Αριθμητική Ανάλυση & Εφαρμογές

Διδάσκων: Δημήτριος Ι. Φωτιάδης

Τμήμα Μηχανικών Επιστήμης Υλικών

Ιωάννινα 2017-2018

Το περιεχόμενο βασίζεται στο βιβλίο: Μ.Ν. Βραχάτης, Αριθμητική Ανάλυση, Κλειδάριθμος, 2012.

Υπολογισμοί και Σφάλματα

Παράσταση Πραγματικών Αριθμών

- Συστήματα Αριθμών
- Παράσταση Ακέραιου Μέρους ενός Πραγματικού Αριθμού
- Παράσταση Δεκαδικού Μέρους ενός Πραγματικού Αριθμού
- Παράσταση Κινητής Υποδιαστολής

Συστήματα Αριθμών

Κάθε πραγματικός αριθμός x μπορεί να παρασταθεί μονοσήμαντα σε ένα σύστημα αριθμών με **βάση** $b \geq 2$ με τη μορφή:

$$x = \pm \sum_{i=n}^{-\infty} d_i b^i$$

όπου οι συντελεστές d_i της σειράς αυτής είναι στοιχεία από το σύνολο των ψηφίων

$$D = \{0, 1, 2, \ldots, b-1\}$$

και ονομάζονται **ψηφία** του αριθμού x. Συμβολικά ο αριθμός x γράφεται ως εξής

$$x = \begin{cases} d_n d_{n-1} \cdots d_0 . d_{-1} d_{-2} \cdots & \text{αν} \quad n \geq 0, \quad d_n \neq 0, \\ \\ .000 \cdots 0 d_n d_{n-1} d_{n-2} \cdots & \text{αν} \quad n < 0. \end{cases}$$

Το περιεχόμενο βασίζεται στο βιβλίο: Μ.Ν. Βραχάτης, Αριθμητική Ανάλυση, Κλειδάριθμος, 2012.

Συστήματα Αριθμών

- Ανάλογα με την τιμή της βάσης b σε ένα σύστημα αριθμών έχουμε αντίστοιχα την ονομασία του συστήματος.
  - $b=2$ **δυαδικό** σύστημα αριθμών
  - $b=10$ **δεκαδικό** σύστημα αριθμών
- Κάθε αριθμός έχει **πεπερασμένη παράσταση** σε ένα σύστημα αριθμών, αν υπάρχει ακέραιος k με $k \le n$:

$$d_i = 0, \quad \text{για όλα τα} \quad i = k, k-1, k-2, \dots$$

- **Σημαντικά ψηφία** ενός πραγματικού αριθμού ονομάζονται όλα τα ψηφία του αριθμού εκτός των μηδενικών ψηφίων που βρίσκονται στην αρχή του αριθμού.
  - Το πρώτο διάφορο του μηδενός ψηφίο αποτελεί το **πρώτο σημαντικό ψηφίο** του αριθμού.
  - π.χ. ο αριθμός του δεκαδικού συστήματος 0.0007374 έχει 4 σημαντικά ψηφία και πρώτο σημαντικό ψηφίο το 7.

Συστήματα Αριθμών

Κανόνες στα συστήματα αριθμών

- Η βάση κάθε συστήματος αριθμών είναι κατά ένα μεγαλύτερη του μεγαλύτερου ψηφίου του συστήματος.
- Πολλαπλασιασμός ή διαίρεση με τη βάση γίνεται με μια μετάθεση της υποδιαστολής δεξιά ή αριστερά αντίστοιχα.
- Αν η βάση του συστήματος αριθμών είναι μεγαλύτερη από το δέκα, τότε χρησιμοποιούνται τα γράμματα A, B, C, D, E, F,… για να παραστήσουν αντίστοιχα τα ψηφία: 10, 11, 12, 13, 14, 15, 16,…
  - π.χ. το δεκαεξαδικό σύστημα αριθμών έχει τα ψηφία 0, 1, 2, …, 9 , A, B, C, D, E, και F.

Συστήματα Αριθμών

Κανόνες στα συστήματα αριθμών

- Για να διαχωρίσουμε τους αριθμούς μεταξύ διαφόρων συστημάτων αριθμών τοποθετούμε στο τέλος του αριθμού
  - ως υποδείκτη τη βάση του συστήματος (π.χ. 101.11_2, 101.11_{10}, 101.11_8, 101.11_{16})ή
  - το κεφαλαίο γράμμα της βάσης του συστήματος αριθμών που ανήκει (π.χ. 101.11B, 101.11D, 101.11O ή 101.11Q, 101.11H).
  - Αν κάποιος αριθμός δε δηλώνεται, αυτό σημαίνει ότι ανήκει στο δεκαδικό σύστημα αριθμών.
- Το πλήθος των σημαντικών ψηφίων ενός ακέραιου αριθμού που απαιτούνται για να παρασταθεί σε ένα σύστημα αριθμών με βάση b_1 είναι μεγαλύτερο ή ίσο από το πλήθος των σημαντικών ψηφίων του ίδιου αριθμού που απαιτούνται για να παρασταθεί σε ένα σύστημα αριθμών με βάση b_2 όταν $b_1<b_2$.

