CSE 840: Computational Foundations of Artificial Intelligence March 10, 2025

Lecture 12: Power Series, Taylor Series

Instructor: Vishnu Boddeti *Scribe: Minh Nguyen*

1 Power Series

Definition 1 *A series of the form*

$$p(x) = \sum_{n=0}^{\infty} a_n x^n$$

is called a ***power series****.*

An infinite sum is called a **series**, and it consists of terms that are **powers** of x.

Theorem 2 (Radius of Convergence) *For every power series*

$$p(x) = \sum_{n=0}^{\infty} a_n x^n$$

there exists a constant r, $0 \leq r \leq \infty$, called the radius of convergence such that:

- *The series converges* ***absolutely*** *for all x with $|x| < r$ (means that $\sum_{n=0}^{\infty} a_n |x|^n$ converges, the sequence of partial sums $P_N(x) := \sum_{n=0}^{N} a_n |x|^n$ converges "in the usual sense" as $N \to \infty$).*
- ***Unclear*** *what happens when $|x| = r$.*
- *If $|x| < r$, the series even converges* ***uniformly****.*

The radius of convergence only depends on the sequence $(a_n)_n$ and can be computed by various formulae (if it exists):

- $\gamma = \frac{1}{L}$ *where* $L = \limsup_{n\to\infty} (|a_n|)^{1/n}$
- $\gamma = \lim_{n\to\infty} \left|\frac{a_n}{a_{n+1}}\right|$

Example 3 *Consider the power series:*

$$p(x) = \sum_{n=0}^{\infty} n^c x^n \quad \textit{for some constant } c.$$

The radius of convergence is given by:

$$\gamma = \lim_{n\to\infty} \left|\frac{a_n}{a_{n+1}}\right| = \lim_{n\to\infty} \frac{n^c}{(n+1)^c} = \lim_{n\to\infty} \left(\frac{n}{n+1}\right)^c = 1.$$

Case 1 *For c = −1:*

$$\sum_{n=1}^{\infty} \frac{1}{n} x^n,$$

the radius of convergence is $r = 1$.

- *For* $x = +1$, *the series diverges:*

$$\sum_{n=1}^{\infty} \frac{1}{n} \cdot x^n = \sum_{n=1}^{\infty} \frac{1}{n} \cdot 1^n = \sum_{n=1}^{\infty} \frac{1}{n} \to \infty.$$

- *For* $x = -1$, *the series converges.*
- *For* $x > 1$, *the series diverges.*

Case 2 *For c = 0:*

$$\sum_{n=0}^{\infty} n^c x^n = \sum_{n=0}^{\infty} x^n,$$

the series diverges for $|x| = r$ *(both* $x = 1$ *and* $x = -1$*).*

Example 4 *Exponential Series:*

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

has a radius of convergence $r = \infty$, *since*

$$\lim_{n\to\infty} \left| \frac{a_n}{a_{n+1}} \right| = \lim_{n\to\infty} \frac{(n+1)!}{n!} = n + 1 \to \infty.$$

Example 5 *The series:*

$$\sum_{n=0}^{\infty} n!\, x^n$$

has a radius of convergence $r = 0$, *since*

$$\lim_{n\to\infty} \left| \frac{a_n}{a_{n+1}} \right| = \lim_{n\to\infty} \frac{n!}{(n+1)!} = \lim_{n\to\infty} \frac{1}{n+1} \to 0.$$

From Power Series to Taylor Series

Observation 6 *Given a power series:*

$$f(x_0 + h) = \sum_{n=0}^{\infty} a_n h^n,$$

its derivative is:

$$f'(x_0 + h) = (a_0 + a_1 h + a_2 h^2 + \cdots)' = a_1 + 2a_2 h + 3a_3 h^2 + \cdots = \sum_{n=1}^{\infty} n \cdot a_n h^{n-1}.$$

For higher-order derivatives of the power series:

$$f^{(k)}(x_0+h)=\sum_{n=k}^{\infty}a_n\big(n\cdot(n-1)\cdot(n-2)\cdots(n-k+1)\big)h^{n-k}.$$

In particular, we have:

$$f^{(k)}(x_0)=a_k\cdot k! \quad\Rightarrow\quad \boxed{a_k=\frac{f^{(k)}(x_0)}{k!}}.$$

Theorem 7 *Let $f(x_0+h)=\sum_{n=0}^{\infty}a_nh^n$ with $r>0$. Then for h with $|h|<r$, we have:*

$$f(x_0+h)=\sum_{n=0}^{\infty}\frac{f^{(n)}(x_0)}{n!}h^n.$$

Remark 8 *Intuition: Start with a power series that converges. Then we have a nice formula that expresses the coefficients in terms of the derivatives of the function.*

