



# **Monitorowanie i Diagnostyka w Systemach Sterowania**

**Wydział Elektrotechniki i Automatyki  
Katedra Elektrotechniki, Systemów Sterowania i Informatyki  
dr hab. inż. Michał Grochowski**

# Monitorowanie i Diagnostyka w Systemach Sterowania

na studiach II stopnia specjalności: Systemy Sterowania i Podejmowania Decyzji

---

---

## Analiza składników podstawowych - wprowadzenie (Principal Components Analysis - PCA)

Opracował: dr inż. Michał Grochowski

[kiss.pg.mg@gmail.com](mailto:kiss.pg.mg@gmail.com)

[michal.grochowski@pg.edu.pl](mailto:michal.grochowski@pg.edu.pl)

tel: 58 347 23 57

## Motywacja:

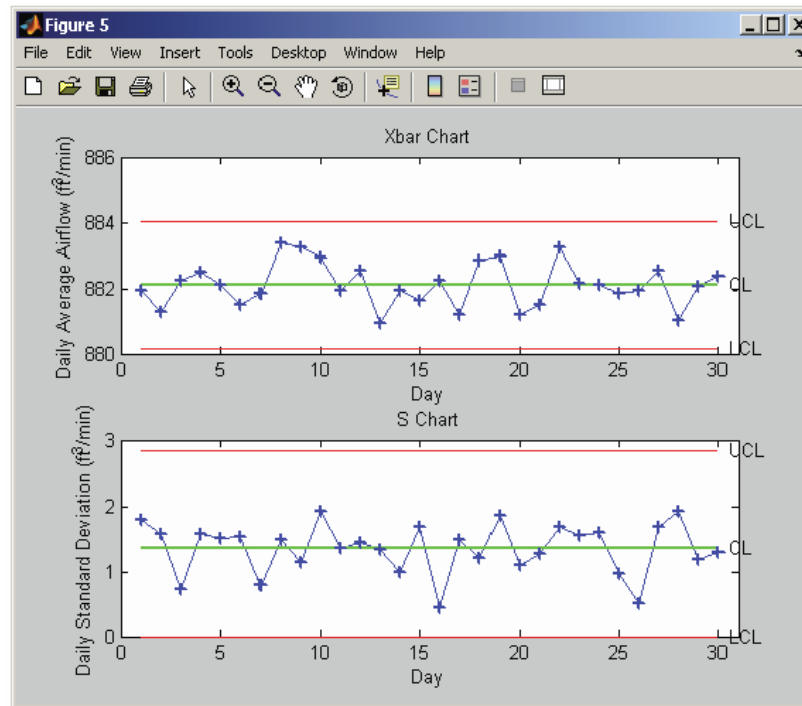
- niemożliwe jest śledzenie przez operatora setek/tysięcy zmiennych;
- w większości przypadków wczesne wykrycie awarii jest niemożliwe „nieuzbrojonym okiem”
- uszkodzenia/awarie/dysfunkcje są łatwiejsze do wykrycia poprzez analizę wielu zmiennych procesowych;

Jedno z możliwych rozwiązań:

**Wielowymiarowa analiza statystyczna,  
np. Analiza Składników Podstawowych  
(Principal Components Analysis – PCA)**

# Porównanie jedno i wielowymiarowych metod statystycznych:

- Najprostszą z metod statystycznych jest jednowymiarowa analiza statystyczna poszczególnych zmiennych procesu;
- dla każdej zmiennej ustalane są dolna i górna granica kontrolna, której przekroczenie wskazuje stan „alarmowy” (np. karty kontrolne Shewharta);
- jednowymiarowe metody statystyczne ustalają próg dla każdej ze zmiennych z osobna;

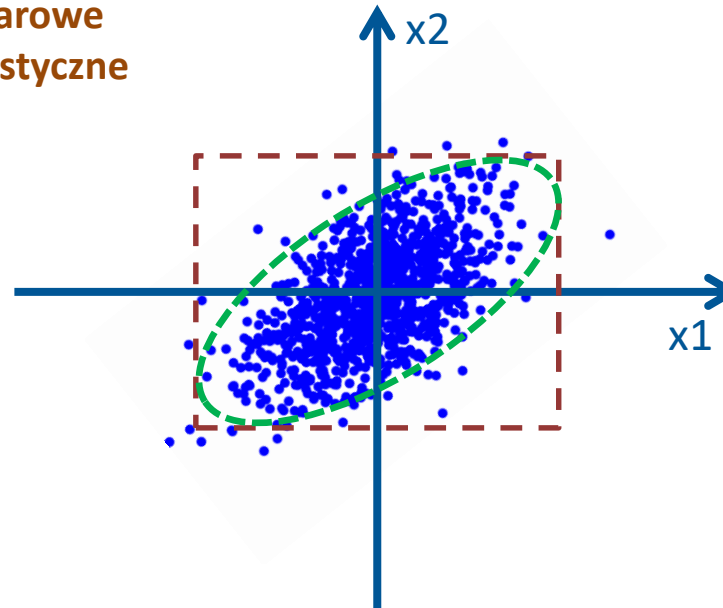


**UCL** – Górna granica kontrolna;  
**LCL** – Dolna granica kontrolna;

## Porównanie jedno i wielowymiarowych metod statystycznych:

- jednowymiarowe metody statystyczne ustalają próg dla każdej ze zmiennych z osobna;
- nie analizują (ignorują) fakt korelacji pomiędzy zmiennymi;
- dane „anormalne” mogą być zaklasyfikowane jako „normalne”

Jednowymiarowe  
metody statystyczne



Wielowymiarowe  
metody statystyczne

# Porównanie jedno i wielowymiarowych metod statystycznych:

## Przykład

Przy pomocy dwóch metod badano stężenie pewnego składnika w analizowanym związku chemicznym.

**Table 1.1. Data for Chemical Example**

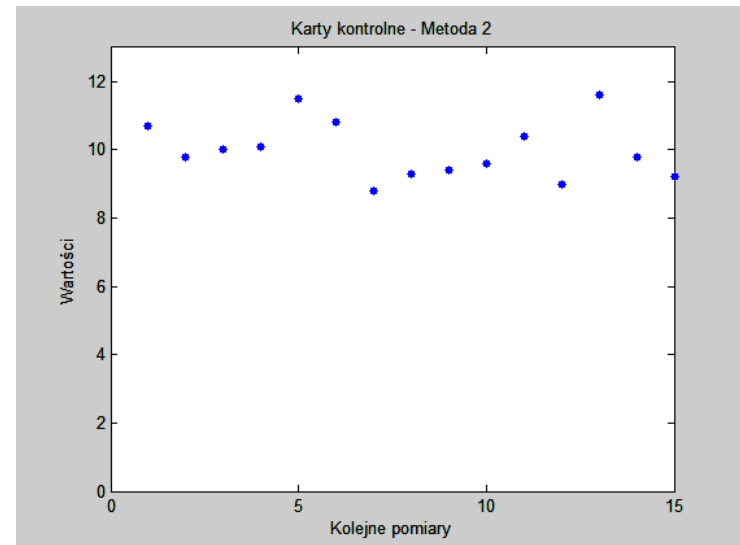
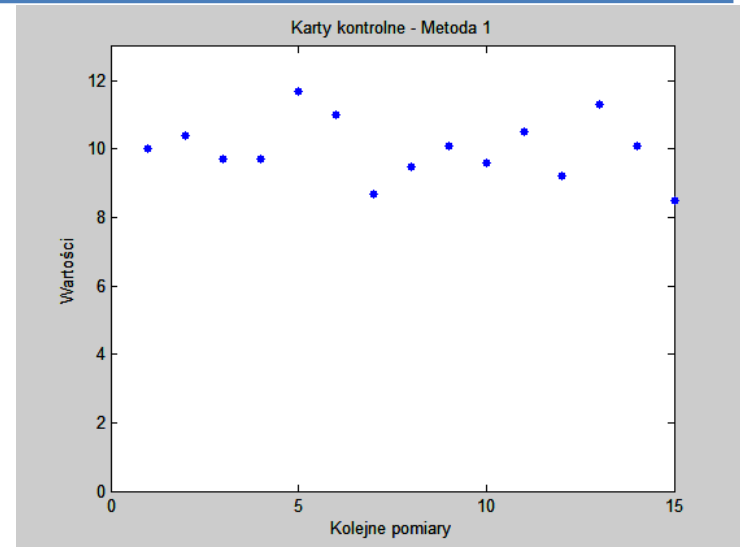
Obs. No.	Method 1	Method 2
1	10.0	10.7
2	10.4	9.8
3	9.7	10.0
4	9.7	10.1
5	11.7	11.5
6	11.0	10.8
7	8.7	8.8
8	9.5	9.3
9	10.1	9.4
10	9.6	9.6
11	10.5	10.4
12	9.2	9.0
13	11.3	11.6
14	10.1	9.8
15	8.5	9.2

źródło: S. Joe Qin. Process Chemometric Techniques and Applications. Department of Chemical Engineering. The University of Texas at Austin

# Porównanie jedno i wielowymiarowych metod statystycznych:

## Przykład

Próbki z analizowanego procesu były pobierane w momencie gdy proces przebiegał „normalnie”.



