

Literatura

Literatura podstawowa

- [1] D. Kincaid, W. Cheney, *Analiza numeryczna*, WNT, 2005.
- [2] A. Bjorck, G. Dahlquist, *Metody numeryczne*, PWN, 1987.
- [3] M. Dryja, J. i M. Jankowscy, *Przegląd metod i algorytmów numerycznych*, cz. 2, WNT, 1988.
- [4] Z. Fortuna, B. Macukow, J. Wąsowski, *Metody numeryczne*, WNT 1998.
- [5] J. i M. Jankowscy, *Przegląd metod i algorytmów numerycznych*, cz. 1, WNT, 1988.
- [6] J. Stoer, R. Bulirsch, *Wstęp do analizy numerycznej*, PWN, 1987.

Literatura uzupełniająca

- [7] Kielbasiński, H. Schwetlick, *Numeryczna algebra liniowa*, WNT, 1993.
- [8] A. Ralston, *Wstęp do analizy numerycznej*, PWN, 1975.

Znaczna część wykładu została przygotowana na podstawie książki [5].



Wielomiany ortogonalne

Lemat 1

Wielomiany ortogonalne $P_0, P_1, \dots, P_n$ tworzą bazę przestrzeni liniowej Π_n .

Lemat 2

Dowolny wielomian Q_j stopnia j niższego od k jest prostopadły do k -tego wielomianu ortogonalnego P_k ; $(Q_j, P_k) = 0$ dla $j < k$.

Lemat 3

Twierdzenie 9

Ciąg wielomianów ortogonalnych w przestrzeni $\mathbb{X} = \left\{ \begin{matrix} L_p^2[a, b] \\ l_{p,N}^2 \end{matrix} \right\}$ jest określony jednoznacznie z dokładnością do czynników stałych, tj. jeśli $\{P_k\}$ i $\{Q_k\}$ są ciągami ortogonalnymi w $\mathbb{X}$, to $P_k = \alpha_k Q_k$, $\alpha_k = \text{const}$, $k = 0, 1, \dots$



Wielomiany ortogonalne

Twierdzenie 1

Ciąg wielomianów ortogonalnych $\{P_k\}$ w przestrzeni $\mathbb{X} = L_p^2[a, b]$ lub $\mathbb{X} = l_{p,N}^2$ spełnia zależność rekurencyjną

$$\begin{aligned}P_k(x) &= (\alpha_k x + \beta_k)P_{k-1}(x) + \gamma_k P_{k-2}(x) \quad k = 1, 2, \dots \\P_{-1}(x) &= 0, \quad P_0(x) = a_0,\end{aligned}$$

gdzie $\alpha_k = \frac{a_k}{a_{k-1}} \neq 0$, $\beta_k, \gamma_k \neq 0$, γ_1 jest nieistotna, ponieważ $P_{-1}(x) = 0$, a_k są współczynnikami wielomianu P_k przy x^k ,
 $P_k(x) = a_k x^k + \dots$,

$$\beta_k = -\frac{\alpha_k (xP_{k-1}, P_{k-1})}{(P_{k-1}, P_{k-1})}, \quad \gamma_k = -\frac{\alpha_k (P_{k-1}, P_{k-1})}{\alpha_{k-1} (P_{k-2}, P_{k-2})}$$

Dla wielomianów ortogonalnych na zbiorze dyskretnym $\{x_1, x_2, \dots, x_N\}$ (w przestrzeni $l_{p,N}^2$) zależność rekurencyjna jest spełniona tylko dla $k \leq N - 1$.



Wielomiany ortogonalne

Twierdzenie 1 umożliwia konstrukcję wielomianów ortogonalnych o zadanych współczynnikach a_k , $k = 0, 1, \dots, n$. Możemy również ustalić, że $a_k = 1$, $k = 0, 1, \dots, n$. Wówczas $\alpha_k = 1$, $P_0(x) = 1$ i

$$\beta_k = -\frac{(xP_{k-1}, P_{k-1})}{(P_{k-1}, P_{k-1})}, \quad \gamma_k = -\frac{(P_{k-1}, P_{k-1})}{(P_{k-2}, P_{k-2})},$$

Wielomiany Czebyszewa $T_k(x) = \cos(k \arccos x)$, $k = 0, 1, \dots$, tworzą układ ortogonalny w przestrzeni $L_p^2[-1, 1]$ z funkcją wagową $p(x) = \frac{1}{\sqrt{1-x^2}}$ i spełniają zależność rekurencyjną

$$\begin{cases} T_0(x) = 1 & T_1(x) = x \\ T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x) & (k \geq 2) \end{cases}$$



