Lezione 35 (ore 59,60) - 25 maggio 2016, 16:30 - 18:30 B23

(35.01) Osservazione (stima numerica della derivata seconda).

Tra le variabili di ingresso della procedura TS_1_pv compare la funzione G2, utilizzata per scegliere il passo, che dunque l'utilizzatore *deve* determinare e realizzare. Questa necessità può essere eliminata modificando la procedura di scelta del passo. Precisamente, il valore G2(t(k),x(k)) = y"(t(k)) può essere *stimato* con le considerazioni seguenti:

(a) Per definizione:

$$\frac{y'(t(k) + \tau) - y'(t(k))}{\tau} \to y''(t(k)) \quad \text{per} \quad \tau \to 0$$

dunque:

$$\textit{per } \tau \textit{ piccolo } \ \frac{y'(t(k) + \tau) - y'(t(k))}{\tau} \ \textit{ è una buona approssimazione di } y''(t(k))$$

(b) Poiché y(t) è la soluzione dell'equazione differenziale che vale x(k) all'istante t(k) si ha:

$$y'(t(k)) = F(\ t(k),y(t(k))\ ) = F(\ t(k),x(k)\ )$$

e:

$$y'(t(k) + \tau) = F(\ t(k) + \tau,y(t(k) + \tau)\ )$$

Quest'ultimo valore *non è calcolabile* perché la procedura non conosce y(t(k) + τ). Allora:

$$\textit{si approssima } y(t(k) + \tau) \textit{ con } y(t(k)) + y'(t(k))\ \tau = x(k) + F(\ t(k),x(k)\ )\ \tau$$

Complessivamente:

$$\textit{scelto } \tau \textit{ piccolo, si stima } G2(\ t(k),x(k)\ ) = y''(t(k)) \textit{ con}$$

$$\frac{F(\ t(k) + \tau,x(k) + F(\ t(k),x(k)\ )\ \tau\ ) - F(\ t(k),x(k)\ )}{\tau}$$

Questa quantità, dato τ, *è calcolabile senza usare* G2.

(35.02) Osservazione (ragionevolezza della stima).

Indicando con F(k) il valore F(t(k),x(k)), si consideri la funzione di τ:

$$H(\tau) = F(\ t(k) + \tau,x(k) + F(k)\ \tau\ )$$

Poiché si suppone che F(t,x) abbia *derivate parziali prime* continue, anche H ha derivata prima continua. Allora:

$$H(\tau) = H(0) + (H'(0) + z(\tau))\ \tau \quad \text{con } z(\tau) \to 0 \text{ per } \tau \to 0$$

Ma: $H(0) = F(k)$ e

$$H'(0) = \partial_t F(t(k),x(k)) + \partial_x F(t(k),x(k)) \cdot F(t(k),x(k)) = G2(t(k),x(k)) = y''(t(k))$$

dunque:

$$H(\tau) = F(k) + (y''(t(k)) + z(\tau))\ \tau$$

e:

$$y''(t(k)) - \frac{H(\tau) - F(k)}{\tau} = z(\tau) \to 0 \text{ per } \tau \to 0$$

(35.03) Osservazione (scelta del passo con stima numerica della derivata).

Utilizzando la stima trovata si ottiene un procedimento di scelta del passo che non richiede l'uso di G2:

- SCELTA di $h(k)$. Dati $E > 0$ e $\lambda > 0$, per ogni k si sceglie τ piccolo e si pone:

$$d(k) = \max \left\{ \lambda \ , \left\| \frac{F(t(k) + \tau, x(k) + F(k)\ \tau) - F(k)}{\tau} \right\| \right\}$$

  e poi:

$$h(k) = \min \left\{ \sqrt{\frac{2E}{d(k)}} \ , tf - t(k) \right\}$$

(35.04) Esercizio (realizzazione in Scilab).

Modificare opportunamente la procedura TS_1_pv per ottenere una function di intestazione:

```
[T, X, PASSO] = TS_1_CNS_pv(x0, t0, tf, F, E, LAMBDA, HMIN, TAU)
```

che realizza il metodo di Eulero esplicito che sceglie il passo come descritto nella Osservazione (35.03).

