# Contents

# 1

# Polynomials, Finite Element and Neural Network Functions

In this chapter, we discuss approximation properties of spaces of polynomials and finite elements consisting of piecewise polynomials.

## 1.1 Approximation by polynomials and Weierstrass Theorem

Let $\alpha = (\alpha_1, \cdots, \alpha_d)$ with $\alpha_i$ being non-negative integers, we note $|\alpha| = \sum_{i=1}^{d} \alpha_i$ and

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_d^{\alpha_d}.$$

We use $\mathbb{P}_m(\mathbb{R}^d)$ to define the polynomials of $d$-variables of degree less than $m$ which consists functions of the form

$$\sum_{|\alpha|\le m} a_\alpha x^\alpha = \sum_{\alpha_1+\alpha_2+\ldots+\alpha_d \le m} a_{\alpha_1,\ldots,\alpha_d} x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_d^{\alpha_d}$$

### 1.1.1 Weierstrass Theorem and Proof

Important property: polynomials can approximate any reasonable function!

- dense in $C(\Omega)$ [Weierstrass theorem]
- dense in all Sobolev spaces: $L^2(\Omega), W^{m,p}(\Omega), \ldots$

**Theorem 1.** *Let $\Omega \subset R^n$ be a closed and bounded set. Given any continuous function $f(x)$ on $\Omega$, there exists a sequence of polynomials $\{P_n(x)\}$ such that*

$$\lim_{n\to\infty} \max_{x\in\Omega} |f(x) - P_n(x)| = 0 \tag{1.1}$$

*Proof.* Let us first give the proof for $d = 1$ and $\Omega = [0,1]$. Given $f : [0,1] \to R$ be a continuous function.

Let

$$\tilde{f}(x) = f(x) - l(x) \tag{1.2}$$

where $l(x) = f(0) + x(f(1) - f(0))$. Then $\tilde{f}(0) = \tilde{f}(1) = 0$. Noting that $l(x)$ is a linear function, hence without lose of generality, we can only consider the case $f : [0, 1] \to R$ with $f(0) = f(1) = 0$.

Since $f$ is continuous on the closed interval $[0, 1]$, then $f$ is uniformly continuous on $[0, 1]$.

First we extend $f$ to be zero outside of $[0, 1]$ and obtain $f : R \to R$, then it is obviously that $f$ is still uniformly continuous.

Next for $0 \le x \le 1$, we construct

$$p_n(x) = \int_{-1}^{1} f(x+t)Q_n(t)dt = \int_{-x}^{1-x} f(x+t)Q_n(t)dt = \int_0^1 f(t)Q_n(t-x)dt \tag{1.3}$$

where $Q_n(x) = c_n(1 - x^2)^n$ and

$$\int_{-1}^{1} Q_n(x)dx = 1. \tag{1.4}$$

Thus $\{p_n(x)\}$ is a sequence of polynomials.

Since

$$\int_{-1}^{1}(1 - x^2)^n dx = 2\int_0^1 (1 - x^2)^n dx = 2\int_0^1 (1 - x)^n(1 + x)^n dx \tag{1.5}$$

$$\ge 2\int_0^1 (1 - x)^n dx = \frac{2}{n+1} > \frac{1}{n}. \tag{1.6}$$

Combing with $\int_{-1}^{1} Q_n(x)dx = 1$, we obtain $c_n < n$ implying that for any $\delta > 0$

$$0 \le Q_n(x) \le n(1 - \delta^2)^n \quad (\delta \le |x| \le 1), \tag{1.7}$$

so that $Q_n \to 0$ uniformly in $\delta \le |x| \le 1$ as $n \to \infty$.

Given any $\epsilon > 0$, since $f$ in uniformly continuous, there exists $\delta > 0$ such that for any $|y - x| < \delta$, we have

$$|f(y) - f(x)| < \frac{\epsilon}{2}. \tag{1.8}$$

Finally, let $M = \max |f(x)|$, using (1.8), (1.4) and (1.7), we have

$$\big|p_n(x) - f(x)\big| = \Big|\int_{-1}^{1}(f(x+t) - f(t))Q_n(t)dt\Big| \le \int_{-1}^{1}\big|f(x+t) - f(t)\big|Q_n(t)dt \tag{1.9}$$

$$\le 2M\int_{-1}^{-\delta} Q_n(t)dt + \frac{\epsilon}{2}\int_{-\delta}^{\delta} Q_n(t)dt + 2M\int_{\delta}^{1} Q_n(t)dt \tag{1.10}$$

$$\le 4Mn(1 - \delta^2)^n + \frac{\epsilon}{2} < \epsilon \tag{1.11}$$

for all large enough $n$, which proves the theorem.

