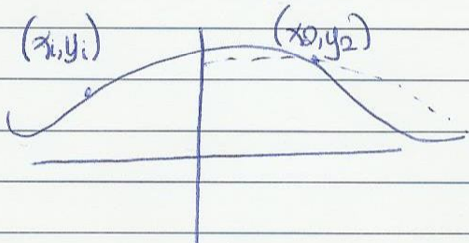
11/2/13

Αριθμητική Ανάλυση   1ο Κεφάλαιο

Περιεχόμενο του μαθήματος
(συνολικά)

1. Αριθμητική κινητής υποδιαστολής
Σφάλματα στρογγυλοποίησης
Ευαισθησία προβλημάτων και αλγορίθμων σε σφάλματα ~~στρογγυλοποίησης~~ στρογγύλευσης
2. Μη γραμμικές εξισώσεις
$f: [a,b] \to \mathbb{R}$
Ζητούνται ρίζες της f ⊥ 1η πρόοδος.
3. Γραμμικά συστήματα
A: $n \times n$ πίνακας
$b \in \mathbb{R}^n$
Ζητείται $x \in \mathbb{R}^n$ τ.ω. $Ax=b$ ⊥ 2η πρόοδος.
4. Παρεμβολή



$(x_i,y_i) \in \mathbb{R}^2$, $i=1,2,\dots,n$
$x_i \neq x_j$, για $i \neq j$

Ζητείται p κάποιας συγκεκριμένης μορφής τέτοια ώστε $p(x_i)=y_i$ $\forall i$, δηλαδή το γράφημα παρεμβάλλεται (περνάει) από τα σημεία.
5. Αριθμητική Ολοκλήρωση

$$\int_a^b f(x)dx = F(b)-F(a), \text{ αν } F'=f$$

$$\int_a^b f(x)dx \approx \sum_{i=1}^{n} w_i f(x_i)$$ ⊥ 3η Πρόοδος.

1. Αριθμητική κινητής υποδιαστολής
Σφάλματα στρογγύλευσης

Ποιότητα αριθμητικής μεθόδου

- απαιτούμενη μνήμη } σχετικά
- απαιτούμενος χρόνος
- ακρίβεια (μικρό σφάλμα) ~> βασικότερο

1ο Παράδειγμα

$$I_n = \int_0^1 x^n e^{x-1} dx \,,\ n \in \mathbb{N}$$

$e^{x-1} > 0$

(ακολουθία που το όριό της συγκλίνει (είναι) 0)

- $(I_n)_{n\in\mathbb{N}}$ φθίνουσα και μηδενική ακολουθία
- $I_n = 1 - n I_{n-1},\ n \geq 2$
- $I_1 = \dfrac{1}{e}$

Απόδειξη

$(I_n)_{n\in\mathbb{N}}$ φθίνουσα:

$x^{n+1} < x^n$ για $x \in (0,1)$

$\Rightarrow x^{n+1} \cdot e^{x-1} < x^n \cdot e^{x-1}$ για $x \in (0,1)$

$\Rightarrow \int_0^1 x^{n+1} \cdot e^{x-1} < \int_0^1 x^n \cdot e^{x-1}$

$\Rightarrow I_{n+1} < I_n$

$I_n > 0 \quad \forall n$

$(I_n)_{n\in\mathbb{N}}$ μηδενική: $0 < I_n = \int_0^1 x^n \underbrace{e^{x-1}}_{\leq 1} dx < \int_0^1 x^n dx = \dfrac{1}{n+1}$

$0 < I_n < \dfrac{1}{n+1}$ $\Rightarrow \boxed{\lim_{n\to\infty} I_n = 0}$

↳ $n\to\infty$ άρα συγκλίνει στο 0

$$I_n = \int_0^1 x^n e^{x-1}\,dx = \int_0^1 x^n (e^{x-1})'\,dx =$$

$$= x^n e^{x-1}\Big|_{x=0}^{x=1} - \int_0^1 (x^n)' e^{x-1}\,dx = 1 - \int_0^1 n x^{n-1} e^{x-1}\,dx =$$

$$= 1 - n \cdot I_{n-1}$$

για $n=1$: $I_1 = 1 - \int_0^1 x^{1-1} e^{x-1}\,dx = 1 - \int_0^1 e^{x-1}\,dx = 1 - e^{x-1}\Big|_{x=0}^{x=1} =$

$$= \cancel{1} - (\cancel{e^{1-1}} - e^{0-1}) = e^{-1} = 1/e$$

Αστάθεια

$$\left.\begin{array}{l} I_n = 1 - nI_{n-1} \\ \tilde{I}_n = 1 - n\tilde{I}_{n-1} \end{array}\right\} \Rightarrow I_n - \tilde{I}_n = -n(I_{n-1} - \tilde{I}_{n-1})$$

$\Rightarrow$ $\quad I_n - \tilde{I}_n = \boxed{(-1)^{n-1} n!}\,(I_1 - \tilde{I}_1)$

επαγωγή

η απόλυτη τιμή του αριθμού αυτού αυξάνει πολύ γρήγορα γεγονός που οδηγεί στην αστάθεια (πολύ μεγάλη)!

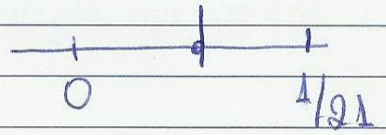
Ζητείται το I_{10}:

Χρησιμοποιώ τη σχέση $\quad I_{n-1} = \dfrac{1 - \tilde{I}_n}{n}$, $\quad n = 20, 19, \ldots 11$

- ευσταθής αλγόριθμος.

Με ποια τιμή θα ξεκινήσω;

$$0 < \tilde{I}_{20} < \frac{1}{21}$$



Ιδανικά ξεκινάω με $\tilde{I}_{20} = \dfrac{1}{42}$

12/2/13

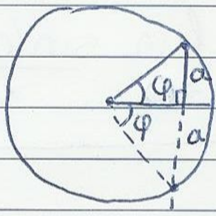
Αριθμητική Ανάλυση

2ο Παράδειγμα (Προσέγγιση του π με τη μέθοδο του Αρχιμήδη)

Έστω $y_n = 2^n \boxed{\sin\frac{\pi}{2^n}}$, $n \in \mathbb{N}$

$y_1 = 2\cdot\sin\frac{\pi}{2} = 2\cdot 1 = 2$

$y_2 = 4\cdot\sin\frac{\pi}{4} = 4\cdot\frac{\sqrt{2}}{2} = 2\sqrt{2}$



$\varphi = \frac{\pi}{2^n}$

$a = \sin\frac{\pi}{2^n}$

$2\varphi = \frac{2\pi}{2^n}$

$2a = 2\sin\frac{\pi}{2^n}$

όλες οι πλευρές και οι γωνίες είναι ίσες

$2a$ είναι η πλευρά του κανονικού πολυγώνου με 2^n πλευρές εγγεγραμμένου στον μοναδιαίο κύκλο

Η περίμετρος αυτού του πολυγώνου είναι $2^n\cdot 2a = 2^n\cdot 2\cdot\sin\frac{\pi}{2^n} =$

$= 2y_n$ άρα το y_n είναι η ημιπερίμετρος

Γεωμετρικά το $(y_n)_{n\in\mathbb{N}}$ είναι αύξουσα και $y_n \to \pi$, $n\to\infty$

