

Numerička matematika

14. predavanje

Saša Singer

`singer@math.hr`

`web.math.pmf.unizg.hr/~singer`

PMF – Matematički odsjek, Zagreb

Sadržaj predavanja

- Rješavanje sustava nelinearnih jednadžbi:
 - Newtonova metoda za nelinearne sustave.
 - Metoda sekante za nelinearne sustave.
- Uvod u optimizaciju bez ograničenja:
 - Formulacija problema.
 - Vrste metoda za optimizaciju bez ograničenja.

Rješavanje sustava nelinearnih jednadžbi

Formulacija problema

Zadana je funkcija

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^m.$$

Radi jednostavnosti, pretpostavljamo da je f definirana na cijelom prostoru \mathbb{R}^n .

Kao i prije, tražimo (jednu ili sve) točke $x \in \mathbb{R}^n$ za koje je

$$f(x) = 0 \in \mathbb{R}^m.$$

Najjednostavniji primjer je tzv. linearna (ili afina) funkcija f

$$f(x) = Ax - b, \quad x \in \mathbb{R}^n$$

gdje je $A \in \mathbb{R}^{m \times n}$ zadana pravokutna matrica s m redaka i n stupaca, a $b \in \mathbb{R}^m$ je zadani vektor.

Nultočke ove funkcije su rješenja linearnog sustava $Ax = b$.

Pojednostavljenja — dodatne pretpostavke

Raspisom vektorske jednadžbe $f(x) = 0$ po **komponentama** u prostoru \mathbb{R}^m , dobivamo **sustav** s m jednadžbi i n nepoznanica

$$f_i(x_1, x_2, \dots, x_n) = 0, \quad i = 1, \dots, m.$$

Dodatne pretpostavke u nastavku:

- funkcija f je, općenito, **nelinearna**,
- broj jednadžbi m **jednak** je broju nepoznanica n ,
tj. funkcija je $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$.

Osim toga, pretpostavljamo

- da f ima samo **izolirane** nultočke $\alpha \in \mathbb{R}^n$, i
- da je f dovoljno **glatka** — na cijelom \mathbb{R}^n , ili barem u nekoj **okolini** nultočke.

O generalizaciji metoda iz jedne dimenzije

Očita ideja za rješavanje sustava nelinearnih jednačbi je

- generalizacija metoda za rješavanje jedne jednačbe.

Problem: U više dimenzija nema uspoređivanja funkcijskih vrijednosti (vektora f) — osim po komponentama f_i .

Zato se “jednostavne” metode — poput bisekcije,

- teško generaliziraju, a i složenost postaje problem (imamo 2^n vrhova “kocke” i n komponentnih funkcija u svakom vrhu). Probajte zamisliti u \mathbb{R}^2 !

S druge strane, većina “iteracijskih” funkcija se

- relativno jednostavno generalizira — iz “skalarnih”, u “vektorske” ili “matrično–vektorske”.

Ilustracija toga = Newtonova metoda i metoda sekante.

Newtonova metoda u \mathbb{R}^n

Newtonovu metodu dobivamo linearizacijom funkcije f oko trenutne aproksimacije $x^{(k)} \in \mathbb{R}^n$ — slično kao za $n = 1$, samo su oznake malo drugačije (indeksi služe za komponente).

Izaberemo neku početnu točku $x^{(0)} \in \mathbb{R}^n$. Onda, induktivno, generiramo niz aproksimacija $x^{(k)}$, za $k \geq 0$, na sljedeći način.

Oko trenutne točke $x^{(k)}$, svaku komponentnu funkciju f_i

aproximiramo linearnim dijelom Taylorovog razvoja, tj. odbacujemo sve članove nakon linearnih. Dobivamo

$$f_i(x) \approx f_i(x^{(k)}) + \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(x^{(k)}) (x_j - x_j^{(k)}), \quad i = 1, \dots, n.$$

Newtonova metoda u \mathbb{R}^n

Iskoristimo **Jacobijevu** matricu $J_f(x)$, koja sadrži **parcijalne derivacije** svih komponentnih funkcija po svim varijablama

$$[J_f(x)]_{ij} = \frac{\partial f_i}{\partial x_j}, \quad i = 1, \dots, n, \quad j = 1, \dots, n,$$

ili

$$J_f(x) = \frac{\partial f}{\partial x}(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \dots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

Newtonova metoda u \mathbb{R}^n

Lineariziranu aproksimaciju za $f(x)$ oko točke $x^{(k)}$ možemo zapisati u **matrično–vektorskom** obliku

$$f(x) \approx f(x^{(k)}) + J_f(x^{(k)}) (x - x^{(k)}).$$

Novu aproksimaciju $x^{(k+1)}$ za nultočku dobivamo iz zahtjeva da je

$$f(x^{(k+1)}) \approx f(x^{(k)}) + J_f(x^{(k)}) (x^{(k+1)} - x^{(k)}) = 0.$$

Ako je $J_f(x^{(k)})$ **regularna** matrica, onda je pripadna **korekcija**

$$d^{(k)} := x^{(k+1)} - x^{(k)}$$

jedinstveno **rješenje linearnog** sustava jednadžbi

$$J_f(x^{(k)}) d^{(k)} = -f(x^{(k)}).$$

Newtonova metoda u \mathbb{R}^n

Za korekciju $d^{(k)}$ onda vrijedi (ali se **ne** računa ovako)

$$d^{(k)} = -[J_f(x^{(k)})]^{-1} f(x^{(k)}),$$

pa izlazi standardni zapis za iteracije u **Newtonovoj** metodi

$$x^{(k+1)} = x^{(k)} - [J_f(x^{(k)})]^{-1} f(x^{(k)}), \quad k = 0, 1, \dots$$

U usporedbi sa “**skalarnim**” oblikom Newtonove metode,

$$x^{(k+1)} = x^{(k)} - \frac{f(x^{(k)})}{f'(x^{(k)})},$$

- **množenje inverzom Jacobijeve** matrice $J_f(x)$, i to **slijeva**, je “zamjena” za **dijeljenje** derivacijom funkcije $f'(x)$.

