Numerička analiza

18. predavanje

Autori: Saša Singer i Nela Bosner

Predavač: Tina Bosner

tinab@math.hr

web.math.hr/~nela/nad.html

PMF – Matematički odsjek, Zagreb

Sadržaj predavanja

- Besselove funkcije i Millerov algoritam:
  - Opći oblik Millerovog algoritma.
  - Rekurzija za Besselove funkcije i stabilnost.
  - Millerov algoritam za Besselove funkcije.
- Izglađivanje podataka i metoda najmanjih kvadrata:
  - Veza izglađivanja i najmanjih kvadrata.
  - Globalno i lokalno izglađivanje.
  - Diskretni ortog. polinomi i Forsytheov algoritam.
  - Lokalno izglađivanje — najmanji kvadrati, integral.

Besselove funkcije i Millerov algoritam

Tročlane rekurzije i izvrednjavanje funkcija

Kod generalizirane Hornerove sheme koristili smo znanje p_0 i p_1 za silaznu varijantu algoritma za računanje p_N.

To nije potrebno: dovoljno je znati neku vezu među funkcijama p_n koja se lako računa.

Primjer. Funkcije izvodnice oblika

$$F(x) = \sum_{n=0}^{\infty} q_n \, p_n(x),$$

gdje se $F(x)$ računa nekom analitičkom formulom bez upotrebe $p_n(x)$, tj. $F(x)$ možemo naći neovisno o funkcijama p_n.

Općenito o Millerovom algoritmu

Millerov algoritam (po J. C. P. Milleru, 1954. godine) primjenjuje se kada

- vrijednosti funkcija $p_n(x)$ vrlo brzo padaju, kad n raste, a greška zaostaje.

Pretpostavimo da funkcije p_n zadovoljavaju neku homogenu rekurziju, na primjer tročlanu,

$$p_{n+1}(x) + \alpha_n(x)p_n(x) + \beta_n(x)p_{n-1}(x) = 0, \quad n = 1, 2, \ldots$$

Poznavanje bilo kojeg p_n (čak ni p_0, niti p_1) nije potrebno. Treba znati samo koeficijente α_n i β_n.

Opća forma Millerovog algoritma

Odaberimo startnu vrijednost indeksa M od koje ćemo početi — ovisno o vrijednosti N indeksa funkcije koju tražimo:

- ako tražimo $p_N(x)$ (ili $p_N(x), \ldots, p_0(x)$, ili samo neke od njih), M se obično odabere tako da je $M > N$ i vrijedi

  $$\frac{p_M(x)}{p_N(x)} \approx \text{točnost računanja.}$$

  To obično garantira i da je

  $$F_M(x) := \sum_{n=0}^{M} q_n \, p_n(x),$$

  barem jednako točna aproksimacija za $F(x)$.

Opća forma Millerovog algoritma (nastavak)

Definiramo $\tilde{p}_{M+1} = 0$, $\tilde{p}_M = 1$ i računamo $\tilde{p}_n$, za $n = M-1, \ldots, 0$, unatrag po rekurziji za $p_n(x)$:

$$\tilde{p}_n = \frac{-(\alpha_{n+1}(x)\tilde{p}_{n+1} + \tilde{p}_{n+2})}{\beta_{n+1}(x)}, \quad n = M-1, \ldots, 0.$$

Zbog homogenosti rekurzije, dobiveni niz vrijednosti

$$\tilde{p}_M, \ldots, \tilde{p}_0$$

je vrlo približno proporcionalan stvarnim vrijednostima

$$p_M(x), \ldots, p_0(x),$$

barem u području od $p_N(x)$ do $p_0(x)$, tj. vrijedi $p_n(x) \approx \tilde{p}_n \cdot c$, za $n \leq N$. Treba još naći normalizacijski faktor c.

Opća forma Millerovog algoritma (nastavak)

Budući da znamo koeficijente q_n u razvoju funkcije izvodnice F po funkcijama p_n

$$F(x) = \sum_{n=0}^{\infty} q_n \, p_n(x),$$

umjesto nepoznatih vrijednosti $p_n(x)$, uvrstimo $\tilde{p}_n$ i računamo aproksimaciju $\tilde{F}_M$

$$\widetilde{F}_M := \sum_{n=0}^{M} q_n \, \tilde{p}_n.$$

Gornji indeks sumacije može biti i bitno manji od M, ako znamo da $p_n(x)$ vrlo brzo padaju, kad n raste.

Opća forma Millerovog algoritma (nastavak)

U sumi za $\widetilde{F}_M$ dovoljno je uzeti toliko članova da se izračunata vrijednost $\widetilde{F}_M$

- stabilizira na točnost računala ili traženu točnost.

Zatim direktno analitički izračunamo $F(x)$ po poznatoj formuli i stavimo

$$c := \frac{F(x)}{\widetilde{F}_M},$$

što je traženi normalizacijski faktor, uz pretpostavku da je $F_M(x)$ dovoljno dobra aproksimacija za $F(x)$.

Opća forma Millerovog algoritma (nastavak)

Na kraju izračunamo

$$p_n(x) = \tilde{p}_n \cdot c$$

za sve one n između 0 i N koji nas zanimaju, jer u tom području vrijedi vrlo dobra proporcionalnost $p_n(x) \sim \tilde{p}_n$.

Vrlo često se startna vrijednost M određuje iz:

- nekih poznatih relacija za familiju funkcija $p_n(x)$, ili
- eksperimentalno, povećavanjem M, sve dok se ne postigne željena točnost za $p_N(x)$.

Općenito o Besselovim funkcijama

Besselove funkcije prvi puta je uveo Bessel, 1824. godine, promatrajući jedan problem iz tzv. dinamičke astronomije, vezan uz zgodan način zapisa položaja planeta koji se kreće po elipsi oko Sunca.

Za praktične potrebe, Bessel je traženu veličinu prikazao kao red funkcija poznatih podataka s koeficijentima koji su funkcije oblika

$$J_n(x) = \frac{1}{\pi} \int_0^{\pi} \cos(x \sin\theta - n\theta)\, d\theta, \quad n \in \mathbb{N}.$$

Te funkcije zovemo Besselovim funkcijama prve vrste.

Općenito o Besselovim funkcijama (nastavak)

Definicija Besselovih funkcija može se proširiti na $n \in \mathbb{Z}$, i tada je

$$J_{-n}(x) = (-1)^n J_n(x).$$

Nažalost, ni za jednu od ovih funkcija ne postoji neka jednostavna "formula" ili oblik za računanje.

Definicijsku relaciju možemo iskoristiti i za numeričko računanje vrijednosti $J_n(x)$, za zadane $n \in \mathbb{N}_0$ i $x \in \mathbb{R}$, tako da upotrijebimo neku od metoda numeričke integracije.

Međutim, postoje i mnogo brži algoritmi za postizanje iste tražene točnosti izračunate vrijednosti $J_n(x)$.

Još jedna definicija Besselovih funkcija

U klasičnom pristupu, preko funkcija izvodnica, Besselove funkcije definiramo kao koeficijente uz t^n, u razvoju

$$\exp\left(x\,\frac{t - 1/t}{2}\right) = \sum_{n=-\infty}^{\infty} J_n(x)\,t^n.$$

Ova definicija ekvivalentna je integralnoj i iz nje se mogu izvesti mnoge važne relacije za Besselove funkcije.

Besselove funkcije zadovoljavaju tročlanu rekurziju

$$J_{n+1}(x) - \frac{2n}{x} J_n(x) + J_{n-1}(x) = 0, \quad n \in \mathbb{N}.$$

Također, vrijedi $J_1(x) = -J_0'(x)$.

Izračunavanje Besselovih funkcija

Razmišljanja . . .

- Kad bismo znali izračunati $J_0(x)$ i $J_1(x)$ (ili $J_0'(x)$), onda bismo mogli izračunati i $J_n(x)$.
- Osim toga, generaliziranom Hornerovom shemom mogli bismo onda računati i razne razvoje po Besselovim funkcijama.
- Nažalost, tročlana rekurzija za Besselove funkcije je izrazito nestabilna unaprijed.
- Da bismo to pokazali, promotrimo ponašanje vrijednosti Besselovih funkcija $J_n(x)$ u ovisnosti o n i x.

Diferencijalna jednadžba i Besselove funkcije

Iz funkcije izvodnice pokazuje se da Besselove funkcije J_n, zadovoljavaju diferencijalnu jednadžbu

$$x^2 y'' + xy' + (x^2 - n^2)y = 0,$$

koja se može promatrati na cijeloj kompleksnoj ravnini, pa se u slučaju kad n nije cijeli broj koristi oznaka ν.

Jedno od rješenja ove jednadžbe su Besselove funkcije prve vrste J_ν, koje imaju svojstvo da su ograničene u 0 kad je $\operatorname{Re}\nu \geq 0$. Analitički im je oblik

$$J_\nu(x) = \left(\frac{1}{2}x\right)^\nu \sum_{k=0}^{\infty} \frac{(-x^2/4)^k}{k!\,\Gamma(\nu + k + 1)},$$

gdje su ν i x, općenito, kompleksni brojevi.

