TRINITY COLLEGE DUBLIN

School of Computer Science and Statistics

Final Assignment 2021-22 CS7CS2 Optimisation for Machine Learning

- **Unlike the rest of the module coursework you must do this assignment entirely yourself - you must not discuss or collaborate on the assignment with other students in any way, you must write answers in your own words and write code entirely yourself. If you use any online or other external content in your report you should take care to cite the source. It is mandatory to complete the declaration that the work is entirely your own and you have not collaborated with anyone - the declaration form is available on Blackboard. All submissions will be checked for plagiarism.**
- **Reports must be typed and submitted as a separate pdf on Blackboard (not as part of a zip file).**
- **Include the source of code written for the assignment as an appendix in your submitted pdf report (the code itself, not a screenshot, so the plagiarism checker can run on it). Also include a separate zip file containing the executable code and any data files needed. Programs should be running code written in Python. Keep code brief and clean with meaningful variable names etc.**
- **Important: Your primary aim is to articulate that you understand what you're doing - not just running a program and quoting numbers it outputs. Generally most of the credit is given for the explanation/analysis as opposed to the code/numerical answer.**
- **Reports should typically be about 5 pages, with 10 pages the upper limit (excluding appendix with code).**

Assignment

Write a short report comparing the performance of SGD with (i) Adam and (ii) a constant step size. To do this you'll need to make a number of important choices:

- How to measure performance. You might, for example, plot the ML loss function vs optimisation iterations and use the lowest value as a performance measure. But this measures the ML model performance on the training data and so not the generalisation performance on unseen data. You might also measure the ML loss function of the trained model on held-out test data, which would measure generalisation performance. Probably its a good idea to look at both measures.
- Remember that SGD involves randomisation and so it may be necessary to collect data from several runs to understand how performance fluctuates from run to run.
- What hyperparameters (SGD mini-batch size, Adam α, β_1, β_2 values, constant step size α) to use and how to choose them. Probably its a good idea to look at the performance both when using default hyperparameter values and also when using optimised values.
- What ML model and data to use for the evaluation. Its probably worth considering two models/datasets since the optimisation algorithm performance may change with different models/data, and be sure include at least one neural net ML model. Note

that you shouldn't choose models/datasets which are too large since training will take too long, and so you need to use some judgement here (your choice will partly depend on what compute power you have available). Its a good idea to use standard datasets e.g. MNIST, CIFAR, Imdb.

- Use one of the existing ML libraries e.g keras/tensorflow or pytorch. Don't implement Adam etc yourself.

To see some existing examples of performance evaluation you might find it helpful to read the following two papers:

- Adam: A Method For Stochastic Optimization, `https://arxiv.org/pdf/1412.6980.pdf`
- The Marginal Value of Adaptive Gradient Methods in Machine Learning, `https://arxiv.org/pdf/1705.08292.pdf`

The first introduces the Adam optimisation algorithm and compares its performance (in terms of training error) against a range of other algorithms, see Figures 1-3. The second compares the performance (in terms of test error i.e. generalisation) of SGD against a range of algorithms including Adam, see Figures 1-2. Some questions you might reflect on are: do these papers clearly address all the choices noted above? And if not, might that be important or not?
100 marks: indicative breakdown (i) experiment design 30 marks, (ii) methodology 30 marks, (iii) evaluation and critical discussion 30 marks, (iv) report organisation and presentation 10 marks