

ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ  
ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

**Ηλεκτρονικοί Υπολογιστές II :**  
**Εισαγωγή στην Αριθμητική Ανάλυση**

Σημειώσεις Διαλέξεων και Εργαστηρίων

Μ. ΓΡΑΜΜΑΤΙΚΑΚΗΣ

Γ. ΚΟΠΙΔΑΚΗΣ – Ν. ΠΑΠΑΔΑΚΗΣ – Σ. ΣΤΑΜΑΤΙΑΔΗΣ

Ηράκλειο  
Ιανουάριος 2008

Copyright © 2005 Σ. Σταματιάδης, (stamatis@materials.uoc.gr), Μ. Γραμματικάκης, Γ. Κοπιδάκης, Ν. Παπαδάκης

Η στοιχειοθεσία έγινε από τον Σ. Σταματιάδη με τη χρήση του L<sup>A</sup>T<sub>E</sub>X2<sub>ε</sub>.  
Χρησιμοποιήθηκε η σειρά χαρακτήρων 'Κέρκης' (© Τμήμα Μαθηματικών, Πανεπιστήμιο Αιγαίου).

Τελευταία τροποποίηση του κειμένου έγινε την 11 Ιανουαρίου 2008. Η πιο πρόσφατη έκδοση βρίσκεται στο <http://www.edu.physics.uoc.gr/~tety213>

# Περιεχόμενα

|                                                                                       |           |
|---------------------------------------------------------------------------------------|-----------|
| <b>1 Σφάλματα</b>                                                                     | <b>1</b>  |
| 1.1 Εισαγωγή . . . . .                                                                | 1         |
| 1.2 Ασκήσεις . . . . .                                                                | 3         |
| <b>2 Αριθμητική Επίλυση μη Γραμμικών Εξισώσεων</b>                                    | <b>5</b>  |
| 2.1 Εισαγωγή . . . . .                                                                | 5         |
| 2.1.1 Χαρακτηρισμοί μεθόδων εύρεσης ρίζας . . . . .                                   | 6         |
| 2.2 Μέθοδος Διχοτόμησης . . . . .                                                     | 6         |
| 2.2.1 Θεώρημα Αλγορίθμου Διχοτόμησης (Φράγμα σφάλματος, Ταχύτητα σύγκλισης) . . . . . | 7         |
| 2.3 Μέθοδος Σταθερού Σημείου $x = g(x)$ . . . . .                                     | 9         |
| 2.3.1 Σταθερά σημεία-Σχετικά Θεωρήματα . . . . .                                      | 9         |
| 2.4 Μέθοδος Newton-Raphson . . . . .                                                  | 11        |
| 2.5 Μέθοδος τεμνουσας . . . . .                                                       | 13        |
| 2.6 Ασκήσεις . . . . .                                                                | 14        |
| <b>3 Επίλυση Γραμμικών Συστημάτων</b>                                                 | <b>17</b> |
| 3.1 Εισαγωγή . . . . .                                                                | 17        |
| 3.2 Μέθοδος Cramer . . . . .                                                          | 17        |
| 3.3 Από ευθείας απαλοιφή Gauss . . . . .                                              | 18        |
| 3.4 Ευστάθεια γραμμικών συστημάτων . . . . .                                          | 21        |
| 3.5 Εφαρμογές . . . . .                                                               | 22        |
| 3.5.1 Υπολογισμός Οριζουσας . . . . .                                                 | 22        |
| 3.5.2 Μέθοδος Gauss-Jordan για τον υπολογισμό του αντίστροφου πίνακα . . . . .        | 22        |
| 3.5.3 Εύρεση ιδιοτιμών και ιδιοδιανυσμάτων . . . . .                                  | 23        |
| 3.6 Ασκήσεις . . . . .                                                                | 24        |
| <b>4 Παρεμβολή και Προσέγγιση Συναρτήσεων</b>                                         | <b>27</b> |
| 4.1 Παρεμβολή με πολυώνυμο Lagrange . . . . .                                         | 27        |
| 4.2 Προσέγγιση με τη μέθοδο ελαχίστων τετραγώνων . . . . .                            | 28        |
| 4.2.1 Πολυωνυμική προσέγγιση . . . . .                                                | 30        |
| 4.3 Ασκήσεις . . . . .                                                                | 31        |
| <b>5 Αριθμητική Ολοκλήρωση</b>                                                        | <b>33</b> |
| 5.1 Εισαγωγή . . . . .                                                                | 33        |
| 5.2 Κανόνας Τραπεζίου . . . . .                                                       | 33        |
| 5.3 Κανόνας Simpson . . . . .                                                         | 35        |

|          |                                                     |           |
|----------|-----------------------------------------------------|-----------|
| 5.3.1    | Κανόνας Simpson των $3/8$                           | 37        |
| 5.3.2    | Κανόνας Simpson των $3/8$ για εκτεταμένο διάστημα   | 38        |
| 5.4      | Ειδικές Περιπτώσεις                                 | 38        |
| 5.4.1    | Ολοκλήρωση σε άνισα τμήματα                         | 38        |
| 5.4.2    | Ολοκληρώματα με μη πεπερασμένα όρια ολοκλήρωσης     | 38        |
| 5.5      | Μέθοδος Gauss                                       | 39        |
| 5.6      | Ασκήσεις                                            | 41        |
| <b>6</b> | <b>Διαφορικές Εξισώσεις</b>                         | <b>45</b> |
| 6.1      | Γενικά                                              | 45        |
| 6.2      | Εισαγωγή                                            | 45        |
| 6.2.1    | Διωνυμικό Ανάπτυγμα                                 | 46        |
| 6.3      | Κατηγορίες και Λύσεις Διαφορικών Εξισώσεων          | 47        |
| 6.3.1    | Πρωτοβάθμιες ΔΕ                                     | 47        |
| 6.3.2    | Δευτεροβάθμιες ΔΕ                                   | 47        |
| 6.3.3    | Σύστημα πρωτοβάθμιων ΔΕ με σταθερούς συντελεστές    | 48        |
| 6.4      | Μέθοδος Σειράς Taylor                               | 49        |
| 6.4.1    | Μέθοδος Euler                                       | 51        |
| 6.4.2    | Σφάλμα Μεθόδου Taylor                               | 52        |
| 6.5      | Μέθοδος Runge–Kutta                                 | 53        |
| 6.5.1    | Μέθοδος Runge–Kutta 2 <sup>ου</sup> βαθμού          | 54        |
| 6.5.2    | Μέθοδος Runge–Kutta 4 <sup>ου</sup> βαθμού          | 55        |
| 6.5.3    | Σχόλια                                              | 56        |
| 6.6      | Τελεστές Διαφορών                                   | 56        |
| 6.6.1    | Ιδιότητες                                           | 57        |
| 6.6.2    | Άλλοι τελεστές                                      | 58        |
| 6.6.3    | Γενικευμένοι τύποι του Newton                       | 60        |
| 6.6.4    | Εφαρμογή των τελεστών στον υπολογισμό ολοκληρωμάτων | 60        |
| 6.7      | Πολυβηματικές Μέθοδοι                               | 61        |
| 6.7.1    | Μέθοδος Adams–Bashforth                             | 62        |
| 6.7.2    | Μέθοδος Adams–Moulton                               | 63        |
| 6.7.3    | Μέθοδοι Πρόβλεψης–Διόρθωσης (Predictor–Corrector)   | 64        |
| 6.8      | Συστήματα Διαφορικών Εξισώσεων                      | 66        |
| 6.9      | Εξισώσεις Διαφορών                                  | 68        |
| 6.9.1    | Εξίσωση διαφορών πρώτου βαθμού                      | 69        |
| 6.9.2    | Εξίσωση διαφορών δεύτερου βαθμού                    | 69        |
| 6.9.3    | Μη ομογενείς εξισώσεις διαφορών                     | 71        |
| 6.9.4    | Σχόλια                                              | 71        |
| 6.10     | Αριθμητική Ευστάθεια                                | 72        |
| 6.11     | Απόλυτη Ευστάθεια                                   | 75        |
| 6.12     | Ασκήσεις                                            | 76        |

# Κεφάλαιο 1

## Σφάλματα

### 1.1 Εισαγωγή

Η αναπαράσταση πραγματικών ή ακεραίων αριθμών δεν είναι πάντα δυνατή με ακρίβεια λόγω της πεπερασμένης μνήμης του Η/Υ. Για παράδειγμα, ο αριθμός  $y \in \mathcal{R}$  αναπαρίσταται ως (βάση δεκαδικών αριθμών)

$$\tilde{y} = \pm 0.d_1 d_2 \dots d_K \times 10^{\pm s},$$

όπου

$$\begin{aligned} 1 \leq d_1 \leq 9, \\ 0 \leq d_i \leq 9, \quad i \in \{2, 3, \dots, K\} \quad \text{και} \\ 0 \leq s \leq M, \end{aligned}$$

με  $K, M$  σταθερές εξαρτώμενες από τον εκάστοτε Η/Υ.

Έτσι έχουμε, αν

$$K = 6, M = 10 \Rightarrow \tilde{\pi} = 0.314159 \times 10^1,$$

ενώ αν

$$K = 5, M = 10 \Rightarrow \tilde{\pi} = 0.31416 \times 10^1,$$

Το *σφάλμα στρογγύλισης* (*round-off error*) ορίζεται ως  $|y - \tilde{y}|$ . Ο αριθμός  $K$  αποτελεί το πλήθος των *σημαντικών ψηφίων* (*significant digits*).

#### Παρατηρήσεις:

1. Αν ο αριθμός  $y$  (ή και αποτέλεσμα ενδιάμεσης πράξης) υπερβαίνει κατ' απόλυτη τιμή το μέγιστο αναπαραστάσιμο στον Η/Υ αριθμό, έχουμε *υπερχείλιση* (*overflow*). Αντίστοιχα, αν είναι κατ' απόλυτη τιμή μικρότερος από το μικρότερο αναπαραστάσιμο στον Η/Υ αριθμό, τότε έχουμε *υπεκχείλιση* (*underflow*). Η τιμή που θα αποκτήσει αυτός και στις δύο περιπτώσεις είναι απροσδιόριστη, ο υπολογισμός όμως μπορεί να συνεχίσει με, σχεδόν σίγουρα, λάθος αποτέλεσμα. Σε υπολογιστές που υλοποιούν το πρότυπο αναπαράστασης αριθμών IEEE οι τιμές είναι αντίστοιχα  $\pm\text{infinity}$  (το πλησιέστερο "άπειρο") και  $\pm 0$ .<sup>1</sup>

---

<sup>1</sup>στο πρότυπο υπάρχει διάκριση μεταξύ των  $+0$  (από τη κατεύθυνση των θετικών αριθμών) και  $-0$  (από τους αρνητικούς).

2. Ο τρόπος αναπαράστασης που περιγράφηκε μπορεί να αποθηκεύσει *ακριβώς* ένα πεπερασμένο πλήθος πραγματικών αριθμών. Οι υπόλοιποι *προσεγγίζονται* με έναν από αυτούς, είτε με αποκοπή είτε με στρογγύλευση, ανάλογα με τον υπολογιστή.

**Παράδειγμα:** Έστω ότι θέλουμε να υπολογίσουμε το άθροισμα των αριθμών  $x = 5891.26$  και  $y = .0773414$  σε υπολογιστή με  $K = 5$  στο μοντέλο αναπαράστασής του. Έστω ακόμα ότι αυτή γίνεται με στρογγύλευση. Οι αριθμοί  $x, y$  επομένως αποθηκεύονται ως

$$\tilde{x} = 0.58913 \times 10^4, \quad \tilde{y} = 0.77341 \times 10^{-1}.$$

Πριν την εκτέλεση της πράξης οι αριθμοί τροποποιούνται ώστε να έχουν τον ίδιο εκθέτη στην αναπαράσταση:

$$\tilde{x} = 0.58913 \times 10^4, \quad \tilde{y} = 0.00001 \times 10^4.$$

Επομένως, το άθροισμα στον υπολογιστή των αριθμών  $x, y$  είναι  $0.58914 \times 10^4 = 5891.4$  ενώ η αλγεβρική πρόσθεσή τους δίνει ως αποτέλεσμα το  $5891.3373414$ , το οποίο στρογγυλεύόμενο σε  $K$  ψηφία είναι  $5891.3$ .

3. Συνέπεια της πεπερασμένης αναπαράστασης είναι ακόμα το ότι το αποτέλεσμα σύνθετων εκφράσεων δεν ακολουθεί απαραίτητα τους κανόνες της άλγεβρας. Π.χ. αν  $x = 5891.26$ ,  $y = .0773414$ ,  $z = -0.07$  και  $K = 5$  έχουμε

$$\tilde{x} = +0.58913 \times 10^4, \quad \tilde{y} = +0.77341 \times 10^{-1}, \quad \tilde{z} = -0.70000 \times 10^{-1}.$$

Η πράξη  $x + y + z$  στον υπολογιστή έχει διαφορετικό αποτέλεσμα αν εκτελεστεί ως  $(x + y) + z$  από αυτό που προκύπτει αν εκτελεστεί ως  $x + (y + z)$  (υπολογίστε τα!).

4. Προσέξτε ότι στο μοντέλο που περιγράψαμε ισχύει  $1 + x = 1$  για κάθε  $x$  με  $|x| < 5 \times 10^{-K}$ . Αυτό σημαίνει ότι υπάρχει ένα όριο κάτω από το οποίο οι αριθμοί συμπεριφέρονται σαν το μηδέν σε προσθέσεις ή αφαιρέσεις με αριθμούς της τάξης του 1. Το όριο αυτό ονομάζεται *έψιλον της μηχανής*: παρατηρήστε ότι είναι πολύ μεγαλύτερο από τον μικρότερο αριθμό που μπορεί να αναπαρασταθεί.

**Παράδειγμα:** Έστω Η/Υ με αναπαράσταση πραγματικών αριθμών με βάση το 2, 4 bit mantissa, 4 bit εκθέτη και με bits προσήμου. Τότε έχουμε:

$$\begin{array}{ll} \text{Μέγιστος θετικός αριθμός} & (+111)_2 \times 2^{(+111)_2} = +7 \times 2^7 \\ \text{Μέγιστος αρνητικός αριθμός} & (-111)_2 \times 2^{(+111)_2} = -7 \times 2^7 \\ \text{Ελάχιστος θετικός αριθμός} & (+001)_2 \times 2^{(-111)_2} = +1 \times 2^{-7} \\ \text{Ελάχιστος αρνητικός αριθμός} & (-001)_2 \times 2^{(-111)_2} = -1 \times 2^{-7} \end{array}$$

Η απόπειρα αναπαράστασης ενός αριθμού  $x$  έξω από τα παραπάνω όρια δίνει

$$\begin{array}{ll} \text{υπερχείλιση αν} & x > 7 \times 2^7 \text{ ή } x < -7 \times 2^7, \\ \text{υπεκχείλιση αν} & -2^{-7} < x < 2^{-7} \text{ και } x \neq 0. \end{array}$$

**Παρατηρήσεις:**

1. Για τον εκθέτη μπορεί να μη χρησιμοποιηθεί bit προσήμου αλλά bias. Έτσι  $-(2^{N_e-1} - 2) \leq x \leq (2^{N_e-1} - 1)$  όπου  $N_e$  ο αριθμός ψηφίων του εκθέτη.
2. Για την mantissa με  $N_m$  bits έχουμε  $N_m$  bits ακρίβειας. Συνήθως ο αριθμός είναι κανονικοποιημένος ως  $1.f_1f_2f_3 \dots$  ή, σπανιότερα,  $0.f_1f_2f_3 \dots$  όπου  $f_1, f_2, f_3, \dots$  τα ψηφία του δυαδικού αριθμού.

**1.2 Ασκήσεις**

1. Υπολογίστε το έψιλον της μηχανής για πραγματικούς αριθμούς απλής και διπλής ακρίβειας με τους εξής τρόπους:

(α') Εφαρμόστε τον αλγόριθμο:

Θέτουμε  $\varepsilon \leftarrow 1$ . Για όσο ισχύει  $1 + \varepsilon \neq 1$  θέτουμε  $\varepsilon \leftarrow \varepsilon/2$  και επαναλαμβάνουμε.

(β') Καλέστε τις ρουτίνες SLAMCH() και DLAMCH() της συλλογής ρουτινών LAPACK.

(γ') Καλέστε την εσωτερική συνάρτηση EPSILON() της FORTRAN 90.

2. Οι ρίζες του τριωνύμου  $ax^2 + bx + c$  δίνονται ως

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a},$$

όταν  $a \neq 0$ .

Έστω  $a = 1$ ,  $b = 3000.001$ ,  $c = 3$ .

(α') Υπολογίστε τα  $x_{1,2}$  με απλή και διπλή ακρίβεια. Συγκρίνετέ τα με τις ακριβείς ρίζες ( $x_1 = -0.001$ ,  $x_2 = -3000.0$ ).

(β') Επαναλάβετε τους υπολογισμούς του προηγούμενου σκέλους εφαρμόζοντας τον αλγεβρικά ισοδύναμο τύπο

$$x_{1,2} = \frac{2c}{-b \mp \sqrt{b^2 - 4ac}}.$$

Τι παρατηρείτε ως προς την ακρίβεια των υπολογισμών σας;

3. Γράψτε κώδικα ώστε να υπολογίσετε την τιμή του  $e^1$  εφαρμόζοντας τη σχέση

$$e^x = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n.$$

Βρείτε, δηλαδή, την προσδιοριζόμενη τιμή για  $n = 1, 2, 3, \dots$ . Τι παρατηρείτε ως προς την ταχύτητα σύγκλισης στην πραγματική τιμή του (2.718281828459045...);

4. Γράψτε κώδικα σε υπορουτίνα της FORTRAN που να υπολογίζει το  $e^x$  εφαρμόζοντας τη σχέση

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Για τη διευκόλυνσή σας παρατηρήστε ότι ο κάθε όρος στο άθροισμα προκύπτει από τον αμέσως προηγούμενο αν αυτός πολλαπλασιαστεί με το  $x/n$ .

Δοκιμάστε τον κώδικά σας για θετικά και αρνητικά  $x$ . Τι παρατηρείτε;

5. Γράψτε κώδικα σε υπορουτίνα της FORTRAN που να υπολογίζει το  $\sin x$  εφαρμόζοντας τη σχέση

$$\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} .$$

Για τη διευκόλυνσή σας παρατηρήστε ότι ο κάθε όρος στο άθροισμα προκύπτει από τον αμέσως προηγούμενο αν αυτός πολλαπλασιαστεί με το  $-\frac{x^2}{2k(2k+1)}$ .

Δοκιμάστε τον κώδικά σας για θετικά και αρνητικά  $x$ . Τι παρατηρείτε ;

6. Γράψτε κώδικα σε υπορουτίνα της FORTRAN που να υπολογίζει το  $\cos x$  εφαρμόζοντας τη σχέση

$$\cos x = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!} .$$

Δοκιμάστε τον κώδικά σας για θετικά και αρνητικά  $x$ . Τι παρατηρείτε ;

## Κεφάλαιο 2

# Αριθμητική Επίλυση μη Γραμμικών Εξισώσεων

### 2.1 Εισαγωγή

Στο κεφάλαιο αυτό θα παρουσιάσουμε κάποιους αλγορίθμους (μεθόδους) εύρεσης των λύσεων μιας εξίσωσης με ένα άγνωστο. Η εξίσωση έχει γενικά τη μορφή

$$f(x) = 0, \quad x \in \mathcal{R}. \quad (2.1)$$

Οι λύσεις της, τα συγκεκριμένα σημεία  $x$  που την ικανοποιούν, λέγονται και *ρίζες* της συνάρτησης  $f(x)$ .

Στην περίπτωση που η συνάρτηση  $f(x)$  είναι γραμμική (δηλαδή, της μορφής  $f(x) = ax + b$ ) η εύρεση της ρίζας είναι τετριμμένη. Οι δυσκολίες εμφανίζονται στην αντίθετη περίπτωση και γι' αυτό θα επικεντρωθούμε στην επίλυση μη γραμμικών εξισώσεων. Όταν η  $f(x)$  είναι γενικό πολυώνυμο μέχρι και 4<sup>ου</sup> βαθμού, υπάρχουν αναλυτικοί τύποι που δίνουν τις ρίζες της. Ήδη, όμως, από τον 3<sup>ο</sup> βαθμό είναι αρκετά δύσχρηστοι. Στη γενική περίπτωση που δεν είναι πολυώνυμο, η εύρεση των ριζών (ή και η απόδειξη της ύπαρξής τους) γενικά δεν είναι δυνατή με αναλυτικούς τύπους.

Η επίλυση με αριθμητικές μεθόδους της εξίσωσης (2.1) βασίζεται στην εύρεση μιας ακολουθίας τιμών  $x_0, x_1, \dots, x_k$ , που συγκλίνουν για  $k \rightarrow \infty$  σε μία ρίζα της εξίσωσης. Μερικά χρήσιμα θεωρήματα είναι:

**Θεώρημα Ενδιαμέσου Τιμής** Έστω  $f(x)$  συνεχής συνάρτηση στο κλειστό διάστημα  $[a, b]$ . Αν  $\lambda$  είναι ένας οποιοσδήποτε πραγματικός αριθμός μεταξύ των  $f(a), f(b)$  (συμπεριλαμβανομένων και αυτών), τότε υπάρχει ένα τουλάχιστον  $c \in [a, b]$  ώστε  $f(c) = \lambda$ .

**Θεώρημα Μέσης Τιμής** Έστω  $f(x)$  συνεχής συνάρτηση για  $x \in [a, b]$ , διαφορίσιμη στο  $(a, b)$ , με παράγωγο  $f'(x)$ . Τότε υπάρχει ένα τουλάχιστον  $c \in [a, b]$  ώστε  $f(b) - f(a) = f'(c)(b - a)$ . Αν επιπλέον ισχύει  $f(a) = f(b)$  τότε σε κάποιο  $c \in [a, b]$  έχουμε  $f'(c) = 0$  (**Θεώρημα Rolle**).

**Θεώρημα Taylor** Έστω ότι η συνάρτηση  $f(x)$ ,  $x \in [a, b]$ , έχει παράγωγο τάξης  $n+1$  και η  $f^{n+1}(x)$  είναι συνεχής στο  $[a, b]$ . Αν  $x, x_0 \in [a, b]$ ,  $x \neq x_0$ , τότε υπάρχει  $\xi = \xi(x)$ ,  $\xi \in (x_0, x)$  ώστε

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots + \frac{f^n(x_0)}{n!}(x - x_0)^n + R_n(x),$$

όπου το υπόλοιπο

$$R_n(x) = \frac{f^{n+1}(\xi(x))}{(n+1)!} (x - x_0)^{n+1} .$$

### 2.1.1 Χαρακτηρισμοί μεθόδων εύρεσης ρίζας

#### Ταχύτητα σύγκλισης

Μια μέθοδος επίλυσης της εξίσωσης  $f(x) = 0$ , παράγει την ακολουθία προσεγγιστικών λύσεων  $x_0, x_1, \dots$  η οποία συγκλίνει στην ρίζα  $\xi$ . Η μέθοδος χαρακτηρίζεται ως  $\alpha$  τάξης όσον αφορά στη σύγκλιση, αν υπάρχουν  $\alpha, \lambda > 0$  ώστε

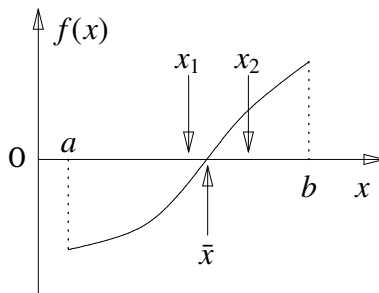
$$\lim_{n \rightarrow \infty} \frac{|x_{n+1} - \xi|}{|x_n - \xi|^\alpha} = \lambda .$$

Ο αριθμός  $\lambda$  αποτελεί την *ταχύτητα* (ή *ρυθμό*) *σύγκλισης*.

#### Ευστάθεια

Όπως θα δούμε, οι περισσότερες μέθοδοι εύρεσης ρίζας χρειάζονται μια αρχική προσέγγιση της λύσης (ή και περισσότερες), την οποία βελτιώνουν σε κάθε στάδιο της επίλυσης. Η αριθμητική τους ευστάθεια προσδιορίζεται από τη συμπεριφορά τους σε μεταβολές αυτής της αρχικής τιμής. Μια μέθοδος είναι *ευσταθής* αν οποιαδήποτε κατάλληλα μικρή μεταβολή της αρχικής τιμής δεν επηρεάζει την εύρεση της ρίζας, ενώ είναι *ασταθής* αν μια μικρή μεταβολή της αρχικής προσέγγισης οδηγεί μακριά από τη ρίζα.

## 2.2 Μέθοδος Διχοτόμησης



Σχήμα 2.1: Σχηματική αναπαράσταση της Μεθόδου Διχοτόμησης για την εύρεση ρίζας

Η μέθοδος βασίζεται στο Θεώρημα Ενδιάμεσης Τιμής (ΘΕΤ). Αν  $f(x)$  συνεχής στο  $[a, b]$  και έχουμε  $f(a)f(b) < 0$ , τότε από το θεώρημα, υπάρχει  $c = \bar{x} \in (a, b)$  ώστε  $f(\bar{x}) = 0$ . Άρα υπάρχει τουλάχιστον μία ρίζα της  $f(x)$  στο  $(a, b)$ . Το συμπέρασμα αυτό αποτελεί το *θεώρημα Weierstrass*.

Η διαδικασία που ακολουθεί η μέθοδος διχοτομεί το διάστημα  $[a, b]$ , εντοπίζει τη ρίζα σε ένα από τα δύο υποδιαστήματα και επαναλαμβάνεται στο επιλεγμένο υποδιάστημα. Παράγεται έτσι μια ακολουθία διαστημάτων  $[a_1, b_1], [a_2, b_2], \dots, [a_N, b_N]$  και μια ακολουθία προσεγγίσεων της ρίζας  $x_1 = (a_1 + b_1)/2, x_2 = (a_2 + b_2)/2, \dots, x_N = (a_N + b_N)/2$ . Αν  $\varepsilon$  είναι σταθερά που δηλώνει το αποδεκτό σφάλμα μπορούμε να θέσουμε ως κριτήριο τερματισμού ένα ή περισσότερα από τα

- $|b_N - a_N| < \varepsilon$

- $|x_N - x_{N-1}| < \varepsilon$
- $\left| \frac{x_N - x_{N-1}}{x_N} \right| < \varepsilon$  αν  $x_N \neq 0$
- $|f(x_N)| < \varepsilon$
- $N \geq N_0$  για να αποφύγουμε άπειρες επαναλήψεις όταν η  $f(x)$  δεν ικανοποιεί τα κριτήρια του ΘΕΤ και ο αλγόριθμος δε συγκλίνει. Σε αυτήν την περίπτωση πρέπει να ελέγξουμε αν το αποτέλεσμα είναι αποδεκτό.

**Αλγόριθμος:** Επίλυση της  $f(x) = 0$  με τη μέθοδο διχοτόμησης:

1. Επιλέγουμε δύο τιμές  $a, b$  έτσι ώστε η  $f(x)$  να είναι συνεχής στο  $[a, b]$  και να ισχύει  $f(a)f(b) < 0$ .
2. Θέτουμε  $x \leftarrow \frac{a+b}{2}$ .
3. Αν το  $x$  είναι ικανοποιητική προσέγγιση της ρίζας πηγαίνουμε στο βήμα 6.
4. Αν ισχύει ότι  $f(a)f(x) < 0$  τότε θέτουμε  $b \leftarrow x$ . Αλλιώς, θέτουμε  $a \leftarrow x$ .
5. Επαναλαμβάνουμε τη διαδικασία από το βήμα 2.
6. Τέλος.

**Παράδειγμα:** Έστω η συνάρτηση  $f(x) = x^3 + 4x^2 - 10$ , η οποία είναι συνεχής στο  $[1, 2]$ . Καθώς  $f(1) = -5$  και  $f(2) = 14$ , έχουμε  $f(1)f(2) < 0$  και, επομένως, υπάρχει μία τουλάχιστον ρίζα της στο  $[1, 2]$ . Παρατηρούμε ακόμα ότι  $f'(x) = 3x^2 + 8x > 0$  για κάθε  $x$  στο συγκεκριμένο διάστημα. Επομένως, η  $f(x)$  είναι αύξουσα σε αυτό και άρα έχει μοναδική ρίζα στο  $[1, 2]$ . Εφαρμόζουμε τη μέθοδο διχοτόμησης για την εύρεσή της και προκύπτουν οι ακολουθίες του Πίνακα 2.1. Μετά από 20 επαναλήψεις το σχετικό σφάλμα είναι  $|x_{20} - \bar{x}| \leq 0.5 |b_{20} - a_{20}| \approx 0.95 \times 10^{-6}$ , άρα έχουμε προσδιορίσει τουλάχιστον 6 σωστά ψηφία της ρίζας. Η προσεγγιστική τιμή είναι 1.36523 ενώ η ακριβής είναι 1.36523001361638...

### 2.2.1 Θεώρημα Αλγορίθμου Διχοτόμησης (Φράγμα σφάλματος, Ταχύτητα σύγκλισης)

Αν η  $f(x)$  είναι συνεχής στο  $[a, b]$  και  $f(a)f(b) < 0$ , τότε η μέθοδος διχοτόμησης παράγει μια ακολουθία  $x_1, x_2, \dots$  με την ιδιότητα  $|x_n - \bar{x}| \leq \frac{1}{2^n}(b-a)$ ,  $n \geq 1$ , όπου  $\bar{x}$  η ρίζα της  $f(x)$  στο  $[a, b]$ . Η σύγκλιση είναι πρώτης τάξης (γραμμική) και αργή, καθώς η ταχύτητά της είναι 0.5.

Απόδειξη:

$$\begin{aligned} \Theta\text{ΕΤ} \Rightarrow \quad b_1 - a_1 &= b - a, & \bar{x} &\in (a_1, b_1) \\ b_2 - a_2 &= \frac{1}{2}(b_1 - a_1) = \frac{1}{2}(b - a), & \bar{x} &\in (a_2, b_2) \\ b_3 - a_3 &= \frac{1}{2}(b_2 - a_2) = \frac{1}{2^2}(b - a), & \bar{x} &\in (a_3, b_3) \\ &\vdots & & \vdots \\ b_n - a_n &= \frac{1}{2^{n-1}}(b - a), & \bar{x} &\in (a_n, b_n) \end{aligned}$$

| $n$ | $a_n$      | $b_n$      | $x_n$      | $f(x_n)$                  |
|-----|------------|------------|------------|---------------------------|
| 1   | 1.00000000 | 2.00000000 | 1.50000000 | 2.3750                    |
| 2   | 1.00000000 | 1.50000000 | 1.25000000 | -1.7969                   |
| 3   | 1.25000000 | 1.50000000 | 1.37500000 | 0.16211                   |
| 4   | 1.25000000 | 1.37500000 | 1.31250000 | -0.84839                  |
| 5   | 1.31250000 | 1.37500000 | 1.34375000 | -0.35098                  |
| 6   | 1.34375000 | 1.37500000 | 1.35937500 | $-0.96409 \times 10^{-1}$ |
| 7   | 1.35937500 | 1.37500000 | 1.36718750 | $0.32356 \times 10^{-1}$  |
| 8   | 1.35937500 | 1.36718750 | 1.36328125 | $-0.32150 \times 10^{-1}$ |
| 9   | 1.36328125 | 1.36718750 | 1.36523438 | $0.72025 \times 10^{-4}$  |
| 10  | 1.36328125 | 1.36523438 | 1.36425781 | $-0.16047 \times 10^{-1}$ |
| 11  | 1.36425781 | 1.36523438 | 1.36474609 | $-0.79893 \times 10^{-2}$ |
| 12  | 1.36474609 | 1.36523438 | 1.36499023 | $-0.39591 \times 10^{-2}$ |
| 13  | 1.36499023 | 1.36523438 | 1.36511230 | $-0.19437 \times 10^{-2}$ |
| 14  | 1.36511230 | 1.36523438 | 1.36517334 | $-0.93585 \times 10^{-3}$ |
| 15  | 1.36517334 | 1.36523438 | 1.36520386 | $-0.43192 \times 10^{-3}$ |
| 16  | 1.36520386 | 1.36523438 | 1.36521912 | $-0.17995 \times 10^{-3}$ |
| 17  | 1.36521912 | 1.36523438 | 1.36522675 | $-0.53963 \times 10^{-4}$ |
| 18  | 1.36522675 | 1.36523438 | 1.36523056 | $0.90310 \times 10^{-5}$  |
| 19  | 1.36522675 | 1.36523056 | 1.36522865 | $-0.22466 \times 10^{-4}$ |
| 20  | 1.36522865 | 1.36523056 | 1.36522961 | $-0.67174 \times 10^{-5}$ |

Πίνακας 2.1: Ακολουθίες των διαστημάτων, της προσεγγιστικής ρίζας και της αντίστοιχης τιμής της  $f(x) = x^3 + 4x^2 - 10$  κατά την εφαρμογή της μεθόδου διχοτόμησης

Καθώς  $x_n = \frac{1}{2}(a_n + b_n)$  και είτε  $x_n \leq \bar{x} \leq b_n$  είτε  $a_n \leq \bar{x} \leq x_n$ , έχουμε:

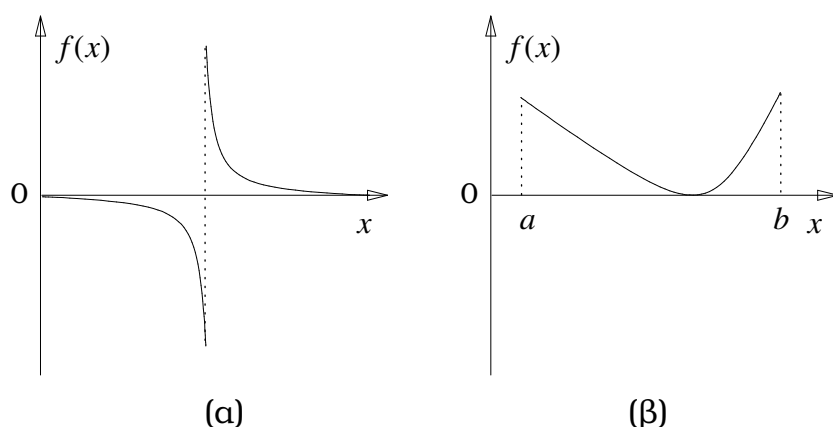
$$|\bar{x} - x_n| = \left| \bar{x} - \frac{1}{2}(a_n + b_n) \right| \leq \frac{1}{2}(b_n - a_n) = \frac{1}{2^n}(b - a).$$

Επομένως,  $\lim_{n \rightarrow \infty} x_n = \bar{x}$  καθώς  $\lim_{n \rightarrow \infty} \frac{1}{2^n}(b - a) = 0$ .

