

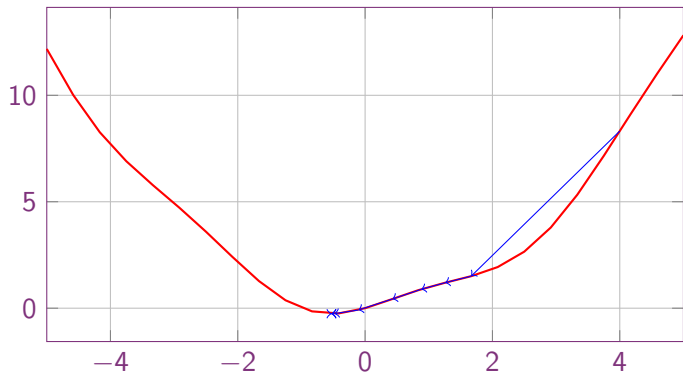
Neural Networks and Backpropagation Lecture

Grégoire Payen de La Garanderie

Durham University

February 17, 2020

Neural Networks and Function Minimisation

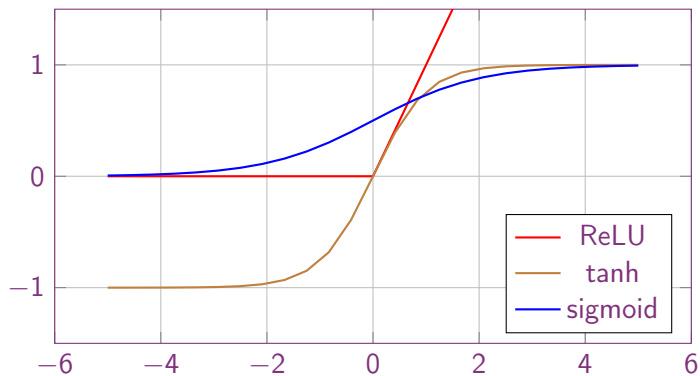


Function of one real-valued variable

$$f: \mathbb{R} \rightarrow \mathbb{R}$$

$$x \mapsto f(x)$$

Activation functions

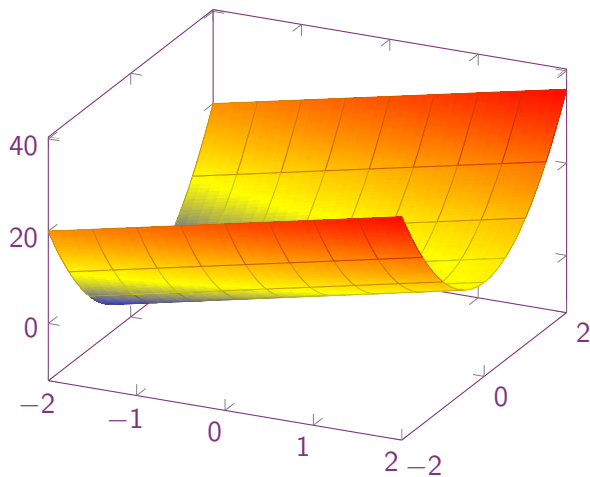


Function of two real-valued variables

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}$$

$$x, y \mapsto f(x, y)$$

Example: $f(x, y) = 4x + 7y^2$



Function of n real-valued variables

$$f: \mathbb{R}^n \rightarrow \mathbb{R}$$

$$x_1, \dots, x_n \mapsto f(x_1, \dots, x_n)$$

Jacobian

$$f: \mathbb{R}^n \rightarrow \mathbb{R}^m$$

$$x_1, \dots, x_n \mapsto (y_1, \dots, y_m) = f(x_1, \dots, x_n)$$

Jacobian

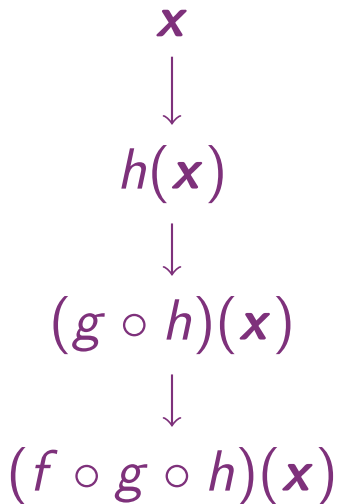
$$J_f = \frac{df}{d\mathbf{x}} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

Chain Rule

Given $h(x) = f(g(x))$

$$h'(x) = f'(g(x)) \times g'(x)$$

Example: A simple network



Function of multiple vectors

$$f: \mathbb{R}^s \times \mathbb{R}^t \rightarrow \mathbb{R}^u$$

$$x, y \mapsto f(x, y)$$

Example: Multilinear Map

$$f: \mathbb{R}^n \times \mathbb{R}^{m \times n} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$$

$$\mathbf{x}, \mathbf{W}, \mathbf{b} \quad \mapsto \quad \mathbf{W}\mathbf{x} + \mathbf{b}$$