# Porównanie jedno i wielowymiarowych metod statystycznych:

## Przykład

Dodatkowo zebrano również 4 próbki (A,B,C,D) podczas których proces przebiegał w sposób „anormalny”.

**Table 1.1. Data for Chemical Example**

Obs. No.	Method 1	Method 2
1	10.0	10.7
2	10.4	9.8
3	9.7	10.0
4	9.7	10.1
5	11.7	11.5
6	11.0	10.8
7	8.7	8.8
8	9.5	9.3
9	10.1	9.4
10	9.6	9.6
11	10.5	10.4
12	9.2	9.0
13	11.3	11.6
14	10.1	9.8
15	8.5	9.2

Point	$x_1$	$x_2$
A	12.3	12.5
B	7.0	7.3
C	11.0	9.0
D	7.3	9.1

źródło: S. Joe Qin. Process Chemometric Techniques and Applications. Department of Chemical Engineering. The University of Texas at Austin

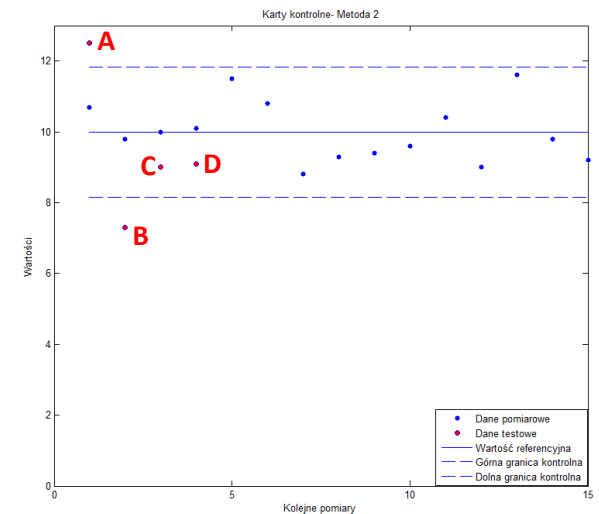
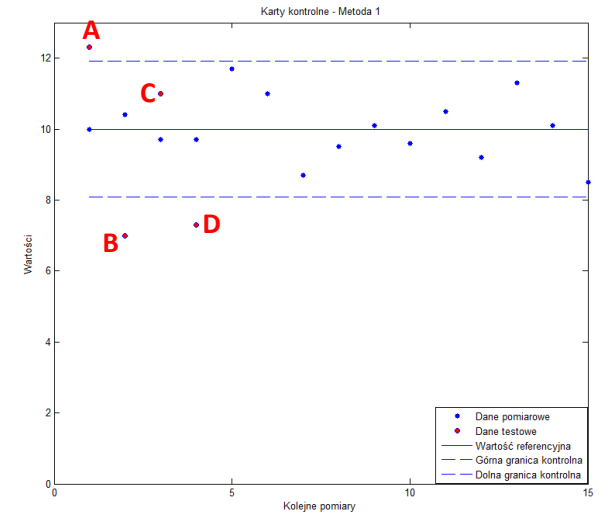


# Porównanie jedno i wielowymiarowych metod statystycznych:

## Przykład

W celu kontroli jakości przebiegu użyto Kart (wykresów) kontrolnych (Control Charts).

W tym celu wyznaczono granice kontrolne poprawności przebiegu procesu (dolną i górną), przy założeniu 95% ufności danym.



Point	$x_1$	$x_2$
A	12.3	12.5
B	7.0	7.3
C	11.0	9.0
D	7.3	9.1

# Porównanie jedno i wielowymiarowych metod statystycznych:

## Przykład

### Granice kontrolne

Odchylenie standardowe zmiennej pomnożone przez 2.145 (rozkład t studenta dla 95% i 14 stopni swobody).

Oznacza to że próbki będą się znajdować w granicach kontrolnych z prawdopodobieństwem 95%.

$$GGK1 = 10 + 2.145 * 0.89 = 11.91$$

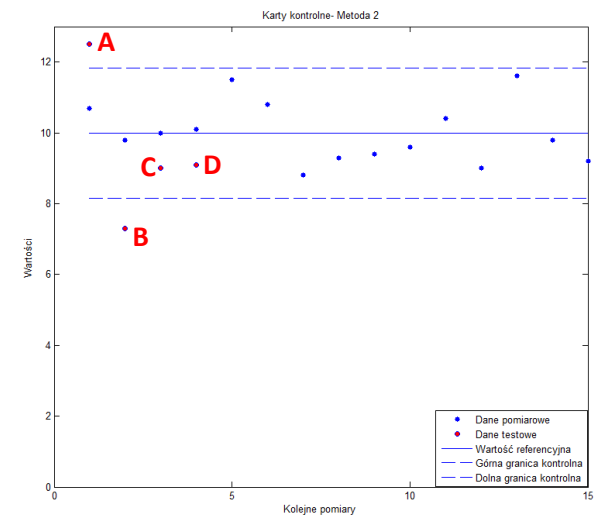
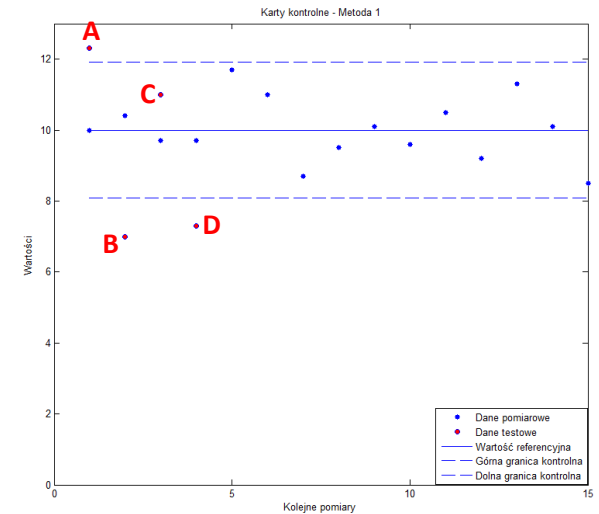
$$DGK1 = 10 - 2.145 * 0.89 = 8.09$$

