

Matematyczne aspekty analizy danych

semestr zimowy 2025/2026

Dr Anna Muranova
UWM w Olsztynie

Wykład 8

Wektory w $\mathbb{R}^n$

Zbiór

$$\mathbb{R}^n = \{\bar{x} = (x_1, \dots, x_n) \mid x_i \in \mathbb{R}\}$$

z dodawaniem definiowanym dla każdych $\bar{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ oraz $\bar{y} = (y_1, \dots, y_n) \in \mathbb{R}^n$ jako:

$$\bar{x} + \bar{y} = (x_1 + y_1, \dots, x_n + y_n)$$

i mnożeniem na liczby, definiowanym jako:

$$a\bar{x} = (a \cdot x_1, \dots, a \cdot x_n) \text{ dla każdego } a \in \mathbb{R},$$

jest *przestrzenią wektorową*.

Uwaga! Nie można dodawać do siebie wektorów z różnych przestrzeni wektorowych!

Wektory: liniowa zależność i baza

Wektory $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_n$ przestrzeni V nazywamy się *liniowo niezależnymi*, gdy zachodzi wynikanie:

Jeśli $a_1\bar{v}_1 + a_2\bar{v}_2 + \dots + a_n\bar{v}_n = \bar{0}$, to $a_i = 0$ dla $i = 1, 2, \dots, n$.

Bazą przestrzeni V nazywa się liniowo niezależny zbiór wektorów $v_1, v_2, \dots, v_n$, który rozpina przestrzeń V , tzn. każdy wektor $v \in V$ może być zapisany jako $\sum_i^n a_i v_i$ dla niektórych $a_i \in \mathbb{R}$.

Na przykład, wektory

$$e_1 = (1, 0, \dots, 0)$$

$$e_2 = (0, 1, \dots, 0)$$

$$e_n = (0, 0, \dots, 1)$$

są bazą w $\mathbb{R}^n$.

Theorem

W $\mathbb{R}^n$ każde n liniowo niezależnych wektorów są bazą.

Przykład 1

Wektory $(1, 1)$ i $(-3, 2)$ z $\mathbb{R}^2$ są liniowo niezależne. Rzeczywiście, niech a_1 oraz a_2 będą takimi liczbami rzeczywistymi, że

$$(1, 1)a_1 + (-3, 2)a_2 = (0, 0).$$

Biorąc każdą współrzędną z osobna, uzyskuje się układ równań z niewiadomymi a_1 i a_2 :

$$\begin{cases} a_1 - 3a_2 = 0, \\ a_1 + 2a_2 = 0. \end{cases}$$

Jedynymi rozwiązaniami tego układu są $a_1 = 0$ i $a_2 = 0$, co potwierdza liniową niezależność wektorów.

Przykład 2

Podzbiór przestrzeni $\mathbb{R}^4$ złożony z wektorów jest liniowo zależny.

$$\begin{pmatrix} 1 \\ 4 \\ 2 \\ -3 \end{pmatrix}, \quad \begin{pmatrix} 7 \\ 10 \\ -4 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} -2 \\ 1 \\ 5 \\ -4 \end{pmatrix}$$

Należy znaleźć takie liczby rzeczywiste a_1, a_2, a_3 , nie wszystkie równe zeru, że

$$a_1 \begin{pmatrix} 1 \\ 4 \\ 2 \\ -3 \end{pmatrix} + a_2 \begin{pmatrix} 7 \\ 10 \\ -4 \\ -1 \end{pmatrix} + a_3 \begin{pmatrix} -2 \\ 1 \\ 5 \\ -4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Rozwiązując układ równań

$$\begin{cases} a_1 + 7a_2 - 2a_3 = 0, \\ 4a_1 + 10a_2 + a_3 = 0, \\ 2a_1 - 4a_2 + 5a_3 = 0, \\ -3a_1 - a_2 - 4a_3 = 0, \end{cases}$$

uzyskuje się

$$\begin{cases} a_1 = -\frac{3}{2}a_3, \\ a_2 = \frac{1}{2}a_3. \end{cases}$$

Wektory: liniowa zależność i wyznacznik

Wektory $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_n$ przestrzeni $\mathbb{R}^n$ są liniowo niezależnymi, wtedy i tylko wtedy, kiedy

