

UNIVERSIDADE DE SÃO PAULO

Análise comparativa de métodos  
numéricos de equações  
algébrico-diferenciais

NEYVA MARIA LOPES ROMEIRO  
ANTONIO CASTELO FILHO

Nº 17

---

NOTAS

---



Instituto de Ciências Matemáticas de São Carlos



Instituto de Ciências Matemáticas de São Carlos

ISSN - 0103-2577

Análise comparativa de métodos  
numéricos de equações  
algébrico-diferenciais

NEYVA MARIA LOPES ROMEIRO  
ANTONIO CASTELO FILHO

Nº 17

NOTAS DO ICMSC  
Série Computação

São Carlos  
Dez./ 1994

**Análise Comparativa de Métodos  
Numéricos de Equações  
Algébrico-Diferenciais**

**Neyva Maria Lopes Romeiro**

**Antonio Castelo Filho**

**ICMSC - Instituto de Ciências Matemáticas de São Carlos, da  
Universidade de São Paulo**

**São Carlos - 1994**

# Conteúdo

---

---

<b>1</b>	<b>Introdução</b>	<b>01</b>
<b>2</b>	<b>Teoria das EAD</b>	<b>03</b>
2.1	Introdução	03
2.2	EAD Lineares com Coeficientes Constantes	04
2.3	EAD Lineares com Coeficientes Variando com o Tempo	05
2.4	EAD não Lineares	07
<b>3</b>	<b>Método BDF</b>	<b>09</b>
3.1	Introdução	09
3.2	Métodos de Passos Múltiplos	09
3.3	Convergência do Método BDF	10
3.3.1	Sistemas Semi Explícitos	11
3.3.2	Sistemas Completamente Implícitos	11
<b>4</b>	<b>Métodos de Runge-Kutta Implícito</b>	<b>13</b>
4.1	Introdução	13
4.2	Sistemas Lineares com Coeficientes Constantes	14
4.3	Sistemas não Lineares	21

<b>5</b>	<b>Algoritmos</b>	<b>23</b>
5.1	Introdução	23
5.2	Algoritmos e Estratégias no DASSL	23
5.2.1	Solução do Sistema não Linear	26
5.3	Algoritmos e Estratégias no RADAU5	27
5.3.1	Solução do Sistema não Linear	27
5.3.2	Valor Inicial para a Iteração de Newton	29
<b>6</b>	<b>Resultados Numéricos</b>	<b>31</b>
6.1	Introdução	31
6.2	Equação do Pêndulo	31
6.3	Equação de Van der Pol	33
6.4	Equação da Esfera	34
6.5	Equação Teste	36
<b>7</b>	<b>Conclusão</b>	<b>39</b>
	<b>Bibliografia</b>	<b>41</b>

# 1 Introdução

---

---

Chamaremos de Equações Algébrico-Diferenciais (EAD) a um sistema que envolve equações algébricas e/ou diferenciais. Este tipo de equação surge naturalmente em muitas aplicações, tais como:

- Equações do Movimento Descrito por Pêndulos;
- Simulação de Trabalhos Elétricos;
- Engenharia de Controle;
- Dinâmica de Fluidos;
- Sistemas Mecânicos Restritos e outros.

As EAD apresentam dificuldades numéricas e analíticas quando não ocorrem por meio de uma Equação Diferencial Ordinária (EDO) ou Equações Puramente Diferenciais.

No Capítulo 2 definiremos o conceito de índice das EAD. Geralmente as EAD de índices elevados tornam o problema, numericamente, mais difícil. Gostaríamos de ressaltar que este trabalho aborda apenas sistemas de índice zero ou um.

Temos ainda outras definições e resultados importantes, neste capítulo, que sem estes não poderíamos concluir nosso trabalho.

No Capítulo 3 estudaremos uma das classes dos métodos de passos múltiplos, conhecido como diferenciação para trás (*backward differentiation formulas*) ou método BDF. Estes métodos são, amplamente, usados para a integração de equações diferenciais stiff.

No Capítulo 4 estudaremos os métodos de Runge-Kutta Implícito (RKI). Recentemente tem sido desenvolvido muitos estudos sobre este método devido as seguintes razões: a combinação de ordem elevada com boa estabilidade; permite uma seleção simples do passo dado; é um método de início próprio; e outras.

No Capítulo 5 analisaremos o desenvolvimento de algoritmos para resolver sistemas de EAD usando os métodos BDF e RKI e, em particular, os algoritmos utilizados em dois códigos, já consagrados, o DASSL e o RADAU5.

O código DASSL utiliza o método BDF. Este código é indicado para resolver sistemas completamente implícito de EAD de índice zero ou um. Podemos, também, aplicar este código em EAD semi explícita, como veremos no Capítulo 6.

O código RADAU5 utiliza o método de RKI em sistemas semi explícitos. Este código, também, obtém resultados satisfatórios para sistemas de EAD semi explícito de índice maior ou igual a dois.

No Capítulo 6 apresentaremos resultados numéricos de alguns exemplos de sistemas de EAD de índice zero ou um. Os resultados destes exemplos são obtidos usando os códigos DASSL e RADAU5.

No Capítulo 7 faremos uma conclusão do trabalho de um modo geral, assim como as comparações dos resultados obtidos no Capítulo 6, avaliando a estabilidade dos métodos BDF e Runge-Kutta Implícito.

# 2 Teoria das EAD

---

---

## 2.1 Introdução

Neste capítulo desenvolveremos uma teoria necessária para compreendermos algumas das diferenças entre Equações Algébrico-Diferenciais e Equações Diferenciais Ordinárias. Esta teoria encontra-se em [BRENAN - 89].

Apresentaremos, inicialmente, algumas definições necessárias para a compreensão de resultados posteriores.

**Definição 2.1.1** *Um Homeomorfismo entre  $X$  e  $Y$  é uma aplicação  $\varphi: X \rightarrow Y$  contínua, inversível com inversa contínua. Se  $\varphi$  e  $\varphi^{-1}$  são de classe  $C^k$ , dizemos que  $\varphi$  é um Homeomorfismo de classe  $C^k$  entre  $X$  e  $Y$ . Um Homeomorfismo de classe  $C^k$ , com  $k \geq 1$  é dito Difeomorfismo de classe  $C^k$ .*

**Definição 2.1.2**  *$M \subset R^k$  é uma Variedade de classe  $C^r$  de dimensão  $n$  se para todo  $x \in M$ , existem vizinhanças abertas  $U \subset R^k$  de  $x$ ,  $V \subset R^n$  e um Homeomorfismo de classe  $C^r$ ,  $\varphi: V \rightarrow U \cap M$ .*

A função  $y(t)$  é uma solução da EAD não linear geral

$$F(t, y, y') = 0 \tag{2.1.1}$$

tem um intervalo  $I$  se  $y$  é continuamente diferenciável em  $I$  e satisfaz (2.1.1) para todo  $t \in I$ .

**Definição 2.1.3** *Seja  $I$  um intervalo aberto de  $R$ ,  $\Omega$  um conjunto conexo aberto de  $R^{2m+1}$  e  $F$  uma função diferenciável de  $\Omega$  em  $R^m$ . Então a EAD (2.1.1) é solúvel em  $I$  sobre  $\Omega$  se existe uma família  $r$ -dimensional de soluções  $\phi(t, c)$  definidas em um conjunto conexo aberto  $I \times \tilde{\Omega}$ ,  $\tilde{\Omega} \subset R^r$  tal que*

- i)  $\phi(t, c)$  está definida em  $I$  para cada  $c \in \tilde{\Omega}$ ;
- ii)  $(t, \phi(t, c), \phi'(t, c)) \in \Omega$  para  $(t, c) \in I \times \tilde{\Omega}$ ;



- iii) Se  $\psi(t)$  é qualquer outra solução e  $(t, \psi(t, c), \psi'(t, c)) \in \Omega$ , então  $\psi(t) = \phi(t, c)$  para algum  $c \in \tilde{\Omega}$ ;
- iv) o gráfico de  $\phi$  como uma função de  $(t, c)$  é uma variedade  $(r+1)$ -dimensional.

**Definição 2.1.4** Denotamos o índice de uma EAD por  $\nu$ , onde  $\nu$  é o número mínimo de vezes que a equação (2.1.1), ou parte dela, deve ser diferenciada em relação a  $t$  para determinar  $y'$  como uma função contínua de  $y$  e  $t$ .

## 2.2 EAD Lineares com Coeficientes Constantes

Consideraremos a EAD linear com coeficiente constante

$$Ax' + Bx = f, \quad (2.2.1)$$

onde  $A$  e  $B$  são matrizes  $m \times m$ .

**Definição 2.2.1** Se  $\lambda$  é um parâmetro complexo, então  $(\lambda A + B)$  é a matriz pencil.

**Definição 2.2.2** Se  $\det(\lambda A + B)$  não é identicamente nulo como uma função de  $\lambda$ , então a matriz pencil (ou pencil) é regular. Quando o pencil é regular, dizemos que a EAD também é regular.

**Teorema 2.2.1** Suponha que  $(\lambda A + B)$  é um pencil regular. Então existe matrizes não singulares  $P$  e  $Q$  tais que

$$PAQ = \begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix}, \quad PBQ = \begin{bmatrix} C & 0 \\ 0 & I \end{bmatrix}, \quad (2.2.2)$$

onde  $N$  é uma matriz nilpotente com índice de nilpotência  $k$  e  $I$  a matriz identidade. No caso especial em que  $A$  é não singular,  $PAQ = I$  e  $PBQ = C$ . Se  $\det(\lambda A + B)$  é identicamente constante, então (2.2.2) reduz a  $PAQ = N$  e  $PBQ = I$ .

Considerando que as mudanças de coordenadas  $P$  e  $Q$  proporcionadas por este teorema sejam aplicadas à EAD (2.2.1), isto é, considerando  $x = Qy$ , então (2.2.1) torna-se  $AQy' + BQy = f$ . Multiplicando a esquerda por  $P$ , temos

$$PAQy' + PBQy = Pf. \quad (2.2.3)$$

Fazendo  $Pf = [f_1, f_2]^t$  e usando o Teorema 2.2.1, obtemos

$$\begin{bmatrix} I & 0 \\ 0 & N \end{bmatrix} \begin{bmatrix} y_1' \\ y_2' \end{bmatrix} + \begin{bmatrix} C & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$$

ou

$$y_1' + Cy_1 = f_1 \quad (2.2.4a)$$

$$Ny_2' + y_2 = f_2, \quad (2.2.4b)$$

onde (2.2.4a) é uma EDO cuja solução existe para qualquer valor inicial de  $y_1$  e qualquer função  $f_1$  contínua e, (2.2.4b) tem apenas uma solução, pois

$$Ny_2' + y_2 = (ND + I)y_2 = f_2,$$

assim  $y_2 = (ND + I)^{-1} f_2$ , onde  $D = \frac{d}{dt}$ .

Mas

$$(ND + I)^{-1} = (I - (-ND))^{-1} = \sum_{i=0}^{\infty} (-ND)^i = \sum_{i=0}^{k-1} (-ND)^i = \sum_{i=0}^{k-1} (-1)^i N^i D^i,$$

onde  $k$  é o índice de nilpotência de  $N$ , ou seja  $N^k = 0$  e  $N^{k-1} \neq 0$ , assim  $(ND)^i = 0$  para  $i \geq k$ . Desta forma

$$y_2 = (ND + I)^{-1} f_2 = \sum_{i=0}^{k-1} (-1)^i N^i f_2^i \quad (2.2.5)$$

onde  $f_2^i = \frac{d^i f_2}{dt^i}$ .