Συστήματα Αριθμών

Παραδείγματα

1. Ο αριθμός 101.1_2 μπορεί να παρασταθεί μονοσήμαντα στο δεκαδικό σύστημα αριθμών ως

$$101.1_2 = 1\times 2^2 + 0\times 2^1 + 1\times 2^0 + 1\times 2^{-1}$$
$$= 4 + 0 + 1 + 0.5 = 5.5_{10}$$

2. Ο αριθμός $F2B_{16}$ μπορεί να παρασταθεί μονοσήμαντα στο δεκαδικό σύστημα αριθμών ως

$$F2B_{16} = F\times 16^2 + 2\times 16^1 + B\times 16^0$$
$$= 15\times 16^2 + 2\times 16^1 + 11\times 16^0$$
$$= 3840 + 32 + 11 = 3883_{10}$$

Παράσταση Ακέραιου Μέρους ενός Πραγματικού Αριθμού

Επαναληπτική διαδικασία για τον υπολογισμό των ψηφίων $d_i, i=0(1)n$ ενός αριθμού στο σύστημα αριθμών με βάση b, ο οποίος μονοσήμαντα παριστάνει το ακέραιο μέρος x ενός αριθμού που ανήκει στο δεκαδικό σύστημα αριθμών:

$$d_i = y_i \bmod b, \quad i = 0, \ldots, n$$

όπου

$$y_i = \begin{cases} x & \text{αν} \quad i = 0, \\ \dfrac{y_{i-1} - d_{i-1}}{b} & \text{αν} \quad i = 1, \ldots, n. \end{cases}$$

Παράσταση Ακέραιου Μέρους ενός Πραγματικού Αριθμού

Αλγόριθμος

Βήμα 1: Είσοδος $I = \{x, b\}$.

Βήμα 2: Θέσε $i = 0$ και $y_0 = x$.

Βήμα 3: Αν ισχύει $i \neq 0$ υπολόγισε το $y_i = (y_{i-1} - d_{i-1})/b$.

Βήμα 4: Αν ισχύει $y_i < b$ τότε θέσε $d_i = y_i$ και πήγαινε στο Βήμα 6, διαφορετικά υπολόγισε το $d_i = y_i \bmod b$.

Βήμα 5: Αντικατάστησε το i με το $i+1$ και πήγαινε στο Βήμα 4.

Βήμα 6: Έξοδος $O = \{d_0, d_1, \ldots\}$.

Παράσταση Ακέραιου Μέρους ενός Πραγματικού Αριθμού

- Όπως έχουμε αναφέρει ένα αριθμητικό πρόβλημα θεωρείται λυμένο με την παράθεση ενός αλγορίθμου που εφαρμοζόμενος δίνει τη λύση του προβλήματος.
- Έτσι , υλοποιώντας τον αντίστοιχο αλγόριθμο σε μία γλώσσα προγραμματισμού μπορούμε να πάρουμε τα αριθμητικά αποτελέσματα με τη χρήση ενός υπολογιστή.

Παράσταση Ακέραιου Μέρους ενός Πραγματικού Αριθμού

Παράδειγμα

$$I=\{x,b\}=\{111_{10},2\}$$
$$O=\{d_0,d_1,\ldots,d_6\}=\{1,1,1,1,0,1,1\}$$

i	y_i		b		y_{i+1}		d_i
0	111	:	2	=	55	+	1
1	55	:	2	=	27	+	1
2	27	:	2	=	13	+	1
3	13	:	2	=	6	+	1
4	6	:	2	=	3	+	0
5	3	:	2	=	1	+	1
6	1	:	2	=	0	+	1

Παράσταση Δεκαδικού Μέρους ενός Πραγματικού Αριθμού

Επαναληπτική διαδικασία για τον υπολογισμό των ψηφίων $d_i, i=-1,-2,\ldots,-k$ ενός αριθμού στο σύστημα αριθμών με βάση b, ο οποίος μονοσήμαντα παριστάνει το δεκαδικό μέρος ενός αριθμού x που ανήκει στο δεκαδικό σύστημα αριθμών:

$$d_{-i} = \lfloor y_{-i} \times b \rfloor, \qquad i = 1, 2, \ldots, k$$

όπου

$$y_{-i} = \begin{cases} x & \text{αν } i = 1, \\ (y_{-i+1} \times b) - d_{-i+1} & \text{αν } i = 2, \ldots, k. \end{cases}$$

Παράσταση Δεκαδικού Μέρους ενός Πραγματικού Αριθμού

Αλγόριθμος

Βήμα 1: Είσοδος $I=\{x,b,k\}$.

