

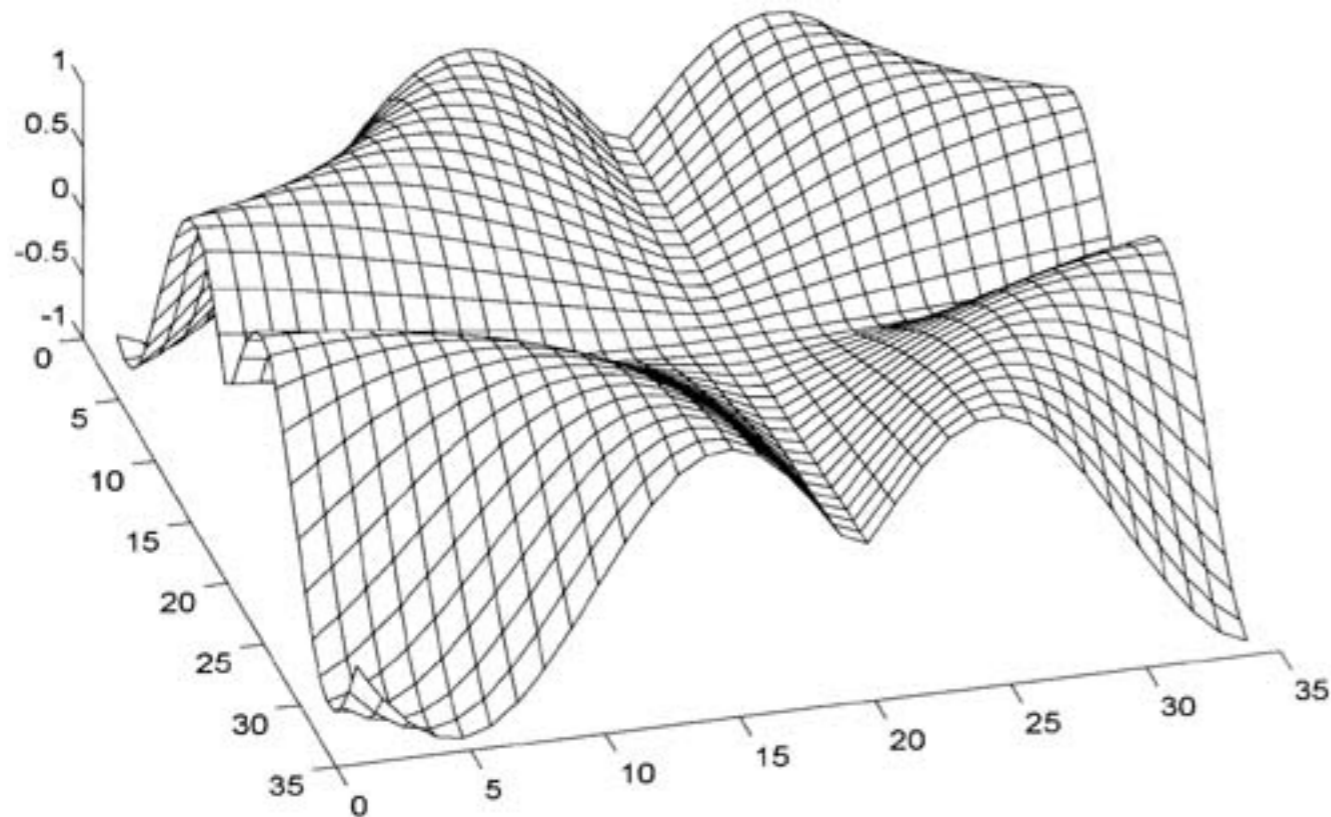
Visualizing Gradient Descent

16-385 Computer Vision (Kris Kitani)

