Solution Manual for:
Numerical Computing with MATLAB
by Cleve B. Moler

John L. Weatherwax*

July 25, 2007

Chapter 7 (Ordinary Differential Equations)

Problem 7.1

Defining the vector y as

$$y = \begin{bmatrix} v \\ u \\ \dot{v} \\ \dot{u} \end{bmatrix} \tag{1}$$

Then we have for its time derivative the following

$$\frac{dy}{dt} = \begin{bmatrix} \dot{v} \\ \dot{u} \\ \ddot{v} \\ \ddot{u} \end{bmatrix} = \begin{bmatrix} y_3 \\ y_4 \\ -\frac{y_2}{1+t^2} + \cos(r) \\ -\frac{y_1}{1+t^2} - \sin(r) \end{bmatrix} \tag{2}$$

Where r given in terms of components of y is

$$r = \sqrt{y_4^2 + y_3^2} \tag{3}$$

Given the initial conditions in terms of u, v, $\dot{u}$, and $\dot{v}$ we have

$$y_0 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \tag{4}$$

*wax@alum.mit.edu

Problem 7.2

See the matlab code in the file `prob_7_2.m`.

Problem 7.3

The algorithm BS23 can be represented by the following sequence of steps

$$\begin{aligned}
s_1 &= f(t_n, y_n) && (5)\\
s_2 &= f(t_n + \frac{h}{2}, y_n + \frac{h}{2} s_1) && (6)\\
s_3 &= f(t_n + \frac{3}{4} h, y_n + \frac{3}{4} s_2) && (7)\\
y_{n+1} &= y_n + \frac{h}{9}(2s_1 + 3s_2 + 4s_3) && (8)
\end{aligned}$$

Part (a): The specific ODE's (and their exact solutions) to consider are given by

$$\begin{aligned}
\frac{dy}{dt} &= 1 \quad \Rightarrow \quad y = t + C && (9)\\
\frac{dy}{dt} &= t \quad \Rightarrow \quad y = \frac{t^2}{2} + C && (10)\\
\frac{dy}{dt} &= t^2 \quad \Rightarrow \quad y = \frac{t^3}{3} + C && (11)\\
\frac{dy}{dt} &= t^3 \quad \Rightarrow \quad y = \frac{t^4}{4} + C && (12)
\end{aligned}$$

For the first system, $dy/dt = 1$, so $f(t, y) = 1$ then the BS23 procedure gives the following

$$\begin{aligned}
s_1 &= 1 && (13)\\
s_2 &= 1 && (14)\\
s_3 &= 1 && (15)\\
y_{n+1} &= y_n + \frac{h}{9}(2 + 3 + 4) = y_n + h && (16)
\end{aligned}$$

Does this equal the true solution at $t_n + h$? We evaluate the above to obtain

$$y(t_n + h) = (t_n + h) + C = t_n + C + h = y_n + h \qquad (17)$$

and the answer is yes. For the second function $f(t, y) = t$ we see that the BS23 procedure is

$$\begin{aligned}
s_1 &= t_n && (18)\\
s_2 &= t_n + \frac{h}{2} && (19)\\
s_3 &= t_n + \frac{3}{4} h && (20)\\
y_{n+1} &= y_n + \frac{h}{2}(2t_n + 3(t_n + \frac{h}{2}) + 4(t_n + \frac{3}{4} h)) && (21)
\end{aligned}$$

Simplifying this last equation we have

$$
\begin{aligned}
&= y_n + \frac{h}{9}(9t_n + (\frac{3}{2}+3)h) && (22)\\
&= y_n + ht_n + \frac{h}{9}(\frac{9}{2})h && (23)\\
&= y_n + ht_n + \frac{h^2}{2} && (24)
\end{aligned}
$$

Again we ask, does this equal the true solution at t_{n+1}? Lets check by evaluating

$$y(t_{n+1}) = y(t_n+h) = \frac{1}{2}(t_n+h)^2 + C = \frac{t_n^2}{2} + C + ht_n + \frac{h^2}{2} = y(t_n) + ht_n + \frac{h^2}{2} \qquad (25)$$

and we see that the answer is yes. For the second function $f(t,y) = t^2$, then the BS23 gives

$$
\begin{aligned}
s_1 &= t_n^2 && (26)\\
s_2 &= (t_n + \frac{h}{2})^2 && (27)\\
s_3 &= (t_n + \frac{3}{4}h)^2 && (28)\\
y_{n+1} &= y_n + \frac{h}{9}(2t_n^2 + 3(t_n+\frac{h}{2})^2 + 4(t_n + \frac{3}{4}h)^2) && (29)
\end{aligned}
$$