Βήμα 2: Θέσε $i=1$ και $y_{-1}=x$.

Βήμα 3: Αν ισχύει $i\neq 1$ υπολόγισε το $y_{-i}=(y_{-i+1}\times b)-d_{-i+1}$.

Βήμα 4: Αν ισχύει $y_{-i}=0$ τότε πήγαινε στο Βήμα 6, διαφορετικά υπολόγισε το $d_{-i}=\lfloor y_{-i}\times b\rfloor$.

Βήμα 5: Αν ισχύει $i\leq k$ αντικατάστησε το i με το $i+1$ και πήγαινε στο Βήμα 3.

Βήμα 6: Έξοδος $O=\{d_{-1},d_{-2},\ldots\}$.

Το περιεχόμενο βασίζεται στο βιβλίο: Μ.Ν. Βραχάτης, Αριθμητική Ανάλυση, Κλειδάριθμος, 2012.

Παράσταση Δεκαδικού Μέρους ενός Πραγματικού Αριθμού

Παράδειγμα

$$I = \{x, b, k\} = \{0.59375_{10}, 2, 50\}$$

$$O = \{d_{-1}, d_{-2}, \ldots\} = \{1, 0, 0, 1, 1\}$$

i	y_{-i}		b		y_{-i-1}		d_{-i}
1	0.59375	x	2	=	0.1875	+	1
2	0.1875	x	2	=	0.375	+	0
3	0.375	x	2	=	0.75	+	0
4	0.75	x	2	=	0.5	+	1
5	0.5	x	2	=	0	+	1

Παράσταση Δεκαδικού Μέρους ενός Πραγματικού Αριθμού

Υπάρχει περίπτωση ένας αριθμός που εκφράζεται με πεπερασμένα στοιχεία σε ένα σύστημα αριθμών μίας συγκεκριμένης βάσης να μην έχει πεπερασμένη παράσταση σε ένα σύστημα αριθμών διαφορετικής βάσης π.χ. $I=\{0.1_{10}, 2, 50\}$

i	y_{-i}		b		y_{-i-1}		d_{-i}
1	0.1	x	2	=	0.2	+	0
2	0.2	x	2	=	0.4	+	0
3	0.4	x	2	=	0.8	+	0
4	0.8	x	2	=	0.6	+	1
5	0.6	x	2	=	0.2	+	1
6	0.2	x	2	=	0.4	+	0
7	0.4	x	2	=	0.8	+	0
8	0.8	x	2	=	0.6	+	1
9	0.6	x	2	=	0.2	+	1

Παράσταση Κινητής Υποδιαστολής

- Δεν μπορούμε να εκφράσουμε με πεπερασμένη παράσταση όλους τους αριθμούς.
- Σε αυτήν την περίπτωση μπορούμε να πάρουμε μόνο προσεγγίσεις, οι οποίες γίνονται ακριβέστερες όσο αυξάνει ο χώρος αποθήκευσης.
- π.χ. ο αριθμός 0.000110011_2 του δυαδικού συστήματος αριθμών αντιστοιχεί στον 0.0996093_{10} και προσεγγίζει τον 0.1_{10}.
- Συχνά και κυρίως στον επιστημονικό υπολογισμό απαιτείται μεγάλη ακρίβεια στην παράσταση των αριθμών, γεγονός το οποίο σημαίνει μεγάλο αποθηκευτικό χώρο σε μία μηχανή επεξεργασίας αριθμών.
- **Τι γίνεται στην περίπτωση που θέλουμε να αποθηκεύσουμε τον αριθμό $\mathbf{0.0000000019_{10}}$ όταν ο αποθηκευτικός μας χώρος μπορεί να αποθηκεύσει μόνο οκτώ δεκαδικά ψηφία;**

Παράσταση Κινητής Υποδιαστολής

- Ένας αριθμός x σε ένα σύστημα αριθμών με βάση b μπορεί να γραφεί ως εξής:

$$x = \pm r \times b^{\pm E}$$

- Όπου E ένας μη αρνητικός ακέραιος και r ένας αριθμός στο σύστημα αριθμών με βάση b.
- Αυτή η παράσταση ονομάζεται **παράσταση κινητής υποδιαστολής** για τον αριθμό x σε ένα σύστημα αριθμών με βάση b, με εκθέτη E και ουρά ή κλάσμα r.
- Η παράσταση κινητής υποδιαστολής δεν είναι μοναδική εφόσον μπορούμε να γράψουμε $x = \pm(rb) \times b^{\pm E-1}$.