Question

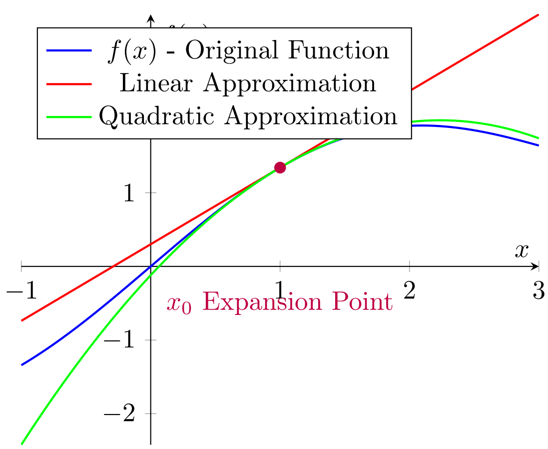
Does the theorem hold the other way around? That is, given any function (possibly with nice assumptions), can we simply build the series:

$$\sum_{n=0}^{\infty}\frac{f^{(n)}(x_0)}{n!}h^n$$

and "hope" that it converges to the function $f(x)$?

2 Taylor Series

Intuition of Taylor's Theorem



- Linear Approximation: $f(x_0+h) = f(x_0) + f'(x_0)\cdot h + r(h)\cdot h$, where $x = x_0 + h$ and $r(h) \xrightarrow{h\to 0} 0$.
- Quadratic Approximation: $f(x_0+h) = f(x_0) + f'(x_0)\cdot h + \frac{1}{2}f''(x_0)\cdot h^2 + r(h)\cdot h^2$, where $r(h) \xrightarrow{h\to 0} 0$.

Theorem 9 *Let $I \subset \mathbb{R}$ be an open interval, and $f : I \to \mathbb{R}$. Suppose $f \in C^{n+1}([a,b])$ and $x_0 \in I$. Define:*

$$T_n(x_0,h) := \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}\cdot h^k \quad \textit{(Taylor series up to degree } n\textit{)}.$$

$$R_n(x_0,h) := \int_{x_0}^{x_0+h} \frac{(x+h-t)^n}{n!} f^{(n+1)}(t)\,dt \quad \textit{(Remainder term)}.$$

Then:

$$f(x_0+h) = T_n(x_0,h) + R_n(x_0,h).$$

Proof Sketch: The proof follows from the fundamental theorem of calculus and proceeds by induction on n.

- **Base case ($n = 0$)**: We need to prove:

$$f(x_0+h) = f(x_0) + \int_{x_0}^{x_0+h} f'(t)\,dt.$$

  This is equivalent to the fundamental theorem of calculus:

$$\int_a^b F'(x)\,dx = F(b) - F(a).$$

- **Inductive step ($n \to n+1$)**: Consider the function:

$$F(x_0+h) = \frac{(x_0+h-t)^{n+1}}{(n+1)!} f^{(n+1)}(t).$$

  The proof proceeds by integrating and applying the fundamental theorem of calculus.
- **T**ake its derivative.
- **I**ntegrate and exploit fundamental theorem.

△

Theorem 10 (Taylor's Theorem with Lagrange Remainder) *Let $I \subset \mathbb{R}$ be an interval, $f : I \to \mathbb{R}$, and $f \in C^{n+1}(I)$, $x_0 \in I$. If $h \in \mathbb{R}$ such that $x_0 + h \in I$. Then:*

$$f(x_0+h) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}\cdot h^k + R_n(h),$$

*where the first part is **n-th order Taylor polynomial**, and $R_n(h)$ is the remainder term of order n. There exists ξ with $\xi \in (x_0, x_0 + h)$ (or $\xi \in (x_0 + h, x_0)$) such that:*

$$R_n(h) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \cdot h^{n+1}.$$

Often, this is written as:

$$f(x_0 + h) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} \cdot h^k + O(h^{n+1}),$$

or, with $x = x_0 + h$:

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + O((x - x_0)^{n+1}).$$

Proof of Theorem 10:

$$F_{n,h}(t) := \sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(x_0 + h - t)^k,$$

Note:

$$F_{n,h}(x_0) = T_n(x_0, h), \quad F_{n,h}(x_0 + h) = f(x_0 + h)$$

Define:

$$g_{n,h}(t) := (x_0 + h - t)^{n+1}, \quad g'_{n,h}(t) = -(n+1)(x_0 + h - t)^n$$

Generalized Mean Value Theorem

Using the generalized mean value theorem, we have:

$$\frac{F_{n,h}(x_0 + h) - F_{n,h}(x_0)}{g_{n,h}(x_0 + h) - g_{n,h}(x_0)} = \frac{F'_{n,h}(\xi)}{g'_{n,h}(\xi)},$$

where $\xi \in (x_0, x_0 + h)$.

