

Κώστας Δ. Κόκκοτας

Εισαγωγή στην Αριθμητική Ανάλυση με Εφαρμογές στη Φυσική

– Σημειώσεις για τους φοιτητές –

13 Φεβρουαρίου 2008

Περιεχόμενα

| | | |
|----------|--|-----------|
| 1 | ΡΙΖΕΣ ΜΗ-ΓΡΑΜΜΙΚΩΝ ΕΞΙΣΩΣΕΩΝ | 1 |
| 1.1 | ΜΕΘΟΔΟΣ ΔΙΧΟΤΟΜΗΣΗΣ (BOLZANO) | 2 |
| 1.2 | Η ΜΕΘΟΔΟΣ ΤΗΣ ΓΡΑΜΜΙΚΗΣ ΠΑΡΕΜΒΟΛΗΣ | 4 |
| 1.3 | Η ΜΕΘΟΔΟΣ ΤΟΥ MULLER | 10 |
| 1.4 | Η ΜΕΘΟΔΟΣ $x = g(x)$ | 11 |
| 1.4.1 | Βελτίωση Aitken | 14 |
| 1.5 | ΜΕΘΟΔΟΣ NEWTON - RAPHSON | 15 |
| 1.5.1 | Δεύτερης τάξης Newton-Raphson (Halley) | 18 |
| 1.5.2 | Πολλαπλές ρίζες | 19 |
| 1.6 | ΣΥΣΤΗΜΑΤΑ ΜΗ-ΓΡΑΜΜΙΚΩΝ ΕΞΙΣΩΣΕΩΝ | 22 |
| 1.6.1 | Η Μέθοδος Newton | 23 |
| 1.6.2 | Μέθοδος τύπου $x = g(x)$ | 25 |
| 1.7 | ΑΣΚΗΣΕΙΣ | 28 |
| 2 | ΓΡΑΜΜΙΚΑ ΣΥΣΤΗΜΑΤΑ | 31 |
| 2.1 | ΜΕΘΟΔΟΣ GAUSS | 32 |
| 2.2 | ΜΕΘΟΔΟΣ GAUSS-JORDAN | 35 |
| 2.3 | ΜΕΘΟΔΟΣ L-U | 36 |
| 2.4 | ΜΕΘΟΔΟΣ JACOBI (Επαναληπτική) | 39 |
| 2.5 | ΜΕΘΟΔΟΣ GAUSS - SEIDEL | 40 |
| 2.6 | Η ΟΡΙΖΟΥΣΑ ΚΑΙ Ο ΑΝΤΙΣΤΡΟΦΟΣ ΕΝΟΣ ΠΙΝΑΚΑ | 42 |
| 2.6.1 | Η ορίζουσα | 42 |
| 2.6.2 | Ο αντίστροφος πίνακας | 42 |
| 2.7 | ΙΔΙΟΤΙΜΕΣ ΚΑΙ ΙΔΙΟΔΙΑΝΥΣΜΑΤΑ | 43 |
| 2.7.1 | Η μέθοδος των δυνάμεων | 44 |
| 2.7.2 | Επιτάχυνση Aitken | 46 |
| 2.7.3 | Αντίστροφη μέθοδος των δυνάμεων | 46 |
| 2.7.4 | Μέθοδος της μετάθεσης | 47 |
| 2.8 | ΑΣΚΗΣΕΙΣ | 48 |

| | | |
|----------|---|-----|
| 3 | ΠΑΡΕΜΒΟΛΗ & ΠΡΟΒΛΕΨΗ | 49 |
| 3.1 | ΠΟΛΥΩΝΥΜΟ LAGRANGE | 50 |
| 3.1.1 | Απόδειξη του τύπου Lagrange | 51 |
| 3.1.2 | Σφάλμα του τύπου Lagrange | 52 |
| 3.2 | ΣΥΜΠΤΩΤΙΚΟ ΠΟΛΥΩΝΥΜΟ ΓΙΑ ΙΣΑΠΕΧΟΝΤΑ x_i | 54 |
| 3.2.1 | Τύποι για ισαπέχοντα x_i | 56 |
| 3.3 | ΕΦΑΠΤΟΜΕΝΟ ΠΟΛΥΩΝΥΜΟ (HERMITE) | 57 |
| 3.4 | ΠΟΛΥΩΝΥΜΟ TAYLOR | 58 |
| 3.5 | ΠΑΡΕΜΒΟΛΗ ΜΕ SPLINES | 59 |
| 3.6 | ΑΣΚΗΣΕΙΣ | 65 |
| 4 | ΑΡΙΘΜΗΤΙΚΗ ΠΑΡΑΓΩΓΙΣΗ | 69 |
| 4.1 | ΑΡΙΘΜΗΤΙΚΗ ΠΑΡΑΓΩΓΙΣΗ ΜΕ ΧΡΗΣΗ ΣΥΜΠΤΩΤΙΚΟΥ ΠΟΛΥΩΝΥΜΟΥ | 69 |
| 4.2 | ΤΥΠΟΙ ΚΕΝΤΡΙΚΩΝ ΔΙΑΦΟΡΩΝ | 72 |
| 5 | ΑΡΙΘΜΗΤΙΚΗ ΟΛΟΚΛΗΡΩΣΗ | 75 |
| 5.1 | ΤΥΠΟΙ NEWTON-COTES | 75 |
| 5.1.1 | Κανόνας τραπεζίου | 78 |
| 5.1.2 | Κανόνας του Simpson | 80 |
| 5.1.3 | Κανόνας του Simpson(3/8) | 80 |
| 5.1.4 | Η βελτίωση του Romberg | 80 |
| 5.1.5 | Εφαρμογή των Splines στην Παραγωγή και Ολοκλήρωση συναρτήσεων | 82 |
| 5.2 | ΜΕΘΟΔΟΣ ΠΡΟΣΔΙΟΡΙΣΤΕΩΝ ΣΥΝΤΕΛΕΣΤΩΝ | 84 |
| 5.2.1 | Μέθοδος του Filon | 86 |
| 5.3 | ΜΕΘΟΔΟΣ GAUSS | 88 |
| 5.3.1 | Μέθοδος Gauss-Legendre | 89 |
| 5.3.2 | Γενίκευση της μεθόδου Gauss | 91 |
| 5.4 | ΑΣΚΗΣΕΙΣ | 93 |
| 6 | ΔΙΑΦΟΡΙΚΕΣ ΕΞΙΣΩΣΕΙΣ | 95 |
| 6.1 | ΜΕΘΟΔΟΙ ΕΝΟΣ ΒΗΜΑΤΟΣ | 95 |
| 6.1.1 | Μέθοδος Σειρών Taylor | 96 |
| 6.1.2 | Μέθοδοι Euler & Euler - Heun | 97 |
| 6.1.3 | Σφάλματα και Ευστάθεια | 98 |
| 6.1.4 | Μέθοδος Runge-Kutta | 100 |
| 6.2 | ΜΕΘΟΔΟΙ ΠΟΛΛΑΠΛΩΝ ΒΗΜΑΤΩΝ | 103 |
| 6.2.1 | Μέθοδος Adams | 103 |
| 6.2.2 | Μέθοδος Milne | 105 |
| 6.3 | ΜΕΘΟΔΟΙ ΠΡΟΒΛΕΨΗΣ - ΔΙΟΡΘΩΣΗΣ | 105 |
| 6.4 | ΣΥΣΤΗΜΑΤΑ ΔΙΑΦΟΡΙΚΩΝ ΕΞΙΣΩΣΕΩΝ | 107 |
| 6.5 | ΑΣΚΗΣΕΙΣ | 109 |

| | | |
|-----------|---|-----|
| 7 | ΠΡΟΣΕΓΓΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ | 111 |
| 7.1 | ΜΕΘΟΔΟΣ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ | 111 |
| 7.1.1 | Ευθεία ελαχίστων τετραγώνων | 111 |
| 7.1.2 | Παραβολή ελαχίστων τετραγώνων | 114 |
| 7.1.3 | Συνεχή δεδομένα | 115 |
| 7.2 | ΠΡΟΣΕΓΓΙΣΗ ΕΛΑΧΙΣΤΟΥ – ΜΕΓΙΣΤΟΥ | 117 |
| 7.3 | ΠΡΟΣΕΓΓΙΣΗ ΜΕ ΡΗΤΕΣ ΣΥΝΑΡΤΗΣΕΙΣ | 119 |
| 7.3.1 | Προσέγγιση με ρητή συνάρτηση για ομάδα σημείων | 121 |
| 7.4 | ΑΣΚΗΣΕΙΣ | 122 |
| 8 | ΔΙΑΦΟΡΙΚΕΣ ΕΞΙΣΩΣΕΙΣ ΜΕ ΜΕΡΙΚΕΣ ΠΑΡΑΓΩΓΟΥΣ | 125 |
| 8.1 | ΕΛΛΕΙΠΤΙΚΕΣ ΕΞΙΣΩΣΕΙΣ | 126 |
| 8.1.1 | Η Εξίσωση Laplace | 126 |
| 8.2 | ΥΠΕΡΒΟΛΙΚΕΣ ΕΞΙΣΩΣΕΙΣ | 131 |
| 8.2.1 | Αριθμητική Επίλυση της Κυματικής Εξίσωσης | 131 |
| 8.3 | ΠΑΡΑΒΟΛΙΚΕΣ ΕΞΙΣΩΣΕΙΣ | 136 |
| 8.3.1 | Αριθμητική Λύση: Εξίσωση Διαφορών | 136 |
| 8.3.2 | Αριθμητική Λύση: Μέθοδος Crank-Nicholson | 138 |
| Α' | ΑΝΑΛΥΣΗ ΣΦΑΛΜΑΤΟΣ | 141 |
| Α'.0.1 | Σφάλμα αποκοπής | 142 |
| Α'.0.2 | Σφάλμα στρογγύλευσης | 143 |
| Α'.0.3 | Απώλεια σημαντικών ψηφίων | 143 |
| Α'.0.4 | Προσέγγιση τάξης $O(h^n)$ | 144 |
| Α'.0.5 | Διάδοση σφάλματος | 144 |
| Α'.0.6 | Αβεβαιότητα στα δεδομένα | 145 |
| | Ευρετήριο | 147 |

ΡΙΖΕΣ ΜΗ-ΓΡΑΜΜΙΚΩΝ ΕΞΙΣΩΣΕΩΝ

Κλασικό πρόβλημα των στοιχειωδών μαθηματικών είναι η εύρεση μιας τιμής ρ τέτοιας, ώστε για μια συνάρτηση $f(x)$, $x \in (a, b)$ να ισχύει:

$$f(\rho) = 0 \quad (1.1)$$

Η βασική διαδικασία για την εύρεση ριζών μη-γραμμικών εξισώσεων είναι η δημιουργία μιας αναδρομικής σχέσης. Η διαδικασία αυτή θα ακολουθηθεί και στην εύρεση ριζών γραμμικών και μη-γραμμικών συστημάτων, αλλά και στην αριθμητική επίλυση διαφορικών εξισώσεων. Επομένως, σε κάθε μέθοδο που θα αναπτύξουμε, στόχος μας είναι η εύρεση αναδρομικών σχέσεων της μορφής

$$x_{n+1} = \sigma(x_n) \quad (1.2)$$

που θα δίνει μια ακολουθία τιμών $x_0, x_1, \dots, x_n, \dots$ και στο όριο $n \rightarrow \infty$ να δίνει τη ρίζα της εξίσωσης (1.1).

Τα βασικά ερωτήματα που καλούμαστε να απαντήσουμε κατά την ανάπτυξη των διαφόρων μεθόδων είναι:

- Κάτω από ποιες συνθήκες συγκλίνει μια μέθοδος
- Αν συγκλίνει, πόσο γρήγορα συγκλίνει
- Πώς θα βρεθεί η αρχική τιμή x_0 .

ΠΑΡΑΔΕΙΓΜΑ

Ένα κλασικό λογιστικό πρόβλημα είναι το εξής:

Εάν ένας καταθέτης τοποθετεί κάθε μήνα στο τραπεζικό λογαριασμό του ένα σταθερό κεφάλαιο K , τότε πόσος χρόνος απαιτείται, ώστε το κεφάλαιο του να γίνει F δοθέντος ότι το επιτόκιο είναι x .

Ο μαθηματικός κανόνας που χρησιμοποιούν οι τράπεζες είναι:

$$F = K \cdot \frac{12}{x} \left[\left(1 + \frac{x}{12} \right)^N - 1 \right] \quad (1.3)$$

όπου N είναι ο αριθμός των μηνών.

Εύκολα μπορεί κανείς να οδηγηθεί στη λύση του προβλήματος, αν λογαριθμήσει κατάλληλα τη σχέση (1.3). Θα μπορούσε όμως να ακολουθήσει και μια διαφορετική διαδικασία, δηλαδή να δοκιμάσει διάφορες τιμές του N , έως ότου επιτύχει το απαιτούμενο αποτέλεσμα, δηλαδή μετά από K προσπάθειες να οδηγηθεί στην τιμή N_K , για την οποία

$$f(N_K) = F - \frac{12K}{x} \left[\left(1 + \frac{x}{12}\right)^{N_K} - 1 \right] = 0 \quad (1.4)$$

επομένως, ουσιαστικά να αναζητεί τη ρίζα της εξίσωσης

$$f(N) = 0. \quad (1.5)$$

Το πρόβλημα, όπως το περιγράψαμε, ίσως δεν απαιτεί τη βοήθεια των μεθόδων εύρεσης ριζών που θα περιγράψουμε στη συνέχεια. Όμως, έστω ότι διαφοροποιούμε τη λογική του, δηλαδή θεωρούμε ότι ο καταθέτης γνωρίζει για πόσους μήνες θα καταθέτει χρήματα και αναζητά το κατάλληλο επιτόκιο. Σε αυτή την περίπτωση, αναγόμεναι στην εύρεση της ρίζας της εξίσωσης

$$f(x) = 0. \quad (1.6)$$

Οι μέθοδοι, που θα αναπτυχθούν στη συνέχεια, θα μας βοηθήσουν να επιλύσουμε το πρόβλημα με ακρίβεια και ταχύτητα.

1.1 ΜΕΘΟΔΟΣ ΔΙΧΟΤΟΜΗΣΗΣ (BOLZANO)

Στο προηγούμενο παράδειγμα, ας προσπαθήσουμε να εφαρμόσουμε μια απλοϊκή ίσως διαδικασία. Υποθέτουμε ότι το κεφάλαιο που κατατίθεται κάθε μήνα είναι €1000 και το ποσό που έχουμε στόχο να συλλέξουμε σε διάστημα 50 μηνών είναι €65000.

Αν δοκιμάσουμε λοιπόν για δυο τυχαίες (αλλά λογικές) τιμές του επιτοκίου, έστω $x_1 = 0.10$ και $x_2 = 0.15$, θα βρούμε:

$$\begin{aligned} f(0.10) &= 3286.4 \\ f(0.15) &= -3881.8 \end{aligned}$$

που σημαίνει ότι με επιτόκιο 0.10 (δηλαδή 10% ετησίως), υπολειπόμεθα του στόχου μας κατά € 3286.4, ενώ με επιτόκιο 0.15 τον ξεπερνούμε κατά € 3881.8. Επομένως, η φυσική μας επιλογή θα είναι να δοκιμάσουμε για επιτόκιο 0.125, δηλαδή μια τιμή στο μέσο του διαστήματος (0.1, 0.15). Για επιτόκιο $x_3 = 0.125$ βρίσκουμε ότι:

$$f(0.125) = -174.5$$

άρα απέχουμε από το στόχο μας μόνο κατά €174.5 ευρώ, ενώ φαίνεται πλέον πως η σωστή τιμή του επιτοκίου βρίσκεται στο διάστημα (0.10, 0.125). Επομένως, επιλέγοντας μια τιμή στο μέσο του διαστήματος, έστω $x_4 = 0.1125$, βρίσκουμε ότι

$$f(0.1125) = 1585.6.$$

Επομένως η σωστή τιμή βρίσκεται στο διάστημα $(0.1125, 0.125)$, άρα μπορούμε να επιλέξουμε ένα νέο επιτόκιο, που θα είναι $x_5 = (0.1125 + 0.125)/2$, και να επαναλάβουμε την παραπάνω διαδικασία. Μετά από μερικές ακόμη προσπάθειες, θα καταλήξουμε στην τιμή του κεφαλαίου που επιθυμούμε να συγκεντρώσουμε μετά από καταθέσεις 50 μηνών.

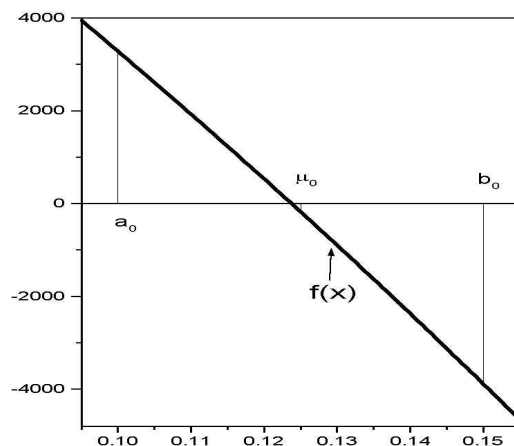
Η ‘απλοϊκή’ μέθοδος που χρησιμοποιήσαμε στο προηγούμενο παράδειγμα αποτελεί μια από τις κλασικές μεθόδους εύρεσης ριζών εξισώσεων και ονομάζεται **μέθοδος διχοτόμησης** ή **μέθοδος του Bolzano**. Η παραπάνω διαδικασία μπορεί να κωδικοποιηθεί ως ακολούθως :

Έστω ότι μια ρίζα βρίσκεται στο διάστημα $[a_0, b_0]$, τότε $f(a_0) \cdot f(b_0) < 0$. Αν $\mu_0 = (a_0 + b_0)/2$, τότε :

- (I) είτε $f(\mu_0) \cdot f(a_0) < 0$
- (II) είτε $f(\mu_0) \cdot f(b_0) < 0$
- (III) είτε $f(\mu_0) = 0$.

Αν ισχύει η (III), τότε έχει υπολογισθεί η ρίζα και σταματά η διαδικασία, αλλιώς ορίζω νέο διάστημα

$$[a_1, b_1] = \begin{cases} [\mu_0, b_0] & \text{αν (II)} \\ [a_0, \mu_0] & \text{αν (I)} \end{cases} \quad (1.7)$$



Σχήμα 1.1. Γραφική απεικόνιση της διαδικασίας που αναπτύχθηκε για τη μέθοδο διχοτόμησης

```

REPEAT
    SET  $x_3 = (x_1 + x_2)/2$ 
    IF  $f(x_3) \cdot f(x_1) < 0$ 
        SET  $x_2 = x_3$ 
    ELSE
        SET  $x_1 = x_3$ 
    ENDIF
UNTIL  $(|x_1 - x_2| < E) \text{ OR } f(x_3) = 0$ 

```

Πίνακας 1.1. Σχέδιο προγράμματος για τη εύρεση ρίζας μη-γραμμικής εξίσωσης με τη μέθοδο της διχοτόμησης.

ΚΡΙΤΙΚΗ

Δύο στοιχεία κάνουν την μέθοδο ελάχιστα ελκυστική:

- Η αργή σύγκλιση
- Επικίνδυνη, όταν υπάρχουν ασυνέχειες

ΣΦΑΛΜΑ

Ως σφάλμα ορίζουμε την “ απόσταση ” $\varepsilon_n = |\rho - x_n|$ της τιμής x_n από τη ρίζα ρ της εξίσωσης. Για τη μέθοδο διχοτόμησης το σφάλμα είναι μικρότερο από το μισό του διαστήματος στο οποίο περικλείεται η ρίζα

$$\varepsilon_n \leq \frac{1}{2} |a_n - b_n| \quad (1.8)$$

Σε κάθε βήμα το σφάλμα μειώνεται στο μισό του προηγούμενου

$$\varepsilon_{n+1} = \frac{\varepsilon_n}{2} = \frac{\varepsilon_{n-1}}{2^2} = \dots = \frac{\varepsilon_0}{2^{n+1}} \quad (1.9)$$

υπάρχει επομένως η δυνατότητα να υπολογίσουμε από την αρχή τον αριθμό των βημάτων που απαιτούνται για την επίτευξη μιας δεδομένης ακρίβειας στην εύρεση της ρίζας. Εστω επομένως ότι E είναι η ζητούμενη ακρίβεια τότε η προ-γούμενη σχέση μας οδηγεί στο ζητούμενο:

$$n = \log_2 \frac{\varepsilon_0}{E} \quad (1.10)$$

Δηλαδή αν δοθεί το εύρος του διαστήματος, ε_0 , εντός του οποίου εντοπίζεται η ρίζα και τη ζητούμενη ακρίβεια E τότε από τη σχέση (1.10) υπολογίζεται άμεσα ο απαιτούμενος αριθμός επαναλήψεων, n , της διαδικασίας.

1.2 Η ΜΕΘΟΔΟΣ ΤΗΣ ΓΡΑΜΜΙΚΗΣ ΠΑΡΕΜΒΟΛΗΣ

Όπως διαπιστώσαμε, η μέθοδος της διχοτόμησης συγκλίνει στην ακριβή τιμή της ρίζας αρκετά αργά. Για το λόγο αυτό, έχουν αναπτυχθεί μέθοδοι με τις

οποίες επιτυγχάνεται ταχύτερη σύγκλιση. Αρκετά δημοφιλής είναι η μέθοδος της **γραμμικής παρεμβολής**.

Στηρίζεται στην εξής λογική: Αν στο διάστημα $[x_1, x_2]$ υπάρχει μία ρίζα της $f(x)$, δηλαδή $f(x_1) \cdot f(x_2) < 0$, τότε φέρνω την ευθεία που διέρχεται από τα σημεία $(x_1, f(x_1))$ και $(x_2, f(x_2))$ με εξίσωση:

$$y(x) = f(x_1) + \frac{f(x_1) - f(x_2)}{x_1 - x_2} (x - x_1) \quad (1.11)$$

η οποία τέμνει τον άξονα Ox έστω στο σημείο x_3 που υπολογίζεται απο την παρακάτω σχέση:

$$\begin{aligned} x_3 &= \frac{x_2 f(x_1) - x_1 f(x_2)}{f(x_1) - f(x_2)} \\ &= x_2 - \frac{f(x_2)}{f(x_2) - f(x_1)} (x_2 - x_1) \end{aligned} \quad (1.12)$$

Επομένως στην πράξη αντικαθιστώ στο διάστημα $[x_1, x_2]$ την συνάρτηση $f(x)$ με μία ευθεία και θεωρώ προσεγγιστικά ότι η τομή αυτής της ευθείας με τον άξονα Ox είναι η ζητούμενη ρίζα. Το νέο σημείο είναι πλησιέστερα στη ρίζα και ως εκ τούτου μπορώ να χρησιμοποιήσω τη σχέση (1.12) ξανά για την αναζήτηση μιας νέας τιμής που να βρίσκεται πλησιέστερα στη ρίζα ρ . Για να χρησιμοποιήσουμε όμως τη σχέση (1.12) απαιτείται η κατάλληλη επιλογή ενός εκ των δύο αρχικών σημείων. Μια απλή (αλλά όχι μοναδική) διαδικασία επιλογής είναι να φροντίζουμε το ζεύγος των σημείων που χρησιμοποιούμε να περικλείει τη ρίζα, δηλαδή:

- (I) είτε $f(x_1) \cdot f(x_3) < 0 \Rightarrow x_2 = x_3$
- (II) είτε $f(x_2) \cdot f(x_3) < 0 \Rightarrow x_1 = x_3$
- (III) είτε $f(x_3) = 0$

ΑΝΑΔΡΟΜΙΚΗ ΣΧΕΣΗ

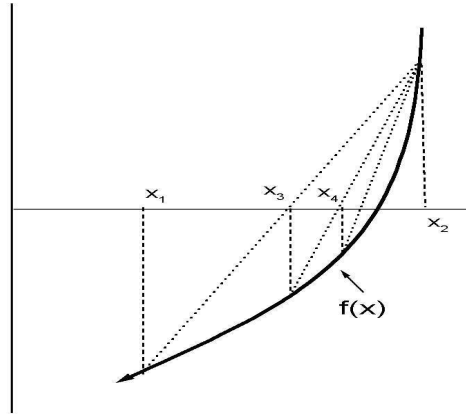
Η σχέση (1.12) με κατάλληλους δείκτες είναι η ζητούμενη αναδρομική σχέση:

$$x_{n+2} = x_{n+1} - \frac{f(x_{n+1})}{f(x_{n+1}) - f(x_n)} (x_{n+1} - x_n). \quad (1.13)$$

Δηλαδή αν δοθούν δύο αρχικές τιμες x_n και x_{n+1} με την παραπάνω σχέση μπορούμε να υπολογίσουμε μια νές τιμή x_{n+2} που βρίσκεται πλησιέστερα στη ρίζα ρ της μη-γραμμικής εξίσωσης.

ΠΑΡΑΔΕΙΓΜΑ

Ας εφαρμόσουμε στο παράδειγμα (ηλογιστικό πρόβλημα) της προηγούμενης μεθόδου τη γραμμική παρεμβολή, θα βρούμε για $x_1 = 0.1$ και $x_2 = 0.15$ ότι



Σχήμα 1.2. Γραφική απεικόνιση της διαδικασίας που αναπτύχθηκε για τη μέθοδο διχοτόμησης

$x_3 = 0.1229$ και $f(x_3) = 122.1$ δηλαδή, ήδη από το πρώτο βήμα έχουμε επιτύχει εξαιρετικά ακριβές αποτέλεσμα και, αν επαναλάβουμε ακόμη μια φορά τη διαδικασία, καταλήγουμε πρακτικά στο ζητούμενο αποτέλεσμα, δηλαδή: $x_4 = 0.12375$ και $f(x_4) = 4.4$

| Αριθ. επαναλήψεων | x_1 | x_2 | x_3 | $f(x_3)$ |
|-------------------|---------|-------|-----------|----------|
| 1 | 0.10 | 0.15 | 0.1229 | 122.114 |
| 2 | 0.1229 | 0.15 | 0.12375 | 4.361 |
| 3 | 0.12375 | 0.15 | 0.1237787 | 0.156 |
| 4 | 0.12378 | 0.15 | 0.1237798 | 0.00571 |

Πίνακας 1.2. Τα αποτελέσματα από τη χρήση της γραμμικής παρεμβολής για το πρόβλημα του επιτοκίου

Στο πρόβλημα που περιγράψαμε, η σύγκλιση είναι ταχύτατη, γιατί η μορφή της καμπύλης στην περιοχή της ρίζας είναι γραμμική, και ως εκ τούτου, η προσέγγισή της με μια ευθεία είναι πολύ καλή.

ΚΡΙΤΙΚΗ

Η μέθοδος της γραμμικής παρεμβολής :

- Συγκλίνει ταχύτερα από τη μέθοδο διχοτόμησης

- Δεν είναι υποχρεωτικό η ρίζα να εσωκλείεται μεταξύ των δύο αρχικών τιμών.

Η διαδικασία εύρεσης της ρίζας με την μεθοδολογία που αναπτύξαμε περιγράφεται από τον ακόλουθο αλγόριθμο:

| ΑΛΓΟΡΙΘΜΟΣ (I) ΓΙΑ ΤΗ ΓΡΑΜΜΙΚΗ ΠΑΡΕΜΒΟΛΗ | |
|---|--|
| Υποθέτει ότι η ρίζα περικλείεται στο αρχικό διάστημα (x_1, x_2) | |
| REPEAT | |
| SET | $x_3 = x_2 - f(x_2) \cdot \frac{x_2 - x_1}{f(x_2) - f(x_1)}$ |
| IF | $f(x_3) \cdot f(x_1) < 0$ |
| SET | $x_2 = x_3$ |
| ELSE | |
| SET | $x_1 = x_3$ |
| ENDIF | |
| UNTIL | $ f(x_3) < E$ |

Πίνακας 1.3. Πρώτη παραλλαγή της μεθόδου γραμμικής παρεμβολής.

Στη διαδικασία που μόλις περιγράψαμε (Πίνακας 1.3) υπάρχει κάποιο αδύνατο σημείο, γιατί, όπως γίνεται αντιληπτό, η τιμή x_2 παραμένει σταθερή. Άρα πλησιάζουμε τη ρίζα " μονόπλευρα " και αυτό έχει ως συνέπεια, μερικές φορές, η σύγκλιση να είναι αρκετά αργή, βλ. Σχήμα 1.2

Υπάρχουν δυο εναλλακτικές διαδικασίες για τη διόρθωση της μονόπλευρης σύγκλισης. Η πρώτη διαδικασία αλλάζει την επιλογή των τιμών για το επόμενο βήμα, δηλαδή, αν από δυο τιμές x_1 και x_2 έχουμε υπολογίσει με τη χρήση της εξίσωσης (1.13) μια τιμή x_3 , χρησιμοποιούμε για το επόμενο βήμα την x_3 και αυτήν από τις x_1, x_2 , για την οποία η $|f(x)|$ είναι μικρότερη. Επομένως, ο τροποποιημένος αλγόριθμος θα είναι:

| ΑΛΓΟΡΙΘΜΟΣ (II) ΓΙΑ ΤΗΝ ΓΡΑΜΜΙΚΗ ΠΑΡΕΜΒΟΛΗ | |
|--|--|
| Η ρίζα δεν περικλείεται στο αρχικό διάστημα (x_1, x_2) | |
| REPEAT | |
| SET | $x_3 = x_2 - f(x_2) \cdot \frac{x_2 - x_1}{f(x_2) - f(x_1)}$ |
| IF | $ f(x_1) < f(x_2) $ |
| SET | $x_2 = x_3$ |
| ELSE SET | $x_1 = x_3$ |
| ENDIF | |
| UNTIL | $ f(x_3) < E$ |

Πίνακας 1.4. Δεύτερη παραλλαγή της μεθόδου γραμμικής παρεμβολής.

Η δεύτερη διαδικασία που βελτιώνει την ακρίβεια και την ταχύτητα σύγκλισης είναι η αντικατάσταση της τιμής της $f(x)$ στο σταθερό σημείο x_2 με την τιμή

$f(x_2)/2$ κοκ. Η διαδικασία αυτή περιγράφεται στον παρακάτω τροποποιημένο αλγόριθμο:

| ΑΛΓΟΡΙΘΜΟΣ (II) (τροποποιημένος) |
|--|
| $F_1 = f(x_1)$ $F_2 = f(x_2)$ $S = f(x_1)$ REPEAT SET $x_3 = x_2 - F_2 \cdot \frac{x_2 - x_1}{F_2 - F_1}$ IF $f(x_3) \cdot F_1 < 0$ SET $x_2 = x_3$ SET $F_2 = f(x_3)$ IF $f(x_3) \cdot S > 0$ SET $F_1 = F_1/2$ ENDIF ELSE SET $x_1 = x_3$ SET $F_1 = f(x_3)$ IF $f(x_3) \cdot S > 0$ SET $F_2 = F_2/2$ ENDIF ENDIF SET save = $f(x_3)$ UNTIL $ f(x_3) < E$ |

Πίνακας 1.5. Εναλλακτικός αλγόριθμος για τη γραμμική παρεμβολή

ΣΥΓΚΛΙΣΗ

Έστω ξ η ακριβής ρίζα της εξίσωσης. Αν θεωρήσουμε ότι $\varepsilon_n = |\xi - x_n|$ είναι το σφάλμα στην εύρεση της ρίζας της $f(x) = 0$ για $x = x_n$, τότε η μέθοδος της γραμμικής παρεμβολής συγκλίνει με βάση τη σχέση

$$\varepsilon_{n+1} = k \cdot \varepsilon_n^{1.618} \quad (1.14)$$

Η απόδειξη αυτής της σχέσης, μπορεί να χρησιμοποιηθεί και ως βάση για τον υπολογισμό της ταχύτητας σύγκλισης και άλλων μεθόδων αυτού του κεφαλαίου.

Έστω η αναδρομική σχέση

$$x_{n+2} = x_{n+1} - \frac{f(x_{n+1})}{f(x_{n+1}) - f(x_n)} (x_{n+1} - x_n) \quad (1.15)$$

τότε

$$x_n = \xi + \varepsilon_n \quad (1.16)$$

όπου $f(\xi) = 0$. Δηλαδή, ε_n είναι η "απόσταση" (το σφάλμα) της x_n από την ακριβή ρίζα της εξίσωσης ξ . Επομένως αν θεωρήσουμε ένα αναπτύγμα Taylor για κάθε μία από τις τρεις τιμές x_1 , x_2 και x_3

$$f(x_i) = f(\xi + \varepsilon_i) = f(\xi) + \varepsilon_i f'(\xi) + \frac{\varepsilon_i^2}{2} f''(\xi) \quad (1.17)$$

και αντικαταστήσουμε τα αναπτύγματα στην αρχική σχέση βρίσκουμε

$$\xi + \varepsilon_{n+2} = \xi + \varepsilon_{n+1} - \frac{\varepsilon_{n+1} f'(\xi) + \varepsilon_{n+1}^2 f''(\xi)/2}{f'(\xi)(\varepsilon_{n+1} - \varepsilon_n) + \frac{1}{2} f''(\xi)(\varepsilon_{n+1}^2 - \varepsilon_n^2)} \cdot (\varepsilon_{n+1} - \varepsilon_n)$$

η σχέση αυτή μετά από μερικές απλοποιήσεις οδηγεί στην:

$$\varepsilon_{n+2} = \varepsilon_{n+1} \left[1 - \left(1 - \frac{\varepsilon_n f''(\xi)}{2 f'(\xi)} \right) \right] = \varepsilon_{n+1} \varepsilon_n \frac{f''(\xi)}{2 f'(\xi)}$$

η οποία συνδέει το σφάλμα στο βήμα $n+2$ με τα σφάλματα στα βήματα $n+1$ και n .

Το ζητούμενο όμως είναι μια σχέση της μορφής $\varepsilon_{n+1} = k \cdot \varepsilon_n^m$ όπου τα k και m είναι άγνωστες σταθερές. Για τον υπολογισμό μιας σχέσης αυτής της μορφής θα εργασθούμε ως εξής, τα σφάλματα στις επαναλήψεις $n+1$ και $n+2$ είναι:

$$\left. \begin{aligned} \varepsilon_{n+1} &= k \cdot \varepsilon_n^m \\ \varepsilon_{n+2} &= k \cdot \varepsilon_{n+1}^m \end{aligned} \right\} \Rightarrow k \cdot \varepsilon_{n+1}^m = \varepsilon_{n+1} \varepsilon_n \left(\frac{f''(\xi)}{2 f'(\xi)} \right)$$

και τελικά λαμβάνουμε

$$\varepsilon_{n+1} = \left(\frac{1}{k} \right)^{\frac{1}{m-1}} \varepsilon_n^{\frac{1}{m-1}} A^{\frac{1}{m-1}} \quad \text{όπου} \quad A = \frac{f''(\xi)}{2 f'(\xi)} \quad (1.18)$$

οπότε αντικαθιστώντας και το ε_{n+1} οδηγούμαστε σε μια σχέση μόνο για το ε_n από την οποία θα προσπαθήσουμε να υπολογίσουμε τις τιμές του k και m . Άρα

$$k \cdot \varepsilon_n^m = \left(\frac{A}{k} \right)^{\frac{1}{m-1}} \varepsilon_n^{\frac{1}{m-1}} \quad \left| \quad \begin{aligned} k &= \left(\frac{A}{k} \right)^{1/(m-1)} \\ m &= \frac{1}{m-1} \Rightarrow m^2 - m - 1 = 0 \end{aligned} \right.$$

Άρα

$$m = \frac{1 \pm \sqrt{5}}{2} = 1.618 \quad \text{και} \quad k^m = A = \frac{f''(\xi)}{2 f'(\xi)}$$

Οποτε τελικά καταλήγουμε στη ζητούμενη σχέση:

$$\varepsilon_{n+1} = k \cdot \varepsilon_n^{1.618} \quad (1.19)$$

Παρατηρούμε ότι η σύγκλιση της μεθόδου δεν είναι γραμμική αλλά περίπου τετραγωνική και προφανώς η εύρεση της ρίζας μιας εξίσωσης απαιτεί σημαντικά μικρότερο αριθμό αριθμητικών πράξεων.

1.3 Η ΜΕΘΟΔΟΣ ΤΟΥ MULLER

Η μέθοδος Muller αποτελεί επέκταση της μεθόδου της γραμμικής παρεμβολής και αντί να προσεγγίζει την συνάρτηση με ευθεία την προσεγγίζει με *παραβολή*.

Αν δοθούν τρεις αρχικές τιμές x_{i-2}, x_{i-1}, x_i , υποθέτουμε ότι η $f(x)$ προσεγγίζεται από ένα 2ο-βάθμιο πολυώνυμο $P(x)$ του οποίου εύκολα υπολογίζουμε τις ρίζες. Για τις τρεις αρχικές τιμές του x_i λαμβάνουμε τρεις εξισώσεις της μορφής :

$$f(x_i) \approx P(x_i) = Ax_i^2 + Bx_i + \Gamma \quad (1.20)$$

από τις οποίες υπολογίζουμε τους 3 συντελεστές του τριωνύμου μέσω των σχέσεων

$$\begin{aligned} A &= qP(x_i) - q(1+q)P(x_{i-1}) + q^2P(x_{i-2}) \\ B &= (2q+1)P(x_i) - (1+q)^2P(x_{i-1}) + q^2P(x_{i-2}) \\ \Gamma &= (1+q)P(x_i) \end{aligned} \quad (1.21)$$

όπου

$$q = \frac{x_i - x_{i-1}}{x_{i-1} - x_{i-2}}. \quad (1.22)$$

Οπότε η επόμενη προσεγγιστική τιμή x_{i+1} βρίσκεται ως η ρίζα της παραβολής

$$Ax^2 + Bx + \Gamma = 0. \quad (1.23)$$

ΑΝΑΔΡΟΜΙΚΗ ΣΧΕΣΗ

Η παρακάτω σχέση μπορεί να χρησιμοποιηθεί αναδρομικά για την εύρεση της ρίζας

$$x_{i+1} = x_i - (x_i - x_{i-1}) \left[\frac{2\Gamma}{B \pm \sqrt{B^2 - 4A\Gamma}} \right] \quad (1.24)$$

Το σημείο του παρονομαστή επιλέγεται ούτως ώστε να γίνεται το κλάσμα μικρότερο απολύτως.

ΣΦΑΛΜΑ

Αποδεικνύεται ότι το σφάλμα της μεθόδου είναι:

$$\varepsilon_{n+1} = k\varepsilon_n^{1.84} \quad (1.25)$$

δηλαδή η συγκλιση είναι σημαντικά καλύτερη από τη μέθοδο της γραμμικής παρεμβολής.

ΚΡΙΤΙΚΗ

Τα πλεονεκτήματα της μεθόδου είναι:

- Συγκλίνει ταχύτατα
- Μπορεί να χρησιμοποιηθεί και για την εύρεση μιγαδικών ριζών.

1.4 Η ΜΕΘΟΔΟΣ $x = g(x)$

Στα Μαθηματικά, ονομάζουμε *σταθερό σημείο* μιας συνάρτησης $g(x)$ έναν πραγματικό αριθμό τέτοιον, ώστε $p = g(p)$. Ενώ η ακολουθία $p_{n+1} = g(p_n)$ για $n = 0, 1, \dots$ καλείται *ακολουθία σταθερού σημείου*.

ΘΕΩΡΗΜΑ: Έστω μια συνεχής συνάρτηση $g(x)$ και μια ακολουθία p_n , που δημιουργείται από μια αναδρομική σχέση σταθερού σημείου. Αν $\lim_{n \rightarrow \infty} p_n = p$, τότε το p είναι ένα σταθερό σημείο της $g(x)$.

Αν δοθεί μια συνάρτηση $f(x)$ της οποίας ζητάμε τη ρίζα $f(\rho) = 0$ και είναι δυνατό να βρεθεί μια γραφή $x = g(x)$ τέτοια, ώστε, αν $f(\rho) = 0$ να είναι $\rho = g(\rho)$, τότε η ακολουθία $x_n = g(x_n)$ οδηγεί στη ρίζα της εξίσωσης.

Επομένως, προσπαθούμε να βρούμε μια σχέση της μορφής

$$x_{n+1} = g(x_n) \quad (1.26)$$

ούτως ώστε

$$\lim_{n \rightarrow \infty} x_n = \rho. \quad (1.27)$$

ΘΕΩΡΗΜΑ: Αν σε ένα διάστημα $I = (\rho - \varepsilon, \rho + \varepsilon)$ περιέχεται μια απλή ρίζα ρ της εξίσωσης $f(x) = 0$ και, αν η εξίσωση αυτή μπορεί να γραφεί στη μορφή $x = g(x)$ με $|g'(x)| \leq \rho < 1$, τότε η ακολουθία που προκύπτει με $x_0 \in I$ συγκλίνει και δίνει τη ρίζα ρ στο όριο $n \rightarrow \infty$. Επιπλέον, αυτή είναι μοναδική ρίζα της $f(x)$ στο διάστημα I .

ΑΠΟΔΕΙΞΗ

Αν $x_n \in I$, τότε

$$x_{n+1} - \rho = g(x_n) - g(\rho) = \frac{g(x_n) - g(\rho)}{x_n - \rho} (x_n - \rho) = g'(\xi_n)(x_n - \rho)$$

Επειδή $|g'(x_n)| \leq \rho < 1$,

$$|x_{n+1} - \rho| \leq \rho |x_n - \rho| < |x_n - \rho|$$

μετά από n βήματα

$$|x_n - \rho| \leq \rho^n |x_0 - \rho| \Rightarrow \lim_{n \rightarrow \infty} |x_n - \rho| = 0$$

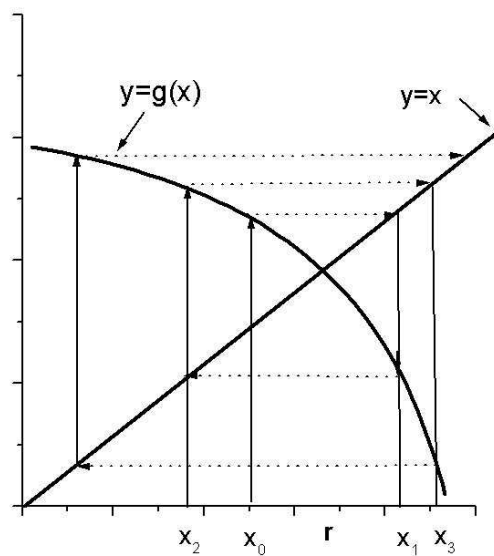
Τέλος, αν ρ_1 μια άλλη ρίζα, τότε

$$\rho - \rho_1 = \frac{g(\rho) - g(\rho_1)}{\rho - \rho_1} (\rho - \rho_1) = g'(\xi)(\rho - \rho_1)$$

Επομένως επειδή $|g'(\xi)| \leq \rho < 1$ οδηγούμαστε στο συμπέρασμα

$$|\rho - \rho_1| \leq \rho |\rho - \rho_1| < |\rho - \rho_1|$$

που είναι **αδύνατον**. Άρα η ρ είναι η μοναδική ρίζα στο διάστημα I .



Σχήμα 1.4. Γραφική απεικόνιση της διαδικασίας σύγκλισης της μεθόδου $x = g(x)$ για $g'(x) > 1$.

| επαναλήψεις | x_n | $(x_n - \rho)/\rho$ |
|-------------|---------|---------------------|
| 1 | 0.1 | - |
| 2 | 0.1043 | -0.157 |
| 3 | 0.1080 | -0.1275 |
| 4 | 0.1111 | -0.1028 |
| 5 | 0.1136 | -0.0825 |
| ... | ... | ... |
| 29 | 0.12376 | -0.000135 |

Πίνακας 1.6.

$$\varepsilon_{n+1} = \rho - x_{n+1} = g(\rho) - g(x_n) = g'(\rho)(\rho - x_n) = g'(\rho)\varepsilon_n$$

Αν η $g'(\rho)$ μεταβάλλεται *αργά*, τότε $y'(\rho) = g'(\rho)$. Δηλαδή το ο ρυθμός ελάττωσης του σφάλματος σε κάθε βήμα είναι γραμμικός και δίνεται από τη σχέση:

$$\varepsilon_{n+1} = g'(\rho)\varepsilon_n \quad (1.28)$$

Επομένως, επιδίωξη μας θα είναι η δημιουργία μιας τέτοιας αναδρομικής σχέσης για την οποία η παράγωγος $g'(x)$ να είναι κατα το δυνατόν μικρότερη.

ΠΑΡΑΔΕΙΓΜΑ

Να εφαρμοσθεί η μέθοδος για την εύρεση μιας ρίζας της εξίσωσης $f(x) = x + \ln(x)$ στο διάστημα $[0.1, 1]$.

Δοκιμάζουμε διάφορες γραφές της εξίσωσης:

- $x_{n+1} = -\ln(x_n)$
αλλιώς $|g'(x)| = \left|\frac{1}{x}\right| \geq 1$ στο διάστημα $[0.1, 1]$ **οπότε δεν συγκλίνει.**
- $x_{n+1} = e^{-x_n}$
οπότε $|y'(x)| = |e^{-x}| \leq e^{-0.1} \approx 0.9 < 1$ **άρα συγκλίνει.**
- Μια άλλη γραφή είναι η εξής: $x = (x + e^{-x})/2$
οπότε $|g'(x)| = \frac{1}{2}|1 - e^{-x}| \leq \frac{1}{2}|1 - e^{-1}| = 0.316$ **άρα συγκλίνει.**
- $x = \frac{x+2e^{-x}}{3}$ οπότε $|g'(x)| = \frac{1}{3}|1 - 2e^{-x}| \leq \frac{1}{3}|1 - 2e^{-1}| = 0.03$ **άρα συγκλίνει.**

Προφανώς θα επιλέξουμε την τελευταία γραφή και η ζητούμενη αναδρομική σχέση θα είναι:

$$x_{n+1} = \frac{x_n + 2e^{-x_n}}{3}.$$

1.4.1 Βελτίωση Aitken

Όταν η σύγκλιση μιας μεθόδου είναι γραμμική, όπως στην προηγούμενη μέθοδο, όπου είχαμε $e_{n+1} = g'(\rho)e_n$, τότε, για $n \rightarrow \infty$, η μέθοδος είναι δυνατόν να επεκταθεί, για να επιτύχουμε ακριβέστερο αποτέλεσμα χωρίς επιπλέον πράξεις.

Το σφάλμα μετά από n εφαρμογές της σχέσης $x_{n+1} = g'(x_n)$ είναι:

$$\rho - x_{n+1} \approx g'(\rho)(\rho - x_n)$$

και μετά από $n+1$ εφαρμογές είναι:

$$\rho - x_{n+2} \approx g'(\rho)(\rho - x_{n+1})$$

οπότε διαιρώντας κατά μέλη βρίσκουμε ότι

$$\frac{\rho - x_{n+1}}{\rho - x_{n+2}} = \frac{g'(\rho)(\rho - x_n)}{g'(\rho)(\rho - x_{n+1})}$$

και λύνοντας ως προς ρ βρίσκουμε:

$$\rho = x_n - \frac{(x_n - x_{n-1})^2}{x_n - 2x_{n-1} + x_{n-2}} \quad (1.29)$$

Η μέθοδος αυτή είναι γνωστή ως *βελτίωση του Aitken*.

ΠΑΡΑΔΕΙΓΜΑ

Αν στα αποτελέσματα του Πίνακα 1.6 της μεθόδου $x = g(x)$ χρησιμοποιήσουμε ως $x_0 = 0.1080$, $x_1 = 0.1111$ και $x_2 = 0.1136$, τότε η βελτίωση Aitken δίνει $r = 0.1240$ με σχετικό σφάλμα $\varepsilon = 0.00024$. Η ακρίβεια αυτή επιτεύχθηκε μετά από **25 εφαρμογές** της μεθόδου $x = g(x)$.

1.5 ΜΕΘΟΔΟΣ NEWTON - RAPHSON

Αν στην περιοχή μιας ρίζας ρ της συνάρτησης $f(x)$ η $f'(x)$ και η $f''(x)$ είναι συνεχείς, τότε μπορούμε να αναπτύξουμε αλγορίθμους, οι οποίοι συγκλίνουν στη ρίζα της εξίσωσης $f(x) = 0$ ταχύτερα από όλες τις μεθόδους που εξετάσαμε ως τώρα. Η μέθοδος Newton-Raphson είναι ευρύτατα διαδεδομένη και η δημοφιλέστερη από όλες τις προηγούμενες μεθόδους, απαιτεί όμως τη γνώση της συναρτησιακής μορφής της 1ης παραγώγου της συνάρτησης $f(x)$.

Έστω ένα σημείο x_0 , το οποίο θεωρούμε ως την πρώτη προσέγγιση στη ρίζα. Η τομή της εφαπτομένης της καμπύλης στο σημείο $(x_0, f(x_0))$ και του άξονα Ox ορίζει ένα νέο σημείο x_1 για το οποίο προφανώς ισχύει $f(x_1) = 0$. Οπότε, από το ορθογώνιο τρίγωνο x_1x_0A (βλ. Σχήμα 1.5) βρίσκουμε:

$$\tan \theta = f'(x_0) = \frac{f(x_0)}{x_1 - x_0} \quad (1.30)$$

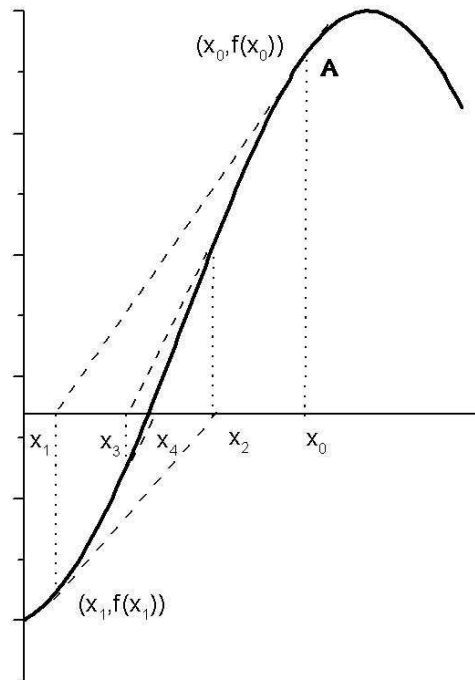
και λύνοντας ως προς x_1 δημιουργούμε την αναδρομική σχέση:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}. \quad (1.31)$$

Αυτή η διαδικασία μπορεί να επαναληφθεί για το x_1 και να βρεθεί ένα νέο σημείο x_2 κοκ.

| επαναλήψεις | x_0 | x_1 | $f(x_1)$ |
|-------------|---------|-----------|----------|
| 1 | 0.15 | 0.1247 | -132.475 |
| 2 | 0.1247 | 0.1237810 | -0.01681 |
| 3 | 0.12378 | 0.1237798 | -0.00101 |

Πίνακας 1.7. Εφαρμογή της μεθόδου Newton-Raphson για το λογιστικό πρόβλημα.



Σχήμα 1.5. Γραφική απεικόνιση της διαδικασίας που αναπτύχθηκε για τη μέθοδο Newton-Raphson.

Ας προσπαθήσουμε τώρα να δημιουργήσουμε τη σχέση (1.31) με πιο αυστηρό τρόπο. Αν υποθέσουμε ότι x_{n+1} είναι η ακριβής λύση της εξίσωσης και μια τιμή x_n βρίσκεται σχετικά κοντά στην x_{n+1} και έστω $x_{n+1} = x_n + \varepsilon_n$. Τότε:

$$f(x_{n+1}) = f(x_n + \varepsilon_n) = f(x_n) + \varepsilon_n f'(x_n) + \frac{\varepsilon_n^2}{2} f''(x_n) + \dots$$

Αλλά, επειδή υποθέσαμε ότι η x_{n+1} είναι ρίζα της $f(x)$, θα ισχύει $f(x_{n+1}) = 0$, και έτσι αναγόμεναι στη σχέση

$$0 = f(x_n) + \varepsilon_n f'(x_n)$$

δηλαδή

$$\varepsilon_n = -\frac{f(x_n)}{f'(x_n)} \quad (1.32)$$

οπότε καταλήγουμε στην αναδρομική σχέση :

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (1.33)$$

που προφανώς προγραμματίζεται πολύ εύκολα.

ΥΠΟΛΟΓΙΣΜΟΣ ΣΦΑΛΜΑΤΟΣ

Έστω ξ η ρίζα της εξίσωσης $f(x) = 0$. Τότε $x_n = \xi + \varepsilon_n$ και $x_{n+1} = \xi + \varepsilon_{n+1}$, οπότε :

$$\xi + \varepsilon_{n+1} = \xi + \varepsilon_n - \frac{f(\xi + \varepsilon_n)}{f'(\xi + \varepsilon_n)} = \xi + \varepsilon_n - \frac{f(\xi) + \varepsilon_n f'(\xi) + \frac{1}{2} \varepsilon_n^2 f''(\xi)}{f'(\xi) + \varepsilon_n f''(\xi)}$$

Επειδή όμως $f(\xi) = 0$ και

$$\frac{1}{1 + \varepsilon f''(\xi)/f'(\xi)} \approx 1 - \varepsilon_n \frac{f''(\xi)}{f'(\xi)}$$

καταλήγουμε στη σχέση

$$\varepsilon_{n+1} = -\frac{f''(\xi)}{2f'(\xi)} \cdot \varepsilon_n^2 \quad (1.34)$$

Όπως παρατηρείτε, η σύγκλιση της μεθόδου είναι “τετραγωνική”, δηλαδή καλύτερη από κάθε άλλη μέθοδο που χρησιμοποιήσαμε ως τώρα.