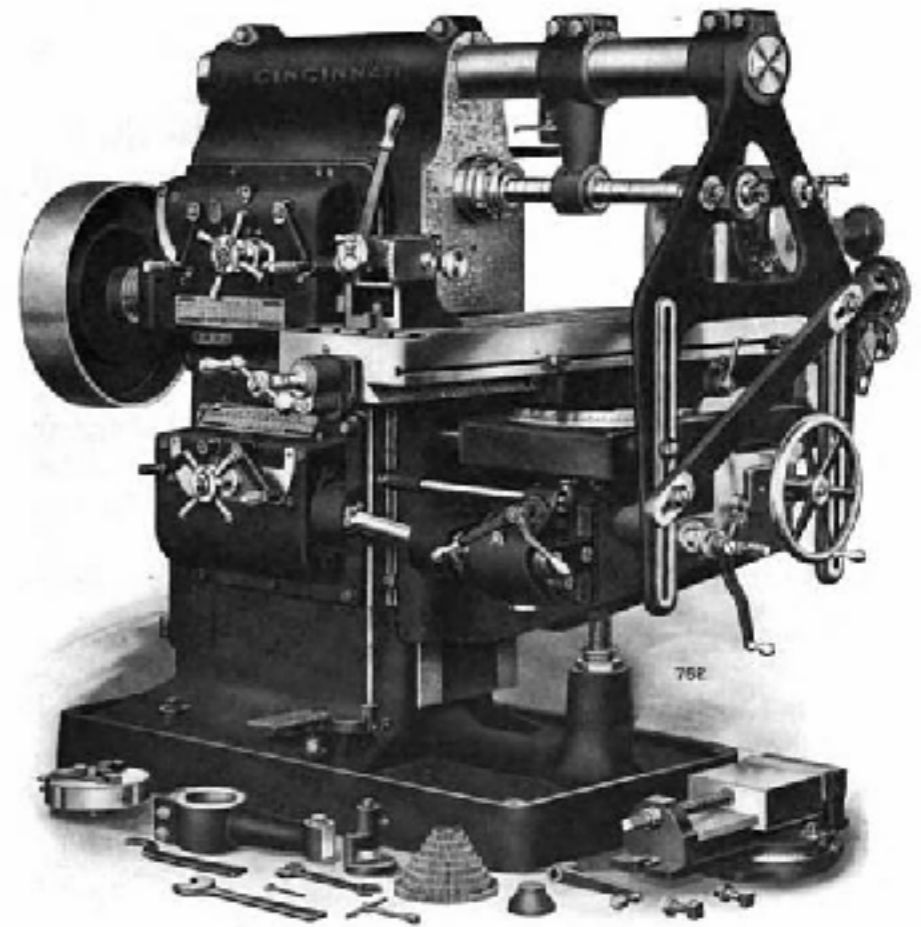
Carnegie Mellon University

(partial) derivatives tell us how much one variable affects another

Two ways to think about them:

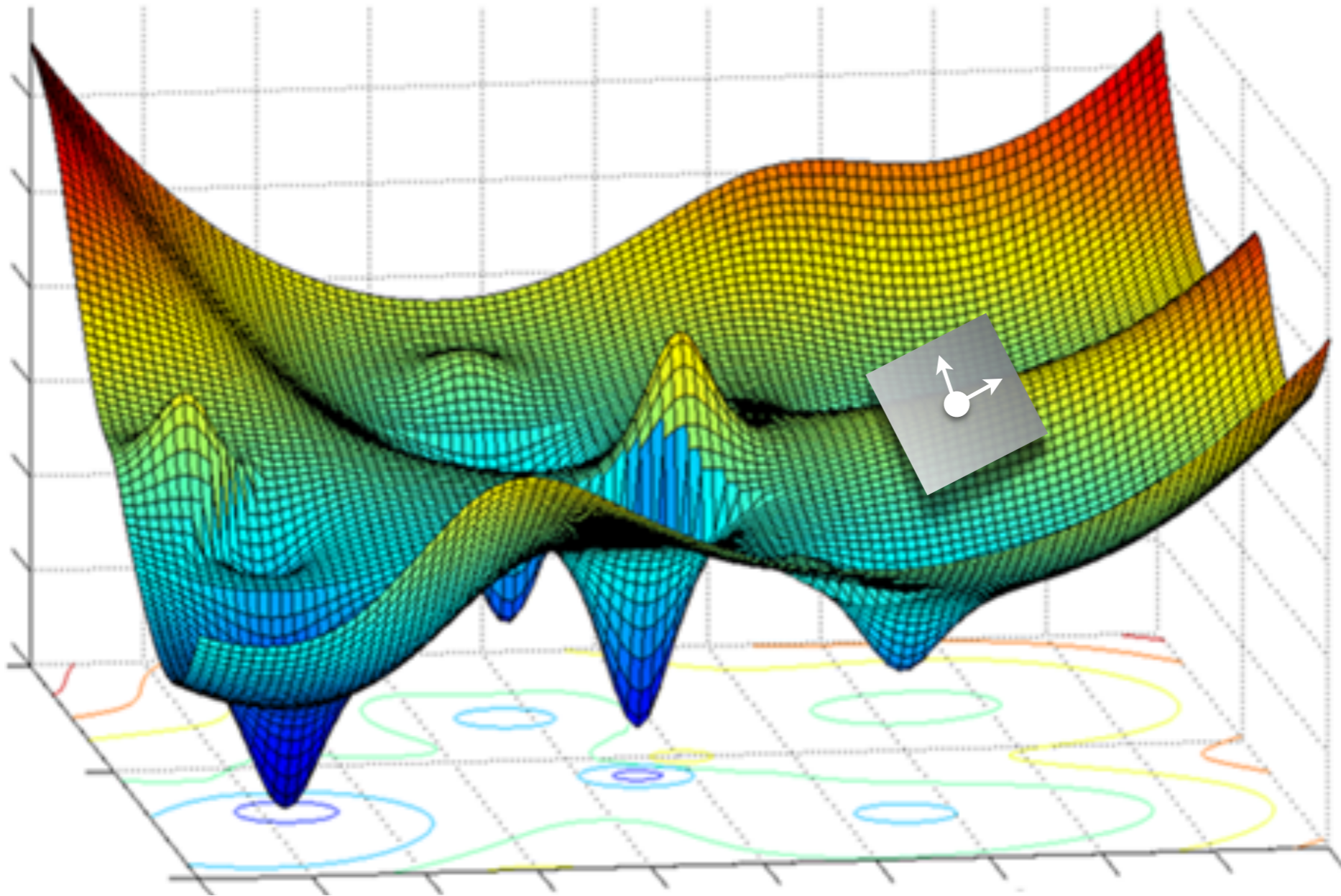


Slope of a function



Knobs on a machine

1. Slope of a function:



$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \left[\frac{\partial f(\mathbf{x})}{\partial x}, \frac{\partial f(\mathbf{x})}{\partial y} \right] \quad \text{describes the slope around a point}$$

