Soluzioni Analitiche e Numeriche Applicate all'Ingegneria Ambientale

Massimiliano Martinelli
massimiliano.martinelli@gmail.com

Università Politecnica delle Marche, Ancona
Facoltà di Ingegneria

11-12 Febbraio 2009

Equazioni Differenziali Ordinarie

Definizione

Sia $I = [x_0, b]$ (o $I = [x_0, +\infty)$) e $y\colon I \to C^p(I)$. Sia inoltre $f : I \times \mathbb{R}^p \to \mathbb{R}$ regolare rispetto ad ognuno dei suoi argomenti. Si definisce Equazione Differenziale Ordinaria (ODE, *Ordinary Differential Equation*) un'equazione del tipo

$$y^{(p)}(x) = f(x, y(x), y'(x), \ldots, y^{(p-1)}(x))$$

dove p è detto *ordine* dell'equazione. $\Rightarrow$ Equazione in cui la derivata p-esima della funzione incognita $y(x)$ dipende da x, $y(x)$ e da tutte le sue derivate fino all'ordine $p-1$

Equazioni Differenziali Ordinarie

Osservazione

L'equazione precedente si può riscrivere come un sistema di ODE di ordine 1

$$\mathbf{y}'(x) = \mathbf{F}(x, \mathbf{y})$$

dove abbiamo posto $\mathbf{y}: I \to \mathbb{R}^p$ e $\mathbf{F}: I \times \mathbb{R}^p \to \mathbb{R}^p$ nel seguente modo

$$\mathbf{y}(x) = \begin{pmatrix} y(x) \\ y'(x) \\ \vdots \\ y^{(p-1)}(x) \end{pmatrix} \qquad \mathbf{F}(x, \mathbf{y}(x)) = \begin{pmatrix} y_2 \\ y_3 \\ \vdots \\ y_p \\ f(x, \mathbf{y}(x)) \end{pmatrix}$$

- Per definire in modo completo un problema differenziale occorre associare all'ODE in esame alcune condizioni supplementari
- Si possono distinguere due tipi diversi di problemi (che devono essere trattati separatamente. Appaiono infatti delle differenze sostanziali sia per l'aspetto teorico che per il trattamento numerico)

- Per definire in modo completo un problema differenziale occorre associare all'ODE in esame alcune condizioni supplementari
- Si possono distinguere due tipi diversi di problemi (che devono essere trattati separatamente. Appaiono infatti delle differenze sostanziali sia per l'aspetto teorico che per il trattamento numerico)

Problemi di Cauchy (o ai valori iniziali)

$$\begin{cases} y^{(p)}(x) = f(x, y(x), y'(x), \ldots, y^{(p)}(x)) & \forall x \in I \\ y(x_0) = y_0 \quad , \quad y'(x_0) = y'_0 \quad , \ldots , \quad y^{(p-1)}(x_0) = y_0^{(p-1)} \end{cases}$$

oppure

$$\begin{cases} \mathbf{y}'(x) = \mathbf{F}(x, \mathbf{y}(x)) & \forall x \in I \\ \mathbf{y}(x_0) = \mathbf{y}_0 \end{cases} \tag{1}$$

- Per definire in modo completo un problema differenziale occorre associare all'ODE in esame alcune condizioni supplementari
- Si possono distinguere due tipi diversi di problemi (che devono essere trattati separatamente. Appaiono infatti delle differenze sostanziali sia per l'aspetto teorico che per il trattamento numerico)

Problemi di Cauchy (o ai valori iniziali)

$$\begin{cases} y^{(p)}(x) = f(x, y(x), y'(x), \ldots, y^{(p)}(x)) & \forall x \in I \\ y(x_0) = y_0 \quad , \quad y'(x_0) = y'_0 \quad , \ldots , \quad y^{(p-1)}(x_0) = y_0^{(p-1)} \end{cases}$$

oppure

$$\begin{cases} \mathbf{y}'(x) = \mathbf{F}(x, \mathbf{y}(x)) & \forall x \in I \\ \mathbf{y}(x_0) = \mathbf{y}_0 \end{cases} \tag{1}$$

Problemi ai limiti

Problemi in cui si impongono delle condizioni in alcuni punti distinti

Problema di Cauchy: esistenza e unicità

- Siano $\mathbf{F}\colon I \times \mathbb{R}^p \to \mathbb{R}^p$ e $l\colon I \to \mathbb{R}$ due funzioni continue tali che per ogni $x \in I$ e per ogni $\mathbf{y}, \mathbf{z} \in \mathbb{R}^p$ valga

$$\langle \mathbf{F}(x,\mathbf{y}) - \mathbf{F}(x,\mathbf{z}), \mathbf{y} - \mathbf{z} \rangle \leq l(x)||\mathbf{y} - \mathbf{z}||^2$$

Allora il problema della ricerca di una funzione $\mathbf{y}\colon I \to \mathbb{R}^p$ continua e derivabile che verifica (1) ammette una ed una sola soluzione

Problema di Cauchy: esistenza e unicità

- Siano $\mathbf{F}\colon I \times \mathbb{R}^p \to \mathbb{R}^p$ e $l\colon I \to \mathbb{R}$ due funzioni continue tali che per ogni $x \in I$ e per ogni $\mathbf{y}, \mathbf{z} \in \mathbb{R}^p$ valga

$$\langle \mathbf{F}(x,\mathbf{y}) - \mathbf{F}(x,\mathbf{z}), \mathbf{y} - \mathbf{z} \rangle \leq l(x)||\mathbf{y} - \mathbf{z}||^2$$

Allora il problema della ricerca di una funzione $\mathbf{y}\colon I \to \mathbb{R}^p$ continua e derivabile che verifica (1) ammette una ed una sola soluzione

Corollario: teorema di Cauchy-Lipschitz

- Sia $\mathbf{F}\colon I \times \mathbb{R}^p \to \mathbb{R}^p$ una funzione continua e $L > 0$ tali che per ogni $x \in I$ e per ogni $\mathbf{y}, \mathbf{z} \in \mathbb{R}^p$ valga

$$||\mathbf{F}(x,\mathbf{y}) - \mathbf{F}(x,\mathbf{z})|| \leq L||\mathbf{y} - \mathbf{z}||$$

Allora il problema della ricerca di una funzione $\mathbf{y}\colon I \to \mathbb{R}^p$ continua e derivabile che verifica (1) ammette una ed una sola soluzione

- Nel caso di un problema a valori iniziali, i metodi numerici definiti nel caso scalare si generalizzano senza difficolta al caso vettoriale

- Nel caso di un problema a valori iniziali, i metodi numerici definiti nel caso scalare si generalizzano senza difficolta al caso vettoriale
- Consideriamo il problema seguente:

Trovare $y\colon I \to C^1(I)$ tale che

$$\begin{cases} y'(x) = f(x, y(x)) & \forall x \in I \\ y(x_0) = y_0 \end{cases} \tag{2}$$

- Nel caso di un problema a valori iniziali, i metodi numerici definiti nel caso scalare si generalizzano senza difficolta al caso vettoriale
- Consideriamo il problema seguente:

  Trovare $y\colon I \to C^1(I)$ tale che

$$\begin{cases} y'(x) = f(x, y(x)) & \forall x \in I \\ y(x_0) = y_0 \end{cases} \tag{2}$$

- Il problema differenziale precedente è equivalente all'equazione integrale seguente:

  Trovare $y\colon I \to C^1(I)$ tale che

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\, dt \tag{3}$$

- Nel caso di un problema a valori iniziali, i metodi numerici definiti nel caso scalare si generalizzano senza difficolta al caso vettoriale
- Consideriamo il problema seguente:

  Trovare $y\colon I \to C^1(I)$ tale che

$$\begin{cases} y'(x) = f(x, y(x)) & \forall x \in I \\ y(x_0) = y_0 \end{cases} \tag{2}$$

- Il problema differenziale precedente è equivalente all'equazione integrale seguente:

  Trovare $y\colon I \to C^1(I)$ tale che

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\, dt \tag{3}$$

- Questa equivalenza si può sfruttare anche a livello numerico, definendo un opportuno metodo di discretizzazione

Problema ben posto

- Consideriamo il seguente problema:

$$\text{Dato } d \text{ trovare } x \text{ tale che} \quad F(x,y) = 0 \tag{4}$$

- Si suppone che il problema (4) sia *ben posto*, ovvero che

$$\forall \eta > 0, \quad \exists K(\eta, d) \quad \text{tale che} \quad ||\delta d|| < \eta \quad \Rightarrow \quad ||\delta x|| \leq K(\eta, d)||\delta d||$$

- Se il problema (4) ammette un'unica soluzione, allora esiste necessariamente una funzione *risolvente* G tale che

$$x = G(d) \quad \text{cioè} \quad F(G(d), d) = 0$$

Consistenza

- Un metodo numerico per la soluzione approssimata del problema $F(x,d)=0$ consiste in una sequenza di problemi approssimati

$$F_n(x_n,d_n)=0 \tag{5}$$

- Si spera che $\lim_{n\to\infty} x_n = x$ (convergenza alla soluzione esatta)
- È necessario che $d_n \to d$ e che F_n "approssima" F per $n\to\infty$
- Se il dato d è ammissibile per F_n diciamo che il metodo numerico (5) è *consistente* se

$$\lim_{n\to\infty} F_n(x,d) = \lim_{n\to\infty} \left[F_n(x,d) - F(x,d)\right] = 0$$

  dove x è la soluzione di (4) corrispondente al dato d
- Il metodo è *fortemente consistente* se $F_n(x,d)=0$ per qualsiasi valore di n

Stabilità

- Si richiede che, per n fissato, il problema

$$F_n(x_n, d_n) = 0$$

ammette un'unica soluzione x_n corrispondente al dato d_n e tale soluzione x_n dipende dai dati in maniera continua, ovvero

$$\forall \eta > 0\,,\ \exists K_n(\eta, d_n) \text{ tale che } ||\delta d_n|| < \eta \Rightarrow ||\delta x_n|| \leq K_n(\eta, d_n)||\delta d_n||$$

Stabilità

- Si richiede che, per n fissato, il problema

$$F_n(x_n, d_n) = 0$$

ammette un'unica soluzione x_n corrispondente al dato d_n e tale soluzione x_n dipende dai dati in maniera continua, ovvero

$$\forall \eta > 0\,,\ \exists K_n(\eta, d_n) \text{ tale che } ||\delta d_n|| < \eta \Rightarrow ||\delta x_n|| \leq K_n(\eta, d_n)||\delta d_n||$$

- In altre parole, significa che la soluzione $x_n + \delta x_n$ del problema perturbato

$$F_n(x_n + \delta x_n, d_n + \delta d_n) = 0$$

è non si discosta troppo dalla soluzione x_n

Convergenza

- Il metodo numerico $F_n(x_n, d_n) = 0$ è *convergente* se e solo se

$$\forall \varepsilon > 0 \quad \exists n_0(\varepsilon), \quad \exists \delta(n_0, \varepsilon) > 0 \quad \text{tale che}$$
$$\forall n > n_0(\varepsilon), \quad \forall ||\delta d_n|| < \delta(n_0, \varepsilon) \quad \Rightarrow \quad ||x(d) - x_n(d + \delta d_n)|| \leq \varepsilon \qquad (6)$$

dove $x_n(d + \delta d_n)$ è la soluzione del problema $F_n(x_n, d + \delta d_n) = 0$

Convergenza

- Il metodo numerico $F_n(x_n, d_n) = 0$ è *convergente* se e solo se

$$\forall \varepsilon > 0 \quad \exists n_0(\varepsilon), \quad \exists \delta(n_0, \varepsilon) > 0 \quad \text{tale che}$$
$$\forall n > n_0(\varepsilon), \quad \forall ||\delta d_n|| < \delta(n_0, \varepsilon) \quad \Rightarrow \quad ||x(d) - x_n(d + \delta d_n)|| \leq \varepsilon \qquad (6)$$

  dove $x_n(d + \delta d_n)$ è la soluzione del problema $F_n(x_n, d + \delta d_n) = 0$
- Se poniamo $e_n = x(d) - x_n(d + \delta d_n)$, la definizione precedente equivale a dire che

$$\lim_{n \to \infty} e_n = 0$$

Convergenza

- Il metodo numerico $F_n(x_n, d_n) = 0$ è *convergente* se e solo se

$$\forall \varepsilon > 0 \quad \exists n_0(\varepsilon), \quad \exists \delta(n_0, \varepsilon) > 0 \quad \text{tale che}$$
$$\forall n > n_0(\varepsilon), \quad \forall ||\delta d_n|| < \delta(n_0, \varepsilon) \quad \Rightarrow \quad ||x(d) - x_n(d + \delta d_n)|| \leq \varepsilon \qquad (6)$$

  dove $x_n(d + \delta d_n)$ è la soluzione del problema $F_n(x_n, d + \delta d_n) = 0$
- Se poniamo $e_n = x(d) - x_n(d + \delta d_n)$, la definizione precedente equivale a dire che

$$\lim_{n \to \infty} e_n = 0$$

- Per un problema ben posto, la condizione (6) equivale a richiedere che

$$||x(d + \delta d_n) - x_n(d + \delta d_n)|| \leq \frac{\varepsilon}{2}$$

Relazioni tra consistenza, stabilità e convergenza

Convergenza $\Rightarrow$ Stabilità

Sia $F(x,d) = 0$ un problema ben posto. Se il problema numerico $F_n(x_n, d_n) = 0$ è convergente, allora è stabile.

Relazioni tra consistenza, stabilità e convergenza

Convergenza $\Rightarrow$ Stabilità

Sia $F(x,d) = 0$ un problema ben posto. Se il problema numerico $F_n(x_n, d_n) = 0$ è convergente, allora è stabile.

Consistenza + Stabilità $\Rightarrow$ Convergenza

Sia $F(x,d) = 0$ un problema ben posto. Se il problema numerico $F_n(x_n, d_n) = 0$ è consistente e stabile allora è convergente.

Relazioni tra consistenza, stabilità e convergenza

Convergenza $\Rightarrow$ Stabilità

Sia $F(x,d) = 0$ un problema ben posto. Se il problema numerico $F_n(x_n, d_n) = 0$ è convergente, allora è stabile.

Consistenza + Stabilità $\Rightarrow$ Convergenza

Sia $F(x,d) = 0$ un problema ben posto. Se il problema numerico $F_n(x_n, d_n) = 0$ è consistente e stabile allora è convergente.

Teorema di Equivalenza o di Lax-Richtmyer

Per un metodo numerico consistente, la stabilità è equivalente alla convergenza

Discretizzazione del problema

- Fissato l'intervallo di integrazione $I = [x_0, x_F]$, si definisce una successione di *nodi di discretizzazione*

$$x_i = x_{i-1} + h_i \qquad \text{con} \quad i = 1, 2, \ldots$$

Discretizzazione del problema

- Fissato l'intervallo di integrazione $I = [x_0, x_F]$, si definisce una successione di *nodi di discretizzazione*

$$x_i = x_{i-1} + h_i \qquad \text{con} \quad i = 1, 2, \ldots$$

Per iniziare, assumiamo un *passo di discretizzazione* costante, cioè $h_i = h$ e quindi

$$x_i = x_0 + ih$$

Discretizzazione del problema

- Fissato l'intervallo di integrazione $I = [x_0, x_F]$, si definisce una successione di *nodi di discretizzazione*

$$x_i = x_{i-1} + h_i \qquad \text{con} \quad i = 1, 2, \ldots$$

Per iniziare, assumiamo un *passo di discretizzazione* costante, cioè $h_i = h$ e quindi

$$x_i = x_0 + ih$$

- Usiamo la notazione compatta $y_i = y(x_i)$

Metodi di sviluppo in serie di Taylor

- Si suppone che $y(x)$, la soluzione esatta del problema (2), sia di classe $C^{k+1}(I)$

Metodi di sviluppo in serie di Taylor

- Si suppone che $y(x)$, la soluzione esatta del problema (2), sia di classe $C^{k+1}(I)$
- Lo sviluppo in serie di Taylor all'ordine k di $y_{n+1} = y(x_{n+1}) = y(x_n + h)$ nell'intorno del punto x_n fornisce:

$$y_{n+1} = y_n + \sum_{i=1}^{k} \frac{h^i}{i!} y^{(i)}(x_n) + \frac{h^{k+1}}{(k+1)!} y^{(k+1)}(\xi_k(x_n))$$

$$= y_n + h\left(\sum_{i=1}^{k} \frac{h^{i-1}}{i!} y^{(i)}(x_n) + \frac{h^k}{(k+1)!} y^{(k+1)}(\xi_k(x_n)) \right)$$

dove $\xi_k(x_n) \in (x_n, x_{n+1})$.

