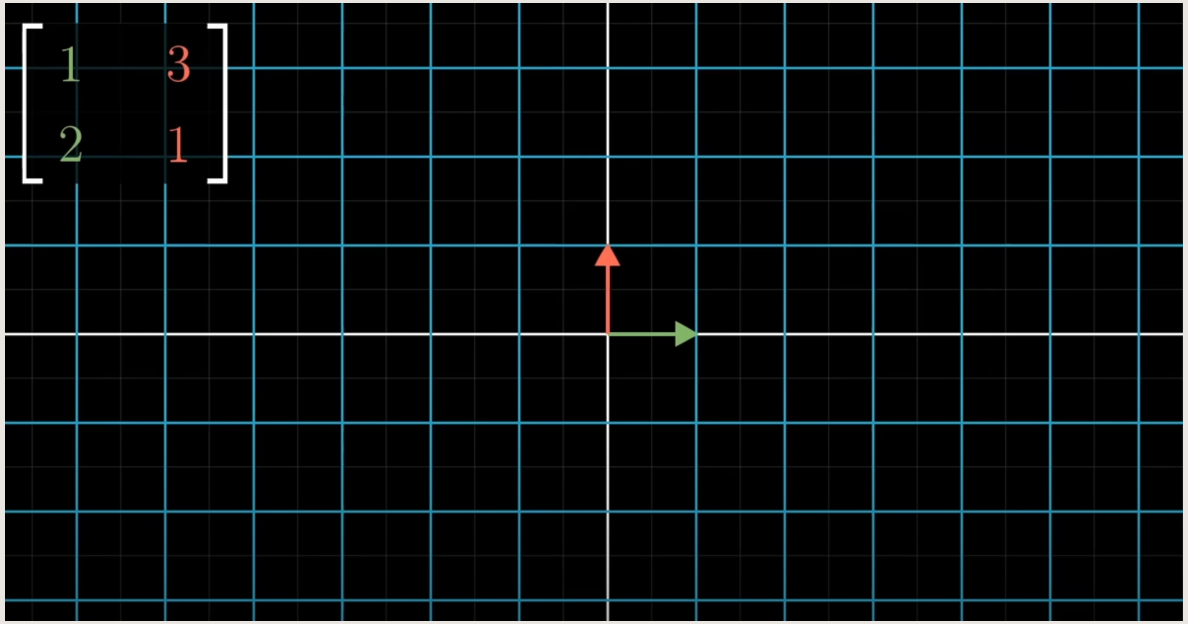
Transformacje można rozpatrywać niezależnie od wektora na jakim ta transformacja jest wykonywana. Transformacja A zrobi coś z wektorem **v**, ale z wektorem **w** zrobić coś innego. Jeżeli pomyślimy o tej transformacji jako o operacji wykonywanej na dowolnym, możliwym wektorze zobaczymy ją jako operację, która w pewien sposób przekształca (transformuje) nie tyle jeden wektor, co całą przestrzeń wektorów.

