**LabB Exercises 2019-20**

These are not the “assignments”. However, they should be submitted as well, following the

procedure agreed during the first lectures.

**1) Gradient descent**

Consider notebook NB2 by Mehta et al.

1. Add the ADAmax algorithm (find its definition outside the review by Mehta)
2. Show a quantitative statistical comparison of the performance of different algorithms:
   1. Gradient descent
   2. Gradient descent with momentum
   3. Nesterov (NAG)
   4. RMSprop
   5. ADAM
   6. ADAmax
   7. eventually any other method one is curious to test…

With a given function (either Beale’s function or another function), choose a set of initial points within a given domain where it has a minimum. The set could be a regular grid of points or a random set. Starting from these points, plot the average value of the function vs time during minimization for each method (with a good value of its own learning rate, chosen after some test). Eventually (a) plot also the standard deviation around the average value; (b) plot data vs real CPU time rather than “t” of the iteration (it could be a better comparison because some methods are more complicated and use more CPU during each step).

1. Show an example where ADAM algorithm starts to become unstable with respect to a minimum of the function that was reached earlier. Choose a simple function for this.

**Due to a new bug in google, the next exercises will appear in separate files**