Izračunavanje Besselovih funkcija

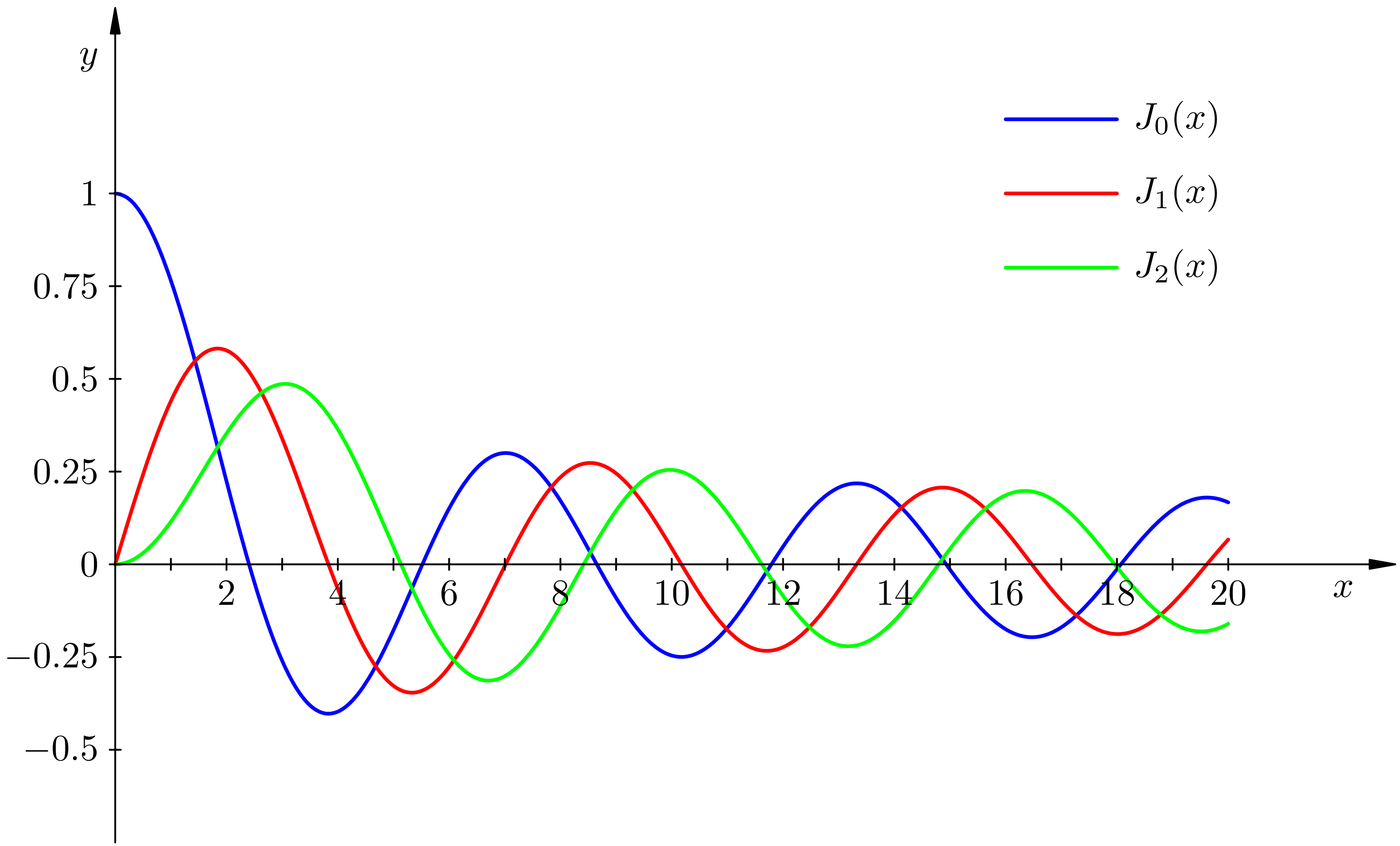
Promatrajmo ponašanje ovih funkcija samo za nenegativne realne indekse $\nu \geq 0$ i argumente $x \geq 0$.

Ako je ν cijeli broj, onda prethodna relacija glasi

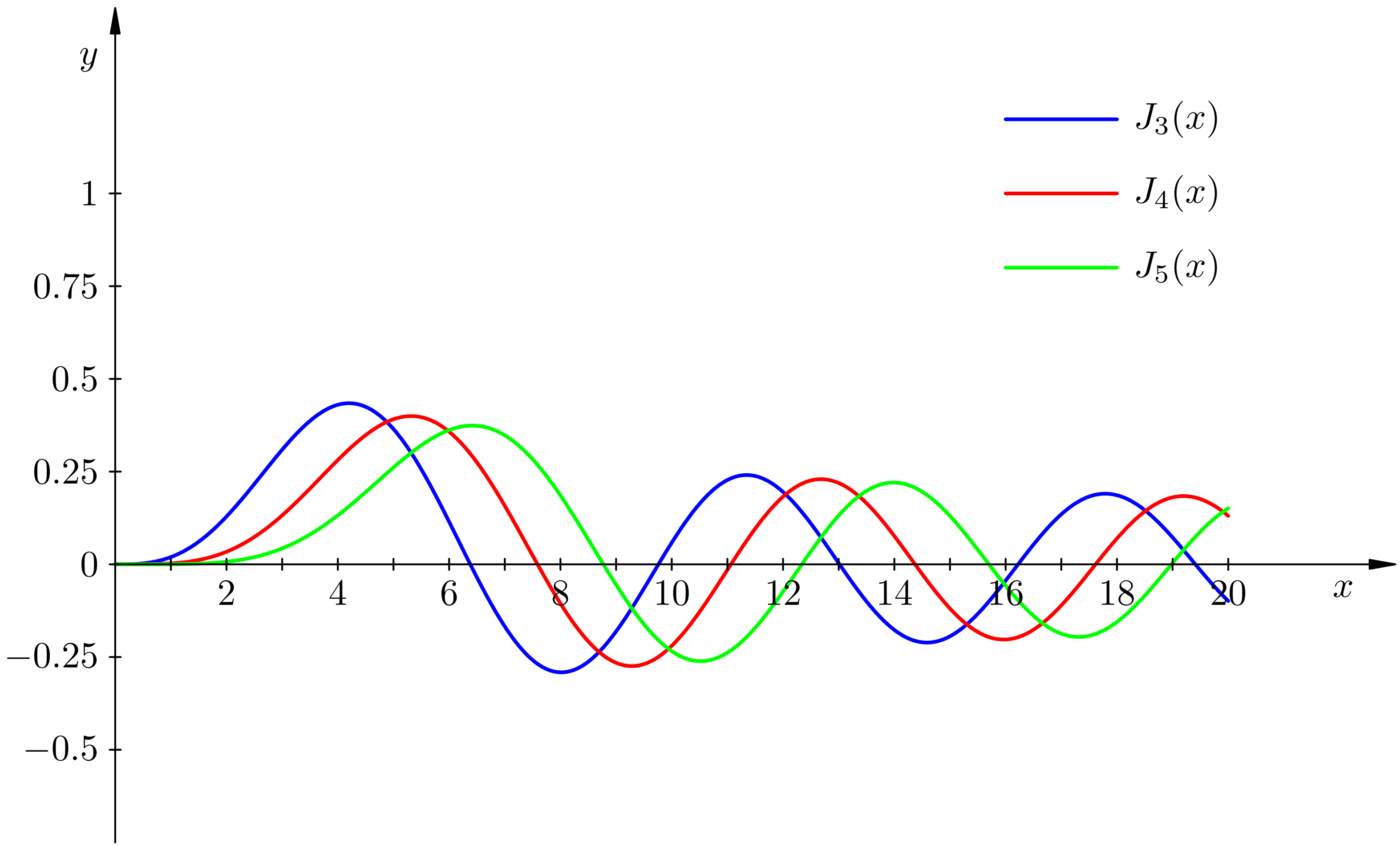
$$J_n(x) = \left(\frac{1}{2}x\right)^n \sum_{k=0}^{\infty} \frac{(-x^2/4)^k}{k!\,(n+k)!}.$$

- Oba reda očito konvergiraju za $x \geq 0$, a donji čak na cijelom skupu $\mathbb{C}$.
- Na prvi pogled izgleda kao da smo time riješili i problem računanja vrijednosti $J_0(x)$ i $J_1(x)$.
- Ovaj red vrlo brzo konvergira za relativno male x.
- Za malo veće x, kad je $x \approx n$ (ili ν) dobivamo sve veće kraćenje zbrajanjem uzastopnih članova reda.

