Numerical Analysis

Chapter 4: Approximation theory

Xiaoming He
Department of Mathematics & Statistics
Missouri University of Science & Technology

Textbook: Classical and Modern Numerical Analysis:
Theory, Methods and Practice
A. S. Ackleh, E. J. Allen, R. B. Kearfott and P. Seshaiyer
CRC Press, Boca Raton, FL, 2010

Outline

1. Introduction
2. Best approximation
3. Polynomial Approximation
4. Piecewise polynomial approximation
5. Other topics

Outline

1. Introduction
2. Best approximation
3. Polynomial Approximation
4. Piecewise polynomial approximation
5. Other topics

What's the target problem?

- Approximate a function $f(x)$ by an elementary function $p(x)$.

What's the target problem?

- Approximate a function $f(x)$ by an elementary function $p(x)$.
- For example, $p(x)$ could be a polynomial, a piecewise polynomial, a trigonometric polynomial, a rational function, or a linear combination of "nice" functions, which are easy to use in numerical computation.

What's the target problem?

- Approximate a function $f(x)$ by an elementary function $p(x)$.
- For example, $p(x)$ could be a polynomial, a piecewise polynomial, a trigonometric polynomial, a rational function, or a linear combination of "nice" functions, which are easy to use in numerical computation.
- How to decide if an approximation is "good" or "bad"?

What's the target problem?

- Approximate a function $f(x)$ by an elementary function $p(x)$.
- For example, $p(x)$ could be a polynomial, a piecewise polynomial, a trigonometric polynomial, a rational function, or a linear combination of "nice" functions, which are easy to use in numerical computation.
- How to decide if an approximation is "good" or "bad"?
- What are the features of different types of approximations?

What's the target problem?

- Approximate a function $f(x)$ by an elementary function $p(x)$.
- For example, $p(x)$ could be a polynomial, a piecewise polynomial, a trigonometric polynomial, a rational function, or a linear combination of "nice" functions, which are easy to use in numerical computation.
- How to decide if an approximation is "good" or "bad"?
- What are the features of different types of approximations?
- Accuracy
- ....

Vector space

Definition (Subspace)

W is a subspace of a real vector space V if $u \in W$, $v \in W$ implies that $\alpha u + \beta v \in W$ for all $\alpha, \beta \in \mathbb{R}$.

Definition (Span)

Let $u_1, u_2, \cdots, u_n \in V$. The set of all linear combination of $u_1, u_2, \cdots, u_n$ is called the span of $u_1, u_2, \cdots, u_n$, which is denoted by $span\{u_1, u_2, \cdots, u_n\}$

Vector space

Example

- P^1 is a subspace of P^2 and P^j ($j = 0, 1, \cdots, k$ is a subspace of P^k.

Vector space

Example

- P^1 is a subspace of P^2 and P^j ($j = 0, 1, \cdots, k$ is a subspace of P^k.
- $P^k[a, b]$ is a subspace of $C[a, b]$.

Vector space

Example

- P^1 is a subspace of P^2 and P^j ($j = 0, 1, \cdots, k$ is a subspace of P^k.
- $P^k[a, b]$ is a subspace of $C[a, b]$.
- If $W = span\{u_1, u_2, \cdots, u_n\}$ and $u_1, u_2, \cdots, u_n \in V$, then it is easy to verify that W is a subspace of V.

Vector space

Definition (Linearly independent/dependent)

Let V be a vector space. Then $u_1, u_2, \cdots, u_n \in V$ are linearly independent if $\sum_{i=1}^{n} \alpha_i u_i = 0$ implies that $\alpha_i = 0$ $(i = 1, 2, \cdots, n)$. Otherwise, $u_1, u_2, \cdots, u_n \in V$ are linearly dependent.

Vector space

Definition (Linearly independent/dependent)

Let V be a vector space. Then $u_1, u_2, \cdots, u_n \in V$ are linearly independent if $\sum_{i=1}^{n} \alpha_i u_i = 0$ implies that $\alpha_i = 0$ $(i = 1, 2, \cdots, n)$. Otherwise, $u_1, u_2, \cdots, u_n \in V$ are linearly dependent.

Example

Consider $V = C[0, 1]$. $u_1 = 1$, $u_2 = x$ and $u_3 = x^2$ are linearly independent, but $u_1 = 1$, $u_2 = x$ and $u_3 = 2 - 3x$ are linearly dependent.

Vector space

Definition (Basis)

Let V be a vector space. If $u_1, u_2, \cdots, u_n \in V$ are linearly independent and $V = span\{u_1, u_2, \cdots, u_n\}$, then $u_1, u_2, \cdots, u_n$ forms a basis for V.

Vector space

Definition (Basis)

Let V be a vector space. If $u_1, u_2, \cdots, u_n \in V$ are linearly independent and $V = span\{u_1, u_2, \cdots, u_n\}$, then $u_1, u_2, \cdots, u_n$ forms a basis for V.

Definition (Dimension)

If V has a basis of n elements, Then, n is called the dimension of V.

Vector space

Lemma (I)

If V is a n-dimensional vector space and $u_1, u_2, \cdots, u_n \in V$ are linearly independent, then $V = span\{u_1, u_2, \cdots, u_n\}$. Moreover, every basis of V has n elements and any collection of $n+1$ elements in V is linearly dependent.

Vector space

Lemma (I)

If V is a n-dimensional vector space and $u_1, u_2, \cdots, u_n \in V$ are linearly independent, then $V = span\{u_1, u_2, \cdots, u_n\}$. Moreover, every basis of V has n elements and any collection of $n+1$ elements in V is linearly dependent.

Example

$P^2 = span\{1, x, x^2\}$. Since 1, x, and x^2 are linearly independent, they form a basis of P^2 and the dimension of P^2 is 3.

Norm

Definition (Norm of a function)

Assume V is a vector space. For any function $v \in V$, a norm of v is defined to be a number $\|v\|$ satisfying

Norm

Definition (Norm of a function)

Assume V is a vector space. For any function $v \in V$, a norm of v is defined to be a number $\|v\|$ satisfying

1. $\|v\| \geq 0$.

Norm

Definition (Norm of a function)

Assume V is a vector space. For any function $v \in V$, a norm of v is defined to be a number $\|v\|$ satisfying

1. $\|v\| \geq 0$.
2. $\|v\| = 0$ if and only if $v = 0$.

Norm

Definition (Norm of a function)

Assume V is a vector space. For any function $v \in V$, a norm of v is defined to be a number $\|v\|$ satisfying

1. $\|v\| \geq 0$.
2. $\|v\| = 0$ if and only if $v = 0$.
3. $\|\lambda v\| = |\lambda| \, \|v\|$ for $\lambda \in \mathbb{R}$ or$\mathbb{C}$.

Norm

Definition (Norm of a function)

Assume V is a vector space. For any function $v \in V$, a norm of v is defined to be a number $\|v\|$ satisfying

1. $\|v\| \geq 0$.
2. $\|v\| = 0$ if and only if $v = 0$.
3. $\|\lambda v\| = |\lambda| \, \|v\|$ for $\lambda \in \mathbb{R}$ or$\mathbb{C}$.
4. $\|u + v\| \leq \|u\| + \|v\|$ for any $u, v \in V$.

Norm

Definition (Norm of a function)

Assume V is a vector space. For any function $v \in V$, a norm of v is defined to be a number $\|v\|$ satisfying

1. $\|v\| \geq 0$.
2. $\|v\| = 0$ if and only if $v = 0$.
3. $\|\lambda v\| = |\lambda| \, \|v\|$ for $\lambda \in \mathbb{R}$ or$\mathbb{C}$.
4. $\|u + v\| \leq \|u\| + \|v\|$ for any $u, v \in V$.

Definition (Normed vector space)

A vector space V is called a normed vector space if a norm $\|v\|$ defined for each $v \in V$.

Norm

Assume $V = C[a, b]$. Here are some examples of norm:

- $\|v\|_\infty = \max\limits_{x \in [a,b]} |v(x)|\,\rho(x)$ is called the max norm with weight function $\rho(x) > 0$. When $\rho(x) = 1$, it is called the max norm.

Norm

Assume $V = C[a, b]$. Here are some examples of norm:

- $\|v\|_\infty = \max_{x \in [a,b]} |v(x)| \, \rho(x)$ is called the max norm with weight function $\rho(x) > 0$. When $\rho(x) = 1$, it is called the max norm.
- $\|v\|_2 = \left(\int_a^b |v(x)|^2 \rho(x) \; dx \right)^{\frac{1}{2}}$ is called the L^2 norm with weight function $\rho(x) > 0$. When $\rho(x) = 1$, it is called the L^2 norm.

Norm

Assume $V = C[a, b]$. Here are some examples of norm:

- $\|v\|_\infty = \max\limits_{x\in[a,b]} |v(x)|\,\rho(x)$ is called the max norm with weight function $\rho(x) > 0$. When $\rho(x) = 1$, it is called the max norm.
- $\|v\|_2 = \left(\int_a^b |v(x)|^2 \rho(x)\ dx\right)^{\frac{1}{2}}$ is called the L^2 norm with weight function $\rho(x) > 0$. When $\rho(x) = 1$, it is called the L^2 norm.
- $\|v\|_1 = \int_a^b |v(x)|\,\rho(x)\ dx$ is called the L^1 norm with weight function $\rho(x) > 0$. When $\rho(x) = 1$, it is called the L^1 norm.

Outline

1. Introduction
2. Best approximation
3. Polynomial Approximation
4. Piecewise polynomial approximation
5. Other topics

Definition of a best approximation

Definition (Best approximation)

Let W be a finite-dimensional subspace of a normed vector space V. Given $v \in V$, a best approximation in W to v with respect to a norm $\|\cdot\|$ is a $w \in W$ such that the distance $\|v - w\|$ is the least among all $w \in W$. That is, $\|v - w\| \leq \|v - u\|$ for any $u \in W$.

Definition of a best approximation

Definition (Best approximation)

Let W be a finite-dimensional subspace of a normed vector space V. Given $v \in V$, a best approximation in W to v with respect to a norm $\|\cdot\|$ is a $w \in W$ such that the distance $\|v - w\|$ is the least among all $w \in W$. That is, $\|v - w\| \leq \|v - u\|$ for any $u \in W$.

Remark

- *A geometric explanation of the best approximation: see Figure 4.1 on page 193 of the textbook.*

Definition of a best approximation

Definition (Best approximation)

Let W be a finite-dimensional subspace of a normed vector space V. Given $v \in V$, a best approximation in W to v with respect to a norm $\|\cdot\|$ is a $w \in W$ such that the distance $\|v - w\|$ is the least among all $w \in W$. That is, $\|v - w\| \leq \|v - u\|$ for any $u \in W$.

Remark

- *A geometric explanation of the best approximation: see Figure 4.1 on page 193 of the textbook.*
- *Example:* $V = C[a, b]$,
  $W = \{$*set of polynomials of degree* $\leq n\}$, *and* $\|\cdot\| = \|\cdot\|_\infty$.

Definition of a best approximation

Definition (Best approximation)

Let W be a finite-dimensional subspace of a normed vector space V. Given $v \in V$, a best approximation in W to v with respect to a norm $\|\cdot\|$ is a $w \in W$ such that the distance $\|v - w\|$ is the least among all $w \in W$. That is, $\|v - w\| \leq \|v - u\|$ for any $u \in W$.

Remark

- *A geometric explanation of the best approximation: see Figure 4.1 on page 193 of the textbook.*
- *Example:* $V = C[a, b]$,
  $W = \{$*set of polynomials of degree* $\leq n\}$, *and* $\|\cdot\| = \|\cdot\|_\infty$.
- *Does such a best approximation* w *always exist?*

Definition of a best approximation

Definition (Best approximation)

Let W be a finite-dimensional subspace of a normed vector space V. Given $v \in V$, a best approximation in W to v with respect to a norm $\|\cdot\|$ is a $w \in W$ such that the distance $\|v - w\|$ is the least among all $w \in W$. That is, $\|v - w\| \leq \|v - u\|$ for any $u \in W$.