Now simplifying the last equation we have

$$
\begin{aligned}
y_{n+1} &= y_n + \frac{h}{9}(2t_n^2 + (3+6)ht_n + (\frac{3}{4}+\frac{9}{4})h^2) && (30)\\
&= y_n + \frac{h}{9}(9t_n^2 + \frac{9}{2}ht_n + \frac{12}{4}h^2) && (31)\\
&= y_n + ht_n^2 + h^2t_n + h^3\frac{1}{3} && (32)
\end{aligned}
$$

Does this equal the exact solution at $t_n + h$? To decide we compute

$$
\begin{aligned}
y(t_n+h) &= \frac{(t_n+h)^3}{3} + C == \frac{1}{3}(t_n^3 + 3t_n^2h + 3t_nh^2 + h^3) + C && (33)\\
&= \frac{t_n^3}{3} + C + (t_n^2h + t_nh^2 + \frac{h^3}{3}) && (34)
\end{aligned}
$$

and we see that the answer is yes. Finally, for the function $f(t,y) = t^3$, the BS23 algorithm gives

$$
\begin{aligned}
s_1 &= t_n^3 && (35)\\
s_2 &= (t_n+\frac{h}{2})^3 = t_n^3 + \frac{3}{2}t_n^2h + \frac{3}{4}t_nh^2 + \frac{1}{8}h^3 && (36)\\
s_3 &= (t_n + \frac{3}{4}h)^3 = t_n^3 + \frac{9}{4}t_n^2h + \frac{27}{16}t_nh^2 + \frac{27}{64}h^3 && (37)\\
y_{n+1} &= y_n + \frac{h}{9}(2t_n^3 + 3(t_n^3 + \frac{3}{2}t_n^2h + \frac{3}{4}t_nh^2 + \frac{1}{8}h^3) && (38)\\
&\quad +4(t_n^3 + \frac{9}{4}t_n^2h + \frac{27}{16}t_nh^2 + \frac{27}{64}h^3)) && (39)
\end{aligned}
$$

Simplifying this last term gives

$$y_{n+1} = y_n + h t_n^3 + \frac{3}{2} t_n^2 h^2 + t_n h^3 + \frac{h^4}{6} \tag{40}$$

The true solution at $t_n + h$ however is equal to

$$\begin{aligned} y(t_n + h) &= \frac{(t_n + h)^4}{4} + C = \frac{1}{4}(t_n^4 + 4t_n^3 h + 6t_n^2 h^2 + 4t_n h_3 + h^4) + C & (41) \\ &= \frac{1}{4} t_n^4 + C + t_n^3 h + \frac{3}{2} t_n^2 h^2 + t_n h^3 + \frac{h^4}{4} & (42) \\ &= y_n + h t_n^3 + \frac{3}{2} t_n^2 h^2 + t_n h^3 + \frac{h^4}{4}\,. & (43) \end{aligned}$$

Since this *differs* from the the approximate solution by $\left(\frac{1}{6} - \frac{1}{4}\right) h^4$ we see that this method is *not* exact for $f(t) = t^3$.

Part (b): The BS23 error estimate is given by

$$e_{n+1} = \frac{h}{72}(-5s_1 + 6s_2 + 8s_3 - 9s_4) \tag{44}$$

where s_i for $i = 1, 2, 3, 4$ is computed as above. When one computes s_1, s_2, s_3, and s_4 for $f(t) = 1$ and $f(t) = t$ one obtains a *zero* estimation error. Considering the function $f(t) = t^2$ we have an estimated error given by

$$e_{n+1} = \frac{h}{72}(-5t_n^2 + 6(t_n + \frac{h}{2})^2 + 8(t_n + \frac{3}{4}h)^2 - 9(t_n + h)^2) \tag{45}$$

when one expands the above expression one obtains

$$e_{n+1} = -\frac{h^3}{24} \tag{46}$$

which is not zero. Thus the estimated error is exact for $f(t) = 1$ and $f(t) = t$ only.

Problem 7.4

See the matlab code in the file `prob_7_4.m`.

Problem 7.5

Part (a): See the matlab code in the file `myrk4.m`.

Part (b): The *local* truncation error of RK4 is given by $O(h^{4+1})$, so the *global* truncation error should be $O(h^4)$. When you decrease the step size h by 2 this should reduce the global error by $1/2^4 = 1/16$. An experiment illustrating this behaviour can be found in the routine/code `prob_7_5.m`. There we sequentially decrease the stepsize parameter h and compute the global error between the final timestep and the known true solution value. We then plot the log of stepsize parameter v.s. the log of the global error. On this plot one can see that in this space we have a linear decrease in error representing a powerlaw. A linear function is fitted to to this data with the Matlab function `polyfit`. The slope of this line is given by 4 empirically displaying the known relation between the error and the stepsize of

$$\text{error} \propto h^4$$

Part (c): For this problem formulation our differential equation is given by $\ddot{y} = -y$. The system formulation for this equation is given by defining

$$Y = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} y(t) \\ \dot{y}(t) \end{bmatrix} \tag{47}$$

then

$$\frac{d}{dt}Y(t) = \frac{d}{dt}\begin{bmatrix} y(t) \\ \dot{y}(t) \end{bmatrix} = \begin{bmatrix} \dot{y}(t) \\ \ddot{y}(t) \end{bmatrix} = \begin{bmatrix} y_2(t) \\ -y_1(t) \end{bmatrix} \tag{48}$$

so in total in terms of y_1 and y_2 the differential equation is

$$\frac{d}{dt}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} y_2 \\ -y_1 \end{bmatrix} \tag{49}$$

