Norwegian University of Science and Technology
Department of Mathematical Sciences

MA8404 Numerical solution of time dependent differential equations
Autumn 2019

Solutions to exercise set 1

1 **a)** We have a system of the form

$$y_1' = ay_1 + by_2$$
$$y_2' = cy_1 + dy_2.$$

Since

$$ad - bc = -100 \cdot (-100) - 1 \cdot 1 = 9999 > 0,$$
$$(a-d)^2 + 4bc = 0^2 + 4 \cdot 1 = 4 > 0,$$

the general solution is given by

$$y_1 = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t},$$
$$y_2 = C_1(\lambda_1 + 100)e^{\lambda_1 t} + C_2(\lambda_2 + 100)e^{\lambda_2 t},$$

where λ_1, λ_2 are solutions of the characteristic equation

$$\lambda^2 + 200\lambda + 100^2 - 1 = 0,$$

that is

$$\lambda_1 = -101, \quad \lambda_2 = -99.$$

Inserting this and finding from the boundary conditions that $C_1 = C_2 = \frac{1}{2}$, we get the exact solution

$$y_1(t) = \frac{1}{2}(e^{-99t} + e^{-101t}),$$
$$y_2(t) = \frac{1}{2}(e^{-99t} - e^{-101t}).$$

b)

$$\|y(0) - v(0)\| = \|(1,0) - (1, \frac{1}{10})\| = \sqrt{0^2 + (\frac{1}{10})^2} = \frac{1}{10}.$$

By the triangle inequality,

$$\|f(t,y) - f(t,v)\| \leq 100\|y(t) - v(t)\| + \|y(t) - v(t)\| = 101\|y(t) - v(t)\|,$$

and hence

$$\|y(1) - v(1)\| \leq \frac{1}{10}e^L = \frac{1}{10}e^{101},$$

which is a terrible bound.

c) We get

$$l = \mu_2(A) = \lambda_{\max}\left(\frac{A + A^T}{2}\right) = \lambda_{\max}(A) = -99.$$

Hence

$$\|y(1) - v(1)\| \leq \frac{1}{10}e^{-99} \approx 1.011 \cdot 10^{-44}.$$

From the boundary conditions, we find that

$$\begin{aligned} v_1(t) &= \frac{11}{20}e^{-99t} + \frac{9}{20}e^{-101t}, \\ v_2(t) &= \frac{11}{20}e^{-99t} - \frac{9}{20}e^{-101t}, \end{aligned}$$

and the exact value of the norm is thus

$$\begin{aligned} \|y(1) - v(1)\| &= \|(\frac{1}{20}e^{-101} - \frac{1}{20}e^{-99}, -\frac{1}{20}e^{-101} - \frac{1}{20}e^{-99})\| \\ &= \frac{1}{20}\sqrt{2(e^{-101} - e^{-99})^2} \approx 7.216 \cdot 10^{-45}. \end{aligned}$$

d) To prove (5), notice that (2) becomes

$$\langle A(y - v), y - v\rangle \leq l\|y - v\|^2, \qquad \forall y, v \in \mathbb{R}^m.$$

By using $x = y - v$, this gives

$$l = \max_{x \neq 0} \frac{x^T A x}{x^T x} = \lambda_{\max}\left(\frac{A + A^T}{2}\right)$$

using the fact that $x^T A x = x^T A^T x$, and well known results from linear algebra.

2 **a)** The eight order conditions for a 4th-order Runge-Kutta method are

(1) $\sum_i b_i = 1$	(4) $\sum_{i,j} b_i a_{ij} c_j = \frac{1}{6}$	(7) $\sum_{i,j} b_i a_{ij} c_j^2 = \frac{1}{12}$
(2) $\sum_i b_i c_i = \frac{1}{2}$	(5) $\sum_i b_i c_i^3 = \frac{1}{4}$	(8) $\sum_{i,j,k} b_i a_{ij} a_{jk} c_k = \frac{1}{24}$
(3) $\sum_i b_i c_i^2 = \frac{1}{3}$	(6) $\sum_{i,j} b_i c_i a_{ij} c_j = \frac{1}{8}$	