Ισχυρισμός $(y_n)_{n\in\mathbb{N}}$ αύξουσα

$y_n = 2^n\sin\frac{\pi}{2^n}$

$\sin(2x) = 2\sin x\cos x$

$= \underbrace{2^n\cdot 2\cdot\sin\frac{\pi}{2^{n+1}}}_{y_{n+1}}\cdot\cos\frac{\pi}{2^{n+1}}$. Άρα $y_n = y_{n+1}\cdot\underbrace{\cos\frac{\pi}{2^{n+1}}}_{\text{θετικός} < 1}$

Συμπέρασμα: $\Rightarrow y_n < y_{n+1}$

<u>Ισχυρισμός.</u> $\lim\limits_{n\to\infty} y_n = \pi$

(+ γεωμετρική απόδειξη αν ζητηθεί)

$$\lim_{x\to 0}\frac{\sin x}{x} = 1 \overset{a\in\mathbb{R}}{\Rightarrow}$$

$$\Rightarrow \lim_{x\to 0}\frac{\sin ax}{ax} = 1 \Rightarrow \boxed{\lim_{x\to 0}\frac{\sin(ax)}{x} = a}$$

$$y_n = \frac{\sin\frac{\pi}{2^n}}{\frac{1}{2^n}}$$

$$\lim_{n\to\infty} y_n = \pi .$$

Ξεκινάω από το y_{n+1} για να βρω μια σχέση μεταξύ y_n και y_{n+1} χωρίς π

$$\left(\sin^2 x = \frac{1-\cos(2x)}{2}\right)$$

με ρίζα και παίρνω >0

$$y_{n+1} = 2^{n+1}\cdot\sin\frac{\pi}{2^{n+1}}$$

$$= 2^{n+1}\cdot\sqrt{\frac{1}{2}\left(1-\cos\frac{\pi}{2^n}\right)} = 2^{n+1}\cdot\sqrt{\frac{1}{2}\left[1-\sqrt{1-\underbrace{\sin^2\frac{\pi}{2^n}}_{\left(\frac{y_n}{2^n}\right)^2}}\right]}$$

$$\Rightarrow \begin{cases} y_{n+1} = 2^{n+1}\cdot\sqrt{\underbrace{\frac{1}{2}\left(1-\sqrt{1-(2^{-n}y_n)^2}\right)}_{\text{αφαίρεση σχεδόν ίσων αριθμών που οδηγεί σε αστάθεια}}}, \quad n\in\mathbb{N} \\ y_1 = 2 \end{cases}$$

$$\left(\sqrt{a}-\sqrt{b} = \frac{(\sqrt{a}-\sqrt{b})(\sqrt{a}+\sqrt{b})}{\sqrt{a}+\sqrt{b}} = \frac{a-b}{\sqrt{a}+\sqrt{b}} \right)$$

$$1-\sqrt{1-(2^{-n}\cdot y_n)^2} = \frac{1-1+(2^{-n}\cdot y_n)^2}{1+\sqrt{1-(2^{-n}\cdot y_n)^2}} = \frac{(2^{-n}\cdot y_n)^2}{1+\sqrt{1-(2^{-n}\cdot y_n)^2}}$$

Άρα

$$y_{n+1} = 2^{n+1}\cdot\sqrt{\frac{1}{2}\cdot\frac{(2^{-n}\cdot y_n)^2}{1+\sqrt{1-(2^{-n} y_n)^2}}} = 2\cdot\sqrt{\frac{1}{2(1+\sqrt{1-(2^{-n}\cdot y_n)^2})}}\cdot y_n =$$

$$= \sqrt{\frac{2}{1+\sqrt{1-(2^{-n}\cdot y_n)^2}}}\cdot y_n$$ → αυτός είναι Ευσταθής αλγόριθμος

Απο μαθητικής αποψης είναι και οι δύο σωστοί αλλά ο ασταθής είναι ευαίσθητος σε λάθη στρογγύλευσης.

14/2/13

Αριθμητική Ανάλυση

Παράσταση αριθμών ως προς οποιαδήποτε βάση

Καθημερινή ζωή: Δεκαδικό σύστημα
Βάση: 10
Ψηφία: 0, 1, ..., 9

Παράδειγμα: $3,14159 = 3\cdot10^{0}+1\cdot10^{-1}+4\cdot10^{-2}+1\cdot10^{-3}+5\cdot10^{-4}+9\cdot10^{-5}$ (, → με τελεία)

$\longrightarrow a_N a_{N-1} a_0 . a_{-1} a_{-2}$

Γενικά Έστω $a_N a_{N-1}, a_0, a_{-1}, a_{-2}, \dots$ δεκαδικά ψηφία (ακέραιοι από 0-9)

Τότε $(a_N a_{N-1}, \dots, a_1 a_0)_{10} =$

$$= \boxed{a_N\cdot10^{N}+a_{-1}\cdot10^{N-1}+\dots+a_1\cdot10^{1}+a_0\cdot10^{0}} + a_{-1}\cdot10^{-1}+a_{-2}\cdot10^{-2}+\dots$$

Ακέραιο μέρος: $a_N a_{N-1} \dots a_1 a_0$

$$P(x) = a_N x^N + a_{N-1}x^{N-1} + \dots + a_1 x + a_0$$

Το ακέραιο μέρος είναι η τιμή του P για $x=10$.

Κλασματικό μέρος $.a_{-1}a_{-2}\dots\dots\dots$

Η τιμή της δυναμοσειράς $\sum_{k=1}^{\infty} a_k x^k$ για $x = 1/10$

Η σειρά μπορεί να έχει είτε πεπερασμένο είτε άπειρο πλήθος όρων

Μοναδικότητα της παράστασης

$4,130 = 4,12\bar{9} = 4,1299999\dots$ → τα 9 αριθ. έχουν ως αποτέλεσμα την αύξηση κατά 1 του προηγούμενου στοιχείου

Αυτό οφείλεται στο γεγονός ότι.

$$9\cdot\underbrace{\sum_{i=1}^{\infty}10^{-i}}_{=\frac{\frac{1}{10}}{1-\frac{1}{10}}} = 1$$

→ γεωμετρική σειρά αθροίσματος, πιο γενικά $\Longrightarrow$

$$\frac{\frac{1}{10}}{1-\frac{1}{10}} = \frac{\frac{1}{10}}{\frac{9}{10}} = \frac{1}{9}$$

$$(\beta-1)\cdot\sum_{i=1}^{\infty}\beta^{-1}=1$$

Αν για κάθε $k_0\in\mathbb{N}$ υπάρχει $k\geq k_0$ τέτοιο ώστε

$a_{-k}\neq 9$ τότε η παράσταση είναι μοναδική

Γενική περίπτωση.

Σύστημα με βάση $\beta\geq 2$, $\beta\in\mathbb{N}$
Βάση: β
Ψηφία: $0,1,2,\ldots \beta-1$

Αν a_k ψηφία τότε $\pm(a_N a_{N-1}\ldots a_1 a_0 \cdot a_{-1}a_{-2})_\beta =$

$$=\pm(a_N\beta^N+a_{N-1}\beta^{N-1}+\ldots+a_1\beta^1+a_0\beta^0+$$
$$+a_{-1}\beta^{-1}+a_{-2}\beta^{-2}\ldots)$$

Παράδειγμα

$$(100110.11)_2=1\cdot2^5+0\cdot2^4+0\cdot2^3+1\cdot2^2+1\cdot2^1+0\cdot2^0+1\cdot2^{-1}+1\cdot2^{-2}=(38,75)_{10}$$

i) Μετατροπή από σύστημα με βάση β στο δεκαδικό

a) Ακεραίων αριθμών

Παράδειγμα $(53473)_8=5\cdot8^4+3\cdot8^3+4\cdot8^2+7\cdot8^1+3\cdot8^0=(22331)_{10}$

Τρόπος Υπολογισμού: $5\cdot8^4+3\cdot8^3+4\cdot8^2+7\cdot8+3=3+8(7+8(4+8(3+5\cdot8)))$

Σχήμα του Horner

Γενικά: $P(x)=a_Nx^N+a_{N-1}x^{N-1}+\ldots+a_1x+a_0=a_0+x(a_1+x(a_2+x(\ldots+x(a_{N-1}+xa_N))))$
και κάναμε τις πράξεις από "μέσα" προς τα "έξω"

$y\leftarrow a_N$ για $i=N-1,\ldots,0$

$\boxed{y\leftarrow a_i+xy}$ flop: διαδικασία να πολλαπλασιάζω δύο αριθμούς και να τους προσθέσω σ' έναν άλλον
Το σχήμα του Horner απαιτεί N flop
ο οικονομικότερος τρόπος.