$$\det(v_1, \dots, v_n) = 0$$

Np. dla wektorów $(1, 1)$ i $(-3, 2)$ z $\mathbb{R}^2$ odpowiednia macierz ma postać

$$A = \begin{pmatrix} 1 & -3 \\ 1 & 2 \end{pmatrix}.$$

Ponieważ

$$\det A = 1 \cdot 2 - 1 \cdot (-3) = 5 \neq 0,$$

więc wektory te są liniowo niezależne.

Macierz przejścia między bazami

Niech $U = \{\bar{u}_1, \dots, \bar{u}_n\}$ i $V = \{\bar{v}_1, \dots, \bar{v}_n\}$ będą dwoma bazami tej samej przestrzeni $\mathbb{R}^n$. Niech $t : \mathbb{R}^n \rightarrow \mathbb{R}^n$ będzie przekształceniem liniowym, które przekształca bazę U na bazę V , tzn. $t(\bar{u}_i) = \bar{v}_i$. Macierz T przekształcenia t w bazie U nazywamy macierzą przejścia od bazy U do bazy V .

Macierz tę wyznaczamy zatem z równości:

$$\begin{aligned}v_1 &= \alpha_1^1 u_1 + \alpha_1^2 u_2 + \dots + \alpha_1^n u_n, \\v_2 &= \alpha_2^1 u_1 + \alpha_2^2 u_2 + \dots + \alpha_2^n u_n, \\&\vdots \\v_n &= \alpha_n^1 u_1 + \alpha_n^2 u_2 + \dots + \alpha_n^n u_n.\end{aligned}$$

Czyli mamy

$$S = \begin{pmatrix} \alpha_1^1 & \alpha_1^2 & \dots & \alpha_1^n \\ \alpha_2^1 & \alpha_2^2 & \dots & \alpha_2^n \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_n^1 & \alpha_n^2 & \dots & \alpha_n^n \end{pmatrix}.$$

Przykład cz.1

W przestrzeni $\mathbb{R}^3$, znaleźć macierz przejścia z bazy U do bazy V , gdzie

$$U = \{(0, 1, 0), (-1, 0, 0), (0, 0, 2)\},$$

$$V = \{(1, 1, 1), (0, 1, 2), (0, 0, -1)\}.$$

Rozwiązanie: Zgodnie z powyższą definicją, obliczamy:

$$v_1 = (1, 1, 1) = \alpha u_1 + \beta u_2 + \gamma u_3 = \alpha(0, 1, 0) + \beta(-1, 0, 0) + \gamma(0, 0, 2).$$

Zatem dostajemy równość

$$(1, 1, 1) = (-\beta, \alpha, 2\gamma),$$

co daje nam

$$\alpha = 1, \quad \beta = -1, \quad \gamma = \frac{1}{2}.$$

Obliczone współczynniki tworzą pierwszą kolumnę szukanej macierzy przejścia S .

Przykład cz.2

Następnie:

$$v_2 = (0, 1, 2) = \alpha u_1 + \beta u_2 + \gamma u_3 = \alpha(0, 1, 0) + \beta(-1, 0, 0) + \gamma(0, 0, 2).$$

Zatem dostajemy równość

$$(0, 1, 2) = (-\beta, \alpha, 2\gamma),$$

co daje nam

$$\alpha = 1, \quad \beta = 0, \quad \gamma = 1.$$

Obliczone współczynniki tworzą drugą kolumnę szukanej macierzy przejścia S .