Metodi di sviluppo in serie di Taylor

- Ricordando che $y'(x) = f(x,y(x)) = \varphi_1(x,y;f)$ abbiamo

$$\begin{aligned} y''(x) &= f_x(x,y) + f(x,y) f_y(x,y) && = \varphi_2(x,y;f) \\ y'''(x) &= f_{xx} + f\left(2f_{xy} + f_{yy}\right) + f_y\left(f_x + f f_y\right) && = \varphi_3(x,y;f) \\ &\dots \\ y^{(i)}(x) &= \dots && = \varphi_i(x,y;f) \end{aligned}$$

Perciò possiamo scrivere

$$\begin{aligned} y_{n+1} &= y_n + h\sum_{i=1}^{k} \frac{h^{i-1}}{i!} y^{(i)}(x_n) + \frac{h^{k+1}}{(k+1)!} y^{(k+1)}(\xi_k(x_n)) \\ &= y_n + h\sum_{i=1}^{k} \frac{h^{i-1}}{i!} \varphi_i(x_n,y_n;f) + \frac{h^{k+1}}{(k+1)!} y^{(k+1)}(\xi_k(x_n)) \end{aligned}$$

- Supponiamo ora di scegliere una funzione $\Phi(x,y;h,f)$ tale che

$$\lim_{h\to 0} \Phi(x,y;h,f) = f(x,y)$$

e di scrivere l'algoritmo ad un passo

$$\begin{cases} u_0 = y_0 \\ u_{n+1} = u_n + h\Phi(x_n,u_n;h,f) \end{cases} \tag{7}$$

- Supponiamo ora di scegliere una funzione $\Phi(x,y;h,f)$ tale che

$$\lim_{h\to 0} \Phi(x,y;h,f) = f(x,y)$$

e di scrivere l'algoritmo ad un passo

$$\begin{cases} u_0 = y_0 \\ u_{n+1} = u_n + h\Phi(x_n,u_n;h,f) \end{cases} \tag{7}$$

Analisi dell'errore

- Si tratta di valutare la differenza tra la soluzione esatta y_{n+1} e la soluzione (approssimata) u_{n+1} calcolata con l'algoritmo (7)

$$e_{n+1} = y_{n+1} - u_{n+1}$$

- Supponiamo ora di scegliere una funzione $\Phi(x,y;h,f)$ tale che

$$\lim_{h\to 0} \Phi(x,y;h,f) = f(x,y)$$

e di scrivere l'algoritmo ad un passo

$$\begin{cases} u_0 = y_0 \\ u_{n+1} = u_n + h\Phi(x_n,u_n;h,f) \end{cases} \tag{7}$$

Analisi dell'errore

- Si tratta di valutare la differenza tra la soluzione esatta y_{n+1} e la soluzione (approssimata) u_{n+1} calcolata con l'algoritmo (7)

$$e_{n+1} = y_{n+1} - u_{n+1}$$

- Sia u^*_{n+1} la soluzione ottenuta applicando un passo dell'algoritmo (7) partendo dal dato iniziale y_n, ovvero

$$u^*_{n+1} = y_n + h\Phi(x_n,y_n;h,f)$$

- L'errore globale e_{n+1} si può decomporre come

$$e_{n+1} = \underbrace{(y_{n+1} - u^*_{n+1})}_{E_{\text{loc}}} + \underbrace{(u^*_{n+1} - u_{n+1})}_{E_{\text{prop}}}$$

- Notiamo che la soluzione esatta può anche essere scritta come

$$y_{n+1} = y_n + h\Phi(x_n, y_n; h, f) + h\tau^k_n(y; h)$$

perciò $E_{\text{loc}} = y_{n+1} - u^*_{n+1} = h\tau^k_n(h; y)$ rappresenta il contributo "locale" dell'errore.

- $e_{\text{loc}} = \frac{1}{h}E_{\text{loc}} = \tau^k_n(h; y)$ è chiamato *errore locale di troncamento*
- Una condizione necessaria per la *convergenza*, cioè

$$\forall n \qquad ||e_n|| \leq C(h) \quad \text{con} \quad \lim_{h\to 0} C(h) = 0$$

è che l'errore locale di troncamento sia infinitesimo con il passo h, ovvero che $\lim_{h\to 0} e_{\text{loc}}(h) = 0$

- La proprietà $\lim_{h\to 0} e_{\text{loc}}(h) = 0$ esprime la *consistenza* del metodo
- Si dice che lo schema è di ordine p se $e_{\text{loc}}(h) = O(h^p)$ per $h \to 0$
- Per un metodo di sviluppo in serie di Taylor di ordine k in cui $\Phi(x,y;h,f) = \sum_{i=1}^{k} \frac{h^{i-1}}{i!}\varphi_i(x,y)$ abbiamo $e_{\text{loc}} = O(h^k)$

$$e_{n+1} = \underbrace{(y_{n+1} - u^*_{n+1})}_{E_{\text{loc}}} + \underbrace{(u^*_{n+1} - u_{n+1})}_{E_{\text{prop}}}$$

- $E_{\text{prop}} = e_n + h\big(\Phi(x_n,y_n;h,f) - \Phi(x_n,u_n;h,f)\big)$ è dovuto alla propagazione degli errori
- Ad ogni passo un nuovo errore locale viene introdotto, per poi essere successivamente propagato
- L'errore globale è formato dall'accumulo successivo degli errori locali
- La proprietà di *stabilità* è legata al fatto che questo accumulo non deve essere "eccessivo" quando $h \to 0$

Metodo di Eulero in avanti

- Nel caso particolare $\Phi(x_n, y_n; h, f) = f(x, y)$ si ottiene il *metodo di Eulero in avanti (o esplicito)*

$$\boxed{u_{n+1} = u_n + hf(x_n, u_n) \quad \text{per } n > 0}$$

Metodo di Eulero in avanti

- Nel caso particolare $\Phi(x_n, y_n; h, f) = f(x, y)$ si ottiene il *metodo di Eulero in avanti (o esplicito)*

$$\boxed{u_{n+1} = u_n + hf(x_n, u_n) \quad \text{per } n > 0}$$

- Se f è lipschitziana rispetto al secondo argomento, cioè $\forall x \in I$ verifica

$$|f(x, y) - f(x, z)| \leq L|y - z| \qquad \forall y, z \in \mathbb{R}$$

e inoltre $y \in C^2(I)$ tale che $M = \max_{\xi \in I} |y''(\xi)| < +\infty$ allora abbiamo la seguente stima dell'errore:

$$|e_n| \leq \frac{Mh}{2L}\left[e^{L(x_n - x_0)} - 1\right] \qquad \forall x \geq 0$$

Metodo di Eulero in avanti

- Nel caso particolare $\Phi(x_n, y_n; h, f) = f(x, y)$ si ottiene il *metodo di Eulero in avanti (o esplicito)*

$$\boxed{u_{n+1} = u_n + hf(x_n, u_n) \quad \text{per } n > 0}$$

- Se f è lipschitziana rispetto al secondo argomento, cioè $\forall x \in I$ verifica

$$|f(x, y) - f(x, z)| \leq L|y - z| \qquad \forall y, z \in \mathbb{R}$$

e inoltre $y \in C^2(I)$ tale che $M = \max_{\xi \in I} |y''(\xi)| < +\infty$ allora abbiamo la seguente stima dell'errore:

$$|e_n| \leq \frac{Mh}{2L}\left[e^{L(x_n - x_0)} - 1\right] \qquad \forall x \geq 0$$

- Il metodo di Eulero esplicito è allora convergente con $e_n = O(h)$

Metodo di Eulero in avanti

- Nel caso particolare $\Phi(x_n, y_n; h, f) = f(x, y)$ si ottiene il *metodo di Eulero in avanti (o esplicito)*

$$\boxed{u_{n+1} = u_n + hf(x_n, u_n) \quad \text{per } n > 0}$$

- Se f è lipschitziana rispetto al secondo argomento, cioè $\forall x \in I$ verifica

$$|f(x,y) - f(x,z)| \leq L|y - z| \qquad \forall y, z \in \mathbb{R}$$

e inoltre $y \in C^2(I)$ tale che $M = \max_{\xi \in I} |y''(\xi)| < +\infty$ allora abbiamo la seguente stima dell'errore:

$$|e_n| \leq \frac{Mh}{2L}\left[e^{L(x_n - x_0)} - 1\right] \qquad \forall x \geq 0$$

- Il metodo di Eulero esplicito è allora convergente con $e_n = O(h)$
- Stesso ordine di infinitesimo tra errore globale ed errore di troncamento locale

Metodo di Eulero in avanti: influenza degli errori di arrotondamento

- Tenendo conto degli errori di arrotondamento, la soluzione $\tilde{u}_{n+1}$ fornita dal calcolatore con il metodo di Eulero in avanti si può scrivere come:

$$\begin{cases} \tilde{u}_0 = y_0 + \eta_0 \\ \tilde{u}_{n+1} = \tilde{u}_n + hf(x_n, \tilde{u}_n) + \eta_{n+1} \quad \text{per } n \geq 0 \end{cases}$$

in cui η_i (con $i \geq 0$) è l'errore dovuto all'arrotondamento

Metodo di Eulero in avanti: influenza degli errori di arrotondamento

- Tenendo conto degli errori di arrotondamento, la soluzione $\tilde{u}_{n+1}$ fornita dal calcolatore con il metodo di Eulero in avanti si può scrivere come:

$$\begin{cases} \tilde{u}_0 = y_0 + \eta_0 \\ \tilde{u}_{n+1} = \tilde{u}_n + hf(x_n, \tilde{u}_n) + \eta_{n+1} \quad \text{per } n \geq 0 \end{cases}$$

  in cui η_i (con $i \geq 0$) è l'errore dovuto all'arrotondamento

- Si ottiene allora la seguente maggiorazione dell'errore

$$|y_n - \tilde{u}_n| \leq e^{L(x_n - x_0)} \left[|\eta_0| + \frac{1}{L} \left(\frac{Mh}{2} + \frac{\eta}{h} \right) \right] \quad \forall n \geq 0$$

  con $\eta = \max_{1 \leq i \leq n} |\eta_i|$

Metodo di Eulero in avanti: influenza degli errori di arrotondamento

- Tenendo conto degli errori di arrotondamento, la soluzione $\tilde{u}_{n+1}$ fornita dal calcolatore con il metodo di Eulero in avanti si può scrivere come:

$$\begin{cases} \tilde{u}_0 = y_0 + \eta_0 \\ \tilde{u}_{n+1} = \tilde{u}_n + hf(x_n, \tilde{u}_n) + \eta_{n+1} \quad \text{per } n \geq 0 \end{cases}$$

  in cui η_i (con $i \geq 0$) è l'errore dovuto all'arrotondamento
- Si ottiene allora la seguente maggiorazione dell'errore

$$|y_n - \tilde{u}_n| \leq e^{L(x_n - x_0)} \left[|\eta_0| + \frac{1}{L} \left(\frac{Mh}{2} + \frac{\eta}{h} \right) \right] \quad \forall n \geq 0$$

  con $\eta = \max_{1 \leq i \leq n} |\eta_i|$
- Esiste allora un valore ottimale di h, h_{opt}, in corrispondenza del quale l'errore $|y_n - \tilde{u}_n|$ è minimo:
  - per $h < h_{\text{opt}}$ l'errore di arrotondamento diventa preponderante
  - per $h > h_{\text{opt}}$ il contributo dell'errore di arrotondamento nell'errore globale diventa trascurabile

- I metodi che utilizzano lo sviluppo in serie di Taylor sono molto interessanti da un punto di vista teorico
  - Nozione di ordine di precisione
  - Analisi accurata dell'errore di discretizzazione

- I metodi che utilizzano lo sviluppo in serie di Taylor sono molto interessanti da un punto di vista teorico
  - Nozione di ordine di precisione
  - Analisi accurata dell'errore di discretizzazione
- L'uso di un metodo di sviluppo in serie di Taylor non è in generale molto conveniente: ad ogni passo richiede la valutazione di f e di tutte le sue derivate parziali

- I metodi che utilizzano lo sviluppo in serie di Taylor sono molto interessanti da un punto di vista teorico
  - Nozione di ordine di precisione
  - Analisi accurata dell'errore di discretizzazione
- L'uso di un metodo di sviluppo in serie di Taylor non è in generale molto conveniente: ad ogni passo richiede la valutazione di f e di tutte le sue derivate parziali
- Il metodo più semplice da implementare (Eulero in avanti) converge lentamente (ordine 1)

- I metodi che utilizzano lo sviluppo in serie di Taylor sono molto interessanti da un punto di vista teorico
  - Nozione di ordine di precisione
  - Analisi accurata dell'errore di discretizzazione
- L'uso di un metodo di sviluppo in serie di Taylor non è in generale molto conveniente: ad ogni passo richiede la valutazione di f e di tutte le sue derivate parziali
- Il metodo più semplice da implementare (Eulero in avanti) converge lentamente (ordine 1)

- L'obiettivo è quello di avere un metodo di ordine > 1 senza avere gli svantaggi dei metodi di sviluppo in serie di Taylor di ordine elevato

- I metodi che utilizzano lo sviluppo in serie di Taylor sono molto interessanti da un punto di vista teorico
  - Nozione di ordine di precisione
  - Analisi accurata dell'errore di discretizzazione
- L'uso di un metodo di sviluppo in serie di Taylor non è in generale molto conveniente: ad ogni passo richiede la valutazione di f e di tutte le sue derivate parziali
- Il metodo più semplice da implementare (Eulero in avanti) converge lentamente (ordine 1)

- L'obiettivo è quello di avere un metodo di ordine > 1 senza avere gli svantaggi dei metodi di sviluppo in serie di Taylor di ordine elevato
- Si possono distinguere due grandi famiglie
  - Metodi di Runge-Kutta
  - Metodi a più passi (o *multistep*)

Varianti dello schema di Eulero

- Abbiamo visto che, definito il metodo ad un passo

$$u_{n+1} = u_n + h\Phi(x_n, u_n; h, f)$$

  la soluzione esatta si può scrivere come

$$y_{n+1} = y_n + h\Phi(x_n, y_n; h, f) + h\tau_n(h; y)$$

- Nel caso del metodo di Eulero $\Phi(x, y; h, f) = f(x, y)$ e quindi $\tau_n(h; y) = O(h)$

Varianti dello schema di Eulero

- Abbiamo visto che, definito il metodo ad un passo

$$u_{n+1} = u_n + h\Phi(x_n, u_n; h, f)$$

  la soluzione esatta si può scrivere come

$$y_{n+1} = y_n + h\Phi(x_n, y_n; h, f) + h\tau_n(h; y)$$

- Nel caso del metodo di Eulero $\Phi(x, y; h, f) = f(x, y)$ e quindi $\tau_n(h; y) = O(h)$
- L'idea è quella di scegliere $\Phi(x, y; h, f)$ in modo tale che $\tau_n(h; y)$ abbia ordine di infinitesimo più elevato possibile
- Scegliendo $\Phi(x, y; h, f) = \sum_{i=1}^{k} \frac{h^{i-1}}{i!} \varphi_i(x, y; f)$ si ha che $\tau_n(h; y) = O(h^k)$, ma questo richiede il calcolo delle derivate parziali di $f(x, y)$

- Nel caso $k = 2$ abbiamo

$$\begin{aligned}\Phi(x,y;h,f) &= \varphi_0(x,y) + h\varphi_1(x,y) \\ &= f(x,y) + h(f_x(x,y) + f_y(x,y)f(x,y))\end{aligned}$$

perciò, se vogliamo evitare il calcolo delle derivate parziali di f, dobbiamo scegliere invece di $\Phi(x,y;h,f)$, una funzione $\widetilde{\Phi}(x,y;h,f)$ (in cui non compaiono derivate di f) tale che

$$\widetilde{\Phi}(x,y;h,f) = \Phi(x,y;h,f) + O(h^2) \tag{8}$$

- Nel caso $k = 2$ abbiamo

$$\begin{aligned}\Phi(x,y;h,f) &= \varphi_0(x,y) + h\varphi_1(x,y) \\ &= f(x,y) + h(f_x(x,y) + f_y(x,y)f(x,y))\end{aligned}$$

perciò, se vogliamo evitare il calcolo delle derivate parziali di f, dobbiamo scegliere invece di $\Phi(x,y;h,f)$, una funzione $\widetilde{\Phi}(x,y;h,f)$ (in cui non compaiono derivate di f) tale che

$$\widetilde{\Phi}(x,y;h,f) = \Phi(x,y;h,f) + O(h^2) \tag{8}$$

- Scegliamo $\widetilde{\Phi}(x,y;h,f) = a_1 f(x,y) + a_2 f(x + p_1 h, y + p_2 h f(x,y))$ in cui le costanti a_1, a_2, p_1 e p_2 devono essere determinate in modo da soddisfare (8)

$$\begin{aligned}\widetilde{\Phi}(x,y;h,f) &= a_1 f(x,y) + a_2 f(x + p_1 h, y + p_2 h f(x,y)) \\ &= a_1 f(x,y) + a_2 \left[f(x, y + p_2 h f(x,y)) + p_1 h f_x(x, y + p_2 h f(x,y)) \right] \\ &= a_1 f + a_2 \left[\left(f + p_2 h f_y f + O(h^2) \right) + p_1 h \left(f_x + p_2 h f_{xy} f + O(h^2) \right) \right] \\ &= (a_1 + a_2) f + h a_2 \left(p_1 f_x + p_2 f_y f \right) + O(h^2)\end{aligned}$$

- Perché sia

$$(a_1 + a_2)f + ha_2\left(p_1 f_x + p_2 f_y f\right) = f + h\left(f_x + f_y f\right)$$

deve essere

$$\begin{cases} a_2 \neq 0 \\ a_1 + a_2 = 1 \\ p_1 = p_2 = \frac{1}{2a_2} \end{cases}$$

- In base alla scelta della costante a_2 abbiamo diversi metodi il cui errore locale di troncamento $\tau_n(h;y)$ è di ordine 2

