LECTURE 3
FIRST ORDER DIFFERENTIAL EQUATIONS

Solution of ordinary differential equations in which time t is the independent variable is usually an important part of solving partial differential equations in atmospheric and oceanic models. In this lecture, we examine finite-difference schemes for first-order ordinary differential equations with t as the independent variable and $U(t)$ as the dependent variable. We consider the equation

$$\frac{dU}{dt} = f(U,t). \tag{1}$$

We divide the time axis into segments of equal interval Δt, and denote by $U^{(n)}$ the approximate solution for $U(n\Delta t)$. We further assume that we know the values $U^{(n)}$, $U^{(n-1)}$, $U^{(n-2)}$, . . ., and wish to construct a scheme for finding $U^{(n+1)}$.

a. Examples of finite-difference scheme

Let $f^{(n)} \equiv f(U^{(n)},\ n\Delta t)$. The following are example of finite-difference schemes for (1):

The forward (Euler) scheme, $$\frac{U^{(n+1)} - U^{(n)}}{\Delta t} = f^{(n)}\ ; \tag{2}$$

The backward (implicit) scheme $$\frac{U^{(n+1)} - U^{(n)}}{\Delta t} = f^{(n+1)}\ ; \tag{3}$$

The centered (leap-frog) scheme $$\frac{U^{(n+1)} - U^{(n-1)}}{2\Delta t} = f^{(n)}\ . \tag{4}$$

The scheme for (3) is called implicit, since the unknown $f^{(n+1)}$ appears on the right hand side, while the schemes for (2) and (4) are explicit. The schemes for (2) and (3) are two-level schemes since two time levels n and $n+1$ are involved. The scheme for (4), on the other hand, is a three-level scheme since three time levels $n-1$, n and $n+1$ are involved.

When more than two time levels are involved, the scheme is not self-starting. In the case of the leapfrog scheme, for example, two initial conditions $U^{(0)}$ and $U^{(1)}$ must be given to determine $U^{(2)}$, and subsequently $U^{(3)}$, $U^{(4)}$, ... The additional initial condition $U^{(1)}$ is called the computational initial condition. The usual procedure for obtaining the computational initial condition is using one of the two-level schemes, which are self-starting.

The trapezoidal scheme,
$$\frac{U^{(n+1)}-U^{(n)}}{\Delta t}=\frac{1}{2}\left(f^{(n+1)}+f^{(n)}\right). \tag{5}$$

The Milne corrector,
$$\frac{U^{(n+1)}-U^{(n-1)}}{2\Delta t}=\frac{1}{6}\left(f^{(n+1)}+4f^{(n)}+f^{(n-1)}\right) \tag{6}$$

Note that (6) is a two-level implicit scheme and that (6) is a three-level implicit scheme.

b. Derivation of a family of schemes for (1)

To derive a broader family of schemes, let us integrate (1) with respect to time from $(n-m)\Delta t$ to $(n+1)\Delta t$. Here m is zero or a positive integer. Then

$$U^{(n+1)}-U^{(n-m)}=\int_{(n-m)\Delta t}^{(n+1)\Delta t} f(U,t)dt \quad \cdot \tag{7}$$

This expression, divided by $(1+m)\Delta t$, may be approximated by

$$\frac{1}{(1+m)\Delta t}\left(U^{(n+1)}-U^{(n-m)}\right)=\beta f^{(n+1)}+\alpha_0 f^{(n)}+\alpha_{-1}f^{(n-1)}+\alpha_{-2}f^{(n-2)}+\ldots+\alpha_{-\ell}f^{(n-\ell)}+\varepsilon. \tag{8}$$

Here, ℓ is either zero or a positive integer. When $\beta\neq 0$, the scheme is implicit while when $\beta=0$ the scheme is explicit.