Remark

- *A geometric explanation of the best approximation: see Figure 4.1 on page 193 of the textbook.*
- *Example:* $V = C[a, b]$,
  $W = \{$*set of polynomials of degree* $\leq n\}$, *and* $\|\cdot\| = \|\cdot\|_\infty$.
- *Does such a best approximation* w *always exist?*
- *Yes!*

Existence of a best approximation

Theorem (I)

Let W be an $n+1$-dimensional subspace of a normed linear space V. Let $u_0, u_1, \cdots, u_n$ be linearly independent elements of W. That is, $W = span\{u_0, u_1, \cdots, u_n\}$. Then there is a $p \in W$, i.e., $p = \sum_{j=0}^{n} \alpha_j u_j$ for a given $f \in V$, such that

$$\|f - p\| = \left\| f - \sum_{j=0}^{n} \alpha_j u_j \right\| = \min_{\gamma_0, \gamma_1, \cdots, \gamma_n} \left\| f - \sum_{j=0}^{n} \gamma_j u_j \right\|$$

That is, $\|f - p\| \leq \|f - q\|$ for all $q \in W$. Hence p is the best approximation in W to $f \in V$ with respect to norm $\|\cdot\|$.

Existence of a best approximation

Proof.

See pages 192-194 of the textbook. (Independent study problem) □

Example of a best approximation

- Example 1: Find the best approximation $p \in P^0$ to $e^x \in C[0,1]$ for the norm infinity norm $\|\cdot\|_\infty$.

Example of a best approximation

- Example 1: Find the best approximation $p \in P^0$ to $e^x \in C[0,1]$ for the norm infinity norm $\|\cdot\|_\infty$.
- Solution: We need to find a p that minimizes $\max_{0 \le x \le 1} |e^x - p|$.

Example of a best approximation

- Example 1: Find the best approximation $p \in P^0$ to $e^x \in C[0,1]$ for the norm infinity norm $\|\cdot\|_\infty$.
- Solution: We need to find a p that minimizes $\max\limits_{0\leq x\leq 1} |e^x - p|$.

  Since $f(x) = e^x - p$ is a monotonic function of x, $\max\limits_{0\leq x\leq 1} |e^x - p|$ is the larger one of $\left|e^0 - p\right| = |1-p|$ and $|e - p|$.

Example of a best approximation

- Example 1: Find the best approximation $p \in P^0$ to $e^x \in C[0,1]$ for the norm infinity norm $\|\cdot\|_\infty$.
- Solution: We need to find a p that minimizes $\max\limits_{0 \le x \le 1} |e^x - p|$.

  Since $f(x) = e^x - p$ is a monotonic function of x, $\max\limits_{0 \le x \le 1} |e^x - p|$ is the larger one of $\left|e^0 - p\right| = |1 - p|$ and $|e - p|$.

  Given any p, $|1 - p|$ is the distance from p to 1 and $|e - p|$ is the distance from p to e.

Example of a best approximation

- Example 1: Find the best approximation $p \in P^0$ to $e^x \in C[0,1]$ for the norm infinity norm $\|\cdot\|_\infty$.
- Solution: We need to find a p that minimizes $\max_{0 \le x \le 1} |e^x - p|$.

  Since $f(x) = e^x - p$ is a monotonic function of x, $\max_{0 \le x \le 1} |e^x - p|$ is the larger one of $\left|e^0 - p\right| = |1-p|$ and $|e - p|$.

  Given any p, $|1-p|$ is the distance from p to 1 and $|e-p|$ is the distance from p to e.

  Which p can minimize the larger one of the two distances?

Example of a best approximation

- Example 1: Find the best approximation $p \in P^0$ to $e^x \in C[0,1]$ for the norm infinity norm $\|\cdot\|_\infty$.
- Solution: We need to find a p that minimizes $\max_{0\le x\le 1} |e^x - p|$.

  Since $f(x) = e^x - p$ is a monotonic function of x, $\max_{0\le x\le 1} |e^x - p|$ is the larger one of $\left|e^0 - p\right| = |1-p|$ and $|e-p|$.

  Given any p, $|1-p|$ is the distance from p to 1 and $|e-p|$ is the distance from p to e.

  Which p can minimize the larger one of the two distances?

  The middle point $p = \frac{1}{2}(e+1)$.

Example of a best approximation

- Verification:
  If $p = \frac{1}{2}(e+1)$, then $|e-p| = |1-p| = \frac{1}{2}(e-1)$;

Example of a best approximation

- Verification:
  If $p = \frac{1}{2}(e+1)$, then $|e-p| = |1-p| = \frac{1}{2}(e-1)$;
  If $p > \frac{1}{2}(e+1)$, then $|1-p| > \frac{1}{2}(e-1)$;

Example of a best approximation

- Verification:

  If $p = \frac{1}{2}(e+1)$, then $|e-p| = |1-p| = \frac{1}{2}(e-1)$;

  If $p > \frac{1}{2}(e+1)$, then $|1-p| > \frac{1}{2}(e-1)$;

  If $p < \frac{1}{2}(e+1)$, then $|e-p| > \frac{1}{2}(e-1)$.

Example of a best approximation

- Verification:
  If $p = \frac{1}{2}(e+1)$, then $|e-p| = |1-p| = \frac{1}{2}(e-1)$;
  If $p > \frac{1}{2}(e+1)$, then $|1-p| > \frac{1}{2}(e-1)$;
  If $p < \frac{1}{2}(e+1)$, then $|e-p| > \frac{1}{2}(e-1)$.

- So the minimum value of the larger one of the two distances is $\frac{1}{2}(e-1)$. So is the minimum value of $\max\limits_{0\le x\le 1} |e^x - p|$.

Example of a best approximation

- Example 2: Find the best approximation $p \in P^0$ to $e^x \in C[0,1]$ for the L^2 norm $\|\cdot\|_2$.

Example of a best approximation

- Example 2: Find the best approximation $p \in P^0$ to $e^x \in C[0,1]$ for the L^2 norm $\|\cdot\|_2$.
- Solution: We need to find a p that minimizes $\|e^x - p\|_2 = \left[\int_0^1 (e^x - p)^2 \ dx\right]^{\frac{1}{2}}$.

Example of a best approximation

- Example 2: Find the best approximation $p \in P^0$ to $e^x \in C[0,1]$ for the L^2 norm $\|\cdot\|_2$.
- Solution: We need to find a p that minimizes $\|e^x - p\|_2 = \left[\int_0^1 (e^x - p)^2 \ dx\right]^{\frac{1}{2}}$.
  Since $\|\cdot\| \geq 0$, it's equivalent to find a p that minimizes $g(p) = \|e^x - p\|_2^2 = \int_0^1 (e^x - p)^2 \ dx$.

Example of a best approximation

- Example 2: Find the best approximation $p \in P^0$ to $e^x \in C[0,1]$ for the L^2 norm $\|\cdot\|_2$.

- Solution: We need to find a p that minimizes $\|e^x - p\|_2 = \left[\int_0^1 (e^x - p)^2 \ dx\right]^{\frac{1}{2}}$.
  Since $\|\cdot\| \geq 0$, it's equivalent to find a p that minimizes $g(p) = \|e^x - p\|_2^2 = \int_0^1 (e^x - p)^2 \ dx$.
  Then

$$
\begin{aligned}
g'(p) &= \int_0^1 \frac{d\,(e^x - p)^2}{dp} \ dx \\
&= \int_0^1 -2\,(e^x - p) \ dx \\
&= -2(e - p - 1).
\end{aligned}
$$

Example of a best approximation

- Then $g'(p) = 0 \Rightarrow p = e - 1$.

Example of a best approximation

- Then $g'(p) = 0 \Rightarrow p = e - 1$.
- Since $g''(p) = -2\int_0^1 -1 \, dx = 2 > 0$, then $p = e - 1$ is the minimizer and the minimum value is $\|e^x - p\| = \frac{1}{2}(4e - e^2 - 3)^{\frac{1}{2}}$.

Example of a best approximation

- Then $g'(p) = 0 \Rightarrow p = e - 1$.
- Since $g''(p) = -2\int_0^1 -1 \, dx = 2 > 0$, then $p = e - 1$ is the minimizer and the minimum value is $\|e^x - p\| = \frac{1}{2}(4e - e^2 - 3)^{\frac{1}{2}}$.

Remark

- *Question: Is there a general framework to find the best approximation?*

Example of a best approximation

- Then $g'(p) = 0 \Rightarrow p = e - 1$.
- Since $g''(p) = -2\int_0^1 -1 \, dx = 2 > 0$, then $p = e - 1$ is the minimizer and the minimum value is $\|e^x - p\| = \frac{1}{2}(4e - e^2 - 3)^{\frac{1}{2}}$.

Remark

- *Question: Is there a general framework to find the best approximation?*
- *Yes!*

Example of a best approximation

- Then $g'(p) = 0 \Rightarrow p = e - 1$.
- Since $g''(p) = -2\int_0^1 -1\ dx = 2 > 0$, then $p = e - 1$ is the minimizer and the minimum value is $\|e^x - p\| = \frac{1}{2}(4e - e^2 - 3)^{\frac{1}{2}}$.

Remark

- *Question: Is there a general framework to find the best approximation?*
- *Yes! The best approximation in inner product spaces.*

Preparation: Inner product space

Definition (Inner product)

Let V be a vector space. An inner product (u, v) for $u, v \in V$ is a real number such that

Preparation: Inner product space

Definition (Inner product)

Let V be a vector space. An inner product (u, v) for $u, v \in V$ is a real number such that

1. $(u, u) \geq 0$ for $u \in V$.

Preparation: Inner product space

Definition (Inner product)

Let V be a vector space. An inner product (u, v) for $u, v \in V$ is a real number such that

1. $(u, u) \geq 0$ for $u \in V$.
2. $(u, u) = 0$ if and only if $u = 0$.

Preparation: Inner product space

Definition (Inner product)

Let V be a vector space. An inner product (u, v) for $u, v \in V$ is a real number such that

1. $(u, u) \geq 0$ for $u \in V$.
2. $(u, u) = 0$ if and only if $u = 0$.
3. $(u, v) = (v, u)$.

Preparation: Inner product space

Definition (Inner product)

Let V be a vector space. An inner product (u, v) for $u, v \in V$ is a real number such that

1. $(u, u) \geq 0$ for $u \in V$.
2. $(u, u) = 0$ if and only if $u = 0$.
3. $(u, v) = (v, u)$.
4. $(\alpha u + \beta v, w) = \alpha(u, w) + \beta(v, w)$ for all $u, v, w \in V$ and $\alpha, \beta \in \mathbb{R}$.

Preparation: Inner product space

Definition (Inner product)

Let V be a vector space. An inner product (u, v) for $u, v \in V$ is a real number such that

1. $(u, u) \geq 0$ for $u \in V$.
2. $(u, u) = 0$ if and only if $u = 0$.
3. $(u, v) = (v, u)$.
4. $(\alpha u + \beta v, w) = \alpha(u, w) + \beta(v, w)$ for all $u, v, w \in V$ and $\alpha, \beta \in \mathbb{R}$.

Definition (Inner product space)

A real vector space V is called a real inner product space if a inner product is defined for each $u, v \in V$.

Preparation: Inner product space

Example

- $V = C[a, b]$ with $(f, g) = \int_a^b \rho(x)f(x)g(x)\ dx$ for $f, g \in V$ where $\rho \in V$ and $\rho(x) > 0$ for $a \leq x \leq b$.