$$GGK2 = 10 + 2.145 * 0.86 = 11.84$$

$$DGK2 = 10 - 2.145 * 0.86 = 8.16$$

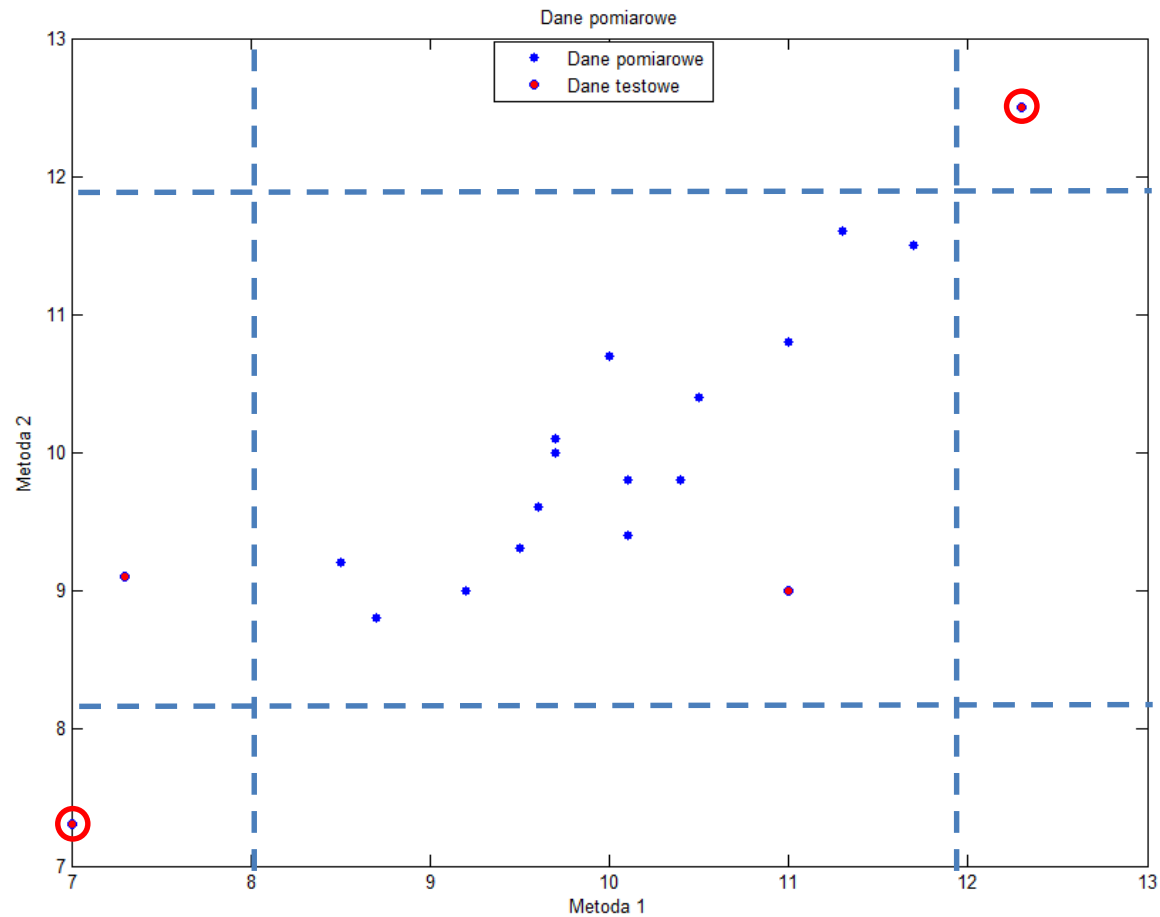
Jeżeli zmienne byłyby nieskorelowane (ale tu są) to prawdopodobieństwo że oba pomiary będą w granicach kontrolny wyniesie  $0.95 * 0.95 = 0.9 \Rightarrow 90\%$ .

Dla 9 nieskorelowanych zmiennych byłoby to  $0.95^9 = 0.63$ .  
Jeżeli zmienne są idealnie skorelowane to pozostaje 0.95,  
Jeżeli są trochę skorelowane to coś pomiędzy 0.9 a 0.95.



# Porównanie jedno i wielowymiarowych metod statystycznych:

## Przykład



GGK1 = 11.91

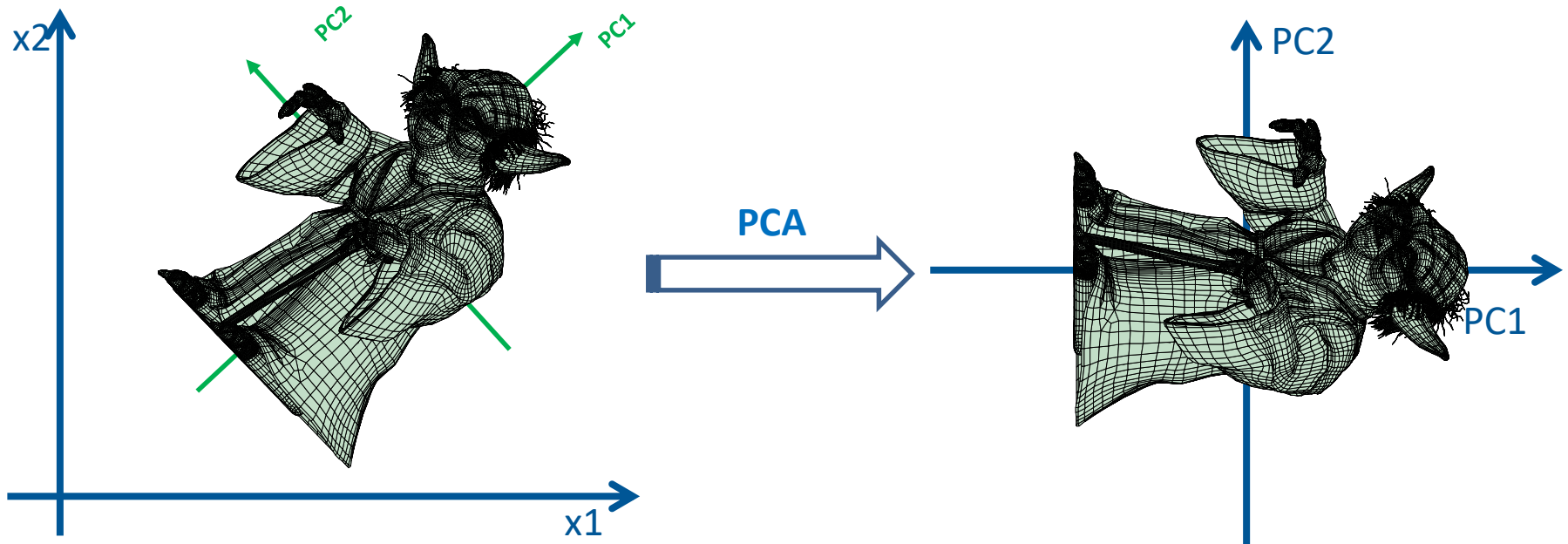
DGK1 = 8.09

GGK2 = 11.84

DGK2 = 8.16

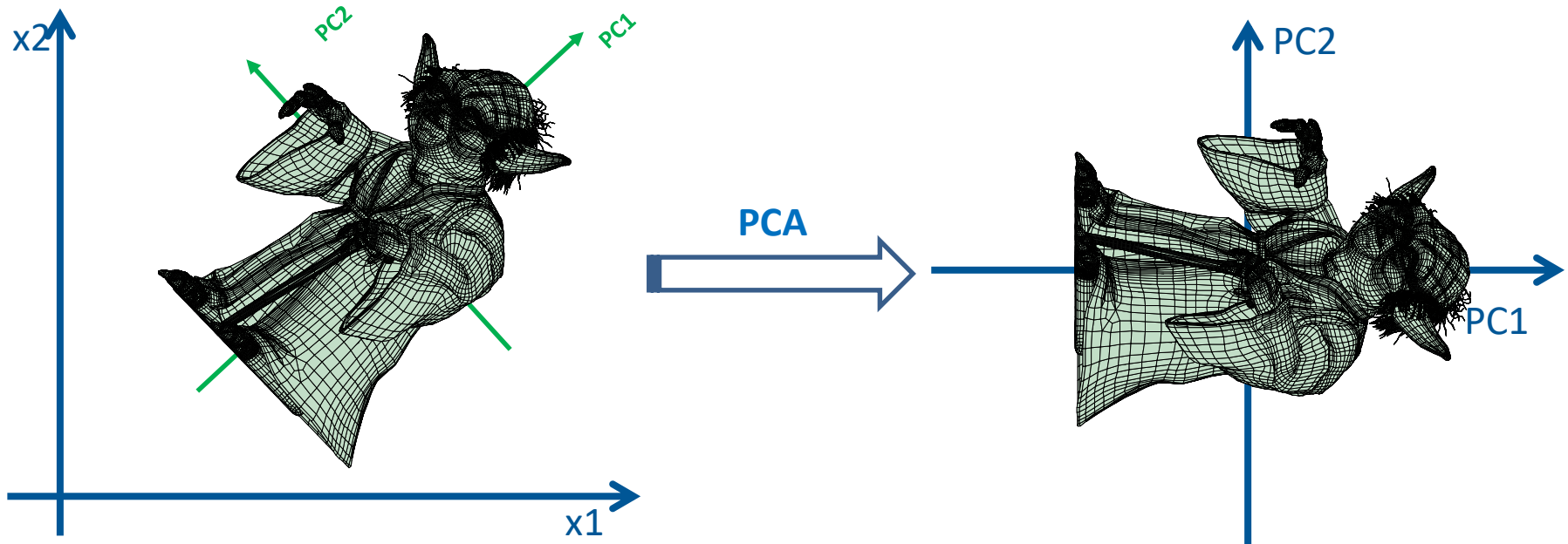
# Idea metody Analizy Składowych Podstawowych:

- Najpopularniejszą metodą wielowymiarowej analizy statystycznej jest metoda Analizy Składowych Głównych/Podstawowych – PCA (Principal Components Analysis );
- Ideą PCA jest dobór minimalnej ilości nowych zmiennych do możliwie dokładnego opisu wariacji (zmienności) danych wejściowych (np. pomiarowych), przy użyciu kombinacji liniowych oryginalnych zmiennych;
- Składowiki główne są dobrane zgodnie z kierunkami opisującymi największą zmienność (wariancję) procesu



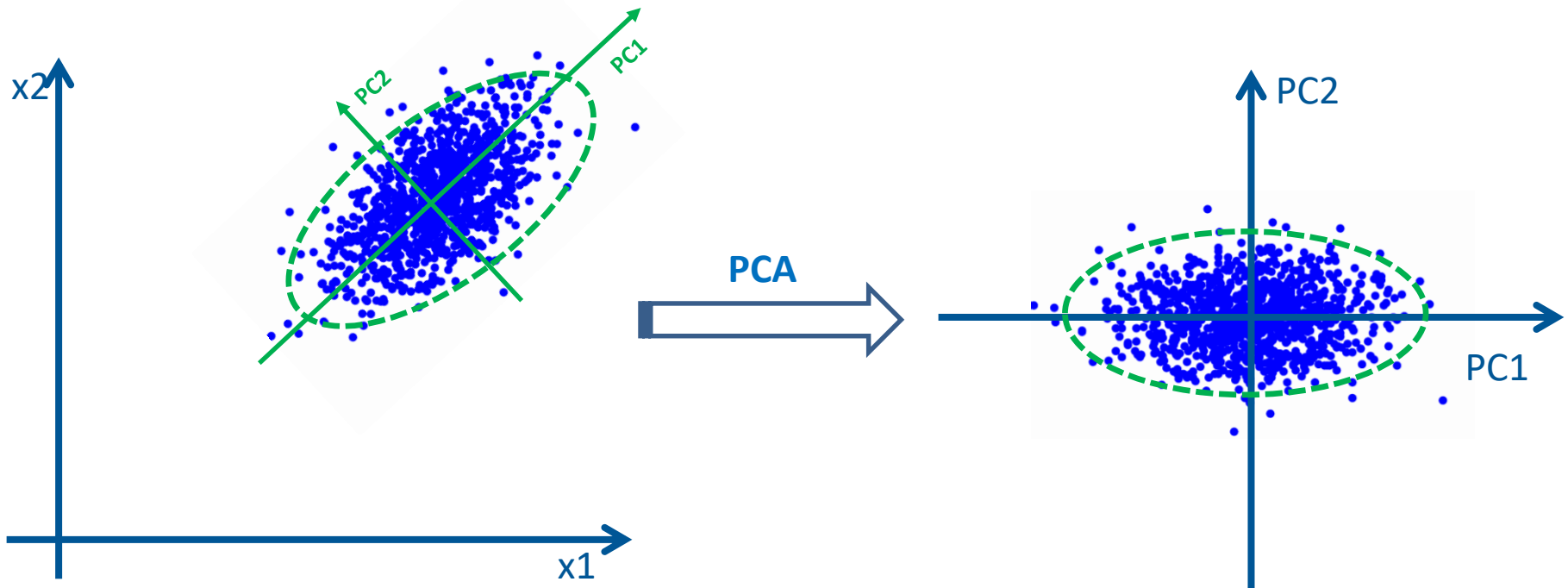
# Idea metody Analizy Składników Podstawowych:

- Tworzymy nowy układ współrzędnych oparty o nowe, nieskorelowane zmienne – składniki główne (PCs);
- Kierunki te są do siebie ortogonalne;
- Dane podczas obróbki wstępnej są normalizowane i sprowadzane do wartości średniej równej 0;
- Analiza PCA ma sens jedynie wtedy gdy dane są ze sobą skorelowane !!!



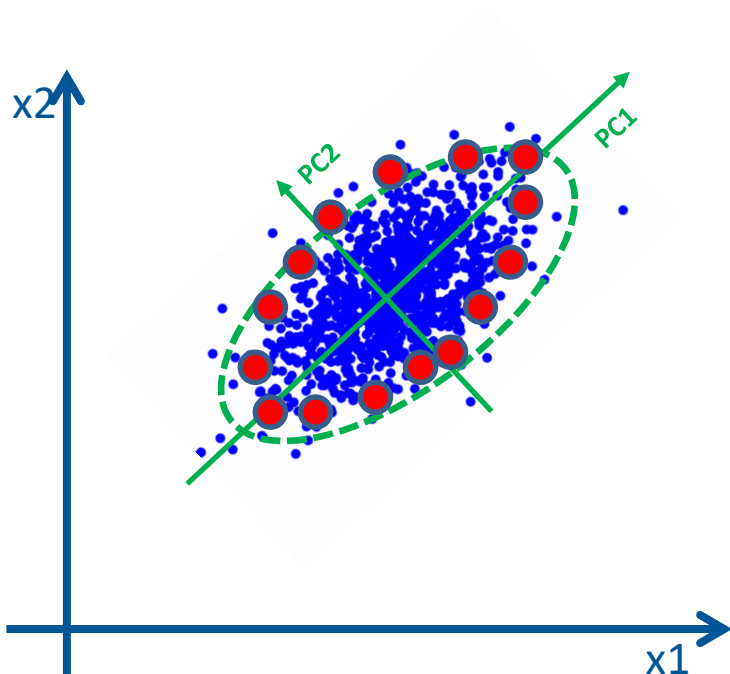
# Idea metody Analizy Składowików Podstawowych:

- Tworzymy nowy układ współrzędnych oparty o nowe, nieskorelowane zmienne – składowiki główne (PCs);
- Kierunki te są do siebie ortogonalne;
- Dane podczas obróbki wstępnej są normalizowane i sprowadzane do wartości średniej równej 0;
- Analiza PCA ma sens jedynie wtedy gdy dane są ze sobą skorelowane !!!



# Idea metody Analizy Składników Podstawowych:

Rozkład Gaussowski 2D

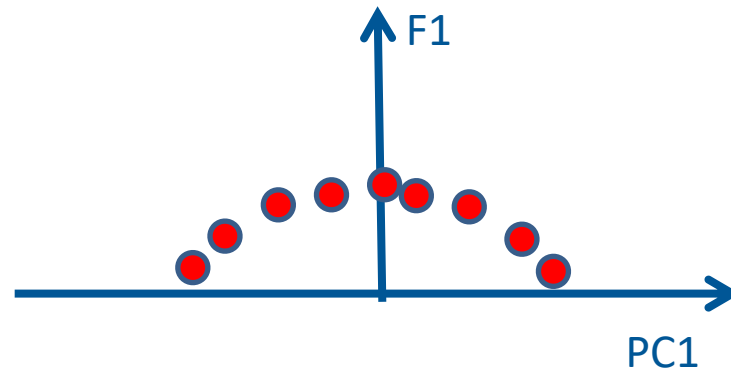


Projekcja  
na PC1  
i PC2

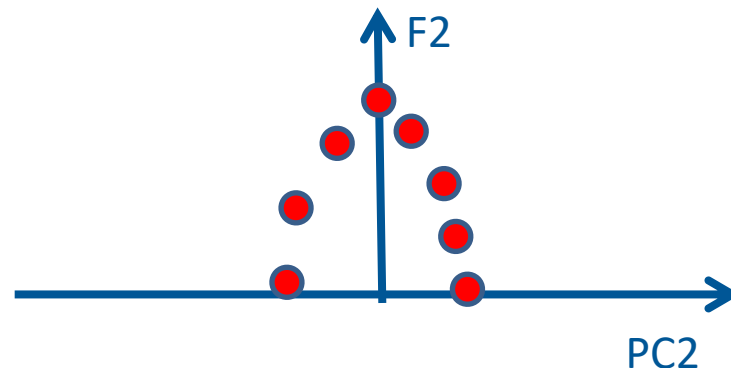


Nieskorelowane  
zmienne !!

