Homework 1.

The Butcher's Theorem states that for a s-stage Runge-Kutta method:

$$y_{n+1} = y_n + h\sum_{i=1}^{s} b_i k_i$$
$$k_i = f(t_n + c_i h, y_n + h\sum_{j=1}^{s} a_{ij} k_j),$$

if there exist p, η, and ξ such that the following equations are satisfied

$$B(p): \quad \sum_{i=1}^{s} b_i c_i^{k-1} = \frac{1}{k} \qquad k=1,2,\cdots p$$
$$C(\eta): \quad \sum_{j=1}^{s} a_{ij} c_j^{k-1} = \frac{c_i^k}{k} \qquad i=1,2,\cdots s,\ k=1,2,\cdots p$$
$$D(\xi): \quad \sum_{i=1}^{s} b_i c_i^{k-1} a_{ij} = \frac{b_j(1-c_j^k)}{k} \qquad j=1,2,\cdots s,\ k=1,2,\cdots \xi$$

and

$$p \le \eta + \xi + 1, \quad p \le 2(\eta + 1),$$

then the Runge-Kutta method is of order p.

Now we use this theorem to find out the order of accuracy of the following methods.

Problem 1.

(a) modified Euler.

For this method, we have

$$b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad c = \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 0 \\ \frac{1}{2} & 0 \end{pmatrix}$$

Let's first check $B(p)$:

$$\sum_{i=1}^{s} b_i c_i^{k-1} = 0 + (\tfrac{1}{2})^{k-1}$$

$k=1$: $(\frac{1}{2})^0 = \frac{1}{1}$

$k=2$: $(\frac{1}{2})^1 = \frac{1}{2}$

$k=3$: $(\frac{1}{2})^2 \neq \frac{1}{3}$

Therefore we have $p=2$.

Next we check $C(\eta)$

① $k=1$ $\quad \sum_{j=1}^{s} a_{ij} c_j^0 = \sum_{j=1}^{s} a_{ij}$

$i=1$: $\sum_{j=1}^{s} a_{ij} = 0 = \frac{c_1}{1}$

$i=2$ $\quad \sum_{j=1}^{s} a_{ij} = \frac{1}{2} = \frac{c_2}{1}$

② $k=2$

$i=1$: $\sum_{j=1}^{s} a_{ij} c_j = 0 = \frac{c_1}{2}$

$i=2$: $\sum_{j=1}^{s} a_{ij} c_j = \frac{1}{2}\cdot 0 + 0\cdot\frac{1}{2} \neq \frac{1/2}{2} = \frac{c_2}{2}$

Therefore, we have $\eta = 1$.

Finally, for $D(\xi)$, we can choose $\xi = 0$ and then we don't need to check anything.

So we have $p=2, \eta=1, \xi=0$, which satisfy

$$p \leq \eta+\xi+1 \quad \text{and} \quad p \leq 2(\eta+1).$$

Hence, the modified Euler method is order 2.

(b) Improved Euler.

We have

$$b = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \quad c = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Let's first check $B(p)$:

$k=1$: $\sum_{i=1}^{s} b_i c_i^0 = \sum_{i=1}^{s} b_i = \frac{1}{2} + \frac{1}{2} = 1 = \frac{1}{1}$

$k=2$ $\quad \sum_{i=1}^{s} b_i c_i = 0 + \frac{1}{2} = \frac{1}{2}$

$k=3$ $\quad \sum_{i=1}^{s} b_i c_i^2 = 0 + \frac{1}{2}\cdot 1 \neq \frac{1}{3}$

Therefore $p=2$.

Next we check $C(\eta)$:

$k=1$ $\quad \sum_{j=1}^{s} a_{ij} c_j^0 = \sum_{j=1}^{s} a_{ij} = \begin{cases} 0 = \frac{c_1}{1} & i=1 \\ 1 = \frac{c_2}{1} & i=2 \end{cases}$

$k=2$ $\quad \sum_{j=1}^{s} a_{ij} c_j = \begin{cases} 0 & i=1 \\ 0 \neq \frac{c_2}{2} & i=2 \end{cases}$

So we have $\eta = 1$.

Again we can pick $\xi=0$ and we don't need to check $D(\xi)$.

