Analiza Numeryczna, Wykład 9, 11/12/2019

Powtórka

Wielomiany Bernsteina: $B_k^n(t) := t^k (1-t)^{n-k}$ $(n \in \mathbb{N}, k = 0,1,\dots,n)$, $B_k^n(t) \in \Pi_n \setminus \Pi_{n-1}$

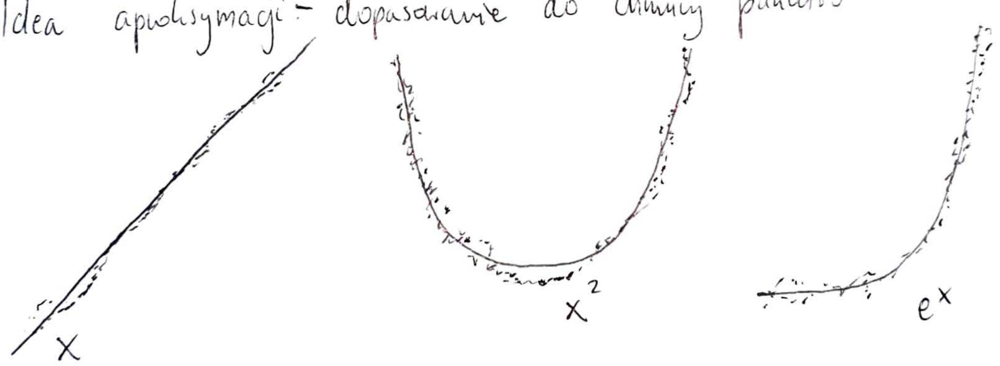
Dzięki własności $\sum_{k=0}^{n} B_k^n(t) \equiv 1$ możemy zdefiniować krzywą Beziera daną wzorem

$P_n(t) := \sum_{k=0}^{n} W_k B_k^n(t)$ $(t \in [0,1]$, W_k - punkty na płaszczyźnie$)$. $\forall t \in [0,1]$ $P_n(t) \in \mathrm{conv}\{W_i\}$,

tzn. $P_n(t)$ jest punktem na płaszczyźnie jako kombinacja wypukła punktów kontrolnych W_i dla $i = 0,1,\dots,n$. Możemy obliczyć punkt $P_n(t)$ używając algorytmu de Casteljau w czasie $O(n^2)$, łatwo przedstawić jego interpretację geometryczną. Kombinacja barycentryczna: $\alpha_0, \alpha_1, \dots, \alpha_n \Rightarrow \sum_{i=0}^{n} \alpha_i = 1$, $W_0,\dots,W_n$ - punkty kontrolne, czyli $\alpha_0 W_0 + \alpha_1 W_1 + \dots + \alpha_n W_n$ jest punktem.

Aproksymacja średniokwadratowa na zbiorze dyskretnym

Idea aproksymacji - dopasowanie do chmury punktów



Rezygnujemy z interpolacji chcemy być „blisko" chmury punktów

Norma średniokwadratowa ($\equiv$ śr□) na zbiorze dyskretnym:

Niech dane będą parami różne punkty $X := \{x_0, x_1, \dots, x_N\}$ oraz funkcja f określona na X. Normę średniokwadratową funkcji f na zbiorze X oznaczamy symbolem $\|f\|_2$ i definiujemy wzorem:

$$\|f\|_2 := \sqrt{\sum_{k=0}^{N} \left(f(x_k)\right)^2}$$

Uwagi:

1° Norma średniokwadratowa to funkcja (funkcjonał): $\|\cdot\|_2 : F \longrightarrow \mathbb{R}$

$\uparrow$ zbiór funkcji

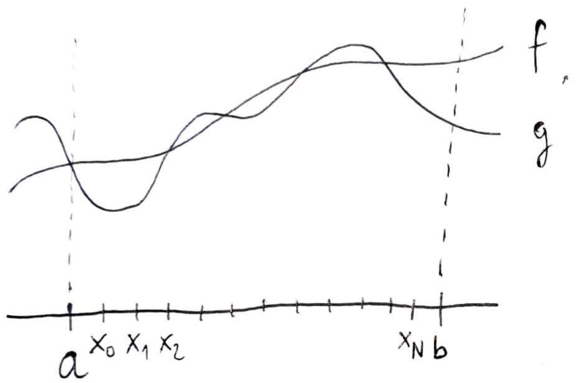
2° Własności normy (f,g – funkcje):

(a) $\|f\|_2 = 0 \iff f(x_k)$ dla $k = 0,1,\dots,N$

(b) $\|\alpha \cdot f\|_2 = |\alpha| \cdot \|f\|_2 \quad (\alpha \in \mathbb{R})$

(c) $\|f+g\|_2 \leqslant \|f\|_2 + \|g\|_2$ (nierówność trójkąta)

Pytanie: Jak sprawdzić, czy funkcje f i g są do siebie podobne (bliskie siebie)?



