Visualizing Gradient Descent

16-385 Computer Vision (Kris Kitani)

Carnegie Mellon University
(partial) derivatives tell us how much one variable affects another
Two ways to think about them:

- Slope of a function
- Knobs on a machine
1. Slope of a function:

\[
\frac{\partial f(x)}{\partial x} = \begin{bmatrix} \frac{\partial f(x)}{\partial x}, & \frac{\partial f(x)}{\partial y} \end{bmatrix}
\]
describes the slope around a point
2. Knobs on a machine:

input $x$ \rightarrow \text{output } f(x; w)$

describes how each ‘knob’ affects the output

\[ \frac{\partial f(x)}{\partial w_1}, \frac{\partial f(x)}{\partial w_2}, \frac{\partial f(x)}{\partial w_3} \]

small change in parameter $\Delta w_1$ \rightarrow output will change by $\frac{\partial f(x)}{\partial w_1} \Delta w_1$
Gradient descent:

Given a fixed-point on a function, move in the direction opposite of the gradient.
Gradient descent:

\[ \mathbf{w} = \mathbf{w} - \nabla \mathbf{w} \]

update rule: