

# Προβλήματα Αρχικών Τιμών

Έστω  $a, b \in \mathbb{R}$ ,  $a < b$ ,  $f: [a, b] \times \mathbb{R} \rightarrow \mathbb{R}$  μια συνάρτηση, και  $y_0 \in \mathbb{R}$ . Ένα τυπικό πρόβλημα αρχικών τιμών είναι:  
Αναζητείται  $y: [a, b] \rightarrow \mathbb{R}$ , τέτοια ώστε

$$\begin{aligned}y'(t) &= f(t, y(t)), & a \leq t \leq b, \\y(a) &= y_0.\end{aligned}\tag{1}$$

- ▶  $f$  συνεχής  $(t, y) \in [a, b] \times \mathbb{R}$ . ( $f \in C([a, b] \times \mathbb{R})$ .)
- ▶ Η συνάρτηση  $y$  ορισμένη στο  $[a, b]$ ,  $y \in C^1[a, b]$ , η οποία ικανοποιεί την (1) και την αρχική συνθήκη  $ay(a) = y_0$ , λέγεται λύση του ΠΑΤ (1).
- ▶ αν η  $f$  είναι  $m$  φορές συνεχώς παραγωγίσιμη στο  $[a, b] \times \mathbb{R}$ , τότε η λύση  $y$  είναι  $m + 1$  συνεχώς παραγωγίσιμη στο  $[a, b]$ .

# Χρόνος / Διακριτά χρονικά βήματα

Σκεφτόμαστε την μεταβλητή  $t$  ως χρόνο.

Οι αριθμητικοί αλγόριθμοι υπολογίζουν προσεγγίσεις σε **διακριτές χρονικές στιγμές**:

- ▶ Έστω

$$a = t^0 < t^1 < \dots < t^N = b$$

μια διάμεριση του  $[a, b]$ .

- ▶ Θεωρούμε τους διακριτούς χρόνους  $t^n$ .
- ▶ Αναζητούμε **προσεγγίσεις του  $y$  στους διακριτούς χρόνους  $t^n$** :

$$y(t^i), \quad i = 0, \dots, N.$$

- ▶ Συμβολισμός: Θα λέμε

$y^i$  την προσέγγιση του  $y$  η οποία προκύπτει από τον αριθμητικό μας αλγόρ

- ▶ Ή

$$y^i \approx y(t^i), \quad i = 0, \dots, N.$$

# Υψηλής τάξης μονο-βηματικές μέθοδοι

Οι μέθοδοι Runge–Kutta (RK) είναι *μονο-βηματικές μέθοδοι*, Δηλαδή, μέθοδοι στις οποίες ο υπολογισμός της προσέγγισης  $y^{n+1}$  χρησιμοποιεί την προσέγγιση στο προηγούμενο βήμα  $y^n$ .

Βασικές αρχές:

- ▶ προσεγγίσεις ενδιάμεσων σημείων του  $[t^n, t^{n+1}]$
- ▶ κανόνες ολοκλήρωσης υψηλής τάξης
- ▶ *συνδυασμός ενδιάμεσων προσεγγίσεων με τους κανόνες ολοκλήρωσης υψηλής τάξης για την κατασκευή της προσέγγισης  $y^{n+1}$  υψηλής τάξης ακρίβειας.*

# Runge-Kutta: κατασκευή / βασικά συστατικά

Έστω  $q \in \mathbb{N}$ ,  $\tau_i \in \mathbb{R}$ ,  $i = 1, \dots, q$  (συνήθως  $0 \leq \tau_i \leq 1$ ),  $a_{ij} \in \mathbb{R}$ ,  $i, j = 1, \dots, q$ , και  $b_i \in \mathbb{R}$ ,  $i = 1, \dots, q$ . Γιά  $\psi : \mathbb{R} \rightarrow \mathbb{R}$  προσεγγίζουμε τα ολοκληρώματα

