Linear Multistep Methods[1]

We consider the Initial Value Problem (IVP)

$$y' = f(x, y), \quad y(a) = y_0, \quad x \in [a, b], \quad y(x), \ f(x, y) \in \mathbb{R}^m \tag{1}$$

where we assume that there exists some *Lipschitz Constant L* such that

$$\|f(x, y) - f(x, z)\| \leq L \|y - z\|, \quad \text{for all} \quad (x, y), (x, z) \in [a, b] \times \mathbb{R}^m \tag{2}$$

Assumption (2) implies that the initial value problem (1) possesses a unique solution.

We use the notation

$$x_n = a + nh, \quad y_n \approx y(x_n), \quad f_n = f(x_n, y_n), \quad n = 0, 1, 2, \cdots \tag{3}$$

Our fundamental problem for the first few weeks of this semester is how to compute the y_n. They may be defined by Linear Multistep Methods (LMMs):

$$\sum_{j=0}^{k} \alpha_j y_{n+j} = h \sum_{j=0}^{k} \beta_j f_{n+j} \tag{4}$$

Here, k is the *step-number*, h is the *step-size* (assumed constant for the moment), $\alpha_k = 1$, $|\alpha_0| + |\beta_0| \neq 0$, the method is *explicit* if $\beta_k = 0$, otherwise it is *implicit*.

We are concerned with the *Global Truncation Error*

$$e_n = y(x_n) - y_n. \tag{5}$$

The LMM (4) is said to be *convergent* if, for all IVPs (1) satisfying (2), all $x \in [a, b]$, and all starting strategies $y_\mu = \eta_\mu(h)$, $\mu = 0, 1, \cdots, k - 1$ satisfying

$$\lim_{h \longrightarrow 0} \eta_\mu(h) = y_0 \tag{6}$$

the following holds:

$$\lim_{h \longrightarrow 0, nh = x - a} y_n = y(x) \tag{7}$$

(Note that the requirement on the starting values is rather weak, it is satisfied e.g. by the "strategy" $y_\mu = y_0$.)

An obvious minimum requirement for a LMM is that it be convergent.

Two simple Examples

To get an idea of what's involved in convergence, let's look at two extremely simple examples. This will illustrate a

Major Technique. *Consider a test problem, compute the general solutions of the numerical method and the analytical problem, and compare them.*

What's the simplest IVP imaginable? How about

$$y' = 0, \quad y(0) = y_0. \tag{8}$$

---

[1] by Peter Alfeld. TEX processed July 9, 2014

Clearly, the solution is
$$y(x) = y_0. \tag{9}$$

Now let's compute the *general* solution of the difference method

$$\sum_{j=0}^{k} \alpha_j y_{n+j} = 0. \tag{10}$$

This is a *homogeneous linear difference equation* whose solutions define a k dimensional linear space of sequences y_n, $n = 0, 1, 2, 3, \cdots$.

To understand the space of solutions better we associate with the LMM (4) its *first* and *second characteristic polynomial* ρ and σ, respectively by

$$\rho(\xi) = \sum_{j=0}^{k} \alpha_j \xi^j, \quad \text{and} \quad \sigma(\xi) = \sum_{j=0}^{k} \beta_j \xi^j. \tag{11}$$

It is obvious that if
$$\rho(r) = 0 \tag{12}$$

then
$$y_n = r^n \tag{13}$$

is a solution of the difference equation (10). Any linear combination of solutions of the form (13) also solves (10). If there are k distinct roots of ρ then all solutions are linear combinations of solutions of the form (13). If there is a multiple root r satisfying

$$\rho(r) = \rho'(r) = \cdots = \rho^{(q)}(r) = 0 \tag{14}$$

then $q + 1$ corresponding linearly independent solutions of (10) are given by[2]

$$y_n = n^j r^n \quad \text{where} \quad j = 0, 1, \cdots, q. \tag{15}$$

In what follows let's suppose for simplicity that all roots r_1, r_2, $\cdots$, r_k of ρ are distinct. The general solution of the difference equation is then given by

$$y_n = \sum_{\mu=1}^{k} \gamma_\mu r_\mu^n. \tag{16}$$

The coefficients γ_j are defined by the starting values.

