Unit design: Numerical Differentiation

Motivation and what was done in the lecture

In system biology on studies processes of change in the format of dynamic systems, i.e. via differential equations. Examples of 1-dimensional systems are models of exponential growth, restricted exponential growth and logistic growth. These differential equations contain parameters whose values must be estimated from measured data.

Let's use the exponential growth model $\frac{dy}{dt} = r \cdot y$ as example. The solution is well-known and the parameter r is usually obtained for measured data by plotting $\ln(y)$ against time t, followed by linear regression. But a nice alternative is to compute the numerical derivative $y'$and plot it versus the quantity y, followed by linear regression with a straight line through the origin.

For the logistic growth model $\frac{dy}{dt} = r \cdot y \cdot (a - y)$ one can use the second method as well: compute the numerical derivative y' and plot it versus the quantity y, followed by quadratic regression.

These example may motivate student to think about computing the numerical derivative of a function when only function values are given. Numerical differentiation is a subject that is suitable for a more open inquiry approach because Dutch students are already familiar with the difference quotient as approximation of a derivative at a point in the domain of some function. Thus one might expect that they can easily come up with the forward finite difference as a numerical approximation of a derivative value.

In the last part of the lecture about differentiation we tried out the inquiry-based approach to let students discover ways to numerically approximate a slope at some point. The task given was:

> Given are the following values of a function $y(t)$ in the neighbourhood of $t = 1$:

t	0.7	0.8	1.0	1.1	1.2
y(t)	0.741	0.819	1.000	1.105	1.221

> What is the best approximation of $y'(1)$? (exact answer = 1 because $y(t) = e^{t-1}$)
> Try several methods and compare the results with each other

The data set was carefully chosen: functions values come from a known function so that students know in advance that the numeric answer should be close to 1. The five points are not at equal distance. This has to be taken into account in comparing various finite difference methods. We list some results that student could obtain;

- Forward finite difference with time step 0.1: $y'(1) \approx \frac{y(1.1)-y(1)}{0.1} = 1.050$
- Forward finite difference with time step 0.2: $y'(1) \approx \frac{y(1.2)-y(1)}{0.2} = 1.105$

- Backward finite difference with time step 0.2: $y'(1) \approx \frac{y(1)-y(0.8)}{0.2} = 0.905$
- Average value of forward and backward finite difference with time step 0.2: $y'(1) \approx \frac{1.105+0.905}{2} = 1.005$
- A waited sum of the forward finite difference with time step 0.1 and the backward finite difference with time step 0.2. As weight factor we use $\frac{1}{3}$ for the forward finite difference and $\frac{2}{3}$ for the backward finite difference: $y'(1) \approx \frac{1}{3} \cdot 1.050 + \frac{2}{3} \cdot 0.905 = 0.953$. This choice is based on the series approximation $f(a+h) - f(a-2h) = 3\,f'(a) \cdot h - \frac{3}{2} f''(a) \cdot h^2 + \cdots$ and then rewriting the quotient $\frac{f(a+h)-f(a-2h)}{3h}$ as $\frac{1}{3}\frac{f(a+h)-f(a)}{h} + \frac{2}{3}\frac{f(a)-f(a-2h)}{2h}$ (in our example $a = 1, h = 0.1$
- A much better waited sum emphasizes the values obtained with a small time step: As weight factor we use $\frac{2}{3}$ for the forward finite difference and $\frac{1}{3}$ for the backward finite difference: $y'(1) \approx \frac{2}{3} \cdot 1.050 + \frac{1}{3} \cdot 0.905 = 1.002$.
- Determine the algebraic expression of the quadratic function whose graph goes through the points (0.8, 0.819), (1,1) and (1.1, 1.105) and use the formula to compute the derivative, which can them be uses to find the derivative at domain value 1. One obtains the value 1.002 because it is in essence the same method as the previous waited sum (this is easily seen from the Lagrange interpolation method).

The lecture was recorded and reflection notes were written shortly after the lecture:

> The task to bring up ideas for estimating the derivative at a point, based on only five data points, functioned above expectation. We thought after 7 hours of lectures and tutorials, students would not be in the mood of thinking and exploration. However, the question seemed to intrigue students, first of all because it puzzled them a bit that they were only allowed to use five data points and no formula (maybe the explanation on how the data had been created, raised some confusion).
>
> After six minutes (due to lack of time and new ideas didn't pop up anymore) we asked for suggested methods:
>
> - The first method mentioned was actually forward difference: this we expected from the explanations in the first part of the lecture.
> - The second suggested method was to take the mean value of two consecutive difference formulas that involves the point of interest. In fact, this is the average of a forward and backward difference. Here, the fact that the distribution of data was not at constant sampling rate was ignored by students or at least taken for granted.
> - The second suggestion provided a starting point in discussing backward difference and the influence of the size of a time step.
> - The discussions led to the suggestion to average more differences, which is a good idea when there is a lot of noise on the data and with more data points available.

- We did not discuss the possible use of weights in the second method to compensate for differences in time steps, because it only came into the lecturer's mind afterwards and we would not have time to discuss it into depth anyway

As mentioned before, the students participated quite well in class despite the late hour. Unfortunately, due to the arrangement of the tables in the lecture room and a lack of time, it was a challenge to have in depth conversations with the students regarding the numerical derivative.

The design of student activities for the tutorial / practice session

Being not sure whether students already y had become rather proficient with basic skills for using R to do computations to write their own R scripts from scratch, we chose to guide students first in plotting mathematical functions. We let them go through the steps required to plot a mathematical function and to combine graphs in one diagram and use scaffolding hints to help them further.

The heart of the students practice session is the implementation of the 2-point forward difference method and the 3-point difference method for computing a numerical derivative from data constructed by students themselves. The goals, i.e. the kind of picture students should create, is clearly set and, if needed, short hints can put students on the right track. This activity is already more open in the sense that students are free how to implement the methods. Students are asked whether to notice difference between the results obtained by the two numerical methods.

The final task is to add nose to the data set constructed earlier and explore how it affect the quality of results for the two numerical methods. Students are asked to investigate how the size of the step size in the numerical method and the strength of the noise affect to result of both method. This kind of student work can be characterized as a combination of guided and structured inquiry: student must determine how to investigate the problem and conclusions are solely based on students' investigations, although as authors and instructors we know in what kind of direction the answers should be. One week later we make our outcomes to the tasks visible in the SOWISO environment for comparison by students with their own results.