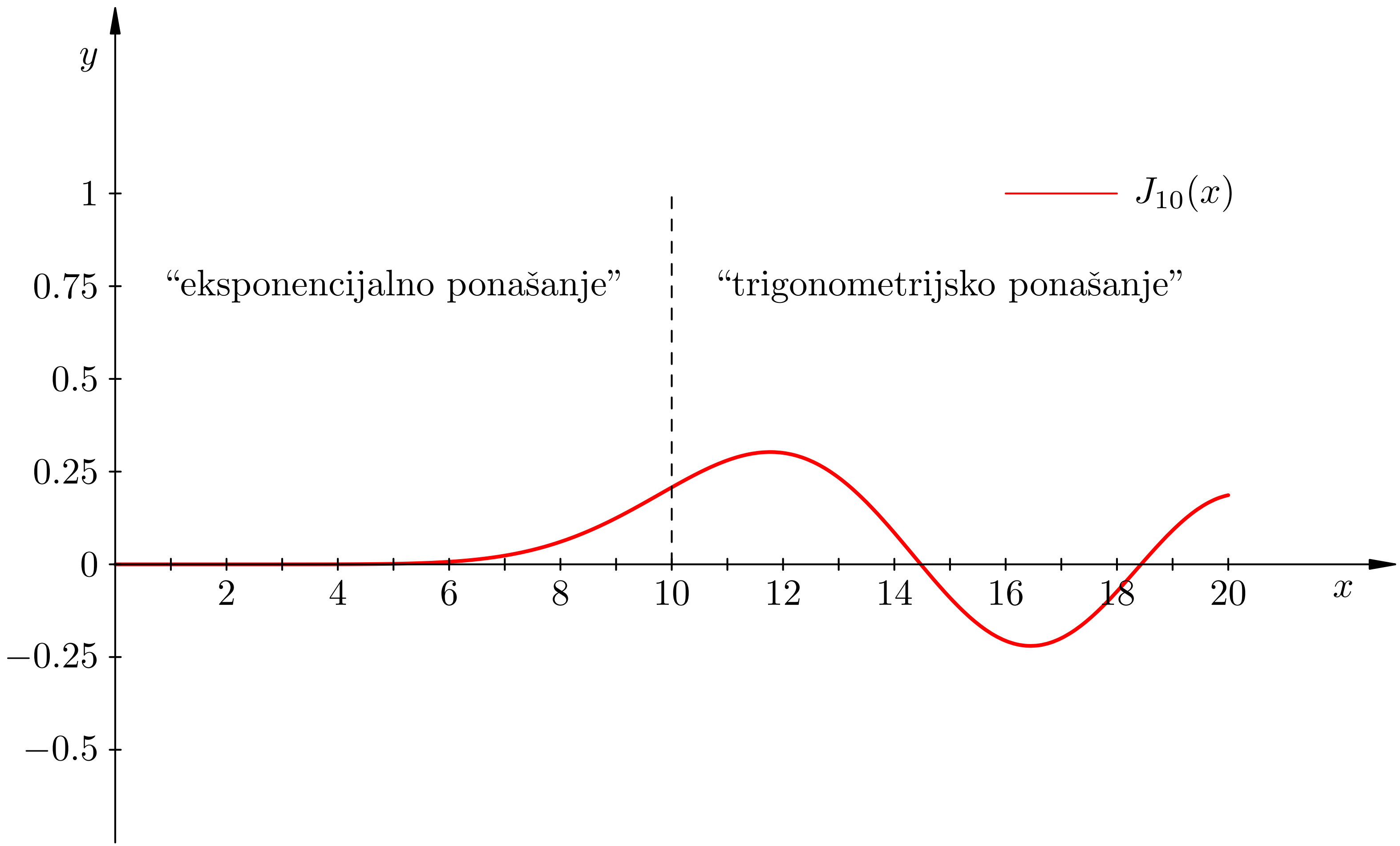
Prve tri Besselove funkcije



Sljedeće tri Besselove funkcije



Deseta Besselova funkcija



Besselove funkcije – drugi pogled

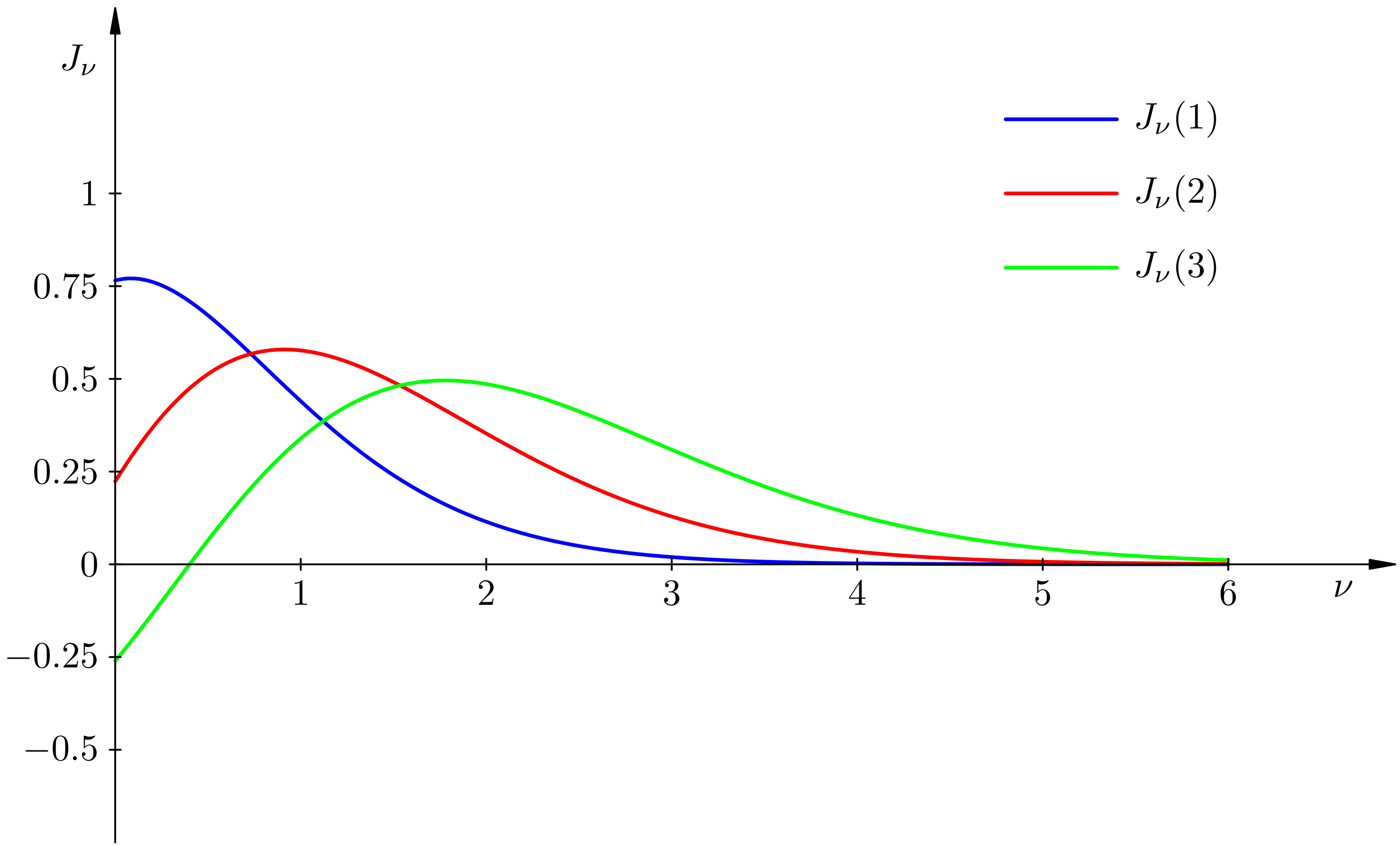
Ako gledamo kao funkcije od x,

- područje eksponencijalnog ponašanja mijenja se u trigonometrijsko područje približno za $x = \nu$ (na pr., iz Taylorovog reda).

