Komputerowa analiza zagadnień różniczkowych

5. Metody Rungego-Kutty

P. F. Góra

http://th-www.if.uj.edu.pl/zfs/gora/

2013

Definicja metody

Poszukujemy rozwiązania problemu zmienności pochodnej (prawej strony standardowego problemu Cauchy'ego) w trakcie wykonywania kroku całkowania. Postulujemy, że należy używać pewnej *średniej pochodnej* — średniej ważonej wyliczonej z pochodnych w pewnych punktach pośrednich:

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h\sum_{i=1}^{s} w_i\mathbf{k}_i + O(h^{p+1})\,, \tag{1a}$$

$$\mathbf{k}_i = \mathbf{f}\left(t_n + h\alpha_i, \mathbf{y}_n + h\sum_{j=1}^{s}\beta_{ij}\mathbf{k}_j\right). \tag{1b}$$

Równania (1) definiują *s-etapową metodę Rungego-Kutty rzędu p*. Liczby w_i, α_i, β_{ij} są parametrami metody. $\mathbf{y}_l, \mathbf{k}_i \in \mathbb{R}^N$, $\mathbf{f} : \mathbb{R}\times\mathbb{R}^N \to \mathbb{R}^N$.

Zgodność metod RK

Najważniejszym kryterium, jakie musi spełniać każda metoda numerycznego całkowania równań różniczkowych, jest zgodność metody — wymaganie, aby metoda odtwrzała wyjściowy problem Cauchy'ego w granicy nieskończenie małych kroków. Dla metod RK z (1a) mamy

$$\lim_{h\to 0^+} \frac{\mathbf{y}_{n+1} - \mathbf{y}_n}{h} = \sum_{i=1}^{s} w_i \lim_{h\to 0^+} \mathbf{k}_i \,. \tag{2}$$

Granicą lewej strony (2) jest oczywiście $\left.\frac{d\mathbf{y}}{dt}\right|_n = \mathbf{f}(t_n, \mathbf{y}_n)$. Z (1b) wynika, że $\lim\limits_{h\to 0^+} \mathbf{k}_i = \mathbf{f}(t_n, \mathbf{y}_n)$, a zatem warunkiem zgodności metod RK jest

$$\sum_{i=1}^{s} w_i = 1 \,. \tag{3}$$

Zapis metod RK w postaci tabeli

Współczynniki metody RK zapisuje się na ogół w postaci tabeli

$$\begin{array}{c|ccccc} \alpha_1 & \beta_{11} & \beta_{12} & \dots & \beta_{1s} \\ \alpha_2 & \beta_{21} & \beta_{22} & \dots & \beta_{2s} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_s & \beta_{s1} & \beta_{s2} & \dots & \beta_{ss} \\ \hline & w_1 & w_2 & \dots & w_s \end{array} = \begin{array}{c|c} \boldsymbol{\alpha} & \mathbf{B} \\ \hline & \mathbf{w}^T \end{array} . \tag{4}$$

Jawne i niejawne metody RK

Jeżeli macierz B ma niezerowe elementy na i powyżej diagonali, metoda jest *niejawna*: do obliczenia pochodnej w punkcie pośrednim trzeba znać tęże pochodną (diagonala) i ewentualnie pochodne w kolejnych punktach pośrednich (wyrazy ponad diagonalą), tak więc proces obliczania wektorów k_i staje się rozwiązywaniem skomplikowanego (na ogół nieliniowego) układu równań na te wielkości. Wymiar tego układu wynosi $s \times N$.

Jeżeli macierz B ma niezerowe elementy tylko pod diagonalą, metoda jest jawna: do obliczenia kolejnych $\mathbf{k}_i$ potrzebna jest wyłącznie znajomość $\mathbf{k}_{j<i}$.

Związek pomiędzy liczbą etapów (s) a rzędem metody (p)

Na pierwszy rzut oka może się wydawać, że każde dodatkowe obliczenie prawej strony równania, czyli każde obliczenie pochodnej w punkcie pośrednim, zwiększa rząd metody o jeden. W rzeczywistości tak nie jest. Dla metod *jawnych* zawsze zachodzi $p \leqslant s$; minimalną liczbę etapów potrzebych do zrealizowania metody jawnej danego rzędu podaje poniższa tabela:

Rząd metody	1	2	3	4	5	6	7	8
Minimalna liczba etapów	1	2	3	4	6	7	9	11

Podana wyżej zależność uzasadnia szczególną popularność jawnych metod czteroetapowych — aby uzyskać metodę rzędu wyższego, niż czwarty, trzeba wyliczyć więcej punktów pośrednich, niż wynosi rząd metody.