How can this solution converge to the true solution of the IVP (8)? Clearly we must have:

— 1 must be a root of ρ.

— No root of ρ can exceed 1 in absolute value.

— In view of what was said above about multiple roots, any root of ρ of absolute value equal to 1 must be simple.

Let's note these properties formally

$$\begin{array}{lll} \text{a.} & \rho(1) = 0 & \\ \text{b.} & \rho(r) = 0 & \implies \quad |r| \leq 1 \\ \text{c.} & \rho(r) = 0 \quad \text{and} \quad |r| = 1 & \implies \quad \rho'(r) \neq 0 \end{array} \tag{17}$$

---

[2] Verifying this is a good exercise.

Definition. *A method that satisfies properties b. and c. is said to be zero-stable.*

To gain more insight, let's look at a slightly more complicated IVP, say

$$y' = 1, \quad y(0) = 0 \tag{18}$$

which obviously has the solution

$$y(x) = x. \tag{19}$$

The difference equation now becomes:

$$\sum_{j=0}^{k} \alpha_j y_{n+j} = h \sum_{j=0}^{k} \beta_j \tag{20}$$

This is an *inhomogeneous* linear difference equation and its *general* solution can be written as any *particular* solution plus the general solution of the homogeneous equation (8). To find a particular solution we have to do some inspired guessing. Considering the analytical solution suggests to try a solution of the form

$$y_n = \gamma x_n = \gamma nh \tag{21}$$

where γ is as yet undetermined. Plugging this into the difference equation and using property a. above yields

$$\gamma \sum_{j=0}^{k} \alpha_j (n+j)h = \gamma \left(nh \sum_{j=0}^{k} \alpha_j + h \sum_{j=0}^{k} \alpha_j j \right)$$

$$= \gamma h \sum_{j=0}^{k} \alpha_j j \tag{22}$$

$$= h \sum_{j=0}^{k} \beta_j$$

Clearly, the last equation requires

$$\gamma = \frac{\displaystyle\sum_{j=0}^{k} \beta_j}{\displaystyle\sum_{j=0}^{k} \alpha_j j} = \frac{\sigma(1)}{\rho'(1)}. \tag{23}$$

It is now almost obvious[3] that for convergence we have to have that

$$\gamma = 1. \tag{24}$$

Let's summarize: For the LMM to converge just for the simple DEs $y' = 0$ and $y' = 1$ we have to have properties a.–c. above, and, in addition:

$$\text{d.} \quad \sigma(1) = \rho'(1). \tag{25}$$

Definition. *A method that satisfies a. and d. is said to be consistent.*

Why did we do all this? Most amazingly, it turns out that these properties are also sufficient for convergence in general! Thus

> **the LMM will converge for all IVPs (1) if**
>
> **it converges just for $y' = 0$ and $y' = 1$!**

Marvel at that! A proof of this fact is given in Henrici's book but is beyond the scope of this course.

Thus it turns out that

> **LMM convergent $\Longleftrightarrow$ it is consistent and zero-stable**

---

[3] Work out the details! Also, what happens if $\rho'(1) = 0$?

Local Truncation Error

Associated with the LMM is its *Local Truncation Error* LTE[4] defined by

$$\text{LTE} = \sum_{j=0}^{k} \alpha_j y(x_{n+j}) - h \sum_{j=0}^{k} \beta_j y'(x_{n+j}) \tag{28}$$

where $y(x)$ is the true solution of the IVP (1). Assuming that the solution can be expanded into a power series about x_n[5] it turns out that

$$\text{LTE} = \sum_{i=0}^{\infty} C_i h^i y^{(i)}(x_n) \tag{29}$$

where

$$C_0 = \sum_{j=0}^{k} \alpha_j \quad \left(= \rho(1) \right)$$

$$C_1 = \sum_{j=1}^{k} j\alpha_j - \sum_{j=0}^{k} \beta_j \quad \left(= \rho'(1) - \sigma(1) \right) \tag{30}$$

$$C_q = \frac{1}{q!} \sum_{j=1}^{k} j^q \alpha_j - \frac{1}{(q-1)!} \sum_{j=1}^{k} j^{q-1} \beta_j, \quad q = 2, 3, \cdots$$

The LMM is *of order p* if

$$C_0 = C_1 = \cdots = C_p = 0, \quad C_{p+1} \neq 0. \tag{31}$$

The number C_{p+1} is the *error constant of the LMM*.