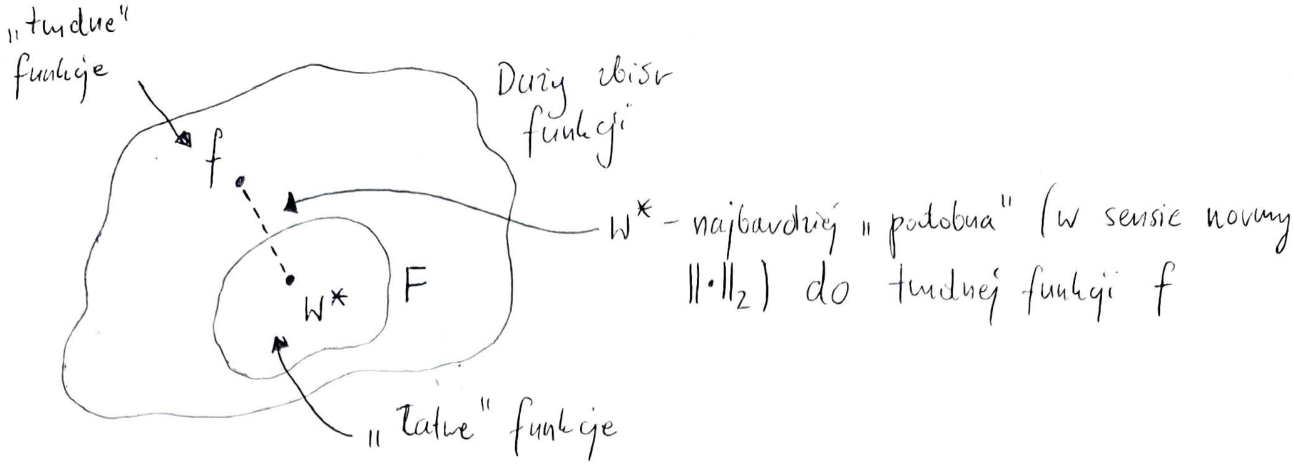
Idea: Sprawdzić, czy $\|f-g\|_2$ jest małe, jeśli tak, to są blisko, jeśli nie, to daleko.

$$\|f-g\|_2 = \sqrt{\sum_{k=0}^{N} \left(f(x_k) - g(x_k)\right)^2}$$

Zadanie: Aproksymacja średniokwadratowa

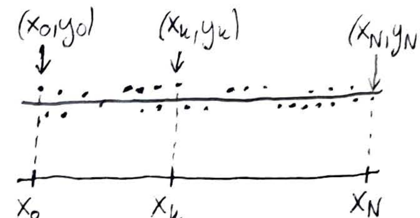
Dla danego zbioru $X := \{x_0, x_1, \dots, x_N\}$ ($x_i \neq x_j$ dla $i \neq j$) oraz funkcji f określonej na X ($f(x_k) =: y_k$) znaleźć taki element $w^* \in F$ (F – ustalony zbiór funkcji) nazywany elementem optymalnym dla funkcji f na zbiorze X w sensie aproksymacji średniokwadratowej, że

$$\|f - w^*\|_2 = \min_{w \in F} \|f - w\|_2 \equiv \min_{w \in F} \sqrt{\sum_{k=0}^{N} \Big(\underbrace{f(x_k)}_{y_k} - w(x_k)\Big)^2}$$



Przykłady:

(a) $F := \{ w(x) = a : a \in \mathbb{R} \} = \Pi_0$ - funkcje stałe



Pomiary $(x_k, y_k) \equiv (x_k, f(x_k))$ dla $0 \leq k \leq N$,
szukamy elementu optymalnego $w^* \in F$ o własności:

$$\|f - w^*\|_2 = \min_{w \in F} \|f - w\|_2 = \min_{a \in \mathbb{R}} \sqrt{\sum_{k=0}^{N} \Big(\underbrace{f(x_k)}_{y_k} - \underbrace{a}_{w(x_k)}\Big)^2}$$

Funkcja błędu: $E(a) = \sum_{k=0}^{N} (y_k - a)^2$ ← szukamy minimum:

