...

Robert Susmaga

Instytut Informatyki

ul. Piotrowo 2

Poznań

kontakt mail'owy

Robert.Susmaga@CS.PUT.Poznan.PL

kontakt osobisty

Centrum Wykładowe, „blok informatyki", pok. 7

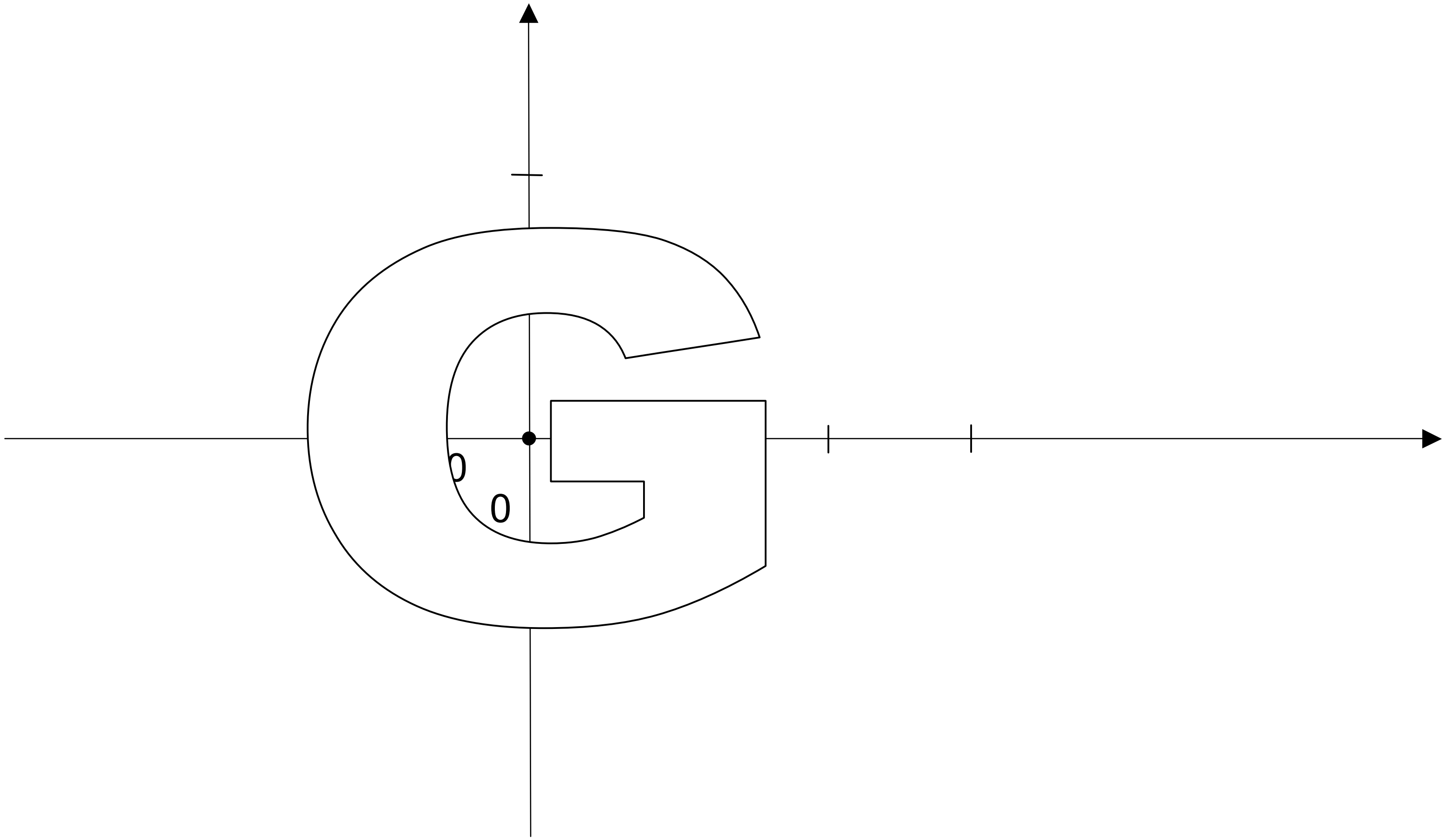
Wyłączenie odpowiedzialności

Prezentowane materiały, będące dodatkiem pomocniczym do wykładów, z konieczności fragmentarycznym i niedopracowanym, należy wykorzystywać z pełną świadomością faktu, że mogą nie być pozbawione przypadkowych błędów, braków, wypaczeń i przeinaczeń :-)

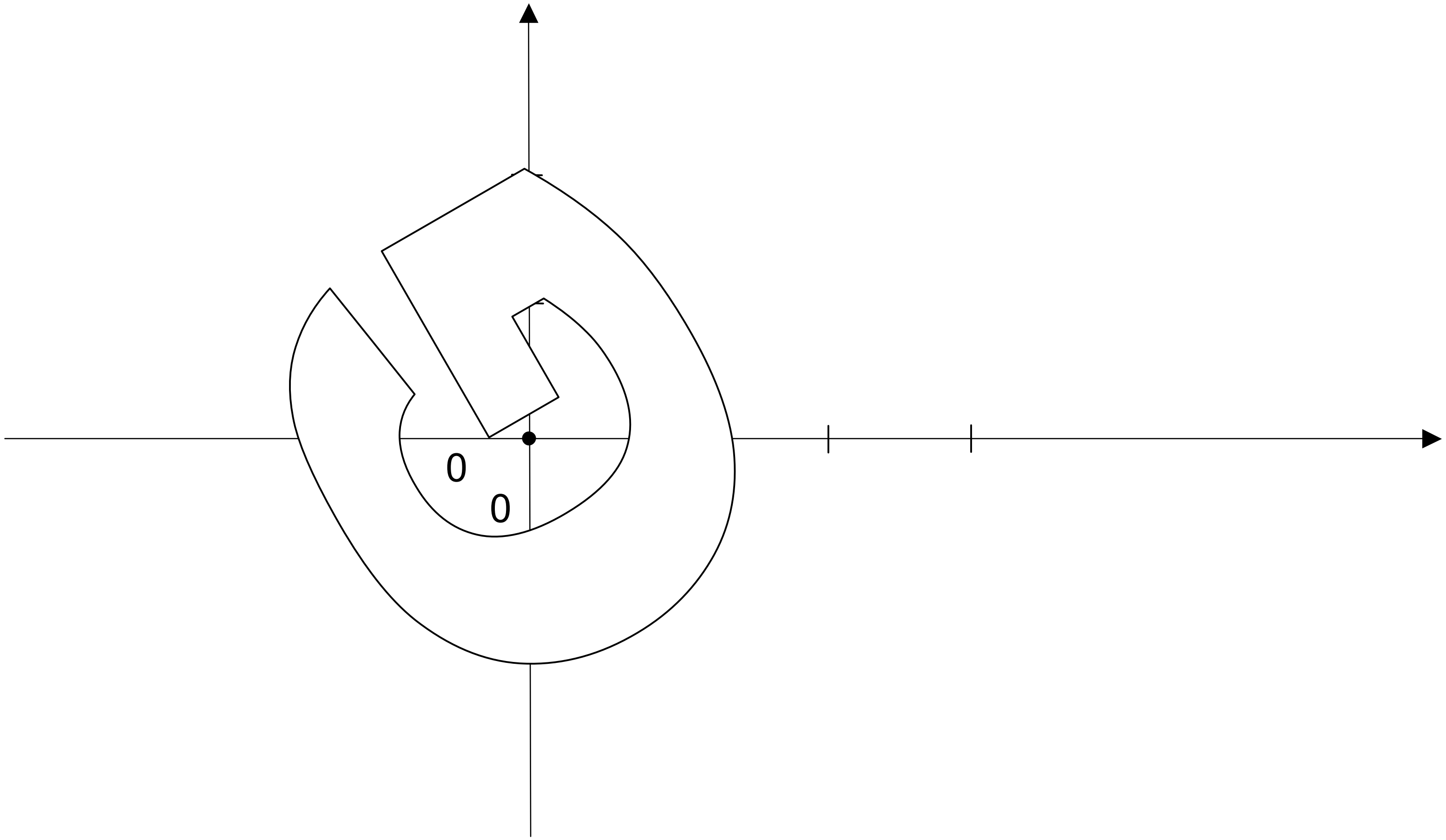
Autor

…

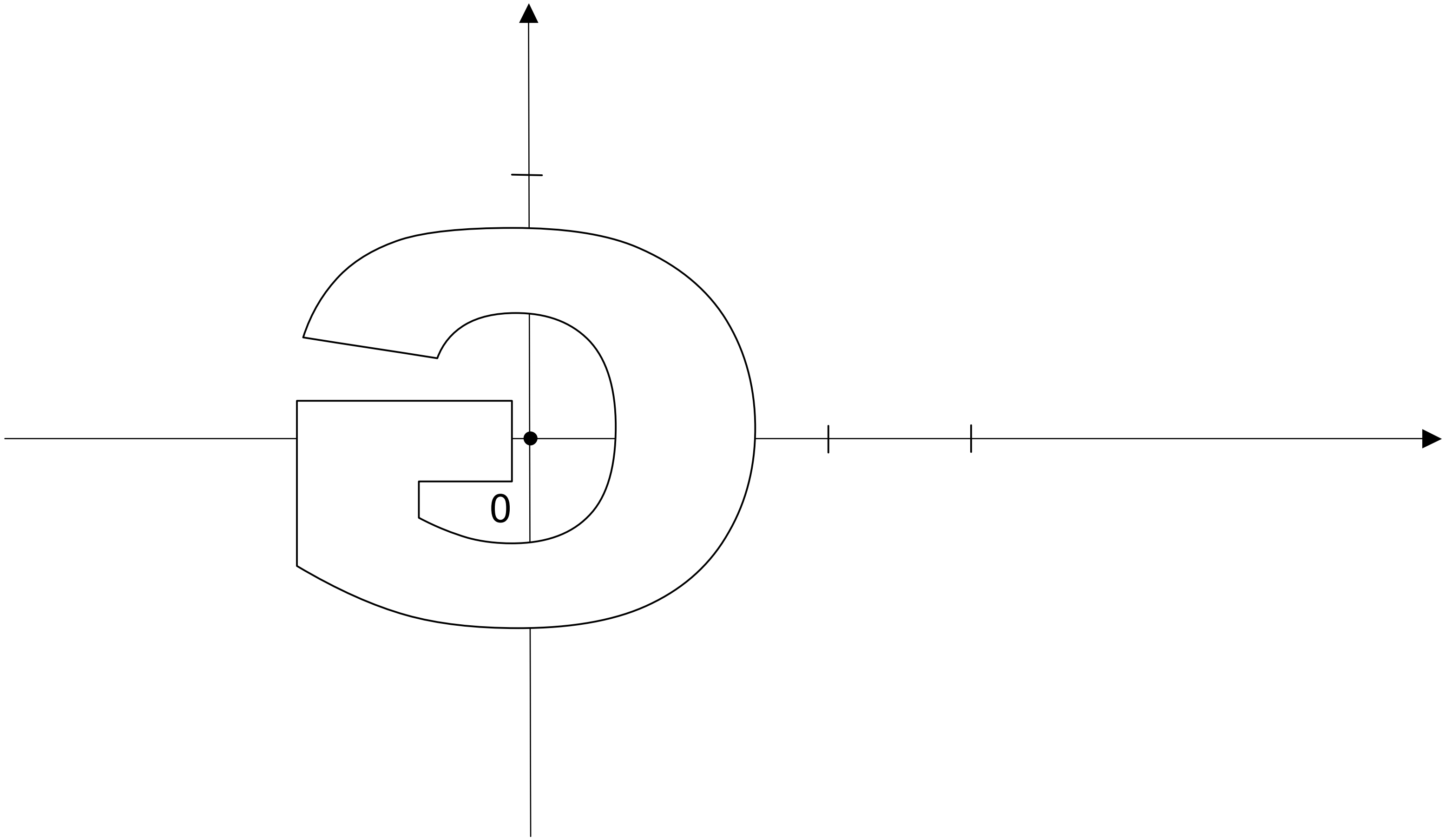
Co ujawnia EVD?



Co ujawnia EVD?