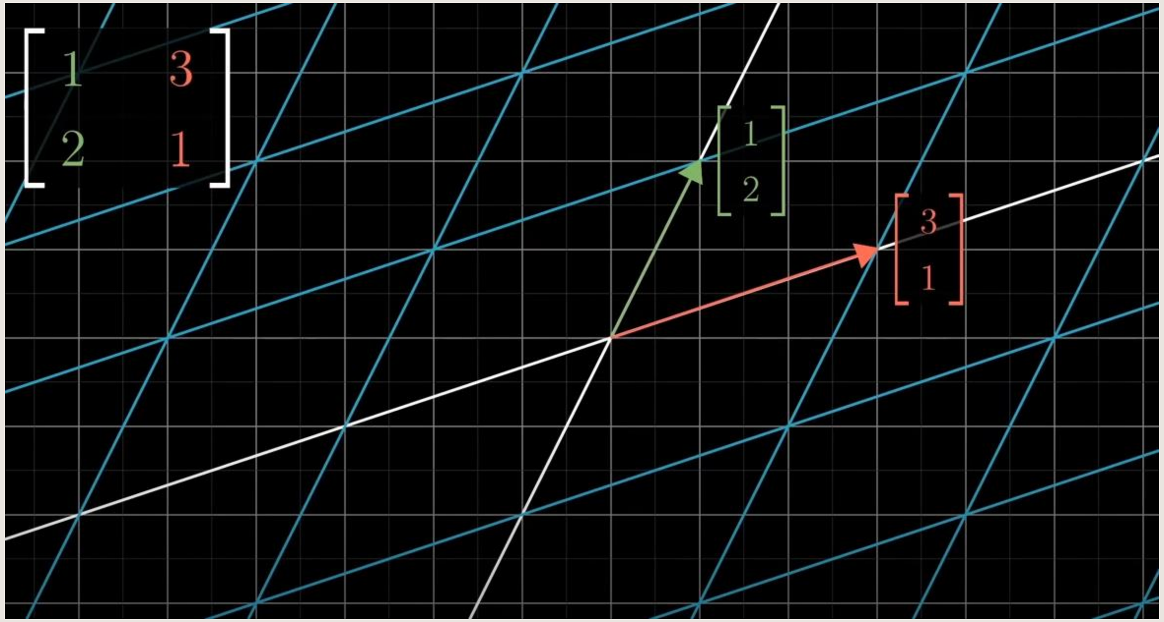
Tylko jak reprezentować przestrzeń wektorów ?
Robi się to za pomocą tzw. wektorów bazowych. Ortonormalnych wektorów zaczepionych w początku układu współrzędnych. A całą przestrzeń reprezentuje się jako siatkę skonstruowaną z wektorów bazowych.



Macierz

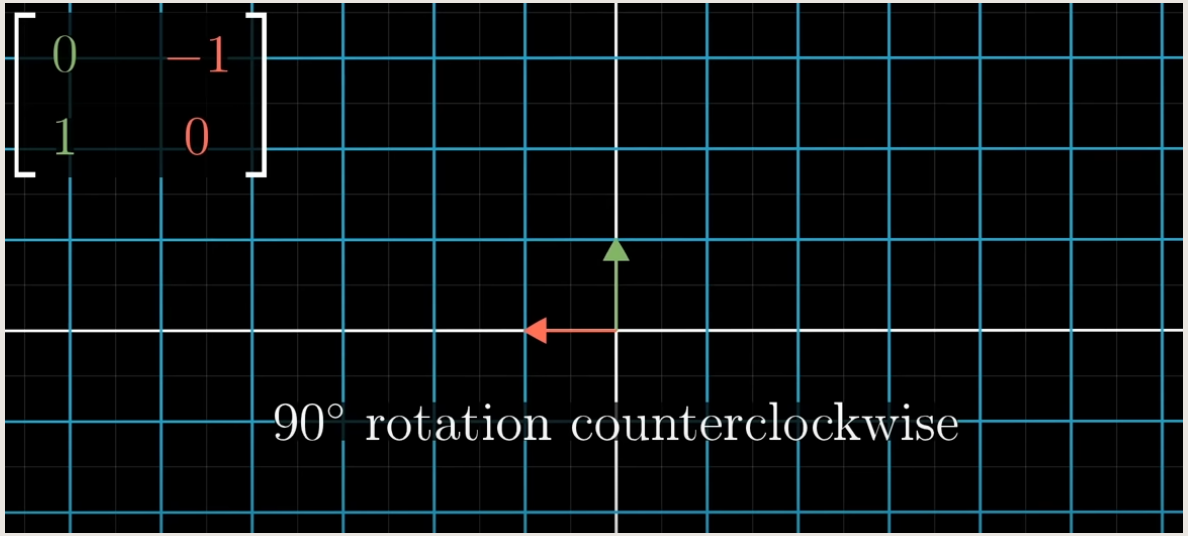
Teraz o macierzy A możemy myśleć jako o transformacji przekształcającej jedną przestrzeń w inną. A skoro przestrzeń jest zdefiniowana przez wektory bazowe to znaczy, że myślimy o niej jako o operacji przekształacjącej właśnie te wektory bazowe (będące budulcem przestrzeni).