$E'(a) = 0$ ← warunek konieczny na istnienie ekstremum

$$E'(a) = -2 \sum_{k=0}^{N} (y_k - a)^1 = 0$$

$$\sum_{k=0}^{N} y_k - \underbrace{\sum_{k=0}^{N} a}_{=(N+1)a} = 0$$

$$a = \frac{1}{N+1} \left(\sum_{k=0}^{N} y_k \right)$$ ← średnia arytmetyczna

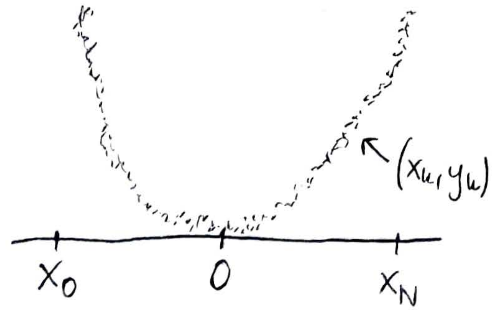
Wniosek: Element optymalny dla pomiarów (x_k, y_k) $(0 \leq k \leq N)$ w sensie aproksymacji średniokwadratowej to $w^*(x) = a^*$, gdzie

$$a^* = \frac{\sum_{k=0}^{N} (y_k)}{N+1}$$ przy założeniu, że model obejmuje jedynie funkcje stałe (tzn. $F \equiv \Pi_0$).

(b) $F := \{w(x) = ax^2 : a \in \mathbb{R}\}$, $F \neq \Pi_2$, ale $F \subset \Pi_2$

Pomiary $(x_k, y_k) \equiv (x_k, f(x_k))$ $\quad (0 \leq k \leq N)$

Szukamy elementu $w^* \in F$ o własności:



$$\|f - w^*\|_2 = \min_{w \in F} \|f - w\|_2 =$$

$$= \min_{\substack{w(x)=ax^2, \\ a \in \mathbb{R}}} \sqrt{\sum_{k=0}^{N} \Big(\underbrace{y_k}_{f(x_k)} - \underbrace{w(x_k)}_{a x_k^2}\Big)^2} = \min_{a \in \mathbb{R}} \sqrt{E(a)}$$

← funkcja błędu

...

Funkcja błędu: $E(a) = \sum_{k=0}^{N} \left(y_k - a x_k^2\right)^2$ ← szukamy minimum

$$E'(a) = -2 \sum_{k=0}^{N} \left(y_k - a x_k^2\right)^1 x_k^2 =$$

$$= -2\left(\sum_{k=0}^{N} y_k x_k^2 - a \sum_{k=0}^{N} x_k^4\right)$$

$$E'(a) = 0$$

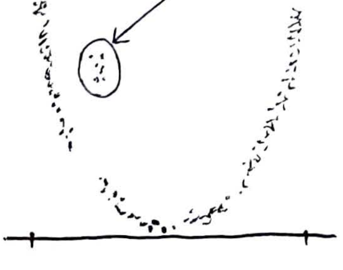
$$a = \frac{\sum_{k=0}^{N} y_k x_k^2}{\sum_{k=0}^{N} x_k^4}$$

Wniosek: Element optymalny w tym wypadku (!) to: $w^*(x) = a^* x^2$,

gdzie $a^* := \dfrac{\sum_{k=0}^{N} y_k x_k^2}{\sum_{k=0}^{N} x_k^4}$.

↳ nasza funkcja to ax^2, czyli wierzchołek ma w punkcie $(0,0)$

Uwaga: „Odstające" obserwacje



Aproksymacja średniokwadratowa radzi sobie całkiem nieźle z „odstającymi" obserwacjami, zwykle współczynnik się minimalnie zmienia.

Znajdowanie ekstremów funkcji wielu zmiennych

POCHODNE CZĄSTKOWE!

np. $f(x,y,z) = ax^2 + (y-z)^7 + \cos\left(\frac{x}{y}\right)$

$$\frac{\partial f(x,y,z)}{\partial x} = 2ax + 0 - \sin\left(\frac{x}{y}\right)\cdot\frac{1}{y}$$

$$\frac{\partial f(x,y,z)}{\partial y} = 0 - 7(y-z)^6 + 0$$

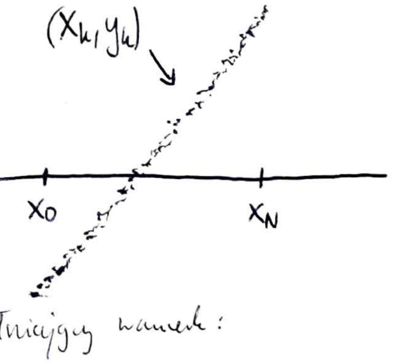
$$\frac{\partial f(x,y,z)}{\partial z} = 0 + 7(y-z)^6 - \sin\left(\frac{x}{y}\right)\cdot\left(-xy^{-2}\right)$$

liczenie pochodnej po danej zmiennej przy założeniu, że pozostałe zmienne to stałe

