Math 4650 Final Exam Solutions

1. For the differential equation $y'(t) = y(t) - t^2$ with $y(0) = 2$, verify that the exact solution is $y(t) = t^2 + 2t + 2$, so that $y(1) = 5$.

   Solution: We have $y'(t) = 2t + 2$, and $y(t) - t^2 = 2t + 2$, so indeed, the differential equation is satisfied. Obviously $y(0) = 2$ as well, so the initial condition is satisfied. Since we know the differential equation is well-posed, this is the unique solution.

   (a) Use Euler's method to approximate $y(1)$ when $h = 0.5$.
   Solution: Euler's method is

   $$w_{i+1} = w_i + hf(t_i, w_i) = w_i + hw_i - ht_i^2.$$

   Therefore when $h = 0.5$ we have

   $$w_1 = (1 + 0.5)(2) - 0.5(0)^2 = 3$$

   and

   $$w_2 = (1 + 0.5)(3) - 0.5(0.5)^2 = 4.375.$$

   The error is 0.625.

   (b) Use the midpoint method to approximate $y(1)$ when $h = 1$.
   Solution: The midpoint method is

   $$w_{i+1} = w_i + hf(t_i + \tfrac{h}{2}, w_i + \tfrac{h}{2} f(t_i, w_i)) = w_i + h\big(w_i + \tfrac{h}{2}(w_i - t_i^2) - (t_i + \tfrac{h}{2})^2\big).$$

   In this case, with $h = 1$, we have

   $$w_1 = 2 + \big(2 + 0.5(2 - 0) - (0 + 0.5)^2\big) = 4.75.$$

   The error is 0.25.

   (c) Use the trapezoid method (the one-step Adams-Moulton method) to approximate $y(1)$ when $h = 1$. (Use the actual Adams-Moulton, not a predictor-corrector.)
   Solution: The general trapezoid method is

   $$w_{i+1} = w_i + \frac{h}{2}\big(f(t_i, w_i) + f(t_{i+1}, w_{i+1})\big).$$

   Plugging in the function, we have

   $$w_{i+1} = w_i + \frac{h}{2}(w_i - t_i^2 + w_{i+1} - t_{i+1}^2).$$

   Now solving for w_{i+1}, we get

   $$w_{i+1} = \frac{1}{1 - \frac{h}{2}}\big(w_i + \tfrac{h}{2}(w_i - t_i^2 - t_{i+1}^2)\big).$$

   With $h = 1$, $t_0 = 0$, $t_1 = 1$, and $w_0 = 2$, we get

   $$w_1 = \frac{1}{0.5}\big(2 + 0.5(2 - 0^2 - 1^2)\big) = \frac{2.5}{0.5} = 5.$$

   This method is exact.

2. The algorithm for Romberg integration is as follows.

```
h = b-a
R[1,1] = h/2 * (f(a)+f(b))
N = 1
for i from 2 to n do
   R[i,1] = 0
   x = a - 0.5*h
   % NOTE: N in the following loop is 2^(i-2)
   for k from 1 to N do
      x = x + h
      R[i,1] = R[i,1] + f(x)
   end do
   R[i,1] = (h*R[i,1]+R[i-1,1])/2
   M = 4
   for j from 2 to i do
      R[i,j] = R[i,j-1] + (R[i,j-1]-R[i-1,j-1])/(M-1)
      M = 4*M
   end do
   h = h/2
   N = 2*N
end do
```

Suppose each function evaluation $f(x)$ requires m multiplications/divisions.

Set up a formula that tells you the number of multiplications/divisions in total, in terms of m and n.

You do not need to evaluate or simplify this formula.

Solution: There are two multiplications outside the i-loop, as well as two function evaluations, so we'll need to add a factor of $2m + 2$ at the end.

Inside the k loop, there are m multiplications at each step (coming from the function evaluation $f(x)$), and since the k loop runs $N = 2^{i-2}$ times, there is a total of $2^{i-2}m$ multiplications in total for the k loop.