Co ujawnia EVD?



Co ujawnia EVD?

- Uogólniona macierz obrotu:
  (macierz obrotów uogólnionych na zmiany znaku)

  macierz ortogonalna

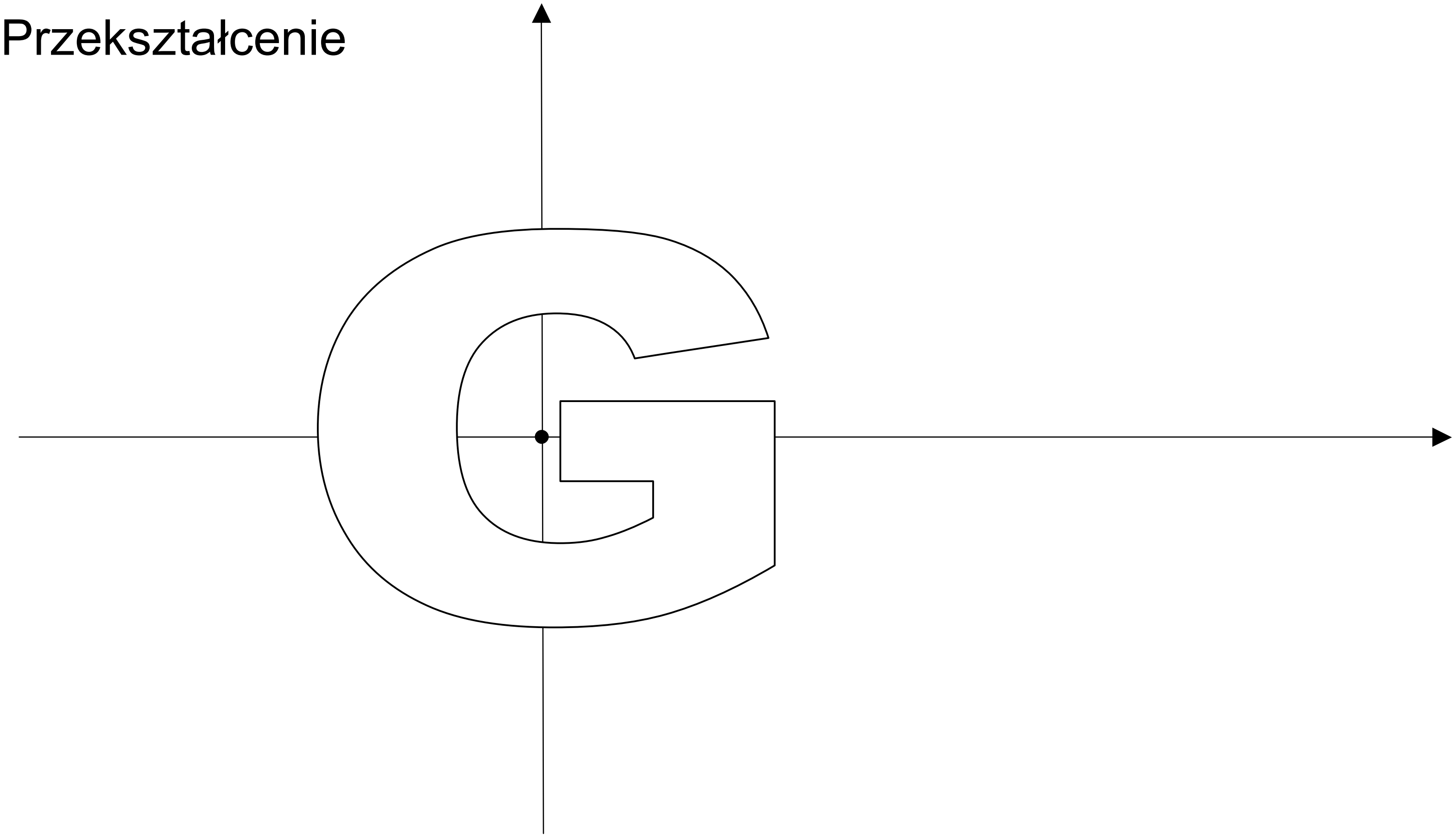
  $\mathbf{K}^T\mathbf{K} = \mathbf{I}$
- Przykłady

  $\mathbf{I}$: $\mathbf{I}^T\mathbf{I} = \mathbf{I}$

  $-\mathbf{I}$: $(-\mathbf{I})^T(-\mathbf{I}) = (-\mathbf{I})(-\mathbf{I}) = \mathbf{I}$
- I podobnie: jeżeli **K** jest ortogonalna, to –**K** jest także ortogonalna
  - operacja –**K**: zmiana znaku wszystkich kolumn/wierszy
- Analogiczne „zjawisko” zachodzi także dla wybranych wierszy/kolumn,
  - jeżeli **K** jest ortogonalna, to macierz **Z**, powstała z **K** przez zmianę znaku wybranych kolumn/wierszy, jest także ortogonalna

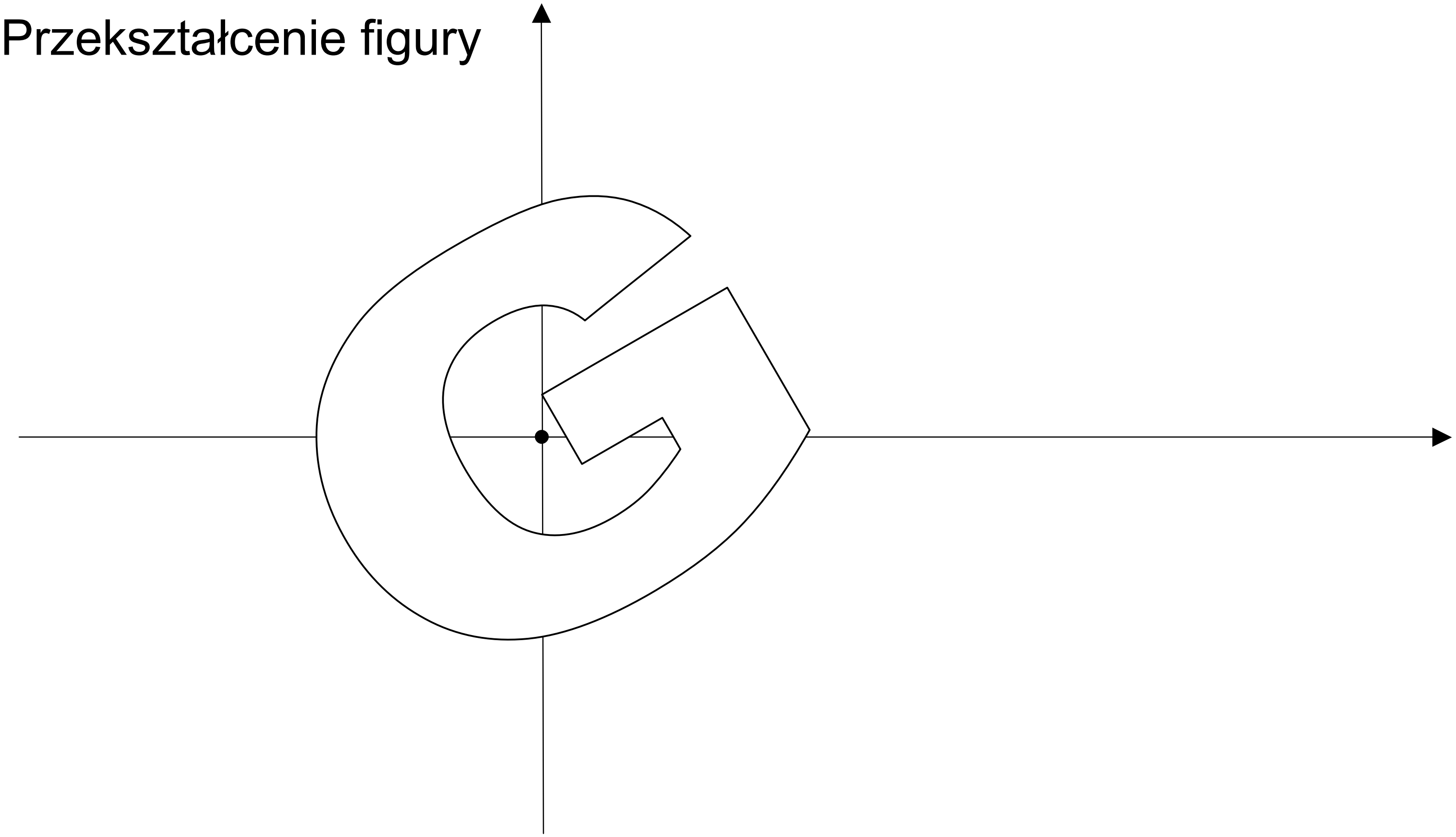
Co ujawnia EVD?

- Przekształcenie



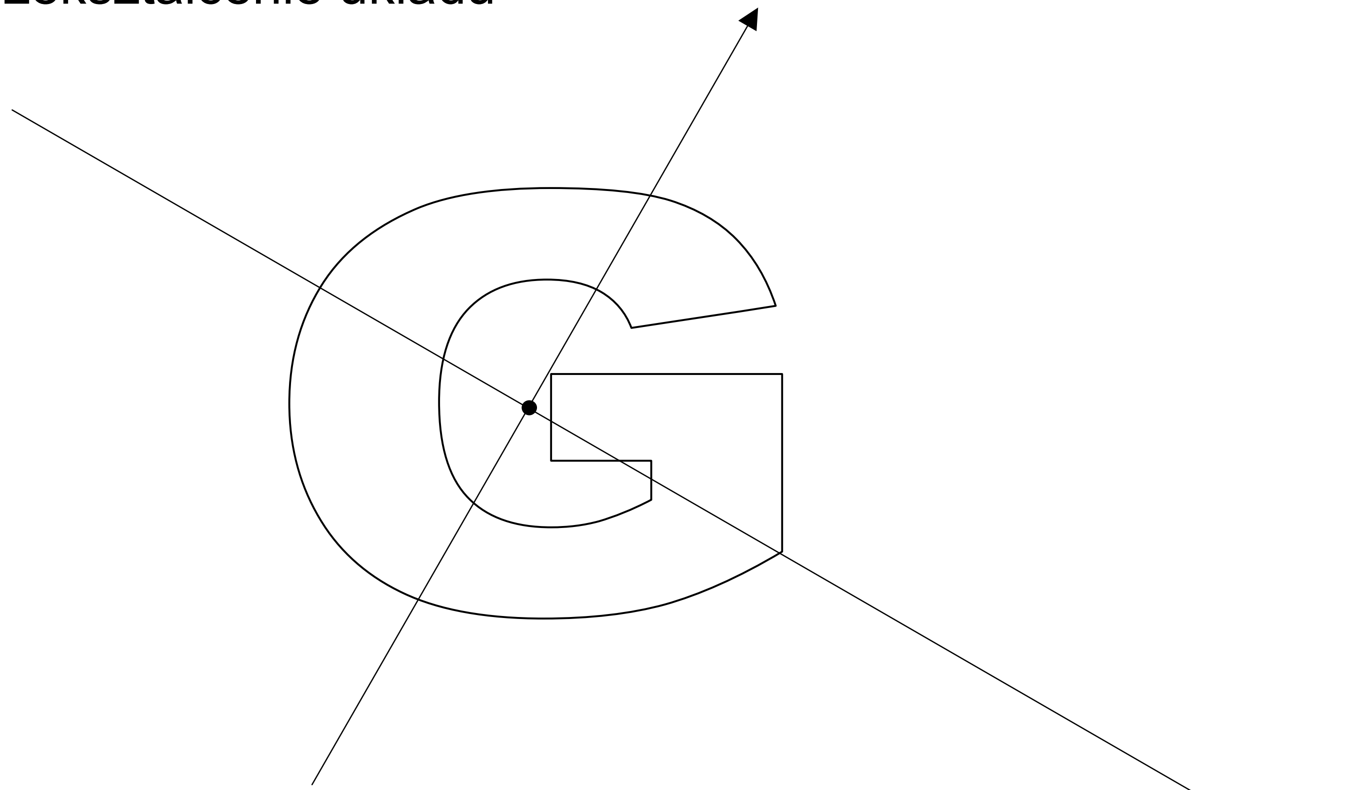
Co ujawnia EVD?