So we have $p=2$, $\eta=1$ and $\xi=0$.

Clearly $p \le \eta+\xi+1$ and $p \le 2(\eta+1)$.

So the improved Euler method is order 2.

The classic explicit Euler method is of order 1. Therefore, the improved Euler method is an improvements over the classic one.

Problem 2.

(a) We have

$$b=\begin{pmatrix}\frac{1}{2}\\ \frac{1}{2}\end{pmatrix} \quad C=\begin{pmatrix}\frac{3-\sqrt{3}}{6}\\ \frac{3+\sqrt{3}}{6}\end{pmatrix} \quad A=\begin{pmatrix}\frac{1}{4} & \frac{3-\sqrt{12}}{12}\\ \frac{3+\sqrt{12}}{12} & \frac{1}{4}\end{pmatrix}$$

Let us first check $B(p)$:

$k=1$ $\quad \sum_{i=1}^{s} b_i C_i^0 = \sum_{i=1}^{s} b_i = \frac{1}{2}+\frac{1}{2}=1$

$k=2$ $\quad \sum_{i=1}^{s} b_i C_i = \frac{1}{2}\left(\frac{3-\sqrt{3}}{6}+\frac{3+\sqrt{3}}{6}\right)=\frac{1}{2}$

$k=3$ $\quad \sum_{i=1}^{s} b_i C_i^2 = \frac{1}{2}\left(\left(\frac{3-\sqrt{3}}{6}\right)^2+\left(\frac{3+\sqrt{3}}{6}\right)^2\right)=\frac{1}{2}\cdot 2\,\frac{3^2+3}{36}=\frac{1}{3}$

$k=4$ $\quad \sum_{i=1}^{s} b_i C_i^3 = \frac{1}{2}\left(\left(\frac{3-\sqrt{3}}{6}\right)^3+\left(\frac{3+\sqrt{3}}{6}\right)^3\right)=\frac{1}{2}\left(\frac{3-\sqrt{3}}{6}+\frac{3+\sqrt{3}}{6}\right)\left(\left(\frac{3-\sqrt{3}}{6}\right)^2-\frac{3-\sqrt{3}}{6}\cdot\frac{3+\sqrt{3}}{6}+\left(\frac{3+\sqrt{3}}{6}\right)^2\right)=\frac{1}{2}\left(\frac{2}{3}-\frac{1}{6}\right)=\frac{1}{4}$

For a 2-stage R-K method, we can have at most 4th order. So we can stop here and choose $p=4$.

Next we check $C(\eta)$

$k=1$ $\quad \sum_{j=1}^{s} a_{ij} C_j^0 = \sum_{j=1}^{s} a_{ij}$

① $i=1$ $\quad \sum_{j=1}^{s} a_{ij} = \frac{1}{4}+\frac{3-\sqrt{12}}{12} = \frac{6-2\sqrt{3}}{12} = \frac{3-\sqrt{3}}{6} = \frac{C_1}{1}$

② $i=2$ $\quad \sum_{j=1}^{s} a_{ij} = \frac{1}{4}+\frac{3+\sqrt{12}}{12} = \frac{6+2\sqrt{3}}{12} = \frac{3+\sqrt{3}}{6} = \frac{C_2}{1}$

$k=2$

① $i=1$: $\quad \sum_{j=1}^{s} a_{ij} C_j = \frac{1}{4}\cdot\frac{3-\sqrt{3}}{6} + \frac{3-2\sqrt{3}}{12}\cdot\frac{3+\sqrt{3}}{6} = \frac{3-\sqrt{3}}{24}+\frac{3-3\sqrt{3}}{72} = \frac{3-\sqrt{3}}{12} = \frac{C_1}{2}$

② $i=2$ $\quad \sum_{j=1}^{s} a_{ij} C_j = \frac{3+\sqrt{12}}{12}\cdot\frac{3-\sqrt{3}}{6} + \frac{1}{4}\cdot\frac{3+\sqrt{3}}{6} = \frac{3+3\sqrt{3}}{12}+\frac{3+\sqrt{3}}{24} = \frac{3+\sqrt{3}}{12} = \frac{C_2}{2}$

Notice that if we choose $\eta=2$, then $4=p\le 2(\eta+2)$. If we have $\xi=1$, then $4 = p \le \eta + \xi + 1$.

