BSc Mathematics for Computer Scientists 2: X. Root Finding, Approximation of Real Functions by Polynomials, Power Series

Péter Hajnal

Bolyai Institute, Faculty of Science and Informatics, University of Szeged, Hungary

Spring semester 2026

Root Finding: The Basic Problem

Problem

Given: $f : [a, b] \to \mathbb{R}$ a continuous function, differentiable on the interval $]a, b[$.
Find the solutions of $f(x) = 0$, i.e., the roots of f.

Root Finding: The Basic Problem

Problem

Given: $f : [a, b] \to \mathbb{R}$ a continuous function, differentiable on the interval $]a, b[$.
Find the solutions of $f(x) = 0$, i.e., the roots of f.

• Several methods are known. Even today, new ideas are being developed to improve existing algorithms.

Problem

Given: $f : [a, b] \to \mathbb{R}$ a continuous function, differentiable on the interval $]a, b[$.
Find the solutions of $f(x) = 0$, i.e., the roots of f.

- Several methods are known. Even today, new ideas are being developed to improve existing algorithms.

- We present two methods.

Root Finding: Binary Search: Preparations

- The first method is naive; it does not use the differentiability of f.

- The first method is naive; it does not use the differentiability of f.

- We compute an interval I and update it iteratively.

Root Finding: Binary Search: Preparations

- The first method is naive; it does not use the differentiability of f.

- We compute an interval I and update it iteratively. I always contains a root.

Root Finding: Binary Search: Preparations

• The first method is naive; it does not use the differentiability of f.

• We compute an interval I and update it iteratively. I always contains a root. The length of I decreases. If it becomes sufficiently small, then any point in the interval gives a good approximation of the root guaranteed to lie within the interval.

Root Finding: Binary Search: Preparations

• The first method is naive; it does not use the differentiability of f.

• We compute an interval I and update it iteratively. I always contains a root. The length of I decreases. If it becomes sufficiently small, then any point in the interval gives a good approximation of the root guaranteed to lie within the interval.

• We ensure that the current interval contains a root by requiring that the function values at the endpoints have opposite signs.

• The first method is naive; it does not use the differentiability of f.

• We compute an interval I and update it iteratively. I always contains a root. The length of I decreases. If it becomes sufficiently small, then any point in the interval gives a good approximation of the root guaranteed to lie within the interval.

• We ensure that the current interval contains a root by requiring that the function values at the endpoints have opposite signs. We assume this already for the initial interval: $f(a)f(b) < 0$.

Root Finding: Binary Search: The Algorithm

Algorithm: Binary Search for Root Finding

Algorithm: Binary Search for Root Finding

Given $f : [a, b] \to \mathbb{R}$ a continuous function.

Algorithm: Binary Search for Root Finding

Given $f : [a, b] \to \mathbb{R}$ a continuous function.

Initialization: Let $n = 0$, $I_0 = I_n = [a, b]$.

Algorithm: Binary Search for Root Finding

Given $f : [a, b] \to \mathbb{R}$ a continuous function.

Initialization: Let $n = 0$, $I_0 = I_n = [a, b]$.

Update rule: Let $m = \frac{a+b}{2}$.

Algorithm: Binary Search for Root Finding

Given $f : [a, b] \to \mathbb{R}$ a continuous function.

Initialization: Let $n = 0$, $I_0 = I_n = [a, b]$.

Update rule: Let $m = \frac{a+b}{2}$. If $f(m) = 0$, then we are done. We have found a root, namely m.

Algorithm: Binary Search for Root Finding

Given $f : [a, b] \to \mathbb{R}$ a continuous function.

Initialization: Let $n = 0$, $I_0 = I_n = [a, b]$.

Update rule: Let $m = \frac{a+b}{2}$. If $f(m) = 0$, then we are done. We have found a root, namely m.

If $f(m) \neq 0$, then let $c \in \{a, b\}$ be that endpoint of I_n for which the sign of the function differs from the sign of $f(m)$.

Algorithm: Binary Search for Root Finding

Given $f : [a, b] \to \mathbb{R}$ a continuous function.

Initialization: Let $n = 0$, $I_0 = I_n = [a, b]$.

Update rule: Let $m = \frac{a+b}{2}$. If $f(m) = 0$, then we are done. We have found a root, namely m.

If $f(m) \neq 0$, then let $c \in \{a, b\}$ be that endpoint of I_n for which the sign of the function differs from the sign of $f(m)$. Let $n \leftarrow n + 1$. Let $I_n \leftarrow$ "the set of points between m and c".

Root Finding: Binary Search: The Algorithm

Algorithm: Binary Search for Root Finding

Given $f : [a, b] \to \mathbb{R}$ a continuous function.

Initialization: Let $n = 0$, $I_0 = I_n = [a, b]$.

Update rule: Let $m = \frac{a+b}{2}$. If $f(m) = 0$, then we are done. We have found a root, namely m.

If $f(m) \neq 0$, then let $c \in \{a, b\}$ be that endpoint of I_n for which the sign of the function differs from the sign of $f(m)$. Let $n \leftarrow n + 1$. Let $I_n \leftarrow$ "the set of points between m and c".

Stopping rule: If the length of I_n is smaller than ϵ, we stop. Otherwise, we return to the Update rule.

Root Finding: Binary Search: Analysis

• In each iteration step (assuming we have not yet found a root), we split the current interval at its midpoint. Whichever subinterval we choose, the length of the current interval is halved.

Root Finding: Binary Search: Analysis

• In each iteration step (assuming we have not yet found a root), we split the current interval at its midpoint. Whichever subinterval we choose, the length of the current interval is halved.

Theorem

Assume that initially the current interval has unit length, i.e., its length is 1.
Then after k updates, the length of the current interval is $\frac{1}{2^k}$.

• In each iteration step (assuming we have not yet found a root), we split the current interval at its midpoint. Whichever subinterval we choose, the length of the current interval is halved.

Theorem

Assume that initially the current interval has unit length, i.e., its length is 1.
Then after k updates, the length of the current interval is $\frac{1}{2^k}$.

Based on the theorem, any point in the current interval gives an approximation of the guaranteed root such that the error is at most $\frac{1}{2^k}$.

Root Finding: Binary Search: Analysis

• In each iteration step (assuming we have not yet found a root), we split the current interval at its midpoint. Whichever subinterval we choose, the length of the current interval is halved.

Theorem

Assume that initially the current interval has unit length, i.e., its length is 1.
Then after k updates, the length of the current interval is $\frac{1}{2^k}$.

Based on the theorem, any point in the current interval gives an approximation of the guaranteed root such that the error is at most $\frac{1}{2^k}$. In particular, after k updates the approximation is "accurate to k digits in the binary number system".

Root Finding: Newton's Method: Ideas

- We compute a point x_0, which we update iteratively.

Root Finding: Newton's Method: Ideas

- We compute a point x_0, which we update iteratively. The update will be such that it can be interpreted as approaching the root.

Root Finding: Newton's Method: Ideas

• We compute a point x_0, which we update iteratively. The update will be such that it can be interpreted as approaching the root.

• **Idea for the update:** After approximating the root by x_n, we proceed using the tangent.

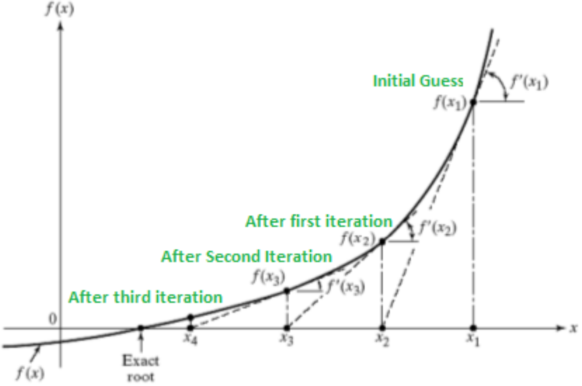
At the point x_n, we write the equation of the tangent:

$$y = f(x_n) + f'(x_n)(x - x_n)$$

We find the root of the tangent equation (a linear equation)$\rightarrow x_{n+1}$. That is, we find the intersection of the tangent with the x-axis ($y = 0$):

$$0 = f(x_n) + f'(x_n)(x_{n+1} - x_n) \quad \Rightarrow \quad x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

Root Finding: Newton's Method: Illustration



Forrás:
https://predictivehacks.com/newton-raphson-method-in-python/

Root Finding: Newton's Method: The Algorithm

Algorithm: Newton's Method for Root Finding

Algorithm: Newton's Method for Root Finding

Given $f : [a, b] \to \mathbb{R}$ a differentiable function that is guaranteed to have a root in $\mathrm{dom}(f)$.

Algorithm: Newton's Method for Root Finding

Given $f : [a, b] \to \mathbb{R}$ a differentiable function that is guaranteed to have a root in $\mathrm{dom}(f)$.

Initialization: Let $n = 0$, $x_0 \in [a, b]$ an "intelligent guess" for a root of f.