To make sure that the given method satisfies these conditions, we have to insert the values from the Butcher-tableau. For instance, (5) and (7) yield

$$\sum_i b_i c_i^3 = b_2 c_2^3 + b_3 c_3^3 + b_4 c_4^3 = \tfrac{1}{3}\left(\tfrac{1}{2}\right)^3 + \tfrac{1}{3}\left(\tfrac{1}{2}\right)^3 + \tfrac{1}{6} \cdot 1^3 = \frac{1}{4}$$

and

$$\sum_{i,j} b_i a_{ij} c_j^2 = b_3 a_{32} c_2^2 + b_4\left(a_{42} c_2^2 + a_{43} c_3^2\right) = \tfrac{1}{3}\tfrac{1}{2}\left(\tfrac{1}{2}\right)^2 + \tfrac{1}{6}\left(0 \cdot \left(\tfrac{1}{2}\right)^2 + 1 \cdot \left(\tfrac{1}{2}\right)^2\right) = \frac{1}{12}.$$

b) The four order conditions for explicit methods of order 3 with 3 stages can be written in matrix form:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & c_2 & c_3 \\ 0 & c_2^2 & c_3^2 \\ 0 & 0 & a_{32}c_2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{1}{2} \\ \frac{1}{3} \\ \frac{1}{6} \end{bmatrix},$$

which is a linear system in b_1, b_2 and b_3. Since b_1 is automatically given by b_2 and b_3 from the first equation, we can focus on the reduced system

$$\begin{bmatrix} c_2 & c_3 \\ c_2^2 & c_3^2 \\ 0 & a_{32}c_2 \end{bmatrix} \begin{bmatrix} b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{3} \\ \frac{1}{6} \end{bmatrix}.$$

In order for this system to have a solution, which must be unique, the right-hand side needs to be in the column space of the matrix. As such, the 3×3 matrix we obtain by adding the right-hand side as a third column must be singular. We hence require

$$\begin{vmatrix} c_2 & c_3 & \frac{1}{2} \\ c_2^2 & c_3^2 & \frac{1}{3} \\ 0 & a_{32}c_2 & \frac{1}{6} \end{vmatrix} = 0, \qquad \text{implying that} \qquad c_2\left(3\,a_{32}c_2^2 - 2\,a_{32}c_2 - c_2c_3 + c_3^2\right) = 0.$$

Note that $c_2 \neq 0$, because if not, the first column would then be exact zero and the second would not be proportional to the third. This would violate the last order condition. We therefore conclude that a necessary and sufficient condition is given by

$$3\,a_{32}c_2^2 - 2\,a_{32}c_2 - c_2c_3 + c_3^2 = 0. \tag{9}$$

c) Insertion of $a_{32} = c_3$ into (9) yields the condition

$$c_3\left[c_3 - 3c_2(1 - c_2)\right] = 0.$$

We cannot have $c_3 = 0$, since then $a_{32} = 0$, which violates the last order condition. Instead we require

$$c_3 = 3c_2(1 - c_2),$$

and so the method is characterized by the single free parameter c_2, where $c_2 = 0$ and $c_2 = 1$ must be excluded. Also, we find that

$$b_1 = \frac{c_2^3 - \frac{3}{2}c_2^2 + \frac{2}{3}c_2 - \frac{1}{9}}{c_2^2(c_2 - 1)}, \qquad b_2 = \frac{c_2 - \frac{1}{3}}{2c_2^2} \qquad \text{and} \qquad b_3 = \frac{1}{18c_2^2(1 - c_2)}.$$

d) For the left tree:

$$\sum_{i,j=1}^{s} b_i c_i^3 a_{ij} c_j^2 = \frac{1}{7 \cdot 3} = \frac{1}{21}.$$

And for the right tree:

$$\sum_{i,j,k,l=1}^{s} b_i a_{ij} c_j a_{ik} c_k a_{kl} c_l = \frac{1}{7 \cdot 2 \cdot 4 \cdot 2} = \frac{1}{112}.$$

3 By Theorem $4.4, iv)$, in *Lecture notes on ODEs*, we have

$$\begin{aligned} e(\bullet)(h) &= h, \\ e(\tau)(h) &= \int_0^h \prod_{k=1}^{\kappa} e(\tau_k)(s)\,ds. \end{aligned}$$

