

Metody Numeryczne, część 4 studia stacjonarne

Aproksymacja w sensie najmniejszych kwadratów

Krystyna Ziętak

30 październik 2018

Spis treści

- 1 Aproksymacja - porównanie z interpolacją
- 2 Aproksymacja w sensie najmniejszych kwadratów
- 3 Charakteryzacja najlepszej aproksymacji
- 4 Przykłady aproksymacji w sensie najmniejszych kwadratów
- 5 Interpretacja układu równań normalnych
 - Funkcje wielu zmiennych - pochodne cząstkowe
 - Inne sformułowanie twierdzenia o wielomianie optymalnym
- 6 Podsumowanie

Interpolacja - przypomnienie

- Wielomian interpolacyjny

$$w_n(a) = a_0x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n$$

stopnia $\leq n$ przyjmuje zadane z góry wartości y_j w $n + 1$ różnych węzłach interpolacji x_j ($j = 0, \dots, n$):

$$w_n(x_j) = y_j, \quad j = 0, \dots, n.$$

- Liczba współczynników, które definiują wielomian interpolacyjny, jest taka sama jak liczba zadanych warunków, które musi on spełniać w węzłach interpolacji.
- Uwaga.** W ogólnym przypadku nie istnieje wielomian $w_n(x)$ stopnia $\leq n$ spełniający więcej niż $n + 1$ warunków $w_n(x_j) = y_j$.

Przykład 1

Czy istnieje wielomian $w_1(x) = a_0x + a_1$ taki, że

$$w_1(-1) = 5, \quad w_1(0) = 6, \quad w_1(1) = 7 ?$$

Trzeba rozwiązać układ

$$-a_0 + a_1 = 5, \quad a_1 = 6, \quad a_0 + a_1 = 7$$

Rozwiązanie: $a_0 = 1, a_1 = 6$ Odp. TAK

Przykład 2

Czy istnieje wielomian $w_1(x) = a_0x + a_1$ taki, że

$$w_1(-1) = 4, \quad w_1(0) = 6, \quad w_1(1) = 7 ?$$

Trzeba rozwiązać układ

$$-a_0 + a_1 = 4, \quad a_1 = 6, \quad a_0 + a_1 = 7$$

układ sprzeczny: $a_0 = 2, \quad a_0 = 1, \quad \text{Odp. NIE}$

Aproksymacja

- Niech danych będzie m różnych liczb x_1, \dots, x_m oraz m liczb y_1, \dots, y_m .
- Niech n liczba naturalna taka, że $m \geq n + 1$.
- Szukamy wielomianu $w_n(x)$ stopnia $\leq n$ takiego, żeby jego wartości dla argumentów x_1, \dots, x_m były możliwie bliskie zadanym wartościom y_1, \dots, y_m .
- Jeśli $m = n + 1$, to jest to interpolacja. Dla $m = n + 1$ istnieje dokładnie jeden wielomian $w_n(x)$ stopnia $\leq n$ taki, że $w_n(x_j) = y_j$ dla $j = 1, \dots, m = n + 1$.
- Wartości y_j mogą być interpretowane jako wartości - dokładne lub obarczone błędem - pewnej funkcji $f(x)$ dla argumentów x_j .

Jakie wybrać kryterium “bliskości” ?

To jest dyskretna aproksymacja, bo wykorzystujemy wartości funkcji tylko dla $x = x_j, j = 1, \dots, m$.

- **Dyskretna aproksymacja w sensie Czebyszewa**
Maksymalna różnica między wartościami $w_n(x_j)$ szukanego wielomianu $w_n(x)$ dla $x = x_j$ i zadanymi liczbami y_j

$$\max_{j \in \{1, \dots, m\}} |w_n(x_j) - y_j|$$

ma być minimalna.

