Géométrie des réseaux ReLU f. connectés

Lemme : *fθ polynomiale par morceaux*

Pour tout $n \in \mathbb{N}^*$, $X \in \mathbb{R}^{n_0 \times n}$, les ensembles $\widetilde{U}_1^X, \dots, \widetilde{U}_{m_X}^X$ sont non-vides, ouverts et disjoints

1) Pour tout $j = 1, \dots, m_X$, $\theta \longrightarrow f_\theta(X)$ est polynomiale de degré inférieur à H sur $\widetilde{U}_j^X$

2) $\left(\bigcup_{j=1}^{m_X} \widetilde{U}_j^X\right)^c$ est fermé et de mesure de Lebesgue nulle.

---

preuve : Les 3 premières propriétés des ensembles $\widetilde{U}_j^X$ découlent de leur définition.

Montrons 1). Soit $j \in \{1, \dots, m_X\}$.

Comme chaque colonne de $f_\theta(X)$ est l'image d'un exemple (une colonne) de X, il suffit de montrer que pour tout $i = 1, \dots, n$,

$$\theta \longrightarrow f_\theta(x^i)$$

est polynomiale sur $\widetilde{U}_j^X$.

Comme sur $\widetilde{U}_j^X$ les activations $a(x^i, \theta)$ sont constantes, il suffit de montrer que, pour tout $h = 0, \dots, H$,

H_h : pour $x \in \mathbb{R}^{E_0}$: $\forall D \subset \mathbb{R}^E \times \mathbb{R}^B$, si $\theta \longmapsto a(x,\theta)$ est constante sur D, alors $\theta \longmapsto f_\theta^h(x)$ est polynomiale sur D, de degré inférieur à h

+ <u>Initialisation</u> : $h = 0$:

La fonction $\theta \longmapsto f_\theta^0(x^{(0)}) = x^{(0)}$ est bien polynomiale de degré 0 sur $\mathbb{R}^{E_0} \times \mathbb{R}^B$.

+ <u>récurrence</u> : Supposons H_{h-1} vraie pour $h \in \{1, \dots, H-1\}$

on a pour $\theta \in D$ et v sur la couche h,

$$f_\theta^h(x^{(h)})_v = \sigma\left(\left[W_h f_\theta^{h-1}(x^{(h-1)}) + b_h^{(h)}\right]_v\right)$$

$$= a(x,\theta)_{i,i} \left[W_h f_\theta^{h-1}(x^{(h-1)}) + b_h\right]_v$$

où i est l'indice correspondant à v. Donc

$$f_\theta^h(x^{(h)})_v = a(x,\theta)_{i,i}\left(\sum_{v' \in V_{h-1}} (W_h)_{v,v'} f_\theta^{h-1}(x^{(h-1)})_{v'} + [b_h]_v\right)$$

qui, comme $\theta \mapsto f_\theta^{h-1}(x^{(h-1)})$ est un polynome de degré inférieur à $h-1$ et $a(x,\theta)_{i,i}$ est constant, est bien un polynôme de degré inférieur à h.

[f_θ polyn. par morc. 2 / 3

Cela termine la preuve que H_k est vraie pour tout $k = 1 \dots H-1$

On en déduit que H_{H-1} est vraie.

On montre H_H comme précédemment, mais sans le terme $\left[a(X, \theta)\right]_{i,j}$.

2) Sketch de preuve de 2).

i) On montre que $\left(\bigcup_{j=1}^{m_r} \tilde{U}_j^r\right)^c$ correspond à l'ensemble des θ pour lesquels au moins une $\left[W_k f_\theta^{k-1}(x^i) + b_k\right]_v$ est nul.

Soit $\theta \in \left(\bigcup_{j=1}^{m_r} \tilde{U}_j^r\right)^c \Rightarrow$
- Soit j' tel que $a(X, \theta) = \Lambda_{j'}^r$
- $\exists\, j$ et $(\theta_k)_{k \in \mathbb{N}}$ tel que

Comme $\theta \notin \tilde{U}_{j'}^r$, $\forall k,\ a(X, \theta_k) = \Lambda_j^r$ et $\lim_{k \to +\infty} \theta_k = \theta$

- Soit v et j tels que $\left[a(X, \theta)\right]_{i,j} \neq \left[a(X, \theta_k)\right]_{i,j}$ où i correspond à v et. (On a choisit v le moins profond possible.) On montre que

$$\left[W_k f_\theta^{k-1}(x^i) + b_k\right]_v = 0$$

On peut montrer l'inclusion dans l'autre sens.

ii) On montre que $\{\theta \in \mathbb{R}^E \times \mathbb{R}^B \mid \exists i, v, \left[W_k f_\theta^{a-1}(x^i) + b_k\right]_v = 0\}$

- correspond aux 0 de polynômes non nuls.

fo polym. par morc. 3 / 3