Przykłady metod RK

Metody Eulera, jawna i niejawna, są, formalnie rzecz biorąc, jednoetapowymi metodami RK rzędu 1. Ich tabelki mają postać

$$
\begin{array}{c|c} 0 & 0 \\ \hline & 1 \end{array}
\qquad\qquad
\begin{array}{c|c} 1 & 1 \\ \hline & 1 \end{array}
\tag{5}
$$

(jawna) (niejawna)

Jawna metoda punktu środkowego jest dwuetapową metodą rzędu drugiego

$$
\begin{aligned}
\mathbf{k}_1 &= \mathbf{f}(x_n, \mathbf{y}_n)\,, && \text{(6a)}\\
\mathbf{k}_2 &= \mathbf{f}\left(x_n + \tfrac{1}{2}h, \mathbf{y}_n + \tfrac{1}{2}h\mathbf{k}_1\right)\,, && \text{(6b)}\\
\mathbf{y}_{n+1} &= \mathbf{y}_n + h\,\mathbf{k}_2 + O(h^3) && \text{(6c)}
\end{aligned}
$$

$$
\begin{array}{c|cc}
0 & 0 & 0 \\
\frac{1}{2} & \frac{1}{2} & 0 \\
\hline
 & 0 & 1
\end{array}
$$

Niejawna metoda punktu środkowego jest jednoetapową metodą rzędu drugiego

$$\mathbf{k}_1 = \mathbf{f}\left(t_n + \tfrac{1}{2}h, \mathbf{y}_n + \tfrac{1}{2}h\mathbf{k}_1\right), \tag{7a}$$

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h\mathbf{k}_1 + O(h^3). \tag{7b}$$

Tabelka dla tej metody ma postać

$$\begin{array}{c|c} \frac{1}{2} & \frac{1}{2} \\ \hline & 1 \end{array}$$

Klasyczna metoda czteroetapowa

Najczęściej stosowaną metodą RK jest *klasyczna metoda czteroetapowa* — większość ludzi uważa, że termin "metoda Rungego-Kutty" określa tylko tę metodę ☺

$$
\begin{aligned}
\mathbf{k}_1 &= \mathbf{f}\left(t_n, \mathbf{y}_n\right), && \text{(8a)} \\
\mathbf{k}_2 &= \mathbf{f}\left(t_n + \tfrac{1}{2}h, \mathbf{y}_n + \tfrac{1}{2}h\mathbf{k}_1\right), && \text{(8b)} \\
\mathbf{k}_3 &= \mathbf{f}\left(t_n + \tfrac{1}{2}h, \mathbf{y}_n + \tfrac{1}{2}h\mathbf{k}_2\right), && \text{(8c)} \\
\mathbf{k}_4 &= \mathbf{f}\left(t_n + h, \mathbf{y}_n + h\mathbf{k}_3\right), && \text{(8d)} \\
\mathbf{y}_{n+1} &= \mathbf{y}_n + h\left(\tfrac{1}{6}\mathbf{k}_1 + \tfrac{1}{3}\mathbf{k}_2 + \tfrac{1}{3}\mathbf{k}_3 + \tfrac{1}{6}\mathbf{k}_4\right) + O\left(h^5\right). && \text{(8e)}
\end{aligned}
$$

Tabelka dla klasycznej metody czteroetapowej

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

Zasady wyprowadzania współczynników metod RK

Współczynniki metod RK rzędu p otrzymuje się rozwijając obie strony wyrażenia

$$y_{n+1} - y_n = h \sum_i w_i k_i \tag{9}$$

w szereg potęgowy względem h z dokładnością do wyrazów rzędu $O(h^{p+1})$ i żądając równości współczynników rozwinięcia. Ponieważ pochodne w punktach pośrednich, k_i, są wyrażone przez funkcję f (prawą stronę równania), żądamy, aby funkcja f była klasy co najmniej C^{p-1}.

