

Interpolacija merskih podatkov

Datum: 07/10/2024
Avtor: Aleksander Grm

Zapiski temeljijo na predlogi OnLine knjige *Numerične metode v ekosistemu Pythona*, Janko Slavič

Uvod

Pri **interpolaciji** izhajamo iz tabele (različnih) vrednosti x_i, y_i :

x	y
x_0	y_0
x_1	y_1
\dots	\dots
x_{n-1}	y_{n-1}

določiti pa želimo vmesne vrednosti. Če želimo določiti vrednosti zunaj območja x v tabeli, govorimo o **ekstrapolaciji**.

V okviru **interpolacije** (angl. *interpolation*) točke povežemo tako, da predpostavimo neko funkcijo in dodamo pogoj, da funkcija *mora* potekati skozi podane točke.

Pri **aproksimaciji** (angl. *approximation* ali tudi *curve fitting*) pa predpostavimo funkcijo, ki se čimbolj (glede na izbrani kriterij) prilga podatkom.

Primer

x	y
1.0	0.54030231
2.5	-0.80114362
4.0	-0.65364362

Pri interpolaciji izhajamo iz tabele vrednosti. Da bomo pozneje lahko enostavno prikazali napako, smo zgornjo tabelo generirali s pomočjo izraza $y = \cos(x)$!

Pripravimo numerični zgled; najprej uvozimo pakete:

```
In [ ]: import numpy as np
import matplotlib.pyplot as plt

Nato pripravimo tabelo samr prikaz:

In [ ]: n = 3 # number of sample points

x = np.linspace(1, 4, n)
f = np.cos # posplošna interpolirano funkcijo (lahko spremenite v drugo funkcijo)
f_line = f.__str__().split('\n')[1] # avtomatsko vzamemo ime funkcije
y = f(x) + np.random.rand(len(x))*0.3

mpl.plot(x, y, 'o', label='Interpolacijske točke');
mpl.legend();
```

Interpolacija s polinomom

Interpolacija s polinomom se zdi najbolj primerna, saj je enostavna!

Polinom stopnje $n - 1$:

$$y = a_0 x^{n-1} + a_1 x^{n-2} + \dots + a_{n-2} x + a_{n-1}.$$

Je definiran z n konstantami a_i . Da določimo n konstant, potrebujemo n (različnih) enačb. Za vsak par x_i, y_i lahko torej zapišemo:

$$y_i = a_0 x_i^{n-1} + a_1 x_i^{n-2} + \dots + a_{n-2} x_i + a_{n-1}.$$

Ker imamo podanih n parov, lahko določimo n neznanih konstant a_i , ki definirajo polinom stopnje $n - 1$. Sistem n linearnih enačb lahko zapišemo:

$$\begin{pmatrix} x_0^{n-1} & x_0^{n-2} & \dots & x_0^0 \\ x_1^{n-1} & x_1^{n-2} & \dots & x_1^0 \\ \vdots & \vdots & \ddots & \vdots \\ x_{n-1}^{n-1} & x_{n-1}^{n-2} & \dots & x_{n-1}^0 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-1} \end{pmatrix}$$

Sistem linearnih enačb zapišemo v obliki:

$$\mathbf{M} \mathbf{a} = \mathbf{b}.$$

Definirajmo matriko koeficientov \mathbf{M} :

```
In [ ]: M = np.asarray([[_**p for p in reversed(range(len(x)))] for _ in x])
M
```

Izračunamo koeficiente a_0, a_1, \dots :

```
In [ ]: resitev = np.linalg.solve(M, y)
resitev
```

Pripravimo interpolacijski polinom kot Pythonovo funkcijo:

```
In [ ]: def y_function(x, resitev):
    A = np.asarray([[_**p for p in reversed(range(len(resitev)))] for _ in x])
    return A.dot(resitev)
```

Izris interpolacijskega polinoma pri bolj gosti mreži točk:

```
In [ ]: xint = np.linspace(np.min(x), np.max(x), 30)
yint = y_function(xint, resitev)
mpl.plot(x, y, 'o', label='Interpolacijske točke')
mpl.plot(xint, yint, '-', label='Interpolacija')
mpl.legend();
```

Slabosti zgornjega postopka so:

- Število numeričnih operacij raste sorazmerno z n^3 .
- problem je lahko slabo pogojen (z večanjem stopnje polinoma slaba pogojenost naglo narašča):

```
In [ ]: np.linalg.cond(M)
```

Navodilo: vrnite se par vrstic nazaj in spremenite število interpolacijskih točk n na višjo vrednost (npr. 10).