- Przekształcenie figury



Co ujawnia EVD?

- Przekształcenie układu



Co ujawnia EVD?

- Co ujawnia EVD o macierzy przekształcającej?
  - niech dane będzie przekształcenie wektora danych **x** polegające na jego (prawostronnym) przemnożeniu przez (przekształcającą) macierz kwadratową **A**, produkujące wektor **y** = **Ax**
    - przekształcenie to jest równoważne przekształceniu $\mathbf{y}^T = \mathbf{x}^T\mathbf{A}^T$
      - gdzie $\mathbf{A}^T$ jest także kwadratową macierzą przekształcającą

Co ujawnia EVD?

- Przypadki szczególne
  - jeżeli **A** jest macierzą diagonalną, to operacja stanowi jednokładność przeskalowanie wzdłuż poszczególnych osi
  - jeżeli **A** jest macierzą ortogonalną, to operacja stanowi uogólnienie obrotu (obrót uogólniony o ewentualne zmiany znaku)
- A co w ogólnym przypadku?
  - (pominięte)

Co ujawnia EVD?

- Pytanie
  - czy rozkład EVD macierzy przekształcającej (o ile istnieje) ujawnia coś więcej o tym przekształceniu?
  - założymy więc, że macierz przekształcająca posiada rozkład EVD

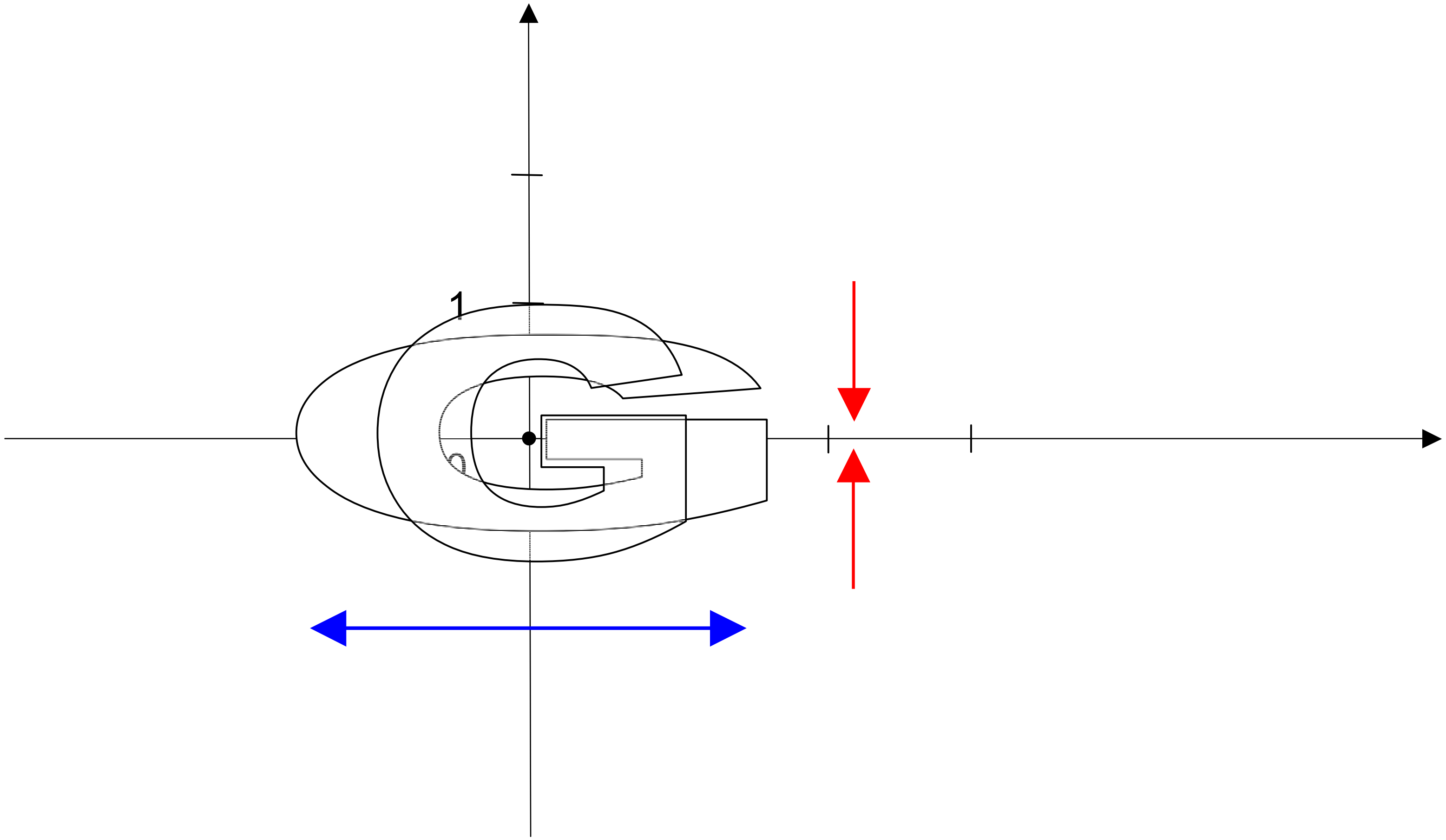
Co ujawnia EVD?

- W przypadku, gdy **A** posiada rozkład EVD: $\mathbf{A} = \mathbf{KLK}^{-1}$
  - (pominięte)
- W przypadku, gdy **A** posiada rozkład EVD: $\mathbf{A} = \mathbf{KLK}^{T}$
  - wtedy $\mathbf{y} = \mathbf{Ax}$ można zapisać jako $\mathbf{y} = \mathbf{KLK}^{T}\mathbf{x}$ i interpretować „po kolei”, wykorzystując $\mathbf{y} = \mathbf{Kv}$, gdzie $\mathbf{v} = \mathbf{Lu}$, gdzie $\mathbf{u} = \mathbf{K}^{T}\mathbf{x}$
    - czyli $\mathbf{y} = \mathbf{K}(\mathbf{L}(\mathbf{K}^{T}\mathbf{x}))$
    - pierwsza operacja: $\mathbf{u} = \mathbf{K}^{T}\mathbf{x}$ (mnożenie **x** przez macierz ortogonalną)
    - druga operacja: $\mathbf{v} = \mathbf{Lu}$ (mnożenie **u** przez macierz diagonalną)
    - trzecia operacja: $\mathbf{y} = \mathbf{Kv}$ (mnożenie **v** przez macierz ortogonalną)
  - ponieważ macierze **K** i $\mathbf{K}^{T}$ są (w tym przypadku) własnymi odwrotnościami, operacja trzecia jest odwrotnością operacji pierwszej
    - pierwsza „przechodzi” do nowego układu współrzędnych
    - pierwsza „wraca” do oryginalnego układu współrzędnych

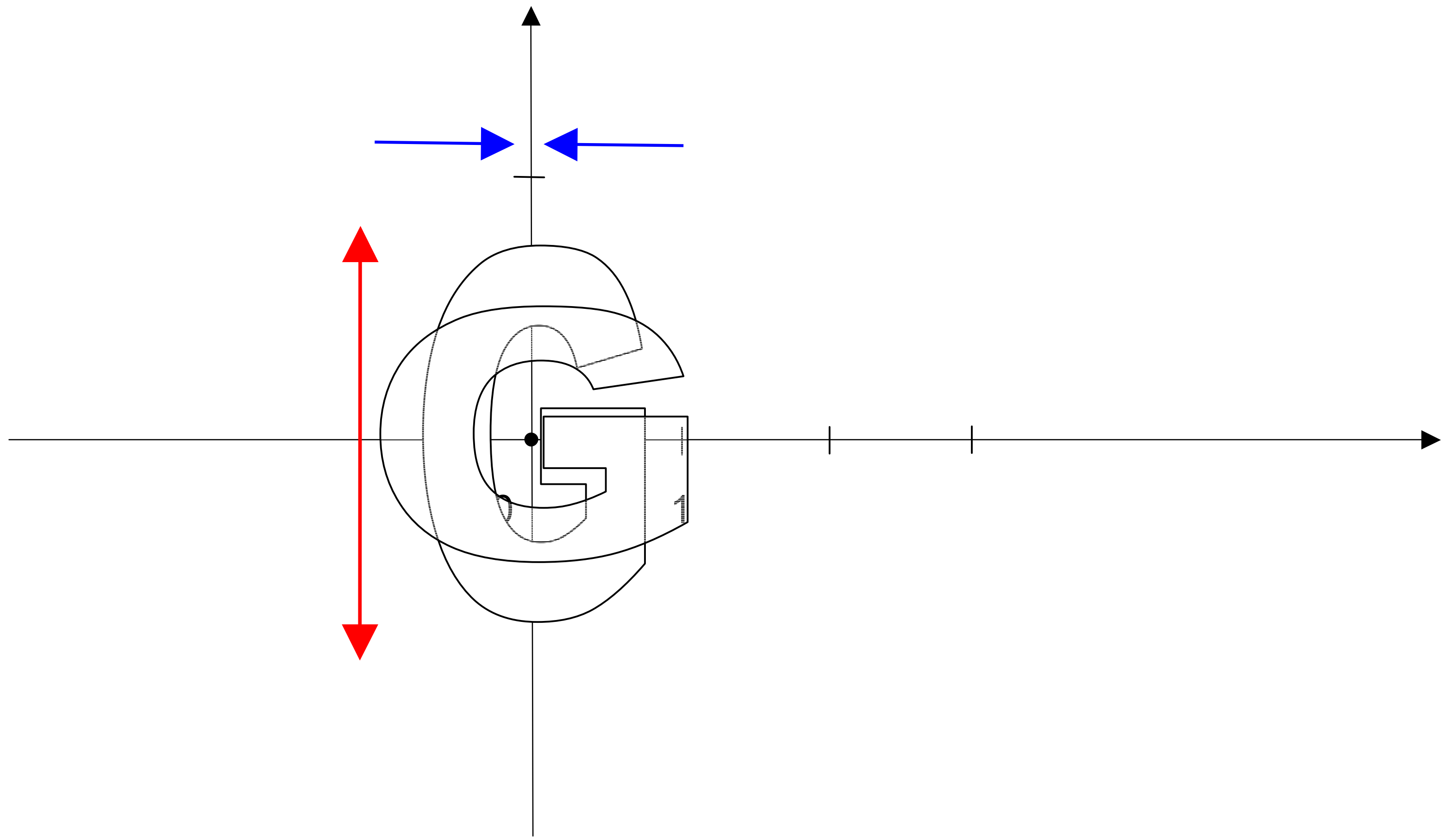
Co ujawnia EVD?

