Numerical Analysis

Final task

Submission date: 22/2/22 23:59 (strict).

This task is individual. No collaboration is allowed. Plagiarism will not be tolerated.

The programming language for this task is Python 3.7. You can use standard libraries coming with Anaconda distribution. In particular limited use of numpy and pytorch is allowed and highly encouraged.

You should not use those parts of the libraries that implement numerical methods taught in this course (unless explicitly stated otherwise in the instructions of the particular assignment)**.** This restriction includes, for example, finding roots and intersections of functions, interpolation, integration, matrix decomposition, eigenvectors, solving linear systems, etc.

The use of the following methods in the submitted code must be clearly announced in the beginning of the explanation of each assignment where it is used and will result in deduction of points. Failure to announce the use of any restricted functions will result in disqualification of the assignment.

numpy.linalg.solve (15% of the assignment score)

(not studied in class) numpy.linalg.cholesky, torch.cholesky, linalg.qr, torch.qr (10% of the assignment score)

numpy.*.polyfit, numpy.*.*fit (40% of the assignment score)

numpy.*.interpolate, torch.*.interpolate (60% of the assignment score)

numpy.*.roots (30% of the assignment 2 score and 15% of the assignment 3 score)

~~TBD – restriction of~~ numeric differentiation functions are allowed!

numpy.linalg.inv, scipy.linalg.inv, torch.inverse,

and all other external libraries for matrix inversion (20% of the assignment score)

Additional functions and penalties may be allowed according to requests in the task forum.

You must not use reflection (self-modifying or self-inspecting code).

Attached are mockups of for 4 assignments where you need to add your code implementing the relevant functions. You can add classes and auxiliary methods as needed. Unittests found within the assignment files must pass before submission. BUT! existing unit tests are provided for demonstration and to encourage you to write additional tests as you go. You can add any number of additional unittests to ensure correctness of your implementation. Passing only the existing unittests does not ensure that your code will not fail in all cases. It is your responsibility to test your code and ensure that it is stable. You should add additional unittests to ensure correctness of your implementation.

In addition, attached are two supplementary python modules. You can use them but you cannot change them.

Formatiert: Schriftart: (Standard) +Überschriften (Times New Roman), 9 Pt.

Formatiert: Schriftart: (Standard) +Überschriften (Times New Roman), 9 Pt.

Formatiert: Schriftart: (Standard) +Überschriften (Times New Roman), 9 Pt.

Formatiert: Schriftart: Fett

Upon the completion of the final task, you should submit the five assignment files and this document with answers to the theoretical questions. The archive should not contain folders, but only the submission files! ~~TBD specific submission instructions.~~

Assignments will be graded according to **error** of the numerical solutions and **running time**. Some assignments have required specific error bounds – they will be graded according to running time. Some assignments limit the running time – they will be graded according to error. For all executions there is 2 minutes running time cap after which the execution will be halted.

Every assignment will be AUTOMATICALLY tested on a number of different functions and different parameters. It may be executed multiple times on the same function with the same parameters. Every execution will start with a clean memory. Any exception throwed during an execution will render the execution invalid and nullify its contribution to the grade. **Test your code!!!**

Any disqualification of an assignment (e.g. due to unannounced use of restricted functions) or an execution (e.g. due to exception) will not contribute to the grade regardless the effort put in the development.

Expect that the assignment will be tested on various combinations of the arguments including function, ranges, target errors, and target time. We advise to use the functions listed below as test cases and benchmarks – add additional unittests with implementations of these functions. At least half of the test functions will be polynomials. Functions 3,8,10,11 will account for at most 5% of the test cases. All test functions are continuous in the given range. If no range is given the function is continuous in $[-\infty, +\infty]$.

1. $f_1(x) = 5$
2. $f_2(x) = x^2 - 3x + 5$
3. $f_3(x) = \sin(x^2)$
4. $f_4(x) = e^{-2x^2}$
5. $f_5(x) = \arctan(x)$
6. $f_6(x) = \frac{\sin(x)}{x}$
7. $f_7(x) = \frac{1}{\ln(x)}$
8. $f_8(x) = e^{e^x}$
9. $f_9(x) = \ln(\ln(x))$
10. $f_{10}(x) = \sin(\ln(x))$
11. $f_{11}(x) = 2^{\frac{1}{x^2}} * \sin\left(\frac{1}{x}\right)$
12. For Assignment 4 see sampleFunction.*

Assignment 1 (~~15pt~~14pt):

(10pt) Implement the function **Assignment1.interpolate(..)** following the pydoc instructions.

