Recitation 1

Gradients and Directional Derivatives

Brett Bernstein

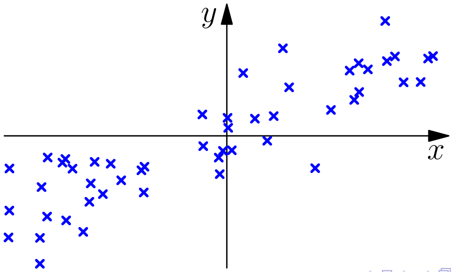
CDS at NYU

January 25, 2017

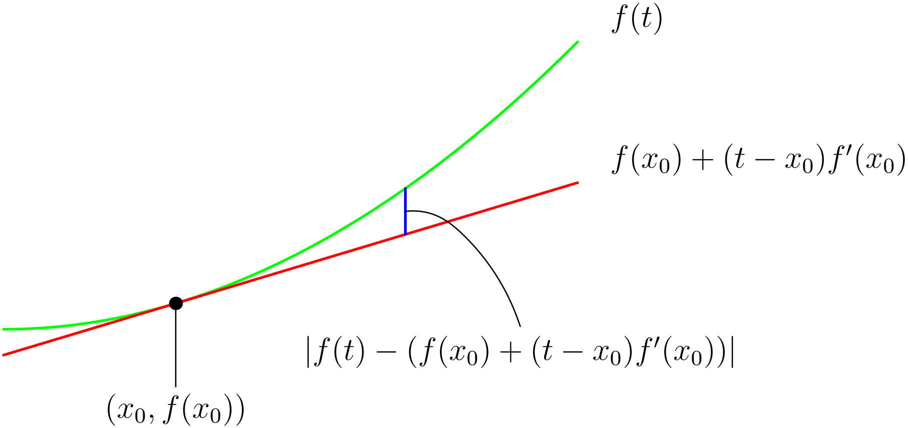
Intro Question

Question

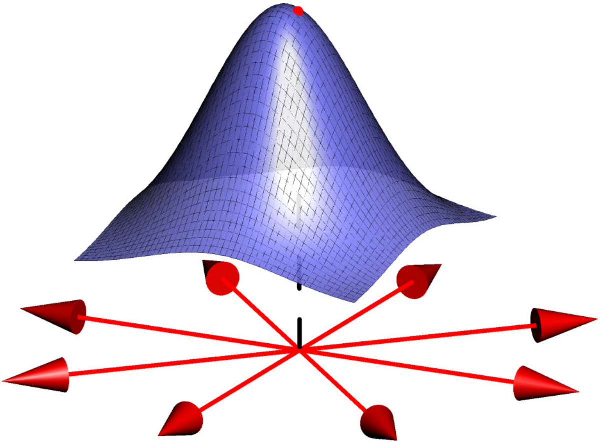
We are given the data set $(x_1, y_1), \ldots, (x_n, y_n)$ where $x_i \in \mathbb{R}^d$ and $y_i \in \mathbb{R}$. We want to fit a linear function to this data by performing empirical risk minimization. More precisely, we are using the hypothesis space $\mathcal{F} = \{f(x) = w^T x \mid w \in \mathbb{R}^d\}$ and the loss function $\ell(a, y) = (a - y)^2$. Given an initial guess $\tilde{w}$ for the empirical risk minimizing parameter vector, how could we improve our guess?



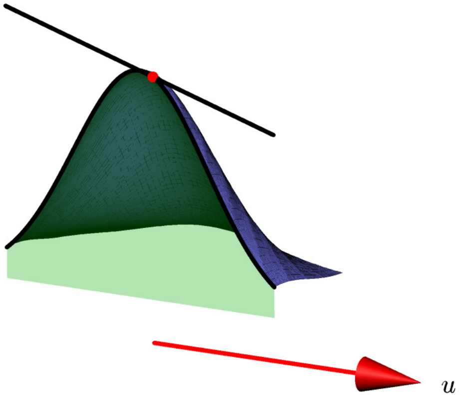
1D Linear Approximation By Derivative



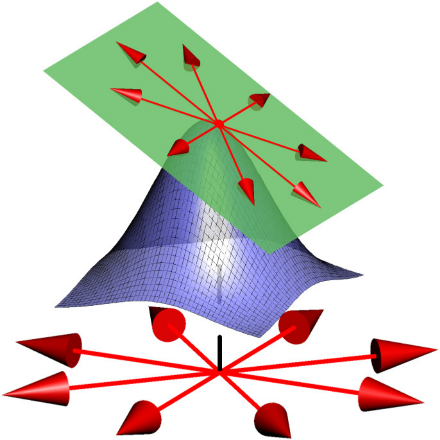
Multiple Possible Directions for $f : \mathbb{R}^2 \to \mathbb{R}$



Directional Derivative as a Slope of a Slice



Tangent Plane for $f : \mathbb{R}^2 \to \mathbb{R}$



Critical Points of $f : \mathbb{R}^2 \to \mathbb{R}$