Przykład cz.3

I na koniec:

$$v_3 = (0, 0, -1) = \alpha u_1 + \beta u_2 + \gamma u_3 = \alpha(0, 1, 0) + \beta(-1, 0, 0) + \gamma(0, 0, 2).$$

Zatem dostajemy równość

$$(0, 0, -1) = (-\beta, \alpha, 2\gamma),$$

co daje nam

$$\alpha = 0, \quad \beta = 0, \quad \gamma = -\frac{1}{2}.$$

Obliczone współczynniki tworzą trzecią kolumnę szukanej macierzy przejścia S . Składając razem powyższe rozwiązania, dostajemy szukaną macierz przejścia z bazy U do bazy V :

$$S = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 0 \\ \frac{1}{2} & 1 & -\frac{1}{2} \end{pmatrix}.$$

Współrzędne wektora w dwóch bazach

Twierdzenie. Niech w przestrzeni $\mathbb{R}^n$ będą dane dwie bazy $U = \{u_i\}$ i $V = \{v_i\}$. Niech wektor $\bar{x} \in \mathbb{R}^n$ ma współrzędne $\{\alpha_i\}$ w bazie U oraz współrzędne $\{\beta_i\}$ w bazie V . Jeśli S jest macierzą przejścia od bazy U do bazy V , to zachodzi wzór:

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = S \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}.$$

Macierz odwrotna S^{-1} (zawsze istnieje) jest macierzą przejścia od bazy V do U , oraz zachodzi wzór:

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = S^{-1} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

- ▶ S_{UV} — macierz przejścia z bazy U do bazy V ;
- ▶ S_{VU} — macierz przejścia z bazy V do bazy U ;
- ▶ $X_U = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$ — współrzędne wektora $\bar{x}$ w bazie U ;
- ▶ $X_V = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}$ — współrzędne wektora $\bar{x}$ w bazie V ;

mamy następujące wzory:

$$X_V = S_{VU}X_U,$$

$$X_V = S_{UV}^{-1}X_U,$$

$$X_U = S_{UV}X_V,$$

$$X_U = S_{VU}^{-1}X_V.$$

Ponadto, zachodzi:

$$S_{UV} = S_{VU}^{-1},$$

$$S_{VU} = S_{UV}^{-1}.$$

Przykład

W przestrzeni $\mathbb{R}^2$ mamy dane dwie bazy:

$$U = \{(0, 1), (1, 1)\},$$

$$V = \{(2, 2), (2, 0)\}.$$

Mając współrzędne wektora $X_U = [1, 2]$ w bazie U przejść do jego współrzędnych X_V w bazie V . Zrobić to na dwa sposoby: wprost oraz licząc macierz przejścia S_{VU} . Wprost:

$$\bar{x} = 1 \cdot (0, 1) + 2 \cdot (1, 1) = (2, 3) = a(2, 2) + b(2, 0), *$$

$$\text{skąd } a = \frac{3}{2}, b = \frac{-1}{2}.$$

$$*\bar{x} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \cdot 1 + 1 \cdot 2 \\ 1 \cdot 1 + 1 \cdot 2 \end{bmatrix}.$$

Rozwiązanie

Obliczamy macierz przejścia S_{VU} :

$$u_1 = (0, 1) = \alpha v_1 + \beta v_2 = \alpha(2, 2) + \beta(2, 0) = (2\alpha + 2\beta, 2\alpha).$$

Zatem mamy równość

$$(0, 1) = (2\alpha + 2\beta, 2\alpha),$$

co daje:

$$\alpha = \frac{1}{2}, \quad \beta = -\frac{1}{2}.$$

Podobnie dla u_2 :

$$u_2 = (1, 1) = \alpha v_1 + \beta v_2 = \alpha(2, 2) + \beta(2, 0) = (2\alpha + 2\beta, 2\alpha).$$

Zatem mamy równość

$$(1, 1) = (2\alpha + 2\beta, 2\alpha),$$

co daje: $\alpha = \frac{1}{2}, \quad \beta = 0$.

