Estudio y construcción de métodos de integración numérica de tipo Runge-Kutta-Hermite-Birkhoff





Diego Navalón Casalod

Trabajo de fin del grado en Matemáticas Universidad de Zaragoza

Summary

0.1. Introduction

In this project, the class of explicit Runge-Kutta-Hermite-Birkhoff (RKHB) methods [BNV] for solving initial value problems is considered. This class of numerical schemes use the first derivative as in Runge-Kutta methods and the second derivative as in Taylor methods. In each step, the second derivative is evaluated only at the initial point of each interval. These methods are useful when the computation of the second derivative, in a exact way or using an automatic differentiation software, has a similar computational cost than the derivative function.

The aim of this work is to give a formal analysis of the order conditions for this type of schemes like Runge-Kutta methods. Moreover, simplifying assumptions are used in order to reduce drastically the large number of the set of non-linear equations.

We consider the IVP

$$\begin{aligned} y'(t) &= f(t, y(t)), \quad t \in (t_0, T), \ y \in \mathbb{R}^m, \\ y(t_0) &= y_0, \end{aligned}$$
 (1)

where $f : \mathbb{R}^{m+1} \to \mathbb{R}^m$ is sufficiently differentiable. By adding the trivial equation t' = 1, the problem (1) can be transformed into an autonomous one, but increasing the dimension by one.

A s-stage explicit RKHB scheme is defined by

$$\mathbf{g}_{i} = f(\mathbf{t}_{n} + \mathbf{c}_{i}h, \mathbf{y}_{n} + h\sum_{j=1}^{i-1} \mathbf{a}_{ij}\mathbf{g}_{j} + h^{2}\gamma_{i}y_{n}''), \quad i = 1, \dots, s,$$
(2)

$$y_{n+1} = y_n + h \sum_{j=1}^{s} b_j g_j + h^2 \gamma_0 y_n'',$$
(3)

with γ_i , c_i , b_i , a_{ij} real coefficients defining the method.

In Butcher's notation [B], the above explicit RKHB method is represented by:

0	0	0								
c_2	γ_2	a ₂₁	0							
c_3	γ3	a ₃₁	a ₃₂	0				c	Γ	A
÷	÷	÷	·	۰.			≡			
c _s	γs	a _{s1}	a_{s2}		$a_{s,s-1}$	0			%	b'
	γo	b ₁	b ₂		b_{s-1}	b _s	-			

with $A \in \mathbb{R}^{s \times s}$, Γ , b, c $\in \mathbb{R}^{s}$.

For the analysis of the order, we consider an autonomous problem y' = f(y). It is worth to note that each stage is a B-serie

$$B(a, \mathbf{y}) = \mathbf{y} + \sum_{\tau \in T} \frac{h^{\rho(\tau)}}{\sigma(\tau)} a(\tau) F(\tau)(\mathbf{y})$$
(4)

where T is the set of (rooted) trees τ , σ is the symmetry coefficient, ρ is the order of the tree τ and F(y) is an elementary differential.

The main result of the theory of B-series is due to Butcher [B]:

Theorem 0.1.1. Let $a : T \cup \{\emptyset\} \to \mathbb{R}$ be a mapping satisfying $a(\emptyset) = 1$. Then if B(a, y) is a B-serie, hf(B(a, y)) is again a B-serie B(a', y) where $a'(\emptyset) = 0$, $a'(\bullet) = 1$, and

$$a'(\tau) = a(\tau_1) \cdot \ldots \cdot a(\tau_k)$$

for $\tau = \{\tau_1, ..., \tau_k\}.$

In this way, the function $\Phi(\tau) = (\Phi_1(\tau), \dots, \Phi_s(\tau)) \in \mathbb{R}^s$, where $\Phi_i(\tau)$ is the B-serie of the *i*-stage of the numerical scheme and it can be defined as $\Phi(\emptyset) = 0 \in \mathbb{R}^s$, $\Phi(\bullet) = \mathbf{e} = (1, \dots, 1)^\top$, and

$$\Phi(\tau) = \Phi(\tau_1) \bullet \ldots \bullet \Phi(\tau_k),$$

for $\tau = {\tau_1, ..., \tau_k}$, where the • indicates the componentwise product between vectors.

The main difference between the B-series of the *i*-stage of a RK and RKHB is the term $h^2 \gamma_i y''$. This term is associated to the tree \clubsuit and then by applying Theorem 0.1.1, the coefficient associated to this tree is $Ac + \Gamma$. Note that in RK schemes, $\Gamma = 0 \in \mathbb{R}^s$ and therefore this class of methods reduce to Runge-Kutta ones when $\gamma_i = 0$, $\forall i$.

Theorem 0.1.2. The RKHB method has order p if and only if

$$b^{ op} \Phi(au) = rac{1}{\gamma(au)}, \ for \ oldsymbol{
ho}(au) \leq p,$$

where $\gamma(\tau) = \rho(\tau)\gamma(\tau_1)\ldots\gamma(\tau_k)$.

In this table we show the order conditions up to fourth-order.

ho(au)	τ	F(au)	$b^{\top} \Phi(\tau)$
1	•	f	$\mathbf{b}^{\top}\mathbf{e} = 1$
2	•	f'f	$\mathbf{b}^{\top}\mathbf{c} + \gamma_0 = 1/2$
	٩		
3		f'f'f	$\mathbf{b}^{\top}(A\mathbf{c}+\Gamma)=1/6$
3		f''(f,f)	$b^{\top}c^{2} = 1/3$
4	V	$f^{\prime\prime\prime}(f,f,f)$	$\mathbf{b}^{\top}\mathbf{c}^3 = 1/4$
4		$f^{\prime\prime\prime}(f,f,f)$	$\mathbf{b}^{\top}(\mathbf{c}\cdot(A\mathbf{c}+\Gamma))=1/8$
4		$f^{\prime\prime\prime}(f,f,f)$	$\mathbf{b}^{\top} A \mathbf{c}^2 = 1/12$
4		f'f'f'f	$\mathbf{b}^{\top} A(A\mathbf{c} + \Gamma) = 1/24$

The following result follows from the analysis of the tall trees:

Estudio y construcción de métodos de integración numérica de tipo Runge-Kutta-Hermite-Birkhoff

Corollary 0.1.3. The maximum attainable order for an explicit s-stage RKHB is s + 1.

In order to manipulate the set of large number of non-linear equations, it is convenient to introduce the so-called simplifying hypothesis

$$A\mathbf{e} = \mathbf{c}, \quad A\mathbf{c} + \Gamma = \frac{\mathbf{c}^2}{2}, \quad A\mathbf{c}^2 = \frac{\mathbf{c}^3}{3} - \frac{\mathbf{c}_2^3}{3}\mathbf{e}_2,$$
 (5)

being $\mathbf{e}_2 = (0, 1, 0, \dots, 0)^{\top}$.

To obtain a high-quality pair, we take s = 5 stages, and we solve the system of the set of fifthorder order conditions with the simplifying hipothesis (5). Obviously, $b_2 = 0$ and $\gamma_2 = \frac{c_2^2}{2}$. Solving sequentially (5) and the set of order conditions relative to the bushy trees

$$\mathbf{b}^{\top}\mathbf{e} = 1, \quad \mathbf{b}^{\top}\mathbf{c} + \gamma_0 = \frac{1}{2}, \quad \mathbf{b}^{\top}\mathbf{c}^i = \frac{1}{i+1}, \quad i = 2, \dots, 4,$$

for b_i and γ_0 , imposing the order conditions $b^{\top}A^2(Ac + \Gamma) = \frac{1}{5!}$ and $b^{\top}A(c \cdot (Ac + \Gamma)) = \frac{1}{40}$ we get a family of fifth-order methods depending on the free parameters γ_5 and c_i , i = 2, ..., 5.

To construct the fourth-order embedded solution of the pair 5(4), we have at our disposal a sufficient number of \hat{b} parameters, therefore the FSAL technique will not be used.

Proposition 0.1.4. The stability function of an RKHB is given as

$$R(z) = 1 + zb^{\top} (I - zA)^{-1} (e + z^{2}\Gamma) + z^{2}\gamma_{0},$$

and for explicit methods is a polynom.