Zadanie aproksymacji średniokwadratowej wielomianami

Dla danej $f \in L^2_\rho[a, b]$ ($I^2_{\rho, N}$), szukamy elementu optymalnego dla f względem podprzestrzeni Π_n , tj. wielomianu $w_f \in \Pi_n$, dla którego zachodzi równość

$$\|f - w_f\|_2 = \inf_{w \in \Pi_n} \|f - w\|_2 = E_n(f).$$

w_f nazywamy n -tym wielomianem optymalnym dla f , a $E_n(f)$ nazywamy n -tym błędem optymalnym.



Zadanie aproksymacji średniokwadratowej wielomianami

Twierdzenie 2

Jeśli $\{P_k\}$ jest ciągiem wielomianów ortogonalnych w $\mathbb{X} = L_p^2[a, b]$ lub $\mathbb{X} = l_{p,N}^2$ i zawiera co najmniej wielomiany $P_0, P_1, \dots, P_n$, to n -ty wielomian optymalny w_f określony jest jednoznacznie i wyraża się wzorem

$$w_f = \sum_{k=0}^n \frac{(f, P_k)}{(P_k, P_k)} P_k,$$

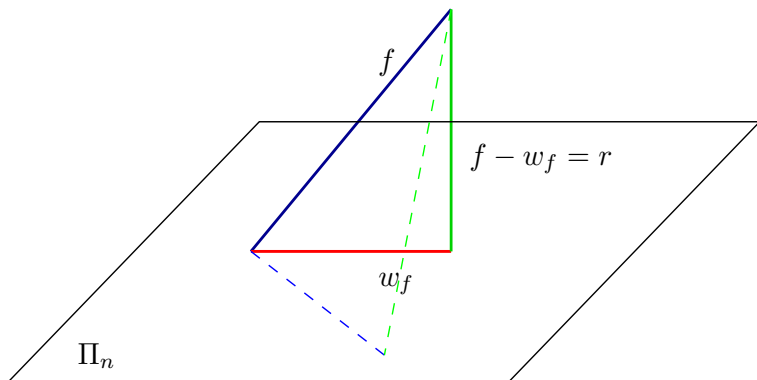
a n -ty błąd optymalny jest równy

$$E_n(f) = \left(\|f\|_2^2 - \sum_{k=0}^n \frac{(f, P_k)^2}{(P_k, P_k)} \right)^{1/2}$$



Zadanie aproksymacji średniokwadratowej wielomianami

Z rysunku poniżej wynika, że w_f jest elementem optymalnym dla f wtedy i tylko wtedy, gdy $f - w_f$ ($f - w_f = r$) jest prostopadły do Π_n ($r \perp \Pi_n$).



Zadanie aproksymacji średniokwadratowej wielomianami

Wielomiany ortogonalne $P_0, P_1, \dots, P_n$ są bazą Π_n na mocy lematu 1. Stąd n -ty wielomian optymalny dla f można przedstawić

$$w_f = \sum_{i=0}^n c_i P_i.$$

w_f jest optymalny, gdy $r \perp \Pi_n$, co jest równoważne $r \perp P_j$, $j = 0, \dots, n$.

$$(f - w_f, P_j) = 0, \quad j = 0, \dots, n$$

$$(w_f, P_j) = (f, P_j), \quad j = 0, \dots, n$$

$$\left(\sum_{i=0}^n c_i P_i, P_j \right) = (f, P_j), \quad j = 0, \dots, n$$

Otrzymujemy układ $n + 1$ równań o $n + 1$ niewiadomych c_i .



$$\begin{array}{rcl} c_0(P_0, P_0) + c_1(P_1, P_0) + \cdots + c_n(P_n, P_0) & = & (f, P_0) \\ c_0(P_0, P_1) + c_1(P_1, P_1) + \cdots + c_n(P_n, P_1) & = & (f, P_1) \\ \vdots & & \vdots \\ c_0(P_0, P_n) + c_1(P_1, P_n) + \cdots + c_n(P_n, P_n) & = & (f, P_n) \end{array}$$