**Παράδειγμα:** Έστω η συνάρτηση  $f(x) = x^3 + 4x^2 - 10$ , συνεχής με μία ρίζα στο  $[1, 2]$ . Ποιός είναι ο αριθμός απαιτούμενων επαναλήψεων της μεθόδου διχοτόμησης ώστε  $|x_n - \bar{x}| \leq \varepsilon = 10^{-5}$ ;

Καθώς  $|x_n - \bar{x}| \leq 2^{-n}(b - a) = 2^{-n}(2 - 1) = 2^{-n}$  συμπεραίνουμε ότι απαιτείται να ισχύει  $2^{-n} \leq \varepsilon$ . Άρα πρέπει να έχουμε  $n \geq -\log_2 \varepsilon = -\frac{\log_{10} \varepsilon}{\log_{10} 2}$ . Για  $\varepsilon = 10^{-5}$  προκύπτει ότι  $n \geq \frac{5}{\log_{10} 2} \approx 16.61$  άρα αρκούν 17 επαναλήψεις για να έχουμε  $|x_n - \bar{x}| \leq 10^{-5}$ .

**Παρατήρηση:** Η μέθοδος διχοτόμησης αποτυγχάνει όταν δεν πληρούνται οι προϋποθέσεις του Θεωρήματος Ενδιάμεσης Τιμής. Π.χ. όταν η συνάρτηση δεν είναι συνεχής, Σχήμα 2.2α, η μέθοδος εντοπίζει για ρίζα το σημείο ασυνέχειας. Αντίστροφα, αν δεν μπορούμε να εντοπίσουμε δύο σημεία στα οποία η συνάρτηση έχει ετερόσημες τιμές, δε σημαίνει ότι δεν έχει ρίζα (Σχήμα 2.2β).



Σχήμα 2.2: Σχηματικές αναπαραστάσεις συναρτήσεων για τις οποίες η μέθοδος διχοτόμησης **(α)** εντοπίζει μη υπαρκτή ρίζα, **(β)** αποτυγχάνει να εντοπίσει ρίζα στο προσδιοριζόμενο διάστημα

### 2.3 Μέθοδος Σταθερού Σημείου $x = g(x)$

Το πρόβλημα εύρεσης (πραγματικής) λύσης της  $f(x) = 0$  είναι ισοδύναμο με την επίλυση της εξίσωσης  $x = g(x)$  όπου  $g(x)$  κατάλληλη συνάρτηση. Ειδικές μορφές της  $g(x)$  δίνουν ευσταθείς και γρήγορους επαναληπτικούς αλγορίθμους για την εύρεση της λύσης.

**Αλγόριθμος:** Έστω η αρχική λύση (προσέγγιση)  $x_0$ . Κατασκευάζουμε την ακολουθία  $x_0, x_1, x_2, \dots, x_n$  ως εξής:

$$x_1 = g(x_0), \quad x_2 = g(x_1), \quad x_3 = g(x_2), \quad \dots, \quad x_n = g(x_{n-1}).$$

Αν η ακολουθία συγκλίνει σε ένα σημείο  $\bar{x}$  και καθώς η  $g(x)$  είναι συνεχής<sup>1</sup> έχουμε

$$\bar{x} \equiv \lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} g(x_{n-1}) = g(\lim_{n \rightarrow \infty} x_{n-1}) \equiv g(\bar{x}).$$

Άρα

1. Θέτουμε στο  $x$  την αρχική προσέγγιση.
2. Ελέγχουμε αν ικανοποιείται το κριτήριο τερματισμού (όποιο έχουμε επιλέξει). Αν ναι, πηγαίνουμε στο βήμα 4.
3. Θέτουμε  $x \leftarrow g(x)$  και επαναλαμβάνουμε από το βήμα 2.
4. Τέλος.

#### 2.3.1 Σταθερά σημεία–Σχετικά Θεωρήματα

**Ορισμός.** Η συνάρτηση  $g(x)$  έχει *σταθερό σημείο* στο  $[a, b]$  αν υπάρχει  $\varrho \in [a, b]$  ώστε  $g(\varrho) = \varrho$ .

<sup>1</sup> $\lim g(x_n) = g(\lim x_n)$ .

**Κριτήριο ύπαρξης σταθερού σημείου.** Έστω  $g(x)$  συνεχής συνάρτηση στο  $[a, b]$ , με  $a \leq g(x) \leq b$ ,  $\forall x \in [a, b]$ . Τότε η  $g(x)$  έχει τουλάχιστον ένα σταθερό σημείο στο  $[a, b]$ .

**Απόδειξη:** Ισχύει  $g(a) \geq a$ ,  $g(b) \leq b$ . Ορίζουμε τη συνεχή συνάρτηση  $h(x) = g(x) - x$ . Τότε  $h(a) \geq 0$ ,  $h(b) \leq 0$ . Το ΘΕΤ εξασφαλίζει ότι υπάρχει  $\bar{x}$  ώστε  $h(\bar{x}) = 0$ .

**Παράδειγμα:** Έστω  $g(x) = 3^{-x}$ ,  $x \in [0, 1]$ . Έχουμε  $g(0) = 1$ ,  $g(1) = 1/3$  και  $g'(x) = -3^{-x} \ln 3 < 0 \forall x \in [0, 1]$ . Η  $g(x)$  είναι φθίνουσα και  $0 \leq g(x) \leq 1 \forall x \in [0, 1]$ . Από το κριτήριο ύπαρξης προκύπτει ότι η  $g(x)$  έχει τουλάχιστον ένα σταθερό σημείο (μοναδικό καθώς είναι φθίνουσα).

**Μοναδικότητα σταθερού σημείου.** Έστω  $g(x)$  συνεχής και διαφορίσιμη συνάρτηση στο  $[a, b]$ , με  $a \leq g(x) \leq b$  και  $|g'(x)| < 1 \forall x \in [a, b]$ . Τότε η  $g(x)$  έχει μοναδικό σταθερό σημείο στο  $[a, b]$ .

**Απόδειξη:** Έστω  $p, r$  δύο σταθερά σημεία στο  $[a, b]$  με  $p \neq r$ . Θα έχουμε τότε  $p - r = g(p) - g(r)$ . Από το Θεώρημα Μέσης Τιμής προβλέπεται ότι υπάρχει  $\xi \in [a, b]$  ώστε  $g(p) - g(r) = g'(\xi)(p - r) < (p - r)$ . Επομένως, στο συγκεκριμένο  $\xi$  έχουμε  $g'(\xi) = 1$ , αντίθετα με την αρχική υπόθεση.

**Παράδειγμα:** Η  $g(x) = \frac{x^2-1}{3}$  έχει μοναδικό σταθερό σημείο στο  $[-1, 1]$  καθώς, όταν  $|x| \leq 1$ , ισχύει α)  $-1/3 \leq g(x) \leq 0$  και κατ' επέκταση,  $-1 \leq g(x) \leq 1$ , και β)  $|g'(x)| = |2x/3| < 1$ .

**Σύγκλιση της μεθόδου.** Έστω  $g(x)$  συνεχής και διαφορίσιμη συνάρτηση στο  $[a, b]$ , με  $a \leq g(x) \leq b$  και  $|g'(x)| \leq k < 1 \forall x \in [a, b]$ . Τότε, αν  $x_0 \in [a, b]$ , η ακολουθία  $x_{n+1} = g(x_n)$ ,  $n = 0, 1, \dots$  συγκλίνει στο μοναδικό σταθερό σημείο,  $\bar{x}$ , της  $g(x)$  στο  $[a, b]$ . Το σφάλμα  $|x_n - \bar{x}| \leq k^n \max(x_0 - a, b - x_0)$ ,  $n \geq 1$ .

### Παραδείγματα :

1. Έστω η συνάρτηση  $f(x) = x^2 - 6x + 5$  με ρίζες 1.0, 5.0. Ας δοκιμάσουμε να τις εντοπίσουμε με την επαναληπτική σχέση

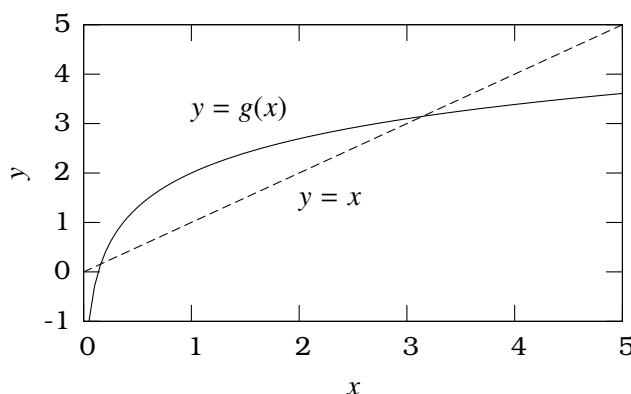
$$g(x) = \frac{x^2 + 5}{6} = x.$$

Για  $x_0 = 2.5$  έχουμε

$$\begin{aligned} x_1 = g(x_0) &= 1.8750 \\ x_2 = g(x_1) &\approx 1.4193 \\ x_3 = g(x_2) &\approx 1.1691 \\ x_4 = g(x_3) &\approx 1.0611 \\ x_5 = g(x_4) &\approx 1.0210 \\ x_6 = g(x_5) &\approx 1.0078 \\ x_7 = g(x_6) &\approx 1.0024 \\ x_8 = g(x_7) &\approx 1.0008 \\ x_9 = g(x_8) &\approx 1.0003 \\ x_{10} = g(x_9) &\approx 1.0001 \\ x_{11} = g(x_{10}) &\approx 1.0000 \end{aligned}$$

Αν δοκιμάσουμε άλλο αρχικό σημείο θα έχουμε πάλι σύγκλιση στο 1 ή απόκλιση στο  $+\infty$ . Μπορεί να αποδειχθεί ότι κανένα σημείο εκτός από το  $x_0 = 5.0$  δε δίνει ακολουθία με όριο την άλλη ρίζα.

2. Ας υπολογίσουμε τις ρίζες της  $f(x) = \ln x - x + 2$ ,  $x > 0$ . Γράφουμε  $g(x) = \ln x + 2 = x$ . Καθώς η  $g(x)$  είναι αύξουσα και  $g(1) = 2$ , υπάρχει ρίζα στο  $[0, 1]$ . Από το γράφημα (Σχήμα 2.3) παρατηρούμε ότι η άλλη ρίζα είναι  $\bar{x} \approx 3.1$ . Αν δοκιμάσουμε με αρχική προσέγγιση



Σχήμα 2.3: Εκτίμηση των σταθερών σημείων της  $g(x) = \ln x + 2$

$x_0 \in \{0.5, 1.0, 1.5, 2.0, 4.0, \dots\}$ , έχουμε σύγκλιση στη ρίζα  $\bar{x} = 3.146193 \dots$ . Αντίθετα, δεν μπορούμε να βρούμε αρχικό σημείο για να εντοπίσουμε την άλλη ρίζα. Παρατηρήστε ότι για  $x_0 \leq e^{-2}$  ή  $x_0 \leq e^{e^{-2}-2}, \dots, x_0 \leq 0.158594339563$  δεν ορίζεται ακολουθία. (Η τιμή 0.158594339563 είναι η άλλη ρίζα· μπορείτε να την εντοπίσετε έχοντας ως  $g(x) = e^{x-2}$ ).

Εξετάστε τη σύγκλιση με διάφορα αρχικά  $x$  για την  $g(x) = x \frac{\ln x + 1}{x-1}$ . Παρατηρήστε ότι διαφορετική επιλογή της  $g(x)$  και της αρχικής προσέγγισης μας δίνει διαφορετική ταχύτητα σύγκλισης (διαφορετικό αριθμό επαναλήψεων).

3. Η  $f(x) = x^3 + 4x^2 - 10 = 0$  έχει μία ρίζα στο  $[1, 1.5]$ . Η μέθοδος  $x = g(x)$  έχει διαφορετική ταχύτητα σύγκλισης ανάλογα με την επιλογή της  $g(x)$ , π.χ.  $g(x) = x - x^3 - 4x^2 + 10$ ,  $g(x) = \sqrt{\frac{10}{x}} - 4x$ ,  $g(x) = \sqrt{\frac{10}{4+x}}$ ,  $g(x) = \frac{1}{2} \sqrt{10 - x^3}$ , κλπ.

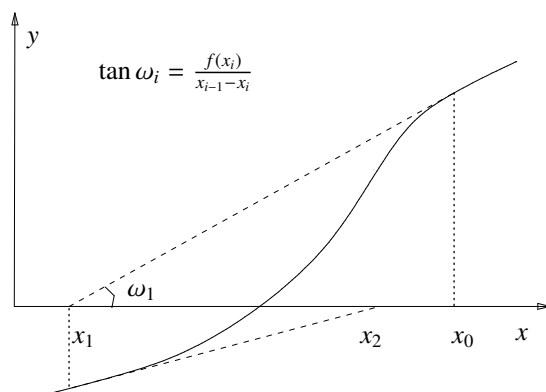
**Παρατήρηση:** Η γενική επαναληπτική μέθοδος  $x_{n+1} = g(x_n)$ ,  $n = 0, 1, \dots$  είναι πρώτης τάξης αν  $g'(x) \neq 0$ , δεύτερης τάξης αν  $g'(x) = 0$  και η  $g''(x)$  είναι συνεχής σε διάστημα που περικλείει τη ρίζα, κλπ.

## 2.4 Μέθοδος Newton-Raphson

Η μέθοδος Newton-Raphson είναι επαναληπτική μέθοδος της μορφής  $x = g(x)$ . Η επιλογή της  $g(x)$  γίνεται ως εξής:

Έστω ότι αναζητούμε τη ρίζα της συνεχούς και διαφορίσιμης, σε διάστημα  $[a, b]$ , συνάρτησης  $f(x)$ . Αν γνωρίζουμε την τιμή αυτής και των παραγώγων της σε κάποιο σημείο  $x_0 \in [a, b]$ , το Θεώρημα Taylor μας εξασφαλίζει ότι στη ρίζα,  $\bar{x}$ , ισχύει

$$f(\bar{x}) = f(x_0) + f'(x_0)(\bar{x} - x_0) + \frac{f''(\xi(x))}{2!}(\bar{x} - x_0)^2.$$



Σχήμα 2.4: Σχηματική εύρεση ρίζας με τη μέθοδο Newton-Raphson

Αγνοώντας τον όρο του υπολοίπου έχουμε

$$0 \approx f(x_0) + f'(x_0)(\bar{x} - x_0) \Rightarrow \bar{x} \approx x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Επομένως, η συνάρτηση  $g(x) = x - \frac{f(x)}{f'(x)}$  μπορεί να παράξει την ακολουθία διαδοχικών προσεγγίσεων στη ρίζα αρκεί να έχουμε  $f'(x_i) \neq 0$ .

**Θεώρημα (χωρίς απόδειξη):** Έστω ότι η  $f(x)$  είναι συνεχής και τουλάχιστον δύο φορές παραγωγίσιμη στο  $[a, b]$ , με συνεχή τη δεύτερη παράγωγό της. Αν  $\bar{x}$  ρίζα της  $f(x)$  στο  $[a, b]$  (δηλ.  $f(\bar{x}) = 0$ ) και  $f'(\bar{x}) \neq 0$  τότε υπάρχει  $\delta > 0$  ώστε η ακολουθία  $\{x_n\}$  που ορίζεται με τη μέθοδο Newton-Raphson συγκλίνει στο  $\bar{x}$ ,  $\forall x_0 \in [\bar{x} - \delta, \bar{x} + \delta]$ .

**Παρατηρήσεις:** Το σφάλμα σε κάθε επανάληψη είναι της τάξης του  $(\bar{x} - x_i)^2$ . οδηγεί σε τετραγωνική σύγκλιση. Αρκούν λίγα βήματα για να έχουμε πολύ ικανοποιητική προσέγγιση της ρίζας, με την προϋπόθεση ότι ξεκινήσουμε από σημείο όχι μακριά από αυτή. Από την άλλη, αν  $f(\bar{x}) \approx 0$  έχουμε πολύ αργή σύγκλιση.

Η μέθοδος είναι δεύτερης τάξης αν  $f'(\bar{x}) \neq 0$  (έχουμε απλή ρίζα). Αυτό προκύπτει αν δούμε τον αλγόριθμο ως εύρεση σταθερού σημείου της  $g(x) = x - \frac{f(x)}{f'(x)}$ . Καθώς

$$\left. \begin{aligned} g'(x) &= \frac{f(x)f''(x)}{[f'(x)]^2} \\ f'(\bar{x}) &\neq 0 \\ f(\bar{x}) &= 0 \end{aligned} \right\} \Rightarrow g'(\bar{x}) = 0.$$

Για πολλαπλή ρίζα ( $f'(\bar{x}) = 0$ ) η μέθοδος είναι πρώτης τάξης. Δεύτερης τάξης είναι ο αλγόριθμος

$$x_{n+1} = x_n - \frac{f(x_n)f'(x_n)}{[f'(x_n)]^2 - f(x_n)f''(x_n)}$$

ή ο

$$x_{n+1} = x_n - m \frac{f(x_n)}{f'(x_n)},$$

όπου  $m$  η πολλαπλότητα της ρίζας.

**Παράδειγμα:** Έστω  $f(x) = x^2 - 6x + 5$ . Έχουμε

$$x_{n+1} = x_n - \frac{x_n^2 - 6x_n + 5}{2x_n - 6}, \quad n = 0, 1, 2, \dots$$

Οι διαδοχικές προσεγγίσεις των ριζών 1.0, 5.0 με αρχικά σημεία 2.0, 6.0 είναι οι εξής

| $n$ | $x_n^{(1)}$       | $x_n^{(2)}$       |
|-----|-------------------|-------------------|
| 0   | 2.0               | 6.0               |
| 1   | 0.5               | 5.166666666666667 |
| 2   | 0.95              | 5.00641025641026  |
| 3   | 0.999390243902439 | 5.00001024002622  |
| 4   | 0.999999907077705 | 5.00000000002621  |
| 5   | 0.999999999999998 | 5.0               |
| 6   | 1.0               |                   |

## 2.5 Μέθοδος τέμνουσας

Σύμφωνα με αυτήν τη μέθοδο, προσεγγίζουμε τη συνάρτηση  $f(x)$  με ευθεία που περνά από δύο σημεία  $(x_{n-1}, f(x_{n-1}))$  και  $(x_n, f(x_n))$ . Τα  $x_{n-1}, x_n$  είναι διαδοχικές προσεγγίσεις της ρίζας. Η νέα προσέγγιση,  $x_{n+1}$ , είναι η τομή με τον άξονα  $x$  (η ρίζα) της προσεγγιστικής ευθείας. Η ευθεία  $y = y(x)$  είναι

$$y = f(x_n) + \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}(x - x_n).$$

Επομένως,

$$x_{n+1} = x_n - \frac{f(x_n)(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})}.$$

Όπως καταλαβαίνετε, πρέπει να επιλέξουμε δύο αρχικά σημεία,  $x_0, x_1$ , ώστε να παράγουμε την ακολουθία. Από την άλλη, η κάθε επανάληψη χρειάζεται ένα μόνο νέο υπολογισμό τιμής της συνάρτησης, πράγμα σημαντικό όταν ο υπολογισμός είναι σχετικά αργός.

**Αλγόριθμος:** Επίλυση της  $f(x) = 0$  με τη μέθοδο της τέμνουσας:

1. Επιλέγουμε δύο τιμές  $a, b$ .
2. Βρίσκουμε την τομή με τον άξονα των  $x$  της ευθείας που περνά από τα σημεία  $(a, f(a)), (b, f(b))$ . Την ονομάζουμε  $c$ .
3. Αν το  $c$  είναι ικανοποιητική προσέγγιση της ρίζας πηγαίνουμε στο βήμα 6.
4. Θέτουμε  $a \leftarrow b, b \leftarrow c$ .
5. Επαναλαμβάνουμε τη διαδικασία από το βήμα 2.
6. Τέλος.

## 2.6 Ασκήσεις

1. Υλοποιήστε τον αλγόριθμο διχοτόμησης με υπορουτίνα της Fortran. Χρησιμοποιήστε τη για να εντοπίσετε τη ρίζα της

- $f(x) = x^3 + 4x^2 - 10$  στο διάστημα  $[1, 2]$ ,
- $f(x) = \sqrt{x} - \cos x$  στο διάστημα  $[0, 1]$ .

2. Δείξτε ότι η  $g(x) = \ln x + 2$  έχει ένα και μοναδικό σταθερό σημείο στο  $[2, 4]$ . Υπολογίστε το μέγιστο αριθμό επαναλήψεων ώστε  $|x_n - \bar{x}| \leq 10^{-3}$ .

3. Παρά το γεγονός ότι η μέθοδος διχοτόμησης είναι μια απολύτως αποδεκτή μέθοδος για τον προσδιορισμό των ριζών συναρτήσεων μιας μεταβλητής, η μέθοδος είναι σχετικά αναποτελεσματική. Ένα μειονέκτημα της μεθόδου διχοτόμησης είναι ότι με τον χωρισμό του διαστήματος από  $x_1$  σε  $x_2$  σε ίσα μισά, δε λαμβάνεται υπόψη η πληροφορία για το μέγεθος των  $f(x_1)$  και  $f(x_2)$ .

Τροποποιήστε τη μέθοδο διχοτόμησης σε μια εναλλακτική μέθοδο όπου η νέα τιμή του  $x$  να εξαρτάται από τις τιμές των  $f(a)$  και  $f(b)$ . Δοκιμάστε να βρίσκετε (σε κάθε επανάληψη) την ευθεία που περνά από τα σημεία  $(a, f(a))$  και  $(b, f(b))$ , και ορίστε σαν νέα προσέγγιση,  $x$ , την τομή αυτής με τον άξονα των  $x$ , (αντί για το μέσο του  $[a, b]$  της μεθόδου διχοτόμησης). Θα καταλήξετε στη μέθοδο *Ψευδούς Θέσης*<sup>2</sup>.

(α') Γράψτε ένα πρόγραμμα Fortran το οποίο να υλοποιεί τη νέα μέθοδο.

(β') Εφαρμόστε την για να βρείτε τη ρίζα της

$$f(x) = -2.0 + 6.2x - 4.0x^2 + 0.7x^3$$

στο διάστημα  $[0.4, 0.6]$ .

(γ') Εφαρμόστε τη μέθοδο *ψευδούς θέσης* και τη μέθοδο διχοτόμησης για να βρείτε τις ρίζες της

$$f(x) = x^{10} - 0.95$$

στο διάστημα  $[0, 1.4]$ . Ποια μέθοδος συγκλίνει πιο γρήγορα με σχετικό σφάλμα  $< 10^{-6}$ ;

Πότε η νέα μέθοδος συμπεριφέρεται χειρότερα από τη μέθοδο διχοτόμησης;

4. Γράψτε κώδικα που να υλοποιεί τη γενική επαναληπτική μέθοδο  $x = g(x)$ . Χρησιμοποιήστε τον για να υπολογίσετε

- μια ρίζα της  $f(x) = x^2 - 6x + 5$ ,
- τη ρίζα της  $f(x) = x - \cos^3 x$  κοντά στο 0.6.

5. Υπολογίστε το  $y = \frac{e^x - 1}{x}$  με ένα ευσταθή αλγόριθμο για μικρό, κατ' απόλυτη τιμή,  $x$ . Για μικρό  $|x|$  χρησιμοποιούμε το ανάπτυγμα Taylor του  $e^x$  ώστε να αποφύγουμε την αλληλοαναίρεση όρων ίδιας τάξης.

6. Υπολογίστε με ευσταθή αλγόριθμο τις λύσεις των εξισώσεων

<sup>2</sup>[http://en.wikipedia.org/wiki/False\\_position\\_method](http://en.wikipedia.org/wiki/False_position_method)

(α)  $1.5x^2 + 13 \times 10^6x + 0.037 = 0$ .

Οι ακριβείς είναι  $x_1 \approx -2.8462 \times 10^{-9}$ ,  $x_2 \approx -8.6667 \times 10^6$ .

(β)  $1.5x^2 - 37 \times 10^6x + 0.057 = 0$ .

Οι ακριβείς είναι  $x_1 \approx 1.5405 \times 10^{-9}$ ,  $x_2 \approx 2.4667 \times 10^7$ .

7. Εφαρμόστε τη μέθοδο Newton–Raphson για να υπολογίσετε τις ρίζες της

(α)  $f(x) = \sin x - x^2$ ,

(β)  $f(x) = 3xe^x - 1$ .

8. Υπολογίστε τις ρίζες της  $f(x) = 4 \cos x - e^{-x}$  με ακρίβεια  $10^{-8}$  με τη μέθοδο διχοτόμησης, τη μέθοδο σταθερού σημείου, τη μέθοδο Newton–Raphson και τη μέθοδο τέμνουσας.

9. Βρείτε με 12 ψηφία σωστά το σημείο τομής των καμπυλών  $e^x$ ,  $\tan(2x)$  στο διάστημα  $[-1, 1]$ . Συμβουλή: σχεδιάστε τις καμπύλες.

10. Υλοποιήστε σε κώδικα Fortran τον αλγόριθμο Muller<sup>3</sup> για την επίλυση μη γραμμικών εξισώσεων. Είναι παρόμοιος με τη μέθοδο τέμνουσας αλλά προσεγγίζει τη συνάρτηση με *παραβολή* (εξίσωση της μορφής  $ax^2 + bx + c$ ) και, επομένως, χρειάζεται *τρία* σημεία για τον προσδιορισμό της. Είναι γενικά πιο γρήγορος από τη μέθοδο τέμνουσας. Εφαρμόστε τον για να βρείτε τη μη μηδενική ρίζα της  $f(x) = \sin x - x^2$ .

11. Υλοποιήστε σε κώδικα τη μέθοδο Newton–Raphson, κατάλληλα τροποποιημένη ώστε να υπολογίζει τις ρίζες πολυωνύμου βαθμού  $n$ ,  $p_n(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ , όταν έχουμε ως δεδομένους τους συντελεστές του  $a_0, a_1, \dots, a_n$ . Το πολυώνυμο και η παράγωγός του να υπολογίζονται με τον αλγόριθμο Horner.

12. Υλοποιήστε σε κώδικα τη μέθοδο τέμνουσας, κατάλληλα τροποποιημένη ώστε να υπολογίζει τις ρίζες πολυωνύμου βαθμού  $n$ ,  $p_n(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ , όταν έχουμε ως δεδομένους τους συντελεστές του  $a_0, a_1, \dots, a_n$ . Το πολυώνυμο και η παράγωγός του να υπολογίζονται με τον αλγόριθμο Horner.

<sup>3</sup><http://mathworld.wolfram.com/MullersMethod.html>



## Κεφάλαιο 3

# Επίλυση Γραμμικών Συστημάτων

### 3.1 Εισαγωγή

Αριθμητική λύση γενικών γραμμικών συστημάτων  $n \times n$   $A_{n \times n} x_{n \times 1} = b_{n \times 1}$ , όπου

$$A = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

**Θεώρημα:** Τα παρακάτω είναι ισοδύναμα:

1. Για κάθε δεύτερο μέλος  $b$ , το σύστημα  $Ax = b$  έχει μοναδική λύση.
2. Ο πίνακας  $A$  έχει αντίστροφο ( $A^{-1}$ ).
3. Η ορίζουσα του  $A$ ,  $\det A$  είναι μη μηδενική.
4. Το ομογενές σύστημα  $Ax = 0$  έχει μοναδική λύση τη  $x = 0$ .
5. Οι στήλες ή οι γραμμές του  $A$  είναι γραμμικά ανεξάρτητες.

Τις βασικές μεθόδους επίλυσης γραμμικών συστημάτων τις διακρίνουμε σε απ' ευθείας και σε επαναληπτικές (iterative).

### 3.2 Μέθοδος Cramer

Η μέθοδος Cramer προσδιορίζει τη λύση του γραμμικού συστήματος ως εξής

$$x_j = \frac{\det B_j}{\det A}, \quad j = 1, 2, \dots, n,$$

όπου ο πίνακας  $B_j$  προκύπτει από τον  $A$  αν αντικαταστήσουμε την  $j$ -στήλη του με το διάνυσμα  $b$ .

Η λύση με αυτή τη μέθοδο απαιτεί  $(n + 1)!$  πολλαπλασιασμούς και γι' αυτό δεν εφαρμόζεται στην πράξη για  $n \geq 4$ .

**Παρατήρηση:** Ο υπολογισμός της ορίζουσας μπορεί να γίνει με τις μεθόδους που παρουσιάζονται στην §3.5.1.

### 3.3 Απ' ευθείας απαλοιφή Gauss

**Θεώρημα:** Αν στο σύστημα γραμμικών εξισώσεων  $Ax = b$

1. πολλαπλασιάσουμε μια εξίσωση με ένα μη μηδενικό αριθμό,
2. ανταλλάξουμε τη σειρά δύο εξισώσεων,
3. αντικαταστήσουμε την εξίσωση  $j$  από το άθροισμα των εξισώσεων  $i$  και  $j$ , (και φυσικά από οποιοδήποτε γραμμικό συνδυασμό τους)

τότε το νέο σύστημα  $A'x = b'$  έχει την ίδια λύση με το αρχικό. Επιπλέον,  $\det A' \neq 0 \Leftrightarrow \det A \neq 0$ . Η απαλοιφή Gauss γίνεται σε δύο στάδια

**Τριγωνοποίηση** Χρήση των παραπάνω ιδιοτήτων και μετατροπή του συστήματος  $Ax = b$  σε άνω τριγωνική μορφή  $A'x = b'$  όπου  $a'_{ij} = 0, i > j$ . Δηλαδή

$$\begin{bmatrix} a'_{11} & a'_{12} & \cdots & a'_{1n} \\ 0 & a'_{22} & \cdots & a'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a'_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b'_1 \\ b'_2 \\ \vdots \\ b'_n \end{bmatrix}. \quad (3.1)$$

Η διαδικασία λέγεται *τριγωνοποίηση* και γίνεται ως εξής:

1. Αν  $a_{11} = 0$  εναλλάσσουμε τις γραμμές του  $A$  ώστε  $a_{11} \neq 0$ . Αν όλα τα  $a_{i1} = 0$  τότε  $\det A = 0$  και το σύστημα έχει άπειρες ( $b = 0$ ) ή καμία λύση ( $b \neq 0$ ).
2. Καθώς  $a_{11} \neq 0$ , πολλαπλασιάζουμε την πρώτη εξίσωση με  $-a_{21}/a_{11}$  και την προσθέτουμε στη δεύτερη. Η δεύτερη εξίσωση αποτελείται πλέον από τα στοιχεία  $a'_{2j}, b'_2$  όπου

$$a'_{21} = 0, \quad a'_{2j} = a_{2j} - a_{1j} \frac{a_{21}}{a_{11}}, \quad b'_2 = b_2 - b_1 \frac{a_{21}}{a_{11}}.$$

Επαναλαμβάνουμε για όλες τις εξισώσεις ώστε να μηδενίσουμε την πρώτη στήλη από το δεύτερο στοιχείο και κάτω. Δηλαδή, πολλαπλασιάζουμε την πρώτη γραμμή με  $-a_{j1}/a_{11}$  και την προσθέτουμε στην  $j$  εξίσωση. Η διαδικασία αυτή μετατρέπει το σύστημα σε νέο,  $A'x = b'$  με την ακόλουθη μορφή

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a'_{22} & \cdots & a'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a'_{n2} & \cdots & a'_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b'_2 \\ \vdots \\ b'_n \end{bmatrix}.$$

Κατόπιν, επαναλαμβάνουμε την ίδια διαδικασία στον πίνακα  $A'_{11}$  (ο  $A'$  χωρίς την πρώτη γραμμή και την πρώτη στήλη) με δεξί μέλος το  $[b'_2, \dots, b'_n]^T$ .

Η κατάληξη αυτής της διαδικασίας φέρνει το αρχικό σύστημα στη μορφή (3.1).

**Οπισθοδρόμηση** Η εύρεση της λύσης γίνεται με τη μέθοδο οπισθοδρόμησης, από την τελευταία προς την πρώτη εξίσωση. Έχουμε

$$x_n = \frac{b'_n}{a'_{nn}}, \quad x_k = \frac{1}{a'_{kk}} \left( b'_k - \sum_{j=k+1}^n a_{kj} x_j \right), \quad k = n-1, n-2, \dots, 1.$$

**Παρατήρηση:** Όλα τα  $a'_{jj} \neq 0$ , διαφορετικά στο αρχικό σύστημα θα είχαμε  $\det A = 0$ .

**Παράδειγμα:** Το σύστημα

$$\begin{bmatrix} 0 & 1 & 2 \\ 5 & 3 & 1 \\ 2 & -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 6 \end{bmatrix}.$$

επιλύεται ως εξής:

τα διαδοχικά στάδια τριγωνοποίησης δίνουν

1. Καθώς  $a_{11} = 0$  και  $a_{21} \neq 0$  εναλλάσσουμε τις δύο πρώτες εξισώσεις

$$\begin{bmatrix} 5 & 3 & 1 \\ 0 & 1 & 2 \\ 2 & -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 6 \end{bmatrix}.$$

Η δεύτερη εξίσωση έχει ήδη  $a_{21} = 0$ , όπως επιδιώκουμε. Πολλαπλασιάζουμε την πρώτη εξίσωση με  $-2/5$  και την προσθέτουμε στην τρίτη

$$\begin{bmatrix} 5 & 3 & 1 \\ 0 & 1 & 2 \\ 0 & -3.2 & 0.6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 4.4 \end{bmatrix}.$$

Πολλαπλασιάζουμε τη δεύτερη γραμμή με 3.2 και την προσθέτουμε στην τρίτη

$$\begin{bmatrix} 5 & 3 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 14 \end{bmatrix}.$$

Με οπισθοδρόμηση έχουμε

$$x_3 = 2, \quad x_2 = -1, \quad x_1 = 1.$$

**Αλγόριθμος:**

**Τριγωνοποίηση** Για  $k = 1, 2, \dots, n-1$ :

1. Αν  $a_{kk} = 0$ ,
  - βρίσκουμε το πρώτο μη μηδενικό  $a_{ik}$ ,  $i = k+1, \dots, n$ . Αν δεν υπάρχει κανένα τέτοιο στοιχείο, το σύστημα δεν έχει μοναδική λύση.
  - Εναλλάσσουμε την  $k$  με την  $i$  εξίσωση (πίνακες  $A$  και  $b$ ).
2. Για κάθε  $i = k+1, \dots, n$ ,

- Θέτουμε

$$\begin{aligned}\lambda &= -\frac{a_{ik}}{a_{kk}}, \\ a_{ij} &\leftarrow a_{ij} + \lambda a_{kj}, \quad j = k, \dots, n \\ b_j &\leftarrow b_j + \lambda b_k, \quad j = k, \dots, n\end{aligned}$$

**Οπισθοδρόμηση** Για  $k = n, n-1, \dots, 1$ , θέτουμε

$$x_k = \frac{1}{a_{kk}} \left( b_k - \sum_{j=k+1}^n a_{kj} x_j \right).$$

Αν  $k+1 > n$  το άθροισμα δεν υπολογίζεται.

