Approximation of Continuous Functions with ReLU Nets Fix

Will Troiani

October 2019

Definition 1. *The* ***inverse modulus of continuity of*** ϵ *of a continuous function* $f : \mathbb{R}^2 \to \mathbb{R}$ *in the domain* $K \subseteq \mathbb{R}^2$ *is*

$$w_{f,K}^{-1}(\epsilon) = \sup\{\delta \in \mathbb{R} \mid \forall x, y \in K,\ |x - y| \leq \delta \Rightarrow |f(x) - f(y)| \leq \epsilon\}$$

Lemma 1. *Let* $f : \mathbb{R}^2 \to \mathbb{R}$ *be continuous,* $\epsilon > 0$*,* $r > r' > 0$*. Let* $S_r := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = r\}$ *be the circle of radius* r*. Let* $X, Y \in S_r$ *be such that* $|X - Y| \leq r$ *and* $X', Y' \in S_{r'}$ *be such that* X', X, Y, Y' *are collinear, and let* $Z \in S_{r'}$ *be the point on the arc connecting* X' *to* Y' *and the straight line which passes through the origin and the midpoint of the straight line connecting* X *and* Y*. See Figure 1. Let* L *denote the minor sector of* S_r *induced by the radii* $0X$ *and* $0Y$ *and let* L' *denote the minor segment of* $S_{r'}$ *induced by the chord which connects* X' *to* Y'*. If*

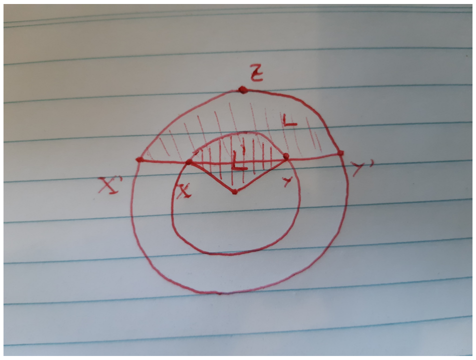

Figure 1: The configuration described in Lemma 1

$$\mathrm{Diam}(L) \leq w_{f,B_{r'}(0)}^{-1}(\epsilon)$$

then there exists an affine function $l : \mathbb{R}^2 \to \mathbb{R}$ *such that*

- $l(x) \leq \epsilon$ *for all* $x \in L'$*,*

- *$|f(x)-f(Z)| \leq l(x)+\epsilon$ for all $x \in L$.*

Proof. Let l be the affine function such that $l(Z)=0$ and $l(X')=l(Y')=\epsilon$. Let $x \in L$. Denote by $Z' \in \mathbb{R}^2$ the point which is collinear to X', Y', X, Y and is also collinear to Z, x. Then as l is affine and $|x-Z| \leq |Z'-Z|$, it follows that $l(x) \leq l(Z')=\epsilon$.

Say $x \in L'$. Then write $|x-Z| = kw^{-1}_{f,B_{r'}(0)}(\epsilon)+\rho$ where $k \in \mathbb{N}$ and $0 \leq \rho < w^{-1}_{f,B_{r'}(0)}(\epsilon)$. Then

$$\begin{aligned} |x-Z| &= kw^{-1}_{f,B_{r'}(0)}(\epsilon)+\rho \\ \Rightarrow |x-Z| &\geq kw^{-1}_{f,B_{r'}(0)}(\epsilon) \end{aligned}$$

and so,

$$\frac{\epsilon|Z-x|}{w^{-1}_{f,B_{r'}(0)}(\epsilon)} \geq k\epsilon \tag{1}$$

Also, consider a sequence of points $Z=x_1, x_2, ..., x_{k+1}=x$ such that $|x_{i+1}-x_i| = w^{-1}_{f,B_{r'}(0)}(\epsilon)$ for $i=1,...,k-1$ and $|x_{k+1}-x_k| < w^{-1}_{f,B_{r'}(0)}(\epsilon)$. See Figure Figure needed. Then

$$\begin{aligned} |f(x)-f(Z)| &= |f(x_1)-f(x_2)+f(x_2)-f(x_3)+...+f(x_2)-f(x_1)| \\ &\leq |f(x_1)-f(x_2)|+...+|f(x_2)-f(x_1)| \\ &< k\epsilon+\epsilon \end{aligned}$$

