Metody numeryczne

Aproksymacja szeregami wielomianów

dr hab. inż. Jerzy Baranowski, prof. AGH

Aproksymacja wielomianami

Co poza interpolacją?

- Z twierdzenia Weierstrassa wiemy, że wielomiany mogą dowolnie dokładnie przybliżyć funkcję ciągłą
- W ogólności możemy

$$f(x) \approx \sum_{i=0}^{n} a_i T_i(x)$$

gdzie $T_i(x)$ są funkcjami bazy.

O co chodzi?

Przestrzeń wektorowa

$(K, + , \cdot\ )$ jest ciałem skalarów (np. ciałem rzeczywistych)

Przestrzenią wektorową nad ciałem nazywamy zbiór wektorów V, z określonymi działaniami:

- Dodawaniem wektorów $+ : V \times V \to V$
- Mnożenie wektorów przez skalar $\cdot : K \times V \to V$

Kombinacja liniowa

Podstawowy element aproksymacji

Niech V będzie przestrzenią liniową nad ciałem K. Niech $x_1, \ldots, x_n$ będzie skończonym układem wektorów przestrzeni V i niech $\alpha_1, \ldots, \alpha_n$ będzie skończonym układem skalarów ciała K.

Kombinacją liniową układu wektorów $x_1, \ldots, x_n$ o współczynnikach $\alpha_1, \ldots, \alpha_n$ nazywa się wektor:

$$x = \sum_{i=1}^{n} \alpha_i x_i = \alpha_1 x_1 + \ldots + \alpha_n x_n$$

Baza przestrzeni wektorowej

- Niech V będzie przestrzenią wektorową
- Układ wektorów $x_1, \ \dots, x_n \in V$ nazywamy bazą przestrzeni V jeżeli dla każdego $y \in V$ istnieje jednoznaczne przedstawienie go w formie kombinacji liniowej, tj.

$$y = \sum_{i=1}^{n} \alpha_i x_i = \alpha_1 x_1 + \ldots + \alpha_n x_n$$

Iloczyn skalarny
Kąt między wektorami

- Iloczyn skalarny, to działanie między dwoma wektorami zwracające skalar, tj. $\circ : V \times V \to K$
- Iloczyn skalarny definiujemy za pomocą zestawu aksjomatów
- Iloczyn skalarny określa nam kąt między wektorami

$$x \circ y = \|x\| \cdot \|y\| \cdot \cos(\theta)$$

- W szczególności jeśli $x \circ y = 0$, to wektory są ortogonalne, tj. $x \perp y$

Baza ortogonalna i ortonormalna

Po co nam iloczyny skalarny cz.1

- O bazie mówimy, że jest **ortogonalna** (względem iloczynu skalarnego), jeżeli dla każdej pary wektorów x_i, x_j zachodzi

$$x_i \circ x_j = \begin{cases} 0 \quad \text{gdy} \quad i \neq j \\ \|x_i\|^2 \quad \text{gdy} \quad i = j \end{cases}$$

- O bazie mówimy, że jest **ortonormalna** (względem iloczynu skalarnego), jeżeli dla każdej pary wektorów x_i, x_j zachodzi

