LECTURE 24

Error Analysis for Multi-step Methods

1. Review

In this lecture we shall study the errors and stability properties for numerical solutions of initial value problems of the form

$$\frac{dx}{dt} = f(t,x) \tag{24.1}$$

$$x(t_0) = x_0 \tag{24.2}$$

Recall that the starting point for multi-step methods such as the Adams-Bashforth and Adams-Moulton methods was actually the integral of equation (24.1)

$$x(t_{n+1}) - x(t_n) = \int_{t_n}^{t_{n+1}} \frac{dx}{dt} dt = \int_{t_n}^{t_{n+1}} f(t,x) dt$$

By replacing $f(t,x)$ by its polynomial intepolation at the points $t_n, t_{n-2}, \ldots, t_{n-4}$ one obtains the fifth order Adams-Bashforth formula

$$\begin{aligned} x_{n+1} &= x_n + \int_{t_n}^{t_{n+1}} \left(\sum_{i=n-4}^{n} f_i \ell_i(t) \right) dt \\ &= x_n + \frac{h}{720} \left(1901 f_n - 2774 f_{n-1} + 2616 f_{n-2} - 1274 f_{n-3} + 251 f_{n-4}\right) \end{aligned}$$

where

$$\begin{aligned} t_n &= t_0 + nh \\ x_n &= x(t_n) \\ f_n &= f(t_n, x_n) \end{aligned}$$

and the numerical coefficients in the second line are just the integrals of the cardinal functions $\ell_i(t)$.

Similarly, by replacing the function $f(x,t)$ by its polynomial interpolation at the points $t_{n+1}, t_{n,}, \ldots, t_{n-3}$. we obtained the Adams-Moulton formula

$$\begin{aligned} x_{n+1} &= x_n + \int_{t_n}^{t_{n+1}} \left(\sum_{i=n-3}^{n+1} f_i \ell_i(t) \right) dt \\ &= x_n + \frac{h}{720} \left(251 f_{n+1} + 646 f_n - 264 f_{n-1} + 106 f_{n-2} - 19 f_{n-3}\right) \end{aligned}$$

2. Linear Multi-step Methods

Of course, there's nothing to prevent us from calculating even higher order analogs of the Adams-Bashforth and Adams-Moulton formulae. But instead of doing so explicitly, we'll now assume that we have in our hands a **k-step linear multi-step method** given by a formula of the form

$$a_k x_n + a_{k=1} x_{n-1} + \cdots + a_1 x_{n-k-1} + a_0 x_{n-k} = h\left[b_k f_n + b_{k-1} f_{n-1} + \cdots + b_1 f_{n-k+1} + b_0 f_{n-k}\right] \tag{24.3}$$

relating k successive values of the $x(t)$ to k successive values of the function $f(t,x)$ (the function that defines the differential equation). The utility of such an equation is, of course, to compute x_n from the k preceding values, $x_{n-k}, x_{n-k+1}, \dots, x_{n-1}$, and so we shall henceforth assume that the constant $a_k \neq 0$. The coefficient b_k on the right hand side may or not equal zero. If $b_k = 0$ then we say the multi-step method corresponding to equation (24.3) is **explicit**; because, in this case we can solve the equation explicitly for x_n

$$x_n = -\frac{1}{a_k}\left(a_{k=1}x_{n-1} + \cdots + a_1 x_{n-k-1} + a_0 x_{n-k}\right) + \frac{h}{a_k}\left[b_k f_n + b_{k-1} f_{n-1} + \cdots + b_1 f_{n-k+1} + b_0 f_{n-k}\right]$$

If $b_k \neq 0$ we say that the corresponding method is **implicit**; in this case, the term $f_n = f(t_n, x_n)$ on the right hand side also depends on x_n and so we have an implicit algebraic equation for x_n.

2.1. Convergence of Multi-Step Methods.

DEFINITION 24.1. *A multi-step method defined by the a formula of the form (24.3) is said to be* ***convergent*** *in a region* $[t_0, t_1]$ *if*

$$\lim_{h\to 0} x_h(t) = x(t) \quad , \quad t \in [t_0, t_1] \tag{24.4}$$

provided only that

$$\lim_{h\to 0} x_h(t + jh) = x_0 \quad , \quad 0 \le j < k \tag{24.5}$$

Here $x_h(t)$ *is the numerical solution computed using a step size of* h *and* $x(t)$ *is the exact solution.*

This definition is natural enough. The following definitions are not so natural, but nevertheless extremely useful.

