

Aproksymacja i Metoda Najmniejszych Kwadratów

Gabriel Tyszk

23 czerwca 2025

Spis treści

1	Aproksymacja Średniokwadratowa	1
1.1	Dlaczego stosuje się aproksymację zamiast interpolacji?	1
1.2	Na czym polega aproksymacja metodą najmniejszych kwadratów?	1
1.2.1	Dla funkcji ciągłych (analitycznych)	2
1.2.2	Dla danych dyskretnych (punktowych)	2
2	Metoda Najmniejszych Kwadratów dla Wielomianu	2
2.1	Układ Równań Normalnych	2
2.2	Interpretacja pomiarowa	3
2.3	Linearyzacja funkcji nieliniowych	3

1 Aproksymacja Średniokwadratowa

Aproksymacja średniokwadratowa polega na przybliżonym zastępowaniu danej funkcji inną funkcją, minimalizując sumę kwadratów różnic między nimi. Jest to odmienne podejście od interpolacji, gdzie poszukuje się funkcji, której wykres przechodzi dokładnie przez zadane punkty (węzły).

1.1 Dlaczego stosuje się aproksymację zamiast interpolacji?

Zastosowanie aproksymacji jest często bardziej uzasadnione niż interpolacja w wielu praktycznych sytuacjach:

- **Efekt Rungego:** W przypadku interpolowania za pomocą wielomianów potęgowych, wraz ze wzrostem liczby węzłów wzrasta rząd wielomianu interpolacyjnego, co jednak nie zawsze prowadzi do ulepszenia przybliżonego przedstawienia funkcji na rozpatrywanym obszarze a może prowadzić do oscylacji wielomianu
- **Funkcje sklejane:** Wymagają one stosowania różnych postaci funkcji na kolejnych przedziałach.
- **Błędy danych eksperymentalnych:** Jeśli wartości funkcji w zadanych punktach są obarczone błędami pomiarowymi (np. wynikami eksperymentów), wymuszanie na funkcji interpolującej przechodzenia dokładnie przez te punkty jest nieuzasadnione, gdyż błędy pomiarowe miałyby znaczący wpływ na wynik. W aproksymacji wpływ ten jest mniej istotny.
- **Wiele pomiarów dla tej samej wartości argumentu:** Zdarza się również, że znamy kilka pomiarów wartości funkcji dla tej samej wartości jej argumentu.

1.2 Na czym polega aproksymacja metodą najmniejszych kwadratów?

W metodzie najmniejszych kwadratów poszukuje się funkcji $P(x)$, która w sposób przybliżony przedstawia funkcję $y = f(x)$ na całym rozpatrywanym obszarze argumentów x .

1.2.1 Dla funkcji ciągłych (analitycznych)

Jeśli funkcja $f(x)$ jest znana w przedziale $[a, b]$, celem jest minimalizacja całki z kwadratu różnicy między funkcjami:

$$\int_a^b [f(x) - P(x)]^2 dx \quad (1)$$

Wielkość:

$$\Delta = \sqrt{\frac{\int_a^b [f(x) - P(x)]^2 dx}{b - a}} \quad (2)$$

nazywa się odchyleniem średniokwadratowym i charakteryzuje błąd przybliżenia.

1.2.2 Dla danych dyskretnych (punktowych)

Gdy analityczne wyrażenie funkcji $f(x)$ jest nieznane, a funkcja jest dana przez swoje wartości $y_i = f(x_i)$ w $n + 1$ punktach x_0, x_1, \dots, x_n , minimalizuje się sumę kwadratów różnic:

$$S = \sum_{i=0}^n [f(x_i) - P(x_i)]^2 \quad (3)$$

Zamiast odchylenia średniokwadratowego stosuje się wtedy

$$\Delta_n = \sqrt{\frac{\sum_{i=0}^n [f(x_i) - P(x_i)]^2}{n + 1}} \quad (4)$$

Ważne jest, że minimalizowana jest różnica wartości rzędnej y punktów i prostej (lub innej funkcji aproksymującej), a nie odległość geometryczna punktu od prostej.

