Univerzitet u Novom Sadu
Prirodno-matematički fakultet
Departman za matematiku i informatiku

Bojan Bašić
Bojana Bojić
Jelena Grubačić

Metod najmanjih kvadrata

Seminarski rad iz Numeričke analize

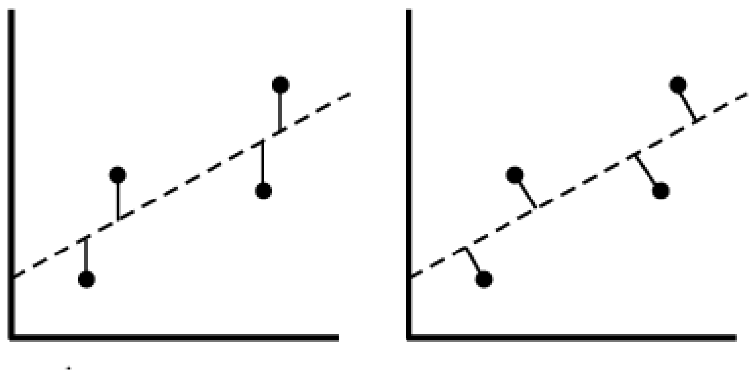
Novi Sad, 2007

Sadržaj

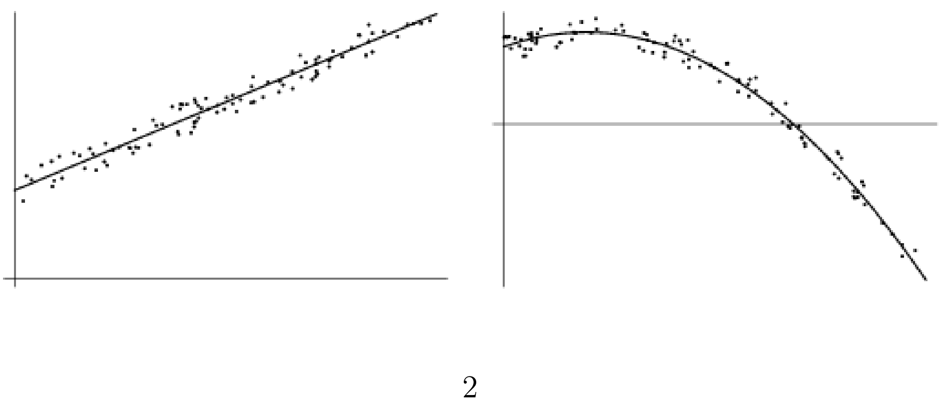
1 Uvod

Neka je postavljen sledeći problem: zagrevati sud s vodom i izmeriti temperaturu vode nakon jednog, dva, tri, odnosno četiri minuta. Na osnovu dobijenih rezultata nacrtati grafik zavisnosti temperature vode od proteklog vremena, znajući da je u pitanju prava linija.

Problem je naizgled jednostavan, ali ubrzo bismo ustanovili da izmerene vrednosti ne sačinjavaju pravu liniju. Drugim rečima, izmereni rezultati nisu egzaktni, već sadrže neku grešku; ove greške mogu nastati kao posledice raznih stvari: nesavršenosti mernih instrumenata, nesavršenosti uslovâ eksperimenta, ljudskog faktora itd. Želimo, dakle, da dobijeni grafik što manje (u izvesnom smislu) odstupa od izmerenih rezultata, tj. da predstavlja dobru **aproksimaciju**.



U praksi se najčešće uzimaju u obzir vertikalna odstupanja (na slici levo), za razliku od normalnih (slika desno). Ovo ima sledeće prednosti: direktno predstavlja zavisnost y od datog x (ono što se najčešće i želi), formule s kojima će se raditi postaju znatno jednostavnije, a metod ćemo moći da koristimo i za traženje aproksimacija polinomima višeg stepena i ostalim funkcijama (naime, u datom primeru reč je o tzv. **linearnoj regresiji**, budući da se traži grafik koji je linearna funkcija, ali tako nešto je u praksi prilično retko; videti sliku dole za primer linearne regresije i primer aproksimacije drugačijom funkcijom). U svakom slučaju, ukoliko su greške svedene na "razumnu meru", razlika između vertikalnih i normalnih odstupanja postaje zanemarljiva.



Postavlja se pitanje kako definisati najbolju aproksimaciju. Mogli bismo zahtevati da suma apsolutnih vrednosti odstupanja (grešaka) bude što manja; ovo se ponekad naziva l_1 aproksimacija. Mogli bismo, s druge strane, zahtevati da najveće odstupanje bude što je manje moguće (naravno, po apsolutnoj vrednosti); ovo se ponekad naziva l_∞ aproksimacija. Ipak, u praksi se najčešće minimizuje zbir kvadrata odstupanja; ovo se naziva l_2 aproksimacija ili aproksimacija najmanjih kvadrata, što je i tema ovog rada.

