CS 357 - 14 Finite Difference Methods

Boyang Li (boyangl3)

Definition: For a differentiable function function $f(x) : \mathbb{R} \rightarrow \mathbb{R}$, the derivative is defined as

$$f'(x) = \lim_{x \to 0} \frac{f(x+h) - f(x)}{h}$$

The finite difference method is used to approximate **derivative at given point**, $df(x)$, define the **forward finite difference method** as

$$f'(x) \approx df(x) = \frac{f(x+h) - f(x)}{h}$$

where h is called **perturbation** and usually h is a small amount.

Introduction of this function: From Taylor's expansion of $f(x+h)$ and then take the derivative.

$$f(x+h) = f(x) + f'(x) \cdot h + \frac{1}{2} f''(x) \cdot h^2 + \frac{1}{n} f^n(x) \cdot h^n = f(x) + f'(x) \cdot h + \mathcal{O}(h^2)$$

Simplify the equation and extract $f'(x)$, we finally get the approximation

$$f'(x) = \frac{f(x+h) - f(x)}{h} + \mathcal{O}(h)$$

Error: The error is equal to the exact derivative and the approximated derivative,

$$Error = |f'(x) - df(x)|$$

3 Types of Finite Difference Methods: There are 3 different types of finite difference methods, all of them can be used to estimate derivatives.

- **Forward** Finite Difference Method: $df(x) = \dfrac{f(x+h) - f(x)}{h}$
- **Backward** Finite Difference Method: $df(x) = \dfrac{f(x) - f(x-h)}{h}$
- **Central** Finite Difference Method: $df(x) = \dfrac{f(x+h) - f(x-h)}{2h}$

Note: The central finite difference method has the smallest error bound.

Gradient Approximation ($f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$): For a function that takes a vector and maps into a real number, the derivative is the **gradient** of that function.

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \\ \vdots \\ \dfrac{\partial f}{\partial x_n} \end{bmatrix}$$

We can approximate the gradient by applying the finite different method to every entry:

$$\nabla_F Df(\mathbf{x}) = \begin{bmatrix} df(x_1) \\ df(x_2) \\ \vdots \\ df(x_n) \end{bmatrix} = \begin{bmatrix} \dfrac{f(\mathbf{x}+h\boldsymbol{\delta}_1) - f(\mathbf{x})}{h} \\ \dfrac{f(\mathbf{x}+h\boldsymbol{\delta}_2) - f(\mathbf{x})}{h} \\ \vdots \\ \dfrac{f(\mathbf{x}+h\boldsymbol{\delta}_3) - f(\mathbf{x})}{h} \end{bmatrix}$$

Note: $\boldsymbol{\delta}_i$ is a vector with a 1 at the ith position and 0 elsewhere. For example, $\mathbf{x} = [1, 3, 5]$ then $\boldsymbol{\delta}_2 = [0, 1, 0]$.

Jacobian Approximation ($f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^m$): For a function that takes a vector and maps to another vector, the derivative is the **Jacobian** of that function.

$$\mathbb{J}(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \dots & \dfrac{\partial f_1}{\partial x_n} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \dots & \dfrac{\partial f_2}{\partial x_n} \\ & \ddots & & \\ \dfrac{\partial f_m}{\partial x_1} & \dfrac{\partial f_m}{\partial x_2} & \dots & \dfrac{\partial f_m}{\partial x_n} \end{bmatrix}$$

We can approximate the gradient by applying the finite different method to every entry:

$$\mathbb{J}_{FD}(\mathbf{x}) = \begin{bmatrix} df_1(x_1) & df_1(x_2) & \dots & df_1(x_n) \\ df_2(x_1) & df_2(x_2) & \dots & df_2(x_n) \\ & \ddots & & \\ df_m(x_1) & df_m(x_2) & \dots & df_m(x_n) \end{bmatrix}$$