Inside the j loop, there are two multiplications at each step, and the loop is run $i - 2 + 1 = i - 1$ times. So the total through the j loop is $2(i - 1)$.

Finally there are five multiplications for each i which are outside both the j loop and the k loop.

Thus in total, the number of multiplications and divisions is

$$2m + 2 + \sum_{i=2}^{n} \left(2^{i-2}m + 2(i-1) + 5\right).$$

Although it was not asked, this formula simplifies to

$$n^2 + 4n - 3 + m(2^{n-1} + 1).$$

3. For each of the following characteristic polynomials arising from a multistep differential equation scheme, indicate whether the corresponding method is consistent or inconsistent, and whether it is stable or unstable. Explain your answers briefly.

   (a) $\lambda^3 - \frac{7}{6}\lambda^2 + \frac{1}{6} = (\lambda - 1)(\lambda - \frac{1}{2})(\lambda + \frac{1}{3})$
   Solution: The only thing we can tell about consistency from the difference equation is that if $\lambda = 1$ is not a root, then the method is certainly not consistent. In this case, $\lambda = 1$ is a root, so (assuming all else works) the method is consistent.

   The other roots are all strictly less than 1 in absolute value, so the method is strongly stable (in particular, stable).

   (b) $\lambda^3 - \frac{3}{2}\lambda^2 + \frac{1}{2} = (\lambda - 1)(\lambda - 1)(\lambda + \frac{1}{2})$
   Solution: Since $\lambda = 1$ is a root, the method is consistent. However, since $\lambda = 1$ is a double-root, the method does not satisfy the root condition and therefore is unstable.

   (c) $\lambda^3 + \frac{21}{20}\lambda^2 - \frac{1}{20} = (\lambda + 1)(\lambda - \frac{1}{5})(\lambda + \frac{1}{4})$
   Solution: Since $\lambda = 1$ is not a root, the method cannot possibly be consistent. However, if it were, the method would be stable, since all roots are simple and have magnitude at most 1.

   (d) $\lambda^3 - \frac{13}{4}\lambda^2 + \frac{9}{4} = (\lambda - 1)(\lambda - 3)(\lambda + \frac{3}{4})$
   Solution: Since $\lambda = 1$ is a root, the method is consistent. However, since $\lambda = 3$ is also a root, the method must be unstable.

4. Suppose we use the two-step difference scheme

$$w_{i+1} = -4w_i + 5w_{i-1} + 4hf(t_i, w_i) + 2hf(t_{i-1}, w_{i-1})$$

   to approximate the equation $y'(t) = f(t, y(t))$.

   (a) Compute the local truncation error of this method.
   Solution: To compute the local truncation error, we assume that $w_i = y(t_i)$ and $w_{i-1} = y(t_{i-1}) = y(t_i - h)$. This assumption tells us that $f(t_i, w_i) = y'(t_i)$ and that $f(t_{i-1}, w_{i-1}) = y'(t_{i-1}) = y'(t_i - h)$. So we have

$$\begin{aligned} w_{i+1} &= -4y(t) + 5y(t-h) + 4hy'(t) + 2hy'(t-h) \\ &= -4y(t) + 5\big[y(t) - hy'(t) + \tfrac{h^2}{2}y''(t) - \tfrac{h^3}{6}y'''(t) + \tfrac{h^4}{24}y^{(iv)}(t) + \cdots\big] \\ &\qquad + 4hy'(t) + 2h\big[y'(t) - hy''(t) + \tfrac{h^2}{2}y'''(t) - \tfrac{h^3}{6}y^{(iv)}(t) + \cdots\big] \\ &= y(t) + hy'(t) + \tfrac{h^2}{2}y''(t) + \tfrac{h^3}{6}y'''(t) - \tfrac{h^4}{8}y^{(iv)}(t) + \cdots \end{aligned}$$