**Teorema 2.2.2** *A EAD linear com coeficiente constante (2.2.1) é solúvel se, e somente se  $(\lambda A + B)$  é um pencil regular.*

## 2.3 EAD Lineares com Coeficientes Variando com o Tempo

As EAD lineares com coeficientes variando com o tempo têm a forma

$$A(t)x'(t) + B(t)x(t) = f(t) \quad (2.3.1)$$

e está definida no intervalo  $I$ .

**Definição 2.3.1** *O sistema (2.3.1) com  $A$  e  $B$  matrizes  $m \times m$  é solúvel em  $I$  se, para toda  $f$   $m$ -vezes diferenciável, existe pelo menos uma solução para (2.3.1) continuamente diferenciável. Além disso, as soluções são definidas em  $I$  e determinadas para algum valor em  $t \in I$ .*

**Definição 2.3.2** *Suponha que (2.3.1) é solúvel em  $I$ , o índice de (2.3.1) é o número mínimo de diferenciações necessárias para determinar  $x'$  em função de  $x$  e  $t$ .*

**Definição 2.3.3** Se (2.3.1) é uma EAD regular, então o índice local em  $t$ , denotado por  $\nu_i(t)$ , é o índice do pencil  $(\lambda A(t) + B(t))$ .

Considerando a mudança de coordenada em (2.3.1) dada por  $x = Q(t)y$  e também multiplicando por  $P(t)$ , obtemos

$$PAQy' + (PAQ' + PBQ)y = Pf, \quad (2.3.2)$$

onde  $P$  e  $Q$  são não singulares no intervalo de interesse e pelo menos tão suave quanto os coeficientes da EAD.

**Definição 2.3.4** O sistema (2.3.1) é semi explícito se estiver na forma

$$\begin{aligned} x_1' + B_{11}(t)x_1 + B_{12}(t)x_2 &= f_1 \\ B_{21}(t)x_1 + B_{22}(t)x_2 &= f_2. \end{aligned} \quad (2.3.3)$$

A EAD semi explícita (2.3.3) tem índice um se, e somente se  $B_{22}$  é não singular para todo  $t$ .

**Definição 2.3.5** A EAD (2.3.1) está na forma Hessenberg de dimensão  $r$  se pode ser escrita como

$$\begin{bmatrix} I & 0 & & & 0 \\ 0 & I & & & \\ & & I & & \\ & & & I & \\ 0 & & & & 0 \end{bmatrix} \begin{bmatrix} x_1' \\ \vdots \\ x_r' \end{bmatrix} + \begin{bmatrix} B_{11} & * & * & B_{1,r-1} & B_{1r} \\ B_{21} & * & * & B_{2,r-2} & 0 \\ 0 & * & * & * & \\ \cdot & \cdot & * & * & \\ 0 & \cdot & 0 & B_{r,r-1} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_r \end{bmatrix} = \begin{bmatrix} f_1 \\ \vdots \\ f_r \end{bmatrix},$$

onde  $x_i$  são vetores,  $B_{ij}$  são matrizes e o produto  $B_{r,r-1}B_{r-1,r-2}\dots B_{1r}$  é não singular.

**Definição 2.3.6** Dizemos que o sistema (2.3.1) está na forma canônica padrão se

$$\begin{bmatrix} I & 0 \\ 0 & N(t) \end{bmatrix} x' + \begin{bmatrix} C(t) & 0 \\ 0 & I \end{bmatrix} x = f(t), \quad (2.3.4)$$

onde  $N$  é estritamente triangular inferior (ou superior).

**Definição 2.3.7** Uma propriedade de uma matriz é uma propriedade estrutural se todo elemento não nulo da matriz pode ser trocado por variáveis independentes e a propriedade ainda vale em um conjunto aberto de valores destas variáveis.

**Teorema 2.3.1** *Suponha que (2.3.1) é solúvel em  $I$ , então (2.3.1) é analiticamente equivalente a*

$$\begin{bmatrix} I & G \\ 0 & N \end{bmatrix} z' + \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} z = \begin{bmatrix} g \\ h \end{bmatrix}, \quad (2.3.5)$$

onde  $Nz_2' + z_2 = h$  tem solução única para cada função  $h$ . Além disso, existe uma família contável de intervalos abertos disjuntos  $I_i$ , tal que  $\cup I_i$  é denso em  $I$ . O sistema  $Nz_2' + z_2 = h$  é analiticamente equivalente a algum sistema na forma canônica padrão  $Mw' + w = f$ , onde  $M$  é estruturalmente nilpotente.

**Definição 2.3.8** *A EAD é uma cadeia triangular de  $r$ -ésima ordem se tem a forma*

$$\begin{bmatrix} A_{11} & 0 & \cdot & \cdot & 0 \\ A_{21} & A_{22} & \cdot & \cdot & \cdot \\ A_{31} & A_{32} & A_{33} & \cdot & \cdot \\ * & * & * & * & 0 \\ A_{r1} & * & * & * & A_{rr} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_r \end{bmatrix} + \begin{bmatrix} B_{11} & 0 & \cdot & \cdot & 0 \\ B_{21} & B_{22} & \cdot & \cdot & \cdot \\ B_{31} & * & * & \cdot & \cdot \\ * & * & * & * & 0 \\ B_{r1} & * & * & * & B_{rr} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_r \end{bmatrix} = \begin{bmatrix} f_1 \\ \vdots \\ f_r \end{bmatrix}$$

e cada subsistema  $A_{ii}x_i' + B_{ii}x_i = f_i - \sum_{j=1}^{i-1} (A_{ij}x_j' + B_{ij}x_j)$  é de índice zero, ou índice um, ou tem a forma Hessenberg ou está na forma canônica padrão.

## 2.4 EAD não Lineares

Muitas das EAD importantes são não lineares. A teoria usada para EAD linear com coeficiente variando com o tempo é de grande importância para desenvolvermos uma teoria análoga em sistemas não lineares. A não linearidade terá apenas resultados locais.

Consideremos a EAD não linear (2.1.1) e o seguinte sistema de equações

$$\begin{aligned} F_{[0]}(t, y, y') &= 0 \\ F_{[1]}(t, y, y', y'') &= 0 \\ &\vdots \\ F_{[j-1]}(t, y, y', \dots, y^{(j-1)}, y^{(j)}) &= 0 \end{aligned} \quad (2.4.1)$$

onde, por exemplo,  $F_{[1]} = F_y y'' + F_{y'} y' + F_t$ . Podemos escrever (2.4.1), na forma

$$\mathbf{F}_j(t, y, y_j) = 0, \quad (2.4.2)$$

onde  $y_j = [y', \dots, y^{(j)}]^t$ .

**Definição 2.4.1** O índice  $v$  de (2.1.1) é o menor número  $v$  tal que  $F_{v+1}$  determina apenas a variável  $y'$  como uma função contínua de  $y$  e  $t$ .

**Definição 2.4.2** O índice local  $v_i$  de (2.1.1) para  $(\hat{t}, \hat{y}, \hat{y}')$  é o índice do pencil  $\lambda F_y(\hat{t}, \hat{y}, \hat{y}') + F_y(\hat{t}, \hat{y}, \hat{y}')$ .

**Definição 2.4.3** A EAD (2.1.1) possui a forma Hessenberg de dimensão  $r$  se pode ser escrita como

$$\begin{aligned} x_1' &= F_1(x_1, x_2, \dots, x_r, t) \\ x_2' &= F_2(x_1, x_2, \dots, x_{r-1}, t) \\ &\vdots \\ x_i' &= F_i(x_{i-1}, x_i, \dots, x_{r-1}, t), \quad 3 \leq i \leq r-1 \\ &\vdots \\ 0 &= F_r(x_{r-1}, t) \end{aligned}$$

e  $(\partial F_r / \partial x_{r-1})(\partial F_{r-1} / \partial x_{r-2}) \dots (\partial F_2 / \partial x_1)(\partial F_1 / \partial x_r)$  é não singular.

**Definição 2.4.4** A EAD (2.1.1) está na forma canônica padrão de dimensão  $r$  se pode ser escrita como

$$\begin{aligned} x_1' &= F_1(x_1, t) \\ N(x_1, x_2, t)x_2' &= x_2 + h(x_1, t), \end{aligned}$$

onde  $N$  é uma matriz bloco  $r \times r$  estritamente triangular inferior.

**Definição 2.4.5** A EAD (2.1.1) é uma cadeia triangular inferior de ordem  $r$  se tem a forma

$$\begin{aligned} F_1(x_1', x_1, t) &= 0 \\ F_2(x_1', x_2', x_1, x_2, t) &= 0 \\ &\vdots \\ F_r(x_1', \dots, x_r', x_1, \dots, x_r, t) &= 0 \end{aligned}$$

e o sistema definido pelas funções  $F_i$  é uma EDO implícita, ou tem índice um, ou está na forma Hessenberg, ou na forma canônica padrão em relação a variável  $x_i$ .

# 3 Método BDF

---

---

## 3.1 Introdução

Neste capítulo estudaremos as propriedades de convergência e ordem dos métodos de passos múltiplos lineares, em particular, o método BDF aplicado em sistemas de Equações Algébrico-Diferenciais na forma

$$F(t, y, y') = 0, \quad (3.1.1)$$

onde  $F$  e  $y$  tem a mesma dimensão. Assumiremos que  $F$  é continuamente diferenciável.

**Definição 3.1.1**  $f(x, h)$  é  $O(h^p)$  se existe  $k > 0$  tal que  $\|f(x, h)\| \leq kh^p$  para  $h$  suficientemente pequeno.

Veremos que a solução do método BDF de  $k$  passos é  $O(h^k)$  para sistemas de índice um.

## 3.2 Métodos de Passos Múltiplos

Para a EDO não linear

$$y' = g(t, y) \quad (3.2.1)$$

o método de passo múltiplo linear tem a forma

$$\sum_{j=0}^k a_j y_{n-j} - h \sum_{j=0}^k b_j g(t_{n-j}, y_{n-j}) = 0. \quad (3.2.2)$$

O método one-leg para (3.2.2) envolve apenas uma avaliação de  $g$  e tem a forma

$$\frac{1}{h} \sum_{j=0}^k a_j y_{n-j} - g\left(\sum_{j=0}^k b_j t_{n-j}, \sum_{j=0}^k b_j y_{n-j}\right) = 0. \quad (3.2.3)$$

O método BDF pode ser considerado como sendo um método de passo múltiplo linear ou um método one-leg, basta tomar nas equações (3.2.2) e (3.2.3),  $b_0 = 1$  e  $b_j = 0$ ,  $j = 1, 2, \dots, k$ .

A primeira técnica geral para soluções numéricas de EAD, proposta por Gear [GEAR - 71], foi utilizando o método BDF. Este método foi utilizado, inicialmente, em sistemas de equações diferenciais acoplados a equações algébricas

$$\begin{aligned} y' &= f(t, x, y) \\ 0 &= g(t, x, y), \end{aligned} \quad (3.2.4)$$

onde  $y$  e  $g$  têm a mesma dimensão. O método BDF trata as variáveis algébricas e diferenciais do problema (3.2.4) de maneira análoga. Assim este método pode ser estendido para a generalização de qualquer sistema completamente implícito (3.1.1)

O método BDF mais simples é o método de Euler implícito, que consiste em substituir a derivada em (3.1.1) por uma diferença para trás

$$F\left(t_n, y_n, \frac{y_n - y_{n-1}}{h}\right) = 0, \quad (3.2.5)$$

onde  $h = t_n - t_{n-1}$ . O sistema de equações não lineares resultante para  $y_n$  em cada passo é usualmente resolvido pelo método de Newton.

