ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ
ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΑΓΡΟΝΟΜΩΝ & ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ
ΤΟΜΕΑΣ ΣΥΓΚΟΙΝΩΝΙΑΚΩΝ & ΥΔΡΑΥΛΙΚΩΝ ΕΡΓΩΝ

ΑΡΙΘΜΗΤΙΚΗ ΑΝΑΛΥΣΗ ΜΕ ΧΡΗΣΗ Η/Υ

Ε. Σιδηρόπουλος
αναπληρωτής καθηγητής

Χ. Φωτιάδης
επίκουρος καθηγητής

Θεσσαλονίκη 2010

ΠΕΡΙΕΧΟΜΕΝΑ

1 ΕΙΣΑΓΩΓΗ σελ. 1

1.1 ΑΠΟΛΥΤΟ ΚΑΙ ΣΧΕΤΙΚΟ ΣΦΑΛΜΑ. ΠΑΡΑΣΤΑΣΗ ΑΡΙΘΜΟΥ ΜΕ ΠΕΠΕΡΑΣΜΕΝΟ ΑΡΙΘΜΟ ΨΗΦΙΩΝ 2

1.2 ΑΠΩΛΕΙΑ ΣΗΜΑΝΤΙΚΩΝ ΨΗΦΙΩΝ. ΑΣΤΑΘΕΙΑ 8

2 ΠΡΟΣΕΓΓΙΣΤΙΚΗ ΕΠΙΛΥΣΗ ΜΗ ΓΡΑΜΜΙΚΩΝ ΕΞΙΣΩΣΕΩΝ 13

2.1 ΜΕΘΟΔΟΣ ΣΥΝΑΡΤΗΣΙΑΚΩΝ ΕΠΑΝΑΛΗΨΕΩΝ 13

2.1.1 ΟΡΙΣΜΟΣ ΚΑΙ ΠΑΡΑΔΕΙΓΜΑΤΑ 13

2.1.2 ΘΕΩΡΙΑ ΣΥΓΚΛΙΣΗΣ 18

2.1.3 ΤΡΟΠΟΠΟΙΗΣΗ ΤΗΣ ΜΕΘΟΔΟΥ ΤΩΝ ΣΥΝΑΡΤΗΣΙΑΚΩΝ ΕΠΑΝΑΛΗΨΕΩΝ 22

2.1.4 ΓΕΝΙΚΕΥΣΗ ΤΩΝ ΘΕΩΡΗΜΑΤΩΝ ΣΥΓΚΛΙΣΗΣ 29

2.2 ΜΕΘΟΔΟΣ NEWTON - RAPHSON 32

2.2.1 ΤΑΧΥΤΗΤΑ ΣΥΓΚΛΙΣΗΣ ΤΗΣ ΜΕΘΟΔΟΥ NEWTON - RAPHSON 36

2.2.2 ΠΑΡΑΛΛΑΓΕΣ ΤΗΣ ΜΕΘΟΔΟΥ NEWTON - RAPHSON 39

2.3 ΕΠΙΤΑΧΥΝΣΗ ΣΥΝΑΡΤΗΣΙΑΚΩΝ ΕΠΑΝΑΛΗΨΕΩΝ ΚΑΤΑ AITKEN 47

2.3.1 ΑΝΑΛΥΣΗ ΤΟΥ ΣΦΑΛΜΑΤΟΣ ΣΤΗΝ ΕΠΙΤΑΧΥΝΣΗ ΚΑΤΑ AITKEN 48

2.4 ΜΕΘΟΔΟΣ ΤΗΣ ΔΙΧΟΤΟΜΗΣΗΣ 51

3 ΥΠΟΛΟΓΙΣΜΟΣ ΡΙΖΩΝ ΠΟΛΥΩΝΥΜΩΝ 52

3.1 ΥΠΟΛΟΓΙΣΜΟΣ ΤΗΣ ΤΙΜΗΣ ΕΝΟΣ ΠΟΛΥΩΝΥΜΟΥ ΚΑΙ ΤΗΣ ΠΑΡΑΓΩΓΟΥ ΤΟΥ 52

3.2 Η ΜΕΘΟΔΟΣ NEWTON - RAPHSON ΓΙΑ ΤΗΝ ΕΥΡΕΣΗ ΠΡΑΓΜΑΤΙΚΩΝ ΡΙΖΩΝ ΠΟΛΥΩΝΥΜΩΝ 58

3.3 Η ΜΕΘΟΔΟΣ BERNOULLI 58

3.4 ΜΙΓΑΔΙΚΕΣ ΡΙΖΕΣ ΠΟΛΥΩΝΥΜΩΝ 62

3.5 ΕΝΤΟΠΙΣΜΟΣ ΡΙΖΩΝ ΠΟΛΥΩΝΥΜΩΝ 68

4 ΕΠΙΛΥΣΗ ΓΡΑΜΜΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ 73

4.1 ΑΠΑΛΟΙΦΗ GAUSS 73

4.2 ΣΥΣΤΗΜΑ ΜΕ ΤΡΙΔΙΑΓΩΝΙΚΟ ΠΙΝΑΚΑ 77

4.3 ΕΠΑΝΑΛΗΠΤΙΚΕΣ ΜΕΘΟΔΟΙ 80

4.4 ΝΟΡΜΕΣ ΔΙΑΝΥΣΜΑΤΩΝ ΚΑΙ ΠΙΝΑΚΩΝ 82

4.5 ΙΔΙΟΤΙΜΕΣ ΚΑΙ ΙΔΙΟΔΙΑΝΥΣΜΑΤΑ σελ. 86
4.5.1 ΠΡΟΣΕΓΓΙΣΤΙΚΟΣ ΥΠΟΛΟΓΙΣΜΟΣ ΤΩΝ ΙΔΙΟΤΙΜΩΝ ΠΙΝΑΚΩΝ 89
4.5.2 ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΙΔΙΟΤΙΜΩΝ 90
4.6 ΠΑΘΟΛΟΓΙΚΕΣ ΚΑΤΑΣΤΑΣΕΙΣ ΣΤΗΝ ΕΠΙΛΥΣΗ ΓΡΑΜΜΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ 97
4.6.1 ΠΑΘΟΛΟΓΙΚΕΣ ΚΑΤΑΣΤΑΣΕΙΣ ΣΕ ΠΡΟΒΛΗΜΑΤΑ ΧΑΡΤΟΓΡΑΦΙΑΣ-ΦΩΤΟΓΡΑΜΜΕΤΡΙΑΣ 101

5 ΠΑΡΕΜΒΟΛΗ 103
5.1 ΠΟΛΥΩΝΥΜΟ ΠΑΡΕΜΒΟΛΗΣ 103
5.2 ΠΟΛΥΩΝΥΜΟ LAGRANGE 106
5.3 ΣΦΑΛΜΑ ΤΟΥ ΠΟΛΥΩΝΥΜΟΥ ΠΑΡΕΜΒΟΛΗΣ 108
5.4 ΕΠΑΝΑΛΑΜΒΑΝΟΜΕΝΗ ΓΡΑΜΜΙΚΗ ΠΑΡΕΜΒΟΛΗ 110
5.5 ΠΟΛΥΩΝΥΜΟ ΠΑΡΕΜΒΟΛΗΣ ΚΑΤΑ NEWTON 113
5.6 SPLINES 120

6 ΑΡΙΘΜΗΤΙΚΗ ΟΛΟΚΛΗΡΩΣΗ 131
6.1 ΣΤΟΙΧΕΙΩΔΕΙΣ ΜΕΘΟΔΟΙ 131
6.2 ΚΑΝΟΝΑΣ ΤΟΥ SIMPSON 136

7 ΑΡΙΘΜΗΤΙΚΗ ΕΠΙΛΥΣΗ ΔΙΑΦΟΡΙΚΩΝ ΕΞΙΣΩΣΕΩΝ 140
7.1 ΜΕΘΟΔΟΣ ΤΟΥ EULER 140
7.2 ΜΕΘΟΔΟΙ RUNGE - KUTTA 144
7.3 ΜΕΘΟΔΟΙ ΠΕΠΕΡΑΣΜΕΝΩΝ ΔΙΑΦΟΡΩΝ 147
7.3.1 ΠΡΟΣΕΓΓΙΣΗ ΠΑΡΑΓΩΓΩΝ ΜΕ ΠΕΠΕΡΑΣΜΕΝΕΣ ΔΙΑΦΟΡΕΣ 147
7.3.2 ΜΕΘΟΔΟΣ ΤΩΝ CRANK - NICOLSON 151

8. ΠΡΟΓΡΑΜΜΑΤΑ 154

ΑΣΚΗΣΕΙΣ ΓΙΑ ΛΥΣΗ 182

ΓΕΝΙΚΗ ΒΙΒΛΙΟΓΡΑΦΙΑ 187

1. ΕΙΣΑΓΩΓΗ

Τα ορισμένα ολοκληρώματα, οι διαφορικές εξισώσεις, οι μη γραμμικές αλγεβρικές εξισώσεις, τα γραμμικά συστήματα, οι ακολουθίες, οι σειρές είναι θέματα των καθαρών Μαθηματικών που οδηγούν σε υπολογιστικά προβλήματα. Η Αριθμητική Ανάλυση σκοπό έχει την εξεύρεση, διαμόρφωση και ανάλυση αριθμητικών λύσεων σε μαθηματικά προβλήματα.

Αυτό σημαίνει ότι η Αριθμητική Ανάλυση ανάγει τελικά τα προβλήματα που πραγματεύεται σε υπολογιστικές διαδικασίες. Αυτές αποτελούνται από σαφώς ορισμένα διαδοχικά βήματα και κάθε βήμα συνίσταται στην εκτέλεση μιας ή περισσοτέρων από τις τέσσερεις αριθμητικές πράξεις. Το αριθμητικό αποτέλεσμα μιας τέτοιας διαδικασίας είναι, στη γενική περίπτωση, προσεγγιστικό. Αυτό οφείλεται στους παρακάτω λόγους :

(α) Η παράσταση αριθμών, είτε γίνεται από τον άνθρωπο, είτε από μια μηχανή, περιλαμβάνει πάντα πεπερασμένο αριθμό ψηφίων. Η αναγκαστική παράλειψη ψηφίων εισάγει ένα σφάλμα, που είναι γνωστό ως **σφάλμα αποκοπής ή στρογγύλευσης.**

(β) Το σφάλμα αυτό μεταδίδεται στις αριθμητικές πράξεις που ακολουθούν. Επιπλέον, το αποτέλεσμα κάθε μιας από αυτές υπόκειται σε περιορισμό των ψηφίων του.

(γ) Πολλά προβλήματα παρουσιάζουν μια θεωρητικά ατέρμονη αλληλουχία επί μέρους βημάτων. Τέτοιο πρόβλημα είναι η εύρεση του αθροίσματος μιας συγκλίνουσας σειράς. Οι περιορισμένες δυνατότητες του ανθρώπου ή της μηχανής επιβάλλουν τη διακοπή της διαδικασίας έπειτα από έναν αριθμό βημάτων και την απόκτηση κάποιου αποτελέσματος προσεγγιστικού. Το σφάλμα που εισάγεται με αυτό τον τρόπο μπορεί να ονομαστεί **σφάλμα διακοπής**.

(δ) Οι πραγματικοί αριθμοί αποτελούν ένα συνεχές. Το συνεχές όμως αυτό δεν μπορεί να εισαχθεί με την ιδεατή μορφή του μέσα σε αριθμητικές διαδικασίες. Για παράδειγμα, το διάστημα [0, 1], ενώ περιλαμβάνει έναν άπειρο αριθμό σημείων, αναγκαστικά αντιπροσωπεύεται σε μία αριθμητική διαδικασία με ένα πεπερασμένο αριθμό εσωτερικών του σημείων. Μια τέτοια διαδικασία είναι η αριθμητική ολοκλήρωση, που θα εκτεθεί σε κατοπινό κεφάλαιο. Συνεπώς, για τις αριθμητικές ανάγκες, το συνεχές διακριτοποιείται και το σφάλμα που εισάγεται με αυτό τον τρόπο στους υπολογισμούς ονομάζεται **σφάλμα διακριτοποίησης**.

1.1 ΑΠΟΛΥΤΟ ΚΑΙ ΣΧΕΤΙΚΟ ΣΦΑΛΜΑ. ΠΑΡΑΣΤΑΣΗ ΑΡΙΘΜΟΥ ΜΕ ΠΕΠΕΡΑΣΜΕΝΟ ΑΡΙΘΜΟ ΨΗΦΙΩΝ

Έστω x πραγματικός αριθμός και X μία προσέγγισή του. Για παράδειγμα, $x = \sqrt{2}$ και X = 1.41. Ονομάζεται **σφάλμα** Δ_x η διαφορά ανάμεσα στον πραγματικό αριθμό x και στην προσέγγιση X :

$$\Delta_x = x - X.$$

Μεγαλύτερη πρακτική αξία έχει το **σχετικό ή ποσοστιαίο σφάλμα** :

$$\delta_x = \frac{x - X}{x}$$

Η γενική παράσταση πραγματικού αριθμού x στο δεκαδικό σύστημα έχει ως εξής :

$$x = a_m.10^m + a_{m-1}.10^{m-1} + \ldots + a_0.10^0 + b_1.10^{-1} + b_2.10^{-2} + \ldots$$

όπου, $a_m = 1, 2, \ldots, 9$

$a_i = 0, 1, 2, \ldots, 9$ για $i = 0, 1, 2, \ldots, m-1$

$b_i = 0, 1, 2, \ldots, 9$ για $i = 1, 2, 3, \ldots$

Παραδείγματα :

$\pi = 3.14159...$ $= 3*10^0 + 1*10^{-1} + 4*10^{-2} + ...$

$\frac{100}{3} = 33.333...$ $= 3*10^1 + 3*10^0 + 3*10^{-1} + 3*10^{-2} + ...$

$\frac{401}{4} = 100.25$ $= 1*10^2 + 0*10^1 + 0*10^0 + 2*10^{-1} + 5*10^{-2}$.

Στους υπολογισμούς οι αριθμοί εμφανίζονται με πεπερασμένο αριθμό δεκαδικών ψηφίων. Τα ψηφία αυτά δεν είναι όλα ακριβή ή **σωστά**, π.χ. αν ο αριθμός x = 32.26 προσεγγίζεται από τον αριθμό X = 32.34, τότε τα τρία πρώτα ψηφία του X είναι σωστά. Για να δώσουμε έναν ακριβή ορισμό, λέμε ότι σε μια πεπερασμένη δεκαδική παράσταση **X** τα πρώτα **n** ψηφία είναι **σωστά**, αν

$|x - X| \leq 0.5*10^{m-n+1}$, όπου τα πρώτα n ψηφία :

$a_m, a_{m-1}, ..., a_0, b_1, ..., b_{m-n-1}$.

Πράγματι, για x = 32.36 και X = 32.34,

$m = 1$ και $|x - X| = 0.02 < 0.05 = 0.05*10^{-1}$.

Άρα $m - n + 1 = -1$ και $n = 3$.

Αν τώρα, x = 32.96 και X = 33.00,

$|x - X| = 0.04 < 0.05 = 0.5*10^{-1}$ και $m - n + 1 = -1$, όπου $m = 1$.

Άρα, και πάλι $n = 3$.

Για x = 10 και X = 9.995,

$|x - X| = 0.005 = 0.5*10^{-2}$ και $m - n + 1 = -2$ με $m = 0$.

Άρα $n = 3$.

Σε μία δεκαδική παράσταση, **σημαντικό** ψηφίο είναι κάθε ψηφίο διάφορο του 0. Αν κάποιο ψηφίο είναι 0, τότε θεωρείται σημαντικό, όταν βρίσκεται ανάμεσα σε ψηφία διάφορα από το 0. Αν βρίσκεται στο τέλος μιας πεπερασμένης δεκαδικής παράστασης, τότε θεωρείται σημαντικό, εάν ως δεκαδικό ψηφίο είναι σωστό με την έννοια του παραπάνω ορισμού. Για παράδειγμα, στη δεκαδική παράσταση 0.0020080 τα δύο πρώτα μηδενικά μετά την υποδιαστολή δεν είναι σημαντικά, το 2 και το 8 είναι σημαντικά όπως επίσης σημαντικά είναι τα δύο μηδενικά ανάμεσα στο 2 και στο 8. Αν το τελευταίο μηδενικό δεν είναι σημαντικό, τότε δεν πρέπει να αναγράφεται : 0.002008.

Αν οι αριθμητικές πράξεις εκτελούνται ομοιόμορφα με αριθμούς **κινητής υποδιαστολής**, τότε ο παραπάνω αριθμός γράφεται $0.20080*10^{-2}$, έτσι ώστε το πρώτο σημαντικό ψηφίο να βρίσκεται αμέσως μετά την υποδιαστολή. Ας σημειωθεί ακόμη ότι η παραμονή στην παραπάνω παράσταση του μηδενικού μετά το 8 υποδηλώνει πως αυτό θεωρείται σημαντικό. Στις περισσότερες υπολογιστικές μηχανές ακόμα και αριθμοί μεγαλύτεροι από τη μονάδα μετατρέπονται στην παραπάνω μορφή, π.χ. $625.122 = 0.625122*10^{3}$. Με τον τρόπο αυτό κάθε αριθμός μπορεί να θεωρηθεί ότι αναλύεται σε δύο παράγοντες, ένα κλάσμα και μία δύναμη του 10. Η αναγωγή αυτή των αριθμών ονομάζεται **κανονικοποίηση.**

Όπως ήδη αναφέρθηκε, η παράσταση ενός πραγματικού αριθμού για υπολογιστικούς σκοπούς δεν μπορεί παρά να περιλαμβάνει πεπερασμένο αριθμό ψηφίων. Από αυτά θεωρούμε μόνο τα σημαντικά και έστω n ο αριθμός των πρώτων σημαντικών ψηφίων. Τα ψηφία από το n+1 και πέρα απλώς παραλείπονται. Η πράξη αυτή λέγεται **αποκοπή**. Για να ληφθεί κάπως υπόψη η επίδραση των παραλειπόμενων ψηφίων είναι δυνατό να εφαρμοστεί μια άλλη πράξη που λέγεται **στρογγύλευση**. Για να

στρογγυλευτεί ένας αριθμός στα πρώτα n σημαντικά του ψηφία, διακρίνονται οι εξής περιπτώσεις :

(α) Αν το n+1 ψηφίο είναι μικρότερο από το 5, τότε απλώς παραλείπονται όλα τα ψηφία από n+1 και πέρα.

(β) Αν το n+1 ψηφίο είναι μεγαλύτερο από 5, τότε προστίθεται μία μονάδα στο n ψηφίο και παραλείπονται όλα τα ψηφία από το n+1 και πέρα.

(γ) Αν το n+1 ψηφίο είναι το 5 και υπάρχουν μη μηδενικά ψηφία μετά από αυτό, τότε εφαρμόζεται η παραπάνω περίπτωση (β).

(δ) Αν το n+1 ψηφίο είναι το 5 και ακολουθείται από μηδενικά, τότε εξετάζεται το n-οστό ψηφίο. Αν αυτό είναι άρτιο εφαρμόζεται η περίπτωση (β).

Για παράδειγμα, ο π = 3.1415926535. . . μπορεί να αποκοπεί ή να στρογγυλευτεί, όπως παρακάτω :

αποκοπή :	3.14	3.141	3.1415	3.14159
στρογγύλευση :	3.14	3.142	3.1416	3.14159

κ.λπ.

Είναι εύκολο να αποδειχτεί ότι, στην περίπτωση της αποκοπής, $|\Delta| \leq 10^{-t}$, όπου t η τελευταία παραμένουσα δεκαδική θέση. Για τη στρογγύλευση, $|\Delta| \leq 0.5*10^{-t}$.

Η πράξη της στρογγύλευσης μπορεί να οριστεί συνοπτικότερα ως εξής : Για να γίνει στρογγύλευση στα πρώτα n σημαντικά ψηφία εξετάζονται τα ψηφία n και n+1. Αν το n+1 ψηφίο είναι άρτιο, τότε εφαρμόζεται απλή αποκοπή από το n+1 ψηφίο και πέρα. Διαφορετικά, προστίθενται πέντε μονάδες στο n+1 ψηφίο και στη συνέχεια, στον αριθμό που προκύπτει,

εφαρμόζεται απλή αποκοπή από το ψηφίο n+1 και πέρα. Εύκολα επαληθεύεται ότι οι δύο ορισμοί της στρογγύλευσης είναι ισοδύναμοι.

Στα παραπάνω χρησιμοποιήθηκε το 10 ως βάση για την παράσταση των αριθμών, γιατί αυτό γίνεται όταν οι υπολογισμοί εκτελούνται από τον άνθρωπο. Στις μεγάλες υπολογιστικές μηχανές όμως, ως βάση χρησιμοποιείται το 2 ή ακόμη το 8 και το 16, γιατί έτσι ταιριάζει καλύτερα στη δομή των ηλεκτρονικών υπολογιστών.

Ακόμα, στους μεγάλους υπολογιστές οι αριθμοί παριστάνονται με μορφή κανονικοποιημένη και με σύστημα κινητής υποδιαστολής. Αν **x** είναι ένας πραγματικός αριθμός, τότε η παράστασή του με τον τρόπο αυτό συμβολίζεται με fl(x) :

$$fl(x) = \pm(d_1.\beta^{-1} + d_2.\beta^{-2} + \ldots + d_n.\beta^{-n}).\beta^{e} ,$$

όπου β είναι η βάση και d_i (i = 1, 2, . . . , n) τα ψηφία. Για τα d_i ισχύει :

$$1 \le d_1 \le \beta-1, \qquad 0 \le d_i \le \beta-1 \text{ για } i = 2, 3, \ldots, n.$$

Ο πεπερασμένος αριθμός n των πρώτων σημαντικών ψηφίων αποτελεί χαρακτηριστικό της μηχανής. Ο εκθέτης e είναι ακέραιος και κυμαίνεται ανάμεσα σε μία κατώτερη δυνατή τιμή -L και σε μία ανώτερη δυνατή τιμή U :

$$-L \le e \le U.$$

Τα L και U εκφράζουν χαρακτηριστικά της μηχανής.

Έστω **F** το σύνολο των αριθμών της παραπάνω μορφής. Το **F** είναι ένα σύνολο με πεπερασμένο αριθμό μελών και χρησιμοποιείται για την αναπαράσταση του συνόλου των πραγματικών αριθμών που έχει άπειρα

μέλη. Εύκολα μπορεί να αποδειχτεί ότι ο αριθμός των μελών του **F** είναι ίσος με $2.(\beta-1).\beta^{n-1}.(U-L+1)+1$.

Επιπλέον, πρέπει να σημειωθεί ότι οι αριθμοί του **F** δεν είναι ισοκατανεμημένοι επάνω στον άξονα των πραγματικών αριθμών.

Για κάθε πραγματικό αριθμό x βρίσκεται ένας αριθμός $fl(x) \in \mathbf{F}$, με αποκοπή ή στρογγύλευση. Προκύπτει συνεπώς η ανάγκη να εκτιμηθεί το σφάλμα που υπεισέρχεται στην προσέγγιση αυτή.

Ο πραγματικός αριθμός x μπορεί να γραφτεί ως εξής, με βάση το β :

$$x = \pm(d_1.\beta^{-1} + d_2.\beta^{-2} + \ldots + d_n.\beta^{-n} + d_{n+1}.\beta^{-n-1} + d_{n+2}.\beta^{-n-2} + \ldots).\beta^e.$$

Τότε για την περίπτωση της αποκοπής,

$$|\Delta_x| = |x - fl(x)| = (d_{n+1}.\beta^{-n-1} + d_{n+2}.\beta^{-n-2} + \ldots).\beta^e =$$

$$= \beta^{-n}.(d_{n+1}.\beta^{-1} + d_{n+2}.\beta^{-2} + \ldots).\beta^e \leq \beta^{-n}.\beta^e .$$

Αλλά $|\Delta_x| = |x|.|\delta_x|$.

Άρα, $$\left|\delta_x \leq \frac{\beta^{-n}.\beta^e}{|x|}\right|$$

Από την παράσταση του x,

$$|x| = \beta^{-1}.(d_1 + d_2.\beta^{-1} + \ldots + d_n.\beta^{1-n} + d_{n+1}.\beta^{-n} + d_{n+2}.\beta^{-1-n} + \ldots).\beta^e$$

$$> d_1.\beta^{-1}.\beta^e$$

Άρα, $\frac{1}{|x|} < \frac{\beta}{d_1.\beta^e}$ και $|\delta_x| < \frac{1}{d_1}.\beta^{1-n}$

Συνεπώς,

Θεώρημα

Το σχετικό σφάλμα που υπεισέρχεται στην αναπαράσταση fl(x) ενός πραγματικού αριθμού x είναι σε απόλυτη τιμή μικρότερο από, για την $\frac{1}{d_1}.\beta^{1-n}$ περίπτωση της αποκοπής.

Επειδή $d_1 \geq 1$, προκύπτει ότι $|\delta_x| < \beta^{1-n}$.

Με την τελευταία ανισότητα συνδέεται το σφάλμα κατευθείαν με τα χαρακτηριστικά της υπολογιστικής μηχανής.

Ο δεύτερος από τους ορισμούς της στρογγύλευσης που δόθηκαν παραπάνω, επιδέχεται τώρα άμεση γενίκευση : Για να γίνει στρογγύλευση στα πρώτα n ψηφία του κανονικοποιημένου ως προς βάση β πραγματικού αριθμού x, εξετάζονται τα ψηφία n και n+1.Αν το n+1 ψηφίο είναι μικρότερο ή ίσο με β/2, και επιπλέον το n ψηφίο είναι άρτιο, τότε εφαρμόζεται απλή αποκοπή από το ψηφίο n+1 και πέρα. Διαφορετικά, προστίθεται $0.5*\beta^{-n}$ στο κλασματικό μέρος της παράστασης του x και στη συνέχεια, στον αριθμό που προκύπτει εφαρμόζεται απλή αποκοπή από το ψηφίο n+1 και πέρα.

1.2 ΑΠΩΛΕΙΑ ΣΗΜΑΝΤΙΚΩΝ ΨΗΦΙΩΝ. ΑΣΤΑΘΕΙΑ

Τα σφάλματα που οφείλονται στην πεπερασμένη αναπαράσταση των πραγματικών αριθμών μεταδίδονται στις αριθμητικές πράξεις. Ιδιαίτερη προσοχή απαιτεί η πράξη της αφαίρεσης, όταν αφαιρούνται αριθμοί που δεν διαφέρουν πολύ μεταξύ τους. Έστω ότι ζητείται η διαφορά μεταξύ των αριθμών x = 48.128 και y = 48.107, που υποτίθενται γνωστοί με πέντε σωστά σημαντικά ψηφία. Αυτό σημαίνει ότι $|\Delta_x| \leq 0.5 *10^{+1-5+1} = 0.5 *10^{-3}$ και $|\Delta_y| \leq 0.5 *10^{-3}$.

Αν z = x - y, τότε προφανώς $|\Delta_z| \leq |\Delta_x| + |\Delta_y| \leq 10^{-3}$ ή $|\Delta_z| \leq 0.001$.

Γράφουμε τους x και y με την κανονικοποιημένη μορφή και αφαιρούμε :

$$z = 0.48128 * 10^2 - 0.48107 * 10^2 = 0.00021 * 10^2 = 0.21 * 10^{-1} .$$

Συνεπώς το z έχει μόνο δύο σημαντικά ψηφία από τα οποία το δεύτερο είναι αμφίβολης ορθότητας, επειδή $|\Delta_z|$ είναι της τάξης του $0.01*10^{-1}$. Αν τώρα παρατηρήσουμε τα επί μέρους σχετικά σφάλματα,

$$|\delta_x| \le \frac{0.5*10^{-3}}{0.48128*10^2} \approx 10^{-5}$$

$$|\delta_y| \le \frac{0.5*10^{-3}}{0.48107*10^2} \approx 10^{-5}$$

$$|\delta_z| \le \frac{10^{-3}}{0.21*10^{-1}} \approx 0.05 ,$$

το $|\delta_z|$ φαίνεται αμέσως ότι είναι αρκετές τάξεις μεγέθους μεγαλύτερο.

Ανάλογο φαινόμενο θα παρουσιαστεί στον υπολογισμό της διαφοράς $\sqrt{2.01}-\sqrt{2}$, ή γενικότερα στον υπολογισμό της παράστασης $\sqrt{x+\varepsilon}-\sqrt{x}$, όπου **ε** μικρός αριθμός. Στην περίπτωση αυτή, το μειονέκτημα της αφαίρεσης μπορεί να διορθωθεί, αν η παράσταση μετασχηματιστεί :

$$\sqrt{x+\varepsilon}-\sqrt{x} = \frac{\varepsilon}{\sqrt{x+\varepsilon}-\sqrt{x}} .$$

Αν εκτελεστούν οι πράξεις στη μετασχηματισμένη παράσταση, δεν θα συμβεί απώλεια σημαντικών ψηφίων. Το ίδιο ισχύει και για τη συνάρτηση f(x) = 1 - cos x, που παρουσιάζει απώλεια ακρίβειας όταν υπολογίζεται για μικρό x. Η f(x) μετασχηματίζεται ως εξής :

$$f(x) = 1 - \cos x = \frac{1-\cos^2 x}{1+\cos x} = \frac{\sin^2 x}{\cos x}$$

Μια γενικότερη μέθοδος για τον χειρισμό της f(x) στηρίζεται στο ανάπτυγμα Taylor του cos x :

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots$$

Χαρακτηριστικό παράδειγμα για την χρήση των σειρών Taylor προσφέρει η συνάρτηση $\dfrac{\tan x - \sin x}{x^3}$

Το θεώρημα της μέσης τιμής μπορεί επίσης να χρησιμοποιηθεί σε πολλές περιπτώσεις, π.χ. αν το x είναι μικρό,

$$\ln(a+x) - \ln a = \frac{x}{a+\theta} \approx \frac{x}{a+\frac{x}{2}}$$ όπου $0 < \theta < x$.

Ο μετασχηματισμός των παραπάνω παραστάσεων προτείνεται σαν άσκηση :

$\dfrac{1}{\sqrt{x+1}} - \dfrac{1}{x}$, $\cos(x+\varepsilon) - \cos x$, $\sqrt[3]{x+1} - \sqrt[3]{x}$,

$\dfrac{1}{\sqrt{x+1}} - \dfrac{1}{\sqrt{x}}$, $(x+\varepsilon)^{2/3} - x^{2/3}$, $\tan(x+\varepsilon) - \tan x$,

$1 - e^{-x}$, $\ln(1 - x)$.

Οι παραπάνω περιπτώσεις χαρακτηρίζονται από το γεγονός ότι μικρό σφάλμα στην ανεξάρτητη μεταβλητή επιφέρει πολύ μεγαλύτερο σφάλμα στον υπολογισμό της συνάρτησης. Εμφανίζεται δηλαδή αυξημένη "ευαισθησία" μιας συνάρτησης f(x) ως προς μεταβολές του x.

Ο λόγος (σε απόλυτη τιμή) του σχετικού σφάλματος της συνάρτησης προς το σχετικό σφάλμα της ανεξάρτητης μεταβλητής δίνει ένα ποσοτικό

κριτήριο και ένα μέτρο της ευαισθησίας της συνάρτησης για μια περιοχή τιμών της ανεξάρτητης μεταβλητής. Ο λόγος αυτός ονομάζεται **βαθμός κατάστασης**. και ορίζεται ως εξής :

Αν x είναι μία τιμή της ανεξάρτητης μεταβλητής και $\bar{x}$ μία προσεγγιστική της τιμή, τότε ο βαθμός κατάστασης (β.κ.) είναι ίσος με

$$\beta.\kappa. = \frac{\left|\dfrac{f(x)-f(\bar{x})}{f(x)}\right|}{\left|\dfrac{x-\bar{x}}{x}\right|} \approx \left|\frac{f'(x).x}{f(x)}\right|$$

Για παράδειγμα, έστω η συνάρτηση f(x) = 1 - cos x για μικρές τιμές του x. Στην περίπτωση αυτή,

$$\left|\frac{f'(x).x}{f(x)}\right| = \left|\frac{x.\sin x}{1-\cos x}\right| = \left|\frac{(1+\cos x).\sin x}{1-\cos^2 x}.x\right| = \left|\frac{(1+\cos x).x}{\sin^2 x}\right| \approx \frac{1+\cos x}{x}.$$

Είναι φανερό συνεπώς ότι για μικρές τιμές του x ο βαθμός κατάστασης μπορεί να πάρει τιμές απεριόριστα μεγάλες. Σαν άσκηση να βρεθεί ο βαθμός κατάστασης που αντιστοιχεί στη μετασχηματισμένη παράσταση

$$f(x) = \frac{\sin^2 x}{1+\cos x}$$

Ένα άλλο παράδειγμα μεγάλης ευαισθησίας προσφέρεται από την $f(x) = \dfrac{1}{1-x^2}$ για τιμές του x κοντά στο 1.

Στην περίπτωση αυτή, $\left|\dfrac{f'(x).x}{f(x)}\right| = \left|\dfrac{2x^2}{1-x^2}\right|$,

που πράγματι αυξάνει απεριόριστα όταν το x βρίσκεται στην περιοχή του 1.

Απεναντίας, η $f(x) = \sqrt{x}$ παρουσιάζει αρκετά μικρή τιμή του β.κ. :

$$\left|\frac{\left(\sqrt{x}\right)'.x}{\sqrt{x}}\right| = \left|\frac{x}{2.\sqrt{x}.\sqrt{x}}\right| = \frac{1}{2}.$$

Μια αριθμητική διαδικασία, ένας αλγόριθμος, αποτελείται από επί μέρους υπολογιστικά βήματα. Αν ένα ή περισσότερα από αυτά τα βήματα συνίστανται στον υπολογισμό μιας συνάρτησης με μεγάλο β.κ., τότε τα αριθμητικά σφάλματα, που υπεισέρχονται στον υπολογισμό της συνάρτησης αυτής, διαδίδονται και μεγεθύνονται ανεξέλεγκτα στην πορεία του αλγορίθμου.

Το φαινόμενο αυτό ονομάζεται **αστάθεια** και αναφέρεται στον αλγόριθμο συνολικά και όχι σε έναν περιορισμένο αριθμητικό υπολογισμό.

Άσκηση. Να αποδειχτεί ότι ο αριθμός των μελών του συνόλου **F** είναι ίσος με $2.(\beta-1).\beta^{n-1}.(U-L+1)+1$.

2. ΠΡΟΣΕΓΓΙΣΤΙΚΗ ΕΠΙΛΥΣΗ ΜΗ ΓΡΑΜΜΙΚΩΝ ΕΞΙΣΩΣΕΩΝ

2.1 ΜΕΘΟΔΟΣ ΣΥΝΑΡΤΗΣΙΑΚΩΝ ΕΠΑΝΑΛΗΨΕΩΝ

2.1.1 ΟΡΙΣΜΟΣ ΚΑΙ ΠΑΡΑΔΕΙΓΜΑΤΑ

Δίνεται η μη γραμμική εξίσωση μιας μεταβλητής

$$f(x) = 0. \tag{2.1}$$

Ζητείται να υπολογιστούν οι ρίζες της (2.1) σ' ένα διάστημα $a \le x \le b$.

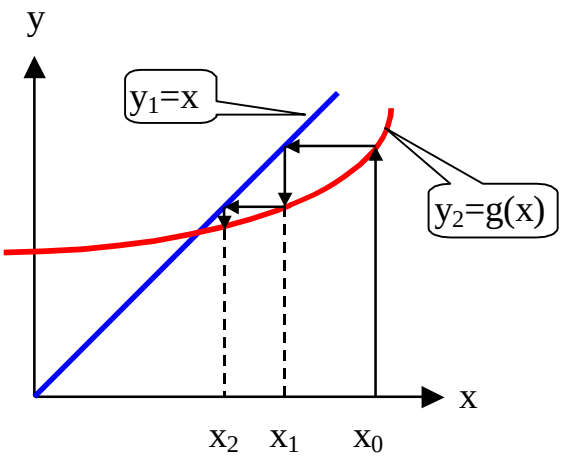
Η εξίσωση (2.1) γράφεται με την ισοδύναμη μορφή:

$$x = g(x) \tag{2.2}$$

Στη συνέχεια, για τον υπολογισμό των ριζών, προτείνεται η παρακάτω επαναληπτική διαδικασία:

$$x_{i+1} = g(x_i), \qquad i = 0, 1, 2, \ldots \tag{2.3}$$

όπου $\mathbf{x_0}$ είναι μια κατάλληλα εκλεγμένη **αρχική τιμή**.



Πρέπει να σημειωθεί ότι είναι πάντα δυνατό να γραφεί η (2.1) με τη μορφή (2.2). Ένας τρόπος για να γίνει αυτό είναι ο εξής :

$$x = x - \varphi(x) \cdot f(x), \qquad (2.4)$$

όπου $\varphi(x)$ συνάρτηση τέτοια που $0 < \varphi(x) < \infty$, για κάθε x πραγματικό.

Είναι προφανές ότι κάθε ρίζα της (2.4) είναι και ρίζα της (2.1) και αντίστροφα.

Ένα απλό παράδειγμα εφαρμογής της διαδικασίας (2.3) είναι το εξής:

Να υπολογιστούν οι ρίζες της δευτεροβάθμιας εξίσωσης

$$f(x) = 2x^2 - 7x + 3 = 0. \qquad (2.5)$$

Οι ρίζες είναι εκ των προτέρων γνωστές ($\rho_1 = 0.5$ και $\rho_2 = 3$) και έτσι θα καταστεί δυνατός ο έλεγχος των αποτελεσμάτων της προσεγγιστικής μεθόδου.

Η παραπάνω εξίσωση γράφεται με τη μορφή : $x = \dfrac{3 + 2x^2}{7}$, δηλαδή

$$g(x) = \frac{3 + 2x^2}{7} \quad . \qquad (2.6)$$

Αν τεθεί $x_0 = 0.2$, τότε η ακολουθία (2.3) συγκλίνει στη ρίζα $\rho_1 = 0.5$, όπως φαίνεται από τις παρατιθέμενες διαδοχικές τιμές. Το ίδιο συμβαίνει και αν ξεκινήσουμε με αρχικές τιμές $x_0 = 0.01$, $x_0 = -0.7$, $x_0 = 1.5$.

ΑΡΧΙΚΗ ΤΙΜΗ = .2

i	x
0	.2
1	.44
2	.4838857
3	.4954701
4	.4987116
5	.4996324
6	.499895
7	.49997
8	.4999914

ΑΡΧΙΚΗ ΤΙΜΗ = .001

i	x
0	.001
1	.4285717
2	.4810496
3	.4946882
4	.4984904
5	.4995693
6	.499877
7	.4999649
8	.49999

ΑΡΧΙΚΗ ΤΙΜΗ = -.7

i	x
0	-.7
1	.5685714
2	.5209353
3	.5061067
4	.5017554
5	.5005024
6	.5001436
7	.500041
8	.5000117

ΑΡΧΙΚΗ ΤΙΜΗ = 1.5

i	x
0	1.5
1	1.071429
2	.7565597
3	.5921093
4	.528741
5	.5084477
6	.502434
7	.5006971
8	.5001993
9	.500057
10	.5000163
11	.5000046

Αντίθετα, αν $x_0 = 3.1$, δηλαδή πολύ κοντά στην άλλη ρίζα, η ακολουθία των x_i αποκλίνει. Επίσης, η $x_0 = 2.98$ οδηγεί σε σύγκλιση, όχι όμως στη ρίζα $\rho_1 = 3$, όπως θα περίμενε κανείς, αλλά και πάλι στη ρίζα $\rho_1 = 0.5$.

ΑΡΧΙΚΗ ΤΙΜΗ = 3.1 **ΑΡΧΙΚΗ ΤΙΜΗ = 2.98**

i	x	i	x
0	**3.1**	0	**2.98**
1	3.174286	1	2.965829
2	3.307454	2	2.941754
3	3.554072	3	2.901119
4	4.037551	4	2.833284
5	5.086234	5	2.722142
6	7.819935	6	2.54573
7	17.90039	7	2.280212
8	91.97832	8	1.914105
9	2417.574	9	1.475371
10	1669905	10	1.050491
11	7.967381 E+11	11	0.743866
		12	0.5866676
		13	0.5269082
		14	0.5078949
		15	0.5022735
		16	0.5006511
		17	0.5001861
		18	0.5000532
		19	0.5000152
		20	0.5000044

Από το απλό αυτό παράδειγμα προκύπτει ότι η εξίσωση (2.3) υποδεικνύει μια μέθοδο, η οποία είναι δυνατό να οδηγήσει στον προσεγγιστικό υπολογισμό των ριζών μιας εξίσωσης, η ακολουθία όμως που ορίζεται από την εξίσωση (2.3) δεν συγκλίνει πάντα. Η συμπεριφορά της ακολουθίας x_i (σύγκλιση και όριο), όπως φάνηκε και από τα παραδείγματα, εξαρτάται από την αρχική τιμή x_0. Εξαρτάται όμως και από τη συνάρτηση g(x).

Πράγματι, για τον υπολογισμό των ριζών της παραπάνω δευτεροβάθμιας εξίσωσης, μπορεί να χρησιμοποιηθεί και η

$$g(x) = 3.5 - \frac{1.5}{x} \quad , \qquad (2.7)$$

Για να εξεταστεί η συμπεριφορά της εξίσωσης (2.7), δίνεται μια σειρά αποτελεσμάτων για την περίπτωση αυτή και για διάφορες αρχικές τιμές.

ΑΡΧΙΚΗ ΤΙΜΗ = 2 ΑΡΧΙΚΗ ΤΙΜΗ = 5 ΑΡΧΙΚΗ ΤΙΜΗ = .495

i	**x**	**i**	**x**	**i**	**x**
0	**2**	0	**5**	0	**.495**
1	2.75	1	3.2	1	.469697
2	2.954545	2	3.03125	2	.3064518
3	2.992308	3	3.005155	3	-1.394734
4	2.998715	4	3.000858	4	4.575474
5	2.999786	5	3.000143	5	3.172165
6	2.999964	6	3.000024	6	3.027137
7	2.999994	7	3.000004	7	3.004482
				8	3.000746
				9	3.000124
				10	3.000021
				11	3.000003

Παρατηρούμε ότι αυτή τη φορά η επαναληπτική διαδικασία συγκλίνει προς τη ρίζα $\rho_2 = 3$, ακόμα κι όταν η αρχική τιμή βρίσκεται πολύ κοντά στην άλλη ρίζα.

2.1.2 ΘΕΩΡΙΑ ΣΥΓΚΛΙΣΗΣ

Από όλα αυτά τα αριθμητικά πειράματα προκύπτει η ανάγκη για την αναζήτηση κριτηρίων, με βάση τα οποία, σε ορισμένες τουλάχιστον περιπτώσεις, θα είναι εκ των προτέρων γνωστή η συμπεριφορά της ακολουθίας $\{x_i\}$.

Για να εξεταστεί η σύγκλιση της ακολουθίας (2.3), σχηματίζεται η ποσότητα:

$$\varepsilon_i = x_i - \xi ,$$

όπου ξ είναι ρίζα της (2.1) και επομένως και της (2.2), δηλαδή $\xi = g(\xi)$. Αν η διαδικασία στοχεύει στη ρίζα ξ, τότε η ε_i παριστάνει το "σφάλμα" που περιέχεται στην i προσέγγιση. Η μέθοδος των διαδοχικών προσεγγίσεων θεωρείται ότι συγκλίνει, αν τα σφάλματα μικραίνουν κατ' απόλυτο τιμή, καθώς η διαδικασία προχωρεί από βήμα σε βήμα.

Εξετάζονται συνεπώς τα διαδοχικά σφάλματα:

$$\varepsilon_{i+1} = x_{i+1} - \xi = g(x_i) - g(\xi) = g'(\xi_i)\cdot(x_i - \xi) = g'(\xi_i)\cdot\varepsilon_i$$

όπου το ξ_i μια τιμή μεταξύ x_i και ξ, σύμφωνα με το θεώρημα της μέσης τιμής. Αν θεωρήσουμε τις απόλυτες τιμές, τότε,

$$|\varepsilon_{i+1}| \leq |g'(\xi_i)|\cdot|\varepsilon_i|.$$

Για να είναι $|\varepsilon_{i+1}| < |\varepsilon_i|$, θα πρέπει $|g'(\xi_i)| < 1$.

