Direction of Steepest Descent

The Directional Derivative Gives the Rate of Change in Some Direction

Definition: Directional Derivative

Suppose $f : \mathbb{R}^n \to \mathbb{R}$ has continuous partial derivatives[a]. Then the rate of change of f at $x \in \mathbb{R}^n$ in the direction $p \in \mathbb{R}^n$ is given by the directional derivative

$$D_p f(x) = (\nabla f(x))^T \frac{p}{\|p\|}$$

[a] One can define the directional derivative without requiring continuous partial derivatives, but this assumption is useful for our purposes.

Which Direction is “Best”?

Which direction gives the most rapid rate of change in f?

The Negative Gradient is the Steepest Descent Direction

Definition: Steepest Descent Direction

The steepest descent direction at $x \in \mathbb{R}^n$ of a function $f : \mathbb{R}^n \to \mathbb{R}$ with continuous partial derivatives is given by

$$p_{\mathrm{SD}} \equiv -\nabla f(x)$$

The Steepest Descent Method: first cut

Algorithm: Steepest Descent Method

Data: $x_0 \in \mathbb{R}^n$ (initial guess)

Result: x^* (local minimum)

for $k = 0, 1, 2, \ldots$ **do**

  if $\|\nabla f_k\| \leq \epsilon_r \|\nabla f_0\| + \epsilon_a$ **then** return

  set $p_k \leftarrow -\nabla f_k / \|\nabla f_k\|$

  update $x_{k+1} \leftarrow x_k + \alpha p_k$

end

Problem: what is α?

The parameter α is called the step length.

The "ideal" choice for α is to solve the following 1-dimensional problem:

$$\min_{\alpha} \phi(\alpha) = f(x_k + \alpha p_k).$$