Project 1: Statistical Computing Methods

Introduction of Computational Statistics/Statistical Computing

- Here are what you need to do for Project 1:

To do individually - Part 1

- In part 1, you are expected to learn the listed methods by yourself:

-Root finding (bisection, Newton-Raphson)

-Optimization (genetic algorithm)

-Numerical integration (Quadrature)

-Monte Carlo integration (distribution sampling, thus need to learn one or two sampling methods)

- **Then** write a concise report (no more than 5 pages) to answer the following questions:

1. What are these methods for?
2. How do these methods work?
3. What are the advantages and disadvantages of these methods?
4. How to evaluate the performance of each of these methods?
5. How can these methods help you in real life? Give an example for each of the methods

Ref: previous links and the book: *Computational Statistics*, by Givens and Hoeting

To do as a group – Part2

- 1. implement both bisection and Newton-Raphson algorithms to find the maximal value of $\frac{log(x+x^2)}{1+x^3}$
- 2. implement a genetic algorithm to maximize the following function with n=10:

$$f(\vec{x}) = \begin{cases} \left| \dfrac{\sum_{i=1}^{n} \cos^4(x_i) - 2\prod_{i=1}^{n} \cos^2(x_i)}{\sqrt{\sum_{i=1}^{n} i x_i^2}} \right|, & if\ (\forall i,\ 0 \le x_i \le 10)\ and\ (\prod_{i=1}^{n} x_i \ge 0.75) \\ 0, & Otherwise\ (i.e. not\ feasible) \end{cases}$$

- 3. implement both a quadrature method and a Monte Carlo method to estimate the integral: $\int_0^{+\infty} \frac{|\cos(x)|}{x} e^{-(\log(x)-3)^2}\, dx$
- 4. calculate the correlation between x&y of the following distribution:

$$p(x,\ y) \propto \sin(\pi x)\sin^{20}(\pi x^2) + \sin(\pi y)\sin^{20}(2\pi y^2), \quad 0 \le x, y \le 1$$