Main question of Approximation Theory

Let $Y$ be a finite dimensional subspace of a normed space $X$. For all $x \in X$ there exists a best approximation $y_{x} \in Y$ of $x$ in $Y$, that is, $\left\|x-y_{x}\right\|=\min _{y \in Y}\|x-y\|$.

Approximation Theory

Definition \& Remark \& (Non)Examples

Strictly convex norm / space

Approximation Theory

Theorem w/ proof idea

Weierstrass: For $f \in \mathcal{C}([a, b])$ and $\varepsilon>0$ there exists a $p \in P$ with $\|f-p\|_{\infty}<\varepsilon$.

For best approximation, closedness and norm are important

Approximation Theory

Theorem w/ Proof
Let $Y$ be subspace of a normed space $X$ and $x \in X$. The set $Y_{x} \subset Y$ of best approximations of $x$ in $Y$ is bounded and convex.
(If $Y$ is finite dimensional, $Y_{x}$ is closed for every $x \in X$.)
Approximation Theory

Corollary w/ proof

Let $Y \subset X$ be a subspace of a strictly convex space $(X,\|\cdot\|)$ and $x \in X$. Then either $Y_{x}=\varnothing$ or $Y_{x}=\left\{y_{x}\right\}$.

Approximation Theory

Definition, Remark

Bernstein (basis) polynomials

Let $X:=\left(\mathbb{R}^{2},\|\cdot\|_{\infty}\right), Y:=\operatorname{span}\left((0,1)^{\top}\right)$ and $x:=(1,0)^{\top}$. For $y=$ $\left(y_{1}, 0\right) \in Y$ we have $\|y-x\|_{\infty}=\left\|\left(0, y_{1}\right)^{\top}-(1,0)^{\top}\right\|_{\infty}=\max \left(\left|y_{1}\right|, 1\right) \geqslant$ 1. If $y_{1} \in[-1,1]$, then $\|x-y\|_{\infty}=1=\min _{y \in Y}\|y-x\|_{\infty}$, so there are infinitely many best approximations to $x$ in $Y$. If we instead consider $\left(\mathbb{R}^{2},\|\cdot\|_{2}\right)$, then the unique best approximation is $y^{*}=0$, as $\left\|x-y^{*}\right\|_{2}=\|x\|_{2}=1<\|x-y\|_{2}$ for all $y \in Y \backslash\{0\}$.
In general, if $Y \subset X$ is not closed, for $x \in \bar{Y} \backslash Y$, there is no best approximation of $x$ in $Y$. Indeed, as $x \in \bar{Y}$, there exists a sequence $\left(y_{n}\right)_{n \in \mathbb{N}} \subset Y$ with $y_{n} \rightarrow x$, that is, $\left\|y_{n}-x\right\| \xrightarrow{n \rightarrow \infty} 0$. If there were a $y^{*} \in Y$ such that $\left\|y^{*}-x\right\|=\min _{y \in Y}\|y-x\|$, then $\left\|y^{*}-x\right\|=0$ and thus $x=y^{*} \in Y$, which is a contradiction to $x \in \bar{Y} \backslash Y$.

If $Y_{x}$ is empty, the statement holds, so assume $Y_{x} \neq \varnothing$.
Boundedness. By definition, $Y_{x}$ is a subset of the ball $S_{d_{x}}:=$ $\left\{y \in Y:\|x-y\| \leqslant d_{x}\right\}$, where $d_{x}:=\min _{y \in Y}\|x-y\|$.
Convexity. Suppose $Y_{x}$ is not a singleton. For $y_{1}, y_{2} \in Y_{x}$ and $\lambda \in(0,1)$ define $\tilde{y}:=\lambda y_{1}+(1-\lambda) y_{2}$. We have $\tilde{y} \in Y$ as $Y$ is a subspace and thus convex. Then $\left\|x-y_{1}\right\|=\left\|x-y_{2}\right\|=$ $\min _{y \in Y}\|x-y\|$ and $\|x-\tilde{y}\|=\left\|\lambda x+(1-\lambda) x-\lambda y_{1}-(1-\lambda) y_{2}\right\| \leqslant$ $\lambda\left\|x-y_{1}\right\|+(1-\lambda)\left\|x-y_{2}\right\|=(X+1 \not \subset) \min _{y \in Y}\|x-y\|=$ $\min _{y \in Y}\|x-y\|$, so $\|x-\tilde{y}\|=\min _{y \in Y}\|x-y\|$ and hence $\tilde{y} \in Y_{x}$, so $Y_{x}$ is convex.

It suffices to check the case where the convex set $Y_{x}$ has infinitely many elements and deduce a contradiction. If $y_{1}, y_{2} \in Y$ with $y_{1} \neq y_{2}$, then by convexity of $Y_{x}$ we have $\lambda y_{1}+(1-\lambda) y_{2} \in$ $Y_{x} \subset S_{d_{x}}$, but $S_{d_{x}}$ can not contain this line $\left\{\lambda y_{1}+(1-\lambda) y_{2}:\right.$ $\lambda \in(0,1)\}$, as $X$ is strictly convex.

For a bounded function $f$ on $[0,1]$, the Bernstein polynomial of degree $n$ is

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

where $\left.\binom{n}{k} x^{k}(1-x)^{n-k}\right)_{k=0}^{n}$ are the BERNSTEIN basis polynomials of degree $n$.

For a sequence of linear positive operators $\left(T_{n}: \mathcal{C}([0,1]) \rightarrow\right.$ $\mathcal{C}([0,1]))_{n \in \mathbb{N}}$, the following are equivalent.

1. $T_{n} f \rightrightarrows f$ for all $f \in \mathcal{C}([0,1])$.
2. $T_{n} f \rightrightarrows f$ for all $f \in\left(f_{k}\right)_{k=0}^{2}$, where $f_{k}(x):=x^{k}$.
3. $T_{n} f_{0} \rightrightarrows f_{0}$ and $\left(t \mapsto\left(T_{n} \varphi_{t}\right)(t)\right) \rightrightarrows 0$, where $\varphi_{t}(x):=$ $(x-t)^{2}$.

The main questions are the ones of existence, uniqueness, construction (how can we find or construct such a approximation) and measure (e.g. choice of norm).

Let $x \in X$. Since $Y$ is a vector space, $0 \in Y$ and thus $\min _{y \in Y}\|x-y\| \leqslant\|x-0\|=\|x\|$. Hence any best approximation must be in the set $K:=\{y \in Y:\|x-y\| \leqslant\|x\|\} \subset Y$, which is bounded and closed and thus compact. The function $f: K \rightarrow \mathbb{R}, y \mapsto\|x-y\|$ is continuous on the compact set $K$ so its attains a minimum in $y_{x} \in K$, which is the best approximation.

A norm $\|\cdot\|$ is strictly convex if for all $x, y \in X$ with $x \neq y$ and $\|x\|=\|y\|=r>0$ and all $\lambda \in(0,1)$ we have $\| \lambda x+(1-$ $\lambda) y \|<r$. We say that $(X,\|\cdot\|)$ is a strictly convex space.

Geometrically, this corresponds the straight line segment between to two distinct points on the boundary of any $\|\cdot\|$-ball being contained in the ball.

The $p$-norm on $\mathbb{R}^{n}$ is strictly convex for $p \in(1, \infty)$ and not strictly convex for $p \in\{1, \infty\}$. The $L^{2}$-norm on $\mathcal{C}([a, b])$ is strictly convex, while $\|\cdot\|_{\infty}$ is not.

It suffices to show that $B_{n}\left(f_{m}\right) \rightrightarrows f_{k}$ for $k \in\{0,1,2\}$ by the Bohman-Korovkin theorem. We have $B_{n}\left(f_{0}\right)=f_{0}$, $B_{n}\left(f_{1}\right)=f_{1}$ and $B_{n}\left(f_{2}\right)=\frac{n-1}{n} f_{2}+\frac{1}{n} f_{1}$.

Corollary: Let $f \in \mathcal{C}^{1}([a, b])$ and $\varepsilon>0$. Then there exists a polynomial $p$ such that $\|f-p\|_{\infty}<\varepsilon$ and $\left\|f^{\prime}-p^{\prime}\right\|_{\infty}<\varepsilon$.

An operator $T: \mathcal{C}([0,1]) \rightarrow \mathcal{C}([0,1])$ is positive if $f(x) \geqslant 0$ for all $x \in[0,1]$ implies $(T f)(x) \geqslant 0$ for all $x \in[0,1]$.