- **Dyskretna aproksymacja w sensie najmniejszych kwadratów**
Suma kwadratów różnic między wartościami $w_n(x_j)$ szukanego wielomianu $w_n(x)$ dla $x = x_j$ i zadanymi liczbami y_j

$$\sum_{j=1}^m (w_n(x_j) - y_j)^2$$

ma być minimalna. **Uwaga.** *To jest celem wykładu.*

Aproksymacja wielomianami w sensie najmniejszych kwadratów

Dane są różne liczby

$$x_1, \dots, x_m$$

oraz liczby

$$y_1, \dots, y_m.$$

Dana jest liczba naturalna $n \leq m - 1$. Szukamy wielomianu

$$w_n(x) = a_0x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n,$$

dla którego wyrażenie

$$\sum_{j=1}^m (w_n(x_j) - y_j)^2$$

przyjmuje minimalną wartość.

- Wielomian ten aproksymuje w sensie najmniejszych kwadratów funkcję $f(x)$ przyjmującą wartości y_j dla $x = x_j$ dla $j = 1, \dots, m$.
- Mówimy, że funkcja $f(x)$ jest aproksymowana w sensie najmniejszych kwadratów na zbiorze $\{x_1, \dots, x_m\}$.
- To jest dyskretna aproksymacja, bo wykorzystujemy wartości funkcji tylko dla $x = x_j, j = 1, \dots, m$.

Dyskretną aproksymację w sensie najmniejszych kwadratów za pomocą wielomianów nazywa się w polskiej literaturze często metodą najmniejszych kwadratów.

Funkcja $n + 1$ zmiennych:

$$F(a_0, \dots, a_n) = \sum_{j=1}^m (w_n(x_j) - y_j)^2 = \sum_{j=1}^m \left(\sum_{k=0}^n a_k x_j^{n-k} - y_j \right)^2$$

Dane: $n, m, x_1, \dots, x_m, y_1, \dots, y_m$.

Wynik: a_0, \dots, a_n .

Szukamy współczynników a_0, \dots, a_n , dla których funkcja $F(a_0, \dots, a_n)$ przyjmuje najmniejszą wartość.

Oznaczenie: $w_n(x) = a_0x^n + \dots + a_{n-1}x + a_n$

$$F(a_0, \dots, a_n) = \sum_{j=1}^m (w_n(x_j) - y_j)^2$$

$$F(a_0, \dots, a_n) = \sum_{j=1}^m \left(\sum_{k=0}^n a_k x_j^{n-k} - y_j \right)^2$$

Jest to funkcja wielu zmiennych. Zmiennymi są współczynniki a_0, \dots, a_n szukanego wielomianu aproksymującego. Liczby x_j, y_j są dane (są ustalone). Liczby te określają m punktów (x_j, y_j) na płaszczyźnie.

Przykład

$$m = 3, \quad n = 1$$

$$F(a_0, a_1) = (a_0x_1 + a_1 - y_1)^2 + (a_0x_2 + a_1 - y_2)^2 + (a_0x_3 + a_1 - y_3)^2$$

Twierdzenie. Niech dane będą różne liczby x_1, \dots, x_m oraz liczby y_1, \dots, y_m , $m \geq n + 1$. Wówczas istnieje dokładnie jeden (optymalny) wielomian $w_n(x)$ stopnia $\leq n$

$$w_n(x) = a_0x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n,$$

dla którego wartość funkcji

$$F(a_0, \dots, a_n) = \sum_{j=1}^m (w_n(x_j) - y_j)^2$$

jest minimalna. Współczynniki a_0, \dots, a_n tego optymalnego wielomianu są rozwiązaniem układu równań liniowych:

$$\sum_{j=1}^m x_j^{n-i} \left(\sum_{k=0}^n a_k x_j^{n-k} - y_j \right) = 0 \quad \text{dla } i = 0, \dots, n$$

$$\sum_{j=1}^m x_j^{n-i} \left(\sum_{k=0}^n a_k x_j^{n-k} - y_j \right) = 0 \quad \text{dla } i = 0, \dots, n$$

Ten układ $n + 1$ równań liniowych nazywamy
układem równań normalnych.