Problemi slični onom sa početka teksta pojavljuju se u najrazličitijim naukama, pod različitim okolnostima. Jedan sveobuhvatan tekst, ako bi ga uopšte i bilo moguće sastaviti, zauzimao bi neuporedivo više prostora od ovog rada. Stoga je cilj autorâ samo da čitaoca uvedu u ovu temu, a za više informacija o metodu najmanjih kvadrata i aproksimaciji uopšte preporučuju literaturu datu na kraju.

2 Istorijski počeci

Osnove metoda najmanjih kvadrata razvio je Gaus[1] 1795. godine, kada je bio osamnaestogodišnjak. Ipak, nije ga objavio do 1809, kada se ovaj metod pojavio u drugom tomu njegovih radova o nebeskoj mehanici, *Theoria Motus Corporum Coelestium in sectionibus conicis solem ambientium.* U međuvremenu je do sličnog metoda nezavisno došao i Adrijen-Mari Ležandr[2], koji je 1805. uz svoju knjigu *Nouvelles methods pour la determination des orbites des cometes* objavio dodatak na devet strana pod nazivom *Sur la metode des moindres quarres* (što je izazvalo i žučne polemike između Gausa i Ležandra, jer je svaki smatrao da on polaže pravo na prvobitno otkriće ovog metoda). Za nepunu deceniju, metod se vrlo brzo raširio po Evropi i postao nezaobilazan, naročito u astronomiji i geodeziji. Govorilo se da je metod najmanjih kvadrata za statistiku ono što je diferencijalni račun za analizu.

Jedna od najranijih potvrda jačine ovog metoda došla je upravo u astronomiji. Đuzepe Pjaci[3], otkrio je 1. januara 1801. patuljastu planetu Cereru i uspeo je da prati njeno kretanje četrdeset dana, dok se nije izgubila u Sunčevom sjaju. Na osnovu podataka koje je on dotad prikupio, trebalo je predvideti kada će Cerera opet doći u vidno polje. Tada dvadesetčetvorogodišnji Gaus izvršio je proračune na osnovu metoda koji je razvio, i upravo na osnovu njegovih proračuna Fon Cah[4] je ponovo locirao Cereru.

Dokaz optimalnosti (u izvesnom smislu) ovog metoda dao je Gaus 1829, a danas je poznat pod nazivom Gaus-Markovljeva[5] teorema.

[1] Carl Friedrich Gauss (1777-1855), nemački matematičar.
[2] Adrien-Marie Legendre (1752-1833), francuski matematičar.
[3] Giuseppe Piazzi (1746-1826), italijanski astronom i matematičar.
[4] Franz Xaver von Zach (1754-1832), austrijski astronom.
[5] Андрей Андреевич Марков (1856-1922), ruski matematičar.

3 Metod najmanjih kvadrata - uopšteno

Formalizujmo razmatranje iz uvoda. Neka su izmerene veličine (x_i, y_i), $i = 1, 2, \ldots, n$, i neka je poznat **oblik** funkcije Φ koja treba da što približnije aproksimira ove podatke. Pod poznavanjem oblika funkcije podrazumevamo poznavanje skupa **baznih funkcija** $\{\varphi_i\}_{i=1}^{m}$ na osnovu kojih konstruišemo funkciju Φ. Zavisnost funkcije Φ od funkcija φ_i data je preko konstanti $a_1, a_2, \ldots, a_m$, a cilj je da njih adekvatno odredimo. Uzima se da je $n > m$.

Aproksimacije l_1, l_∞ i l_2 pomenute u uvodu sada se mogu interpretirati kao nalaženje minimuma funkcija

$$\sum_{i=1}^{n} |\Phi(x_i; a_1, a_2, \ldots, a_m) - y_i|,$$

$$\max_{1 \leqslant i \leqslant n} |\Phi(x_i; a_1, a_2, \ldots, a_m) - y_i|,$$

odnosno

$$\sum_{i=1}^{n} \left(\Phi(x_i; a_1, a_2, \ldots, a_m) - y_i\right)^2$$

redom. Odmah se primećuje da prve dve funkcije nisu diferencijabilne, što je otežavajuća okolnost za nalaženje minimuma. Zato je upravo metod najmanjih kvadrata najkorišćeniji, kako je i rečeno u uvodu.

