Lecture Summary: Direction of Steepest Ascent and Descent

Source: Week 11-Lec 01.pdf

Key Points

- **Motivation and Example:**
  - The problem of determining how water flows down a hill is modeled by studying the direction in which altitude decreases most rapidly.
  - In one dimension, the direction of steepest descent is determined by the sign of the derivative of the altitude function $h(x)$.
  - Example: A cross-section of the Deccan Plateau illustrates water flowing towards regions of lower altitude based on the slope.
- **From One Dimension to Two Dimensions:**
  - In two dimensions, altitude is modeled using $h(x, y)$.
  - The rate of change in any direction is described using the **directional derivative**.
  - The direction of steepest descent corresponds to minimizing the directional derivative, while steepest ascent maximizes it.
- **Directional Derivative and Gradient:**
  - The directional derivative of f at $\vec{a}$ in the direction of a unit vector $\vec{u}$ is:

    $$D_{\vec{u}}f(\vec{a}) = \nabla f(\vec{a}) \cdot \vec{u} = \|\nabla f(\vec{a})\| \cos(\theta),$$

    where θ is the angle between $\nabla f(\vec{a})$ and $\vec{u}$.
  - **Steepest Ascent:** Occurs when $\cos(\theta) = 1$, i.e., $\vec{u}$ is in the same direction as $\nabla f(\vec{a})$.
  - **Steepest Descent:** Occurs when $\cos(\theta) = -1$, i.e., $\vec{u}$ is in the opposite direction to $\nabla f(\vec{a})$.
  - **No Change:** Occurs when $\cos(\theta) = 0$, i.e., $\vec{u}$ is orthogonal to $\nabla f(\vec{a})$.
- **Examples:**
  - Example 1: For $f(x, y) = \sin(xy)$ at $(\pi, 1)$:
    * Gradient: $\nabla f(\pi, 1) = (-1, -\pi)$.
    * Steepest ascent: Direction is $-\nabla f/\|\nabla f\| = (1/\sqrt{1 + \pi^2}, \pi/\sqrt{1 + \pi^2})$.
    * Steepest descent: Direction is $-\nabla f(\pi, 1) = (-1, -\pi)/\sqrt{1 + \pi^2}$.
    * No change: Perpendicular directions such as $(\pi, -1)$ or $(-\pi, 1)$.
  - Example 2: For $f(x, y, z) = x^2 + y^2 + z^2$ at $(1, 1, 1)$:
    * Gradient: $\nabla f(1, 1, 1) = (2, 2, 2)$.
    * Steepest ascent: Direction is $\nabla f/\|\nabla f\| = (2/\sqrt{12}, 2/\sqrt{12}, 2/\sqrt{12})$.
    * Steepest descent: Opposite to the gradient.
    * No change: Perpendicular directions such as $(1, -1, 0)$ or $(1/\sqrt{2}, -1/\sqrt{2}, 0)$.

- **Continuity of Gradient:**
  - The gradient must be continuous in a neighborhood around the point of interest for these results to hold.
  - If the gradient is not continuous, the directional derivatives may not exist in certain directions, and the results cannot be applied.
- **Application in Machine Learning:**
  - Gradient Descent: An optimization technique that iteratively moves in the direction of steepest descent to minimize a function.
  - This method is widely used in training machine learning models.

Simplified Explanation

Steepest Ascent/Descent: The gradient ∇f points in the direction of steepest ascent, while $-\nabla f$ points in the direction of steepest descent.

Example: For $f(x, y) = \sin(xy)$ at $(\pi, 1)$: - Gradient: $(-1, -\pi)$. - Steepest ascent: $(1/\sqrt{1+\pi^2}, \pi/\sqrt{1+\pi^2})$. - Steepest descent: Opposite direction to ascent.

Conclusion

In this lecture, we:

- Analyzed steepest ascent, descent, and directions of no change using the gradient.
- Applied these concepts to practical examples.
- Highlighted their significance in optimization and machine learning.

Understanding steepest ascent/descent and their relation to gradients is foundational for optimization in mathematics and applied sciences.