Παράσταση Κινητής Υποδιαστολής

- Για να διατηρήσουμε την παράσταση κινητής υποδιαστολής μοναδική ως προς την παράστασή της, έχει επικρατήσει να χρησιμοποιείται η κανονικοποιημένη παράσταση κινητής υποδιαστολής.

- Η **κανονικοποιημένη παράσταση κινητής υποδιαστολής** είναι η παράσταση εκείνη σύμφωνα με την οποία ένας αριθμός x σε σύστημα αριθμών με βάση το b γράφεται ως εξής:

$$x = \pm r \times b^{\pm E}, \quad \frac{1}{b} \le r < 1$$

Απόλυτο και Σχετικό Σφάλμα

- Υποθέτουμε ότι x είναι η αληθής (ακριβή) τιμή ενός μεγέθους ή μιας ποσότητας και ότι x^* μια προσεγγιστική τιμή που προσεγγίστηκε με έναν οποιοδήποτε τρόπο, τότε η παρακάτω διαφορά:

$$\varepsilon = x^* - x$$

  ονομάζεται **σφάλμα** (error), ενώ η ποσότητα η οποία είναι αντίθετη προς το σφάλμα:

$$r = -\varepsilon = x - x^*$$

  ονομάζεται **διόρθωση** (correction).

- Η απόλυτη τιμή του σφάλματος ε $|\varepsilon| = |x^* - x|$ ονομάζεται **απόλυτο σφάλμα**.

Το περιεχόμενο βασίζεται στο βιβλίο: Μ.Ν. Βραχάτης, Αριθμητική Ανάλυση, Κλειδάριθμος, 2012.

Απόλυτο και Σχετικό Σφάλμα

- Η ποσότητα που εκφράζει το λόγο του σφάλματος ε προς την ακριβή τιμή $x \neq 0$, δηλαδή η ποσότητα
$$\delta = \frac{\varepsilon}{x} = \frac{x^* - x}{x}$$
ονομάζεται **σχετικό σφάλμα**.
- Το σχετικό σφάλμα κατά προσέγγιση είναι ίσο με:
$$\delta \approx \frac{\varepsilon}{x^*} = \frac{x^* - x}{x^*}.$$
- Η ποσότητα $|\delta|$ εκφράζει το **απόλυτο σχετικό σφάλμα**.

Απόλυτο και Σχετικό Σφάλμα

- Το σχετικό σφάλμα είναι ανεξάρτητο από τη μονάδα της προσεγγιστικής μέτρησης x^*, σε αντίθεση με το απόλυτο σφάλμα το οποίο εξαρτάται από τη χρησιμοποιούμενη μονάδα μέτρησης.
- Όταν χρησιμοποιούμε το σχετικό σφάλμα, λαμβάνουμε υπόψη και το μέγεθος της ποσότητας που μετράμε, σε αντίθεση με το απόλυτο σφάλμα στο οποίο το μέγεθος δε συμμετέχει.

Παράδειγμα

$$x = 100, \quad x^* = 101, \quad |\varepsilon| = |x^* - x| = |101 - 100| = 1, \quad \delta = \frac{\varepsilon}{x} = \frac{1}{100} = 0.01$$

$$x = 10000, \quad x^* = 9999, \quad |\varepsilon| = |x^* - x| = |9999 - 10000| = 1, \quad \delta = \frac{\varepsilon}{x} = \frac{1}{10000} = 0.0001$$

Πηγές και Είδη Σφαλμάτων

- Η λύση ενός προβλήματος με τη βοήθεια αριθμητικών μεθόδων διαφέρει πάντοτε από την ακριβή λύση λόγω της παρουσίας σφαλμάτων.
  - Αρχικά σφάλματα
  - Σφάλμα του μαθηματικού προβλήματος ή σφάλμα της μαθηματικής περιγραφής
- **Προβλήματα κακής κατάστασης**
  - Η λύση των αριθμητικών προβλημάτων είναι πολύ ευαίσθητη σε μικρές μεταβολές των δεδομένων του προβλήματος και
  - μικρές μεταβολές των δεδομένων του προβλήματος δημιουργούν μεγάλη μεταβολή στα αποτελέσματα.