- Perché sia
$$(a_1+a_2)f+ha_2\left(p_1f_x+p_2f_yf\right)=f+h\left(f_x+f_yf\right)$$
deve essere
$$\begin{cases} a_2\neq 0 \\ a_1+a_2=1 \\ p_1=p_2=\frac{1}{2a_2} \end{cases}$$
- In base alla scelta della costante a_2 abbiamo diversi metodi il cui errore locale di troncamento $\tau_n(h;y)$ è di ordine 2

Metodo di Heun: $a_1=a_2=\frac{1}{2}$, $p_1=p_2=1$

$$u_{n+1}=u_n+\frac{h}{2}\left[f(x_n,u_n)+f(x_{n+1},u_n+hf(x_n,u_n))\right]$$

Metodo di Eulero modificato: $a_1=0$, $a_2=1$, $p_1=p_2=\frac{1}{2}$

$$u_{n+1}=u_n+hf(x_n+\tfrac{1}{2}h,u_n+\tfrac{1}{2}hf(x_n,u_n))$$

- I metodi visti finora, permettono di calcolare u_{n+1} utilizzando i valori u_n (e più in generale $u_n, u_{n-1}, \ldots, u_{n-r}$) calcolati precedentemente: per questo sono chiamati *espliciti*
- Un metodo è chiamato *implicito* se il valore incognito u_{n+1} dipende implicitamente da se stesso attraverso f

- I metodi visti finora, permettono di calcolare u_{n+1} utilizzando i valori u_n (e più in generale $u_n, u_{n-1}, \ldots, u_{n-r}$) calcolati precedentemente: per questo sono chiamati *espliciti*
- Un metodo è chiamato *implicito* se il valore incognito u_{n+1} dipende implicitamente da se stesso attraverso f

Metodo di Eulero implicito (o all'indietro)

Invece di approssimare y_{n+1} nell'intorno di y_n si approssima y_n nell'intorno di y_{n+1}, ottenendo

$$u_{n+1} = u_n + hf(x_{n+1}, u_{n+1})$$

perciò ad ogni passo si deve risolvere un'equazione non-lineare

- I metodi visti finora, permettono di calcolare u_{n+1} utilizzando i valori u_n (e più in generale $u_n, u_{n-1}, \ldots, u_{n-r}$) calcolati precedentemente: per questo sono chiamati *espliciti*
- Un metodo è chiamato *implicito* se il valore incognito u_{n+1} dipende implicitamente da se stesso attraverso f

Metodo di Eulero implicito (o all'indietro)

Invece di approssimare y_{n+1} nell'intorno di y_n si approssima y_n nell'intorno di y_{n+1}, ottenendo

$$u_{n+1} = u_n + hf(x_{n+1}, u_{n+1})$$

perciò ad ogni passo si deve risolvere un'equazione non-lineare

Metodo di Crank-Nicholson:

Si fa una media tra i metodi di Eulero implicito ed esplicito

$$u_{n+1} = u_n + \frac{h}{2}\left[f(x_n, u_n) + f(x_{n+1}, u_{n+1})\right]$$

Metodi di Runge-Kutta

- L'idea è di costruire delle formule per $\widetilde{\Phi}(x,y;h,f)$ in cui lo sviluppo in serie di Taylor coincida, per un fissato numero di termini con $\sum_{i=1}^{k} \frac{h^{i-1}}{i!}\varphi_i(x,y)$, senza usare esplicitamente le derivate di f
- Il prezzo per questo approccio è un aumento del numero di valutazioni di f ad ogni passo

Metodi di Runge-Kutta

- L'idea è di costruire delle formule per $\widetilde{\Phi}(x,y;h,f)$ in cui lo sviluppo in serie di Taylor coincida, per un fissato numero di termini con $\sum_{i=1}^{k} \frac{h^{i-1}}{i!}\varphi_i(x,y)$, senza usare esplicitamente le derivate di f
- Il prezzo per questo approccio è un aumento del numero di valutazioni di f ad ogni passo

Esempi

- Metodo di Heun: $u_{n+1} = u_n + \frac{h}{2}\left(f(x_n,u_n) + f(x_{n+1},u_n + hf(x_n,u_n))\right)$
- Metodo di Eulero modificato: $u_{n+1} = u_n + hf(x_n + \frac{h}{2}, u_n + \frac{h}{2}f(x_n,u_n))$

Metodi di Runge-Kutta espliciti a s stadi

Un metodo di Runge-Kutta esplicito a s stadi può essere scritto come

$$u_{n+1} = u_n + h \sum_{i=1}^{s} b_i K_i \qquad n \geq 0$$

dove i coefficienti K_i sono definiti da

$$\begin{cases} K_1 = f(x_n, u_n) \\ K_i = f\left(x_n + c_i h, u_n + h \sum_{j=1}^{i-1} a_{ij} K_j\right) \qquad i = 2, \ldots, s \end{cases}$$

Metodi di Runge-Kutta espliciti a s stadi

- I coefficienti a_{ij}, b_i e c_i caratterizzano un particolare metodo e vengono raccolti sotto forma di tabella (*matrice di Butcher*)

$$\begin{array}{c|ccccc} 0 & & & & & \\ c_2 & a_{21} & & & & \\ c_3 & a_{31} & a_{32} & & & \\ \vdots & \vdots & \vdots & \ddots & & \\ c_s & a_{s1} & a_{s2} & \dots & a_{s,s-1} & \\ \hline & b_1 & b_2 & \dots & b_{s-1} & b_s \end{array}$$

- Per ottenere i valori dei coefficienti, si utilizza lo stesso procedimento visto per il caso $s = 2$, considerando un numero maggiore di termini dello sviluppo in serie di Taylor
- Consistenza se e solo se $\sum_{i=1}^{s} b_i = 1$

Metodi di Runge-Kutta espliciti

Barriere di Butcher

- Un metodo di Runge-Kutta esplicito a s stadi non può avere ordine maggiore di s

Metodi di Runge-Kutta espliciti

Barriere di Butcher

- Un metodo di Runge-Kutta esplicito a s stadi non può avere ordine maggiore di s
- Per $p \geq 5$ non esiste nessun metodo di Runge-Kutta esplicito di ordine p con $s = p$ stadi
- Per $p \geq 7$ non esiste nessun metodo di Runge-Kutta esplicito di ordine p con $s = p + 1$ stadi
- Per $p \geq 8$ non esiste nessun metodo di Runge-Kutta esplicito di ordine p con $s = p + 2$ stadi

Metodi di Runge-Kutta espliciti

Barriere di Butcher

- Un metodo di Runge-Kutta esplicito a s stadi non può avere ordine maggiore di s
- Per $p \geq 5$ non esiste nessun metodo di Runge-Kutta esplicito di ordine p con $s = p$ stadi
- Per $p \geq 7$ non esiste nessun metodo di Runge-Kutta esplicito di ordine p con $s = p + 1$ stadi
- Per $p \geq 8$ non esiste nessun metodo di Runge-Kutta esplicito di ordine p con $s = p + 2$ stadi

ordine	1	2	3	4	5	6	7	8
$s_{\min}$	1	2	3	4	6	7	9	11

Metodi a 2 stadi (con $p = 2$)

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

Metodo di Heun

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

Eulero modificato

Metodi a 3 stadi (con $p = 3$)

$$\begin{array}{c|ccc} 0 & & & \\ \frac{1}{3} & \frac{1}{3} & & \\ \frac{2}{3} & 0 & \frac{1}{3} & \\ \hline & \frac{1}{4} & 0 & \frac{3}{4} \end{array}$$

Formula di Heun

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

Formula di Kutta

Metodo a 4 stadi (con $p = 4$)

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

Metodo di Runge-Kutta “classico”

Metodo a 4 stadi (con $p = 4$)

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

Metodo di Runge-Kutta “classico”

Equivale a:

$$u_{n+1} = u_n + \frac{h}{6}(K_1 + 2K_2 + 2K_3 + K_4)$$

dove i coefficienti K_i sono

$$\begin{cases} K_1 = f(x_n, u_n) \\ K_2 = f(x_n + \frac{h}{2}, u_n + \frac{h}{2}K_1) \\ K_3 = f(x_n + \frac{h}{2}, u_n + \frac{h}{2}K_2) \\ K_4 = f(x_n + h, u_n + hK_3) \end{cases}$$

Metodi adattivi

Stima *a posteriori ed adattività del passo*

- Si vuole scegliere il passo h di volta in volta, in modo che l'errore locale commesso sia inferiore ad una certa soglia ε

Metodi adattivi

Stima *a posteriori ed adattività del passo*

- Si vuole scegliere il passo h di volta in volta, in modo che l'errore locale commesso sia inferiore ad una certa soglia ε
- Si cerca una stima *a posteriori* dell'errore locale commesso al singolo passo

Metodi adattivi

Stima *a posteriori ed adattività del passo*

- Si vuole scegliere il passo h di volta in volta, in modo che l'errore locale commesso sia inferiore ad una certa soglia ε
- Si cerca una stima *a posteriori* dell'errore locale commesso al singolo passo
- Sia u^*_{n+1} la soluzione ottenuta applicando un passo di un metodo di Runge-Kutta di ordine p a partire dal dato iniziale y_n

$$y_{n+1} - u^*_{n+1} = \Psi(y_n)h^{p+1} + O(h^{p+2})$$

Metodi adattivi

Stima *a posteriori ed adattività del passo*

- Si vuole scegliere il passo h di volta in volta, in modo che l'errore locale commesso sia inferiore ad una certa soglia ε
- Si cerca una stima *a posteriori* dell'errore locale commesso al singolo passo
- Sia u^*_{n+1} la soluzione ottenuta applicando un passo di un metodo di Runge-Kutta di ordine p a partire dal dato iniziale y_n

$$y_{n+1} - u^*_{n+1} = \Psi(y_n)h^{p+1} + O(h^{p+2})$$

- Sia $\hat{u}^*_{n+1}$ la soluzione ottenuta con lo stesso metodo di Runge-Kutta ma con un passo $2h$ a partire dal dato iniziale y_{n-1}

$$\begin{aligned} y_{n+1} - \hat{u}^*_{n+1} &= \Psi(y_{n-1})(2h)^{p+1} + O(h^{p+2}) \\ &= \Psi(y_n)(2h)^{p+1} - \Psi'(\xi)2^{p+1}h^{p+2} + O(h^{p+2}) \qquad \text{con } \xi \in (y_{n-1}, y_n) \\ &= \Psi(y_n)(2h)^{p+1} + O(h^{p+2}) \end{aligned}$$

Metodi adattivi

Stima *a posteriori ed adattività del passo*

- Facendo la sottrazione $\left(y_{n+1} - \hat{u}^*_{n+1}\right) - \left(y_{n+1} - u^*_{n+1}\right)$ si ha

$$u^*_{n+1} - \hat{u}^*_{n+1} = \Psi(y_n) h^{p+1} \left(2^{p+1} - 1\right) + O(h^{p+2})$$

da cui possiamo ricavare il termine $\Psi(y_n) h^{p+1} + O(h^{p+2}) = \frac{u^*_{n+1} - \hat{u}^*_{n+1}}{2^{p+1}-1}$ che ci permette di calcolare la seguente stima dell'errore locale

$$y_{n+1} - u^*_{n+1} = \underbrace{\frac{u^*_{n+1} - \hat{u}^*_{n+1}}{2^{p+1} - 1}}_{\varepsilon} + O(h^{p+2})$$

Metodi adattivi

Stima *a posteriori ed adattività del passo*

- Facendo la sottrazione $\left(y_{n+1} - \hat{u}^*_{n+1}\right) - \left(y_{n+1} - u^*_{n+1}\right)$ si ha

$$u^*_{n+1} - \hat{u}^*_{n+1} = \Psi(y_n) h^{p+1} \left(2^{p+1} - 1\right) + O(h^{p+2})$$

da cui possiamo ricavare il termine $\Psi(y_n) h^{p+1} + O(h^{p+2}) = \frac{u^*_{n+1} - \hat{u}^*_{n+1}}{2^{p+1}-1}$ che ci permette di calcolare la seguente stima dell'errore locale

$$y_{n+1} - u^*_{n+1} = \underbrace{\frac{u^*_{n+1} - \hat{u}^*_{n+1}}{2^{p+1} - 1}}_{\varepsilon} + O(h^{p+2})$$

- La stima dell'errore produce un considerevole aumento del costo computazionale: per calcolare $\hat{u}^*_{n+1}$ servono $s - 1$ valutazioni addizionali di f

Metodi adattivi immersi

Stima *a posteriori ed adattività del passo*

- Vogliamo fornire una stima *a posteriori* dell'errore locale senza fare valutazioni addizionali di f

Metodi adattivi immersi

Stima *a posteriori ed adattività del passo*

- Vogliamo fornire una stima *a posteriori* dell'errore locale senza fare valutazioni addizionali di f
- Una possibilità è quella di utilizzare due metodi Runge-Kutta di ordine p e $p+1$ che hanno lo stesso insieme di valori K_i

Metodi adattivi immersi

Stima *a posteriori ed adattività del passo*

- Vogliamo fornire una stima *a posteriori* dell'errore locale senza fare valutazioni addizionali di f
- Una possibilità è quella di utilizzare due metodi Runge-Kutta di ordine p e $p+1$ che hanno lo stesso insieme di valori K_i
- Metodo di ordine p: $u_{n+1} = u_n + h\sum_{i=1}^{s} b_i K_i$
- Metodo di ordine $p+1$: $\hat{u}_{n+1} = \hat{u}_n + h\sum_{i=1}^{s} \hat{b}_i K_i$

Metodi adattivi immersi

Stima *a posteriori ed adattività del passo*

- Vogliamo fornire una stima *a posteriori* dell'errore locale senza fare valutazioni addizionali di f
- Una possibilità è quella di utilizzare due metodi Runge-Kutta di ordine p e $p+1$ che hanno lo stesso insieme di valori K_i
- Metodo di ordine p: $u_{n+1} = u_n + h\sum_{i=1}^{s} b_i K_i$
- Metodo di ordine $p+1$: $\hat{u}_{n+1} = \hat{u}_n + h\sum_{i=1}^{s} \hat{b}_i K_i$
- Stima dell'errore locale:

$$y_{n+1} - u^*_{n+1} = \underbrace{(y_{n+1} - \hat{u}^*_{n+1})}_{O(h^{p+1})} + (\hat{u}^*_{n+1} - u^*_{n+1})$$

$$= O(h^{p+1}) + h\sum_{i=1}^{s} (\hat{b}_i - b_i) K_i$$

Metodi adattivi immersi

- La stima dell'errore locale è pari a $y_{n+1} - u^*_{n+1} = h\sum_{i=1}^{s}(\hat{b}_i - b_i)K_i + O(h^{p+1})$
- Poiché vogliamo che l'errore locale sia minore di una certa soglia ε, dobbiamo usare un passo h tale che

$$h \lesssim \frac{\varepsilon}{\sum_{i=1}^{s}(\hat{b}_i - b_i)K_i}$$

Metodi adattivi immersi

- La stima dell'errore locale è pari a $y_{n+1} - u^*_{n+1} = h\sum_{i=1}^{s}(\hat{b}_i - b_i)K_i + O(h^{p+1})$
- Poiché vogliamo che l'errore locale sia minore di una certa soglia ε, dobbiamo usare un passo h tale che

$$h \lesssim \frac{\varepsilon}{\sum_{i=1}^{s}(\hat{b}_i - b_i)K_i}$$

- Questi metodi vengono rappresentati tramite la matrice di Butcher modificata

$$\begin{array}{c|ccccc} 0 & & & & & \\ c_2 & a_{21} & & & & \\ c_3 & a_{31} & a_{32} & & & \\ \vdots & \vdots & \vdots & \ddots & & \\ c_s & a_{s1} & a_{s2} & \dots & a_{s,s-1} & \\ \hline & b_1 & b_2 & \dots & b_{s-1} & b_s \\ & \hat{b}_1 & \hat{b}_2 & \dots & \hat{b}_{s-1} & \hat{b}_s \end{array}$$

Esempio di metodo immerso: RKF4(5)

- Il metodo di Runge-Kutta Fehlberg di ordine 4 (RKF4(5)) è uno degli schemi immersi più noti
- La matrice di Butcher modificata per questo metodo è data da

$$
\begin{array}{c|cccccc}
0 & & & & & & \\
\frac{1}{4} & \frac{1}{4} & & & & & \\
\frac{3}{8} & \frac{3}{32} & \frac{9}{32} & & & & \\
\frac{12}{13} & \frac{1932}{2197} & -\frac{7200}{2197} & \frac{7296}{2197} & & & \\
1 & \frac{439}{216} & -8 & \frac{3680}{513} & -\frac{845}{4104} & & \\
\frac{1}{2} & -\frac{8}{27} & 2 & -\frac{3544}{2565} & \frac{1869}{4104} & \frac{11}{40} & \\
\hline
b^T & \frac{25}{216} & 0 & -\frac{1408}{2565} & \frac{2197}{4104} & -\frac{1}{5} & 0 \\
\hat{b}^T & \frac{16}{135} & 0 & \frac{6656}{12825} & \frac{28561}{56430} & -\frac{9}{50} & \frac{2}{55}
\end{array}
$$

Metodi di Runge-Kutta impliciti

Un metodo di Runge-Kutta implicito a s stadi può essere scritto come

$$\begin{cases} u_{n+1} = u_n + h\sum_{i=1}^{s} b_i K_i \quad n \geq 0 \\ K_i = f\left(x_n + c_i h, u_n + h\sum_{j=1}^{s} a_{ij} K_j\right) \qquad i = 1, \ldots, s \end{cases} \tag{9}$$