Preparation: Inner product space

Example

- $V = C[a, b]$ with $(f, g) = \int_a^b \rho(x) f(x) g(x)\, dx$ for $f, g \in V$ where $\rho \in V$ and $\rho(x) > 0$ for $a \leq x \leq b$.
- Can you see any relationship between this inner product and the L^2 norm?

Preparation: Inner product space

Example

- $V = C[a,b]$ with $(f,g) = \int_a^b \rho(x) f(x) g(x)\, dx$ for $f, g \in V$ where $\rho \in V$ and $\rho(x) > 0$ for $a \le x \le b$.
- Can you see any relationship between this inner product and the L^2 norm?
- $\|f\|_2^2 = \int_a^b |f(x)|^2 \rho(x)\, dx = (f, f)$.

Preparation: Inner product space

Example

- $V = C[a, b]$ with $(f, g) = \int_a^b \rho(x)f(x)g(x)\ dx$ for $f, g \in V$ where $\rho \in V$ and $\rho(x) > 0$ for $a \leq x \leq b$.
- Can you see any relationship between this inner product and the L^2 norm?
- $\|f\|_2^2 = \int_a^b |f(x)|^2\ \rho(x)\ dx = (f, f)$.

Theorem (II)

Any real inner product space V is a real normal linear space with norm defined by $\|v\| = (v, v)^{\frac{1}{2}}$.

Preparation: more topics

- Complex inner product spaces.
- Hilbert and Banach spaces.
- Cauchy-Schwarz inequality.
- .....

Definition of the best approximation in inner product spaces

Definition

Let $W = span\{w_1, w_2, \cdots, w_n\}$ be a finite-dimensional subspace of an inner product space V. Here $\{w_i\}_{i=1}^{n}$ is a linearly independent set. Given $v \in V$, a best approximation in W to v with respect to the norm $\|v\| = (v, v)^{\frac{1}{2}}$ is a $w \in W$ such that the distance $\|v - w\|$ is the least among all $w \in W$. That is, $\|v - w\| \leq \|v - u\|$ for any $u \in W$.

Definition of the best approximation in inner product spaces

Definition

Let $W = span\{w_1, w_2, \cdots, w_n\}$ be a finite-dimensional subspace of an inner product space V. Here $\{w_i\}_{i=1}^n$ is a linearly independent set. Given $v \in V$, a best approximation in W to v with respect to the norm $\|v\| = (v, v)^{\frac{1}{2}}$ is a $w \in W$ such that the distance $\|v - w\|$ is the least among all $w \in W$. That is, $\|v - w\| \leq \|v - u\|$ for any $u \in W$.

Remark

- *Can we use the specified W, V, and $\|v\|$ to find a general formulation of the best approximation?*

Definition of the best approximation in inner product spaces

Definition

Let $W = span\{w_1, w_2, \cdots, w_n\}$ be a finite-dimensional subspace of an inner product space V. Here $\{w_i\}_{i=1}^{n}$ is a linearly independent set. Given $v \in V$, a best approximation in W to v with respect to the norm $\|v\| = (v, v)^{\frac{1}{2}}$ is a $w \in W$ such that the distance $\|v - w\|$ is the least among all $w \in W$. That is, $\|v - w\| \leq \|v - u\|$ for any $u \in W$.

Remark

- *Can we use the specified W, V, and $\|v\|$ to find a general formulation of the best approximation?*
- *Yes!*

Definition of the best approximation in inner product spaces

Definition

Let $W = span\{w_1, w_2, \cdots, w_n\}$ be a finite-dimensional subspace of an inner product space V. Here $\{w_i\}_{i=1}^n$ is a linearly independent set. Given $v \in V$, a best approximation in W to v with respect to the norm $\|v\| = (v, v)^{\frac{1}{2}}$ is a $w \in W$ such that the distance $\|v - w\|$ is the least among all $w \in W$. That is, $\|v - w\| \leq \|v - u\|$ for any $u \in W$.

Remark

- *Can we use the specified W, V, and $\|v\|$ to find a general formulation of the best approximation?*
- *Yes!*
- *But how?*

Derivation of the best approximation in inner product spaces

- Since $W = span\{w_1, w_2, \cdots, w_n\}$, then $w = \sum_{j=1}^{n} \alpha_j w_j$ for some α_j $(j = 1, 2, \cdots, n)$.

Derivation of the best approximation in inner product spaces

- Since $W = span\{w_1, w_2, \cdots, w_n\}$, then $w = \sum_{j=1}^{n} \alpha_j w_j$ for some α_j $(j = 1, 2, \cdots, n)$.
- Question: how to find these α_j $(j = 1, 2, \cdots, n)$?

Derivation of the best approximation in inner product spaces

- By $\|v\| = (v, v)^{\frac{1}{2}}$, we have

$$
\begin{aligned}
\|w - v\|^2 &= (w - v, w - v) \\
&= \left(\sum_{j=1}^{n} \alpha_j w_j - v, \sum_{k=1}^{n} \alpha_k w_k - v \right) \\
&= \sum_{j=1}^{n} \sum_{k=1}^{n} \alpha_j \alpha_k (w_j, w_k) - \sum_{j=1}^{n} \alpha_j (v, w_j) \\
&\quad - \sum_{k=1}^{n} \alpha_k (v, w_k) + (v, v) \\
&= F(\alpha_1, \alpha_2, \cdots, \alpha_n).
\end{aligned}
$$

Derivation of the best approximation in inner product spaces

- Thus, the problem reduces to finding the minimum of F as a function of $\alpha_1, \alpha_2, \cdots, \alpha_n$.

Derivation of the best approximation in inner product spaces

- Thus, the problem reduces to finding the minimum of F as a function of $\alpha_1, \alpha_2, \cdots, \alpha_n$.
- In order to find the minimum of F, we need to compute $\frac{\partial F}{\partial \alpha_i}$.

Derivation of the best approximation in inner product spaces

- Thus, the problem reduces to finding the minimum of F as a function of $\alpha_1, \alpha_2, \cdots, \alpha_n$.
- In order to find the minimum of F, we need to compute $\frac{\partial F}{\partial \alpha_i}$.
- It is not hard to get

$$
\begin{aligned}
\frac{\partial \left(\sum_{j=1}^{n} \alpha_j (v, w_j) \right)}{\partial \alpha_i} &= (v, w_i) \\
\frac{\partial \left(\sum_{k=1}^{n} \alpha_k (v, w_k) \right)}{\partial \alpha_i} &= (v, w_i) \\
\frac{\partial (v, v)}{\partial \alpha_i} &= 0.
\end{aligned}
$$

Derivation of the best approximation in inner product spaces

- As for $\sum_{j=1}^{n}\sum_{k=1}^{n}\alpha_j\alpha_k(w_j,w_k)$, we have

$$\begin{aligned}
&\sum_{j=1}^{n}\sum_{k=1}^{n}\alpha_j\alpha_k(w_j,w_k)\\
= &\sum_{j=1,j\neq i}^{n}\ \sum_{k=1,k\neq i}^{n}\alpha_j\alpha_k(w_j,w_k)+\sum_{k=1,k\neq i}^{n}\alpha_i\alpha_k(w_i,w_k)\\
&+\sum_{j=1,j\neq i}^{n}\alpha_j\alpha_i(w_j,w_i)+\alpha_i^2(w_i,w_i).
\end{aligned}$$

Derivation of the best approximation in inner product spaces

- Hence

$$\frac{\partial\left(\sum_{j=1}^{n}\sum_{k=1}^{n}\alpha_j\alpha_k(w_j,w_k)\right)}{\partial\alpha_i}$$

$$= 0+\sum_{k=1,k\neq i}^{n}\alpha_k(w_i,w_k)+\sum_{j=1,j\neq i}^{n}\alpha_j(w_j,w_i)+2\alpha_i(w_i,w_i)$$

$$= 2\sum_{j=1}^{n}\alpha_j(w_j,w_i).$$

Derivation of the best approximation in inner product spaces

- Then

$$\frac{\partial F}{\partial \alpha_i} = 2\sum_{j=1}^{n} \alpha_j (w_j, w_i) - 2(v, w_i).$$

Derivation of the best approximation in inner product spaces

- Then

$$\frac{\partial F}{\partial \alpha_i} = 2\sum_{j=1}^{n} \alpha_j (w_j, w_i) - 2(v, w_i).$$

- Setting $\frac{\partial F}{\partial \alpha_i} = 0$, then

$$\sum_{j=1}^{n} \alpha_j (w_j, w_i) = (v, w_i), \text{ for } i = 1, 2, \cdots, n.$$

Derivation of the best approximation in inner product spaces

- Then

$$\frac{\partial F}{\partial \alpha_i} = 2\sum_{j=1}^{n} \alpha_j (w_j, w_i) - 2(v, w_i).$$

- Setting $\frac{\partial F}{\partial \alpha_i} = 0$, then

$$\sum_{j=1}^{n} \alpha_j (w_j, w_i) = (v, w_i), \text{ for } i = 1, 2, \cdots, n.$$

- Given v and w_i ($i = 1, 2, \cdots, n$), we can solve the above linear system to obtain α_j ($j = 1, 2, \cdots, n$), hence $w = \sum_{j=1}^{n} \alpha_j w_j$.

Another way to understand the best approximation in inner product spaces

Definition (Projection)

Let W be a finite-dimensional subspace of an inner product space V. An operator P that maps V into W such that $P^2 = P$ is called a projection operator form V into W.

Another way to understand the best approximation in inner product spaces

Remark

Assume $W = span\{w_1, w_2, \cdots, w_n\}$ *where* $\{w_i\}_{i=1}^n$ *is a linearly independent set. Then define* $P: V \to W$ *as* $Pv = \sum_{j=1}^n \alpha_j w_j$ *where the coefficients* α_j $(j = 1, 2, \cdots, n)$ *satisfy*

$$\sum_{j=1}^n \alpha_j (w_j, w_i) \;=\; (v, w_i), \;\; \text{for } i = 1, 2, \cdots, n.$$

Then the projection Pv *of* v *is actually the best approximation in* W *to* v *with respect to the norm* $\|v\| = (v, v)^{\frac{1}{2}}$.

Further topics of the best approximation in inner product spaces

Remark

- *How to solve the above linear system?*

Further topics of the best approximation in inner product spaces

Remark

- *How to solve the above linear system?*
- *Chapter 3!*

Further topics of the best approximation in inner product spaces

Remark

- *How to solve the above linear system?*
- *Chapter 3!*
- *Is it possible to find some special basis w_i $(i = 1, 2, \cdots, n)$ to dramatically simplify the above linear system?*

Further topics of the best approximation in inner product spaces

Remark

- *How to solve the above linear system?*
- *Chapter 3!*
- *Is it possible to find some special basis w_i $(i = 1, 2, \cdots, n)$ to dramatically simplify the above linear system?*
- *Yes! Orthogonal basis w_i $(i = 1, 2, \cdots, n)$!*

Orthogonal basis

Definition (Orthogonal/orthonormal)

Let V be an inner product space. Two vectors u and v in V are called orthogonal if $(u, v) = 0$. A set of such vectors that are pairwise orthogonal is called orthonormal if $(u, u) = 1$ for each u in that set.

Definition (Orthogonal/orthonormal basis)

Let $W = span\{w_1, w_2, \cdots, w_n\}$ be a finite-dimensional subspace of an inner product space V. If $w_1, w_2, \cdots, w_n$ are orthogonal (or orthonormal), then they form an orthogonal (or orthonormal) basis of W.

Orthogonal basis

Remark

- *If $w_1, w_2, \cdots, w_n$ are pairwise orthogonal, then*

$$(w_j, w_i) = \delta_{ji} = \begin{cases} 0, & \text{if } j \neq i, \\ 1, & \text{if } j = i. \end{cases}$$

Hence

$$\begin{aligned} & \sum_{j=1}^{n} \alpha_j (w_j, w_i) = (v, w_i), \ \ i = 1, 2, \cdots, n, \\ \Rightarrow \quad & \alpha_i = \frac{(v, w_i)}{(w_i, w_i)}, \ \ i = 1, 2, \cdots, n, \\ \Rightarrow \quad & w = Pv = \sum_{j=1}^{n} \frac{(v, w_j)}{(w_j, w_j)} w_j. \end{aligned}$$

Orthogonal basis

Remark

- *If* $w_1, w_2, \cdots, w_n$ *are orthonormal, then*

$$
\begin{aligned}
& (w_i, w_i) = 1, \;\; i = 1, 2, \cdots, n, \\
\Rightarrow \quad & \alpha_i = (v, w_i), \;\; i = 1, 2, \cdots, n, \\
\Rightarrow \quad & w = Pv = \sum_{j=1}^{n} (v, w_j) w_j.
\end{aligned}
$$

Orthogonal basis

Theorem (III)

Let $W = span\{w_1, w_2, \cdots, w_n\}$ *be a finite-dimensional subspace of an inner product space* V*. If* $w_1, w_2, \cdots, w_n$ *are orthonormal, then* $w = Pv = \sum_{j=1}^{n}(v, w_j)w_j$ *is the best approximation in* W *to* v *with respect to the norm* $\|v\| = (v, v)^{\frac{1}{2}}$*. That is,* $\|v - Pv\| \le \|v - u\|$ *for any* $u \in W$*.*

Proof.

See pages 199-201 of the textbook for the proof and the related materials, such as orthogonal complement and the orthogonal decomposition. (Independent study problem) $\square$

Orthogonal basis

Remark

- *Usually we can easily obtain a basis* $\{u_1, u_2, \cdots, u_n\}$ *which are linearly independent.*

Orthogonal basis

Remark

- *Usually we can easily obtain a basis* $\{u_1, u_2, \cdots, u_n\}$ *which are linearly independent.*
- *Question: how do we find an orthonormal (or at least orthogonal) basis* $\{w_1, w_2, \cdots, w_n\}$*?*

Orthogonal basis

Remark

- *Usually we can easily obtain a basis* $\{u_1, u_2, \cdots, u_n\}$ *which are linearly independent.*
- *Question: how do we find an orthonormal (or at least orthogonal) basis* $\{w_1, w_2, \cdots, w_n\}$*?*
- *Answer: Gram-Schmidt process!*

Orthogonal basis

Gram-Schmidt process:

1. $v_1 = u_1$

Orthogonal basis

Gram-Schmidt process:

1. $v_1 = u_1$
2. $v_j = u_j - \sum_{k=1}^{j-1} \frac{(u_j, v_k)}{(v_k, v_k)} v_k$ for $j = 2, 3, \cdots, m$.

Orthogonal basis

Gram-Schmidt process:

1. $v_1 = u_1$
2. $v_j = u_j - \sum_{k=1}^{j-1} \frac{(u_j, v_k)}{(v_k, v_k)} v_k$ for $j = 2, 3, \cdots, m$.
3. $w_j = \frac{v_j}{\|v_j\|}$ for $j = 1, 2, 3, \cdots, m$.

Orthogonal basis

Theorem (IV)

$\{v_1, v_2, \cdots, v_n\}$ is an orthogonal basis and $\{w_1, w_2, \cdots, w_n\}$ is an orthonormal basis. Moreover, $span\{u_1, u_2, \cdots, u_n\} = span\{v_1, v_2, \cdots, v_n\} = span\{w_1, w_2, \cdots, w_n\}$.

Proof.

See page 202 of the textbook. (Independent study problem) $\square$

Orthogonal basis

- Example 3: Let $V = C[-1,1]$, $W = span\{1, x, x^2\}$, $(f,g) = \int_{-1}^{1} f(x)g(x)\,dx$ for $f, g \in V$, and $\|f\| = (f,f)^{\frac{1}{2}}$. What are the orthonormal basis $\{w_1, w_2, w_3\}$ of W? Find the best approximation in W to $f(x) = e^x \in V$.

Orthogonal basis

- Example 3: Let $V = C[-1,1]$, $W = span\{1, x, x^2\}$, $(f,g) = \int_{-1}^{1} f(x)g(x)\,dx$ for $f, g \in V$, and $\|f\| = (f,f)^{\frac{1}{2}}$. What are the orthonormal basis $\{w_1, w_2, w_3\}$ of W? Find the best approximation in W to $f(x) = e^x \in V$.
- Solution:

$$v_1 = u_1 = 1,$$

Orthogonal basis

- Example 3: Let $V = C[-1,1]$, $W = span\{1, x, x^2\}$, $(f,g) = \int_{-1}^{1} f(x)g(x)\, dx$ for $f, g \in V$, and $\|f\| = (f,f)^{\frac{1}{2}}$. What are the orthonormal basis $\{w_1, w_2, w_3\}$ of W? Find the best approximation in W to $f(x) = e^x \in V$.
- Solution:

$$
\begin{aligned}
v_1 &= u_1 = 1, \\
\|v_1\|^2 &= (v_1, v_1) = \int_{-1}^{1} 1\, dx = 2,
\end{aligned}
$$

Orthogonal basis

- Example 3: Let $V = C[-1,1]$, $W = span\{1, x, x^2\}$, $(f,g) = \int_{-1}^{1} f(x)g(x)\,dx$ for $f, g \in V$, and $\|f\| = (f,f)^{\frac{1}{2}}$. What are the orthonormal basis $\{w_1, w_2, w_3\}$ of W? Find the best approximation in W to $f(x) = e^x \in V$.
- Solution:

$$
\begin{aligned}
v_1 &= u_1 = 1, \\
\|v_1\|^2 &= (v_1, v_1) = \int_{-1}^{1} 1\,dx = 2, \\
w_1 &= \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}},
\end{aligned}
$$

Orthogonal basis

- Example 3: Let $V = C[-1,1]$, $W = span\{1, x, x^2\}$, $(f,g) = \int_{-1}^{1} f(x)g(x)\,dx$ for $f, g \in V$, and $\|f\| = (f,f)^{\frac{1}{2}}$. What are the orthonormal basis $\{w_1, w_2, w_3\}$ of W? Find the best approximation in W to $f(x) = e^x \in V$.
- Solution:

$$\begin{aligned}
v_1 &= u_1 = 1, \\
\|v_1\|^2 &= (v_1, v_1) = \int_{-1}^{1} 1\,dx = 2, \\
w_1 &= \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}}, \\
v_2 &= u_2 - \frac{(u_2, v_1)}{(v_1, v_1)} v_1 = x - \frac{1}{2}\int_{-1}^{1} x\,dx = x,
\end{aligned}$$

Orthogonal basis

- Example 3: Let $V = C[-1,1]$, $W = span\{1, x, x^2\}$, $(f,g) = \int_{-1}^{1} f(x)g(x)\,dx$ for $f, g \in V$, and $\|f\| = (f,f)^{\frac{1}{2}}$. What are the orthonormal basis $\{w_1, w_2, w_3\}$ of W? Find the best approximation in W to $f(x) = e^x \in V$.
- Solution:

$$
\begin{aligned}
v_1 &= u_1 = 1, \\
\|v_1\|^2 &= (v_1, v_1) = \int_{-1}^{1} 1\,dx = 2, \\
w_1 &= \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}}, \\
v_2 &= u_2 - \frac{(u_2, v_1)}{(v_1, v_1)} v_1 = x - \frac{1}{2}\int_{-1}^{1} x\,dx = x, \\
\|v_2\|^2 &= (v_2, v_2) = \int_{-1}^{1} x^2\,dx = 2/3,
\end{aligned}
$$

Orthogonal basis

- Continued solution:

$$w_2 = \frac{v_2}{\|v_2\|} = \frac{x}{\sqrt{2/3}},$$

Orthogonal basis

- Continued solution:

$$
\begin{aligned}
w_2 &= \frac{v_2}{\|v_2\|} = \frac{x}{\sqrt{2/3}}, \\
v_3 &= u_3 - \frac{(u_3, v_1)}{(v_1, v_1)} v_1 - \frac{(u_3, v_2)}{(v_2, v_2)} v_2
\end{aligned}
$$

Orthogonal basis

- Continued solution:

$$
\begin{aligned}
w_2 &= \frac{v_2}{\|v_2\|} = \frac{x}{\sqrt{2/3}}, \\
v_3 &= u_3 - \frac{(u_3, v_1)}{(v_1, v_1)} v_1 - \frac{(u_3, v_2)}{(v_2, v_2)} v_2 \\
&= x^2 - \frac{1}{2} \int_{-1}^{1} x^2 \, dx - \frac{x}{2/3} \int_{-1}^{1} x^3 \, dx
\end{aligned}
$$

Orthogonal basis

- Continued solution:

$$
\begin{aligned}
w_2 &= \frac{v_2}{\|v_2\|} = \frac{x}{\sqrt{2/3}}, \\
v_3 &= u_3 - \frac{(u_3, v_1)}{(v_1, v_1)} v_1 - \frac{(u_3, v_2)}{(v_2, v_2)} v_2 \\
&= x^2 - \frac{1}{2} \int_{-1}^{1} x^2 \, dx - \frac{x}{2/3} \int_{-1}^{1} x^3 \, dx \\
&= x^2 - \frac{1}{2} \cdot \frac{2}{3} - \frac{x}{2/3} \cdot 0 = x^2 - \frac{1}{3}
\end{aligned}
$$

Orthogonal basis

- Continued solution:

$$
\begin{aligned}
w_2 &= \frac{v_2}{\|v_2\|} = \frac{x}{\sqrt{2/3}}, \\
v_3 &= u_3 - \frac{(u_3, v_1)}{(v_1, v_1)} v_1 - \frac{(u_3, v_2)}{(v_2, v_2)} v_2 \\
&= x^2 - \frac{1}{2} \int_{-1}^{1} x^2 \, dx - \frac{x}{2/3} \int_{-1}^{1} x^3 \, dx \\
&= x^2 - \frac{1}{2} \cdot \frac{2}{3} - \frac{x}{2/3} \cdot 0 = x^2 - \frac{1}{3} \\
\|v_3\|^2 &= (v_3, v_3) = \int_{-1}^{1} \left(x^2 - \frac{1}{3} \right)^2 dx = 8/45,
\end{aligned}
$$

Orthogonal basis

- Continued solution:

$$
\begin{aligned}
w_2 &= \frac{v_2}{\|v_2\|} = \frac{x}{\sqrt{2/3}}, \\
v_3 &= u_3 - \frac{(u_3, v_1)}{(v_1, v_1)} v_1 - \frac{(u_3, v_2)}{(v_2, v_2)} v_2 \\
&= x^2 - \frac{1}{2} \int_{-1}^{1} x^2 \; dx - \frac{x}{2/3} \int_{-1}^{1} x^3 \; dx \\
&= x^2 - \frac{1}{2} \cdot \frac{2}{3} - \frac{x}{2/3} \cdot 0 = x^2 - \frac{1}{3} \\
\|v_3\|^2 &= (v_3, v_3) = \int_{-1}^{1} \left(x^2 - \frac{1}{3} \right)^2 \; dx = 8/45, \\
w_3 &= \frac{v_3}{\|v_3\|} = \frac{x^2 - \frac{1}{3}}{\sqrt{8/45}}.
\end{aligned}
$$

Orthogonal basis

- Continued solution: Thus, $\{\frac{1}{\sqrt{2}}, \frac{x}{\sqrt{2/3}}, \frac{x^2-\frac{1}{3}}{\sqrt{8/45}}\}$ is an orthonormal basis of W.