Συνεπώς μπορούμε να ξεκινήσουμε με κάποιο διάστημα $I = [a, b]$, που να περιέχει τη ρίζα ξ, και στο οποίο $|g'(x)| < 1 \ \forall x \in I$. Ακριβέστερα, υποθέτουμε ότι $|g'(x)| \leq K \ \forall x \in I$, όπου K θετική σταθερά με $K < 1$.
Στο σημείο αυτό μπορούμε να αποδείξουμε ότι η ξ είναι η μοναδική ρίζα της $x = g(x)$ μέσα στο διάσημα I. Πράγματι, αν υποθέσουμε ότι υπάρχουν δύο ρίζες ξ_1 και ξ_2 στο I, τότε

$$|\xi_1 - \xi_2| = |g(\xi_1) - g(\xi_2)| = |g'(\xi_{1,2})| \cdot |\xi_1 - \xi_2| \leq K \cdot |\xi_1 - \xi_2| < |\xi_1 - \xi_2|$$

όπου $\xi_{1,2}$ μια τιμή ανάμεσα στις δύο ρίζες και συνεπώς $\xi_{1,2} \in I$.

Προκύπτει αμέσως ότι $1 < 1$, πράγμα άτοπο.

Έχει συνεπώς αποδειχτεί το θεώρημα:

Θεώρημα 1:

Έστω διάστημα $I = [a, b]$*, που να περιέχει μια ρίζα* ξ *της* $x = g(x)$*, και στο οποίο* $|g'(x)| \leq K < 1 \ \forall x \in I$.

Τότε δεν υπάρχει άλλη ρίζα μέσα στο ίδιο διάστημα.

Μέσα στο διάστημα I του Θεωρήματος 1, λογικό είναι να τοποθετήσουμε την αρχική τιμή x_0. Εξετάζουμε τώρα, αν και οι μετέπειτα προσεγγίσεις βρίσκονται μέσα στο I. Υποθέτουμε ότι $x_i \in I$. Τότε,

$$|x_{i+1} - \xi| = |g(x_i) - g(\xi)| = |g'(\xi_i)| \cdot |x_i - \xi| \leq K \cdot |x_i - \xi|.$$

Συνεπώς, $|x_{i+1} - \xi| < |x_i - \xi|$, πράγμα που σημαίνει ότι η απόσταση του x_{i+1} από τη ρίζα ξ είναι μικρότερη από την απόσταση του x_i από την ξ. Είναι όμως δυνατό τα x_i και x_{i+1} να βρίσκονται το ένα αριστερά και το άλλο δεξιά από την ξ. Για τον λόγο αυτό γίνεται και η πρόσθετη υπόθεση

ότι το διάστημα I είναι συμμετρικό ως προς τη ρίζα ξ, οπότε προκύπτει ότι $x_{i+1} \in I$. Σύμφωνα δε με τη μέθοδο της επαγωγής, επειδή $x_0 \in I$, έπεται ότι $x_i \in I$, i = 0, 1, 2, . . .

Επίσης, από την παραπάνω ανισότητα προκύπτει ότι:

$$|x_{i+1} - \xi| \leq K . |x_i - \xi| \leq K^2 \cdot |x_{i-1} - \xi| \leq \ldots \leq K^{i+1} \cdot |x_0 - \xi| .$$

Αλλά, $K^{i+1} \to 0$, καθώς το $i \to \infty$, επειδή $K < 1$. Επομένως,

$$x_i \to \xi, \text{ καθώς το } i \to \infty.$$

Έχει συνεπώς αποδειχτεί το παρακάτω θεώρημα:

Θεώρημα 2:

Έστω ξ μια ρίζα της f(x) = 0*, ή της ισοδύναμής της* x = g(x) *και ένα διάστημα ακτίνας* h *γύρω από το σημείο* x = ξ*. Ακόμη,* οι g(x), g΄(x) *υποτίθενται συνεχείς στο διάστημα* I *με* $|g'(x)| \leq K < 1$*. Τότε, αν* $x_0 \in I$*, η διαδικασία* $x_{i+1} = g(x_i)$ *συγκλίνει στο ξ. Επίσης, όλες οι ενδιάμεσες προσεγγίσεις βρίσκονται μέσα στο διάστημα* I.

Τα θεωρήματα 1 και 2 μπορούν να χρησιμοποιηθούν για την πρακτική εφαρμογή της μεθόδου των συναρτησιακών επαναλήψεων.

Πράγματι, για την περίπτωση της $g(x) = \dfrac{3 + 2x^2}{7}$

από τη συνθήκη $|g'(x)| < 1$ προκύπτει $\left|\dfrac{4x}{7}\right| < 1$ ή $|x| < 1.75$.

Εξάλλου, f(-1.74) . f(1.74) < 0, όπου η f δίνεται από την (2. 5).

Συνεπώς, το διάστημα I = [-1.74, 1.74] περιέχει μια τουλάχιστον ρίζα και ακόμη $|g'(x)| < 1 \quad \forall\, x \in I$. Από το Θεώρημα 1 προκύπτει αμέσως ότι υπάρχει μόνο μια ρίζα στο διάστημα I. Η θέση όμως της ρίζας προφανώς δεν είναι γνωστή, άρα δεν είναι γνωστό αν ικανοποιούνται οι προϋποθέσεις του Θεωρήματος 2 και έτσι δεν είναι βέβαιο αν όλες οι ενδιάμεσες προσεγγίσεις θα βρίσκονται μέσα στο διάστημα I.

Ήδη έχουν δοκιμαστεί οι αρχικές τιμές $x_0 = 0.01$, $x_0 = -0.7$, $x_0 = 1.5$. Και οι τρεις βρίσκονται μέσα στο I και σε καμιά περίπτωση δεν βρέθηκε ενδιάμεση προσέγγιση έξω από το διάστημα I. Το ίδιο θα συμβεί και με τις αρχικές τιμές $x_0 = -1.0$ και $x_0 = -1.5$, όπως φαίνεται παρακάτω.

ΑΡΧΙΚΗ ΤΙΜΗ = -1 ΑΡΧΙΚΗ ΤΙΜΗ = -1.5

i	**x**	**i**	**x**
0	-1	0	-1.5
1	.7142857	1	1.071429
2	.574344	2	.7565597
3	.5228203	3	.5921093
4	.5066689	4	.528741
5	.5019181	5	.5084477
6	.5005491	6	.502434
7	.5001569	7	.5006971
8	.5000448	8	.5001993
9	.5000128	9	.500057
		10	.5000163
		11	.5000046

Ας σημειωθεί ότι η ρίζα είναι $\rho_1 = 0.5$, άρα εκ των υστέρων γνωρίζουμε ότι ένα διάστημα $I_1 = [-0.74, 1.74]$, ως υποσύνολο του I και ως συμμετρικό ως προς τη ρίζα, ικανοποιεί τις απαιτήσεις του Θεωρήματος 2 και επομένως εξηγείται το γεγονός ότι οι τρεις πρώτες προσεγγίσεις οδηγούν σε σύγκλιση και ότι όλες οι ενδιάμεσες προσεγγίσεις βρίσκονται μέσα στο I_1. Δεν καλύπτεται θεωρητικά η συμπεριφορά των δύο τελευταίων αρχικών τιμών, που βρίσκονται έξω από το I_1. Το θέμα αυτό όμως θα αντιμετωπιστεί με βάση γενικότερα θεωρήματα που θα αναπτυχθούν παρακάτω.

2.1.3 ΤΡΟΠΟΠΟΙΗΣΗ ΤΗΣ ΜΕΘΟΔΟΥ ΤΩΝ ΣΥΝΑΡΤΗΣΙΑΚΩΝ ΕΠΑΝΑΛΗΨΕΩΝ

Προς το παρόν χαρακτηριστικό είναι το παράδειγμα που ακολουθεί και για το οποίο ενδιάμεσες προσεγγίσεις πέφτουν έξω από ένα διάστημα για το οποίο $|g'(x)| < 1$. Πράγματι, έστω η εξίσωση

$$x = e^{-0.1x} + \frac{\cos x}{1.09} \tag{2.8}$$

Η εξίσωση (2.8) είναι ήδη γραμμένη με τη μορφή $x = g(x)$. Η παράγωγος της $g(x)$ είναι:

$$g'(x) = -0.1 * e^{-0.1x} - \frac{\sin x}{1.09} \tag{2.9}$$

Από την (2.9) προκύπτει εύκολα ότι αν $I = [0.7, 1.4]$, τότε $|g'(x)| < 1$.

Επίσης, $[0.7 - g(0.7)] \cdot [1.4 - g(1.4)] < 0$.

Συνεπώς το διάστημα I περιέχει μια τουλάχιστον ρίζα της εξίσωσης (2.8) και, σύμφωνα με το Θεώρημα 1, μία μόνο ρίζα.

Έστω τώρα $x_0 = 0.87$, δηλ. $x_0 \in I$. Παρατηρούμε τότε ότι η τρίτη προσέγγιση έχει ήδη ξεφύγει από το διάστημα I, όπως φαίνεται στον παρακάτω πίνακα, όπου επίσης σημειώνονται και οι τιμές της $g'(x)$ για κάθε προσέγγιση. Η μέθοδος των συναρτησιακών επαναλήψεων δεν σημειώνει σύγκλιση ούτε μετά από την 100η προσέγγιση.

ΑΡΧΙΚΗ ΤΙΜΗ = .87

	x	g(x)	g´(x)	x_1
0	.87	1.508261	-.7928869	1.508261
1	1.508261	.9173316	-1.001638	.9173316
2	.9173316	1.470093	-.8196589	1.470093
3	1.470093	.9555187	-.9991118	.9555187
4	.9555187	1.438399	-.8400736	1.438399
5	1.438399	.9871373	-.9960047	.9871373
6	.9871373	1.411581	-.8561523	1.411581
7	1.411581	1.013805	-.9926627	1.013805
8	1.013805	1.388575	-.8691198	1.388575
9	1.388575	1.036604	-.989277	1.036604
10	1.036604	1.368638	-.8797678	1.368638
11	1.368638	1.056294	-.9859571	1.056294
12	1.056294	1.351227	-.8886344	1.351227
13	1.351227	1.073434	-.9827657	1.073434
14	1.073434	1.335932	-.8961006	1.335932
15	1.335932	1.088442	-.9797385	1.088442
16	1.088442	1.322435	-.9024442	1.322435
17	1.322435	1.101646	-.976894	1.101646
18	1.101646	1.310483	-.9078741	1.310483
19	1.310483	1.113307	-.9742399	1.113307
87	1.207366	1.212398	-.9461339	1.212398

88	1.212398	1.207633	-.9477198	1.207633
89	1.207633	1.212146	-.9462186	1.212146
90	1.212146	1.207872	-.9476406	1.207872
91	1.207872	1.211919	-.9462945	1.211919
92	1.211919	1.208087	-.9475696	1.208087
93	1.208087	1.211716	-.9463625	1.211716
94	1.211716	1.20828	-.9475059	1.20828
95	1.20828	1.211534	-.9464235	1.211534
96	1.211534	1.208452	-.9474487	1.208452
97	1.208452	1.21137	-.9464781	1.21137
98	1.21137	1.208607	-.9473975	1.208607
99	1.208607	1.211224	-.9465271	1.211224
100	1.211224	1.208746	-.9473515	1.208746

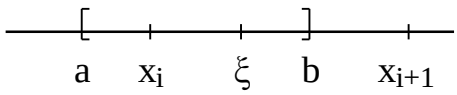
Η προβληματική κατάσταση που προέκυψε στο παραπάνω παράδειγμα μπορεί να διορθωθεί με την εξής τροποποίηση της μεθόδου των συναρτησιακών επαναλήψεων, όπως αυτή αρχικά ορίστηκε από τις εξισώσεις (2.3) :

Έστω ότι $x_i \in I$ και $x_{i+1} \notin I$. Τότε το x_{i+1} αντικαθίσταται με τη διορθωμένη τιμή

$$x_{i+1}^{*} = \frac{x_i + x_{i+1}}{2} . \qquad (2.10)$$

Αποδεικνύεται ότι $x_{i+1}^{*} \in I$.

Πράγματι, μπορούμε να υποθέσουμε χωρίς περιορισμό της γενικότητας, ότι $x_i < \xi$, όπου ξ η ρίζα που περιέχεται στο I.



Τότε διακρίνονται δύο περιπτώσεις:

(α) $x_{i+1} > \xi$, οπότε $x_{i+1} > b$, διότι $x_{i+1} \notin I$.

(β) $x_{i+1} < \xi$.

Η περίπτωση αυτή αποκλείεται επειδή ισχύει ότι

$$|x_{i+1} - \xi| < |x_i - \xi|. \qquad (2.11)$$

Αυτό όμως θα σήμαινε ότι $x_{i+1} \in I$, ενώ υποτέθηκε ότι $x_{i+1} \notin I$.

Συνεπώς, $x_{i+1} > \xi$ και επομένως $|x_{i+1} - \xi| = x_{i+1} - \xi$ και $|x_i - \xi| = \xi - x_i$.

Τότε, η παραπάνω ανισότητα (2.11) γράφεται ως εξής:

$$x_{i+1} - \xi < \xi - x_i .$$

Προκύπτει συνεπώς ότι $x_{i+1} + x_i < 2\xi$. Άρα,

$$x_{i+1}^* = \frac{x_i + x_{i+1}}{2} < \xi \qquad \text{και τελικά, } x_{i+1}^* \in I.$$

Απομένει να εξεταστεί η συμπεριφορά του σφάλματος ε_i για την περίπτωση κατά την οποία εφαρμόζεται η διορθωτική εκτίμηση της εξίσωσης (2.10).

Έστω $\varepsilon_i^* = x_i^* - \xi$. Τότε,

$$2\varepsilon_{i+1}^* = 2(x_{i+1}^* - \xi) = (x_i - \xi) + (x_{i+1} - \xi) = (x_i - \xi) + g(x_i) - g(\xi) =$$

$$= (x_i - \xi) + (x_i - \xi).g'(\xi_i) = (x_i - \xi) .[1 + g'(\xi_i)] = \varepsilon_i .[1 + g'(\xi_i)].$$

$$\text{Άρα,} \quad \varepsilon_{i+1}^* = \varepsilon_i [1 + g'(\xi_i)] / 2.$$

Στα παραπάνω, η ξ_i είναι μια τιμή της x ανάμεσα στα x_i και ξ, σύμφωνα με το θεώρημα της μέσης τιμής. Συνεπώς $\xi_i \in I$ και $|g'(\xi_i)| \leq K < 1$, αν ακολουθήσουμε τις προϋποθέσεις του Θεωρήματος 2. Άρα,

$$|1 + g'(\xi_i)| \le 1 + |g'(\xi_i)| \le 1 + K.$$

Επομένως,

$$|\varepsilon_{i+1}^*| < K^*.|\varepsilon_i|, \text{ όπου } \quad K^* = \frac{1+K}{2}.$$

Προφανώς, $K < K^* < 1$.

Αν εισαχθεί τώρα το ενιαίο σύμβολο e_i για το σφάλμα του i βήματος, τότε μπορεί να γραφεί η ανισότητα:

$$|e_{i+1}| < K^* .|e_i| < K^{*2} .|e_{i-1}| < \ldots\ldots < K^{*i+1}. |e_0|.$$

Εντελώς ανάλογα με το Θεώρημα 2,

$e_{i+1} \to 0$, καθώς $i \to \infty$. Άρα, $x_i \to \xi$.

Έχει συνεπώς αποδειχτεί το παρακάτω θεώρημα:

Θεώρημα 3:

Έστω ξ *μια ρίζα της* $f(x) = 0$*, ή της ισοδύναμής της* $x = g(x)$ *και ένα διάστημα* I*, που περιέχει την* ξ*. Ακόμη, οι* $g(x)$, $g'(x)$ *υποτίθενται συνεχείς στο διάστημα* I *με* $|g'(x)| \le K < 1$*. Τότε, έστω* $x_0 \in I$ *και*

$$x_{i+1} = \begin{cases} g(x_i), & \alpha\nu\ g(x_i) \in I \\ \dfrac{x_i + g(x_i)}{2}, & \alpha\nu\ g(x_i) \notin I \end{cases}$$

Τότε η ακολουθία $\{x_i\}$ *συγκλίνει στο* ξ*. Επίσης, όλες οι ενδιάμεσες προσεγγίσεις βρίσκονται μέσα στο διάστημα* I.

Ο αλγόριθμος του Θεωρήματος 3 εφαρμόζεται τώρα στο παραπάνω πρόβλημα με την ίδια αρχική τιμή. Η διαδικασία αυτή τη φορά συγκλίνει.

ΑΡΧΙΚΗ ΤΙΜΗ = .87

	x	g(x)	g´(x)	x_1
0	.87	1.508261	-.7928869	1.508261
1	1.508261	.9173316	-1.001638	1.212796
2	1.212796	1.207256	-.9478442	1.207256
3	1.207256	1.212502	-.946099	1.212502
4	1.212502	1.207534	-.9477523	1.207534
5	1.207534	1.212239	-.9461873	1.212239
6	1.212239	1.207784	-.9476699	1.207784
7	1.207784	1.212003	-.9462665	1.212003
8	1.212003	1.208008	-.9475958	1.208008
9	1.208008	1.211791	-.9463374	1.211791

. . . . . .

100	1.209937	1.209965	-.9469468	1.209965
101	1.209965	1.209938	-.9469556	1.209938
102	1.209938	1.209963	-.9469472	1.209963
103	1.209963	1.20994	-.9469551	1.20994
104	1.20994	1.209962	-.9469476	1.209962
105	1.209962	1.209941	-.9469547	1.209941
106	1.209941	1.209961	-.9469481	1.209961
107	1.209961	1.209942	-.9469544	1.209942
108	1.209942	1.20996	-.9469483	1.20996
109	1.20996	1.209943	-.9469541	1.209943
110	1.209943	1.209959	-.9469486	1.209959
111	1.209959	1.209944	-.9469538	1.209944
112	1.209944	1.209958	-.9469489	1.209958
113	1.209958	1.209944	-.9469535	1.209944
114	1.209944	1.209957	-.9469492	1.209957
115	1.209957	1.209945	-.9469533	1.209945

116	1.209945	1.209957	-.9469494	1.209957
117	1.209957	1.209946	-.9469531	1.209946
118	1.209946	1.209956	-.9469496	1.209956
119	1.209956	1.209946	-.9469529	1.209946

Παρατηρούμε ότι, για το συγκεκριμένο τουλάχιστον πρόβλημα, η σύγκλιση είναι βραδεία. Και αυτό μπορεί να διορθωθεί, αν υιοθετηθεί για όλες τις ενδιάμεσες προσεγγίσεις ο τύπος

$$x_{i+1} = \frac{x_i + g(x_i)}{2}$$

Στην περίπτωση αυτή λαμβάνονται οι ακόλουθες διαδοχικές προσεγγίσεις:

ΑΡΧΙΚΗ ΤΙΜΗ = .87

	x	**g(x)**	**g΄(x)**	**x_1**
0	.87	1.508261	-.7928869	1.189131
1	1.189131	1.229598	-.9402063	1.209364
2	1.209364	1.210507	-.9467662	1.209936
3	1.209936	1.209966	-.9469464	1.209951
4	1.209951	1.209952	-.9469512	1.209951

Η ταχεία σύγκλιση εξηγείται από το γεγονός ότι μέσα στο υπό μελέτη διάστημα I η παράγωγος g′(x) είναι αρνητική. Για τον λόγο αυτό, οι διαδοχικές προσεγγίσεις παρουσιάζουν μια ταλάντωση γύρω από τη ρίζα ξ και έτσι ο μέσος όρος των δύο τελευταίων προσεγγίσεων επιταχύνει

δραστικά τη σύγκλιση. Η συμπεριφορά αυτή μπορεί να γίνει κατανοητή με ένα κατάλληλο διάγραμμα.

2.1.4 ΓΕΝΙΚΕΥΣΗ ΤΩΝ ΘΕΩΡΗΜΑΤΩΝ ΣΥΓΚΛΙΣΗΣ

Τα Θεωρήματα 1 και 2 που αποδείχτηκαν παραπάνω δεν παρέχουν εξήγηση της συμπεριφοράς της συνάρτησης g(x) σε όλες τις περιπτώσεις. Αυτό ήδη φάνηκε στην περίπτωση της εξίσωσης (2.6). Η ακολουθία των x_i παρουσίασε σύγκλιση ακόμα και για αρχικές τιμές έξω από το διάστημα I, όπως αυτό προέκυπτε από το Θεώρημα 2. Το γεγονός αυτό μας οδηγεί στην αναζήτηση γενικότερων θεωρημάτων.

Το Θεώρημα 1 εξετάζει τη μοναδικότητα της ρίζας, όχι όμως και την ύπαρξή της. Ένα άλλο πρόβλημα, όπως ήδη συζητήθηκε, ήταν η παραμονή των ενδιαμέσων προσεγγίσεων μέσα στο διάστημα I. Το πρόβλημα αυτό αντιμετωπίστηκε με την τροποποίηση της μεθόδου των απλών συναρτησιακών επαναλήψεων. Μια άλλη αντιμετώπιση συνίσταται στην επιβολή άνω και κάτω φράγματος στη συνάρτηση g(x):

Αν $x \in I = [a, b]$, τότε $a \leq g(x) \leq b$.

Πράγματι, είναι αμέσως προφανές, ότι αν $x_i \in I$, τότε $x_{i+1} = g(x_i) \in I$.

Η συνθήκη αυτή θα μπορούσε συνεπώς να χρησιμοποιηθεί για τη διατύπωση ενός θεωρήματος εναλλακτικού ως προς το Θεώρημα 2. Σημαντικότερο όμως είναι το γεγονός ότι με βάση τη συνθήκη αυτή μπορεί να αποδειχτεί η ύπαρξη ρίζας μέσα στο διάστημα I. Το θεώρημα που ακολουθεί είναι γνωστό ως θεώρημα σταθερού σημείου:

Θεώρημα 4:

Έστω g(x) *συνάρτηση συνεχής στο διάστημα* I = [a, b], *με* a □ g(x) □b. *Τότε υπάρχει τουλάχιστον μία τιμή* ξ □I, *τέτοια που* ξ = g(ξ).

Απόδειξη:

$$a \le g(\alpha) \Rightarrow a - g(a) \le 0$$

$$g(b) \le b \Rightarrow b - g(b) \le 0.$$

Θέτουμε τώρα $\varphi(x) = x - g(x)$. Τότε, $\varphi(a) \le 0$ και $\varphi(b) \ge 0$.
Επειδή τώρα η συνάρτηση φ είναι συνεχής, υπάρχει μία τουλάχιστον τιμή $\xi \in I$, τέτοια που $\varphi(\xi) = 0$, ή $\xi = g(\xi)$, ο.ε.δ.

Τα θεωρήματα και οι παρατηρήσεις που αναπτύχθηκαν ως εδώ μπορούν να ενσωματωθούν στο παρακάτω θεώρημα.

Θεώρημα 5

Έστω διάστημα $I = [a, b]$ *και συνεχής συνάρτηση* $g(x)$ *που ορίζεται στο διάστημα αυτό και έχει τις εξής ιδιότητες:*

(i) $a \le g(x) \le b \quad \forall x \in I$.
(ii) $|g'(x)| \le K \ \forall x \in I$, *όπου* K *θετική σταθερά με* $K < 1$.

Τότε υπάρχει ρίζα ξ *της* $x = g(x)$ *μέσα στο διάστημα* I, *η ρίζα αυτή είναι μοναδική και, ακόμα, η ακολουθία* $\{x_i\}$ *που ορίζεται από την αναδρομική σχέση* $x_{i+1} = g(x_i)$ *με* $x_0 \in I$, *συγκλίνει με όριο την* ξ *και με όλα τα ενδιάμεσα* x_i *μέσα στο I.*

Η απόδειξη παραλείπεται, δεδομένου ότι τα επί μέρους στοιχεία που απαιτούνται γι' αυτήν υπάρχουν ήδη στις αποδείξεις των θεωρημάτων που προηγήθηκαν.

Το Θεώρημα 5 παρέχει πλήρη πρόβλεψη της συμπεριφοράς της συνάρτησης

$$g(x) = \frac{3 + 2x^2}{7}$$

σε ολόκληρο το διάστημα [-1.74, 1.74], δεδομένου ότι $0.428 \le g(x) \le 1.294$ για $x \in [-1.74, 1.74]$.

Η συνθήκη *(ii)* του τελευταίου Θεωρήματος επιδέχεται γενίκευση. Πρώτα όμως δίνεται ο εξής ορισμός:

Ορισμός: *Έστω διάστημα* $I = [a, b]$ *και συνάρτηση* $g(x)$ *που ορίζεται στο διάστημα αυτό. Αν* $|g(x) - g(y)| \leq L.|x - y|$*,όπου* L *θετική σταθερά, τότε η* $g(x)$ *ονομάζεται συνεχής κατά Lipschitz.*

Εύκολα αποδεικνύεται το εξής λήμμα:

Λήμμα: *Αν οι* $g(x)$, $g'(x)$ *είναι συνεχείς στο* I*, τότε είναι και συνεχείς κατά Lipschitz στο ίδιο διάστημα.*

<u>*Απόδειξη*</u> : Πράγματι, για $x, y \in I$, $g(x) - g(y) = g'(\eta).(x - y)$, όπου η μια τιμή ανάμεσα στα x και y, σύμφωνα με το θεώρημα της μέσης τιμής.
Αλλά η $g'(x)$, ως συνεχής, είναι φραγμένη στο διάστημα I, δηλ. $|g'(x)| \leq K$ $\forall x \in I$, όπου K θετική σταθερά.
Συνεπώς, $|g(x) - g(y)| \leq |g'(\eta)|.|x - y| \leq K.|x - y|$, ο.ε.δ.
Το θεώρημα που ακολουθεί βασίζεται στην έννοια της συνέχειας κατά Lipschitz και αποδεικνύεται εύκολα:

Θεώρημα 6
Έστω διάστημα $I = [a, b]$ *και συνάρτηση* $g(x)$ *που ορίζεται στο διάστημα αυτό και έχει τις εξής ιδιότητες:*
(i) $a \leq g(x) \leq b$ $\forall x \in I$.
(ii) Η $g(x)$ *είναι συνεχής κατά Lipschitz στο* I*, με σταθερά* $L < 1$.
Τότε υπάρχει ρίζα ξ *της* $x = g(x)$ *μέσα στο διάστημα* I*, η ρίζα αυτή είναι μοναδική και, ακόμα, η ακολουθία* $\{x_i\}$ *που ορίζεται από την αναδρομική*

σχέση $x_{i+1} = g(x_i)$ *με* $x_0 \in I$, *συγκλίνει με όριο την* ξ *και με όλα τα ενδιάμεσα* x_i *μέσα στο* I.

Η απόδειξη είναι εντελώς ανάλογη με αυτήν του προηγούμενου θεωρήματος. Παρατηρούμε, ότι με βάση το παραπάνω Λήμμα, το Θεώρημα 5 εμφανίζεται τώρα ως ειδική περίπτωση του Θεωρήματος 6.

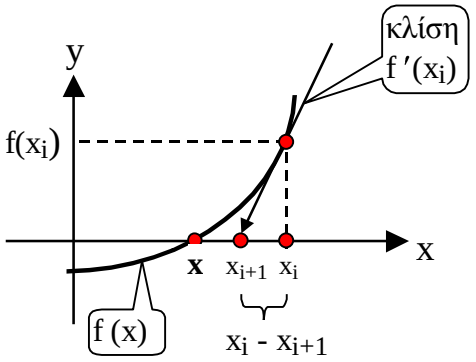
2.2 ΜΕΘΟΔΟΣ NEWTON - RAPHSON

Η μέθοδος αυτή συνίσταται στην παρακάτω επαναληπτική διαδικασία :

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)} \ ,$$

$$i = 0, 1, 2, \ldots \qquad (2.12)$$

όπου $\mathbf{x_0}$ κατάλληλα εκλεγμένη **αρχική τιμή**.



Γραφική απεικόνιση της μεθόδου Newton - Raphson

Η διαδικασία (2.12) μπορεί να γραφτεί με τον συμβολισμό της προηγούμενης παραγράφου, αν τεθεί

$$g(x) = x - \frac{f(x)}{f'(x)} \ .$$

Παρατηρούμε ότι

$$g'(x) = \frac{f(x).f''(x)}{[f'(x)]^2} \ .$$

Αν λοιπόν ξ είναι μια ρίζα της $f(x) = 0$, τότε $g'(\xi) = 0$, επειδή $f(\xi) = 0$ και υπό την προϋπόθεση ότι $f'(\xi) \neq 0$.

Ας υποθέσουμε ότι οι $f(x)$, $f'(x)$, $f''(x)$ είναι συνεχείς σε ένα διάστημα I που περιέχει την $x = \xi$.

Τότε, εξαιτίας της συνέχειας της $f'(x)$, υπάρχει ένα I_1 που περιέχει την ξ και στο οποίο $f'(x) \neq 0$, $\forall\ x \in I_1$.

Είναι όμως συνεχής και η $g'(x)$ μέσα στο διάστημα I, πράγμα που προκύπτει από τη συνέχεια των f, f ΄ και f ΄΄ στο ίδιο διάστημα και από το γεγονός ότι $f'(\xi) \neq 0$. Συνεπώς, επειδή $g'(\xi) = 0$, προκύπτει ότι υπάρχει διάστημα I_2 που περιέχει την ξ και στο οποίο $|g'(x)| < 1$ $x \in I_2$.

Τα διαστήματα I_1 και I_2 μπορούν να θεωρηθούν ότι περιέχονται εξ ολοκλήρου στο I. Έστω τώρα $I_3 = I_1 \cap I_2$. Τότε $x \in I_3$, η $g(x)$ είναι συνεχής και, επί πλέον, $|g'(x)| < 1$. Επίσης, το I_3 μπορεί να θεωρηθεί συμμετρικό ως προς τη ρίζα ξ.

Άρα, αν $x_0 \in I_3$, τότε σύμφωνα με το Θεώρημα 2 της παραγράφου (2.1.2), η ακολουθία $\{x_i\}$ που ορίζεται από την εξίσωση (2.12) συγκλίνει στο ξ.

Καταλήγουμε συνεπώς στο παρακάτω θεώρημα σχετικά με τη σύγκλιση της μεθόδου Newton - Raphson :

Θεώρημα

Έστω ότι οι f, f ′, f ′′ *είναι συνεχείς συναρτήσεις του* x *σε ένα διάστημα* I *που περιέχει τη ρίζα* $x = \xi$ *της* $f(x) = 0$ *και ότι* $f'(x) \neq 0$. *Τότε υπάρχει ένα διάστημα* $I_3 \subset I$ *και τέτοιο που* $\xi \in I_3$, *έτσι που αν* $x_0 \in I_3$, *τότε η ακολουθία που ορίζεται από την (2.12) συγκλίνει στο* ξ.

Αριθμητική εφαρμογή : Να υπολογιστούν οι ρίζες της εξίσωσης $f(x) = 2x^2 - 7x + 3 = 0$ με τη μέθοδο Newton - Raphson :

$$f'(x) = 4x - 7, \qquad g(x) = x - \frac{2x^2 - 7x + 3}{4x - 7}$$

Παρακάτω δίνονται αποτελέσματα για διάφορες αρχικές τιμές x0. Παρατηρούμε ότι η μέθοδος Newton - Raphson συγκλίνει σε διαφορετική ρίζα για τις διάφορες αρχικές τιμές.

ΑΡΧΙΚΗ ΤΙΜΗ x = 0.1 **ΑΡΧΙΚΗ ΤΙΜΗ x = 1**

I=0	x=	.4515151	I=0	x=	.3333333
I=1	x=	.4990948	I=1	x=	.490196
I=2	x=	.4999996	I=2	x=	.4999619
I=3	x=	.5	I=3	x=	.5
			I=4	x=	.5

ΑΡΧΙΚΗ ΤΙΜΗ x = 1.7 **ΑΡΧΙΚΗ ΤΙΜΗ x = 1.8**

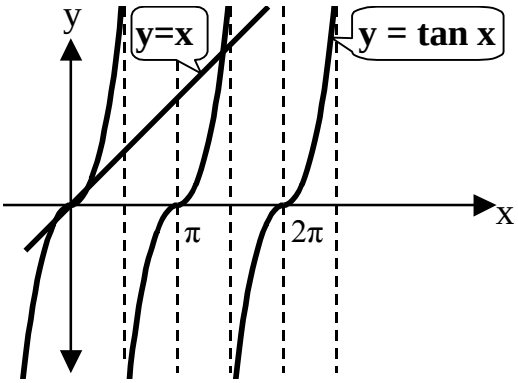
I=0	x=	-13.90002	I=0	x=	17.40001
I=1	x=	-6.124928	I=1	x=	9.624928
I=2	x=	-2.286672	I=2	x=	5.786671
I=3	x=	-.461874	I=3	x=	3.961874
I=4	x=	.2908557	I=4	x=	3.209144
I=5	x=	.4850113	I=5	x=	3.014988
		.4999112	I=6	x=	3.000089
		.5	I=7	x=	3
		.5	I=8	x=	3

ΑΡΧΙΚΗ ΤΙΜΗ x = 3.2 **ΑΡΧΙΚΗ ΤΙΜΗ x = 5**

I=0	x=	3.013793	I=0	x=	3.615385
I=1	x=	3.000075	I=1	x=	3.101507
I=2	x=	3	I=2	x=	3.003812
I=3	x=	3	I=3	x=	3.000006
			I=4	x=	3

Αν η αρχική τιμή δεν εκλεγεί έτσι που να βρίσκεται κοντά σε μια ρίζα, τότε, όχι μόνο είναι δυνατό να παρατηρηθεί σύγκλιση σε άλλη ρίζα, αλλά και η μέθοδος να αποκλίνει.

Για παράδειγμα, ας υπολογιστούν οι ρίζες της x = tan x. Για τη μέθοδο Newton - Raphson,



f(x) = x - tan x, $$f'(x) = 1 - \frac{1}{\cos^2 x} = -\tan^2 x$$

$$g(x) = x - \frac{x - \tan x}{\tan^2 x} \ .$$

Παρακάτω δίνονται αποτελέσματα, από τα οποία φαίνεται, ότι στη μέθοδο Newton - Raphson η αρχική τιμή πρέπει σε ορισμένες περιπτώσεις να βρίσκεται πολύ κοντά στη ρίζα, π.χ. για $x_0 = 4.35$ η μέθοδος συγκλίνει στη ρίζα $\xi = 4.4934095$, ενώ για $x_0 = 4.28$ η μέθοδος αποκλίνει και για $x_0 = 4.72$ συγκλίνει σε άλλη ρίζα.

ΑΡΧΙΚΗ ΤΙΜΗ x = 4.35 **ΑΡΧΙΚΗ ΤΙΜΗ x = 4.28**

I=0	x=	4.596144	I=0	x=	4.73011
I=1	x=	4.542043	I=1	x=	4.749318
I=2	x=	4.504422	I=2	x=	4.792747
I=3	x=	4.493979	I=3	x=	4.904361
I=4	x=	4.493411	I=4	x=	5.284002
I=5	x=	4.49341	I=5	x=	8.113594
			I=6	x=	8.95159
			I=7	x=	45.05458
			I=8	x=	57.86177
			I=9	x=	61.61732
			. . .	. .	. . . .
			I=21	x=	7.523681E+09

2.2.1 ΤΑΧΥΤΗΤΑ ΣΥΓΚΛΙΣΗΣ ΤΗΣ ΜΕΘΟΔΟΥ NEWTON-RAPHSON

$$x_{i+1} - \xi = g(x_i) - g(\xi).$$

Από το θεώρημα του Taylor,

$$g(x_i) = g(\xi) + g'(\xi) \bullet (x_i - \xi) + g''(\eta_i) \bullet \frac{(x_i - \xi)^2}{2}$$

όπου η_i ένα σημείο ανάμεσα στα x_i και ξ.

Για τη μέθοδο Newton - Raphson ισχύει $g'(\xi) = 0$.

Άρα

$$x_i - \xi = g''(\eta_i).\frac{(x_i - \xi)^2}{2} \qquad \Rightarrow \varepsilon_{i+1} = \frac{1}{2} g''(\eta_i).\varepsilon_i^{\,2} \ .$$

Συνεπώς, $$\lim_{i \to \infty} \frac{\varepsilon_{i+1}}{\varepsilon_i} = \frac{g''(\xi)}{2} \ .$$

Η σύγκλιση λοιπόν είναι **τετραγωνική** σε αντίθεση με τη μέθοδο των απλών συναρτησιακών επαναλήψεων, για την οποία ισχύει, όπως ήδη αποδείχτηκε,

$$\lim_{i \to \infty} \frac{\varepsilon_{i+1}}{\varepsilon_i} = \frac{g'(\xi)}{2} \ . \qquad \text{(γραμμική σύγκλιση).}$$

Σύγκριση της ταχύτητας σύγκλισης των δύο μεθόδων γίνεται μέσα από ένα συγκεκριμένο παράδειγμα :

Να βρεθούν οι ρίζες της εξίσωσης $x = \sin x + 0.25$ στο διάστημα $[\pi/4, \pi/2]$.

Για τη μέθοδο των απλών συναρτησιακών επαναλήψεων,

$g(x) = \sin x + 0.25$ και $g'(x) = \cos x$.

Επειδή $|g'(x)| \le \frac{1}{\sqrt{2}} < 1$ στο διάστημα $I = [\pi/4, \pi/2]$, η μέθοδος θα συγκλίνει, αν $x_0 \in I$.

Παρακάτω δίνονται αποτελέσματα για $x_0 = 1.4$ και $x_0 = 1.2$.

ΑΡΧΙΚΗ ΤΙΜΗ $x_0 = 1.4$ **ΑΡΧΙΚΗ ΤΙΜΗ $x_0 = 1.2$**

I=0	x=	1.23545	I=0	x=	1.182039
I=1	x=	1.194296	I=1	x=	1.175381
I=2	x=	1.179957	I=2	x=	1.172837
I=3	x=	1.17459	I=3	x=	1.171854
I=4	x=	1.172532	I=4	x=	1.171472
I=5	x=	1.171735	I=5	x=	1.171324
I=6	x=	1.1714261	I=6	x=	1.171266
I=7	x=	1.171306	I=7	x=	1.171244
I=8	x=	1.171259	I=8	x=	1.171235
I=9	x=	1.171241			
I=10	x=	1.171234			

Αν για το ίδιο πρόβλημα εφαρμοστεί η μέθοδος Newton - Raphson, τότε $f(x) = x - \sin x - 0.25$, $f'(x) = 1 - \cos^2 x$ και

$$g(x) = x - \frac{x - \sin x - 0.25}{1 - \cos x} .$$

Παρακάτω δίνονται τα αποτελέσματα για τις ίδιες αρχικές τιμές $x_0 = 1.4$ και $x_0 = 1.2$ και φαίνεται αμέσως η ταχύτερη σύγκλιση.

ΑΡΧΙΚΗ ΤΙΜΗ $x_0 = 1.4$ **ΑΡΧΙΚΗ ΤΙΜΗ $x_0 = 1.2$**

I=0	x=	1.201755	I=0	x=	1.171832
I=1	x=	1.171907	I=1	x=	1.17123
I=2	x=	1.17123	I=2	x=	1.17123
I=3	x=	1.17123			

2.2.2 ΠΑΡΑΛΛΑΓΕΣ ΤΗΣ ΜΕΘΟΔΟΥ NEWTON - RAPHSON

Η μέθοδος Newton - Raphson υπάγεται σε μια γενικότερη κατηγορία μεθόδων, η οποία μπορεί να περιγραφεί ως εξής :

$$x_{i+1} = x_i - \frac{f(x_i)}{k} \ ,$$

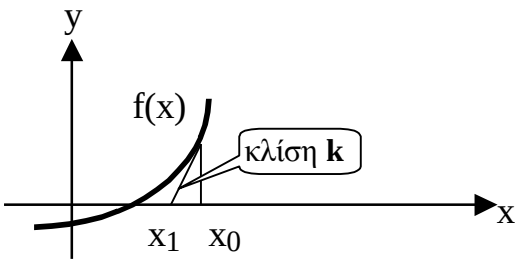
όπου το k μπορεί να είναι σταθερό (δηλαδή ανεξάρτητο από το x_i) ή μεταβλητό.

Από το επόμενο σχήμα, έστω x_0 μια αρχική τιμή : Θεωρούμε την ευθεία.

$$\frac{y - y_0}{x - x_0} = k$$

Για y = 0, x = x_1. Άρα $x_1 = \frac{x_0 - y_0}{k}$ και γενικότερα,

$$x_{i+1} = x_i - \frac{y_i}{k} \ , \qquad \text{η} \qquad x_{i+1} = x_i - \frac{f(x_i)}{k} \ .$$



Ειδικές περιπτώσεις.

(1) $k = \frac{y_1 - y_0}{x_1 - x_0}$: Σταθερή χορδή.

(2) $k = f'(x_0)$: Σταθερή εφαπτομένη.

(3) $k = \frac{y_i - y_{i-1}}{x_i - x_{i-1}}$: Μεταβλητή χορδή.

(4) $k = f'(x_i)$: Μεταβλητή χορδή.

Η περίπτωση (4) ταυτίζεται με τη μέθοδο Newton - Raphson. Η περίπτωση (3) εφαρμόζεται αντί της (4), όταν αποφεύγεται ο αναλυτικός υπολογισμός της παραγώγου f′.

Ταχύτητα σύγκλισης

Για τις παραπάνω μεθόδους μπορεί να γίνει μια ενιαία ανάλυση, αν τεθεί $x_i = \xi + \varepsilon_i$. Τότε προκύπτει ότι,

$$\varepsilon_{i+1} = \varepsilon_i - \frac{f(\xi + \varepsilon_i)}{k} \quad .$$

Αλλά $f(\xi + \varepsilon_i) = f(\xi) + \varepsilon_i \,.\, f'(\xi) + \ldots \approx \varepsilon_i.f(\xi)$. Άρα,

$$\varepsilon_{i+1} = \varepsilon_i.\frac{1 - f'(\xi)}{k}$$

Κατά συνέπεια, το ε_{i+1} θα είναι σημαντικά μικρότερο από το ε_i, αν το k δεν διαφέρει πολύ από το f′ (ξ).

Για την περίπτωση (3),

$$x_{i+1} = x_i - y_i \,/\, \frac{y_i - y_{i-1}}{x_i - x_{i-1}} \qquad \text{ή} \qquad x_{i+1} = \frac{x_{i-1}.y_i - x_i.y_{i-1}}{y_i - y_{i-1}}$$

Αντικαθιστώντας και πάλι $x_{i-1} = \xi + \varepsilon_{i-1}$, $x_i = \xi + \varepsilon_i$, $x_{i+1} = \xi + \varepsilon_{i+1}$, προκύπτει τελικά ότι

$$\varepsilon_{i+1} = \frac{[\varepsilon_{i-1} \cdot f(\xi + \varepsilon_i) - \varepsilon_i \cdot f(\xi + \varepsilon_{i-1})]}{[f(\xi + \varepsilon_i) - f(\xi + \varepsilon_{i-1})]} \tag{2.13}$$

Αλλά,

$f(\xi + \varepsilon_i) \quad = f(\xi) + \varepsilon_i \,.\, f'(\xi) + \dfrac{\varepsilon_i^2}{2}.f''(\xi) \quad + \ldots$ και

$f(\xi + \varepsilon_{i-1}) \quad = f(\xi) + \varepsilon_{i-1}.\, f'(\xi) + \frac{\varepsilon_{i-1}^2}{2}.f''(\xi) \quad + \ldots$

Συνεπώς,

$\varepsilon_{i-1} \cdot f(\xi + \varepsilon_i) - \varepsilon_i \cdot f(\xi + \varepsilon_{i-1}) = \quad \frac{1}{2}.\varepsilon_i.\varepsilon_{i-1} \cdot (\varepsilon_i - \varepsilon_{i-1}) \cdot f''(\xi) + \ldots$ και

$f(\xi + \varepsilon_i) - f(\xi + \varepsilon_{i-1}) = (\varepsilon_i - \varepsilon_{i-1}) \cdot f'(\xi) + \ldots$

Άρα, $$\varepsilon_{i+1} = \varepsilon_{i-1} \cdot \varepsilon_i \cdot \frac{f''(\xi)}{[2f'(\xi)]} \tag{2.14}$$

$$\varepsilon_{i+1} \approx K \cdot \varepsilon_{i-1} \cdot \varepsilon_i \quad . \tag{2.15}$$

Ζητείται τώρα ένας σταθερός εκθέτης m, τέτοιος που το ε_{i+1} να είναι της μορφής :

$$\varepsilon_{i+1} = A \cdot \varepsilon_i^m. \tag{2.16}$$

Από τις (2.15) και (2.16) προκύπτει ότι $\varepsilon_{i+1} = K \cdot A \cdot \varepsilon_{i-1}^{m+1}$.