ΚΡΙΤΙΚΗ

Η μέθοδος Newton-Raphson είναι η απλούστερη από όλες τις άλλες μεθόδους και συνιστάται η χρήση της όταν είναι δυνατός ο υπολογισμός της παραγώγου της συνάρτησης. Γενικά αξίζει να θυμόμαστε ότι :

- Έχει ταχύτατη σύγκλιση
- Απαιτεί γνώση της 1ης παραγώγου της συνάρτησης
- Είναι θεωρητικά ‘ισοδύναμη’ με την γραμμική παρεμβολή

ΠΑΡΑΔΕΙΓΜΑ

Να χρησιμοποιηθεί η μέθοδος Newton-Raphson για την εύρεση της τετραγωνικής ρίζας ενός αριθμού a .

Αν θεωρήσουμε ότι το a είναι λύση της εξίσωσης :

$$f(x) = x^2 - a \quad \text{τότε} \quad f'(x) = 2x$$

οπότε αντικαθιστώντας στην αναδρομική σχέση (1.33) καταλήγω στην αναδρομική σχέση από την οποία με μόνο λίγες επαναλήψεις μπορεί να υπολογισθεί με ακρίβεια η ρίζα του αριθμού

$$x_{n+1} = x_n - \frac{x_n^2 - a}{2x_n}$$

ή σε μια καλύτερη γραφή :

$$x_{n+1} = \frac{1}{2} \left(x_n + \frac{a}{x_n} \right). \quad (1.35)$$

```

COMPUTE  $f(x_1), f'(x_1)$ 
SET  $x_2 = x_1$ 
IF ( $f(x_1) \neq 0$ ) END ( $f'(x_1) \neq 0$ )
  REPEAT
    SET  $x_1 = x_1$ 
    SET  $x_2 = x_1 - f(x_1)/f'(x_1)$ 
  UNTIL ( $|x_1 - x_2| < E$ ) OR ( $|f(x_2)| < E'$ )
ENDIF

```

Πίνακας 1.8. Αλγόριθμος για τον προγραμματισμό της μεθόδου Newton-Raphson.

1.5.1 Δεύτερης τάξης Newton-Raphson (Halley)

Θα δοκιμάσουμε μια βελτίωση της μεθόδου Newton-Raphson προσπαθώντας να επιτύχουμε ταχύτερη σύγκλιση. Για να το επιτύχουμε θα χρησιμοποιήσουμε ξανά το ανάπτυγμα Taylor κοντά στη ρίζα της εξίσωσης.

Έστω ότι $\varepsilon_n = x_{n+1} - x_n$, τότε

$$f(x_{n+1}) = f(x_n + \varepsilon_n) = f(x_n) + \varepsilon_n f'(x_n) + \frac{\varepsilon_n^2}{2} f''(x_n) + \dots$$

οπότε, αν θεωρήσουμε ότι $f(x_{n+1}) \approx 0$, τότε

$$f(x_n) + \varepsilon_n \left[f'(x_n) + \frac{\varepsilon_n}{2} f''(x_n) \right] = 0$$

οπότε λύνοντας ως το πρώτο από τα ε_n που βρίσκεται εκτός της αγκύλης καταλήγουμε στην παρακάτω σχέση

$$\varepsilon_n = - \frac{f(x_n)}{f'(x_n) + \frac{\varepsilon_n}{2} f''(x_n)}$$

και αντικαθιστώντας το ε_n του παρανομαστή από την αντίστοιχη σχέση (1.32) της απλής Newton - Raphson που αναπτύξαμε προηγουμένως, δηλαδή:

$$\varepsilon_n \approx -\frac{f(x_n)}{f'(x_n)}$$

καταλήγουμε στη σχέση:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n) - \frac{f''(x_n) \cdot f(x_n)}{2f'(x_n)}} = x_n - \frac{2f(x_n)f'(x_n)}{2f'^2(x_n) - f''(x_n)f(x_n)} \quad (1.36)$$

Είναι προφανές ότι αν θέσουμε $f''(x_n) = 0$ καταλήγουμε στη σχέση Newton-Raphson, εξίσωση (1.33). Η μέθοδος αυτή αναφέρεται στη βιβλιογραφία ως **μέθοδος του Halley**.

ΣΥΓΚΛΙΣΗ

Η σύγκλιση της μεθόδου Halley είναι εντυπωσιακή και επιτυγχάνει “κυβική” σύγκλιση (να αποδειχθεί ως άσκηση).

$$\varepsilon_{n+1} = -\left[\frac{1}{6} \frac{f'''(\xi)}{f'(\xi)} - \frac{1}{4} \left(\frac{f''(\xi)}{f'(\xi)} \right)^2 \right] \cdot \varepsilon_n^3 \quad (1.37)$$

Στο παράδειγμα που ακολουθεί κατανοούμε εύκολα την ταχύτητα σύγκλισης της μεθόδου Halley.

ΠΑΡΑΔΕΙΓΜΑ

Να αποδείξετε ότι η τετραγωνική ρίζα ενός αριθμού a δίνεται από τη σχέση:

$$x_{n+1} = \frac{x_n^3 + 3x_n a}{3x_n^2 + a} \quad (1.38)$$

Αν θέσω $f(x) = x^2 - a$ τότε $f'(x) = 2x$ και $f''(x) = 2$ οπότε αντικαθιστώντας στην αναδρομική σχέση Halley καταλήγω στη ζητούμενη αναδρομική σχέση.

Ας χρησιμοποιήσουμε τις δύο σχέσεις για να υπολογίσουμε την ρίζα του 9. Αρα θα βρούμε τις ρίζες της εξίσωσης $f(x) = x^2 - 9 = 0$. Ας θεωρήσουμε μία αρχική τιμή $x_0 = 15$ (αρκετά μακριά από την πραγματική ρίζα). Ο παρακάτω πίνακας μας δείχνει τα αποτελέσματα από τις δύο μεθόδους.

1.5.2 Πολλαπλές ρίζες

Είναι δυνατόν η ρίζα μιάς εξίσωσης $f(x) = 0$ να είναι πολλαπλή, δηλαδή $f(x) = (x - \rho)^m q(x)$ όπου m είναι η πολλαπλότητα της ρίζας ρ . Επομένως, $f'(x) = (x - \rho)^{m-1} [mq(x) + (x - \rho)q'(x)]$ άρα οι $f(x)$ και $f'(x)$ μηδενίζονται συγχρόνως για $x = \rho$ οπότε ο λόγος $f(x)/f'(x)$ θα υπολογίζεται με μεγάλο σφάλμα διότι ουσιαστικά θα έχουμε τη διαίρεση δύο μικρών αριθμών.

| Newton-Raphson | Σφάλμα | Halley | Σφάλμα |
|----------------|---------------------------------------|---------------------|--|
| $x_0=15$ | $\varepsilon_0 = 12$ | $x_0=15$ | $\varepsilon_0 = 12$ |
| $x_1=7.8$ | $\varepsilon_1 = 4.8$ | $x_1=5.526$ | $\varepsilon_1 = 2.5$ |
| $x_2=4.477$ | $\varepsilon_2 = 1.477$ | $x_2=3.16024$ | $\varepsilon_2 = 0.16$ |
| $x_3=3.2436$ | $\varepsilon_3 = 0.243$ | $x_3=3.00011$ | $\varepsilon_3 = 1.05 \times 10^{-4}$ |
| $x_4=3.0092$ | $\varepsilon_4 = 9.15 \times 10^{-3}$ | $x_4=3.0000000...$ | $\varepsilon_4 = 3.24 \times 10^{-14}$ |
| $x_5=3.00001$ | $\varepsilon_5 = 1.39 \times 10^{-5}$ | $x_5= 3.0000000...$ | $\varepsilon_5 = 0.0$ |

Πίνακας 1.9. Σύγκριση των μεθόδων Newton-Raphson και Halley για την εύρεση της τετραγωνικής ρίζας του αριθμού 9.

Για να ξεπερασθεί το πρόβλημα δημιουργούμε μια νέα συνάρτηση

$$\phi(x) = \frac{f(x)}{f'(x)} = \frac{(x - \rho)q(x)}{mq(x) + (x - \rho)q'(x)}$$

η οποία έχει τη ρ ως ρίζα με πολλαπλότητα $m = 1$ και η αναδρομική σχέση είναι:

$$\begin{aligned} x_{n+1} &= x_n - \frac{\phi(x_n)}{\phi'(x_n)} = x_n - \frac{f(x_n)/f'(x_n)}{\{[f'(x_n)]^2 - f(x_n)f''(x_n)\}/[f'(x_n)]^2} \\ &= x_n - \frac{f(x_n)f'(x_n)}{[f'(x_n)]^2 - f(x_n)f''(x_n)} \end{aligned} \quad (1.39)$$

Η ταχύτητα σύγκλισης της μεθόδου αυτής είναι τετραγωνική, διότι πρακτικά η μέθοδος Newton-Raphson εφαρμόστηκε για την εύρεση των ριζών της εξίσωσης $\phi(x) = 0$ για την οποία η ρ είναι απλή ρίζα.

ΠΑΡΑΔΕΙΓΜΑ

Η εξίσωση $f(x) = x^4 - 4x^2 + 4 = 0$ έχει προφανή διπλή ρίζα την $\rho = \sqrt{2} = 1.414213...$. Χρησιμοποιώντας την παραπάνω μέθοδο επαληθεύστε το αναλυτικό αποτέλεσμα αλλιά και την ταχύτητα σύγκλιση της μεθόδου.

Η μέθοδος Newton-Raphson δίνει την αναδρομική σχέση

$$x_{n+1} = x_n - \frac{x_n^2 - 2}{4x_n} \quad (A)$$

ενώ η μέθοδος για πολλαπλές ρίζες, εξίσωση (1.39), δίνει

$$x_{n+1} = x_n - \frac{(x_n^2 - 2)x_n}{x_n^2 + 2} \quad (B).$$

Η σύγκριση των δύο μεθόδων φαίνεται στον Πίνακα

| x | Σχέση (A) | Σφάλμα (A) | Σχέση (B) | Σφάλμα (B) |
|-------|-------------|----------------------|-------------|------------------------|
| x_0 | 1.5 | 8.6×10^{-2} | 1.5 | 8.6×10^{-2} |
| x_1 | 1.458333333 | 4.4×10^{-2} | 1.411764706 | -2.4×10^{-3} |
| x_2 | 1.436607143 | 2.2×10^{-2} | 1.414211439 | -2.1×10^{-6} |
| x_3 | 1.425497619 | 1.1×10^{-2} | 1.414213562 | -1.6×10^{-12} |

που δείχνει την ταχύτητα σύγκλισης των δύο σχέσεων.

ΕΦΑΡΜΟΓΗ

Θα συγκρίνουμε τον αριθμό των επαναλήψεων (ταχύτητα σύγκλισης) που απαιτούνται για την επίτευξη μιας δοσμένης ακρίβειας μεταξύ δύο μεθόδων. Ως εφαρμογή θα χρησιμοποιήσουμε τη γραμμική σύγκλιση της μεθόδου διχοτόμησης και της τετραγωνικής σύγκλισης της μεθόδου Newton-Raphson.

Για **γραμμική σύγκλιση** θα έχουμε:

$$\lim_{n \rightarrow \infty} \frac{|\varepsilon_{n+1}|}{|\varepsilon_n|} = a$$

οπότε θα ισχύει ότι:

$$|\varepsilon_n| \approx a|\varepsilon_{n-1}| \approx a^2|\varepsilon_{n-2}| \approx \dots \approx a^n|\varepsilon_0|$$

οπότε επιλύοντας την παραπάνω σχέση ως προς n καταλήγουμε στη γενική σχέση για τον αριθμό των απαιτούμενων επαναλήψεων ώστε να επιτευχθεί η ακρίβεια $|\varepsilon_n|$:

$$n \approx \frac{\log_{10} |\varepsilon_n| - \log_{10} |\varepsilon_0|}{\log_{10} |a|} \quad (1.40)$$

Κατ' αντιστοιχία για **τετραγωνική σύγκλιση** θα έχουμε:

$$\lim_{n \rightarrow \infty} \frac{|\tilde{\varepsilon}_{n+1}|}{|\tilde{\varepsilon}_n|^2} = b$$

και θα ισχύει ότι:

$$|\tilde{\varepsilon}_n| \approx b|\tilde{\varepsilon}_{n-1}|^2 \approx b^3|\tilde{\varepsilon}_{n-2}|^4 \approx b^7|\tilde{\varepsilon}_{n-3}|^8 \approx \dots \approx b^{2^{n+1}-1}|\tilde{\varepsilon}_0|^{2^{n+1}}$$

Επιλύοντας την παραπάνω σχέση ως προς n καταλήγουμε στη γενική σχέση για τον αριθμό των απαιτούμενων επαναλήψεων ώστε να επιτευχθεί η ακρίβεια $|\varepsilon_n|$:

$$2^{n+1} \approx \frac{\log_{10} |\tilde{\varepsilon}_n| + \log_{10} |b|}{\log_{10} |\tilde{\varepsilon}_0| + \log_{10} |b|} \quad (1.41)$$

Αν υποθέσουμε ότι σε κάποιο πρόβλημα ζητούμε μετά από n επαναλήψεις το σφάλμα να μην υπερβαίνει το 10^{-6} δηλαδή $\varepsilon_n \leq 10^{-6}$ με αρχικό σφάλμα $\varepsilon_0 = 0.5$ και $a = b = 0.7$ τότε η σχέση (1.40) δίνει

$$n \approx \frac{\log_{10} |10^{-6}| - \log_{10} |0.5|}{\log_{10} |0.7|} \approx 37$$

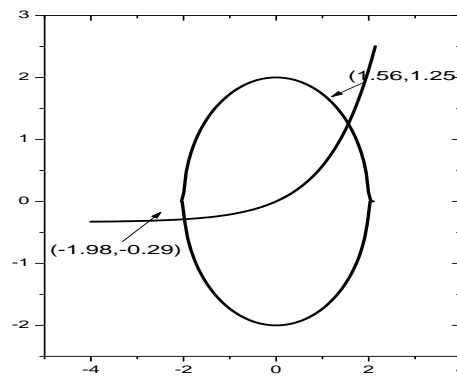
άρα απαιτούνται τουλάχιστον 37 επαναλήψεις ώστε να προσεγγισθεί η ζητούμενη ακρίβεια.

Οι ίδιες απαιτήσεις ακρίβειας για τετραγωνική σύγκλιση απαιτούν

$$2^{n+1} \approx \frac{\log_{10} |10^{-6}| + \log_{10} |0.7|}{\log_{10} |0.5| + \log_{10} |0.7|} \approx 13.5$$

άρα το η ζητούμενη ακρίβεια θα επιτευχθεί μετά από 3 επαναλήψεις.

Η διαφορά στην ταχύτητα σύγκλισης γίνεται ακόμη πιο μεγάλη αν για ζητήσουμε ακόμη μεγαλύτερη τελική ακρίβεια, έστω για παράδειγμα $\varepsilon_n \leq 10^{-14}$ τότε η γραμμική σύγκλιση θα επιτύχει αυτή την ακρίβεια μετά από 89 επαναλήψεις ενώ η τετραγωνική σύγκλιση θα την επιτύχει μετά από 4 επαναλήψεις!



Σχήμα 1.6. Γραφική απεικόνιση του μη-γραμμικού συστήματος (1.42).

1.6 ΣΥΣΤΗΜΑΤΑ ΜΗ-ΓΡΑΜΜΙΚΩΝ ΕΞΙΣΩΣΕΩΝ

Οι τεχνικές που αναπτύξαμε στις προηγούμενες ενότητες αυτού του Κεφαλαίου για τη λύση μη-γραμμικών εξισώσεων μπορούν να εφαρμοστούν και στη λύση συστημάτων μη-γραμμικών εξισώσεων.

Στη συνέχεια θα αναπτύξουμε δύο μεθόδους για την αριθμητική επίλυση συστημάτων που βασίζονται η πρώτη στην τεχνική της μεθόδου Newton-Raphson και η δεύτερη της μεθόδου $x = g(x)$.

Ένα παράδειγμα συστήματος μη-γραμμικών εξισώσεων είναι το εξής. Έστω δύο συναρτήσεις :

$$f(x, y) = e^x - 3y - 1 \tag{1.42}$$

$$g(x, y) = x^2 + y^2 - 4$$

Ζητούμε τα πιθανά σημεία για τα οποία ταυτόχρονα ικανοποιούνται οι σχέσεις

$$\begin{aligned} f(x, y) &= 0 \\ g(x, y) &= 0 \end{aligned}$$

Προφανώς οι εξισώσεις $f(x, y) = 0$ και $g(x, y) = 0$ ορίζουν καμπύλες στο επίπεδο xy και η τομή τους είναι η ζητούμενη λύση του παραπάνω συστήματος (βλ. Σχήμα 1.6).

1.6.1 Η Μέθοδος Newton

Η μέθοδος του Newton αποτελεί στην ουσία επέκταση της μεθόδου Newton-Raphson που μελετήσαμε στο προηγούμενο κεφάλαιο. Ως πρώτο βήμα θα μελετήσουμε την διαδικασία εφαρμογής της σε σύστημα δύο μη-γραμμικών εξισώσεων και στη συνέχεια θα την επεκτείνουμε σε συστήματα N εξισώσεων με N αγνώστους.

ΣΥΣΤΗΜΑΤΑ ΔΥΟ ΕΞΙΣΩΣΕΩΝ

Έστω το σύστημα

$$\begin{aligned} f(x, y) &= 0 \\ g(x, y) &= 0 \end{aligned}$$

Εάν εφαρμόσουμε μια επαναληπτική διαδικασία που συγκλίνει σε μια από τις πιθανές λύσεις του συστήματος, τότε μετά από $n + 1$ επαναλήψεις το σημείο (x_{n+1}, y_{n+1}) μπορεί να θεωρηθεί ότι προσεγγίζει ικανοποιητικά σε μία λύση του συστήματος, δηλαδή $f(x_{n+1}, y_{n+1}) \cong 0$ και $g(x_{n+1}, y_{n+1}) \cong 0$. Εάν θεωρήσουμε ότι στο προηγούμενο βήμα (x_n, y_n) το σφάλμα προσέγγισης της λύσης είναι τέτοιο ώστε $x_{n+1} = x_n + \varepsilon_n$ και $y_{n+1} = y_n + \delta_n$, τότε αναπτύσσοντας σε σειρά Taylor γύρω από τη λύση, βρίσκουμε:

$$\begin{aligned} 0 &\cong f(x_{n+1}, y_{n+1}) = f(x_n + \varepsilon_n, y_n + \delta_n) \cong f(x_n, y_n) + \varepsilon_n \frac{\partial f}{\partial x} + \delta_n \frac{\partial f}{\partial y} \\ 0 &\cong g(x_{n+1}, y_{n+1}) = g(x_n + \varepsilon_n, y_n + \delta_n) \cong g(x_n, y_n) + \varepsilon_n \frac{\partial g}{\partial x} + \delta_n \frac{\partial g}{\partial y} \end{aligned}$$

και λύνοντας το παραπάνω σύστημα ως προς τα ε_n και δ_n καταλήγουμε στις σχέσεις

$$\varepsilon_n = \frac{-f \frac{\partial g}{\partial y} + g \frac{\partial f}{\partial y}}{\frac{\partial f}{\partial x} \frac{\partial g}{\partial y} - \frac{\partial g}{\partial x} \frac{\partial f}{\partial y}} \quad (1.43)$$

$$\delta_n = \frac{-g \frac{\partial f}{\partial x} + f \frac{\partial g}{\partial x}}{\frac{\partial f}{\partial x} \frac{\partial g}{\partial y} - \frac{\partial g}{\partial x} \frac{\partial f}{\partial y}} \quad (1.44)$$

Από τον ορισμό των ε_n και δ_n , καταλήγουμε στις παρακάτω σχέσεις που αποτελούν γενίκευση της μεθόδου Newton-Raphson για συστήματα 2 μη-γραμμικών εξισώσεων:

$$x_{n+1} = x_n - \frac{f \cdot g_y - g \cdot f_y}{f_x \cdot g_y - g_x \cdot f_y} \quad (1.45)$$

$$y_{n+1} = y_n - \frac{g \cdot f_x - f \cdot g_x}{f_x \cdot g_y - g_x \cdot f_y} \quad (1.46)$$

όπου έχουμε χρησιμοποιήσει το συμβολισμό $f_x = \partial f / \partial x$.

ΣΥΣΤΗΜΑΤΑ N ΕΞΙΣΩΣΕΩΝ

Η παραπάνω μέθοδος μπορεί εύκολα να γενικευθεί και για συστήματα N εξισώσεων:

$$\begin{aligned} f_1(x^1, x^2, \dots, x^N) &= 0 \\ f_2(x^1, x^2, \dots, x^N) &= 0 \\ &\vdots \\ f_n(x^1, x^2, \dots, x^N) &= 0 \end{aligned}$$

με N αγνώστους (x^1, x^2, \dots, x^N) . Αν θεωρήσουμε ότι οι τιμές $(x_{n+1}^1, x_{n+1}^2, \dots, x_{n+1}^N)$ προσεγγίζουν ικανοποιητικά μια πιθανή λύση του συστήματος, τότε με βάση τα προηγούμενα θα υπάρχει μια N -αδα τιμών $(x_n^1, x_n^2, \dots, x_n^N)$ για τις οποίες θα ισχύουν οι σχέσεις

$$\begin{aligned} x_{n+1}^1 &= x_n^1 + \Delta x_n^1 \\ &\vdots \\ x_{n+1}^N &= x_n^N + \Delta x_n^N \end{aligned}$$

Αρα μπορούμε να θεωρήσουμε αναπτύγματα της μορφής :

$$\begin{aligned} 0 &\cong f_1(x_{n+1}^1, \dots, x_{n+1}^N) = f_1(x_n^1 + \Delta x_n^1, \dots, x_n^N + \Delta x_n^N) \\ &\approx f_1 + \frac{\partial f_1}{\partial x^1} \Delta x_n^1 + \dots + \frac{\partial f_1}{\partial x^N} \Delta x_n^N \\ &\vdots \\ 0 &\cong f_N(x_{n+1}^1, \dots, x_{n+1}^N) = f_N(x_n^1 + \Delta x_n^1, \dots, x_n^N + \Delta x_n^N) \\ &\approx f_N + \frac{\partial f_N}{\partial x^1} \Delta x_n^1 + \dots + \frac{\partial f_N}{\partial x^N} \Delta x_n^N \end{aligned} \quad (1.47)$$

Οπότε, τα Δx_n^i (όπου $i = 1 \dots N$) θα υπολογισθούν από τη λύση του γραμμικού συστήματος

$$\begin{pmatrix} \frac{\partial f_1}{\partial x^1} & \frac{\partial f_1}{\partial x^2} & \dots & \frac{\partial f_1}{\partial x^N} \\ \vdots & & & \vdots \\ \frac{\partial f_N}{\partial x^1} & \frac{\partial f_N}{\partial x^2} & \dots & \frac{\partial f_N}{\partial x^N} \end{pmatrix} \cdot \begin{pmatrix} \Delta x_n^1 \\ \vdots \\ \Delta x_n^N \end{pmatrix} = - \begin{pmatrix} f_1 \\ \vdots \\ f_N \end{pmatrix} \quad (1.48)$$

με αγνώστους τις N τιμές Δx_n^i .

Άρα, εάν ξεκινήσουμε με μια N -άδα αρχικών τιμών $(x_0^1, x_0^2, \dots, x_0^N)$, τότε από τη λύση του παραπάνω συστήματος θα υπολογίσουμε τις ποσότητες Δx_n^i που οδηγούν σε μια νέα N -άδα τιμών $(x_1^1, x_1^2, \dots, x_1^N)$ μέσω των σχέσεων :

$$\begin{aligned} x_1^1 &= x_0^1 + \Delta x_0^1 \\ &\vdots \\ x_1^N &= x_0^N + \Delta x_0^N \end{aligned} \quad (1.49)$$

Εφόσον συγκλίνει, η παραπάνω διαδικασία μπορεί να επαναληφθεί όσες φορές απαιτείται ώστε να επιτευχθεί η ζητούμενη ακρίβεια, για παράδειγμα έως ότου $\max |\Delta x_n^i| < E$ όπου E μια δοθείσα επιθυμητή ακρίβεια.

1.6.2 Μέθοδος τύπου $x = g(x)$

Αποτελεί επέκταση της μεθόδου που μελετήσαμε στην Κεφάλαιο 1.4, αλλά μεταφέρει και όλα τα πρόβλήματα που αναλύσαμε στο συγκεκριμένο κεφάλαιο.

Έστω το σύστημα των N εξισώσεων

$$\begin{aligned} f_1(x_1, x_2, \dots, x_N) &= 0 \\ &\vdots \\ f_N(x_1, x_2, \dots, x_N) &= 0 \end{aligned} \quad (1.50)$$

Εάν είναι δυνατόν το σύστημα να γραφεί στη μορφή :

$$\begin{aligned} x_1 &= F_1(x_1, x_2, \dots, x_N) \\ &\dots \\ x_N &= F_N(x_1, x_2, \dots, x_N) \end{aligned} \quad (1.51)$$

τότε μπορούμε να εφαρμόσουμε την μεθοδολογία που αναπτύχθηκε για τη μέθοδο $x = g(x)$. Υπάρχουν δύο διαδικασίες για την εύρεση λύσεων του συστήματος:

Α' ΜΕΘΟΔΟΣ

Στο αρχικό βήμα, δίνω N αρχικές τιμές $(x_1, \dots, x_N)_0$ μόνο στο δεξιό μέλος όλων των εξισώσεων (1.51) συγχρόνως και υπολογίζω μια νέα N -άδα τιμών $(x_1, \dots, x_N)_1$ από το αριστερό μέλος των εξισώσεων. Επαναλαμβάνω τη διαδικασία για τη νέα N -άδα τιμών και υπολογίζω μια καλύτερη προσέγγιση της λύσης κ.ο.κ.

Β' ΜΕΘΟΔΟΣ

Δίνω N αρχικές τιμές $(x_1, \dots, x_N)_0$ στο δεξιό μέλος της πρώτης από τις εξισώσεις (1.51) και υπολογίζω (από το αριστερό της μέλος) τη νέα τιμή $(x_1)_1$. Στη συνέχεια, στο δεξιό μέλος της δεύτερης εξίσωσης δίνω τις τιμές $(x_1)_1, (x_2, \dots, x_N)_0$ και υπολογίζω τη νέα τιμή $(x_2)_1$. Στη τρίτη εξίσωση δίνω $(x_1, x_2)_1, (x_3, \dots, x_N)_0$ κ.ο.κ.

ΚΡΙΤΗΡΙΟ ΣΥΓΚΛΙΣΗΣ

Το σύστημα $x_i = f_i(x_1, x_2, \dots, x_N)$ με $i = 1, \dots, N$ θα συγκλίνει σε μια περιοχή γύρω από μια λύση, αν ικανοποιούνται τα κριτήρια σύγκλισης

$$\begin{aligned} \left| \frac{\partial f_1}{\partial x_1} \right| + \left| \frac{\partial f_1}{\partial x_2} \right| + \dots + \left| \frac{\partial f_1}{\partial x_N} \right| &< 1 \\ &\vdots \\ \left| \frac{\partial f_N}{\partial x_1} \right| + \left| \frac{\partial f_N}{\partial x_2} \right| + \dots + \left| \frac{\partial f_N}{\partial x_N} \right| &< 1 \end{aligned} \quad (1.52)$$

Τα κριτήρια αυτά αποτελούν επέκταση του κριτηρίου που ισχύει για τη σύγκλιση στη ρίζα μίας μη-γραμμικής εξίσωσης, χρησιμοποιώντας τη μέθοδο $x = g(x)$. Είναι προφανές, πόσο σύνθετη γίνεται πλέον η μελέτη της σύγκλισης με αυτή τη μέθοδο, στην περίπτωση ενός συστήματος εξισώσεων.

ΠΑΡΑΔΕΙΓΜΑ

Εστω το σύστημα

$$\begin{aligned} x^2 + y^2 &= 4 \\ e^x - 3y &= 1 \end{aligned}$$

του οποίου οι ακριβείς λύσεις είναι $(1.5595, 1.2522)$ και $(-1.9793, -0.2873)$ (σχήμα 1.6). Το σύστημα μπορεί να γραφεί στη μορφή

$$\begin{aligned} x_{n+1} &= -\sqrt{4 - y_n^2} \\ y_{n+1} &= \frac{1}{3}(e^{x_n} - 1) \end{aligned}$$

(όπου επιλέξαμε το πρόσημο (-) στην πρώτη εξίσωση). Αν χρησιμοποιήσουμε την πρώτη μέθοδο, με αρχικές τιμές $(-1, 0)$, δημιουργούμε την ακολουθία τιμών που φαίνεται στον Πίνακα 1.10 η οποία μετά από 5 επαναλήψεις έχει προσεγγίσει ικανοποιητικά τη μία από τις ρίζες της εξίσωσης.

Αν χρησιμοποιηθεί η δεύτερη μέθοδο, τότε δημιουργούμε την ακολουθία τιμών που φαίνεται στον Πίνακα 1.11 Δηλαδή, για το συγκεκριμένο σύστημα, ο αριθμός

| | | | | | | |
|-----|----|---------|---------|---------|---------|---------|
| n | 0 | 1 | 2 | 3 | 4 | 5 |
| x | -1 | -2 | -1.9884 | -1.9791 | -1.9792 | -1.9793 |
| y | 0 | -0.2107 | -0.2882 | -0.2877 | -0.2873 | -0.2873 |

Πίνακας 1.10.

| | | | | |
|-----|----|---------|---------|---------|
| n | 0 | 1 | 2 | 3 |
| x | -1 | -2 | -1.9791 | -1.9793 |
| y | 0 | -0.2882 | -0.2873 | -0.2873 |

Πίνακας 1.11.

των επαναλήψεων που απαιτείται για την επίτευξη της παραπάνω ακρίβειας είναι περίπου ο μισός απ' ό,τι με την πρώτη μέθοδο.

Ας προσπαθήσουμε τώρα να και βρούμε τη δεύτερη λύση του συστήματος, που θα προκύψει από τη λύση του συστήματος

$$\begin{aligned}x_{n+1} &= \sqrt{4 - y_n^2} \\ y_{n+1} &= \frac{1}{3}(e^{x_n} - 1)\end{aligned}$$

(επιλέγοντας, αυτή τη φορά, το πρόσημο (+) στην πρώτη από τις εξισώσεις). Ας ξεκινήσουμε από μια τιμή που είναι αρκετά κοντά στην ακριβή λύση, έστω λοιπόν $x_0 = 1.5$ και $y_0 = 1$. Η ακολουθία τιμών δίνεται στον Πίνακα 1.12

| | | | | | | |
|-----|-----|--------|--------|--------|--------|--------|
| n | 0 | 1 | 2 | 3 | 4 | 5 |
| x | 1.5 | 1.7321 | 1.2630 | 1.8126 | 1.0394 | 1.9050 |
| y | 1 | 1.5507 | 0.8453 | 1.7087 | 0.6092 | 1.9064 |

Πίνακας 1.12.

Παρατηρούμε ότι αποκλίνουμε από τη λύση του συστήματος, και τούτο διότι, αν εφαρμόσουμε τα κριτήρια σύγκλισης στην περιοχή της λύσης, θα παρατηρήσουμε ότι δεν ικανοποιούνται.

Αν όμως γράψουμε το σύστημα στη μορφή

$$\begin{aligned}y_{n+1} &= \sqrt{4 - x_n^2} \\ x_{n+1} &= \ln(1 + 3y_n)\end{aligned}$$

τότε η ακολουθία τιμών που δημιουργούμε συγκλίνει (αλλά αργά) στην δεύτερη λύση του συστήματος μετά από μεγάλο αριθμό επαναλήψεων.

1.7 ΑΣΚΗΣΕΙΣ

1. Να βρεθεί η ρίζα της εξίσωσης $e^x - \sin(x) = 0$ ($\rho = -3.183063012$). Αν το αρχικό διάστημα είναι $[-4, -3]$, πόσες εφαρμογές της μεθόδου διχοτόμησης απαιτούνται, για να επιτύχουμε ακρίβεια τεσσάρων δεκαδικών ψηφίων (δηλαδή) $|x_n - \rho| < 0.00005$.
2. Υπολογίστε το σημείο τομής των καμπυλών $y = e^x$ και $y = 2x + 1$ με ακρίβεια τριών δεκαδικών ψηφίων με τη μέθοδο της διχοτόμησης.
3. Εφαρμόστε τη μέθοδο της γραμμικής παρεμβολής στη λύση του προβλήματος 1. Πόσες επαναλήψεις απαιτούνται για την επίτευξη της ακρίβειας των τεσσάρων δεκαδικών ψηφίων.
4. Επαναλάβετε την άσκηση 2 με τη χρήση γραμμικής παρεμβολής.
5. Εφαρμόστε τη μέθοδο Muller στις ασκήσεις 1 και 2.
6. Βρείτε τις τρεις ρίζες της $e^x - 2x^2 = 0$, με κατάλληλη χρήση της μεθόδου $x = g(x)$.
7. Να βρεθούν οι ρίζες της $3x - \sin(x) - e^x = 0$ με τη μέθοδο της γραμμικής παρεμβολής και τη μέθοδο Newton-Raphson.
8. Βρείτε τις ρίζες της εξίσωσης $x^2 - 2x - 3 = 0$ (προφανώς είναι 3 και -1) με τη μέθοδο $x = g(x)$. Δοκιμάστε τις κατάλληλες γραφές για να επιτύχεται σύγκλιση.
9. Εφαρμόστε τη μέθοδο Newton-Raphson στις ασκήσεις 1 και 2.
10. Να βρεθεί η n -οστή ρίζα ενός αριθμού (αναδρομική σχέση)
11. Να βρεθεί η 3η ρίζα ενός αριθμού
12. Να βρεθεί η ιδιοτιμή λ της εξίσωσης $y'' + \lambda^2 y = 0$ με οριακές συνθήκες $y(0) = 0$ και $y(1) = y'(1)$.
13. Γνωρίζουμε ότι αν η $f(x)$ έχει διπλή ρίζα στο $x = r$ τότε $f'(r) = 0$. Επίσης αν η $f(x)$ έχει μια ρίζα με πολλαπλότητα m στο $x = r$ τότε $f^{(i)}(r) = 0$, για $i = 1, 2, \dots, m-1$. Με βάση τα προηγούμενα να δειχθεί ότι αν $f(x)$, $f'(x)$ και $f''(x)$ είναι συνεχείς και φραγμένες σ'ένα διάστημα που περιέχει μια ρίζα πολλαπλότητας m δηλαδή είναι $f(r) = f'(r) = \dots = f^{(m-1)}(r) = 0$ αλλά $f^{(m)}(r) \neq 0$ τότε η

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

συγκλίνει τετραγωνικά.

14. Να βρεθεί ο αντίστροφος ενός αριθμού χωρίς χρήση διαίρεσης.
15. Χρησιμοποιώντας τη μέθοδο $x = g(x)$ να βρεθεί μια ρίζα του πολωνύμου $2x^3 + 4x^2 - 2x - 5 = 0$ κοντά στο 1. Χρησιμοποιείτε τουλάχιστον δύο διαφορετικές αναδιατάξεις του πολωνύμου και συγκρίνετε το αποτέλεσμα. Τέλος χρησιμοποιήστε τη μέθοδο Aitken για τη βελτίωση του τελικού αποτελέσματος.
16. Να βρεθεί αριθμητικά η τιμή του $1/\sqrt{a}$
17. Η $f(x) = e^x - 3x^2 = 0$ έχει τρεις ρίζες. Αν γραφεί στη μορφή $x = \pm \sqrt{e^x/3}$ να βρεθούν οι ρίζες της, όσες είναι δυνατόν, και αν δεν μπορεί να βρεθεί κάποια με αυτό το σχήμα τότε δοκιμάστε εναλλακτικά σχήματα.

18. Να γίνουν δύο βήματα με τη μέθοδο Newton-Raphson για το παρακάτω σύστημα, με αρχικές τιμές $(0,1)$:

$$4x^2 - y^2 = 0$$

$$4xy^2 - x = 1$$

19. Με αρχικές τιμές $(0,0,1)$ να γίνει ένα βήμα με τη μέθοδο Newton-Raphson για το σύστημα

$$xy - z^2 = 1$$

$$xyz - x^2 + y^2 = 2$$

$$e^x - e^y + z = 3$$

Να εξηγηθεί το αποτέλεσμα.

20. Να λυθούν τα παρακάτω συστήματα με χρήση των δύο αριθμητικών μεθόδων και συγκρίνεται την ταχύτητα σύγκλισης τους.

$$e^x - y = 0$$

$$xy - e^x = 0$$

$$xyz - x^2 + y^2 = 1.34$$

$$xy - z^2 = 0.09$$

$$e^x - e^y + z = 0.41$$

$$x^3 - y = 0$$

$$x^2 + y^3 = 1$$

ΓΡΑΜΜΙΚΑ ΣΥΣΤΗΜΑΤΑ

Η λύση συστημάτων γραμμικών εξισώσεων απαντάται σε μια πλειάδα προβλημάτων κάθε επιστημονικού κλάδου. Σε πολλές περιπτώσεις, χρειάζεται να λύσει κανείς γραμμικά συστήματα πολύ μεγάλου μεγέθους, π.χ. 1000×1000 συντελεστών. Η λύση ενός τέτοιου συστήματος απαιτεί ένα πολύ μεγάλο αριθμό πράξεων. Ακόμη και με τη χρήση γρήγορου υπολογιστή, η λύση μπορεί να είναι χρονοβόρα και να απαιτεί μεγάλη υπολογιστική μνήμη. Επίσης, όσο μεγαλύτερος είναι ο αριθμός των βημάτων, τόσο αυξάνει το συνολικό σφάλμα στρογγύλευσης των αριθμητικών πράξεων στη μνήμη του υπολογιστή.

Οι μέθοδοι επίλυσης γραμμικών συστημάτων κατατάσσονται σε δύο κατηγορίες, στις **άμεσες** και στις **επαναληπτικές** μεθόδους. Οι άμεσες μέθοδοι μπορούν να οδηγήσουν σε ακριβή λύση του συστήματος, ενώ οι επαναληπτικές μέθοδοι ξεκινούν με κάποια αρχική εκτίμηση της λύσης και βελτιώνουν την εκτίμηση μέχρι κάποια επιθυμητή ακρίβεια. Για πολύ μεγάλα γραμμικά συστήματα, οι επαναληπτικές μέθοδοι είναι ασύγκριτα πιο γρήγορες απ' ό,τι οι άμεσες μέθοδοι. Για ορισμένους τύπους γραμμικών συστημάτων π.χ. για συστήματα που ανάγονται σε συμμετρικό πίνακα ή σε πίνακα με διαγώνια δομή, έχουν αναπτυχθεί ειδικές μέθοδοι που επιτυγχάνουν την επίλυση του συστήματος με λιγότερες πράξεις απ' ό,τι με τη χρήση γενικών μεθόδων.

Τα προβλήματα που συνήθως πρέπει να λύσουμε είναι:

- Να λυθεί το σύστημα: $\mathbf{A} \cdot \mathbf{x} = \mathbf{B}$, όπου

$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ a_{12} & & a_{2N} \\ \vdots & & \vdots \\ a_{1N} & \cdots & a_{NN} \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{pmatrix}$$

- Να βρεθεί ο \mathbf{A}^{-1} (αντίστροφος πίνακας)
- Να βρεθεί η ορίζουσα του \mathbf{A}
- Να βρεθούν οι ιδιοτιμές και τα ιδιοδιανύσματα του \mathbf{A}

2.1 ΜΕΘΟΔΟΣ GAUSS

Σε αυτή την ενότητα θα παρουσιάσουμε μια βασική μέθοδο για την επίλυση γραμμικών συστημάτων N εξισώσεων με N αγνώστους. Το βασικό βήμα είναι η μετατροπή του συστήματος σε ένα άνω-τριγωνικό σύστημα οπότε στη συνέχεια η διαδικασία υπολογισμού των λύσεων είναι απλή.

Υπενθυμίζουμε ότι στα γραμμικά συστήματα επιτρέπονται οι παρακάτω πράξεις που δεν αλλοιώνουν τις λύσεις του αρχικού συστήματος :

- Αλλαγή της σειράς δύο εξισώσεων
- Πολλαπλασιασμός μιας εξίσωσης με μία μη-μηδενική σταθερά
- Μια εξίσωση μπορεί να αντικατασταθεί από το άθροισμα αυτής της εξίσωσης και ένα πολλαπλάσιο κάποιας από τις υπόλοιπες.

Ας υποθέσουμε ότι δίδεται το γραμμικό σύστημα: $\mathbf{A} \cdot \mathbf{x} = \mathbf{B}$ με $\det(\mathbf{A}) \neq 0$.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1N}x_N &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2N}x_N &= b_2 \\ &\vdots \\ a_{N1}x_1 + a_{N2}x_2 + a_{N3}x_3 + \dots + a_{NN}x_N &= b_N \end{aligned} \quad (2.1)$$

Για τη μετατροπή του παραπάνω συστήματος σε άνω τριγωνικό θα ακολουθήσουμε τα επόμενα βήματα:

• ΒΗΜΑ 1ο

Πολλαπλασιάζω την πρώτη εξίσωση με a_{21}/a_{11} και την αφαιρώ από τη δεύτερη εξίσωση. Ομοίως, πολλαπλασιάζω την πρώτη με a_{31}/a_{11} και την αφαιρώ από την τρίτη κ.ο.κ., οπότε λαμβάνω:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1N}x_N &= b_1 \\ 0 + a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \dots + a_{2N}^{(1)}x_N &= b_2^{(1)} \\ &\vdots \\ 0 + a_{N2}^{(1)}x_2 + a_{N3}^{(1)}x_3 + \dots + a_{NN}^{(1)}x_N &= b_N^{(1)} \end{aligned} \quad (2.2)$$

όπου, για παράδειγμα, είναι

$$a_{22}^{(1)} = \frac{a_{12}a_{21}}{a_{11}} - a_{22}$$

- **ΒΗΜΑ 2ο**

Η πρώτη γραμμή και η πρώτη στήλη παραμένουν οπότε πολλαπλασιάζω αντίστοιχα τη δεύτερη εξίσωση με $a_{32}^{(1)}/a_{22}^{(1)}$ και την αφαιρώ από την τρίτη εξίσωση κ.ο.κ., οπότε

$$\begin{aligned}
 a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1N}x_N &= b_1 \\
 0 + a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \cdots + a_{2N}^{(1)}x_N &= b_2^{(1)} \\
 0 + 0 + a_{33}^{(2)}x_3 + \cdots + a_{3N}^{(2)}x_N &= b_3^{(2)} \\
 &\vdots \\
 0 + 0 + a_{N3}^{(2)}x_3 + \cdots + a_{NN}^{(2)}x_N &= b_N^{(2)}
 \end{aligned} \tag{2.3}$$

Αντίστοιχα, για παράδειγμα, είναι:

$$a_{33}^{(2)} = \frac{a_{32}^{(1)}a_{33}^{(1)}}{a_{22}^{(1)}} - a_{33}^{(1)}$$

- Μετά από **N - 1 ΒΗΜΑΤΑ** καταλήγω στο σύστημα:

$$\begin{aligned}
 a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1N}x_N &= b_1 \\
 a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \cdots + a_{2N}^{(1)}x_N &= b_2^{(1)} \\
 a_{33}^{(2)}x_3 + \cdots + a_{3N}^{(2)}x_N &= b_3^{(2)} \\
 &\vdots \\
 0 + 0 + \cdots + a_{NN}^{(N-1)}x_N &= b_N^{(N-1)}
 \end{aligned} \tag{2.4}$$

Το παραπάνω σύστημα είναι τριγωνικό και η λύση του μπορεί να δοθεί εύκολα ακολουθώντας την παρακάτω διαδικασία.

- Προφανώς για από τη N-οστή εξίσωση του παραπάνω συστήματος (2.5) λαμβάνουμε την τιμή του x_N :

$$x_N = \frac{b_N^{(N-1)}}{a_{NN}^{(N-1)}} \tag{2.5}$$

- ενώ οι υπόλοιπες ρίζες του συστήματος υπολογίζονται εύκολα από την σχέση:

$$x_i = \frac{b_i^{(i-1)} - \sum_{k=i+1}^N a_{ik}^{(i-1)}x_k}{a_{ii}^{(i-1)}} \tag{2.6}$$

Ο αριθμός των πράξεων που απαιτεί η λύση ενός συστήματος N γραμμικών εξισώσεων με τη μέθοδο Gauss είναι $(4N^3 + 9N^2 - 7N)/6$ (γιατί).

Χρήση της Οδήγησης (Pivoting)

Για τη βελτίωση της ακρίβειας των υπολογισμών φροντίζουμε στο σύστημα (2.3) το a_{11} να είναι το απολύτως μεγαλύτερο στοιχείο. Αντίστοιχα, στο σύστημα (2.4) το $a_{22}^{(1)}$ πρέπει να είναι το απολύτως μεγαλύτερο στοιχείο κ.ο.κ.

Το pivoting μας προφυλάσσει και από μια πιθανή παθολογική συμπεριφορά της μεθόδου Gauss. Η παθολογική συμπεριφορά οφείλεται στην ιδιότητα της μεθόδου Gauss σύμφωνα με την οποία διαιρούμε σε κάθε βήμα με το στοιχείο a_{NN} , επομένως, αν κάποιο από αυτά τα διαγώνια στοιχεία είναι μηδέν ή γενικά ένας πολύ μικρός αριθμός, η ακρίβεια της μεθόδου είναι πολύ κακή. Με το pivoting αποφεύγουμε τέτοιες παθολογικές καταστάσεις.

ΠΑΡΑΔΕΙΓΜΑ

Εστω ότι, μετά από τη χρήση της μεθόδου Gauss στη λύση ενός $N \times N$ συστήματος, οι δυο τελευταίες εξισώσεις, η $N - 1$ και η N είναι:

$$\begin{aligned} 0x_{N-1} + x_N &= 1 \\ 2x_{N-1} + x_N &= 3 \end{aligned}$$

Προφανώς, έχουν λύση: $x_{N-1} = x_N = 1$. Αλλά, λόγω σφαλμάτων αποκοπής, στην πράξη το σύστημα θα έχει τη μορφή:

$$\begin{aligned} \epsilon x_{N-1} + x_N &= 1 \\ 2x_{N-1} + x_N &= 3 \end{aligned}$$

οπότε συνεχίζοντας στο τελευταίο βήμα θα πάρουμε:

$$\begin{aligned} \epsilon x_{N-1} + x_N &= 1 \\ \left(1 - \frac{2}{\epsilon}\right) x_N &= 3 - \frac{2}{\epsilon} \end{aligned}$$

με λύσεις

$$\begin{aligned} x_N &= \frac{3 - \frac{2}{\epsilon}}{1 - \frac{2}{\epsilon}} \approx 1 \quad \text{σωστό} \\ x_{N-1} &= \frac{1 - x_N}{\epsilon} ! \end{aligned}$$

Παρατηρούμε δηλαδή, πως το x_{N-1} είναι απροσδιόριστο, γιατί αποτελεί το λόγο δυο μικρών αριθμών, των οποίων η ακρίβεια εξαρτάται από την ακρίβεια με την οποία εκτελεί τις πράξεις ο Η/Υ. Στη συνέχεια, ο όρος x_{N-1} θα χρησιμοποιηθεί για τον υπολογισμό των x_{N-2}, \dots, x_1 με καταστροφικά αποτελέσματα.

Αν όμως χρησιμοποιηθεί **οδήγηση**, το σύστημα γράφεται:

$$2x_{N-1} + x_N = 3$$

$$\epsilon x_{N-1} + x_N = 1$$

οπότε πολλαπλασιάζοντας επί ϵ την πρώτη και επί 2 την δεύτερη καταλήγουμε

$$\begin{aligned} 2x_{N-1} + x_N &= 3 \\ \left(1 - \frac{\epsilon}{2}\right) x_N &= 1 - \frac{3\epsilon}{2} \end{aligned}$$

οπότε η λύση του είναι $x_N \approx 1.0$ και $x_{N-1} = \frac{3-x_N}{2} \approx 1.0$.

2.2 ΜΕΘΟΔΟΣ GAUSS-JORDAN

Η μέθοδος Gauss-Jordan αποτελεί παραλλαγή της μεθόδου Gauss. Η βασική της διαφορά έγκειται στο ότι το τελικό σύστημα, δεν είναι τριγωνικό αλλά πρακτικά διαγωνιοποιείται ο πίνακας \mathbf{A} και η λύση του συστήματος είναι προφανής.

Η μέθοδος βασίζεται στην εξής λογική: Σε κάθε βήμα απαλείφουμε κι έναν όρο από την πρώτη εξίσωση. Δηλαδή, στο σύστημα (2.2) του προηγούμενου κεφαλαίου, πολλαπλασιάζω τη δεύτερη εξίσωση με $a_{12}/a_{22}^{(1)}$ και την αφαιρώ από την πρώτη, οπότε το σύστημα 2.3 γράφεται:

$$\begin{aligned} a_{11}x_1 + 0 + a_{13}^{(2)}x_3 + \dots + a_{1N}^{(2)}x_N &= b_1^{(2)} \\ 0 + a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \dots + a_{2N}^{(1)}x_N &= b_2^{(1)} \\ 0 + 0 + a_{33}^{(2)}x_3 + \dots + a_{3N}^{(2)}x_N &= b_3^{(2)} \\ &\vdots \\ 0 + 0 + a_{N3}^{(2)}x_3 + \dots + a_{NN}^{(2)}x_N &= b_N^{(2)} \end{aligned} \quad (2.7)$$

στο επόμενο βήμα θα απαλειφθούν οι όροι με συντελεστές κ.ο.κ. Οπότε τελικά μετά από N βήματα καταλήγω στο σύστημα:

$$\begin{aligned} a_{11}x_1 &= b_1^{(N-1)} \\ a_{22}^{(1)}x_2 &= b_2^{(N-1)} \\ &\vdots \\ a_{NN}^{(N-1)}x_N &= b_N^{(N-1)} \end{aligned} \quad (2.8)$$

του οποίου η λύση του δίνεται από την προφανή σχέση

$$x_i = \frac{b_i^{(N-1)}}{a_{ii}^{(i-1)}}. \quad (2.9)$$

2.3 ΜΕΘΟΔΟΣ L-U

Προηγουμένως είδαμε πόσο εύκολα μπορεί να λυθεί ένα άνω-τριγωνικό σύστημα. Εδώ θα εξετάσουμε τη διαδικασία της 'παραγοντοποίησης' ενός πίνακα **A**. Κάθε πίνακας **A** μπορεί να γραφεί ως γινόμενο δυο πινάκων που περιέχουν τη διαγώνιο και τα πάνω ή κάτω απ' αυτή στοιχεία του πίνακα. Για παράδειγμα :

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} = \begin{pmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{pmatrix} \cdot \begin{pmatrix} 1 & u_{12} & u_{13} & u_{14} \\ 0 & 1 & u_{23} & u_{24} \\ 0 & 0 & 1 & u_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{A} = \underset{(lower)}{\mathbf{L}} \cdot \underset{(upper)}{\mathbf{U}}$$

Στόχος μας είναι ο υπολογισμός των στοιχείων των πινάκων **L** και **U**. Αυτό θα γίνει με τη λύση ενός συστήματος $N \times N$ εξισώσεων που στην περίπτωση μας όπου ο πίνακας **A** είναι 4×4 αναγόμαστε στη λύση ενός συστήματος 16 εξισώσεων.

Πολλαπλασιάζοντας τις γραμμές του **L** με την πρώτη στήλη του **U**, λαμβάνουμε:

$$l_{11} = a_{11}, \quad l_{21} = a_{21}, \quad l_{31} = a_{31}, \quad l_{41} = a_{41} \quad (2.10)$$

Δηλαδή, η πρώτη στήλη του **L** είναι η ίδια με την πρώτη στήλη του **A**. Στη συνέχεια, πολλαπλασιάζουμε την πρώτη γραμμή του **L** με τις στήλες του **U**, οπότε λαμβάνουμε:

$$u_{12} = \frac{a_{12}}{l_{11}}, \quad u_{13} = \frac{a_{13}}{l_{11}}, \quad u_{14} = \frac{a_{14}}{l_{11}} \quad (2.11)$$

Άρα και η πρώτη γραμμή του **U** υπολογίστηκε. Συνεχίζοντας υπολογίζουμε τη δεύτερη στήλη του **L**.

$$\left. \begin{aligned} l_{21}u_{12} + l_{22} &= a_{22} \\ l_{31}u_{12} + l_{32} &= a_{32} \\ l_{41}u_{12} + l_{42} &= a_{42} \end{aligned} \right\} \Rightarrow \begin{aligned} l_{22} &= a_{22} - l_{21}u_{12} \\ l_{32} &= a_{32} - l_{31}u_{12} \\ l_{42} &= a_{42} - l_{41}u_{12} \end{aligned} \quad (2.12)$$

Όπως παρατηρούμε, τα l_{21}, l_{31}, l_{41} και u_{12} έχουν ήδη υπολογισθεί, επομένως και η δεύτερη στήλη του πίνακα **L** υπολογίστηκε. Με ανάλογο τρόπο, οι σχέσεις

$$\begin{aligned} u_{23} &= \frac{a_{23} - l_{21}u_{13}}{l_{22}} \\ l_{33} &= a_{33} - l_{31}u_{13} - l_{32}u_{23} \\ u_{24} &= \frac{a_{24} - l_{21}u_{14}}{l_{22}} \\ l_{43} &= a_{43} - l_{41}u_{13} - l_{42}u_{23} \\ u_{34} &= \frac{a_{34} - l_{31}u_{14} - l_{32}u_{24}}{l_{33}} \\ l_{44} &= a_{44} - l_{41}u_{14} - l_{42}u_{24} - l_{43}u_{34} \end{aligned} \quad (2.13)$$

μας δίνουν τα υπόλοιπα στοιχεία των δυο πινάκων **U** και **L**.