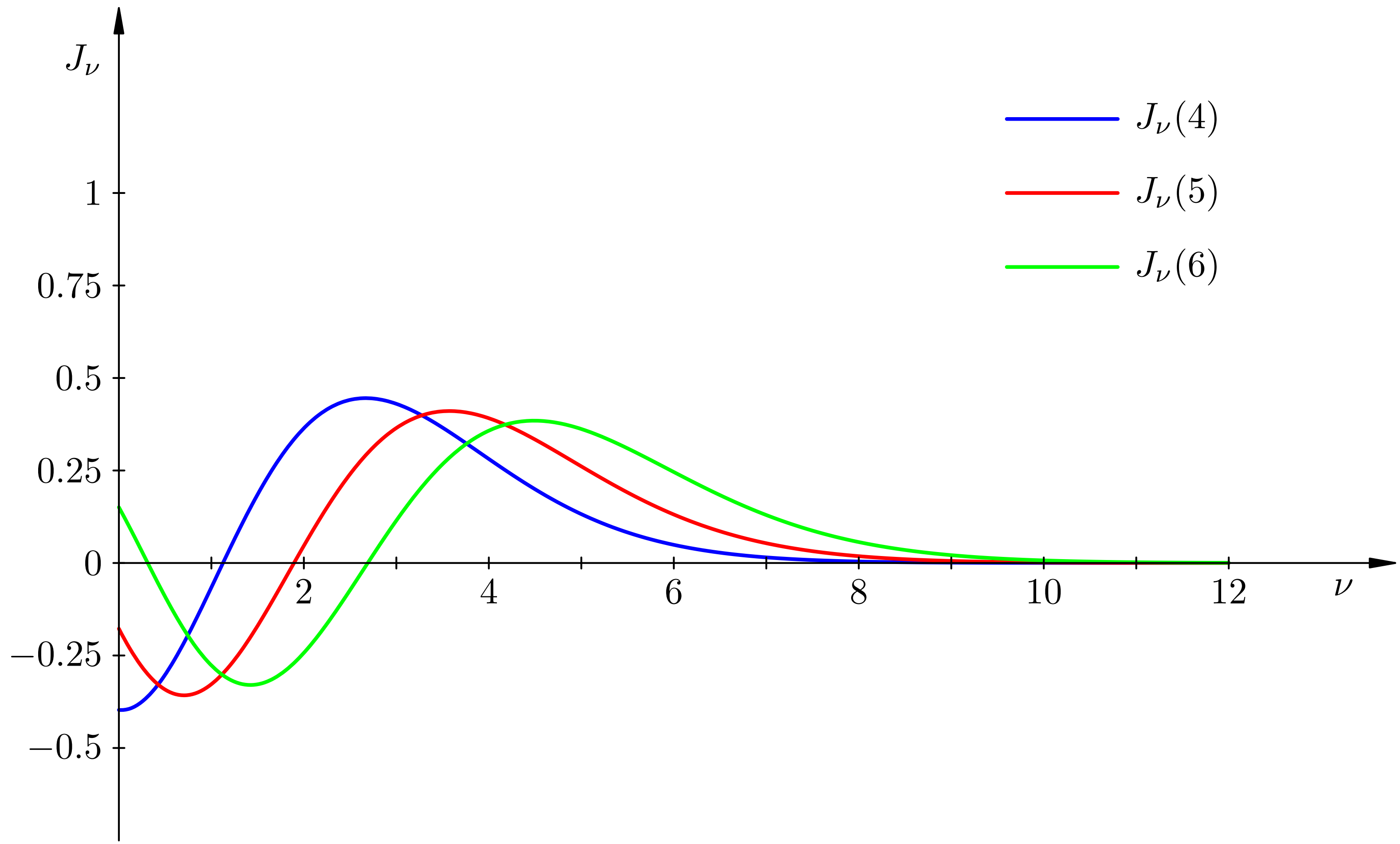
Gledamo li Besselove funkcije, ne kao funkcije od x, nego za fiksni x, kao funkcije indeksa ν, onda Besselove funkcije

- pokazuju slično ponašanje, samo po ν,
- područje trigonometrijskog ponašanja za $x \approx \nu$ trne u eksponencijalno.

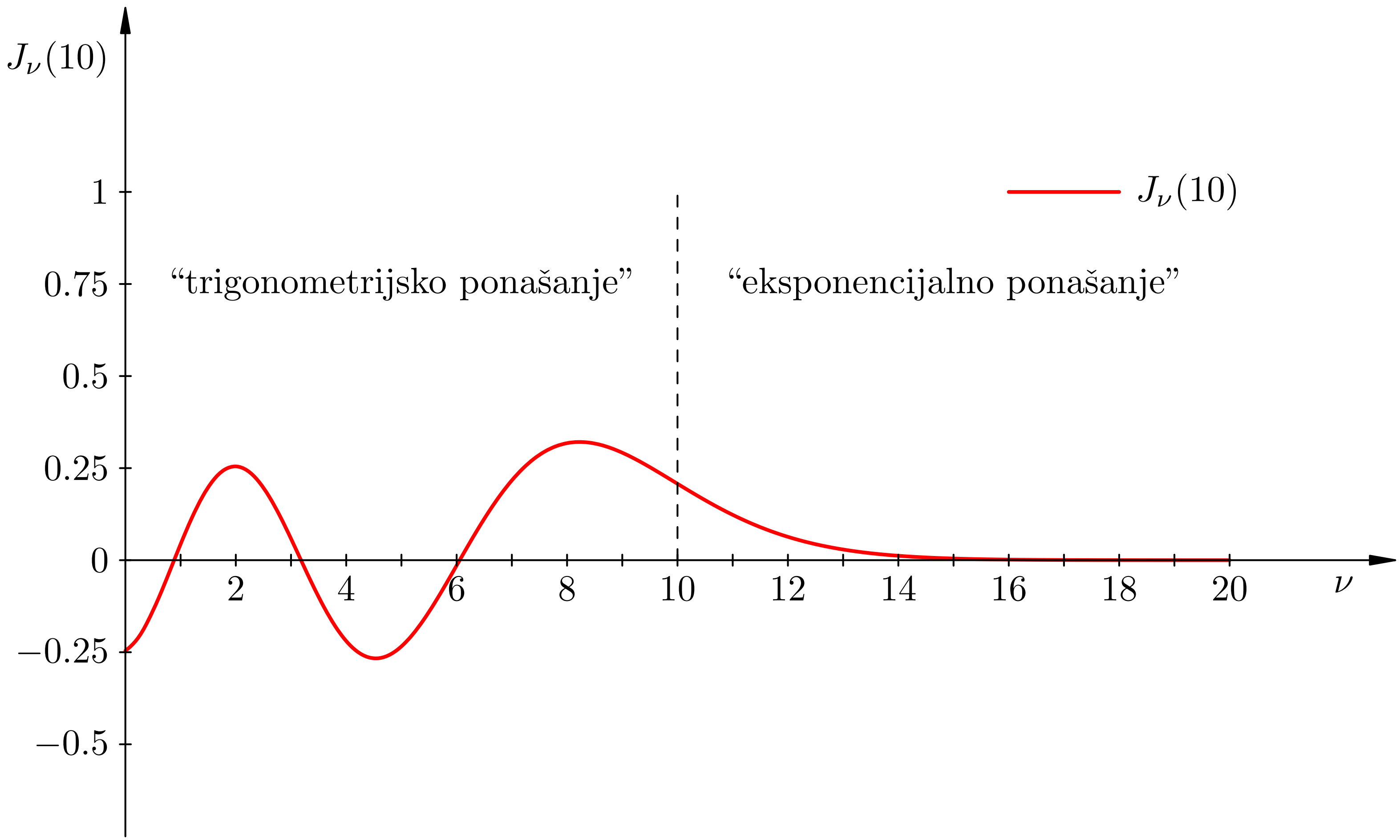
Besselove funkcije $J_\nu(k)$, za $k = 1, 2, 3$



Besselove funkcije $J_\nu(k)$, za $k = 4, 5, 6$



Besselova funkcija $J_\nu(10)$



Računanje Besselovih funkcija

Kad n raste, u rekurziji dobivamo sve manje i manje brojeve, što znači da mora doći do kraćenja.

To pokazuje da je rekurzija

$$J_{n+1}(x) - \frac{2n}{x} J_n(x) + J_{n-1}(x) = 0, \quad n \in \mathbb{N},$$

nestabilna u rastućem smjeru po n, čim uđemo u eksponencijalno područje $n > x$ (veza s cos i ch).

Ilustracija nestabilnosti za $x = 1$:

- računamo vrijednosti $J_n(x)$ korištenjem rekurziju uzlazno po n, u `extended` preciznosti.
- Dobiveni rezultati na 18 decimala (apsolutno) dani su u sljedećoj tablici.

Primjer za $J_n(1)$

n	izračunati $J_n(1)$	točni $J_n(1)$
0	0.7651976865579666552	0.7651976865579666552
1	0.4400505857449335116	0.4400505857449335116
2	0.1149034849319004481	0.1149034849319004481
3	0.0195633539826684406	0.0195633539826684406
4	0.0024766389641099555	0.0024766389641099555
5	0.0002497577302112347	0.0002497577302112344
6	0.0000209383380024188	0.0000209383380023889
7	0.0000015023258177779	0.0000015023258174437
8	0.0000000942234446486	0.0000000942234417726
9	0.0000000052493259991	0.0000000052492501880
10	0.0000000002644213522	0.0000000002630615122
11	0.0000000000391010558	0.0000000000119800667
12	0.0000000005958019177	0.0000000000004999722

Komentari

Možda je dobro uočiti:

- Kraćenje, a time i gubitak relativne točnosti počinje odmah — za $n = 2$, ulaskom u eksponencijalno područje.
- Međutim, to se ne vidi u ovoj tablici, jer su rezultati prikazani apsolutno, a ne relativno.
- Za $n = 11$ nemamo više niti jednu točnu znamenku.
- Za $n = 12$ gubimo i monotoni pad po n.