**Παρατηρήσεις:**

**Πολλαπλά δεξιά μέλη,**  $b = b_{n \times m}$  Όταν θέλουμε να επιλύσουμε πολλές φορές το σύστημα με ίδιο πίνακα  $A$  αλλά  $m$  διαφορετικά δεξιά μέλη  $b$ , είναι προτιμότερο να εκτελέσουμε συγχρόνως την διαδικασία για όλα τα  $b$ , δηλαδή, να σχηματίσουμε ένα πίνακα  $b$  με  $m$  στήλες και να επεκτείνουμε τις πράξεις που υπαγορεύει ο αλγόριθμος για το  $b$  σε όλες τις στήλες.

**Απαιτούμενες πράξεις.** Αν μετρήσουμε πολλαπλασιασμούς και διαιρέσεις (που είναι πιο χρονοβόρες από τις προσθέσεις και αφαιρέσεις) η μέθοδος Gauss χρειάζεται

$$\sum_{k=1}^{n-1} \left[ \underbrace{(n-k)^2}_{\text{για } a_{ij}} + \underbrace{2(n-k)}_{\text{για } b_i} \right] + \sum_{k=1}^n (n-k+1) = \frac{n^3}{3} + n^2 - \frac{n}{3}$$

πράξεις, πολύ λιγότερες από τις  $(n+1)!$  που χρειάζεται η μέθοδος Cramer.

**Απαιτήσεις μνήμης.** Ο γενικός πίνακας  $A$  χρειάζεται  $n^2$  θέσεις μνήμης για πραγματικούς ή μιγαδικούς (ό,τι τύπου είναι τα στοιχεία του). Επιπλέον  $n$  θέσεις απαιτεί ο  $b$ .

**Μερική οδήγηση κατά γραμμές.** Για να ελαχιστοποιήσουμε τα αριθμητικά σφάλματα κατά την τριγωνοποίηση, είναι σημαντικό να επιλέξουμε το διαγώνιο στοιχείο  $a_{kk}$  (που διαιρεί την  $k$  γραμμή) ώστε να είναι αρκετά μεγάλο κατ' απόλυτη τιμή. Μπορούμε να κάνουμε κατάλληλη εναλλαγή γραμμών (της  $k$  με κάποια από τις επόμενες, με  $i > k$ ) ώστε να μεταφερθεί στη διαγώνιο το μεγαλύτερο κατ' απόλυτη τιμή στοιχείο από τα  $a_{ik}$ ,  $i \geq k$ .

**Παράδειγμα:** Το σύστημα

$$\begin{bmatrix} 0.0003 & 1.566 \\ 0.3454 & -2.436 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1.569 \\ 1.018 \end{bmatrix}$$

έχει λύση  $x_1 = 10$ ,  $x_2 = 1$ . Όμως, αν υποθέσουμε Η/Υ με αναπαράσταση αριθμών

$$\pm 0.f_1 f_2 \dots f_n \times 10^{\pm |s|}, \quad |s| \leq 10, \quad n = 5,$$

η απλή απαλοιφή Gauss δίνει προσεγγιστικά μετά την τριγωνοποίηση

$$\begin{bmatrix} 0.3 \times 10^{-3} & 0.1566 \times 10^1 \\ 0 & -0.1804 \times 10^1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0.1569 \times 10^1 \\ -0.1805 \times 10^0 \end{bmatrix}$$

και τότε,  $x_2 = 1.0006$ ,  $x_1 = 6.868$ . Η οδήγηση με εναλλαγή γραμμών είναι απαραίτητη για να βρούμε τα ακριβή  $x_1, x_2$ . Έτσι, αν εναλλάξουμε την πρώτη με τη δεύτερη εξίσωση, αν, δηλαδή, ξεκινήσουμε με το σύστημα

$$\begin{bmatrix} 0.3454 & -2.436 \\ 0.0003 & 1.566 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1.018 \\ 1.569 \end{bmatrix}$$

η τριγωνοποίηση δίνει

$$\begin{bmatrix} 0.3454 \times 10^0 & -0.2436 \times 10^1 \\ 0 & 0.1568 \times 10^1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0.1018 \times 10^1 \\ 0.1568 \times 10^1 \end{bmatrix}.$$

Συνεπώς,  $x_2 = 1$ ,  $x_1 = 10$ .

**Ολική οδήγηση (κατά γραμμές και στήλες).** Φέρνουμε στο  $a_{kk}$  με κατάλληλη εναλλαγή γραμμών και στηλών το μεγαλύτερο κατ' απόλυτη τιμή στοιχείο όλου του πίνακα. Προσέξτε ότι η εναλλαγή στηλών απαιτεί και εναλλαγή στοιχείων στο διάνυσμα  $x$ .

### 3.4 Ευστάθεια γραμμικών συστημάτων

Αν έχουμε μικρές μεταβολές στα  $A, b$  πόσο αλλάζει η λύση;

**Παράδειγμα:**

$$\begin{bmatrix} 1 & 3 \\ 1 & 3.01 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4.01 \end{bmatrix}$$

έχει λύση  $x_1 = x_2 = 1$ .

Το ελαφρά διαφορετικό σύστημα

$$\begin{bmatrix} 1 & 3 \\ 1 & 2.99 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4.02 \end{bmatrix}$$

έχει λύση  $x_1 = 10$ ,  $x_2 = -2$ , τελείως διαφορετική.

**Ορισμός:** Το σύστημα  $Ax = b$  χαρακτηρίζεται ως ασταθές αν έχουμε μεγάλη απόκλιση στη λύση για μικρές αλλαγές στα  $A, b$ .

Κριτήριο ευστάθειας (well-conditioned system) είναι ο δείκτης κατάστασης  $\mu$

$$\mu = \|A\| \cdot \|A^{-1}\|$$

του πίνακα  $A$  ως προς τη νόρμα  $\|\cdot\|$ . Π.χ. μία νόρμα είναι η “νόρμα αθροίσματος γραμμών”

$$\|A\|_{\infty} = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|.$$

Αν  $\mu \gg 1$  το σύστημα είναι ασταθές.

### 3.5 Εφαρμογές

Η διαδικασία τριγωνοποίησης με τη μέθοδο απαλοιφής Gauss που περιγράφηκε παραπάνω βρίσκει εφαρμογή και σε άλλα προβλήματα γραμμικής άλγεβρας, πέρα από την επίλυση γραμμικών συστημάτων.

#### 3.5.1 Υπολογισμός Ορίζουσας

Η πρόσθεση σε μία γραμμή ενός (τετραγωνικού) πίνακα του πολλαπλάσιου μίας άλλης είναι διαδικασία που διατηρεί την ορίζουσα. Ο άνω τριγωνικός πίνακας που παράγεται με την απαλοιφή Gauss όπως περιγράφηκε, αν η τριγωνοποίηση περιοριστεί μόνο σε τέτοιες μεταβολές, έχει πολύ εύκολο υπολογισμό της ορίζουσας του και, επομένως, και της ορίζουσας του αρχικού. Η εφαρμογή της σχέσης (3.2), με ανάπτυξη κατά την πρώτη στήλη, δίνει ως ορίζουσα το γινόμενο των διαγώνιων στοιχείων του :

$$\det A = \prod_{i=1}^n a'_{ii} .$$

Προσέξτε ότι τα στοιχεία  $a'_{ii}$  είναι τα διαγώνια στοιχεία του τριγωνικού πίνακα. Επίσης, σε περίπτωση που εφαρμόσουμε οδήγηση (δηλ. εναλλαγή γραμμών ή στηλών) πρέπει να λάβουμε υπόψη ότι κάθε τέτοια μεταβολή αλλάζει το πρόσημο της ορίζουσας.

Επομένως, η ορίζουσα  $\det A$  μπορεί να υπολογιστεί ως το γινόμενο των στοιχείων της διαγωνίου του τελικού πίνακα (μετά την άνω ή κάτω τριγωνοποίηση),  $A'$ , επί  $(-1)^s$  όπου  $s$  είναι ο συνολικός αριθμός εναλλαγών γραμμών (ή στηλών) που έγιναν κατά την απαλοιφή.

Εναλλακτικά, ο υπολογισμός της ορίζουσας μπορεί να γίνει αναδρομικά από τη σχέση

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det A_{ij} , \quad (3.2)$$

όπου  $A_{ij}$  είναι ο πίνακας διαστάσεων  $(n-1) \times (n-1)$  που προκύπτει από τον  $A$  διαγράφοντας τη γραμμή  $i$  και τη στήλη  $j$ .

#### 3.5.2 Μέθοδος Gauss–Jordan για τον υπολογισμό του αντίστροφου πίνακα

Μια μέθοδος επίλυσης γραμμικών συστημάτων, εναλλακτική της τριγωνοποίησης και οπισθοδρόμησης είναι η εξής: Αφού φέρουμε τον πίνακα σε άνω τριγωνική μορφή με την απαλοιφή Gauss επαναλαμβάνουμε τη διαδικασία από την τελευταία γραμμή προς την πρώτη ώστε να μηδενίσουμε και τα στοιχεία πάνω από τη διαγώνιο. Επομένως, με αυτήν τη διαδικασία, ένα σύστημα της μορφής

$$Ax = B$$

γίνεται

$$A'x = B' ,$$

όπου ο  $A'$  είναι διαγώνιος πίνακας, ο οποίος με πολύ εύκολη τροποποίηση μπορεί να γίνει ο μοναδιαίος,

$$Ix = B'' .$$

Η μέθοδος αυτή παράγει απ' ευθείας τη λύση του συστήματος, απαιτεί όμως περισσότερες πράξεις από την τριγωνοποίηση σε συνδυασμό με την οπισθοδρόμηση, και γι' αυτό δε χρησιμοποιείται συνήθως. Όμως, προσέξτε ότι οποιαδήποτε μέθοδος επίλυσης παράγει τελικά το

$$x = A^{-1}B .$$

Συνεπώς, αν επιλέξουμε για πίνακα  $B$  διαδοχικά τα  $n$  διανύσματα  $(1, 0, \dots, 0)^T$ ,  $(0, 1, \dots, 0)^T$ ,  $\dots$ ,  $(0, 0, \dots, 1)^T$  θα έχουμε ως λύσεις τις αντίστοιχες στήλες του πίνακα  $A^{-1}$ . Η μέθοδος αυτή για την εύρεση του αντιστρόφου ενός πίνακα  $A_{n \times n}$  είναι η μέθοδος Gauss-Jordan και απαιτεί την επίλυση  $n$  γραμμικών συστημάτων  $Ax = B$  με διαφορετικά δεξιά μέλη. Κατά τη διαδικασία της τριγωνοποίησης, οποιαδήποτε μεταβολή των συστημάτων καθορίζεται αποκλειστικά από τα στοιχεία του  $A$  και, συνεπώς, μπορούν να επιλυθούν ταυτόχρονα. Επομένως, για την εύρεση του αντιστρόφου ενός τετραγωνικού πίνακα  $A$  ξεκινούμε από τον πίνακα

$$\left[ \begin{array}{cccc|cccc} a_{11} & a_{12} & \cdots & a_{1n} & 1 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 1 \end{array} \right]$$

και εκτελούμε πλήρη τριγωνοποίηση στο αριστερό τμήμα του ώστε να παράγουμε τον πίνακα

$$\left[ \begin{array}{cccc|cccc} 1 & 0 & \cdots & 0 & \tilde{a}_{11} & \tilde{a}_{12} & \cdots & \tilde{a}_{1n} \\ 0 & 1 & \cdots & 0 & \tilde{a}_{21} & \tilde{a}_{22} & \cdots & \tilde{a}_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & \tilde{a}_{n1} & \tilde{a}_{n2} & \cdots & \tilde{a}_{nn} \end{array} \right].$$

Το δεξί τμήμα αυτού αποτελεί τον αντίστροφο πίνακα,  $A^{-1}$ .

### 3.5.3 Εύρεση ιδιοτιμών και ιδιοδιανυσμάτων

ΑΣ θυμίσουμε τον ορισμό των εννοιών του ιδιοδιανύσματος και της ιδιοτιμής ενός πίνακα  $A$ .

Αν υπάρχει ένας αριθμός  $\lambda$ , εν γένει *μγαδικός*, και ένα διάνυσμα (πίνακας-στήλη)  $x$ , διάφορο του  $(0, 0, \dots, 0)$  για τα οποία ισχύει

$$Ax = \lambda x, \quad (3.3)$$

τότε το  $x$  λέγεται *ιδιοδιάνυσμα* του  $A$  ενώ το  $\lambda$  είναι η αντίστοιχη *ιδιοτιμή*. Παρατηρήστε ότι το  $x$  δεν είναι μοναδικό καθώς οποιοδήποτε πολλαπλάσιό του αποτελεί επίσης λύση του συστήματος (3.3) για την ίδια ιδιοτιμή.

Η (3.3) μπορεί να γραφεί ως εξής

$$Ax = \lambda x \Rightarrow Ax = \lambda Ix \Rightarrow (A - \lambda I)x = 0.$$

Το σύστημα έχει μοναδική λύση, την  $x = (0, 0, \dots, 0)^T$ , αν και μόνο αν ο πίνακας  $A - \lambda I$  αντιστρέφεται. Καθώς δεν ενδιαφερόμαστε για τη μηδενική λύση, οδηγούμαστε στην απαίτηση να ισχύει  $\det(A - \lambda I) = 0$ . Παρατηρήστε ότι η έκφραση  $\det(A - \lambda I)$  είναι ένα πολυώνυμο  $n$  βαθμού ως προς  $\lambda$ . Η εύρεση των  $n$  ριζών του μπορεί να γίνει αναλυτικά (για  $n < 5$ ) ή, γενικότερα, αριθμητικά με τις μεθόδους που περιγράψαμε στο Κεφάλαιο 2.

Εάν επιθυμούμε, μπορούμε να υπολογίσουμε τους συντελεστές του πολυωνύμου:

$$(-1)^n (\lambda^n - p_{n-1} \lambda^{n-1} \cdots - p_1 \lambda - p_0) = 0.$$

Η μέθοδος Faddeev-Leverrier είναι μια αποτελεσματική διαδικασία για τον υπολογισμό των συντελεστών  $p_i$  του πολυωνύμου. Η μέθοδος χρησιμοποιεί μια ακολουθία πινάκων  $\{B\}$  για τον υπολογισμό των συντελεστών  $p_i$ .

**Αλγόριθμος Faddeev–Leverrier:**

1. Θέτουμε  $k \leftarrow 1$ ,  $B_{n-k} \leftarrow A$ ,  $p_{n-k} \leftarrow \text{Tr}B_{n-k}^1$ .
2. Υπολογίζουμε τον πίνακα  $B_{n-(k+1)} \leftarrow A(B_{n-k} - p_{n-k}I)$  και τον συντελεστή  $p_{n-k} \leftarrow \frac{1}{k}\text{Tr}B_{n-k}$ .
3. Θέτουμε  $k \leftarrow k + 1$ . Αν ισχύει  $k < n$  συνεχίζουμε με το βήμα 2.

Ένα επιπλέον πλεονέκτημα της μεθόδου είναι ότι ο αντίστροφος πίνακας μπορεί κατόπιν να υπολογιστεί πολύ εύκολα ως:

$$A^{-1} = \frac{1}{p_0} (B_1 - p_1 I) .$$

Αφού προσδιοριστούν οι ιδιοτιμές, η επίλυση του γραμμικού συστήματος  $(A - \lambda I)x = 0$  μπορεί να γίνει με τις μεθόδους που παρουσιάστηκαν. Προσέξτε ότι το σύστημα έχει άπειρες λύσεις, οπότε τουλάχιστον μία από τις συνιστώσες του διανύσματος  $x$  είναι “ελεύθερη”. Αυτό προκύπτει καθώς κατά την επίλυση του συστήματος καταλήγουμε σε εξίσωση της μορφής  $0x_k = 0$  για κάποιο  $k$ . Τότε, η συγκεκριμένη συνιστώσα μπορεί να τεθεί αυθαίρετα 1. Δεν θα αναφερθούμε στην περίπτωση που εμφανιστεί και δεύτερη “ελεύθερη” συνιστώσα.

**3.6 Ασκήσεις**

1. Υλοποιήστε σε πρόγραμμα FORTRAN το πρώτο στάδιο (τριγωνοποίηση) της απαλοιφής Gauss. Θεωρήστε ότι τα διαγώνια στοιχεία του πίνακα είναι μη μηδενικά.  
Υπόδειξη: Δημιουργήστε ένα πίνακα  $4 \times 4$  με τυχαία στοιχεία για να ελέγξετε το πρόγραμμά σας.
2. Συμπληρώστε το πρόγραμμα της προηγούμενης άσκησης με το δεύτερο στάδιο (οπισθοδρόμηση) της απαλοιφής Gauss.
3. Να γράψετε συνάρτηση της FORTRAN που να υπολογίζει την ορίζουσα ενός πίνακα. Θα δέχεται ως ορίσματα τον πίνακα και, αν σας χρειάζεται, την τάξη του και θα επιστρέφει την ορίζουσα. Μπορείτε να εφαρμόσετε την απαλοιφή Gauss για να φέρετε τον πίνακα σε τριγωνική μορφή. Η διαδικασία αυτή διατηρεί την ορίζουσα, η οποία υπολογίζεται εύκολα από το γινόμενο των διαγώνιων στοιχείων του τριγωνικού πίνακα.

Χρησιμοποιήστε τη για να υπολογίσετε την ορίζουσα του

$$\begin{bmatrix} 2.1 & 3.9 & 0.3 & -4.1 \\ 4.3 & -1.3 & 0.8 & 1.5 \\ 1.0 & -2.8 & 4.3 & -8.1 \\ 2.4 & 6.1 & -1.1 & 12.5 \end{bmatrix} .$$

4. Υλοποιήστε σε πρόγραμμα FORTRAN τη μέθοδο Cramer.

<sup>1</sup>Το Tr συμβολίζει το ίχνος του πίνακα (το άθροισμα των στοιχείων της διαγωνίου).



**Gauss–Seidel.** Στη δεύτερη παραλλαγή, οι “νέες” τιμές των  $x_i$ ,  $x_i^{(k+1)}$ , χρησιμοποιούνται στον τύπο αμέσως μόλις υπολογιστούν:

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left( b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)} \right), \quad \forall i = 1, \dots, n.$$

Ο υπολογισμός του  $x_i^{(k+1)}$  χρειάζεται τις τιμές  $x_j^{(k)}$  για  $j > i$  και τις τιμές  $x_j^{(k+1)}$  για  $j < i$ .

Να γράψετε δύο υπορουτίνες Fortran, με ονόματα JACOBI, SEIDEL που να υλοποιούν τους αντίστοιχους αλγορίθμους. Να τις χρησιμοποιήσετε για την εύρεση της λύσης του συστήματος  $Ax = B$  όπου

$$A = \begin{bmatrix} 12.1 & 3.9 & 0.3 & -4.1 \\ 4.3 & -11.3 & 0.8 & 1.5 \\ 1.0 & -2.8 & 14.3 & -8.1 \\ 2.4 & 6.1 & -1.1 & 12.5 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \quad B = \begin{bmatrix} 1.2 \\ 2.3 \\ 3.4 \\ 4.5 \end{bmatrix}.$$

Να χρησιμοποιήσετε το ακόλουθο κριτήριο τερματισμού των επαναλήψεων: κάθε στοιχείο του διανύσματος  $Ax - B$  να είναι κατ' απόλυτη τιμή μικρότερο του  $10^{-7}$ .

## Κεφάλαιο 4

# Παρεμβολή και Προσέγγιση Συναρτήσεων

### 4.1 Παρεμβολή με πολυώνυμο Lagrange

Έστω ότι γνωρίζουμε τις τιμές μιας συνάρτησης  $f(x)$ ,  $f_0, f_1, \dots, f_n$  σε σημεία  $x_0, x_1, \dots, x_n$ , και ζητάμε να υπολογίσουμε την τιμή  $f(x')$ ,  $x' \neq x_i$  όπου  $x'$  ένα ενδιάμεσο σημείο,  $\min\{x_i\} < x' < \max\{x_i\}$ .

Η παρεμβολή Lagrange προσεγγίζει την  $f(x)$  με ένα πολυώνυμο  $n$ -βαθμού  $p_n(x)$  που περνά από τα σημεία  $(x_i, f_i)$ .

**Θεώρημα:** Έστω η συνάρτηση  $f(x)$ , ορισμένη στο  $[a, b]$ , και  $n+1$  διαφορετικά σημεία  $x_0, x_1, \dots, x_n$  σε αυτό. Τότε υπάρχει μοναδικό πολυώνυμο  $p(x) \in \{\Pi_n(x)\}$  ώστε  $p(x_i) = f_i$ ,  $i = 0, 1, \dots, n$ , το οποίο λέγεται *πολυώνυμο παρεμβολής Lagrange*. Δίνεται από τον τύπο

$$p(x) = \sum_{i=0}^n \ell_i(x) f_i,$$

όπου για  $i = 0, 1, 2, \dots, n$ ,

$$\ell_i(x) = \prod_{j=0, j \neq i}^n \frac{x - x_j}{x_i - x_j}. \quad (4.1)$$

**Παράδειγμα:** Το πολυώνυμο παρεμβολής για τη συνάρτηση  $f(x) = \frac{1}{x}$  στα σημεία παρεμβολής  $x_0 = 2.0$ ,  $x_1 = 2.5$ ,  $x_2 = 4.0$  είναι

$$p(x) = \ell_0(x)f_0 + \ell_1(x)f_1 + \ell_2(x)f_2 = 0.5\ell_0(x) + 0.4\ell_1(x) + 0.25\ell_2(x),$$

όπου

$$\begin{aligned} \ell_0(x) &= \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} = \frac{(x - 2.5)(x - 4)}{(2 - 2.5)(2 - 4)} = x^2 - 6.5x + 10, \\ \ell_1(x) &= \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} = \frac{(x - 2)(x - 4)}{(2.5 - 2)(2.5 - 4)} = -\frac{4}{3}x^2 + 8x + 6.4, \\ \ell_2(x) &= \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} = \frac{(x - 2)(x - 2.5)}{(4 - 2)(4 - 2.5)} = -\frac{1}{3}x^2 + 1.5x + \frac{5}{3}. \end{aligned}$$

Επομένως, το  $p(x) = 0.05x^2 - 0.425x + 1.15$  αποτελεί καλή προσέγγιση της  $f(x) = \frac{1}{x}$  στο διάστημα  $[2, 4]$ .

**Θεώρημα:** Έστω η συνάρτηση  $f(x)$  που είναι συνεχής με  $\nu + 1$  συνεχείς παραγώγους στο  $[a, b]$  για τις οποίες ισχύει  $|f^{(\nu+1)}(\xi)| \leq M, \forall \xi \in [a, b]$ . Έστω ακόμα ότι  $x_0 \equiv a, x_1, x_2, \dots, x_\nu \equiv b$  είναι  $\nu + 1$  διαφορετικά σημεία στο διάστημα  $[a, b]$  και  $p(x)$  το πολυώνυμο παρεμβολής Lagrange για τη συνάρτηση  $f(x)$ . Τότε

$$|f(x) - p(x)| \leq \frac{M}{(\nu + 1)!} \prod_{i=0}^{\nu} |x - x_i|, \quad \forall x \in [a, b].$$

**Παράδειγμα:** Έστω  $f(x) = e^x$ . Ζητείται να υπολογιστεί το σφάλμα παρεμβολής Lagrange στα σημεία  $x_0, x_1$ .

Έχουμε  $\nu = 1$  και  $f''(x) = e^x$ . Άρα  $|f''(x)| \leq e^{x_1}, \forall x \in [x_0, x_1]$ . Από το παραπάνω θεώρημα έχουμε:

$$|f(x) - p(x)| \leq \frac{e^{x_1}}{2!} (x - x_0)(x - x_1).$$

Επομένως, για συγκεκριμένο  $x \in [x_0, x_1]$  μπορούμε να υπολογίσουμε το σφάλμα  $\varepsilon$ , ή, αντίστροφα, για συγκεκριμένο σφάλμα να υπολογίσουμε το  $\Delta x = x_1 - x_0$  ως εξής: πρέπει

$$\max_{x_0 \leq x \leq x_1} |f(x) - p(x)| \leq \frac{e^{x_1}}{2} \max[(x - x_0)(x - x_1)] < \varepsilon.$$

Όμως το  $(x - x_0)(x - x_1)$  γίνεται μέγιστο στο  $x = \frac{x_0 + x_1}{2}$ . Άρα

$$\frac{e^{x_1}}{2} \left( \frac{x_1 - x_0}{2} \right)^2 < \varepsilon \Rightarrow (\Delta x)^2 < \frac{8\varepsilon}{e^{x_1}}.$$

Αν  $x_1 = 1.0$ ,  $\varepsilon = 10^{-6}$ , τότε  $\Delta x < 1.716 \times 10^{-3}$ . Παρατηρούμε ότι αν το διάστημα  $[a, b]$  είναι μικρό και χρησιμοποιούμε μικρού βαθμού πολυώνυμα έχουμε καλή προσέγγιση. Αντίθετα, πολυώνυμα μεγάλου βαθμού τείνουν να έχουν έντονη ταλαντωτική συμπεριφορά στα διαστήματα μεταξύ των σημείων, Σχήμα 4.1.

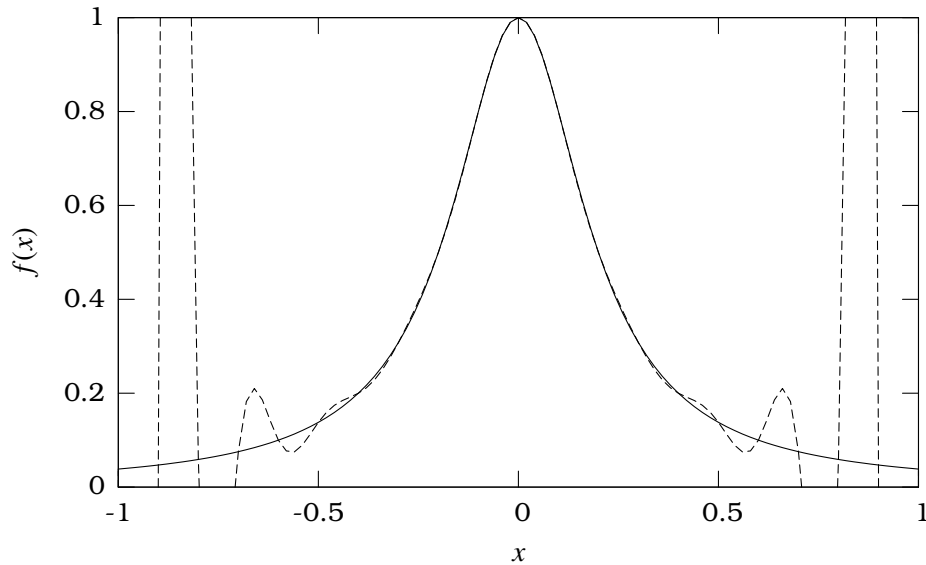
## 4.2 Προσέγγιση με τη μέθοδο ελαχίστων τετραγώνων

Πολλές φορές θέλουμε καλή προσέγγιση σε μία συνάρτηση  $f(x)$  όταν έχουμε ένα σύνολο σημείων  $(x_i, y_i)$  αλλά δεν ισχύει απαραίτητα  $y_i = f(x_i)$ . Τέτοια περίπτωση αποτελούν οι μετρήσεις πειραματικών δεδομένων καθώς περιέχουν σφάλματα.

Με τη μέθοδο ελαχίστων τετραγώνων προσαρμόζουμε στα δεδομένα μας μια συνάρτηση  $f(x)$  προκαθορισμένης μορφής, με παραμέτρους, ώστε το άθροισμα των τετραγώνων των αποκλίσεων από τα  $y_i$ ,  $\sum_{i=1}^n [f(x_i) - y_i]^2$ , να γίνεται ελάχιστο ως προς αυτές τις παραμέτρους.

Έστω ότι η ζητούμενη συνάρτηση είναι γραμμική (πολυώνυμο βαθμού 1). Ορίζουμε τότε τη συνάρτηση  $f(x) = \alpha x + \beta$  όπου οι συντελεστές  $\alpha, \beta$  προσδιορίζονται από τη μέθοδο ελαχίστων τετραγώνων από τις εξής σχέσεις

$$\alpha = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2}, \quad (4.2a)$$



Σχήμα 4.1: Προσέγγιση της συνάρτησης  $f(x) = (1 + 25x^2)^{-1}$  στο  $[-1, 1]$  με πολυώνυμο 20<sup>ου</sup> βαθμού. Παρατηρήστε τις έντονες ταλαντώσεις στα άκρα του διαστήματος

$$\beta = \frac{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2}. \quad (4.26)$$

**Απόδειξη:** Το άθροισμα

$$E(\alpha, \beta) = \sum_{i=1}^n (\alpha x_i + \beta - y_i)^2$$

γίνεται ακρότατο όταν  $\partial E / \partial \alpha = 0$ ,  $\partial E / \partial \beta = 0$ . Οι εξισώσεις οδηγούν στο σύστημα

$$\begin{bmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n y_i \end{bmatrix}$$

Το σύστημα έχει μοναδική λύση, την (4.2α), όταν η ορίζουσα είναι μη μηδενική.

Ο συντελεστής  $r^2$  που προσδιορίζει την ποιότητα της ευθείας (δηλ. το πόσο καλή είναι η προσέγγιση) είναι

$$r^2 \equiv \frac{\sum_{i=1}^n \left[ (\alpha x_i + \beta) - \frac{1}{n} \sum_{j=1}^n y_j \right]^2}{\sum_{i=1}^n \left( y_i - \frac{1}{n} \sum_{j=1}^n y_j \right)^2}.$$

Ισχύει πάντα ότι  $0 < r^2 \leq 1$ . Το  $r^2 = 1$  υποδηλώνει τέλεια προσαρμογή (η ευθεία περνά από όλα τα σημεία), ενώ η τιμή γίνεται μικρότερη από 1 όσο πιο διασκορπισμένα είναι τα σημεία γύρω από την ευθεία.

### 4.2.1 Πολυωνυμική προσέγγιση

Έστω  $f(x) = p(x) = \sum_{i=0}^m \alpha_i x^i$ . Τότε, η προσέγγιση ελαχίστων τετραγώνων οδηγεί στο γραμμικό σύστημα

$$\begin{bmatrix} S_0 & S_1 & \cdots & S_m \\ S_1 & S_2 & \cdots & S_{m+1} \\ \vdots & \vdots & & \vdots \\ S_m & S_{m+1} & \cdots & S_{2m} \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}$$

όπου

$$S_k = \sum_{i=0}^n x_i^k, \quad k = 0, 1, \dots, 2m,$$

$$\beta_k = \sum_{i=0}^n x_i^k y_i, \quad k = 0, 1, \dots, m.$$

Το σύστημα αυτό έχει μοναδική λύση στο  $\mathcal{R}$  αν  $m < n$ ,  $x_i \neq x_j \forall i \neq j$ . Αν  $n = m + 1$ , η λύση αντιστοιχεί στο πολυώνυμο Lagrange  $m$  βαθμού.

Αν το άθροισμα  $\sum_{i=1}^n [f(x_i) - y_i]^2$  έχει τη μορφή σφάλματος που προκύπτει από την αντικατάσταση  $f(x_i) = \alpha x_i + \beta$  ή  $f(x_i) = p(x_i)$ , τότε οι συνθήκες για ελάχιστο είναι και ικανές ( $\Leftarrow$ ) και αναγκαίες ( $\Rightarrow$ ) και η λύση είναι μοναδική.

**Παράδειγμα:** Έστω ότι τα σημεία  $(x_i, y_i)$  είναι  $\{(0.0, 1.0), (0.25, 1.284), (0.5, 1.6487), (0.75, 2.117), (1.0, 2.7183)\}$ ,  $i = 1, \dots, 5$ . Το δευτεροβάθμιο πολυώνυμο που εξάγεται από τη μέθοδο ελαχίστων τετραγώνων προκύπτει ως λύση της

$$\begin{bmatrix} 5.0 & 3.5 & 1.8750 \\ 2.5 & 1.875 & 1.5625 \\ 1.875 & 1.5625 & 1.3828 \end{bmatrix} \cdot \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 8.768 \\ 5.4514 \\ 4.4015 \end{bmatrix} \Rightarrow \begin{cases} \alpha_0 = 1.0052 \\ \alpha_1 = 0.8641 \\ \alpha_2 = 0.8437 \end{cases}.$$

Επομένως,

$$p(x) = 1.0052 + 0.8641x + 0.8437x^2.$$

Προσδιορίστε το αντίστοιχο πρωτοβάθμιο πολυώνυμο. Τι παρατηρείτε;

#### Παρατηρήσεις:

- Συχνά η άγνωστη συνάρτηση  $f(x)$  εξαρτάται γραμμικά από άλλες συναρτήσεις  $f(x) = \sum_{i=0}^m \alpha_i f_i(x)$ . Η μέθοδος οδηγεί (με τις συνθήκες  $\partial E / \partial \alpha_i = 0 \forall i = 0, 1, \dots, m$ ), σε μοναδική λύση υπό ορισμένες συνθήκες.
- Αν η  $f(x)$  δεν είναι γραμμική συνάρτηση, η λύση βρίσκεται με παρόμοιο τρόπο. Για παράδειγμα, αν  $f(x) = \beta e^{\alpha x}$ , το άθροισμα

$$E(\alpha, \beta) = \sum_{i=1}^n (\beta e^{\alpha x_i} - y_i)^2,$$

ελαχιστοποιείται όταν  $\partial E/\partial \alpha = \partial E/\partial \beta = 0$ . Οι παραπάνω συνθήκες οδηγούν στο μη γραμμικό σύστημα

$$\sum_{i=1}^n 2(\beta e^{\alpha x_i} - y_i) \beta x_i e^{\alpha x_i} = 0,$$

$$\sum_{i=1}^n 2(\beta e^{\alpha x_i} - y_i) e^{\alpha x_i} = 0.$$

Προσέξτε, όμως, ότι αν υπολογίσουμε το λογάριθμο των δεδομένων,  $\{x_i, \ln y_i\}$ ,  $i = 1, \dots, n$ , η αντίστοιχη  $f(x)$  που τα προσεγγίζει γίνεται  $\ln \beta + \alpha x$ , η οποία είναι γραμμική και η μορφή της έχει ήδη εξεταστεί.

