Department of Engineering/Informatics, King's College London
Nature-Inspired Learning Algorithms (7CCSMBIM)
Tutorial 2

Q1. What are the advantages and disadvantages of gradient descent method?

Q2. Show how gradient descent method works using pseudo code.

Q3. Consider a least-squares problem, $\min_{\mathbf{x}} f(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{B}\|_2^2$ where $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$. Find the optimal solution $\mathbf{x}$.

Q4. Consider the minimisation problem of the function: $f(x,y) = (x-1)x + (y+1)y$ using Nelder-Mead Downhill Simplex Method.

    a. Starting with 3 vertices $(0, 0)$, $(1, 2)$, $(3, 4)$, determine the points $\mathbf{B}$, $\mathbf{G}$ and $\mathbf{W}$.

    b. Find the midpoint $\mathbf{M}$ and $f(\mathbf{M})$.

    c. Find the reflection $\mathbf{R}$ and $f(\mathbf{R})$.

    d. Determine the 3 vertices ($\mathbf{B}$, $\mathbf{G}$ and $\mathbf{W}$) in the next iteration. If Case (ii) is performed in the pseudo code of Nelder-Mead downhill simplex method, $\mathbf{C} = \frac{\mathbf{W}+\mathbf{M}}{2}$ is used.

Q5. Considering the minimisation problem of the function: $f(x,y) = (x-1)x + (y+1)y$ using the gradient descent method, find x and y, and $f(x,y)$ for the first 3 iterations with step size $h_k = 0.1$ and initial guess $(7,8)$.

Q6. Consider the minimisation problem of the function: $f(x,y) = (x-1)x + (y+1)y$ using the gradient descent method.

    a. Determine the optimal step size h_k.

    b. Show that it requires one iteration for the gradient descent method to converge with the optimal step size h_k obtained in Q6a.

Q7. Consider the minimisation problem of the function: $f(x,y) = (x-1)x + (y+1)y$ using the random walk optimisation algorithm. The *Threshold* (for accepting the worse solution) is 0.75 and the random number generator will generate a repeating sequence of $\{0.8, 0.7, 0.2, 0.6, 0.9, 0.7, 0.5, 0.6\}$ (The left-most number is the first number generated). The diagonal entries of $\mathbf{D}_k$ are all 1 and all the entries of $\mathbf{h}_k$ are 0.5. Fill in the content of the following table.

k	$\mathbf{x}_k^T$	$f(\mathbf{x}_k)$	$\mathbf{x}_{k+1}^T$	$f(\mathbf{x}_{k+1})$	rand()	$\mathbf{x}_{best}$	$f(\mathbf{x}_{best})$
0	$\begin{bmatrix} -1 & -2 \end{bmatrix}$						
1							
2							
3							
4							
5							

Q8. Consider a system taking two input variables x_1 and x_2 and generate an output y. A set of M input-output data is collected in an experiment, e.g., the dataset is $(x_1(1), x_2(1), y(1))$, $(x_1(2), x_2(2), y(2))$, $\cdots$, $(x_1(M), x_2(M), y(M))$. Design a linear regressor in the form of $\hat{y} = w_1 x_1 + w_2 x_2 + b$ to best fit the data in terms of Mean Squared Error using the gradient descent method, where w_1 and w_2 and b are the parameters to be determined.

  a. Formulate the data fitting problem as a minimisation problem.

  b. Denote the step size as h_k. Derive the update rule for each parameter.

  c. Use $h_k = 0.1$ and the initial guess for all variables is 1. Considering the dataset: $(1, -2, 3)$, $(2, 4, -1)$, $(3, 0, 5)$, obtain the best set of parameters for the linear regressor.

  d. Plot "iteration k" against "MSE" for 200 iterations. Is the choice of $h_k = 0.1$ right or wrong?

Q9. Find the least-squares solution of $\mathbf{A}\mathbf{x} = \mathbf{B}$ where:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix} \text{ and } \mathbf{B} = \begin{bmatrix} 0 \\ 1 \\ 3 \\ 4 \end{bmatrix}.$$

Q10. Consider the three data points (x, y) given as follows:
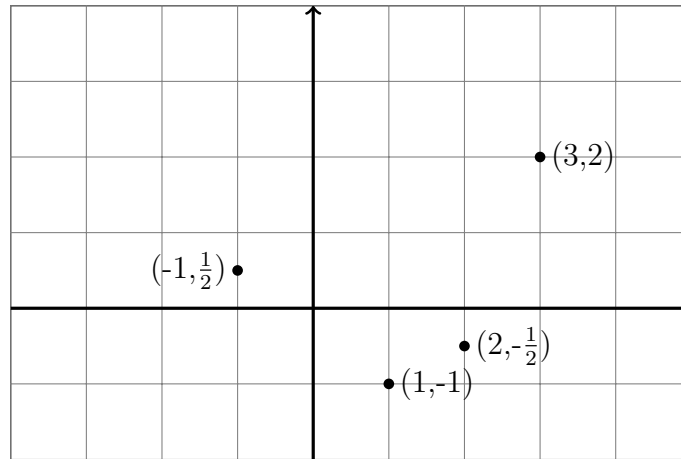
$$(0, 6), (1, 0), (2, 0),$$

which are the data obtained from a single line but have errors due to measurement. Find a straight line that best fits these points in the sense of least-squares? Explain the quantity that is being minimized.

