...

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

...

Interpretacje geometryczne równań liniowych

- Geometryczna interpretacja rozwiązania układu równań
  - jako punkt wspólny hiperpłaszczyzn: tzw. przestrzeń wierszowa (inaczej: przestrzeń rozwiązań)
  - jako kombinacja wektorów: tzw. przestrzeń kolumnowa (inaczej: przestrzeń żądań)

Interpretacje geometryczne równań liniowych

- Dany jest układ oznaczony równań:

  $2x_1 – 1x_2 = 0$

  $2x_1 + 1x_2 = 4$

  którego rozwiązaniem jest wektor: $[1,2]^T$

  – dwa równania / dwie zmienne
- Jaka jest interpretacja geometryczna tego układu oraz jego rozwiązania?

Interpretacje geometryczne równań liniowych

- Aby przedstawić wymagane interpretacje, układ równań warto zapisać w postaci macierzowej: **Ax** = **b**,

gdzie: $\mathbf{A} = \begin{bmatrix} 2 & -1 \\ 2 & 1 \end{bmatrix}$ $\quad \mathbf{b} = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$

Interpretacje geometryczne równań liniowych

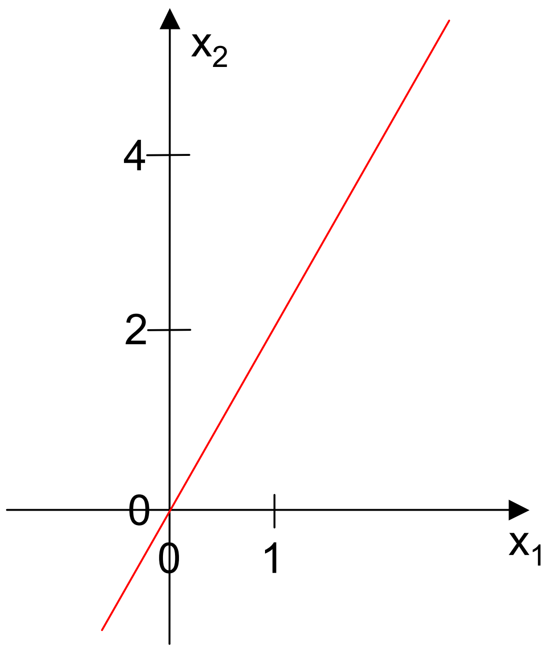
- Interpretacja rozwiązania w przestrzeni rozwiązań
  - postać macierzowa układu równań:

$$\begin{bmatrix} 2 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$$

- Po podziale macierzy $\mathbf{A} = [\mathbf{w}_1; \mathbf{w}_2]$ na wiersze interpretujemy je jako równania opisujące proste (prawe strony równań powstają z podzielonego na wiersze wektora $\mathbf{b} = [b_1; b_2]$)
  - równanie: $(\mathbf{w}_1)^T\mathbf{x} = b_1$, czyli $2x_1 - 1x_2 = 0$
  - równanie: $(\mathbf{w}_2)^T\mathbf{x} = b_2$, czyli $2x_1 + 1x_2 = 4$
    - równania reprezentują hiperpłaszczyzny (które w tym przypadku mają postać prostych)

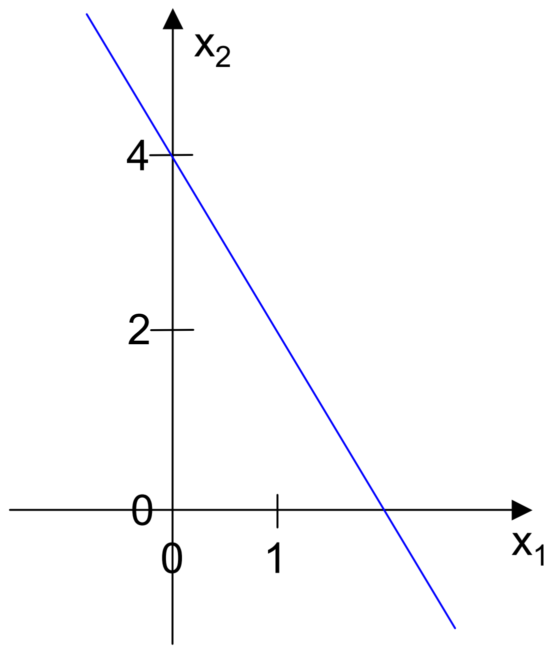
Interpretacje geometryczne równań liniowych

- Interpretacja geometryczna równania: $2x_1 – 1x_2 = 0$
  - wektory spełniające to równanie są elementami prostej P1



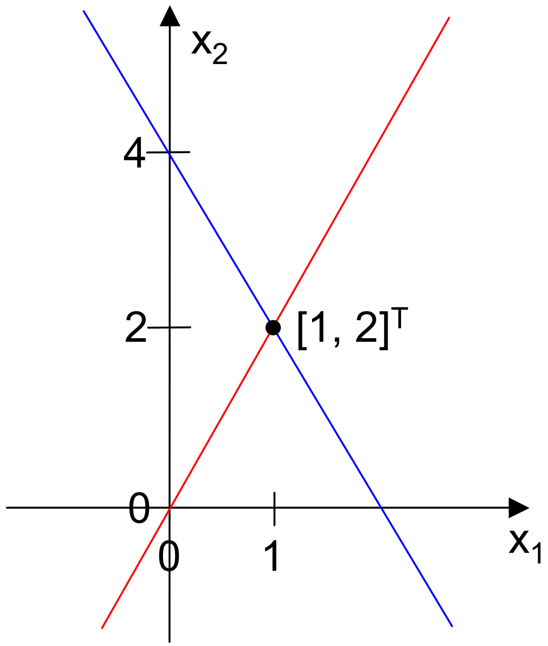
Interpretacje geometryczne równań liniowych

- Interpretacja geometryczna równania: $2x_1+1x_2 = 4$
  - wektory spełniające to równanie są elementami prostej P2



Interpretacje geometryczne równań liniowych

- Interpretacja geometryczna rozwiązania:
  - rozwiązaniem są wektory będące elementami obu prostych (w tym przypadku jest to tylko jeden wektor: $[1, 2]^T$)



Interpretacje geometryczne równań liniowych

- Interpretacja rozwiązania w przestrzeni żądań
  - postać macierzowa układu równań:

$$\begin{bmatrix} 2 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$$

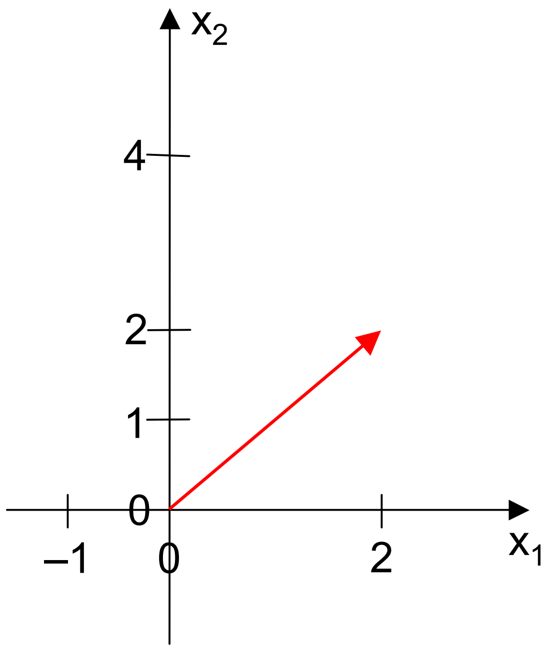
- Po podziale macierzy **A** = [$\mathbf{k}_1$, $\mathbf{k}_2$] na kolumny interpretujemy je jako elementy kombinacji liniowej, której współczynnikami są zmienne x_j (a wynikiem – wektor **b**)

$$x_1 \begin{bmatrix} 2 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$$

  - w kombinacji występują wektory (które w tym przypadku mają postać punktów na płaszczyźnie)

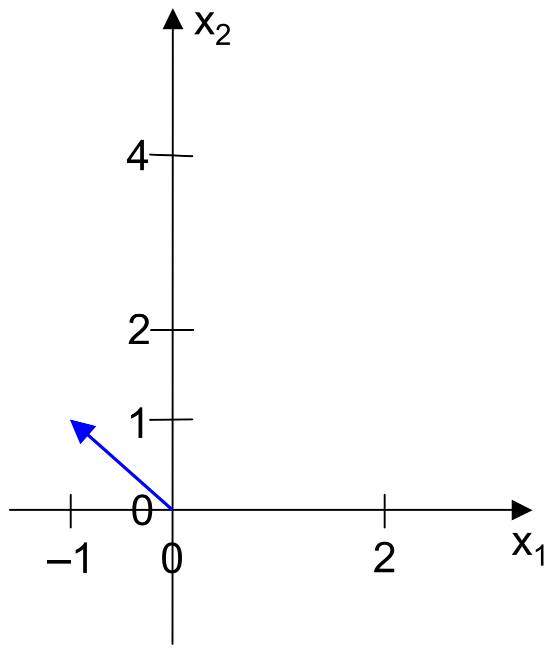
Interpretacje geometryczne równań liniowych

- Interpretacja geometryczna wektora $\mathbf{k}_1$: $[2, 2]^T$



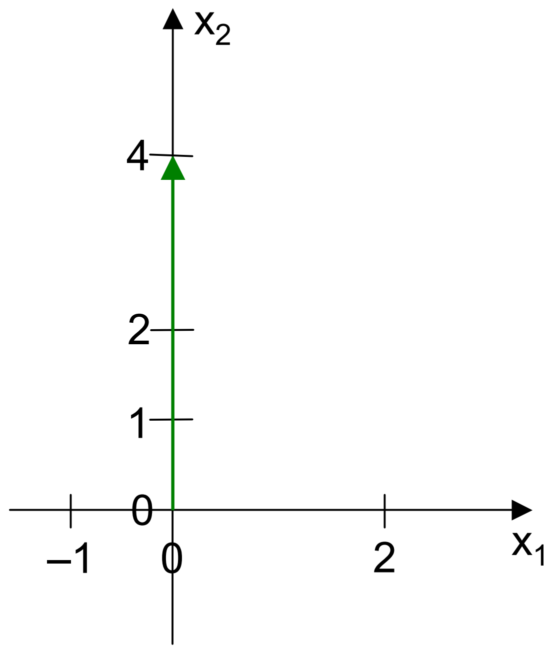
Interpretacje geometryczne równań liniowych

- Interpretacja geometryczna wektora $\mathbf{k}_2$: $[-1, 1]^T$



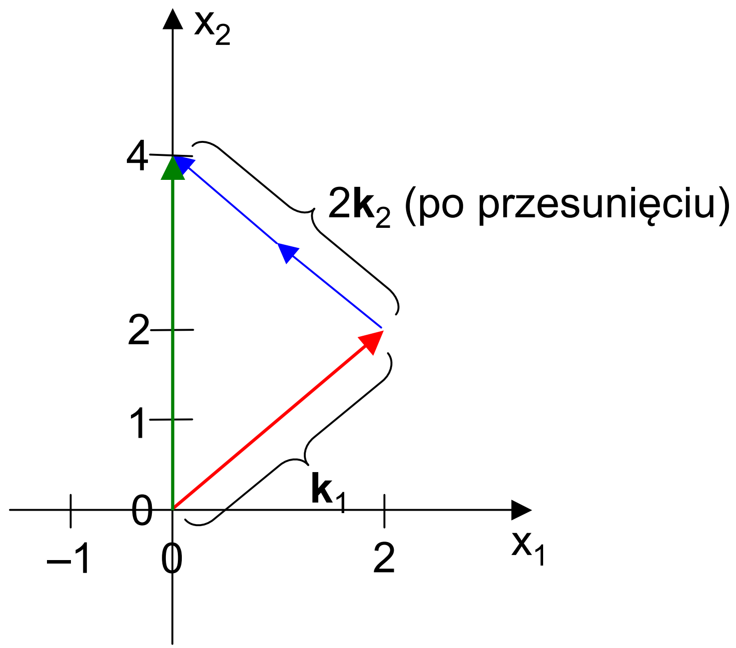
Interpretacje geometryczne równań liniowych

- Interpretacja geometryczna wektora **b**: $[0, 4]^T$



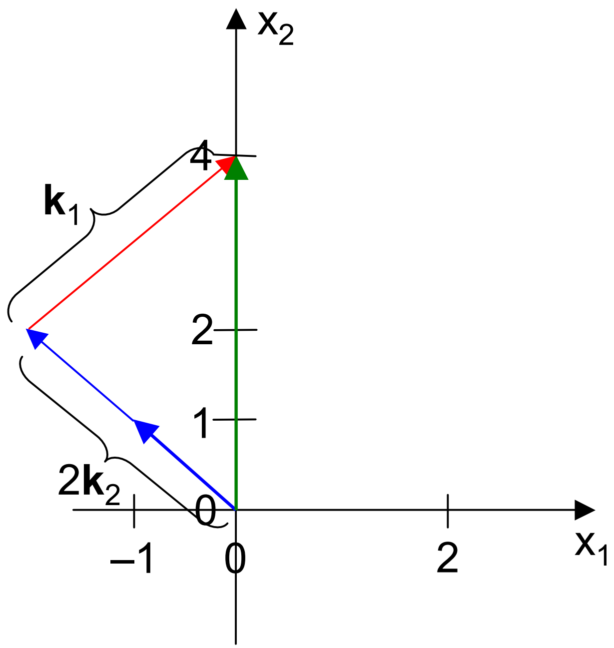
Interpretacje geometryczne równań liniowych

- Interpretacja geometryczna rozwiązania:
  – rozwiązaniem jest kombinacja liniowa o współczynnikach $x_1 = 1$ oraz $x_2 = 2$ (ponieważ wektor **b** można wyrazić jako $1\mathbf{k}_1 + 2\mathbf{k}_2$)



Interpretacje geometryczne równań liniowych

- Interpretacja geometryczna rozwiązania c.d.:
  - oczywiście także $\mathbf{b} = 2\mathbf{k}_2 + 1\mathbf{k}_1$ (w obu przypadkach wektor $\mathbf{k}_1$ jest mnożony przez współczynnik 1 a wektor $\mathbf{k}_2$ przez współczynnik 2)



Interpretacje geometryczne równań liniowych

- Dalszy przykład: sprzeczny układ równań:

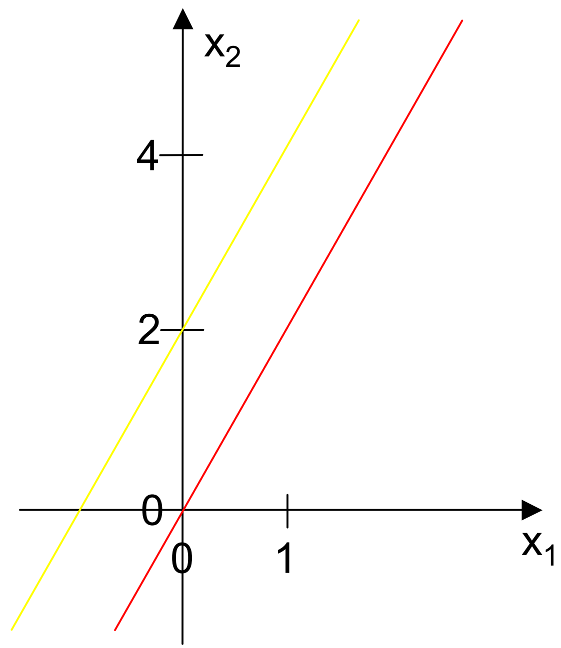
  $2x_1 – 1x_2 = 0$

  $2x_1 – 1x_2 = 2$

  (układ ten nie ma rozwiązania)
  - dwa równania / dwie zmienne
- Jaka jest interpretacja geometryczna braku rozwiązania tego układu w przestrzeni rozwiązań/żądań?