Every linear positive operator is bounded.
For every $f \in \mathcal{C}([0,1])$ we have $\pm f \leqslant\|f\|_{\infty} \cdot 1$. As $T$ is linear and positive, we have $\pm T(f) \leqslant\|f\|_{\infty} T(1)$ and thus $|T(f)| \leqslant\|f\|_{\infty} T(1)$. Hence $\|T\| \leqslant\|T(1)\|_{\infty}$, so $T$ is bounded.

## Modulus of continuity

Approximation Theory

Lemma w/ Proof

Bounds on the modulus of continuity: $w_{f}$ is not linear, but ...

## Convergence rate Bernstein approximation

Approximation Theory

Theorem w/ Proof

If $\left(x_{k}\right)_{k=0}^{n}$ are distinct points and $\left(y_{k}\right)_{k=0}^{n} \subset \mathbb{R}$, then there exists a unique polynomial $p \in P_{n}$ such that $p\left(x_{k}\right)=y_{k}$,

$$
k \in\{0, \ldots, n\} .
$$

Approximation Theory

Theorems w/o proof, Example

Properties of the modulus of continuity

Definition, Remark

$\alpha$-LIPSCHITZ continuity

Remark

Advantages and disadvantages of Bernstein polynomials

Definition \& Remark

Lagrange interpolating operator

Approximation Theory

Remark

Uniform convergence of Lagrange approximation?

Advantages and disadvantages of
Lagrange interpolation

- $f$ is uniformly continuous $\Longleftrightarrow \lim _{\delta \backslash 0} w_{f}(\delta)=0$, that is, for all $\varepsilon>0$ there exists a $\delta>0$ such that $w_{f}(\delta)<\varepsilon$.
- For $0<\delta<\delta_{1}$ we have $w_{f}(\delta) \leqslant w_{f}\left(\delta^{\prime}\right)$.
- The set $A$ of continuous functions that have the same modulus of continuity are uniformly equicontinuous.

A function $f:[a, b] \rightarrow \mathbb{R}$ is LIPSChitz continuous of order $\alpha$ and we write $f \in \operatorname{Lip}_{K}^{\alpha}$ if there exists a positive constant $K>0$ such that

$$
|f(x)-f(y)| \leqslant K|x-y|^{\alpha} \quad \forall x, y \in[a, b]
$$

We have $f \in \operatorname{Lip}_{K}^{\alpha}$ if and only if $w_{f}(\delta) \leqslant K \delta^{\alpha}$, that is $w_{f}(\delta) \in$ $O\left(\delta^{-\alpha}\right)$ for $\delta \searrow 0$.

The disadvantages of the Bernstein polynomials are that the convergence $B_{n}(f) \rightrightarrows f$ is too slow (not optimal by JACKson and not improvable) to be useful in applications and we have $B_{n}(f) \neq f$ for $f \in P$ (so $B_{n}^{2} \neq B_{n}$ ), e.g. if $f(x)=x^{2}$ and $\varepsilon=10^{-4}$, then we need $n>2500$ for $\left\|f-B_{n}(f)\right\|<\varepsilon$. $B_{n}$ is not self-adjoint.
The advantages are that $B_{n}(f)$ is linear and positive, we have an error bound and that $B_{n}(f) \rightrightarrows f$ for all $f \in \mathcal{C}([a, b])$ and that if $f \in \mathcal{C}^{m}([a, b])$, then $\left(B_{n}(f)\right)^{(k)} \rightrightarrows f^{(k)}$ for all $k \in\{0, \ldots, m\}$. The basis functions are a partition of unity.

The (self-adjoint!) LAGRANGE interpolating operator is

$$
L_{n}: \mathcal{C}([a, b]) \rightarrow P_{n}, \quad f \mapsto \sum_{k=0}^{n} f\left(x_{k}\right) \ell_{k}(x)
$$

where

$$
\left.\ell_{k}(x)\right):=\prod_{\substack{j=0 \\ j \neq k}}^{n} \frac{x-x_{j}}{x_{k}-x_{j}}
$$

has degree $n$. is independent of $\left(y_{k}\right)_{k=0}^{n}$. Since $\ell_{k}\left(x_{j}\right)=\delta_{j, k}$, the $L_{n}(f)$ interpolates $f$ at $\left(x_{k}\right)_{k=0}^{n} . \ell_{k}$ are partition of unity.

Advantages: they are linear projections and that the basis polynomials only depend on $\left(x_{k}\right)_{k=0}^{n}$, so interpolating multiple functions at the same points is easy.
Disadvantage: not positive, removing or adding one point $x_{k}$ yields completely different basis functions. Furthermore, $L_{n}(f) \not \ddagger f$. For $f=|\cdot|$ on $[-1,1], L_{n}(f) \rightarrow f$ only for $x \in\{ \pm 1,0\}$ (better with rational approximation).

The modulus of continuity for a bounded function $f$ on $[a, b]$ is
$w_{f}:[0, \infty) \rightarrow[0, \infty), \quad \delta \mapsto \sup \{|f(x)-f(y)|:|x-y| \leqslant \delta\}$.
$|f(x)-f(y)| \leqslant w_{f}(|x-y|), w_{f}(0)=0$.
If $f \in \mathcal{C}^{1}([a, b])$, then for all $x, y \in[a, b]$ we have $\frac{|f(x)-f(y)|}{|x-y|} \leqslant$ $\left\|f^{\prime}\right\|_{\infty}$, so $w_{f}(\delta) \leqslant\left\|f^{\prime}\right\|_{\infty} \delta$. (For LIPSCHITZ functions replace by LIPSCHITZ constant.)

For $n \in \mathbb{N}, w_{f}(n \cdot) \leqslant n w_{f}$ and $w_{f}(\lambda \cdot) \leqslant(1+\lambda) w_{f}$ for all $\lambda>0$. Take $x, y \in[0,1]$ with $|x-y|<n \delta$ and then consider the following equidistant partition of the interval between them: $n_{k}:=x+\frac{k}{n}(y-x)$ for $k \in\{0, \ldots, n\}$. Then $\left|n_{k+1}-n_{k}\right|=\frac{|y-x|}{n}|k+1-k|<\delta$ and thus

$$
\begin{aligned}
|f(x)-f(y)| & =\left|\sum_{k=0}^{n-1} f\left(x_{k+1}\right)-f\left(x_{k}\right)\right| \stackrel{\Delta \neq \sum_{k=0}^{n-1}\left|f\left(x_{k+1}\right)-f\left(x_{k}\right)\right|}{\leqslant} \\
& \leqslant \sum_{k=0}^{n-1} w_{f}(\delta)=n w_{f}(\delta) .
\end{aligned}
$$

For $\lambda \in \mathbb{R}_{+} \backslash \mathbb{N}$ there exists a $n \in \mathbb{N}$ with $\lambda \in(n, n+1)$. As $w_{f}$ is increasing, $w_{f}(\lambda \delta) \leqslant w_{f}((n+1) \delta) \leqslant(n+1) w_{f}(\delta)<(\lambda+1) w_{f}(\delta)$ for all $\delta>0$.

