

You are to write a report on two optimization algorithms: Nesterov's accelerated gradient method and mirror descent (both are covered in the Wright and Recht textbook). The TAs and the instructor will *not* provide help on the project. You have to find appropriate sources on the Internet and write a report, providing precise proofs of convergence. You have to cite the resources used. You can use ChatGPT or other AI assistants to find resources or to learn about the material or to do coding, but you will be responsible for any errors produced by ChatGPT. The report has to be typewritten by you using your own words (not verbatim from any source), it should not be more than 15 pages long (but can be shorter if you are able to convey everything that is needed in fewer than 15 pages) and must be in 11 pt. font or bigger. The report should be submitted as a pdf file via gradescope. Details on how to submit the final report via gradescope will be provided later. Your project report should include the following:

- The report should contain an introduction section, a section on accelerated gradient descent, a section on mirror descent, a section on numerical examples, and a conclusion section.
- The introduction should state what the project is about and what the reader can learn by reading the report. The conclusion should contain a concise summary of the findings in the report.
- The sections on accelerated gradient descent and mirror descent should contain a clear description of the corresponding algorithm, a proof of convergence, and a comparison of the theoretical rate of convergence of the algorithm with the theoretical rate of convergence of standard gradient descent or standard projected gradient descent as appropriate. The section on mirror descent should also clearly show how standard projected gradient descent is a special case of mirror descent.
- The section on numerical examples should contain two plots. One plot should provide a comparison of the numerical performance of accelerated gradient descent with that of standard gradient descent, showing that accelerated gradient descent is clearly better. The other plot should provide a comparison of the numerical performance of mirror descent with that of standard projected gradient descent, showing that mirror descent is clearly better. In each plot, the x -axis should be the number of iterations and the y -axis should be value of the objective function that is being minimized. Each of these plots can be for a different optimization problem; you have to make up your own optimization problems guided by the theory, since the theory will tell you the type of optimization problems for which each algorithm is better than standard gradient descent. You have to include the code used to generate the numerical examples, you can use any programming language for this and any built-in libraries for the purpose, but you are responsible for the accuracy of the code. The code should be included in your fifteen pages.