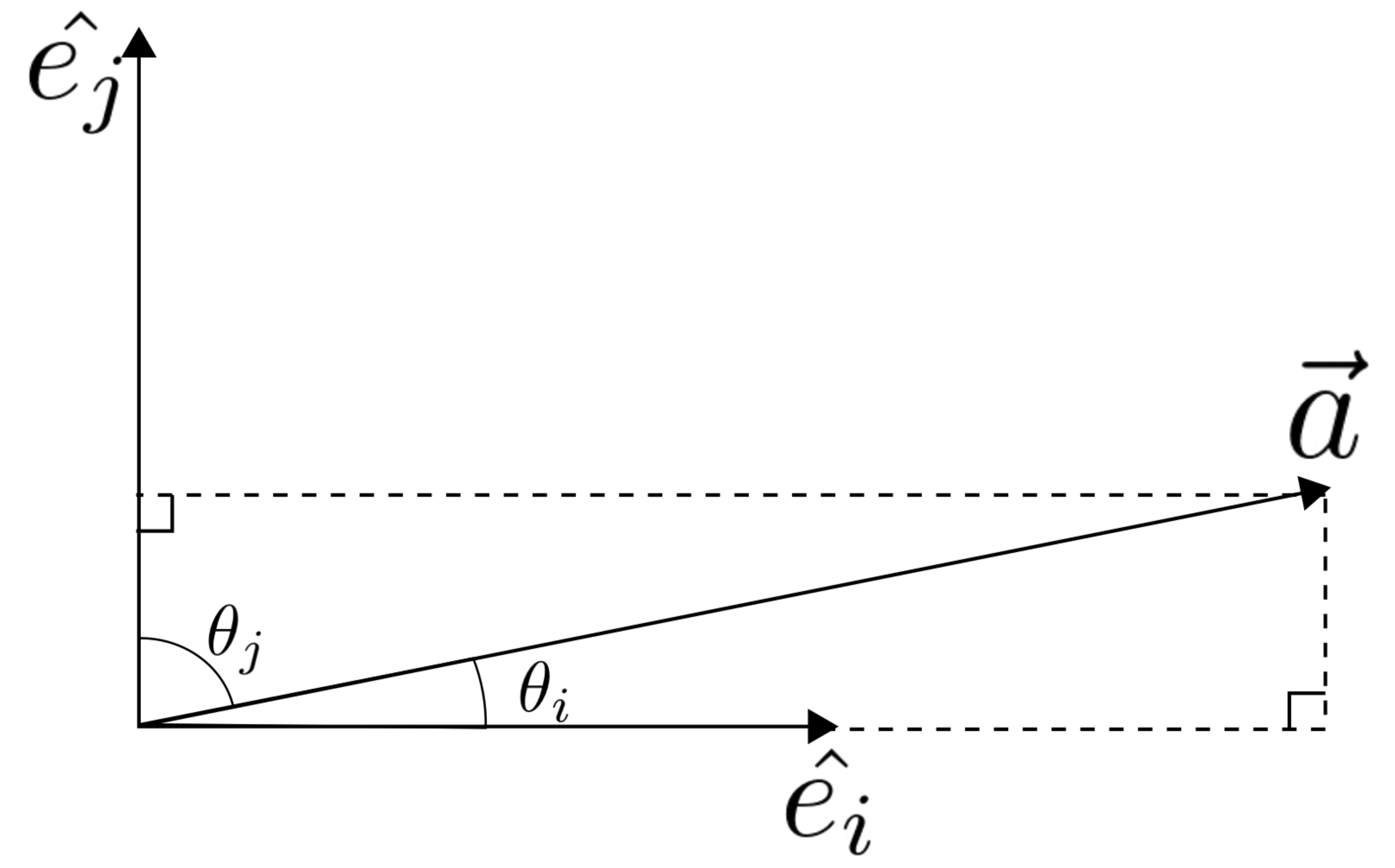
$$x_i \circ x_j = \begin{cases} 0 \quad \text{gdy} \quad i \neq j \\ 1 \quad \text{gdy} \quad i = j \end{cases}$$

Rzutowanie wektorów na bazę

Projekcja

- Niech będzie dana ortonormalna baza e_i w przestrzeni V
- Wektor $a \in V$ da się jednoznacznie przestawić jako $y = \sum_{i=1}^{n} \alpha_i e_i$ gdzie

  $\alpha_i = a \circ e_i$
- Jak baza nie będzie ortonormalna, wzory będą nieco brzydsze



Rzutowanie wektora na podprzestrzeń

Redukcja wymiaru

- Niech V będzie przestrzenią wektorową z bazą (x_i), a $V_1 \subset V$ jej podprzestrzenią
- Niech układ wektorów $(\hat{x}_i) \subset (x_i)$ będzie bazą w V_1
- Niech $y_1 \in V_1$ będzie przedstawieniem wektora $y \in V$ w bazie $(\hat{x}_i)$
- Wtedy wektor $r = y - y_1$ jest ortogonalny do bazy $(\hat{x}_i)$

Co to ma wspólnego z wielomianami?