For a bounded function $f$ on $[0,1],\left\|f-B_{n}(f)\right\|_{\infty} \leqslant \frac{3}{2} w_{f}\left(\frac{1}{\sqrt{n}}\right)$.

```
|f(x)-(\mp@subsup{B}{n}{}f)(x)|}\stackrel{\Delta\not=}{\leqslant}\mp@subsup{\sum}{k=0}{n}|f(x)-f(\frac{k}{n})|(\begin{array}{c}{n}\\{k}\end{array})\mp@subsup{x}{}{k}(1-x\mp@subsup{)}{}{n-k}
wf}(|x-\frac{k}{n}|)=\mp@subsup{w}{f}{}(\frac{1}{\sqrt{}{n}}\sqrt{}{n}|x-\frac{k}{n}|)\leqslant(1+\sqrt{}{n}|x-\frac{k}{n}|)\mp@subsup{w}{f}{}(\frac{1}{\sqrt{}{n}})
|f(x)-(\mp@subsup{B}{n}{}f)(x)|=\mp@subsup{w}{f}{}(\frac{1}{\sqrt{}{n}})(1+\sqrt{}{n}\mp@subsup{\sum}{k=0}{n}|x-\frac{k}{n}|(\begin{array}{l}{n}\\{k}\end{array})\mp@subsup{x}{}{k}(1-x\mp@subsup{)}{}{k}).
(\mp@subsup{\sum}{k=0}{n}|x-\frac{k}{n}||(\begin{array}{l}{n}\\{k}\end{array})\mp@subsup{x}{}{k}(1-x\mp@subsup{)}{}{k}\mp@subsup{)}{}{2}}\leqslant=\quad(\mp@subsup{\sum}{k=0}{n}|x-\frac{k}{n}\mp@subsup{|}{}{2}(\begin{array}{l}{n}\\{k}\end{array})\mp@subsup{x}{}{k}(1-x\mp@subsup{)}{}{k}
( 㳸 (\begin{array}{l}{n}\\{k}\end{array})\mp@subsup{x}{}{k}(1-x\mp@subsup{)}{}{k})=(\mp@subsup{\sum}{k=0}{n}(\mp@subsup{x}{}{2}+\frac{\mp@subsup{k}{}{2}}{\mp@subsup{n}{}{2}}-2x\frac{k}{n})(\begin{array}{l}{n}\\{k}\end{array})\mp@subsup{x}{}{k}(1-x\mp@subsup{)}{}{k})=
x + +(Bn}\mp@subsup{f}{2}{\prime})(x)-2x(\mp@subsup{B}{n}{}\mp@subsup{f}{1}{})(x)=\frac{x-\mp@subsup{x}{}{2}}{n}\leqslant\frac{1}{4n}
|f(x)-(\mp@subsup{B}{n}{}f)(x)|\leqslant\mp@subsup{w}{f}{}(\frac{1}{\sqrt{}{n}})(1+\sqrt{}{n}\frac{1}{2\sqrt{}{n}})\leqslant\frac{3}{2}\mp@subsup{w}{f}{}(\frac{1}{\sqrt{}{n}}).
```

Let $p(x):=\sum_{k=0}^{n} c_{k} x^{k}$. Formulating the interpolation condition as matrix multiplication yields

$$
V c=:\left(\begin{array}{cccc}
1 & x_{0} & \ldots & x_{0}^{n} \\
1 & x_{1} & \ldots & x_{1}^{n} \\
1 & x_{n} & \ldots & x_{n}^{n}
\end{array}\right)\left(\begin{array}{c}
c_{0} \\
\vdots \\
c_{n}
\end{array}\right)=\left(\begin{array}{c}
y_{0} \\
\vdots \\
y_{n}
\end{array}\right)=: y .
$$

The matrix $V$ is the Vandermonde matrix with $\operatorname{det}(V)=$ $\prod_{1 \leqslant j<i \leqslant n} x_{i}-x_{j}$. Since the $\left(x_{k}\right)_{k=0}^{n}$ are distinct, $\operatorname{det}(V) \neq 0$, so $V$ is invertible and there exists a unique solution $c=V^{-1} y$.

Given an array of points $x$, there exists an $f \in \mathcal{C}([a, b])$ such that $\left\|L_{n}(f)-f\right\|_{\infty} \rightarrow \infty$, e.g. $\frac{1}{x^{2}+1}$ on $[a, b]=[-5,5]$ with equidistant nodes. Adding more points makes it worse (oscillation at the endpoints). Better: ChEBYCHEv nodes. This is due to:
Given a sequence $\left(T_{n}: \mathcal{C}([a, b]) \rightarrow P_{n}\right)_{n \in \mathbb{N}}$ of linear continuous projections, there exists a $f \in \mathcal{C}([a, b])$ such that $\| T_{n}(f)-$ $f \|_{\infty} \rightarrow \infty$ (Kharshiladze, Lozinski (1941)).

We have $\left\|L_{n}\right\|=\left\|\sum_{k=0}^{n} \mid \ell_{k}(\cdot)\right\|_{\infty}=: \Lambda_{n}$ and $\left\|f-L_{n}(f)\right\|_{\infty} \leqslant(1+$ $\left.\Lambda_{n}\right) E_{n}(f)$.

Newton's Method

Approximation Theory

Theorem w/ Proof

There exists a unique $p \in P_{m}$ such that $p^{(j)}\left(x_{i}\right)=f^{(j)}\left(x_{i}\right), 0 \leqslant i \leqslant n, 0 \leqslant j \leqslant J_{i}$ for every data set of distinct $\left(x_{k}\right)_{k=0}^{n}$ and $\left(f^{(j)}\left(x_{i}\right)\right)_{j=0}^{J_{i}}$, where $m=\sum_{i=0}^{n}\left(J_{i}+1\right)-1$.

For all $f \in \mathcal{C}[a, b]$ we have $L_{n}(f) \rightrightarrows f$ if $\left(x_{k}\right)_{k=1}^{n}$ are the zeros of the $n$-th Chebyshev polynomial. (Fejér-Hermite operator)

Characterisation of best approximations

Upper bound on polynomial approximation error

## Definition

## Fejér-Hermite Operator

Approximation Theory

Definition, Lemma w/o proof

## Chebychev polynomials

Definition

HaAr space

Theorem w/o Proof

Characterisation / Alternation Theorem

de La Vallée Poussin

If $f \in \mathcal{C}^{n+1}([a, b])$ is interpolated by $p \in P_{n}$ at $\left(x_{k}\right)_{k=0}^{n}$, then $\|f-p\|_{\infty} \leqslant \frac{1}{(n+1)!}\left\|f^{(n+1)}\right\|_{\infty}\|W\|_{\infty}$, where $W:=\prod_{k=0}^{n}\left(\cdot-x_{k}\right)$. We show that for all $x \in[a, b]$ there exists a $\xi_{x} \in[a, b]$ such that $f(x)-p(x)=$ $\frac{1}{(n+1)!} f^{(n+1)}\left(\xi_{x}\right) W(x)$. If $x=x_{k}$ for some $k \in\{0, \ldots, n\}$, then the LHS is zero as $p$ interpolates $f$ but the RHS is also zero since $W$ vanishes on $\left(x_{k}\right)_{k=0}^{n}$. If $x \neq x_{k}$ for all $k \in\{0, \ldots, n\}$, define the scalar $\lambda_{x}:=\frac{f(x)-p(x)}{W(x)}$ and the function $\varphi:=f-p-\lambda_{x} W$. As $f, p, W \in \mathcal{C}^{n+1}([a, b])$, so is $\varphi$. We have $\varphi\left(x_{k}\right)=0$ for all $k \in\{0, \ldots, n\}$ and $\varphi(x)=0$, so $\varphi$ has at least $n+2$ zeros. By Rolle's Theorem, $\varphi^{\prime}$ has at least $n+1$ zeros, so, inductively, $\varphi^{(n+1)}$ has at least one zero $\xi_{x}$. Hence $0=\varphi^{(n+1)}\left(\xi_{x}\right)=f^{(n+1)}\left(\xi_{x}\right)-0-\cdot \lambda_{x}(n+1)$ ! and so $\lambda_{x}=\frac{1}{(n+1)!} f^{(n+1)}\left(\xi_{x}\right)$.
$\|W\|_{\infty}$ is minimal for Chebychev nodes.

The $n$-th Fejér-Hermite operator is

$$
L_{n}: \mathcal{C}([a, b]) \rightarrow P_{2 n-1}, \quad f \mapsto \sum_{i=1}^{n} f\left(x_{i}\right) A_{i}
$$

where

$$
A_{k}(x):=\left(1-2\left(x-x_{k}\right) \ell_{k}^{\prime}\left(x_{k}\right)\right) \ell_{k}^{2}(x)
$$

for distinct $\left(x_{k}\right)_{k=1}^{n} \subset[a, b]$ such that $\left(L_{n} f\right)\left(x_{k}\right)=f\left(x_{k}\right)$ and $\left(L_{n} f\right)^{\prime}\left(x_{k}\right)=0$ for all $k \in\{1, \ldots, n\}$.

For $n \in \mathbb{N}_{\geqslant 1}$, the polynomials $T_{n}$ and $U_{n-1}$ of degree exactly $n$ and $n-1$ such that $\cos (n x)=T_{n}(\cos (x))$ and $\sin (n x)=$ $U_{n-1}(\cos (x)) \sin (x)$ are the Chebychev polynomials of first and second order.
Roots of $T_{n}$ are $\left(\cos \left(\frac{2 k+1}{2 n} \pi\right)\right)_{k=0}^{n} \subset[-1,1]$, extrema are $\left(\cos \left(\frac{k}{n} \pi\right)\right)_{k=0}^{n} \subset[-1,1]$.
The leading coefficient of $T_{n}$ is $2^{n-1}$.
We have $\left(T_{n}, T_{m}\right)_{w}=\frac{\pi}{2} \delta_{n, m}($ for $n \neq 0)$ for $w(x):=\frac{1}{\sqrt{1-x^{2}}}$ and $\left(U_{n}, U_{m}\right)_{w}=\frac{\pi}{2} \delta_{n, m}$ for $w(x):=\sqrt{x^{2}-1}$.