Ukoliko je funkcija

$$\sum_{i=1}^{n} \left(\Phi(x_i; a_1, a_2, \ldots, a_m) - y_i\right)^2, \tag{1}$$

(označimo je kraće sa $H(a_1, a_2, \ldots, a_m)$) zaista diferencijabilna, i ako su svi parcijalni izvodi neprekidni, moraju biti ispunjeni sledeći uslovi:

$$\left.\begin{array}{l} \dfrac{\partial H}{\partial a_1} = 0 \\[2mm] \dfrac{\partial H}{\partial a_2} = 0 \\ \vdots \\ \dfrac{\partial H}{\partial a_m} = 0 \end{array}\right\},$$

pa vrednosti $a_1, a_2, \ldots, a_m$ možemo dobiti rešavanjem ovog sistema jednačina. Ove jednačine nazivaju se **normalne jednačine**.

Posebno, neka je za dati skup baznih funkcija $\{\varphi_i\}_{i=1}^{n}$ funkcija Φ definisana na sledeći način:

$$\Phi(x; a_1, \ldots, a_m) = a_1\varphi_1(x) + a_2\varphi_2(x) + \cdots + a_m\varphi_m(x).$$

Funkcija Φ je, dakle, linearna kombinacija funkcija φ_i, pa se zbog toga ovaj specijalan slučaj naziva **linearni najmanji kvadrati**. Najpre primetimo da sve moguće ovako definisane funkcije Φ čine vektorski prostor nad poljem $\mathbb{R}$, dok funkcije φ_i, $i = 1, 2, \ldots, m$, pripadaju bazi tog vektorskog prostora. Zato je prirodno podrazumevati da su funkcije φ_i linearno nezavisne, a posmatraćemo i specijalan slučaj kada su φ_i i φ_j međusobno ortogonalne za sve $i \neq j$.

Pre nego što pređemo na detaljnije ispitivanje osobina linearnih najmanjih kvadrata, definisaćemo ortogonalnost i druge osnovne pojmove iz linearne algebre.

4 Skalarni proizvod i ortogonalnost vektora

Neka je V vektorski prostor nad poljem $\mathbb{R}$.

Skalarni proizvod na V jeste svaka funkcija $(\cdot,\cdot) : V \times V \to \mathbb{R}$ sa sledećim osobinama: $(\forall\ u, v, w \in V)(\forall\ \alpha \in \mathbb{R})$

1. $(u, v) = (v, u)$,
2. $(u + v, w) = (u, w) + (v, w)$,
3. $(\alpha u, v) = \alpha(u, v)$,
4. $(u, u) \geqslant 0$,
5. $(u, u) = 0 \iff u = 0$.

Norma vektora je funkcija $||\cdot|| : V \to \mathbb{R}$,

$$||u|| \stackrel{\text{def}}{=} \sqrt{(u, u)}.$$

Vektori su, po definiciji, **ortogonalni** ako im je skalarni proizvod jednak 0.

Ako $u_1, u_2, \ldots, u_k \in V$, matrica

$$G(u_1, u_2, \ldots, u_k) = \begin{bmatrix} (u_1, u_1) & (u_1, u_2) & \cdots & (u_1, u_k) \\ (u_2, u_1) & (u_2, u_2) & \cdots & (u_2, u_k) \\ \vdots & \vdots & \ddots & \vdots \\ (u_k, u_1) & (u_k, u_2) & \cdots & (u_k, u_k) \end{bmatrix}$$

naziva se **Gramova**[6] **matrica**, i može se pokazati da je ona regularna ako i samo ako su vektori $u_1, u_2, \ldots, u_k$ linearno nezavisni. Primetimo još toliko da za uzajamno ortogonalne vektore $u_1, u_2, \ldots, u_k$ Gramova matrica izgleda ovako:

$$\begin{bmatrix} ||u_1||^2 & 0 & \cdots & 0 \\ 0 & ||u_2||^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & ||u_k||^2 \end{bmatrix}.$$

Vratimo se na naš primer. Ukoliko posmatramo vektorski prostor neprekidnih funkcija na intervalu $[a, b]$, za dve funkcije f i g skalarni proizvod (f, g) najčešće se definiše na sledeći način:

$$(f, g) \stackrel{\text{def}}{=} \int\limits_a^b f(x)g(x)\,dx.$$

[6] Jørgen Pedersen Gram (1850-1916), danski matematičar.

Ako za $[a, b]$ uzmemo takav interval da $x_i \in [a, b]$ za sve $i = 1, 2, \dots, n$, mogli bismo (uz izvesne dodatne uslove) upravo ovako definisati skalarni proizvod za funkcije s kojima radimo.