Lagrangeva metoda

Lagrangeva metoda ne zahteva reševanja sistema enačb in je s stališča števila računskih operacij (narašča sorazmerno z n^2) boljša od predhodno predstavljene polinomske interpolacije (število operacij narašča sorazmerno z n^3), kjer smo reševali sistem linearnih enačb. Rešitev pa je seveda popolnoma enaka!

```
In [ ]: from IPython.display import YouTubeVideo
YouTubeVideo('c_TB0uXB9w', width=800, height=300)
```

Lagrangev interpolacijski polinom stopnje $n - 1$ je definiran kot:

$$L_{n-1}(x) = \sum_{i=0}^{n-1} y_i l_i(x),$$

kjer je $l_i(x)$ Lagrangev polinom stopnje i

$$l_i(x) = \prod_{j=0, j \neq i}^{n-1} \frac{x - x_j}{x_i - x_j}.$$

Poglejmo si interpolacijo za zgoraj prikazane x in y podatke.

Definirajmo najprej Lagrangev polinom, kjer imamo i znanih točk

$$l_i(x) := \prod_{j=0, j \neq i}^{n-1} \frac{x - x_j}{x_i - x_j}.$$

Primer

Imamo dane tri meritve $\{(x_0, y_0), (x_1, y_1), (x_2, y_2)\}$. V tem primeru, kjer želimo izdelati interpolacijski polinom čez celoten interval merilnih točk, imamo tri različne Lagrangeve interpolacijske polinome, ki so stopnje 2

$$l_0(x) = \frac{x - x_1}{x_0 - x_1} \frac{x - x_2}{x_0 - x_2}, \quad l_1(x) = \frac{x - x_0}{x_1 - x_0} \frac{x - x_2}{x_1 - x_2}, \quad l_2(x) = \frac{x - x_0}{x_2 - x_0} \frac{x - x_1}{x_2 - x_1},$$

kar pomeni, da lahko skozi tri točke napejemo Lagrangejev polinom stopnje 2.

Zapišimo splošen Lagrangejev polinom $l_i(x)$ v obliki Python funkcije:

```
In [ ]: def lagrange(x, x_int, i):
    """ Vrne vrednosti i-tega Lagrangevega polinoma

    x: neodvisna spremenljivka
    x_int: seznam interpolacijskih točk
    i: indeks polinoma
    """
    Lx = 1.0
    for j in range(len(x_int)):
        if j != i:
            Lx *= (x-x_int[j]) / (x_int[i]-x_int[j])

    return Lx
```

in izris lepege grafa:

```
In [ ]: def slika(i=0):
    xint = np.linspace(np.min(x), np.max(x), 30)
    mpl.plot(x, y, 'o', label='Interpolacijske točke')
    mpl.axhline(0, color='k', linewidth=0.3);
    mpl.plot(xint, lagrange(xint, x_int=x, i=i), '-', label=f'Lagrangev polinom i={i}');
    for _ in x:
        mpl.axhline(_, color='r', linewidth=0.5);
    mpl.legend();
    mpl.show()

In [ ]: slika(i=1)
```

Opazimo, da ima i -ti Lagrangev polinom v x_i vrednost 1, v ostalih podanih točkah pa nič!

