Universality of Parameterizations

Seohong Park

July 2024

Inner-product parameterization. We know that any continuous function $f : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ defined on a compact topological space $\mathcal{S}$ can be uniformly approximated by $\varphi(s)^\top \psi(g)$. In other words, for any $\varepsilon > 0$, there exist $n \geq 1$, $\varphi : \mathcal{S} \to \mathbb{R}^n$, and $\psi : \mathcal{S} \to \mathbb{R}^n$ such that $|f(s,g) - \varphi(s)^\top \psi(g)| < \varepsilon$ for all $s, g \in \mathcal{S}$. This can be easily proven by the Stone–Weierstrass theorem. The intuition is that any continuous function can be approximated by a polynomial to arbitrary accuracy, and any bivariate polynomial can be represented in the form $\varphi(s)^\top \psi(g)$.

Asymmetric norm parameterization. Then, what about $\|\varphi(s) - \psi(g)\|$? Is it a universal approximator for any non-negative continuous function f? We show that this is not the case, regardless of the norm $\|\cdot\|$. Let f be a non-negative continuous function satisfying $f(s,s) = 0$ for all $s \in \mathcal{S}$, and $f(x,y) = 1$ and $f(y,x) = 2$ for some $x, y \in \mathcal{S}$. Assume that a universal approximator exists and let φ and ψ be functions satisfying $|f(s,g) - \|\varphi(s) - \psi(g)\|| < \varepsilon$ for all $s, g \in \mathcal{S}$, for a given $\varepsilon > 0$. Then, $\|\varphi(s) - \psi(s)\| < \varepsilon$ for all $s \in \mathcal{S}$. However, we have

$$\begin{aligned}
2 &= f(y,x) \\
&< \|\varphi(y) - \psi(x)\| + \varepsilon \\
&\leq \|\varphi(y) - \psi(y)\| + \|\psi(y) - \varphi(x)\| + \|\varphi(x) - \psi(x)\| + \varepsilon \\
&< \|\psi(y) - \varphi(x)\| + 3\varepsilon \\
&< f(x,y) + 4\varepsilon \\
&= 1 + 4\varepsilon,
\end{aligned}$$

and setting $\varepsilon = 1/4$ leads to a contradiction.

"Rotated" symmetric inner-product parameterization. What about $\varphi(s)^\top A \varphi(g)$ with $n \geq 1$, $A \in \mathbb{R}^{n \times n}$, and $\varphi : \mathcal{S} \to \mathbb{R}^n$? Is this symmetric inner-product parameterization a universal approximator for any (potentially asymmetric) continuous function by additionally having a fixed $n \times n$ "rotation" matrix A? This is true, and here's a one-line proof. Let $n = 2m$,

$$\begin{aligned}
\varphi(s) &= \begin{bmatrix} \psi(s) \\ \theta(s) \end{bmatrix}, \\
A &= \begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix},
\end{aligned}$$

where $\psi, \theta : \mathcal{S} \to \mathbb{R}^m$, I is the $m \times m$ identity matrix, and 0 is the $m \times m$ zero matrix. Then, $\varphi(s)^\top A \varphi(g) = \psi(s)^\top \theta(g)$, and we know that this is universal.