   On the other hand, we have

$$y(t+h) = y(t) + hy'(t) + \tfrac{h^2}{2}y''(t) + \tfrac{h^3}{6}y'''(t) + \tfrac{h^4}{24}y^{(iv)}(t) + \cdots$$

   Therefore the local truncation error is

$$\tau_{i+1} = \frac{y(t+h) - w_{i+1}}{h} = \frac{\frac{h^4}{24}y^{(iv)}(t) + \frac{h^4}{8}y^{(iv)}(t)}{h} + \cdots = \frac{h^3}{6}y^{(iv)}(t) + \cdots = \frac{h^3}{6}y^{(iv)}(\xi).$$

(b) Is the method convergent? Explain.

Solution: The method is convergent if and only if it is both consistent and stable. The computation above proves that the method is consistent (since $\lim_{h\to 0}\tau_{i+1} = 0$). For stability, we check the characteristic polynomial, which is

$$\lambda^2 + 4\lambda - 5 = 0.$$

The roots of this are $\lambda = 1$ and $\lambda = -5$. So the method is unstable, and therefore it is not convergent.

5. Define tridiagonal $n \times n$ matrices A_n as follows:

$$A_1 = \begin{pmatrix} 3 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 3 & 2 \\ 1 & 3 \end{pmatrix}, \qquad A_3 = \begin{pmatrix} 3 & 2 & 0 \\ 1 & 3 & 2 \\ 0 & 1 & 3 \end{pmatrix}, \qquad A_4 = \begin{pmatrix} 3 & 2 & 0 & 0 \\ 1 & 3 & 2 & 0 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 1 & 3 \end{pmatrix},$$

etc.

(a) If $d_n = \det A_n$, show (for example, by expanding on the first row) that

$$d_n = 3d_{n-1} - 2d_{n-2},$$

for any $n \geq 3$.

Solution: The notation is a bit difficult in general, but let's see it for d_5:

$$\begin{vmatrix} 3 & 2 & 0 & 0 & 0 \\ 1 & 3 & 2 & 0 & 0 \\ 0 & 1 & 3 & 2 & 0 \\ 0 & 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 1 & 3 \end{vmatrix} = 3\begin{vmatrix} 3 & 2 & 0 & 0 \\ 1 & 3 & 2 & 0 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 1 & 3 \end{vmatrix} - 2\begin{vmatrix} 1 & 2 & 0 & 0 \\ 0 & 3 & 2 & 0 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 1 & 3 \end{vmatrix}$$

$$= 3\begin{vmatrix} 3 & 2 & 0 & 0 \\ 1 & 3 & 2 & 0 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 1 & 3 \end{vmatrix} - 2\begin{vmatrix} 3 & 2 & 0 \\ 1 & 3 & 2 \\ 0 & 1 & 3 \end{vmatrix}$$

It's clear that it works the same way no matter how large the matrix is, and thus we get the desired difference equation.

(b) Find a general formula for d_n for any n by solving the difference equation.

Solution: The characteristic polynomial is $\lambda^2 - 3\lambda + 2 = 0$, with solutions $\lambda = 1$ and $\lambda = 2$. So the general solution is

$$d_n = A + B2^n.$$

To find A and B, we need initial conditions. Obviously $d_1 = 3$, and it's easy to compute that $d_2 = 3 \cdot 3 - 2 \cdot 1 = 7$. Therefore we get the equations

$$d_1 = A + 2B = 3, \qquad d_2 = A + 4B = 7,$$

and solving these we find $B = 2$ and $A = -1$. So the general formula is

$$d_n = -1 + 2^{n+1}.$$

6. (a) Give the Cholesky factorization $A = LL^t$ for the positive-definite matrix

$$A = \begin{pmatrix} 4 & -2 & 6 \\ -2 & 2 & -1 \\ 6 & -1 & 17 \end{pmatrix}.$$

You may use any method to find L, but check your work. (Hint: if you do it correctly, the entries of L are all integers.)