Przykład

- Dane: $m = 3, n = 1$,
 $x_1 = -1, x_2 = 3, x_3 = 5$
 $y_1 = -1, y_2 = 0, y_3 = 4$
- Niech $w_1(x) = a_0x + a_1$ będzie wielomianem stopnia ≤ 1 .
Szukamy współczynników a_0, a_1 takich, żeby wyrażenie

$$(a_0x_1 + a_1 - y_1)^2 + (a_0x_2 + a_1 - y_2)^2 + (a_0x_3 + a_1 - y_3)^2$$

było minimalne.

Oznaczenie

$$F(a_0, a_1) = (a_0x_1 + a_1 - y_1)^2 + (a_0x_2 + a_1 - y_2)^2 + (a_0x_3 + a_1 - y_3)^2$$

czyli dla naszych danych mamy

$$F(a_0, a_1) = ((-1)a_0 + a_1 + 1)^2 + (3a_0 + a_1)^2 + (5a_0 + a_1 - 4)^2$$

Musimy wyznaczyć wartości a_0 i a_1 , dla który jest osiągnięte minimum funkcji $F(a_0, a_1)$ dwóch zmiennych a_0 i a_1 .

$$F(a_0, a_1) = ((-1)a_0 + a_1 + 1)^2 + (3a_0 + a_1)^2 + (5a_0 + a_1 - 4)^2$$

Rozwiązanie - minimum jest osiągnięte dla

$$a_0 = \frac{3}{4}, \quad a_1 = -\frac{3}{4}$$

$$w_1(x) = \frac{3}{4}x - \frac{3}{4}$$

Interpretacja geometryczna wielomianu stopnia 1:
prosta

$$y = \frac{3}{4}x - \frac{3}{4}$$

Uzasadnienie odpowiedzi

Dla $m = 3, n = 1$ układ równań normalnych, zdefiniowany w twierdzeniu, ma następującą postać:

$$x_1(a_0x_1 + a_1 - y_1) + x_2(a_0x_2 + a_1 - y_2) + x_3(a_0x_3 + a_1 - y_3) = 0$$

$$(a_0x_1 + a_1 - y_1) + (a_0x_2 + a_1 - y_2) + (a_0x_3 + a_1 - y_3) = 0$$

Dla danych: $x_1 = -1, x_2 = 3, x_3 = 5$

$y_1 = -1, y_2 = 0, y_3 = 4$ otrzymujemy układ

$$-(-a_0 + a_1 + 1) + 3(3a_0 + a_1) + 5(5a_0 + a_1 - 4) = 0$$

$$(-a_0 + a_1 + 1) + (3a_0 + a_1) + (5a_0 + a_1 - 4) = 0$$

Zatem mamy do rozwiązania układ

$$-(-a_0 + a_1 + 1) + 3(3a_0 + a_1) + 5(5a_0 + a_1 - 4) = 0$$

$$(-a_0 + a_1 + 1) + (3a_0 + a_1) + (5a_0 + a_1 - 4) = 0$$

czyli układ

$$35a_0 + 7a_1 - 21 = 0, \quad 7a_0 + 3a_1 - 3 = 0$$

czyli układ

$$5a_0 + a_1 = 3, \quad 7a_0 + 3a_1 = 3$$

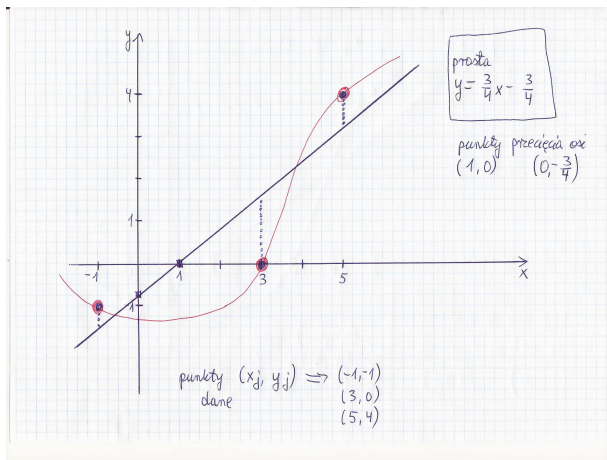
Układ równań liniowych

$$5a_0 + a_1 = 3, \quad 7a_0 + 3a_1 = 3$$

ma rozwiązanie

$$a_0 = \frac{3}{4}, \quad a_1 = -\frac{3}{4}$$

Graficzna ilustracja



Przypomnij sobie, jak znajduje się punkty przecięcia prostej z osiami układu współrzędnych!