Algorithm: Newton's Method for Root Finding

Given $f : [a, b] \to \mathbb{R}$ a differentiable function that is guaranteed to have a root in $\mathrm{dom}(f)$.

Initialization: Let $n = 0$, $x_0 \in [a, b]$ an "intelligent guess" for a root of f.

Update rule: Let $n \leftarrow n + 1$. $x_n \leftarrow x_n - \frac{f(x_n)}{f'(x_n)}$.

Algorithm: Newton's Method for Root Finding

Given $f : [a, b] \to \mathbb{R}$ a differentiable function that is guaranteed to have a root in $\mathrm{dom}(f)$.

Initialization: Let $n = 0$, $x_0 \in [a, b]$ an "intelligent guess" for a root of f.

Update rule: Let $n \leftarrow n + 1$. $x_n \leftarrow x_n - \frac{f(x_n)}{f'(x_n)}$.

Stopping rule: If the difference $|x_n - x_{n-1}|$ is smaller than ϵ, we stop. We report the current x_n as an approximation of the root.

Algorithm: Newton's Method for Root Finding

Given $f : [a, b] \to \mathbb{R}$ a differentiable function that is guaranteed to have a root in $\mathrm{dom}(f)$.

Initialization: Let $n = 0$, $x_0 \in [a, b]$ an "intelligent guess" for a root of f.

Update rule: Let $n \leftarrow n + 1$. $x_n \leftarrow x_n - \frac{f(x_n)}{f'(x_n)}$.

Stopping rule: If the difference $|x_n - x_{n-1}|$ is smaller than ϵ, we stop. We report the current x_n as an approximation of the root. If not, we return to the Update rule.

Algorithm: Newton's Method for Root Finding

Given $f : [a, b] \to \mathbb{R}$ a differentiable function that is guaranteed to have a root in $\mathrm{dom}(f)$.

Initialization: Let $n = 0$, $x_0 \in [a, b]$ an "intelligent guess" for a root of f.

Update rule: Let $n \leftarrow n + 1$. $x_n \leftarrow x_n - \frac{f(x_n)}{f'(x_n)}$.

Stopping rule: If the difference $|x_n - x_{n-1}|$ is smaller than ϵ, we stop. We report the current x_n as an approximation of the root. If not, we return to the Update rule.

Warning: It does not work if $f'(x_n) = 0$, for example if x_n is a local minimum/maximum. The precise formulation of the algorithm is more technical. Our goal is not completeness.

Root Finding: Newton's Method: Analysis

- A precise analysis is technical. It belongs to numerical mathematics, and we do not discuss it.

- A precise analysis is technical. It belongs to numerical mathematics, and we do not discuss it.

- We note that if x_n enters a "basin of attraction" of a root r, then the method converges to the root very quickly.

• A precise analysis is technical. It belongs to numerical mathematics, and we do not discuss it.

• We note that if x_n enters a "basin of attraction" of a root r, then the method converges to the root very quickly.

• The success of the algorithm depends on how quickly we get a sufficiently good approximation of the root.

Example

Let us approximate the value of $\sqrt{2}$.

Example

Let us approximate the value of $\sqrt{2}$.

- The example asks for approximating the positive root of $f(x) = x^2 - 2$.

Example

Let us approximate the value of $\sqrt{2}$.

- The example asks for approximating the positive root of $f(x) = x^2 - 2$.
- $x_0 = 1.5$ is a good "guess" for the initial value.

Root Finding: Newton's Method: Application

Example

Let us approximate the value of $\sqrt{2}$.

- The example asks for approximating the positive root of $f(x) = x^2 - 2$.
- $x_0 = 1.5$ is a good "guess" for the initial value.
- $$x_{n+1} = x_n - \frac{x_n^2 - 2}{2x_n} = \frac{x_n + \frac{2}{x_n}}{2}$$

Based on this, define a sequence:

$x_1 = \frac{1.5 + \frac{2}{1.5}}{2} = \frac{17}{12} = 1.416666667$, $x_2 = \frac{577}{408} = 1.414215686$, $x_3 = \frac{665857}{470832} = 1.414213562$.

Example

Let us approximate the value of $\sqrt{2}$.

- The example asks for approximating the positive root of $f(x) = x^2 - 2$.
- $x_0 = 1.5$ is a good "guess" for the initial value.
- $$x_{n+1} = x_n - \frac{x_n^2 - 2}{2x_n} = \frac{x_n + \frac{2}{x_n}}{2}$$

Based on this, define a sequence:
$x_1 = \frac{1.5+\frac{2}{1.5}}{2} = \frac{17}{12} = 1.416666667$, $x_2 = \frac{577}{408} = 1.414215686$, $x_3 = \frac{665857}{470832} = 1.414213562$.

- $|x_2 - x_3|$ is already very small. $\sqrt{2} \approx \frac{665857}{470832}$. In fact,

$$\left(\frac{665857}{470832}\right)^2 = \frac{443365544449}{221682772224} = 2 + \frac{1}{221682772224}.$$

Example

Let us approximate the value of $\sqrt{2}$.

- The example asks for approximating the positive root of $f(x) = x^2 - 2$.
- $x_0 = 1.5$ is a good "guess" for the initial value.
- $$x_{n+1} = x_n - \frac{x_n^2 - 2}{2x_n} = \frac{x_n + \frac{2}{x_n}}{2}$$

Based on this, define a sequence:
$x_1 = \frac{1.5 + \frac{2}{1.5}}{2} = \frac{17}{12} = 1.416666667$, $x_2 = \frac{577}{408} = 1.414215686$,
$x_3 = \frac{665857}{470832} = 1.414213562$.

- $|x_2 - x_3|$ is already very small. $\sqrt{2} \approx \frac{665857}{470832}$. In fact,

$$\left(\frac{665857}{470832}\right)^2 = \frac{443365544449}{221682772224} = 2 + \frac{1}{221682772224}.$$

- This method was already known to the Babylonians, even though derivatives were not yet known at that time.

Break

Optimization: The Basic Question

Optimization: The Basic Question

• One of the most natural applications of differential calculus is optimization.

Optimization: The Basic Question

• One of the most natural applications of differential calculus is optimization.

• With our current tools, we can only handle functions of one variable.

Optimization: The Basic Question

• One of the most natural applications of differential calculus is optimization.

• With our current tools, we can only handle functions of one variable.

• We consider an example from optics.

Optimization: The Basic Question

• One of the most natural applications of differential calculus is optimization.

• With our current tools, we can only handle functions of one variable.

• We consider an example from optics.

Fermat's Principle

A light ray in an arbitrary optical system always follows a path for which the travel time between the starting and ending points is minimal.

Optimization: The Basic Question

• One of the most natural applications of differential calculus is optimization.

• With our current tools, we can only handle functions of one variable.

• We consider an example from optics.

Fermat's Principle

A light ray in an arbitrary optical system always follows a path for which the travel time between the starting and ending points is minimal.

• Light chooses a path between points P and Q located on two sides of a line. On one side of the line it travels with speed v_1, on the other side with speed v_2.

Optimization: The Basic Question

• One of the most natural applications of differential calculus is optimization.

• With our current tools, we can only handle functions of one variable.

• We consider an example from optics.

Fermat's Principle

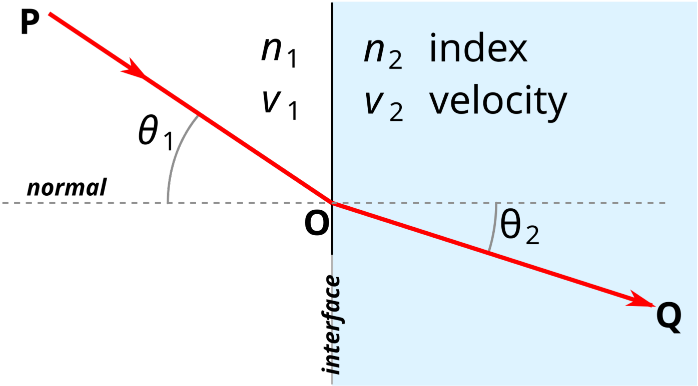
A light ray in an arbitrary optical system always follows a path for which the travel time between the starting and ending points is minimal.

• Light chooses a path between points P and Q located on two sides of a line. On one side of the line it travels with speed v_1, on the other side with speed v_2. What path does it follow?

Optimization: Observations

- The light ray must reach the line at a point O. The path consists of segments PO and OQ, each lying in a homogeneous medium. Here minimal time is equivalent to minimal distance. The path of minimal length is the straight segment PO and the straight segment OQ.

- The light ray must reach the line at a point O. The path consists of segments PO and OQ, each lying in a homogeneous medium. Here minimal time is equivalent to minimal distance. The path of minimal length is the straight segment PO and the straight segment OQ.



Optimization: Formalization

- Let the equation of the line be $x = 0$. Let $P(-a, 0)$ and $Q(b, -c)$, where $a, b, c > 0$.

