Institutionen för datavetenskap
Umeå universitet

1 december 2010
Numeriska metoder för civilingenjörer

Solutions to review exercises for Themes 4 and 5

1 Theme 4

1. False. The condition for stability is that the solution curves for different initial conditions should not diverge as $t \to +\infty$.

2. A scalar, linear ODE written on the form $y' = \lambda y + f(t)$ is stable when $\mathrm{Re}(\lambda) \le 0$, asymptotically stable when $\mathrm{Re}(\lambda) < 0$, and unstable when $\mathrm{Re}(\lambda) > 0$. Thus, (a) and (b) are unstable, (c) is asymptotically stable, and (d) is stable.

3. (a) Introduce $z = y'$, and write as the system

$$\begin{pmatrix} y \\ z \end{pmatrix}' = \begin{pmatrix} z \\ \gamma \sin \omega t - \delta z - \sigma(y^3 - y) \end{pmatrix}$$

   (b) Introduce $g = f'$, $h = g' (= f'')$, and write

$$\begin{pmatrix} f \\ g \\ h \end{pmatrix}' = \begin{pmatrix} g \\ h \\ -\frac{1}{2} f h \end{pmatrix}$$

   (c) Introduce $z = y'$, and write

$$\begin{pmatrix} y \\ z \end{pmatrix}' = \begin{pmatrix} z \\ z(1 - y^2) + y \end{pmatrix}.$$

4. (a) An equation, typically nonlinear, has to be solved at each time step for an implicit method. No equation has to be solved for the explicit method.

   (b) The truncation error is the difference between the numerical and exact solution after one time step.

   (c) The method has the order of accuracy p if the truncation error is $O(\Delta t^{p+1})$.

5. True, by definition.

6. An implicit method requires generally much more floating-point operations at each time step, compared to an explicit method, due to the need to solve equations. An explicit method is thus more efficient as long as the time step required for stability is not too small (that is, much smaller than the time step required for sufficient accuracy).

7. (a) Forward Euler (*Euler framåt*), Backward Euler (*Euler bakåt*), and the trapezoidal method (*trapetsmetoden*).

   (b) Substitute $y_k = y(t_k)$, where y is the solution to $y' = f(t, y)$, and compute LHS−RHS of the scheme:

$$\begin{aligned}
y(t_{k+1}) - y(t_k) &- \Delta t \left[\alpha f(t_{k+1}, y(t_{k+1})) + (1-\alpha) f(t_k, y(t_k))\right] \\
&= y(t_{k+1}) - y(t_k) - \Delta t \left[\alpha y'(t_{k+1}) + (1-\alpha) y'(t_k)\right] = \text{[Taylor expansion]} \\
&= y(t_k) + y'(t_k)\Delta t + y''(t_k)\frac{\Delta t^2}{2} + y'''(t_k)\frac{\Delta t^3}{6} + O(\Delta t^4) - y(t_k) \\
&\quad - \Delta t \left[\alpha \left(y'(t_k) + y''(t_k)\Delta t + y'''(t_k)\frac{\Delta t^2}{2} + O(\Delta t^3)\right) + (1-\alpha) y'(t_k)\right] \\
&= y''(t_k)\left(\frac{1}{2} - \alpha\right)\Delta t^2 - y'''(t_k)\left(\frac{1}{6} - \frac{\alpha}{2}\right)\Delta t^3 + O(\Delta t^4).
\end{aligned}$$

   Thus, we have the order of accuracy 2 for $\alpha = 1/2$ and 1 otherwise.

(c) Applying the scheme on the model problem yields

$$y_{k+1} = y_k + \Delta t\left(\alpha\lambda y_{k+1} + (1-\alpha)\lambda y_k\right),$$

that is,

$$(1-\alpha\Delta t\lambda)y_{k+1} = [1+\Delta t(1-\alpha)\lambda]\, y_k.$$

Stability requires $|y_{k+1}| \le |y_k|$, which holds if and only if

$$\frac{\left|1+\Delta t(1-\alpha)\lambda\right|}{|1-\alpha\Delta t\lambda|} \le 1. \tag{1}$$

Since $\lambda < 0$, we write $\lambda = -|\lambda|$ and multiply both sides of (**??**) with $|1-\alpha\Delta t\lambda| = 1+\alpha\Delta t|\lambda| \ge 0$, to obtain

$$\left|1-\Delta t(1-\alpha)|\lambda|\right| \le 1+\alpha\Delta t|\lambda|,$$

that is,

$$-1-\alpha\Delta t|\lambda| \le 1-\Delta t(1-\alpha)|\lambda| \le 1+\alpha\Delta t|\lambda|.$$

The right inequality is always satisfied, whereas the left inequality yields that

$$\Delta t(1-2\alpha)|\lambda| \le 2,$$

which always is satisfied for $1/2 \le \alpha \le 1$. Thus, the scheme is unconditionally stable (*ovillkorligt stabil*) for $1/2 \le \alpha \le 1$. However, for $0 \le \alpha < 1/2$, we get the stability condition

$$\Delta t|\lambda| \le \frac{2}{1-2\alpha}.$$

(d) Choosing $\alpha = 1/2$ and substituting $f(t_{n+1}, y_{n+1})$ with $f(t_{n+1}, \kappa))$, where $\kappa = y_n + \Delta t f(t_n, y_n)$ (forward Euler extrapolation), we obtain Heun's method.
the forward Euler estimate

8. (a) The scheme is explicit.