Warunkiem koniecznym istnienia ekstremum funkcji f wielu zmiennych $x_0, x_1, \dots, x_L$ jest zerowanie się wszystkich pochodnych cząstkowych:

$$\begin{cases} \dfrac{\partial f(x_0,\dots,x_L)}{\partial x_0} = 0 \\ \dfrac{\partial f(x_0,\dots,x_L)}{\partial x_1} = 0 \\ \vdots \\ \dfrac{\partial f(x_0,\dots,x_L)}{\partial x_L} = 0 \end{cases}$$

układ równań nazywany układem równań normalnych



(c) $F := \{ax+b : a,b \in \mathbb{R}\} \equiv \Pi_1$

Pomiary (x_k, y_k) $(0 \le k \le N)$,

Chcemy znaleźć element optymalny $w^* \in F$ spełniający warunek:

$$\|f-w^*\|_2 = \min_{w\in F} \|f-w\|_2 = \min_{a,b\in\mathbb{R}} \sqrt{E(a,b)}$$

Funkcja błędu: $E(a,b) := \sum_{k=0}^{N} \left(f(x_k) - w(x_k)\right)^2 = \sum_{k=0}^{N} \left(y_k - ax_k - b\right)^2$

Szukamy minimum funkcji dwóch zmiennych:

$$\begin{cases} \frac{\partial E(a,b)}{\partial a} = -2\sum_{k=0}^{N}(y_k - ax_k - b)^1 x_k = 0 \\ \frac{\partial E(a,b)}{\partial b} = -2\sum_{k=0}^{N}(y_k - ax_k - b)^1 = 0 \end{cases} \Rightarrow \begin{cases} a\left(\sum_k x_k^2\right) + b\left(\sum_k x_k\right) = \sum_k x_k y_k \\ a\left(\sum_k x_k\right) + b(N+1) = \sum_k y_k \end{cases}$$

$$\Rightarrow \begin{cases} a = \frac{(N+1)S_4 - S_1 S_3}{(N+1)S_2 - S_1^2} \\ b = \frac{S_2 S_3 - S_1 S_4}{(N+1)S_2 - S_1^2} \end{cases} \qquad S_1 := \sum_k x_k \quad S_2 := \sum_k x_k^2 \quad S_3 := \sum_k y_k \quad S_4 := \sum_k x_k y_k$$

Wniosek: Element optymalny $w^*(x) = a^*(x) + b^*$ (regresja liniowa), gdzie

$$\begin{cases} a^* = \frac{(N+1)S_4 - S_1 S_3}{(N+1)S_2 - S_1^2} \\ b^* = \frac{S_2 S_3 - S_1 S_4}{(N+1)S_2 - S_1^2} \end{cases}$$

Sytuacja ogólna:

Wybieramy pewne funkcje (podstawowe) $g_0(x), g_1(x), \dots, g_m(x)$ (np. $g_i(x) = x^i$, wtedy $\mathrm{lin}\{g_0, g_1, \dots, g_m\} \equiv \Pi_m$) i za model przyjmujemy $F := \{a_0 g_0(x) + a_1 g_1(x) + \dots + a_m g_m(x) : a_0, \dots, a_m \in \mathbb{R}\}$. Dla pomiarów (x_k, y_k) $(0 \le k \le N)$ szukamy elementu optymalnego $w^* \in F$ o własności:

$$\|f - w^*\|_2 = \min_{w \in F}\|f - w\|_2 \equiv \min_{a_0,\dots,a_m \in \mathbb{R}} \sqrt{E(a_0, a_1, \dots, a_m)}, \; E \text{ - funkcja błędu},$$

$$E(a_0, \dots, a_m) = \sum_{k=0}^{N}\Big(y_k - \overbrace{\sum_{i=0}^{m} a_i g_i(x_k)}^{w(x_k)}\Big)^2,$$ następnie szukamy minimum E za pomocą pochodnych cząstkowych (przyrównujemy je do zera), którego rozwiązanie definiuje element optymalny w^* w sensie aproksymacji średniokwadratowej na zbiorze dyskretnym.