$$\int_0^{\tau_i} \psi(s) ds$$

από αθροίσματα (κανόνες αριθμητικής ολοκλήρωσης)

$$\sum_{j=1}^q a_{ij} \psi(\tau_j), \quad i = 1, \dots, q,$$

και το ολοκλήρωμα

$$\int_0^1 \psi(s) ds$$

από τον κανόνα

$$\sum_{j=1}^q b_j \psi(\tau_j).$$

Δηλαδή, οι συντελεστές  $a_{ij}$ ,  $\tau_i$ ,  $b_i$  περιγράφουν  $q + 1$  κανόνες αριθμητικής ολοκλήρωσης. Σ' αυτούς τους κανόνες τα  $\tau_i$  είναι οι *κόμβοι*, και τα  $b_i$  είναι τα *βάρη* του κανόνα ολοκλήρωσης στο  $[0, 1]$ , και  $a_{ij}$ ,  $j = 1, \dots, q$ , είναι τα *βάρη* των κανόνων ολοκλήρωσης στα  $[0, \tau_j]$ .

# Runge-Kutta: κατασκευή / βασικά συστατικά II

Κάθε τέτοιο σύνολο σταθερών περιγράφει πλήρως μια μέθοδο Runge-Kutta (RK). Γράφουμε τις σταθερές αυτές στην μορφή πίνακα *Runge-Kutta tableau* (Συμβολισμός του J. Butcher)

$$\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1q} & \tau_1 \\ a_{21} & a_{22} & \dots & a_{2q} & \tau_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{q1} & a_{q2} & \dots & a_{qq} & \tau_q \\ \hline b_1 & b_2 & \dots & b_q & \end{array} = \frac{Q}{b^T} \Big| \tau, \quad (2)$$

με  $Q = (a_{ij}) \in \mathbb{R}^{q,q}$ ,  $b = (b_1, \dots, b_q)^T$  και  $\tau = (\tau_1, \dots, \tau_q)^T$ .

# Runge-Kutta: Ορισμός

## Συνοψίζοντας:

Η μέθοδος RK η οποία αντιστοιχεί στον πίνακα (tableau) (2) δίδει προσεγγίσεις  $y^0, \dots, y^N$ , οι οποίες υπολογίζονται διαδοχικά από τις σχέσεις

$$\begin{aligned}y^0 &:= y_0 \\ y^{n,i} &= y^n + h \sum_{j=1}^q a_{ij} f(t^{n,j}, y^{n,j}), \quad 1 \leq i \leq q, \\ &, \quad n = 0, \dots, N-1. \\ y^{n+1} &:= y^n + h \sum_{i=1}^q b_i f(t^{n,i}, y^{n,i})\end{aligned} \tag{3}$$

# Παράδειγμα

Η μέθοδος δύο σταδίων

$$\begin{array}{cc|c} \frac{1}{4} & \frac{1}{4} - \mu & \frac{1}{2} - \mu \\ \frac{1}{4} + \mu & \frac{1}{4} & \frac{1}{2} + \mu \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}, \quad \mu = \frac{\sqrt{3}}{6}, \quad (4)$$

έχει πολύ ενδιαφέρουσες ιδιότητες. Η αντίστοιχη μέθοδος RK είναι η μόνη μέθοδος RK δύο σταδίων με τάξη ακρίβειας 4. Όλες οι άλλες μέθοδοι RK δύο σταδίων έχουν τάξη ακρίβειας το πολύ τρία. Η μέθοδος αυτή λέγεται *RK μέθοδος Gauss–Legendre δύο σταδίων*.

- ▶ Τα  $\tau_i = \frac{1}{2} \pm \mu$  και  $b_i = 1/2$  της (4) είναι οι κόμβοι και τα βάρη αντιστοίχα του κανόνα αριθμητικής ολοκλήρωσης του Gauss στο διάστημα  $[0, 1]$  με συνάρτηση βάρους  $w(x) = 1$ .