For the first case,

$$\rho(\bullet) = 1 \quad \Rightarrow \quad \gamma(\bullet) = \frac{h^{\rho(\bullet)}}{e(\bullet)(h)} = \frac{h^1}{h} = 1.$$

Furthermore,

$$\begin{aligned} e(\tau)(h) = \int_0^h \prod_{k=1}^{\kappa} e(\tau_k)(s)\,ds &= \int_0^h \prod_{k=1}^{\kappa} \frac{1}{\gamma(\tau_k)} s^{\rho(\tau_k)}\,ds \\ &= \prod_{k=1}^{\kappa} \frac{1}{\gamma(\tau_k)} \int_0^h s^{\sum_{k=1}^{\kappa} \rho(\tau_k)}\,ds \\ &= \prod_{k=1}^{\kappa} \frac{1}{\gamma(\tau_k)} \frac{1}{1+\sum_{k=1}^{\kappa} \rho(\tau_k)} h^{1+\sum_{k=1}^{\kappa} \rho(\tau_k)}\,ds. \end{aligned}$$

We use that

$$\rho(\tau) = 1 + \sum_{k=1}^{\kappa} \rho(\tau_k)$$

and get

$$\gamma(\tau) = h^{\rho(\tau)} / \left(\prod_{k=1}^{\kappa} \frac{1}{\gamma(\tau_k)} \frac{h^{\rho(\tau)}}{\rho(\tau)} \right) = \rho(\tau) \prod_{k=1}^{\kappa} \gamma(\tau_k).$$

4 **a)** For ODEs of the form

$$y' = y^2$$

we have $f(y) = y^2$ and

$$\begin{aligned} f'(y) &= 2y, \\ f''(y) &= 2, \\ f^{(m)}(y) &= 0, \quad m \geq 3. \end{aligned}$$

So the elementary differential $F(\tau) = 0$ for all trees with one or more nodes, the root included, from which there are 3 or more branches.

b) Given

$$y' = Ay + g(t), \qquad A \text{ constant matrix.}$$

Repeated differentiations of y' gives us:

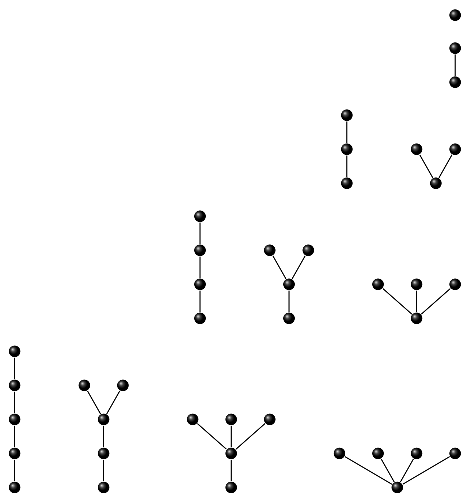
$$y' = Ay + g(t)$$

$$y'' = Ay' + g'(t)$$

$$y^{(3)} = Ay'' + g''(t),$$

$$y^{(4)} = Ay^{(3)} + g^{(3)}(t),$$

$$y^{(5)} = Ay^{(4)} + g^{(4)}(t),$$



and so on. So, going from $y^{(p)}$ to $y^{(p+1)}$, we add a new root to the trees of $y^{(p)}$, corresponding to multiplying by A, that is $f'F(\tau)$ for all τ of $y^{(p)}$, and one extra bushy tree from g. In conclusion, we see that the elementary differentials are non-zero only for tall trees with a bushy top.

5 (*Optional*) The strategy is more or less the same as the one outlined in the Lecture Note, section 4.2:

The split ODE can bw written on integral form as

$$y(t_0 + h) = y_0 + \int_0^h f(y(t_0 + s))ds + \int_0^h g(y(t_0 + s))ds. \tag{10}$$

Assume the exact solution can we written as a (so far unknown) B-series, developed around $y(t_0) = y_0$:

$$y(t_0 + h) = B(e, y_0; h) = y_0 + \sum_{\tau \in T \setminus \{\emptyset\}} \alpha(\tau) \cdot e(\tau)(h) \cdot F(\tau)(y_0) \tag{11}$$