---

* METODI RUNGE-KUTTA *

(35.05) Esempio.

Nel metodo TS(2) è richiesta all'utilizzatore la determinazione e realizzazione di G2 (utilizzata dalla procedura per il calcolo di x(k+1)) e G3 (utilizzata dalla procedura di scelta di h(k)). In generale il compito è tanto più gravoso quanto più alto è l'ordine del metodo: nel metodo TS(p) l'utilizzatore deve determinare e realizzare G2,...,G(p) per il calcolo di x(k+1) e G(p+1) per la scelta di h(k). I *metodi Runge-Kutta* sono pensati per eliminare questo onere.

Per introdurre la *struttura* dei metodi, vediamo come si trasforma il calcolo di x(k+1) nel metodo TS(2) utilizzando la stima della derivata seconda dell'Osservazione (35.01).

In TS(2):

$$x(k+1) = x(k) + F(\ t(k),x(k)\ )\ h(k) + \frac{1}{2}\ G2(\ t(k),x(k)\ )\ h(k)^2$$

Scegliendo τ = h(k) nella stima dell'Osservazione (35.01) si ottiene:

$$G2(\ t(k),x(k)\ ) = \frac{F(\ t(k) + h(k),x(k) + F(\ t(k),x(k)\ )\ h(k)\ ) - F(\ t(k),x(k)\ )}{h(k)}$$

e sostituendo (posto F(k) = F(t(k),x(k))):

$$\begin{aligned} x(k+1) &= x(k) + F(k)\ h(k) + \frac{1}{2}\ [\ F(\ t(k) + h(k),x(k) + F(k)\ h(k)\ ) - F(k)\ ]\ h(k) \\ &= x(k) + \frac{1}{2}\ [\ F(k) + F(\ t(k) + h(k),x(k) + F(k)\ h(k)\ )\ ]\ h(k) \end{aligned}$$

Questa procedura di calcolo di x(k+1) definisce (indipendentemente dalla procedura adottata per la scelta di h(k)) il *metodo di Heun* (noto anche come *metodo dei trapezi esplicito* o *metodo di Eulero esplicito modificato*) che risulta (dimostrazione omessa) di *ordine* 2 *per* h → 0 e quindi *convergente per* E → 0 di *ordine* 2/3.

(35.06) Definizione (metodi RK a due e tre stadi).

Si chiamano *metodi Runge-Kutta* (RK) *a due stadi* quelli nei quali, scelti opportunamente numeri reali c(2), a(21), b(1) e b(2), il valore x(k+1) si ottiene, dopo aver scelto h(k), ponendo:

- ST(1) = F(t(k),x(k))
- ST(2) = F(t(k) + c(2) h(k),x(k) + a(21) ST(1) h(k))

e poi

- x(k+1) = x(k) + [b(1) ST(1) + b(2) ST(2)] h(k)

Si chiamano *metodi Runge-Kutta* (RK) *a tre stadi* quelli nei quali, scelti opportunamente numeri reali c(2), c(3), a(21), a(31), a(32), b(1), b(2) e b(3), il valore x(k+1) si ottiene, dopo aver scelto h(k), ponendo:

- ST(1) = F(t(k),x(k))
- ST(2) = F(t(k) + c(2) h(k),x(k) + a(21) ST(1) h(k))
- ST(3) = F(t(k) + c(3) h(k),x(k) + [a(31) ST(1) + a(32) ST(2)] h(k))

e poi

- $x(k+1) = x(k) + [\, b(1)\, ST(1) + b(2)\, ST(2) + b(3)\, ST(3)\,]\, h(k)$

I valori dei *parametri* c(i), a(i,j), b(i) sono determinati (*non univocamente*) dalla condizione di *ottenere* metodi di *ordine per* $h \to 0$ *più alto possibile*.

(35.07) Osservazione.

