UNIVERSITY OF SCIENCE AND TECHNOLOGY
DEPARTMENT OF APPLIED COMPUTER SCIENCE
AND MODELLING

FUNDAMENTALS OF OPTIMIZATION

Multi-variable optimization using gradient mehods

1. Aim of the exercise.

The aim of the exercise is using gradient, multi-variable optimization methods to find the minimum of the given objective function.

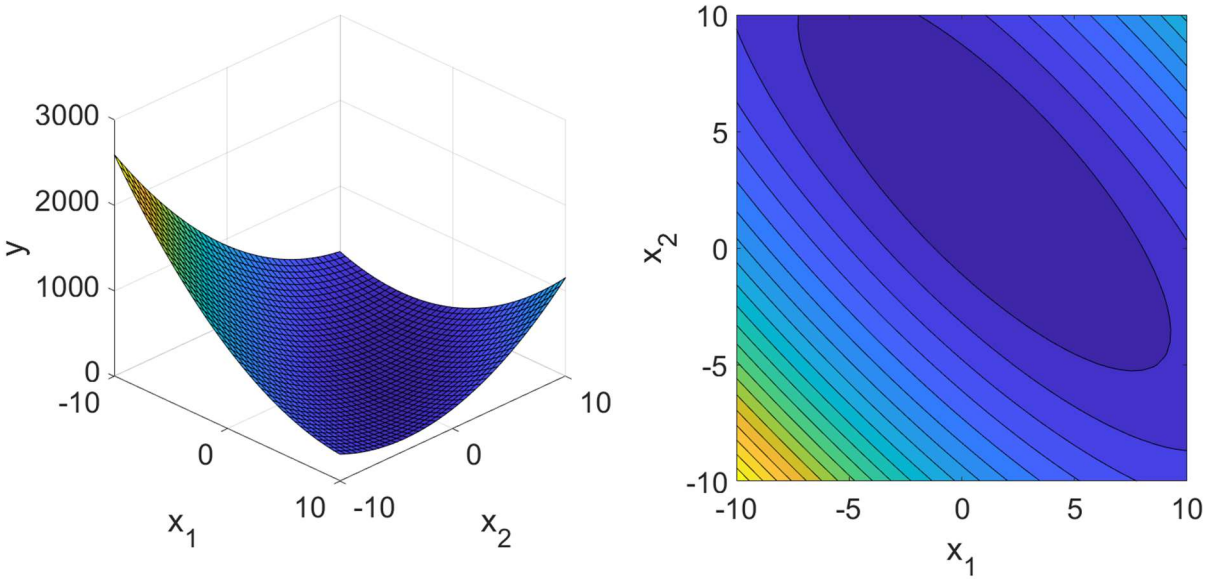
2. Test function.

The test function is given by following equation:

$$f(x_1, x_2) = (x_1 + 2x_2 - 7)^2 + (2x_1 + x_2 - 5)^2$$

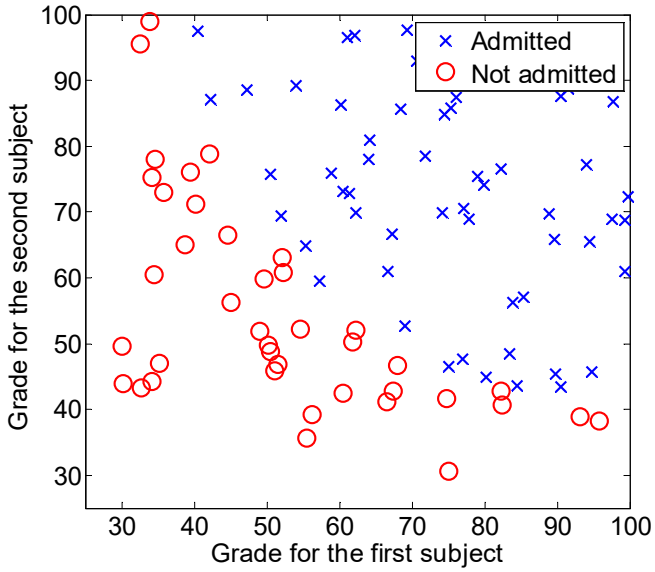
where: $x_1 \in [-10, 10], x_2 \in [-10, 10]$.