Τελικά: $y=P(x)$

β) <u>Κλασματικών αριθμών</u> x. $(0<x<1)$

<u>Παράδειγμα</u>: $(.11)_2 = 1\cdot 2^{-1} + 1\cdot 2^{-2} = (0{,}75)_{10}$

ii) Μετατροπή από το δεκαδικό σ'ένα σύστημα με βάση β.

α) <u>Ακεραίων αριθμών</u>.

Βασίζεται στον αλγόριθμο της διαίρεσης.

<u>Παράδειγμα</u> $(369)_{10} \rightsquigarrow$ στο οκταδικό σύστημα.

$(369)_{10} = (\ldots a_2 a_1 a_0)_8 =$

$$= \underbrace{a_0}_{\text{ακέραιος}} + \underbrace{8(a_1 + 8(a_2 + \ldots)))}_{\text{ακέραιος}}$$

$0 \le a_0 \le 7$

<u>Άρα</u>: Το a_0 είναι το υπόλοιπο της διαίρεσης $369:8$ και το $a_1 + 8(a_2\ldots)$ είναι το αντίστοιχο πηλίκο.

369	8
1	46

Άρα $a_0 = 1$ και $46 = a_1 + 8(a_2 + \ldots)$

46	8
6	5

Άρα $a_1 = 6$ και $5 = a_2 + 8(a_3\ldots)$

Τώρα $a_2 = 5$ και $\underbrace{a_3 + 8(a_4\ldots)}$ $= 0$

$\Rightarrow a_3 = a_4 = \ldots = 0.$

<u>Συμπέρασμα</u>: $(369)_{10} = (561)_8$.

<u>Επαλήθευση</u>: $(561)_8 = 5\cdot 8^2 + 6\cdot 8 + 1 = 369$

• $0<x<1$ x στο δεκαδικό σύστημα

$$x=(.a_{-1}a_{-2}\ldots)_\beta = a_{-1}\cdot\beta^{-1}+a_{-2}\beta^{-2}+\ldots\ldots$$

$$\Rightarrow \beta x = a_{-1}+a_{-2}\beta^{-1}+\ldots..$$
$$= a_{-1}.a_{-2}\ldots\ldots$$

Συμπέρασμα: Το a_{-1} είναι το ακέραιο μέρος του βx

Παράδειγμα

$x=(.372)_{10} \rightsquigarrow$ στο δυαδικό σύστημα.

$$(.372)_{10}=(.a_{-1}a_{-2}\quad)_2$$

$2x=.744$, άρα $a_{-1}=0$, $y_1=.744$
$2y_1=1.488$, άρα $a_{-2}=1$, $y_2=.488$
$2y_2=0.976$, άρα $a_{-3}=0$, $y_3=.976$
$2y_3=1.952$, άρα $a_{-4}=1$, $y_4=.952$

Επομένως $(.372)_{10}=(.0101\ldots)_2$

Παράδειγμα: Ισχυρισμός $\displaystyle\sum_{n=1}^{\infty}\left(\frac{1}{2^{4n}}+\frac{1}{2^{4n+1}}\right)=\frac{1}{10}$

υπολογισμός γεωμετρικής σειράς

$$\sum_{n=1}^{\infty}\frac{3}{2}\cdot\frac{1}{2^{4n}}=\frac{3}{2}\sum_{n=1}^{\infty}\underbrace{\frac{1}{2^{4n}}}_{=\left(\frac{1}{2^4}\right)^n}=\frac{3}{2}\,\frac{\frac{1}{2^4}}{1-\frac{1}{2^4}}=\frac{3}{2}\,\frac{\frac{1}{16}}{\frac{15}{16}}=\frac{3}{2}\cdot\frac{1}{15}=\frac{1}{10}$$

$$\frac{1}{10}=\sum_{n=1}^{\infty}\left(\frac{1}{2^{4n}}+\frac{1}{2^{4n+1}}\right)=2^{-4}+2^{-5}+2^{-8}+2^{-9}+\ldots\ldots$$

$$\Rightarrow (.1)_{10}=(.\underset{2^{4}\ 2^{-5}}{00011}00110011)_2$$

• Αν κάποιο $\gamma_i = 0$, τότε σταματάμε, αφού όλα τα ψηφία από εκεί και πέρα θα είναι μηδέν.
• Στην περίπτωση $\gamma_{i+j} = \gamma_i$, για τα μικρότερα δυνατά i και j η παράσταση είναι περιοδική με περίοδο j
(Από πεπερασμένη μορφή θα είναι πεπερασμένη ή άπειρη η μορφή του και στο άλλο σύστημα.

Αριθμοί Μηχανής

Έστω $x \in \mathbb{R}, x \neq 0$. Υποθέτω σ' ένα σύστημα με βάση β ότι ο x μπορεί να γραφεί ως

$$x = \pm\, .d_1 d_2 \ldots \cdot \beta^e \quad \text{με } d_1 \neq 0$$

με d_i ψηφία του συστήματος και e κατάλληλος ακέραιος
Η μορφή (x) λέγεται (κανονική) μορφή κινητής υποδιαστολής.

Το σύνολο αριθμών μηχανής $M = M(\beta, t, L, U)$ χαρακτηρίζεται από τις παραμέτρους:
• β = βάση του αριθμητικού συστήματος
• t = ακρίβεια = πλήθος των στοιχείων του κλάσματος των αριθμών.
• L κάτω φράγμα } του εκθέτη e του β
• U άνω φράγμα } δηλαδή $L \leq e \leq U$

L, U ακέραιοι και $L \approx -U$

Κάθε $x \in M$, $x \neq 0$, είναι της μορφής $\boxed{x = \pm\, .d_1 d_2 \ldots d_t \cdot \beta^e}$ με $d_1 \neq 0$ και ο εκθέτης e $L \leq e \leq U$

Αριθμητική Ανάλυση 18/2/13

Αριθμοί μηχανής

$M = M(\beta, t, L, U)$

$x \in M$: · $x = 0$.
· $x = .d_1 d_2 \ldots d_t \cdot \beta^e$
$d_1 \neq 0$, $L \leq e \leq U$.

το M έχει πεπερασμένο πλήθος αριθμών.

Μέγιστο σε απόλυτη τιμή στοιχείο του M: (προκύπτει αν μεγαλώσω όλα τα στοιχεία και τον εκθέτη)

$d_i = \beta - 1$, $i = 1, \ldots, t$, και $e = U$

Ελάχιστο σε απόλυτη τιμή, μη μηδενικό στοιχείο του M:

$.10000 \cdot \beta^L$

Η απόσταση δύο διαδοχικών στοιχείων του M δεν είναι σταθερή

αν πολ/σω 2 αρ. μηχανής δεν παίρνω αναγκαστικά αρ. μηχανής

Το M δεν είναι κλειστό ως προς τον πολ/σμό.