### 4.3 Ασκήσεις

1. Υλοποιήστε σε υποπρόγραμμα Fortran την παρεμβολή Lagrange. Αυτό θα δέχεται ως ορίσματα εισόδου δύο πίνακες  $x, y$  που θα περιέχουν τα ζεύγη σημείων  $(x, y)$  καθώς και την τιμή στην οποία θέλουμε να υπολογίζει το πολυώνυμο Lagrange· την τιμή αυτού θα την επιστρέφει.

Δημιουργήστε με ξεχωριστό πρόγραμμα ένα αρχείο “points.dat” με τα σημεία  $(x, y)$  για μια γνωστή συνάρτηση, π.χ.  $\sin x$  στο διάστημα  $[2, 4]$ , παίρνοντας 15 ισαπέχοντα σημεία στον άξονα  $x$ .

Εφαρμόστε το υποπρόγραμμα που γράψατε για να υπολογίσετε τις προσεγγιστικές τιμές της “άγνωστης” συνάρτησης του αρχείου “points.dat” σε 100 ισαπέχοντα σημεία μεταξύ των  $\min\{x_i\}$  και  $\max\{x_i\}$ .

2. Τροποποιήστε τον κώδικα που γράψατε για την προηγούμενη άσκηση ώστε να υπολογίζει και να επιστρέφει, εκτός από τις προσεγγιστικές τιμές της συνάρτησης [δηλαδή, τις τιμές του πολυωνύμου Lagrange,  $p(x)$  της (4.1)], και τις αντίστοιχες τιμές της *πρώτης παραγώγου*,  $p'(x)$ .
3. Γράψτε πρόγραμμα που να προσεγγίζει άγνωστη συνάρτηση με τη μέθοδο ελαχίστων τετραγώνων. Η συνάρτηση θα δίνεται ως ζεύγη σημείων, σε δύο πίνακες  $x, y$ . Δώστε τη δυνατότητα στο χρήστη του προγράμματος να επιλέγει την προσεγγιστική καμπύλη μεταξύ των

(α)  $y = ax + b$  (γραμμική),

(β)  $y = ax^b$  (δύναμη),

(γ)  $y = a + be^x$  (εκθετική),

(δ)  $y = a + b \ln x$  (λογαριθμική).

*Υπόδειξη:* Δημιουργήστε με ξεχωριστό κώδικα τους πίνακες  $x, y$  για μια γνωστή συνάρτηση, π.χ.  $\sin x$  στο διάστημα  $[2, 4]$ , παίρνοντας 15 ισαπέχοντα σημεία.

4. Η συνολική φωτεινή ισχύς  $P$  που εκπέμπεται από ένα μέλαν σώμα επιφάνειας  $A$  δίνεται συναρτησί της απόλυτης θερμοκρασίας του  $T$  από τη σχέση

$$P = \sigma AT^4,$$

όπου  $\sigma$  η σταθερά Stefan-Boltzmann. Πειραματικές μετρήσεις για ένα νήμα ηλεκτρικού λαμπτήρα (που θεωρούμε ότι προσεγγίζει το μέλαν σώμα) σε θερμοκρασίες 300 K–2300 K έδωσαν τις ακόλουθες τιμές

| $T$ (K) | $P$ (W) |
|---------|---------|
| 300     | 0.0013  |
| 400     | 0.0162  |
| 500     | 0.0297  |
| 600     | 0.0318  |
| 700     | 0.0484  |
| 800     | 0.0965  |
| 900     | 0.1357  |
| 1000    | 0.2947  |
| 1100    | 0.4563  |
| 1200    | 0.5398  |
| 1300    | 0.8884  |
| 1400    | 1.0031  |
| 1500    | 1.4193  |
| 1600    | 1.9052  |
| 1700    | 2.4026  |
| 1800    | 2.5031  |
| 1900    | 3.9072  |
| 2000    | 4.3156  |
| 2100    | 5.5060  |
| 2200    | 6.9044  |
| 2300    | 7.6370  |

Αν υποθέσουμε ότι η επιφάνεια του νήματος είναι  $0.05 \text{ cm}^2$ , να επαληθεύσετε από τα δεδομένα το νόμο Stefan-Boltzmann (ότι πράγματι η δύναμη στην οποία υψώνεται το  $T$  είναι 4) και να εκτιμήσετε τη σταθερά  $\sigma$ .

## Κεφάλαιο 5

# Αριθμητική Ολοκλήρωση

### 5.1 Εισαγωγή

Για τη συντριπτική πλειοψηφία των συναρτήσεων  $f(x)$  δεν υπάρχουν ή είναι πολύ δύσκολοι οι τύποι της αντιπαραγώγου της  $f(x)$ , δηλαδή της  $F(x)$  η οποία ικανοποιεί τη σχέση  $F'(x) = f(x)$ , ώστε να υπολογιστεί ακριβώς το

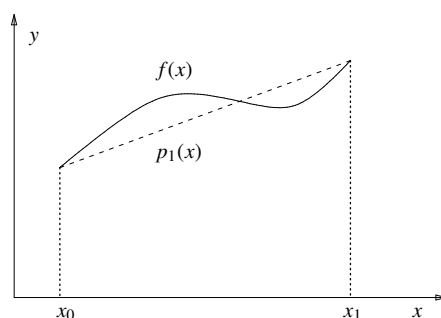
$$\int_a^b f(x) dx = F(b) - F(a) .$$

Συχνά η  $f(x)$  δεν είναι γνωστή παρά μόνο σε συγκεκριμένα σημεία. Και σε αυτήν την περίπτωση καταφεύγουμε σε αριθμητικές μεθόδους ολοκλήρωσης, σύμφωνα με τις οποίες έχουμε

$$\int_a^b f(x) dx \approx \sum_{i=1}^n c_i f(x_i) ,$$

όπου  $x_i, i = 1, 2, \dots, n$ , συγκεκριμένα σημεία στο  $[a, b]$ , και  $c_i$  προσδιοριζόμενες σταθερές.

### 5.2 Κανόνας Τραπεζίου



Σχήμα 5.1: Γραμμική προσέγγιση συνάρτησης για την εφαρμογή του τύπου ολοκλήρωσης τραπεζίου

Για ένα μικρό διάστημα  $[x_0, x_1]$  έχουμε προσεγγιστικά

$$I = \int_{x_0}^{x_1} f(x) dx \approx \frac{x_1 - x_0}{2} [f(x_0) + f(x_1)] ,$$

που αντιστοιχεί σε πρωτοβάθμιο πολυώνυμο παρεμβολής Lagrange, Σχήμα 5.1. Το σφάλμα ολοκλήρωσης είναι

$$I - \frac{x_1 - x_0}{2} [f(x_0) + f(x_1)] = -\frac{1}{12}(x_1 - x_0)^3 f''(\xi), \quad \xi \in [x_0, x_1].$$

Η επανάληψη του τύπου για πολλά διαδοχικά διαστήματα δίνει την προσεγγιστική έκφραση για το ολοκλήρωμα σε εκτεταμένο διάστημα. Έτσι, αν έχουμε χωρίσει το  $[a \equiv x_0, b \equiv x_n]$  σε  $n$  ίσα διαστήματα  $[x_i, x_{i+1}]$  με  $x_j = x_0 + jh$ ,  $j = 0, 1, \dots, n$  και  $h = \frac{b-a}{n}$ , έχουμε

$$\int_{x_0}^{x_n} f(x) dx \approx \frac{h}{2} (f_0 + 2f_1 + 2f_2 + \dots + 2f_{n-1} + f_n), \quad (5.1)$$

όπου  $f_i \equiv f(x_i)$ .

**Θεώρημα:** Αν  $f(x)$  συνεχής με δύο συνεχείς παραγώγους στο  $[a, b]$ , και

$$\max_{x \in [a, b]} |f''(x)| \leq M,$$

τότε το σφάλμα ολοκλήρωσής της,  $\varepsilon$ , με τη σύνθετη μέθοδο τραπεζίου είναι

$$|\varepsilon| \leq \frac{b-a}{12} M h^2 = \frac{nM}{12} h^3.$$

Απόδειξη:

$$\begin{aligned} E &= \int_{x_0}^{x_1} f(x) dx - \frac{h}{2} \left( f_0 + 2 \sum_{i=1}^{n-1} f_i + f_n \right) \\ &= \sum_{i=0}^{n-1} \left[ \int_{x_i}^{x_{i+1}} f(x) dx - \frac{h}{2} (f_i + f_{i+1}) \right] = \sum_{i=0}^{n-1} \varepsilon_i. \end{aligned}$$

Σε κάθε διάστημα  $[x_i, x_{i+1}]$  έχουμε:

$$\varepsilon_i = -\frac{1}{12}(x_{i+1} - x_i)^3 f''(\xi_i) = -\frac{h^3}{12} f''(\xi_i).$$

Επομένως,

$$|\varepsilon_i| \leq \frac{h^3}{12} M$$

και

$$|E| = \sum_{i=0}^{n-1} |\varepsilon_i| \leq \frac{nM}{12} h^3.$$

- Αν  $f''(x)$  συνεχής και φραγμένη στο  $[a, b]$ , τότε  $|E| \propto h^2$  καθώς  $h \rightarrow 0$ . Άρα,  $|E| \leq Ch^2$  (δεύτερης τάξης).
- Αν  $f'(x)$  είναι συνεχής στο  $[a, b]$ , τότε το σφάλμα  $|E| \propto h$  καθώς  $h \rightarrow 0$ . Άρα,  $|E| \leq Ch$  (πρώτης τάξης, αργότερη σύγκλιση).

**Παράδειγμα:** Ας υπολογίσουμε αριθμητικά το  $I = \int_0^\pi \sin x \, dx$  και να το συγκρίνουμε με την πραγματική του τιμή, 2. Έστω  $n + 1$  ισαπέχοντα σημεία στο  $[0, \pi]$ ,  $x_k = k\frac{\pi}{n}$ ,  $k = 0, 1, \dots, n$ . Τότε

$$I \approx I_n = \frac{\pi}{n} \left( \frac{\sin x_0 + \sin x_n}{2} + \sum_{i=1}^{n-1} \sin x_k \right).$$

Επομένως,

| $n$ | $I_n$      | $\varepsilon = I - I_n$ |
|-----|------------|-------------------------|
| 1   | 0.00000000 | 2.00000000              |
| 2   | 1.57079633 | 0.42920367              |
| 3   | 1.81379936 | 0.18620064              |
| 4   | 1.89611890 | 0.10388110              |
| 5   | 1.93376560 | 0.06623440              |
| 6   | 1.95409723 | 0.04590277              |
| 7   | 1.96631668 | 0.03368332              |
| 8   | 1.97423160 | 0.02576840              |
| 9   | 1.97965081 | 0.02034919              |
| 10  | 1.98352354 | 0.01647646              |
| 11  | 1.98638699 | 0.01361301              |
| 12  | 1.98856378 | 0.01143622              |
| 13  | 1.99025718 | 0.00974282              |
| 14  | 1.99160043 | 0.00839957              |
| 15  | 1.99268383 | 0.00731617              |
| 16  | 1.99357034 | 0.00642966              |
| 17  | 1.99430494 | 0.00569506              |
| 18  | 1.99492046 | 0.00507954              |
| 19  | 1.99544132 | 0.00455868              |
| 20  | 1.99588597 | 0.00411403              |

Παρατηρήστε ότι το σφάλμα τείνει στο 0 ως  $h^2$ .

Το ελάχιστο  $n$  για να έχουμε  $\varepsilon \leq 10^{-4}$ , προσδιορίζεται ως εξής:

$$|f''(x)| = |-\sin x| \leq 1 \quad \forall x \in [0, \pi].$$

Επομένως,

$$|E| \leq \frac{\pi}{12} h^2 = \frac{\pi}{12} \frac{\pi^2}{n^2} \leq 10^{-4} \Rightarrow n \geq 161.$$

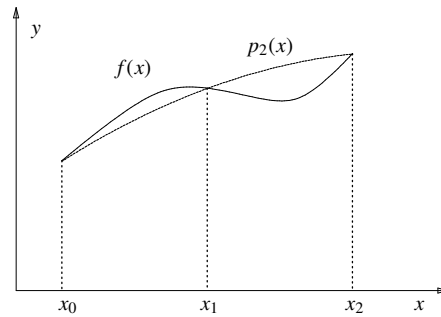
### 5.3 Κανόνας Simpson

Στη μέθοδο Simpson προσεγγίζουμε το ολοκλήρωμα χρησιμοποιώντας πολυώνυμο Lagrange δευτέρου βαθμού που παρεμβάλεται σε τρία σημεία, Σχήμα 5.2.

$$\int_{x_0}^{x_2} f(x) \, dx \approx \frac{h}{3} (f_0 + 4f_1 + f_2),$$

καθώς το  $p_2(x)$  που προσδιορίζεται από τα  $(x_0, f_0)$ ,  $(x_1, f_1)$ ,  $(x_2, f_2)$  είναι

$$p_2(x) = \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} f_0 + \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} f_1 + \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} f_2.$$



Σχήμα 5.2: Προσέγγιση συνάρτησης με παραβολή για την εφαρμογή του τύπου ολοκλήρωσης Simpson

Επομένως,

$$\begin{aligned} \int_{x_0}^{x_2} p_2(x) dx &= f_0 \int_{x_0}^{x_2} \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} dx + f_1 \int_{x_0}^{x_2} \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} dx \\ &\quad + f_2 \int_{x_0}^{x_2} \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} dx \\ &= f_0 \frac{h}{3} + f_1 \frac{4h}{3} + f_2 \frac{h}{3}. \end{aligned}$$

Αν η  $f(x)$  έχει τέσσερις συνεχείς παραγώγους στο  $[x_0, x_2]$  έχουμε για το σφάλμα:

$$\int_{x_0}^{x_2} f(x) dx - \frac{h}{3}(f_0 + 4f_1 + f_2) = -\frac{1}{90}h^5 f^{(4)}(\xi), \quad \xi \in [x_0, x_2].$$

Παρόμοια με τον τύπο τραπέζιου, μπορούμε να υπολογίσουμε το σύνθετο τύπο Simpson στο διάστημα  $[a, b]$  υποθέτοντας ότι  $b-a = 2kh$ . Προκύπτει ότι

$$\int_a^b f(x) dx \approx \frac{h}{3} \left( f_0 + f_{2k} + 4 \sum_{j=1}^k f_{2j-1} + 2 \sum_{j=1}^{k-1} f_{2j} \right). \quad (5.2)$$

Επιπλέον, αν η τέταρτη παράγωγος είναι φραγμένη στο  $[a, b]$ ,

$$\max_{x \in [a, b]} |f^{(4)}(x)| \leq M,$$

το σφάλμα  $\varepsilon$  του σύνθετου τύπου Simpson είναι

$$|\varepsilon| \leq \frac{b-a}{90} M h^4. \quad (5.3)$$

**Παράδειγμα:** Ο υπολογισμός του

$$I = \int_0^\pi \sin x dx$$

με τον σύνθετο τύπο Simpson δίνει

| $n$ | $I_n$     | $\varepsilon = I - I_n$ |
|-----|-----------|-------------------------|
| 2   | 2.0943951 | -0.0943951              |
| 4   | 2.0045598 | -0.0045598              |
| 6   | 2.0008632 | -0.0008632              |
| 8   | 2.0002692 | -0.0002692              |
| 10  | 2.0001095 | -0.0001095              |
| 12  | 2.0000526 | -0.0000526              |
| 14  | 2.0000283 | -0.0000283              |
| 16  | 2.0000166 | -0.0000166              |
| 18  | 2.0000103 | -0.0000103              |
| 20  | 2.0000068 | -0.0000068              |

Παρατηρήστε ότι για να επιτύχουμε σφάλμα κάτω από 0.005 χρειαζόμαστε 4+1 σημεία· αντίθετα, για ίδιο σφάλμα με τον τύπο τραπεζίου απαιτούνται 19+1. Γενικότερα, καθώς  $|f^{(4)}(x)| = |\sin x| \leq 1 \forall x$ , το σφάλμα συνδέεται με τον αριθμό διαστημάτων  $n$  με τη σχέση

$$|\varepsilon| \leq \frac{\pi}{90} 1 \left(\frac{\pi}{n}\right)^4 .$$

Επομένως, σφάλμα  $< 10^{-6}$  απαιτεί  $n \geq 43$ .

Η γενικευμένη μορφή Simpson για  $k + 1$  σημεία αντιστοιχεί σε πολυώνυμο παρεμβολής Lagrange βαθμού  $k$  και ονομάζεται μέθοδος Newton-Cotes κλειστού τύπου.

**Παρατήρηση:** Η τάξη ενός κανόνα ολοκλήρωσης είναι κατά 1 μεγαλύτερη από το μεγαλύτερο ακέραιο  $k$  τέτοιο ώστε ο κανόνας να δίνει το ακριβές αποτέλεσμα για την ολοκλήρωση των μονωνύμων  $1, x, x^2, \dots, x^k$  αλλά δεν ολοκληρώνει ακριβώς το  $x^{k+1}$ .

### 5.3.1 Κανόνας Simpson των $3/8$

Ο κανόνας των  $3/8$  δημιουργήθηκε με την αντικατάσταση ενός  $3^{\text{ης}}$  τάξης πολυωνύμου στο ολοκλήρωμα. Για 4 δεδομένα σημεία μπορεί να αποδειχθεί ότι η προσέγγιση του ολοκληρώματος δίνεται από τον τύπο:

$$I \approx \frac{3h}{8} [f(a) + 3f(x_1) + 3f(x_2) + f(b)] ,$$

όπου:

$a, b$  τα όρια ολοκλήρωσης,

$x_1, x_2$  τιμές του  $x$  που χωρίζουν το διάστημα  $[a, b]$  σε 3 ίσα τμήματα, και

$$h = \frac{b-a}{3} .$$

Ο τύπος του Simpson  $3/8$  έχει σφάλμα:

$$E_t = -\frac{3}{80} h^5 f^{(4)}(\xi) = -\frac{(b-a)^5}{6480} f^{(4)}(\xi) .$$

Ο τύπος των  $3/8$  είναι ελάχιστα πιο ακριβής από τον τύπο του Simpson με το  $1/3$ , παρά το γεγονός ότι χρησιμοποιεί ένα παραπάνω σημείο. Η μεγάλη χρησιμότητα του τύπου του  $3/8$  είναι ότι ο αριθμός των διαστημάτων που χρειάζεται για τον υπολογισμό είναι περιττός (και πολλαπλάσιο του 3), συνεπώς επιτρέπει την εκτίμηση με ακρίβεια τρίτης τάξης για όλο το διάστημα ολοκλήρωσης.

### 5.3.2 Κανόνας Simpson των $3/8$ για εκτεταμένο διάστημα

Αν και ο τύπος έχει περιορισμένη εφαρμογή, για διαστήματα με πλήθος  $n$  πολλαπλάσιο του 3, μπορεί να αποδειχθεί ότι:

$$I \approx \frac{3h}{8} \left( f_0 + 3 \sum_{i=0}^{k-1} f_{3i+1} + 3 \sum_{i=0}^{k-1} f_{3i+2} + 2 \sum_{i=0}^{k-2} f_{3i+3} + f_n \right),$$

όπου  $f_i = f(x_i)$ ,  $x_0 \equiv a$ ,  $x_n \equiv b$  και  $k = n/3$ .

## 5.4 Ειδικές Περιπτώσεις

### 5.4.1 Ολοκλήρωση σε άνισα τμήματα

Στην περίπτωση που η συνάρτηση  $f(x)$  δεν είναι δεδομένη, αλλά τα σημεία της δίνονται με την μορφή πινάκων (π.χ. από πειραματικές μετρήσεις), τότε είναι πιθανό τα σημεία στα οποία ορίζεται η συνάρτησή μας να μην ισαπέχουν. Τότε έχουμε δύο εναλλακτικές δυνατότητες:

- Ολοκλήρωσης του πολυωνύμου Lagrange στο διάστημα ορισμού των δεδομένων (ή ολοκλήρωσης κατά υποδιαστήματα των αντίστοιχων πολυωνύμων).
- Χρήση της μεθόδου του τραpezίου.

Η πρώτη μέθοδος, για μεγάλο αριθμό σημείων  $N$ , μπορεί να μην είναι ακριβής εξαιτίας της υψηλής τάξης πολυωνύμου που δημιουργείται. Η μέθοδος του τραpezίου μπορεί να προσαρμοστεί εύκολα ώστε να εφαρμοστεί σε κάθε διάστημα ο τύπος:

$$I = h_1 \frac{f(x_1) + f(x_2)}{2} + h_2 \frac{f(x_2) + f(x_3)}{2} + \dots + h_N \frac{f(x_{N-1}) + f(x_N)}{2}.$$

Εάν γειτονικά τμήματα είναι ίσα τότε μπορεί να εφαρμοστεί ένας τύπος Newton-Cotes (π.χ. Simpson) υψηλότερης τάξης.

### 5.4.2 Ολοκληρώματα με μη πεπερασμένα όρια ολοκλήρωσης

Ως τώρα παρουσιάστηκαν μέθοδοι για ολοκληρώματα με πεπερασμένα όρια. Σε πραγματικά προβλήματα υπάρχει περίπτωση να χρειαστεί να υπολογιστούν ανώμαλα ολοκληρώματα (που να συμπεριλαμβάνουν στο πεδίο τιμών τους το  $+\infty$  ή το  $-\infty$ ).

Για να υπολογιστούν αυτά τα ολοκληρώματα μπορούμε να κάνουμε μια κατάλληλη αλλαγή μεταβλητής, π.χ.  $x = 1/t$ . Τότε:

$$\int_a^b f(x) dx = \int_{1/b}^{1/a} \frac{1}{t^2} f\left(\frac{1}{t}\right) dt$$

για  $a, b$  ομόσημα. Συνεπώς η συγκεκριμένη αλλαγή μεταβλητής μπορεί να χρησιμοποιηθεί όταν  $a = -\infty$  και  $b < 0$  είτε  $b = +\infty$  και  $a > 0$ .

Εάν τα  $a, b$  είναι ετερόσημα, τότε μπορούμε ορίσουμε ένα σημείο  $c$ , τέτοιο ώστε να μπορεί να χρησιμοποιηθεί η παραπάνω αλλαγή μεταβλητής. Π.χ. για τον υπολογισμό της  $f(x)$  στο  $[-3, +\infty)$ , μπορούμε να κάνουμε τα ακόλουθα:

$$\int_{-3}^{+\infty} f(x) dx = \int_{-3}^{+3} f(x) dx + \int_{+3}^{+\infty} f(x) dx = \int_{-3}^{+3} f(x) dx + \int_0^{1/3} \frac{1}{t^2} f\left(\frac{1}{t}\right) dt.$$

Μερικές φορές υπάρχει περίπτωση να περιλαμβάνεται στο διάστημα ολοκλήρωσης ένα ή περισσότερα διακριτά σημεία στα οποία η συνάρτηση δεν ορίζεται. Σε τέτοια περίπτωση μπορούμε να εφαρμόσουμε τμηματικά τους ανοικτούς τύπους Newton–Cotes.

## 5.5 Μέθοδος Gauss

Αν θεωρήσουμε την προσέγγιση

$$\int_{-1}^1 f(x) dx \approx \sum_{i=1}^m c_i f(\xi_i), \quad (5.4)$$

όπου  $\xi_i$  σταθερά σημεία στο  $[-1, 1]$  και  $c_i$  συντελεστές, τότε

- για τον κανόνα τραπεζίου έχουμε  $c_1 = c_2 = 1$ ,  $\xi_1 = -1$ ,  $\xi_2 = 1$ .
- για τον κανόνα Simpson έχουμε  $c_1 = c_3 = 1/3$ ,  $c_2 = 4/3$ ,  $\xi_1 = -1$ ,  $\xi_2 = 0$ ,  $\xi_3 = 1$ .

Η ερώτηση είναι: για δεδομένο αριθμό σημείων  $m$ , ποια είναι τα  $c_i, \xi_i$   $i = 1, 2, \dots, m$  ώστε ο κανόνας (5.4) να έχει τη μέγιστη δυνατή ακρίβεια.

Έστω ότι η μέγιστη δυνατή ακρίβεια σημαίνει πως ο κανόνας δίνει το ακριβές αποτέλεσμα στην ολοκλήρωση των μονωνύμων  $1, x, x^2, \dots, x^{2m-1}$ . Αυτή η συνθήκη οδηγεί στους κανόνες ολοκλήρωσης Gauss.

**Κανόνας Gauss με  $m = 1$ .** Έχουμε

$$\int_{-1}^1 f(x) dx \approx c_1 f(\xi_1).$$

Καθώς πρέπει να προκύπτει το ακριβές αποτέλεσμα για  $f(x) = 1$ ,  $f(x) = x$  έχουμε

$$\begin{aligned} f(x) = 1 &\rightsquigarrow c_1 \cdot 1 = \int_{-1}^1 1 dx = 2 \Rightarrow c_1 = 2, \\ f(x) = x &\rightsquigarrow c_1 \xi_1 = \int_{-1}^1 x dx = 0 \Rightarrow \xi_1 = 0. \end{aligned}$$

Επομένως, ο κανόνας Gauss με ένα σημείο είναι

$$\int_{-1}^1 f(x) dx \approx 2f(0).$$

Ο κανόνας αυτός ολοκληρώνει ακριβώς τα  $1, x$  αλλά όχι το  $x^2$ .

**Κανόνας Gauss με  $m = 2$ .** Ζητώντας να παράγεται το ακριβές αποτέλεσμα για  $f(x) = 1$ ,  $f(x) = x$ ,  $f(x) = x^2$ ,  $f(x) = x^3$ , έχουμε

$$\begin{aligned} c_1 + c_2 &= 2, \\ c_1 \xi_1 + c_2 \xi_2 &= 0, \\ c_1 \xi_1^2 + c_2 \xi_2^2 &= \frac{2}{3}, \\ c_1 \xi_1^3 + c_2 \xi_2^3 &= 0. \end{aligned}$$

Η λύση του συστήματος είναι μοναδική και είναι η εξής

$$\begin{aligned}c_1 &= 1, \\c_2 &= 1, \\ \xi_1 &= -\frac{1}{\sqrt{3}}, \\ \xi_2 &= \frac{1}{\sqrt{3}}.\end{aligned}$$

Επομένως

$$\int_{-1}^1 f(x) dx \approx f\left(-\frac{1}{\sqrt{3}}\right) + f\left(\frac{1}{\sqrt{3}}\right).$$

**Κανόνας Gauss με  $m = 3$ .** Η απαίτηση για ακριβές αποτέλεσμα για  $x^k$ ,  $k = 0, \dots, 6$  σχηματίζει το μη γραμμικό σύστημα

$$\begin{aligned}c_1 + c_2 + c_3 &= 2, \\c_1\xi_1 + c_2\xi_2 + c_3\xi_3 &= 0, \\c_1\xi_1^2 + c_2\xi_2^2 + c_3\xi_3^2 &= \frac{2}{3}, \\c_1\xi_1^3 + c_2\xi_2^3 + c_3\xi_3^3 &= 0, \\c_1\xi_1^4 + c_2\xi_2^4 + c_3\xi_3^4 &= \frac{2}{5}, \\c_1\xi_1^5 + c_2\xi_2^5 + c_3\xi_3^5 &= 0.\end{aligned}$$

Η λύση του είναι

$$\begin{aligned}c_1 = \frac{5}{9}, \quad c_2 = \frac{8}{9}, \quad c_3 = \frac{5}{9}, \\ \xi_1 = -\sqrt{0.6}, \quad \xi_2 = 0, \quad \xi_3 = \sqrt{0.6}.\end{aligned}$$

### Παρατηρήσεις:

- Το σφάλμα του σύνθετου τύπου για τη μέθοδο Gauss είναι
  1. με ένα σημείο  $|\varepsilon| \propto h^2$  (όπως ο τύπος τραπεζίου με δύο σημεία).
  2. με δύο σημεία  $|\varepsilon| \propto h^4$  (όπως ο τύπος Simpson με τρία σημεία).
  3. με τρία σημεία  $|\varepsilon| \propto h^6$ .
  4. με  $m$  σημεία  $|\varepsilon| \propto h^{2m}$ .
- Τα σημεία  $\xi_i$ ,  $i = 1, \dots, m$  είναι οι ρίζες του πολυωνύμου Legendre,  $P_m(x)$ .
- Στη γενική περίπτωση που έχουμε να υπολογίσουμε το

$$I = \int_a^b f(x) dx$$

κάνουμε τον κατάλληλο γραμμικό μετασχηματισμό ώστε το διάστημα  $[a, b]$  να γίνει  $[-1, 1]$ . Θέτουμε  $\tilde{x} = \lambda x + \mu$  και ζητούμε να ισχύει  $\tilde{x} = -1$  όταν  $x = a$  και  $\tilde{x} = 1$  όταν  $x = b$ . Επομένως

$$\begin{aligned}\tilde{x} &= \frac{2}{b-a}x - \frac{b+a}{b-a}, \\ d\tilde{x} &= \frac{2}{b-a} dx.\end{aligned}$$

Έτσι

$$\begin{aligned}I = \int_a^b f(x) dx &= \frac{b-a}{2} \int_{-1}^1 f\left(\frac{b-a}{2}\tilde{x} + \frac{b+a}{2}\right) d\tilde{x} \\ &\approx \frac{b-a}{2} \sum_{i=1}^m c_i f\left(\frac{b-a}{2}\xi_i + \frac{b+a}{2}\right).\end{aligned}$$

## 5.6 Ασκήσεις

- (α) Υλοποιήστε τον αλγόριθμο τραπεζίου σε υποπρόγραμμα FORTRAN. Το υποπρόγραμμα θα δέχεται ως ορίσματα τουλάχιστον τα όρια της ολοκλήρωσης και το πλήθος των διαστημάτων. Θα επιστρέφει την προσεγγιστική τιμή του ολοκληρώματος.  
(β) Χρησιμοποιήστε το υποπρόγραμμα για να υπολογίσετε το ολοκλήρωμα

$$\int_0^{\pi} \sin x dx$$

διαδοχικά με  $N = 2, 4, 8, 16, \dots, 512$  διαστήματα. Το πρόγραμμά σας να τυπώνει για κάθε  $N$  την υπολογιζόμενη τιμή και το απόλυτο σφάλμα ως προς την ακριβή τιμή.

- (α) Υλοποιήστε τον αλγόριθμο Simpson σε υποπρόγραμμα FORTRAN. Το υποπρόγραμμα θα δέχεται ως ορίσματα τουλάχιστον τα όρια της ολοκλήρωσης και το πλήθος των διαστημάτων. Θα επιστρέφει την προσεγγιστική τιμή του ολοκληρώματος.  
(β) Χρησιμοποιήστε την για να υπολογίσετε το ολοκλήρωμα

$$\int_0^{\pi} \sin x dx$$

με όσα διαστήματα χρειάζεται ώστε να έχετε ακρίβεια τουλάχιστον 6 ψηφίων.

*Υπόδειξη:* Επιλέξτε κατάλληλα το βήμα (άρα και το πλήθος των διαστημάτων) ώστε το σφάλμα (§5.3) να είναι μικρότερο από  $10^{-6}$ .

- Γράψτε κώδικες FORTRAN που να υπολογίζουν με κάθε μία από τις μεθόδους που παρουσιάστηκαν τα ολοκληρώματα

$$(α) \int_1^2 \log x dx,$$

$$(β) \int_1^2 e^x \cos x dx,$$

$$(γ) \int_1^2 \frac{1}{x+5} dx.$$

[Σωστές τιμές: (α') 0.167766..., (β') -0.0560659..., (γ') 0.154151...]

4. Υπολογίστε προσεγγιστικά με ακρίβεια  $10^{-6}$  τα ολοκληρώματα στο διάστημα  $[0, 3]$  των συναρτήσεων

(α')  $f(x) = 2x + 1,$

(β')  $f(x) = x^2 \sqrt{x},$

(γ')  $f(x) = \frac{1}{1+x^2},$

(δ')  $f(x) = \frac{1}{1+(x-\pi)^2},$

(ε')  $f(x) = \frac{1}{2+\cos x},$

(ϛ')  $f(x) = \cos(4x)e^x,$

(ζ')  $f(x) = e^{\cos x},$

(η')  $f(x) = \sqrt{x}.$

5. Έστω

$$f(x) = \begin{cases} -x & -1 \leq x \leq 0, \\ x^2 & 0 \leq x \leq 1, \end{cases}$$

συνεχής συνάρτηση στο  $[-1, 1]$  χωρίς παράγωγο στο  $x = 0$ . Υπολογίστε το σφάλμα

$$\varepsilon_n = I - I_n = \int_{-1}^1 f(x) dx - I_n,$$

όπου  $I_n$  ο τύπος τραπεζίου με  $n$  υποδιαίρεσεις, και δείξτε ότι  $|\varepsilon_n| \leq Ch$ , όπου  $h = 2/n$ . (Υποθέστε ότι το  $n$  είναι άρτιος ή περιττός.)