DEFINITION 24.2. *Consider a multi-step method corresponding to a relation of the form (24.3). Set*

$$\begin{aligned} P(z) &= a_k z^k + a_{k-1} z^{k-1} + \cdots + a_1 z + a_0 \\ Q(z) &= b_k z^k + b_{k-1} z^{k-1} + \cdots + b_1 z + b_0 \end{aligned}$$

We shall say that the method (24.3) is ***stable*** *if the roots of the polynomial* $P(z)$ *lie in disk* $|z| \le 1$, *and if each root such that* $|z| = 1$ *is simple. The method (24.3) is said to be* ***consistent*** *if* $P(1) = 0$ *and* $P'(1) = Q(1)$.

THEOREM 24.3. *Consider a multi-step method corresponding to a relation of the form (24.3). Then this method is convergent if and only if it is both stable and consistent.*

A partial proof (that of the necessity of stability and consistency for convergence) is given in the text.

2.2. The Order of Multi-step Method.

The **order** of a multi-step method is an interger that corresponds to the number of terms in the Taylor expansion of the solution that a multi-step method simulates. Let us represent the multi-step method (24.3) as a linear functional

$$\begin{aligned} L[x] &= \sum_{j=0}^{k} \left[a_j x(jh) - h b_j f(jh)\right] \\ &= \sum_{j=0}^{k} \left[a_j x(jh) - h b_j x'(jh)\right] \end{aligned}$$

Here we let $k = n$ to simplify our notation and assume that the first value of in equation (24.3) begins at $t = 0$, rather than $t = (n-k)h$. Now

$$x(jh) = \sum_{i=0}^{\infty} \frac{(jh)^i}{i!} x^{(i)}(0)$$

$$x'(jh) = \sum_{i=0}^{\infty} \frac{(jh)^i}{i!} x^{(i+1)}(0)$$

and so we can write

$$L[x] = \sum_{j=0}^{k} \left[a_j \sum_{i=0}^{\infty} \frac{(jh)^i}{i!} x^{(i)}(0) - hb_j \sum_{i=0}^{\infty} \frac{(jh)^i}{i!} x^{(i+1)}(0) \right]$$

Collecting terms proportional to $x(0), x'(0), \ldots$ (or equivalently by their degree in h) we have

$$L[x] = d_0 x(0) + d_1 h x'(0) + d_2 h^2 x''(0) + \cdots$$

where

$$d_0 = \sum_{i=0}^{k} a_i \tag{24.6}$$

$$d_1 = \sum_{i=0}^{k} (i a_i - b_i) \tag{24.7}$$

$$d_2 = \sum_{i=0}^{k} \left(\frac{1}{2} i^2 a_i - i b_i \right) \tag{24.8}$$

$$\vdots$$

$$d_j = \sum_{i=0}^{k} \left(\frac{i^j}{j!} a_i - \frac{i^{j-1}}{(j-1)!} b_j \right) \tag{24.9}$$

THEOREM 24.4. *The following three properties of the multi-step method (24.3) are equivalent:*

1. $d_0 = d_1 = \cdots = d_m = 0$.
2. $L[P] = 0$ *for each polynomial* P *of degree* $\leq m$.
3. $L[x]$ *is* $\mathcal{O}\left(h^{m+1}\right)$ *for all* $x \in C^{m+1}$.

Proof.

- (1) $\Rightarrow$ (2):
  If (1) is true then
  $$L[x] = 0 + 0 + \cdots + 0 + d_{m+1} h^{m+1} x^{(m+1)}(0) + d_{m+2} h^{m+2} x^{(m+2)}(0) + \cdots$$
  But if P is a polynomial of degree $\leq m$ then $P^{(m+1)}(x) = 0$. Therefore, for such a polynomial
  $$\begin{aligned} L[P] &= 0 + 0 + \cdots + 0 + d_{m+1} h^{m+1} P^{(m+1)}(0) + d_{m+2} h^{m+2} P^{(m+2)}(0) + \cdots \\ &= 0 \end{aligned}$$
- (2) $\Rightarrow$ (3):
  If $x \in C^{m+1}$, then by Taylor's Theorem we can write
  $$x(t) = P(t) + R(t)$$
  where $P(t)$ is a polynomial of degree m and
  $$R(t) = \frac{x^{(m+1)}(\xi)}{(m+1)!} t^{m+1}$$