2 Metoda Najmniejszych Kwadratów dla Wielomianu

Szukamy wielomianu stopnia k (gdzie $k < n$) postaci:

$$P_k(x) = a_0 + a_1x + a_2x^2 + \dots + a_kx^k \quad (5)$$

Wyrażenie S przyjmuje wówczas postać:

$$S = \sum_{i=0}^n (y_i - a_0 - a_1x_i - a_2x_i^2 - \dots - a_kx_i^k)^2 \quad (6)$$

Aby znaleźć minimum funkcji S względem współczynników a_0, a_1, \dots, a_k , należy obliczyć pochodne cząstkowe $\frac{\partial S}{\partial a_j}$ dla $j = 0, 1, \dots, k$ i przyrównać je do zera.

2.1 Układ Równań Normalnych

Obliczając pochodne cząstkowe i dzieląc każde równanie przez 2, otrzymujemy następujący układ równań liniowych, zwany **układem równań normalnych**:

$$\begin{aligned} a_0(n+1) + a_1 \sum x_i + a_2 \sum x_i^2 + \dots + a_k \sum x_i^k &= \sum y_i \\ a_0 \sum x_i + a_1 \sum x_i^2 + a_2 \sum x_i^3 + \dots + a_k \sum x_i^{k+1} &= \sum x_i y_i \\ a_0 \sum x_i^2 + a_1 \sum x_i^3 + a_2 \sum x_i^4 + \dots + a_k \sum x_i^{k+2} &= \sum x_i^2 y_i \\ &\vdots \\ a_0 \sum x_i^k + a_1 \sum x_i^{k+1} + a_2 \sum x_i^{k+2} + \dots + a_k \sum x_i^{2k} &= \sum x_i^k y_i \end{aligned}$$

Mnożnik $(n+1)$ w pierwszym równaniu wynika z sumowania stałej a_0 $n+1$ razy.

Ten układ równań można zapisać w postaci macierzowej:

$$A^T A \cdot \mathbf{a} = A^T \mathbf{y} \quad (7)$$

gdzie:

- \mathbf{a} jest wektorem niewiadomych współczynników: $\mathbf{a} = [a_0, a_1, \dots, a_k]^T$.
- \mathbf{y} jest wektorem wartości funkcji: $\mathbf{y} = [y_0, y_1, \dots, y_n]^T$.
- Macierz A (macierz Vandermonde'a dla wielomianów) ma postać:

$$A = \begin{bmatrix} 1 & x_0 & x_0^2 & \dots & x_0^k \\ 1 & x_1 & x_1^2 & \dots & x_1^k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^k \end{bmatrix} \quad (8)$$

Macierz współczynników $A^T A$ w tym układzie jest macierzą symetryczną i dodatnio określoną. Należy zaznaczyć, że układu równań normalnych nie rozwiązuje się poprzez pomnożenie obu stron równania przez $(A^T)^{-1}$.

2.2 Interpretacja pomiarowa

Metoda najmniejszych kwadratów może być również wyprowadzona z perspektywy pomiarowej, gdzie wartości y_i są modelowane jako suma wielomianu $P_k(x_i)$ i błędu e_i (nazywanego resztą modelu lub błędem losowym):

$$y_i = a_0 + a_1 x_i + a_2 x_i^2 + \dots + a_k x_i^k + e_i \quad (9)$$

Korzystając z tego podejścia finalnie również otrzymujemy układ równań normalnych.

2.3 Linearyzacja funkcji nieliniowych

W przypadku funkcji nieliniowych, takich jak funkcje potęgowe ($y = ax^b$) lub wykładnicze ($y = ae^{bx}$), można je zlinearyzować poprzez transformacje zmiennych (np. logarytmowanie), a następnie zastosować metodę najmniejszych kwadratów do zlinearyzowanego problemu:

- Dla $y = ax^b$: $\ln y = \ln a + b \ln x$.
- Dla $y = ae^{bx}$: $\ln y = \ln a + bx$.