For the five-stage, fifth order scheme under consideration,

$$R(z) = 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \frac{z^4}{24} + \frac{z^5}{120} + \frac{(2c_2 - 1)c_3z^6}{720(c_2 - c_3)}$$

For the optimization of these methods, a compromise between accuracy and stability properties is $c \mid A$

taken into account. A set of optimized parameters leads to the high-order scheme $\frac{c}{b^T}$

and

$$\gamma_0 = 0, \Gamma = \left(0, \frac{1}{128}, \frac{-115 - 49r}{300}, \frac{115 - 41r}{300}, \frac{-4 + 9r}{6}\right)^{\top}, r = \sqrt{5}$$

and in figure 1 we plot the stability domain for our scheme comparing it against the fifth-order approximation of the Dormand and Prince scheme. For the sake of comparison, the $\|\cdot\|_2$ norm of the coefficients of the elementary differentials in the leading term of the local error expansion of the formula of high order in our approximation is 2.59×10^{-4} and 3.99×10^{-4} for the Dormand and Prince one.

0.2. Numerical experiments

Here we present the results obtained on a representative numerical problem and we compare the behavior of the new methods developped against the embedded pair 5(4) of Dormand and Prince. The criterion used is the classical test of the maximum global error in the solution over the whole integration interval.

Example. The rigid solid equations

$$\begin{aligned} y_1' &= (\alpha - \beta) y_2 y_3 \\ y_2' &= (1 - \alpha) y_3 y_1 \\ y_3' &= (\beta - 1) y_1 y_2, \quad t \in [0, 20], \\ y(0) &= (0, 1, 1)^\top, \end{aligned}$$

with $\alpha = 1 + \frac{1}{\sqrt{1.51}}$, $\beta = 1 - \frac{0.51}{\sqrt{1.51}}$. The exact solution is given by:

$$y(t) = \left(\sqrt{1.51}\operatorname{sn}(t, 0.51), \operatorname{cn}(t, 0.51), \operatorname{dn}(t, 0.51)\right)^{\top}$$

where sn, cn, dn are the Jacobi's elliptical functions.

We have integrated this problem implementing both methods in a variable step mode with tolerances $tol = 10^{-i}$, i = 3, ..., 8. In figure 2 we show the efficiency plot (global error *vs.* number of functions evaluations) for each considered method.



Fig. 1 Stability regions for fifth order methods (continuous line RKHB5(4), dashed line DOPRI5(4))

Estudio y construcción de métodos de integración numérica de tipo Runge-Kutta-Hermite-Birkhoff



From the numerical results obtained in several problems and particularly in the problem presented here, it follows that the new pair RKHB5(4) is usually more efficient than the Dormand and Prince pair of the same order.

Índice general

Su	mma	ry	III
	0.1.	Introduction	III
	0.2.	Numerical experiments	VI
1.	Intro	oducción a los métodos Runge-Kutta	1
	1.1.	Estimación del error local. Pares encajados	6
	1.2.	Estabilidad Absoluta	7
2.	Méte	odos Runge-Kutta-Hermite-Birkhoff	9
	2.1.	Notación RKHB	9
	2.2.	Condiciones de orden	10
	2.3.	Estabilidad Absoluta	11
	2.4.	Construcción de métodos RKHB a paso fijo	11
		2.4.1. Métodos de orden 3 con 2 etapas	12
		2.4.2. Métodos de orden 4 con 3 etapas	12
	2.5.	Construcción de pares encajados	14
		2.5.1. Par encajado 4(3) con 3 etapas	14
		2.5.2. Par encajado 5(3) con 4 etapas	16
		2.5.3. Par encajado 5(4) con 5 etapas	17
	2.6.	Métodos continuos	17
3.	Exp	erimentos Numéricos	19
	3.1.	Conclusiones	23
Bi	bliogr	rafía	25
Ar	iexo I		27

Capítulo 1

Introducción a los métodos Runge-Kutta

Un método Runge-Kutta (RK) [B] [CMR] explícito de s etapas para la resolución numérica de un Problema de Valor Inicial (PVI)

$$(PVI) = \begin{cases} y'(t) = f(t, y(t)) \\ y(t_n) = y_n \end{cases}$$
(1.1)

donde $f : \mathbb{R}^{m+1} \to \mathbb{R}^m$ es suficientemente diferenciable, es un algoritmo en el que a partir de (t_n, y_n) se obtiene una aproximación y_{n+1} a la solución exacta y(t) en $t_n + h$ dada por

$$g_{1} = f(t_{n}, y_{n}),$$

$$g_{2} = f(t_{n} + c_{2}h, y_{n} + ha_{21}g_{1}),$$

$$g_{3} = f(t_{n} + c_{3}h, y_{n} + ha_{31}g_{1} + ha_{32}g_{2}),$$

$$\vdots$$

$$g_{s} = f(t_{n} + c_{s}h, y_{n} + h(a_{s1}g_{1} + a_{s2}g_{2} + \dots + a_{s,s-1}g_{s-1})),$$
scion en $t_{n} + h$

siendo la aproximación en $t_n + b$

$$y_{n+1} = y_n + h(b_1g_1 + \ldots + b_sg_s).$$
 (1.2)

Utilizando la notación de Butcher, el esquema anterior se puede expresar de la siguiente manera:

$$\begin{array}{c|c} c & A \\ \hline & \\ \hline & \\ b^T \end{array}$$

donde los vectores b y c representan los pesos y nodos respectivamente de la fórmula de cuadratura.

No es necesario considerar ecuaciones diferenciales ordinarias (EDOs) de orden p del tipo y^{(p} = $g(t, y, ..., y^{(p-1)})$, pues éstas pueden reescribirse como un sistema de primer orden y dimensión p con el cambio de variable

$$y_{1} = y, y_{2} = y', \dots, y_{p} = y^{(p-1)}$$

$$\begin{cases} y'_{1} = y_{2}, \\ y'_{2} = y_{3}, \\ \vdots \\ y'_{p} = g(t, y_{1}, y_{2}, \dots, y_{p-1}). \end{cases}$$
(1.3)

obteniendo

También es posible transformar un problema no autónomo y' = f(t, y) a un problema autónomo z' = F(z). En efecto, llamando $z^{\top} = (t, y^{\top})$ se tiene

$$z' = (1, f(\mathbf{t}, \mathbf{y})) = (1, f(z)) \equiv F(z),$$

y, para que el problema autónomo y el no autónomo asociado sean equivalentes, es necesario y suficiente que

$$\mathbf{b}^{\top}\mathbf{e} = 1$$
 \mathbf{y} $\mathbf{A}\mathbf{e} = c$, $\mathbf{e} = (1, \dots, 1)^{\top}$.

Estas condiciones las cumplen la inmensa mayoría de los RK.

Un concepto fundamental en la resolución numérica de problemas de valor inicial es el de orden de convergencia del esquema de integración numérico empleado.

Definición 1.0.1. Denotando por $y(\hat{t}+h;\hat{t},\hat{y})$ a la solución exacta local del PVI y' = f(t,y), $y(\hat{t}) = \hat{y}$, se define el error local de la fórmula (1.2) en el punto (\hat{t},\hat{y}) con paso de integración h a la diferencia

$$e(\widehat{t},\widehat{y},h) = y\left(\widehat{t}+h;\widehat{t},\widehat{y}\right) - \left[\widehat{y}+h\phi(\widehat{t},\widehat{y},h)\right]$$

donde y $(\hat{t}+h;\hat{t},\hat{y})$ es la solución exacta local e $\hat{y}+h\phi(\hat{t},\hat{y},h)$ es la aproximación numérica, siendo ϕ la función incremento del esquema numérico (1.2).

Para cuantificar la diferencia anterior en potencias de h, es usual desarrollar en serie de Taylor cada uno de los sumandos. Primeramente, para hallar el desarrollo de la solución teórica, es preciso calcular las sucesivas derivadas de la misma, que para el problema autónomo son

$$\begin{aligned} \mathbf{y}' &= f(\mathbf{y}) = f \\ \mathbf{y}'' &= f'\mathbf{y}' = f'f \\ \mathbf{y}''' &= f''(f,f) + f'f'f \\ \mathbf{y}^{IV} &= f'''(f,f,f) + f''(f'f,f) + f''(f,f'f) + f'f''(f,f) + f'f'f'f \\ &= f'''(f,f,f) + 3f''(f'f,f) + f'f''(f,f) + f'f'f'f \\ &\vdots \end{aligned}$$

donde se ha denotado por f' a la forma lineal $\frac{\partial f}{\partial y}$ (matriz Jacobiana), por f'' la derivada segunda como forma bilineal, y así sucesivamente, estando todas ellas evaluadas en la solución exacta del PVI (1.1).