Substituting the true solution $U(t)$ and corresponding $f(U,t)$ into (8), expanding into a Taylor series about $t=n\Delta t$, and denoting d/dt by $(\ )'$, we obtain

$$\frac{1}{(1+m)\Delta t}\left[\left\{U+\Delta tU'+\frac{\Delta t^2}{2!}U''+\frac{\Delta t^3}{3!}U'''+\frac{\Delta t^4}{4!}U''''+\ldots\right\}\right.$$
$$\left.-\left\{U-m\Delta tU'+\frac{(m\Delta t)^2}{2!}U''-\frac{(m\Delta t)^3}{3!}U'''+\frac{(m\Delta t)^4}{4!}U''''+-\ldots\right\}\right]$$
$$=\beta\left\{f+\Delta tf'+\frac{\Delta t^2}{2!}f''+\frac{\Delta t^3}{3!}f'''+\ldots\right\}$$
$$+\alpha_0 f$$
$$+\alpha_{-1}\left\{f-\Delta tf'+\frac{\Delta t^2}{2!}f''-\frac{\Delta t^3}{3!}f'''+-\ldots\right\}$$
$$+\alpha_{-2}\left\{f-2\Delta tf'+\frac{(2\Delta t)^2}{2!}f''-\frac{(2\Delta t)^3}{3!}f'''+-\ldots\right\}$$
$$+\alpha_{-3}\left\{f-3\Delta tf'+\frac{(3\Delta t)^2}{2!}f''-\frac{(3\Delta t)^3}{3!}f'''+-\ldots\right\}$$

$$+\alpha_{-\ell}\left\{f-\ell t\Delta f'+\frac{(\ell\Delta t)^2}{2!}f''-\frac{(\ell\Delta t)^3}{3!}f'''+-\ldots\right\}$$
$$+\varepsilon. \tag{9}$$

Since the finite-difference scheme for (9) only approximates the differential equation, the true solution $U(t)$ generally does not satisfy (9). Therefore ε in (9) is generally non-zero. The quantity ε is the truncation error of the finite-difference scheme given by (9). Rearranging and using $U'=f$ we obtain

$$\varepsilon=U'\{1-(\beta+\alpha_0+\alpha_{-1}+\ldots+\alpha_{-\ell})\}$$
$$+\Delta t\,U''\left\{\frac{1}{2}\frac{1-m^2}{1+m}-\beta+\alpha_{-1}+2\alpha_{-2}+3\alpha_{-3}+\ldots.+\ell\alpha_{-\ell}\right\}$$
$$+\frac{(\Delta t)^2}{2!}U'''\left\{\frac{1}{3}\frac{1-m^3}{1+m}-\beta+\alpha_{-1}+4\alpha_{-2}+9\alpha_{-3}+\ldots.+\ell^2\alpha_{-\ell}\right\}$$
$$+\frac{(\Delta t)^3}{3!}U''''\left\{\frac{1}{4}\frac{1-m^4}{1+m}-\beta+\alpha_{-1}+8\alpha_{-2}+27\alpha_{-3}+\ldots.+\ell^3\alpha_{-\ell}\right\} \tag{10}$$

A finite-difference scheme is consistent if $\varepsilon\to 0$ as $\Delta t\to 0$. For the scheme given by (9), the consistency condition can be obtained from (11) as

$$1=\beta+\alpha_0+\alpha_{-1}+\alpha_{-2}+\ldots.+\alpha_{-\ell}. \tag{11}$$

The order of accuracy of a consistent finite-difference scheme is the order or the infinitesimal $\varepsilon(\Delta t)$. The order of accuracy can be made higher than the first by an appropriate choice of the $\ell+1$ coefficients. Generally, order of accuracy can be made at least as high as $\ell+2$ for implicit schemes and $\ell+1$ for explicit schemes.

(i) $m=0$

The scheme (8) now takes the form

$$\frac{U^{(n+1)}-U^{(n)}}{\Delta t}$$
$$=\beta f^{(n+1)}+\alpha_0 f^{(n)}+\alpha_{-1}f^{(n-1)}+\alpha_{-2}f^{(n-2)}+\ldots+\alpha_{-\ell}f^{(n-\ell)}. \tag{12}$$

(i.1) $\ell=0$

The consistency condition (11) becomes $\beta+\alpha_0=1$. All other α's are zero. Then (12) becomes

$$\frac{U^{(n+1)}-U^{(n)}}{\Delta t}=\beta f^{(n+1)}+(1-\beta)f^{(n)}. \tag{13}$$

Since only two time levels are involved, this family of schemes are two-level schemes. The truncation error (10) becomes $\Delta t U''(1/2-\beta)+0\left[(\Delta t)^2\right]$.

