Metody Numeryczne, część 7
studia stacjonarne

Rozkład Cholesky'ego

Zastosowanie do wyznaczania aproksymacji
w sensie najmniejszych kwadratów

Krystyna Ziętak

listopad 2018

Spis treści

Aproksymacja wielomianami w sensie najmniejszych kwadratów

Dane są różne liczby

$$x_1, \ldots, x_m$$

oraz liczby

$$y_1, \ldots, y_m.$$

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

$$w_n(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n,$$

dla którego wyrażenie

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

przyjmuje minimalną wartość.

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

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

Charakteryzacja wielomianu optymalnego

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

$$w_n(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n = \sum_{k=0}^{n} a_k x^{n-k},$$

aproksymujący funkcję o wartościach $y_1, \ldots, y_m$ dla m różnych argumentów $x_1, \ldots, x_m$ w sensie najmniejszych kwadratów. Współczynniki tego optymalnego wielomianu są rozwiązaniem następującego układu równań liniowych

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

Układ równań liniowych z twierdzenia o charakteryzacji wielomianu optymalnego nazywamy **układem równań normalnych**.

Jest to $n+1$ równań liniowych z niewiadomymi $a_0, a_1, \ldots, a_n$.

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

Pomocnicze oznaczenia

X - macierz $m \times (n+1)$

$$X = \begin{bmatrix} x_1^n & x_1^{n-1} & \cdots & x_1^2 & x_1 & 1 \\ x_2^n & x_2^{n-1} & \cdots & x_2^2 & x_2 & 1 \\ . & . & \cdots & . & \ddots & \\ x_m^n & x_m^{n-1} & \cdots & x_m^2 & x_m & 1 \end{bmatrix}$$

$$a = \begin{bmatrix} a_0 \\ a_1 \\ . \\ . \\ . \\ a_n \end{bmatrix}, \qquad y = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_m \end{bmatrix}$$

Układ równań normalnych, zobacz Twierdzenie

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

Inny zapis układu równań normalnych

$$X^T X a = X^T y$$

Jest to układ równań liniowych o macierzy układu $B = X^T X$ stopnia $n + 1$. Niewiadomymi są współczynniki $a_0, \ldots, a_n$ wielomianu optymalnego.

Macierz $B = X^T X$ jest symetryczna i dodatnio określona. Dzięki temu można ją przedstawić w następującej postaci

$$B = CC^T,$$

gdzie macierz C jest macierzą trójkątną dolną z rozkładu Cholesky'ego macierzy B (zobacz następny rozdział).

Macierz dodatnio określona

Symetryczna macierz A jest dodatnio określona wtedy i tylko wtedy, gdy wszystkie jej wartości własne są dodatnie.

Niech macierz $A = [a_{ij}]$ stopnia n będzie symetryczna, dodatnio określona. Wówczas ma ona rozkład Cholesky'ego $A = LL^T$, gdzie L jest nieosobliwą macierzą trójkątną dolną stopnia n.

Elementy dolnego trójkąta macierzy $L = [\ell_{ij}]$ wyznacza się z wzorów dla $j = 1, 2, \ldots, n$:

$$\ell_{jj} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} \ell_{jk}^2},$$

$$\ell_{ij} = \frac{a_{ij} - \sum_{k=1}^{j-1} \ell_{ik}\ell_{jk}}{\ell_{jj}}, \qquad i = j+1, j+2, \ldots, n.$$

Uwaga. Jeśli $j = 1$, to suma $\sum_{k=1}^{j-1} \beta_k$ jest równa zero. Dlatego powyższe wzory trzeba oddzielnie implementować dla $j = 1$.

Podsumowanie

- Współczynniki wielomianu aproksymującego w sensie najmniejszych kwadratów są rozwiązaniem układu równań normalnych
$$X^T X a = X^T y$$
(zobacz inny zapis tezy z twierdzenia, strona 8).
- Macierz $B = X^T X$ jest symetryczna i dodatnio określona, ponieważ macierz prostokątna $X \in \mathbb{R}^{m \times (n+1)}$ ma rząd $n + 1$. Dlatego istnieje rozkład Cholesky'ego macierzy B:
$$B = CC^T$$
(C jest macierzą trójkątną dolną). Ten rozkład zastosujemy do rozwiązania układu równań normalnych $X^T X a = X^T y$.

Wyznacz rozkład Cholesky'ego macierzy $B = X^T X$:

$$B = CC^T.$$

Macierz C jest trójkątna dolna. Zastosuj rozkład Cholesky'ego macierzy B do rozwiązania układu równań normalnych $X^T X a = X^T y$. Układ ten możemy zapisać tak:

$$CC^T a = X^T y.$$

Algorytm rozwiązania układu równań normalnych $C(C^T a) = X^T y$

Oznaczenie $w = C^T a$. Macierz C jest macierzą trójkątną dolną z rozkładu Cholesky'ego macierzy $B = X^T X$.

- Wyznacz rozwiązanie w układu $Cw = X^T y$. Jest to układ o nieosobliwej macierzy trójkątnej dolnej C.
- Wyznacz rozwiązanie a układu $C^T a = w$. Elementami wektora a są współczynniki szukanego wielomianu optymalnego.

Uwaga. Trzeba rozwiązać dwa układy równań o macierzach trójkątnych. Rozwiązywanie takich układów jest proste.

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

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

$$X = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \\ x_4 & 1 \\ x_5 & 1 \end{bmatrix}, \qquad y = [y_1, y_2, y_y, y_4, y_5]^T, \qquad a = [a_0, a_1]^T$$

$$B = X^T X = \ldots,$$
$$b = X^T y = \ldots$$

Rozwiąż układ $Ba = b$ za pomocą rozkładu Cholesky'ego macierzy B.