The functions $\left(g_{k}\right)_{k=0}^{n} \subset \mathcal{C}([a, b])$ satisfy the HAAR condition if every $n+1$ vectors $\left(g_{k}\left(x_{j}\right)\right)_{k=0}^{n}$ for $j \in\{0, \ldots, n\}$ are linearly independent, that is, the matrix $\left(g_{k}\left(x_{j}\right)\right)_{j, k=0}^{n}$ is invertible for all sets of distinct points $\left(x_{j}\right)_{j=0}^{n} \subset[a, b]$. Then $\mathcal{A}:=\operatorname{span}\left(g_{0}, \ldots, g_{n}\right)$ is a HAAR space and $\left(g_{k}\right)_{k=0}^{n}$ is a CheBYCHEV system or HAAR system.

Let $\mathcal{A} \subset \mathcal{C}([a, b])$ be a $(n+1)$-dimensional HAAR space. Let $f \in \mathcal{C}([a, b])$ and $p^{*} \in \mathcal{A}$ such that $f-p^{*}$ alternates in sign at $n+2$ points $a \leqslant \xi_{0}<\ldots<\xi_{n+1} \leqslant b$. Then

$$
\begin{equation*}
E_{n}(f):=\min _{p \in \mathcal{A}}\|f-p\|_{\infty} \geqslant \min _{i \in\{0, \ldots, n+1\}}\left|f\left(\xi_{i}\right)-p^{*}\left(\xi_{i}\right)\right| . \tag{2}
\end{equation*}
$$

Given distinct $\left(x_{k}\right)_{k=0}^{n} \subset[a, b]$ and $\left(y_{k}\right)_{k=0}^{n} \subset \mathbb{R}$, define $u_{0} \equiv 1$ and

$$
\begin{aligned}
p(x):= & \sum_{k=0}^{n}\left[y_{0}, \ldots, y_{k}\right] u_{k}(x), \quad \text { where } u_{k}(x):=\prod_{j=0}^{k-1}\left(x-x_{j}\right), \\
& {\left[y_{k}\right]:=y_{k}, \quad\left[y_{k}, y_{j}\right]:=\frac{y_{j}-y_{k}}{x_{j}-x_{k}} \text { for } k \neq j, }
\end{aligned}
$$

and
$\left[y_{j_{0}}, \ldots, y_{j_{m}}\right]:=\frac{\left[y_{j_{1}}, \ldots, y_{j_{m}}\right]-\left[y_{j_{0}}, \ldots, y_{j_{m}-1}\right]}{x_{j_{m}}-x_{j_{0}}}$ for $j_{0} \neq j_{m}$.

The system has a unique solution if the matrix $A$ of this linear system is invertible. Let $p \in P_{m}$ such that

$$
p^{(j)}\left(x_{i}\right)=0 \quad \forall i \in\{0, \ldots, n\}, j \in\left\{0, \ldots, J_{i}\right\}
$$

which corresponds to the homogeneous system $A c=0$, where $c$ are the coefficients of $p$. Then $p$ has $n+1$ roots $\left(x_{i}\right)_{i=0}^{n}$ with multiplicities $J_{i}+1$. Hence $P$ has at least $\sum_{j=0}^{n}\left(J_{i}+1\right)=m+1$ roots, so $p \equiv 0$ and thus $c=0$. Hence $A$ is injective, so it is invertible.

We can without loss of generality assume that $[a, b]=[-1,1]$. We show that for this special choice of $\left(x_{k}\right)_{k=1}^{n}$, we have

$$
L_{n}(f)(x)=\frac{1}{n^{2}} T_{n}(x)^{2} \sum_{i=1}^{n} f\left(x_{i}\right) \frac{1-x x_{i}}{\left(x-x_{i}\right)^{2}}
$$

because then $L_{n}(f)$ is clearly linear and positive.

Let $f \in \mathcal{C}([a, b]), \mathcal{A} \subset \mathcal{C}([a, b])$ be a linear subspace and $p^{*} \in$ $\mathcal{A}$. Then $p^{*}$ is a best approximation of $f$ in $\mathcal{A}$ if and only if there is no $p \in \mathcal{A}$ such that

$$
\begin{equation*}
\left(f(x)-p^{*}(x)\right) p(x)>0 \quad \forall x \in E_{M} \tag{1}
\end{equation*}
$$

where

$$
E_{M}:=\left\{x \in[a, b]:\left|f(x)-p^{*}(x)\right|=\left\|f-p^{*}\right\|_{\infty}\right\}
$$

Let $\mathcal{A} \subset \mathcal{C}([a, b])$ be a HAAR space of dimension $n+1$ and $f \in \mathcal{C}([a, b])$. Then $p^{*} \in \mathcal{A}$ is a best approximation to $f$ in $\mathcal{A}$ if and only if there exist $n+2$ points $\left\{\xi_{0}, \ldots, \xi_{n+1}\right\}$ such that

- $a \leqslant \xi_{0}<\xi_{1}<\ldots<\xi_{n+1} \leqslant b$
- $\left|f\left(\xi_{i}\right)-p^{*}\left(\xi_{i}\right)\right|=\left\|f-p^{*}\right\|_{\infty}$ for all $i \in\{0, \ldots, n+1\}$,
- $f\left(\xi_{i+1}\right)-p^{*}\left(\xi_{i+1}\right)=-\left(f\left(\xi_{i}\right)-p^{*}\left(\xi_{i}\right)\right)$ for all $i \in\{0, \ldots, n\}$.


## Uniqueness theorem

## Approximation Theory

Theorem w/o Proof

Markovs inequalities

Approximation Theory

Proof

Markov inequality

Approximation Theory

Lemma w/ proof, Remark

The weighted norm $\|\cdot\|_{w}$ is strictly convex on $\mathcal{C}([a, b] ; \mathbb{R})$.

Approximation Theory

Theorem w/o Proof

Nonexistence theorem

Theorem w/o proof, Remark

Markov-Bernstein inequality

Definition, Remarks

Least squares approximation

Least squares characterisation theorem

Approximation Theory

Theorem, Lemma w/o proofs

Explicit representation of the best
approximation

Recurrence relation for orthogonal polynomials

Let $\mathcal{A}:=\operatorname{span}\left\{\varphi_{0}, \varphi_{1}, \ldots\right\} \subset \mathcal{C}([a, b])$ be a subspace of $\mathcal{C}([a, b])$ such that for every $n$, the set $\left\{\varphi_{0}, \ldots, \varphi_{n}\right\}$ satisfies the HAAR condition. Then no point outside $\mathcal{A}$ has a best approximation from $\mathcal{A}$.

Such a system $\left\{\varphi_{0}, \varphi_{1}, \ldots\right\}$ is called a Markov system.

Corollary: 1st, 2nd Weierstrass theorem.

The analogue of Markov's inequality for the unit disk in $\mathbb{C}$ due to Bernstein states that if $S(z)=\sum_{k=0}^{n} a_{k} z^{k}$ is a polynomial for $z \in \mathbb{C}$, then

$$
\max _{|z|=1}\left|S^{\prime}(z)\right| \leqslant n \max _{|z|=1}|S(z)|
$$

and we have equality for $p(z)=\lambda z^{n}$, where $\lambda \in \mathbb{C}$.
Using the bijective substitution $z=e^{i \theta}$, we obtain a trigonometric polynomial $p(\theta):=\sum_{k=0}^{n} a_{k} e^{i \theta k} \in T_{n}$ with

$$
\max _{\theta \in[0,2 \pi]}\left|S^{\prime}(\theta)\right| \leqslant n \max _{\theta \in[0,2 \pi]}|S(\theta)| .
$$

Let $\mathcal{A} \subset \mathcal{C}([a, b] ; \mathbb{R})$ be set of approximating functions and $w:[a, b] \rightarrow(0, \infty)$ a fixed positive integrable weight function.
The best weighted least squares approximation from $\mathcal{A}$ to $f \in \mathcal{C}([a, b] ; \mathbb{R})$ is $\operatorname{argmin}_{p \in \mathcal{A}}\|f-p\|_{w}$, where the weighted scalar product and the induced norm are (for $f, g \in$ $\mathcal{C}([a, b] ; \mathbb{R}))(f, g)_{w}:=\int_{a}^{b} w(x) f(x) g(x) \mathrm{d} x,\|f\|_{w}:=\sqrt{(f, f)_{w}}$. Hence least squares approximation is best approximation with a weighted norm, so if e.g. $\mathcal{A}$ is a finite dimensional linear subspace, then the best weighted least approximation exists.

