Σύντομες Σημειώσεις στην Αριθμητική Ανάλυση των Διαφορικών Εξισώσεων

ΓΕΩΡΓΟΣ ΠΑΠΑΓΕΩΡΓΙΟΥ

Αν. Καθηγητής

Τμήμα Εφαρμοσμένων Μαθηματικών & Φυσικών Επιστημών

Ε. Μ. Πολυτεχνείο

ΑΘΗΝΑ 2001-2002

Πρόλογος

Οι σημειώσεις αυτές απευθύνονται στους σπουδαστές του τρίτου έτους του Τμήματος Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών του Ε. Μ. Πολυτεχνείου, για τις ανάγκες του μαθήματος Αριθμητική Ανάλυση ΙΙ και Εργαστήριο, και ειδικότερα το κομμάτι της ύλης που αφορά την προσέγγιση της λύσης του προβλήματος αρχικών τιμών 1ης τάξης της μορφής:

$$y' = f(x, y),\ x \in [a, b]$$
$$y(a) = y_0$$

για μία απλή ή ένα σύστημα συνήθων διαφορικών εξισώσεων, αν και η πρόθεση είναι να επεκταθούν και στην μελέτη της αριθμητικής επίλυσης άλλων τύπων διαφορικών εξισώσεων.

Τέτοια προβλήματα εμφανίζονται συχνά στους διάφορους κλάδους της Επιστήμης και της Τεχνολογίας. Συνήθως, αναλυτική λύση, δηλαδή λύση αυτών των προβλημάτων σε κλειστή μορφή, δύσκολα υπολογίζεται και στις περισσότερες περιπτώσεις δεν μπορεί να υπολογιστεί. Επομένως, η εφαρμογή αριθμητικών (προσεγγιστικών) μεθόδων, με την βοήθεια των υπολογιστών, αποτελεί τον συνηθισμένο τρόπο επίλυσής τους.

Οι προσεγγιστικές μέθοδοι για συνήθεις διαφορικές εξισώσεις, αν και είναι μία από τις πλέον παλιές περιοχές της Αριθμητικής Ανάλυσης, αποτελούν και σήμερα μία από τις πλέον σημαντικές περιοχές έρευνας και εφαρμογών. Ένας μεγάλος αριθμός ερευνητών, ασχολείται με την επινόηση και υπολογισμό νέων μεθόδων, την βελτίωση και ανάπτυξη της θεωρίας, την βελτίωση υπαρχόντων μεθόδων, και την συστηματική σύγκριση των μεθόδων κάθε κατηγορίας. Στην παραπέρα ανάπτυξη αυτής της περιοχής των Εφαρμοσμένων Μαθηματικών, σημαντικό ρόλο έχει παίξει η ραγδαία ανάπτυξη των υπολογιστών τα τελευταία χρόνια.

Οι σημειώσεις αυτές διαπραγματεύονται δύο βασικές κατηγορίες αριθμητικών μεθόδων για την επίλυση συνήθων διαφορικών εξισώσεων. Οι δύο βασικές κατηγορίες είναι οι μέθοδοι απλού βήματος, και οι μέθοδοι πολλών βημάτων γενικά k-βημάτων. Πέρα από την περιγραφή του τρόπου κατασκευής των μεθόδων αυτών, μελετώνται συστηματικά βασικές έννοιες όπως, η έννοια της **σύγκλισης,** της **τάξης ακρίβειας,** της **συνέπειας,** της **ευστάθειας** και της εκτίμησης του **τοπικού σφάλματος αποκοπής,** οι οποίες παίζουν καθοριστικό ρόλο για την ποιοτική μελέτη των μεθόδων και τον έλεγχο των αριθμητικών αποτελεσμάτων.

Γιώργος Παπαγεωργίου

Ιανουάριος 2002

Μέρος Πρώτο

Συνήθεις Διαφορικές Εξισώσεις

Γενική θεωρία και βασικές έννοιες

1. Εισαγωγή

Οι συνήθεις διαφορικές εξισώσεις εμφανίζονται συχνά σαν μαθηματικά μοντέλα σε πολλούς κλάδους της επιστήμης μηχανικής και οικονομίας

Δυστυχώς, είναι σπάνιο αυτές οι εξισώσεις να έχουν λύσεις οι οποίες μπορούν να εκφραστούν σε κλειστή μορφή. Είναι επομένως αναγκαίο να αναζητήσουμε προσεγγιστικές λύσεις με κατάλληλες αριθμητικές μεθόδους. Σήμερα αυτό μπορεί να γίνει με πολύ οικονομικό τρόπο, με υψηλή ακρίβεια και ένα αξιόπιστο φράγμα στο σφάλμα μεταξύ της αναλυτικής λύσης και της αριθμητικής προσέγγισης.

Στην συνέχεια θα ασχοληθούμε με την κατασκευή και ανάλυση των αριθμητικών μεθόδων για την επίλυση της διαφορικής εξίσωσης της μορφής:

$$y' = f\left(x, y\right),\ x \in \left[a, b\right] \tag{1.1}$$

όπου y είναι μία πεπερασμένη συνάρτηση με πραγματικές τιμές, και $y' = dy/dx$. Η επίλυση ενός συστήματος διαφορικών εξισώσεων 1ης τάξης της μορφής:

$$\mathbf{y}' = \mathbf{f}\left(x, \mathbf{y}\right),\ x \in \left[a, b\right] \tag{1.2}$$

όπου $\mathbf{y} = \left[y_1, y_2, \ldots, y_n\right]^{\mathrm{T}}$ και $\mathbf{f} = \left[f_1, f_2, \ldots, f_n\right]^{\mathrm{T}}$ μπορεί να αντιμετωπιστεί με τις ίδιες προσεγγιστικές μεθόδους όπως και η εξίσωση (1.1). Ανάλογη αντιμετώπιση μπορούμε να έχουμε για την διαφορική εξίσωση τάξης n της μορφής:

$$y^{(n)} = f\left(x, y, y', y'', \ldots, y^{(n)}\right),\ x \in \left[a, b\right] \tag{1.3}$$

ή ενός συστήματος n διαφορικών εξισώσεων τάξης n της μορφής:

$$\mathbf{y}^{(n)} = \mathbf{f}\left(x, \mathbf{y}, \mathbf{y}', \mathbf{y}'', \ldots \mathbf{y}^{(n)}\right),\ \ x \in \left[a, b\right]. \tag{1.4}$$

2. Πρόβλημα αρχικών τιμών 1ης τάξης

Η διαφορική εξίσωση (1.1) ή (1.2) αν έχει λύση τότε η λύση αυτή ανήκει σε μία οικογένεια με άπειρες λύσεις και ορίζεται σαν η γενική λύση της διαφορικής εξίσωσης.

Στην πράξη αναζητούμε συνήθως να υπολογίσουμε ένα μέλος αυτής της οικογένειας, και αυτό προσδιορίζεται αν η διαφορική εξίσωση συνοδεύεται με την λεγόμενη αρχική συνθήκη. Έτσι ορίζεται το πρόβλημα αρχικών τιμών 1ης τάξης της μορφής:

$$\begin{aligned} &y' = f\left(x, y\right),\ x \in \left[a, b\right] \\ &y\left(a\right) = y_0 \end{aligned} \tag{2.1}$$

και το πρόβλημα αρχικών τιμών 1ης τάξης για συστήματα της μορφής:

$$\begin{aligned} &\mathbf{y}'=\mathbf{f}(x,\mathbf{y}),\ x\in[a,b] \\ &\mathbf{y}(a)=\mathbf{y}_0 \end{aligned} \tag{2.2}$$

αντίστοιχα.

Εν γένει, ακόμη και όταν η $f(.,.)$ είναι μία συνεχής συνάρτηση, δεν είναι σίγουρο ότι το πρόβλημα (2.1) έχει μοναδική λύση.

Για παράδειγμα, το πρόβλημα,

$$y'=y^{2/3},\ y(0)=0,$$

έχει δύο λύσεις τις

$$y(x)\equiv 0 \text{ και } y(x)=x^3/27.$$

Αν η συνάρτηση f πληροί επί πλέον υποθέσεις, μπορεί να αποδειχθεί η ύπαρξη και μοναδικότητα της λύσης του προβλήματος (2.1) σύμφωνα με το επόμενο θεώρημα.

Θεώρημα 2.1 (Picard)

Υποθέτουμε ότι η συνάρτηση f είναι συνεχής στο σύνολο:

$$D=\{(x,y):a\le x\le b,\ |y-y_0|<Y_m\}$$

όπου $Y_m>0$ σταθερά. Επί πλέον υποθέτουμε ότι υπάρχει μία σταθερά L έτσι ώστε η ανισότητα:

$$|f(x,y)-f(x,z)|\le L|y-z| \tag{2.3}$$

να ισχύει για κάθε επιλογή των ζευγών (x,y) και (x,z) του D. Τότε υπάρχει μία μοναδική συνάρτηση $y(x)$ ορισμένη στο διάστημα $a\le x\le b$, η οποία ικανοποιεί το πρόβλημα (2.1).

Η συνθήκη (2.3) ονομάζεται συνθήκη Lipschitz, και το L σταθερά Lipschitz.

Το θεώρημα (2.1) έχει μία φυσική επέκταση για το πρόβλημα (2.2) όπου $\mathbf{y}_0\in\mathrm{R}^n$, και η $\mathbf{f}:[a,b]\times\mathrm{R}^n\to\mathrm{R}^n$. Εισάγοντας την Ευκλείδεια νόρμα $\|\bullet\|$ στο R^n με την σχέση:

$$\|\mathbf{u}\|=\left(\sum_{i=1}^{n}|u_i|^2\right)^{1/2},\quad \mathbf{u}\in\mathrm{R}^n$$

το προηγούμενο θεώρημα μπορεί να διατυπωθεί ως εξής:

Θεώρημα 2.2

Υποθέτουμε ότι η συνάρτηση $\mathbf{f}$ είναι συνεχής στο σύνολο:

$$D=\{(x,y):a\le x\le b,\ \|\mathbf{y}-\mathbf{y}_0\|<Y_m\}$$

όπου $Y_m>0$ σταθερά. Επί πλέον υποθέτουμε ότι υπάρχει μία σταθερά L έτσι ώστε η ανισότητα:

$$\left\| \mathbf{f}(x,\mathbf{y}) - \mathbf{f}(x,\mathbf{z}) \right\| \le L \left\| \mathbf{y} \text{-} \mathbf{z} \right\| \tag{2.4}$$

να ισχύει για κάθε επιλογή των ζευγών $(x,\mathbf{y})$ και $(x,\mathbf{z})$ του D.όπου $\|\bullet\|$ είναι η Ευκλείδεια νόρμα Τότε υπάρχει μία μοναδική συνάρτηση $\mathbf{y}(x)$ ορισμένη στο διάστημα $a \le x \le b$, η οποία ικανοποιεί το πρόβλημα (2.2).

Μία ικανή συνθήκη για να ισχύει η (2.4) είναι η $\mathbf{f}$ να είναι συνεχής στο D, παραγωγίσιμη σε κάθε εσωτερικό σημείο $(x,\mathbf{y})$ του D, και να υπάρχει σταθερά $L > 0$ έτσι ώστε:

$$\left\| \frac{\partial \mathbf{f}}{\partial \mathbf{y}}(x,\mathbf{y}) \right\| \le L \tag{2.5}$$

για κάθε $(x,\mathbf{y})$ εσωτερικό του D, όπου $\partial \mathbf{f} / \partial \mathbf{y}$ είναι ο $n \times n$ Ιακωβίνος πίνακας της

$$y \in \mathrm{R}^n \to \mathbf{f}(x,\mathbf{y}) \in \mathrm{R}^n$$

και $\|\bullet\|$ είναι μία νόρμα πίνακα συμβατή στην Ευκλείδεια νόρμα διανύσματος.

3. Πρόβλημα αρχικών τιμών ανώτερης τάξης

Πολλά φυσικά προβλήματα μοντελοποιούνται με διαφορικές εξισώσεις ανώτερης τάξης. Το πρόβλημα αρχικών τιμών για μία διαφορική εξίσωση τάξης n, εκφράζεται ως εξής:

$$\begin{aligned} & y^{(n)} = f\left(x, y, y^{(1)}, \dots, y^{(n-1)}\right), \; x \in [a,b] \\ & y^{(i)}(a) = \eta_i, \; i = 0,1,\dots n-1 \end{aligned} \tag{3.1}$$

Προβλήματα αυτής της μορφής μπορούν εύκολα να αναχθούν σε προβλήματα αρχικών τιμών για συστήματα διαφορικών εξισώσεων 1ης τάξης με n εξισώσεις. Η αναγωγή γίνεται ως εξής:

Ορίζουμε τις μεταβλητές,

$$y_1 = y, \; y_2 = y^{(1)}, \; y_3 = y^{(2)}, \dots, y_{n-1} = y^{(n-2)}, \; y_n = y^{(n-1)}$$

οπότε οι συναρτήσεις, $y_1(x), \; y_2(x), \dots, y_n(x)$ ικανοποιούν το σύστημα n εξισώσεων πρώτης τάξης της μορφής:

$$\begin{aligned} & y_1' = y_2 \\ & y_2' = y_3 \\ & . \\ & . \\ & y_{n-1}' = y_n \\ & y_n' = f(x, y_1, y_2, \dots, y_n) \end{aligned} \tag{3.2}$$

με αρχικές συνθήκες:

$$y_1(a) = \eta_0, \; y_2(a) = \eta_1, \; y_3(a) = \eta_2, \dots, y_n(a) = \eta_{n-1}. \tag{3.3}$$

Έτσι το πρόβλημα (3.1) ανάγεται τελικά στο πρόβλημα:

$$\mathbf{y}' = \mathbf{f}\left(x, \mathbf{y}(x)\right), \; x \in [a, b] \\ \mathbf{y}(a) = \eta \tag{3.4}$$

όπου $\mathbf{f} = \left[y_2 , y_3 , ..., y_n , f\left(x, y_1, y_2, ..., y_n\right)\right]^{\mathrm{T}}$.

Παράδειγμα 3.1 Η διαφορική εξίσωση,

$$y^{(3)} = -y + x^2$$

είναι 3ης τάξης με ενδιάμεσες παραγώγους $y^{(1)}$ και $y^{(2)}$. Θέτουμε:

$$y_1 = y, \; y_2 = y^{(1)}, \; y_3 = y^{(2)}$$

και παραγωγίζοντας προκύπτει τελικά

$$\begin{aligned} y_1' &= y^{(1)} = y_2 \\ y_2' &= y^{(2)} = y_3 \\ y_3' &= y^{(3)} = -y + x^2 = -y_1 + x^2 \end{aligned}.$$

Τότε το ζητούμενο σύστημα διαφορικών εξισώσεων 3ης τάξης γράφεται:

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= y_3 \\ y_3' &= -y_1 + x^2 \end{aligned}$$

Γενικά, μία διαφορική εξίσωση τάξης n, ανάγεται σε σύστημα n διαφορικών εξισώσεων 1ης τάξης. Ένα σύστημα διαφορικών εξισώσεων ανώτερης τάξης, μπορεί να αναχθεί σε σύστημα διαφορικών εξισώσεων 1ης τάξης, αν εφαρμόσουμε κατάλληλα τον προηγούμενο μετασχηματισμό για κάθε μία εξίσωση του συστήματος. Είναι προφανές ότι ένα σύστημα k εξισώσεων τάξης n, ανάγεται σε ένα $k \times n$ σύστημα διαφορικών εξισώσεων 1ης τάξης. Για παράδειγμα ένα σύστημα με δύο εξισώσεις 2ης τάξης μετασχηματίζεται σε ισοδύναμο σύστημα τεσσάρων εξισώσεων 1ης τάξης.

4. Ευσταθείς και ασταθείς διαφορικές εξισώσεις

Θεωρούμε μία διαφορική εξίσωση πρώτης τάξης. Ως γνωστόν, η εξίσωση αν έχει λύση, θα έχει μία οικογένεια λύσεων. Αν τα μέλη της οικογένειας των λύσεων απομακρύνονται το ένα από το άλλο όταν το x αυξάνει, τότε λέμε ότι η εξίσωση είναι **ασταθής**. Αν όμως τα μέλη της οικογένειας αυτής πλησιάζουν το ένα το άλλο όταν το x αυξάνει, τότε λέμε ότι η εξίσωση είναι **ευσταθής**.

Μιλώντας πιο αυστηρά, μπορούμε να πούμε ότι μία μικρή διαταραχή σε μία λύση μιας ευσταθούς διαφορικής εξίσωσης θα εξουδετερωθεί με τον χρόνο, διότι οι καμπύλες των λύσεων συγκλίνουν, ενώ για μία ασταθή διαφορική εξίσωση, η διαταραχή θα μεγαλώνει με τον χρόνο, επειδή οι καμπύλες των λύσεων θα αποκλίνουν.

Αυτή η ποιοτική έννοια της ευστάθειας για μία διαφορική εξίσωση μπορεί να μελετηθεί πιο αναλυτικά ως εξής.

Θεωρούμε το σύστημα των διαφορικών εξισώσεων:

$$\mathbf{y}'=\mathbf{f}(x,\mathbf{y}),\;\; x\in[a,b]$$

και τον Ιακωβιανό πίνακα $\mathbf{J}$, με τα στοιχεία

$$\{J\}_{i,j}=\partial f_i/\partial y_j\,. \tag{4.1}$$

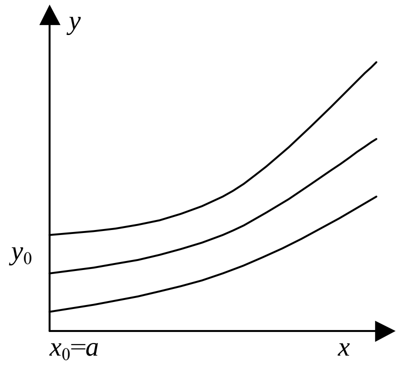
i. Αν κάποια από τις ιδιοτιμές του πίνακα $\mathbf{J}$ έχουν θετικά πραγματικά μέρη, τότε η διαφορική εξίσωση είναι ασταθής.

ii. Αν όλες οι ιδιοτιμές του πίνακα $\mathbf{J}$ έχουν αρνητικά πραγματικά μέρη τότε η διαφορική εξίσωση είναι ευσταθής.

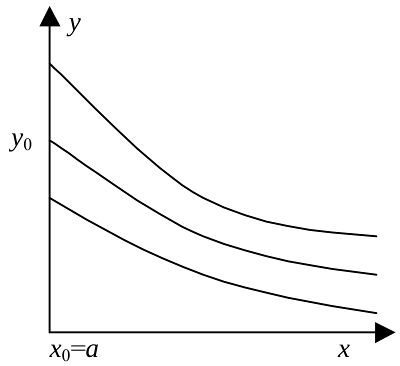
iii. Αν μία ή περισσότερες ιδιοτιμές του πίνακα $\mathbf{J}$ έχουν μηδενικά πραγματικά μέρη και όλες οι υπόλοιπες αρνητικά, τότε η εξίσωση είναι ουδέτερα ευσταθής.

Τα στοιχεία του Ιακωβιανού πίνακα $\mathbf{J}$, είναι συναρτήσεις ως προς x και y. Επομένως οι ιδιοτιμές του μπορεί να μεταβάλλονται όταν μεταβάλλεται το x, συνεπώς η ευστάθεια της διαφορικής εξίσωσης μπορεί να αλλάζει από περιοχή σε περιοχή.

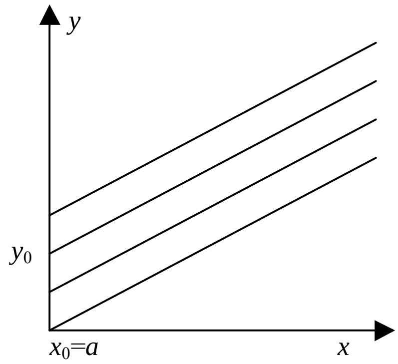
Παράδειγμα 4.1 Η διαφορική εξίσωση $y'=y$, έχει λύση την οικογένεια $y(x)=ce^{x}$, όπου c οποιαδήποτε πραγματική σταθερά. Εδώ η $\partial f/\partial y=1>0$, άρα η εξίσωση είναι ασταθής, και παρατηρούμε από το σχήμα ότι οι καμπύλες των λύσεων αποκλίνουν όταν το x αυξάνει



Η διαφορική εξίσωση $y'=-y$, έχει λύση την οικογένεια $y(x)=ce^{-x}$. Στην περίπτωση αυτή η $\partial f/\partial y=-1<0$, άρα η διαφορική εξίσωση είναι ευσταθής, και παρατηρούμε από το επόμενο σχήμα ότι οι καμπύλες των λύσεων συγκλίνουν μεταξύ τους όταν το x αυξάνει.



Η διαφορική εξίσωση $y'=\alpha$, α σταθερά, έχει σαν λύση την οικογένεια $y(x)=\alpha x+c$. Εδώ $\partial f/\partial y=0$, οπότε η διαφορική εξίσωση είναι ουδέτερα ευσταθής και οι καμπύλες των λύσεων θα είναι παράλληλες μεταξύ τους όπως φαίνεται στο επόμενο σχήμα.



5. Αριθμητικές μέθοδοι – γενικά

Μία αναλυτική λύση μιας διαφορικής εξίσωσης είναι ένας κλειστός τύπος μέσω του οποίου μπορούμε να υπολογίσουμε την τιμή της συνάρτησης της λύσης σε κάθε σημείο x. Αντίθετα, η αριθμητική λύση της διαφορικής εξίσωσης, είναι ένας πίνακας από προσεγγιστικές τιμές της λύσης σε διακεκριμένα σημεία του πεδίου ορισμού της. Μία τέτοια αριθμητική λύση προκύπτει προσομοιάζοντας την συμπεριφορά του συστήματος που αντιστοιχεί στην διαφορική εξίσωση.

Προσεγγιστικές τιμές της λύσης κατασκευάζονται βήμα προς βήμα (step by step) στα διακεκριμένα σημεία κινούμενοι πάνω στο διάστημα στο οποίο ζητάμε την λύση.

Θεωρούμε το πρόβλημα αρχικών τιμών (2.1) και την ακολουθία των διακεκριμένων σημείων $\{x_n\}$, τα οποία ορίζονται από την σχέση:

$$x_n=a+nh,\ n=0,1,\dots,N \quad \text{όπου} \quad h=(b-a)/N,$$

N θετικός ακέραιος. Ο θετικός πραγματικός αριθμός h, ονομάζεται βήμα.

Μία αριθμητική μέθοδος, σε κάθε σημείο της διαμέρισης x_n, υπολογίζει μία προσέγγιση y_n της $y(x_n)$, που είναι η τιμή της αναλυτικής λύσης στο σημείο x_n. Δύο είναι οι βασικές κατηγορίες αριθμητικών μεθόδων για τον υπολογισμό των προσεγγίσεων y_n.

a) **Μέθοδοι απλού βήματος** (one-step methods). Στην περίπτωση αυτή, αν θεωρήσουμε ότι γνωρίζουμε την προσέγγιση y_n στο σημείο x_n, τότε μπορούμε να υπολογίσουμε την προσέγγιση y_{n+1} στο επόμενο σημείο $x_{n+1}=x_n+h$, χρησιμοποιώντας την προηγούμενη προσέγγιση. Στην περίπτωση αυτή η γενική μορφή της μεθόδου μπορεί να εκφραστεί με την σχέση:

$$y_{n+1}=y_n+h\Phi(x_n,\,y_n,\,h) \tag{5.1}$$

όπου $\Phi(x,\,y,\,h)$, μία κατάλληλα υπολογισμένη έκφραση.

b) **Μέθοδοι πολλαπλού βήματος** (multistep methods). Μία μέθοδος αυτής της κατηγορίας, για να προσεγγίσει την λύση σε κάποιο σημείο της διαμέρισης, χρησιμοποιεί πληροφορίες από περισσότερα του ενός προηγούμενα σημεία. Αν η μέθοδος είναι γενικά k-βημάτων, η προσέγγιση y_{n+k} της λύσης στο σημείο x_{n+k}, υπολογίζεται χρησιμοποιώντας πληροφορίες από την προσέγγιση της λύσης στα k προηγούμενα σημεία x_n, $x_{n+1},\ldots,x_{n+k-1}$. Η γενική μορφή μιας μεθόδου k-βημάτων μπορεί να εκφραστεί γενικά από την σχέση:

$$\sum_{j=0}^{k} \alpha_j \, y_{n+j} = h\,\Phi_f\left(x_n\,,\, y_{n+k}\,,\; y_{n+k-1}\,,\ldots,\, y_n\,,\, h\right) \tag{5.2}$$

όπου ο δείκτης f στην Φ δηλώνει ότι η εξάρτηση της Φ από τις προσεγγίσεις y_{n+k}, $y_{n+k-1},\ldots,y_n$ γίνεται μέσω της συνάρτησης $f\left(x,y\right)$, και $\left\{\alpha_j\right\},\; j=0,1,\ldots,k$ είναι προσδιοριστέοι συντελεστές.

Αν η συνάρτηση Φ_f είναι ανεξάρτητη της ζητούμενης προσέγγισης y_{n+k}, η μέθοδος ονομάζεται **άμεση** (explicit), διαφορετικά η μέθοδος ονομάζεται **έμμεση** (implicit).

6. Ακρίβεια των αριθμητικών μεθόδων

Αρχικά θα αναφερθούμε στα διάφορα σφάλματα που προκύπτουν όταν εφαρμόζουμε μία αριθμητική μέθοδο για την προσέγγιση της λύσης του προβλήματος (2.1), και εν συνεχεία θα υπολογίσουμε φράγματα για το βασικό σφάλμα που ακολουθεί μία αριθμητική μέθοδο, θεωρώντας μία απλή μέθοδο μοντέλο.

Γενικά, μπορούμε να πούμε ότι μία αριθμητική μέθοδος για την προσέγγιση της λύσης του προβλήματος (2.1) υποφέρει από δύο βασικές πηγές σφαλμάτων.

