Scientific Computation
Autumn 2023
Project 3

Due: Friday December 15th, 1pm

In addition to this project description, there are 3 files for this assignment:

- project3.py: a Python file which you will complete and submit on Blackboard (see below for details)
- report3.tex: a Latex template file for the short report which you will submit. The discussion and figure(s) described below should be placed in this report.
- data1.npz: a data file needed for part 1

Part 1

(6 pts) You have been provided with the data file *data1.npz* which contains atmospheric wind speed data for a band of latitudes in the southern hemisphere for an entire year. Code is provided in the function, *part1*, to load the entire dataset and display one day of data. This code creates a 365 x 16 x 144 numpy array, u, where $u[i, j, k]$ corresponds to the daily average zonal wind speed on day i at latitude, $lat[j]$, and longitude, $lon[k]$, where lat and lon are arrays that are also loaded. The zonal wind speed is the component of the wind velocity parallel to a line of latitude (positive zonal wind corresponds to eastward air flow). Analyze this wind data with a focus on fluctuations in space and time. As one part of your analysis, you should apply PCA and use the results to discuss meaningful trends associated with large variations in u with respect to time. You should use your judgement to decide how to apply PCA to this data. Generally, your analysis should develop insights into nontrivial trends using quantitative results obtained from appropriately designed computations. Examples of *trivial* trends would be those that can be identified simply by plotting the data (e.g. it was relatively windy at location x at time t). You are not expected to assess if or to what extent the data is chaotic. You should present your analysis along with supporting figures in your report, and the accompanying code should be placed in the function, *part1*.
Note: You are not expected to use/acquire knowledge of atmospheric science for this question.

Part 2

Often when working with atmospheric data, it is useful to interpolate data to locations where data is unavailable (e.g. at longitudes not included in the data for part 1). However, interpolating on a spherical surface introduces unwanted complications, so we will instead consider data on a two-dimensional Cartesian grid with the grid defined as follows:

$$x_j = j\Delta x,\ j = 0, 1, 2, ..., n-1,\ \Delta x = \frac{1}{n-1} \tag{1}$$

$$y_i = i\Delta y,\ i = 0, 1, 2, ..., m-1,\ \Delta y = \frac{1}{m-1} \tag{2}$$

$$\tag{3}$$

and the data at location (x_j, y_i) is $f_{i,j}$. In this question you will investigate the effectiveness of two methods for using interpolation to estimate $f_{i,j}$ at:

$$\tilde{y}_i = i\Delta y,\ i = 1/2, 3/2, 5/2, ..., m-3/2.$$

The interpolated data at $(x_j, \tilde{y}_i)$ will be denoted as $\tilde{f}_{i,j}$. Method 1 is relatively simple and can be viewed as averaging the given data:

$$\tilde{f}_{i,j} = \frac{1}{2}\left(f_{i-1/2,j} + f_{i+1/2,j}\right),\ i = 1/2, 3/2, 5/2, ..., m-3/2$$

Method 2 can be viewed as an *implicit* scheme:

$$\alpha\tilde{f}_{i-1,j} + \tilde{f}_{i,j} + \alpha\tilde{f}_{i+1,j} = \left[\frac{b}{2}\left(f_{i-3/2,j} + f_{i+3/2,j}\right) + \frac{a}{2}\left(f_{i+1/2,j} + f_{i-1/2,j}\right)\right]$$

$$\tilde{f}_{1/2,j} = \left[\tilde{a}f_{0,j} + \tilde{b}f_{1,j} + \tilde{c}f_{2,j} + \tilde{d}f_{3,j}\right]$$

$$\tilde{f}_{m-3/2,j} = \left[\tilde{a}f_{m-1,j} + \tilde{b}f_{m-2,j} + \tilde{c}f_{m-3,j} + \tilde{d}f_{m-4,j}\right]$$

$$i \in \{1/2, 3/2, 5/2, ..., m-3/2\},$$

The first equation is used for all i except $i = 1/2$ and $i = m - 3/2$.

1. (2 pts) Complete *part2* so that it efficiently implements method 2 to compute $\tilde{f}$ for the two-dimensional m x n array provided as input. The jth column of the input array should contain the data, $f(x_j, y_i)$, $i = 0, 1, 2, ..., m-1$, and this data should be interpolated to $\tilde{y}_i$. The coefficients, α, a, b, $\tilde{a}$, $\tilde{b}$, $\tilde{c}$, $\tilde{d}$, have been included in *part2*. Note that code for method 1 has been provided in the function.

2. (4 pts) Analyze method 2, critically compare it to method 1, and assess the effectiveness of the two methods for multiscale problems. Assume that m is the same order-of-magnitude as n, and include a discussion of results from carefully-designed numerical tests which illustrate how well the methods do or do not work. Ultimately, your analysis should reflect a clear understanding of the cost, accuracy, and efficiency of method 2 and your implementation. Add your analysis and accompanying figures to your report. Place the code used for the analysis in *part2_analyze*.

