Ε Ν Ο Τ Η Τ Α 1

ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ

1.1. Κάθε αριθμός στο *δεκαδικό σύστημα* μπορεί να γραφεί στη μορφή

(i) $\left(\alpha_n \cdot \alpha_{n-1} \cdots \alpha_2\alpha_1\alpha_0\right)_{10} = \sum_{i=0}^{n} \alpha_i \text{ x } 10^i$, α_i =0, 1, 2, ... , 9, i = 0, 1, 2, ..., n-1, α_n = 1, 2, ... , 9.

(ακέραιος αριθμός (n+1)- ψηφίων)

(ii) $\left(\alpha_n \cdot \alpha_{n-1} \cdots \alpha_2\alpha_1\alpha_0 \cdot \alpha_{-1}\alpha_{-2} \cdots \right)_{10} = \alpha_n \cdot 10^n + \cdots + \alpha_0 \cdot 10^0 + \alpha_{-1} \cdot 10^{-1} + \alpha_{-2} \cdot 10^{-2} + \cdots$

Το ακέραιο μέρος του αριθμού είναι η τιμή του πολυωνύμου

$$p(x) = \alpha_n \cdot x^n + \alpha_{n-1} \cdot x^{n-1} + \cdots + \alpha_1 \cdot x + \alpha_0$$

για x = 10, ενώ το κλασματικό του μέρος η τιμή της δυναμοσειράς $\sum_{j=1}^{\infty} \alpha_{-j} \cdot x^j$

για x = $1/10$ (πεπερασμένη ή με άπειρο πλήθος όρων).

1.2. Κάθε αριθμός στο *δυαδικό σύστημα* μπορεί να γραφεί στη μορφή

$\left(b_n \cdot b_{n-1} \cdots b_2b_1b_0 \cdot b_{-1}b_{-2} \cdots \right)_2 = b_n \cdot 2^n + b_{n-1} \cdot 2^{n-1} + \cdots + b_0 \cdot 2^0 + b_{-1} \cdot 2^{-1} + b_{-2} \cdot 2^{-2} + \dots$,

b_i =0, 1, i=…, -2, -1, 0, 1, 2, …, n-1, b_n =1.

1.3. Κάθε μη μηδενικός πραγματικός αριθμός γραμμένος σ' ένα αριθμητικό σύστημα με βάση το β μπορεί να γραφεί στην κανονική μορφή *κινητής υποδιαστολής* ως

$$x = \pm\left(\cdot\, \alpha_1\alpha_2 \cdots\right) \cdot \beta^e \; , \quad \alpha_1 \neq 0 \; ,$$

όπου α_i ψηφία του αριθμητικού συστήματος με βάση το β και e κατάλληλος ακέραιος.

π.χ. . $-\left(\cdot\, 00698\right)_{10} = -\cdot 698 \text{ x } 10^{-2}$, $\left(101.001\right)_2 = \cdot 101001 \text{ x } 2^3$.

1.4. Οι αριθμοί μηχανής $x \neq 0$ ενός υπολογιστή είναι ένα σύνολο ρητών αριθμών, γραμμένων σε μορφή κινητής υποδιαστολής της μορφής

$$x = \pm \cdot \alpha_1\alpha_2 \cdots \alpha_t \cdot \beta^e \qquad (1)$$

όπου

(i) το *κλάσμα* $\cdot\, \alpha_1\alpha_2 \cdots \alpha_t$ αποτελείται από t ψηφία ως προς τη βάση β με $\alpha_1 \neq 0$,

(ii) ο *εκθέτης* e είναι ακέραιος με $L \leq e \leq U$ (L, U ακέραιοι και L ; - U).

Το σύνολο Μ των αριθμών μηχανής με τις παραπάνω ιδιότητες της αριθμητικής μονάδας ενός υπολογιστή είναι οι ρητοί αριθμοί (1) και ο 0.

1.5. Καλούμε *αριθμητική ακρίβεια* ενός αριθμού που χρησιμοποιείται για την προσέγγιση μιας πραγματικής τιμής μέσω μετρήσεων ή υπολογισμών το πλήθος των ψηφίων που θα χρησιμοποιήσουμε για την αναπαράσταση του αριθμού.