# Παράδειγμα

Η κλασική μέθοδος *Runge–Kutta* είναι άμεση, έχει 4 στάδια και τάξη ακρίβειας 4. Περιγράφεται από το RK tableau

$$\begin{array}{cccc|c} 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 & 1 \\ \hline \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} & \end{array}$$

(5)



# Ανάλυση Μεθόδων Runge-Kutta

- ▶ Επιλυσιμότητα (οι προσεγγίσεις είναι καλώς ορισμένες)
- ▶ Ευστάθεια
- ▶ Συνέπεια
- ▶ Σύγκλιση

# Επιλυσιμότητα και Ευστάθεια Μεθόδων Runge-Kutta

- ▶ Υποθέτουμε ότι η  $f$  ικανοποιεί την ΟΣ-Lipschitz

$$|f(t, y_1) - f(t, y_2)| \leq L|y_1 - y_2|. \quad (6)$$

## Πρόταση 1

(Υπαρξη και μοναδικότητα των προσεγγίσεων .) Υποθέτουμε ότι η (6) ικανοποιείται, και έστω  $h < \frac{1}{\gamma}$ , όπου

$$\gamma := L \max_{1 \leq i \leq q} \sum_{j=1}^q |a_{ij}|.$$

Τότε το σύστημα (3) έχει μοναδική λύση  $y^{n,1}, \dots, y^{n,q}$ .

# Απόδειξη.

Θεωρούμε την απεικόνιση  $F: \mathbb{R}^q \rightarrow \mathbb{R}^q$ ,

$$F_i(x) := y^n + h \sum_{j=1}^q a_{ij} f(t^{n,j}, x_j), \quad 1 \leq i \leq q,$$

με  $x = (x_1, \dots, x_q)^T$  και  $F(x) = (F_1(x), \dots, F_q(x))^T$ .

Προφανώς, κάθε σταθερό σημείο της  $F$  είναι λύση του συστήματος (3) και, αντίστροφα, κάθε λύση του συστήματος (3) είναι σταθερό σημείο της  $F$ . Συνεπώς, αρκεί να δείξουμε ότι, για  $h < \frac{1}{\gamma}$ , η  $F$  έχει ακριβώς ένα σταθερό σημείο.

Χρησιμοποιώντας την (6) έχουμε, για  $x, \tilde{x} \in \mathbb{R}^q$  και  $i = 1, \dots, q$ ,

$$\begin{aligned} |F_i(x) - F_i(\tilde{x})| &= \left| h \sum_{j=1}^q a_{ij} [f(t^{n,j}, x_j) - f(t^{n,j}, \tilde{x}_j)] \right| \\ &\leq hL \sum_{j=1}^q |a_{ij}| |x_j - \tilde{x}_j|, \end{aligned}$$

άρα

$$|F_i(x) - F_i(\tilde{x})| \leq \gamma h \max_j |x_j - \tilde{x}_j|. \quad (7)$$

# Απόδειξη (συνέχεια)

Συμβολίζουμε με  $\|\cdot\|_\infty$  την νόρμα απείρου του  $\mathbb{R}^q$ . Η (7) συνεπάγεται

$$\|F(\mathbf{x}) - F(\tilde{\mathbf{x}})\|_\infty \leq \gamma h \|\mathbf{x} - \tilde{\mathbf{x}}\|_\infty. \quad (8)$$

Δεδομένου ότι έχουμε υποθέσει ότι  $\gamma h < 1$ , η απεικόνιση  $F$  είναι συστολή στο  $(\mathbb{R}^q, \|\cdot\|_\infty)$ , και έχει συνεπώς ακριβώς ένα σταθερό σημείο, και η απόδειξη είναι πλήρης. □

# Ευστάθεια (Βασική Ευστάθεια)

## Ορισμός

(Ευστάθεια Μεθόδων Runge-Kutta.) Θα λέμε ότι η μέθοδος Runge–Kutta είναι ευσταθής, αν