Međutim, nas zanimaju samo vrednosti ovih funkcija u tačkama $x_1, x_2, \dots, x_n$, te će biti mnogo praktičnije da posmatramo sledeće n-dimenzionalne vektore:

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix},$$

$$\boldsymbol{\varphi}_k = \begin{bmatrix} \varphi_k(x_1) \\ \varphi_k(x_2) \\ \vdots \\ \varphi_k(x_n) \end{bmatrix}, \quad k = 1, 2 \dots, m,$$

i nad njima definišemo skalarni proizvod kao

$$(\boldsymbol{u}, \boldsymbol{v}) \stackrel{\text{def}}{=} \boldsymbol{u}^T \boldsymbol{v} = [u_1 \ u_2 \ \cdots \ u_n] \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n. \tag{2}$$

Ukoliko su φ_i i φ_j ortogonalne, važi:

$$(\boldsymbol{\varphi}_i, \boldsymbol{\varphi}_j) = 0,$$

tj.

$$\varphi_i(x_1)\varphi_j(x_1) + \varphi_i(x_2)\varphi_j(x_2) + \cdots + \varphi_i(x_n)\varphi_j(x_n) = 0.$$

5 Linearni najmanji kvadrati

Pogledajmo kako glase normalne jednačine za ovaj slučaj.

$$H(a_1, a_2, \ldots, a_m) = \sum_{i=1}^{n} \left(a_1\varphi_1(x_i) + a_2\varphi_2(x_i) + \cdots + a_m\varphi_m(x_i) - y_i\right)^2$$

$$\begin{aligned}
\frac{\partial H}{\partial a_k} &= \sum_{i=1}^{n} 2\,\varphi_k(x_i)\left(a_1\varphi_1(x_i) + a_2\varphi_2(x_i) + \cdots + a_m\varphi_m(x_i) - y_i\right) = \\
&= 2\sum_{i=1}^{n} \varphi_k(x_i)\left(\sum_{j=1}^{m} a_j\varphi_j(x_i) - y_i\right) = \\
&= 2\left(\sum_{i=1}^{n}\sum_{j=1}^{m} \varphi_k(x_i)a_j\varphi_j(x_i) - \sum_{i=1}^{n}\varphi_k(x_i)y_i\right)
\end{aligned}$$

Izjednačavajući s nulom ovakve izraze za svako $k = 1, 2, \ldots, m$ dobijamo **sistem linearnih jednačina**:

$$\sum_{j=1}^{m} a_j \sum_{i=1}^{n} \varphi_k(x_i)\varphi_j(x_i) - \sum_{i=1}^{n}\varphi_k(x_i)y_i = 0, \quad k = 1, 2, \ldots, m. \tag{3}$$

Dokažimo da ovaj sistem ima jedinstveno rešenje $(a_1^*, a_2^*, \ldots, a_m^*)$. To je tačno akko sistem

$$\sum_{j=1}^{m} a_j \sum_{i=1}^{n} \varphi_k(x_i)\varphi_j(x_i) = 0, \quad k = 1, 2, \ldots, m \tag{4}$$

ima jedinstveno rešenje. Pretpostavimo suprotno, dakle da postoje $a_1, a_2, \ldots, a_m$ koji nisu svi jednaki nuli i koji zadovoljavaju (4). Primetimo da tada važi:

$$\left\|\sum_{j=1}^{m} a_j\varphi_j\right\|^2 = \sum_{i=1}^{n}\sum_{j=1}^{m} a_j\varphi_j(x_i)\sum_{k=1}^{m} a_k\varphi_k(x_i) = \sum_{k=1}^{m} a_k \underbrace{\sum_{j=1}^{m} a_j \sum_{i=1}^{n}\varphi_k(x_i)\varphi_j(x_i)}_{=0} = 0,$$

što je u kontradikciji s pretpostavkom da su φ_i nezavisne. Time je dokaz završen.

Pogledajmo još i šta se dešava ukoliko su φ_i i φ_j ortogonalni za sve $i \neq j$. Tada dalje uprošćavamo (3):

$$0 = \sum_{j=1}^{m} a_j \underbrace{\sum_{i=1}^{n}\varphi_k(x_i)\varphi_j(x_i)}_{=0 \text{ za } k \neq j} - \sum_{i=1}^{n}\varphi_k(x_i)y_i = a_k\sum_{i=1}^{n}\varphi_k(x_i)^2 - \sum_{i=1}^{n}\varphi_k(x_i)y_i,$$

iz čega odmah dobijamo rešenja:

$$a_k^* = \frac{\sum_{i=1}^{n} \varphi_k(x_i) y_i}{\sum_{i=1}^{n} \varphi_k(x_i)^2}, \quad k = 1, 2, \ldots, m. \tag{5}$$

Primetimo da je imenilac sigurno različit od nule, budući da bi suprotno impliciralo $\varphi_k(x_i) = 0$ za sve $i = 1, 2, \ldots, n$, a to je nemoguće.

