

# Principal Component Analysis - PCA

Mariusz Godlewski, Piotr Kędzierski

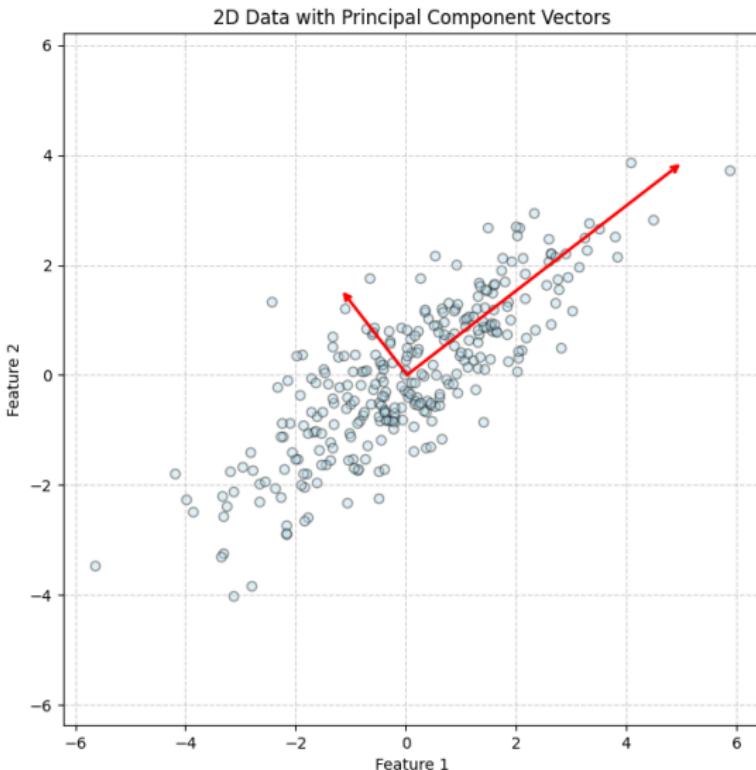
22 lutego 2026

# Konstrukcja Macierzy Danych

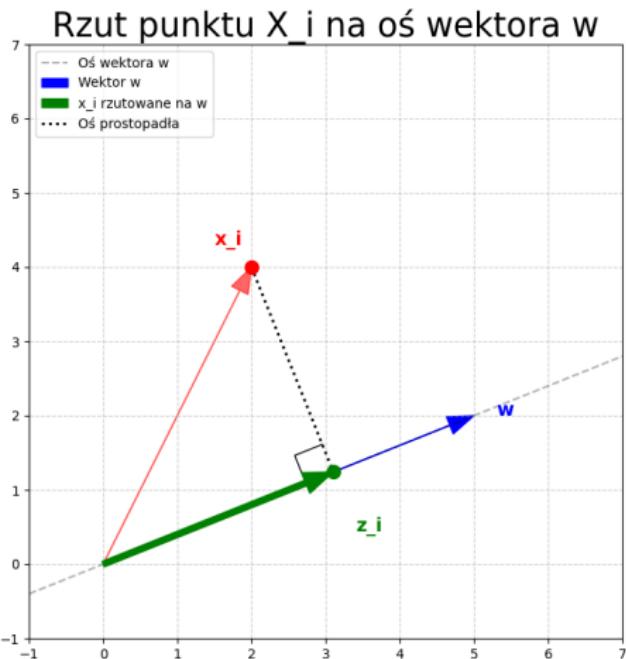
Na początku zdefiniujemy, jak wyglądają nasze dane. Wektory  $X_i$ , o wymiarach  $d \times 1$ , reprezentują obserwacje o d cechach. Natomiast  $\mathbb{X}$  ( $n \times d$ ) jest tablicą składającą się z transponowanych wektorów  $X_i$ .

$$X_i = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix} \quad \mathbb{X} = \begin{bmatrix} \quad & X_1^T & \quad \\ \quad & X_2^T & \quad \\ \vdots & & \vdots \\ \quad & X_n^T & \quad \end{bmatrix}$$

# Wizualizacja głównych składowych w 2D



# Rzut prostokątny



$$\|w\| = 1$$

$$z_i = \langle X_i, w \rangle = X_i^T w$$

# Konstrukcja macierzy kowariancji 1

$$\text{Wzór na kowariancję: } \text{cov}(X, Y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