Πηγές και Είδη Σφαλμάτων

Παράδειγμα προβλήματος κακής κατάστασης

- Η λύση του συστήματος γραμμικών εξισώσεων

$$x + 2y = 3$$
$$0.5x + 1.001y = 1.5$$

είναι $\left(x, y\right) = \left(3, 0\right)$.

- Αν ο συντελεστής 0.5 του x στη 2η εξίσωση διαταραχθεί με μία ποσότητα -0.001 και γίνει 0.499, τότε η λύση του προβλήματος είναι: $\left(x, y\right) = \left(1, 1\right)$.

Πηγές και Είδη Σφαλμάτων

- Το **σφάλμα αποκοπής** είναι σφάλμα που δημιουργείται από το χρησιμοποιούμενο αλγόριθμο και προσέγγιση με την υπόθεση ότι όλες οι αριθμητικές πράξεις είναι ακριβείς.
- **Παράδειγμα:** Υπολογισμός μίας σειράς απείρων όρων χρησιμοποιώντας πεπερασμένο πλήθος όρων.

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \cdots$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \frac{x^{11}}{11!} + \ldots$$

- ❖ Ανάλογα με την επιθυμητή προσέγγιση χρησιμοποιείται ένα ορισμένο πλήθος όρων της σειράς.
- ❖ Οι άπειροι στο πλήθος όροι που αποκόπτονται εισάγουν το σφάλμα αποκοπής.

Πηγές και Είδη Σφαλμάτων

Παράδειγμα

Υπολογισμός του e^x χρησιμοποιώντας τους k όρους της σειράς

- Πραγματική τιμή $x = e^x$
- Προσεγγιστική τιμή x^*
- Σφάλμα ε

$$\varepsilon = x^* - x = \sum_{i=0}^{k} \frac{x^i}{i!} - \sum_{i=0}^{\infty} \frac{x^i}{i!} = -\sum_{i=k+1}^{\infty} \frac{x^i}{i!}$$

Πηγές και Είδη Σφαλμάτων

- **Σφάλμα στρογγυλοποίησης:** Κάθε αριθμός που χρησιμοποιείται σε έναν υπολογισμό ως δεδομένο ή ως αποτέλεσμα μιας πράξης πρέπει να προσεγγιστεί με έναν αριθμό που έχει περιορισμένο πλήθος ψηφίων.

$$\pi = 3.141592353589793 2\cdots$$

$$\frac{4}{3} = 1.3333333333333333\ldots$$

- Ο μεγαλύτερος επιτρεπτός αριθμός ψηφίων εξαρτάται από το μέσο επεξεργασίας αριθμών.

- Κατά τη διαδικασία της στρογγυλοποίησης το ζητούμενο είναι η ελαχιστοποίηση του σφάλματος στρογγυλοποίησης.

Πηγές και Είδη Σφαλμάτων

Διαδικασία στρογγυλοποίησης ενός αριθμού σε *k* δεκαδικά ψηφία

- Παραλείπονται όλα τα δεκαδικά ψηφία που υπάρχουν μετά την k δεκαδική θέση
- Αν το πρώτο ψηφίο που αποκόπτεται είναι μεγαλύτερο του 5, το τελευταίο ψηφίο που παραμένει αυξάνεται κατά μία μονάδα.
- Αν το πρώτο ψηφίο που αποκόπτεται είναι μικρότερο του 5, το τελευταίο ψηφίο που παραμένει δεν αλλάζει.
- Διαφορετικά το τελευταίο ψηφίο που παραμένει:
  - δεν αλλάζει, αν αυτό είναι άρτιο.
  - αυξάνεται κατά μία μονάδα, αν αυτό είναι περιττό.

Πηγές και Είδη Σφαλμάτων

Αν x είναι η ακριβής τιμή που στρογγυλοποιείται και x^* είναι η προσεγγιστική τιμή μετά την στρογγυλοποίηση της x σε k δεκαδικά ψηφία, τότε για το απόλυτο σφάλμα θα ισχύει ότι:

$$|\varepsilon| \leq \frac{1}{2}10^{-k}$$

Πηγές και Είδη Σφαλμάτων

Παράδειγμα: Στρογγυλοποίηση του αριθμού $\sqrt{7} \approx 2.6457513$ σε $k = 6(-1)1$ δεκαδικά ψηφία και υπολογισμός των αντίστοιχων απόλυτων σχετικών σφαλμάτων.

k	x^*	$\lvert\varepsilon\rvert$	0.5×10^{-k}
6	2.645751	0.0000003	0.0000005
5	2.64575	0.0000013	0.000005
4	2.6458	0.0000486	0.00005
3	2.646	0.0002486	0.0005
2	2.65	0.0042486	0.005
1	2.6	0.0457513	0.05

Πηγές και Είδη Σφαλμάτων

- Η ανάγκη της χρήσης αριθμών με πεπερασμένο πλήθος ψηφίων έχει ως συνέπεια να μην ισχύουν πάντα οι ιδιότητες των πραγματικών αριθμών.
  - π.χ. αντιμεταθετική, προσεταιριστική, επιμεριστική
- Η σειρά εκτέλεσης των πράξεων για την επίλυση ενός αριθμητικού προβλήματος έχει μεγάλη σημασία.

