Numerical Algorithm: Convex Geometry of ReLU-Layers, Injectivity on the Ball and Local Reconstruction

SHREYAS ADIGA, KUVAM PAHUJA

May 8, 2024

Recap

- Rectified Linear Unit (ReLU) is a type of activation function widely used in artificial neural networks
- Introduces non-linearity by outputting input if positive, otherwise zero
- Authors propose adjusting biases to make activation function injective
- Convex geometry of weight matrix plays essential role in injectivity analysis
- Frame Theory used to interpret weight matrices
- Need to find upper bound of bias for injectivity

Recap

- We looked at the following theorem to estimate bias $\alpha^{\mathbb{B}}$ on $\mathbb{R}^n$

Theorem 4.4. *(PBE for $\mathbb{B}$) If $X \subseteq \mathbb{S}$ is omnidirectional, then X is $\alpha^{\mathbb{B}}$-rectifying on $\mathbb{B}$ and $\alpha_i^{\mathbb{B}}$, given in* (13) *can be computed as*

$$\alpha_i^{\mathbb{B}} = \begin{cases} 0 & \textit{if } \alpha_i^X \geq 0 \\ \alpha_i^{\mathbb{S}} & \textit{otherwise.} \end{cases}$$

If $\alpha_i^X < 0$, then $\alpha_i^{\mathbb{S}}$ given in (15) *is the minimum over $j : x_i \in F_j$ of the solutions of the convex linear programs*

$$\begin{aligned} \min \ & \left(x_i^\top D_{I_{F_j}}\right) d \\ & \textit{subject to } d \geq 0 \\ & \quad \|D_{I_{F_j}} d\|_2 \leq 1, \end{aligned}$$

where $D_{I_{F_j}}$ is the synthesis operator of $X_{I_{F_j}}$.

Polytope Bias Estimation for $\mathbb{B}_r^+$

We approach the PBE for $\mathbb{B}_r^+$ by restricting the computations of the PBE introduced in Theorem 4.4 to only those facets, that actively contribute to the estimation. In this sense, we only consider those frame elements whose associated facets have a non-trivial intersection with $\mathbb{R}_+^n$. We denote the corresponding index sets as

$$J^+ = \{j : F_j \cap \mathbb{R}_+^n \neq \emptyset\}, \qquad I^+ = \bigcup_{j \in J^+} I_{F_j}.$$

According to this, instead of omnidirectionality, we only have to require

$$\mathbb{R}_+^n \subseteq \bigcup_{j \in J^+} \mathrm{cone}(F_j), \quad \text{and}$$

$$0 \notin F_j \ \text{ for all } j \in J^+,$$

which we shall refer to as *non-negative omnidirectionality*.

Polytope Bias Estimation for $\mathbb{B}_r^+$

Theorem 4.6 (PBE for $\mathbb{B}^+$). *If $X \subseteq \mathbb{S}$ is non-negatively omnidirectional, then X is $\alpha^{\mathbb{B}^+}$-rectifying on $\mathbb{B}^+$ with*

$$\alpha_i^{\mathbb{B}^+} = \begin{cases} \alpha_i^{\mathbb{B}} & \textit{for } i \in I^+ \\ s_i & \textit{else,} \end{cases}$$

where $s_i \in \mathbb{R}$ is arbitrary.

Optimality of the PBE

- We cannot expect the proposed PBE to yield upper biases that are maximal
- In the special cases when the polytopes are regular and simplicial (every facet has exactly n vertices), e.g. Mercedes-Benz, Tetrahedron and Icosahedron frame, we expect that the estimated upper bias is indeed maximal
- PBE is also stable to perturbations as long as the combinatorial structure is preserved

Local Reconstruction via Facets

- The frame operator of XI_{Fj} is denoted by SI_{Fj}, and its canonical dual frame is given by $\tilde{X}I_{Fj} = \left(\tilde{S}I_{Fj}{}^{-1}x_i\right)_{i\in I_{Fj}}$.

Theorem 4.8. *Let $X \in \mathbb{S}$ be α-rectifying on $\mathbb{B}$ and omnidirectional. For every $x \in \mathbb{B}$ there is j such that*

$$\tilde{D}_{I_{F_j}} C_\alpha x = x,$$

where

$$\begin{aligned}\tilde{D}_{I_{F_j}} &: \mathbb{R}^m \to \mathbb{R}^n \\ &(c_i)_{i\in I} \mapsto \sum_{i\in I_{F_j}} (c_i + \alpha_i) \cdot S^{-1}_{I_{F_j}} x_i.\end{aligned}$$

In other words, $\tilde{D}_{I_{F_j}}$ is a left-inverse of C_α for all $x \in F_j^{\mathbb{B}}$.

Algorithm

Algorithm 2 Reconstruction via Facets

Get I_{F_j} via computing V_X

for $j = 1, \ldots, J$ **do**

$$S_{I_{F_j}}^{-1} \leftarrow \left((C_{I_{F_j}})^\top C_{I_{F_j}} \right)^{-1}$$

$$\overline{X} \leftarrow X_{I_{F_j}}$$

$$\tilde{D}_{I_{F_j}} \leftarrow \begin{pmatrix} | & | & & | \\ S_{I_{F_j}}^{-1} \overline{x}_1 & S_{I_{F_j}}^{-1} \overline{x}_2 & \cdots & S_{I_{F_j}}^{-1} \overline{x}_{|I_{F_j}|} \\ | & | & & | \end{pmatrix}$$

end for

$z = C_\alpha x$

$\overline{z} \leftarrow z + \alpha$

while $j = 1, \ldots, J$ **do**

if $I_{F_j} \in I_x^\alpha$ **then**

$$\tilde{D}_{I_{F_j}} \overline{z}|_{I_{F_j}} = x$$

end if

end while

Results

- $\alpha^{\mathbb{B}}$ for the Icosahedron frame: [0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]
- $\alpha^{\mathbb{B}}$ for a random frame: [-0.19850917 -0.2 -0.17862086 -0.11276229 -0.11326834 -0.15744438 -0.2 -0.19971033 -0.17862086 0. 0. 0.]
- $\alpha^{\mathbb{B}}$ for the tetrahedron frame: [-0.57735027 -0.57735027 -0.57735027 -0.57735027]

Results

Polytope	Facet obtained with relu	Error with relu	Facet obtained with z cos z	Error with z cos z
Icosahedron	[9, 3, 1]	1.942e-16	[5, 0, 10]	0.1889
Tetrahedron	[3, 2, 0]	3.43e-16	[2, 0, 1]	1.824
Random polytope in $\mathbb{R}^3$	[9, 5, 2]	1.328e-15	[9, 5, 2]	3.661

Table: Facets obtained and errors for 2 activation functions