Από την (2.16), $\varepsilon_{i+1} = A \cdot (A \cdot \varepsilon_{i-1}^m)^m = A \,.\, A^m \cdot \varepsilon_{i-1}^{m^2}$.

Οι δύο διαφορετικές παραστάσεις για το ε_{i+1} είναι εκ ταυτότητος ίσες και συνεπώς,

$K = A^m$ και $m^2 = m + 1$. Άρα, $\quad m = \frac{1+\sqrt{5}}{2} \approx 1.618$.

Άρα, $$\varepsilon_{i+1} = A \,.\, \varepsilon_i^{1.618} \,. \tag{2.17}$$

Υπενθυμίζεται εδώ, ότι στη μέθοδο Newton - Raphson, ο εκθέτης m ήταν ίσος με 2.

Απόδειξη της σύγκλισης

Η αυστηρότερη όμως μαθηματική απόδειξη απόδειξη για τη σύγκλιση της μεθόδου της χορδής έχει ως εξής:

Το δεξιό μέλος της (2.13) γράφεται ως συνάρτηση των ε_i και ε_{i-1}:

$$F(\varepsilon_i, \varepsilon_{i-1}) = \frac{\varepsilon_{i-1} \cdot f(\xi + \varepsilon_i) - \varepsilon_i \cdot f(\xi + \varepsilon_{i-1})}{f(\xi + \varepsilon_i) - f(\xi + \varepsilon_{i-1})}$$

Προκύπτει όμως από την (2.14), ότι το ανάπτυγμα κατά Taylor (με υπόλοιπο) της F ως προς ε_i και ε_{i-1} και με αφετηρία τις τιμές $\varepsilon_i = 0$ και $\varepsilon_{i-1} = 0$, θα είναι της μορφής:

$$F(\varepsilon_i, \varepsilon_{i-1}) = \frac{1}{2}\varepsilon_i \cdot \varepsilon_{i-1} \frac{\partial^2 F}{\partial \varepsilon_i \partial \varepsilon_{i-1}}(\bar{\varepsilon}_i, \bar{\varepsilon}_{i-1}) \tag{2.18}$$

αν το υπόλοιπο συνίσταται από όρους δεύτερης τάξης.

Στην (2.18) οι τιμές $\bar{\varepsilon}_i$ και $\bar{\varepsilon}_{i-1}$ βρίσκονται ανάμεσα στις τιμές 0 και ε_i και 0 και ε_{i-1} αντίστοιχα.

Έστω τώρα ότι τα x_i και x_{i-1} βρίσκονται μέσα σ' ένα διάστημα I που περιέχει την ρίζα ξ: I = [ξ-ρ, ξ+ρ].

Έστω επίσης ότι μέσα στο διάστημα αυτό,

$$\left| \frac{\partial^2 F}{\partial \varepsilon_i \partial \varepsilon_{i-1}} \right| \le M \tag{2.19}$$

πράγμα που δικαιολογείται, αν υποτεθεί ότι η παραπάνω δεύτερη μικτή παράγωγος είναι συνεχής, επειδή στην περίπτωση αυτή θα είναι και φραγμένη.

Το εύρος ρ επιλέγεται τώρα έτσι που

$$\frac{1}{2}\rho^2 M < \rho \qquad \text{ή} \qquad \frac{1}{2}\rho . M < 1 .$$

Τότε προκύπτει ότι

$$|\varepsilon_{i+1}| = \frac{1}{2}|\varepsilon_i||\varepsilon_{i-1}|\left|\frac{\partial^2 F}{\partial\varepsilon_i \partial\varepsilon_{i-1}}\right| \le \frac{1}{2}\rho^2 M < \rho \ .$$

Επομένως, αν τα x_{i-1}, x_i βρίσκονται μέσα στο I ($\varepsilon_i < \rho$, $\varepsilon_{i-1} < \rho \Rightarrow$ $x_{i-1}, x_i \in I$), τότε το ίδιο θα ισχύει και για το x_{i+1}: $\varepsilon_{i+1} < \rho \Rightarrow x_{i+1} \in I$. Αυτό επαγωγικά σημαίνει ότι αν τα x_0 και $x_1 \in I$, τότε και $x_i \in I \ \forall i$.
Επίσης,

$$|\varepsilon_{i+1}| \le |\varepsilon_i|\left[\frac{1}{2}|\varepsilon_{i-1}|M\right] \le \frac{1}{2}\rho.M|\varepsilon_i| < |\varepsilon_i| \ .$$

Συνεπώς, $|\varepsilon_{i+1}| < |\varepsilon_i|$.

Η ανισότητα αυτή, σε συνδυασμό με το γεγονός, ότι το εύρος ρ μπορεί να αναπροσαρμοστεί και να λάβει τιμές οσοδήποτε μικρές, αποδεικνύει τη σύγκλιση της μεθόδου της χορδής, υπό τον όρο οι δύο αρχικές τιμές να βρίσκονται αρκετά κοντά στη ρίζα ξ.

Η σύγκλιση της μεθόδου επίσης σημαίνει ότι η ακολουθία των σφαλμάτων μπορεί να παρασταθεί ως εξής με βάση τις (2.13) και (2.18):

$$|\varepsilon_{i+1}| = C_i|\varepsilon_i||\varepsilon_{i-1}| \quad \text{, όπου } C_i \to C. \tag{2.20}$$

Εξετάζουμε τώρα την ταχύτητα σύγκλισης.
Θεωρούμε την ακολουθία

$$u_i = \frac{|\varepsilon_{i+1}|}{|\varepsilon_i|^m} \tag{2.21}$$

όπου i = 0,1,2,…

Θα αποδείξουμε ότι $u_i \to c$ για κάποιο $m > 1$.

Από την (2.20),

$$\frac{|\varepsilon_{i+1}|}{|\varepsilon_i|^m} = C_i \left(\frac{|\varepsilon_i|}{|\varepsilon_{i-1}|^m} \right)^{\lambda} \tag{2.22}$$

όπου $\lambda=1- m$ και $\lambda=-1/m$.

Έπεται ότι $m^2 - m - 1 = 0$

Συνεπώς, $m = (1+\sqrt{5})/2$, δεδομένου ότι επιλέγουμε τη ρίζα για την οποία $m > 1$.

Από την (2.21) και την (2.22)

$u_i = C_i .u_{i-1}^{\lambda}$, όπου $C_i \to C$ και $i =1,2,\ldots$

Πρέπει να αποδείξουμε ότι η ακολουθία $\{ u_i \}$ τείνει σε μία σταθερά $c \neq 0$. Παίρνουμε τους λογαρίθμους και των δύο μελών:

$$\ln u_i = \ln C_i + \lambda \ln u_{i-1}.$$

Θέτουμε $v_i = \ln u_i$ και $b_i = \ln C_i$, οπότε

$$v_i = b_i + \lambda .v_{i-1} , \; i = 1,2,\ldots \tag{2.23}$$

Θεωρούμε την ακολουθία των διαφορών

$$\omega_i = v_i - v_{i-1} \text{ με } i=1,2,\ldots. \tag{2.24}$$

Τότε,

$\omega_i = \alpha_i + \lambda\, \omega_{i-1}$, όπου $\alpha_i = b_i - b_{i-1}$ και $i = 1,2,\ldots$.

Η $\{\alpha_i\}$, προφανώς, είναι μηδενική ακολουθία.

Προκύπτει εύκολα ότι

$$\omega_{i+1} = \alpha_i + \lambda a_i + \lambda^2 \alpha_{i-1} + \ldots + \lambda^{i-1}\alpha_2 + \lambda^i \omega_1$$

Θα αποδείξουμε ότι $\omega_i \to 0$.

Για $\varepsilon > 0$ υπάρχει N αρκετά μεγάλο έτσι που για $i \geq N$, $|\alpha_i| < \varepsilon$ και $|\lambda|^i < \varepsilon$.

Αυτό ισχύει, επειδή η $\{\alpha_i\}$ είναι μηδενική ακολουθία και επειδή $|\lambda| < 1$.

Επομένως,

$$\begin{aligned} |\omega_{i+1}| \leq\ & \varepsilon(1 + |\lambda| + |\lambda|^2 + \ldots + |\lambda|^N) + \\ & + |\lambda|^{N+1}(|\alpha_{i-N}| + |\lambda||\alpha_{i-(N+1)}| + \ldots + |\lambda|^{i-(N+2)}|\alpha_2|) + |\lambda|^i |\omega_1| \end{aligned}$$

Από την παραπάνω ανισότητα προκύπτει τελικά ότι

$$|\omega_{i+1}| \leq \varepsilon \frac{1}{1-|\lambda|} + \varepsilon \frac{1}{1-|\lambda|} M + \varepsilon |\omega_1| = \text{σταθερά} \times \varepsilon$$

Στην παραπάνω ανισότητα έχει ληφθεί υπόψη το γεγονός ότι η ακολουθία $\{\alpha_i\}$ είναι φραγμένη, ως μηδενική, και M είναι φράγμα της ακολουθίας αυτής.

Συνεπώς, η $|\omega_{i+1}|$ μπορεί να γίνει οσοδήποτε μικρή, ή $\omega_i \to 0$.

Από την (2.24),

$v_i - v_{i-1} = \omega_i \to 0.$

Από την (2.23),

$v_i - v_{i-1} = b_i + (\lambda - 1)\, v_{i-1}$

Από τις δύο τελευταίες σχέσεις προκύπτει ότι

$(1 - \lambda)\, v_i \to b,$

όπου

$$b = \lim b = \lim \ln C = \ln C$$

Επομένως,

$(1-\lambda).\ln u_i \to \ln C \;\Rightarrow u_i^{1-\lambda} \to C \Rightarrow u_i \to C^{1/(1-\lambda)}$

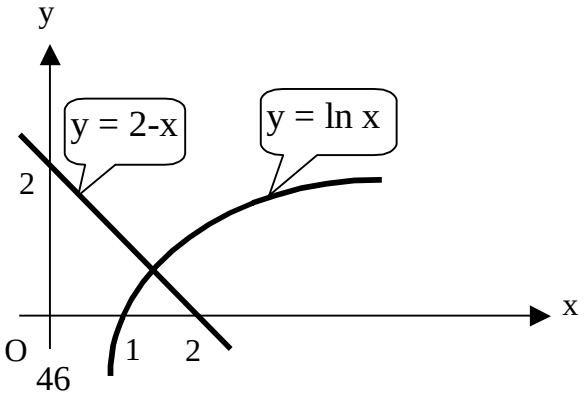
Τελικά, $u_i \to C^{1/m}$.

Άσκηση

Να υπολογιστεί η ρίζα της εξίσωσης ln x + x - 2 = 0 με ακρίβεια τριών δεκαδικών ψηφίων και τη μέθοδο Newton - Raphson.

Λύση

Η εξίσωση μπορεί να γραφτεί ln x = 2 - x και μια πρόχειρη γραφική παράσταση οδηγεί στο συμπέρασμα ότι σαν αρχική τιμή μπορούμε να πάρουμε $x_0 = 1.5$.



$f(x) = \ln x + x - 2$

Άρα, $f'(x) = \dfrac{1}{x} + 1$ και

$$x_{i+1} = x_i - \frac{\ln x_i + x_i - 2}{\frac{1}{x_i} + 1}$$

Για $x_0 = 1.5 \rightarrow x_1 = 1.556$

$x_1 = 1.556 \rightarrow x_2 = 1.557$

. . . . . κ.λ.π.

2.3 ΕΠΙΤΑΧΥΝΣΗ ΣΥΝΑΡΤΗΣΙΑΚΩΝ ΕΠΑΝΑΛΗΨΕΩΝ ΚΑΤΑ ΑΙΤΚΕΝ

Για τις απλές συναρτησιακές επαναλήψεις $x_{i+1} = g(x_i)$ έχει αποδειχτεί ότι, ασυμπτωτικά ισχύει $\varepsilon_{i+1} = K \cdot \varepsilon_i$, όπου $\varepsilon_i = x_i - \xi$. Αυτό σημαίνει ότι $\varepsilon_i = K^i \cdot \varepsilon_0$. Διακρίνουμε λοιπόν τρεις αγνώστους, ξ, K, ε_0. Προσδιορίζουμε τους τρεις αγνώστους από τρεις διαδοχικές προσεγγίσεις:

$x_{i+2} - \xi = K^{i+2} \cdot \varepsilon_0$, $x_{i+1} - \xi = K^{i+1} \cdot \varepsilon_0$, $x_i - \xi = K^i \cdot \varepsilon_0$.

Προκύπτει ότι $$\frac{x_{i+1} - \xi}{x_i - \xi} = \frac{x_{i+2} - \xi}{x_{i+1} - \xi} = K \ .$$

Άρα $$\xi = \frac{x_i . x_{i+2} - x_{i+1}^2}{x_{i+2} - 2x_{i+1} + x_i}$$

Η παραπάνω σχέση θα ίσχυε με ακρίβεια, αν και αυτή η σχέση $\varepsilon_{i+1} = K \cdot \varepsilon_i$ ήταν ακριβής. Παρόλα αυτά, η σχέση αυτή υποδεικνύει μια άλλη προσεγγιστική διαδικασία :

$$x^*_i = \frac{x_i . x_{i+2} - x_{i+1}^2}{x_{i+2} - 2x_{i+1} + x_i} \qquad (2.25)$$

Αυτό μπορεί να γραφτεί ως εξής :

$$x^*_i = G(x_i, x_{i+1}, x_{i+2}).$$

Με τον τρόπο αυτό προκύπτει ο εξής αλγόριθμος :

Επιλέγεται μια αρχική τιμή x_0 και υπολογίζονται τα x_1 και x_2 από τις σχέσεις $x_1 = g(x_0)$ και $x_2 = g(x_1)$. Στη συνέχεια το x_0 βελτιώνεται από τη σχέση $x^*_0 = G(x_0, x_1, x_2)$.

Στο επόμενο βήμα, $x_1 = g(x^*_0)$, $x_2 = g(x_1)$ και $x^*_0 = G(x_0, x_1, x_2)$, κ.ο.κ. Παρακάτω λύνεται η εξίσωση $x = \sin x + 0.25$ με αρχική τιμή $x_0 = 1.4$. Το ίδιο πρόβλημα λύθηκε ήδη στην προηγούμενη παράγραφο με δύο άλλες μεθόδους.

$x = \sin x + 0.25 = g(x)$

$x_0 = 1.4 \quad \rightarrow \quad x_1 = g(x_0) = 1.2354497$

$x_2 = g(x_1) = 1.194296$

$$x^*_0 = \frac{x_0.x_2 - x_1^2}{x_2 - 2x_1 + x_0} = \frac{0.1456784}{0.1233966} = 1.180571$$

$x_1 = g(x^*_0) = 1.174824$

$x_2 = g(x_1) = 1.172622 \quad \ldots\ldots\ldots$

Στο κεφάλαιο 9 δίνεται πρόγραμμα που υπολογίζει τις ρίζες της προηγούμενης εξίσωσης, με αρχική τιμή $x_0 = 1.4$.

2.3.1 ΑΝΑΛΥΣΗ ΤΟΥ ΣΦΑΛΜΑΤΟΣ ΣΤΗΝ ΕΠΙΤΑΧΥΝΣΗ ΚΑΤΑ ΑΙΤΚΕΝ.

Η ανάλυση του σφάλματος για την επιτάχυνση των συναρτησιακών επαναλήψεων στηρίζεται στην αντίστοιχη βασική ανάλυση του

σφάλματος των επαναλήψεων αυτών (κεφ. 2.1.2). Υπενθυμίζουμε ότι για το σχήμα $x_{i+1} = g(x_i)$ έχει αποδειχτεί ότι

$$\varepsilon_{i+1} = g'(\xi_i).\varepsilon_i \qquad (2.26),$$

όπου ξ_i είναι μια τιμή ανάμεσα στα x_i και ξ.

Η (2.26) προέκυψε από ανάπτυγμα Taylor που περιέλαβε μόνο τον όρο πρώτης τάξης. Αν περιληφθεί και όρος δεύτερης τάξης, τότε

$$x_{i+1} = g(x_i) = g(\xi) + (x_i - \xi)g'(\xi) + \frac{1}{2}(x_i - \xi)^2 g''(\xi_i)$$

όπου ξ_i είναι και πάλι μια τιμή ανάμεσα στα x_i και ξ και ξ είναι ρίζα:

$$g(\xi)=\xi.$$

Προκύπτει συνεπώς ότι

$$\varepsilon_{i+1} = [c + g''(\xi_i)\varepsilon_i].\varepsilon_i \qquad (2.27)$$

Άρα, η (2.26) μπορεί να γραφεί με τη μορφή

$$\varepsilon_{i+1} = (c + \sigma_i)\varepsilon_i \qquad (2.28)$$

όπου $c=g'(\xi)$ και $\sigma_i=g''(\xi_i)\varepsilon_i$. Προφανώς, $\sigma_i \to 0$ καθώς το $i\to\infty$.

Έστω τώρα

$\varepsilon_i^* = x_i^* - \xi$ το σφάλμα το οποίο υπεισέρχεται κατά τη διαδικασία επιτάχυνσης. Τότε, από την (2.25) προκύπτει εύκολα η σχέση του σφάλματος αυτού με τα ε_i, ε_{i+1} και ε_{i+2}:

$$\varepsilon_i^* = \frac{\varepsilon_i.\varepsilon_{i+2} - \varepsilon_{i+1}^2}{\varepsilon_{i+2} - 2\varepsilon_{i+1} + \varepsilon_i} \qquad (2.29)$$

Επίσης, από τις (2.28) και (2.29) αποδεικνύεται ότι

$$\frac{\varepsilon_i^*}{\varepsilon_i} = (c+\sigma_i)\frac{\sigma_{i+1}^2 - \sigma_i^2}{(c-1)^2 - (c-2)\sigma_i + c\sigma_{i+1} + \sigma_i\sigma_{i+1}} \qquad (2.30)$$

Η παραπάνω παράσταση τείνει προς το 0 καθώς τα σ_i και σ_{i+1} τείνουν προς το 0, δεδομένου ότι και το $c < 1$. Υπενθυμίζεται εδώ ότι $c=g'(\xi)<1$ για να έχουμε σύγκλιση στις απλές επαναλήψεις.

Συνεπώς,

$$\frac{\varepsilon_i^*}{\varepsilon_i} \to 0 \text{ καθώς το } i\to\infty,$$

πράγμα που σημαίνει ότι το σφάλμα της επιταχυνόμενης διαδικασίας πράγματι τείνει προς το 0 «ταχύτερα» από ότι το σφάλμα των απλών επαναλήψεων.

Η συμπεριφορά του σ_i καθώς το $i\to\infty$, μπορεί να γίνει πιο συγκεκριμένη, αν ληφθεί υπόψη ότι $\sigma_i=g''(\xi_i)\varepsilon_i$. Τότε, το σ_i γράφεται με τη μορφή

$$\sigma_i=\omega_i.\varepsilon_i, \qquad (2.31)$$

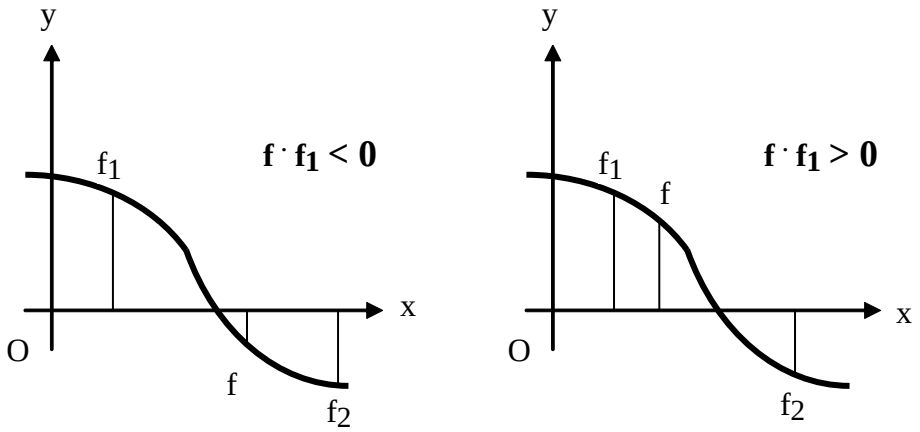
όπου το ω_i τείνει προς ένα όριο $\omega = g''(\xi)$.

Από τις (2.30) και (2.31) προκύπτει τελικά ότι

$$\frac{\varepsilon_i^*}{\varepsilon_i^2} \to c\omega/(c-1) \qquad (2.32)$$

2.4 ΜΕΘΟΔΟΣ ΤΗΣ ΔΙΧΟΤΟΜΗΣΗΣ

Έστω η εξίσωση f(x) = 0. Ορίζεται ένα βήμα Δx και υπολογίζονται οι τιμές της f για x_0, $x_0 + \Delta x$, $x_0 + 2\Delta x$, κ.λ.π. Έστω f_1 , f_2 δύο τέτοιες διαδοχικές τιμές της f : $f_1 = f(x_1)$, $f_2 = f(x_1+\Delta x)$. Αν $f_1 \cdot f_2 > 0$ τότε η διαδικασία συνεχίζεται Αν $f_1 \cdot f_2 < 0$, αυτό μαρτυρεί την ύπαρξη μιας τουλάχιστον ρίζας στο διάστημα $(x_1, x_1 + \Delta x)$. Τότε το διάστημα αυτό διχοτομείται, δηλαδή λαμβάνεται το σημείο $x_1 + \Delta x/2$ και υπολογίζεται η τιμή της f στο σημείο αυτό. Αν η |f| είναι μικρότερη από μια δεδομένη μικρή τιμή ε, τότε η διαδικασία τερματίζεται. Αλλιώς, εξετάζεται το γινόμενο $f \cdot f_1$. Αν $f \cdot f_1 < 0$, τότε $f_2 = f$ και η θέση 2 μετατίθεται στο f. Επακολουθεί νέα διχοτόμηση κ.ο.κ.



Αν υποτεθεί ότι στο διάστημα I = [a, b] υπάρχει μία μόνο ρίζα ξ της f(x), τότε η μέθοδος της διχοτόμησης συγκλίνει.

Έστω $I_i = [a_i, b_i]$ το διάστημα μέσα στο οποίο εντοπίζεται η ρίζα ξ κατά το βήμα i. Αν l_i είναι το μήκος του διαστήματος αυτού, τότε το διάστημα του επόμενου βήματος θα έχει μήκος $l_{i+1} = l_i / 2$. Θα ισχύει επομένως

$l_i = l / 2^i$, όπου l το μήκος του αρχικού διαστήματος I.

Θεωρούμε ότι η ρίζα ξ προσεγγίζεται με ακρίβεια ε, όταν $l_i \le \varepsilon$, ή όταν

$$2^i \geq l / \varepsilon$$

Αυτό σημαίνει ότι ο αριθμός των διαδοχικών διχοτομήσεων πρέπει να είναι

$$i \geq \log(l/\varepsilon) / \log 2$$

Για παράδειγμα, αν $l = 1$ και $\varepsilon = 10^{-5}$, $i \geq 5 / \log 2 = 16.6$. Συνεπώς, $i \geq 17$.

Η μέθοδος της διχοτόμησης χρησιμεύει κυρίως για τον εντοπισμό ριζών μη γραμμικών εξισώσεων. Ο ακριβέστερος υπολογισμός γίνεται με μια από τις άλλες μεθόδους.

Στο κεφάλαιο 9 δίνεται πρόγραμμα που εντοπίζει και υπολογίζει τις ρίζες της $f(x) = 2x^2 - 7x + 3 = 0$ ($\rho_1 = 0.5$ και $\rho_2 = 3.0$).

3. ΥΠΟΛΟΓΙΣΜΟΣ ΡΙΖΩΝ ΠΟΛΥΩΝΥΜΩΝ

3.1 ΥΠΟΛΟΓΙΣΜΟΣ ΤΗΣ ΤΙΜΗΣ ΕΝΟΣ ΠΟΛΥΩΝΥΜΟΥ ΚΑΙ ΤΗΣ ΠΑΡΑΓΩΓΟΥ ΤΟΥ

Έστω πολυώνυμο n βαθμού :

$$p_n(x) = a_0 + a_1.x + a_2.x^2 + \ldots + a_n.x^n.$$

Για μεγάλες όπως και για μικρές τιμές του x, η ύψωση σε δυνάμεις εισάγει αριθμητικά σφάλματα. Επιπλέον, πρέπει να αναζητηθεί μέθοδος με μικρότερο αριθμό αριθμητικών πράξεων. Αυτό μπορεί να γίνει αν το παραπάνω πολυώνυμο μετασχηματιστεί ως εξής :

$$p_n(x) = a_0 + a_1.x + a_2.x^2 + \ldots + a_n.x^n =$$

$= a_0 + x.(a_1 + a_2.x + a_3.x^2 + \ldots + a_n.x^{n-1}) =$

$= a_0 + x.(a_1 + x.(a_2 + a_3.x + a_4.x^2 + \ldots + a_n.x^{n-2})) =$

$= a_0 + x.(a_1 + x.(a_2 + x.(a_3 + a_4.x + \ldots + a_n.x^{n-3}))) =$

$= . \quad . \quad .$

$= a_0 + x.(a_1 + x.(a_2 + x.(a_3 + \ldots + x.(a_{n-1} + x.a_n)) \ldots)$

Η παραπάνω παράσταση οδηγεί σε μια μέθοδο "**διαδοχικών παρενθέσεων**". Για να υπολογιστεί η τιμή του πολυωνύμου για x = z υπολογίζονται διαδοχικά οι παραστάσεις μέσα στις παρενθέσεις :

b_n	=	a_n		
b_{n-1}	=	a_{n-1}	+	$z.b_n$
.	.		.	
b_{n-i}	=	a_{n-i}	+	$z.b_{n-i+1}$
.	.		.	
b_0	=	a_0	+	$z.b_1$

. (3.1)

Η τελική ποσότητα b_0 είναι η τιμή του πολυωνύμου στη θέση x = z :

$$b_0 = p_n(z). \tag{3.2}$$

Εξετάζεται τώρα η διαίρεση του πολυωνύμου $p_n(x)$ με το **(x - z)** :

$$p_n(x) = q_{n-1}(x).(x - z) + R_0 \tag{3.3}$$

Το $q_{n-1}(x)$, πολυώνυμο n-1 βαθμού, είναι το πηλίκο της διαίρεσης και το R_0 το υπόλοιπο. Παρατηρούμε ότι $R_0 = p_n(z) = b_0$.

Το $q_{n-1}(x)$ γράφεται ως εξής :

$$q_{n-1}(x) = b_1 + b_2.x + b_3.x^2 + \ldots + b_n.x^{n-1}.$$

Θα αποδειχτεί ότι οι συντελεστές του $q_{n-1}(x)$ είναι οι ίδιοι με τις τιμές b_i που προκύπτουν από τις εξισώσεις (3.1). Πράγματι, από την (3.3) έπεται ότι

$a_0 + a_1.x + a_2.x^2 + \ldots + a_n.x^n =$

$= (b_1 + b_2.x + b_3.x^2 + \ldots + b_n.x^{n-1}).(x - z) + R_0 =$

$= (R_0 - b_1.z) + (b_1 - b_2.z).x + (b_2 - b_3.z).x^2 + \ldots + (b_{n-1} - b_n.z).x^{n-1} + b_n.x^n.$

Εξισώνοντας τους συντελεστές των ομοιόβαθμων όρων, προκύπτει :

b_0	=	a_0	+	$z.b_1$
b_1	=	a_1	+	$z.b_2$
.	.	.	.	.
b_{n-1}	=	a_{n-1}	+	$z.b_n$
b_n	=	a_n		

Οι τελευταίες σχέσεις είναι ίδιες με τις (3.1). Ο υπολογισμός της τιμής του q_{n-1} για $x = z$ γίνεται και πάλι με τη μέθοδο των διαδοχικών παρενθέσεων σύμφωνα με τον αλγόριθμο των εξισώσεων (3.1) :

c_n	=	b_n		
c_{n-1}	=	b_{n-1}	+	$z.c_n$
.	.	.	.	.
c_1	=	b_1	+	$z.c_2$

Όπως και παραπάνω,

$$c_1 = q_{n-1}(z). \tag{3.4}$$

Αν τώρα παραγωγίσουμε την (3.3), προκύπτει :

$$p'_n(x) = q'_{n-1}(x).(x - z) + q_{n-1}(x)$$

και $p'_n(z) = q_{n-1}(z)$. Σε συνδυασμό με την (3.4)

$$c_1 = p'_n(z).$$

Στο κεφάλαιο 9 παρουσιάζεται πρόγραμμα που δίνει την τιμή ενός πολυωνύμου και της παραγώγου του για μια συγκεκριμένη τιμή της ανεξάρτητης μεταβλητής. Δεδομένα στο πρόγραμμα είναι οι συντελεστές των δυνάμεων του πολυωνύμου και η τιμή z της ανεξάρτητης μεταβλητής.

Για παράδειγμα υπολογίζεται η τιμή του πολυωνύμου

$p(x) = -5 + 2x - x^2 + 2x^3 + 3x^4$ καθώς και η τιμή της παραγώγου $p'(x)$ για $x = 2$. Τα αποτελέσματα είναι $p(2) = 59$ και $p'(2) = 118$.

Η ανάλυση που παρουσιάστηκε μπορεί να γενικευτεί ως εξής με αφετηρία την εξίσωση (3.3) :

$p_n(x)$	=	$(x - z).q_{n-1}(x)$	+	R_0	R_0	=	$p_n(z)$
$q_{n-1}(x)$	=	$(x - z).q_{n-2}(x)$	+	R_1	R_1	=	$q_{n-1}(z)$
$q_{n-2}(x)$	=	$(x - z).q_{n-3}(x)$	+	R_2	R_2	=	$q_{n-2}(z)$
. . .	. . . . .	. .	. .	. . . . .	. . .		

Τελικά, το πολυώνυμο $p_n(x)$ γράφεται με τη μορφή :

$p_n(x)$	$= (x-z).q_{n-1}(x)$	$+ R_0$	$=$			
	$= (x-z)^2.q_{n-2}(x)$	$+ R_1.(x-z)$	$+ R_0$	$=$		
	$= (x-z)^3.q_{n-3}(x)$	$+ R_2.(x-z)^2$	$+ R_1.(x-z)$	$+ R_0$	$=$	
	$= . . .$					
	$= (x-z)^n.R_n$	$+ R_{n-1}.(x-z)^{n-1}$	$+ . . .$	$+ R_1.(x-z)$	$+ R_0$	.

Έπειτα από i παραγωγίσεις,

$$p_n^{(i)} = i!\ R_i\ . \tag{3.5}$$

Διαμορφώνεται επομένως ένας ενιαίος αλγόριθμος, που επιτρέπει τον υπολογισμό της τιμής ενός πολυωνύμου και των παραγώγων του. Ο αλγόριθμος αυτός συνίσταται στον διαδοχικό υπολογισμό των συντελεστών b_i, c_i κ.λπ. σύμφωνα με τη διαδικασία (3.1). Θα είναι $R_0 = b_0$, $R_1 = b_1$, κ.λπ. Οι υπολογισμοί διατάσσονται όπως παρακάτω :

a_0	$b_0 = R_0$				
a_1	b_1	$c_1 = R_1$			
a_2	b_2	c_2	.		
a_3	b_3	c_3	.		
.	.	.	.		
a_{n-1}	b_{n-1}	c_{n-1}	.	R_{n-1}	
a_n	b_n	c_n	.	.	R_n

Παράδειγμα

Να βρεθούν οι τιμές των παραγώγων του πολυωνύμου

$p_4(x) = 3x^4 + 2x^3 - x^2 + 2x - 5$ για $x = 2$.

-5	59	$=R_0$					
2	32		118	$=R_1$			
-1	15	43		83	$=R_2$		
2	8	14	20		26	$=R_3$	
3	3	3	3	3		3	$=R_4$

Σύμφωνα με την (3.5), οι τιμές των παραγώγων είναι :

$p'_4(2) = 118$

$p''_4(2) = 2 * 83 = 166$

$p'''_4(2) = 6 * 26 = 156$ και

$p^{(4)}_4(2) = 24 * 3 = 72$.

Με τον παραπάνω αλγόριθμο μπορεί να μετατραπεί ένας θετικός ακέραιος N από το δεκαδικό σύστημα στο δυαδικό.

Έστω $(\alpha_n\ \alpha_{n-1} \ldots \alpha_0)_2$ η δυαδική παράσταση του N.

Τότε $N = a_0.2^0 + a_1.2^1 + \ldots + a_{n-1}.2^{n-1} + a_n.2^n = p_n(2) = b_0$.

Άρα για $z = 2$

$$\boxed{\begin{array}{ll} b_0 & = N \\ b_1 & = (b_0 - a_0)/2 \\ b_2 & = (b_1 - a_1)/2 \\ . & . \ . \\ b_k & = (b_{k-1} - a_{k-1})/2 \end{array}}$$

όπου $a_k = 1$ αν b_k περιττός

όπου $a_k = 0$ αν b_k άρτιος

.

Εφαρμογή

N = 187

k	b_k	a_k
0	187	1
1	93	1
2	46	0
3	23	1
4	11	1
5	5	1
6	2	0
7	1	1

Άρα $187_{10} = (10111011)_2$.

3.2 Η ΜΕΘΟΔΟΣ NEWTON-RAPHSON ΓΙΑ ΤΗΝ ΕΥΡΕΣΗ ΠΡΑΓΜΑΤΙΚΩΝ ΡΙΖΩΝ ΠΟΛΥΩΝΥΜΩΝ

Το σχήμα Newton-Raphson απαιτεί σε κάθε βήμα τον υπολογισμό της τιμής του πολυωνύμου και της παραγώγου του. Αλλά, σύμφωνα με τα παραπάνω η τιμή του πολυωνύμου δίνεται από την ποσότητα b_0 και η τιμή της παραγώγου από την c_1 :

$$x_{i+1} = x_i - \frac{p(x_i)}{p'(x_i)} = x_i - \frac{b_0(x_i)}{c_1(x_i)}$$

Αν εφαρμοστεί η μέθοδος αυτή για τον υπολογισμό της ρίζας $\xi = 1.324718$ του πολυωνύμου $p(x) = x^3 - x - 1$ με αρχική τιμή $x_0 = 1.1$, τότε θα προκύψουν οι εξής διαδοχικές προσεγγίσεις :

$x_1 = 1.3923954$,

$x_2 = 1.3286261$,

$x_3 = 1.3247321$,

$x_4 = 1.324718$,

$x_4 = 1.324718$.

3.3 Η ΜΕΘΟΔΟΣ BERNOULLI

Έστω πολυώνυμο n βαθμού

$$p_n(x) = a_0 + a_1.x + a_2.x^2 + \ldots + a_n.x^n$$

και έστω ότι έχει n διακεκριμένες πραγματικές ρίζες, $\xi_1, \xi_2, \ldots, \xi_n$.

Η ρητή παράσταση $R(x) = \dfrac{1}{p_n(x)}$ αναλύεται τότε σε άθροισμα κλασμάτων

$$R(x) = \frac{\alpha_1}{x-\xi_1} + \frac{\alpha_2}{x-\xi_2} + \ldots + \frac{\alpha_n}{x-\xi_n} .$$

Το κάθε ένα από τα κλάσματα γράφεται με τη μορφή σειράς :

$$\frac{\alpha_i}{x-\xi_i} = -\frac{\alpha_i}{\xi_i}\frac{1}{1-\dfrac{x}{\xi_i}} = -\alpha_i\left(\frac{1}{\xi_i} + \frac{x}{\xi_i^2} + \frac{x^2}{\xi_i^3} + \ldots\right).$$

Η παραπάνω σειρά είναι γεωμετρική πρόοδος και συγκλίνει για $|x| < \xi_i$. Αν γίνει η αναγωγή των ομοίων όρων στην παράσταση R(x), τότε η R(x) γράφεται :

$$R(x) = \sum_{k=0}^{\infty} \beta_k.x^k \quad \text{, όπου}$$

$$\beta_k = -\left(\frac{\alpha_1}{\xi_1^{k+1}} + \frac{\alpha_2}{\xi_2^{k+1}} + \ldots + \frac{\alpha_n}{\xi_n^{k+1}} \right).$$

Σχηματίζουμε την ποσότητα q_k :

$$q_k = \frac{\beta_k}{\beta_{k-1}} = \frac{\frac{\alpha_1}{\xi_1^{k+1}} + \frac{\alpha_2}{\xi_2^{k+1}} + \ldots + \frac{\alpha_n}{\xi_n^{k+1}}}{\frac{\alpha_1}{\xi_1^{k}} + \frac{\alpha_2}{\xi_2^{k}} + \ldots + \frac{\alpha_n}{\xi_n^{k}}} = \frac{1}{\xi_1} \frac{1 + \frac{\alpha_2}{\alpha_1}\left(\frac{\xi_1}{\xi_2}\right)^{k+1} + \ldots + \frac{\alpha_n}{\alpha_1}\left(\frac{\xi_1}{\xi_n}\right)^{k+1}}{1 + \frac{\alpha_2}{\alpha_1}\left(\frac{\xi_1}{\xi_2}\right)^{k} + \ldots + \frac{\alpha_n}{\alpha_1}\left(\frac{\xi_1}{\xi_n}\right)^{k}}$$

Επειδή $\left|\frac{\xi_1}{\xi_i}\right| < 1$ για i = 1, 2, 3, . . . , n, $\lim_{k\to\infty} q_k \to \frac{1}{\xi_1}$.

Επομένως, με την ακολουθία q_k προσεγγίζεται η ρίζα ξ_1.

Για την εκτίμηση του σφάλματος της παραπάνω προσεγγιστικής μεθόδου, μελετούμε το σφάλμα $\varepsilon_k = \frac{1}{\xi_1} - q_k$. Προκύπτει ότι

$$\frac{\frac{1}{\xi_1} - q_k}{\left(\frac{\xi 1}{\xi 2}\right)^k} = \frac{1}{\xi_1} \frac{\frac{\alpha_2}{\alpha_1}\left(1 - \frac{\xi_1}{\xi_2}\right) + \frac{\alpha_3}{\alpha_1}\left(1 - \frac{\xi_1}{\xi_3}\right)\left(\frac{\xi_2}{\xi_3}\right)^k + \ldots + \frac{\alpha_n}{\alpha_1}\left(1 - \frac{\xi_1}{\xi_n}\right)\left(\frac{\xi_2}{\xi_n}\right)^k}{1 + \frac{\alpha_2}{\alpha_1}\left(\frac{\xi_1}{\xi_2}\right)^k + \frac{\alpha_3}{\alpha_1}\left(\frac{\xi_1}{\xi_3}\right)^k + \ldots + \frac{\alpha_n}{\alpha_1}\left(\frac{\xi_1}{\xi_n}\right)^k}$$

και

$$\frac{\dfrac{\frac{1}{\xi_1}-q_{k+1}}{\left(\frac{\xi_1}{\xi_2}\right)^{k+1}}}{\dfrac{\frac{1}{\xi_1}-q_k}{\left(\frac{\xi_1}{\xi_2}\right)^{k}}} \underset{k\to\infty}{\to} 1 \ .$$

Άρα $\lim\limits_{k\to\infty}\dfrac{\varepsilon_{k+1}}{\varepsilon_k}=\dfrac{\xi_1}{\xi_2}$ (γραμμική σύγκλιση).

Σχηματίζουμε τώρα τις ποσότητες :

$$e_k = q_{k+1} - q_k, \qquad e_{k-1} = q_k - q_{k-1}$$

.

$$\frac{e_k}{e_{k-1}}=\frac{\left(q_{k+1}-\frac{1}{\xi_1}\right)-\left(q_k-\frac{1}{\xi_1}\right)}{\left(q_k-\frac{1}{\xi_1}\right)-\left(q_{k-1}-\frac{1}{\xi_1}\right)}=\frac{\varepsilon_{k+1}-\varepsilon_k}{\varepsilon_k-\varepsilon_{k-1}}=\frac{\frac{\varepsilon_{k+1}}{\varepsilon_k}-1}{1-\frac{\varepsilon_{k-1}}{\varepsilon_k}} \ .$$

Άρα, $$\lim_{k\to\infty}\frac{e_k}{e_{k-1}}=\frac{\frac{\xi_1}{\xi_2}-1}{1-\frac{\xi_2}{\xi_1}}=\frac{\xi_1}{\xi_2} \ .$$

Έστω $p_k=\dfrac{e_k}{e_{k-1}}q_k$. Τότε $\lim\limits_{k\to\infty}p_k=\dfrac{1}{\xi_2}$.

Επομένως, με την ακολουθία p_k προσεγγίζεται η δεύτερη ρίζα.

Συνοψίζουμε τον αλγόριθμο :

(a) Αναπτύσσεται δυναμοσειρά για την R(x), όπου $R(x)=\dfrac{1}{p_n(x)}$:

$$R(x) = \sum_{k=0}^{\infty} \beta_k . x^k \ .$$

(b) Από την ακολουθία των β σχηματίζονται οι ακολουθίες

$$q_k = \frac{\beta_k}{\beta_{k-1}},\ e_k = q_{k+1} - q_k,\ p_k = \frac{e_k}{e_{k-1}} q_k \ .$$

(c) Οι δύο πρώτες ρίζες προσεγγίζονται για μεγάλο **k** με τις ακολουθίες p_k και q_k .

$$(q_k \to \frac{1}{\xi_1},\ p_k \to \frac{1}{\xi_2}).$$

Οι συντελεστές του παραπάνω αναπτύγματος υπολογίζονται ως εξής :

$$\frac{1}{a_0 + a_1.x + a_2.x^2 + ... + a_n.x^n} = \beta_0 + \beta_1.x + \beta_2.x^2 + \ldots$$

Χωρίς περιορισμό της γενικότητας μπορεί να τεθεί $a_0 = 1$, οπότε

$$1 = (1 + a_1.x + a_2.x^2 + \ldots + a_n.x^n).(\beta_0 + \beta_1.x + \beta_2.x^2 + \ldots).$$

Εξισώνονται οι συντελεστές των ομοίων όρων και τότε προκύπτουν οι εξισώσεις :

$\beta_0 = 1$

$\beta_1 + \beta_0.a_1 = 0$

$\beta_2 + \beta_1.a_1 + \beta_0.a_2 = 0$

$\beta_3 + \beta_2.a_1 + \beta_1.a_2 + \beta_0.a_3 = 0$

. . . . . .

Παράδειγμα

Να βρεθούν οι δύο πρώτες ρίζες του πολυωνύμου

$p_3(x) = -0.5x^3 + 2.75x^2 - 3.25x + 1.$

k	**β_k**	**q_k**	**e_k**	**p_k**
0	1			
1	3.25	3.25	-0.8462	
2	7.8125	2.4038	-0.2338	0.6642
3	16.953	2.1700	-0.0914	0.8483
4	.	.	.	.

Σαν άσκηση να συμπληρωθεί ο παραπάνω πίνακας.

3.4 ΜΙΓΑΔΙΚΕΣ ΡΙΖΕΣ ΠΟΛΥΩΝΥΜΩΝ

Αν ένα πολυώνυμο με πραγματικούς συντελεστές έχει μία μιγαδική ρίζα, τότε έχει ρίζα και τη συζυγή της. Αυτό σημαίνει ότι στην ανάλυση πολυωνύμου σε γινόμενο προκύπτει ένα δευτεροβάθμιο πολυώνυμο ως παράγοντας.