Orthogonal basis

- Continued solution: Thus, $\{\frac{1}{\sqrt{2}}, \frac{x}{\sqrt{2/3}}, \frac{x^2-\frac{1}{3}}{\sqrt{8/45}}\}$ is an orthonormal basis of W.

  And the best approximation in W to $f(x) = e^x \in V$ is

$$
\begin{aligned}
Pf &= \sum_{j=1}^{3}(f, w_j)w_j \\
&= \frac{1}{\sqrt{2}}\int_{-1}^{1}\frac{1}{\sqrt{2}}e^x\ dx + \frac{x}{\sqrt{2/3}}\int_{-1}^{1}\frac{x}{\sqrt{2/3}}e^x\ dx \\
&\quad + \frac{x^2-\frac{1}{3}}{\sqrt{8/45}}\int_{-1}^{1}\frac{x^2-\frac{1}{3}}{\sqrt{8/45}}e^x\ dx \\
&\approx 1.1752 + 1.1036x + 0.5367x^2.
\end{aligned}
$$

Outline

1. Introduction
2. Best approximation
3. Polynomial Approximation
4. Piecewise polynomial approximation
5. Other topics

Basic idea of polynomial approximation

- A simple choice for the subspace W of the best approximation: a set of polynomials!

Basic idea of polynomial approximation

- A simple choice for the subspace W of the best approximation: a set of polynomials!
- Question: can polynomials always provide good approximation?

Basic idea of polynomial approximation

- A simple choice for the subspace W of the best approximation: a set of polynomials!
- Question: can polynomials always provide good approximation?
- Answer: Yes! But we need to be careful about the choices of polynomials.

Basic idea of polynomial approximation

Theorem (I: Weierstrass Approximation Theorem)

Given a function $f \in C[a,b]$ *and* $\varepsilon > 0$, *there exists a polynomial* $p(x)$ *such that*

$$\|p - f\|_\infty = \max_{a \le x \le b} |p(x) - f(x)| \le \varepsilon.$$

Proof.

See pages 205-209 of the textbooks for the proof and the related materials for the Bernstein polynomials. (Independent study problem) $\square$

Taylor polynomial approximation

Theorem (II: Taylor's expansion)

Suppose that $f \in C^{n+1}[a,b]$ *and* $x_0 \in [a,b]$. *Then for any* $x \in [a,b]$, *we have the following Taylor's expansion of* $f(x)$ *at* x_0:

$$f(x) = P_n(x) + R_n(x),$$

where

$$\begin{aligned} P_n(x) &= \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x_0)(x-x_0)^k \\ &= f(x_0) + f'(x_0)(x-x_0) + \frac{1}{2} f''(x_0)(x-x_0)^2 + \cdots \\ &\quad + \frac{1}{n!} f^{(n)}(x_0)(x-x_0)^n, \end{aligned}$$

Taylor polynomial approximation

Theorem (Continued)

$$R_n = \frac{1}{(n+1)!} f^{(n+1)}(\xi)(x - x_0)^{n+1}$$

for some $\xi \in [x_0, x]$ *(Lagrange form of the remainder)* ,

or

$$R_n = \frac{1}{n!} \int_{x_0}^{x} f^{(n+1)}(s)(x - s)^n \, ds$$

for some $\xi \in [x_0, x]$ *(Integral form of the remainder)* .

Lagrange polynomial approximation/interpolation

- What if we don't know $f(x)$ explicitly, but just the values y_j $(j = 0, 1, 2, \cdots, n)$ of $f(x)$ at some points x_j $(j = 0, 1, 2, \cdots, n)$ due to the measurement limitation or the information availability?

Lagrange polynomial approximation/interpolation

- What if we don't know $f(x)$ explicitly, but just the values y_j $(j = 0, 1, 2, \cdots, n)$ of $f(x)$ at some points x_j $(j = 0, 1, 2, \cdots, n)$ due to the measurement limitation or the information availability?

- Basic idea: Given $n + 1$ distinct real numbers $x_0, x_1, x_2, \cdots, x_n$ and $n + 1$ arbitrary numbers $y_0, y_1, \cdots, y_n$, find an interpolating polynomial $p(x)$ of degree at most n such that $y_j = p(x_j)$ $(j = 0, 1, 2, \cdots, n)$.

Lagrange polynomial approximation/interpolation

- What if we don't know $f(x)$ explicitly, but just the values y_j $(j = 0, 1, 2, \cdots, n)$ of $f(x)$ at some points x_j $(j = 0, 1, 2, \cdots, n)$ due to the measurement limitation or the information availability?

- Basic idea: Given $n+1$ distinct real numbers $x_0, x_1, x_2, \cdots, x_n$ and $n+1$ arbitrary numbers $y_0, y_1, \cdots, y_n$, find an interpolating polynomial $p(x)$ of degree at most n such that $y_j = p(x_j)$ $(j = 0, 1, 2, \cdots, n)$.

- Questions: Does such an interpolating polynomial $p(x)$ exist? If it does, then is it unique?

Lagrange polynomial approximation/interpolation

- What if we don't know $f(x)$ explicitly, but just the values y_j $(j = 0, 1, 2, \cdots, n)$ of $f(x)$ at some points x_j $(j = 0, 1, 2, \cdots, n)$ due to the measurement limitation or the information availability?
- Basic idea: Given $n+1$ distinct real numbers $x_0, x_1, x_2, \cdots, x_n$ and $n+1$ arbitrary numbers $y_0, y_1, \cdots, y_n$, find an interpolating polynomial $p(x)$ of degree at most n such that $y_j = p(x_j)$ $(j = 0, 1, 2, \cdots, n)$.
- Questions: Does such an interpolating polynomial $p(x)$ exist? If it does, then is it unique?
- Yes!

Lagrange polynomial approximation/interpolation

Theorem (III)

For any $n+1$ distinct real numbers $x_0, x_1, x_2, \cdots, x_n$ and for $n+1$ arbitrary numbers $y_0, y_1, y_2, \cdots, y_n$, let

$$L_k(x) = \prod_{i=0, i\neq k}^{n} \frac{x - x_i}{x_k - x_i}.$$

Then there exists a unique interpolating polynomial

$$p(x) = \sum_{k=0}^{n} y_k L_k(x)$$

of degree at most n such that $y_j = p(x_j)$ $(j = 0, 1, 2, \cdots, n)$.

Lagrange polynomial approximation/interpolation

Proof:

- First, since each $L_k(x)$ is of degree n, then $p(x)$ is of degree n.

Lagrange polynomial approximation/interpolation

Proof:

- First, since each $L_k(x)$ is of degree n, then $p(x)$ is of degree n.
- Second, it is easy to verify that

$$L_k(x_j) = \delta_{jk} = \begin{cases} 0, & \text{if } j \neq k, \\ 1, & \text{if } j = k. \end{cases}$$

Lagrange polynomial approximation/interpolation

Proof:

- First, since each $L_k(x)$ is of degree n, then $p(x)$ is of degree n.
- Second, it is easy to verify that

$$L_k(x_j) = \delta_{jk} = \begin{cases} 0, & \text{if } j \neq k, \\ 1, & \text{if } j = k. \end{cases}$$

- Then

$$p(x_j) = \sum_{k=0}^{n} y_k L_k(x_j) = \sum_{k=0}^{n} y_k \delta_{jk} = y_j.$$

Lagrange polynomial approximation/interpolation

Proof:

- First, since each $L_k(x)$ is of degree n, then $p(x)$ is of degree n.
- Second, it is easy to verify that

$$L_k(x_j) = \delta_{jk} = \begin{cases} 0, & \text{if } j \neq k, \\ 1, & \text{if } j = k. \end{cases}$$

- Then

$$p(x_j) = \sum_{k=0}^{n} y_k L_k(x_j) = \sum_{k=0}^{n} y_k \delta_{jk} = y_j.$$

- Hence we complete the proof of the existence of $p(x)$.

Lagrange polynomial approximation/interpolation

Continued proof:

- Now let's turn to the uniqueness of $p(x)$.

Lagrange polynomial approximation/interpolation

Continued proof:

- Now let's turn to the uniqueness of $p(x)$.
- Assume there are two such polynomials $p(x)$ and $q(x)$ of degree at most n that their values at x_j $(j = 0, 1, 2, \cdots, n)$ are y_j $(j = 0, 1, 2, \cdots, n)$. That is,

$$y_j = p(x_j) = q(x_j), \; j = 0, 1, 2, \cdots, n.$$

Lagrange polynomial approximation/interpolation

Continued proof:

- Now let's turn to the uniqueness of $p(x)$.
- Assume there are two such polynomials $p(x)$ and $q(x)$ of degree at most n that their values at x_j $(j = 0, 1, 2, \cdots, n)$ are y_j $(j = 0, 1, 2, \cdots, n)$. That is,

$$y_j = p(x_j) = q(x_j), \; j = 0, 1, 2, \cdots, n.$$

- Since $p(x)$ and $q(x)$ are polynomials of degree at most n, then we may let

$$\begin{aligned} p(x) &= \sum_{k=0}^{n} \alpha_k x^k, \\ q(x) &= \sum_{k=0}^{n} \beta_k x^k \end{aligned}$$

for some constants α_k and β_k $(k = 0, 1, 2, \cdots, n)$.

Lagrange polynomial approximation/interpolation

Continued proof:

- Let $\gamma_k = \alpha_k - \beta_k$. Then

$$p(x) - q(x) = \sum_{k=0}^{n} (\alpha_k - \beta_k) x^k = \sum_{k=0}^{n} \gamma_k x^k.$$

Lagrange polynomial approximation/interpolation

Continued proof:

- Let $\gamma_k = \alpha_k - \beta_k$. Then

$$p(x) - q(x) = \sum_{k=0}^{n} (\alpha_k - \beta_k) x^k = \sum_{k=0}^{n} \gamma_k x^k.$$

- Since $y_j = p(x_j) = q(x_j)$ $(j = 0, 1, 2, \cdots, n)$, then $p(x_j) - q(x_j) = 0$ $(j = 0, 1, 2, \cdots, n)$.