Te zasady są słuszne dla równań skalarnych. Okazuje się, że *niekiedy*, dla układów równań otrzymuje się inne wyrażenia na rząd metody, niż w przypadku skalarnym. W niniejszym wykładzie pominiemy to zastrzeżenie.

Lewa strona*

Rozwijając lewą stronę wyrażenia (9) otrzymujemy

$$y_{n+1} - y_n = \sum_{l=1}^{p} \frac{1}{l!} \left.\frac{d^l y}{dt^l}\right|_{t_n} h^l + O(h^{p+1})\,. \tag{10}$$

Mamy teraz

$$\frac{dy}{dt} = f(t,y) \tag{11a}$$

$$\frac{d^2y}{dt^2} = \frac{d}{dt} f(t,y) = \frac{\partial f}{\partial t} + \frac{dy}{dt}\frac{\partial f}{\partial y} = \frac{\partial f}{\partial t} + f\frac{\partial f}{\partial y}$$

$$= \left(\frac{\partial}{\partial t} + f\frac{\partial}{\partial y}\right) f(t,y) \tag{11b}$$

$$
\begin{aligned}
\frac{d^3y}{dt^3} &= \left(\frac{\partial}{\partial t} + f\frac{\partial}{\partial y}\right)\left(\frac{\partial}{\partial t} + f\frac{\partial}{\partial y}\right) f(t,y) \\
&= \frac{\partial^2 f}{\partial t^2} + 2f\frac{\partial^2 f}{\partial t \partial y} + f^2\frac{\partial^2 f}{\partial y^2} + \frac{\partial f}{\partial t}\frac{\partial f}{\partial y} + f\left(\frac{\partial f}{\partial y}\right)^2 \qquad \text{(11c)} \\
\frac{d^4y}{dt^4} &= \left(\frac{\partial}{\partial t} + f\frac{\partial}{\partial y}\right)\left(\frac{\partial}{\partial t} + f\frac{\partial}{\partial y}\right)\left(\frac{\partial}{\partial t} + f\frac{\partial}{\partial y}\right) f(t,y) \\
&= \frac{\partial^3 f}{\partial t^3} + 3f\frac{\partial^3 f}{\partial t^2 \partial y} + 3f^2\frac{\partial^3 f}{\partial t \partial y^2} + f^3\frac{\partial^3 f}{\partial y^3} \\
&+ \frac{\partial f}{\partial y}\frac{\partial^2 f}{\partial t^2} + 3\frac{\partial f}{\partial t}\frac{\partial^2 f}{\partial t \partial y} + 5f\frac{\partial f}{\partial y}\frac{\partial^2 f}{\partial t \partial y} + 3f\frac{\partial f}{\partial t}\frac{\partial^2 f}{\partial y^2} + 4f^2\frac{\partial f}{\partial y}\frac{\partial^2 f}{\partial y^2} \\
&+ \frac{\partial f}{\partial t}\left(\frac{\partial f}{\partial y}\right)^2 + f\left(\frac{\partial f}{\partial y}\right)^3 . \qquad \text{(11d)}
\end{aligned}
$$

Prawa strona*

$$
\begin{aligned}
h\sum_i w_i k_i &= h\sum_i w_i \left(\sum_{l=0}^{p-1} \frac{1}{l!} \left.\frac{d^l k_i}{dh^l}\right|_0 h^l + O(h^p)\right) \\
&= \sum_{l=0}^{p-1} \frac{1}{l!} \left(\sum_i w_i \left.\frac{d^l k_i}{dh^l}\right|_0\right) h^{l+1} + O(h^{p+1})\,, \qquad (12)
\end{aligned}
$$

a zatem musimy obliczyć pochodne k_i tylko do rzędu $p-1$, nie zaś p. Różniczkując (1b) i korzystając z twierdzenia o funkcjach uwikłanych otrzymujemy

$$
\frac{dk_i}{dh} = \alpha_i \frac{\partial f}{\partial t} + \frac{\partial f}{\partial y}\left(\sum_j \beta_{ij} k_j + h\sum_j \beta_{ij}\frac{dk_j}{dh}\right)\,. \qquad (13a)
$$