Definición 1.0.2. Se denominan **diferenciales elementales** de f de orden q a cada uno de los sumandos anteriores que aparecen en la derivada $y^{(q)}$.

En el análisis de las diferenciales elementales, Butcher [B] destacó la existencia de una correspondencia biunívoca entre éstas y ciertos grafos que denominaremos árboles con raíz.

Definición 1.0.3. *El conjunto de árboles con raíz T se define recursivamente:*

- *El grafo con un vértice, llamado raíz, pertenece a T y lo denotaremos por* τ_0 *.*
- Si $\tau_1, \ldots, \tau_k \in T$, entonces el árbol $\tau = \{\tau_1, \ldots, \tau_k\} \in T$

Estudio y construcción de métodos de integración numérica de tipo Runge-Kutta-Hermite-Birkhoff



donde los argumentos que aparecen en la forma *p*-lineal $f^{(p}(*,...,*)$ se asocian a árboles que se añaden al final de cada rama del árbol correspondiente.

A cada diferencial elemental de $y^{(p)}$ se le denota por $F(\tau)(y)$, siendo τ el árbol con raíz correspondiente. Ejemplos de dichas diferenciales elementales vienen dados en la Tabla 1.1



Tabla 1.1 Diferenciales elementales

Es interesante notar que las distintas diferenciales elementales que aparecen en $y^{(p)}$, verifican que la suma de todos los índices de derivación de cada una de ellas es igual a p - 1.

Definición 1.0.4. *El orden de un árbol* τ *, denotado por* $\rho(\tau)$ *, corresponde al número de vértices que posee dicho árbol.*

Debido a esto, la derivada p-ésima de la solución exacta se puede escribir como

$$\mathbf{y}^{(p}(\mathbf{t}) = \sum_{\rho(\tau)=p} \alpha(\tau) F(\tau)(\mathbf{y}), \tag{1.4}$$

donde $\alpha(\tau)$ es el coeficiente que acompaña a la diferencial elemental asociada al árbol τ .

Definición 1.0.5. *Dado un árbol* $\tau \in T$, *la diferencial elemental* $F(\tau) : \mathbb{R}^m \to \mathbb{R}^m$ *se define recursivamente:*

- $F(\tau_0)(y) = f(y)$,
- $F(\tau)(y) = f^{(k)}(F(\tau_1)(y), \dots, F(\tau_k)(y))$ si $\tau = \{\tau_1, \dots, \tau_k\}.$

Entonces, desarrollando por Taylor y por (1.4) es evidente que

$$\mathbf{y}(\mathbf{t}_0 + h) = \mathbf{y}_0 + \sum_{\tau \in \mathbf{T}} \frac{h^{\rho(\tau)}}{\rho(\tau)!} \alpha(\tau) F(\tau)(\mathbf{y}).$$

Por último, será de gran utilidad el coeficiente de simetría de un árbol τ para la definición de las series normalizadas de Butcher.

Definición 1.0.6. *El coeficiente de simetría, que se denotará por* $\sigma(\tau)$ *, se define como:*

- $\sigma(\bullet) = 1$,
- $\sigma(\tau) = \sigma(\tau_1)\sigma(\tau_2)\dots\mu_1!\mu_2!\dots$ si $\tau = \{\tau_1,\dots,\tau_k\}$ y μ_i es el número de veces que se repite el árbol τ_i .

Así, ejemplos ilustrativos del coeficiente de simetría de diferentes árboles son

$$\sigma\left(\stackrel{\bullet}{\bullet}\right) = \{\bullet, \bullet, \bullet\} = \sigma(\bullet) \cdot 3! = 1 \cdot 6 = 6,$$

$$\sigma\left(\stackrel{\bullet}{\bullet}\right) = \{\stackrel{\bullet}{\bullet}, \stackrel{\bullet}{\bullet}\} = \sigma\left(\stackrel{\bullet}{\bullet}\right) \cdot 2! = \{\bullet\} \cdot 2 = 2.$$

Gran parte del desarrollo teórico de los métodos Runge-Kutta fue desarrollado por John C. Butcher [B], quien introdujo el concepto de B-serie y el Teorema Fundamental de B-series, el cual sirve como herramienta básica para el desarrollo de la aproximación numérica.

Definición 1.0.7. Dada una aplicación $a : T \cup \{\phi\} \to \mathbb{R}$, una *B*-serie B(a,y) es una expresión de la forma

$$B(a,y) \equiv a(\phi)y + \sum_{\tau \in T} \frac{a(\tau)}{\sigma(\tau)} F(\tau)(y) h^{\rho(\tau)}.$$

A continuación se enuncia el Teorema Fundamental de B-series:

Teorema 1.0.8. Sea $a: T \cup \{\phi\} \to \mathbb{R}$ una aplicación tal que $a(\phi) = 1$ definiendo una B-serie B(a,y). Entonces hf(B(a,y)) es una serie de Butcher B(a',y) con:

$$\left\{ egin{array}{ll} a'(\phi) = 0, \ a'(ullet) = 1, \ a'(au) = a(au_1) \dots a(au_k) & si \ au = \{ au_1, \dots, au_k\}. \end{array}
ight.$$

Como aplicación práctica calcularemos los coeficientes de la serie de Butcher de la solución exacta en el punto $t_0 + h$. Supongamos que

$$\mathbf{y}(\mathbf{t}_0+h) = B(e,\mathbf{y}_0) = e(\phi)\mathbf{y}_0 + \sum_{\tau \in T} \frac{e(\tau)}{\sigma(\tau)} h^{\rho(\tau)} F(\tau)(\mathbf{y}_0); \quad e(\phi) = 1,$$

derivando formalmente y multiplicando por h,

$$hy'(t_0 + h) = hf(y(t_0 + h)) = \sum_{\tau \in T} \frac{e(\tau)}{\sigma(\tau)} \rho(\tau) h^{\rho(\tau)} F(\tau)(y_0),$$
(1.5)

Estudio y construcción de métodos de integración numérica de tipo Runge-Kutta-Hermite-Birkhoff

y, como y(t₀ + *h*) es una B-serie, por el teorema Fundamental de B-series se tiene que $hf(y(t_0 + h)) = hf(B(e,y))$ es una B-serie B(e',y) con coeficientes:

$$\begin{cases} e'(\phi) = 0, \\ e'(\bullet) = 1, \\ e'(\tau) = e(\tau_1) \dots e(\tau_k); \quad \tau = \{\tau_1, \dots, \tau_k\}. \end{cases}$$
(1.6)

Por la ecuación (1.5) se cumple que $e'(\tau) = e(\tau)\rho(\tau)$, y por (1.6) $e'(\tau) = e(\tau_1)\dots e(\tau_k)$, por lo que $e(\tau) = \frac{1}{\rho(\tau)}e(\tau_1)\dots e(\tau_k)$ si $\tau = \{\tau_1, \dots, \tau_k\}$.

Usualmente, en lugar de considerar $e(\tau)$, se emplea el producto de inundación $\gamma(\tau) = \frac{1}{e(\tau)}$, definido como $\gamma(\bullet)=1$ y $\gamma(\tau) = \rho(\tau)\gamma(\tau_1)\ldots\gamma(\tau_k)$ si $\tau = \{\tau_1,\ldots,\tau_k\}$. **Ejemplos:**

$$\gamma \begin{pmatrix} \bullet & \bullet \\ \bullet \\ \bullet \end{pmatrix} = 4 \cdot \gamma (\bullet) \gamma (\bullet) \gamma (\bullet) = 4$$
$$\gamma \begin{pmatrix} \bullet \\ \bullet \\ \bullet \\ \bullet \end{pmatrix} = 4 \cdot \gamma \begin{pmatrix} \bullet \\ \bullet \\ \bullet \\ \bullet \end{pmatrix} = 4 \cdot 3 \cdot \gamma \begin{pmatrix} \bullet \\ \bullet \\ \bullet \\ \bullet \end{pmatrix} = 4 \cdot 3 \cdot 2 \cdot \gamma (\bullet) = 4!$$

En cuanto al desarrollo en serie de la aproximación numérica, si aplicamos sucesivamente el Teorema Fundamental de B-series en cada etapa y si $G_i(\tau) \equiv G(a_i, \tau)$ es el desarrollo en serie de Butcher de la etapa i-ésima y $\Phi(\tau) = (G_1(\tau), G_2(\tau), \dots, G_s(\tau))^T$, se tiene que

$$\begin{cases} \Phi(\bullet) = (1, 1, \dots, 1)^T = \mathbf{e}, \\ \Phi(\phi) = (0, \dots, 0)^T, \\ \Phi(\tau) = \mathbf{A}\Phi(\tau_1) \bullet \mathbf{A}\Phi(\tau_2) \bullet \dots \bullet \mathbf{A}\Phi(\tau_k) & \text{si } \tau = \{\tau_1, \dots, \tau_k\}, \end{cases}$$

siendo • el producto componente a componente entre dos vectores.