Če torej Lagrangev polinom za $i = 0$ pomnožimo z y_0 , bomo pri $x = x_0$ dobili pravo vrednost, v ostalih interpolacijskih točkah pa nič; implementirajmo torej Lagrangev interpolacijski polinom:

$$L_{n-1}(x) = \sum_{i=0}^{n-1} y_i l_i(x),$$

```
In [ ]: def lagrange_interpolacija(x, x_int, y_int):
    """ Vrne vrednosti Lagrangeve interpolacije

    x: neodvisna spremenljivka (skalar ali numerično polje)
    x_int: abscisa interpolacijskih točk
    y_int: ordinata interpolacijskih točk
    """
    y = 0.
    for i in range(len(x_int)):
        Lx = 1.0
        for j in range(len(x_int)):
            if j != i:
                Lx *= (x-x_int[j]) / (x_int[i]-x_int[j])
        y += y_int[i] * Lx
    return y
```

Pripravimo sliko:

```
In [ ]: def slika(i=0):
    xint = np.linspace(np.min(x), np.max(x), 30)
    mpl.plot(x, y, 'o', label='Interpolacijske točke')
    mpl.plot(xint, lagrange(xint, x_int=x, i=i), '-', label=f'Lagrangev polinom i={i}');
    mpl.plot(xint, lagrange_interpolacija(xint, x_int=x, y_int=y), '-', label=f'Lagrangev int polinom');
    mpl.axhline(0, color='k', linewidth=0.3);
    for _ in x:
        mpl.axhline(_, color='r', linewidth=0.5);
    mpl.legend();
    mpl.show()

In [ ]: slika(i=1)
```

Iz `ipywidgets` uvozimo `Interact`, ki je močno orodje za avtomatsko generiranje (preprostega) uporabniškega vmesnika znotraj `Jupyter` okolja. Tukaj bomo uporabili relativno preprosto interakcijo s sliko; za pregled vseh zmožnosti pa radovednega bralca naslavljamo na dokumentacijo.

Uvoz funkcije `Interact`

```
In [ ]: from ipywidgets import interact
```

```
In [ ]: interact(slika);
```

Iz slike vidimo, da ima Lagrangev polinom i samo pri x_i vrednost 1 v ostalih točkah $\neq i$ pa ima vrednost nič; za Lagrangev polinom $l_i(x)$ pomnožimo z y_i zadostimo i -ti točki iz tabele. Posledično Lagrangeva interpolacija z vsoto Lagrangevih polinomov interpolira tabelo.

Polinomska interpolacija pri velikem številu točk je lahko slabo pogojena naloga in zato jo odsvetujemo.

Ocena napake

Če je $f(x)$ funkcija, ki jo interpoliramo in je $P_{n-1}(x)$ interpolacijski polinom stopnje $n - 1$, potem se lahko pokaže (glejte npr.: Burden, Faires, Burden: Numerical Analysis), da je napaka interpolacije s polinomom:

$$e = f(x) - P_{n-1}(x) = \frac{f^{(n)}(\xi)}{n!} (x - x_0)(x - x_1) \dots (x - x_{n-1}),$$

kjer je $f^{(n)}$ odvod funkcije, $n - 1$ stopnja interpolacijskega polinoma in ξ vrednost na interpoliranem intervalu $[x_0, x_{n-1}]$.

Zgled

Tukaj si bomo ogledali interpolacijo točk:

```
In [ ]: x = np.array([0., 1., 2.14285714, 3.28571429, 4.42857143, 5.57142857,
                  6.71428571, 7.85714286, 9., 10.])
y = np.array([0., 0.5, 0.52359878, 1.04719755, 1.57079633],
              [0., 0.5, 0.5600234, 1., 1.5])

Točke prikazimo:

In [ ]: mpl.plot(x, y, 'o');

Linearna interpolacija za vrednost pri x=1,57079633/2:
```

```
In [ ]: y_linearna = lagrange_interpolacija(x=[1]/2, x_int=x[1:3], y_int=y[1:3])
y_linearna
```

Kvadratna:

```
In [ ]: y_kvadratna = lagrange_interpolacija(x=[1]/2, x_int=x[0:3], y_int=y[0:3])
y_kvadratna
```

Kubična

```
In [ ]: y_kubična = lagrange_interpolacija(x=[1]/2, x_int=x, y_int=y)
y_kubična
```