1. **Σφάλμα στρογγύλευσης** (rounding error)

Το σφάλμα αυτό οφείλεται στην πεπερασμένη ακρίβεια της αριθμητικής κινητής υποδιαστολής του υπολογιστή. Αν θεωρήσουμε ότι η μέθοδος (5.1) εφαρμόζεται για την προσέγγιση της λύσης του προβλήματος (2.1), και $\overline{y}_n$ είναι η αριθμητική τιμή της προσέγγισης y_n της $y\left(x_n\right)$ στο σημείο x_n, είναι φανερό ότι το σφάλμα στρογγύλευσης ορίζεται από την σχέση:

$$R_{n+1} = \overline{y}_{n+1} - \overline{y}_n - h\,\Phi\left(x_n\,,\,\overline{y}_n\,,\,h\right) \tag{6.1}$$

και μετράει σε κάθε βήμα, πόσο η αριθμητική τιμή της λύσης που προκύπτει από τον υπολογιστή αποτυγχάνει να ικανοποιήσει την μέθοδο.

2. **Σφάλμα αποκοπής** (truncation error)

Το σφάλμα αυτό οφείλεται στην αριθμητική μέθοδο που εφαρμόζουμε, και θα υφίσταται ακόμη και αν όλες οι αριθμητικές πράξεις μπορούν να εκτελεστούν ακριβώς.

Το σφάλμα αποκοπής για την μέθοδο (5.1) ορίζεται από την σχέση:

$$T_{n+1} = y\left(x_{n+1}\right) - y\left(x_n\right) - h\,\Phi\left(x_n\,,\, y\left(x_n\right),\, h\right) \tag{6.2}$$

και μετράει σε ένα βήμα, πόσο αποτυγχάνει η θεωρητική λύση του προβλήματος να ικανοποιήσει την μέθοδο, και μπορεί να θεωρηθεί σαν ένα πρώτο μέτρο για την ακρίβεια της μεθόδου.

Τα δύο αυτά σφάλματα, αν και ξεκινούν από διαφορετικές πηγές, δεν είναι ανεξάρτητα μεταξύ τους. Για παράδειγμα, το σφάλμα αποκοπής συνήθως μπορεί να ελαττωθεί χρησιμοποιώντας ένα μικρότερο h, αλλά αυτό μπορεί να έχει σαν συνέπεια την αύξηση του σφάλματος στογγύλευσης. Στις περισσότερες όμως περιπτώσεις, το σφάλμα αποκοπής αποτελεί τον κυρίαρχο παράγοντα υπολογισμού της ακρίβειας των αριθμητικών λύσεων των συνήθων διαφορικών εξισώσεων, και επομένως μπορούμε να αγνοούμε το σφάλμα στρογγύλευσης.

Στην συνέχεια, θα ασχοληθούμε πιο εμπεριστατωμένα με το σφάλμα αποκοπής της μεθόδου. Το σφάλμα αυτό γενικά μπορεί να χωριστεί σε δύο τύπους. Το **τοπικό** σφάλμα αποκοπής και το **ολικό** σφάλμα αποκοπής.

1. **Τοπικό σφάλμα αποκοπής** T_k (local truncation error), το οποίο είναι το σφάλμα το οποίο γίνεται σε ένα βήμα της αριθμητικής μεθόδου, έστω το k βήμα. Ειδικότερα:

$$T_k = y_k - u_{k-1}\left(x_k\right) \tag{6.3}$$

όπου y_k είναι η προσεγγιστική λύση στο σημείο x_k, και u_{k-1} είναι το μέλος της οικογένειας της αναλυτικής λύσης της διαφορικής εξίσωσης το οποίο διέρχεται από το σημείο $\left(x_{k-1}, y_{k-1}\right)$.

2. **Τοπικό σφάλμα αποκοπής** E_k (global truncation error), το οποίο είναι η διαφορά της προσεγγιστικής λύσης, και της ακριβούς λύσης η οποία ορίζεται από την αρχική συνθήκη στο αρχικό σημείο x_0. Ειδικότερα:

$$E_k = y_k - u_0\left(x_k\right) = y_k - y\left(x_k\right). \tag{6.4}$$

Το ολικό σφάλμα δεν είναι απαραίτητα ίσο με το άθροισμα των τοπικών σφαλμάτων. Το ολικό σφάλμα γενικά θα είναι μεγαλύτερο από το άθροισμα των τοπικών σφαλμάτων, αν η διαφορική εξίσωση είναι ασταθής, αλλά μπορεί να είναι μικρότερο αν η διαφορική εξίσωση είναι ευσταθής.

Είναι φανερό ότι ενδιαφερόμαστε για μικρό ολικό σφάλμα, αλλά μπορούμε να ελέγχουμε απ' ευθείας μόνο το τοπικό σφάλμα.

Ορισμός 6.1 Μία αριθμητική μέθοδος για την προσέγγιση της λύσης του προβλήματος αρχικών τιμών, θα λέμε ότι έχει ακρίβεια τάξης p, αν

$$T_{n+1} = O\left(h^{p+1}\right)$$

όπου T_{n+1} το τοπικό σφάλμα αποκοπής στο σημείο x_{n+1}, που σημαίνει ότι υπάρχει $h_0 > 0$ και σταθερά $C \neq 0$:

$$\left|T_{n+1}\right| \leq C h^{p+1}, \ \forall \ h \ \mu\varepsilon \ 0 < h < h_0. \tag{6.5}$$

7. Ευστάθεια των αριθμητικών μεθόδων

Η έννοια της ευστάθειας για τις αριθμητικές λύσεις μιας συνήθους διαφορικής εξίσωσης είναι ανάλογη, αλλά διαφορετική, από την ευστάθεια της ίδιας της διαφορικής εξίσωσης. Όπως είναι γνωστό, μία διαφορική εξίσωση είναι ευσταθής αν οι καμπύλες των λύσεών της δεν αποκλίνει η μία από την άλλη, όταν το x αυξάνει.

Ομοίως μία αριθμητική μέθοδος είναι **ευσταθής** αν μικρές διαταραχές δεν επιτρέπουν οι αριθμητικές λύσεις που προκύπτουν να αποκλίνει η μία από την άλλη χωρίς κάποιο φράγμα, διαφορετικά θα είναι **ασταθής**.

Τέτοιες αποκλίσεις των αριθμητικών λύσεων μπορεί να οφείλονται είτε στην αστάθεια της διαφορικής εξίσωσης που επιλύεται, είτε και στην ίδια την αριθμητική μέθοδο, ακόμη και αν η διαφορική εξίσωση είναι ευσταθής.

Σε επόμενη παράγραφο, θα αναφερθούμε στην ευστάθεια συγκεκριμένων αριθμητικών μεθόδων και την οποία θα αναλύσουμε περισσότερο.

Μέθοδοι απλού βήματος

8. Γενική μέθοδος απλού βήματος

Είναι ήδη γνωστό ότι μία μέθοδος απλού βήματος προσεγγίζει την λύση του προβλήματος (2.1) στο σημείο x_{n+1}, όταν έχει υπολογιστεί η προσέγγιση y_n στο σημείο x_n, έχει επιλεγεί το βήμα h και δίνεται ο τύπος της διαφορικής εξίσωσης.

Έστω $y(x)$ η ακριβής λύση της διαφορικής εξίσωσης $y' = f(x, y)$. Η τιμή της λύσης στο σημείο $x_{n+1} = x_n + h$ είναι $y(x_{n+1})$, και από το θεώρημα Taylor έχουμε:

$$y(x_{n+1}) = y(x_n + h) = y(x_n) + h\,y'(x_n) + \frac{h^2}{2!}y''(x_n) + \dots$$
$$y(x_n) + h\left\{ y'(x_n) + \frac{h}{2!}y''(x_n) + \frac{h^2}{3!}y'''(x_n) + \dots \right\},$$

ή αν ονομάσουμε:

$$\Delta\big(x_n, y(x_n), h\big) = y'(x_n) + \frac{h}{2!}y''(x_n) + \frac{h^2}{3!}y'''(x_n) + \dots$$

θα προκύψει η σχέση:

$$y(x_{n+1}) = y(x_n) + h\,\Delta\big(x_n, y(x_n), h\big) \tag{8.1}$$

Η παράσταση $\Delta\big(x_n, y(x_n), h\big)$, ονομάζεται ακριβής σχετική αύξηση, και αν η ακριβής λύση είναι ένα πολυώνυμο δεδομένου βαθμού, τότε η σειρά τερματίζεται αλλά αυτό εν γένει δεν συμβαίνει.

Συνεπώς, μία μέθοδος απλού βήματος προκύπτει αν υπολογιστεί μία προσέγγιση της ακριβούς σχετικής αύξησης Δ, δηλαδή μία συνάρτηση Φ όπου:

$$\Phi\big(x_n, y(x_n), h\big) \approx \Delta\big(x_n, y(x_n), h\big) .$$

Τότε η (8.1) γράφεται:

$$y(x_{n+1}) \approx y(x_n) + h\,\Phi\big(x_n, y(x_n), h\big) . \tag{8.2}$$

Η (8.2) δεν είναι ακριβής για την θεωρητική λύση, αλλά γίνεται ακριβής για την αντίστοιχη προσεγγιστική. Συνεπώς, η γενική μορφή μιας μεθόδου απλού βήματος μπορεί να γραφεί:

$$\begin{aligned} &y_{n+1} = y_n + h\,\Phi(x_n, y_n, h) \\ &y_0 \;\; \delta o\theta\varepsilon\nu,\; n = 0, 1, 2, \dots \end{aligned} \tag{8.3}$$

Η προσέγγιση της $\Delta\left(x_n, y\left(x_n\right), h\right)$ με την $\Phi\left(x_n, y\left(x_n\right), h\right)$, εισάγει το τοπικό σφάλμα αποκοπής στο οποίο έχουμε ήδη αναφερθεί, οπότε γενικά ισχύει:

$$T_{n+1}=y\left(x_{n+1}\right)-y\left(x_n\right)-h\Phi\left(x_n, y\left(x_n\right), h\right)$$

$$\left\{y\left(x_n\right)+h\Delta\left(x_n, y\left(x_n\right), h\right)\right\}-\left\{y\left(x_n\right)+h\Phi\left(x_n, y\left(x_n\right), h\right)\right\}$$

ή

$$T_{n+1}=h\left\{\Delta\left(x_n, y\left(x_n\right), h\right)-\Phi\left(x_n, y\left(x_n\right), h\right)\right\} \tag{8.4}$$

9. Μέθοδος Euler

Θεωρούμε το πρόβλημα αρχικών τιμών:

$$y'=f\left(x,y\right),\ x\in\left[a,b\right],\ y\left(a\right)=y_0$$

και την διαμέριση του $\left[a,b\right]$ με τα σημεία $x_n=a+nh,\ n=0,1,2,...,N$ με $h=\left(b-a\right)/N$, N θετικός ακέραιος.

Η πιο απλή μέθοδος απλού βήματος η οποία μπορεί να υπολογίζει διαδοχικά τις προσεγγίσεις $y_1,\ y_2,...,y_N$ στα σημεία $x_1,\ x_2,...,x_N$ της διαμέρισης, της λύσης $y\left(x\right)$ του παραπάνω προβλήματος, έχει προταθεί από τον Euler, και μπορεί να υπολογιστεί με διάφορους απλούς τρόπους

a) **Ολοκλήρωση**

Ολοκληρώνουμε την $y'=f\left(x,y\right)$ στο διάστημα $\left[x_n,\ x_{n+1}\right]$, οπότε προκύπτει:

$$\int_{x_n}^{x_{n+1}} y'\left(x\right)dx=\int_{x_n}^{x_{n+1}} f\left(x,\ y\left(x\right)\right)dx$$

και ισχύει για κάθε $n=0,1,2,...,N-1$. Σύμφωνα με την θεωρία της ολοκλήρωσης έχουμε:

$$y\left(x_{n+1}\right)-y\left(x_n\right)=\int_{x_n}^{x_{n+1}} f\left(x,\ y\left(x\right)\right)dx$$

ή

$$y\left(x_{n+1}\right)=y\left(x_n\right)+\int_{x_n}^{x_{n+1}} f\left(x,\ y\left(x\right)\right)dx,\quad n=0,1,2,...,N-1 \tag{9.1}$$

Το ολοκλήρωμα στο δεύτερο μέλος της (9.1) το προσεγγίζουμε με τον κανόνα του ορθογωνίου, ή η $f\left(x,y\left(x\right)\right)$ προσεγγίζεται με πολυώνυμο βαθμού μηδέν. Τότε ισχύει ότι:

$$\int_{x_n}^{x_{n+1}} f\left(x,\ y\left(x\right)\right)dx\approx f\left(x_n, y\left(x_n\right)\right)$$

οπότε προκύπτει η σχέση:

$$y\left(x_{n+1}\right)\approx y\left(x_n\right)+h\,f\left(x_n, y\left(x_n\right)\right),\ n=0,1,...,N-1,\ y\left(a\right)=y_0.$$

Η σχέση αυτή δεν είναι ακριβής για τις τιμές της θεωρητικής λύσης στα σημεία της διαμέρισης, αλλά γίνεται ακριβής για τις αντίστοιχες προσεγγίσεις. Συνεπώς η μέθοδος Euler που προκύπτει γράφεται:

$$\begin{aligned} & y_{n+1} = y_n + h\, f\left(x_n, y_n\right) \\ & n = 0,1,\ldots,N-1,\ y\left(a\right) = y_0 \end{aligned} \tag{9.2}$$

και ονομάζεται άμεση μέθοδος Euler.

b) **Ανάπτυγμα Taylor**

Θεωρούμε το ανάπτυγμα Taylor για την συνάρτηση $y\left(x\right)$ στο σημείο $x+h$, οπότε έχουμε:

$$y\left(x+h\right) = y\left(x\right) + h\, y'\left(x\right) + \frac{h^2}{2!} y''\left(x\right) + \ldots\ ,$$

και κρατάμε τους δύο πρώτους όρους οπότε:

$$y\left(x+h\right) \approx y\left(x\right) + h\, y'\left(x\right)$$

ή

$$y\left(x+h\right) \approx y\left(x\right) + h\, f\left(x, y\left(x\right)\right)\ .$$

Αν θεωρήσουμε τα διαδοχικά σημεία x_n, x_{n+1} της διαμέρισης, η προηγούμενη σχέση γράφεται:

$$y\left(x_{n+1}\right) \approx y\left(x_n\right) + h\, f\left(x_n, y\left(x_n\right)\right)$$

η οποία γίνεται ακριβής για τις αντίστοιχες προσεγγίσεις, οπότε προκύπτει η μέθοδος (9.2).

Παράδειγμα 9.1 Θεωρούμε την απλή διαφορική εξίσωση $y' = y$, η οποία έχει την λύση $y\left(x\right) = c\, e^x$. Θα εφαρμόσουμε την άμεση μέθοδο Euler για να τη επιλύσουμε αριθμητικά. Για κάποιο h, θα προωθήσουμε την λύση από το σημείο $x_0 = 0$, στο σημείο $x_1 = x_0 + h$ οπότε προκύπτει η σχέση:

$$y_1 = y_0 + h\, y'_0 = y_0 + h\, y_0 = \left(1+h\right) y_0\, .$$

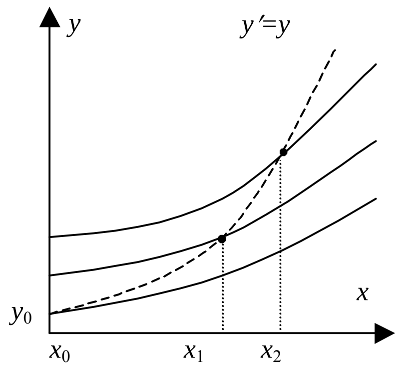
Η τιμή της λύσης στο σημείο x_1 δεν είναι ακριβής, δηλαδή $y_1 \neq y\left(x_1\right)$. Έτσι αν $x_0 = 0$, $y_0 = 1$ και $h = 0.5$ τότε έχουμε $y_1 = 1.5$, ενώ η ακριβής λύση είναι $y\left(x_1\right) = y\left(0.5\right) \approx 1.649$.

Από τον υπολογισμό της μεθόδου με την σειρά Taylor μπορούμε να διαπιστώσουμε ότι το σφάλμα της προσέγγισης είναι ανάλογο του h^2. Επομένως, αν μειώσουμε το βήμα κατά 1/2 το σφάλμα θα μειωθεί κατά 1/4 αν θεωρήσουμε ότι το σφάλμα στρογγύλευσης είναι αμελητέο. Για κάθε μη μηδενικό σφάλμα, η προσέγγιση y_1 βρίσκεται σε διαφορετικό μέλος της οικογένειας των λύσεων από εκείνο από το οποίο ξεκινήσαμε.

Για να συνεχίσουμε την διαδικασία επίλυσης, παίρνουμε το επόμενο σημείο $x_2 = 1.0$. Μετά τις πράξεις θα πάρουμε:

$$y_2 = y_1 + h\, y_1 = 1.5 + \left(1.5\right)\left(0.5\right) = 2.25\, .$$

Η ακριβής λύση στο $x_2 = 1.0$ είναι $y(x_2) = y(1.0) \approx 2.718$. Η προσέγγιση $y_2 = 2.25$ διαφέρει από την $y(x_2)$, αλλά διαφέρει επίσης και από την καμπύλη της λύσης η οποία διέρχεται από το προηγούμενο σημείο (x_1, y_1) η οποία στο σημείο $x = 1$ έχει την προσεγγιστική τιμή $y = 2.473$. Έτσι μετακινούμεθα σε ένα ακόμη μέλος της οικογένειας των λύσεων της διαφορικής εξίσωσης.



Αν συνεχίσουμε την διαδικασία εκτελώντας και άλλα βήματα, θα δημιουργήσουμε έναν πίνακα διακεκριμένων τιμών της προσεγγιστικής λύσης. Έτσι όμως σε κάθε βήμα θα πηγαίνουμε από το ένα μέλος της οικογένειας των λύσεων στο άλλο. Για την συγκεκριμένη ασταθή διαφορική εξίσωση, τα σφάλματα που γίνονται στην αριθμητική μέθοδο μεγαλώνουν από βήμα σε βήμα, σαν αποτέλεσμα της απόκλισης των καμπύλων της λύσης.

10. Γενίκευση μεθόδου Euler

Η διαδικασία υπολογισμού της μεθόδου Euler με ολοκλήρωση, μπορεί να γενικευθεί αντικαθιστώντας τον κανόνα του ορθογωνίου για την προσέγγιση του ολοκληρώματος, με μία μονοπαραμετρική οικογένεια κανόνων ολοκλήρωσης της μορφής:

$$\int_{x_n}^{x_{n+1}} f\left(x,\, y(x)\right) dx \approx h\left\{(1-\theta)\, f\left(x_n, y(x_n)\right) + \theta\, f\left(x_{n+1}, y(x_{n+1})\right)\right\} \tag{10.1}$$

με $\theta \in [0,1]$ την παράμετρο. Τότε η σχέση:

$$y(x_{n+1}) = y(x_n) + \int_{x_n}^{x_{n+1}} f\left(x,\, y(x)\right) dx, \quad n = 0,1,2,\ldots,N-1$$

γράφεται

$$\begin{aligned} & y(x_{n+1}) \approx y(x_n) + h\left\{(1-\theta)\, f\left(x_n, y(x_n)\right) + \theta\, f\left(x_{n+1}, y(x_{n+1})\right)\right\} \\ & y(a) = y_0, \quad n = 0,1,\ldots,N-1 \end{aligned} \tag{10.2}$$

από την οποία προκύπτει η ονομαζόμενη **θ-μέθοδος** της μορφής:

$$\begin{aligned} & y_{n+1} = y_n + h\left\{(1-\theta)\, f\left(x_n, y_n\right) + \theta\, f\left(x_{n+1}, y_{n+1}\right)\right\} \\ & y(a) = y_0, \quad n = 0,1,\ldots,N-1, \ \theta \in [0,1] \end{aligned} \tag{10.3}$$

Για συγκεκριμένες τιμές της σταθεράς θ προκύπτει και μία αντίστοιχη μέθοδος της οικογένειας αυτής.

a) Για $\theta=0$ προκύπτει η απλή ή άμεση μέθοδος Euler (9.2).

b) Για $\theta=1$ προκύπτει η μέθοδος,

$$\begin{aligned} &y_{n+1}=y_n+h\, f\left(x_{n+1}, y_{n+1}\right) \\ &y(a)=y_0,\ n=0,1,\ldots N-1 \end{aligned} \tag{10.4}$$

η οποία ονομάζεται **έμμεση μέθοδος Euler.**

Η εφαρμογή αυτής της μεθόδου απαιτεί την επίλυση μιας, έμμεσης εξίσωσης για τον υπολογισμό του y_{n+1}, δοθέντος του y_n, το οποίο συνεπάγεται ενδιάμεσες διαδικασίες απαραίτητες σε τέτοιες περιπτώσεις.

Η μέθοδος (10.4) είναι έμμεση επειδή πρέπει να υπολογίσουμε την τιμή της f στο σημείο y_{n+1} πριν το υπολογίσουμε. Αυτό σημαίνει ότι μία τιμή για το y_{n+1} η οποία ικανοποιεί την (10.4) πρέπει να υπολογιστεί. Αν η f είναι μία μη γραμμική συνάρτηση ως προς y, όπως είναι συνήθως, τότε θα πρέπει να εφαρμόσουμε μία επαναληπτική μέθοδο, όπως η μέθοδος σταθερού σημείου ή η μέθοδος Newton. Μία καλή αρχική τιμή για την επαναληπτική μέθοδο μπορεί να υπολογιστεί από μία άμεση μέθοδο, όπως η απλή μέθοδος Euler, ή από την λύση στο προηγούμενο βήμα.

c) Για $\theta=1/2$ προκύπτει μία ενδιαφέρουσα επίσης μέθοδος της μορφής:

$$\begin{aligned} &y_{n+1}=y_n+\frac{h}{2}\left\{ f\left(x_n, y_n\right)+f\left(x_{n+1}, y_{n+1}\right)\right\} \\ &y(a)=y_0,\ n=0,1,\ldots N-1 \end{aligned} \tag{10.5}$$

η οποία ονομάζεται **μέθοδος τραπεζίου**.

Είναι φανερό ότι η μέθοδος του τραπεζίου υπολογίζει τις προσεγγίσεις της λύσης στα διάφορα σημεία της διαμέρισης με μεγαλύτερη ακρίβεια από την άμεση μέθοδο Euler. Όμως είναι ολιγότερο εύχρηστη, επειδή απαιτεί την επίλυση μιας έμμεσης μεθόδου σε κάθε σημείο x_{n+1} για τον υπολογισμό της y_{n+1}. Μία εύκολη αντιμετώπιση αυτής της δυσκολίας είναι να χρησιμοποιήσουμε την άμεση μέθοδο Euler για μία πρώτη προσέγγιση της τιμής $y\left(x_{n+1}\right)$, και εν συνεχεία να χρησιμοποιήσουμε αυτή την προσέγγιση μέσα στον τύπο του τραπεζίου για να υπολογίσουμε μία περισσότερο ακριβή προσέγγιση για το $y\left(x_{n+1}\right)$. Τότε προκύπτει η μέθοδος:

$$\begin{aligned} &y_{n+1}=y_n+\frac{h}{2}\left\{ f\left(x_n, y_n\right)+f\left(x_{n+1}, y_n+h\, f\left(x_n, y_n\right)\right)\right\} \\ &y(a)=y_0,\ n=0,1,\ldots N-1 \end{aligned} \tag{10.6}$$

η οποία ονομάζεται **βελτιωμένη μέθοδος Euler.**

Η προηγούμενη μέθοδος μπορεί να εκφραστεί και ως εξής:

$$
\begin{aligned}
&k_1 = f\left(x_n, y_n\right)\\
&k_2 = f\left(x_n + h, y_n + h k_1\right)\\
&y_{n+1} = y_n + \frac{h}{2}\left(k_1 + k_2\right)\\
&y(a) = y_0, \quad n = 0, 1, \ldots, N-1
\end{aligned}
\tag{10.7}
$$

και είναι μία έκφραση συμβατή με την οικογένεια των μεθόδων Runge-Kutta που θα μελετήσουμε σε επόμενες παραγράφους.

Συμπερασματικά, η θ-μέθοδος για $\theta = 0$ ορίζει μία άμεση μέθοδο, και για $0 < \theta \le 1$ ορίζει έμμεσες μεθόδους απλού βήματος. Στην συνέχεια θα εξετάσουμε βασικές έννοιες της ποιοτικής θεωρίας των αριθμητικών μεθόδων απλού βήματος βασιζόμενοι σε μερικά απλά μέλη αυτής της οικογένειας.

11. Τάξη ακρίβειας αριθμητικής μεθόδου

Στην παράγραφο αυτή θα εξετάσουμε την τάξη ακρίβειας της απλής μεθόδου Euler.

Θεωρούμε το ανάπτυγμα Taylor της μορφής:

$$y(x+h) = y(x) + h\, y'(x) + O\left(h^2\right),$$

όπου το $O\left(h^2\right)$ συμβολίζει ότι παραλείπουμε από την σειρά Taylor όλους τους όρους που εξαρτώνται από δυνάμεις του h^2 και πάνω. Για $x = x_k$, η παραπάνω σχέση γράφεται:

$$y\left(x_k + h\right) = y\left(x_k\right) + h\, y'\left(x_k\right) + O\left(h^2\right)$$

ή

$$y\left(x_{k+1}\right) = y\left(x_k\right) + h\, f\left(x_k, y\left(x_k\right)\right) + O\left(h^2\right). \tag{11.1}$$

Η μέθοδος Euler ως γνωστόν ορίζεται από την σχέση:

$$y_{k+1} = y_k + h\, f\left(x_k, y_k\right), \tag{11.2}$$

και αφαιρώντας από την (11.2) την (11.1) προκύπτει η σχέση:

$$y_{k+1} - y\left(x_{k+1}\right) = \left\{y_k - y\left(x_k\right)\right\} + h\left\{f\left(x_k, y_k\right) - f\left(x_k, y\left(x_k\right)\right)\right\} - O\left(h^2\right).$$

Αν υποθέσουμε ότι δεν υπάρχουν σφάλματα στα προηγούμενα σημεία, θα ισχύει η σχέση $y_k = y\left(x_k\right)$, οπότε:

$$y_k - y\left(x_k\right) = 0 \quad \text{και} \quad f\left(x_k, y_k\right) - f\left(x_k, y\left(x_k\right)\right) = 0,$$

άρα προκύπτει ότι:

$$y_{k+1} - y\left(x_{k+1}\right) = O\left(h^2\right),$$

ή

$$T_{k+1} = O\left(h^2\right),$$

T_{k+1} το τοπικό σφάλμα αποκοπής, άρα η μέθοδος Euler είναι 1ης τάξης.