δηλαδή

$x, x^* \in M \not\Rightarrow xx^* \in M$.

Παράδειγμα

$.1000\ldots0 \cdot \beta^L \cdot .1000\ldots0\,\beta^L \notin M$ (αφού προκύπτει μικρότερη απόλυτη τιμή από τη μικρότερη δυνατή)

Το M δεν είναι κλειστό ως προς την πρόσθεση

Παράδειγμα

$\beta = 10$, $t = 5$ $\quad 1, 10^{-5} \in M$.

διαφορετική τάξη

$1 + 10^{-5} = 1{,}00001$ · (δεν είναι στο M γιατί ξεπερνάνε τα ψηφία του καινούργιου αριθμού το t άρα δεν μπορεί να αναπαρασταθεί στη μηχανή

Μας ενδιαφέρει το M να είναι όσο πιο πυκνό και όσο πιο ευρύ γίνεται.

δηλαδή να έχει μεγάλο t και μεγάλο διάστημα [L, U]

πυκνό → αποστάσεις απόλυτων τιμών.

ευρύ → διαφορά μεγαλύτερο με μικρότερο

Προσέγγιση πραγματικών αριθμών με αριθμούς μηχανής

i. $|x| > .d_1 d_2 \dots d_t \cdot \beta^{U}$, με $d_i = \beta - 1$, $\forall\, i = 1, \dots, t$.

Υπερχείλιση (overflow) και σταματούν οι υπολογισμοί
(σπάνιο φαινόμενο και συνήθως εξαιτίας ασταθούς αλγόριθμου)

ii. $0 < |x| < .10\dots0 \cdot \beta^{+L}$

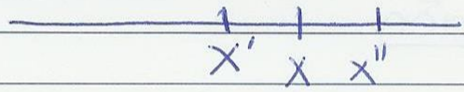
underflow → παλιότερα

Κατά κανόνα, σήμερα, ο x προσεγγίζεται με το μηδέν και συνεχίζονται οι υπολογισμοί

iii.

$$.1 \cdot \beta^{L} \le |x| \le .d_1 d_2 \dots d_t \cdot \beta^{U},\ d_i = \beta - 1$$

O x προσεγγίζεται με έναν αριθμό $fl(x) \in M$.



x', x'' διαδοχικοί αριθμοί μηχανής.

Υπόθεση: $|x - fl(x)| \le |x - y|\ \forall\, y \in M$

Ισχυρισμός: σχετικό σφάλμα είναι σημαντικότερο από το απόλυτο σφάλμα $(x - fl(x))$

$$\left|\frac{x - fl(x)}{x}\right| \le \frac{1}{2} \cdot \beta^{1-t}$$

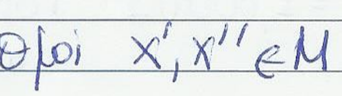
Απόδειξη

α) $fl(x) = x$ τότε η επίλυση είναι προφανής $(x \in M)$

β)

αν $x \notin M$, τότε υπάρχουν διαδοχικοί αριθμοί $x', x'' \in M$ τ.ω. $x' < x < x''$

Προφανώς ισχύει $|x - fl(x)| \le \frac{1}{2} |x' - x''|$



Άρα:

$\cdot \left|\frac{x-fl(x)}{x}\right| \leq \frac{1}{2}\frac{|x'-x''|}{|x|}$, Έστω ότι $x>0$. (αφού δεν παίζουν ρόλο τα πρόσημα)

$x = .d_1d_2...d_t d_{t+1}\cdot\beta^k$ → αυτό κάνει το x να μην είναι αριθμός μηχανής

Τότε

$x' = .d_1d_2...d_t\cdot\beta^k$

$x'' = (.d_1d_2...d_t+\beta^{-t})\cdot\beta^k$

$x''-x' = \beta^{k-t}$

δύο διαδοχικοί αριθμοί δεν έχουν σταθερή διαφορά αφού $x''-x'$ εξαρτάται από k

Άρα $\left|\frac{x-fl(x)}{x}\right| \leq \frac{1}{2}\frac{\beta^{k-t}}{x}$.

Τώρα $x > .1\cdot\beta^k$

άρα $\left|\frac{x-fl(x)}{x}\right| \leq \frac{1}{2}\frac{\beta^{k-t}}{.1\cdot\beta^k} = \frac{1}{2}\frac{\beta^k\cdot\beta^{-t}}{\beta^{-1}\cdot\beta^k} = \frac{1}{2}\beta^{1-t}$

Στρογγύλευση (αυτό που γίνεται στην πράξη)

π.χ. $\beta=10$, $t=5$

$x = .a_1a_2...a_5a_6...\cdot 10^k$

(Επιλέγω τον αριθμό ανάλογα με το τι κάνει ο a_6.
Μέχρι τον a_5 κρατάμε όλα τα στοιχεία από εκεί και πέρα τα πετάω όλα)

• Αν $a_6 \geq 5$, τότε

$fl(x) = x'' = (.a_1...a_5+10^{-5})\cdot 10^k$

• Αν $a_6 < 5$, τότε $fl(x) = x' = a_1...a_5\cdot 10^k$.

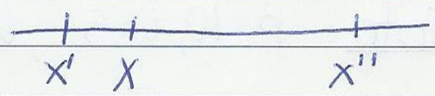
(στην περίπτωση που το $a_6=5$ και όλα τα άλλα $a_i=0$, $i\geq 7$, μπορούμε να επιλέξουμε το x' αντί του x'')

Δεύτερη δυνατότητα (δεν ισχύει η υπόθεση που κάναμε πριν).

Αποκοπή

Τότε:

$\left|\frac{x-fl(x)}{x}\right| \leq \beta^{1-t}$



Γενικά $\left|\frac{x-fl(x)}{x}\right| \le u$

$u = \frac{1}{2}\beta^{1-t}$ για στρογγύλευση

$u = \beta^{1-t}$ για αποκοπή

u: μοναδιαίο σφάλμα στρογγύλευσης (εδώ η μπορεί να είναι και η αποκοπή)

Πράξεις: $* \in \{+, -, \times, :\}$ συμβολίζουμε μια πράξη από τις τέσσερις

$x * y$ → μια πράξη.

Υπόθεση

(Ένας υπολογιστής μου δίνει έναν αριθμό) $z = fl(fl(x) \underset{(*)}{*} fl(y))$

Το λάθος στην υπόθεση είναι ότι υπέθεσα ότι η πράξη γίνεται σωστά/ακριβώς. αλλά γίνονται τουλάχιστον με διπλάσια ακρίβεια.

Παράδειγμα: $\beta=10$, $t=5$, $U=-L=10$. , στρογγύλευση.

$x = 5891,26$, $y = .0773414$

$fl(x) = .589126 \cdot 10^4$
$= .58913 \cdot 10^4$ (επειδή $6>5$)

$fl(y) = .773414 \cdot 10^{-1}$ (επειδή $4<5$) $= .000077341 \cdot 10^4$

$fl(x)+fl(y) = .589377341 \cdot 10^4$

$z = fl(fl(x)+fl(y)) = .58914 \cdot 10^4$

Παρατήρηση: $fl(fl(x)+fl(y)) \neq x+y$
$\neq fl(x+y)$
$\neq fl(x)+fl(y)$

Παράδοση: $\beta=10$, $t=5$, $U=-L=10$. , στρογγύλευση

$a_1 = 1$, $a_2 = a_3 = 3 \cdot 10^{-5}$

Τότε $a_1 + a_2$

$fl(a_1+a_2) = fl(1.00003) = 1$

Άρα παρόμοια

$fl(fl(a_1+a_2)+a_3) = 1$

$fl(a_2+a_3) = 6 \cdot 10^{-5}$

$fl(a_1 + fl(a_2+a_3)) = fl(1.00006) = 1{,}0001$

Συμπέρασμα: Έχει σημασία η σειρά με την οποία γίνονται οι πράξεις.