- Wniosek:
  – przekształcenie EVD ujawnia, że w mnożenie przez (zakładaną) macierz przekształcającą „zawiera w sobie” mnożenie przez macierz diagonalną utworzoną z wartości własnych macierzy przekształcającej, co ma oczywiste konsekwencje dla rezultatów przekształcenia
    - wartość własna > 1: zwiększanie rozmiaru figury wzdłuż danego wymiaru
    - wartość własna < 1: zmniejszanie rozmiaru figury wzdłuż danego wymiaru
    - wartość własna = 0: redukcja danego wymiaru

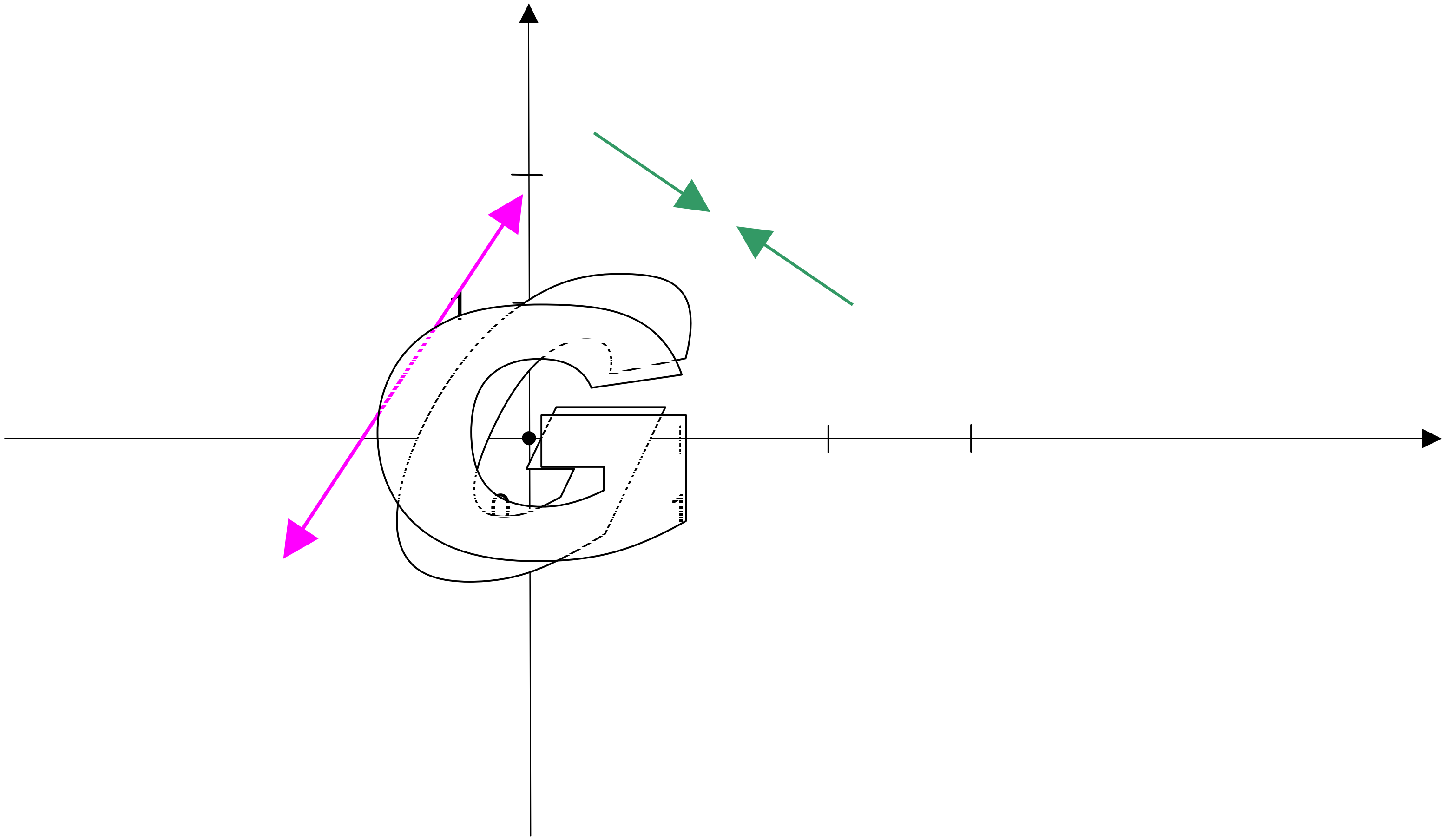
Co ujawnia EVD?



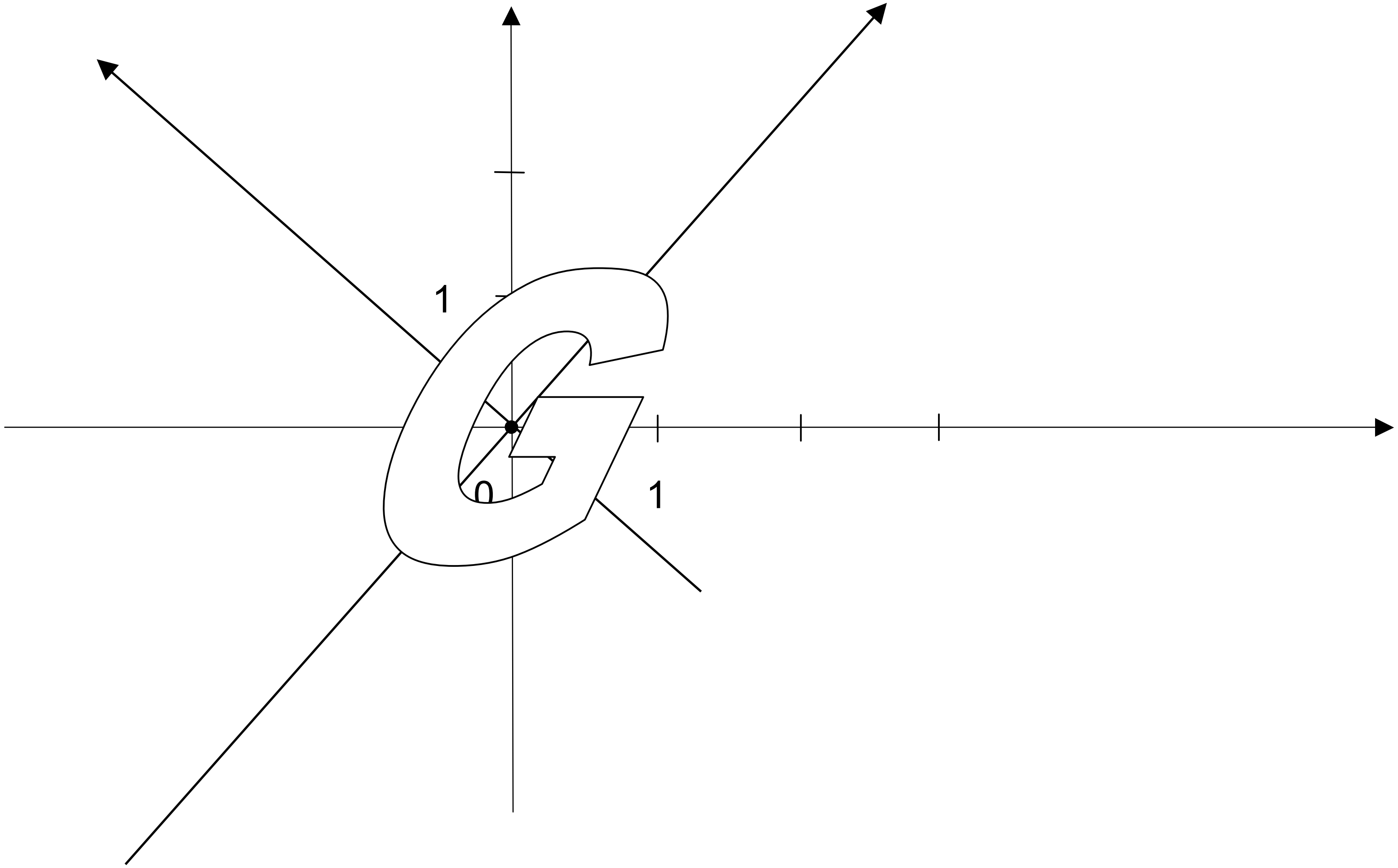
Co ujawnia EVD?



Co ujawnia EVD?



Co ujawnia EVD?



…

PCA jako transformacja układu współrzędnych

- Właściwości wykorzystywanej w przekształceniu PCA macierzy przekształcającej **K**:
  - macierz **K** charakteryzuje się:
    - unormowaniem kolumn: $(\mathbf{k}_i)^T\mathbf{k}_i = 1$ dla każdego i
    - ortogonalnością kolumn: $(\mathbf{k}_i)^T\mathbf{k}_j = 0$ dla każdej pary $i \neq j$
  - wniosek: kolumny macierzy **K** definiują pewien układ współrzędnych (bazę) – kolumny $\mathbf{k}_i$ są wersorami tego układu
    - ponieważ kolumny $\mathbf{k}_i$ (a tym samym reprezentowane przez nie wersory) są unormowane i ortogonalne, to definiowany układ współrzędnych jest układem kartezjańskim
    - kolumny/wersory $\mathbf{k}_i$ tego układu są tak dobrane, aby reprezentować kierunki wzdłuż których wariancja zmiennych jest maksymalna
- Zadaniem procedury PCA jest więc przekształcenie danych do układu zdefiniowanego przez macierz **K**

PCA jako transformacja układu współrzędnych

- Niech będzie dana baza **A** oraz pewien wektor **y**
  - pytanie: jak zdobyć współrzędne wektora **y** w bazie **B**?
  - odpowiedź: współrzędne wektora **y** w bazie **A** przestawia taki wektor **x**, który spełnia równanie macierzowe $\mathbf{Ax} = \mathbf{b}$
    - wektor ten jest rozwiązaniem tego równania, a więc można go obliczyć np. jako $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$
- Procedura PCA dokonuje przekształcenia danych z macierzy **X** do nowego układu współrzędnych zdefiniowanego przez macierz **K**, jest jednak realizowana przez zwykłe mnożenie macierzy: $\mathbf{Y} = \mathbf{XK}$
  - zachodzą pytania:
    - na jakim etapie następuje poszukiwanie nowych współrzędnych?
    - kiedy następuje poszukiwanie odwrotności macierzy **K**?

PCA jako transformacja układu współrzędnych

- Jakie są współrzędne danego wektora $\mathbf{x}$ w układzie reprezentowanym przez ortogonalną macierz $\mathbf{K}$?
    - (czyli przez wersory będące kolumnami tej macierzy)
  - w ogólności, poszukiwane współrzędne (oznaczenie: $\mathbf{y}$) spełniają

    $\mathbf{Ky} = \mathbf{x}$
  - czyli

    $\mathbf{K}^{-1}\mathbf{Ky} = \mathbf{K}^{-1}\mathbf{x}$

    $\mathbf{Iy} = \mathbf{K}^{-1}\mathbf{x}$

    $\mathbf{y} = \mathbf{K}^{-1}\mathbf{x}$

PCA jako transformacja układu współrzędnych

- Jakie są współrzędne danego wektora $\mathbf{x}$ w układzie...
  - ale ortogonalna $\mathbf{K}$ spełnia $\mathbf{K}^{-1} = \mathbf{K}^{T}$, zatem

    $\mathbf{y} = \mathbf{K}^{T}\mathbf{x}$
  - powyższe zachodzi dla wektora $\mathbf{x}$ (kolumnowego), jak to będzie wyglądało dla tego samego wektora w postaci $\mathbf{x}^{T}$ (wierszowej)?
  - wykorzystując $\mathbf{A} = \mathbf{B} \Leftrightarrow \mathbf{A}^{T} = \mathbf{B}^{T}$ mamy

    $\mathbf{y}^{T} = (\mathbf{K}^{T}\mathbf{x})^{T}$

    $\mathbf{y}^{T} = \mathbf{x}^{T}(\mathbf{K}^{T})^{T}$

    $\mathbf{y}^{T} = \mathbf{x}^{T}\mathbf{K}$

PCA jako transformacja układu współrzędnych

- Jakie są współrzędne danego wektora **x** w układzie...
  - czyli dla wektorów $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ zachodzi

    $(\mathbf{y}_1)^T = (\mathbf{x}_1)^T\mathbf{K}$

    $(\mathbf{y}_2)^T = (\mathbf{x}_2)^T\mathbf{K}$

    ...