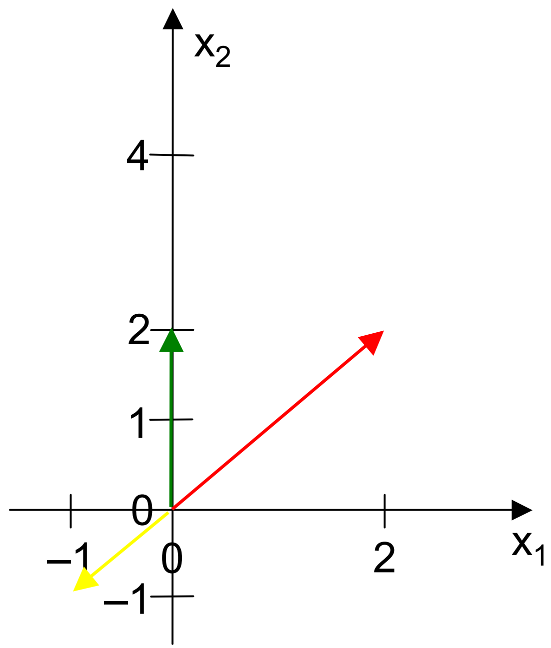
Interpretacje geometryczne równań liniowych

- Interpretacja w przestrzeni rozwiązań:
  - brak punktu, w którym proste P1 i P2 przecinałyby się (co wynika z faktu, że proste te są równoległe)



Interpretacje geometryczne równań liniowych

- Interpretacja w przestrzeni żądań:
  – brak kombinacji wektorów **k_1** oraz **k_2**, której wynikiem byłby **b** (co wynika z faktu, że **k_1** i **k_2** są zależne, a **b** od nich niezależny)



Interpretacje geometryczne równań liniowych

- Dalszy przykład: nieoznaczony układ równań:

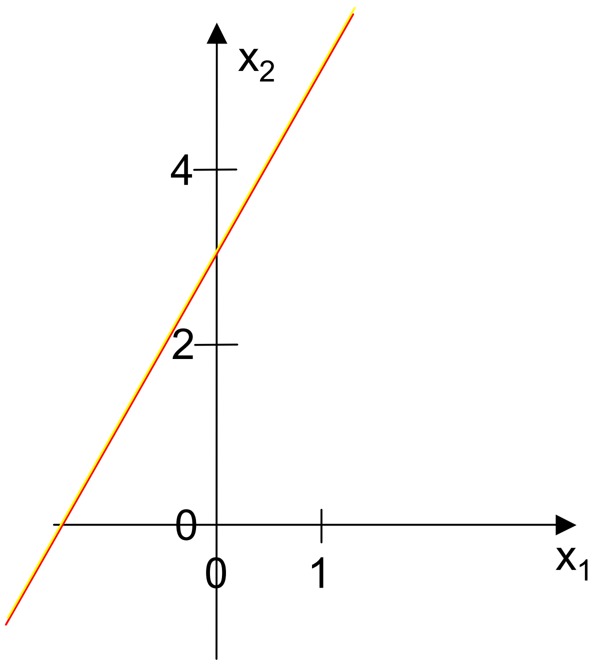
  $2x_1 – 1x_2 = 3$

  $2x_1 – 1x_2 = 3$

  (układ ten ma nieskończenie wiele rozwiązań)
  - dwa równania / dwie zmienne
- Jaka jest interpretacja geometryczna nieskończonej liczby rozwiązań tego układu w przestrzeni rozwiązań/żądań?

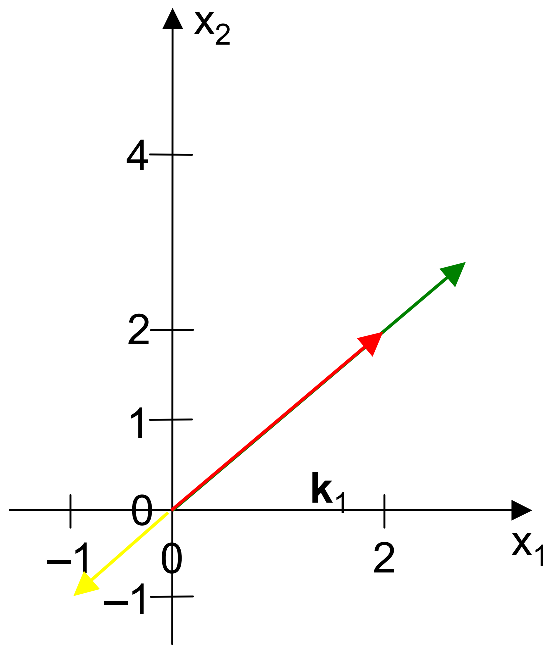
Interpretacje geometryczne równań liniowych

- Interpretacja w przestrzeni rozwiązań:
  - proste P1 i P2 pokrywają się (co wynika z faktu, że każda z nich jest opisana tym samym równaniem)



Interpretacje geometryczne równań liniowych

- Interpretacja w przestrzeni żądań:
  - wiele różnych kombinacji wektorów $\mathbf{k}_1$ i $\mathbf{k}_2$ daje w wyniku wektor $\mathbf{b}$ (co wynika z faktu, że $\mathbf{k}_1$, $\mathbf{k}_2$ i $\mathbf{b}$ są zależne)



Interpretacje geometryczne równań liniowych

- Charakterystyka interpretacji geometrycznej układu równań $\mathbf{A}_{mxn}\mathbf{x}_{nx1} = \mathbf{b}_{mx1}$ w przestrzeni rozwiązań
  - macierz **A** oraz wektor **b** są dzielone na wiersze
    - każdy wiersz macierzy **A** z odpowiednim elementem wektora **b** tworzy równanie jednej hiperpłaszczyzny
  - rozwiązaniem układu równań jest punkt wspólny wszystkich hiperpłaszczyzn
  - liczba
    - hiperpłaszczyzn: m
    - wymiarów przestrzeni, w której opisane są hiperpłaszczyzny: n

Interpretacje geometryczne równań liniowych

- Układ sprzeczny, oznaczony i nieoznaczony w przestrzeni rozwiązań
  – układ sprzeczny (0 rozwiązań)
    - nie istnieje punkt wspólny wszystkich hiperpłaszczyzn
    - pewne hiperpłaszczyzny się nie przecinają (są równoległe) lub przecięcia różnych hiperpłaszczyzn zachodzą w różnych punktach
  – układ oznaczony (1 rozwiązanie)
    - istnieje punkt wspólny wszystkich hiperpłaszczyzn
    - wszystkie hiperpłaszczyzny przecinają się w jednym punkcie
  – układ nieoznaczony ($\propto$ rozwiązań):
    - istnieje nieskończenie wiele punktów wspólnych wszystkich hiperpłaszczyzn
    - wszystkie hiperpłaszczyzny się pokrywają

Interpretacje geometryczne równań liniowych

- Charakterystyka interpretacji geometrycznej układu równań $\mathbf{A}_{mxn}\mathbf{x}_{nx1} = \mathbf{b}_{mx1}$ w przestrzeni żądań
  - macierz **A** jest dzielona na kolumny, wektor wektor **b** stanowi dodatkowa kolumnę
    - kolumny macierzy **A** stają się elementami kombinacji liniowej
  - rozwiązanie układu polega na znalezieniu współczynników kombinacji liniowej
  - liczba
    - wektorów po lewej stronie kombinacji liniowej: n
    - wymiarów przestrzeni, w której opisane są wektory: m

Interpretacje geometryczne równań liniowych

- Układ sprzeczny, oznaczony i nieoznaczony w przestrzeni żądań
  - układ sprzeczny (0 rozwiązań)
    - nie istnieją współczynniki, które utworzyłyby wektor **b** jako kombinację liniową danych wektorów (kolumn macierzy **A**)
    - wektor **b** nie leży w przestrzeni rozpinanej przez kolumny macierzy **A** (**b** jest niezależny od kolumn macierzy **A**)
  - układ oznaczony (1 rozwiązanie)
    - istnieje jednoznaczna kombinacja współczynników wyrażających wektor **b** jako kombinację liniową kolumn macierzy **A**
    - kolumny macierzy **A** są niezależne ale wektor **b** jest od nich zależny
  - układ nieoznaczony ($\propto$ rozwiązań):
    - istnieją różne jednoznaczne kombinacje współczynników wyrażających wektor **b** jako kombinację liniową kolumn macierzy **A**
    - kolumny macierzy **A** są zależne i wektor **b** jest od nich zależny

…

Układy współrzędnych

- W kontekście układów współrzędnych można dopatrywać się zasadniczo dwojakiego zastosowania wektorów:
  – wektory jako wersory (wektory jednostkowe) układów współrzędnych
  – wektory jako zbiory współczynników kombinacji liniowych
- Podstawowe operacje w tym kontekście
  – ustalanie właściwości przekształceń danych z jednego układu współrzędnych do innego

Układy współrzędnych

- Umowny system normalizacyjny pozwalający na dokonywanie pomiarów
  - układy opisujące różne cechy ciał
  - układy opisujące cechy przestrzeni
- Cechy możliwych przestrzeni:
  - przestrzenie płaskie (o krzywiźnie zerowej)
  - przestrzenie elipsoidalne (o krzywiźnie dodatniej)
  - przestrzenie hiperboliczne (o krzywiźnie ujemnej)

Układy współrzędnych

- Rodzaje układów współrzędnych
  - przyrodnicze
    - geograficzne
    - elipsoidalne
  - matematyczne
    - „nieliniowe”
      - biegunowe
      - sferyczne
      - walcowe
    - „liniowe”
      - kartezjańskie
      - niekartezjańskie
      - barycentryczne

Układy współrzędnych

- Pojęcie i definicja bazy w układach kartezjańskich
  - przestrzeń
  - wektory rozpinające
  - baza przestrzeni

Macierz wersorów

- W kartezjańskim układzie współrzędnych wersory charakteryzują się:
  - wzajemną prostopadłością (ortogonalnością)
  - równą długością (unormowaniem) – standardowo 1
- Role wersorów układu kartezjańskiego może odgrywać taki zbiór wektorów, który spełnia tzw. warunek ortonormalności:
  - zbiór wektorów $\{\mathbf{k}_1, \mathbf{k}_2, ..., \mathbf{k}_N\}$ jest ortonormalny, jeżeli:
    - każdy z wektorów jest unormowany: $\langle \mathbf{k}_i, \mathbf{k}_i \rangle = 1$
    - każda para (różnych) wektorów jest ortogonalna: $\langle \mathbf{k}_i, \mathbf{k}_j \rangle = 0$ dla $i \neq j$
- Właściwości:
  - niezerowe wektory ortogonalne są niezależne
  - wektory ortonormalne są niezależne

Macierz wersorów

- Wektory jako macierz
  - ortonormalnośc zbioru n wektorów n-elementowych $\mathbf{k}_i$ (i = 1..n) można sprawdzić tworząc macierz $\mathbf{K} = [\mathbf{k}_1, \mathbf{k}_2, ..., \mathbf{k}_n]$ i obliczając iloczyn $\mathbf{K}^T\mathbf{K}$
    - ponieważ wynik mnożenia macierzy **AB** jest tak naprawdę złożeniem iloczynów skalarnych wektorów wierszowych z macierzy **A** i wektorów kolumnowych z macierzy **B**, a więc $\mathbf{A}^T\mathbf{B}$ jest złożeniem iloczynów skalarnych wektorów kolumnowych z macierzy **A** i wektorów kolumnowych z macierzy **B**
  - jeżeli $\mathbf{K}^T\mathbf{K} = \mathbf{I}$ to wyniki przemnożenia przez siebie tych samych wektorów kolumnowych są równe 1 a wyniki przemnożenia przez siebie różnych wektorów są równe 0
    - wniosek: gdy dla macierzy $\mathbf{K} = [\mathbf{k}_1, \mathbf{k}_2, ..., \mathbf{k}_N]$ zachodzi $\mathbf{K}^T\mathbf{K} = \mathbf{I}$ to zbiór wektorów $[\mathbf{k}_1, \mathbf{k}_2, ..., \mathbf{k}_N]$ jest ortonormalny