(13a) nie jest jawnym wyrażeniem na pochodne dk_i/dh, ale układem równań, z którego pochodne te można wyliczyć. W wyrażeniu tym wszystkie k_j oraz pochodne f określone są w punkcie $\left(t_n + \alpha_l h, y_n + h \sum\limits_m \beta_{lm} k_m\right)$, gdzie indeks l jest równy, odpowiednio j, i. Postępując analogicznie otrzymujemy

$$
\begin{aligned}
\frac{d^2 k_i}{dh^2} &= \alpha_i \frac{d}{dh}\left(\frac{\partial f}{\partial t}\right) + \left[\frac{d}{dh}\left(\frac{\partial f}{\partial y}\right)\right]\left(\sum_j \beta_{ij} k_j + h \sum_j \beta_{ij} \frac{dk_j}{dh}\right) \\
&+ \frac{\partial f}{\partial y}\frac{d}{dh}\left(\sum_j \beta_{ij} k_j + h \sum_j \beta_{ij} \frac{dk_j}{dh}\right) \\
&= \alpha_i^2 \frac{\partial^2 f}{\partial t^2} + 2\alpha_i \frac{\partial^2 f}{\partial t \partial y}\left(\sum_j \beta_{ij} k_j + h \sum_j \beta_{ij} \frac{dk_j}{dh}\right) \\
&+ \frac{\partial^2 f}{\partial y^2}\left(\sum_j \beta_{ij} k_j + h \sum_j \beta_{ij} \frac{dk_j}{dh}\right)^2 \\
&+ \frac{\partial f}{\partial y}\left(2 \sum_j \beta_{ij} \frac{dk_j}{dh} + h \sum_j \beta_{ij} \frac{d^2 k_j}{dh^2}\right) \qquad (13b)
\end{aligned}
$$

$$
\begin{aligned}
\frac{d^3k_i}{dh^3} &= \alpha_i^2\frac{d}{dh}\left(\frac{\partial^2 f}{\partial t^2}\right) + 2\alpha_i\left[\frac{d}{dh}\left(\frac{\partial^2 f}{\partial t\partial y}\right)\right]\left(\sum_j \beta_{ij}k_j + h\sum_j \beta_{ij}\frac{dk_j}{dh}\right) \\
&+ 2\alpha_i\frac{\partial^2 f}{\partial t\partial y}\frac{d}{dh}\left(\sum_j \beta_{ij}k_j + h\sum_j \beta_{ij}\frac{dk_j}{dh}\right) \\
&+ \left[\frac{d}{dh}\left(\frac{\partial^2 f}{\partial y^2}\right)\right]\left(\sum_j \beta_{ij}k_j + h\sum_j \beta_{ij}\frac{dk_j}{dh}\right)^2 + \frac{\partial^2 f}{\partial y^2}\frac{d}{dh}\left(\sum_j \beta_{ij}k_j + h\sum_j \beta_{ij}\frac{dk_j}{dh}\right)^2 \\
&+ \left[\frac{d}{dh}\left(\frac{\partial f}{\partial y}\right)\right]\left(2\sum_j \beta_{ij}\frac{dk_j}{dh} + h\sum_j \beta_{ij}\frac{d^2k_j}{dh^2}\right) \\
&+ \frac{\partial f}{\partial y}\frac{d}{dh}\left(2\sum_j \beta_{ij}\frac{dk_j}{dh} + h\sum_j \beta_{ij}\frac{d^2k_j}{dh^2}\right)
\end{aligned}
$$

$$
\begin{aligned}
= \;& \alpha_i^3 \frac{\partial^3 f}{\partial t^3} + 3\alpha_i^2 \frac{\partial^3 f}{\partial t^2 \partial y} \left(\sum_j \beta_{ij} k_j + h \sum_j \beta_{ij} \frac{dk_j}{dh} \right) \\
+ \;& 3\alpha_i \frac{\partial^3 f}{\partial t \partial y^2} \left(\sum_j \beta_{ij} k_j + h \sum_j \beta_{ij} \frac{dk_j}{dh} \right)^2 + \frac{\partial^3 f}{\partial y^3} \left(\sum_j \beta_{ij} k_j + h \sum_j \beta_{ij} \frac{dk_j}{dh} \right)^3 \\
+ \;& 3\alpha_i \frac{\partial^2 f}{\partial t \partial y} \left(2 \sum_j \beta_{ij} \frac{dk_j}{dh} + h \sum_j \beta_{ij} \frac{d^2 k_j}{dh^2} \right) \\
+ \;& 3 \frac{\partial^2 f}{\partial y^2} \left(\sum_j \beta_{ij} k_j + h \sum_j \beta_{ij} \frac{dk_j}{dh} \right) \left(2 \sum_j \beta_{ij} \frac{dk_j}{dh} + h \sum_j \beta_{ij} \frac{d^2 k_j}{dh^2} \right) \\
+ \;& \frac{\partial f}{\partial y} \left(3 \sum_j \beta_{ij} \frac{d^2 k_j}{dh^2} + h \sum_j \beta_{ij} \frac{d^3 k_j}{dh^3} \right)
\end{aligned}
\tag{13c}
$$