ΓΕΝΙΚΟΣ ΤΥΠΟΣ ΓΙΑ 'L-U' ΑΝΑΛΥΣΗ

Με βάση τα προηγούμενα οδηγούμαστε στη δημιουργία ενός 'κανόνα' υπολογισμού των στοιχείων των δύο πινάκων **L** και **U**.

$$l_{ij} = a_{ij} - \sum_{k=1}^{j-1} l_{ik} u_{kj} \quad \text{για } j \leq i \quad \text{και } i = 1, 2, \dots, N \quad (2.14)$$

$$u_{ji} = \frac{a_{ji} - \sum_{k=1}^{j-1} l_{jk} u_{ki}}{l_{jj}} \quad \text{για } j \leq i \quad \text{και } j = 2, 3, \dots, N \quad (2.15)$$

$$(2.16)$$

Για την ειδική περίπτωση του $j = 1$ τα l_{i1} θα είναι $l_{i1} = a_{i1}$ ενώ για $i = 1$ ο κανόνας για τα u_{1j} είναι

$$u_{1j} = \frac{a_{1j}}{l_{11}} = \frac{a_{1j}}{a_{11}}.$$

ΠΑΡΑΔΕΙΓΜΑ

Θα εφαρμόσθει η μέθοδος LU για τον πίνακα

$$\mathbf{A} = \begin{pmatrix} 3 & -1 & 2 \\ 1 & 2 & 3 \\ 2 & -2 & 1 \end{pmatrix}$$

οπότε με βάση τις παραπάνω σχέσεις βρίσκουμε:

$$l_{11} = 3, \quad l_{21} = 1, \quad l_{31} = 2, \quad u_{12} = -\frac{1}{3}, \quad u_{13} = \frac{2}{3}$$

$$l_{22} = 2 - 1 \cdot \left(-\frac{1}{3}\right) = \frac{7}{3}, \quad l_{32} = -\frac{4}{3}, \quad u_{23} = 1, \quad l_{33} = -1$$

οπότε τελικά οι ζητούμενοι πίνακες είναι:

$$\mathbf{L} = \begin{pmatrix} 3 & 0 & 0 \\ 1 & 7/3 & 0 \\ 2 & 4/3 & -1 \end{pmatrix} \quad \mathbf{U} = \begin{pmatrix} 1 & -1/3 & 2/3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Εφαρμογή της 'L-U' στα γραμμικά συστήματα

Με τη μέθοδο L-U ένα γραμμικό σύστημα της μορφής $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ ανάγεται στο σύστημα

$$\mathbf{L} \cdot \mathbf{U} \cdot \mathbf{x} = \mathbf{b} \quad (2.17)$$

Λύνουμε πρώτα το κάτω-τριγωνικό σύστημα $\mathbf{L} \cdot \mathbf{b}' = \mathbf{b}$ σύμφωνα με τον κανόνα:

ΑΛΓΟΡΙΘΜΟΣ για τη μέθοδο LU

INPUT N, a_{ij}
 OUTPUT (αναμενόμενο) l_{ij}, u_{ij}

STEP 1
 $l_{11} = a_{11}$ IF $l_{11} = 0$ STOP (Αδύνατη η εφαρμογή της LU)

STEP 2
 FOR $j=2, \dots, N$ SET $u_{1j} = a_{1j}/l_{11}$ (υπολογισμός της πρώτης σειράς του U)
 SET $l_{j1} = a_{j1}$ (υπολογισμός της πρώτης στήλης του L)

STEP 3
 FOR $I=2, \dots, N-1$ DO STEPS 4 & 5

STEP 4
 $l_{ii} = a_{ii} - \sum_{k=1}^{i-1} l_{ik}u_{ki}$
 IF $l_{ii} = 0$ STOP (αδύνατη η εφαρμογή της LU)

STEP 5
 FOR $j=i+1, \dots, N$
 SET $u_{ij} = \frac{1}{l_{ii}} \left(a_{ij} - \sum_{k=1}^{i-1} l_{ik}u_{kj} \right)$ (i γραμμή του U)
 SET $l_{ji} = a_{ji} - \sum_{k=1}^{i-1} l_{jk}u_{ki}$ (i γραμμή του L)

STEP 6
 SET $l_{NN} = a_{NN} - \sum_{k=1}^{N-1} l_{Nk}u_{kN}$ (Αν $l_{NN} = 0$ τότε $A = L \cdot U$ αλλά είναι singular)

STEP 7
 OUTPUT l_{ij}, u_{ij}

Πίνακας 2.1. Λογικό διάγραμμα εφαρμογής της μεθόδου LU

$$b'_1 = \frac{b_1}{l_{11}}, \quad b'_i = \frac{b_i - \sum_{k=1}^{i-1} l_{ik}b'_k}{l_{ii}} \quad \text{για } i = 2, 3, \dots, N \quad (2.18)$$

οπότε απομένει να λύσουμε το τριγωνικό σύστημα $U \cdot x = b'$ με τις σχέσεις που δόθηκαν στη μέθοδο Gauss, για την επίλυση τριγωνικών συστημάτων. Η μέθοδος αυτή χρησιμοποιείται όταν έχουμε να λύσουμε ένα μεγάλο αριθμό γραμμικών συστημάτων που διαφέρουν μόνο στο διάνυσμα b .

ΠΑΡΑΔΕΙΓΜΑ

Αν το διάνυσμα b είναι $b = (12, 11, 2)^T$ και A ο πίνακας του προηγούμενου παραδείγματος, τότε το διάνυσμα b' θα έχει τη μορφή:

$$b'_1 = \frac{12}{3} = 4, \quad b'_2 = \frac{11 - 1 \cdot 4}{7/3}, \quad b'_3 = \dots = -2$$

Άρα καταλήγουμε ότι:

$$\begin{pmatrix} 1 & -1/3 & 2/3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 2 \end{pmatrix}$$

δηλαδή $x_3 = 2$, $x_2 = 1$ και $x_1 = 3$.

2.4 ΜΕΘΟΔΟΣ JACOBI (Επαναληπτική)

Η μέθοδος αυτή είναι γενίκευση της μεθόδου $x = g(x)$ που έχουμε χρησιμοποιήσει για την εύρεση ριζών μη-γραμμικών εξισώσεων και συστημάτων στα Κεφάλαια 1.4 και 1.7. Προφανώς, η μέθοδος αυτή δεν είναι ακριβής αλλά υπο κατάλληλες συνθήκες μπορεί να επιτύχει υψηλή ακρίβεια με μικρό αριθμό υπολογιστικών πράξεων και πολύ απλό, προγραμματιστικά, κώδικα.

Έστω το σύστημα των N εξισώσεων με N αγνώστους :

$$\begin{aligned} f_1(x_1, x_2, \dots, x_N) &= 0 \\ f_2(x_1, x_2, \dots, x_N) &= 0 \\ &\dots \\ f_n(x_1, x_2, \dots, x_N) &= 0 \end{aligned} \quad (2.19)$$

που εύκολα μπορεί να γραφεί στη μορφή :

$$\begin{aligned} x_1 &= g_1(x_2, x_3, \dots, x_N) \\ x_2 &= g_2(x_1, x_3, \dots, x_N) \\ &\dots \\ x_N &= g_N(x_1, x_2, \dots, x_{N-1}) \end{aligned} \quad (2.20)$$

με αυτό τον τρόπο το αντιμετωπίζουμε όπως τα μη γραμμικά συστήματα του κεφαλαίου 1.7, με τη διαφορά ότι απουσιάζουν οι μη-γραμμικοί όροι. Πρακτικά κάθε μια από τις N εξισώσεις θα μπορούσε να γραφεί στη μορφή :

$$x_i = \frac{b_i}{a_{ii}} - \frac{1}{a_{ii}} \sum_{\substack{j=1 \\ j \neq i}}^N a_{ij} x_j \quad (2.21)$$

και στη συνέχεια δίνοντας N 'αυθαίρετες' αρχικές τιμές $x_1^{(0)}, x_2^{(0)}, \dots, x_N^{(0)}$, δημιουργώ μια γενικευμένη αναδρομική σχέση της μορφής

$$x_i^{(k+1)} = g_i(x_1^{(k)}, \dots, x_N^{(k)}) \quad (2.22)$$

Ικανή **συνθήκη σύγκλισης** είναι η :

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^N |a_{ij}| \quad (2.23)$$

που αν ικανοποιείται η σύγκλιση είναι βέβαιη και ανεξάρτητη από τις αρχικές τιμές $x_1^{(0)}, x_2^{(0)}, \dots, x_N^{(0)}$.

Σε μορφή πινάκων η αναδρομική σχέση μπορεί να γραφεί ως εξής:

$$\mathbf{x}^{(k+1)} = \mathbf{D}^{-1}\mathbf{B} - \mathbf{D}^{-1}\mathbf{C}\mathbf{x}^{(k)}$$

όπου $\mathbf{A} = \mathbf{D} + \mathbf{C}$ δηλ. ο \mathbf{D} περιέχει τα διαγώνια στοιχεία του \mathbf{A} και ο \mathbf{C} όλα τα υπόλοιπα στοιχεία έχοντας θέσει 0 τα διαγώνια.

ΠΑΡΑΔΕΙΓΜΑ

Από το σύστημα

$$\begin{aligned} 4x - y + z &= 7 \\ 4x - 8y + z &= -21 \\ -2x + y + 5z &= 15 \end{aligned}$$

(που έχει λύσεις $x = 2, y = 4, z = 3$) μπορώ να δημιουργήσω τις αναδρομικές σχέσεις:

$$\begin{aligned} x^{(k+1)} &= \frac{7 + y^{(k)} - z^{(k)}}{4} \\ y^{(k+1)} &= \frac{21 + 4x^{(k)} + z^{(k)}}{8} \\ z^{(k+1)} &= \frac{15 + 2x^{(k)} - y^{(k)}}{5} \end{aligned}$$

και θέτοντας ως αρχικές τιμές $(1, 2, 2)$ παίρνω την ακολουθία προσεγγιστικών λύσεων

$$\begin{aligned} (1, 2, 2) &\rightarrow (1.75, 3.375, 3) \\ &\rightarrow (1.844, 3.875, 3.025) \\ &\rightarrow (1.963, 3.925, 2.963) \\ &\rightarrow (1.991, 3.977, 3.0) \\ &\rightarrow (1.994, 3.995, 3.001) \end{aligned}$$

Δηλαδή, απαιτήθηκαν 5 συνολικά επαναλήψεις, για να βρούμε τη λύση του συστήματος με ακρίβεια 3 δεκαδικών ψηφίων.

2.5 ΜΕΘΟΔΟΣ GAUSS – SEIDEL

Είναι επαναληπτική μέθοδος και αποτελεί βελτίωση της μεθόδου Jacobi ώστε να επιτυγχάνεται ταχύτερα η σύγκλιση στις ρίζες του συστήματος.

Αν θεωρήσω κάποιες αρχικές τιμές για το διάνυσμα x_i έστω τις $x_i^{(0)}$ υπολογίζω από την πρώτη εξίσωση το $x_1^{(1)}$ και στη συνέχεια το χρησιμοποιώ στη δεύτερη εξίσωση μαζί με τα αρχικά $(x_2^{(0)}, x_3^{(0)}, \dots, x_N^{(0)})$ για τον υπολογισμό του $x_2^{(1)}$. Στη συνέχεια τα $x_1^{(1)}$ και $x_2^{(1)}$ χρησιμοποιούνται στον υπολογισμό του $x_3^{(1)}$ με αυτό τον τρόπο από τις υπόλοιπες $N - 3$ εξισώσεις λαμβάνω τα $x_i^{(1)}$.

Στη γενική περίπτωση για τον υπολογισμό του x_1 μετά από k επαναλήψεις θα χρησιμοποιώ μια σχέση της μορφής :

$$x_1^{(k+1)} = \frac{1}{a_{11}} \left(b_1 - \sum_{j=2}^N a_{1j} x_j^{(k)} \right) \quad (2.24)$$

στη δεύτερη εξίσωση από την οποία θα υπολογίζω το x_2 μετά από k επαναλήψεις αντικαθιστώ το $(x_1^{(k+1)}, x_2^{(k)}, \dots, x_N^{(k)})$ και υπολογίζω το $x_2^{(k+1)}$ από τη σχέση

$$x_2^{(k+1)} = \frac{1}{a_{22}} \left(b_2 - a_{21} x_1^{(k+1)} - \sum_{j=3}^N a_{2j} x_j^{(k)} \right) \quad (2.25)$$

οπότε στην τρίτη εξίσωση θέτω $(x_1^{(k+1)}, x_2^{(k+1)}, x_3^{(k)}, \dots, x_N^{(k)})$ κ.ο.κ. Άρα, ο γενικός τύπος θα είναι :

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^N a_{ij} x_j^{(k)} \right) \quad (2.26)$$

Και εδώ, η συνθήκη σύγκλισης είναι :

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^N |a_{ij}| \quad i = 1, 2, \dots, N \quad (2.27)$$

$$\mathbf{x}^{(k+1)} = \mathbf{D}^{-1} \left(\mathbf{B} - \mathbf{L}\mathbf{x}^{(k+1)} - \mathbf{U}\mathbf{x}^{(k)} \right) \quad (2.28)$$

όπου $\mathbf{A} = \underset{\text{lower}}{\mathbf{L}} + \underset{\text{diagonal}}{\mathbf{D}} + \underset{\text{upper}}{\mathbf{U}}$. Ο πίνακας \mathbf{L} περιέχει τα στοιχεία του πίνακα \mathbf{A} που βρίσκονται κάτω από τη διαγώνιο, ο πίνακας \mathbf{D} μόνο τα διαγώνια στοιχεία του \mathbf{A} και τέλος ο πίνακας \mathbf{U} τα στοιχεία του πίνακα \mathbf{A} που βρίσκονται πάνω από τη διαγώνιο.

ΠΑΡΑΔΕΙΓΜΑ

Εφαρμόστε την μέθοδο Gauss-Seidel στο παράδειγμα της προηγούμενης παραγράφου και συγκρίνετε την ταχύτητα σύγκλισης.

Πρακτικά οι προηγούμενες αναδρομικές σχέσεις θα γραφούν ως

$$\begin{aligned}x^{(k+1)} &= \frac{7 + y^{(k)} - z^{(k)}}{4} \\y^{(k+1)} &= \frac{21 + 4x^{(k)} + z^{(k)}}{8} \\z^{(k+1)} &= \frac{15 + 2x^{(k)} - y^{(k)}}{5}\end{aligned}$$

Παίρνω την ακολουθία τιμών:

$$\begin{aligned}(1, 2, 2) &\rightarrow (1.75, 3.75, 2.95) \\&\rightarrow (1.95, 3.97, 2.99) \\&\rightarrow (1.996, 3.996, 2.999)\end{aligned}$$

δηλαδή, σε τρεις επαναλήψεις υπολογίστηκαν οι ακριβείς λύσεις του συστήματος, που είναι $(2, 4, 3)$. Σημειώστε ότι με τη μέθοδο Jacobi απαιτήθηκαν 5 συνολικά επαναλήψεις, για να επιτευχθεί η ίδια ακρίβεια, δηλαδή η μέθοδος Gauss-Seidel απαιτεί σχεδόν τα μισά βήματα, για να συγκλίνει στην ακριβή λύση με την ίδια ακρίβεια.

2.6 Η ΟΡΙΖΟΥΣΑ ΚΑΙ Ο ΑΝΤΙΣΤΡΟΦΟΣ ΕΝΟΣ ΠΙΝΑΚΑ

2.6.1 Η ορίζουσα

Η **ορίζουσα** ενός πίνακα υπολογίζεται εύκολα, αν έχουμε ήδη εφαρμόσει τη μέθοδο Gauss σε έναν πίνακα. Τότε η ορίζουσα θα είναι το γινόμενο των διαγωνίων στοιχείων του πίνακα. Για παράδειγμα:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{12} & a_{23} \\ 0 & a_{33} \end{vmatrix} = a_{11} \cdot a_{22} \cdot a_{33}$$

2.6.2 Ο αντίστροφος πίνακας

Ο υπολογισμός του **αντίστροφου** ενός πίνακα διαστάσεων $N \times N$, απαιτεί ουσιαστικά την επίλυση N γραμμικών συστημάτων.

Δηλαδή, αν δοθεί ο πίνακας $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ και $\begin{pmatrix} x & y \\ z & w \end{pmatrix}$ είναι ο αντίστροφός του.

Τότε προφανώς θα είναι:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

και για τον υπολογισμό των αγνώστων x, y, z και w θα πρέπει να λυθούν τα παρακάτω δύο συστήματα:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{και} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} y \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Παρατηρούμε όμως ότι αν εφαρμόσουμε τον κανόνα Gauss-Jordan για το ένα σύστημα, δεν χρειάζεται να το επαναλάβουμε και για το δεύτερο. Το μόνο που αλλάζει είναι οι πράξεις που γίνονται στις αντίστοιχες στήλες του μοναδιαίου πίνακα. Οπότε, στην ουσία, προσπαθούμε να κάνουμε στο μοναδιαίο πίνακα τις ίδιες ακριβώς πράξεις που κάνουμε και στο δοθέντα πίνακα. Δηλαδή θα γράψουμε (σχηματικά) τους πίνακες ως εξής :

$$\begin{array}{cccc|cccc} a_{11} & a_{12} & \dots & a_{1N} & 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2N} & 0 & 1 & \dots & 0 \\ & & & & & & & \\ & & & & & & & \\ a_{N1} & a_{N2} & \dots & a_{NN} & 0 & 0 & \dots & 1 \\ \mathbf{A} & & & & \mathbf{I} & & & \end{array} \quad (2.29)$$

και θα εφαρμόσω στον πίνακα \mathbf{I} ακριβώς τις ίδιες πράξεις που απαιτεί ο κανόνας Gauss - Jordan για τη διαγωνοποίηση του πίνακα \mathbf{A} .

$$\begin{array}{cccc|cccc} 1 & 0 & \dots & 0 & \tilde{a}_{11} & \tilde{a}_{12} & \dots & \tilde{a}_{1N} \\ 0 & 1 & \dots & 0 & \tilde{a}_{21} & \tilde{a}_{22} & \dots & \tilde{a}_{2N} \\ & & & & & & & \\ & & & & & & & \\ 0 & 0 & \dots & 1 & \tilde{a}_{N1} & \tilde{a}_{N2} & \dots & \tilde{a}_{NN} \end{array} \quad (2.30)$$

Οπότε τελικά στην αριστερή πλευρά δημιουργούμε τον μοναδιαίο πίνακα και στη δεξιά τον αντίστροφο του $\tilde{a}_{ij} = \mathbf{A}^{-1}$.

2.7 ΙΔΙΟΤΙΜΕΣ ΚΑΙ ΙΔΙΟΔΙΑΝΥΣΜΑΤΑ

Έστω ένας $N \times N$ πίνακας (τα στοιχεία του οποίου μπορεί να είναι και μιγαδικοί αριθμοί) και έστω λ ένας πραγματικός (ή και μιγαδικός) αριθμός, για τον οποίο η εξίσωση

$$\mathbf{A} \cdot \mathbf{x} = \lambda \mathbf{x} \quad (2.31)$$

έχει μια μη τετριμμένη λύση (δηλαδή, $\mathbf{x} \neq \mathbf{0}$). Η τιμή λ ορίζεται ως μια **ιδιοτιμή** του πίνακα \mathbf{A} . Ένα μη μηδενικό διάνυσμα \mathbf{x} , που ικανοποιεί την εξίσωση (2.31), ονομάζεται **ιδιοδιάνυσμα** του \mathbf{A} , που αντιστοιχεί στην τιμή λ .

Για παράδειγμα, στην εξίσωση

$$\begin{pmatrix} 1 & 2 & 3 \\ -1 & 3 & 1 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} = -1 \cdot \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$$

$\mathbf{A} \quad \mathbf{u}_1 \quad \lambda_1 \quad \mathbf{u}_1$

το διάνυσμα $\mathbf{u}_1 = (2, 1, -2)^\top$ είναι ένα ιδιοδιάνυσμα και $\lambda_1 = -1$ η αντίστοιχη ιδιοτιμή του \mathbf{A} .

Η εξίσωση (2.31) είναι ισοδύναμη με την εξίσωση

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0 \quad (2.32)$$

οπότε, λύνοντας την (2.32), που ονομάζεται **χαρακτηριστική εξίσωση**, μπορούμε να υπολογίσουμε τις ιδιοτιμές του πίνακα \mathbf{A} . Για παράδειγμα, αν την εφαρμόσουμε στον πίνακα της εξίσωσης (2.7), δημιουργούμε την εξίσωση

$$\det \begin{vmatrix} 1 - \lambda & 2 & 3 \\ -1 & 3 - \lambda & 1 \\ 2 & 0 & 1 - \lambda \end{vmatrix} = \lambda^3 - 5\lambda^2 + 3\lambda + 9 = 0 \quad (2.33)$$

Το πολυώνυμο $\lambda^3 - 5\lambda^2 + 3\lambda + 9 = 0$ ονομάζεται **χαρακτηριστικό πολυώνυμο** και οι ρίζες του $\lambda_i = -1, 3$ και 3 είναι οι ιδιοτιμές του πίνακα \mathbf{A} .

Αυτή η κλασική διαδικασία εύρεσης των ιδιοτιμών, και στη συνέχεια, των ιδιοδιανυσμάτων ενός πίνακα συνίσταται μόνο για πίνακες με μικρό αριθμό στοιχείων. Για πίνακες με πολλά στοιχεία, το χαρακτηριστικό πολυώνυμο θα είναι μεγάλου βαθμού και οι συντελεστές του θα έχουν υπολογιστεί με ακρίβεια που θα ελαττώνεται, καθώς θα αυξάνεται η τάξη του πίνακα. Όμως οι ρίζες των πολυωνύμων είναι πολύ ευαίσθητες έστω και σε μικρά σφάλματα των συντελεστών του, επομένως, η ακρίβεια, με την οποία θα υπολογιστούν οι ιδιοτιμές, θα είναι περιορισμένη. Για το λόγο αυτό, θα παρουσιάσουμε στη συνέχεια μια διαδικασία κατάλληλη για τον ακριβή υπολογισμό των ιδιοτιμών των πινάκων.

2.7.1 Η μέθοδος των δυνάμεων

Η **μέθοδος των δυνάμεων** και οι διάφορες παραλλαγές της έχουν σχεδιασθεί ειδικά για τον υπολογισμό των ιδιοτιμών και ιδιοδιανυσμάτων με τη χρήση Η/Υ. Η μέθοδος εφαρμόζεται αν υπάρχει μια ιδιοτιμή λ_1 που είναι απολύτως μεγαλύτερη από όλες τις υπόλοιπες, δηλαδή

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_N| \quad (2.34)$$

και επίσης, αν κάθε διάνυσμα \mathbf{x} μπορεί να γραφεί ως γραμμικός συνδυασμός των N ιδιοδιανυσμάτων $\{\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \dots, \mathbf{u}^{(N)}\}$. Δηλαδή, για κάθε $\mathbf{u}^{(i)}$ ισχύει:

$$\mathbf{A}\mathbf{u}^{(i)} = \lambda_i \mathbf{u}^{(i)} \quad (1 \leq i \leq N) \quad (2.35)$$

και

$$\mathbf{x} = a_1 \mathbf{u}^{(1)} + a_2 \mathbf{u}^{(2)} + \dots + a_N \mathbf{u}^{(N)} \quad (2.36)$$

Αν πολλαπλασιάσω και τα δυο μέλη της (2.36) με τον πίνακα \mathbf{A} , θα έχω:

$$\mathbf{x}^{(1)} \equiv \mathbf{A}\mathbf{x} = a_1 \lambda_1 \mathbf{u}^{(1)} + a_2 \lambda_2 \mathbf{u}^{(2)} + \dots + a_N \lambda_N \mathbf{u}^{(N)} \quad (2.37)$$

Αν πολλαπλασιάσω k φορές την (2.37) με τον πίνακα \mathbf{A} , θα καταλήξω σε μια σχέση της μορφής

$$\begin{aligned} \mathbf{x}^{(k)} &\equiv \mathbf{A}^k \mathbf{x} = a_1 \lambda_1^k \mathbf{u}^{(1)} + a_2 \lambda_2^k \mathbf{u}^{(2)} + \dots + a_N \lambda_N^k \mathbf{u}^{(N)} \\ &= \lambda_1^k \left(a_1 \mathbf{u}^{(1)} + a_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \mathbf{u}^{(2)} + \dots + a_N \left(\frac{\lambda_N}{\lambda_1} \right)^k \mathbf{u}^{(N)} \right) \end{aligned} \quad (2.38)$$

επειδή όμως η λ_1 είναι η απολύτως μεγαλύτερη ιδιοτιμή (σχέση 2.34), οι όροι $(\lambda_j/\lambda_1)^k$ τείνουν στο μηδέν, καθώς το $k \rightarrow \infty$.

Αρα, για $k \rightarrow \infty$ λαμβάνουμε προσεγγιστικά ότι:

$$\mathbf{x}^{(k)} = \mathbf{A}^k \mathbf{x} \approx \lambda_1^k a_1 \mathbf{u}^{(1)} \quad (2.39)$$

και επομένως, ο λόγος

$$\mathbf{r}_k \equiv \frac{\mathbf{x}^{(k+1)}}{\mathbf{x}^{(k)}} = \frac{\mathbf{A}^{k+1} \mathbf{x}}{\mathbf{A}^k \mathbf{x}} \approx \frac{\lambda_1^{k+1} a_1 \mathbf{u}^{(1)}}{\lambda_1^k a_1 \mathbf{u}^{(1)}} \rightarrow \lambda_1 \quad (2.40)$$

τείνει στην τιμή λ_1 , καθώς $k \rightarrow \infty$. Παρατηρούμε επίσης ότι από την (2.39) υπολογίζουμε προσεγγιστικά και το αντίστοιχο ιδιοδιάνυσμα $\mathbf{u}^{(1)} \equiv \mathbf{x}^{(k)}$.

ΠΑΡΑΔΕΙΓΜΑ

Έστω ο πίνακας $\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 2 & 2 \\ 1 & 0 & 3 \end{pmatrix}$ με ιδιοτιμές $\lambda_1 = 3.4142$, $\lambda_2 = 2$ και $\lambda_3 = 0.5858$ και αντίστοιχα ιδιοδιανύσματα $\mathbf{u}^{(1)} = (0.3694, 1, 0.8918)^\top$, $\mathbf{u}^{(2)} = (0, 1, 0)^\top$ και $\mathbf{u}^{(3)} = (0.7735, 1, -0.3204)^\top$. Έστω, το τυχαίο διάνυσμα $\mathbf{x} = (1, 2, 1)^\top$, οπότε, για παράδειγμα, είναι:

$$\mathbf{A}^5 = \begin{pmatrix} 68 & 0 & 164 \\ 136 & 32 & 428 \\ 164 & 0 & 396 \end{pmatrix}$$

$$\text{και } \mathbf{x}^{(5)} = \mathbf{A}^5 \mathbf{x} = (232, 628, 560)^\top.$$

Ανάλογα

$$\mathbf{A}^6 = \begin{pmatrix} 232 & 0 & 560 \\ 532 & 64 & 1484 \\ 560 & 0 & 1352 \end{pmatrix}$$

$$\text{και } \mathbf{x}^{(6)} = \mathbf{A}^6 \mathbf{x} = (792, 2144, 1912)^\top$$

Οπότε, με βάση τη σχέση (2.40) προσεγγιστική τιμή της απολύτως μεγαλύτερης ιδιοτιμής θα είναι

$$\lambda_1 \approx \frac{x^{(6)}}{x^{(5)}} = \frac{2144}{628} \approx 3.4140$$

Το ίδιο προσεγγιστικό αποτέλεσμα θα λαμβάναμε και αν χρησιμοποιούσαμε τους λόγους $792/232$ και $1912/560$. Αντίστοιχα, το ιδιοδιάνυσμα $\mathbf{u}^{(1)}$ θα είναι ίσο με το $\mathbf{x}^{(6)}$, που με κανονικοποίηση (διαιρούμε με το μεγαλύτερο στοιχείο του) γίνεται $(0.3694, 1, 0.8918)$. Δηλαδή, έχουμε υπολογίσει τη μεγαλύτερη ιδιοτιμή με ακρίβεια τριών δεκαδικών ψηφίων και το αντίστοιχο ιδιοδιάνυσμα με ακρίβεια τεσσάρων δεκαδικών ψηφίων.

2.7.2 Επιτάχυνση Aitken

Ο λόγος r_k^{-1} της σχέσης (2.40) αποτελεί μια προσέγγιση στην ακριβή τιμή λ_1 και το σφάλμα θα είναι $\epsilon_k = |r_k - \lambda_1|$. Αντίστοιχα, για το r_{k+1} θα ισχύει $\epsilon_{k+1} = |r_{k+1} - \lambda_1|$. Επειδή η σύγκλιση της μεθόδου είναι γραμμική, δηλαδή

$$\epsilon_{k+1} = A\epsilon_k \quad (2.41)$$

μπορώ να εφαρμόσω τη μέθοδο του Aitken για την επιτάχυνση της σύγκλισης. Δηλαδή, στην ουσία, τη σχέση του κεφαλαίου 1.5 που μας οδηγεί σε μια τιμή για το λ_1 , που είναι συνάρτηση των r_k, r_{k+1}, r_{k+2} . Η σχέση αυτή είναι:

$$\lambda_1 \approx \frac{r_k r_{k+2} - r_{k+1}^2}{r_{k+2} - 2r_{k+1} + r_k} \quad (2.42)$$

και βελτιώνει σημαντικά την ακρίβεια της ιδιοτιμής χωρίς να απαιτεί επιπλέον αριθμητικές πράξεις.

ΕΦΑΡΜΟΓΗ

Στο προηγούμενο παράδειγμα, είναι $r_5 = 3.414$, $r_4 = 3.413$ και $r_3 = 3.407$. Εφαρμόζοντας την επιτάχυνση Aitken υπολογίζουμε ότι $\lambda_1 = 3.4142$, που είναι ακριβέστερο κατά δυο δεκαδικά ψηφία από το r_5 .

2.7.3 Αντίστροφη μέθοδος των δυνάμεων

Η μέθοδος αυτή βασίζεται στη μέθοδο των δυνάμεων, που αναπτύχθηκε στο προηγούμενο κεφάλαιο, και μας δίνει τη δυνατότητα να υπολογίσουμε την *αποπλύτως μικρότερη ιδιοτιμή*. Αυτό είναι δυνατόν λόγω του παρακάτω θεωρήματος:

ΘΕΩΡΗΜΑ: Αν λ είναι μια ιδιοτιμή ενός πίνακα \mathbf{A} , τότε η λ^{-1} είναι μια ιδιοτιμή του \mathbf{A}^{-1} .

Απόδειξη: Έστω $\mathbf{Ax} = \lambda\mathbf{x}$. Τότε $\mathbf{x} = \mathbf{A}^{-1}(\lambda\mathbf{x}) = \lambda\mathbf{A}^{-1}\mathbf{x}$. Οπότε $\mathbf{A}^{-1}\mathbf{x} = \lambda^{-1}\mathbf{x}$, άρα, η τιμή λ^{-1} είναι ιδιοτιμή του πίνακα \mathbf{A}^{-1} .

Επομένως, εάν ένας πίνακας \mathbf{A} έχει N ιδιοτιμές $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_{N-1}| > |\lambda_N| > 0$, τότε, οι ιδιοτιμές του \mathbf{A}^{-1} είναι οι τιμές λ_i^{-1} , για τις οποίες θα ισχύει:

$$|\lambda_N^{-1}| > |\lambda_{N-1}^{-1}| \geq \dots \geq |\lambda_1^{-1}| > 0 \quad (2.43)$$

Επομένως, μπορούμε να υπολογίσουμε την ιδιοτιμή λ_N^{-1} εφαρμόζοντας τη μέθοδο των δυνάμεων στον \mathbf{A}^{-1} . Η λογική διαδικασία θα ήταν να εφαρμόσουμε

¹ Το r_k είναι διάνυσμα αλλά όλα τα στοιχεία του συγκλίνουν στην ίδια τιμή για το λ_1 , οπότε εδώ θα το χρησιμοποιήσουμε ως βαθμωτή ποσότητα

για τον \mathbf{A}^{-1} τη διαδικασία της μεθόδου δυνάμεων αλλά, όπως αναφέραμε στο κεφάλαιο 2.6, η αντιστροφή ενός πίνακα είναι αρκετά χρονοβόρα διαδικασία για τον H/Y . Έτσι, είναι προτιμότερο για τον υπολογισμό του $(\mathbf{A}^{-1})^{(k+1)} \mathbf{x}$ να λύσουμε με χρήση της μεθόδου Gauss το σύστημα $\mathbf{A} \cdot \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)}$, όπου με $\mathbf{x}^{(k)}$ συμβολίζουμε το αποτέλεσμα του πολλαπλασιασμού k -φορές του πίνακα \mathbf{A} με ένα διάνυσμα \mathbf{x} , δηλαδή, $\mathbf{x}^{(k)} = \mathbf{A}^k \mathbf{x}$. Πιο συγκεκριμένα, για ένα αρχικό διάνυσμα $\mathbf{x}^{(0)}$, θα πάρουμε $\mathbf{x}^{(1)} = \mathbf{A}^{-1} \mathbf{x}^{(0)}$. Αρα, $\mathbf{A} \mathbf{x}^{(1)} = \mathbf{x}^{(0)}$. Επομένως, η λύση του συστήματος αυτού καθορίζει την τιμή του $\mathbf{x}^{(1)}$. Με αυτή τη διαδικασία, για το διάνυσμα $\mathbf{x}^{(k+1)}$ θα ισχύει:

$$\mathbf{x}^{(k+1)} = \mathbf{A}^{-1} \mathbf{x}^{(k)} \Rightarrow \mathbf{A} \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} \quad (2.44)$$

ΠΑΡΑΔΕΙΓΜΑ

Έστω ο πίνακας $\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 2 & 2 \\ 1 & 0 & 3 \end{pmatrix}$ του προηγούμενου παραδείγματος και ένα αρχικό διάνυσμα $\mathbf{x}^{(0)} = (1, 2, 1)^T$, τότε θα πρέπει να λύσουμε το σύστημα $\mathbf{A} \mathbf{x}^{(1)} = \mathbf{x}^{(0)}$ με τη μέθοδο του Gauss. Εύκολα υπολογίζουμε ότι: $\mathbf{x}^{(1)} = (1, \frac{3}{2}, 0)^T$. Συνεχίζουμε λύνοντας το σύστημα $\mathbf{A} \mathbf{x}^{(2)} = \mathbf{x}^{(1)}$ με $\mathbf{x}^{(2)} = (\frac{3}{2}, 2, \frac{-1}{2})^T$ κ.ο.κ. Τελικά, μπορούμε να σχηματίσουμε το λόγο $r_6 = \frac{x^{(7)}}{x^{(6)}} \approx 1.70707$, οπότε απολύτως μικρότερη ιδιοτιμή του \mathbf{A} θα είναι $\lambda_3 = 1/1.70707 = 0.5858$.

2.7.4 Μέθοδος της μετάθεσης

Η μέθοδος αυτή χρησιμοποιείται συμπληρωματικά των δυο προηγούμενων μεθόδων των δυνάμεων, η εφαρμογή των οποίων μας οδήγησε στην εύρεση της απολύτως μικρότερης και μεγαλύτερης ιδιοτιμής. Με τη μέθοδο της μετάθεσης, θα υπολογίσουμε τις ενδιάμεσες ιδιοτιμές.

Η εφαρμογή της μεθόδου βασίζεται στην ιδιότητα των πινάκων που περιγράφεται από το παρακάτω θεώρημα:

ΘΕΩΡΗΜΑ: Αν οι N τιμές λ_i με $(i = 1, \dots, N)$ είναι ιδιοτιμές ενός $N \times N$ πίνακα \mathbf{A} , τότε δοθέντος ενός μιγαδικού αριθμού μ ο πίνακας $\mathbf{A} - \mu \mathbf{I}$ (\mathbf{I} : ο μοναδιαίος πίνακας) θα έχει ως ιδιοτιμές τις $(\lambda_i - \mu)$ για $(i = 1, \dots, N)$.

Επομένως, ως θεωρήσουμε στο διάστημα (λ_N, λ_1) ένα σημείο μ τέτοιο, ώστε για μια ιδιοτιμή λ_k να ισχύει $0 < |\lambda_k - \mu| < \epsilon$ και όλες οι άλλες ιδιοτιμές να ικανοποιούν την ανισότητα $|\lambda_i - \mu| > \epsilon$. Τότε, λόγω του παραπάνω θεωρήματος, η $|\lambda_k - \mu|$ θα είναι η απολύτως μικρότερη ιδιοτιμή του $\mathbf{A} - \mu \mathbf{I}$, επομένως, μπορώ να την υπολογίσω χρησιμοποιώντας την αντιστροφή μέθοδο των δυνάμεων. Υπενθυμίζουμε ότι συνίσταται ο υπολογισμός των $\mathbf{x}^{(k+1)}$ από τη λύση του συστήματος $(\mathbf{A} - \mu \mathbf{I}) \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)}$. Οπότε, τελικά, αν υπολογίσουμε το r_k , η ιδιοτιμή λ_i θα είναι:

$$\lambda_i = \frac{1}{r_k} + \mu \quad (2.45)$$

Αναλόγως, μπορούμε να υπολογίσουμε την ιδιοτιμή λ_i , για την οποία η απόσταση να είναι μέγιστη. Σ' αυτήν την περίπτωση, υπολογίζουμε με τη μέθοδο των δυνάμεων την τιμή $r_k = \lambda_j - \mu$, οπότε $\lambda_j = r_k + \mu$.

2.8 ΑΣΚΗΣΕΙΣ

1. Αν

$$\mathbf{A} = \begin{pmatrix} -0.002 & 4.000 & 4.000 \\ -2.000 & 2.906 & -5.387 \\ 3.000 & -4.031 & -3.112 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 7.998 \\ -4.481 \\ -4.413 \end{pmatrix}$$

να βρεθούν η $\det(\mathbf{A})$, ο \mathbf{A}^{-1} αφού πρώτα λυθεί το σύστημα $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$. Στη συνέχεια να υπολογισθούν οι ιδιοτιμές και τα ιδιοδιανύσματα του \mathbf{A} .

ΠΑΡΕΜΒΟΛΗ & ΠΡΟΒΛΕΨΗ

Υπάρχει μια ολόκληρη κατηγορία μαθηματικών προβλημάτων αλλά και προβλημάτων της καθημερινής ζωής στα οποία απαιτείται από συγκεκριμένα δεδομένα μιας παρατήρησης να προβλέψουμε την εξέλιξη μιας διαδικασίας (ΠΡΟΒΛΕΨΗ) ή να εκτιμήσουμε πια ήταν η κατάσταση ενός φαινομένου στο ενδιάμεσο διάστημα δύο μετρήσεων (ΠΑΡΕΜΒΟΛΗ). Για παράδειγμα αν ένας μετεωρολογικός σταθμός καταγράφει τη μεταβολή της θερμοκρασίας κάθε μια ώρα θα μπορούσαμε να υπολογίσουμε ποιά είναι η θερμοκρασία σε μια τυχαία χρονική στιγμή (π.χ. στις 12:15) ή να προβλέψουμε ποιά θα είναι η θερμοκρασία στις 20:00 αν έχουμε μετρήσεις μόνο ως τις 19:00.

Στα μαθηματικά πολλές φορές απαιτείται ο υπολογισμός μιας συνάρτησης (στην ουσία ενός κανόνα) που να ακολουθούν ζεύγη δοθέντων σημείων. Αλλά ακόμη και αν γνωρίζουμε τη μορφή μιας πολύπλοκης συνάρτησης προτιμούμε να την αντικαταστήσουμε με μια απλούστερη συνάρτηση την οποία στη συνέχεια μπορούμε ευκολότερα να χρησιμοποιήσουμε στους υπολογισμούς μας π.χ. για ολοκληρώσεις ή παραγωγίσεις κλπ. Φυσικά, θα πρέπει η νέα και απλή συνάρτηση να αποτελεί κατά το δυνατόν σωστή αναπαράσταση της ακριβούς, και γι' αυτό σε κάθε περίπτωση είναι σημαντικό να γνωρίσουμε το σφάλμα αυτής της αντικατάστασης. Η απλούστερη συνάρτηση είναι το πολυώνυμο και γι' αυτό όλες οι μέθοδοι που θα αναπτυχθούν σε αυτό το Κεφάλαιο θα έχουν ως σκοπό τον προσδιορισμό ενός πολυωνύμου το οποίο θα υπακούει σε συγκεκριμένους κανόνες που θα ορίσουμε.

ΜΕΘΟΔΟΙ

Οι μέθοδοι που θα αναπτύξουμε χαρακτηρίζονται από το είδος της αρχικής πληροφορίας που θα είναι διαθέσιμη. Έτσι, συνοπτικά αναφέρουμε τις παρακάτω μεθόδους που θα αναπτυχθούν στη συνέχεια διεξοδικά:

- **ΠΟΛΥΩΝΥΜΟ LAGRANGE:** δοθείσης μιας n -άδας ζευγών σημείων (x_i, y_i) , όπου τα $\Delta x_i = x_{i+1} - x_i$ δεν είναι κατ' ανάγκη ίσα, θα υπολογίσουμε ένα

πολυώνυμο $n - 1$ βαθμού, που θα διέρχεται από αυτά τα n σημεία. Το πολυώνυμο αυτό ονομάζεται **συμπτωτικό πολυώνυμο**.

- **ΣΥΜΠΤΩΤΙΚΟ ΠΟΛΥΩΝΥΜΟ ΓΙΑ ΙΣΑΠΕΧΟΝΤΑ** x_i : η προηγούμενη διαδικασία απλοποιείται σημαντικά αν τα n σημεία ισαπέχουν και τα συμπτωτικά πολυώνυμα υπολογίζονται με στοιχειώδεις πράξεις.
- **ΕΦΑΠΤΟΜΕΝΟ ΠΟΛΥΩΝΥΜΟ**: Στην ουσία εδώ υπολογίζουμε ένα πολυώνυμο από μια n -άδα τριάδων (x_i, y_i, y'_i) . Το πολυώνυμο αυτό είναι βαθμού $2n - 1$.
- **ΠΟΛΥΩΝΥΜΟ TAYLOR**: πρόκειται για το γνωστό πολυώνυμο Taylor με το οποίο αντικαθιστούμε μια συνάρτηση με ένα πολυώνυμο βαθμού n , απαιτώντας το πολυώνυμο και η συνάρτηση, όπως επίσης και οι παράγωγοί του μέχρι τάξης n να ταυτίζονται σε ένα δοθέν σημείο.
- **ΠΑΡΕΜΒΟΛΗ ΜΕ SPLINES**: μια εξαιρετική μέθοδος συνδυασμού των παραπάνω μεθόδων. Στην ουσία, για μια n -άδα ζευγών (x_i, y_i) δημιουργούμε $n - 1$ τριτοβάθμια πολυώνυμα (**κυβική spline**), ένα για κάθε ζεύγος σημείων $(x_i, y_i) - (x_{i+1}, y_{i+1})$ αντί ενός $(n - 1)$ -βάθμιο πολυώνυμο που διέρχεται από τα n σημεία.

3.1 ΠΟΛΥΩΝΥΜΟ LAGRANGE

Όπως προαναφέραμε, το **συμπτωτικό πολυώνυμο** ορίζεται ως ένα πολυώνυμο βαθμού n που διέρχεται από $n + 1$ σημεία.

Έστω ότι θέλουμε να υπολογίσουμε το συμπτωτικό πολυώνυμο 3ου βαθμού που διέρχεται από τα σημεία:

| | x_0 | x_1 | x_2 | x_3 | x_4 |
|--------|-------|-------|-------|-------|-------|
| x | 3.2 | 2.7 | 1.0 | 4.8 | 5.6 |
| $f(x)$ | 22.0 | 17.8 | 14.2 | 38.3 | 51.7 |
| | f_0 | f_1 | f_2 | f_3 | f_4 |

Πίνακας 3.1.

Προφανώς, πρέπει να επιλέξουμε τέσσερα από τα πέντε σημεία που δίνονται, έστω λοιπόν τα τέσσερα πρώτα. Αν το πολυώνυμο έχει τη μορφή $ax^3 + bx^2 + cx + d = P(x)$, τότε δημιουργούμε 4 εξισώσεις με 4 αγνώστους, αγνώστοι προφανώς είναι τα a, b, c και d . Λύνοντας το σύστημα με τις μεθόδους του προηγούμενου κεφαλαίου βρίσκουμε: $a = -0.5275$, $b = 6.4952$, $c = -16.117$ και $d = 24.3499$. Δηλαδή, το συμπτωτικό πολυώνυμο είναι το:

$$P(x) = -0.5275x^3 + 6.4952x^2 - 16.117x + 24.3499$$

οπότε, αν ζητηθεί, για παράδειγμα, η τιμή της $f(3)$ βρίσκουμε ότι $f(3) = 20.21$.

Το **πολυώνυμο Lagrange** δίνει απ' ευθείας το συμπτωτικό πολυώνυμο, χωρίς τη λύση συστήματος. Στο συγκεκριμένο πρόβλημα, θα είναι:

$$P_3(x) = \frac{(x-x_1)(x-x_2)(x-x_3)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)}f_0 + \frac{(x-x_0)(x-x_2)(x-x_3)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)}f_1 \\ + \frac{(x-x_0)(x-x_1)(x-x_3)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)}f_2 + \frac{(x-x_0)(x-x_1)(x-x_2)}{(x_3-x_0)(x_3-x_2)(x_3-x_1)}f_3$$

το οποίο εύκολα προγραμματίζεται και επόμενως για κάθε τιμή x_i θα βρίσκουμε αμέσως την τιμή της $f(x_i)$.

3.1.1 Απόδειξη του τύπου Lagrange

Η απόδειξη θα γίνει με χρήση της επαγωγικής μεθόδου, δηλαδή:

I. Το πολυώνυμο 0 (μηδενικού) βαθμού είναι η σταθερή συνάρτηση: $P_0(x) = a_0 = f_1$ για όλα τα x , δηλαδή μια ευθεία παράλληλη στον άξονα Ox που διέρχεται από το (x_1, f_1) .

II. Το πολυώνυμο 1ου βαθμού είναι: $P_1(x) = a_0 + a_1x$, για το οποίο ισχύει:

$$P_1(x_1) = a_0 + a_1x_1 = f_1$$

$$P_1(x_2) = a_0 + a_1x_2 = f_2$$

οπότε λύνοντας το σύστημα βρίσκουμε:

$$a_0 = \frac{x_2f_1 - x_1f_2}{x_2 - x_1} \quad \text{και} \quad a_1 = \frac{f_2 - f_1}{x_2 - x_1}$$

Αρα:

$$P_1(x) = \frac{x_2f_1 - x_1f_2}{x_2 - x_1} + \frac{f_2 - f_1}{x_2 - x_1} \cdot x = f_1 \frac{x - x_2}{x_1 - x_2} + f_2 \frac{x - x_1}{x_2 - x_1} \\ = f_1 L_1(x) + f_2 L_2(x)$$

όπου

$$L_1(x) = \frac{x - x_2}{x_1 - x_2} \quad \text{και} \quad L_2(x) = \frac{x - x_1}{x_2 - x_1}$$

Για τα $L_i(x)$, που ονομάζονται **συντελεστές Lagrange**, ισχύει προφανώς

$$L_1(x_1) = 1, \quad L_2(x_1) = 0 \\ L_1(x_2) = 0, \quad L_2(x_2) = 1$$

Στη γενική περίπτωση λόγω της παραπάνω θα είναι:

$$L_j(x_k) = \delta_{jk} = \begin{cases} 0 & \text{για } j \neq k \\ 1 & \text{για } j = k \end{cases}$$

όπου με δ_{jk} συμβολίζουμε το δέλτα του Kronecker. Επομένως ένα πολυώνυμο βαθμού n θα βρεθεί από τη σχέση:

$$P_n(x) = f_0 L_0(x) + f_1 L_1(x) + \dots + f_n L_n(x) = \sum_{i=0}^n f_i L_i(x) \quad (3.1)$$

όπου

$$L_j(x) = \frac{(x-x_1)(x-x_2)\dots(x-x_{j-1})(x-x_{j+1})\dots(x-x_n)}{(x_j-x_1)(x_j-x_2)\dots(x_j-x_{j-1})(x_j-x_{j+1})\dots(x_j-x_n)} \quad (3.2)$$

3.1.2 Σφάλμα του τύπου Lagrange

Το σφάλμα που γίνεται στον υπολογισμό μιας τιμής $f(x)$ για κάποιο δοθέν x μπορεί να εκτιμηθεί από τη σχέση:

$$E(x) = (x - x_0)(x - x_1) \dots (x - x_n) \frac{f^{(n+1)}(\xi)}{(n+1)!} \quad (3.3)$$

όπου n ο βαθμός του συμπτωτικού πολυωνύμου και το ξ είναι κατάλληλο σημείο μεταξύ των x_0, \dots, x_n .

Απόδειξη

Δημιουργώ μια νέα βοηθητική συνάρτηση

$$F(x) = f(x) - P(x) - C \cdot \pi(x)$$

όπου C είναι μια σταθερά και

$$\pi(x) = (x - x_0)(x - x_1)(x - x_2) \dots (x - x_n) \quad (3.4)$$

Επομένως, επειδή $f(x_i) = P(x_i)$ η $F(x)$ θα μηδενίζεται για κάθε τιμή του $x = x_i$. Αν επιλέξω την σταθερά C να είναι:

$$C = \frac{f(t) - P(t)}{\pi(t)}$$

όπου t είναι μια τυχαία τιμή τότε θα είναι και $F(t) = 0$ άρα η $F(x)$ έχει $n+2$ τουλάχιστον ρίζες. Σύμφωνα με το θεώρημα του Rolle η $F'(x)$ έχει $n+1$ ρίζες μεταξύ των ριζών της $F(x)$, η $F''(x)$ έχει n ρίζες μεταξύ των ριζών της $F'(x)$ κοκ. Συνεχίζοντας έτσι συμπεραίνουμε ότι η $F^{(n+1)}(x)$ έχει τουλάχιστον μία ρίζα σε ένα σημείο $x = \xi$. Επειδή όμως η $P^{(n+1)}(x) = 0$ θα έχουμε

$$F^{(n+1)}(\xi) \equiv 0 = f^{(n+1)}(\xi) - C \cdot (n+1)!$$

Άρα

$$C = \frac{f^{(n+1)}(\xi)}{(n+1)!}$$

και κατά συνέπεια:

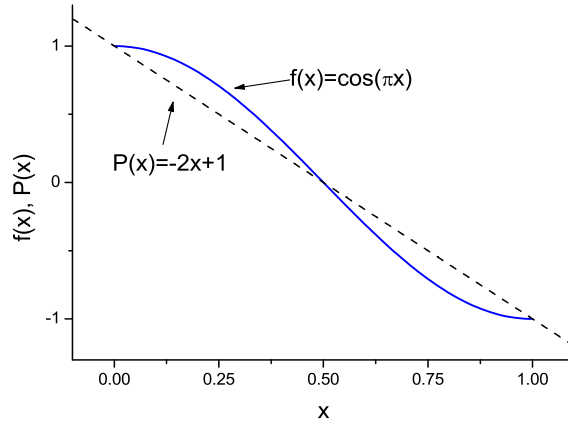
$$f(t) - P(t) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \pi(t)$$

Επειδή όμως το t ήταν ένα τυχαίο σημείο διάφορο των x_0, x_1, \dots, x_n η παραπάνω σχέση θα ισχύει για κάθε x . Άρα τελικά

$$f(x) - P(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \pi(x). \quad (3.5)$$

ΠΑΡΑΔΕΙΓΜΑ

Να προσεγγισθεί η συνάρτηση $y = \cos(\pi x)$ με ένα δευτεροβάθμιο συμπτωτικό πολυώνυμο που παίρνει τις ίδιες τιμές με την y για $x_0 = 0$, $x_1 = 0.5$, $x_2 = 1$. Να μελετηθεί το σφάλμα.