Zgled ocene napake

Pri interpolaciji ponavadi funkcije $f(x)$ ne poznamo in napako ocenimo s pomočjo formule:

$$e = \frac{f^{(n)}(\xi)}{n!} (x - x_0)(x - x_1) \dots (x - x_{n-1}).$$

Pri tem vrednost ξ ni znana; ker je v primeru linearne interpolacije ($n = 2$) drugi odvod sinusne funkcije ($f^{(n)}$) med -1 in $+1$, velja:

$$|e| \leq \frac{-1}{2!} (\pi/4 - \pi/6) (\pi/4 - \pi/3) = \frac{1}{2} \frac{\pi}{12} \frac{\pi}{12} = \frac{\pi^2}{288} = 0.034$$

Poleg **Lagrangeve metode** bi si tukaj lahko pogledali še **Newtonovo metodo** interpolacije.

Interpolacija z uporabo `scipy`

Poglejmo si interpolacijo v okviru modula `scipy.interpolate` (dokumentacija).

Uporabili bomo funkcijo za interpoliranje tabele z zlepk, `scipy.interpolate.interp1d` (dokumentacija):

`Interp1d(x, y, kind='linear', axis=-1, copy=True, bounds_error=None, fill_value=nan, assume_sorted=False)`
Podati moramo vsaj dva parametra: seznama interpolacijskih točk x in y . Privzeti parameter `kind='linear'` pomeni, da interpoliramo z odsekom linearno funkcijo. `interp1d` vrne funkcijo `f`, ki jo kličemo (npr. `y = f(x)`) za izračun interpolirane vrednosti.

Parameter `kind` ima lahko npr. tudi: `'zero'`, `'linear'`, `'quadratic'` in `'cubic'`; takrat se uporabi interpolacijski zlepek (ang. *spine*) reda 0, 1, 2 oz. 3. Zlepke si bomo pogledali v naslednjem poglavju.

```
In [ ]: from scipy.interpolate import interp1d
```

Definirajmo tabelo podatkov:

```
In [ ]: x = np.array([1., 2.14285714, 3.28571429, 4.42857143, 5.57142857,
                  6.71428571, 7.85714286, 9., 10.])
y = np.array([0.52359878, 1.04719755, 1.57079633, 2.09461541, 2.61745445,
              3.14159265, 3.66519143, 4.18879181, 4.71238918])

len(x)

In [ ]: from scipy.interpolate import interp1d
```

```
In [ ]: f = interp1d(x, y, kind='linear')
x_g = np.linspace(0, 10, 11); 20*(len(x)-1)
mpl.plot(x, y, 'o', label='Interpolacijske točke')
mpl.plot(x_g, f(x_g), '-', label='Kubični zlepek')
mpl.plot(x_g, np.sin(x_g), label='sin', alpha=0.5)
mpl.legend();
```

Kubični zlepi

Preden gremo v teorijo zlepkov, si pogledimo rezultat, ki ga dobimo s pomočjo formule `interp1d` s parametrom `kind='cubic'` (rezultat je kubični zlepek).

```
In [ ]: f = interp1d(x, y, kind='cubic')
x_g = np.linspace(0, 10, 11); 20*(len(x)-1)
mpl.plot(x, y, 'o', label='Interpolacijske točke')
mpl.plot(x_g, f(x_g), '-', label='Kubični zlepek')
mpl.plot(x_g, np.sin(x_g), label='sin', alpha=0.5)
mpl.legend();
```

Kubični zlepek so pogost način interpolacije.

Zahtevamo, da je: $x_0 < x_1 < \dots < x_n$.

Od točke x_i do x_{i+1} naj bo zlepek polinom:

$$f_{i,i+1}(x) = a_{i,3}x^3 + a_{i,2}x^2 + a_{i,1}x + a_{i,0},$$

pri čemer so neznane vrednosti konstant $a_{i,j}$.

Če imamo na primer $n + 1$ točk, potem je treba določiti n polinomov.

Celotni zlepek čez $n + 1$ točk je definiran z:

$$f(x) = \begin{cases} f_{i,0}(x); & x \in [x_0, x_1) \\ f_{i,1}(x); & x \in [x_1, x_2) \\ \vdots \\ f_{n-1,n}(x); & x \in [x_{n-1}, x_n] \end{cases}$$

Vsak polinom $f_{i,i+1}$ je definiran s 4 konstantami $a_{i,j}$; skupaj torej moramo izračunati $4n$ konstant $a_{i,j}$.