Finalmente puede enunciarse el resultado fundamental sobre el orden de un método RK.

Definición 1.0.9. Se dice que un método RK tiene orden p si se verifica

$$e(t, y, h) = h^{p+1} \sum_{\rho(\tau) = p+1} C(\tau) F(\tau)(y) + O(h^{p+2}), \qquad (1.7)$$

siendo $C(\tau) = \frac{\alpha(\tau)}{\rho(\tau)!} \left(1 - \gamma(\tau)b^{\top}\phi(\tau)\right)$. Al primer término de (1.7) se le denomina término principal del error local (TPEL).

Teorema 1.0.10. Un método RK tiene orden p si y sólo si

$$b^T \Phi(\tau) = \frac{1}{\gamma(\tau)}, \quad \forall \tau \, tal \, que \, \rho(\tau) \leq p.$$

1.1. Estimación del error local. Pares encajados

Cuando se utiliza un método RK se está avanzando la integración en un paso, es decir, se determina una aproximación y_{n+1} a la solución del PVI (1.1) en el punto t_{n+1} partiendo de (t_n, y_n) . Pero habitualmente se integra sobre un intervalo lo suficientemente grande para que sea necesario realizar numerosos pasos. Para mejorar el código es interesante dotarlo de un mecanismo que le permita variar cada paso, por ejemplo, de forma que el error local cometido fuera menor que una cierta tolerancia TOL fijada por el usuario. De esta forma, si al avanzar un paso en la integración se comprueba que el error local se mantiene por debajo de la tolerancia se da por bueno y se permite aumentar el paso, mientras que si es más grande se rechaza y se repite la integración con un paso menor.

Evidentemente, en la práctica no se conoce el error local cometido exactamente, por lo que se debe recurrir a ciertas estimaciones del mismo.

La técnica más ampliamente implementada en los códigos modernos es la de pares encajados, que debido a que los métodos que se consideran hacen uso de este procedimiento se verá brevemente en qué consiste.

Se suponen dos métodos RK, un primer método (A,b) de orden p y otro segundo $(\widehat{A}, \widehat{b})$ de orden $q \le p-1$, los cuales son utilizados para avanzar la integración de un paso. De esta forma se cometen dos errores locales distintos:

$$e(\mathbf{t}_{n},\mathbf{y}_{n},h) = \mathbf{y}(\mathbf{t}_{n+1};\mathbf{t}_{n},\mathbf{y}_{n}) - \mathbf{y}_{n+1} = \boldsymbol{\eta}(\mathbf{y}_{n})h^{p+1} + O(h^{p+2}),$$

para el método (A,b) y

$$\widehat{e}(\mathbf{t}_n, \mathbf{y}_n, h) = \mathbf{y}(\mathbf{t}_{n+1}; \mathbf{t}_n, \mathbf{y}_n) - \widehat{\mathbf{y}}_{n+1} = \widehat{\eta}(\mathbf{y}_n)h^{q+1} + O\left(h^{q+2}\right),$$

para el método $(\widehat{A}, \widehat{b})$, siendo η y $\widehat{\eta}$ los correspondientes TPEL de ambos métodos.

Por lo tanto, la diferencia entre ambas aproximaciones

$$est(h, y_n) = \hat{y}_{n+1} - y_{n+1} = \hat{\eta}(t_n, y_n)h^{q+1} + O(h^{q+2}), \qquad (1.8)$$

supone una estimación asintóticamente correcta del error local del método de orden bajo.

Por tanto, atendiendo a (1.8), una estimación del error local del método de orden bajo asintóticamente correcta es

$$\mathrm{EST}(h) = h \sum_{i=1}^{s} (\widehat{\mathbf{b}}_i - \mathbf{b}_i) \mathbf{g}_i.$$
(1.9)

Es claro que, para realizar la estimación del error local hay que efectuar todas las evaluaciones de función derivada que correspondan a los dos métodos. Debido a ello, la matriz \widehat{A} se toma como una submatriz de A de forma que puedan ser compartidas la mayor cantidad de etapas.

Definición 1.1.1. Una pareja de métodos de la forma anterior se llama **par de métodos encajados** y se denotan por **RKp(q)**.

Estudio y construcción de métodos de integración numérica de tipo Runge-Kutta-Hermite-Birkhoff

1.2. Estabilidad Absoluta

La estabilidad absoluta lineal estudia el comportamiento de un método de integración numérica cuando se aplica al problema modelo lineal

$$\mathbf{y}' = \lambda \mathbf{y}, \ \mathbf{y}(0) = \mathbf{y}_0, \ \lambda \in \mathbb{C}, \ \operatorname{Re}(\lambda) \le 0,$$
 (1.10)

exigiendo que la sucesión de aproximaciones numéricas $\{y_n\}$ converja hacia cero cuando $n \to \infty$ con h λ fijo.

Definición 1.2.1. Dado un método numérico, si y_n denota la aproximación a la solución del PVI anterior en tiempo t_n con paso h>0 constante, entonces se define la región de estabilidad como

$$D = \{ z = \lambda h \in \mathbb{C} : y_n \xrightarrow[n \to \infty]{} 0 \}.$$

Cuando un RK se aplica al problema (1.10), se obtiene

$$\mathbf{y}_{n+1} = R(\mathbf{h}\boldsymbol{\lambda}) \cdot \mathbf{y}_n,$$

donde

$$R(z) = 1 + z\mathbf{b}^{\top}(I - zA)^{-1}\mathbf{e},$$

es la función de amplificación del esquema utilizado.

Capítulo 2

Métodos Runge-Kutta-Hermite-Birkhoff

En este capítulo se introducirán los métodos concernientes a este trabajo. Como ocurría con los métodos Runge-Kutta clásicos, los métodos Runge-Kutta-Hermite-Birkhoff (RKHB) utilizan la primera derivada, pero la novedad es que estos esquemas utilizan también la segunda derivada que solo será evaluada en el punto inicial de cada paso de integración. Destacar que este tipo de esquemas serán útiles cuando, a la hora de calcular la segunda derivada, ya sea de manera exacta o utilizando diferenciadores automáticos, tenga un coste computacional muy similar al de la función derivada.

En primer lugar se analizará formalmente las condiciones de orden de este tipo de esquemas, de manera similiar a los métodos Runge-Kutta, observando si hay alguna particularidad especial en algún árbol. Además, debido a la alta cantidad de ecuaciones no lineales de las ecuaciones de orden, se emplearán hipótesis simplificadoras adecuadas para facilitar el manejo y resolución de las mismas.

2.1. Notación RKHB

Se considera, como antes, el siguiente problema de valor inicial:

$$(PVI) \begin{cases} \mathbf{y}'(\mathbf{t}) = f(\mathbf{t}, \mathbf{y}(\mathbf{t})) & \mathbf{t} \in (\mathbf{t}_0, T), \mathbf{y} \in \mathbb{R}^m, \\ \\ \mathbf{y}(\mathbf{t}_0) = \mathbf{y}_0 \end{cases}$$

donde $f : \mathbb{R}^{m+1} \to \mathbb{R}^m$ es suficientemente diferenciable. Añadiendo también la ecuación trivial t' = 1, el problema, como en el capítulo anterior, se puede transformar en un problema autónomo con una dimensión más.