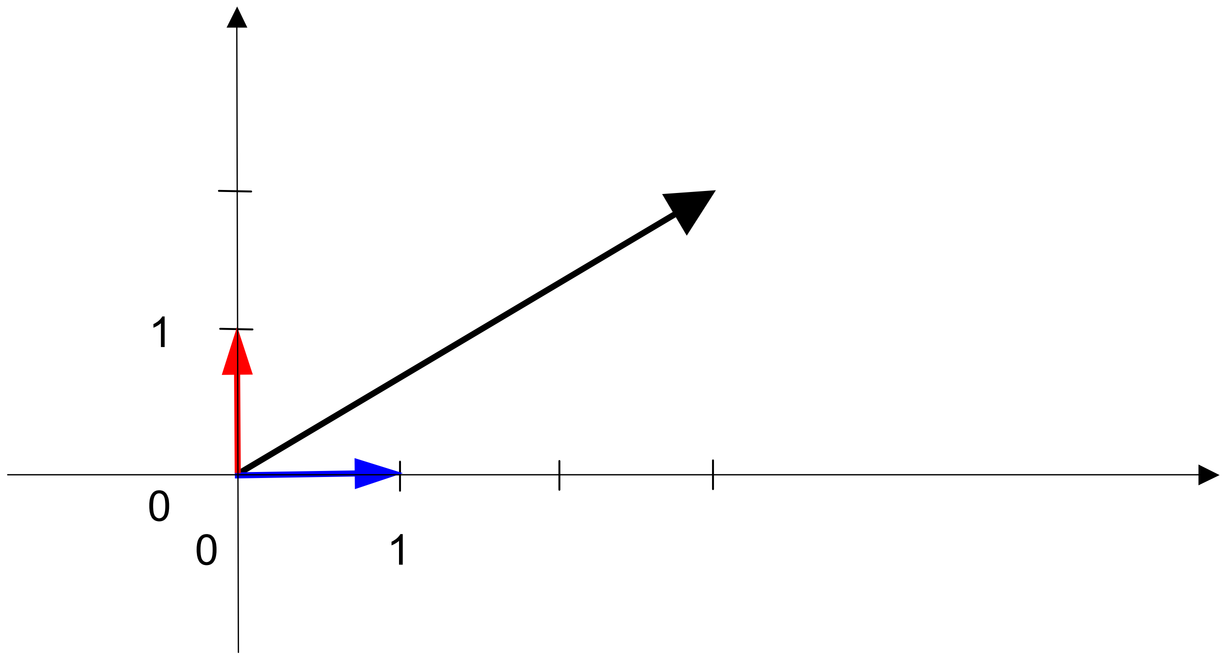
Macierz wersorów

- Macierz, której kolumny tworzą zbiór wektorów ortonormalnych nazywa się macierzą ortogonalną
  - macierz ortogonalna spełnia więc równanie: $\mathbf{K}^T\mathbf{K} = \mathbf{I}$
  - kolumny takiej macierzy mogą odgrywać rolę wersorów pewnego kartezjańskiego układu współrzędnych
    - macierz wersorów

Przekształcenia współrzędnych

- Jaki problem algebraiczny stawiamy i rozwiązujemy?

Przekształcenia współrzędnych

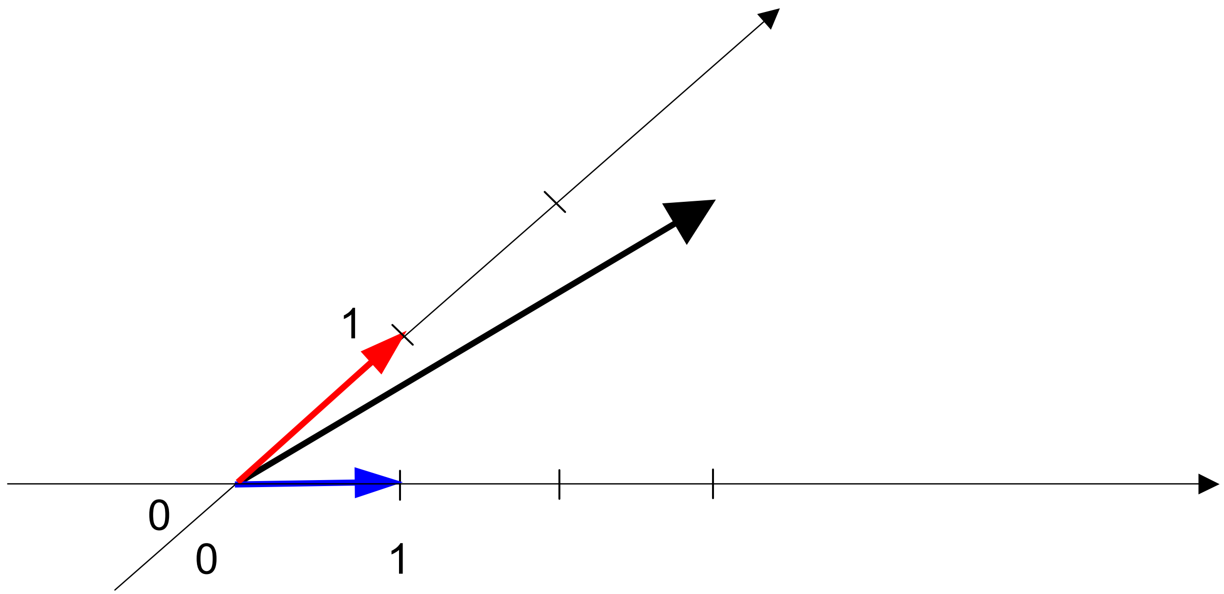


$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Przekształcenia współrzędnych



$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Przekształcenia współrzędnych

- Dane są:
  macierz $\mathbf{A}_{mxn}$, wektor $\mathbf{x}_{nx1}$ i wektor $\mathbf{b}_{mx1}$ spełniające: $\mathbf{Ax} = \mathbf{b}$
  – jakie sa implikacje geometryczne tego równania (układu równań)?

Przekształcenia współrzędnych

- Problem: przekształcenie punktu do innego układu współrzędnych (układy liniowe)
  - dany jest punkt o współrzędnych (y_1, y_2) w układzie współrzędnych o środku w punkcie (0,0) i wersorach $\mathbf{e}_1 = [1, 0]^T$, $\mathbf{e}_2 = [0, 1]^T$
    - punkt (y_1, y_2) jest oczywiście dalej traktowany jako wektor $[y_1,y_2]^T$
  - jakie będą współrzędne tego punktu w innym układzie, o środku w punkcie (0,0) i wersorach $\mathbf{a}_1 = [1, 1]^T$, $\mathbf{a}_2 = [0, 2]^T$?
- Wykorzystywane założenia:
  - wersory są wektorami niezależnymi, z czego wynika, że:
    - każdy punkt ma jednoznaczne współrzędne (jest jednoznacznie wyznaczony poprzez swoje współrzędne względem wersorów)
    - macierz wersorów jest nieosobliwa (a tym samym posiada odwrotność)

Przekształcenia współrzędnych

- Jaki jest związek współrzędnych punktu z wersorami odpowiedniego układu współrzędnych?
  - fakt, że wektor $\mathbf{y} = [y_1, y_2]^T$ jest wyrażony w układzie współrzędnych o wersorach $\mathbf{e}_1 = [1, 0]^T$, $\mathbf{e}_2 = [0, 1]^T$ jest równoważny z faktem, że może on być wyrażony jako kombinacja liniowa wersorów:
    - współczynniki kombinacji: x_1 i x_2

      $\mathbf{y} = x_1 \cdot \mathbf{e}_1 + x_2 \cdot \mathbf{e}_2$
    - x_1 i x_2 można zapisać jako wektor $\mathbf{x}$

      $\mathbf{y} = [\mathbf{e}_1, \mathbf{e}_2][x_1, x_2]^T$

      $\mathbf{y} = \mathbf{I}\mathbf{x}$
  - współczynniki kombinacji uzyskujemy rozwiązując powyższy układ równań

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

Przekształcenia współrzędnych

- W ogólnym przypadku:
  - dla innych (liniowych) układów współrzędnych, zdefiniowanych z użyciem innych wersorów, np.:
    - $\mathbf{a}_1 = [1, 1]^T$ i $\mathbf{a}_2 = [0, 2]^T$
  - należy rozwiązać układ:
    - współczynniki kombinacji liniowej: $\mathbf{w} = [w_1, w_2]^T$

      $[\mathbf{a}_1, \mathbf{a}_2] \cdot \mathbf{w} = \mathbf{y}$ (gdzie $\mathbf{y} = [\mathbf{e}_1, \mathbf{e}_2][x_1, x_2]^T$)

      czyli macierzowo:

      $\mathbf{Aw} = \mathbf{y}$
    - rozwiązaniem jest:

      $\mathbf{w} = \mathbf{A}^{-1}\mathbf{y}$

Przekształcenia współrzędnych

- Dla przykładowych danych:

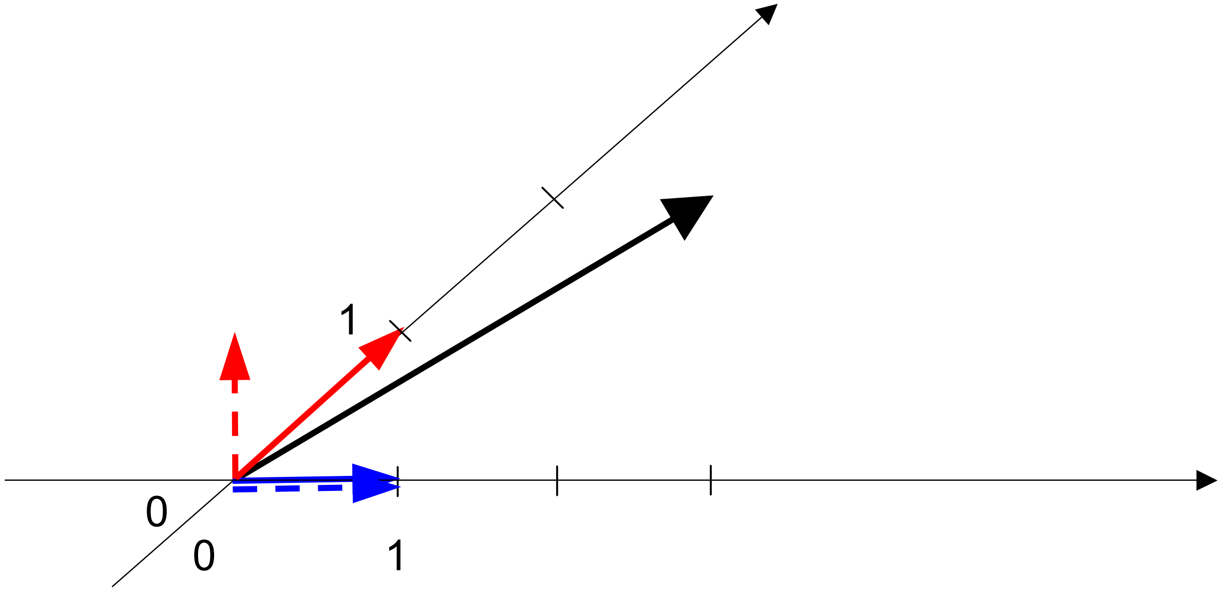
  $\mathbf{A} = [\ 1\ ,\ 0\ ;\ 1\ ,\ 2\ ]$

  $\mathbf{A}^{-1} = [\ 1\ ,\ 0\ ;\ -0.5\ ,\ 0.5\ ]$

  $\mathbf{w} = \mathbf{A}^{-1}\cdot\mathbf{y} = [\ 1\ ,\ 0\ ;\ -0.5\ ,\ 0.5\ ]\cdot\mathbf{y} = [\ y_1\ ,\ (y_2-y_1)/2\ ]^T$

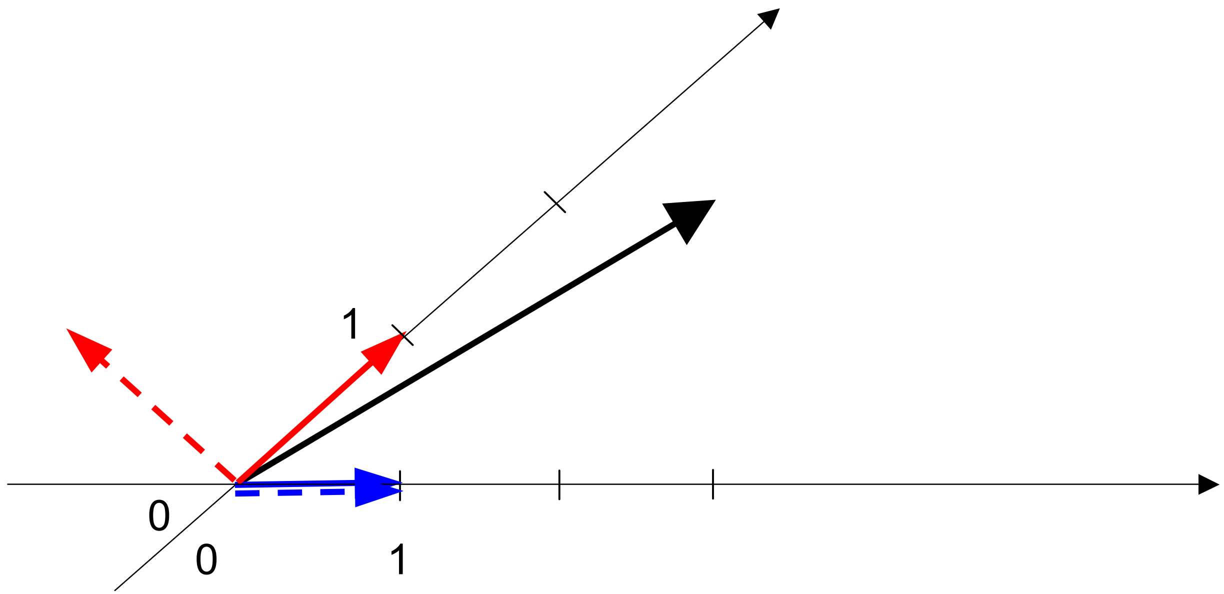
  – ostatecznie:
    - punkt, którego
      współrzędnymi względem wersorów **I** były wartości (y_1,y_2)
      względem wersorów **A** będzie miał współrzędne $(y_1\ ,\ (y_2-y_1)/2)$