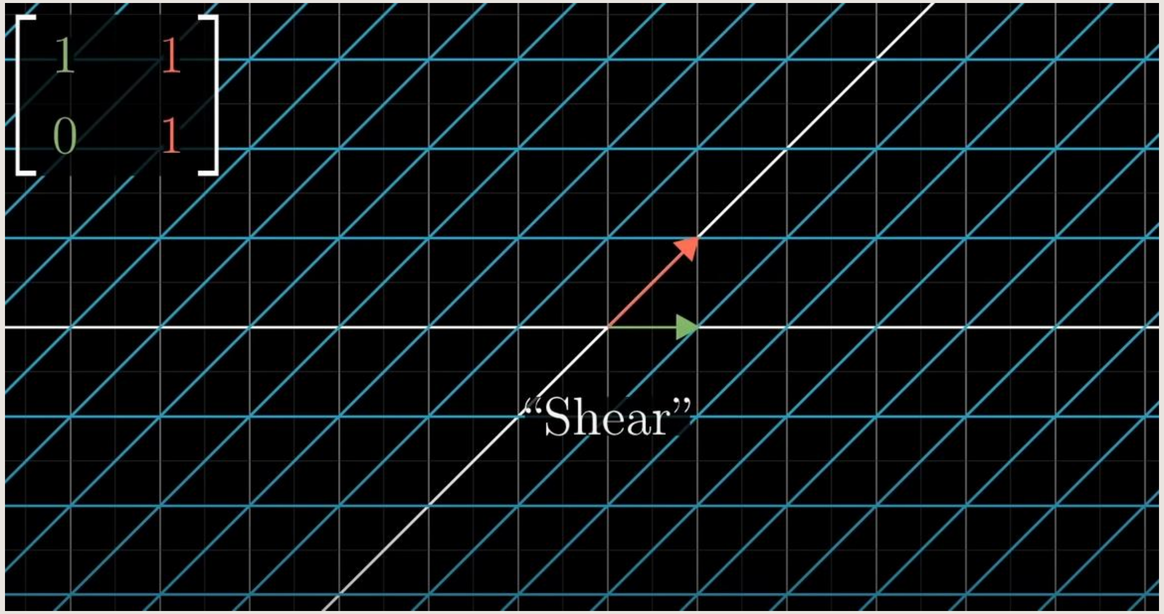




Macierz

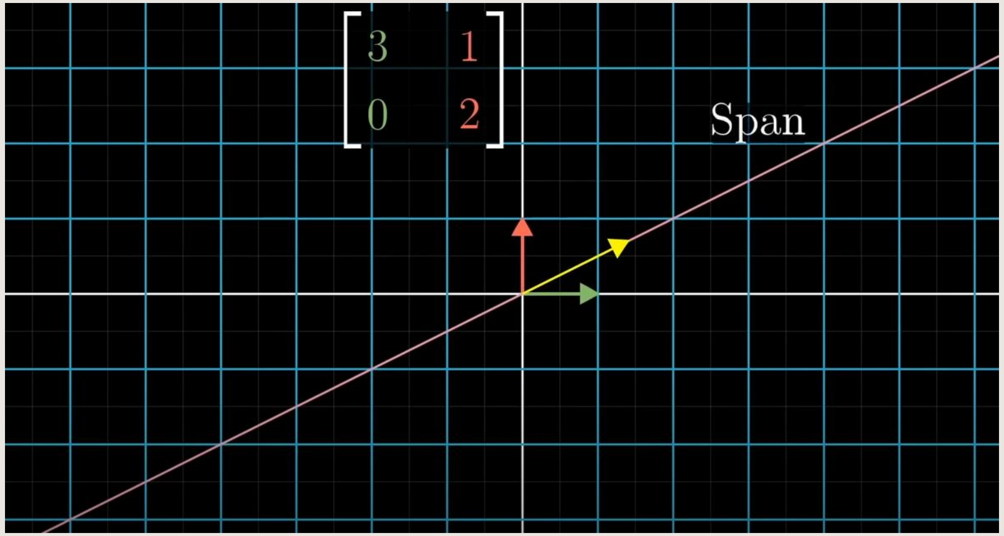
Istnieją wiele ciekawych transformacji.
Po prawej, od góry: obrót, ścinanie.