- Let the equation of the line be $x = 0$. Let $P(-a, 0)$ and $Q(b, -c)$, where $a, b, c > 0$.
- Let $O(0, x)$, where x is the only variable.

Optimization: Formalization

- Let the equation of the line be $x = 0$. Let $P(-a, 0)$ and $Q(b, -c)$, where $a, b, c > 0$.

- Let $O(0, x)$, where x is the only variable.

- The total time:

$$T(x) = \frac{\sqrt{x^2 + a^2}}{v_1} + \frac{\sqrt{(c - x)^2 + b^2}}{v_2}.$$

By Fermat's principle, we seek the minimum of $T(x)$.

Optimization: Differentiation

Optimization: Differentiation

- The optimal x is among the roots of the derivative.

Optimization: Differentiation

- The optimal x is among the roots of the derivative.
- The derivative:

$$T'(x) = \frac{1}{v_1}\frac{1}{2}\frac{2x}{\sqrt{x^2+a^2}} + \frac{1}{v_2}\frac{1}{2}\frac{-2(c-x)}{\sqrt{(c-x)^2+b^2}}.$$

Optimization: Differentiation

- The optimal x is among the roots of the derivative.
- The derivative:

$$T'(x) = \frac{1}{v_1}\frac{1}{2}\frac{2x}{\sqrt{x^2+a^2}} + \frac{1}{v_2}\frac{1}{2}\frac{-2(c-x)}{\sqrt{(c-x)^2+b^2}}.$$

- At the root of the derivative:

$$\frac{\sin\theta_1}{v_1} = \frac{\sin\theta_2}{v_2},$$

where θ_1 is the angle of incidence and θ_2 is the angle of refraction.

- The optimal x is among the roots of the derivative.
- The derivative:

$$T'(x) = \frac{1}{v_1}\frac{1}{2}\frac{2x}{\sqrt{x^2+a^2}} + \frac{1}{v_2}\frac{1}{2}\frac{-2(c-x)}{\sqrt{(c-x)^2+b^2}}.$$

- At the root of the derivative:

$$\frac{\sin\theta_1}{v_1} = \frac{\sin\theta_2}{v_2},$$

where θ_1 is the angle of incidence and θ_2 is the angle of refraction.

Theorem [Snell–Descartes Law]

For the path of the refracted light ray:

$$n_1 \sin\theta_1 = n_2 \sin\theta_2,$$

where $n = c/v$ is the refractive index.

Break

Polynomials

Polynomials

Definition

$$\mathbb{R}[x] = \{0\} \cup \{p(x) : p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_d x^d, \text{ where } a_0, a_1, a_2, \ldots, a_d \in \mathbb{R}, a_d \neq 0\}$$

Polynomials

Definition

$$\mathbb{R}[x] = \{0\} \cup \{p(x) : p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_d x^d, \text{ where } a_0, a_1, a_2, \ldots, a_d \in \mathbb{R}, a_d \neq 0\}$$

- Polynomials are simple and elegant.

Polynomials

Definition

$$\mathbb{R}[x] = \{0\} \cup \{p(x) : p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_d x^d, \text{ where } a_0, a_1, a_2, \ldots, a_d \in \mathbb{R}, a_d \neq 0\}$$

- Polynomials are simple and elegant.
- Polynomials are continuous and infinitely differentiable.

Polynomials

Definition

$$\mathbb{R}[x] = \{0\} \cup \{p(x) : p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_d x^d, \text{ where } a_0, a_1, a_2, \ldots, a_d \in \mathbb{R}, a_d \neq 0\}$$

- Polynomials are simple and elegant.

- Polynomials are continuous and infinitely differentiable. The derivative can be computed easily and efficiently.

Polynomials

Definition

$$\mathbb{R}[x] = \{0\} \cup \{p(x) : p(x) = a_0 + a_1x + a_2x^2 + \ldots + a_dx^d, \text{ where } a_0, a_1, a_2, \ldots, a_d \in \mathbb{R}, a_d \neq 0\}$$

- Polynomials are simple and elegant.

- Polynomials are continuous and infinitely differentiable. The derivative can be computed easily and efficiently.

- We approximate arbitrary (even complicated) functions $f : [a, b] \to \mathbb{R}$ by polynomials.

Polynomials

Definition

$$\mathbb{R}[x] = \{0\} \cup \{p(x) : p(x) = a_0 + a_1x + a_2x^2 + \ldots + a_dx^d, \text{ where } a_0, a_1, a_2, \ldots, a_d \in \mathbb{R}, a_d \neq 0\}$$

- Polynomials are simple and elegant.

- Polynomials are continuous and infinitely differentiable. The derivative can be computed easily and efficiently.

- We approximate arbitrary (even complicated) functions $f : [a, b] \to \mathbb{R}$ by polynomials.

- When do we consider an approximation good? Which of two approximations is better?

Polynomials: Interpolation: Basic question

Given a function $f : [a, b] \to \mathbb{R}$.

Given a function $f : [a, b] \to \mathbb{R}$.

Given places $t_1 < t_2 < \ldots < t_n$ and values $\varphi_1 = f(t_1)$, $\varphi_2 = f(t_2)$, $\ldots$, $\varphi_n = f(t_n)$ (function values / measurement data).

Given a function $f : [a, b] \to \mathbb{R}$.

Given places $t_1 < t_2 < \ldots < t_n$ and values $\varphi_1 = f(t_1)$, $\varphi_2 = f(t_2)$, $\ldots$, $\varphi_n = f(t_n)$ (function values / measurement data).

Find a polynomial $p \in \mathbb{R}[x]$ such that

$$p(t_1) = \varphi_1, \quad p(t_2) = \varphi_2, \quad \ldots \quad , p(t_n) = \varphi_n.$$

Polynomials: Interpolation: Basic question

Given a function $f : [a, b] \to \mathbb{R}$.

Given places $t_1 < t_2 < \ldots < t_n$ and values $\varphi_1 = f(t_1)$, $\varphi_2 = f(t_2)$, $\ldots$, $\varphi_n = f(t_n)$ (function values / measurement data).

Find a polynomial $p \in \mathbb{R}[x]$ such that

$$p(t_1) = \varphi_1, \quad p(t_2) = \varphi_2, \quad \ldots \quad , p(t_n) = \varphi_n.$$

That is, fit a "low-degree" polynomial to the data coming from f.

Theorem

Given places $t_1 < t_2 < \ldots < t_n$ and values φ_1, φ_2, ..., φ_n. Then there exists exactly one polynomial p of degree at most $n-1$ such that

$$\varphi_1 = f(t_1), \quad \varphi_2 = f(t_2), \quad \ldots, \quad \varphi_n = f(t_n).$$

Polynomials: Interpolation: Lagrange

Theorem

Given places $t_1 < t_2 < \ldots < t_n$ and values φ_1, φ_2, …, φ_n. Then there exists exactly one polynomial p of degree at most $n-1$ such that

$$\varphi_1 = f(t_1), \quad \varphi_2 = f(t_2), \quad \ldots, \quad \varphi_n = f(t_n).$$

- The theorem contains two statements: existence and uniqueness.

Theorem

Given places $t_1 < t_2 < \ldots < t_n$ and values φ_1, φ_2, ..., φ_n. Then there exists exactly one polynomial p of degree at most $n-1$ such that

$$\varphi_1 = f(t_1), \quad \varphi_2 = f(t_2), \quad \ldots, \quad \varphi_n = f(t_n).$$

- The theorem contains two statements: existence and uniqueness.
- We outline two approaches to prove the theorem.

Polynomials: Interpolation: Lagrange: First proof

Given $t_1 < t_2 < \ldots < t_n$ and φ_1, φ_2, …, φ_n real numbers.

Given $t_1 < t_2 < \ldots < t_n$ and φ_1, φ_2, ..., φ_n real numbers.

We seek a polynomial $p(x)$ / coefficients $a_0, a_1, \ldots, a_{n-1} \in \mathbb{R}$ such that

$$\begin{array}{ccc} a_0 + a_1 t_1 + a_2 t_1^2 + & \ldots & + a_{n-1} t_1^{n-1} = \varphi_1 \\ a_0 + a_1 t_2 + a_2 t_2^2 + & \ldots & + a_{n-1} t_2^{n-1} = \varphi_2 \\ & \vdots & \\ a_0 + a_1 t_n + a_2 t_n^2 + & \ldots & + a_{n-1} t_n^{n-1} = \varphi_n. \end{array} \tag{1}$$

Given $t_1 < t_2 < \ldots < t_n$ and φ_1, φ_2, …, φ_n real numbers.