Lagrange polynomial approximation/interpolation

Continued proof:

- Let $\gamma_k = \alpha_k - \beta_k$. Then

$$p(x) - q(x) = \sum_{k=0}^{n} (\alpha_k - \beta_k) x^k = \sum_{k=0}^{n} \gamma_k x^k.$$

- Since $y_j = p(x_j) = q(x_j)$ $(j = 0, 1, 2, \cdots, n)$, then $p(x_j) - q(x_j) = 0$ $(j = 0, 1, 2, \cdots, n)$.
- Hence

$$\sum_{k=0}^{n} \gamma_k x_j^k = 0, \; j = 0, 1, 2, \cdots, n.$$

Lagrange polynomial approximation/interpolation

Continued proof:

- The above linear system can be rewritten as

$$A\begin{pmatrix}\gamma_0\\ \gamma_1\\ \vdots\\ \gamma_n\end{pmatrix}=\begin{pmatrix}1 & x_0 & x_0^2 & x_0^3 & \cdots & x_0^n\\ 1 & x_1 & x_1^2 & x_1^3 & \cdots & x_1^n\\ & \vdots & & \ddots & & \\ 1 & x_n & x_n^2 & x_n^3 & \cdots & x_n^n\end{pmatrix}\begin{pmatrix}\gamma_0\\ \gamma_1\\ \vdots\\ \gamma_n\end{pmatrix}=\begin{pmatrix}0\\ 0\\ \vdots\\ 0\end{pmatrix}.$$

Lagrange polynomial approximation/interpolation

Continued proof:

- The above linear system can be rewritten as

$$A\begin{pmatrix}\gamma_0\\ \gamma_1\\ \vdots\\ \gamma_n\end{pmatrix}=\begin{pmatrix}1 & x_0 & x_0^2 & x_0^3 & \cdots & x_0^n\\ 1 & x_1 & x_1^2 & x_1^3 & \cdots & x_1^n\\ & \vdots & & \ddots & & \\ 1 & x_n & x_n^2 & x_n^3 & \cdots & x_n^n\end{pmatrix}\begin{pmatrix}\gamma_0\\ \gamma_1\\ \vdots\\ \gamma_n\end{pmatrix}=\begin{pmatrix}0\\ 0\\ \vdots\\ 0\end{pmatrix}.$$

- The above matrix A is a Vandermonde matrix and the determinant of the matrxi A is

$$det(A)=\prod_{k=1}^{n}\prod_{j=0}^{k-1}(x_k-x_j).$$

Lagrange polynomial approximation/interpolation

Continued proof:

- The above linear system can be rewritten as

$$A\begin{pmatrix}\gamma_0\\ \gamma_1\\ \vdots\\ \gamma_n\end{pmatrix}=\begin{pmatrix}1 & x_0 & x_0^2 & x_0^3 & \cdots & x_0^n\\ 1 & x_1 & x_1^2 & x_1^3 & \cdots & x_1^n\\ & \vdots & & \ddots & & \\ 1 & x_n & x_n^2 & x_n^3 & \cdots & x_n^n\end{pmatrix}\begin{pmatrix}\gamma_0\\ \gamma_1\\ \vdots\\ \gamma_n\end{pmatrix}=\begin{pmatrix}0\\ 0\\ \vdots\\ 0\end{pmatrix}.$$

- The above matrix A is a Vandermonde matrix and the determinant of the matrxi A is

$$det(A)=\prod_{k=1}^{n}\prod_{j=0}^{k-1}(x_k-x_j).$$

- Since $x_0, x_1, x_2, \cdots, x_n$ are distinct real numbers, then $det(A)\neq 0$.

Lagrange polynomial approximation/interpolation

Continued proof: Recall

Theorem

Let A be an $n \times n$ matrix. Then the following are equivalent:

- *A is nonsingular/invertible.*
- *$det(A) \neq 0$.*
- *$\overrightarrow{x} = 0$ is the unique solution of $A\overrightarrow{x} = 0$.*
- *$A\overrightarrow{x} = \overrightarrow{b}$ has a unique solution.*
- *The columns (and rows) of A are linearly independent.*

Lagrange polynomial approximation/interpolation

Continued proof:

- Hence the above linear system has a unique solution $\gamma_k = 0$ $(k = 1, 2, \cdots, n)$, which implies $\alpha_k = \beta_k$ $(k = 0, 1, 2, \cdots, n)$.

Lagrange polynomial approximation/interpolation

Continued proof:

- Hence the above linear system has a unique solution $\gamma_k = 0$ $(k = 1, 2, \cdots, n)$, which implies $\alpha_k = \beta_k$ $(k = 0, 1, 2, \cdots, n)$.
- Then $p(x) = q(x)$. This completes the proof.

Lagrange polynomial approximation/interpolation

Continued proof:

- Hence the above linear system has a unique solution $\gamma_k = 0$ $(k = 1, 2, \cdots, n)$, which implies $\alpha_k = \beta_k$ $(k = 0, 1, 2, \cdots, n)$.
- Then $p(x) = q(x)$. This completes the proof.

Remark

In fact, the above theory from linear algebra can be used to prove the existence and uniqueness at the same time. (Homework)

Lagrange polynomial approximation/interpolation

Definition (Lagrange basis)

The set of functions $\{L_k(x)\}_{k=0}^{n}$ is called the Lagrange basis for the space of polynomials of degree n associated with the set of points $\{x_k\}_{k=0}^{n}$. And L_k $(k = 0, 1, 2, \cdots, n)$ are called the Lagrange basis functions.

Lagrange polynomial approximation/interpolation

Definition (Lagrange basis)

The set of functions $\{L_k(x)\}_{k=0}^n$ is called the Lagrange basis for the space of polynomials of degree n associated with the set of points $\{x_k\}_{k=0}^n$. And L_k $(k = 0, 1, 2, \cdots, n)$ are called the Lagrange basis functions.

Definition (Lagrange form)

$p(x) = \sum_{k=0}^{n} y_k L_k(x)$ is called the Lagrange form of the interpolating polynomial $p(x)$, which satisfies $y_k = p(x_k)$ $(j = 0, 1, 2, \cdots, n)$.

Lagrange polynomial approximation/interpolation

Remark

What if pick up $y_k = f(x_k)$ $(k = 0, 1, 2, \cdots, n)$ *for a given function* $f(x)$*?*

Lagrange polynomial approximation/interpolation

Remark

What if pick up $y_k = f(x_k)$ $(k = 0, 1, 2, \cdots, n)$ *for a given function* $f(x)$*?*

Definition (Lagrange polynomial approximation)

$p(x) = \sum_{k=0}^{n} f(x_k) L_k(x)$ is called the Lagrange polynomial approximation/interpolation of a given function $f(x)$.

Lagrange polynomial approximation/interpolation

The error of the Lagrange polynomial approximation:

Theorem (IV)

If $x_0, x_1, x_2, \cdots, x_n$ are $n+1$ distinct points in $[a, b]$ and $f \in C^{n+1}[a, b]$, then for each $x \in [a, b]$, there exists a number $\xi = \xi(x) \in (a, b)$ such that

$$f(x) - p(x) = R(x) = \frac{f^{(n+1)}(\xi(x))W(x)}{(n+1)!}$$

where $W(x) = \prod_{i=0}^{n}(x - x_i)$ and $p(x)$ is the Lagrange polynomial approximation/interpolation.

Proof.

See pages 215-216 of the textbook. (Independent study problem) $\square$

Lagrange polynomial approximation/interpolation

Remark

- *What happen if f is a n-th order polynomial?*

Lagrange polynomial approximation/interpolation

Remark

- *What happen if f is a n-th order polynomial?*
- $f^{(n+1)}(\xi(x)) = 0$ *since* $f^{(n+1)} \equiv 0$.

Lagrange polynomial approximation/interpolation

Remark

- *What happen if f is a n-th order polynomial?*
- $f^{(n+1)}(\xi(x)) = 0$ *since* $f^{(n+1)} \equiv 0$.
- *The error* $f(x) - p(x) = 0$.

Lagrange polynomial approximation/interpolation

Remark

- *What happen if f is a n-th order polynomial?*
- $f^{(n+1)}(\xi(x)) = 0$ *since* $f^{(n+1)} \equiv 0$.
- *The error* $f(x) - p(x) = 0$.
- *Hence the n-th order Lagrange polynomial approximation is exact for n-th order polynomial.*

Lagrange polynomial approximation/interpolation

Remark

- *Define* $\|f\|_\infty = \max_{a \le x \le b} |f(x)|$. *Then the error bound is*

$$\begin{aligned}
\|f - p\|_\infty &= \max_{a \le x \le b} |f(x) - p(x)| \\
&\le \frac{1}{(n+1)!} \left\| f^{(n+1)} \right\|_\infty \|W\|_\infty \\
&\le \frac{1}{(n+1)!} \left\| f^{(n+1)} \right\|_\infty (b-a)^{n+1}.
\end{aligned}$$

Lagrange polynomial approximation/interpolation

Remark

- *Define* $\|f\|_\infty = \max_{a \le x \le b} |f(x)|$. *Then the error bound is*

$$\begin{aligned} \|f - p\|_\infty &= \max_{a \le x \le b} |f(x) - p(x)| \\ &\le \frac{1}{(n+1)!} \left\| f^{(n+1)} \right\|_\infty \|W\|_\infty \\ &\le \frac{1}{(n+1)!} \left\| f^{(n+1)} \right\|_\infty (b-a)^{n+1}. \end{aligned}$$

- *In fact, we have*

$$\|f - p\|_\infty \le C h^{n+1} \left\| f^{(n+1)} \right\|_\infty .$$

where $h = \max_{0 \le j \le n-1} (x_{j+1} - x_j)$.

Lagrange polynomial approximation/interpolation

Remark

- *The error bound depends on the nodes* $\{x_i\}_{i=0}^n$ *since* $W(x) = \prod_{i=0}^n (x - x_i)$.

Lagrange polynomial approximation/interpolation

Remark

- *The error bound depends on the nodes* $\{x_i\}_{i=0}^n$ *since* $W(x) = \prod_{i=0}^n (x - x_i)$.
- *Can we choose* $\{x_i\}_{i=0}^n$ *suitably to minimize the error bound?*

Lagrange polynomial approximation/interpolation

Remark

- *The error bound depends on the nodes* $\{x_i\}_{i=0}^n$ *since* $W(x) = \prod_{i=0}^n (x - x_i)$.
- *Can we choose* $\{x_i\}_{i=0}^n$ *suitably to minimize the error bound?*
- *Yes! Chebyshev Polynomials!*

Chebyshev Polynomials approximation/interpolation

Theorem (V)

The uniform norm of $W(x) = \prod_{i=0}^{n}(x - x_i)$ is minimized on $[a, b]$ when

$$x_i = \frac{1}{2}\left[(b-a)cos\left(\frac{2i+1}{n+1}\cdot\frac{\pi}{2}\right) + a + b\right], \; i = 0, 1, 2, \cdots, n,$$

and the minimum value of the norm is

$$\|W\|_\infty = \frac{1}{2^{2n+1}}(b-a)^{n+1}.$$

Proof.

See pages 217-219 of the textbook for the proof and the related materials on Chebyshev polynomials. (Independent study problem) □

Chebyshev polynomials approximation/interpolation

Definition

If the Lagrange polynomial approximation/interpolation of a function $f(x)$ on $[a, b]$ uses the following roots of Chebyshev polynomials

$$x_i = \frac{1}{2}\left[(b-a)cos\left(\frac{2i+1}{n+1}\cdot\frac{\pi}{2}\right)+a+b\right],\ i = 0, 1, 2, \cdots, n,$$

then it is called the Chebyshev polynomials approximation/interpolation of the function $f(x)$ on $[a, b]$.

Chebyshev polynomials approximation/interpolation

Remark

Compared with the error bound of the regular Lagrange polynomial approximation

$$\|f-p\|_\infty \le \frac{1}{(n+1)!}\left\|f^{(n+1)}\right\|_\infty (b-a)^{n+1},$$

the error bound of the Chebyshev polynomial approximation is

$$\|f-p\|_\infty \le \frac{1}{(n+1)!}\left\|f^{(n+1)}\right\|_\infty (b-a)^{n+1}\frac{1}{2^{2n+1}}.$$

Hermit polynomial approximation/interpolation

Definition

For any n distinct real numbers $x_1, x_2, \cdots, x_n$, the Hermit polynomial approximation/interpolation of a given function $f(x)$ is a polynomial $p(x)$ of degree at most $2n-1$ such that

$$p(x_i) = f(x_i), \quad p'(x_i) = f'(x_i), \; i = 1, 2, \cdots, n.$$

Hermit polynomials approximation/interpolation

Theorem (VI)

If $x_1, x_2, \cdots, x_n$ are $n+1$ distinct real numbers in $[a,b]$ and $f \in C^1[a,b]$. Then there exists a unique Hermit polynomial approximation/interpolation $H_n(x)$ of $f(x)$. And it is given by

$$H_n(x) = \sum_{k=1}^{n} f(x_k) h_k(x) + \sum_{k=1}^{n} f'(x_k) \tilde{h}_k(x),$$

where

$$\begin{aligned} h_k(x) &= \left[1 - 2L_k'(x_k)(x - x_k)\right] (L_k(x))^2 , \\ \tilde{h}_k(x) &= (x - x_k) (L_k(x))^2 . \end{aligned}$$

Hermit polynomials approximation/interpolation

Theorem (VI: Continued)

Moreover, if $f \in C^{2n}[a,b]$, then there exists a $\xi = \xi(x) \in [a,b]$ such that

$$f(x) - H_n(x) = \frac{\left(\prod_{k=1}^{n}(x - x_k)\right)^2}{(2n)!} f^{2n}(\xi),$$

Proof.