Przypadek dowolnego m i $n = 1$, $m \geq n + 1$

$$w_1(x) = a_0x + a_1$$

$$F(a_0, a_1) = \sum_{j=1}^m (w_1(x_j) - y_j)^2 = \sum_{j=1}^m (a_0x_j + a_1 - y_j)^2$$

Układ dwóch równań normalnych ma następującą postać

$$\sum_{j=1}^m x_j (a_0x_j + a_1 - y_j) = 0$$

$$\sum_{j=1}^m (a_0x_j + a_1 - y_j) = 0$$

Uwaga. Niewiadomymi w tym układzie są a_0 i a_1 .

Musimy rozwiązać układ dwóch równań liniowych z dwiema niewiadomymi a_0 i a_1

$$\sum_{j=1}^m x_j (a_0 x_j + a_1 - y_j) = 0, \quad \sum_{j=1}^m (a_0 x_j + a_1 - y_j) = 0$$

Po przekształceniu ten układ ma postać

$$\left(\sum_{j=1}^m x_j^2 \right) a_0 + \left(\sum_{j=1}^m x_j \right) a_1 = \sum_{j=1}^m (x_j y_j)$$

$$\left(\sum_{j=1}^m x_j \right) a_0 + m a_1 = \sum_{j=1}^m y_j$$

Własności sumowania

$$\sum_{j=1}^m (b_j + c_j) = \sum_{j=1}^m b_j + \sum_{j=1}^m c_j$$

$$\sum_{j=1}^m (d \times b_j) = d \sum_{j=1}^m b_j$$

$$\sum_{j=1}^m d = m \times d$$

W poprzednim przykładzie mieliśmy $m = 3, n = 1$
 $x_1 = -1, x_2 = 3, x_3 = 5 \quad y_1 = -1, y_2 = 0, y_3 = 4$

$$\sum_{j=1}^m x_j^2 = 1 + 9 + 25 = 35$$

$$\sum_{j=1}^m x_j = -1 + 3 + 5 = 7$$

$$\sum_{j=1}^m (x_j y_j) = (-1)(-1) + 3 \cdot 0 + 5 \cdot 4 = 21$$

$$\sum_{j=1}^m y_j = -1 + 0 + 4 = 3$$

Wobec tego układ

$$\left(\sum_{j=1}^m x_j^2\right) a_0 + \left(\sum_{j=1}^m x_j\right) a_1 = \sum_{j=1}^m (x_j y_j)$$
$$\left(\sum_{j=1}^m x_j\right) a_0 + m a_1 = \sum_{j=1}^m y_j$$

ma następującą postać

$$35a_0 + 7a_1 = 21, \quad 7a_0 + 3a_1 = 3$$

porównaj poprzednie sleidy z tym przykładem !!!

Przykład dla $m = 4, n = 1$

$x_1 = -1$	$x_2 = 1$	$x_3 = 4$	$x_4 = 5$
$y_1 = 3$	$y_2 = 1$	$y_3 = 4$	$y_4 = 0$

$$\sum_{j=1}^m x_j^2 = 43,$$

$$\sum_{j=1}^m x_j = 9,$$

$$\sum_{j=1}^m (x_j y_j) = 14,$$

$$\sum_{j=1}^m y_j = 8$$

układ równań

$$43a_0 + 9a_1 = 14, \quad 9a_0 + 4a_1 = 8$$

Przykład dla $m = 5, n = 1$, $y_j = \sin(x_j)$

$x_1 = -\pi/2$	$x_2 = -\pi/4$	$x_3 = 0$	$x_4 = \pi/4$	$x_5 = \pi/2$
$y_1 = -1$	$y_2 = -1/\sqrt{2}$	$y_3 = 0$	$y_4 = 1/\sqrt{2}$	$y_5 = 1$

$$\sum_{j=1}^m x_j^2 = \frac{5}{8}\pi^2, \quad \sum_{j=1}^m x_j = 0,$$

$$\sum_{j=1}^m (x_j y_j) = \left(1 + \frac{1}{2\sqrt{2}}\right) \pi, \quad \sum_{j=1}^m y_j = 0$$

układ równań

$$\frac{5}{8}\pi^2 a_0 + 0 \cdot a_1 = \left(1 + \frac{1}{2\sqrt{2}}\right) \pi, \quad 0 \cdot a_0 + 5a_1 = 0$$

