關於 Lemma 3.1 與 Lemma 3.2 的說明

背景

神經網路之所以能逼近各種複雜函數，是因為它能把簡單的「積木函數」組合起來。這篇論文的目標，是要證明 **使用 tanh 激活函數的淺層網路（只有一個隱藏層）也能非常有效率地逼近光滑函數**。

第一步就是檢查網路能不能逼近 **多項式**。

- 為什麼要從多項式開始？因為多項式是很多近似理論的基礎。Taylor 展開就是把函數用多項式來近似。
- 如果能證明 tanh 網路可以準確表示多項式，那麼進一步就能近似更廣泛的平滑函數。

Lemma 3.1 和 Lemma 3.2 就是在處理這個問題：如何用淺層 tanh 網路近似 **奇數次方**和**偶數次方**。

Lemma 3.1 —— 近似奇數次方

對於任意奇數次方，存在一個淺層 tanh 網路，它的寬度大約是次方數的一半，可以把這些多項式在一個區間內近似到任意精度。當誤差要求更小時，權重會變大，但仍然在可控範圍。

主要想法：

- 作者用 **有限差分公式**（數值分析的工具）從 tanh 函數中「抽出」多項式的形狀。
- 因為 tanh 是 **奇函數**（$\tanh(-x)=-\tanh(x)$），所以它天然適合近似奇次方。
- 把不同位置、不同縮放的 tanh 函數組合起來，就能拼出 x^p 這樣的曲線。

👉 **重點結論**：奇數次方單項式可以直接用小型淺層 tanh 網路構造出來。

Lemma 3.2 —— 近似偶數次方

除了奇數次方以外，也可以構造一個淺層 tanh 網路，把偶數次方（x2,x4,...x^2, x^4,...x2,x4,...）逼近到任意精度。網路會稍微大一些，但仍然只有一個隱藏層。

主要想法：

- 偶次方比較麻煩，因為 tanh 本身是奇函數，直接做不到。
- 作者用了代數技巧：把偶次方寫成 **奇次方的組合**。
  例如：

  $$x^2 = \frac{1}{2\alpha(3)}\left((x+\alpha)^3 - (x-\alpha)^3 - 6\alpha x\right),$$

  也就是把偶數次方轉換成奇數次方的線性組合。

- 因為 Lemma 3.1 已經處理了奇數次方，所以偶數次方也能藉此間接完成。
- 他們用遞迴的方法，逐步得到 2,4,6,⋯次 的近似。

👉 **重點結論**：利用奇數次方的近似結果，可以構造出偶數次方。需要更多神經元（大約 3 倍），但仍然是淺層網路。

為什麼這兩個 Lemma 重要？

- Lemma 3.1 + Lemma 3.2 = 能近似所有多項式（奇次、偶次）。
- 因為多項式是泰勒展開的基礎 → 代表 tanh 網路能近似所有光滑函數。
- 這為後續的結果鋪路：只要兩層的 tanh 網路，就能達到跟 ReLU 深層網路一樣甚至更好的近似效果。

直觀小例子

想像我們想用 tanh 來「畫」出 y = x^2。

- 單獨一個 tanh 函數看起來像 S 曲線，完全不像 x^2。
- 但如果你把幾個平移、縮放過的 tanh 疊加起來，就能拼湊出類似 x^2 的形狀。
- Lemma 3.1 告訴你怎麼做出.x^3, x^5,⋯。
- Lemma 3.2 告訴你怎麼用這些結果再拼湊出.x^2, x^4,..。

換句話說，tanh 網路就像積木，可以拼出任意次方的曲線。

☑️ 總結（給修過微積分的學生）：

- **Lemma 3.1**：淺層 tanh 網路能近似奇數次多項式。
- **Lemma 3.2**：用技巧把奇數次組合起來，也能近似偶數次多項式。
- **合併結果**：任何多項式都能用一層隱藏層的 tanh 網路逼近 → 奠定了後面近似光滑函數的基礎。