Replacing $F_{n,h}(x_0 + h)$ with $f(x_0 + h)$, we have:

$$\begin{aligned} f(x_0 + h) - T_n(x_0, h) &= (g_{n,h}(x_0 + h) - g_{n,h}(x_0)) \cdot \frac{F'_{n,h}(\xi)}{g'_{n,h}(\xi)} \\ &= \frac{h^{n+1} \cdot F'_{n,h}(\xi)}{(n+1) \cdot (h + x_0 - \xi)^n} \end{aligned}$$

Left clause:

$$\begin{aligned} f(x_0 + h) - T_n(x_0, h) &= F'_{n,h}(t) \\ &= \frac{d}{dt} \sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(h + x_0 - t)^k \\ &= \sum_{k=0}^{n} \frac{f^{(k+1)}(t)}{k!}(h + x_0 - t)^k - \sum_{k=1}^{n} \frac{f^{(k)}(t)}{(k-1)!}(h + x_0 - t)^{k-1} \\ &= \frac{f^{(n+1)}(t)}{n!}(h + x_0 - t)^n \end{aligned}$$

Substituting to right clause:

$$f(x_0+h) - T_n(x_0,h) = \frac{h^{n+1} \cdot \frac{f^{(n+1)}(\xi_p)}{n!} \cdot (h+x_0-\xi)^n}{(n+1)\cdot(h+x_0-\xi)^n}$$
$$= \frac{h^{n+1} \cdot f^{(n+1)}(\xi)}{(n+1)!}$$

Result:

$$f(x_0+h) = T_n(x_0,h) + \frac{h^{n+1} \cdot f^{(n+1)}(\xi)}{(n+1)!}$$

□

Theorem 11 *Let $f \in C^\infty(I)$, $x_0 \in I$, and $h \in \mathbb{R}$ such that $x_0 + h \in I$. Define*

$$T(x_0,h) := \lim_{n\to\infty} T_n(x_0,h) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} \cdot h^n.$$

Then we have $f(x) = T(x)$ if $R_n(x_0,h) \xrightarrow{n\to\infty} 0$.

For example, this is the case if there exist constants $\alpha, c > 0$ such that

$$\left|\frac{f^{(n)}(t)}{f(t)}\right| \leq \alpha \cdot c^n, \quad \forall t \in I, \quad \forall n \in \mathbb{N}.$$

This follows directly from the Lagrangian remainder.

Examples

- Exponential series:

  $$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad \text{(Power series with } r = \infty\text{)}$$

  The exponential function always coincides with its Taylor series.

  Other examples include sin, cos, polynomials, and power series (analytic functions).

- $f(x) = \log(x+1)$, Taylor series around zero.

  To prove: The convergence radius for the Taylor series is $r = 1$. For x outside of $(-1,1)$, the Taylor series does not make sense at all.

- $$f(x) = \begin{cases} \exp\left(-\frac{1}{x^2}\right) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

  This function has the peculiar property that for all $n \in \mathbb{N}$:

  $$f^{(n)}(0) = 0.$$

Consider the Taylor series about $x_0 = 0$. All terms will be zero, i.e., $\forall n : T_n(0,h) = 0$ and $r = \infty$.

$$f(x_0+h) = T_n(x_0,h) + R_n(x_0,h)$$

$$T_n(x_0 = 0, h) = 0 \quad \text{but} \quad f(0+h) = \exp\left(-\frac{1}{h^2}\right)$$

Taylor series around $x_0 = 0$ is zero. Function value around $x_0 = 0$ is not zero.

$$\forall (x_0 + h) \neq 0, \quad T_n(x_0, h) \neq f(x_0 + h)$$

2.1 Taylor Series in Deep Learning

Taylor series can be used to approximate activation functions in deep learning.

- **Softmax Activation**:
$$\text{softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$$
Taylor expansions can approximate the exponential terms for efficient computation. The power series for e^x is:
$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$
By substituting the power series expansion of e^{z_i} and e^{z_j}, we can approximate the softmax function:
$$\text{softmax}(z_i) \approx \frac{1 + z_i + \frac{z_i^2}{2!} + \frac{z_i^3}{3!} + \cdots}{\sum_{j=1}^{K} \left(1 + z_j + \frac{z_j^2}{2!} + \frac{z_j^3}{3!} + \cdots\right)}$$
- **Sigmoid Activation**:
$$\sigma(x) = \frac{1}{1 + e^{-x}} \approx \frac{1}{2} + \frac{x}{4} - \frac{x^3}{48} + \cdots$$
- **Tanh Activation**:
$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \approx x - \frac{x^3}{3} + \frac{2x^5}{15} - \cdots$$
- **ReLU Activation**:
$$\text{ReLU}(x) = \max(0, x)$$
(No Taylor expansion needed, but approximations may be used for variations like Leaky ReLU.)

These approximations simplify calculations, especially in resource-constrained environments.