Jak już powiedziano, powyższe wzory nie dają jawnych wyrażeń na pochodne k_i, a tylko układ równań, z których pochodne te możnaby wyliczyć. Jednak dla $h = 0$ sytuacja znacznie się upraszcza: położenie $h = 0$ usuwa z prawej strony człony najwyższego rzędu, wobec czego otrzymujemy jawne wyrażenia na pochodne w zerze. I tak

$$
\begin{aligned}
k_i(0) &= f\,, &\text{(14a)}\\
\left.\frac{dk_i}{dh}\right|_0 &= \alpha_i \frac{\partial f}{\partial t} + \sum_j \beta_{ij} f \frac{\partial f}{\partial y}\,, &\text{(14b)}\\
\left.\frac{d^2k_i}{dh^2}\right|_0 &= \alpha_i^2 \frac{\partial^2 f}{\partial t^2} + 2\alpha_i \sum_j \beta_{ij}\, f \frac{\partial^2 f}{\partial t \partial y} + \left(\sum_j \beta_{ij}\right)^2 f^2 \frac{\partial^2 f}{\partial y^2} \\
&+ 2\sum_j \beta_{ij}\alpha_j\, \frac{\partial f}{\partial t}\frac{\partial f}{\partial y} + 2\sum_j \sum_q \beta_{ij}\beta_{jq}\, f \left(\frac{\partial f}{\partial y}\right)^2\,, &\text{(14c)}
\end{aligned}
$$

$$
\begin{aligned}
\left.\frac{d^3k_i}{dh^3}\right|_0 &= \alpha_i^3\,\frac{\partial^3 f}{\partial t^3} + 3\alpha_i^2\sum_j\beta_{ij}\,f\frac{\partial^3 f}{\partial t^2\partial y} + 3\alpha_i\left(\sum_j\beta_{ij}\right)^2 f^2\frac{\partial^3 f}{\partial t\partial y^2} + \left(\sum_j\beta_{ij}\right)^3 f^3\frac{\partial^3 f}{\partial y^3} \\
&+ 3\sum_j\beta_{ij}\alpha_j^2\,\frac{\partial f}{\partial y}\frac{\partial^2 f}{\partial t^2} + 6\alpha_i\sum_j\beta_{ij}\alpha_j\,\frac{\partial f}{\partial t}\frac{\partial^2 f}{\partial t\partial y} \\
&+ 6\sum_{j,q}\beta_{ij}(\alpha_i+\alpha_j)\beta_{jq}\,f\frac{\partial f}{\partial y}\frac{\partial^2 f}{\partial t\partial y} + 6\sum_j\beta_{ij}\sum_q\beta_{iq}\alpha_q\,f\frac{\partial f}{\partial t}\frac{\partial^2 f}{\partial y^2} \\
&+ 3\left(\sum_j\beta_{ij}\left(\sum_q\beta_{jq}\right)^2 + 2\sum_j\beta_{ij}\sum_{q,r}\beta_{iq}\beta_{qr}\right) f^2\frac{\partial f}{\partial y}\frac{\partial^2 f}{\partial y^2} \\
&+ 6\sum_{j,q}\beta_{ij}\beta_{jq}\alpha_q\,\frac{\partial f}{\partial t}\left(\frac{\partial f}{\partial y}\right)^2 + 6\sum_{j,q,r}\beta_{ij}\beta_{jq}\beta_{qr}\,f\left(\frac{\partial f}{\partial y}\right)^3. \qquad (14d)
\end{aligned}
$$

Funkcję f i jej pochodne należy teraz obliczać w punkcie (t_n, y_n).