We seek a polynomial $p(x)$ / coefficients $a_0, a_1, \ldots, a_{n-1} \in \mathbb{R}$ such that

$$\begin{array}{cccc} a_0 + a_1 t_1 + a_2 t_1^2 + & \ldots & + a_{n-1} t_1^{n-1} = \varphi_1 \\ a_0 + a_1 t_2 + a_2 t_2^2 + & \ldots & + a_{n-1} t_2^{n-1} = \varphi_2 \\ & \vdots & \\ a_0 + a_1 t_n + a_2 t_n^2 + & \ldots & + a_{n-1} t_n^{n-1} = \varphi_n. \end{array} \tag{1}$$

• We have n linear conditions for the unknowns $a_0, a_1, a_2, \ldots, a_{n-1}$ (n unknowns).

Given $t_1 < t_2 < \ldots < t_n$ and $\varphi_1, \varphi_2, \ldots, \varphi_n$ real numbers.

We seek a polynomial $p(x)$ / coefficients $a_0, a_1, \ldots, a_{n-1} \in \mathbb{R}$ such that

$$\begin{aligned} a_0 + a_1 t_1 + a_2 t_1^2 + \quad &\ldots \quad + a_{n-1} t_1^{n-1} = \varphi_1 \\ a_0 + a_1 t_2 + a_2 t_2^2 + \quad &\ldots \quad + a_{n-1} t_2^{n-1} = \varphi_2 \\ &\vdots \\ a_0 + a_1 t_n + a_2 t_n^2 + \quad &\ldots \quad + a_{n-1} t_n^{n-1} = \varphi_n. \end{aligned} \tag{1}$$

- We have n linear conditions for the unknowns $a_0, a_1, a_2, \ldots, a_{n-1}$ (n unknowns).
- The coefficient matrix of the system is

$$M_L = \begin{pmatrix} 1 & t_1 & t_1^2 & \ldots & t_1^{n-1} \\ 1 & t_2 & t_2^2 & \ldots & t_2^{n-1} \\ \vdots & & & & \vdots \\ 1 & t_n & t_n^2 & \ldots & t_n^{n-1} \end{pmatrix}.$$

Polynomials: Interpolation: Lagrange: First proof

Theorem

(i) The system of equations (1) has a unique solution.

(ii)
$$\det M_L = \prod_{1 \le i < j \le n} (t_i - t_j) \neq 0.$$

Theorem

(i) The system of equations (1) has a unique solution.

(ii)

$$\det M_L = \prod_{1\le i<j\le n} (t_i - t_j) \neq 0.$$

- The remaining part of the proof follows from linear algebra.

Polynomials: Interpolation: Lagrange: Second proof

Theorem

Given places $t_1 < t_2 < \ldots < t_n$. Then there exist polynomials $p_i \in \mathbb{R}[x]$ ($i = 1, 2, \ldots, n$) such that

$$p_i(t_1) = p_i(t_2) = \ldots = p_i(t_{i-1}) = p_i(t_{i+1}) = p_i(t_{i+2}) = \ldots = p_i(t_n) = 0,$$

and moreover $p_i(t_i) \neq 0$.

Theorem

Given places $t_1 < t_2 < \ldots < t_n$. Then there exist polynomials $p_i \in \mathbb{R}[x]$ $(i = 1, 2, \ldots, n)$ such that

$$p_i(t_1) = p_i(t_2) = \ldots = p_i(t_{i-1}) = p_i(t_{i+1}) = p_i(t_{i+2}) = \ldots = p_i(t_n) = 0,$$

and moreover $p_i(t_i) \neq 0$.

Proof: $p_i(x) = (x - t_1)(x - t_2) \ldots (x - t_{i-1})(x - t_{i+1}) \ldots (x - t_n)$.

Theorem

Given places $t_1 < t_2 < \ldots < t_n$. Then there exist polynomials $p_i \in \mathbb{R}[x]$ $(i = 1, 2, \ldots, n)$ such that

$$p_i(t_1) = p_i(t_2) = \ldots = p_i(t_{i-1}) = p_i(t_{i+1}) = p_i(t_{i+2}) = \ldots = p_i(t_n) = 0,$$

and moreover $p_i(t_i) \neq 0$.

Proof: $p_i(x) = (x - t_1)(x - t_2)\ldots(x - t_{i-1})(x - t_{i+1})\ldots(x - t_n)$.

Theorem

Given places $t_1 < t_2 < \ldots < t_n$. Then there exist polynomials $p_i \in \mathbb{R}[x]$ $(i = 1, 2, \ldots, n)$ such that

$$p_i(t_1) = p_i(t_2) = \ldots = p_i(t_{i-1}) = p_i(t_{i+1}) = p_i(t_{i+2}) = \ldots = p_i(t_n) = 0,$$

and moreover $p_i(t_i) = 1$.

Polynomials: Interpolation: Lagrange: Second proof

Theorem

Given places $t_1 < t_2 < \ldots < t_n$. Then there exist polynomials $p_i \in \mathbb{R}[x]$ $(i = 1, 2, \ldots, n)$ such that

$$p_i(t_1) = p_i(t_2) = \ldots = p_i(t_{i-1}) = p_i(t_{i+1}) = p_i(t_{i+2}) = \ldots = p_i(t_n) = 0,$$

and moreover $p_i(t_i) \neq 0$.

Proof: $p_i(x) = (x - t_1)(x - t_2)\ldots(x - t_{i-1})(x - t_{i+1})\ldots(x - t_n)$.

Theorem

Given places $t_1 < t_2 < \ldots < t_n$. Then there exist polynomials $p_i \in \mathbb{R}[x]$ $(i = 1, 2, \ldots, n)$ such that

$$p_i(t_1) = p_i(t_2) = \ldots = p_i(t_{i-1}) = p_i(t_{i+1}) = p_i(t_{i+2}) = \ldots = p_i(t_n) = 0,$$

and moreover $p_i(t_i) = 1$.

Proof: $\widehat{p}_i(x) = \frac{1}{(t_i - t_1)(t_i - t_2)\ldots(t_i - t_{i-1})(t_i - t_{i+1})\ldots(t_i - t_n)} p_i(x)$.

Polynomials: Interpolation: Lagrange: Second proof

Theorem

Given places $t_1 < t_2 < \ldots < t_n$. Then there exists a polynomial $p \in \mathbb{R}[x]$ such that

$$p(t_1) = \varphi_1, \quad p(t_2) = \varphi_2, \ldots, p(t_n) = \varphi_n.$$

Theorem

Given places $t_1 < t_2 < \ldots < t_n$. Then there exists a polynomial $p \in \mathbb{R}[x]$ such that

$$p(t_1) = \varphi_1, \quad p(t_2) = \varphi_2, \ldots, p(t_n) = \varphi_n.$$

Proof:

$$p(x) = \varphi_1 \widehat{p}_1(x) + \varphi_2 \widehat{p}_2(x) + \ldots + \varphi_n \widehat{p}_n(x).$$

Theorem

Given places $t_1 < t_2 < \ldots < t_n$. Then there exists a polynomial $p \in \mathbb{R}[x]$ such that

$$p(t_1) = \varphi_1, \quad p(t_2) = \varphi_2, \ldots, p(t_n) = \varphi_n.$$

Proof:

$$p(x) = \varphi_1 \widehat{p}_1(x) + \varphi_2 \widehat{p}_2(x) + \ldots + \varphi_n \widehat{p}_n(x).$$

- Uniqueness can be proven independently.

Break

Weierstrass approximation: The basic question

Weierstrass approximation: The basic question

Given a function $f : [a, b] \to \mathbb{R}$ and an arbitrary parameter $\varepsilon > 0$.

Weierstrass approximation: The basic question

Given a function $f : [a, b] \to \mathbb{R}$ and an arbitrary parameter $\varepsilon > 0$.

Find a real polynomial p_ε such that

$$|f(x) - p_\varepsilon(x)| < \varepsilon, \quad \text{for all } x \in [a, b] = \mathrm{dom}(f).$$

Given a function $f : [a, b] \to \mathbb{R}$ and an arbitrary parameter $\varepsilon > 0$.

Find a real polynomial p_ε such that

$$|f(x) - p_\varepsilon(x)| < \varepsilon, \qquad \text{for all } x \in [a, b] = \text{dom}(f).$$

Observation

If f is not continuous, then such a polynomial certainly does not exist.

Weierstrass approximation: The basic question

Given a function $f : [a, b] \to \mathbb{R}$ and an arbitrary parameter $\varepsilon > 0$.

Find a real polynomial p_ε such that

$$|f(x) - p_\varepsilon(x)| < \varepsilon, \qquad \text{for all } x \in [a, b] = \mathrm{dom}(f).$$

Observation

If f is not continuous, then such a polynomial certainly does not exist.

Why?

Weierstrass approximation: The basic question

Given a function $f : [a, b] \to \mathbb{R}$ and an arbitrary parameter $\varepsilon > 0$.

Find a real polynomial p_ε such that

$$|f(x) - p_\varepsilon(x)| < \varepsilon, \qquad \text{for all } x \in [a, b] = \mathrm{dom}(f).$$

Observation

If f is not continuous, then such a polynomial certainly does not exist.