Παρατήρηση: Για κάθε x, $|x| < 10^{-5}$ έχουμε ότι $fl(1+x) = 1$

Γενικά: $|x| < \frac{1}{2} \beta^{1-t}$ ... έχουμε ότι $fl(1+x) = 1$.

$\frac{1}{2}\beta^{1-t}$ λέγεται έψιλον ή μηδέν μιας μηχανής. Και είναι το ίδιο με το μοναδιαίο σφάλμα στρογγύλευσης

Αριθμητική Ανάλυση 19/2/13

Άσκηση 1.2 (βιβλίο)

α) $1-\cos x$, $|x|$ μικρή (χωρίς ανάπτυγμα Taylor)

$1-\cos x = 2\sin^2 \frac{x}{2}$

β) $e^{x-y} = \frac{e^x}{e^y}$ γ) $\log x - \log y = \log \frac{x}{y}$

δ) $\sin(a+x) - \sin a$, $|x|$ μικρή

$\sin(a+x) - \sin a = 2\sin\frac{x}{2}\cos\left(\frac{2a+x}{2}\right)$

Άσκηση 1.3

$x^2 - 2ax + \beta = 0$ $x_1 = a + \sqrt{a^2-b}$ → ευσταθής τρόπος

$a, b > 0$, $a^2 >> b$ $x_2 = a - \sqrt{a^2-b}$ → ασταθής τρόπος

Αφού x_1, x_2 ρίζες τότε. (ή πολ/ζω το x_2 με τη συζυγή παράσταση)

$_1x^2 - 2ax + b = {}_1(x-x_1)(x-x_2)$ συντελεστής

$\Rightarrow \boxed{x_1 \cdot x_2 = b} \Rightarrow \boxed{x_2 = \frac{b}{x_1}}$ → ευσταθής τρόπος για το x_2

Άσκηση 1.4

$x^3+3px+2q=0$, $p,q \in \mathbb{R}$ $p^3+q^2>0$.

α) ν.δ.ο. η εξίσωση έχει ακριβώς μια πραγματική ρίζα ρ που δίνεται από τον τύπο του Cardano: $\rho = u-v$, με $u=(\sqrt{p^3+q^2}-q)^{1/3}$

και $v=(\sqrt{p^3+q^2}+q)^{1/3}$

α) $f(x)=x^3+3px+2q$

Ύπαρξη μιας πραγματικής ρίζας:

$\lim_{x\to\infty} f(x)=\infty$, $\lim_{x\to-\infty} f(x)=-\infty$

Άρα η f λαμβάνει τόσο θετικές όσο και αρνητικές τιμές.
Αφού η f είναι συνεχής, λαμβάνει και την τιμή 0, (λόγω θεωρήματος ενδιάμεσων τιμών), άρα έχει τουλάχιστον μια ρίζα

→ δεν έχει παραπάνω από ένα
Μοναδικότητα της ρίζας:

$f'(x)=3x^2+3p=3(x^2+p)$

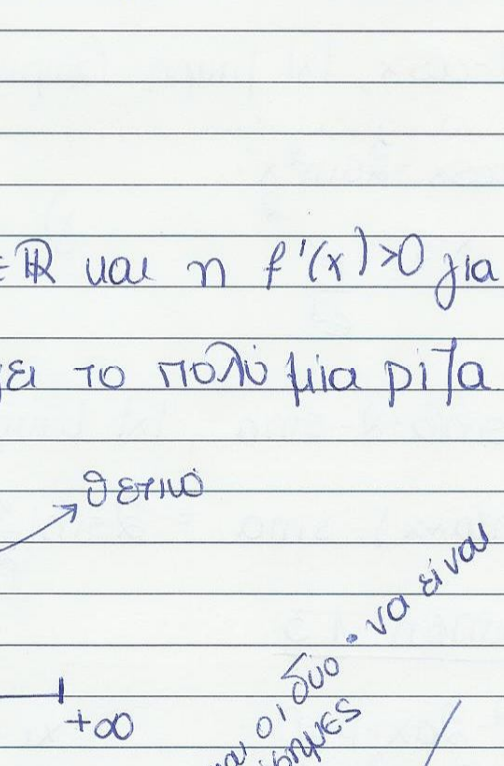
1η περίπτωση: $p\geq 0$, τότε $f'(x)\geq 0$ $\forall$ $x\in\mathbb{R}$ και η $f'(x)>0$ για $x\neq 0$

άρα η f είναι γνήσια αύξουσα, οπότε έχει το πολύ μια ρίζα

2η περίπτωση: $p<0$.

$f'(x)=0 \Leftrightarrow x^2+p=0 \Leftrightarrow x^2=-p \Leftrightarrow x_{1,2}=\pm\sqrt{-p}$ → θετικό

x	$-\infty$		$-\sqrt{-p}$		$\sqrt{-p}$		$+\infty$
f'		+	0	−	0	+	
f		↗		↘		↗	



Τώρα: $f(-\sqrt{-p}) = 2(q - p\sqrt{-p})$
$f(\sqrt{-p}) = 2(q + p\sqrt{-p})$

Οπότε

$$f(-\sqrt{-p})\cdot f(\sqrt{-p}) = 4(q^2+p^3) > 0.$$

i. $f(+\sqrt{-p}) > 0$: Τότε η f έχει ακριβώς μία ρίζα στο $(-\infty, -\sqrt{-p})$, καμμία ρίζα στο $[-\sqrt{-p}, \sqrt{-p}]$ και καμμία ρίζα στο $(\sqrt{-p}, +\infty)$

ii. $f(\sqrt{-p}) < 0$: Τότε η f έχει ακριβώς μία ρίζα στο $[\sqrt{-p}, \infty)$ καμμία στο $[-\sqrt{-p}, \sqrt{-p}]$ και καμμία στο $(-\infty, -\sqrt{-p}]$

Ισχυρισμός: $p = u - v$ τύπος Cardano.

$u = (\sqrt{p^3+q^2} - q)^{1/3}$
$v = (\sqrt{p^3+q^2} + q)^{1/3}$

$$f(p) = (u-v)^3 + 3p(u-v) + 2q = u^3 - 3u^2v + 3uv^2 - u^3 + 3p(u-v) + 2q =$$

$$= \cancel{-2q} - 3uv(u-v) + 3p(u-v) \cancel{+2q} = 3(p-uv)\cdot(u-v)$$

Τώρα:
$$u\cdot v = \{\underbrace{(\sqrt{p^3+q^2}-q)(\sqrt{p^3+q^2}+q)}_{p^3+\cancel{q^2}-\cancel{q^2}}\}^{1/3} = (p^3)^{1/3} = p$$ άρα κάνει 0.