Metodi di Runge-Kutta impliciti

Un metodo di Runge-Kutta implicito a s stadi può essere scritto come

$$\begin{cases} u_{n+1} = u_n + h\sum_{i=1}^{s} b_i K_i \quad n \geq 0 \\ K_i = f\left(x_n + c_i h, u_n + h\sum_{j=1}^{s} a_{ij} K_j\right) \qquad i = 1, \ldots, s \end{cases} \tag{9}$$

- Sono potenzialmente più accurate delle corrispondenti formule esplicite con lo stesso numero di stadi. Però il calcolo dei K_i richiede la risoluzione di un sistema non lineare di dimensione s
- Un metodo di Runge-Kutta implicito a s stadi non può avere ordine maggiore di $2s$

Metodi di Runge-Kutta impliciti

Un metodo di Runge-Kutta implicito a s stadi può essere scritto come

$$\begin{cases} u_{n+1} = u_n + h\sum_{i=1}^{s} b_i K_i \quad n \geq 0 \\ K_i = f\left(x_n + c_i h, u_n + h\sum_{j=1}^{s} a_{ij} K_j\right) \qquad i = 1,\ldots,s \end{cases} \tag{9}$$

- Sono potenzialmente più accurate delle corrispondenti formule esplicite con lo stesso numero di stadi. Però il calcolo dei K_i richiede la risoluzione di un sistema non lineare di dimensione s
- Un metodo di Runge-Kutta implicito a s stadi non può avere ordine maggiore di $2s$
- Problema di esistenza della soluzione di (9): se $f(x,y)$ è continua e lipschitziana rispetto a y con constante L e se $h < (L \max_i \sum_j |a_{ij}|)^{-1}$ allora esiste un'unica soluzione di (9)

- I metodi impliciti possono essere costruiti utilizzando la forma integrale del problema di Cauchy, ovvero

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\, dt$$

  e ricorrendo ad opportune interpolazioni e formule di quadratura
- I coefficienti b_i e c_i sono i pesi e i nodi della formula di quadratura usata, mentre a_{ij} dipende sia dalla formula di quadratura che dall'interpolazione usata

Esempi di metodi a 2 stadi

$$\begin{array}{c|cc} 0 & 0 & 0 \\ \frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ \hline & \frac{1}{4} & \frac{3}{4} \end{array}$$

Hammer-Hollingsworth

$$\begin{array}{c|cc} \frac{3-\sqrt{3}}{3} & \frac{1}{4} & \frac{3-2\sqrt{3}}{12} \\ \frac{3+\sqrt{3}}{6} & \frac{3+2\sqrt{3}}{12} & \frac{1}{4} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}$$

Gauss-Legendre (p=4)

Metodi di Runge-Kutta semi-impliciti

- La difficoltà computazionale di uno schema implicito, può essere ridotta considerando i metodi *semi-impliciti*, ovvero i metodi in cui $a_{ij} = 0$ per $j > i$, ovvero

$$K_i = f\left(x_n + c_i h, u_n + h \sum_{j=1}^{i-1} a_{ij} K_j + h a_{ii} K_i\right)$$

- Uno schema semi-implicito richiede la risoluzione di s equazioni non lineari indipendenti

Metodi a più passi (*multistep* o MS)

- Nei metodi di Runge-Kutta, l'approssimazione u_n e le s valutazioni della funzione f nell'intervallo $[x_n, x_{n+1}]$ (utilizzate per la determinazione di u_{n+1}) non sono più utilizzate per i passi successivi
- Nei metodi a più passi, per calcolare u_{n+1} si utilizzano alcune approssimazioni precedentemente calcolate $u_n, u_{n-1}, \ldots, u_{n-r}$

Metodi a più passi (*multistep* o MS)

- Nei metodi di Runge-Kutta, l'approssimazione u_n e le s valutazioni della funzione f nell'intervallo $[x_n, x_{n+1}]$ (utilizzate per la determinazione di u_{n+1}) non sono più utilizzate per i passi successivi
- Nei metodi a più passi, per calcolare u_{n+1} si utilizzano alcune approssimazioni precedentemente calcolate $u_n, u_{n-1}, \ldots, u_{n-r}$

- Un metodo MS lineare a r passi è definito dalla seguente relazione

$$\sum_{q=0}^{r} \alpha_q u_{n+q} = h \sum_{q=0}^{r} \beta_q f(x_{n+q}, u_{n+q}) \tag{10}$$

  in cui $\alpha_r = 1$ e $|\alpha_0| + |\beta_0| \neq 0$.
- Per inizializzare un metodo a $r > 1$ passi sono necessari $r - 1$ valori aggiuntivi al valore iniziale y_0
- I dati relativi ai r passi precedenti devono essere memorizzati (potrebbe essere un problema nel caso vettoriale)

- Un metodo MS si può anche scrivere come

$$u_{n+1} = \sum_{j=0}^{r-1} a_j u_{n-j} + h \sum_{j=-1}^{r-1} b_j f(x_{n-j}, u_{n-j})$$

- L'errore di troncamento locale è dato da $y_{n+1} - u^*_{n+1}$

$$h\tau_{n+1} = y_{n+1} - \left(\sum_{j=0}^{r-1} a_j y_{n-j} + h \sum_{j=-1}^{r-1} b_j y'_{n-j} \right)$$

- Se $y \in C^{p+1}(I)$ per qualche $p \geq 1$, il metodo è di ordine p se e solo se sono soddisfatte le seguenti condizioni (deve essere $\tau_n = O(h^p)$)

$$\sum_{j=0}^{r-1} (-j)^i a_j + i \sum_{j=-1}^{r-1} (-j)^{i-1} b_j = 1 \qquad i = 0, 1, \ldots, q$$

- In questo caso abbiamo $h\tau_{n+1} = C_{p+1} y^{p+1}_{n-r+1} + O(h^{p+2})$ dove il primo termine è chiamato errore di troncamento locale principale

Metodi a più passi: i metodi di Adams

- Questi metodi vengono direttamente derivati dalla forma integrale

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\, dt$$

- L'integrale è approssimato utilizzando invece di f un opportuno polinomio di interpolazione
- Si possono definire sia metodi espliciti (detti di Adams-Bashforth) che impliciti (detti di Adams-Moulton)

Metodi di Adams-Bashforth (AB)

- Consideriamo lo schema seguente

$$u_{n+1} = u_n + \int_{x_n}^{x_{n+1}} P_n(x)\,dx$$

in cui P_n è un polinomio di interpolazione di grado al più $r-1$ definito da:

$$P_n(x_{n-j}) = f(x_{n-j}, u_{n-j}) = f_{n-j} \qquad \forall j = 0, \ldots, r-1$$

Metodi di Adams-Bashforth (AB)

- Consideriamo lo schema seguente

$$u_{n+1} = u_n + \int_{x_n}^{x_{n+1}} P_n(x)\,dx$$

in cui P_n è un polinomio di interpolazione di grado al più $r-1$ definito da:

$$P_n(x_{n-j}) = f(x_{n-j}, u_{n-j}) = f_{n-j} \qquad \forall j = 0, \ldots, r-1$$

- Utilizzando la rappresentazione di Lagrange si ha

$$P_n(x) = \sum_{j=0}^{r-1} f_{n-j} L_j(x)$$

con $L_j(x) = \prod_{\substack{i=1 \\ i \neq j}}^{r-1} \frac{(x - x_{n-i})}{(x_{n-j} - x_{n-i})}$ per $j = 0, \ldots, r-1$ (notare che $L_j(x_{n-q}) = \delta_{jq}$)

Metodi di Adams-Bashforth (AB)

$$u_{n+1} = u_n + \int_{x_n}^{x_{n+1}} P_n(x)\,dx = u_n + \int_{x_n}^{x_{n+1}} \left(\sum_{j=0}^{r-1} f_{n-j} L_j(x) \right) dx$$

$$= u_n + \sum_{j=0}^{r-1} f_{n-j} \int_{x_n}^{x_{n+1}} L_j(x)\,dx$$

$$= u_n + h \sum_{j=0}^{r-1} b_j f_{n-j}$$

in cui $b_j = \dfrac{1}{h} \displaystyle\int_{x_n}^{x_{n+1}} L_j(x)\,dx$

Metodi di Adams-Bashforth (AB)

$$u_{n+1} = u_n + \int_{x_n}^{x_{n+1}} P_n(x)\,dx = u_n + \int_{x_n}^{x_{n+1}} \left(\sum_{j=0}^{r-1} f_{n-j} L_j(x) \right) dx$$

$$= u_n + \sum_{j=0}^{r-1} f_{n-j} \int_{x_n}^{x_{n+1}} L_j(x)\,dx$$

$$= u_n + h \sum_{j=0}^{r-1} b_j f_{n-j}$$

in cui $b_j = \dfrac{1}{h} \displaystyle\int_{x_n}^{x_{n+1}} L_j(x)\,dx$

- I coefficienti b_j (per h costante) dipendono soltanto da r e caratterizzano un particolare metodo di Adams-Bashforth
- Per $r = 1$ si ritrova il metodo di Eulero in avanti

Metodi di Adams-Bashforth (AB)

$$u_{n+1} = u_n + \int_{x_n}^{x_{n+1}} P_n(x)\,dx = u_n + \int_{x_n}^{x_{n+1}} \left(\sum_{j=0}^{r-1} f_{n-j} L_j(x) \right) dx$$

$$= u_n + \sum_{j=0}^{r-1} f_{n-j} \int_{x_n}^{x_{n+1}} L_j(x)\,dx$$

$$= u_n + h \sum_{j=0}^{r-1} b_j f_{n-j}$$

in cui $b_j = \dfrac{1}{h} \displaystyle\int_{x_n}^{x_{n+1}} L_j(x)\,dx$

- I coefficienti b_j (per h costante) dipendono soltanto da r e caratterizzano un particolare metodo di Adams-Bashforth
- Per $r = 1$ si ritrova il metodo di Eulero in avanti
- Gli schemi di Adams-Bashforth a r passi sono di ordine r

Metodi di Adams-Bashforth (AB)

Metodo AB2 ($r = 2$)

$$u_{n+1} = u_n + \frac{h}{2}(3f_n - f_{n-1})$$

Metodo AB3 ($r = 3$)

$$u_{n+1} = u_n + \frac{h}{12}(23f_n - 16f_{n-1} + 5f_{n-2})$$

Metodo AB4 ($r = 4$)

$$u_{n+1} = u_n + \frac{h}{24}(55f_n - 59f_{n-1} + 37f_{n-2} - 9f_{n-3})$$

Metodi di Adams-Moulton (AM)

- Come per il caso esplicito si considera uno schema della forma seguente

$$u_{n+1} = u_n + \int_{x_n}^{x_{n+1}} Q_n(x)\,dx$$

in cui Q_n è un polinomio di interpolazione di grado al più r definito da:

$$\begin{cases} Q_n(x_{n-j}) = f_{n-j} & \forall j = 0, \ldots, r-1 \\ Q_n(x_{n+1}) = f_{n+1} & \text{(valore incognito)} \end{cases}$$

Metodi di Adams-Moulton (AM)

- Come per il caso esplicito si considera uno schema della forma seguente

$$u_{n+1} = u_n + \int_{x_n}^{x_{n+1}} Q_n(x)\,dx$$

in cui Q_n è un polinomio di interpolazione di grado al più r definito da:

$$\begin{cases} Q_n(x_{n-j}) = f_{n-j} & \forall j = 0, \ldots, r-1 \\ Q_n(x_{n+1}) = f_{n+1} & \text{(valore incognito)} \end{cases}$$

- Integrando, si ottiene lo schema implicito seguente

$$u_{n+1} = u_n + h \sum_{j=-1}^{r-1} \tilde{b}_j f_{n-j}$$

Metodi di Adams-Moulton (AM)

- Come per il caso esplicito si considera uno schema della forma seguente

$$u_{n+1} = u_n + \int_{x_n}^{x_{n+1}} Q_n(x)\,dx$$

in cui Q_n è un polinomio di interpolazione di grado al più r definito da:

$$\begin{cases} Q_n(x_{n-j}) = f_{n-j} & \forall j = 0, \ldots, r-1 \\ Q_n(x_{n+1}) = f_{n+1} & \text{(valore incognito)} \end{cases}$$

- Integrando, si ottiene lo schema implicito seguente

$$u_{n+1} = u_n + h \sum_{j=-1}^{r-1} \tilde{b}_j f_{n-j}$$

- Lo schema di Adams-Moulton a r passi è di ordine $r+1$

Metodi di Adams-Moulton (AM)

Metodo AM0 ($r = 0$)

$$u_{n+1} = u_n + hf_{n+1} \qquad \text{(Eulero implicito)}$$

Metodo AM1 ($r = 1$)

$$u_{n+1} = u_n + \frac{h}{2}(f_{n+1} + f_n) \qquad \text{(Crank-Nicholson)}$$

Metodo AM2 ($r = 2$)

$$u_{n+1} = u_n + \frac{h}{12}(5f_{n+1} + 8f_n - f_{n-1})$$

Metodo AM3 ($r = 3$)

$$u_{n+1} = u_n + \frac{h}{24}(9f_{n+1} + 19f_n - 5f_{n-1} + f_{n-2})$$

Metodi a più passi: generalizzazione

- I metodi di Adams si possono generalizzare integrando tra x_{n-q} e x_{n+1} ($q \geq 1$):

$$u_{n+1} = u_{n-q} + h \sum_{j=-1}^{r-1} \tilde{b}_j f_{n-j}$$

- Per $q = 1$ si ottengono i metodi di Nyström nel caso esplicito e i metodi di Milne-Simpson nel caso implicito

Metodi a più passi: generalizzazione

- I metodi di Adams si possono generalizzare integrando tra x_{n-q} e x_{n+1} ($q \geq 1$):

$$u_{n+1} = u_{n-q} + h \sum_{j=-1}^{r-1} \tilde{b}_j f_{n-j}$$

- Per $q = 1$ si ottengono i metodi di Nyström nel caso esplicito e i metodi di Milne-Simpson nel caso implicito
- Il metodo di Nyström per $r = 1$ è il metodo che utilizza la formula di integrazione del punto medio

$$u_{n+1} = u_{n-1} + 2hf_n$$

- Il metodo di Milne-Simpson per $r = 2$ è il metodo che utilizza la formula di integrazione di Cavalieri-Simpson

$$u_{n+1} = u_{n-1} + \frac{h}{3}(f_{n+1} + 4f_n + f_{n-1})$$

- La risoluzione di un problema di Cauchy non-lineare con uno schema implicito richiede ad ogni iterazione la risoluzione di un equazione non-lineare

$$u_{n+1} = \Psi_n(u_{n+1}) \qquad n \geq 0 \tag{11}$$

- Per il metodo di Adams-Moulton a r passi si ha

$$\Psi_n(\xi) = h\tilde{b}_{-1} f(x_{n+1}, \xi) + u_n + h \sum_{j=0}^{r-1} \tilde{b}_j f_{n-j}$$

- La risoluzione di un problema di Cauchy non-lineare con uno schema implicito richiede ad ogni iterazione la risoluzione di un equazione non-lineare

$$u_{n+1} = \Psi_n(u_{n+1}) \qquad n \geq 0 \tag{11}$$

- Per il metodo di Adams-Moulton a r passi si ha

$$\Psi_n(\xi) = h\tilde{b}_{-1} f(x_{n+1}, \xi) + u_n + h \sum_{j=0}^{r-1} \tilde{b}_j f_{n-j}$$

- Per risolvere (11) si può usare un metodo di punto fisso

$$u_{n+1}^{(k+1)} = \Psi_n(u_{n+1}^{(k)}) \qquad k = 0, 1, \ldots \tag{12}$$

- Per i metodi di Adams-Moulton, affinché il metodo (12) converga, bisogna che il passo di discretizzazione soddisfi $h \leq \frac{1}{|\tilde{b}_{-1}|L}$ dove L è la costante di Lipschitz di $f(x,y)$ rispetto a y. A parte casi particolari (problemi "stiff") questa restrizione è poco penalizzante

Approccio predittore-correttore

- Per limitare il costo computazionale è importante definire una buona stima iniziale $u_{n+1}^{(0)}$

Approccio predittore-correttore

- Per limitare il costo computazionale è importante definire una buona stima iniziale $u_{n+1}^{(0)}$
- Una scelta ragionevole è quella di prendere come $u_{n+1}^{(0)}$ il valore fornito da un metodo esplicito
- In questo modo lo schema implicito "corregge" il valore $u_{n+1}^{(0)}$ inizialmente "predetto" dallo schema esplicito

Approccio predittore-correttore

- Per limitare il costo computazionale è importante definire una buona stima iniziale $u_{n+1}^{(0)}$
- Una scelta ragionevole è quella di prendere come $u_{n+1}^{(0)}$ il valore fornito da un metodo esplicito
- In questo modo lo schema implicito "corregge" il valore $u_{n+1}^{(0)}$ inizialmente "predetto" dallo schema esplicito
- Approccio predittore-correttore (algoritmi $P(EC)^m E$ e $P(EC)^m$):
  1. una iterazione con lo schema esplicito (inizializzazione del metodo di punto fisso)
  2. m iterazioni con lo schema implicito

