PHYS 325 Scientific Computing Fall 2018

Solutions Exam 2

Problem 1: General questions

a) The interpolating polynomial through 6 points has order 5. [1 point]

b) The function $f(x) = a_1 + a_2 x + a_3 x^2$ is not linear in x [1 point]
The function $f(x) = a_1 + a_2 x + a_3 x^2$ is linear in the a_i [1 point]
The function $f(x, y) = x \log(y)$ is linear in x [1 point]
The function $f(x, y) = x \log(y)$ is not linear in y [1 point]

c) $\delta f = \sqrt{(\delta x)^2 + (\delta y)^2}$ or $|\delta x| + |\delta y|$ depending on the method you use. [1 point]
$\frac{\delta f}{f} = \sqrt{\left(\frac{\delta x}{x}\right)^2 + \left(\frac{\delta y}{y}\right)^2}$ or $\left|\frac{\delta x}{x}\right| + \left|\frac{\delta y}{y}\right|$ depending on the method you use. [2 points]
$\delta f = \delta x/|x|$ [1 point]

d) The output of the program is 15, because we add $0+1+2+3+4+5$ via the loop. [1 point]

Problem 2: Function optimization

a) To find the extrema we need to calculate the derivative of the function and set it to zero:

$$f'(x) = x^3 + x^2 - 6x = 0$$

The first solution is trivial to see: $x = 0$. Factoring out x we can get the remaining solutions with the quadratic equation formula:

$$x^2 + x - 6 = 0 \Rightarrow x = \frac{1}{2}\left(-1 \pm \sqrt{1+24}\right) = \frac{-1 \pm 5}{2}$$

Hence the remaining solutions are $x = 2$ and $x = -3$. We now need to determine which ones are the minima and which one is the maximum. For this we need the second derivative:

$$f''(x) = 3x^2 + 2x - 6$$

Evaluated at the positions of the extrema: $f''(x = 0) = -6$ is negative, hence $x = 0$ is the maximum; $f''(2) = 10$ is positive, hence $x = 2$ is a minimum; $f''(-3) = 27$ is positive, hence $x = -3$ is also a minimum.

b) The value $x_0 = -1$ lies between the minimum at $x = -3$ and the maximum at $x = 0$. This means the function is sloped towards the first minimum at $x = -3$ there. To find the $x = 2$ minimum we need to start to the right of the maximum, for example at $x = 1$ or $x = 10$ etc.

c) We need any three points around the maximum, such that the middle point is higher than the two points at the boundary, for example $x_0 = -1$, $x_1 = 0$ and $x_2 = 1$.

Problem 3: Data analysis

a) Polynomial interpolation means finding a single polynomial that goes through all data points. Spline interpolation means finding piecewise polynomials (typically low order) between each pair of successive data points and then matching the pieces (function values and derivatives up to a certain degree) at these points.

b) Interpolation is used for constructing values between the original data points, while extrapolation is used for contructing values outside of the range of the original data points.

c) It might be a coincidence, or B might cause A, or A and B might have a common cause C.

d) The error is proportional to $1/\sqrt{N}$.

Problem 4: Data fitting

a)

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{y_i - f(x_i)}{\sigma_i} \right)^2$$

b) $\nu = N - M = 200 - 3 = 197$

c) Yes, χ^2_{min}/ν is close to 1.

d) No, Q is too small.

e) There might be systematic errors which average out over the whole data set. For example, if the data is systematically below the fit in certain parts of the domain and systematically above in other parts, these effects might cancel when we sum over all data points in the χ^2 formula.

f) We typically have many more data points than fit parameters. Also there is statistical noise in the data, which means many different fits may be almost equally suitable to describe the data.

g) We need to exclude singular values that are much smaller (typically about 12 orders of magnitude or more or double precision) than the highest singular value. In this case this means the last 3 values: `9.0e-12, 2.5e-22, 7.3e-23`

Problem 5: Numerical integration

a) The idea behind the trapezoidal rule is to approximate the integrand by a straight line between each pair of points x_i and x_{i+1}.

b) The integration error is of $\mathcal{O}(\Delta x^2)$.

c) Adaptive method: start with a small N and successively double N until desired accuracy is reached. There is a formula for the error, namely $\epsilon = |I_{2N} - I_N|/3$ for the trapezoidal rule. Note that when doubling N we do not need to recalculate previously evaluated x_i values, only the new ones.

d) The trapezoidal rule is very simple to code, works with equally spaced points, works even for noisy or otherwise poorly behaved functions. However, it is less efficient than other methods.

e) Gaussian quadrature is suitable for smooth, well-behaved functions. An advantage is that it is extremely accurate for such functions and it only needs a very small number of points. A disadvantage is that you need a non-uniform grid of points and it is expensive to change the number of points.

Problem 6: Numerical integration with Python

```
import numpy as np

def f(x):
    return np.sqrt(x) - x*np.log(x)

numberSteps = input('Number of steps: ')

xmax = 2
xmin = 1
binWidth = (xmax - xmin)/float(numberSteps)     # this could also be done earlier

xi = xmin

# initialize numericalInt for numerically suitable formula, otherwise set to 0
numericalInt = (f(xmax) + f(xmin))/2

# the -1 below is for the numerically suitable formula
for i in range(numberSteps-1):
    # numericalInt += ( f(xi) + f(xi + binWidth) )/2.0     # telescoping sum!!!
    xi = xi + binWidth
    numericalInt += f(xi)              # numerically suitable formula

numericalInt *= binWidth

print("Numerical value integral: ", numericalInt)
```

The errors were corrected in the code in red:

- formula for the integrand is incorrect
- `numberSteps` needs to be cast to a numeric type, either right after input, or when used in a formula
- forgot to initialize `xi`
- forgot to initialize `numericalInt`
- forgot `:` in the for loop
- telescoping sum needs to be replaced by numerically suitable formula [4 points for correct implementation]
- forgot to multiply the numerical integral by the bin width
- misspelled a variable name