Η διαίρεση ενός πολυωνύμου n βαθμού με ένα τέτοιο τετραγωνικό παράγοντα δίνει ένα πηλίκο q(x) βαθμού n-2 και ένα υπόλοιπο 1ου βαθμού :

$$p(x) = (x^2 - \alpha.x - \beta).q(x) + b_1.(x - \alpha) + b,$$

όπου $p(x) = a_n.x^n + a_{n-1}.x^{n-1} + \ldots + a_0,$ a_i = πραγματικοί

και $q(x) = b_n.x^{n-2} + b_{n-1}.x^{n-3} + \ldots + b_2.$

Για να έχει το p(x) μιγαδική ρίζα, θα πρέπει να διαιρείται με τον τετραγωνικό παράγοντα $x^2 - \alpha.x - \beta$. Συνεπώς θα πρέπει $b_1 = 0$, $b_0 = 0$. Προσδιορίζονται λοιπόν τα α και β, έτσι που $b_1 = 0$ και $b_0 = 0$.

$$p(x) = b_n.x^n + (b_{n-1} - \alpha.b_n).x^{n-1} + (b_{n-2} - \alpha.b_{n-1} - \beta.b_n).x^{n-2} +$$

$$+ (b_{n-3} - \alpha.b_{n-2} - \beta.b_{n-1}).x^{n-3} + \ldots + (b_1 - \alpha.b_2 - \beta.b_3).x + (b_0 - \alpha.b_1 - \beta.b_2)$$

Εξισώνοντας τους ομοιόβαθμους όρους, προκύπτει ότι :

b_n	=	a_n					
b_{n-1}	=	a_{n-1}	+	$\alpha.b_n$			
b_{n-2}	=	a_{n-2}	+	$\alpha.b_{n-1}$	+	$\beta.b_n$	
b_{n-3}	=	a_{n-3}	+	$\alpha.b_{n-2}$	+	$\beta.b_{n-1}$	
. .	.	. .	.	. . .	.	. . .	(3.6)
b_1	=	a_1	+	$\alpha.b_2$	+	$\beta.b_3$	
b_0	=	a_0	+	$\alpha.b_1$	+	$\beta.b_2$	

Από τις εξισώσεις (3.6) φαίνεται η εξάρτηση των b_0, b_1 από τα **α** και **β**. Πρέπει τώρα,

$$b_0(\alpha, \beta) = 0 \text{ και } b_1(\alpha, \beta) = 0. \qquad (3.7)$$

Οι (3.7) αποτελούν σύστημα δύο μη γραμμικών εξισώσεων με δύο αγνώστους. Για την επίλυσή του μπορούμε να ξεκινήσουμε με αρχικές

τιμές α_0 και β_0, οι οποίες δεν ικανοποιούν κατ' ανάγκη τις (3.7). Θεωρούμε νέες τιμές α_1, β_1 με $\alpha_1 = \alpha_0 + \delta\alpha_0$, $\beta_1 = \beta_0 + \delta\beta_0$ και υποθέτουμε ότι οι α_1 και β_1 αποτελούν καλύτερη προσέγγιση για τη λύση του συστήματος (3.7). Αν θέσουμε $b_0(\alpha_1, \beta_1) = 0$ και $b_1(\alpha_1, \beta_1) = 0$, αναπτύξουμε κατά Taylor και κρατήσουμε μόνο τους όρους πρώτης τάξης, τότε

$$b_0(\alpha, \beta) = b_0(\alpha_0, \beta_0) + b_{0\alpha}(\alpha_0, \beta_0).(\alpha - \alpha_0) + b_{0\beta}(\alpha_0, \beta_0).(\beta - \beta_0)$$

$$b_1(\alpha, \beta) = b_1(\alpha_0, \beta_0) + b_{1\alpha}(\alpha_0, \beta_0).(\alpha - \alpha_0) + b_{1\beta}(\alpha_0, \beta_0).(\beta - \beta_0)$$

όπου $\alpha - \alpha_0 = \delta\alpha_0$ και $\beta - \beta_0 = \delta\beta_0$

$$b_{0\alpha} = \frac{\partial b_0}{\partial \alpha} \qquad b_{0\beta} = \frac{\partial b_0}{\partial \beta}$$

$$b_{1\alpha} = \frac{\partial b_1}{\partial \alpha} \qquad b_{1\beta} = \frac{\partial b_1}{\partial \beta}$$

και $b_0(\alpha, \beta) = 0, \qquad b_1(\alpha, \beta) = 0.$

Άρα καταλήγουμε στην

$$\frac{\partial b_0}{\partial \alpha}.\delta\alpha_0 + \frac{\partial b_0}{\partial \beta}.\delta\beta_0 = -b_0(\alpha_0, \beta_0)$$

$$\frac{\partial b_1}{\partial \alpha}.\delta\alpha_0 + \frac{\partial b_1}{\partial \beta}.\delta\beta_0 = -b_1(\alpha_0, \beta_0) \ . \tag{3.8}$$

Οι μερικές παράγωγοι που εμφανίζονται υπολογίζονται για $\alpha = \alpha_0$ και $\beta = \beta_0$. Οι (3.8) αποτελούν ένα σύστημα δύο γραμμικών εξισώσεων με δύο αγνώστους, τους $\delta\alpha_0$ και $\delta\beta_0$. Οι τιμές α_1 και β_1, που προκύπτουν,

εισάγονται στο σύστημα (3.8) στη θέση των α_0 και β_0, υπολογίζονται έτσι νέες τιμές α_1, β_1 κ.λ.π.

Για τον υπολογισμό των μερικών παραγώγων που εμφανίζονται στις (3.8), παραγωγίζονται οι (3.7) ως προς α και β :

$$\frac{\partial b_{n-1}}{\partial \alpha} = b_n$$

$$\frac{\partial b_{n-2}}{\partial \alpha} = b_{n-1} + \alpha \frac{\partial b_{n-1}}{\partial \alpha}$$

$$\frac{\partial b_{n-3}}{\partial \alpha} = b_{n-2} + \alpha \frac{\partial b_{n-2}}{\partial \alpha} + \beta \frac{\partial b_{n-1}}{\partial \alpha}$$

. . . . . .

$$\frac{\partial b_1}{\partial \alpha} = b_2 + \alpha \frac{\partial b_2}{\partial \alpha} + \beta \frac{\partial b_3}{\partial \alpha}$$

$$\frac{\partial b_0}{\partial \alpha} = b_1 + \alpha \frac{\partial b_1}{\partial \alpha} + \beta \frac{\partial b_2}{\partial \alpha}$$

Έστω $c_{k+1} = \frac{\partial b_k}{\partial \alpha}$, k = 0, 1, 2, . . . , n-1. Τότε,

c_n	=	b_n				
c_{n-1}	=	b_{n-1}	+	$\alpha.c_n$		
c_{n-2}	=	b_{n-2}	+	$\alpha.c_{n-1}$	+	$\beta.c_n$
. .	.	.	.	.		
c_2	=	b_2	+	$\alpha.c_3$	+	$\beta.c_4$
c_1	=	b_1	+	$\alpha.c_2$	+	$\beta.c_3$

$$c_2 = \frac{\partial b_1}{\partial \alpha}$$

$$c_1 = \frac{\partial b_0}{\partial \alpha}$$

Έστω $d_{i+2} = \frac{\partial b_i}{\partial \beta}$, i = 0, 1, 2, . . . , n-2. Τότε,

d_n	=	b_n				
d_{n-1}	=	b_{n-1}	+	$\alpha.d_n$		
d_{n-2}	=	b_{n-2}	+	$\alpha.d_{n-1}$	+	$\beta.d_n$
. .	.	. .	.			
d_3	=	b_3	+	$\alpha.d_4$	+	$\beta.d_5$
d_2	=	b_2	+	$\alpha.d_3$	+	$\beta.d_4$

$$d_1 = \frac{\partial b_1}{\partial \beta}$$

$$d_2 = \frac{\partial b_0}{\partial \beta}$$

Προκύπτει αμέσως ότι, $c_k = d_k$, $k = 2, 3, \ldots, n$.

Άρα, $d_2 = c_2$, $d_3 = c_3$ και $\frac{\partial b_0}{\partial \beta} = c_2$, $\frac{\partial b_1}{\partial \beta} = c_3$.

Το σύστημα των εξισώσεων (3.8) γράφεται τώρα ως εξής για τις αρχικές τιμές α_0 και β_0 :

$$c_1(\alpha_0, \beta_0).\delta\alpha_0 + c_2(\alpha_0, \beta_0).\delta\beta_0 = -b_0(\alpha_0, \beta_0)$$

$$c_2(\alpha_0, \beta_0).\delta\alpha_0 + c_3(\alpha_0, \beta_0).\delta\beta_0 = -b_1(\alpha_0, \beta_0). \qquad (3.10)$$

Παρακάτω συνοψίζουμε τον **αλγόριθμο** :

(1) Εκλέγονται αρχικές τιμές α_0, β_0.

(2) Υπολογίζονται τα b_k από την (3.6).

(3) Υπολογίζονται τα c_k από την (3.9).

(4) Υπολογίζονται τα $d\alpha_0$ και $d\beta_0$ από το σύστημα (3.10).

(5) Οι νέες τιμές δίνονται από τις εξισώσεις

$$\alpha_1 = \alpha_0 + \delta\alpha_0 \text{ και } \beta_1 = \beta_0 + \delta\beta_0.$$

(6) Ονομάζονται $\alpha_0 \leftarrow \alpha_1$ και $\beta_0 \leftarrow \beta_1$ και η διαδικασία επιστρέφει στο βήμα (2).

Οι ρίζες είναι $x_{1,2} = \dfrac{\alpha \pm \sqrt{\alpha^2 + 4\beta}}{2}$.

Ο αλγόριθμος που μόλις περιγράφτηκε ονομάζεται μέθοδος Newton - Bairstow. Στο κεφάλαιο 9 δίνεται πρόγραμμα που εφαρμόζει τον αλγόριθμο αυτό.

Για παράδειγμα απομονώνεται ο τετραγωνικός παράγοντας $(1 + x + x^2)$ από το πολυώνυμο $1 + 2x + 2x^2 + x^3$. Αρχικές τιμές παίρνονται οι $\alpha_0 = -2$, $\beta_0 = -3$.

3.5 ΕΝΤΟΠΙΣΜΟΣ ΡΙΖΩΝ ΠΟΛΥΩΝΥΜΟΥ

Έστω σύνολο συνεχών συναρτήσεων $\{ f_0(x), f_1(x), \ldots, f_n(x) \}$.

Ορισμός

Το παραπάνω σύνολο ονομάζεται ***ακολουθία Sturm*** *στο διάστημα* [a, b], *αν έχει τις εξής ιδιότητες :*

(i) Η $f_0(x)$ *έχει το πολύ, απλές ρίζες στο διάστημα* [a, b].

(ii) Η $f_{\mathbf{n}}(x) \neq 0$ *στο διάστημα* [a, b].

(iii) Αν $f_i(\xi) = 0$, *τότε* $f_{i-1}(\xi).f_{i+1}(\xi) < 0$ *για κάθε ρίζα* ξ *της* $f_i(x)$ *στο διάστημα* [a, b] *και για* $i = 1, 2, \ldots, n$.

(iv) Αν $f_0(\xi) = 0$, *τότε* $f_0'(\xi).f_1(\xi) > 0$ *για κάθε ρίζα* ξ *της* $f_0(x)$ *στο διάστημα* [a, b].

Για μια ακολουθία Sturm ισχύει η πρόταση :

Θεώρημα

Έστω S(x) *ο αριθμός των αλλαγών προσήμου στην ακολουθία Sturm* { $f_0(x)$, $f_1(x)$, . . . , $f_n(x)$ }. *Τότε ο αριθμός των ριζών της* f_0 *στο διάστημα* [a, b] *δίνεται από την διαφορά* S(a) - S(b).

Απόδειξη : Καθώς το x μεταβαίνει από το a στο b, ο αριθμός των αλλαγών προσήμου μπορεί να μεταβληθεί, αν κάποια από τις συναρτήσεις αλλάξει πρόσημο. Από τη συνθήκη (ii) αυτό δεν μπορεί να συμβεί στη συνάρτηση f_n. Εξετάζουμε λοιπόν τις άλλες συναρτήσεις. Έστω ότι το x περνά από μία ρίζα ξ της $f_i(x)$, i = 1, 2, 3, . . . , n, δηλαδή $f_i(\xi) = 0$. Αν ληφθεί υπόψη η συνθήκη (iii) και η συνέχεια των συναρτήσεων στην γειτονιά της ξ, σχηματίζονται οι παρακάτω πίνακες για τις δυνατές περιπτώσεις προσήμων των f_{i-1}, f_i, f_{i+1} στη θέση x = ξ και στις γειτονικές θέσεις ξ-ε και ξ+ε.

x	$f_{i-1}(x)$	$f_i(x)$	$f_{i+1}(x)$	x	$f_{i-1}(x)$	$f_i(x)$	$f_{i+1}(x)$
ξ-ε	+	±	-	ξ-ε	-	±	+
ξ	+	0	-	ξ	-	0	+
ξ+ε	+	∓	-	ξ+ε	-	∓	+

Παρατηρούμε ότι δεν συμβαίνει μεταβολή στον αριθμό αλλαγών προσήμου πριν και μετά το πέρασμα του x από τη θέση ξ. Απομένει η περίπτωση της f_0. (Συνθήκη (iv)).

x	$f_0(x)$	$f_1(x)$	x	$f_0(x)$	$f_1(x)$
ξ-ε	+	-	ξ-ε	-	+
ξ	0	-	ξ	0	+
ξ+ε	-	-	ξ+ε	+	+

Από τους δύο τελευταίους πίνακες παρατηρούμε ότι, καθώς το x περνά από μια ρίζα της f_0, οι αλλαγές προσήμου ελαττώνονται κατά μία. Από τη συγκέντρωση των παραπάνω παρατηρήσεων συνάγεται η ισχύς του θεωρήματος.

Αν τώρα η f_0 είναι πολυώνυμο, τότε κατασκευάζεται ως εξής μια ακολουθία Sturm.

Έστω $f_1(x) = f_0'(x)$. Διαιρούμε την f_0 με την f_1. Προκύπτει ένα πηλίκο $\Pi_1(x)$ και ένα υπόλοιπο $-f_2(x)$. Διαιρούμε την f_1 με την f_2. Προκύπτει ένα πηλίκο $\Pi_2(x)$ και ένα υπόλοιπο $-f_3(x)$:

$$
\begin{array}{lcll}
f_0(x) & = & f_1(x).\Pi_1(x) & - \; f_2(x) \\
f_1(x) & = & f_2(x).\Pi_2(x) & - \; f_3(x) \\
\ldots & \ldots & \ldots\ldots & \ldots \\
f_{i-1}(x) & = & f_i(x).\Pi_i(x) & - \; f_{i+1}(x) \\
f_i(x) & = & f_{i+1}(x).\Pi_{i+1}(x) & - \; f_{i+2}(x) \\
\ldots & \ldots & \ldots\ldots & \ldots \\
f_{n-2}(x) & = & f_{n-1}(x).\Pi_{n-1}(x) & - \; f_n(x) \\
f_{n-1}(x) & = & f_n(x).\Pi_n(x) &
\end{array}
\qquad (3.11)
$$

Η διαδικασία που περιγράφτηκε από τις εξισώσεις (3.11) τερματίζεται όταν βρεθεί ο μέγιστος κοινός διαιρέτης f_n των πολυωνύμων $f_0(x)$ και $f_1(x)$. Πράγματι, εύκολα φαίνεται από τις εξισώσεις (3.11) ότι η f_n είναι κοινός διαιρέτης όλων των f_i και ότι, επίσης, κάθε άλλος διαιρέτης των f_0 και f_1 θα διαιρεί και την f_n. Μπορεί συνεπώς να υποτεθεί ότι $f_n(x)$ = σταθερά $\neq 0$. Γιατί αν αυτό δεν συμβαίνει, μπορούμε να διαιρέσουμε τα $f_0, f_1, \ldots, f_{n-1}$ με το f_n. Για τον ίδιο λόγο μπορεί να υποτεθεί ότι η $f_0(x)$ έχει το πολύ απλές ρίζες στο διάστημα [a, b]. Κατά συνέπεια ικανοποιούνται οι συνθήκες (i) και (ii). Ακόμα, αν $f_i(\xi) = 0$, τότε από τις (3.11) είναι προφανές ότι $f_{i-1}(\xi).f_{i+1}(\xi) = - f_{i+1}^2(\xi) < 0$. Επίσης αν $f_0(\xi) = 0$, τότε $f_0'(\xi).f_1(\xi) = f_0'^2(\xi) > 0$ επειδή $f_0'(x) = f_1(x)$. Συνεπώς ικανοποιούνται και οι συνθήκες (iii) και (iv).

Αυτό σημαίνει ότι ο αλγόριθμος (3.11) ορίζει μια ακολουθία Sturm, που μπορεί να χρησιμεύσει για τον προσδιορισμό του αριθμού των ριζών ενός πολυωνύμου μέσα σε ένα δεδομένο διάστημα.

Παράδειγμα

Έστω το πολυώνυμο $p(x) = x^3 - 6x^2 + 3x - 9$.

Σχηματίζεται η ακολουθία Sturm :

$f_0(x) = x^3 - 6x^2 + 3x - 9$, $\qquad f_1(x) = 3x^2 - 12x + 3$,

$f_2(x) = 6x + 7$, $\qquad f_3(x) = -\dfrac{253}{12}$.

Παρακάτω δίνεται ο πίνακας για την αλλαγή προσήμων σε διάφορα διαστήματα.

x	$f_0(x)$	$f_1(x)$	$f_2(x)$	$f_3(x)$	S(x)
$-\infty$	$-\infty$	$+\infty$	$-\infty$	$-253/12$	2
0	-9	3	7	$-253/12$	2
10	421	183	67	$-253/12$	1
$+\infty$	$+\infty$	$+\infty$	$+\infty$	$-253/12$	1

$S(-\infty) - S(0) = 0$, $S(10) - S(\infty) = 0$, $S(0) - S(10) = 1$.

Άρα η f(x) έχει μία πραγματική ρίζα στο [0, 10].

4. ΕΠΙΛΥΣΗ ΓΡΑΜΜΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

4.1 ΑΠΑΛΟΙΦΗ GAUSS

Στην ανάπτυξη που ακολουθεί υποτίθενται γνωστές από τη Γραμμική Άλγεβρα οι στοιχειώδεις πράξεις των πινάκων καθώς και οι συνθήκες κάτω από τις οποίες ένα σύστημα γραμμικών εξισώσεων έχει λύση.

Ένα γραμμικό σύστημα **n** εξισώσεων με **n** αγνώστους γράφεται ως εξής :

$$
\begin{aligned}
a_{11}.x_1 &+ a_{12}.x_2 + a_{13}.x_3 + \ldots + a_{1n}.x_n = b_1 \\
a_{21}.x_1 &+ a_{22}.x_2 + a_{23}.x_3 + \ldots + a_{2n}.x_n = b_2 \\
a_{n1}.x_1 &+ a_{n2}.x_2 + a_{n3}.x_3 + \ldots + a_{nn}.x_n = b_n
\end{aligned}
\tag{4.1}
$$

Η λύση του συστήματος (4.1) είναι εύκολη, αν το σύστημα έχει τριγωνική μορφή :

$$
\begin{aligned}
a_{11}.x_1 + a_{12}.x_2 + \ldots &+ a_{1,n-1}.x_{n-1} + a_{1,n}.x_n = b_1 \\
a_{22}.x_2 + \ldots &+ a_{2,n-1}.x_{n-1} + a_{2,n}.x_n = b_2 \\
\ldots \quad &\ldots \quad \ldots \\
a_{n-2,n-2}.x_{n-2} &+ a_{n-2,n-1}.x_{n-1} + a_{n-2,n}.x_n = b_{n-2} \\
a_{n-1,n-1}.x_{n-1} &+ a_{n-1,n}.x_n = b_{n-1} \\
a_{n,n}.x_n &= b_n
\end{aligned}
\tag{4.2}
$$

Ξεκινώντας από την τελευταία εξίσωση του συστήματος (4.2), προκύπτει:

$$x_n = b_n/a_{n,n}$$

$x_{n-1} = (b_{n-1} - a_{n-1,n}.x_n)/a_{n-1,n-1}$

$x_{n-2} = (b_{n-2} - a_{n-2,n}.x_n - a_{n-2,n-1}.x_{n-1})/a_{n-2,n-2}$

. . .

$x_{n-k} = (b_{n-k} - a_{n-k,n}.x_n - a_{n-k,n-1}.x_{n-1} - a_{n-k,n-2}.x_{n-2} - . \quad . \quad . -$

$- a_{n-k,n-k+1}.x_{n-k+1})/a_{n-k,n-k}$

για k = 1, 2, . . . , n-1.

Σκοπός της απαλοιφής Gauss είναι να ανάγει το σύστημα (4.1) στη μορφή (4.2). Αυτό γίνεται σε διαδοχικά στάδια. Μετά το πέρας του k-1 σταδίου η μορφή του συστήματος θα είναι :

$$a_{11}^{(k-1)}.x_1 + a_{12}^{(k-1)}.x_2 + \ldots + a_{1,k-1}^{(k-1)}.x_{k-1} + a_{1,k}^{(k-1)}.x_k + \ldots + a_{1,n}^{(k-1)}.x_n = b_1^{(k-1)}$$

$$a_{22}^{(k-1)}.x_2 + \ldots + a_{2,k-1}^{(k-1)}.x_{k-1} + a_{2,k}^{(k-1)}.x_k + \ldots + a_{2,n}^{(k-1)}.x_n = b_2^{(k-1)}$$

$$a_{k-1,k-1}^{(k-1)}.x_{k-1} + a_{k-1,k}^{(k-1)}.x_k + \ldots + a_{k-1,n}^{(k-1)}.x_n = b_{k-1}^{(k-1)}$$

. . . . . .

(4.3)

$$a_{k,k}^{(k-1)}.x_k + \ldots + a_{k,n}^{(k-1)}.x_n = b_k^{(k-1)}$$

$$a_{k+1,k}^{(k-1)}.x_k + \ldots + a_{k+1,n}^{(k-1)}.x_n = b_{k+1}^{(k-1)}$$

$$a_{n,k}^{(k-1)}.x_k + \ldots + a_{n,n}^{(k-1)}.x_n = b_n^{(k-1)}$$

Ορίζεται τώρα ως w_{ij} ο πίνακας διαστάσεων $n * (n+1)$:

$w_{ij} = a_{ij}$, για i = 1, 2, . . . , n και j = 1, 2, . . . ,n,

ενώ $w_{i,n+1} = b_i$, i = 1, 2,. . . , n.

Στο στάδιο k η απαλοιφή συνεχίζεται κατά τον εξής τρόπο (βλέπε σύστημα εξισώσεων (4.3)). Η εξίσωση k διαιρείται με το w_{kk}, πολλαπλασιάζεται με w_{ik} και αφαιρείται από τη σειρά i. Αυτό γίνεται για όλες τις σειρές από i = k+1 ως i = n. Ο αλγόριθμος που εκτελεί αυτή την απαλοιφή γράφεται ως εξής :

```
FOR i = k+1 TO n
        m = w_ik/w_kk
        FOR j = k+1 TO n+1
                w_ij = w_ij - m.w_kj
        NEXT j
NEXT i.
```

Ο παραπάνω αλγόριθμος δεν αντιμετωπίζει το πρόβλημα της εμφάνισης κάποιου $w_{kk} = 0$. Πρέπει λοιπόν, πριν γίνει η απαλοιφή να εξεταστεί αν $w_{kk} = 0$. Αν όχι, τότε να βρεθεί ένα $i > k$ έτσι που $w_{ik} \neq 0$ και να γίνει αμοιβαία αλλαγή των σειρών i και k :

```
IF w_kk ≠ 0 THEN (**)
FOR i = k+1 TO n
                IF w_ik ≠ 0 THEN (*)
NEXT i
(*)     REM αμοιβαία αλλαγή των σειρών i και k
(**)    REM CONTINUE
```

Υπάρχει όμως ακόμα και το πρόβλημα ενός w_{kk} πολύ μικρού που να προκαλεί αριθμητικά σφάλματα. Για το λόγο αυτό, για i = k, k+1, . . . , n, βρίσκεται εκείνο το $i = i_1$ για το οποίο $|w_{i1}| = \max_{k \le i \le n} |w_{ik}|$.

Στη συνέχεια γίνεται αμοιβαία αλλαγή των σειρών k και i_1 :

```
m = 0
FOR i= k TO n
        IF |w_ik| > m THEN
            m = |w_ik|
            i_1 = i
        END IF
NEXT i
REM αμοιβαία αλλαγή των σειρών i_1 και k.
```

Για καλύτερη εξισορρόπηση των τάξεων μεγέθους των w_{ik}, ακολουθείται η παρακάτω βελτιωμένη διαδικασία :

Για i = k, k+1, . . . , n βρίσκεται το $d_i = \max\limits_{k \le j \le n} |w_{ij}|$.

Σχηματίζονται τα $|w_{ik}| / d_i$ (σχετικό μέγεθος του $|w_{ik}|$).

Βρίσκεται εκείνο το i_1 για το οποίο $|w_{i1,k}| / d_i = \max\limits_{k \le i \le n} |w_{ik}| / d_i$.

Άσκηση

Να λυθεί το παρακάτω σύστημα εξισώσεων.

(1)	3x	+	6y	+	9z	=	39
(2)	2x	+	5y	-	2z	=	3
(3)	x	+	3y	-	z	=	2

Λύση

Διαίρεση της (1) με το 3

(1’)	x	+	2y	+	3z	=	13
(2’)	2x	+	5y	-	2z	=	3
(3’)	x	+	3y	-	z	=	2

Αφαίρεση της (1’) * 2 από την (2’) και της (1’) από την (3’).

(1’’)	x	+	2y	+	3z	=	13
(2’’)			y	-	8z	=	-23
(3’’)			y	-	4z	=	-11

Διαίρεση της (2’’) με το 1 και αφαίρεση από την (3’’).

(1’’’)	x	+	2y	+	3z	=	13
(2’’’)			y	-	8z	=	-23
(3’’’)					4z	=	12

Οπισθοδρομική επίλυση.

z	=	3						
y	=	-23	+	24	=	1		
x	=	13	-	2	-	9	=	2

Άρα τελική λύση : x = 2, y = 1, z = 3.

4.2 ΣΥΣΤΗΜΑ ΜΕ ΤΡΙΔΙΑΓΩΝΙΚΟ ΠΙΝΑΚΑ

Το σύστημα των εξισώσεων (4.1) γράφεται ως εξής με το συμβολισμό των πινάκων : A.x = b. Ας υποτεθεί ότι έχουν βρεθεί πίνακες L και U, τέτοιοι που A = L.U και

$$L = \begin{bmatrix} \ell_{11} & 0 & 0 & . & 0 \\ \ell_{21} & \ell_{22} & 0 & . & . \\ \ell_{31} & \ell_{32} & \ell_{33} & 0 & . \\ . & . & . & . & 0 \\ \ell_{n1} & \ell_{n2} & \ell_{n3} & . & \ell_{nn} \end{bmatrix} \qquad U = \begin{bmatrix} 1 & u_{12} & u_{13} & . & u_{1n} \\ 0 & 1 & u_{23} & . & u_{2n} \\ . & 0 & 1 & u_{34} & . \\ . & . & . & . & . \\ 0 & . & . & 0 & 1 \end{bmatrix}$$

Τότε L.U.x = b και L.z = b, όπου z = U.x.

Λύνεται λοιπόν πρώτα το σύστημα :

$$\begin{array}{l} \ell_{11}.z_1 = b_1 \\ \ell_{21}.z_1 + \ell_{22}.z_2 = b_2 \\ \ell_{31}.z_1 + \ell_{32}.z_2 + \ell_{33}.z_3 = b_3 \\ . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\ \ell_{n1}.z_1 + \ell_{n2}.z_2 + \ell_{n3}.z_3 + \ldots + \ell_{nn}.z_n = b_n \end{array}$$

και στη συνέχεια, με γνωστά τώρα τα z_i, απομένει το παρακάτω σύστημα άνω τριγωνικής μορφής :

$$\begin{array}{r} x_1 + u_{12}.x_2 + u_{13}.x_3 + \ldots + u_{1n}.x_n = z_1 \\ x_2 + u_{23}.x_3 + \ldots + u_{2n}.x_n = z_2 \\ x_3 + \ldots + u_{3n}.x_n = z_3 \\ . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\ x_n = z_n \end{array}$$

Ενδιαφέρουσα ειδική περίπτωση αποτελεί πίνακας A τριδιαγωνικός :

$$\begin{bmatrix} B_1 & C_1 & 0 & . & . & . & . & . & 0 & 0 \\ A_2 & B_2 & C_2 & 0 & . & . & . & . & . & 0 \\ . & A_3 & B_3 & C_3 & & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & A_i & B_i & C_i & 0 & . \\ . & . & . & . & . & . & . & . & . & . \\ 0 & . & . & . & . & . & 0 & A_{n-1} & B_{n-1} & C_{n-1} \\ 0 & 0 & . & . & . & . & . & 0 & A_n & B_n \end{bmatrix} =$$

$$\begin{bmatrix} \beta_1 & 0 & . & . & . & . & . & 0 \\ \alpha_2 & \beta_2 & 0 & . & . & . & . & . \\ 0 & \alpha_3 & \beta_3 & 0 & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & 0 & \alpha_i & \beta_i & 0 & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & 0 & \alpha_{n-1} & \beta_{n-1} & 0 \\ 0 & . & . & . & . & 0 & \alpha_n & \beta_n \end{bmatrix} \text{x} \begin{bmatrix} 1 & c_1 & 0 & . & . & . & . & 0 \\ 0 & 1 & c_2 & 0 & . & . & . & . \\ . & 0 & 1 & c_3 & 0 & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & 0 & 1 & c_i & 0 & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & 0 & 1 & c_{n-1} \\ 0 & . & . & . & . & . & 0 & 1 \end{bmatrix}$$

Προκύπτει ότι

$\beta_1 = B_1, \qquad c_1 = C_1/B_1$

$\alpha_i = A_i, \qquad \beta_i = B_i - c_{i-1}.\alpha_i, \qquad c_i = C_i/\beta_i, \qquad i=2, 3, \ldots, n-1$

και $\qquad \alpha_n = A_n, \qquad \beta_n = B_n - c_{n-1}.\alpha_n.$

Η λύση του συστήματος A.x = b γίνεται σε δύο στάδια :

$z_1 = b_1/\beta_1, \;\; z_i = (b_i - \alpha_i.z_{i-1})/\beta_i, \qquad i = 2, 3, \ldots, n.$

$x_n = z_n$, $x_{n-k} = z_{n-k} - c_{n-k}.x_{n-k+1}$, $k = 1, 2, \ldots, n-1$.

Σαν άσκηση να λυθεί το σύστημα :

$$
\begin{array}{rcrcrcrcr}
x_1 & + & 2x_2 & & & & & = & 5 \\
x_1 & + & 3x_2 & + & x_3 & & & = & 10 \\
 & & x_2 & + & 3x_3 & + & x_4 & = & 15 \\
 & & & & x_3 & + & 2x_4 & = & 11
\end{array}
$$

4.3 ΕΠΑΝΑΛΗΠΤΙΚΕΣ ΜΕΘΟΔΟΙ

Οι γενικές εξισώσεις (4.1) μετασχηματίζονται ως εξής :

$$
\begin{array}{lllllll}
x_1 & = & -(a_{12}.x_2 & + a_{13}.x_3 & + \ldots & + & a_{1n}.x_n)/a_{11} + b_1/a_{11} \\
x_2 & = & -(a_{21}.x_1 & + a_{23}.x_3 & + \ldots & + & a_{2n}.x_n)/a_{22} + b_2/a_{22} \\
. & . & . & . & . & . & . \\
x_n & = & -(a_{n1}.x_1 & + a_{n2}.x_2 & + \ldots & + & a_{n,n-1}.x_{n-1})/a_{nn} + b_n/a_{nn}
\end{array}
\qquad (4.4)
$$

Οι εξισώσεις (4.4) γράφονται συνοπτικά ως εξής :

$\mathbf{x} = \mathbf{B}.\mathbf{x} + \mathbf{c}$, όπου $(\mathbf{c})_i = c_i = b_i/a_{ii}$, $i = 1, 2, \ldots, n$ και

$(\mathbf{B})_{ij} = -a_{ij}/a_{ii}$ για $i \neq j$ και $(\mathbf{B})_{ij} = 0$ για $i = j$.

Οι εξισώσεις (4.4) υποδεικνύουν ένα σχήμα διαδοχικών βελτιώσεων ενός αρχικού διανύσματος $x^{(0)}$:

$$\mathbf{x}^{(k+1)} = \mathbf{B}.\mathbf{x}^{(k)} + \mathbf{c}, \qquad k = 0, 1, 2, \ldots \qquad (4.5)$$

Η επαναληπτική διαδικασία σταματά όταν $\sum_{i=1}^{n} \left| x_i^{(k+1)} - x_i^{(k)} \right| < \varepsilon$,
όπου ε δεδομένος μικρός αριθμός.

Το υπολογιστικό σχήμα (4.5) είναι γνωστό ως **μέθοδος Jacobi**. Η μέθοδος αυτή μπορεί να τροποποιηθεί αν οι βελτιωμένες τιμές x_i χρησιμοποιούνται αμέσως μόλις προκύψουν. Τότε το επαναληπτικό σχήμα ονομάζεται **μέθοδος Gauss-Seidel**. Αν χρησιμοποιηθεί συμβολισμός αθροισμάτων και όχι πινάκων όπως στην (4.5), τότε οι δύο αυτές μέθοδοι περιγράφονται από τους τύπους :

$$x_i^{(k+1)} = \sum_{j=1}^{n} b_{ij} x_j^{(k)} + c_i \ , \qquad k = 0, 1, \ldots \quad \text{(Jacobi)}$$

$$x_i^{(k+1)} = \sum_{j=1}^{i-1} b_{ij} x_j^{(k+1)} + \sum_{j=i}^{n} b_{ij} x_j^{(k)} + c_i \ , \qquad k = 0, 1, \ldots \quad \text{(Gauss-Seidel)}$$

Εξετάζονται τώρα οι συνθήκες κάτω από τις οποίες συγκλίνει η επαναληπτική διαδικασία Jacobi. Ορίζεται το σφάλμα στο βήμα k :

$\boldsymbol{\varepsilon}^{(k)} = \mathbf{x}^{(k)} - \mathbf{x}$. Τότε από την (4.5) προκύπτει ότι

$$\boldsymbol{\varepsilon}^{(k+1)} = \mathbf{B} \cdot \boldsymbol{\varepsilon}^{(k)} \tag{4.6}$$

Με βάση την (4.6) μπορεί να γίνει ανάλυση των διαδοχικών σφαλμάτων και να βρεθούν ικανές συνθήκες, κάτω από τις οποίες η μέθοδος Jacobi συγκλίνει.

Η ανάλυση αυτή καταλήγει στο παρακάτω θεώρημα:

Θεώρημα

Αν ισχύει μια από τις παρακάτω συνθήκες (i) ή (ii), τότε η μέθοδος Jacobi συγκλίνει για οποιοδήποτε αρχικό διάνυσμα x_0.

(i) $\sum_i |a_{ij} / a_{ii}| < 1, \quad j = 1, 2, \ldots, N$ (ii) $\sum_j |a_{ij}| < a_{ii}|, \quad i = 1, 2, \ldots, N$

Σημειώνεται εδώ ότι η τελευταία συνθήκη εκφράζει διαγώνια υπεροχή για όλες τις σειρές του πίνακα.

Το θεώρημα αυτό αποδεικνύεται συνοπτικότερα με τη βοήθεια της έννοιας της νόρμας διανύσματος, η οποία εισάγεται στην επόμενη παράγραφο.

4.4 ΝΟΡΜΕΣ ΔΙΑΝΥΣΜΑΤΩΝ ΚΑΙ ΠΙΝΑΚΩΝ

Η έννοια του μεγέθους σε διανύσματα εκφράζεται και ως μέτρο, μήκος ή απόσταση, ανάλογα με την εκάστοτε γεωμετρική εποπτεία. Ο όρος «**νόρμα**» (**norm**) υπερβαίνει τις διάφορες γεωμετρικές θεωρήσεις και έχει γενικότερη αναλυτική σημασία.

Η νόρμα ενός διανύσματος x συμβολίζεται με ||x||. Για τον διδιάστατο ευκλείδειο χώρο, ως νόρμα διανύσματος x με συντεταγμένες x_1, x_2 μπορεί να ληφθεί το γνωστό μήκος:

$$\|\mathbf{x}\| = (x_1^2+x_2^2)^{1/2}$$

Η επέκταση στον τριδιάστατο και N-διάστατο χώρο είναι προφανής:

$$\|\mathbf{x}\| = (\sum_{i=1}^{N} x_i^2)^{1/2} \qquad (4.10)$$

Είναι εύκολο να αποδειχτεί ότι η παράσταση (4.10) ικανοποιεί τις παρακάτω συνθήκες:

(i) $\|\mathbf{x}\| \geq 0$ και $\|\mathbf{x}\| = 0$ αν και μόνο αν x=0.

(ii) $\|k\mathbf{x}\| \leq |k|\ \|\mathbf{x}\|$, όπου k πραγματικός αριθμός.

(iii) $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\|+\|\mathbf{y}\|$

Η τελευταία συνθήκη είναι γνωστή ως τριγωνική ανισότητα. Το εσωτερικό γινόμενο δύο διανυσμάτων ορίζεται ως εξής:

$$\mathbf{x} \bullet \mathbf{y} = \sum_{i=1}^{N} x_i \cdot y_i \ .$$

Μια τέταρτη ιδιότητα είναι η εξής:

(iv) $$\|\mathbf{x} \bullet \mathbf{y}\| \le \|\mathbf{x}\| \cdot \|\mathbf{y}\| \ .$$

Επειδή δε το εσωτερικό γινόμενο είναι βαθμωτό μέγεθος,

$$\|\mathbf{x} \bullet \mathbf{y}\| = |\mathbf{x} \bullet \mathbf{y}|$$

Οι συνθήκες (i), (ii), (iii) και (iv) αποτελούν τον ορισμό της νόρμας. Πρέπει όμως να σημειωθεί ότι η παράσταση (4.10) δεν είναι η μόνη που ικανοποιεί τις συνθήκες αυτές, δηλαδή δεν είναι η μόνη νόρμα του διανύσματος x. Για παράδειγμα, η

$$\|\mathbf{x}\| = \sum_{i=1}^{N} |x_i|$$

επίσης ικανοποιεί τις συνθήκες ορισμού της νόρμας. Αν τώρα η (4.10) γενικευτεί κατά τον εξής τρόπο:

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{N} |x_i|^p\right)^{1/p} \qquad (4.13)$$

μπορεί και πάλι να αποδειχτεί ότι και η $\|.\|_p$ είναι νόρμα. Μπορεί ακόμα εύκολα να επαληθευτεί ότι οι νόρμες (4.10) και (4.12) είναι ειδικές περιπτώσεις της (4.13):

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{N} |x_i| \qquad (4.14)$$

$$\|\mathbf{x}\|_2 = \left(\sum_{i=1}^{N} x_i^2\right)^{1/2} \qquad (4.15)$$

Ακόμη αποδεικνύεται ότι

$$\|\mathbf{x}\|_\infty = \max_{1\le i\le N}|x_i| \tag{4.16}$$

Πράγματι, από την (4.13), $\|\mathbf{x}\|_p{}^p = \sum_{i=1}^{N}|x_i|^p$

Αλλά, $\max_i|x_i|^p \le \sum_{i=1}^{N}|x_i|^p \le N.\max_i|x_i|^p$

Συνεπώς, $\max_i|x_i| \le \|\mathbf{x}\|_p \le N^{1/p}.\max_i|x_i|$

Επειδή δε, $N^{1/p} \to 1$, καθώς το $p \to \infty$, προκύπτει τελικά η (4.16).

Η έννοια της νόρμας επεκτείνεται σε πίνακες. Το μέγεθος $\|\mathbf{A}\|$ είναι νόρμα του πίνακα **A**, αν είναι συνάρτηση των στοιχείων του πίνακα A και αν ικανοποιεί τις συνθήκες (i), (ii), (iii) και (iv). Στις συνθήκες αυτές τα διανύσματα **x** και **y** αντικαθίστανται με **A** και **B**, όπου **A** και **B** πίνακες και όπου οι πράξεις που σημειώνονται είναι τώρα οι αντίστοιχες πράξεις μεταξύ πινάκων.

Νόρμες πινάκων συμβατές με τις νόρμες διανυσμάτων (4.14) και (4.16) μπορούν να διαμορφωθούν ως εξής, με βάση τη συνθήκη (iv), κατά την οποία θα πρέπει:

$$\|\mathbf{Ax}\| \le \|\mathbf{A}\|\,\|\mathbf{x}\|,$$

όπου A πίνακας και x διάνυσμα.

Θεωρούμε λοιπόν το διάνυσμα $\mathbf{Ax} = \sum_{j=1}^{N} A_{ij}.x_j$.

Τότε, η νόρμα $\|\,.\,\|_1$ του διανύσματος αυτού θα είναι:

$$\|\mathbf{Ax}\|_1 = \sum_{i=1}^{N}\left|\sum_{j=1}^{N} A_{ij}.x_j\right| \le \sum_i\sum_j |A_{ij}||x_j| = \sum_j \sum_i |A_{ij}||x_j| = \sum_j |x_j| \sum_i |A_{ij}| .$$

Συνεπώς, $\|\mathbf{Ax}\| \le \|\mathbf{A}\|_1 . \|\mathbf{x}\|$

όπου, εξ ορισμού, $\|\mathbf{A}\|_1 = \max_{1\le j\le N} \sum_{i=1}^{N} |A_{ij}| .$

Η παράσταση $\|A\|_1$ ικανοποιεί την συνθήκη (iv). Μπορεί επίσης εύκολα να αποδειχτεί ότι ικανοποιεί και τις συνθήκες (i), (ii) και (iii). Συνεπώς είναι νόρμα πίνακα και μάλιστα συμβατή με τη νόρμα διανύσματος $\| . \|_1$. Εντελώς ανάλογα, για τη νόρμα $\|A\|_\infty$ μπορεί να συναχθεί ότι

$$\|\mathbf{A}\|_\infty = \max_{1\le i\le N} \sum_{j=1}^{N} |A_{ij}| .$$

Ως εφαρμογή της έννοιας της νόρμας, αποδεικνύουμε το θεώρημα της προηγούμενης παραγράφου.

Από την (4.6) προκύπτει: $\|\boldsymbol{\varepsilon}^{(\kappa+1)}\| \le \|\mathbf{B}\| . \|\boldsymbol{\varepsilon}^{(\kappa)}\|$,

όπου $\|\mathbf{B}\|$ νόρμα του πίνακα B συμβατή με τη νόρμα $\|\boldsymbol{\varepsilon}\|$ του διανύσματος $\boldsymbol{\varepsilon}$.

Αν $\|\mathbf{B}\| \le M < 1$, τότε,

$$\left\|\boldsymbol{\varepsilon}^{(k+1)}\right\| \le M.\left\|\boldsymbol{\varepsilon}^{(k)}\right\| \le M^2.\left\|\boldsymbol{\varepsilon}^{(k-1)}\right\| \le \ldots \le M^{(k+1)}\left\|\boldsymbol{\varepsilon}^{(0)}\right\| \to 0$$

και η ακολουθία $\|\boldsymbol{\varepsilon}(k)\| \to 0$, καθώς το $k \to \infty$.

Αν $\|\mathbf{B}\| = \|\mathbf{B}\|_1$, τότε η παραπάνω συνθήκη γράφεται ως εξής, με βάση τον ορισμό της νόρμας $\| \, . \, \|_1$:

$$\max_j \sum_i \left| b_{ij} \right| < 1 \ . \qquad (\alpha)$$

Αν $\|\mathbf{B}\| = \|\mathbf{B}\|_\infty$, τότε η συνθήκη για τη σύγκλιση της μεθόδου γράφεται ως εξής με βάση τον ορισμό της νόρμας $\| \, . \| \, \infty$:

$$\max_i \sum_j \left| b_{ij} \right| < 1 \ . \qquad (\beta)$$

Οι δύο παραπάνω ανισότητες είναι ισοδύναμες αντίστοιχα με τις συνθήκες (i) και (ii) του θεωρήματος της παραγράφου (4.3).