- La definizione si estende a metodi con un numero intero qualsiasi di stadi.
- Per i metodi a due stadi l'ordine più alto che si può ottenere è *due*, per quelli a tre stadi è *tre*. In generale *non* è lecito aspettarsi di poter ottenere metodi di ordine q scegliendo opportunamente i valori dei parametri di un metodo a q stadi.
- Il metodo di Heun è il metodo RK *a due stadi* che si ottiene scegliendo:

$$c(2) = 1 \quad , \quad a(21) = 1 \quad , \quad b(1) = b(2) = 1/2$$

- Il metodo di Eulero esplicito è il metodo RK *ad uno stadio* ottenuto scegliendo b(1) = 1.

(35.08) Osservazione (scelta di h(k) nei metodi Runge-Kutta).

Coerentemente con l'intento di eliminare l'onere della determinazione e realizzazione di funzioni G(j), la scelta di h(k) nei metodi Runge-Kutta è usualmente effettuata come segue.

Siano: RK il metodo di Runge-Kutta, di ordine p per $h \to 0$, scelto per il calcolo di x(k+1) e RK' un altro metodo di Runge-Kutta, di ordine p' = p+1. Allora:

- SCELTA di h(k). Dati $E > 0$ e $\lambda > 0$, per ogni k si sceglie τ piccolo, si calcolano:

  (1) XX = un passo di RK a partire da x(k), di lunghezza τ
  (2) XX' = un passo di RK' a partire da x(k), di lunghezza τ

  si pone:

$$d(k) = \max \{\, \lambda \;,\; \| XX - XX' \| \,\}$$

  e poi:

$$h(k) = \min \left\{ \sqrt[p+1]{\frac{E}{d(k)}}\; \tau \;,\; tf - t(k) \right\}$$

Questa procedura di scelta si spiega considerando che:

(a) Il metodo RK ha ordine p per $h \to 0$ dunque, determinato opportunamente un numero reale C:

$$\textit{si stima } s(h) \textit{ con } C\, h^{p+1}$$

(b) Poiché:

$$XX - y(t(k) + \tau) =$$

$$= (C + z(\tau))\ \tau^{p+1}, \text{ con } z(\tau) \to 0 \text{ per } \tau \to 0 \quad \text{(RK ha ordine p)}$$

$$XX' - y(t(k) + \tau) =$$

$$= (C' + w(\tau))\ \tau^{p+2}, \text{ con } w(\tau) \to 0 \text{ per } \tau \to 0 \quad \text{(RK' ha ordine p+1)}$$

allora:

$$\frac{XX - XX'}{\tau^{p+1}} = C + [\ z(\tau) - (C' + w(\tau))\ \tau\ ] \to C \quad \text{per } \tau \to 0$$

Dunque:

$$\textit{si stima } C \textit{ con } \frac{XX - XX'}{\tau^{p+1}}$$

(c) Complessivamente:

$$\textit{si stima } s(h) \textit{ con } \frac{XX - XX'}{\tau^{p+1}}\ h^{p+1}$$

dunque:

$$\left|\left| \frac{XX - XX'}{\tau^{p+1}}\ h^{p+1} \right|\right| = E \quad \Leftrightarrow \quad h = \sqrt[p+1]{\frac{E}{||\ XX - XX'\ ||}}\ \tau$$

(35.09) Realizzazione Scilab (RK12_pv).