- ▶ για κάθε ΠΑΤ στο οποίο η  $f$  ικανοποιεί την ΟΣ- Lipschitz
- ▶ και δυο ακολουθίες προσεγγίσεων  $y^0, \dots, y^N, z^0, \dots, z^N$ , οι οποίες παράγονται από την ίδια μέθοδο Runge Kutta αλλά, γενικά, με διαφορετικές αρχικές τιμές  $y^0$  και  $z^0$ ,

ισχύει

$$\max_{1 \leq n \leq N} |y^n - z^n| \leq C|y^0 - z^0|. \quad (9)$$

# Οι Μέθοδοι Runge-Kutta είναι ευσταθείς

Θα δείξουμε ένα πιο γενικό αποτέλεσμα το οποίο θα χρησιμοποιηθεί και στην απόδειξη σύγκλισης. Υποθέτουμε ότι τα  $z^n$  ικανοποιούν τροποποιημένες εξισώσεις της μορφής: Για δεδομένο  $z^0 \in \mathbb{R}$ , τα  $z^n$  ικανοποιούν

$$z^{n,i} = z^n + h \sum_{j=1}^q a_{ij} f(t^{n,j}, z^{n,j}), \quad i = 1, \dots, q, \quad (10)$$

$$z^{n+1} = z^n + h \sum_{i=1}^q b_i f(t^{n,i}, z^{n,i}) + \rho^n, \quad \rho^n \text{ δεδομένοι αριθμοί.}$$

## Πρόταση 2

- ▶ Υποθέτουμε ότι η  $f$  ικανοποιεί την ΟΣ- Lipschitz.
- ▶ Έστω  $y^0, \dots, y^N$ , οι προσεγγίσεις που παράγονται από την μέθοδο RK και τα  $z^0, \dots, z^N$ , ορίζονται παραπάνω

Τότε, υπάρχουν σταθερές  $C_1$  και  $C_2$ , ανεξάρτητες του  $h$ , έτσι ώστε

$$\max_{1 \leq n \leq N} |y^n - z^n| \leq C_1 |y^0 - z^0| + \frac{C_2}{h} \max_{0 \leq n \leq N-1} |\rho^n|. \quad (11)$$

# Απόδειξη

Αφαιρώντας κατά μέλη τις σχέσεις που ορίζουν τα  $y^{n,i}$  και  $z^{n,i}$  και χρησιμοποιώντας την συνθήκη Lipschitz παίρνουμε, για κάθε  $i = 1, \dots, q$ ,

$$|y^{n,i} - z^{n,i}| \leq |y^n - z^n| + hL \sum_{j=1}^q |a_{ij}| |y^{n,j} - z^{n,j}|,$$

συνεπώς

$$|y^{n,i} - z^{n,i}| \leq |y^n - z^n| + h\gamma \max_j |y^{n,j} - z^{n,j}|.$$

Άρα,

$$\max_i |y^{n,i} - z^{n,i}| \leq |y^n - z^n| + h\gamma \max_j |y^{n,j} - z^{n,j}|.$$

Δεδομένου ότι έχουμε υποθέσει  $\gamma h < 1$ , η τελευταία σχέση δίνει

$$|y^{n,i} - z^{n,i}| \leq \frac{1}{1 - \gamma h} |y^n - z^n|, \quad i = 1, \dots, q, \quad n = 0, \dots, N-1.$$

Συνεπώς, αν  $h \leq h_0 < 1/\gamma$ , υπάρχει σταθερά  $C$ , ανεξάρτητη του  $h$ , τέτοια ώστε

$$|y^{n,i} - z^{n,i}| \leq C |y^n - z^n|, \quad i = 1, \dots, q, \quad n = 0, \dots, N-1. \quad (12)$$

# Απόδειξη (συνέχεια)

Αφαιρώντας τις τελευταίες σχέσεις των (3) και (10), και χρησιμοποιώντας πάλι την συνθήκη Lipschitz, παίρνουμε

$$|y^{n+1} - z^{n+1}| \leq |y^n - z^n| + hL \sum_{i=1}^q |b_i| |y^{n,i} - z^{n,i}| + |\rho^n|,$$