4.5 ΙΔΙΟΤΙΜΕΣ ΚΑΙ ΙΔΙΟΔΙΑΝΥΣΜΑΤΑ ΠΙΝΑΚΩΝ

Ιδιοτιμή (eigenvalue) ενός τετραγωνικού πίνακα **A**, n-οστής τάξης, λέγεται ο αριθμός **λ** που ικανοποιεί τη σχέση

$$\mathbf{Ax} = \lambda \mathbf{x}, \qquad (4.17)$$

όπου $\mathbf{x} \neq 0$ το χαρακτηριστικό διάνυσμα ή **ιδιοδιάνυσμα** (**eigenvector**) του **A** που αντιστοιχεί στην τιμή **λ**. Οι ιδιοτιμές του A είναι οι ρίζες της χαρακτηριστικής εξίσωσης

$$D(\lambda) = \det(\mathbf{A} - \lambda.\mathbf{I}) = 0, \qquad (4.18)$$

όπου **I** ο μοναδιαίος πίνακας n-οστής τάξης. Η D(λ) λέγεται χαρακτηριστική ορίζουσα και μπορεί να γραφτεί σαν πολυώνυμο ως προς λ. Επομένως ο **A** έχει τουλάχιστον μία και το πολύ n διαφορετικές ιδιοτιμές.

Γνωρίζοντας τον πίνακα **A** μπορούμε να προσδιορίσουμε τους συντελεστές του χαρακτηριστικού πολυωνύμου και στη συνέχεια τις ρίζες, π.χ. με τη μέθοδο Newton. Όταν όμως ο αριθμός n είναι αρκετά

μεγάλος, απαιτούνται πολλά χρονοβόρα υπολογιστικά βήματα και έτσι προτιμούνται άλλες μέθοδοι με τις οποίες προσδιορίζονται τα όρια των ιδιοτιμών καθώς και οι προσεγγιστικές τους τιμές.

Το θεώρημα του Gershgorin που αναφέρεται παρακάτω δίνει έναν τρόπο προσδιορισμού των ορίων των ιδιοτιμών ενός πίνακα.

Θεώρημα (Gershgorin)

Έστω **λ** *μία ιδιοτιμή ενός τετραγωνικού πίνακα* $\mathbf{A} = (\mathbf{a_{ij}})$, n-*οστής τάξης. Τότε για κάποιο ακέραιο* k ($1 \le k \le n$) *ισχύει :*

$$|a_{kk} - \lambda| \le |a_{k1}| + |a_{k2}| + \ldots + |a_{k,k-1}| + |a_{k,k+1}| + \ldots + |a_{kn}| \quad (4.19)$$

Απόδειξη

Έστω x το ιδιοδιάνυσμα που αντιστοιχεί στην ιδιοτιμή λ του πίνακα A. Τότε

$$\mathbf{Ax} = \lambda\mathbf{x} \qquad \text{ή} \qquad (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0} \qquad (4.20)$$

Αν x_k είναι η συνιστώσα του **x** με τη μεγαλύτερη απόλυτη τιμή, θα ισχύει προφανώς $|x_m|/|x_k| \le 1$ ($m = 1, 2, \ldots, n$). Η (4.20) αποτελεί ένα σύστημα n εξισώσεων, εκ των οποίων η k μπορεί να γραφτεί

$$a_{k1}.x_1 + \ldots + a_{k,k-1}.x_{k-1} + (a_{kk} - \lambda).x_k + a_{k,k+1}.x_{k+1} + \ldots + a_{kn}.x_n = 0.$$

Επομένως

$a_{kk} - \lambda = -a_{k1}.(x_1/x_k) - \ldots - a_{k,k-1}.(x_{k-1}/x_k) - a_{k,k+1}.(x_{k+1}/x_k) - \ldots - a_{kn}.(x_n/x_k)$.

Είναι φανερό βέβαια ότι

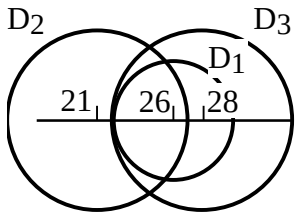
$$|x_1/x_k| \leq 1, \ldots, |x_n/x_k| \leq 1.$$

Θεωρώντας τις απόλυτες τιμές για τα δύο μέλη της παραπάνω εξίσωσης και εφαρμόζοντας την τριγωνική ανισότητα $|\alpha + \beta| \leq |\alpha| + |\beta|$ καταλήγουμε στη σχέση (4.19).

Η παραπάνω ανισότητα (4.19) προσδιορίζει για κάθε k = 1, 2, . . . , n έναν κυκλικό δίσκο στο μιγαδικό λ-επίπεδο με κέντρο το a_{kk} και ακτίνα που δίνεται από το δεξιό μέλος της ανισότητας. Το θεώρημα δείχνει ότι κάθε μία από τις ιδιοτιμές του **A** βρίσκεται σε έναν από τους n δίσκους.

Άσκηση

Οι ιδιοτιμές του πίνακα $\begin{vmatrix} 26 & -2 & 2 \\ 2 & 21 & 4 \\ 4 & 2 & 28 \end{vmatrix}$ βρίσκονται στους τρείς δίσκους του σχήματος.



D_1 : κέντρο στο 26, ακτίνα $|-2| + 2 = 4$

D_2 : κέντρο στο 21, ακτίνα $2 + 4 = 6$

D_3: κέντρο στο 28, ακτίνα $4 + 2 = 6$.

Μπορεί να επαληθευτεί εύκολα από τη Γραμμική Άλγεβρα ότι οι ιδιοτιμές του A είναι 30, 25 και 29. (Άσκηση !!).

4.5.1 ΠΡΟΣΕΓΓΙΣΤΙΚΟΣ ΥΠΟΛΟΓΙΣΜΟΣ ΤΩΝ ΙΔΙΟΤΙΜΩΝ ΠΙΝΑΚΩΝ

Μια αριθμητική διαδικασία για τον υπολογισμό προσεγγιστικών τιμών των ιδιοτιμών ενός n · n τετραγωνικού πίνακα $\mathbf{A} = (\mathbf{a_{ij}})$ είναι η εξής :

Θεωρείται ένα αρχικό διάνυσμα $x \neq 0$ με n συνιστώσες και υπολογίζονται διαδοχικά

$$\mathbf{x}_1 = \mathbf{A}\mathbf{x}_0, \quad \mathbf{x}_2 = \mathbf{A}\mathbf{x}_1, \quad \ldots, \quad \mathbf{x}_s = \mathbf{A}\mathbf{x}_{s-1}. \qquad (4.21)$$

Για απλούστευση γράφονται $\mathbf{x}_{s-1} = \mathbf{x}$ και $\mathbf{x}_s = \mathbf{y}$. Αν ο πίνακας Α είναι πραγματικός και συμμετρικός τότε ισχύει το παρακάτω θεώρημα.

Θεώρημα

Έστω **A** *πίνακας* n-*οστής τάξης πραγματικός και συμμετρικός και* $x \neq 0$ *πραγματικό διάνυσμα με* **n** *συνιστώσες.*

Έστω δε ότι $\mathbf{y} = \mathbf{A}\mathbf{x}$, $\quad m_0 = \mathbf{x}^T\mathbf{x}, \quad m_1 = \mathbf{x}^T\mathbf{y}, \quad m_2 = \mathbf{y}^T\mathbf{y}$.

Τότε ο λόγος $q = m_1/m_0$ *(λόγος του Rayleigh) είναι μία προσεγγιστική τιμή της ιδιοτιμής* λ *του πίνακα* **A**, *και αν* $q = \lambda + \varepsilon$, *όπου* **ε** *το σφάλμα στο λόγο* q, *τότε*

$$|\varepsilon| = \sqrt{\frac{m_2}{m_0} - q^2}$$

Παράδειγμα

Θεωρούμε τον πραγματικό και συμμετρικό πίνακα $\mathbf{A} = \begin{vmatrix} 8 & 1 & 1 \\ 1 & 5 & 2 \\ 1 & 2 & 5 \end{vmatrix}$

και έστω το αρχικό διάνυσμα $x_0 = \begin{vmatrix} 1 \\ 1 \\ 1 \end{vmatrix}$.

Τότε από τη σχέση (4.21) μπορούν να υπολογιστούν διαδοχικά τα

$$\mathbf{x}_1 = \begin{vmatrix} 10 \\ 8 \\ 8 \end{vmatrix} \qquad \mathbf{x}_2 = \begin{vmatrix} 96 \\ 66 \\ 66 \end{vmatrix} \qquad \mathbf{x}_3 = \begin{vmatrix} 900 \\ 558 \\ 558 \end{vmatrix} \qquad \mathbf{x}_4 = \begin{vmatrix} 8316 \\ 4806 \\ 4806 \end{vmatrix}$$

Θεωρώντας $\mathbf{x} = \mathbf{x}_3$ και $\mathbf{y} = \mathbf{x}_4$, έχουμε

$m_0 = \mathbf{x}^T\mathbf{x} = 1432728$, $\qquad m_1 = \mathbf{x}^T\mathbf{y} = 12847896$,

$m_2 = \mathbf{y}^T\mathbf{y} = 115351128$.

Άρα,

$q = m_1/m_0 = 8.967$ και $|\varepsilon| \le \sqrt{\dfrac{m_2}{m_0} - q^2} = 0.311$.

Επομένως η 8.967 είναι η προσεγγιστική τιμή μιας ιδιοτιμής που βρίσκεται μεταξύ 8.565 και 9.278.

Η ακριβής τιμή είναι λ = 9. (Άσκηση !!).

4.5.2. ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΙΔΙΟΤΙΜΩΝ

Μέθοδος των δυνάμεων

Έστωσαν $\mathbf{x}_1$, $\mathbf{x}_2$, …..,$\mathbf{x}_n$ n γραμμικώς ανεξάρτητα ιδιοδιανύσματα του πίνακα Α.

Λαμβάνεται τυχαίο διάνυσμα $\mathbf{y}_0$, το οποίο εκφράζεται ως γραμμικός συνδυασμός των $\mathbf{x}_1$, $\mathbf{x}_2$, …..,$\mathbf{x}_n$:

$$\mathbf{y}_0 = \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + ... + \alpha_n\mathbf{x}_n$$

Ακολουθούνται τα εξής βήματα:

$$\mathbf{z}_1 = \mathbf{A}\mathbf{y}_0 = \lambda_1\alpha_1\mathbf{x}_1 + \lambda_2\alpha_2\mathbf{x}_2 + ... + \lambda_n\alpha_n\mathbf{x}_n =$$

$$\lambda_1\left(\alpha_1\mathbf{x}_1 + \left(\frac{\lambda_2}{\lambda_1}\right)\alpha_2\mathbf{x}_2 + ... + \left(\frac{\lambda_n}{\lambda_1}\right)\alpha_n\mathbf{x}_n\right)$$

$$\mathbf{y}_1 = \mathbf{z}_1 / \|\mathbf{z}_1\|_\infty$$

……

$$\mathbf{z}_k = \mathbf{A}\mathbf{y}_{k-1} = \frac{\lambda_1^k}{\|\mathbf{z}_1\|_\infty \|\mathbf{z}_2\|_\infty ... \|\mathbf{z}_{k-1}\|_\infty}\left(\alpha_1\mathbf{x}_1 + \left(\frac{\lambda_2}{\lambda_1}\right)^k \alpha_2\mathbf{x}_2 + ... + \left(\frac{\lambda_n}{\lambda_1}\right)^k \alpha_n\mathbf{x}_n\right)$$

$$\mathbf{y}_k = \mathbf{z}_k / \|\mathbf{z}_k\|_\infty .$$

Τελικά

$$\frac{\mathbf{z}_{k+1\ \ i}}{\mathbf{y}_{k\ \ i}} = \lambda_1 \frac{\left(\alpha_1\mathbf{x}_1 + \left(\frac{\lambda_2}{\lambda_1}\right)^{k+1} \alpha_2\mathbf{x}_2 + ... + \left(\frac{\lambda_n}{\lambda_1}\right)^{k+1} \alpha_n\mathbf{x}_n\right)_i}{\left(\alpha_1\mathbf{x}_1 + \left(\frac{\lambda_2}{\lambda_1}\right)^{k} \alpha_2\mathbf{x}_2 + ... + \left(\frac{\lambda_n}{\lambda_1}\right)^{k} \alpha_n\mathbf{x}_n\right)_i}$$

Προκύπτει ότι

$$\frac{\mathbf{z}_{k+1\ \ i}}{\mathbf{y}_{k\ \ i}} \underset{k\to\infty}{\longrightarrow} \lambda_1$$

Το $\mathbf{y}_k$ τείνει στο αντίστοιχο ιδιοδιάνυσμα.

ΠΑΡΑΔΕΙΓΜΑ

$$\mathbf{A} = \begin{pmatrix} 1 & -3 & 2 \\ 4 & 4 & -1 \\ 6 & 3 & 5 \end{pmatrix}$$

$\mathbf{y}_0 = 1 \quad 1 \quad 1^T$

$\mathbf{y}_5 = .3276 \quad .0597 \quad 1^T$

$\mathbf{y}_{10} = .3007 \quad .0661 \quad 1^T$

$\mathbf{y}_{15} = .300 \quad .0667 \quad 1^T$

$\mathbf{z}_{16} = 2.0999 \quad .4668 \quad 7.001^T$

$(\mathbf{z}_{16}/\mathbf{y}_{15})_3 = 7 = \lambda_1$.

$\mathbf{y}_{15}$ = ιδιοδιάνυσμα

Η μέθοδος Jacobi για συμμετρικούς πίνακες.

Έχει αποδειχτεί ότι ένας συμμετρικός πίνακας είναι διαγωνιοποιήσιμος, Δηλαδή για πίνακα **A** συμμετρικό υπάρχει πίνακας **U** ορθομοναδιαίος ($\mathbf{U}^T\mathbf{U} = \mathbf{I}$) έτσι που

$\mathbf{U}^T\mathbf{A}\,\mathbf{U} = \mathbf{D}$, όπου **D** διαγώνιος πίνακας. Τα στοιχεία του πίνακα **D** είναι οι ιδιοτιμές του **A**. Με τη μέθοδο Jacobi επιδιώκεται η διαγωνιοποίηση του πίνακα **A**. Επιφέρονται διαδοχικοί μετασχηματισμοί ομοιότητας και προκύπτει η εξής ακολουθία πινάκων:

$\mathbf{A}^{(0)} = \mathbf{A}$, $\mathbf{A}^{(r)} = \mathbf{U}_r^T\mathbf{A}^{(r-1)}\mathbf{U}_r$, $r = 1,2,\ldots$

Κάθε πίνακας $\mathbf{U}_r$ κατασκευάζεται έτσι που να μηδενίζεται ένα συγκεκριμένο μη διαγώνιο στοιχείο του **A**, έστω a_{pq}. Το a_{pq} επιλέγεται ως το στοιχείο με τη μεγαλύτερη απόλυτη τιμή μεταξύ των μη διαγωνίων στοιχείων:

$$a_{pq} = \max_{i,k,i\neq k} |a_{ik}|$$

Ο πίνακας $\mathbf{U}_r$ ορίζεται ως εξής (παραλείπεται ο δείκτης r):

$u_{pp} = \cos\varphi$, $u_{pq} = \sin\varphi$, $u_{qq} = \cos\varphi$, $u_{qp} = -\sin\varphi$, $p<q$

$u_{ii} = 1$, $i \neq p$, $i \neq q$.

$u_{ik} = 0$, $i \neq k$, $i \neq p$, $i \neq q$.

Θεώρημα: Η ακολουθία των πινάκων $\mathbf{A}^{(r)}$, r = 1,2,.. συγκλίνει.
Απόδειξη:
Σύμφωνα με τον ορισμό του, ο $\mathbf{U}_r$ μπορεί να γραφεί ως το άθροισμα $\mathbf{U}_r = \mathbf{I} + \mathbf{V}_r$, του μοναδιαίου πίνακα $\mathbf{I}$ και ενός πίνακα $\mathbf{V}_r$, ο οποίος ορίζεται ως εξής (με παράλειψη και πάλι του δείκτη r):

$v_{pp} = \cos\varphi - 1$, $v_{pq} = \sin\varphi$, $v_{qq} = \cos\varphi - 1$, $v_{qp} = -\sin\varphi$, $p>q$

$v_{ik} = 0$, $i \neq p$, $i \neq q$.

$$\mathbf{U}(p,q,\phi) = \begin{array}{c} \begin{pmatrix} 1 & & & & & & & & & & & & & \\ & \cdot & & & & & & & & & & & & \\ & & \cdot & & & & & & & & & & & \\ & & & \cos\phi & & & & & \sin\phi & & & & & \\ & & & & 1 & & & & & & & & & \\ & & & & & \cdot & & & & & & & & \\ & & & & & & \cdot & & & & & & & \\ & & & & & & & \cdot & & & & & & \\ & & & & & & & & 1 & & & & & \\ & & & -\sin\phi & & & & & \cos\phi & & & & & \\ & & & & & & & & & \cdot & & & & \\ & & & & & & & & & & \cdot & & & \\ & & & & & & & & & & & 1 & & \\ & & & & & & & & & & & & 1 & \end{pmatrix} \end{array} \begin{array}{l} \\ \\ \\ \leftarrow p \\ \\ \\ \\ \\ \\ \leftarrow q \\ \\ \\ \\ \\ \\ \end{array}$$

Αναλυτικά ο πίνακας $\mathbf{U}_r$ γράφεται ως εξής:

$$u_{ij} = \delta_{ij} + (\cos\phi - 1)\delta_{ip}\delta_{jp} + \sin\varphi\delta_{ip}\delta_{jq} - \sin\phi\delta_{iq}\delta_{jp} + (\cos\phi - 1)\delta_{iq}\delta_{jq}$$

όπου το δ_{ij} είναι το δέλτα του Kronecker. Το γινόμενο $\mathbf{U}_i^T\mathbf{A}^{(r-1)}\mathbf{U}_r$ μπορεί τώρα να εκτελεστεί αναλυτικά. Συμβολίζεται με

$$a'_{ik} = \sum_j u_{ij}^T a_{jk}$$

το τυπικό στοιχείο το γινομένου $\mathbf{U}^T\mathbf{A}^{(r-1)}$ και με

$$a_{ik}^{//} = \sum_j \sum_l u_{ij}^T a_{jl} u_{lk}$$

το τυπικό στοιχείο το γινομένου $\mathbf{U}^T\mathbf{A}^{(r-1)}\,\mathbf{U}$

Αποδεικνύεται εύκολα (Άσκηση!) ότι

$$a_{pk}^{/} = a_{pk}\cos\phi - a_{qk}\sin\phi$$

$$a_{qk}^{/} = a_{pk}\sin\phi + a_{qk}\cos\phi$$

για k =1,2,..,n

$$a_{ik}^{/} = a_{ik}, \qquad i = 1, 2, ..., n, \qquad i \neq p, q \lim_{x\to\infty}$$

$$a_{pk}^{//} = a_{pk}\cos\phi - a_{qk}\sin\phi$$

$$a_{qk}^{//} = a_{pk}\sin\phi + a_{qk}\cos\phi$$

$$a_{ip}^{//} = a_{ip}\cos\phi - a_{iq}\sin\phi$$

$$a_{iq}^{//} = a_{ip}\sin\phi + a_{iq}\cos\phi$$

$$a_{ik}^{//} = a_{ik}, \qquad i, k \neq p, q$$

$$a_{pp}^{//} = a_{pp}\cos^2\phi + a_{qq}\sin^2\phi - 2a_{pq}\sin\varphi\cos\phi$$

$$a_{qq}^{//} = a_{pp}\sin^2\phi + a_{qq}\cos^2\phi + 2a_{pq}\sin\varphi\cos\phi$$

$$a_{pq}^{//} = a_{qp}^{//} = \frac{1}{2}(a_{pp} - a_{qq})\sin 2\phi + a_{pq}\cos 2\phi$$

Σημειώνεται εδώ ότι και ο πίνακας $\mathbf{A}^{//}$ είναι συμμετρικός.

$$a_{pq}^{//} = 0 \Rightarrow \begin{cases} \tan 2\varphi = \dfrac{2a_{pq}}{(a_{qq} - a_{pp})}, & \alpha\nu \quad a_{qq} \neq a_{pp} \\ \phi = \dfrac{\pi}{4}, & \alpha\nu \quad a_{qq} = a_{pp} \end{cases}$$

Για να έχουμε μονοσήμαντο αποτέλεσμα, η γωνία φ υπόκειται στον περιορισμό

$-\frac{\pi}{2} \leq 2\phi \leq \frac{\pi}{2}$. Αυτό σημαίνει ότι $\cos 2\phi \geq 0$.

Για τον λόγο αυτό οι τιμές των cos2φ και sin2φ λαμβάνονται ως εξής:

$\cos 2\varphi = \sigma(a_{qq} - a_{pp}) / R$ και

$\sin 2\varphi = 2\sigma a_{pq} / R$

όπου

$R = \sqrt{(a_{qq} - a_{pp}) + 4a_{pq}^2}$ και $\sigma = \begin{cases} 1 & \alpha\nu \quad a_{qq} > a_{pp} \\ -1 & \alpha\nu \quad a_{qq} < a_{pp} \end{cases}$

Από τα παραπάνω προκύπτει εύκολα ότι

$$a_{pk}^{//2} + a_{qk}^{//2} = a_{pk}^2 + a_{qk}^2, \quad k \neq p,q$$

$$a_{ip}^{//2} + a_{iq}^{//2} = a_{ip}^2 + a_{iq}^2, \quad i \neq p,q$$

$$a_{pp}^{//2} + a_{qq}^{//2} + a_{pq}^{//2} + a_{qp}^{//2} = a_{pp}^2 + a_{qq}^2 + a_{pq}^2 + a_{qp}^2, \quad p < q$$

ή

$$a_{pp}^{//2} + a_{qq}^{//2} = a_{pp}^2 + a_{qq}^2 + 2a_{pq}^2, \quad p < q$$

$$a_{ik}^{//} = a_{ik}, \qquad i,k \neq p,q$$

Ορίζεται ως $N^2(\mathbf{A}) = \sum_i \sum_j a_{ij}^2$ (νόρμα Frobenius)

Τότε είναι προφανές ότι

$$N^2(\mathbf{A}) = N^2(\mathbf{A}^{//}).$$

Ενδιαφέρον όμως παρουσιάζουν οι μη διαγώνιοι όροι, δεδομένου ότι προσπαθούμε να αποδείξουμε ότι οι όροι αυτοί τείνουν στο μηδέν. Για τον λόγο αυτό στο παραπάνω άθροισμα εξαιρούμε τους διαγωνίους όρους και έχουμε την νέα ποσότητα:

$$N_d^2(\mathbf{A}) = \sum_i \sum_{\substack{j \\ j \neq i}} a_{ij}^2$$

Προφανώς,

$$N_d^2(\mathbf{A}^{//}) = N^2(\mathbf{A}^{//}) - \sum_i a_{ii}^{//2} = N^2(\mathbf{A}) - \sum_{\substack{i \\ i \neq p,q}} a_{ii}^{//2} - a_{pp}^{//2} - a_{qq}^{//2} =$$

$$= N^2(\mathbf{A}) - \sum_{\substack{i \\ i \neq p,q}} a_{ii}^2 - a_{pp}^2 - a_{qq}^2 - 2a_{pq}^2 = N^2(\mathbf{A}) - \sum_i a_{ii}^2 - 2a_{pq}^2 = N_d^2(\mathbf{A}) - 2a_{pq}^2$$

Η τελευταία σχέση γράφεται ως

$$N_d^2(\mathbf{A}^{(r+1)}) = N_d^2(\mathbf{A}^{(r)}) - 2a_{pq}^{(r)2}$$

Επειδή το στοιχείο a_{pq} επιλέγεται κάθε φορά, έτσι που να έχει τη μεγαλύτερη απόλυτη τιμή από όλα τα μη διαγώνια στοιχεία, δηλ. $\left|a_{pq}^{(r)}\right| \geq \left|a_{ik}^{(r)}\right| \quad \forall i,k \quad i < k$.

$a_{pq}^{(r)2} n(n-1) \geq N_d^2(\mathbf{A}^{(r)})$. Συνεπώς,

$$N_d^2(\mathbf{A}^{(r+1)}) \leq N_d^2(\mathbf{A}^{(r)})\left(1 - \frac{2}{n(n-1)}\right)$$

Για n > 2, έστω $K = \left(1 - \frac{2}{n(n-1)}\right)$. Τότε $0 < K < 1$. Συνεπώς,

$$N_d^2(\mathbf{A}^{(r+1)}) \leq N_d^2(\mathbf{A}^{(r)})K \leq ..... \leq N_d^2(\mathbf{A}^{(0)})K^{r+1} \underset{r \to \infty}{\to} 0$$

Άρα κάθε ένα από τα μη διαγώνια στοιχεία τείνει στο 0, ο.ε.δ.

Παράδειγμα:

Να υπολογιστούν με τη μέθοδο Jacobi όλες οι ιδιοτιμές και αντίστοιχα ιδιοδιανύσματα του πίνακα

$$\begin{pmatrix} 6 & 4 & -2 \\ 4 & 12 & -4 \\ -2 & -4 & 13 \end{pmatrix}$$

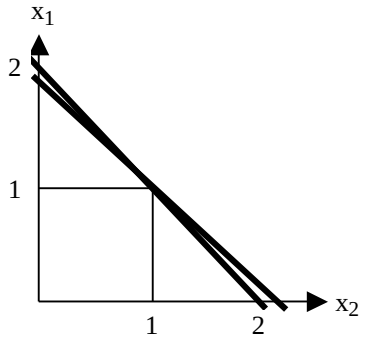
4.6 ΠΑΘΟΛΟΓΙΚΕΣ ΚΑΤΑΣΤΑΣΕΙΣ ΣΤΗΝ ΕΠΙΛΥΣΗ ΓΡΑΜΜΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ (ILL-CONDITIONED LINEAR SYSTEMS)

Έστω το σύστημα $\begin{vmatrix} 1.01 & 0.99 \\ 0.99 & 1.01 \end{vmatrix} . \begin{vmatrix} x_1 \\ x_2 \end{vmatrix} = \begin{vmatrix} 2 \\ 2 \end{vmatrix}$ με προφανή ακριβή λύση την $x_1 = 1$, $x_2 = 1$.

Θεωρούμε το σύστημα $\begin{vmatrix} 1.02 & 0.99 \\ 0.99 & 1.01 \end{vmatrix} . \begin{vmatrix} x_1 \\ x_2 \end{vmatrix} = \begin{vmatrix} 2 \\ 2 \end{vmatrix}$ στο οποίο υπάρχει μια μικρή μόνο διαφοροποίηση σε σχέση με το προηγούμενο, στον συντελεστή a_{11}.

Η επίλυση του δεύτερου συστήματος οδηγεί, με ακρίβεια τριών δεκαδικών ψηφίων, στην εξής λύση $x_1 = 0.798$, $x_2 = 1.197$.

Παρατηρούμε ότι η μεταβολή των x_1 και x_2 είναι δυσανάλογα μεγάλη σε σύγκριση με τη μεταβολή στον πίνακα Α του συστήματος. Για το φαινόμενο αυτό μπορεί να δοθεί μια εξήγηση, αν οι δύο εξισώσεις του αρχικού συστήματος θεωρηθούν ότι παριστάνουν δύο ευθείες σε ένα σύστημα αξόνων x_1, x_2, όπως φαίνεται στο Σχήμα.



Οι δύο αυτές ευθείες είναι σχεδόν παράλληλες και το σημείο τομής τους, δηλαδή η λύση του συστήματος, υφίσταται μεγάλες μεταβολές και όταν ακόμη οι μεταβολές της θέσης των ευθειών είναι μικρές.

Σε συστήματα περισσοτέρων εξισώσεων, η κατάσταση αυτή μπορεί να μην είναι προφανής από τη μορφή του συστήματος. Για παράδειγμα έστω το σύστημα :

$$\begin{vmatrix} 0.78 & -3.22 & 2.41 \\ 2.33 & 4.28 & -1.25 \\ 0.34 & 2.92 & -1.60 \end{vmatrix} . \begin{vmatrix} x_1 \\ x_2 \\ x_3 \end{vmatrix} = \begin{vmatrix} 0.82 \\ -0.55 \\ -0.62 \end{vmatrix}$$

Μπορεί να επαληθευτεί άμεσα, ότι η ακριβής λύση του συστήματος είναι $x_1 = -1$, $x_2 = 1$, $x_3 = 2$.

Ας επιχειρήσουμε αριθμητική επίλυση του συστήματος με απαλοιφή Gauss και οδήγηση (pivoting), κρατώντας όμως σε κάθε αριθμητική πράξη μόνο τρία σημαντικά ψηφία. Τότε τα διαδοχικά βήματα στον επαυξημένο πίνακα είναι τα εξής :

$$\begin{vmatrix} 2.33 & 4.28 & -1.25 & : & -0.55 \\ 0.78 & -3.22 & 2.41 & : & 0.82 \\ 0.34 & 2.92 & -1.60 & : & -0.62 \end{vmatrix}$$

$$\begin{vmatrix} 2.33 & 4.28 & -1.25 & : & -0.55 \\ 0 & -4.65 & 2.83 & : & 1.00 \\ 0 & 2.29 & -1.42 & : & -0.54 \end{vmatrix}$$

$$\begin{vmatrix} 2.33 & 4.28 & -1.25 & : & -0.55 \\ 0 & -4.65 & 2.83 & : & 1.00 \\ 0 & 0 & -0.0276 & : & -0.048 \end{vmatrix}$$

Η λύση που προκύπτει τώρα είναι $x_1^* = -0.853$, $x_2^* = 0.844$, $x_3^* = 1.74$.

Η διαφορά από την ακριβή λύση είναι πολύ μεγάλη, πράγμα που οφείλεται στη δυσανάλογη μεγέθυνση των σφαλμάτων που προέρχονται από τη στρογγύλευση στα τρία σημαντικά ψηφία.

Αν κρατηθούν τέσσερα ψηφία θα παρατηρηθεί κάποια βελτίωση, αλλά και πάλι οι διαφορές θα είναι σημαντικές. Η περίπτωση αυτή μπορεί να επαληθευτεί εύκολα. (Άσκηση !!).

Και στα δύο συστήματα που εξετάστηκαν υπάρχει ένα εγγενές πρόβλημα, για την ανίχνευση του οποίου πρέπει να αναζητηθεί ένα ποσοτικό κριτήριο.

Τα συστήματα αυτά εμφανίζουν μία παθολογική κατάσταση, η οποία χαρακτηρίζεται από μικρή τιμή της ορίζουσας του πίνακα Α. Η τιμή όμως της ορίζουσας δεν θα μπορούσε να αποτελέσει ένα καθολικό κριτήριο, επειδή είναι δυνατόν να εμφανίσει μικρή τιμή όταν τα στοιχεία του Α είναι μικρά, χωρίς να υπάρχει παθολογική κατάσταση. Μια άλλη σκέψη θα ήταν να αντικατασταθεί στα αριστερά μέλη των εξισώσεων η λύση που προκύπτει, για να διαπιστωθεί πόσο καλά αναπαράγονται τα δεξιά μέλη. Ορίζονται τα **υπολείμματα** (**residuals**) **r** ως εξής :

$\mathbf{r} = \mathbf{A}\bar{\mathbf{x}} - \mathbf{b}$, όπου $\bar{\mathbf{x}}$ προσεγγιστική λύση.

Τότε

$$\mathbf{r} = \mathbf{A}\bar{\mathbf{x}} - \mathbf{A}\mathbf{x} = \mathbf{A}(\bar{\mathbf{x}} - \mathbf{x})$$

Για τον πίνακα 3 x 3 του παραδείγματος, προκύπτει ότι :

$$\mathbf{Ax} = \mathbf{A}\begin{vmatrix} -1 \\ 1 \\ 2 \end{vmatrix} = \begin{vmatrix} -0.55 \\ 0.82 \\ -0.62 \end{vmatrix} \qquad \text{(ακριβής λύση)}$$

$$\mathbf{A}\bar{\mathbf{x}} = \mathbf{A}\begin{vmatrix} -0.853 \\ 0.844 \\ 1.740 \end{vmatrix} = \begin{vmatrix} -0.55 \\ 0.81 \\ -0.601 \end{vmatrix}$$

Προκύπτει ότι τα υπολείμματα **r** δεν είναι πολύ μεγάλα, παρόλο που υπάρχει παθολογική κατάσταση και συνεπώς δεν μπορούν να αποτελέσουν κριτήριο.

Κατάλληλο κριτήριο βρίσκεται με τη βοήθεια της νόρμας πινάκων και διανυσμάτων.

Έστω $\mathbf{e} = \mathbf{x} - \bar{\mathbf{x}}$, το διάνυσμα των διαφορών ανάμεσα στην ακριβή λύση **x** και την προσεγγιστική $\bar{\mathbf{x}}$.

Τότε,

$$\mathbf{r} = \mathbf{A}(\mathbf{x} - \bar{\mathbf{x}}) = \mathbf{Ae} \quad \text{η} \quad \mathbf{e} = \mathbf{A}^{-1}\mathbf{r}$$

Ακολουθούν οι ανισότητες

$$\|\mathbf{r}\| \le \|\mathbf{A}\|.\|\mathbf{e}\|$$

$$\|\mathbf{e}\| \le \|\mathbf{A}^{-1}\|.\|\mathbf{r}\|.$$

Από αυτές προκύπτει :

$$\frac{\|\mathbf{r}\|}{\|\mathbf{A}\|} \le \|\mathbf{e}\| \le \|\mathbf{A}^{-1}\| \cdot \|\mathbf{r}\|$$

Εντελώς ανάλογα αποδεικνύεται ότι

$$\frac{\|\mathbf{b}|}{\|\mathbf{A}\|} \le \|\mathbf{x}\| \le \|\mathbf{A}^{-1}\| \cdot \|\mathbf{b}\|$$

Τελικά,

.
$$\frac{1}{\|\mathbf{A}\|\cdot\|\mathbf{A}^{-1}\|}\frac{\|\mathbf{r}\|}{\|\mathbf{b}\|} \le \frac{\|\mathbf{e}\|}{\|\mathbf{x}\|} \le \|\mathbf{A}\|\cdot\|\mathbf{A}^{-1}\|\frac{\|\mathbf{r}\|}{\|\mathbf{b}\|}$$

Στην παραπάνω ανισότητα σημειώνουμε την ποσότητα ||r||/||b|| που αντιπροσωπεύει το σχετικό υπόλειμμα, επίσης την ποσότητα ||e||/||x|| που αντιπροσωπεύει το σχετικό σφάλμα στη λύση του συστήματος.

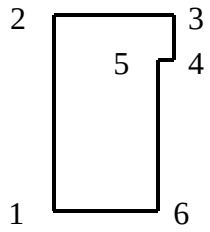
Η ποσότητα $\|\mathbf{A}\|.\|\mathbf{A}^{-1}\|$ μπορεί να οριστεί ως **βαθμός κατάστασης** β.κ. (**condition number**). Πράγματι, αν ο β.κ. είναι μεγάλος, τότε το σχετικό υπόλειμμα δεν αποτελεί μέτρο ενδεικτικό της παθολογικής κατάστασης, όπως φαίνεται από την παραπάνω δεξιά ανισότητα. Αν όμως ο β.κ. είναι κοντά στη μονάδα, τότε το σχετικό υπόλειμμα αποτελεί καλό μέτρο της ευαισθησίας της λύσης, όπως φαίνεται από την παραπάνω ανισότητα.

4.6.1 ΠΑΘΟΛΟΓΙΚΕΣ ΚΑΤΑΣΤΑΣΕΙΣ ΣΕ ΠΡΟΒΛΗΜΑΤΑ ΧΑΡΤΟΓΡΑΦΙΑΣ - ΦΩΤΟΓΡΑΜΜΕΤΡΙΑΣ

Κατά τη διάρκεια ψηφιοποίησης κτιρίων από αναλογικούς χάρτες ή από στερεοζεύγη φωτογραφιών, συμβαίνει συχνά οι ψηφιοποιημένες κορυφές των κτιρίων να μη σχηματίζουν μεταξύ τους ορθές γωνίες, είτε λόγω σφαλμάτων παρατήρησης είτε λόγω τυχαίων σφαλμάτων των χρησιμοποιουμένων οργάνων. Μια κλασική λύση του προβλήματος αυτού είναι η συνόρθωση των παρατηρουμένων συντεταγμένων με τη συνθήκη οι πλευρές να σχηματίζουν μεταξύ τους ορθές γωνίες. Όταν το μήκος μιας πλευράς σε σχέση με τις διπλανές είναι πολύ μικρό τότε ο πίνακας των κανονικών εξισώσεων παρουσιάζουν παθολογική κατάσταση (ill-conditioned).

Παράδειγμα

	ακριβή		**παρατηρούμενα**		**συνορθωμένα**	
	x	**y**	**x**	**y**	**x**	**y**
1	0.00	0.00	0.00	0.00	0.016	0.117
2	0.00	100.00	0.10	100.10	0.071	100.050
3	10.00	100.00	10.00	100.00	9.951	100.044
4	10.00	75.00	9.98	75.02	9.937	75.027
5	9.90	75.00	9.80	74.99	9.937	75.027
6	9.90	0.00	10.00	0.10	9.950	0.111



Λόγω του παθολογικού (ill-conditioned) συστήματος παρατηρούμε ότι μετά τη συνόρθωση το κτίριο αποτελείται από 5 κορυφές και όχι από 6 (!!). Οι συνορθωμένες συντεταγμένες για τις κορυφές 4 και 5 είναι ίδιες (!!). Ο βαθμός κατάστασης στο παραπάνω πρόβλημα ήταν 6.10^8.

5. ΠΑΡΕΜΒΟΛΗ

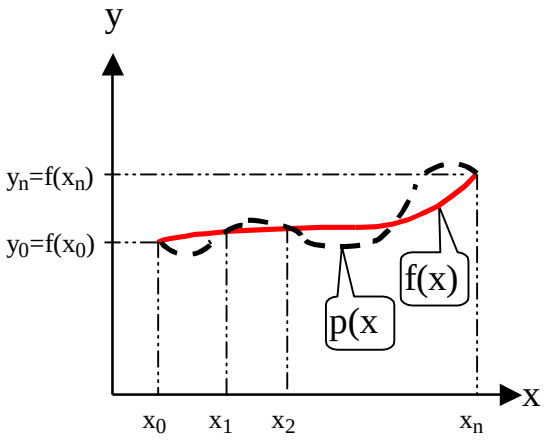
5.1 ΠΟΛΥΩΝΥΜΟ ΠΑΡΕΜΒΟΛΗΣ

Έστω σύνολο $x_0, x_1, \ldots, x_n$, n+1 διακεκριμένων σημείων σε ένα διάστημα I και έστω f(x) συνάρτηση, τέτοια που

$$f(x_i) = f_i,$$

όπου f_i (i = 0, 1, . . ., n) είναι δεδομένες τιμές της συνάρτησης στα σημεία $x_0, x_1, \ldots, x_n$.

Ζητείται ένα πολυώνυμο p(x), n βαθμού, έτσι που $p(x_i) = f_i$, i = 0, 1, . . ., n. Το πολυώνυμο αυτό θα ονομαστεί **πολυώνυμο παρεμβολής (interpolating polynomial)**.



Έστω,

$$p(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n.$$

Πρέπει να ισχύει :

$$\begin{aligned}
f_0 &= a_0 + a_1.x_0 + a_2.x_0^2 + \ldots + a_n.x_0^n \\
f_1 &= a_0 + a_1.x_1 + a_2.x_1^2 + \ldots + a_n.x_1^n \\
&\ldots\ldots \\
&\ldots\ldots \\
f_n &= a_0 + a_1.x_n + a_2.x_n^2 + \ldots + a_n.x_n^n.
\end{aligned} \tag{5.1}$$

Οι συντελεστές του πολυωνύμου, $a_0, a_1, \ldots, a_n$ προσδιορίζονται από το σύστημα (5.1). Για να έχει το σύστημα μοναδική λύση πρέπει και αρκεί η ορίζουσα των συντελεστών των αγνώστων να είναι διάφορη του μηδενός. Η ορίζουσα αυτή είναι :

$$V_n = \begin{pmatrix} 1 & x_0 & x_0^2 & \ldots & x_0^n \\ 1 & x_1 & x_1^2 & \ldots & x_1^n \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & x_n & x_n^2 & \ldots & x_n^n \end{pmatrix}$$

Η ορίζουσα $V_{\mathbf{n}}$ είναι γνωστή ως **ορίζουσα Vandermonde τάξης n**.

Για $x_n = x_0$ δύο σειρές της ορίζουσας είναι ίδιες, άρα $V_n = 0$. Ομοίως για $x_n = x_1, \ldots, x_n = x_{n-1}$. Άρα η ορίζουσα V_n μπορεί να θεωρηθεί πολυώνυμο ως προς x_n το οποίο μηδενίζεται για $x_n = x_{n-1}, x_{n-2}, \ldots, x_0$ και μπορεί να γραφτεί :

$$V_n = A_n.(x_n - x_{n-1}).(x_n - x_{n-2}). \ldots (x_n - x_1).(x_n - x_0).$$

Αν η ορίζουσα V_n αναπτυχθεί κατά τα στοιχεία της n σειράς, τότε είναι προφανές ότι ο συντελεστής του x_n^n είναι ορίζουσα Vandermonde τάξης **n-1**, δηλαδή $A = V_{n-1}$.

Άρα,

$$V_n = V_{n-1}.(x_n - x_{n-1}).(x_n - x_{n-2}). \ldots (x_n - x_1).(x_n - x_0).$$

Συνεπώς,

$V_1 = (x_1 - x_0)$

$V_2 = (x_2 - x_1).\quad (x_2 - x_0).$

$\qquad\qquad\quad (x_1 - x_0).$

$V_3 = (x_3 - x_2).\quad (x_3 - x_1).\quad (x_3 - x_0).$

$\qquad\qquad\quad (x_2 - x_1).\quad (x_2 - x_0).$

$\qquad\qquad\qquad\qquad (x_1 - x_0).$

. . . . . .

. . . . . .

. . . . . .

$V_n = (x_n - x_{n-1}).\ (x_n - x_{n-2}). . .\quad (x_n - x_1).\quad (x_n - x_0)$

$\qquad\qquad\quad (x_{n-1} - x_{n-2}). . .\quad (x_{n-1} - x_1).\ (x_{n-1} - x_0).$

. . . . . .

$\qquad\qquad\qquad\qquad (x_2 - x_1).\quad (x_2 - x_0).$

$\qquad\qquad\qquad\qquad\qquad\qquad (x_1 - x_0)$.

Το σύστημα (5.1) έχει μοναδική λύση αν η ορίζουσα $V_n \neq 0$, δηλαδή αν, $x_i \neq x_j$ για $i, j = 0, 1, . . . ,n$.

Το πολυώνυμο παρεμβολής είναι μοναδικό.

Πράγματι, έστω ότι υπάρχουν δύο διαφορετικά πολυώνυμα $p_1(x)$ και $p_2(x)$, βαθμού n, έτσι ώστε $p_1(x_i) = f_i$ και $p_2(x_i) = f_i$, $i = 0, 1, . . ., n$.

Θεωρούμε τη διαφορά των δύο πολυωνύμων $D(x) = p_1(x) - p_2(x)$. Το $D(x)$ είναι πολυώνυμο βαθμού μικρότερου ή το πολύ ίσου με n.

Για $i = 0, 1, . . . , n$

$$D(x_i) = p_1(x_i) - p_2(x_i) = f_i - f_i = 0.$$

Άρα το D(x) έχει τουλάχιστον n + 1 ρίζες. Και επειδή ο βαθμός του είναι το πολύ n, D(x) είναι εκ ταυτότητος ίσο με 0.

Άρα $p_1(x) \equiv p_2(x)$.