Πηγές και Είδη Σφαλμάτων

Παράδειγμα:

Υπολογισμός του αθροίσματος $x = \sum_{i=1}^{40} a_i,\ a_i = 0.5$ σε μέσο επεξεργασίας αριθμών που αποθηκεύει 2 μόνο ψηφία.

Εφαρμόζοντας τα γνωστό τρόπο πρόσθεσης αριθμών

$$x = \left(\cdots\left(\left(\left(0.5 + 0.5\right) + 0.5\right) + 0.5\right) + \cdots + 0.5\right)$$

έχουμε διαδοχικά $x = 1,\ x = 1.5,\ x = 2, \ldots, x = 10,\ x = 10, \ldots, x = 10.$

Λύση: Υπολογισμός δύο μερικών αθροισμάτων που το καθένα είναι ίσο με 10 και πρόσθεση αυτών.

Πηγές και Είδη Σφαλμάτων

- Ένα **σημαντικό ψηφίο** ενός προσεγγιστικού αριθμού x^* είναι **ακριβές**, εάν το απόλυτο σφάλμα δεν υπερβαίνει τη μισή μονάδα της τάξεως που αντιστοιχεί σε αυτό το ψηφίο.
  - Αν k σημαντικά ψηφία του x^* είναι ακριβή, τότε θα λέμε ότι ο αριθμός είναι **ακριβής σε k σημαντικά ψηφία**.
  - Εάν όλα τα σημαντικά ψηφία του x^* είναι ακριβή, τότε ο x^* είναι **ακριβής στο δοθέντα αριθμό ψηφίων**.
- **Δύο αριθμοί συμφωνούν σε k δεκαδικά ψηφία**, αν η απόλυτη τιμή της διαφοράς τους είναι μικρότερη ή ίση από 0.5×10^{-k}.
- **Δύο αριθμοί συμφωνούν** σε k **σημαντικά ψηφία**, αν μετά τη στρογγυλοποίησή τους σε k σημαντικά ψηφία οι αριθμοί ταυτίζονται.

Μετάδοση Σφαλμάτων κατά τους Υπολογισμούς

Θεώρημα 1: Το απόλυτο σφάλμα του αθροίσματος δύο αριθμών είναι μικρότερο ή ίσο με το άθροισμα των απολύτων σφαλμάτων των αριθμών αυτών.

Απόδειξη:

$$\begin{aligned}\varepsilon &= \left(x_1^* + x_2^*\right) - \left(x_1 + x_2\right) \\ &= \left(x_1^* - x_1\right) + \left(x_2^* - x_2\right) \\ &= \varepsilon_1 + \varepsilon_2\end{aligned}$$

Οπότε $\left|\varepsilon\right| = \left|\varepsilon_1 + \varepsilon_2\right| \le \left|\varepsilon_1\right| + \left|\varepsilon_2\right|$

Μετάδοση Σφαλμάτων κατά τους Υπολογισμούς

Θεώρημα 2: Το απόλυτο σφάλμα της διαφοράς δύο αριθμών είναι μικρότερο ή ίσο με το άθροισμα των απολύτων σφαλμάτων των αριθμών αυτών.

Απόδειξη:

$$\begin{aligned}\varepsilon &= \left(x_1^* - x_2^*\right) - \left(x_1 - x_2\right) \\ &= \left(x_1^* - x_1\right) - \left(x_2^* - x_2\right) \\ &= \varepsilon_1 - \varepsilon_2\end{aligned}$$

Οπότε $\left|\varepsilon\right| = \left|\varepsilon_1 - \varepsilon_2\right| \le \left|\varepsilon_1\right| + \left|\varepsilon_2\right|$

Μετάδοση Σφαλμάτων κατά τους Υπολογισμούς

- Πρέπει να αποφεύγεται, εκεί που είναι δυνατόν, η αφαίρεση δύο περίπου ίσων προσεγγιστικών αριθμών, διότι αυτή η αφαίρεση οδηγεί στη μείωση της ακρίβειας του αποτελέσματος.