Part 3

The function *part3q1* computes numerical solutions to the following system of $2n$ ODEs:

$$\frac{du_i}{dt} = \alpha u_i + 2\beta b v_i - r_i^2\left(u_i - c v_i\right) + \beta\left(u_{i-1} + u_{i+1}\right) - \beta b\left(v_{i-1} + v_{i+1}\right), \tag{4}$$

$$\frac{dv_i}{dt} = \alpha v_i - 2\beta b u_i - r_i^2\left(v_i + c u_i\right) + \beta\left(v_{i-1} + v_{i+1}\right) + \beta b\left(u_{i-1} + u_{i+1}\right), \tag{5}$$

$$r_i^2 = u_i^2 + v_i^2,\ i = 0, 1, 2, ..., n-1 \tag{6}$$

Here, α, β, b, and c are model parameters, and if $b = c = 0$, the model from project 2 is recovered. Note that when $i = 0$ in the equations above, $i - 1$ should be replaced with $n - 1$, and when $i = n - 1$, $i + 1$ should be replaced with 0.

1. (6 pts) The parameters, n, α, β, and b along with the initial condition have been set for you in *part3_analyze*. Carefully explore the dynamics that can be generated for $0.5 \leq c \leq 1.5$. Depending on your computer and the details of the computation, simulations may take a few minutes to run, so running very large numbers of simulations for a very large number of values of c may be difficult and should be avoided. After some initial exploration, you should make a sensible decision about which values of c to present results for and how to present those results. Note that simulations typically contain an initial transient as the system responds to the initial conditions followed by a relatively settled dynamical state, and the provided code has been set up to help you discard these transients. Additionally, you should only use results for $100 \leq i \leq n-100$ in your analysis, and it is sufficient to only use results for u. Your analysis should focus on the global qualitative dynamics (e.g. the system is steady (no time-dependence), simple sinusoidal oscillations in time, ...) and comparing dynamics for different values of c. You do not need to provide a detailed analysis of the dynamics for several values of c. You *should* provide a detailed analysis of the dynamical state observed when $c = 1.3$. Qualitative observations should be supported by quantitative results and well-designed figures. Add the code used in your analysis to the function, *part3_analyze*, and add the discussion of your findings along with supporting figures to your report. While it is fine to re-use some of the tools/ideas used in part 1, the questions that you are answering here with your analysis should be substantially different, and there should also be some substantive differences in the way tools/ideas from part 1 are applied here if they are re-used.
   Note: You are not required to use the default values that have been set for function input variables unless the value is explicitly required by the question.

2. (2 pts) The function *part3q2* computes a solution for this system and then four lines of code are used to create a numpy array using the computed solution. Provide a clear and concise explanation of 1) what these four lines of code are doing (i.e. what is the problem being solved, and what is the strategy used to solve it?), and whether or not the four lines of code should be considered to be efficient.

Further guidance

- You should submit both your completed python file (*project3.py*),a pdf containing your discussion and figure(s) (*report3.pdf*). You are not required to use the provided latex template, any well-organized pdf is fine. To submit your assignment, go to the module Blackboard page and click on"Project 3". There will be an option to attach your files to your submission. (these should be named *project3.py* and *report3.pdf*). After attaching the files, submit your assignment.

- Please do not modify the input/output of the provided functions without permission. You may use numpy, scipy, matplotlib, networkx, time, timeit, collections, and heapq as needed. Otherwise, please do not import any modules without permission. You may create additional functions as needed, and you may use any code that I have provided during the term. Some functions have input variables with specified default values. You are not required to use these values unless the relevant question explicitly states that you are.

- Marking will be based on the correctness of your work and the degree to which your submission reflects a good understanding of the material covered up through lecture 17. Excluding figures, you should aim to keep the pdf version of your report to less than 6 pages of text.
- Open-ended questions require sensible time-management on your part. Do not spend so much time on this assignment that it interferes substantially with your other modules. If you are concerned that your approach to the assignment may require an excessive amount of time, please get in touch with the instructor.
- Questions on the assignment should be asked in private settings. This can be a "private" question on Ed (which is distinct from "anonymous") or by arrangement with the instructor.
- Please regularly backup your work. For example, you could keep an updated copy of your files on OneDrive and use Overleaf for your report.
- In order to assign partial credit, we need to understand what your code is doing, so please add comments to the code to help us.
- You have been asked to submit code in Python functions, but it may be helpful to initially develop code outside of functions so that you can easily check the values of variables in a Python terminal.
- The weightings for assignments can be found in the lecture 1 slides.
- Your submission, including codes you develop (aside from codes I have provided during the term), should be your own work.