Math 5604 Homework 3

Jacob Hauck

March 11, 2024

Problem 1.

Consider the IVP

$$\begin{aligned} x' &= x^2 y - e^{-t} - e^{-2t}\cos(t) \\ y' &= yz - \sin(t) - t^2\cos(t) \\ z' &= x + y + 2t - e^{-t} - \cos(t) \\ x(0) = 1, &\quad y(0) = 1, \quad z(0) = 0. \end{aligned} \tag{1}$$

(a) Assuming a time step of $k > 0$ with time nodes $\{t_n\}_{n=0}^N$, with $t_0 = 0$ and $t_N = 1$, we can discretize this IVP on the interval $[0, 1]$ using the following backward Euler scheme:

$$\begin{aligned} x^{n+1} &= x^n + k\left[\left(x^{n+1}\right)^2 y^{n+1} - e^{-t_{n+1}} - e^{-2t_{n+1}}\cos(t_{n+1})\right] \\ y^{n+1} &= y^n + k\left[y^{n+1}z^{n+1} - \sin(t_{n+1}) - t_{n+1}^2\cos(t_{n+1})\right] \qquad\qquad n = 0, 1, \ldots N-1 \\ z^{n+1} &= z^n + k\left[x^{n+1} + y^{n+1} + 2t_{n+1} - e^{-t_{n+1}} - \cos(t_{n+1})\right] \\ x^0 = 1, &\quad y^0 = 1, \quad z^0 = 0. \end{aligned} \tag{2}$$

Since $\left(x^{n+1}, y^{n+1}, z^{n+1}\right)^T$ is a root of $f_n(u, v, w)$, where

$$f_n(u, v, w) = \begin{bmatrix} u - x^n - k\left[u^2 v - e^{-t_{n+1}} - e^{-2t_{n+1}}\cos(t_{n+1})\right] \\ v - y^n - k\left[vw - \sin(t_{n+1}) - t_{n+1}^2\cos(t_{n+1})\right] \\ w - z^n - k\left[u + v + 2t_{n+1} - e^{-t_{n+1}} - \cos(t_{n+1})\right] \end{bmatrix}, \tag{3}$$

we can use Newton's method to find $\left(x^{n+1}, y^{n+1}, z^{n+1}\right)^T$ by finding the root of f_n using an initial guess of $(x^n, y^n, z^n)^T$. In order to use Newton's method, we will need the Jacobian Df_n of f_n:

$$Df_n(u, v, w) = \begin{bmatrix} 1 - 2kuv & -ku^2 & 0 \\ 0 & 1 - kw & -kv \\ -k & -k & 1 \end{bmatrix}. \tag{4}$$

The implementation of the backward Euler method for this problem can be found in `problem1.m`, and the implementation of Newton's method can be found in `newton.m`.

(b) Using `problem1_calculations.m` to calculate the numerical values of $x(1)$, $y(1)$, and $z(1)$ with step size $k \in \{1/16, 1/64\}$, we get

$$\begin{aligned} (0.400273, 0.540425, 1.075813)^T, &\qquad k = \frac{1}{16} \\ (0.375735, 0.539848, 1.018419)^T, &\qquad k = \frac{1}{64} \end{aligned}$$

(c) Using `problem1_calculations.m` to calculate the numerical errors at $t = 1$ from the exact solution $(e^{-t}, \cos(t), t^2)^T$, we get the results in Table 1, which are copied from `p1_output.txt`. We notice that the convergence rate for each component and in ℓ^∞ seems to be 1. The $y(t)$ convergence, however, doesn't start to follow a pattern until the step size is small (in particular, the first 3 or 4 rate entries are all over the place).

	x		y		z		ℓ^∞	
k	Error	Rate	Error	Rate	Error	Rate	Error	Rate
1/4	0.158586	-	0.061537	-	0.362011	-	0.362011	-
1/8	0.068026	1.221101	0.006829	3.171723	0.158577	1.190850	0.158577	1.190850
1/16	0.032393	1.070401	0.000123	5.795091	0.075813	1.064661	0.075813	1.064661
1/32	0.015864	1.029925	0.000605	-2.298793	0.037176	1.028067	0.037176	1.028067
1/64	0.007856	1.013926	0.000455	0.412177	0.018419	1.013176	0.018419	1.013176
1/128	0.003910	1.006730	0.000264	0.785388	0.009169	1.006394	0.009169	1.006394
1/256	0.001950	1.003310	0.000141	0.905477	0.004574	1.003150	0.004574	1.003150
1/512	0.000974	1.001641	0.000073	0.955407	0.002285	1.001564	0.002285	1.001564

Table 1: Errors and convergence rates of backward Euler using different error metrics

Problem 2.

Recall the backward Euler method for the IVP

$$y' = f(t, y), \quad t > 0; \qquad y(t_0) = a \tag{5}$$

is given implicitly by the scheme

$$y^{n+1} = y^n + kf\left(t_{n+1}, y^{n+1}\right), \qquad n = 0, 1, 2, \dots \tag{6}$$

$$y^0 = a, \tag{7}$$

where $\{t_n\}$ is a sequence of evenly-spaced times (with the same t_0 from (5)) with $t_{n+1} - t_n = k$. The value y^n is meant to be an approximation of $y(t_n)$.