Vidimo da se događa nešto slično kao kod računanja e^{-nx}, što upućuje na okretanje rekurzije i primjenu Millerovog algoritma.

Besselove funkcije i Millerov algoritam

Millerov algoritam daje dobre rezultate (drugi stupac tablice, koji je točan), a funkcija izvodnica koja se pritom koristi za normalizaciju je

$$J_0(x) + 2(J_2(x) + J_4(x) + \cdots + J_{2k}(x) + \cdots) = 1.$$

Ova relacija izlazi direktno iz funkcije izvodnice za $t = 1$, kad iskoristimo parnost i neparnost Besselovih funkcija po n, tj.

$$J_{-n} = (-1)^n J_n.$$

Za praktičnu primjenu Millerovog algoritma poželjno je znati precizno ponašanje rekurzije.

- Za fiksni x zanima nas ponašanje $J_n(x)$ za velike n.

Besselove funkcije i Millerov algoritam (nast.)

- Može se pokazati da u eksponencijalnom području vrijedi tzv. asimptotska relacija

$$J_\nu(x) \approx \frac{1}{\sqrt{2\pi\nu}} \left(\frac{ex}{2\nu}\right)^\nu,$$

  za fiksni x i velike ν, tj. za $\nu \to \infty$.
- Gledano po n, za velike n, $J_n(x)$ se ponaša kao

$$\frac{c_n}{n^{n+0.5}},$$

  gdje je $c_n = (ex/2)^n/\sqrt{2\pi}$, a to vrlo brzo trne kad n raste.
- Korist: odavde se može izračunati početni indeks M za Millerov algoritam, tako da osiguramo potrebnu točnost.

Besselove funkcije i Millerov algoritam (nast.)

Na sličan način može se opisati i ponašanje Besselovih funkcija J_ν kada je ν fiksan, a gledamo male ili velike argumente x.

Za fiksni ν, kad $x \to 0$ iz prvog člana Taylorovog reda dobivamo i asimptotsku relaciju

$$J_\nu(x) \approx \left(\frac{1}{2}x\right)^\nu \frac{1}{\Gamma(\nu+1)},$$

koja je, očito, dobra aproksimacija za x u eksponencijalnom području — za x blizu nule.

Besselove funkcije i Millerov algoritam (nast.)

S druge strane, za $x \gtrsim n$, $J_n(x)$ se ponaša poput kosinusa, tj. oscilira. Prava relacija u trigonometrijskom području je

$$J_\nu(x) \approx \sqrt{\frac{2}{\pi x}} \cos\left(x - \frac{\pi}{4} - \frac{\nu\pi}{2}\right),$$

za fiksni ν, kad $x \to \infty$.

Izglađivanje podataka i diskretni najmanji kvadrati

Uvod u izglađivanje podataka

Zadan je skup podataka (ili točaka u ravnini)

$$(x_i, f_i), \quad i = 0, \ldots, n,$$

pri čemu

- vrijednosti x_i smatramo točnim — bez grešaka,
- a vrijednosti f_i su "izmjerene" vrijednosti i nose u sebi neku grešku, označimo ju s ε_i.

Na primjer, f_i možemo interpretirati kao

- približne vrijednosti neke funkcije f u točkama x_i.

Prave vrijednosti $f(x_i)$, naravno, ne znamo!

Točke x_i, bar zasad, ne moraju biti međusobno različite, isto kao i kod diskretnih najmanjih kvadrata.
Međutim, pretpostavimo da jesu međusobno različite.

Uvod u izglađivanje podataka (nastavak)

Dakle, imamo izmjereni uzorak funkcijskih vrijednosti funkcije f na diskretnom skupu točaka $x_0, \ldots, x_n$, za koji vrijedi

$$f_i = f(x_i) + \varepsilon_i, \quad i = 0, \ldots, n.$$

Ideja: uz određene statističke pretpostavke na "slučajnost" grešaka ε_i,

- zamijeniti f_i vrijednošću neke aproksimacijske funkcije φ u točki x_i,
- tako da očekivane greške u $\varphi_i := \varphi(x_i)$ budu "manje".

Drugim riječima, cilj je

- ukloniti slučajne greške u f_i, ili
- "izgladiti" ove izmjerene podatke!

Statističke pretpostavke na model

Za izglađivanje trebamo dvije statističke pretpostavke na greške.

Prva i ključna pretpostavka:

- U mjerenjima nema sistematskih grešaka, tj. imamo dobro “kalibrirane” instrumente, ili,
- srednja vrijednost ili očekivanje greške je nula

$$E(\varepsilon_i) = 0, \quad i = 0, \ldots, n.$$

Pisano vektorski u prostoru $\mathbb{R}^{n+1}$, uz oznaku $\varepsilon = (\varepsilon_0, \ldots, \varepsilon_n)^T$,

$$E(\varepsilon) = 0.$$

Uz tu pretpostavku, nepoznata vrijednost $f(x_i)$ je ista kao i očekivanje izmjerene vrijednosti f_i, za $i = 0, \ldots, n$.

Statističke pretpostavke na model (nastavak)

Druga pretpostavka, koja bitno olakšava računanje:

- Greške u različitim mjerenjima (od točke do točke) su statistički nezavisne,
- tj., kovarijacijska matrica $V := V(\varepsilon)$ je dijagonalna

$$V(\varepsilon) = \mathrm{diag}(\sigma_0^2, \ldots, \sigma_n^2),$$

a $\sigma_0, \ldots, \sigma_n$ su standardne devijacije u pojedinim mjerenjima (kvadrati su varijance).

Sasvim općenito, kovarijacijska matrica V može biti

- bilo koja (simetrična) pozitivno definitna matrica.

No, dijagonalnost bitno ubrzava računanje, a često je istinita u praksi.

Skalarni produkt generiran s $W = V^{-1}$

Na vektorskom prostoru $\mathbb{R}^{n+1}$ definiramo skalarni produkt generiran matricom $W := V^{-1}$ na sljedeći način

$$\langle u, v\rangle_W := v^T W u = \langle Wu, v\rangle, \quad u, v \in \mathbb{R}^{n+1}.$$

Iz pozitivne definitnosti matrice V^{-1} lako slijedi da je ovo zaista skalarni produkt.

Pripadnu W-normu vektora definiramo na standardni način

$$\|v\|_W := \sqrt{\langle v, v\rangle_W} = \sqrt{\langle Wv, v\rangle}, \quad v \in \mathbb{R}^{n+1}.$$

Obično ćemo koristiti kvadrat norme

$$\|v\|_W^2 = \langle v, v\rangle_W = \langle Wv, v\rangle, \quad v \in \mathbb{R}^{n+1}.$$

Linearna aproksimacija

Aproksimacijsku funkciju φ prikazujemo kao

- nepoznatu linearnu kombinaciju
- poznatih (izabranih) funkcija — tzv. funkcija "baze" u izabranom modelu.

Neka su $\varphi_0, \ldots, \varphi_m$ izabrane funkcije "baze" u modelu. Aproksimacijska funkcija φ ima oblik

$$\varphi(x) = c_0\varphi_0(x) + \cdots c_m\varphi_m(x) = \sum_{j=0}^{m} c_j\varphi_j(x),$$

Koeficijenti $c_0, \ldots, c_m$ u ovoj linearnoj kombinaciji su nepoznati parametri koje treba odrediti.

Dovoljno je uzeti $m \leq n$. U praksi, točaka $x_0, \ldots, x_n$ ima mnogo više nego nepoznatih parametara $c_0, \ldots, c_m$, tj. $n \gg m$.

Linearna aproksimacija (nastavak)

Iz vrijednosti funkcija baze φ_j u točkama x_i formiramo

- (općenito, pravokutnu) matricu $A \in \mathbb{R}^{(n+1)\times(m+1)}$

na sljedeći način

$$a_{ij} := \varphi_j(x_i), \quad i = 0, \ldots, n, \quad j = 0, \ldots, m.$$

Ako su funkcije baze “dobro” izabrane, onda su

- stupci matrice A linearno nezavisni,

tako da A ima puni stupčani rang (zato nam treba $m \leq n$).

Uz ove oznake, imamo

$$\varphi(x_i) = \sum_{j=0}^{m} c_j \varphi_j(x_i) = \sum_{j=0}^{m} a_{ij} c_j, \quad i = 0, \ldots, n.$$

Linearni model i procjena parametara

Nepoznate parametre $c_0, \ldots, c_m$ određujemo iz "modela"

$$f_i = \varphi(x_i) + \varepsilon_i, \quad i = 0, \ldots, n,$$

Označimo s

- $\tilde{f} := (f_0, \ldots, f_n)^T \in \mathbb{R}^{n+1}$ vektor izmjerenih vrijednosti,
- $c := (c_0, \ldots, c_m)^T \in \mathbb{R}^{m+1}$ vektor nepoznatih parametara.