Why?

From now on, assume that the function f is continuous.

Weierstrass approximation: The main theorem

Theorem: Weierstrass approximation theorem

Let $f : [a, b] \to \mathbb{R}$ be any continuous function. Given an arbitrary $\varepsilon > 0$. Then there exists a polynomial $p_\varepsilon(x) \in \mathbb{R}[x]$ such that

$$|f(x) - p_\varepsilon(x)| < \varepsilon, \qquad \text{for all } x \in [a, b] = \mathrm{dom}(f).$$

Theorem: Weierstrass approximation theorem

Let $f : [a, b] \to \mathbb{R}$ be any continuous function. Given an arbitrary $\varepsilon > 0$. Then there exists a polynomial $p_\varepsilon(x) \in \mathbb{R}[x]$ such that

$$|f(x) - p_\varepsilon(x)| < \varepsilon, \qquad \text{for all } x \in [a, b] = \mathrm{dom}(f).$$

Proof:

Theorem: Weierstrass approximation theorem

Let $f : [a, b] \to \mathbb{R}$ be any continuous function. Given an arbitrary $\varepsilon > 0$. Then there exists a polynomial $p_\varepsilon(x) \in \mathbb{R}[x]$ such that

$$|f(x) - p_\varepsilon(x)| < \varepsilon, \qquad \text{for all } x \in [a, b] = \mathrm{dom}(f).$$

Proof:

- Without loss of generality, we may assume that $[a, b] = [0, 1]$.

Theorem: Weierstrass approximation theorem

Let $f : [a, b] \to \mathbb{R}$ be any continuous function. Given an arbitrary $\varepsilon > 0$. Then there exists a polynomial $p_\varepsilon(x) \in \mathbb{R}[x]$ such that

$$|f(x) - p_\varepsilon(x)| < \varepsilon, \qquad \text{for all } x \in [a, b] = \mathrm{dom}(f).$$

Proof:

- Without loss of generality, we may assume that $[a, b] = [0, 1]$.
- Let (the n-th degree Bernstein polynomial of f on $[0, 1]$):

$$B_{f,n}(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}x^k(1-x)^{n-k}.$$

Theorem: Weierstrass approximation theorem

Let $f : [a, b] \to \mathbb{R}$ be any continuous function. Given an arbitrary $\varepsilon > 0$. Then there exists a polynomial $p_\varepsilon(x) \in \mathbb{R}[x]$ such that

$$|f(x) - p_\varepsilon(x)| < \varepsilon, \qquad \text{for all } x \in [a, b] = \mathrm{dom}(f).$$

Proof:

- Without loss of generality, we may assume that $[a, b] = [0, 1]$.
- Let (the n-th degree Bernstein polynomial of f on $[0, 1]$):

$$B_{f,n}(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-x)^{n-k}.$$

- We verify that for any $\varepsilon > 0$, a sufficiently large degree $B_{f,n}(x)$ is appropriate.

Weierstrass approximation: The main theorem

Theorem: Weierstrass approximation theorem

Let $f : [a, b] \to \mathbb{R}$ be any continuous function. Given an arbitrary $\varepsilon > 0$. Then there exists a polynomial $p_\varepsilon(x) \in \mathbb{R}[x]$ such that

$$|f(x) - p_\varepsilon(x)| < \varepsilon, \qquad \text{for all } x \in [a, b] = \mathrm{dom}(f).$$

Proof:

- Without loss of generality, we may assume that $[a, b] = [0, 1]$.
- Let (the n-th degree Bernstein polynomial of f on $[0, 1]$):

$$B_{f,n}(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-x)^{n-k}.$$

- We verify that for any $\varepsilon > 0$, a sufficiently large degree $B_{f,n}(x)$ is appropriate. This is technical, so we omit the verification.

Weierstrass approximation: Example

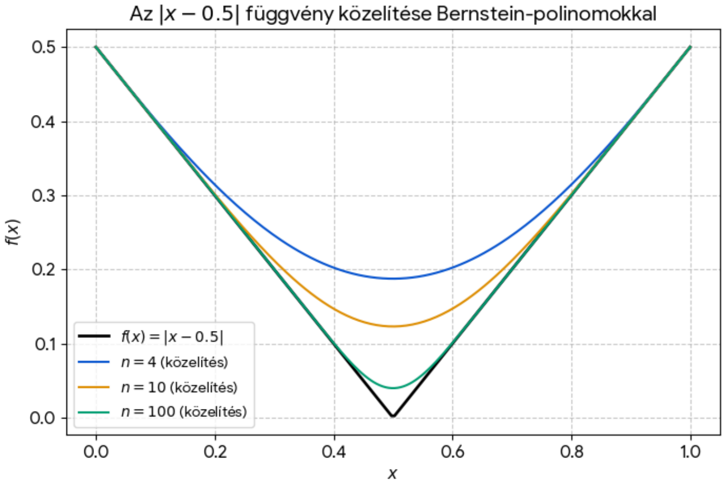


Figure: Three approximations of the function $f(x) = |x - 0.5|$ on the interval $[0, 1]$.

Weierstrass approximation: Example: Explanation

- 4th-degree approximating polynomial (blue graph).

- 4th-degree approximating polynomial (blue graph).

Bernstein formula:

$$\begin{aligned} P_4(x) &= \sum_{k=0}^{4} \left|\frac{k}{4} - 0.5\right| \binom{4}{k} x^k (1-x)^{4-k} \\ &= 0.5(1-x)^4 + 0.25 \cdot 4x(1-x)^3 + 0.25 \cdot 4x^3(1-x) + 0.5x^4. \end{aligned}$$

- 4th-degree approximating polynomial (blue graph).

Bernstein formula:

$$\begin{aligned} P_4(x) = & \sum_{k=0}^{4} \left|\frac{k}{4} - 0.5\right| \binom{4}{k} x^k (1-x)^{4-k} \\ = & \ 0.5(1-x)^4 + 0.25 \cdot 4x(1-x)^3 + 0.25 \cdot 4x^3(1-x) + 0.5x^4. \end{aligned}$$

$\epsilon \approx 0.125.$

- 4th-degree approximating polynomial (blue graph).

Bernstein formula:

$$\begin{aligned} P_4(x) &= \sum_{k=0}^{4} \left|\frac{k}{4} - 0.5\right| \binom{4}{k} x^k (1-x)^{4-k} \\ &= 0.5(1-x)^4 + 0.25 \cdot 4x(1-x)^3 + 0.25 \cdot 4x^3(1-x) + 0.5x^4. \end{aligned}$$

$\epsilon \approx 0.125$.

- Let the degree be 10 (orange graph).

• 4th-degree approximating polynomial (blue graph).
Bernstein formula:

$$\begin{aligned} P_4(x) &= \sum_{k=0}^{4} \left|\frac{k}{4} - 0.5\right| \binom{4}{k} x^k (1-x)^{4-k} \\ &= 0.5(1-x)^4 + 0.25 \cdot 4x(1-x)^3 + 0.25 \cdot 4x^3(1-x) + 0.5x^4. \end{aligned}$$

$\epsilon \approx 0.125$.

• Let the degree be 10 (orange graph).
The formula ($\epsilon \approx 0.079$):

$$P_{10}(x) = \sum_{k=0}^{10} \left|\frac{k}{10} - 0.5\right| \binom{10}{k} x^k (1-x)^{10-k}.$$

Weierstrass approximation: Example: Explanation

• 4th-degree approximating polynomial (blue graph).
Bernstein formula:

$$\begin{aligned} P_4(x) &= \sum_{k=0}^{4} \left|\frac{k}{4}-0.5\right| \binom{4}{k} x^k (1-x)^{4-k} \\ &= 0.5(1-x)^4 + 0.25\cdot 4x(1-x)^3 + 0.25\cdot 4x^3(1-x) + 0.5x^4. \end{aligned}$$

$\epsilon \approx 0.125$.

• Let the degree be 10 (orange graph).
The formula ($\epsilon \approx 0.079$):

$$P_{10}(x) = \sum_{k=0}^{10} \left|\frac{k}{10}-0.5\right| \binom{10}{k} x^k (1-x)^{10-k}.$$

• Let the degree be 100 (green graph).

Weierstrass approximation: Example: Explanation

- 4th-degree approximating polynomial (blue graph).

Bernstein formula:

$$\begin{aligned} P_4(x) = & \sum_{k=0}^{4} \left|\frac{k}{4} - 0.5\right| \binom{4}{k} x^k (1-x)^{4-k} \\ = & \ 0.5(1-x)^4 + 0.25 \cdot 4x(1-x)^3 + 0.25 \cdot 4x^3(1-x) + 0.5x^4. \end{aligned}$$

$\epsilon \approx 0.125$.

- Let the degree be 10 (orange graph).