Warunki

Pozostaje nam teraz uzgodnić współczynniki rozwinięcia (9) w kolejnych potęgach h. Zauważmy, że ponieważ metoda ma działać dla dowolnych f spełniających podane założenia, współczynniki występujące po obu stronach (9) przy pewnej kombinacji f i jej pochodnych muszą być sobie równe.

$p = 1$

W najniższym rzędzie uzgadniamy tylko współczynniki przy wyrazach rzędu h^1. Otrzymujemy

$$\frac{1}{1!} f = \frac{1}{0!} \sum_{i=1}^{s} w_i \, f \,,$$

czyli

$$\sum_{i=1}^{s} w_i = 1 \,, \tag{15}$$

co, jak już wiemy, jest warunkiem zgodności metody RK. Widać zatem, że każda zgodna metoda RK, bez względu na to, czy zachodzą jakieś dodatkowe związki pomiędzy jej współczynnikami, czy nie, jest metodą rzędu co najmniej pierwszego.

$p = 2$

Aby metoda była rzędu drugiego, żądamy spełnienia (15) oraz następujących warunków dodatkowych, wynikających z postulatu zgodności współczynników przy wyrazach rzędu h^2:

$$\sum_{i=1}^{s} w_i \alpha_i = \frac{1}{2}\,, \tag{16a}$$

$$\sum_{i=1}^{s} w_i \sum_{j=1}^{s} \beta_{ij} = \frac{1}{2}\,. \tag{16b}$$

Punkt ten wymaga pewnej dyskusji. Odejmując stronami (16a), (16b) otrzymujemy

$$\sum_{i=1}^{s} \left(\alpha_i - \sum_{j=1}^{s} \beta_{ij} \right) w_i = 0\,. \tag{17}$$

Gdyby w_i były liniowo niezależne, powyższy warunek oznaczałby, że koniecznie $\alpha_i = \sum\limits_{j=1}^{s} \beta_{ij}$. Ale w_i *nie* są liniowo niezależne, jako że wiąże je warunek zgodności (15)! W tej sytuacji (17) narzuca pewne dodatkowe więzy na współczynniki metody: można go potraktować jako dodatkowy więz na wagi w_i (tylko $s-2$ jest wówczas liniowo niezależnych), jednak znacznie wygodniej jest potraktować to jako warunek znikania współczynników kombinacji liniowej stojącej po prawej stronie (17) i tak też uczynimy.

Można też na to spojrzeć i od innej strony: Zgodność wyrazów rozwinięcia ma zachodzić w każdym rzędzie *niezależnie*, co jest możliwe tylko, gdy $\alpha_i = \sum_{j=1}^{s} \beta_{ij}$.

Wszystkie metody RK rzędu drugiego opisane są zatem układem równań

$$\sum_{i=1}^{s} w_i = 1\,, \tag{18a}$$

$$\sum_{i=1}^{s} w_i \alpha_i = \frac{1}{2}\,, \tag{18b}$$

$$\sum_{j=1}^{s} \beta_{ij} = \alpha_i\,. \tag{18c}$$

Zauważmy, iż z (18c) wynika, że we wszystkich metodach jawnych $\alpha_1 = 0$.

$p = 3$

W trzecim rzędzie dostajemy

$$\sum_{i=1}^{s} w_i \alpha_i^2 = \frac{1}{3}, \tag{19a}$$

$$\sum_{i=1}^{s} w_i \sum_{j=1}^{s} \beta_{ij} \alpha_j = \frac{1}{6}, \tag{19b}$$

gdzie skorzystaliśmy z (18c) w celu zmniejszenia ilości niezależnych równań. Równania te, wraz z równaniami (18a)–(18c), opisują wszystkie metody RK rzędu trzeciego. Widać, że o ile równania (18a)–(18c) narzucały związki liniowe na współczynniki α, β, równania (19a)–(19b) narzucają związki kwadratowe.

