Title

Author

March 30, 2023

Abstract

Hello blah bldsjlfdsjk sjdklfs

1 Introduction

More text

2 Main Results

We start with a definition:

Definition 1. *Define a (k,d,n)-generative net with layer widths $k_0, k_1, \ldots, k_d$ with $k = k_0$ and $k_d = n$ and coefficients $\{l_i\}_{i\in[d]} \subseteq \mathbb{N}^d$ to be a function $G :^k \to^n$ of the form*

$$G(z) = \sigma_d \left(W^{(d)} \sigma_{d-1} \left(\cdots W^{(2)} \sigma_1 \left(W^{(1)} z \right) \right) \right),$$

where for each $i \in [d]$, $\sigma_i : \mathbb{R}^{k_i} \to \mathbb{R}^{k_i}$ is a piecewise affine function with pieces generated by no more than $l_i k_i$ hyperplanes.

Our main theorem is the following:

Theorem 1. *The self-difference of the range of a (k,d,n)-Generative Network*?? *G is contained in the range of a positive-homogeneous piecewise linear function $\tilde{G} :^{2k+1} \to^n$ with a number of linear conical pieces M such that*

$$\log M \leq 2kd \log \left(\frac{es}{k} \right),$$

where

$$s = \left(\prod_{i=1}^{d} \max(k, k_i) l_i \right)^{1/d}.$$

3 Proofs

We now prove the main result:

Proof. Notice that G is a piecewise affine function with N pieces and
??

Lemma 2. *A **??** will have at most*

$$\log N \leq kd \log \left(\frac{es}{k} \right)$$

affine pieces where

$$s := \left(\prod_{i=1}^{d} \max(k, k_i) l_i \right)^{1/d}$$

is a geometric mean of the weighted widths of the layers of the network.

It follows that $(x, y) \to G(x) - G(y)$ is also piecewise affine and has no more than $2N$ pieces.

Then from
??

Lemma 3. *Let $f : \mathbb{R}^k \to \mathbb{R}^n$ be piecewise affine with N pieces. Then there exists a function $\tilde{f} : \mathbb{R}^{k+1} \to \mathbb{R}^n$ which is piecewise linear with conical pieces such that $(\tilde{f}) \supseteq (f)$. Furthermore, $\tilde{f}((x, 1)) \equiv f(x)$ and in the case that f has polyhedral pieces, $\tilde{f}$ will have pieces that are polyhedral cones.*

there exists a function $\tilde{G} : \mathbb{R}^{2k+1} \to \mathbb{R}^n$, and one can check that it has all the properties in the statement.

□

We used two lemmas in the previous proof; **??**, **??**. These are based on [**vershyninHighDimensionalProbabilityIntroduction2018**]. We now show **??**.

Proof. From the definition of a **??**, the *ith* layer has an activation that is piecewise affine with no more than $l_i k_i$ affine pieces. This *ith* layer will have a in its domain contained in a finite number of subspaces that are of dimension of no more than $\min(k, k_i)$. Therefore consider the *ith* activation acting on a $\min(k, k_i)-$ dimensional subspace. By
?? lemma::

Lemma 4. *Let U be a $k-$ dimensional hyperplane in general position in ^{n} . Let there be P hyperplanes $\{H_i\}_{i \in [P]}$ in ^{n} . These hyperplanes induce connected components between the hyperplanes that we call "cells". The subspace U may intersect at most*

$$\sum_{i=0}^{k} \binom{P}{i} \leq \left(\frac{eP}{k}\right)^k$$

cells.

, the layer will generate no more than $\left(\frac{el_i k_i}{\min(k,k_i)}\right)^{\min(k,k_i)}$ pieces for each subspace in its domain. Notice that this expression is monotonically increasing in k_i . Therefore, by increase k_i to $\max(k, k_i)$ we get the upper bound

$$\left(\frac{el_i \max(k_i, k)}{k}\right)^k.$$

It follows that the number of affine pieces N of the full network is such that

$$N \leq \prod_{i=1}^{d} \left(\frac{el_i \max(k, k_i)}{k}\right)^k \leq \left(\frac{es}{k}\right)^{kd},$$

where

$$s := \left(\prod_{i=1}^{d} \max(k, k_i) l_i\right)^{1/d}.$$

□

And we now give the proof of lemma **??**;

Proof. Letting S be the finite partition of $dom(f)$ corresponding to the affine pieces of f , we have that $f(x) = \sum_{s \in S} W_s x + b_s$ for some $R^{n\times k}$ matrices $\{W_s\}$ and some vectors $\{b_s\} \subseteq \mathbb{R}^k$. Let $s_0 \in S$. For any vector $\tilde{x} \in \mathbb{R}^{k+1}$ with $\tilde{x} = (\tilde{x}_{[k]}, \tilde{x}_{k+1})$, let

$$\tilde{f}(\tilde{x}) :=_{\tilde{x}_k \leq 0} (W_{s_0} \tilde{x}_[k] + b_{s_0} \tilde{x}_k) +_{\tilde{x}_k > 0} \sum_{s \in S} {}_{\frac{1}{\tilde{x}_k} \tilde{x}_{[k]} \in s} (W_s \tilde{x}_[k] + b_s \tilde{x}_k).$$

One can check that for this construction all the properties described in the statement hold.

□

Created: 2023-03-27**??**