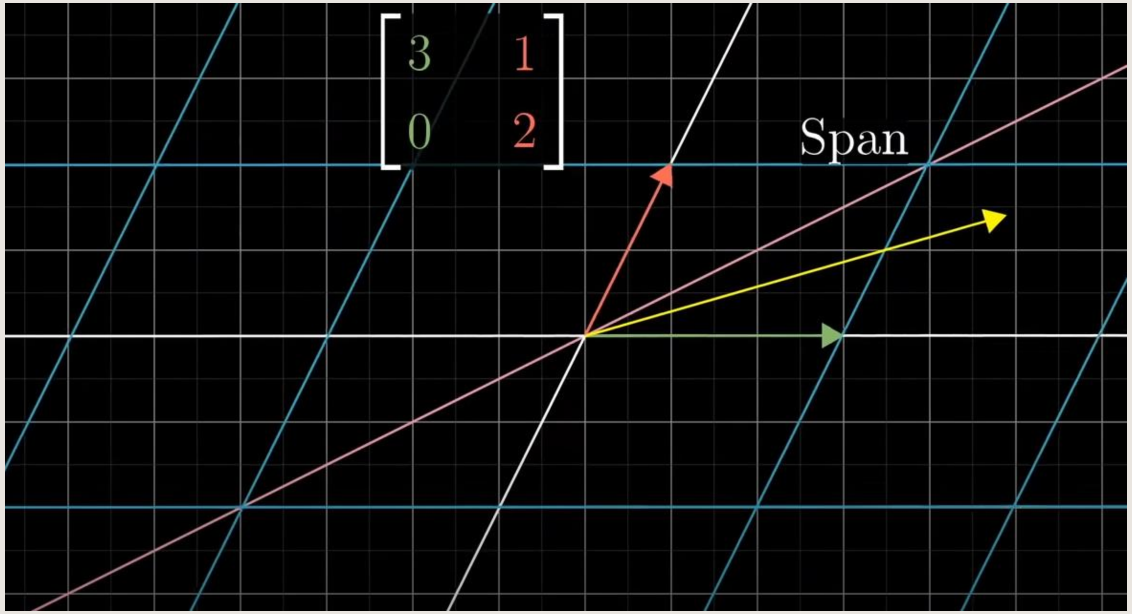




Macierz

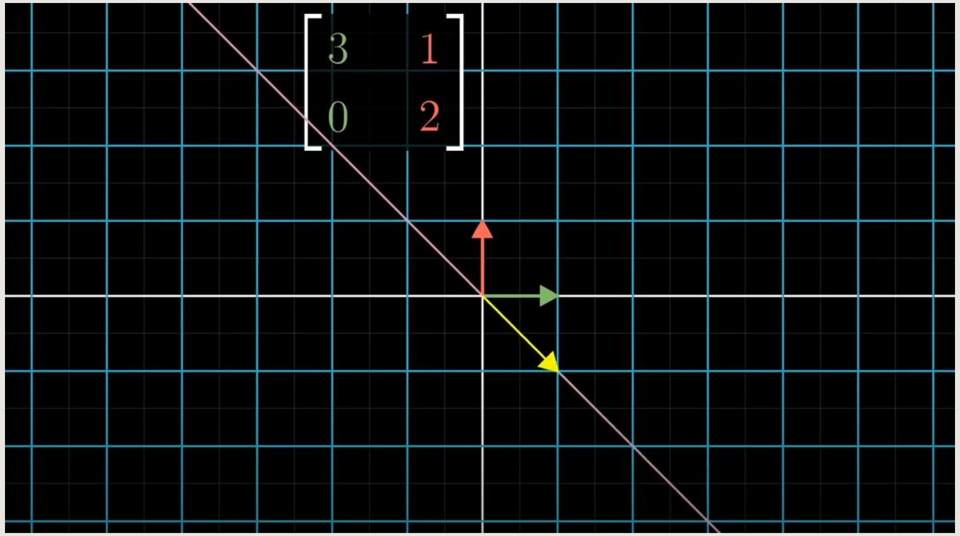
W zależności od transformacji, w transformowanej przestrzeni mogą znajdować się takie wektory, który po transformacji nie zmienią swojego kierunku. Są to szczególne wektory, ponieważ stanowią one swoistą charakterystykę macierzy. Nie ważne z jakiej przestrzeni w jaką przechodzimy. Te wektory będą miały zawsze ten sam kierunek. Jedynie w wyniku transformacji zostaną przeskalowane. Takie wektory nazywamy wektorami własnymi macierzy, a współczynnik o jaki są skalowane podczas transformacji nazywamy wartościami własnymi macierzy (np. w przypadku macierzy obrotu będą to wektory znajdujące się na osi obrotu).

