

**Homework 4. Due Oct. 7.**

1. **(10 pts)** The goal of this problem is to experiment with the accelerated gradient descend algorithms, namely Nesterov's accelerated gradient descend

$$\mathbf{x}_{k+1} = \mathbf{y}_k - \alpha \nabla f(\mathbf{y}_k), \quad (1)$$

$$\mathbf{y}_{k+1} = (1 + \mu_k) \mathbf{x}_{k+1} - \mu_k \mathbf{x}_k, \quad \mu_k = 1 - \frac{3}{5 + k} \quad (2)$$

( $\alpha$  is constant, you need to choose it from experimentation), and Adam (adaptive moment estimation) – see the original paper by D. P. Kingma and J. L. Ba “[Adam: A Method for Stochastic Optimization](#)”. For each of these methods, implement their deterministic and stochastic versions.

Apply both of these methods to the problem from HW3, problem 2 (optimizing a neural network for solving a boundary value problem for the Poisson equation). Compare performance of the deterministic versions of Nesterov and Adam with the deterministic gradient descend, and their stochastic versions with the stochastic gradient descend. As in HW3, try to adjust parameters for the methods to make the error as small as you can.

Write a report summarizing your observations and supplement it with all appropriate graphs. Copy/paste and insert code lines for Nesterov and Adam into your report.