Artificial Neural Networks

John Stachurski

2025

Topics

- Foo
- Bar

Aim: approximate an unknown functional relationship

$$y = f(x) \qquad (x \in \mathbb{R}^k,\ y \in \mathbb{R})$$

Examples.

- $x =$ cross section of returns, $y =$ return on oil futures tomorrow
- $x =$ weather sensor data, $y =$ max temp tomorrow

Problem:

- observe $(x_i, y_i)_{i=1}^n$ and seek f such that $y_{n+1} \approx f(x_{n+1})$

Nonlinear regression: choose model $\{f_\theta\}_{\theta\in\Theta}$ and minimize the empirical loss

$$\ell(\theta) := \sum_{i=1}^{n}(y_i - f_\theta(x_i))^2 \quad \text{s.t.} \quad \theta \in \Theta$$

In the case of ANNs, we consider all f_θ having the form

$$f_\theta = \sigma \circ A_m \circ \cdots \circ \sigma \circ A_2 \circ \sigma \circ A_1$$

where

- $A_j x = W_j x + b_j$ is an affine map
  - $W_j =$ "weights" and $b_j =$ "bias"
- σ is a nonlinear "activation" function

Nonlinear regression: choose model $\{f_\theta\}_{\theta\in\Theta}$ and minimize the empirical loss

$$\ell(\theta) := \sum_{i=1}^{n}(y_i - f_\theta(x_i))^2 \quad \text{s.t.} \quad \theta \in \Theta$$

In the case of ANNs, we consider all f_θ having the form

$$f_\theta = \sigma \circ A_m \circ \cdots \circ \sigma \circ A_2 \circ \sigma \circ A_1$$

where

- $A_j x = W_j x + b_j$ is an affine map
  - $W_j =$ "weights" and $b_j =$ "bias"
- σ is a nonlinear "activation" function