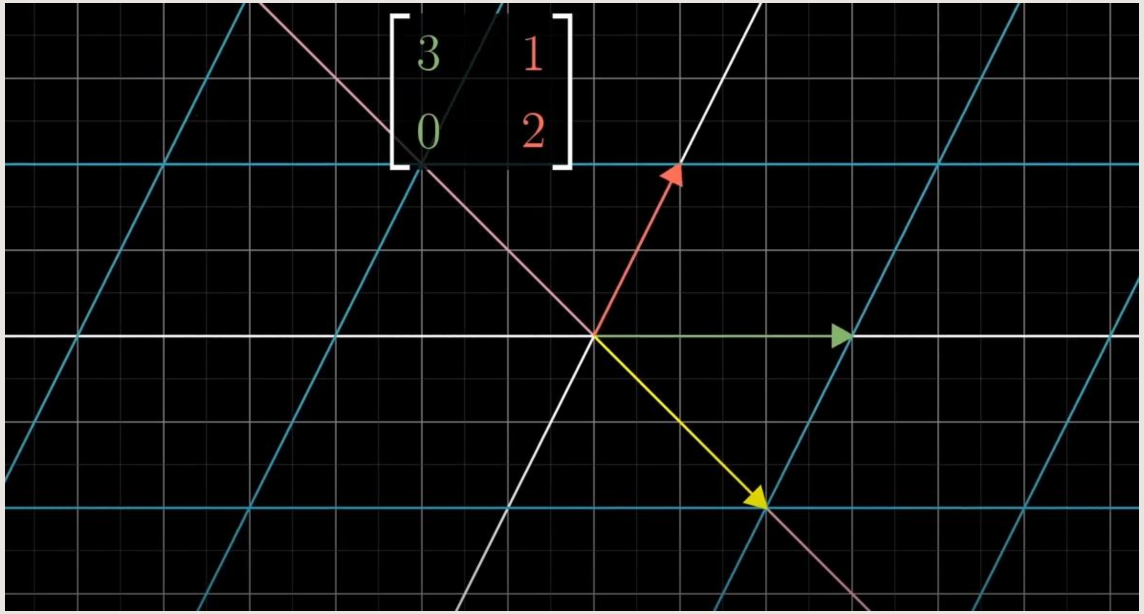




Macierz

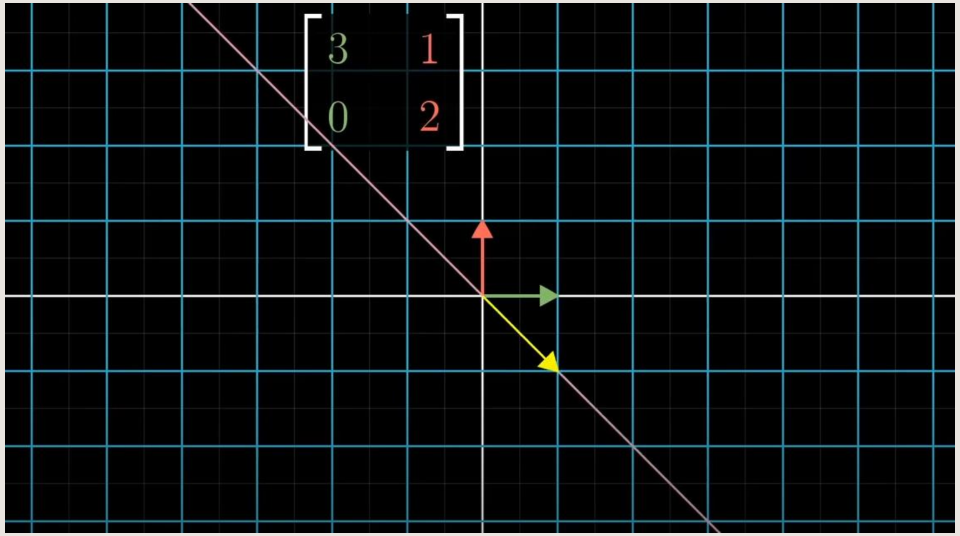
Te wektory są wyjątkowe, ponieważ za ich pomocą możemy reprezentować całą macierz. Szukanie składowych głównych jest właśnie poszukiwaniem wektorów własnych macierzy zbudowanej z naszych danych wejściowych.