Plug this into (10) and apply Lemma 4.1 in the note on both f and g:

$$\begin{aligned} y(t_0 + h) &= y_0 + \int_0^h f(B(e, y_0; s)ds + \int_0^h g(B(e, y_0; s)ds \\ &= y_0 + \sum_{u \in U_f} \beta(u) \cdot \int_0^h \psi_e(u)(s)ds \cdot G(u)(x_0) + \sum_{u \in U_g} \beta(u) \cdot \int_0^h \psi_e(u)(s)ds \cdot G(u)(x_0) \end{aligned} \tag{12}$$

Notice that the two sums only differ in the summation sets U_f and U_g. In Lemma 4.1, only point **b)** will be affected. We get

$$\begin{aligned} G([\emptyset]_f) &= f(y_0), & G(u)(y_0) &= f^{(\kappa)}(y_0)\big(F(\tau_1)(y_0), \cdots, F(\tau_\kappa)(y_0)\big), \\ & & &\text{for} \quad u = [\tau_1, \cdots, \tau_\kappa]_f \in U_f \\ G([\emptyset]_g) &= g(y_0), & G(u)(y_0) &= g^{(\kappa)}(y_0)\big(F(\tau_1)(y_0), \cdots, F(\tau_\kappa)(y_0)\big), \\ & & &\text{for} \quad u = [\tau_1, \cdots, \tau_\kappa]_g \in U_g. \end{aligned}$$

Compare (11) and (12), noticing that each (nonempty) tree in T has to correspond to one term in either U_f or U_g. To keep track of this, assign a black root $\bullet$ to the trees in T corresponding to terms in U_f and a white one $\circ$ to those in U_g, and call the two sets T_f and T_g (or whatever you want). Then $T = T_f \cup T_g \cup \emptyset$. As in Theorem 4.4 of the Lecture note, we can conlude that $y(t_0+h)$ can be written as a B-series, the terms is given by comparing (11) and (12), and by using Lemma 4.1.

To distinguish between the two sets, let $\tau = [\tau_1, \cdots, \tau_\kappa]_\bullet \in T_f$ be a tree with a black root, and $\tau = [\tau_1, \cdots, \tau_\kappa]_\circ \in T_g$ a tree with a white root. The subtrees can be of both kind.

We can summarize this: The exact solution of (10) can be written as a formal series of the form (11) with

The set of trees T are defined as:

$$\begin{aligned}
&\bullet \in T_f, \qquad \circ \in T_g \\
&\tau = [\tau_1, \ldots, \tau_\kappa]_\bullet \in T_f \text{ and } \tau = [\tau_1, \ldots, \tau_\kappa]_\circ \in T_g \quad \text{for all } \tau_1, \ldots, \tau_\kappa \in T_f \cup T_g. \\
&T = T_f \cup T_g \cup \emptyset
\end{aligned}$$

In the following we use $\tau = [\tau_1, \cdots, \tau_\kappa]_*$ where $* \in \{\bullet, \circ\}$.

The elementary differentials $F(\tau)$ are given by:

$$\begin{aligned}
&F(\emptyset)(y_0) = y_0, \qquad F(\bullet) = f, \qquad F(\circ) = g, \\
&\quad F(\tau) = f^{(\kappa)}\Big(F(\tau_1), \ldots, F(\tau_\kappa)\Big), \qquad \forall \tau \in T_f, \\
&\quad F(\tau) = g^{(\kappa)}\Big(F(\tau_1), \ldots, F(\tau_\kappa)\Big), \qquad \forall \tau \in T_g.
\end{aligned}$$

The elementary weight functions $e(\tau)(h)$ are:

$$\begin{aligned}
&e(\emptyset) = 1, \qquad e(\bullet) = e(\circ) = h, \\
&e(\tau) = \int_0^h \prod_{k=1}^{\kappa} e(\tau_k)(s) ds \quad \text{for all } \tau \in T.
\end{aligned}$$

And finally, the combinatorial terms $\alpha(\tau)$ are given by

$$\alpha(\emptyset) = 1, \quad \alpha(\bullet) = \alpha(\circ) = 1, \qquad \alpha(\tau) = \frac{1}{r_1! r_2! \cdots r_q!} \prod_{k=1}^{\kappa} \alpha(\tau_k),$$

where $r_1, \ldots, r_q$ counts equal trees among the subtrees $\tau_1, \ldots, \tau_\kappa$.