Przekształcenia współrzędnych



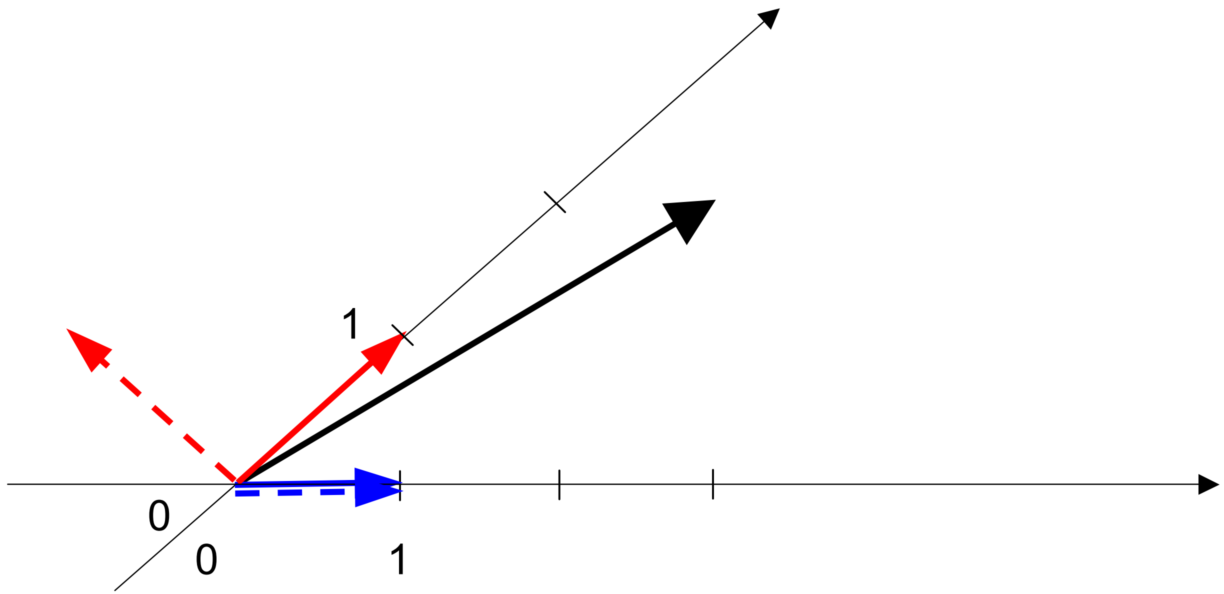
$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Przekształcenia współrzędnych



$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} ? \\ ? \end{bmatrix}$$

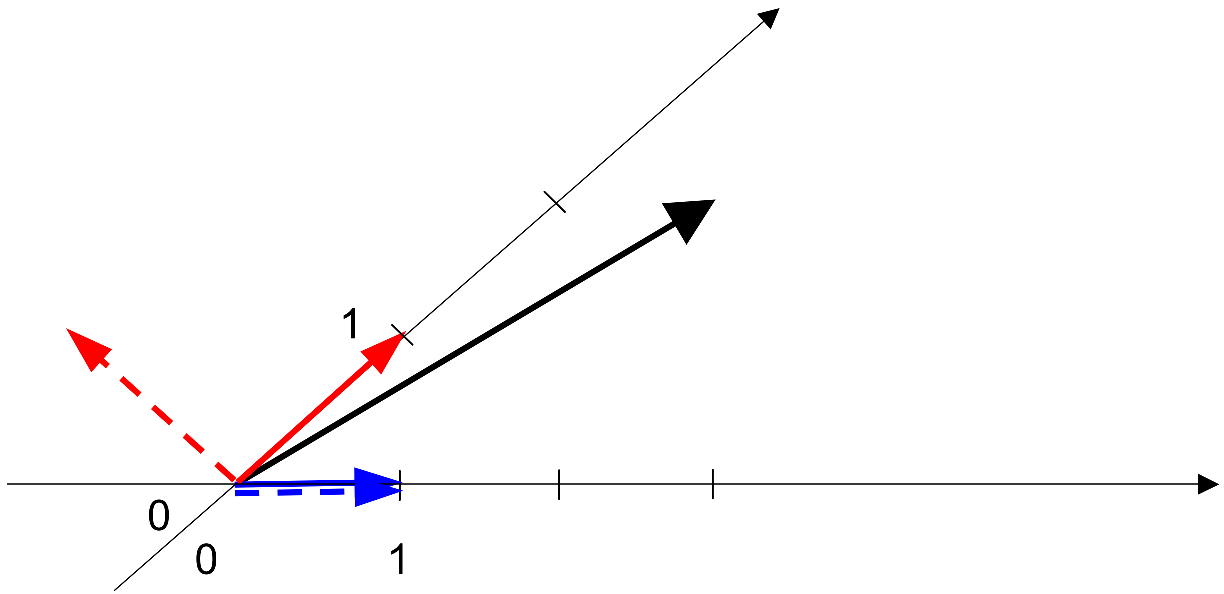
Przekształcenia współrzędnych



$$\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$$

Przekształcenia współrzędnych



$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \end{bmatrix}$$

…

Analiza nienadzorowana

Analiza nienadzorowana

- Ocena zmiennych/atrybutów pod względem przydatności do analiz
  - (szeroko rozumianych) kosztów pozyskania
  - (szeroko rozumianej) ilości zawartej w nich informacji
- Podstawowy podział metod oceny
  - ocena indywidualna (pojedyncze atrybuty)
  - ocena zbiorowa (zbiory atrybutów)

Analiza nienadzorowana

- Ocena indywidualna
  - jeżeli zmienne (atrybuty) danych są podzielone na warunkowe i decyzyjne (analiza nadzorowana), to ilość informacji „zawartej” w pewnym atrybucie warunkowym można wyrażać poprzez miary zależności pomiędzy tym atrybutem a atrybutem decyzyjnym (entropia)
  - w sytuacji gdy w zbiorze nie zdefiniowano atrybutów decyzyjnych, atrybuty muszą być oceniane inaczej (czyli bez udziału atrybutów decyzyjnych)
    - w tym przypadku za miarę informacyjności przyjmuje się wariancję atrybutu, co można wytłumaczyć tym, że atrybut nie wykazujący żadnej zmienności (wariancja 0) nie niesie (prawie) żadnej informacji
      - dokładniej, niesie pewną ilość informacji, ale nie niesie żadnej informacji nadającej się do różnicowania obiektów (odróżniania ich od siebie)

Analiza nienadzorowana

- Przykład: opis
  - 15 samochodów (obiekty)
    w kategoriach
  - 4 parametrów (zmienne/atrybuty)

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

Analiza nienadzorowana

- Przykład: opis...

Marka / Model	**Zużycie**	**Przyspieszenie**	**Cena**	**Liczba kół**
Alfa Romeo 156	8,1	9,3	71,14	4
Audi A4	7,9	11,9	93,35	4
BMW 316I	7,5	12,3	81,79	4
Daewoo Lanos	8,4	12,2	34,90	4
Honda Civic	6,7	10,8	48,90	4
Hyunday Accent	6,4	11,7	35,30	4
Lada Samara	7,2	13,0	24,90	4
Mitsubishi Carisma	7,2	12,0	60,60	4
Opel Astra II	7,0	12,0	56,95	4
Peugeot 206 XR	6,6	13,2	38,36	4
Renault Megane	7,0	9,9	50,05	4
Saab 9 3S	9,7	11,0	85,35	4
Seat Cordoba	8,3	10,9	44,99	4
Toyota Corrola	7,7	10,2	50,36	4
Volkswagen Golf IV	8,3	9,9	61,62	4
	-------------	-------------	-------------	-------------
średnia	7,6	11,4	55,9	4,0
wariancja	0,8	1,4	399,7	0,0
odch. stand.	0,9	1,2	20,0	0,0

Analiza nienadzorowana

- Problem mocno zróżnicowanych jednostek
  - rozwiązanie: „wyrównywanie” jednostek (tam, gdzie to możliwe)
    - w praktyce: normalizacja zmiennych (w różnych postaciach)

Analiza nienadzorowana

- Normalizowanie zmiennych (kolumn macierzy danych)
  - operacja mająca na celu ujednolicenie jednostek
  - jej wykonanie realizuje przyjęte założenie o równym traktowaniu wszystkich zmiennych

Analiza nienadzorowana

- Jeżeli dane oryginalne są zapisane jako elementy x_{ij} macierzy **X** (o rozmiarach MxN), to przez dane znormalizowane z_{ij} rozumie się:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}$$

  gdzie:
  - $\bar{x}_j$ jest średnią arytmetyczną zmiennej $\mathbf{x}_j$ (czyli kolumny j macierzy **X**)
  - s_j jest odchyleniem standardowym zmiennej $\mathbf{x}_j$ (czyli kolumny j macierzy **X**)
  - operacja ta może być przedstawiona w zapisie macierzowym jako:

$$\mathbf{Z} = \mathbf{JXS}^{-1}$$

  - gdzie:
    - $\mathbf{J} = (\mathbf{I} - \mathbf{E}/m)$ jest macierzą skalującą
      - m (skalar) jest liczbą wierszy macierzy **X**
    - **S** jest macierzą diagonalną wartości s_j (odchyleń standardowych)
  - normalizacja jest w praktyce złożeniem centrowania i skalowania

Analiza nienadzorowana

- Cechy danych znormalizowanych
  - zero odpowiada średniej arytmetycznej danej zmiennej (kolumny)
    - wartość zero oznacza wartość średnią
    - wartość dodatnia oznacza wartość większą od średniej
    - wartość ujemna oznacza wartość mniejszą od średniej
  - jedynka odpowiada odchyleniu standardowemu danej zmiennej (kolumny)
    - wartość plus jeden oznacza wartość większą średniej o jedno odchylenie standardowe
    - wartość jeden oznacza wartość większą średniej o jedno odchylenie standardowe

Analiza nienadzorowana

- Przykład: opis... (zmienne znormalizowane)

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

…

Przekształcenie PCA

PCA

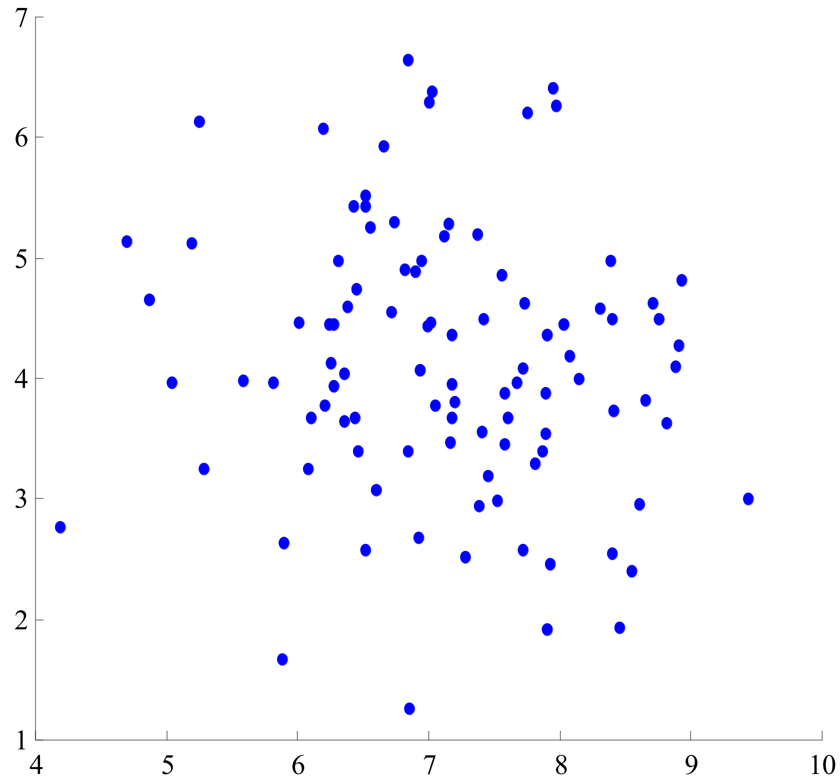
- PCA: Principal Components Analysis
  (ang. analiza składowych głównych)
- PCA działa w oparciu o macierz kowariancji zmiennych
  (ewentualnie: macierz korelacji zmiennych)

Macierz kowariancji

- Niech $\mathbf{S}_X = \mathbf{X}^T\mathbf{X}$
  - jeżeli **X** jest macierzą wycentrowaną, to
    - $(1/(m-1))\cdot\mathbf{S}_X$ jest macierzą kowariancji (z próby)
    - $(1/m)\cdot\mathbf{S}_X$ jest macierzą kowariancji (z populacji)
  - jeżeli **X** jest macierzą znormalizowaną, to
    - $(1/(m-1))\cdot\mathbf{S}_X$ jest macierzą korelacji
- Uwaga
  - mnożenie przez skalary typu $(1/(m-1))$ czy $(1/m)$ może być w dalszych rozważaniach pominięte (ponieważ odpowiada jednakowemu traktowaniu wszystkich kolumn macierzy **X**)