The test function is shown in the figure below.



3. Logistic regression.

Admission to a certain university is determined by the grades obtained in two subjects. The figure below shows some examples of grades together with the decision to admit or not a candidate.



The aim of optimization is to find the optimal parameters of the classifier, which, on the basis of the obtained grades, will be able to predict whether a given person will be admitted to university. The classifier will use the hypothesis defined as follows:

$$h_\theta(\mathbf{x}) = \frac{1}{1 + e^{-\boldsymbol{\theta}^T \mathbf{x}}}$$

where: vector $\boldsymbol{\theta} = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix}$ contains the parameters of the classifier, vector $\mathbf{x} = \begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix}$ contains the grades for the first (x_1) and the second (x_2) subject.

The classifier will indicate that the *i*-th candidate will be admitted to the university if the value of the hypothesis will be greater than or equal to 0.5, i.e.:

$$h_\theta(\mathbf{x}^i) \geq 0{,}5 \implies y^i = 1$$

Training data (shown in the figure above) are available in the **data.mat** file in the form of matrix X and vector Y:

$$X = [\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^m], \mathbf{x}^i = \begin{bmatrix} 1 \\ x_1^i \\ x_2^i \end{bmatrix}, i = 1, \dots, m$$

$$Y = [y^1, y^2, \dots, y^m],\ y^i = \{0{,}1\}, i = 1, \dots, m$$

where: number of training data $m = 100$, $y^i = 1$ means that the *i*-th candidate has been admitted, $y^i = 0$ means that the *i*-th candidate was not admitted.

The search for the optimal parameters of the classifier should be performed by minimizing the cost function given by the formula:

$$J(\boldsymbol{\theta}) = -\frac{1}{m}\sum_{i=1}^{m}\left(y^i \cdot \log_e\left(h_\theta(\mathbf{x}^i)\right) + (1-y^i)\cdot \log_e\left(1-h_\theta(\mathbf{x}^i)\right)\right)$$

The partial derivatives of the cost function $J(\boldsymbol{\theta})$ are defined as follows:

$$\frac{\partial J(\boldsymbol{\theta})}{\partial \theta_j} = \frac{1}{m}\sum_{i=1}^{m}\left(h_\theta(\mathbf{x}^i) - y^i\right)\cdot x_j^i \text{ dla } j = 1, \dots, n$$

where n is the length of gradient vector.

4. Optimization methods.

To perform optimization use **fminunc** function. Starting point for the test function should be chosen randomly from given range, while for the logistic regression, it should be equal to zero $\boldsymbol{\theta}^{(0)} = \begin{bmatrix}0\\0\\0\end{bmatrix}$.

5. Realization of the exercise.

During the exercise four m-files should be written:

- **start.m** – a script which runs all computations. It should:
  - display the names of the Authors of the code,
  - display the optimum found for the test function ($\mathbf{x}$, y and the number of objective function calls),
  - plot the figure showing the test function (as a contour plot), starting point and found optimum,
  - display the optimum found for the logistic regression ($\boldsymbol{\theta}$, J and the number of objective function calls),
  - display the number of correct prediction of admittance calculated on the basis of training dataset,
  - plot figure showing the candidates grades and the decision boundary.
- **ff_test.m** – a function which calculates and returns the test function value and its gradient at given point: [**y,dy]=ff_test(x)**
- **ff_logistic_reg.m** – a function which calculates and returns the cost function and its gradient: **[J,dJ]=ff_logistic_reg(theta,X,Y)**;
- **predict.m** – a function which returns the predicted admission decision (0 or 1): **y=predict(x,theta)**.

To validate if the and **ff_logistic_reg** function is written correctly compute value of cost function and its gradient for $\boldsymbol{\theta} = \begin{bmatrix}0\\0\\0\end{bmatrix}$. The results should be:

```
>> load('data.mat');
>> [J,dJ]=ff_logistic_reg(zeros(3,1),X,Y)
J =
      0.69315
dJ =
```

```
   -0.1
-12.009
-11.263
```

6. Report.

As the report, four m-files should be sent via UPeL platform (by one Author only).