Let $\mathcal{A} \subset(H,(\cdot, \cdot))$ be linear subspace of an inner product space and $f \in H$. Then $p^{*} \in \mathcal{A}$ is best approximation from $\mathcal{A}$ to $f$ if and only if the error $e^{*}:=f-p^{*}$ is orthogonal to $\mathcal{A}$, that is, $\left(e^{*}, p\right)=0$ for all $p \in \mathcal{A}$.


We have Pythagoras' Theorem: for $p^{*}, q^{*} \in \mathcal{A}$ and $f \in H,\left\|f-q^{*}\right\|^{2}=$ $\left\|f-p^{*}\right\|^{2}+\left\|q^{*}-p^{*}\right\|^{2}$. For $q^{*}=0 \in \mathcal{A}$ we get $\|f\|^{2}=\left\|p^{*}\right\|^{2}+\left\|f-p^{*}\right\|^{2}$.

The monic orthogonal polynomials with respect to $w$ are uniquely determined by $Q_{0} \equiv 1, Q_{1}(x):=x-a_{0}$ and for $j \geqslant 1$ by

$$
Q_{j+1}(x):=\left(x-a_{j}\right) Q_{j}(x)-b_{j} Q_{j-1}(x)
$$

where

$$
a_{j}:=\frac{\left(Q_{j}, x Q_{j}(\cdot)\right)_{w}}{\left(Q_{j}, Q_{j}\right)_{w}}, \quad b_{j}:=\frac{\left(x Q_{j}, Q_{j-1}\right)_{w}}{\left(Q_{j-1}, Q_{j-1}\right)_{w}}=\frac{\left\|Q_{j}\right\|_{w}^{2}}{\left\|Q_{j-1}\right\|_{w}^{2}} .
$$

All roots of $Q_{n}$ are simple, real and contained in $(a, b)$.

Let $\mathcal{A} \subset \mathcal{C}([a, b])$ be a $n$-dimensional subspace. Then a function $f \in \mathcal{C}([a, b])$ has a unique best approximation from $\mathcal{A}$ if and only if $\mathcal{A}$ is a HaAR space.

If $p \in P_{n}$, then $\max _{x \in[-1,1]}\left|p^{\prime}(x)\right| \leqslant n^{2} \max _{x \in[-1,1]}|p(x)|$ and we have equality for $p=\alpha T_{n}$ for any $\alpha \in \mathbb{R}$. This bounded is optimal. (As $P_{n}$ is finite-dimensional, and $p \mapsto p^{\prime}$ is a linear operation, there exists a constant $M_{n} \in \mathbb{R}$ such that $\max _{x \in[-1,1]}\left|p^{\prime}(x)\right| \leqslant M_{n} \max _{x \in[-1,1]}|p(x)|$.) One can even show that (proven by Markov's brother)
$\max _{x \in[-1,1]}\left|p^{(k)}(x)\right| \leqslant \frac{n^{2}\left(n^{2}-1^{2}\right) \ldots\left(n^{2}-(k-1)^{2}\right)}{1 \cdot 3 \cdot \ldots \cdot(2 k-1)} \max _{x \in[-1,1]}|p(x)|$ where equality holds for any multiple of $T_{n}$.

By a lemma we have
$\max _{x \in[-1,1]}\left|p^{\prime}(x)\right| \leqslant n \max _{x \in[-1,1]}\left|\sqrt{1-x^{2}} p^{\prime}(x)\right|=n \max _{\theta \in[0,2 \pi]}\left|\sin (\theta) p^{\prime}(\cos (\theta))\right|$.
Let $S(\theta):=p(\cos (\theta))$. Then $S^{\prime}(\theta)=-p^{\prime}(\cos (\theta)) \sin (\theta)$. By the Bernstein-Markov inequality for $S \in T_{n}$ we have

$$
\begin{aligned}
\max _{x \in[-1,1]}\left|p^{\prime}(x)\right| & \stackrel{(3)}{\leqslant} n \max _{\theta \in[0,2 \pi]}\left|\sin (\theta) p^{\prime}(\cos (\theta))\right|=n \max _{\theta \in[0,2 \pi]}\left|S^{\prime}(\theta)\right| \\
& \stackrel{\text { Bernstein }}{\leqslant} n^{2} \max _{\theta \in[0,2 \pi]}|S(\theta)| \\
& =n^{2} \max _{\theta \in[0,2 \pi]}|p(\cos (\theta))|=n^{2} \max _{x \in[-1,1]}|p(x)| .
\end{aligned}
$$

$f, g \in \mathcal{C}([a, b] ; \mathbb{R}), f \neq g,\|f\|_{w}=\|g\|_{w}=1$ (wlog midpoint strict convexity):

$$
\begin{aligned}
\left(\frac{1}{2}\|f+g\|_{w}\right)^{2} & =\frac{1}{4} \int_{a}^{b} w(x)|f(x)+g(x)|^{2} \mathrm{~d} x \\
& <\frac{1}{2} \int_{a}^{b} w(x)|f(x)|^{2} \mathrm{~d} x+\frac{1}{2} \int_{a}^{b} w(x)|g(x)|^{2} \mathrm{~d} x \\
& =\frac{1}{2}\|f\|_{w}^{2}+\frac{1}{2}\|g\|_{w}^{2}=1,
\end{aligned}
$$

using that $(x+y)^{2}<2\left(x^{2}+y^{2}\right)$ for $x \neq y$.
Hence if $\mathcal{A} \subset \mathcal{C}([a, b] ; \mathbb{R})$ is a linear subspace, either the least squares approximation does not exist or it is unique.

Let $\mathcal{A} \subset H$ be a linear subspace of an inner product space spanned by basis functions $\left(\varphi_{i}\right)_{i=0}^{n}$ and $f \in H$. If the orthogonality condition

$$
\left(\varphi_{i}, \varphi_{j}\right)=0 \quad \forall i \neq j, i, j \in\{0, \ldots, n\}
$$

is satisfied, then the best approximation from $\mathcal{A}$ to $f$ is

$$
p^{*}=\sum_{i=0}^{n} \frac{\left(\varphi_{i}, f\right)}{\left\|\varphi_{i}\right\|^{2}} \varphi_{i} .
$$

Orthogonal polynomials

Approximation Theory

Lemma w/ Proof

Least squares approximation in $T_{n}$

Approximation Theory

3 Theorems w/o proofs

Uniform convergence of the partial Fourier sums

Approximation Theory

Definition, Theorem w/o proof

Chebychev series

Approximation Theory

Theorem w/o proof

Jackson's Theorem for $\mathcal{C}_{2 \pi}^{1}$ and for $\mathcal{C}_{2 \pi}^{k}$

Fourier series

Definition, Remark, Lemma w/o proof

Partial Fourier sum

Approximation Theory

Definition, Remark, Lemma w/o proof

Dirichlet kernel and its properties

Definition, Properties, Theorem w/o proof

FEJÉr kernel

Approximation Theory

Theorem w/o proof

Jackson's Theorem for $\mathcal{C}^{1}[-1,1], \mathcal{C}_{2 \pi}$ and for $\operatorname{Lip}_{K}^{1}$

The Fourier series of $f \in L_{2 \pi}^{1}$ is
$\frac{a_{0}}{2}+\sum_{k=1}^{\infty} a_{k}[f] \cos (k x)+b_{k}[f] \sin (k x)=\sum_{k \in \mathbb{Z}} c_{k}[f] e^{i k x}$, where $a_{k}[f]:=\frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos (k t) \mathrm{d} t, b_{k}[f]:=\frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin (k t) \mathrm{d} t$ and $c_{k}[f]:=\frac{1}{2 \pi} \int_{-\pi}^{\pi} f(t) e^{-i k t} \mathrm{~d} t$ are the Fourier coefficients with $2 c_{0}[f]=a_{0}[f], 2 c_{k}[f]=a_{k}[f]-i b_{k}[f], 2 c_{-k}[f]=$ $a_{k}[f]+i b_{k}[f]$ for $k \in \mathbb{N}_{>0}$.

