DERIVATIVES, GRADIENTS AND HESSIANS

Derivatives

The **derivative** of a function $f(x)$, written as $f'(x)$ or $\frac{df}{dx}(x)$ measures how rapidly the function is increasing

Formally, this measures the slope of the tangent to the function at any point

- A negative derivative at x means the function is decreasing at x

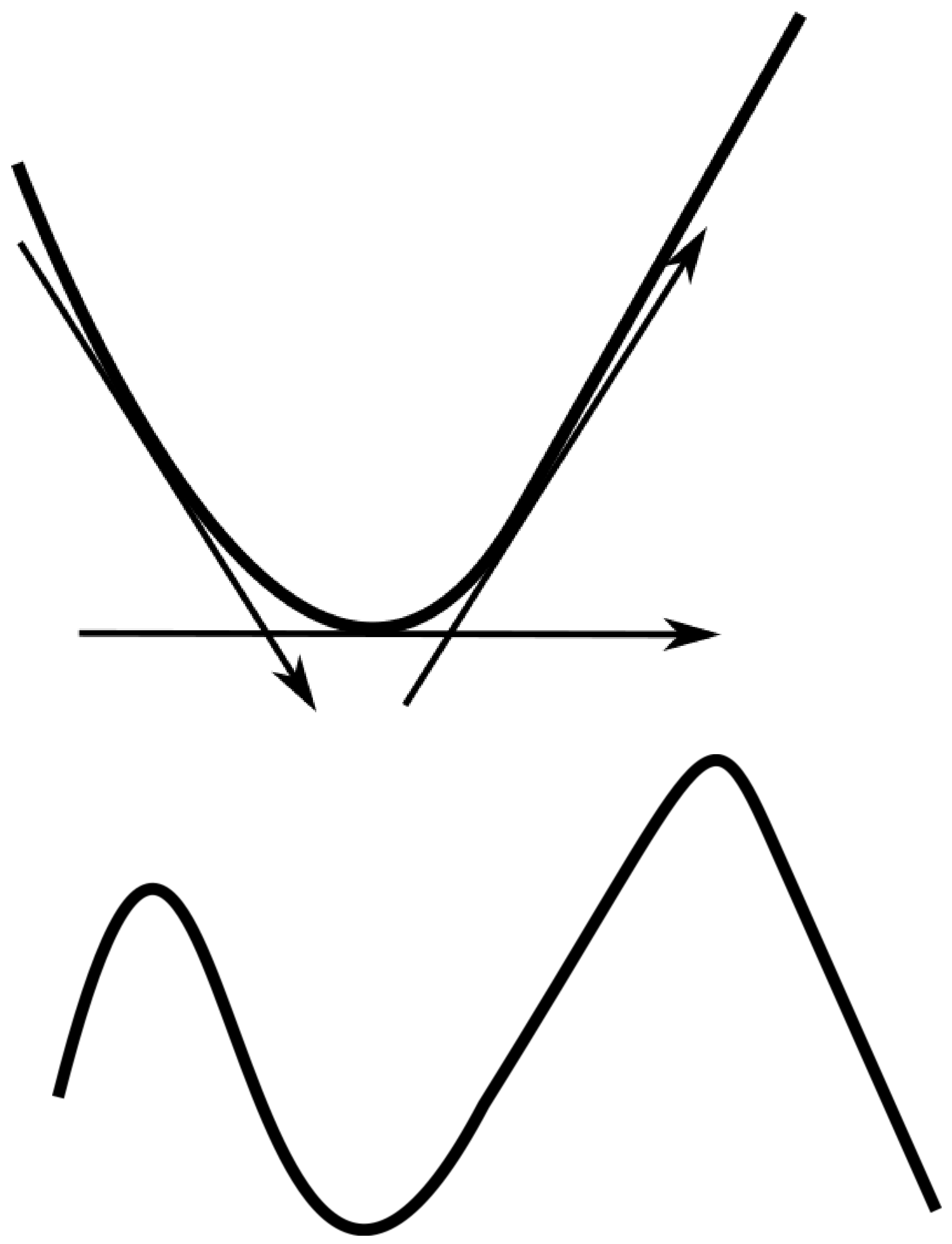
At the maximum or minimum of a function, the derivative is 0

These ideas allow algorithms to optimize/find extrema of functions

Note: these usually can only find **local** (not **global**) optima

The second derivative is the derivative of the derivative, and measures the curvature of the function