$p = 4$

W czwartym rzędzie, po skorzystaniu z (18c), jako nietrywialne równania dostajemy

$$\sum_{i=1}^{s} w_i \alpha_i^3 = \frac{1}{4}\,, \tag{20a}$$

$$\sum_{i=1}^{s} w_i \alpha_i \sum_{j=1}^{s} \beta_{ij} \alpha_j = \frac{1}{8}\,, \tag{20b}$$

$$\sum_{i=1}^{s} w_i \sum_{j=1}^{s} \beta_{ij} \alpha_j^2 = \frac{1}{12}\,, \tag{20c}$$

$$\sum_{i=1}^{s} w_i \sum_{j=1}^{s} \beta_{ij} \sum_{l=1}^{s} \beta_{jl} \alpha_l = \frac{1}{24}\,. \tag{20d}$$

Warto zaznaczyć, iż po zastosowaniu (18c) niektóre z równań na współczynniki są spełnione tożsamościowo, co stanowi niejako test wewnętrznej spójności teorii. Równania (20a)–(20d), łącznie z równaniami (19a)–(19b) oraz (18a)–(18c), opisują wszystkie metody RK rzędu czwartego.

Równania określające metody czwartego rzędu narzucają sześcienne związki pomiędzy współczynnikami α, β.

Wyrażenia, które muszą spełniać współczynniki metod wyższego rzędu, wyprowadza się na tej samej zasadzie co powyższe, są one jednak bardziej skomplikowane. W dowolnym rzędzie p równania te zawierają człon

$$\sum_{i=1}^{s} w_i \alpha_i^{p-1} = \frac{1}{p} \tag{21}$$

i narzucają związki rzędu $p-1$.

Przykład: Rząd jawnej metody punktu środkowego (6)

$$
\begin{array}{c|cc}
0 & 0 & 0 \\
\frac{1}{2} & \frac{1}{2} & 0 \\
\hline
 & 0 & 1
\end{array}
$$

Sprawdzamy: $w_1 + w_2 = 0 + 1 = 1$ — metoda jest co najmniej rzędu pierwszego

$w_1\alpha_1 + w_2\alpha_2 = 0 \cdot 0 + 1 \cdot \frac{1}{2} = \frac{1}{2}$
$\beta_{11} + \beta_{12} = 0 = \alpha_1; \qquad \beta_{21} + \beta_{22} = \frac{1}{2} + 0 = \frac{1}{2} = \alpha_2$ — metoda jest co najmniej rzędu drugiego

$w_1\alpha_1^2 + w_2\alpha_2^2 = 0 \cdot 0 + 1 \cdot \left(\frac{1}{2}\right)^2 = \frac{1}{4} \neq \frac{1}{3}$ — metoda nie jest rzędu trzeciego (ani wyższego).

Przykład: Rząd niejawnej metody punktu środkowego (7)

$$\begin{array}{c|c} \frac{1}{2} & \frac{1}{2} \\ \hline & 1 \end{array}$$

$w_1 = 1 = 1$ — metda jest rzędu co najmniej pierwszego

$w_1\alpha_1 = 1 \cdot \frac{1}{2} = \frac{1}{2}$
$\beta_{11} = \frac{1}{2} = \alpha_1$ — metoda jest co najmniej rzędu drugiego

$w_1\alpha_1^2 = 1 \cdot \left(\frac{1}{2}\right)^2 = \frac{1}{4} \neq \frac{1}{3}$ — metoda <u>nie</u> jest rzędu trzeciego (ani wyższego).

Metody jawne tego samego rzędu i o tej samej liczbie punktów pośrednich

Można udowodnić, że wszystkie jawne s-etapowe metody rzędu p, są, w pewnym sensie, równoważne: Wymagają takiej samej liczby obliczeń prawej strony równania, mają taki sam rząd i — dla ustalonych s i p — mają takie same obszary stabilności (patrz następny wykład). Na przykład jawna metoda punktu środkowego i jawna metoda trapezowa są w tym sensie równoważne. Jaki jest więc sens używania obu tych metod?

Metody co prawda są równoważne pod wymienionymi wyżej względami, ale uzyskiwane wyniki są numerycznie inne. Z rozmaitych powodów jedna z metod może być wygodniejsza w użyciu od innych — najczęściej chodzi o inną obsługę błędów. Możliwe są także inne powody, dla których któraś z metod może być preferowana względem innych.