The $n$-th partial Fourier sum of $f \in L^{1}(\mathbb{T})$ is the trigonometric polynomial of degree $n$
$S_{n}[f](x):=\sum_{|k| \leqslant n} c_{k}[f] e^{i k x}=\frac{a_{0}}{2}+\sum_{k=1}^{n} a_{k}[f] \cos (k x)+b_{k}[f] \sin (k x)$.
The operator $S_{n}: L^{2}(\mathbb{T}) \rightarrow T_{n}$ is linear and bounded with $\left\|S_{n}\right\|=1$ and a projection as well as injective.

For $n \in \mathbb{N}_{0}$, the $n$-th Dirichlet kernel is

$$
D_{n}(x):=\sum_{|k| \leqslant n} e^{i k x}= \begin{cases}\frac{\sin \left((2 n+1) \frac{x}{2}\right)}{\sin \left(\frac{x}{2}\right)}, & \text { if } x \in[-\pi, \pi] \backslash\{0\}, \\ 2 n+1, & \text { if } x=0 .\end{cases}
$$

$D_{n}$ is even, $\int_{-\pi}^{\pi} D_{n}(x) \mathrm{d} x=2 \pi$ and $\left|D_{n}(x)\right| \leqslant 2 n+1$ with equality only for $x=0 . c_{k}\left[D_{n}\right]=\mathbb{1}_{|k| \leqslant n}$ and

$$
\frac{4}{\pi^{2}} \log (n) \leqslant \frac{1}{2 \pi} \int_{-\pi}^{\pi}\left|D_{n}(x)\right| \mathrm{d} x \leqslant 3+\log (n) .
$$

$S_{n}[f]=f * D_{n}$ and $D_{n}=1+2 \sum_{k=1}^{n} \cos (k \cdot)$.
The Fejér kernel is

$$
K_{n}:=\frac{1}{n} \sum_{k=0}^{n-1} D_{k} .
$$

Then $K_{n}$ is even and nonnegative, $\frac{1}{2 \pi} \int_{-\pi}^{\pi} K_{n}(t) \mathrm{d} t=1$ and for $t \neq 0, K_{n}(t)=\frac{\sin ^{2}\left(\frac{n}{2} \cdot\right)}{n \sin ^{2}(\dot{\overline{2}})}$. Further, $c_{\ell}\left[K_{n}\right]=\frac{n-|\ell|}{n} \mathbb{1}_{|\ell| \leqslant n}$.
Let $\sigma_{n}[f]:=\frac{1}{n} \sum_{k=0}^{n-1} S_{k}[f]=K_{n} * f$. Then $\sigma_{n}: \mathcal{C}(\mathbb{T}) \rightarrow T_{n}$ is linear, bounded, but not a projection with $\left\|f-\sigma_{n}[f]\right\|_{2} \xrightarrow{n \rightarrow \infty}$ 0, proven with Bohman-Korovkin, proves Weierstrass.

For $f \in \mathcal{C}^{1}([-1,1])$ we have $E_{n}(f) \leqslant \frac{\pi}{2(n+1)}\left\|f^{\prime}\right\|_{\infty}$.

For $f \in \mathcal{C}_{2 \pi}$ we have $E_{n}^{T}(f) \leqslant \frac{3}{2} w_{f}\left(\frac{\pi}{n+1}\right)$.

The orthogonal monic polynomials $P_{n}^{\alpha, \beta}$ corresponding to the weight function $w(x)=(1-x)^{\alpha}(1+x)^{\beta}$ for $\alpha, \beta>-1$ are the Jacobi polynomials.
In particular, $P_{n}^{-\frac{1}{2},-\frac{1}{2}}$ are the normalised CHEBYCHEV polynomials of the first kind, $P_{n}^{\frac{1}{2}, \frac{1}{2}}$ are the normalised CHEBYCHEV polynomials of the second kind.
If we choose $w \equiv 1$, we get the Legendre polynomials. The polynomials corresponding to $w(x):=\exp \left(-x^{2}\right)$ for $x \in(0, \infty)$ are the Hermite polynomials defined by $H_{0} \equiv 1, H_{1}(x)=2 x$ and $H_{n+1}(x)=2 x H_{n}(x)-2 n H_{n-1}(x)$ for $n \geqslant 1$.

The $n$-th Fourier partial sum is the best least squares approximation from $T_{n}$ to $f:\left\|f-S_{n}[f]\right\|_{2}=\min _{p \in T_{n}}\|f-p\|_{2}$ $\forall f \in L^{2}(\mathbb{T}) \supset \mathcal{C}(\mathbb{T})$

$$
\text { Let } p:=\sum_{|k| \leqslant n} d_{k} e^{i k} \in T_{n} . \text { Then }
$$

$\begin{aligned}\|f-p\|_{2}^{2} & =\|f\|_{2}^{2}-\langle p, f\rangle-\langle f, p\rangle+\sum_{|k| \leqslant n}\left|d_{k}\right|^{2} \\ & =\|f\|_{2}^{2}-\sum_{|k| \leqslant n}\left|c_{k}[f]\right|^{2}+\underbrace{\sum_{|k| \leqslant n}\left|c_{k}[f]-d_{k}\right|^{2} \geqslant\|f\|_{2}^{2}-\sum_{|k| \leqslant n}\left|c_{k}[f]\right|^{2}}_{\geqslant 0} .\end{aligned}$
with equality if and only if $c_{k}[f]=d_{k}$ for all $|k| \leqslant n$.

If $f \in \mathcal{C}^{2}(\mathbb{T})$, then $S_{n}[f] \rightrightarrows f$ absolutely.
If $f \in \mathcal{C}(\mathbb{T})$ and $\sum_{k \in \mathbb{Z}} c_{k}[f]<\infty$, then $S_{n}[f] \rightrightarrows f$.
(Bernstein:) For $f \in \mathcal{C}^{r}(\mathbb{T})$ and $n>1$ we have

$$
\left\|f-S_{n}[f]\right\|_{\infty} \leqslant c\left\|f^{(r)}\right\|_{\infty} \frac{\ln (n)}{n^{r}},
$$

where $c$ is a constant independent of $f$ and $n$.

The Chebychev series of $f:[-1,1] \rightarrow \mathbb{R}$ is

$$
\sum_{k=0}^{\infty} \frac{\left(f, T_{k}\right)_{w}}{\left(T_{k}, T_{k}\right)_{w}} T_{k},
$$

where $w:[-1,1] \rightarrow \mathbb{R}, x \mapsto\left(1-x^{2}\right)^{-\frac{1}{2}}$.
If $f \in \mathcal{C}^{2}([-1,1])$, then the Chebychev series of $f$ converges uniformly to $f$.

For $f \in \mathcal{C}_{2 \pi}^{1}$ we have

$$
E_{n}^{T}=\min _{p \in T_{n}}\|f-p\|_{\infty} \leqslant \frac{\pi}{2(n+1)}\left\|f^{\prime}\right\|_{\infty}
$$

This is optimal.
If $f \in \mathcal{C}_{2 \pi}^{k}$, then

$$
E_{n}^{T}(f) \leqslant\left(\frac{\pi}{2(n+1)}\right)^{k}\left\|f^{(k)}\right\|_{\infty}
$$

The optimal bound is $\frac{\pi}{2} \frac{1}{(n+1)^{k}}\left\|f^{(k)}\right\|_{\infty}$.

Jackson's Theorem III: $C_{2 \pi}$

Approximation Theory

3 Theorems w/o proof, Remark

Bernstein Theorems I, II, III

Approximation Theory

## Rational functions

Approximation Theory

Theorem w/ proof, Corollaries

## Stone-Weierstrass theorem

Approximation Theory

2 Theorems w/o proof

JACKSON's Theorem V for $f:[-1,1] \rightarrow \mathbb{R}$

Definition, Remark, Theorem

ZYGMUND modulus of continuity

Theorem w/ proof, Remark

Existence of a rational best approximation

Approximation Theory

Remark, Theorem w/o proof

Multivariate polynomials

Approximation Theory

Theorem w/ Proof

There are no HaAr spaces of continuous functions on $\mathbb{R}^{d}$ for $d \geqslant 2$, except one dimensional ones.