Zatem

$$S_{VU} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & 0 \end{bmatrix}.$$

Ostatecznie obliczamy:

$$X_V = S_{VU}X_U = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{3}{2} \\ -\frac{1}{2} \end{bmatrix}.$$

Rozwiązanie w numpy

```
import numpy as np

xu = np.array([1,2])
u = np.array([[0,1],[1,1]])
v = np.array([[2,2],[2,0]])
x = u@xu
print(x)
#u @ xu = x = v @ xv
xv = np.linalg.inv(v)@x
print(xv)
svu = np.linalg.inv(v)@u
print(svu)
```

Rozwiązanie w sympy

```
import sympy as sp

xu = sp.Matrix([1,2])
xu.transpose()
u = sp.Matrix([[0,1],[1,1]])
v = sp.Matrix([[2,2],[2,0]])
x = u*xu
print(x)
#u * xu = x = v * xv
xv = v.inv()*x
print(xv)
svu = v.inv()*u
print(svu)
```


Przekształcenie liniowe

Przekształcenie liniowe jest to funkcja $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ między przestrzeniami liniowymi zachowująca ich działania w tym sensie, że:

- ▶ odwzorowanie sumy wektorów z jednej przestrzeni w drugą jest równe sumie odwzorowań poszczególnych wektorów tej sumy,
- ▶ odwzorowanie iloczynu wektora przez skalar jest równe iloczynowi skalara przez odwzorowanie danego wektora.

Powyższe warunki można połączyć w jeden, równoważny z nimi warunek liniowości

$$f(c\bar{x} + d\bar{y}) = c f(\bar{x}) + d f(\bar{y}).$$

Przekształcenie liniowe

Każde przekształcenie liniowe między $\mathbb{R}^n$ oraz $\mathbb{R}^m$ można przedstawić za pomocą macierzy $m \times n$. Np.: jeśli

$$(1, 0, 0) \rightarrow (1, 2), (0, 1, 0) \rightarrow (1, 1), (0, 0, 1) \rightarrow (1, 0),$$

to dla

$$\bar{x} = (2, 3, 1) = 2 \cdot (1, 0, 0) + 3 \cdot (0, 1, 0) + 1 \cdot (0, 0, 1)$$

mamy

$$f(\bar{x}) = 2 \cdot (1, 2) + 3 \cdot (1, 1) + 1 \cdot (1, 0) = (6, 7)$$

Co można zapisać jako:

$$f(\bar{x}) = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}.$$

Tzn. macierz

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \end{pmatrix}$$

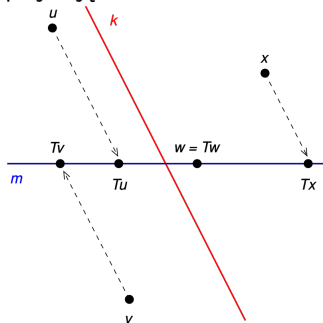
odpowiada przekształceniu f .

Rzut

Niech dana będzie przestrzeń liniowa $\mathbb{R}^n$ (nad ustalonym ciałem).
Przekształcenie liniowe $P: \mathbb{R}^n \rightarrow \mathbb{R}^n$ tej przestrzeni w siebie spełniające warunek idempotentności

$$P^2 = P,$$

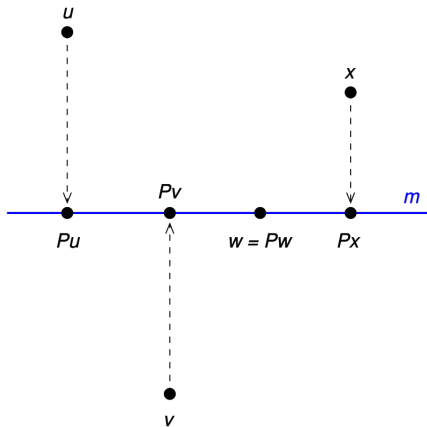
czyli $P(P(\bar{v})) = P(\bar{v})$ dla każdego $\bar{v} \in V$ nazywa się rzutem (ukośnym) lub projekcją.



Rzut T wzdłuż prostej k na prostą m .

Źródło: <https://commons.wikimedia.org/w/index.php?curid=1329674>

Rzut



Rzut ortogonalny P na prostą m .

Źródło: <https://commons.wikimedia.org/w/index.php?curid=1329670>

W tym kontekście rzut ukośny nazywa się *operatorem idempotentnym*, a rzut ortogonalny znany jest jako operator rzutowy.