5.2 ΠΟΛΥΩΝΥΜΟ LAGRANGE

Έστω $p(x) = a_0.L_0(x) + a_1.L_1(x) + \ldots + a_n.L_n(x) \quad = \sum_{k=0}^{n} a_k.L_k(x)$.

όπου $$L_k(x) = \prod_{\substack{i=0 \\ i\neq k}}^{n} \frac{x - x_i}{x_k - x_i} \quad , \qquad k = 0, 1, \ldots, n.$$

$$L_k(x) = \frac{(x - x_0).(x - x_1). \;\; . \;\; .(x - x_{k-1}).(x - x_{k+1}). \;\; . \;\; .(x - x_n)}{(x_k - x_0).(x_k - x_1). \;\; . \;\; .(x_k - x_{k-1}).(x_k - x_{k+1}). \;\; . \;\; .(x_k - x_n)} \quad .$$

Η συνάρτηση $L_k(x)$ είναι γινόμενο n γραμμικών παραγόντων, άρα πολυώνυμο βαθμού n. Επομένως το p(x) είναι πολυώνυμο βαθμού ≤ n. Επιπλέον είναι φανερό ότι

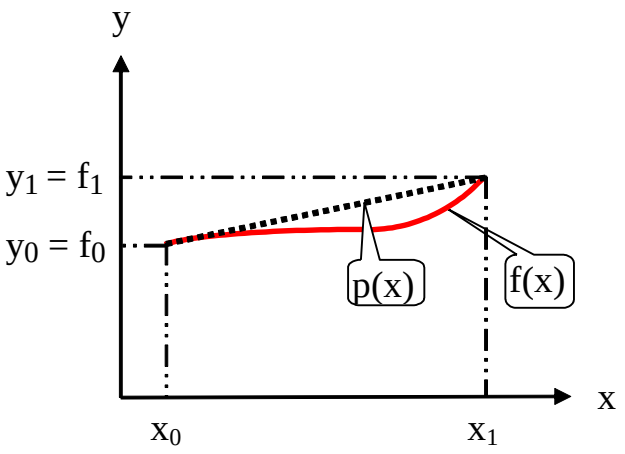
$$L_k(x_i) = \begin{Bmatrix} 1, & i=k \\ 0, & i\neq k \end{Bmatrix}, \quad i = 0, 1, \ldots, n$$

Δηλαδή, $$p(x_i) = \sum_{k=0}^{n} a_k.L_k(x_i) = a_i \quad , \qquad i = 0, 1, \ldots, n.$$

και επομένως οι συντελεστές του πολυωνύμου Lagrange είναι οι τιμές του p(x) στα σημεία $x_0, x_1, \ldots, x_n$. Έτσι

$$p(x) = \sum_{k=0}^{n} f(x_k).L_k(x) = f_0.L_0(x) + f_1.L_1(x) + . \quad . \quad . \quad + f_n.L_n(x) \quad . \qquad (5.2)$$

Ειδική περίπτωση : n = 1. (γραμμική παρεμβολή)



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

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

$$p(x) = f_0.L_0(x) + f_1.L_1(x) = \frac{x - x_1}{x_0 - x_1}.f_0 + \frac{x - x_0}{x_1 - x_0}.f_1 =$$

$$= \frac{(x_1 - x).f_0 - (x_0 - x).f_1}{x_1 - x_0} = \frac{1}{x_1 - x_0}\begin{vmatrix} f_0 & x_0 - x \\ f_1 & x_1 - x \end{vmatrix}.$$

Παράδειγμα

Για μία συνάρτηση f(x) δίνονται τα

f(1) = 1.5708, f(3) = 1.5719, f(5) = 1.5738.

Να προσεγγιστεί η f(3.5) με τη βοήθεια του πολυωνύμου παρεμβολής Lagrange.

Λύση

$$L_0(x) = \frac{(3.5-3).(3.5-5)}{(1-3).(1-5)} = -0.09375$$

$$L_1(x) = \frac{(3.5-1).(3.5-5)}{(3-1).(3-5)} = 0.93750$$

$$L_2(x) = \frac{(3.5-1).(3.5-3)}{(5-1).(5-3)} = 0.15625$$

f(3.5) $\approx$ -0.09375 * 1.5708 + 0.9375 * 1.5719 + 0.15625 * 1.5738

= **1.5723**

5.3 ΣΦΑΛΜΑ ΤΟΥ ΠΟΛΥΩΝΥΜΟΥ ΠΑΡΕΜΒΟΛΗΣ

Έστω f(x) συνάρτηση με $f(x_i) = f_i$, i = 0, 1, . . . , n και p(x) το πολυώνυμο παρεμβολής n βαθμού, τέτοιο που $p(x_i) = f_i$ για i = 0, 1, . . . , n. Επειδή η f(x) και το p(x) συμπίπτουν στα σημεία $x_0, x_1, \ldots, x_n$ η διαφορά τους μπορεί να γραφτεί :

$$f(x) - p(x) = (x - x_0).(x - x_1) \ldots (x - x_n) . K(x) .$$

Για μία συγκεκριμένη τιμή x^* :

$$f(x^*) - p(x^*) - (x^* - x_0).(x^* - x_1) \ldots (x^* - x_n) . K(x^*) = 0. \tag{5.3}$$

Ορίζουμε τη συνάρτηση Φ(x) :

$$\Phi(x) = f(x) - p(x) - (x - x_0).(x - x_1) \ldots (x - x_n) . K(x^*).$$

Παραγωγίζουμε την Φ(x) (n + 1) φορές :

$$\Phi^{(n+1)}(x) = f^{(n+1)}(x) - (n + 1)!\, K(x^*) \qquad \text{(Γιατί ;)}$$

Παρατηρούμε ότι η Φ(x) μηδενίζεται στα σημεία $x_0, x_1, \ldots, x_n$ και x^* δηλαδή (n + 2) φορές.

Άρα η $\Phi'(x)$ έχει τουλάχιστον n + 1 ρίζες,

η $\Phi^{(k)}(x)$ έχει τουλάχιστον n + 2 - k ρίζες,

η $\Phi^{(n+1)}(x)$ έχει τουλάχιστον 1 ρίζα.

Έστω $\bar{x}$ ρίζα της $\Phi^{(n+1)}(x)$, δηλαδή $\Phi^{(n+1)}(\bar{x}) = 0$

Έπεται ότι

$$f^{(n+1)}(\bar{x}) = (n + 1)!\ K(x^*).$$

Από την τελευταία σχέση προσδιορίζεται η σταθερά

$K(x^*) = f^{(n+1)}(\bar{x}) / (n + 1)!$ και αντικαθίσταται στη σχέση (5.3) :

$$f(x^*) - p(x^*) = (x^* - x_0).(x^* - x_1) \ldots (x^* - x_n).\ f^{(n+1)}(\bar{x}) / (n + 1)! \tag{5.4}$$

Επειδή το x* είναι αυθαίρετο, η τελευταία σχέση ισχύει για κάθε x

$$f(x) - p(x) = \frac{(x-x_0).(x-x_1) \ldots (x-x_n)}{(n+1)!} f^{(n+1)}(\bar{x}) \quad . \tag{5.5}$$

Η εξίσωση (5.5) εκφράζει το **σφάλμα του πολυωνύμου παρεμβολής**.

Σαν παράδειγμα μπορεί να εκτιμηθεί ένα άνω φράγμα για τη γραμμική παρεμβολή :

$$p(x) = \frac{(x_1 - x).f_0 - (x_0 - x).f_1}{x_1 - x_0}$$

$$f(x) - p(x) = \frac{(x - x_0).(x - x_1)}{2!}.f''(\xi) \ ,$$

όπου $x_0 < \xi < x_1$.

Αν $f''(\xi) < M$, τότε $f(x) - p(x) \le \frac{|(x - x_0).(x - x_1)|}{2}.M$.

Το μέγιστο του $(x - x_0).(x - x_1)$ πραγματοποιείται για $x = (x_0 + x_1)/2$ και είναι ίσο με $(x_1 - x_0)^2/4$.

Άρα $$f(x) - p(x) \le \frac{(x_1 - x_0)^2}{8}.M \ .$$

Η πρακτική σημασία της σχέσης (5.5) είναι μικρή μια και δεν ξέρουμε ποτέ την $f^{(n+1)}(x)$ και το ξ.

Η θεωρητική της σημασία θα φανεί στον υπολογισμό του σφάλματος ολοκλήρωσης σε τύπους ολοκλήρωσης.

5.4 ΕΠΑΝΑΛΑΜΒΑΝΟΜΕΝΗ ΓΡΑΜΜΙΚΗ ΠΑΡΕΜΒΟΛΗ

Έστω p_{01} το πολυώνυμο 1ου βαθμού που περνά από τα σημεία (x_0, f_0) και (x_1, f_1)

$$p_{01}(x) = \frac{(x_1 - x).f_0 - (x_0 - x).f_1}{x_1 - x_0} = \frac{\begin{vmatrix} f_0 & x_0 - x \\ f_1 & x_1 - x \end{vmatrix}}{x_1 - x_0}$$

και p_{0j} το πολυώνυμο 1ου βαθμού που περνά από τα σημεία (x_0, f_0), (x_j, f_j), για τα οποία πολυώνυμα ισχύει προφανώς $p_{01}(x_0) = p_{0j}(x_0) = f_0$.

Τότε το πολυώνυμο

$$p_{01j}(x) = \frac{(x_j - x).p_{01} - (x_1 - x).p_{0j}}{x_j - x_1}$$

είναι πολυώνυμο παρεμβολής $2^{ου}$ βαθμού για τα σημεία (x_0, f_0), (x_1, f_1) και (x_j, f_j) γιατί όπως φαίνεται

$$p_{01j}(x_0) = \frac{(x_j - x_0).f_0 - (x_1 - x_0).f_0}{x_j - x_1} = f_0$$

και ομοίως $p_{01j}(x_1) = f_1$ και $p_{01j}(x_j) = f_j$.

Γενικότερα, συμβολίζεται με $p_{0,1,\ldots,k}$ το πολυώνυμο παρεμβολής για τα σημεία $x_0, x_1, \ldots, x_{k-1}, x_k$. Τότε

$$p_{0,1,\ldots,k-1,k,j} = \frac{(x_j - x).p_{0,1,\ldots,k-1,k} - (x_k - x).p_{0,1,\ldots,k-1,j}}{x_j - x_k} \quad . \qquad (5.6)$$

($k = 0, 1, \ldots, n-1$ και $j = k+1, \ldots, n$.)

Το $p_{0,1,\ldots,k-1,k,j}$ είναι το πολυώνυμο παρεμβολής για τα σημεία $x_0, x_1, \ldots, x_{k-1}, x_k, x_j$.

Η σχέση (5.6) υποδεικνύει μία επαναληπτική διαδικασία για την εύρεση του πολυωνύμου παρεμβολής. Η διαδικασία αυτή φαίνεται στον παρακάτω πίνακα :

x_j f_j	p_{0j}	p_{01j}	p_{012j}			$x_j - x$
$x_0 p_0$						$x_0 - x$
$x_1 p_1$	p_{01}					$x_1 - x$
$x_2 p_2$	p_{02}	p_{012}				$x_2 - x$
. .	. .	. .				. .
x_j p_j	p_{0j}	p_{01j}	p_{012j}			$x_j - x$
. .	.	. .	. .	. .	. . .	. .
$x_n p_n$	p_{0n}	p_{01n}	p_{012n}	. .	$p_{0,1,2,\ldots,n}$	$x_n - x$

Άσκηση

Δίνεται f(1) = 1.5708, f(3) = 1.5719 και f(5) = 1.5738. Να προσεγγιστεί η f(3.5) με τη μέθοδο της επαναλαμβανόμενης γραμμικής παρεμβολής.

Λύση

$$p_{\cdot} = \frac{\begin{vmatrix} f_{\cdot} & x_{\cdot} - x \\ f_{\cdot} & x_{\cdot} - x \end{vmatrix}}{x_{\cdot} - x_{\cdot}} = \frac{\begin{vmatrix} 1.5708 & -2.5 \\ 1.5719 & -0.5 \end{vmatrix}}{3-1} = 1.57217$$

$$p_{02} = \frac{\begin{vmatrix} 1.5708 & -2.5 \\ 1.5738 & 1.5 \end{vmatrix}}{5-1} = 1.57267$$

$$p_{012} = \frac{\begin{vmatrix} 1.57217 & -0.5 \\ 1.57267 & 1.5 \end{vmatrix}}{2} = 1.57232$$

x_j	f_j	p_{0j}	p_{01j}	x_j - x
1	1.5708			-2.5
3	1.5719	1.57217		-0.5
5	1.5738	1.57267	1.57232	1.5

Με την παραπάνω διαδικασία λύνεται και το πρόβλημα της αντίστροφης παρεμβολής. Έστω f^{-1} η συνάρτηση η αντίστροφη της f(x), δηλαδή $x = f^{-1}(y)$. Η αντίστροφη συνάρτηση υπάρχει αν η y = f(x) είναι μονότιμη και αν η dy/dx δεν μηδενίζεται στο διάστημα που εξετάζουμε.

Άλλος τρόπος επίλυσης είναι να εκφράσουμε την y σαν συνάρτηση του x και να λύσουμε στη συνέχεια με κάποια επαναληπτική διαδικασία το πολυώνυμο που προκύπτει.

Άσκηση

Για τη συνάρτηση y = sin x δίνονται τα ακόλουθα τέσσερα ζεύγη αντιστοίχων τιμών (x_j, y_j).

n	0	1	2	3
y	0.0998	0.1987	0.2955	0.3894
x	0.1	0.2	0.3	0.4

Για y = 0.2800 να εκτιμηθεί η x = arc sin (0.2800).

Λύση

y_j	x_j	x_{0j}	x_{01j}	x_{012j}	y_j - y
0.0998	0.1				-0.1802
0.1987	0.2	0.28220			-0.0813
0.2955	0.3	0.28416	0.28384		0.0155
0.3894	0.4	0.28667	0.28411	0.28379	0.1094

5.5 ΠΟΛΥΩΝΥΜΟ ΠΑΡΕΜΒΟΛΗΣ ΚΑΤΑ NEWTON

Έστω και πάλι $p_n(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$ με

$$p_n(x_i) = f_i, \quad i = 0, 1, 2, \ldots, n.$$

Γράφουμε το πολυώνυμο παρεμβολής με την εξής μορφή :

$p_n(x) = A_0 + A_1.(x - x_0) + A_2.(x - x_0).(x - x_1) + \ldots +$

$+ A_n.(x - x_0).(x - x_1)\ldots(x - x_{n-1}).$

Για n = k

$p_k(x) = A_0 + A_1.(x - x_0) + A_2.(x - x_0).(x - x_1) + \ldots +$

$+ A_k.(x - x_0).(x - x_1)\ldots(x - x_{k-1}),$

Για n = k - 1

$p_{k-1}(x) = A_0 + A_1.(x - x_0) + A_2.(x - x_0).(x - x_1) + \ldots +$

$+ A_{k-1}.(x - x_0).(x - x_1)\ldots(x - x_{k-2}).$

Έπεται ότι,

$$p_k(x) = p_{k-1}(x) + A_k.(x - x_0).(x - x_1)\ldots(x - x_{k-1}).$$

Από την παραπάνω σχέση είναι προφανές ότι για $x = x_0, x_1, \ldots, x_{k-1}$

$$p_k(x_i) = p_{k-1}(x_i).$$

Αλλά,

το $p_k(x)$ είναι το πολυώνυμο παρεμβολής για τα σημεία $x_0, x_1, \ldots, x_k$, δηλαδή $p_k(x_i) = f_i$. Άρα $p_{k-1}(x_i) = f_i$ για $i = 0, 1, \ldots, k-1$.

Συνεπώς το p_{k-1} είναι το πολυώνυμο παρεμβολής για τα $x_0, x_1, \ldots, x_{k-1}$.

Το A_k είναι ο συντελεστής του x^k δηλαδή της μεγαλύτερης δύναμης και συμβολίζεται με

$$A_k = f[x_0, x_1, \ldots, x_k].$$

Επομένως το πολυώνυμο κατά Newton γράφεται

$p_n(x) = f[x_0] + f[x_0, x_1].(x - x_0) + f[x_0, x_1, x_2].(x - x_0).(x - x_1) + \ldots$

$+ f[x_0, x_1, x_2, \ldots, x_n].(x - x_0).(x - x_1)\ldots(x - x_{n-1})$

ή $$p_n(x) = \sum_{i=0}^{n} f[x_0, x_1, x_2, \ldots, x_i] \prod_{j=0}^{i-1}(x - x_j) .$$

Για n = 1

$$p_1(x) = f[x_0] + f[x_0, x_1].(x - x_0).$$

Αλλά είναι ήδη γνωστό ότι

$$p_1(x) = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_0) \ .$$

Άρα $$f[x_0] = f(x_0)$$

και

$$f[x_0, x_1] = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

Γενικά ισχύει η επόμενη σχέση (5.7)

$$f[x_0, x_1, x_2, ..., x_k] = \frac{f[x_1, x_2, x_3, \ldots, x_k] - f[x_0, x_1, x_2, \ldots, x_{k-1}]}{x_k - x_0}$$

Απόδειξη

Συμβολίζεται,

με $p_k(x)$ το πολυώνυμο παρεμβολής για τα σημεία $x_0, x_1, x_2, \ldots, x_k$,

με $p_{k-1}(x)$ το πολυώνυμο παρεμβολής για τα σημεία $x_0, x_1, x_2, \ldots, x_{k-1}$ και

με $q_{k-1}(x)$ το πολυώνυμο παρεμβολής για τα σημεία $x_1, x_2, \ldots, x_k$.

Τότε εύκολα φαίνεται ότι ένα πολυώνυμο $p(x)$

$$p(x) = \frac{x - x_0}{x_k - x_0} q_{k-1}(x) + \frac{x_k - x}{x_k - x_0} p_{k-1}(x) \qquad (5.8)$$

περνά από τα σημεία (x_i, f_i) για $i = 0, 1, 2, \ldots, k-1, k$.

Πράγματι,

Για $x = x_0 \rightarrow$ $p(x_0) = p_{k-1}(x_0) = f_0$.

Για $x = x_i \rightarrow \quad p(x_i) = \dfrac{x_i - x_0}{x_k - x_0} q_{k-1}(x_i) + \dfrac{x_k - x_i}{x_k - x_0} p_{k-1}(x_i) =$

$$= \frac{x_i - x_0}{x_k - x_0} f_i + \frac{x_k - x_i}{x_k - x_0} f_i =$$

$$= \frac{x_i - x_0 + x_k - x_i}{x_k - x_0} f_i = f_i$$

Για $x = x_k \rightarrow \qquad p(x_k) = q_{k-1}(x_k) = f_k$.

Άρα $p(x) \equiv p_k(x)$.

Από την (5.8) προκύπτει ότι :

Συντελεστής του x^k στο $p(x) = \dfrac{1}{x_k - x_0}$ *(συντελεστής του x^{k-1} στο

q_{k-1}) - $\dfrac{1}{x_k - x_0}$ * (συντελεστής του x^{k-1} στο p_{k-1}).

Η τελευταία σχέση είναι ταυτόσημη με την (5.7).

Τώρα, ο υπολογισμός των συντελεστών του πολυωνύμου Newton διατάσσεται σε πίνακα :

	x \ k	(0)	(1)	(2)	(3)	(4)
(d_0)	x_0	$f[x_0]$				
			$f[x_0,x_1]$			
(d_1)	x_1	$f[x_1]$		$f[x_0,x_1,x_2]$		
			$f[x_1,x_2]$		$f[x_0,x_1,x_2,x_3]$	
(d_2)	x_2	$f[x_2]$		$f[x_1,x_2,x_3]$		$f[x_0,x_1,x_2,x_3,x_4]$
			$f[x_2,x_3]$		$f[x_1,x_2,x_3,x_4]$	
(d_3)	x_3	$f[x_3]$		$f[x_2,x_3,x_4]$		
			$f[x_3,x_4]$			
(d_4)	x_4	$f[x_4]$				

Για εξοικονόμηση μνήμης μπορεί να χρησιμοποιηθεί στον προγραμματισμό ένα μόνο διάνυσμα (d_0, d_1, d_2, d_3, d_4) που να μεταφέρεται στις τέσσερις στήλες του πίνακα.

Αρχικά, $d_i = f[x_i]$, $i = 0, 1, 2, 3, 4$. Στη συνέχεια,

k=1	k=2	k=3
$d_4 \leftarrow (d_4 - d_3)/(x_4 - x_3)$	$d_4 \leftarrow (d_4 - d_3)/(x_4 - x_2)$	$d_4 \leftarrow (d_4 - d_3)/(x_4 - x_1)$
$d_3 \leftarrow (d_3 - d_2)/(x_3 - x_2)$	$d_3 \leftarrow (d_3 - d_2)/(x_3 - x_1)$	$d_3 \leftarrow (d_3 - d_2)/(x_3 - x_0)$
$d_2 \leftarrow (d_2 - d_1)/(x_2 - x_1)$	$d_2 \leftarrow (d_2 - d_1)/(x_2 - x_0)$	k=4
$d_1 \leftarrow (d_1 - d_0)/(x_1 - x_0)$		$d_4 \leftarrow (d_4 - d_3)/(x_4 - x_0)$

Μετά τον υπολογισμό των συντελεστών, η τιμή του πολυωνύμου για μία συγκεκριμένη τιμή του x, υπολογίζεται σύμφωνα με το σχήμα διαδοχικού εγκλωβισμού :

$$
\begin{aligned}
p_n(x) &= A_0 + A_1.(x - x_0) + A_2.(x - x_0).(x - x_1) + \ldots + \\
&\quad + A_n.(x - x_0).(x - x_1)\ldots(x - x_{n-1}) = \\
&= A_0 + (x - x_0).\{A_1 + (x - x_1).\{A_2 + (x - x_2).(\ldots \\
&\quad + A_{n-2} + (x - x_{n-2}).\{A_{n-1} + A_n .(x - x_{n-1})\}\}\} \ldots \}.
\end{aligned}
$$

Το σφάλμα στο πολυώνυμο Newton υπολογίζεται ως εξής :

$$\varepsilon_k(x) = f(x) - p_k(x).$$

Για $x = x_0, x_1, \ldots, x_k$

$$\varepsilon_k(x_i) = f(x_i) - p_k(x_i) = 0.$$

Άρα, το $\varepsilon_k(x)$ έχει τουλάχιστον k+1 ρίζες

το $\varepsilon_k'(x)$ έχει τουλάχιστον k ρίζες

το $\varepsilon_k^{(k)}(x)$ έχει τουλάχιστον 1 ρίζα.

Έστω $x = \xi$ η ρίζα αυτή : $\varepsilon_k^{(k)}(\xi) = 0$,

ή $f^{(k)}(\xi) - p_k^{(k)}(\xi) = 0$.

Αλλά $p_k^{(k)}(\xi) = k!\, A_k$, όπου $A_k = f[x_0, x_1, \ldots, x_k]$.

Άρα $f[x_0, x_1, \ldots, x_k].k! = f^{(k)}(\xi)$.

Με τον τρόπο οι πεπερασμένες διαφορές f[ ] συνδέονται με παραγώγους της συνάρτησης f, ως εξής :

$$f[x_0, x_1, \ldots, x_k] = f^{(k)}(\xi) / k!, \tag{5.9}$$

όπου ξ μία τιμή της x μέσα στο διάστημα των τιμών $x_0, x_1, \ldots, x_k$.

Για k = n $\varepsilon_n(x) = f(x) - p_n(x)$, $p_n(x_i) = f_i$, $i = 0, 1, 2, \ldots n$.

Έστω x* διάφορο από τα $x_0, x_1, \ldots, x_n$.

Έστω $p_{n+1}(x)$ το πολυώνυμο παρεμβολής για τα σημεία $x_0, x_1, \ldots, x_n$, x*:

$$p_{n+1}(x) = p_n(x) + f[x_0, x_1, \ldots, x^*].\prod_{j=0}^{n}(x - x_j) \ .$$

Τότε,

$$f(x^*) = p_{n+1}(x^*) = p_n(x^*) + f[x_0, x_1, \ldots, x^*].\prod_{j=0}^{n}(x^* - x_j) \ .$$

$$\varepsilon_n(x^*) = f(x^*) - p_n(x^*) = f[x_0, x_1, \ldots, x^*].\prod_{j=0}^{n}(x^* - x_j) \ .$$

Άρα το σφάλμα στο πολυώνυμο Newton είναι όπως ο τελευταίος όρος. Λαμβάνοντας υπόψη την (5.9) καταλήγουμε στη μορφή

$$f(x^*) - p(x^*) = \frac{f^{(n+1)}(\bar{x})}{(n+1)!} . \prod_{j=0}^{n} (x^* - x_j) \ .$$

όπου $\bar{x}$ σημείο μέσα στο διάστημα των $x_0, x_1, \ldots, x_n, x^*$. Η τελευταία σχέση είναι η ίδια με την εξίσωση (5.4).

Παράδειγμα

Να βρεθεί το πολυώνυμο Newton το οποίο διέρχεται από τα σημεία

x	-4	-1	0	2	5
y	1245	33	5	9	1335

Λύση

x	f				
-4	1245				
		-404			
-1	33		94		
		-28		-14	
0	5		10		3
		2		13	
2	9		88		
		442			
5	1335				

$$f(x) = 1245 - 404(x + 4) + 94(x + 4).(x + 1) - 14(x + 4).(x + 1).x +$$
$$+ 3(x + 4).(x + 1).x.(x - 2) = 3x^4 - 5x^3 + 6x^2 - 14x + 5.$$

5.6 SPLINES

Η συνάρτηση f(x) στο διάστημα I είναι δυνατόν να προσεγγιστεί από n πολυώνυμα. Με τον τρόπο αυτό αποφεύγονται οι ταλαντώσεις που παρουσιάζει ένα μόνο πολυώνυμο παρεμβολής μεταξύ των n+1 διακεκριμένων σημείων $(x_0, f_0), (x_1, f_1), \ldots, (x_n, f_n)$.

Τα πολυώνυμα αυτά ονομάζονται **splines**. Από πρακτική άποψη οι συνηθέστερες είναι οι κυβικές splines. Απαραίτητη προϋπόθεση να είναι συνεχείς, με συνεχείς πρώτες και δεύτερες παραγώγους.

Η κυβική καμπύλη (spline) για το i διάστημα μεταξύ των σημείων (x_i, f_i) και (x_{i+1}, f_{i+1}) μπορεί να γραφτεί

$$p(x) = a_i.(x - x_i)^3 + b_i.(x - x_i)^2 + c_i.(x - x_i) + d_i \qquad (5.10)$$

και

$$p(x_i) = f_i = a_i.(x_i - x_i)^3 + b_i.(x_i - x_i)^2 + c_i.(x_i - x_i) + d_i = d_i$$

(5.11)

$$p(x_{i+1}) = f_{i+1} = a_i.(x_{i+1} - x_i)^3 + b_i.(x_{i+1} - x_i)^2 + c_i.(x_{i+1} - x_i) + d_i \quad \Rightarrow$$

$$p(x_{i+1}) = f_{i+1} = a_i.h_i^3 + b_i.h_i^2 + c_i.h_i + d_i, \text{ όπου } h_i = x_{i+1} - x_i.$$

(5.12)

Η πρώτη και η δεύτερη παράγωγος του p(x) είναι

$$p'(x)= 3a_i.(x - x_i)^2 + 2b_i.(x - x_i) + c_i\ , \qquad (5.13)$$

$$p''(x)= 6a_i.(x - x_i) + 2b_i\ . \qquad (5.14)$$

Αν συμβολίσουμε με S_i και S_{i+1} την τιμή της δεύτερης παραγώγου στα σημεία (x_i, f_i) και (x_{i+1}, f_{i+1}) αντίστοιχα, η (5.14) γίνεται :

$$S_i = 6a_i.(x_i - x_i) + 2b_i = 2b_i$$

$$S_{i+1} = 6a_i.(x_{i+1} - x_i) + 2b_i = 6a_i.h_i + 2b_i \ .$$

Έτσι οι συντελεστές b_i, a_i εκφράζονται ως συναρτήσεις των S_i και S_{i+1}

$$b_i = S_i/2 \quad , \tag{5.15}$$

$$a_i = (S_{i+1} - S_i) / 6h_i \ . \tag{5.16}$$

Αντικατάσταση των a_i, b_i, d_i από τις σχέσεις (5.11), (5.15) και (5.16) στην (5.12) οδηγεί στην παρακάτω παράσταση για το c_i :

$$p(x_{i+1}) = \frac{S_{i+1} - S_i}{6h_i} h_i^3 + \frac{S_i}{2} h_i^2 + c_i.h_i + p(x_i) \ .$$

Άρα,

$$c_i = \frac{p(x_{i+1}) - p(x_i)}{h_i} - \frac{2h_i.S_i + h_i.S_{i+1}}{6} = \frac{f_{i+1} - f_i}{h_i} - \frac{2h_i.S_i + h_i.S_{i+1}}{6} \ .$$

Η κλίση και η καμπυλότητα για ένα ζεύγος splines, στο κοινό τους σημείο είναι η ίδια.

Η κλίση της $p(x)$ στο σημείο x_i δίνεται από την (5.13)

$$p'(x_i) = 3a_i.(x_i - x_i)^2 + 2b_i.(x_i - x_i) + c_i = c_i \ .$$

Αν θεωρηθεί το προηγούμενο διάστημα $[x_{i-1}, x_i]$, η κλίση στο x_i δίνεται ομοίως

$$p'(x_i) = 3a_{i-1}.(x_i - x_{i-1})^2 + 2b_{i-1}.(x_i - x_{i-1}) + c_{i-1} =$$

$$= 3a_{i-1}.h_{i-1}^2 + 2b_{i-1}.h_{i-1} + c_{i-1} \ .$$

Εξισώνοντας τις παραπάνω σχέσεις και αντικαθιστώντας τα a_i, b_i, c_i, d_i συναρτήσει των S και p

$$c_i = p'(x_i) = \frac{p(x_{i+1}) - p(x_i)}{h_i} - \frac{2h_i.S_i + h_i.S_{i+1}}{6} =$$

$$3\frac{S_i - S_{i-1}}{6h_{i-1}}.h_{i-1}{}^2 + 2\frac{S_{i-1}}{2}.h_{i-1} + \frac{p(x_i) - p(x_{i-1})}{h_{i-1}} - \frac{2h_{i-1}.S_{i-1} + h_{i-1}.S_i}{6} .$$

Απλοποίηση της σχέσης οδηγεί στην εξίσωση

$$h_{i-1}.S_{i-1} + (2h_{i-1} + 2h_i).S_i + h_i.S_{i+1} = 6\left\{\frac{f_{i+1} - f_i}{h_i} - \frac{f_i - f_{i-1}}{h_{i-1}}\right\} \quad (5.17)$$

Η εξίσωση (5.17) εφαρμόζεται για κάθε ένα εσωτερικό σημείο από **i = 1** έως **i = n - 1**. Αυτό οδηγεί σε ένα σύστημα n - 1 εξισώσεων για τις n + 1 τιμές της S_i. Δύο επιπλέον εξισώσεις μπορούν να προσδιοριστούν με βάση τις συνθήκες στα ακραία σημεία 0, n.

1) $S_0 = S_n = 0$. (Οι ακραίες καμπύλες είναι γραμμικές στα όριά τους).(**φυσική ή ελεύθερη καμπύλη**).

2) $S_0 = S_1$, $S_n = S_{n-1}$. (Οι ακραίες καμπύλες προσεγγίζουν την παραβολή).

3) Η S_0 είναι γραμμική παρεμβολή των S_1, S_2 και η S_n γραμμική παρεμβολή των S_{n-1} και S_{n-2}.

Οι εξισώσεις (6.8) μπορούν να γραφτούν σε μορφή πίνακα ως εξής :

$$\begin{bmatrix} h_0 & 2(h_0+h_1) & h_1 & 0 & . & . & 0 \\ 0 & h_1 & 2(h_1+h_2) & h_2 & 0 & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & . & . & 0 & h_{n-2} & 2(h_{n-2}+h_{n-1}) & h_{n-1} \end{bmatrix} \begin{bmatrix} S_0 \\ S_1 \\ . \\ . \\ . \\ S_{n-1} \\ S_n \end{bmatrix}$$

$$= 6 \begin{bmatrix} \frac{f_2 - f_1}{h_1} - \frac{f_1 - f_0}{h_0} \\ \frac{f_3 - f_2}{h_2} - \frac{f_2 - f_1}{h_1} \\ . \\ . \\ \frac{f_n - f_{n-1}}{h_{n-1}} - \frac{f_{n-1} - f_{n-2}}{h_{n-2}} \end{bmatrix}$$

Οι άγνωστοι S_0, S_n μπορούν να απαλειφθούν, με βάση τις συνθήκες στα ακραία σημεία και το σύστημα των εξισώσεων γίνεται (n - 1) x (n - 1). Το μητρώο των συντελεστών των αγνώστων είναι τριδιαγωνικό. Το σύστημα επιλύεται με τη βοήθεια του αλγόριθμου Thomas (κεφ. 4.2). Μετά τα S_i, μπορούν να υπολογιστούν οι τιμές των a, b_i, c_i, d_i.

$$a_i = \frac{S_{i+1} - S_i}{6h_i}, \qquad b_i = \frac{S_i}{2} \tag{5.18}$$

$$c_i = \frac{f_{i+1} - f_i}{h_i} - \frac{2h_i.S_i + h_i.S_{i+1}}{6}, \qquad d_i = f_i$$

Μια εναλλακτική διατύπωση της παρεμβολής με splines, η οποία δεν απαιτεί την επίλυση συστήματος είναι η εξής:

Έστω ότι, όπως και προηγουμένως, έχουμε τα n+1 ζεύγη τιμών (x_i, f_i), i = 0, 1, 2,…, n-1, n και υποθέτουμε ότι για κάθε ένα από τα διαστήματα

$x_i \le x \le x_{i+1}$, i = 0, 1, 2,...., n-1

το αντίστοιχο πολυώνυμο spline θα δίνεται από την εξίσωση (5.10).
Για όλα τα εσωτερικά σημεία i = 1, 2, …., n-1 εξισώνονται, εκτός από τις τεταγμένες, οι πρώτες και οι δεύτερες παράγωγοι στα σημεία αυτά. Με τον τρόπο αυτό προκύπτουν για i = 1, 2,…., n-1 οι εξής εξισώσεις:

$$a_{i-1}h_{i-1}^3 + b_{i-1}h_{i-1}^2 + c_{i-1}h_{i-1} = f_i - f_{i-1}$$

$$3a_{i-1}h_{i-1}^2 + 2b_{i-1}h_{i-1} + c_{i-1} = c_i$$

$$3a_{i-1}h_{i-1} + b_{i-1} = b_i$$

Το σύστημα των παραπάνω τριών εξισώσεων λύνεται ως προς α_{i-1}, b_{i-1}, c_{i-1}:

$$a_{i-1} = \frac{b_i}{h_{i-1}} - \frac{c_i}{{h_{i-1}}^2} + \frac{f_i - f_{i-1}}{{h_{i-1}}^3}$$

$$b_{i-1} = -2b_i + \frac{3c_i}{h_{i-1}} - \frac{3(f_i - f_{i-1})}{{h_{i-1}}^2}$$

$$c_{i-1} = h_{i-1}b_i - 2c_i + \frac{3(f_i - f_{i-1})}{h_{i-1}}$$

Οι τελευταίες τρεις εξισώσεις προφανώς είναι της μορφής:

$$\left.\begin{aligned} a_{i-1} &= k_i^{(a)} b_i + \lambda_i^{(a)} c_i + \mu_i^{(a)} \\ b_{i-1} &= k_i^{(b)} b_i + \lambda_i^{(b)} c_i + \mu_i^{(b)} \\ c_{i-1} &= k_i^{(c)} b_i + \lambda_i^{(c)} c_i + \mu_i^{(c)} \end{aligned}\right\} \qquad (5.19)$$

Με συμβολισμό πινάκων, οι δύο τελευταίες εξισώσεις γράφονται:

$$\mathbf{b_{i-1} = K_i b_i + \mu_i}$$

όπου

$$\mathbf{b_i} = \begin{pmatrix} b_i \\ c_i \end{pmatrix}, \qquad \mathbf{K}_i = \begin{pmatrix} k_i^{(b)} & \lambda_i^{(b)} \\ k_i^{(c)} & \lambda_i^{(c)} \end{pmatrix}, \qquad \boldsymbol{\mu}_i = \begin{pmatrix} \mu_i^{(b)} \\ \mu_i^{(c)} \end{pmatrix}$$

Επαγωγικά προκύπτει ότι

$$\begin{aligned} \mathbf{b_0} = {} & \mathbf{K_1 K_2 K_3 \ldots\ldots K_{n-1} b_{n-1}} + \\ & \mathbf{K_1 K_2 \ldots K_{n-2}\, \mu_{n-1}} + \\ & \mathbf{K_1 .. K_{n-3}\, \mu_{n-2}} + \\ & \mathbf{\ldots\ldots\ldots\ldots\ldots\ldots} + \\ & \mathbf{K_1 K_2 \mu_3} + \\ & \mathbf{K_1 \mu_2} + \\ & \boldsymbol{\mu_1} \end{aligned}$$

Επομένως το διάνυσμα $\mathbf{b}_0$ είναι της μορφής

$$\mathbf{b}_0 = \mathbf{L}\mathbf{b}_{n-1} + \mathbf{m}$$

όπου

$$\mathbf{L} = \mathbf{K}_1 \mathbf{K}_2 \mathbf{K}_3 \ldots\ldots \mathbf{K}_{n-1}$$

και

$$
\begin{aligned}
\mathbf{m} = \quad & \mathbf{K}_1 K_2 ... \mathbf{K}_{n-2}\, \boldsymbol{\mu}_{n-1} + \\
& \mathbf{K}_1 .. \mathbf{K}_{n-3}\, \boldsymbol{\mu}_{n-2} + \\
& ..... + \\
& \mathbf{K}_1 \mathbf{K}_2 \boldsymbol{\mu}_3 + \\
& \mathbf{K}_1 \boldsymbol{\mu}_2 + \\
& \boldsymbol{\mu}_1
\end{aligned}
$$

Ο πίνακας **L** και το διάνυσμα **m** γράφονται ως εξής:

$$
\mathbf{L} = \begin{pmatrix} l_1^{(b)} & l_2^{(b)} \\ l_1^{(c)} & l_2^{(c)} \end{pmatrix}, \qquad \mathbf{m} = \begin{pmatrix} m^{(b)} \\ m^{(c)} \end{pmatrix}
$$

Υπενθυμίζουμε εδώ ότι

$$
\mathbf{b}_0 = \begin{pmatrix} b_0 \\ c_0 \end{pmatrix}
$$

Επομένως ο συντελεστής b_0 προκύπτει ως γραμμική συνάρτηση των b_{n-1} και c_{n-1}:

$$
b_0 = l_1^{(b)} b_{n-1} + l_2^{(b)} c_{n-1} + m^{(b)} \tag{5.20}
$$

Σχετικά τώρα με την τιμή του συντελεστή b_0, θεωρούμε το spline $p_0(x)$ (εξίσωση. 5.10) και επιβάλλουμε τη συνθήκη

$$
p_0''(x_0) = 0
$$

Η συνθήκη αυτή αντιστοιχεί με τη συνθήκη $S_0 = 0$ της προηγούμενης παραγράφου.

Αλλά $p_0''(x_0) = 2b_0$

Συνεπώς, $b_0 = 0$ και έτσι, από την εξίσωση (5.20), προκύπτει η παρακάτω εξίσωση για τα b_{n-1} και c_{n-1}:

$$l_1^{(b)} b_{n-1} + l_2^{(b)} c_{n-1} = -m^{(b)} \tag{5.21}$$

Σχετικά με τους συντελεστές b_n και c_n, χρησιμοποιούνται οι εξής συνθήκες:

$$p_{n-1}(x_n) = f_n \qquad \text{και} \qquad p''_{n-1}(x_n) = 0$$

Η πρώτη συνθήκη αφορά στην εξίσωση των τεταγμένων και η δεύτερη αντιστοιχεί με τη συνθήκη $S_n = 0$ της προηγούμενης παραγράφου.

Από τις συνθήκες αυτές, σε συνδυασμό με την (5.10), προκύπτουν οι εξισώσεις:

$$a_{n-1} h_{n-1}^3 + b_{n-1} h_{n-1}^2 + c_{n-1} h_{n-1} = f_n - f_{n-1} \tag{5.22}$$

$$3a_{n-1} h_{n-1} + b_{n-1} = 0 \tag{5.23}$$

Από τις δύο τελευταίες εξισώσεις (5.22) και (5.23) απαλείφεται το α_{n-1}:

$$\frac{4}{3} h_{n-1} b_{n-1} + c_{n-1} = \frac{f_n - f_n - 1}{h_{n-1}} \tag{5.24}$$

Οι εξισώσεις (5.21) και (5.24) οδηγούν στον προσδιορισμό των b_n και c_n και η εξίσωση (5.23) στον προσδιορισμό του α_n. Οι υπόλοιποι συντελεστές α_i, b_i, c_i για i=n-2, n-3, ….0 υπολογίζονται απευθείας με τη βοήθεια των αναδρομικών σχέσεων (5.19).

Άσκηση

Τα δεδομένα του παρακάτω πίνακα αποτελούν μεταβολές του φαινόμενου ύψους ενός αστέρα συναρτήσει του χρόνου. Η συνθήκη στα όρια είναι

$S_0 = S_1, \qquad S_n = S_{n-1}.$

n	0	1	2	3	4	5	6	7	8
χρόνος	.0	.2	.3	.4	.5	.6	.7	.8	1.
φαινόμενο ύψος	.302	.185	.106	.093	.240	.579	.561	.468	.302

Λύση

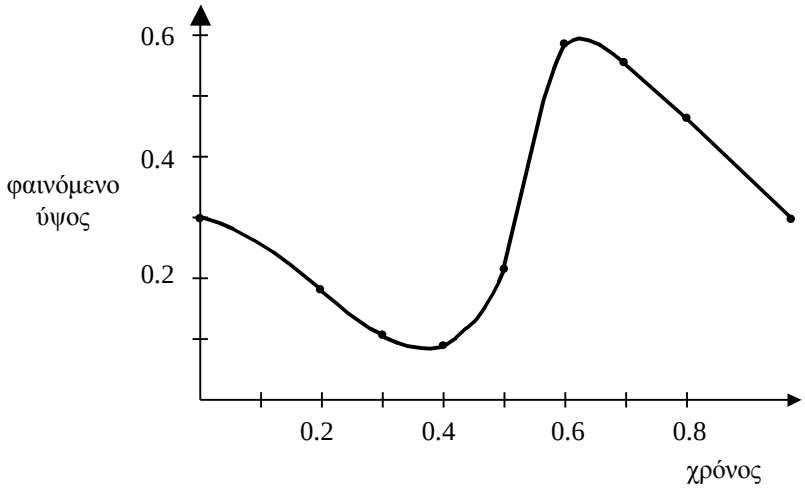
Το σύστημα των εξισώσεων για το παράδειγμα μπορεί να γραφτεί με μορφή πίνακα ως εξής :

$$\begin{bmatrix} .2 & .6 & .1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .1 & .4 & .1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & .1 & .4 & .1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & .1 & .4 & .1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & .1 & .4 & .1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & .1 & .4 & .1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & .1 & .6 & .2 \end{bmatrix} \begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \\ S_6 \\ S_7 \\ S_8 \end{bmatrix} = \begin{bmatrix} -1.23 \\ 3.96 \\ 9.60 \\ 11.52 \\ -21.42 \\ -4.50 \\ 0.60 \end{bmatrix}$$

Το σύστημα αυτό τροποποιείται με βάση τις οριακές συνθήκες $S_0 = S_1$ και $S_8 = S_7$ και γίνεται :

$$\begin{bmatrix} .8 & .1 & 0 & 0 & 0 & 0 & 0 \\ .1 & .4 & .1 & 0 & 0 & 0 & 0 \\ 0 & .1 & .4 & .1 & 0 & 0 & 0 \\ 0 & 0 & .1 & .4 & .1 & 0 & 0 \\ 0 & 0 & 0 & .1 & .4 & .1 & 0 \\ 0 & 0 & 0 & 0 & .1 & .4 & .1 \\ 0 & 0 & 0 & 0 & 0 & .1 & .8 \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \\ S_6 \\ S_7 \end{bmatrix} = \begin{bmatrix} -1.23 \\ 3.96 \\ 9.60 \\ 11.52 \\ -21.42 \\ -4.50 \\ 0.60 \end{bmatrix}$$

Η λύση του παραπάνω συστήματος εξισώσεων οδηγεί στον υπολογισμό των τιμών $S_1, S_2, \ldots, S_7$ και στη συνέχεια, των κυβικών πολυωνύμων από τις σχέσεις (5.18).