Καλούμε *ακρίβεια* το μέγεθος που περιγράφει το πόσο κοντά βρίσκεται η τιμή που έχουμε χρησιμοποιήσει για την προσέγγιση ενός μεγέθους στην πραγματική του τιμή.

1.6. Καλούμε *ακρίβεια* t ενός αριθμού μηχανής το πλήθος των ψηφίων (ως προς τη βάση του αριθμητικού συστήματος που χρησιμοποιούμε) του κλάσματος του συγκεκριμένου αριθμού.

1.7. Κατά την προσέγγιση των πραγματικών αριθμών με αριθμούς μηχανής διακρίνουμε τις ακόλουθες περιπτώσεις:

(i) *«υπερχείλιση»* (overflow) $\rightarrow$ στην προσπάθεια αναπαράστασης ενός αριθμού απολύτως μεγαλύτερο από το μέγιστο στοιχείο του Μ, με αποτέλεσμα να σταματήσει ο υπολογισμός,

(ii) *«υπεκχείλιση»* (underflow) $\rightarrow$ στην προσπάθεια αναπαράστασης π.χ. ενός αριθμού x με $0 < x < \cdot 1 \cdot \beta^{L}$. Στην περίπτωση αυτή ο αριθμός x αντικαθίσταται από το 0 και οι πράξεις συνεχίζονται.

1.8. Συνήθως κάθε αριθμός x με $\cdot 1 \cdot \beta^{L} < |x| < \beta^{U}$ προσεγγίζεται για αναπαράσταση στον υπολογιστή και για εκτέλεση πράξεων από έναν κοντινό του αριθμό μηχανής fl(x) (floating point), δηλαδή έναν αριθμό μηχανής (στοιχείο του συνόλου Μ) έτσι, ώστε $|x-fl(x)| \leq |x-y|$, για κάθε $y \in M$ (για $x = 0 \rightarrow fl(x) = x$). Στην περίπτωση αυτή το *σχετικό σφάλμα* αυτής της προσέγγισης είναι

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

Η επιλογή του fl(x) ως τον πλησιέστερο αριθμό μηχανής προς τον x καλείται

(i) *στρογγύλευση* και πραγματοποιείται μέσω στρογγύλευσης του t-στού ψηφίου του x προς τα πάνω ή προς τα κάτω:

π.χ. για $\beta = 10$, $t = 5$, $x = (.\ \alpha_1 \ldots \alpha_5\ \alpha_6\ \alpha_7 \ldots) \,.\, 10^{k}$ έχουμε ότι:

- αν $\alpha_6 > 5$, τότε $fl(x) = (.\ \alpha_1\alpha_2 \ldots \alpha_5 + 10^{-5}) \,.\, 10^{k}$,
- αν $\alpha_6 < 5$, τότε $fl(x) = (.\ \alpha_1\alpha_2 \ldots \alpha_5) \,.\, 10^{k}$,
- αν $\alpha_6 = 5$, τότε είτε αποφασίζουμε με βάση το α_7 είτε επιλέγουμε από τους παραπάνω fl(x) εκείνον που ικανοποιεί μία συγκεκριμένη ιδιότητα (π.χ. έχει άρτιο 5^{o} δ.ψ.).

(ii) *αποκοπή* όταν αποκόπτουμε όλα τα ψηφία μετά το ψηφίο α_t

(δηλαδή για $x = (.\ \alpha_1\alpha_2 \ldots \alpha_t\ \alpha_{t+1} \ldots) \,.\, \beta^{k} \rightarrow fl(x) = (.\ \alpha_1\alpha_2 \ldots \alpha_t) \,.\, \beta^{k}$). Τότε το σχετικό σφάλμα αυτής της προσέγγισης είναι

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

Συνοψίζοντας ισχύει ότι:

κάθε μη μηδενικός πραγματικός αριθμός x με $\cdot 1 \cdot \beta^L < |x| < \beta^U$ μπορεί να παρασταθεί από τον αριθμό μηχανής f |(x) με σχετικό σφάλμα προσέγγισης

$$\frac{|fl(x)-x|}{|x|} \le \begin{cases} \frac{1}{2}\beta^{1-t} & , \text{ στρογγύλευση} \\ \beta^{1-t} & , \text{ αποκοπή} \end{cases}$$

1.9. Κατά την προσέγγιση μιας πραγματικής τιμής x από μια προσεγγιστική τιμή x_e διακρίνουμε τα ακόλουθα είδη σφαλμάτων

(i) το *απόλυτο σφάλμα* $E_{\alpha\pi} = \|x - x_e\|$,

(ii) το *σχετικό σφάλμα* $E_{\sigma\chi} = \dfrac{\|x - x_e\|}{\|x\|}$.