Note. *According to our earlier definition, a LMM is consistent if it is of order at least 1.*

Note. *We can now reinterpret the two requirements for convergence: Stability means that errors do not get unduly amplified, and consistency means that the error introduced at each step is not too large.*

The following table gives the maximum order of convergent LMMs

	explicit	implicit
k even	k	$k + 2$
k odd	k	$k + 1$

---

[4] There is also an "error under the localizing assumption"

$$y_{n+j} = y(x_{n+j}), \quad j = 0, 1, \cdots, k - 1. \tag{26}$$

It follows by a simple application of the Mean Value Theorem that under that assumption

$$y(x_{n+k}) - y_{n+k} = LTE + O(h^{p+1}). \tag{27}$$

Thus, at least asymptotically, the local truncation error, and the error under the localizing assumption, are equivalent.

[5] as seen in class, the order and error constant are independent of the point about which the LTE is expanded.

Examples

Of particular importance among the following examples are the Adams Methods and the Backward Differentiation formulas. These are the ones used in the many versions of Gear's package.

Here are some examples for *explicit convergent* LMMs:

1. **Euler's Method.** $p = k = 1, \quad C_2 = \frac{1}{2}$:

$$y_{n+1} - y_n = h f_n \tag{32}$$

2. **The Midpoint Rule.** $p = k = 2, \quad C_3 = \frac{1}{3}$:

$$y_{n+2} - y_n = 2h f_{n+1} \tag{33}$$

3. **Adams-Bashforth Methods.** $p = k, \quad k = 1, 2, 3, \cdots$:

$$y_{n+k} - y_{n+k-1} = h \sum_{j=0}^{k-1} \beta_j^* f_{n+j} \tag{34}$$

(Note that the requirement $p = k$ uniquely defines the β_j^*.)

Here are some examples for *implicit convergent* LMMs:

4. **Backward (or Implicit) Euler Method.** $p = k = 1, \quad C_2 = -\frac{1}{2}$:

$$y_{n+1} - y_n = h f_{n+1} \tag{35}$$

5. **The Trapezoidal Rule.** $p = 2, \quad k = 1, \quad C_3 = -\frac{1}{12}$:

$$y_{n+1} - y_n = \frac{h}{2} \left(f_{n+1} + f_n \right) \tag{36}$$

6. **Simpson's Rule.** $p = 4, \quad k = 2, \quad C_5 = -\frac{1}{90}$:

$$y_{n+2} - y_n = \frac{h}{3} \left(f_{n+2} + 4 f_{n+1} + f_n \right) \tag{37}$$

In spite of being convergent, the performance of Simpson's Rule is mediocre (due to the spurious root of the first characteristic polynomial being -1).

7. **Adams-Moulton Methods.** $p = k + 1, \quad k = 1, 2, 3, \cdots$:

$$y_{n+k} - y_{n+k-1} = h \sum_{j=0}^{k} \beta_j f_{n+j} \tag{38}$$

(Note that the requirement $p = k + 1$ uniquely defines the β_j.)

8. **Backward Differentiation Methods.** $p = k, \quad k = 1, 2, 3, 4, 5, 6$:

$$\sum_{j=0}^{k} \alpha_j y_{n+j} = h \beta_k f_{n+k} \tag{39}$$

These methods are not zero-stable if $k > 6$. Note that the requirement $p = k$ uniquely defines the α_j.