Rozwiązanie: $a_0 = \frac{8\left(1 + \frac{1}{2\sqrt{2}}\right)}{5\pi}, \quad a_1 = 0$

Odpowiedź ($n = 1$)

Wielomian optymalny stopnia 1

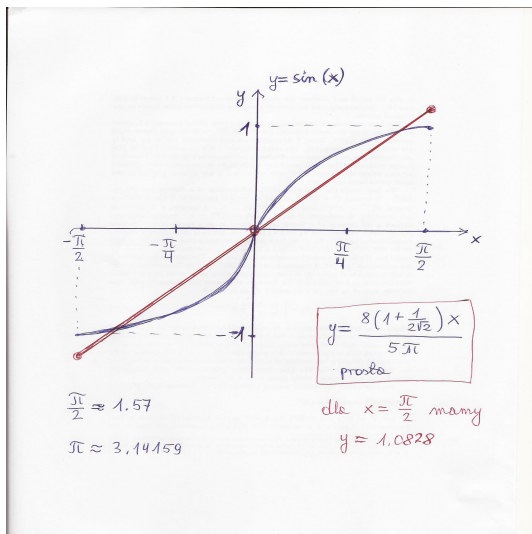
$$w_1(x) = \frac{8 \left(1 + \frac{1}{2\sqrt{2}}\right)}{5\pi} x$$

$$\frac{8 \left(1 + \frac{1}{2\sqrt{2}}\right)}{5\pi} \approx 0.68936$$

Dla $n = 2$ ten sam wielomian jest optymalny

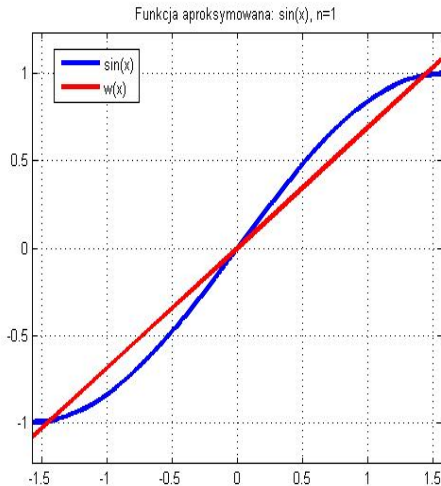
"ręczny" wykres sinusa na przedziale $[-\pi/2, \pi/2]$

i prostej $y = \frac{8(1+\frac{1}{2\sqrt{2}})}{5\pi} x$



wykresy sinusa na przedziale $[-\pi/2, \pi/2]$

i prostej $y = \frac{8(1+\frac{1}{2\sqrt{2}})}{5\pi} x$



Przykład dla $m = 5, n = 3$, $y_j = \sin(x_j)$

$x_1 = -\pi/2$	$x_2 = -\pi/4$	$x_3 = 0$	$x_4 = \pi/4$	$x_5 = \pi/2$
$y_1 = -1$	$y_2 = -1/\sqrt{2}$	$y_3 = 0$	$y_4 = 1/\sqrt{2}$	$y_5 = 1$

Wielomian optymalny stopnia 3

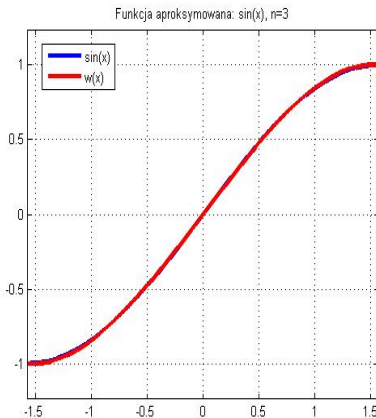
$$w_3(x) = -0.142496x^3 + (0.61006 \times 10^{-16})x^2 + 0.988215x + 0$$