Rozkład Gaussowski  
(średnia1, wariancja1)

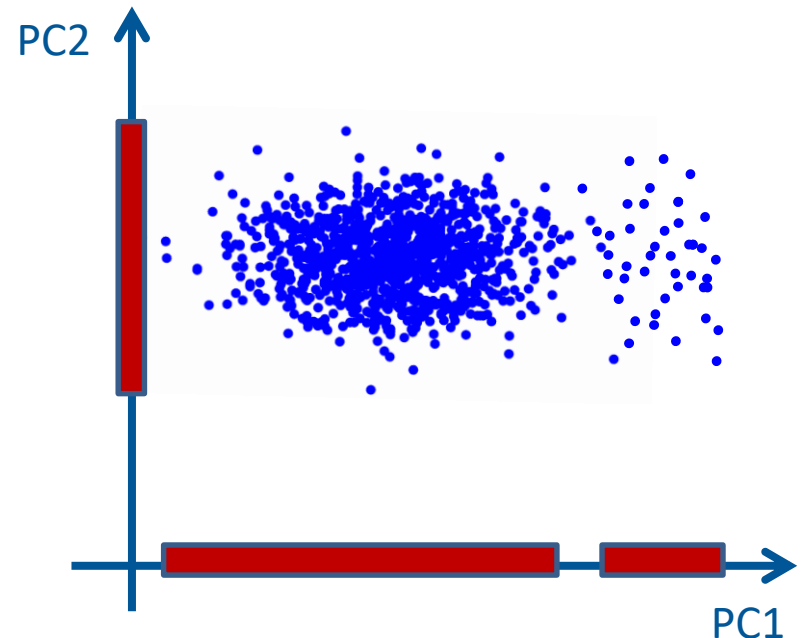
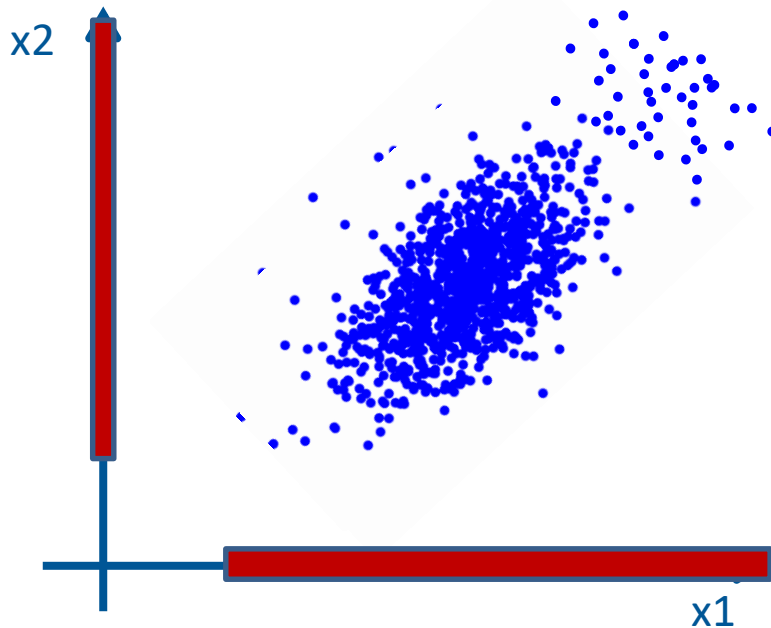


Rozkład Gaussowski  
(średnia2, wariancja2)



## Idea metody Analizy Składników Podstawowych:

Poprzez obrót przestrzeni, PCA ułatwia (a często umożliwia) wyodrębnienie/rozdzielenie grup danych (pierwszy krok do klasteryzacji, klasyfikacji itp...)

