Optimization Techniques in Machine Learning

Your Name

1 Introduction

This document covers the application of optimization techniques in machine learning.

2 Gradient Descent

2.1 Example Problem

Use gradient descent to minimize the function $f(x) = x^2 + 4x + 4$.

Solution

The gradient of $f(x)$ is:

$$\nabla f(x) = \frac{d}{dx}(x^2 + 4x + 4) = 2x + 4.$$

Starting from an initial guess $x_0 = 0$ and using a learning rate $\alpha = 0.1$, the update rule for gradient descent is:

$$x_{n+1} = x_n - \alpha \nabla f(x_n).$$

First iteration:

$$x_1 = x_0 - 0.1 \cdot \nabla f(x_0) = 0 - 0.1 \cdot (2 \cdot 0 + 4) = 0 - 0.4 = -0.4.$$

Second iteration:

$$x_2 = x_1 - 0.1 \cdot \nabla f(x_1) = -0.4 - 0.1 \cdot (2 \cdot -0.4 + 4) = -0.4 - 0.1 \cdot (0.8 + 4) = -0.4 - 0.48 = -0.88.$$

Third iteration:

$$x_3 = x_2 - 0.1 \cdot \nabla f(x_2) = -0.88 - 0.1 \cdot (2 \cdot -0.88 + 4) = -0.88 - 0.1 \cdot (-1.76 + 4) = -0.88 - 0.1 \cdot 2.24 = -1.104.$$

Continuing this process, the value of x will converge towards the minimum point. For $f(x) = x^2 + 4x + 4$, the minimum occurs at $x = -2$, as verified by setting the gradient to zero:

$$2x + 4 = 0 \implies x = -2.$$

Thus, using gradient descent with the given learning rate, we approach the solution $x = -2$ iteratively.