    $(\mathbf{y}_m)^T = (\mathbf{x}_m)^T\mathbf{K}$
  - to samo prościej
    - (traktując wektory $(\mathbf{x}_1)^T, (\mathbf{x}_2)^T, \ldots, (\mathbf{x}_m)^T$ jako wiersze macierzy **X**)

    $\mathbf{Y} = \mathbf{XK}$
- Wniosek:
  - tworząc nowe zmienne (jako $\mathbf{Y} = \mathbf{XK}$), metoda PCA znajduje w praktyce współrzędne wektorów reprezentowanych przez wiersze macierzy **X** w układzie reprezentowanym przez kolumny macierzy **K**

PCA jako transformacja układu współrzędnych

- Jakie są współrzędne danego wektora **x** w układzie...
  - istotną rolę w uproszczeniu powyższego zapisu odgrywa fakt, że macierz **K** jest ortogonalna
  - co by było, gdyby nie była to macierz ortogonalna?
    - gdyby **K** była osobliwa, to nie mogłaby reprezentować wersorów układu współrzędnych
    - gdyby **K** była nieosobliwa (co gwarantuje istnienie $\mathbf{K}^{-1}$), ale nie byłaby ortogonalna, to zachodziłoby w ogólności

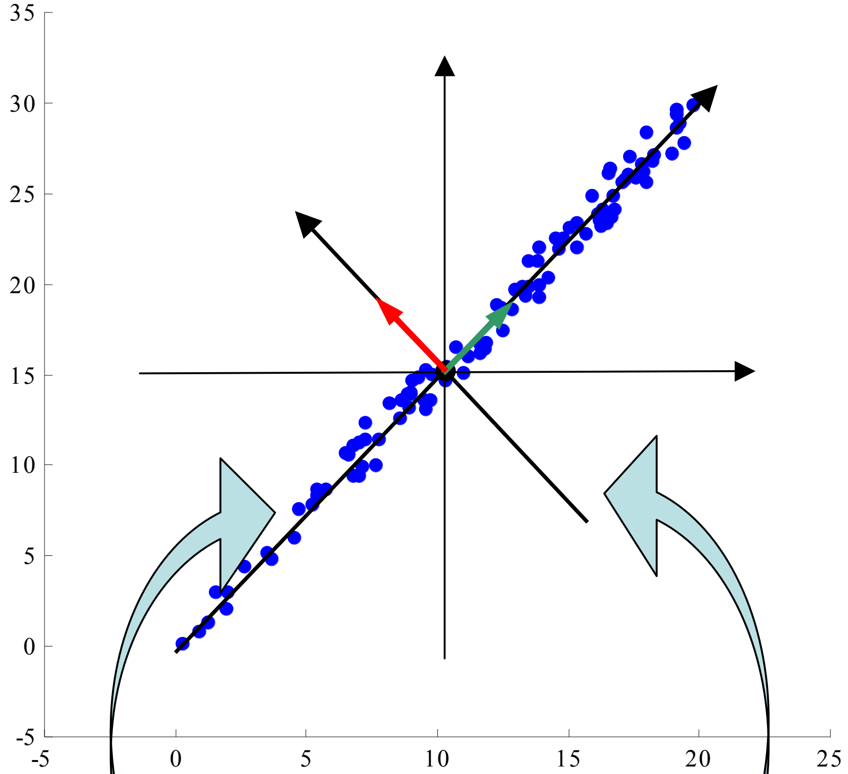
      $\mathbf{y} = \mathbf{K}^{-1}\mathbf{x}$
      - i wtedy

        $\mathbf{y}^T = (\mathbf{K}^{-1}\mathbf{x})^T$

        $\mathbf{y}^T = \mathbf{x}^T(\mathbf{K}^{-1})^T$
      - a w ogólności

        $\mathbf{Y} = \mathbf{X}(\mathbf{K}^{-1})^T$

Przykład: PCA dla danych dwuwymiarowych



K

0.55	0.83
0.83	–0.55

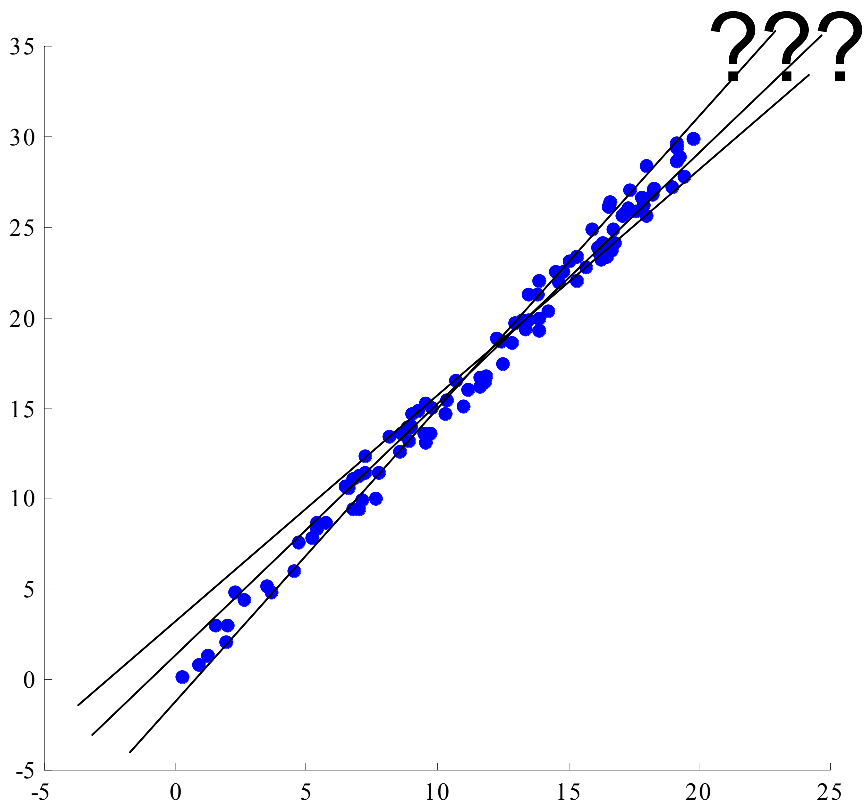
× 20

20·**K**

11.0	16.6
16.6	–11.0

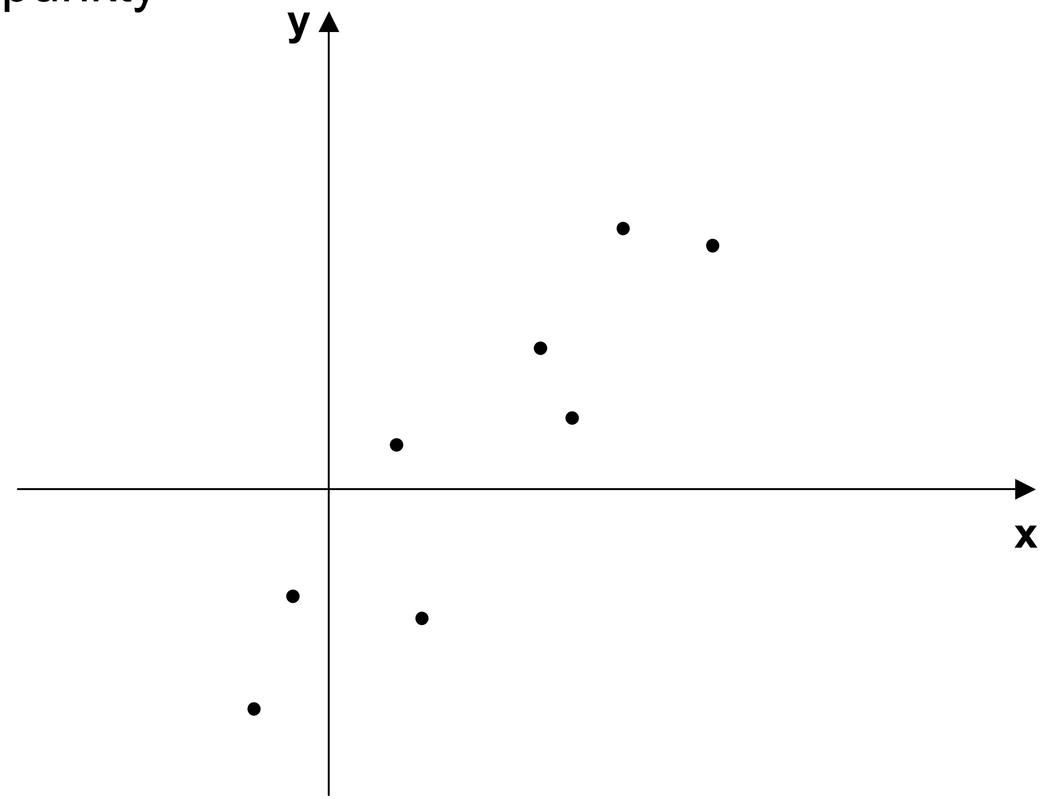
…

PCA a MNK



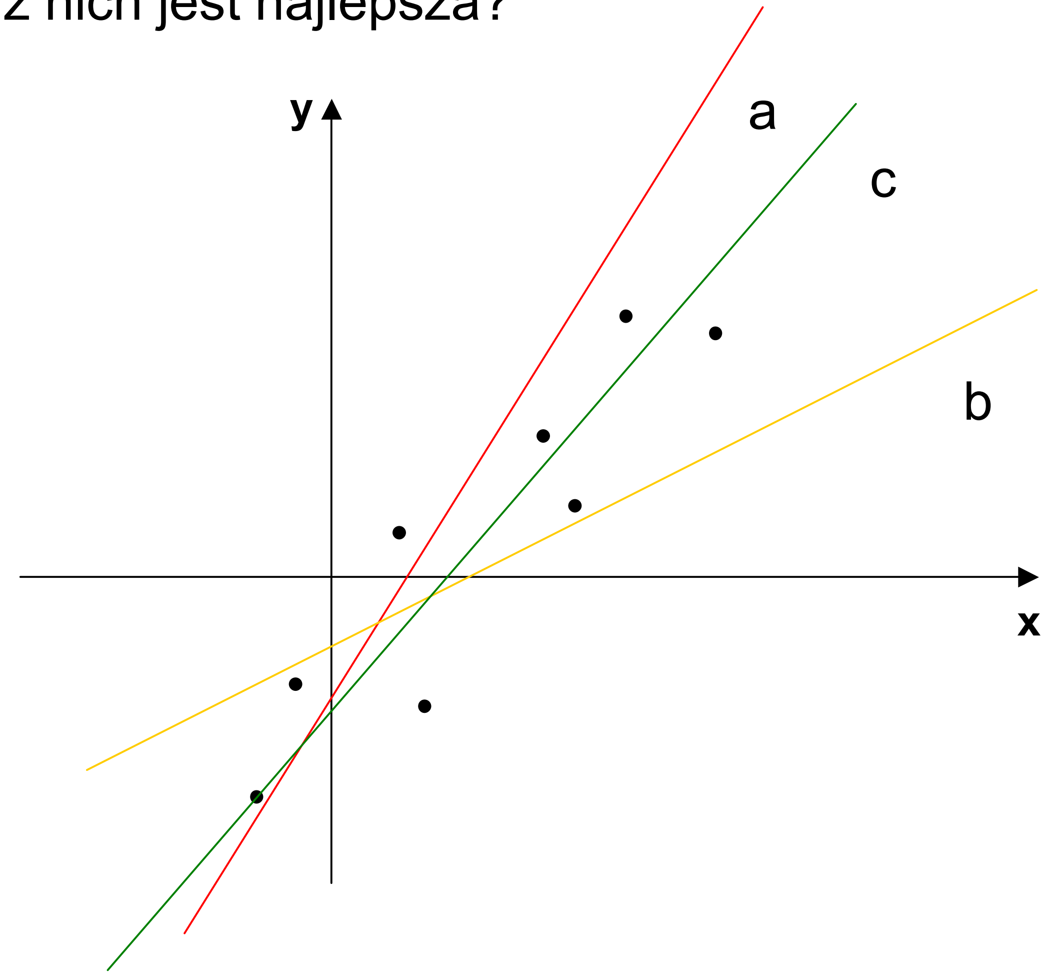
Metoda najmniejszych kwadratów

- Interpretacja geometryczna najmniejszych kwadratów w przestrzeni zmiennych
  - dane punkty