β) $p^3 \gg q^2$

Τότε $u \approx \sqrt{p}$ και $v \approx \sqrt{p}$ οπότε στην αφαίρεση $u-v$ χάνεται ακρίβεια πράγμα που σημαίνει ότι ο αλγόριθμος είναι ασταθής. (για άρτιες δυνάμεις διαιρούνται με $(u-v)$ κι' $(u+v)$
για περιττές — " — — " — μόνο με $(u-v)$)

γ) $u^3 - v^3 = \underbrace{(u-v)}_{p}(u^2 + u\cdot v + v^2)$

$$p = \frac{u^3-v^3}{u^2+u\cdot v+u^2} = \frac{-2q}{u^2+p+u^2}$$ (τώρα δε χάνεται ακρίβεια)

θετικό, $p>0$.

Άσκηση 1.7 (βασική ιδέα)

a) $n \geq 3$, $y_n = n \sin\frac{\pi}{n}$, $Y_n = n \tan\frac{\pi}{n}$

ν.δ.ο. $y_n = \pi - \frac{\pi^3}{6n^2} + O\left(\frac{1}{n^4}\right)$ (για να βρεθεί αναπτύσσω το y_n κατά Taylor)

→ μεγαλύτερος εκθέτης → γρηγορότερη σύγκλιση.

$Y_n = \pi + \frac{\pi^3}{3n^2} + O\left(\frac{1}{n^4}\right)$ (οι δυνάμεις είναι περιττές παρόλο που δε φαίνεται γιατί πολ/ζουμε με n)

Ακολουθεί η πρόταση κατά Richardson.
(extrapolation)

$2y_n = 2\pi - \frac{\pi^3}{3n^2} + O\left(\frac{1}{n^4}\right)$. (θα προσθέσω κατά μέλη)

$Y_n = \pi + \frac{\pi^3}{3n^2} + O\left(\frac{1}{n^4}\right)$

$2y_n + Y_n = 3\pi + O\left(\frac{1}{n^4}\right)$

$\frac{2y_n + Y_n}{3} = \pi + O\left(\frac{1}{n^4}\right)$ (καλύτερη προσέγγιση για μεγάλα n από ότι αν παρέμειναν μόνα τους)

Αριθμητική Ανάλυση 22/2/13

Επιρροή των σφαλμάτων στρογγύλευσης στους υπολογισμούς.

$* \in \{+,-,\times,:\}$

x,y στο εύρος των αριθμών μηχανής, $x,y \neq 0$.

Ερώτημα: Τι μπορούμε να πούμε για το σχετικό σφάλμα;

$$\frac{fl(fl(x)*fl(y)) - x*y}{x*y} \;;$$

$\left|\frac{x-fl(x)}{x}\right| \le u$

a) $\frac{fl(x)-x}{x} = \varepsilon \Leftrightarrow fl(x) = x(1+\varepsilon)$, με $\overbrace{\varepsilon=\varepsilon(x)}^{\text{το } \varepsilon \text{ εξαρτάται από το } x.}$, $|\varepsilon| \le u$.

β) $\forall\, m \in \mathbb{N}$ και $\varepsilon_1, \dots, \varepsilon_m$ τέτοιο ώστε $|\varepsilon_i| \le u < 1$,
ισχύει: υπάρχει ε, $|\varepsilon| \le u$, τ.ω.

$$\underbrace{\prod_{i=1}^{m}(1+\varepsilon_i)}_{\text{έστω } \lambda} = (1+\varepsilon)^m.$$

$(1-u)^m \le \lambda \le (1+u)^m$

Θεωρούμε τη συνάρτηση $\varphi: [-u,u] \to \mathbb{R}$, $\varphi(x) = (1+x)^m$.

Τότε $\varphi(-u) \le \lambda \le \varphi(u)$

Σύμφωνα με το Θεώρημα της ενδιάμεσης τιμής, υπάρχει $\varepsilon \in [-u,u]$ τ.ω.
$\lambda = \varphi(\varepsilon) = (1+\varepsilon)^m$

$x = x(1+\varepsilon_1)$ $\quad |\varepsilon_i| \le u$.
$y = y(1+\varepsilon_2)$

a) Πολλαπλασιασμός

$z = fl(fl(x) \cdot fl(y)) =$
$= fl(x(1+\varepsilon_1) \cdot y(1+\varepsilon_2)).$
$= xy\underbrace{(1+\varepsilon_1)(1+\varepsilon_2)(1+\varepsilon_3)}_{=(1+\varepsilon)^3}$, $|\varepsilon_3| \le u$ $\quad |\varepsilon| \le u$
$= x \cdot y(1+\varepsilon)^3$

$$\left|\frac{z-xy}{xy}\right| = \left|(1+\varepsilon)^3 - 1\right| = |3\varepsilon + 3\varepsilon^2 + \varepsilon^3| \le 3u + 3u^2 + u^3$$
$$\le 3u + 4u^2$$ επειδή είναι πολύ μικρό

Άρα λέμε.
Το σχετικό σφάλμα στον πολλ/σμό είναι το πολύ τριπλάσιο του μοναδιαίου σφάλματος στρογγύλευσης.

β). Διαίρεση

$$z = fl\left(\frac{fl(x)}{fl(y)}\right) = \frac{x(1+\varepsilon_1)}{y\boxed{(1+\varepsilon_2)}} \cdot (1+\varepsilon_3) = \frac{x}{y}\underbrace{(1+\varepsilon_1)(1+\varepsilon_3)}_{\shortparallel \atop (1+\varepsilon)^2}(1+\delta)$$

το αφήνω προς το παρόν έτσι γιατί δεν ξέρω αν το $|1+\delta|<u$.

$$\frac{1}{1+\varepsilon_2} = \delta + 1 \Rightarrow \delta = \frac{-\varepsilon_2}{1+\varepsilon_2} \Rightarrow |\delta| \le \frac{u}{1-u} = u(1+u+O(u^2)) = u + O(u^2)$$

$$\left|\frac{z-\frac{x}{y}}{\frac{x}{y}}\right| = \left|(1+\varepsilon)^2(1+\delta) - 1\right| = |2\varepsilon + \delta + \varepsilon^2 + 2\varepsilon\delta + \delta\varepsilon^2|$$
$$\le 3u + a, \text{ με } a = O(u^2)$$

γ) Πρόσθεση - Αφαίρεση

$$z = fl(fl(x) + fl(y))$$
$$= fl(x(1+\varepsilon_1) + y(1+\varepsilon_2))$$
$$= x\underbrace{(1+\varepsilon_1)(1+\varepsilon_3)}_{(1+\varepsilon)^2} + y\underbrace{(1+\varepsilon_2)(1+\varepsilon_3)}_{(1+\delta)^2}$$

$|\delta|, |\varepsilon| \le u$

Άρα

$$z = (x+y) + (2x\varepsilon + 2y\delta) + \underbrace{(x\varepsilon^2 + y\delta^2)}_{O(u^2)}$$ (τάξης u^2 άρα το ξεχνάω)

$$\Rightarrow \frac{z-(x+y)}{x+y} \simeq 2\,\frac{x\varepsilon + y\delta}{x+y}$$

ξέρω ότι είναι $|\varepsilon|, |\delta| \le u$.

$$\Rightarrow \left|\frac{z-(x+y)}{x+y}\right| \simeq 2\left|\frac{x\varepsilon + y\delta}{x+y}\right| \le 2\,\frac{|x|+|y|}{|x+y|}\cdot u$$

1η Περίπτωση: x,y ομόσημοι

Τότε $|x+y| = |x|+|y|$ άρα

$$\left|\frac{z-(x+y)}{x+y}\right| \le 2u + O(u^2)$$

2η Περίπτωση: x,y ετερόσημοι, στη χειρότερη περίπτωση θα έχουμε

$\varepsilon \approx -\delta \approx u$, οπότε:

$$\left|\frac{z-(x+y)}{x+y}\right| \approx \frac{|x-y|}{|x+y|} \cdot u$$

Το υλάσμα μπορεί να είναι πολύ μεγάλο όταν αφαιρούμε σχεδόν ίσους αριθμούς.