Szereg potęgowy (Taylora)

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \dots$$

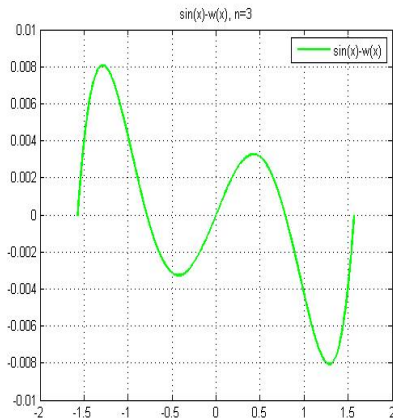
$$1/3! = 1/6 = 0.166\dots$$

wykresy sinusa i wielomianu optymalnego stopnia 3



Dokładność jest tak dobra, że wykresy w praktyce pokryły się.

wykres błędu aproksymacji sinusa przez wielomian stopnia 3,
czyli różnicy $\sin(x) - w_3(x)$



$$|\sin(x) - w_3(x)| \leq 0.01$$

- Jaka jest interpretacja układu równań normalnych podanych w twierdzeniu?
- Co ten układ ma wspólnego z warunkami na minimum funkcji $F(a_0, \dots, a_n)$ wielu zmiennych?
- Porównaj z warunkami na minimum funkcji jednej zmiennej!

Przykłady funkcji wielu zmiennych

$$f(x, y) = 2xy + \sin(x) \cos(xy) + x^2 + 5y^3$$

$$f(x, y, z) = 4xy - zy + z - xyz$$

$$f(x, z, y) = (2x - 3y + z)^2 + (-3x + 2y + 5z)^2 + (x + y + z)^2$$

- Pochodną rzędu pierwszego funkcji jednej zmiennej $f(x)$ w punkcie $x = x_0$ nazywamy granicę:

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

Pochodne cząstkowe

- Niech f będzie funkcją dwóch zmiennych: $f(x, y)$.
Pochodną cząstkową rzędu pierwszego funkcji f względem zmiennej x w punkcie (x_0, y_0) nazywamy granicę

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h}$$

Pochodną cząstkową rzędu pierwszego funkcji f względem zmiennej y w punkcie (x_0, y_0) nazywamy granicę

$$\lim_{h \rightarrow 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h}$$

Przykłady

funkcja f jednej zmiennej x

$$f(x) = (3x + 4)^2 = 9x^2 + 24x + 16$$

$$f'(x) = 2 \cdot 3 \cdot (3x + 4) = 18x + 24$$

funkcja f dwóch zmiennych x, y

$$f(x, y) = (3x + 4)^2 + (5y - 3)^2 = 9x^2 + 24x + 16 + 25y^2 - 30y + 9$$

pochodne cząstkowe

$$\frac{\partial f}{\partial x}(x, y) = 2 \cdot 3 \cdot (3x + 4) = 18x + 24$$

$$\frac{\partial f}{\partial y}(x, y) = 2 \cdot 5 \cdot (5y - 3) = 50y - 30$$

Twierdzenie. Niech dane będą różne liczby x_1, \dots, x_m oraz liczby y_1, \dots, y_m , $m \geq n + 1$. Wówczas istnieje dokładnie jeden wielomian $w_n(x)$ stopnia $\leq n$

$w_n(x) = a_0x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n = \sum_{k=0}^n a_k x^{n-k}$,
dla którego wartość wyrażenia

$$F(a_0, \dots, a_n) = \sum_{j=1}^m (w_n(x_j) - y_j)^2 = \sum_{j=1}^m \left(\sum_{k=0}^n a_k x_j^{n-k} - y_j \right)^2$$

jest minimalna. Współczynniki tego optymalnego wielomianu są rozwiązaniem następującego układu równań liniowych

$$\frac{\partial F}{\partial a_i} = 0, \quad \text{dla } i = 0, \dots, n$$

To jest układ równań normalnych podany w poprzednim sformułowaniu tego twierdzenia z dokładnością do czynnika 2.