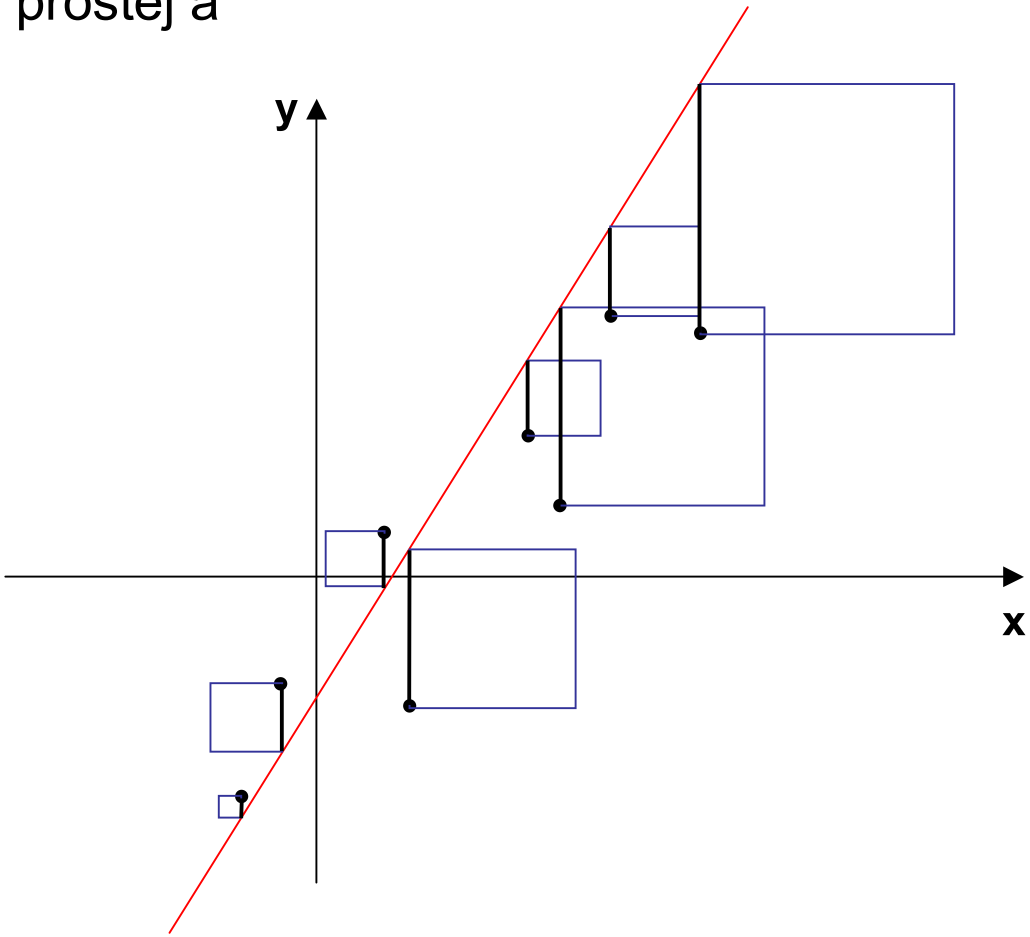
Metoda najmniejszych kwadratów

- Przykładowe modele (proste)
  - która z nich jest najlepsza?



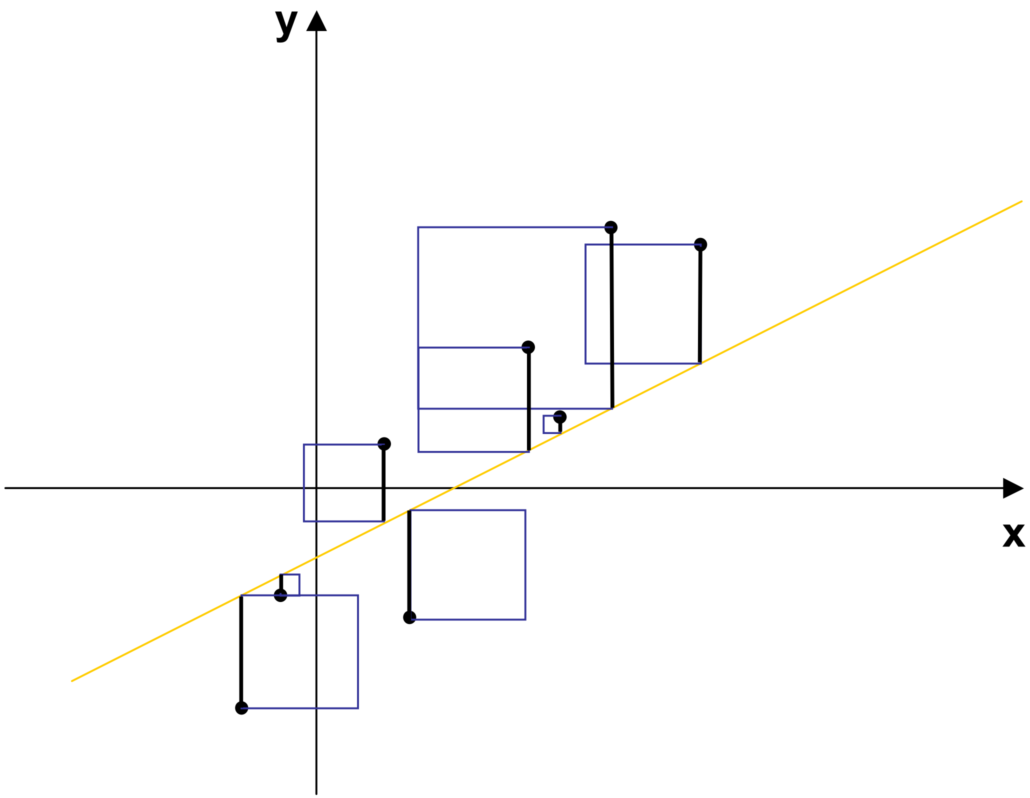
Metoda najmniejszych kwadratów

- Przykładowe modele (proste)
  - ocena prostej a



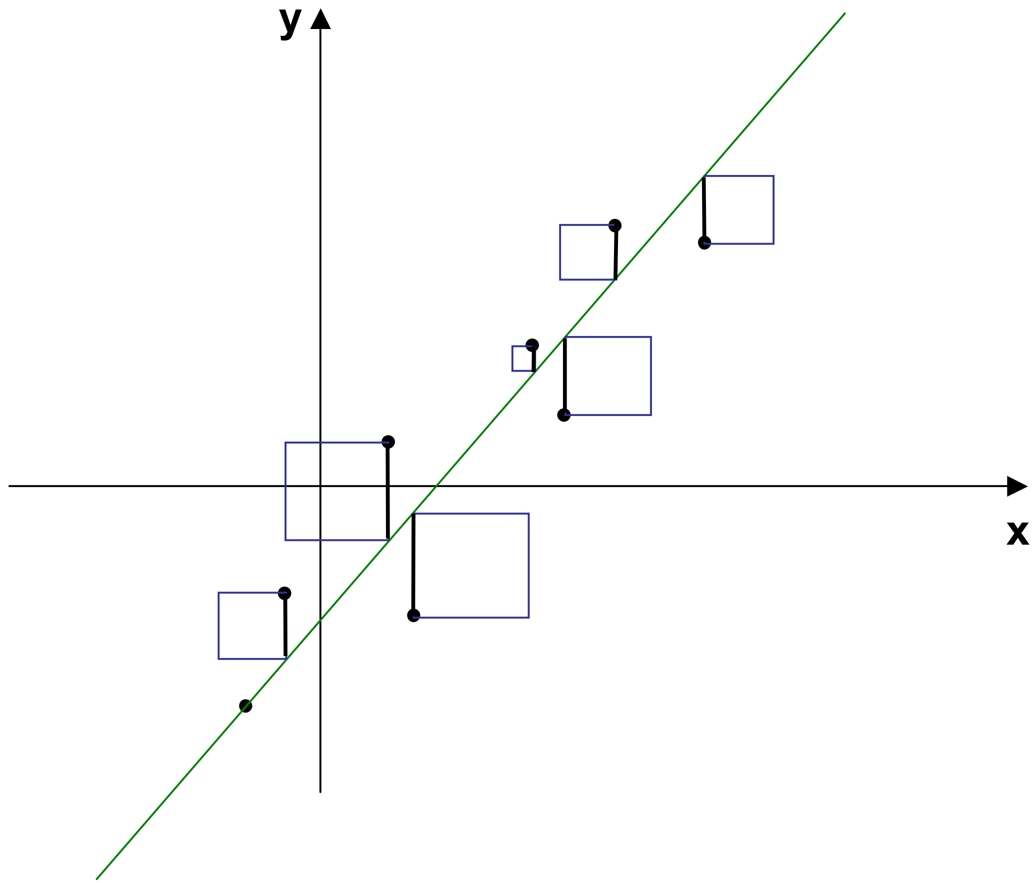
Metoda najmniejszych kwadratów

- Przykładowe modele (proste)
  – ocena prostej b



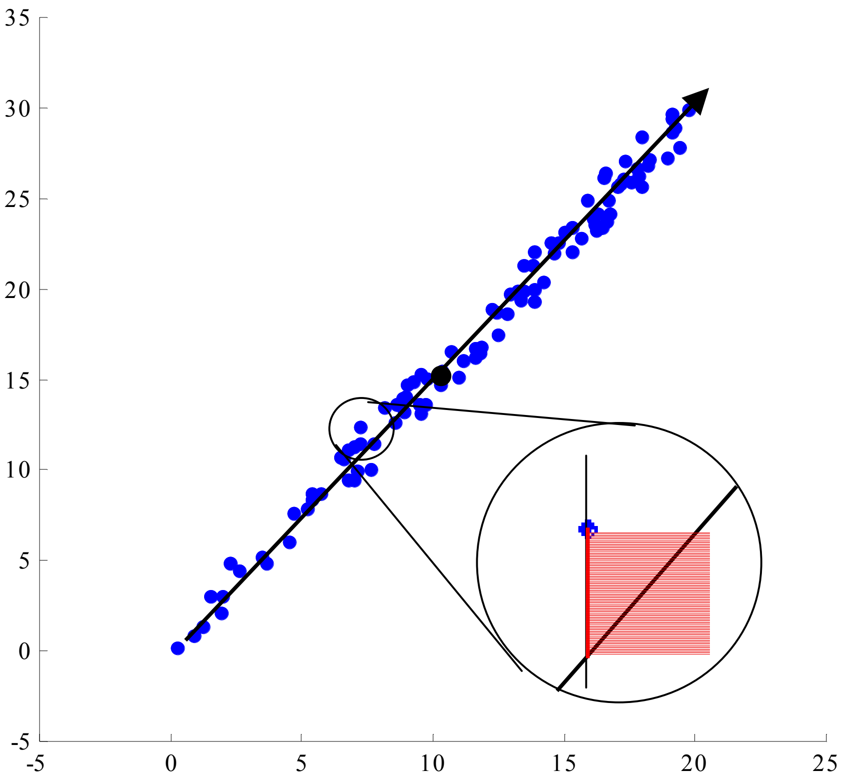
Metoda najmniejszych kwadratów

- Przykładowe modele (proste)
  - ocena prostej c



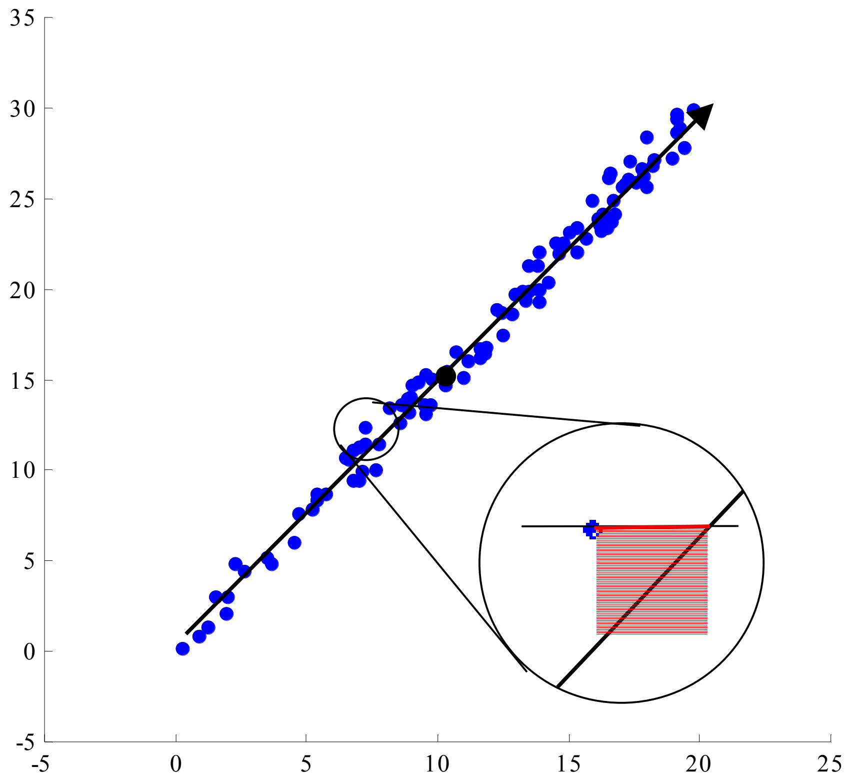
PCA a MNK

- Klasyczna regresja: y = f(x)