Notice that

$$\left.\frac{d^j R}{dt^j}\right|_{t=0} = 0$$

Hence,

$$\begin{aligned} L[x] &= L[P] + L[R] \\ &= 0 + d_{m+1} h^{m+1} x^{(m+1)}(0) + d_{m+2} h^{m+2} x^{(m+2)}(0) + \cdots \\ &= \mathcal{O}\left(h^{m+1}\right) \end{aligned}$$

- (3) $\Rightarrow$ (1)
  If (3) is true, then we must have $d_0 = d_1 = \cdots = d_m = 0$. Hence, (3) implies (1).

Definition 24.5. *The* ***order*** *of a multi-step method is the unique natural number m such that*

$$0 = d_0 = d_1 = \cdots = d_m \neq d_{m+1}$$

Example 24.6. What is the order of the following multistep method

$$x_n - x_{n-2} = \frac{h}{3}\left(f_n + 4f_{n-1} + f_{n-2}\right)$$

- This is a three step method ($k = 2$), with

$$\begin{array}{ccc} a_2 = 1 & , \quad a_1 = 0 & , \quad a_0 = -1 \\ b_2 = \frac{1}{3} & , \quad b_1 = \frac{4}{3} & , \quad b_0 = \frac{1}{3} \end{array}$$

We have

$$\begin{aligned} d_0 &= \sum_{k=0}^{2} a_i = 1 + 0 - 1 = 0 \\ d_1 &= \sum_{i=0}^{k} (i a_i - b_i) = \left(0 - \frac{1}{3}\right) + \left(0 - \frac{4}{3}\right) + \left(2 - \frac{1}{3}\right) = 0 \\ d_2 &= \sum_{i=0}^{k} \left(\frac{1}{2} i^2 a_i - i b_i\right) = (0 - 0) + \left(0 - \frac{4}{3}\right) + \left(\frac{4}{2} - \frac{2}{3}\right) = 0 \\ d_3 &= \sum_{i=0}^{k} \left(\frac{1}{6} i^3 a_i - \frac{1}{2} i^2 b_i\right) = (0 - 0) + \left(0 - \frac{4}{6}\right) + \left(\frac{8}{6} - \frac{4}{6}\right) = 0 \\ d_4 &= \sum_{i=0}^{k} \left(\frac{1}{24} i^4 a_i - \frac{1}{6} i^3 b_i\right) = (0 - 0) + \left(0 - \frac{4}{18}\right) + \left(\frac{(16)(1)}{(24)} - \frac{(8)(1)}{(6)(3)}\right) = 0 \\ d_5 &= \sum_{i=0}^{k} \left(\frac{1}{120} i^5 a_i - \frac{1}{24} i^4 b_i\right) = (0 - 0) + \left(0 - \frac{(16)(4)}{24}\right) + \left(\frac{(32)(1)}{120} - \frac{(16)(1)}{(24)(3)}\right) = -\frac{118}{45} \end{aligned}$$

And so the order is $m = 4$.

2.3.

2.4. Local Truncation Error. By a **local trunction error** we mean the error induced at a particular stage of an iterative numerical procedure. In the case at hand, this means the error induced by using a difference relation of the form (24.3) instead to obtain an estimate x_n of $x(t_n)$ instead of evaluating the exact solution at t_n.

THEOREM 24.7. *Consider a multi-step method corresponding to a relation of the form (24.3). Then if $x(t) \in C^{k+2}(\mathbb{R})$ and if $\frac{\partial f}{\partial x}$ is continuous, we have*

$$x(t_n) - x_n = \frac{d_{k+1}}{a_k} h^{m+1} x^{(m+1)}(t_{n-k}) + \mathcal{O}\left(h^{m+2}\right)$$

where the coefficients are defined by equations (??) - (24.9)

Proof. It suffices to prove the statement for $n = k$, since x_n can be expressed as a solution with initial conditions imposted at t_{n-k}. Using the linear function L of the preceding section we have for the exact solution $x(t)$

$$L[x] = \sum_{i=0}^{k} \left[a_i x(t_i) - h b_i x'(t_i)\right] = \sum_{i=0}^{k} \left[a_i x(t_i) - h b_i f(t_i, x(t_i))\right] \tag{24.10}$$