O método BDF de  $k$  passos consiste em substituir  $y'$  pela derivada do polinômio que interpola a solução sobre os pontos  $t_n, t_{n-1}, \dots, t_{n-k}$  avaliados em  $t_n$ . Assim,

$$F\left(t_n, y_n, \frac{\rho y_n}{h}\right) = 0, \quad (3.2.6)$$

onde  $\rho y_n = \sum_{i=0}^k \alpha_i y_{n-i}$  e  $\alpha_i$ ,  $i = 0, 1, \dots, k$ , são os coeficientes do método BDF.

Da teoria de equações diferenciais ordinárias temos que o método BDF de  $k$  passos é estável para  $k \leq 6$  e instável para  $k \geq 7$ , [LAMBERT - 73].

### 3.3 Convergência do Método BDF

Para o resultado de convergência do método BDF aplicado ao sistema de EAD linear com coeficiente constante, temos o seguinte teorema, cuja demonstração encontra-se em [BRENAN - 89], assim como as demais demonstrações deste capítulo.

**Teorema 3.3.1** *O método BDF de  $k$  passos, com  $k < 7$  e  $h$  fixo, aplicado ao sistema de EAD linear com coeficiente constante e índice  $v$  é convergente de ordem  $O(h^k)$  depois de  $(v - 1)k + 1$  passos.*

### 3.3.1 Sistemas Semi Explícitos

Sistemas semi explícitos de índice um são escritos na forma

$$y' = f(t, x, y) \quad (3.3.1a)$$

$$0 = g(t, x, y), \quad (3.3.1b)$$

onde

$$g_y(t, x, y) \quad (3.3.2)$$

é inversível na vizinhança da solução de (3.3.1).

**Definição 3.3.1** *Um método de passo múltiplo linear aplicado a (3.3.1) é dado por*

$$\sum_{j=0}^k a_j y_{n-j} = h \sum_{j=0}^k b_j f(x_{n-j}, y_{n-j}, t_{n-j}) \quad (3.3.3a)$$

$$0 = g(x_n, y_n, t_n). \quad (3.3.3b)$$

Os métodos de passos múltiplos aplicados às EAD semi explícitas de índice um são estáveis e convergente para EAD de mesma ordem, assim como no caso de EDO. Este resultado segue do teorema da função implícita. Especificamente,  $g$  pode ser resolvida para  $y_n$  em termos de  $x_n$ , isto é  $y_n = G(t_n, x_n)$ , então,  $y_n$  pode ser inserida na fórmula de passos múltiplos (3.3.3a) para produzir a mesma equação diferença das obtidas quando aplicamos o método linear de passos múltiplos na EDO

$$y' = f(x, G(x, t), t). \quad (3.3.4)$$

Em princípio, a EAD semi explícita pode ser resolvida por qualquer método de passo múltiplo linear que seja apropriado para a EDO (3.3.4).

### 3.3.2 Sistemas Completamente Implícitos

Sistemas completamente implícitos de índice um, têm a forma

$$\begin{aligned} F(t, y, y') &= 0 \\ y(t_0) &= y_0. \end{aligned} \quad (3.3.5)$$



Assumiremos que  $F$  é uma função suficientemente suave e que, existe uma solução suave  $y(t)$  satisfazendo os valores iniciais dados.

**Definição 3.3.2** A EAD não linear (3.3.5) é de índice uniforme se o índice do sistema de coeficientes constantes  $Aw'(t) + Bw(t) = g(t)$ , onde  $A = F_y(\hat{t}, \hat{y}, \hat{y}')$  e  $B = F_{y'}(\hat{t}, \hat{y}, \hat{y}')$ , é um para todo  $(\hat{t}, \hat{y}, \hat{y}')$  na vizinhança do gráfico da solução e, além disso, se

- i) as derivadas parciais de  $A$  com respeito a  $t, y, y'$  existem e são limitadas na vizinhança da solução;
- ii) o posto de  $A$  é constante na vizinhança da solução.

Para a análise de convergência do método BDF, assumirmos que  $F$  têm a forma

$$F(t, y, y') = \begin{bmatrix} f(t, y, y') \\ g(t, y) \end{bmatrix} \quad (3.3.6)$$

**Lema 3.3.1** Seja  $y$  um vetor  $m$ -dimensional, então se  $F$  é uniforme de índice um na forma (3.3.6), sempre existirá matrizes  $P(t, y(t), y'(t))$  e  $Q(t, y(t), y'(t))$  tais que

$$PAQ = \begin{bmatrix} I_{m_1} & 0 \\ 0 & 0 \end{bmatrix}, \quad PBQ = \begin{bmatrix} C(t, y(t), y'(t)) & 0 \\ 0 & I_{m_2} \end{bmatrix} \quad (3.3.7)$$

onde  $A(t, y(t), y'(t)) = F_y(t, y(t), y'(t))$ ,  $B(t, y(t), y'(t)) = F_{y'}(t, y(t), y'(t))$ ,  $m = m_1 + m_2$  e satisfazem

- i)  $Q(t, y(t), y'(t))$  e  $Q^{-1}(t, y(t), y'(t))$  existem e são limitadas para todo  $(t, y(t), y'(t))$  resolvendo (3.3.5),
- ii)  $Q^{-1}(t_1, y(t_1), y'(t_1)) Q(t_2, y(t_2), y'(t_2)) = I_m + O(t_2 - t_1)$ ,
- iii)  $C(t_1, y(t_1), y'(t_1)) = C(t_2, y(t_2), y'(t_2)) + O(t_2 - t_1)$ .

**Teorema 3.3.2** Considere (3.3.5) sendo uma EAD uniforme de índice um no intervalo  $I = [t_0, t_0 + T]$ . Então a solução numérica de (3.3.5) pelo método BDF de  $k$  passos, com  $k < 7$  e  $h$  fixo, converge com  $O(h^k)$  de exatidão se todos valores iniciais são  $O(h^k)$  e se a iteração de Newton em cada passo é resolvida com  $O(h^{k+1})$  de exatidão.

A demonstração deste teorema encontra-se com mais detalhes em [PIRES - 94].

# 4 Método Runge-Kutta Implícito

---

---

## 4.1 Introdução

Estudaremos aqui, as propriedades de ordem, estabilidade e convergência do método de RKI aplicado a sistemas de EAD na forma

$$F(t, y, y') = 0, \quad (4.1.1)$$

onde os valores iniciais  $y(t_0)$  e  $y'(t_0)$  são dados.

Muitas vezes, estes métodos não alcançam a mesma ordem de exatidão em sistemas algébricos-diferenciais como o fazem em sistemas puramente diferenciais.

Muitos resultados deste capítulo, foram obtidos em [PETZOLD - 86], [HAIRER - 89] e [HAIRER - 91].

O método RKI para solução da EDO

$$y' = f(t, y) \quad (4.1.2)$$

é dado por

$$Y_i = f\left(t_{n-1} + c_i h, y_{n-1} + h \sum_{j=1}^M a_{ij} Y_j\right), \quad i = 1, 2, \dots, M \quad (4.1.3)$$
$$y_n = y_{n-1} + h \sum_{i=1}^M b_i Y_i$$

onde  $h = t_n - t_{n-1}$ .

Considerando este método aplicado ao sistema de EAD (4.1.1), temos

$$F\left(t_{n-1} + c_i h, y_{n-1} + h \sum_{j=1}^M a_{ij} Y_j, Y_i\right) = 0, \quad i = 1, 2, \dots, M \quad (4.1.4)$$

$$y_n = y_{n-1} + h \sum_{i=1}^M b_i Y_i.$$

O método de RKI, também, pode ser denotado pelo diagrama de Butcher

$$\begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \cdots & a_{1M} \\ c_2 & a_{21} & a_{22} & \cdots & a_{2M} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_M & a_{M1} & a_{M2} & \cdots & a_{MM} \\ \hline & b_1 & b_2 & \cdots & b_M \end{array} = \frac{c}{b'} \quad (4.1.5)$$

onde  $\mathcal{A} = (a_{ij})$ ,  $b = [b_1, b_2, \dots, b_M]^t$  e  $c = [c_1, c_2, \dots, c_M]^t$ .

Consideraremos apenas métodos cuja matriz Runge-Kutta  $\mathcal{A} = (a_{ij})$ , em (4.1.5), seja não singular.

Nos preocuparemos aqui, apenas com sistemas de EAD de índice um. Uma EAD linear

$$A(t)y'(t) + B(t)y(t) = g(t), \quad (4.1.6)$$

tem índice um se existem matrizes, dependendo do tempo, não singulares  $P$  e  $Q$  tais que

$$P(t)A(t)Q(t) = \begin{bmatrix} I_{m1} & 0 \\ 0 & 0 \end{bmatrix}, \quad P(t)B(t)Q(t) = \begin{bmatrix} C(t) & 0 \\ 0 & I_{m2} \end{bmatrix} \quad (4.1.7)$$

Esta transformação separa o sistema em partes "diferencial" e "algébrica". No caso de sistemas não lineares, associamos as matrizes  $A$  e  $B$  com  $F_y$  e  $F_y$ , respectivamente.

## 4.2 Sistemas Lineares com Coeficientes Constantes

Considere o sistema de EAD (4.1.1) e o método de RKI de M-estágios(4.1.4) aplicado a este sistema onde a matriz  $\mathcal{A} = (a_{ij})$  é não singular. Sendo  $y_n$  determinado por (4.1.4),  $y_{n-1}$  e  $t_{n-1}$  conhecidos, o método de RKI pode ser escrito na forma

$$y_n = y_{n-1} + h\psi(t_{n-1}, y_{n-1}, h). \quad (4.2.1)$$

Partindo deste procedimento necessitamos de algumas definições.

**Definição 4.2.1** O erro local  $d_n$  do método (4.1.4) é dado por

$$y(t_n) = y(t_{n-1}) + h\psi(t_{n-1}, y(t_{n-1}), h) - d_n. \quad (4.2.2)$$

**Definição 4.2.2** O método de RKI (4.1.4) é estritamente estável para a EAD (4.1.1) se a diferença entre o passo Runge-Kutta perturbado

$$F\left(t_{n-1} + c_i h, z_{n-1} + h \sum_{j=1}^M a_{ij} Z_j' + \delta_n^{(i)}, Z_i'\right) = 0, \quad i = 1, 2, \dots, M \quad (4.2.3)$$

$$z_n = z_{n-1} + h \sum_{i=1}^M b_i Z_i' + \delta_n^{(M+1)}$$

sendo  $z_0 = y_0 + \delta_0^{(M+1)}$ ,  $\|\delta_n^{(i)}\| = \Delta$ ,  $i = 1, 2, \dots, (M+1)$ , e o passo Runge-Kutta (4.1.4) satisfazem  $\|z_n - y_n\| \leq K_0 \Delta$ , onde  $0 < h \leq h_0$  e  $K_0, h_0$  são constantes que dependem apenas do método e da EAD.