Un método RKHB [BNV] explícito de s etapas se define por

$$\begin{cases} g_{1} = f(t_{n}, y_{n}) \\ g_{2} = f(t_{n} + c_{2}h, y_{n} + ha_{21}g_{1} + h^{2}\gamma_{2}y_{n}'') \\ g_{3} = f(t_{n} + c_{3}h, y_{n} + ha_{31}g_{1} + ha_{32}g_{2} + h^{2}\gamma_{3}y_{n}'') \\ \vdots \\ g_{s} = f(t_{n} + c_{s}h, y_{n} + h(a_{s1}g_{1} + a_{s2}g_{2} + \dots + a_{s,s-1}g_{s-1}) + h^{2}\gamma_{s}y_{n}'') \end{cases}$$
(2.1)

y la aproximación en $t_n + h$ utilizada en estos métodos viene dada por

$$y_{n+1} = y_n + h(b_1g_1 + \ldots + b_sg_s) + h^2\gamma_0y_n'' = y_n + h\sum_{j=1}^s b_jg_j + h^2\gamma_0y_n''$$

Este tipo de métodos también se puede escribir utilizando la notación de Butcher de la siguiente manera

 $\operatorname{con} A \in \mathbb{R}^{s \times s}, \Gamma, \mathbf{b}, c \in \mathbb{R}^{s}.$

2.2. Condiciones de orden

Como puede apreciarse de la definición (2.1), la principal diferencia entre las B-series que quedan definidas en la etapa *i*-ésima entre los métodos Runge-Kutta y los métodos Runge-Kutta-Hermite-Birkhoff es el término $h^2 \gamma_i y''$ en la evaluación interna de la función derivada. Notar que, al ser y'' = f'f la diferencial elemental asociada, el árbol que le acompaña será \mathbf{s} , y utilizando el Teorema 1.0.8, el coeficiente asociado a este árbol es ahora $Ac + \Gamma$, siendo $\Gamma = (0, \gamma_2, \dots, \gamma_s)^{\top}$. Es interesante notar que los métodos RKHB son una generalización de los RK, ya que si $\Gamma = 0$, estos esquemas se reducen a los Runge-Kutta tradicionales.

Igual que con los métodos RK, estos métodos también tendrán orden $\geq p$ si se cumple que $b^T \Phi(\tau) = \frac{1}{\gamma(\tau)} \forall \tau \text{ tal } \text{que } \rho(\tau) \leq p.$

La evaluación de la segunda derivada en estos métodos nos otorga una ventaja en cuanto al orden máximo alcanzable, ya que los métodos RK explícitos de *s* etapas pueden alcanzar orden máximo *s*, mientras que en los métodos RKHB explícitos se tiene el siguiente resultado:

Proposición 2.2.1. El orden máximo alcanzable por un método RKHB explícito de s etapas es s+1.

Demostración. Se considera el árbol de *s*+2 nodos de altura *s*+2. La condición de orden asociada a dicho árbol es $b^T A^{s-1}(Ac + \Gamma) = \frac{1}{(s+2)!}$. Ahora bien, como el método es de orden s y A es una matriz nilpotente de orden s por ser el método explícito, se verifica que $b^T A^{s-1}(Ac) = 0 \neq \frac{1}{(s+2)!}$. En cuanto al otro término, como

$$\mathbf{A}^{s-1} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ * & 0 & \dots & 0 \end{bmatrix} \mathbf{y} \, \Gamma = \begin{bmatrix} 0 \\ \gamma_2 \\ \vdots \\ \gamma_s \end{bmatrix}$$

se tiene que $b^T A^{s-1} \Gamma = 0$ y es imposible cumplir la condición de orden para dicho árbol. Por tanto, el orden máximo alcanzable por un método RKHB de s etapas es s+1.

2.3. Estabilidad Absoluta

A continuación se va a construir la función de amplificación para estos métodos y, con ella, se podrá estudiar la estabilidad absoluta de los mismos. Aplicando el esquema RKHB a la ecuación test de prueba y' = λ y se tiene el siguiente resultado:

Proposición 2.3.1. La función de amplificación de un método RKHB viene dada por

$$R(z) = 1 + zb^T (I - zA)^{-1} (\mathbf{e} + z^2 \Gamma) + z^2 \gamma_0.$$

Demostración. Considerando el problema de valor inicial antes mencionado, las etapas para un método RKHB se pueden reescribir como

$$Y = y_n \mathbf{e} + \mathbf{h}\lambda \mathbf{A}Y + h^2 \lambda^2 \Gamma y_n,$$

	$\begin{bmatrix} Y_1 \end{bmatrix}$		0	0		0		0		1	
donde $Y =$	$\begin{array}{c} Y_2 \\ \vdots \end{array}$, A=	a ₂₁	0 ·	···· ·.	0	, $\Gamma =$	γ ₂ :	y e =	1	
	Y _s		a _{s1}		$a_{s,s-1}$	0		γs		1	

Despejando *Y* de la ecuación anterior se tiene que $Y = (I - h\lambda A)^{-1} (\mathbf{e} + h^2 \lambda^2 \Gamma) \mathbf{y}_n$.

Como la aproximación que se busca es

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h\mathbf{b}^T \boldsymbol{\lambda} \mathbf{Y} + \gamma_0 h^2 \boldsymbol{\lambda}^2 \mathbf{y}_n,$$

utilizando la expresión de Y obtenida

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h\mathbf{b}^T \lambda (I - h\lambda \mathbf{A})^{-1} (\mathbf{e} + h^2 \lambda^2 \Gamma) \mathbf{y}_n + \gamma_0 h^2 \lambda^2 \mathbf{y}_n$$
$$\mathbf{y}_{n+1} = \left(1 + h\lambda b^T (I - h\lambda \mathbf{A})^{-1} (\mathbf{e} + h^2 \lambda^2 \Gamma) + \gamma_0 h^2 \lambda^2\right) \mathbf{y}_n.$$

Llamando $z = \lambda h$, se tiene

$$\mathbf{y}_{n+1} = \left(1 + zb^T (I - z\mathbf{A})^{-1} (\mathbf{e} + z^2 \Gamma) + \gamma_0 z^2\right) \mathbf{y}_n = \mathbf{R}(z) \mathbf{y}_n.$$

En el caso de métodos explícitos $|I - h\lambda A| = 1$, luego R(z) es un polinomio de grado *s*+1, lo que implica que la región de estabilidad es finita.

2.4. Construcción de métodos RKHB a paso fijo

En esta sección se van a construir diferentes métodos RKHB de órdenes 3 y 4 que utilizarán una etapa menos que el orden que se quiere alcanzar.

2.4.1. Métodos de orden 3 con 2 etapas

Para orden 3 se tiene

$$A = \begin{bmatrix} 0 & 0 \\ a_{21} & 0 \end{bmatrix}, \mathbf{c} = \begin{bmatrix} 0 \\ c_2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix}, \Gamma = \begin{bmatrix} 0 \\ \gamma_2 \end{bmatrix} \mathbf{y} \gamma_0.$$

Las condiciones para orden 3 quedan reflejadas en la Tabla 2.1 y

Tabla 2.1 Condiciones de orden 3

ho(au)	τ	F(au)	$b^ op \Phi(au)$
1	•	f	$\mathbf{b}^{\top}\mathbf{e} = 1$
2		f'f	$\mathbf{b}^{\top}\mathbf{c} + \gamma_0 = 1/2$
	•		
31		f'f'f	$\mathbf{b}^{\top}(A\mathbf{c}+\Gamma)=1/6$
32		f''(f,f)	$\mathbf{b}^{\top}\mathbf{c}^2 = 1/3$

para estos métodos se va a utilizar la hipótesis simplificadora básica

$$A\mathbf{e} = \mathbf{c},\tag{2.2}$$

condición para que los problemas autónomo y no autónomo sean equivalentes, que implica que $a_{21} = c_2$, por tanto se tienen 5 parámetros libres, c_2 , b_1 , b_2 , γ_2 y γ_0 . En cuanto a las condiciones de orden

$$\begin{array}{ccc} \hline 1 & b^T \Phi(\bullet) = \frac{1}{\gamma(\bullet)} \Rightarrow b^T \mathbf{e} = 1 \Rightarrow b_1 + b_2 = 1 \\ \hline 2 & b^T \Phi(\bullet) = \frac{1}{\gamma(\bullet)} \Rightarrow b^T c + \gamma_0 = \frac{1}{2} \Rightarrow b_1 c_1 + b_2 c_2 = \frac{1}{2} \\ \hline 3_1 & b^T (\mathrm{Ac} + \Gamma) = \frac{1}{6} \Rightarrow b^T \Gamma = \frac{1}{6} \Rightarrow b_2 \gamma_2 = \frac{1}{6} \\ \hline 3_2 & b^T c^2 = \frac{1}{3} \Rightarrow b_2 c_2 = \frac{1}{3} \end{array}$$

Resolviendo las ecuaciones anteriores

$$b_2 = \frac{1}{3c_2^2}, \ \gamma_2 = \frac{c_2^2}{2}, \ \gamma_0 = \frac{1}{2} - \frac{1}{3c_2}y \ b_1 = 1 - \frac{1}{3c_2^2},$$

por tanto, se dispondrá de una familia uniparamétrica de métodos RKHB de 2 etapas y de orden 3, dependiendo de $c_2 \in (0, 1]$.