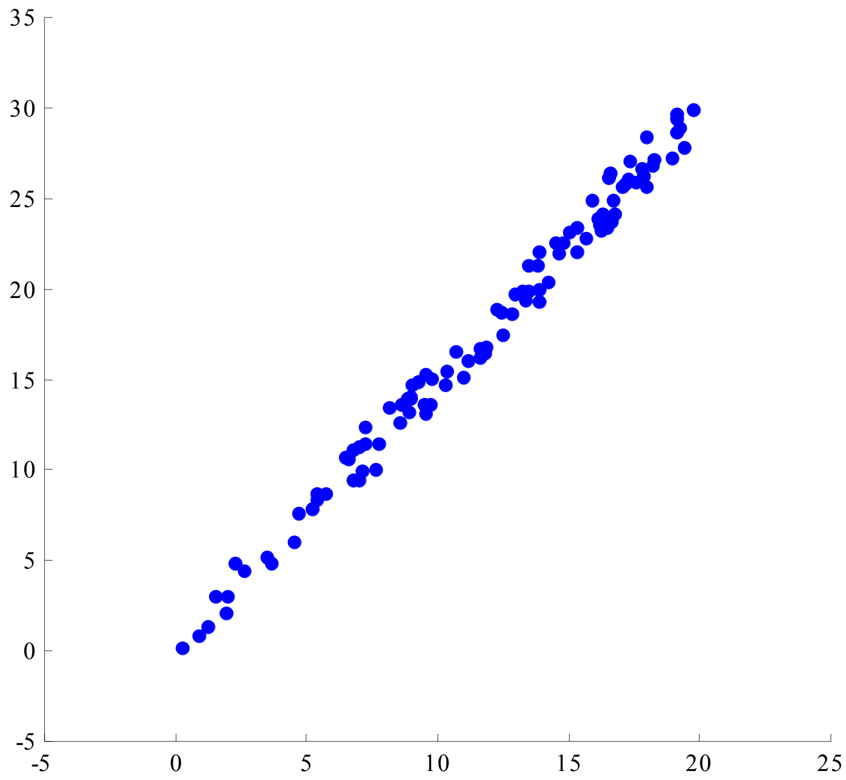
Przykład: PCA dla danych dwuwymiarowych

$\mathbf{x}_1$	$\mathbf{x}_2$
7.8669	3.3972
9.4316	3.0066
7.1102	5.1889
7.0264	6.3880
7.9703	6.2655
6.9947	6.3011
8.4095	3.7299
8.7579	4.5028
7.8850	3.8808
8.1409	3.9981
7.4032	3.5674
7.1910	3.8052
6.3064	4.9854
7.0110	4.4686
5.8944	2.6351
...	...
...	...
...	...

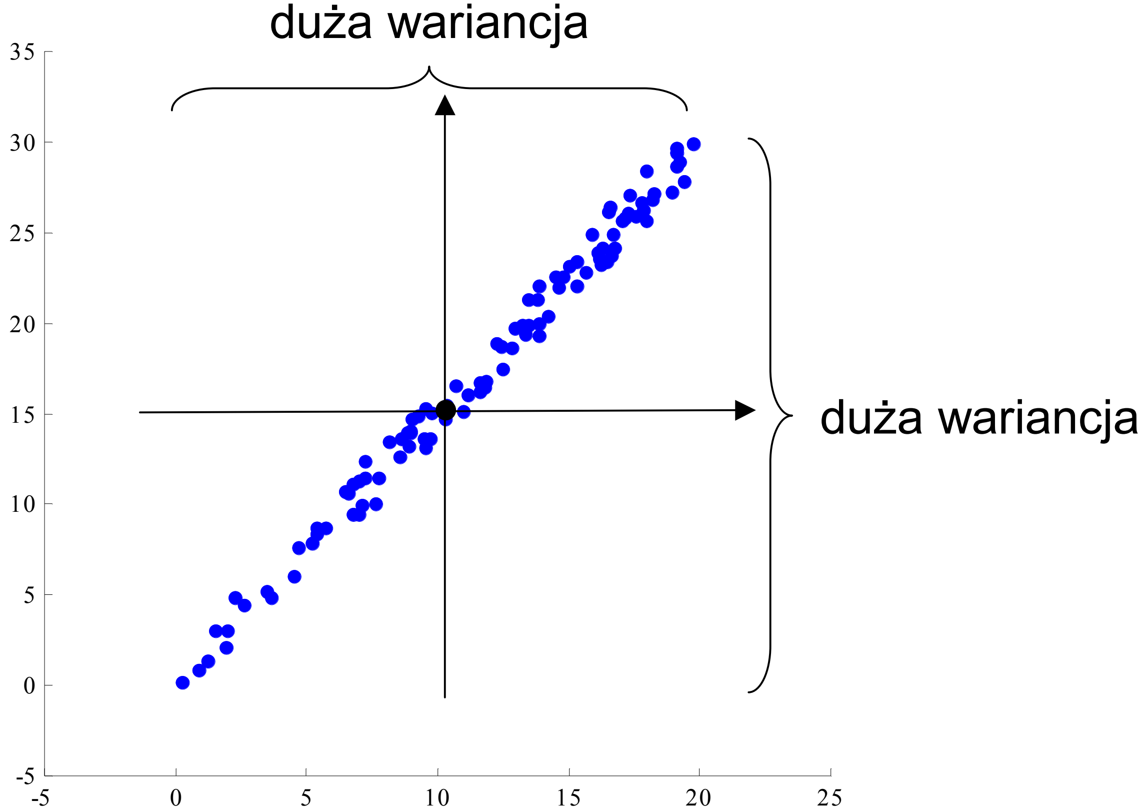


Przykład: PCA dla danych dwuwymiarowych

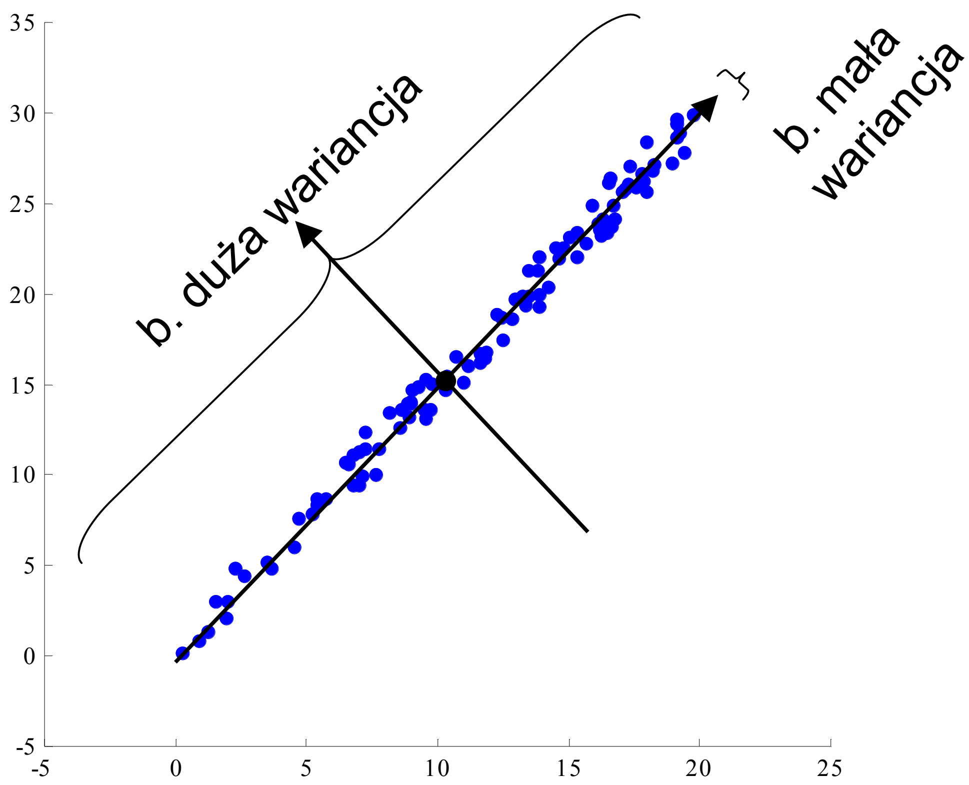
$\mathbf{x}_1$	$\mathbf{x}_2$
13.34	19.37
17.57	25.89
5.40	8.33
14.63	21.99
1.95	3.00
11.60	16.71
11.76	16.49
7.25	12.36
18.24	27.17
8.99	14.70
17.31	27.05
6.77	9.45
10.29	14.67
7.02	9.44
1.52	3.04
...	...
...	...
...	...