Onda dobivamo tzv. linearni model za procjenu parametara c u obliku

$$\tilde{f} = Ac + \varepsilon.$$

Uz sve navedene pretpostavke, najbolju procjenu parametara dobivamo iz Gauss–Markov teorema:

- to je ona s najmanjom "varijancom"!

Veza izglađivanja i najmanjih kvadrata

Kod linearnog modela $\tilde{f} = Ac + \varepsilon$ za procjenu parametara c, pretpostavke Gauss–Markovog teorema su sljedeće:

- $E(\varepsilon) = 0$ $(E(\varepsilon_i) = 0,\ i = 0, \ldots n)$,
- $V(\varepsilon) = E(\varepsilon\varepsilon^T) = \sigma^2 I_{n+1}$ $(Var(\varepsilon_i) = \sigma^2,\ i = 0, \ldots n)$,
- $cov(\varepsilon_i, \varepsilon_j) = 0,\ i \neq j, i, j = 0, \ldots, n.$

Tada taj teorem tvrdi da je rješenje problema najmanjih kvadrata $\hat{c} = (A^T A)^{-1} A^T \tilde{f}$ najbolji linearni nepristrani procjenitelj (BLUE) za c.

- Budući da su podaci u vektoru $\tilde{f}$ poznati, tražimo procjenitelje koji su linearna funkcija od $\tilde{f}$, tj.

$$\tilde{c} = M\tilde{f} + g, \qquad \text{gdje su } M \in \mathbb{R}^{(m+1)\times(n+1)},\ g \in \mathbb{R}^{m+1}.$$

Veza izglađivanja i najmanjih kvadrata (nast.)

- Sljedeće, zahtijevamo da procijenitelj bude nepristran, tj. da je
$$E(\tilde{c}) = c,$$
za bilo koju valjani c. Da bi $\tilde{c}$ bio nepristran tada mora vrijediti
$$\begin{aligned} E(\tilde{c}) &= ME(\tilde{f}) + g = ME(Ac + \varepsilon) + g \\ &= MAc + ME(\varepsilon) + g = MAc + g = \text{(po pretpostavci)} \\ &= c \end{aligned}$$

- Dakle, mora vrijediti
$$g = 0, \qquad MA = I_{m+1}.$$

Veza izglađivanja i najmanjih kvadrata (nast.)

Primijetimo da rješenje problema najmanjih kvadrata je nepristrano, jer je za $\hat{c}$

$$M = (A^T A)^{-1} A^T, \quad i \quad MA = (A^T A)^{-1} A^T A = I_{m+1}.$$

Dakle potraga za optimalnim nepristranim linearnim procjeniteljem se svodi na potragu za procjeniteljima oblika

$$\tilde{c} = M\tilde{f}, \qquad MA = I_{m+1}.$$

- Bez smanjenja općenitosti možemo redefinirati matricu M kao

  $$M = (A^T A)^{-1} A^T + C,$$

  gdje je $C \in \mathbb{R}^{(m+1)\times(n+1)}$.

Veza izglađivanja i najmanjih kvadrata (nast.)

Ponovo, da bi naš procijenitelj bio nepristran treba vrijediti

$$\begin{aligned} I_{m+1} &= MA \\ &= [(A^T A)^{-1} A^T + C] A = I_{m+1} + CA, \;\Rightarrow \\ CA &= 0_{m+1}. \end{aligned}$$

Sada možemo izračunati kovarijacijsku matricu za sve alternativne procijenitelje $\tilde{c}$:

$$\tilde{c} = M\tilde{f} = M(Ac + \varepsilon) = c + M\varepsilon.$$

Tada je

$$\tilde{c} - c = M\varepsilon.$$

Budući da je nepristran po konstrukciji vrijedi $E(\tilde{c} - c) = 0$.

Veza izglađivanja i najmanjih kvadrata (nast.)

Odavde slijedi da je kovarijacijska matrica procjenitelja oblika

$$\begin{aligned} V(\tilde{c}) &= E((\tilde{c} - c)(\tilde{c} - c)^T) \\ &= E(M\varepsilon(M\varepsilon)^T) = E(M\varepsilon\varepsilon^T M^T) \\ &= ME(\varepsilon\varepsilon^T)M^T = M\sigma^2 I_{n+1} M^T = \sigma^2 M M^T. \end{aligned}$$

Dalje, vrijedi

$$\begin{aligned} MM^T &= [(A^T A)^{-1} A^T + C][(A^T A)^{-1} A^T + C]^T \\ &= (A^T A)^{-1} A^T A (A^T A)^{-1} + (A^T A)^{-1} A^T C^T + \\ &\quad + CA(A^T A)^{-1} + CC^T. \end{aligned}$$

Budući da je $CA = 0$, dobivamo

$$MM^T = (A^T A)^{-1} + CC^T.$$

Veza izglađivanja i najmanjih kvadrata (nast.)

Matrica CC^T je pozitivno semidefinitna.

- Najbolji procjenitelj u klasi procjenitelja koje promatramo je onaj sa “najmanjom” kovarijacijskom matricom.
- To znači da kada od kovarijacijske matrice bilo kojeg drugog procjenitelja iz klase (linearan i nepristran) oduzmemo kovarijacijsku matricu najboljeg procjenitelja dobijemo pozitivno semidefinitnu matricu:

$$\sigma^2 MM^T - V(\hat{c}) \geq 0, \quad \text{tj. pozitivno semidefinitna.}$$

- Budući da će se MM^T minimizirati kada je $C = 0$, najbolji procjenitelj u klasi je $\hat{c}$.
- Bilo koji drugi procjenitelj iz klase ima “veću” kovarijacijsku matricu.

Veza izglađivanja i najmanjih kvadrata (nast.)

Dakle, možemo zaključiti da je rješenje problema najmanjih kvadrata $\hat{c}$ najbolji linearni nepristrani procjenitelj (BLUE) koji zadovoljava tri uvjeta Gauss–Markovog teorem.

Posljedice Gauss–Markovog teorem su sljedeće:

- $$\begin{aligned} V(\hat{c}) &= \sigma^2 (A^T A)^{-1} A^T [(A^T A)^{-1} A^T]^T \\ &= \sigma^2 (A^T A)^{-1} A^T A (A^T A)^{-1} = \sigma^2 (A^T A)^{-1}. \end{aligned}$$
- $$\begin{aligned} \|\tilde{f} - A\hat{c}\|_2^2 &= (\tilde{f} - A\hat{c})^T (\tilde{f} - A\hat{c}) \\ &= (\tilde{f} - A\hat{c})^T \tilde{f} \quad (\text{rezidual okomit na } \mathcal{R}(A)) \\ &= \tilde{f}^T \tilde{f} - \hat{c}^T A^T \tilde{f} = \tilde{f}^T \tilde{f} - \hat{c}^T A^T A \hat{c} \\ &\quad (\text{iz normalnog sustava jednadžbi}). \end{aligned}$$

Veza izglađivanja i najmanjih kvadrata (nast.)

- Za sljedeći rezultat nam treba ovaj pomoćni rezultat iz teorije vjerojatnosti: neka je M simetrična matrica tada vrijedi

$$E(y^T M y) = E(y)^T M E(y) + \mathrm{tr}(M V(y)).$$

- Tada imamo

$$E((\tilde{f} - A\hat{c})^T(\tilde{f} - A\hat{c})) = E(\tilde{f}^T \tilde{f}) - E(\hat{c}^T A^T A\hat{c}).$$

- Izračunat ćemo svako od ovih očekivanja posebno.

$$\begin{aligned} E(\tilde{f}^T \tilde{f}) &= E(\tilde{f})^T E(\tilde{f}) + \mathrm{tr}(I \cdot \sigma^2 I_{n+1}) \\ &= (Ac)^T (Ac) + \sigma^2 \mathrm{tr}(I_{n+1}) = c^T A^T A c + (n+1)\sigma^2. \end{aligned}$$

jer je $V(\tilde{f}) = V(\varepsilon) = \sigma^2 I_{n+1}$.

Veza izglađivanja i najmanjih kvadrata (nast.)

$$\begin{aligned} E(\hat{c}^T A^T A\hat{c}) &= E(\hat{c})^T A^T A E(\hat{c}) + \mathrm{tr}(A^T A V(\hat{c})) \\ &= c^T A^T A c + \mathrm{tr}(A^T A \cdot \sigma^2 (A^T A)^{-1}) \\ &= c^T A^T A c + \sigma^2 \mathrm{tr}(I_{m+1}) \\ &= c^T A^T A c + (m+1)\sigma^2 \end{aligned}$$

- Napokon dobivamo

$$\begin{aligned} E((\tilde{f} - A\hat{c})^T (\tilde{f} - A\hat{c})) &= c^T A^T A c + (n+1)\sigma^2 - \\ &\quad - c^T A^T A c - (m+1)\sigma^2 \\ &= (n-m)\sigma^2. \end{aligned}$$

Veza izglađivanja i najmanjih kvadrata (nast.)