6. **Μέθοδος Romberg.** Ο εκτεταμένος τύπος τραπεζίου για το ολοκλήρωμα

$$I_0 = \int_{x_0}^{x_n} f(x) dx$$

δίνει

$$I_0 = I_h + \alpha_2 h^2 + \alpha_4 h^4 + \dots, \quad (5.5)$$

όπου

$$I_h = \frac{h}{2} (f_0 + 2f_1 + 2f_2 + \dots + 2f_{n-1} + f_n),$$

$h = (x_n - x_0)/n$  και  $\alpha_i$  οι συντελεστές των όρων  $h^i$  του σφάλματος.

Γράψτε τη (5.5) για τρία διαφορετικά βήματα, π.χ.  $h, h/2, h/4$ . Παρατηρήστε ότι σχηματίζεται ένα σύστημα τριών γραμμικών εξισώσεων με αγνώστους τα  $I_0, \alpha_2, \alpha_4$ . Βρείτε τη λύση του συστήματος ως προς  $I_0$ : ο τύπος στον οποίο θα καταλήξετε—γραμμικός συνδυασμός των  $I_h, I_{h/2}, I_{h/4}$  που έχουν σφάλματα  $O(h^2)$ —δίνει την ακριβή τιμή του ολοκληρώματος με σφάλμα  $O(h^6)$ .

Υλοποιήστε σε κώδικα FORTRAN τον παραπάνω αλγόριθμο ολοκλήρωσης (μέθοδος Romberg).

7. Υλοποιήστε σε κώδικα FORTRAN μια υπορουτίνα που να υπολογίζει ολοκληρώματα ανεξάρτητα με το πλήθος των σημείων στα οποία είναι γνωστή η συνάρτηση. Αν το πλήθος των διαστημάτων είναι περιττό (και μεγαλύτερο του 3), να χρησιμοποιεί τον τύπο  $3/8$  Simpson για τα πρώτα 3 και για τα υπόλοιπα τον τύπο  $1/3$  Simpson. Αν είναι άρτιο, να χρησιμοποιεί μόνο τον  $1/3$  Simpson.
8. Υλοποιήστε σε κώδικα FORTRAN τη μέθοδο ολοκλήρωσης Gauss για 2 και για 3 σημεία. Εφαρμόστε τον για να υπολογίσετε το ολοκλήρωμα

$$\int_{2.1}^{5.2} x^3 e^{-x} dx .$$

[Σωστή τιμή ολοκληρώματος: 3.60346...]

9. **Μέθοδος Gauss-Hermite**<sup>1</sup>. Η μέθοδος ολοκλήρωσης Gauss μπορεί να επεκταθεί και στον υπολογισμό ολοκληρωμάτων ειδικής μορφής. Έτσι το

$$\int_{-\infty}^{\infty} e^{-x^2} f(x) dx \approx \sum_{i=1}^n w_i f(x_i) , \quad (5.6)$$

όπου  $x_i$  είναι οι ρίζες του πολυωνύμου Hermite τάξης  $n$ ,  $H_n(x)$ , και  $w_i$  τα αντίστοιχα βάρη, τα οποία είναι τα

$$w_i = \frac{2^{n-1} n! \sqrt{\pi}}{[n H_{n-1}(x_i)]^2} .$$

Τα πρώτα πολυώνυμα Hermite είναι τα

$$\begin{aligned} H_0(x) &= 1 \\ H_1(x) &= 2x \\ H_2(x) &= 4x^2 - 2 \\ H_3(x) &= 8x^3 - 12x \\ H_4(x) &= 16x^4 - 48x^2 + 12 \end{aligned}$$

Να γράψετε υπορουτίνα FORTRAN που να υλοποιεί τον τύπο (5.6) για  $n = 4$ . Χρησιμοποιήστε τη για να υπολογίσετε το ολοκλήρωμα

$$\int_{-\infty}^{\infty} e^{-x^2} x^2 dx .$$

Συγκρίνετε με την ακριβή τιμή ( $\sqrt{\pi}/2$ ).

10. **Μέθοδος Gauss-Laguerre**<sup>2</sup>. Η μέθοδος ολοκλήρωσης Gauss μπορεί να επεκταθεί και στον υπολογισμό ολοκληρωμάτων ειδικής μορφής. Έτσι, ισχύει ότι

$$\int_0^{\infty} e^{-x} f(x) dx \approx \sum_{i=1}^n w_i f(x_i) , \quad (5.7)$$

<sup>1</sup><http://mathworld.wolfram.com/Hermite-GaussQuadrature.html>

<sup>2</sup><http://mathworld.wolfram.com/Laguerre-GaussQuadrature.html>

όπου  $x_i$  είναι οι ρίζες του πολυωνύμου Laguerre τάξης  $n$ ,  $L_n(x)$ , και  $w_i$  τα αντίστοιχα βάρη, τα οποία είναι τα

$$w_i = \frac{x_i}{[(n+1)L_{n+1}(x_i)]^2}.$$

Τα πρώτα πολυώνυμα Laguerre είναι τα

$$\begin{aligned} L_0(x) &= 1 \\ L_1(x) &= -x + 1 \\ L_2(x) &= (x^2 - 4x + 2)/2 \\ L_3(x) &= (-x^3 + 9x^2 - 18x + 6)/6 \\ L_4(x) &= (x^4 - 16x^3 + 72x^2 - 96x + 24)/24 \\ L_5(x) &= (-x^5 + 25x^4 - 200x^3 + 600x^2 - 600x + 120)/120 \end{aligned}$$

Να γράψετε υπορουτίνα FORTRAN που να υλοποιεί τον τύπο (5.7) για  $n = 4$ . Χρησιμοποιήστε τη για να υπολογίσετε το ολοκλήρωμα

$$\int_0^{\infty} e^{-x}(x^6 - 3\sqrt{x} + 2) dx.$$

Συγκρίνετε με την ακριβή τιμή ( $6! - 3\sqrt{\pi}/2 + 2 \times 0!$ ).

Υπόδειξη: Το  $L_4(x)$  έχει τις 4 ρίζες του πραγματικές στο διάστημα  $[0, 10]$ .

11. **Μέθοδος Gauss-Chebyshev**<sup>3</sup>. Η μέθοδος ολοκλήρωσης Gauss μπορεί να επεκταθεί και στον υπολογισμό ολοκληρωμάτων ειδικής μορφής. Έτσι, ισχύει ότι

$$\int_{-1}^1 \frac{1}{\sqrt{1-x^2}} f(x) dx \approx \sum_{i=1}^n w_i f(x_i), \quad (5.8)$$

όπου  $x_i$  είναι οι ρίζες του πολυωνύμου Chebyshev πρώτου είδους τάξης  $n$ ,  $T_n(x)$ , και  $w_i$  τα αντίστοιχα βάρη, τα οποία είναι τα

$$w_i = \frac{\pi}{n}.$$

Τα πρώτα πολυώνυμα Chebyshev πρώτου είδους είναι τα

$$\begin{aligned} T_0(x) &= 1 \\ T_1(x) &= x \\ T_2(x) &= 2x^2 - 1 \\ T_3(x) &= 4x^3 - 3x \\ T_4(x) &= 8x^4 - 8x^2 + 1 \end{aligned}$$

Να γράψετε υπορουτίνα FORTRAN που να υλοποιεί τον τύπο (5.8) για  $n = 4$ . Χρησιμοποιήστε τη για να υπολογίσετε το ολοκλήρωμα

$$\int_{-1}^1 \frac{x^2 e^{-x}}{\sqrt{1-x^2}} dx.$$

<sup>3</sup><http://mathworld.wolfram.com/Chebyshev-GaussQuadrature.html>

## Κεφάλαιο 6

# Διαφορικές Εξισώσεις<sup>1</sup>

### 6.1 Γενικά

Για τη επίλυση των διαφόρων προβλημάτων υπάρχουν γενικά δύο τύποι μαθηματικών μοντέλων. 1. *Στατικά μοντέλα*, π.χ. το κυκλοφοριακό σύστημα μιας πόλης, ελαχιστοποίηση κόστους, κλπ. Για να λύσουμε τέτοια προβλήματα χρειαζόμαστε Γραμμικό Προγραμματισμό, Γραμμική Άλγεβρα, κλπ. και 2. *Δυναμικά Μοντέλα*, π.χ. η μετάδοση της θερμότητας, ταλαντώσεις δοκού, οποιαδήποτε μεταβολή ενός μεγέθους συναρτήσει του χρόνου, κλπ. Τέτοια προβλήματα γενικά περιγράφονται με *διαφορικές εξισώσεις (ΔΕ)*, συνήθειες ΔΕ και ΔΕ με μερικές παραγώγους. Για την επίλυσή τους χρειάζονται γνώσεις Γραμμικής Άλγεβρας, πρόβλημα ιδιοτιμών, θεωρία ΔΕ, κλπ.

Στο κεφάλαιο αυτό θα ασχοληθούμε με την αριθμητική λύση ΔΕ. Θα υπέθετε κανείς ότι για όλες τις ΔΕ υπάρχουν αναλυτικοί τύποι που μας δίνουν τη λύση τους. Αυτό όμως δε συμβαίνει στην πράξη· οι περισσότερες ΔΕ δεν επιδέχονται αναλυτική λύση. Εκτός όμως από αυτό, για πολλά προβλήματα δε μας ενδιαφέρει τόσο η αναλυτική λύση όσο οι αριθμητικές τιμές της σε ορισμένα σημεία. Για το λόγο αυτό οι πιο κατάλληλες μέθοδοι για τη λύση ΔΕ είναι οι αριθμητικές.

### 6.2 Εισαγωγή

Θα παραθέσουμε εδώ μερικές βασικές προτάσεις και ορισμούς από τη θεωρία των ΔΕ και από την Ανάλυση που θα χρησιμοποιήσουμε παρακάτω.

**Ορισμός:** Μια εξίσωση που περιγράφει μια σχέση μεταξύ μιας ανεξάρτητης μεταβλητής, μιας εξαρτημένης συνάρτησης και μίας ή περισσότερων παραγώγων της θα τη λέμε ΔΕ, δηλ.

$$y^{(n)} = f(x, y(x), y'(x), \dots, y^{(n-1)}(x)) . \quad (6.1)$$

Λύση της ΔΕ θα λέμε μια συνάρτηση  $\phi(x)$  παραγωγίσιμη  $n$  φορές σε ένα διάστημα, η οποία ικανοποιεί την (6.1).

$$\phi^{(n)} = f(x, \phi(x), \phi'(x), \dots, \phi^{(n-1)}(x)) .$$

Μια γενική λύση της (6.1) περιέχει  $n$  αυθαίρετες σταθερές, επομένως υπάρχει μια  $n$ -παραμετρική οικογένεια λύσεων.

---

<sup>1</sup>Το κεφάλαιο βασίζεται στις σημειώσεις του Καθηγητή Θ. Καλαμπούκη (1982).

Αν  $y(x_0), y'(x_0), \dots, y^{(n-1)}(x_0)$  είναι γνωστά για το σημείο  $x_0$ , τότε λέμε ότι έχουμε ένα πρόβλημα αρχικών τιμών.

**Ορισμός:** Θα λέμε ότι η συνάρτηση  $f(t, y)$ , συνεχής για  $t \in [a, b]$  και  $-\infty < y < \infty$  ικανοποιεί μια συνθήκη Lipschitz όταν ισχύει

$$|f(t, y_1) - f(t, y_2)| \leq L |y_1 - y_2| \quad (6.2)$$

για κάθε  $t \in [a, b]$  και  $y_1, y_2 \in \mathcal{R}$ .

**Θεώρημα:** Αν  $f(t, y)$  είναι συνεχής για  $t \in [a, b]$  και  $-\infty < y < \infty$  και ικανοποιεί μια συνθήκη Lipschitz τότε το πρόβλημα αρχικών τιμών

$$\begin{aligned} y' &= f(t, y) \\ y(0) &= y_0 \end{aligned}$$

έχει μία μοναδική λύση  $y(t)$ ,  $t \in [a, b]$ .

Στα παρακάτω προϋποθέτουμε ότι η ΔΕ πληροί όλες εκείνες τις συνθήκες που τις εξασφαλίζουν την ύπαρξη και το μονοσήμαντο της λύσης.

### 6.2.1 Διωνυμικό Ανάπτυγμα

$$(1+x)^p = 1 + \frac{p}{1!}x + \frac{p(p-1)}{2!}x^2 + \dots + \binom{p}{n}x^n + \dots, \quad |x| < 1, \quad (6.3)$$

όπου

$$\binom{p}{n} \equiv \frac{p!}{n!(p-n)!}.$$

**Θεώρημα Taylor:** Αν η  $f(x)$  έχει συνεχείς παραγώγους σ' ένα διάστημα, τότε το ανάπτυγμα της στην περιοχή ενός σημείου  $\alpha$  δίνεται από τον τύπο

$$f(\alpha+h) = f(\alpha) + \frac{h}{1!}f'(\alpha) + \frac{h^2}{2!}f''(\alpha) + \dots + \frac{h^{n-1}}{(n-1)!}f^{(n-1)}(\alpha) + R_n, \quad (6.4)$$

όπου

$$R_n = \frac{h^n}{n!}f^{(n)}(\alpha + \xi) \quad \text{για } \xi \in (0, h).$$

Για δύο διαστάσεις:

Αν η  $f(x, y)$  έχει συνεχείς μερικές παραγώγους σ' ένα διάστημα τότε

$$f(\alpha+h, \beta+k) = f(\alpha, \beta) + \frac{1}{1!}df(\alpha, \beta) + \frac{1}{2!}d^2f(\alpha, \beta) + \dots + \frac{1}{(n-1)!}d^{n-1}f(\alpha, \beta) + R'_n, \quad (6.5)$$

όπου

$$R'_n = \frac{1}{n!}d^n f(\alpha + \theta h, \beta + \theta k) \quad \text{για } \theta \in (0, 1)$$

και

$$\begin{aligned} df &= h \frac{\partial f}{\partial x} + k \frac{\partial f}{\partial y} \\ d^2 f &= h^2 \frac{\partial^2 f}{\partial x^2} + 2kh \frac{\partial^2 f}{\partial x \partial y} + k^2 \frac{\partial^2 f}{\partial y^2} \\ &\vdots \quad \quad \quad \vdots \end{aligned}$$

## 6.3 Κατηγορίες και Λύσεις Διαφορικών Εξισώσεων

### 6.3.1 Πρωτοβάθμιες ΔΕ

Η ΔΕ  $y' = f(x, y)$  είναι

**Διαχωρίσιμος** Όταν η  $f(x, y)$  μπορεί να γραφεί ως  $U(x)V(y)$ . Π.χ.

$$\frac{dy}{dx} = \lambda xy \Rightarrow \int \frac{dy}{y} = \int \lambda x dx \Rightarrow \ln y = \lambda \frac{x^2}{2} + c \Rightarrow y = A \exp\left(\frac{\lambda x^2}{2}\right).$$

Η σταθερά  $A$  υπολογίζεται από τις αρχικές συνθήκες.

**Ομογενής** Όταν ισχύει  $f(\lambda x, \lambda y) = f(x, y)$  για όλα τα κατάλληλα  $\lambda \in \mathcal{R}^*$ . Π.χ.

$$f(x, y) = \frac{x - y}{x + y}.$$

**Γραμμική πρώτου βαθμού** Όταν η  $f(x, y)$  είναι συνάρτηση της μορφής  $-P(x)y + Q(x)$ , γραμμική, δηλαδή, ως προς  $y$ . Τότε

$$\frac{dy}{dx} + P(x)y = Q(x).$$

Ολοκληρώσιμος παράγων είναι το  $\exp\left(\int P(x) dx\right)$ , ποσότητα που δεν είναι εύκολο να υπολογιστεί στην πράξη.

### 6.3.2 Δευτεροβάθμιες ΔΕ

Η διαφορική εξίσωση

$$\alpha y'' + \beta y' + \gamma y = f(x) \tag{6.6}$$

χαρακτηρίζεται ως γραμμική δεύτερου βαθμού με σταθερούς συντελεστές. Αν η  $f(x) \neq 0$  τότε η (6.6) λέγεται *μη ομογενής*. Η αντίστοιχη ομογενής είναι η

$$\alpha y'' + \beta y' + \gamma y = 0. \tag{6.7}$$

**Θεώρημα:** Αν  $y = G(x)$  είναι η γενική λύση της (6.7) και  $y = y_1(x)$  είναι μια μερική λύση της (6.6), τότε η  $y = G(x) + y_1(x)$  είναι η γενική λύση της (6.6).

**Γενική λύση της (6.7)** Αναζητούμε λύσεις της μορφής  $y = e^{mx}$  με  $m$  σταθερό, οπότε έχουμε

$$\begin{aligned} \alpha m^2 e^{mx} + \beta m e^{mx} + \gamma e^{mx} &= 0, \quad \text{ή} \\ \alpha m^2 + \beta m + \gamma &= 0. \end{aligned}$$

Η τελευταία εξίσωση λέγεται *χαρακτηριστική εξίσωση* της (6.7).

Για τη λύση της (6.6) διακρίνουμε λοιπόν τρεις περιπτώσεις.

1. Η χαρακτηριστική εξίσωση έχει δύο πραγματικές ρίζες  $m_1, m_2$ . Η γενική λύση της (6.7) είναι

$$y = c_1 \exp(m_1 x) + c_2 \exp(m_2 x).$$

**Παράδειγμα :**

$$\frac{d^2y}{dx^2} - \lambda^2 y = 0 .$$

Για  $y = e^{mx}$  έχουμε  $m^2 - \lambda^2 = 0$  ή  $m = \pm \lambda$  και η λύση θα είναι

$$y = c_1 e^{\lambda x} + c_2 e^{-\lambda x} .$$

2. Η χαρακτηριστική εξίσωση έχει μία διπλή ρίζα  $m_1 = m_2 = m$ . Τότε η  $e^{mx}$  είναι μία λύση της (6.7). Μια άλλη λύση, γραμμικά ανεξάρτητη από την πρώτη, είναι η  $x e^{mx}$ . Η γενική λύση της (6.7) είναι

$$y = c_1 e^{mx} + c_2 x e^{mx} .$$

3. Η χαρακτηριστική εξίσωση έχει μιγαδικές ρίζες  $\alpha \pm i\beta$ . Σε αυτήν την περίπτωση οι δύο λύσεις είναι  $e^{(\alpha \pm i\beta)x}$ . Κατάλληλοι γραμμικοί συνδυασμοί τους δίνουν δύο άλλες, γραμμικά ανεξάρτητες, πραγματικές λύσεις

$$e^{(\alpha+i\beta)x} + e^{(\alpha-i\beta)x} \propto e^{\alpha x} \cos \beta x$$

και

$$e^{(\alpha+i\beta)x} - e^{(\alpha-i\beta)x} \propto e^{\alpha x} \sin \beta x .$$

Η γενική λύση της (6.7) είναι

$$y = c_1 e^{\alpha x} \cos \beta x + c_2 e^{\alpha x} \sin \beta x .$$

**Παράδειγμα :**

$$\frac{d^2y}{dx^2} + \lambda^2 y = 0 .$$

Για  $y = e^{mx}$  έχουμε  $m^2 + \lambda^2 = 0$  ή  $m = \pm i\lambda$ . Η λύση θα είναι

$$y = c_1 \cos \lambda x + c_2 \sin \lambda x .$$

**6.3.3 Σύστημα πρωτοβάθμιων ΔΕ με σταθερούς συντελεστές**

Ένα σύστημα ΔΕ πρώτου βαθμού με σταθερούς συντελεστές μπορεί να γραφεί σε διανυσματική μορφή ως εξής

$$\dot{q} \equiv \frac{dq}{dt} = Aq , \quad q(t=0) = q_0 , \quad (6.8)$$

όπου  $A$  ένας πραγματικός πίνακας με σταθερά στοιχεία. Η λύση της (6.8) είναι

$$q = q_0 e^{At}$$

όπου

$$e^{At} = I + At + \frac{1}{2!}(At)^2 + \frac{1}{3!}(At)^3 + \dots$$

Θα δώσουμε παρακάτω μια μέθοδο επίλυσης της (6.8) με τη βοήθεια του προβλήματος ιδιοτιμών.

Έστω  $P^{-1}AP = \Lambda$  όπου  $\Lambda$  είναι ο διαγώνιος πίνακας με στοιχεία τις ιδιοτιμές του  $A$  και  $P$  ο πίνακας με στήλες τα ιδιοδιανύσματα. Πολλαπλασιάζουμε από αριστερά την (6.8) με  $P^{-1}$  οπότε έχουμε

$$\frac{dP^{-1}q}{dt} = P^{-1}APP^{-1}q. \quad (6.9)$$

Αν  $z(t) = P^{-1}q(t)$  τότε από την (6.9) έχουμε

$$\dot{z} \equiv \frac{dz}{dt} = \Lambda z,$$

και η λύση της δίνεται από τον τύπο  $z = z_0 e^{\Lambda t}$  ή  $z_i = (z_0)_i \exp(\lambda_i t)$ . Από τη σχέση  $q = Pz$  βρίσκουμε τη λύση  $q(t)$  της (6.8).

**Παράδειγμα:** Έχουμε το ακόλουθο το σύστημα διαφορικών εξισώσεων

$$\begin{aligned} \frac{dy_1}{dt} &= 2y_1 + 3y_2, \\ \frac{dy_2}{dt} &= -4y_1 - 5y_2, \end{aligned}$$

με αρχικές συνθήκες  $y_1(0) = 1, y_2(0) = 0$ .

Το σύστημα αυτό γράφεται στη μορφή  $\dot{y} = Ay$  όπου

$$A = \begin{bmatrix} 2 & 3 \\ -4 & -5 \end{bmatrix}$$

Οι ιδιοτιμές του  $A$  είναι  $-1$  και  $-2$ , με αντίστοιχα ιδιοδιανύσματα  $[1, -1]^T$  και  $[-0.75, 1]^T$ . Οι  $P, P^{-1}$  επομένως είναι

$$P = \begin{bmatrix} 1 & -0.75 \\ -1 & 1 \end{bmatrix}, \quad P^{-1} = \begin{bmatrix} 4 & 3 \\ 4 & 4 \end{bmatrix}.$$

Άρα

$$\dot{z} = P^{-1}y_0 = \begin{bmatrix} 4 & 3 \\ 4 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix},$$

οπότε

$$z = \begin{bmatrix} 4e^{-t} \\ 4e^{-2t} \end{bmatrix},$$

και

$$y = Pz = \begin{bmatrix} 1 & -0.75 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 4e^{-t} \\ 4e^{-2t} \end{bmatrix} = \begin{bmatrix} 4e^{-t} - 3e^{-2t} \\ -4e^{-t} + 4e^{-2t} \end{bmatrix}.$$

## 6.4 Μέθοδος Σειράς Taylor

Έστω το πρόβλημα αρχικών τιμών

$$y' = f(x, y), \quad x \in [a, b] \quad (6.10\alpha)$$

$$y(a) = y_0. \quad (6.10\beta)$$

Το πρώτο βήμα σε μία αριθμητική μέθοδο επίλυσης του (6.10α) είναι η διαμέριση του διαστήματος  $[a, b]$  σε πεπερασμένο αριθμό υποδιαστημάτων

$$a \equiv x_0 < x_1 < \dots < x_n \equiv b .$$

Κατόπιν, ως εύρεση αριθμητικής λύσης του (6.10α) θα εννοούμε τον προσδιορισμό των τιμών της  $y(x)$  στα σημεία  $x_i$ ,  $i = 1, 2, \dots, n$ . Για ευκολία, έστω ότι τα σημεία είναι ισαπέχοντα. Η απόσταση διαδοχικών σημείων είναι  $h = \frac{b-a}{n}$  όπου  $n$  ο αριθμός των υποδιαστημάτων και  $x_k = a + kh$ ,  $k = 0, 1, \dots, n$ . Στα επόμενα θα συμβολίζουμε με  $y(x_k)$  την ακριβή λύση στο  $x_k$  και με  $y_k$  την αντίστοιχη προσεγγιστική, όπως θα προκύπτει από μία αριθμητική μέθοδο.

Θα υπολογίσουμε τη λύση του (6.10α) χρησιμοποιώντας το ανάπτυγμα Taylor της  $y$  στο  $x_r$ . Έστω ότι η τιμή της λύσης στο  $x_r$  είναι γνωστή. Τότε

$$y(x_{r+1}) = y(x_r) + hy'(x_r) + \frac{h^2}{2!}y''(x_r) + \dots + \frac{h^{m-1}}{(m-1)!}y^{(m-1)}(x_r) + R_m \quad (6.11)$$

όπου

$$R_m = \frac{h^m}{m!}y^{(m)}(\xi) \quad \xi \in (x_r, x_{r+1}) .$$

Ισχύουν τα ακόλουθα

$$\begin{aligned} y' &= f(x, y) \equiv f^{[0]} \\ y'' &= \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \frac{dy}{dx} = f_x + f_y f \equiv f^{[1]} \\ y''' &= f_{xx} + f_{xy}f + f(f_{yx} + f_{yy}f) + f_y(f_x + f_y f) = f_x^{[1]} + f_y^{[1]}f \equiv f^{[2]} . \end{aligned}$$

Αποδεικνύεται ότι ισχύει ο αναδρομικός τύπος

$$f^{[r+1]} = f_x^{[r]} + f_y^{[r]}f, \quad r = 0, 1, \dots . \quad (6.12)$$

Επομένως

$$y^{(m)} = f^{[m-1]} . \quad (6.13)$$

Από τις δύο προηγούμενες σχέσεις μπορούμε να υπολογίσουμε παραγώγους υψηλού βαθμού. Όταν έχουμε υπολογίσει ένα ικανό αριθμό παραγώγων τότε μπορούμε να υπολογίσουμε την τιμή της λύσης στο σημείο  $x_r + h \equiv x_{r+1}$  από την (6.11). Με τον ίδιο τρόπο προχωράμε από το  $x_r + h$  στο  $x_r + 2h$  κλπ. μέχρι ότου καλύψουμε όλο το διάστημα  $[a, b]$ .

**Παράδειγμα:** Έστω το πρόβλημα αρχικών τιμών

$$y' = \lambda y, \quad y(0) = 1 .$$

Η αναλυτική λύση είναι  $y = e^{\lambda x}$ .

Ισχύουν

$$f^{[0]} = \lambda y, \quad f^{[1]} = \lambda^2 y, \quad f^{[2]} = \lambda^3 y, \dots$$

και

$$y(0+h) = y(h) = 1 + \lambda h + \frac{1}{2!}(\lambda h)^2 + \frac{1}{3!}(\lambda h)^3 + \dots = e^{\lambda h} .$$

Γενικά, η  $y^{(m)}(x)$  εξαρτάται από όλες τις μερικές παραγώγους της  $f$  βαθμού  $\leq m-1$ . Επομένως, όταν υπολογίζουμε τη σειρά Taylor χρησιμοποιώντας όρους μέχρι και  $m$  βαθμού, θα πρέπει να

υπολογίσουμε  $m$  παραγώγους βαθμού  $m - 1$ ,  $m - 1$  παραγώγους βαθμού  $m - 2$  κλπ., χρειάζεται, δηλαδή, ο υπολογισμός  $\frac{m(m+1)}{2}$  συναρτήσεων. Αν η  $f$  είναι πολύπλοκη συνάρτηση, όπως είναι συνήθως, τότε ο υπολογισμός όλων αυτών των παραγώγων είναι μια επίπονη και χρονοβόρα δουλειά. Άρα, αντί να χρησιμοποιήσουμε μια υψηλού βαθμού σειρά Taylor για τον υπολογισμό του  $y(x)$  σε ένα σχετικά μεγάλο διάστημα, είναι προτιμότερο να διαιρέσουμε το διάστημα  $[a, b]$  σε μικρά τμήματα και να χρησιμοποιήσουμε μια σειρά Taylor μικρότερου βαθμού.

#### Αλγόριθμος μεθόδου Taylor βαθμού $m - 1$ :

1. Εκλογή κατάλληλου βήματος  $h = \frac{b-a}{n}$ .
2.  $y(a) = y_0$ .
3. Για  $r = 0, 1, \dots, n$  θέτουμε

$$y_{r+1} = y_r + hf_r^{[0]} + \dots + \frac{h^{m-1}}{(m-1)!} f_r^{[m-2]}, \quad (6.14)$$

όπου τα  $f_r^{[i]}$  θα υπολογιστούν από την (6.12).

#### 6.4.1 Μέθοδος Euler

Η απλούστερη από τις μεθόδους Taylor είναι η μέθοδος Euler ( $m = 1$ ) που προκύπτει από την (6.11) αν αποκόψουμε τους όρους της σειράς μετά το δεύτερο όρο, δηλαδή,

$$y_{r+1} = y_r + hf(x_r, y_r). \quad (6.15)$$

Η μέθοδος Euler είναι πολύ εύκολο να προγραμματιστεί. Σε κάθε βήμα υπολογίζουμε το  $f(x_r, y_r)$ , το οποίο χρησιμοποιούμε για τον υπολογισμό του  $y_{r+1}$  κλπ.

**Παράδειγμα:** Έστω  $y' = -y$ ,  $y(0) = 1$ ,  $x \in [0, 1]$ , της οποίας η αναλυτική λύση είναι  $y(x) = e^{-x}$ . Η μέθοδος Euler δίνει

$$y_{r+1} = (1 - h)y_r.$$

Για  $h = 0.1$  δίνουμε στον παρακάτω πίνακα, μερικές τιμές της λύσης και τις αντίστοιχες ακριβείς.

| $x$ | $y$    | $y(x)$ |
|-----|--------|--------|
| 0.0 | 1.0000 | 1.0000 |
| 0.1 | 0.9000 | 0.9048 |
| 0.2 | 0.8100 | 0.8187 |
| 0.3 | 0.7290 | 0.7408 |
| 0.4 | 0.6561 | 0.6703 |
| 0.5 | 0.5905 | 0.6065 |
| 0.6 | 0.5314 | 0.5488 |
| 0.7 | 0.4783 | 0.4966 |
| 0.8 | 0.4305 | 0.4493 |
| 0.9 | 0.3874 | 0.4066 |
| 1.0 | 0.3486 | 0.3679 |

### 6.4.2 Σφάλμα Μεθόδου Taylor

Το σφάλμα στη μέθοδο Taylor προέρχεται από δύο πηγές: 1. *Σφάλμα αποκοπής*, που προέρχεται από την προσέγγιση (6.14) (global discretization error ή global truncation error) και 2. το *σφάλμα στρογγύλευσης*, που οφείλεται στο πεπερασμένο της αναπαράστασης των πραγματικών αριθμών από τον ΗΥ. Στα επόμενα αγνοούμε το σφάλμα στρογγύλευσης και θεωρούμε ότι όλες οι πράξεις γίνονται χωρίς σφάλμα.

Ορίζουμε ως

$$E(h) = \max_{1 \leq r \leq n} |y_r - y_x(r)|, \quad (6.16)$$

το *ολικό σφάλμα αποκοπής* ( $E(h) \rightarrow 0$  όταν  $h \rightarrow 0$ ), και

$$L(x, h) = \frac{1}{h} [y(x+h) - y(x)] - f(x, y(x)), \quad (6.17)$$

το *τοπικό σφάλμα αποκοπής*, που χαρακτηρίζει τη διαφοροποίηση της προσεγγιστικής παραγώγου  $y'$  από την ακριβή τιμή.

Υποθέτουμε ότι η  $y''$  είναι φραγμένη στο  $[a, b]$ , δηλαδή

$$\max_{a \leq x \leq b} |y''(x)| = M,$$

και ότι  $y = y(x_r)$  για κάποιο  $r$ . Τότε

$$y(x_{r+1}) - y_{r+1} = y(x_{r+1}) - y(x_r) = hf(x_r, y(x_r)) = hL(x_r, h),$$

δηλαδή το σφάλμα σε ένα βήμα στη μέθοδο Euler είναι  $h$  φορές το τοπικό σφάλμα αποκοπής αρχίζοντας από την αρχική λύση.

Ενδιαφερόμαστε για το  $\max L(x, h)$  για κάθε τιμή του  $x$ , συνεπώς ορίζουμε ως *τοπικό σφάλμα αποκοπής για τη μέθοδο Euler* το

$$L(h) = \max_{a \leq x \leq b-h} |L(x, h)|.$$

Από το ανάπτυγμα Taylor, (6.11), και την (6.4.2) έχουμε ότι

$$L(h) \leq \frac{h}{2} M = O(h).$$

Με  $O(h)$  υποδηλώνουμε ότι η ποσότητα τείνει στο 0 ανάλογα με το  $h$ .

Το πρόβλημά μας είναι να βρούμε μια σχέση μεταξύ του τοπικού και του ολικού σφάλματος αποκοπής. Αν θέσουμε  $z_r = y(x_r) - y_r$  τότε

$$\begin{aligned} z_{r+1} &\equiv y(x_{r+1}) - y_{r+1} = y(x_r) + hf(x_r, y(x_r)) + hL(x_r, h) - y_r - hf(x_r, y_r) \\ &= z_r + h[f(x_r, y(x_r)) - f(x_r, y_r)] + hL(x_r, h). \end{aligned} \quad (6.18)$$

Έστω ότι

$$\left| \frac{\partial f(x, y)}{\partial y} \right| \leq M_1 \quad x \in [a, b], \quad |y| < \infty.$$

Τότε, από το Θεώρημα μέσης τιμής για  $0 < \theta < 1$  έχουμε

$$|f(x_r, y(x_r)) - f(x_r, y_r)| = \left| \frac{\partial f}{\partial y}(x_r, \theta y(x_r) + (1-\theta)y_r)(y(x_r) - y_r) \right| \leq M_1 z_r. \quad (6.19)$$

Από τις (6.19) και (6.18) έχουμε ότι

$$|z_{r+1}| \leq (1 + hM_1)|z_r| + h|L(h)| .$$

Αν  $c = 1 + hM_1$  τότε

$$\begin{aligned} |z_{r+1}| &\leq c|z_r| + hL(h) \\ &\leq c^2|z_{r-1}| + ch|L(h)| + h|L(h)| \\ &\leq \dots \\ &\leq c^r|z_1| + c^{r-1}h|L(h)| + \dots + ch|L(h)| + h|L(h)| . \end{aligned}$$

Το φράγμα αυτό περιέχει το άθροισμα  $n$  όρων της μορφής  $O(h^2)$  (γιατί;), και καθώς  $n = \frac{b-a}{h}$ , το άθροισμα είναι της μορφής  $O(h)$ . Συνεπώς, αν η συνάρτηση  $f$  έχει φραγμένη μερική παράγωγο ως προς  $y$  και αν η λύση του προβλήματος αρχικών τιμών έχει φραγμένη δεύτερη παράγωγο, τότε το ολικό σφάλμα αποκοπής στη μέθοδο Euler ικανοποιεί το  $E(h) \propto h$  (Η μέθοδος Euler είναι πρώτου βαθμού). Για την ολοκλήρωση της απόδειξης χρειάζεται να δειχτεί ότι το  $c^n \equiv (1 + hM_1)^n$  είναι φραγμένο όταν  $h \rightarrow 0$ .