Σχήμα 3.1. Γραφική παράσταση της συνάρτησης $y = \cos(\pi x)$ και του συμπτωτικού πολυωνύμου $P(x) = -2x + 1$.

ΛΥΣΗ

Ο πίνακας τιμών είναι:

| | | | |
|-------|---|-----|----|
| x_i | 0 | 0.5 | 1 |
| f_i | 1 | 0.0 | -1 |

Οπότε οι συντελεστές Lagrange θα είναι:

$$\begin{aligned}
 L_1 &= \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)} = \frac{(x - 0.5)(x - 1)}{(0 - 0.5)(0 - 1)} = 2x^2 - 3x + 1, \\
 L_2 &= \frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)} = \frac{(x - 0)(x - 1)}{(0.5 - 0)(0.5 - 1)} = -4(x^2 - x) \\
 L_3 &= \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)} = \frac{(x - 0)(x - 0.5)}{(1 - 0)(1 - 0.5)} = 2x^2 - x
 \end{aligned} \tag{3.6}$$

επομένως

$$P(x) = 1 \cdot (2x^2 - 3x + 1) - 0.4 \cdot (x^2 - x) + (-1) \cdot (2x^2 - x) = -2x + 1$$

Το πολυώνυμο τελικά δεν είναι δευτεροβάθμιο αλλ'ήλ πρωτοβάθμιο, το Σχήμα 3.1 δείχνει τις δύο συναρτήσεις. Το σφάλμα θα είναι:

$$E(x) = x \cdot (x - 0.5) \cdot (x - 1) \frac{\pi^3 \sin(\pi \xi)}{3!}$$

και για $x = 0.25$ είναι: $E(0.25) \leq 0.24$.

3.2 ΣΥΜΠΤΩΤΙΚΟ ΠΟΛΥΩΝΥΜΟ ΓΙΑ ΙΣΑΠΕΧΟΝΤΑ x_i

Όταν οι τιμές των x_i ισαπέχουν, τότε η εύρεση του συμπτωτικού πολυωνύμου απλοποιείται, αν χρησιμοποιήσουμε κατάλληλα τους **τελεστές διαφοράς** και τους **πίνακες διαφορών**.

Ορίζουμε **ως τελεστή διαφοράς προς τα εμπρός** τον τελεστή Δ , με την εξής ιδιότητα $\Delta f_i = f_{i+1} - f_i$. Αναλόγως μπορώ να ορίσω ως **συντελεστή διαφοράς προς τα πίσω** τον τελεστή $\nabla f_i = f_i - f_{i-1}$

Μπορούν να οριστούν ως **διαφορές 1ης τάξης** τα

$$\Delta f_0 = f_1 - f_0, \Delta f_1 = f_2 - f_1, \dots$$

άρα

$$\Delta f_i = f_{i+1} - f_i \quad (3.7)$$

Ως **διαφορές 2ης τάξης** τα

$$\Delta^2 f_1 = \Delta(\Delta f_1) = \Delta(f_2 - f_1) = \Delta f_2 - \Delta f_1 = (f_3 - f_2) - (f_2 - f_1) = f_3 - 2f_2 + f_1$$

άρα

$$\Delta^2 f_i = f_{i+2} - 2f_{i+1} + f_i \quad (3.8)$$

Ως **διαφορές 3ης τάξης** τα

$$\Delta^3 f_1 = \Delta^2(\Delta f_1) = f_4 - 3f_3 + 3f_2 - f_1$$

$$\Delta^3 f_i = f_{i+3} - 3f_{i+2} + 3f_{i+1} - f_i \quad (3.9)$$

και τέλος ως **διαφορές n-οστής τάξης**

$$\Delta^n f_i = f_{i+n} - n f_{i+n-1} + \frac{n(n-1)}{2!} f_{i+n-2} - \frac{n(n-1)(n-2)}{3!} f_{i+n-3} + \dots \quad (3.10)$$

Με βάση τα παραπάνω, είναι εύκολο να δημιουργήσουμε **πίνακες διαφορών** για κάθε n -άδα σημείων (x_i, y_i) . Οι πίνακες διαφορών διευκολύνουν σημαντικά την εύρεση του συμπτωτικού πολυωνύμου. Για παράδειγμα, παραθέτουμε τον πίνακα 3.2.

Θα αποδείξουμε τώρα ότι το συμπτωτικό πολυώνυμο n βαθμού θα δίνεται από τη σχέση:

$$\begin{aligned} P_n(x_s) &= f_0 + s \cdot \Delta f_0 + \frac{s(s-1)}{2!} \cdot \Delta^2 f_0 + \frac{s(s-1)(s-2)}{3!} \cdot \Delta^3 f_0 + \dots \\ &= f_0 + \binom{s}{1} \cdot \Delta f_0 + \binom{s}{2} \cdot \Delta^2 f_0 + \binom{s}{3} \cdot \Delta^3 f_0 + \dots \\ &= \sum_{i=0}^n \binom{s}{i} \cdot \Delta^i f_0 \end{aligned} \quad (3.11)$$

όπου $x_s = x_0 + s \cdot h$. Προφανώς ισχύει ότι:

| x | $f(x)$ | $\Delta f(x)$ | $\Delta^2 f(x)$ | $\Delta^3 f(x)$ | $\Delta^4 f(x)$ | $\Delta^5 f(x)$ |
|-----|--------|---------------|-----------------|-----------------|-----------------|-----------------|
| 0.0 | 0.000 | | | | | |
| | | 0.203 | | | | |
| 0.2 | 0.203 | | 0.017 | | | |
| | | 0.220 | | 0.024 | | |
| 0.4 | 0.423 | | 0.041 | | 0.020 | |
| | | 0.261 | | 0.044 | | 0.032 |
| 0.6 | 0.684 | | 0.085 | | 0.052 | |
| | | 0.246 | | 0.096 | | |
| 0.8 | 1.030 | | 0.181 | | | |
| | | 0.527 | | | | |
| 1.0 | 1.557 | | | | | |

Πίνακας 3.2.

$$s = 0 \rightarrow P_n(x_0) = f_0$$

$$s = 1 \rightarrow P_n(x_1) = f_0 + \Delta f_0 = f_1$$

$$s = 2 \rightarrow P_n(x_2) = f_0 + 2\Delta f_0 + \Delta^2 f_0 = f_2$$

ΑΠΟΔΕΙΞΗ Για $s = 0$ και $s = 1$ ο τύπος ισχύει. Έστω ότι ισχύει και για $s = m$, δηλαδή:

$$P(x_m) = \sum_{i=0}^n \binom{m}{i} \cdot \Delta^i f_0$$

Θα δείξουμε ότι ισχύει για $s = m + 1$, δηλαδή

$$\begin{aligned}
 P(x_{m+1}) &= P(x_m) + \Delta P(x_m) \\
 &= \sum_{i=0}^n \binom{m}{i} \cdot \Delta^i f_0 + \sum_{i=0}^n \binom{m}{i} \cdot \Delta^{i+1} f_0 \\
 &= f_0 + \sum_{i=1}^n \binom{m}{i} \cdot \Delta^i f_0 + \sum_{i=1}^n \binom{m}{i-1} \cdot \Delta^i f_0 + \Delta^{n+1} f_0 \\
 &= f_0 + \sum_{i=1}^n \binom{m+1}{i} \cdot \Delta^i f_0 + \Delta^{n+1} f_0 \\
 &= \sum_{i=0}^{n+1} \binom{m+1}{i} \cdot \Delta^i f_0
 \end{aligned} \tag{3.12}$$

οεδ

Στη παραπάνω απόδειξη χρησιμοποιήθηκε η σχέση

$$\binom{k+1}{n} = \binom{k}{n} + \binom{k}{n-1}.$$

ΠΑΡΑΤΗΡΗΣΗ: Γενικά το s μπορεί να πάρει και μη ακέραιες τιμές, δηλαδή για μία τυχαία τιμή x θα είναι:

$$s = \frac{x - x_0}{h}. \quad (3.13)$$

3.2.1 Τύποι για ισαπέχοντα x_i

- **Newton - Gregory προς τα μπρός** (από το σημείο $x_0 \rightarrow x_n$)

$$P_n(x) = f_0 + \binom{s}{1} \cdot \Delta f_0 + \binom{s}{2} \cdot \Delta^2 f_0 + \dots + \binom{s}{n} \cdot \Delta^n f_0 \quad (3.14)$$

- **Newton - Gregory προς τα πίσω** (από το σημείο $x_{-n} \rightarrow x_0$)

$$P_n(x) = f_0 + \binom{s}{1} \cdot \Delta f_{-1} + \binom{s+1}{2} \cdot \Delta^2 f_{-2} + \binom{s+2}{3} \cdot \Delta^3 f_{-3} + \dots + \binom{s+n-1}{n} \cdot \Delta^n f_{-n} \quad (3.15)$$

ΠΑΡΑΔΕΙΓΜΑ

Δίνεται ο Πίνακας Διαφορών 3.3, να υπολογισθούν τα συμπτωτικά πολυώνυμα Newton-Gregory.

| x | $F(x)$ | $\Delta f(x)$ | $\Delta^2 f(x)$ | $\Delta^3 f(x)$ | $\Delta^4 f(x)$ | $\Delta^5 f(x)$ |
|-----|---------|---------------|-----------------|-----------------|-----------------|-----------------|
| 0.2 | 1.06894 | | | | | |
| 0.5 | 1.18136 | 0.11242 | | | | |
| 0.8 | 1.30561 | 0.12425 | 0.01183 | 0.00123 | | |
| 1.1 | 1.44292 | 0.13731 | 0.01306 | 0.00138 | 0.00015 | |
| 1.4 | 1.59467 | 0.15175 | 0.01444 | 0.0152 | 0.00014 | 0.000001 |
| 1.7 | 1.76238 | 0.16771 | 0.01596 | | | |

Πίνακας 3.3.

Newton - Gregory προς τα εμπρός για τα σημεία από $x = 0.5 \rightarrow x = 1.4$, δηλαδή $x_0 = 0.5$

$$\begin{aligned} P_3(x) &= 1.18136 + 0.12425 \cdot s + 0.01306 \cdot \binom{s}{2} + 0.00138 \cdot \binom{s}{3} \\ &= 1.18136 + 0.12425 \cdot s + 0.01306 \cdot s \cdot (s-1)/2 + 0.00138 \cdot s \cdot (s-1) \cdot (s-2)/6 \\ &= 0.9996 + 0.3354 \cdot x + 0.052 \cdot x^2 + 0.085 \cdot x^3 \end{aligned}$$

Newton – Gregory προς τα πίσω για τα σημεία από $x = 1.1 \rightarrow x = 0.2$, δηλαδή $x_0 = 1.1$

$$\begin{aligned} P_3(x) &= 1.44292 + 0.13731 \cdot s + 0.01306 \cdot \binom{s+1}{2} + 0.00123 \cdot \binom{s+2}{3} \\ &= 1.44292 + 0.13731 \cdot s + 0.01306 \cdot s \cdot (s+1)/2 + 0.00123 \cdot s \cdot (s+1) \cdot (s+2)/6 \\ &= 0.99996 + 0.33374 \cdot x + 0.05433 \cdot x^2 + 0.007593 \cdot x^3 \end{aligned}$$

3.3 ΕΦΑΠΤΟΜΕΝΟ ΠΟΛΥΩΝΥΜΟ (HERMITE)

Είναι δυνατόν ένα συμπτωτικό πολυώνυμο να ‘ μοιάζει ’ κάπως παραπάνω σε μια δοθείσα συνάρτηση ή ένα σύνολο τιμών, εφόσον απαιτηθεί και οι παράγωγοι του πολυωνύμου να παίρνουν τις ίδιες τιμές με τη συνάρτηση ή το σύνολο τιμών στα σημεία επαφής.

Επομένως, αν απαιτήσουμε το πολυώνυμο με τη συνάρτηση ή το σύνολο τιμών να παίρνουν τις ίδιες τιμές σε n σημεία, όπως και οι παράγωγοί του

$$\begin{aligned} P(x) &= y(x) \\ P'(x) &= y'(x) \end{aligned}$$

για x_0, x_1, \dots, x_n καταλήγουμε σε $2n$ εξισώσεις, οι οποίες καθορίζουν ένα πολυώνυμο το πολύ $2n - 1$ βαθμού. Το πολυώνυμο αυτό δίνεται από τον **τύπο του Hermite**.

$$P_{2n-1}(x) = \sum_{i=1}^n A_i(x) y_i + \sum_{i=1}^n B_i(x) y'_i \quad (3.16)$$

όπου

$$A_i(x) = [1 - 2(x - x_i)L'_i(x_i)] \cdot [L_i(x)]^2 \quad (3.17)$$

$$B_i(x) = (x - x_i) \cdot [L_i(x)]^2 \quad (3.18)$$

όπου $L_i(x)$ είναι οι συντελεστές Lagrange.

ΣΦΑΛΜΑ

Το σφάλμα προφανώς θα είναι ισοδύναμο με αυτό ενός συμπτωτικού πολυωνύμου βαθμού $2n - 1$

$$y(x) - p(x) = \frac{y^{(2n+1)}(\xi)}{(2n+1)!} [(x - x_1)(x - x_2) \dots (x - x_n)]^2 \quad (3.19)$$

| k | x_k | y_k | y'_k |
|-----|-------|-------|--------|
| 0 | 0 | 0 | 0 |
| 1 | 4 | 2 | 0 |

ΠΑΡΑΔΕΙΓΜΑ

Να υπολογισθεί το εφαπτόμενο πολυώνυμο για τα εξής δεδομένα
Οι συντελεστές *Lagrange* θα είναι:

$$L_0(x) = \frac{x - x_1}{x_0 - x_1} = \frac{x - 4}{0 - 4} = -\frac{x - 4}{4}$$

$$L_1(x) = \frac{x - x_0}{x_1 - x_0} = \frac{x}{4}$$

$$L'_0(x) = \frac{1}{x_0 - x_1} = -\frac{1}{4}$$

$$L'_1(x) = \frac{1}{x_1 - x_0} = \frac{1}{4}$$

Άρα:

$$A_0(x) = [1 - 2 \cdot L'_0(x - x_0)] \cdot L_0^2 = \left[1 - 2 \cdot \left(-\frac{1}{4}\right)(x - 0)\right] \cdot \left(\frac{x - 4}{4}\right)^2$$

$$A_1(x) = [1 - 2 \cdot L'_0(x - x_1)] \cdot L_1^2 = \left[1 - 2 \cdot \frac{1}{4}(x - 4)\right] \cdot \left(\frac{x}{4}\right)^2 = \left(3 - \frac{x}{2}\right) \cdot \left(\frac{x}{4}\right)^2$$

$$B_0(x) = (x - 0) \cdot \left(\frac{x - 4}{4}\right)^2 = x \left(\frac{x - 4}{4}\right)^2 \quad B_1(x) = (x - 4) \cdot \left(\frac{x}{4}\right)^2$$

Επομένως το πολυώνυμο *Hermite* έχει τη μορφή:

$$P(x) = (6 - x) \frac{x^2}{16}.$$

3.4 ΠΟΛΥΩΝΥΜΟ TAYLOR

Μέχρι τώρα προσπαθούσαμε να υπολογίσουμε πολυώνυμα $P(x)$ τα οποία ταυτιζόταν σε ένα αριθμό σημείων με μια συνάρτηση $f(x)$ (συμπτωτικό πολυώνυμο) ή επιπλέον είχαν και 'επαφή', δηλαδή ταυτιζόταν και η πρώτη παράγωγος της συνάρτησης με τις τιμές του πολυωνύμου στα κοινά σημεία (εφαπτόμενο πολυώνυμο). Το πολυώνυμο Taylor έχει την ιδιότητα να ταυτίζεται σε ένα μόνο σημείο x_0 με τη δοθείσα συνάρτηση, αλλά, επιπλέον, εάν είναι βαθμού n , τότε και οι πρώτοι παράγωγοί τους μέχρι τάξης n θα ταυτίζονται σε αυτό το σημείο, δηλαδή:

$$P^{(i)}(x_0) = f^{(i)}(x_0) \quad \text{για } i = 0, 1, \dots, n$$

Το πολυώνυμο Taylor μπορεί να γραφεί:

$$P(x) = \sum_{i=0}^n \frac{f^{(i)}(x_0)}{i!} (x - x_0)^i \quad (3.20)$$

Το σφάλμα του 'μοιάζει' με το σφάλμα του συμπτωτικού πολυωνύμου και δίνεται (όπως είναι γνωστό από το διαφορικό λογισμό) από τη σχέση:

$$E_N(x) = \frac{x^{N+1} f^{(N+1)}(\xi)}{(N+1)!} \quad (3.21)$$

ΠΑΡΑΔΕΙΓΜΑ

Να βρεθεί ένα πολυώνυμο που μαζί με τις πρώτες n παραγώγους του παίρνει στο $x_0 = 0$ τις τιμές που παίρνουν οι αντίστοιχες ποσότητες της συνάρτησης e^x .

ΛΥΣΗ

Όλες οι παράγωγοι της e^x είναι προφανώς e^x , οπότε

$$y_0 = y_0^{(1)} = y_0^{(2)} = \dots = y_0^{(n)} = 1$$

και επομένως:

$$p(x) = \sum_{i=1}^n \frac{1}{i!} x^i = 1 + x + \frac{x^2}{2} + \dots + \frac{1}{n!} x^n$$

Το σφάλμα θα είναι:

$$E = x^{n+1} \frac{e^\xi}{(n+1)!}$$

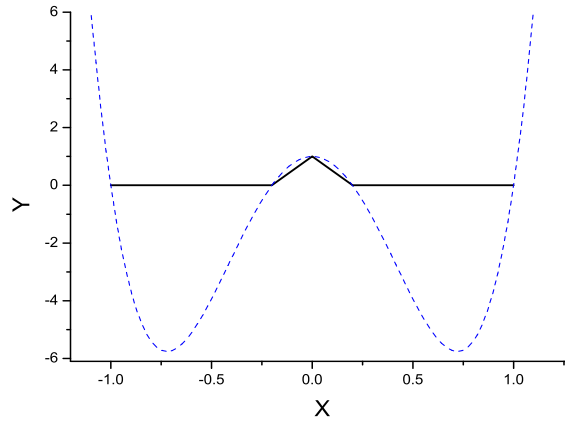
Μπορούμε στη συνέχεια να επεκτείνουμε το ερώτημα ζητώντας να υπολογίσουμε πόσους όρους πρέπει να κρατήσουμε στο ανάπτυγμα Taylor ώστε να επιτύχουμε ακρίβεια 6 σημαντικών ψηφίων ή καλύτερα σφάλμα μικρότερο του 5×10^{-7} για $x \in [0, 1]$. Οπότε, επειδή η μεγαλύτερη δυνατή τιμή του σφάλματος είναι για $x = 1$ θα έχουμε:

$$E = 0.00000005 \leq \frac{e}{n+1}$$

οπότε η ανισότητα θα ικανοποιείται για $n \geq 10$. Δηλαδή, θα πρέπει να χρησιμοποιήσω τουλάχιστον 10 όρους στο ανάπτυγμα για να επιτύχω τη ζητούμενη ακρίβεια.

3.5 ΠΑΡΕΜΒΟΛΗ ΜΕ SPLINES

Πολλές φορές το συμπτωτικό πολυώνυμο δεν αποτελεί την πλέον ιδανική αναπαράσταση της συνάρτησης που αντιπροσωπεύει ένα σύνολο σημείων, ειδικότερα όταν υπάρχουν τοπικές 'ασυνέχειες' στα δεδομένα.



Σχήμα 3.2. Γραφική παράσταση της συνάρτησης (3.22) και του συμπτωτικού πολυωνύμου 4ου βαθμού (3.23).

ΠΑΡΑΔΕΙΓΜΑ

Ας θεωρήσουμε τη συνάρτηση:

$$f(x) = \begin{cases} 0 & -1 \leq x \leq -0.2 \\ 1 - 5|x| & -0.2 < x < 0.2 \\ 0 & 0.2 \leq x \leq 1.0 \end{cases} \quad (3.22)$$

Αν υποθέσουμε ότι βρίσκουμε στο διάστημα $(-1, 1)$ το συμπτωτικό πολυώνυμο τέταρτου βαθμού

$$P(x) = 1 - 26x^2 + 25x^4 \quad (3.23)$$

τότε εύκολα παρατηρούμε ότι δεν αποτελεί σωστή αναπαράσταση της συνάρτησης (3.22). Το αυτό θα συνέβαινε αν είχαμε χρησιμοποιήσει και πολυώνυμο μεγαλύτερου βαθμού.

Γίνεται προφανής η ανάγκη αλγορίθμου που να παίρνει υπόψη του τις τοπικές ‘ασυνέχειες’. Για να το επιτύχουμε αυτό, εμφυτεύουμε ένα κατάλληλο πολυώνυμο σε κάθε ζεύγος σημείων.

Στη συνέχεια, θα προσπαθήσουμε να εμφυτεύσουμε ένα πολυώνυμο 3ου βαθμού για κάθε ζεύγος σημείων (x_i, y_i) και (x_{i+1}, y_{i+1}) , αυτό απαιτεί πληροφορία από 4 σημεία, επομένως για να το επιτύχουμε, απαιτείται, για παράδειγμα, η κλίση και η καμπυλότητα των πολυωνύμων δεξιά και αριστερά του σημείου (x_i, y_i) να ταυτίζονται.

Έστω το τριτοβάθμιο πολυώνυμο που διέρχεται από τα σημεία (x_i, y_i) και (x_{i+1}, y_{i+1})

$$y(x) = a_i \cdot (x - x_i)^3 + b_i \cdot (x - x_i)^2 + c_i \cdot (x - x_i) + d_i$$

Επειδή διέρχεται από τα δυο άκρα του διαστήματος, θα είναι:

$$y_i = a_i \cdot (x_i - x_i)^3 + b_i \cdot (x_i - x_i)^2 + c_i \cdot (x_i - x_i) + d_i = d_i$$

και

$$\begin{aligned} y_{i+1} &= a_i \cdot (x_{i+1} - x_i)^3 + b_i \cdot (x_{i+1} - x_i)^2 + c_i \cdot (x_{i+1} - x_i) + d_i \\ &= a_i h_i^3 + b_i h_i^2 + c_i h_i + d_i \end{aligned}$$

όπου $h_i = x_{i+1} - x_i$.

Χρειαζόμαστε και την πρώτη και δεύτερη παράγωγο, για να συσχετίσουμε τις κλίσεις και καμπυλότητες των πολυωνύμων, δηλαδή:

$$\begin{aligned} y'(x) &= 3a_i \cdot (x - x_i)^2 + 2b_i \cdot (x - x_i) + c_i \\ y''(x) &= 6a_i \cdot (x - x_i) + 2b_i \end{aligned}$$

Έστω S_i η δεύτερη παράγωγος στο x_i και S_{i+1} στο x_{i+1} . Τότε:

$$\begin{aligned} S_i &= 6a_i \cdot (x_i - x_i) + 2b_i = 2b_i \\ S_{i+1} &= 6a_i \cdot (x_{i+1} - x_i) + 2b_i = 6a_i h_i + 2b_i \end{aligned}$$

Άρα

$$b_i = \frac{S_i}{2}, \quad a_i = \frac{S_{i+1} - S_i}{6h_i} \quad (3.24)$$

οπότε

$$y_{i+1} = \frac{S_{i+1} - S_i}{6h_i} h_i^3 + \frac{S_i}{2} h_i^2 + c_i h_i + y_i$$

επομένως

$$c_i = \frac{y_{i+1} - y_i}{h_i} - \frac{2h_i S_i + h_i S_{i+1}}{6} \quad (3.25)$$

Αν απαιτήσουμε τώρα οι κλίσεις από δεξιά και αριστερά του σημείου x_i να είναι ίσες, τότε επειδή

$$\begin{aligned} y'_i &= 3a_i \cdot (x_i - x_i)^2 + 2b_i \cdot (x_i - x_i) + c_i = c_i \\ y'_i &= 3a_{i-1} \cdot (x_i - x_{i-1})^2 + 2b_{i-1} \cdot (x_i - x_{i-1}) + c_{i-1} \\ &= 3a_{i-1} h_{i-1}^2 + 2b_{i-1} h_{i-1} + c_{i-1} \end{aligned}$$

καταλήγουμε στη σχέση:

$$\begin{aligned} y'_i &= \frac{y_{i+1} - y_i}{h_i} - \frac{2h_i S_i + h_i S_{i+1}}{6} \\ &= 3 \left(\frac{S_i - S_{i-1}}{6h_{i-1}} \right) h_{i-1}^2 + 2 \frac{S_{i-1}}{2} h_{i-1} + \frac{y_i - y_{i-1}}{h_{i-1}} - \frac{2h_{i-1} S_{i-1} + h_{i-1} S_i}{6} \end{aligned}$$

και απλοποιώντας βρίσκουμε

$$h_{i-1}S_{i-1} + 2(h_{i-1} + h_i)S_i + h_iS_{i+1} = 6 \left(\frac{y_{i+1} - y_i}{h_i} - \frac{y_i - y_{i-1}}{h_{i-1}} \right) \quad (3.26)$$

Αν έχουμε $n+1$ σημεία, η παραπάνω σχέση μπορεί να εφαρμοσθεί στα $n-1$ εσωτερικά σημεία του διαστήματος (x_0, x_n) . Άρα, θα έχουμε ένα σύστημα $n-1$ εξισώσεων για τους $n+1$ αγνώστους S_i . Οπότε πρέπει να ορίσω δύο 'αυθαίρετες' συνθήκες για τα S_0 και S_n στα άκρα του συστήματος.

ΕΠΙΛΟΓΗ ΣΥΝΘΗΚΩΝ ΣΤΑ ΑΚΡΑ ΤΟΥ ΔΙΑΣΤΗΜΑΤΟΣ

Όπως προαναφέραμε υπάρχει μια ελευθερία στον υπολογισμό των συνθηκών στα άκρα του διαστήματος. Υπάρχουν όμως μερικές επιλογές που είναι αρκετά αξιόπιστες και χρησιμοποιούνται ευρύτατα.

I. $S_0 = 0$ και $S_n = 0$

Αυτή η επιλογή είναι ισοδύναμη με την υπόθεση ότι στα άκρα του διαστήματος η κυβική σπλινε που μελετάμε να είναι γραμμική (στη βιβλιογραφία αναφέρεται ως **φυσική spline**).

II. $S_0 = S_1$ και $S_n = S_{n-1}$

Η επιλογή αυτή είναι ισοδύναμη με την υπόθεση ότι στα άκρα έχουμε παραβολική προσέγγιση.

III. Προβλέποντας την τιμή του S_0 από τα S_1, S_2 και του S_n από τα S_{n-1}, S_{n-2} , δηλαδή υποθέτουμε ότι η κλίση της S_i είναι ίδια μεταξύ των σημείων x_0 και x_1

$$\frac{S_1 - S_0}{h_0} = \frac{S_2 - S_1}{h_1} \Rightarrow S_0 = \frac{(h_0 + h_1)S_1 - h_0S_2}{h_1}$$

και μεταξύ των σημείων x_{n-1} και x_n

$$\frac{S_n - S_{n-1}}{h_{n-1}} = \frac{S_{n-1} - S_{n-2}}{h_{n-2}} \Rightarrow S_n = \frac{(h_{n-2} + h_{n-1})S_{n-1} - h_{n-1}S_{n-2}}{h_{n-2}}$$

IV. Εξαναγκασμός των κλίσεων στα άκρα να πάρουν συγκεκριμένες τιμές. Εστω $f'(x_0) = A$ και $f'(x_n) = B$

$$\begin{aligned} 2h_0S_0 + h_1S_1 &= 6 \left(\frac{y_1 - y_0}{h_0} - A \right) \\ h_{n-1}S_{n-1} + 2h_nS_n &= 6 \left(B - \frac{y_n - y_{n-1}}{h_{n-1}} \right) \end{aligned}$$

Το σύστημα των $n+1$ εξισώσεων γράφεται γενικά ως :

$$= 6 \begin{pmatrix} \dots \\ \frac{y_2 - y_1}{h_1} - \frac{y_1 - y_0}{h_0} \\ \frac{y_3 - y_2}{h_2} - \frac{y_2 - y_1}{h_1} \\ \dots \\ \frac{y_n - y_{n-1}}{h_{n-1}} - \frac{y_{n-1} - y_{n-2}}{h_{n-2}} \\ \dots \end{pmatrix} = \mathbf{Y}$$

επιλογή I: $S_0 = 0$ και $S_n = 0$, δημιουργούμε τον παρακάτω σύστημα:

$$\begin{pmatrix} 2(h_0 + h_1) & h_1 & & & \\ h_1 & 2(h_1 + h_2) & h_2 & & \\ & h_2 & 2(h_2 + h_3) & h_3 & \\ & & \dots & & \\ & & h_{n-2} & 2(h_{n-2} + h_{n-1}) & \end{pmatrix} \cdot \begin{pmatrix} S_1 \\ S_2 \\ S_3 \\ \vdots \\ S_{n-1} \end{pmatrix} = \mathbf{Y}$$

επιλογή II: $S_0 = S_1$ και $S_n = S_{n-1}$, δημιουργούμε τον παρακάτω πίνακα συντελεστών:

$$\begin{pmatrix} 3h_0 + 2h_1 & h_1 & & & \\ h_1 & 2(h_1 + h_2) & h_2 & & \\ & h_2 & 2(h_2 + h_3) & h_3 & \\ & & \dots & \dots & \\ & & & h_{n-2} & 2h_{n-2} + 3h_{n-1} \end{pmatrix} \cdot \begin{pmatrix} S_1 \\ S_2 \\ S_3 \\ \vdots \\ S_{n-1} \end{pmatrix} = \mathbf{Y}$$

επιλογή III : τα S_0 και S_n υπολογίζονται με γραμμική πρόβλεψη οπότε δημιουργούμε τον παρακάτω πίνακα συντελεστών:

$$\begin{pmatrix} \frac{(h_0+h_1)(h_0+2h_1)}{h_1} & \frac{h_1^2-h_0^2}{h_1} & & & \\ h_1 & 2(h_1+h_2) & h_2 & & \\ & h_2 & 2(h_2+h_3) & h_3 & \\ & & \dots & \dots & \\ & & \frac{h_{n-2}^2-h_{n-1}^2}{h_{n-2}} & \frac{(h_{n-1}+h_{n-2})(h_{n-1}+2h_{n-2})}{h_{n-2}} & \end{pmatrix} \cdot \begin{pmatrix} S_1 \\ S_2 \\ S_3 \\ \vdots \\ S_{n-1} \end{pmatrix} = \mathbf{Y}$$

τα S_0 και S_n υπολογίζονται μετά τη λύση του συστήματος.

επιλογή IV : όπου $f'(x_0) = A$, $f'(x_n) = B$

$$\begin{pmatrix} 2h_0 & h_1 & & & \\ h_0 & 2(h_0 + h_1) & h_1 & & \\ & h_1 & 2(h_1 + h_2) & h_2 & \\ & & \dots & & \\ & & & h_{n-2} & 2h_{n-1} \end{pmatrix} \cdot \begin{pmatrix} S_1 \\ S_2 \\ S_3 \\ \vdots \\ S_{n-1} \end{pmatrix} = \mathbf{Y}$$

Μετά τη λύση του συστήματος υπολογίζουμε τους συντελεστές a_i , b_i , c_i και d_i για τα τριτοβάθμια πολυώνυμα σε κάθε διάστημα από τις σχέσεις:

$$a_i = \frac{S_{i+1} - S_i}{6h_i} \quad (3.27)$$

$$b_i = \frac{S_i}{2} \quad (3.28)$$

$$c_i = \frac{y_{i+1} - y_i}{h_i} - \frac{2h_i S_i + h_i S_{i+1}}{6} \quad (3.29)$$

$$d_i = y_i \quad (3.30)$$

ΠΑΡΑΤΗΡΗΣΕΙΣ

Η επιλογή **I** πλαταίνει την καμπύλη στα άκρα.

Η επιλογή **III** δημιουργεί μεγάλη καμπυλότητα στα άκρα.

Η επιλογή **IV** είναι πιθανότατα καλύτερη, αν υπολογίσουμε σωστά τις κλίσεις στα άκρα.

ΠΑΡΑΔΕΙΓΜΑ

Να προσαρμοσθεί μια κυβική *spline* στα δεδομένα (τα δεδομένα προέρχονται από

| | | | | | |
|-----|----|----|---|----|----|
| x | 0 | 1 | 2 | 3 | 4 |
| y | -8 | -7 | 0 | 19 | 56 |

τη συνάρτηση $y = x^3 - 8$.)

ΛΥΣΗ

Θα εξετάσουμε το πρόβλημα και για τις τέσσερις συνθήκες για τα άκρα.

ΣΥΝΘΗΚΗ I: Δημιουργούμε το σύστημα

$$\begin{pmatrix} 4 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 4 \end{pmatrix} \cdot \begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} 36 \\ 72 \\ 108 \end{pmatrix} \Rightarrow \begin{matrix} S_0 = 0 \\ S_1 = 6.4285 \\ S_2 = 10.2857 \\ S_3 = 24.4285 \\ S_4 = 0 \end{matrix}$$

ΣΥΝΘΗΚΗ II: Δημιουργούμε το σύστημα

$$\begin{pmatrix} s & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & s \end{pmatrix} \cdot \begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} 36 \\ 72 \\ 108 \end{pmatrix} \Rightarrow \begin{aligned} S_1 &= S_0 = 4.8 \\ S_2 &= 1.2 \\ S_3 &= 19.2 = S_4 \end{aligned}$$

ΣΥΝΘΗΚΗ III: Δημιουργούμε το σύστημα

$$\begin{pmatrix} 6 & 0 & 0 \\ 1 & 4 & 1 \\ 0 & 0 & 6 \end{pmatrix} \cdot \begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} 36 \\ 72 \\ 108 \end{pmatrix} \Rightarrow \begin{aligned} S_0 &= 0 \quad S_1 = 6 \\ S_2 &= 12 \quad S_3 = 18 \\ S_4 &= 24 \end{aligned}$$

ΣΥΝΘΗΚΗ IV: Δημιουργούμε το σύστημα

$$\begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 1 & 4 & 1 & 0 & 0 \\ 0 & 1 & 4 & 1 & 0 \\ 0 & 0 & 1 & 4 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \\ S_4 \end{pmatrix} = \begin{pmatrix} 6 \\ 36 \\ 72 \\ 108 \\ 66 \end{pmatrix} \Rightarrow \begin{aligned} S_0 &= 0 \\ S_1 &= 6 \\ S_2 &= 12 \\ S_3 &= 18 \\ S_4 &= 24 \end{aligned}$$

Παρατηρήστε ότι οι τιμές των S_i ταυτίζονται με αυτές της 2ης παραγράφου της $y = x^3 - 8$ όταν χρησιμοποιήθηκαν οι συνθήκες III και IV.

3.6 ΑΣΚΗΣΕΙΣ

1. Προσεγγίστε τη συνάρτηση $y(x) = \sin \frac{\pi}{2}x$ με ένα δευτεροβάθμιο πολυώνυμο (συμπτωτικό), το οποίο παίρνει τις ίδιες τιμές με τη συνάρτηση για $x = 0$, $x_1 = 0.5$ και $x_2 = 1$. Να μελετηθεί το σφάλμα $y(x) - p(x)$.
2. Προσεγγίστε τη συνάρτηση $y(x) = \log(x + 1)$ με ένα δευτεροβάθμιο πολυώνυμο (συμπτωτικό), το οποίο παίρνει τις τιμές με τη συνάρτηση για $x = 0$, $x_1 = 4$, $x_2 = 9$. Να μελετηθεί το σφάλμα $y(x) - p(x)$.
3. Χρησιμοποιώντας τον τύπο του Lagrange βρείτε ένα συμπτωτικό πολυώνυμο τρίτου βαθμού για τα παρακάτω ζεύγη x_k, y_k .

| | | | | |
|-------|---|----|---|----|
| x_k | 0 | 1 | 4 | 6 |
| y_k | 1 | -1 | 1 | -1 |

- Μετά, υπολογίστε τις τιμές του πολυωνύμου για $x = 2, 3, 5$.
4. Ξερισμοποιώντας τον τύπο του Lagrange βρείτε ένα συμπτωτικό πολυώνυμο τετάρτου βαθμού για τα παρακάτω ζεύγη x_k, y_k .
Μετά, υπολογίστε τις τιμές του πολυωνύμου για $x = 2$ και $x = 3$.
 5. Εφαρμόζοντας τον τύπο του Newton προσδιορίστε ένα συμπτωτικό πολυώνυμο τετάρτου βαθμού, που παίρνει τις τιμές :

| | | | | | |
|-------|---|----|----|----|---|
| x_k | 0 | 1 | 4 | 5 | 6 |
| y_k | 0 | 16 | 48 | 88 | 0 |

| | | | | | |
|-------|---|----|---|----|---|
| x_k | 1 | 2 | 3 | 4 | 5 |
| y_k | 1 | -1 | 1 | -1 | 1 |

6. Να βρεθεί ένα συμπτωτικό πολυώνυμο δευτέρου βαθμού, που να παίρνει τις ίδιες τιμές με την $y(x) = x^4$ για $x = 0, 1, 2$.
7. Να βρεθεί ένα συμπτωτικό πολυώνυμο δευτέρου βαθμού, που να παίρνει τις ίδιες τιμές με τη συνάρτηση $y(x) = \sqrt{x}$ για $x = 0, 1, 4$.
8. Να βρεθεί ένα συμπτωτικό πολυώνυμο τετάρτου βαθμού, που να παίρνει τις τιμές :

| | | | | | |
|-------|---|---|---|---|----|
| x_k | 2 | 4 | 6 | 8 | 10 |
| y_k | 0 | 0 | 1 | 0 | 0 |

9. Να βρεθεί ένα συμπτωτικό πολυώνυμο εβδόμου βαθμού, που να παίρνει τις τιμές :

| | | | | | | | | |
|-------|---|---|---|---|----|----|----|----|
| x_k | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| y_k | 1 | 2 | 4 | 7 | 11 | 16 | 22 | 29 |

10. Από τον τύπο του Hermite βρείτε ένα εφαπτόμενο πολυώνυμο 3ου βαθμού για τα δεδομένα (Απάντηση $p(x) = 2x^2 - x^3$):

| | | |
|-------|-------|--------|
| x_k | y_k | y'_k |
| 0 | 0 | 0 |
| 1 | 1 | 1 |

11. Από τον τύπο του Hermite βρείτε ένα εφαπτόμενο πολυώνυμο με δεδομένα (Απάντηση $p(x) = x^4 - 4x^3 + 4x^2$):

| | | |
|-------|-------|--------|
| x_k | y_k | y'_k |
| 0 | 0 | 0 |
| 1 | 1 | 0 |
| 2 | 0 | 0 |

12. Βρείτε ένα εφαιπτόμενο πολυώνυμο τρίτου βαθμού για τα δεδομένα (Απάντηση $p(x) = x^3 - x^2 + 1$):

| x_k | y_k | y'_k |
|-------|-------|--------|
| 0 | 1 | -2 |
| 1 | 1 | 4 |

13. Βρείτε δυο πολυώνυμα δευτέρου βαθμού, το ένα με $p_1(0) = p'_1(0) = 0$ και το άλλο με $p_2(4) = 2$ και $p'_2(4) = 0$, που να περνούν και τα δυο από το σημείο $(2, 1)$. Δείξτε ότι $p'_1(2) = p'_2(2)$. Τι σημαίνει αυτό.
14. Βρείτε τα πολυώνυμα Taylor βαθμού n για τις συναρτήσεις $\sin x$ και $\cos x$, με $x_0 = 0$. Για ποιο n το πολυώνυμο Taylor προσεγγίζει αυτές τις συναρτήσεις σωστά μέχρι και το τρίτο δεκαδικό ψηφίο, για $x \in (0, \pi/2)$.
15. Ένας μεταβλητός αστέρας παρουσιάζει τις εξής μεταβολές φαινομένου μεγέθους ως συνάρτηση του χρόνου

| Χρονos (sec) | 0.0 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 1.0 |
|-------------------|-------|-------|-------|-------|------|-------|-------|-------|-------|
| Φαινόμενο μέγεθος | 0.302 | 0.185 | 0.106 | 0.093 | 0.24 | 0.579 | 0.561 | 0.468 | 0.302 |

Ποιο είναι το φαινόμενο μέγεθος τη χρονική στιγμή $t = 0.35$. Χρησιμοποιείτε splines για να συμπληρώσετε τον πίνακα ώστε να γνωρίζουμε την τιμή του φαινομένου μεγέθους κάθε 0.05.

16. Ο παρακάτω πίνακας δίνει το μήκος του τόξου ενός λεπτού της μοίρας στο γεωγραφικό παράλληλο (προσέξτε ότι στον ισημερινό αντιστοιχεί στο ναυτικό μίλι) και την σταθερά της βαρύτητας σε διάφορα γεωγραφικά πλάτη.

| Γεωγραφικό Πλάτος | Μήκος 1' της μοίρας στον γεωγραφικό παράλληλο | Σταθερά της βαρύτητας g |
|-------------------|---|---------------------------|
| 0° | 1855.4 m | 9.7805 m/sec ² |
| 15° | 1792.0 m | 9.7839 m/sec ² |
| 30° | 1608.2 m | 9.7934 m/sec ² |
| 45° | 1314.2 m | 9.8063 m/sec ² |
| 60° | 930.0 m | 9.8192 m/sec ² |
| 75° | 481.7 m | 9.8287 m/sec ² |
| 90° | 0.0 m | 9.8322 m/sec ² |

Να βρεθεί πόσο είναι το μήκος l' της μοίρας στο γεωγραφικό πλάτος της Θεσ/νίκης ($40^0\ 37'$) και ποία είναι η τοπική τιμή της σταθεράς της βαρύτητας g .

ΑΡΙΘΜΗΤΙΚΗ ΠΑΡΑΓΩΓΙΣΗ

Η αριθμητική παραγωγή είναι ιδιαίτερα χρήσιμη, καθώς πολλά φυσικά μεγέθη αντιπροσωπεύουν τη χρονική ή χωρική μεταβολή κάποιας μετρήσιμης ποσότητας. Για παράδειγμα, σ' ένα πείραμα μπορεί να μετρήσουμε τη θέση ενός κινούμενου αντικειμένου σε διάφορες χρονικές στιγμές. Για να υπολογίσουμε και την ταχύτητα του αντικειμένου στις διάφορες χρονικές στιγμές, θα πρέπει να βρούμε την παράγωγο των μετρήσεων της θέσης ως προς το χρόνο, με κάποια αριθμητική μέθοδο. Άλλες εφαρμογές των μεθόδων που θα αναπτυχθούν στο παρόν Κεφάλαιο περιλαμβάνουν την ανάπτυξη αλγορίθμων για τη λύση προβλημάτων οριακών τιμών, την ανάπτυξη μεθόδων για την επίλυση διαφορικών εξισώσεων κλπ.

Γενικά, η αριθμητική παράγωγος χρησιμοποιείται όταν δεν είναι γνωστή η αναλυτική μορφή μιας συνάρτησης $y = y(x)$ και γνωρίζουμε μόνο τις τιμές της συνάρτησης σε συγκεκριμένα σημεία.

4.1 ΑΡΙΘΜΗΤΙΚΗ ΠΑΡΑΓΩΓΙΣΗ ΜΕ ΧΡΗΣΗ ΣΥΜΠΤΩΤΙΚΟΥ ΠΟΛΥΩΝΥΜΟΥ

Η πλέον προφανής διαδικασία για την αριθμητική παραγωγή μιας σύνθετης συνάρτησης $y(x)$ είναι κατ' αρχάς ο υπολογισμός του συμπτωτικού πολυωνύμου $P(x)$ και στη συνέχεια η παραγωγή του. Για παράδειγμα, αν χρησιμοποιήσουμε το συμπτωτικό πολυώνυμο Newton προς τα εμπρός, σε κάποιο σημείο $x = x_0 + sh$ δηλαδή

$$y(x) \rightarrow P(x) = y_0 + s\Delta y_0 + \frac{s(s-1)}{2!}\Delta^2 y_0 + \frac{s(s-1)(s-2)}{3!}\Delta^3 y_0 + \dots \quad (4.1)$$

και στη συνέχεια το παραγωγίσουμε, θα πάρουμε:

$$\frac{dy(x)}{dx} = \frac{dP(x)}{dx} = \frac{1}{s} \frac{dP(s)}{ds} = \frac{1}{h} \left[\Delta y_0 + \frac{2s-1}{2!}\Delta^2 y_0 + \frac{3s^2-6s+2}{3!}\Delta^3 y_0 + \dots \right] \quad (4.2)$$

Επομένως, αν θελήσουμε να υπολογίσουμε την παράγωγο στη θέση x_0 αρκεί να θέσουμε στην παραπάνω σχέση $s = 0$, για τη θέση x_1 αρκεί να θέσουμε στην παραπάνω σχέση $s = 1$ κ.ο.κ., άρα:

$$y'_0 = \frac{1}{h} \left[\Delta y_0 - \frac{1}{2} \Delta^2 y_0 + \frac{1}{3} \Delta^3 y_0 + \dots \right] \quad (4.3)$$

και ανάλογα με τον αριθμό των όρων που θέλουμε να διατηρήσουμε, μπορούμε να δημιουργήσουμε τις παρακάτω σχέσεις :

$$y'_0 = \frac{y_1 - y_0}{h} + O(h) \quad (4.4)$$

$$y'_0 = -\frac{3y_0 - 4y_1 + y_2}{2h} + O(h^2) \quad (4.5)$$

$$y'_0 = -\frac{11y_0 - 18y_1 + 9y_2 - 2y_3}{6h} + O(h^3) \quad (4.6)$$

...

Με ανάλογο τρόπο θα υπολογισθεί και η δεύτερη παράγωγος της συνάρτησης $y(x)$ στη θέση x_0 . Είναι:

$$\frac{d^2 P(x)}{dx^2} = \frac{1}{h^2} \frac{d^2 P(s)}{ds^2} = \frac{1}{h^2} [\Delta^2 y_0 + (s-1)\Delta^3 y_0 + \dots] \quad (4.7)$$

οπότε

$$y''_0 = \frac{y_0 - 2y_1 + y_2}{h^2} + O(h) \quad (4.8)$$

και

$$y''_0 = \frac{2y_0 - 5y_1 + 4y_2 - y_3}{h^2} + O(h^2) \quad (4.9)$$

κ.ο.κ.

Το σφάλμα υπολογίζεται εύκολα από τον τύπο του σφάλματος του συμπτωτικού πολυωνύμου. Δηλαδή, γνωρίζουμε από την σχέση (3.3) ότι το σφάλμα του συμπτωτικού πολυωνύμου n βαθμού (που προκύπτει από $n+1$ σημεία) είναι:

$$E(x) = (x - x_0)(x - x_1)(x - x_2) \dots (x - x_n) \frac{y^{n+1}(\xi)}{(n+1)!} \quad (4.10)$$

οπότε το σφάλμα της παραγώγου του συμπτωτικού πολυωνύμου για τη θέση π.χ. $x = x_0$ θα βρεθεί παραγωγίζοντας την παραπάνω σχέση, δηλαδή :

$$E'(x_0) = (x_0 - x_1)(x_0 - x_2) \dots (x_0 - x_n) \frac{y^{n+1}(\xi)}{(n+1)!} \quad (4.11)$$

οπότε αν υποθέσουμε ισαπέχοντα σημεία $h = x_{i+1} - x_i$, καταλήγουμε στη σχέση

$$E'(x_0) = -h(-2h) \dots (-nh) \frac{y^{n+1}(\xi)}{(n+1)!} = (-1)^n h^n \frac{y^{n+1}(\xi)}{n+1}. \quad (4.12)$$

Επομένως το σφάλμα στη σχέση (4.4) θα είναι $O(h)$, στην (4.5) θα είναι $O(h^2)$ και στην (4.6) $O(h^3)$. Ανάλογα, το σφάλμα στη δεύτερη παράγωγο θα είναι $O(h^{(n-1)})$ (βλ. άσκηση 4.1). Επομένως, στη σχέση (4.8) θα είναι $O(h)$ και στη σχέση (4.9), $O(h^2)$.

Αντίστοιχες σχέσεις μπορώ να δημιουργήσω και με τη χρήση του πολυωνύμου Newton προς τα πίσω, οπότε υπολογίζω την παράγωγο μιας συνάρτησης στο τέλος του διαστήματος.

ΕΦΑΡΜΟΓΗ

Υπολογίστε τη δεύτερη παράγωγο μιας συνάρτησης στη θέση $x = x_1$ χρησιμοποιώντας τις τιμές y_0, y_1, y_2 και y_3 .

Θα χρησιμοποιήσουμε το συμπτωτικό πολυώνυμο Lagrange υποθέτοντας όμως ότι τα σημεία x_0, x_1, x_2, \dots είναι ισαπέχοντα. Το πολυώνυμο Lagrange είναι:

$$P(x) = \frac{(x-x_1)(x-x_2)(x-x_3)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)}y_0 + \frac{(x-x_0)(x-x_2)(x-x_3)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)}y_1 \\ + \frac{(x-x_0)(x-x_1)(x-x_3)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)}y_2 + \frac{(x-x_0)(x-x_1)(x-x_2)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)}y_3 \quad (4.13)$$

οπότε παραγωγίζοντας δυο φορές και αντικαθιστώντας καταλήγω:

$$P''(x) = \frac{2y_0}{-6h^3}[(x-x_1) + (x-x_2) + (x-x_3)] \\ + \frac{2y_1}{2h^3}[(x-x_0) + (x-x_2) + (x-x_3)] \\ + \frac{2y_2}{-2h^3}[(x-x_0) + (x-x_1) + (x-x_3)] \\ + \frac{2y_3}{6h^3}[(x-x_0) + (x-x_1) + (x-x_2)] \quad (4.14)$$

και θέτοντας $x = x_1$ καταλήγουμε εύκολα στη σχέση:

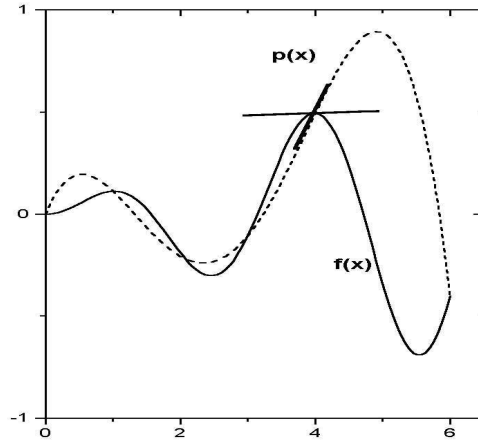
$$P''(x) = \frac{y_0 - 2y_1 + y_2}{h^2} \quad (4.15)$$

Ο όρος y_3 , αν και είχε χρησιμοποιηθεί, απουσιάζει από την τελική σχέση, παρ' όλα αυτά η ακρίβεια της μεθόδου είναι $O(h^2)$, όπως προβλέπεται από τη χρήση συμπτωτικού πολυωνύμου για τέσσερα σημεία. Στη συνέχεια, θα δούμε ότι η παραπάνω σχέση μπορεί να δημιουργηθεί και με τη χρήση της τεχνικής των κεντρικών διαφορών.

ΑΚΡΙΒΕΙΑ

Θα πρέπει να επισημανθεί ότι γενικά το συμπτωτικό πολυώνυμο μπορεί να προσεγγίζει με ικανοποιητική ακρίβεια μια συνάρτηση και να συμπίπτει με

αυτήν σε ορισμένα σημεία, δεν συμβαίνει όμως συνήθως το ίδιο και με την παράγωγό του. Όπως φαίνεται από το Σχήμα 4.1 είναι προφανές το σφάλμα στον υπολογισμό της $y'(x)$ με τη χρήση του $P'(x)$.



Σχήμα 4.1. Γραφική απεικόνιση μιας συνάρτησης $f(x)$ και του συμπτωτικού της πολυωνύμου $p(x)$. Στην τιμή $x = 4$ συμπίπτουν αλλά οι παράγωγοι τους παρουσιάζουν σημαντική απόκλιση.