Άσκηση

Να υπολογιστεί η καμπύλη 3^{ου} βαθμού (spline) η οποία διέρχεται από τα παρακάτω σημεία.

n	**0**	**1**	**2**	**3**	**4**
x	0	1	2	3	4
f	-8	-7	0	19	56

Λύση

α) Για οριακή συνθήκη $S_0 = 0$, $S_4 = 0$, το σύστημα των εξισώσεων γράφεται

$$\begin{bmatrix} 4 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} 36 \\ 72 \\ 108 \end{bmatrix}.$$

Από τη λύση του συστήματος έχουμε

$$S_0 = 0,\ S_1 = 6.4285,\ S_2 = 10.2857,\ S_3 = 24.4285,\ S_4 = 0.$$

Επομένως η καμπύλη η οποία διέρχεται από τα σημεία π.χ. 0 και 1 υπολογίζεται (i = 0)

$a_0 = (S_1 - S_0)/6h_0 = 1.0714,$ $\qquad b_0 = S_0/2 = 0$

$$c_0 = \frac{f_1 - f_0}{h_0} - \frac{2h_0.S_0 + h_0.S_1}{6} = \{-7-(-8)\}/1- (2*1*0+1*6.4285)/6 = 2.0714$$

$d_0 = p(x_0) = -8$

$$p(x) = 1.0714x^3 + 2.0714x - 8$$

α) Για οριακή συνθήκη $S_0 = S_1$, $S_3 = S_4$ το σύστημα των εξισώσεων γράφεται

$$\begin{bmatrix} 5 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 5 \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} 36 \\ 72 \\ 108 \end{bmatrix} .$$

Από τη λύση του συστήματος έχουμε

$$S_0 = 4.80,\ S_1 = 4.80,\ S_2 = 12,\ S_3 = 19.20,\ S_4 = 19.20 \ .$$

6. ΑΡΙΘΜΗΤΙΚΗ ΟΛΟΚΛΗΡΩΣΗ

6.1 ΣΤΟΙΧΕΙΩΔΕΙΣ ΜΕΘΟΔΟΙ

Με την αριθμητική ολοκλήρωση επιδιώκεται ο προσεγγιστικός υπολογισμός του ολοκληρώματος

$$\int_a^b f(x)dx \ .$$

Η συνάρτηση f(x) αντικαθίσταται από ένα πολυώνυμο στο διάστημα [a, b] και υπολογίζεται το ολοκλήρωμα του πολυωνύμου στο διάστημα αυτό. Μία μορφή πολυωνύμου που χρησιμοποιείται για τον σκοπό αυτό είναι το πολυώνυμο παρεμβολής κατά Newton :

$$p_n(x) = f[x_0] + f[x_0, x_1].(x - x_0) + \ldots$$
$$+ f[x_0, x_1, \ldots, x_n].(x - x_0).(x - x_1)\ldots(x - x_{n-1}) \ .$$

Για την παρακάτω ανάλυση υποτίθεται ότι τα σημεία $x_0, x_1, \ldots, x_n$ ισαπέχουν κατά ένα σταθερό βήμα h, δηλαδή

$x_{i+1} - x_i = h$ για i = 1, 2, . . ., n-1. Αποδεικνύεται ότι

$$f[x_0, x_1, \ldots, x_n] = \frac{\Delta^n f(x_0)}{n!h^n}, \qquad (6.1)$$

όπου ο **τελεστής (operator) Δ** ορίζεται

$\Delta^0 f(x_0) = f(x_0)$

$\Delta^1 f(x_0) = f(x_1) - f(x_0)$

$\Delta^2 f(x_0) = \Delta^1 f(x_1) - \Delta^1 f(x_0) = f(x_2) - 2f(x_1) + f(x_0)$

. . . . . .

$\Delta^k f(x_0) = \Delta^{k-1} f(x_1) - \Delta^{k-1} f(x_0) =$

$$= f(x_k) - \binom{k}{1} f(x_{k-1}) + \binom{k}{2} f(x_{k-2}) - \ldots + (-1)^k f(x_0).$$

Στην τελευταία σχέση τα $\binom{k}{1}, \binom{k}{2}, \ldots$

είναι οι συντελεστές του διωνύμου του Newton και ορίζονται ως εξής :

$$\binom{x}{0} = 1,$$

$$\binom{x}{r} = \frac{x.(x-1)\ldots(x-r+1)}{r!} \qquad \text{για x πραγματικό, r ακέραιο} > 0$$

$$\binom{x}{r} = 0 \qquad \text{για x πραγματικό, r ακέραιο} < 0.$$

Η σχέση (6.1) θα αποδειχτεί επαγωγικά.

Για n = 0,

$f[x_0] = f(x_0) = \Delta^0 f(x_0)$.

Για n = 1,

$f[x_0, x_1] = \{f(x_1) - f(x_0)\}/(x_1 - x_0) = \dfrac{\Delta^1 f(x_0)}{h}$.

Υποτίθεται ότι η εξίσωση (6.1) ισχύει για n = k - 1 και θα αποδειχτεί ότι ισχύει για n = k. Από την εξίσωση (5.7)

$$f[x_0, x_1, \ldots, x_k] = \frac{f[x_1, x_2, x_3, \ldots, x_k] - f[x_0, x_1, x_2, \ldots, x_{k-1}]}{x_k - x_0} =$$

$$= \frac{\Delta^{k-1} f(x_1) - \Delta^{k-1} f(x_0)}{(x_k - x_0).(k-1)!h^{k-1}} = \frac{\Delta^k f(x_0)}{k!h^k}$$

$(x_k - x_0 = k.h)$.

Συνεπώς η (6.1) ισχύει για κάθε ακέραιο n.

Το πολυώνυμο Newton γράφεται κατά συνέπεια ως εξής :

$$p_n(x) = \sum_{i=0}^{n} f[x_0, x_1, \ldots, x_i] \prod_{j=0}^{i-1} (x - x_j) = \sum_{i=0}^{n} \frac{\Delta^i f(x_0)}{i!h^i} \prod_{j=0}^{i-1} (x - x_j) \ .$$

Η ισαποχή των σημείων x_i, i = 0, 1, . . ., n, υποδεικνύει την αλλαγή μεταβλητής :

$$x = x_0 + s \cdot h.$$

Συνεπώς, $x_j = x_0 + j \cdot h$ και $x - x_j = (s - j) \cdot h$

Άρα,

$$\frac{1}{i!h^i}\prod_{j=0}^{i-1}(x - x_j) = \prod_{j=0}^{i-1}\frac{s-j}{j+1} = \frac{s.(s-1).(s-2)...(s-i+1)}{1.2.3...i} = \binom{s}{i}.$$

Τελικά το πολυώνυμο Newton γράφεται με τη μορφή

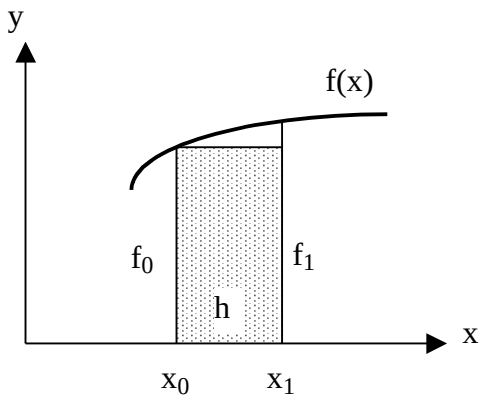
$$p_n(x) = p_n(x_0 + s.h) = \sum_{i=0}^{n}\Delta^i f_0.\binom{s}{i} = f_0 + s.\Delta f_0 + \binom{s}{2}.\Delta^2 f_0 + . . . + \binom{s}{n}.\Delta^n f_0 .$$

Για την εκτέλεση της αριθμητικής ολοκλήρωσης από x_0 έως x_1,

$$\int_{x_0}^{x_1} f(x)\,dx \approx \int_{x_0}^{x_1} p_n(x)dx = h\int_0^1 p_n(s)ds .$$

Για n = 0, $p_n(s) = f_0$, $\int_{x_0}^{x_1} f(x)\,dx \approx h\int_0^1 f_0\,ds = h.f_0$.

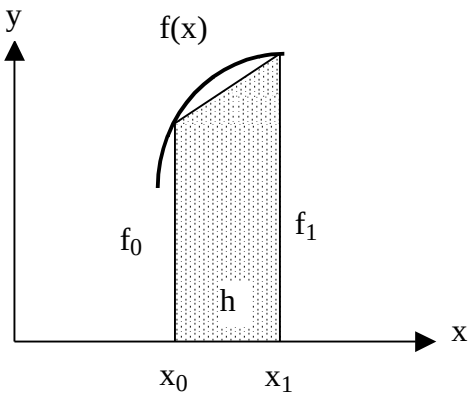
Ο παραπάνω τύπος εκφράζει τον "**ορθογώνιο**" κανόνα ολοκλήρωσης και ερμηνεύεται γεωμετρικά στο σχήμα.



Για n = 1, $p_n(s) = f_0 + s\,.\Delta f_0$,

$$\int_{x_0}^{x_1} f(x)\,dx \approx h\int_0^1 (f_0 + s.\Delta f_0)\,ds = \frac{h}{2}.(f_0 + f_1)\,.$$

Ο παραπάνω τύπος εκφράζει τον "**τραπεζοειδή**" κανόνα ολοκλήρωσης.



Για να υπολογιστεί το σφάλμα στην αριθμητική ολοκλήρωση, γράφεται πρώτα το σφάλμα του πολυωνύμου παρεμβολής (5.5):

$$f(x) - p(x) = \psi(x)\,.\,f^{(n+1)}(\xi) / (n+1)! \qquad (6.2)$$

όπου $\psi(x) = (x - x_0).(x - x_1)\,.\,.\,.\,(x - x_n)$ και η ξ μία τιμή που βρίσκεται ανάμεσα στον μεγαλύτερο και τον μικρότερο από τους αριθμούς $x_0, x_1, \ldots, x_n, x$. Είναι δηλαδή η ξ συνάρτηση του x, $\xi = \xi(x)$.

Η εξίσωση (6.2) ολοκληρώνεται :

$$\int_a^b f(x)\,dx - \int_a^b p(x)\,dx = \int_a^b \frac{\psi(x)\,f^{(n+1)}(\xi)}{(n+1)!}dx\ .$$

Αν η ψ(x) δεν αλλάζει πρόσημο στο διάστημα [a, b] τότε μπορεί να εφαρμοστεί το σχετικό θεώρημα μέσης τιμής του ολοκληρωτικού λογισμού :

$$\int_a^b \psi(x)\,.\varphi(x)\,dx = \varphi(\eta)\int_a^b \psi(x)\,dx \qquad , \qquad \text{όπου } a < \eta < b.$$

Αν η ολοκλήρωση εκτείνεται ανάμεσα σε δύο διαδοχικά σημεία x_i και x_{i+1}, τότε η ψ(x) δεν αλλάζει πρόσημο και το σφάλμα της αριθμητικής ολοκλήρωσης γράφεται :

$$\int_a^b f(x)\,dx - \int_a^b p(x)\,dx = \frac{f^{(n+1)}(\eta)}{(n+1)!}\int_a^b \psi(x)\,dx \ . \qquad (6.3)$$

Προκύπτει κατά συνέπεια το παρακάτω θεώρημα :

Αν η f(x) έχει συνεχείς παραγώγους n+1 τουλάχιστον και αν το πρόσημο της ψ(x) δεν αλλάζει μέσα στο διάστημα από a ως b, τότε το σφάλμα της αριθμητικής ολοκλήρωσης δίνεται από την εξίσωση (6.3).

Συγκεκριμένα, για τον ορθογώνιο τύπο (n = 0) το σφάλμα δίνεται από την σχέση :

$$E_1^R = f'(\eta_0)\int_{x_0}^{x_1}(x - x_0)\,dx = f'(\eta_0)\int_0^1 (s.h)\,h.ds = \frac{f'(\eta_0).h^2}{2} \ ,$$

όπου $x_0 < \eta_0 < x_1$.

Αντίστοιχα για τον τραπεζοειδή τύπο (n = 1)

$$E_1^T = \frac{f''(\eta_0)}{2!}\int_{x_0}^{x_1}(x - x_0).(x - x_1)\,dx = \frac{f''(\eta_0)}{2!}\int_0^1 (s.h)\,[(s - 1).h]\,h.ds = \frac{f''(\eta_0).h^3}{12} \ ,$$

όπου $x_0 < \eta_0 < x_1$.

Οι παραπάνω εκφράσεις αποτελούν τα τοπικά σφάλματα των τύπων ολοκλήρωσης. Για να εφαρμοστεί η αριθμητική ολοκλήρωση, δίνεται ένα διάστημα [a, b] και αυτό χωρίζεται σε υποδιαστήματα, σε κάθε ένα από τα οποία εφαρμόζεται ένας τύπος ολοκλήρωσης. Το πρόβλημα συνεπώς είναι να προσδιοριστεί το ολικό σφάλμα σε ένα διάστημα ολοκλήρωσης [a, b] που υποδιαιρείται σε N ίσα υποδιάστηματα.

Έστω

$$h = (b - a)/N.$$

Τότε, εφαρμογή του **ορθογώνιου τύπου** σε κάθε υποδιάστημα δίνει την εξής προσέγγιση για το ολοκλήρωμα

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

$$R_N = h.(f_0 + f_1 + \ldots + f_{N-1}).$$

Σύμφωνα με τα παραπάνω το συνολικό σφάλμα είναι

$$E_N^R = f'(\eta_0) + f'(\eta_1) + \ldots + f'(\eta_{N-1}) \cdot \frac{h^2}{2}$$

ή

$$E_N^R = \frac{h^2}{2} N.f'(\eta) = \frac{b-a}{2} h.f'(\eta) , \qquad a < \eta < b.$$

Αντίστοιχα για τον τραπεζοειδή κανόνα προκύπτει ότι

$$T_N = h.(f_0/2 + f_1 + f_2 + \ldots + f_{N-1} + f_N/2) .$$

$$E_N^T = -\frac{b-a}{12} h^2 f''(\eta) , \qquad a < \eta < b.$$

Από τις παραπάνω εκφράσεις φαίνεται εύκολα ότι ο ορθογώνιος και ο τραπεζοειδής κανόνας απαιτούν μεγάλο αριθμό υποδιαιρέσεων για να δώσουν ικανοποιητική ακρίβεια.

7.2 ΚΑΝΟΝΑΣ ΤΟΥ SIMPSON

Αν η τοπική ολοκλήρωση γίνει στο διάστημα (x_0, x_2) τότε n = 2 και

$$\int_{x_0}^{x_2} f(x)dx = \int_0^2 p(s)\, ds = h\int_0^2 \left\{ f_0 + s.\Delta f_0 + \binom{s}{2}.\Delta^2 f_0 + \binom{s}{3}.\Delta^3 f_0 + \ldots \right\} ds.$$

$$S_2 = h.(2f_0 + 2\Delta f_0 + 1/3\ \Delta^2 f_0 + \varnothing\ \Delta^3 f_0) .$$

$S_2 = (h/3).(f_2 + 4f_1 + f_0)$.

Για το διάστημα [a, b] πρέπει ο αριθμός **N** των υποδιαιρέσεων να είναι **άρτιος.** Έτσι τελικά,

$$S = (h/3) \,.\, (f_0 + 4f_1 + 2f_2 + 4f_3 + \ldots + 4f_{N-1} + f_N) \,.$$

Για την εκτίμηση του τοπικού σφάλματος του κανόνα του Simpson δεν μπορεί να εφαρμοστεί το προηγούμενο θεώρημα, επειδή η συνάρτηση $\psi(x)$ δεν διατηρεί το ίδιο πρόσημο στο διάστημα $[x_0, x_2]$. Αν η αρχή των αξόνων τοποθετηθεί στο μέσο του (x_0, x_2), τότε προκύπτει η παρακάτω έκφραση για το τοπικό σφάλμα :

$$E(h) = (h/3)\,[f(-h) + 4f(0) + f(h)] - \int_{-h}^{h} f(x)\,dx$$

Έπειτα από διαδοχικές παραγωγίσεις της E(h), βρίσκεται η τρίτη παράγωγος,

$$E'''(h) = \frac{h}{3}[f'''(h) - f'''(-h)] \qquad (6.4)$$

και επίσης ότι, $E'(0) = 0$, $E''(0) = 0$, $E'''(0) = 0$.

Σύμφωνα με το θεώρημα της μέσης τιμής, (διαφ. λογισμού)

$E'''(2h^2/3)\, f^{(4)}(\theta.h)$, όπου $-1 < \theta < 1$.

Συνεπώς μπορεί να οριστεί μία συνεχής συνάρτηση g(h), έτσι που

$$g(h) = f^{(4)}(\theta.h).$$

$$E'''(h) = (2h^2/3)\, g(h). \qquad (6.5)$$

Η E(h) μπορεί να γραφτεί με τη μορφή :

$$E(h) = \int_0^h \frac{(h-t)^2}{2}\,\frac{2}{3}t^2\ g(t)\,dt\ . \qquad (6.6)$$

Με την παράσταση (6.6) ικανοποιούνται οι (6.4) και (6.5). Επειδή η συνάρτηση $(h - t)^2\ t^2/3$ δεν αλλάζει πρόσημο στο διάστημα (-h, h), μπορεί τώρα να εφαρμοστεί το θεώρημα μέσης τιμής του ολοκληρωτικού λογισμού :

$$E(h) = g(\bar{\theta}.h)\int_0^h (h - t)^2 (t^2/3)\,dt = \frac{h^2}{90} g(\bar{\theta}.h)\ , \qquad 0 < \bar{e} < 1\ .$$

Συνεπώς το τοπικό σφάλμα γράφεται :

$$E(h) = \frac{h^5}{90} f^{(4)}(\xi) \quad \text{όπου } -h < \xi < h\ .$$

Εύκολα αποδεικνύεται ότι το **ολικό σφάλμα** για N υποδιαστήματα του διαστήματος [a, b] παίρνει τη μορφή :

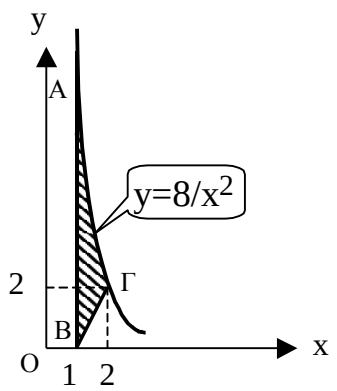
$$E_N^S = \frac{b-a}{180} h^4 . f^{(4)}(\eta), \qquad \text{όπου } \mathbf{a < \eta < b}.$$

Στο κεφάλαιο 8 δίνεται το πρόγραμμα που υπολογίζει σαν παράδειγμα το ολοκλήρωμα. $\int_1^2 \frac{dx}{x}$

Ο αριθμός των υποδιαιρέσεων του διαστήματος (1, 2) είναι 20. Η ακριβής τιμή του ολοκληρώματος είναι I = 0.6931472.

Άσκηση

Να υπολογιστεί με τη μέθοδο Simpson το εμβαδόν του γραμμοσκιασμένου τμήματος ΑΒΓ με ακρίβεια δύο δεκαδικών ψηφίων.



Λύση

Η ευθεία που διέρχεται από τα σημεία Β, Γ έχει εξίσωση y = 2(x - 1). Επομένως το εμβαδόν ΑΒΓ δίνεται από το ολοκλήρωμα

$$\int_1^2 f(x)\,dx = \int_1^2 [8/x^2 - 2(x-1)]\,dx \ .$$

Άρα,

$f'(x) = -16/x^3 - 2$

. . . . . .

. . . . . .

$f^{(4)}(x) = 960/x^6$, η οποία γίνεται μέγιστη για x = 1 (φθίνουσα συνάρτηση).

Άρα, $$E_N^S = \frac{2-1}{180}.h^4.(960/x^6)\Big|_{x=1} < 5*10^{-3} \ .$$

Επομένως h < 0.175 → N > 5.715 και επιλέγεται N = 6.

x	**f(x)**	
1	8.	* 1
7/6	5.54421	* 4
8/6	3.83333	* 2
9/6	2.55556	* 4
10/6	1.54667	* 2
11/6	0.71349	* 4
2	0.	* 1

Το εμβαδόν είναι λοιπόν

S = (1/18) . (8. + 4 * 5.54421 + 2 * 3.83333 + 4 * 2.55556 + 2 * 1.54667 + 4 * 0.71349 + 0.) = 3.0007.

7. ΑΡΙΘΜΗΤΙΚΗ ΕΠΙΛΥΣΗ ΔΙΑΦΟΡΙΚΩΝ ΕΞΙΣΩΣΕΩΝ

7.1 ΜΕΘΟΔΟΣ ΤΟΥ EULER

Το πρόβλημα είναι η επίλυση της διαφορικής εξίσωσης πρώτης τάξης

$$y' = f(x, y), \qquad y(x_0) = y_0. \tag{7.1}$$

Το πρόβλημα (7.1) έχει μοναδική λύση, αν η $\frac{\partial f}{\partial y}$ είναι συνεχής στο διάστημα ολοκλήρωσης.

Για την αριθμητική επίλυση εκλέγεται ένα σταθερό βήμα ολοκλήρωσης h. Συνεπώς για δύο διαδοχικά σημεία x_n και x_{n+1},

$$x_{n+1} - x_n = h \qquad \text{και} \qquad x_n = x_0 + n \cdot h.$$

Από το θεώρημα του Taylor,

$$y(x_{n+1}) = y(x_n) + h.y'(x_n) + \frac{h^2}{2}.y''(\xi_n) \qquad , x_n \le \xi_n \le x_{n+1} \tag{7.2}$$

Από την εξίσωση (7.2), κατά προσέγγιση,

$$y(x_{n+1}) = y(x_n) + h.y'(x_n) . \tag{7.3}$$

Από την (7.1),

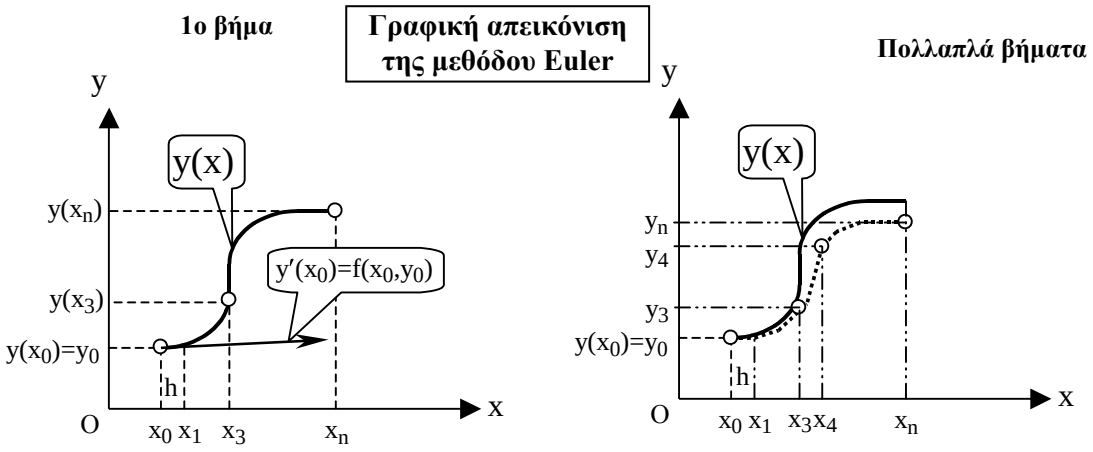
$$y'(x_n) = f[x_n, y(x_n)] .$$

Συμβολίζεται με $y(x_n)$ η ακριβής τιμή της λύσης του προβλήματος (7.1) για $x = x_n$ και με y_n η προσεγγιστική τιμή που προκύπτει στο n - οστό βήμα. Τότε,

$$y'(x_n) \approx f(x_n, y_n)$$

και από την εξίσωση (7.3) προκύπτει το παρακάτω υπολογιστικό σχήμα που είναι γνωστό σαν μέθοδος του Euler :

$$y_{n+1} = y_n + h \cdot f(x_n, y_n) = y_n + \Delta y_n \text{ , όπου } n = 0, 1, \ldots \qquad (7.4)$$



Αγνοείται το σφάλμα στρογγυλοποίησης στην παρακάτω ανάλυση και εξετάζεται το σφάλμα διακριτοποίησης ε_n.

$$\varepsilon_n = y_n - y(x_n). \qquad (7.5)$$

Από τις εξισώσεις (7.2), (7.4) και (7.5) προκύπτει ότι :

$$\varepsilon_{n+1} = y_{n+1} - y(x_{n+1}) = y_n + h.f(x_n, y_n) - [y(x_n) + h.\, y'(x_n) + \frac{h^2}{2} y''(\xi_n)] \Rightarrow$$

$$\varepsilon_{n+1} = y_n - y(x_n) + h\ \ f(x_n, y_n) - f[x_n, y(x_n)] \ - \frac{h^2}{2} y''(\xi_n) \Rightarrow$$

$$\varepsilon_{n+1} = \varepsilon_n + h.f_y(x_n, \bar{y}_n).[y_n - y(x_n)] - \frac{h^2}{2} y''(\xi_n) \Rightarrow \quad \text{(θ. μέσης τιμής)}$$

$$\varepsilon_{n+1} = \varepsilon_n + h \cdot f_y(x_n, \bar{y}_n) \cdot \varepsilon_n - \frac{h^2}{2} y''(\xi_n) \qquad (7.7)$$

Στο σημείο αυτό γίνεται η υπόθεση ότι στο διάστημα ολοκλήρωσης $f_y(x, y) < L$ και $y''(x) < M$, όπου L και M σταθερές. Τελικά,

$$\varepsilon_{n+1} \leq (1 + h.L).\varepsilon_n + \frac{h^2}{2}.M \ . \qquad (7.8)$$

Η εξίσωση (7.8) γράφεται με τη μορφή

$$\varepsilon_{n+1} \le A.\varepsilon_n + B,$$

από την οποία επαγωγικά αποδεικνύεται ότι

$$\varepsilon_n \le A^n.\varepsilon_0 + \frac{A^n - 1}{A - 1}.B, \quad \text{ή} \quad \varepsilon_n = \frac{A^n - 1}{A - 1}.B \qquad \text{επειδή } \varepsilon_0 = 0.$$

Αλλά $A^n = (1 + h.L)^n < e^{nhL} = e^{L(x_n - x_0)}$ και $B = \frac{h^2}{2}.M$.

Άρα
$$\varepsilon_n < \frac{1}{2} h.M \frac{e^{(x_n - x_0)} - 1}{L} \ . \tag{7.9}$$

Η εξίσωση (7.9) αποδεικνύει ότι το σφάλμα είναι $\varnothing(h^2)$. Ακόμα φαίνεται ότι για $h \to 0$ το σφάλμα τείνει στο 0.

Άσκηση

Δίνεται η διαφορική εξίσωση $dy = x.y^2.dx$ με αρχική συνθήκη $y(1) = 1$. Να λυθεί στο διάστημα [1.0, 1.3] με τη μέθοδο Euler και με $h = 0.1$. Να εκτιμηθεί το σφάλμα εφαρμόζοντας την ανάλυση σφάλματος της μεθόδου και να υπολογιστεί το πραγματικό σφάλμα.

Λύση

Η αναλυτική λύση της εξίσωσης υπολογίζεται $-\frac{1}{y} = \frac{x^2}{2} + C$ και με την συνθήκη $y(1) = 1$ έχουμε τελικά

$$y = \frac{2}{3 - x^2} \ .$$

Από τη σχέση (7.7) **το σφάλμα της μεθόδου** είναι

$$\varepsilon_{n+1} = \varepsilon_n \left[1 + h.f_y(x_n, \bar{y}_n) \right] - \frac{h^2}{2} y''(\xi_n) \ .$$

Έχουμε επομένως

$y' = f(x, y) = x \,.\, y^2$,

$f_y = 2x.y$, $f_x = y^2$ και επομένως

$y'' = f_x + f_y \,.\, y' = f_x + f_y \,.\, f = y^2 + 2x^2 \,.\, y^3$.

προσεγγιστική τιμή - ακριβής τιμή

$$\frac{2}{3-1.1^2}$$

x	**y**	**f(x, y)**	**1+h.f$_y$**	**0.5 h^2.y″**	**σφάλμα μεθόδου**	**σφάλμα πραγματικό**
1.0	1.000	1.000	1.200	0.015	0	0
1.1	1.100	1.331	1.242	0.022	-0.019	-0.017
1.2	1.233	1.825	1.296	0.035	-0.052	-0.049
1.3	1.416	2.605	1.368	0.058	-0.114	-0.111

$$y_{n+1} = y_n + h.f(x_n, y_n)$$

1.416 = 1.233 + 0.1 * 1.825

$$\varepsilon_{n+1} = \varepsilon_n \left[1 + h.f_y(x_n, \bar{y}_n) \right] - \frac{h^2}{2} y''(\xi_n)$$

- 0.052 = (-0.019) * [1.242] – (0.022 + 0.035)/2

7.2 ΜΕΘΟΔΟΙ RUNGE - KUTTA

Η μέθοδος Euler δεν προσφέρει μεγάλη ακρίβεια. Από την άλλη, μία παραπέρα ανάπτυξη της σειράς Taylor, σύμφωνα με την εξίσωση (7.2) θα είχε σαν συνέπεια την εισαγωγή παραγώγων ανώτερης τάξης και την επιβάρυνση των υπολογισμών. Οι μέθοδοι Runge - Kutta έχουν σαν επιδίωξη την αποφυγή παραγώγων ανώτερης τάξης, ενώ ταυτόχρονα πετυχαίνουν την ίδια ακρίβεια με σειρά Taylor όσο το δυνατόν μεγαλύτερης τάξης.

Το $y(x_{n+1})$ αναπτύσσεται σε σειρά Taylor

$$y(x_{n+1}) = y(x_n) + h.y'(x_n) + \frac{h^2}{2!}.y''(x_n) + \frac{h^3}{3!}.y'''(x_n) + \ldots$$

$$= y(x_n) + h.f(x_n, y_n) + \frac{h^2}{2!}.(f_x + f.f_y)\Big|_{(x_n, y_n)} + \tag{7.10}$$

$$+\frac{h^3}{6}.\ f_{xx} + 2f.f_{xy} + f_{yy}.f^3 + f_x.f_y + f_y{}^2.f\ \Big|_{(x_n, y_n)} + \varnothing(h^4)$$

Το σχήμα Runge - Kutta διαμορφώνεται ως εξής :

$$y_{n+1} = y_n + a.K_1 + b.K_2$$

$$K_1 = h.f(x_n, y_n)$$

$$K_2 = h.f(x_n + \alpha.h, y_n + \beta.K_1), \tag{7.11}$$

όπου a, b, α, β θα προσδιοριστούν παρακάτω. Με τη βοήθεια και πάλι του θεωρήματος του Taylor,

$K_2/h = f(x_n + \alpha.h, y_n + \beta.K_1) =$

$= f(x_n, y_n) + \alpha.h.f_x + \beta.K_1.f_y + (\alpha^2.h^2/2).f_{xx} + \alpha.\beta.h.K_1.f_{xy} +$

$+ (\beta^2.K_1^2/2).f_{yy} + \varnothing(h^3)$

Η παραπάνω έκφραση αντικαθίσταται στην (7.11) :

$y_{n+1} = y_n + (a + b).h.f + b.h^2.(\alpha.f_x + \beta.f.f_y) + b.h^3.[(\alpha^2/2).f_{xx} +$

$+ \alpha.\beta.f.f_{xy} + (\beta^2/2).f^2.f_{yy})] + \varnothing(h^4)$. (7.12)

Για να συμφωνήσουν οι (7.10) και (7.12) κατά τους όρους h και h^2, πρέπει

$a + b = 1$ και $b.\alpha = b.\beta = 1/2.$

Με $a = b = 1/2$ και $\alpha = \beta = 1$, η μέθοδος **Runge - Kutta 2ης τάξης** γράφεται:

$$y_{n+1} = y_n + (1/2).(K_1 + K_2) ,$$

$$K_1 = h.f(x_n, y_n)$$

$$K_2 = h.f(x_n + h, y_n + K_1) .$$

Πιο συνηθισμένη είναι η μέθοδος **Runge - Kutta 4ης τάξης** :

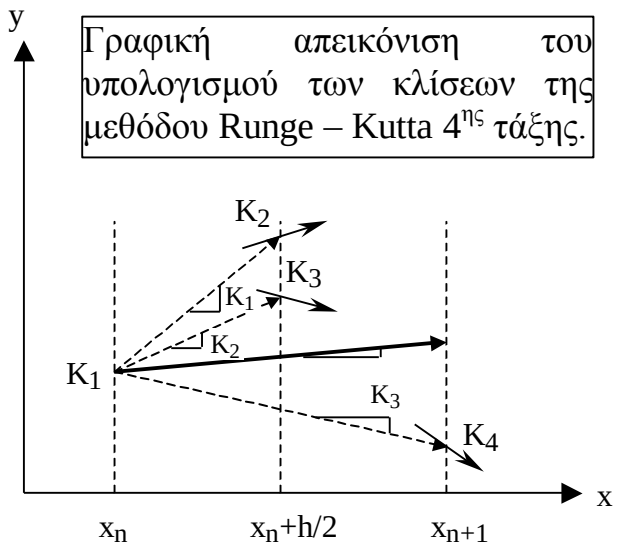
$y_{n+1} = y_n + (1/6).(K_1 + 2K_2 + 2K_3 + K_4) = y_n + \Delta y_n$ (7.13)

$K_1 = h.f(x_n, y_n)$

$K_2 = h.f(x_n + h/2, y_n + K_1/2)$

$K_3 = h.f(x_n + h/2, y_n + K_2/2)$

$K_4 = h.f(x_n + h, y_n + K_3)$ **.**



Γραφική απεικόνιση του υπολογισμού των κλίσεων της μεθόδου Runge – Kutta 4ης τάξης.

Ο παρακάτω πίνακας μπορεί να χρησιμοποιηθεί για ευκολία κατά την επίλυση μιας διαφορικής εξίσωσης με τη μέθοδο Runge - Kutta.

n	x	y	K = h.f(x, y)	Δy
0	$\mathbf{x_0}$	$\mathbf{y_0}$	$K_1^{(0)}$	$K_1^{(0)}$
	$x_0 + h/2$	$y_0 + K_1^{(0)}/2$	$K_2^{(0)}$	$2K_2^{(0)}$
	$x_0 + h/2$	$y_0 + K_2^{(0)}/2$	$K_3^{(0)}$	$2K_3^{(0)}$
	$x_0 + h$	$y_0 + K_3^{(0)}$	$K_4^{(0)}$	$K_4^{(0)}$
				$\Delta y_0 =$ $(1/6).\sum\Delta y$
				$y_1 = y_0 + \Delta y_0$
1	$\mathbf{x_1}$	$\mathbf{y_1}$	$K_1^{(1)}$	$K_1^{(1)}$
. . .	. . .	. . .	. . .	. . .

Άσκηση

Να λυθεί η διαφορική εξίσωση $y' = 0.25y^2 + x^2$ στο διάστημα [0, 0.5], με βήμα h = 0.1 και αρχική συνθήκη y(0) = -1.

n	x	y	K =h.f(x,y)	Δy
0	**0**	**-1**	0.025	0.025
	0.05	-0.98750	0.024629	0.049258
	0.05	-0.98769	0.024638	0.049276
	0.1	-0.97536	0.024783	0.024783
				Δy=0.02472
1	**0.1**	**-0.97528**	. . .	. . .
. . .	. . .	. . .	. . .	. . .

7.3 ΜΕΘΟΔΟΙ ΠΕΠΕΡΑΣΜΕΝΩΝ ΔΙΑΦΟΡΩΝ

Οι μέθοδοι πεπερασμένων διαφορών είναι από τις πλέον συνηθισμένες για την επίλυση διαφορικών εξισώσεων με οριακές συνθήκες, που δεν μπορούν να λυθούν με αναλυτικές μεθόδους. Οι παράγωγοι σε κάποιο σημείο προσεγγίζονται από πηλίκα διαφορών, π. χ. η $\partial u / \partial x$ προσεγγίζεται από $\Delta u / \Delta x$, όπου το Δx έχει μια πολύ μικρή τιμή.

7.3.1 ΠΡΟΣΕΓΓΙΣΗ ΠΑΡΑΓΩΓΩΝ ΜΕ ΠΕΠΕΡΑΣΜΕΝΕΣ ΔΙΑΦΟΡΕΣ.

Αν η συνάρτηση u(x) και οι παράγωγοί της είναι συνεχείς και μονότιμες τότε σύμφωνα με το θεώρημα του Taylor

$$u(x+h) = u(x) + h.u'(x) + \frac{1}{2}h^2.u''(x) + \frac{1}{6}h^3.u'''(x) + \ldots$$

καθώς επίσης και

$$u(x-h) = u(x) - h.u'(x) + \frac{1}{2}h^2.u''(x) - \frac{1}{6}h^3.u'''(x) + \ldots$$

Προσθέτοντας τις δύο αυτές εξισώσεις έχουμε

$$u(x+h) + u(x-h) = 2u(x) + h^2 u''(x) + \varnothing(h^4)$$

ή προσεγγιστικά

$$u''(x) = \left(\frac{d^2u}{dx^2}\right)_{x=x} = \frac{u(x+h) - 2u(x) + u(x-h)}{h^2} \qquad (7.14)$$

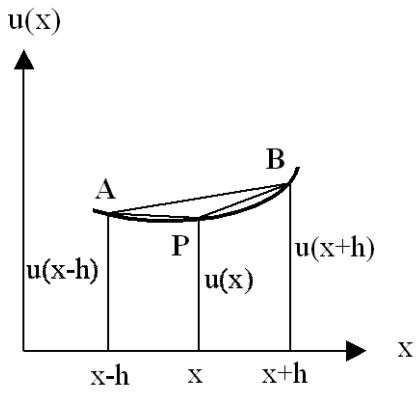
με σφάλμα της τάξης h^2.

Αφαιρώντας τις ίδιες εξισώσεις καταλήγουμε :

$$u'(x) = \left(\frac{du}{dx}\right)_{x=x} = \frac{u(x+h) - u(x-h)}{2h} \qquad (7.15)$$

με σφάλμα της τάξης h^2.

Ουσιαστικά, με την εξίσωση (7.15) προσεγγίζεται η κλίση της εφαπτομένης στο σημείο P (βλέπε σχήμα), από την κλίση της χορδής AB και έχουμε μία προσέγγιση **κεντρικής διαφοράς (central difference).**

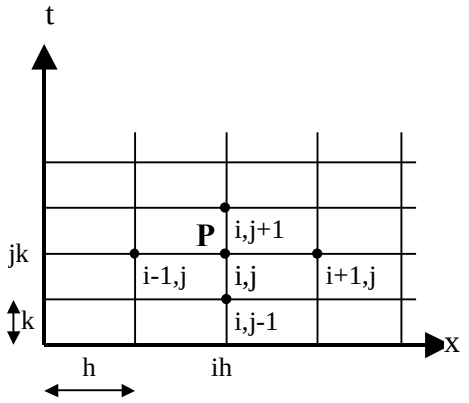


Δύο παρόμοιες προσεγγίσεις είναι οι εξής :

$$u'(x) = \frac{u(x+h)-u(x)}{h} \qquad \textbf{(forward-difference)}$$

$$u'(x) = \frac{u(x)-u(x-h)}{h} \qquad \textbf{(backward-difference)}.$$

Ας υποθέσουμε τώρα ότι η u είναι συνάρτηση των x και t και ας χωρίσουμε το επίπεδο x - t σε ορθογώνια διαστάσεων Δx = h και Δt = k, όπως φαίνεται στο επόμενο σχήμα.



Οι συντεταγμένες του σημείου P λοιπόν θα είναι x = ih και t = jk, όπου i και j ακέραιοι.

Αν συμβολίσουμε την τιμή της u στο P με

$u_P = u(ih, jk) = u_{i,j}$,

θα έχουμε από την εξίσωση (7.14)

$$\left(\frac{\partial^2 u}{\partial x^2}\right)_P = \left(\frac{\partial^2 u}{\partial x^2}\right)_{i,j} = \frac{u[(i+1)h, jk] - 2u[ih, jk] + u[(i-1)h, jk]}{h^2}$$

ή

$$\left(\frac{\partial^2 u}{\partial x^2}\right)_{i,j} = \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} \tag{7.16}$$

Χρησιμοποιώντας επίσης την προσέγγιση forward-difference θα έχουμε για την παράγωγο της u ως προς t, στο σημείο P

$$\frac{\partial u}{\partial t} = \frac{u_{i,j+1} - u_{i,j}}{k} \ . \tag{7.17}$$

Άσκηση.

Να λυθεί η διαφορική εξίσωση παραβολικού τύπου

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \tag{F0}$$

με τις συνθήκες :

α) <u>Οριακές συνθήκες</u>

u = 0 όταν x = 0 και x = 1, για κάθε τιμή του t.

β) <u>Αρχική συνθήκη (t = 0)</u>

u = 2x για $0 \le x \le 1$,

u = 2(1-x) για $1/2 \le x \le 1$.

Λύση

Σύμφωνα με τις εξισώσεις (7.16) και (7.17), η παραπάνω διαφορική εξίσωση μπορεί να γραφτεί

$$\frac{u_{i,j+1} - u_{i,j}}{k} = \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} \tag{F1}$$

όπου $x = ih$, $(i = 0, 1, 2, \ldots)$

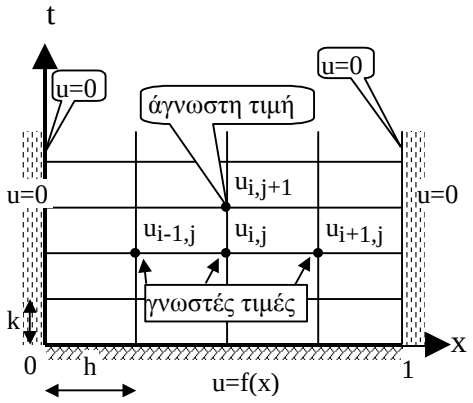
και $t = jk$, $(j = 0, 1, 2, \ldots)$.

Η σχέση (F1) μπορεί να γραφτεί

$$u_{i,j+1} = u_{i,j} + r(u_{i-1,j} - 2u_{i,j} + u_{i+1,j}) \qquad \text{(F2)}$$

όπου $r = \dfrac{\Delta t}{(\Delta x)^2} = \dfrac{k}{h^2}$,

και δίνει τη δυνατότητα υπολογισμού της άγνωστης τιμής $u_{i,j+1}$ στον κόμβο (i, j+1) με βάση τις γνωστές τιμές στο προηγούμενο j βήμα του χρόνου.



Υπολογίζονται λοιπόν οι άγνωστες τιμές u του 1ου βήματος του χρόνου (t = k), με βάση τις γνωστές οριακές και αρχικές τιμές για t = 0. Κατόπιν, υπολογίζονται οι άγνωστες τιμές του 2ου βήματος του χρόνου, με βάση τις ήδη υπολογισμένες τιμές του πρώτου, κλπ.

Μία σχέση της μορφής της (F2), όπου η άγνωστη τιμή u υπολογίζεται άμεσα με βάση γνωστές τιμές, ονομάζεται **ρητή (explicit).**

Το συγκεκριμένο πρόβλημα είναι όπως φαίνεται συμμετρικό ως προς x = 1/2.