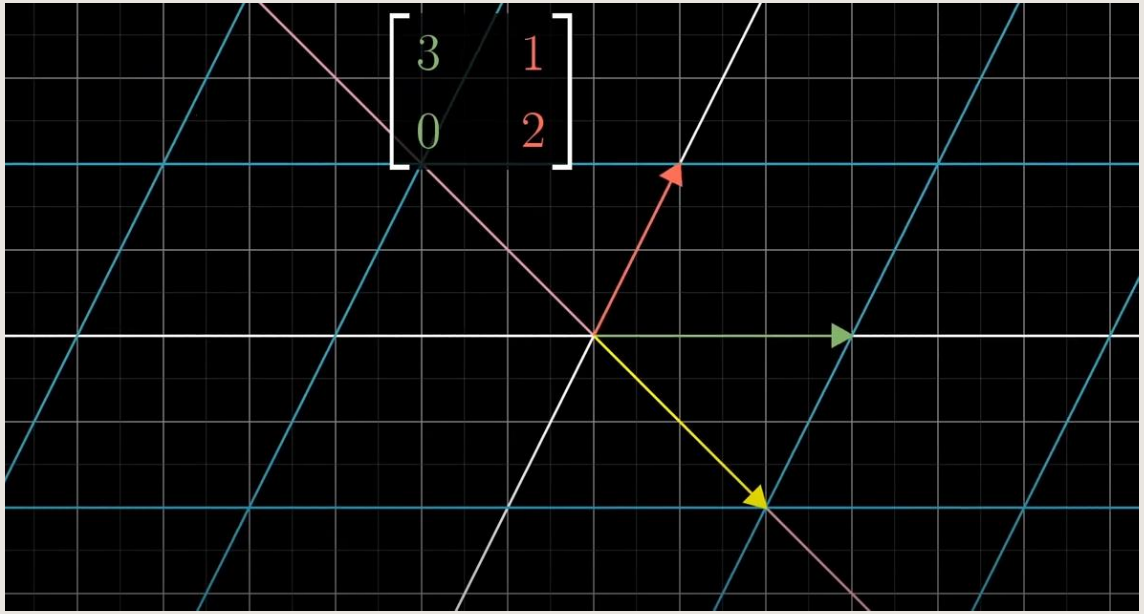




Macierz

Teraz, gdyby przedstawić naszą macierz w taki sposób, żeby pokazywała jak są transponowane wektory własne otrzymalibyśmy macierz diagonalną (czyli taką, która ma same zera poza wartościami na przekątnej głównej - diagonali). Byłaby macierzą diagonalną, ponieważ te wektory własne w wyniku działania macierzy są wyłącznie przeskalowywane, a takie działanie mogą mieć wyłącznie macierze diagonalne.