At a minimum, the second derivative is positive, and at a maximum it is negative



(top) Derivatives of a function, (bottom) A function with two local maxima and one global maximum

Gradients

Consider functions with d-dimensional inputs and scalar outputs

E.g. a loss function that takes as input a vector of weights/parameters and returns a measure of model fit

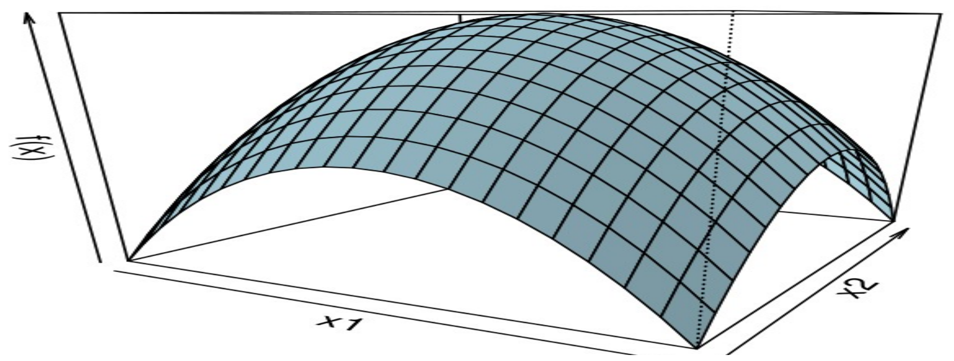
The analog of a derivative for functions with multidimensional inputs is the gradient $\nabla f(x)$

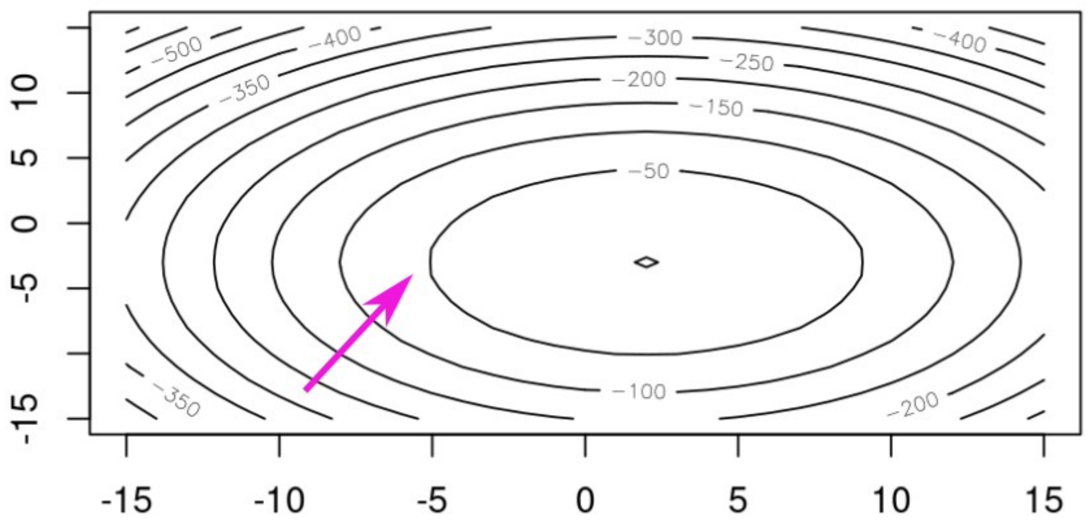
This is a d-dimensional vector of partial derivatives $\left[\frac{\partial f}{\partial x_1}(x), \frac{\partial f}{\partial x_2}(x) \cdots \frac{\partial f}{\partial x_d}(x)\right]$

- The i^{th} component is the derivative with respect to x_i keeping other components fixed

The inner product of the gradient with any unit-length vector is the rate of change of f in the direction of that vector

- The gradient itself points in the direction of steepest ascent





(above) two representations of a function with a 2-d input, along with a gradient vector shown in magenta

Hessians

Analogous to the second-order derivative is the **Hessian**

This is a $d \times d$ symmetric matrix whose (j,k)th element is $\frac{\partial^2 f}{\partial x_j \partial x_k}(x)$

At the minima and maxima of f, all components of the gradient vector are 0

At any minimum, the Hessian is **positive definite**, at any maximum it is **negative definite**

A matrix H is positive definite if

- All its eigenvalues are 0, or
- For any vector, v we have $v^T H v^T$ > 0