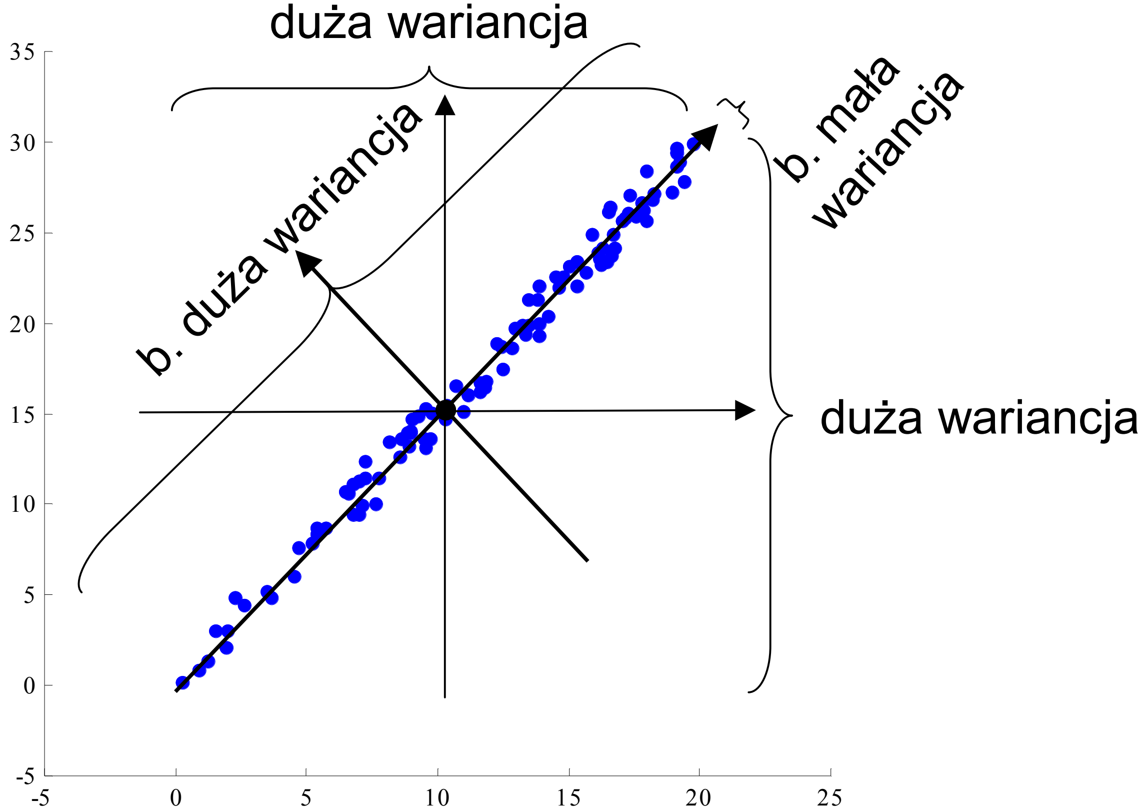
Przykład: PCA dla danych dwuwymiarowych



Przykład: PCA dla danych dwuwymiarowych

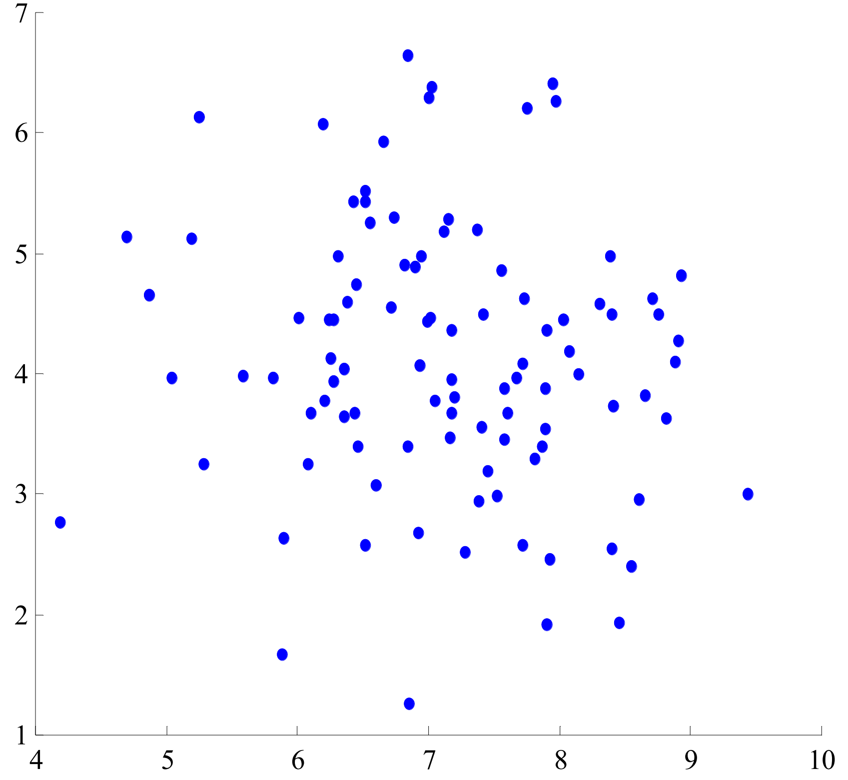


Przykład: PCA dla danych dwuwymiarowych

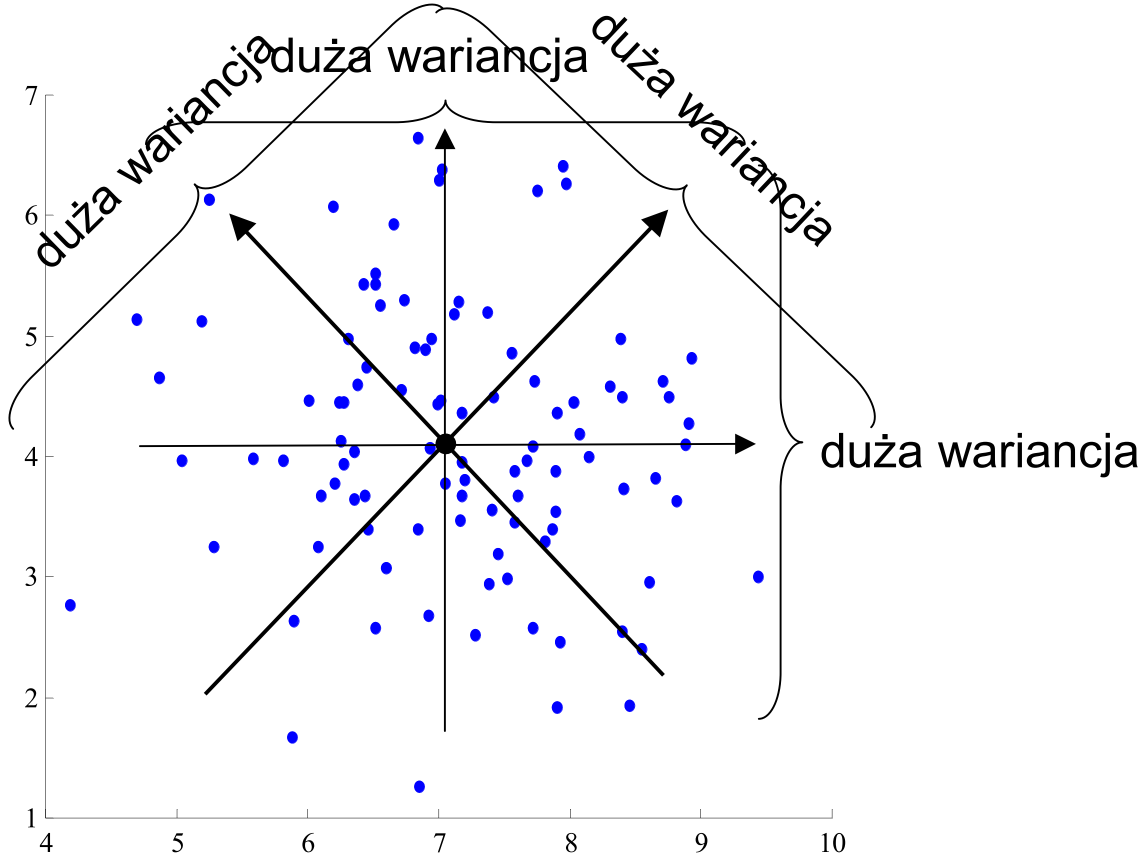


- (pod względem wariancji)
  - stare zmienne: małe zróznicowanie
  - nowe zmienne: duże zróznicowanie

Przykład: PCA dla danych dwuwymiarowych



Przykład: PCA dla danych dwuwymiarowych



- (pod względem wariancji)
  - stare zmienne: małe zróżnicowanie
  - nowe zmienne: małe zróżnicowanie

Procedura PCA

- Dana jest macierz danych: **X** (obserwacje w wierszach)
  - oblicz macierz kowariancji $\mathbf{S}_x = \mathbf{X}^T\mathbf{X}$
    - (symetryczna) macierz $\mathbf{S}_x$ zawiera wartości pozwalające ocenić
      - wariancje zmiennych (elementy na głównej przekątnej)
      - zależności pomiędzy zmiennymi (elementy poza główną przekątną)

Procedura PCA

- Dana jest macierz danych: $\mathbf{X}$ (obserwacje w wierszach)
  - oblicz macierz kowariancji $\mathbf{S}_x = \mathbf{X}^T\mathbf{X}$
  - dokonaj rozkładu EVD: $\mathbf{S}_x = \mathbf{KLK}^T$
  - utwórz nowe zmienne wykonując operację $\mathbf{Y} = \mathbf{XK}$

Procedura PCA

- Założenia dotyczące macierzy danych **X**, wynikające z faktu, że przekształcenie PCA jest wrażliwe na
  - położenie środka układu współrzędnych
    - wszystkie obserwacje macierzy **X** powinny być wycentrowane (wartości średnie zmiennych wynoszą wtedy zero)
      - wycentrowanie obserwacji pozwala na obliczanie macierzy wariancji (z pominięciem współczynnika skalującego) jako $\mathbf{X}^T\mathbf{X}$
  - skalę jednostek układu współrzędnych

Procedura PCA

- Dodatkowe uwagi dotyczące przekształcenia PCA
  - wykorzystywana macierz kowariancji jest tak naprawdę macierzą
    - (z punktu widzenia skalarnego) sum iloczynów skalarów
    - (z punktu widzenia wektorowego) iloczynów skalarnych wektorów
    - (z punktu widzenia macierzowego) Grama
  - przekształcenie jest skuteczne tylko wtedy gdy pomiędzy zmiennymi macierzy danych **X** występują zależności liniowe (co przejawia się dodatnimi kowariancjami zmiennych)

Procedura PCA

- Dane są macierze:
  - macierz danych oryginalnych **X**
  - macierz kowariancji danych oryginalnych $\mathbf{S}_x = \mathbf{X}^T\mathbf{X}$
  - macierz danych przekształconych $\mathbf{Y} = \mathbf{XK}$, gdzie **K** jest macierzą wektorów własnych macierzy kowariancji $\mathbf{S}_x = \mathbf{X}^T\mathbf{X}$, czyli $\mathbf{S}_x = \mathbf{KLK}^T$
- Wtedy po pierwsze:
  - macierz kowariancji danych przekształconych $\mathbf{S}_\mathbf{y}$ można wyrazić jako:

$$\mathbf{S}_y = \mathbf{Y}^T\mathbf{Y} = (\mathbf{XK})^T\mathbf{XK} = \mathbf{K}^T\mathbf{X}^T\mathbf{XK} = \mathbf{K}^T\mathbf{S}_x\mathbf{K} = \mathbf{K}^T\mathbf{KLK}^T\mathbf{K} = \mathbf{ILI} = \mathbf{L}$$

Procedura PCA

- Wniosek 1:
  macierz kowariancji $\mathbf{S}_Y$ zmiennych przekształconych **Y** wyraża się macierzą diagonalną **L** (utworzoną z wartości własnych macierzy $\mathbf{S}_x$)
  - kowariancje zmiennych przekształconych (elementy pozadiagonalne macierzy $\mathbf{S}_Y$) są równe zero
    - nowe zmienne są niezależne liniowo
  - wariancje zmiennych przekształconych (elementy diagonalne macierzy $\mathbf{S}_Y$) są równe wartościom własnym macierzy $\mathbf{S}_x$ (zapisanym na przekątnej macierzy **L**)
    - można je poznać tuż po wyliczeniu wartości własnych macierzy $\mathbf{S}_x$,
      a więc jeszcze przed wyliczeniem nowej macierzy danych **Y** = **XK**

Procedura PCA

- Dane są macierze:
  - macierz danych oryginalnych $\mathbf{X}$
  - macierz kowariancji danych oryginalnych $\mathbf{S}_x = \mathbf{X}^T\mathbf{X}$
  - macierz danych przekształconych $\mathbf{Y} = \mathbf{XK}$, gdzie $\mathbf{K}$ jest macierzą wektorów własnych macierzy kowariancji $\mathbf{S}_x = \mathbf{X}^T\mathbf{X}$, czyli $\mathbf{S}_x = \mathbf{KLK}^T$
- Wtedy po drugie:
  - wykorzystując: $\mathbf{S}_y = \mathbf{K}^T\mathbf{S}_x\mathbf{K}$ (z poprzedniego wyprowadzenia) mamy:
    $\mathrm{tr}(\mathbf{S}_y) = \mathrm{tr}(\mathbf{K}^T\mathbf{S}_x\mathbf{K})$
    - ale $\mathrm{tr}(\mathbf{K}^T\mathbf{S}_x\mathbf{K}) = \mathrm{tr}(\mathbf{KK}^T\mathbf{S}_x) = \mathrm{tr}(\mathbf{IS}_x) = \mathrm{tr}(\mathbf{S}_x)$
      - ponieważ $\forall_{\mathbf{A},\mathbf{B},\mathbf{C}}$: $\mathrm{tr}(\mathbf{ABC}) = \mathrm{tr}(\mathbf{CAB})$
  - czyli
    $\mathrm{tr}(\mathbf{S}_y) = \mathrm{tr}(\mathbf{S}_x)$

Procedura PCA

- Wniosek 2:
  suma wariancji zmiennych przekształconych **Y** jest równa sumie wariancji zmiennych oryginalnych **X**
  - czyli suma wariancji przed przekształceniem jest taka sama jak po przekształceniu
    - nic w przyrodzie nie ginie! :-)

Procedura PCA

- Uwaga dotycząca zależności „suma wariancji przed przekształceniem jest taka sama jak po przekształceniu”
  - macierz ortogonalna **K** będąca macierzą wektorów własnych macierzy $\mathbf{S}_X$ nie jest jedyną macierzą zapewniającą powyższą zależność
  - zależność ta jest spełniona dla każdej macierzy ortogonalnej
    - a w szczególności: dla macierzy **K**
  - uzasadnienie
    - niech $\mathbf{Y} = \mathbf{XU}$, gdzie **U** jest ortogonalna, wtedy:

$$
\begin{aligned}
\mathrm{tr}(\mathbf{S}_Y) &= \mathrm{tr}(\mathbf{Y}^T\mathbf{Y}) = \mathrm{tr}((\mathbf{XU})^T\mathbf{XU}) = \mathrm{tr}(\mathbf{U}^T\mathbf{X}^T\mathbf{XU}) = \\
&= \mathrm{tr}(\mathbf{UU}^T\mathbf{X}^T\mathbf{X}) = \mathrm{tr}(\mathbf{IX}^T\mathbf{X}) = \mathrm{tr}(\mathbf{X}^T\mathbf{X}) = \mathrm{tr}(\mathbf{S}_X)
\end{aligned}
$$

Procedura PCA

- Pytanie:
  dlaczego to właśnie taka a nie inna macierz **K** jest poszukiwaną w metodzie PCA macierzą przekształcającą stare zmienne w nowe?
- Odpowiedź:
  ponieważ $\mathbf{S}_Y = \mathbf{Y}^T\mathbf{Y}$, gdzie $\mathbf{Y} = \mathbf{XK}$, jest macierzą diagonalną
  - dzięki czemu kolumny macierzy **Y** stanowią zbiór wektorów niezależnych (liniowo)
- Jest tak dlatego, że **K** jest macierzą, która diagonalizuje macierz kowariancji $\mathbf{S}_X = \mathbf{X}^T\mathbf{X}$
  - tzn. zapewnia, że $\mathbf{K}^T\mathbf{S}_x\mathbf{K}$ jest macierzą diagonalną
  - diagonalność $\mathbf{S}_Y$ wynika z faktu, że $\mathbf{S}_Y = \mathbf{K}^T\mathbf{S}_X\mathbf{K}$
    - ponieważ $\mathbf{S}_Y = \mathbf{Y}^T\mathbf{Y} = \mathbf{K}^T\mathbf{X}^T\mathbf{XK} = \mathbf{K}^T\mathbf{S}_X\mathbf{K}$

…

Procedura PCA

- Nowe zmienne (kolumny macierzy **Y**) powstają ze starych zmiennych (kolumn macierzy **X**) w rezultacie operacji $\mathbf{Y} = \mathbf{XK}$
  - gdzie **K** jest macierzą wektorów własnych macierzy $\mathbf{S}_x = \mathbf{X}^T\mathbf{X}$

Procedura PCA

- Dane są macierze:
  - macierz danych oryginalnych **X** (dane w wierszach)
  - macierz przekształcająca **K**
- Nowe dane (wiersze macierzy **Y**) powstają w rezultacie **Y** = **XK**

$$\begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \mathbf{x}_3^T \\ \dots \\ \mathbf{x}_N^T \end{bmatrix} \times \begin{bmatrix} \mathbf{k}_1 & \mathbf{k}_2 & \mathbf{k}_3 \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} \\ \dots & \dots & \dots \\ y_{N1} & y_{N2} & y_{N3} \end{bmatrix}$$

Procedura PCA

- Element macierzy y_{ij} jest iloczynem skalarnym wiersza $(\mathbf{x}_i)^T$ macierzy **X** oraz kolumny $\mathbf{k}_j$ macierzy **K**: $y_{ij} = (\mathbf{x}_i)^T\mathbf{k}_j$
  - liczba nowych obserwacji = liczba starych obserwacji
  - liczba nowych zmiennych = liczba starych zmiennych