Przykłady

- ▶ Przekształcenie tożsamościowe jest rzutem ortogonalnym reprezentowanym przez macierz jednostkową, np. $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ (operator jednostkowy jest operatorem rzutowym).
- ▶ Przekształcenie liniowe, którego macierz ma postać $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, jest rzutem ortogonalnym, podczas gdy zadane macierzą $\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$ jest rzutem (ukośnym), ale nie ortogonalnym (pierwsza macierz opisuje operator rzutowy, druga – tylko idempotentny).

Przykład

Ortogonalny rzut na prostą $x = y$ ma macierz

$$\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$

Wektory i wartości własne.

Niezerowy wektor v nazywamy *wektorem własnym* macierzy M , jeśli $Mv = \lambda v$ dla pewnej liczby λ . Liczbę λ nazywamy *wartością własną* macierzy M .
Mówimy też, że wektor v należy do wartości własnej λ .

Inaczej, wektor własny macierzy, to taki niezerowy wektor, który nie zmienia kierunku po przekształceniu.

Przykład 1

Znaleźć wartości i wektory własne macierzy

$$\begin{pmatrix} 6 & 3 \\ 2 & 5 \end{pmatrix}.$$

Rozwiązanie. Wartości własne:

$$\begin{pmatrix} 6 & 3 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}$$

jest równoważne układowi równań

$$\begin{cases} 6x + 3y = \lambda x, \\ 2x + 5y = \lambda y, \end{cases}$$

czyli

$$\begin{cases} (6 - \lambda)x + 3y = 0, \\ 2x + (5 - \lambda)y = 0. \end{cases}$$

Układ równań ma niezerowe rozwiązanie, o ile równania są do siebie proporcjonalne (liniowo zależne), co można ująć w postaci równania

$$0 = \begin{vmatrix} 6 - \lambda & 3 \\ 2 & 5 - \lambda \end{vmatrix} = (6 - \lambda)(5 - \lambda) - 2 \cdot 3 = \lambda^2 - 11\lambda + 24,$$

skąd $\lambda = 3$ lub $\lambda = 8$.

Przykład 1: wektory własne

Dla każdej z wartości własnych znajdziemy wektor własny (tutaj wystarczy jeden).

Dla $\lambda = 3$:

$$6x + 3y = 3x \implies 3x + 3y = 0 \implies x = -y,$$

więc wektor własny to

$$v = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Dla $\lambda = 8$:

$$6x + 3y = 8x \implies -2x + 3y = 0 \implies 2x = 3y,$$

więc wektor własny to

$$v = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

Algorytm: wartości własne

Szukamy λ taki ze $|A - \lambda I| = 0$. Szukamy wektory własne $|A - \lambda I|\vec{x} = \vec{0}$.

Znaleźć wartości i wektory własne macierzy

$$\begin{pmatrix} 5 & 4 & 4 \\ -3 & -3 & -5 \\ 1 & 2 & 4 \end{pmatrix}.$$

(na tablice, $\lambda = 1, 2, 3$)

Python

```
import numpy as np

A = np.array([[5,4, 4],[-3, -3, -5],[1,2, 4]])
print(np.linalg.eigvals(A))

import sympy as sp

A = sp.Matrix([[5,4, 4],[-3, -3, -5],[1,2, 4]])
print(A.eigenvals())
```

Algorytm: wektory własne

Szukamy λ taki ze $|A - \lambda I| = 0$. Szukamy wektory własne $|A - \lambda I|\bar{x} = \bar{b}$.

Znaleźć wartości i wektory własne macierzy

$$\begin{pmatrix} 5 & 4 & 4 \\ -3 & -3 & -5 \\ 1 & 2 & 4 \end{pmatrix}.$$

(na tablice)

```
import sympy as sp

A = sp.Matrix([[5,4, 4],[-3, -3, -5],[1,2, 4]])
y,z = sp.symbols('y,z')
X = sp.Matrix([[1],[y],[z]])
for i in 1,2,3:
    print(sp.solve([(A-i*sp.eye(3))*X],y,z,dict=True))
X = sp.Matrix([[y], [1], [z]])#uwaga
print(sp.solve([(A - 2 * sp.eye(3)) * X], y, z, dict=True))
```

Python

```
import numpy as np

A = np.array([[5,4, 4],[-3, -3, -5],[1,2, 4]])
print(np.linalg.eig(A))

import sympy as sp

A = sp.Matrix([[5,4, 4],[-3, -3, -5],[1,2, 4]])
print(A.eigenvects())
#(eigenvalue, algebraic_multiplicity, [eigenvectors])
```

Czy zawsze istnieją wektory własne?

