

# 1141ML Week 3 - Hand Writting 02

313652018 王宣璋

September 2025

## 1 Unsolving problem: Weight Magnitude in Approximation

In the reference paper, in both Lemma 3.1 and Lemma 3.2, the construction shows that shallow tanh networks can approximate monomials with arbitrary accuracy. However, the required weights often grow rapidly as the error tolerance  $\varepsilon \rightarrow 0$  or as the polynomial degree  $s$  increases.

What are the implications of such large weights in practice? For instance, could excessively large weights lead to numerical instability, gradient explosion or vanishing during training, or poor generalization?

How might one balance theoretical approximation guarantees with practical considerations in optimization?