and so,

$$|f(x)-f(Z)| < k\epsilon+\epsilon \tag{2}$$

Equations 1 and 2 together imply

$$|f(x)-f(Z)| < \frac{\epsilon|x-Z|}{w^{-1}_{f,B_{r'}(0)}(\epsilon)}+\epsilon$$

Thus it remains to show:

$$\frac{\epsilon|x-Z|}{w^{-1}_{f,B_{r'}(0)}(\epsilon)} \leq l(x)$$

This comes down to the fact that the slope of l in the direction of the vector $x-Z$ can be bounded below by $\frac{\epsilon}{w^{-1}_{f,B_{r'}(0)}(\epsilon)}$. More precisely, consider the following parametrisation of the straight line which intersects Z and x:

$$\begin{aligned} c &: \mathbb{R} \to \mathbb{R}^2 \\ s &\mapsto Z+\frac{s}{|x-Z|}(x-Z) \end{aligned}$$

Then the function $lc : \mathbb{R} \to \mathbb{R}$ is affine, and in fact is linear as

$$lc(0) = l(Z) = 0$$

So for all $y \in \mathbb{R}$, $lc(y) = my$ for some gradient m. Let $W \in \mathbb{R}^2$ be the point which intersects the line parametrised by c and the line which passes through X and Y. Then since $|W - Z| \leq w^{-1}_{f,B_{r'}(0)}(\epsilon)$ and $l(W) = \epsilon$, it follows that:

$$m = \frac{l(W) - l(Z)}{|W - Z|} = \frac{\epsilon}{|W - Z|} \geq \frac{\epsilon}{w^{-1}_{f,B_{r'}(0)}(\epsilon)}$$

and so:

$$\frac{\epsilon|x - Z|}{w^{-1}_{f,B_{r'}(0)}(\epsilon)} \leq m|x - Z| = lc(|x - Z|) = l(x)$$

which completes the proof. $\square$

Corollary 1. *In the setting of Lemma 1, but with every instance of ϵ replaced by $\frac{\epsilon}{2}$, if there exists a max-min string $g : \mathbb{R}^2 \to \mathbb{R}$ which approximates f on L' then there exists a max-min string $\hat{g} : \mathbb{R}^2 \to \mathbb{R}$ which approximates f on $L \cup L'$.*

Proof. Let $l : \mathbb{R}^2 \to \mathbb{R}$ be an affine function such that

- $l(x) \leq \frac{\epsilon}{2}$, for all $x \in L'$,
- $|f(x) - f(Z)| \leq l(x) + \frac{\epsilon}{2}$, for all $x \in L$

the existence of which is guaranteed by Lemma 4. Then define the max-min string

$$\hat{g} = \max\{\min\{g, f(Z) + l\}, f(Z) - l\}$$

On L, we have $f \leq g + \epsilon$ and

$$f \leq f(Z) + l + \epsilon$$

so

$$\begin{aligned} f &\leq \min\{g + \epsilon, f(Z) + l + \epsilon\} \\ &= \min\{g, f(Z) + l\} + \epsilon \\ &\leq \max\big\{\min\{g, f(Z) + l\}, f(Z) - l\big\} + \epsilon \\ &= \hat{g} + \epsilon \end{aligned}$$

Similarly, since $g - \epsilon \leq f$ and

$$f(Z) - l - \epsilon \leq f$$

it follows that

$$\begin{aligned} \hat{g} - \epsilon &= \max\big\{\min\{g, f(Z) + l - \epsilon\}, f(Z) - l\big\} - \epsilon \\ &= \max\{\min\{g - \epsilon, f(Z) + l - \epsilon\}, f(Z) - l - \epsilon\} \\ &\leq \max\{g - \epsilon, f(Z) - l - \epsilon\} \\ &\leq f \end{aligned}$$

On L', by construction:

$$f(Z) - \frac{\epsilon}{2} \leq f(Z) - l \leq \hat{g} \leq f(Z) + l \leq f(Z) + \frac{\epsilon}{2}$$

and for all $x \in L'$

$$|f(x) - f(Z)| \leq \frac{\epsilon}{2}$$

by definition of $w^{-1}_{f,B_{r'}(0)}(\frac{\epsilon}{2})$. Thus:

$$\begin{aligned}
|\hat{g}(x) - f(x)| &= |\hat{g}(x) - f(Z) + f(Z) - f(x)| \\
&\leq |\hat{g}(x) - f(Z)| + |f(Z) - f(x)| \\
&\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} \\
&= \epsilon
\end{aligned}$$