Macierz może nie mieć wektorów własnych. Rozważmy macierz

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Aby znaleźć wartości własne, obliczamy wyznacznik:

$$0 = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1.$$

Rozwiązań możemy jednak szukać wśród liczb zespolonych:

$$\lambda = \pm i, \quad v = \begin{pmatrix} \pm i \\ 1 \end{pmatrix}.$$

Dopuszczając liczby zespolone, zawsze będziemy mieli co najmniej jeden wektor własny.

Uwagi

- ▶ W zastosowaniach wygodnie jest mieć bazę złożoną z wektorów własnych rozważanej macierzy.
- ▶ Wektory własne należące do różnych wartości własnych są liniowo niezależne. Dlatego, gdy wszystkie wartości własne są różne, istnieje baza złożona z wektorów własnych.
- ▶ W przypadku macierzy symetrycznej zawsze znajdziemy bazę złożoną z wektorów własnych.
- ▶ Czasem taka baza nie istnieje. Ale dla większości macierzy – istnieje ($\Leftrightarrow$ macierz jest diagonalizowalna $\Leftrightarrow A = BDB^{-1}$).

Rozważmy macierz $M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Dla tej macierzy mamy $M^2 = 0$, dlatego wartości własne mogą być tylko zerami (można to sprawdzić metodą zastosowaną w zadaniach). Gdyby wektory własne u i v tworzyły bazę, to dla dowolnego wektora w mielibyśmy:

$$Mw = M(au + bv) = aMu + bMv = 0.$$

Oznaczałoby to, że $M = 0$, co jest sprzeczne z definicją macierzy M .

Rozdział 6. Regresja liniowa

Regresja liniowa

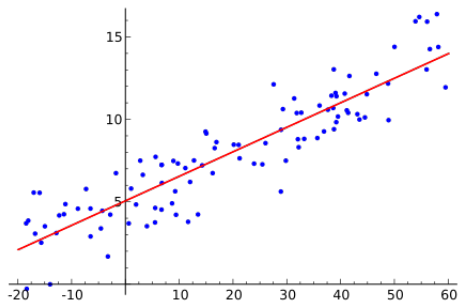
Niech mamy dwie cechy x i y . Regresja jest metodą budowania krzywej zależności Y od X , na podstawie ich wartości w próbie, $X_1, \dots, X_n$ oraz $Y_1, \dots, Y_n$. Założenie modelu:

$$Y_i = \beta_1 + \beta_2 X_i + \varepsilon_i,$$

gdzie β_1, β_2 są parametrami, a ε_i – wartość losowa błędu. Regresja liniowa – metody oparte o liniowe kombinacje zmiennych i parametrów dopasowujących model do danych. Dopasowana linia lub krzywa regresji reprezentuje oszacowaną wartość oczekiwaną zmiennej y przy konkretnych wartościach innej zmiennej lub zmiennych x . W najprostszym przypadku dopasowana jest stała lub funkcja liniowa, tzn. że regresja liniowa jest metodą wyznaczenia parametrów najlepiej dopasowanej prostej

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i, \quad \hat{\beta}_1 = ?, \hat{\beta}_2 = ?$$

Regresja liniowa



Rysunek: źródło: https://en.wikipedia.org/wiki/Linear_regression

Regresja liniowa

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i, \quad \hat{\beta}_1 = ?, \hat{\beta}_2 = ?$$

Zmienna y jest tradycyjnie nazywana zmienną *objaśnianą* lub *zależną*. Zmienne x nazywa się zmiennymi *objaśniającymi* lub *niezależnymi*. Zarówno zmienne objaśniane i objaśniające mogą być wielkościami skalarnymi lub wektorami. Niech $(X_i, Y_i), i = 1, 2, \dots, n$ — wartości doświadczalne ($x_i \neq x_j$ dla $i \neq j$). Szukamy $y = \hat{\beta}_1 + \hat{\beta}_2 x = f(x)$, która by przechodziła możliwie najbliżej wszystkich punktów doświadczalnych (X_i, Y_i) .