Non-Convergent Maximal Linear Multistep Methods

The coefficients of an LMM can be chosen so as to maximize the order. Matching the number of parameters with the number of conditions suggests that an explicit method can have order $2k - 1$, and implicit one order $2k$. It turns out that the linear equations are always consistent, and that you don't get a higher order for free. Thus these values are indeed attained. LMMs of maximum possible order are called *maximal*. Note that they may not be convergent. Examples of maximal methods listed above include: Euler's Method, the Trapezoidal Rule, and Simpson's Rule. Following are some additional *divergent* methods:

<center>Explicit LMMs</center>

9. $k = 2$, $p = 3$, $C_4 = \frac{1}{6}$:

$$y_{n+2} + 4y_{n+1} - 5y_n = h\left(4f_{n+1} + 2f_n\right) \tag{40}$$

10. $k = 3$, $p = 5$, $C_6 = \frac{1}{2}$.

$$y_{n+3} + 18y_{n+2} - 9y_{n+1} - 10y_n = h\left(9f_{n+2} + 18f_{n+1} + 3f_n\right) \tag{41}$$

<center>Implicit LMMs</center>

11. $k = 3$, $p = 6$, $C_7 = \frac{-3}{140}$.

$$11y_{n+3} + 27y_{n+2} - 27y_{n+1} - 11y_n = h\left(3f_{n+3} + 27f_{n+2} + 27f_{n+1} + 3f_n\right) \tag{42}$$

Absolute Stability

We now know how to find convergent methods, but these are not necessarily *good* methods (because they may be inefficient). To get better insight we need another *test equation*. However, that equation is not as obvious as the previous ones ($y' = 0$ and $y' = 1$). It can be motivated as follows:

$$y' = f(x, y)$$
<center>**Linearize**</center>
$$\approx y'(x) + f_y\left(x, y(x)\right)\left(y - y(x)\right)$$
$$= A(x)y + g(x)$$
<center>**Freeze**</center>
$$\approx Ay + g \tag{43}$$
<center>**Homogenize**</center>
$$\longrightarrow y' = Ay$$
<center>**Diagonalize**</center>
$$\longrightarrow y' = \lambda y \quad \text{where} \quad \lambda \in \mathbf{C}$$

The individual steps can be justified as follows:

Linearize This is simply a Taylor expansion of first order about the exact solution of the differential equation. It is reasonable as long as we are close to the exact solution.

Freeze We move along the x-axis in small steps, so this assumption is locally valid. Globally it is of course totally unrealistic.

Homogenize The inhomogeneous term g only adds a constant term to the solution, and can be ignored (since we already dealt with constant components of the solution by considering the test equation $y' = 0$).

<center>6</center>

Diagonalize Suppose there is a similarity transform

$$D = C^{-1}AC \tag{44}$$

where C and D are complex, and D is diagonal. This is possible whenever A is non-defective, i.e., it possesses a full set of linearly independent eigenvectors. In that case, the columns of C are the *eigenvectors* of A, and the diagonal entries of D are the corresponding eigenvalues.

We are thus led to the test equation

$$y' = \lambda y \tag{45}$$

where λ is a *complex* number that plays the role of one of the eigenvalues of the Jacobian of f. Note that this closely models the situation in solving the problem $y' = Ay$ analytically[6].

Plugging the differential equation (45) into the LMM (4) yields the difference equation

$$\sum_{j=0}^{k} (\alpha_j - \kappa\beta_j)\, y_{n+j} = 0 \tag{46}$$

where

$$\kappa = h\lambda \in \mathbf{C}. \tag{47}$$

This is once more a *homogeneous linear difference equation* whose solutions are governed by the roots of the stability polynomial

$$\pi(\kappa, \xi) = \sum_{j=0}^{k} (\alpha_j - \kappa\beta_j)\, \xi^j = \rho(\xi) - \kappa\sigma(\xi). \tag{48}$$

The roots of a polynomial depend continuously on the coefficients of the polynomial. The roots of the stability polynomial therefore depend continuously on κ.