Come esempio di realizzazione, consideriamo il metodo RK che utilizza Eulero esplicito (ordine 1) per il calcolo di x(k+1) e che sceglie h(k) affiancandolo con il metodo di Heun (ordine 2). Ne risulta un metodo di ordine 1 per h → 0 e quindi convergente di ordine 1/2 per E → 0.

```
function [T, X, PASSO] = RK12_pv(x0, t0, tf, F, E, LAMBDA, HMIN, TAU)
//
// Integra numericamente, sull'intervallo [t0,tf], il Problema
// di Cauchy in R(n):
//
// x' = F(t,x)
// x(t0) = x0
//
// con il metodo di Eulero esplicito (RK di ordine 1) - a passo
// variabile - affiancato, per la scelta del passo, dal metodo di
// Heun (RK di ordine 2).
//
// x0: condizione iniziale (colonna di n elementi)
// t0: istante iniziale (numero reale)
// tf: istante finale (numero reale)
// F: function che definisce l'equazione differenziale - F(t,x) deve
//    essere una colonna di n numeri reali
// E: valore massimo della stima dell'errore locale (numero reale)
```

```
// LAMBDA: numero reale che stabilisce il valore massimo del passo
//         (OPZIONALE - valore predefinito: 1d-5)
// HMIN: valore minimo consentito del passo
//       (OPZIONALE - valore predefinito: (tf - t0) / 1d6)
// TAU: valore del passo per il calcolo delle stime utilizzate
//      nel calcolo di h(k) (OPZIONALE - valore predefinito: (tf - t0) / 1d3)
//
// T = [t(0),...,t(N)]: riga contenente gli istanti di integrazione
// X = [x(0),...,x(N)]: matrice n x (N+1) contenente le approssimazioni
// PASSO = [h(0),...,h(N-1)]: riga contenente i passi di integrazione
//
// Valore degli argomenti opzionali
//
if ~exists('LAMBDA','l') then LAMBDA = 1d-5; end;
if ~exists('HMIN','l') then HMIN = (tf - t0) / 1d6; end;
if ~exists('TAU','l') then TAU = (tf - t0) / 1d3; end;
//
// Inizializzazione delle variabili di uscita
//
T(1,1) = t0;
X(:,1) = x0;
PASSO = [];
//
// ciclo principale
//
while (T(1,$) < tf), // arresta la costruzione se ha raggiunto tf
  //
  // scelta del passo
  //
  // XX1 = X(:,$) + F(T(1,$),X(:,$)) * TAU;
  ST1 = F(T(1,$),X(:,$));
  ST2 = F( T(1,$) + TAU,X(:,$) + ST1 * TAU );
  // XX2 = X(:,$) + ( (ST1 + ST2)/2 ) * TAU;
  //
  //     XX1 - XX2 = (ST1 - ST2)/2 * TAU
  //
  d = max(LAMBDA, norm( ((ST1 - ST2)/2) * TAU ));
  PASSO(1,$+1) = min(sqrt(E/d) * TAU, tf - T(1,$));
  //
  // calcolo approssimazione e nuovo istante di integrazione
  //
  X(:,$+1) = X(:,$) + F(T(1,$),X(:,$)) * PASSO(1,$);
  T(1,$+1) = T(1,$) + PASSO(1,$);
  //
  // arresta la costruzione se il passo calcolato risulta troppo
  // piccolo e non ha raggiunto tf
  //
  if (PASSO(1,$) < HMIN) & (T(1,$) < tf) then break; end;
  //
end;
//
// Verifica se l'integrazione ha raggiunto tf
```

```
71 //
72 if T(1,$) < tf then
73    printf("\n\nIntegrazione interrotta a T = %3.2e", T(1,$));
74 end;
75 //
76 endfunction
77 //
78 // Esempio per assegnare valori ai parametri opzionali:
79 //
80 //    [T,X,PASSO] = TS_1_pv(x0,t0,tf,F,G,E,HMIN = y);
81 //
82 //         => LAMBDA = valore predefinito, HMIN = y, TAU = valore predefinito
83 //
```

Si osservi che:

- Nella scelta del passo la differenza XX1 - XX2 può essere determinata *senza calcolare* XX1 ed XX2 (righe 48-55). Risulta infatti:

$$\mathrm{XX1} - \mathrm{XX2} = \frac{\mathrm{ST1} - \mathrm{ST2}}{2}\,\mathrm{TAU}$$

- Per la scelta del passo si è utilizzato *lo stesso valore di* τ *ad ogni iterazione.*

Il file che contiene la procedura, insieme ad un esempio di uso, si può trovare nella pagina web del corso, sezione "altro materiale didattico".

(35.10) Esercizio.

1. Verificare che la scelta del passo con stima numerica della derivata (Osservazione (35.03) è praticamente identica alla scelta del passo adottata nella procedura RK12_pv.

2. Cercare (su web) un metodo RK a tre stadi di ordine tre per $h \to 0$.
