Programming Assignment 1

Issued: Tuesday 16$^{\text{th}}$ October, 2018   **Due:** Thursday 25$^{\text{th}}$ October, 2018

1.1. (5 points) *Linear regression*   Consider the linear observation model

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{n}$$

where $\boldsymbol{A}$ is a $10\,000 \times 10$ matrix, and $\boldsymbol{x}, \boldsymbol{n}$ are column vectors with lengths 10 and $10\,000$. Use gradient descent to find the $\boldsymbol{x}$ that minimizes the loss $\frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2^2$.

1.2. (5 points) *Softmax regression of MNIST*   The MNIST database contains $60\,000$ training images and $10\,000$ testing images. In this task, each image of MNIST has been reshape to a column vector of length $28^2 = 784$. You need to calculate the gradient `grad` and update the weights `soft max_weights` in softmax regression.

`Pa1_2018.py` will walk you through this exercise.

Notice:

- Use matrix operations other than loops for efficiency. If the running time exceeds 5 minutes, you will get point deductions.

- In the second task, test accuracy is about 80% after 10 training epochs.