For $f:[-1,1] \rightarrow \mathbb{R}$ we have
$E_{n}(f) \leqslant \begin{cases}\frac{3}{2} w_{f}\left(\frac{\pi}{n+1}\right), & \text { if } f \in \mathcal{C}([-1,1]), \\ \frac{\pi M}{2(n+1)}, & \text { if } f \in \operatorname{Lip}_{k}^{M}, \\ \left(\frac{\pi}{2}\right)^{k}\left(\prod_{j=n-k+2}^{n+1} \frac{1}{j}\right)\left\|f^{(k)}\right\|_{\infty}, & \text { if } f \in \mathcal{C}^{k}([-1,1]),\end{cases}$

The Zygmund modulus of continuity of a bounded function $f$ is

$$
w_{f}^{*}(\delta):=\sup _{x} \sup _{|h|<\delta}|f(x+h)-2 f(x)+f(x-h)| .
$$

We have $w_{f}^{*}(\delta) \leqslant 2 w_{f}(\delta)$.
Let $f \in \mathcal{C}_{2 \pi}$. Then $E_{n}^{T}(f) \in O\left(\frac{1}{n}\right)$ if and only if $\delta \mapsto \frac{w_{f}^{*}(\delta)}{\delta}$ is bounded.
$\left\{R \in R_{m}^{n}([a, b]):\|R-f\|_{\infty} \leqslant\|f\|_{\infty}\right\}$ is closed, bounded, but not compact, e.g. $\frac{1}{k x+1} \rightarrow \mathbb{1}_{\{0\}}$.
For $f \in \mathcal{C}([a, b])$ there exists a a best approximation from $R_{m}^{n}([a, b])$.
$\delta:=\inf \left\{\|f-R\|_{\infty}: R \in R_{m}^{n}([a, b])\right\} . \exists\left(R_{k}=\frac{P_{k}}{Q_{k}}\right) \subset R_{m}^{n}([a, b])$ s.t. $\left\|f-R_{k}\right\|_{\infty} \rightarrow \delta$, assume $\left\|Q_{k}\right\|_{\infty}=1$. Take subsequence such that $\| R_{k}-$ $f \|_{\infty} \leqslant \delta+1$. Then $\left\|R_{k}\right\|_{\infty} \leqslant\left\|R_{k}-f\right\|_{\infty}+\|f\|_{\infty} \leqslant \delta+1+\|f\|_{\infty}=: \varepsilon$. Hence $\left|P_{k}(x)\right|=\left|R_{k}(x)\left\|Q_{k}(x) \mid \leqslant\right\| Q_{k}\left\|_{\infty}\right\| R_{k} \|_{\infty} \leqslant \varepsilon\right.$, so $\forall k \in \mathbb{N}$ we have $\left(P_{k}, Q_{k}\right) \in\left\{(P, Q) \in \mathcal{P}_{n} \times \mathcal{P}_{m}:\|P\|_{\infty} \leqslant \varepsilon,\|Q\|_{\infty} \leqslant 1\right\}$, which is compact. Up to a subsequence, $P_{k} \rightarrow P, Q_{k} \rightarrow Q$. Then $\|Q\|_{\infty}=1$. As $Q \in \mathcal{P}_{m}, \exists$ at most $m$ zeros of $Q$. As $|P(x)| \leqslant \varepsilon|Q(x)|$, zeros of $Q$ are also zeros of $P$, so we can get rid of zeros, so $\frac{P_{k}}{Q_{k}} \rightarrow \frac{P}{Q} \in R_{m}^{n}([a, b])$.

Let $\mathbb{N}_{0}^{d}:=\left\{\alpha=\left(\alpha_{1}, \ldots, \alpha_{d}\right): \alpha_{i} \in \mathbb{N}_{0}\right\}$ and $|\alpha|:=\sum_{i=1}^{d} \alpha_{i}$ as well as $x^{\alpha}=\prod_{k=1}^{d} x_{k}^{\alpha_{k}}$ for $\alpha \in \mathbb{N}_{0}^{d}$. The function $x \mapsto$ $x^{\alpha}$ is called monomial. A polynomial $p$ can be represented as $p(x)=\sum_{\alpha \in I} c_{\alpha} x^{\alpha}$, where $I \subset \mathbb{N}_{0}^{d}$ is finite. The degree of $P$ is $\max \left(\left\{|\alpha|: \alpha \in I, c_{\alpha} \neq 0\right\}\right)$. (Then the degree of $p(x) \equiv 0$ is $-\infty$.)
The linear space of all polynomials of degree at most $n$ in $\mathbb{R}^{d}$ is denoted by $P_{n}\left(\mathbb{R}^{d}\right)$. If $p \in P_{n}\left(\mathbb{R}^{d}\right)$, then $p(x)=\sum_{|\alpha| \leqslant n} c_{\alpha} x^{\alpha}$. The monomials $x \mapsto x^{\alpha}$ with $|\alpha| \leqslant n$ form a basis for $P_{n}\left(\mathbb{R}^{d}\right)$.

Assume we have a HAAR space with $\operatorname{dim} n \geqslant 2$ and HAAR system $\left\{u_{1}, \ldots, u_{n}\right\}$. Thus if $x_{1}, \ldots, x_{n} \subset \mathbb{R}^{d}$ are distinct, then $A:=\left[u_{i}\left(x_{j}\right)\right]_{i, j}$ is invertible. Select closed path in $\mathbb{R}^{d}$ containing $x_{1}$ and $x_{2}$ but no other points $x_{3}, \ldots, x_{n}$. Move $x_{1}$ and $x_{2}$ continuously towards each other along this path, s.t. $x_{1}$ and $x_{2}$ exchange places. This corresponds to exchanging the first and second column in $A$. Hence det changes sign, so it has to be zero somewhere on the path, which contradicts that $\left\{u_{1}, \ldots, u_{n}\right\}$ is a HAAR system.
Hence $\left\{x^{\alpha}:|\alpha| \leqslant n\right\}$ can't be a HAAR system of $P_{n}\left(\mathbb{R}^{d}\right)$ if $n, d \geqslant$ 2 , so interpolation for sets of $\operatorname{dim}\left(P_{n}\left(\mathbb{R}^{d}\right)\right)=\binom{n+d}{d}$ distinct points is not possible!

For $f \in \mathcal{C}_{2 \pi}$ we have $E_{n}^{T}(f) \leqslant \frac{3}{2} w_{f}\left(\frac{\pi}{n+1}\right)$.
One can get rid of the factor $\frac{3}{2}$.
This implies the second WeIerstrass approximation theorem.
Corollary: Dini-Lipschitz. If $f \in \mathcal{C}_{2 \pi}$ and $w_{f}(\delta) \ln \left(\frac{1}{\delta}\right) \xrightarrow{\delta \backslash 0}$ 0 , then $S_{n}[f] \rightrightarrows f$.

If $f \in \mathcal{C}_{2 \pi}$ and $E_{n}^{T}(f) \in O\left(n^{-\alpha}\right)$ for some $\alpha \in(0,1)$, then $f \in \operatorname{Lip}^{\alpha}$.

For $\alpha=1$ we have that $W_{f}(\delta) \leqslant k \delta$ implies that $E_{n}(f) \in$ $O\left(\frac{1}{n}\right)$ by Jackson's Theorem II, but the converse does not hold.
If $f \in \mathcal{C}_{2 \pi}$ and $E_{n}^{T}(f) \in O\left(\frac{1}{n}\right)$, then $w(\delta) \leqslant k \delta|\ln (\delta)|$ for small $\delta>0$.

Let $f \in \mathcal{C}_{2 \pi}, E_{n}^{T}(f) \in O\left(n^{-\alpha-p}\right)$, where $p \in \mathbb{N}$ and $\alpha \in(0,1)$. Then $f^{\prime}, \ldots, f^{(p)}$ exist and we have $f^{(p)} \in \operatorname{Lip}^{\alpha}$.

A rational function is a quotient of two polynomials:

$$
\begin{equation*}
x \mapsto \frac{P(x)}{Q(x)}=\frac{a_{0}+a_{1} x+\ldots+a_{n} x^{n}}{b_{0}+b_{1} x+\ldots+b_{m} x^{m}} . \tag{4}
\end{equation*}
$$

The set of bounded rational functions on an interval $[a, b]$ is

$$
R_{m}^{n}([a, b]):=\left\{(4): P \in \mathcal{P}_{n}, Q \in \mathcal{P}_{m}, Q(x)>0 \forall x \in[a, b]\right\} .
$$

The condition $Q(x)>0$ ensures continuity.