η οποία μαζί με την (12) δίνει, για  $n = 0, \dots, N-1$ ,

$$|y^{n+1} - z^{n+1}| \leq \left(1 + hLC \sum_{i=1}^q |b_i|\right) |y^n - z^n| + |\rho^n|. \quad (13)$$

Χρησιμοποιώντας το Βασικό Λήμμα Ευστάθειας (βλ slides: Stability Euler) και θέτοντας  $C' := LC \sum_{i=1}^q |b_i|$  λαμβάνουμε

$$\max_{1 \leq n \leq N} |y^n - z^n| \leq e^{C'(b-a)} |y^0 - z^0| + \frac{e^{C'(b-a)} - 1}{C'h} \max_{0 \leq n \leq N-1} |\rho^n|, \quad (14)$$

η οποία είναι η (11). □



# Βασική Ευστάθεια Μεθόδων RK

- ▶ Αν θέσουμε  $\rho^n = 0$  το παραπάνω αποτέλεσμα δίδει

$$\max_{1 \leq n \leq N} |y^n - z^n| \leq C_1 |y^0 - z^0|, \quad (15)$$

δηλαδή, υπό τις προϋποθέσεις της Πρότασης η Μέθοδος Runge-Kutta είναι ευσταθής.

- ▶ Εντελώς αντίστοιχα αποτελέσματα ισχύουν και για *συστήματα* ΣΔΕ.

# Τάξη ακρίβειας μεθόδων RK

- ▶ Σφάλμα συνέπειας
- ▶ Σύγκλιση

# Σφάλμα συνέπειας

Υποθέτουμε ότι η λύση της ΣΔΕ  $y$ ,  $y' = f(t, y)$ , είναι ομαλή συνάρτηση. Ορίζουμε, για  $n = 0, \dots, N-1$ , το *Σφάλμα Συνέπειας*  $E^n \in \mathbb{R}$  ως εξής:

$$\zeta^{n,i} = y(t^n) + h \sum_{j=1}^q a_{ij} f(t^{n,j}, \zeta^{n,j}), \quad 1 \leq i \leq q, \quad (16)$$
$$E^n := \left[ y(t^n) + h \sum_{i=1}^q b_i f(t^{n,i}, \zeta^{n,i}) \right] - y(t^{n+1})$$

- ▶ Η ποσότητα  $E^n$  λέγεται *σφάλμα συνέπειας*.
- ▶ Ερμηνεία του ορισμού: Ξεκινώντας από την τιμή της ακριβούς λύσης  $y(t^n)$ , αντί για την προσέγγιση  $y^n$ , και εκτελώντας ένα βήμα της μεθόδου παίρνουμε στο  $t^{n+1}$   $\left[ y(t^n) + h \sum_{i=1}^q b_i f(t^{n,i}, \zeta^{n,i}) \right]$  (βλ. (17))
- ▶  $E^n$  είναι η διαφορά ανάμεσα σ' αυτήν την τιμή και την τιμή της ακριβούς λύσης  $y(t^{n+1})$  στο  $t^{n+1}$ .
- ▶ Οι ποσότητες  $\zeta^{n,i}$ ,  $i = 1, \dots, q$ , άρα και το  $E^n$  είναι καλώς ορισμένες με βάση την Πρόταση 1.