Definition. *The root $r = r(\kappa)$ of $\pi(\kappa, \xi)$ satisfying $r(0) = 1$ is said to be the essential root of $\pi(\kappa, \xi)$. The other roots are called the spurious roots of $\pi(\kappa, \xi)$.*

Definition. *We say that the LMM (4) is absolutely stable[7]. for a given $\kappa = h\lambda$ if all solutions of the test equation (45) with stepsize h converge to zero as n tends to infinity.*

Equivalently, the LMM is absolutely stable for a given κ if all roots of the stability polynomial π are less than 1 in absolute value. The set of all κ for which the LMM is absolutely stable is *the region of absolute stability* of the LMM. This is a subset of the complex plane.

Note. *It is important to realize that the LMM being absolutely stable does not imply a value judgment. The analytic solutions of the test equation (45) will tend to zero as x tends to infinity iff the real part of λ is negative, and it will grow unboundedly if the real part of λ is positive. Thus the ideal region of absolute stability of an LMM is the open left half plane, no larger or smaller.*

The boundary of the region of absolute stability is a subset of the set

$$\mathrm{BL}(\rho, \sigma) = \left\{ z \in \mathbf{C} \mid \quad z = \frac{\rho(e^{i\theta})}{\sigma(e^{i\theta})}, \quad i = \sqrt{-1}, \quad \theta \in [0, 2\pi] \right\}. \tag{49}$$

(The notation BL means "boundary locus").

---

[6] As an exercise you may wish to work out the connection.

[7] There is also a much less important concept of *relative stability* which requires that the spurious roots grow no faster than the essential root

Example. *For* **Euler's Method** *we obtain*

$$z = \frac{e^{i\theta} - 1}{1} = \cos\theta + i\sin\theta - 1 \tag{50}$$

which is the unit circle shifted to the left by 1. For $\kappa = 0$ the only root of the stability polynomial is $\xi = 1$. Euler's method is thus absolutely stable for κ inside the circle. Clearly, it is not absolutely stable outside of the circle.

Example. *For the* **Trapezoidal Rule** *we obtain*

$$\rho(\xi) = \xi - 1, \quad and \quad \sigma(\xi) = \frac{1}{2}(\xi + 1). \tag{51}$$

Thus

$$z = \frac{2\left(e^{i\theta} - 1\right)}{e^{i\theta} + 1} = \frac{2\left(e^{i\theta} - 1\right)\left(e^{-i\theta} + 1\right)}{\left(e^{i\theta} + 1\right)\left(e^{-i\theta} + 1\right)} = \frac{2\left(1 + e^{i\theta} - e^{-i\theta} - 1\right)}{\left|e^{i\theta} + 1\right|^2} = \frac{4i\sin\theta}{\left|e^{i\theta} + 1\right|^2}. \tag{52}$$

Thus the boundary locus is the imaginary axis. A simple check shows that the region of absolute stability is the open half plane, and that the method is absolutely unstable for positive real parts of κ. Thus the Trapezoidal Rule has precisely the "right" region of absolute stability.

We need a more powerful tool to verify that all roots of a given polynomial are in the unit circle. A polynomial Φ with complex coefficients is said to be a *Schur Polynomial* if all of its roots have absolute value less than 1. The following criterion[8] may be used to identify Schur polynomials:

(Schur Criterion.) Let

$$\Phi(r) = c_k r^k + c_{k-1} r^{k-1} + \cdots + c_1 r + c_0, \tag{53}$$

where $c_0 c_k \neq 0$, and the c_i denote complex coefficients. Define

$$\hat{\Phi}(r) = \bar{c}_0 r^k + \bar{c}_1 r^{k-1} + \cdots + \bar{c}_{k-1} r + \bar{c}_k \tag{54}$$

(where $\bar{c}$ denotes the conjugate complex of c) and

$$\Phi_1(r) = \frac{1}{r}\left(\hat{\Phi}(0)\Phi(r) - \Phi(0)\hat{\Phi}(r)\right). \tag{55}$$

Clearly, Φ_1 has degree at most $k - 1$. Then, Φ is a Schur Polynomial if and only if

$$|\hat{\Phi}(0)| > |\Phi(0)| \tag{56}$$

and Φ_1 is a Schur Polynomial.

The set $\mathrm{BL}(\rho, \sigma)$ divides the complex plane into a finite number of subsets. The Schur Criterion can be applied to one point in each of these subsets, and the region of absolute stability can be identified in this manner.

Next is a (casual) argument that shows that for a convergent LMM the essential root of ρ behaves as it should for small real values of $h\lambda$. (How should it behave? It should be greater than 1 for positive κ, and smaller than 1 for negative κ.)

---

[8] An excellent reference on this and related topics is: J.J.H. Miller, *On the location of zeros of certain classes of polynomials with applications to numerical analysis*, J. Inst. Math. Appl., **8**, 397–406 (1971). The journal (JIMA) is available in the Marriott Library.

The roots of the stability polynomial are defined by

$$\pi\big(\kappa, r(\kappa)\big) = \rho\big(r(\kappa)\big) - \kappa\sigma\big(r(\kappa)\big) = 0. \tag{57}$$

Differentiating implicitly with respect to κ yields

$$\rho'\big(r(\kappa)\big)r'(\kappa) - \sigma\big(r(\kappa)\big) - \kappa\sigma'\big(r(\kappa)\big)r'(\kappa) = 0. \tag{58}$$

Solving for $r'(\kappa)$ yields the differential equation

$$r'(\kappa) = \frac{\sigma\big(r(\kappa)\big)}{\rho'\big(r(\kappa)\big) - \kappa\sigma'\big(r(\kappa)\big)}. \tag{59}$$

The essential root r_1 satisfies $r_1(0) = 1$. Thus we obtain

$$r_1(0) = 1 \quad \text{and} \quad r_1'(0) = \frac{\sigma(1)}{\rho'(1)} = 1 \tag{60}$$

Increasing κ somewhat will increase r_1 (to beyond 1) and decreasing κ will decrease r_1 (below 1).

This kind of argument can be extended to the complex plane.

The following consequences are immediate:

1. A convergent LMM cannot be absolutely stable for small positive values of $\Re(\kappa)$.

2. A convergent LMM for which all spurious roots of the first characteristic polynomial are strictly within the unit circle will have a non-empty region of absolute stability to the left of the origin.

Note. *Item 2. is the reason for the success of the Adams methods. Here the spurious roots are all equal to zero. Hence it will take a while for them to move out of the unit circle as κ changes. As a consequence the Adams methods have fairly large regions of absolute stability.*

Following is another application of the above kind of reasoning:

Example. *Consider the spurious root of Simpson's Rule. It satisfies*

$$r_2(0) = -1 \quad \text{and} \quad r_2'(0) = \frac{\sigma(-1)}{\rho'(-1)} = \frac{1/3\left(\xi^2 + 4\xi + 1\right)}{2\xi}\bigg|_{\xi=-1} = +\frac{1}{3}. \tag{61}$$

Thus the spurious root will become more negative and thus larger than 1 in absolute value as κ becomes negative. Simpson's Rule is therefore absolutely unstable for small negative (as well as for small positive) values of κ. In fact, the region of absolute stability of Simpson's Rule is empty (exercise!).

We next ask for methods with infinite regions of absolute stability, including the entire left half plane if possible.

However, we are stymied immediately by the following

Observation. *The region of absolute stability of an explicit method is bounded.*

To see this note that as κ tends to infinity (in some direction in the complex plane) the leading coefficient of the stability polynomial will converge to zero relative to some of the others. So in a sense the stability polynomial approaches a polynomial of degree less than k. Thus at least one root somehow has to "vanish". The only way this can happen is that one of the roots of the stability polynomial tends to infinity as κ tends to infinity. (Working out the details will be a home-work problem.)

We are now at a crossroads. We could either

— Insist on using explicit methods because of their simplicity, or

— Explore the use of implicit methods.

Both branches have their merits, and we will explore them both, beginning with the first.