Define $e_n = y(t_n) - y^n$. On a given interval $[t_0, t_0 + T]$, suppose we use a step size $k = \frac{T}{N}$, so that $t_N = t_0 + T$. Then the global truncation error (GTE) is given by $\max_{0 \le n \le N} |e_n|$.

Assume that f is L-Lipschitz in y uniformly for $t \in [t_0, t_0 + T]$, and assume that $y \in C^2([t_0, t_0 + T])$, with $|y''(t)| \le C$ for all $t \in [t_0, t_0 + T]$.

By Taylor's Theorem, for all $n = 0, 1, 2, \dots N - 1$, there exists $\tau_n \in [t_n, t_{n+1}]$ such that

$$y(t_{n+1}) = y(t_n) + ky'(t_{n+1}) + \frac{1}{2}k^2 y''(\tau_n).$$

Then

$$\begin{aligned} y(t_{n+1}) &= y(t_n) - y_n + y_n + kf\left(t_{n+1}, y^{n+1}\right) + k\left[f(t_{n+1}, y(t_{n+1})) - f\left(t_{n+1}, y^{n+1}\right)\right] + \frac{1}{2}k^2 y''(\tau_n) \\ &= e_n + y^{n+1} + k\left[f(t_{n+1}, y(t_{n+1})) - f\left(t_{n+1}, y^{n+1}\right)\right] + \frac{1}{2}k^2 y''(\tau_n). \end{aligned}$$

Hence, by the assumptions on y and f,

$$\begin{aligned} |e_{n+1}| &\le |e_n| + k\left|f(t_{n+1}, y(t_{n+1})) - f\left(t_{n+1}, y^{n+1}\right)\right| + \frac{1}{2}k^2 |y''(\tau_n)| \\ &\le |e_n| + kL\left|y(t_{n+1}) - y^{n+1}\right| + \frac{1}{2}Ck^2 \\ &= |e_n| + kL|e_{n+1}| + \frac{1}{2}Ck^2 \end{aligned}$$

This holds for all $n = 0, 1, 2, \dots N-1$. Noting that $y^0 = a = y(t_0)$, we have $e_0 = 0$, so this gives us a recurrent set of inequalities for $|e_n|$. Since we are only interested in proving GTE $\to 0$ as $k \to 0$, we can safely assume that $k < \frac{1}{L}$. In this case, we have

$$|e_{n+1}| \leq \frac{|e_n| + \frac{1}{2}Ck^2}{1-kL}, \qquad n = 0, 1, 2, \dots, N-1. \tag{8}$$

Using the fact that $e_0 = 0$ and iterating (8), we get

$$|e_n| \leq \sum_{j=0}^{n-1} \frac{\frac{1}{2}Ck^2}{(1-kL)^{j+1}} = \frac{\frac{1}{2}Ck^2}{1-kL} \sum_{j=0}^{n-1} \left(\frac{1}{1-kL}\right)^j = \frac{\frac{1}{2}Ck^2}{1-kL} \frac{\left(\frac{1}{1-kL}\right)^n - 1}{\frac{1}{1-kL} - 1} = \frac{Ck}{2L}\left[\left(\frac{1}{1-kL}\right)^n - 1\right].$$

Since $1 - kL > 0$ and $kL \geq 0$, it follows that $\left(\frac{1}{1-kL}\right)^n \leq \left(\frac{1}{1-kL}\right)^N$ for $n = 0, 1, \dots, N$. Recalling that $k = \frac{T}{N}$, we have

$$\text{GTE} = \max_{0 \leq n \leq N} |e_n| \leq \frac{Ck}{2L}\left[\left(1 - \frac{TL}{N}\right)^{-N} - 1\right].$$

If $kL = \frac{TL}{N}$ is close to 1, then this bound doesn't say much. Since we are interested in bounding the error as $k \to 0$, and we have already assumed that $k < \frac{1}{L}$, there is no harm in further assuming that $k < \frac{1}{2L}$. Thus, $\frac{TL}{N} \leq \frac{1}{2}$. Note that by the Taylor series for $\log(1-x)$,

$$-\log(1-x) = x + \frac{x^2}{2} + \frac{x^3}{3} + \dots \leq x + x^2, \qquad 0 \leq x \leq \frac{1}{2}$$

because

$$\frac{x^2}{2} + \frac{x^3}{3} + \dots \leq \frac{x^2}{2}\left(1 + x + x^2 + \dots\right) = \frac{x^2}{2} \cdot \frac{1}{1-x} \leq x^2, \qquad 0 \leq x \leq \frac{1}{2}$$

Therefore,

$$\text{GTE} \leq \frac{Ck}{2L}\left[e^{-N\log\left(1 - \frac{TL}{N}\right)} - 1\right] \leq \frac{Ck}{2L}\left[e^{TL + \frac{(TL)^2}{N}} - 1\right] \leq \frac{Ck}{2L}\left[e^{\frac{3TL}{2}} - 1\right],$$

which shows that GTE $= \mathcal{O}(k)$ as $k \to 0$. Thus, the Backward Euler method is convergent, and the convergence order is 1.