

# Homework on Newton's methods

Leave your name and uni here

## Problem 1: Univariate optimizations.

In this problem, you will compare three common methods for univariate minimization: Newton's Method, Bisection Method and Golden-Section Search.

### Function A:

$$f_1(x) = \ln(1 + x^2) + x^2$$

### Function B:

$$f_2(x) = x^4 - 6x^2 + 4x + 8$$

### Answer the following questions:

1. Compute  $f'(x)$  and  $f''(x)$
2. Are  $f_1(x)$  and  $f_2(x)$  unimodal functions? If not, how many modes it contains?
3. Implementing Bisection, Golden Search and Newton's methods to find global minimum of  $f_1(x)$  and  $f_2(x)$
4. Discuss your results – which method is fastest in terms of iteration counts? which method is easiest to apply if you only know a broad interval containing the minimum? which method fails or converges poorly if started badly? how the shape of the function influences the performance of these algorithms?
5. Please show all relevant R code for bisection, golden-section, and Newton's methods.

### Answer: your answer starts here...

```
#R codes:
```

## Problem 2: Newton's Method in Two Dimensions

$g(x, y)$  is a 2D function

$$g(x, y) = x^2 + xy + y^2 - 4x - 3y + 7$$

## Answer the following questions:

1. Derive the gradient  $\Delta g(x, y)$  and the Hessian matrix  $\Delta^2 g(x, y)$ .
2. Implement a Newton's algorithm to find its minimizer.
3. Choose two different starting values, and compare the resulting solutions, iteration counts and path to convergence

4. Create a contour plot of  $g$  around its minimum, and overlay the sequence of iterates from Newton's method to show the path to the minimum.

**Answer: your answer starts here...**

*#R codes:*

### Problem 3

Suppose we have data  $(x_i, y_i), i = 1, \dots, n$ , with  $y_i$  follows a conditional exponential distribution

$$Y_i | x_i \sim \exp(\lambda_i), \text{ where } \log(\lambda_i) = \alpha + \beta x_i.$$

**Please complete the following tasks:**

1. Derive the log-likelihood of  $(x_i, y_i)$ , as well as its Gradient and Hessian Matrix
2. Generate a synthetic data with true  $\alpha = 0.5$ , true  $\beta = 1.2$  and sample size  $n = 200$ .
3. Implement an Newton's algorithm to its MLE
4. Implement a modified Newton's algorithm, where you incorporate both of the step-having and ascent direction check. If a direction is not ascent, you can switch to a simpler gradient descent.
5. For a generalized linear models, one can replace the observed Hessian with the expected Hessian (the Fisher information), which might lead to stable updates akin to a Fisher scoring approach. Implement another modified Newton's algorithm with Fisher scoring to estimate MLE.
6. Compare the final parameter estimates, iteration counts and convergences of optimization, and summarize your findings.