Naglasimo da iz same činjenice da je nešto rešenje normalnih jednačina **ne** sledi da je to ujedno i minimum funkcije H (znamo samo da je to stacionarna tačka). Ipak, lako se primećuje da je funkcija H konveksna (među sabircima od kojih je sačinjena svaki jeste konveksna funkcija, a zbir konveksnih funkcija opet je konveksna funkcija), te funkcija

$$\Phi^*(x) = a_1^* \varphi_1(x) + a_2^* \varphi_2(x) + \cdots + a_m^* \varphi_m(x)$$

jeste najbolja aproksimacija, i pri tom jedinstvena.

Sistem (3) možemo izraziti i preko matrica. Uočimo, pre svega, da ga možemo zapisati i kao:

$$\sum_{j=1}^{m} a_j (\boldsymbol{\varphi}_k, \boldsymbol{\varphi}_j) - (\boldsymbol{\varphi}_k, \boldsymbol{y}) = 0, \quad k = 1, 2, \ldots, m.$$

Dalje, neka je

$$B = \begin{bmatrix} \varphi_1(x_1) & \varphi_2(x_1) & \cdots & \varphi_m(x_1) \\ \varphi_1(x_2) & \varphi_2(x_2) & \cdots & \varphi_m(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1(x_n) & \varphi_2(x_n) & \cdots & \varphi_m(x_n) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\varphi}_1 & \boldsymbol{\varphi}_2 & \cdots & \boldsymbol{\varphi}_m \end{bmatrix} \in \mathbb{R}^{n \times m},$$

i primetimo da je $B^T B = G(\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \ldots, \boldsymbol{\varphi}_m)$ (Gramova matrica). Vidimo da sistem (3) predstavlja upravo jednakost

$$B^T B \boldsymbol{a} = B^T \boldsymbol{y}.$$

Iz ovoga sledi da sistem ima jedinstveno rešenje kada je $B^T B$ regularna matrica, što znači da je jedinstvenost rešenja zapravo posledica činjenice da je $G(\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \ldots, \boldsymbol{\varphi}_m)$ regularna ako i samo ako su vektori $\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \ldots, \boldsymbol{\varphi}_m$ linearno nezavisni (kao u našem slučaju). Ukoliko, pored toga, pretpostavimo i da su međusobno ortogonalni, sistem se svodi na:

$$\begin{bmatrix} ||\boldsymbol{\varphi}_1||^2 & 0 & \cdots & 0 \\ 0 & ||\boldsymbol{\varphi}_2||^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & ||\boldsymbol{\varphi}_m||^2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{\varphi}_1 \\ \boldsymbol{\varphi}_2 \\ \vdots \\ \boldsymbol{\varphi}_m \end{bmatrix} \boldsymbol{y},$$

što nam omogućuje da izrazimo

$$||\boldsymbol{\varphi}_k||^2 a_k = (\boldsymbol{\varphi}_k, \boldsymbol{y}),$$

a to je upravo zapis (5) preko vektora.

Za kraj, napomenimo da je sve ovo specijalan slučaj opštije teoreme iz linearne algebre, za skalarni proizvod definisan kao (2). U opštem slučaju može se reći:

Neka su funkcije $\varphi_1, \varphi_2, \ldots, \varphi_m$ baza prostora U. Tada postoji jedinstveno određena funkcija

$$\Phi^* = \sum_{j=1}^{m} a_j^* \varphi_j$$

takva da je

$$||\Phi^* - y||^2 \leqslant ||\Phi - y||,$$

za sve $\Phi \in U$. Funkcija Φ^* je određena jednačinama

$$(y - \Phi^*, \varphi_j) = 0, \quad j = 1, 2, \ldots, m.$$

Dokaz ove teoreme nećemo davati, a zainteresovane čitaoce upućujemo na [1], odakle je formulacija preuzeta.

6 Implementacija u programskom paketu *Mathematica*

Pokazaćemo kako se može pronaći aproksimacija linearnih najmanjih kvadrata korišćenjem programskog paketa *Mathematica*. Da bismo izračunali koeficijente $a_1, a_2, \ldots, a_m$ koristićemo se sistemom (3), koji ćemo rešavati komandom `Solve`. Sledi kod kojim se to postiže.

```
NajmanjiKvadrati[merenja_, baza_] :=
 Module[
  {n = Length[merenja], m = Length[baza], sistem, promenljive,
   resenje},
  sistem =
   Table[
    Sum[
      (baza[[k]] /.
         x -> merenja[[i, 1]]) (Sum[
          Subscript[a, j] (baza[[j]] /. x -> merenja[[i, 1]]), {j, 1,
           m}] - merenja[[i, 2]]),
      {i, 1, n}
      ] == 0,
    {k, 1, m}
    ];
  promenljive = Table[Subscript[a, l], {l, 1, m}];
  resenje = Solve[sistem, promenljive];
  Sum[resenje[[1, i, 2]] baza[[i]], {i, 1, m}]
  ]
```

Za $\varphi_1(x) \equiv 1$, $\varphi_2(x) = x$, $\varphi_3(x) = x^2$ i

x	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
y	0.65	0.35	0.15	0.04	0.01	0.035	0.155	0.36	0.645

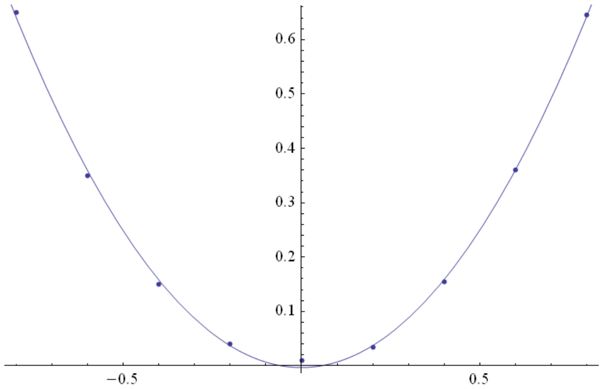
funkcija se poziva komandom

```
NajmanjiKvadrati[{{-0.8, 0.65}, {-0.6, 0.35}, {-0.4, 0.15}, {-0.2,
   0.04}, {0, 0.01}, {0.2, 0.035}, {0.4, 0.155}, {0.6, 0.36}, {0.8,
   0.645}}, {1, x, x^2}]
```

i vraća rezultat

```
-0.00311688 + 0.00125 x + 1.0096 x^2.
```

Pogledajmo grafik ovog primera:



Jasno je, ukoliko su funkcije φ_i međusobno ortogonalne, možemo upotrebiti i gotovu formulu (5). U tom slučaju imamo nešto jednostavniji kod:

```
NajmanjiKvadratiOrtogonalni[merenja_, baza_] :=
 Module[
  {n = Length[merenja], m = Length[baza]},
  Sum[baza[[
     k]] (Sum[(baza[[k]] /. x -> merenja[[i, 1]]) merenja[[i, 2]], {i,
         1, n}])/(Sum[(baza[[k]] /. x -> merenja[[i, 1]])^2, {i, 1,
        n}]),
   {k, 1, m}
   ]
  ]
```

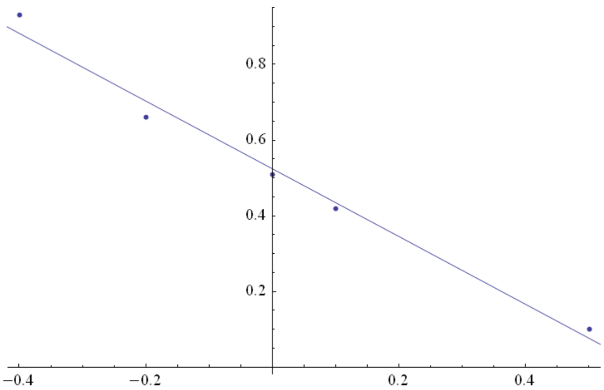
Primera radi, komanda

```
NajmanjiKvadratiOrtogonalni[{{-0.4, 0.93}, {-0.2, 0.66}, {0,
   0.51}, {0.1, 0.42}, {0.5, 0.1}}, {1, x}]
```

vraća rezultat

```
0.524 - 0.895652 x,
```

a grafik ovog primera je



Naravno, isto se dobija i korišćenjem funkcije `NajmanjiKvadrati`.

7 Nelinearni najmanji kvadrati

Ne postoji postupak koji rešava problem nelinearnih najmanjih kvadrata u opštem slučaju. Takvi problemi moraju se rešavati iterativnim postupcima dok ne postignemo zadovoljavajuću tačnost. Bez želje da detaljnije ulazimo u ovu problematiku, u ovom poglavlju dajemo samo uvodnu priču, onoliko koliko je (po našem mišljenju) potrebno da čitaoce zainteresuje za dalja istraživanja.