Metoda Gilla

$$
\begin{array}{c|cccc}
0 & 0 & 0 & 0 & 0 \\
\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\
\frac{1}{2} & \frac{\sqrt{2}-1}{2} & \frac{2-\sqrt{2}}{2} & 0 & 0 \\
1 & 0 & -\frac{\sqrt{2}}{2} & \frac{2+\sqrt{2}}{2} & 0 \\
\hline
 & \frac{1}{6} & \frac{2-\sqrt{2}}{6} & \frac{2+\sqrt{2}}{6} & \frac{1}{6}
\end{array}
\tag{22}
$$

Jest to metoda jawna, tego samego rzędu, co klasyczna metoda czteroetapowa (8), o takiej samej złożoności obliczeniowej i takim samym obszarze stabilności. Co zatem, potencjalnie, przemawia za używaniem metody (22)? Okazuje się, iż metodę tę można zaprogramować tak, aby "zaoszczędzić" pamięć przydzielaną na jeden wektor pośredni.

Rozpiszmy metodę Gilla jawnie:

$$\mathbf{k}_1 = \mathbf{f}\left(x_n, \mathbf{y}_n\right), \tag{23a}$$

$$\mathbf{k}_2 = \mathbf{f}\left(x_n + \frac{1}{2}h, \mathbf{y}_n + \frac{1}{2}h\mathbf{k}_1\right), \tag{23b}$$

$$\mathbf{k}_3 = \mathbf{f}\left(x_n + \frac{1}{2}h, \mathbf{y}_n + \frac{1}{2}h\left((\sqrt{2}-1)\mathbf{k}_1 + (2-\sqrt{2})\mathbf{k}_2\right)\right) \tag{23c}$$

$$\mathbf{k}_4 = \mathbf{f}\left(x_n + h, \mathbf{y}_n + \frac{1}{2}h\left(-\sqrt{2}\mathbf{k}_2 + (2+\sqrt{2})\mathbf{k}_3\right)\right), \tag{23d}$$

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h\left(\frac{1}{6}\mathbf{k}_1 + \frac{2-\sqrt{2}}{6}\mathbf{k}_2 + \frac{2+\sqrt{2}}{6}\mathbf{k}_3 + \frac{1}{6}\mathbf{k}_4\right). \tag{23e}$$

Do prawidłowego działania algorytmu trzeba, jak się wydaje, zarezerwować pamięć potrzebną do przechowywania wektorów $\mathbf{k}_{1,2,3,4}$. Można się jednak obyć bez zapamiętywania jednego z tych wektorów. Klucz polega na obserwacji, iż ostatnie wyrażenie można przepisać w ten sposób:

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \frac{1}{6}h\left(\mathbf{k}_1 + 2\mathbf{k}_2 + \left(-\sqrt{2}\mathbf{k}_2 + (2+\sqrt{2})\mathbf{k}_3\right) + \mathbf{k}_4\right). \quad (24)$$

W wyrażeniu (24) występuje taka sama kombinacja wektorów, co w wyrażeniu (23d). Możemy wobec tego metodę Gilla zaprogramować przy użyciu następującego pseudokodu:

$$\mathbf{k}_1 = \mathbf{f}\left(x_n, \mathbf{y}_n\right), \tag{25a}$$

$$\mathbf{k}_2 = \mathbf{f}\left(x_n + \frac{1}{2}h, \mathbf{y}_n + \frac{1}{2}h\mathbf{k}_1\right), \tag{25b}$$

$$\mathbf{k}_3 = \mathbf{f}\left(x_n + \frac{1}{2}h, \mathbf{y}_n + \frac{1}{2}h\left((\sqrt{2}-1)\mathbf{k}_1 + (2-\sqrt{2})\mathbf{k}_2\right)\right) \tag{25c}$$

$$\mathbf{k}_3 = -\sqrt{2}\mathbf{k}_2 + (2+\sqrt{2})\mathbf{k}_3, \tag{25d}$$

$$\mathbf{k}_2 = \mathbf{k}_1 + 2\mathbf{k}_2, \tag{25e}$$

$$\mathbf{k}_1 = \mathbf{f}\left(x_n + h, \mathbf{y}_n + \frac{1}{2}h\mathbf{k}_3\right), \tag{25f}$$

$$\mathbf{y}_{n+1} = \mathbf{y} + \frac{1}{6}h\left(\mathbf{k}_2 + \mathbf{k}_3 + \mathbf{k}_1\right). \tag{25g}$$

W ten sposób nie musimy przydzielać osobnej pamięci do zapamiętania $\mathbf{k}_4$, jako że “recyklujemy” pamięć użytą do zapamiętania $\mathbf{k}_1$. W czasach, gdy pamięci (ferrytowe) były koszmarnie drogie, miało to istotne znaczenie.