12. Ευστάθεια αριθμητικής μεθόδου

Στην παράγραφο αυτή θα μελετήσουμε την ευστάθεια των αριθμητικών μεθόδων, δηλαδή την συμπεριφορά της μετάδοσης του σφάλματος από βήμα σε βήμα, θεωρώντας αρχικά την άμεση μέθοδο Euler.

Το ολικό σφάλμα είναι το άθροισμα των τοπικών σφαλμάτων σε κάθε βήμα, και μπορεί να αποκαλείται το **μεταδιδόμενο σφάλμα** (propagated error). Για να μελετήσουμε την ευστάθεια μιας αριθμητικής μεθόδου, θα πρέπει να εξετάσουμε το αντίστοιχο μεταδιδόμενο σφάλμα.

Για την περίπτωση της μεθόδου Euler, θεωρούμε την σχέση:

$$y_{k+1} - y\left(x_{k+1}\right) = \left\{ y_k - y\left(x_k\right) \right\} + h_k \left\{ f\left(x_k, y_k\right) - f\left(x_k, y\left(x_k\right)\right) \right\} - O\left(h^2\right). \quad (12.1)$$

Ο δείκτης k στο βήμα υποδηλώνει ότι το βήμα μπορεί να είναι μεταβλητό και όχι σταθερό από σημείο σε σημείο.

Με βάση το θεώρημα της μέσης τιμής ισχύει η σχέση:

$$f\left(x_k, y_k\right) - f\left(x_k, y\left(x_k\right)\right) = J(\xi)\left(y_k - y\left(x_k\right)\right)$$

όπου J γενικά μπορεί να είναι ο Ιακωβιανός πίνακας της συνάρτησης f, και ειδικότερα η παράγωγος της f σε κάποιο άγνωστο σημείο ξ. Τότε η σχέση (12.1) γράφεται:

$$y_{k+1} - y\left(x_{k+1}\right) = \left(y_k - y\left(x_k\right)\right) + h_k J(\xi)\left(y_k - y\left(x_k\right)\right) - O\left(h^2\right)$$

ή

$$E_{k+1} = \left(1 + h_k J(\xi)\right) E_k + T_{k+1} \quad (12.2)$$

όπου E_k σημαίνει ολικό σφάλμα και T_k τοπικό σφάλμα αποκοπής. Η σχέση (12.2) δείχνει ότι το ολικό σφάλμα πολλαπλασιάζεται σε κάθε βήμα με τον παράγοντα $\left(1 + h_k J(\xi)\right)$, ο οποίος ονομάζεται **παράγοντας αύξησης** του σφάλματος (growth or amplification factor).

Επομένως, αν

$$\left| 1 + h_k J \right| < 1$$

τότε τα σφάλματα δεν θα αυξάνουν, και η μέθοδος λέμε ότι είναι **ευσταθής.** Αν αυτό δεν ισχύει, τα σφάλματα θα αυξάνουν από βήμα σε βήμα, και τότε θα λέμε ότι η μέθοδος είναι **ασταθής.**

Η προηγούμενη ανισότητα ισοδυναμεί με:

$$-1 < 1 + h_k J < 1 \Leftrightarrow h_k J \in (-2, 0).$$

Η τελευταία σχέση συνεπάγεται ότι για να έχουμε ομαλή μετάδοση του σφάλματος από βήμα σε βήμα, θα πρέπει να χρησιμοποιούμε κατάλληλο βήμα h_k, έτσι ώστε το $h_k\,J$, να βρίσκεται πάντα μέσα στο διάστημα $(-2,0)$, το οποίο ονομάζεται **διάστημα ευστάθειας** της μεθόδου. Είναι προφανές ότι η τιμή του h_k εξαρτάται και από την τιμή του J, η οποία εξαρτάται από την ίδια την διαφορική εξίσωση.

Σημειώνουμε, ότι αυτού του είδους η αστάθεια μπορεί να συμβεί όταν η διαφορική εξίσωση είναι ασταθής $(J>0)$, αλλά μπορεί επίσης να συμβεί για ευσταθή διαφορική εξίσωση $(J<0)$, αν $h_k>-2/J$.

Για ένα σύστημα διαφορικών εξισώσεων, ο παράγοντας αύξησης του σφάλματος ορίζεται από τον πίνακα, $\mathbf{I}+h_k\,\mathbf{J}$, όπου $\mathbf{J}$ είναι ο Ιακωβιανός πίνακας της συνάρτησης $\mathbf{f}$, και $\mathbf{I}$ ο αντίστοιχος μοναδιαίος. Στην περίπτωση αυτή θα λέμε ότι η αριθμητική μέθοδος είναι ευσταθής αν ισχύει η συνθήκη:

$$\rho\left(\mathbf{I}+h_k\,\mathbf{J}\right)<1,$$

η οποία ικανοποιείται αν οι ιδιοτιμές του πίνακα $h_k\,\mathbf{J}$, βρίσκονται μέσα στον κύκλο του μιγαδικού επιπέδου $\mathrm{K}(-1,1)$. Το διάστημα $(-2,0)$ που υπολογίσαμε στην περίπτωση μιας απλής διαφορικής εξίσωσης, θα είναι η τομή του παραπάνω μοναδιαίου κύκλου με τον αρνητικό ημιάξονα. Το χωρίο το οποίο ορίζεται με τον κύκλο αυτόν συνήθως ονομάζεται **χωρίο ευστάθειας** της μεθόδου.

Γενικά, ο παράγοντας αύξησης του σφάλματος εξαρτάται από την διαφορική εξίσωση η οποία ορίζει τον Ιακωβιανό πίνακα, την αριθμητική μέθοδο που χρησιμοποιούμε η οποία ορίζει την μορφή του παράγοντα, και το βήμα h_k.

Ένας εναλλακτικός τρόπος ανάλυσης για την μελέτη της ακρίβειας και της ευστάθειας μιας αριθμητικής μεθόδου, βασίζεται στην εφαρμογή της μεθόδου σε ένα γραμμικό πρόβλημα μοντέλο της μορφής:

$$y'=\lambda\,y,\ \ y(0)=y_0 \tag{12.3}$$

του οποίου η ακριβής λύση είναι η συνάρτηση $y(x)=y_0\,e^{\lambda x}$.

Για να εξετάσουμε την ακρίβεια μιας μεθόδου θα πρέπει να συγκρίνουμε τις προσεγγιστικές λύσεις με τις αντίστοιχες αναλυτικές βασιζόμενοι στο ανάπτυγμα Taylor αυτών. Για να μελετήσουμε την ευστάθεια μιας μεθόδου, θα πρέπει να υπολογίσουμε τον παράγοντα αύξησης του σφάλματος της αριθμητικής λύσης, και να επιβάλλουμε σε αυτόν τις κατάλληλες συνθήκες ώστε η αριθμητική λύση να συμπεριφέρεται με τρόπο ανάλογο της αναλυτικής. Τα συμπεράσματα που προκύπτουν με βάση το γραμμικό μοντέλο θεωρούμε ότι ισχύουν και για την περίπτωση της επίλυσης μιας οποιασδήποτε διαφορικής εξίσωσης.

Θα εφαρμόσουμε την ανάλυση αυτή για να μελετήσουμε την ευστάθεια και ακρίβεια της απλής μεθόδου Euler. Αν εφαρμόσουμε την μέθοδο αυτή στο πρόβλημα (12.3), θα προκύψει ότι:

$$y_{k+1}=y_k+h\left(\lambda\,y_k\right)=\left(1+h\,\lambda\right)y_k,\ \ k\ge 0. \tag{12.4}$$

Αν εφαρμόσουμε την (12.4) διαδοχικά για τις διάφορες τιμές του k, θα προκύψει η σχέση:

$$y_k=\left(1+h\lambda\right)^k y_0,\quad k=1,2,\ldots \tag{12.5}$$

Αν $\lambda<0$, η ακριβής λύση συγκλίνει εκθετικά στο μηδέν όταν το $x\to\infty$. Για να ισχύει κάτι ανάλογο και για την προσεγγιστική λύση, θα πρέπει λόγω της (12.5) να ισχύει η ανισότητα $\left|1+h\lambda\right|<1$. Αυτό το αποτέλεσμα συμφωνεί με την προηγούμενη ανάλυση της μεθόδου Euler, επειδή για την διαφορική εξίσωση (12.3), $J=\lambda$. Με βάση την παραπάνω συνθήκη η ευστάθεια της μεθόδου είναι εξασφαλισμένη αν $h\lambda\in\left(-2,0\right)$.

Για να εξετάσουμε την τάξη ακρίβειας της μεθόδου, συγκρίνουμε τον παράγοντα αύξησης του σφάλματος της προσεγγιστικής λύσης με το ανάπτυγμα Taylor του όρου $e^{h\lambda}$ της αναλυτικής λύσης. Ο παράγοντας αύξησης του σφάλματος της μεθόδου είναι $1+h\lambda$, και μπορεί να συμφωνήσει με το ανάπτυγμα της σειράς:

$$e^{h\lambda}=1+\left(h\lambda\right)+\frac{\left(h\lambda\right)^2}{2!}+\frac{\left(h\lambda\right)^3}{3!}+\ldots$$

μόνο στους δύο πρώτους όρους, δηλαδή όρους που εξαρτώνται από την πρώτη δύναμη του h, και επομένως συνάγεται εύκολα ότι η τάξη ακρίβειας της μεθόδου είναι 1.

13. Ανάλυση σφάλματος της μεθόδου Euler

Στην παράγραφο αυτή θα εξετάσουμε την σχέση του ολικού σφάλματος μιας αριθμητικής μεθόδου και του βήματος ολοκλήρωσης, και πως συνδέεται η μείωση του ολικού σφάλματος με την μείωση του βήματος.

Η ανάλυση θα βασιστεί στην μέθοδο Euler, και τα αντίστοιχα συμπεράσματα για την θ-μέθοδο μπορούν να εξαχθούν βασιζόμενοι στα συμπεράσματα της μεθόδου Euler.

Θεωρούμε την άμεση μέθοδο Euler,

$$y_{n+1}=y_n+h\,f\left(x_n,y_n\right),\quad n=0,1,\ldots,N-1,\quad y\left(a\right)=y_0. \tag{13.1}$$

Είναι γνωστό ότι αν αντικαταστήσουμε τις προσεγγιστικές τιμές με τις αντίστοιχες αναλυτικές, θα ισχύει:

$$y\left(x_{n+1}\right)\neq y\left(x_n\right)+h\,f\left(x_n,y\left(x_n\right)\right)$$

ή

$$\frac{y\left(x_{n+1}\right)-y\left(x_n\right)}{h}-f\left(x_n,y\left(x_n\right)\right)\neq 0.$$

Η ποσότητα

$$T_n=\frac{y\left(x_{n+1}\right)-y\left(x_n\right)}{h}-f\left(x_n,y\left(x_n\right)\right) \tag{13.2}$$

ορίζει το σφάλμα αποκοπής της μεθόδου Euler, και θα παίξει ρόλο κλειδί στην ανάλυση. Είναι γνωστό επίσης ότι η ποσότητα αυτή αντιστοιχεί στην "ποσότητα" η οποία μετράει πόσο η αναλυτική λύση αποτυγχάνει να ικανοποιήσει την μέθοδοEuler.

Σύμφωνα με το θεώρημα Taylor, υπάρχει $\xi_n \in \left(x_n, x_{n+1}\right)$:

$$y\left(x_{n+1}\right) = y\left(x_n\right) + h\,y'\left(x_n\right) + \frac{1}{2}h^2 y''\left(\xi_n\right)$$

οπότε η (13.2) γράφεται:

$$T_n = \frac{1}{h}\left\{y\left(x_n\right) + h\,y'\left(x_n\right) + \frac{1}{2}h^2 y''\left(\xi_n\right) - y\left(x_n\right)\right\} - y'\left(x_n\right)$$

ή

$$T_n = \frac{1}{2}h\,y''\left(\xi_n\right) \tag{13.3}$$

όπου υποθέτουμε ότι η f είναι μία λεία συνάρτηση δύο μεταβλητών, ώστε να υπάρχει η y'' και να είναι φραγμένη στο $\left[a, b\right]$.

Η μέθοδος Euler μπορεί να γραφεί:

$$0 = \frac{y_{n+1} - y_n}{h} - f\left(x_n, y_n\right) \tag{13.4}$$

Αφαιρούμε κατά μέλη τις σχέσεις (13.2) και (13.4) οπότε τελικά προκύπτει η σχέση:

$$\left\{y\left(x_{n+1}\right) - y\left(x_n\right)\right\} - \left\{y_{n+1} - y_n\right\} - h\left\{f\left(x_n, y\left(x_n\right)\right) - f\left(x_n, y_n\right)\right\} = hT_n$$

ή

$$y\left(x_{n+1}\right) - y_{n+1} = y\left(x_n\right) - y_n + h\left\{f\left(x_n, y\left(x_n\right)\right) - f\left(x_n, y_n\right)\right\} + hT_n$$

ή

$$E_{n+1} = E_n + h\left\{f\left(x_n, y\left(x_n\right)\right) - f\left(x_n, y_n\right)\right\} + hT_n$$

ή

$$\left|E_{n+1}\right| \le \left|E_n\right| + h\left|f\left(x_n, y\left(x\right)\right) - f\left(x_n, y_n\right)\right| + h\left|T_n\right| \tag{13.5}$$

Υποθέτουμε ότι $\left|y_n - y_0\right| < Y_m$, οπότε από την συνθήκη Lipschitz προκύπτει ότι:

$$\left|f\left(x_n, y\left(x\right)\right) - f\left(x_n, y_n\right)\right| \le L\left|y\left(x_n\right) - y_n\right| = L\left|E_n\right|,$$

και επομένως η σχέση (13.5) γράφεται:

$$\left|E_{n+1}\right| \le \left|E_n\right| + hL\left|E_n\right| + h\left|T_n\right|$$

ή

$$\left|E_{n+1}\right| \le \left(1 + hL\right)\left|E_n\right| + h\left|T_n\right|, \; n = 0, 1, \ldots, N-1. \tag{13.6}$$

Έστω ότι $T = \max\limits_{0 \le n \le N-1}\left|T_n\right|$, οπότε η (13.6) γράφεται:

$$\left|E_{n+1}\right| \le (1+hL)\left|E_n\right| + h\left|T\right|, \; n=0,1,\dots,N-1 \tag{13.7}$$

Μπορεί να δειχθεί επαγωγικά ότι:

$$\left|E_n\right| \le \frac{T}{L}\left\{(1+hL)^n - 1\right\} + (1+hL)^n \left|E_0\right| \tag{13.8}$$

Είναι φανερό ότι $1+hL \le e^{hL}$, οπότε $(1+hL)^n \le e^{nhL}$. Όμως ισχύει ότι $nh = x_n - x_0$, άρα $(1+hL)^n \le e^{L(x_n - x_0)}$ και η (13.8) γράφεται:

$$\begin{aligned}&\left|E_n\right| \le \frac{T}{L}\left\{e^{L(x_n-x_0)} - 1\right\} + e^{L(x_n-x_0)}\left|E_0\right| \\ &n=1,2,\dots,N\end{aligned} \tag{13.9}$$

Αν υποθέσουμε ότι $M_2 = \max_{x\in[\alpha,\beta]}\left|y''(x)\right|$, οπότε $\left|T\right| \le \frac{1}{2}hM_2$, τότε η (13.9) γράφεται:

$$\begin{aligned}&\left|E_n\right| \le \frac{hM_2}{2L}\left\{e^{L(x_n-x_0)} - 1\right\} + e^{L(x_n-x_0)}\left|E_0\right| \\ &n=1,2,\dots,N\end{aligned} \tag{13.10}$$

και έτσι ορίζεται ένα φράγμα για το ολικό σφάλμα. Σύμφωνα με την παραπάνω ανάλυση μπορούμε να διατυπώσουμε το επόμενο θεώρημα.

Θεώρημα 13.1 Θεωρούμε το πρόβλημα αρχικών τιμών,

$$y' = f(x,y), \; x\in[\alpha,\beta], \; y(\alpha) = y_0$$

για το οποίο υποθέτουμε ότι ισχύει το θεώρημα Picard, και την ακολουθία των διακεκριμένων σημείων,

$$x_n = \alpha + nh, \; n=0,1,...,N \; \mu\varepsilon \; h = (\beta - \alpha)/N .$$

Επίσης υποθέτουμε ότι η συνάρτηση f, είναι μία λεία συνάρτηση δύο μεταβλητών ώστε να υπάρχει η y'', και να είναι φραγμένη στο διάστημα $[\alpha,\beta]$. Να δειχθεί για την άμεση μέθοδο Euler, ότι το ολικό σφάλμα E_n, φράσσεται από την σχέση:

$$\left|E_n\right| \le e^{L(x_n-x_0)}\left|E_0\right| + \frac{T}{L}\left\{e^{L(x_n-x_0)} - 1\right\}, \; n=1,2,...,N$$

όπου L είναι η σταθερά Lipschitz, και $T = \max_{0\le n\le N-1}\left|T_n\right|$, όπου T_n είναι τα τοπικά σφάλματα αποκοπής.

Για την γενική θ-μέθοδο μπορεί να δειχθεί ότι η αντίστοιχη σχέση της (13.10) είναι:

$$\begin{aligned}&\left|E_n\right| \le \frac{h}{L}\left\{\left|\tfrac{1}{2} - \theta\right| M_2 + \tfrac{1}{3}hM_3\right\}\left\{e^{L\left(\frac{x_n-x_0}{1-\theta Lh}\right)} - 1\right\} + e^{L\left(\frac{x_n-x_0}{1-\theta Lh}\right)}\left|E_0\right| \\ &n=1,2,\dots,N\end{aligned} \tag{13.11}$$

όπου $M_3 = \max_{x\in[\alpha,\beta]}\left|y'''(x)\right|$.

Θεωρούμε ότι δεν υπάρχει σφάλμα στρογγύλευσης στο αρχικό σημείο, $y(a)=y(x_0)=y_0$, οπότε μπορούμε να υποθέσουμε ότι $|E_0|=y(x_0)-y_0=0$. Τότε, από την (13.11) μπορούμε να διαπιστώσουμε ότι $|E_n|=O(h^2)$ για $\theta=1/2$, ενώ για $\theta=0$ και $\theta=1$, $|E_n|=O(h)$. Το ίδιο ισχύει για κάθε $\theta=\pm 1/2$.

Από την παραπάνω ανάλυση μπορούμε να διαπιστώσουμε ότι κάθε φορά που το βήμα h υποδιπλασιάζεται, το σφάλμα αποκοπής και το ολικό σφάλμα μειώνεται κατά ένα παράγοντα ίσο με το 2 όταν $\theta\neq 1/2$, και κατά ένα παράγοντα ίσο με το 4 όταν $\theta = 1/2$. Συνεπώς, η μέθοδος του Τραπεζίου οδηγεί σε ακριβέστερες προσεγγίσεις από την μέθοδο Euler, αλλά είναι περισσότερο πολύπλοκή στην εφαρμογή της διότι είναι έμμεση.

14. Ανάλυση σφάλματος της γενικής μεθόδου απλού βήματος

Στην παράγραφο 8 ορίσαμε ότι η γενική μέθοδος απλού βήματος έχει την μορφή:

$$\begin{aligned} &y_{n+1}=y_n+h\Phi(x_n,y_n,h) \\ &y(a)=y_0,\ n=0,1,\dots,N-1 \end{aligned} \tag{14.1}$$

όπου η συνάρτηση $\Phi(.,.,.)$ είναι συνεχής ως προς τις μεταβλητές της.

Το ολικό σφάλμα ως γνωστόν ορίζεται από την σχέση,

$$E_n=y(x_n)-y_n \tag{14.2}$$

και το σφάλμα αποκοπής μπορεί να γραφεί στην μορφή:

$$T_n=\frac{y(x_{n+1})-y(x_n)}{h}-\Phi\big(x_n,y(x_n),h\big). \tag{14.3}$$

Το επόμενο θεώρημα υπολογίζει ένα φράγμα για το ολικό σφάλμα συναρτήσει του τοπικού σφάλματος.

Θεώρημα 14.1 Θεωρούμε την γενική μέθοδο απλού βήματος (14.1), όπου η συνάρτηση Φ υποθέτουμε ότι είναι συνεχής ως προς τις μεταβλητές της. Επίσης υποθέτουμε ότι η συνάρτηση αυτή ικανοποιεί μία συνθήκη Lipschitz ως προς την δεύτερη μεταβλητή, δηλαδή υπάρχει μία θετική σταθερά L_Φ, έτσι ώστε για $0<h<h_0$ και για το ίδιο χωρίο R όπως στο θεώρημα Picard να ισχύει:

$$\left|\Phi(x,y,h)-\Phi(x,z,h)\right|\le L_\Phi\,|y-z| \tag{14.4}$$

για κάθε ζεύγος (x,y), (x,z) του χωρίου. Επίσης υποθέτουμε ότι $|y_n, y_0|\le Y_m$. Τότε για το ολικό σφάλμα ισχύει:

$$\begin{aligned} &|E_n|\le e^{L_\Phi(x_n,x_0)}|E_0|+\left\{\frac{e^{L_\Phi(x_n,x_0)}-1}{L_\Phi}\right\}T \\ &n=1,2,\dots,N \end{aligned} \tag{14.5}$$

όπου $T = \max_{0 \le n \le N-1} \left| T_n \right|$.

Απόδειξη Αν αφαιρέσουμε την (14.1) από την (14.3) τελικά προκύπτει η σχέση

$$E_{n+1} = E_n + h\{\Phi(x_n, y(x_n), h) - \Phi(x_n, y_n, h)\} + hT_n .$$

Τα σημεία $(x_n, y(x_n))$ και (x_n, y_n), ανήκουν στο χωρίο R , και η συνθήκη (14.4) συνεπάγεται ότι:

$$\left| E_{n+1} \right| \le \left| E_n \right| + hL_\Phi \left| E_n \right| + h\left| T_n \right|, \ \ n = 0,1,....,N-1$$

ή

$$\left| E_{n+1} \right| \le (1 + hL_\Phi)\left| E_n \right| + h\left| T_n \right|, \ \ n = 0,1,....,N-1 .$$

Επομένως θα έχουμε, αν $T = \max_{0 \le n \le N-1} \left| T_n \right|$,

$$\left| E_1 \right| \le (1 + hL_\Phi)\left| E_0 \right| + hT ,$$

$$\left| E_2 \right| \le (1 + hL_\Phi)^2\left| E_0 \right| + h\{1 + (1 + hL_\Phi)\}T ,$$

$$\left| E_3 \right| \le (1 + hL_\Phi)^3\left| E_0 \right| + h\{1 + (1 + hL_\Phi) + (1 + hL_\Phi)^2\}T ,$$

.

.

$$\left| E_n \right| \le (1 + hL_\Phi)^n\left| E_0 \right| + h\{(1 + hL_\Phi)^n - 1\}T / L_\Phi ,$$

Ισχύει ότι $1 + hL_\Phi \le e^{hL_\Phi}$, και $nh = x_n - x_0$, οπότε εύκολα προκύπτει το ζητούμενο.

Σημειώνουμε ότι το φράγμα του σφάλματος (13.10)της μεθόδου Euler είναι ειδική περίπτωση του φράγματος (14.5).

Παράδειγμα 14.1 Θεωρούμε το πρόβλημα αρχικών τιμών

$$y' = \tan^{-1} y , \ \ y(0) = y_0$$

και υποθέτουμε ότι εφαρμόζουμε την άμεση μέθοδο Euler. Σκοπός του παραδείγματος είναι να εφαρμόσουμε την σχέση (13.10), ώστε να υπολογίσουμε το μέγεθος του αντίστοιχου ολικού σφάλματος. Άρα πρέπει να υπολογίσουμε τις ποσότητες L και M_2.

Έχουμε ότι $f(x, y) = \tan^{-1} y$, οπότε από το θεώρημα της μέσης τιμής θα πάρουμε ότι:

$$\left| f(x, y) - f(x, z) \right| = \left| \frac{\partial f}{\partial y}(x, \xi)(y - z) \right|$$

όπου το ξ βρίσκεται μεταξύ y και z. Στην περίπτωση μας ισχύει ότι:

$$\left| \frac{\partial f}{\partial y} \right| = \left| (1 + y^2)^{-1} \right| \le 1,$$

οπότε $L=1$. Για να υπολογίσουμε το M_2, χρειαζόμαστε ένα φράγμα για την $|y''|$. Άρα πρέπει πρώτα να υπολογίσουμε την δεύτερη παράγωγο της y.

Έχουμε:

$$y'=f(x,y) \;\Rightarrow\; y''=\frac{d}{dx}\left(\tan^{-1} y\right) \;\Rightarrow$$

$$y''=\left(1+y^2\right)^{-1} y' \;\Rightarrow\; y''=\left(1+y^2\right)^{-1}\tan^{-1} y.$$

Επομένως, $|y''|\le M_2 = \tfrac{1}{2}\pi$. Η σχέση (13.10) αν αντικαταστήσουμε τα L και M_2 γράφεται:

$$|E_n|\le e^{x_n}|E_0|+\tfrac{1}{4}\pi\left(e^{x_n}-1\right)h,\quad n=0,1,2,\ldots,N \tag{14.6}$$

Ειδικότερα, αν υποθέσουμε ότι $E_0=y(x_0)-y_0$, θα έχουμε ότι:

$$|E_n|\le \tfrac{1}{4}\pi\left(e^{x_n}-1\right)h,\quad n=0,1,2,\ldots,N \tag{14.7}$$

Αν θέλουμε να ισχύει ότι $|E_n|\le TOL$, για κάθε $n=0,1,2,\ldots,N$, τότε υπολογίζεται ότι για να συμβαίνει αυτό θα πρέπει να ισχύει η ανισότητα

$$h\le\frac{4}{\pi}\left(e^{b}-1\right)^{-1}TOL. \tag{14.8}$$

Επομένως μπορούμε να υπολογίζουμε την αριθμητική λύση με όση ακρίβεια θέλουμε, αρκεί να επιλέξουμε την κατάλληλη τιμή για το βήμα h. Αυτό στην πράξη δεν μπορεί να επιλέγεται αυθαίρετα όσο μικρό θέλουμε, διότι πολύ μικρές τιμές για το h μπορεί να δίνουν μεγάλα σφάλματα στρογγύλευσης. Τα σφάλματα στρογγύλευσης μπορούν να φραγούν με ανάλογο τρόπο.