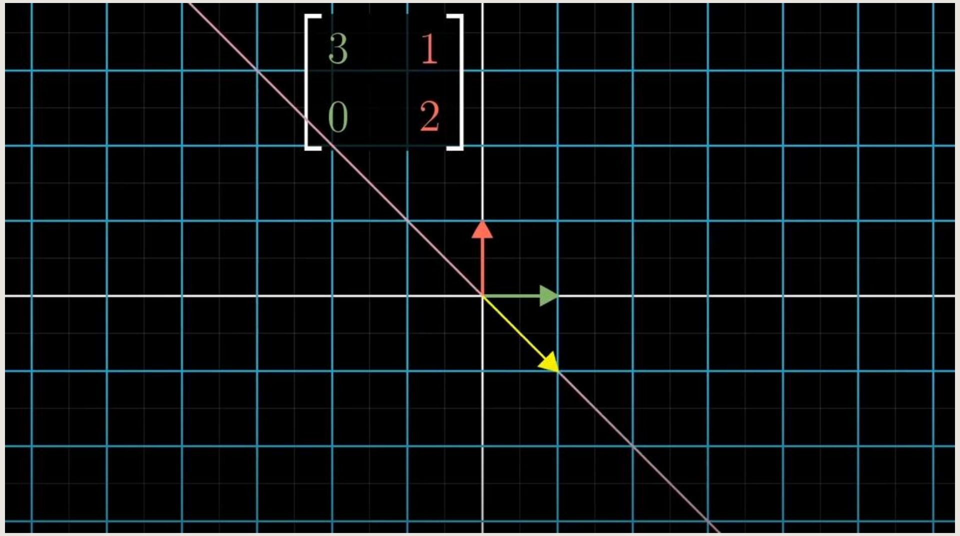


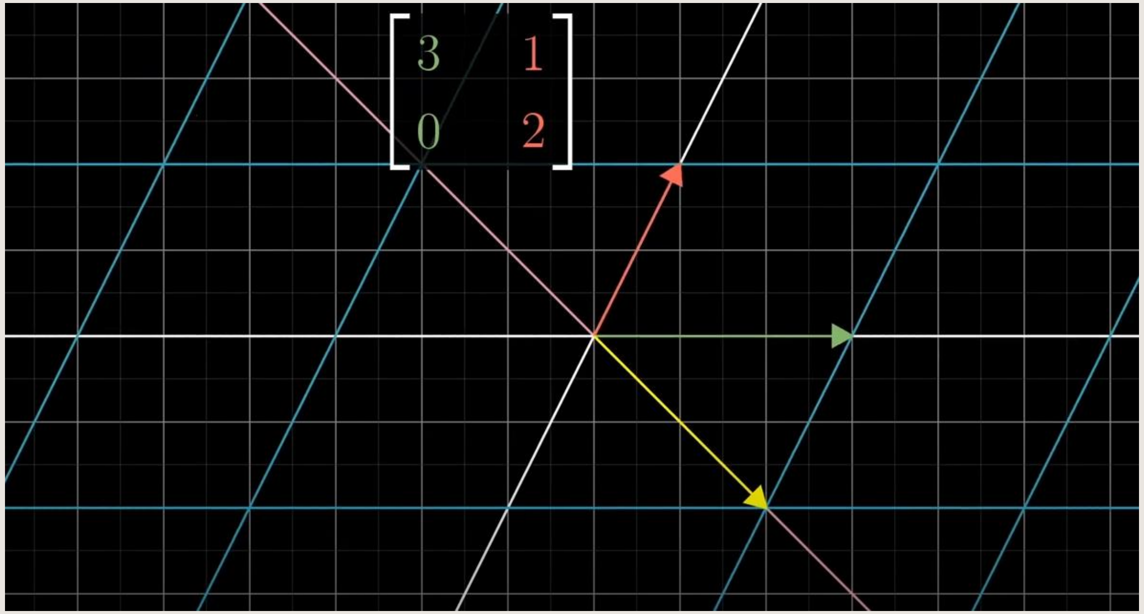


Macierz

Teraz, gdyby przedstawić naszą macierz w taki sposób, żeby pokazywała jak są transponowane wektory własne otrzymalibyśmy macierz diagonalną (czyli taką, która ma same zera poza wartościami na przekątnej głównej - diagonali). Byłaby macierzą diagonalną, ponieważ te wektory własne w wyniku działania macierzy są wyłącznie przeskalowywane, a takie działanie mogą mieć wyłącznie macierze diagonalne.

$$\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3x \\ 2y \end{bmatrix}$$





Macierz

Bardzo lubimy macierze diagonalne, ponieważ bardzo prosto liczy się na nich operacje. Powiedzmy, że chcemy 100 razy przemnożyć tę macierz przez wektor. Wynik znamy natychmiast

$$\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \cdot \ldots \cdot \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3^{100} \cdot x \\ 2^{100} \cdot y \end{bmatrix}$$

Macierz

Tylko w jaki sposób przekształcić macierz tak, żeby reprezentowała transformacje wektorów własnych. Tutaj wchodzi temat faktoryzacji macierzy, czyli przedstawiania macierzy na różne sposoby, rozbijania na różne macierze pośrednie, których złożenie daje naszą macierz (operacje pośrednie). Istnieje wiele sposób faktoryzacji macierzy. Do najpopularniejszych należą:

- Metoda eliminacji $A = LU$
- Metoda Grama-Schmidta (ortogonalizacji) $A = QR$
- Rozkład na wektory własne $S = Q \lambda Q^T$
- Rozkład na wartości osobliwe (SVD, Singular Value Decomposition) $A = U \Sigma U^T$

Implementując PCA najczęściej korzysta się właśnie z rozkładu na wartości własne lub z rozkładu na wartości osobliwe (SVD)