Solution: Gaussian row reduction leads to $l_{21} = -\frac{1}{2}$ and $l_{31} = \frac{3}{2}$. Eliminating the first column, we get

$$A \sim \begin{pmatrix} 4 & -2 & 6 \\ 0 & 1 & 2 \\ 0 & 2 & 8 \end{pmatrix}.$$

Now we read off $l_{32} = 2$, and the last reduction gives

$$U = \begin{pmatrix} 4 & -2 & 6 \\ 0 & 1 & 2 \\ 0 & 0 & 4 \end{pmatrix}.$$

From this, we see that $D = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$, which means that the Cholesky matrix will be

$$L' = L\sqrt{D} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ \frac{3}{2} & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 1 & 0 \\ 3 & 2 & 2 \end{pmatrix}.$$

Finally we do a quick verification:

$$(L')(L')^t = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 1 & 0 \\ 3 & 2 & 2 \end{pmatrix} \begin{pmatrix} 2 & -1 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 4 & -2 & 6 \\ -2 & 2 & -1 \\ 6 & -1 & 17 \end{pmatrix} = A.$$

(b) Use your factorization to solve the system

$$\begin{pmatrix} 4 & -2 & 6 \\ -2 & 2 & -1 \\ 6 & -1 & 17 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -4 \\ 7 \\ 8 \end{pmatrix}.$$

(Hint: again, the solutions should all be integers.)

Solution: First solve $L'u = b$:

$$\begin{pmatrix} 2 & 0 & 0 \\ -1 & 1 & 0 \\ 3 & 2 & 2 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} -4 \\ 7 \\ 8 \end{pmatrix},$$

so that $u = -2$, $v = 5$, and $w = 2$. Next we solve $L'^t x = b$:

$$\begin{pmatrix} 2 & -1 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -2 \\ 5 \\ 2 \end{pmatrix},$$

so that $z = 1$, $y = 3$, and $x = -1$.

7. For each of the following two situations, describe by name the method you would use to solve the problem numerically, if the goal is to minimize the amount of time taken by the program.

   Be specific: if the method is adaptive, specify how the error is estimated; if not, specify the step size.

   Also explain why you expect your method to be preferable to any other.

   (a) $\frac{dy}{dt} = g(t) + g(y)$ with $y(1) = 1$, and the function g is the solution of $g(x)^{g(x)} = x$. All values of the solution must be known on the interval $[1, 3]$ to within an accuracy of 0.0001.

   Solution: The function is extremely difficult to compute (it would require Newton's method for many iterations to be accurate to within 0.0001 for each value of x plugged in). As a result, whatever method we employ should minimize the number of function evaluations. So we want to avoid Runge-Kutta, and use instead an Adams method (since they only involve one new function evaluation per step).

   Since we require a certain accuracy, we will need to use an adaptive method. The one that makes the most sense is using a predictor-corrector method, with a four-step Adams-Bashforth as predictor and three-step Adams-Moulton as corrector. We start with $h = 0.1$ (both methods are order $O(h^4)$, so this should get us close to the desired accuracy of 10^{-4}), use fourth-order Runge-Kutta to get the method started, and estimate the error at each step, cutting h in half if it's larger than the tolerance.

   (b) $\frac{dy}{dt} = y^2 - f(t)$ with $y(0) = 1$, where the function f is known to within 10^{-4}, but only at the values $f(0)$, $f(0.01)$, $f(0.02)$, etc. The solution is desired on the interval $[0, 1]$ to within the best possible accuracy.

   Solution: The function is only known for a few values of t, which means those are the only ones we can possibly plug in. Since we want the best possible accuracy, we should use the fourth-order Runge-Kutta method with fixed step size. Runge-Kutta will require evaluations at time t_i, $t_i + \frac{h}{2}$, and $t_i + h$, so all of these will have to be one of the points 0.01, 0.02, etc. Thus we choose $h = 0.02$ and use the standard fourth-order Runge-Kutta.