The formula ($\epsilon \approx 0.079$):

$$P_{10}(x) = \sum_{k=0}^{10} \left|\frac{k}{10} - 0.5\right| \binom{10}{k} x^k (1-x)^{10-k}.$$

- Let the degree be 100 (green graph).

The formula ($\epsilon \approx 0.035$):

$$P_{100}(x) = \sum_{k=0}^{100} \left|\frac{k}{100} - 0.5\right| \binom{100}{k} x^k (1-x)^{100-k}.$$

Break

Taylor Polynomials: Introduction

Taylor Polynomials: Introduction

Reminder

Let $f : (a, b) \to \mathbb{R}$ be a differentiable function. Let $x_0 \in \mathrm{dom}(f)$. The linear function ℓ that best approximates f ($\ell(x_0) = f(x_0)$, $\ell'(x_0) = f'(x_0)$):

$$\ell(x) = f(x_0) + f'(x_0)(x - x_0) \approx f(x).$$

Reminder

Let $f : (a, b) \to \mathbb{R}$ be a differentiable function. Let $x_0 \in \mathrm{dom}(f)$. The linear function ℓ that best approximates f ($\ell(x_0) = f(x_0)$, $\ell'(x_0) = f'(x_0)$):

$$\ell(x) = f(x_0) + f'(x_0)(x - x_0) \approx f(x).$$

- Let us look for a quadratic polynomial $q(x) = a(x - x_0)^2 + b(x - x_0) + c$ such that $q(x_0) = f(x_0)$, $q'(x_0) = f'(x_0)$ and $q''(x_0) = f''(x_0)$

Reminder

Let $f : (a, b) \to \mathbb{R}$ be a differentiable function. Let $x_0 \in \mathrm{dom}(f)$. The linear function ℓ that best approximates f ($\ell(x_0) = f(x_0)$, $\ell'(x_0) = f'(x_0)$):

$$\ell(x) = f(x_0) + f'(x_0)(x - x_0) \approx f(x).$$

- Let us look for a quadratic polynomial $q(x) = a(x - x_0)^2 + b(x - x_0) + c$ such that $q(x_0) = f(x_0)$, $q'(x_0) = f'(x_0)$ and $q''(x_0) = f''(x_0)$
- Writing out the conditions:

$$\begin{aligned} q(x_0) &= c = f(x_0), \\ q'(x_0) &= b = f'(x_0), \\ q''(x_0) &= 2a = f''(x_0). \end{aligned}$$

Taylor Polynomials: The first steps

Observation

Let $f : (a, b) \to \mathbb{R}$ be a twice differentiable function. Let $x_0 \in \mathrm{dom}(f)$.
Then there exists exactly one polynomial of degree at most 2 such that $q(x_0) = f(x_0)$, $q'(x_0) = f'(x_0)$ and $q''(x_0) = f''(x_0)$. This polynomial is:

$$q(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2.$$

Observation

Let $f : (a, b) \to \mathbb{R}$ be a twice differentiable function. Let $x_0 \in \mathrm{dom}(f)$.
Then there exists exactly one polynomial of degree at most 2 such that $q(x_0) = f(x_0)$, $q'(x_0) = f'(x_0)$ and $q''(x_0) = f''(x_0)$. This polynomial is:

$$q(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2.$$

- We were looking for a polynomial of degree at most 2, i.e., three coefficients.

Observation

Let $f:(a,b)\to\mathbb{R}$ be a twice differentiable function. Let $x_0\in\mathrm{dom}(f)$.
Then there exists exactly one polynomial of degree at most 2 such that $q(x_0)=f(x_0)$, $q'(x_0)=f'(x_0)$ and $q''(x_0)=f''(x_0)$. This polynomial is:

$$q(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{f''(x_0)}{2}(x-x_0)^2.$$

- We were looking for a polynomial of degree at most 2, i.e., three coefficients. We had three conditions.

Observation

Let $f : (a, b) \to \mathbb{R}$ be a twice differentiable function. Let $x_0 \in \mathrm{dom}(f)$.
Then there exists exactly one polynomial of degree at most 2 such that $q(x_0) = f(x_0)$, $q'(x_0) = f'(x_0)$ and $q''(x_0) = f''(x_0)$. This polynomial is:

$$q(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2.$$

• We were looking for a polynomial of degree at most 2, i.e., three coefficients. We had three conditions. Writing them out gave three linear conditions.

Taylor Polynomials: The first steps

Observation

Let $f : (a, b) \to \mathbb{R}$ be a twice differentiable function. Let $x_0 \in \mathrm{dom}(f)$.
Then there exists exactly one polynomial of degree at most 2 such that $q(x_0) = f(x_0)$, $q'(x_0) = f'(x_0)$ and $q''(x_0) = f''(x_0)$. This polynomial is:

$$q(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2.$$

• We were looking for a polynomial of degree at most 2, i.e., three coefficients. We had three conditions. Writing them out gave three linear conditions. Solving the system of equations was straightforward.

Observation

Let $f : (a, b) \to \mathbb{R}$ be a twice differentiable function. Let $x_0 \in \mathrm{dom}(f)$.
Then there exists exactly one polynomial of degree at most 2 such that $q(x_0) = f(x_0)$, $q'(x_0) = f'(x_0)$ and $q''(x_0) = f''(x_0)$. This polynomial is:

$$q(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2.$$

- We were looking for a polynomial of degree at most 2, i.e., three coefficients. We had three conditions. Writing them out gave three linear conditions. Solving the system of equations was straightforward.
- Notice the similarity with the linear algebraic solution of the Lagrange interpolation theorem.

Taylor Polynomials: The first steps

Observation

Let $f : (a, b) \to \mathbb{R}$ be a twice differentiable function. Let $x_0 \in \mathrm{dom}(f)$.
Then there exists exactly one polynomial of degree at most 2 such that $q(x_0) = f(x_0)$, $q'(x_0) = f'(x_0)$ and $q''(x_0) = f''(x_0)$. This polynomial is:

$$q(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2.$$

- We were looking for a polynomial of degree at most 2, i.e., three coefficients. We had three conditions. Writing them out gave three linear conditions. Solving the system of equations was straightforward.
- Notice the similarity with the linear algebraic solution of the Lagrange interpolation theorem.
- The generalization is straightforward.

Taylor Polynomials: Definition

Taylor Polynomials: Definition

Observation

Let $f : (a, b) \to \mathbb{R}$ be an n times differentiable function. Let $x_0 \in \mathrm{dom}(f)$. Then there exists exactly one polynomial of degree at most n such that $p_{f,n,x_0}(x_0) = f(x_0)$, $p'_{f,n,x_0}(x_0) = f'(x_0)$, $\ldots$, $p^{(n)}_{f,n,x_0}(x_0) = f^{(n)}(x_0)$. This polynomial is:

$$\begin{aligned} p_{f,n,x_0}(x) = \quad & f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 \\ & + \frac{f'''(x_0)}{3!}(x - x_0)^3 + \ldots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n. \end{aligned}$$

Taylor Polynomials: Definition

Observation

Let $f : (a, b) \to \mathbb{R}$ be an n times differentiable function. Let $x_0 \in \mathrm{dom}(f)$. Then there exists exactly one polynomial of degree at most n such that $p_{f,n,x_0}(x_0) = f(x_0)$, $p'_{f,n,x_0}(x_0) = f'(x_0)$, $\ldots$, $p^{(n)}_{f,n,x_0}(x_0) = f^{(n)}(x_0)$. This polynomial is:

$$\begin{aligned} p_{f,n,x_0}(x) = \quad & f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 \\ & + \frac{f'''(x_0)}{3!}(x - x_0)^3 + \ldots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n. \end{aligned}$$

• Finding a polynomial of degree at most n means finding $n + 1$ coefficients under $n + 1$ linear conditions.

Taylor Polynomials: Definition

Observation

Let $f:(a,b)\to\mathbb{R}$ be an n times differentiable function. Let $x_0\in\mathrm{dom}(f)$. Then there exists exactly one polynomial of degree at most n such that $p_{f,n,x_0}(x_0)=f(x_0)$, $p'_{f,n,x_0}(x_0)=f'(x_0)$, $\ldots$, $p^{(n)}_{f,n,x_0}(x_0)=f^{(n)}(x_0)$. This polynomial is:

$$\begin{aligned}p_{f,n,x_0}(x) = \quad & f(x_0)+f'(x_0)(x-x_0)+\frac{f''(x_0)}{2}(x-x_0)^2\\ & +\frac{f'''(x_0)}{3!}(x-x_0)^3+\ldots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n.\end{aligned}$$

• Finding a polynomial of degree at most n means finding $n+1$ coefficients under $n+1$ linear conditions. The system has a unique solution.