## 6.5 Μέθοδος Runge–Kutta

Μια επαναληπτική μέθοδος θα λέμε ότι είναι βαθμού  $p$  όταν το τοπικό σφάλμα αποκοπής μπορεί γενικά να γραφεί ως

$$R(x, h) = ch^{p+1}y^{(p+1)}(\xi) , \quad \xi \in (x, x + h) .$$

Όπως είδαμε, το τοπικό σφάλμα αποκοπής στη μέθοδο του Euler είναι ανάλογο του  $h^2$ , επομένως, για να έχουμε καλή ακρίβεια, θα πρέπει το  $h$  να είναι πολύ μικρό. Αυτό είναι και το βασικότερο μειονέκτημα της μεθόδου του Euler γιατί όταν το  $h$  μικραίνει τότε αυξάνει το σφάλμα στρογγύλευσης. Αντίθετα, στη μέθοδο Euler  $n$  βαθμού, το τοπικό σφάλμα είναι ανάλογο του  $h^{n+1}$ , οπότε μπορούμε να χρησιμοποιήσουμε μεγαλύτερο βήμα  $h$ . Το μειονέκτημα είναι ότι θα πρέπει να υπολογίσουμε υψηλές παραγώγους, βαθμού  $n - 1$ , πράγμα που είναι δύσκολο στην πράξη. Η μέθοδος Runge–Kutta (RK) που θα περιγράψουμε παρακάτω ξεπερνά όλες τις δυσκολίες της μεθόδου Taylor (υψηλές παράγωγοι) και της Euler (μικρό  $h$ ). Η ιδέα της μεθόδου RK είναι να δώσουμε στο  $y_{r+1} = y(x_r + h)$  μια μορφή η οποία να ταυτίζεται με το ανάπτυγμα Taylor για ένα συγκεκριμένο αριθμό όρων της σειράς (βαθμός του  $h$ ) χωρίς, όμως, να υπολογίζουμε στην πράξη τις παραγώγους της  $f(x, y)$  όπως απαιτείται στη μέθοδο Taylor.

Έστω το πρόβλημα αρχικών τιμών

$$y' = f(x, y) \tag{6.20α}$$

$$y(x_0) = y_0 . \tag{6.20β}$$

Η μέθοδος RK ουσιαστικά αντικαθιστά την προσέγγιση της  $y_{r+1}$  από τη σειρά Taylor

$$y_{r+1} = y_r + hy'_r + \frac{h^2}{2!}y''_r + \frac{h^3}{3!}y'''_r + \dots \tag{6.21}$$

με τη σχέση

$$\begin{aligned} y_{r+1} = & y_r + [\omega_1 f(x_r, y_r) + \omega_2 f(x_r + c_2h, y_r + d_2h) + \omega_3 f(x_r + c_3h, y_r + d_3h) \\ & + \dots + \omega_p f(x_r + c_ph, y_r + d_ph)] . \end{aligned} \tag{6.22}$$

Οι σταθερές  $\omega_i$ ,  $c_i$ ,  $d_i$  στην (6.22) υπολογίζονται κατά τέτοιο τρόπο ώστε όταν αναπτύξουμε στο δεύτερο μέλος της εξίσωσης τις συναρτήσεις  $f$  κατά Taylor και εξισώσουμε με την (6.21) τότε οι συντελεστές των ομοιοβαθμων όρων να συμπίπτουν. Για να απλοποιήσουμε τις πράξεις και να συστηματοποιήσουμε τη μέθοδο, εκφράζουμε τα  $d_i$  ως γραμμικούς συνδυασμούς των προηγούμενων τιμών της  $f$ . Δηλαδή, αντί της (6.22) γράφουμε

$$y_{r+1} = y_r + \omega_1 k_1 + \omega_2 k_2 + \cdots + \omega_p k_p \quad (6.23)$$

με

$$k_1 = hf(x_r, y_r) \quad (6.24\alpha)$$

$$k_2 = hf(x_r + c_2 h, y_r + a_{21} k_1) \quad (6.24\beta)$$

$$k_3 = hf(x_r + c_3 h, y_r + a_{31} k_1 + a_{32} k_2) \quad (6.24\gamma)$$

$$\vdots \quad \vdots \quad (6.24\delta)$$

$$k_p = hf(x_r + c_p h, y_r + a_{p1} k_1 + a_{p2} k_2 + \cdots + a_{p(p-1)} k_{p-1}), \quad (6.24\epsilon)$$

όπου οι σταθερές  $\omega_i$ ,  $c_i$ ,  $a_{ij}$  θα πρέπει να υπολογιστούν. Παρακάτω θα εξετάσουμε την περίπτωση για  $p = 2$  που είναι η απλούστερη.

### 6.5.1 Μέθοδος Runge–Kutta 2<sup>ου</sup> βαθμού

Για  $p = 2$  οι (6.23, 6.24α') γίνονται

$$y_{r+1} = y_r + \omega_1 k_1 + \omega_2 k_2 \quad (6.25\alpha)$$

$$k_1 = hf(x_r, y_r) \quad (6.25\beta)$$

$$k_2 = hf(x_r + c_2 h, y_r + a_{21} k_1). \quad (6.25\gamma)$$

Αναπτύσσουμε την  $f(x_r + c_2 h, y_r + a_{21} k_1)$  κατά Taylor:

$$k_2 = h \left[ f + c_2 h f_x + a_{21} k_1 f_y + \frac{1}{2!} \left( (c_2 h)^2 f_{xx} + 2c_2 h a_{21} k_1 f_{xy} + (a_{21} k_1)^2 f_{yy} + \cdots \right) \right],$$

όπου  $f \equiv f(x_r, y_r)$ ,  $f_x \equiv \frac{\partial f}{\partial x}$ ,  $f_y \equiv \frac{\partial f}{\partial y}$ ,  $f_{xx} \equiv \frac{\partial^2 f}{\partial x^2}$ ,  $f_{xy} \equiv \frac{\partial^2 f}{\partial x \partial y}$ ,  $f_{yy} \equiv \frac{\partial^2 f}{\partial y^2}$ .

Αντικαθιστούμε τα  $k_1, k_2$  στην (6.25α')

$$\begin{aligned} y_{r+1} &= y_r + \omega_1 h f \\ &\quad + \omega_2 h \left[ f + c_2 h f_x + a_{21} k_1 f_y + \frac{1}{2!} \left( (c_2 h)^2 f_{xx} + 2c_2 h a_{21} k_1 f_{xy} + (a_{21} k_1)^2 f_{yy} + \cdots \right) \right] \\ &= y_r + (\omega_1 + \omega_2) h f + \omega_2 c_2 h^2 f_x + \omega_2 a_{21} h^2 f_y \\ &\quad + \frac{\omega_2 h^3}{2!} \left( c_2^2 f_{xx} + 2c_2 a_{21} f f_{xy} + a_{21}^2 f^2 f_{yy} \right) + \cdots \end{aligned} \quad (6.26)$$

Το ανάπτυγμα της  $y$  κατά Taylor στο  $x_r$  δίνεται από την (6.21). Ισχύουν

$$y' = f, \quad (6.27\alpha)$$

$$y'' = f_x + f_y f, \quad (6.27\beta)$$

$$y''' = f_{xx} + 2f f_{xy} + f^2 f_{yy} + f_y (f_x + f f_y). \quad (6.27\gamma)$$

Αντικαθιστώντας τις (6.27α') στην (6.21) έχουμε:

$$y_{r+1} = y_r + hf + \frac{h^2}{2!}(f_x + f_y f) + \frac{h^3}{3!} [f_{xx} + 2ff_{xy} + f^2 f_{yy} + f_y(f_x + ff_y)] + \dots \quad (6.28)$$

Αν στις σχέσεις (6.26,6.28) εξισώσουμε τους ομοιόβαθμους όρους μέχρι και δεύτερου βαθμού ως προς  $h$  έχουμε το σύστημα

$$\begin{aligned} \omega_1 + \omega_2 &= 1, \\ \omega_2 c_2 &= \frac{1}{2}, \\ \omega_2 a_{21} &= \frac{1}{2}. \end{aligned}$$

Το σύστημα αυτό είναι τριών εξισώσεων με τέσσερις αγνώστους, δηλαδή έχουμε ένα βαθμό ελευθερίας (και άρα άπειρες λύσεις). Αν  $\omega_2 = \alpha \neq 0$  με  $\alpha$  σταθερά, τότε  $\omega_1 = 1 - \alpha$ ,  $c_2 = a_{21} = \frac{1}{2\alpha}$ . Για τις τιμές αυτές των  $\omega_1$ ,  $\omega_2$ ,  $c_2$ ,  $a_{21}$  η (6.26) γίνεται

$$y_{r+1} = y_r + hf + \frac{h^2}{2} (f_x + ff_y) + \frac{h^3}{8\alpha} (f_{xx} + 2ff_{xy} + f^2 f_{yy}) + O(h^4). \quad (6.29)$$

Το τοπικό σφάλμα αποκοπής της μεθόδου θα είναι (6.28)–(6.29) δηλαδή

$$\left( \frac{h^3}{6} - \frac{h^3}{8\alpha} \right) (f_{xx} + 2ff_{xy} + f^2 f_{yy}) + \frac{h^3}{6} f_y (f_x + ff_y) + O(h^4). \quad (6.30)$$

Από την (6.30) βλέπουμε ότι δεν μπορούμε να καθορίσουμε την αυθαίρετη σταθερά  $\alpha$  έτσι ώστε το τοπικό σφάλμα αποκοπής να είναι ανάλογο του  $h^4$ . Μια απλή μορφή στους σχετικούς τύπους της μεθόδου RK δίνει η επιλογή  $\alpha = 1/2$  οπότε  $\omega_1 = \omega_2 = 1/2$ ,  $c_2 = a_{21} = 1$  με αντίστοιχους τύπους δεύτερου βαθμού

$$\begin{aligned} y_{r+1} &= y_r + \frac{1}{2}(k_1 + k_2), \\ k_1 &= hf(x_r, y_r) \\ k_2 &= hf(x_r + h, y_r + k_1). \end{aligned}$$

Το τοπικό σφάλμα αποκοπής για τη συγκεκριμένη επιλογή γίνεται

$$-\frac{h^3}{12} (f_{xx} + 2ff_{xy} + f^2 f_{yy} + 2f_x f_y + 2ff_y^2) + O(h^4).$$

Παρατηρούμε ότι το τοπικό σφάλμα αποκοπής της RK ( $p = 2$ ) είναι ανάλογο του  $h^3$ , σε αντίθεση με τη μέθοδο Euler στην οποία είναι ανάλογο του  $h^2$ . Συνεπώς, μπορούμε να χρησιμοποιήσουμε μεγαλύτερο βήμα  $h$ .

### 6.5.2 Μέθοδος Runge–Kutta 4<sup>ου</sup> βαθμού

Με το ίδιο τρόπο όπως για τη RK δεύτερου βαθμού μπορούμε να κατασκευάσουμε μια RK υψηλότερου βαθμού. Σε μια RK  $n$  βαθμού το τοπικό σφάλμα αποκοπής είναι ανάλογο του  $h^{n+1}$ .

Η πιο εύχρηστη είναι η RK τετάρτου βαθμού :

$$\begin{aligned} y_{r+1} &= y_r + \frac{1}{6} (k_1 + 2k_2 + 2k_3 + k_4) , \\ k_1 &= hf(x_r, y_r) , \\ k_2 &= hf\left(x_r + \frac{h}{2}, y_r + \frac{k_1}{2}\right) , \\ k_3 &= hf\left(x_r + \frac{h}{2}, y_r + \frac{k_2}{2}\right) , \\ k_4 &= hf(x_r + h, y_r + k_3) . \end{aligned}$$

Στη περίπτωση αυτή το σφάλμα είναι ανάλογο του  $h^5$  αλλά χρειάζονται τέσσερις υπολογισμοί της  $f$  σε κάθε επανάληψη.

### 6.5.3 Σχόλια

**Πλεονεκτήματα RK** Η μέθοδος γράφεται εύκολα σε κώδικα, είναι οικονομική σε μνήμη, είναι εύκολο να ξεκινήσει (χρειάζεται μία αρχική τιμή) και το μήκος του βήματος  $h$  μπορεί εύκολα να μεταβληθεί από επανάληψη σε επανάληψη χωρίς επιπλέον δουλειά.

**Μειονεκτήματα RK** Είναι δύσκολο να υπολογίσουμε το σφάλμα αποκοπής. Ο υπολογισμός ενός φράγματος για το σφάλμα εξαρτάται από παράγοντες οι οποίοι δεν παρουσιάζονται απ' ευθείας στη μέθοδο και συνεπώς απαιτεί επιπλέον υπολογισμούς. Επίσης, μια RK βαθμού  $n$  απαιτεί  $n$  υπολογισμούς της συνάρτησης  $f$  όταν αυτή είναι πολύπλοκη, όπως συνήθως, χρειάζεται σεβαστός αριθμός πράξεων σε κάθε επανάληψη.

## 6.6 Τελεστές Διαφορών

Ο σκοπός αυτής της παραγράφου είναι διπλός: πρώτον, για να δούμε τη δύναμη και την απλότητα των μεθόδων με τελεστές στην κατασκευή διαφορών τύπων που είναι χρήσιμοι σε πολλά θέματα της Αριθμητικής Ανάλυσης, όπως στις ΔΕ. Δεύτερον, για να εξαγάγουμε μερικούς από τους τελεστές που θα χρησιμοποιήσουμε παρακάτω.

Αν  $f(x)$  είναι πραγματική συνάρτηση, ορισμένη σε ένα διάστημα  $[a, b]$ , χωρίζουμε το διάστημα αυτό σε  $n$  ίσα τμήματα μήκους  $h$  το καθένα, δηλαδή,  $[a \equiv x_0, x_1], [x_1, x_2], \dots, [x_{n-1}, x_n \equiv b]$ . Τότε  $x_r = x_0 + rh$ . Θεωρούμε ότι είναι γνωστές οι τιμές της  $f$  στα σημεία  $x_i$ ,  $i = 0, 1, \dots, n$ ,  $y_i = f(x_i)$ . Ορίζουμε ως *τελεστή διαφοράς προς τα εμπρός*, και συμβολίζουμε με  $\Delta$  το

$$\Delta y_r = y_{r+1} - y_r , \quad y_{r+1} = f(x_{r+1}) = f(x_r + h) .$$

Ισχύει ότι

$$\Delta^2 y_r = \Delta(\Delta y_r) = y_{r+2} - 2y_{r+1} + y_r .$$

Γενικά ισχύει  $\Delta^p y_r = \Delta^{p-1}(\Delta y_r)$ .

**Παράδειγμα:** Έστω  $y = x^n$ . Τότε

$$\begin{aligned} \Delta^n x^n &= \Delta^{n-1}(\Delta x^n) = \Delta^{n-1}[(x+h)^n - x^n] = \Delta^{n-1}(nhx^{n-1} + \dots) \\ &= \Delta^{n-2}[nh(n-1)x^{n-2} + \dots] = \dots \\ &= n!h^n . \end{aligned}$$

Επομένως, η  $n$ -ιοστή διαφορά του  $x^n$  είναι ένας σταθερός αριθμός. Γενικά ισχύει ότι η  $n$ -ιοστή διαφορά ενός πολυωνύμου  $n$  βαθμού είναι σταθερός αριθμός.

**Παράδειγμα:** Να υπολογιστούν οι διαφορές της  $y = x^2$  στο  $[0, 2]$ , με  $h = 0.5$ .

| $r$ | $x$ | $y = x^2$ | $\Delta y$ | $\Delta^2 y$ | $\Delta^3 y$ | $\Delta^4 y$ |
|-----|-----|-----------|------------|--------------|--------------|--------------|
| 0   | 0.0 | 0.0       |            |              |              |              |
|     |     |           | 0.25       |              |              |              |
| 1   | 0.5 | 0.25      |            | 0.5          |              |              |
|     |     |           | 0.75       | 0.0          |              |              |
| 2   | 1.0 | 1.0       |            | 0.5          | 0.0          |              |
|     |     |           | 1.25       | 0.0          |              |              |
| 3   | 1.5 | 2.25      |            | 0.5          |              |              |
|     |     |           | 1.75       |              |              |              |
| 4   | 2.0 | 4.0       |            |              |              |              |

Από το παράδειγμα βλέπουμε ότι οι διαφορές  $\Delta^3 y$  είναι μηδέν. Γενικά ισχύει ότι οι  $n+1$  διαφορές ενός πολυωνύμου βαθμού  $n$  είναι μηδέν. Αν η  $f(x)$  είναι συνάρτηση και όχι πολυώνυμο, τότε μπορούμε να βρούμε ένα  $h$  τέτοιο ώστε οι διαφορές υψηλότερου βαθμού συνεχώς να ελαττώνονται σε μέγεθος.

**Παράδειγμα:** Έστω  $f(x) = e^{kx}$ . Τότε

$$\begin{aligned}\Delta f(x) &= e^{k(x+h)} - e^{kx} = e^{kx} (e^{kh} - 1) \\ \Delta^2 f(x) &= e^{kx} (e^{kh} - 1)^2 \\ &\vdots \\ \Delta^n f(x) &= e^{kx} (e^{kh} - 1)^n\end{aligned}$$

Για να ελαττώνονται οι διαφορές θα πρέπει  $e^{kh} - 1 < 1$  ή  $h < \frac{\ln 2}{k}$ .

### 6.6.1 Ιδιότητες

Ο τελεστής  $\Delta$ , όπως και οι άλλοι τελεστές που θα δούμε παρακάτω, υπακούουν στους νόμους της Άλγεβρας.

**Ισότητα τελεστών** Δυο τελεστές  $T_1, T_2$ , είναι ίσοι όταν για κάθε  $f(x)$  ισχύει  $T_1 f(x) = T_2 f(x)$ .

**Τελεστής μετατόπισης** Ο τελεστής μετατόπισης  $E$  ορίζεται από τη σχέση  $y_{r+1} = E y_r$ . Εύκολα μπορούμε να δούμε ότι  $E = 1 + \Delta$ , όπου 1 είναι ο ταυτοτικός τελεστής, δηλαδή ο τελεστής που αφήνει μια συνάρτηση  $f$  αμετάβλητη.

Για τον τελεστή  $E$  η δύναμη  $E^\alpha$  για κάθε  $\alpha$  ορίζεται από

$$E^\alpha y_r = y_{r+\alpha}, \quad \text{ή} \quad E^\alpha f(x) = f(x + ah).$$

**Διαφορικός τελεστής** Ο διαφορικός τελεστής  $D$  ορίζεται από τη σχέση  $y'_r = D y_r$ . Συνδέεται με τον τελεστή μετατόπισης με τη σχέση  $E = e^{hD}$ . Πράγματι

$$y_{r+1} = y_r + h D y_r + \frac{h^2}{2!} D^2 y_r + \dots = \left[ 1 + hD + \frac{1}{2!} (hD)^2 + \dots \right] y_r = e^{hD} y_r.$$

Συνεπώς,  $E = e^{hD}$  και  $hD = \ln E = \ln(1 + \Delta)$ . Από την τελευταία σχέση μπορούμε να υπολογίσουμε την τιμή της παραγώγου μιας συνάρτησης σε ένα σημείο  $x_r$ , όταν είναι γνωστές οι πεπερασμένες διαφορές της συνάρτησης στα  $x_i$ , από τη σχέση:

$$hy' = hDy = \ln(1 + \Delta)y = \Delta y - \frac{1}{2}\Delta^2 y + \frac{1}{3}\Delta^3 y - \dots$$

**Παράδειγμα:** Να υπολογιστεί η παράγωγος της  $f(x) = e^x$  στο  $x = 1.0$ . Οι διαφορές της  $e^x$  δίνονται στον Πίνακα 6.1, ( $h = 0.1$ ). Άρα

$$0.1y'_{x=1.0} \approx 0.285884 - \frac{1}{2} \times 0.030067 + \frac{1}{3} \times 0.003162 - \frac{1}{4} \times 0.000333 + \frac{1}{5} \times 0.000035 = 0.271828 .$$

Επομένως,  $y'_{x=1.0} \approx 2.718280$ . Η ακριβής τιμή είναι  $e^1 = 2.718281828 \dots$

Με τον ίδιο τρόπο μπορούμε να υπολογίσουμε την τιμή της δεύτερης παραγώγου της  $y$  σε ένα σημείο. Από τη σχέση  $y''_r = D^2 y_r$  έχουμε

$$h^2 y''_r = h^2 D^2 y_r = [\ln(1 + \Delta)]^2 y_r = \Delta^2 y_r - \Delta^3 y_r + \frac{11}{12} \Delta^4 y_r + \dots$$

### 6.6.2 Άλλοι τελεστές

**Τελεστής διαφοράς προς τα πίσω** Συμβολίζεται με  $\nabla$  και ορίζεται από τη σχέση

$$\nabla y_r = y_r - y_{r-1} .$$

**Τελεστής κεντρικής διαφοράς** Συμβολίζεται με  $\delta$  και ορίζεται από τη σχέση

$$\delta y_r = y_{r+\frac{1}{2}} - y_{r-\frac{1}{2}} .$$

**Τελεστής μέσης τιμής** Συμβολίζεται με  $\mu$  και ορίζεται από τη σχέση

$$\mu y_r = \frac{y_{r+\frac{1}{2}} + y_{r-\frac{1}{2}}}{2} .$$

Ισχύουν οι σχέσεις

$$\nabla = 1 - E^{-1} , \quad \delta = E^{1/2} - E^{-1/2} , \quad \delta = e^{hD/2} - e^{-hD/2} = 2 \sinh\left(\frac{hD}{2}\right) , \quad \delta = \Delta(1 + \Delta)^{-1/2} .$$

Από τη σχέση  $\delta = 2 \sinh(hD/2)$  μπορούμε να υπολογίσουμε την τιμή της δεύτερης παραγώγου μιας συνάρτησης όταν είναι γνωστές οι κεντρικές διαφορές της,  $\delta y_r$

$$h^2 y''_r = h^2 D^2 y_r = \left[2 \sinh^{-1}\left(\frac{\delta}{2}\right)\right]^2 y_r .$$

Καθώς

$$\sinh^{-1} z = z - \frac{1}{6}z^3 + \frac{3}{40}z^5 - \frac{5}{56}z^7 + \dots ,$$

έχουμε

$$h^2 y''_r = \left(\delta^2 - \frac{1}{2}\delta^4 + \frac{1}{90}\delta^6 + \dots\right) y_r .$$

| $x$ | $f(x) = e^x$ | $\Delta f(x)$ | $\Delta^2 f(x)$ | $\Delta^3 f(x)$ | $\Delta^4 f(x)$ | $\Delta^5 f(x)$ | $\Delta^6 f(x)$ | $\Delta^7 f(x)$ |
|-----|--------------|---------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 0.0 | 1.000000     |               |                 |                 |                 |                 |                 |                 |
| 0.1 | 1.105171     | 0.105171      | 0.011061        |                 |                 |                 |                 |                 |
| 0.2 | 1.221403     | 0.116232      | 0.012224        | 0.001163        | 0.000122        |                 |                 |                 |
| 0.3 | 1.349859     | 0.128456      | 0.013510        | 0.001286        | 0.000135        | 0.000013        |                 |                 |
| 0.4 | 1.491825     | 0.141966      | 0.014931        | 0.001421        | 0.000149        | 0.000014        | 0.000001        | 0.000000        |
| 0.5 | 1.648721     | 0.156897      | 0.016501        | 0.001570        | 0.000165        | 0.000016        | 0.000002        | 0.000000        |
| 0.6 | 1.822119     | 0.173398      | 0.018236        | 0.001735        | 0.000183        | 0.000017        | 0.000002        | 0.000000        |
| 0.7 | 2.013753     | 0.191634      | 0.020154        | 0.001918        | 0.000202        | 0.000019        | 0.000002        | 0.000000        |
| 0.8 | 2.225541     | 0.211788      | 0.022274        | 0.002120        | 0.000223        | 0.000021        | 0.000002        | 0.000000        |
| 0.9 | 2.459603     | 0.234062      | 0.024617        | 0.002343        | 0.000246        | 0.000023        | 0.000002        | 0.000000        |
| 1.0 | 2.718282     | 0.258679      | 0.027205        | 0.002589        | 0.000272        | 0.000026        | 0.000003        | 0.000000        |
| 1.1 | 3.004166     | 0.285884      | 0.030067        | 0.002861        | 0.000301        | 0.000029        | 0.000003        | 0.000000        |
| 1.2 | 3.320117     | 0.315951      | 0.033229        | 0.003162        | 0.000333        | 0.000032        | 0.000003        | 0.000000        |
| 1.3 | 3.669297     | 0.349180      | 0.036724        | 0.003495        | 0.000368        | 0.000035        |                 |                 |
| 1.4 | 4.055200     | 0.385903      | 0.040586        | 0.003862        |                 |                 |                 |                 |
| 1.5 | 4.481689     | 0.426489      |                 |                 |                 |                 |                 |                 |

Πίνακας 6.1: Πίνακας διαφορών της συνάρτησης  $f(x) = e^x$  με ακρίβεια 6 δεκαδικών ψηφίων

### 6.6.3 Γενικευμένοι τύποι του Newton

Έστω  $y(x)$  μια συνάρτηση η οποία γράφεται  $y(x) = y(x_r + ah) = y_{r+\alpha}$ , με  $\alpha = (x - x_r)/h$ ,  $\alpha$  συνεχής. Τότε

$$\begin{aligned} y(x) &= y(x_r + ah) = y_{r+\alpha} = E^\alpha y_r = (1 + \Delta)^\alpha y_r \\ &= \left( 1 + \alpha\Delta + \frac{\alpha(\alpha-1)}{2!}\Delta^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}\Delta^3 + \dots \right) y_r. \end{aligned} \quad (6.31)$$

Με τον ίδιο τρόπο, από τη σχέση  $E = (1 - \nabla)^{-1}$  έχουμε:

$$\begin{aligned} y_{r+\alpha} &= E^\alpha y_r = (1 - \nabla)^{-\alpha} y_r \\ &= \left( 1 + \alpha\nabla + \frac{\alpha(\alpha+1)}{2!}\nabla^2 + \frac{\alpha(\alpha+1)(\alpha+2)}{3!}\nabla^3 + \dots \right) y_r. \end{aligned} \quad (6.32)$$

### 6.6.4 Εφαρμογή των τελεστών στον υπολογισμό ολοκληρωμάτων

#### Παραδείγματα:

1. Να υπολογιστεί το ολοκλήρωμα

$$I = \int_{x_r}^{x_r+2h} y(x) dx.$$

Για  $x \in (x_r, x_r + 2h)$ ,  $x = x_r + \alpha h$ . Από την (6.31) έχουμε:

$$\begin{aligned} I &= \int_{x_r}^{x_r+2h} \left( y_r + \alpha\Delta y_r + \frac{\alpha(\alpha-1)}{2!}\Delta^2 y_r + \dots \right) dx \\ &= \int_0^2 \left( y_r + \alpha\Delta y_r + \frac{\alpha(\alpha-1)}{2!}\Delta^2 y_r + \dots \right) h d\alpha \\ &= 2h \left( y_r + \Delta y_r + \frac{1}{6}\Delta^2 y_r + 0\Delta^3 y_r - \frac{1}{180}\Delta^4 y_r + \dots \right). \end{aligned}$$

Όταν στην παραπάνω εξίσωση αποκόψουμε τους όρους τρίτου και μεγαλύτερου βαθμού ως προς  $\Delta$  έχουμε:

$$I = \frac{h}{3} (y_r + 4y_{r+1} + y_{r+2}).$$

Η σχέση αυτή είναι η μέθοδος Simpson, §5.3, και ο κυρίαρχος όρος του σφάλματος αποκοπής θα είναι:

$$R = -\frac{h}{90}\Delta^4 y_r \Rightarrow R = -\frac{h^5}{90}y_r^{(4)}.$$

2. Να υπολογιστεί το ολοκλήρωμα

$$I = \int_{x_r}^{x_{r+1}} y(x) dx$$

με τη βοήθεια του τελεστή  $\nabla$ . Από την (6.32) έχουμε:

$$\begin{aligned} \int_{x_r}^{x_{r+1}} y(x) dx &= \int_{x_r}^{x_{r+1}} \left( y_r + \alpha\nabla y_r + \frac{\alpha(\alpha+1)}{2!}\nabla^2 y_r + \dots \right) dx \\ &= h \left( y_r + \frac{1}{2}\nabla y_r + \frac{5}{12}\nabla^2 y_r + \frac{3}{8}\nabla^3 y_r + \dots \right). \end{aligned} \quad (6.33)$$

3. Να υπολογιστεί το ολοκλήρωμα

$$I = \int_{x_r}^{x_{r+1}} y(x) dx$$

με τη βοήθεια του τελεστή  $\nabla$  και του  $y_{r+1}$  αντί του  $y_r$ .

$$\begin{aligned} \int_{x_r}^{x_{r+1}} y(x) dx &= \int_{-1}^0 \left( y_{r+1} + \alpha \nabla y_{r+1} + \frac{\alpha(\alpha+1)}{2!} \nabla^2 y_{r+1} + \dots \right) h d\alpha \\ &= h \left( y_{r+1} - \frac{1}{2} \nabla y_{r+1} - \frac{1}{12} \nabla^2 y_{r+1} - \frac{1}{24} \nabla^3 y_{r+1} - \dots \right). \end{aligned} \quad (6.34)$$

Από τις (6.33, 6.34) παρατηρούμε ότι οι συντελεστές του  $\nabla$  μικραίνουν γρηγορότερα στη δεύτερη. Για το λόγο αυτό, όταν αποκόπτουμε σε αυτές τους όρους από ένα συγκεκριμένο όρο και μετά, θα πρέπει να αναμένουμε καλύτερη προσέγγιση του ολοκληρώματος από την (6.34).

Όταν αποκόψουμε από την (6.34) τους όρους δεύτερου και μεγαλύτερου βαθμού, τότε παίρνουμε τον κανόνα τραπεζίου, §5.2,

$$I = \frac{h}{2}(y_{r+1} + y_r),$$

και ο κυρίαρχος όρος του σφάλματος αποκοπής θα είναι  $-\frac{1}{12}h^3 y_{r+1}''$ .

4. Να υπολογιστεί το ολοκλήρωμα

$$I = \int_{x_r}^{x_{r+1}} y(x) dx$$

με τη βοήθεια του τελεστή  $\delta$ .

Ισχύει ότι

$$D = \frac{2}{h} \sinh^{-1} \left( \frac{\delta}{2} \right) = \frac{1}{h} \left( \delta - \frac{1}{24} \delta^3 + \frac{3}{640} \delta^5 - \dots \right). \quad (6.35)$$

Για ένα  $x \in (x_r - h, x_r + h)$ ,  $x = x_r - h + \alpha h$  και

$$y(x) = E^{\alpha-1} y_r = e^{h(\alpha-1)D} y_r.$$

Η παραπάνω σχέση, σε συνδυασμό με την (6.35), δίνει μετά από πράξεις:

$$\begin{aligned} \int_{x_r}^{x_{r+1}} y(x) dx &= \int_0^1 \left[ 1 + (\alpha-1) \left( \delta - \frac{1}{24} \delta^3 + \dots \right) \right. \\ &\quad \left. + \frac{(\alpha-1)^2}{2!} \left( \delta - \frac{1}{24} \delta^3 + \dots \right) + \dots \right] h d\alpha \\ &= 2h \left( 1 + \frac{1}{6} \delta^2 - \frac{1}{180} \delta^4 + \frac{1}{1512} \delta^6 - \dots \right) y_r. \end{aligned}$$

## 6.7 Μέθοδοι βασιζόμενες σε αριθμητική ολοκλήρωση. Πολυβηματικές μέθοδοι

Στις μεθόδους που εξετάσαμε για την επίλυση μιας ΔΕ, Taylor, Euler, Runge-Kutta, μέσω εξίσωσης διαφορών, η ιδέα ήταν να αναπτύξουμε τη λύση σε σειρά Taylor. Στις μεθόδους αυτές,

για να υπολογίσουμε την τιμή της λύσης  $y_{r+1}$  σε ένα σημείο  $x_{r+1}$ , χρησιμοποιήσαμε την τιμή της λύσης  $y_r$  σε ένα μόνο σημείο  $x_r$ . γι' αυτό, οι μέθοδοι αυτές λέγονται μονοβηματικές. Μια μέθοδος στην οποία για τον υπολογισμό της  $y_{r+1}$  χρειαζόμαστε περισσότερες από μία τιμές της λύσης σε προηγούμενα σημεία,  $y_r, y_{r-1}, y_{r-2}, \dots$  θα τη λέμε πολυβηματική. Με άλλα λόγια, όταν για μία ΔΕ η αντίστοιχη εξίσωση διαφορών έχει βαθμό  $\geq 2$ , τότε η μέθοδος επίλυσής της είναι πολυβηματική. Στα επόμενα θα εξετάσουμε μεθόδους για την επίλυση μιας ΔΕ χρησιμοποιώντας αριθμητική ολοκλήρωση.

