

Wykład 2 - Regresja Wielomianowa

dr inż. Grzegorz Sarwas

Wydział Elektryczny Politechniki Warszawskiej

Regresja

- Regresja - przewidywanie wartości danej zmiennej o ciągłej wartości na podstawie wartości innych zmiennych, zakładając liniowy lub nieliniowy model zależności.
- Miara korelacji Pearsona odpowiada na pytanie **czy** (w jakim stopniu) dane atrybuty (x, y) są od siebie zależne liniowo. Odpowiedzi na pytanie **jak** od siebie zależą, udzieli model regresyjny.

Regresja wielomianowa

Celem regresji wielomianowej jest dopasowanie do danych doświadczalnych/historycznych wykresu funkcji, która oddaje charakter tych danych.

Zadanie polega na znalezieniu wielomianu $\hat{f}(x)$ postaci:

$$\hat{f}(x) = \omega_0 + \omega_1 x^1 + \omega_2 x^2 + \cdots + \omega_n x^n,$$

gdzie $x \in R$, a $\omega \in R^{n+1}$ jest wektorem zawierającym współczynniki wielomianu.

Przykład

Mając dane historyczne $X = x_1, x_2, \dots, x_k$ oraz odpowiadające im wartości $Y = y_1, y_2, \dots, y_k$ będziemy poszukiwać najlepiej dopasowanego wektora ω .

W efekcie musimy znaleźć takie parametry modelu (wielomianu), które zminimalizują nam sumę kwadratów błędu (RSE).

Czyli takiego, dla którego wartości zwracane przez funkcję \hat{f} i odpowiadające im wartości rzeczywiste najmniej się różnią.

Przykład cd.

Założmy że mamy już wybrany wektor $\omega = [\omega_0, \omega_1, \omega_2, \dots, \omega_n]$
możemy zatem obliczyć:

$$\begin{aligned}\hat{f}(x_1) &= \omega_0 + \omega_1 x_1^1 + \omega_2 x_1^2 + \dots + \omega_n x_1^n \\ \hat{f}(x_2) &= \omega_0 + \omega_1 x_2^1 + \omega_2 x_2^2 + \dots + \omega_n x_2^n \\ &\vdots \\ \hat{f}(x_k) &= \omega_0 + \omega_1 x_k^1 + \omega_2 x_k^2 + \dots + \omega_n x_k^n\end{aligned}$$

Błąd dopasowania modelu

Błąd dopasowania możemy wyliczyć poprzez porównanie wartości otrzymanej dla danego argumentu $\hat{f}(x_i)$ z wartością rzeczywistą y_i odpowiadającą temu argumentowi, stąd mamy:

$$e^1 = \hat{f}(x_1) - y_1$$

$$e^2 = \hat{f}(x_2) - y_2$$

$$\vdots$$

$$e^k = \hat{f}(x_k) - y_k$$

Błąd dopasowania modelu cd.

Błąd dopasowania możemy wyliczyć poprzez porównanie wartości otrzymanej dla danego argumentu $\hat{f}(x_i)$ z wartością rzeczywistą y_i odpowiadającą temu argumentowi, stąd mamy:

$$e^1 = \hat{f}(x_1) - y_1$$

$$e^2 = \hat{f}(x_2) - y_2$$

$$\vdots$$

$$e^k = \hat{f}(x_k) - y_k$$

Następnie sumując kwadraty błędów e^1, \dots, e^k otrzymujemy wartość błędów dla danych parametrów modelu:

$$E(\omega) = \sum_{i=1}^k \left(\hat{f}(x_i) - y_i \right)^2$$

Chcemy aby suma kwadratów błędów była jak najmniejsza dla danego stopnia wielomianu, a to co możemy zmieniać to wartości współczynników wektora ω .

Macierzowe przedstawienie równania regresji

Powyższe analizy możemy zapisać w postaci macierzowej.

$$X\omega = Y,$$

$$\text{gdzie } X = \begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^n \\ 1 & x_2 & x_2^2 & \dots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_k & x_k^2 & \dots & x_k^n \end{vmatrix}, \omega = \begin{vmatrix} \omega_0 \\ \omega_1 \\ \dots \\ \omega_n \end{vmatrix}, Y = \begin{vmatrix} y_0 \\ y_1 \\ \dots \\ y_k \end{vmatrix}$$

Rozwiązaniu równania regresji

Ponieważ w ogólnym przypadku macierz X jest macierzą prostokątną, więc nie można jej odwrócić (nie da się wyzerować błędu aproksymacji), za to możemy znaleźć rozwiązanie przybliżone (takie, które minimalizuje nasz błąd):

$$\min_{\omega} ||X\omega - Y||_2^2$$

Rozwiązaniem, które jest w stanie zminimalizować nam sumę kwadratów błędów jest równanie:

$$w = X^+ Y,$$

gdzie X^+ jest pseudoodwrotnością macierzy X .

Równanie Moore'a-Perose'a

W naszym przypadku macierz X jest macierzą pełnego rzędu kolumnowego (tzn. istnieje niezerowy wyznacznik minor o wymiarze $n_{kolumn} \times n_{kolumn}$ lub macierz posiada 2 niezerowe wartości singularne lub wszystkie kolumny są od siebie liniowo niezależne). X^+ zdefiniowana jest wtedy jako:

$$X^+ = (X^T X)^{-1} X^T$$

Odwzorowanie to jest odwzorowaniem jednoznacznym.

Macierz Vandermonda.

Jeśli do aproksymacji naszego modelu wykorzystamy wielomian stopnia $k - 1$ (o 1 mniejszego niż ilość próbek), wtedy macierz X będzie macierzą kwadratową o wymiarze $k \times k$ postaci:

$$X = \begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{k-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{k-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_k & x_k^2 & \dots & x_k^{k-1} \end{vmatrix}.$$

Macierz taka nazywana jest macierzą Vandermonda.

Ponieważ macierz ta jest kwadratowa oraz jest pełnego rzędu kolumnowego, oznacza to, że jest ona macierzą nieosobliwą, a co za tym idzie pozwala się odwrócić. Wtedy zachodzi równanie: $X^+ = X^{-1}$ Co oznacza, że otrzymujemy dokładne rozwiązania (błąd $E(\hat{\omega}) = 0$)