# COMS10013 - Analysis - WS4

## Questions

These are the questions you should make sure you work on in the workshop.

1. **Gradient descent:** In this question we're going to study the function

$$E(x,y) = x^2 + y^2$$

   (a) Calculate $\nabla E(x,y)$

   (b) Using the initial value $(x,y) = (1,2)$ and step $\eta = 0.1$, calculate $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ of the gradient descent algorithm. What is $E(\mathbf{x}_3)$?

   (c) Using the same initial value, calculate $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ taking

      (i) $\eta = 0.5$

      (ii) $\eta = 1$

   Extrapolate what value you'd get for $\mathbf{x}_n$ and $E(\mathbf{x}_n)$; discuss whether this is a good choice of $\eta$.

   (d) Find all minima of $E$ using calculus

   (e) Pick a new initial starting point (not a minimum value that you found in (d)!), and calculate $E(\mathbf{x}_3)$ using $\eta = 0.1$. Compare this to what you found in (b) – is it a better starting point? Is this what you'd expect given what you found in (d)?

2. **Nelder-Mead Method:** In this question, we're going to study the function

$$E(x,y) = x^2 + y^2 \,.$$

The eagle-eyed amongst you will notice that this is the same function as in question 1. This means that we already know what the minimum should be; the purpose of this question is to use the downhill simplex method to find the minimum.

We'll start with the initial simplex points: $(0,0), (1,0), (0,1)$.

   (a) Draw the initial simplex and assign values $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ to the initial simplex points so that $E(\mathbf{x}_1) \leq E(\mathbf{x}_2 \leq E(\mathbf{x}_3$.

   (b) Calculate the centroid of the simplex and find $\mathbf{x}_r$ and $E(\mathbf{x}_r)$, where $\mathbf{x}_r$ is the reflected point.

   (c) What happens next in the algorithm? Determine the next point that the algorithm will create, and assign the values $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ to the new simplex that is created. Draw the new simplex.

   (d) Carry out the algorithm so that you find the next two new points in the algorithms (and draw the new simplicies that arise).

   (e) Based on what you're seeing in this algorithm (and also the fact that you know the minimal point of $E$), how would you characterise a sensible stopping condition for the algorithm for this choice of $E$?

## Extra questions

These are extra questions you might attempt in the workshop or at a later time. Some parts of these questions require a computer (e.g. using Python).

1. **Gradient descent:** In this question we're going to aim to minimise the function

$$f(x, y) = e^{-x} \cos(x) y^2$$

using Gradient descent.

(a) Calculate $\nabla f(x, y)$

(b) Choose three random initial vectors.

(c) Using $\eta = 0.01, 0.1, 0.5$, calculate $\mathbf{x}_{100}$ and $f(\mathbf{x}_{100})$ for each of your initial starting points. Interpret your result: given what you've found, what do you expect the minimum value of $f$ to be? Which values of $\eta$ lead to the quickest minimisation? Which initial points seem better or worse?

(d) With the same values of $\eta$ as in (c), calculate $\bar{x}$ for the first 100 outputs, and also $f(\bar{x})$. Interpret your result and compare it with the outputs found in (c).

(e) Look again at $f$; what do you expect the minimum value to be? Comment on how well the outputs from (c) and (d) perform compared to this minimum value.

2. **Nelder-Mead:** Using your favourite computer program of choice, write a function that carries out the Nelder-Mead algorithm on a two dimensional function. If you're feeling confident, try it on a three dimensional function.