2.4.2. Métodos de orden 4 con 3 etapas

Actuando como en la construcción de métodos de orden 3 optimales y entendiendo por optimal el método de orden máximo con el mínimo número posible de etapas. En este caso, los parámetros

Estudio y construcción de métodos de integración numérica de tipo Runge-Kutta-Hermite-Birkhoff

disponibles son:

$$A = \begin{bmatrix} 0 & 0 & 0 \\ a_{21} & 0 & 0 \\ a_{31} & a_{32} & 0 \end{bmatrix}, c = \begin{bmatrix} 0 \\ c_2 \\ c_3 \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}, \Gamma = \begin{bmatrix} 0 \\ \gamma_2 \\ \gamma_3 \end{bmatrix} y \gamma_0,$$

teniendo, junto a las condiciones de orden 3 de la Tabla 2.1, las condiciones adicionales para orden 4 de la Tabla 2.2:

ho(au)	τ	F(au)	$b^{\top} \Phi(\tau)$
41		$f^{\prime\prime\prime}(f,f,f)$	$\mathbf{b}^{\top}\mathbf{c}^3 = 1/4$
42		f''(f'f,f)	$\mathbf{b}^{\top}(\mathbf{c}\cdot(A\mathbf{c}+\Gamma))=1/8$
43		f'f''(f,f)	$\mathbf{b}^{\top} A \mathbf{c}^2 = 1/12$
44		f'f'f'f	$\mathbf{b}^{\top} A (A\mathbf{c} + \Gamma) = 1/24$

Tabla 2.2 Condiciones de orden 4

En estos métodos, además de la hipótesis (2.2), se utilizará la hipótesis simplificadora adicional

$$Ac + \Gamma = \frac{c^2}{2}.$$
 (2.3)

Utilizando estas hipótesis simplificadoras conseguimos reducir el número de las condiciones, pasando de las 8 iniciales a tan solo 5, pues $3_2 \equiv 3_1$, $4_4 \equiv 4_3$ y $4_2 \equiv 4_1$. De (2.3) se cumple que $\gamma_2 = \frac{c_2^2}{2}$ y $\gamma_3 = \frac{c_3^2}{2} - a_{32}c_2$ y de (2.2) $a_{21} = c_2$ y $a_{31} = c_3 - a_{32}$.

Las ecuaciones de orden resultantes son:

- $\begin{bmatrix} 1 \end{bmatrix} b_1 + b_2 + b_3 = 1$
- 2 $b_2c_2 + b_3c_3 + \gamma_0 = \frac{1}{2}$
- 31 $b^T(Ac + \Gamma) = \frac{1}{6} \equiv b^T c^2 = \frac{1}{3}$, debido a la hipótesis simplificadora (2.3)
- $\begin{array}{|c|c|c|c|c|}\hline 3_2 & b_2 c_2^2 + b_3 c_3^2 = \frac{1}{3} \end{array}$
- 4₄ $b^T A(Ac + \Gamma) = \frac{1}{24} \equiv b^T Ac^2 = \frac{1}{12}$, por la hipótesis simplificadora (2.3)

$$\begin{vmatrix} 4_3 \\ b_3 a_{32} c_2^2 = \frac{1}{12} \end{vmatrix}$$

$$\begin{array}{|c|c|c|c|c|}\hline \hline & 4_2 \\ \hline & b^T (c \cdot (Ac + \Gamma)) = \frac{1}{8} \equiv b^T c^3 = \frac{1}{4}, \text{ debido a (2.3)} \\ \hline \hline & 4_1 \\ \hline & b_2 c_2^3 + b_3 c_3^3 = \frac{1}{4} \end{array}$$

Resolviendo el sistema no lineal de ecuaciones anterior se tiene

Tabla 2.3 Métodos de orden 4

$$\begin{array}{ll} a_{21} = c_2 & a_{31} = c_3 - a_{32} \\ \gamma_2 = \frac{c_2^2}{2} & \gamma_3 = \frac{c_3^2}{2} - a_{32}c_2 \\ b_1 = 1 - (b_2 + b_3) & \gamma_0 = \frac{1}{2} - (b_2c_2 + b_3c_3) \\ b_2 = \frac{\frac{1}{3} - b_3c_3^2}{c_2^2} & a_{32} = \frac{1}{12c_2^2b_3} \\ b_3 = \frac{4c_2 - 3}{12c_3^2(c_2 - c_3)} & \end{array}$$

De las ecuaciones de la Tabla 2.3 se tiene una familia biparamétrica de métodos optimales de orden 4 con 3 etapas, tomando $c_2, c_3 \in (0, 1], c_2 \neq c_3, c_2 \neq \frac{3}{4}$.

2.5. Construcción de pares encajados

En esta sección se van a construir varios pares encajados de órdenes distintos utilizando los cálculos de la sección anterior. Además, se desarrollará un par encajado de órdenes 5(4) para competir con el clásico DOPRI5(4) [CMR].

2.5.1. Par encajado 4(3) con 3 etapas

Puesto que lo interesante en los pares encajados es minimizar el TPEL de la aproximación de orden alto, se calculan los coeficientes que acompañan a las diferenciales elementales de orden 5 dados en la Tabla 2.4 y, con ellos, la norma euclídea que se minimizará.

$$||TPEL||^{2} = \frac{5c_{2}^{2}(287 + 184c_{3}(-3 + 2c_{3})) + 27(48 + 13c_{3}(-8 + 5c_{3})) - 6c_{2}(356 + 23c_{3}(-31 + 20c_{3}))}{207360}$$

Minimizando la expresión anterior para $c_2 y c_3$, se obtienen que los parámetros que optimizan el método son $c_2 = 0.59216$ y $c_3 = 0.811228$, resultando ||TPEL|| = 1.21×10^{-2} .

Por simplicidad a la hora de calcular coeficientes se tomará $c_2 = \frac{3}{5}$ y $c_3 = \frac{4}{5}$, puesto que el TPEL no varía demasiado y los coeficientes del método son números racionales sencillos.

τ	F(au)	$(rac{1}{\gamma(au)} - \mathbf{b}^ op \mathbf{\Phi}(au)) rac{1}{\sigma(au)}$
	$f^{\prime\prime\prime\prime}(f,f,f,f)$	$(\tfrac{1}{5} - \mathbf{b}^\top \mathbf{c}^4) \tfrac{1}{24}$
	$f^{\prime\prime\prime}(f,f,f^{\prime}f)$	$(\frac{1}{10} - \mathbf{b}^{\top}(\mathbf{c}^2 \cdot (A\mathbf{c} + \Gamma)))\frac{1}{2}$
	f''(f,f''(f,f))	$(\frac{1}{15} - \mathbf{b}^{\top}(\mathbf{c} \cdot A\mathbf{c}^2))\frac{1}{2}$
	$f^{\prime\prime}(f^{\prime}f,f^{\prime}f)$	$\left(\frac{1}{20} - \mathbf{b}^{\top} (A\mathbf{c} + \Gamma) \cdot (A\mathbf{c} + \Gamma)\right) \frac{1}{2}$
	$f^{\prime\prime}(f,f^{\prime}f^{\prime}f)$	$\frac{1}{30} - \mathbf{b}^{\top} (\mathbf{c} \cdot A(A\mathbf{c} + \Gamma))$
	f'f'''(f,f,f)	$(\frac{1}{20} - \mathbf{b}^{\top} A \mathbf{c}^3) \frac{1}{6}$
	f'f''(f,f'f)	$\frac{1}{40} - \mathbf{b}^{\top} A (\mathbf{c} \cdot (A\mathbf{c} + \Gamma))$
	f'f'f''(f,f)	$(\frac{1}{60} - \mathbf{b}^{\top} A^2 \mathbf{c}^2) \frac{1}{2}$
	f'f'f'f'f	$\frac{1}{120} - \mathbf{b}^{\top} A^2 (A\mathbf{c} + \Gamma)$

Tabla 2.4 Términos del TPEL de orden 5

 $\begin{array}{c|c} 0 & 0 \\ 3 & 3 \end{array}$

Con esto, el tablero de Butcher para el método de orden 4 optimizado es:

$$\gamma_0 = \frac{\frac{7}{5}}{\frac{4}{5}} \frac{\frac{3}{5}}{\frac{28}{135}} \frac{\frac{16}{27}}{\frac{653}{1728}} \frac{\frac{25}{25}}{\frac{25}{108}} \frac{\frac{25}{64}}{\frac{653}{1728}}$$