Kako določimo konstante $a_{i,j}$?

Za določitev $4n$ neznak potrebujemo $4n$ enačb. Poglejmo si, kako jih dobimo:

- n enačb dobimo iz interpolacijskega pogoja:

$$y_i = f_{i,i+1}(x_i), \quad i = 0, 1, 2, \dots, n - 1$$

- 1 enačbo iz zadnje točke:

$$y_n = f_{n-1,n}(x_n)$$

- $3(n - 1)$ enačb dobimo iz pogoja C^2 zveznosti:

$$\lim_{x \rightarrow x_i^-} f(x) = \lim_{x \rightarrow x_i^+} f(x),$$

$$\lim_{x \rightarrow x_i^-} f'(x) = \lim_{x \rightarrow x_i^+} f'(x)$$

in

$$\lim_{x \rightarrow x_i^-} f''(x) = \lim_{x \rightarrow x_i^+} f''(x).$$

Skupaj imamo definiranih $4n - 2$ enačbi, manjkata torej še dve!

Različni tipi zlepkov se ločijo po tem, kako ti dve enačbi določimo. V nadaljevanju si bomo pogledali naravne kubične zlepeke.

Naravni kubični zlepi

Naravni kubični zlepi temeljijo na ideji Eulerjevega nosilca:

$$EI \frac{d^4 y}{dx^4} = q(x),$$

kjer je E elastični modul, I drugi moment preseka in $q(x)$ zunanja porazdeljena sila. Ker zunanje porazdeljene sile ni ($q(x) = 0$), velja:

$$EI \frac{d^4 y}{dx^4} = 0.$$

Sledi, da lahko v vsaki točki tanek nosilec popišemo s polinomom tretje stopnje.

C^2 zveznost je zagotovljena v kolikor so vmesne podpore nosilca členki (moment zato nima nezvezne spremenbe).

Manjkajoči 2 neznanjki pri naravnih kubičnih zlepkih določimo iz pogoja, da je moment na koncih enak nič (členkasto vpetje):

$$f''(x_0) = 0 \quad \text{in} \quad f''(x_n) = 0$$

Izpeljava je natančneje prikazana v knjigi Kussalaas J: Numerical Methods in Engineering with Python 3, 2013, stran 120 (glejte tudi J. Petričič: Interpolacija, Fakulteta za strojništvo, 1999); podrobna izpeljava presega namen te knjige.

Tukaj si bomo pogledali samo končni rezultat, ki ga lahko izpeljemo od zgornjih pogojev. V primeru ekvidistantne delitve $h = x_{i+1} - x_i$ tako izpeljemo sistem enačb ($i = 1, \dots, n - 1$):

$$k_{i-1} + 4k_i + k_{i+1} = \frac{6}{h^2} (y_{i-1} - 2y_i + y_{i+1}),$$

kjer je neznanaka k_i drugi odvod odsekovne funkcije $k_i = f''_{i,i+1}(x_i)$.

Rešivj sistem enačb dobimo, če dodamo še robna pogoja za naravne kubične zlepeke:

$$k_0 = k_n = 0.$$

Ko določimo neznanke k_i , jih uporabimo v odsekom definirani funkciji:

$$f_{i,i+1}(x) = \frac{k_i}{6} \left(\frac{(x - x_{i+1})^3}{h} - (x - x_{i+1}) \right) - \frac{k_{i+1}}{6} \left(\frac{(x - x_i)^3}{h} - (x - x_i) \right) + \frac{y_i (x - x_{i+1}) - y_{i+1} (x - x_i)}{h}.$$