# Σφάλμα συνέπειας II

Ισοδύναμος τρόπος να γραφεί το Σφάλμα Συνέπειας

$$\begin{aligned}\zeta^{n,i} &= y(t^n) + h \sum_{j=1}^q a_{ij} f(t^{n,j}, \zeta^{n,j}), & 1 \leq i \leq q, \\ \zeta^{n+1} &= y(t^n) + h \sum_{i=1}^q b_i f(t^{n,i}, \zeta^{n,i}) \\ E^n &= \zeta^{n+1} - y(t^{n+1})\end{aligned}\tag{17}$$

- ▶ Ξεκινώντας από την τιμή της ακριβούς λύσης  $y(t^n)$ , αντί για την προσέγγιση  $y^n$ , και εκτελώντας ένα βήμα της μεθόδου παίρνουμε στο χρονικό βήμα  $t^{n+1}$ , το  $\zeta^{n+1}$ . Αυτή η ποσότητα είναι το αποτέλεσμα της εφαρμογής του αλγορίθμου (output) της μεθόδου RK με είσοδο (input)  $y(t^n)$ .
- ▶ το  $E^n$  μετρά την "απόσταση" αυτού του αποτελέσματος (output:  $\zeta^{n+1}$ ) από την ακριβή τιμή  $y(t^{n+1})$
- ▶ το  $E^n$  ικανοποιεί

$$y(t^{n+1}) = y(t^n) + h \sum_{i=1}^q b_i f(t^{n,i}, \zeta^{n,i}) - E^n$$

# Τάξη ακρίβειας μεθόδων Runge–Kutta

Λέμε ότι η *τάξη ακρίβειας* της μεθόδου Runge–Kutta είναι  $p$ , αν  $p$  είναι ο μεγαλύτερος ακέραιος για τον οποίο υπάρχει σταθερά  $\tilde{C}$ , ανεξάρτητη του  $h$  και του  $N$ , (η οποία εξαρτάται όμως από τα δεδομένα του προβλήματος και τις παραμέτρους της μεθόδου RK) έτσι ώστε

$$\max_{0 \leq n \leq N-1} |E^n| \leq \tilde{C}h^{p+1}. \quad (18)$$

- ▶ Σημειώστε ότι απαιτείται να ισχύει η (18) για *κάθε* ΠΑΤ με αρκετά ομαλές συναρτήσεις  $f$  και  $y$ .

# Σύγκλιση

## Θεώρημα

Υποθέτουμε ότι η ΟΣ-Lipschitz ικανοποιείται και το ΠΑΤ έχει μια αρκετά ομαλή λύση. Θέτουμε  $\gamma := L \max_{1 \leq i \leq q} \sum_{j=1}^q |a_{ij}|$ , και έστω ότι το  $h_0 > 0$  ικανοποιεί

$$\gamma h_0 < 1, \quad 0 < h \leq h_0.$$

Αν η μέθοδος RK ικανοποιεί την εκτίμηση συνέπειας (18), τότε έχουμε

$$\max_{0 \leq n \leq N} |y(t^n) - y^n| \leq \frac{\tilde{C}}{C'} [e^{C'(b-a)} - 1] h^p, \quad (19)$$

όπου οι σταθερές  $\tilde{C}$  και  $C, C'$ , είναι ανεξάρτητες του  $h$  (οι  $C', \tilde{C}$  έχουν οριστεί παραπάνω.)

# Σύγκλιση II

**Απόδειξη.** Από τον ορισμό του σφάλματος συνέπειας,

$$\zeta^{n,i} = y(t^n) + h \sum_{j=1}^q a_{ij} f(t^{n,j}, \zeta^{n,j}), \quad 1 \leq i \leq q,$$

$$y(t^{n+1}) = \left[ y(t^n) + h \sum_{i=1}^q b_i f(t^{n,i}, \zeta^{n,i}) \right] - E^n.$$

Σύμφωνα με το αποτέλεσμα ευστάθειας (Πρόταση 2), χρησιμοποιώντας το γεγονός ότι  $y^0 = y_0 = y(t^0)$ , προκύπτει ότι,

$$\max_{0 \leq n \leq N} |y(t^n) - y^n| \leq C_2 h^{-1} \max_{0 \leq n \leq N-1} |E^n|,$$

Εφαρμόζοντας την (18) προκύπτει το επιθυμητό αποτέλεσμα.