Przykład dyskretnej aproksymacji średniokwadratowej wielomianami

- Dane: $m = 3, n = 1$,
 $x_1 = -1, x_2 = 3, x_3 = 5$
 $y_1 = -1, y_2 = 0, y_3 = 4$
- Niech $w(x) = a_0x + a_1$. Szukamy współczynników a_0, a_1 takich, żeby wyrażenie

$$(a_0x_1 + a_1 - y_1)^2 + (a_0x_2 + a_1 - y_2)^2 + (a_0x_3 + a_1 - y_3)^2$$

było minimalne.

Oznaczenie

$$F(a_0, a_1) = (a_0x_1 + a_1 - y_1)^2 + (a_0x_2 + a_1 - y_2)^2 + (a_0x_3 + a_1 - y_3)^2$$

czyli dla naszych danych mamy

$$F(a_0, a_1) = ((-1)a_0 + a_1 + 1)^2 + (3a_0 + a_1)^2 + (5a_0 + a_1 - 4)^2$$

Zastosowanie drugiego sformułowania twierdzenia

$$F(a_0, a_1) = (a_0x_1 + a_1 - y_1)^2 + (a_0x_2 + a_1 - y_2)^2 + (a_0x_3 + a_1 - y_3)^2$$

$$\frac{\partial F}{\partial a_0} = 2x_1(a_0x_1 + a_1 - y_1) + 2x_2(a_0x_2 + a_1 - y_2) + 2x_3(a_0x_3 + a_1 - y_3)$$

$$\frac{\partial F}{\partial a_1} = 2(a_0x_1 + a_1 - y_1) + 2(a_0x_2 + a_1 - y_2) + 2(a_0x_3 + a_1 - y_3)$$

Trzeba rozwiązać układ równań

$$\frac{\partial F}{\partial a_0} = 0, \quad \frac{\partial F}{\partial a_1} = 0$$

niewiadome: a_0 i a_1

Dane: $x_1 = -1$, $x_2 = 3$, $x_3 = 5$ $y_1 = -1$, $y_2 = 0$, $y_3 = 4$

$$\frac{\partial F}{\partial a_0} = -2(-a_0 + a_1 + 1) + 6(3a_0 + a_1) + 10(5a_0 + a_1 - 4)$$

$$\frac{\partial F}{\partial a_1} = 2(-a_0 + a_1 + 1) + 2(3a_0 + a_1) + 2(5a_0 + a_1 - 4)$$

Zatem mamy do rozwiązania układ

$$-2(-a_0 + a_1 + 1) + 6(3a_0 + a_1) + 10(5a_0 + a_1 - 4) = 0$$

$$2(-a_0 + a_1 + 1) + 2(3a_0 + a_1) + 2(5a_0 + a_1 - 4) = 0$$

czyli układ

$$70a_0 + 14a_1 - 42 = 0, \quad 14a_0 + 6a_1 - 6 = 0$$

czyli układ

$$5a_0 + a_1 = 3, \quad 7a_0 + 3a_1 = 3$$

Posumowanie

- **Dane:**
 m punktów na płaszczyźnie (x_j, y_j) dla $j = 1, \dots, m$,
 n liczba naturalna
- **Wynik:** wielomian $w_n(x) = a_0x^n + \dots + a_{n-1}x + a_n$,
stop. $(w_n(x)) \leq n$, aproksymujący dane w sensie
najmniejszych kwadratów.
- **Algorytm.** Aby wyznaczyć współczynniki a_0, \dots, a_n
wielomianu aproksymującego (optymalnego) $w_n(x)$,
trzeba rozwiązać układ $n + 1$ równań liniowych

$$\sum_{j=1}^m x_j^{n-i} \left(\sum_{k=0}^n a_k x_j^{n-k} - y_j \right) = 0 \quad \text{dla } i = 0, \dots, n.$$

Uwaga. Niewiadomymi są współczynniki a_0, \dots, a_n
wielomianu aproksymującego. Liczby x_j i y_j są dane.