The function will receive a function f, a range, and a number of points to use.

The function will return another "interpolated" function g. During testing, g will be called with various floats x to test for the interpolation errors.

Grading policy:

Running time complexity > O(n^2): 0-20%

Running time complexity = O(n^2): 20-80%

Running time complexity = O(n): 50-100%

Running time complexity will be measured empirically as a function of n.

The grade within the above ranges is a function of the average relative error of the interpolation function at random test points. Correctly implemented linear splines will give you 50% of the assignment value.

Solutions will be tested with $n \in \{1,10,20,50,100,200,500,1000\}$ on variety of functions at least half of which are polynomials of various degrees with coefficients ranging in $[-1,1]$.

Restricted functions I used:

Formatiert: Schriftart: Fett

(~~5pt~~4pt) Question 1.1: Explain the key points in your implementation.

Formatiert: Rechts

Formatierte Tabelle

כדי למצוא את פונקציית האינטרפולציה השתמשתי בעקומות ביזייר. השתמשתי בביזייר מכיוון שהוא מתקרב בצורה טובה יותר מאשר פולינום.

תחילה מצאתי n נקודות בטווח שבין a ל-b ומצאתי את ה-control points בעזרת האלגוריתם של תומאס.

השתמשתי באלגוריתם של תומאס מכיוון שהוא מוצא את ה-control points בצורה יעילה וחוסך פתירה של מטריצה באופן מלא – המטריצה שמשתמשים בה היא כולה אפסים מלבד 3 אלכסונים, מה שמאפשר לנו לפתור אותה בצורה יעילה.

לאחר מכן הגדרתי פונקציה שמקבלת x ובדקתי בין אילו שתי נקודות מ-n הנקודות ה-x נמצא. מצאתי את t בעזרת מציאת המיקום היחסי על הישר המתאים. השתמשתי ב-control points המתאימות ובעזרת הנוסחה מצאתי את עקומת ביזייר.

Formatiert: Rechts

Formatiert: Rechts

Assignment 2 (~~15pt~~14pt):

(10pt) Implement the function **Assignment2.intersections(..)** following the pydoc instructions.

The function will receive 2 functions- f_1, f_2, and a float maxerr.

The function will return an iterable of approximate intersection Xs, such that:

$$\forall x \in X, |f_1(x) - f_2(x)| < maxerr$$

$$\forall x_i x_j \in X, |x_i - x_j| > maxerr$$

Grading policy: The grade will be affected by the number of correct and incorrect intersection points found, the running time of **itr = Assignment2.intersections(..)** followed by **list(itr).**

Restricted functions I used:

(~~5pt~~4pt) Question 2.1: Explain the key points in your implementation in particular explain how did you address the problem of finding multiple roots.

כדי למצוא את נקודות החיתוך של שתי הפוקנציות חיסרתי בין הפונקציות ומצאתי את השורשים של פונקציית ההפרש בעזרת השיטה של ניוטון.
ביצעתי את השיטה של ניוטון בעזרת ניחושים החל מ-a ועד b עם קפיצות של (b-a)/100 מכיוון שמצאתי שכך מקבלים את היחס הכי טוב בין השגיאה לזמן הריצה.

Formatierte Tabelle

Formatiert: Rechts

Formatiert: Rechts

Assignment 3 (~~35pt~~31pt):

Implement a function **Assignment3.integrate(…)** and **Assignment3.areabetween(..)** following the pydoc instructions and answer two theoretical questions.

(5pt) Assignment3.integrate(…) receives a function f, a range, and a number of points n.

It must return approximation to the integral of the function f in the given range.

You may call f at most n times.

Grading policy: The grade is affected by the integration error only, provided reasonable running time e.g., no more than 2 minutes for n=100.

Restricted functions I used:

(~~5pt~~4pt) Question 3.1: Explain the key points in your implementation of Assignment3.integrate(…).