Έστω το πρόβλημα αρχικών τιμών

$$y' = f(x, y), \quad (6.36\alpha)$$

$$y(x_0) = y_0. \quad (6.36\beta)$$

Ολοκληρώνουμε τα δύο μέλη της (6.36α) μεταξύ  $x_r$  και  $x_{r+1}$  οπότε έχουμε:

$$y_{r+1} - y_r = \int_{x_r}^{x_{r+1}} f(x, y) dx = \int_{x_r}^{x_{r+1}} y'(x) dx \Rightarrow y_{r+1} = y_r + \int_{x_r}^{x_{r+1}} y'(x) dx. \quad (6.37)$$

Η ιδέα εδώ είναι να αντικαταστήσουμε το ολοκλήρωμα με μια κατάλληλη προσέγγιση ή, καλύτερα, να αντικαταστήσουμε την  $f$  με ένα πολυώνυμο παρεμβολής. Υποθέτουμε ότι έχουμε ήδη υπολογίσει προσεγγίσεις  $y_0, y_1, \dots, y_n$ , της  $y(x)$  στα ισαπέχοντα σημεία  $x_r = x_0 + rh$ ,  $r = 0, 1, \dots, n$  με  $f_r$  θα συμβολίζουμε μια προσέγγιση της  $f(x_r, y(x_r))$ ,  $r = 0, 1, \dots, n$ .

### 6.7.1 Μέθοδος Adams–Bashforth

Αν για τον υπολογισμό του ολοκληρώματος (6.37) χρησιμοποιήσουμε την (6.33), παίρνουμε έναν αλγόριθμο για τη λύση του (6.36α) ο οποίος λέγεται *αλγόριθμος των Adams–Bashforth*:

$$y_{r+1} = y_r + \left(1 + \frac{1}{2}\nabla + \frac{5}{12}\nabla^2 + \frac{3}{8}\nabla^3 + \dots\right)y'_r. \quad (6.38)$$

Ο τύπος αυτός λέγεται και *τύπος της πρόβλεψης* (Predictor) ή, ακόμη, *ανοικτός τύπος* (explicit). Αν στον τύπο παρεμβολής της  $f$  χρησιμοποιήσουμε τα  $x_r$  και τα προηγούμενά τους  $x_{r-1}, x_{r-2}, \dots$  τότε η μέθοδος που παίρνουμε λέγεται ανοικτή (explicit) ενώ αν για την προσέγγιση της  $y$  στο  $x_{r+1}$  χρησιμοποιήσουμε το  $x_{r+1}$  και τα προηγούμενά του η μέθοδος λέγεται κλειστή.

Από την (6.38) έχουμε όλους τους τύπους της μεθόδου Adams–Bashforth ανάλογα με τους όρους που αποκόπτουμε κάθε φορά. Έτσι

1. Αν αποκόψουμε τους όρους μετά το σταθερό όρο τότε προκύπτει η μέθοδος Euler, §6.4.1,

$$y_{r+1} = y_r + hy'_r.$$

2. Αν κρατήσουμε και τον όρο  $\nabla$  τότε

$$y_{r+1} = y_r + \frac{h}{2}(3y'_r - y'_{r-1}),$$

με τον κυρίαρχο όρο στο σφάλμα αποκοπής τον  $\frac{5}{12}h^3 y'''(\xi)$ .

3. Αν αποκόψουμε τους όρους μετά το  $\nabla^3$  έχουμε τον τύπο

$$y_{r+1} = y_r + \frac{h}{24}(55y'_r - 59y'_{r-1} + 37y'_{r-2} - 9y'_{r-3}).$$

Όπως βλέπουμε από τα παραδείγματα αυτά, σε μία μέθοδο Adams–Bashforth βαθμού  $p$  χρειαζόμαστε  $p$  τιμές της  $y$  για να ξεκινήσουμε τον αλγόριθμο. Αυτό αποτελεί και το σοβαρότερο μειονέκτημα της μεθόδου. Για το λόγο αυτό, η μέθοδος χρησιμοποιείται σε συνδυασμό με μια άλλη μέθοδο για τον προσδιορισμό των  $y(x_0), \dots, y(x_{p-1})$ : η τελευταία είναι συνήθως μια Runge–Kutta του ίδιου βαθμού. Έχοντας γνωστή την αρχική συνθήκη  $y(x_0) = y_0$ , ξεκινάμε τη λύση της (6.36α) με μία Runge–Kutta βαθμού  $p$  και υπολογίζουμε τις τιμές της λύσης στα  $x_0, x_1, \dots, x_{p-1}$ , έστω  $u_0, u_1, \dots, u_{p-1}$ . Κατόπιν, με τη βοήθεια των υπολογισμένων τιμών συνεχίζουμε τη λύση του προβλήματος με μια Adams–Bashforth βαθμού  $p$ . Το πλεονέκτημα της Adams–Bashforth σε αντίθεση με τη Runge–Kutta είναι ότι χρειαζόμαστε μόνο ένα υπολογισμό της  $f$ , ενώ στη Runge–Kutta  $p$  βαθμού χρειαζόμαστε  $p$  υπολογισμούς της  $f$  σε κάθε επανάληψη.

### 6.7.2 Μέθοδος Adams–Moulton

Θα προσεγγίσουμε και πάλι το ολοκλήρωμα στην (6.37) με τη βοήθεια της (6.34). Έχουμε

$$y_{r+1} = y_r + \left(1 - \frac{1}{2}\nabla - \frac{1}{12}\nabla^2 - \frac{1}{24}\nabla^3 - \dots\right)y'_{r+1}. \quad (6.39)$$

Η σχέση αυτή είναι ένας *κλειστός τύπος* (implicit) με την έννοια ότι χρειαζόμαστε πληροφορίες για την τιμή της λύσης στο  $x_{r+1}$ , (καθώς  $y'_{r+1} = f(x_{r+1}, y_{r+1})$ ). Γενικότερα, μη γραμμικές εξισώσεις προκύπτουν για τον υπολογισμό του  $y_{r+1}$ . Ο τύπος (6.39) είναι γνωστός και ως *τύπος της διόρθωσης* (Corrector) ή *αλγόριθμος των Adams–Moulton*.

Όταν στην (6.39) αποκόψουμε τους όρους μετά το  $\nabla$  έχουμε τον κανόνα τραπεζίου

$$y_{r+1} = y_r + \frac{h}{2}(y'_r + y'_{r+1}). \quad (6.40)$$

Ο κυρίαρχος όρος του σφάλματος αποκοπής είναι  $-\frac{1}{12}h^3y'''(\xi)$ .

#### Παραδείγματα:

1. Να λυθεί η

$$y' = xy.$$

Χρησιμοποιώντας τον τύπο (6.40) έχουμε

$$y_{r+1} = \frac{1 + \frac{1}{2}hx_r}{1 - \frac{1}{2}hx_{r+1}}y_r.$$

2. Να λυθεί η

$$y' = xy^2$$

με τον κανόνα τραπεζίου.

$$y_{r+1} = y_r + \frac{h}{2}(x_r y_r^2 + x_{r+1} y_{r+1}^2).$$

Εδώ έχουμε μια μη γραμμική εξίσωση διαφορών ως προς  $y_{r+1}$  και θα πρέπει σε κάθε επανάληψη, να λύνουμε μια μη γραμμική εξίσωση, πράγμα το οποίο είναι πολύ δύσκολο στην πράξη.

### 6.7.3 Μέθοδοι Πρόβλεψης–Διόρθωσης (Predictor–Corrector)

Με τη μέθοδο πρόβλεψης υπολογίζουμε το  $y_{r+1}$  από τις γνωστές τιμές  $y_r, y_{r-1}, \dots, y_0$ , αλλά οι συντελεστές της σειράς των πεπερασμένων διαφορών (6.38) μικραίνουν πολύ αργά. Δηλαδή, για να έχουμε μια ικανοποιητική ακρίβεια θα πρέπει να πάρουμε ένα σχετικά μεγάλο αριθμό όρων στη (6.38), πράγμα το οποίο σημαίνει ότι σε κάθε επανάληψη θα έχουμε περισσότερη δουλειά, περισσότερες πράξεις, περισσότερα λάθη στρογγύλευσης. Με τη μέθοδο της διόρθωσης, οι συντελεστές της σειράς πεπερασμένων διαφορών (6.39) μπορεί να μικραίνουν πολύ γρήγορα αλλά προκύπτουν μη γραμμικές εξισώσεις.

#### Παράδειγμα:

##### Μέθοδος πρόβλεψης

$$y_{r+1} = y_r + h y'_r.$$

Το σφάλμα αποκοπής είναι

$$T_p = \frac{h^2}{2} y''(\xi).$$

##### Μέθοδος διόρθωσης

$$y_{r+1} = y_r + \frac{h}{2} (y'_r + y'_{r+1}).$$

Το σφάλμα αποκοπής είναι

$$T_c = -\frac{h^3}{12} y'''(\xi).$$

Παρακάτω θα περιγράψουμε έναν αλγόριθμο, ο οποίος λέγεται *αλγόριθμος πρόβλεψης-διόρθωσης*. Αποτελεί ένα συνδυασμό των παραπάνω μεθόδων και ξεπερνά όλες τις δυσκολίες της καθεμίας χωριστά.

**Αλγόριθμος:** Έστω ότι δίνεται η τιμή του  $y$  στο  $x_r$ ,  $y_r = y(x_r)$ .

1. Πρόβλεψε την τιμή  $\bar{y}_{r+1}$  με μια μέθοδο Adams–Bashforth.
2. Υπολόγισε το  $\bar{y}'_{r+1} = f(x_{r+1}, \bar{y}_{r+1})$ .
3. Διόρθωσε, υπολόγισε το  $y_{r+1}$  χρησιμοποιώντας το  $\bar{y}'_{r+1}$  στο δεξί μέλος μιας μεθόδου Adams–Moulton.
4. Υπολόγισε το  $y'_{r+1} = f(x_{r+1}, y_{r+1})$ .
5. Επανάλαβε για το επόμενο διάστημα.

**Υπολογισμός του σφάλματος αποκοπής**

Διαλέγουμε ένα αλγόριθμο Πρόβλεψης-Διόρθωσης του ίδιου βαθμού και έστω  $T_p$  το σφάλμα αποκοπής από την πρόβλεψη και  $T_c$  από τη διόρθωση.

Όταν μια μέθοδος είναι βαθμού  $s$  τότε

$$T_p = A_p h^{s+1} y^{(s+1)}(\xi_p), \quad \text{και} \quad (6.41)$$

$$T_c = A_c h^{s+1} y^{(s+1)}(\xi_c), \quad (6.42)$$

όπου τα  $A_p, A_c$  είναι σταθεροί συντελεστές. Υποθέτουμε ότι το διάστημα  $h = x_{r+1} - x_r$  είναι τόσο μικρό ώστε

$$y^{(s+1)}(\xi_p) \approx y^{(s+1)}(\xi_c) \approx y^{(s+1)}(\xi). \quad (6.43)$$

Ισχύουν

$$\bar{y}_{r+1} - Y_{r+1} = T_p = A_p h^{s+1} y^{(s+1)}(\xi_p), \quad (6.44\alpha)$$

$$y_{r+1} - Y_{r+1} = T_c = A_c h^{s+1} y^{(s+1)}(\xi_c), \quad (6.44\beta)$$

όπου  $Y_{r+1}$  είναι η πραγματική λύση και  $\bar{y}_{r+1}$  η υπολογισμένη τιμή από την πρόβλεψη και  $y_{r+1}$  η υπολογισμένη από τη διόρθωση. Από τις (6.43, 6.44α) έχουμε

$$y_{r+1} - \bar{y}_{r+1} = A h^{s+1} y^{(s+1)}(\xi),$$

με  $A = A_p - A_c$ . Από τη σχέση αυτή είναι εύκολο να υπολογίσουμε τα  $T_p, T_c$ .

**Παράδειγμα:** Ως τύπο πρόβλεψης διαλέγουμε μια μέθοδο δεύτερου βαθμού

$$y_{r+1} = y_r + \frac{h}{2} (3y'_r - y'_{r-1}), \quad (6.45\alpha)$$

$$\text{με} \quad T_p = \frac{5}{12} h^3 y'''(\xi_p) \quad (6.45\beta)$$

και ως τύπο διόρθωσης την αντίστοιχη σχέση

$$y_{r+1} = y_r + \frac{h}{2} (y'_r + y'_{r+1}), \quad (6.46\alpha)$$

$$\text{με} \quad T_c = -\frac{1}{12} h^3 y'''(\xi_c). \quad (6.46\beta)$$

Τότε

$$T_p = \frac{5}{6} (y_{r+1} - \bar{y}_{r+1}) \quad (6.47\alpha)$$

$$\text{και} \quad T_c = -\frac{1}{6} (y_{r+1} - \bar{y}_{r+1}). \quad (6.47\beta)$$

**Επιλογή του βήματος  $h$** 

Σκοπός μας είναι να χρησιμοποιήσουμε ένα βήμα  $h$  το οποίο θα κρατά το σφάλμα αποκοπής, το οποίο εξαρτάται από το  $h$ , μέσα σε ορισμένα όρια. Όταν το  $h$  είναι πολύ μεγάλο, τότε η λύση δεν είναι ακριβής (μεγάλο σφάλμα αποκοπής)· όταν το  $h$  είναι πολύ μικρό τότε και πάλι η λύση δεν είναι καλή (πολλά λάθη στρογγύλευσης).

Όμως, μπορούμε να ξεκινήσουμε μια μέθοδο Πρόβλεψης-Διόρθωσης με βήμα  $h$  το οποίο θα μεταβάλλεται σε κάθε επανάληψη ανάλογα με το σφάλμα αποκοπής. Όταν το σφάλμα είναι μεγάλο θα παίρνουμε μικρότερο βήμα στην επόμενη επανάληψη και αντίστροφα. Θα πρέπει, βέβαια, να καθορίσουμε τι είναι μεγάλο και τι μικρό σφάλμα αποκοπής. Αλλάζοντας την τιμή του  $h$  σε κάθε επανάληψη εμφανίζονται πολλές δυσκολίες· το πρόβλημα απλοποιείται αν πάρουμε ως νέο βήμα το διπλάσιο ή το μισό του παλαιού, ανάλογα με το σφάλμα.

Στην πρώτη περίπτωση, όταν δηλαδή, πάρουμε ως νέο  $h$  το διπλάσιο του προηγούμενου, το πρόβλημα είναι απλό καθώς  $\tilde{x}_{r-1} \equiv x_{r-2}$ ,  $\tilde{x}_{r-2} \equiv x_{r-4}$ , κλπ. και οι τιμές της  $y$  είναι γνωστές στα σημεία  $\tilde{x}_{r-1}$ ,  $\tilde{x}_{r-2}$ ,  $\tilde{x}_{r-3}$ , ... Όταν το νέο  $h$  είναι ίσο με  $h/2$  τότε  $\tilde{x}_{r-2} \equiv x_{r-1}$ ,  $\tilde{x}_{r-4} \equiv x_{r-2}$ , ... το νέο  $\tilde{x}_{r-1}$  είναι το σημείο στο μέσο μεταξύ των  $x_{r-1}$  και  $x_r$ , και αντίστοιχα για τα υπόλοιπα. Στην περίπτωση αυτή, θα πρέπει να υπολογίσουμε με παρεμβολή την τιμή της  $y$  στο  $\tilde{x}_{r-1} \equiv x_{r-\frac{1}{2}}$ ,  $y_{r-\frac{1}{2}} = y(x_{r-\frac{1}{2}})$ , κ.ο.κ.

Όταν στον αλγόριθμο Πρόβλεψης-Διόρθωσης που δώσαμε παραπάνω, ενσωματώσουμε τη μέθοδο υπολογισμού του σφάλματος και τη μέθοδο επιλογής του διαστήματος  $h$ , ο αλγόριθμος γίνεται:

**Αλγόριθμος Πρόβλεψης-Διόρθωσης:** Κατ' αρχήν αποφασίζουμε ότι  $E_{\min} < |E_r| < E_{\max}$ .

1. Πρόβλεψε το  $\bar{y}_{r+1}$ , υπολόγισε το  $\bar{y}'_{r+1} = f(x_{r+1}, \bar{y}_{r+1})$ .
2. Διόρθωσε βρίσκοντας το  $y_{r+1}$ , υπολόγισε το  $y'_{r+1} = f(x_{r+1}, y_{r+1})$ .
3. Υπολόγισε το σφάλμα από τα  $\bar{y}_{r+1}$ ,  $y_{r+1}$ .
4. Αν  $|E_r| > E_{\max}$  τότε  $h \leftarrow h/2$  και επανάλαβε από το  $x_r$ .
5. Αν  $|E_r| < E_{\min}$  τότε  $h \leftarrow 2h$  και επανάλαβε από το  $x_r$ .
6. Αλλιώς, συνέχισε με το ίδιο βήμα  $h$  ξανά από την πρώτη εντολή.

Η μέθοδος Πρόβλεψης-Διόρθωσης είναι πολυθηματική μέθοδος, επομένως, θα πρέπει να γνωρίζουμε περισσότερες της μίας αρχικές συνθήκες για να ξεκινήσουμε. Ο πιο συνηθισμένος τρόπος είναι να εφαρμόσουμε μια Runge-Kutta ανάλογου βαθμού για το ξεκίνημα του αλγορίθμου και μόλις υπολογιστούν οι απαιτούμενες αρχικές συνθήκες να συνεχίσουμε με μέθοδο Πρόβλεψης-Διόρθωσης. Ένα άλλο σημείο στο οποίο πρέπει να παρεμβάλουμε τη μέθοδο Runge-Kutta είναι το βήμα 4. Στο βήμα αυτό θα υπολογίσουμε το  $y_{r-\frac{1}{2}}$  με Runge-Kutta από τις γνωστές τιμές της  $y$  στα προηγούμενα σημεία.

## 6.8 Συστήματα Διαφορικών Εξισώσεων

Οι μέθοδοι αριθμητικής επίλυσης ΔΕ που εξετάσαμε, μπορούν εύκολα να εφαρμοστούν στην περίπτωση συστημάτων ΔΕ ή μιας ΔΕ υψηλότερου βαθμού ( $\geq 2$ ). Έστω το σύστημα των ΔΕ πρώτου βαθμού

$$y'_1 = f_1(x, y_1, y_2, \dots, y_n) \quad (6.48\alpha)$$

$$y'_2 = f_2(x, y_1, y_2, \dots, y_n) \quad (6.48\beta)$$

$$\vdots \quad \vdots \quad (6.48\gamma)$$

$$y'_n = f_n(x, y_1, y_2, \dots, y_n) \quad (6.48\delta)$$

όπου οι  $f_i$  είναι πραγματικές συναρτήσεις, ορισμένες για  $x \in [a, b]$  και για κάθε πραγματικό  $y_1, y_2, \dots, y_n$ . Οι τιμές των  $y_i$  στο  $(n + 1)$  βήμα θα υπολογιστούν από τις τιμές των  $y'_i$  και τις προηγούμενες τιμές των  $y_i, y'_i$ , με τον ίδιο τρόπο όπως στις απλές ΔΕ. Ο υπολογισμός των  $y'_i$  από την (6.48α) είναι το μόνο σημείο στο οποίο υπάρχει διαφορά από την απλή περίπτωση καθώς έχουμε  $n$  τιμές  $y_i$  αντί για μία, όπως στις απλές ΔΕ. Γενικά, η λύση του (6.48α), αν υπάρχει, δε θα είναι μοναδική, εκτός αν προσδιοριστούν  $n$  αρχικές συνθήκες:

$$y_i(x_0) = s_i, \quad i = 1, \dots, n, \quad (6.49)$$

όπου τα  $s_i$  είναι γνωστά και  $x_0 \in [a, b]$ . Οι (6.48α, 6.49) συνιστούν ένα πρόβλημα αρχικών τιμών που σε διανυσματική μορφή γράφεται:

$$\mathbf{y}' = \mathbf{f}(x, \mathbf{y}), \quad (6.50\alpha)$$

$$\mathbf{y}(x_0) = \mathbf{s}. \quad (6.50\beta)$$

**Παράδειγμα:** Να λυθεί το πρόβλημα αρχικών τιμών

$$\begin{aligned} y'_1 &= xy_1 - y_2 \\ y'_2 &= -y_1 + y_2 \\ y_1(0) &= s_1 \\ y_2(0) &= s_2 \end{aligned}$$

Ορίζουμε  $y_{10} \equiv y_1(0)$  και  $y_{20} \equiv y_2(0)$ . Η σειρά Taylor 2<sup>ου</sup> βαθμού είναι:

$$\begin{aligned} y_{1r+1} &= y_{1r} + h(x_r y_{1r} - y_{2r}) + \frac{h^2}{2} [(x_r^2 + 2)y_{1r} - (1 + x_r)y_{2r}] \\ y_{2r+1} &= y_{2r} + h(-y_{1r} + y_{2r}) + \frac{h^2}{2} [-(1 + x_r)y_{1r} + 2y_{2r}], \end{aligned}$$

με  $r = 0, 1, \dots$

Η Runge-Kutta δεύτερου βαθμού δίνει:

$$\begin{aligned} y_{1r+1} &= y_{1r} + \frac{h}{2} [x_r y_{1r} - y_{2r} + (x_r - h)(y_{1r} + x_r y_{1r} - 2y_{2r}) - 2y_{2r} + y_{1r}] \\ y_{2r+1} &= y_{2r} + \frac{h}{2} (4y_{2r} - 3y_{1r} - x_r y_{1r}). \end{aligned}$$

Η μέθοδος Πρόβλεψης-Διόρθωσης εφαρμόζεται με παρόμοιο τρόπο.

**Παρατήρηση:** Όπως είναι γνωστό, μια ΔΕ βαθμού  $\geq 2$  μπορεί να γραφτεί ως ένα σύστημα ΔΕ πρώτου βαθμού, οπότε για την αριθμητική επίλυση μιας ΔΕ βαθμού ανώτερου του πρώτου, μετατρέπουμε την εξίσωση σε ισοδύναμο σύστημα ΔΕ πρώτου βαθμού και συνεχίζουμε κατά τα γνωστά.

Το (6.50α) έχει μία μοναδική λύση αν η  $\mathbf{f}$  ικανοποιεί μια συνθήκη Lipschitz, §6.2, δηλαδή,  $\exists L > 0$  τέτοιο ώστε  $\forall \mathbf{y}, \mathbf{z} \in \mathcal{R}^n$  και για κάθε  $x \in [a, b]$  να ισχύει

$$\|\mathbf{f}(x, \mathbf{y}) - \mathbf{f}(x, \mathbf{z})\|_\infty \leq L \|\mathbf{y} - \mathbf{z}\|_\infty.$$

Οι διάφορες μέθοδοι επίλυσης ΔΕ πρώτης τάξης γενικεύονται εύκολα για συστήματα ΔΕ πρώτου βαθμού. Η μέθοδος Taylor θα είναι:

$$\mathbf{y}(x_{r+1}) = \mathbf{y}(x_r) + h\mathbf{f}(x_r, \mathbf{y}(x_r)) + \frac{h^2}{2!}\mathbf{f}'(x_r, \mathbf{y}(x_r)) + \cdots + \frac{h^p}{p!}\mathbf{f}^{(p)}(x_r, \mathbf{y}(x_r)) + \mathbf{R}_{p+1},$$

όπου

$$\mathbf{f}'(x, \mathbf{y}(x)) \equiv \frac{d}{dx}\mathbf{f}(x, \mathbf{y}(x)) = \frac{\partial \mathbf{f}}{\partial x} + \sum_{j=1}^n \frac{dy_j(x)}{dx} \frac{\partial \mathbf{f}}{\partial y_j} = \frac{\partial \mathbf{f}}{\partial x} + \sum_{j=1}^n f_j \frac{\partial \mathbf{f}}{\partial y_j},$$

$$\frac{\partial \mathbf{f}}{\partial y_j} \equiv \left( \frac{\partial f_1}{\partial y_j}, \frac{\partial f_2}{\partial y_j}, \dots, \frac{\partial f_n}{\partial y_j} \right).$$

Η μέθοδος Taylor δεύτερου βαθμού θα είναι:

$$\mathbf{y}(x_{r+1}) = \mathbf{y}(x_r) + h\mathbf{f}(x_r, \mathbf{y}_r) + \frac{h^2}{2!}\mathbf{f}'(x_r, \mathbf{y}_r),$$

$$\mathbf{y}_0 = \mathbf{s}.$$

Η Runge-Kutta 4<sup>ου</sup> βαθμού θα είναι:

$$\mathbf{y}_{r+1} = \mathbf{y}_r + \frac{h}{6}(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4)$$

με

$$\begin{aligned} \mathbf{k}_1 &= \mathbf{f}(x_r, \mathbf{y}_r), \\ \mathbf{k}_2 &= \mathbf{f}(x_r + h/2, \mathbf{y}_r + h/2\mathbf{k}_1), \\ \mathbf{k}_3 &= \mathbf{f}(x_r + h/2, \mathbf{y}_r + h/2\mathbf{k}_2), \\ \mathbf{k}_4 &= \mathbf{f}(x_r + h, \mathbf{y}_r + h\mathbf{k}_3). \end{aligned}$$

## 6.9 Εξισώσεις Διαφορών

Για να μπορέσουμε στα επόμενα να μιλήσουμε για ακρίβεια και ευστάθεια των προσεγγίσεων των λύσεων που παίρνουμε από διάφορες μεθόδους όπως η Euler, Runge-Kutta κλπ. για την επίλυση ενός προβλήματος αρχικών τιμών θα πρέπει να χρησιμοποιήσουμε ορισμένες ιδιότητες των εξισώσεων διαφορών (ΕΔ). Στην §6.2 δώσαμε τον ορισμό μιας ΔΕ ως μια σχέση μεταξύ μιας ανεξάρτητης μεταβλητής  $x$ , μιας εξαρτημένης συνάρτησης  $y$  και μιας ή περισσότερων παραγώγων της. Ομοίως, ορίζουμε μια ΕΔ ως μια σχέση μεταξύ μιας ανεξάρτητης μεταβλητής  $x$ , μιας εξαρτημένης συνάρτησης  $y$  και μιας ή περισσότερων διαφορών της  $\Delta y, \Delta^2 y, \dots$ . Δηλαδή, εκεί που μια ΔΕ περιέχει μια συνάρτηση σε ένα διάστημα πραγματικών αριθμών και τις παραγώγους της, μια ΕΔ περιέχει τη συνάρτηση ορισμένη σε ένα διάστημα ακεραίων και τις διαφορές της. Όπως φαίνεται από τον ορισμό παραπάνω, υπάρχει μεγάλη ομοιότητα μεταξύ ΔΕ και ΕΔ, και ό,τι ισχύει στη θεωρία για τις ΔΕ μπορούμε να το μεταφέρουμε στις ΕΔ. Μια ΕΔ βαθμού  $n$  μπορεί να γραφεί ως

$$\Delta^n y_n = f(n, y_n, \Delta y_n, \Delta^2 y_n, \dots, \Delta^{n-1} y_n). \quad (6.51)$$

Μια ακολουθία  $y_m, y_{m+1}$  που ικανοποιεί την (6.51) για κάθε  $n = m, m+1, \dots$  θα τη λέμε *λύση της ΕΔ*. Βαθμός μιας ΕΔ είναι η μεγαλύτερη διαφορά μεταξύ των δεικτών της άγνωστης συνάρτησης. Η πιο απλή ΕΔ βαθμού  $m$  είναι της μορφής

$$\alpha_0 y_n + \alpha_1 y_{n+1} + \cdots + \alpha_m y_{n+m} = 0,$$

όπου τα  $\alpha_0, \alpha_1, \dots, \alpha_m$  είναι ανεξάρτητα του  $n$ . Η παραπάνω σχέση αποτελεί μια *ομογενή γραμμική ΕΔ βαθμού  $n$* . Παρακάτω θα εξετάσουμε τις ΕΔ πρώτου και δεύτερου βαθμού.

### 6.9.1 Εξίσωση διαφορών πρώτου βαθμού

Έστω η ΕΔ

$$\alpha_0 y_{r+1} + a_1 y_r = 0.$$

Αναζητούμε λύσεις της μορφής  $\beta^r$ . Αντικαθιστώντας στην ΕΔ έχουμε

$$\alpha_0 \beta^{r+1} + a_1 \beta^r = 0.$$

Η εξίσωση αυτή λέγεται *χαρακτηριστική εξίσωση της ΕΔ* και η λύση της είναι  $\beta = -\frac{a_1}{\alpha_0}$ . Η λύση της ΕΔ επομένως είναι

$$y_r = c \left( -\frac{a_1}{\alpha_0} \right)^r,$$

όπου το  $c$  είναι μια αυθαίρετη σταθερά που υπολογίζεται από τις αρχικές συνθήκες.

**Παράδειγμα:**

$$\begin{aligned} y_{r+1} + y_r &= 0, \\ y(0) &= 1. \end{aligned}$$

Η γενική της λύση είναι  $y_r = c(-1)^r$ . Από την αρχική συνθήκη  $c = 1$ , συνεπώς  $y_r = (-1)^r$ .

### 6.9.2 Εξίσωση διαφορών δεύτερου βαθμού

Έστω η ΕΔ

$$\alpha_0 y_{r+2} + a_1 y_{r+1} + a_2 y_r = 0. \quad (6.52)$$

Αναζητούμε και πάλι λύσεις της μορφής  $\beta^r$ . Η χαρακτηριστική εξίσωση θα είναι

$$\alpha_0 \beta^2 + a_1 \beta + a_2 = 0.$$

Αν συμβολίσουμε με  $\Delta$  το  $a_1^2 - 4\alpha_0 a_2$ , διακρίνουμε τις περιπτώσεις

1.  $\Delta > 0$ . Τότε

$$\beta_{1,2} = \frac{-a_1 \pm \sqrt{\Delta}}{2\alpha_0}$$

και η γενική λύση της (6.52) είναι

$$y_r = c_1 \beta_1^r + c_2 \beta_2^r,$$

όπου τα  $c_1, c_2$  θα υπολογιστούν από τις αρχικές συνθήκες.

2.  $\Delta = 0$ . Τότε

$$\beta = -\frac{a_1}{2\alpha_0} \quad (\text{διπλή}).$$

Μια άλλη λύση της ΕΔ θα είναι η  $r\beta^r$ . Πράγματι

$$\begin{aligned} &\alpha_0(r+2)\beta^{r+2} + a_1(r+1)\beta^{r+1} + a_2\beta^r \\ &= \beta^r [r(\alpha_0\beta^2 + a_1\beta + a_2) + \beta(2\alpha_0\beta + a_1)] = 0, \end{aligned}$$

και η γενική λύση είναι

$$y_r = c_1 \beta_1^r + c_2 r \beta_2^r.$$

3.  $\Delta < 0$ . Τότε

$$\beta_{1,2} = \frac{-\alpha_1 \pm i\sqrt{|\Delta|}}{2\alpha_0} = \varrho e^{\pm i\theta},$$

όπου

$$\varrho = \frac{\sqrt{\alpha_1^2 + \Delta^2}}{2|\alpha_0|}, \quad \theta = \arctan\left(\frac{\sqrt{|\Delta|}}{\alpha_1}\right).$$

Η γενική λύση θα είναι:

$$\begin{aligned} y_r &= c_1\beta_1^r + c_2\beta_2^r = c_1\varrho^r e^{ir\theta} + c_2\varrho^r e^{-ir\theta} \\ &= \varrho^r [c_1(\cos r\theta + i \sin r\theta) + c_2(\cos r\theta - i \sin r\theta)] \\ &= \varrho(C_1 \cos r\theta + C_2 \sin r\theta) \end{aligned}$$

όπου οι σταθερές  $C_1 = c_1 + c_2$  και  $C_2 = i(c_1 - c_2)$  υπολογίζονται από τις αρχικές συνθήκες.

Τα παραπάνω μπορούν να γενικευτούν για ΕΔ βαθμού  $n > 2$ . Όταν μια ρίζα του χαρακτηριστικού πολυωνύμου είναι  $p$  βαθμού τότε

$$y_r = (c_1 + c_2 r + c_3 r^2 + \dots + c_p r^{p-1})\beta^r.$$

### Παραδείγματα:

1. Έστω η εξίσωση διαφορών

$$y_{r+1} = y_r + y_{r-1}, \quad y(0) = 1, \quad y(1) = 1.$$

Η λύση είναι:

$$y_r = \frac{1}{\sqrt{5}} \left( \frac{1 + \sqrt{5}}{2} \right)^r - \frac{1}{\sqrt{5}} \left( \frac{1 - \sqrt{5}}{2} \right)^r.$$

2. Έστω η εξίσωση διαφορών

$$y_{r+2} - 4y_{r+1} + 4y_r = 0.$$

Η γενική λύση είναι:

$$y_r = c_1 2^r + c_2 r 2^r.$$

3. Έστω η εξίσωση διαφορών

$$y_{r+2} - 2y_{r+1} + 10y_r = 0.$$

Η γενική λύση είναι:

$$y_r = \sqrt{10}^r (c_1 \cos r\theta + c_2 \sin r\theta),$$

όπου

$$\cos \theta = \frac{1}{\sqrt{10}}, \quad \sin \theta = \frac{3}{\sqrt{10}}.$$

### 6.9.3 Μη ομογενείς εξισώσεις διαφορών

Η μη ομογενής ΕΔ

$$\alpha_0 y_n + \alpha_1 y_{n+1} + \dots + \alpha_m y_{n+m} = b_n ,$$

με  $b_n \neq 0$  λύνεται υπολογίζοντας πρώτα τη γενική λύση της αντίστοιχης ομογενούς

$$\alpha_0 y_n + \alpha_1 y_{n+1} + \dots + \alpha_m y_{n+m} = 0 .$$

Κατόπιν, βρίσκουμε μια μερική λύση της μη ομογενούς: η γενική λύση της ισούται με το άθροισμα αυτής της μερικής λύσης και της γενικής λύσης της ομογενούς.

**Παράδειγμα:**

$$\Delta y_r + 2y_r = r \rightarrow y_{r+1} + y_r = r .$$

Η γενική λύση της ομογενούς είναι  $y_r = c(-1)^r$ . Για να βρούμε μια λύση της μη ομογενούς αναζητούμε λύσεις της μορφής  $y_r = ar + \beta$ , οπότε έχουμε

$$ar + \alpha + \beta + ar + \beta = r ,$$

άρα  $\alpha = 0.5$ ,  $\beta = -0.25$ . Η γενική λύση της μη ομογενούς ΕΔ είναι

$$y_r = c(-1)^r + 0.5r - 0.25 .$$

Αν το δεξι μέλος της μη ομογενούς ΕΔ είναι μια σταθερά, αναζητούμε λύσεις της μορφής  $y_r = c$ .

