

# Problem Set: Nonlinear Programming

## Instructions

In this problem set, you will implement three numerical methods for solving nonlinear equations:

- Bisection Method
- Fixed Point Iteration
- Newton's Method

For each problem:

- Write your own implementation from scratch (no use of built-in solvers like `uniroot()` or `nleqslv()`).
  - Use a convergence criterion of  $|x_{k+1} - x_k| < 10^{-6}$ .
  - Include a plot of the function over an appropriate interval.
  - Report the final root estimate, number of iterations, and whether convergence was successful.
- 

## 1 Bisection Method

Let

$$f(x) = \cos(x) - x$$

Use the bisection method to solve  $f(x) = 0$  on the interval  $[0, 1]$ .

### 1.1 Tasks

- a. Confirm that  $f(0) \cdot f(1) < 0$ .
- b. Implement the bisection method.
- c. Plot  $f(x)$  over  $[0, 1]$  and mark the root.
- d. Report: number of iterations, final estimate, and value of  $f(x)$  at the root.

**Comment:** Explain why is convergence guaranteed given the function?

---

## 2 Fixed Point Iteration

Consider the function:

$$x = g(x) = \sqrt{1+x}$$

## 2.1 Tasks

- Define the iteration  $x_{n+1} = g(x_n)$ .
  - Use  $x_0 = 1$  as the starting value.
  - Plot  $g(x)$  and the 45-degree line  $y = x$  over  $[0, 3]$ .
  - Determine whether the iteration converges.
  - Try other initial values (e.g.,  $x_0 = 0.5$ ,  $x_0 = 3$ ) and compare behavior.
  - Check the derivative  $g'(x)$  at the fixed point. What does this tell you about convergence?
- 

## 3 Problem 3: Newton's Method

Let

$$f(x) = x^3 - 2x + 2$$

### 3.1 Tasks

- Implement Newton's method.
- Use  $x_0 = 0$  and  $x_0 = -1.5$  as two different initial values.
- Plot  $f(x)$  and its derivative over  $[-3, 3]$ .
- Report the outcome of the iteration for each starting point.
- Compare the convergence (or divergence) behavior. What role does the choice of initial guess play?
- Why does the method fail to move when starting at  $x_0 = 0$ ? explain.

## 4 Problem 4: Pollution Abatement under an Emissions Cap

A firm produces output  $q$  which generates pollution. For each unit of output, one unit of emissions is generated unless the firm invests in abatement  $a$  (e.g., cleaner technology, filters). Abatement reduces emissions one-for-one, but it is costly.

- Profit function (before regulation):

$$\pi(q, a) = 100q - 2q^2 - \frac{1}{2}a^2$$

- Emissions constraint:

$$q - a \leq E$$

where  $E$  is the regulatory emissions cap.

- Non-negativity:  $q \geq 0$ ,  $a \geq 0$ .

Questions:

- Lagrangian: Construct the Lagrangian for the problem, introducing a multiplier  $\lambda$  on the emissions constraint.
- KKT conditions: Derive the first-order conditions (stationarity, complementary slackness, feasibility).
- Solution:
  - Solve for the unconstrained optimum  $(q^u, a^u)$  and show whether the emissions constraint binds.
  - Solve for the constrained optimum  $(q^a, a^a)$  when the cap is  $E = 20$ .
- Verify that the second-order conditions for a maximum are satisfied at the constrained optimum.

- e. Interpretation of  $\lambda$ :
- What is the economic meaning of  $\lambda^*$  in this problem?
  - How does it relate to the marginal cost of tightening the emissions cap?
- f. Comparative statics: Suppose  $E$  decreases by one unit (stricter regulation). Use  $\lambda^*$  to predict the change in the firm's profit.
- g. Dual problem: Formulate the dual problem and interpret its objective function in this context.

## 5 Problem 5: Extend the Simulated Method of Moments Problem

Recall the SMM problem from class and the course notes. See more information on SMM [here](#)

The r script for the class example is in the modules material.

- a. Estimate the standard errors of the parameters. You can use asymptotic theory (Hessian) or bootstrap methods.
- b. The current SMM estimator uses an identity weighting matrix. Modify the code to use an optimal weighting matrix. You can use the sample variance-covariance matrix of the moments as an estimate of the optimal weighting matrix. Compare the estimates and standard errors of the parameters using the identity and optimal weighting matrices. Which one is more efficient? Why?