Αν ληφθεί $\Delta x = h = 1/10$, $\Delta t = k = 1/1000$, έχουμε

$$r = \frac{\Delta t}{(\Delta x)^2} = \frac{k}{h^2} = \frac{1}{10}$$

και η σχέση (F2) γίνεται

$$u_{i,j+1} = 1/10(u_{i-1,j} + 8u_{i,j} + u_{i+1,j}) .$$

Έτσι υπολογίζονται τελικά οι τιμές της u σε κάθε κόμβο, όπως φαίνεται στον παρακάτω πίνακα. Για παράδειγμα,

$u_{5,1} = 1/10[0.8 + (8 * 1) + 0.8] = 0.9600$

$u_{4,2} = 1/10[0.6 + (8 * 0.8) + 0.96] = 0.7960.$

		i = 0	**i = 1**	**i = 2**	**i = 3**	**i = 4**	**i = 5**	**i = 6**
		x = 0	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**
j = 0	**t = 0**	0	0.2000	0.4000	0.6000	0.8000	1.0000	0.8000
j = 1	**0.001**	0	0.2000	0.4000	0.6000	0.8000	0.9600	0.8000
j = 2	**0.002**	0	0.2000	0.4000	0.6000	0.7960	0.9280	0.7960
j = 3	**0.003**	0	0.2000	0.4000	0.5996	0.7896	0.9016	0.7896
j = 4	**0.004**	0	0.2000	0.4000	0.5986	0.7818	0.8792	0.7818
j = 5	**0.005**	0	0.2000	0.3999	0.5971	0.7732	0.8597	0.7732
.....	**.....**	.....	.....	.....	.....	.....	.....	.....
j = 10	**0.01**	0	0.1996	0.3968	0.5822	0.7281	0.7867	0.7281
.....	**.....**	.....	.....	.....	.....	.....	.....	.....
j = 20	**t=0.02**	0	0.1938	0.3781	0.5373	0.6486	0.6891	0.6486

7.3.2 ΜΕΘΟΔΟΣ ΤΩΝ CRANK - NICOLSON.

Η επίλυση μιας διαφορικής εξίσωσης με την μέθοδο που περιγράφτηκε είναι απλή, αλλά η ακρίβειά της εξαρτιέται από τη τιμή του λόγου r. Έχει

αποδειχτεί ότι δίνει ικανοποιητική ακρίβεια για τιμές του r, $0 < r \leq 0.5$, δηλαδή για πολύ μικρές τιμές του dx.

Οι Crank και Nicolson πρότειναν την αντικατάσταση της παραγώγου $\frac{\partial^2 u}{\partial x^2}$ από την μέση τιμή της πεπερασμένης διαφοράς στο (j + 1) και j βήμα του χρόνου. Έτσι η διαφορική εξίσωση (F0) γίνεται

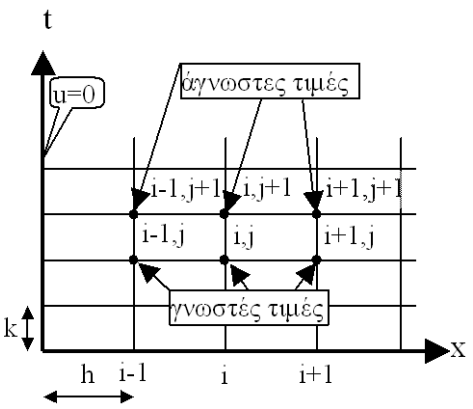
$$\frac{u_{i,j+1} - u_{i,j}}{k} = \frac{1}{2}\left\{ \frac{u_{i+1,j+1} - 2u_{i,j+1} + u_{i-1,j+1}}{h^2} + \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} \right\} \qquad (7.18)$$

η οποία μπορεί να γραφτεί τελικά

$$- ru_{i-1,j+1} + (2 + 2r)u_{i,j+1} - ru_{i+1,j+1} = ru_{i-1,j} + (2 - 2r)u_{i,j} + ru_{i+1,j}\,, \qquad (7.19)$$

όπου $r = k/h^2$.

Η σχέση (F3) περιέχει όπως φαίνεται στο αριστερό μέλος τρεις άγνωστες και στο δεξιό μέλος τρεις γνωστές τιμές της u. Αν υπάρχουν N εσωτερικοί κόμβοι σε κάθε βήμα του χρόνου, τότε για j = 0 και i = 1, 2, 3, . . ., N η (F3) δίνει ένα σύστημα N εξισώσεων για τις N άγνωστες τιμές του 1ου βήματος του χρόνου, συναρτήσει των γνωστών αρχικών και οριακών τιμών. Ομοίως, για j = 1 οι N άγνωστες τιμές της u εκφράζονται συναρτήσει των υπολογισμένων κατά το προηγούμενο βήμα τιμών της u, κτλ.



Ένα αριθμητικό σχήμα της μορφής που περιγράφηκε, κατά την οποία οι άγνωστες τιμές προκύπτουν από την

επίλυση ενός συστήματος εξισώσεων ονομάζεται **έμμεσο (implicit).**
Το μητρώο των συντελεστών των αγνώστων που προκύπτει είναι τριδιαγωνικό και επιλύεται με μία από τις μεθόδους που έχουν ήδη αναφερθεί.

Η άσκηση που λύθηκε προηγουμένως, για $\Delta x = h = 1/10$ και $\Delta t = k = 1/100$ οδηγεί τη σχέση (F3) στη μορφή

$$- u_{i-1,j+1} + 4u_{i,j+1} - u_{i+1,j+1} = u_{i-1,j} + u_{i+1,j}. \qquad \text{(F4).}$$

Κατά το πρώτο βήμα του χρόνου ($j = 0$) το σύστημα των εξισώσεων (όπως διαμορφώνεται και με βάση τις αρχικές και οριακές συνθήκες) είναι:

$$\begin{array}{rcrcrcrcl}
-0 & + & 4u_1 & - & u_2 & = & 0 & + & 0.4, \\
-u_1 & + & 4u_2 & - & u_3 & = & 0.2 & + & 0.6, \\
-u_2 & + & 4u_3 & - & u_4 & = & 0.4 & + & 0.8, \\
-u_3 & + & 4u_4 & - & u_5 & = & 0.6 & + & 1.0, \\
-2u_4 & + & 4u_5 & & & = & 0.8 & + & 0.8.
\end{array}$$

Από τη λύση του συστήματος προκύπτουν οι τιμές της u κατά το 1ο βήμα του χρόνου ($t = 0.01$),

$$u_1 = 0.1989,$$
$$u_2 = 0.3956,$$
$$u_3 = 0.5834,$$
$$u_4 = 0.7381,$$
$$u_5 = 0.7691.$$

Η ίδια διαδικασία επαναλαμβάνεται για το επόμενο βήμα ($j = 1$) κτλ.

8. ПРОГРАММАТА

1. ΕΠΙΤΑΧΥΝΣΗ ΣΥΝΑΡΤΗΣΙΑΚΩΝ ΕΠΑΝΑΛΗΨΕΩΝ ΚΑΤΑ ΑΙΤΚΕΝ

```
'FAITKEN.BAS
'ΠΡΟΓΡΑΜΜΑ ΓΙΑ ΤΗ ΜΕΘΟΔΟ ΑΙΤΚΕΝ
DECLARE SUB FROUTINE (X!, F!)
DECLARE SUB GROUTINE (X0, X1, X2)
OPEN "c:aitken.out" FOR OUTPUT AS #1
CLS
EPS = .000001: IMAX% = 25
'INITIALIZE
I% = 0
PRINT : INPUT " ΑΡΧΙΚΗ ΤΙΜΗ = "; X0
PRINT
PRINT #1, : PRINT #1, " ΑΡΧΙΚΗ ΤΙΜΗ = "; X0
PRINT #1,
DO
  CALL FROUTINE(X0, G): X1 = G
  CALL FROUTINE(X1, G): X2 = G
  I% = I% + 1
  PRINT " I="; I%; TAB(10); " X0="; X0
  PRINT TAB(10); " X1="; X1; TAB(28); " X2="; X2
  PRINT
  PRINT #1, " I="; I%; TAB(10); " X0="; X0
  PRINT #1, TAB(10); " X1="; X1; TAB(28); " X2="; X2
  PRINT #1,
  IF I% > IMAX% THEN
    PRINT
    PRINT " ΔΕΝ ΕΧΟΥΜΕ ΣΥΓΚΛΙΣΗ ΜΕΤΑ ΑΠΟ "; IMAX%; "
ΠΡΟΣΕΓΓΙΣΕΙΣ"
```

```
    EXIT DO
  END IF
  IF ABS(X2 - X1) < EPS THEN
    PRINT
    PRINT " FINAL X="; X2
    PRINT #1,
    PRINT #1, " FINAL X="; X2
    EXIT DO
   END IF
   'COMPUTE NEW X0
   CALL GROUTINE(X0, X1, X2)
LOOP
END

SUB FROUTINE (X, F)
F = SIN(X) + .25
END SUB

SUB GROUTINE (X0, X1, X2)
D1 = X1 - X0
D2 = X2 - 2 * X1 + X0
X0 = X0 - D1 * D1 / D2
END SUB
```

<u>ΑΠΟΤΕΛΕΣΜΑΤΑ ΠΡΟΓΡΑΜΜΑΤΟΣ</u>

ΑΡΧΙΚΗ ΤΙΜΗ = 1.4

```
 I= 1     X0= 1.4
```

X1= 1.23545 X2= 1.194296

I= 2 X0= 1.180571

X1= 1.174824 X2= 1.172622

I= 3 X0= 1.171255

X1= 1.171239 X2= 1.171233

I= 4 X0= 1.17123

X1= 1.17123 X2= 1.17123

FINAL X= 1.17123

2. ΜΕΘΟΔΟΣ ΔΙΧΟΤΟΜΗΣΗΣ (σελ. 34)

```
'BISECT.BAS
'MODULE CONTAINING
'SUBROUTINE BISECT(X0, DX, EPS)
DECLARE SUB FROUTINE (X!, F!)
'INITIALIZE
X1 = X0
CALL FROUTINE(X1, F1)
'SEARCH LOOP
DO
  X2 = X1 + DX
  CALL FROUTINE(X2, F2)
  IF F1 * F2 < 0 THEN
    EXIT DO
  END IF
  X1 = X2
  F1 = F2
LOOP
'BISECTION LOOP
I% = 0
PRINT : PRINT
LPRINT : LPRINT
DO
   I% = I% + 1
   X = .5 * (X1 + X2)
   CALL FROUTINE(X, F)
   PRINT " X="; X; " F="; F
   LPRINT " X="; X; " F="; F
   IF ABS(F) < EPS THEN
     EXIT DO
   END IF
   IF F * F1 < 0 THEN
     X2 = X: F2 = F
   ELSE
     X1 = X: F1 = F
   END IF
LOOP
END



SUB FROUTINE (X, F)
F = 2 * X * X - 7 * X + 3
```

```
END SUB
```

3. ΥΠΟΛΟΓΙΣΜΟΣ ΤΗΣ ΤΙΜΗΣ ΠΟΛΥΩΝΥΜΟΥ ΚΑΙ ΤΗΣ ΠΑΡΑΓΩΓΟΥ ΤΟΥ (σελ. 35)

```
'BRACKETS.BAS
'MODULE CONTAINING
'SUBROUTINE BRACKETS(A(), N%, Z, B, C)
'FOR EVALUATING A POLYNOMIAL AND ITS
'DERIVATIVE FOR X=Z

DECLARE SUB BRACKETS (A(), N%, Z, B, C)

READ Z, N%: DATA 2, 4
DIM A(N% + 1)
DATA -5,2,-1,2,3
CLS
FOR I = 0 TO N%
   READ A(I)
NEXT I

PRINT " ΒΑΘΜΟΣ ΠΟΛΥΩΝΥΜΟΥ N="; N%
PRINT
PRINT " ΣΥΝΤΕΛΕΣΤΕΣ ΠΟΛΥΩΝΥΜΟΥ: A0, A1,...,AN:"
PRINT "                        ";
FOR I = 0 TO N%
   PRINT A(I);
NEXT I
LPRINT " ΒΑΘΜΟΣ ΠΟΛΥΩΝΥΜΟΥ N="; N%
LPRINT
LPRINT " ΣΥΝΤΕΛΕΣΤΕΣ ΠΟΛΥΩΝΥΜΟΥ : A0, A1,...,AN:"
LPRINT "                         ";
FOR I = 0 TO N%
   LPRINT A(I);
NEXT I

CALL BRACKETS(A(), N%, Z, B, C)

PRINT : PRINT
PRINT " B="; B; " C="; C
LPRINT : PRINT
LPRINT " B="; B; " C="; C

END
```

```
SUB BRACKETS (A(), N%, Z, B, C)
B = A(N%)
C = B
FOR I% = 1 TO N% - 1
   B = A(N% - I%) + Z * B
   C = B + Z * C
NEXT I%
B = A(0) + Z * B
END SUB
```

4. ΜΙΓΑΔΙΚΕΣ ΡΙΖΕΣ ΠΟΛΥΩΝΥΜΩΝ ΜΕΘΟΔΟΣ NEWTON - BAIRSTOW (σελ. 43)

```
'BAIRSTOW.BAS
'PROGRAM APPLYING
'NEWTON - BAIRSTOW METHOD
DECLARE SUB BAIRSTOW (A!(), N%, P!, Q!, B0!, B1!, C1!, C2!, C3!)
'READ IN POLYNOMIAL DATA
OPEN “c:bairstow.out” FOR OUTPUT AS#1
READ N%: DATA 3
DIM A(N% + 1)
DATA 1,2,2,1
CLS
FOR I = 0 TO N%
   READ A(I)
NEXT I
'ECHO POLYNOMIAL DATA
PRINT " ΒΑΘΜΟΣ ΠΟΛΥΩΝΥΜΟΥ N="; N%
PRINT
PRINT " ΣΥΝΤΕΛΕΣΤΕΣ ΠΟΛΥΩΝΥΜΟΥ: A0, A1,...,AN:"
PRINT "                        ";
FOR I = 0 TO N%
   PRINT A(I);
NEXT I
PRINT #1, " ΒΑΘΜΟΣ ΠΟΛΥΩΝΥΜΟΥ N="; N%
PRINT #1,
PRINT #1, " ΣΥΝΤΕΛΕΣΤΕΣ ΠΟΛΥΩΝΥΜΟΥ : A0, A1,...,AN:"
PRINT #1, "                        ";
FOR I = 0 TO N%
   PRINT #1, A(I);
NEXT I
READ EPS: DATA 1.E-05
'INITIAL VALUES:
P = -2!: Q = -3!
PRINT : PRINT
PRINT " INITIAL VALUES:"
PRINT " P0="; P; " Q0="; Q
PRINT
PRINT #1, : PRINT
PRINT #1, " INITIAL VALUES:"
PRINT #1, " P0="; P; " Q0="; Q
PRINT #1,
DO
    CALL BAIRSTOW(A(), N%, P, Q, B0, B1, C1, C2, C3)
```

```
    DP = (-B1 * C2 + B0 * C3) / (C2 * C2 - C1 * C3)
    DQ = (-B0 * C2 + B1 * C1) / (C2 * C2 - C1 * C3)
    P = P + DP
    Q = Q + DQ
    PRINT " P="; P; TAB(15); " Q="; Q
    PRINT #1, " P="; P; TAB(15); " Q="; Q
LOOP UNTIL ABS(DP) + ABS(DQ) < EPS

SUB BAIRSTOW (A(), N%, P, Q, B0, B1, C1, C2, C3)
'INITIALIZE
B1 = A(N%)
B0 = A(N% - 1) + P * B1
'
C2 = B1
C1 = B0 + P * C2
'
FOR I% = 2 TO N%
   B = A(N% - I%) + P * B0 + Q * B1
   B1 = B0
   B0 = B
   IF I% = N% THEN
      EXIT FOR
   END IF
   C = B + P * C1 + Q * C2
   C3 = C2
   C2 = C1
   C1 = C
NEXT I%
END SUB
```

ΑΠΟΤΕΛΕΣΜΑΤΑ ΠΡΟΓΡΑΜΜΑΤΟΣ

ΒΑΘΜΟΣ ΠΟΛΥΩΝΥΜΟΥ N= 3

ΣΥΝΤΕΛΕΣΤΕΣ ΠΟΛΥΩΝΥΜΟΥ: A0, A1,...,AN:
1 2 2 1
INITIAL VALUES:
P0=-2 Q0=-3

P=-1.75 Q=-1.5
P=-1.176471 Q=-.7022059
P= .6492103 Q=-.386784
P=-.8796158 Q= 1.322817
P=-.839319 Q=-1.024194
P=-.9759283 Q=-.9819173
P=-1.000287 Q=-.9994066
P=-.9999995 Q=-.9999998
P=-.9999999 Q=-1

5. ΑΠΑΛΟΙΦΗ GAUSS (σελ. 50)

```
'GAUSS.BAS
'MODULE CONTAINING
'GAUSS(W(), N%, X()),
'A ROUTINE FOR SOLVING A LINEAR SYSTEM
'OF N EQUATIONS IN N UNKNOWNS

DECLARE SUB GAUSS (W(), N%, X())
DECLARE SUB MATPRINT (W!(), N%)
COMMON SHARED K%
OPEN “c:gauss.out” FOR OUTPUT AS#1
CLS

'READ IN MATRIX DATA

READ N%: DATA 3
DATA 1,2,3,14,2,3,1,11,0,1,2,8

'NOTE: THE FOLLOWING DATA CORRESPOND TO AN ILL-
'CONDITIONED SYSTEM
'DATA .78,-3.22,2.41,.82,2.33,4.28,-1.25,-.55,.34,2.92,-1.6,-.62

DIM W(N%, N% + 1), X(N%)
FOR I% = 1 TO N%
   FOR J% = 1 TO N% + 1
      READ W(I%, J%)
   NEXT J%
NEXT I%

'DATA ECHO

CALL MATPRINT(W(), N%)

CALL GAUSS(W(), N%, X())

DO
LOOP WHILE INKEY$ = ""

'ΕΚΤΥΠΩΣΗ ΤΗΣ ΛΥΣΗΣ
PRINT
PRINT : PRINT " ΛΥΣΗ": PRINT
PRINT #1,
PRINT #1, : PRINT #1, " ΛΥΣΗ": PRINT #1,
```

```
FOR I% = 1 TO N%
   PRINT "X("; I%; ")="; X(I%)
NEXT I%
FOR I% = 1 TO N%
   PRINT #1, "X("; I%; ")="; X(I%)
NEXT I%
END
SUB GAUSS (W(), N%, X())

FOR K% = 1 TO N% - 1
   '
   M = 0
   FOR I% = K% TO N%
      D = 0
      FOR J% = K% TO N%
         IF ABS(W(I%, J%)) > D THEN
            D = ABS(W(I%, J%))
         END IF
       NEXT J%
       '
       IF ABS(W(I%, K%)) / D > M THEN
         M = ABS(W(I%, K%)) / D
         I1% = I%
       END IF
   NEXT I%
   'ΑΝΤΑΛΛΑΓΗ ΣΕΙΡΩΝ I1 ΚΑΙ K
   FOR J% = K% TO N% + 1
      T = W(K%, J%)
      W(K%, J%) = W(I1%, J%)
      W(I1%, J%) = T
    NEXT J%
    CALL MATPRINT(W(), N%)
    'ΑΠΑΛΟΙΦΗ
    FOR I% = K% + 1 TO N%
       M = W(I%, K%) / W(K%, K%)
       FOR J% = K% TO N% + 1
          W(I%, J%) = W(I%, J%) - M * W(K%, J%)
       NEXT J%
    NEXT I%
    CALL MATPRINT(W(), N%)
NEXT K%

'ΛΥΣΗ ΤΡΙΓΩΝΙΚΟΥ ΣΥΣΤΗΜΑΤΟΣ
X(N%) = W(N%, N% + 1) / W(N%, N%)
FOR K% = 1 TO N% - 1
```

```
   S = 0
   FOR J% = 0 TO K% - 1
      S = S + X(N% - J%) * W(N% - K%, N% - J%)
   NEXT J%
   X(N% - K%) = (W(N% - K%, N% + 1) - S) / W(N% - K%, N% - K%)
NEXT K%

END SUB

SUB MATPRINT (W(), N%)

PRINT : PRINT " K="; K%; "  MATRIX W(I,J)": PRINT
PRINT #1, : PRINT #1, " K="; K%; "  MATRIX W(I,J)": PRINT #1,

FOR I% = 1 TO N%
   FOR J% = 1 TO N%
      JT% = (J% - 1) * 10 + 1
      PRINT ; TAB(JT%); W(I%, J%);
      PRINT #1, ; TAB(JT%); W(I%, J%);
   NEXT J%
   PRINT TAB(JT% + 15); W(I%, N% + 1)
   PRINT #1, TAB(JT% + 15); W(I%, N% + 1)
NEXT I%

END SUB
```

ΑΠΟΤΕΛΕΣΜΑΤΑ ΠΡΟΓΡΑΜΜΑΤΟΣ

K= 0 MATRIX W(I,J)

1	2	3	14
2	3	1	11
0	1	2	8

K= 1 MATRIX W(I,J)

2	3	1	11
1	2	3	14
0	1	2	8

K= 1 MATRIX W(I,J)

2	3	1	11
0	.5	2.5	8.5
0	1	2	8

K= 2 MATRIX W(I,J)

2	3	1	11
0	1	2	8
0	.5	2.5	8.5

K= 2 MATRIX W(I,J)

2	3	1	11
0	1	2	8
0	0	1.5	4.5

ΛΥΣΗ

X(1)= 1
X(2)= 2
X(3)= 3

6. ΣΥΣΤΗΜΑ ΜΕ ΤΡΙΔΙΑΓΩΝΙΚΟ ΠΙΝΑΚΑ ΑΛΓΟΡΙΘΜΟΣ THOMAS (σελ. 54)

```
'THOMAS.BAS
'A PROGRAM FOR THOMAS ALGORITHM
OPEN “c:thomas.out” FOR OUTPUT AS#1
CLS
READ N%: DATA 4
DIM A(N%), B(N%), C(N%), D(N%), X(N%)
DATA 1,1,1
DATA 1,3,3,2
DATA 2,1,1
DATA 5,10,15,11

FOR I% = 2 TO N%
   READ A(I%)
NEXT I%

FOR I% = 1 TO N%
   READ B(I%)
NEXT I%

FOR I% = 1 TO N% - 1
   READ C(I%)
NEXT I%

FOR I% = 1 TO N%
   READ D(I%)
NEXT I%

PRINT
PRINT #1,
FOR I% = 1 TO N%
   PRINT "I%="; I%; " A="; A(I%); " B="; B(I%); " C="; C(I%); " D="; D(I%)
   PRINT #1, "I%="; I%; " A="; A(I%); " B="; B(I%); " C="; C(I%); " D=";
D(I%)
NEXT I%

C(1) = C(1) / B(1)
FOR I% = 2 TO N% - 1
   B(I%) = B(I%) - C(I% - 1) * A(I%)
   C(I%) = C(I%) / B(I%)
NEXT I%
B(N%) = B(N%) - C(N% - 1) * A(N%)
```

```
'
D(1) = D(1) / B(1)
FOR I% = 2 TO N%
   D(I%) = D(I%) - A(I%) * D(I% - 1)
   D(I%) = D(I%) / B(I%)
NEXT I%
'

FOR I% = 1 TO N% - 1
   D(N% - I%) = D(N% - I%) - C(N% - I%) * D(N% - I% + 1)
NEXT I%
',,ΕΚΤΥΠΩΣΗ ΛΥΣΗΣ
PRINT
PRINT #1,
FOR I% = 1 TO N%
   PRINT "X("; I%; ")="; D(I%)
   PRINT #1, "X("; I%; ")="; D(I%)
NEXT I%

END
```

<u>ΑΠΟΤΕΛΕΣΜΑΤΑ ΠΡΟΓΡΑΜΜΑΤΟΣ</u>

```
I%= 1  A= 0  B= 1  C= 2  D= 5
I%= 2  A= 1  B= 3  C= 1  D= 10
I%= 3  A= 1  B= 3  C= 1  D= 15
I%= 4  A= 1  B= 2  C= 0  D= 11

X( 1 )= 1
X( 2 )= 2
X( 3 )= 3
X( 4 )= 4
```

7. ΕΠΑΝΑΛΗΠΤΙΚΕΣ ΜΕΘΟΔΟΙ
ΜΕΘΟΔΟΣ GAUSS - SEIDEL (σελ. 56)

```
'SEIDEL.BAS
'PROGRAM FOR GAUSS - SEIDEL METHOD.
DECLARE SUB MATECHO (A!(), B!(), N%)
DECLARE SUB XPRINT (K%, X!(), N%)
COMMON SHARED K%
OPEN “c:seidel.out” FOR OUTPUT AS#1
CLS
READ EPS, KMAX%
DATA 1.E-05, 25
'READ IN MATRIX DATA
READ N%: DATA 3
DATA 12,1,1,1,12,1,1,1,12
DATA 14,14,14
DATA 0,0,0
'READ N%: DATA 3
'DATA 37,6,0,6,37,6,0,6,37
'DATA 19,15,8
'DATA 0,0,0
DIM A(N%, N%), B(N%), X(N%)

FOR I% = 1 TO N%
   FOR J% = 1 TO N%
      READ A(I%, J%)
   NEXT J%
NEXT I%

FOR I% = 1 TO N%
   READ B(I%)
NEXT I%

'DATA ECHO
CALL MATECHO(A(), B(), N%)

PRINT
PRINT " PRESS ANY KEY TO CONTINUE"
DO
LOOP WHILE INKEY$ = ""
'INITIAL X()
FOR I% = 1 TO N%
   READ X(I%)
```

```
NEXT I%

K% = 0
CALL XPRINT(K%, X(), N%)
DO
   D = 0
   FOR I% = 1 TO N%
       TEMP = X(I%)
       S = 0
       FOR J% = 1 TO N%
          IF J% <> I% THEN
             S = S + A(I%, J%) * X(J%)
          END IF
       NEXT J%
       X(I%) = (B(I%) - S) / A(I%, I%)
       D = D + ABS(X(I%) - TEMP)
    NEXT I%
    K% = K% + 1
    CALL XPRINT(K%, X(), N%)
    IF K% > KMAX% THEN
      PRINT "ΔΕΝ ΣΥΓΚΛΙΝΕΙ ΜΕΤΑ ΑΠΟ "; K%; " ΠΡΟΣΕΓΓΙΣΕΙΣ"
      EXIT DO
    END IF
LOOP UNTIL D < EPS

END



SUB MATECHO (A(), B(), N%)
PRINT : PRINT " SYSTEM MATRIX ": PRINT
PRINT #1, : PRINT #1, " SYSTEM  MATRIX ": PRINT #1,

FOR I% = 1 TO N%
   FOR J% = 1 TO N%
      JT% = (J% - 1) * 10 + 1
      PRINT ; TAB(JT%); A(I%, J%);
      PRINT #1, ; TAB(JT%); A(I%, J%);
   NEXT J%
   PRINT TAB(JT% + 15); B(I%)
   PRINT #1, TAB(JT% + 15); B(I%)
NEXT I%
END SUB

SUB XPRINT (K%, X(), N%)
```

```
PRINT : PRINT " ITER. NO="; K%: PRINT
PRINT #1, : PRINT #1, " ITER. NO="; K%: PRINT #1,
FOR I% = 1 TO N%
   PRINT "X("; I%; ")="; X(I%)
NEXT I%
FOR I% = 1 TO N%
   PRINT #1, "X("; I%; ")="; X(I%)
NEXT I%
END SUB
```

ΑΠΟΤΕΛΕΣΜΑΤΑ ΠΡΟΓΡΑΜΜΑΤΟΣ

SYSTEM MATRIX

12	1	1	14
1	12	1	14
1	1	12	14

ITER. NO= 0

X(1)= 0
X(2)= 0
X(3)= 0

ITER. NO= 1

X(1)= 1.166667
X(2)= 1.069444
X(3)= .9803241

ITER. NO= 2

X(1)= .9958526
X(2)= 1.001985
X(3)= 1.00018

ITER. NO= 3

X(1)= .9998196
X(2)= 1
X(3)= 1.000015

ITER. NO= 4

X(1)= .9999987
X(2)= .9999989
X(3)= 1

ITER. NO= 5

X(1)= 1
X(2)= 1
X(3)= 1

8. ΠΟΛΥΩΝΥΜΟ ΠΑΡΕΜΒΟΛΗΣ ΚΑΤΑ NEWTON (σελ. 76)

```
'NEWTON.BAS
READ N%: DATA 2
DIM X(N% + 1), D(N% + 1)
'READ IN DATA

FOR I% = 0 TO N%
   READ X(I%)
NEXT I%
FOR I% = 0 TO N%
   READ D(I%)
NEXT I%
DATA 1, 3, 5
DATA 1.5708, 1.5719, 1.5738
'DATA ECHO
CLS
PRINT
PRINT #1,
PRINT " ΔΕΔΟΜΕΝΑ ΖΕΥΓΗ ΤΙΜΩΝ"
PRINT
PRINT #1, " ΔΕΔΟΜΕΝΑ ΖΕΥΓΗ ΤΙΜΩΝ"
PRINT #1,

FOR I% = 0 TO N%
   PRINT " X="; X(I%); " F="; D(I%)
   PRINT #1, " X="; X(I%); " F="; D(I%)
NEXT I%
PRINT : PRINT
PRINT #1, : PRINT #1,

INPUT " ΤΙΜΗ ΤΟΥ X ="; Z
PRINT #1, " ΤΙΜΗ ΤΟΥ X ="; Z

FOR K% = 1 TO N%
   FOR I% = N% TO K% STEP -1
      D(I%) = (D(I%) - D(I% - 1)) / (X(I%) - X(I% - K%))
   NEXT I%
NEXT K%
'
B = D(N%)
FOR I% = 1 TO N%
   B = D(N% - I%) + B * (Z - X(N% - I%))
NEXT I%
```

```
'
PRINT
PRINT " ΤΙΜΗ ΤΟΥ ΠΟΛΥΩΝΥΜΟΥ NEWTON ="; B
PRINT #1,
PRINT #1, " ΤΙΜΗ ΤΟΥ ΠΟΛΥΩΝΥΜΟΥ NEWTON ="; B

END
```

ΑΠΟΤΕΛΕΣΜΑΤΑ ΠΡΟΓΡΑΜΜΑΤΟΣ

ΔΕΔΟΜΕΝΑ ΖΕΥΓΗ ΤΙΜΩΝ

X= 1 F= 1.5708
X= 3 F= 1.5719
X= 5 F= 1.5738

ΤΙΜΗ ΤΟΥ X = 3.5

ΤΙΜΗ ΤΟΥ ΠΟΛΥΩΝΥΜΟΥ NEWTON = 1.5723

9. ΚΑΝΟΝΑΣ ΤΟΥ SIMPSON (σελ. 91)

```
'SIMPSON.BAS
'MODULE CONTAINING SUBROUTINE SIMPSON (A, B, N, S) TO
'PERFORM NUMERICAL INTEGRATION
'ACCORDING TO SIMPSON'S METHOD
'A ΕΙΝΑΙ ΤΟ ΚΑΤΩ ΟΡΙΟ ΤΗΣ ΟΛΟΚΛΗΡΩΣΗΣ
'B ΕΙΝΑΙ ΤΟ ΑΝΩ ΟΡΙΟ ΤΗΣ ΟΛΟΚΛΗΡΩΣΗΣ
'N ΕΙΝΑΙ Ο ΑΡΙΘΜΟΣ ΤΩΝ ΥΠΟΔΙΑΣΤΗΜΑΤΩΝ
DECLARE SUB FFUNCTION (X#, F#)
COMMON SHARED BETA AS DOUBLE

SUB FFUNCTION (X#, F#)
DEFDBL A-Z
'PROBLEM DATA
'QKs is Q/Ks
QKs = .5#
hb = 1#
lambda = 2#
p = 1#
miu = 2# + lambda * (2# + p)
BETA = (1 / QKs) ^ (1# / miu) * hb
U = X / BETA
F = 1# / (1# + U ^ miu)

END SUB


DEFSNG A-Z
SUB SIMPSON (A#, B#, N%, S#)
DEFDBL A-Z

DX = (B - A) / N%
X = A: I% = 0
CALL FFUNCTION(X#, F#)
S = F: C = 4
DO
  I% = I% + 1: X = X + DX

  CALL FFUNCTION(X#, F#)
  IF I% = N% THEN
    EXIT DO
  END IF
  S = S + C * F
```

```
   C = 6 - C
LOOP

S = S + F
S = S * DX / 3

END SUB
```

10. ΜΕΘΟΔΟΣ RUNGE - KUTTA (σελ. 97)

```
'RUNGE.BAS
'ΠΡΟΓΡΑΜΜΑ ΓΙΑ ΤΗ ΜΕΘΟΔΟ
'RUNGE-KUTTA 4ΗΣ ΤΑΞΗΣ

DECLARE SUB FROUTINE (X#, Y#, F#)
DECLARE SUB PRINTXY (X#, Y#, YE#)
DEFDBL A-Z
'Η ΕΙΝΑΙ ΤΟ ΒΗΜΑ ΟΛΟΚΛΗΡΩΣΗΣ
CLS
READ H, XM
DATA 0.02#, 2.3#

'ΑΡΧΙΚΕΣ ΣΥΝΘΗΚΕΣ
READ X0, Y0
DATA 2#, 2#
X = X0: Y = Y0

DO
   YE = X / 2# + 2# / X
   CALL PRINTXY(X, Y, YE)

   IF X >= XM THEN
     EXIT DO
   END IF

   CALL FROUTINE(X, Y, F)
   K1 = H * F

   X = X0 + .5# * H
   Y = Y0 + .5# * K1
   CALL FROUTINE(X, Y, F)
   K2 = H * F

   X = X0 + .5# * H
   Y = Y0 + .5# * K2
   CALL FROUTINE(X, Y, F)
   K3 = H * F

   X = X0 + H
   Y = Y0 + K3
   CALL FROUTINE(X, Y, F)
   K4 = H * F
```

```
    Y = Y0 + (K1 + 2# * K2 + 2# * K3 + K4) / 6#

    X0 = X
    Y0 = Y

LOOP

DEFSNG A-Z
SUB FROUTINE (X#, Y#, F#)
DEFDBL A-Z

F = 1 - Y / X

END SUB


DEFSNG A-Z
SUB PRINTXY (X#, Y#, YE#)
DEFDBL A-Z

PRINT "X="; X; TAB(23); " Y="; Y; TAB(47); " YE="; YE
LPRINT "X="; X; TAB(23); " Y="; Y; TAB(47); " YE="; YE

END SUB
```

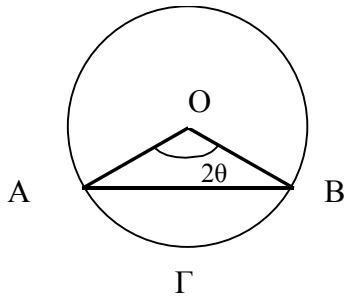
10. ΑΣΚΗΣΕΙΣ ΓΙΑ ΛΥΣΗ

1. Να υπολογιστεί με ακρίβεια τεσσάρων δεκαδικών ψηφίων η μικρότερη ρίζα της εξίσωσης $5e^{-0.2x} = \sin x$.

2. Να υπολογιστεί με ακρίβεια τεσσάρων δεκαδικών ψηφίων η θετική ρίζα της εξίσωσης $x = 2\sin^2 x$.

3. Με τη μέθοδο Newton - Raphson, να υπολογιστεί η μεγαλύτερη θετική ρίζα της εξίσωσης $0.15x - \cos x = 0$.

4. Με τη μέθοδο Newton - Raphson και με ακρίβεια 10^{-3} να υπολογιστεί η γωνία **θ**, ώστε το εμβαδόν του κυκλικού τμήματος ΑΒΓΑ να είναι διπλάσιο του εμβαδού του τριγώνου ΑΟΒ.
(<u>Απάντηση</u> θ = 1.139 rad)



5. Με τη μέθοδο Newton - Raphson και με ακρίβεια 10^{-3} να υπολογιστεί η θετική ρίζα της εξίσωσης $x^3 - \cosh x = 0$.

$$(\cosh x = \frac{e^x + e^{-x}}{2})$$

6. Με τη μέθοδο Newton - Raphson και με ακρίβεια 10^{-3} να υπολογιστεί η δεύτερη κατά σειρά μεγέθους θετική ρίζα της εξίσωσης $e^{-x^2} - \sin 2x = 0$.

7. Να αποδειχτεί ότι ο αλγόριθμος

$$x_{i+1} = \frac{1}{2}(x_i + \frac{N}{x_i})$$

για την εύρεση της τετραγωνικής ρίζας ενός αριθμού, βασίζεται στη μέθοδο Newton - Raphson.

8. Δίνονται τα παρακάτω ζεύγη αντιστοίχων τιμών.

x	14	17	31	35
y	68.7	64	44	39.1

Να υπολογιστεί η τιμή y(27) με τη μέθοδο της επαναλαμβανόμενης γραμμικής παρεμβολής.

9. Να υπολογιστεί το πολυώνυμο παρεμβολής Newton το οποίο διέρχεται από τα σημεία (-1, 0), (0, -5), (1, -4), (3, 16), (6, 721).

(<u>Απάντηση</u> $x^4 - 3x^3 + 2x^2 + x - 5$)

10. Δίνονται τα παρακάτω τρία ζεύγη αντιστοίχων τιμών $[x_i, f(x_i)]$, $i = 0, 1, 2, \ldots$ όπου $f(x) = \sin 2x$.

[0.75, 0.9975], [1.0, 0.9093], [1.5, 0.1411].

Να προσεγγιστεί με βάση τις τιμές αυτές η f(1.2) με τη βοήθεια του πολυωνύμου Lagrange και στη συνέχεια να γίνει εκτίμηση του σφάλματος της παρεμβολής.

11. Δίνονται τα παρακάτω ζεύγη τιμών και ζητείται να υπολογιστεί η τιμή y(2.7) με τη βοήθεια του πολυωνύμου παρεμβολής Lagrange.

x	2.0	2.3	3.0	3.5
y	5.838	6.127	6.694	7.047

12. Με πόση ακρίβεια θα μπορούσατε να υπολογίσετε την τιμή $\cos\frac{\pi}{5}$ χρησιμοποιώντας πολυώνυμα παρεμβολής, αν γνωρίζατε τις τιμές

$\cos\frac{\pi}{6}, \cos\frac{\pi}{4}, \cos\frac{\pi}{3}$ και $\cos\frac{\pi}{2}$;

13. Να υπολογιστεί το ολοκλήρωμα $\int_2^3 x \ln x \, dx$

με τον τραπεζοειδή κανόνα και ακρίβεια 4 δεκαδικών για h = 0.1. Στη συνέχεια να βρεθεί το πραγματικό σφάλμα και να συγκριθεί με το μέγιστο.

(<u>Απάντηση</u> I = 2.3078)

14. Να υπολογιστεί η τιμή του ολοκληρώματος $\int_0^1 (1+x)^{-1/2} dx$

με την αριθμητική μέθοδο Simpson έτσι ώστε το ολικό σφάλμα της μεθόδου να είναι μικρότερο του $5*10^{-6}$. (<u>Απάντηση</u> N= 8)

15. Να υπολογιστεί το ολοκλήρωμα $\int_0^2 x\, e^x dx$

με τη μέθοδο Simpson και ακρίβεια 5 δεκαδικών ψηφίων, για N = 6 υποδιαστήματα και να εκτιμηθεί το σφάλμα. (<u>Απάντηση</u> I = 8.39134)

16. Δίνονται τα παρακάτω ζεύγη τιμών $[x_i, f(x_i)]$, i = 0, 1, 2, 3.

[0.5, -0.75], [1, 0], [1.5, 1.25], [2, 3].

Να υπολογιστεί η καμπύλη 3ου βαθμού (spline) που περνά από τα σημεία

[1, 0], [1.5, 1.25].

Οριακές συνθήκες στα ακραία σημεία $S_0 = S_1$ και $S_3 = S_2$.

(<u>Απάντηση</u> $a_1 = 0$, $b_1 = 1$, $c_1 = 2$, $d_1 = 0$)

17. Να λυθεί το παρακάτω σύστημα των εξισώσεων.

$$\begin{aligned} 5x_1 + x_2 - x_3 &= 4 \\ 2x_1 + 6x_2 - x_3 &= 11 \\ -2x_1 - x_2 + 8x_3 &= 20 \end{aligned}$$

18. Με τη χρησιμοποίηση κατάλληλου ολοκληρώματος, της αριθμητικής μεθόδου Simpson και υπολογισμό του ακριβούς βήματος ολοκλήρωσης να υπολογιστεί η τιμή του $\ln \sqrt[3]{2}$ με ακρίβεια τεσσάρων δεκαδικών ψηφίων.

19. Με τη μέθοδο Runge - Kutta 4ης τάξης να λυθεί η εξίσωση

$$\frac{y'}{y} = \frac{1}{x} - y$$

με οριακή συνθήκη y(1) = 1 στο διάστημα [1, 2] και με h = 0.2.

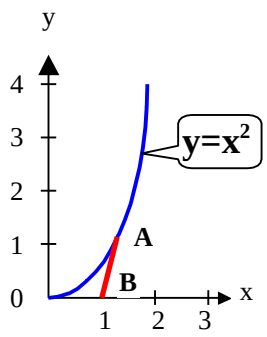
20. Έστω η διαφορική εξίσωση $y' = a.x$ με αρχική συνθήκη y(0) = 0, όπου a σταθερά. Να αποδειχτεί ότι η μέθοδος Runge - Kutta 2ης τάξης δίνει την ακριβή αναλυτική λύση, αν εφαρμοστεί στην παραπάνω εξίσωση.

21. Δίνεται η διαφορική εξίσωση $y' = -100y + 100$ με αρχική συνθήκη $y(0) = y_0$. Να δοθεί αναλυτική λύση για την προσέγγιση y_n της μεθόδου Euler στο n-οστό βήμα.

22. Με χρησιμοποίηση αριθμητικής μεθόδου, να γραφτεί αλγόριθμος για τον υπολογισμό του αντίστροφου ενός θετικού αριθμού Α. Στη συνέχεια χρησιμοποιώντας ως αρχικές τιμές 0.2 και 0.4 να υπολογιστεί ο αντίστροφος του αριθμού 6 με ακρίβεια τεσσάρων δεκαδικών ψηφίων. Τι παρατηρείτε;

23. Το άθροισμα δύο αριθμών είναι ίσο με **20**. Αν σε κάθε αριθμό προστεθεί η τετραγωνική του ρίζα, το γινόμενο των δύο αθροισμάτων είναι ίσο με **155.55**. Με χρήση προσεγγιστικών μεθόδων να υπολογιστούν οι δύο αριθμοί με ακρίβεια $\mathbf{10^{-4}}$.
(<u>Απάντηση</u> $x_1 = 6.5128$, $x_2 = 13.4872$)

24. Με τη μέθοδο Newton - Raphson και με ακρίβεια τεσσάρων δεκαδικών ψηφίων να υπολογιστούν οι συντεταγμένες ενός σημείου Α της καμπύλης $\mathbf{y = x^2}$, ώστε η απόστασή του από το Β(1,0) να είναι ελάχιστη.



(<u>Απάντηση</u> x = 0.5898, y = 0.3478)

ΒΑΣΙΚΗ ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Numerical Solution of Partial Differential Equations, Smith, G.D., Oxford University Press, London, 1965.
2. Applied Numerical Methods, Carnahan, Brice, et. al., Wiley, New York, 1969.
3. Applied Numerical Analysis, Gerald, C.F., Addison-Wesley Publishing Co., 1970.
4. Introduction to Applied Numerical Analysis, Hamming, R.W., McGraw-Hill, New York, 1971.
5. Elementary Numerical Analysis, Conte, S.D. and Carl de Boor, McGraw-Hill, 1972.
6. Iterative Methods for the Solution of Equations, Traub ,J.F. Chelsea, 1982.
7. Computational Mathematics, Kopchenova, N.V. and I.A. Maron, Mir Publishers, Moskow, 1984.
8. Numerical Analysis, Burden, R.L. and J.D. Faires, Brooks/Cole Publeshing Company, 1997
9. Εισαγωγή στην Αριθμητική Ανάλυση, Ακρίβης, Γ.Δ., Β.Α. Δουγλής, 1997
10. Numerical Methods for Engineers, Chapra, S.C. and R.P. Canale, WCB/McGraw Hill, 1998.