Sada, da bi naš model sa općenitom $V = V(\varepsilon)$ prilagodili uvjetima Gauss–Markovog teorema, najjednostavniji postupak je cijeli model pomnožiti sa $V^{-\frac{1}{2}}$, i promatrati

$$V^{-\frac{1}{2}}\tilde{f} = V^{-\frac{1}{2}}Ac + V^{-\frac{1}{2}}\varepsilon,$$

jer je u tom slučaju

$$V(V^{-\frac{1}{2}}\varepsilon) = V^{-\frac{1}{2}}VV^{-\frac{1}{2}} = I$$

Ovaj postupak je ekvivalentan računanju rješenja problema najmanjih kvadrata

$$\|\tilde{f} - Ac\|_W^2 = \langle W(\tilde{f} - Ac), \tilde{f} - Ac\rangle \to \min,$$

uz $W = V^{-1}$.

Veza izglađivanja i najmanjih kvadrata (nast.)

Pripadni sustav normalnih jednadžbi je tada

$$(A^T W A)c = A^T W \tilde{f}.$$

Dodatno, kovarijacijska matrica najbolje procjene $\hat{c}$ je

$$V(\hat{c}) = (A^T W A)^{-1}.$$

i još vrijedi

$$E(\|\tilde{f} - Ac\|_W^2) = n - m.$$

Standardni zapis skalarnog produkta

U praksi je kovarijacijska matrica V često dijagonalna

$$V(\varepsilon) = \mathrm{diag}(\sigma_0^2, \ldots, \sigma_n^2),$$

a $\sigma_0, \ldots, \sigma_n$ su standardne devijacije u pojedinim mjerenjima (kvadrati su varijance). Tada se koristi standardna oznaka

$$w_i = \frac{1}{\sigma_i^2}, \quad i = 0, \ldots, n,$$

a pripadni težinski skalarni produkt ima oblik

$$\langle u, v \rangle_W = \sum_{i=0}^{n} w_i u_i v_i, \quad u, v \in \mathbb{R}^{n+1}.$$

Složenost pada s $O(n^2)$ na $O(n)$.

Težinski najmanji kvadrati

Pripadni problem najmanjih kvadrata

$$\|\tilde{f} - Ac\|_W^2 = \langle W(\tilde{f} - Ac), \tilde{f} - Ac\rangle \to \min,$$

onda ima "puni" oblik

$$\|\tilde{f} - Ac\|_W^2 = \sum_{i=0}^{n} w_i \bigg(f_i - \sum_{j=0}^{m} c_j \varphi_j(x_i) \bigg)^2 \to \min\,.$$

Praktični problemi kod izglađivanja

- Praktična uloga varijanci σ_i, odnosno težina w_i.
- Skaliranje slično klasičnoj pretpostavci — greške su nezavisne, normalna distribucija $N(0,1)$.
- Izbacivanje istih točaka (zbog dimenzije).

U primjeni kod izglađivanja — m se traži!

- Ideja — diži m (dimenziju prostora), sve dok se $\|\tilde{f} - Ac\|_W^2/(n-m)$ približno ne stabilizira.
- Matrica A varira (raste) s m.
- To ima smisla koristiti ako uzimamo ortogonalne baze, tako da A ima ortogonalne stupce ($A^T W A = I_m$).
- Realizacija za polinome = Forsytheov algoritam s detekcijom stupnja.

Forsytheov algoritam

Diskretni ortogonalni polinomi

Za ekvidistantne (ili uniformne) mreže (fiksni korak h)

- ne treba generirati diskretne ortogonalne polinome rekurzivnim formulama,
- jer postoje eksplicitne formule — tzv. Gramovi polinomi i njihove afine transformacije.

Za neekvidistantne mreže

- treba generirati pripadne diskretne ortogonalne polinome.

To se isplati i za ekvidistantne mreže (isti algoritam).

Dodatno, u oba slučaja, tražimo još i

- koeficijente u razvoju zadane funkcije f po tim ortogonalnim polinomima,

a sve se dobiva jednim algoritmom — u paketu.

Diskretni ortogonalni polinomi — rekurzija

Svi ortogonalni polinomi (pa i diskretni) zadovoljavaju tročlanu, homogenu rekurziju oblika:

$$p_{k+1}(x) = (a_k x + b_k) p_k(x) - c_k p_{k-1}(x).$$

Ako su polinomi monični (vodeći koeficijent je jednak 1), onda prethodnu rekurziju možemo zapisati kao

$$p_{k+1}(x) = (x - \alpha_k) p_k(x) - \beta_k p_{k-1}(x).$$

U nastavku korstimo samo ovaj oblik rekurzije.

Oznaka c_j nadalje označava koeficijente u razvoju.

Forsytheov algoritam

Forsytheov algoritam

- iz uvjeta ortogonalnosti na zadanoj mreži nalazi koeficijente α_k i β_k,
- računa sljedeći koeficijent c_j u razvoju funkcije,
- zajedno s generaliziranom Hornerovom shemom daje efikasan algoritam za nalaženje polinoma po metodi najmanjih kvadrata.

Algoritam je pronašao George E. Forsythe, 1957. godine.

Algoritam se sastoji od dvije faze:

- generiranja diskretnih ortogonalnih polinoma,
- računanja koeficijenata aproksimacije.

Usput još, računa i pripadnu srednjekvadratnu grešku.

Problem težinske aproksimacije polinomima

Zadan je skup podataka:

$$(x_i, f_i), \quad i = 1, \ldots, n,$$

s težinama w_i. Idealno je da su težine inverzno proporcionalne varijancama mjerenja (v. Gauss–Markovljev model), no može i drugačije.

Taj niz podataka želimo

- aproksimirati polinomom p stupnja m, uz $m < n$,
- kojeg želimo prikazati kao linearnu kombinaciju diskretnih ortogonalnih polinoma na mreži

$$\{x_1,\ x_2,\ \ldots,\ x_n\},$$

uz zadane težine w_i u pripadnom skalarnom produktu.

Oblik aproksimacijske funkcije

Dakle, polinom p tražimo u obliku

$$p(x) = \sum_{j=0}^{m} c_j p_j(x),$$

pri čemu su $p_0, \ldots, p_m$ ortogonalni polinomi na zadanoj mreži,

$$\langle p_k, p_\ell \rangle = \sum_{i=1}^{n} w_i \, p_k(x_i) \, p_\ell(x_i) = 0, \quad \text{za } k \neq \ell.$$

Takav niz polinoma generiran je već spomenutom rekurzijom

$$p_k(x) = (x - \alpha_{k-1}) p_{k-1}(x) - \beta_{k-1} p_{k-2}(x), \quad k = 1, 2, \ldots, m.$$

uz start rekurzije $p_{-1}(x) = 0$, $p_0(x) = 1$ (pomak indeksa za 1).

Računanje koeficijenata ortogonalnih polinoma

Množenjem prethodne jednadžbe s $w_i p_{k-1}$, uvrštavanjem x_i i sumiranjem, dobivamo

$$\alpha_{k-1} = \frac{\displaystyle\sum_{i=1}^{n} w_i x_i p_{k-1}^2(x_i)}{\displaystyle\sum_{i=1}^{n} w_i p_{k-1}^2(x_i)}, \quad k = 1, 2, \ldots, m.$$

ili

$$\alpha_{k-1} = \frac{\langle \mathrm{id}\, p_{k-1}, p_{k-1} \rangle}{\langle p_{k-1}, p_{k-1} \rangle}, \quad k = 1, 2, \ldots, m,$$

gdje je id identiteta, tj. $\mathrm{id}(x) = x$.

Računanje koeficijenata ortogonalnih polinoma

Na sličan način, samo množenjem $w_i p_{k-2}$, uvrštavanjem x_i i sumiranjem, dobivamo

$$\beta_{k-1} = \frac{\displaystyle\sum_{i=1}^{n} w_i x_i p_{k-1} p_{k-2}(x_i)}{\displaystyle\sum_{i=1}^{n} w_i p_{k-2}^2(x_i)}, \quad k = 1, 2, \ldots, m.$$

ili

$$\beta_{k-1} = \frac{\langle \mathrm{id}\, p_{k-1}, p_{k-2} \rangle}{\langle p_{k-2}, p_{k-2} \rangle}, \quad k = 1, 2, \ldots, m.$$

Koeficijenti aproksimacijske funkcije

Budući da je razvoj polinoma p imao oblik,

$$p(x) = \sum_{j=0}^{m} c_j p_j(x),$$

želimo da norma reziduala tog sustava jednadžbi

$$f_i = \sum_{j=0}^{m} c_j p_j(x_i), \quad i = 1, \ldots, n$$

bude minimalna za parove točaka (x_i, f_i) (vidjeti izvod matrične formulacije najmanjih kvadrata).