Metoda najmniejszych kwadratów 1

W przypadku metody najmniejszych kwadratów dopasowanie polega na minimalizacji sumy:

$$S(\beta_1, \beta_2) = \sum_{i=1}^n [Y_i - f(X_i)]^2 = \sum_{i=1}^n [Y_i - (\beta_1 + \beta_2 X_i)]^2.$$

Wiadomo, że funkcja wielu zmiennych ma minimum w punkcie, dla którego pochodne cząstkowe tej funkcji po wszystkich zmiennych są równe zero, a zatem w tym przypadku muszą być spełnione warunki

$$\begin{cases} \frac{\partial S(\beta_1, \beta_2)}{\partial \beta_1} = 0 \\ \frac{\partial S(\beta_1, \beta_2)}{\partial \beta_2} = 0 \end{cases}$$

Metoda najmniejszych kwadratów 2

$$\begin{cases} \frac{\partial S(\beta_1, \beta_2)}{\partial \beta_1} = 0 \\ \frac{\partial S(\beta_1, \beta_2)}{\partial \beta_2} = 0 \end{cases} \Leftrightarrow \begin{cases} -2 \sum_{i=1}^n (Y_i - (\hat{\beta}_1 + \hat{\beta}_2 X_i)) = 0 \\ -2 \sum_{i=1}^n (Y_i - (\hat{\beta}_1 + \hat{\beta}_2 X_i)) X_i = 0 \end{cases} \Leftrightarrow \begin{cases} \hat{\beta}_1 n + \hat{\beta}_2 \sum_{i=1}^n X_i = \sum_{i=1}^n Y_i \\ \hat{\beta}_1 \sum_{i=1}^n X_i + \hat{\beta}_2 \sum_{i=1}^n X_i^2 = \sum_{i=1}^n X_i Y_i \end{cases}$$

Metoda najmniejszych kwadratów 3

Skąd według Wzory Cramera (dla rozwiązywania układu równań przy pomocy wyznaczników) dla β_2 wynika:

$$\left\{ \begin{aligned} \hat{\beta}_2 &= \frac{n \sum_{i=1}^n X_i Y_i - \left(\sum_{i=1}^n X_i \right) \left(\sum_{i=1}^n Y_i \right)}{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2} \\ \hat{\beta}_1 &= \frac{1}{n} \left(\sum_{i=1}^n Y_i - \hat{\beta}_2 \left(\sum_{i=1}^n X_i \right) \right) \end{aligned} \right.$$

Oszacowana wartość błędu w takim przypadku wynosi:

$$\hat{\sigma} = \frac{1}{n} \left(\sum_{i=1}^n (Y_i - (\hat{\beta}_1 + \hat{\beta}_2 X_i))^2 \right)$$

Twierdzenie Gaussa-Markowa

Jeżeli próba posiada następujące właściwości model danych jest poprawnie określony, wtedy estymator najmniejszych kwadratów jest najlepszym (tj. mającym najmniejszą wariancję) estymatorem spośród liniowych, nieobciążonych estymatorów liniowego modelu regresji (ang. *best linear unbiased estimator, BLUE*).

Przykład

Dane są w tablicy. Obliczyć regresję liniową.

Pojemność RAM	Cena
2	12
4	16
8	28
16	62

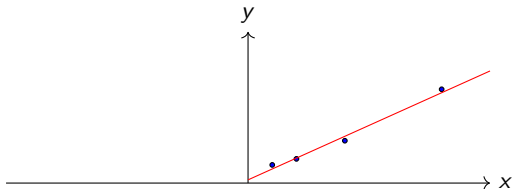
Tablica: Dane RAM

Rozwiązanie

$$\hat{\beta}_2 = \frac{n \sum_{i=1}^n X_i Y_i - \left(\sum_{i=1}^n X_i \right) \left(\sum_{i=1}^n Y_i \right)}{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2} = \frac{4(24 + 64 + 224 + 992) - 30 \cdot 118}{4(340) - 900}$$
$$= \frac{1676}{460} \approx 3.65$$

oraz

$$\hat{\beta}_1 = \frac{1}{n} \left(\sum_{i=1}^n Y_i - \hat{\beta}_2 \left(\sum_{i=1}^n X_i \right) \right) = \frac{1}{4} \left(118 - \frac{1676}{460} \cdot 30 \right) \approx 2.12$$