Στην περίπτωση που τα x και x_e δεν είναι αριθμοί αλλά διανύσματα ή πίνακες (μητρώα) η απόλυτη τιμή στον ορισμό των σφαλμάτων αντικαθίσταται από τις ακόλουθες *νόρμες* (ή *στάθμες)* :

- **για διανύσματα** $x=(x_1\ x_2, \ldots, x_n)^T$

(i) την *Ευκλείδεια νόρμα* $\|x\|_E = \left[\sum_{i=1}^{n} |x_i|^2\right]^{1/2}$,

(ii) τις νόρμες $\|\ell\|_p$ με $p \in [0, +\infty)$ $\|x\|_p = \left[\sum_{i=1}^{n} |x_i|^p\right]^{1/p}$.

Συνήθως χρησιμοποιούμε τις νόρμες με

- **p=1** $\rightarrow$ $\|x\|_1 = \sum_{i=1}^{n} |x_i|$,
- **p=2** $\rightarrow$ $\|x\|_2 = \|x\|_E = \left[\sum_{i=1}^{n} |x_i|^2\right]^{1/2}$,
- **p=∞** $\rightarrow$ $\|x\|_\infty = \max_i |x_i|$.

- **για πίνακες** $A = (\alpha_{ij})_{i,j}$, i, j=1,2,…,n.

(i) $\|A\|_1 = \max_j \sum_{i=1}^{n} |\alpha_{ij}|$,

(ii) $\|A\|_\infty = \max_i \sum_{j=1}^{n} |\alpha_{ij}|$.

1.10 Θα λέμε ότι η προσεγγιστική τιμή x_e πλησιάζει την πραγματική τιμή x κατά t *σημαντικά ψηφία* όταν το t είναι ο μεγαλύτερος ακέραιος αριθμός για τον οποίο ισχύει

$$\frac{|x - x_e|}{|x|} \le \frac{1}{2} \cdot 10^{1-t} .$$

(στην περίπτωση αριθμητικής συστήματος με βάση το β = 10).

Παρατήρηση. Αν το σχετικό σφάλμα προσέγγισης $\frac{|x - x_e|}{|x|}$ είναι μεγαλύτερο του 0.5, τότε δεν υπάρχει κανένα σημαντικό ψηφίο που να προσεγγίζει τους δύο αριθμούς.

1.11. Καλούμε **αλγόριθμο** μία σειρά εντολών που υλοποιούν μία αριθμητική μέθοδο. Τόσο η *ευστάθεια* όσο και η *ακρίβεια* (πόσο μικρό είναι το αντίστοιχο θεωρητικό σφάλμα προσέγγισης) είναι τα βασικά χαρακτηριστικά που καθορίζουν την ποιότητά του. Άλλα επιμέρους χαρακτηριστικά ενός αλγορίθμου είναι η *βεβαιότητα* (ότι το αποτέλεσμα σε κάθε βήμα εκτέλεσης του αλγορίθμου θα είναι εντός της επιθυμητής ακρίβειας), το *κόστος,* ο *χώρος* αποθήκευσης του τελικού αποτελέσματος και των ενδιάμεσων αποτελεσμάτων καθώς και η *αξιοπιστία.*

Εντούτοις σε αρκετές περιπτώσεις οι αλγόριθμοι είναι ιδιαίτερα *ευαίσθητοι* σε σφάλματα στρογγύλευσης, τα οποία μπορούν να επηρεάσουν το τελικό αποτέλεσμα. Θα λέμε ότι ένας αλγόριθμος είναι (αριθμητικά) *ευσταθής* (αντίθετα *ασταθής*) αν τα τελικά αποτελέσματά του δεν επηρεάζονται πολύ από τις μικρές διαταραχές που τα σφάλματα στρογγύλευσης δημιουργούν σε κάθε βήμα του.