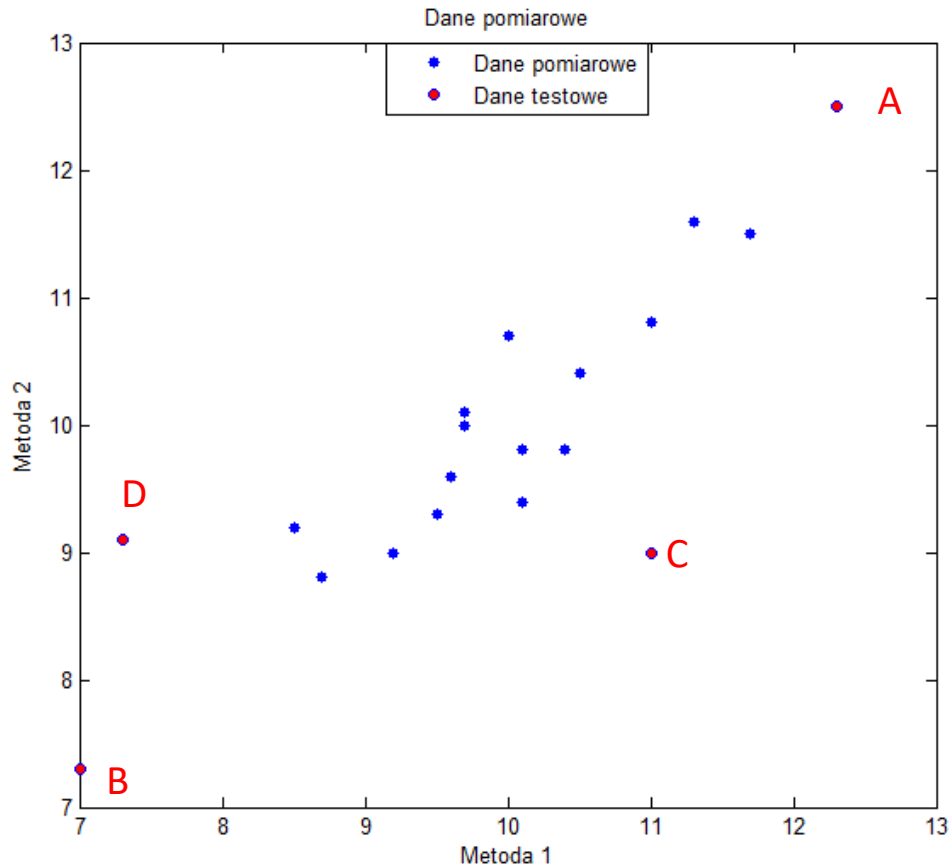




# Porównanie jedno i wielowymiarowych metod statystycznych:

## Przykład cd...

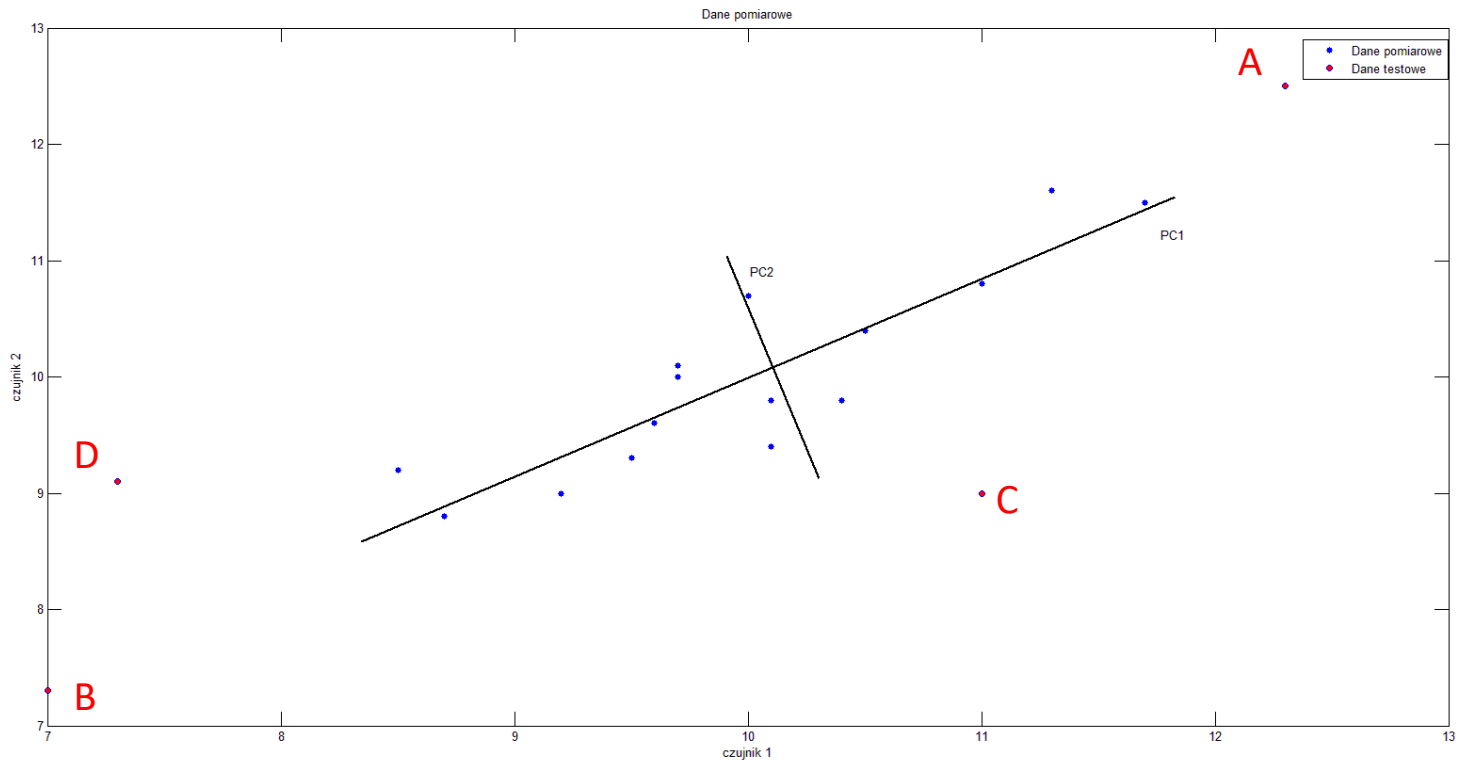
### Dane pomiarowe oraz testowe



# Porównanie jedno i wielowymiarowych metod statystycznych:

## Przykład cd...

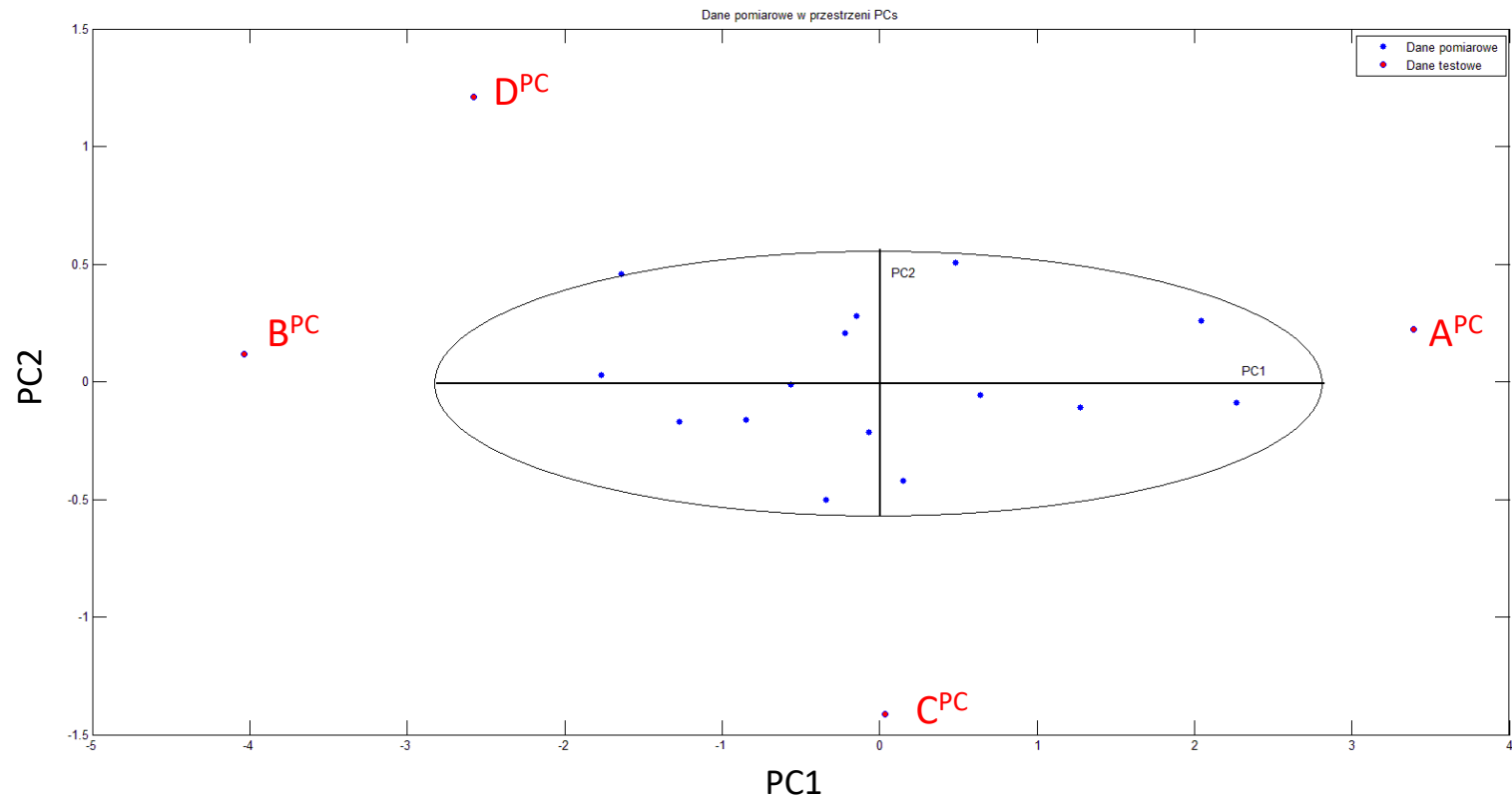
### Składniki główne procesu



# Porównanie jedno i wielowymiarowych metod statystycznych:

## Przykład cd...

Dane pomiarowe oraz testowe przetransponowane do przestrzeni składowych głównych (PCs)



# Monitoring procesu

---

## Miara Hotellinga ( $T^2$ )

Jest ważoną odwrotnością wariancji pomiarów, sumą kwadratów euklidesowej odległości od punktu przecięcia się osi PCs, przeładowanej serii pomiarów:

$$T^2 = X^{PC} \cdot \text{Variancja}^{-1} \cdot (X^{PC})^T$$

Punkty o jednakowej wartości  $T^2$  tworzą hipereliptyczne powierzchnie o promieniach proporcjonalnych do rozłożonych wzdłuż nich wariancji sygnału.

## Graniczna miara Hotellinga

$$T_{\text{lim}}^2 = \frac{M(N-1)}{N-M} F(\alpha, M, N-M)$$

$M$  – ilość (rozmiar) zmiennych pomiarowych (liczba stopni swobody licznika);