When $\beta=0$, the scheme reduces to the forward or Euler scheme given by (2). It has first-order accuracy as expected from the general rule ($\ell+1$ for explicit schemes). When $\beta=1$, the scheme reduces to the backward scheme given by (3), which again has first-order accuracy. The highest order of accuracy can be obtained by choosing $\beta=1/2$. Then the scheme reduces to the trapezoidal scheme given by (5), which has second-order accuracy as expected from the general rule ($\ell+2$ for implicit schemes).

(i.2) $\ell \geq 1$

This family of schemes use the information on f's at earlier time levels to obtain higher-order accuracy.

(i.2.a) Explicit schemes ($\beta=0$)

With $\ell=1$, the consistency condition (11) becomes $\alpha_0+\alpha_{-1}=1$. All other α's are zero. Then the scheme reduces to

$$\frac{U^{(n+1)}-U^{(n)}}{\Delta t}=\alpha_0 f^{(n)}+(1-\alpha_0)f^{(n-1)}. \tag{14}$$

The right hand side has the form of linear extrapolation when $\alpha_0>1$. The truncation error is $\Delta t U''(1/2+\alpha_{-1})+0\left[(\Delta t)^2\right]=\Delta t U''(3/2-\alpha_o)+0\left[(\Delta t)^2\right]$. Second-order accuracy is then obtained by choosing $\alpha_0=3/2$. This is the second-order Adams Bashforth scheme. Similarly, order of accuracy $\ell+1$ can be obtained with $\ell \geq 2$ if we choose α's such that

ℓ	α_0	α_{-1}	α_{-2}	α_{-3}	Truncation Error
2	$23/12$	$-4/3$	$5/12$		$0\left[(\Delta t)^3\right]$
3	$55/24$	$-59/24$	$37/24$	$-9/24$	$0\left[(\Delta t)^4\right]$

These are (higher-order) Adams Bashforth schemes.

(i.2.b) Implicit schemes ($\beta \neq 0$)

Then the order of accuracy $\ell+2$ can be obtained if we choose

ℓ	β	α_0	α_{-1}	α_{-2}	α_{-3}	Truncation Error
1	5/12	8/12	-1/12			$0[(\Delta t)^3]$
2	9/24	19/24	-5/24	1/24		$0[(\Delta t)^4]$
3	251/720	646/720	-246/720	-19/720	-19/720	$0[(\Delta t)^5]$

These are Adams–Moulton schemes.

(ii) $m = 1$

The scheme (8) now takes the form

$$\frac{U^{(n+1)} - U^{(n-1)}}{2\Delta t} = \beta f^{(n+1)} + \alpha_o f^{(n)} + \alpha_{-1} f^{(n-1)} + \alpha_{-2} f^{(n-2)} + \ldots + \alpha_{-\ell} f^{(n-\ell)}. \tag{15}$$

(ii.1) $\ell = 0$

The consistency condition is $\beta + \alpha_0 = 1$. The scheme reduces to

$$\frac{U^{(n+1)} - U^{(n-1)}}{2\Delta t} = \beta f^{(n+1)} + (1-\beta) f^{(n)}. \tag{16}$$

The truncation error (10) becomes $\Delta t U''(-\beta) + 0\left[(\Delta t)^2\right]$.

When $\beta = 0$, the scheme reduces to the leapfrog scheme given by (4). It has second-order accuracy, which is higher than expected from the general rule $\ell + 1$ for explicit schemes). Even when we allow the possibility of $\beta \neq 0$, second-order accuracy expected from the general rule ($\ell + 2$ for implicit schemes) is obtained with $\beta = 0$ (again the leapfrog scheme).

(ii.2) $\ell \geq 1$

As in the case of (i.2), we can increase order of accuracy by properly choosing α's.

First let $\beta = 0$. Even formally with $\ell = 1$, the highest order of accuracy (second order) is obtained with $\alpha_{-1} = 0$ (once again the leapfrog scheme). Higher-order schemes obtained with $\ell \geq 2$ are the Nyström schemes.

Now let $\beta \neq 0$. With $\ell = 1$, the consistency conditon is $\beta + \alpha_0 + \alpha_{-1} = 1$ and the truncation error is

$$\Delta t\, U''(-\beta+\alpha_{-1})+(\Delta t)^2/2!\, U'''(1/3-\beta-\alpha_{-1})+(\Delta t)^3/3!\, U''''(-\beta+\alpha_{-1})+0(\Delta t)^4.$$

From this we can see that $\beta=\alpha_{-1}=1/6$ and $\alpha_0=4/6$ give fourth-order accuracy. This is higher than what we expect from the general rule ($\ell+2$ for the implicit schemes). The scheme reduces to the Milne corrector* given by (7). With $\ell=2$, there is no gain in accuracy since the highest order of accuracy is obtained for $\alpha_{-2}=0$.