Taylor Polynomials: Definition

Observation

Let $f:(a,b)\to\mathbb{R}$ be an n times differentiable function. Let $x_0\in\mathrm{dom}(f)$. Then there exists exactly one polynomial of degree at most n such that $p_{f,n,x_0}(x_0)=f(x_0)$, $p'_{f,n,x_0}(x_0)=f'(x_0)$, $\ldots$, $p^{(n)}_{f,n,x_0}(x_0)=f^{(n)}(x_0)$. This polynomial is:

$$\begin{aligned} p_{f,n,x_0}(x) = \quad & f(x_0)+f'(x_0)(x-x_0)+\frac{f''(x_0)}{2}(x-x_0)^2 \\ & +\frac{f'''(x_0)}{3!}(x-x_0)^3+\ldots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n. \end{aligned}$$

- Finding a polynomial of degree at most n means finding $n+1$ coefficients under $n+1$ linear conditions. The system has a unique solution.

Definition

The polynomial in the previous theorem is called the n-th order Taylor polynomial of f around x_0. Its notation is T_{f,n,x_0}. If f and x_0 are clear from the context, we simply write $T_n(x)$.

Taylor Polynomials: Remainder

Definition: Remainder (error term) of the Taylor polynomial

$$R_{f,n,x_0}(x) = f(x) - T_{f,n,x_0}(x),$$

that is $f(x) = T_{f,n,x_0}(x) + R_{f,n,x_0}(x)$.

Taylor Polynomials: Remainder

Definition: Remainder (error term) of the Taylor polynomial

$$R_{f,n,x_0}(x) = f(x) - T_{f,n,x_0}(x),$$

that is $f(x) = T_{f,n,x_0}(x) + R_{f,n,x_0}(x)$.

- $R_{f,n,x_0}(x)$ measures the deviation between the original function $f(x)$ and the n-th order Taylor polynomial $T_n(x)$.

Taylor Polynomials: Remainder

Definition: Remainder (error term) of the Taylor polynomial

$$R_{f,n,x_0}(x) = f(x) - T_{f,n,x_0}(x),$$

that is $f(x) = T_{f,n,x_0}(x) + R_{f,n,x_0}(x)$.

- $R_{f,n,x_0}(x)$ measures the deviation between the original function $f(x)$ and the n-th order Taylor polynomial $T_n(x)$.

- Estimating $R_{f,n,x_0}(x)$ helps us determine how good the Taylor polynomial approximates the original function.

Taylor Polynomials: Remainder

Definition: Remainder (error term) of the Taylor polynomial

$$R_{f,n,x_0}(x) = f(x) - T_{f,n,x_0}(x),$$

that is $f(x) = T_{f,n,x_0}(x) + R_{f,n,x_0}(x)$.

- $R_{f,n,x_0}(x)$ measures the deviation between the original function $f(x)$ and the n-th order Taylor polynomial $T_n(x)$.

- Estimating $R_{f,n,x_0}(x)$ helps us determine how good the Taylor polynomial approximates the original function.

- If f and x_0 are clear from the context, we simply write $R_n(x)$.

Taylor Polynomials: Remainder Estimates

Taylor Polynomials: Remainder Estimates

Theorem [Lagrange]

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}, \quad \text{where } \xi \text{ lies between } x_0 \text{ and } x$$

Taylor Polynomials: Remainder Estimates

Theorem [Lagrange]

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}, \quad \text{where } \xi \text{ lies between } x_0 \text{ and } x$$

Theorem [Cauchy]

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^n(x-x_0), \quad \text{where } \xi \text{ lies between } x_0 \text{ and } x.$$

Taylor Polynomials: Remainder Estimates

Theorem [Lagrange]

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}, \quad \text{where } \xi \text{ lies between } x_0 \text{ and } x$$

Theorem [Cauchy]

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^n(x-x_0), \quad \text{where } \xi \text{ lies between } x_0 \text{ and } x.$$

Theorem [Peano]

$$R_n(x) = o((x-x_0)^n).$$

Taylor Polynomials: Formal Power Series

Definition

$$a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \ldots + a_n(x - x_0)^n + \ldots$$

is called a (real) formal power series. Its set is denoted by $\mathbb{R}[[x]]$.

Definition

$$a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \ldots + a_n(x - x_0)^n + \ldots$$

is called a (real) formal power series. Its set is denoted by $\mathbb{R}[[x]]$.

Thus, for now, $f \in \mathbb{R}[[x]]$ is a formal expression, encoding a sequence of coefficients. It has no functional meaning yet.

Definition

$$a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \ldots + a_n(x - x_0)^n + \ldots$$

is called a (real) formal power series. Its set is denoted by $\mathbb{R}[[x]]$.

Thus, for now, $f \in \mathbb{R}[[x]]$ is a formal expression, encoding a sequence of coefficients. It has no functional meaning yet.

Definition

Let

$f = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \ldots + a_n(x - x_0)^n + \ldots \in \mathbb{R}[[x]]$.

The radius of convergence of f is R, where $R = \sup\{\rho\} \in \widehat{\mathbb{R}}$. Here $\{\rho\}$ is the set of non-negative ρ such that for $x_1 \in (x_0 - \rho, x_0 + \rho)$ the

$$a_0 + a_1(x_1 - x_0) + a_2(x_1 - x_0)^2 + \ldots + a_n(x_1 - x_0)^n + \ldots$$

infinite numerical sum is defined.

Taylor Polynomials: Formal Power Series: Radius of Convergence

The radius of convergence R of a formal power series $f \in \mathbb{R}[[x]]$ lies in $[0, \infty]$. We give another, more computable description.

The radius of convergence R of a formal power series $f \in \mathbb{R}[[x]]$ lies in $[0, \infty]$. We give another, more computable description.

Theorem

$$R = \frac{1}{\limsup_{n\to\infty} |c_n|^{1/n}}.$$

Here $\limsup_{n\to\infty} s_n$ denotes the supremum of the accumulation points (of the sequence s_n in $\widehat{\mathbb{R}}$). It may be $0^+ = 0$ (the $+$ indicates non-negative sequences) or ∞. Recall that $\frac{1}{\infty}$ has been defined. It is natural to adopt the convention $\frac{1}{0^+} = \infty$.

Taylor Polynomials: Formal Power Series: Radius of Convergence: Background

The theorem on the radius of convergence follows easily from results on numerical series. Let

$$\sum_{n=0}^{\infty} s_n$$

be an infinite series.

Taylor Polynomials: Formal Power Series: Radius of Convergence: Background

The theorem on the radius of convergence follows easily from results on numerical series. Let

$$\sum_{n=0}^{\infty} s_n$$

be an infinite series.

Theorem [Ratio Test]

Assume that $\lim_{n\to\infty} \frac{|s_{n+1}|}{|s_n|} = q$. Then

(i) If $|q| < 1$, then the series is absolutely convergent,

(ii) If $|q| > 1$, then the series is divergent.

The theorem on the radius of convergence follows easily from results on numerical series. Let

$$\sum_{n=0}^{\infty} s_n$$

be an infinite series.

Theorem [Ratio Test]

Assume that $\lim_{n\to\infty} \frac{|s_{n+1}|}{|s_n|} = q$. Then

(i) If $|q| < 1$, then the series is absolutely convergent,

(ii) If $|q| > 1$, then the series is divergent.

• If $q = 1$, the theorem gives no information. In fact, no general statement is possible: there are examples where the series converges and others where it diverges.

Taylor Polynomials: Formal Power Series: Radius of Convergence: Background (continued)

Taylor Polynomials: Formal Power Series: Radius of Convergence: Background (continued)

Theorem [Root Test]

Assume that $\lim_{n\to\infty} \sqrt[n]{|s_n|} = q$. Then

(i) If $|q| < 1$, then the series is absolutely convergent,

(ii) If $|q| > 1$, then the series is divergent.

Theorem [Root Test]

Assume that $\lim_{n\to\infty} \sqrt[n]{|s_n|} = q$. Then

(i) If $|q| < 1$, then the series is absolutely convergent,

(ii) If $|q| > 1$, then the series is divergent.

• If $q = 1$, the theorem gives no information. In fact, no general statement is possible: there are examples where the series converges and others where it diverges.

Let f be a function that is differentiable arbitrarily many times (we say infinitely many times).

Let f be a function that is differentiable arbitrarily many times (we say infinitely many times).

Definition

The Taylor series of f (around x_0) is

$$f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2}(x-x_0)^2 + \frac{f'''(x_0)}{3!}(x-x_0)^3 + \ldots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + \ldots \in \mathbb{R}[[x]].$$

Let f be a function that is differentiable arbitrarily many times (we say infinitely many times).

Definition

The Taylor series of f (around x_0) is

$$f(x_0)+ \quad f'(x_0)(x-x_0)+\frac{f''(x_0)}{2}(x-x_0)^2$$
$$+\quad \frac{f'''(x_0)}{3!}(x-x_0)^3+\ldots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n+\ldots\in\mathbb{R}[[x]].$$

• On the interval (x_0-R,x_0+R) the Taylor series defines a function.

Let f be a function that is differentiable arbitrarily many times (we say infinitely many times).