Przestrzenie funkcyjne

- Przestrzeń funkcji jest przestrzenią wektorową nad ciałem liczb rzeczywistych
  - Można zdefiniować dodawanie funkcji
  - Można zdefiniować mnożenie funkcji przez skalar
- Oznacza to, że jeżeli przestrzeń funkcji ma przeliczalną bazę to funkcję można przedstawić w formie liniowej kombinacji funkcji bazy.
- Komplikacja - baza w przestrzeni funkcji ma nieskończenie wiele elementów

Przestrzenie Hilberta

Czyli tam gdzie wszystko działa

- Przestrzeń Hilberta jest zupełną przestrzenią metryczną (i ośrodkową)
- W przestrzeniach Hilberta mamy
  - Zdefiniowany iloczyn skalarny $x \circ y$ oraz normę $||x|| = \sqrt{x \circ x}$
  - Przeliczalne bazy
- Przestrzeniami Hilberta są np. R^n, przestrzeń funkcji całkowalnych z kwadratem na odcinku, przestrzeń ciągów sumowalnych z kwadratem i wiele innych

No a gdzie wielomiany?

Dochodzimy do konkretów

- W przestrzeniach Hilberta można zdefiniować różne bazy
- W szczególności mogą być to wielomiany
- Iloczyny skalarne definiuje się zazwyczaj jako

$$p \circ q = \int_{-1}^{1} p(x)q(x)f(x)\mathrm{d}x$$

gdzie $f(x)$ to tzw. funkcja wagowa, w szczególności może być równa 1.

Wielomiany ortogonalne

Funkcja wagowa decyduje

Znając iloczyn skalarny możemy sprawdzić ortogonalność wielomianów.

W zależności od wybranej funkcji wagowej dostajemy różne rodzaje wielomianów.

- Jeżeli $f(x) = 1$, są to wielomiany Legendre'a
- Jeżeli $f(x) = \dfrac{1}{\sqrt{1-x^2}}$, są to wielomiany Czebyszewa

Wielomiany Czebyszewa

$T_i(x)$

- Są wzory analityczne na wielomiany Czebyszewa

$$T_n(x) = \cos(n \operatorname{arc}\cos(x))$$

$$\text{lub}$$

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$

$$T_0(x) = 1$$

$$T_1(x) = x$$

Aproksymacja szeregiem Czebyszewa
Nieskończony szereg

Ponieważ wielomiany Czebyszewa tworzą bazę w oryginalnej nieskończonej przestrzeni Hilberta możemy funkcję przestawić jako

$$f(x) = \sum_{i=0}^{\infty} a_i T_i(x)$$

Współczynniki szeregu są dane wzorem

$$a_i = \frac{2}{\pi} \int_{-1}^{1} \frac{f(x) T_i(x)}{\sqrt{1 - x^2}}$$

Przy czym dla $i = 0$ współczynnik trzeba dodatkowo podzielić przez 2

Od pewnego i wartości współczynników są ograniczone z góry i malejące

Czym jest aproksymacja wielomianowa?