On the other hand, the numerical solution $\{x_i\}$ should satisfy

$$0 = \sum_{i=0}^{k} \left[a_i x_i - h b_i f(t_i, x_i)\right] \tag{24.11}$$

Subtracting (24.11) from (24.10) yields

$$L[x] = \sum_{i=0}^{k} \left[a_i \left(x(t_i) - x_i\right) - h b_i \left(f(t_i, x(t_i)) - f(t_i, x_i)\right)\right]$$

Since we are interested here only in the error induced by the current iteration of the multi-step method we shall assume that the previously determined x_i are all exactly correct: $x_i = x(t_i)$. In this case, all but the final terms of the summand vanish and we have

$$L[x] = a_k \left(x(t_k) - x_k\right) - h b_k \left(f(t_k, x(t_k)) - f(t_k, x_k)\right)$$

Writing

$$f(t_k, x_k) = f(t_k, x(t_k)) + \frac{\partial f}{\partial x}(\xi)\left(x(t_k) - x_k\right) \quad , \quad \text{for some } \xi \in [x(t_k), x_k]$$

we have

$$\begin{aligned} L[x] &= a_k \left(x(t_k) - x_k\right) - h b_k \frac{\partial f}{\partial x}(\xi)\left(\left(x(t_k) - x_k\right)\right) \\ &= \left[a_k - h b_k F\right]\left(x(t_k) - x_k\right) \end{aligned}$$

We thus have

$$x(t_k) - x_k = \frac{L[x]}{a_k - h b_k F} = \frac{d_{m+1} h^{m+1} x^{(m+1)}(t_{n-k}) + \cdots}{a_k - h b_k F} \approx \frac{d_{m+1} h^{m+1} x^{(m+1)}(t_{n-k})}{a_k} + \mathcal{O}\left(h^{m+2}\right)$$

3. Global Truncation Error

We shall now establish a bound on total error induced by the local truncation errors that occur at each stage of a multi-step numerical method. As before we let x_n represent the value of the numerical solution after n iterations of the multi-step method and let $x(t_n)$ denote the corresponding value of the exact solution (for the same time t_n). Now the first thing one should realize is that the global truncation error is **not** simply the sum of the local truncation errors that occur at each stage of the iterative method. For the accuracy of a successive stage depends crucially on the accuracy of the stage that preceded it.

To get a handle on how the error of a preceding stage effects the error of a latter stage we consider a family of initial value problems:

$$\frac{dx}{dt} = f(t, x) \tag{24.12}$$

$$x(0) = s \quad , \quad s \in \mathbb{R} \tag{24.13}$$

We shall assume that $f_x \equiv \frac{\partial f}{\partial x}$ is continuous and satisfies

$$f_x(t,x) \le \lambda \quad , \quad \text{for all } t \in [0,T] \text{ and all } x \in \mathbb{R}$$

The exact solution of (24.12) and (24.13) is, of course, a function of t, but in order to make its dependence on the initial condtion (24.13) explicit, we shall denote it by $x(t,s)$. Thus, we write

$$\begin{aligned} \frac{\partial}{\partial t} x\,(t,s) &= f\,(t, x(t,s)) \\ x(0,s) &= s \end{aligned}$$

If we differentiate these equations with respect to s we obtain

$$\begin{aligned} \frac{\partial}{\partial t}\frac{\partial x}{\partial s} &= \frac{\partial}{\partial s}\frac{\partial x}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} \\ \frac{\partial x}{\partial s}(0,s) &= 1 \end{aligned}$$

Writing $u = \frac{\partial x}{\partial x}$, and noting that the second equation tells us that $u(0,s)$ does not depend on s, we have

$$u' = f_x u \tag{24.14}$$

$$u(0) = 1 \tag{24.15}$$

THEOREM 24.8. *If $f_x < \lambda$ then the solution of (24.14) and (24.15) satisfies the inequality*

$$|u(t)| \le e^{\lambda t} \quad , \quad t \ge 0$$

Proof. Write

$$\alpha(t) = \lambda - f_x(t)$$

Since $f_x(t) < \lambda$, $\alpha(t)$ is a positive function. And so we have

$$\frac{u'}{u} = f_x = \lambda - \alpha(t)$$

or

$$\frac{d}{dt}(\ln|u|) \le \lambda - \alpha(t)$$

Integrating both sides between 0 and t yields

$$\ln|u(t)| - \ln|u(0)| = \int_0^t \frac{d}{dt}(\ln|u|)\,dt = \int_0^t (\lambda - \alpha(t))\,dt = \lambda t - \int_0^t \alpha(t)dt$$