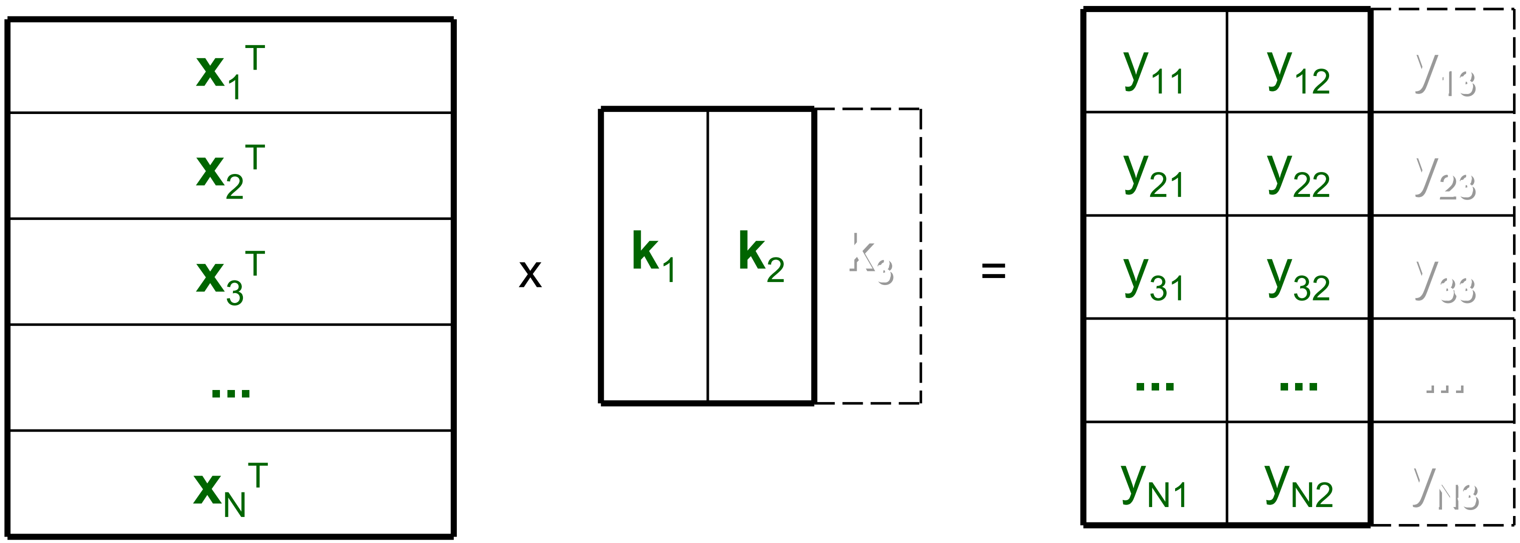
$$\begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \mathbf{x}_3^T \\ \dots \\ \mathbf{x}_N^T \end{bmatrix} \times \begin{bmatrix} \mathbf{k}_1 & \mathbf{k}_2 & \mathbf{k}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{x}_1^T\mathbf{k}_1 & \mathbf{x}_1^T\mathbf{k}_2 & \mathbf{x}_1^T\mathbf{k}_3 \\ \mathbf{x}_2^T\mathbf{k}_1 & \mathbf{x}_2^T\mathbf{k}_2 & \mathbf{x}_2^T\mathbf{k}_3 \\ \mathbf{x}_3^T\mathbf{k}_1 & \mathbf{x}_3^T\mathbf{k}_2 & \mathbf{x}_3^T\mathbf{k}_3 \\ \dots & \dots & \dots \\ \mathbf{x}_N^T\mathbf{k}_1 & \mathbf{x}_N^T\mathbf{k}_2 & \mathbf{x}_N^T\mathbf{k}_3 \end{bmatrix}$$

Procedura PCA

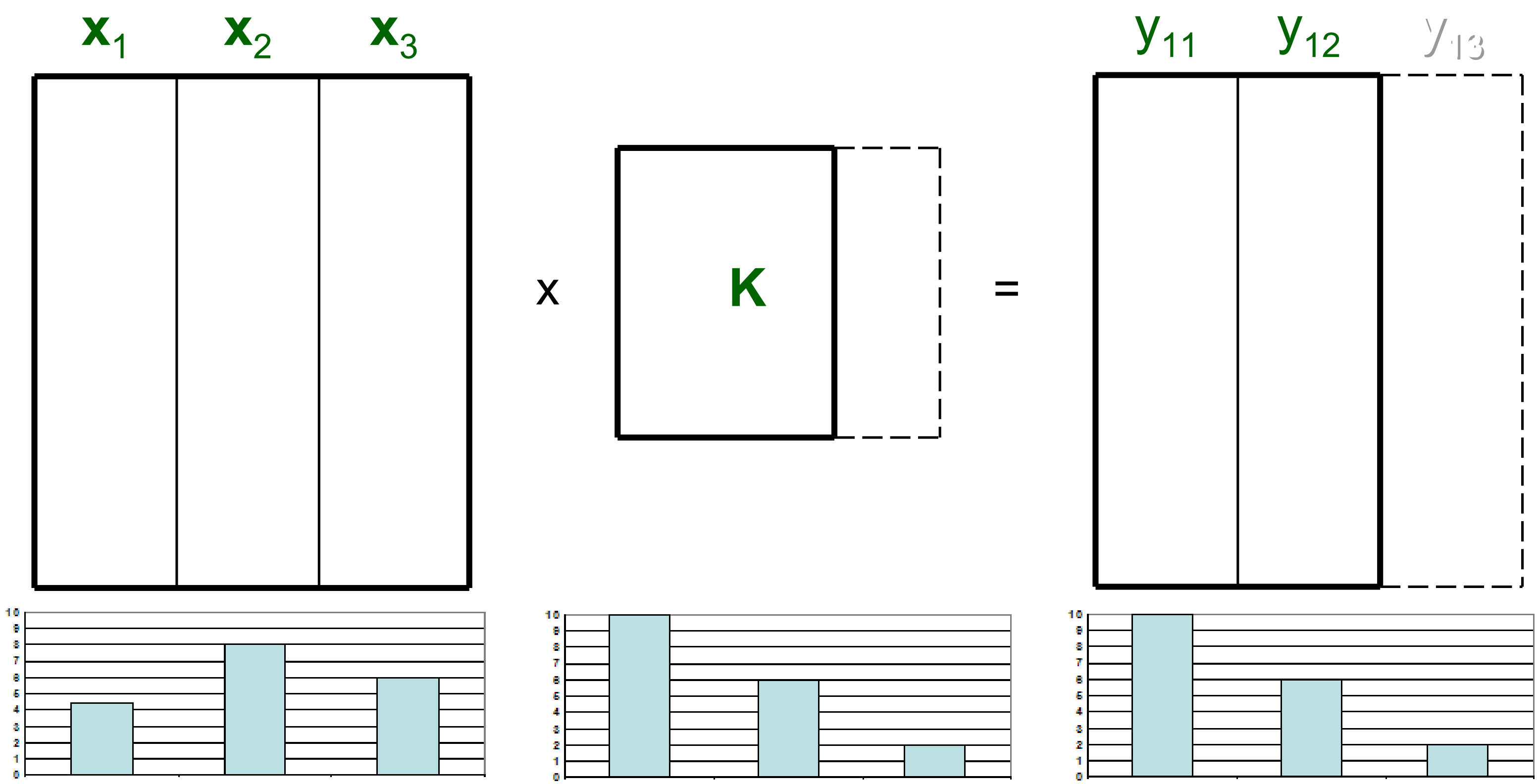
- Ważne spostrzeżenie:
  - jeżeli jakieś nowe zmienne miałyby mieć wariancję bliską zeru, to można pominąć tworzenie tych zmiennych, oszczędzając
    - moc obliczeniową (przekształcenie będzie mniej wymagające)
    - czas (przekształcenie potrwa krócej)
    - pamięć (nowa macierz zmiennych zajmie mniej miejsca)

Procedura PCA

- Liczba nowych zmiennych zależy od liczby kolumn macierzy przekształcającej **K**
  - zmniejszenie (uwzględnionych) liczby kolumn tej macierzy prowadzi do zmniejszenia nowych zmiennych



Zasada zachowania informacji (wariancji)



var(x_1) var(x_2) var(x_3) λ_1 λ_2 λ_3 var(y_1) var(y_2) var(y_3)

λ_1 = var(y_1) λ_2 = var(y_2) λ_3 = var(y_3)

var(x_1)+var(x_2)+var(x_3) = $\lambda_1 + \lambda_2 + \lambda_3$

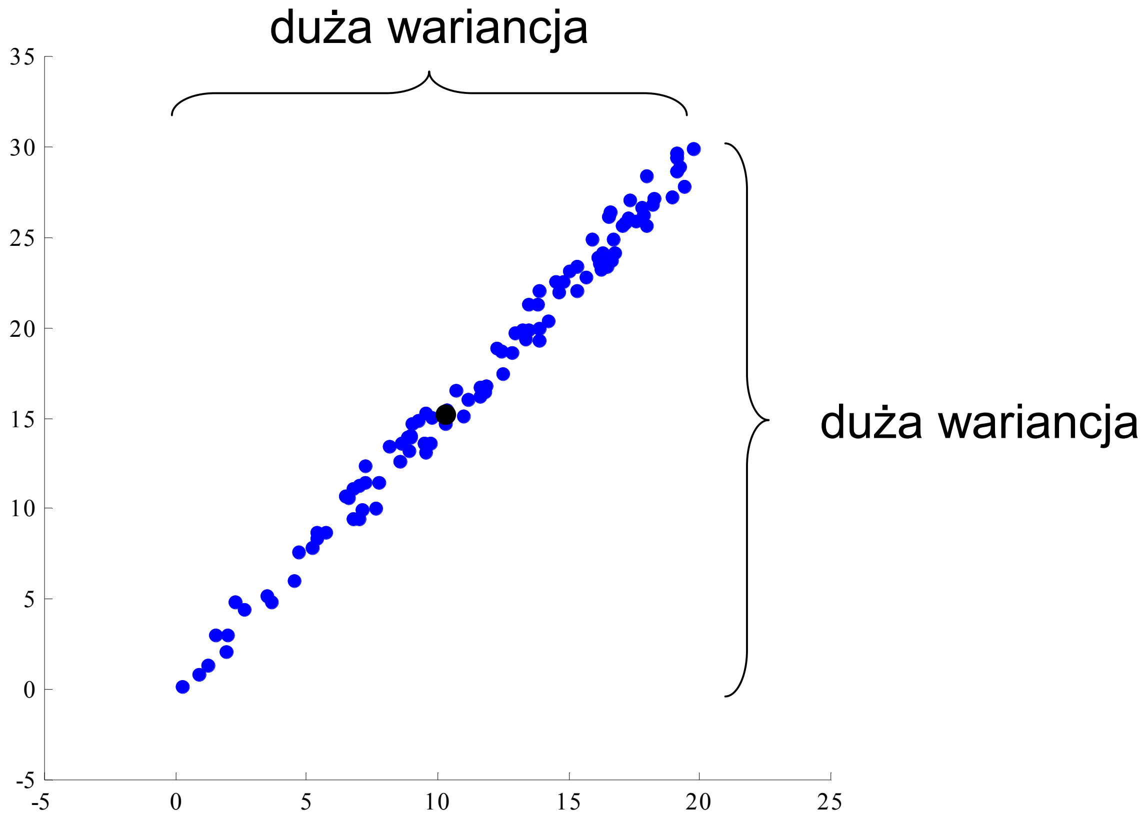
var(x_1)+var(x_2)+var(x_3) = var(y_1)+var(y_2)+var(y_3)

Przekształcenie odwrotne

- Pytania:
  - jak na podstawie nowych zmiennych odtworzyć stare zmienne?
  - jak to zrobić, jeżeli zredukowano liczbę nowych zmiennych?
- Odpowiedzi:
  - odtworzenia starych zmiennych na podstawie nowych można dokonać dokonując przekształcenia odwrotnego
    - ponieważ $\mathbf{Y} = \mathbf{XK}$, to $\mathbf{YK}^{-1} = \mathbf{XKK}^{-1}$, i wtedy $\mathbf{YK}^{-1} = \mathbf{XI}$, czyli $\mathbf{X} = \mathbf{YK}^{-1}$
    - ponieważ w rozważanym przypadku $\mathbf{K}^{-1} = \mathbf{K}^{\mathrm{T}}$, wystarczy wykonać $\mathbf{YK}^{\mathrm{T}}$
    - odtworzenie wszystkich zmiennych oryginalnych jest (formalnie) tylko możliwe z wykorzystaniem całej, niezredukowanej macierzy **Y**
  - jeżeli pewne zmienne przekształcone (czyli kolumny macierzy **Y**) zostały zredukowane, to należy odtworzyć je przed wykonaniem mnożenia
    - ponieważ oryginalne wartości zredukowanych kolumn nie są znane, odtwarzamy je z wykorzystaniem wartości średnich tych kolumn
      - dla wycentrowanych danych będą to oczywiście wartości zerowe

Przykład: PCA dla danych dwuwymiarowych

$\mathbf{x}_1$	$\mathbf{x}_2$
13.34	19.37
17.57	25.89
5.40	8.33
14.63	21.99
1.95	3.00
11.60	16.71
11.76	16.49
7.25	12.36
18.24	27.17
8.99	14.70
17.31	27.05
6.77	9.45
10.29	14.67
7.02	9.44
1.52	3.04
...	...
...	...
...	...



