ROB 101: Computational Linear Algebra

Noah Peters

March 29, 2023

Abstract

Lecture notes for ROB 101 at the University of Michigan. LaTeXtemplate by Pingbang Hu.

Contents

Chapter 11

Solutions of Nonlinear Equations

Lecture 18: Review and Approximating Nonlinear Equations

11.1 Bisection Algorithm

14 Mar. 12:00

Theorem 11.1.1 (Intermediate Value Theorem). Suppose $f : \mathbb{R} \to \mathbb{R}$ is a continuous function and you know two real numbers, $a < b$, such that $f(a) \cdot f(b) < 0$. Then $\exists c \in \mathbb{R}$ such that:

$$a < c < b$$
$$f(c) = 0$$

Using the midpoint between two numbers to find the root may not give us the root right away. Further, the root isn't always exactly in between a and b. So, we can use the **bisection algorithm** to *approximate* roots:

Algorithm 11.1: Bisection Algorithm

```
Data: a < b ∈ ℝ such that f(a) · f(b) < 0
1 c = (a + b)/2 if f(c) = 0 then
2     return x* = c
3 else
4     if f(c) · f(a) < 0 then
5         b = c
6     else
7         a = c
8 Loop back to 1.
```

11.2 Derivatives and Approximation

Definition 11.2.1 (Derivative). A **derivative** is the slope of a function at a specific point.

There are 3 ways to represent the numerical approximation of a derivative:

1. **Forward Difference Approximation**: $\frac{df(x_0)}{dx} = \frac{f(x_0+h)-f(x_0)}{h}$
2. **Backward Difference Approximation**: $\frac{df(x_0)}{dx} = \frac{f(x_0)-f(x_0-h)}{h}$
3. **Symmetric Difference Approximation**: $\frac{df(x_0)}{dx} = \frac{f(x_0+h)-f(x_0-h)}{2h}$

Remark. If the 3 approximations above don't agree, then the limit does not exist and the function is not differentiable.

For a differentiable function, $f(x)$, $f(x) \approx f(x_0) + \frac{df(x_0)}{dx}(x - x_0)$ near point $(x_0, f(x_0))\,\forall x_0$. We can use this idea to *find roots* using linear approximations to nonlinear functions.

11.3 Newton's Method

To find the roots using linear approximations to nonlinear functions we can use **Newton's Method**:

Definition 11.3.1 (Newton's Method). Assume $f : \mathbb{R} \to \mathbb{R}$ is differentiable everywhere. Let x_k be our current estimate of a root, then:

$$f(x) \approx f(x_k) + \frac{df(x_k)}{dx}(x - x_k).$$

We want x_{k+1} such that $f(x_{k+1}) = 0$.

If we solve for x_{k+1} we get:

$$x_{k+1} = x_k - \frac{df(x_k)}{dx}^{-1} f(x_k).$$

This method takes very big "steps", so it may be more beneficial to take smaller "steps". This leads to the **damped Newton Method**,

$$x_{k+1} = x_k - \epsilon\Big(\frac{df(x_k)}{dx}\Big)^{-1} f(x_k),$$

where $0 < \epsilon < 1$ A typical value may be $\epsilon = 0.1$.

Lecture 19: Vectors and Approximating Nonlinear Equations

11.4 Vectors and Nonlinear Equations

16 Mar. 9:00

We can use vectors for linear approximations by understanding partial derivatives. Given nonlinear funcitons $f : \mathbb{R}^m \to \mathbb{R}^m$ we want the linear approximation at $x_0 \in \mathbb{R}^m$.

$$f(x) \approx f(x_0) + A(x - x_0)$$

where $A_{n\times m}, x, x_0 \in \mathbb{R}^m$, and $f(x_0), f(x) \in \mathbb{R}^n$. Here, A represents a matrix made up of partial derivatives.

Remark. Everything is the same as finding a linear approximation at a point. We are just replacing the slope with a matrix and the xs with vectors.

11.5 Partial Derivatives

Set $x = x_0 + he_i$ where he_i is some small outside adjustment to x_0 such that all x_0 remain the same except the i-th component.

$$x_0 + he_i = \begin{bmatrix} x_{01} \\ x_{02} \\ \vdots \\ x_{0m} \end{bmatrix} + h \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} x_{01} \\ x_{02} \\ \vdots \\ x_{0i} + h \\ \vdots \\ x_{0m} \end{bmatrix} \text{ where } h \text{ is small}$$

Equivalently, we have:

$$f(x_0 + he_i) \approx f(x_0) + A(x_0 + he_i - x_0) = f(x_0) + ha_i^{col}$$

where we can now solve for a_i^{col}, which represents the derivative of f with respect to x_i.

We can represent the numerical approximation of a partial derivative similar to how we represented the numerical approximation of a standard derivative. A partial derivative is represented with the mathematical symbol del: ∂.

11.6 Jacobian

Given the nonlinear functions $f : \mathbb{R}^m \to \mathbb{R}^m$, the **Jacobian** is

$$\frac{\partial f(x)}{\partial x} := \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \frac{\partial f(x)}{\partial x_2} \cdots \frac{\partial f(x)}{\partial x_m} \end{bmatrix}_{n \times m}.$$

- The partial derivatives are stacked to form a matrix
- For each $x \in \mathbb{R}^m$, $\frac{\partial f(x)}{\partial x}$ is an $n \times m$ matrix
- The **gradient** of $f : \mathbb{R}^m \to \mathbb{R}$ is a special Jacobian that for each $x \in \mathbb{R}^m$, $\nabla f(x)$ is a $1 \times m$ matrix

$f : \mathbb{R}^m \to \mathbb{R}^n$ looks like:

$$f(x) = \begin{bmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{bmatrix}$$

$\frac{\partial f(x)}{\partial x}$ looks like:

$$\frac{\partial f(x)}{\partial x} = \begin{bmatrix} \frac{\partial f_1(x)}{\partial x_1} & \frac{\partial f_1(x)}{\partial x_2} & \cdots & \frac{\partial f_1(x)}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n(x)}{\partial x_1} & \frac{\partial f_n(x)}{\partial x_2} & \cdots & \frac{\partial f_n(x)}{\partial x_m} \end{bmatrix}$$

The linear approximation of nonlinear functions $f : \mathbb{R}^m \to \mathbb{R}^n$ at point $x_0 \in \mathbb{R}^m$ is

$$f(x) \approx f(x_0) + A(x - x_0) = f(x_0) + \frac{\partial f(x_0)}{\partial x}(x - x_0).$$

Problem 11.6.1. We have two functions:

$$f_1(x_1, x_2) = \log(x_1) + \sqrt{x_2}$$
$$f_2(x_1, x_2) = x_1 \cdot x_2$$

and let $x_0 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Answer. $f : \mathbb{R}^2 \to \mathbb{R}^2$ so ⊛

Appendix