This is implemented in the Matlab code `prob_7_5_fn.m`. In addition, in the driver routine `prob_7_5.m` we have implemented the numerical experiments mentioned in the book and have verified the statements there.

Problem 7.6

Part (a): The differential equation

$$\dot{y} = -1000y + 1000\sin(t) + \cos(t) \quad \text{with} \quad y(0) = 1 \tag{50}$$

has a homogeneous solution given by $y(t) = Ce^{-1000t}$ and a *particular* solution given by a function of the form $A\sin(t) + B\cos(t)$. By putting this particular solution into the equation above we can solve for the coefficients A and B. When we do this we get the following

$$A\cos(t) - B\sin(t) = -1000A\sin(t) - 1000B\cos(t) + 1000\sin(t) + \cos(t) \tag{51}$$

Grouping the coefficients of $\cos(t)$ and $\sin(t)$ we obtain

$$(A + 1000B - 1)\cos(t) + (-B + 1000A - 1000)\sin(t) = 0\,. \tag{52}$$

When we require that the coefficients of $\cos(t)$ and $\sin(t)$ vanish we obtain the following system of equations

$$\begin{aligned} A + 1000B - 1 &= 0 && (53) \\ -B + 1000A - 1000 &= 0\,. && (54) \end{aligned}$$

The solution to the first equation in terms of B is given by $A = 1 - 1000B$. When substituted into the second equation we obtain

$$1000(1 - 1000B) - B = 1000 \tag{55}$$

Which has as its solution $B = 0$. Thus $A = 1$ giving in total the solution y of

$$y(t) = \sin(t) + Ce^{-1000t} \tag{56}$$

using the initial condition $y(0) = 1$ we obtain $C = 1$ and the total analytical solution is given by

$$y(t) = e^{-1000t} + \sin(t) \tag{57}$$

I should mention that heuristically this problem is midly stiff because of the various time scales present in the solution. The exponential term has a timescale $O(1/1000)$ while the sin term has a timescale like $O(1)$ or effectivly three orders of magnitude different.

Part (b)-(f): See the routines `prob_7_6.m` and `prob_7_6_fn.m` for an implementation of this problem.

Problem 7.7

Part (a): Each equation has as its solution $y(t) = \sin(t)$.

Part (b): The corresponding first order system for the third given differential equations is (defining $y_1(t) = y(t)$)

$$\frac{d}{dt}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} y_2 \\ -y_1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \tag{58}$$

The corresponding system for the fourth differential equation is given by

$$\frac{d}{dt}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} y_2 \\ -\sin(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} 0 \\ -\sin(t) \end{bmatrix} \tag{59}$$

Note that the first representation above is *autonomous* while the second is not.

Part (c): The Jacobian for each of the functions f involved in the definitions of each differential equation are given by

$$\begin{aligned}
\frac{\partial f_1}{\partial y} &= 0 \\
\frac{\partial f_2}{\partial y} &= -\frac{y}{\sqrt{1-y^2}} \\
\frac{\partial f_3}{\partial y} &= \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \\
\frac{\partial f_4}{\partial y} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}
\end{aligned}$$

From the above we can see that as $t \to \frac{\pi}{2}$ that WWX finish

Problem 7.9

Consider the matrix A given by

$$A = \begin{bmatrix} -\beta & 0 & \eta \\ 0 & -\sigma & \sigma \\ -\eta & \rho & -1 \end{bmatrix} \tag{60}$$