Utworzenie macierzy przekształcającej

$$\mathbf{S}_x = \begin{bmatrix} 27.4 & 40.6 \\ 40.6 & 60.9 \end{bmatrix}$$

$$\begin{array}{cc} \lambda_1 & \lambda_2 \\ 0.2 & 88.1 \end{array}$$

$$\begin{array}{cc} \mathbf{k}_1 & \mathbf{k}_2 \\ 0.83 & 0.55 \\ -0.55 & 0.83 \end{array}$$

$$\underset{\mathbf{K}}{\begin{bmatrix} 0.55 & 0.83 \\ 0.83 & -0.55 \end{bmatrix}} \times \underset{\mathbf{L}}{\begin{bmatrix} 88.1 & 0 \\ 0 & 0.2 \end{bmatrix}} \times \underset{\mathbf{K}^T}{\begin{bmatrix} 0.55 & 0.83 \\ 0.83 & -0.55 \end{bmatrix}} = \underset{\mathbf{S}_x}{\begin{bmatrix} 27.4 & 40.6 \\ 40.6 & 60.9 \end{bmatrix}}$$

Wykorzystanie macierzy przekształcającej

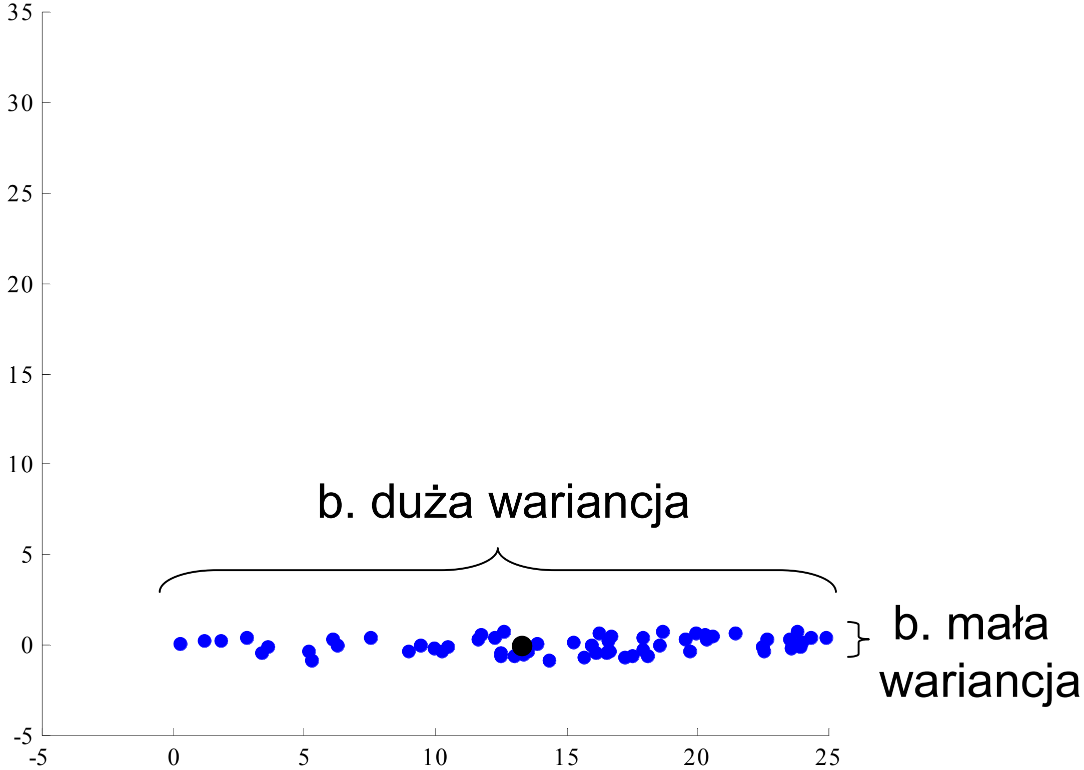
x_1	x_2
13.34	19.37
17.57	25.89
5.40	8.33
14.63	21.99
1.95	3.00
11.60	16.71
11.76	16.49
7.25	12.36
18.24	27.17
8.99	14.70
17.31	27.05
6.77	9.45
10.29	14.67
7.02	9.44
1.52	3.04
...	...
...	...
...	...

$$\times \quad \mathbf{K} = \begin{bmatrix} 0.55 & 0.83 \\ 0.83 & -0.55 \end{bmatrix} \quad =$$

y_1	y_2
23.13	-0.68
30.73	-2.30
9.71	1.51
25.89	-1.63
3.50	2.84
20.02	-0.03
19.98	0.23
13.86	-0.04
32.11	-2.73
16.75	-0.45
31.37	-3.30
11.47	1.68
17.65	0.48
11.64	1.86
3.22	2.50
...	...
...	...
...	...

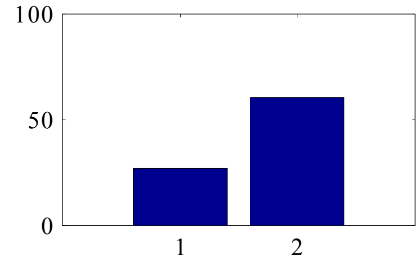
Nowe zmienne (zmienne przekształcone)

y_1	y_2
23.13	-0.68
30.73	-2.30
9.71	1.51
25.89	-1.63
3.50	2.84
20.02	-0.03
19.98	0.23
13.86	-0.04
32.11	-2.73
16.75	-0.45
31.37	-3.30
11.47	1.68
17.65	0.48
11.64	1.86
3.22	2.50
...	...
...	...
...	...

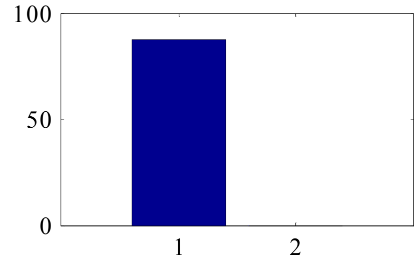


Porównanie wariancji i redukcja zmiennych

- Wariancje zmiennych oryginalnych:
  - Var($\mathbf{x}_1$) = 27.4, Var($\mathbf{x}_2$) = 60.9
  - suma: 27.4 + 60.9 = 88.3



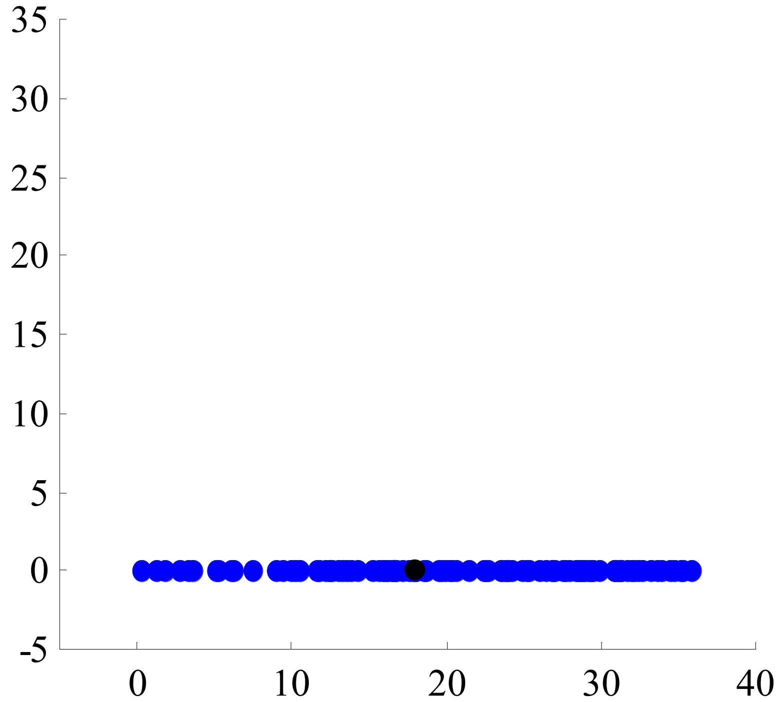
- Wariancje zmiennych przekształconych:
  - Var($\mathbf{y}_1$) = 88.1, Var($\mathbf{y}_2$) = 0.2
  - suma: 88.1 + 0.2 = 88.3



- Wniosek: ze względu na małą pojemność informacyjną (wyrażającą się małą wariancją) zmienna $\mathbf{y}_2$ może zostać pominięta w dalszych analizach
- W praktyce redukowanie zmiennych sprowadza się do utworzenia nowych zmiennych, takich, że ich wartości są wartościami średnimi zmiennej redukowanej
- Wszystkie zmienne po redukcji będą oznaczane przez $\mathbf{z}_i$

Nowe zmienne (zmienne zredukowane)

$\mathbf{z}_1$	$\mathbf{z}_2$
23.13	0.00
30.73	0.00
9.71	0.00
25.89	0.00
3.50	0.00
20.02	0.00
19.98	0.00
13.86	0.00
32.11	0.00
16.75	0.00
31.37	0.00
11.47	0.00
17.65	0.00
11.64	0.00
3.22	0.00
...	...
...	...
...	...



Wykorzystanie macierzy odwrotnej: odtwarzanie zmiennych oryginalnych ze zredukowanych

z_1	z_2
23.13	0.00
30.73	0.00
9.71	0.00
25.89	0.00
3.50	0.00
20.02	0.00
19.98	0.00
13.86	0.00
32.11	0.00
16.75	0.00
31.37	0.00
11.47	0.00
17.65	0.00
11.64	0.00
3.22	0.00
...	...
...	...
...	...

x

$\mathbf{K}^T$

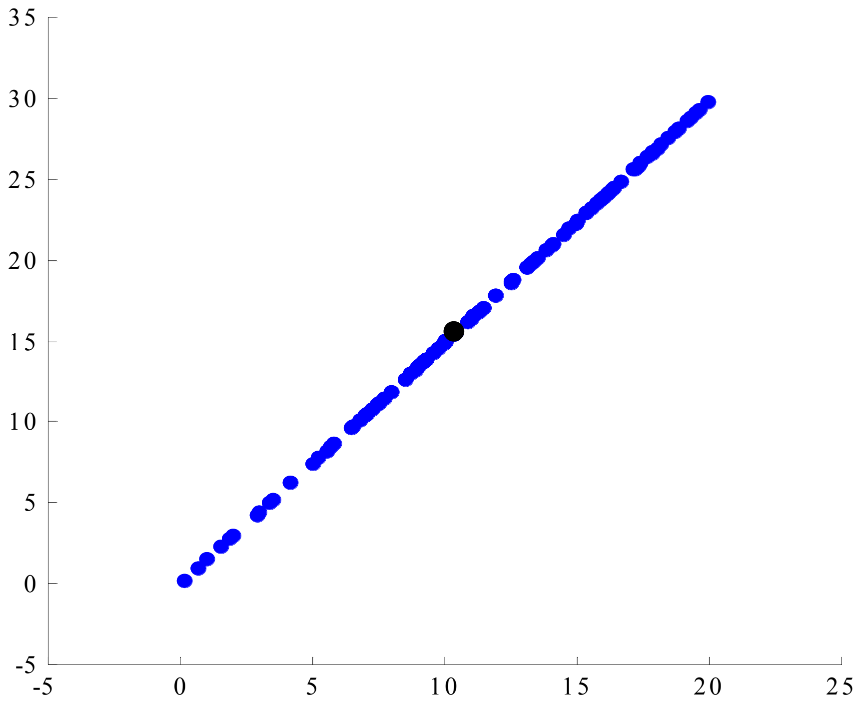
0.55	0.83
0.83	–0.55

=

u_1	u_2
13.48	19.22
18.86	24.60
3.99	9.73
15.44	21.18
-0.39	5.35
11.29	17.03
11.26	17.00
6.93	12.67
19.83	25.57
8.97	14.72
19.31	25.05
5.24	10.98
9.61	15.35
5.36	11.10
-0.59	5.15
...	...
...	...
...	...

Zmienne odtworzone w przybliżeniu (wygładzone)

u_1	u_2
13.48	19.22
18.86	24.60
3.99	9.73
15.44	21.18
-0.39	5.35
11.29	17.03
11.26	17.00
6.93	12.67
19.83	25.57
8.97	14.72
19.31	25.05
5.24	10.98
9.61	15.35
5.36	11.10
-0.59	5.15
...	...
...	...
...	...



Wykorzystanie macierzy odwrotnej: odtwarzanie zmiennych oryginalnych z przekształconych

$\mathbf{y}_1$	$\mathbf{y}_2$
23.13	-0.68
30.73	-2.30
9.71	1.51
25.89	-1.63
3.50	2.84
20.02	-0.03
19.98	0.23
13.86	-0.04
32.11	-2.73
16.75	-0.45
31.37	-3.30
11.47	1.68
17.65	0.48
11.64	1.86
3.22	2.50
…	…
…	…
…	…

x

$\mathbf{K}^T$

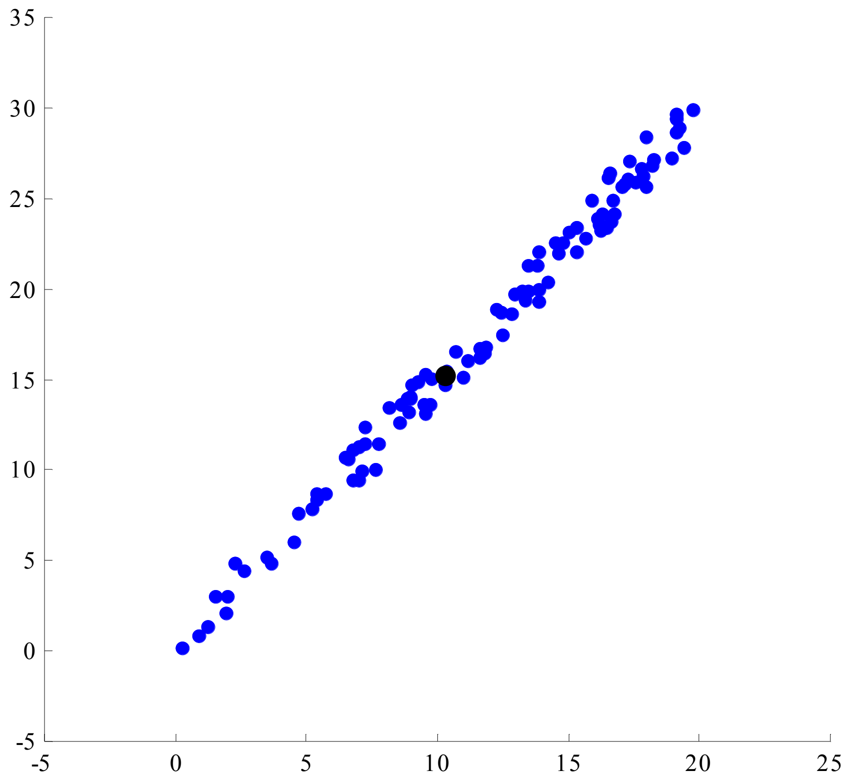
0.55	0.83
0.83	–0.55

=

$\mathbf{x}_1$	$\mathbf{x}_2$
13.34	19.37
17.57	25.89
5.40	8.33
14.63	21.99
1.95	3.00
11.60	16.71
11.76	16.49
7.25	12.36
18.24	27.17
8.99	14.70
17.31	27.05
6.77	9.45
10.29	14.67
7.02	9.44
1.52	3.04
…	…
…	…
…	…

Zmienne odtworzone w pełni

$\mathbf{x}_1$	$\mathbf{x}_2$
13.34	19.37
17.57	25.89
5.40	8.33
14.63	21.99
1.95	3.00
11.60	16.71
11.76	16.49
7.25	12.36
18.24	27.17
8.99	14.70
17.31	27.05
6.77	9.45
10.29	14.67
7.02	9.44
1.52	3.04
...	...
...	...
...	...



…