- **Καταστροφική ακύρωση σημαντικών ψηφίων:** Σχετίζεται με την απώλεια σωστών σημαντικών ψηφίων μικρών αριθμών, οι οποίοι απορρέουν από πράξεις μεταξύ μεγάλων αριθμών.

Μετάδοση Σφαλμάτων κατά τους Υπολογισμούς

Παράδειγμα 1: Υπολογισμός του $e^{-\alpha}$ με χρήση της σειράς Taylor όταν $\alpha > 0$. Στην περίπτωση αυτή ο υπολογισμός του $e^{-\alpha}$ πραγματοποιείται με αθροίσεις των παρακάτω όρων:

$$e^{-a} = 1 - a + \frac{a^2}{2!} - \frac{a^3}{3!} + \frac{a^4}{4!} + ...$$

- Όταν το α είναι σχετικά μεγάλο, τότε το σφάλμα στρογγυλοποίησης μπορεί να είναι μεγαλύτερο από την πραγματική τιμή του $e^{-\alpha}$.
- Για $\alpha = -5.5$ αν αθροίσουμε τους 25 πρώτους χρησιμοποιώντας 5 σημαντικά ψηφία, τότε η άθροιση θα μας δώσει:

$$e^{-5.5} = 0.0026363,$$

ενώ το σωστό αποτέλεσμα είναι:

$$e^{-5.5} = 0.0040868,$$

δηλαδή δεν έχουμε κανένα σημαντικό ψηφίο σωστό.

Μετάδοση Σφαλμάτων κατά τους Υπολογισμούς

- Το φαινόμενο αυτό μπορεί να αντιμετωπιστεί υπολογίζοντας τους μεγάλους σε μέγεθος όρους με μεγαλύτερη ακρίβεια σημαντικών ψηφίων, έτσι ώστε να είναι δυνατόν να συνεισφέρουν στην απάντηση.
  - Κοστίζει σε μνήμη καθώς και σε υπολογιστικό χρόνο.
- Εναλλακτικά, αντί του $e^{-\alpha}$ μπορεί να υπολογιστεί το e^{α} και μετά να αντιστραφεί το τελικό εξαγόμενο.
  - Το αποτέλεσμα της άθροισης δεν είναι μικρός αριθμός και έτσι δεν υφίσταται το φαινόμενο της καταστροφικής ακύρωσης.

Μετάδοση Σφαλμάτων κατά τους Υπολογισμούς

Παράδειγμα 2: Υπολογισμός της διαφοράς $\sqrt{x+1}-\sqrt{x}$ για $x=707,$ χρησιμοποιώντας τέσσερα σημαντικά ψηφία.

- Τότε θα έχουμε:

$$\sqrt{708}-\sqrt{707}=26.61-26.59=0.02.$$

- Αφού οι αριθμοί είναι στρογγυλοποιημένοι σε τέσσερα σημαντικά ψηφία, τότε τα απόλυτα σφάλματα των αριθμών αυτών είναι μικρότερα ή ίσα με 0.5×10^{-2} και επομένως το απόλυτο σφάλμα της διαφοράς τους είναι μικρότερο ή ίσο με $10^{-2}=0.01>0.5\times10^{-2}$.
- Έτσι η τιμή 0.02 δεν είναι αρκετά ακριβής.

Μετάδοση Σφαλμάτων κατά τους Υπολογισμούς

Εναλλακτική μέθοδος:

$$\sqrt{708}-\sqrt{707}=\left(\sqrt{708}-\sqrt{707}\right)\left(\frac{\sqrt{708}+\sqrt{707}}{\sqrt{708}+\sqrt{707}}\right)=$$

$$=\frac{1}{\sqrt{708}+\sqrt{707}}=\frac{1}{26.61+26.59}=0.01880$$

Προφανώς, η παραπάνω διαδικασία δίνει πιο ακριβείς τιμές και ο λόγος είναι ότι αποφεύχθηκε η αφαίρεση δύο περίπου ίσων αριθμών.

Μετάδοση Σφαλμάτων κατά τους Υπολογισμούς

Θεώρημα 3: Το απόλυτο σχετικό σφάλμα του γινομένου δύο αριθμών είναι κατά προσέγγιση μικρότερο ή ίσο από το άθροισμα των απολύτων σχετικών σφαλμάτων των αριθμών αυτών.

Θεώρημα 4: Το απόλυτο σχετικό σφάλμα του πηλίκου δύο αριθμών είναι κατά προσέγγιση μικρότερο ή ίσο από το άθροισμα των απολύτων σχετικών σφαλμάτων των αριθμών αυτών.

