1. The statements and ideas behind Lemma 3.1 and Lemma 3.2 :

此論文在探討如何藉單一隱藏層的 tanh 神經網路逼近各種函數，此 2 Lemma 主要是針對單變多項式逼近的結果，lemmas 提供一種建構性的方法建立一個單一隱藏層的 tanh 網路，精確近似給定範圍內的各階次冪函數，同時量化所需網路規模與參數大小。建構完多項式的逼近方法後，後續就能延伸探討一般連續函數的逼近 (By the Weierstrass approximation thm, Remark 3.3)

Lemma 3.1. Let $k \in \mathbb{N}_0$ and $s \in 2\mathbb{N}-1$. Then it holds that for all $\varepsilon > 0$, there exists a shallow tanh neural network $\Psi_{s,\varepsilon} : [-M, M] \to \mathbb{R}^{\frac{s+1}{2}}$ of width $\frac{s+1}{2}$ s.t.

$$\max_{\substack{p \le s \\ p\text{-odd}}} \| f_p - (\Psi_{s,\varepsilon})_{\frac{p+1}{2}} \|_{W^{k,\infty}} \le \varepsilon,$$

Moreover, the weights of $\Psi_{s,\varepsilon}$ scale as $O\left(\varepsilon^{-s/2} \left(2(s+2)\sqrt{2M}\right)^{s(s+3)}\right)$ for small ε and large s.

Lemma 3.1 說明：處理奇數次方的單項式函數的逼近。

Given any $k \in \mathbb{N} \cup \{0\}$ (導數階層) and 奇數次方指數上界 s，

for any $\varepsilon > 0$, there exists a 單一隱藏層 tanh 神經網路 $\Psi_{s,\varepsilon}$，具備 $\frac{s+1}{2}$ 個隱藏神經元 s.t. for all $p \le s$ ($p \in 2\mathbb{N}-1$)，該網路的第 $\frac{p+1}{2}$ 個輸出分量可以近似函數 $f_p(x) = x^p$ on $[-M, M]$，其逼近誤差在 $W^{k,\infty}$ 範數下不超過 ε。

簡單來說，Lemma 3.1 保證存在一個只有單一隱藏層的 tanh 神經網路，就能在一個區間上把奇數次單項多項式逼近到任意誤差 $\varepsilon > 0$.

Lemma 3.1 所需背景知識：

1. tanh(x) 在 $x=0$ 附近可以展開為泰勒級數：

$$\tanh(x) \approx x - \frac{x^3}{3} + \frac{2x^5}{15} - \cdots$$

2、線性組合

簡單例子：想逼近 $f(x)=x^3$.

想法：1. $\tanh(x)$ 有 x 和 $\frac{-x^3}{3}$

2. 透過線性組合 $\tan(\alpha x)$, $\tan(\beta x)$ 抵消一次項。

Lemma 3.2. Let $k \in \mathbb{N}_0$, $s \in 2\mathbb{N}-1$ and $M>0$. For every $\varepsilon>0$, there exists a shallow tanh neural network $\psi_{s,\varepsilon} : [-M, M] \to \mathbb{R}^s$ of width $\frac{3(s+1)}{2}$ s.t.

$$\max_{p\le s} \| f_p - (\psi_{s,\varepsilon})_p \|_{W^{k,\infty}} < \varepsilon .$$

Furthermore, the weights scale as $O(\varepsilon^{-s/2}(\sqrt{M}(s+2))^{3s(s+3)/2})$ for small ε and large s.

Lemma 3.2 說明：在 Lemma 3.1 的基礎上，進一步處理偶數次方單項多項式的逼近。

Let $s \in 2\mathbb{N}-1$，則次方上界為 s 的所有單項多項式（奇 or 偶），for any $\varepsilon>0$, there exists width $\frac{3(s+1)}{2}$ 的單一隱藏層 tanh 神經網路 $\psi_{s,\varepsilon}$，滿足 for all $p \le s$:

$$\| f_p(x) - (\psi_{s,\varepsilon}(x))_p \|_{W^{k,\infty}} \le \varepsilon .$$

簡單來說：Lemma 3.2 保證在給定一個偶數次方單項多項式時，

可以透過「奇數次方的近似」進一步建構偶數次方的近似。

讓我們只要一個稍寬的單一隱藏層 tanh 神經網路，就能逼近

$1, x, x^2, \dots, x^s$, where $s \in 2\mathbb{N}-1$。

Lemma 3.2 所需背景知識：

1. 代數恆等式：用於將偶數次方多項式以奇數次方多項式組合出來。
2. 數學歸納法：假設能逼近低次方，便能一步步建構高次方。

簡單例子：近似 $f(x)=x^2$：

1. by Lemma 3.1，可近似 x^3
2. by 代數恆等式，可用 x^3 組合出 x^2 的近似。

2、希望未來在作業繳交要求可以提早說明，如 Assignment 2 的 program part，在作業 deadline 後教授才說明作業中要包含哪些內容。希望未來如果有比較個以往的作業類型，能一併說明要包含哪些內容。