

Numerične metode 1 - definicije, trditve in izreki

Oskar Vavtar

2020/21

Kazalo

1	NUMERIČNO RAČUNANJE	3
1.1	Uvod	3
1.2	Premična pika	3
1.3	Občutljivost problema	4
1.4	Vrste napak pri numeričnem računanju	4
1.5	Stabilnost metode	5
1.6	Analiza zaokrožitvenih napak	5

1 NUMERIČNO RAČUNANJE

1.1 Uvod

Definicija 1.1 (Napaka). Pri numeričnem računanju izračunamo numerični približek za točno rešitev. Razlika med približkom in točno vrednostjo je *napaka* približka. Ločimo *absolutno* in *relativno* napako.

- absolutna napaka = približek – točna vrednost
- relativna napaka = $\frac{\text{absolutna napaka}}{\text{točna vrednost}}$

Naj bo x točna vrednost, \hat{x} pa približek za x .

- Če je $\hat{x} = x + d_a$, potem je $d_a = \hat{x} - x$ *absolutna napaka*.
- Če je $\hat{x} = x(1 + d_r)$, potem je $d_r = \frac{\hat{x} - x}{x}$ *relativna napaka*.

1.2 Premična pika

Definicija 1.2. Velja $\text{fl}(x) = x(1 + \delta)$ za $|\delta| \leq u$, kjer je

$$u = \frac{1}{2}b^{1-t}$$

osnovna zaokrožitvena napaka:

- single: $u = 2^{-24} = 6 \times 10^{-8}$
- double: $u = 2^{-53} = 1 \times 10^{-16}$

Izrek 1. Če za število x velja, da $|x|$ leži na intervalu med najmanjšim in največjim pozitivnim predstavljivim normaliziranim številom, potem velja

$$\frac{|\text{fl}(x) - x|}{|x|} \leq u.$$

1.3 Občutljivost problema

Definicija 1.3 (Občutljivost). Če se rezultat pri majhni spremembi argumentov (*motnji* oz. *perturbaciji*) ne spremeni veliko, je problem *neobčutljiv*, sicer pa je *občutljiv*.

1.4 Vrste napak pri numeričnem računanju

Definicija 1.4. Pri numeričnem računanju se pojavijo 3 vrste napak:

1. NEODSTRANLJIVE NAPAKE: Npr. ko podatek ni predstavljivo število. Namesto $y = f(x)$ lahko v najboljšem primeru izračunamo $\bar{y} = f(\bar{x})$, kjer je \bar{x} najbližje predstavljivo število.

$$D_n = y - \bar{y} = f(x) - f(\bar{x})$$

2. NAPAKA METODE: Npr. ko na voljo nimamo željene operacije. Namesto $f(\bar{x})$ potem izračunamo $\tilde{y} = g(\bar{x})$, kjer je $g(x)$ približek za $f(x)$, kjer znamo vrednost g izračunati s končnim številom operacij.

$$D_m = \bar{y} - \tilde{y} = f(\bar{x}) - g(\bar{x})$$

3. ZAOKROŽITVENE NAPAKE: Pri izračunu $\tilde{y} = g(\bar{x})$ lahko pri vsaki osnovni operaciji pride do zaokrožitvene napake, zato na koncu kot numeričen rezultat dobimo \hat{y} .

$$D_z = \tilde{y} - \hat{y}$$

Skupna napaka:

$$D = D_n + D_m + D_z$$

V splošnem lahko ocenimo:

$$|D| \leq |D_n| + |D_m| + |D_z|$$

1.5 Stabilnost metode

Definicija 1.5. Če metoda za $\forall x$ vrne \hat{y} , ki je *absolutno (relativno)* blizu točnemu y , je metoda *direktno stabilna*.

Če metoda za $\forall x$ vrne tak \hat{y} , da $\exists \hat{x}$ absolutno (relativno) blizu x , da je $\hat{y} = f(\hat{x})$ (točno), je metoda *obratno stabilna*.

V splošnem:

$$|\text{direktna napaka}| \leq |\text{občutljivost}| \cdot |\text{obratna napaka}|$$

1.6 Analiza zaokrožitvenih napak

Algoritem. Dana so predstavljljiva števila x_0, x_1, \dots, x_n ; računamo $p = x_0 \cdot x_1 \cdot \dots \cdot x_n$.

TOČEN ALGORITEM:

```
p0 = x0
i = 1, ..., n
    pi = pi-1 · xi
p = pn
```

DEJANSKI ALGORITEM:

```
 $\hat{p}_0 = x_0$ 
i = 1, ..., n
     $\hat{p}_i = \hat{p}_{i-1} \cdot x_i \cdot (1 + \delta_i) \quad |\delta_i| \leq u$ 
 $\hat{p} = \hat{p}_n$ 
```