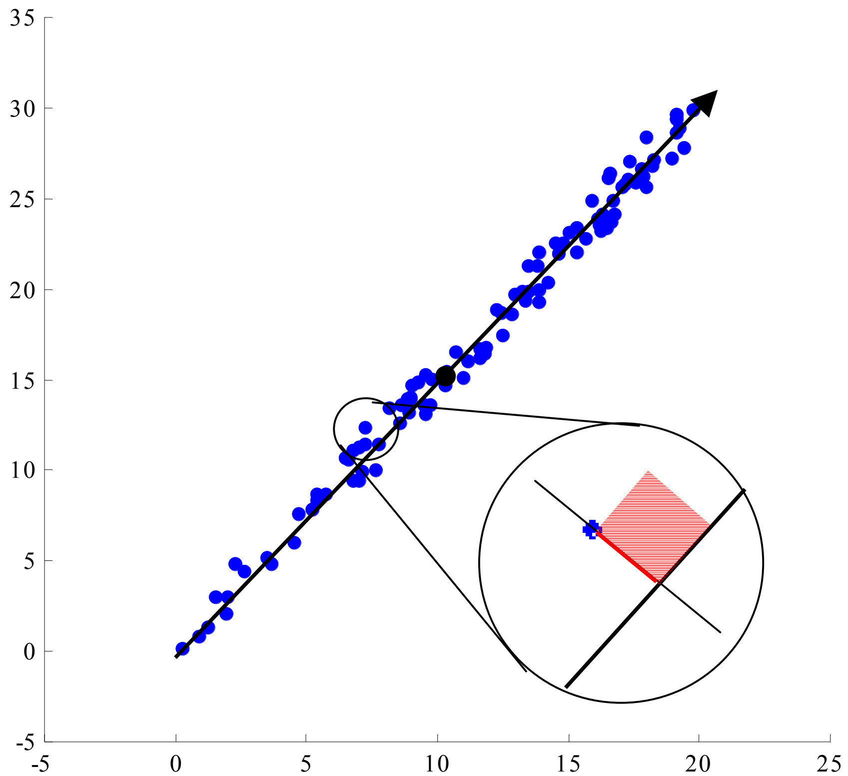
PCA a MNK

- Klasyczna regresja po zamianie zmiennych rolami:
  x = g(y)



PCA a MNK

- PCA



PCA a MNK

- MNK
  - wynik (propozycja prostej) uzyskany dla

    y = f(x) (y zależne od x)

    x = g(y) (x zależne od y)

    będzie identyczny dla $f^{-1} = g$
    (jest to jednakże bardzo szczególny przypadek)
  - w sytuacji braku jasnej sytuacji co do tego, która zmienna jest zależna od której (czyli: która jest zależna, a która niezależna) mamy dwa różne rozwiązania
    - co więcej, jeżeli występują między nimi różnice, to jedno może być lepsze, a drugie gorsze; niestety, zazwyczaj nie wiadomo, które jest tym lepszym rozwiązaniem!
- PCA
  - jeden wynik (jedna propozycja prostej),
    (zarówno dla [**x y**] jak i [**y x**])

…

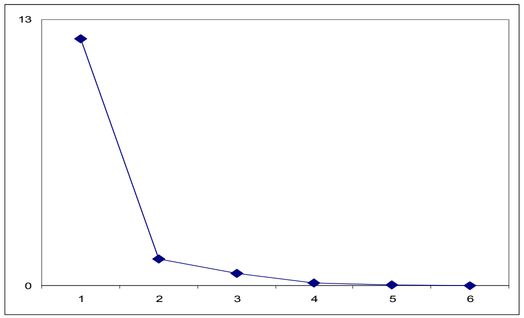
PCA a EVD

- PCA a EVD
  - podsumowanie
    - PCA: procedura przekształcająca dane: $\mathbf{X} \rightarrow \mathbf{Y}$
    - EVD: procedura rozkładająca macierz **A** na iloczyn $\mathbf{KLK}^{-1}$ (inaczej: procedura diagonalizująca macierz **A**)
  - relacja PCA–EVD
    - w procedurze PCA wykorzystuje się procedurę EVD do zdiagonalizowania macierzy kowariancji $\mathbf{S}_X$

…

Metody doboru liczby zmiennych redukowanych

- Niech $\lambda = \{\lambda_1, \lambda_2, ..., \lambda_n\}$ będzie zbiorem posortowanych nierosnąco wartości własnych macierzy kowariancji
  - znajdź średnią wartość $\bar{\lambda}$ wszystkich λ_i i utwórz tylko zmienne odpowiadające wartościom $\lambda_i \geq \bar{\lambda}$
  - utwórz tylko zmienne 1..s, przy minimalnym s, dla którego zachodzi: $\frac{\sum_{i=1}^{s} \lambda_i}{\sum_{i=1}^{n} \lambda_i} \geq p$, przy zadanym p
  - dobierz wizualnie zmienne na podstawie tzw. wykresu osypiska:



…

Zastosowania PCA: przykład biologiczny

- Jolicoeur i Mosiman (1960) dokonywali pomiarów skorupy żółwi, otrzymując oryginalne zmienne:
  - długość
  - szerokość
  - wysokość
- Ze względu na względnie stałe proporcje powyższych wielkości (duża korelacja) zmienne oryginalne można przekształcić wykorzystując PCA i otrzymując:
  - wielkość (98.64% informacji)
  - kształt-A (0.94%) i kształt-B (0.41%)
- Podobne badania
  - białe leghorny (Wright, 1954)

Zastosowania PCA: przykład psychologiczny

- Birren i Morrison (1961) badali wyniki testów Wechslera (testy na inteligencję dla dorosłych). Obserwowano:
  - wyniki testu (11 zmiennych), oraz dodatkowo
  - wiek i wykształcenie
- W rezultacie przekształcenia PCA otrzymano zmienne, które (po zanalizowaniu korelacji z oryginalnymi wynikami testów) zinterpretowano następująco:
  - ogólna wydajność intelektualna (51.47%)
  - doświadczenie (10.90%)
  - miernik wyobraźni przestrzennej (6.15%)
  - miernik umiejętności rachunkowych (5.48%)

…

Samochody osobowe

- Warianty
  - 15 samochodów osobowych
- Kryteria
  - 4 wybrane parametry
    - zużycie paliwa [1/100 l/km], zakres: [6,10], typ: strata
    - przyspieszenie do 100km/h [s], zakres: [9,14], typ: strata
    - cena [tys. PLN], zakres: [24, 94], typ: strata
    - wyposażenie [punkty uznaniowe], zakres: [1, 10], typ: zysk

Wizualizacja danych wielowymiarowych

- Oryginalne opisy 15 samochodów (tabela informacyjna)

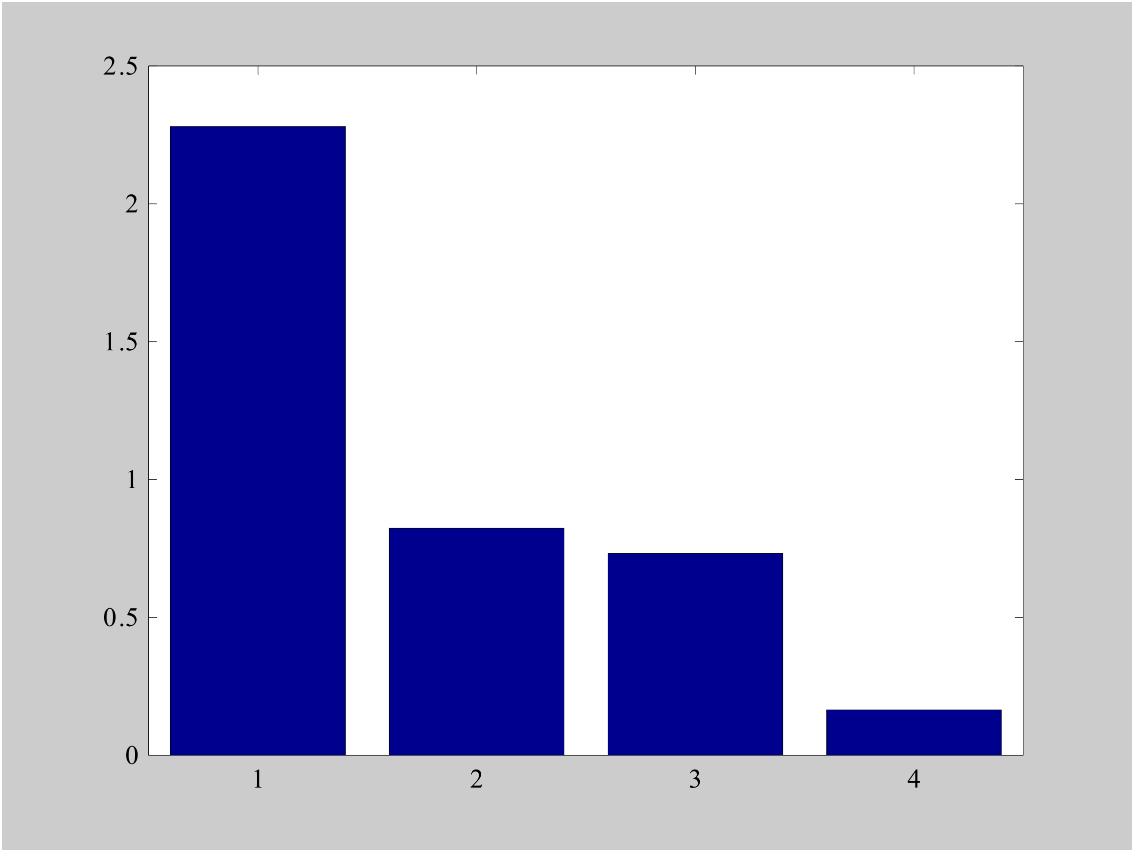
Marka	Model	Zużycie	Przyspieszenie	Cena	Wyposażenie
Alfa Romeo	156	8,1	9,3	71,14	9
Audi	A4	7,9	11,9	93,35	10
BMW	316I	7,5	12,3	81,79	8
Daewoo	Lanos	8,4	12,2	34,90	3
Honda	Civic	6,7	10,8	48,90	7
Hyunday	Accent	6,4	11,7	35,30	2
Lada	Samara	7,2	13,0	24,90	2
Mitsubishi	Carisma	7,2	12,0	60,60	9
Opel	Astra II	7,0	12,0	56,95	8
Peugeot	206 XR	6,6	13,2	38,36	4
Renault	Megane	7,0	9,9	50,05	9
Saab	9 3S	9,7	11,0	85,35	8
Seat	Cordoba	8,3	10,9	44,99	8
Toyota	Corrola	7,7	10,2	50,36	4
Volkswagen	Golf IV	8,3	9,9	61,62	6