Ważona metoda najmniejszych kwadratów

Ta metoda (WLS), znana również jako ważona regresja liniowa, to uogólnienie zwykłych metod najmniejszych kwadratów i regresji liniowej, w których do regresji włączana jest wiedza o nierównej wariancji błędu (heteroscedastyczność), przy nieskorelowanych błędach. W tej metodzie dopasowanie liniowej regresji polega na minimalizacji sumy:

$$S(\beta_1, \beta_2) = \sum_{i=1}^n W_{ii} [Y_i - f(X_i)]^2 = \sum_{i=1}^n [W_{ii} Y_i - (\beta_1 + \beta_2 X_i)]^2,$$

gdzie W_{ii} są wagami. Probie z mniejszą wariancją muszą mieć większą wagę, nprz.

$$W_{ii} = \frac{1}{\sigma_i^2}.$$

Uogólniona metoda najmniejszych kwadratów

Uogólniona metoda najmniejszych kwadratów po raz pierwszy została opisana przez Alexandra Aitkena w 1935 roku. Ta metoda używa się, jeżeli błędy są skorelowane. W tej metodzie dopasowanie liniowej regresji polega na minimalizacji

$$S(\beta) = (Y - \beta_1 - \beta_2 X)^T \Omega^{-1} (Y - \beta_1 - \beta_2 X),$$

gdzie Ω jest macierzą wariancji błędów: $\Omega = \text{Cov}(\varepsilon|X)$.

Regresja liniowa dla kilka zmiennych niezależnych

Niech dany będzie zbiór danych zaobserwowanych $(Y_i, X_{i1}, \dots, X_{im})_{i=1}^n$. Model regresji liniowej zakłada, że istnieje liniowa (afiniczna) relacja pomiędzy zmienną zależną Y_i a wektorem X_i . Zależność ta jest modelowana przez uwzględnienie składnika losowego (błędu) ε_i , który jest zmienną losową. Dokładniej, model ten jest postaci

$$Y_i = \beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{im} + \varepsilon_i = \mathbf{X}_i^T \boldsymbol{\beta} + \varepsilon_i, \quad i = 1, \dots, n,$$

Powyższe n równań można zapisać w sposób macierzowy: $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, gdzie:

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} \mathbf{X}_1^T \\ \mathbf{X}_2^T \\ \vdots \\ \mathbf{X}_n^T \end{pmatrix} = \begin{pmatrix} 1 & X_{11} & \dots & X_{1m} \\ 1 & X_{21} & \dots & X_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & \dots & X_{nm} \end{pmatrix},$$

oraz

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix}, \quad \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

Regresja liniowa dla kilka zmiennych niezależnych

Najczęściej do obliczenia $\hat{\beta}$ wykorzystuje się, jak i w przypadku jednej zmiennej niezależnej, klasyczna metoda najmniejszych kwadratów i jej pochodne.

Dla klasycznej metody najmniejszych kwadratów:

$$\hat{\beta} = (X^T X)^{-1} X^T Y.$$

Kwartet Anscombe'a

Kwartet Anscombe'a to cztery zestawy danych o identycznych cechach statystycznych, takich jak średnia arytmetyczna, wariancja, współczynnik korelacji czy równanie regresji liniowej, jednocześnie wyglądających zgoła różnie przy przedstawieniu graficznym. Układ tych danych został stworzony w 1973 roku przez brytyjskiego statystyka Francisca Anscombe'a aby ukazać znaczenie graficznej reprezentacji danych przy okazji ich analizy statystycznej.

I		II		III		IV	
X	Y	X	Y	X	Y	X	Y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

Tablica: Kwartet Anscombe'a

Kwartet Anscombe'a