у

El método de orden bajo del par encajado tendrá que verificar las condiciones de orden dadas en la Tabla 2.1 con A, c y Γ fijados.

$$\begin{array}{ccc} \hline 1 & \widehat{\mathbf{b}}^{T} \mathbf{e} = \widehat{\mathbf{b}}_{1} + \widehat{\mathbf{b}}_{2} + \widehat{\mathbf{b}}_{3} = 1 \\ \\ \hline 2 & \widehat{\mathbf{b}}^{T} \mathbf{c} + \widehat{\gamma}_{0} = \widehat{\mathbf{b}}_{2} \mathbf{c}_{2} + \widehat{\mathbf{b}}_{3} \mathbf{c}_{3} + \widehat{\gamma}_{0} = \frac{1}{2} \\ \\ \hline 3_{1} & \widehat{\mathbf{b}}^{T} \mathbf{c}^{2} = \widehat{\mathbf{b}}_{2} \mathbf{c}_{2}^{2} + \widehat{\mathbf{b}}_{3} \mathbf{c}_{3}^{2} = \frac{1}{3} \end{array}$$

resultando

$$\widehat{b}_2 = \frac{\frac{1}{3} - \widehat{b}_3 c_3^2}{c_2^2}, \ \ \widehat{\gamma}_0 = \frac{1}{2} - (\widehat{b}_2 c_2 + \widehat{b}_3 c_3) \ \ y \ \ \widehat{b}_1 = 1 - (\widehat{b}_2 + \widehat{b}_3).$$

Autor: Diego Navalón Casalod

La condición 3₂ no es necesario imponerla debido a la hipótesis simplificadora (2.3). Así, tomando $\hat{b}_3 = \frac{1}{10}$ se obtiene el método encajado dado por:

$$\widehat{\mathbf{b}} = \left(\frac{41}{270}, \frac{101}{135}, \frac{1}{10}\right) \quad \mathbf{y} \quad \widehat{\gamma}_0 = \frac{-13}{450}$$

2.5.2. Par encajado 5(3) con 4 etapas

Para la construcción de este par encajado se utilizará un método óptimo de orden 5 con 4 etapas y, como antes, se construirá el de menor orden manteniendo constante la matriz A y los vectores c y Γ .

En la obtención de los coeficientes del método de orden 5 que se pueden deducir de la Tabla 2.4, se ha utilizado, además de las condiciones de orden vistas en la Sección 2.4 y las hipótesis (2.2) y (2.3), la nueva hipótesis simplificadora

$$Ac^{2} = \frac{c^{3}}{3} - \frac{c_{2}^{3}}{3}\mathbf{e}_{2},$$
(2.4)

de la que se puede deducir que $b_2 = 0$, ya que si multiplicamos (2.4) por b^{\top} queda $b^{\top}Ac^2 = b^{\top}\frac{c^3}{3} - \frac{c_2^3}{3}b^{\top}\mathbf{e}_2$, de donde $b_2c_2 = 0 \rightarrow b_2 = 0$, pues $b^{\top}Ac^2 = b^{\top}\frac{c^3}{3} = \frac{1}{12}$. Además, destacar que el término $\frac{c_2^3}{3}\mathbf{e}_2$ es necesario restarlo pues, si no estuviese dicho término tendría que ser $\frac{c_2^3}{3} = 0$, luego $c_2 = 0$ que es imposible, pues la primera y segunda etapa serían iguales.

Sin entrar en detalles técnicos de la construcción y optimización del método por su extensión, solo se ofrece el TPEL (ver Anexo I) y los coeficientes del método de orden alto, calculando después los parámetros del método de orden bajo.

$$\begin{split} \text{TPEL=} & \left(-\frac{1}{720},-\frac{1}{720},\frac{1}{240}(2c_2-1),\frac{1}{720}(2c_2-1),-\frac{1}{1800},-\frac{1}{1800},-\frac{1}{2400},-\frac{1}{1200},-\frac{1}{7200},\frac{1}{720},\frac{1}{720},\frac{1}{720},\frac{1}{720},\frac{1}{240},\frac{1}{2400},\frac{1}{2400},\frac{1}{3600},\frac{1}{3600},\frac{1}{3600},\frac{1}{3600},\frac{1}{3600},\frac{1}{3600},\frac{1}{3600},\frac{1}{3600},\frac{1}{3600},\frac{1}{3600},\frac{1}{3600},\frac{1}{3600},\frac{1}{3600},\frac{1}{3600},\frac{1}{3600},\frac{1}{3600},\frac{1}{36000}\right), \text{ donde } \\ & \text{|TPEL|=} 3.13 \times 10^{-3} + \frac{(1-2c_2)^2}{25920}, \text{ que se minimiza cuando } c_2 = \frac{1}{2}. \end{split}$$

Por tanto, el método de orden 5 queda de la siguiente manera

y

$$\gamma_0 = \frac{1}{36}, \Gamma = \left(0, \frac{1}{8}, \frac{9}{250}, \frac{1}{18}\right)^{\top}.$$

Procediendo como antes para la obtención del método de orden 3:

$$\boxed{1} \quad \widehat{\mathbf{b}}^T \mathbf{e} = \widehat{\mathbf{b}}_1 + \widehat{\mathbf{b}}_2 + \widehat{\mathbf{b}}_3 + \widehat{\mathbf{b}}_4 = 1,$$

2.6. Métodos continuos

$$\begin{array}{|c|c|c|c|c|} \hline 2 & \widehat{b}^T c + \widehat{\gamma}_0 = \widehat{b}_2 c_2 + \widehat{b}_3 c_3 + \widehat{b}_4 c_4 + \widehat{\gamma}_0 = \frac{1}{2} \\ \hline \hline 3_1 & \widehat{b}^T c^2 = \widehat{b}_2 c_2^2 + \widehat{b}_3 c_3^2 + \widehat{b}_4 c_4^2 = \frac{1}{3}, \end{array}$$

que se puede reescribir como:

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & c_3 & 1 \\ 0 & c_3^2 & 1 \end{bmatrix} \begin{bmatrix} \hat{b_1} \\ \hat{b_3} \\ \hat{\gamma_0} \end{bmatrix} = \begin{bmatrix} 1 - \hat{b_2} - \hat{b_4} \\ \frac{1}{2} - \hat{b_2}c_2 - \hat{b_4}c_4 \\ \frac{1}{3} - \hat{b_2}c_2^2 - \hat{b_4}c_4^2 \end{bmatrix}$$

quedando $\hat{b}_2 y \hat{b}_4$ como parámetros libres. Tomando $\hat{b}_2 = 0 y \hat{b}_4 = \frac{1}{10}$ se obtiene el método de orden bajo:

$$\widehat{\boldsymbol{b}}^{\top} = (\frac{34}{135}, 0, \frac{35}{54}, \frac{1}{10}) \quad \boldsymbol{y} \quad \widehat{\boldsymbol{\gamma}}_0 = \frac{1}{90}$$

2.5.3. Par encajado 5(4) con 5 etapas

En la construcción de este par encajado se trabajará, al igual que en el par encajado anterior, con todo el conjunto de hipótesis simplificadoras (2.2)-(2.4).

En la construcción de la solución de orden alto, se utilizan hasta las condiciones de orden 5 que se pueden deducir de la Tabla 2.4 y las de orden 6 necesarias para el cálculo del TPEL (ver Anexo I).

Además, como se ha hecho en el par encajado anterior, no se entrará en detalles de la construcción y optimización, dando únicamente los parámetros del par encajado.

$$\operatorname{Asi} \mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{8} & 0 & 0 & 0 & 0 \\ \frac{-565 - 241r}{150} & \frac{64(5 + 2r)}{75} & 0 & 0 & 0 \\ \frac{965 - 299r}{150} & \frac{32(-565 + 199r)}{2175} & \frac{69 - 30r}{29} & 0 & 0 \\ \frac{-37}{3} + 18r & \frac{-32(-55 + 63r)}{87} & \frac{-545 + 271r}{58} & \frac{5 + r}{2} & 0 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1\\ \frac{1}{12}\\ \frac{5}{12}\\ \frac{1}{12} \end{bmatrix}, \mathbf{c} = \begin{bmatrix} 0\\ \frac{1}{12}\\ \frac{1}{12} \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 0\\ \frac{5}{132}\\ \frac{5}{12}\\ \frac{1}{12} \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 0\\ \frac{1}{12}\\ \frac{5}{12}\\ \frac{1}{12} \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 0\\ \frac{5}{12}\\ \frac{5}{12}\\ \frac{1}{12} \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 0\\ \frac{1}{12}\\ \frac{1}{12} \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 0\\ \frac{1}{12$$

Lo que constituye el par encajado RKHB5(4), que se utilizará en el capítulo siguiente para comparar con el DOPRI5(4).

