Gradients and Gradient Descent

Gradient Descent Algorithm

- Pick an initial point x_0
- Iterate until convergence

$$x_{t+1} = x_t - \gamma_t \nabla f(x_t)$$

where γ_t is the t^{th} step size (sometimes called learning rate)

Empirical Gradient

In other cases, the objective / evaluation function might not be available in a differentiable form

Gradients and Learning Rate / Step

- The value of the gradient

Searching in Partially Obserable and Nondeterministic Environments