15. Συνέπεια αριθμητικής μεθόδου

Στην παράγραφο αυτή θα συζητήσουμε με συντομία την έννοια της συνέπειας, μία ιδιότητα την οποία πρέπει να έχει μία αριθμητική μέθοδος.

Ορισμός 15.1 Η αριθμητική μέθοδος (14.1) θα λέμε ότι είναι **συνεπής** (consistent) με την διαφορική εξίσωση $y'=f(x,y)$, αν το σφάλμα αποκοπής (14.3) είναι τέτοιο ώστε για κάθε $\varepsilon>0$, υπάρχει ένας αριθμός $h(\varepsilon)$ για τον οποίο ισχύει:

$$|T_n|<\varepsilon \quad \text{για} \quad 0<h<h(\varepsilon)$$

για οποιοδήποτε ζεύγος σημείων $(x_n, y(x_n))$, $(x_{n+1}, y(x_{n+1}))$ σε οποιαδήποτε καμπύλη της λύσης στο R.

Για την γενική μέθοδο (14.1) έχουμε υποθέσει ότι η συνάρτηση $\Phi(.,.,.)$ είναι συνεχής. Επίσης η y' είναι μία συνεχής συνάρτηση στο $[a,b]$. Επομένως από την (14.3) θα έχουμε:

$$\lim_{h\to 0} T_n = y'(x_n)-\Phi(x_n, y_n, 0).$$

Αυτό σημαίνει ότι η μέθοδος απλού βήματος (14.1) είναι συνεπής, αν και μόνο αν

$$\Phi(x_n, y_n, 0) \equiv f(x, y). \tag{15.1}$$

Μπορούμε τώρα να διατυπώσουμε το επόμενο θεώρημα για την σύγκλιση της γενικής μεθόδου απλού βήματος (14.1).

Θεώρημα 15.1 Υποθέτουμε ότι η λύση του προβλήματος αρχικών τιμών

$$y' = f(x, y), \; x \in [a, b], \; y(a) = y_0,$$

ανήκει στο R όπως και η προσέγγιση αυτής που προκύπτει από την εφαρμογή της μεθόδου (14.1) όταν το $h < h_0$. Υποθέτουμε επίσης ότι η συνάρτηση $\Phi(.,.,.)$ είναι ομοιόμορφα συνεχής στο $R \times [0, h_0]$, και ικανοποιεί την συνθήκη συνέπειας (15.1) και την συνθήκη Lipschitz

$$|\Phi(x, y, h) - \Phi(x, z, h)| \le L_\Phi \, |y - z| \quad \text{στο} \quad R \times [0, h_0].$$

Τότε, αν υπολογίσουμε διαδοχικά τις ακολουθίες προσεγγίσεων $\{y_n\}$, για τα σημεία $x_n = a + nh$, $n = 1, 2, \ldots, N$ με την μέθοδο (14.1) με διαδοχικά μικρότερα βήματα h, κάθε ένα μικρότερο του h_0, θα έχουμε σύγκλιση της αριθμητικής λύσης προς την λύση του προβλήματος αρχικών τιμών με την έννοια ότι:

$$|y(x_n) - y_n| \to 0 \quad \text{όταν} \quad h \to 0, \; x_n \to x \in [a, b].$$

16. Έμμεσες μέθοδοι

Στην παράγραφο 10 υπολογίσαμε μερικές απλές μεθόδους απλού βήματος, από τις οποίες μερικές είναι έμμεσες. Στην παράγραφο αυτή θα ασχοληθούμε πιο διεξοδικά με αυτές.

Η άμεση μέθοδος Euler προφανώς εφαρμόζεται εύκολα, αλλά έχει ένα μάλλον περιορισμένο διάστημα ευστάθειας $(-2, 0)$. Μπορούμε να έχουμε μεγαλύτερο χωρίο ευστάθειας όταν η μέθοδος είναι έμμεση, με κόστος βέβαια την πολυπλοκότητα της εφαρμογής.

Έτσι, αν θεωρήσουμε την μέθοδο,

$$y_{n+1} = y_n + h \; f(x_{n+1}, y_{n+1}) \tag{16.1}$$

θα πρέπει να υπολογίσουμε την τιμή της f στο σημείο y_{n+1} πριν το υπολογίσουμε. Αυτό σημαίνει ότι πρέπει να υπολογίσουμε μία αρχική τιμή για το y_{n+1} που ικανοποιεί την (16.1), και αν η f είναι μη γραμμική ως προς y, όπως συνήθως συμβαίνει, τότε μία επαναληπτική μέθοδος, όπως η μέθοδος Newton, θα πρέπει να εφαρμοστεί για την βελτίωση αυτής της τιμής. Μία καλή αρχική προσέγγιση για την αρχική τιμή της επαναληπτικής μεθόδου μπορεί να υπολογιστεί είτε από μία έμμεση μέθοδο, είτε από την προσέγγιση της λύσης στο προηγούμενο σημείο.

Παράδειγμα 16.1 Θεωρούμε την μη γραμμική διαφορική εξίσωση,

$$y' = -y^3$$

με αρχική συνθήκη $y(0) = 1$. Αν εφαρμόσουμε την μέθοδο (16.1) με βήμα $h = 0.5$, θα προκύψει η εξίσωση

$$y' = y_0 + h\, f(x_1, y_1) = 1 - 0.5\, y_1^3 \tag{16.2}$$

για τον υπολογισμό της λύσης στο επόμενο σημείο. Η εξίσωση αυτή είναι σε τέτοια μορφή που μπορούμε να εφαρμόσουμε την επαναληπτική μέθοδο σταθερού σημείου άμεσα.

Πράγματι, από την (16.2) προκύπτει η σχέση της μορφής:

$$y_1^{(s+1)} = 1 - 0.5\left(y_1^{(s)}\right)^3, \; s = 0, 1, \ldots \tag{16.3}$$

όπου η αρχική προσέγγιση $y_1^{(0)}$ μπορεί να υπολογιστεί εφαρμόζοντας την άμεση μέθοδο Euler στο βήμα αυτό. Πράγματι έχουμε:

$$y_1^{(0)} = y_0 - 0.5\, y_0^3 = 1 - 0.5 \cdot 1^3 \Rightarrow y_1^{(0)} = 0.5\,.$$

Με βάση αυτή την αρχική τιμή υπολογίζουμε διαδοχικά:

$$y_1^{(1)} = 1 - 0.5\left(y_1^{(0)}\right)^3 = 0.937$$

$$y_1^{(2)} = 1 - 0.5\left(y_1^{(1)}\right)^3 = 0.589$$

$$y_1^{(3)} = 1 - 0.5\left(y_1^{(2)}\right)^3 = 0.897$$

.

.

κ.τ.λ. οπότε μετά από μερικές εφαρμογές της επαναληπτικής σχέσης (16.3) υπολογίζουμε ότι $y_1 \approx 0.7709$.

Η μη γραμμική εξίσωση (16.2) μπορεί επίσης σε κάθε βήμα να επιλυθεί, αν εφαρμόσουμε την γνωστή επαναληπτική μέθοδο Newton. Τώρα, στην εξίσωση (16.3) αντιστοιχεί η εξίσωση

$$y = 1 - \tfrac{1}{2} y^3$$

ή η εξίσωση

$$f(y) = y + \tfrac{1}{2} y^3 - 1 = 0\,.$$

Η πρώτη παράγωγος είναι

$$f'(y) = 1 + \tfrac{3}{2} y^2\,,$$

οπότε η προσέγγιση της y σε κάθε σημείο με $h = 0.5$, θα μπορεί να υπολογιστεί με τον επαναληπτικό τύπο:

$$y_{n+1} = y_n - \frac{y_n + \tfrac{1}{2} y_n^3 - 1}{1 + \tfrac{3}{2} y_n^2}\,, \; n = 0, 1, \ldots$$

με το $y_0 = 0.5$, υπολογισμένο πάλι με την άμεση μέθοδο Euler. Η εφαρμογή του τύπου θα υπολογίσει την αριθμητική τιμή της προσέγγισης στα διάφορα σημεία της διαμέρισης.

Είναι φανερό ότι η εφαρμογή μιας έμμεσης μεθόδου δημιουργεί σε κάθε βήμα πρόσθετο υπολογιστικό κόστος με τους απαραίτητους υπολογισμούς. Επομένως, στο ερώτημα γιατί να εφαρμόσουμε μία έμμεση μέθοδο, η απάντηση είναι ότι μία έμμεση μέθοδος έχει μεγαλύτερο χωρίο ευστάθειας συγκρινόμενη με αντίστοιχες άμεσες, συνεπώς περισσότερο ικανή για συγκεκριμένες κατηγορίες διαφορικών εξισώσεων.

17. Ευστάθεια και ακρίβεια έμμεσων μεθόδων

a) Έμμεση μέθοδος Euler

Για να μελετήσουμε κατ' αρχήν την ευστάθεια αυτής της μεθόδου, θα την εφαρμόσουμε στην γραμμική εξίσωση μοντέλο της μορφής:

$$y' = \lambda y \tag{17.1}$$

οπότε θα έχουμε:

$$y_{k+1} = y_k + h\lambda y_{k+1}$$

ή

$$y_{k+1} = \frac{1}{1-h\lambda} y_k\,,\ \ k = 0,1,\ldots$$

Για τις διάφορες τιμές του k, ορίζεται ο αναγωγικός τύπος:

$$y_k = \left(\frac{1}{1-h\lambda}\right)^k y_0 \tag{17.2}$$

Η θεωρητική λύση του προβλήματος είναι η συνάρτηση $y(x) = y_0 e^{\lambda x}$, η οποία όταν το $\lambda < 0$, (ευσταθής διαφορική εξίσωση), συγκλίνει εκθετικά στο μηδέν όταν το x αυξάνει.

Είναι φανερό ότι ενδιαφερόμαστε η προσεγγιστική λύση να έχει μία ανάλογη συμπεριφορά, δηλαδή η ακολουθία $\{y_k\}$ να είναι μηδενική. Σύμφωνα με την (17.2) αυτό θα ισχύει αν ικανοποιείται η ανισότητα:

$$|\,1/(1-h\lambda)| < 1 \tag{17.3}$$

Είναι εύκολο να διαπιστώσουμε ότι η (17.3) ισχύει για κάθε θετική τιμή του h όταν το $\lambda < 0$, και επομένως το διάστημα ευστάθειας της μεθόδου είναι $(-\infty, 0\,)$. Όταν έχουμε ένα σύστημα διαφορικών εξισώσεων 1ης τάξης θα μιλάμε για **χωρίο ευστάθειας**, το οποίο στην περίπτωση της έμμεσης μεθόδου Euler, θα είναι ολόκληρο το αριστερό ημιεπίπεδο του μιγαδικού επιπέδου.

Άρα, για κάθε ευσταθή διαφορική εξίσωση, η έμμεση μέθοδος Euler είναι ευσταθής για κάθε θετική τιμή του h. Μία τέτοια μέθοδος λέμε ότι είναι 'ευσταθής χωρίς περιορισμούς' (unconditionally stable) ή **Α-ευσταθής** (A-stable) ή **απόλυτα ευσταθής** (absolutely stable).

Για μία τέτοια μέθοδο, μόνο η επιθυμητή τοπική ακρίβεια μπορεί να περιορίσει την τιμή του h. Αυτό σημαίνει ότι θα μπορούμε να επιλέγουμε μεγαλύτερες τιμές για το βήμα h, απ' ότι σε μία άμεση μέθοδο της ίδιας τάξης, και έτσι να επιτύχουμε υψηλότερη ικανότητα, καίτοι σε κάθε βήμα απαιτούνται περισσότεροι υπολογισμοί.

Για να εξετάσουμε την τάξη ακρίβειας της έμμεσης μεθόδου Euler, θεωρούμε τον παράγοντα αύξησης του σφάλματος (growth factor), και τον αναπτύσσουμε σε σειρά οπότε θα έχουμε την σχέση:

$$\frac{1}{1-h\lambda}=1+h\lambda+(h\lambda)^2+\ldots \tag{17.4}$$

Θεωρούμε επίσης το ανάπτυγμα σε σειρά του όρου $e^{h\lambda}$ της αναλυτικής λύσης, οπότε θα έχουμε την σχέση:

$$e^{h\lambda}=1+h\lambda+\frac{(h\lambda)^2}{2}+\ldots \tag{17.5}$$

Παρατηρούμε ότι τα δύο αναπτύγματα (17.4) και (17.5) συμφωνούν μέχρι τους δύο πρώτους όρους, επομένως η έμμεση μέθοδος Euler έχει τάξη ακρίβειας ένα.

Τελικά, αν και η έμμεση μέθοδος Euler είναι ευσταθής χωρίς περιορισμούς, η μικρή τάξη ακρίβειας της μεθόδου περιορίζει πολύ την χρήση της.

b) Έμμεση μέθοδος τραπεζίου

Θεωρούμε τώρα την έμμεση μέθοδο τραπεζίου:

$$y_{k+1}=y_k+\frac{h}{2}\{f(x_k,y_k)+f(x_{k+1},y_{k+1})\} \tag{17.6}$$

την οποία πάλι εφαρμόζουμε στην εξίσωση μοντέλο (17.1), και προκύπτει η σχέση:

$$y_{k+1}=y_k+\frac{h}{2}\{\lambda y_k+\lambda y_{k+1}\}$$

ή

$$\left(1-{h\lambda}/{2}\right)y_{k+1}=\left(1+{h\lambda}/{2}\right)y_k\ ,\ \ k=0,1,2,\ldots$$

Για τις διάφορες τιμές του k προκύπτει ο αναγωγικός τύπος της μορφής:

$$y_k=\left(\frac{1+h\lambda/2}{1-h\lambda/2}\right)^k y_0 \tag{17.7}$$

Η μέθοδος είναι ευσταθής αν, για τον παράγοντα αύξησης του σφάλματος ισχύει

$$\left|(1+h\lambda/2)/(1-h\lambda/2)\right|<1,$$

και η ανισότητα αυτή θα ισχύει, όταν το $\lambda<0$, για κάθε θετική τιμή του h. Αυτό συνεπάγεται ότι το διάστημα ευστάθειας της μεθόδου για μία απλή διαφορική εξίσωση είναι $(-\infty,0)$ ενώ το χωρίο ευστάθειας για ένα σύστημα διαφορικών εξισώσεων είναι όλο το αριστερό ημιεπίπεδο του μιγαδικού επιπέδου. Άρα, η έμμεση μέθοδος τραπεζίου είναι Α-ευσταθής.

Για την τάξη ακρίβειας της μεθόδου, θα αναπτύξουμε τον παράγοντα αύξησης του σφάλματος σε σειρά, οπότε προκύπτει η σχέση:

$$\frac{1+h\lambda/2}{1-h\lambda/2}=\left(1+{}^{h\lambda}\!/_{2}\right)\left\{1+\frac{h\lambda}{2}+\left(\frac{h\lambda}{2}\right)^2+\left(\frac{h\lambda}{2}\right)^3+\ldots\right\}$$
$$=1+h\lambda+\frac{(h\lambda)^2}{2}+\frac{(h\lambda)^3}{4}+\ldots \tag{17.8}$$

Οι όροι του αναπτύγματος (17.8) συμφωνούν με τους όρους του αναπτύγματος (17.5) μέχρι και τον 3^o που εξαρτάται από το h^2. Άρα η μέθοδος του τραπεζίου έχει τάξη ακρίβειας 2.

Στην παράγραφο αυτή εξετάσαμε δύο έμμεσες μεθόδους απλού βήματος που είναι ευσταθείς χωρίς περιορισμούς. Την ιδιότητα αυτή δεν την έχουν όλες οι έμμεσες μέθοδοι. Όμως οι έμμεσες μέθοδοι γενικά έχουν μεγαλύτερα χωρία ευστάθειας από τις άμεσες, αλλά το βήμα ολοκλήρωσης που μπορεί να χρησιμοποιηθεί δεν είναι απεριόριστο. Συνεπώς, η εμμεσότητα δεν είναι αρκετή για να εξασφαλίσει την ευστάθεια, και η ευστάθεια δεν μπορεί να εξασφαλίσει την ακρίβεια της μεθόδου.

Τέλος, υπάρχουν διαφορικές εξισώσεις οι οποίες χαρακτηρίζονται σαν stiff, οι οποίες για να επιλυθούν ικανοποιητικά με μία αριθμητική μέθοδο, η μέθοδος πρέπει να είναι A-ευσταθής ή να έχει μεγάλο χωρίο ευστάθειας. Αυτό οδηγεί στο συμπέρασμα ότι οι έμμεσες μέθοδοι είναι οι πλέον κατάλληλες για την επίλυση τέτοιων διαφορικών εξισώσεων.

Τυπικά, ένα ευσταθές σύστημα συνήθων διαφορικών εξισώσεων χαρακτηρίζεται σαν stiff, αν οι ιδιοτιμές του Ιακωβιανού πίνακα $\mathbf{J}$ είναι τέτοιες ώστε απολύτως ο λόγος των μεγεθών τους να είναι πολύ μεγάλος. Δηλαδή, υπάρχουν ιδιοτιμές με πολύ μικρό και πολύ μεγάλο μέγεθος. Μία περισσότερο συστηματική ανάλυση των stiff διαφορικών εξισώσεων είναι πέρα από τους σκοπούς αυτού του μαθήματος.

18. Μέθοδος σειράς Taylor

Είναι ήδη γνωστό ότι η μέθοδος Euler μπορεί να κατασκευαστεί θεωρώντας το ανάπτυγμα Taylor. Αν $y(x)$ είναι η θεωρητική λύση του προβλήματος αρχικών τιμών, τότε η τιμή αυτής στο σημείο x_{n+1} της διαμέρισης μπορεί να εκφραστεί με το ανάπτυγμα Taylor με την σχέση:

$$y(x_{n+1})=y(x_n+h)=y(x_n)+h\,y'(x_n)+\frac{h^2}{2!}y''(x_n)+\ldots+\frac{h^p}{p!}y^{(p)}(x_n)+\ldots$$

Αν θεωρήσουμε τους p πρώτους όρους, τότε θα προκύψει η προσεγγιστική σχέση:

$$y(x_{n+1})\approx y(x_n)+h\,y'(x_n)+\frac{h^2}{2!}y''(x_n)+\ldots+\frac{h^p}{p!}y^{(p)}(x_n) \tag{18.1}$$

η οποία είναι ακριβής για τις αντίστοιχες προσεγγίσεις. Έτσι ορίζεται η μέθοδος Taylor τάξης p της μορφής:

$$y_{n+1}=y_n+h\,y'_n+\frac{h^2}{2!}y''_n+\ldots+\frac{h^p}{p!}y_n^{(p)} \tag{18.2}$$

y_0 γνωστό, $n=0,1,2,\ldots,N-1$.

Σημειώνουμε, ότι μία μέθοδος αυτής της μορφής απαιτεί τον υπολογισμό παραγώγων ανώτερης τάξης της y, πράγμα το οποίο οδηγεί σε πολύπλοκους αναλυτικούς υπολογισμούς αν η $y' = f(x, y)$ είναι μη γραμμική, και αυτός είναι ένας βασικός λόγος και δεν εφαρμόζεται στην πράξη.

Για $p=2$, προκύπτει η μέθοδος τάξης 2 της μορφής:

$$y_{n+1} = y_n + h y'_n + \frac{h^2}{2!} y''_n \tag{18.3}$$

όπου

$$y'_n = f(x_n, y_n) \quad \text{και} \quad \frac{d}{dx} f(x_n, y_n) = f_x(x_n, y_n) + f(x_n, y_n) f_y(x_n, y_n),$$

και ο δείκτης στην f δηλώνει μερική παράγωγο ως προς την αντίστοιχη μεταβλητή. Έτσι η (18.3) τελικά γράφεται:

$$y_{n+1} = y_n + h\ f(x_n, y_n) + \frac{h^2}{2!}\left\{f_x(x_n, y_n) + f(x_n, y_n) f_y(x_n, y_n)\right\}$$

y_0 γνωστό, $n = 0,1,2,\ldots,N-1$.

Για $p=3$, προκύπτει η μέθοδος τάξης 3της μορφής:

$$y_{n+1} = y_n + h y'_n + \frac{h^2}{2!} y''_n + \frac{h^3}{3!} y'''_n \tag{18.4}$$

y_0 γνωστό, $n = 0,1,2,\ldots,N-1$,

όπου

$$y'_n = f(x_n, y_n)$$

$$y'' = f_x(x_n, y_n) + f(x_n, y_n) f_y(x_n, y_n)$$

$$y'''_n = \left\{f_{xx} + f\ f_{xy} + f^2\ f_{yy} + f_x\ f_y + f_y^2\ f\right\}_{(x_n, y_n)}.$$

Τελευταία, η δυνατότητα των υπολογιστών σε συμβολικές πράξεις και αναλυτική παραγώγιση των συναρτήσεων, δίνει στις μεθόδους αυτού του τύπου μεγαλύτερη ευχέρεια χρήσης.

Παράδειγμα 18.1 Δίνεται το πρόβλημα, $y' = -2xy$, $y(0) = 1$. Η εφαρμογή της μεθόδου Taylor τάξης 2 απαιτεί τον υπολογισμό της y'', σύμφωνα με τα γνωστά. Άρα θα έχουμε την σχέση:

$$y'' = -2y^2 + (-2xy^2)(-4xy) = 2y^2(4x^2 y - 1)$$

και η μέθοδος γράφεται:

$$y_{n+1} = y_n - 2h x_n y_n^2 + h^2 y_n^2 (4x_n^2 y_n - 1), \quad n = 0,1,2\ldots$$

Θα εφαρμόσουμε την μέθοδο δύο φορές με $h = 0.5$

Βήμα 1: $y_1 = y_0 - 2h x_0 y_0^2 + h^2 y_0^2 (4x_0^2 y_0 - 1)$, $x_0 = 0$, $h = 0.25$, $y_0 = 1$

Αντικαθιστούμε τα δεδομένα και υπολογίζεται η προσέγγιση $y_1 = 0.9375$.

Βήμα 2: $y_2 = y_1 - 2hx_1y_1^2 + h^2y_1^2(4x_1^2y_1 - 1)$, $x_1 = 0.25$, $h = 0.25$, $y_1 = 0.9375$

Αντικαθιστούμε τα δεδομένα και υπολογίζεται η προσέγγιση $y_2 = 0.8247$.

Για να ελέγξουμε τα αποτελέσματα, θεωρούμε την θεωρητική λύση του προβλήματος $y(x) = 1/(1+x^2)$, και υπολογίζουμε:

$$y(x_1) = y(0.25) = 0.9412, \quad \text{και} \quad y(x_2) = y(0.5) = 0.8$$

οπότε μπορούμε να κάνουμε τις συγκρίσεις με τις αντίστοιχες προσεγγιστικές τιμές.

19. Άμεση μέθοδος τύπου Runge-Kutta

Οι μέθοδοι τύπου Runge-Kutta που θα ασχοληθούμε στην συνέχεια, στοχεύουν σε μεγαλύτερη ακρίβεια θυσιάζοντας την απλή και οικονομική εφαρμογή της μεθόδου Euler. Εδώ η $f(.,.)$ υπολογίζεται σε ενδιάμεσα σημεία των σημείων $(x_n, y(x_n))$ και $(x_{n+1}, y(x_{n+1}))$.

Η γενική άμεση μέθοδος Runge-Kutta me r υπολογισμούς (στάδια) της f ορίζεται από τις παρακάτω σχέσεις:

$$y_{n+1} = y_n + h\Phi(x_n, y_n, h) \tag{19.1}$$

$$\Phi(x_n, y_n, h) = \sum_{i=1}^{r} b_i k_i \tag{19.2}$$

όπου

$$\begin{aligned} k_1 &= f(x_n, y_n) \\ k_i &= f\left(x_n + hc_i, y_n + h\sum_{j=1}^{i-1} a_{ij} k_j\right), \; i = 2, 3, ..., r \end{aligned} \tag{19.3}$$

$$c_i = \sum_{j=1}^{i-1} a_{ij} \;\; i = 2, 3, ..., r \tag{19.4}$$

Οι συντελεστές της γενικής αυτής μεθόδου μπορούν να παρασταθούν με τον γνωστό πίνακα Butcher,

c_i	a_{ij}					b_i
0						b_1
c_2	a_{21}					b_2
c_3	a_{31}	a_{32}				b_3
.	.	.				.
.	.	.				.
.	.	.				.
c_r	a_{r1}	a_{r2}	.	.	. $a_{r,r-1}$	b_r

Οι ιδιότητες της μεθόδου ορίζονται από τους προσδιοριστέους συντελεστές c_i, a_{ij}, b_i και τον αριθμό r των υπολογισμών της f. Η μέθοδος είναι άμεση με την έννοια ότι κάθε υπολογισμός k_i εξαρτάται από τους προηγούμενους υπολογισμούς k_j, $j=1,2,...,i-1$. Αυτό δεν αποτελεί ουσιαστική ιδιότητα, αλλά απλοποιεί τους υπολογισμούς και ο αλγόριθμος προγραμματίζεται πιο εύκολα.

Η μέθοδος (19.1) – (19.4) απαιτεί r υπολογισμούς της $f(x,y)$ σε κάθε βήμα ολοκλήρωσης σε διάφορα ενδιάμεσα σημεία του βήματος. Κάθε υπολογισμός k_j, $j=2,3,...,r$ μπορεί να θεωρηθεί σαν προσέγγιση κάποιας παραγώγου.

Για να κατασκευάσουμε μία μέθοδο Runge-Kutta γενικά τάξης p, θα πρέπει να απαιτήσουμε το ανάπτυγμα του δεύτερου μέλους της σχέσης

$$y_{n+1}=y_n+h\sum_{i=1}^{r}b_i k_i \tag{19.5}$$

σε δυνάμεις του h, να διαφέρει από το ανάπτυγμα ή την μέθοδο Taylor τάξης p:

$$y_{n+1}=y_n+h\,f(x_n,y_n)+\frac{h^2}{2!}f^{(1)}(x_n,y_n)+\ldots+\frac{h^p}{p!}f^{(p-1)}(x_n,y_n) \tag{19.6}$$

όπου

$$f^{(q)}(x_n,y_n)=\frac{d^q}{dx^q}f(x_n,y_n),\ q=1,2,\ldots,(p-1),$$

σε όλους τους όρους με δυνάμεις του h μεγαλύτερες του p. Η απαίτηση αυτή, να ταυτίζονται οι αντίστοιχοι όροι των αναπτυγμάτων μέχρι κάποια δύναμη του h, οδηγεί σε ένα μη γραμμικό αλγεβρικό σύστημα με αγνώστους τους προσδιοριστέους συντελεστές. Ο αριθμός των αγνώστων υπερβαίνει τον αριθμό των εξισώσεων, οπότε οι λύσεις είναι παραμετρικές, και με κατάλληλη επιλογή των παραμέτρων προκύπτουν συγκεκριμένες μέθοδοι από την αντίστοιχη οικογένεια μεθόδων.

Στην συνέχεια, θα υπολογίσουμε μερικές απλές και κλασσικές μεθόδους αυτής της κατηγορίας για διάφορες τιμές του r, και θα αναφερθούμε σε μερικές βασικές και ικανές μεθόδους τις οποίες χρησιμοποιούμε στις εφαρμογές.

Τέλος, η μέθοδος Euler μπορεί να ενταχθεί στην κατηγορία των μεθόδων Runge-Kutta me $r=1$, και θα ονομάζεται μέθοδος Runge-Kutta με έναν υπολογισμό της f ή ένα στάδιο.

20. Άμεση μέθοδος Runge-Kutta δύο σταδίων

Όταν το $r=2$, η μέθοδος γράφεται:

$$y_{n+1}=y_n+h\left(b_1k_1+b_2k_2\right) \tag{20.1}$$

όπου

$$\begin{aligned} k_1&=f(x_n,y_n)\\ k_2&=f(x_n+hc_2,y_n+a_{21}hk_1)\end{aligned} \tag{20.2}$$

Κατ' αρχήν, η μέθοδος θα πρέπει να είναι συνεπής. Δηλαδή πρέπει $\Phi(x,y,0)=f(x,y)$. Από την (20.1) προκύπτει ότι:

$$\Phi(x,y,h)=b_1\ f(x,y)+b_2\ f\big(x+h\,c_2\,,y+a_{21}\,h\,f\,(x,y)\big),$$

άρα

$$\Phi(x,y,0)=b_1\ f(x,y)+b_2\ f(x,y)=\left(b_1+b_2\right)f(x,y).$$

Επομένως, η συνέπεια εξασφαλίζεται αν και μόνο αν

$$b_1+b_2=1 \tag{20.3}$$

Επί πλέον συνθήκες για τους συντελεστές προκύπτουν, αν επιχειρήσουμε να μεγιστοποιήσουμε την τάξη ακρίβειας της μεθόδου. Για την ευκολία της ανάλυσης υιοθετούμε τους συμβολισμούς:

$$f=f(x,y)\,,\ \ f_x=\frac{\partial}{\partial x}f(x,y)\,,\ \ f_{xx}=\frac{\partial^2}{\partial x^2}f(x,y)\,,\ \ f_{xy}=\frac{\partial^2}{\partial x\partial y}f(x,y)\,,\,.. \tag{20.4}$$

Ως γνωστόν, η μέθοδος Taylor τάξης p ορίζεται αν στην (19.1) η

$$\Phi_{\mathrm{T}}(x_n,y_n,h)=f(x_n,y_n)+\frac{h^2}{2!}f^{(1)}(x_n,y_n)+\ldots+\frac{h^p}{p!}f^{(p-1)}(x_n,y_n) \tag{20.5}$$

Υπολογίζοντας παραγώγους της f μέχρι και τάξη 2, και εφαρμόζοντας τους συμβολισμούς (20.4), η $\Phi_{\mathrm{T}}(x_n,y_n,h)$ γράφεται:

$$\Phi_{\mathrm{T}}(x_n,y_n,h)=f+\frac{1}{2}h\left\{f_x+f\ f_y\right\}+\frac{1}{6}h^2\left\{\left(f_x+f\ f_y\right)f_y+f_{xx}+2f\ f_{xy}+f^2\ f_{yy}\right\}+O\!\left(h^3\right)$$

ή

$$\Phi_{\mathrm{T}}(x_n,y_n,h)=f+\frac{1}{2}h\,F_1+\frac{1}{6}h^2\left\{F_1\ f_y+F_2\right\}+O\!\left(h^3\right) \tag{20.6}$$

όπου

$$F_1=f_x+f\ f_y\ ,\ \ F_2=f_{xx}+2\ f\ f_{xy}+f^2 f_{yy} \tag{20.7}$$

και όλες οι συναρτήσεις υπολογίζονται στο σημείο (x_n,y_n).

Για να ορίσουμε τις συνθήκες που καθορίζουν την τάξη ακρίβειας της μεθόδου, θα αναπτύξουμε το k_2 από την (20.2) σε δυνάμεις του h, εφαρμόζοντας το θεώρημα Taylor για 2 μεταβλητές.

Σύμφωνα με αυτό θα έχουμε:

$$\begin{aligned} k_2 &= f\left(x_n+h\,c_2\,,y_n+a_{21}\,h\,k_1\right)\\ &= f+\left\{h\,c_2\ f_x+a_{21}\,h\,k_1\ f_y\right\}+\left\{\frac{1}{2}h^2\,c_2^2\ f_{xx}+c_2\,a_{21}\,h^2\,k_1\ f_{xy}+\frac{1}{2}a_{21}^2\,h^2\,k_1^2\ f_{yy}\right\}\\ &\quad+O\!\left(h^3\right)\end{aligned}$$

Επειδή $k_1 = f\left(x_n, y_n\right)$, και $c_2 = a_{21}$ λόγω της (19.4), τελικά προκύπτει:

$$k_2 = f + h c_2 F_1 + \frac{1}{2} h^2 c_2^2 F_2 + O\left(h^3\right) \tag{20.8}$$

Αντικαθιστούμε στην (20.1) τις εκφράσεις των k_1 και k_2, οπότε για την ζητούμενη μέθοδο θα προκύψει:

$$y_{n+1} = y_n + h\left(b_1 + b_2\right) f + h^2 b_2 c_2 F_1 + \frac{1}{2} h^3 b_2 c_2^2 F_2 . \tag{20.9}$$

Η αντίστοιχη μέθοδος Taylor γράφεται:

$$y_{n+1} = y_n + h f + \frac{1}{2} h^2 F_1 + \frac{1}{6} h^3 \left(F_1 f_y + F_2\right) \tag{20.10}$$

Η μέγιστη τάξη που μπορεί να έχει η ζητούμενη μέθοδος είναι δύο, και αυτό προκύπτει ότι τα αναπτύγματα (20.9) και (20.10) μπορούν να συμφωνήσουν μέχρι και τους όρους που εξαρτώνται από το h^2, αν ισχύουν οι εξισώσεις:

$$\begin{aligned} &b_1 + b_2 = 1 \\ &b_2 c_2 = \frac{1}{2} \\ &c_2 = a_{21} \end{aligned} \tag{20.11}$$

Το σύστημα (20.11) των συντελεστών μπορεί να ορίσει μία μονοπαραμετρική οικογένεια λύσεων της μορφής:

$$a_{21} = c_2 \,,\ b_2 = 1/2c_2 \,,\ b_1 = 1 - 1/2c_2$$

από την οποία, για επιλεγμένες τιμές της ελεύθερης παραμέτρου προκύπτουν συγκεκριμένες μέθοδοι.

a) Για $c_2 = \frac{1}{2} \Rightarrow a_{21} = \frac{1}{2}$, $b_2 = 1$, $b_1 = 0$.

Τότε προκύπτει η μέθοδος:

$$\begin{aligned} &k_1 = f\left(x_n, y_n\right) \\ &k_2 = f\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}h k_1\right) \\ &y_{n+1} = y_n + h k_2 \end{aligned} \tag*{((20.12)}$$

η οποία θα αναφέρεται σαν **τροποποιημένη** (modified) μέθοδος Euler.

b) Για $c_2 = 1 \Rightarrow a_{21} = 1$, $b_2 = \frac{1}{2}$, $b_1 = \frac{1}{2}$.

Τότε προκύπτει η μέθοδος:

$$
\begin{aligned}
&k_1 = f(x_n, y_n)\\
&k_2 = f(x_n + h, y_n + hk_1)\\
&y_{n+1} = y_n + \frac{h}{2}(k_1 + k_2)
\end{aligned}
\tag{20.13}
$$

η οποία αντιστοιχεί στην **βελτιωμένη** (improved) μέθοδο Euler.

21. Άμεση μέθοδος Runge-Kutta τριών σταδίων

Όταν το $r=3$, η μέθοδος γράφεται:

$$y_{n+1} = y_n + h(b_1 k_1 + b_2 k_2 + b_3 k_3) \tag{21.1}$$

όπου

$$
\begin{aligned}
&k_1 = f(x_n, y_n)\\
&k_2 = f(x_n + hc_2, y_n + a_{21} hk_1)\\
&k_3 = f(x_n + hc_3, y_n + ha_{31} k_1 + ha_{32} k_2)\\
&c_2 = a_{21}, \quad c_3 = a_{31} + a_{32}
\end{aligned}
\tag{21.2}
$$

Στις εκφράσεις των k_2 και k_3 κάνουμε τις αντικαταστάσεις $a_{21}=c_2$ και $a_{32}=c_3-a_{32}$, και εν' συνεχεία τα αναλύουμε σε ανάπτυγμα σύμφωνα με το θεώρημα Taylor για δύο μεταβλητές. Τότε προκύπτουν οι παρακάτω εκφράσεις για τα k_2 και k_3, όπου όλες οι συναρτήσεις υπολογίζονται στο σημείο (x_n, y_n),

$$
\begin{aligned}
k_2 &= f + hc_2(f_x + k_1 f_y) + \frac{1}{2}h^2 c_2^2 (f_{xx} + 2k_1 f_{xy} + k_1^2 f_{yy}) + O(h^3)\\
&= f + hc_2(f_x + f f_y) + \frac{1}{2}h^2 c_2^2 (f_{xx} + 2f f_{xy} + f^2 f_{yy}) + O(h^3)
\end{aligned}
$$

Αν χρησιμοποιήσουμε τις εκφράσεις (20.7), τελικά προκύπτει:

$$k_2 = f + hc_2 F_1 + \frac{1}{2}h^2 c_2^2 F_2 + O(h^3) \tag{21.3}$$

και

$$
\begin{aligned}
k_3 = f &+ h\{c_3 f_x + [(c_3 - a_{32})k_1 + a_{32} k_2] f_y\} +\\
&\frac{1}{2}h^2\{c_3^2 f_{xx} + 2c_3[(c_3 - a_{32})k_1 + a_{32}k_2] f_{xy} + [(c_3 - a_{32})k_1 + a_{32}k_2]^2 f\}\\
&+ O(h^3)
\end{aligned}
$$

ή τελικά:

$$k_3 = f + hc_3 F_1 + h^2\left(c_2 a_{32} F_1 f_y + \frac{1}{2}c_3^2 F_2\right) + O(h^3) \tag{21.4}$$

Αν αντικαταστήσουμε τις εκφράσεις (21.3) και (21.4) στην σχέση (19.2) για $r=3$, θα προκύψει:

$$\Phi(x_n, y_n, h) = (b_1 + b_2 + b_3) f + h(b_2 c_2 + b_3 c_3) F_1 + \frac{1}{2} h^2 \left\{ 2 b_3 c_2 a_{32} F_1 f_y + (b_2 c_2^2 + b_3 c_3^2) F_2 \right\} + O(h^3) \tag{21.5}$$

Άρα η (21.1) γράφεται:

$$y_{n+1} = y_n + h(b_1 + b_2 + b_3) f + h^2 (b_2 c_2 + b_3 c_3) F_1 + \frac{1}{2} h^3 \left\{ 2 b_3 c_2 a_{32} F_1 f_y + (b_2 c_2^2 + b_3 c_3^2) F_2 \right\} \tag{21.6}$$

Η αντίστοιχη μέθοδος Taylor γράφεται:

$$y_{n+1} = y_n + h f + \frac{1}{2} h^2 F_1 + \frac{1}{6} h^3 (F_1 f_y + F_2) + O(h^3) \tag{21.7}$$

Η μέγιστη τάξη που μπορεί να επιτύχει η (21.6) είναι τρία, και αυτό προκύπτει από το γεγονός ότι τα αναπτύγματα (21.6) και (21.7) μπορούν να συμφωνήσουν μέχρι και τους όρους που εξαρτώνται από το h^3, αν ικανοποιούνται οι εξισώσεις:

$$\begin{aligned} & b_1 + b_2 + b_2 = 1 \\ & b_2 c_2 + b_3 c_3 = \frac{1}{2} \\ & b_2 c_2^2 + b_3 c_3^2 = \frac{1}{3} \\ & b_3 c_2 a_{32} = \frac{1}{6} \end{aligned} \tag{21.8}$$

Το σύστημα (21.8) έχει τέσσερις εξισώσεις με έξι αγνώστους $b_1, b_2, b_3, c_2, c_3, a_{32}$. Έτσι μπορούμε να υπολογίσουμε μία οικογένεια μεθόδων Runge-Kutta με δύο ελεύθερες παραμέτρους. Κατάλληλες τιμές των ελεύθερων παραμέτρων ορίζουν συγκεκριμένες μεθόδους της οικογένειας. Δύο αξιοσημείωτα παραδείγματα μεθόδων από την οικογένεια αυτή είναι:

a) Μέθοδος Heun

Οι συντελεστές της μεθόδου είναι:

$$b_1 = \frac{1}{4},\ b_2 = 0,\ b_3 = \frac{3}{4},\ c_2 = \frac{1}{3},\ c_3 = \frac{2}{3},\ a_{21} = \frac{1}{3},\ a_{31} = 0,\ a_{32} = \frac{2}{3}$$

οπότε η μέθοδος θα έχει την μορφή:

$$\begin{aligned} & y_{n+1} = y_n + \frac{h}{4}(k_1 + 3k_3) \\ & k_1 = f(x_n, y_n) \\ & k_2 = f\left(x_n + \frac{1}{3} h, y_n + \frac{1}{3} h k_1 \right) \\ & k_3 = f\left(x_n + \frac{2}{3} h, y_n + \frac{2}{3} h k_2 \right) \end{aligned} \tag{21.9}$$

b) Κλασσική μέθοδος Runge-Kutta τρίτης τάξης

Οι συντελεστές της μεθόδου είναι:

$$b_1=\frac{1}{6},\, b_2=\frac{2}{3},\, b_3=\frac{1}{6}\,,\, c_2=\frac{1}{2},\, c_3=1,\; a_{21}=\frac{1}{2},\; a_{31}=-1,\, a_{32}=2\,,$$

οπότε η μέθοδος θα έχει την μορφή:

$$\begin{aligned}
y_{n+1}&=y_n+\frac{h}{6}\left(k_1+4k_2+k_3\right)\\
k_1&=f\left(x_n,y_n\right)\\
k_2&=f\left(x_n+\frac{1}{2}h\,,y_n+\frac{1}{2}hk_1\right)\\
k_3&=f\left(x_n+h\,,y_n-hk_1+2hk_2\right)
\end{aligned}\tag{21.10}$$

22. Άμεση μέθοδος Runge-Kutta τεσσάρων σταδίων

Ας υποθέσουμε ότι $r=4$, τότε η μέθοδος γράφεται:

$$y_{n+1}=y_n+h\left(b_1k_1+b_2k_2+b_3k_3+b_4k_4\right)\tag{22.1}$$

όπου

$$\begin{aligned}
k_1&=f\left(x_n,y_n\right)\\
k_2&=f\left(x_n+hc_2,y_n+a_{21}hk_1\right)\\
k_3&=f\left(x_n+hc_3,y_n+ha_{31}k_1+ha_{32}k_2\right)\\
k_3&=f\left(x_n+hc_3,y_n+ha_{31}k_1+ha_{32}k_2\right)\\
k_4&=f\left(x_n+hc_4,y_n+ha_{41}k_1+ha_{42}k_2+ha_{43}k_3\right)\\
c_2&=a_{21}\,,\quad c_3=a_{31}+a_{32}\,,\; c_4=a_{41}+a_{42}+a_{43}
\end{aligned}\tag{22.2}$$

Αν εργαστούμε με ανάλογο τρόπο, θα καταλήξουμε μετά από πολύπλοκους και βαρετούς υπολογισμούς, σε μία οικογένεια μεθόδων Runge-Kutta με τέσσερις υπολογισμούς της f, η οποία έχει δύο ελεύθερες παραμέτρους. Μία δημοφιλής μέθοδος αυτής της οικογένειας είναι:

$$\begin{aligned}
y_{n+1}&=y_n+\frac{h}{6}\left(k_1+2k_2+2\,k_3+k_4\right)\\
k_1&=f\left(x_n,y_n\right)\\
k_2&=f\left(x_n+\frac{1}{2}h\,,y_n+\frac{1}{2}hk_1\right)\\
k_3&=f\left(x_n+\frac{1}{2}h\,,y_n+\frac{1}{2}hk_2\right)\\
k_4&=f\left(x_n+h\,,y_n+hk_3\right)
\end{aligned}\tag{22.3}$$

Μέχρι τώρα παρουσιάσαμε μεθόδους Runge-Kutta r-σταδίων με $r=1,2,3,4$. Για τις μεθόδους αυτές ισχύει ότι η τάξη ακρίβειας είναι αντίστοιχα $p=1,2,3,4$. Μία φυσική ερώτηση είναι αν υπάρχουν μέθοδοι Runge-Kutta με $r\geq 5$, και ποία είναι η αντίστοιχη τάξη αυτών.

Έχει αποδειχθεί ότι για $r=5,6,7,8,9$ η μέγιστη τάξη ακρίβειας η οποία μπορεί να επιτευχθεί είναι αντίστοιχα $p=4,5,6,6,7$, και για $r\geq 10$, η μέγιστη τάξη είναι $p\leq r-2$, αν και μπορεί αυτά τα αποτελέσματα να ανατραπούν.

23. Ευστάθεια των μεθόδων Runge-Kutta

Για να μελετήσουμε την ιδιότητα της ευστάθειας για τις μεθόδους Runge-Kutta, είναι χρήσιμο να θεωρήσουμε το πρόβλημα μοντέλο της μορφής:

$$y'=\lambda y,\ \ y(0)=y_0\ (\neq 0) \qquad (23.1)$$

όπου λ είναι ένας πραγματικός και αρνητικός αριθμός. Η αναλυτική λύση του προβλήματος εύκολα προκύπτει ότι είναι η $y(x)=y_0\, e^{\lambda x}$, η οποία συγκλίνει εκθετικά στο μηδέν όταν $x\to+\infty$.

Το ερώτημα που θα θέλαμε να διερευνήσουμε εδώ είναι, κάτω από ποίες συνθήκες ή περιορισμούς στο βήμα h, η μέθοδος Runge-Kutta παράγει προσεγγιστικές λύσεις με ανάλογη συμπεριφορά. Η κατανόηση στο θέμα αυτό θα μας δώσει χρήσιμες πληροφορίες για την καλύτερη επιλογή του h, όταν εφαρμόζεται για την αριθμητική προσέγγιση ενός προβλήματος αρχικών τιμών στο διάστημα $[a,b]$, όπου $b >> a$.

Για λόγους καθαρά εκπαιδευτικούς θα μελετήσουμε την ευστάθεια για τις μεθόδους Runge-Kutta τάξης $p=1,2,3,4$.

a) Μέθοδος Euler

Η μέθοδος Euler μπορεί να θεωρηθεί σαν μία μέθοδος Runge-Kutta με $r=1$ και $p=1$. Αν την εφαρμόσουμε στο πρόβλημα (23.1) τότε θα προκύψει η σχέση:

$$y_{n+1}=(1+h\lambda)y_n\ ,\ \ n\geq 0$$

Αν ο παραπάνω αναδρομικός τύπος εφαρμοστεί για τις διάφορες τιμές του n, προκύπτει η σχέση:

$$y_{n+1}=(1+h\lambda)^n\ y_0\ ,\ \ n=0,1,2,\ldots$$

η οποία ορίζει την ακολουθία των προσεγγίσεων $\{y_n\}_{n=0}^{\infty}$. Η ακολουθία αυτή θα συγκλίνει στο μηδέν, δηλαδή θα έχει μία συμπεριφορά ανάλογή της αναλυτικής λύσης, αν και μόνο αν ικανοποιείται η ανισότητα:

$$|1+h\lambda|<1$$

από την οποία προκύπτει το γνωστό αποτέλεσμα $h\lambda\in(-2,0)$. Η προηγούμενη συνθήκη μας δίνει την δυνατότητα να υπολογίζουμε κατάλληλη τιμή για το h, ώστε να έχουμε την επιθυμητή συμπεριφορά της προσεγγιστικής λύσης. Όταν η τιμή του h είναι τέτοια ώστε η τελευταία συνθήκη να ισχύει, λέμε ότι η μέθοδος είναι **απόλυτα ευσταθής** και το διάστημα $(-2,0)$, ονομάζεται **διάστημα απόλυτης ευστάθειας** της μεθόδου.

b) Μέθοδος Runge-Kutta με $r=2$ ή βελτιωμένη μέθοδος Euler

Η μέθοδος αυτή ορίζεται από τις σχέσεις:

$$k_1 = f(x_n, y_n)$$
$$k_2 = f(x_n + h, y_n + hk_1)$$
$$y_{n+1} = y_n + \frac{h}{2}(k_1 + k_2)$$

Εφαρμόζουμε την μέθοδο στην εξίσωση (23.1), και προκύπτουν οι σχέσεις:

$$k_1 = \lambda y_n$$
$$k_2 = \lambda y_n + h\lambda^2 y_n$$

οπότε η προσέγγιση παίρνει την μορφή:

$$y_{n+1} = \left(1 + h\lambda + \frac{1}{2}(h\lambda)^2\right) y_n \ , \ n \geq 0$$

από την οποία προκύπτει για τις διάφορες τιμές του n ο αναδρομικός τύπος:

$$y_n = \left(1 + h\lambda + \frac{1}{2}(h\lambda)^2\right) y_0 \ , \ n = 0, 1, 2, \ldots$$

Η μέθοδος θα είναι απόλυτα ευσταθής αν ικανοποιείται η ανισότητα:

$$\left| 1 + h\lambda + \frac{1}{2}(h\lambda)^2 \right| < 1,$$

και αυτή ικανοποιείται αν $h\lambda \in (-2, 0)$. Συνεπώς το διάστημα της απόλυτης ευστάθειας της μεθόδου είναι $(-2, 0)$.

c) Μέθοδος Runge-Kutta με $r = 3$

Μία κλασσική μέθοδος αυτής της οικογένειας εκφράζεται με τις σχέσεις:

$$y_{n+1} = y_n + \frac{h}{6}(k_1 + 4k_2 + k_3)$$
$$k_1 = f(x_n, y_n)$$
$$k_2 = f\left(x_n + \frac{1}{2}h, y_n + \frac{1}{2}hk_1\right)$$
$$k_3 = f(x_n + h, y_n - hk_1 + 2hk_2)$$

Αν η μέθοδος εφαρμοστεί στην εξίσωση (23.1), προκύπτουν οι εξισώσεις:

$$k_1 = \lambda y_n$$
$$k_2 = \lambda y_n + \frac{1}{2}h\lambda^2 y_n$$
$$k_3 = \lambda y_n + h\lambda^2 y_n + h^2\lambda^3 y_n$$

με τις οποίες η προσέγγιση γράφεται:

$$y_{n+1} = \left(1 + h\lambda + \frac{1}{2}(h\lambda)^2 + \frac{1}{6}(h\lambda)^3\right) y_n \ , \ n \geq 0 .$$

Από την προηγούμενη προκύπτει ο αναδρομικός τύπος:

$$y_n=\left(1+h\lambda+\frac{1}{2}(h\lambda)^2+\frac{1}{6}(h\lambda)^3\right)y_0\ ,\ n=0,1,2,\ldots$$

Η συνθήκη ευστάθειας είναι:

$$\left|1+h\lambda+\frac{1}{2}(h\lambda)^2+\frac{1}{6}(h\lambda)^3\right|<1,$$

από την οποία υπολογίζεται ότι $h\lambda\in(-2.51,0)$. Άρα για τις μεθόδους Runge-Kutta με $r=3$, το διάστημα απόλυτης ευσταθείς είναι $(-2.51,0)$.

d) Μέθοδος Runge-Kutta με $r=4$

Θεωρούμε την παρακάτω μέθοδο της οικογένειας αυτής:

$$\begin{aligned}
y_{n+1}&=y_n+\frac{h}{6}\left(k_1+2k_2+2k_3+k_4\right)\\
k_1&=f\left(x_n,y_n\right)\\
k_2&=f\left(x_n+\frac{1}{2}h\ ,y_n+\frac{1}{2}hk_1\right)\\
k_3&=f\left(x_n+\frac{1}{2}h\ ,y_n+\frac{1}{2}hk_2\right)\\
k_4&=f\left(x_n+h\ ,y_n+hk_3\right)
\end{aligned}$$

την οποία εφαρμόζουμε στο πρόβλημα (23.1). Με διαδικασία ανάλογη των προηγούμενων περιπτώσεων η προσέγγιση της λύσης εκφράζεται με την εξίσωση:

$$y_{n+1}=\left(1+h\lambda+\frac{1}{2}(h\lambda)^2+\frac{1}{6}(h\lambda)^3+\frac{1}{24}(h\lambda)^4\right)y_n\ ,\ n\geq 0$$

από την οποία προκύπτει ο αναδρομικός τύπος:

$$y_n=\left(1+h\lambda+\frac{1}{2}(h\lambda)^2+\frac{1}{6}(h\lambda)^3+(h\lambda)^4\right)y_0\ ,\ n=0,1,2,\ldots$$

Για την ευστάθεια της μεθόδου απαιτείται να ισχύει η συνθήκη:

$$\left|1+h\lambda+\frac{1}{2}(h\lambda)^2+\frac{1}{6}(h\lambda)^3+\frac{1}{24}(h\lambda)^4\right|<1,$$

από την οποία υπολογίζεται ότι αυτή θα ικανοποιείται όταν $h\lambda\in(-2.78,0)$, επομένως το διάστημα ευστάθειας των μεθόδων Runge-Kutta με $r=4$, είναι $(-2.78,0)$.

Η παραπάνω ανάλυση μπορεί να εφαρμοστεί για οποιαδήποτε μέθοδο Runge-Kutta με $r>4$ και γενικά προκύπτει για την προσέγγιση, μία σχέση της μορφής:

$$y_{n+1}=P_r\left(h\lambda\right)y_n\ ,\ n\geq 0,$$

όπου το πολυώνυμο $P_r\left(h\lambda\right)$, ονομάζεται πολυώνυμο απόλυτης ευστάθειας της μεθόδου, και εξαρτάται εκτός από το $h\lambda$, και από τους συντελεστές της μεθόδου.

Γραμμικές πολυβηματικές μέθοδοι

24. Εισαγωγή

Στο κεφάλαιο αυτό θα μελετήσουμε την κατηγορία των αριθμητικών μεθόδων οι οποίες είναι γνωστές σαν γραμμικές πολυβηματικές μέθοδοι (Linear multistep methods), ή πιο συγκεκριμένα, γραμμικές μέθοδοι $k-$βημάτων., για την επίλυση του προβλήματος αρχικών τιμών της μορφής:

$$\begin{aligned} &y' = f(x,y), \ x\in[a,b] \\ &y(a)=y_0 \end{aligned} \tag{24.1}$$

Οι μέθοδοι Runge-Kutta που μελετήσαμε στο προηγούμενο κεφάλαιο, αποτελούν μία βελτίωση των μεθόδων Euler ως προς την ακρίβεια, με πρόσθετο όμως υπολογιστικό κόστος. Στην πραγματικότητα, οι μέθοδοι Runge-Kutta απαιτούν περισσότερους υπολογισμούς της συνάρτησης $f(.,.)$ απ' ότι φαίνεται αναγκαίο. Για παράδειγμα, η μέθοδος Runge-Kutta $4^{ης}$ τάξης απαιτεί τέσσερις υπολογισμούς σε κάθε βήμα, οι οποίοι γίνονται όμως σε ενδιάμεσα σημεία του βήματος.

Μία πολυβηματική μέθοδος απαιτεί υπολογισμούς της συνάρτησης f σε περισσότερα του ενός σημεία της διαμέρισης, άρα θα πρέπει να γνωρίζουμε την προσέγγιση της λύσης στα σημεία αυτά.

Θεωρούμε μία ακολουθία από ισαπέχοντα σημεία $\{x_n\}$ με απόσταση h. Η γενική γραμμική μέθοδος k-βημάτων γράφεται:

$$\sum_{j=0}^{k} a_j \, y_{n+j} = h \sum_{j=0}^{k} \beta_j \, f\left(x_{n+j}, y_{n+j}\right) \tag{24.2}$$

όπου οι συντελεστές $a_0, a_1, ..., a_k$ και $\beta_0, \beta_1, ..., \beta_k$ είναι πραγματικές σταθερές και πρέπει να προσδιοριστούν. Υποθέτουμε ότι $a_k \neq 0$, με $a_k = 1$ και $|a_0| + |\beta_0| \neq 0$.

1). Αν $\beta_k = 0$, τότε η προσέγγιση y_{n+k} υπολογίζεται άμεσα από τις k προηγούμενες τιμές y_j και $f(x_j, y_j)$, $j = 0,1,...,k-1$, και τότε η μέθοδος k-βημάτων λέμε ότι είναι άμεση.

2). $\beta_k \neq 0$, τότε η άγνωστη προσέγγιση y_{n+k}, εμφανίζεται και στα δύο μέλη της ισότητας, και ειδικότερα απαιτείται η f να υπολογιστεί στο σημείο (x_{n+k}, y_{n+k}), για το οποίο το y_{n+k} είναι το ζητούμενο. Στην περίπτωση αυτή η μέθοδος είναι έμμεση. Η μέθοδος (24.2) ονομάζεται γραμμική, επειδή περιέχει γραμμικούς συνδυασμούς των τιμών $\{y_n\}$ και $\{f(x_n, y_n)\}$.

Παράδειγμα 24.1 Θα δώσουμε μερικά χαρακτηριστικά παραδείγματα πολυβηματικών μεθόδων διαφόρων βημάτων

a) Η μέθοδος:

$$y_{n+1} = y_n + h\, f\left(x_{n+1}, y_{n+1}\right)$$

Είναι μία έμμεση γραμμική μέθοδος ενός βήματος, γνωστή σαν έμμεση μέθοδος Euler.

b) Επίσης η μέθοδος:

$$y_{n+1} = y_n + \frac{h}{2}\left\{ f_{n+1} + f_n \right\}$$

είναι μία έμμεση γραμμική μέθοδος ενός βήματος, γνωστή ήδη σαν έμμεση μέθοδος τραπεζίου.

c) Η μέθοδος της μορφής:

$$y_{n+4} = y_{n+3} + \frac{h}{24}\left\{9 f_{n+4} + 19 f_{n+3} - 5 f_{n+2} - 9 f_{n+1}\right\}$$

είναι μία έμμεση γραμμική μέθοδος τεσσάρων βημάτων, γνωστή σαν μέθοδος AdamsMoulton τεσσάρων βημάτων.

d) Η μέθοδος της μορφής:

$$y_{n+4} = y_{n+3} + \frac{h}{24}\left\{9 f_{n+3} - 519 f_{n+2} + 37 f_{n+1} - 9 f_n\right\}$$

είναι μία άμεση γραμμική μέθοδος τεσσάρων βημάτων, γνωστή σαν μέθοδος AdamsBashforth τεσσάρων βημάτων.

Στην συνέχεια θα μελετήσουμε την κατηγορία αυτή των μεθόδων ως προς τα διάφορα βασικά χαρακτηριστικά, καθώς και τον τρόπο κατασκευής τους

25. Κατασκευή των πολυβηματικών μεθόδων

Υπάρχουν διάφοροι τρόποι για να κατασκευάσουμε μία πολυβηματική μέθοδο δοσμένης μορφής. Στην συνέχεια θα αναφερθούμε με συντομία σε μερικούς απ' αυτούς.

a) Κατασκευή με ανάπτυγμα Taylor

Θεωρούμε το ανάπτυγμα Taylor της συνάρτησης $y(x)$, στο σημείο $x_n + h$, οπότε προκύπτει η σχέση:

$$y(x_{n+1}) = y(x_n + h) = y(x_n) + h\, y'(x_n) + \frac{h^2}{2!} y''(x_n) + \ldots + \frac{h^p}{p!} y^{(p)}(x_n) + \ldots \qquad (25.1)$$

Αν στο ανάπτυγμα αυτό κρατήσουμε τους δύο πρώτους όρους, και αντικαταστήσουμε την $y'(x_n)$ από την διαφορική εξίσωση, θα προκύψει η προσεγγιστική σχέση:

$$y(x_n + h) \approx y(x_n) + h\, f\left(x_n, y(x_n)\right)$$

η οποία είναι ακριβής για τις αντίστοιχες προσεγγιστικές τιμές, οπότε προκύπτει η μέθοδος:

$$y_{n+1} - y_n = h\, f_n,$$

γνωστή σαν άμεση μέθοδος Euler.

Ας θεωρήσουμε τώρα τα αναπτύγματα Taylor για τις συναρτήσεις $y(x_n+h)$ και $y(x_n-h)$ γύρω από το σημείο x_n. Τότε θα έχουμε τις εκφράσεις:

$$y(x_n+h)=y(x_n)+h\,y'(x_n)+\frac{h^2}{2!}y''(x_n)+\frac{h^3}{3!}y'''(x_n)+\ldots$$

και

$$y(x_n-h)=y(x_n)-h\,y'(x_n)+\frac{h^2}{2!}y''(x_n)-\frac{h^3}{3!}y'''(x_n)+\ldots$$

Αφαιρώντας κατά μέλη προκύπτει η σχέση:

$$y(x_n+h)-y(x_n-h)=2h\,y'(x_n)+2\cdot\frac{h^3}{3!}y'''(x_n)+\ldots\ ,$$

από την οποία αν κρατήσουμε μόνο τον πρώτο όρο στο δεύτερο μέλος θα προκύψει η προσεγγιστική σχέση:

$$y(x_n+h)-y(x_n-h)\approx 2h\,y'(x_n)$$

από την οποία εύκολα συνάγεται η μέθοδος δύο βημάτων της μορφής:

$$y_{n+1}-y_{n-1}=2h\,f_n\,.$$

b) Κατασκευή με ολοκλήρωση

Θεωρούμε τα σημεία:

$$x_{n-1}\ ,\quad x_n=x_{n-1}+h\ ,\quad x_{n+1}=x_{n-1}+2h\,,$$

και την ταυτότητα:

$$y(x_{n+1})-y(x_{n-1})=\int_{x_{n-1}}^{x_{n+1}}y'(x)dx \tag{25.2}$$

Αν προσεγγίσουμε το ολοκλήρωμα στην (25.2) με τον κανόνα του Simpson, θα προκύψει η σχέση:

$$y(x_{n+1})\approx y(x_{n-1})+\frac{h}{3}\{f(x_{n-1},y(x_{n-1}))+4\,f(x_n,y(x_n))+f(x_{n+1},y(x_{n+1}))\}$$

η οποία οδηγεί στην μέθοδο:

$$y_{n+1}=y_{n-1}+\frac{h}{3}\{f_{n-1}+4\,f_n+f_{n+1}\} \tag{25.3}$$

όπου $f_{n+j}=f(x_{n+j},y_{n+j}),\quad j=-1,0,1\,.$

Η μέθοδος είναι έμμεση δύο βημάτων, και απαιτείται να γνωρίζουμε τις προσεγγίσεις της λύσης y_{n-1}, y_n στα σημεία x_{n-1}, x_n, για να προχωρήσουμε στον υπολογισμό της προσέγγισης της λύσης στο σημείο x_{n+1}.

c) Κατασκευή με παρεμβολή

Μία ιδιαιτέρως εύκολη διαδικασία για την κατασκευή πολυβηματικών μεθόδων βασίζεται στην χρήση παρεμβολικών πολυωνύμων και την ολοκλήρωση. Οι οικογένειες μεθόδων που κατασκευάζονται με αυτόν τον τρόπο, έχουν πολύ χρήσιμα χαρακτηριστικά τα οποία κάνουν την εφαρμογή τους στον υπολογιστή πολύ ελκυστική. Ειδικότερα οι συντελεστές τους προκύπτουν με έναν επαναληπτικό τρόπο, και επομένως μία αλλαγή της τάξης σημαίνει την πρόσθεση ή την αφαίρεση ενός όρου.

Στην παράγραφο αυτήν θα ασχοληθούμε με το γενικό πλαίσιο κατασκευής μεθόδων με ολοκλήρωση και ειδικότερα με τις μεθόδους τύπου Adams.

Ας θεωρήσουμε το γνωστό πρόβλημα αρχικών τιμών (24.2), και την ακολουθία $\{x_k\}$, έτσι ώστε:

$$x_k = x_{k-1} + h,\; k = 1, 2, \ldots$$

Από την (24.2) ορίζεται η σχέση:

$$\int_{x_n}^{x_{n+1}} y'(x)dx = \int_{x_n}^{x_{n+1}} f(x, y(x))dx$$

οπότε προκύπτει:

$$y(x_{n+1}) = y(x_n) + \int_{x_n}^{x_{n+1}} f(x, y(x))\, dx\ . \tag{25.4}$$

Συνήθως το παραπάνω ολοκλήρωμα δεν μπορεί να υπολογιστεί ακριβώς. Είναι εύκολο όμως να προσεγγίσουμε την $f(x, y(x))$, χρησιμοποιώντας ένα πολυώνυμο παρεμβολής το οποίο βασίζεται σε γνωστές τιμές $(x_k, y(x_k))$. Η ολοκλήρωση αυτού του παρεμβολικού πολυωνύμου ορίζει ένα τύπο ο οποίος ονομάζεται ***quadrature***.

Έστω $P_q(x)$ ένα πολυώνυμο βαθμού q το οποίο ικανοποιεί τις παρεμβολικές συνθήκες:

$$P_q(x_k) = f(x_k, y(x_k)),\quad k = n-q, n-q+1, \ldots, n-1, n\ .$$

Τότε ισχύει η σχέση:

$$f(x, y(x)) = P_q(x) + E_q(x),\quad E_q(x_k) = 0,\quad k = n-q, \ldots, n-1, n\ ,$$

όπου $E_q(x)$ είναι το σφάλμα της παρεμβολής. Από την (25.4) προκύπτει η σχέση:

$$y(x_{n+1}) = y(x_n) + \int_{x_n}^{x_{n+1}} P_q(x)\,dx + \int_{x_n}^{x_{n+1}} E_q(x)\,dx \tag{25.5}$$

η οποία αποτελεί την βάση για την κατασκευή της μεθόδου πολλών βημάτων:

$$y_{n+1} = y_n + \int_{x_n}^{x_{n+1}} U_q(x)dx\ , \tag{25.6}$$

όπου η

$$U_q(x_k) = f_k = f(x_k, y_k),\quad n-q \le k \le n,$$

παρεμβάλλει τις προσεγγιστικές τιμές της f οι οποίες ορίζονται από την αριθμητική λύση.

Αν θεωρήσουμε για παράδειγμα την έκφραση του πολυωνύμου Lagrange για το $U_q(x)$ και μεταβάλλουμε τις τιμές q, μπορεί από τον τύπο (25.6) να προκύψει εύκολα μία οικογένεια πολυβηματικών μεθόδων τύπου Adams. Άλλου τύπου πολυβηματικές μέθοδοι μπορούν να προκύψουν αν αλλάξουμε τα όρια της ολοκλήρωσης.

Τέλος το τοπικό σφάλμα αποκοπής μίας πολυβηματικής μεθόδου ορίζεται από την σχέση:

$$T_{n+1} = y(x_n) + \int_{x_n}^{x_{n+1}} P_q(x)\,dx - y(x_{n+1})$$
$$= -\int_{x_n}^{x_{n+1}} E_q(x)\,dx .$$

26 Μηδενική ευστάθεια (zero-stability)

Από την σχέση (24.2) γίνεται φανερό ότι χρειαζόμαστε k αρχικές προσεγγίσεις $y_0,\ldots,y_{k-1}$ της λύσης, για να μπορέσουμε να εφαρμόσουμε μία μέθοδο k- βημάτων για την προσέγγιση της λύσης του προβλήματος (24.1). Απ' αυτές, η y_0 δίνεται από την αρχική συνθήκη του προβλήματος, αλλά οι υπόλοιπες $y_1,\ldots,y_{k-1}$ πρέπει να υπολογιστούν με κάποιο τρόπο. Υπάρχουν διάφοροι τρόποι, ο πλέον συνηθισμένος είναι να εφαρμόσουμε μία μέθοδο απλού βήματος, κυρίως μία μέθοδο Runge-Kutta, η οποία θα πρέπει να είναι ανάλογης ακρίβειας με την μέθοδο πολλών βημάτων. Σε κάθε περίπτωση, οι αρχικές προσεγγίσεις θα εμπεριέχουν αριθμητικά σφάλματα, και είναι ενδιαφέρον να γνωρίζουμε πως αυτά θα επηρεάσουν τις επόμενες προσεγγίσεις y_n, $n \geq k$, οι οποίες θα υπολογιστούν εφαρμόζοντας την μέθοδο (24.2). Αυτό μας οδηγεί στο να μελετήσουμε την **ευστάθεια** της αριθμητικής μεθόδου, όταν υπάρχουν μικρές διαταραχές στις αρχικές προσεγγίσεις. Αρχικά διατυπώνουμε τον ορισμό.

Ορισμός 26.1 Μία γραμμική μέθοδος k- βημάτων για την επίλυση του προβλήματος (24.1), λέμε ότι είναι **μηδενικά-ευσταθής** (zero-stable), αν υπάρχει σταθερά K, έτσι ώστε για δύο οποιεσδήποτε ακολουθίες $\{y_n\}$ και $\{\hat{y}_n\}$ προσεγγίσεων, οι οποίες έχουν υπολογιστεί με την ίδια μέθοδο αλλά διαφορετικές αρχικές προσεγγίσεις $y_0, y_1,\ldots,y_{k-1}$ και $\hat{y}_0, \hat{y}_1, \ldots, \hat{y}_{k-1}$ αντίστοιχα, να ισχύει:

$$|\, y_n - \hat{y}_n| \leq K \cdot \max\left\{|\, y_0 - \hat{y}_0|, |\, y_1 - \hat{y}_1|, \ldots, |\, y_{k-1} - \hat{y}_{k-1}|\right\}$$

για $x_n \leq X_M \leq b$.

Συνήθως, για να δείξουμε ότι μία αριθμητική μέθοδος είναι ή όχι ευσταθής, παρατηρούμε την συμπεριφορά της, εφαρμόζοντας την μέθοδο στην στοιχειώδη διαφορική εξίσωση $y' = 0$. Αυτός είναι και ο λόγος για τον οποίο το είδος της ευστάθειας που εκφράζεται με τον ορισμό 26.1, ονομάζεται μηδενική ευστάθεια.

Όμως, η μελέτη της μηδενικής-ευστάθειας μιας μεθόδου με βάση τον ορισμό 26.1 είναι μία ανίερη διαδικασία και αποφεύγεται. Στην συνέχεια, θα ορίσουμε ένα αλγεβρικό ισοδύναμο του ορισμού της μηδενική; Ευστάθειας, το οποίο είναι γνωστό σαν **συνθήκη ρίζας** (root condition), η οποία θα απλοποιεί την διαδικασία. Για τον σκοπό ορίζουμε δύο **χαρακτηριστικά πολυώνυμα**, το $1^{\text{ο}}$ και $2^{\text{ο}}$, τα οποία ορίζονται από τους συντελεστές της μεθόδου και είναι της μορφής:

$$\rho(z)=\sum_{j=0}^{k} a_j z^j \tag{26.1}$$

$$\sigma(z)=\sum_{j=0}^{k} \beta_j z^j \tag{26.2}$$

αντίστοιχα. Υποθέτουμε ότι $a_k \neq 0$ και $a_0^2+\beta_0^2 \neq 0$. Το βασικό αποτέλεσμα αυτής της παραγράφου, δίνεται από τα συμπεράσματα του επόμενου θεωρήματος.

Θεώρημα 26.1 Μία γραμμική μέθοδος πολλών βημάτων είναι μηδενικά-ευσταθής για κάθε συνήθη διαφορική εξίσωση της μορφής (24.1), όπου η f ικανοποιεί μία συνθήκη Lipschitz της μορφής:

$$| f(x,y)-f(x,z)| \leq L\, | y-z|,$$

αν και μόνο αν το πρώτο χαρακτηριστικό πολυώνυμο της μεθόδου έχει όλες τις ρίζες μέσα στον κλειστό μοναδιαίο δίσκο, και όσες βρίσκονται στον μοναδιαίο κύκλο να είναι απλές.

Παράδειγμα 26.1 Να εξετασθεί η μηδενική ευστάθεια των μεθόδων του παραδείγματος 24.1.

a) Το πρώτο χαρακτηριστικό πολυώνυμο της έμμεσης μεθόδουEuler είναι $\rho(z)=z-1$. Η εξίσωση $\rho(z)=0$, έχει μία ρίζα $z=1$, που βρίσκεται στον μοναδιαίο κύκλο και είναι απλή. Άρα η μέθοδος είναι μηδενικά-ευσταθής.

b) Η έμμεση μέθοδος τραπεζίου έχει το ίδιο χαρακτηριστικό πολυώνυμο με την προηγούμενη περίπτωση, άρα είναι μηδενικά-ευσταθής

c) Οι μέθοδοι Adams-Bashforth και Adams-Moulton έχουν πρώτο χαρακτηριστικό πολυώνυμο της μορφής:

$$\rho(z)=z^4-z^3=z^3(z-1),$$

του οποίου οι ρίζες βρίσκονται μέσα στον μοναδιαίο δίσκο, και αυτή που βρίσκεται στον μοναδιαίο κύκλο ($z=1$) είναι απλή. Άρα και οι δύο μέθοδοι είναι μηδενικά-ευσταθείς.

Παράδειγμα 26.2 Η μέθοδος τριών βημάτων της μορφής:

$$11y_{n+3}+27y_{n+2}-27y_{n+1}-11y_n=3h\left\{ f_{n+3}+9f_{n+2}+5f_{n+1}+f_n \right\},$$

έχει πρώτο χαρακτηριστικό πολυώνυμο το

$$\rho(z)=11z^3+27z^2-27z-11,$$

του οποίου οι ρίζες είναι $z_1=1$, $z_2 \approx -0.3189$ και $z_2 \approx -3.1356$. Άρα, $|z_3|>1$, επομένως η μέθοδος δεν είναι μηδενικά-ευσταθής

27 Ακρίβεια και Συνέπεια

Στην παράγραφο αυτή θα μελετήσουμε την **τάξη ακρίβειας** και την **συνέπεια** της μεθόδου (24.2). Για τον σκοπό αυτό θα θεωρήσουμε, όπως και στην περίπτωση των μεθόδων απλού βήματος, το τοπικό σφάλμα αποκοπής.

Αν θεωρήσουμε ότι η $y(x)$ είναι μία λύση της διαφορικής εξίσωσης $y' = f(x, y)$, τότε το τοπικό σφάλμα αποκοπής της μεθόδου (24.2) ορίζεται από την σχέση:

$$T_n = \sum_{j=0}^{k} a_j\, y(x_{n+j}) - h \sum_{j=0}^{k} \beta_j\, f(x_{n+j}, y(x_{n+j}))$$

ή

$$T_n = \sum_{j=0}^{k} a_j\, y(x_{n+j}) - h \sum_{j=0}^{k} \beta_j\, y'(x_{n+j}) \tag{27.1}$$

Η σχέση (27.1) μπορεί να κανονικοποιηθεί κατάλληλα, ώστε το T_n να παραβάλλεται με το $y'(x) - f(x, y(x))$, οπότε μπορούμε να γράψουμε:

$$T_n = \frac{\sum_{j=0}^{k} a_j\, y(x_{n+j}) - h \sum_{j=0}^{k} \beta_j\, y'(x_{n+j})}{h \sum_{j=0}^{k} \beta_j} \tag{27.2}$$

Στην περίπτωση αυτή θα πρέπει $\sum_{j=0}^{k} \beta_j \neq 0$, και με βάση την (26.2) $\sigma(1) \neq 0$.

Υποθέτουμε ότι η λύση είναι επαρκώς λεία, και εφαρμόζουμε το ανάπτυγμα Taylor στις συναρτήσεις $y(x_{n+j})$ και $y'(x_{n+j})$. Τότε ισχύουν οι σχέσεις:

$$y(x_{n+j}) = y(x_n) + (jh)\, y'(x_n) + \frac{(jh)^2}{2!} y''(x_n) + \ldots + \frac{(jh)^p}{p!} y^{(p)}(x_n) + \ldots \tag{27.3}$$

$$y'(x_{n+j}) = y'(x_n) + (jh)\, y''(x_n) + \frac{(jh)^2}{2!} y'''(x_n) + \ldots + \frac{((j-1)h)^{p-1}}{(p-1)!} y^{(p)}(x_n) + \ldots \tag{27.4}$$

Αν αντικαταστήσουμε τις (27.3) και (27.4) στην (27.2), και εκφράσουμε τον αριθμητή σε δυνάμεις του h, θα προκύψει η σχέση:

$$T_n = \frac{1}{h\,\sigma(1)} \left\{ C_0\, y(x_n) + C_1\, h\, y'(x_n) + \ldots + C_p h^p y^{(p)}(x_n) + \ldots \right\} \tag{27.5}$$

όπου υπολογίζεται ότι:

$$C_0 = \sum_{j=0}^{k} a_j$$

$$C_1 = \sum_{j=1}^{k} j\, a_j - \sum_{j=0}^{k} \beta_j$$

$$C_2 = \sum_{j=1}^{k} \frac{j^2}{2!} a_j - \sum_{j=1}^{k} j\beta_j \qquad (27.6)$$

.

.

$$C_p = \sum_{j=1}^{k} \frac{j^p}{p!} a_j - \sum_{j=1}^{k} \frac{j^{p-1}}{(p-1)!} \beta_j$$

Μπορούμε τώρα να διατυπώσουμε τους παρακάτω ορισμούς.

Ορισμός 27.1 Η αριθμητική μέθοδος (24.2) θα λέμε ότι είναι **συνεπής** με την διαφορική εξίσωση (24.1), αν το σφάλμα αποκοπής που ορίζεται από την σχέση (27.2) είναι τέτοιο ώστε, για κάθε $\varepsilon > 0$, υπάρχει αριθμός $h(\varepsilon)$ για τον οποίον ισχύει:

$$|T_n| < \varepsilon \quad \text{για} \quad 0 < h < h(\varepsilon),$$

για οποιαδήποτε $(k+1)$ σημεία $(x_n, y(x_n)), \ldots, (x_{n+k}, y(x_{n+k}))$ πάνω σε οποιαδήποτε καμπύλη της λύσης στο R, του προβλήματος αρχικών τιμών (24.1).