כדי להעריך את האינטגרל הנוצר ע"י הפונקציה השתמשתי בשיטת סימפסון. לקחתי n נקודות מהפונקציה בטווח הנתון והצבתי בנוסחה של סימפסון.

Formatiert

(10pt) Assignment3.areabetween(..) receives two functions f_1, f_2 .

It must return the area between f_1, f_2 .

In order to correctly solve this assignment you will have to find all intersection points between the two functions. You may ignore all intersection points outside the range $x \in [1,100]$.

Note: there is no such thing as negative "area".

Grading policy: The assignment will be graded according to the integration error and running time.

Restricted functions I used:

(~~5pt~~4pt) **Question 3.2:** Explain the key points in your implementation of Assignment3. areabetween (…).

כדי למצוא את השטח בין שתי הפונקציות תחילה מצאתי את נקודות החיתוך של הפוקנציות בעזרת שאלה 2.
לאחר מכן סכמתי את השטחים בין כל שתי נקודות חיתוך של הפונקציות. מצאתי את השטחים בעזרת סעיף א - חישבתי את האינטגרל של הפרש הפונקציות בין כל שתי נקודות חיתוך.

(~~4~~5pt) **Question 3.3:** Explain why is the function $2^{\frac{1}{x^2}} * \sin\left(\frac{1}{x}\right)$ is difficult for numeric integration with equally spaced points?

האינטגרל של הפונקציה $2^{\frac{1}{x^2}} * \sin\left(\frac{1}{x}\right)$ קשה לחישוב מכיוון שסביב 0 ישנן תנודות גדולות של הפונקציה – עבור xים קרובים הyים מאוד רחוקים. השיטות שלמדנו לחישוב אינטגרל מסתמכות על כך שעבור xים קרובים ה-yים יחסית קרובים גם הם.
הצעה לחישוב האינטגרל יכלה להיות לקחת מספר רב מאוד של נקודות על הפונקציה אבל כאשר המספרים כל כך קטנים המחשב לא יודע לחשב במדוייק את החישובים.

(~~5pt~~4pt) **Question 3.4:** What is the maximal integration error of the $2^{\frac{1}{x^2}} * \sin\left(\frac{1}{x}\right)$ in the range [0.1, 10]? Explain.

הנוסחה לחישוב טעות מקסימלית בחישוב אינטגרל היא: $e(f) = -\frac{h^2}{12} * (f'(b) - f'(a))$
נחשב לפי הפונקציה הנתונה:

$$h = \frac{b-a}{n} = \frac{10-0.1}{n}$$

$$f'(10) = -0.01, f'(0.1) = 1.06 * 10^{33}$$

$$e(f) = -\frac{n^2}{12} * (-0.01 - 1.06 * 10^{33})$$

ניתן לראות כי על מנת למקסם את ערך השגיאה הערך של *n* יהיה 1.
לכן השגיאה המקסימלית היא $e(f) = 8.6575 * 10^{33}$

Formatiert

Formatiert

Formatiert: Schriftart: Nicht Fett, Kursiv

Formatiert

Formatiert: Schriftart: Nicht Fett, Kursiv

Assignment 4 (~~15pt~~14pt)

(10pt) Implement the function **Assignment4.fit(…)** following the pydoc instructions.

The function will receive an input function that returns noisy results. The noise is normally distributed.

Assignment4.fit should return a function g fitting the data sampled from the noisy function. Use least squares fitting such that g will exactly match the clean (not noisy) version of the given function.

To aid in the fitting process the arguments a and b signify the range of the sampling. The argument d is the expected degree of a polynomial that would match the clean (not noisy) version of the given function.

You have no constrains on the number of invocations of the noisy function but the maximal running time is limited. Invocation of f may take some time but will never take longer than 0.5 sec.

Additional parameter to **Assignment4.fit** is maxtime representing the maximum allowed runtime of the function, if the function will execute more than the given amount of time, the execution will not contribute to the grade causing significant deduction. You should consider the risk of failure vs gains in accuracy when you get close to the time limit.