4.2 ΤΥΠΟΙ ΚΕΝΤΡΙΚΩΝ ΔΙΑΦΟΡΩΝ

Μια κατηγορία σχέσεων για τον υπολογισμό αριθμητικών παραγώγων είναι αυτή των κεντρικών διαφορών. Εδώ, η παράγωγος οποιασδήποτε τάξης στο σημείο x_0 υπολογίζεται από την τιμή της συνάρτησης σ' αυτό το σημείο και σε ζεύγη σημείων εκατέρωθεν του x_0 συμμετρικά. Τέτοιες σχέσεις είναι οι παρακάτω:

$$y(x_0)' = y_0' = \frac{y_1 - y_{-1}}{2h} + O(h^2) \quad (4.16)$$

$$y(x_0)' = y_0' = \frac{-y_2 + 8y_1 - 8y_{-1} + y_{-2}}{12h} + O(h^4) \quad (4.17)$$

$$y(x_0)'' = y_0'' = \frac{y_1 - 2y_0 + y_{-1}}{h^2} + O(h^2) \quad (4.18)$$

$$y(x_0)'' = y_0'' = \frac{-y_2 + 16y_1 - 30y_0 + 16y_{-1} - y_{-2}}{12h^2} + O(h^4) \quad (4.19)$$

$$y(x_0)''' = y_0''' = \frac{y_2 - 2y_1 + 2y_{-1} - y_{-2}}{2h^3} + O(h^2) \quad (4.20)$$

$$y(x_0)^{(4)} = y_0^{(4)} = \frac{y_2 - 4y_1 + 6y_0 - 4y_{-1} + y_{-2}}{h^4} + O(h^2) \quad (4.21)$$

Στη συνέχεια θα δείξουμε τον τρόπο δημιουργίας μερικών από τις παραπάνω σχέσεις. Ας θεωρήσουμε τα αναπτύγματα Taylor δεξιά και αριστερά του σημείου x_0 , στις θέσεις $x_0 \pm h$ και $x_0 \pm 2h$

$$y(x_0 + h) \equiv y_1 = y_0 + hy'_0 + \frac{h^2}{2}y''_0 + \frac{h^3}{6}y'''_0 + \frac{h^4}{24}y_0^{(4)} + \dots \quad (4.22)$$

$$y(x_0 - h) \equiv y_{-1} = y_0 - hy'_0 + \frac{h^2}{2}y''_0 - \frac{h^3}{6}y'''_0 + \frac{h^4}{24}y_0^{(4)} - \dots \quad (4.23)$$

$$y(x_0 + 2h) \equiv y_2 = y_0 + 2hy'_0 + 2h^2y''_0 + \frac{4}{3}h^3y'''_0 + \frac{2}{3}h^4y_0^{(4)} + \dots \quad (4.24)$$

$$y(x_0 - 2h) \equiv y_{-2} = y_0 - 2hy'_0 + 2h^2y''_0 - \frac{4}{3}h^3y'''_0 + \frac{2}{3}h^4y_0^{(4)} - \dots \quad (4.25)$$

μπορώ με κατάλληλους συνδυασμούς τους να δημιουργήσω τις ζητούμενες σχέσεις. Για παράδειγμα, αφαιρώντας την (4.23) από την (4.22) λαμβάνω τη σχέση (4.16), ενώ προσθέτοντας έχω τη σχέση (4.18) την οποία είχαμε αποδείξει και νωρίτερα με τη χρήση του συμπληρωτικού πολυωνύμου Lagrange, σχέση (4.15).

Ενεργώντας ανάλογα και πολλαπλασιάζοντας την (4.22) και την (4.23) με 8 και κάνοντας χρήση των (4.24) και (4.25) δημιουργώ μια σχέση της μορφής

$$y_{-2} - y_2 + 8(y_1 - y_{-1}) \quad (4.26)$$

που εύκολα μας οδηγεί στη σχέση (4.17) η οποία έχει σφάλμα $\sim O(h^4)$. Η συγκεκριμένη επιλογή έγινε με τέτοιο τρόπο ώστε πέραν των όρων με άρτιες δυνάμεις του h να απειλειφθούν και οι όροι που περιέχουν το h^3 , οπότε οι μόνοι όροι που απέμειναν είναι αυτοί που δημιουργούν τη ζητούμενη σχέση και οι όροι σφάλματος που στο συγκεκριμένο ανάπτυγμα είναι αυτοί που περιέχουν το h^5 . Με ανάλογο τρόπο δημιουργούνται και οι υπόλοιπες σχέσεις.

ΠΑΡΑΔΕΙΓΜΑ

Μια συνάρτηση $f(x)$ παίρνει τις τιμές του Πίνακα 4.1 Να βρεθεί η $f'(1)$ (Οι τιμές του πίνακα προέρχονται από τη συνάρτηση $f(x) = e^x$).

| | | | |
|--------|--------|--------|--------|
| x | 0.90 | 1.00 | 1.11 |
| $f(x)$ | 2.4596 | 2.7183 | 3.0344 |

Πίνακας 4.1.

Παρατηρούμε ότι τα τρία σημεία δεν ισαπέχουν ($h = x_0 - x_{-1} = 0.1$, $h_1 = x_1 - x_0 = 0.11$), οπότε θα πρέπει να δημιουργήσουμε μια νέα σχέση υπολογισμού της παραγώγου με βάση την προηγούμενη θεωρία. Θα μπορούσαν να

χρησιμοποιηθούν δυο διαφορετικές μέθοδοι. Η προφανής επιλογή είναι να υπολογίσουμε με τον τύπο του Lagrange το δευτεροβάθμιο συμπτωτικό πολυώνυμο και, στη συνέχεια, να το παραγωγίσουμε. Η άλλη επιλογή είναι να χρησιμοποιήσουμε μια παραλλήλα της μεθόδου των κεντρικών διαφορών, όπως θα δείξουμε και στη συνέχεια.

Είναι $y(x_0 + h) = y(1 + 0.11)$ και $y(x_0 - h_1) = y(1 - 0.1)$. Γενικά, θα ισχύει:

$$y_1 = y(x_0 + h) = y_0 + hy'_0 + \frac{h^2}{2}y''_0 + \frac{h^3}{6}y'''_0 + \dots \quad (4.27)$$

$$y_{-1} = y(x_0 - h_1) = y_0 - h_1y'_0 + \frac{h_1^2}{2}y''_0 - \frac{h_1^3}{6}y'''_0 + \dots \quad (4.28)$$

οπότε διαιρώντας κατ' αρχάς την πρώτη με h^2 και τη δεύτερη με h_1^2 και στη συνέχεια αφαιρώντας τις δύο σχέσεις καταλήγω σε μια σχέση υπολογισμού της πρώτης παραγώγου με ακρίβεια $O(h_1h) \sim O(h^2)$

$$y'_0 = \frac{\frac{1}{h^2}y_1 - \frac{1}{h_1^2}y_{-1}}{\frac{1}{h_1} + \frac{1}{h}} - \left(\frac{1}{h} - \frac{1}{h_1}\right)y_0 \quad (4.29)$$

Θέτοντας τις κατάλληλες τιμές $h_1 = 0.11$ και $h = 0.1$, βρίσκουμε, $y'_0 = 2.7233$ (η ακριβής τιμή είναι 2.7183).

ΑΣΚΗΣΕΙΣ

1. Να βρεθεί ο τύπος του σφάλματος για τη δεύτερη παράγωγο του συμπτωτικού πολυωνύμου.
2. Αποδείξτε τις σχέσεις (4.19), (4.20) και (4.21)
3. Με τη χρήση ΗΥ υπολογίστε την $f'(x_0)$ και $f''(x_0)$, για $f(x) = \tan^{-1}(x)$ και $x_0 = \sqrt{2}$. Ξερεσιμοποιήστε τις σχέσεις (4.16) και (4.18) με βήμα $h = 0.1$

ΑΡΙΘΜΗΤΙΚΗ ΟΛΟΚΛΗΡΩΣΗ

Ενας μεγάλος αριθμός προβλημάτων στη φυσική και τα μαθηματικά απαιτεί τον υπολογισμό ορισμένων ολοκληρωμάτων. Όταν όμως το αόριστο ολοκλήρωμα δεν μπορεί να υπολογισθεί είτε όταν η υπό ολοκλήρωση συνάρτηση δίνεται σε διακριτή μορφή, η μόνη διέξοδος είναι ο προσεγγιστικός υπολογισμός της τιμής του ορισμένου ολοκληρώματος. Η αριθμητική ολοκλήρωση είναι εν γένει ευσταθής διαδικασία και η σύγκλιση σε ένα ακριβές αποτέλεσμα δεν είναι δύσκολη. Εν τούτοις, υπάρχουν πολλές ‘ παθολογικές ’ καταστάσεις αλλά και δυνατότητες επίτευξης υψηλότερης ακρίβειας με ελάχιστο αριθμό πράξεων αν χρησιμοποιηθεί κατάλληλη διαδικασία.

5.1 ΤΥΠΟΙ NEWTON-COTES

Η τεχνική της αριθμητικής ολοκλήρωσης είναι ανάλογη μ’ αυτή της αριθμητικής παραγωγής. Κατ’ αρχάς, είτε δίδεται μία συνάρτηση $f(x)$ προς ολοκλήρωση είτε δίνονται τα σημεία (x_i, y_i) από τα οποία διέρχεται η συνάρτηση. Θα υπολογίσουμε ένα κατάλληλο συμπτωτικό πολυώνυμο και στη συνέχεια αντί της συνάρτησης ολοκληρώνουμε το συμπτωτικό πολυώνυμο. Δηλαδή:

$$\int_a^b y(x)dx \rightarrow \int_a^b P_n(x)dx \quad (5.1)$$

Το συμπτωτικό πολυώνυμο μπορεί να υπολογισθεί με χρήση του τύπου του Newton

$$P_n(x_s) = f_0 + s\Delta f_0 + \frac{s(s-1)}{2!}\Delta^2 f_0 + \frac{s(s-1)(s-2)}{3!}\Delta^3 f_0 + \dots \quad (5.2)$$

Υπενθυμίζουμε ότι $x_s = x_0 + sh$.

Οπότε ανάλογα με τον αριθμό των όρων του πολυωνύμου που διατηρούμε λαμβάνουμε διάφορετικής ακρίβειας μεθόδους προσεγγιστικού υπολογισμού

της τιμής των ορισμένων ολοκληρωμάτων. Το σφάλμα στην αριθμητική ολοκλήρωση θα υπολογισθεί από την ολοκλήρωση του σφάλματος του συμπτωτικού πολυωνύμου, δηλαδή

$$E = \int_a^b E_n(x_s) dx \quad (5.3)$$

όπου

$$E_n(x_s) = \frac{s(s-1)(s-2)\dots(s-n)}{(n+1)!} h^{n+1} f^{(n+1)}(\xi) \quad (5.4)$$

για $\xi \in [a, b]$.

Στη συνέχεια θα δημιουργήσουμε τύπους αριθμητικής ολοκλήρωσης με βάση το βαθμό του συμπτωτικού πολυωνύμου.

- **Συμπτωτικό πολυώνυμο 1ου βαθμού**

Βρίσκουμε:

$$\begin{aligned} \int_{x_0}^{x_1} f(x) dx &\rightarrow \int_{x_0}^{x_1} P_1(x_s) dx = \int_{x_0}^{x_1} (f_0 + s \Delta f_0) dx = h \int_{s=0}^{s=1} (f_0 + s \Delta f_0) ds \\ &= h f_0 s \Big|_0^1 + h \Delta f_0 \frac{s^2}{2} \Big|_0^1 = h \left(f_0 + \frac{1}{2} \Delta f_0 \right) \\ &= \frac{h}{2} (f_0 + f_1) \end{aligned} \quad (5.5)$$

Είναι προφανές ότι έχουμε αντικαταστήσει στο διάστημα ολοκλήρωσης, από x_0 έως x_1 τη συνάρτηση $f(x)$ με μια ευθεία που ενώνει τα σημεία $(x_0, f(x_0))$ και $(x_1, f(x_1))$. Στη συνέχεια υπολογίζουμε το εμβαδό του τραπέζιου που σχηματίζεται.

ΣΦΑΛΜΑ

Το σφάλμα στον προσεγγιστικό υπολογισμό του ορισμένου ολοκληρώματος λόγω της αντικατάστασης της συνάρτησης $f(x)$ με το συμπτωτικό πολυώνυμο $P(x)$ είναι:

$$f(x) - P(x) = \frac{1}{2} s(s-1) h^2 f''(\xi) \quad \text{για } x_0 \leq \xi \leq x_1 \quad (5.6)$$

όποτε το σφάλμα στην προηγούμενη προσεγγιστική ολοκλήρωση υπολογίζεται από την ολοκλήρωση του παραπάνω σφάλματος, δηλαδή

$$\begin{aligned} E &= \int_{x_0}^{x_1} \frac{1}{2} s(s-1) h^2 f''(\xi) dx = \frac{h^3}{2} \int_{s=0}^{s=1} s(s-1) f''(\xi) ds \\ &= h^3 f''(\xi_1) \left(\frac{s^3}{6} - \frac{s^2}{4} \right) \Big|_0^1 = -\frac{1}{12} h^3 f''(\xi_1) \end{aligned} \quad (5.7)$$

όπου $\xi_1 \in [x_0, x_1]$.

- **Συμπτωτικό πολυώνυμο 2ου βαθμού**

Ακολουθώντας την προηγούμενη διαδικασία βρίσκουμε ότι:

$$\begin{aligned}
 \int_{x_0}^{x_2} f(x) dx &\rightarrow \int_{x_0}^{x_2} P_2(x_s) dx = \int_{x_0}^{x_2} \left(f_0 + s\Delta f_0 + \frac{1}{2}s(s-1)\Delta^2 f_0 \right) dx \\
 &= h \int_{s=0}^{s=2} \left(f_0 + s\Delta f_0 + \frac{1}{2}s(s-1)\Delta^2 f_0 \right) ds \\
 &= h \left(2f_0 + 2\Delta f_0 + \frac{1}{3}\Delta^2 f_0 \right) \\
 &= \frac{h}{3} (f_0 + 4f_1 + f_2)
 \end{aligned} \tag{5.8}$$

ΣΦΑΛΜΑ

Μια σημαντική διαπίστωση είναι ότι αν ολοκληρώσουμε τον επόμενο όρο του συμπτωτικού πολυωνύμου για τον υπολογισμό του σφάλματος βρίσκουμε ότι μηδενίζεται. Δηλαδή

$$\frac{1}{6} \int_{x_0}^{x_2} s(s-1)(s-2)\Delta^3 f_0 dx = 0. \tag{5.9}$$

Το αποτέλεσμα αυτό αποτελεί ‘ευτυχή σύμπτωση’ διότι με αυτό τον τρόπο το σφάλμα θα προέλθει από τον ανώτερης τάξης όρο του συμπτωτικού πολυωνύμου και επομένως θα είναι μικρότερο από ότι θα αναμέναμε. Επομένως θα ολοκληρώσουμε το μεθεπόμενο όρο και το σφάλμα στην αριθμητική ολοκλήρωση είναι:

$$E = \frac{1}{24} \int_{x_0}^{x_2} s(s-1)(s-2)(s-3)h^4 f^{(4)}(\xi) dx = \dots = -\frac{1}{90} h^5 f^{(4)}(\xi_1) \tag{5.10}$$

για $x_0 \leq \xi_1 \leq x_2$.

- **Προσέγγιση με συμπτωτικό πολυώνυμο 3ου βαθμού**

Με βάση τα προηγούμενα βρίσκουμε:

$$\int_{x_0}^{x_3} f(x) dx \rightarrow \int_{x_0}^{x_3} P_3(x_s) dx = \frac{3h}{8} (f_0 + 3f_1 + 3f_2 + f_3). \tag{5.11}$$

ΣΦΑΛΜΑ

Με ανάλογο τρόπο βρίσκουμε ότι:

$$E = -\frac{3}{80} h^5 f^{(4)}(\xi_1) \quad \text{για } x_0 \leq \xi_1 \leq x_3 \tag{5.12}$$

Δηλαδή, το σφάλμα είναι της αυτής τάξης, $O(h^5)$, με την προηγούμενη μέθοδο όπου χρησιμοποιήσαμε συμπτωτικό πολυώνυμο 2ου βαθμού. Η ακρίβεια

βελτιώνεται εάν διατηρήσουμε και τον επόμενο όρο στο συμπτωτικό πολυώνυμο, δηλαδή θα παρατηρήσουμε σημαντική βελτίωση στην ακρίβεια αν χρησιμοποιήσουμε συμπτωτικό πολυώνυμο 4ου βαθμού, το οποίο έχει ίδιας τάξης σφάλμα με την ολοκλήρωση συμπτωτικού πολυωνύμου 5ου βαθμού, δηλαδή $O(h^7)$.

Στη βιβλιογραφία, οι παραπάνω σχέσεις για την αριθμητική ολοκλήρωση, που εξάγονται με τη χρήση του συμπτωτικού πολυωνύμου, είναι γνωστοί ως **τύποι Newton-Cotes**.

$$\begin{aligned}\int_{x_0}^{x_1} f(x)dx &= \frac{h}{2} (f_0 + f_1) - \frac{1}{12} h^3 f^{(2)}(\xi_1) \\ \int_{x_0}^{x_2} f(x)dx &= \frac{h}{3} (f_0 + 4f_1 + f_2) - \frac{1}{90} h^5 f^{(4)}(\xi_1) \\ \int_{x_0}^{x_3} f(x)dx &= \frac{3h}{8} (f_0 + 3f_1 + 3f_2 + f_3) - \frac{3}{80} h^5 f^{(4)}(\xi_1)\end{aligned}$$

Πίνακας 5.1. Τύποι Newton-Cotes για αριθμητική ολοκλήρωση με χρήση 1ης, 2ης και 3ης τάξης συμπτωτικού πολυωνύμου Newton.

5.1.1 Κανόνες τραπεζίου

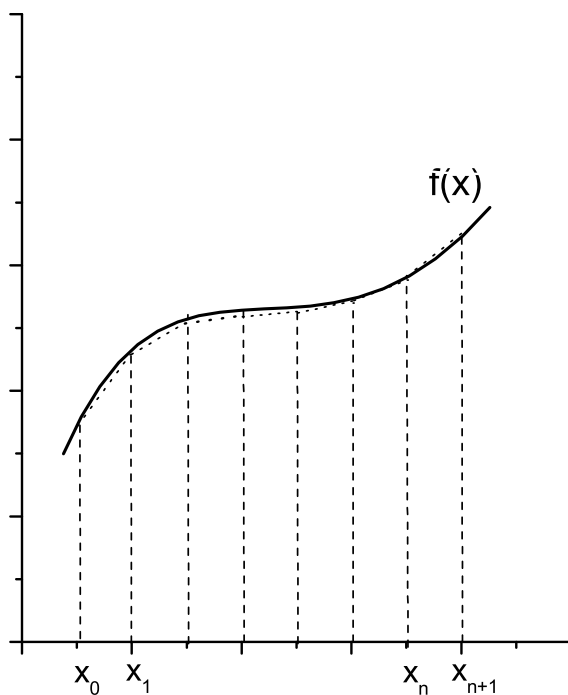
Αν το διάστημα ολοκλήρωσης (a, b) είναι μεγάλο, τότε ορίζουμε μια διαμέριση $\{a = x_0, \dots, x_n = b; n\}$, όπως φαίνεται στο Σχήμα 5.1. Στην ουσία παίρνουμε το άθροισμα των εμβαδών των επιμέρους τραπεζίων. Αν τα σημεία είναι ισαπέχοντα, τότε με βάση τα προηγούμενα:

$$\begin{aligned}\int_a^b f(x)dx &= \int_{x_0}^{x_n} f(x)dx = \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} P_1(x)dx \\ &= \int_{x_0}^{x_1} P_1(x)dx + \int_{x_1}^{x_2} P_1(x)dx + \dots + \int_{x_{n-1}}^{x_n} P_1(x)dx \\ &= \sum_{i=0}^{n-1} \frac{h}{2} (f_i + f_{i+1})\end{aligned}$$

Οπότε καταλήγουμε στη σχέση:

$$\int_a^b f(x)dx = \frac{h}{2} (f_0 + 2f_1 + \dots + 2f_n + f_{n+1}) \quad (5.13)$$

Ο παραπάνω τύπος για τον αριθμητικό υπολογισμό της τιμής ορισμένων ολοκληρωμάτων ονομάζεται **κανόνες του τραπεζίου**.



Σχήμα 5.1. Γραφική απεικόνιση της μέθοδου του τραπεζίου.

ΣΦΑΛΜΑ

Το συνολικό σφάλμα θα είναι το άθροισμα των σφαλμάτων σε κάθε ένα από τα υποδιαστήματα (x_i, x_{i+1}) , δηλαδή

$$E = -\frac{h^3}{12} [f''(\xi_1) + f''(\xi_2) + \dots + f''(\xi_n)] \quad (5.14)$$

αν όμως η $f''(x)$ είναι φραγμένη και συνεχής στο διάστημα $[a = x_0, b = x_n]$, τότε υπάρχει ένα κατάλληλο ξ τέτοιο, ώστε το άθροισμα στην παραπάνω σχέση να γίνει $nf''(\xi)$. Οπότε το συνολικό σφάλμα θα είναι:

$$E = -\frac{h^3}{12} nf''(\xi) = -\frac{h^2}{12} (b-a)f''(\xi) \quad (5.15)$$

επειδή $h = (b-a)/n$.

5.1.2 Κανόνas του Simpson

Αν το υπό ολοκλήρωση διάστημα το χωρίσουμε σε τριάδες σημείων, τότε για κάθε τριάδα μπορούμε να χρησιμοποιήσουμε τον τύπο (5.8) για την προσέγγιση της συνάρτησης με συμπτωτικό πολυώνυμο 2ου βαθμού. Επομένως, απαιτείται **άρτιος αριθμός υποδιαϊρέσεων** του διαστήματος (a, b) . Δηλαδή

$$\begin{aligned}\int_{a=x_0}^{b=x_n} f(x)dx &= \sum_{i=0}^{n-2} \int_{x_i}^{x_{i+2}} P_2(x)dx = \sum_{i=0}^{n-2} \frac{h}{3} (f_i + 4f_{i+1} + f_{i+2}) \\ &= \frac{h}{3} (f_1 + 4f_2 + 2f_3 + 4f_4 + \dots + 2f_{n-2} + 4f_{n-1} + f_n) \quad (5.16)\end{aligned}$$

ΣΦΑΛΜΑ

Το συνολικό σφάλμα θα είναι προφανώς, το άθροισμα των επιμέρους σφαλμάτων, σχέση (5.10). Επομένως, το συνολικό σφάλμα θα είναι:

$$E = -\frac{h^5}{90} \frac{n}{2} f^{(4)}(\xi) = -\frac{b-a}{180} h^4 f^{(4)}(\xi) \quad \text{για } x_0 \leq \xi \leq x_n \quad (5.17)$$

5.1.3 Κανόνas του Simpson (3/8)

Κατ' αναλογία με την προηγούμενη μέθοδο, αν για κάθε τετράδα σημείων δημιουργήσω ένα συμπτωτικό πολυώνυμο τότε θα καταλήξω στον τύπο

$$\int_{a=x_0}^{b=x_n} f(x) dx = \frac{3h}{8} (f_0 + 3f_1 + 3f_2 + 2f_3 + \dots + 2f_{n-3} + 3f_{n-2} + 3f_n + f_{n+1}) \quad (5.18)$$

που προφανώς θα ισχύει για αριθμό υποδιαϊρέσεων που διαιρείται με το 3.

ΣΦΑΛΜΑ

Το συνολικό σφάλμα θα είναι:

$$E = -\frac{b-a}{80} h^4 f^{(4)}(\xi_1) \quad \text{για } x_0 \leq \xi_1 \leq x_n \quad (5.19)$$

5.1.4 Η βελτίωση του Romberg

Αν με τον κανόνα του τραπεζίου έχουμε υπολογίσει αριθμητικά μια τιμή προσεγγιστική τιμή ενός ολοκληρώματος, έστω I_1 , για βήμα $h = x_{i+1} - x_i$ και στη συνέχεια υπολογίσουμε μια νέα προσεγγιστική τιμή I_2 για βήμα kh , τότε μπορούμε να κάνουμε μια πρόβλεψη για την ακριβή τιμή.

Για παράδειγμα, αν χρησιμοποιούμε τη μέθοδο τραπεζίου με βήμα h θα υπολογίσουμε μια τιμή για το ολοκλήρωμα έστω την I_1 , ενώ για βήμα kh μια άλλη τιμή του ολοκληρώματος έστω την I_2 , οπότε αν η ακριβής τιμή είναι η A θα έχουμε ότι:

$$\begin{aligned}
\int_{x_0}^{x_n} f(x)dx &= \frac{h}{2} (f_0 + 2f_1 + 2f_2 + \dots + 2f_{n-1} + f_n) \\
&\quad - \frac{b-a}{12} h^2 f^{(2)}(\xi_1) \\
\int_{x_0}^{x_n} f(x)dx &= \frac{h}{3} (f_0 + 4f_1 + 2f_2 + 4f_3 + \dots + 2f_{n-2} + 4f_{n-1} + f_n) \\
&\quad - \frac{b-a}{180} h^5 f^{(4)}(\xi_1) \\
\int_{x_0}^{x_n} f(x)dx &= \frac{3h}{8} (f_0 + 3f_1 + 3f_2 + 2f_3 + 3f_4 + 3f_5 + \dots + 3f_{n-2} + 3f_{n-1} + f_n) \\
&\quad - \frac{b-a}{80} h^4 f^{(4)}(\xi_1)
\end{aligned}$$

Πίνακας 5.2. Οι τύποι αριθμητικής ολοκλήρωσης και τα σφάλματά τους.

- για βήμα h : Ακριβής Τιμή $A = I_1 + ch^2$
- για βήμα kh : Ακριβής Τιμή $A = I_2 + ck^2h^2$

όπου c μια σταθερά. Δηλαδή, δημιουργήσαμε ένα σύστημα δύο εξισώσεων με δύο αγνώστους το A και το c , του οποίου η λύση είναι:

$$A = \frac{k^2 I_1 - I_2}{k^2 - 1} \quad \text{και} \quad c = \frac{I_2 - I_1}{h^2(1 - k^2)} \quad (5.20)$$

Επομένως η *ακριβής τιμή* καθορίζεται από τις τιμές των I_1 , I_2 και k . Αν για παράδειγμα θέσουμε $k = 1/2$, τότε

$$A = I_2 + \frac{1}{3} (I_2 - I_1) \quad (5.21)$$

Γενικότερα, αν υποθέσουμε μέθοδο με σφάλμα τάξης $O(h^n)$, ο παραπάνω τύπος γενικεύεται ως ακολούθως :

$$A = I_2 + \frac{I_2 - I_1}{2^n - 1}. \quad (5.22)$$

Προφανώς η τιμή A δεν είναι η ακριβής τιμή του ολοκληρώματος αλλά μια βελτίωση ή πρόβλεψη της τιμής του τελικού αποτελέσματος με χρήση των δύο αρχικών αποτελεσμάτων I_1 και I_2 . Η μέθοδος μπορεί να χρησιμοποιηθεί για να βελτιώσει αριθμητικά αποτελέσματα που βρέθηκαν με χρήση είτε της μεθόδου του *τραπεζίου* είτε του Simpson. Πρέπει να σημειώσουμε ότι δεν απαιτούνται στην πράξη δύο ολοκληρώσεις αλλά δύο διαφορετικές αθροίσεις των τιμών της υπό ολοκλήρωση συνάρτησης. Αρκεί στην πρώτη άθροιση να πάρουμε όλες τις $n+1$ τιμές της που απαιτεί πχ η μέθοδος του τραπεζίου, και συγχρόνως να αθροίσουμε χωριστά και τις τιμές $(n+1)/2$ της συνάρτησης δηλαδή χρησιμοποιούμε

κάθε δεύτερο x_i . Με αυτό τον τρόπο δημιουργούμε τις δύο προσεγγιστικές τιμές του ολοκληρώματος, τα I_1 και I_2 , και στη συνέχεια κάνουμε χρήση της σχέσης (5.22).

ΕΦΑΡΜΟΓΗ

Αποδεικνύεται εύκολα ότι η μέθοδος Romberg για τον κανόνα τραπεζίου με $k = 1/2$ είναι ισοδύναμη της μεθόδου Simpson.

Ας δοκιμάσουμε κατάρχας για 3 σημεία. Πράγματι, το ολοκλήρωμα I_1 για βήμα $2h$ και το ολοκλήρωμα I_2 για βήμα h θα είναι:

$$\begin{aligned} I_1 &= \frac{2h}{2} (f_0 + f_2) = h(f_0 + f_2) \\ I_2 &= \frac{h}{2} (f_0 + f_1) + \frac{h}{2} (f_1 + f_2) = \frac{h}{2} (f_0 + 2f_1 + f_2) \end{aligned}$$

οπότε η «ακριβής τιμή» βάσει τα προηγούμενα, σχέση (5.21), θα είναι:

$$A = \frac{2h(f_0 + 2f_1 + f_2) - h(f_0 + f_2)}{3} = \frac{h}{3} (f_0 + 4f_1 + f_2)$$

που προφανώς αντιστοιχεί στη μέθοδο Simpson, εξίσωση (5.8). Δηλαδή, η ακρίβεια των αποτελεσμάτων με τη χρήση της μεθόδου Romberg από $O(h^3)$ έχει γίνει $O(h^5)$.

Προφανώς η παραπάνω διαδικασία μπορεί να χρησιμοποιηθεί και για βελτίωση των αποτελεσμάτων της μεθόδου Simpson οπότε η ακρίβεια των υπολογισμών από $O(h^5)$ γίνεται $O(h^7)$.

5.1.5 Εφαρμογή των Splines στην Παραγωγή και Ολοκλήρωση συναρτήσεων

Η προσέγγιση με splines αποτελεί μια εξαιρετική μέθοδο προσέγγισης μιας συνάρτησης ή ενός συνόλου (x_i, y_i) διακριτών δεδομένων με 3ο-βάθμια πολυώνυμα. Αν επομένως για ένα σύνολο διακριτών δεδομένων (x_i, y_i) έχουμε υπολογίσει τους συντελεστές των 3ο-βαθμίων πολυωνύμων τότε οι splines μπορούν να χρησιμοποιηθούν για τον αριθμητικό υπολογισμό παραγώγων και ολοκληρωμάτων. Η κυβική spline που προσεγγίζει μια συνάρτηση σε ένα διάστημα $x_i \leq x \leq x_{i+1}$ γράφεται:

$$f(x) = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + d_i \quad (5.23)$$

και οι συντελεστές a_i , b_i , c_i και d_i δίνονται από τις σχέσεις:

$$\begin{aligned} a_i &= \frac{S_{i+1} - S_i}{6(x_{i+1} - x_i)}, & b_i &= \frac{S_i}{2}, & d_i &= f(x_i) \\ c_i &= \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} - \frac{2(x_{i+1} - x_i)S_i + (x_{i+1} - x_i)S_{i+1}}{6} \end{aligned} \quad (5.24)$$

Ετσι, η **πρώτη** και η **δεύτερη παράγωγος** σε κάποιο τυχαίο σημείο (x, y) υπολογίζονται αυτόματα από την παραγωγή του 3ο-βάθμιου πολυωνύμου (5.23). Δηλαδή, από τις σχέσεις

$$f'(x) = 3a_i(x - x_i)^2 + 2b_i(x - x_i) + c_i \quad (5.25)$$

$$f''(x) = 6a_i(x - x_i) \quad (5.26)$$

Ενώ στα $n + 1$ σημεία (x_i, y_i) οι παράγωγοι δίνονται πολύ απλά ως

$$f'(x_i) = c_i, \quad \text{και} \quad f''(x_i) = 2b_i \quad (5.27)$$

Είναι προφανές ότι η κυβική spline δίνει ικανοποιητικά αποτελέσματα μόνο για τον υπολογισμό παραγώγων πρώτης και δεύτερης τάξης. Ενώ για ανώτερης τάξης παραγώγους απαιτούνται ανώτερης τάξης splines.

Η αριθμητική ολοκλήρωση μιας συνάρτησης $f(x)$ ή μιας n -άδας σημείων (x_i, y_i) μπορεί να προσεγγιστεί με αντικατάσταση της συνάρτησης με το 3ο-βάθμιο πολυώνυμο σε κάθε διάστημα $[x_i, x_{i+1}]$, δηλαδή:

$$\begin{aligned} \int_{x_0}^{x_n} f(x)dx &= \sum_{i=0}^{n-1} \left[\frac{a_i}{4}(x - x_i)^4 + \frac{b_i}{3}(x - x_i)^3 + \frac{c_i}{2}(x - x_i)^2 + d_i(x - x_i) \right]_{x_i}^{x_{i+1}} \\ &= \sum_{i=0}^{n-1} \left[\frac{a_i}{4}(x_{i+1} - x_i)^4 + \frac{b_i}{3}(x_{i+1} - x_i)^3 + \frac{c_i}{2}(x_{i+1} - x_i)^2 + d_i(x_{i+1} - x_i) \right]. \end{aligned}$$

Στην πράξη χρησιμοποιούμε ισαπέχοντα σημεία οπότε αν $h = x_{i+1} - x_i$, τότε η ολοκλήρωση ανάγεται στον υπολογισμό των αθροισμάτων των συντελεστών της κυβικής spline. Δηλαδή,

$$\int_{x_0}^{x_n} f(x)dx = \frac{h^4}{4} \sum_{i=0}^{n-1} a_i + \frac{h^3}{3} \sum_{i=0}^{n-1} b_i + \frac{h^2}{2} \sum_{i=0}^{n-1} c_i + h \sum_{i=0}^{n-1} d_i \quad (5.28)$$

Επομένως, αρκεί ο υπολογισμός των παραπάνω αθροισμάτων για να έχουμε μια εξαιρετικά ακριβή τιμή του ολοκληρώματος.

ΚΡΙΤΙΚΗ

Προφανώς, η μέθοδος παρά την ακρίβειά της δεν είναι η βέλτιστη από άποψη αριθμού πράξεων, και τούτο διότι ο υπολογισμός των συντελεστών των πολυωνύμων των splines είναι χρονοβόρα διαδικασία. Εάν όμως οι συντελεστές a_i , b_i , c_i και d_i έχουν ήδη υπολογισθεί για κάποιον άλλο υπολογισμό, τότε η μέθοδος συνιστάται ανεπιφύλακτα.

ΠΑΡΑΔΕΙΓΜΑ

Να υπολογισθεί το ολοκλήρωμα της συνάρτησης $f(x) = \sin(\pi x)$ στο διάστημα $0 \leq x \leq 1$ με χρήση splines που εφαρμόζεται στα $x = 0, 0.25, 0.5, 0.75, 1.0$.

| i | x | S_i | a_i | b_i | c_i | d_i |
|-----|------|----------|---------|---------|---------|--------|
| 1 | 0 | 0 | -4.8960 | 0 | 3.1340 | 0 |
| 2 | 0.25 | -7.344 | -2.0288 | -3.6720 | 2.2164 | 0.7071 |
| 3 | 0.5 | -10.3872 | 2.0288 | -3.1936 | 0 | 1.0 |
| 4 | 0.75 | -7.344 | 4.8960 | -3.6720 | -2.2164 | 0.7071 |

Χρησιμοποιώντας τη συνθήκη στα άκρα $S_1 = 0$ και $S_5 = 0$ βρίσκουμε τους εξής συντελεστές :

Οπότε το ολοκλήρωμα γίνεται:

$$\int_0^1 \sin(\pi x) dx = \frac{6.251^4}{4} (0) + \frac{(0.25)^3}{3} (-12.5376) + \frac{(0.25)^2}{2} (3.1340) + 0.25 (2.4142) = 0.6362$$

Η ακριβής τιμή είναι 0.6366 οπότε το σφάλμα θα είναι $\varepsilon \approx 0.0004$ ενώ με τη μέθοδο Simpson για τον ίδιο αριθμό σημείων το σφάλμα είναι ~ 0.0015 , δηλαδή 4 φορές μεγαλύτερο.

5.2 ΜΕΘΟΔΟΣ ΠΡΟΣΔΙΟΡΙΣΤΕΩΝ ΣΥΝΤΕΛΕΣΤΩΝ

Γνωρίζουμε από τον Ολοκληρωτικό Λογισμό ότι για μια συνάρτηση $f(x)$ υπάρχει μια κατάλληλη τιμή ξ στο διάστημα ολοκλήρωσης (a, b) ούτως ώστε να ισχύει (θεώρημα μέσης τιμής):

$$\int_a^b f(x) dx = (b - a) f(\xi) \quad (5.29)$$

Αν το γενικεύσουμε, θα μπορούσαμε να προσεγγίσουμε ένα διάστημα με ένα γραμμικό συνδυασμό τιμών της $f(x)$ εντός του διαστήματος (a, b) . Επίσης, θα μπορούσαμε να συνδυάσουμε ακόμη και παραγώγους της $f(x)$ σ' αυτό το διάστημα. Στην πιο απλή περίπτωση, προσεγγίζουμε το ολοκλήρωμα με δυο μόνο τιμές της συνάρτησης, την $f(a)$ και την $f(b)$. Οπότε γράφουμε:

$$I = \int_a^b f(x) dx \equiv c_0 f(a) + c_1 f(b) \quad (5.30)$$

Οι τιμές των c_0, c_1 εξαρτώνται από την $f(x)$ αλλά και από τις τιμές των a και b .

Για να είναι αληθής η παραπάνω σχέση, θα πρέπει η $f(x)$ να είναι είτε σταθερή συνάρτηση είτε γραμμική ως προς x . Οπότε, αν αντικαταστήσουμε μια φορά για $f(x) = 1$ και στη συνέχεια για $f(x) = x$, τα ολοκληρώματα υπολογίζονται ακριβώς και επομένως θα δημιουργήσουμε δύο εξισώσεις για τους δύο αγνώστους c_0 και c_1 . Δηλαδή:

$$f(x) = x \Rightarrow \int_a^b x \cdot dx = \frac{x^2}{2} \Big|_a^b = \frac{1}{2} (b^2 - a^2) \equiv c_0 \cdot a + c_1 \cdot b \quad (5.31)$$

$$f(x) = 1 \Rightarrow \int_a^b 1 \cdot dx = x \Big|_a^b = (b - a) \equiv c_0 \cdot 1 + c_1 \cdot 1 \quad (5.32)$$

Οπότε, εύκολα βρίσκουμε ότι:

$$c_0 = \frac{b-a}{2} \quad \text{και} \quad c_1 = \frac{b-a}{2} \quad (5.33)$$

Αρα

$$\int_a^b f(x) dx \equiv \frac{b-a}{2} [f(a) + f(b)] \quad (5.34)$$

Ουσιαστικά, έχουμε δημιουργήσει ξανά τον «κανόνα του τραpezίου», που δίνει ακριβή αποτελέσματα για συναρτήσεις που παριστάνουν ευθείες.

Αν συνεχίσουμε δημιουργώντας μία σχέση με τρεις όρους παρόμοια με την (5.30), δηλαδή,

$$\int_a^b f(x) dx \equiv c_0 f(a) + c_1 f\left(\frac{a+b}{2}\right) + c_2 f(b) \quad (5.35)$$

τότε η παραπάνω σχέση μπορεί να είναι ακριβής μόνο αν η $f(x)$ ήταν πολυώνυμο το πολύ 2ου βαθμού. Οπότε, θέτω $f(x) = 1$, $f(x) = x$ και $f(x) = x^2$ και ολοκληρώνοντας για την κάθε μια συνάρτηση δημιουργώ ένα σύστημα τριών εξισώσεων για τους τρεις αγνώστους c_0 , c_1 και c_2 , από τη λύση του οποίου καταλήγω στην εξίσωση:

$$\int_a^b f(x) dx = \frac{b-a}{6} \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right] \quad (5.36)$$

που στην ουσία είναι ο κανόνας του Simpson για τρία σημεία.

Θα μπορούσαμε εύκολα να επεκτείνουμε την παραπάνω διαδικασία, προσθέτοντας και τις πρώτες παραγώγους της συνάρτησης σε κάποια σημεία του διαστήματος ολοκλήρωσης, για παράδειγμα:

$$\int_a^b f(x) dx = c_0 f(a) + c_1 f(b) + c_2 f'(a) + c_3 f'(b) \quad (5.37)$$

οπότε θα πρέπει να καταλήξω σε ένα σύστημα με τέσσερις άγνωστες ποσότητες που θα μας οδηγήσει στη σχέση:

$$\int_a^b f(x) dx = \frac{(b-a)}{2} (f(a) + f(b)) + \frac{(b-a)^2}{12} (f'(a) - f'(b)) \quad (5.38)$$

Δηλαδή, θα υπολογίσουμε το ολοκλήρωμα μόνο με βάση τις τιμές της συνάρτησης και της παραγώγου της στα άκρα του διαστήματος $[a, b]$. Η σχέση αυτή γενικεύεται στον τύπο των Euler-Maclaurin (πώς;)

$$\begin{aligned}
\int_{x_0}^{x_n} f(x) dx &= \frac{h}{2} [f(x_0) + 2f(x_1) + \dots + 2f(x_{n-1}) + f(x_n)] \\
&\quad - \frac{h^2}{12} [f'(x_n) - f'(x_0)] \\
&\quad + \frac{h^4}{720} [f^{(3)}(x_n) - f^{(3)}(x_0)] \\
&\quad - \frac{h^6}{30240} [f^{(5)}(x_n) - f^{(5)}(x_0)]
\end{aligned} \tag{5.39}$$

που είναι μια *εξαιρετικά ακριβής* μέθοδος και στην ουσία βελτιώνει το αποτέλεσμα με μόνο 6 επιπλέον υπολογισμούς της 1ης, 3ης και 5ης παραγώγου της συνάρτησης στα άκρα του διαστήματος.

ΠΑΡΑΔΕΙΓΜΑ

Εφαρμόστε τη μέθοδο Euler-Maclaurin για τον υπολογισμό του ολοκληρώματος

$$I = \int_0^{\pi/2} \sin(x) dx$$

Χρησιμοποιώντας τη σχέση (5.39) μόνο για δυο σημεία (της αρχής και του τέλους του διαστήματος ολοκλήρωσης) βρίσκω:

$$\begin{aligned}
\int_0^{\pi/2} \sin(x) dx &= \frac{\pi}{4} \left(\sin 0 + \sin \frac{\pi}{2} \right) \\
&\quad + \frac{\pi^2}{2^2 \cdot 12} \left(\cos 0 - \cos \frac{\pi}{2} \right) \\
&\quad - \frac{\pi^4}{2^4 \cdot 720} \left(\sin 0 - \sin \frac{\pi}{2} \right) \\
&\quad + \frac{\pi^6}{2^6 \cdot 30240} \left(\cos 0 - \cos \frac{\pi}{2} \right) \\
&= \frac{\pi}{4} + \frac{\pi^2}{48} + \frac{\pi^4}{16 \cdot 720} + \frac{\pi^6}{2^6 \cdot 30240} = 0.99996732
\end{aligned}$$

5.2.1 Μέθοδος του Filon

Είναι μια ιδιαίτερα χρήσιμη μέθοδος που βασίζεται στη τεχνική των προσδιοριστέων συντελεστών για ολοκληρώματα της μορφής

$$\int_a^b f(x) \sin(x) dx \quad \text{και} \quad \int_a^b f(x) \cos(x) dx$$

δηλαδή για ολοκληρώματα με περιοδικές συναρτήσεις. Για παράδειγμα θα τους θα εξετάσουμε το παρακάτω ολοκλήρωμα:

$$\int_0^{2\pi} f(x) \sin(x) dx \approx A_1 f(0) + A_2 f(\pi) + A_3 f(2\pi)$$

Η σχέση αυτή θα δίνει ακριβές αποτέλεσμα για $f(x) = 1$, $f(x) = x$ και $f(x) = x^2$ οπότε δημιουργούμε το παρακάτω σύστημα για τους τρεις άγνωστους συντελεστές

$$\begin{aligned} f(x) = 1 &\Rightarrow 0 = A_1 + A_2 + A_3 \\ f(x) = x &\Rightarrow -2\pi = \pi A_2 + 2\pi A_3 \\ f(x) = x^2 &\Rightarrow -4\pi^2 = \pi^2 A_2 + 4\pi^2 A_3 \end{aligned}$$

με λύση $A_1 = 1$, $A_2 = 0$ και $A_3 = -1$. Άρα

$$\int_0^{2\pi} f(x) \sin x dx \approx f(0) - f(2\pi).$$

ΓΕΝΙΚΟΣ ΤΥΠΟΣ

Ο γενικός τύπος για τη μέθοδο ολοκλήρωσης του Filon είναι :

$$\int_a^b y(x) \sin(kx) dx \approx h [Ay(a) \cos(ka) - Ay(b) \cos(kb) + BS_e + DS_o] \quad (5.40)$$

$$\int_a^b y(x) \cos(kx) dx \approx h [Ay(a) \cos(ka) - Ay(b) \cos(kb) + BC_e + DC_o] \quad (5.41)$$

$$A = \frac{1}{q} + \frac{\sin(2q)}{2q^2} - \frac{2\sin^2(q)}{q^3} \quad (5.42)$$

$$B = \frac{1}{q^2} + \frac{\cos^2(q)}{q^2} - \frac{\sin(2q)}{q^3} \quad (5.43)$$

$$D = \frac{4\sin(q)}{q^3} - \frac{4\cos(q)}{q^2} \quad (5.44)$$

$$S_e = -y(a) \sin(ka) - y(b) \sin(kb) + 2 \sum_{i=0}^n y(a + 2ih) \sin(ka + 2iq) \quad (5.45)$$

$$S_o = \sum_{i=1}^n y[a + (2i-1)h] \sin[ka + (2i-1)q] \quad (5.46)$$

$$C_e = -y(a) \sin(ka) - y(b) \sin(kb) + 2 \sum_{i=0}^n y(a + 2ih) \sin(ka + 2iq) \quad (5.47)$$

$$C_o = \sum_{i=1}^n y[a + (2i-1)h] \sin[ka + (2i-1)q] \quad (5.48)$$

όπου $q = kh$.

ΠΑΡΑΔΕΙΓΜΑ

Για το ολοκλήρωμα

$$I = \int_0^{2\pi} e^{-x/2} \cos(100x) dx \quad (5.49)$$

του οποίου η ακριβής τιμή είναι $4.783810813 \times 10^{-5}$, λαμβανουμε τα αποτελέσματα του Πίνακα 5.3. Παρατηρούμε ότι η μέθοδος του Filon δίνει ικανοποιητικό αποτέλεσμα ακόμη και με χρήση μόνο 4 σημείων ενώ η μέθοδος του Simpson απαιτεί τουλάχιστον 1000 σημεία για να επιτύχει την ίδια ακρίβεια.

| n | Simpson | Filon |
|------|----------------|---------------|
| 4 | 1.91733833E+0 | 4.77229440E-5 |
| 8 | -5.73192992E-2 | 4.72338540E-5 |
| 16 | 2.42801799E-2 | 4.72338540E-5 |
| 128 | 5.55127202E-4 | 4.78308678E-5 |
| 256 | -1.30263888E-4 | 4.78404787E-5 |
| 1024 | 4.77161559E-5 | 4.78381120E-5 |
| 2048 | 4.78309107E-5 | 4.78381084E-5 |

Πίνακας 5.3. Σύγκριση των μεθόδων Simpson και Filon για το ολοκλήρωμα που δίνεται από τη σχέση 5.49.

5.3 ΜΕΘΟΔΟΣ GAUSS

Στην μέθοδο των προσδιοριστέων συντελεστών επιλέγαμε συγκεκριμένα σημεία στο διάστημα ολοκλήρωσης $[a, b]$ και στη συνέχεια προσπαθούσαμε να υπολογίσουμε μόνο τους συντελεστές της συνάρτησης $f(x)$ σ' αυτά τα σημεία. Αν τώρα επεκτείνουμε την τεχνική και ζητήσουμε και τον υπολογισμό όχι μόνο των συντελεστών αλλά και συγκεκριμένων σημείων $a \leq x_i \leq b$ εντός του διαστήματος ολοκλήρωσης τότε αναγώμασε σε μια εξαιρετική μέθοδο αριθμητικής ολοκλήρωσης γνωστή ως η *μέθοδος του Gauss*.

Στη συνέχεια θα εξετάσου μια απλή περίπτωση εφαρμογής της μεθόδου. Έστω ότι το διάστημα ολοκλήρωσης είναι το $[-1, 1]$ τότε υποθέτουμε ότι :

$$\int_{-1}^1 f(x) dx \approx af(x_1) + bf(x_2) \quad (5.50)$$

δηλαδή, έχουμε τέσσερις άγνωστες ποσότητες a , b , x_1 και x_2 . Για να τις υπολογίσουμε απαιτείται ένα σύστημα 4 εξισώσεων, οπότε θα απαιτήσουμε η μέθοδός μας να είναι ακριβής για πολυώνυμα έως και τρίτου βαθμού δηλαδή $f(x) = 1$, $f(x) = x$, $f(x) = x^2$ και $f(x) = x^3$. Με αυτό τον τρόπο δημιουργούμε, αντικαθιστώντας κάθε μία από αυτές τις συναρτήσεις στη σχέση (5.50), ένα σύστημα τεσσάρων εξισώσεων με τέσσερις αγνώστους, δηλαδή:

$$\left. \begin{aligned} f(x) = 1 &\Rightarrow 2 = a + b \\ f(x) = x &\Rightarrow 0 = ax_1 + bx_2 \\ f(x) = x^2 &\Rightarrow \frac{2}{3} = ax_1^2 + bx_2^2 \\ f(x) = x^3 &\Rightarrow 0 = ax_1^3 + bx_2^3 \end{aligned} \right\} \Rightarrow \begin{aligned} a &= b = 1 \\ x_1 &= -x_2 = -\left(\frac{1}{3}\right)^{1/2} = -0.5773 \end{aligned}$$

Επομένως, η σχέση (5.50) γράφεται ως:

$$\int_{-1}^1 f(x)dx \approx f(-0.5773) + f(0.5773)$$

Δηλαδή αρκεί ο υπολογισμός οποιασδήποτε συνάρτησης $f(x)$ σε δύο μόνο κατάλληλα σημεία τα $x_1 = -0.5773$ και $x_2 = 0.5773$ για ένα σχετικά ακριβή αριθμητικό υπολογισμό της τιμής του ολοκληρώματος.

ΠΑΡΑΤΗΡΗΣΗ

Αν τα όρια ολοκλήρωσης δεν είναι $[-1, 1]$, κάνουμε την αντικατάσταση

$$t = \frac{1}{2}(b-a)x + \frac{1}{2}(b+a) \quad \text{και} \quad dt = \frac{b-a}{2}dx. \quad (5.51)$$

ΠΑΡΑΔΕΙΓΜΑ

Να υπολογισθεί αριθμητικά η τιμή του ολοκληρώματος:

$$I = \int_0^{\pi/2} \sin(x)dx$$

Και'αρχάς, αλλαζουμε τη μεταβλητή και τα όρια ολοκλήρωσης

$$x = \frac{1}{2}\left(\frac{\pi}{2}t + \frac{\pi}{2}\right) = \frac{\pi}{4}(t+1) \quad \text{και} \quad dx = \frac{\pi}{4}dt$$

οπότε, για τα νέα όρια, το ολοκλήρωμα γράφεται:

$$\begin{aligned} I &= \frac{\pi}{4} \int_{-1}^1 \sin\left[\frac{\pi}{4}(t+1)\right] dt = \frac{\pi}{4} [1.0 \cdot \sin(0.10566 \cdot \pi) + 1.0 \cdot \sin(0.39434 \cdot \pi)] \\ &= 0.99847 \end{aligned}$$

Για σύγκριση αναφέρουμε ότι ο κανόνας τραπεζίου 2 σημείων δίνει: 0.7854 και ο κανόνας Simpson 3 σημείων δίνει: 1.0023.

5.3.1 Μέθοδος Gauss-Legendre

Η μέθοδος Gauss μπορεί να γενικευθεί για αναπτύγματα περισσότερων όρων. Δηλαδή, αν χρησιμοποιήσουμε για n σημεία θα έχουμε μια σχέση της μορφής :

$$\int_{-1}^1 f(x)dx \approx \sum_{i=1}^n A_i f(x_i) \quad (5.52)$$

Με αυτό τον τρόπο θα πρέπει να λυθεί ένα σύστημα $2n$ εξισώσεων για τον υπολογισμό των $2n$ ποσοτήτων A_i και x_i . Η μέθοδος με βάση τα παραπάνω θα είναι ακριβής για πολυώνυμα έως και βαθμού $2(n-1)$, οπότε οι $2n$ εξισώσεις αποδεικνύεται ότι μπορούν να γραφούν συνοπτικά ως :

$$A_1 x_1^k + \dots + A_n x_n^k = \begin{cases} 0 & \text{για } k = 1, 3, 5, \dots, 2n-1 \\ \frac{2}{k+1} & \text{για } k = 2, 4, 6, \dots, 2n-2 \end{cases} \quad (5.53)$$

Αποδεικνύεται ότι τα x_i είναι ρίζες των πολυωνύμων Legendre βαθμού n των οποίων οι ρίζες βρίσκονται πάντα εντός του διαστήματος $(-1, 1)$. Τα πολυώνυμα Legendre δημιουργούνται από την αναδρομική σχέση:

$$(n+1)L_{n+1}(x) - (2n+1)xL_n(x) + nL_{n-1}(x) = 0 \quad (5.54)$$

όπου τα 3 πρώτα πολυώνυμα είναι:

$$L_0(x) = 1, \quad L_1(x) = x \quad \text{και} \quad L_2(x) = \frac{3}{2}x^2 - \frac{1}{2} \quad (5.55)$$

Κατ' αναλογία τα A_i δίνονται από τις σχέσεις:

$$A_i = \frac{2(1-x^2)}{n^2 [L_{n-1}(x_i)]^2} \quad (5.56)$$

Για παράδειγμα αν $n = 4$ θα πρέπει να βρούμε τις ρίζες του 4ο-βάθμιου πολυωνύμου Legendre

$$P_4 = \frac{1}{8}(35x^4 - 30x^2 + 3)$$

που είναι $x_i = \pm [(15 \pm 2\sqrt{30})/35]^{1/2}$ και στη συνέχεια χρησιμοποιώντας τη σχέση (5.56) υπολογίζω τα A_i . Οι ακριβείς τιμές τους δίνονται στον Πίνακα 5.4.