□

Next we turn to two geometric Lemmas:

Lemma 2. *Consider two concentric circles, S_R and $S_{R'}$ with $R' > R$. Consider a chord intersecting points X and Y on S_R and extend this chord so that it intercepts $S_{R'}$, at X' and Y' say (see figure 2). Then*

$$|X'Y'|^2 = 4(R'^2 - R^2) + |XY|^2$$

Proof. Let P denote the midpoint of the line XY. Then $|P| = \sqrt{R^2 - \frac{1}{4}|XY|^2}$ (again, see figure 2). Thus

$$\left(\sqrt{R^2 - \frac{1}{4}|XY|^2}\right)^2 + \frac{1}{4}|X'Y'|^2 = R'^2$$

from which, the result follows. □

Lemma 3. *Let X, Y be points on S_R such that $0 < |X - Y| \leq R$. Let τ_X and τ_Y be straight lines tangent to S_R respectively at X and Y. Let T denote the intersection of τ_X and τ_Y. Then $\mathrm{Diam}(\Delta XTY) = |X - Y|$*

Proof. If $|X - Y| = R$ then the triangle XOY is equilateral. It follows from this that angle $\angle TXY$ and $\angle TYX$ are both equal to $30°$. As $|X - Y|$ decreases, both angles $\angle TXY$ and $\angle TYX$ decrease. So if $|X - Y| \leq R$ the angle $\angle XTY$ is obtuse, which implies that $\mathrm{Diam}(\Delta XTY) = |X - Y|$. Figure needed. □

Lemma 4. *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be continuous, $\epsilon > 0$, $R > 0$. Assume $w^{-1}_{f,B_R(0)}(\frac{\epsilon}{2})$ and let $R' > w^{-1}_{f,B_R(0)}(\frac{\epsilon}{2})$. Assume further that there exists a max-min string $g : \mathbb{R}^2 \to \mathbb{R}$ which approximates f on $B_{R'}(0)$. Let R'' be*

$$R'' := \sqrt{R'^2 + \frac{w^{-1}_{f,B_R(0)}(\frac{\epsilon}{2})}{\sqrt{2}R'}}$$

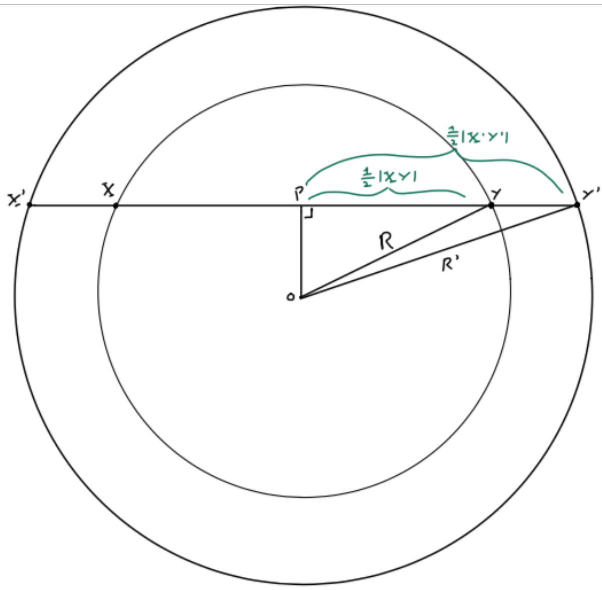

Figure 2: The geometry of Lemma 2

If $R'' < R$ then there exists a max-min string $\hat{g} : \mathbb{R}^2 \to \mathbb{R}$ which approximates f on $B_{R''}(0)$, and if $R'' \geq R$ then there exists a max-min string $\hat{g} : \mathbb{R}^2 \to \mathbb{R}$ which approximates f on $B_R(0)$.