Neka je, kao i dosad,

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix},$$

i uvedimo još:

$$\boldsymbol{\Phi}(\boldsymbol{a}) = \begin{bmatrix} \Phi(x_1; a_1, a_2, \ldots, a_m) \\ \Phi(x_2; a_1, a_2, \ldots, a_m) \\ \vdots \\ \Phi(x_n; a_1, a_2, \ldots, a_m) \end{bmatrix}.$$

Cilj je da nađemo

$$\boldsymbol{a}^* = \begin{bmatrix} a_1^* \\ a_2^* \\ \vdots \\ a_m^* \end{bmatrix}$$

koje minimizuje izraz

$$||\boldsymbol{\Phi}(\boldsymbol{a}) - \boldsymbol{y}||,$$

što je, primetimo, vektorski zapis izraza (1).

Ako je funkcija $\Phi(x_i; a_1, a_2, \ldots, a_m)$ neprekidno diferencijabilna po a_i za sve $i = 1, 2, \ldots, m$, neka je

$$D\boldsymbol{\Phi}(\boldsymbol{a}) = \begin{bmatrix} \dfrac{\partial \Phi(x_1; a_1, a_2, \ldots, a_m)}{\partial a_1} & \dfrac{\partial \Phi(x_1; a_1, a_2, \ldots, a_m)}{\partial a_2} & \cdots & \dfrac{\partial \Phi(x_1; a_1, a_2, \ldots, a_m)}{\partial a_m} \\ \dfrac{\partial \Phi(x_2; a_1, a_2, \ldots, a_m)}{\partial a_1} & \dfrac{\partial \Phi(x_2; a_1, a_2, \ldots, a_m)}{\partial a_2} & \cdots & \dfrac{\partial \Phi(x_2; a_1, a_2, \ldots, a_m)}{\partial a_m} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial \Phi(x_n; a_1, a_2, \ldots, a_m)}{\partial a_1} & \dfrac{\partial \Phi(x_n; a_1, a_2, \ldots, a_m)}{\partial a_2} & \cdots & \dfrac{\partial \Phi(x_n; a_1, a_2, \ldots, a_m)}{\partial a_m} \end{bmatrix}$$

jakobijan funkcije $\mathbf{\Phi}$, i neka je

$$\boldsymbol{r}(\boldsymbol{a}) = \boldsymbol{y} - \mathbf{\Phi}(\boldsymbol{a}).$$

Navešćemo **Gaus-Njutnov**[7] **algoritam**, koji svodi problem na traženje niza linearnih najmanjih kvadrata, dok se ne dostigne zadovoljavajuća tačnost. Postupamo na sledeći način. Neka je data početna tačka $\boldsymbol{a}^{(0)}$. Iduće aproksimacije $\boldsymbol{a}^{(1)}, \boldsymbol{a}^{(2)}, \ldots$ dobijamo narednim procesom:

1. Za postojeće $\boldsymbol{a}^{(i)}$ izračunavamo $\boldsymbol{s}^{(i)}$ koje minimizuje izraz

$$\left|\left|D\mathbf{\Phi}\left(\boldsymbol{a}^{(i)}\right)\boldsymbol{s}^{(i)} - \boldsymbol{r}\left(\boldsymbol{a}^{(i)}\right)\right|\right|^2.$$

   Primetimo da ovde, u stvari, rešavamo problem linearnih najmanjih kvadrata.

2. Posmatramo $\xi(\tau) = \left|\left|\mathbf{\Phi}\left(\boldsymbol{a}^{(i)} + \tau\boldsymbol{s}^{(i)}\right) - \boldsymbol{y}\right|\right|^2$ i biramo najmanji nenegativan ceo broj k takav da važi:

$$\xi\left(2^{-k}\right) < \xi(0) = \left|\left|\boldsymbol{r}\left(\boldsymbol{a}^{(i)}\right)\right|\right|^2.$$

3. Uzimamo

$$\boldsymbol{a}^{(i+1)} = \boldsymbol{a}^{(i)} + 2^{-k}\boldsymbol{s}^{(i)}.$$

[7]Isaac Newton (1643-1727), engleski fizičar, matematičar i astronom.

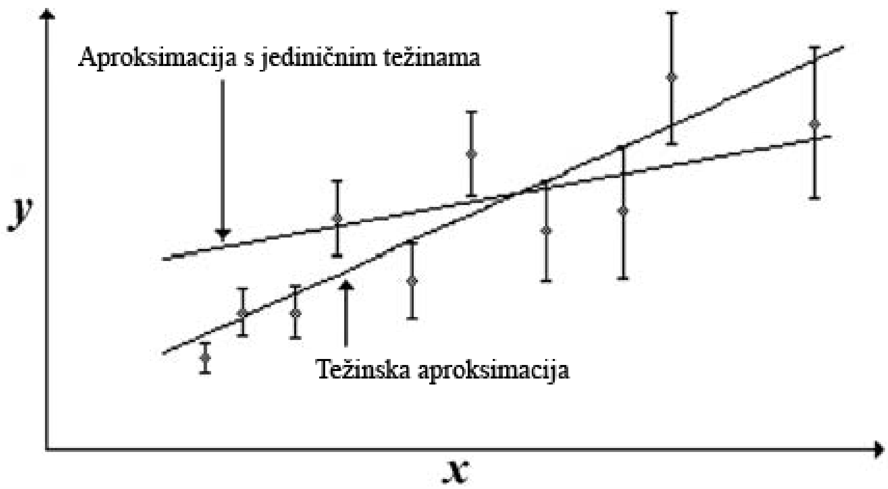
8 Težinski najmanji kvadrati

Na kraju, može se reći još nekoliko reči o **težinskim najmanjim kvadratima**. Prilikom vršenja nekog eksperimenta može se dogoditi da su neki dobijeni rezultati pouzdaniji od drugih i da im treba pridati veću važnost (ili veću "težinu") prilikom traženja dobre aproksimacije. Drugim rečima, u izraz (1) dodajemo još jedan set parametara, i tražimo minimum izraza

$$\sum_{i=1}^{n} w_i \left(\Phi(x_i; a_1, a_2, \ldots, a_m) - y_i\right)^2,$$

gde su $w_1, w_2, \ldots, w_n$ pozitivni realni brojevi. Oni su unapred zadati na osnovu očekivanog odstupanja rezultata merenja od prave vrednosti, tako što većem očekivanom odstupanju odgovara manja težina (nećemo se baviti detaljima). Jasno, ukoliko su sve težine jedinične, dobijamo dosad razmatrani slučaj.

Radi ilustracije efikasnosti ovog metoda navodimo donji grafik.



Očekivana odstupanja ubeležena su pomoću linija kod svakog rezultata merenja. Sa slike se jasno vidi koliko je težinska aproksimacija preciznija. Posebno obratimo pažnju na to koliko je prava dobijena aproksimacijom s jediničnim težinama udaljena od nekoliko prvih tačaka.

Većina dosad izloženog može se bez većih poteškoća uopštiti na težinske najmanje kvadrate. Zadržaćemo se samo na sistemu (3) i na formuli (5) (za međusobno ortogonalne bazne funkcije).

Zaista, u ovom slučaju, potpuno analogno kao u poglavlju 5, dobijamo sledeći sistem:

$$\sum_{j=1}^{m} a_j \sum_{i=1}^{n} w_i \varphi_k(x_i)\varphi_j(x_i) - \sum_{i=1}^{n} w_i \varphi_k(x_i) y_i = 0, \quad k = 1, 2, \ldots, m.$$

Sada ćemo zastati i razmotriti može li se skalarni proizvod definisati pogodnije nego sa (2). Ispostavlja se da može, naime:

$$(\boldsymbol{u}, \boldsymbol{v}) \stackrel{\text{def}}{=} w_1 u_1 v_1 + w_2 u_2 v_2 + \cdots + w_n u_n v_n.$$

Uz ovako definisan skalarni proizvod dokaz jedinstvenosti rešenja gornjeg sistema teče potpuno analogno onom datom u poglavlju 5, a za međusobno ortogonalne φ_i (jasno, s obzirom na novi skalarni proizvod) dobija se rešenje:

$$a_k^* = \frac{\displaystyle\sum_{i=1}^{n} w_i \varphi_k(x_i) y_i}{\displaystyle\sum_{i=1}^{n} w_i \varphi_k(x_i)^2}, \quad k = 1, 2, \ldots, m.$$

9 Literatura

1. D. Herceg, N. Krejić, *Numerička analiza za informatičare*, Prirodno-matematički fakultet, Departman za matematiku i informatiku, Novi Sad, 2006.

2. J. Stoer, R. Bulirsch, *Introduction to Numerical Analysis*, Springer, 1992.

3. P. Vanicek, D. E. Wells, *The Least Squares Approximation*, University of New Brunswick, Department of Geodesy and Geomatics Engineering, 1972.

4. J. Wolberg, *Data Analysis Using the Method of Least Squares*, Springer, 2006.

5. http://ceee.rice.edu/Books/LA/leastsq/index.html

6. http://en.wikipedia.org/wiki/Least_squares_method

7. http://mathworld.wolfram.com/LeastSquaresFitting.html