So before we check if $\eta = 3$ holds, let's first check $D(1)$.

If $\xi = 1$ holds, then there is no need to check $C(3)$.

For $D(\xi)$, if $k = 1$

$$\sum_{i=1}^{s} b_i c_i^0 a_{ij} = \begin{cases} \frac{1}{2}\cdot\frac{1}{4} + \frac{1}{2}\cdot\frac{3+\sqrt{12}}{12} = \frac{3+\sqrt{3}}{12} = \frac{\frac{1}{2}\cdot(1-\frac{3-\sqrt{3}}{6})}{1} & j=1 \\ \frac{1}{2}\cdot\frac{3-\sqrt{12}}{12} + \frac{1}{2}\cdot\frac{1}{4} = \frac{3-\sqrt{3}}{12} = \frac{\frac{1}{2}(1-\frac{3+\sqrt{3}}{6})}{1} & j=2 \end{cases}$$

So $\xi = 1$ holds.

So we have $p = 4$, $\eta = 2$, $\xi = 1$ and $p \le \eta + \xi + 1$, $p \le 2(\eta + 1)$.

Hence the method is of order 4.

(b) For the second method, we have

$$b = \begin{pmatrix} 3/4 \\ 1/4 \end{pmatrix} \quad c = \begin{pmatrix} 1/3 \\ 1 \end{pmatrix} \quad A = \begin{pmatrix} 5/12 & -1/12 \\ 3/4 & 1/4 \end{pmatrix}$$

Let us first check $B(p)$:

$k=1$: $\sum_{i=1}^{s} b_i c_i^0 = 3/4 + 1/4 = 1$

$k=2$: $\sum_{i=1}^{s} b_i c_i = \frac{3}{4}\cdot\frac{1}{3} + \frac{1}{4}\cdot 1 = \frac{1}{2}$

$k=3$: $\sum_{i=1}^{s} b_i c_i^2 = \frac{3}{4}\cdot\frac{1}{9} + \frac{1}{4}\cdot 1 = \frac{1}{12} + \frac{1}{4} = \frac{1}{3}$

$k=4$: $\sum_{i=1}^{s} b_i c_i^3 = \frac{3}{4}\cdot\frac{1}{27} + \frac{1}{4}\cdot 1 = \frac{1}{36} + \frac{9}{36} = \frac{5}{18} \ne \frac{1}{4}$

Therefore, we have $p = 3$.

Next we check $C(\eta)$:

$k=1$: $\sum_{j=1}^{s} a_{ij} c_j^0 = \sum_{j=1}^{s} a_{ij} = \begin{cases} \frac{1}{3} = c_1 & i=1 \\ 1 = c_2 & i=2 \end{cases}$

$k=2$: $\sum_{j=1}^{s} a_{ij} c_j = \begin{cases} \frac{5}{12}\cdot\frac{1}{3} - \frac{1}{12}\cdot 1 = \frac{1}{18} = c_1^2/2 & i=1 \\ \frac{3}{4}\cdot\frac{1}{3} + \frac{1}{4}\cdot 1 = \frac{1}{2} = \frac{c_2^2}{2} & i=2 \end{cases}$

So we have $\eta \ge 2$.

Now we just take $\xi = 0$ and we have $3 = p \le 2 + 1 \le \eta + \xi + 1$ and $3 = p \le 2(\eta + 1)$.

So we know the second method is of order 3.

Therefore, the first method is more accurate.

Problem 3.

First let us use vector and matrix notation to describe the Runge-Kutta method.

$$y_{n+1} = y_n + h b^T k$$

where $k = \begin{pmatrix} k_1 \\ \vdots \\ k_s \end{pmatrix}$ and $k_i = f(t_n + hc_i, y_0 + h(Ak)_i)$

If $f(t,y) = \lambda y$, then we know

$$k = \lambda(y_n \vec{1} + hAk) \quad \text{where } \vec{1} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

$$\Rightarrow (I - \lambda h A) k = \lambda y_n \vec{1}$$

$$k = \lambda y_n (I - \lambda h A)^{-1} \vec{1}$$

Hence,

$$y_{n+1} = y_n + h b^T \lambda y_n (I - \lambda h A)^{-1} \vec{1}$$

$$= (1 + h\lambda b^T (I - \lambda h A)^{-1} \vec{1}) y_n$$

Let $R(x) = 1 + x b^T (I - xA)^{-1} \vec{1}$. Then we have

$$y_{n+1} = R(\lambda h) y_n$$

and $R(\lambda h)$ only depends on A and b, not on C.