Expanding the determinant of the above matrix we have

$$|A| = -\beta \left| \begin{array}{cc} -\sigma & \sigma \\ \rho & -1 \end{array} \right| - \eta \left| \begin{array}{cc} 0 & \eta \\ -\sigma & \sigma \end{array} \right| \tag{61}$$

or

$$|A| = -\beta(\sigma - \sigma\rho) - \eta(\sigma\eta) \tag{62}$$

or

$$|A| = \eta^2 - \rho\beta + \beta \tag{63}$$

Setting this equal to zero to find the values where the matrix is singular we see that

$$\eta = \pm\sqrt{\beta(\rho - 1)}\,. \tag{64}$$

Which was the expression we were looking to find. To find the vectors in the nullity of A when η is one of these special values, insert this into the matrix above and reduce it to row-echelon form. Inserting this expression into A means we are looking for a vector v such that

$$Av = \begin{bmatrix} -\beta & 0 & \pm\sqrt{\beta(\rho-1)} \\ 0 & -\sigma & \sigma \\ \mp\sqrt{\beta(\rho-1)} & \rho & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = 0 \tag{65}$$

dividing the first row in the above expression by β and the second row by $-\sigma$ we obtain

$$\begin{bmatrix} 1 & 0 & \mp\sqrt{\frac{(\rho-1)}{\beta}} \\ 0 & 1 & -1 \\ \mp\sqrt{\beta(\rho-1)} & \rho & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = 0 \tag{66}$$

Now multiplying the first row by $\pm\sqrt{\beta(\rho-1)}$ and adding it to the third row we obtain

$$\begin{bmatrix} 1 & 0 & \mp\sqrt{\frac{(\rho-1)}{\beta}} \\ 0 & 1 & -1 \\ 0 & \rho & -1 - \sqrt{\frac{(\rho-1)}{\beta}}\sqrt{\beta(\rho-1)} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = 0 \tag{67}$$

which simplifies to

$$\begin{bmatrix} 1 & 0 & \mp\sqrt{\frac{(\rho-1)}{\beta}} \\ 0 & 1 & -1 \\ 0 & \rho & -\rho \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = 0\,. \tag{68}$$

which demonstrates that the third row is indeed a multiple of the second row. Eliminating this row we obtain the following requirements among the coefficients of v

$$v_1 = \pm\sqrt{\frac{(\rho-1)}{\beta}}v_3 \quad (69)$$

$$v_2 = v_3 \quad (70)$$

Letting $v_3 = \eta = \sqrt{\beta(\rho-1)}$ and assuming that $\rho - 1 > 0$ we see that the above null vector has components given by

$$v_1 = \rho - 1 \quad (71)$$

$$v_2 = \eta \quad (72)$$

$$v_3 = \eta \quad (73)$$

as desired.

Problem 7.10 (the Jacobian of the Lorenz equations)

The Lorenz equations are given by $\dot{y} = Ay$, with

$$y(t) = \left[\begin{array}{c} y_1(t) \\ y_2(t) \\ y_3(t) \end{array}\right] \quad \text{and} \quad A = \left[\begin{array}{ccc} -\beta & 0 & y_2(t) \\ 0 & -\sigma & \sigma \\ -y_2(t) & \rho & -1 \end{array}\right],$$

Performing the matrix multiplication we have for the full non-linear system representing the Lorenz equations, the following

$$\begin{aligned} \dot{y_1} &= -\beta y_1(t) + y_2(t)y_3(t) \\ \dot{y_2} &= -\sigma y_2(t) + \sigma y_3(t) \\ \dot{y_3} &= -y_1(t)y_2(t) + \rho y_2(t) - y_3(t)\,. \end{aligned}$$

This expression makes it easier to compute the required Jacobian of our mapping $f(\cdot)$. We find that

$$J = \left[\begin{array}{ccc} -\beta & y_3(t) & y_2(t) \\ 0 & -\sigma & \sigma \\ -y_2(t) & -y_1(t)+\rho & -1 \end{array}\right].$$

The fixed points of the Lorentz equations are given by solving for y in $f(y) = 0$, which gives the following system of equations

$$\begin{aligned} -\beta y_1 + y_2 y_3 &= 0 \\ -\sigma y_2 + \sigma y_3 &= 0 \\ -y_1 y_2 + \rho y_2 - y_3 &= 0\,. \end{aligned}$$

The second equation above has the simple solution of $y_2 = y_3$, which when used in the first and third equaions above gives

$$\begin{aligned} -\beta y_1 + y_2^2 &= 0 \\ -y_1 y_2 + \rho y_2 - y_1 &= 0\,. \end{aligned}$$

The last equation above can be factored as $y_2(-y_1+\rho-1)=0$, which has as solutions $y_2=0$, and $y_1=\rho-1$. Following the consequences of each of these choices, if $y_2=0$, then $y_1=0$, while if $y_1=\rho-1$ then the first equation above gives

$$-\beta(\rho-1)+y_2^2=0\,,$$

which has a solution given by

$$y_2=\pm\sqrt{\beta(\rho-1)}\,.$$

Thus in summary then the three fixed points of the Jacobian of the Lorentz equations are

$$\begin{aligned}(y_1,y_2,y_3)&=(0,0,0)\\(y_1,y_2,y_3)&=(\rho-1,\sqrt{\beta(\rho-1)},\sqrt{\beta(\rho-1)})\\(y_1,y_2,y_3)&=(0,0,0)\,.\end{aligned}$$