Consideraremos a EAD linear com coeficiente constante

$$Ay' + By = g(t). \quad (4.2.4)$$

Assumindo que o índice é um, então existem matrizes não singulares  $P$  e  $Q$  tais que

$$PAQ = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \quad PBQ = \begin{bmatrix} C & 0 \\ 0 & I \end{bmatrix}. \quad (4.2.5)$$

Aplicando o método de RKI em (4.2.4), obtemos

$$AY_i' + B\left(y_{n-1} + h \sum_{j=1}^M a_{ij} Y_j'\right) = g(t_{n-1} + c_i h), \quad i = 1, 2, \dots, M \quad (4.2.6)$$

$$y_n = y_{n-1} + h \sum_{i=1}^M b_i Y_i'.$$

Considerando  $\tilde{y}_n = Q^{-1}y_n$ ,  $\tilde{Y}_i' = Q^{-1}Y_i'$ ,  $[\tilde{g}_1, \tilde{g}_2]^t = \tilde{g}(t) = Pg(t)$  e pré-multiplicando (4.2.6) por  $P$ , obtemos

$$PAQ\tilde{Y}_i' + PBQ\left(\tilde{y}_{n-1} + h \sum_{j=1}^M a_{ij}\tilde{Y}_j'\right) = \tilde{g}(t_{n-1} + c_i h), \quad (4.2.7)$$

$$\tilde{y}_n = \tilde{y}_{n-1} + h \sum_{i=1}^M b_i \tilde{Y}_i'.$$

Desta forma as partes diferencial e algébrica do sistema são completamente desacopladas. Assim, é suficiente estudarmos separadamente estas partes para obtermos um entendimento geral das EAD lineares com coeficientes constantes.

Desde que tenhamos assumido que o índice é um, o sistema (4.2.4) divide-se em um sistema de equações diferenciais e em um sistema de equações algébricas, onde o sistema de equações algébricas é

$$\tilde{y}_2 = \tilde{g}_2(t), \quad (4.2.8)$$

que por simplicidade denotaremos  $\tilde{y}_2 = y$  e  $\tilde{g}_2 = g$ .

Aplicando o método de RKI, definido em (4.1.4), na equação (4.2.8), temos

$$y_{n-1} + h \sum_{j=1}^M a_{ij} Y'_j = g(t_{n-1} + c_i h), \quad i = 1, 2, \dots, M \quad (4.2.9a)$$

$$y_n = y_{n-1} + h \sum_{i=1}^M b_i Y'_i. \quad (4.2.9b)$$

Devido a  $\mathcal{A} = (a_{ij})$  ser não singular, determinamos  $Y' = [Y'_1, Y'_2, \dots, Y'_M]^T$  unicamente por (4.2.9a).

Para encontrar o erro local consideraremos

$$y_{n-1} = g(t_{n-1}), \quad (4.2.10)$$

assim, substituindo (4.2.10) em (4.2.9a), obtemos

$$\sum_{j=1}^M a_{ij} Y'_j = \frac{1}{h} (g(t_{n-1} + c_i h) - g(t_{n-1})),$$

que fornece o sistema de equações

$$\begin{aligned} a_{11} Y'_1 + a_{12} Y'_2 + \dots + a_{1M} Y'_M &= (g(t_{n-1} + c_1 h) - g(t_{n-1}))/h \\ a_{21} Y'_1 + a_{22} Y'_2 + \dots + a_{2M} Y'_M &= (g(t_{n-1} + c_2 h) - g(t_{n-1}))/h \\ &\vdots \\ a_{M1} Y'_1 + a_{M2} Y'_2 + \dots + a_{MM} Y'_M &= (g(t_{n-1} + c_M h) - g(t_{n-1}))/h \end{aligned}$$

cuja forma matricial é

$$\begin{bmatrix} a_{11} & \dots & a_{1M} \\ \vdots & \ddots & \vdots \\ a_{M1} & \dots & a_{MM} \end{bmatrix} \begin{bmatrix} Y'_1 \\ \vdots \\ Y'_M \end{bmatrix} = \begin{bmatrix} (g(t_{n-1} + c_1 h) - g(t_{n-1}))/h \\ \vdots \\ (g(t_{n-1} + c_M h) - g(t_{n-1}))/h \end{bmatrix}$$

Sendo  $\mathcal{A} = (a_{ij})$  não singular e considerando  $\mathcal{A}^{-1} = (\tilde{a}_{ij})$  temos

$$\begin{bmatrix} Y'_1 \\ \vdots \\ Y'_M \end{bmatrix} = \begin{bmatrix} \tilde{a}_{11} & \cdots & \tilde{a}_{1M} \\ \vdots & \ddots & \vdots \\ \tilde{a}_{M1} & \cdots & \tilde{a}_{MM} \end{bmatrix} \begin{bmatrix} (g(t_{n-1} + c_1 h) - g(t_{n-1}))/h \\ \vdots \\ (g(t_{n-1} + c_M h) - g(t_{n-1}))/h \end{bmatrix},$$

que pode ser escrito na forma

$$Y' = \mathcal{A}^{-1}G, \quad (4.2.11a)$$

onde

$$G = \begin{bmatrix} (g(t_{n-1} + c_1 h) - g(t_{n-1}))/h \\ \vdots \\ (g(t_{n-1} + c_M h) - g(t_{n-1}))/h \end{bmatrix}. \quad (4.2.11b)$$

Agora, substituindo (4.2.10) na equação (4.2.9b), obtemos

$$\begin{aligned} y_n &= g(t_{n-1}) + h \sum_{i=1}^M b_i Y'_i = g(t_{n-1}) + h(b_1 Y'_1 + b_2 Y'_2 + \cdots + b_M Y'_M) \\ &= g(t_{n-1}) + h[b_1, \dots, b_M] \begin{bmatrix} Y'_1 \\ \vdots \\ Y'_M \end{bmatrix} = g(t_{n-1}) + hb' \mathcal{A}^{-1}G. \end{aligned} \quad (4.2.12)$$

Assim, o erro local é dado por

$$d_n = g(t_{n-1}) + hb' \mathcal{A}^{-1}G - g(t_n). \quad (4.2.13)$$

Expandindo os termos de (4.2.13) em série de Taylor em torno de  $t_{n-1}$  temos

$$\begin{aligned} d_n &= g(t_{n-1}) + hb' \mathcal{A}^{-1} \begin{bmatrix} \left[ \left( g(t_{n-1}) + g'(t_{n-1})c_1 h + g''(t_{n-1}) \frac{(c_1 h)^2}{2} + \cdots \right) - g(t_{n-1}) \right] / h \\ \vdots \\ \left[ \left( g(t_{n-1}) + g'(t_{n-1})c_M h + g''(t_{n-1}) \frac{(c_M h)^2}{2} + \cdots \right) - g(t_{n-1}) \right] / h \end{bmatrix} \\ &\quad - \left( g(t_{n-1}) + g'(t_{n-1})(t_n - t_{n-1}) + g''(t_{n-1}) \frac{(t_n - t_{n-1})^2}{2} + \cdots \right) \end{aligned}$$

sendo  $h = t_n - t_{n-1}$  obtemos

$$d_n = - \left( hg' + \frac{h^2}{2} g'' + \frac{h^3}{6} g''' + \dots \right) + b' A^{-1} \begin{bmatrix} c_1 hg' + \frac{(c_1 h)^2}{2} g'' + \frac{(c_1 h)^3}{6} g''' + \dots \\ \vdots \\ c_M hg' + \frac{(c_M h)^2}{2} g'' + \frac{(c_M h)^3}{6} g''' + \dots \end{bmatrix}$$

$$= - \sum_{r=1}^k \frac{h^r}{r!} g^{(r)}(t_{n-1}) + b' A^{-1} \sum_{r=1}^k \frac{c^j h^r}{r!} g^{(r)}(t_{n-1}) + R_d$$

onde  $R_d = - \sum_{r=k+1}^{\infty} \frac{h^r}{r!} g^{(r)}(t_n) (1 - b' A^{-1} c^j)$  e  $c^j = [c_1^j, c_2^j, \dots, c_M^j]^t$ ,  $j = 1, 2, \dots, k$ , assim

$$d_n = - \sum_{r=1}^k \frac{h^r}{r!} g^{(r)}(t_{n-1}) (1 - b' A^{-1} c^j) + R_d$$

supondo  $b' A^{-1} c^j = 1$ ,  $j = 1, 2, \dots, k$  (faremos um comentário mais geral ao definirmos as condições algébricas mais a frente), temos  $d_n = R_d$  e

$$R_d = - \sum_{r=k+1}^{\infty} \frac{h^r}{r!} g^{(r)}(t_{n-1}) (1 - b' A^{-1} c^j) = \frac{h^{k+1}}{(k+1)!} g^{(k+1)}(\xi) (1 - b' A^{-1} c^{k+1}).$$

Considerando que  $\|g^{(k+1)}\| \leq M_{k+1}$ , podemos majorar  $R_d$  da seguinte forma

$$\|R_d\| \leq \frac{h^{k+1}}{(k+1)!} M_{k+1} (1 - |b' A^{-1} c^{k+1}|). \quad (4.2.14)$$

Portanto,  $d_n = O(h^{k+1})$  se  $b' A^{-1} c^j = 1$ ,  $j = 1, 2, \dots, k$ .

**Definição 4.2.3** A ordem algébrica do método de RKI (4.2.11) é igual a  $k_a$  se  $d_n = O(h^{k_a+1})$ , para toda equação (4.2.8) com  $g(t)$  suficientemente suave.

**Teorema 4.2.1** A ordem algébrica do método de RKI (4.1.4) é  $k_a$  se, e somente se os coeficientes do método satisfazem

$$b' A^{-1} c^j = 1, \quad j = 1, 2, \dots, k_a. \quad (4.2.15)$$

A seguir, examinaremos a propagação dos erros para o método de RKI aplicado em sistemas lineares com coeficientes constantes [PETZOLD - 86]. Resolvendo (4.2.10) para o método de RKI com perturbação (4.2.3) temos

$$z_{n-1} + h \sum_{j=1}^M a_{ij} Z_j' + \delta_n^{(i)} = g(t_{n-1} + c_i h), \quad i = 1, 2, \dots, M$$

$$z_n = z_{n-1} + h \sum_{i=1}^M b_i Z_i' + \delta_n^{(M+1)}.$$
(4.2.16)

Subtraindo (4.2.16) de (4.2.9), considerando  $e_n = y_n - z_n$  e  $E_i' = Y_i' - Z_i'$ , obtemos

$$e_{n-1} + h \sum_{j=1}^M a_{ij} E_j' - \delta_n^{(i)} = 0, \quad i = 1, 2, \dots, M$$
(4.2.17a)

$$e_n = e_{n-1} + h \sum_{i=1}^M b_i E_i' - \delta_n^{(M+1)}.$$
(4.2.17b)

Resolvendo (4.2.17a) para  $E_i'$  temos

$$h[a_{11}E_1' + a_{12}E_2' + \dots + a_{1M}E_M'] = \delta_n^{(1)} + e_{n-1}$$

$$\vdots$$

$$h[a_{M1}E_1' + a_{M2}E_2' + \dots + a_{MM}E_M'] = \delta_n^{(M)} + e_{n-1}.$$

Matricialmente temos

$$h \begin{bmatrix} a_{11} & \dots & a_{1M} \\ \vdots & \ddots & \vdots \\ a_{M1} & \dots & a_{MM} \end{bmatrix} \begin{bmatrix} E_1' \\ \vdots \\ E_M' \end{bmatrix} = \begin{bmatrix} \delta_n^{(1)} \\ \vdots \\ \delta_n^{(M)} \end{bmatrix} + \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} e_{n-1},$$

ainda, podendo ser escrito como

$$E_i' = \frac{1}{h} \mathcal{A}^{-1} (\delta_n - \epsilon_M e_{n-1}),$$
(4.2.18)

onde  $E_i' = [E_1', E_2', \dots, E_M']^t$ ,  $\delta_n = [\delta_n^{(1)}, \delta_n^{(2)}, \dots, \delta_n^{(M)}]^t$  e  $\epsilon_M = [1, 1, \dots, 1]^t$ .

Substituindo em (4.2.17b) temos

$$e_n = (1 - b^t \mathcal{A}^{-1} \epsilon_M) e_{n-1} + b^t \mathcal{A}^{-1} \delta_n - \delta_n^{(M+1)}$$
(4.2.19)

Assim, mostramos o seguinte resultado:



**Teorema 4.2.2** *O método de RKI (4.1.4) é estritamente estável para EAD lineares com coeficientes constantes de índice um se, e somente se os coeficientes do método satisfazem*

$$|1 - b' \mathcal{A}^{-1} \epsilon_M| < 1. \quad (4.2.20)$$

Definindo a constante de estabilidade  $r$  por  $r = 1 - b' \mathcal{A}^{-1} \epsilon_M$ , dizemos que o método de RKI satisfaz a condição estrita de estabilidade se  $|r| < 1$  e é satisfeita a condição de estabilidade se  $|r| \leq 1$ .

A estimativa para o erro local torna-se simples, para sistemas de índice um, agora que conhecemos o erro de truncamento local e a propriedade de estabilidade do método. Considerando que  $\delta_n$  e  $\delta_n^{(M+1)}$  representam, em cada estágio, o erro de truncamento local, excluindo o erro de arredondamento e, de acordo com a suposição do erro local (4.2.10), segue por (4.2.20) que

$$e_n = (b' \mathcal{A}^{-1} \delta_n - \delta_n^{(M+1)}). \quad (4.2.21)$$

Como o erro local  $d_n$  foi denotado como sendo  $O(h^{k_a+1})$  e de acordo com a suposição de que  $e_n = d_n$  temos

$$b' \mathcal{A}^{-1} \delta_n - \delta_n^{(M+1)} = O(h^{k_a+1}).$$

Em seguida, reescrevemos a expressão para o erro global dado na equação (4.2.19), isto é,

$$e_n = r e_{n-1} + (b' \mathcal{A}^{-1} \delta_n - \delta_n^{(M+1)}).$$

Assim,

$$\begin{aligned} e_{n-1} &= r e_{n-2} + (b' \mathcal{A}^{-1} \delta_{n-1} - \delta_{n-1}^{(M+1)}) \\ e_{n-2} &= r e_{n-3} + (b' \mathcal{A}^{-1} \delta_{n-2} - \delta_{n-2}^{(M+1)}) \\ &\vdots \\ e_1 &= r e_0 + (b' \mathcal{A}^{-1} \delta_1 - \delta_1^{(M+1)}) \end{aligned}$$

Portanto,  $e_n$  torna-se

$$e_n = r^n e_0 + \sum_{i=0}^{n-1} r^i (b' \mathcal{A}^{-1} \delta_{n-i} - \delta_{n-i}^{(M+1)}). \quad (4.2.22)$$

Observe que, o erro global será igual ao erro local se  $r = 0$  em (4.2.22). Para  $|r| < 1$ , segue de (4.2.22) que

$$\|e_n\| \leq |r|^n \|e_0\| + \left| \frac{1-r^n}{1-r} \right| M h^{k_a+1},$$

onde  $M$  é uma constante positiva. Visto que  $r$  independe de  $h$  e  $\lim_{n \rightarrow \infty} \frac{1-r^n}{1-r} = \frac{1}{1-r}$  temos, neste caso, que a ordem do erro global é a mesma do erro local na parte "algébrica do sistema". A partir disto temos os seguintes resultados.

**Definição 4.2.4** *A ordem do método de RKI (4.1.4) aplicado a sistemas de EAD com coeficientes constantes, é igual a  $k_c$  se o método converge com erro global  $O(h^{k_c})$  para todo sistema linear com coeficientes constantes (4.2.4) de índice um e com  $g(t)$  suficientemente suave.*

**Teorema 4.2.3** *Suponha que o método de RKI (4.1.4) satisfaça a condição de estabilidade (4.2.20) e a matriz  $\mathcal{A}$  dos coeficientes do método é não singular. Então a ordem do erro global do método de RKI, aplicado a sistemas de EAD lineares com coeficientes constantes, é  $k_c$ , dado por*

$$k_c = \min(k_a + 1, k_d), \quad (4.2.23)$$

onde  $k_d$  é a ordem do método para sistemas puramente diferenciais.

### 4.3 Sistemas não Lineares

Os métodos de RKI, em geral, não obtém a mesma ordem de exatidão para sistemas completamente implícitos de índice um como ocorrem em sistemas lineares. Esta perda adicional de exatidão deve-se ao fato da mistura que pode ocorrer entre a solução da parte diferencial e da parte algébrica do sistema.

**Definição 4.3.1** *O  $i$ -ésimo erro de truncamento local  $\delta_i^{(n)}$ , em  $t_n$ , do método de RKI de  $M$ -estágios é dado por*

$$\delta_i^{(n)} = y(t_{n-1}) + h \sum_{j=1}^M a_{ij} y'(t_{n-1} + c_j h) - y(t_{n-1} + c_i h), \quad i = 1, 2, \dots, M \quad (4.3.1a)$$

$$\delta_{M+1}^{(n)} = y(t_{n-1}) + h \sum_{i=1}^M b_i y'(t_{n-1} + c_i h) - y(t_n). \quad (4.3.1b)$$

**Definição 4.3.2** A condição de ordem  $k_j$ , do método de RKI de  $M$ -estágios(4.1.4), é dada por

$$k_j = \min(k_1, k_2, \dots, k_M, k_{M+1}), \quad (4.3.2)$$

com  $\delta_i^{(n)} = O(h^{k_i+1})$ ,  $i = 1, 2, \dots, (M+1)$ .

Considere  $c_i = \sum_{j=1}^M a_{ij}$  e as condições algébricas  $C(q)$ ,  $B(q)$ ,  $D(q)$  e  $A_1(q)$

$$C(q): \sum_{j=1}^M a_{ij} c_j^{k-1} = \frac{c_i^k}{k}, \quad i = 1, 2, \dots, M, \quad k = 1, 2, \dots, q$$

$$B(q): \sum_{i=1}^M b_i c_i^{k-1} = \frac{1}{k}, \quad k = 1, 2, \dots, q \quad (4.3.3)$$

$$D(q): \sum_{i=1}^M b_i c_i^{k-1} a_{ij} = \frac{b_j}{k} (1 - c_j^k), \quad j = 1, 2, \dots, M, \quad k = 1, 2, \dots, q$$

$$A_1(q): b^T A^{-1} c^k = 1, \quad k = 1, 2, \dots, q.$$

O método de RKI tem condição de ordem  $k_j$  se  $C(k_j)$  e  $B(k_j)$  forem válidos. Além disso, se  $C(q)$  e  $B(q+1)$  forem válidos, a ordem  $k_d$  da EDO é pelo menos  $q+1$ .  $A_1(q)$  corresponde a condição de ordem para sistemas lineares com coeficientes constantes de índice um (ver seção 4.2). Isto pode ser visto, facilmente pela expansão de (4.3.1) em série de Taylor em torno de  $t_{n-1}$  (veja [FRANK - 85a] e [FRANK - 85b]).

Por meio das equações algébricas  $C(q)$ ,  $B(q)$  e  $D(q)$  podemos obter várias classes de métodos. Uma classe destes métodos é o RADAU IIA, que será discutido no Capítulo 5.

**Teorema 4.3.1** Suponha que  $F(y, y', t) = 0$  tenha índice uniforme um, que o método de RKI satisfaça a condição de estabilidade  $|r| \leq 1$ , que os erros nas condições iniciais sejam  $O(h^G)$  e que os erros na iteração de Newton sejam  $O(h^{G+\delta})$  onde  $\delta = 1$  se  $|r| = 1$ , caso contrário,  $\delta = 0$  e  $G > 1$ . Então o erro global satisfaz  $\|e_n\| = O(h^G)$ , onde

$$G = \begin{cases} q, & \text{se } C(q) \text{ e } B(q) \text{ são válidos} \\ q+1, & \text{se } C(q), B(q+1) \text{ são válidos e } -1 \leq r < 1 \\ q+1, & \text{se } C(q), B(q+1), A_1(q+1) \text{ são válidos e } r = 1 \end{cases}$$

A demonstração deste teorema encontra-se com mais detalhes em [ROMEIRO - 94].

# 5 Algoritmos

---

---

## 5.1 Introdução

Neste capítulo analisaremos o desenvolvimento de algoritmos para resolver sistemas de EAD, usando os métodos BDF e RKI e, em particular, os algoritmos utilizados em dois códigos já consagrados, o DASSL e o RADAU5.

O código DASSL, de Petzold [PETZOLD - 82], utiliza o método BDF.

O código RADAU5, de Hairer e Wainner [HAIRER - 91], utiliza o método RADAU IIA de 3-estágios.

O método RADAU IIA é uma das várias classes existentes do método de RKI, obtido a partir das condições algébricas  $B(q)$ ,  $C(q)$  e  $D(q)$  (ver equações (4.3.3)), onde estas condições satisfazem

$$B(2M-1), C(M), D(M-1), c_M = 1 \text{ e } b_i = a_{M_i}, \quad i = 1, \dots, M.$$

## 5.2 Algoritmos e Estratégias no DASSL

DASSL é um código para resolver sistemas de equações algébrico-diferenciais, de índice zero ou um, na forma

$$\begin{aligned} F(t, y, y') &= 0 \\ y(t_0) &= y_0 \\ y'(t_0) &= y'_0 \end{aligned} \tag{5.2.1}$$

onde  $F$ ,  $y$  e  $y'$  são vetores  $n$ -dimensionais.

O DASSL aproxima a derivada usando fórmulas de diferenças para trás (BDF) de  $k$ -ésima ordem,  $k = 1, \dots, 5$  e, em todo passo, é escolhido a ordem  $k$  e o tamanho do passo  $h_{n+1}$  baseado no procedimento da solução.

Aqui, descreveremos a fórmula básica usada no DASSL. Para encontrar uma aproximação para a solução em  $t_{n+1}$ , primeiramente, supomos um valor inicial para a solução através da avaliação do polinômio preditor e a sua derivada em  $t_{n+1}$ . O polinômio preditor  $w_{n+1}^P$  é um polinômio que interpola  $y_{n-i}$ , isto é,

$$w_{n+1}^P(t_{n-i}) = y_{n-i}, \quad i = 0, 1, 2, \dots, k. \quad (5.2.2)$$

Os valores previstos para  $y$  e  $y'$  são obtidos pela avaliação  $w_{n+1}^P(t_{n+1})$  e  $w_{n+1}'(t_{n+1})$ . Assim,

$$\begin{aligned} y_{n+1}^{(0)} &= w_{n+1}^P(t_{n+1}) \\ y_{n+1}'^{(0)} &= w_{n+1}'(t_{n+1}). \end{aligned} \quad (5.2.3)$$

A aproximação  $y_{n+1}$ , para a solução em  $t_{n+1}$ , é a solução da fórmula do corretor e é, finalmente, reconhecida pelo DASSL. A fórmula usada é de coeficiente principal fixo de  $k$ -ésima ordem do método BDF.