2
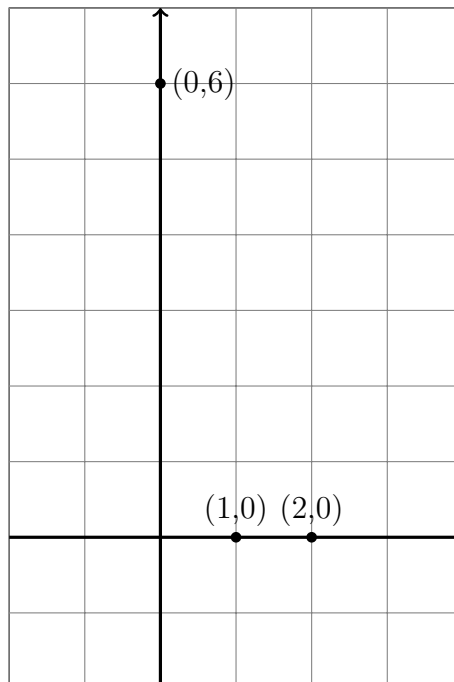
Q11. Find a parabola equation that best fits the below data points (x, y) in the sense of least-squares.
$$(-1, \tfrac{1}{2}), (1, -1), (2, -\tfrac{1}{2}), (3, 2).$$
What quantity is being minimized?

Q12. According to the following data, find the best approximated linear function $f(x, y)$ in the sense of least-squares and identify the quantity being minimized.

x	y	$f(x,y)$
1	0	0
0	1	1
-1	0	3
0	-1	4

Q13. There are three categories of machines I, II and III installed in a factory. Machines have different requirements on operating time. Machines I and II can be operated for most 12 hours whereas machine III needs to be operated for at least 5 hours per day. This factory produces two items M and N. Both require the use of all the three machines. The operating time of producing 1 unit of each item on the three machines are presented in Table 1 as follows:

Items	Machine I	Machine II	Machine III
M	1	2	1
N	2	1	1.25

Table 1: Operating time of producing one unit of each item on each machine.

The profit of item M and N are £600 and £400 each per unit, respectively. What would be the most reasonable strategy to produce the items M and N as to maximise the profit of the factory? What would be the maximum profit in this case?

Q14. There are two factories located in place P and Q. The commodity produced by these two factories will be delivered to three depots situated at A, B and C. Each week, the depots at A, B and C requires 5, 5 and 4 units of the commodity, while the production capacity of the factories at P and Q are 8 and 6 units correspondingly. The commodities produced at both factories will be all transported. The cost of transportation varies from different factories to different depots where the details of the cost of transportation per unit commodity are in Table 2.

From/To	Cost of A	Cost of B	Cost of C
P	160	100	150
Q	100	120	100

Table 2: Cost of transportation per unit commodity.

Find the units of commodity from factories to depots so the cost of transportation is minimum? What will be the minimum cost of transportation?

Q15. A factory wishes to produce a new type of food which has two ingredients denoted as Ingredient I and Ingredient II. This new type of food need to contains vitamin A and vitamin C with minimum value 8 and 10 respectively. Ingredient I contains 2 units/kg of vitamin A and 1 units/kg of vitamin C. Ingredient II contains 1 units/kg of vitamin A and 2 units/kg of vitamin C. The price of ingredient I and ingredient II are £50/kg and £70/kg respectively. Decide the formula (the weight of each ingredient) of this new type of food so that the cost is at the minimum?

Q16. Explain how "Recursive least-squares" works.

Q17. Consider a least-squares problem as $\mathbf{A}\mathbf{x} = \mathbf{B}$, where

$$\mathbf{A} = \begin{bmatrix} 5 & 6 \\ 5 & 8 \\ 4 & 5 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix}$$

Given that there are five new samples coming in the the following order,

$$\mathbf{a}_1^T = \begin{bmatrix} 3 & 4 \end{bmatrix}, \quad b_1 = 7;$$
$$\mathbf{a}_2^T = \begin{bmatrix} 10 & 10 \end{bmatrix}, \quad b_2 = 5;$$
$$\mathbf{a}_3^T = \begin{bmatrix} 3 & 8 \end{bmatrix}, \quad b_3 = 8;$$
$$\mathbf{a}_4^T = \begin{bmatrix} 8 & 8 \end{bmatrix}, \quad b_4 = 2;$$
$$\mathbf{a}_5^T = \begin{bmatrix} 7 & 5 \end{bmatrix}, \quad b_5 = 3.$$

calculate the initial solution, denoted as $\mathbf{x_0}$, for $\mathbf{A}\mathbf{x} = \mathbf{B}$ and update the solution using the recursive least-squares technique with the given new samples. Show the working and the solution after adding the samples one by one according to the order shown in the above.