* Η αφαίρεση σχεδόν ίσων αριθμών (μπορεί να) έχει καταστροφική επιρροή στο αποτέλεσμα

Ειδική Περίπτωση: $x,y \in M$ (να είναι αριθμοί μηχανής)

$z = fl(x+y)$
$= (x+y)(1+\varepsilon) \quad |\varepsilon| \le u.$

$\Rightarrow z-(x+y) = (x+y)\varepsilon.$

$$\Rightarrow \left|\frac{z-(x+y)}{x+y}\right| = |\varepsilon| \le u$$

Παράδειγμα: $\beta=10, t=5, U=-L=10$, στρογγύλευση

$x = .\boxed{451}42708$
$y = -.\boxed{451}15944$

Επειδή τα τρία πρώτα ψηφία είναι ίσα και αλληλοαναιρούνται χάνεται η ακρίβεια από 5 σε 2 ψηφία, γι'αυτό και τα 3 ψηφία είναι 0

$x+y = .26764 \cdot 10^{-3}$

$z = fl(fl(x)+fl(y))$
$= Fl(.45143 - .45116) = .00027 = .27000 \cdot 10^{-3}$

$\left|\dfrac{z-(x+y)}{x+y}\right| \simeq 88 \cdot 10^{-4}$ Αυτό το σφάλμα είναι 88 φορές μεγαλύτερο από ότι αν οι αριθμοί ήταν ομόσημοι

$2u = 10^{-4} \left(\frac{1}{2}\beta^{1-t}\right)$

Πως μπορούμε να αποφύγουμε την αφαίρεση σχεδόν ίσων αριθμών;

1º Παράδειγμα

$\beta=10, t=10.$

ένδειξη ότι χάνεται ακρίβεια

$\sqrt{7298}-\sqrt{7297} = .562847\boxed{0000} \cdot 10^{-2}$

$\left[\left(\sqrt{x}-\sqrt{y}\right) = \left(\dfrac{x-y}{\sqrt{x}+\sqrt{y}}\right)\right]$

πολύ καλή(τερη) προσέγγιση

$\sqrt{7298}-\sqrt{7297} = \dfrac{1}{\sqrt{7298}+\sqrt{7297}} = .5628468294 \cdot 10^{-2}$

2º Παράδειγμα

$f(x) = x - \sin x$, $|x|$ μικρή

Ανάπτυγμα Taylor:

$\varphi(x) = \varphi(x^*) + (x-x^*)\varphi'(x^*) + \dfrac{(x-x^*)^2}{2!}\varphi''(x^*) + \dfrac{(x-x^*)^3}{3!}\varphi'''(\theta)$

με θ μεταξύ x και y.

$\sin x = x - \dfrac{x^3}{6} + \varepsilon(x)$ με $|\varepsilon(x)| < \dfrac{|x|^5}{120.} \to 5!$

$\Rightarrow f(x) = \dfrac{x^3}{6} - \varepsilon(x) \approx \dfrac{x^3}{6}.$

Σφάλματα στον υπολογισμό αθροισμάτων.

Παράδειγμα: $S_n = 1+\sum_{k=1}^{n} \frac{1}{k^2+k}$, $n \in \mathbb{N}$

$\frac{1}{k^2+k} = \frac{1}{k} - \frac{1}{k+1}$

Τηλεσκοπικά τα αθροίσματα στα οποία αλληλοαναιρούνται οι όροι (trivia)

$\Rightarrow S_n = 1+\sum_{k=1}\left(\frac{1}{k}-\frac{1}{k+1}\right)$

$= 1+\left(1-\cancel{\frac{1}{2}}\right)+\left(\cancel{\frac{1}{2}}-\cancel{\frac{1}{3}}\right)+\left(\cancel{\frac{1}{3}}-\frac{1}{4}\right)+\dots+\left(\frac{1}{n}-\frac{1}{n+1}\right)$

Άρα $S_n = 2-\frac{1}{n+1}$ άρα (S_n) γνήσια αύξουσα.

και $S_n \to 2, n\to\infty$

$S_{9999} = 1.9999$, ακριβής τιμή

Αναδρομικός τύπος:

$S_0=1, S_k = S_{k-1}+\frac{1}{k(k+1)}$, $k=1,\dots,n$.

B=10, t=10: $\tilde{S}_{9999} = 1.999899972$

Προσθέσαμε θετικούς αριθμούς αρχίζοντας από το μεγαλύτερο προς το μικρότερο

Με $T_0=\frac{1}{n(n+1)}$, $T_k = T_{k-1}+\frac{1}{(n-k)(n-k+1)}$, $k=1,\dots,n-1,\dots$

$T_n = T_{n-1}+1$. Τότε θεωρητικά $\boxed{T_n=S_n}$

Στον υπολογιστή παίρνουμε:

$\tilde{T}_{9999} = 1.999900000$ που είναι το ακριβές αποτέλεσμα.

Αθροίσαμε θετικούς αριθμούς από το μικρότερο προς το μεγαλύτερο

Αριθμητική Ανάλυση 26/2/13

(Βοηθητικό για να λύσω το πρόβλημα)

Παρατήρηση: Για $\lambda, \mu \in \mathbb{R}$ και $\varepsilon_1, \varepsilon_2 \in [-u,u]$, υπάρχει $\varepsilon_3 \in [-u,u]$: $\lambda\varepsilon_1 + \mu\varepsilon_2 = (|\lambda|+|\mu|)\varepsilon_3$. $\varepsilon_3 = \dfrac{\lambda\varepsilon_1+\mu\varepsilon_2}{(|\lambda|+|\mu|)} \Rightarrow |\varepsilon_3| = \dfrac{|\lambda\varepsilon_1+\mu\varepsilon_2|}{|\lambda|+|\mu|} \leq$

$\leq \dfrac{|\lambda|u+|\mu|u}{|\lambda|+|\mu|} = u$

Πρόβλημα: Έστω $a_1, \ldots, a_N \in \mathbb{R}$. Θέλουμε να υπολογίσουμε το άθροισμα $S_N = \sum_{i=1}^{N} a_i$. Θεωρούμε τον αλγόριθμο: $S_1 = a_1$ (και προσθέτω στο S_{k-1} τον a_k και το ονομάζω S_k) $S_k = S_{k-1} + a_k$, $k = 2, \ldots N$.

Παίρναμε τις προσεγγίσεις:

$\tilde{S}_1 = a_1$, $\tilde{S}_k = fl(\tilde{S}_{k-1} + a_k)$, (παίρνω το fl γιατί δεν ξέρω αν το άθροισμα είναι αριθμός μηχανής.)
$k = 2, \ldots N$.

Έχουμε $\tilde{S}_2 = fl(\tilde{S}_1 + a_2) = \underbrace{(a_1+a_2)}_{S_2}(1+\delta) = S_2 + S_2\delta = S_2 + |S_2|\varepsilon_2$

με $|\delta|, |\varepsilon_2| \leq u$.

$\tilde{S}_2 = S_2 + |S_2|\varepsilon_2$.