See page 221 of the textbook. (Independent study problem) □

Some thoughts for the polynomial approximation

Remark

- *In order to obtain a desired accuracy from the polynomial approximation, enough number of interpolating nodes* $\{x_i\}_{i=0}^{n}$ *need to be selected.*

Some thoughts for the polynomial approximation

Remark

- *In order to obtain a desired accuracy from the polynomial approximation, enough number of interpolating nodes* $\{x_i\}_{i=0}^n$ *need to be selected.*
- *When* n *is large, the corresponding polynomial has very hight order.*

Some thoughts for the polynomial approximation

Remark

- *In order to obtain a desired accuracy from the polynomial approximation, enough number of interpolating nodes* $\{x_i\}_{i=0}^{n}$ *need to be selected.*
- *When* n *is large, the corresponding polynomial has very hight order.*
- *Can we use lower order polynomial to achieve high accuracy?*

Some thoughts for the polynomial approximation

Remark

- *In order to obtain a desired accuracy from the polynomial approximation, enough number of interpolating nodes* $\{x_i\}_{i=0}^{n}$ *need to be selected.*
- *When n is large, the corresponding polynomial has very hight order.*
- *Can we use lower order polynomial to achieve high accuracy?*
- *Yes! Piecewise polynomials!*

Outline

1. Introduction
2. Best approximation
3. Polynomial Approximation
4. Piecewise polynomial approximation
5. Other topics

Basic idea of piecewise polynomial approximation

- Divide the interval $[a, b]$ into many sub-intervals.

Basic idea of piecewise polynomial approximation

- Divide the interval $[a, b]$ into many sub-intervals.

- On each sub-interval, a lower order polynomial is used to approximate the given function $f(x)$. Then we assemble all of the pieces on all of the sub-intervals together to obtain the piecewise polynomial approximation of $f(x)$ on $[a, b]$.

Basic idea of piecewise polynomial approximation

- Divide the interval $[a, b]$ into many sub-intervals.
- On each sub-interval, a lower order polynomial is used to approximate the given function $f(x)$. Then we assemble all of the pieces on all of the sub-intervals together to obtain the piecewise polynomial approximation of $f(x)$ on $[a, b]$.
- Continuous piecewise polynomial approximation requires that the piecewise polynomial approximation to be continuous. That is, the polynomial on each sub-interval must match the polynomials on the neighboring sub-intervals.

Basic idea of piecewise polynomial approximation

- Divide the interval $[a, b]$ into many sub-intervals.
- On each sub-interval, a lower order polynomial is used to approximate the given function $f(x)$. Then we assemble all of the pieces on all of the sub-intervals together to obtain the piecewise polynomial approximation of $f(x)$ on $[a, b]$.
- Continuous piecewise polynomial approximation requires that the piecewise polynomial approximation to be continuous. That is, the polynomial on each sub-interval must match the polynomials on the neighboring sub-intervals.
- Many numerical methods use piecewise polynomial approximation, such as the finite element method.

Continuous piecewise linear approximation

Definition (continuous piecewise linear polynomial)

Given a partition

$$\Delta : a = x_0 < x_1 < x_2 \cdots < x_{n-1} < x_N = b$$

of $[a, b]$, the set L_Δ of all continuous piecewise linear polynomials on $[a, b]$ with respect to Δ is

$$\begin{aligned} L_\Delta &= \{\varphi \in C[a, b] : \ \varphi(x) \text{ is linear on each } [x_i, x_{i+1}] \\ & \quad (i = 0, 1, 2, \cdots, N-1) \text{ of } \Delta\}. \end{aligned}$$

Continuous piecewise linear approximation

Theorem (I)

L_Δ is an $(N+1)-$dimensional subspace of $C[a, b]$.

Continuous piecewise linear approximation

Theorem (I)

L_Δ is an $(N+1)-$dimensional subspace of $C[a,b]$.

Proof:

- First, it is easy to verify that L_Δ is a subspace of $C[a,b]$.

Continuous piecewise linear approximation

Theorem (I)

L_Δ is an $(N+1)-$dimensional subspace of $C[a,b]$.

Proof:

- First, it is easy to verify that L_Δ is a subspace of $C[a,b]$.
- If we can find a continuous piecewise linear basis of $N+1$ functions for L_Δ, then the proof is completed.

Continuous piecewise linear approximation

Theorem (I)

L_Δ is an $(N+1)-$dimensional subspace of $C[a,b]$.

Proof:

- First, it is easy to verify that L_Δ is a subspace of $C[a,b]$.
- If we can find a continuous piecewise linear basis of $N+1$ functions for L_Δ, then the proof is completed.
- Consider $\varphi_i(x) \in L_\Delta$ $(i = 0, \cdots, N)$ such that

$$\varphi_i(x_j) = \delta_{ij} = \begin{cases} 0, & \text{if } j \neq i, \\ 1, & \text{if } j = i. \end{cases}$$

Continuous piecewise linear approximation

Theorem (I)

L_Δ is an $(N+1)-$dimensional subspace of $C[a,b]$.

Proof:

- First, it is easy to verify that L_Δ is a subspace of $C[a,b]$.
- If we can find a continuous piecewise linear basis of $N+1$ functions for L_Δ, then the proof is completed.
- Consider $\varphi_i(x) \in L_\Delta$ $(i = 0, \cdots, N)$ such that

$$\varphi_i(x_j) = \delta_{ij} = \begin{cases} 0, & \text{if } j \neq i, \\ 1, & \text{if } j = i. \end{cases}$$

- Linear independence: consider $\sum_{i=0}^{N} c_i \varphi_i(x) = 0$ for any $x \in [a,b]$.

Continuous piecewise linear approximation

Theorem (I)

L_Δ is an $(N+1)-$dimensional subspace of $C[a,b]$.

Proof:

- First, it is easy to verify that L_Δ is a subspace of $C[a,b]$.
- If we can find a continuous piecewise linear basis of $N+1$ functions for L_Δ, then the proof is completed.
- Consider $\varphi_i(x) \in L_\Delta$ $(i = 0, \cdots, N)$ such that

$$\varphi_i(x_j) = \delta_{ij} = \begin{cases} 0, & \text{if } j \neq i, \\ 1, & \text{if } j = i. \end{cases}$$

- Linear independence: consider $\sum_{i=0}^{N} c_i \varphi_i(x) = 0$ for any $x \in [a,b]$.
- Let $x = x_j$, then $c_j = 0$ $(j = 0, \cdots, N)$.

Continuous piecewise linear approximation

Theorem (I)

L_Δ is an $(N+1)-$dimensional subspace of $C[a,b]$.

Proof:

- First, it is easy to verify that L_Δ is a subspace of $C[a,b]$.
- If we can find a continuous piecewise linear basis of $N+1$ functions for L_Δ, then the proof is completed.
- Consider $\varphi_i(x) \in L_\Delta$ $(i = 0, \cdots, N)$ such that

$$\varphi_i(x_j) = \delta_{ij} = \begin{cases} 0, & \text{if } j \neq i, \\ 1, & \text{if } j = i. \end{cases}$$

- Linear independence: consider $\sum_{i=0}^{N} c_i \varphi_i(x) = 0$ for any $x \in [a,b]$.
- Let $x = x_j$, then $c_j = 0$ $(j = 0, \cdots, N)$.
- So $\varphi_i(x)$ $(i = 0, \cdots, N)$ are linearly independent.

Continuous piecewise linear approximation

Continued proof:

- Span: Given any $f \in L_\Delta$, let $c_i = f(x_i)$ and consider $g(x) = \sum_{i=0}^{N} c_i \varphi_i(x)$.

Continuous piecewise linear approximation

Continued proof:

- Span: Given any $f \in L_\Delta$, let $c_i = f(x_i)$ and consider $g(x) = \sum_{i=0}^{N} c_i \varphi_i(x)$.
- First, $g(x_j) = c_j = f(x_j)$ $(j = 0, \cdots, N)$.

Continuous piecewise linear approximation

Continued proof:

- Span: Given any $f \in L_\Delta$, let $c_i = f(x_i)$ and consider $g(x) = \sum_{i=0}^{N} c_i \varphi_i(x)$.
- First, $g(x_j) = c_j = f(x_j)$ $(j = 0, \cdots, N)$.
- Second, both $f(x)$ and $g(x)$ are linear in each piece $[x_j, x_{j+1}]$ $(j = 0, \cdots, N-1)$.

Continuous piecewise linear approximation

Continued proof:

- Span: Given any $f \in L_\Delta$, let $c_i = f(x_i)$ and consider $g(x) = \sum_{i=0}^{N} c_i \varphi_i(x)$.
- First, $g(x_j) = c_j = f(x_j)$ $(j = 0, \cdots, N)$.
- Second, both $f(x)$ and $g(x)$ are linear in each piece $[x_j, x_{j+1}]$ $(j = 0, \cdots, N-1)$.
- Hence $f(x) = g(x)$ in each piece $[x_j, x_{j+1}]$ $(j = 0, \cdots, N-1)$.

Continuous piecewise linear approximation

Continued proof:

- Span: Given any $f \in L_\Delta$, let $c_i = f(x_i)$ and consider $g(x) = \sum_{i=0}^{N} c_i \varphi_i(x)$.
- First, $g(x_j) = c_j = f(x_j)$ $(j = 0, \cdots, N)$.
- Second, both $f(x)$ and $g(x)$ are linear in each piece $[x_j, x_{j+1}]$ $(j = 0, \cdots, N-1)$.
- Hence $f(x) = g(x)$ in each piece $[x_j, x_{j+1}]$ $(j = 0, \cdots, N-1)$.
- Then $f(x) = g(x) = \sum_{i=0}^{N} c_i \varphi_i(x)$.

Continuous piecewise linear approximation

Continued proof:

- Span: Given any $f \in L_\Delta$, let $c_i = f(x_i)$ and consider $g(x) = \sum_{i=0}^{N} c_i \varphi_i(x)$.
- First, $g(x_j) = c_j = f(x_j)$ $(j = 0, \cdots, N)$.
- Second, both $f(x)$ and $g(x)$ are linear in each piece $[x_j, x_{j+1}]$ $(j = 0, \cdots, N-1)$.
- Hence $f(x) = g(x)$ in each piece $[x_j, x_{j+1}]$ $(j = 0, \cdots, N-1)$.
- Then $f(x) = g(x) = \sum_{i=0}^{N} c_i \varphi_i(x)$.
- This implies $L_\Delta = span\{\varphi_0(x), \cdots, \varphi_N(x)\}$.