Notice that the solution to $\frac{dy}{dt} = \lambda y$ is $y = y_0 e^{\lambda t}$. Hence the local truncation error of our R-K method is

$$|y_{n+1} - y(t_{n+1})| = |R(\lambda h) y(t_n) - y(t_n + h)|$$

$$= |R(\lambda h) y(t_n) - e^{\lambda h} y(t_n)|$$

$$= |(R(\lambda h) - e^{\lambda h}) y(t_n)|$$

$$\leq O(h^{p+1})$$

Since the LTE is $O(h^{p+1})$,

we know the method is of order p.

To verify the previous methods, we need to compute $R(\lambda h)$ and compare it with the Taylor series of $e^{\lambda h}$.

① Modified Euler

$$b=\begin{pmatrix}0\\1\end{pmatrix} \quad c=\begin{pmatrix}0\\ \frac{1}{2}\end{pmatrix} \quad A=\begin{pmatrix}0&0\\ \frac{1}{2}&0\end{pmatrix}.$$

$$R(x) = 1 + x b^T (I - xA)^{-1}\vec{1}$$
$$= 1 + x + \frac{x^2}{2}$$

So $R(\lambda h) = 1 + \lambda h + \frac{(\lambda h)^2}{2}$

$$e^{\lambda h} = 1 + \lambda h + \frac{(\lambda h)^2}{2} + O(h^3)$$

$$\Rightarrow e^{\lambda h} = R(\lambda h) + O(h^3)$$

Hence the modified Euler is of order 2.

② Improved Euler

$$b=\begin{pmatrix}\frac{1}{2}\\ \frac{1}{2}\end{pmatrix} \quad c=\begin{pmatrix}0\\ 1/2\end{pmatrix} \quad A=\begin{pmatrix}0&0\\1&0\end{pmatrix}$$

Then $R(x) = 1 + x b^T (I - xA)^{-1}\vec{1}$
$$= 1 + x + \frac{x^2}{2}$$

So $R(\lambda h) = 1 + \lambda h + \frac{(\lambda h)^2}{2}$

and $e^{\lambda h} = R(\lambda h) + O(h^3)$

Hence the improved Euler is of order 2.

Notice that the modified Euler and improved Euler methods have the same $R(\lambda h) = 1 + \lambda h + \frac{(\lambda h)^2}{2}$. So there are actually the same method. As you can see in the next exercise, their numerical results are exactly the same.

③ $b = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$ $C = \begin{pmatrix} \frac{3-\sqrt{3}}{6} \\ \frac{3+\sqrt{3}}{6} \end{pmatrix}$ $A = \begin{pmatrix} \frac{1}{4} & \frac{3-\sqrt{12}}{12} \\ \frac{3+\sqrt{12}}{12} & \frac{1}{4} \end{pmatrix}$

$$R(x) = 1 + x b^T (I - xA)^{-1} \vec{1}$$
$$= \frac{12+6x+x^2}{12-6x+x^2}$$
$$= 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{144} + O(x^6)$$

Notice that $e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120} + O(x^6)$

Hence $e^{\lambda h} = R(\lambda h) + O(h^5)$

So the method is of order 4.

④ $b = \begin{pmatrix} 3/4 \\ 1/4 \end{pmatrix}$ $C = \begin{pmatrix} 1/3 \\ 1 \end{pmatrix}$ $A = \begin{pmatrix} 5/12 & -1/12 \\ 3/4 & 1/4 \end{pmatrix}$

$$R(x) = 1 + x b^T (I - xA)^{-1} \vec{1}$$
$$= \frac{6+2x}{6-4x+x^2}$$
$$= 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{36} + O(x^5)$$

Notice that $e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + O(x^5)$

Therefore we know

$$e^{\lambda h} = R(\lambda h) + O(h^4)$$

So the method is order of 3.