$(P_i, P_j) = 0$ dla $i \neq j$, $(P_i, P_j) \neq 0$ dla $i = j$. Z układu

wyznaczamy $c_i = \frac{(f, P_i)}{(P_i, P_i)}$, $i = 0, 1, \dots, n$ i wielomian

$w_f = \sum_{i=0}^n \frac{(f, P_i)}{(P_i, P_i)} P_i$. Z faktu, że w przestrzeniach unitarnych zachodzi twierdzenie Pitagorasa mamy

$$\|f\|_2^2 = \|w_f\|_2^2 + \|f - w_f\|_2^2$$

$$\|f - w_f\|_2 = (\|f\|_2^2 - \|w_f\|_2^2)^{1/2} = \left(\|f\|_2^2 - \sum_{k=0}^n \frac{(f, P_k)^2}{(P_k, P_k)} \right)^{1/2}$$



n -ty wielomian optymalny

krok 1. Wyznaczyć wielomiany ortogonalne $P_0, \dots, P_n$ np. ze związku rekurencyjnego lub wykorzystać znane wielomiany ortogonalne.

krok 2. Obliczyć współczynniki $c_k = \frac{(f, P_k)}{(P_k, P_k)}$, $k = 0, 1, \dots, n$

Cała informacja o wielomianach ortogonalnych powinna wyrażać się współczynnikami związku rekurencyjnego, $\alpha_k, \beta_k, \gamma_k$.

Nie powinno się wyznaczać postaci naturalnej n -tego wielomianu optymalnego w_f bez potrzeby

$$w_f(x) = \sum_{k=0}^n c_k P_k(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0.$$



n -ty wielomian optymalny

Wartość n -tego wielomianu optymalnego w_f

$$w_f(x) = \sum_{k=0}^n c_k P_k(x)$$

w punkcie x liczymy za pomocą algorytmu **Clenshow'a** wykorzystującego fakt, że wielomiany ortogonalne spełniają związek rekurencyjny:

$$\begin{cases} Y_{n+1} := Y_{n+2} := 0 \\ Y_k := c_k + (\alpha_{k+1}x - \beta_{k+1})Y_{k+1} - \gamma_{k+2}Y_{k+2}, \quad k = n, n-1, \dots, 0 \\ w_f(x) = a_0 Y_0 \end{cases}$$



Aproksymacja w przestrzeniach unitarnych – ogólne podejście

Dla danej $f \in \mathbb{X}$, $\mathbb{X}$ jest przestrzenią unitarną, szukamy elementu optymalnego dla f względem podprzestrzeni $\mathbb{Y} \subset \mathbb{X}$, tj. $g_f \in \mathbb{Y}$, dla którego zachodzi równość

$$\|f - g_f\| = \inf_{g \in \mathbb{Y}} \|f - g\|.$$

Twierdzenie 3

Niech $\mathbb{X}$ będzie przestrzenią unitarną, $\mathbb{Y}$ będzie jej podprzestrzenią, $\mathbb{Y} \subset \mathbb{X}$.

Dla danej $f \in \mathbb{X}$, $g_f \in \mathbb{Y}$ jest elementem optymalnym dla f względem $\mathbb{Y}$ wtedy i tylko wtedy, gdy $f - g_f$ jest ortogonalny do $\mathbb{Y}$.



Aproksymacja w przestrzeniach unitarnych – ogólne podejście

Niech układ, niekoniecznie ortogonalny, $f_0, f_1, \dots, f_n$ będzie bazą $\mathbb{Y}$. Wówczas szukany element g_f możemy przedstawić:

$$g_f = \sum_{i=0}^n c_i f_i.$$

Podobnie jak dla przestrzeni Π_n , g_f jest optymalny, gdy $f - g_f \perp \mathbb{Y}$, co jest równoważne $f - g_f \perp f_j, j = 0, \dots, n$.

$$\left(\sum_{i=0}^n c_i f_i, f_j \right) = (f, f_j), \quad j = 0, \dots, n$$



podjęcie

Otrzymujemy układ równań normalnych o $n + 1$ niewiadomych c_j .

$$\begin{array}{rcl} c_0(f_0, f_0) + c_1(f_1, f_0) + \cdots + c_n(f_n, f_0) & = & (f, f_0) \\ c_0(f_0, f_1) + c_1(f_1, f_1) + \cdots + c_n(f_n, f_1) & = & (f, f_1) \\ \vdots & & \vdots \\ c_0(f_0, f_n) + c_1(f_1, f_n) + \cdots + c_n(f_n, f_n) & = & (f, f_n) \end{array}$$

Macierz $n + 1 \times n + 1$ powyższego układu nazywamy macierzą Grama.



Aproksymacja w przestrzeniach unitarnych – ogólne podejście

Przykład (Układ równań normalnych źle uwarunkowany)

Niech $\mathbb{X} = L_p^2[0, 1]$, $p(x) = 1$, $\mathbb{Y} = \Pi_n$. Iloczyn skalarny w tej przestrzeni jest postaci $(f, g) = \int_0^1 f(x)g(x)dx$.