### 6.9.4 Σχόλια

Παρακάτω δίνουμε ένα παράδειγμα στο οποίο φαίνεται πόσο καταστροφικό μπορεί να είναι το αποτέλεσμα της λύσης μιας ΕΔ όταν υπεισέρχονται λάθη στρογγύλευσης κατά τη λύση της.

**Παράδειγμα:** Έστω η ΕΔ

$$y_{r+1} - 10.1y_r + y_{r-1} = 0 ,$$

της οποίας η γενική λύση είναι

$$y_r = A_1 10^r + A_2 10^{-r} .$$

Παρατηρούμε ότι στη λύση αυτή κυριαρχεί ο όρος  $10^r$ . Όταν υποθέσουμε ότι ισχύουν οι αρχικές συνθήκες  $y_0 = 10$ ,  $y_1 = 1$  τότε  $A_1 = 0$  και  $A_2 = 10$ . Επομένως, η λύση που ζητάμε είναι  $y_r = 10 \times 10^{-r}$ . Στην πράξη, όμως, το  $A_1$  δε θα είναι ακριβώς 0 αλλά θα είναι π.χ. της τάξης του  $10^{-16}$ , δηλαδή, η λύση που θα υπολογίζουμε πραγματικά θα είναι

$$y_r = 10^{-16} \times 10^r + 10 \times 10^{-r} .$$

Έτσι για  $r = 10$ , έχουμε  $y_r = 10^{-6} + 10^{-9}$ . για  $r = 20$ ,  $y_r = 10^4 + 10^{-19}$ . Οι τιμές αυτές είναι πράγματι πολύ “κακές” προσεγγίσεις των πραγματικών τιμών.

## 6.10 Αριθμητική Ευστάθεια

Η ευστάθεια μιας αριθμητικής μεθόδου αναφέρεται στη συμπεριφορά του σφάλματος μεταξύ της τιμής που υπολογίζουμε και της πραγματικής τιμής της λύσης. Αν το σφάλμα μεγαλώνει όπως ο αριθμός των βημάτων και τελικά κυριαρχεί της λύσης, η μέθοδος που ακολουθείται είναι ασταθής. Ας θεωρήσουμε εδώ τον υπολογισμό των  $y_r$  από την επαναληπτική μέθοδο που ορίζεται από την ΕΔ

$$y_{n+1} = F(y_n, y_{n-1}, \dots, y_{n-m}), \quad n = m, m+1, \dots, \quad (6.53)$$

όταν δίνονται οι αρχικές τιμές  $y_0, y_1, \dots, y_m$ . Έστω ότι για κάποιο αυθαίρετο  $r$  έχουμε ένα σφάλμα στρογγύλευσης  $\varepsilon$ , δηλαδή υπολογίζουμε το

$$y_r^* = y_r + \varepsilon.$$

Θα έχουμε, τότε, την ακολουθία  $y_0, y_1, \dots, y_{r-1}, y_r^*, y_{r+1}^*, \dots$ .

Θα λέμε ότι η επαναληπτική μέθοδος (6.53) είναι *αριθμητικά ευσταθής* αν, για  $y_n \neq 0$ , τα σχετικά σφάλματα

$$\left| \frac{y_n - y_n^*}{y_n} \right|$$

είναι φραγμένα όταν  $n \rightarrow \infty$ .

**Παράδειγμα:** Έστω ότι θέλουμε να υπολογίσουμε τα  $y_n$  από την

$$y_{n+1} = 100.01y_n - y_{n-1},$$

με  $y_0 = 1$  και  $y_1 = 0.01$ . Η γενική λύση της είναι

$$y_n = c_1 100^n + c_2 0.01^n,$$

και η λύση που ικανοποιεί τις αρχικές συνθήκες είναι

$$y_n = 0.01^n, \quad (c_1 = 0, c_2 = 1).$$

Η εισαγωγή λαθών στρογγύλευσης κατά τη διάρκεια των υπολογισμών έχει ως αποτέλεσμα να αλλάξει τη λύση που θέλουμε να υπολογίσουμε. Δηλαδή, το  $c_1$  δε θα είναι ακριβώς 0· επομένως, θα πρέπει να προσθέσουμε τον όρο  $c_1 100^n$  στο  $y_n$ . Αυτό έχει ως αποτέλεσμα το  $c_1 100^n$  πολύ γρήγορα να κυριαρχεί της πραγματικής λύσης (που είναι  $c_2 0.01^n$ ). Έτσι, π.χ., αν  $y_0^* = y_0 + 10^{-6} = 1.000001$  και  $y_1^* = y_1 = 0.01$  τότε

$$\begin{aligned} y_2^* &\approx 0.0000990214, \\ y_3^* &\approx -0.0000968733, \\ y_4^* &\approx -0.0097873247, \\ y_5^* &\approx -0.9787334946, \\ y_6^* &\approx -97.8733515650. \end{aligned}$$

Προφανώς, σε μια προσέγγιση ενός προβλήματος αρχικών τιμών, η συνάρτηση  $F$  στην (6.53) θα εξαρτάται από το  $h$  και την  $f(x, y)$ . Για τον ορισμό της αριθμητικής ευστάθειας δεν ελαττώνουμε το  $h$  όταν  $n \rightarrow \infty$  όπως στην περίπτωση της σειράς Taylor. Καθώς μιας προσέγγιση της ΔΕ από μία

ΕΔ εξαρτάται από τη μορφή της  $f(x, y)$ , είναι πολύ δύσκολο να εξάγουμε γενικά συμπεράσματα όσον αφορά την αριθμητική ευστάθεια. Παρακάτω θα εξετάσουμε λεπτομερώς τη ΔΕ  $y' = -\lambda y$  με  $y(0) = 1$ ,  $\lambda > 0$ , της οποίας η λύση είναι η  $y(x) = e^{-\lambda x}$ . Ο τρόπος διερεύνησης αυτής της ΔΕ και τα αποτελέσματα του μπορούν να εφαρμοστούν και σε άλλα προβλήματα αρχικών τιμών.

Για τη λύση της ΔΕ θα χρησιμοποιήσουμε τη μέθοδο που δίνεται από τον τύπο

$$y_{n+1} = y_{n-1} + 2hf_n . \quad (6.54)$$

Καθώς  $f_n = -\lambda y_n$  η παραπάνω σχέση γίνεται

$$y_{n+1} + 2\lambda h y_n - y_{n-1} = 0 ,$$

της οποίας η γενική λύση είναι

$$y_n = c_1 \beta_1^n + c_2 \beta_2^n ,$$

με  $\beta_{1,2} = -\lambda h \pm \sqrt{1 + \lambda^2 h^2}$ .

Από την παραπάνω σχέση βλέπουμε ότι η  $\beta_1^n$  συμπεριφέρεται σαν την  $y(x)$ , όταν  $h \rightarrow 0$  με  $x_n$  σταθερό, ενώ η  $\beta_2^n$  έχει τελείως διαφορετική συμπεριφορά. Πράγματι, όταν αναπτύξουμε την τετραγωνική ρίζα έχουμε

$$\beta_1 = -\lambda h + 1 + \frac{1}{2}(\lambda h)^2 + O(h^4) .$$

Συγκρίνοντας με την

$$e^{-\lambda h} = 1 - \lambda h + \frac{1}{2}(\lambda h)^2 + O(h^3) ,$$

βλέπουμε ότι

$$\beta_1 = e^{-\lambda h} + O(h^3) = e^{-\lambda h} [1 + O(h^3)] .$$

Η τελευταία ισότητα προκύπτει από τη σχέση  $e^{-\lambda h} = 1 + O(h)$ . Άρα

$$\beta_1^n = e^{-\lambda n h} [1 + O(h^3)]^n = e^{-\lambda x_n} [1 + nO(h^3)] = e^{-\lambda x_n} [1 + O(h^2)] ,$$

και  $x_n = nh$ . Όταν  $h \rightarrow 0$  με το  $x_n$  σταθερό τότε

$$\beta_1^n = e^{-\lambda x_n} = y(x_n) ,$$

δηλαδή, για μικρές τιμές του  $h$  η  $\beta_1^n$  συμπεριφέρεται σαν μια καλή προσέγγιση της  $y(x)$ . Από την άλλη πλευρά, όμως,

$$\begin{aligned} \beta_2 &= -\lambda h - \sqrt{1 + \lambda^2 h^2} = -\lambda h - 1 - \frac{1}{2}(\lambda h)^2 + O(h^4) \\ &= -e^{\lambda h} + O(h^3) = -e^{\lambda h} [1 + O(h^3)] \end{aligned}$$

Άρα

$$\beta_2^n = (-1)^n e^{\lambda x_n} [1 + O(h^2)] .$$

Όταν  $h \rightarrow 0$  με το  $x_n$  σταθερό η  $\beta_2^n$  δεν προσεγγίζει το  $y(x)$ . Γενικά, η λύση της (6.54) είναι

$$y_n = c_1 e^{-\lambda x_n} [1 + O(h^2)] + c_2 (-1)^n e^{\lambda x_n} [1 + O(h^2)] . \quad (6.55)$$

Για να είναι αυτή η τιμή του  $y_n$  καλή προσέγγιση του  $y(x_n)$  θα πρέπει  $c_1 = 1$ ,  $c_2 = 0$ , δηλαδή, πρέπει να υπολογίσουμε τη λύση  $y_n = \beta_1^n$ . Στην πράξη, όμως, επειδή δεν μπορούμε να αποφύγουμε τα λάθη στρογγύλευσης, υπολογίζουμε μία τιμή  $y_n^*$  σαν προσέγγιση του  $y_n$

$$y_n^* = (1 + \delta_1) \beta_1^n + \delta_2 \beta_2^n ,$$

οπότε το σχετικό σφάλμα είναι

$$\left| \frac{y_n - y_n^*}{y_n} \right| = \left| \frac{\delta_1 \beta_1^n + \delta_2 \beta_2^n}{\beta_1^n} \right| = \left| \delta_1 + \delta_2 \left( \frac{\beta_2}{\beta_1} \right)^n \right|.$$

Όταν  $|\beta_2| < |\beta_1|$  το σχετικό σφάλμα είναι φραγμένο και συνεπώς, η αντίστοιχη επαναληπτική μέθοδος είναι ευσταθής. Στην αντίθετη περίπτωση η μέθοδος είναι ασταθής. Αν για το παραπάνω παράδειγμα έχουμε

$$|\beta_2| \equiv \left| \lambda h + \sqrt{1 + \lambda^2 h^2} \right| > \left| \lambda h - \sqrt{1 + \lambda^2 h^2} \right| \equiv |\beta_1|,$$

για  $h$  σταθερό το  $\left( \frac{\beta_2}{\beta_1} \right)^n$  (και το σχετικό σφάλμα) δεν είναι φραγμένο όταν  $n \rightarrow \infty$ .

Η  $\beta_2^n$  λέγεται παρασιτική ή απορριπτή λύση της ΕΔ και εμφανίζεται επειδή έχουμε αντικαταστήσει μια ΔΕ πρώτου βαθμού με μια ΕΔ δεύτερου βαθμού. Στην περίπτωση αυτή, η παρασιτική λύση υπερισχύει της ζητούμενης λύσης.

### Παραδείγματα :

1. Έστω το πρόβλημα αρχικών τιμών  $y' = -3y$ ,  $y(0) = 1$ , το οποίο έχει λύση την  $y(x) = e^{-3x}$ . Αν για την επίλυση αυτής της ΔΕ χρησιμοποιήσουμε τον τύπο (6.54) με  $h = 0.1$  και  $y_0 = y(0) = 1$ ,  $y_1 = 0.7408$  έχουμε τις τιμές του Πίνακα 6.2. Όπως παρατηρούμε, η μέθοδος είναι ασταθής για  $x_n > 1$  και οι εναλλαγές του προσήμου οφείλονται στον όρο  $(-1)^n$  της (6.55).

| $x_n$ | $y_n^*$   | $y_n$    |
|-------|-----------|----------|
| 0.0   | 1.0       | 1.0      |
| 0.1   | 0.7408    | 0.740818 |
| 0.2   | 0.555520  | 0.548812 |
| 0.3   | 0.407488  | 0.406570 |
| 0.4   | 0.311027  | 0.301194 |
| 0.5   | 0.220872  | 0.223130 |
| 0.6   | 0.178504  | 0.165299 |
| 0.7   | 0.113769  | 0.122456 |
| 0.8   | 0.110243  | 0.090718 |
| 0.9   | 0.047624  | 0.067206 |
| 1.0   | 0.081669  | 0.049787 |
| 1.1   | -0.001378 | 0.036883 |
| 1.2   | 0.082495  | 0.027324 |
| 1.3   | -0.050875 | 0.020242 |
| 1.4   | 0.113020  | 0.014996 |
| 1.5   | -0.118687 | 0.011109 |

Πίνακας 6.2: Επίλυση της ΔΕ  $y' = -3y$ ,  $y(0) = 1$ , με εφαρμογή της (6.54) με  $h = 0.1$

2. Θα εξετάσουμε ως προς την ευστάθεια τη μέθοδο Adams-Bashforth δεύτερου βαθμού

$$y_{n+1} = y_n + \frac{h}{2}(3f_n - f_{n-1}),$$

όταν εφαρμοστεί για την επίλυση του προβλήματος αρχικών τιμών  $y' = -\lambda y$ ,  $y(0) = 1$ ,  $\lambda > 0$ . Ο τύπος Adams–Bashforth γράφεται

$$y_{n+1} - \left(1 - \frac{3}{2}\lambda h\right)y_n - \frac{1}{2}\lambda h y_{n-1} = 0. \quad (6.56)$$

Οι ρίζες του χαρακτηριστικού πολυωνύμου του είναι

$$\beta_{1,2} = \frac{1}{2} \left[ \left(1 - \frac{3}{2}\lambda h\right) \pm \sqrt{1 - \lambda h + \frac{9}{4}\lambda^2 h^2} \right].$$

Έχουμε

$$\begin{aligned} \beta_1 &= \frac{1}{2} \left[ \left(1 - \frac{3}{2}\lambda h\right) + 1 + \frac{1}{2} \left(-\lambda h + \frac{9}{4}\lambda^2 h^2\right) - \frac{1}{8} \left(-\lambda h + \frac{9}{4}\lambda^2 h^2\right) + O(h^3) \right] \\ &= 1 - \lambda h + \frac{1}{2}\lambda^2 h^2 + O(h^3) = e^{-\lambda h} \left[ 1 + O(h^3) \right]. \end{aligned}$$

Επομένως,

$$\beta_1^n = e^{-\lambda x_n} \left[ 1 + O(h^2) \right].$$

Από τη σχέση αυτή έχουμε ότι για μικρό  $h$  η  $\beta_1^n$  είναι μια καλή προσέγγιση του  $y(x_n)$ .

Στην πράξη, η λύση της (6.56) είναι της μορφής

$$y_n^* = (1 + \delta_1)\beta_1^n + \delta_2\beta_2^n$$

και

$$\frac{\beta_2}{\beta_1} = \frac{1 - \frac{3}{2}\lambda h - \sqrt{1 - \lambda h + \frac{9}{4}\lambda^2 h^2}}{1 - \frac{3}{2}\lambda h + \sqrt{1 - \lambda h + \frac{9}{4}\lambda^2 h^2}}.$$

Η ποσότητα κάτω από την τετραγωνική ρίζα είναι θετική και συνεπώς για  $1 - 3/2\lambda h \geq 0$  θα ισχύει  $\left|\frac{\beta_2}{\beta_1}\right| \leq 1$ . Το σχετικό σφάλμα θα είναι φραγμένο—και επομένως η μέθοδος είναι ευσταθής—όταν  $h \leq \frac{2}{3\lambda}$ .

Δυσκολίες ευστάθειας δεν υπάρχουν όταν για μια ΔΕ πρώτου βαθμού χρησιμοποιήσουμε μια ΕΔ πρώτου βαθμού καθώς δεν προκύπτει παρασιτική λύση. Υπάρχει, όμως, ένα άλλου τύπου πρόβλημα ευστάθειας που μπορεί να παρουσιαστεί για ΕΔ πρώτου βαθμού.

## 6.11 Απόλυτη Ευστάθεια

Έστω ότι μια ΔΕ αντικαθίσταται από μια ΕΔ βαθμού  $k$ . Τότε η γενική λύση της ΕΔ είναι

$$y_n = c_1\beta_1^n + \dots + c_k\beta_k^n.$$

Αν μια λύση της ΕΔ, έστω η  $\beta_1^n$  αντιπροσωπεύει την πραγματική λύση της ΔΕ, τότε οι άλλες  $k - 1$  ρίζες της ΕΔ θα είναι παρασιτικές. Αν  $|\beta_i| > |\beta_1|$  για  $i \neq 1$ , τότε η παρασιτική λύση  $\beta_i^n$  θα κυριαρχεί στη λύση της ΔΕ και θα έχουμε αστάθεια. Για τη λύση της ΔΕ  $y' = \lambda y$ , η ΕΔ είναι

**Απόλυτα ευσταθής** όταν  $|\beta_i| < 1$  για όλα τα  $i$ ,

**Σχετικά ευσταθής** όταν  $|\beta_1| > |\beta_i|$  για  $i = 2, \dots, k$ .

**Παράδειγμα:** Έστω  $y' = -\lambda y$ ,  $\lambda > 0$  και ενδιαφερόμαστε για την απόλυτη ευστάθεια της μεθόδου Euler όταν εφαρμοστεί για τη λύση αυτής της ΔΕ. Ο τύπος Euler είναι

$$y_{r+1} = (1 - \lambda h)y_r = \beta y_r, \quad \text{με } \beta = 1 - \lambda h.$$

Για να έχουμε απόλυτη ευστάθεια πρέπει να έχουμε  $|\beta| < 1$  ή  $h < 2/\lambda$ . Η συνθήκη αυτή όταν το  $\lambda$  είναι πολύ μεγάλο είναι πολύ δαπανηρή. Π.χ. αν  $y' = -1000y$ ,  $y(0) = 1$ · για απόλυτη ευστάθεια θα πρέπει να έχουμε  $h < 2/1000 = 0.002$ . Για να ολοκληρώσουμε από  $x = 0$  έως  $x = 100$  χρειαζόμαστε τουλάχιστον 50000 βήματα.

**Παρατήρηση:** Οι ανοικτές (explicit) μέθοδοι δεν είναι ποτέ απόλυτα ευσταθείς για όλες τις τιμές του  $h$ . Εξαιρέση αποτελεί ο κανόνας του τραpezίου. Όταν τον εφαρμόσουμε στην  $y' = -\lambda y$ ,  $y(0) = 1$ ,  $\lambda > 0$ , βλέπουμε ότι είναι απόλυτα ευσταθής για όλα τα  $h$ .

## 6.12 Ασκήσεις

1. Εφαρμόστε τη μέθοδο Euler για την επίλυση της ΔΕ

$$y' = \cos x - \sin y + x^2$$

στο διάστημα  $[-1, 1]$ , με  $y(-1) = 3.0$ . Τυπώστε τη λύση ανά  $h = 0.01$ .

2. Εφαρμόστε τη μέθοδο Taylor με 5 όρους για την επίλυση της ΔΕ της προηγούμενης άσκησης.
3. Να λυθεί το σύστημα

$$\begin{aligned} y_1' &= 2y_1 - 2y_2 + 3y_3 \\ y_2' &= y_1 + y_2 + y_3 \\ y_3' &= y_1 + 3y_2 - y_3 \end{aligned}$$

με αρχικές συνθήκες (στο  $t = 0$ )  $y_1 = -2$ ,  $y_2 = 30$ ,  $y_3 = 0$ . Δίνεται ότι

$$\begin{bmatrix} 1 & 11 & 1 \\ -1 & 1 & 1 \\ -1 & -14 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1/2 & -5/6 & 1/3 \\ 0 & 1/15 & -1/15 \\ 1/2 & 1/10 & 2/5 \end{bmatrix}.$$

4. Να γράψετε κώδικα για την επίλυση της ΔΕ  $y' = -y$ , με  $y(0) = 1$  στο διάστημα  $[0, 1]$  με τη μέθοδο Euler και βήματα  $h = 0.2, 0.02, 0.002, 0.0002, 0.00002$ .
5. Να δείξετε ότι η μέθοδος Runge-Kutta 3<sup>ου</sup> βαθμού για την επίλυση της ΔΕ  $y' = f(x, y)$ ,  $y(x_0) = y_0$  είναι:

$$\begin{aligned} y_{r+1} &= y_r + \frac{1}{6}(k_1 + 4k_2 + k_3) \\ k_1 &= hf(x_r, y_r) \\ k_2 &= hf(x_r + 0.5h, y_r + 0.5k_1) \\ k_3 &= hf(x_r + h, y_r + k_2) \end{aligned}$$

6. Εφαρμόστε τη μέθοδο Taylor με 4 όρους για την επίλυση του συστήματος ΔΕ

$$\begin{aligned}y' &= y + z^2 - x^3 \\z' &= z + y^3 + \cos x\end{aligned}$$

με αρχικές συνθήκες (στο  $x = 0$ )  $y = 0.3$ ,  $z = 0.1$ . Τυπώστε τις τιμές των  $y, z$  στο διάστημα  $[0, 1]$  με βήμα 0.1.

7. Να γράψετε κώδικα σε Fortran που να υλοποιεί τη μέθοδο Runge-Kutta 2<sup>ου</sup> βαθμού.
8. Δώστε έναν αλγόριθμο για τον έλεγχο του σφάλματος για τη Runge-Kutta βαθμού  $p$ . Διπλασιάστε το βήμα  $h$  αν το σφάλμα είναι “πολύ μικρό” ή υποδιπλασιάστε το αν το σφάλμα είναι “πολύ μεγάλο”.
9. Δείξτε ότι όταν εφαρμοστεί η Runge-Kutta 2<sup>ου</sup> βαθμού σε ΔΕ της μορφής  $y' = \lambda x$  δίνει την ακριβή λύση.
10. Δώστε ένα αλγόριθμο για τη μέθοδο Taylor τέταρτου βαθμού και υλοποιήστε τον σε κώδικα Fortran. Εφαρμόστε τον για τη ΔΕ  $y' = y^3 + x + y$ , στο  $[0, 0.5]$ , με  $y(0) = 1$ ,  $h = 0.1$ .
11. Να λυθεί η ΔΕ  $y' = x^2 + x - y$ ,  $y(0) = 0$  με τη Runge-Kutta 2<sup>ου</sup> βαθμού. Υπολογίστε τη λύση στο  $x = 0.6$  με  $h = 0.2$ .
12. Το πρόβλημα αρχικών τιμών  $y' = ax + b$ ,  $y(0) = 0$ , έχει λύση

$$y(x) = \frac{a}{2}x^2 + bx.$$

Αν εφαρμοστεί η μέθοδος Euler δείξτε ότι η ΕΔ που προκύπτει έχει λύση

$$y_n = \frac{1}{2}(ax_n + 2b - ah)x_n,$$

όπου  $x_r = rh$ , και, επομένως,

$$y(x_n) - y_n = \frac{a}{2}hx_n.$$

13. Να εφαρμόσετε την Runge-Kutta 2<sup>ου</sup> βαθμού για δύο διαδοχικά διαστήματα  $h$ , για τη ΔΕ  $y' = \lambda y$ ,  $y(0) = 1$ . Να δείξετε ότι ο κυρίαρχος όρος του σφάλματος αποκοπής στο  $y(2h)$  είναι  $(\lambda h)^3/3$ .
14. Χρησιμοποιήστε τη μέθοδο Runge-Kutta 4<sup>ου</sup> βαθμού για να επιλύσετε τη ΔΕ

$$y' = \frac{y}{x} \left(1 - \frac{y}{x}\right)$$

στο διάστημα  $[1, 3]$ , με  $y(1) = 2$ . Τυπώστε τις τιμές με βήμα  $h = 1/128$ , καθώς και το σφάλμα ως προς την ακριβή λύση

$$y(x) = \frac{x}{0.5 + \ln x}.$$

15. Να εφαρμόσετε την Runge-Kutta 2<sup>ου</sup> βαθμού για την εύρεση της κίνησης σώματος μάζας  $m = 2$  Kg, εξαρτώμενου από ελατήριο με δύναμη επαναφοράς  $F(x) = x^2 + 0.01x^3$ . Το σώμα αφήνεται για  $t = 0$  ελεύθερο, χωρίς αρχική ταχύτητα, στη θέση  $x = 2.5$  cm.

16. Να βρείτε τη κίνηση εκκρεμούς για το οποίο ισχύει

$$\ddot{\theta} = -\sin \theta ,$$

όπου  $\theta$  η γωνία απομάκρυνσης από την κάθετο. Το εκκρεμές αφήνεται ελεύθερο, χωρίς αρχική ταχύτητα σε γωνία  $\theta = 45^\circ$ .

Για μικρές γωνίες  $\theta$  ισχύει  $\sin \theta \approx \theta$ . Εφαρμόστε την προσέγγιση αυτή και συγκρίνετε τη λύση της νέας διαφορικής εξίσωσης με τη λύση της ακριβούς ΔΕ.

17. Να λύσετε τη ΔΕ  $\psi'' = (x^2 - 5)\psi$  με αρχική συνθήκη  $\psi(0) = -(2\sqrt{\pi})^{-1/2}$ . Τυπώστε 100 ισαπέχουσες τιμές στο διάστημα  $[-2, 2]$ .

**Υπόδειξη** Να λύσετε δύο προβλήματα αρχικών τιμών, τη ΔΕ στα διαστήματα  $[0, 2]$  και  $[-2, 0]$ .

18. Να λυθούν οι ΕΔ

- $y_{n+1} - ay_n = 0$ ,
- $y_{n+2} - 4y_{n+1} + 4y_n = 0$ ,
- $y_{n+2} - 4y_{n+1} + 4y_n = 1$ .

19. Να δειχθούν οι σχέσεις:

- $\Delta = E - 1$ ,
- $E = e^{hD}$ ,
- $\nabla = \frac{E - 1}{E}$ ,
- $\delta = E^{1/2} - E^{-1/2}$ ,
- $\mu = 0.5(E^{1/2} + E^{-1/2})$ ,
- $\mu = \cosh\left(\frac{hD}{2}\right)$ ,
- $E = 1 + \mu\delta + \frac{\delta^2}{2}$ ,
- $\mu^2 = 1 + \frac{\delta^2}{4}$ .

20. Να λυθούν οι ΕΔ

- $y_{r+2} - 5y_{r+1} + 6y_r = 0, \quad y_0 = 0, y_1 = 1$ ,
- $y_{r+2} - 4y_{r+1} + 4y_r = 0, \quad y_0 = 1, y_1 = 6$ ,
- $y_{r+2} + 6y_{r+1} + 25y_r = 0, \quad y_0 = 0, y_1 = 4$ .

21. Να δειχτεί ότι η γενική λύση της ΕΔ

$$y_{r+2} - (2 + h^2)y_{r+1} + y_r = h^2, \quad h > 0,$$

είναι

$$y_n = c_1 \left[ 1 + h + \frac{h^2}{2} + O(h^3) \right]^n + c_2 \left[ 1 - h + \frac{h^2}{2} + O(h^3) \right]^n - 1.$$

22. Να δειχτεί ότι η γενική λύση της ΕΔ

$$y_{r+2} + 4hy_{r+1} - y_r = 2h, \quad h > 0,$$

είναι

$$y_n = c_1 \left[ 1 - 2h + O(h^2) \right]^n + c_2 (-1)^n \left[ 1 + 2h + O(h^2) \right]^n + \frac{1}{2}.$$

23. Να δειχτεί ότι

$$\frac{h}{2} (y'_r + y'_{r+1}) = \left( \delta + \frac{1}{12} \delta^3 - \frac{1}{120} \delta^5 + \dots \right) y_{r+\frac{1}{2}}.$$

Δίνεται ότι

$$\sinh^{-1} z = z - \frac{1}{6} z^3 + \frac{3}{4} z^5 + \dots.$$

24. Να αποδείξετε ότι το

$$y_0 - y_1 + y_2 - \dots = \sum_{r=0}^{\infty} (-1)^r E^r y_0$$

είναι ίσο με

$$\frac{1}{2} y_0 - \frac{1}{4} \Delta y_0 + \frac{1}{8} \Delta^2 y_0 - \dots.$$

25. Να υλοποιήσετε σε κώδικα τον αλγόριθμο Adams-Bashforth 3<sup>ου</sup> βαθμού και το αντίστοιχο σφάλμα αποκοπής.

26. Να υπολογίσετε μια προσέγγιση του  $y(1.0)$  για τη ΔΕ  $y' = 1 - y$ ,  $y(0) = 0$ , χρησιμοποιώντας τον αλγόριθμο Adams-Bashforth 2<sup>ου</sup> βαθμού με  $h = 0.2$  και αρχικές τιμές  $y(0) = 0$ ,  $y_1 = y(0.2) = 0.18127$ . Συγκρίνετε τα αποτελέσματα με την αναλυτική λύση  $y(x) = 1 - e^{-x}$ .

27. Υλοποιήστε σε κώδικα Fortran τον αλγόριθμο Πρόβλεψης-Διόρθωσης στον οποίο το  $h$  να μεταβάλλεται αυτόματα υπολογίζοντας κάθε φορά το σφάλμα αποκοπής  $T_c$ .

28. Να εξετάσετε ως προς την ευστάθεια τις παρακάτω μεθόδους για τη ΔΕ  $y' = -\lambda y$ ,  $\lambda > 0$ .

(α) Adams-Moulton 3<sup>ου</sup> βαθμού

$$y_{n+1} = y_n + \frac{h}{12} (5y'_{n+1} + 8y'_n - y'_{n-1}),$$

(β) Milne-Simpson

$$y_{n+1} = y_{n-1} + \frac{h}{3} (y'_{n+1} + 4y'_n + y'_{n-1}),$$

(γ)

$$y_{n+1} = y_{n-1} + \frac{h}{2} (y'_n + 3y'_{n-1}).$$

Να δείξετε ότι το απόλυτο σφάλμα για αρκετά μικρό  $h$  συνεχώς μικραίνει παρόλο που η συγκεκριμένη μέθοδος είναι ασταθής.

29. Έστω η ΔΕ  $y' = -\lambda y$ ,  $\lambda > 0$ . Να δείξετε ότι  $y_n \rightarrow 0$  όταν  $n \rightarrow \infty$  και  $h$  σταθερό (δηλαδή είναι απόλυτα ευσταθής), αν

- $|1 - \lambda h + \frac{1}{2}\lambda^2 h^2| < 1$  για τη Runge-Kutta 2<sup>ου</sup> βαθμού,
- $h < 1/\lambda$  για τη Adams-Bashforth 2<sup>ου</sup> βαθμού.

30. Να δείξετε ότι η μέθοδος

$$y_{n+1} = y_{n-1} + \frac{h}{2} (y'_{n+1} + 2y'_n + y'_{n-1})$$

είναι ασταθής για τη ΔΕ  $y' = -\lambda y$  και ευσταθής για την  $y' = -\lambda y + \mu$ , όπου  $\lambda > 0$  και  $\mu \neq 0$ .

31. Να περιγράψετε τις ΕΔ για τη μέθοδο Taylor 2<sup>ου</sup> βαθμού, όταν εφαρμοστεί στο σύστημα

$$\begin{aligned} y'_1 &= x^2 y_1 - y_2, & y_1(0) &= s_1, \\ y'_2 &= -y_1 + x y_2, & y_2(0) &= s_2. \end{aligned}$$

32. Να λυθεί η ΔΕ

$$y'' + 4xyy' + 2y^2 = 0, \quad y(0) = 1, \quad y'(0) = 0, \quad (6.57)$$

με τη μέθοδο Euler και με  $h = 0.1$ . Να υπολογιστούν τα  $y(0.5)$ ,  $y'(0.5)$ .

33. Να λυθεί η (6.57) με μια μέθοδο πρόβλεψης-διόρθωσης 2<sup>ου</sup> βαθμού. Να γίνει περιγραφή των ΕΔ και του αλγορίθμου μόνο.

34. Έστω η ΔΕ  $y' - 2y$ ,  $y(a) = y_0$ . Να εφαρμοστεί η μέθοδος Euler και να βρεθεί η έκφραση του  $y_k$  ως συνάρτηση του  $y_0$ . Να εξεταστεί η συμπεριφορά της μεθόδου όταν  $h \rightarrow 0$ .

35. Να εφαρμοστεί δευτεροβάθμια μέθοδος πρόβλεψης-διόρθωσης για τη ΔΕ  $y' = xy^2$ , στο  $[0, 0.2]$ , με  $y(0) = 0.25$  και  $h = 0.1$ .

36. Να βρείτε όρια του  $\lambda h$  για μια ευσταθή ολοκλήρωση της ΔΕ  $y' = \lambda y$ ,  $y(0) = 1$  με τη μέθοδο Runge-Kutta 3<sup>ου</sup> βαθμού.

37. Να δείξετε ότι το σφάλμα αποκοπής στη μέθοδο τραπεζίου είναι

$$-\frac{h^3}{12} y'''(\xi), \quad \xi \in (x_r, x_{r+1}).$$

38. Να διερευνήσετε την ευστάθεια της Adams-Moulton 3<sup>ου</sup> βαθμού

$$y_{r+1} = y_r + \frac{h}{12} (5y'_{r+1} + 8y'_r - y'_{r-1}).$$

39. Να δείξετε ότι η προσέγγιση της παραγώγου μιας συνάρτησης από τον τύπο των κεντρικών διαφορών είναι καλύτερη από την προσέγγιση με τον τύπο διαφορών προς τα εμπρός.

40. Να εξεταστεί ως προς την ευστάθεια ο κανόνας τραπεζίου για τη ΔΕ  $y' = \lambda y$ .

# Κατάλογος Πινάκων

|     |                                                                                                                                                                 |    |
|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 2.1 | Ακολουθίες των διαστημάτων, της προσεγγιστικής ρίζας και της αντίστοιχης τιμής της $f(x) = x^3 + 4x^2 - 10$ κατά την εφαρμογή της μεθόδου διχοτόμησης . . . . . | 8  |
| 6.1 | Πίνακας διαφορών της συνάρτησης $f(x) = e^x$ με ακρίβεια 6 δεκαδικών ψηφίων . .                                                                                 | 59 |
| 6.2 | Επίλυση της ΔΕ $y' = -3y$ , $y(0) = 1$ , με εφαρμογή της (6.54) με $h = 0.1$ . . . . .                                                                          | 74 |