Μετάδοση Σφαλμάτων κατά τους Υπολογισμούς

Παράδειγμα

Έστω ότι οι αριθμοί x_1 και x_2 δίνονται στρογγυλοποιημένοι σε 2 δεκαδικά ψηφία. Θα δείξουμε ότι το απόλυτο σφάλμα της έκφρασης $3.1x_1 + 2.9x_2$ είναι μικρότερο ή ίσο από 0.03, με την υπόθεση ότι οι συντελεστές 3.1 και 3.2 είναι ακριβείς.

Λύση

$$\varepsilon = x^* - x = \left(3.1x_1^* + 2.9x_2^*\right) - \left(3.1x_1 + 2.9x_2\right)$$

$$= 3.1\left(x_1^* - x_1\right) - 2.9\left(x_2^* - x_2\right) = 3.1\varepsilon_1 + 2.9\varepsilon_2$$

Επειδή οι αριθμοί δίνονται στρογγυλοποιημένοι σε 2 δεκαδικά ψηφία , τότε θα ισχύει:

$$\left|\varepsilon_1\right| \le \frac{1}{2}10^{-2}, \quad \left|\varepsilon_2\right| \le \frac{1}{2}10^{-2}$$

$$\left|\varepsilon\right| = \left|3.1\varepsilon_1 + 2.9\varepsilon_2\right| \le \left|3.1\varepsilon_1\right| + \left|2.9\varepsilon_2\right| \le \frac{3.1}{2}\times 10^{-2} + \frac{2.9}{2}\times 10^{-2} = 0.03$$

Ολικό Σφάλμα

Έστω ότι το προς εξέταση αριθμητικό πρόβλημα είναι η εύρεση μιας τιμής της συνάρτησης *f(x)* για κάποια τιμή της ανεξάρτητης μεταβλητής *x.*

1. Πολλές φορές στην πράξη αντί της συνάρτησης *f(x)* χρησιμοποιείται μια απλούστερη και πιο εύχρηστη συνάρτηση *g(x).*
   – σφάλμα αποκοπής $\varepsilon_\alpha = g\left(x\right) - f\left(x\right)$
2. Αντί όμως της ακριβούς τιμής της ανεξάρτητης μεταβλητής x χρησιμοποιείται μια προσεγγιστική τιμή x*
   – σφάλμα διάδοσης $\varepsilon_\delta = g\left(x^*\right) - g\left(x\right)$
3. Όμως μετά του αριθμητικούς υπολογισμούς και λόγω των διαφορετικών προσεγγιστικών πράξεων που γίνονται κατά τη διάρκεια αυτών, η τελική τιμή είναι μια προσέγγιση *g*(x*)* της πραγματικής τιμή *g(x*).*
   – παραχθέν σφάλμα $\varepsilon_\pi = g^*\left(x^*\right) - g\left(x^*\right)$

Το περιεχόμενο βασίζεται στο βιβλίο: Μ.Ν. Βραχάτης, Αριθμητική Ανάλυση, Κλειδάριθμος, 2012.

Ολικό Σφάλμα

Με βάση τα παραπάνω για το ολικό σφάλμα ε_o στην τελική τιμή *g*(x*)* θα ισχύουν διαδοχικά τα παρακάτω:

$$\varepsilon_o = g^*\left(x^*\right) - f\left(x\right) = \left[g^*\left(x^*\right) - g\left(x^*\right)\right] + \left[g\left(x^*\right) - g\left(x\right)\right] + \left[g\left(x\right) - f\left(x\right)\right]$$

και τελικά θα έχουμε:

$$\varepsilon_o = \varepsilon_\pi + \varepsilon_\delta + \varepsilon_\alpha$$

Τέλος, αν λάβουμε υπόψη το υπολογιστικό σφάλμα ε_υ, που δημιουργείται από τη διαφορά της προσεγγιστικής τιμής *g*(x*)* από την πραγματική τιμή *g(x),* τότε για το υπολογιστικό σφάλμα θα ισχύει ότι:

$$\varepsilon_\upsilon = g^*\left(x^*\right) - g\left(x\right) = \left[g^*\left(x^*\right) - g\left(x^*\right)\right] + \left[g\left(x^*\right) - g\left(x\right)\right]$$

Ολικό Σφάλμα

Επομένως:

$$\varepsilon_{\upsilon} = \varepsilon_{\pi} + \varepsilon_{\delta}$$

Με βάση το παραπάνω υπολογιστικό σφάλμα το ολικό ε_o στην τελική τιμή *g*(x*)* θα είναι:

$$\varepsilon_{o} = \varepsilon_{\upsilon} + \varepsilon_{\alpha}$$

Έτσι το ολικό σφάλμα δίνεται από το άθροισμα του υπολογιστικού σφάλματος και του σφάλματος αποκοπής.