Definition

The Taylor series of f (around x_0) is

$$\begin{aligned} f(x_0) + {} & f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 \\ + {} & \frac{f'''(x_0)}{3!}(x - x_0)^3 + \ldots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \ldots \in \mathbb{R}[[x]]. \end{aligned}$$

- On the interval $(x_0 - R, x_0 + R)$ the Taylor series defines a function. What is the relation between this function and f?

Taylor Polynomials: Taylor Series: Example

Example

$$f(x) = \begin{cases} e^{-1/x^2}, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

Example

$$f(x) = \begin{cases} e^{-1/x^2}, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

For this function:

- $\mathrm{dom}(f) = \mathbb{R}$,
- it is infinitely differentiable on $\mathbb{R} \setminus \{0\}$,
- it is ALSO infinitely differentiable at 0,
- $0 = f(0) = f'(0) = f''(0) = f'''(0) = \ldots$.

Example

$$f(x) = \begin{cases} e^{-1/x^2}, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

For this function:

- $\mathrm{dom}(f) = \mathbb{R}$,
- it is infinitely differentiable on $\mathbb{R} \setminus \{0\}$,
- it is ALSO infinitely differentiable at 0,
- $0 = f(0) = f'(0) = f''(0) = f'''(0) = \ldots$.

Theorem

The Taylor series of f is

$$0 + 0 \cdot x + 0 \cdot x^2 + 0 \cdot x^3 + 0 \cdot x^4 + \ldots \equiv 0.$$

Example

$$f(x) = \begin{cases} e^{-1/x^2}, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

For this function:

- $\mathrm{dom}(f) = \mathbb{R}$,
- it is infinitely differentiable on $\mathbb{R} \setminus \{0\}$,
- it is ALSO infinitely differentiable at 0,
- $0 = f(0) = f'(0) = f''(0) = f'''(0) = \ldots$.

Theorem

The Taylor series of f is

$$0 + 0 \cdot x + 0 \cdot x^2 + 0 \cdot x^3 + 0 \cdot x^4 + \ldots \equiv 0.$$

This Taylor series is absolutely convergent everywhere,

Example

$$f(x) = \begin{cases} e^{-1/x^2}, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

For this function:

- $\mathrm{dom}(f) = \mathbb{R}$,
- it is infinitely differentiable on $\mathbb{R} \setminus \{0\}$,
- it is ALSO infinitely differentiable at 0,
- $0 = f(0) = f'(0) = f''(0) = f'''(0) = \ldots$.

Theorem

The Taylor series of f is

$$0 + 0 \cdot x + 0 \cdot x^2 + 0 \cdot x^3 + 0 \cdot x^4 + \ldots \equiv 0.$$

This Taylor series is absolutely convergent everywhere, but it coincides with $f(x)$ only at $x = 0$.

Definition

Let f be a real-valued function defined on an open interval. We say that f is (real) analytic if around every $x_0 \in \mathrm{dom}(f)$ it can be represented as a function defined by a power series.

Definition

Let f be a real-valued function defined on an open interval. We say that f is (real) analytic if around every $x_0 \in \mathrm{dom}(f)$ it can be represented as a function defined by a power series.

Theorem

If an analytic function f can be represented around $x_0 \in \mathrm{dom}(f)$ by a power series, then this series must be the Taylor series of f at x_0.

Analytic Functions: Examples

Analytic Functions: Examples

Example

$\frac{1}{1-x}$ is analytic on $]-1,1[$ and here

$$\frac{1}{1-x}=\sum_{n=0}^{\infty}x^n=1+x+x^2+x^3+\dots$$

Analytic Functions: Examples

Example

$\frac{1}{1-x}$ is analytic on $]-1,1[$ and here

$$\frac{1}{1-x}=\sum_{n=0}^{\infty}x^n=1+x+x^2+x^3+\dots$$

Example

e^x is analytic on $\mathbb{R}$ and here

$$e^x=\sum_{n=0}^{\infty}\frac{x^n}{n!}=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\dots$$

Analytic functions: Examples

Example

$\sin x$ is analytic on $\mathbb{R}$ and here

$$\sin(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots$$

Analytic functions: Examples

Example

$\sin x$ is analytic on $\mathbb{R}$ and here

$$\sin(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots$$

Example

$\cos x$ is analytic on $\mathbb{R}$ and here

$$\cos(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots$$

Example

$\ln x$ is analytic on $]-1,1[$ and here

$$\ln(1+x)=\sum_{n=1}^{\infty}(-1)^{n+1}\frac{x^n}{n}=x-\frac{x^2}{2}+\frac{x^3}{3}-\dots$$

Example

$\ln x$ is analytic on $]-1,1[$ and here

$$\ln(1+x)=\sum_{n=1}^{\infty}(-1)^{n+1}\frac{x^n}{n}=x-\frac{x^2}{2}+\frac{x^3}{3}-\dots$$

Example

$\frac{x}{e^x-1}$ is analytic on $\mathbb{R}$ and here

$$\frac{x}{e^x-1}=\sum_{n=0}^{\infty}B_n\frac{x^n}{n!}=B_0+B_1x+\frac{B_2}{2!}x^2+\frac{B_3}{3!}x^3+\dots,$$

where $\{B_n\}$ is the sequence of Bernoulli numbers.

Analytic functions: Examples

Analytic functions: Examples

Example

$\tan x$ is analytic on $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ and here

$$\tan(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1} 2^{2n} (2^{2n}-1) B_{2n}}{(2n)!} x^{2n-1}$$

Analytic functions: Examples

Example

$\tan x$ is analytic on $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ and here

$$\tan(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1} 2^{2n} (2^{2n}-1) B_{2n}}{(2n)!} x^{2n-1}$$

Example

$(1+x)^\alpha$ $(\alpha \in \mathbb{R})$ is analytic on $]-1,1[$ and here

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!} x^2 + \dots$$

Analytic functions: "They are nice"

Analytic functions: "They are nice"

Theorem

Let $f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots$ and $g(x) = b_0 + b_1x + b_2x^2 + b_3x^3 + \dots$ be two analytic functions ($0 \in \mathrm{dom}(f) = \mathrm{dom}(g)$). Then

(i) fg is analytic, and

$$fg(x) = a_0b_0 + (a_0b_1 + a_1b_0)x + (a_0b_2 + a_1b_1 + a_2b_0)x^2 + \dots$$

(ii) f' is analytic, and

$$f'(x) = a_1 + 2a_2x + 3a_2x^2 + 4a_3x^3 + \dots$$

That is, analytic functions can be handled formally and conveniently. In particular, they can be differentiated term by term.

Analytic functions: Turning them into pennies

$$e^{0.001} \approx 1 + 0.001,$$

Analytic functions: Turning them into pennies

$$e^{0.001} \approx 1 + 0.001,$$

$$e^{0.1} \approx 1 + 0.1 + \frac{1}{2}0.1^2,$$

Analytic functions: Turning them into pennies

$$e^{0.001} \approx 1 + 0.001,$$

$$e^{0.1} \approx 1 + 0.1 + \frac{1}{2}0.1^2,$$

$$\sin 0.001 \approx 0.001,$$

Analytic functions: Turning them into pennies

$$e^{0.001} \approx 1 + 0.001,$$

$$e^{0.1} \approx 1 + 0.1 + \frac{1}{2}0.1^2,$$

$$\sin 0.001 \approx 0.001,$$

$$\sin 0.1 \approx 0.1 - \frac{1}{6}0.1^3,$$

Analytic functions: Turning them into pennies

$$e^{0.001} \approx 1 + 0.001,$$

$$e^{0.1} \approx 1 + 0.1 + \frac{1}{2}0.1^2,$$

$$\sin 0.001 \approx 0.001,$$

$$\sin 0.1 \approx 0.1 - \frac{1}{6}0.1^3,$$

$$\tan 0.001 \approx 0.001,$$

Analytic functions: Turning them into pennies

$$e^{0.001} \approx 1 + 0.001,$$

$$e^{0.1} \approx 1 + 0.1 + \frac{1}{2}0.1^2,$$

$$\sin 0.001 \approx 0.001,$$

$$\sin 0.1 \approx 0.1 - \frac{1}{6}0.1^3,$$

$$\tan 0.001 \approx 0.001,$$

$$\sqrt{1 + 0.001} \approx 1 + \frac{1}{2}0.001,$$

Analytic functions: Turning them into pennies

$$e^{0.001} \approx 1 + 0.001,$$

$$e^{0.1} \approx 1 + 0.1 + \frac{1}{2}0.1^2,$$

$$\sin 0.001 \approx 0.001,$$

$$\sin 0.1 \approx 0.1 - \frac{1}{6}0.1^3,$$

$$\tan 0.001 \approx 0.001,$$

$$\sqrt{1 + 0.001} \approx 1 + \frac{1}{2}0.001,$$

$$\ln(1 + 0.1) \approx 0.1 - \frac{1}{2}0.1.$$

Thank you for your attention!