Za računanje, **Jacobijeva** matrica $J_f(x^{(k)})$ mora biti **regularna**.

Lokalna konvergencija Newtonove metode u \mathbb{R}^n

Uočite da se rješavanje **jednog nelinearnog** sustava jednažbi svodi na **niz** rješavanja **linearnih** sustava.

Što se **konvergencije** tiče, vrijedi potpuni analogon rezultata o **lokalnoj** konvergenciji i **brzini** konvergencije u jednoj dimenziji.

Teorem. Neka je α nultočka funkcije $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ i neka je f klase C^2 u okolini od α .

Ako je $J_f(\alpha)$ **regularna** matrica (to je analogon **jednostruke** nultočke), onda **postoji** okolina nultočke α , takva da

- za bilo koju **početnu** točku $x^{(0)}$ iz te okoline,
- niz iteracija generiran **Newtonovom** metodom, **konvergira** prema α , i konvergencija je (barem) **kvadratna**,

tj. vrijedi
$$\|\alpha - x^{(k+1)}\| \in O(\|\alpha - x^{(k)}\|^2).$$



Metoda sekante u \mathbb{R}^n

Metodu **sekante** dobivamo iz **Newtonove** metode, slično kao u jednoj dimenziji.

Izaberemo **dvije početne** točke $x^{(0)}$ i $x^{(1)}$.

U točki $x^{(k)}$, za $k \geq 1$, sve **parcijalne derivacije** u **Jacobijevoj** matrici $J_f(x^{(k)})$,

- aproximiramo pripadnim **podijeljenim razlikama** kroz **zadnje** dvije aproksimacije $x^{(k)}$ i $x^{(k-1)}$.

Dobivamo “**sekantnu**” matricu

$$[S_f(x^{(k)})]_{ij} = \frac{f_i(x^{(k)}) - f_i(x^{(k-1)})}{x_j^{(k)} - x_j^{(k-1)}}, \quad i, j = 1, \dots, n.$$

Dalje postupamo isto kao u **Newtonovoj** metodi.

Metoda sekante u \mathbb{R}^n

Ako je $S_f(x^{(k)})$ regularna matrica, onda je pripadna korekcija

$$d^{(k)} := x^{(k+1)} - x^{(k)}$$

jedinstveno rješenje linearnog sustava jednadžbi

$$S_f(x^{(k)}) d^{(k)} = -f(x^{(k)}).$$

Oдавде izlazi standardni zapis za iteracije u metodi sekante

$$x^{(k+1)} = x^{(k)} - [S_f(x^{(k)})]^{-1} f(x^{(k)}), \quad k = 1, 2, \dots$$

Napomena. Korektnija oznaka za matricu je $S_f(x^{(k)}, x^{(k-1)})$, tako da se vidi ovisnost o zadnje dvije aproksimacije.

Za metodu sekante u \mathbb{R}^n vrijede slični rezultati kao u \mathbb{R} .

Uvod u optimizaciju bez ograničenja

Formulacija problema optimizacije

Zadana je funkcija

$$f : \mathbb{R}^n \rightarrow \mathbb{R}.$$

Opet, radi jednostavnosti, pretpostavljamo da je f definirana na cijelom prostoru \mathbb{R}^n .

Tražimo (jednu ili sve) točke $x^* \in \mathbb{R}^n$ za koje je

$$f(x^*) = \min f(x), \quad \text{ili} \quad x^* = \arg \min f(x).$$

Takve točke x^* zovu se točke **minimuma** funkcije f .

Problem **maksimizacije** ili traženja točaka **maksimuma** za funkciju f

• ekvivalentan je **minimizaciji** funkcije $g = -f$.

Zato se često kaže: “**optimizacija**” = “**minimizacija**”.

Vrste problema optimizacije

Ovisno o tome za koje x vrijedi ranija relacija, razlikujemo

- točke lokalnog minimuma — relacija vrijedi za sve x iz neke okoline od x^* ,
- točke globalnog minimuma — relacija vrijedi za sve x iz domene (kod nas je to cijeli \mathbb{R}^n).

Ovaj problem — bez restrikcija u domeni, zove se optimizacija bez ograničenja. Ako minimum tražimo na nekom podskupu $D \subset \mathbb{R}^n$, dobivamo problem optimizacije s ograničenjima.

Napomena. Većina numeričkih metoda za optimizaciju, nalazi (ili garantira) samo točke lokalnog minimuma.

Globalnost ili jedinstvenost se dobiva iz dodatnih svojstava funkcije f — na primjer, konveksnost od f (plus još ponešto).

Metode za optimizaciju bez ograničenja

Uočite da je **kodomena** funkcije f skup \mathbb{R} .

- 🔴 Zato ovdje **smijemo uspoređivati** funkcijske vrijednosti.

Metode za rješavanje problema optimizacije (bez ograničenja), ugrubo, dijelimo po tome **koje informacije** o funkciji f koriste (glatkoća).

- 🔴 Metode **direktnog pretraživanja** koriste samo **funkcijske** vrijednosti za f .
- 🔴 Metode **silaska** koriste i **derivacije** (gradijente) funkcije f , ili neke aproksimacije za njih.
- 🔴 Metode **Newtonovog** tipa koriste i **druge derivacije** (Hesseovu matricu) od f , ili neke aproksimacije za njih.