If $X$ is a compact metric space and $A \subset \mathcal{C}(X)$ is an subalgebra such that $1 \in A$ and $A$ separates points of $X$, that is for $x \neq y \in X$, then there exists a $f \in A$ with $f(x) \neq f(y)$, then $\bar{A}=\mathcal{C}(X)$.

Corollary: multidimensional WeIERSTRASS theorem: if $X \subset$ $\mathbb{R}^{d}$ is compact, then the polynomials in d variables on $X$ are dense in $\mathcal{C}(X)$.

For $x, y \in \mathbb{R}^{d}$ and $\alpha \in \mathbb{N}_{0}^{d}$ we have $(x+y)^{\alpha}=$ $\sum_{0 \leqslant \beta \leqslant \alpha}\binom{\alpha}{\beta} x^{\beta} y^{\alpha-\beta}$, where $\beta \leqslant \alpha$ holds if $\beta_{i} \leqslant \alpha_{i}$ for all $i \in\{1, \ldots, d\}$. If $\beta \leqslant \alpha$, we let $\binom{\alpha}{\beta}:=\frac{\alpha!}{\beta!(\alpha-\beta)!}$ and 0 else, where $\alpha!:=\prod_{k=1}^{d} \alpha_{k}$ !.
(Strictly) positive definite function

Approximation Theory

Definition, Examples

Radial function

Approximation Theory

Theorem w/o Proof

Schoenberg (1938)

Explanation

Which proofs of the Weierstrass
approximation theorem did we see?

Approximation Theory

Strictly positive definite functions are related to polynomial interpolation in higher dimensions.

Approximation Theory

Definition, Remark, Examples

Completely monotone function

Examples

Data $\left(\left(x_{k}, f_{k}\right)\right)_{k=1}^{n} \subset \mathbb{R}^{d} \times \mathbb{R}$ can be interpolated by the functions

$$
\begin{array}{cc}
\sum_{j=1}^{n} \frac{c_{j}}{\sqrt{1+\left\|\cdot-x_{j}\right\|_{2}^{2}}} & \sum_{j=1}^{n} \frac{c_{j}}{1+\left\|\cdot-x_{j}\right\|_{2}^{2}} \\
\sum_{j=1}^{n} c_{j} e^{-\left\|\cdot-x_{j}\right\|_{2}^{2}} & \sum_{j=1}^{n} c_{j} e^{\cdot x_{j}}
\end{array}
$$

Approximation Theory

Assume $X$ is a linear space, $f: X \rightarrow \mathbb{R}$ is a function and $x_{1}, \ldots, x_{n} \in X$ are distinct. We want to find a function $g: X \rightarrow \mathbb{R}$ such that $g\left(x_{k}\right)=f\left(x_{k}\right)$ for all $k \in\{1, \ldots, n\}$ of the form $g=\sum_{j=1}^{m} a_{j} \varphi\left(\cdot-\nu_{j}\right)$, where we suppose that $\nu_{1}, \ldots, \nu_{m} \in X$ are known, $\varphi: X \rightarrow \mathbb{R}$ is fixed, and the $a_{1}, \ldots, a_{m} \in \mathbb{R}$ are unknown. Then $g\left(x_{k}\right)=f\left(x_{k}\right)$ for all $k \in\{1, \ldots, n\}$ is a system of $n$ linear combinations with $m$ unknowns.
For $X=\mathbb{R}^{d}, m=n$ and $\nu_{j}=x_{j}$, this can be rewritten as $A^{(\varphi)}\left(a_{j}\right)_{j=1}^{n}=\left(f\left(x_{j}\right)\right)_{j=1}^{n}$, so its uniquely solvable if $f$ is strictly positive definite.

A function $\varphi:[0, \infty) \rightarrow \mathbb{R}$ is complete monotone if $\varphi \in$ $\mathcal{C}([0, \infty)) \cap \mathcal{C}^{\infty}((0, \infty))$ and $(-1)^{k} \varphi^{(k)}$ is nonnegative for all $k \in \mathbb{N}_{0}$.

The class of completely monotone functions is closed under addition, multiplication and scalar multiplication (like the positive definite functions).

The functions $e^{-\alpha x}$ and $\alpha$ for $\alpha \geqslant 0, \ln \left(\frac{x+2}{x+1}\right)$ as well as $(x+\beta)^{-\alpha}$ for $\beta>0$ and $\alpha \geqslant 0$ are completely monotone.
E.g.: set $\Phi\left(\cdot-x_{j}\right):=\frac{1}{\sqrt{1+\left\|\cdot-x_{j}\right\|_{2}^{2}}} \stackrel{!}{=} \varphi\left(\left\|\cdot-x_{j}\right\|_{2}\right)$, so $\varphi(x)=$ $\left(1+x^{2}\right)^{-1}$ and $\varphi(\sqrt{x})=(1+x)^{-1}$ is completely monotone and nonconstant.
Last point: since $A:=\left[e^{x_{i} x_{j}}\right]_{i, j=1}^{n}$ is positive definite: as GAUSSIANS are positive definite,

$$
0<\left[e^{-\left\|x_{i}-x_{j}\right\|_{2}^{2}}\right]_{i, j=1}^{n}=\left[e^{-\left\|x_{i}\right\|_{2}^{2}} e^{2 x_{i} x_{j}} e^{-\left\|x_{j}\right\|_{2}^{2}}\right]_{i, j=1}^{n}=D A D,
$$

where $D:=\operatorname{diag}\left(e^{-\left\|x_{1}\right\|_{2}^{2}}, \ldots, e^{-\left\|x_{n}\right\|_{2}^{2}}\right)$, which is invertible (If $A>0$ and $\operatorname{det}(B) \neq 0$, then $B^{*} A B>0$.).

A function $\varphi: X \rightarrow \mathbb{C}$ is positive definite if for all $n \in \mathbb{N}$ and every $n$ distinct points $x_{1}, \ldots, x_{n}$ and $\alpha \in \mathbb{C}^{n}$ we have

$$
\begin{equation*}
\sum_{j, k=1}^{n} \alpha_{j} \overline{\alpha_{k}} \varphi\left(x_{j}-x_{k}\right) \geqslant 0 \tag{5}
\end{equation*}
$$

and strictly positive definite if the above inequality is strict for all $\alpha \in \mathbb{C}^{n} \backslash\{0\}$.

Examples: cos, $x \mapsto \exp (i\langle y, x\rangle)$, non-negative linear combination and products of positive definite functions, Gaussians.

A function $\Phi: \mathbb{R}^{d} \rightarrow \mathbb{R}$ is radial if

$$
\Phi(x)=\Phi(y) \quad \text { for all } x, y \in \mathbb{R}^{d} \text { with }\|x\|_{2}=\|y\|_{2}
$$

Hence $\Phi$ is radial if there exists a function $\varphi:[0, \infty) \rightarrow \mathbb{R}$ such that $\Phi=\varphi \circ\|\cdot\|_{2}$.

For $d=1$, all even functions are radial. Gaussians such as $e^{-\alpha\|x\|_{2}^{2}}$ are radial.

A function $\varphi:[0, \infty) \rightarrow \mathbb{R}$ is completely monotone if and only if

$$
\Phi: \mathbb{R}^{d} \rightarrow \mathbb{R}, \quad x \mapsto \varphi\left(\|x\|_{2}^{2}\right)
$$

is positive definite for all $d \in \mathbb{N}_{>0}$.
But only positive definiteness is not enough for interpolation:
Let $\varphi:[0, \infty) \rightarrow \mathbb{R}$. Then

$$
\Phi: \mathbb{R}^{d} \rightarrow \mathbb{R}, \quad x \mapsto \varphi\left(\|x\|_{2}\right)
$$

is strictly positive definite for all $d \in \mathbb{N}_{>0}$ if and only if $\varphi \circ \sqrt{ }$. is completely monotone and non-constant.

Bernstein's proof with Bernstein polynomials, Fejér-Hermite

Non-existence theorem (for second Weierstrass theorem)
Fejér theorem (for second Weierstrass theorem)
Jackson's Theorem III for $\mathcal{C}_{2 \pi} /$ Jackson's theorem V for
$\mathcal{C}([-1,1])$.
Stone-Weierstrass