W dalszej części zakładamy, że zmienne są wyśrodkowane względem swoich średnich ( $\bar{x} = \bar{y} = 0$ ). Wtedy w naszej macierzy  $\mathbb{X}$ , kowariancję j-tej i k-tej zmiennej możemy zapisać w postaci

$$\text{cov}(j, k) = \frac{1}{n} \sum_{i=0}^n x_{ij}x_{ik}$$

Gdzie element  $x_{ij}$  jest i-tą obserwacją j-tej zmiennej

## Konstrukcja macierzy kowariancji 2

Przyjrzyjmy się teraz macierzy

$$M = \mathbb{X}^T \mathbb{X}$$

a konkretnie elementowi znajdującejemu się w j-tym wierszu i k-tej kolumnie, który oznaczamy jako  $M_{jk}$ . Ze wzoru na mnożenie macierzy wiemy, że jest to iloczyn skalarny k-tej kolumny  $\mathbb{X}$  oraz j-tego wiersza  $\mathbb{X}^T$ , czyli j-tej kolumny  $\mathbb{X}$ . Dzięki temu otrzymujemy

$$M_{jk} = x_{1j}x_{1k} + x_{2j}x_{2k} + \cdots + x_{nj}x_{nk} = \sum_{i=0}^n x_{ij}x_{ik} = n \text{cov}(j, k)$$

Łatwo wtedy zauważyc, że macierz kowariancji  $S$ , to

$$S = \frac{1}{n} M = \frac{1}{n} \mathbb{X}^T \mathbb{X}$$

# Nasz Cel

Naszym zadaniem będzie znalezienie takiego wektora  $w$ , który zmaksymalizuje wariancję, gdy rzutujemy na niego punkty  $X_i$ . Przy tym założymy, że dane są znormalizowane (wszystkie cechy mieszczą się w przedziale  $[0, 1]$ ) i mają średnią  $\bar{X}_i = 0$ .

Dodatkowo, ustalimy długość wektora  $\|w\| = 1$ , żeby ona sama w sobie nie wpływała na otrzymywana wariancję.

# Obliczanie Wariancji 1

$$\begin{aligned}\text{wariancja} = \sigma^2 &= \frac{1}{n} \sum_{i=1}^n (z_i)^2 = \frac{1}{n} \sum_{i=1}^n (X_i^T w)^2 = \frac{1}{n} (\mathbb{X}w)^T \mathbb{X}w \\ &= w^T \left( \frac{1}{n} \mathbb{X}^T \mathbb{X} \right) w = w^T S w\end{aligned}$$

$S$  jest macierzą kowariancji. Jest ona symetryczna, dlatego

$$v_1 \perp v_2 \perp \dots \perp v_d, \quad \|v_i\| = 1$$

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d, \quad \lambda_i \in \mathbb{R}$$

Gdzie  $v_i$  są wektorami własnymi  $S$ , a  $\lambda_i$  odpowiadającymi im wartościami własnymi.

## Obliczanie Wariancji 2

Wektory  $v_i$  tworzą przestrzeń ortonormalną, dlatego możemy zapisać wektor  $w$  jako kombinację liniową:

$$w = \sum_{i=1}^d c_i v_i$$

Korzystając z warunku długości wektora  $w$ , możemy ograniczyć wartości współczynników  $c_i$

$$\|w\| = 1 \implies \left( \sum_{i=1}^d c_i v_i \right)^T \sum_{i=1}^d c_i v_i = 1 \implies \sum_{i=1}^d c_i^2 = 1$$

## Obliczanie Wariancji 3

Podstawiamy nowy wzór na  $w$ , pamiętając, że wektory własne spełniają równanie  $Sv = \lambda v$ .

$$Sw = S \left( \sum_{i=1}^d c_i v_i \right) = \sum_{i=1}^d \lambda_i c_i v_i$$

$$w^T Sw = \left( \sum_{i=1}^d c_i v_i \right)^T \left( \sum_{i=1}^d \lambda_i c_i v_i \right) = \sum_{i=1}^d \lambda_i c_i^2$$

Co daje nam nowy wzór na wariancję:

$$\sigma^2 = \sum_{i=1}^d \lambda_i c_i^2$$

# Maksymalizacja Wariancji

Przypomnijmy sobie, że naszym celem jest znalezienie takiego wektora  $w$ , który zmaksymalizuje wariancję, dlatego

$$\sup_{c_i} \sigma^2 = \sup_{c_i} \sum_{i=1}^d \lambda_i c_i^2 = \max\{\lambda_i; \quad i \in \{1, 2, \dots, d\}\}$$

Z tego wynika, że  $c_i = 1$ , gdy  $\lambda_i = \max\{\lambda\}$ , w przeciwny razie  $c_i = 0$ .

Dzięki temu, wnioskujemy, że wektor  $w$  maksymalizujący wariancję, jest wektorem własnym macierzy  $S$ , odpowiadającym największej wartości własnej.

$$w = v_1 \quad \text{oraz} \quad \sigma^2 = \lambda_1$$