Approccio predittore-correttore

- Per limitare il costo computazionale è importante definire una buona stima iniziale $u_{n+1}^{(0)}$
- Una scelta ragionevole è quella di prendere come $u_{n+1}^{(0)}$ il valore fornito da un metodo esplicito
- In questo modo lo schema implicito "corregge" il valore $u_{n+1}^{(0)}$ inizialmente "predetto" dallo schema esplicito
- Approccio predittore-correttore (algoritmi $P(EC)^m E$ e $P(EC)^m$):
  1. una iterazione con lo schema esplicito (inizializzazione del metodo di punto fisso)
  2. m iterazioni con lo schema implicito
- Il metodo di Heun può essere visto come un metodo predittore-correttore nel quale il predittore è il metodo di Eulero e il correttore è il metodo di Crank-Nicholson

$$\begin{cases} u_{n+1}^{(0)} = u_n + hf(x_n, u_n) \\ u_{n+1} = u_n + \frac{h}{2}\left[f(x_n, u_n) + f(x_{n+1}, u_{n+1}^{(0)})\right] \end{cases}$$

Algoritmo $P(EC)^m E$

1. (Prediction): metodo esplicito a r passi

$$u_{n+1}^{(0)} = \sum_{j=0}^{r-1} a_j u_{n-j}^{(m)} + h \sum_{j=0}^{r-1} b_j f_{n-j}^{(m)}$$

2. Per $k = 0, \ldots, m-1$
   1. (Evaluation): valutazione di f nel nuovo punto

$$f_{n+1}^{(k)} = f(x_{n+1}, u_{n+1}^{(k)})$$

   2. (Correction): metodo implicito a $\tilde{r}$ passi

$$u_{n+1}^{(k+1)} = \sum_{j=0}^{\tilde{r}-1} \tilde{a}_j u_{n-j}^{(m)} + h \sum_{j=0}^{\tilde{r}-1} \tilde{b}_j f_{n-j}^{(m)} + h\tilde{b}_{-1} f_{n+1}^{(k)}$$

3. (Evaluation): nuova valutazione di f

$$f_{n+1}^{(m)} = f(x_{n+1}, u_{n+1}^{(m)})$$

Algoritmo $P(EC)^m$

1. (Prediction): metodo esplicito a r passi

$$u_{n+1}^{(0)} = \sum_{j=0}^{r-1} a_j u_{n-j}^{(m)} + h \sum_{j=0}^{r-1} b_j f_{n-j}^{(m)}$$

2. Per $k = 0, \ldots, m-1$
   1. (Evaluation): valutazione di f nel nuovo punto

$$f_{n+1}^{(k)} = f(x_{n+1}, u_{n+1}^{(k)})$$

   2. (Correction): metodo implicito a $\tilde{r}$ passi

$$u_{n+1}^{(k+1)} = \sum_{j=0}^{\tilde{r}-1} \tilde{a}_j u_{n-j}^{(m)} + h \sum_{j=0}^{\tilde{r}-1} \tilde{b}_j f_{n-j}^{(m)} + h\tilde{b}_{-1} f_{n+1}^{(k)}$$

Consistenza dei metodi predittore-correttore

Come definire (al più basso costo possibile) un predittore in modo da mantenere l'ordine di precisione dello schema originale?

Consistenza dei metodi predittore-correttore

Come definire (al più basso costo possibile) un predittore in modo da mantenere l'ordine di precisione dello schema originale?

- Supponiamo di avere un predittore di ordine q e un correttore di ordine $\tilde{q}$
  - Se $\tilde{q} \leq q$, o se $\tilde{q} > q$ con $m > \tilde{q} - q$ allora il metodo ha lo stesso ordine del correttore e anche lo stesso errore di troncamento locale principale
  - Se $\tilde{q} > q$ con $m = \tilde{q} - q$ allora il metodo ha lo stesso ordine del correttore ma l'errore di troncamento locale principale è diverso

Consistenza dei metodi predittore-correttore

Come definire (al più basso costo possibile) un predittore in modo da mantenere l'ordine di precisione dello schema originale?

- Supponiamo di avere un predittore di ordine q e un correttore di ordine $\tilde{q}$
  - Se $\tilde{q} \leq q$, o se $\tilde{q} > q$ con $m > \tilde{q} - q$ allora il metodo ha lo stesso ordine del correttore e anche lo stesso errore di troncamento locale principale
  - Se $\tilde{q} > q$ con $m = \tilde{q} - q$ allora il metodo ha lo stesso ordine del correttore ma l'errore di troncamento locale principale è diverso
- In particolare, se il predittore ha ordine $q = \tilde{q} - 1$, allora il metodo predittore-correttore ha ordine $\tilde{q}$
- In generale si prende $q = \tilde{q}$ o $q = \tilde{q} - 1$

Consistenza dei metodi predittore-correttore

Come definire (al più basso costo possibile) un predittore in modo da mantenere l'ordine di precisione dello schema originale?

- Supponiamo di avere un predittore di ordine q e un correttore di ordine $\tilde{q}$
  - Se $\tilde{q} \leq q$, o se $\tilde{q} > q$ con $m > \tilde{q} - q$ allora il metodo ha lo stesso ordine del correttore e anche lo stesso errore di troncamento locale principale
  - Se $\tilde{q} > q$ con $m = \tilde{q} - q$ allora il metodo ha lo stesso ordine del correttore ma l'errore di troncamento locale principale è diverso
- In particolare, se il predittore ha ordine $q = \tilde{q} - 1$, allora il metodo predittore-correttore ha ordine $\tilde{q}$
- In generale si prende $q = \tilde{q}$ o $q = \tilde{q} - 1$
- Metodo ABM di ordine p: metodo di Adams-Bashforth (predittore) + metodo di Adams-Moulton (correttore), entrambi di ordine p

Metodi alle differenze all'indietro (BDF)

- I metodi BDF (*Backward Differentiation Formulae*) sono metodi impliciti a più passi che sono particolarmente adatti per la risoluzione dei problemi "stiff"
- Invece di approssimare l'integrale di f, si approssima direttamente il valore di y'_{n+1} tramite la derivata prima del polinomio di interpolazione di y nei nodi $x_{n+1}, x_n, \ldots, x_{n-r+1}$
- Gli schemi che si ottengono sono della forma

$$u_{n+1} = \sum_{j=0}^{r} a_j u_{n-j} + h b_{-1} f_{n+1}$$

- I metodi BDF sono di ordine $p = r+1$ (numero di passi) e sono stabili solo per $p \leq 6$

Coefficienti per i metodi BDF stabili

r	a_0	a_1	a_2	a_3	a_4	a_5	b_{-1}
0	1						1
1	$\frac{4}{3}$	$-\frac{1}{3}$					$\frac{2}{3}$
2	$\frac{18}{11}$	$-\frac{9}{11}$	$\frac{2}{11}$				$\frac{6}{11}$
3	$\frac{48}{25}$	$-\frac{36}{25}$	$\frac{16}{25}$	$\frac{3}{25}$			$\frac{12}{25}$
4	$\frac{300}{137}$	$-\frac{300}{137}$	$\frac{200}{137}$	$\frac{75}{137}$	$\frac{12}{137}$		$\frac{60}{137}$
5	$\frac{360}{147}$	$-\frac{450}{147}$	$\frac{400}{147}$	$-\frac{225}{147}$	$\frac{72}{147}$	$-\frac{10}{147}$	$\frac{60}{147}$

- Consideriamo il metodo a k passi seguente:

$$\begin{cases} u_j = g_j(h) & j = 0, 1, \ldots, k-1 \\ \sum_{j=0}^{k} \alpha_j u_{n+j} = h\Psi(x_n; u_{n+k}, \ldots, u_n; f) \end{cases} \tag{13}$$

- In particolare, i metodi a più passi lineari (Adams, BDF), i metodi $P(EC)^m E$ e i metodi Runge-Kutta possono essere scritti nella forma (13)

- Consideriamo il metodo a k passi seguente:

$$\begin{cases} u_j = g_j(h) & j = 0, 1, \ldots, k-1 \\ \sum_{j=0}^{k} \alpha_j u_{n+j} = h\Psi(x_n; u_{n+k}, \ldots, u_n; f) \end{cases} \tag{13}$$

- In particolare, i metodi a più passi lineari (Adams, BDF), i metodi $P(EC)^m E$ e i metodi Runge-Kutta possono essere scritti nella forma (13)

Convergenza

- Un metodo della classe (13) è detto *convergente* se, applicato a un *qualunque* problema di Cauchy (per il quale abbiamo esistenza e unicita della soluzione), soddisfa la proprietà

$$\lim_{h \to 0} \max_{0 \le n \le N_h} ||y_n - u_n|| = 0$$

  in cui $I = [a, b]$ e $N_h h = b - a$

- In particolare, dobbiamo avere un insieme compatibile di dati iniziali, cioé:

$$\lim_{h \to 0} u_j = y_0 \qquad j = 0, \ldots, k-1$$

Zero-stabilità

- Fissato l'intervallo $I = [a, b]$ si studia il comportamento del metodo numerico per il caso limite $h \to 0$

Zero-stabilità

- Fissato l'intervallo $I = [a, b]$ si studia il comportamento del metodo numerico per il caso limite $h \to 0$
- La riduzione di h per un dato intervallo può produrre un aumento dei contributi all'errore totale dovuto all'accumulo degli errori locali di arrotondamento
- $\Rightarrow$ Vogliamo un metodo poco sensibile alle piccole perturbazioni per potere sotto controllo gli errori di arrotondamento

Zero-stabilità

- Un metodo della classe (13) è detto *zero-stabile* se esistono due costanti $h_0 > 0$ e $C > 0$ tali che:

$$\forall h \in (0, h_0] \Rightarrow |z_n^{(h)} - u_n^{(h)}| \leq C\varepsilon \qquad n = 0, \ldots, N_h$$

in cui $u_n^{(h)}$ è la soluzione del problema (13), mentre $z_n^{(h)}$ è la soluzione del problema perturbato seguente

$$\begin{cases} z_j^{(h)} = g_j(h) + \delta_j & j = 0, 1, \ldots, k-1 \\ \sum_{j=0}^{k} \alpha_j z_{n+j}^{(h)} = h\Psi(x_n; z_{n+k}^{(h)}, \ldots, z_n^{(h)}; h; f) + h\delta_{n+k} \end{cases}$$

dove $|\delta_j| \leq \varepsilon$ per $j = 0, \ldots, N_h$

Zero-stabilità

- Un metodo della classe (13) è detto *zero-stabile* se esistono due costanti $h_0 > 0$ e $C > 0$ tali che:

$$\forall h \in (0, h_0] \Rightarrow |z_n^{(h)} - u_n^{(h)}| \leq C\varepsilon \qquad n = 0, \ldots, N_h$$

in cui $u_n^{(h)}$ è la soluzione del problema (13), mentre $z_n^{(h)}$ è la soluzione del problema perturbato seguente

$$\begin{cases} z_j^{(h)} = g_j(h) + \delta_j & j = 0, 1, \ldots, k-1 \\ \sum_{j=0}^{k} \alpha_j z_{n+j}^{(h)} = h\Psi(x_n; z_{n+k}^{(h)}, \ldots, z_n^{(h)}; h; f) + h\delta_{n+k} \end{cases}$$

dove $|\delta_j| \leq \varepsilon$ per $j = 0, \ldots, N_h$

Teorema di equivalenza

Un metodo (13) è convergente se e solo se è consistente, zero-stabile e se l'errore sui dati iniziali tende a zero per $h \to 0$

Come si verifica la zero-stabilità?

- Si deve verificare che $\lim_{h\to 0} |w_n^{(h)}| = \lim_{h\to 0} |z_n^{(h)} - u_n^{(h)}| = 0$ dove $w_n^{(h)}$ soddisfa

$$\begin{cases} w_j^{(h)} = \delta_j \qquad j = 0, 1, \ldots, k-1 \\ \Delta\Psi_{n+k} = \Psi(x_n; z_{n+k}^{(h)}, \ldots, z_n^{(h)}; h; f) - \Psi(x_n; u_{n+k}^{(h)}, \ldots, u_n^{(h)}; h; f) \\ \sum_{j=0}^{k} \alpha_j w_{n+j}^{(h)} = h\Delta\Psi_{n+k} + h\delta_{n+k} \end{cases}$$

con $|\delta_j| \leq \varepsilon$ per $j = 0, \ldots, N_h$

Come si verifica la zero-stabilità?

- Si deve verificare che $\lim_{h\to 0} |w_n^{(h)}| = \lim_{h\to 0} |z_n^{(h)} - u_n^{(h)}| = 0$ dove $w_n^{(h)}$ soddisfa

$$\begin{cases} w_j^{(h)} = \delta_j \qquad j = 0, 1, \ldots, k-1 \\ \Delta\Psi_{n+k} = \Psi(x_n; z_{n+k}^{(h)}, \ldots, z_n^{(h)}; h; f) - \Psi(x_n; u_{n+k}^{(h)}, \ldots, u_n^{(h)}; h; f) \\ \sum_{j=0}^{k} \alpha_j w_{n+j}^{(h)} = h\Delta\Psi_{n+k} + h\delta_{n+k} \end{cases}$$

  con $|\delta_j| \leq \varepsilon$ per $j = 0, \ldots, N_h$

- È utile esprimere $w_n^{(h)}$ esplicitamente in funzione di h, n e dei vari δ_j (e non per ricorrenza)
- Una volta ottenuta una formula esplicita per $w_n^{(h)}$, si verifica se $\lim_{h\to 0} |w_n^{(h)}| \leq C\varepsilon$ (con $C > 0$)

Come si verifica la zero-stabilità?

- Si deve verificare che $\lim_{h\to 0} |w_n^{(h)}| = \lim_{h\to 0} |z_n^{(h)} - u_n^{(h)}| = 0$ dove $w_n^{(h)}$ soddisfa

$$\begin{cases} w_j^{(h)} = \delta_j \qquad j = 0, 1, \ldots, k-1 \\ \Delta\Psi_{n+k} = \Psi(x_n; z_{n+k}^{(h)}, \ldots, z_n^{(h)}; h; f) - \Psi(x_n; u_{n+k}^{(h)}, \ldots, u_n^{(h)}; h; f) \\ \sum_{j=0}^{k} \alpha_j w_{n+j}^{(h)} = h\Delta\Psi_{n+k} + h\delta_{n+k} \end{cases}$$

  con $|\delta_j| \leq \varepsilon$ per $j = 0, \ldots, N_h$

- È utile esprimere $w_n^{(h)}$ esplicitamente in funzione di h, n e dei vari δ_j (e non per ricorrenza)
- Una volta ottenuta una formula esplicita per $w_n^{(h)}$, si verifica se $\lim_{h\to 0} |w_n^{(h)}| \leq C\varepsilon$ (con $C > 0$)
- Il problema principale è passare da una formula per $w_n^{(h)}$ definita per ricorrenza, a una formula esplicita in funzione di h, n e dei vari δ_j

Equazioni alle differenze

- Si considera l'*equazione lineare alle differenze* a coefficienti costanti di ordine k

$$\sum_{j=0}^{k} \alpha_j w_{n+j} = \psi_{n+k} \qquad n = 0, 1, \ldots \tag{14}$$

Equazioni alle differenze

- Si considera l'*equazione lineare alle differenze* a coefficienti costanti di ordine k

$$\sum_{j=0}^{k} \alpha_j w_{n+j} = \psi_{n+k} \qquad n = 0, 1, \ldots \tag{14}$$

- Si associa all'equazione alle differenze omogenea

$$\sum_{j=0}^{k} \alpha_j w_{n+j} = 0 \qquad n = 0, 1, \ldots \tag{15}$$

il polinomio caratteristico $\Pi \in \mathbb{P}_k$ definito da $\Pi(r) = \sum_{j=0}^{k} \alpha_j r^j$

Equazioni alle differenze

- Si considera l'*equazione lineare alle differenze* a coefficienti costanti di ordine k

$$\sum_{j=0}^{k} \alpha_j w_{n+j} = \psi_{n+k} \qquad n = 0, 1, \ldots \tag{14}$$

- Si associa all'equazione alle differenze omogenea

$$\sum_{j=0}^{k} \alpha_j w_{n+j} = 0 \qquad n = 0, 1, \ldots \tag{15}$$

il polinomio caratteristico $\Pi \in \mathbb{P}_k$ definito da $\Pi(r) = \sum_{j=0}^{k} \alpha_j r^j$

- Se denotiamo con r_j, $j = 0, \ldots, k-1$ le radici di $\Pi(r)$, allora ogni sequenza della forma

$$\{r_j^n \ , \ n = 0, 1, \ldots\} \qquad \text{per} \quad j = 0, \ldots, k-1$$

è soluzione dell'equazione (15)

Equazioni alle differenze

- Se $\Pi(r) = 0$ (che è di grado k) ha esattamente k *radici distinte* $r_0, \ldots, r_{k-1}$ allora le k successioni $\{r_j^n\,,\ n = 0, 1, \ldots\}$ sono soluzioni fondamentali dell'equazione omogenea (15) e la soluzione generale è data dalla loro combinazione lineare

$$w_n^{(0)} = \sum_{j=0}^{k-1} \gamma_j r_j^n$$

Equazioni alle differenze

- Se $\Pi(r) = 0$ (che è di grado k) ha esattamente k *radici distinte* $r_0, \ldots, r_{k-1}$ allora le k successioni $\{r_j^n\,,\ n = 0, 1, \ldots\}$ sono soluzioni fondamentali dell'equazione omogenea (15) e la soluzione generale è data dalla loro combinazione lineare

$$w_n^{(0)} = \sum_{j=0}^{k-1} \gamma_j r_j^n$$

- Se $\Pi(r) = 0$ (che è di grado k) ha $p+1$ *radici distinte* $r_0, \ldots, r_p$ aventi rispettivamente molteplicità $m_0, \ldots, m_p$ ($p < k$ e $m_0 + \cdots + m_p = k$) allora tutte le soluzioni dell'equazione omogenea (15) si possono scrivere come

$$w_n^{(0)} = \sum_{j=0}^{p} \left(\sum_{i=0}^{m_i - 1} \gamma_{ij} n^i \right) r_j^n$$