(b) Substitute $y_k = y(t_k)$, where y is the solution to $y' = f(y,t)$, into the scheme and compute LHS−RHS:

$$\begin{aligned}
y(t_{k+1}) - y(t_k) - \frac{\Delta t}{2}\left[3f(t_k, y(t_k)) - f(t_{k-1}, y_{k-1})\right]& \\
= y(t_{k+1}) - y(t_k) - \Delta t\left[\frac{3}{2}y'(t_k) - \frac{1}{2}y'(t_{k-1})\right] &= \text{[Taylor expansion]} \\
= y(t_k) + y'(t_k)\Delta t + y''(t_k)\frac{\Delta t^2}{2} + y'''(t_k)\frac{\Delta t^3}{6} + O(\Delta t^4) - y(t_k)& \\
-\Delta t\left[\frac{3}{2}y'(t_k) - \frac{1}{2}\left(y'(t_k) - y''(t_k)\Delta t + y'''(t_k)\frac{\Delta t^2}{2} + O(\Delta t^3)\right)\right]& \\
= y'''(t_k)\frac{\Delta t^3}{6} + y'''(t_k)\frac{\Delta t^3}{4} + O(\Delta t^4) = \frac{5}{12}y'''(t_k)\Delta t^3 + O(\Delta t^4).&
\end{aligned}$$

The order of accuracy is thus 2.

9. A stiff system is one where there are vastly different time scales, such as for a system of ODEs $\mathbf{u}' = \mathbf{A}\mathbf{u}$ in which the real parts of the eigenvalues of matrix $\mathbf{A}$ are of vastly different size.

Time step restrictions for explicit methods are dictated by the fastest time scales (the largest real port of the eigenvalues of $\mathbf{A}$), which means that many time steps will be needed to capture the slow scales when using explicit methods. If the main interest is in the slow time scales, it may be much more computationally efficient to use implicit methods.

2 Theme 5

1. (a) The point set can always be interpolated with polynomials if all x_i are distinct.

   (b) Interpolation with high-order polynomials yield often strong oscillations between the interpolation points.

2. Cubic splines are well suited for the task. An entirely inappropriate method is to use a polynomial of degree 24.

3. (a) The maximum occurs somewhere in the interval $(0.5, 0.7)$. By interpolate at the points 0.5, 0.6, and 0.7 with a parabola we can use its maximum to estimate the maximum of the underlying function.

   (b) $x_{\max} \approx 0.609$

4. Make an equidistant division of the unit square $(0,1) \times 0,1$ into n intervals of size $h = 1/n$ in each direction and let $x_k = kh$, $y_l = lh$, k, $l = 0, \ldots, n$. Applying the trapezoidal rule first in the y- and then in the x-direction yields

$$\begin{aligned}\int_0^1 \int_0^1 f(x,y)\,\mathrm{d}y\mathrm{d}x &\approx \frac{h}{2} \sum_{l=0}^{n-1} \int_0^1 \left[f(x,y_l) + f(x,y_{l+1})\right] \mathrm{d}x \\ &\approx \frac{h^2}{4} \sum_{l=0}^{n-1} \sum_{k=0}^{n-1} \left[f(x_k,y_l) + f(x_k,y_{l+1}) + f(x_{k+1},y_l) + f(x_{k+1},y_{l+1})\right].\end{aligned}$$

5. The trapezoidal rule yields

$$\begin{aligned}T &= \int_0^{7800} \frac{1}{v(x)}\,\mathrm{d}x \\ &\approx \frac{1300}{2}\left(\frac{1}{750} + \frac{2}{680} + \frac{2}{630} + \frac{2}{640} + \frac{2}{690} + \frac{2}{760} + \frac{1}{830}\right) = 11.25089\ldots.\end{aligned}$$

   The Simpson rule can also be used, of course. The estimate will then be

$$\begin{aligned}T &= \int_0^{7800} \frac{1}{v(x)}\,\mathrm{d}x \\ &\approx \frac{1300}{3}\left(\frac{1}{750} + \frac{4}{680} + \frac{2}{630} + \frac{4}{640} + \frac{2}{690} + \frac{4}{760} + \frac{1}{830}\right) = 11.26962\ldots.\end{aligned}$$

   Thus, with both methods we get an approximate flight time of 11 hours and 15 minutes.