Επομένως, αρκεί να υπολογισθούν μια φορά τα x_i και A_i για διάφορους αριθμούς σημείων και στη συνέχεια χρησιμοποιούνται σε κάθε πρόβλημα.

ΠΑΡΑΔΕΙΓΜΑ

Θα εφαρμόσουμε τη μέθοδο Gauss για 4 σημεία:

$$\begin{aligned} I &= \int_0^{\pi/2} \sin(x)dx = \frac{\pi}{4} \int_{-1}^1 \sin\left[\frac{\pi}{4}(t+1)\right] dt \\ &= \frac{\pi}{4} \left[0.34785 \cdot \sin\left[\frac{\pi}{4} \cdot (1 + 861136)\right] + 0.34785 \cdot \sin\left[\frac{\pi}{4}(1 - 0.861136)\right] \right. \\ &\quad \left. + 0.652145 \cdot \sin\left[\frac{\pi}{4}(1 + 0.33998)\right] + 0.652145 \cdot \sin\left[\frac{\pi}{4}(1 - 0.33998)\right] \right] \\ &= 1.000000... \end{aligned}$$

Αντίστοιχα, η μέθοδος Simpson με 32 σημεία σημεία δίνει: 1.0000003 και για 64 σημεία: 9.99999983.

| n | x_i | A_i |
|-----|------------------|------------|
| 2 | ± 0.57735027 | 1.0000000 |
| 4 | ± 0.86113631 | 0.34785485 |
| | ± 0.33948104 | 0.62214515 |
| 8 | ± 0.96028986 | 0.10122854 |
| | ± 0.79666648 | 0.22381034 |
| | ± 0.52553241 | 0.31370665 |
| | ± 0.18343464 | 0.36268378 |

Πίνακας 5.4. Οι τιμές των x_i και A_i της μεθόδου Gauss-Legendre για 2, 4 και 8 σημεία.

5.3.2 Γενίκευση της μεθόδου Gauss

Αν ένα ολοκλήρωμα είναι της μορφής :

$$I = \int_a^b w(x)y(x)dx \quad (5.57)$$

όπου $w(x)$ είναι μια συνάρτηση βάρους, τότε μπορούμε να επεκτείνουμε τη μέθοδο Gauss χρησιμοποιώντας όχι μόνο τα πολυώνυμα Legendre αλλά και άλλου είδους ορθογώνια πολυώνυμα ανάλογα με τη μορφή της υπο ολοκλήρωση συνάρτησης. Επομένως το παραπάνω ολοκλήρωμα θα γραφεί ως :

$$\int_a^b w(x)y(x)dx = \sum_{i=1}^n A_i y(x_i) \quad (5.58)$$

όπου τα x_i και τα A_i μπορούν να βρεθούν σε μορφή πινάκων.¹ Αναλόγως με τη μορφή της συνάρτησης βάρους $w(x)$ θα έχουμε τις παρακάτω επιλογές.

- **Μέθοδος Gauss-Legendre** για συνάρτηση βάρους $w(x) = 1$ που είναι η μέθοδος της προηγούμενης ενότητας.
- **Μέθοδος Gauss-Laguerre** για ολοκληρώματα της μορφής

$$\int_0^{\infty} e^{-x} y(x) dx \approx \sum_{i=1}^n A_i y(x_i) \quad (5.59)$$

προφανώς η συνάρτηση βάρους είναι $w(x) = e^{-x}$ και τα x_i είναι ρίζες των πολυωνύμων Laguerre που δημιουργούνται από την παρακάτω αναδρομική σχέση :

$$L_n(x) = e^x \frac{d^n}{dx^n} (e^{-x} x^n) \quad (5.60)$$

ενώ οι συντελεστές A_i υπολογίζονται από τη σχέση :

¹ Μια μεγάλη συλλογή καταλλήλων πινάκων για τις μεθόδους αυτού του κεφαλαίου μπορεί να βρεθεί στο Abramowitz-Stegun.

$$A_i = \frac{(n!)^2}{x_i [L'_n(x_i)]^2} \quad (5.61)$$

και συνήθως δίνονται σε πίνακες, (βλέπε Πίνακα 5.5).

| n | x_i | A_i |
|-----|-------------|------------|
| 2 | 0.58578644 | 0.85355339 |
| | 3.41421356 | 0.14644661 |
| 4 | 0.32254769 | 0.60315410 |
| | 1.74576110 | 0.35741869 |
| | 4.53662030 | 0.03888791 |
| | 9.39507091 | 0.00053929 |
| 6 | 0.22284660 | 0.10122854 |
| | 1.18893210 | 0.41700083 |
| | 2.99273633 | 0.11337338 |
| | 5.77514357 | 0.01039920 |
| | 9.83746742 | 0.00026102 |
| | 15.98287398 | 0.00000090 |

Πίνακας 5.5. Οι τιμές των x_i και A_i της μεθόδου Gauss-Laguerre για 2, 4 και 6 σημεία.

- **Μέθοδος Gauss-Hermite** για ολοκληρώματα της μορφής

$$\int_{-\infty}^{\infty} e^{-x^2} y(x) dx \approx \sum_{i=1}^n A_i y(x_i) \quad (5.62)$$

Δηλαδή, η συνάρτηση βάρους είναι $w(x) = e^{-x^2}$ ενώ τα x_i είναι οι ρίζες των πολυωνύμων Hermite. Τα πολυώνυμα Hermite μπορούν να δημιουργηθούν από τη σχέση

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} (e^{-x^2}) \quad (5.63)$$

ενώ οι συντελεστές A_i υπολογίζονται από τη σχέση:

$$A_i = \frac{2^{n+1} n! \sqrt{\pi}}{[H'_n(x_i)]^2} \quad (5.64)$$

Τα x_i , A_i δίνονται σε πίνακες, (βλέπε Πίνακα 5.6).

- **Μέθοδος Gauss-Chebyshev** για ολοκληρώματα της μορφής

$$\int_{-1}^1 \frac{y(x)}{\sqrt{1-x^2}} dx \approx \frac{\pi}{n} \sum_{i=1}^n y(x_i) \quad (5.65)$$

δηλαδή η συνάρτηση βάρους είναι

$$w(x) = \frac{1}{\sqrt{1-x^2}} \quad (5.66)$$

| n | x_i | A_i |
|-----|--|--|
| 2 | ± 0.70710678 | 0.88622693 |
| 4 | ± 0.52464762 ± 1.65068012 | 0.80491409 0.08131284 |
| 6 | ± 0.43607741 ± 1.33584907 ± 2.35060497 | 0.72462960 0.15706732 0.00453001 |

Πίνακας 5.6. Οι τιμές των x_i και A_i της μεθόδου Gauss-Hermite για 2, 4 και 6 σημεία.

ενώ τα x_i είναι ρίζες των πολυωνύμων Chebyshev

$$T_n(x) = \cos[n \arccos(x)] \quad (5.67)$$

που δίνονται από τις σχέσεις

$$x_i = \cos \left[\frac{\pi}{2n} (2i - 1) \right]. \quad (5.68)$$

Τέλος οι συντελεστές A_i δίνονται από μία απλή έκφραση

$$A_i = \frac{\pi}{n} \quad (5.69)$$

όπου n είναι ο βαθμός του πολυωνύμου που χρησιμοποιούμε.

5.4 ΑΣΚΗΣΕΙΣ

1. Να αποδειχθεί ο τύπος του σφάλματος για τη μέθοδο Simpson 3/8 (σχέση 5.12).
2. Με τη μέθοδο του τραpezίου ολοκληρώστε τις συναρτήσεις \sqrt{x} , $\log(x)$, και $1/x$ με βήμα $h = 0.1$ και $h = 0.05$ για το διάστημα $[1, 1.3]$. Συγκρίνετε με τα ακριβή αποτελέσματα και επίσης πόσο μεταβάλλεται η ακρίβεια με τον υποδιπλασιασμό του βήματος.
3. Στο προηγούμενο πρόβλημα εφαρμόστε τη μέθοδο Romberg και εξετάστε την βελτίωση των αποτελεσμάτων.
4. Επαναλάβετε, την άσκηση 1 για τη μέθοδο του Simpson και συγκρίνετε την ακρίβειά της με αυτή της μεθόδου του τραpezίου.
5. Υπολογίστε την τιμή του ολοκληρώματος $\int_0^{\pi/2} \sin x dx$ διαιρώντας το διάστημα ολοκλήρωσης σε 6 υποδιαστήματα, χρησιμοποιώντας τον κανόνα του τραpezίου και στη συνέχεια τον κανόνα του Simpson. Συγκρίνετε το αποτέλεσμα σας με το ακριβές και εξετάστε αν το σφάλμα είναι το αναμενόμενο.
6. Εφαρμόστε τη μέθοδο Romberg στο παραπάνω πρόβλημα.
7. Εάν θέλαμε να υπολογίσουμε την τιμή του ολοκληρώματος $\int_{1.8}^{2.4} dx$ χρησιμοποιώντας τον κανόνα του τραpezίου, και απαιτούσαμε ακρίβεια 5 δεκαδικών ψηφίων, πόσο μικρό έπρεπε να είναι το βήμα h ; Υπολογίστε το θεωρητικά και στη συνέχεια εξετάστε αν συμφωνεί με το αριθμητικό αποτέλεσμα.

8. Επαναλάβετε το προηγούμενο πρόβλημα χρησιμοποιώντας τη μέθοδο του Simpson.
9. Υπολογίστε το ελλειπτικό ολοκλήρωμα $\int_0^{\pi/2} \sqrt{1 - k \sin^2(x)} dx$ για $k = 0.5$ και $k = 0.25$ με ακρίβεια 4 δεκαδικά ψηφίων.
10. Υπολογίστε το μήκος της περιφέρειας της έλλειψης $x^2/16 + y^2/25 = 1$ με ακρίβεια 6 δεκαδικών ψηφίων.
11. Με τη μέθοδο των προσδιοριστέων συντελεστών υπολογίστε τους συντελεστές στον τύπο

$$\int_{-h}^h y(x) dx = h [a_{-1}y(-h) + a_0y(0) + a_1y(h)] + h^2 [b_{-1}y'(-h) + b_0y'(0) + b_1y'(h)] \quad (5.70)$$

έτσι ώστε να είναι ακριβής για όσο το δυνατόν μεγαλύτερου βαθμού πολυώνυμο.

12. Με τη μέθοδο των προσδιοριστέων συντελεστών δείξτε ότι ισχύει

$$\int_0^h y(x) dx = \frac{h}{2} (y_0 + y_1) - \frac{h^3}{24} (y_0'' + y_1'') \quad (5.71)$$

για κάθε πολυώνυμο έως και 4ου βαθμού.

13. Υπολογίστε με τη μέθοδο του Filon το ολοκλήρωμα $\int_0^{2\pi} e^{-x} \sin(10x) dx$. Συγκρίνετε την ακρίβεια της μεθόδου με την ακρίβεια της μεθόδου του Simpson στον ίδιο αριθμό σημείων [Η αναλυτική επίλυση δίνει $(1 - e^{-2\pi})/101$].
14. Εφαρμόστε τη μέθοδο Gauss-Legendre και κατόπιν τη μέθοδο Simpson-Romberg για τον υπολογισμό του ολοκληρώματος $\int_0^{\pi/2} \log(1+x) dx$ (η ακριβής τιμή είναι 0.856589940). Τι παρατηρείτε ;

ΔΙΑΦΟΡΙΚΕΣ ΕΞΙΣΩΣΕΙΣ

Οι περισσότεροι νόμοι σε διάφορες επιστήμες γράφονται με τη μορφή διαφορικών εξισώσεων. Οι κανονικές διαφορικές εξισώσεις (ΔΕ) αποτελούν σημαντικό κεφάλαιο των μαθηματικών και υπάρχει πλειάδα μεθόδων για την αναλυτική επίλυση τους. Εν τούτοις, στην πράξη, ένας μεγάλος αριθμός από χρήσιμες διαφορικές εξισώσεις δεν επιλύονται αναλυτικά και απαιτείται η αριθμητική τους επίλυση. Τις περισσότερες φορές ακόμη και πολύ απλά φυσικά μοντέλα, που χρησιμοποιούνται για την πρόβλεψη της εξέλιξης ενός φαινομένου, δεν επιλύονται αναλυτικά.

Ένα απλό παράδειγμα για τη σχέση διαφορικών εξισώσεων με φυσικά φαινόμενα αποτελεί το μοντέλο 'κυνηγού-θηράματος'. Έστω ότι $x(t)$ είναι ο αριθμός των κυνηγών και $y(t)$ ο αριθμός των θηραμάτων σε μια χρονική στιγμή. Τότε, κάτω από κατάλληλες (και απλουστευτικές) συνθήκες, η εξέλιξη των δύο πληθυσμών θα δίνεται από το σύστημα των διαφορικών εξισώσεων

$$\begin{aligned}\frac{dx}{dt} &= a(y + b)x \\ \frac{dy}{dt} &= c(x + d)y\end{aligned}$$

όπου τα a, b, c, d είναι σταθερές, που δίνονται από το μελετητή του συστήματος.

Η αριθμητική επίλυση προβλημάτων αυτού του είδους θα αποτελέσει το αντικείμενο αυτού του κεφαλαίου. Θα αναπτύξουμε μεθόδους για την αριθμητική επίλυση διαφορικών εξισώσεων και θα αναφέρουμε τα πλεονεκτήματα και τα μειονεκτήματα της κάθε μεθόδου. Στη συνέχεια θα δείξουμε μέσω παραδειγμάτων πως με τη χρήση της Mathematica είναι δυνατή η αριθμητική αλλά και αναλυτική επίλυση διαφορικών εξισώσεων.

6.1 ΜΕΘΟΔΟΙ ΕΝΟΣ ΒΗΜΑΤΟΣ

Όταν για τον υπολογισμό της τιμής μιας συνάρτησης που αποτελεί λύση μιας διαφορικής εξίσωσης στη θέση $x_i + h$ (όπου h είναι ένα μικρό βήμα) χρησιμο-

ποιούμε μόνο την πληροφορία για την τιμή της συνάρτησης στην προηγούμενη θέση x_i τότε οι τεχνικές για την αριθμητική επίλυση των διαφορικών εξισώσεων κατατάσσονται ως **μέθοδοι ενός βήματος**

6.1.1 Μέθοδος Σειρών Taylor

Η πλέον προφανής μέθοδος για την επίλυση μιας διαφορικής εξίσωσης πρώτης τάξης είναι το ανάπτυγμα Taylor.

Έστω, λοιπόν, μια διαφορική εξίσωση της μορφής

$$y' = f(x, y) \quad \text{με} \quad y(x_0) = y_0 \quad (6.1)$$

τότε, για να υπολογίσουμε την τιμή της $y(x)$ σε μια θέση $x = x_0 + h$, θα χρησιμοποιήσουμε το ανάπτυγμα Taylor, δηλαδή:

$$y(x_0 + h) = y(x_0) + hy'(x_0) + \frac{h^2}{2}y''(x_0) + \dots \quad (6.2)$$

Επειδή όμως γνωρίζουμε την $y' = f(x, y)$, μπορούμε εύκολα να υπολογίσουμε τις παραγώγους της $y(x)$ οποιασδήποτε τάξης:

$$y'' = f'(x, y), \quad y''' = f''(x, y), \text{ κ.ο.κ.} \quad (6.3)$$

άρα ο υπολογισμός της τιμής της $y(x_0 + h)$ είναι μια απλή διαδικασία.

ΠΑΡΑΔΕΙΓΜΑ

Έστω λοιπόν η διαφορική εξίσωση:

$$y' = x + y \quad (6.4)$$

με αρχικές τιμές $y(0) = 1$. Να βρεθεί η τιμή της $y(1)$ (η ακριβής λύση δίνεται για σύγκριση και είναι: $y = 2e^x - x - 1$).

Για την επίλυσή της βρίσκω τις παραγώγους έως και 4ης τάξης:

$$\begin{aligned} y'(x_0) &= y'(0) = 0 + 1 = 1 \\ y'' &= 1 + y' = 1 + x + y = 1 + y' \\ y''(x_0) &= 1 + y'(0) = 2 \\ y''' &= 1 + y' \\ y'''(x_0) &= 1 + y'(0) = 2 \\ y^{(4)} &= 1 + y' \\ y^{(4)}(x_0) &= 1 + y'(0) = 2 \end{aligned}$$

Οπότε, χρησιμοποιώντας την εξίσωση (6.2) βρίσκω:

$$y(0+h) = 1 + h + h^2 + \frac{h^3}{3} + \frac{h^4}{12} + \dots \quad (6.5)$$

Το σφάλμα είναι προφανώς αυτό που προβλέπεται από το ανάπτυγμα Taylor, δηλαδή:

$$E = \frac{y^{(5)}(\xi)}{5!} h^5 \quad \text{για } 0 < \xi < h \quad (6.6)$$

| x | y | y Αναλυτικό | Σφάλμα | Σφάλμα* |
|-----|----------|---------------|----------------------|----------------------|
| 0 | 1 | 1 | | |
| 0.1 | 1.110342 | 1.110342 | 1.7×10^{-7} | |
| 0.2 | 1.24280 | 1.24281 | 5.5×10^{-6} | 3.7×10^{-7} |
| 0.3 | 1.39968 | 1.39972 | 4.3×10^{-5} | 6.2×10^{-7} |
| 0.4 | 1.383467 | 1.383649 | 1.8×10^{-4} | 9.1×10^{-7} |
| : | | | | |
| 1.0 | 3.416667 | 3.436564 | 2×10^{-2} | 4.2×10^{-6} |

Πίνακας 6.1. Στον παραπάνω πίνακα, οι τιμές της δεύτερης στήλης βρίσκονται, αν χρησιμοποιήσουμε βήμα $h = 0.1$ την πρώτη φορά, $h = 0.2$ τη δεύτερη φορά, κοκ. Παρατηρούμε επομένως ότι το σφάλμα αυξάνεται, καθώς αυξάνεται το h (4η στήλη). Αντίθετα, αν χρησιμοποιήσουμε σε κάθε βήμα τις τιμές που έχουμε υπολογίσει στο προηγούμενο βήμα, τότε η ακρίβεια παραμένει αρκετά καλή (5η στήλη).

6.1.2 Μέθοδοι Euler & Euler - Heun

Η **μέθοδος Euler** αποτελεί ουσιαστικά περιορισμένη εφαρμογή της μεθόδου Taylor. Διατηρούμε όρους μόνο μέχρι 1ης τάξης ως προς h . Είναι:

$$y(x_0 + h) = y(x_0) + hy'(x_0) \quad (6.7)$$

με προφανές σφάλμα:

$$E = \frac{y''(\xi)}{2} h^2 \quad \text{για } 0 < \xi < h \quad (6.8)$$

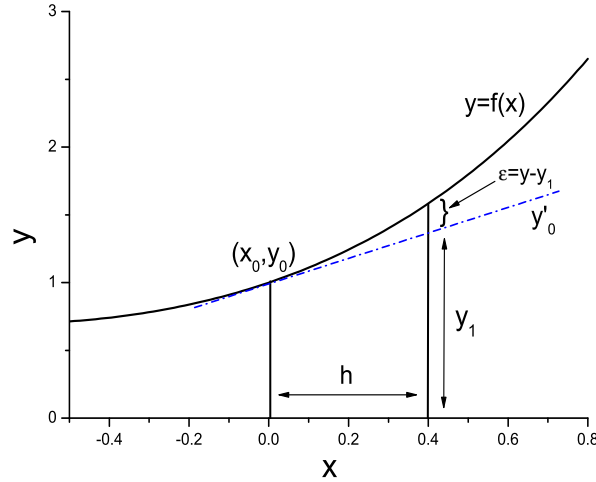
Αν πρόκειται να τη χρησιμοποιήσουμε για λίγα μόνο βήματα, είναι αρκετά καλή και απλή στη χρήση. Για τη διαφορική εξίσωση στο προηγούμενο παράδειγμα έχουμε:

$$y(x_0 + h) = y(x_0) + h(x_0 + y(x_0)) \quad (6.9)$$

ή καλύτερα, σε μορφή αναδρομικής σχέσης:

$$y_{n+1} = y_n + h(x_n + y_n) \quad (6.10)$$

Η **μέθοδος Euler-Heun** αποτελεί μια απλή περίπτωση μεθόδου **πρόδωξης - διόρθωσης** (κατηγορία μεθόδων που θα μελετήσουμε στη συνέχεια).



Σχήμα 6.1. Γεωμετρική απεικόνιση της μεθόδου Euler για την αριθμητική επίλυση της διαφορικής εξίσωσης $y' = x + y$ με $y_0 = y(0) = 1$ και βήμα $h = 0.4$. Σε αυτό το βήμα υπολογίζεται η αριθμητική τιμή της $y = f(x)$, το $y(x_0 + h) = y_1$, στη θέση $x_1 = x_0 + h = 0 + 0.4 = 0.4$. Προφανώς υπάρχει ένα τοπικό σφάλμα αποκοπής $\epsilon = y - y_1$. Ένα ανάλογο σφάλμα απαντάται σε κάθε βήμα, ενώ επιπλέον θα πρέπει να λάβουμε υπόψη μας πως για τον υπολογισμό του y_2 θα χρησιμοποιηθούν τα “λανθασμένα δεδομένα” του προηγούμενου βήματος οπότε δεν αρκεί μόνο η εκτίμηση του τοπικού σφάλματος αλλά απαιτείται έλεγχος του σφάλματος στο σύνολο των βημάτων.

Ουσιαστικά, χρησιμοποιούμε τη μέθοδο Euler (6.7), για να υπολογίσουμε αρχικά την τιμή της $y(x)$ στη θέση $y(x_0 + h)$, δηλαδή την τιμή y_1 , και, στη συνέχεια, υπολογίζουμε την τιμή της y_1 αντικαθιστώντας την $y'(x_0)$ από το μέσο όρο των τιμών της παραγώγου στα σημεία x_0 και $x_0 + h$. Δηλαδή, χρησιμοποιώ για κάθε βήμα τις δύο σχέσεις :

$$y_{n+1} = y_n + hy'_n + O(h^2) \quad (6.11)$$

$$y_{n+1} = y_n + \frac{h}{2} (y'_n + y'_{n+1}) + O(h^3) \quad (6.12)$$

Η μέθοδος Euler-Heun (τοπικό σφάλμα $\sim h^3$) είναι προφανώς ακριβέστερη της απλής μεθόδου Euler που έχει τοπικό σφάλμα $\sim h^2$ (γιατί).

6.1.3 Σφάλματα και Ευστάθεια

Υπάρχει μια πληθώρα σφαλμάτων που υπεισέρχονται στη διαδικασία αριθμητικής επίλυσης μιας διαφορικής εξίσωσης. Ενδεικτικά, θα αναφέρουμε μερικά.

1. Αρχικά σφάλματα δεδομένων

Εάν οι αρχικές τιμές δεν είναι ακριβώς γνωστές, π.χ. τις γνωρίζουμε με συγκεκριμένη ακρίβεια, τότε το αρχικό σφάλμα διαδίδεται σε όλα τα βήματα της ολοκλήρωσης. Έτσι, για παράδειγμα, είναι δυνατό να παρατηρήσουμε το χάος σε διάφορα δυναμικά συστήματα.

2. Σφάλματα Στρογγυλοποίησης

- Από τη συγκράτηση μόνο συγκεκριμένου αριθμού δεκαδικών ψηφίων
- Βεβαιώνεται με την αύξηση των δεκαδικών
- Είναι εξαιρετικά σημαντικό, όταν αφαιρούνται δυο λύσεις της ίδιας περιόδου τάξης

3. Σφάλμα αποκοπής

- Από τη χρήση σειρών συγκεκριμένης τάξης
- Διορθώνεται μικραίνοντας το h .

ΔΙΑΔΟΣΗ ΣΦΑΛΜΑΤΩΝ & ΕΥΣΤΑΘΕΙΑ

Τα προαναφερθέντα σφάλματα διαδίδονται από βήμα σε βήμα. Έστω για παράδειγμα μια τιμή y_n που είναι η αληθής και Y_n αυτή που υπολογίζουμε με κάποια από τις μεθόδους μας. Δηλαδή, το σφάλμα είναι:

$$\varepsilon_n = y_n - Y_n \quad (6.13)$$

οπότε, για παράδειγμα, ας εξετάσουμε τη διάδοση του σφάλματος στη μέθοδο Euler. Θα είναι:

$$\begin{aligned} \varepsilon_{n+1} &= y_{n+1} - Y_{n+1} = y_n + hf(x_n, y_n) - Y_n - hf(x_n, Y_n) \\ &= \varepsilon_n + h \frac{(f(x_n, y_n) - f(x_n, Y_n))}{y_n - Y_n} \varepsilon_n = \varepsilon_n (1 + hf_y(x_n, y_n)) \\ &= (1 + hk) \varepsilon_n \quad \text{όπου} \quad k = \frac{\partial f}{\partial y} \end{aligned} \quad (6.14)$$

δηλαδή η διάδοση του σφάλματος είναι γραμμική. Μια επιπλέον προφανής παρατήρηση από την παραπάνω σχέση είναι ότι αν $|1 + hk| \geq 1$ τότε το σφάλμα θα μεγαλώνει καθώς θα αυξάνεται ο αριθμός των βημάτων και σ αυτή την περίπτωση η μέθοδος του Euler είναι **ασταθής**. Αντίθετα, για $|1 + hk| < 1$ το σφάλμα ελαττώνεται διαρκώς και αποσβένεται και η μέθοδος είναι **απολύτως ευстаθής**. Το κριτήριο αυτό οδηγεί στην αναγκαία σχέση για απόλυτη ευστάθεια που είναι:

$$-2 < hk < 0 \quad (6.15)$$

Από αυτή την ανισότητα συνάγεται ότι για απόλυτη ευστάθεια απαιτείται $\partial f / \partial y < 0$ επειδή το h είναι πάντα θετικό.

ΣΥΓΚΛΙΣΗ

Ας υποθέσουμε τώρα ότι η υπο μελέτη διαφορική εξίσωση είναι της μορφής

$$\frac{dy}{dx} = Ay \quad (6.16)$$

με προφανή λύση $y = e^{Ax}$. Η μέθοδος του Euler δίνει τις προσεγγιστικές λύσεις μέσω μιας αναδρομικής σχέσης της μορφής :

$$y_{n+1} = y_n + hAy_n = (1 + hA)y_n \quad \text{για } n = 0, 1, 2, \dots \quad (6.17)$$

Οπότε αν κάνουμε επαναληπτική χρήση της σχέσης αυτής καταλήγουμε

$$y_{n+1} = (1 + hA)^{n+1}y_0 \quad \text{για } n = 0, 1, 2, \dots \quad (6.18)$$

Για μικρά h ισχύει ότι $1 + hA \approx e^{hA}$ άρα

$$y_{n+1} = (1 + hA)^{n+1}y_0 \approx e^{(n+1)hA}y_0 = e^{(x_{n+1}-x_0)A}y_0 = e^{Ax_{n+1}} \quad (6.19)$$

υπενθυμίζουμε ότι $h = (x_{n+1} - x_0)/(n + 1)$. Δηλαδή η αριθμητική τιμή για μικρά h προσεγγίζει ή καλύτερα **συγκλίνει** στην ακριβή λύση $y = e^{Ax}$.

6.1.4 Μέθοδος Runge-Kutta

Η μέθοδος Runge - Kutta είναι μια κλασική μέθοδος με πολύ καλή ακρίβεια και χρησιμοποιείται ευρύτατα.

Με τη μέθοδο Runge - Kutta προσπαθούμε να αντικαταστήσουμε τις παραγώγους ανώτερης τάξης που εμφανίζονται στη μέθοδο Taylor με κατάλληλους συνδυασμούς των x_i, y_i, y'_i , τα οποία είναι γνωστά. Έτσι, αποφεύγουμε τον υπολογισμό των παραγώγων ανώτερης τάξης.

Στη συνέχεια, αποδεικνύουμε τους τύπους για τη μέθοδο Runge-Kutta δεύτερης τάξης.

Έστω μια διαφορική εξίσωση της μορφής:

$$\frac{dy}{dx} = f(x, y) \quad (6.20)$$

Μια αναδρομική σχέση για τον υπολογισμό του y_n είναι

$$y_{n+1} = y_n + ak_1 + bk_2 \quad (6.21)$$

όπου

$$k_1 = hf(x_n, y_n) \quad (6.22)$$

$$k_2 = hf(x_n + Ah, y_n + Bk_1) \quad (6.23)$$

Ας προσπαθήσουμε να την ταυτοποιήσουμε με μια σειρά Taylor δεύτερης τάξης:

$$y_{n+1} = y_n + hf(x_n, y_n) + \frac{h^2}{2}f'(x_n, y_n) + \dots \quad (6.24)$$

Επειδή όμως

$$\frac{df}{dx} = f_x + f_y \frac{dy}{dx} = f_x + f_y f \quad (6.25)$$

θα είναι:

$$y_{n+1} = y_n + hf_n + \frac{h^2}{2} (f_x + f f_y)_{x=x_n} \quad (6.26)$$

Συγκρίνοντας με την εξίσωση (6.21), που γράφεται:

$$y_{n+1} = y_n + ahf(x_n, y_n) + bhf[x_n + Ah, y_n + Bhf(x_n, y_n)] \quad (6.27)$$

την οποία αναπτύσσω σε σειρά Taylor¹

$$y_{n+1} = y_n + h(a+b)f_n + h^2(Abf_x + Bbf_y)_n \quad (6.28)$$

καταλήγω στις παρακάτω σχέσεις για τις σταθερές a, b, A και B :

$$a+b=1, \quad A \cdot b = \frac{1}{2} \quad \text{και} \quad B \cdot b = \frac{1}{2} \quad (6.29)$$

μπορώ να θέσω αυθαίρετα: π.χ. $a = \frac{2}{3}, \frac{1}{2}, \frac{5}{4} \dots$ και να υπολογίσω τις υπόλοιπες σταθερές.

Αν επιλέξουμε $a = \frac{1}{2}$ τότε θα είναι $b = \frac{1}{2}$ και $A = B = 1$, δηλαδή:

$$y_{n+1} = y_n + \frac{1}{2}(k_1 + k_2) \quad (6.30)$$

όπου

$$k_1 = hf(x_n, y_n), \quad k_2 = hf(x_n + h, y_n + k_1) \quad (6.31)$$

Η σχέση που καταλήξαμε είναι ουσιαστικά η μέθοδος Euler-Heun.

Αν επαναλάβω την ίδια διαδικασία για ανάπτυγμα Taylor έως και h^4 , θα καταλήξω σε ένα σύστημα 11 εξισώσεων με 13 αγνώστους. Με κατάλληλη επιλογή, των δύο από τις άγνωστες ποσότητες, μπορώ να καταλήξω στη μορφή:

$$y_{n+1} = y_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \quad (6.32)$$

όπου

$$k_1 = hf(x_n, y_n) \quad (6.33)$$

$$k_2 = hf\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}k_1\right) \quad (6.34)$$

$$k_3 = hf\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}k_2\right) \quad (6.35)$$

$$k_4 = hf(x_n + h, y_n + k_3) \quad (6.36)$$

Η μέθοδος αυτή είναι γνωστή ως **Runge - Kutta τέταρτης τάξης** με τοπικό σφάλμα: $E \approx 0(h^5)$ και γενικό σφάλμα μετά από n βήματα $E \approx 0(h^4)$. Η

¹ Εδώ έγινε χρήση της: $f(x_n + Ah, y_n + Bhf(x_n, y_n)) \approx (f + f_x Ah + f_y Bhf)_{x=x_n}$

Runge-Kutta τέταρτης τάξης χρησιμοποιείται ευρύτατα αλλά υπάρχουν και ανώτερης τάξης μέθοδοι Runge-Kutta, όπως, για παράδειγμα, η **μέθοδος Runge-Kutta-Fehlberg**, που συνίστουμε ανεπιφύλακτα.

Η μέθοδος Runge-Kutta-Fehlberg δίνεται από τις παρακάτω σχέσεις:

$$k_1 = hf(x_n, y_n) \quad (6.37)$$

$$k_2 = hf\left(x_n + \frac{1}{4}h, y_n + \frac{1}{4}k_1\right) \quad (6.38)$$

$$k_3 = hf\left(x_n + \frac{3}{8}h, y_n + \frac{3}{32}k_1 + \frac{9}{32}k_2\right) \quad (6.39)$$

$$k_4 = hf\left(x_n + \frac{12}{13}h, y_n + \frac{1932}{2197}k_1 - \frac{7200}{2197}k_2 + \frac{7296}{2197}k_3\right) \quad (6.40)$$

$$k_5 = hf\left(x_n + h, y_n + \frac{439}{216}k_1 - 8k_2 + \frac{3680}{513}k_3 - \frac{845}{4104}k_4\right) \quad (6.41)$$

$$k_6 = hf\left(x_n + \frac{1}{2}h, y_n - \frac{8}{27}k_1 + 2k_2 - \frac{3544}{2565}k_3 + \frac{1859}{4104}k_4 - \frac{11}{40}k_5\right) \quad (6.42)$$

οπότε εύκολα υπολογίζουμε μια πρώτη εκτίμηση για την τιμή της y_{n+1}

$$\bar{y}_{n+1} = y_n + \left(\frac{25}{216}k_1 + \frac{1408}{2565}k_3 + \frac{2197}{4104}k_4 - \frac{1}{5}k_5\right) \quad (6.43)$$

Το *τοπικό σφάλμα* είναι τάξης $\approx h^5$. Παρατηρούμε ότι δεν έχει χρησιμοποιηθεί η τιμή k_6 , την οποία όμως θα χρησιμοποιήσουμε στη συνέχεια για τον υπολογισμό ακριβέστερης τιμής με τη σχέση

$$y_{n+1} = y_n + \left(\frac{16}{135}k_1 + \frac{6656}{12825}k_3 + \frac{28561}{56430}k_4 - \frac{9}{50}k_5 + \frac{2}{55}k_6\right) \quad (6.44)$$

με *τοπικό σφάλμα* $\approx h^6$ και γενικό $\approx h^5$.

Παρόλο που το \bar{y}_{n+1} της σχέσης (6.43) δεν έχει σχέση με το y_{n+1} της σχέσης (6.44), εν τούτοις αποτελεί σημαντικό στοιχείο της μεθόδου. Πιο συγκεκριμένα, είναι σύνηθες στην αριθμητική επίλυση των διαφορικών εξισώσεων να επαναλαμβάνουμε την ίδια διαδικασία δυο φορές, μια για βήμα h , και μια για βήμα $h/2$. Αν η μέθοδος έχει σφάλμα $\approx h^n$ την πρώτη φορά, τότε τη δεύτερη φορά το σφάλμα θα είναι $\approx h^n/2^n$. Επομένως, αν το δεύτερο αποτέλεσμα έχει σφάλμα μικρότερο κατά $1/2^n$, τότε έχουμε την πεποίθηση ότι η μέθοδος συγκλίνει, αυτή η διαδικασία όμως είναι πολλές φορές χρονοβόρα.

Στη μέθοδο Runge-Kutta-Fehlberg, δεν επαναλαμβάνουμε την παραπάνω διαδικασία, αντίθετα, συγκρίνουμε τα αποτελέσματα σε κάθε βήμα των σχέσεων (6.43) και (6.44) και, αν το σχετικό σφάλμα τους είναι τάξης $\approx h$, όπως προβλέπει η θεωρία, τότε πιστεύουμε ότι η μέθοδος συγκλίνει. Επίσης είναι εξαιρετικά σημαντικό για την ταχύτητα εκτέλεσης του προγράμματος ότι τα k_1, k_2, \dots, k_6 υπολογίζονται μια μόνο φορά και για τις δυο σχέσεις.

Τέλος, για τη μέθοδο Runge-Kutta-Fehlberg υπάρχει και *προσεγγιστική σχέση για το σφάλμα*. Το σφάλμα δίνεται από τη σχέση

$$E = \frac{1}{360}k_1 - \frac{128}{4275}k_3 - \frac{2197}{7524}k_4 + \frac{1}{50}k_5 + \frac{2}{55}k_6 \quad (6.45)$$

Επειδή τα k_1, k_2, \dots, k_6 είναι γνωστά σε κάθε βήμα, είναι δυνατό να εκτιμήσουμε άμεσα το σφάλμα και, αν είναι μεγαλύτερο από τη ζητούμενη ακρίβεια, υποδιπλασιάζουμε το h , ωστότου να την πετύχουμε.

Η προαναφερθείσα διαδικασία ουσιαστικά αναδεικνύει τη Runge - Kutta - Fehlberg σε **μέθοδο μεταβλητού βήματος**. Το πλεονέκτημα αυτής της μεθόδου είναι ότι έχουμε τη δυνατότητα να **αυξάνουμε** ή να **ελαττώνουμε** το βήμα ανάλογα με την συμπεριφορά της αριθμητικής λύσης σε κάθε σημείο, έτσι επιτυγχάνουμε ταχύτητα και, συγχρόνως, εξασφαλίζουμε τη ζητούμενη ακρίβεια.

6.2 ΜΕΘΟΔΟΙ ΠΟΛΛΑΠΛΩΝ ΒΗΜΑΤΩΝ

Οι μέθοδοι ενός βήματος, που αναλύσαμε λίγο πριν, είναι αρκετά καλές και αξιόπιστες αλλά έχουν ένα βασικό μειονέκτημα: Όταν υπολογίζουμε την τιμή της y_{n+1} , αυτό γίνεται μόνο με χρήση της πληροφορίας που έχουμε στο βήμα n για το σημείο (x_n, y_n) . Με αυτό τον τρόπο, χάνεται όλη η γνώση που αποκτήσαμε στα προηγούμενα $n - 1$ βήματα για τη συμπεριφορά της λύσης της διαφορικής εξίσωσης.

Οι **μέθοδοι πολλαπλού βήματος** έχουν το πλεονέκτημα ότι χρησιμοποιούν κατάλληλα την πληροφορία για την $y(x)$ από τα προηγούμενα 3-5 βήματα. Ένα προφανές μειονέκτημα αυτών των μεθόδων είναι η ανάγκη της γνώσης των αρχικών βημάτων, π.χ. των y_1, y_2, y_3 για τον υπολογισμό του y_4 , οπότε είναι αναγκαία η χρήση μιας από τις προηγούμενες μεθόδους ενός βήματος στην αρχή της διαδικασίας.

Η πλέον αξιόλογη και ευρύτατα χρησιμοποιούμενη μέθοδος πολλαπλών βημάτων είναι η **μέθοδος του Adams** που παρουσιάζεται στη συνέχεια.

6.2.1 Μέθοδος Adams

Είναι από τις πλέον σημαντικές σύγχρονες μεθόδους, γι' αυτό θα δείξουμε τον τρόπο εύρεσης της αναδρομικής σχέσης της.²

Αν δοθεί η διαφορική εξίσωση:

$$\frac{dy}{dx} = f(x, y) \quad \text{τότε} \quad dy = f(x, y) dx \quad (6.46)$$

ολοκληρώνοντας και τα δυο μέρη της παραπάνω ισότητας παίρνω:

² Σημειώνουμε ότι είναι η προκαθορισμένη μέθοδος που χρησιμοποιεί η Mathematica όταν καλούμε την εντολή NDSolve.

$$\int_{x_n}^{x_{n+1}} dy = y_{n+1} - y_n = \int_{x_n}^{x_{n+1}} f(x, y) dx \quad (6.47)$$

Στη συνέχεια, υπολογίζουμε το ολοκλήρωμα προσεγγίζοντας την $f(x, y)$ με κατάλληλης τάξης συμπτωτικό πολυώνυμο χρησιμοποιώντας τα 3, 4, 5, ... προηγούμενα σημεία. Εδώ, θα χρησιμοποιήσουμε ένα πολυώνυμο δευτέρου βαθμού, που το υπολογίσαμε από τα τρία προηγούμενα σημεία. Οπότε χρησιμοποιώντας τη μέθοδο του Newton προς τα πίσω,

$$\begin{aligned} \int_{x_n}^{x_{n+1}} dy &= y_{n+1} - y_n \\ &= \int_{x_n}^{x_{n+1}} \left(f_n + s\Delta f_{n-1} + \frac{(s+1)s}{2}\Delta^2 f_{n-2} + \text{σφάλμα} \right) dx \\ &= h \int_{s=0}^{s=1} \left(f_n + s\Delta f_{n-1} + \frac{(s+1)s}{2}\Delta^2 f_{n-2} \right) ds \\ &\quad + h \int_{s=0}^{s=1} \frac{s(s+1)(s+2)}{6} h^3 f'''(\xi) ds \end{aligned} \quad (6.48)$$

οπότε

$$\begin{aligned} y_{n+1} - y_n &= h \left(f_n + \frac{1}{2}\Delta f_{n-1} + \frac{5}{12}\Delta^2 f_{n-2} \right) + O(h^4) \\ &= h \left[f_n + \frac{1}{2}(f_n - f_{n-1}) + \frac{5}{12}(f_n - 2f_{n-1} + f_{n-2}) \right] + O(h^4) \end{aligned} \quad (6.49)$$

και τελικά

$$y_{n+1} = y_n + \frac{h}{12} [23f_n - 16f_{n-1} + 5f_{n-2}] + O(h^4) \quad (6.50)$$

Αν χρησιμοποιούσαμε συμπτωτικό πολυώνυμο τρίτου βαθμού, θα οδηγούμαστε στη σχέση:

$$y_{n+1} = y_n + \frac{h}{24} [55f_n - 59f_{n-1} + 37f_{n-2} - 9f_{n-3}] + O(h^5) \quad (6.51)$$

Παρατηρούμε ότι έχουμε δημιουργήσει πολύ απλά μεθόδους υψηλής ακρίβειας, που προγραμματίζονται εξαιρετικά εύκολα (συγκρίνετε με τη σχέση 6.44). Εν τούτοις, επαναλαμβάνουμε ότι, απαιτείται η χρήση της μεθόδου Runge-Kutta-Fehlberg για τον υπολογισμό των τιμών της $y(x)$ στις θέσεις $n-3$, $n-2$, $n-1$ και n , αν χρησιμοποιήσουμε την εξίσωση (6.51). Αν δεν είναι εύκολο να προγραμματίσουμε τη μέθοδο Runge-Kutta-Fehlberg, τότε μπορούμε να χρησιμοποιήσουμε την απλοϊκή μέθοδο Euler, όμως με σφάλμα $\approx h^2$ σε κάθε βήμα. Αυτό όμως θα έχει ως αποτέλεσμα την ελάττωση της ακρίβειας για τον υπολογισμό της $y(x)$ στα τέσσερα πρώτα βήματα. Θα μπορούσαμε να διορθώσουμε αυτό το αδύνατο σημείο της Euler μειώνοντας σημαντικά το βήμα, οπότε το συνολικό σφάλμα μετά από μεγαλύτερο προφανώς αριθμό βημάτων να είναι της τάξης $\approx h^5$.

6.2.2 Μέθοδος Milne

Η μέθοδος του Milne χρησιμοποιείται ευρύτατα, αποτελεί αξιόλογη μέθοδο και είναι απλή στο προγραμματισμό της. Επίσης, όπως θα φανεί, είναι εξαιρετικά ακριβής.

Για την εύρεση της αναδρομικής σχέσης ακολουθούμε παρόμοια διαδικασία με αυτή της μεθόδου Adams. Αν υποθέσουμε ότι η λύση είναι γνωστή στα σημεία x_n, x_{n-1}, x_{n-2} και x_{n-3} , τότε η διαφορική εξίσωση:

$$\frac{dx}{dy} = f(x, y) \quad (6.52)$$

μπορεί να γραφεί κατά τα προηγούμενα ως

$$\int_{x_{n-3}}^{x_{n+1}} dy = y_{n+1} - y_{n-3} = \int_{x_{n-3}}^{x_{n+1}} f(x, y) dx \quad (6.53)$$

Αν τώρα αντικαταστήσουμε την $f(x, y)$ με ένα δευτεροβάθμιο συμπτωτικό πολυώνυμο και ολοκληρώσουμε, καταλήγουμε στη σχέση

$$y_{n+1} = y_{n-3} + \frac{4h}{3} (2f_n - f_{n-1} + 2f_{n-2}) \quad (6.54)$$

με σφάλμα

$$E \approx \frac{28}{90} h^5 y^{(5)}(\xi), \quad x_{n-3} \leq \xi \leq x_{n+1} \quad (6.55)$$

6.3 ΜΕΘΟΔΟΙ ΠΡΟΒΛΕΨΗΣ - ΔΙΟΡΘΩΣΗΣ

Σ' αυτές τις μεθόδους, ο υπολογισμός της τιμής της y_{n+1} στο σημείο x_{n+1} γίνεται σε δυο βήματα, δηλαδή υπάρχουν δυο σχέσεις για τον υπολογισμό των y_{n+1} .

ΤΥΠΟΣ ΠΡΟΒΛΕΨΗΣ

$$y_{k+1} = \Pi(y_{k-m}, y_{k-m+1}, \dots, y_k) \quad (6.56)$$

δηλ. στην πρώτη προσπάθεια χρησιμοποιούμε τα m προηγούμενα σημεία.

ΤΥΠΟΣ ΔΙΟΡΘΩΣΗΣ

$$y_{k+1} = \Delta(y_{k-m}, y_{k-m+1}, \dots, y_k, y_{k+1}) \quad (6.57)$$

στη συνέχεια χρησιμοποιείται η πρόβλεψη y_{k+1} σε ένα τύπο που συνδυάζει τις τιμές της συνάρτησης στα m προηγούμενα σημεία (αυτό δεν είναι πάντα υποχρεωτικό).

Αυτή η διαδικασία γίνεται επαναληπτικά, ωστόσο οι δυο διαδοχικές τιμές της y_{k+1} να είναι αρκετά κοντά: $|y_{k+1}^{(\lambda)} - y_{k+1}^{(\lambda+1)}| < E$, όπου λ είναι ο αριθμός επαναλήψεων της διαδικασίας και E η ζητούμενη ακρίβεια.

Στη συνέχεια, θα δείξουμε ένα τύπο διόρθωσης για τη μέθοδο Milne. Για τη διαφορική εξίσωση

$$\frac{dx}{dy} = f(x, y) \quad (6.58)$$

μπορούμε, κατά τα γνωστά, να γράψουμε:

$$\int_{x_{n-1}}^{x_{n+1}} dy = y_{n+1} - y_{n-1} = \int_{x_{n-1}}^{x_{n+1}} f(x, y) dx \quad (6.59)$$

οπότε αντικαθιστώντας το ολοκλήρωμα του δεύτερου μέρους με τον κανόνα του Simpson δημιουργούμε τη σχέση

$$y_{n+1} = y_{n-1} + \frac{h}{3} (f_{n+1} + 4f_n + f_{n-1}) \quad (6.60)$$

με σφάλμα

$$E \approx \frac{h^5}{90} y^{(5)}(\xi) \quad \text{όπου} \quad x_{n-1} < \xi < x_{n+1} \quad (6.61)$$

Παρατηρούμε ότι το σφάλμα είναι μεν $\approx h^5$, όπως και στη Milne, αλλά κατά 28 φορές μικρότερο, που δείχνει την αποτελεσματικότητα της *διορθωτικής διαδικασίας*. Ανάλογα αποδεικνύεται και ο τύπος διόρθωσης για τη μέθοδο Adams.

Συνοπτικά έχουμε τους παρακάτω συνδυασμούς τύπων πρόβλεψης-διόρθωσης για διαφορικές εξισώσεις της μορφής $y = f(x, y)$:

- **ΜΕΘΟΔΟΣ Adams-Multon**

ΤΥΠΟΣ ΠΡΟΒΛΕΨΗΣ $(x_{k-3} < \xi_1 < x_{k+1})$

$$y_{k+1} = y_k + \frac{h}{24} (55f_k - 59f_{k-1} + 37f_{k-2} - 9f_{k-3}) + \frac{251}{720} h^5 y^{(5)}(\xi_1) \quad (6.62)$$

ΤΥΠΟΣ ΔΙΟΡΘΩΣΗΣ $(x_{k-2} < \xi_2 < x_{k+1})$

$$y_{k+1} = y_k + \frac{h}{24} (9f_{k+1} + 19f_k - 5f_{k-1} + f_{k-2}) - \frac{19}{720} h^5 y^{(5)}(\xi) \quad (6.63)$$

- **ΜΕΘΟΔΟΣ Hamming**

Μια επίσης συχνά χρησιμοποιούμενη μέθοδος πρόβλεψης - διόρθωσης είναι η μέθοδος Hamming. Η απόδειξη των σχέσεων επαφίεται στον αναγνώστη.

ΤΥΠΟΣ ΠΡΟΒΛΕΨΗΣ $(x_{k-3} < \xi_1 < x_{k+1})$

$$y_{k+1} = y_{k-3} + \frac{4h}{3} (2f_{k-2} - f_{k-1} + 2f_k) + \frac{28}{90} h^5 y^{(5)}(\xi_1) \quad (6.64)$$

ΤΥΠΟΣ ΔΙΟΡΘΩΣΗΣ $(x_{k-2} < \xi_2 < x_{k+1})$

$$y_{k+1} = \frac{1}{8} (9y_k - y_{k-2}) + \frac{3h}{8} (-f_{k-1} + 2f_k + f_{k+1}) \quad (6.65)$$

• **ΜΕΘΟΔΟΣ Milne**

ΤΥΠΟΣ ΠΡΟΒΛΕΨΗΣ ($x_{k-3} < \xi_1 < x_{k+1}$)

$$y_{k+1} = y_{k-3} + \frac{4h}{3} (2f_{k-2} - f_{k-1} + 2f_k) + \frac{28}{90} h^5 y^{(5)}(\xi_1) \quad (6.66)$$

ΤΥΠΟΣ ΔΙΟΡΘΩΣΗΣ ($x_{k-1} < \xi_2 < x_{k+1}$)

$$y_{k+1} = y_{k-1} + \frac{h}{3} (f_{k-1} + 4f_k + f_{k+1}) - \frac{1}{90} h^5 y^{(5)}(\xi_2) \quad (6.67)$$

Για τον υπολογισμό των αρχικών σημείων χρησιμοποιείται συνήθως Runge-Kutta τέταρτης τάξης ή καλύτερα Runge-Kutta-Fehlberg.

ΠΑΡΑΤΗΡΗΣΗ

Οι μέθοδοι ενός βήματος και πολλών βημάτων μπορούν κατάρχας να μας οδηγήσουν στη λύση με την επιθυμητή ακρίβεια, εν τούτοις οι μέθοδοι πολλών βημάτων είναι ταχύτερες στην εκτέλεση διότι ο αριθμός των πράξεων που απαιτούνται σε κάθε βήμα είναι σημαντικά μικρότερος. Συγκρίνετε για παράδειγμα των αριθμό πράξεων που απαιτούνται σε κάθε βήμα για τη μέθοδο Runge-Kutta-Fehlberg και για τη μέθοδο Adams.

6.4 ΣΥΣΤΗΜΑΤΑ ΔΙΑΦΟΡΙΚΩΝ ΕΞΙΣΩΣΕΩΝ

Οι περισσότερες από τις διαφορικές εξισώσεις στη Φυσική δεν είναι πρώτης τάξης, ωστόσο η μεθοδολογία που αναπτύχθηκε στα προηγούμενα κεφάλαια εφαρμόζεται εύκολα σε διαφορικές εξισώσεις ανώτερης τάξης και συστήματα διαφορικών εξισώσεων.

Οι διαφορικές εξισώσεις δεύτερης τάξης είναι αυτές που συναντώνται ευρύτερα, δηλαδή της μορφής :

$$\frac{d^2 y}{dx^2} = f(x, y, y') \quad (6.68)$$

οπότε απαιτείται η γνώση αρχικών συνθηκών, όχι μόνο για την $y(x)$ αλλά και για την παράγωγό της στην αρχική θέση x_0 . Το σύνηθες είναι να ορίζουμε μια νέα συνάρτηση $z(x)$ τέτοια, ώστε $z(x) = y'$. Οπότε η δεύτερης τάξης διαφορική εξίσωση ανάγεται σε δυο πρωτοβάθμιες με τις ανάλογες αρχικές συνθήκες. Δηλαδή

$$\frac{dy}{dx} = z \quad (6.69)$$

$$\frac{dz}{dx} = f(x, y, z) \quad (6.70)$$

και

$$y(x_0) = y_0 \quad z(x_0) = y'_0 \quad (6.71)$$

Το παραπάνω σύστημα διαφορικών εξισώσεων είναι ισοδύναμο με την διαφορική εξίσωση (6.68). Η διαδικασία αυτή μπορεί να ακολουθηθεί για κάθε βαθμού διαφορική εξίσωση. Με όσα προαναφέρθηκαν είναι προφανής και ο τρόπος με τον οποίο μπορούν να χρησιμοποιηθούν οι μέθοδοι Adams, Milne κλπ. για συστήματα διαφορικών εξισώσεων.