Equazioni alle differenze

- La soluzione generale dell'equazione $\sum_{j=0}^{k} \alpha_j u_{n+j} = \psi_{n+k}$ (con $n = 0, 1, \ldots$) è data da

$$w_n^{(h)} = w_n^{(0)} + w_n^{(\psi)}$$

in cui $w_n^{(0)}$ è la soluzione dell'equazione omogenea associata e $w_n^{(\psi)}$ è una particolare soluzione dell'equazione non omogenea

Equazioni alle differenze

- La soluzione generale dell'equazione $\sum_{j=0}^{k} \alpha_j u_{n+j} = \psi_{n+k}$ (con $n = 0, 1, \ldots$) è data da

$$w_n^{(h)} = w_n^{(0)} + w_n^{(\psi)}$$

in cui $w_n^{(0)}$ è la soluzione dell'equazione omogenea associata e $w_n^{(\psi)}$ è una particolare soluzione dell'equazione non omogenea

- Per calcolare una soluzione particolare $w_n^{(\psi)}$ si può utilizzare il metodo della *variazione delle costanti*. Se per esempio $\psi_n = cQ(n)$ dove c è una costante e Q è un polinomio di grado s, $w_n^{(\psi)}$ sarà un polinomio di grado s (in funzione di n) del tipo

$$w_n^{(\psi)} = c \sum_{j=0}^{s} b_j n^j$$

in cui le costanti $b_0, \ldots, b_s$ devono essere calcolare in modo tale che $w_n^{(\psi)}$ sia una soluzione l'equazione alle differenze non omogenea

Esempio: verifica di zero-stabilità

- Consideriamo il metodo seguente:

$$\begin{cases} u_0\,,\, u_1 \in \mathbb{R} \\ u_{n+2} - 3u_{n+1} + 2u_n = h(f_{n+1} - 2f_n) \end{cases}$$

  applicato al problema di Cauchy $y' = 2x$ con $y(0) = 0$
- Poiché $x_0 = 0$ possiamo scrivere il generico x_n come $x_n = nh$, perciò

$$(f_{n+1} - 2f_n) = 2x_{n+1} - 4x_n = [2(n+1)h - 4nh] = 2h(1-n)$$

- L'equazione alle differenze è quindi $u_{n+2} - 3u_{n+1} + 2u_n = 2h^2(1-n)$
- Applicando lo stesso metodo per il sistema perturbato abbiamo $z_{n+2} - 3z_{n+1} + 2z_n = 2h^2(1-n) + h\delta_{n+2}$ e le perturbazioni sono definite da

$$\begin{cases} w_0 = \delta_0\,,\, w_1 = \delta_1 \\ w_{n+2} - 3w_{n+1} + 2w_n = h\delta_{n+2} \end{cases}$$

Esempio: verifica di zero-stabilità (2)

- Il polinomio caratteristico associato all'equazione omogenea

$$w_{n+2} - 3w_{n+1} + 2w_n = 0$$

è dato da

$$r^2 - 3r + 2 = 0$$

che possiede due radici distinte $r_0 = 1$ e $r_1 = 2$

- La soluzione generale dell'equazione omogenea è

$$w_n^{(0)} = \gamma_0 r_0^n + \gamma_1 r_1^n = \gamma_0 + \gamma_1 2^n$$

- Poiché $\psi_{n+2} = h\delta_{n+2}$, si cerca una soluzione particolare dell'equazione alle differenze tra le funzioni del tipo

$$w_n^{(\psi)} = \sum_{j=0}^{n} a_j \delta_j$$

Esempio: verifica di zero-stabilità (3)

- Siccome deve valere $w_{n+2}^{(\psi)} - 3w_{n+1}^{(\psi)} + 2w_n^{(\psi)} = h\delta_{n+2}$ abbiamo (per $n \geq 0$)

$$\begin{aligned} h\delta_{n+2} &= \sum_{j=0}^{n+2} a_j\delta_j - 3\sum_{j=0}^{n+1} a_j\delta_j + 2\sum_{j=0}^{n} a_j\delta_j \\ &= a_{n+2}\delta_{n+2} - 2a_{n+1}\delta_{n+1} \end{aligned}$$

ovvero $a_{n+2} = h + 2a_{n+1}\frac{\delta_{n+1}}{\delta_{n+2}}$ e quindi la soluzione particolare si scrive come

$$\begin{cases} a_0 = a_1 = 1 \\ a_j = h + 2a_{j-1}\dfrac{\delta_{j-1}}{\delta_j} & j \geq 2 \\ w_n^{(\psi)} = \delta_0 + \delta_1 + \displaystyle\sum_{j=2}^{n} a_j\delta_j & n \geq 2 \end{cases}$$

- La soluzione generale dell'equazione alle differenze non omogenea perciò si scrive come

$$w_n = w_n^{(0)} + w_n^{(\psi)} = \gamma_0 + \gamma_1 2^n + \delta_0 + \delta_1 + \sum_{j=2}^{n} a_j \delta_j$$

- Le costanti γ_0 e γ_1 si determinano utilizzando le condizioni iniziali $w_0 = \delta_0$ e $w_1 = \delta_1$:

$$\begin{cases} w_0 = \gamma_0 + \gamma_1 + \delta_0 = \delta_0 \\ w_1 = \gamma_0 + 2\gamma_1 + \delta_0 + \delta_1 = \delta_1 \end{cases} \Rightarrow \begin{cases} \gamma_0 = \delta_0 \\ \gamma_1 = -\delta_0 \end{cases}$$

- Quindi la soluzione calcolata con il metodo numerico si può scrivere come

$$\begin{cases} a_0 = a_1 = 1 & \\ a_j = h + 2a_{j-1}\dfrac{\delta_{j-1}}{\delta_j} & j \geq 2 \\ w_n = \delta_0(2 - 2^n) + \delta_1 + \displaystyle\sum_{j=2}^{n} a_j \delta_j & n \geq 2 \end{cases}$$

- Utilizzando la relazione $a_j\delta_j = h\delta_j + 2a_{j-1}\delta_{j-1}$ (e dopo un pò di algebra) si ottiene

$$\sum_{j=2}^{n} a_j\delta_j = \delta_1 \sum_{j=1}^{n-1} 2^j + h\sum_{k=2}^{n}\left[\delta_k\left(\sum_{j=0}^{n-k} 2^j\right)\right]$$

$$= \delta_1(2^n - 2) + h\sum_{k=2}^{n}\delta_k(2^{n-k+1} - 1)$$

e quindi le perturbazioni si scrivono come

$$w_n^{(h)} = \delta_1 + (\delta_1 - \delta_0)(2^n - 2) + h\sum_{k=2}^{n}\delta_k(2^{n-k+1} - 1) \qquad n \geq 2$$

- Utilizzando la relazione $a_j\delta_j = h\delta_j + 2a_{j-1}\delta_{j-1}$ (e dopo un pò di algebra) si ottiene

$$\sum_{j=2}^{n} a_j\delta_j = \delta_1 \sum_{j=1}^{n-1} 2^j + h\sum_{k=2}^{n}\left[\delta_k\left(\sum_{j=0}^{n-k} 2^j\right)\right]$$

$$= \delta_1(2^n - 2) + h\sum_{k=2}^{n}\delta_k(2^{n-k+1} - 1)$$

  e quindi le perturbazioni si scrivono come

$$w_n^{(h)} = \delta_1 + (\delta_1 - \delta_0)(2^n - 2) + h\sum_{k=2}^{n}\delta_k(2^{n-k+1} - 1) \qquad n \geq 2$$

- Poiché in genere sarà $\delta_1 \neq \delta_0$, se fissiamo $x_n = nh$ e facciamo tendere $h \to 0$ abbiamo che

$$\lim_{\substack{h\to 0 \\ x_n = nh}} |w_n^{(h)}| = +\infty$$

  e quindi il metodo non è zero-stabile

Assoluta stabilità

- Al contrario della zero-stabilità (in cui si tiene fisso x_n e si fa tendere $h \to 0$), si considera il comportamento con h fissato e $n \to \infty$ (facendo riferimento ad un particolare problema di Cauchy)

Assoluta stabilità

- Al contrario della zero-stabilità (in cui si tiene fisso x_n e si fa tendere $h \to 0$), si considera il comportamento con h fissato e $n \to \infty$ (facendo riferimento ad un particolare problema di Cauchy)
- Più precisamente, si considera il problema modello seguente:

$$\begin{cases} y'(x) = \lambda y(x) & x > 0 \text{ e } \lambda \in \mathbb{C} \\ y(0) = 1 \end{cases} \tag{16}$$

la cui soluzione è $y(x) = e^{\lambda x}$ (notiamo che $\lim_{x\to\infty} |y(x)| = 0$ se $\mathrm{Re}(\lambda) < 0$)

Assoluta stabilità

- Al contrario della zero-stabilità (in cui si tiene fisso x_n e si fa tendere $h \to 0$), si considera il comportamento con h fissato e $n \to \infty$ (facendo riferimento ad un particolare problema di Cauchy)
- Più precisamente, si considera il problema modello seguente:

$$\begin{cases} y'(x) = \lambda y(x) & x > 0 \text{ e } \lambda \in \mathbb{C} \\ y(0) = 1 \end{cases} \tag{16}$$

  la cui soluzione è $y(x) = e^{\lambda x}$ (notiamo che $\lim_{x\to\infty} |y(x)| = 0$ se $\mathrm{Re}(\lambda) < 0$)
- Un metodo numerico per la risoluzione di (16) in cui $\mathrm{Re}(\lambda) < 1$ è *assolutamente stabile* se (per h fissato) $\lim_{n\to\infty} |u_n| = 0$

Assoluta stabilità

- Al contrario della zero-stabilità (in cui si tiene fisso x_n e si fa tendere $h \to 0$), si considera il comportamento con h fissato e $n \to \infty$ (facendo riferimento ad un particolare problema di Cauchy)
- Più precisamente, si considera il problema modello seguente:

$$\begin{cases} y'(x) = \lambda y(x) & x > 0 \text{ e } \lambda \in \mathbb{C} \\ y(0) = 1 \end{cases} \tag{16}$$

  la cui soluzione è $y(x) = e^{\lambda x}$ (notiamo che $\lim_{x\to\infty} |y(x)| = 0$ se $\mathrm{Re}(\lambda) < 0$)
- Un metodo numerico per la risoluzione di (16) in cui $\mathrm{Re}(\lambda) < 1$ è *assolutamente stabile* se (per h fissato) $\lim_{n\to\infty} |u_n| = 0$
- La soluzione u_n dipende da h e λ.
- Si definisce *regione di assoluta stabilità* il seguente sottoinsieme del piano complesso

$$\mathscr{A} = \left\{ z = \lambda h \in \mathbb{C} \text{ tali che } \lim_{n\to\infty} |u_n| = 0 \right\}$$

Analizziamo l'assoluta stabilità di alcuni metodi visti finora

Analizziamo l'assoluta stabilità di alcuni metodi visti finora

Metodo di Eulero in avanti: $u_{n+1} = u_n + hf(x_n, u_n)$

Il metodo di Eulero in avanti per il problema (16) si scrive:

$$u_{n+1} = u_n + h\lambda u_n = u_n(1 + \lambda h) = (1 + \lambda h)^{n+1}$$

perciò $\lim_{n\to\infty} |u_n| = 0$ se $|1 + \lambda h| < 1$

Analizziamo l'assoluta stabilità di alcuni metodi visti finora

Metodo di Eulero in avanti: $u_{n+1} = u_n + hf(x_n, u_n)$

Il metodo di Eulero in avanti per il problema (16) si scrive:

$$u_{n+1} = u_n + h\lambda u_n = u_n(1+\lambda h) = (1+\lambda h)^{n+1}$$

perciò $\lim_{n\to\infty} |u_n| = 0$ se $|1+\lambda h| < 1$

Metodo di Eulero all'indietro: $u_{n+1} = u_n + hf(x_{n+1}, u_{n+1})$

Il metodo di Eulero all'indietro per il problema (16) si scrive:

$$u_{n+1} = u_n + h\lambda u_{n+1} \quad \Rightarrow \quad u_n = \frac{1}{(1-\lambda h)^n}$$

perciò $\lim_{n\to\infty} |u_n| = 0$ se $|1-\lambda h| > 1$

Metodo di Crank-Nicholson: $u_{n+1} = u_n + \frac{h}{2}\left(f(x_n, u_n) + f(x_{n+1}, u_{n+1})\right)$

$$u_{n+1} = u_n + \frac{h}{2}\lambda(u_n + u_{n+1}) \quad \Rightarrow \quad u_n = \left(\frac{1+\frac{\lambda h}{2}}{1-\frac{\lambda h}{2}}\right)^n$$

perciò $\lim_{n\to\infty} |u_n| = 0$ se $\mathrm{Re}(\lambda h) < 0$

Metodo di Crank-Nicholson: $u_{n+1} = u_n + \frac{h}{2}\left(f(x_n,u_n) + f(x_{n+1},u_{n+1})\right)$

$$u_{n+1} = u_n + \frac{h}{2}\lambda(u_n + u_{n+1}) \quad \Rightarrow \quad u_n = \left(\frac{1+\frac{\lambda h}{2}}{1-\frac{\lambda h}{2}}\right)^n$$

perciò $\lim_{n\to\infty} |u_n| = 0$ se $\mathrm{Re}(\lambda h) < 0$

Metodo di Heun: $u_{n+1} = u_n + \frac{h}{2}\left(f(x_n,u_n) + f(x_{n+1},u_n + hf(x_n,u_n))\right)$

$$u_{n+1} = u_n + \frac{\lambda h}{2}(2u_n + \lambda h u_n) \quad \Rightarrow \quad u_n = \left[1 + \lambda h + \frac{(\lambda h)^2}{2}\right]^n$$

perciò $\lim_{n\to\infty} |u_n| = 0$ se $\left|1 + \lambda h + \frac{(\lambda h)^2}{2}\right| < 1$

Metodo di Crank-Nicholson: $u_{n+1} = u_n + \frac{h}{2}\left(f(x_n,u_n) + f(x_{n+1},u_{n+1})\right)$

$$u_{n+1} = u_n + \frac{h}{2}\lambda(u_n + u_{n+1}) \quad \Rightarrow \quad u_n = \left(\frac{1+\frac{\lambda h}{2}}{1-\frac{\lambda h}{2}}\right)^n$$

perciò $\lim_{n\to\infty} |u_n| = 0$ se $\mathrm{Re}(\lambda h) < 0$

Metodo di Heun: $u_{n+1} = u_n + \frac{h}{2}\left(f(x_n,u_n) + f(x_{n+1},u_n + hf(x_n,u_n))\right)$

$$u_{n+1} = u_n + \frac{\lambda h}{2}(2u_n + \lambda h u_n) \quad \Rightarrow \quad u_n = \left[1 + \lambda h + \frac{(\lambda h)^2}{2}\right]^n$$

perciò $\lim_{n\to\infty} |u_n| = 0$ se $\left|1 + \lambda h + \frac{(\lambda h)^2}{2}\right| < 1$

Un metodo è definito essere *A-stabile* se $\mathscr{A} \cap \mathbb{C}^- = \mathbb{C}^-$, cioè se $\lim_{n\to\infty} |u_n| = 0$ per $\mathrm{Re}(\lambda) < 0$ indipendentemente da h.