Με τον ίδιο τρόπο

$\tilde{S}_3 = (\tilde{S}_2 + a_3)(1+\delta') = (S_2 + |S_2|\varepsilon_2 + a_3)(1+\delta') = (S_3 + |S_2|\varepsilon_2)(1+\delta')$

Άρα $\tilde{S}_3 = S_3 + |S_2|\varepsilon_2 + S_3\delta + \underbrace{|S_2|\varepsilon_2\delta'}_{O(u^2)} \approx S_3 + |S_2|\varepsilon_2 + S_3\delta' \xrightarrow{\text{Παρατήρηση}}$ (τα $\approx$ σημαίνει ότι παραλείπουμε όρους της τάξης $O(u^2)$)

$\Rightarrow \tilde{S}_3 \approx S_3 + (|S_2|+|S_3|)\varepsilon_3$, συνεχίζοντας ανάλογα παίρνουμε

$\tilde{S}_N \approx S_N + (|S_2|+|S_3|+\ldots+|S_n|)\varepsilon_N$ με $|\varepsilon_N| \leq u$

ο λόγος είναι μεγάλος αν οποιοδήποτε ενδιάμεσου αθροίσματος είναι πολύ μεγάλη.

$\Rightarrow \left|\dfrac{\tilde{S}_N - S_N}{S_N}\right| \approx \left|\dfrac{|S_2|+|S_3|+\ldots+|S_N|}{|S_N|}\cdot\varepsilon_N\right| \leq \dfrac{|S_2|+|S_3|+\ldots+|S_N|}{|S_N|}\, u$

Θέτουμε $\gamma_N := |s_2| + |s_3| + \dots + |s_N|$ και $p_N := \frac{\gamma_N}{|s_N|}$.

Το p_N είναι ο συντελεστής μεγέθυνσης στον υπολογισμό του αθροίσματος. Για μεγάλο p_N ο αλγόριθμος είναι ασταθής, για μικρό p_N είναι ευσταθής. Ιδιαίτερα, αν η απόλυτη τιμή $|s_i|$ ενός ενδιάμεσου αθροίσματος είναι πολύ μεγαλύτερη της $|s_N|$, ο αλγόριθμος είναι ασταθής.

Ειδική Περίπτωση: $a_i > 0, \forall i = 1 \dots N$. Τότε (θετικοί)

$$\gamma_N = (N-1)a_1 + (N-1)a_2 + (N-2)a_3 + \dots + 2a_{N-1} + a_N.$$

καλή επιλογή

Το γ_N ελαχιστοποιείται για $a_1, a_2 \le a_2 \le a_3 \le \dots \le a_N$, και

μεγιστοποιείται για $a_1, a_2 \ge a_3 \ge a_4 \ge \dots \ge a_N$

Παράδειγμα: Προσέγγιση του e^{-x} για $x \gg 1$. Θεωρούμε

$$s_N(x) = 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \dots + (-1)^{N-1}\frac{x^{N-1}}{(N-1)!}$$

Για αρκετά μεγάλο N έχουμε $s_N(x) \approx e^{-x}$

Για $x = 100$ έχουμε $e^{-100} \approx 0$, ενώ

$s_1 = 1$, $s_2 = 1 - 100 = -99$, $s_3 = 4901$, $s_4 \approx -161766 \dots$ κλπ

Το p_N είναι τεράστιο, αποτυχία καταστροφική.

$e^{-x} = \frac{1}{e^x}$. προσεγγίζουμε το e^x και διαιρούμε το 1 με το αποτέλεσμα (ευσταθής αλγόριθμος 2η περίπτωση.

Ευστάθεια αλγορίθμων

Ορισμός: Ένας αλγόριθμος λέγεται ασταθής αν είναι ευαίσθητος σε σφάλματα στρογγυλεύσης, δηλαδή αν μικρά σφάλματα σφάλματα στρογγυλεύσης που γίνονται στην παράσταση των αριθμών κατά τις πράξεις, μπορούν να επηρεάσουν πολύ (να αλλοιώσουν) το τελικό αποτέλεσμα. Σε αντίθετη περίπτωση λέγεται ευσταθής.

Κατάσταση Προβλημάτων

Λέμε ότι ένα πρόβλημα έχει <u>καλή κατάσταση</u> αν μικρές μεταβολές στα δεδομένα του έχουν ως αποτέλεσμα μικρή μεταβολή της λύσης του.
Λέμε ότι ένα πρόβλημα έχει <u>κακή κατάσταση</u> αν μικρές μεταβολές στα δεδομένα του μπορούν να έχουν ως αποτέλεσμα μεγάλη μεταβολή της λύσης του.

Παραδείγματα

$(x-2)^6=0 \Rightarrow x^*=2$

$(x-2)^6=10^{-6} \Rightarrow x_k-2=10^{-1}\underbrace{\sqrt[6]{1}}_{e^{\frac{2i\pi k}{6}}}$

($\sqrt[n]{1}$, $n\in\mathbb{N}$, η αριθμοί οι κορυφές του κανονικού η-γώνου, εγγεγραμμένου στον μοναδιαίο κύκλο, κορυφή τη μονάδα)

$\boxed{|x_k-x^*|=\frac{1}{10}}$ → κακή κατάσταση γιατί $\frac{1}{10} >>>> 10^{-6}$.
(η μεταβολή της λύσης είναι 10^5 φορές η μεταβολή στα δεδομένα)

$(x-2)^6=0 \Leftrightarrow x-2=0$ → καλή κατάσταση (χάνεται η πολλαπλότητα)

Αριθμητική Ανάλυση 26/2/13

Άσκηση (όλες οι ποσότητες είναι μη αρνητικές)

$$y_n=\int_0^1 \frac{x^n}{x+a}\,dx\ ,\ n=0,1,2,\ldots$$

θα μικρύνει αν x^{n+1} γιατί το $x\in(0,1)$.

α) Ν.Δ.Ο. για $a>0$ η $(y_n)_{n\in\mathbb{N}_0}$ είναι φθίνουσα και μηδενική

για $x\in(0,1)$ ισχύει $x^{n+1}<x^n$. Άρα $\frac{x^{n+1}}{x+a}<\frac{x^n}{x+a}$

$$\Rightarrow \underbrace{\int_0^1\frac{x^{n+1}}{x+a}dx}_{y_{n+1}} < \underbrace{\int_0^1\frac{x^n}{x+a}dx}_{y_n} \Rightarrow y_{n+1}<y_n$$

$$0\le y_n=\int_0^1\frac{x^n}{x+a}dx$$ (θ.ν.δ.ο. το $\int$ πάει στο 0)

$$\int_0^1\frac{x^n}{x+a}dx\le\int_0^1\frac{x^n}{a}dx=\frac{1}{a}\cdot\frac{1}{n+1}$$

$$\left.\begin{array}{l}\Rightarrow 0\le y_n\le\underbrace{\frac{1}{a}\cdot\frac{1}{n+1}}_{\to 0 \text{ για } n\to\infty}\end{array}\right\}\Rightarrow y_n\to 0,\ n\to\infty$$

β) Για $n\ge1$, ισχύει

$$y_n=\sum_{k=0}^{n-1}(-1)^k\boxed{\binom{n}{k}a^k\frac{(1+a)^{n-k}-a^{n-k}}{n-k}}+(-a)^n\log\frac{1+a}{a}$$

→ καθένας από αυτούς τους όρους έχει μεγάλη απόλυτη τιμή → άρα είναι ασταθής.

$$y_n=\int_0^1\frac{x^n}{x+a}dx \overset{\boxed{y=x+a}}{=} \int_a^{1+a}\frac{(y-a)^n}{y}dy$$

$$\left((a+b)^n=\sum_{k=0}^{n}\binom{n}{k}a^n b^{n-k}\right.$$
$$\left.\binom{n}{k}=n!/[k!(n-k)!]\right)$$