Σύμφωνα με τον ορισμό (27.1), για την συνέπεια θα πρέπει $T_n \to 0$ όταν $h \to 0$, και αυτό απαιτεί ότι πρέπει $C_0 = 0$ και $C_1 = 0$. Τότε, σύμφωνα με τα δύο χαρακτηριστικά πολυώνυμα η μέθοδος θα είναι συνεπής αν ισχύουν οι συνθήκες:

$$\rho(1) = 0 \text{ και } \rho'(1) = \sigma(1) \neq 0. \qquad (27.7)$$

Μπορούμε τώρα να παρατηρήσουμε ότι, σύμφωνα με την συνθήκη (27.7), αν μία γραμμική πολυβηματική μέθοδος είναι συνεπής, τότε θα έχει μία απλή ρίζα πάνω στον μοναδιαίο κύκλο στο $z = 1$, επομένως η συνθήκη της ρίζας δεν παραβιάζεται.

Ορισμός 27.1 Η αριθμητική μέθοδος (24.2) θα λέμε ότι έχει **τάξη ακρίβειας** p, αν p είναι ο μεγαλύτερος θετικός ακέραιος έτσι ώστε, για οποιαδήποτε λεία λύση στο R, του προβλήματος (24.1), υπάρχουν σταθερές K και h_0, έτσι ώστε:

$$|T_n| < K h^p \quad \text{για} \quad 0 < h < h_0,$$

για οποιαδήποτε $(k+1)$ σημεία $(x_n, y(x_n)), \ldots, (x_{n+k}, y(x_{n+k}))$ πάνω στην καμπύλη της λύσης.

Από την (27.5) μπορούμε εύκολα να συμπεράνουμε ότι η μέθοδος είναι τάξης p, αν και μόνο αν

$$C_0 = C_1 = \ldots = C_p = 0 \text{ και } C_{p+1} \neq 0. \qquad (27.8)$$

Στην περίπτωση αυτή το σφάλμα αποκοπής δίνεται από την σχέση:

$$T_n = \frac{C_{p+1}}{\sigma(1)} h^{p}\, y^{(p+1)}(x_n), \tag{27.9}$$

και ο αριθμός $C_{p+1} \neq 0$, ονομάζεται **σταθερά σφάλματος** της μεθόδου.

Παρατήρηση Οι σχέσεις (27.6) μπορούν να χρησιμοποιηθούν για να υπολογίσουμε την τάξη και την σταθερά σφάλματος για οποιαδήποτε δεδομένη μέθοδο πολλών βημάτων. Επίσης, μπορούμε να χρησιμοποιήσουμε αυτές τις εκφράσεις για να κατασκευάσουμε γραμμικές πολυβηματικές μεθόδους γνωστής δομής.

Παράδειγμα 27.1 Να υπολογιστεί μία έμμεση γραμμική μέθοδος δύο βημάτων μεγίστης τάξης, με μία ελεύθερη παράμετρο. Να υπολογιστεί επίσης η τάξη και η σταθερό σφάλματος της μεθόδου.

Λύση Μία μέθοδος δύο βημάτων μπορεί να εκφραστεί από την σχέση:

$$y_{n+2} + a_1 y_{n+1} + a_0 y_n = h\left\{\beta_2 f_{n+2} + \beta_1 f_{n+1} + \beta_0 f_n\right\}$$

όπου το $a_2 = 1,\ \beta_2 \neq 0$. Θεωρούμε το $a_0 = a$ ως παράμετρο, οπότε θα έχουμε την μορφή:

$$y_{n+2} + a_1 y_{n+1} + a\, y_n = h\left\{\beta_2 f_{n+2} + \beta_1 f_{n+1} + \beta_0 f_n\right\}$$

Πρέπει να υπολογίσουμε τους άγνωστους συντελεστές a_1, β_2, β_1, β_0. Άρα απαιτούνται τέσσερις εξισώσεις, οι οποίες προκύπτουν από τους τύπους (27.6),απαιτώντας να ισχύουν οι συνθήκες $C_0 = C_1 = C_2 = C_3 = 0$, ώστε να μεγιστοποιηθεί η τάξη ακρίβειας της μεθόδου.

Με βάση τους τύπους αυτούς θα έχουμε:

$$\begin{aligned}
C_0 &= a_0 + a_1 + a_2 \\
C_1 &= \left(a_1 + 2a_2\right) - \left(\beta_0 + \beta_1 + \beta_2\right) \\
C_2 &= \left(\frac{1}{2}a_1 + 2a_2\right) - \left(\beta_1 + 2\beta_2\right) \\
C_3 &= \left(\frac{1}{6}a_1 + \frac{8}{6}a_2\right) - \left(\frac{1}{2}\beta_1 + 2\beta_2\right)
\end{aligned}$$

Λαμβάνοντας υπ’ όψιν ότι $a_2 = 1$ και $a_0 = a$ προκύπτουν οι εξισώσεις:

$$\begin{aligned}
& a_0 + a_1 + a_2 = 0 \\
& \left(a_1 + 2\right) - \left(\beta_0 + \beta_1 + \beta_2\right) = 0 \\
& \left(\frac{1}{2}a_1 + 2\right) - \left(\beta_1 + 2\beta_2\right) = 0 \\
& \left(\frac{1}{6}a_1 + \frac{4}{3}\right) - \left(\frac{1}{2}\beta_1 + 2\beta_2\right) =
\end{aligned}$$

Η λύση του συστήματος αυτού μας δίνει τις τιμές:

$$a_1 = -\left(1 + a\right),\ \ \beta_0 = -\frac{1}{12}\left(1 + 5a\right),\ \ \beta_1 = \frac{2}{3}\left(1 - a\right),\ \ \beta_2 = \frac{1}{12}\left(5 + a\right),$$

οπότε προκύπτει η μονοπαραμετρική έμμεση μέθοδος δύο βημάτων της μορφής:

$$y_{n+2}-(1+a)y_{n+1}+a\,y_n=\frac{h}{12}\{(5+a)\,f_{n+2}+8(1-a)\,f_{n+1}-(1+5a)f_n\}.$$

Επί πλέον υπολογίζεται ότι:

$$C_4=\frac{1}{4!}(a_1+16)-\frac{1}{3!}(\beta_1+8\beta_2)=-\frac{1}{4!}(1+a)$$

$$C_5=\frac{1}{5!}(a_1+32)-\frac{1}{4!}(\beta_1+16\beta_2)=-\frac{1}{3\cdot 5!}(17+13a)$$

a) Αν $a\neq -1$, τότε $C_4\neq 0$ και θα λέμε ότι η μέθοδος είναι τρίτης τάξης με σταθερά σφάλματος το $C_4=-\frac{1}{4!}(1+a)$.

b) Αν $a=-1$, τότε $C_4=0$ αλλά $C_5\neq 0$, και η μέθοδος παίρνει την μορφή:

$$y_{n+2}-y_n=\frac{h}{3}\{f_{n+2}+4f_{n+1}+f_n\},$$

η οποία είναι η μέθοδος Simpson, μία μέθοδος 2 βημάτων τάξης 4 και με σταθερά σφάλματος τον αριθμό $C_5=-\frac{4}{3\cdot 4!}$.

28 Σύγκλιση

Οι έννοιες μηδενική ευστάθεια και συνέπεια έχουν μεγάλο θεωρητικό ενδιαφέρον. Όμως, από την πρακτική πλευρά, εκείνο που έχει σημασία είναι οι αριθμητικές προσεγγίσεις y_n, στα σημεία της διαμέρισης x_n, $n=0,1,\ldots,N$ να βρίσκονται κοντά στις αντίστοιχες τιμές $y(x_n)$ της αναλυτικής λύσης. Επίσης, το ολικό σφάλμα $E_n=y(x_n)-y_n$, πρέπει να μειώνεται όταν το βήμα h μικραίνει. Τώρα μπορούμε να διατυπώσουμε το παρακάτω τυπικό ορισμό σύγκλισης.

Ορισμός 28.1 Η γραμμική μέθοδος k-βημάτων (24.2) λέμε ότι είναι συγκλίνουσα, αν για όλα τα προβλήματα αρχικών τιμών της μορφής (24.1) τα οποία πληρούν τις υποθέσεις του θεωρήματος (2.1), η σχέση

$$\lim_{\substack{h\to 0\\ nh=x-x_0}} y_n=y(x_n) \tag{28.1}$$

Ισχύει για όλα τα $x\in[x_0,X_M]$, και όλες τις λύσεις $\{y_n\}_{n=0}^{N}$ της εξίσωσης διαφορών (24.2), με συνεπείς αρχικές συνθήκες, δηλαδή, με αρχικές συνθήκες $y_s=y_s(h)$, $s=0,1,\ldots,k-1$, για τις οποίες ισχύει ότι:

$$\lim_{h\to 0} y_s(h)=y_0,\ \ s=0,1,\ldots,k-1.$$

Η τελευταία υπόθεση του ορισμού, προκύπτει από το γεγονός ότι οι αρχικές τιμές $y_s(h)$ στην πράξη, δεν είναι δεδομένες και πρέπει να υπολογιστούν με κάποια άλλη προσεγγιστική μέθοδο αριθμητικά. Στην συνέχεια θα αναφερθούμε με συντομία στην σχέση μεταξύ μηδενικής ευστάθειας, συνέπειας και σύγκλισης. Για τον σκοπό θα αναφερθούμε σε μερικά βασικά θεωρήματα (χωρίς απόδειξη) τα οποία θεμελιώνουν την σχέση αυτή.

Θεώρημα 28.1 Μία αναγκαία συνθήκη για την σύγκλιση της μεθόδου (24.2) είναι αυτή να είναι μηδενικά ευσταθής.

Θεώρημα 28.2 Μία αναγκαία συνθήκη για την σύγκλιση της μεθόδου (24.2) είναι αυτή να είναι συνεπής.

Θεώρημα 28.3 Για μία γραμμική πολυβηματική μέθοδο, η οποία είναι συνεπής με την διαφορική εξίσωση (24.1), όπου η συνάρτηση f πληροί μία συνθήκη Lipschitz, και η οποία εφαρμόζεται με συνεπείς αρχικές συνθήκες, η μηδενική ευστάθεια είναι ικανή για την σύγκλιση της μεθόδου.

Ο συνδυασμός των τριών προηγούμενων θεωρημάτων οδηγεί στο πολύ βασικό θεώρημα αυτής της περιοχής.

Θεώρημα 28.4 (Dahlquist) Για μία γραμμική πολυβηματική μέθοδο, η οποία είναι συνεπής με την διαφορική εξίσωση (24.1), όπου η συνάρτηση f πληροί μία συνθήκη Lipschitz, και η οποία εφαρμόζεται με συνεπείς αρχικές συνθήκες, η μηδενική ευστάθεια είναι αναγκαία και ικανή συνθήκη για την σύγκλιση της μεθόδου. Επί πλέον, αν η λύση $y(x)$ έχει συνεχείς παραγώγους μέχρι τάξη $(p+1)$ και $T_n = O(h^p)$, τότε το ολικό σφάλμα θα είναι:

$$E_n = y(x_n) - y_n = O(h^p).$$

Με βάση το θεώρημα Dahlquist, αν μία πολυβηματική μέθοδος δεν είναι μηδενικά ευσταθής, το ολικό σφάλμα της μεθόδου δεν μπορεί να γίνει όσο θέλουμε μικρό όταν το h είναι αρκετά μικρό, για οποιεσδήποτε επαρκώς ακριβείς αρχικές συνθήκες. Στην πραγματικότητα, αν η συνθήκη της ρίζας παραβιαστεί, τότε υπάρχει μία λύση της γραμμικής πολυβηματικής μεθόδου, η οποία αυξάνει ανεξέλεγκτα σε ένα δεδομένο διάστημα του x, ανεξάρτητα αν οι αρχικές συνθήκες είναι ακριβείς. Αυτό το αποτέλεσμα φωτίζει την σημασία της μηδενικής ευστάθειας, και την σχέση της με τους υπολογισμούς στην πράξη.

29 Μέγιστη τάξη πολυβηματικής μεθόδου μηδενικά-ευσταθούς

Υποθέτουμε ότι έχουν ήδη επιλεγεί οι συντελεστές a_j, $j = 0, 1, \ldots, k$ μιας πολυβηματικής μεθόδου της μορφής (24.2). Το ερώτημα που τίθεται είναι πως μπορούμε να επιλέξουμε τους συντελεστές β_j, $j = 0, 1, \ldots, k$, έτσι ώστε η τάξη ακρίβειας της μεθόδου να είναι όσο το δυνατόν μεγαλύτερη.

Με βάση το θεώρημα 28.4, ενδιαφερόμαστε για μεθόδους συνεπείς, επομένως είναι φυσικό να υποθέσουμε ότι το πρώτο και δεύτερο χαρακτηριστικό πολυώνυμο $\rho(z)$ και $\sigma(z)$ αντίστοιχα, τα οποία αντιστοιχούν στην μέθοδο (24.2), ικανοποιούν τις συνθήκες:

$$\rho(1) = 0 \text{ και } \rho'(1) = \sigma(1) \neq 0.$$

Θεωρούμε την συνάρτηση Φ της μιγαδικής μεταβλητής z, η οποία ορίζεται με την σχέση:

$$\Phi(z) = \frac{\rho(z)}{\log z} - \sigma(z).$$

Τότε μπορεί να αποδειχθεί το επόμενο θεμελιώδες λήμμα.

Λήμμα 29.1 Υποθέτουμε ότι p είναι ένας θετικός ακέραιος. Η γραμμική μέθοδος (24.2) με χαρακτηριστικά πολυώνυμα τα $\rho(z)$ και $\sigma(z)$, έχει τάξη ακρίβειας p, αν και μόνο αν ,η συνάρτηση $\Phi(z)$ έχει μία ρίζα με πολλαπλότητα p στο $z=1$.

Με την βοήθεια αυτού του λήμματος, μπορεί να αποδειχθεί το επόμενο θεώρημα το οποίο υπολογίζει ένα κάτω φράγμα για την μέγιστη τιμή μιας γραμμικής πολυβηματικής μεθόδου, με προδιαγεγραμμένο το πρώτο χαρακτηριστικό πολυώνυμο.

Θεώρημα 29.1 Υποθέτουμε ότι $\rho(z)$ είναι ένα πολυώνυμο βαθμού k, έτσι ώστε $\rho(1)=0$ και $\rho'(1)\neq 0$, και $\hat{k}$ ένας θετικός ακέραιος ώστε $0\leq\hat{k}\leq k$. Τότε υπάρχει ένα μοναδικό πολυώνυμο $\sigma(z)$ βαθμού $\hat{k}$, έτσι ώστε $\rho'(1)-\sigma(1)=0$, και η τάξη της γραμμικής πολυβηματικής μεθόδου η οποία ορίζεται με τα πολυώνυμα $\rho(z)$ και $\sigma(z)$ είναι $p\geq\hat{k}+1$.

Παρατήρηση Σε συνδυασμό με το θεώρημα 29.1, στην πλειοψηφία των μεθόδων με πρακτικό ενδιαφέρον, είτε $\hat{k}=k-1$ για μία άμεση μέθοδο, είτε $\hat{k}=k$ για μία έμμεση μέθοδο.

Εξετάζοντας τα πράγματα πιο αναλυτικά, η γραμμική μέθοδος k-βημάτων (24.2) έχει $2k+2$ συντελεστές a_j, β_j, $j=0,1,\ldots,k$, από τους οποίους ο $a_k=1$. Αυτό μας αφήνει με $2k+1$ ελεύθερες παραμέτρους αν η μέθοδος είναι έμμεση, και $2k$ αν είναι άμεση. Σύμφωνα με την σχέση (27.5), αν απαιτήσουμε η μέθοδος (24.2) να έχει τάξη p, θα πρέπει να ικανοποιούνται οι $(p+1)$ συνθήκες $C_0=C_1=\ldots=C_p=0$, οι οποίες εμπεριέχουν τους συντελεστές a_j, β_j, $j=0,1,\ldots,k$.

Επομένως, για την περίπτωση της έμμεσης μεθόδου μπορούμε να απαιτήσουμε να ικανοποιούνται οι $p+1=2k+1$ γραμμικές συνθήκες $C_0=C_1=\ldots=C_{2k+1}=0$, για να υπολογίσουμε τους άγνωστους συντελεστές, πράγμα που οδηγεί σε μία μέθοδο τάξης $p=2k$. Ομοίως, για μία άμεση μέθοδο η μέγιστη τάξη που μπορούμε να περιμένουμε είναι $p=2k-1$.

Δυστυχώς, δεν υπάρχει καμία σιγουριά ότι τέτοιες μέθοδοι είναι μηδενικά ευσταθείς. Αντίθετα, ο Dahlquist έχει αποδείξει ότι δεν υπάρχει συνεπής και μηδενικά ευσταθής μέθοδος k-βημάτων με τάξη ακρίβειας $p>k+2$. Επομένως, οι μέγιστες τάξεις $p=2k$ και $p=2k-1$ δεν μπορούν να επιτευχθούν χωρίς να παραβιαστεί η συνθήκη της μηδενικής ευστάθειας. Μπορούμε τώρα να διατυπώσουμε το επόμενο θεώρημα.

Θεώρημα 29.2 Δεν υπάρχει μηδενικά-ευσταθείς γραμμική μέθοδος k-βημάτων, της οποίας η τάξη να υπερβαίνει το $k+1$ αν k είναι περιττός, ή το $k+2$ αν k είναι άρτιος.

Μία μηδενικά-ευσταθής γραμμική μέθοδος k-βημάτων τάξης $p=k+2$, λέμε ότι είναι μία **βέλτιστη μέθοδος** (optimal method). Μπορεί να δειχθεί, (βλέπε απόδειξη του θεωρήματος 29.2), ότι όλες οι ρίζες του πρώτου χαρακτηριστικού πολυωνύμου $\rho(z)$ το οποίο αντιστοιχεί σε μία βέλτιστη γραμμική πολυβηματική μέθοδο, έχουν μέτρο ίσο με το 1. Όμως, οι βέλτιστες μέθοδοι έχουν διάφορα μειονεκτήματα σε σχέση με τις ιδιότητες ευστάθειας αυτών.

Οι γραμμικές μέθοδοι k-βημάτων των οποίων το πρώτο χαρακτηριστικό πολυώνυμο είναι της μορφής:

$$\rho(z)=z^{k}-z^{k-1} \tag{29.1}$$

ονομάζονται μέθοδοι Adams. Οι άμεσες μέθοδοι Adams αναφέρονται σαν μέθοδοι Adams-Bashforth και οι έμμεσες σαν μέθοδοι Adams-Moulton.

Οι γραμμικές μέθοδοι k-βημάτων των οποίων το πρώτο χαρακτηριστικό πολυώνυμο είναι της μορφής:

$$\rho(z) = z^{k} - z^{k-2} \tag{29.2}$$

ονομάζονται μέθοδοι Nystrom αν είναι άμεσες, και μέθοδοι Milne-Simpson αν είναι έμμεσες. Όλες αυτές είναι μηδενικά-ευσταθείς.

30 Απόλυτη ευστάθεια των πολυβηματικών μεθόδων

Μέχρι τώρα αναφερθήκαμε στην μηδενική ευστάθεια και συνέπεια μιας γραμμικής πολυβηματικής μεθόδου, και είδαμε ότι ο ρόλος της συνέπειας είναι να περιορίσει το μέγεθος των τοπικών σφαλμάτων, ενώ ο ρόλος της μηδενικής ευστάθειας είναι να εξασφαλίσει ότι τα τοπικά σφάλματα δεν θα μεταδίδονται από βήμα σε βήμα κατά τέτοιο τρόπο ώστε οι προσεγγίσεις μετά από ένα ορισμένο αριθμό βημάτων να είναι απαράδεκτες. Στα προηγούμενα, η ανάπτυξη της θεωρίας βασιζόταν στην υπόθεση ότι $h \to 0$, $n \to +\infty$, $n\,h$ σταθερό.

Στην παράγραφο αυτή θα θεωρήσουμε ότι το βήμα h είναι ένας σταθερός θετικός αριθμός, και θα μελετήσουμε την μετάδοση του σφάλματος όταν ο αριθμός n των βημάτων αυξάνει. Ειδικότερα, θέλουμε να διαπιστώσουμε ότι όταν η μέθοδος εφαρμόζεται σε ένα πρόβλημα αρχικών τιμών, του οποίου η λύση ελαττούται προς το μηδέν όταν το $x \to +\infty$, η γραμμική πολυβηματική μέθοδος παράγει λύσεις με ανάλογη συμπεριφορά, για $h > 0$ σταθερό και $x_n = x_0 + n\,h \to +\infty$.

Όπως είναι ήδη γνωστό, ένα πρόβλημα μοντέλο με την παραπάνω συμπεριφορά είναι της μορφής:

$$y' = \lambda\, y, \quad x > 0, \quad y(0) = y_0 \neq 0 \tag{30.1}$$

όπου $\mathrm{Re}\,\lambda < 0$. Για ευκολία θεωρούμε ότι ο αριθμός λ είναι ένας πραγματικός αριθμός, αλλά τα συμπεράσματα επεκτείνονται άμεσα για την γενική περίπτωση όπου λ είναι ένας μιγαδικός αριθμός με $\mathrm{Re}\,\lambda < 0$.

Αν εφαρμόσουμε την μέθοδο:

$$\sum_{j=0}^{k} a_j\, y_{n+j} = h \sum_{j=0}^{k} \beta_j\, f_{n+j} \tag{30.2}$$

στο πρόβλημα (30.1), θα προκύψει μία γραμμική εξίσωση διαφορών της μορφής:

$$\sum_{j=0}^{k} \left(a_j - h\,\lambda\,\beta_j\right) y_{n+j} = 0 \tag{30.3}$$

Η γενική λύση y_n της (30.3) μπορεί να εκφραστεί σαν ένας γραμμικός συνδυασμός δυνάμεων των ριζών της πολυωνυμικής εξίσωσης:

$$\sum_{j=0}^{k} \left(a_j - h\,\lambda\,\beta_j\right) r^{j} = 0 \tag{30.4}$$

Αν θεωρήσουμε τα χαρακτηριστικά πολυώνυμα της μεθόδου:

$$\rho(r)=\sum_{j=0}^{k} a_j r^j \quad \text{και} \quad \sigma(r)=\sum_{j=0}^{k} \beta_j r^j ,$$

τότε η εξίσωση (30.4) γράφεται:

$$\pi\left(r,\bar{h}\right)=\rho(r)-\bar{h}\,\sigma(r) \qquad (30.5)$$

όπου $\bar{h}=h\lambda$. Το πολυώνυμο $\pi\left(r,\bar{h}\right)$, ονομάζεται **πολυώνυμο ευστάθειας** της μεθόδου (30.2). Είναι φανερό ότι $\lim y_n \to 0$ για $h>0$ σταθερό και $n\to+\infty$, αν και μόνο αν όλες οι ρίζες του $\pi\left(r,\bar{h}\right)$ έχουν μέτρο < 1. Έτσι οδηγούμεθα στον επόμενο ορισμό.

Ορισμός 30.1 Η μέθοδος (30.2) λέμε ότι είναι **απόλυτα ευσταθής** για δοθέν $\bar{h}=h\lambda$, αν και μόνο αν για αυτό το $\bar{h}$, όλες οι ρίζες $r_s=r_s(\bar{h})$ του πολυωνύμου ευστάθειας $\pi\left(r,\bar{h}\right)$, ικανοποιούν την ανισότητα $\left|r_s\right|<1,\ s=1,2,...,k$. Διαφορετικά, η μέθοδος λέμε ότι είναι **απόλυτα ασταθής.** Ένα διάστημα $\left(a,b\right)$ της πραγματικής ευθείας ονομάζεται **διάστημα απόλυτης ευστάθειας**, αν η μέθοδος είναι απόλυτα ευσταθής για όλα τα $\bar{h}\in\left(a,b\right)$. Αν η μέθοδος είναι απόλυτα ασταθής για όλα τα $\bar{h}$, λέμε ότι η μέθοδος δεν έχει διάστημα απόλυτης ευστάθειας.

Επειδή για $\lambda>0$ η λύση της (30.1) εμφανίζει εκθετική αύξηση, είναι λογικό να αναμένουμε ότι μία συνεπής και μηδενικά-ευσταθής (και επομένως συγκλίνουσα) γραμμική πολυβηματική μέθοδος, θα υπολογίζει προσεγγίσεις με ανάλογη συμπεριφορά για $h>0$ αρκετά μικρό, και θα είναι επομένως απόλυτα ασταθής για μικρά $\bar{h}=h\lambda$. Τώρα μπορούμε να διατυπώσουμε το επόμενο θεώρημα.

Θεώρημα 30.1 Κάθε συνεπής και μηδενικά-ευσταθής γραμμική πολυβηματική μέθοδος είναι απόλυτα ασταθής για μικρά θετικά $\bar{h}$.

Από το θεώρημα 30.1 προκύπτει ότι, αν μία συνεπής και μηδενικά-ευσταθής γραμμική πολυβηματική μέθοδος εφαρμοστεί για την αριθμητική λύση του προβλήματος αρχικών τιμών

$$\begin{aligned} &y'=f\left(x,y\right),\ x\in\left[a,b\right] \\ &y\left(a\right)=y_0 \end{aligned} ,$$

όπου $\dfrac{\partial f}{\partial x}>0$, το σφάλμα της μεθόδου θα αυξάνει όσο προχωρούν οι υπολογισμοί.

Υπάρχουν διάφορες μέθοδοι με τις οποίες μπορούμε να υπολογίσουμε το διάστημα ή χωρίο απόλυτης ευστάθειας μιας γραμμικής πολυβηματικής μεθόδου. Μεταξύ άλλων, το κριτήριο του Schur και το κριτήριο Routh-Hurwitz είναι τα πλέον γνωστά.

a) Το κριτήριο Schur

Θεωρούμε το πολυώνυμο:

$$\Phi(r)=c_k r^k+c_{k-1} r^{k-1}+...+c_1 r+c_0\ ,\ c_k\neq 0\ ,\ c_0\neq 0 \qquad (30.6)$$

με μιγαδικούς συντελεστές. Το πολυώνυμο $\Phi(r)$, λέμε ότι είναι ένα Schur πολυώνυμο, αν κάθε μία από τις ρίζες του r_s ικανοποιεί την ανισότητα $\left|r_s\right|<1\ ,\ s=1,2,...,k$.