Koeficijenti aproksimacijske funkcije (nastavak)

Množenjem prethodne jednadžbe s $p_k(x_i)w_i$ i korištenjem ortogonalnosti polinoma na zadanoj mreži, dobivamo

$$c_j = \frac{\displaystyle\sum_{i=1}^{n} w_i f_i p_j(x_i)}{\displaystyle\sum_{i=1}^{n} w_i p_j^2(x_i)}, \qquad j = 0, \ldots, m,$$

ili

$$c_j = \frac{\langle f, p_j \rangle}{\langle p_j, p_j \rangle}.$$

Kad su izračunati $c_0, \ldots, c_m$, računanje aproksimacije $p(x)$ u točkama x provodi se generaliziranom Hornerovom shemom, pa potprogram treba vratiti i koeficijente rekurzije α_k i β_k.

Detekcija stupnja polinoma

Primjer: Detekcija stupnja polinoma.

Zadan je polinom

$$f(x) = x^2 + 4x + 7,$$

i tabeliramo ga u točkama

$$x_i = \frac{i-1}{2}, \quad i = 1, \ldots, 5.$$

Tabelirane vrijednosti su:

x_i	0	0.5	1	1.5	2
f_i	9	9.25	12	15.25	19

Detekcija stupnja polinoma (nastavak)

Perturbiramo malo podatke u točkama 0.5 i 1.5:

f_i	0	0.5	1	1.5	2
x_i	9	9.3	12	15.2	19

Za sve težine uzmimo

$$w_i = 1,$$

jer su svi polazni podaci jednako pouzdani.

Treba naći aproksimaciju polinomima po metodi najmanjih kvadrata i odrediti (vjerojatni) stupanj polaznog polinoma f.

Detekcija stupnja polinoma (nastavak)

Redom imamo:

$$p_0(x) = 1$$

i definiramo $\beta_0 = 0$.

Računamo p_1:

$$\alpha_0 = \frac{\displaystyle\sum_{i=1}^{5} x_i \cdot 1}{\displaystyle\sum_{i=1}^{5} 1} = 1 \qquad \Longrightarrow \qquad p_1(x) = x - 1.$$

Detekcija stupnja polinoma (nastavak)

Računamo p_2:

$$\beta_1 = \frac{\displaystyle\sum_{i=1}^{5} x_i\, p_1(x_i)\cdot 1}{\displaystyle\sum_{i=1}^{5} 1^2} = 0.5$$

$$\alpha_1 = \frac{\displaystyle\sum_{i=1}^{5} x_i\, p_1^2(x_i)\cdot 1}{\displaystyle\sum_{i=1}^{5} p_1^2(x_i)} = 1$$

$$\Longrightarrow \quad p_2(x) = (x-1)\cdot(x-1) - 0.5\cdot$$

$$p_2(x) = x^2 - 2x + 0.5$$

Detekcija stupnja polinoma (nastavak)

Pogledajmo vrijednosti ovih polinoma u čvorovima:

x_i	$p_0(x_i)$	$p_1(x_i)$	$p_2(x_i)$
0	1	-1	0.5
0.5	1	-0.5	-0.25
1	1	0	-0.5
1.5	1	0.5	-0.25
2	1	1	0.5

.

Detekcija stupnja polinoma (nastavak)

Nađimo još koeficijente c_j u razvoju:

$$c_0 = \frac{\sum_{i=1}^{5} f_i}{\sum_{i=1}^{5} 1} = 12.5$$

$$c_1 = \frac{\sum_{i=1}^{5} f_i \, p_1(x_i)}{\sum_{i=1}^{5} p_1^2(x_i)} = 5.98$$

$$c_2 = \frac{\sum_{i=1}^{5} f_i \, p_2(x_i)}{\sum_{i=1}^{5} p_2^2(x_i)} = 1.$$

Detekcija stupnja polinoma (nastavak)

Konačno, nađimo još aproksimacijske polinome:

polinom stupnja 0 je: $P_0 = a_0 p_0 = 12.5$

polinom stupnja 1 je: $P_1 = P_0 + a_1 p_1 = 5.98x + 6.52$

polinom stupnja 2 je: $P_2 = P_1 + a_2 p_2 = x^2 + 3.98x + 7.02,$

što je i po koeficijentima jako blizu f.

Pripadne sume kvadrata apsolutnih grešaka su:

$$S_0 = 91.155, \qquad S_1 = 0.879, \qquad S_2 = 0.004.$$

Zaključujemo da je riječ o blago perturbiranim podacima u kvadratnom polinomu, jer je greška u S naglo opala. Za S_3 dobili bismo manje, ali ne bitno, pa ovdje stajemo.

Detekcija stupnja polinoma (nastavak)

Na kraju, napišimo još i tablicu vrijednosti u čvorovima za sva tri polinoma P_j:

x_i	f_i	$P_0(x_i)$	$P_1(x_i)$	$P_2(x_i)$
0	7	12.5	6.52	7.02
0.5	9.3	12.5	9.51	9.26
1	12	12.5	12.50	12.00
1.5	15.2	12.5	15.49	15.24
2	19	12.5	18.48	18.98

.

Lokalno izglađivanje funkcija

Globalno i lokalno izglađivanje

Diskretna metoda najmanjih kvadrata koristi se još i za

- globalno izglađivanje funkcija,
- lokalno izglađivanje funkcija.

O globalnom izglađivanju je već bilo govora.

Lokalno izglađivanje najčešće se provodi povlačenjem

- pravca po metodi najmanjih kvadrata kroz 3, 5 ili 7 točaka,
- parabole po metodi najmanjih kvadrata kroz 5 ili 7 točaka.

Katkad se koriste i

- "lokalne" integracijske formule (v. Multigrid).

Parabola kroz 5 točaka

Primjer. Nađite parabolu po metodi najmanjih kvadrata za 5 ekvidistantnih točaka (x_i, f_i), gdje je $i = k-2, \ldots, k+2$.

Da bismo lakše računali, napravimo transformaciju koja te točke prevodi u $t_{k-2} = -2, \ldots, t_{k+2} = 2$. Označimo novu varijablu s

$$t = \frac{x - x_k}{h},$$

pri čemu je $x_{i+1} - x_i = h$.

Parabola u novoj varijabli t je

$$p(t) = c_2 t^2 + c_1 t + c_0.$$

Parabola kroz 5 točaka (nastavak)

Zbog simetrije točaka oko 0, linearni sustav koji rješava problem najmanjih kvadrata za parabolu p je iznimno jednostavan:

$$\begin{aligned}
5c_0 \qquad\qquad + 10c_2 &= \sum_{i=k-2}^{k+2} f_i \\
10c_1 \qquad\qquad &= \sum_{i=k-2}^{k+2} t_i f_i \\
10c_0 \qquad\qquad + 34c_2 &= \sum_{i=k-2}^{k+2} t_i^2 f_i
\end{aligned}$$

Parabola kroz 5 točaka (nastavak)

Rješenje tog sustava je:

$$c_0 = \frac{1}{35}(-3f_{k-2} + 12f_{k-1} + 17f_k + 12f_{k+1} - 3f_{k+2})$$

$$c_1 = \frac{1}{10}(-2f_{k-2} - f_{k-1} + f_{k+1} + 2f_{k+2})$$

$$c_2 = \frac{1}{14}(2f_{k-2} - f_{k-1} - 2f_k - f_{k+1} + 2f_{k+2}).$$

Parabola kroz 5 točaka (nastavak)

Aproksimacije $\hat{f}_i$ = "izglađene" vrijednosti funkcije dobivamo izvrednjavanjem parabole p u točkama $-2, \ldots, 2$. Matrični zapis je:

$$\begin{bmatrix} \widehat{f}_{k-2} \\ \widehat{f}_{k-1} \\ \widehat{f}_{k} \\ \widehat{f}_{k+1} \\ \widehat{f}_{k+2} \end{bmatrix} = \frac{1}{35} \begin{bmatrix} 3 & -5 & -3 & 9 & 31 \\ -5 & 6 & 12 & 13 & 9 \\ -3 & 12 & 17 & 12 & -3 \\ 9 & 13 & 12 & 6 & -5 \\ 31 & 9 & -3 & -5 & 3 \end{bmatrix} \begin{bmatrix} f_{k-2} \\ f_{k-1} \\ f_{k} \\ f_{k+1} \\ f_{k+2} \end{bmatrix}.$$

Stvar se može napraviti i na neekvidistantnim mrežama, samo su formule kompliciranije.

Usrednjavanje integracijom

Napomena. Metoda najmanjih kvadrata nije jedini lokalni način izglađivanja. Podatke možemo izglađivati i lokalnim integralnim usrednjavanjem.

Primjer. Korištenjem, redom, trapezne i Simpsonove formule, dobivamo

$$\widehat{f_k} = \frac{1}{4}(f_{k-1} + 2f_k + f_{k+1})$$

$$\widehat{f_k} = \frac{1}{6}(f_{k-1} + 4f_k + f_{k+1}).$$

Slično se može napraviti i na neekvidistantnim mrežama.

- Sve ove formule svode se na “težinsko” usrednjavanje.