The above proof generalize the high dimensional case easily. We consider the case that

$$\Omega = [0, 1]^d.$$

By extension and using cut off function, W.L.O.G. that we assume that $f = 0$ on the boundary of $\Omega$ and we then extending this function to be zero outside of $\Omega$.

Let us consider the special polynomial functions

$$Q_n(x) = c_n \prod_{k=1}^{d} (1 - x_k^2) \tag{1.12}$$

Similar proof can then be applied. □

### 1.1.2 Some issues with polynomial approximations

#### Curse of dimensionality

Number of coefficients for polynomials of degrees $n$ in $\mathbb{R}^d$ is

$$N = \binom{d+n}{n} = \frac{(n+d)!}{d!n!}.$$

For example $n = 100$:

| $d =$ | 2 | 4 | 8 |
|---|---|---|---|
| $N =$ | $5 \times 10^3$ | $4.6 \times 10^6$ | $3.5 \times 10^{11}$ |

#### Runge's phenomenon

Consider the case where one desires to interpolate through $n+1$ equispaced points of a function $f(x)$ using the n-degree polynomial $P_n(x)$ that passes through those points. Naturally, one might expect from Weierstrass' theorem that using more points would lead to a more accurate reconstruction of $f(x)$. However, this particular set of polynomial functions $P_n(x)$ is not guaranteed to have the property of uniform convergence; the theorem only states that a set of polynomial functions exists, without providing a general method of finding one.

The $P_n(x)$ produced in this manner may in fact diverge away from $f(x)$ as n increases; this typically occurs in an oscillating pattern that magnifies near the ends of the interpolation points. This phenomenon is attributed to Runge.

Problem: Consider the Runge function

$$f(x) = \frac{1}{1 + 25x^2}$$

(a scaled version of the Witch of Agnesi). Runge found that if this function is interpolated at equidistant points $x_i$ between $-1$ and 1 such that:

$$x_i = \frac{2i}{n} - 1, \quad i \in \{0, 1, \ldots, n\}$$

with a polynomial $P_n(x)$ of degree $\leq n$, the resulting interpolation oscillates toward the ends of the interval, i.e. close to $-1$ and 1. It can even be proven that the interpolation error increases (without bound) when the degree of the polynomial is increased:

$$\lim_{n\to\infty} \left( \max_{-1\leq x\leq 1} |f(x) - P_n(x)| \right) = +\infty.$$

This shows that high-degree polynomial interpolation at equidistant points can be troublesome.

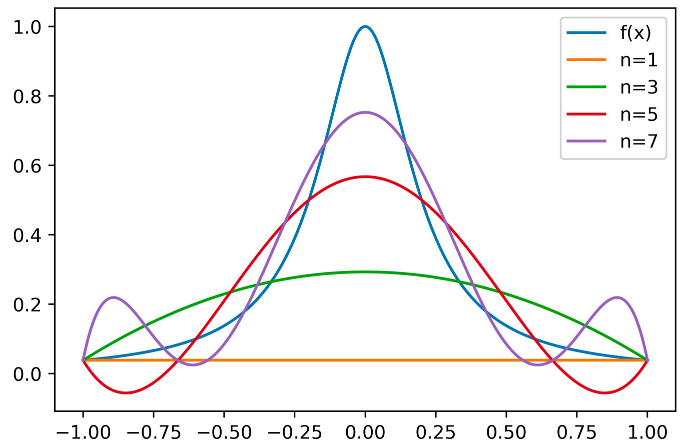


**Fig. 1.1.** Runge's phenomenon: Runge function $f(x) = \frac{1}{1+25x^2}$ and its polynomial interpolation $P_n(x)$.