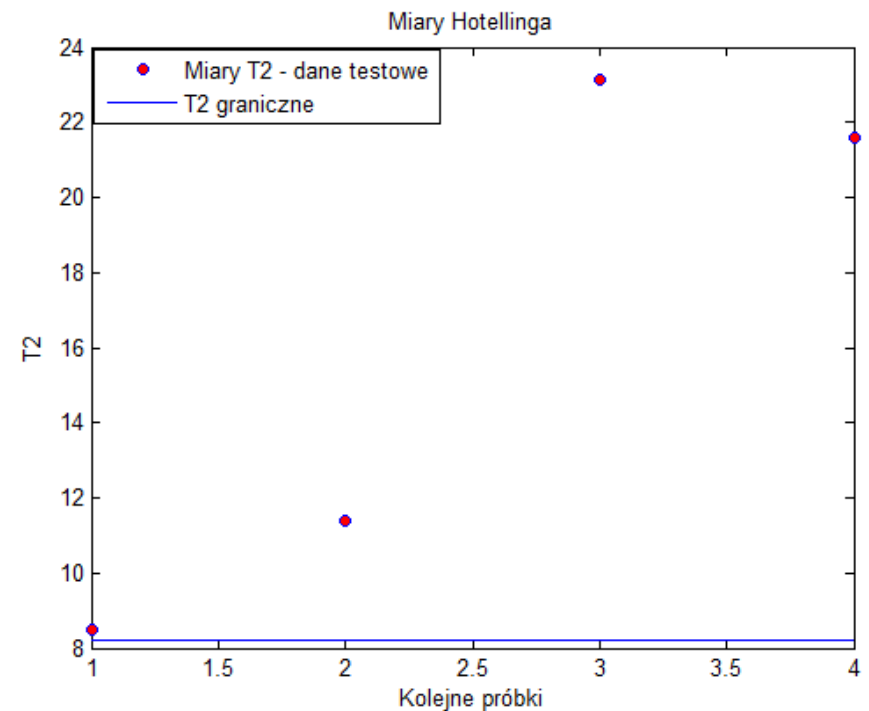
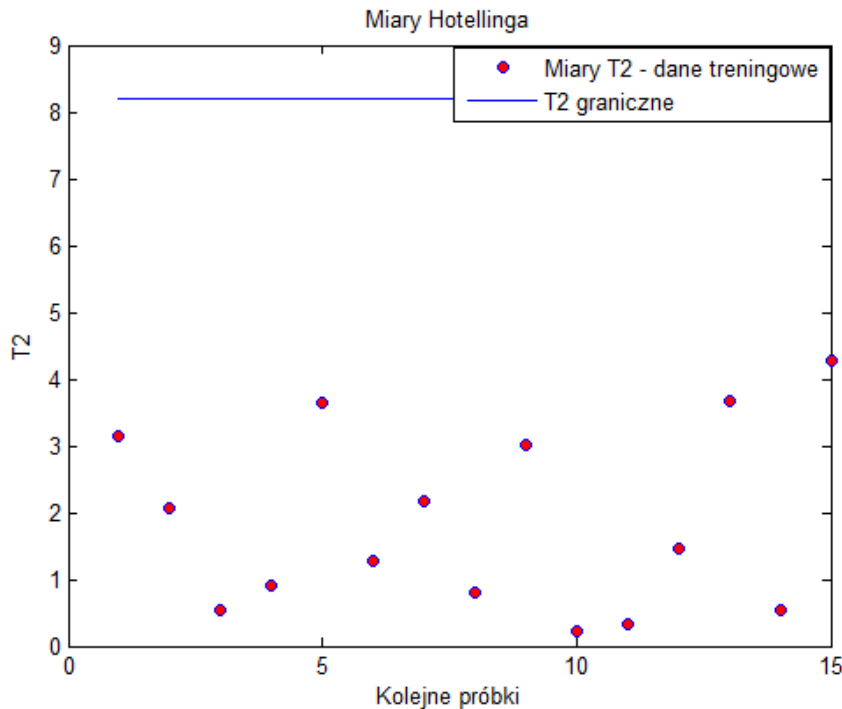
$N$  – ilość danych pomiarowych (liczba stopni swobody mianownika);

$\alpha$  – przedział ufności danych;

$F$  – funkcja rozkładu gęstości prawdopodobieństwa

# Porównanie jedno i wielowymiarowych metod statystycznych:

## Monitoring procesu przy użyciu miary Hotellinga



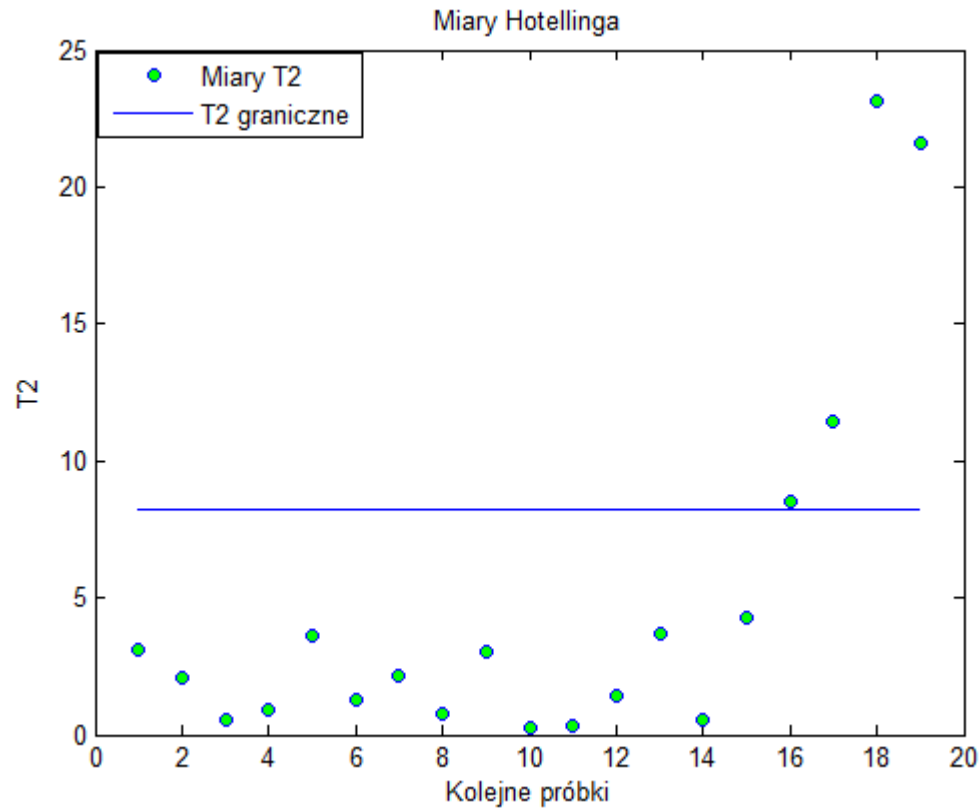
$$T_{\text{lim}}^2 = \frac{M(N-1)}{N-M} F(\alpha, M, N-M) =$$

$$= \frac{2(15-1)}{15-2} F(0.95, 2, 15-2) = 8.197$$

M – ilość (rozmiar) zmiennych pomiarowych (liczba stopni swobody licznika);  
 N – ilość danych pomiarowych (liczba stopni swobody mianownika);  
 $\alpha$  – przedział ufności danych;  
 F – funkcja rozkładu gęstości prawdopodobieństwa

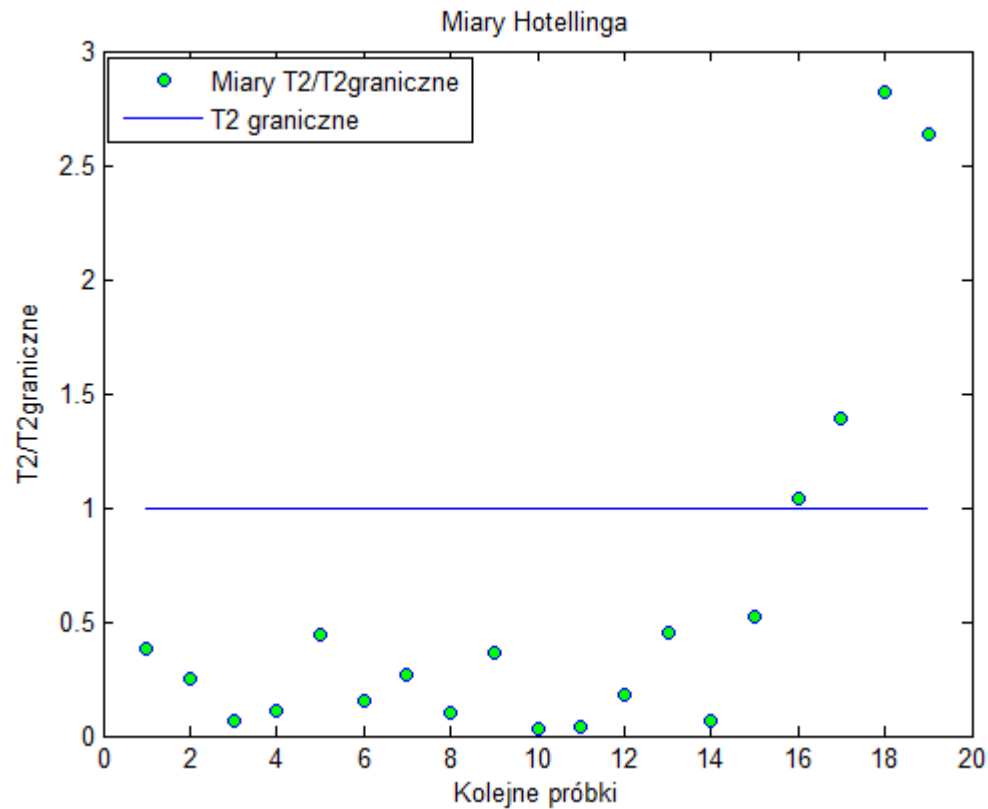
# Porównanie jedno i wielowymiarowych metod statystycznych:

