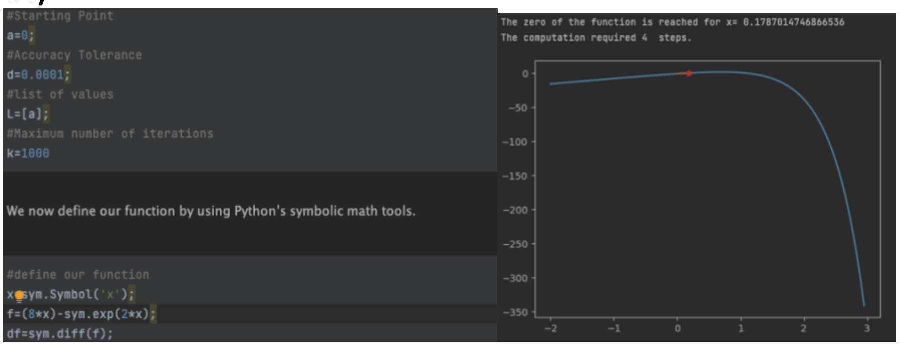
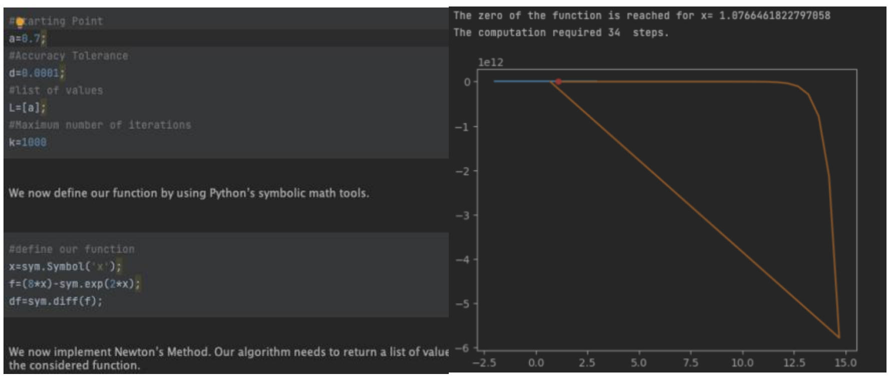
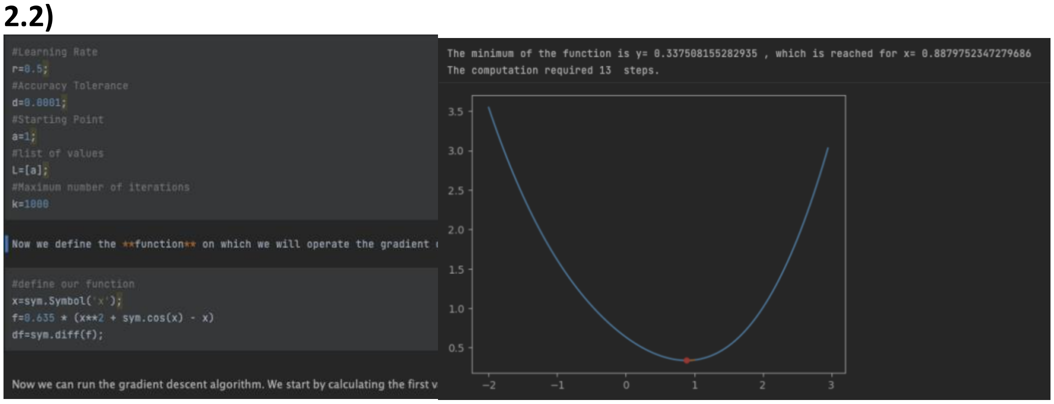
Code Screenshots

1.4)