Podsumowanie

- Zbiór wielomianów stopnia n jest podprzestrzenią przestrzeni funkcyjnej V
- Baza wielomianów z przestrzeni V (ograniczona do stopnia n) jest bazą przestrzeni wielomianów
- Aproksymacja funkcji za pomocą skończonego szeregu wielomianów bazy jest więc rzutowaniem funkcji na podprzestrzeń wielomianów stopnia n

Skończony szereg Czebyszewa

Aproksymacja w praktyce

- Ograniczamy się do n pierwszych wielomianów

$$f(x) \approx \sum_{i=0}^{n} a_i T_i(x)$$

- Wyliczanie współczynników kłopotliwe
- Zbieżność taka sama jak w przypadku wielomianów tyle, że 2 razy szybsza (stała 2 zamiast 4)

Szereg Czebyszewa z wielomianu interpolacyjnego

Skoro tylko 2 razy, to może da się prościej

- Każdy wielomian stopnia n można przedstawić jako sumę wielomianów Czebyszewa stopni od 0 do n.
- Wielomian interpolacyjny na węzłach Czebyszewa jest więc dany wzorem

$$f(x) \approx \sum_{i=0}^{n} c_i T_i(x)$$

- Jak wyznaczyć współczynniki c_i?

Dyskretna Transformata Fouriera (DFT)

Interludium

- DFT jest jednym z podstawowych narzędzi analizy sygnałowej
- Pozwala przekształcić sygnał czasowy do postaci częstotliwościowej (zespolonej)
- W postaci liczbowej zamienia ciąg rzeczywisty $x_0, x_1, \ldots x_{N-1}$ na ciąg zespolony $X_0, X_1, \ldots X_{N-1}$ wg. wzoru

$$X_k = \sum_{n=0}^{N-1} x_n \cdot e^{-\frac{i2\pi}{N}kn}$$

Złożoność $O(N^2)$

Szybka Transformata Fouriera (FFT)

Jeden z najważniejszych algorytmów XX wieku

- Istnieje algorytm (a w zasadzie kilka) pozwalających na wyliczenie FFT w krótszym czasie - $O(N \log N)$
- Najbardziej znany jest algorytm Cooley'a-Tukey'a, który korzystając z dekompozycji $N = N_1 N_2$ redukuje problem DFT wymiaru N do wyliczenia N_1 DFT o wymiarze N_2
- Metoda ta jest stosowana rekursywnie
- Jeżeli $N = 2^M$ sytuacja się najbardziej upraszcza.

Współczynniki szeregu Czebyszewa dla wielomianu interpolacyjnego

Powrót do konkretów

- Korzystając z własności wielomianów Czebyszewa (Ahmed i Fisher, 1968) można pokazać, że

$$c_n = 2 \sum_{s=0}^{N}{}' f\left(\cos\frac{\pi s}{N}\right)\cos\frac{\pi s n}{N}$$

$'$ po sumie oznacza, że pierwszy i ostatni wyraz dzielimy przez 2.

- Alternatywnie możemy to samo zapisać jako

$$c_n = \sum_{s=0}^{2N-1} f\left(\cos\frac{\pi s}{N}\right)\exp\left(\frac{j2\pi s n}{2N}\right)$$

N. Ahmed & P.S. Fisher (1968) Study of algorithmic properties of chebyshev coefficients, International Journal of Computer Mathematics, 2:1-4, 307-317, DOI: 10.1080/00207167008803043

Współczynniki szeregu Czebyszewa dla wielomianu interpolacyjnego

FFT na ratunek

- Opracowany wzór to DFT 2 razy dłuższego ciągu - można zastosować FFT
- Nie potrzebujemy 2 razy więcej wartości, bo

$$f\left(\cos\frac{\pi(2N-s)}{N}\right) = f\left(\cos\frac{\pi s}{N}\right)$$

Obliczanie szeregu Czebyszewa w praktyce

Dobór stopnia interpolacji

Aby dokonać aproksymacji funkcji szeregiem/wielomianem interpolacyjnym Czebyszewa:

1. Wyliczamy wartości funkcji w N węzłach Czebyszewa
2. Wyznaczamy szereg Czebyszewa
3. Sprawdzamy czy współczynniki maleją do poziomu zera maszynowego, jak nie $N = 2N$. Powrót do kroku 1.
4. Jeśli tak, znajdujemy największe $n*$ takie, że $c_{n*} > \varepsilon_m$, $N = n*$