Line Search Methods

The Steepest Descent Method: second 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 $p_k \leftarrow -\nabla f_k / \|\nabla f_k\|$

  find α_k satisfying the Wolfe Conditions

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

end

The Steepest Descent Method is an Example of a Line-Search Method

Algorithm: Generic Line-Search Method

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

Result: x (local minimum)

for $k = 1, 2, 3, \ldots$ **do**

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

  compute search direction p_k that is a descent direction

  find α_k satisfying the Armijo or Wolfe Conditions

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

end

Whatever Direction We Choose, It Should Be a Descent Direction

Definition: Descent Direction

The direction $p \in \mathbb{R}^n$ is a descent direction for the function $f : \mathbb{R}^n \to \mathbb{R}$ at the point $x \in \mathbb{R}^n$ if

$$\nabla f(x)^T p < 0.$$

Illustration of Descent Directions

Why Would We Want Step Directions Other Than $-\nabla f_k$?