## Monitoring procesu przy użyciu miary Hotellinga



# Porównanie jedno i wielowymiarowych metod statystycznych:

## Monitoring procesu przy użyciu miary Hotellinga



# Inne zastosowania PCA

## Kompresja obrazu

### Obraz oryginalny (512x512)





# Inne zastosowania PCA

## Kompresja obrazu

### Kompresja 30 %



# Inne zastosowania PCA

## Kompresja obrazu

### Kompresja 50 %



# Inne zastosowania PCA

## Kompresja obrazu

### Kompresja 75 %



# Inne zastosowania PCA

## Kompresja obrazu

### Kompresja 90 %



# Inne zastosowania PCA

## Kompresja obrazu

### Kompresja 95 %

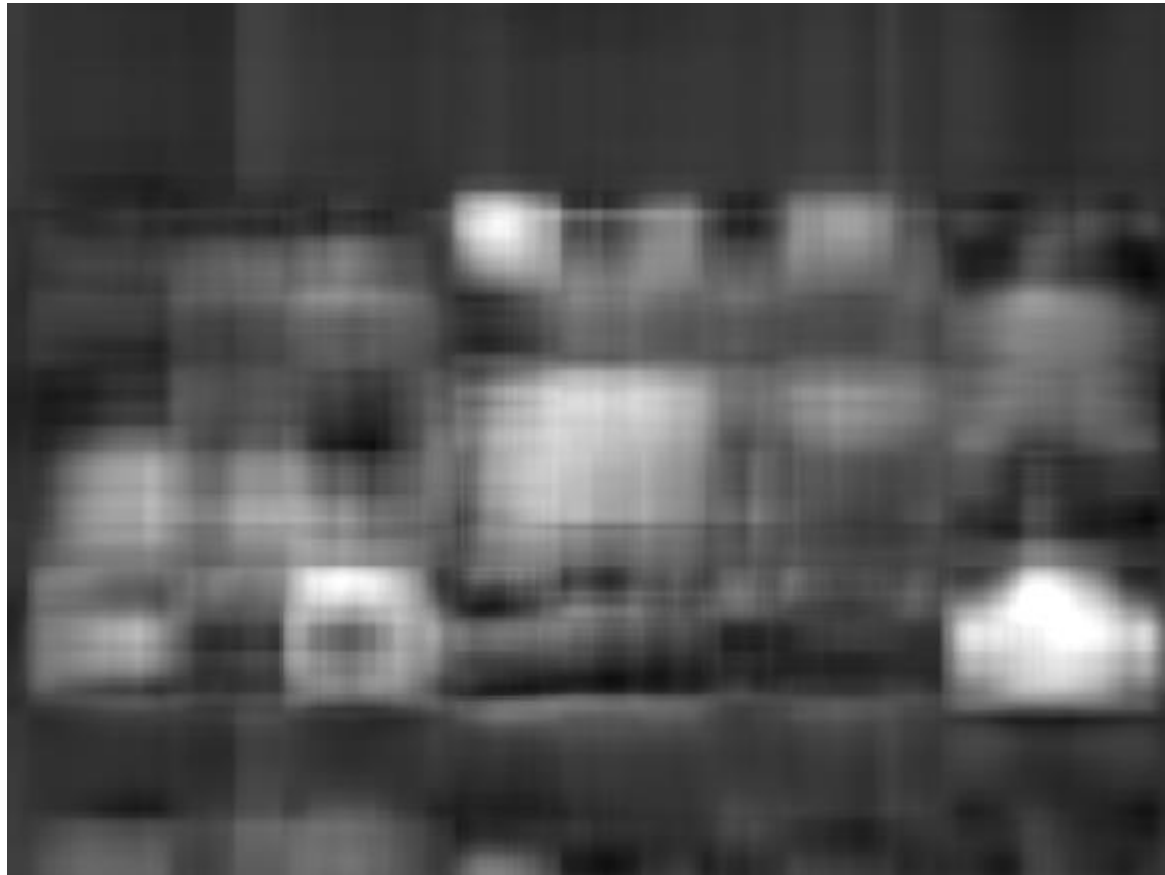


# Inne zastosowania PCA

---

## Kompresja obrazu

Kompresja 99 %



# Inne zastosowania PCA

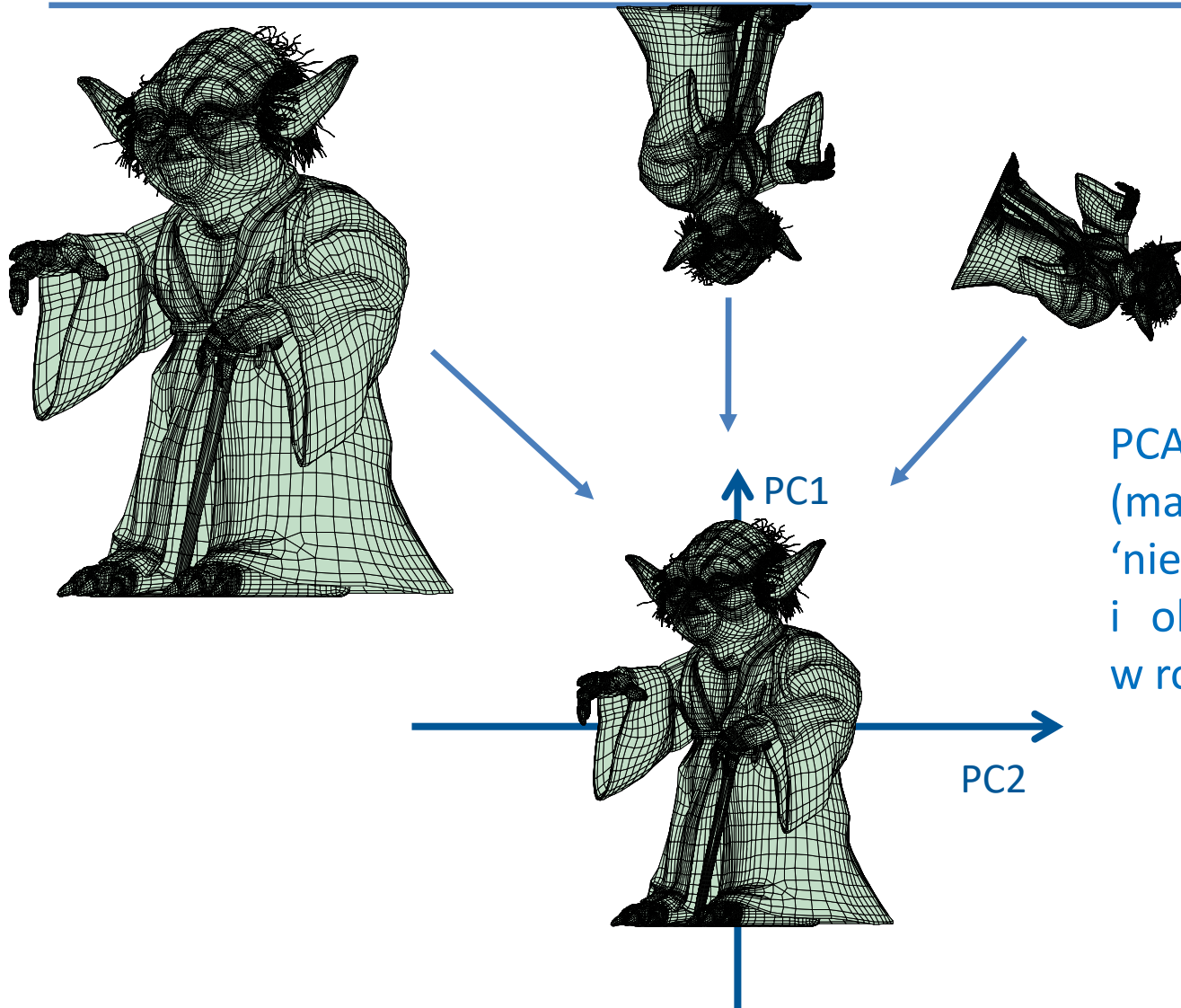
## Rozpoznawanie twarzy – Metoda Eigenfaces





## Inne zastosowania PCA

### Rozpoznawanie twarzy



PCA jest transformacją (mapowaniem) które jest 'niewrażliwe' na skalę i obrót danych => cenne w rozpoznawaniu obiektów



**Dziękuję za uwagę**