Grading policy: the grade is affected by the error between g (that you return) and the clean (not noisy) version of the given function, much like in Assignment1. 60% of the test cases for grading will be polynomials with degree up to 3, with the correct degree specified by d. 30% will be polynomials of degrees 4-12, with the correct degree specified by d. 10% will be non-polynomials with random $d \in [1..12]$.

Restricted functions I used:

(~~5pt~~4pt) Question 4.1: Explain the key points in your implementation.

כדי לבנות את הפונקציה דגמתי נקודות מהפונקציה הרועשת ובניתי את פונקציית השגיאה בעזרת שיטת mean square error. גזרתי את הפונקציה לפי כל אחד מהנעלמים והשוותי ל-0 על מנת למצוא את השגיאה המינימלית. שמתי לב שעבור כל d ניתן להבין בקלות מהם מקדמי המטריצה של מערכת המשוואות הנ"ל. יצרתי את המטריצה המתאימה ופתרתי אותה בעזרת השיטה lu-decomposiotion.
שיטה זו מסתמכת על כך שהמטריצה ריבועית (ישנם d+1 נעלמים וכל פעם גוזרים לפי נעלם אחר לכן d+1 משוואות). היא מחלקת את המטריצה למטריצה עליונה ומטריצה תחתונה וכך פותרת את המטריצה ביעילות.

Formatiert

Formatiert: Schriftart: Kursiv

Assignment 5 (~~30pt~~27pt).

(~~10pt~~9pt) Implement the function **Assignment5.area(…)** following the pydoc instructions.

The function will receive a shape contour and should return the approximate area of the shape. Contour can be sampled by calling with the desired number of points on the contour as an argument. The points are roughly equally spaced.

Naturally, the more points you request from the contour the more accurately you can compute the area. Your error will converge to zero for large n. You can assume that 10,000 points are sufficient to precisely compute the shape area. Your challenge is stopping earlier than that according to the desired error in order to save running time.

Grading policy: the grade is affected by your running time.

Restricted functions I used:

(~~5pt~~4pt) Question 5.1: Explain the key points in your implementation.

כדי לחשב את השטח דגמתי נקודות מהצורה בעזרת ה-contour. חישבתי את השטח בחלקים – עבור כל שתי נקודות xi, xi+1 מהקונטור (x1 , x2… xn) כך ש-xi < xi+1 הוספתי את שטח הטרפז הנוצר ע"י הנקודות עם ציר ה-x , ועבור כל שתי נקודות xi , xi+1 מהקונטור כך ש-xi > xi+1 חיסרתי את שטח הטרפז הנ"ל.

Formatiert

(10pt) Implement the function **Assignment5.fit_shape(…)** and the class **MyShape** following the pydoc instructions.

The function will receive a generator (a function that when called), will return a point (tuple) (x,y), a that is close to the shape contour.

Assume the sampling method might be noisy- meaning there might be errors in the sampling.

The function should return an object which extends **AbstractShape**
When calling the function **AbstractShape.contour(n)**, the return value should be array of n equally spaced points (tuples of x,y). When calling the function **AbstractShape.area()**, the return value should be the area of the shape. You may use your solution to **Assignment5.area** to implement the area function.

Additional parameter to **Assignment5.fit_shape** is maxtime representing the maximum allowed runtime of the function, if the function will execute more than the given amount of time the execution will be halted.

In this assignment only, you may use any numeric optimization libraries and tools. Reflection is not allowed.

Grading policy: the grade is affected by the error of the area function of the shape returned by Assignment4.fit_shape.

There are no restricted functions. The use of any library is allowed.

(~~5pt~~4pt) Question 4B.2: Explain the key points in your implementation.

כדי למצוא את שטח הצורה, ראשית דגמתי נקודות רועשות. כדי לדעת את הסדר שבו הנקודות נמצאות על הצורה מצאתי את הנקודה הממוצעת של הנקודות הרועשות, בדקתי מה הזווית של כל אחת מהנקודות הרועשות מהנקודה הממוצעת ומיינתי את הנקודות לפי גודל הזוית. אח"כ החלקתי את הצורה בעזרת הפונקציה savgol_filter ולבסוף חישבתי את השטח בעזרת הנקודות שחזרו מהפונקציה בצורה דומה לחישוב של סעיף א.

Formatiert: Schriftart: Fett

Formatiert

Formatiert