Problem 4.

Numerical Results:

n	$e_{\theta,n}$	$\alpha_{\theta,n}$	$C_{\theta,n}$	$e_{\omega,n}$	$\alpha_{\omega,n}$	$C_{\omega,n}$
2	4.637×10^{-2}	–	–	5.910×10^{-2}	–	–
4	1.418×10^{-2}	1.710	1.517×10^{-1}	1.232×10^{-2}	2.262	2.835×10^{-1}
8	3.821×10^{-3}	1.891	1.951×10^{-1}	2.731×10^{-3}	2.174	2.510×10^{-1}
16	9.872×10^{-4}	1.953	2.216×10^{-1}	6.360×10^{-4}	2.102	2.161×10^{-1}
32	2.506×10^{-4}	1.978	2.377×10^{-1}	1.530×10^{-4}	2.056	1.901×10^{-1}

scheme #1
(modified Euler)

n	$e_{\theta,n}$	$\alpha_{\theta,n}$	$C_{\theta,n}$	$e_{\omega,n}$	$\alpha_{\omega,n}$	$C_{\omega,n}$
2	4.637×10^{-2}	–	–	5.910×10^{-2}	–	–
4	1.418×10^{-2}	1.710	1.517×10^{-1}	1.232×10^{-2}	2.262	2.835×10^{-1}
8	3.821×10^{-3}	1.891	1.951×10^{-1}	2.731×10^{-3}	2.174	2.510×10^{-1}
16	9.872×10^{-4}	1.953	2.216×10^{-1}	6.360×10^{-4}	2.102	2.161×10^{-1}
32	2.506×10^{-4}	1.978	2.377×10^{-1}	1.530×10^{-4}	2.056	1.901×10^{-1}

scheme #2
(improved Euler)

n	$e_{\theta,n}$	$\alpha_{\theta,n}$	$C_{\theta,n}$	$e_{\omega,n}$	$\alpha_{\omega,n}$	$C_{\omega,n}$
2	1.463×10^{-4}	–	–	8.447×10^{-4}	–	–
4	9.256×10^{-6}	3.982	3.312×10^{-3}	5.344×10^{-6}	3.982	1.335×10^{-3}
8	5.803×10^{-7}	3.996	2.355×10^{-3}	3.550×10^{-7}	4.00	1.360×10^{-3}
16	3.630×10^{-8}	4.00	2.372×10^{-3}	2.096×10^{-8}	4.00	1.369×10^{-3}

scheme #3

n	$e_{\theta,n}$	$\alpha_{\theta,n}$	$C_{\theta,n}$	$e_{\omega,n}$	$\alpha_{\omega,n}$	$C_{\omega,n}$
2	1.197×10^{-3}	–	–	2.976×10^{-3}	–	–
4	1.787×10^{-4}	2.744	8.020×10^{-3}	3.665×10^{-4}	3.022	2.417×10^{-2}
8	2.402×10^{-5}	2.895	9.88×10^{-3}	4.519×10^{-5}	3.020	2.412×10^{-2}
16	4.104×10^{-6}	2.952	1.113×10^{-2}	5.600×10^{-6}	3.012	2.373×10^{-2}
32	3.942×10^{-7}	2.977	1.194×10^{-2}	9.968×10^{-7}	3.007	2.337×10^{-2}

scheme #4

For each method, the numerical results show that $C_{\theta,n}$ (or $C_{w,n}$) is almost a constant number as n changes. This means our assumption about the form of the error is reasonable.

For each method, $\alpha_{\theta,n}$ (or $\alpha_{w,n}$) clearly converges to the order of the accuracy of the method. So the numerical results match our previous theorem.

When considering what method is more efficient, we can examine many aspects such as the accuracy, the stability, the computational cost and the complexity of the implementation of the method.

For our 4 R-K methods here, all of them are 2-stage method. Therefore, the computational cost and implementation of each method are very close. But the first 2 methods are explicit method with order 2 and the last 2 methods are implicit methods with order 4 and 3. So if we take the accuracy and stability into account, the 4th order implicit R-K method is more efficient than other methods.