- Usando il metodo visto in precedenza riguardo alla soluzione dell'equazione alle differenze si ha un approccio generale per verificare l'assoluta stabilità

- Usando il metodo visto in precedenza riguardo alla soluzione dell'equazione alle differenze si ha un approccio generale per verificare l'assoluta stabilità
- Notiamo infatti che applicando un metodo a k passi

$$\sum_{j=0}^{k} \alpha_j u_{n+j} = h\Psi(x_n; u_{n+k}, \ldots, u_n; f)$$

all'equazione differenziale $y'(x) = \lambda y(x)$, si può ottenere l'equazione alle differenze

$$\sum_{j=0}^{k} \alpha_j u_{n+j} - h\lambda \sum_{j=0}^{k} C_j(h\lambda) u_{n+j} = 0 \tag{17}$$

dove $C_j(h\lambda)$ dipende dal metodo usato

- Metodi multistep: $C_j(h\lambda) = \beta_j \in \mathbb{R}$
- Metodi di Runge-Kutta: $C_j(h\lambda) = 0$ per $j \neq 0$ e $C_0(h\lambda)$ è un polinomio (metodi espliciti) o una funzione razionale (metodi impliciti)

- Data l'equazione alle differenze (17), possiamo scrivere la soluzione u_n come combinazione lineare di r_i^n dove $r_i(h\lambda)$ sono le radici del polinomio caratteristico associato all'eq. (17) (chiamato *polinomio di stabilità*)

$$\Pi(r) = \rho(r) - h\lambda\sigma(r) \tag{18}$$

dove

$$\rho(r) = \sum_{j=0}^{k} \alpha_j r^j \qquad \text{e} \qquad \sigma(r) = \sum_{j=0}^{k} C_j(h\lambda) r^j$$

- Data l'equazione alle differenze (17), possiamo scrivere la soluzione u_n come combinazione lineare di r_i^n dove $r_i(h\lambda)$ sono le radici del polinomio caratteristico associato all'eq. (17) (chiamato *polinomio di stabilità*)

$$\Pi(r) = \rho(r) - h\lambda\sigma(r) \tag{18}$$

dove

$$\rho(r) = \sum_{j=0}^{k} \alpha_j r^j \qquad \text{e} \qquad \sigma(r) = \sum_{j=0}^{k} C_j(h\lambda) r^j$$

- La verifica della proprietà $\lim_{n\to\infty} |u_n| = 0$ si riduce alla verifica della condizione $|r_i(h\lambda)| < 1$ dove $r_i(h\lambda)$ sono le radici del polinomio di stabilità (18)

- Data l'equazione alle differenze (17), possiamo scrivere la soluzione u_n come combinazione lineare di r_i^n dove $r_i(h\lambda)$ sono le radici del polinomio caratteristico associato all'eq. (17) (chiamato *polinomio di stabilità*)

$$\Pi(r) = \rho(r) - h\lambda\sigma(r) \tag{18}$$

dove

$$\rho(r) = \sum_{j=0}^{k} \alpha_j r^j \qquad \text{e} \qquad \sigma(r) = \sum_{j=0}^{k} C_j(h\lambda) r^j$$

- La verifica della proprietà $\lim_{n\to\infty} |u_n| = 0$ si riduce alla verifica della condizione $|r_i(h\lambda)| < 1$ dove $r_i(h\lambda)$ sono le radici del polinomio di stabilità (18)

Condizione di assoluta stabilità

Un metodo la cui equazione alle differenze è della forma (17) è assolutamente stabile per un valore assegnato di λh se e solo se per tale valore tutte le radici $r_i(h\lambda)$ del polinomio di stabilità $\Pi(r) = \rho(r) - h\lambda\sigma(r)$ verificano $|r_i(h\lambda)| < 1$

Condizione delle radici

Un metodo del tipo $\sum_{j=0}^{k} \alpha_j u_{n+h} = h\Psi(x_n; u_{n+k}, \ldots, u_n; h; f)$ verifica la *condizione delle radici* se tutte le radici r_i del polinomio caratteristico $\rho(r)$ verificano $|r_i| \leq 1$. Inoltre le radici per cui $|r_i| = 1$ devono essere semplici (ovvero $\rho'(r_i) \neq 0$)

Condizione delle radici

Un metodo del tipo $\sum_{j=0}^{k} \alpha_j u_{n+h} = h\Psi(x_n; u_{n+k}, \ldots, u_n; h; f)$ verifica la *condizione delle radici* se tutte le radici r_i del polinomio caratteristico $\rho(r)$ verificano $|r_i| \leq 1$. Inoltre le radici per cui $|r_i| = 1$ devono essere semplici (ovvero $\rho'(r_i) \neq 0$)

Osservazione

Le radici r_i di $\rho(r)$, sono le radici $r_i(\lambda h)$ del polinomio di stabilità $\Pi(r) = \rho(r) - h\lambda\sigma(r)$ valutate per $\lambda h = 0$

Condizione delle radici

Un metodo del tipo $\sum_{j=0}^{k} \alpha_j u_{n+h} = h\Psi(x_n; u_{n+k}, \ldots, u_n; h; f)$ verifica la *condizione delle radici* se tutte le radici r_i del polinomio caratteristico $\rho(r)$ verificano $|r_i| \leq 1$. Inoltre le radici per cui $|r_i| = 1$ devono essere semplici (ovvero $\rho'(r_i) \neq 0$)

Osservazione

Le radici r_i di $\rho(r)$, sono le radici $r_i(\lambda h)$ del polinomio di stabilità $\Pi(r) = \rho(r) - h\lambda\sigma(r)$ valutate per $\lambda h = 0$

Relazione tra polinomio caratteristico e consistenza

Se un metodo del tipo $\sum_{j=0}^{k} \alpha_j u_{n+h} = h\Psi(x_n; u_{n+k}, \ldots, u_n; h; f)$ è consistente allora il suo polinomio caratteristico $\rho(r)$ ammette la radice $r = 1$

Condizione forte delle radici

Un metodo del tipo $\sum_{j=0}^{k} \alpha_j u_{n+h} = h\Psi(x_n; u_{n+k}, \ldots, u_n; h; f)$ verifica la *condizione forte delle radici* se verifica la condizione delle radici e $r = 1$ è l'unica radice (semplice) tale che $|r_i| = 1$.

Condizione assoluta delle radici

Un metodo del tipo $\sum_{j=0}^{k} \alpha_j u_{n+h} = h\Psi(x_n; u_{n+k}, \ldots, u_n; h; f)$ verifica la *condizione assoluta delle radici* se esiste $h_0 > 0$ tale che

$$|r_i(h\lambda)| < 1 \qquad i = 0, \ldots, k, \quad 0 < h \leq h_0$$

Relazione tra polinomio caratteristico e zero-stabilità

Un metodo del tipo $\sum_{j=0}^{k} \alpha_j u_{n+h} = h\Psi(x_n; u_{n+k}, \ldots, u_n; h; f)$ consistente, è zero-stabile se e solo se verifica la condizione delle radici

Relazione tra polinomio caratteristico e zero-stabilità

Un metodo del tipo $\sum_{j=0}^{k} \alpha_j u_{n+h} = h\Psi(x_n; u_{n+k}, \ldots, u_n; h; f)$ consistente, è zero-stabile se e solo se verifica la condizione delle radici

Esempio

- Nell'esercizio precedente avevamo $u_{n+2} - 3u_{n+1} + 2u_n = h(f_{n+1} - 2f_n)$
- Il polinomio caratteristico $\rho(r) = r^2 - 3r + 2$ ha radici $r_1 = 1$ e $r_2 = 2$
- Poiché $|r_2| > 1$ non è verificata la condizione delle radici $\Rightarrow$ Il metodo non è zero-stabile

Relazione tra polinomio caratteristico e zero-stabilità

Un metodo del tipo $\sum_{j=0}^{k} \alpha_j u_{n+h} = h\Psi(x_n; u_{n+k}, \ldots, u_n; h; f)$ consistente, è zero-stabile se e solo se verifica la condizione delle radici

Esempio

- Nell'esercizio precedente avevamo $u_{n+2} - 3u_{n+1} + 2u_n = h(f_{n+1} - 2f_n)$
- Il polinomio caratteristico $\rho(r) = r^2 - 3r + 2$ ha radici $r_1 = 1$ e $r_2 = 2$
- Poiché $|r_2| > 1$ non è verificata la condizione delle radici $\Rightarrow$ Il metodo non è zero-stabile

Prima barriera di Dahlquist

Non c'è nessun metodo lineare a k passi zero-stabile con ordine $q > k+1$ se k è dispari e $q > k+2$ se k è pari

- Un metodo numerico è detto *A-stabile* se la sua regione di stabilità assoluta contiene tutto il semipiano dei complessi a parte reale negativa
- Un metodo numerico è detto θ-*stabile* (o $A(\theta)$-stabile) se la sua regione di stabilità assoluta contiene la regione angolare degli $z \in \mathbb{C}$ tali che $-\theta < \pi - \arg(z) < \theta$ con $0 < \theta < \frac{\pi}{2}$
- Un metodo numerico è detto A_0-stabile se la sua regione di stabilità assoluta contiene il semiasse dei reali negativi

- Un metodo numerico è detto *A-stabile* se la sua regione di stabilità assoluta contiene tutto il semipiano dei complessi a parte reale negativa
- Un metodo numerico è detto θ-*stabile* (o $A(\theta)$-stabile) se la sua regione di stabilità assoluta contiene la regione angolare degli $z \in \mathbb{C}$ tali che $-\theta < \pi - \arg(z) < \theta$ con $0 < \theta < \frac{\pi}{2}$
- Un metodo numerico è detto A_0-stabile se la sua regione di stabilità assoluta contiene il semiasse dei reali negativi

Seconda barriera di Dahlquist

- Un metodo a più passi lineare esplicito non può essere A_0-stabile
- Non esistono metodi a più passi lineari A-stabili con ordine superiore a 2
- Per ogni $\theta \in (0, \pi/2)$, esistono solo metodi θ-stabili a k passi lineari di ordine k per $k = 3$ e $k = 4$

Stabilità forte

- Un metodo del tipo $\sum_{j=0}^{k} \alpha_j u_{n+h} = h\Psi(x_n; u_{n+k}, \ldots, u_n; h; f)$ è fortemente stabile se è consistente e verifica la condizione forte delle radici. (In particolare, un metodo fortemente stabile è sempre zero-stabile)

Stabilità forte

- Un metodo del tipo $\sum_{j=0}^{k} \alpha_j u_{n+h} = h\Psi(x_n; u_{n+k}, \ldots, u_n; h; f)$ è fortemente stabile se è consistente e verifica la condizione forte delle radici. (In particolare, un metodo fortemente stabile è sempre zero-stabile)

Metodi di Adams: $u_{n+1} = u_n + h\sum_{j=-1}^{k-1} b_j f_{n-j}$

- Nel caso $y'(x) = \lambda x$ si ha $f_j = \lambda u_j$ perciò $u_{n+1} = u_n + h\lambda \sum_{j=-1}^{k-1} b_j u_{n-j}$
- $\Pi(r) = r^k - r^{k-1} - h\lambda \sum_{j=0}^{k} b_j r^{k-j}$
- $\rho(r) = r^k - r^{k-1} = r^{k-1}(r-1)$
- $r = 1$ è una radice semplice del polinomio caratteristico e le altre radici sono $|r_i| = 0 < 1 \Rightarrow$ il metodo è fortemente stabile

Stabilità relativa

- La nozione di assoluta stabilità è legata al fatto che siamo interessati a problemi con soluzione $y(x)$ che verifica

$$\lim_{x \to \infty} y(x) = 0$$

Questo corrisponde per il problema modello $y'(x) = \lambda y(x)$ a scegliere λ tale che $\mathrm{Re}(\lambda) < 0$

Stabilità relativa

- La nozione di assoluta stabilità è legata al fatto che siamo interessati a problemi con soluzione $y(x)$ che verifica

$$\lim_{x \to \infty} y(x) = 0$$

  Questo corrisponde per il problema modello $y'(x) = \lambda y(x)$ a scegliere λ tale che $\mathrm{Re}(\lambda) < 0$
- Non si può richiedere l'assoluta stabilità per i problemi con soluzione tale che

$$\lim_{x \to \infty} |y(x)| = +\infty$$

Stabilità relativa

- La nozione di assoluta stabilità è legata al fatto che siamo interessati a problemi con soluzione $y(x)$ che verifica

$$\lim_{x\to\infty} y(x) = 0$$

  Questo corrisponde per il problema modello $y'(x) = \lambda y(x)$ a scegliere λ tale che $\mathrm{Re}(\lambda) < 0$
- Non si può richiedere l'assoluta stabilità per i problemi con soluzione tale che

$$\lim_{x\to\infty} |y(x)| = +\infty$$

- L'errore non deve aumentare più velocemente della soluzione al crescere di n
- La stabilità relativa corrisponde al fatto che l'errore relativo $\frac{|y_n - u_n|}{|y_n|}$ rimane abbastanza piccolo al crescere di n

- Le radici del polinomio di stabilità $\Pi(r)$ tendono alle radici di $\rho(r)$ quando $h\lambda \to 0$
- Per avere consistenza si deve avere $\rho(1) = 0$, e per la zero stabilità $r = 1$ deve essere radice semplice
- Allora esiste una radice del polinomio di stabilità che tende a 1 quando $h \to 0$ ed inoltre essa è unica. Questa radice è chiamata *radice principale*. Le altre radici si chiamano *radici spurie* (o *parassite*).
- Per h fissato e al crescere di n, vogliamo che i termini spuri rimangano piccoli rispetto al termine principale

- Le radici del polinomio di stabilità $\Pi(r)$ tendono alle radici di $\rho(r)$ quando $h\lambda \to 0$
- Per avere consistenza si deve avere $\rho(1) = 0$, e per la zero stabilità $r = 1$ deve essere radice semplice
- Allora esiste una radice del polinomio di stabilità che tende a 1 quando $h \to 0$ ed inoltre essa è unica. Questa radice è chiamata *radice principale*. Le altre radici si chiamano *radici spurie* (o *parassite*).
- Per h fissato e al crescere di n, vogliamo che i termini spuri rimangano piccoli rispetto al termine principale

Stabilità relativa

Un metodo è detto *relativamente stabile* per un dato $\lambda h \in \mathbb{C}$ se le radici del polinomio di stabilità verificano le condizioni seguenti

$$|r_i(\lambda h)| < |r_1(\lambda h)| \qquad i = 2, \ldots, k$$

dove $r_1(\lambda h)$ è la radice principale

- Consideriamo un metodo $\sum_{j=0}^{k} \alpha_j u_{n+h} = h\Psi(x_n; u_{n+k}, \ldots, u_n; h; f)$ applicato al problema $y'(x) = \lambda y(x)$ con $y(0) = y_0$ e $\mathrm{Re}(\lambda) < 0$
- Siano $r_i(\lambda h)$, $i = 1, \ldots, k$ le radici del polinomio di stabilità $\Pi(r)$

- Consideriamo un metodo $\sum_{j=0}^{k} \alpha_j u_{n+h} = h\Psi(x_n; u_{n+k}, \ldots, u_n; h; f)$ applicato al problema $y'(x) = \lambda y(x)$ con $y(0) = y_0$ e $\mathrm{Re}(\lambda) < 0$
- Siano $r_i(\lambda h)$, $i = 1, \ldots, k$ le radici del polinomio di stabilità $\Pi(r)$

$$\underline{\text{Convergenza}} \quad \Leftrightarrow \quad \begin{cases} \underline{\text{Zero stabilità}} \\ |r_i(0)| \leq 1 \quad i = 1, \ldots k \\ \text{se } |r_i(0)| = 1 \quad \Rightarrow \quad \rho'(r_i(0)) \neq 0 \\ \\ \underline{\text{Consistenza}} \\ \text{condizione necessaria: } r_1(0) = 1 \text{ ovvero } \sum_{j=0}^{k} \alpha_j = 0 \\ \\ \underline{\text{Dati iniziali}} \\ \lim_{h \to 0} u_n = y_0 \qquad n = 0, \ldots, k-1 \end{cases}$$

Stabilità: riepilogo

Dato un metodo consistente abbiamo:

$$A\text{-stabilità}$$
$$\forall h > 0 \quad |r_i(\lambda h)| < 1 \quad i = 1, \ldots, k$$
$$\Downarrow$$
$$\text{Stabilità assoluta}$$
$$\exists h_0 > 0 \text{ tale che } \forall 0 < h \leq h_0 \quad |r_i(\lambda h)| < 1 \quad i = 1, \ldots, k$$
$$\Downarrow$$
$$\text{Stabilità forte}$$
$$r_1(0) = 1 \text{ e } |r_i(\lambda h)| < 1 \quad i = 2, \ldots, k$$
$$\Downarrow$$
$$\text{Stabilità relativa}$$
$$|r_i(\lambda h)| < |r_1(\lambda h)| \quad i = 2, \ldots, k$$
$$\Downarrow$$
$$\text{Zero stabilità}$$

Regioni di assoluta stabilità: metodi di Adams

- L'ampiezza delle regioni di assoluta stabilità per i metodi di Adams diminuisce al crescere del numero di passi
- I metodi impliciti hanno un intervallo di stabilità più ampio di quelli espliciti
- I metodi di Crank-Nicholson e di Eulero all'indietro sono A-stabili $\Rightarrow$ Per $\mathrm{Re}(\lambda) < 0$ sono assolutamente incondizionatamente stabili, cioè la stabilità è assicurata indipendentemente dal passo h

Regioni di assoluta stabilità: metodi di Adams

- L'ampiezza delle regioni di assoluta stabilità per i metodi di Adams diminuisce al crescere del numero di passi
- I metodi impliciti hanno un intervallo di stabilità più ampio di quelli espliciti
- I metodi di Crank-Nicholson e di Eulero all'indietro sono A-stabili $\Rightarrow$ Per $\mathrm{Re}(\lambda) < 0$ sono assolutamente incondizionatamente stabili, cioè la stabilità è assicurata indipendentemente dal passo h

Regioni di assoluta stabilità: metodi BDF

- Non si possono definire metodi BDF zero-stabili con più di 6 passi
- Il metodo BDF zero-stabile a 2 passi è A-stabile
- I metodi BDF zero-stabili a 3, 4, 5 e 6 passi sono θ-stabili

Regioni di assoluta stabilità: metodi di Adams

- L'ampiezza delle regioni di assoluta stabilità per i metodi di Adams diminuisce al crescere del numero di passi
- I metodi impliciti hanno un intervallo di stabilità più ampio di quelli espliciti
- I metodi di Crank-Nicholson e di Eulero all'indietro sono A-stabili $\Rightarrow$ Per $\mathrm{Re}(\lambda) < 0$ sono assolutamente incondizionatamente stabili, cioè la stabilità è assicurata indipendentemente dal passo h