Continuous piecewise linear approximation

Continued proof:

- Span: Given any $f \in L_\Delta$, let $c_i = f(x_i)$ and consider $g(x) = \sum_{i=0}^{N} c_i \varphi_i(x)$.
- First, $g(x_j) = c_j = f(x_j)$ $(j = 0, \cdots, N)$.
- Second, both $f(x)$ and $g(x)$ are linear in each piece $[x_j, x_{j+1}]$ $(j = 0, \cdots, N-1)$.
- Hence $f(x) = g(x)$ in each piece $[x_j, x_{j+1}]$ $(j = 0, \cdots, N-1)$.
- Then $f(x) = g(x) = \sum_{i=0}^{N} c_i \varphi_i(x)$.
- This implies $L_\Delta = span\{\varphi_0(x), \cdots, \varphi_N(x)\}$.
- Therefore $\varphi_i(x)$ $(i = 0, \cdots, N)$ form a basis of L_Δ

Continuous piecewise linear approximation

Continued proof:

- Span: Given any $f \in L_\Delta$, let $c_i = f(x_i)$ and consider $g(x) = \sum_{i=0}^{N} c_i \varphi_i(x)$.
- First, $g(x_j) = c_j = f(x_j)$ $(j = 0, \cdots, N)$.
- Second, both $f(x)$ and $g(x)$ are linear in each piece $[x_j, x_{j+1}]$ $(j = 0, \cdots, N-1)$.
- Hence $f(x) = g(x)$ in each piece $[x_j, x_{j+1}]$ $(j = 0, \cdots, N-1)$.
- Then $f(x) = g(x) = \sum_{i=0}^{N} c_i \varphi_i(x)$.
- This implies $L_\Delta = span\{\varphi_0(x), \cdots, \varphi_N(x)\}$.
- Therefore $\varphi_i(x)$ $(i = 0, \cdots, N)$ form a basis of L_Δ if they exist.

Continuous piecewise linear approximation

Continued proof:

- The existence of $\varphi_i(x)$ $(i = 0, \cdots, N)$:

$$\varphi_0(x) = \begin{cases} \frac{x_1 - x}{x_1 - x_0}, & \text{if } x_0 \le x \le x_1, \\ 0, & \text{otherwise}, \end{cases}$$

$$\varphi_N(x) = \begin{cases} \frac{x - x_{N-1}}{x_N - x_{N-1}}, & \text{if } x_{N-1} \le x \le x_N, \\ 0, & \text{otherwise}, \end{cases}$$

$$\varphi_i(x) = \begin{cases} \frac{x - x_{i-1}}{x_i - x_{i-1}}, & \text{if } x_{i-1} \le x \le x_i, \\ \frac{x_{i+1} - x}{x_{i+1} - x_i}, & \text{if } x_i \le x \le x_{i+1}, \\ 0, & \text{otherwise}. \end{cases}$$

$$(i = 1, \cdots, N-1)$$

Continuous piecewise linear approximation

Continued proof:

- The existence of $\varphi_i(x)$ $(i = 0, \cdots, N)$:

$$\varphi_0(x) = \begin{cases} \frac{x_1 - x}{x_1 - x_0}, & if\ x_0 \le x \le x_1, \\ 0, & otherwise, \end{cases}$$

$$\varphi_N(x) = \begin{cases} \frac{x - x_{N-1}}{x_N - x_{N-1}}, & if\ x_{N-1} \le x \le x_N, \\ 0, & otherwise, \end{cases}$$

$$\varphi_i(x) = \begin{cases} \frac{x - x_{i-1}}{x_i - x_{i-1}}, & if\ x_{i-1} \le x \le x_i, \\ \frac{x_{i+1} - x}{x_{i+1} - x_i}, & if\ x_i \le x \le x_{i+1}, \\ 0, & otherwise. \end{cases}$$

$$(i = 1, \cdots, N-1)$$

- A geometric illustration of $\varphi_i(x)$ $(i = 0, \cdots, N)$: see Figure 4.8 on page 240 of the textbook.

Continuous piecewise linear approximation

Remark

- *The continuous piecewise linear basis functions* $\varphi_i(x)$ $(i = 0, 1, 2, \cdots, N)$ *are the well-known hat functions.*

Continuous piecewise linear approximation

Remark

- *The continuous piecewise linear basis functions* $\varphi_i(x)$ $(i = 0, 1, 2, \cdots, N)$ *are the well-known hat functions.*
- *They are actually the linear finite element basis functions for the finite element method in 1D!*

Continuous piecewise linear approximation

Theorem (II)

Given an $f \in C[a, b]$, there is a unique $\Phi \in L_\Delta$ which satisfies $\Phi(x_i) = f(x_i)$ $(i = 0, 1, 2, \cdots, N)$.

Continuous piecewise linear approximation

Theorem (II)

Given an $f \in C[a,b]$, there is a unique $\Phi \in L_\Delta$ which satisfies $\Phi(x_i) = f(x_i)$ $(i = 0, 1, 2, \cdots, N)$.

Proof:

- Existence: Define $\Phi(x) = \sum_{i=0}^{N} f(x_i)\varphi_i(x)$.

Continuous piecewise linear approximation

Theorem (II)

Given an $f \in C[a,b]$, there is a unique $\Phi \in L_\Delta$ which satisfies $\Phi(x_i) = f(x_i)$ $(i = 0, 1, 2, \cdots, N)$.

Proof:

- Existence: Define $\Phi(x) = \sum_{i=0}^{N} f(x_i)\varphi_i(x)$.
- Then $\Phi \in L_\Delta$ and $\Phi(x_j) = f(x_j)$ $(j = 0, \cdots, N)$ since $\varphi_i(x_j) = \delta_{ij}$.

Continuous piecewise linear approximation

Theorem (II)

Given an $f \in C[a, b]$, there is a unique $\Phi \in L_\Delta$ which satisfies $\Phi(x_i) = f(x_i)$ $(i = 0, 1, 2, \cdots, N)$.

Proof:

- Existence: Define $\Phi(x) = \sum_{i=0}^{N} f(x_i)\varphi_i(x)$.
- Then $\Phi \in L_\Delta$ and $\Phi(x_j) = f(x_j)$ $(j = 0, \cdots, N)$ since $\varphi_i(x_j) = \delta_{ij}$.
- Uniqueness: assume there are two such $\Phi' s$, say Φ_1 and Φ_2.

Continuous piecewise linear approximation

Theorem (II)

Given an $f \in C[a,b]$, there is a unique $\Phi \in L_\Delta$ which satisfies $\Phi(x_i) = f(x_i)$ $(i = 0, 1, 2, \cdots, N)$.

Proof:

- Existence: Define $\Phi(x) = \sum_{i=0}^{N} f(x_i)\varphi_i(x)$.
- Then $\Phi \in L_\Delta$ and $\Phi(x_j) = f(x_j)$ $(j = 0, \cdots, N)$ since $\varphi_i(x_j) = \delta_{ij}$.
- Uniqueness: assume there are two such $\Phi' s$, say Φ_1 and Φ_2.
- Then $\Phi_1(x_j) - \Phi_2(x_j) = f(x_j) - f(x_j) = 0$ $(j = 0, \cdots, N)$.

Continuous piecewise linear approximation

Theorem (II)

Given an $f \in C[a,b]$, there is a unique $\Phi \in L_\Delta$ which satisfies $\Phi(x_i) = f(x_i)$ $(i = 0, 1, 2, \cdots, N)$.

Proof:

- Existence: Define $\Phi(x) = \sum_{i=0}^{N} f(x_i)\varphi_i(x)$.
- Then $\Phi \in L_\Delta$ and $\Phi(x_j) = f(x_j)$ $(j = 0, \cdots, N)$ since $\varphi_i(x_j) = \delta_{ij}$.
- Uniqueness: assume there are two such $\Phi's$, say Φ_1 and Φ_2.
- Then $\Phi_1(x_j) - \Phi_2(x_j) = f(x_j) - f(x_j) = 0$ $(j = 0, \cdots, N)$.
- Also, $\Phi_1 - \Phi_2$ is linear in $[x_j, x_{j+1}]$ $(j = 0, \cdots, N-1)$.

Continuous piecewise linear approximation

Theorem (II)

Given an $f \in C[a,b]$, there is a unique $\Phi \in L_\Delta$ which satisfies $\Phi(x_i) = f(x_i)$ $(i = 0, 1, 2, \cdots, N)$.

Proof:

- Existence: Define $\Phi(x) = \sum_{i=0}^{N} f(x_i)\varphi_i(x)$.
- Then $\Phi \in L_\Delta$ and $\Phi(x_j) = f(x_j)$ $(j = 0, \cdots, N)$ since $\varphi_i(x_j) = \delta_{ij}$.
- Uniqueness: assume there are two such $\Phi's$, say Φ_1 and Φ_2.
- Then $\Phi_1(x_j) - \Phi_2(x_j) = f(x_j) - f(x_j) = 0$ $(j = 0, \cdots, N)$.
- Also, $\Phi_1 - \Phi_2$ is linear in $[x_j, x_{j+1}]$ $(j = 0, \cdots, N-1)$.
- Then $\Phi_1 - \Phi_2 = 0$ in $[x_j, x_{j+1}]$ $(j = 0, \cdots, N-1)$.

Continuous piecewise linear approximation

Theorem (II)

Given an $f \in C[a,b]$, there is a unique $\Phi \in L_\Delta$ which satisfies $\Phi(x_i) = f(x_i)$ $(i = 0, 1, 2, \cdots, N)$.

Proof:

- Existence: Define $\Phi(x) = \sum_{i=0}^{N} f(x_i)\varphi_i(x)$.
- Then $\Phi \in L_\Delta$ and $\Phi(x_j) = f(x_j)$ $(j = 0, \cdots, N)$ since $\varphi_i(x_j) = \delta_{ij}$.
- Uniqueness: assume there are two such $\Phi's$, say Φ_1 and Φ_2.
- Then $\Phi_1(x_j) - \Phi_2(x_j) = f(x_j) - f(x_j) = 0$ $(j = 0, \cdots, N)$.
- Also, $\Phi_1 - \Phi_2$ is linear in $[x_j, x_{j+1}]$ $(j = 0, \cdots, N-1)$.
- Then $\Phi_1 - \Phi_2 = 0$ in $[x_j, x_{j+1}]$ $(j = 0, \cdots, N-1)$.
- Hence $\Phi_1(x) = \Phi_2(x) \Rightarrow$ uniqueness of Φ.

Continuous piecewise linear approximation

Definition (Interpolation)

$\Phi(x) = \sum_{i=0}^{N} f(x_i)\varphi_i(x)$ is called the L_Δ interpolation of f, which is denoted by $I_N f(x)$.

Continuous piecewise linear approximation

Definition (Interpolation)

$\Phi(x) = \sum_{i=0}^{N} f(x_i)\varphi_i(x)$ is called the L_Δ interpolation of f, which is denoted by $I_N f(x)$.

Remark

- *$I_N : C[a,b] \to L_\Delta$ is a linear operator, i.e.,*

$$I_N(a_1 f_1(x) + a_2 f_2(x)) = a_1 I_N f_1(x) + a_2 I_N f_2(x).$$

Continuous piecewise linear approximation

Theorem (III)

If $f \in C^2[a, b]$, *then*

$$\begin{aligned} \|f - I_N f\|_\infty &\leq \frac{1}{8} h^2 \|f''\|_\infty \,, \\ \|(f - I_N f)'\|_\infty &\leq \frac{1}{2} h \|f''\|_\infty \,. \end{aligned}$$

If $f \in C^1[a, b]$, *then*

$$\|f - I_N f\|_\infty \leq \frac{1}{2} h \|f'\|_\infty \,.$$

If $f \in C[a, b]$, *then*

$$\|f - I_N f\|_\infty \to 0.$$

Proof: See pages 241-242 of textbook (Independent study problem).

Outline

1. Introduction
2. Best approximation
3. Polynomial Approximation
4. Piecewise polynomial approximation
5. Other topics

Topics: Independent study problems

- Newton form of the Lagrange polynomials (section 4.3.4)
- Least square approximation (section 4.3.6 and 4.3.7)
- Minmax approximation (section 4.3.8)
- Interval bounds on the errors (section 4.3.9)
- Cubic spline interpolation and B-splines (section 4.4.2 and 4.4.3).
- Trigonometric Approximation (section 4.5)
- Rational approximation (section 4.6)
- Wavelet (section 4.7)
- ......