or

$$\ln|u(t)| < \lambda t \tag{24.16}$$

where on the left hand side we have used the initial condition $u(0) = 1$ and on the right hand side we have used the fact that

$$\lambda t - \int_0^t \alpha(t)dt < \lambda t$$

since $\alpha(t)$ is a positive function. Noting that the exponential function is monotonically increasing, we can maintain the inequality if we exponentiate both sides of (24.16). We thus have

$$u(t) < e^{\lambda t}$$

THEOREM 24.9. *If the inital value problem corresponding to (24.12) and (24.13) is solved with initial values s and $s + \delta$, then the solution curves at time t differ by at most $|\delta| e^{\lambda t}$.*

Proof. By the Mean Value Theorem we have

$$\begin{aligned}
|x(t,s) - x(t,s+\delta)| &= \left|\frac{\partial x}{\partial s}x(t+\theta\delta)\right| |\delta| \quad , \quad \text{for some } \theta \in [0,1] \\
&\equiv u(t+\theta\delta)|\delta| \\
&< |\delta|\, e^{\lambda t}
\end{aligned}$$

THEOREM 24.10. *If the local truncation errors at $t_1, t_2, \ldots, t_n = t_0 + nh$ do not exceed δ in magnitude, then the global truncation error ε does not exceed*

$$\frac{e^{\lambda t_n} - 1}{e^{\lambda h} - 1}\delta$$

Proof.

At the first iteration the total error is just the local truncation error so

$$\varepsilon_1 = \delta$$

At the second iteration the total error is the sum of the local truncation error and the error arising from the fact that initial condition might be off by ε_1. Thus

$$\varepsilon_2 = \delta + e^{\lambda h}\delta$$

Similarly at level 3 we would have

$$\varepsilon_3 = \delta + e^{\lambda h}\varepsilon_2 = \delta + \left(\delta + \delta e^{\lambda h}\right)e^{\lambda h} = \delta + \delta e^{\lambda h} + \delta e^{2\lambda h}$$

and so on and so on until one has

$$\begin{aligned}
\varepsilon_n &= \delta + \delta e^{\lambda h} + \delta e^{2\lambda h} + \cdots + \delta e^{(n-1)\lambda h} \\
&= \delta \sum_{i=0}^{n-1} e^{i\lambda h} \\
&= \delta \frac{e^{n\lambda h} - 1}{e^{\lambda h} - 1} \\
&= \delta \frac{e^{\lambda t_n} - 1}{e^{\lambda h} - 1}
\end{aligned}$$

where in passing from the second to the third equation we have used the fact that

$$x^n - 1 = (x-1)(x^{n-1} + x^{n-2} + \cdots + x + 1)$$

THEOREM 24.11. *If the local truncation error in a numerical solution of an initial value problem is of order $\mathcal{O}(h^m)$ then the global truncation error is of order $\mathcal{O}(h^{m-1})$.*

Proof. According to the preceding theorem,

$$\varepsilon_n = \delta \frac{e^{\lambda t_n} - 1}{e^{\lambda h} - 1}$$

Now for, assuming we keep the final time t_n fixed, we have for $\lambda h << 1$

$$\begin{aligned}
\frac{e^{\lambda t_n} - 1}{e^{\lambda h} - 1} &= \frac{e^{\lambda t_n} - 1}{\left(1 + \lambda h + \frac{1}{2}(\lambda h)^2 + \cdots\right) - 1} \\
&= \frac{e^{\lambda t_n} - 1}{\lambda h + \frac{1}{2}(\lambda h)^2 + \cdots} \\
&= \frac{1}{\lambda h}\left(\frac{e^{\lambda t_n} - 1}{1 + \frac{1}{2}(\lambda h) + \cdots}\right) \\
&\approx \mathcal{O}(h^{-1})
\end{aligned}$$

and so

$$\varepsilon_n \approx \mathcal{O}\left(h^m\right) \cdot \mathcal{O}\left(h^{-1}\right) = \mathcal{O}\left(h^{m-1}\right)$$