KAN Tutorial Slides

Daniel Precioso Garcelán

October 3, 2025

Kolmogorov-Arnold Networks

Kolmogorov-Arnold Representation Theorem

Let $\Omega \subset \mathbb{R}^d$ be a bounded domain and let $f : \Omega \to \mathbb{R}$ be a continuous function; i.e. $f \in C(\Omega)$.

Then there exist continuous univariate functions

$$\Phi_q : \mathbb{R} \to \mathbb{R}, \quad q = 1, \ldots, 2d+1;$$

and continuous univariate functions

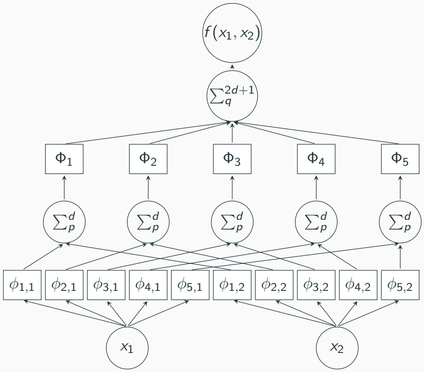
$$\phi_{pq} : \mathbb{R} \to \mathbb{R}, \quad p = 1, \ldots, d; \quad q = 1, \ldots, 2d+1;$$

such that for every $\mathbf{x} = (x_1, \ldots, x_d) \in \Omega$,

$$f(\mathbf{x}) = \sum_{q=1}^{2d+1} \Phi_q \left(\sum_{p=1}^{d} \phi_{pq}(x_p) \right).$$

Kolmogorov-Arnold Representation Theorem

For a bivariate function $f(x_1, x_2)$, we then have:



Which is a simple KAN!

B-Splines

Φ_q and $\phi_{p,q}$ can be chosen from any family of continuous univariate functions. A common choice is the **B-spline** family.

A B-spline of degree k is defined as:

$$B_k(x) = \sum_{i=0}^{n-k-1} P_i N_{i,k}(x)$$

where n is the number of control points (length of the knot vector), and P_i are the basis function weights.

And $N_{i,k}$ are the basis functions, following the standard **Cox-de Boor recursive definition**:

$$N_{i,0}(x) = \begin{cases} 1, & t_i \leq x < t_{i+1} \end{cases}$$

Grid Extension

Continuous Learning

B-splines are made of basin functions that operate in small bounds of X. Learning new information in a part of X does not alter other regions of X. In contrary, traditional MLP risk **catastrohpic forgetting**.

Symbolic Regression

Sparsibility