These schemes with highest-order accuracy for various combinations of m and • are shown in the Table 1.

Table 1. Schemes with Highest-Order Accuracy

		$\ell=0$	$\ell=1$	$\ell=2$	$\ell>2$
Explicit	$m=0$	Euler (1st)	Adams-Bashforth ($\ell+1$ th)		
	$m=1$	Leapfrog (2nd)		Nyström ($\ell+1$ th)	
Implicit	$m=0$	Trapezoidal (2nd)	Adams-Moulton ($\ell+2$ th)		
	$m=1$	Leapfrog (2nd)	Milne-Corrector (4th)		

c. Predictor-Corrector and Runge-Kutta Methods

The schemes with a non-zero β given in section b are implicit since they involve $f^{(n+1)}$, which depends on the unknown $U^{(n+1)}$. The predictor-corrector method uses $f^{(n+1)^*}\equiv f(U^{(n+1)^*},(n+1)\Delta t)$ in place of $f^{(n+1)}$. Here $U^{(n+1)^*}$ is an approximation to $U^{(n+1)}$.

Let us consider the trapezoidal scheme for (5), for example. A predictor-corrector scheme corresponding to (5) is

$$\frac{U^{(n+1)^*}-U^{(n)}}{\Delta t}=f^{(n)} \quad : \text{ predictor} \tag{17}$$

$$\frac{U^{(n+1)}-U^{(n)}}{\Delta t}=\frac{1}{2}\left(f^{(n+1)^*}+f^{(n)}\right) \quad : \text{ corrector,} \tag{18}$$

---

* The Milne predictor: An explicit scheme with m = 3 and • = 3 with $\alpha_0 = 2/3$ $\alpha_{-1} = -1/3$ and α_{-2} = 2/3 and $\alpha_{-3} = 0$.

where the Euler scheme is used as predictor. This scheme may be called the Euler-trapezoidal scheme, but is usually known as the improved Euler scheme (Collatz, 1960) or the Heun scheme (Heun, 1900). It can be shown that this scheme has second-order accuracy.

When the backward scheme is preferred to the trapezoidal scheme, corresponding predictor-corrector schemes are obtained by replacing the corrector (18) by

$$\frac{U^{(n+1)} - U^{(n)}}{\Delta t} = f^{(n+1)^*} \tag{19}$$

where $f^{(n+1/2)^*} \equiv f(U^{(n+1/2)^*}, (n+1/2)\Delta t)$. Note that $U^{(n+1/2)^*}$ is an approximation to U at $t = (n+1/2)\Delta t$ obtained by applying the Euler scheme over the time interval $\Delta t/2$. This scheme is called the Modified Euler scheme (Collatz, 1960) or the simplified Runge-Kutta scheme. It has second-order accuracy and is equivalent to the Heun scheme when f is a linear function of U only.

When (18) is used as the predictor, we have the Euler-backward scheme, usually known as the Matsuno scheme (Matsuno, 1966).

The Milne predictor--corrector scheme combines the Milne predictor and Milne corrector given earlier. As in the Milne corrector, the Milne predictor-corrector scheme has fourth-order accuracy.

The fourth-order Runge-Kutta scheme is given by

$$\begin{aligned}
k_1 &= \Delta t \quad f(U^{(n)}, n\Delta t), \\
k_2 &= \Delta t\, f(U^{(n)} + k_1/2, (n+1/2)\Delta t) \\
k_3 &= \Delta t\, f(U^{(n)} + k_2/2, (n+1/2)\Delta t) \\
k_4 &= \Delta t\, f\left[U^{(n)} + k_3, (n+1)\Delta t\right] \\
U^{(n+1)} &= U^{(n)} + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4).
\end{aligned} \tag{20}$$

This is a very good scheme although it is expensive when f is a complicated function.

Introducción a la Modelación Numérica de la Atmósfera.

Prof. Dr. Carlos Roberto Mechoso

(Mayo 2000)