2.6. Métodos continuos

Por último, en esta sección se trata de aproximar la solución exacta del PVI en cualquier punto interior del intervalo $[t_n, t_n + h]$ sin coste computacional adicional mediante un método continuo. Para ello, utilizando la aproximación de orden 5 del par encajado, la aproximación a las derivadas, que supuesto que son suficientemente regulares, en particular de Lipchitz, que también serán del mismo

orden, por lo que se puede construir un polinomio de interpolación de grado ≤ 5 con los datos existentes. Si se quisiera obtener un polinomio de interpolación de orden 5 utilizando el Dopri 5(4) se necesitarian dos evaluaciones adicionales de la función derivada en el interior del intervalo $[t_n, t_n + h]$.

Utilizando el método de diferencias divididas de Newton se obtiene fácilmente el polinomio de interpolación para el intervalo $[t_n, t_n + h]$ de grado ≤ 5 :

$$p_{n}(t_{n} + \theta h) = (1 - \theta)^{3}(6\theta^{2} + 3\theta + 1)y_{n} +h(1 - \theta)^{3}\theta(3\theta) + 1)f(t_{n}, y_{n}) +\frac{1}{2}h^{2}(1 - \theta)^{3}\theta^{2}y_{n}'' +\theta^{3}(3\theta(2\theta - 5) + 10)y_{n+1} +h\theta^{3}((7 - 3\theta)\theta - 4)f(t_{n+1}, y_{n+1}) +\frac{1}{2}h^{2}(1 - \theta)^{3}\theta^{2}y_{n+1}''$$

Habiendo conseguido así un método continuo en el intervalo $[t_n, t_{n+1}], \forall \theta \in [0, 1].$

Capítulo 3

Experimentos Numéricos

Este último capítulo de la memoria está dedicado a los experimentos numéricos realizados con los métodos desarrollados en el capítulo anterior. Para ello se han considerado pares encajados de métodos Runge-Kutta de características similares y se han implementado en lenguaje Python. El criterio empleado en la comparación ha sido la norma infinito del máximo error global de la solución numérica en todo el intervalo de integración y el costo computacional se ha medido con el número de evaluaciones de función derivada (primera y segunda).

Destacar que en todos los experimentos numéricos realizados siempre se avanza con la solución de orden más alto (extrapolación local).

	orden	etapas	n ^o eval. paso acep.	n ^o eval. paso rech.	TPEL	IEA
DOPRI4(3)[DLMP]	4	4	5	5	5.86×10^{-4}	-3.15
<i>OWZ</i> 4(3)[OZ]	4	5	5	4	3.16×10^{-3}	-2.87
<i>RKHB</i> 4(3)	4	3	4	2	1.21×10^{-2}	-2.78
<i>RKHB</i> 5(3)	5	4	5	3	3.13×10^{-3}	-3.21

Tabla 3.1 Pares encajados 4(3), 5(3)

	etapas	n ^o eval. paso acep.	n ^o eval. paso rech.	TPEL	IEA					
DOPRI5(4)	6	6	6	3.99×10^{-4}	-3.30					
<i>RKHB</i> 5(4)	5	6	4	$2.59 imes 10^{-4}$	-3.72					

Tabla 3.2 Pares encajados 5(4)

Los problemas considerados en esta memoria son: **Problema 1.**

El primer problema considerado ha sido el A3 del paquete DETEST, que viene dado por

$$\begin{cases} y' = \cos(t)y, & t \in [0, 20], \\ y(0) = 1, \end{cases}$$

cuya solución exacta es

$$\mathbf{y}(t) = \exp(\sin(t)).$$

Problema 2.

El segundo problema que se ha considerado ha sido la ecuación de Duffing, dada por

$$\begin{cases} y'' + (\lambda^2 + k^2)y = 2k^2y^3, & t \in [0, 20], \\ y(0) = 0, & y'(0) = \lambda, \end{cases}$$

con k = 0.03 y $\lambda = 1$. La solución analítica de este problema viene dada por:

$$\mathbf{y}(t) = \mathrm{sn}\left(\lambda t, \left(\frac{k}{\lambda}\right)^2\right).$$

donde sn es la función elíptica de Jacobi.

Problema 3.

El último problema en esta comparativa es el del sólido rígido, dado por

$$\begin{cases} y_1' = (\alpha - \beta)y_2y_3\\ y_2' = (1 - \alpha)y_3y_1\\ y_3' = (\beta - 1)y_1y_2\\ y(0) = (0, 1, 1)^\top, \end{cases}$$

con $\alpha = 1 + \frac{1}{\sqrt{1.51}}$, $\beta = 1 - \frac{0.51}{\sqrt{1.51}}$, cuya solución exacta viene dada por:

$$\mathbf{y}(t) = \left(\sqrt{1.51}\mathrm{sn}(t, 0.51), \mathrm{cn}(t, 0.51), \mathrm{dn}(t, 0.51)\right)^{\top}$$

donde sn, cn y dn son las funciones elípticas de Jacobi.





Estudio y construcción de métodos de integración numérica de tipo Runge-Kutta-Hermite-Birkhoff







Fig. 3.3 Problema 3 Pares encajados 4(3) y 5(3)

Autor: Diego Navalón Casalod



Fig. 3.5 Problema 2 Pares encajados 5(4)

Estudio y construcción de métodos de integración numérica de tipo Runge-Kutta-Hermite-Birkhoff



Pares 4(3), 5(3)

En este caso se han utilizado tolerancias 10^{-i} , i = 2, ..., 6 y, de las Figuras 3.1, 3.2 y 3.3, puede deducirse que en general el método más eficiente entre los considerados es el RKHB5(3), lo que supone que una buena optimización de la solución numérica de mayor orden es fundamental en el desarrollo de estos esquemas numéricos. Por lo tanto, parece razonable invertir una etapa adicional y lograr ó conseguir un orden mayor.

Pares 5(4)

Ahora, el rango de tolerancias empleado es $tol = 10^{-i}$, i = 3, ..., 8 y de las Figuras 3.4, 3.5 y 3.6 se aprecia que la eficiencia del nuevo método RKHB5(4) desarrollado en esta memoria es claramente superior al ampliamente utilizado DOPRI5(4) en todos los problemas considerados.

3.1. Conclusiones

En esta memoria se ha llevado a cabo la construcción de pares encajados de métodos del tipo Runge-Kutta-Hermite-Birkhoff para la integración numérica de problemas donde el cálculo de la derivada segunda, bien analíticamente o usando diferenciadores automáticos, sea similar al de la función derivada. Este estudio se ha basado principalmente en la minimización de las componentes del término principal del error local. Se han construido tres nuevos pares encajados de órdenes 4(3), 5(3) y 5(4) y se han comparado con integradores estándar ampliamente utilizados en la literatura científica. Los experimentos numéricos llevados a cabo muestran que el nuevo par RKHB5(4) presenta una clara mejora en eficiencia respecto al ampliamente utilizado Dopri5(4).

Bibliografía

- [BNV] Bozic V.; Nguyen-Ba T.; Vaillancourt R. A three-stage, VSVO, Hermite-Birkhoff-Taylor, ODE solver, Applied Mathematics and Computation 216 (2010), pp. 598-610
- [B] Butcher J.C., Numerical Methods for Ordinary Differential Equations 2nd Edition, 2008 John Wiley & Sons, Ltd
- [CMR] Calvo, M.; Montijano, J.I.; Rández, L. Curso de Análisis Numérico. Métodos Runge-Kutta para la resolución numérica de ecuaciones diferenciales ordinarias, Prensas Universidad de Zaragoza, Zaragoza 1990
- [DLMP] Dormand, J. R.; Lockyer, M. A.; Mcgorrigan, N. E.; Prince, P. J. Global Error Estimation with Runge-Kutta Triples, Computers Math. Applic. Vol. 18, No. 9, pp. 835-846, 1989
- [OZ] Owren, B.; Zennaro, M. Derivation of efficient, continuous, explicit Runge-Kutta methods, SIAM J. Sci. Statist. Comput. 13 (1992), 6, 1488-1501

Anexo I

En este anexo se adjuntan los términos del TPEL de orden 6 utilizados en los pares encajados RKHB 5(3) y 5(4).