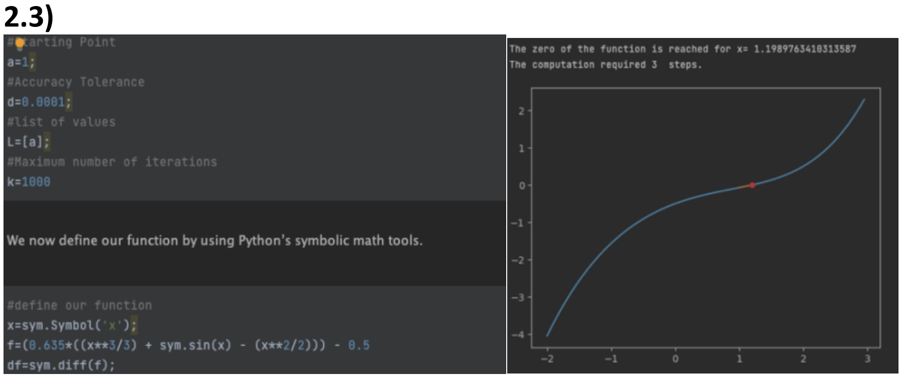




2.2)



2.3)



COMP0199 - Assessment 2

1.1) Stationary point is when $\nabla f = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$\frac{df}{dy} = 0$ | $\frac{df}{dx} = 0$

$\therefore 8x - e^{2x} = 0$ | $\therefore 8y - 2ye^{2x} = 0$

$E_1 : 8y - 2ye^{2x} = 0$

$E_2 : 8x - e^{2x} = 0$

1.2) $8y - 2ye^{2x} = 0$

$2y(4 - e^{2x}) = 0$

$2y(2 + e^{x})(2 - e^{x}) = 0$

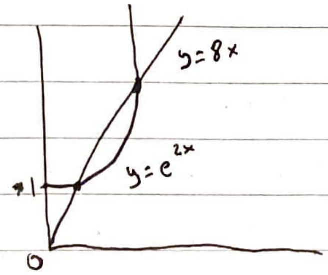
Solutions : $y = 0$ or $x = \ln(2)$

Set of Solutions = $\{(\ln(2), y) \mid y \in \mathbb{R} \cup (x, 0) \mid x \in \mathbb{R}\}$

1.3) $g(x) = 8x - e^{2x}$

1.4) We need to find the zeros of $g(x)$, so we use Newton's method. To implement it, we need to set a few constants. These are the starting point, max number of iterations, and the accuracy tolerance. For the algorithm itself, we use a simple implementation of newtons method. We start a while loop where we keep iterating until ~~our function is~~ the value of our function is within the accuracy tolerance of 0, or until we exceed the max num of iterations. At each iteration, we do the following operations. Take the last value from the ~~calculated~~ list of previously calculated values that we maintain. For the first iteration, this will be the starting point that we defined. Then, we simply subtract this from the value of the function at the previous value divided by the derivative of the function at the previous value. We append this to our list of calculated values and continue iterating.

Choosing the right parameters is an extremely important task. For the accuracy tolerance, I decided upon 0.0001 to be within 3 decimal place accuracy. For the starting point, there's a few observations we can use. Finding the zeroes of $g(x)$ is equivalent to finding the intersection of $8x$ and e^{2x}. Since, $8x < 0$ when $x < 0$ and $e^{2x} > 0$ always, we know that our starting point can't be negative. So we can take $x=0$ as our first starting point. Doing a quick sketch of the functions:



We see that there is a maximum second intersection, since after the second intersection, the derivative of e^{2x} will always be increasing while the derivative of $8x$ is constant, which means they will not intersect again. For the starting point of the second intersection, we can use the x value where the derivative of e^{2x} is greater than 8; since this is where the function will start to grow faster than $8x$ and move towards their intersection. This happens at $x = 0.693$. We can round to $x = 0.7$ since ~~in [illegible]~~ it will be closer to the second intersection anyway.

Using these two starting points, we get the following solutions for E_2: $x = 0.179$ or $x = 1.077$, which corresponds to the following ~~[illegible]~~ pairs of coords:

$$\{(0.179, y) \mid y \in \mathbb{R} \cup (1.077, y) \mid y \in \mathbb{R}\}$$

1.5) For the stationary points we need the intersection of the sets of solutions for E_1 and E_2. This yields two coordinates: $(0.179, 0)$ and $(1.077, 0)$

1.6) Hessian of $f(x,y)$

$$= \begin{bmatrix} -4ye^{2x} & 8-2e^{2x} \\ 8-2e^{2x} & 0 \end{bmatrix}$$

determinant of hessian $= 0 - (8-2e^{2x})^2$

$(8-2e^{2x})^2$ is always positive unless $x = \ln(2)$. Since it is not in our set of stationary points, $(8-2e^{2x})^2$ is always strictly positive for our stationary points and thus $0-(8-2e^{2x})^2$ is always negative for the stationary points. Therefore, both stationary points are saddle points, since the determinant of the hessian at those points is negative.

2.1) $h(x)=\lambda(x^2+\cos(x)-x)$

For $h(x)$ to be a PDF, $\int_0^2 h(x) = 1$

$\lambda\int_0^2 x^2+\cos(x)-x = 1$

$\lambda\left[\frac{x^3}{3}+\sin(x)-\frac{x^2}{2}\right]_0^2 = 1$

$\lambda\left(\frac{8}{3}+\sin(2)-2\right)=1$

$\therefore \boxed{\lambda = 0.635}$

2.2) Since we need to find the local minimum, we can use gradient descent. Similarly to newton's method, we define constants for the accuracy tolerance, starting point, max number of iterations, and for gradient descent we also define the learning rate. We start a while loop and keep iterating until the difference between consecutive values is less than the accuracy tolerance, or we exceed the max num of iterations. At each iteration, we perform a simple calculation. We take the difference of ~~subtract~~ the last calculated value and the product of the derivative at the last calculated value with the learning rate. We append this to our list of values and continue the iteration. For the first iteration, we use the starting point..

For the parameters, we choose 0.0001 for the accuracy tolerance to stay with 3 decimal figure accuracy. ~~Since there is only we are trying to find the one local minimum~~ For the starting point, we can make a few observations. $x^2+\cos(x)-x$ will look like the parabola x^2-x, with a $\cos(x)$ term added. Since $\cos(x)$ is strictly decreasing between 0 and 2 the general shape of the parabola will stay the same and not have any weird jumps between 0 and 2. Thus, since there is only one local minimum for the parabola, the starting point doesn't really matter, and we can choose the middle of the range $(0,2)$ to be on average closest to the possible minimum. Thus, our starting point will be 1. For the learning rate, we can use the fact that if $\|\nabla f(x)-\nabla f(y)\| \leq L\|x-y\|$ for all x,y, then a learning rate of $r \leq \frac{1}{L}$ converges with order $O(\frac{1}{k})$ where k is the number of iterations. Between 0 and 2, the derivative of $x^2+\cos(x)-x$ is bounded between -2 and 2. Thus the maximum difference between derivatives is 4, while the max difference in x-values is 2. Thus $\|\nabla f(x)-\nabla f(y)\| \leq 2\|x-y\|$, and our learning rate can be chosen to be $1/2$. For the max num of iterations, we'll stick with the default of 1000.

Using these parameters, we find $x_0 = 0.888$ and $y_0 = 0.338$ within 13 steps.

2.3) To find a such that $Pr(X < a) = \frac{1}{2}$,

we need to find a such that $\int_0^a h(x) = \frac{1}{2}$

$$\int_0^a 0.635\,(x^2 + \cos(x) - x) = \frac{1}{2}$$

$$0.635\left(\frac{a^3}{3} + \sin(a) - \frac{a^2}{2}\right) = \frac{1}{2}$$

$$\frac{0.635\,a^3}{3} + 0.635\sin(a) - 0.635\,\frac{a^2}{2} - \frac{1}{2} = 0$$

Finding a is the same as finding the zeros of the function above. This can be done with Newton's method with respect to a. Details regarding the implementation were discussed in 1.4, so we don't need to repeat them. We just need to figure out the relevant constants. For accuracy tolerance, we'll stick with 0.0001 to stay within 3 decimal places accuracy. For the starting point we can make a few observations. The derivative of the above function is equivalent to $h(a)$. Since h is a PDF, it will always be positive from 0-2, and so the derivative of the above function is always positive. This means the function is monotonically increasing from 0-2. The function also takes the value of -0.5 at $a=0$ and 0.501 at $a=2$. Since the function goes from positive to negative and is monotonically increasing from 0 - 2, we know that there is ~~at least~~ exactly 1 zero. Thus, the starting point doesn't really matter since we converge to the same 0, and we can just take the midpoint of 0 and 2 to stay within the domain of h. Thus, our starting point will be 1. For max iterations, we'll stick with the default of 1000.

Using these parameters, we find the value of a is 1.149 within 3 steps.