Ας θεωρήσουμε τώρα το πολυώνυμο:

$$\hat{\Phi}(r)=\bar{c}_0 r^k+\bar{c}_1 r^{k-1}+\ldots+\bar{c}_{k-1} r+\bar{c}_k \tag{30.7}$$

όπου $\bar{c}_j$, οι συζυγείς μιγαδικοί των c_j , $j=1,2\ldots,k$. Επί πλέον ορίζουμε το πολυώνυμο:

$$\Phi_1(r)=\frac{1}{r}\left\{\hat{\Phi}(0)\,\Phi(r)-\Phi(0)\hat{\Phi}(r)\right\} \tag{30.8}$$

όπου προφανώς το $\Phi_1(r)$ είναι βαθμού $\leq k-1$. Ισχύει το επόμενο θεώρημα.

Θεώρημα 30.2 Το πολυώνυμο $\Phi(r)$ είναι ένα Schur πολυώνυμο, αν και μόνο αν ισχύουν:

$$\left|\hat{\Phi}(0)\right|>\left|\Phi(0)\right|$$

και το $\Phi_1(r)$ είναι επίσης ένα Schur πολυώνυμο.

Είναι φανερό με βάση το θεώρημα 30.2, ότι το διάστημα (a,b) είναι ένα διάστημα απόλυτης ευστάθειας μιας γραμμικής πολυβηματικής μεθόδου, αν για όλα τα $\bar{h}\in(a,b)$, το πολυώνυμο ευστάθειας $\pi(r,\bar{h})$ της μεθόδου είναι ένα Schur πολυώνυμο. Το κριτήριο αυτό μπορεί να εφαρμοστεί επαναληπτικά, όπου σε κάθε βήμα εφαρμογής προκύπτει ένα πολυώνυμο βαθμού κατά ένα μικρότερο από το προηγούμενο, οπότε διαδοχική χρήση του κριτηρίου οδηγεί σε ένα κοινό διάστημα του $\bar{h}$, το οποίο είναι και το διάστημα απόλυτης ευστάθειας.

Παράδειγμα 30.1 Να εφαρμοστεί το κριτήριο του Schur για τον υπολογισμό του διαστήματος απόλυτης ευστάθειας της γραμμικής μεθόδου:

$$y_{n+2}-y_n=\frac{h}{2}\left(f_{n+1}+3f_n\right)$$

Λύση: Τα χαρακτηριστικά πολυώνυμα της μεθόδου είναι:

$$\rho(r)=r^2-1 \quad \text{και} \quad \sigma(r)=\frac{1}{2}(r+3).$$

Άρα:

$$\pi(r,\bar{h})=r^2-\frac{1}{2}\bar{h}\,r-\left(1+\frac{3}{2}\bar{h}\right),$$

συνεπώς,

$$\hat{\pi}(r,\bar{h})=-\left(1+\frac{3}{2}\bar{h}\right)r^2-\frac{1}{2}\bar{h}\,r+1.$$

Εύκολα συνάγεται ότι $\left|\hat{\pi}(0,\bar{h})\right|>\left|\pi(0,\bar{h})\right|$, αν $\left|1+\frac{3}{2}\bar{h}\right|<1$, το οποίο ισοδυναμεί με τον περιορισμό $\bar{h}\in\left(-\frac{4}{3},0\right)$. Τώρα υπολογίζεται ότι το πολυώνυμο $\pi_1(r,\bar{h})$ είναι:

$$\pi_1(r,\bar{h})=-\frac{1}{2}\bar{h}\left(2+\frac{3}{2}\bar{h}\right)(3r+1),$$

το οποίο έχει την μοναδική ρίζα $r=-\frac{1}{3}$. Επομένως, είναι ένα Schur πολυώνυμο για κάθε τιμή του $\bar{h}$. Εύκολα συνάγουμε από το κριτήριο του Schur ότι το $\pi\left(r,\bar{h}\right)$ είναι ένα Schur πολυώνυμο, αν και μόνο αν $\bar{h}\in\left(-\frac{4}{3},0\right)$, συνεπώς το διάστημα απόλυτης ευστάθειας της μεθόδου είναι $\left(-\frac{4}{3},0\right)$.

b) Το κριτήριο Routh-Hurwitz

Το κριτήριο αυτό βασίζεται στον μετασχηματισμό $z=\frac{(r-1)}{(r+1)}$, ο οποίος απεικονίζει τον ανοικτό μοναδιαίο δίσκο $|r|<1$ του μιγαδικού $r-$επιπέδου, στο αριστερό ανοικτό ημιεπίπεδο $\operatorname{Re} z<0$ του μιγαδικού z-επιπέδου. Ο αντίστροφος μετασχηματισμός δίνεται από την σχέση:

$$r=\frac{1+z}{1-z} \tag{30.9}$$

Κάτω από αυτόν τον μετασχηματισμό, το πολυώνυμο ευστάθειας της γραμμικής πολυβηματικής μεθόδου $\pi\left(r,\bar{h}\right)=\rho(r)-\bar{h}\,\sigma(r)$, γράφεται:

$$\pi(z)=\rho\left(\frac{1+z}{1-z}\right)-\bar{h}\,\sigma\left(\frac{1+z}{1-z}\right) \tag{30.10}$$

Πολλαπλασιάζοντας με τον όρο $(1-z)^k$, προκύπτει ένα πολυώνυμο της μορφής

$$\ddot{Q}(z)=a_0 z^k+a_1 z^{k-1}+\ldots+a_k \tag{30.11}$$

Οι ρίζες του πολυωνύμου ευστάθειας $\pi\left(r,\bar{h}\right)$, βρίσκονται μέσα στον ανοικτό μοναδιαίο δίσκο $|r|<1$, αν και μόνο αν οι ρίζες του πολυωνύμου (30.11) βρίσκονται μέσα στο αριστερό ανοικτό ημιεπίπεδο $\operatorname{Re} z<0$. Μπορούμε να διατυπώσουμε το επόμενο θεώρημα.

Θεώρημα 30.3 Οι ρίζες του πολυωνύμου (30.11), βρίσκονται στο ανοικτό αριστερό ημιεπίπεδο του z-επιπέδου, δηλαδή οι ρίζες του πολυωνύμου $\pi\left(r,\bar{h}\right)$ να βρίσκονται μέσα στον μοναδιαίο δίσκο, αν και μόνο αν όλες οι ηγούμενες κύριες υποορίζουσες του πίνακα

$$A=\begin{pmatrix} a_1 & a_3 & a_5 & . & . & . & a_{2k-1} \\ a_0 & a_2 & a_4 & . & . & . & a_{2k-2} \\ 0 & a_1 & a_3 & . & . & . & a_{2k-3} \\ 0 & a_0 & a_2 & . & . & . & a_{2k-4} \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & . & a_k \end{pmatrix}$$

είναι θετικές με $a_0>0$, και όπου $a_j=0$ αν $j>k$.

Ειδικότερα, μπορεί να δειχθεί ότι αυτό συνεπάγεται ότι $a_j > 0$, $j = 0,1,2,...,k$. Όμως, οι θετικές τιμές των συντελεστών του πολυωνύμου (30.11) είναι μία αναγκαία αλλά όχι και ικανή συνθήκη για απόλυτη ευστάθεια.

Για $k = 2,3,4$, οι αναγκαίες και ικανές συνθήκες για απόλυτη ευστάθεια μιας αντίστοιχης πολυβηματικής μεθόδου, που προκύπτουν με βάση το κριτήριο Routh-Hurwitz είναι

$$k=2: \quad a_0 > 0,\ a_1 > 0,\ a_2 > 0$$

$$k=3: \quad a_0 > 0,\ a_1 > 0,\ a_2 > 0,\ a_3 > 0,\ a_1 a_2 - a_3 a_0 > 0$$

$$k=4: \quad a_0 > 0,\ a_1 > 0,\ a_2 > 0,\ a_3 > 0,\ a_4 > 0,\ a_1 a_2 a_3 - a_0 a_3^2 - a_4 a_1^2 > 0.$$

Άσκηση 30.1 Να εφαρμοστεί το κριτήριο Routh-Hurwitz για να υπολογιστεί το διάστημα απόλυτης ευστάθειας της μεθόδου του παραδείγματος 30.1

31 Μέθοδοι τύπου Adams

Στην παράγραφο αυτή θα ασχοληθούμε ιδιαίτερα με την κατασκευή μεθόδων τύπου Adams. Οι μέθοδοι Adams αποτελούν μία υποκατηγορία των γραμμικών πολυβηματικών μεθόδων και ορίζονται από τον τύπο:

$$y_{n+1} - y_n = h \sum_{j=0}^{k} \beta_j f_{n+j-k+1}. \tag{31.1}$$

Οι μέθοδοι του τύπου αυτού έχουν μία μακρά ιστορία, και οι άμεσες, Adams-Bashforth, έχουν αρχικά εισαχθεί από τους Bashforth και Adams το 1883, ενώ οι έμμεσες, Adams-Moulton εμφανίστηκαν για πρώτη φορά από τον Moulton το 1926. Σήμερα παραμένουν η πιο δημοφιλής οικογένεια γραμμικών πολυβηματικών μεθόδων, και αποτελούν την βάση όλων σχεδόν των μεθόδων Πρόβλεψης-Διόρθωσης (*Predictor-Corrector*), για **μη-δύσκαμπτα** (*non-stiff*) προβλήματα αρχικών τιμών. Οι λόγοι γι' αυτήν την προτίμηση είναι:

a) Σε σύγκριση με άλλες κατηγορίες πολυβηματικών μεθόδων, οι μέθοδοι τύπου Adams έχουν καλά χωρία απόλυτης ευστάθειας.

b) Οι μέθοδοι Adams έχουν ένα συγκεκριμένο πλεονέκτημα όταν το βήμα ολοκλήρωσης αλλάζει στην διάρκεια των υπολογισμών. Τότε συνήθως εμφανίζεται κάποιο πρόβλημα, διότι οι προηγούμενες τιμές δεν συνεχίζουν να αντιστοιχούν στις ενδεδειγμένες τιμές του x. Μία λύση στο πρόβλημα αυτό θα ήταν η χρήση της παρεμβολής για τον υπολογισμό των αναγκαίων προηγούμενων τιμών. Για τις γενικές γραμμικές πολυβηματικές μεθόδους, αυτό σημαίνει ότι κάνουμε παρεμβολή για τις υπάρχουσες προηγούμενες τιμές του y, και εν συνεχεία υπολογίζουμε τις τιμές της f. Για τις μεθόδους Adams, κατ' ευθείαν κάνουμε παρεμβολή στις προηγούμενες τιμές της f.

c) Οι μέθοδοι Adams μπορούν να εκφραστούν με προς τα πίσω διαφορές, σε μία μορφή η οποία διευκολύνει πάρα πολύ την χρήση τους σε αυτόματους κώδικες (automatic codes).

Η κατασκευή των μεθόδων αυτών μπορεί να βασιστεί στο πολυώνυμο παρεμβολής Newton με προς τα πίσω διαφορές. Στην περίπτωση των μεθόδων Adams Bashforth το πολυώνυμο παρεμβολής παρεμβάλλει τα δεδομένα:

$$\left(x_n, f_n\right),\ \left(x_{n-1}, f_{n-1}\right), \ldots, \left(x_{n-k+1}, f_{n-k+1}\right) \tag{31.2}$$

και στην περίπτωση των μεθόδων Adams-Moulton η παρεμβολή γίνεται στα δεδομένα:

$$\left(x_{n+1}, f_{n+1} \right),\ \left(x_n, f_n \right), \ldots, \left(x_{n-k+1}, f_{n-k+1} \right). \tag{31.3}$$

Για να διαχωρίζουμε τις άμεσες από τις έμμεσες μεθόδους, θα χρησιμοποιούμε τον χαρακτήρα, (*), σε όλα τα σύμβολα που σχετίζονται με τις άμεσες μεθόδους, και θα αρχίσουμε, υπολογίζοντας τις άμεσες μεθόδους Adams-Bashforth.

Μέθοδος Adams-Bashforth

Το πολυώνυμο παρεμβολής Newton με διαφορές ∇, που παρεμβάλλει τα δεδομένα (31.2), μπορεί να γραφεί στην μορφή:

$$P_{k-1}^*(x) = P_{k-1}^*(x_n + sh) := P_{k-1}^*(s) = \sum_{i=0}^{k-1} (-1)^i \binom{-s}{i} \nabla^i f_n, \tag{31.4}$$

όπου $s = (x - x_n)/h$. Τότε από την (25.4) προκύπτει η προσεγγιστική σχέση,

$$y(x_{n+1}) - y(x_n) \approx \int_0^1 P_{k-1}^*(s)\, h\, ds,$$

η οποία γίνεται ακριβής για τις προσεγγίσεις y_{n+1}, y_n. Τότε η μέθοδος η οποία προκύπτει μπορεί να γραφεί::

$$y_{n+1} - y_n = \int_0^1 P_{k-1}^*(s)\, h\, ds = h \sum_{i=0}^{k-1} \gamma_i^* \nabla^i f_n\ , \tag{31.5}$$

όπου

$$\gamma_i^* = (-1)^i \int_0^1 \binom{-s}{i}\, ds\ . \tag{31.6}$$

Οι συντελεστές γ_i^* , $i = 0,1,\ldots$ είναι ανεξάρτητοι του *k*. Ο υπολογισμός των συντελεστών αυτών μπορεί να γίνει και ως εξής.

Γεννήτρια συνάρτηση

Ζητάμε να υπολογίσουμε μία **γεννήτρια συνάρτηση** (generating function), για τους συντελεστές γ_i^*, δηλαδή, μία συνάρτηση μίας μεταβλητής έστω r , η οποία όταν αναλυθεί σε δυνάμεις του r , τα γ_i^* θα αποτελούν τους συντελεστές αυτού του αναπτύγματος. Συνεπώς, ζητάμε να υπολογίσουμε μία συνάρτηση $\Gamma^*(r)$ έτσι ώστε:

$$\Gamma^*(r) = \sum_{i=0}^{\infty} \gamma_i^* r^i = \sum_{i=0}^{\infty} (-r)^i \int_0^1 \binom{-s}{i} ds = \int_0^1 \left[\sum_{i=0}^{\infty} (-r)^i \binom{-s}{i} \right] ds\ .$$

Όμως,

$$\sum_{i=0}^{\infty} (-r)^i \binom{-s}{i} = (1-r)^{-s},$$

και ισχύει ότι:

$$\int (1-r)^{-s}\, ds = -(1-r)^{-s} / \ln(1-r).$$

Επομένως,

$$\Gamma^*(r) = \frac{-r}{(1-r)\ln(1-r)} ,$$

από την οποίαν προκύπτει ο τύπος:

$$\Gamma^*(r)\left[\frac{-\ln(1-r)}{r}\right] = \frac{1}{1-r} . \tag{31.7}$$

Είναι γνωστό ότι:

$$\frac{-\ln(1-r)}{r} = 1 + \frac{r}{2} + \frac{r^2}{3} + \frac{r^3}{4} + \dots$$

και

$$\frac{1}{1-r} = 1 + r + r^2 + r^3 + \dots$$

Από την (31.7) προκύπτει ότι τα γ_i^* μπορούν να υπολογιστούν από την σχέση:

$$\left(\gamma_0^* + \gamma_1^* r + \gamma_2^* r^2 + \gamma_3^* r^3 + \dots\right)\left(1 + \frac{r}{2} + \frac{r^2}{3} + \frac{r^3}{4} + \dots\right) = 1 + r + r^2 + r^3 + \dots \ .$$

Εξισώνοντας τους συντελεστές r^i, προκύπτει η σχέση

$$\gamma_i^* + \frac{\gamma_{i-1}^*}{2} + \frac{\gamma_{i-2}^*}{3} + \dots + \frac{\gamma_0^*}{i+1} = 1 \ , \quad i = 0,1,2,\dots \ , \tag{31.8}$$

από την οποίαν υπολογίζονται τα γ_i^* .

Βασικοί τύποι

Εύκολα προκύπτει ότι:

$$\gamma_0^* = 1, \quad \gamma_1^* = \tfrac{1}{2}, \quad \gamma_2^* = \tfrac{5}{12}, \quad \gamma_3^* = \tfrac{3}{8}.$$

Επομένως η οικογένεια των άμεσων γραμμικών μεθόδων τύπου Adams-Bashforth μπορεί να γραφεί:

$$y_{n+1} - y_n = h\left\{ f_n + \tfrac{1}{2}\nabla f_n + \tfrac{5}{12}\nabla^2 f_n + \tfrac{3}{8}\nabla^3 f_n + \dots \right\} . \tag{31.9}$$

Έτσι, για τις διάφορες τιμές του k και αναλύοντας τις διαφορές ∇^k σε τιμές της f η (31.9) δίνει τις παρακάτω άμεσες μεθόδους:

k = 1: $\quad y_{n+1} - y_n = h\, f_n$

k = 2: $\quad y_{n+1} - y_n = \dfrac{h}{2}\left(3 f_n - f_{n-1}\right)$

k = 3: $\quad y_{n+1} - y_n = \dfrac{h}{12}\left(23 f_n - 16 f_{n-1} + 5 f_{n-2}\right)$

k = 4: $y_{n+1} - y_n = \frac{h}{24}\left(55 f_n - 59 f_{n-1} + 37 f_{n-2} - 9 f_{n-3}\right).$

Τάξη και σταθερά σφάλματος

Έχουμε δείξει ότι η γενική μέθοδος Adams-Bashforth *k*-βημάτων, δίνεται από τον τύπο:

$$y_{n+1} - y_n = h \sum_{i=0}^{k-1} \gamma_i^* \nabla^i f_n .$$

Η διαφορά μεταξύ των τιμών της προσέγγισης y_{n+1} οι οποίες προκύπτουν από μεθόδους Adams-Bashforth, *(k+1)*-βημάτων, και, *k*-βημάτων, θα είναι ίση με:

$$h\gamma_k^* \nabla^k f_n = h^{k+1} \gamma_k^* \, y^{(k+1)}\left(x_n\right) + \mathrm{O}\left(h^{k+2}\right). \tag{31.10}$$

Από την (31.10) μπορούμε να συνάγουμε το συμπέρασμα, ότι η μέθοδος Adams-Bashforth *k*-βημάτων, έχει τάξη *k* και σταθερά σφάλματος γ_k^* .

Μέθοδος Adams-Moulton

Με ανάλογο τρόπο υπολογίζονται οι έμμεσες μέθοδοι Adams-Moulton οι οποίες μπορούν επίσης να εκφραστούν με προς τα πίσω διαφορές. Το πολυώνυμο παρεμβολής Newton με διαφορές ∇, που παρεμβάλλει τα δεδομένα (31.3), μπορεί να γραφεί στην μορφή:

$$P_k(x) = P_k\left(x_{n+1} + sh\right) := P_k(s) = \sum_{i=0}^{k} (-1)^i \binom{-s}{i} \nabla^i f_{n+1} .$$

Τότε προκύπτει η προσεγγιστική μέθοδος,

$$y_{n+1} - y_n = \int_{-1}^{0} P_k(s) h \, ds = h \sum_{i=0}^{k} \gamma_i \nabla^i f_{n+1} , \tag{31.11}$$

όπου

$$\gamma_i = (-1)^i \int_{-1}^{0} \binom{-s}{i} ds . \tag{31.12}$$

Γεννήτρια συνάρτηση

Η γεννήτρια συνάρτηση $\Gamma(r)$, για τα γ_i , ορίζεται από την:

$$\Gamma(r) = \sum_{i=0}^{\infty} \gamma_i r^i = \sum_{i=0}^{\infty} (-r)^i \int_{-1}^{0} \binom{-s}{i} ds = \int_{-1}^{0} \left[\sum_{i=0}^{\infty} (-r)^i \binom{-s}{i} \right] ds.$$

Τελικά προκύπτει ότι η γεννήτρια συνάρτηση είναι της μορφής:

$$\Gamma(r) = \frac{-r}{\ln(1-r)},$$

από την οποίαν προκύπτει ότι:

$$\left(\gamma_0 + \gamma_1 r + \gamma_2 r^2 + \gamma_3 r^3 + \ldots\right)\left(1 + \frac{r}{2} + \frac{r^2}{3} + \frac{r^3}{4} + \ldots\right) = 1,$$

οπότε:

$$\gamma_i + \frac{\gamma_{i-1}}{2} + \frac{\gamma_{i-2}}{3} + \ldots + \frac{\gamma_0}{i+1} = \begin{cases} 1 & \text{αν } i = 0 \\ 0 & \text{αν } i = 1,2,\ldots \end{cases}$$

Οι πρώτες τιμές για τα γ_i είναι:

$$\gamma_0 = 1,\ \gamma_1 = -\tfrac{1}{2},\ \gamma_2 = -\tfrac{1}{12},\ \gamma_3 = -\tfrac{1}{24},\ \gamma_4 = -\tfrac{19}{720}.$$

Συνεπώς η οικογένεια των μεθόδων Adams-Moulton μπορεί να γραφεί:

$$y_{n+1} - y_n = h\begin{Bmatrix} f_{n+1} - \frac{1}{2}\nabla f_{n+1} - \frac{1}{12}\nabla^2 f_{n+1} - \frac{1}{24}\nabla^3 f_{n+1} - \\ -\frac{19}{720}\nabla^4 f_{n+1} + \ldots \end{Bmatrix}. \qquad (31.13)$$

Βασικοί τύποι

Από την (31.13) για τις διάφορες τιμές του *k* προκύπτουν οι παρακάτω μέθοδοι:

k = 1: $y_{n+1} - y_n = \frac{h}{2}\left(f_{n+1} + f_n\right)$

k = 2: $y_{n+1} - y_n = \frac{h}{12}\left(5 f_{n+1} + 8 f_n - f_{n-1}\right)$

k = 3: $y_{n+1} - y_n = \frac{h}{24}\left(9 f_{n+1} + 19 f_n - 5 f_{n-1} + f_{n-2}\right)$

k = 4: $y_{n+1} - y_n = \frac{h}{720}\left(251 f_{n+1} + 646 f_n - 264 f_{n-1} + 106 f_{n-2} - 19 f_{n-3}\right).$

Η Adams-Moulton *k*-βημάτων αποδεικνύεται ότι έχει τάξη *k+1* και σταθερά σφάλματος γ_{k+1}.

31 Μέθοδος Πρόβλεψης-Διόρθωσης (Predictor-Corrector)

Γενικά, μία μέθοδος Adams-Moulton δίνει καλλίτερα αποτελέσματα από μία μέθοδο Adams-Bashforth της ίδιας τάξης. Όμως μία έμμεση μέθοδος έχει το μειονέκτημα ότι πρώτα πρέπει να μετατραπεί αλγεβρικά σε άμεση ως προς την άγνωστη προσέγγιση, πράγμα που γίνεται δύσκολα ή πολλές φορές είναι αδύνατο.

Στην πράξη, μία έμμεση μέθοδος δεν εφαρμόζεται μόνη της, αλλά συνήθως χρησιμοποιείται για να βελτιώσει την προσέγγιση που προκύπτει από την εφαρμογή μίας άμεσης μεθόδου. Ο συνδυασμός μίας άμεσης και μίας έμμεσης μεθόδου, δημιουργεί μία νέα κατηγορία πολυβηματικών μεθόδων γνωστή σαν μέθοδος Πρόβλεψης - Διόρθωσης (Predictor-Corrector).

Η εφαρμογή μίας μεθόδου Πρόβλεψης -Διόρθωσης μπορεί να γίνει με δύο κυρίως τρόπους.

Στην μία περίπτωση η άμεση μέθοδος εφαρμόζεται σε κάθε βήμα μία φορά και η έμμεση χρησιμοποιείται σαν επαναληπτική μέθοδος μέχρι να ικανοποιηθεί κάποιο κριτήριο της μορφής:

$$\left| y_{n+1}^{(s+1)} - y_{n+1}^{(s)} \right| < \varepsilon,$$

όπου s παριστάνει τον αριθμό των επαναλήψεων και ε είναι ένας προκαθορισμένος θετικός αριθμός, το μέγεθος του οποίου εξαρτάται από την ακρίβεια με την οποίαν θέλουμε να υπολογίσουμε την λύση. Τότε η προσέγγιση $y_{n+1}^{(s+1)}$, θεωρείται σαν μία παραδεκτή τιμή της λύσης y_{n+1}. Στην άλλη περίπτωση, εκ των προτέρων καθορίζουμε πόσες φορές θα εφαρμόσουμε την έμμεση μέθοδο σε κάθε βήμα. Έστω P δηλώνει μία εφαρμογή της άμεσης μεθόδου, C μία απλή εφαρμογή της έμμεσης μεθόδου και E δηλώνει έναν υπολογισμό της συνάρτησης f σε γνωστές τιμές των μεταβλητών της. Τότε ένα πολύ συνηθισμένο σχήμα μεθόδου Πρόβλεψης Διόρθωσης συμβολίζεται με PECE που σημαίνει ότι σε κάθε βήμα εφαρμόζουμε μία φορά την πρόβλεψη και μία φορά την διόρθωση. Το πιο γενικό σχήμα συμβολίζεται με $\mathrm{P(EC)^M E}$ που σημαίνει ότι σε κάθε βήμα εφαρμόζουμε μία φορά την μέθοδο της πρόβλεψης και M φορές την μέθοδο της διόρθωσης. Μία κλασσική και αρκετά ικανή μέθοδος Πρόβλεψης-Διόρθωσης είναι η Adams-Bashforth-Moulton 4ης τάξης στην μορφή PECE, βασίζεται στους τύπους Adams-Bashforth με k=4 και Adams-Moulton με k=3, και είναι της μορφής:

$$
\begin{aligned}
&\text{P:} \quad y_{n+1}^{(0)} = y_n + \frac{h}{24}\left(55 f_n - 59 f_{n-1} + 37 f_{n-2} - 9 f_{n-3}\right) \\
&\text{E:} \quad f_{n+1}^{(0)} = f\left(x_{n+1}, y_{n+1}^{(0)}\right) \\
&\text{C:} \quad y_{n+1}^{(1)} = y_n + \frac{h}{24}\left(9 f_{n+1}^{(0)} + 19 f_n - 5 f_{n-1} + f_{n-2}\right) \\
&\text{E:} \quad f_{n+1}^{(1)} = f\left(x_{n+1}, y_{n+1}^{(1)}\right).
\end{aligned}
\tag{32.1}
$$