Wizualizacja danych wielowymiarowych

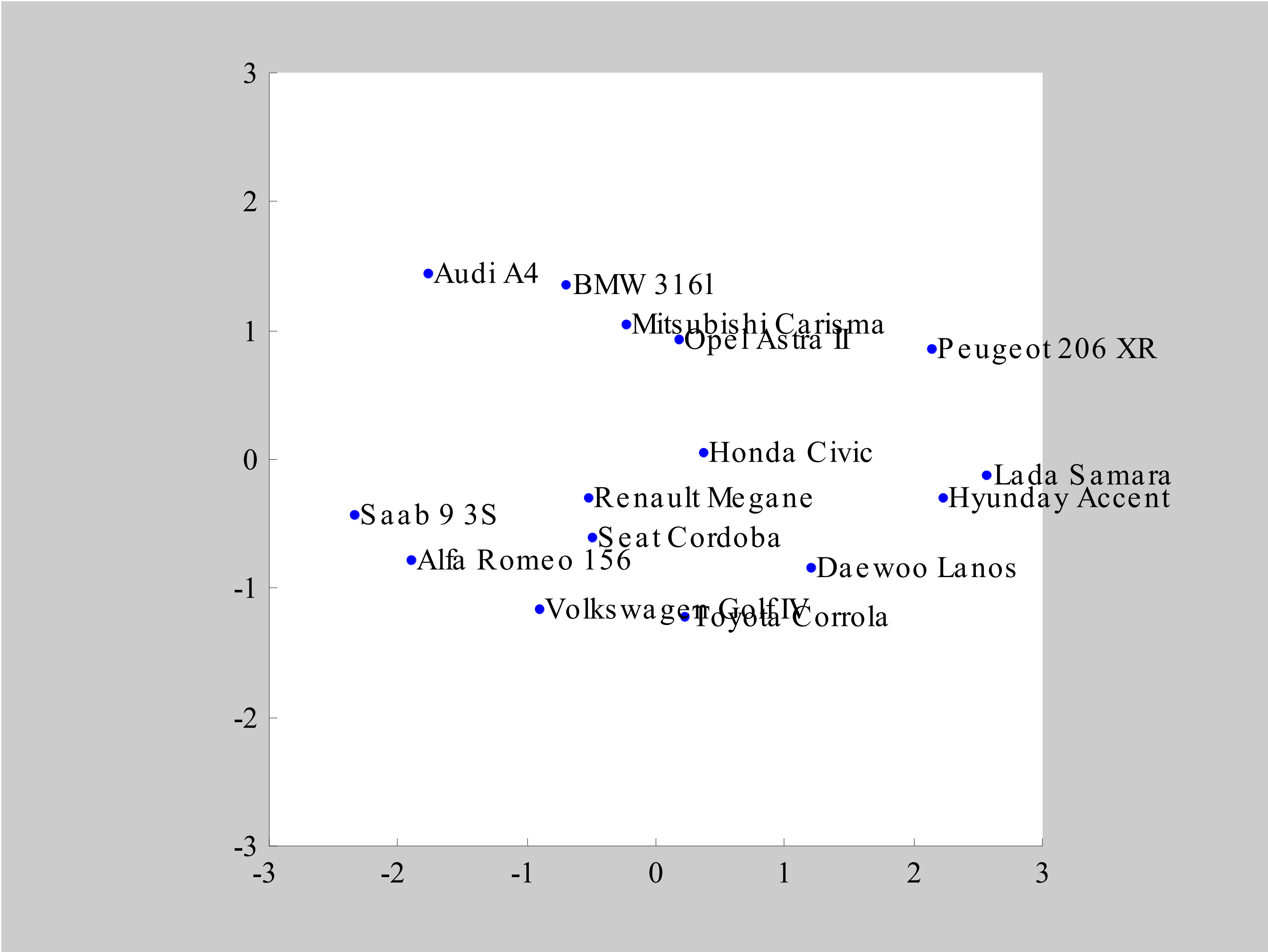
- Znormalizowane opisy 15 samochodów (tabela informacyjna)

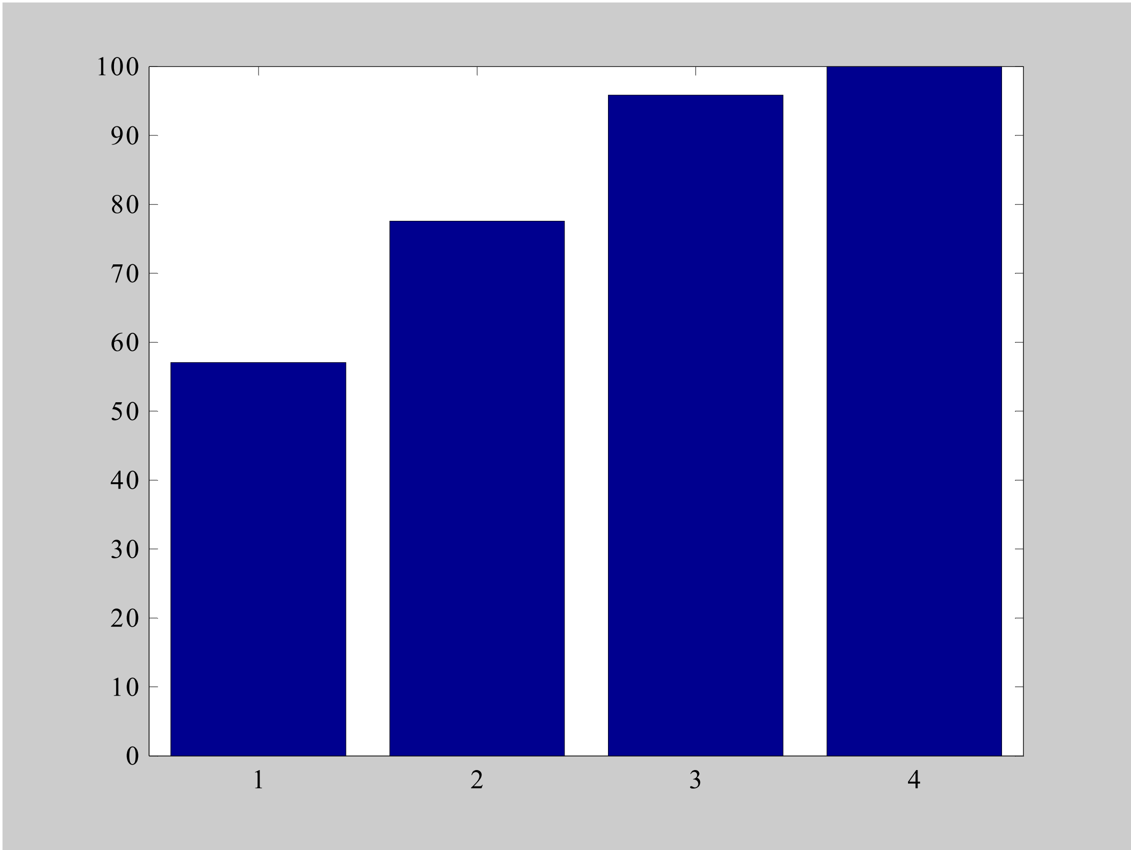
Marka	**Model**	**Zużycie**	**Przyspieszenie**	**Cena**	**Wyposażenie**
Alfa Romeo	156	0,57	-1,74	0,76	0,93
Audi	A4	0,34	0,46	1,87	1,31
BMW	316I	-0,11	0,80	1,29	0,56
Daewoo	Lanos	0,92	0,72	-1,05	-1,34
Honda	Civic	-1,03	-0,47	-0,35	0,18
Hyunday	Accent	-1,37	0,29	-1,03	-1,34
Lada	Samara	-0,46	1,40	-1,55	-1,72
Mitsubishi	Carisma	-0,46	0,55	0,23	0,93
Opel	Astra II	-0,69	0,55	0,05	0,56
Peugeot	206 XR	-1,14	1,57	-0,88	-0,96
Renault	Megane	-0,69	-1,23	-0,29	0,93
Saab	9 3S	2,40	-0,30	1,47	0,56
Seat	Cordoba	0,80	-0,39	-0,55	0,56
Toyota	Corrola	0,11	-0,98	-0,28	-0,96
Volkswagen	Golf IV	0,80	-1,23	0,29	-0,20

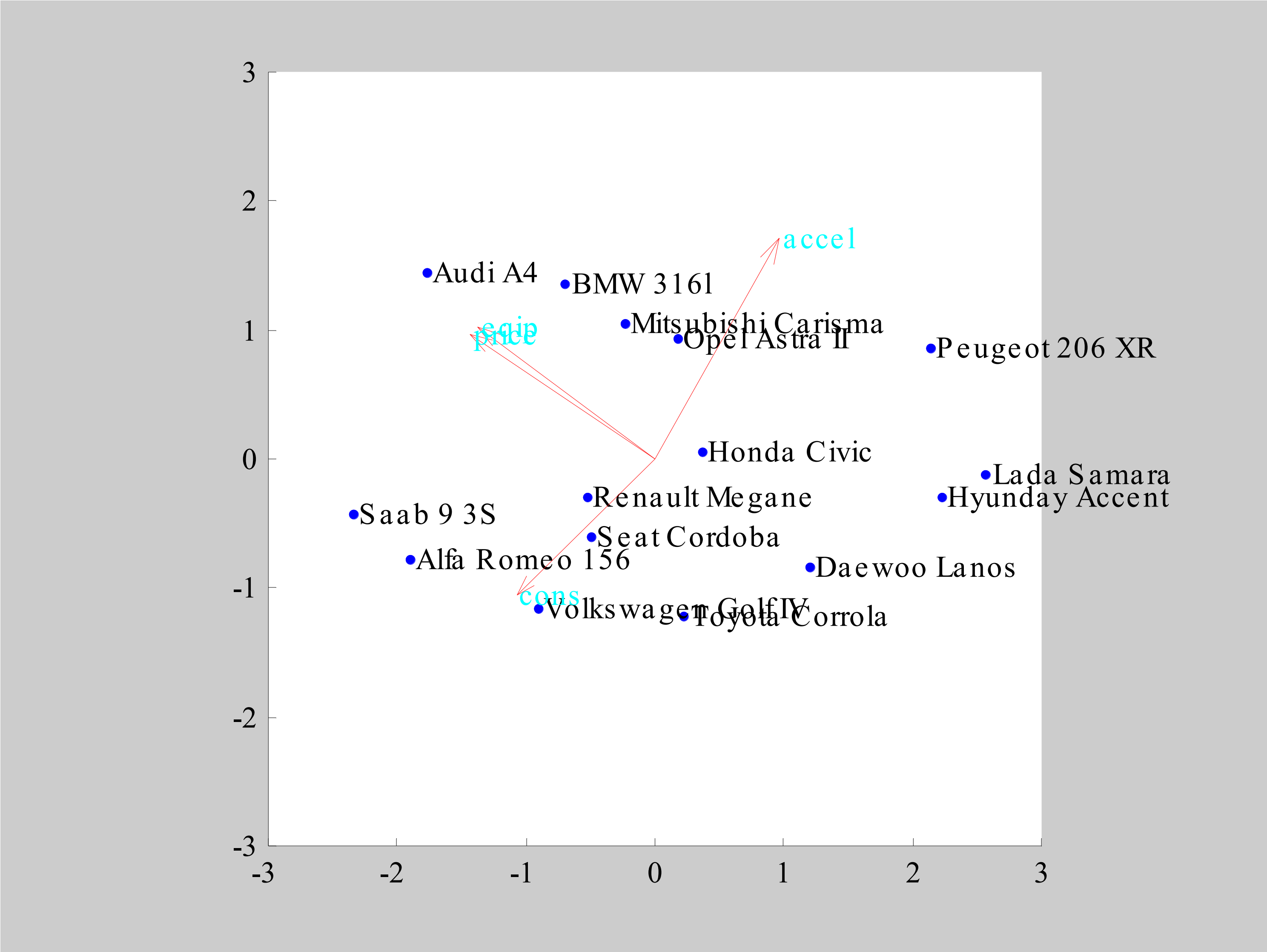
1.0000	-0.3245	0.4872	0.2619
-0.3245	1.0000	-0.2609	-0.3814
0.4872	-0.2609	1.0000	0.7691
0.2619	-0.3814	0.7691	1.0000


2.5
2
1.5
1
0.5
0
1
2
3
4

-1.9007	-0.7758
-1.7763	1.4540
-0.6978	1.3548
1.2019	-0.8323
0.3599	0.0557
2.2258	-0.2916
2.5666	-0.1197
-0.2338	1.0564
0.1759	0.9305
2.1353	0.8628
-0.5332	-0.2984
-2.3404	-0.4234
-0.4961	-0.5953
0.2245	-1.2162
-0.9117	-1.1617


Audi A4
BMW 3161
Mitsubishi Carisma
Opel Astra II
Peugeot 206 XR
Honda Civic
Lada Samara
Hyundai Accent
Renault Megane
Saab 9 3S
Seat Cordoba
Alfa Romeo 156
Daewoo Lanos
Volkswagen Golf IV
Toyota Corrola
3
2
1
0
-1
-2
-3


100
90
80
70
60
50
40
30
20
10
0
1
2
3
4


3
2
1
0
-1
-2
-3
-3
-2
-1
0
1
2
3
accel
Audi A4
BMW 3161
equip
price
Mitsubishi Carisma
Opel Astra II
Peugeot 206 XR
Honda Civic
Lada Samara
Hyunday Accent
Renault Megane
Saab 9 3S
Seat Cordoba
Alfa Romeo 156
Daewoo Lanos
cons
Volkswagen Golf IV
Toyota Corrola

…