A solução para a fórmula do corretor é o vetor  $y_{n+1}$  dado pelo polinômio corretor que interpola o polinômio preditor em  $k$  pontos igualmente espaçados  $t_{n+1} - ih$ ,  $i = 1, \dots, k$  e satisfaz (5.2.1), isto é,

$$\begin{aligned} w_{n+1}^C(t_{n+1}) &= y_{n+1} \\ w_{n+1}^C(t_{n+1} - ih_{n+1}) &= w_{n+1}^P(t_{n+1} - ih_{n+1}), \quad 1 \leq i \leq k \\ F(t_{n+1}, w_{n+1}^C(t_{n+1}), w_{n+1}'(t_{n+1})) &= 0. \end{aligned} \quad (5.2.4)$$

Os valores do preditor  $y_{n+1}^{(0)}$ ,  $y_{n+1}'^{(0)}$  e do corretor  $y_{n+1}$  em  $t_{n+1}$  são definidos em termos dos polinômios que interpola a solução em poucos passos. Segundo as idéias de Krogh [KROGH - 74] e Shampine e Gordon [SHAMPINE - 75], estes polinômios são representados no DASSL em termos de diferenças divididas modificada em  $y$ . Mais precisamente, as quantidades descritas passo a passo são dadas por

$$\psi_i(n+1) = h_{n+1} + h_n + \dots + h_{n+2-i} = t_{n+1} - t_{n+1-i}, \quad i \geq 1$$

$$\alpha_i(n+1) = h_{n+1} / \psi_i(n+1), \quad i \geq 1$$

$$\beta_1(n+1) = 1$$

$$\beta_i(n+1) = \frac{\psi_1(n+1) \psi_2(n+1) \dots \psi_{i-1}(n+1)}{\psi_1(n) \psi_2(n) \dots \psi_{i-1}(n)}, \quad i > 1$$

$$\phi_1(n) = y_n$$

$$\phi_i(n) = \psi_1(n) \psi_2(n) \dots \psi_{i-1}(n) [y_n, y_{n-1}, \dots, y_{n-i+1}], \quad i > 1$$

$$\phi_i^*(n) = \beta_i(n+1)\phi_i(n), \quad i \geq 1 \quad (5.2.5)$$

$$\rho_1(n+1) = 1$$

$$\sigma_i(n+1) = \frac{h_{n+1}^i(i-1)!}{\psi_1(n+1)\psi_2(n+1)\dots\psi_i(n+1)}, \quad i > 1$$

$$\gamma_1(n+1) = 0$$

$$\gamma_i(n+1) = \gamma_{i-1}(n+1) + \alpha_{i-1}(n+1)/h_{n+1}, \quad i > 1$$

$$\alpha_s = -\sum_{j=1}^k \frac{1}{j}$$

$$\alpha^0(n+1) = -\sum_{j=1}^k \alpha_j(n+1).$$

As diferenças divididas são definidas pelas relações de recorrências

$$[y_n] = y_n$$

$$[y_n, y_{n-1}, \dots, y_{n-k}] = \frac{[y_n, y_{n-1}, \dots, y_{n-k+1}] - [y_{n-1}, y_{n-2}, \dots, y_{n-k}]}{t_n - t_{n-k}}.$$

O polinômio do preditor é dado em termos de diferenças divididas por

$$\begin{aligned} w_{n+1}^P(t) = & y_n + (t-t_n)[y_n, y_{n-1}] + (t-t_n)(t-t_{n-1})[y_n, y_{n-1}, y_{n-2}] + \dots \\ & + (t-t_n)(t-t_{n-1})\dots(t-t_{n-k+1})[y_n, y_{n-1}, \dots, y_{n-k}]. \end{aligned} \quad (5.2.6)$$

Avaliando  $w_{n+1}^P$  em  $t_{n+1}$  e reescrevendo (5.2.6) em termos da notação de (5.2.5) temos

$$w_{n+1}^P(t_{n+1}) = \sum_{i=1}^{k+1} \phi_i^*(n)y_n$$

Logo, a fórmula do preditor é  $y_{n+1}^{(0)} = \sum_{i=1}^{k+1} \phi_i^*(n)$ .

Agora, diferenciando (5.2.6) e avaliando em  $t_{n+1}$ , encontramos, após algumas manipulações,

$$y_{n+1}^{(0)} = \sum_{i=1}^{k+1} \gamma_i(n+1)\phi_i^*(n).$$

Para encontrarmos a fórmula do corretor temos, das equações (5.2.2), (5.2.3) e (5.2.4), que

$$w_{n+1}^C(t) - w_{n+1}^P(t) = b(t)(y_{n+1} - y_{n+1}^{(0)}) \quad (5.2.7)$$

onde

$$\begin{aligned} b(t_{n+1} - ih_{n+1}) &= 0 \quad i = 1, 2, \dots, k \\ b(t_{n+1}) &= 1. \end{aligned}$$

Diferenciando a equação (5.2.7) e avaliando em  $t_{n+1}$  obtemos a equação do corretor para a implementação do coeficiente principal fixo,

$$\alpha_s (y_{n+1} - y_{n+1}^{(0)}) + h_{n+1} (y'_{n+1} - y'_{n+1}{}^{(0)}) = 0, \quad (5.2.8)$$

onde  $\alpha_s$  está definido em (5.2.5).

Resolvendo (5.2.8) para  $y'_{n+1}$  e substituindo em (5.2.2) temos

$$F\left(t_{n+1}, y_{n+1}, y_{n+1}^{(0)} - \frac{\alpha_s}{h_{n+1}}(y_{n+1} - y_{n+1}^{(0)})\right) = 0. \quad (5.2.9)$$

### 5.2.1 Solução do Sistema não Linear

Para simplificar a notação, reescreveremos a equação do corretor (5.2.9) como sendo

$$F(t, y, \alpha y + \beta) = 0, \quad (5.2.10)$$

com  $\alpha = \frac{\alpha_s}{h_{n+1}}$  e  $\beta = y_{n+1}^{(0)} - \alpha y_{n+1}^{(0)}$ , sendo todas as variáveis avaliadas em  $t_{n+1}$ .

O corretor é resolvido usando o método de Newton modificado. Isto nos dá

$$y^{(m+1)} = y^{(m)} - G^{-1}F(t, y^{(m)}, \alpha y^{(m)} + \beta), \quad (5.2.11)$$

onde  $y^{(0)}$  é dado por (5.2.6) e  $G$  é a matriz

$$G = \alpha F_y(t, y^{(0)}, \alpha y^{(0)} + \beta) + F_y(t, y^{(0)}, \alpha y^{(0)} + \beta).$$

Em DASSL o critério de parada é dado por

$$\|y^* - y^{(m+1)}\| \leq \frac{\rho}{1-\rho} \|y^{(m+1)} - y^{(m)}\|,$$

onde  $y^*$  é a solução da equação do corretor e  $\rho$  é uma estimativa do raio de convergência, sendo que esta estimativa é dada por

$$\rho = \left( \frac{\|y^{(m+1)} - y^{(m)}\|}{\|y^{(1)} - y^{(0)}\|} \right)^{1/m}.$$

### 5.3 Algoritmos e Estratégias no RADAU5

O código RADAU5 é indicado para resolver sistemas semi explícitos de equações algébrico-diferenciais cuja forma é

$$By' = f(x, y), \quad y(x_0) = y_0,$$

onde  $B$  é uma matriz constante quadrada.

#### 5.3.1 Solução do Sistema não Linear

Aplicando o método de RKI na EAD (4.1.1) obtemos

$$F\left(t_{n-1} + c_i h, y_{n-1} + h \sum_{j=1}^M a_{ij} Y_j, Y_i\right) = 0 \quad i = 1, \dots, M. \quad (5.3.1)$$

O método de Newton, aplicado a este sistema, é dado por

$$Z^{(k+1)} = Z^{(k)} - DG^{-1}(Z^{(k)})G(Z^{(k)}) \quad (5.3.2)$$

sendo

$$Z^{(k)} = (Y_1^{(k)}, Y_2^{(k)}, \dots, Y_M^{(k)})^t;$$

$$G(Z^{(k)}) = \left[ F\left(t_{n-1} + c_1 h, y_{n-1} + h \sum_{j=1}^M a_{1j} Y_j^{(k)}, Y_1^{(k)}\right), \dots, F\left(t_{n-1} + c_M h, y_{n-1} + h \sum_{j=1}^M a_{Mj} Y_j^{(k)}, Y_M^{(k)}\right) \right]^t;$$



$$DG(Z^{(k)}) = \begin{bmatrix} F_y^1 + ha_{11}F_y^1 & \cdots & ha_{1M}F_y^1 \\ \vdots & \ddots & \vdots \\ ha_{M1}F_y^M & \cdots & F_y^M + ha_{MM}F_y^M \end{bmatrix},$$

onde

$$F_y^i = F_y \left( t_{n-1} + c_i h, y_{n-1} + h \sum_{j=1}^M a_{ij} Y_j^{(k)}, Y_i^{(k)} \right)$$

e

$$F_y^i = F_y \left( t_{n-1} + c_i h, y_{n-1} + h \sum_{j=1}^M a_{ij} Y_j^{(k)}, Y_i^{(k)} \right),$$

para  $i = 1, \dots, M$ .

Ao calcularmos a matriz  $DG(Z^{(k)})$ , em cada iteração de Newton, obtemos um custo muito alto, para evitarmos isto, usamos o método de Newton modificado. Este método faz com que a matriz  $DG(Z^{(k)})$  seja a mesma em todas as iterações, isto é  $F_y^* = f_y(t_{n-1}, y_{n-1})$  e  $F_y^* = B$ . Desta forma  $DG(Z^{(k)})$  torna-se

$$DG(Z^{(k)}) = \begin{bmatrix} F_y^* + ha_{11}F_y^* & \cdots & ha_{1M}F_y^* \\ \vdots & \ddots & \vdots \\ ha_{M1}F_y^* & \cdots & F_y^* + ha_{MM}F_y^* \end{bmatrix} = I \otimes F_y^* + hA \otimes F_y^*$$

Logo, para resolvermos (5.3.1), usamos a iteração de Newton modificado, isto é

$$\begin{aligned} (I \otimes F_y^* + hA \otimes F_y^*) S^{(k)} &= -G(Z^{(k)}) \\ Z^{(k+1)} &= Z^{(k)} + S^{(k)}. \end{aligned}$$

A matriz  $(I \otimes F_y^* + hA \otimes F_y^*)$  é inversível, isto pode ser visto em usando o Lema 3.3.1.

Pelo fato da convergência ser linear, isto é,  $\|S^{(k+1)}\| \leq \rho \|S^{(k)}\|$ ,  $\rho < 1$ , o critério de parada usado em RADAU5 é dado por

$$\|Z^{(k+1)} - Z^*\| \leq \frac{\rho}{1-\rho} \|Z^{(k+1)} - y^{(k)}\|, \quad (5.3.3)$$

onde  $\rho$  é uma estimativa do raio de convergência

$$\rho_k = \frac{\|Z^{(k+1)} - Z^{(k)}\|}{\|Z^{(k)} - Z^{(k-1)}\|}, \quad k \geq 1, \quad (5.3.4)$$

O erro iteração não deve ser maior que a discretização do erro local. Sendo, usualmente, considerado como sendo próximo de uma certa tolerância (*Tol*). Por esta razão, a iteração para quando

$$\eta_k \|S^{(k)}\| \leq k \cdot Tol, \quad (5.3.5)$$

com  $\eta_k = \frac{\rho_k}{1 - \rho_k}$  e aceitando  $Z^{(k+1)}$  como sendo uma aproximação de  $Z^*$ . Esta estratégia só pode ser aplicada após, no mínimo, duas iterações.