Regioni di assoluta stabilità: metodi BDF

- Non si possono definire metodi BDF zero-stabili con più di 6 passi
- Il metodo BDF zero-stabile a 2 passi è A-stabile
- I metodi BDF zero-stabili a 3, 4, 5 e 6 passi sono θ-stabili

Regioni di assoluta stabilità: metodi predittore-correttore

Non si può sapere a priori le proprietà di stabilità di un metodo predittore-correttore soltanto con la conoscenza della stabilità dei metodi che lo compongono

Regioni di assoluta stabilità: metodi di Runge-Kutta

- Un metodo di Runge-Kutta a s stadi si scrive come

$$\begin{cases} u_{n+1} = u_n + h\sum_{i=0}^{s} b_i K_i \\ K_i = f(x_n + c_i h, u_n + h\sum_{j=0}^{s} a_{ij} K_j) \quad i = 1, \ldots, s \end{cases}$$

che, applicato al problema modello $y'(x) = \lambda y(x)$ diventa

$$\begin{cases} u_{n+1} = u_n + h\sum_{i=0}^{s} b_i K_i \\ K_i = \lambda(u_n + h\sum_{j=0}^{s} a_{ij} K_j) \quad i = 1, \ldots, s \end{cases}$$

Regioni di assoluta stabilità: metodi di Runge-Kutta

- Introducendo i vettori $\mathbf{b} = (b_1, \ldots, b_s)^T$, $\mathbf{K} = (K_1, \ldots, K_s)^T$ e $\mathbf{1} = (1, \ldots, 1)^T$, il sistema precedente si scrive come

$$\begin{cases} u_{n+1} = u_n + h\mathbf{b}^T\mathbf{K} \\ \mathbf{K} = \lambda(u_n\mathbf{1} + hA\mathbf{K}) \end{cases}$$

Regioni di assoluta stabilità: metodi di Runge-Kutta

- Introducendo i vettori $\mathbf{b} = (b_1, \ldots, b_s)^T$, $\mathbf{K} = (K_1, \ldots, K_s)^T$ e $\mathbf{1} = (1, \ldots, 1)^T$, il sistema precedente si scrive come

$$\begin{cases} u_{n+1} = u_n + h\mathbf{b}^T\mathbf{K} \\ \mathbf{K} = \lambda(u_n\mathbf{1} + hA\mathbf{K}) \end{cases} \Rightarrow \begin{cases} \mathbf{K} = \lambda u_n (I - \lambda hA)^{-1}\mathbf{1} \\ u_{n+1} = u_n \left[1 + \lambda h\mathbf{b}^T(I - \lambda hA)^{-1}\mathbf{1}\right] \end{cases}$$

Regioni di assoluta stabilità: metodi di Runge-Kutta

- Introducendo i vettori $\mathbf{b} = (b_1, \ldots, b_s)^T$, $\mathbf{K} = (K_1, \ldots, K_s)^T$ e $\mathbf{1} = (1, \ldots, 1)^T$, il sistema precedente si scrive come

$$\begin{cases} u_{n+1} = u_n + h\mathbf{b}^T\mathbf{K} \\ \mathbf{K} = \lambda(u_n\mathbf{1} + hA\mathbf{K}) \end{cases} \Rightarrow \begin{cases} \mathbf{K} = \lambda u_n (I - \lambda hA)^{-1}\mathbf{1} \\ u_{n+1} = u_n\left[1 + \lambda h\mathbf{b}^T(I - \lambda hA)^{-1}\mathbf{1}\right] \end{cases}$$

- Perciò abbiamo

$$u_{n+1} = \left[1 + \lambda h\mathbf{b}^T(I - \lambda hA)^{-1}\mathbf{1}\right]u_n = R(\lambda h)u_n$$

  dove $R(\lambda h)$ è chiamata *funzione di stabilità*
- Il metodo di Runge-Kutta è perciò è assolutamente stabile se e solo se $|R(\lambda h)| \leq 1$

Regioni di assoluta stabilità: metodi di Runge-Kutta

- In generale $R(\lambda h) = \frac{P(\lambda h)}{Q(\lambda h)}$ dove $P(\lambda h)$ e $Q(\lambda h)$ sono polinomi algebrici a coefficienti reali di grado $\leq s$

Regioni di assoluta stabilità: metodi di Runge-Kutta

- In generale $R(\lambda h) = \frac{P(\lambda h)}{Q(\lambda h)}$ dove $P(\lambda h)$ e $Q(\lambda h)$ sono polinomi algebrici a coefficienti reali di grado $\leq s$
- Nel caso di un metodo esplicito $Q(\lambda h)$ è costante $\Rightarrow R(\lambda h)$ è un polinomio.
  - Se il metodo è di ordine $p \leq s$

$$R(\lambda h) = 1 + \lambda h + \frac{(\lambda h)^2}{2} + \cdots + \frac{(\lambda h)^p}{p!} + O((\lambda h)^{p+1})$$

  - Per i metodi di Runge-Kutta espliciti tali che $p = s$ (soltano per $s \leq 4$)

$$R(\lambda h) = 1 + \lambda h + \frac{(\lambda h)^2}{2} + \cdots + \frac{(\lambda h)^p}{p!}$$

Esempio: metodi di Runge-Kutta impliciti A-stabili

- Eulero all'indietro: $R(\lambda h) = \dfrac{1}{1-\lambda h}$
- Crank-Nicholson: $R(\lambda h) = \dfrac{1-\frac{\lambda h}{2}}{1+\frac{\lambda h}{2}}$
- Gauss-Legendre: $R(\lambda h) = \dfrac{1+\frac{\lambda h}{2}+\frac{(\lambda h)^2}{12}}{1-\frac{\lambda h}{2}+\frac{(\lambda h)^2}{12}}$

Sistemi di ODE

- Siano $\mathbf{F}\colon \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{y}\colon \mathbb{R} \to \mathbb{R}^n$ e consideriamo il problema di Cauchy non lineare

$$\begin{cases} \mathbf{y}'(x) = \mathbf{F}(x, \mathbf{y}(x)) \\ \mathbf{y}(x_0) = \mathbf{y}_0 \end{cases}$$

Sistemi di ODE

- Siano $\mathbf{F}\colon \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{y}\colon \mathbb{R}\to\mathbb{R}^n$ e consideriamo il problema di Cauchy non lineare

$$\begin{cases} \mathbf{y}'(x) = \mathbf{F}(x,\mathbf{y}(x)) \\ \mathbf{y}(x_0) = \mathbf{y}_0 \end{cases}$$

- Per studiare la stabilità di un metodo numerico applicato al sitema differenziale precedente, facciamo una linearizzazione in un intorno di $\bar{x}$

$$\begin{aligned} \mathbf{y}'(x) &\simeq \mathbf{F}(\bar{x},\mathbf{y}(\bar{x})) + \frac{\partial \mathbf{F}}{\partial x}\Big|_{(\bar{x},\mathbf{y}(\bar{x}))}(x-\bar{x}) + \frac{\partial \mathbf{F}}{\partial \mathbf{y}}\Big|_{(\bar{x},\mathbf{y}(\bar{x}))}(\mathbf{y}(x)-\mathbf{y}(\bar{x})) \\ &= \underbrace{\frac{\partial \mathbf{F}}{\partial \mathbf{y}}\Big|_{(\bar{x},\mathbf{y}(\bar{x}))}}_{A}\mathbf{y}(x) + \underbrace{\left[\mathbf{F}(\bar{x},\mathbf{y}(\bar{x})) + \frac{\partial \mathbf{F}}{\partial x}\Big|_{(\bar{x},\mathbf{y}(\bar{x}))}(x-\bar{x}) - \frac{\partial \mathbf{F}}{\partial \mathbf{y}}\Big|_{(\bar{x},\mathbf{y}(\bar{x}))}\mathbf{y}(\bar{x})\right]}_{\mathbf{b}(x)} \\ &= A\mathbf{y}(x) + \mathbf{b}(x) \end{aligned}$$

Sistemi di ODE

- Siano $\mathbf{F}\colon \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{y}\colon \mathbb{R}\to\mathbb{R}^n$ e consideriamo il problema di Cauchy non lineare

$$\begin{cases} \mathbf{y}'(x) = \mathbf{F}(x,\mathbf{y}(x)) \\ \mathbf{y}(x_0) = \mathbf{y}_0 \end{cases}$$

- Per studiare la stabilità di un metodo numerico applicato al sitema differenziale precedente, facciamo una linearizzazione in un intorno di $\bar{x}$

$$\begin{aligned} \mathbf{y}'(x) &\simeq \mathbf{F}(\bar{x},\mathbf{y}(\bar{x})) + \frac{\partial \mathbf{F}}{\partial x}\Big|_{(\bar{x},\mathbf{y}(\bar{x}))}(x-\bar{x}) + \frac{\partial \mathbf{F}}{\partial \mathbf{y}}\Big|_{(\bar{x},\mathbf{y}(\bar{x}))}(\mathbf{y}(x)-\mathbf{y}(\bar{x})) \\ &= \underbrace{\frac{\partial \mathbf{F}}{\partial \mathbf{y}}\Big|_{(\bar{x},\mathbf{y}(\bar{x}))}}_{A}\mathbf{y}(x) + \underbrace{\left[\mathbf{F}(\bar{x},\mathbf{y}(\bar{x})) + \frac{\partial \mathbf{F}}{\partial x}\Big|_{(\bar{x},\mathbf{y}(\bar{x}))}(x-\bar{x}) - \frac{\partial \mathbf{F}}{\partial \mathbf{y}}\Big|_{(\bar{x},\mathbf{y}(\bar{x}))}\mathbf{y}(\bar{x})\right]}_{\mathbf{b}(x)} \\ &= A\mathbf{y}(x) + \mathbf{b}(x) \end{aligned}$$

- Si studia la stabilità del metodo numerico applicato al problema $\mathbf{y}'(x) = A\mathbf{y}(x)$

- Se la matrice A ha n autovalori distinti λ_i, la soluzione di $\mathbf{y}'(x) = A\mathbf{y}(x)$ si scrive come

$$\mathbf{y}(x) = \sum_{i=1}^{n} c_i e^{\lambda_i x} \mathbf{v}_i$$

  dove $V = (\mathbf{v}_1, \ldots, \mathbf{v}_n)^T$ è la matrice degli autovettori e i coefficienti c_i sono determinati dalle condizioni iniziali
- La matrice A si può fattorizzare come $A = V\Lambda V^{-1}$, in cui Λ è la matrice diagonale formata dagli autovalori λ_i

$$\mathbf{y}'(x) = A\mathbf{y}(x) = V\Lambda V^{-1}\mathbf{y}(x)$$

- Facendo il cambio di variabile $\mathbf{z}(x) = V^{-1}\mathbf{y}(x)$ otteniamo $\mathbf{z}'(x) = \Lambda\mathbf{z}$ ovvero (essendo Λ una matrice diagonale) un sistema di n equazioni differenziali lineari scalari

$$\begin{cases} z_i'(x) = \lambda_i z_i(x) \\ z_i(x_0) = z_i^0 \end{cases} \qquad i = 1, \ldots, n$$

Problemi “stiff”

- Si parla di sistemi stiff quando si ha a che fare con un sistema di ODE che descrive un sistema fisico caratterizzate da velocità caratteristiche molto differenti tra loro
- Non esiste una definizione univoca. Alcuni autori danno una definizione basata sulle difficoltà incontrate nell’implementazione numericha

Problemi “stiff”

- Si parla di sistemi stiff quando si ha a che fare con un sistema di ODE che descrive un sistema fisico caratterizzate da velocità caratteristiche molto differenti tra loro
- Non esiste una definizione univoca. Alcuni autori danno una definizione basata sulle difficoltà incontrate nell’implementazione numericha

Definizione di “problema *stiff*”

Un sistema di ODE è *stiff* se, quando approssimato da uno schema numerico caratterizzato da una regione di stabilità assoluta di dimensione finita, si deve comunque scegliere un passo di discretizzazione eccessivamente piccolo rispetto alla regolarità della soluzione esatta

Problemi "stiff"

- Si parla di sistemi stiff quando si ha a che fare con un sistema di ODE che descrive un sistema fisico caratterizzate da velocità caratteristiche molto differenti tra loro
- Non esiste una definizione univoca. Alcuni autori danno una definizione basata sulle difficoltà incontrate nell'implementazione numericha

Definizione di "problema *stiff*"

Un sistema di ODE è *stiff* se, quando approssimato da uno schema numerico caratterizzato da una regione di stabilità assoluta di dimensione finita, si deve comunque scegliere un passo di discretizzazione eccessivamente piccolo rispetto alla regolarità della soluzione esatta

- I metodi condizionatamente assolutamente stabili non sono appropriati per l'approssimazione dei problemi *stiff*
- I metodi devono essere almeno θ-stabili o A-stabili, ma spesso queste proprietà non bastano $\Rightarrow$ Si introduce il concetto di L-stabilità

Problemi *stiff*

- Cosideriamo il sistema $\mathbf{y}'(x) = A\mathbf{y}(x) + \mathbf{b}(x)$ la cui soluzione si può scrivere come

$$\mathbf{y}(x) = \sum_{i=1}^{n} c_i e^{\lambda_i x} \mathbf{v}_i + \boldsymbol{\varphi}(x)$$

Problemi *stiff*

- Cosideriamo il sistema $\mathbf{y}'(x) = A\mathbf{y}(x) + \mathbf{b}(x)$ la cui soluzione si può scrivere come

$$\mathbf{y}(x) = \sum_{i=1}^{n} c_i e^{\lambda_i x} \mathbf{v}_i + \boldsymbol{\varphi}(x)$$

- Se $\mathrm{Re}(\lambda_i) < 0$, $i = 1, \ldots, n$ allora $\mathbf{y}(x)$ tende alla soluzione particolare $\boldsymbol{\varphi}(x)$ quando $x \to +\infty$
- $\boldsymbol{\varphi}(x)$ può essere interpretata come la componente "stazionaria" mentre l'altro termine corrisponde alla componente "transitoria" della soluzione

Problemi *stiff*

- Cosideriamo il sistema $\mathbf{y}'(x) = A\mathbf{y}(x) + \mathbf{b}(x)$ la cui soluzione si può scrivere come
$$\mathbf{y}(x) = \sum_{i=1}^{n} c_i e^{\lambda_i x} \mathbf{v}_i + \boldsymbol{\varphi}(x)$$
- Se $\mathrm{Re}(\lambda_i) < 0$, $i = 1, \ldots, n$ allora $\mathbf{y}(x)$ tende alla soluzione particolare $\boldsymbol{\varphi}(x)$ quando $x \to +\infty$
- $\boldsymbol{\varphi}(x)$ può essere interpretata come la componente "stazionaria" mentre l'altro termine corrisponde alla componente "transitoria" della soluzione
- Se utilizziamo un metodo condizionatamente assolutamente stabile la limitazione del passo h dipende dall'autovalore di modulo massimo. D'altra parte, l'intervallo di "tempo" durante il quale la corrispondente soluzione darà un contributo significativo è piccolo
- Se siamo interessati alla soluzione stazionaria (ovvero per x grandi), il metodo scelto non è efficiente in quanto deve utilizzare un passo piccolo per poter descrivere una componente della soluzione che per x grandi svanisce

Problemi *stiff*

- Se si vuole arrivare alla soluzione stazionaria, occorre che la componente transitoria decada rapidamente

Problemi *stiff*

- Se si vuole arrivare alla soluzione stazionaria, occorre che la componente transitoria decada rapidamente
- Un metodo A-stabile assicura che $|R(\lambda h)| < 1$ anche per $|\mathrm{Re}(\lambda h)|$ grande. Però niente impedisce che $|R(\lambda h)| \lesssim 1$. Per esempio con il metodo di Crank-Nicholson abbiamo che

$$\lim_{\mathrm{Re}(\lambda h)\to-\infty} |R(\lambda h)| = \lim_{\mathrm{Re}(\lambda h)\to-\infty} \left| \frac{1-\frac{\lambda h}{2}}{1-\frac{\lambda h}{2}} \right| = 1$$

L-stabilità

Un metodo di Runge-Kutta è L-stabile se è A-stabile ed inoltre verifica la condizione

$$\lim_{\mathrm{Re}(\lambda h)\to-\infty} |R(\lambda h)| = 0$$

Esempi di metodi *A*-stabili ma non *L*-stabili

- Metodo di Crank-Nicholson
- Metodo di Gauss-Legendre

Esempi di metodi L-stabili

- Metodo di Eulero all'indietro
- Metodo di Radau IIA (di ordine 5), definito da

$$\begin{array}{c|ccc} \frac{4-\sqrt{6}}{10} & \frac{88-7\sqrt{6}}{360} & \frac{296-169\sqrt{6}}{1800} & \frac{-1+3\sqrt{6}}{225} \\ \frac{4+\sqrt{6}}{10} & \frac{296-169\sqrt{6}}{1800} & \frac{88-7\sqrt{6}}{360} & \frac{-1-3\sqrt{6}}{225} \\ 1 & \frac{16-\sqrt{6}}{36} & \frac{16+\sqrt{6}}{36} & \frac{1}{9} \\ \hline & \frac{16-\sqrt{6}}{36} & \frac{16+\sqrt{6}}{36} & \frac{1}{9} \end{array}$$

con

$$R(\lambda h) = \frac{1 + \frac{2}{5}\lambda h + \frac{1}{20}(\lambda h)^2}{1 - \frac{3}{5}\lambda h + \frac{3}{20}(\lambda h)^2 - \frac{1}{60}(\lambda h)^3}$$