STA 5106
Computational Methods in Statistics I

Department of Statistics
Florida State University

Class 9
September 24, 2019

Chapter 3

Non-Linear Statistical Methods

3.1 Introduction

Non-linear Optimization

- This chapter is devoted to solving for the optimal points of given functions, linear or nonlinear.

- An instance of nonlinear optimization problems in statistics occurs in cases of **maximum likelihood estimation**.

- Example:

$$y = F(x, b) + \varepsilon$$

  where x and y are random variables of interest, b is a constant vector of parameters, and ε is a random error with the density function given by f_ε.

- In many practical situations, we have the sampled values of y, x, and our goal is to estimate b according to some criterion.

Maximum Likelihood Estimation (MLE)

- Assuming sample size is *n* and different measurements are independent of each other, *b* can be estimated as follows:

$$\hat{b} = \arg\max_{b} \prod_{i=1}^{n} f_{\varepsilon}(y_i - F(x_i, b))$$

- The quantity on the right side of the above equation is called the *likelihood function*.

- $\hat{b}$ is called the *maximum likelihood estimate* (*MLE*) of *b*.

- In general, a value of *b* which maximizes the likelihood function can be found out **by seeking the roots of the first derivative of the likelihood function**.

3.2 Numerical Root Finding: Scalar Case

Goal

- We are given a real-valued function $f(x)$ where $x \in \mathbf{R}$ and the goal is to find the roots of f, i.e., find x^* such that $f(x^*) = 0$.

- Due to the non-linearity of the cost function its roots are found in an iterative way: an initial estimate is chosen and is modified iteratively until some stopping criterion is met.

- There are many different methods for root-finding. They all have the following three essential elements:

  1. Some of determining the starting value, x_0.

  2. Given the i-th iterate some formula for calculating the $(i+1)$-th iterate.

  3. Some stopping criterion.

Order of Convergence

- For each of these algorithms we are interested in finding their rate of convergence towards the roots of the function.
- The rate of convergence is defined by a quantity called the **order of convergence**.
- **Definition 10** For a converging sequence $\{x_i\}$, the order of convergence is defined by β if,

$$\lim_{i \to \infty} \frac{E_{i+1}}{E_i^{\beta}} = K,$$

where $E_i = |x_i - x_{i-1}|$ and K is a constant.

Simple Iterations

- Instead of solving for the roots of f we solve an equivalent problem of finding the fixed point of another function g. g is found in such a way that

$$f(x^*) = 0 \Leftrightarrow g(x^*) = x^*.$$

- The selection depends on which g is easier to handle and solve for a fixed point. One choice which is always valid is

$$g(x) = x + f(x) .$$

- Let x_0 be some starting value for the iterative search of the roots of f. At the i-th stage of the algorithm, the next iterate is given by the formula:

$$x_{i+1} = g(x_i) = x_i + f(x_i) .$$

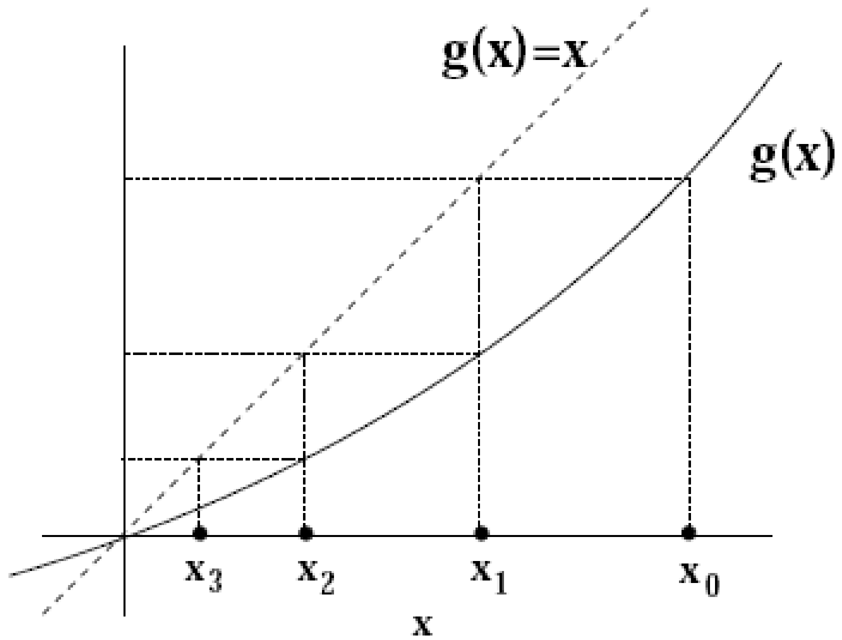
Algorithm

- **Algorithm:** Given a function $f(x)$ find its roots using simple iterations by defining $g(x) = x + f(x)$, and:

- **Algorithm 21** (Simple Iterations)

```
x(1) = x0;
gx = g(x(1));
i = 1;
while (abs(x(i) − gx) > ε)
        gx = x(i);
        x(i + 1) = g(x(i));
        i = i + 1;
end
```

Illustration



- The order of convergence in simple iterations is 1, or the convergence is linear.