### 5.3.2 Valor Inicial para a Iteração de Newton

Seja  $Y_i = y_{n-1} + h \sum_{j=1}^M a_{ij} Y_j$ ,  $i = 1, \dots, M$  e considere  $z_i = Y_i - y_{n-1}$ , onde  $z_i = O(h)$ , pois

$$z_i = h \sum_{j=1}^M a_{ij} Y_j, \quad i = 1, \dots, M. \quad (5.3.6)$$

Logo, uma escolha natural e simples do valor inicial para a iteração de Newton (5.3.2) pode ser

$$z_i^{(0)} = 0, \quad i = 1, \dots, M. \quad (5.3.7)$$

Porém, em geral, escolhas melhores são possíveis. Se o método de RKI satisfaz a condição  $C(q)$  (Capítulo 4), para algum  $q < M$ , então

$$z_i = y(t_{n-1} + c_i h) - y_{n-1} + O(h^{q+1}). \quad (5.3.8)$$

Suponha, agora que  $c_i \neq 0$ ,  $i = 1, 2, 3$ , pois estamos usando o método RADAU IIA de 3-estágios, e considere o polinômio interpolador de grau 3 definido por

$$\begin{aligned} Q(0) &= 0 \\ Q(c_i) &= z_i^{\text{old}}, \quad i = 1, 2, 3, \end{aligned}$$

onde  $z_i^{\text{old}}$  são determinados no passo anterior.

Desde que o erro na interpolação é  $O(h^{q+1})$ , obtemos junto com (5.3.8)  $y(t_{n-1} + c_i h) - y_{n-1} - Q(t) = O(h^{q+1})$ .

Usando valores de  $Q(t)$ , além do intervalo  $[0,1]$  (ver Figura 5.3.1), obtemos

$$z_i^{(0)} = Q(1 + w c_i) + y_n - y_{n-1}, \quad i = 1, 2, 3 \quad \text{e} \quad w = \frac{h_{\text{new}}}{h_{\text{old}}}. \quad (5.3.9)$$

Experimentos numéricos com o método RADAU IIA de 3-estágios mostram que (5.3.9) produz uma convergência mais segura que (5.3.7).

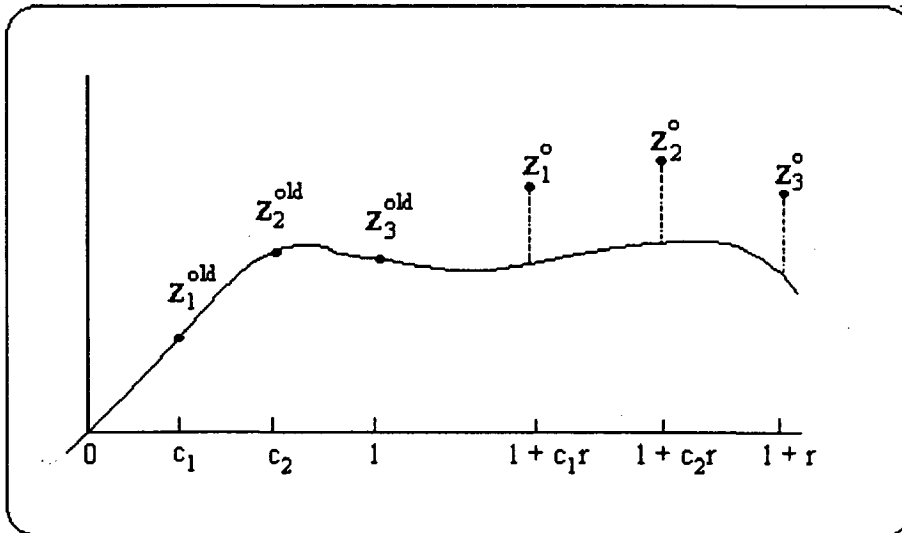


Figura 5.3.1 - Polinômio  $Q(t)$

# 6 Resultados Numéricos

---

---

## 6.1 Introdução

Neste capítulo usaremos os códigos DASSL e RADAU5 para obtermos a solução de sistemas de EAD de índice zero ou um das equações:

- Equação do Pêndulo;
- Equação de Van der Pol;
- Equação da Esfera.
- Equação Teste.

## 6.2 Equação do Pêndulo

Seja a EAD semi explícita não linear

$$p' = u \quad (6.2.1a)$$

$$q' = v \quad (6.2.1b)$$

$$mu' = -p\lambda \quad (6.2.1c)$$

$$mv' = -q\lambda - g \quad (6.2.1d)$$

$$0 = p^2 + q^2 - l^2. \quad (6.2.1e)$$

Nesta formulação o sistema (6.2.1) tem índice três. Diferenciando (6.2.1e) temos

$$0 = pu + qv. \quad (6.2.2)$$

Assim, as equações (6.2.1a)-(6.2.1d) e (6.2.2) formam a EAD não linear semi explícita de índice dois. Diferenciado mais uma vez e usando (6.2.1e) obtemos

$$0 = m(u^2 + v^2) - gq - l^2\lambda. \quad (6.2.3)$$

Assim (6.2.1a)-(6.2.1d) e (6.2.3) formam a EAD não linear semi explícita de índice um.

Para nossa análise usaremos o intervalo  $[0,10]$ , consideraremos  $l = 1$ ,  $g = 1$ ,  $m = 1$  e valores iniciais consistentes  $p(0) = 1$ ,  $q(0) = 0$ ,  $u(0) = 0$ ,  $v(0) = 0$  e  $\lambda(0) = 0$ .

A solução obtida pelo DASSL com uma tolerância para o erro local  $ATOL = RTOL = 1.0D - 5$  encontra-se na Tabela 6.2.1,

$x$	$p$	$q$
0.00	0.1000000000E+01	0.0000000000E+00
1.00	0.8795015006E+00	-0.4757984996E+00
2.00	-0.2041872563E+00	-0.9789636645E+00
3.00	-0.9688347838E+00	-0.2477426641E+00
4.00	-0.9990188721E+00	-0.4276027400E-01
5.00	-0.6851283926E+00	-0.7283317595E+00
6.00	0.5645477174E+00	-0.8255042201E+00
7.00	0.9961859922E+00	-0.8689733379E-01
8.00	0.9853129756E+00	-0.1698778608E+00
9.00	0.3682141699E+00	-0.9296837137E+00
10.00	-0.8115772951E+00	-0.5844822734E+00

Tabela 6.2.1: Solução Equação do Pêndulo usando o DASSL

cujas estatísticas do resultado do método são:

2337 avaliações da função;

172 avaliações da jacobiana;

1692 passos computados, sendo que 21 foram rejeitados.

A solução obtida pelo RADAU5 com uma tolerância para o erro local  $ATOL = RTOL = 1.0D - 5$  encontra-se na Tabela 6.2.2,

$x$	$p$	$q$
0.00	0.1000000000E+01	0.0000000000E+00
1.00	0.8795328332E+00	-0.4758128248E+00
2.00	-0.2042128917E+00	-0.9789203122E+00
3.00	-0.9688684308E+00	-0.2476518781E+00
4.00	-0.9991221412E+00	-0.4267367969E-01
5.00	-0.6852563897E+00	-0.7283108499E+00
6.00	0.5646441095E+00	-0.8253279488E+00
7.00	0.9962728515E+00	-0.8655024711E-01
8.00	0.9855612854E+00	-0.1695384548E+00
9.00	0.3685838004E+00	-0.9295984457E+00
10.00	-0.8117276987E+00	-0.5839949224E+00

Tabela 6.2.2: Solução Equação do Pêndulo usando o RADAU5

cujas estatísticas do resultado do método são:

- 538 avaliações da função;
- 53 avaliações da jacobiana;
- 62 passos computados, sendo que 57 foram aceitos e 2 rejeitados;
- 62 decomposições LU foram calculadas, ou seja, quantidade de vezes que o sistema linear é resolvido;
- 160 substituições para frente e para trás de ambos sistemas dados por L e U.

### 6.3 Equação de Van der Pol

Seja a EAD não linear semi explícita de índice zero

$$\begin{aligned} y' &= z \\ \mathcal{E}z' &= (1 - y^2)z - y. \end{aligned} \quad (6.3.1)$$

Para a nossa análise usaremos o intervalo  $[0,2]$  e consideraremos valores iniciais consistentes  $y(0) = 2$ ,  $z(0) = -0.66$ , com  $\mathcal{E} = 1.0D - 6$ .

A solução obtida pelo DASSL com  $ATOL = RTOL = 1.0D - 5$  encontra-se na Tabela 6.3.1,

$x$	$y$	$z$
0.00	0.2000000000E+01	-0.6600000000E+00
0.20	0.1858188832E+01	-0.7575585837E+00
0.40	0.9069492018E+00	-0.9069492018E+00
0.60	0.1484551264E+01	-0.1233128055E+01
0.80	0.1083748147E+01	-0.6207686279E+01
1.00	-0.1863586036E+01	0.7535888277E+00
1.20	-0.1699644783E+01	0.8998517064E+00
1.40	-0.1493272169E+01	0.1214168340E+01
1.60	-0.1120353612E+01	0.4389615940E+01
1.80	0.1868953098E+01	-0.7496773787E+00
2.00	0.1706022520E+01	-0.8929650991E+00

Tabela 6.3.1: Equação de Van der Pol usando o DASSL

cujas estatísticas do resultado do método são:

- 14365 avaliações da função;
- 1233 avaliações da jacobiana;
- 8385 passos computados, sendo que 273 foram rejeitados.

A solução obtida pelo RADAU5 com  $ATOL = RTOL = 1.0D - 5$  encontra-se na Tabela 6.3.2,

$x$	$y$	$z$
0.00	0.2000000000E+01	-0.6600000000E+00
0.20	0.1858204676E+01	-0.7575443999E+00
0.40	0.1693209270E+01	-0.9069428212E+00
0.60	0.1484575388E+01	-0.1233048474E+01
0.80	0.1083921569E+01	-0.6195343946E+01
1.00	-0.1863646716E+01	0.7535455796E+00
1.20	-0.1699713795E+01	0.8997617000E+00
1.40	-0.1493387337E+01	0.1213935300E+01
1.60	-0.1120811876E+01	0.4373845558E+01
1.80	0.1869058815E+01	-0.7496098398E+00
2.00	0.1706167729E+01	-0.8928088236E+00

Tabela 6.3.2: Equação de Van der Pol usando o RADAU5

cujas estatísticas do resultado do método são:

- 3473 avaliações da função;
- 294 avaliações da jacobiana;
- 476 passos computados, sendo que 463 foram aceitos e 12 rejeitados;
- 379 decomposições LU foram calculadas, ou seja, quantidade de vezes que o sistema linear é resolvido;
- 1003 substituições para frente e para trás de ambos sistemas dados por L e U.

## 6.4 Equação da Esfera

Seja a EAD não linear implícita

$$x^2 + y^2 + y'^2 = 1. \quad (6.4.1)$$

Como RADAU5 resolve apenas sistemas na forma semi explícita, reescreveremos (6.4.1) como sendo

$$y' = \pm \sqrt{1 - x^2 - y^2}, \quad (6.4.2)$$

que é a equação da esfera na forma semi explícita.

Para a nossa análise usaremos o intervalo  $[0,1]$  e consideraremos valor inicial consistente  $y(0) = 0$ . Veremos que a equação da esfera tem singularidade quando  $x = \pm 0.7975005D+00$ , também observaremos que o código DASSL passa por esta singularidade o mesmo não ocorre com o código RADAU5.