Szukamy n -tego wielomianu optymalnego dla f względem Π_n .

Jako bazę Π_n bierzemy $f_0 = 1, f_1 = x, f_2 = x^2, \dots, f_n = x^n$.

Wyznaczamy elementy macierzy Grama

$$(f_i, f_j) = (x^i, x^j) = \int_0^1 x^i x^j dx = \frac{1}{i+j+1}, \quad i, j = 0, 1, \dots, n.$$

Macierz układu jest zatem macierzą Hilberta.



Dowód Lematu 1

Dowód wynika bezpośrednio z faktu, że dla dowolnego układu $f_1, f_2, \dots, f_n$ liniowo niezależnych elementów przestrzeni unitarnej istnieje układ ortogonalny $g_1, g_2, \dots, g_n$ taki, że każdy z elementów g_i , $i = 1, \dots, n$, jest kombinacją liniową $f_1, \dots, f_i$. Układ ortogonalny konstruujemy metodą ortogonalizacji Grama-Schmidta. □



Dowód Lematu 2

Przedstawmy wielomian Q_j jako kombinację liniową wielomianów ortogonalnych (z lematu 1 wynika, że $P_0, \dots, P_j$ są bazą Π_j).

$$Q_j = \sum_{i=0}^j \alpha_i P_i.$$

Biorąc iloczyn skalarny obu stron z wielomianem P_k , otrzymujemy

$$(Q_j, P_k) = \sum_{i=0}^j \alpha_i (P_i, P_k) = 0,$$

ponieważ $(P_i, P_k) = 0$ dla $i < k$.



Dowód Lematu 3

Niech $P_0, P_1, \dots, P_n$ i $Q_0, Q_1, \dots, Q_n$ będą różnymi ciągami ortogonalnymi. Z lematu 1 wynika, że wielomian P_k możemy przedstawić jako kombinację liniową wielomianów $Q_0, \dots, Q_k$

$$P_k = \sum_{j=0}^k \alpha_j Q_j.$$

Biorąc iloczyn skalarny obu stron powyższego równania z wielomianem Q_l , $l < k$, stosując lemat 2 i korzystając z faktu, że ciąg $\{Q_j\}$ jest ortogonalny otrzymujemy

$$0 = (Q_l, P_k) = \sum_{j=0}^k \alpha_j (Q_l, Q_j) = \alpha_l (Q_l, Q_l).$$

Z definicji $(Q_l, Q_l) \neq 0$. Stąd i z powyższego równania otrzymujemy $\alpha_l = 0$ dla $l < k$. Zatem

$$P_k = \sum_{l=0}^k \alpha_l Q_l = \alpha_k Q_k. \quad \square$$



Dowód Twierdzenia 1

P_k jest wielomianem dokładnie k -tego stopnia, czyli $a_k \neq 0$. Przyjmijmy $\alpha_k = \frac{a_k}{a_{k-1}}$. Wówczas wielomian $P_k - \alpha_k x P_{k-1}$ jest stopnia mniejszego niż k . Przedstawmy ten wielomian w postaci kombinacji liniowej wielomianów ortogonalnych (zob. lemat 1)

$$P_k - \alpha_k x P_{k-1} = \sum_{i=0}^{k-1} b_i P_i.$$

Pokażemy, że $(P_k - \alpha_k x P_{k-1}, P_j) = 0$ dla $j < k - 2$. Możemy go zapisać

$$(P_k - \alpha_k x P_{k-1}, P_j) = (P_k, P_j) - \alpha_k (x P_{k-1}, P_j)$$

$(P_k, P_j) = 0$ z ortogonalności układu. Ponieważ w przestrzeniach $L_p^2[a, b]$ i $l_{p,N}^2$ spełniona jest równość

$$(x P_{k-1}, P_j) = (P_{k-1}, x P_j)$$

(zob. postać iloczynów skalarnych). Z lematu 2 natychmiast dostajemy $(P_{k-1}, x P_j) = 0$, co daje $(x P_{k-1}, P_j) = 0$. Zatem

$$0 = (P_k - \alpha_k x P_{k-1}, P_j) = \sum_{i=0}^{k-1} b_i (P_i, P_j) = b_j (P_j, P_j).$$

Z definicji $(P_j, P_j) \neq 0$. Stąd i powyższego równania otrzymujemy $b_j = 0$ dla $j < k - 2$.