2. Knobs on a machine:



describes how each
'knob' affects the
output



$$\frac{\partial f(x)}{\partial w_1}$$



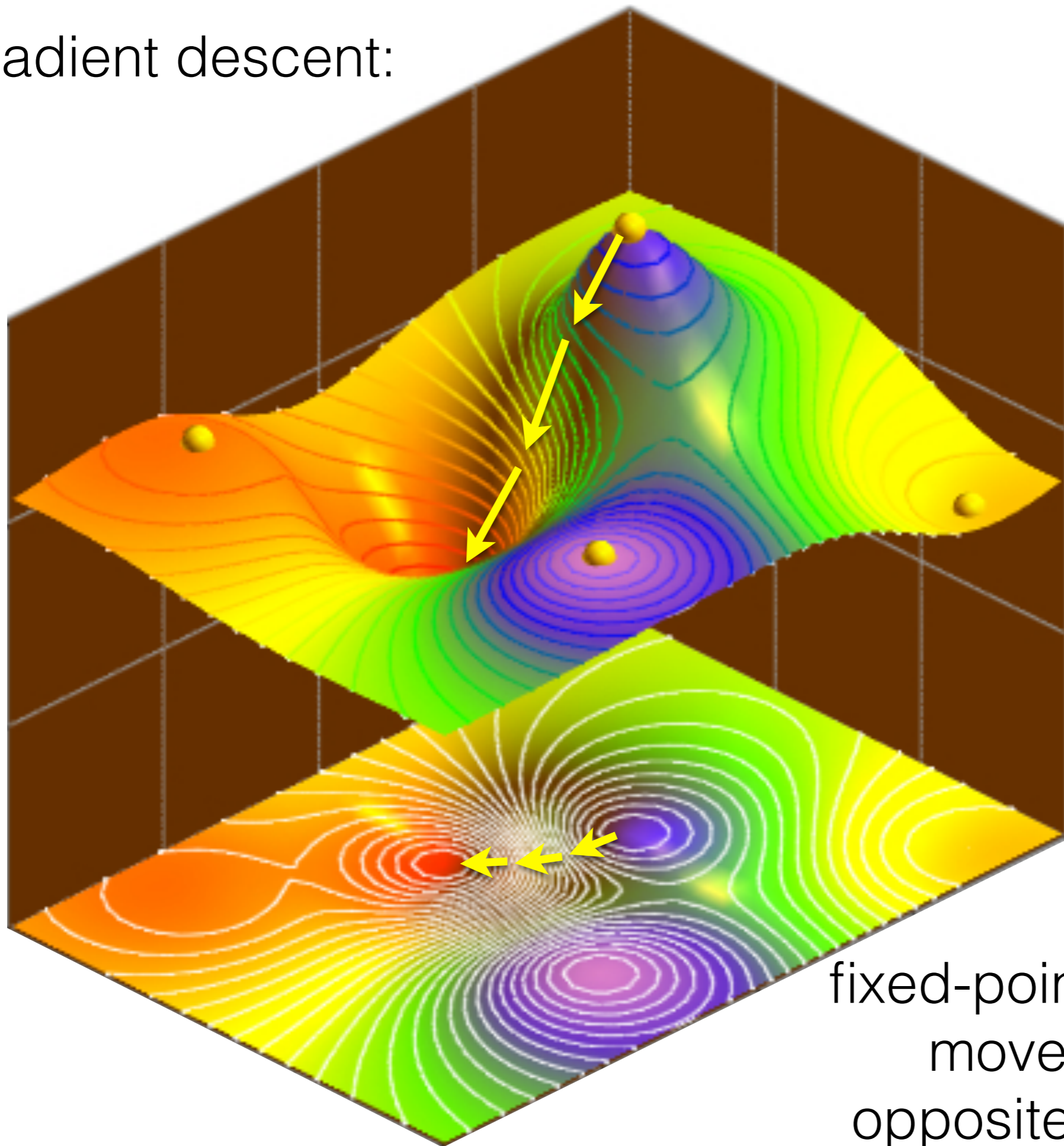
$$\frac{\partial f(x)}{\partial w_2}$$



$$\frac{\partial f(x)}{\partial w_3}$$

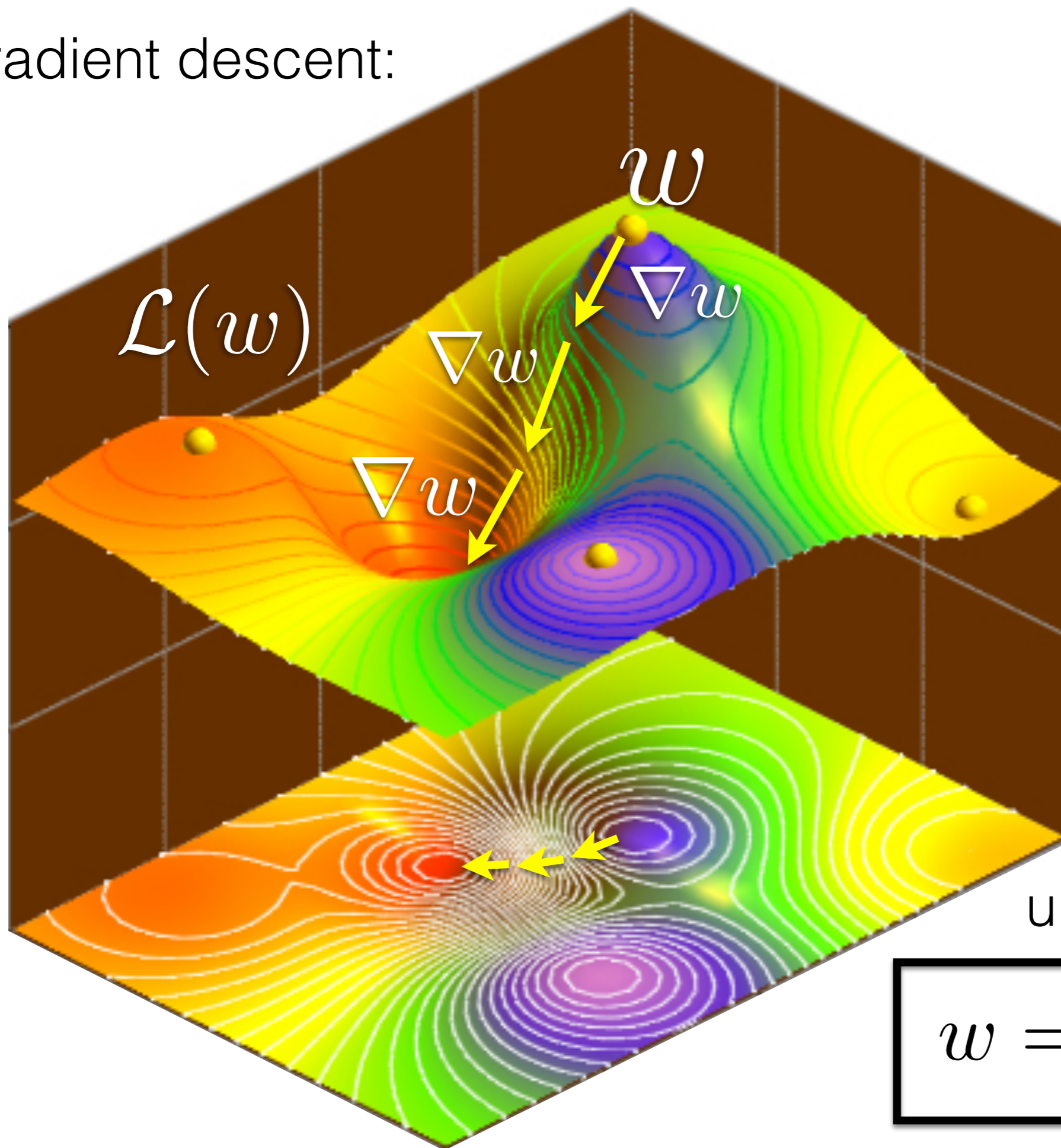
small change in parameter Δw_1 → output will change by $\frac{\partial f(x)}{\partial w_1} \Delta w_1$

Gradient descent:



Given a
fixed-point on a function,
move in the direction
opposite of the gradient

Gradient descent:



update rule:

$$w = w - \nabla w$$