Έστω λοιπόν ότι έχουμε το παρακάτω σύστημα

$$\frac{dy}{dx} = zy + x \quad \text{με} \quad y(0) = 1 \quad (6.72)$$

$$\frac{dz}{dx} = xz + y \quad \text{με} \quad z(0) = 1 \quad (6.73)$$

Αν χρησιμοποιήσουμε τη **μέθοδο Euler**, το σύστημα θα δίνεται από αναδρομικές σχέσεις της μορφής

$$y'_{n+1} = y_n + hy'_n = y_n + h(z_n y_n + x_n) \quad (6.74)$$

$$z'_{n+1} = z_n + hz'_n = z_n + h(x_n z_n + y_n) \quad (6.75)$$

Είναι προφανής λοιπόν η αναλογία της μεθόδου με όσα αναφέραμε στην αρχή του κεφαλαίου για τις μεθόδους Euler και κατ' επέκταση Taylor. Αντίστοιχα, αν χρησιμοποιήσουμε την μέθοδο Runge - Kutta, θα πρέπει να υπολογίσουμε πέραν των k_1, k_2, \dots, k_6 και τα l_1, l_2, \dots, l_6 . Εδώ, αναφέρουμε τη **μέθοδο Runge - Kutta τέταρτης τάξης**, οπότε και η επέκτασή της γίνεται αναλόγως.

Θα είναι λοιπόν για ένα σύστημα $y' = f(x, y, z)$ και $z' = g(x, y, z)$

$$y_{n+1} = y_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \quad (6.76)$$

$$z_{n+1} = z_n + \frac{1}{6}(l_1 + 2l_2 + 2l_3 + l_4) \quad (6.77)$$

όπου

$$k_1 = hf(x_n, y_n, z_n) \quad (6.78)$$

$$k_2 = hf\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}k_1, z_n + \frac{1}{2}l_1\right) \quad (6.79)$$

$$k_3 = hf\left(x_n + \frac{h}{2}, y_n + \frac{1}{2}k_2, z_n + \frac{1}{2}l_2\right) \quad (6.80)$$

$$k_4 = hf(x_n + h, y_n + k_3, z_n + l_3) \quad (6.81)$$

και

$$l_1 = hg(x_n, y_n, z_n) \quad (6.82)$$

$$l_2 = hg\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}k_1, z_n + \frac{1}{2}l_1\right) \quad (6.83)$$

$$l_3 = hg\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}k_2, z_n + \frac{1}{2}l_2\right) \quad (6.84)$$

$$l_4 = hg(x_n + h, y_n + k_3, z_n + l_3) \quad (6.85)$$

6.5 ΑΣΚΗΣΕΙΣ

1. Ξερεσιμοποιείστε μία από τις μεθόδους Euler, Euler-Heun, Runge-Kutta και Runge-Kutta-Felberg για την αριθμητική επίλυση των παρακάτω διαφορικών εξισώσεων και συγκρίνετε με τις ακριβείς τιμές :

α) $y' = xy^{1/3}$, με $y(1) = 1$ (ακριβής λύση $y = \left(\frac{x^2+2}{3}\right)^{3/2}$).

Υπόδειξη: ολοκληρώστε από $x = 1$ ως $x = 2$

β) $y' = -xy^2$ με $y(0) = 2$ (ακριβής λύση $y = \frac{2}{x^2}$)

Υπόδειξη: ολοκληρώστε από $x = 0$ ως $x = 1$

γ) $y' = -2xy$ με $y(1) = 1$ (ακριβής λύση $y = e^{-x^2}$)

Υπόδειξη: ολοκληρώστε από $x = 0$ ως $x = 1$

δ)

$$y' = \frac{y(1 - x^2y^4)}{x(1 + x^2y^4)} \quad \text{με} \quad y(1) = 1.$$

Υπόδειξη: ολοκληρώστε από $x = 1$ ως $x = 2$

ε)

$$y' = \frac{x - e^x}{y + e^y} \quad \text{με} \quad y(0) = 0$$

Βρείτε το $y(1)$.

2. Χρησιμοποιήστε μία από τις μεθόδους πρόβλεψης διόρθωσης Milne, Hamming και Adams-Multon για την επίλυση των προηγούμενων διαφορικών εξισώσεων.
3. Αποδείξτε την πρόβλεψη με τον τύπο του Nystrom

$$y_{k+1} = y_{k-1} + 2hy'_k \quad (6.86)$$

και στη συνέχεια χρησιμοποιήστε τη διόρθωση Euler - Heun

$$y_{k+1} = y_k + \frac{h}{2} (y'_k + y'_{k+1}) \quad (6.87)$$

για τη λύση της διαφορικής εξίσωσης του παραδείγματος 6.1.

4. Να λυθεί αριθμητικά το σύστημα διαφορικών εξισώσεων

$$\begin{aligned} x' &= 1195x - 1995y & x(0) &= 2 \\ y' &= 1197x - 1997y & y(0) &= -2 \end{aligned}$$

Να βρεθούν: οι τιμές $y(1)$, $x(1)$ και οι τιμές $y(-1)$, $x(-1)$ Τι παρατηρείτε αν η ακριβής λύση είναι:

$$x(t) = 10e^{-2t} - 8e^{-800t}, y(t) = 6e^{-2t} - 8e^{-800t}$$

5. Να λυθούν τα συστήματα:

$$\begin{aligned} y' &= yz + x & y(0) &= 0 \\ z' &= y - x & z(0) &= 0 \end{aligned}$$

και

$$\begin{aligned} y' &= z & y(0) &= 0 \\ z' &= -y & z(0) &= 1 \end{aligned}$$

6. Θεωρήστε τη μη-γραμμική ΔΕ

$$q''(t) - aq'(t) + bq^3(t) = 0$$

υποθέστε ότι $b > 0$ και a θετικό ή αρνητικό. Η εξίσωση αυτή είναι γνωστή ως *εξίσωση Duffing*.

- α) Λύστε αριθμητικά την ΔΕ και δημιουργήστε τις γραφικές παραστάσεις των $q(t)$ και $q'(t)$. Δώστε αρχικές τιμές π.χ. $q(0) = 0$ και $q'(0) = 0.001$. Ψποθέστε για τις παραμέτρους τα παρακάτω ζεύγη τιμών ($b = 0.05, a = -1$) ($b = 0.05, a = 4$). Τι παρατηρείτε για κάθε ένα απο τα ζεύγη των παραμέτρων.
- β) Δημιουργήστε διαγράμματα φάσης για κάθε μια απο τις παραπάνω περιπτώσεις.
- γ) Δημιουργήστε προσομοιώσεις των παραπάνω περιπτώσεων.
- δ) (Προαιρετικό πρόβλημα) Θα μπορούσατε να δημιουργήσετε γραφήματα του δυναμικού για την παραπάνω εξίσωση σαν συνάρτηση του q και να εξετάσετε την εξάρτηση απο το a .

7. Θεωρήστε τη μη-γραμμική ΔΕ

$$q''(t) = aq(t) - bq^3(t) - \gamma q'(t) + Q_0 \cos(\omega t)$$

με αρχικές συνθήκες $q(0) = 0$ $q'(0) = 0.001$. Υποθέστε $a = 0.4$, $b = 0.5$, $\gamma = 0.2$ και $\omega = 1/8$ και θεωρήστε δύο τροχίες μία για $Q_0 = 0$ και μια για $Q_0 = 0.1$.

- α) Να λύσετε αριθμητικά τις εξισώσεις κίνησης και για τις δύο παραπάνω περιπτώσεις και να δημιουργήσετε γραφικές παραστάσεις για τα $q(t)$ και $q'(t)$.
- β) Δημιουργήστε τα διαγράμματα φάσης, και χρησιμοποιήστε το μετασχηματισμό Fourier για τον υπολογισμό της συχνότητας.

ΠΡΟΣΕΓΓΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ

Σ' αυτό το κεφάλαιο θα μελετήσουμε μεθόδους για την ανάλυση μεγάλου πλήθους δεδομένων. Πιο συγκεκριμένα, θα προσπαθήσουμε να εκτιμήσουμε τη συνολική συμπεριφορά μιας n -άδας σημείων (x_i, y_i) . Οι τεχνικές που θα εξετάσουμε είναι η μέθοδος των ελαχίστων τετραγώνων και η προσέγγιση ελαχίστου-μεγίστου. Στο τέλος θα εξετάσουμε και μια μέθοδο για τον υπολογισμό 'συμπτωτικών' ρητών συναρτήσεων.

7.1 ΜΕΘΟΔΟΣ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ

Όταν θέλουμε να μελετήσουμε μεγάλο αριθμό σημείων, τότε η προσέγγιση τους με συμπτωτικό πολυώνυμο είναι και δύσκολη αλλά και άνευ νοήματος τις περισσότερες φορές. Σ' αυτή την περίπτωση προσπαθούμε να βρούμε ένα πολυώνυμο $P(x)$ ή γενικότερα μια συνάρτηση $f(x)$ με συμπεριφορά παρόμοια με αυτή της ομάδας των σημείων που διαθέτουμε και για την οποία η διαφορά $y_i - f(x_i)$ να είναι ελάχιστη.

Στην πράξη, ζητάμε την ελαχιστοποίηση του αθροίσματος

$$S = \sum_{i=0}^n [y_i - P(x_i)]^2 \quad (7.1)$$

όπου $P(x)$ το ζητούμενο πολυώνυμο συγκεκριμένου βαθμού (συνήθως όχι μεγαλύτερου του τέταρτου ή πέμπτου).

7.1.1 Ευθεία ελαχίστων τετραγώνων

Το ζητούμενο πολυώνυμο ελαχίστων τετραγώνων θα είναι πρωτοβάθμιο, δηλαδή: $P(x) = ax + b$. Άρα, θα προσπαθήσουμε να ελαχιστοποιήσουμε το άθροισμα:

$$S = \sum_{i=0}^n [y_i - ax_i - b]^2 \quad (7.2)$$

που ανάγεται στον προσδιορισμό των a και b ώστε το S να γίνει **ελάχιστο**.

Εφαρμόζοντας τη θεωρία εύρεσης μεγίστου ή ελαχίστου συναρτήσεων πολλών μεταβλητών απαιτούμε να είναι:

$$\frac{\partial S}{\partial a} = 0 \quad \text{και} \quad \frac{\partial S}{\partial b} = 0 \quad (7.3)$$

Οπότε:

$$\frac{\partial S}{\partial a} = -2 \sum_{i=0}^n x_i (y_i - ax_i - b) = 0 \quad (7.4)$$

$$\frac{\partial S}{\partial b} = -2 \sum_{i=0}^n (y_i - ax_i - b) = 0$$

Οι σχέσεις αυτές μπορούν να γραφούν ως :

$$\begin{aligned} (n+1)b + \left(\sum x_i\right)a &= \sum y_i \\ \left(\sum x_i\right)b + \left(\sum x_i^2\right)a &= \sum x_i y_i \end{aligned} \quad (7.5)$$

και, αν θέσουμε

$$\begin{aligned} s_0 &= n+1 & s_1 &= \sum x_i & s_2 &= \sum x_i^2 \\ u_0 &= \sum y_i & u_1 &= \sum x_i y_i \end{aligned} \quad (7.6)$$

τότε το σύστημα (7.5) απλοποιείται και βρίσκουμε:

$$s_0 b + s_1 a = u_0 \quad \text{και} \quad s_1 b + s_2 a = u_1$$

οπότε

$$a = \frac{s_0 u_1 - s_1 u_0}{s_0 s_2 - s_1^2} \quad \text{και} \quad b = \frac{s_2 u_0 - s_1 u_1}{s_0 s_2 - s_1^2}. \quad (7.7)$$

Επιπλέον, αποδεικνύεται εύκολα ότι:

$$s_0 s_2 - s_1^2 = (n+1) \left(\sum x_i^2 \right) - \left(\sum x_i \right)^2 = \sum_{i < j} (x_i - x_j)^2 > 0 \quad (7.8)$$

Για να είναι το S ελάχιστο για τις τιμές των a και b , θα πρέπει:

$$\Delta = \frac{\partial^2 S}{\partial a^2} \frac{\partial^2 S}{\partial b^2} - \left(\frac{\partial^2 S}{\partial a \partial b} \right)^2 > 0 \quad (7.9)$$

και, επιπλέον, μια από τις $\frac{\partial^2 S}{\partial a^2}$, $\frac{\partial^2 S}{\partial b^2}$ να είναι μεγαλύτερη του μηδενός για τις τιμές των a και b , που δίνονται από την εξίσωση (7.7). Επομένως, βρίσκουμε εύκολα ότι:

$$\frac{\partial^2 S}{\partial a^2} = 2s_2 > 0, \quad \frac{\partial^2 S}{\partial b^2} = 2s_0 > 0 \quad \text{και} \quad \frac{\partial^2 S}{\partial a \partial b} = 2s_1 \quad (7.10)$$

επομένως

$$\Delta = 2s_2 \cdot 2s_0 - (2s_1)^2 = 4(s_0 s_2 - s_1^2) > 0 \quad (7.11)$$

Άρα, οι τιμές που υπολογίζονται στις σχέσεις (7.7) είναι τα ζητούμενα ελάχιστα.

ΠΑΡΑΔΕΙΓΜΑ

Βρείτε ένα τύπο της μορφής $P(x) = Ae^{Mx}$ (όπου A και M είναι άγνωστες σταθερές που πρέπει να υπολογισθούν) για τα δεδομένα:

| | | | | |
|--------|---|----|----|----|
| x | 1 | 2 | 3 | 4 |
| $P(x)$ | 7 | 11 | 17 | 27 |

Θέτω $z = \log(P)$ και $B = \log(A)$ οπότε η ζητούμενη σχέση είναι η $z = B + Mx$, επομένως δημιουργώ ένα νέο πίνακα στοιχείων για τα νέα ζεύγη (x_i, z_i) δηλαδή

| | | | | |
|-----|--------|--------|--------|--------|
| x | 1 | 2 | 3 | 4 |
| z | 1.9459 | 2.3979 | 2.8332 | 3.2958 |

Οπότε εύκολα βρίσκουμε $s_0 = 4$, $s_1 = 10$, $s_2 = 30$, $v_0 = 10.48$ και $v_1 = 28.44$ άρα $M \approx 0.45$, $B = 1.5$ και επομένως $A = 4.48$. Τελικά η ζητούμενη σχέση είναι: $P(x) = 4.48e^{0.45x}$.

ΠΑΡΑΔΕΙΓΜΑ

Δίνονται τα παρακάτω ζεύγη σημείων:

| | | | | | | | | | |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| x | 0.05 | 0.11 | 0.15 | 0.31 | 0.46 | 0.52 | 0.70 | 0.74 | 0.82 |
| y | 0.956 | 0.890 | 0.832 | 0.717 | 0.571 | 0.539 | 0.378 | 0.370 | 0.306 |

Να υπολογισθεί η ευθεία ελαχίστων τετραγώνων.

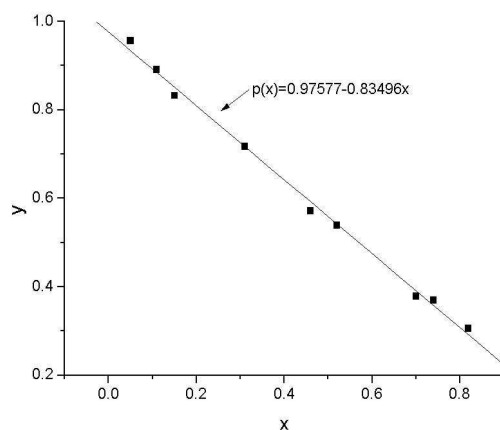
$$\text{Είναι } n = 9, s_0 = 10, s_1 = \sum x_i = 3.86, s_2 = \sum x_i^2 = 2.3252$$

$$u_0 = \sum y_i = 5.559, \quad \text{και} \quad u_1 = \sum x_i y_i = 1.82506$$

οπότε

$$a = -0.83496, \quad \text{και} \quad b = 0.97577$$

Επομένως, $P(x) = 0.97577 - 0.83496x$.



Σχήμα 7.1. Γραφική απεικόνιση της ευθείας ελαχίστων τετραγώνων για το προηγούμενο παράδειγμα.

7.1.2 Παραβολή ελαχίστων τετραγώνων

Η διαδικασία εύρεσης της παραβολής ελαχίστων τετραγώνων είναι παρόμοια με αυτή της ευθείας ελαχίστων τετραγώνων. Επομένως, ελαχιστοποιούμε το άθροισμα:

$$S = \sum_{i=0}^n (y_i - ax_i^2 - bx_i - c)^2 \quad (7.12)$$

Από τις σχέσεις για την ύπαρξη ακροτάτου θα είναι

$$\frac{\partial S}{\partial a} = \frac{\partial S}{\partial b} = \frac{\partial S}{\partial c} = 0 \quad (7.13)$$

και έτσι καταλήγουμε στο σύστημα

$$\begin{pmatrix} s_0 & s_1 & s_2 \\ s_1 & s_2 & s_3 \\ s_2 & s_3 & s_4 \end{pmatrix} \cdot \begin{pmatrix} c \\ b \\ a \end{pmatrix} = \begin{pmatrix} u_0 \\ u_1 \\ u_2 \end{pmatrix} \quad (7.14)$$

όπου

$$s_3 = \sum x_i^3, \quad s_4 = \sum x_i^4 \quad \text{και} \quad u_2 = \sum x_i^2 y_i$$

ενώ τα υπόλοιπα στοιχεία s_0, s_1, s_2, v_0, v_1 δίνονται από τις σχέσεις (7.6).

Στη γενική περίπτωση πολυωνύμου ελαχίστων τετραγώνων βαθμού $n - 1$ οι συντελεστές του πολυωνύμου προσδιορίζονται από τη λύση του παρακάτω συστήματος:

$$\begin{pmatrix} s_0 & s_1 & \dots & \dots & s_n \\ s_1 & s_2 & \dots & \dots & s_{n+1} \\ \vdots & \vdots & & & \vdots \\ \vdots & \vdots & & & \vdots \\ s_n & s_{n+1} & \dots & \dots & s_{2n} \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} u_0 \\ u_1 \\ \vdots \\ \vdots \\ u_n \end{pmatrix} \quad (7.15)$$

όπου τα s_k και u_k δίνονται από τις σχέσεις :

$$s_k = \sum x_i^k u_k = \sum x_i^k y_i$$

και το πολυώνυμο είναι:

$$P(x) = a_0 + a_1 x + \dots + a_n x^n.$$

Στο 2ο παράδειγμα της προηγούμενης ενότητας αν ζητούσαμε πολυώνυμο δευτέρου βαθμού, τότε θα βρίσκαμε

$$P(x) = 0.996 - 0.995x + 0.186x^2$$

ενώ το πολυώνυμο ελαχίστων τετραγώνων τρίτου βαθμού θα είναι:

$$P(x) = 1.0 - 1.06x + 0.37x^2 - 0.14x^3.$$

7.1.3 Συνεχή δεδομένα

Εστω η $y(x)$ μια συνεχής συνάρτηση που επιθυμούμε να προσεγγίσουμε στο διάστημα $[-1, 1]$ με ένα πολυώνυμο της μορφής :

$$P(x) = a_m P_m(x) + a_{m-1} P_{m-1}(x) + \dots + a_0 P_0(x)$$

όπου $P_k(x)$ το πολυώνυμο Legendre βαθμού k και τα a_k σταθερές που πρέπει να υπολογισθούν.¹

Το τετραγωνικό σφάλμα αυτής της προσέγγισης θα είναι:

¹ Τα πολυώνυμα Legendre δίνονται από τον τύπο του Rodrigues

$$P_k(x) = \frac{1}{2^k k!} \frac{d^k}{dx^k} (x^2 - 1)^k$$

Τα πρώτα από αυτά είναι: $P_0(x) = 1$, $P_1(x) = x$, $P_2 = \frac{1}{2}(3x^2 - 1)$ κ.ο.κ.

$$S = \int_{-1}^1 [y(x) - a_m P_m(x) - a_{m-1} P_{m-1}(x) - \dots - a_0 P_0(x)]^2 dx \quad (7.16)$$

Οπότε αν θελήσω να ελαχιστοποιήσω το σφάλμα θα πρέπει να απαιτήσω:

$$\frac{\partial S}{\partial a_k} = 0$$

επομένως θα είναι:

$$\frac{\partial S}{\partial a_k} = \int_{-1}^1 [y(x) - a_m P_m(x) - a_{m-1} P_{m-1}(x) - \dots - a_0 P_0(x)] P_k(x) dx = 0$$

Επειδή όμως τα πολυώνυμα Legendre είναι ορθογώνια, ισχύει:

$$\int_{-1}^1 P_m(x) \cdot P_k(x) dx = \begin{cases} 0 & \text{αν } k \neq m \\ \frac{2}{k+1} & \text{αν } k = m \end{cases} \quad (7.17)$$

οπότε

$$\frac{\partial S}{\partial a_k} = \int_{-1}^1 [y(x) - a_k P_k(x)] P_k(x) dx = 0$$

και εύκολα αποδεικνύεται ότι οι συντελεστές a_k δίνονται από τις σχέσεις:

$$a_k = \frac{2k+1}{2} \int_{-1}^1 y(x) P_k(x) dx \quad (7.18)$$

Τέλος, αν το διάστημα δεν είναι $[-1, 1]$ αλλά τυχαίο $[a, b]$, τότε χρησιμοποιώ την αντικατάσταση:

$$x = \frac{1}{2} (b-a) t + \frac{1}{2} (b+a) \quad (7.19)$$

ΠΑΡΑΔΕΙΓΜΑ

Με τη μέθοδο ελαχίστων τετραγώνων προσδιορίστε την παραβολή που προσεγγίζει τη συνάρτηση $y(t) = \sin(t)$ στο διάστημα $[0, \pi]$.

Κατ'αρχάς αλλιάζω το διάστημα με την εξής αλλαγή μεταβλητής: $t = \frac{\pi}{2}(x+1)$ οπότε $[0, \pi] \rightarrow [-1, 1]$ άρα η νέα συνάρτηση είναι:

$$y = \sin \left[\frac{\pi}{2}(x+1) \right].$$

και οι ζητούμενοι συντελεστές θα είναι:

$$a_0 = \frac{1}{2} \int_{-1}^1 \sin \left[\frac{\pi}{2}(x+1) \right] dx = \frac{2}{\pi}$$

$$a_1 = \frac{3}{2} \int_{-1}^1 \sin \left[\frac{\pi}{2}(x+1) \right] \cdot x \cdot dx = 0$$

$$a_2 = \frac{5}{2} \int_{-1}^1 \sin \left[\frac{\pi}{2}(x+1) \right] \cdot \frac{1}{2}(3x^2 - 1) \cdot dx = \frac{10}{\pi} \left(1 - \frac{12}{\pi^2} \right)$$

τελικά η ζητούμενη παραβολή είναι:

$$y(x) = \frac{2}{\pi} + \frac{10}{\pi} \left(1 - \frac{12}{\pi^2} \right) \frac{1}{2} (3x^2 - 1)$$

οπότε αλληλάζοντας μεταβλητή καταλήγω στην εξής συνάρτηση:

$$y = \frac{2}{\pi} + \frac{10}{\pi} \left(1 - \frac{12}{\pi^2} \right) \left[\frac{6}{\pi^2} \left(t - \frac{\pi}{2} \right)^2 - \frac{1}{2} \right]$$

7.2 ΠΡΟΣΕΓΓΙΣΗ ΕΛΑΧΙΣΤΟΥ - ΜΕΓΙΣΤΟΥ

Έστω ότι δίνεται μια συνάρτηση $y(x)$ ή ένα σύνολο σημείων (x_i, y_i) και ζητούμε τον υπολογισμό μιας νέας προσεγγιστικής συνάρτησης $\bar{y}(x)$. Το κριτήριο επιλογής της συνάρτησης είναι ότι για αυτή τη συνάρτηση **το μέγιστο σφάλμα γίνεται ελάχιστο**.

Το σφάλμα στην αντικατάσταση μιας συνάρτησης $y(x)$ από μια προσεγγιστική συνάρτηση $\bar{y}(x)$ σε ένα τυχαίο σημείο x είναι:

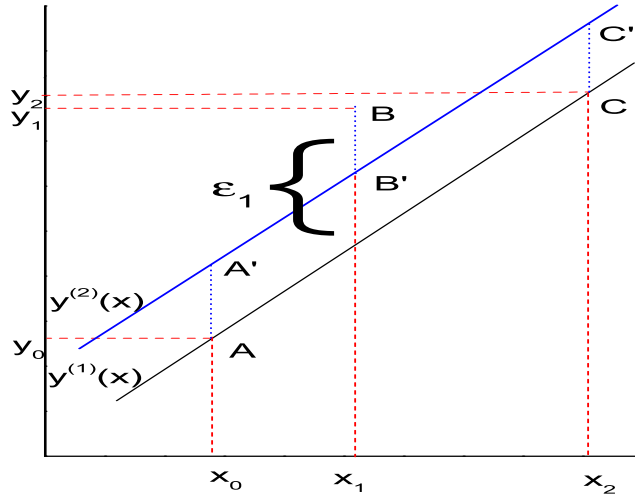
$$\varepsilon = |y(x) - \bar{y}(x)| \quad (7.20)$$

Για απλότητα θα δείξουμε τη διαδικασία για την εύρεση μιας γραμμικής συνάρτησης $\bar{y}(x) = ax + b$ που θα ελαχιστοποιεί το μέγιστο σφάλμα, η ευθεία αυτή ονομάζεται *ευθεία ίσων σφαλμάτων ή Chebyshev* και η προσέγγιση θα γίνει μόνο για τρία σημεία (x_0, y_0) , (x_1, y_1) και (x_2, y_2) . Επομένως το σφάλμα στο τυχαίο σημείο x_i θα είναι:

$$|\varepsilon_i| = |y_i - (ax_i + b)| \quad (7.21)$$

Έστω ότι η εξίσωση για μια ευθεία που διέρχεται από τα σημεία (x_0, y_0) , και (x_2, y_2) είναι:

$$y^{(1)}(x) = \frac{y_2 - y_0}{x_2 - x_0} x + \frac{y_0 x_2 - y_2 x_0}{x_2 - x_0} \quad (7.22)$$



Σχήμα 7.2. Γραφική απεικόνιση της μεθόδου εύρεσης της ευθείας ίσων σφαλμάτων.

Τα σφάλματα για τα τρία σημεία θα είναι:

$$\begin{aligned} \varepsilon_0 &= 0, \quad \text{για το } (x_0, y_0) \\ \varepsilon_1 &= |y_1 - y^{(1)}(x_1)| = |y_1 - ax_1 - b| \quad \text{για το } (x_1, y_1) \\ \varepsilon_2 &= 0 \quad \text{για το } (x_2, y_2) \end{aligned} \quad (7.23)$$

Αν μετατοπίσω την ευθεία παράλληλα έτσι, ώστε το σφάλματα να είναι $\varepsilon_0 = \varepsilon_1 = \varepsilon_2$ θα σχηματίσω μια νέα ευθεία, την

$$y^{(2)} = \frac{y_2 - y_0}{x_2 - x_0}x + c \quad (7.24)$$

Η ποσότητα c είναι μια άγνωστη σταθερά που υπολογίζεται από τη σχέση $AA' = BB'$, δηλαδή:

$$y_0 - \left(\frac{y_2 - y_0}{x_2 - x_0}x_0 + c \right) = \left(\frac{y_2 - y_0}{x_2 - x_0}x_1 + c \right) - y_1 \quad (7.25)$$

οπότε τελικά βρίσκουμε τη νέα **ευθεία ίσων σφαλμάτων**:

$$y^{(2)}(x) = \frac{y_2 - y_0}{x_2 - x_0}x + \frac{y_0(x_1 + x_2) + y_1(x_2 - x_0) - y_2(x_0 + x_1)}{2(x_2 - x_0)} \quad (7.26)$$

Προφανώς τα σφάλματα είναι $\varepsilon_0 = y_0 - y_0^{(2)}$, $\varepsilon_1 = y_1 - y_1^{(2)}$ και $\varepsilon_2 = y_2 - y_2^{(2)}$ για τα οποία ισχύει ότι:

$$|\varepsilon_0| = |\varepsilon_1| = |\varepsilon_2| \quad (7.27)$$

ενώ το σφάλμα σε κάθε ένα από τα τρία σημεία θα είναι:

$$\varepsilon = y_0 - y_0^{(2)} = \frac{y_0(x_2 - x_1) - y_1(x_2 - x_0) + y_2(x_1 - x_0)}{2(x_2 - x_0)} \quad (7.28)$$

Αν έχω περισσότερα σημεία τότε, παίρνω τρία τυχαία σημεία υπολογίζω την ευθεία ίσων σφαλμάτων για αυτά τα σημεία και υπολογίζω στη συνέχεια τα σφάλματα των υπολοίπων σημείων. Απο αυτά βρίσκω το σημείο με το μεγαλύτερο σφάλμα, οπότε επαναλαμβάνω την ίδια διαδικασία κρατώντας τα δύο αρχικά σημεία και το νέο κ.ο.κ.

7.3 ΠΡΟΣΕΓΓΙΣΗ ΜΕ ΡΗΤΕΣ ΣΥΝΑΡΤΗΣΕΙΣ

Επιθυμούμε να προσεγγίσουμε μια συνάρτηση $f(x)$ με μια **ρητή συνάρτηση**, που θα είναι ο λόγος δυο πολυωνύμων n και m βαθμού. Εστω:

$$f(x) \approx R_N(x) \approx \frac{a_0 + a_1x + a_2x^2 + \dots + a_nx^n}{1 + b_1x + b_2x^2 + \dots + b_mx^m} \quad (7.29)$$

όπου $N = n + m$ και $R_N(x)$ είναι το ανάπτυγμα Maclaurin της $f(x)$ (δηλαδή οι $R_N(x)$ και $f(x)$ ταυτίζονται και για τις N πρώτες παραγώγους τους). Οπότε

$$\begin{aligned} f(x) - R_N(x) &\approx (c_0 + c_1x + \dots + c_Nx^N) - \frac{a_0 + a_1x + \dots + a_nx^n}{1 + b_1x + \dots + b_mx^m} \\ &= \frac{(c_0 + c_1x + \dots + c_Nx^N)(1 + b_1x + \dots + b_mx^m) - (a_0 + a_1x + \dots + a_nx^n)}{1 + b_1x + \dots + b_mx^m} \end{aligned}$$

για να είναι οι πρώτες N παράγωγοι των $f(x)$ και $R_N(x)$ ίσες για $x = 0$, θα πρέπει οι συντελεστές του πολυωνύμου στον αριθμητή ως και τάξης N να είναι μηδέν. Αυτό έχει ως συνέπεια να υπάρχουν $N + 1$ εξισώσεις για τους άγνωστους συντελεστές. Η πρώτη εξάγεται από την απαίτηση για $x = 0$ να ισχύει $f(0) = R_N(0)$ οπότε $c_0 - a_0 = 0$ ενώ μέσω των παρακάτω N εξισώσεων θα καθορισθούν οι υπόλοιποι N συντελεστές

$$\begin{aligned} b_1c_0 + c_1 - a_1 &= 0 \\ b_2c_0 + b_1c_1 + c_2 - a_2 &= 0 \\ b_3c_0 + b_2c_1 + b_1c_2 + c_3 - a_3 &= 0 \\ &\vdots \end{aligned} \quad (7.30)$$

$$\begin{aligned} b_m c_{n-m} + b_{m-1} c_{n-m+1} + \dots + c_n - a_n &= 0 \\ b_m c_{n-m+1} + b_{m-1} c_{n-m+2} + \dots + c_{n+1} &= 0 \\ b_m c_{n-m+2} + b_{m-1} c_{n-m+3} + \dots + c_{n+2} &= 0 \\ &\vdots \end{aligned} \quad (7.31)$$

Απο τις σχέσεις αυτές λύνοντας το σύστημα των N εξισώσεων για τους N αγνώστους συντελεστές υπολογίζω τη ρητή συνάρτηση. Προσέξτε ότι σε κάθε γραμμή

οι δείκτες κάθε όρου έχουν άθροισμα N . Η προσέγγιση αυτή είναι γνωστή ως **προσέγγιση Padé**.

ΠΑΡΑΔΕΙΓΜΑ

Να βρεθεί προσέγγιση Padé, $R_9(x)$, για τη συνάρτηση $\tan^{-1}(x)$ με πολυώνυμο πέμπτου βαθμού στον αριθμητή και τετάρτου στον παρονομαστή².

Το ανάπτυγμα Maclaurin της συνάρτησης $\tan^{-1}(x)$ είναι:

$$\tan^{-1}(x) = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \frac{1}{7}x^7 + \frac{1}{9}x^9$$

οπότε

$$\frac{f(x) - R_9(x)}{1 + b_1x + b_2x^2 + b_3x^3 + b_4x^4} = \frac{(x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \frac{1}{7}x^7 + \frac{1}{9}x^9) - (a_0 + a_1x + a_2x^2 + \dots + a_5x^5)}{1 + b_1x + b_2x^2 + b_3x^3 + b_4x^4}$$

Αρα καταλήγω στο παρακάτω σύστημα εξισώσεων:

$$\begin{aligned} a_0 &= 0, & a_1 &= 1, & a_2 &= b_1, & a_3 &= -\frac{1}{3} + b_2 \\ a_4 &= -\frac{1}{3}b_1 + b_3, & a_5 &= \frac{1}{5} - \frac{1}{3}b_2 + b_4 \\ \frac{1}{5}b_1 - \frac{1}{2}b_3 &= 0, & -\frac{1}{7} + \frac{1}{5}b_2 - \frac{1}{3}b_4 &= 0 \\ -\frac{1}{7}b_1 + \frac{1}{5}b_3 &= 0, & \frac{1}{9} - \frac{1}{7}b_2 + \frac{1}{5}b_4 &= 0 \end{aligned} \quad (7.32)$$

που δέχεται τις παρακάτω τιμές ως λύσεις:

$$\begin{aligned} a_0 &= 0, & a_1 &= 1, & a_2 &= 0 \\ a_3 &= \frac{7}{9}, & a_4 &= 0, & a_5 &= \frac{64}{945} \\ b_1 &= 0, & b_2 &= \frac{10}{9}, & b_3 &= 0, & b_4 &= \frac{5}{21}. \end{aligned}$$

Δηλαδή η ζητούμενη ρητή συνάρτηση είναι:

$$\tan^{-1}x \approx R_9(x) = \frac{x + \frac{7}{9}x^3 + \frac{64}{945}x^5}{1 + \frac{10}{9}x^2 + \frac{5}{21}x^4}$$

Αν τώρα θελήσουμε να εξετάσουμε την ακρίβεια της προσέγγισης με την παραπάνω ρητή συνάρτηση μπορούμε να θέσουμε μια τυχαία τιμή, έστω $x = 1$ στην παραπάνω σχέση τότε η τιμή της $R_9(1) = 0.7856$, η ακριβής τιμή είναι 0.7854 ενώ το ανάπτυγμα Maclaurin δίνει 0.8349!

Παρατηρούμε δηλαδή ότι ενώ χρησιμοποιήσαμε το ανάπτυγμα Maclaurin για τον υπολογισμό της ρητής συνάρτησης το τελικό αποτέλεσμα είναι πολύ ακριβέστερο.

² Θα μπορούσε να γραφεί και ως $R_{5,4}(x)$ αν και η επιλογή μας είναι αυθαίρετη. Θα μπορούσαμε για παράδειγμα να θεωρήσουμε πολυώνυμο 3ου βαθμού στον αριθμητή και 6ου στον παρονομαστή.

7.3.1 Προσέγγιση με ρητή συνάρτηση για ομάδα σημείων

Αν δεν έχουμε αναλυτική μορφή μιας συνάρτησης $f(x)$ αλλά δίνεται ένας συγκεκριμένος αριθμός σημείων $(x_i, f(x_i))$, έστω k , τότε στην ουσία θα ακολουθήσουμε μια διαδικασία που είναι παρόμοια με την εύρεση του πολυωνύμου Lagrange στο Κεφάλαιο 3. Δηλαδή κατασκευάζουμε ένα σύστημα k εξισώσεων της μορφής :

$$\frac{a_0 + a_1 x_i + a_2 x_i^2 + \dots + a_n x_i^n}{1 + b_1 x_i + b_2 x_i^2 + \dots + b_m x_i^m} = f(x_i) \quad (7.33)$$

με $k \geq m + n + 1$. Που τελικά οδηγεί στην ανάγκη επίλυσης ενός συστήματος της μορφής :

$$\begin{aligned} a_0 + a_1 x_1 + \dots + a_n x_1^n - (f_1 x_1) b_1 - \dots - (f_1 x_1^m) b_m &= f_1 \\ \vdots \\ a_0 + a_1 x_i + \dots + a_n x_i^n - (f_i x_i) b_1 - \dots - (f_i x_i^m) b_m &= f_i \\ \vdots \\ a_0 + a_1 x_k + \dots + a_n x_k^n - (f_k x_k) b_1 - \dots - (f_k x_k^m) b_m &= f_k \end{aligned}$$

δηλαδή k εξισώσεις για τις k άγνωστες ποσότητες a_0, a_1, \dots, a_n και b_1, b_2, \dots, b_m . Το σύστημα αυτό λύνεται σχετικά εύκολα με τις μεθόδους που αναπτύξαμε στο 2ο Κεφάλαιο.

ΠΑΡΑΔΕΙΓΜΑ

Να βρεθεί μια ρητή συνάρτηση που να επαληθεύει τα παρακάτω σημεία :

| | | | |
|-----|----|---|----|
| x | -1 | 0 | 1 |
| y | 1 | 2 | -1 |

Προφανώς ο βαθμός των πολυωνύμων στον αριθμητή και παρανομαστή θα πρέπει να είναι, αθροιστικά, το πολύ 3 ($n + m + 1 = 3$). Οπότε αν υποθέσω γραμμικά πολυώνυμα στον αριθμητή και παρανομαστή θα έχω μια ρητή συνάρτηση της μορφής :

$$R_{1,1}(x) = \frac{a_0 + a_1 x}{1 + b_1 x}$$

Θέτοντας τις τιμές του παραπάνω πίνακα δημιουργώ τις τρεις εξισώσεις που ακολουθούν :

$$\left. \begin{aligned} a_0 + (-1) a_1 - (-1) b_1 &= 1 \\ a_0 + 0 \cdot a_1 - 0 \cdot b_1 &= 2 \\ a_0 + 1 \cdot a_1 - (-1) b_1 &= -1 \end{aligned} \right\} \Rightarrow \begin{aligned} a_0 &= 2 \\ a_1 &= -1 \\ b_1 &= -2 \end{aligned}$$

Άρα η ζητούμενη ρητή συνάρτηση θα είναι :

$$R_{1,1}(x) = \frac{2 - x}{1 - 2x}.$$

Εναλλακτικά θα μπορούσα να οδηγηθώ και στην παρακάτω ρητή συνάρτηση (εάν είχα θεωρήσει δευτεροβάθμιο πολυώνυμο στον παρανομαστή)

$$R_{0,2}(x) = \frac{a_0}{1 + b_1x + b_2x^2} \Rightarrow R(x) = \frac{2}{1 - 2x - x^2}$$

ενώ προφανώς για δευτεροβάθμιο πολυώνυμο στον αριθμητή οδηγούμαστε στο συμπτωτικό πολυώνυμο:

$$R_{2,0}(x) = \frac{a_0 + a_1x + a_2x^2}{1} \Rightarrow R_{2,0}(x) = 2 - x - 2x^2.$$

7.4 ΑΣΚΗΣΕΙΣ

1. Προσδιορίστε την ευθεία ελαχίστων τετραγώνων για τα παρακάτω σημεία:

| | | | | | | | | | |
|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| x_i | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| y_i | 1.2 | 1.5 | 1.4 | 1.6 | 2.1 | 2.0 | 2.2 | 2.5 | 3.0 |

2. Προσδιορίστε την ευθεία μεγίστου-ελαχίστου (ίσων σφαλμάτων) για τα δεδομένα του προηγούμενου προβλήματος.
3. Προσδιορίστε την ευθεία μεγίστου-ελαχίστου (ίσων σφαλμάτων) για τα σημεία $(0, 0)$, $(1, 0)$ και $(2, 1)$.
4. Προσδιορίστε μια παραβολή ελαχίστων τετραγώνων που να προσεγγίζει τη συνάρτηση $y = e^{-x^2/2}$.
5. Αν εξέλξη ενός φαινομένου ακολουθεί ένα νόμο της μορφής $y(t) = ke^{at^2+bt}$ να βρεθούν οι τιμές των σταθερών k, a, b για τα δεδομένα του παρακάτω πίνακα:

| | | | | | | | | | | |
|--------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| t | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1.0 |
| $y(t)$ | 2.7 | 2.9 | 3.1 | 3.4 | 3.6 | 3.9 | 4.3 | 4.6 | 5.0 | 5.6 |

(Υπόδειξη: Λογαριθμήστε την δοθείσα σχέση και στη συνέχεια χρησιμοποιήστε παραβολή ελαχίστων τετραγώνων).

6. Υπολογίστε τη ρητή προσέγγιση της συνάρτησης $y = e^x$ με μια ρητή συνάρτηση $R_{3,3}(x)$ (δηλαδή 3ο βαθμιο πολυώνυμο σε αριθμητή και παρανομαστή). Συγκρίνετε με το ανάπτυγμα Maclaurin 6ης τάξης για $x = 1$. Τι παρατηρείτε;
7. Επαναλάβετε την παραπάνω διαδικασία για τις συναρτήσεις $y = \cos(x)$ και $y = \sin(x)$ και να βρείτε το $R_{3,5}(x)$ ενώ για την $y = 1/x \sin(x)$ να βρείτε το $R_{4,6}(x)$.
8. Βρείτε μια ρητή συνάρτηση που να επαληθεύει τις τιμές:

| | | | |
|------|------|------|------|
| 0 | 1 | 2 | 4 |
| 0.83 | 1.06 | 1.25 | 4.15 |

ΔΙΑΦΟΡΙΚΕΣ ΕΞΙΣΩΣΕΙΣ ΜΕ ΜΕΡΙΚΕΣ ΠΑΡΑΓΩΓΟΥΣ

Πολλά σημαντικά επιστημονικά προβλήματα στο χώρο της φυσικής περιγράφονται από διαφορικές εξισώσεις με μερικές παραγώγους (ΔΕΜΠ). Συνήθως το φυσικό φαινόμενο που μελετάμε παριστάνεται από μια συνάρτηση περισσοτέρων της μιας μεταβλητών που ικανοποιεί συγκεκριμένη μορφή εξίσωσης. Τα περισσότερα επιστημονικά προβλήματα μάλιστα περιγράφονται από διαφορικές εξισώσεις των οποίων η ανώτερης τάξης παράγωγος είναι δεύτερης τάξης. Για παράδειγμα αν ψ είναι μια συνάρτηση δύο μεταβλητών x και y τότε υπάρχουν τρεις μερικές παράγωγοι της δεύτερης τάξης

$$\frac{\partial^2 \psi}{\partial x^2}, \quad \frac{\partial^2 \psi}{\partial x \partial y}, \quad \frac{\partial^2 \psi}{\partial y^2} \quad (8.1)$$

Με βάση τις τιμές των συντελεστών των παραγώγων δεύτερης τάξης κατατάσσουμε τις διαφορικές εξισώσεις με μερικές παραγώγους σε **ελλειπτικές**, **παραβολικές**, **υπερβολικές**. Μια ΔΕΜΠ δεύτερης τάξης είναι της μορφής

$$A \frac{\partial^2 \psi}{\partial x^2} + B \frac{\partial^2 \psi}{\partial x \partial y} + C \frac{\partial^2 \psi}{\partial y^2} + D \left(x, y, \psi, \frac{\partial \psi}{\partial x}, \frac{\partial \psi}{\partial y} \right) = 0 \quad (8.2)$$

Ανάλογα με την τιμή της ποσότητας $B^2 - 4AC$ οι ΔΕΜΠ κατατάσσονται ως:

- **Ελλειπτικές**, αν $B^2 - 4AC < 0$
- **Παραβολικές**, αν $B^2 - 4AC = 0$
- **Υπερβολικές**, αν $B^2 - 4AC > 0$

Αν οι συντελεστές A , B και C είναι συναρτήσεις των x , y τότε είναι δυνατόν μια ΔΕΜΠ να αλλάζει 'κατηγορία' σε διάφορες περιοχές του πεδίου ορισμού της.

Δυο κλασσικές διαφορικές εξισώσεις με μερικές παραγώγους η *εξίσωση Laplace*

$$\nabla^2 \psi = 0 \quad (8.3)$$

και η *εξίσωση Poisson*

$$\nabla^2 \psi = f(x, y) \quad (8.4)$$

έχουν $B = 0$, $A = C = 1$, και επομένως είναι πάντοτε **ελλειπτικές**.

Η κυματική εξίσωση

$$\nabla^2 \psi - \frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} = 0 \quad (8.5)$$

είναι κλασσική περίπτωση **υπερβολικής** ΔΕΜΠ

Ενώ η εξίσωση της θερμότητας

$$\sigma \nabla^2 \psi - \frac{\partial \psi}{\partial t} = 0 \quad (8.6)$$

είναι χαρακτηριστική περίπτωση **παραβολικής** ΔΕΜΠ.

Για την επίλυση των παραπάνω προβλημάτων απαιτούνται κατ' αρχάς οι κατάλληλες συνοριακές συνθήκες ή/και κατάλληλες αρχικές συνθήκες. Στη συνέχεια υπάρχουν διάφορες τεχνικές αναλυτικής, ημι-αναλυτικής και αριθμητικής επίλυσης τους. Σε αυτό το κεφάλαιο θα ασχοληθούμε με χαρακτηριστικά προβλήματα που συναντώνται στη Φυσική και περιγράφονται με ΔΕΜΠ. Θα παρουσιάσουμε τεχνικές ημι-αναλυτικής και αριθμητικής επίλυσης τους.

8.1 ΕΛΛΕΙΠΤΙΚΕΣ ΕΞΙΣΩΣΕΙΣ

8.1.1 Η Εξίσωση Laplace

Η εξίσωση Laplace σε δύο διαστάσεις

$$\nabla^2 u \equiv u_{xx} + u_{yy} = 0 \quad (8.7)$$

είναι μια κλασσική εξίσωση της Φυσικής και απαντάται σε όλα τα προβλήματα που σχετίζονται με πεδία και δυναμικά. Με βάση τους ορισμούς που δώσαμε στη αρχή του κεφαλαίου είναι μια **ελλειπτική διαφορική εξίσωση με μερικές παραγώγους**. Και όσα θα αναπτύξουμε μπορούν εύκολα να γενικευθούν στη **εξίσωση Poisson**

$$\nabla^2 u \equiv u_{xx} + u_{yy} = g(x, y) \quad (8.8)$$

είτε στην **εξίσωση Helmholtz**

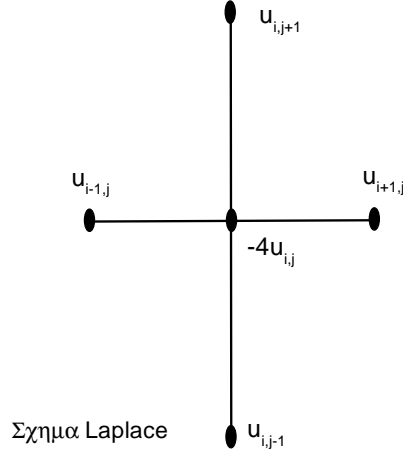
$$u_{xx} + u_{yy} + f(x, y)u = g(x, y) \quad (8.9)$$

όπου η $g(x, y)$ και η $f(x, y)$ είναι δοθείσες συναρτήσεις.

Στο κεφάλαιο αυτό θα ασχοληθούμε με την αριθμητική επίλυση της εξίσωσης Laplace. Αριθμητική επίλυση σημαίνει, αριθμητικό προσδιορισμό των παραγώγων της $u(x, y)$. Αν χρησιμοποιήσουμε τη σχέση (4.18) τότε η u_{xx} στο σημείο (x_i, y_j) θα γραφεί ως

$$[u_{xx}]_{i,j} = \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{h^2} \quad (8.10)$$

και αντίστοιχα η u_{yy}



Σχήμα 8.1. Το σχήμα των πέντε σημείων για την αριθμητική επίλυση της εξίσωσης Laplace.

$$[u_{yy}]_{i,j} = \frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{h^2} \quad (8.11)$$

οπότε η εξίσωση Laplace θα γραφεί ως

$$\nabla^2 u \approx \frac{u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1} - 4u_{i,j}}{h^2} = 0 \quad (8.12)$$

όπου $i = 2, \dots, n-1$ και $j = 2, \dots, m-1$. Αυτή η σχέση είναι γνωστή ως **τύπος των πέντε σημείων για την εξίσωση Laplace**, και δίνει αποτελέσματα με σφάλμα της τάξης του $\sim h^2$. Η σχέση (8.12) συνδέει στη ουσία την τιμή της συνάρτησης $u(x, y)$ στη θέση (x_i, y_j) με τα τέσσερα γειτονικά της σημεία (x_{i+1}, y_j) , (x_{i-1}, y_j) , (x_i, y_{j-1}) και (x_i, y_{j+1}) , όπως φαίνεται στο Σχήμα 8.1. Βασικά, η σχέση (8.12) μπορεί να γραφεί και στη μορφή

$$u_{i,j} = \frac{1}{4} (u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1}) \quad (8.13)$$

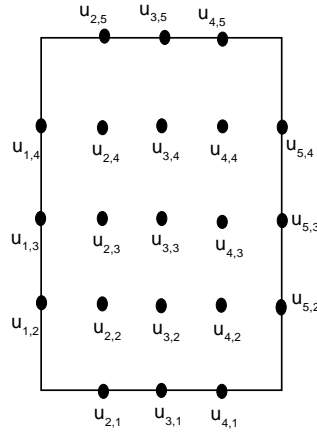
που σημαίνει ότι η τιμή της $u(x, y)$ στη θέση (x_i, y_j) είναι ο αριθμητικός μέσος των τεσσάρων γειτονικών της σημείων (x_{i+1}, y_j) , (x_{i-1}, y_j) , (x_i, y_{j-1}) και (x_i, y_{j+1}) .

Ας υποθέσουμε τώρα πως δίνονται οι συνοριακές τιμές στα ακόλουθα σημεία του πλέγματος

$$\begin{aligned} u(x_1, y_j) &= u_{1,j} \quad \text{για } 2 \leq j \leq m-1 \\ u(x_i, y_1) &= u_{i,1} \quad \text{για } 2 \leq i \leq n-1 \end{aligned}$$

$$\begin{aligned} u(x_n, y_j) &= u_{n,j} \quad \text{για } 2 \leq j \leq m-1 \\ u(x_i, y_m) &= u_{i,m} \quad \text{για } 2 \leq i \leq n-1 \end{aligned} \quad (8.14)$$

Με βάση τη γνώση των παραπάνω συνοριακών τιμών μπορούμε να υπολογίσουμε όλες τις υπόλοιπες τιμές της συνάρτησης $u(x, y)$, λύνοντας στην ουσία ένα σύστημα $(n-2) \times (n-2)$ εξισώσεων για $(n-2)^2$ αγνώστους.



Σχήμα 8.2. Το 5×5 πλέγμα για την αριθμητική επίλυση της εξίσωσης Laplace.

Για παράδειγμα στο πρόβλημα του σχήματος 8.2 έχουμε να υπολογίσουμε τις 9 τιμές της συνάρτησης $u(x, y)$, δηλαδή τα $u_{2,2}$, $u_{2,3}$, $u_{2,4}$, $u_{3,2}$, $u_{3,3}$, $u_{3,4}$, $u_{4,2}$, $u_{4,3}$ και $u_{4,4}$. Αν εφαρμόσουμε τη σχέση (8.13) για κάθε ένα από τα παραπάνω σημεία καταλήγουμε σε ένα σύστημα 9 εξισώσεων για τους 9 αγνώστους που προαναφέραμε.

$$\begin{aligned} -4u_{2,2} + u_{3,2} + u_{2,3} &= -u_{2,1} - u_{1,2} \\ u_{2,2} - 4u_{3,2} + u_{4,2} + u_{3,3} &= -u_{3,1} \\ u_{3,2} - 4u_{4,2} + u_{4,3} &= -u_{4,1} - u_{5,2} \\ u_{2,2} - 4u_{2,3} + u_{3,3} + u_{2,4} &= -u_{1,3} \\ u_{3,2} + u_{2,3} - 4u_{3,3} + u_{4,3} + u_{3,4} &= 0 \\ u_{4,2} + u_{3,3} - 4u_{4,3} + u_{4,4} &= -u_{5,3} \\ u_{2,3} - 4u_{2,4} + u_{3,4} &= -u_{2,5} - u_{1,4} \\ u_{3,3} + u_{2,4} - 4u_{3,4} + u_{4,4} &= -u_{3,5} \\ u_{4,3} + u_{3,4} - 4u_{4,4} &= -u_{4,5} - u_{5,4} \end{aligned} \quad (8.15)$$

Η λύση αυτού του συστήματος θα μας δώσει τις ζητούμενες τιμές στα εσωτερικά σημεία του πλέγματος.

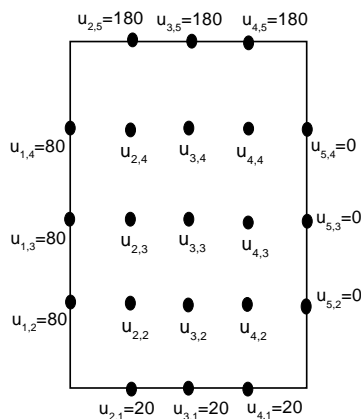
ΠΑΡΑΔΕΙΓΜΑ

Να βρεθεί η λύση της εξίσωσης Laplace $\nabla^2 u = 0$ για την περιοχή $\{(x, y) : 0 \leq x \leq 4, 0 \leq y \leq 4\}$ όπου $u(x, y)$ συμβολίζει τη θερμοκρασία στο σημείο (x, y) . Οι συνοριακές τιμές είναι:

$$\begin{aligned} u(x, 0) = 20 \quad \text{και} \quad u(x, 4) = 180 \quad \text{για} \quad 0 < x < 4 \\ u(0, y) = 80 \quad \text{και} \quad u(4, y) = 0 \quad \text{για} \quad 0 < y < 4 \end{aligned}$$

ΛΥΣΗ

Κατάρχας δημιουργούμε το πλέγμα του σχήματος 3 απο όπου καταλήγουμε εύκολα στο παρακάτω σύστημα 9 εξισώσεων με 9 αγνώστους

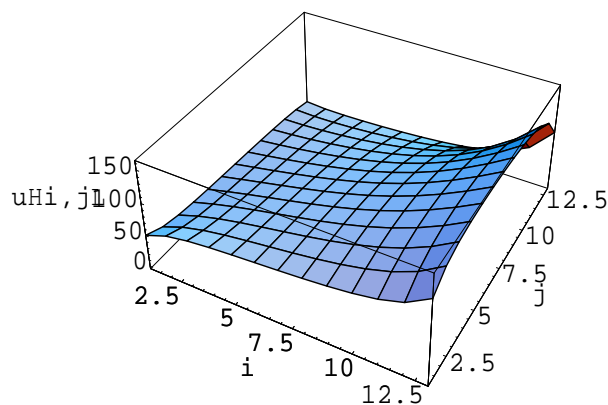


Σχήμα 8.3. Το 5×5 πλέγμα για το συγκεκριμένο παράδειγμα.

$$\begin{aligned} -4u_{2,2} + u_{3,2} + u_{2,3} &= -100 \\ u_{2,2} - 4u_{3,2} + u_{4,2} + u_{3,3} &= -20 \\ u_{3,2} - 4u_{4,2} + u_{4,3} &= -20 \\ u_{2,2} - 4u_{2,3} + u_{3,3} + u_{2,4} &= -80 \\ u_{3,2} + u_{2,3} - 4u_{3,3} + u_{4,3} + u_{3,4} &= 0 \\ u_{4,2} + u_{3,3} - 4u_{4,3} + u_{4,4} &= 0 \\ u_{2,3} - 4u_{2,4} + u_{3,4} &= -260 \\ u_{3,3} + u_{2,4} - 4u_{3,4} + u_{4,4} &= -180 \\ u_{4,3} + u_{3,4} - 4u_{4,4} &= -180 \end{aligned} \quad (8.16)$$

Η λύση του συστήματος μας οδηγεί στη λύση:

$$u_{2,2} = 55.7143, u_{3,2} = 43.2143, u_{4,2} = 27.1429, u_{2,3} = 79.6429, u_{3,3} = 70.000, \\ u_{4,3} = 45.3571, u_{2,4} = 112.857, u_{3,4} = 111.786 \text{ και } u_{4,4} = 84.2857.$$



Σχήμα 8.4. Το γράφημα της λύσης της εξίσωσης Laplace για το παράδειγμα. Έχει χρησιμοποιηθεί πλέγμα 15×15 .

Σημαντική βελτίωση στην ακρίβεια της μεθόδου μπορεί να επιτευχθεί με την αντικατάσταση των παραγώγων της εξίσωσης Laplace με βάση τη σχέση (4.18). Οπότε η εξίσωση Laplace θα γραφεί ως

$$\nabla^2 u \approx \frac{u_{i-1,j+1} + 4u_{i,j+1} + u_{i+1,j+1} + 4u_{i-1,j} - 20u_{i,j} + 4u_{i+1,j} + u_{i-1,j-1} + 4u_{i,j-1} + u_{i+1,j-1}}{6h^2} = 0 \quad (8.17)$$

όπου $i = 2, \dots, n-1$ και $j = 2, \dots, m-1$. Αυτή η σχέση είναι γνωστή ως **τύπος των εννέα σημείων για την εξίσωση Laplace**, και δίνει αποτελέσματα με σφάλμα της τάξης του $\sim h^4$.

ΠΡΟΒΛΗΜΑΤΑ

1. Επιλύστε αριθμητικά την εξίσωση Laplace για ένα τετράγωνο με συνοριακές συνθήκες

$$u(0, y) = 0 \quad \text{όταν} \quad 0 \leq y \leq 1$$

$$u(1, y) = 0 \quad \text{όταν} \quad 0 \leq y \leq 1$$

$$u(x, 1) = 0 \quad \text{όταν} \quad 0 \leq x \leq 1$$

$$u(x, 0) = x(1 - x) \quad \text{όταν} \quad 0 \leq x \leq 1$$

(Υπόδειξη: για πλέγμα 4×4 σημείων η λύση είναι $u_{2,2} = 1/12$, $u_{3,2} = 1/12$, $u_{2,3} = 1/36$, $u_{3,3} = 1/36$). Βρείτε τη λύση για πλέγμα 8×8 σημείων.

- Χρησιμοποιήστε τη σχέση (4.18) για να αποδείξετε το σχήμα 9 σημείων, δηλαδή την εξίσωση (8.17), για την επίλυση της εξίσωσης Laplace. Στη συνέχεια να τη χρησιμοποιήσετε για την επίλυση του λυμένου παραδείγματος.
- Χρησιμοποιήστε τη μέθοδο των 9 σημείων για την επίλυση του λυμένου παραδείγματος με ένα πλέγμα 20×20 σημείων.
- Σε μια ορθογώνια περιοχή $6\text{cm} \times 8\text{cm}$ η συνάρτηση $u(x, y)$ στα σύνορα της είναι μηδέν (0) και υπακούει στο νόμο

$$\nabla^2 u + 2 = 0.$$

Υπολογίστε τις τιμές της $u(x, y)$, ανά 1cm στο εσωτερικό της περιοχής. (Ουσιαστικά εδώ λύνουμε την εξίσωση Poisson). Υπόδειξη: αν την λύνετε για πλέγμα εύρους 2cm τότε θα βρίσκατε τις λύσεις $u_{2,2} = u_{3,2} = u_{2,4} = u_{3,4} = 4.56$ και $u_{2,3} = u_{3,3} = 5.72$

8.2 ΥΠΕΡΒΟΛΙΚΕΣ ΕΞΙΣΩΣΕΙΣ

8.2.1 Αριθμητική Επίλυση της Κυματικής Εξίσωσης

Σε αυτή την ενότητα θα μελετήσουμε μια 'πλήρως' αριθμητική επίλυση της κυματικής εξίσωσης και για λόγους απλότητας θα περιορισθούμε σε ένα μονοδιάστατο πρόβλημα. Πιο συγκεκριμένα θα μελετήσουμε την εξίσωση που περιγράφει μία παλλόμενη χορδή μήκους a με πακτωμένα τα άκρα της. Η κίνηση της περιγράφεται από την εξίσωση:

$$\frac{\partial^2 u(x, t)}{\partial t^2} = c^2 \frac{\partial^2 u(x, t)}{\partial x^2} \quad \text{για} \quad 0 < x < a \quad \text{και} \quad 0 < t < b \quad (8.18)$$

με οριακές συνθήκες

$$\begin{aligned} u(0, t) = 0 \quad \text{και} \quad u(a, t) = 0 & \quad \text{για} \quad 0 \leq t \leq b \\ u(x, 0) = f(x) & \quad \text{για} \quad 0 \leq x \leq a \\ u_t(x, 0) = g(x) & \quad \text{για} \quad 0 < x < a \end{aligned} \quad (8.19)$$

Παρότι υπάρχει αναλυτική λύση της εξίσωσης με βάση τη μέθοδο που χρησιμοποιήσαμε στην προηγούμενη ενότητα, με χρήση των σειρών Fourier, εδώ θα ακολουθήσουμε μια πλήρως αριθμητική τεχνική που αποτελεί βάση για τη μέλξη και πλέον σύνθετων υπερβολικών ΔΕΜΠ που δεν επιδέχονται αναλυτικών λύσεων.

Δημιουργία των εξισώσεων διαφορών

Ας διαμερίσουμε ένα ορθογώνιο με πλευρές μήκους a και b σε $(m-1) \times (n-1)$ υποδιαίρεσεις μήκους $\Delta x = h$ και $\Delta t = k$ όπως στη Σχήμα 8.5.

Προφανώς με βάση τις αρχικές συνθήκες ισχύει $u(x_i, t_1) = f(x_1)$. Με βάση την εξίσωση των κεντρικών διαφορών (4.18) οι δεύτεροι παράγωγοι u_{tt} και u_{xx} γράφονται

$$u_{tt}(x, t) = \frac{u(x, t+k) - 2u(x, t) + u(x, t-k))}{k^2} + O(k^2) \quad (8.20)$$

$$u_{xx}(x, t) = \frac{u(x+h, t) - 2u(x, t) + u(x-h, t))}{h^2} + O(h^2) \quad (8.21)$$

και επειδή διαμερίσαμε τα διαστήματα $(0, a)$ και $(0, b)$ σε ίσα τμήματα δηλαδή: $x_{i+1} = x_i + h$ και $t_{j+1} = t_j + k$ γράφουμε την εξίσωση (8.18) ως εξίσωση διαφορών στη μορφή:

$$\frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{k^2} = c^2 \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} \quad (8.22)$$

που αποτελεί την προσεγγιστική μορφή της εξίσωσης (8.18). Για λόγους απλότητας θα χρησιμοποιήσουμε την αντικατάσταση $r = ck/h$ και η παραπάνω σχέση γράφεται

$$u_{i,j+1} - 2u_{i,j} + u_{i,j-1} = r^2 (u_{i+1,j} - 2u_{i,j} + u_{i-1,j}). \quad (8.23)$$

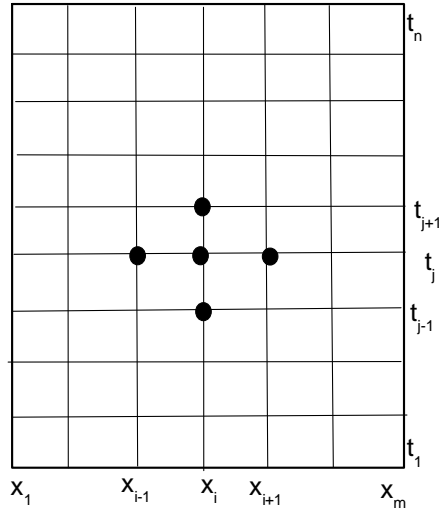
Η εξίσωση αυτή θα χρησιμοποιηθεί τώρα για τον υπολογισμό των τιμών της $u(x, t)$ στη γραμμή $j+1$ με βάση την πληροφορία που έχουμε από τις γραμμές j και $j-1$, Σχήμα 8.6. Οπότε αναδιατάσσοντας την παραπάνω σχέση λαμβάνουμε:

$$u_{i,j+1} = 2(1 - r^2)u_{i,j} + r^2 (u_{i+1,j} + u_{i-1,j}) - u_{i,j-1} \quad \text{για } i = 2, 3, \dots, n-1 \quad (8.24)$$

Επομένως χρησιμοποιώντας την σχέση αυτή μπορούμε να υπολογίσουμε τις τιμές της συνάρτησης $u(x, t)$ στο ορθογώνιο $R = \{(x, t) : 0 \leq x \leq a, 0 \leq t \leq b\}$.

Στις αριθμητικές μεθόδους υπάρχουν 'αστάθειες' με αποτέλεσμα η λύση που υπολογίζουμε να μην αντιστοιχεί πάντα στην πραγματική. Στη συγκεκριμένη περίπτωση η συνθήκη που εξασφαλίζει αριθμητική ευστάθεια είναι $r = ck/h \leq 1$. Αυτό το κριτήριο ευστάθειας είναι γνωστό στη βιβλιογραφία ως *Courant-Friedrichs-Lewy* κριτήριο ή πιο απλά *κριτήριο Courant*. Παρότι υπάρχει μαθηματική απόδειξη για το συγκεκριμένο κριτήριο θα προσπαθήσουμε να το κατανοήσουμε με βάση μια απλή ανάλυση. Πιο συγκεκριμένα, η ποσότητα $u_{i,j+1}$ υπολογίζεται από την πληροφορία που έχουμε στα σημεία $(i-1, j)$, (i, j) , $(i, j+1)$ και $(i, j-1)$, με άλλα λόγια **μόνο** από αυτά τα σημεία επιτρέπεται να μεταφέρεται πληροφορία στο σημείο $u_{i,j+1}$. Αν παρατηρήσουμε προσεκτικά αυτό που απαιτεί το κριτήριο Courant είναι:

$$|c| \leq \frac{\Delta x}{\Delta t} \quad (8.25)$$



Σχήμα 8.5. Ο τρόπος με τον οποίο ‘συνδέονται’ τα 5 σημεία της σχέσης

Δηλαδή η ταχύτητα με την οποία ταξιδεύει μια κύμανση (c) δεν θα πρέπει να υπερβαίνει τη μέγιστη ταχύτητα διάδοσης που επιτρέπεται από την διαμέριση που επιλέξαμε. Αυτό αποτελεί ένα γενικό κριτήριο για την ευστάθεια των αριθμητικών μεθόδων π.χ. δεν επιτρέπεται να επιλέγουμε μεγάλα αριθμητικά βήματα για να μελετήσουμε φαινόμενα που παρουσιάζουν σημαντικές αλλαγές σε μικρότερη χρονική κλίμακα. Ένα τέτοιο παραδειγμα θα αποτελούσε η αριθμητική μελέτη της τροχιάς της Γης με χρονικό βήμα μεγαλύτερο του ενός έτους, ή η μελέτη της κίνησης ενός εκκρεμούς με χρονικό βήμα μεγαλύτερο της περιόδου του.

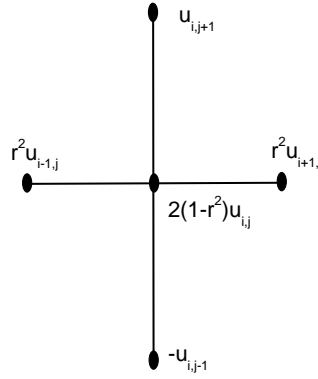
Οι αρχικές τιμές

Όπως προαναφέραμε απαιτούνται οι αρχικές τιμές της $u(x, t)$ στις γραμμές $j = 1$ και $j = 2$. Επειδή όμως η δεύτερη γραμμή δεν δίνεται πάντα την υπολογίζουμε με βάση τις αρχικές συνθήκες. Αν χρησιμοποιήσουμε πρώτης τάξης ανάπτυγμα Taylor της $u(x, t)$ περι το σημείο $(x_i, 0)$ λαμβάνουμε

$$u(x_i, k) = u(x_i, 0) + ku_t(x_i, 0) + O(k^2) \quad (8.26)$$

Οπότε επειδή $u(x_i, 0) = f(x_i) = f_i$ και $u_t(x_i, 0) = g(x_i) = g_i$ η παραπάνω σχέση γράφεται

$$u_{i,2} = f_i + kg_i \quad \text{για} \quad i = 2, 3, \dots, n-1. \quad (8.27)$$



Σχήμα 8.6. Η διαμέριση(πλέγμα) για την αριθμητική επίλυση της κυματικής εξίσωσης $u_{tt} = c^2 u_{xx}$.

Προφανώς η παραπάνω σχέση είναι προσεγγιστική και το αριθμητικό σφάλμα που υπεισέρχεται μέσω της παραπάνω σχέσης είναι τάξης $O(k^2)$ το οποίο διαδίδεται κατά την διάρκεια της αριθμητικής εξέλιξης της εξίσωσης και γιαυτό είναι σημαντικό να θέτουμε απο την αρχή αρκετά μικρή τιμή στο k .

ΠΑΡΑΔΕΙΓΜΑ

Να λυθεί η κυματική εξίσωση της παλλόμενης χορδής

$$u_{tt}(x, t) = 4u_{xx}(x, t) \quad \text{για} \quad 0 < x < 1 \quad \text{και} \quad 0 < t < 0.5 \quad (8.28)$$

με οριακές συνθήκες

$$\begin{aligned} u(0, t) &= 0 \quad \text{και} \quad u(1, t) = 0 \quad \text{για} \quad 0 \leq t \leq 0.5 \\ u(x, 0) &= f(x) = \sin(\pi x) + \sin(2\pi x) \quad \text{για} \quad 0 \leq x \leq 1 \\ u_t(x, 0) &= g(x) = 0 \quad \text{για} \quad 0 < x < 1 \end{aligned} \quad (8.29)$$

ΛΥΣΗ

Για απλότητα θα χρησιμοποιήσουμε $h = 0.1$ και $k = 0.05$. Επειδή $c = 2$ λαμβάνουμε $r = ck/h = 1$, επίσης επειδή $g(x) = 0$ και $r = 1$ η σχέση (8.28) γράφεται ως:

$$u_{i,2} = \frac{f_{i-1} + f_{i+1}}{2} \quad \text{για} \quad i = 2, 3, \dots, 9. \quad (8.30)$$

Αντικαθιστώντας $r = 1$ στην εξίσωση (8.24) δημιουργούμε την αρκετά απλή αναδρομική σχέση

$$u_{i,j+1} = u_{i+1,j} + u_{i-1,j} - u_{i,j-1} \quad (8.31)$$

ΠΡΟΒΛΗΜΑΤΑ

1. Να λυθεί η κυματική εξίσωση της παλλόμενης χορδής

$$u_{tt}(x, t) = 4u_{xx}(x, t) \quad \text{για } 0 < x < 1 \quad \text{και} \quad 0 < t < 0.5 \quad (8.32)$$

με οριακές συνθήκες

$$\begin{aligned} u(0, t) &= 0 \quad \text{και} \quad u(1, t) = 0 && \text{για } 0 \leq t \leq 1 \\ u(x, 0) &= f(x) = \frac{5x}{3} && \text{για } 0 \leq x \leq \frac{3}{5} \\ u(x, 0) &= f(x) = \frac{5(1-x)}{2} && \text{για } \frac{3}{5} \leq x \leq 1 \\ u_t(x, 0) &= g(x) = 0 && \text{για } 0 < x < 1 \end{aligned}$$

2. Να επιβεβαιωθεί αναλυτικά ότι η $u(x, t) = \sin(n\pi x) \cos(cn\pi t)$ είναι λύση της εξίσωσης $u_{tt}(x, t) = c^2 u_{xx}(x, t)$ για κάθε θετικό ακέραιο $n = 1, 2, \dots$
 3. Να λυθεί η κυματική εξίσωση της παλλόμενης χορδής

$$u_{tt}(x, t) = u_{xx}(x, t) \quad \text{για } 0 < x < 1 \quad \text{και} \quad 0 < t < 1 \quad (8.33)$$

με οριακές συνθήκες

$$\begin{aligned} u(0, t) &= 0 \quad \text{και} \quad u(1, t) = 0 && \text{για } 0 \leq t \leq 1 \\ u(x, 0) &= f(x) = \sin(\pi x) && \text{για } 0 \leq x \leq 1 \\ u_t(x, 0) &= g(x) = 0 && \text{για } 0 < x < 1 \end{aligned}$$

4. Δοκιμάστε στα παραπάνω προβλήματα τιμές του r μεγαλύτερες και μικρότερες του 1 για να μελετήσετε την ευστάθεια της μεθόδου.

8.3 ΠΑΡΑΒΟΛΙΚΕΣ ΕΞΙΣΩΣΕΙΣ

Σε αυτό το κεφάλαιο θα δώσουμε μια μέθοδο αναλυτικής επίλυσης μιας μονοδιάστατης παραβολικής διαφορική εξίσωσης με μερικές παραγώγους.

Εστω η ΔΕΜΠ

$$\frac{\partial u}{\partial t} = \alpha^2 \frac{\partial^2 u}{\partial x^2} \quad \text{για } 0 \leq x \leq 1 \quad \text{και} \quad 0 \leq t < \infty \quad (8.34)$$

με οριακές συνθήκες :

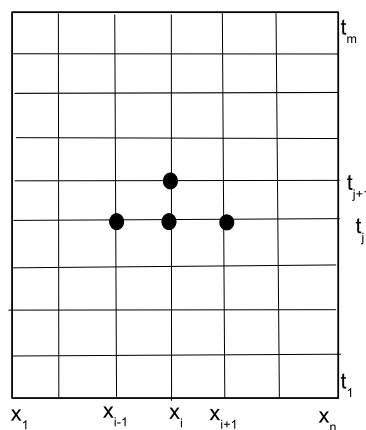
$$u(0, t) = 0, \quad u(1, t) = 0 \quad \text{για } 0 \leq t < \infty \quad (8.35)$$

και αρχικές συνθήκες :

$$u(x, 0) = \phi(x), \quad \text{για } 0 \leq x \leq 1. \quad (8.36)$$

8.3.1 Αριθμητική Λύση: Εξίσωση Διαφορών

Ας υποθέσουμε ότι το διάστημα $0 \leq x \leq 1$ υποδιαιρείται σε $n-1$ υποδιαστήματα $\Delta x = h$ και ας υποθέσουμε ότι στη χρονική διεύθυνση ότι θα λύσουμε τη ΔΕΜΠ έως τη χρονική στιγμή $t = T$ οπότε θα θεωρήσουμε $m-1$ υποδιαίρεσεις στο διάστημα $0 \leq t \leq T$, Σχήμα 8.7. Αντικαθιστώντας τις μερικές παραγώγους u_t



Σχήμα 8.7. Η διαμέριση(πλέγμα) για την αριθμητική επίλυση της παραβολικής εξίσωσης.

και u_{xx} από τις εξισώσεις διαφορών

$$u_t(x, t) = \frac{u(x, t+k) - u(x, t)}{k} + O(k) \quad (8.37)$$

$$u_{xx}(x, t) = \frac{u(x+h, t) - 2u(x, t) + u(x-h, t)}{h^2} + O(h^2) \quad (8.38)$$

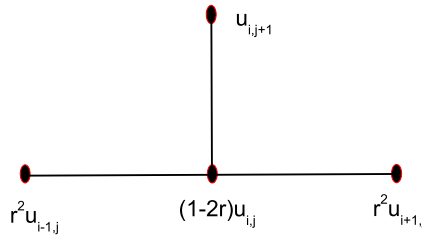
οπότε αντικαθιστούμε την εξίσωση (8.34) με την εξίσωση διαφορών

$$\frac{u_{i,j+1} - u_{i,j}}{k} = \alpha^2 \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{h^2} \quad (8.39)$$

οπότε αν θέσουμε $r = \alpha^2 k / h^2$ στην παραπάνω σχέση δημιουργούμε την εξίσωση διαφορών

$$u_{i,j+1} = (1 - 2r) u_{i,j} + r (u_{i-1,j} + u_{i+1,j}) \quad (8.40)$$

Η εξίσωση αυτή δίνει τις τιμές της συνάρτησης $u(t, x)$ στην γραμμή $j+1$ σαν συνάρτηση των τιμών της συνάρτησης στη γραμμή j , όπως φαίνεται στο Σχήμα 8.8.



Σχήμα 8.8. Το αριθμητικό σχήμα για την αριθμητική επίλυση της παραβολικής εξίσωσης.

Το παραπάνω αριθμητικό σχήμα είναι εξαιρετικά απλό και ο προγραμματισμός του σε οποιαδήποτε γλώσσα σχεδόν τετριμμένος. Παρ' όλα ταύτα θα πρέπει να εξετάσουμε την ευστάθειά του, πιο συγκεκριμένα το σχήμα είναι ευσταθές για $0 \leq r \leq 1/2$. Δηλαδή, το χρονικό βήμα θα πρέπει να καθορίζεται από τη σχέση $k \leq h^2 / (2\alpha^2)$. Ειδικά η επιλογή $r = 1/2$ απλοποιεί το αριθμητικό σχήμα (8.40) που γράφεται στη μορφή:

$$u_{i,j+1} = \frac{u_{i-1,j} + u_{i+1,j}}{2} \quad (8.41)$$

Θα πρέπει να παρατηρήσουμε ότι το σχήμα που παρουσιάσαμε είναι πρώτης τάξης, $O(k)$, ως προς το χρόνο και δεύτερης τάξης, $O(h^2)$, ως προς το χώρο.

8.3.2 Αριθμητική Λύση: Μέθοδος Crank-Nicholson

Η μέθοδος των Crank και Nicholson είναι μια **έμμεση** μέθοδος αριθμητικής επίλυσης παραβολικών ΔΕΜΠ.

Σ' αυτήν τη μέθοδο η τιμή της συνάρτησης $u(x, y)$ σε κάποιο σημείο (x_i, y_j) δεν υπολογίζεται απεύθείας αλλά μέσω της επίλυσης ενός γραμμικού συστήματος εξισώσεων. Πιο συγκεκριμένα στην διακριτοποίηση της 2ης χωρικής παραγώγου της $u(x, y)$ (δηλαδή της u_{xx}) αντί να χρησιμοποιήσουμε τα στοιχεία της γραμμής j που είναι γνωστά, χρησιμοποιούμε και τα άγνωστα στοιχεία της γραμμής $j + 1$ όπως φαίνεται στο Σχήμα 8.9. Για την ακρίβεια παίρνουμε το μέσο όρο των διακριτοποιημένων τιμών της u_{xx} όπως φαίνεται στην παρακάτω σχέση:

$$\frac{u_{i,j+1} - u_{i,j}}{k} = \frac{1}{2} \alpha^2 \left(\frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{h^2} + \frac{u_{i-1,j+1} - 2u_{i,j+1} + u_{i+1,j+1}}{h^2} \right) \quad (8.42)$$

Είναι προφανές επομένως ότι στη σχέση αυτή υπάρχουν τρεις άγνωστες ποσότητες, τα $u_{i-1,j+1}$, $u_{i,j+1}$ και $u_{i+1,j+1}$. Αν αναδιατάξουμε την παραπάνω σχέση διατηρώντας από αριστερά τις άγνωστες τιμές και από δεξιά τις γνωστές τιμές καταλήγουμε στη σχέση:

$$-ru_{i-1,j+1} + 2(1+r)u_{i,j+1} - ru_{i+1,j+1} = 2(1-r)u_{i,j} + r(u_{i-1,j} + u_{i+1,j}) \quad (8.43)$$

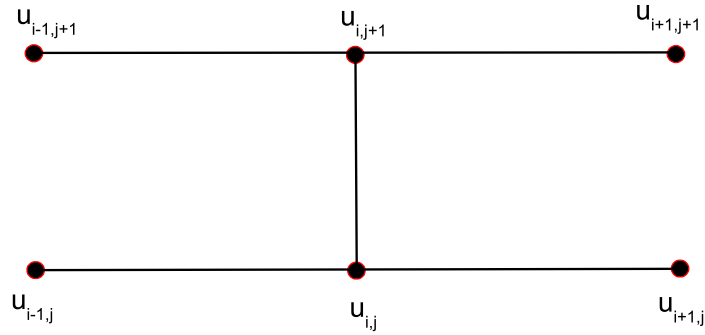
που σημαίνει ότι για τον υπολογισμό $u(x, y)$ στη χρονική στιγμή $j \cdot k$ θα πρέπει να λύσουμε ένα σύστημα $n - 2$ τέτοιων εξισώσεων. Για απλότητα θα υποθέσουμε ότι $r = 1$ οπότε η παραπάνω σχέση παίρνει την απλή μορφή:

$$-u_{i-1,j+1} + 4u_{i,j+1} - u_{i+1,j+1} = u_{i-1,j} + u_{i+1,j} \quad (8.44)$$

που οδηγεί στο σύστημα:

$$\begin{pmatrix} 4 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 4 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 4 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 4 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 4 \end{pmatrix} \begin{pmatrix} u_{2,j+1} \\ u_{3,j+1} \\ \dots \\ u_{k,j+1} \\ \dots \\ u_{n-2,j+1} \\ u_{n-1,j+1} \end{pmatrix} = \begin{pmatrix} 2c_1 + u_{3,j} \\ u_{2,j} + u_{4,j} \\ \dots \\ u_{k-1,j} + u_{k+1,j} \\ \dots \\ u_{n-3,j} + u_{n-1,j} \\ u_{n-2,j} + 2c_2 \end{pmatrix} \quad (8.45)$$

του οποίου η λύση μας δίνει τις τιμές της $u(x, y)$ για όλα τα x_i τη χρονική στιγμή $j \cdot k$.



Σχήμα 8.9. Το αριθμητικό σχήμα για την αριθμητική επίλυση της παραβολικής εξίσωσης $u_t = \alpha^2 u_{xx}$ με τη μέθοδο Crank-Nicholson.

Προφανώς, η διαδικασία αυτή είναι σημαντικά πιο χρονοβόρα από αυτήν που παρουσιάσαμε στην προηγούμενη ενότητα, **αλλά** είναι ευστάθης για κάθε τιμή, του r . Υπενθυμίζουμε, ότι στην προηγούμενη μέθοδο το κριτήριο ευστάθειας επέτρεπε τιμές $0 < r < 1/2$. Αρα με την έμμεση μέθοδο των Crank-Nicholson μπορούμε να χρησιμοποιήσουμε σημαντικά μεγαλύτερο χρονικό βήμα που επιταχύνει σημαντικά την επίλυση του προβλήματος.

ΠΡΟΒΛΗΜΑΤΑ

Δίνεται η παραβολική ΔΕΜΠ

$$u_t = u_{xx} \quad \text{για} \quad 0 < x < 1 \quad \text{και} \quad 0 < t < 2$$

με αρχικές συνθήκες

$$u(x, 0) = 4x(1 - x) \quad \text{για} \quad 0 \leq x \leq 1$$

και οριακές συνθήκες

$$u(0, t) = u(1, t) = 0 \quad \text{για} \quad 0 \leq t \leq 2.$$

1. Δημιουργήστε ένα κώδικα στη Mathematica για την αναλυτική λύση της παραπάνω ΔΕΜΠ.
2. Χρησιμοποιήστε το αριθμητικό σχήμα (8.40) για την επίλυση της παραπάνω ΔΕΜΠ.
3. Χρησιμοποιήστε την έμμεση μέθοδο των Crank-Nicholson για την επίλυση της παραπάνω ΔΕΜΠ.

Α΄

ΑΝΑΛΥΣΗ ΣΦΑΛΜΑΤΟΣ

Στην επίλυση μαθηματικών προβλημάτων με χρήση των μεθόδων της Αριθμητικής Ανάλυσης χρησιμοποιούμε προσεγγιστικές μεθόδους και οι λύσεις που βρίσκουμε δεν είναι προφανώς ακριβείς. Η ακρίβεια των αριθμητικών μεθόδων εξαρτάται από διάφορους παράγοντες, η κατανόηση των οποίων πολλές φορές μας βοηθά στη δημιουργία ακριβέστερων μεθόδων αλλά και στην εκτίμηση της ακρίβειας του αριθμητικού αποτελέσματος.

Ας υποθέσουμε ότι το p' είναι μια προσεγγιστική τιμή της ακριβούς τιμής p , τότε:

- Ως **σφάλμα** θα ορίσουμε την ποσότητα $E_p = p - p'$, δηλαδή τη διαφορά της ακριβούς από την προσεγγιστική τιμή, και
- Ως **σχετικό σφάλμα** θα ορίσουμε την ποσότητα $R_p = (p - p')/p$, δηλαδή το ποσοστό του σφάλματος σε σχέση με την ακριβή τιμή (προφανώς υποθέτουμε ότι $p \neq 0$).

ΠΑΡΑΔΕΙΓΜΑ

1. Η ακριβής τιμή του π είναι $\pi = 3.14159\dots$ και έστω μια προσεγγιστική τιμή του $\pi' = 3.14$. Το σφάλμα είναι:

$$E_\pi = \pi - \pi' = 3.14159 - 3.14000 = 0.00159$$

και το σχετικό σφάλμα:

$$R_\pi = \frac{\pi - \pi'}{\pi} = \frac{0.00159}{3.14159} = 0.000506.$$

2. Αν ακριβής τιμή είναι $y = 1.000.000$ και έστω μια προσεγγιστική τιμή $y' = 999.996$. Το σφάλμα είναι:

$$E_y = y - y' = 1.000.000 - 999.996 = 4$$

και το σχετικό σφάλμα:

$$R_y = \frac{y - y'}{y} = \frac{4}{1.000.000} = 0.000004.$$

3. Αν ακριβής τιμή είναι $x = 0.000012$ και έστω μια προσεγγιστική τιμή $x' = 0.000009$. Το σφάλμα είναι:

$$E_x = x - x' = 0.000012 - 0.000009 = 0.000003$$

και το σχετικό σφάλμα:

$$R_x = \frac{x - x'}{x} = \frac{0.000003}{0.000012} = 0.25.$$

Στην πρώτη περίπτωση δεν υπήρχε σημαντική διαφορά μεταξύ των E_π και R_π και αμφότερα μπορούν να χρησιμοποιηθούν για την εκτίμηση της ακρίβειας του π' . Στη δεύτερη περίπτωση το σφάλμα $E_y = 4$ είναι μεγάλο αλλ'η κοινή λογική μας οδηγεί στην εκτίμηση ότι η προσέγγιση δεν είναι κακή και αυτό εκφράζεται καλύτερα μέσω του σχετικού σφάλματος $R_y = 4 \times 10^{-6}$. Τέλος στην τρίτη περίπτωση το σφάλμα είναι πολύ μικρό 3×10^{-6} αλλ'η το σχετικό σφάλμα είναι μεγάλο και αντιστοιχεί στο 25% της ακριβούς τιμής. Από την παραπάνω συζήτηση γίνεται προφανές ότι το σχετικό σφάλμα είναι ένας καλύτερος δείκτης της ακρίβειας μιας προσέγγισης.

ΟΡΙΣΜΟΣ: Ένας αριθμός y' προσεγγίζει μια ακριβή τιμή y με d σημαντικά ψηφία αν, ο d είναι ο μεγαλύτερος αριθμός για τον οποίο ισχύει:

$$\frac{|y - y'|}{|y|} < \frac{10^{-d}}{2}. \quad (\text{Α'.1})$$

ΠΑΡΑΔΕΙΓΜΑ

Θα υπολογίσουμε τα δεκαδικά ψηφία στις προσεγγίσεις του προηγούμενου παραδείγματος

1. $|\pi - \pi'|/|\pi| = 0.000506 \approx 10^{-3}/2$ δηλαδή $d = 3$ (τρία σημαντικά ψηφία).
2. $|y - y'|/|y| = 0.000004 < 10^{-5}/2$ δηλαδή $d = 5$ (πέντε σημαντικά ψηφία).
3. $|x - x'|/|x| = 0.25 < 10^{-0}/2$ δηλαδή $d = 0$ (κανένα σημαντικό ψηφίο).

Α'.0.1 Σφάλμα αποκοπής

Τα σφάλματα αποκοπής εισάγονται όταν για λόγους ευκολίας αντικαθιστούμε μια πιο σύνθετη μαθηματική έκφραση με μία απλούστερη. Για παράδειγμα το ανάπτυγμα Taylor της συνάρτησης του ημιτόνου:

$$\sin(x) \approx x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots$$

ΠΑΡΑΔΕΙΓΜΑ

Εστω ότι ζητούμε να υπολογίσουμε το ορισμένο ολοκλήρωμα $I = \int_0^1 \arctan(x) dx$. Για τον υπολογισμό της τιμής του θα μπορούσαμε να χρησιμοποιήσουμε το ανάπτυγμα Taylor της συνάρτησης $\arctan(x)$, δηλαδή

$$\arctan(x) \approx x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \frac{x^9}{9} - \dots$$

οπότε αν διατηρήσουμε όρους ως και $O(x^7)$ λαμβάνουμε

$$\begin{aligned} I' &= \int_0^1 \left(x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} \right) dx = \left(\frac{x^2}{2} - \frac{x^4}{12} + \frac{x^6}{30} - \frac{x^8}{56} \right) \Big|_0^1 \\ &= \frac{121}{280} = 0.4321428571. \end{aligned}$$

Η ακριβής τιμή είναι $I = 0.4388245732$, άρα το σχετικό σφάλμα είναι $R_I = (I - I')/I = 0.0152264$. Αν στην παραπάνω προσέγγιση διατηρούσαμε όρους ως και $O(x^9)$ τότε η προσεγγιστική τιμή του ολοκληρώματος θα ήταν $I' = 0.4432539683$ και το σχετικό σφάλμα $R_I = -0.0100938$. Δηλαδή όπως αναμενόταν η ακρίβεια βελτιώθηκε.

Α'.0.2 Σφάλμα στρογγύλευσης

Η ακρίβεια με την οποία αποθηκεύεται ένας πραγματικός αριθμός στη μνήμη του ΗΥ είναι συγκεκριμένη και επομένως μόνο ένας συγκεκριμένος αριθμός σημαντικών ψηφίων του αριθμού διατηρείται. Τούτο έχει ως αποτέλεσμα οι εκτελούμενες αριθμητικές πράξεις από τιον ΗΥ να δίδουν αποτέλεσμα με σκρίβεια που είναι ανάλογη αυτής που αποθηκεύονται οι αριθμοί.

Α'.0.3 Απώλεια σημαντικών ψηφίων

Κατά τη διάρκεια εκτέλεσης αριθμητικών πράξεων ελοχεύει ο κίνδυνος να **απολεσθούν σημαντικά ψηφία**. Έστω δύο σχετικά ίσοι αριθμοί π.χ. $x = 1.11112$ και $y = 1.11113$ που δίδονται με ακρίβεια 5 δεκαδικών ψηφίων. Θα αναμέναμε ότι οι αριθμητικές πράξεις μεταξύ τους είναι της αυτής ακρίβειας, τούτο όμως δεν είναι αληθές για την αφαίρεση τους διότι $x - y = 0.00001$ αλλά στον ΗΥ αποθηκεύεται ως $0.1bbb \times 10^{-5}$. Οπότε για τον αριθμό που προέκυψε γνωρίζουμε μόνο ένα ψηφίο ενώ τα bbb ψηφία είναι άγνωστα (και πιθανότατα θα χρησιμοποιηθούν τυχαίες τιμές από τον ΗΥ για αυτά). Δηλαδή, στη συνέχεια οι πράξεις που θα χρησιμοποιούν την παραπάνω διαφορά θα έχουν ακρίβεια ενός μόνο σημαντικού ψηφίου.

ΠΑΡΑΔΕΙΓΜΑ

Ας εξετάσουμε το αριθμητικό αποτέλεσμα που θα πάρουμε συγκρίνοντας δύο 'γραφές' της ίδιας συνάρτησης. Έστω λοιπόν οι συναρτήσεις

$$f_1(x) = x^2 \left(\sqrt{x+1} - \sqrt{x} \right) \quad \text{και} \quad f_2(x) = \frac{x^2}{\sqrt{x+1} + \sqrt{x}}$$

για $x = 300$ αν εκτελέσουμε την πράξη σε ένα ΗΥ που 'κρατά' 8 δεκαδικά ψηφία θα πάρουμε $f_1(300) = 2596.0$ και $f_2(300) = 2595.9147$ οπότε η διαφορά τους είναι $f_1(300) - f_2(300) = 0.0853$ δηλαδή αντί η διαφορά τους να είναι της τάξης του 10^{-8} είναι περίπου 0.1 δηλαδή έχουμε απώλεια της τάξης των 7 σημαντικών ψηφίων. Αντίστοιχα, αν οι πράξεις εκτελούνταν με ακρίβεια 16 σημαντικών ψηφίων τότε αντί ακρίβειας της τάξης του 10^{-16} το αποτέλεσμα είναι $f_1(300) - f_2(300) \approx 4.7 \times 10^{-11}$, δηλαδή απώλεια 5 σημαντικών ψηφίων.

Α'.0.4 Προσέγγιση τάξης $O(h^n)$

Πολλές φορές στην αριθμητική επίλυση μαθηματικών προβλημάτων αντικαθιστούμε μια σύνθετη συνάρτηση $f(h)$ με μια απλούστερη (πχ ένα πολυώνυμο) έστω την $p(h)$ με ένα μέγιστο σφάλμα $M|h^n|$. Οπότε μπορούμε να χρησιμοποιήσουμε τον παρακάτω ορισμό:

ΟΡΙΣΜΟΣ: Αν μια συνάρτηση $f(h)$ προσεγγίζεται από τη συνάρτηση $p(h)$ και υπάρχει μια πραγματική σταθερά $M > 0$ και ένας ακέραιος n τέτοιοι ώστε

$$\frac{|f(h) - p(h)|}{|h|^n} \leq M \quad \text{όπου} \quad h \ll 1 \quad (\text{Α'.2})$$

τότε λέμε ότι η $p(h)$ προσεγγίζει την $f(h)$ και η προσέγγιση αυτή είναι τάξης $O(h^n)$ και γράφουμε:

$$f(h) = p(h) + O(h^n) \quad (\text{Α'.3})$$

ΘΕΩΡΗΜΑ: Αν οι συναρτήσεις $f(x)$ και $g(x)$ μπορούν να γραφούν ως : $f(h) = p(h) + O(h^n)$ και $g(h) = q(h) + O(h^m)$ και υπάρχει $p = \min(m, n)$. Τότε θα ισχύει:

$$f(h) + g(h) = p(h) + g(h) + O(h^p) \quad (\text{Α'.4})$$

$$f(h) \cdot g(h) = p(h) \cdot q(h) + O(h^p) \quad (\text{Α'.5})$$

$$\frac{f(h)}{g(h)} = \frac{p(h)}{q(h)} + O(h^p) \quad \text{εφόσον } g(h) \neq 0, q(h) \neq 0 \quad (\text{Α'.6})$$

ΠΑΡΑΔΕΙΓΜΑ**Α'.0.5 Διάδοση σφάλματος**

Ας προσπαθήσουμε να εξετάσουμε τον τρόπο με τον οποίο ένα σφάλμα διαδίδεται κατά τη διάρκεια επαναλαμβανόμενων επαναλήψεων. Ας θεωρήσουμε την

πρόσθεση δύο αριθμών x και y . Προφανώς, στον ΗΥ αντί των ακριβών τιμών θα γίνει πρόσθεση των προσεγγιστικών τους τιμών x' και y' , που εμπεριέχουν σφάλμα ε_x και ε_y αντίστοιχα. Δηλαδή, θα είναι $x = x' + \varepsilon_x$ και $y = y' + \varepsilon_y$ οπότε το άθροισμα τους θα είναι

$$x + y = (x' + \varepsilon_x) + (y' + \varepsilon_y) = (x' + y') + (\varepsilon_x + \varepsilon_y). \quad (\text{Α'.7})$$

Άρα στην πράξη της πρόσθεσης το συνολικό σφάλμα θα είναι το άθροισμα των επι μέρους σφαλμάτων.

Στον πολλαπλασιασμό η εκτίμηση του σφάλματος είναι δυσκολότερη. Το γινόμενο είναι

$$xy = (x' + \varepsilon_x)(y' + \varepsilon_y) = x'y' + x'\varepsilon_y + y'\varepsilon_x + \varepsilon_x\varepsilon_y \quad (\text{Α'.8})$$

Αν ισχύει ότι $|x'| > 1$ και $|y'| > 1$ τότε οι όροι $x'\varepsilon_y$ και $y'\varepsilon_x$ θα επαυξήσουν τα αρχικά σφάλματα, αλλά δεν είναι δυνατόν μάλιστα τον τρόπο να γίνει ορθή εκτίμηση του σφάλματος. Αντ' αυτού είναι πιο αξιόπιστο κριτήριο η μελέτη του σχετικού σφάλματος. Η παραπάνω σχέση μπορεί εύκολα να γραφεί ως

$$\frac{xy - x'y'}{xy} = \frac{x'\varepsilon_y + y'\varepsilon_x + \varepsilon_x\varepsilon_y}{xy} = \frac{x'\varepsilon_y}{xy} + \frac{y'\varepsilon_x}{xy} + \frac{\varepsilon_x\varepsilon_y}{xy}. \quad (\text{Α'.9})$$

Η σχέση αυτή μπορεί να απλοποιηθεί σημαντικά αν υποθέσουμε ότι $x'/x \approx 1$, $y'/y \approx 1$ και $(\varepsilon_y/x)(\varepsilon_x/y) = R_x R_y \approx 0$. Οπότε καταλήγουμε στην απλοποιημένη ορφή της παραπάνω σχέσης

$$\frac{xy - x'y'}{xy} \approx \frac{\varepsilon_y}{y} + \frac{\varepsilon_x}{x} = R_x + R_y. \quad (\text{Α'.10})$$

Συνήθως έχουμε πολλές επαναλήψεις των απλών πράξεων, οπότε το ζητούμενο είναι τα οποιαδήποτε σφάλματα στα αρχικά δεδομένα μετά από μια ακολουθία αριθμητικών πράξεων να έχουν το μικρότερο δυνατό σφάλμα στο τελικό αποτέλεσμα. Ένας αλγόριθμος που διατηρεί το σφάλμα μικρό κατά τη διάρκεια των πράξεων και το συνολικό τελικό σφάλμα δεν αυξάνεται θα λέγεται **ευσταθής**, αλλιώς θα λέγεται **ασταθής**. Επομένως, οι μέθοδοι που χρησιμοποιούμε θα πρέπει πέραν την ταχύτητας σύγκλισης στο ορθό αποτέλεσμα να είναι και ευσταθείς.

ΟΡΙΣΜΟΣ: Ας παραστήσουμε με $E(n)$ την αύξηση του σφάλματος μετά από n βήματα. Αν ισχύει ότι $|E(n)| \approx n\varepsilon$ θα λέμε ότι η αύξηση του σφάλματος είναι **γραμμική**. Αν $|E(n)| \approx K^n \varepsilon$ τότε η αύξηση του σφάλματος είναι **εκθετική**. Προφανώς αν $|K| > 1$ το σφάλμα αυξάνεται χωρίς όρια καθώς $n \rightarrow \infty$, ενώ αν $0 < |K| < 1$ τότε το σφάλμα ελαχιστοποιείται καθώς $n \rightarrow \infty$.

Α'.0.6 Αβεβαιότητα στα δεδομένα

Πολλές φορές τα δεδομένα που χρησιμοποιούμε σε διάφορα προβλήματα προέρχονται από μετρήσεις με περιορισμένη ακρίβεια ή υπάρχει αβεβαιότητα ως προς

τις τιμές που θα χρησιμοποιηθούν. Σε αυτή την περίπτωση τα οποιαδήποτε αριθμητικά αποτελέσματα έχουν ακρίβεια μικρότερη ή ίση των αρχικών δεδομένων. Ως εκτούτου ανεξάρτητα από την ακρίβεια της μεθόδου που χρησιμοποιούμε θα πρέπει να γίνεται χρήση των αποτελεσμάτων με τόσα σημαντικά ψηφία όσα είναι των αρχικών δεδομένων.

Ευρετήριο

Aitken μέθοδος, ιδιοτιμές, 46
Aitken, βελτίωση, 14
Crank-Nicholson μέθοδος, 138
Euler-Maclaurin, τύπος, 85
Filon, μέθοδος, 86
Gauss, μέθοδος ολοκλήρωσης, 88
Hermite πολυώνυμο, 57
Lagrange πολυώνυμο, 50
Newton-Cotes, τύποι, 75
Newton-Gregory τύπος, 56
Romberg, βελτίωση, 80
Simpson (3/8), κανόνας, 80
Simpson 3/8 μέθοδος, σφάλμα, 80
Simpson μέθοδος, σφάλμα, 80
Simpson, κανόνας, 80
Splines παρεμβολή, 59
Splines, ολοκλήρωση, 82
Splines, παραγωγή, 82
Taylor Πολυώνυμο, 58

Αποκοπή, σφάλμα, 142

Βελτίωση του Romberg, 80

Διαφορικές εξισώσεις, 95
Διαφορικές εξισώσεις : Συστήματα, 107
Διαφορικές Εξισώσεις με Μερικές Παραγώγους : Ελλειπτικές, 126
Διαφορικές Εξισώσεις με Μερικές Παραγώγους : Παραβολικές, 136
Διαφορικές Εξισώσεις με Μερικές Παραγώγους : Υπερβολικές, 131
Διαφορικές εξισώσεις, ευστάθεια, 99
Διαφορικές εξισώσεις, σύγκλιση, 99

Διαφορικές εξισώσεις, σφάλματα διάδοσης, 99
Διαφορικές εξισώσεις: Μέθοδοι πρόβλεψης-διόρθωσης, 105
Διαφορικές εξισώσεις: Μέθοδος Adams-Moulton, 106
Διαφορικές εξισώσεις: Μέθοδος Adams, 103
Διαφορικές εξισώσεις: Μέθοδος Euler-Heun, 97
Διαφορικές εξισώσεις: Μέθοδος Euler, 97
Διαφορικές εξισώσεις: Μέθοδος Hamming, 106
Διαφορικές εξισώσεις: Μέθοδος Milne, 105, 106
Διαφορικές εξισώσεις: Μέθοδος Nystrom, 109
Διαφορικές εξισώσεις: Μέθοδος Runge-Kutta-Fehlberg, 102
Διαφορικές εξισώσεις: Μέθοδος Runge-Kutta, 100, 107
Διαφορικές εξισώσεις: Μέθοδος Runge-Kutta: Συστήματα, 108
Διαφορικές εξισώσεις: Μέθοδος Σειρών Taylor, 96
Διαφορικές εξισώσεις: Σφάλματα, 98
Δυναμικών μέθοδος, ιδιοτιμές, 46

Εφαπτόμενο πολυώνυμο, 57
Ελλειπτικές Διαφορικές Εξισώσεις με Μερικές Παραγώγους, 126
Ευστάθεια, Διαφορικές εξισώσεις, 99

Γραμμικά Συστήματα, 31

- Γραμμικά Συστήματα, Ιδιοτιμές, μέθοδος δυνάμεων, 44
 Γραμμικά Συστήματα: Αντίστροφος Πίνακας, 42
 Γραμμικά Συστήματα: Ιδιοδιανύσματα, 43
 Γραμμικά Συστήματα: Ιδιοτιμές, 43
 Γραμμικά Συστήματα: Μέθοδος Gauss-Jordan, 35
 Γραμμικά Συστήματα: Μέθοδος Gauss-Seidel, 40
 Γραμμικά Συστήματα: Μέθοδος Gauss, 32
 Γραμμικά Συστήματα: Μέθοδος Jacobi, 39
 Γραμμικά Συστήματα: Μέθοδος L-U, 36
 Γραμμικά Συστήματα: Οδήγηση pivoting, 34
 Γραμμικά Συστήματα: Ορίζουσα, 42
- Ιδιοτιμές, Aitken μέθοδος, 46
 Ιδιοτιμές, δυνάμεων μέθοδος, 46
 Ιδιοτιμές, Μεταθεσης μέθοδος, 47
- Κανόνας του Simpson, 80
 Κανόνας του Simpson (3/8), 80
 Κανόνας τραπεζίου, 78
 Κεντρικές διαφορές, 72
- Μέθοδος $x = g(x)$, 11
 Μέθοδος Bolzano, 2
 Μέθοδος Muller, 10
 Μέθοδος Newton-Raphson, 15
 Μέθοδος διχοτόμησης, 2
 Μέθοδος γραμμικής παρεμβολής, 4
 Μέθοδος ολοκλήρωσης Gauss, 88
 Μέθοδος του Filon, 86
 Μεταθεσης μέθοδος, ιδιοτιμές, 47
- Ολοκλήρωση με Splines, 82
- Πίνακες, Ιδιοτιμές, μέθοδος δυνάμεων, 44
 Πίνακες: Αντίστροφος, 42
 Πίνακες: Ιδιοδιανύσματα, 43
 Πίνακες: Ιδιοτιμές, 43
 Πίνακες: Ορίζουσα, 42
 Παραβολικές Διαφορικές Εξισώσεις με Μερικές Παραγώγους, 136
 Παραγωγή αριθμητική, 69
 Παραγωγή με Splines, 82
 Παρεμβολή, 49
 Παρεμβολή με Splines, 59
 Πολυώνυμο Lagrange, 50
 Πολυώνυμο Taylor, 58
 Πρόβλεψη, 49
- Ρίζες εξισώσεων, 1
- Σύγκλιση, Διαφορικές εξισώσεις, 99
 Σφάλμα, 141
 Σφάλμα αποκοπής, 142
 Σφάλμα μεθόδου Muller, 10
 Σφάλμα μεθόδου Simpson, 80
 Σφάλμα μεθόδου Simpson 3/8, 80
 Σφάλμα μεθόδου διχοτόμησης, 4
 Σφάλμα μεθόδου γραμμικής παρεμβολής, 8
 Σφάλμα μεθόδου τραπεζίου, 79
 Σφάλμα της $x = g(x)$, 12
 Σφάλμα της Newton-Raphson, 17
 Σφάλμα, σχετικό, 141
 Σφάλματα διάδοσης, Διαφορικές εξισώσεις, 99
 Συμπτωτικό πολυώνυμο, ισαπέχοντα, 54
- Τύπος Euler-Maclaurin, 85
 Τύπος Newton-Gregory, 56
 Τραπεζίου μέθοδος, σφάλμα, 79
 Τραπεζίου, κανόνας, 78
- Υπερβολικές Διαφορικές Εξισώσεις με Μερικές Παραγώγους, 131