Numerična implementacija

Najprej pripravimo funkcijo, katera za podane interpolacijske točke reši sistem linearnih enačb in vrne koeficiente k_i :

```
In [ ]: def kubicni_zlepki_koeficient(x, y):
    """ Vrne koeficiente kubičnih zlepkov 'k', matriko koeficientov 'A' in konstant.

    x in y predstavljata seznam znanih vrednosti; x mora biti ekvidistent.
    """
    n = len(x)
    A = np.zeros((n, n)) # pripravimo matriko koeficientov
    b = [1/x[i]*8] * korak h
    for i in range(n):
        if i==0 or i==n-1:
            A[i,i] = 1. + k_0 in k_n sta nič zato tukaj damo 1
            # pri vektorju konstant pa bomo dali 0, k_0 in k_n bosta torej 0
        else:
            A[i,i-1:i+2] = np.asarray([1., 4., 1.])
    b = np.zeros(n)
    b[1:-1] = (6/h**2)*(y[i:-2] - 2*y[i-1] + y[i+2]) # desna stran zgornje enačbe
    k = np.linalg.solve(A,b)
    return k, A, b

Opomba: pri zgornjem linearnem problemu, lahko izračun zelo pohitimo, če upoštevamo tridiagonalnost matrike koeficientov! (glejte odgovor na stackoverflow.com).

Poglejmo si primer izračuna koeficientov:

In [ ]: x = np.asarray([1, 2, 3, 4, 5])
y = np.asarray([0, 1, 0, 1, 0])

k, A, b = kubicni_zlepki_koeficient(x, y)
print('Matrika koeficientov A lin. sistema:', A)
print('Vektor konstant b lin. sistema: ', b)
print('Koeficienti k so:', k)
```

Nato potrebujemo še kubični polinom v določenem intervalu; implementirajmo izraz:

$$f_{i,i+1}(x) = \frac{k_i}{6} \left(\frac{(x - x_{i+1})^3}{h} - (x - x_{i+1}) \right) - \frac{k_{i+1}}{6} \left(\frac{(x - x_i)^3}{h} - (x - x_i) \right) + \frac{y_i (x - x_{i+1}) - y_{i+1} (x - x_i)}{h}$$

```
In [ ]: def kubicni_zlepki(k, x, y, x_najdi):
    """ Vrne kubični zlepek pri določeni 'xint'.

    param k: koeficienti kubičnih zlepkov
    param x in y: znane vrednosti, x mora biti ekvidistent
    param x_najdi: vrednosti, kjer želimo izračunati kubični zlepek
    """
    h = x[0] - x[1]
    i = int((x_najdi-x[0])/((-h)))
    if i != len(k)-1:
        i = len(k)-2
    out = ((x_najdi - x[i+1])**3/h - (x_najdi - x[i+1])*h)*k[i]/6.0 \
          - ((x_najdi - x[i])**3/h - (x_najdi - x[i])*h)*k[i+1]/6.0 \
          + (y[i]*(x_najdi - x[i+1]) \
            + y[i+1]*(x_najdi - x[i]))/h
    return out
```

Izračunamo interpolirane vrednosti:

```
In [ ]: xint = np.linspace(np.min(x), np.max(x), 50)
yint = np.asarray(kubicni_zlepki(k, x, y, _ in xint))

In [ ]: mpl.plot(x, y, 'o', label='Interpolacijske točke')
mpl.plot(xint, yint, label='Naravni kubični zlepek')
mpl.legend();
```

- ### Nekaj vprašanj za razmislek!
- Preštudirajte Lagrangevo polinomsko interpolacijo in pripravite funkcijo za Lagrangeve polinome. Pojasnite (z grafičnim prikazom) Lagrangeve polinome.
 - Definirajte funkcijo za Lagrangevo polinomsko interpolacijo. Na primeru pojasnite, kako deluje.
 - Pojasnite teoretično ozadje naravnih kubičnih zlepkov.
 - Naravne kubične zlepeke smo izpeljali pod pogojem, da momenta na koncu ni, včasih želimo drugačnega pogoja na koncih (npr. znani naklon ali znani momenti).
Modificirajte na predvajanje predstavljeni kodo za primer, da je na koncih moment $\neq 0$ (predpostavite neko numerično vrednost).
 - Podatke:

```
x = np.linspace(0, 10, 10)
y = np.random.rand(10)*0.5
```

interpolirajte z uporabo `scipy.interpolate.UnivariateSpline`. Podatke prikžite.
 - Za zgoraj definirane podatke preučite pomoč in najдите vse ničle. Prikžite jih na predhod