Proof. To ease notation let $a_{R'} := \frac{w^{-1}_{f,B_R(0)}(\frac{\epsilon}{2})}{\sqrt{2}R'}$. Let $X, Y \in \mathbb{R}^2$ lie on the circle of radius R' which satisfy

$$|XY| = w^{-1}_{f,B_R(0)}(\frac{\epsilon}{2})(1 - a_{R'})^{\frac{1}{2}}$$

such a pair (X, Y) exists as

$$|XY|^2 = w^{-1}_{f,B_R(0)}(\frac{\epsilon}{2})^2(1 - a_{R'}^2) < w^{-1}_{f,B_R(0)}(\frac{\epsilon}{2}) \leq R'$$

We consider first the case when $R'' < R$. Let $X', Y' \in \mathbb{R}^2$ lie on the circle of radius R'' which intersect the line segment connecting X and Y. Let L denote the minor segment on the circle of radius R' induced by the chord which connects X' to Y' and let L' denote the minor sector on the circle of radius R' induced by the radii $0X$ and $0Y$. Let Z denote the intersection of the circle of radius R'' and the line which connects 0 and the midpoint of the line XY. See Figure 3. By Lemma 2 and a direct calculation,

$$|X'Y'| = \sqrt{4(R''^2 - R'^2) + |XY|^2} = w^{-1}_{f,B_R(0)}(\frac{\epsilon}{2})$$

Moreover, $w^{-1}_{f,B_R(0)}(\frac{\epsilon}{2}) < R' < R''$, so by Lemma 3:

$$\text{Diam}(L) \leq w^{-1}_{f,B_R(0)}(\frac{\epsilon}{2})$$

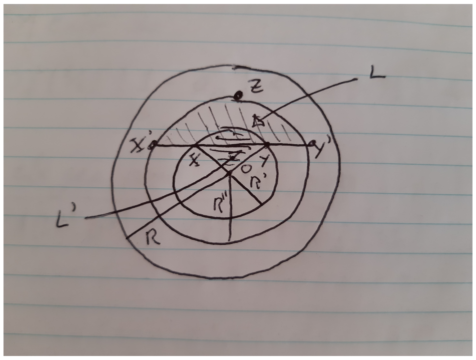

Figure 3: Configuration of the balls in Lemma 4

Thus the hypotheses of Corollary 1 are satisfied.

The case when $R'' \geq R$ is almost identical but the pair (X', Y') are taken to intersect the ball of radius R instead of the ball of radius R''. $\square$

This brings us to the main Theorem:

Theorem 1. *Let $\epsilon \in \mathbb{R}_{>0}$ and $f : \mathbb{R}^2 \to \mathbb{R}$ be continuous, and $K \subseteq \mathbb{R}^2$ compact. Then there exists a max-min string $g : \mathbb{R}^2 \to \mathbb{R}$ such that*

$$\forall x \in K, \ |g(x) - f(x)| \leq \epsilon$$

Proof. Since compact subsets of $\mathbb{R}^2$ are bounded, it suffices to consider the case when $K = B_R := \{x \in \mathbb{R}^2 \mid |x| \leq R\}$, for some $R \in \mathbb{R}$. If $R \leq w^{-1}_{f,B_R(0)}\left(\frac{\epsilon}{2}\right)$ then the max-min string $g : \mathbb{R}^2 \to \mathbb{R}$ which is the constant function $g \equiv f(0)$ suffices. Suppose $w^{-1}_{f,B_R(0)}\left(\frac{\epsilon}{2}\right) < R$ and $f : B_R(0) \to \mathbb{R}$ are given. Consider the following set

$$\mathscr{X} := \{R' \in \mathbb{R} \mid \text{there exists a max-min string } g \text{ which approximates } f \text{ on } B_R(0)\}$$

We claim that $R \in \mathscr{X}$, recall the notation from Lemma 4 that $a_R := \frac{w_f^{-1}\left(\frac{\epsilon}{2}\right)^2}{\sqrt{2}R}$. Suppose for a contradiction that $R \notin \mathscr{X}$, then $\mathscr{X}$ is bounded above. Let $s^2 = \sup \mathscr{X}^2$ and $R' \in \mathscr{X}$ to be such that

$$s^2 - a_s < R'^2$$

Notice that $s \notin \mathscr{X}$ as if $s \in \mathscr{X}$ then Lemma 4 can be used to find a strictly greater value in $\mathscr{X}$. As $R' < s$, it follows that $a_s < a_{R'}$ and

$$s^2 < R'^2 + a_s < R'^2 + a_{R'}$$

thus

$$s < \sqrt{R'^2 + a_{R'}}$$

but by Lemma 4, $\sqrt{R'^2 + a_{R'}} \in \mathscr{X}$, contradicting that s is a supremum. $\square$