A solução obtida pelo DASSL, usando a parte positiva da equação (6.4.2), isto é,  $y' = \sqrt{1-x^2-y^2}$ , com  $ATOL = RTOL = 1.0D-6$  encontra-se na Tabela 6.4.1,

$x$	$y$
0.00	0.0000000000E+00
0.10	0.9983347389E-01
0.20	0.1986698677E+00
0.30	0.2955219490E+00
0.40	0.3894216083E+00
0.50	0.4794299565E+00
0.60	0.5646476071E+00
0.70	0.6442235985E+00
0.80	0.7173620816E+00
0.90	0.7833323466E+00
1.00	0.8414759471E+00

Tabela 6.4.1: Equação da Esfera usando o DASSL

cujas estatísticas do resultado do método são:

1327 avaliações da função;

677 avaliações da jacobiana;

1011 passos computados, sendo que 10 foram rejeitados.

A solução obtida pelo DASSL, usando equação implícita (6.4.1), com  $ATOL = RTOL = 1.0D-6$  encontra-se na Tabela 6.4.2,

$x$	$y$
0.00	0.0000000000E+00
0.10	0.9983347389E-01
0.20	0.1986698593E+00
0.30	0.2955218577E+00
0.40	0.3894212026E+00
0.50	0.4794295539E+00
0.60	0.5646474364E+00
0.70	0.6442231276E+00
0.80	0.7173610736E+00
0.90	0.7833314321E+00
1.00	0.8414749838E+00

Tabela 6.4.2: Equação da Esfera usando o DASSL



cujas estatísticas do resultado do método são:

1185 avaliações da função;  
677 avaliações da jacobiana;  
1009 passos computados, e nenhum passo rejeitado.

A solução obtida pelo RADAU5, usando a parte positiva da equação (6.4.2), com  $ATOL = RTOL = 1.0D - 6$  encontra-se na Tabela 6.4.3,

$x$	$y$
0.00	0.0000000000E+00
0.10	0.9966642867E-01
0.20	0.1973223552E+00
0.30	0.2909126948E+00
0.40	0.3782748086E+00
0.50	0.4570282005E+00
0.60	0.5243338768E+00
0.70	0.5762208042E+00

Tabela 6.4.3: Equação da Esfera usando o RADAU5

cujas estatísticas do resultado do método são:

474 avaliações da função;  
42 avaliações da jacobiana;  
114 passos computados, sendo que 65 foram aceitos e 48 rejeitados;  
87 decomposições LU foram calculadas,, ou seja, quantidade de vezes que o sistema linear é resolvido;  
136 substituições para frente e para trás de ambos sistemas dados por L e U.

Como comentamos, no início desta seção, o código RADAU5 para quando  $x = 0.7975005D+00$  e  $h = 0.3455723D-15$ .

## 6.5 Equação Teste

Seja a EAD linear de índice zero

$$y' = \lambda y \quad (6.5.1)$$

Para a nossa análise usaremos o intervalo  $[-10,10]$  com  $\lambda = 50$  e consideraremos valor inicial consistente  $y(0) = 1$ .

Neste exemplo, verificaremos a estabilidade dos métodos.

A solução obtida pelo DASSL com  $ATOL = RTOL = 1.0D - 6$  encontra-se na Tabela 6.5.1,

$x$	$y$	Erro
-10.00	-0.1860210341E-16	0.2610976758E+200
-8.00	-0.2570236933E-16	0.1342041424E+158
-6.00	-0.2365326141E-14	0.4594471877E+116
-4.00	-0.1319363747E-13	0.9533687826E+73
-2.00	-0.4420550209E-11	0.1188295646E+33
0.00	1.0000000000E+01	0.0000000000E+00
2.00	0.2689010223E+44	0.3322328557E-03
4.00	0.7230807133E+87	0.6688877475E-03
6.00	0.1944379863+131	0.1005684491E-02
8.00	0.5228479980+174	0.1342589501E-02
10.00	0.1405949675+218	0.1679588226E-02

Tabela 6.5.1: Equação Teste usando o DASSL

cujas estatísticas do resultado do método são:

- 58279 avaliações da função;
- 879 avaliações da jacobiana;
- 55741 passos computados, sendo que 140 foram rejeitados;

A solução obtida pelo RADAU5 com  $ATOL = RTOL = 1.0D - 6$  encontra-se na Tabela 6.5.2,

$x$	$y$	Erro
-10.00	0.4747760941E-14	0.6663920309E+203
- 8.00	-0.4903333626E-13	0.2560260791E+161
- 6.00	0.4055127546E-14	0.7876786782E+116
- 4.00	0.2045322112E-11	0.1477944393E+76
- 2.00	-0.4479671822E-11	0.1204188262E+33
0.00	0.1000000000E+01	0.0000000000E+00
2.00	0.2688117384E+44	0.9001393152E-07
4.00	0.7225975277E+87	0.2088606472E-06
6.00	0.1942427112+131	0.3691675860E-06
8.00	0.5221472676+174	0.5720042493E-06
10.00	0.1403593230+218	0.7212839936E-06

Tabela 6.5.2: Equação Teste usando o RADAU5

cujas estatísticas do resultado do método são:

- 22042 avaliações da função;
- 2 avaliações da jacobiana;
- 5038 passos computados, sendo que 5038 foram aceitos e nenhum rejeitado;
- 69 decomposições LU foram calculadas,, ou seja, quantidade de vezes que o sistema linear é resolvido;
- 5668 substituições para frente e para trás de ambos sistemas dados por L e U.

# 7 Conclusão

---

---

Neste trabalho estudamos EAD, apenas, de índices zero ou um. O motivo disto é que queríamos fazer comparações com os resultados dos métodos BDF e RKI usando os códigos DASSL e RADAU5, respectivamente.

Os métodos de passos múltiplos, em particular o método BDF, aplicados em EAD semi explícita de índice um são estáveis e convergentes para EAD com a mesma ordem de convergência de uma EDO.

Muitas vezes os métodos de Runge-Kutta Implícito não alcançam a mesma ordem de exatidão em EAD como o fazem em EDO.

Vimos que o os códigos DASSL e RADAU5 podem ser aplicados para resolver sistemas de EAD de índices zero ou um. Ambos os códigos resolvem os sistemas não lineares usando o método de Newton com algumas modificações.

Ao fazermos a análise dos resultados dos exemplos apresentados no Capítulo 6 observamos que:

*i)* O código DASSL resolve sistemas de EAD tanto na forma semi explícita quanto na forma implícita, o mesmo não ocorre com o RADAU5, que resolve apenas sistemas de EAD na forma semi explícita;

*ii)* O DASSL apresenta melhores resultados quando usamos a EAD na forma implícita, como pode ser visto na equação da esfera. Também, pode ser observado que o custo na forma implícita é menor;

*iii)* No caso da equação da esfera, o código RADAU5 é muito mais estável perto da singularidade, enquanto que o DASSL passa por esta singularidade.

Ao rodarmos os exemplos numéricos vimos que o código RADAU5 exige muito mais informações, sobre a EAD, que o DASSL, tornando assim o código mais estável, em compensação não é necessária tantas informações para utilizar o código DASSL, facilitando o trabalho do usuário.

Os códigos em discussão podem ser calculados usando o seguinte processo:

- o próprio código calcula as derivadas parciais;
- o usuário fornece as derivadas parciais,

em ambos os casos foi observado que não há diferenças significativas do resultado.

Para os exemplos apresentados no Capítulo 6 podemos observar que o custo do método BDF em relação ao método de Runge-Kutta é mais alto.

# 8 Bibliografia

---

---

[BRENAN - 89]: BRENAN, K. E., CAMPBELL, S. L. & PETZOLD, L. R., *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*, North-Holland (1989).

[FRANK - 85a]: FRANK, R., SCHNEID, J. & UEBERHUBER, C.W., *Stability properties of implicit Runge-Kutta methods*, SIAM J. Numer. Anal., 22 (1985), 497 - 514.

[FRANK - 85b]: FRANK, R., SCHNEID, J. & UEBERHUBER, C.W., *Order results for implicit Runge-Kutta methods applied to stiff systems*, SIAM J. Numer. Anal., 22 (1985), 515 - 534.

[GEAR - 71]: GEAR, C. W., *The simultaneous numerical solution of differential-algebraic equations*, IEEE Trans. Circuit Theory, CT - 18 (1971), 89 - 95.

[HAIRER - 89]: HAIRER, E., LUBICH, C. & ROCHE, M., *The Numerical Solution of Differential-Algebraic Systems by Runge-Kutta Methods*, Springer-Verlag (1989).

[HAIRER - 91]: HAIRER, E. & WANNER, G., *Solving Ordinary Differential Equations II - Stiff and Differential-Algebraic Problems*, Springer-Verlag (1991).

[KROGH - 74]: KROGH, F.T., *Changing step in the integration of differential equations using modified divided differences*, Proc. Conf. Num. Solution of ODEs, Lecture Notes in Mathematics No. 362, Springer-Verlag, New York, (1974).

- [LAMBERT - 73]: LAMBERT, J. D., *Computational Methods in Ordinary Differential Equations*, Wiley (1973).
- [PETZOLD - 82]: PETZOLD, L., *Differential / algebraic equations are not ODEs*, SIAM J. Sci. Stat. Comput., 3 (1982), 367 - 384.
- [PETZOLD - 86]: PETZOLD, L., *Order results for implicit Runge-Kutta methods applied to differential / algebraic systems*, SIAM J. Numer. Anal., 23 (1986), 836 - 852.
- [PIRES - 94]: PIRES, A.A., *Métodos Numéricos para a Solução de Equações Algébrico-Diferenciais*, Instituto de Ciências Matemáticas de São Carlos, da Universidade de São Paulo, São Carlos, (1994).
- [ROMEIRO - 94]: ROMEIRO, N.M.L., *Análise Comparativa de Métodos Numéricos para Equações Algébrico-Diferenciais*, Instituto de Ciências Matemáticas de São Carlos, da Universidade de São Paulo, São Carlos, (1994).
- [SHAMPINE - 75]: L.F. SHAMPINE, L.F. & GORDON, M.K., *Computer Solution of Ordinary Differential Equations*, W.H. Freeman and Co., (1975).

# NOTAS DO ICMSC

## SÉRIE COMPUTAÇÃO

- 016/94 MAGALHÃES, A.L.C.C.; SIQUEIRA, M.F.; OLIVEIRA, M.C.F. - Operadores de Euler na modelagem por fronteira: conceito, aplicação, estudos de casos.
- 015/94 ODA, C.S.; MOREIRA, E.S. - ASNMP graphical network monitor with automatic topology discovery.
- 014/94 FELIPE, L.S.G.; FRANCO, N.M.B. - Sobre a ordem de convergência para as equações integrais de volterra de segunda espécie tipo Abel com soluções não suaves.
- 013/94 PIMENTEL, M.G.C. - A framework for user-hypertext interaction.
- 012/94 TURINE, M.A.S.; MENDES, M.D.C.; NUNES, M.G.V. - TEGRAM: a geometry tutoring system based on Tangram.
- 011/94 SPOLON, R.; SPOLON, R.; SANTANA, M.J.; SANTANA, R.H.C. - Desenvolvimento de um gerador de aplicação para simulação de sistemas discretos.
- 010/94 SAWAKI, J.; MONARD, M.C.; RODRIGUES, S.R. SABNAG: - um sistema baseado em conhecimento para suporte aos usuários da biblioteca NAG.
- 009/94 NICOLETTI, M.C.; MONARD, M.C. - Learning restricted Horn clauses: some considerations on the ij-determination concept.
- 008/94 TOME, M.F.; DUFFY, B. GENSMAC: - a numerical method for solving unsteady non-newtonian free surface flows.
- 007/94 TOME, M.F.; MCKEE, S. - Numerical simulation of viscous fluid: buckling of planar jets.