Dowód Twierdzenia 1 ...

Tak więc

$$P_k = \alpha_k xP_{k-1} + b_{k-1}P_{k-1} + b_{k-2}P_{k-2} = \alpha_k xP_{k-1} + \beta_k P_{k-1} + \gamma_k P_{k-2}, \quad (1)$$

gdzie $\beta_k = b_{k-1}$, $\gamma_k = b_{k-2}$. Stosując iloczyn skalarny z P_{k-1} do powyższej równości otrzymujemy

$$0 = (P_k, P_{k-1}) = \alpha_k (xP_{k-1}, P_{k-1}) + \beta_k (P_{k-1}, P_{k-1})$$

Wyznaczamy

$$\beta_k = -\frac{\alpha_k (xP_{k-1}, P_{k-1})}{(P_{k-1}, P_{k-1})}.$$

Podobnie wyznaczamy γ_k biorąc iloczyn skalarny z P_{k-2}

$$\gamma_k = -\frac{\alpha_k (xP_{k-1}, P_{k-2})}{(P_{k-2}, P_{k-2})} = -\frac{\alpha_k (P_{k-1}, P_{k-1})}{\alpha_{k-1} (P_{k-2}, P_{k-2})}.$$

$(P_{k-1}, P_{k-1}) = \alpha_{k-1} (xP_{k-2}, P_{k-1})$ (zob. (1)), a

$(xP_{k-2}, P_{k-1}) = (P_{k-2}, xP_{k-1})$.



Dowód Twierdzenia 1 ...

Dla wielomianów ortogonalnych na zbiorze dyskretnym $\{x_1, x_2, \dots, x_N\}$ zależność rekurencyjna jest spełniona tylko dla $k \leq N - 1$.

Następujący wielomian

$$a_N(x - x_1)(x - x_2) \cdots (x - x_N)$$

zeruje się w x_k ze zbioru. Czyli jest on ortogonalny do wszystkich wielomianów ortogonalnych P_j niższego stopnia. Z lematu 3 wynika, że ciąg wielomianów ortogonalnych jest określony jednoznacznie z dokładnością do czynników stałych, więc N -ty wielomian ortogonalny musi być postaci $a_N(x - x_1)(x - x_2) \cdots (x - x_N)$. Zatem $(P_N, P_N) = 0$, co dowodzi, że P_N jest zerowym elementem $l_{p,N}^2$. Więc nie może być elementem układu ortogonalnego. □



Dowód Twierdzenia 3

($\Rightarrow$) Załóżmy nie wprost, że istnieje $h \in \mathbb{Y}$, takie, że $(f - g_f, h) = \alpha \neq 0$. Wtedy $h \neq 0$ i

$$g := g_f + \frac{\alpha}{(h, h)} h \in \mathbb{Y}.$$

Zatem

$$\begin{aligned} \|f - g\|^2 &= (f - g, f - g) = \left(f - g_f - \frac{\alpha}{(h, h)} h, f - g_f - \frac{\alpha}{(h, h)} h\right) \\ &= \|f - g_f\|^2 - 2 \frac{\alpha}{(h, h)} (f - g_f, h) + \frac{\alpha^2}{(h, h)^2} (h, h) \\ &= \|f - g_f\|^2 - 2 \frac{\alpha}{(h, h)} \alpha + \frac{\alpha^2}{(h, h)} = \|f - g_f\|^2 - \frac{\alpha^2}{(h, h)} < \|f - g_f\|^2. \end{aligned}$$

Sprzeczność, że g_f jest elementem optymalnym dla f względem $\mathbb{Y}$.



Dowód Twierdzenia 3...

($\Leftarrow$) Załóżmy, że $f - g_f$ jest ortogonalny do $\mathbb{Y}$. Stąd

$$(f - g_f, h) = 0 \text{ dla wszystkich } h \in \mathbb{Y}.$$

Dowolny element $g \in \mathbb{Y}$ możemy przedstawić

$$g = g_f - (g_f - g), \quad g_f - g \in \mathbb{Y}.$$

Zatem $(f - g_f, g_f - g) = 0$. Z twierdzenia Pitagorasa (zob. wykład 11) mamy

$$\begin{aligned} \|f - g\|^2 &= \|(f - g_f) + (g_f - g)\|^2 = \|f - g_f\|^2 + \|g_f - g\|^2 \\ &\geq \|f - g_f\|^2. \end{aligned}$$

Stąd g_f jest elementem optymalnym dla f względem $\mathbb{Y}$. □

