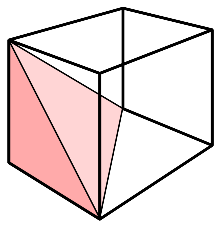
Measuring the "size" of hypothesis spaces over $\mathbb{R}$ from the perspective of No Free Lunch

Yan King Yin

May 2, 2024

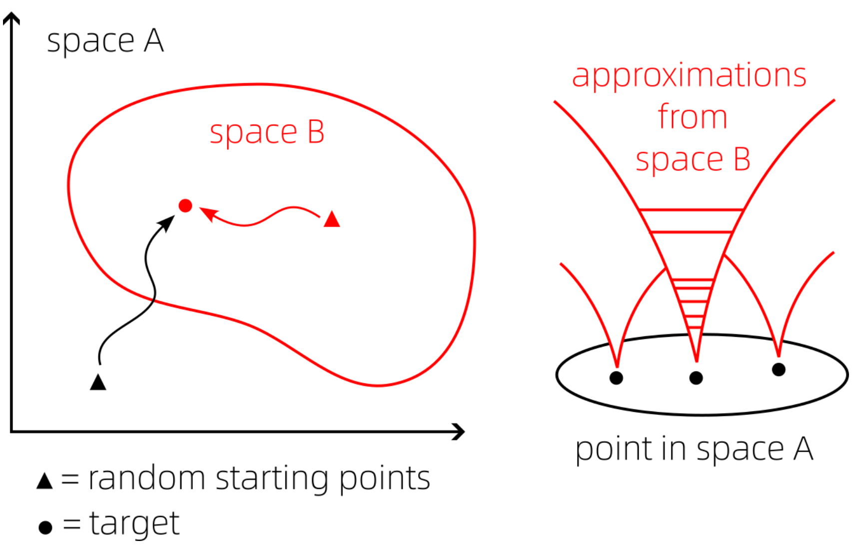
Motivating example



(1)

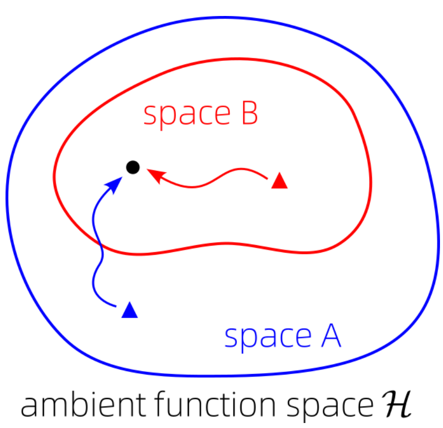
The **input space** is $X = \mathbb{R}^n = n$-dimensional hypercube. Permutation symmetry implies that only one corner of the hypercube need be considered, this is the **fundamental domain**. The **volume** of this domain over the entire hypercube shrinks exponentially as n grows, which appears to suggest that this symmetry is very significant for efficiency consideration. We want to quantify the notion of this efficiency.

We want to compare the **hypothesis spaces** A and B: A = unconstrained, fully-connected neural network with L layers and N total weights. The parameter space $A = \Theta_A = \mathbb{R}^N$. Case B = symmetric neural network with a special structure but nonetheless its parameter space is also of the form $B = \Theta_B = \mathbb{R}^M$.



(2)

It's easier to consider A and B as both embedded in an ambient function space $\mathcal{H}$ that is "fine" enough to contain all points of A and B:



Embeddings:
$\Theta_A \hookrightarrow \mathcal{H}$
$\Theta_B \hookrightarrow \mathcal{H}$ (3)

We also need a measure of the "roughness" of the landscapes A and B in the function space $\mathcal{H}$. The objective function $J = \sum d(f(x_i;\Theta), y_i) + \text{regularization}(f)$ where d is some distance function.

1

This

- Here

2

This

- Here

③

This

- Here